

# Integrative Structure Validation Report ?

April 10, 2025 - 12:57 PM PDT

*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	8ZZB
PDB-Dev ID	PDBDEV_00000011
Structure Title	Integrative structure and functional anatomy of three spokes of a nuclear pore complex
Structure Authors	Kim SJ; Fernandez-Martinez J; Nudelman I; Shi Y; Zhang W; Raveh B; Herricks T; Slaughter BD; Hogan JA; Upla P; Chemmama IE; Pellarin R; Echeverria I; Shivaraju M; Chaudhury AS; Wang J; Williams R; Unruh JR; Greenberg CH; Jacobs EY; Yu Z; de la Cruz MJ; Mironska R; Stokes DL; Aitchison JD; Jarrold MF; Gerton JL; Ludtke SJ; Akey CW; Chait BT; Sali A; Rout MP
Deposited on	2018-01-16

*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)*

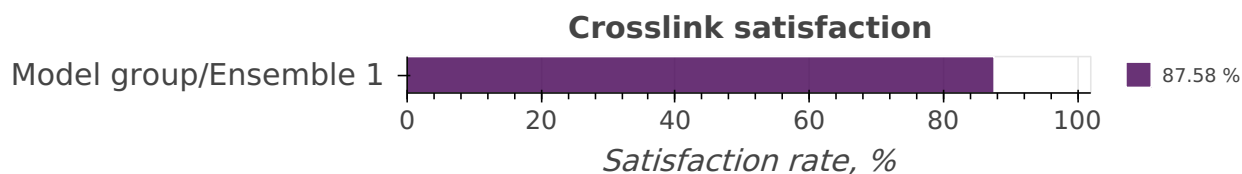
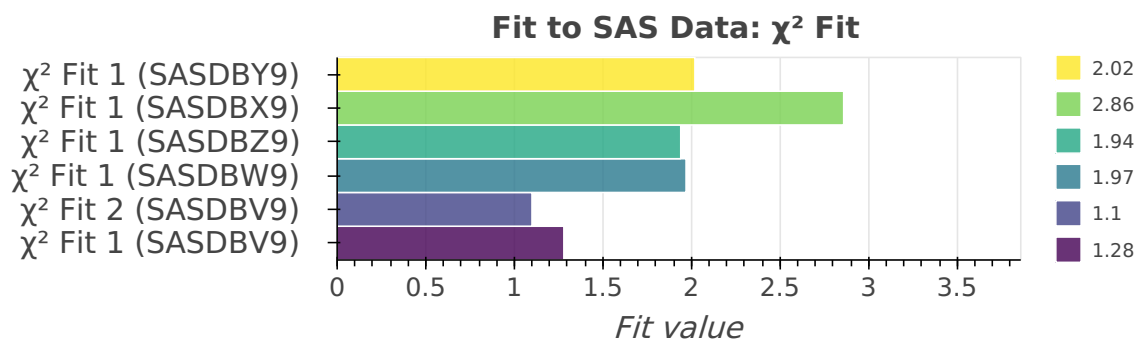
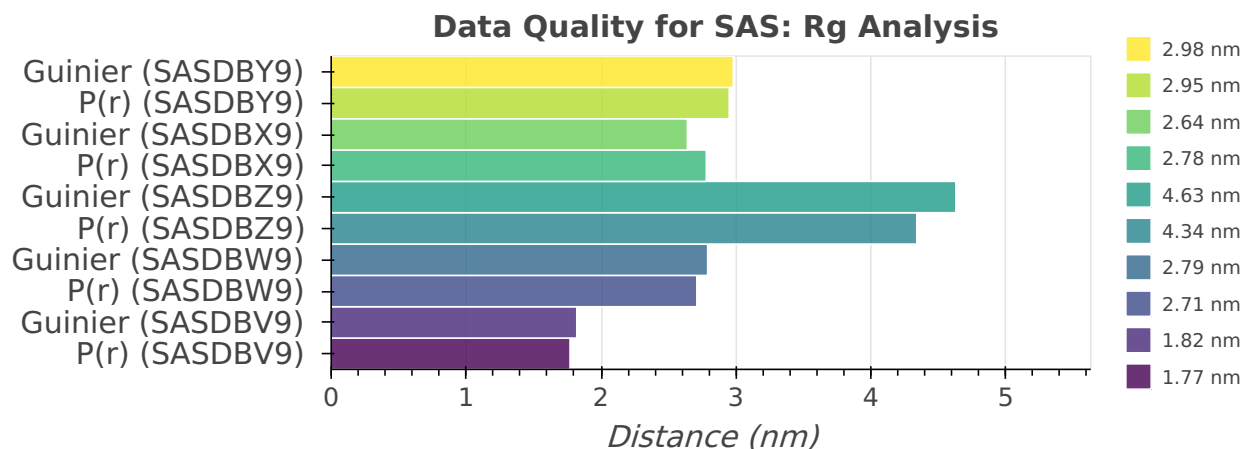
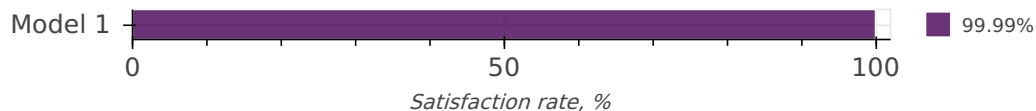
*A user guide is available at [https://pdb-ihm.org/validation\\_help.html](https://pdb-ihm.org/validation_help.html) with specific help available everywhere you see the ? symbol.*

*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: Excluded Volume Analysis



## Ensemble information ?

*This entry consists of 1 distinct ensemble(s).*

## Summary ?

*This entry consists of 1 model(s). A total of 65 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	Nup84	A	726	-	1-6, 7-20, 21-26, 27-80, 81-95, 96-126, 127-135, 136-364, 365-371, 372-483, 484-505, 506-562, 563-574, 575-726	100.00 / 89.39	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				H					
				O					
				V					
				AC					
				AJ					
		2	Nup85	B	744	-	1-46, 47-126, 127-131, 132-230, 231-234, 235-436, 437-450, 451-492, 493-495, 496-544, 545-552, 553-560, 561-566, 567-585, 586-589, 590-597, 598-602, 603-612, 613-615, 616-634, 635-637, 638-655, 656-660, 661-675, 676-684, 685-699, 700-706, 707-719, 720-724, 725-744	100.00 / 82.93	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				I					
				P					
				W					
				AD					
				AK					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		3	Nup120	C	1037	-	1, 2-29, 30-52, 53-305, 306-310, 311-711, 712-714, 715-726, 727-732, 733-746, 747-753, 754-766, 767-769, 770-781, 782-806, 807-818, 819-820, 821-833, 834-837, 838-853, 854-861, 862-879, 880-883, 884-895, 896-900, 901-913, 914-916, 917-931, 932-942, 943-955, 956-959, 960-971, 972-975, 976-987, 988-993, 994-1008, 1009-1024, 1025-1036, 1037	100.00 / 86.40	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				J					
				Q					
				X					
				AE					
				AL					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		4	Nup133	D	1157	-	1-55, 56-78, 79-85, 86-125, 126-132, 133-144, 145-161, 162-184, 185-192, 193-200, 201-205, 206-249, 250-257, 258-480, 481-489, 490-763, 764-771, 772-1155, 1156-1157	100.00 / 89.11	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				K					
				R					
				Y					
				AF					
				AM					
		5	Nup145c	E	712	-	1-91, 92-99, 100-125, 126-144, 145-148, 149-550, 551-553, 554-560, 561-565, 566-576, 577-586, 587-602, 603-611, 612-624, 625-630, 631-645, 646-653, 654-673, 674-680, 681-689, 690-702, 703-712	100.00 / 74.44	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				L					
				S					
				Z					
				AG					
				AN					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		6	Seh1	F	349	-	1-248, 249-287, 288-346, 347-349	100.00 / 87.97	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				M					
				T					
				AA					
				AH					
				AO					
		7	Sec13	G	297	-	1-9, 10-158, 159-165, 166-296, 297	100.00 / 94.28	Multiscale: Coarse-grained: 1 - 9 residue(s) per bead
				N					
				U					
				AB					
				AI					
				AP					
		8	Dyn2	AQ	92	-	1-6, 7-92	100.00 / 93.48	Multiscale: Coarse-grained: 1 - 6 residue(s) per bead
				AR					
				AY					
				AZ					
				BG					
				BH					
		9	Nup82	AS	713	-	1-6, 7-16, 17-22, 23-120, 121-122, 123-452, 453-521, 522-612, 613-624, 625-669, 670-677, 678-713	100.00 / 85.55	Multiscale: Coarse-grained: 1 - 10 residue(s) per bead
				AT					
				BA					
				BB					
				BI					
				BJ					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		10	Nup159	AU	1460	-	1082-1116, 1117-1126, 1127-1210, 1211-1239, 1240-1265, 1266-1321, 1322-1331, 1332-1372, 1373-1381, 1382-1412, 1413-1428, 1429-1456, 1457-1460	25.96 / 51.45	Multiscale: Coarse-grained: 1 - 35 residue(s) per bead
				AV					
				BC					
				BD					
				BK					
				BL					
		11	Nsp1	AW	823	-	601-636, 637-727, 728-741, 742-778, 779-787, 788-823	27.10 / 73.54	Multiscale: Coarse-grained: 1 - 36 residue(s) per bead
				AX					
				BE					
				BF					
				BM					
				BN					
				BX					
				CB					
				CF					
				CJ					
				CN					
				CR					
				CV					
				CZ					
				DD					
				DH					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		12	Nic96	BO	839	20-56	1-19, 57-204, 205-360, 361-365, 366-374, 375-404, 405-444, 445-454, 455-515, 516-532, 533-747, 748-752, 753-835, 836-839	100.00 / 71.63	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				BS					
		11	Nsp1	BP	823	637-727, 742-778, 788-823	601-636, 728-741, 779-787	27.10 / 73.54	Multiscale: Coarse-grained: 1 - 36 residue(s) per bead
				BT					
		13	Nup49	BQ	472	270-359, 369-407, 433-472	201-269, 360-368, 408-432	57.63 / 62.13	Multiscale: Coarse-grained: 1 - 69 residue(s) per bead
				BU					
		14	Nup57	BR	541	287-423, 433-476, 505-540	201-286, 424-432, 477-504, 541	63.03 / 63.64	Multiscale: Coarse-grained: 1 - 86 residue(s) per bead
				BV					



ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		12	Nic96	BW	839	-	1-19, 20-56, 57-204, 205-360, 361-365, 366-374, 375-404, 405-444, 445-454, 455-515, 516-532, 533-747, 748-752, 753-835, 836-839	100.00 / 71.63	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				CA					
				CE					
				CI					
				CM					
				CQ					
				CU					
				CY					
				DC					
				DG					
		13	Nup49	BY	472	-	201-269, 270-359, 360-368, 369-407, 408-432, 433-472	57.63 / 62.13	Multiscale: Coarse-grained: 1 - 69 residue(s) per bead
				CC					
				CG					
				CK					
				CO					
				CS					
				CW					
				DA					
				DE					
				DI					
		14	Nup57	BZ	541	-	201-286, 287-423, 424-432, 433-476, 477-504, 505-540, 541	63.03 / 63.64	Multiscale: Coarse-grained: 1 - 86 residue(s) per bead
				CD					
				CH					
				CL					
				CP					
				CT					
				CX					
				DB					
				DF					
				DJ					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		15	Nup157	DK	1391	-	1-87, 88-289, 290-300, 301-309, 310-338, 339-457, 458-480, 481-515, 516-534, 535-679, 680-703, 704-730, 731-743, 744-775, 776-785, 786-830, 831-835, 836-892, 893-899, 900-916, 917-920, 921-933, 934-943, 944-1016, 1017-1038, 1039-1141, 1142-1154, 1155-1390, 1391	100.00 / 80.01	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				DO					
				DS					
				DW					
				EA					
				EE					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		16	Nup170	DL	1502	-	1-97, 98-299, 300-310, 311-319, 320-352, 353-471, 472-504, 505-537, 538-573, 574-717, 718-764, 765-791, 792-830, 831-862, 863-883, 884-916, 917-918, 919-930, 931-935, 936-992, 993-999, 1000-1016, 1017-1020, 1021-1033, 1034-1043, 1044-1116, 1117-1140, 1141-1191, 1192-1194, 1195-1243, 1244-1256, 1257-1502	100.00 / 74.37	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				DP					
				DT					
				DX					
				EB					
				EF					
		17	Nup188	DM	1655	-	1-11, 12-34, 35-39, 40-91, 92-100, 101-123, 124-130, 131-166, 167-173, 174-224, 225-255, 256-282, 283- --- ---	100.00 / 82.18	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				DQ					
				DU					
				DY					
				EC					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	287, 288-304, 305-317, 318-434, 435-458, 459-479, 480-492, 493-508, 509-514, 515-530, 531-550, 551-577, 578-583, 584-605, 606-607, 608-619, 620-631, 632-785, 786-792, 793-889, 890-891, 892-1100, 1101-1118, 1119-1133, 1134-1156, 1157-1241, 1242-1246, 1247-1265, 1266-1275, 1276-1292, 1293-1302, 1303-1322, 1323-1331, 1332-1354, 1355-1382, 1383-1567, 1568-1592, 1593-1628, 1629-1632, 1633-1652, 1653-1655	Model coverage/ Starting model coverage (%)	Scale
				EG					

ID	Model(s)	Entity ID	Molecule name	Chain(s)	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				DN [auth]					
				DV	1683	-	1-362, 363-416, 417-574, 575-601, 602-798, 799-813, 814-849, 850-856, 857-953, 954-960, 961-1126, 1127-1136, 1137-1226, 1227-1233, 1234-1258, 1259-1271, 1272-1366, 1367-1370, 1371-1418, 1419-1420, 1421-1502, 1503-1510, 1511-1559, 1560-1583, 1584-1590, 1591-1596, 1597-1619, 1620-1622, 1623-1644, 1645-1650, 1651-1683	100.00 / 88.53	
				DZ					
				ED					
				EH					
		19	Nup53	EI	475	-	1-247, 248-284, 285-303, 304-360, 361-475	100.00 / 19.79	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
				EN					
				ES					
				EX					
				FC					
				FH					
		20	Nup59	EJ	528	-	1-265, 266-302, 303-345, 346-402, 403-528	100.00 / 17.80	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
				EO					
				ET					
				EY					
				FD					
				FI					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		21	Ndc1	EK	655	-	1-655	100.00 / 0.00	Multiscale: Coarse-grained: 55 - 100 residue(s) per bead
				EP					
				EU					
				EZ					
				FE					
				FJ					
		22	Pom34	EL	299	-	1-299	100.00 / 0.00	Multiscale: Coarse-grained: 49 - 50 residue(s) per bead
				EQ					
				EV					
				FA					
				FF					
				FK					
		23	Pom152	EM	1337	-	1-378, 379-472, 473-519, 520-611, 612-615, 616-714, 715-721, 722-818, 819-823, 824-918, 919-930, 931-1026, 1027-1035, 1036-1141, 1142-1149, 1150-1229, 1230-1243, 1244-1337	100.00 / 63.80	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
				ER					
				EW					
				FB					
				FG					
				FL					
		24	Nup100	FM	959	816-958	551-815, 959	42.65 / 34.96	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				FN					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		25	Nup116	FO	1113	-	751-965, 966-1111, 1112-1113	32.61 / 40.22	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				FP					
				FU					
				FV					
				GA					
				GB					
		27	Gle1	FR	538	-	1-120	22.30 / 0.00	Multiscale: Coarse-grained: 20 - 50 residue(s) per bead
				FX					
				GD					
		24	Nup100	FS	959	-	551-815, 816-958, 959	42.65 / 34.96	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				FT					
				FY					
				FZ					
		28	Nup145	GE	1317	459-605	201-458	30.75 / 36.30	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				GF					
		29	Nup1	GG	1076	-	1-351	32.62 / 0.00	Multiscale: Coarse-grained: 1 - 50 residue(s) per bead
				GL					
				GQ					
		30	Nup60	GH	539	-	1-398	73.84 / 0.00	Multiscale: Coarse-grained: 48 - 50 residue(s) per bead
				GI					
				GM					
				GN					
				GR					
				GS					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
		28	Nup145	GJ	1317	-	201-458, 459-605	30.75 / 36.30	Multiscale: Coarse-grained: 1 - 25 residue(s) per bead
				GK					
				GO					
				GP					
		31	Mlp1	GT	1875	-	238-716	25.55 / 0.00	Multiscale: Coarse-grained: 29 - 50 residue(s) per bead
				GV					
				GX					
		32	Mlp2	GU	1679	-	215-690	28.35 / 0.00	Multiscale: Coarse-grained: 26 - 50 residue(s) per bead
				GW					
				GY					

## Datasets used for modeling ?

There are 65 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Integrative model	Not available	<a href="#">10.1016/j.cell.2016.10.028</a>
2	Integrative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
3	Experimental model	PDB	5CWS
4	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
5	Experimental model	PDB	2QX5
6	Experimental model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
7	Experimental model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
8	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
9	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
10	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
11	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
12	Comparative model	Zenodo	<a href="#">10.5281/zenodo.1194547</a>



ID	Dataset type	Database name	Data access code
13	Comparative model	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
14	Integrative model	Not available	<a href="https://doi.org/10.1016/j.str.2017.01.006">10.1016/j.str.2017.01.006</a>
15	Integrative model	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
16	Experimental model	PDB	3NF5
17	Comparative model	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
18	Comparative model	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
19	Experimental model	PDB	3KEP
20	Experimental model	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
21	Mass Spectrometry data	Zenodo	<a href="https://zenodo.org/record/1149746">10.5281/zenodo.1149746</a>
22	Crosslinking-MS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
23	Crosslinking-MS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
24	EM raw micrographs	EMPIAR	EMPIAR-10155
25	3DEM volume	EMDB	EMD-7321
26	3DEM volume	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
27	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
28	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
29	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
30	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
31	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
32	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
33	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
34	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
35	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
36	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
37	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
38	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
39	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
40	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
41	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
42	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
43	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
44	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
45	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
46	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>
47	SAS data	Zenodo	<a href="https://zenodo.org/record/1194547">10.5281/zenodo.1194547</a>

ID	Dataset type	Database name	Data access code
48	SAS data	SASBDB	<a href="#">SASDBV9</a>
49	SAS data	SASBDB	<a href="#">SASDBW9</a>
50	SAS data	SASBDB	<a href="#">SASDBZ9</a>
51	SAS data	SASBDB	<a href="#">SASDBX9</a>
52	SAS data	SASBDB	<a href="#">SASDBY9</a>
53	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
54	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
55	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
56	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
57	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
58	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
59	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
60	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
61	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
62	SAS data	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
63	EM raw micrographs	EMPIAR	<a href="#">EMPIAR-10162</a>
64	2DEM class average	Zenodo	<a href="#">10.5281/zenodo.1194547</a>
65	2DEM class average	Zenodo	<a href="#">10.5281/zenodo.1194547</a>

## 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	500	False	True
2	1	Sampling	Replica exchange monte carlo	None	3000	False	True
3	1	Sampling	Replica exchange monte carlo	None	1000	False	True

*There are 13 software packages reported in this entry.*

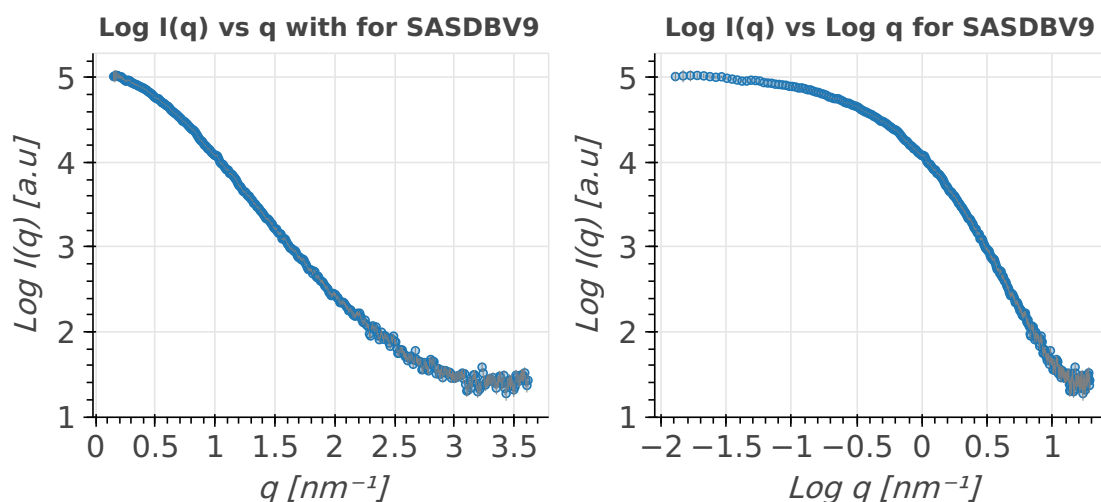
ID	Software name	Software version	Software classification	Software location
1	<a href="#">Integrative Modeling Platform (IMP)</a>	develop-0a5706e202	integrative model building	<a href="https://integrativemodeling.org">https://integrativemodeling.org</a>
2	<a href="#">IMP PMI module</a>	67456c0	integrative model building	<a href="https://integrativemodeling.org">https://integrativemodeling.org</a>
3	<a href="#">HHPred</a>	2.0.16	protein homology detection	<a href="https://toolkit.tuebingen.mpg.de/hhpred">https://toolkit.tuebingen.mpg.de/hhpred</a>
4	<a href="#">PSIPRED</a>	4.00	secondary structure prediction	<a href="http://bioinf.cs.ucl.ac.uk/psipred/">http://bioinf.cs.ucl.ac.uk/psipred/</a>
5	<a href="#">DISOPRED</a>	3	disorder prediction	<a href="http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1">http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1</a>
6	<a href="#">DomPred</a>	Not available	domain boundary prediction	<a href="http://bioinf.cs.ucl.ac.uk/dompred">http://bioinf.cs.ucl.ac.uk/dompred</a>
7	<a href="#">COILS/PCOILS</a>	Not available	coiled-coil prediction	<a href="https://toolkit.tuebingen.mpg.de/#/tools/pcoils">https://toolkit.tuebingen.mpg.de/#/tools/pcoils</a>
8	<a href="#">EMAN2</a>	2.20	image processing	<a href="http://blake.bcm.edu/emanwiki/EMAN2">http://blake.bcm.edu/emanwiki/EMAN2</a>
9	<a href="#">RELION</a>	1.40	image processing	<a href="https://www2.mrc-lmb.cam.ac.uk/relion/">https://www2.mrc-lmb.cam.ac.uk/relion/</a>
10	<a href="#">SGD</a>	Not available	database	<a href="https://www.yeastgenome.org/">https://www.yeastgenome.org/</a>
11	<a href="#">HeliQuest</a>	Not available	helix prediction	<a href="http://heliquet.ipmc.cnrs.fr/">http://heliquet.ipmc.cnrs.fr/</a>
12	<a href="#">MODELLER</a>	9.15	comparative modeling	<a href="https://salilab.org/modeller/">https://salilab.org/modeller/</a>
13	<a href="#">MODELLER</a>	9.13	comparative modeling	<a href="https://salilab.org/modeller/">https://salilab.org/modeller/</a>

## Data quality ?

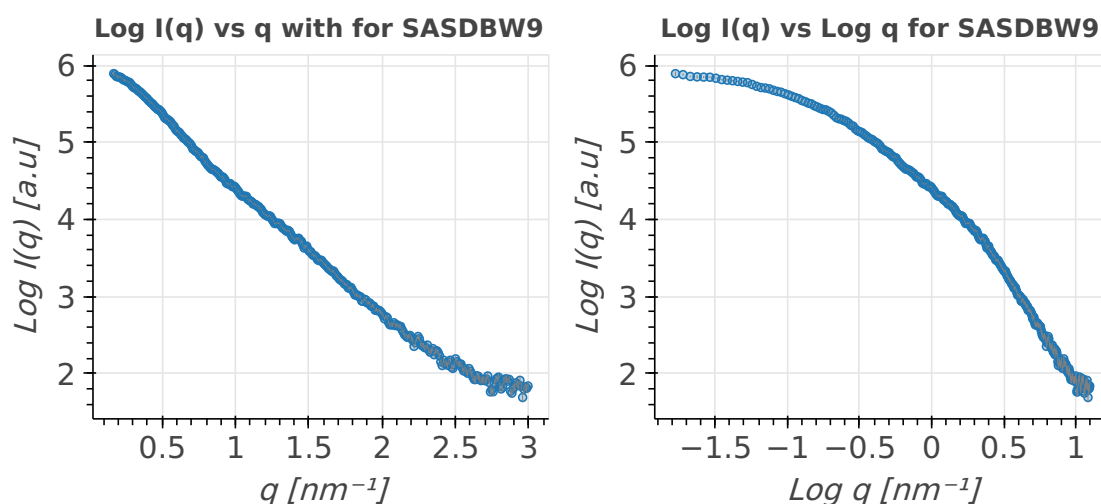
### Scattering profile ?

SAS data used in this integrative model was obtained from 5 deposited SASBDB entry (entries).

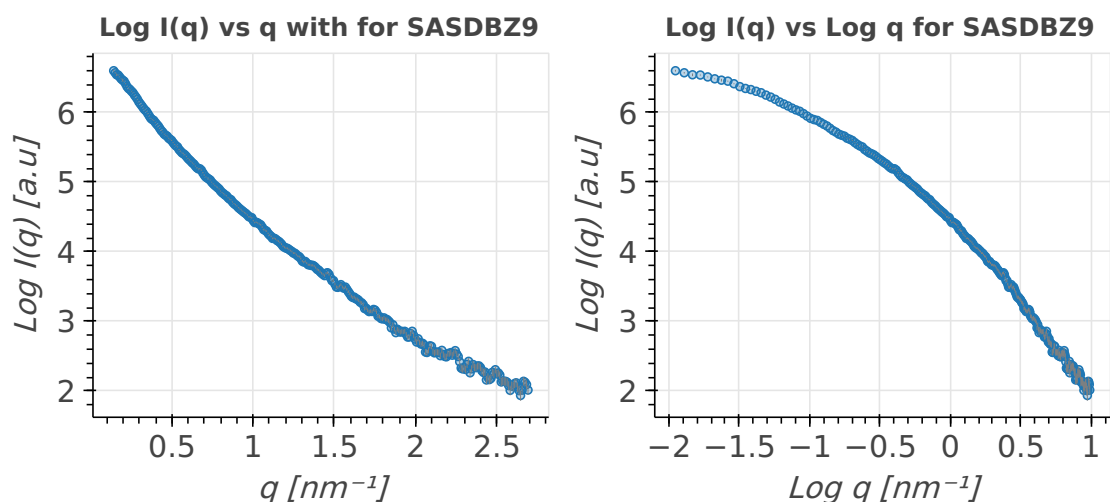
Scattering profile for [SASDBV9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



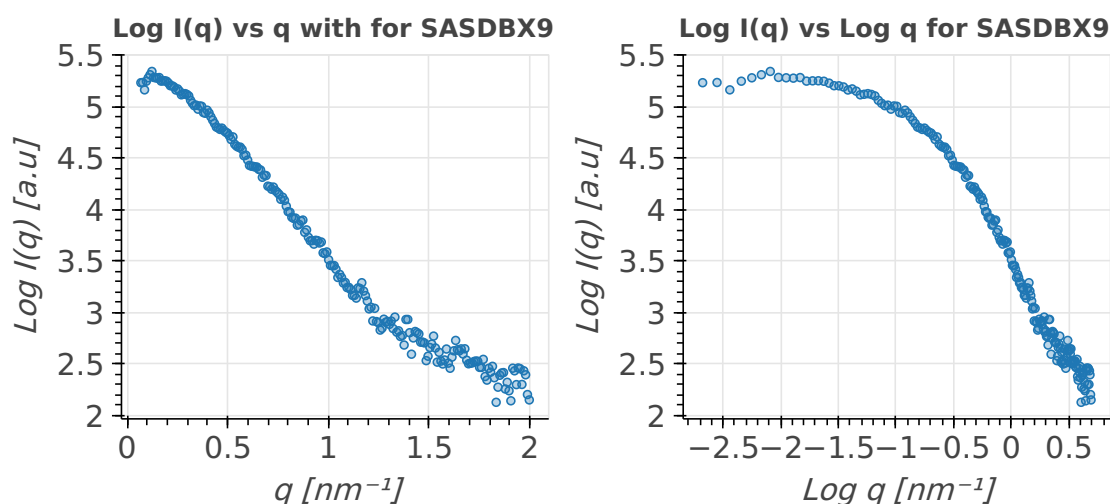
Scattering profile for [SASDBW9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



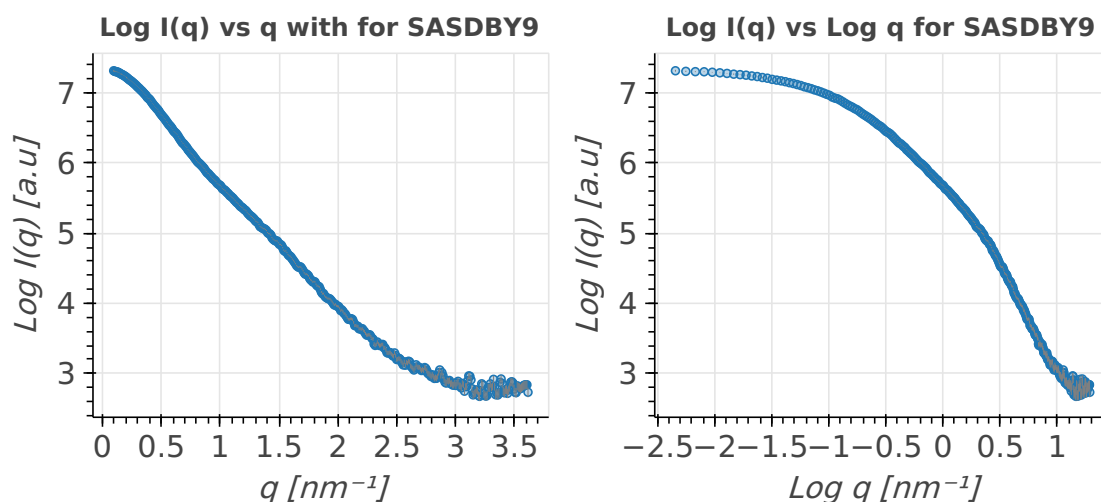
Scattering profile for [SASDBZ9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



Scattering profile for [SASDBX9](#): data from solutions of biological macromolecules are presented as both  $\text{log } I(q)$  vs  $q$  and  $\text{log } I(q)$  vs  $\text{log } (q)$  based on [SAS validation task force \(SASvtf\) recommendations](#).  $I(q)$  is the intensity (in arbitrary units) and  $q$  is the modulus of the scattering vector.



Scattering profile for [SASDBY9](#): data from solutions of biological macromolecules are presented as both  $\text{log } I(q)$  vs  $q$  and  $\text{log } I(q)$  vs  $\text{log } (q)$  based on [SAS validation task force \(SASvtf\) recommendations](#).  $I(q)$  is the intensity (in arbitrary units) and  $q$  is the modulus of the scattering vector.



### Key experimental estimates ?

Molecular weight (MW) estimates from experiments and analysis: true molecular weight can be compared to the Porod estimate from scattering profiles.

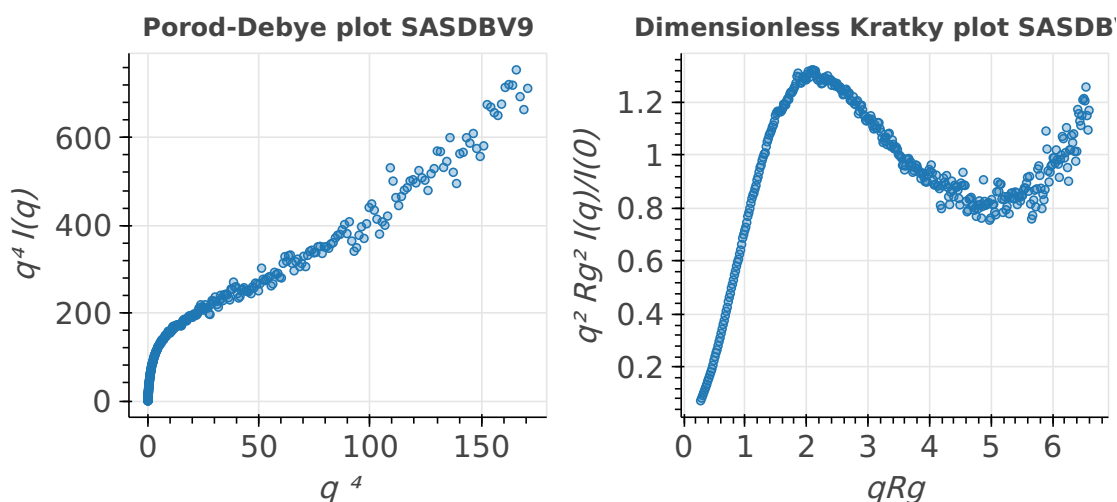
SASDB ID	Chemical composition MW	Standard MW	Porod Volume/MW
SASDBV9	12.6 kDa	12.2 kDa	Not available
SASDBW9	24.1 kDa	25.2 kDa	Not available
SASDBZ9	49.4 kDa	48.3 kDa	Not available
SASDBX9	12.5 kDa	14.7 kDa	Not available
SASDBY9	25.9 kDa	25.2 kDa	Not available

Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

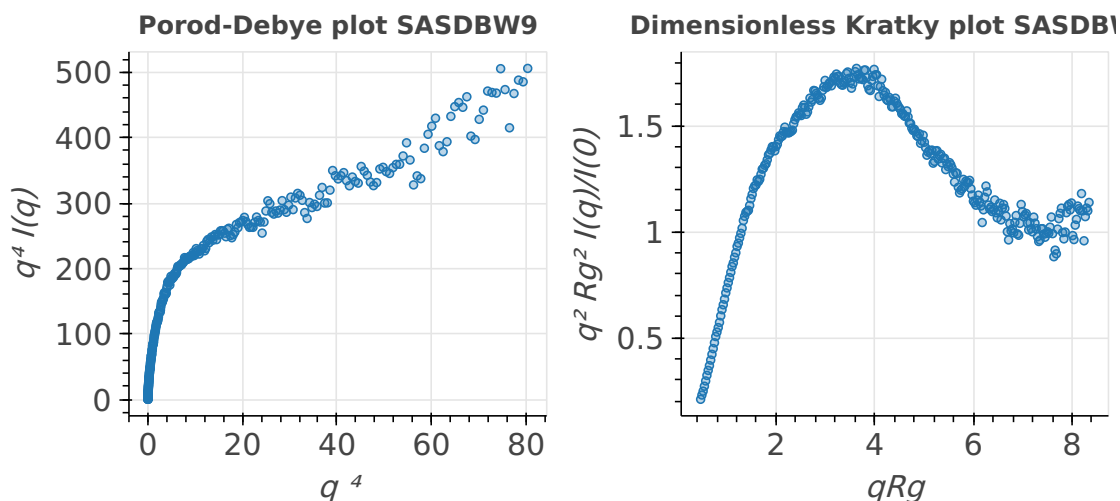
SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDBV9	Not available	17.94 $\text{nm}^3$	Not available	Not available	Not available
SASDBW9	Not available	22.50 $\text{nm}^3$	Not available	Not available	Not available
SASDBZ9	Not available	66.59 $\text{nm}^3$	Not available	Not available	Not available
SASDBX9	Not available	56.68 $\text{nm}^3$	Not available	Not available	Not available
SASDBY9	Not available	27.97 $\text{nm}^3$	Not available	Not available	Not available

### Flexibility analysis ?

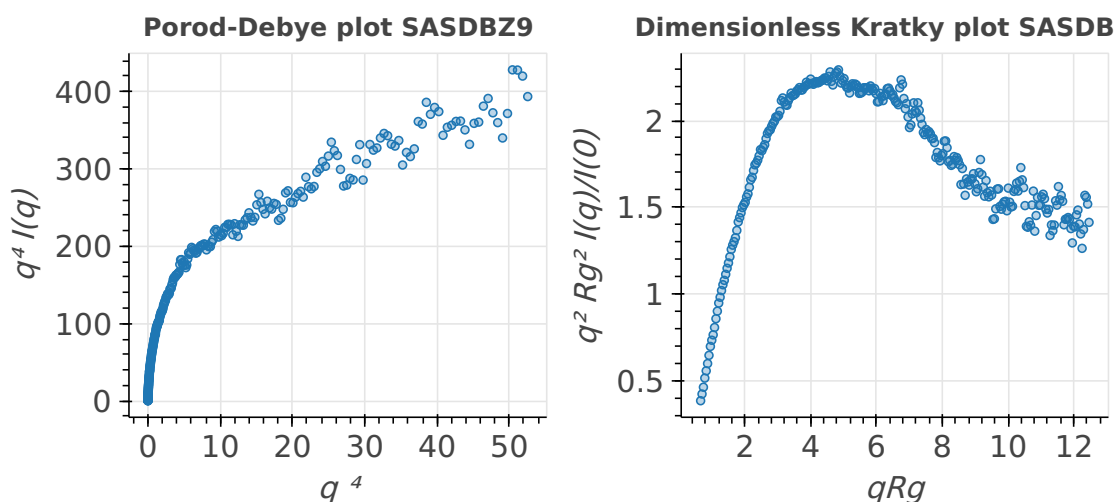
Flexibility analysis for SASDBV9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.



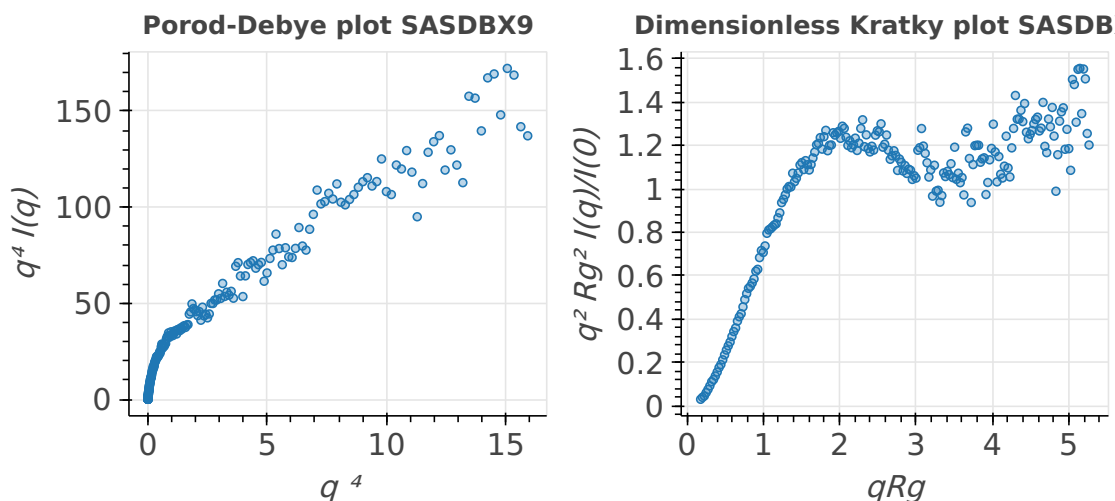
Flexibility analysis for SASDBV9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.



Flexibility analysis for SASDBZ9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.

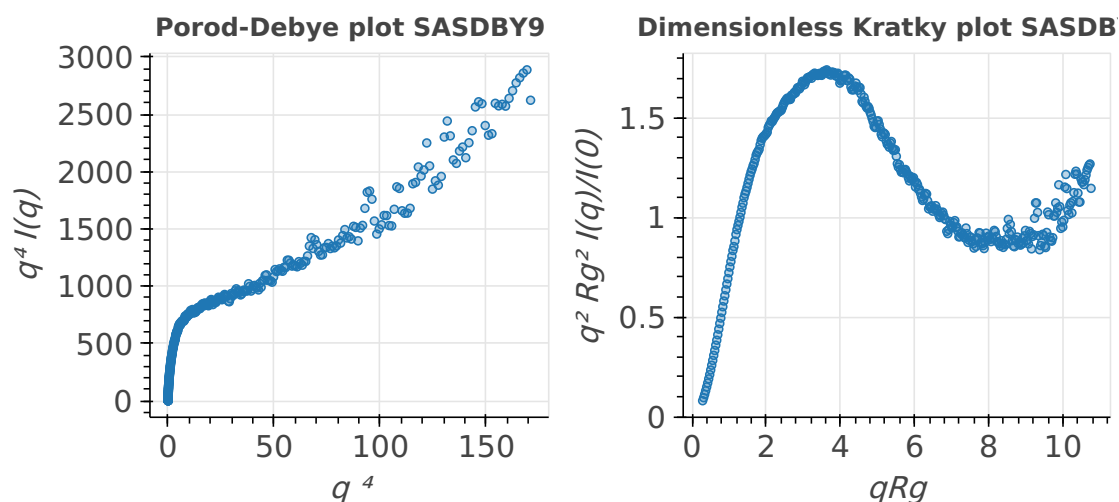


Flexibility analysis for SASDBX9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.



Flexibility analysis for SASDBY9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.



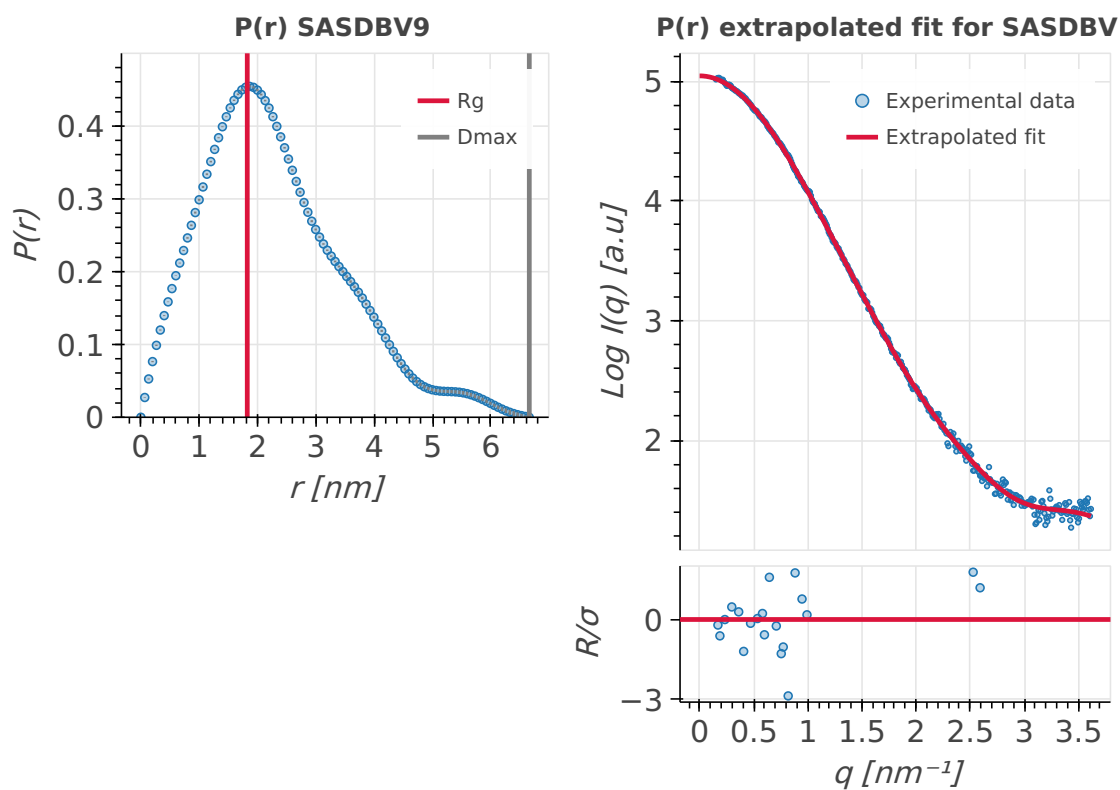


### Pair-distance distribution analysis ?

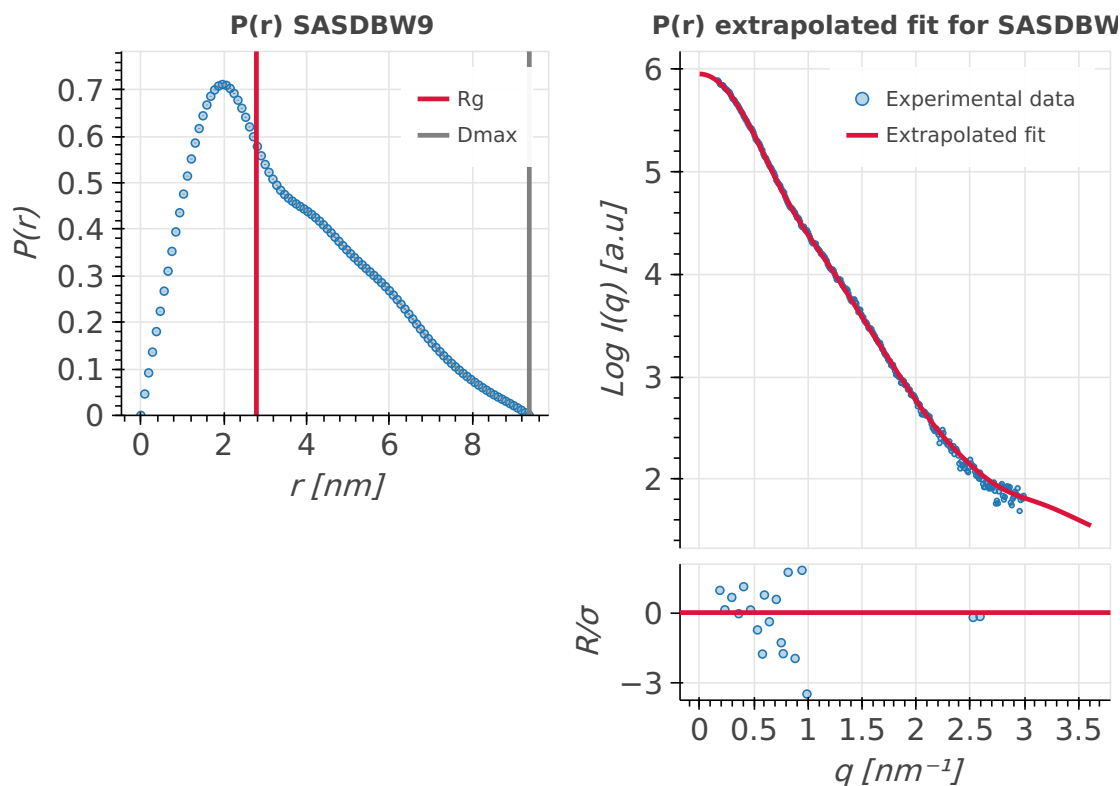
P(r) analysis: P(r) represents the distribution of distances between all pairs of atoms within the particle weighted by the respective electron densities. P(r) is the Fourier transform of I(s) (and vice versa).  $R_g$  can be estimated from integrating the P(r) function. Agreement between the P(r) and Guinier-determined  $R_g$  (table below) is a good measure of the self-consistency of the SAS profile.  $R_g$  is a measure for the overall size of a macromolecule; e.g. a protein with a smaller  $R_g$  is more compact than a protein with a larger  $R_g$ , provided both have the same molecular weight (MW). The point where P(r) is decaying to zero is called  $D_{\max}$  and represents the maximum size of the particle.

SASDB ID	Software used	$D_{\max}$	$D_{\max}$ error	$R_g$	$R_g$ error
SASDBV9	GNOM 4.5a	6.660 nm	Not available	1.824 nm	0.006 nm
SASDBW9	GNOM 4.5a	9.370 nm	Not available	2.787 nm	0.007 nm
SASDBZ9	GNOM 4.5a	15.430 nm	Not available	4.629 nm	0.011 nm
SASDBX9	GNOM 4.5a	7.930 nm	Not available	2.636 nm	0.008 nm
SASDBY9	GNOM 4.5a	10.450 nm	Not available	2.976 nm	0.005 nm

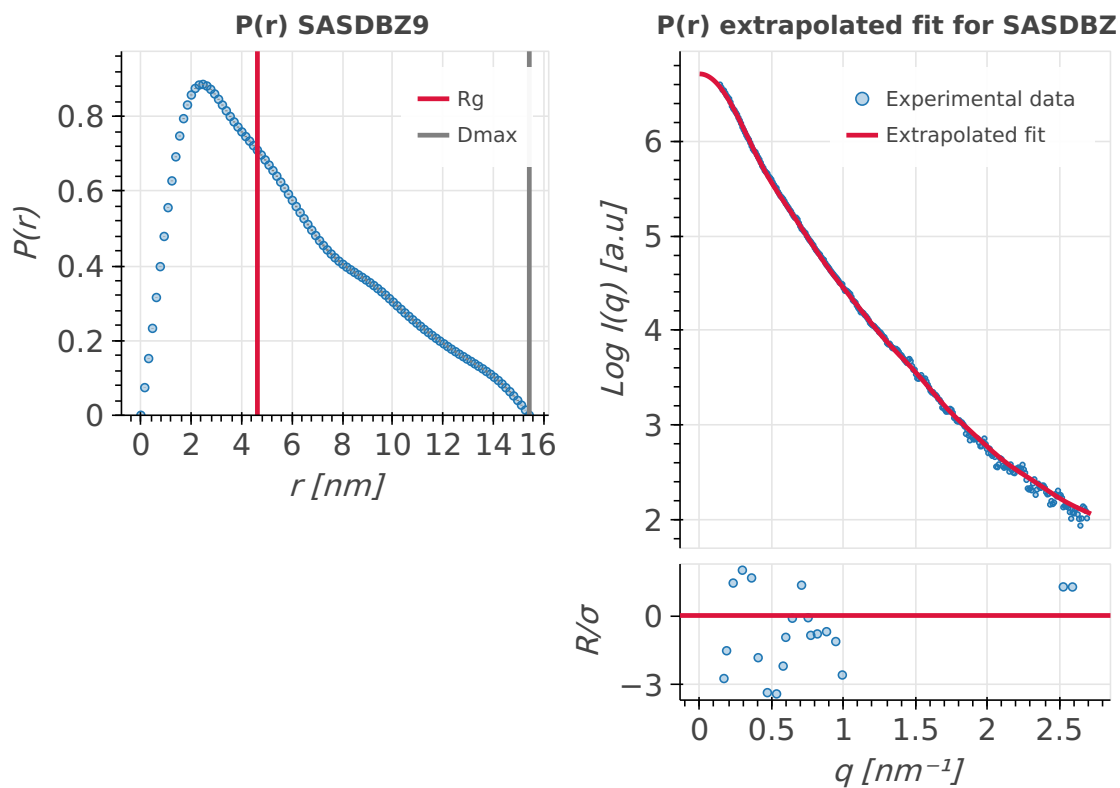
P(r) for SASDBV9: The value of P(r) should be zero beyond  $r=D_{\max}$ .



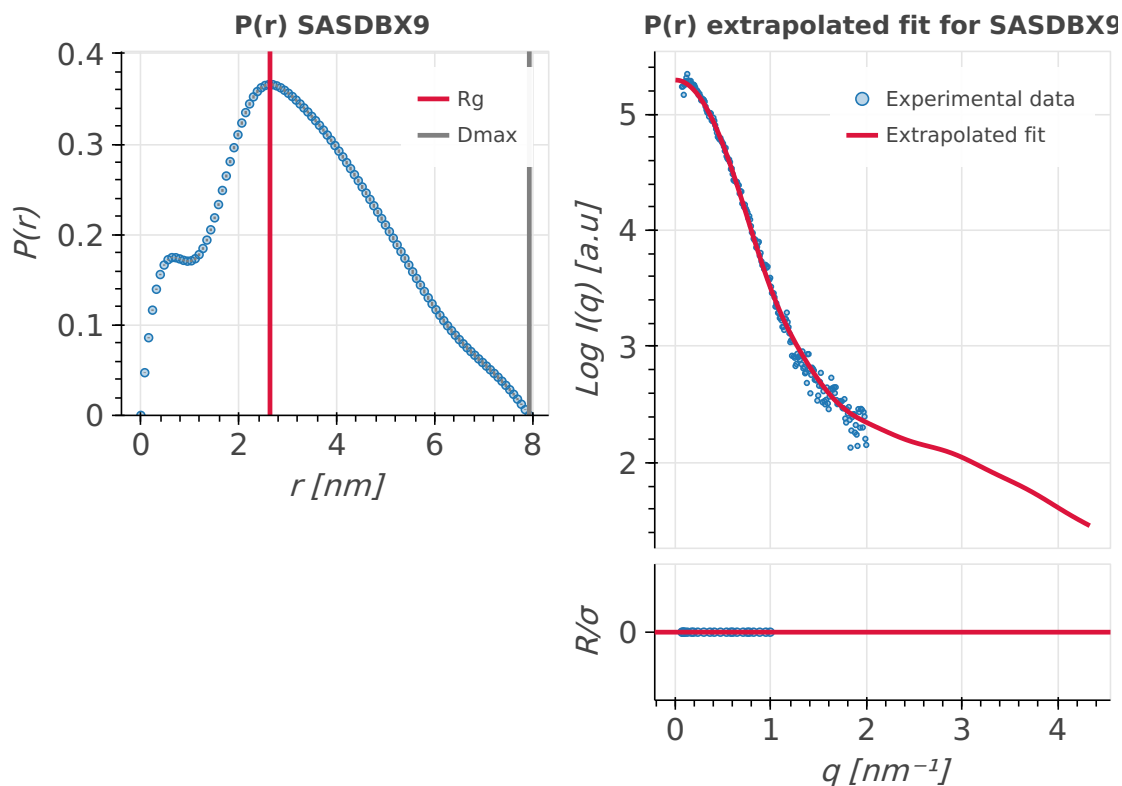
P(r) for SASDBW9: The value of  $P(r)$  should be zero beyond  $r=D_{max}$ .



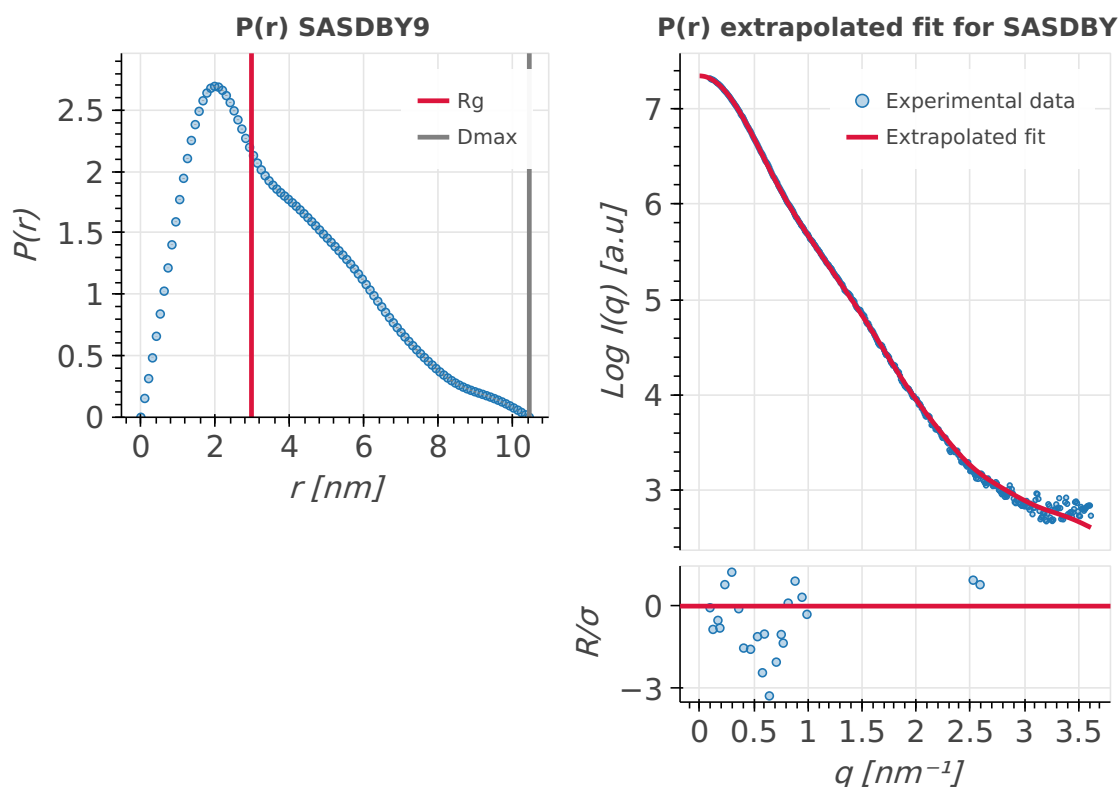
P(r) for SASDBZ9: The value of  $P(r)$  should be zero beyond  $r=D_{max}$ .



P(r) for SASDBX9: The value of  $P(r)$  should be zero beyond  $r=D_{max}$ .



P(r) for SASDBY9: The value of  $P(r)$  should be zero beyond  $r=D_{max}$ .

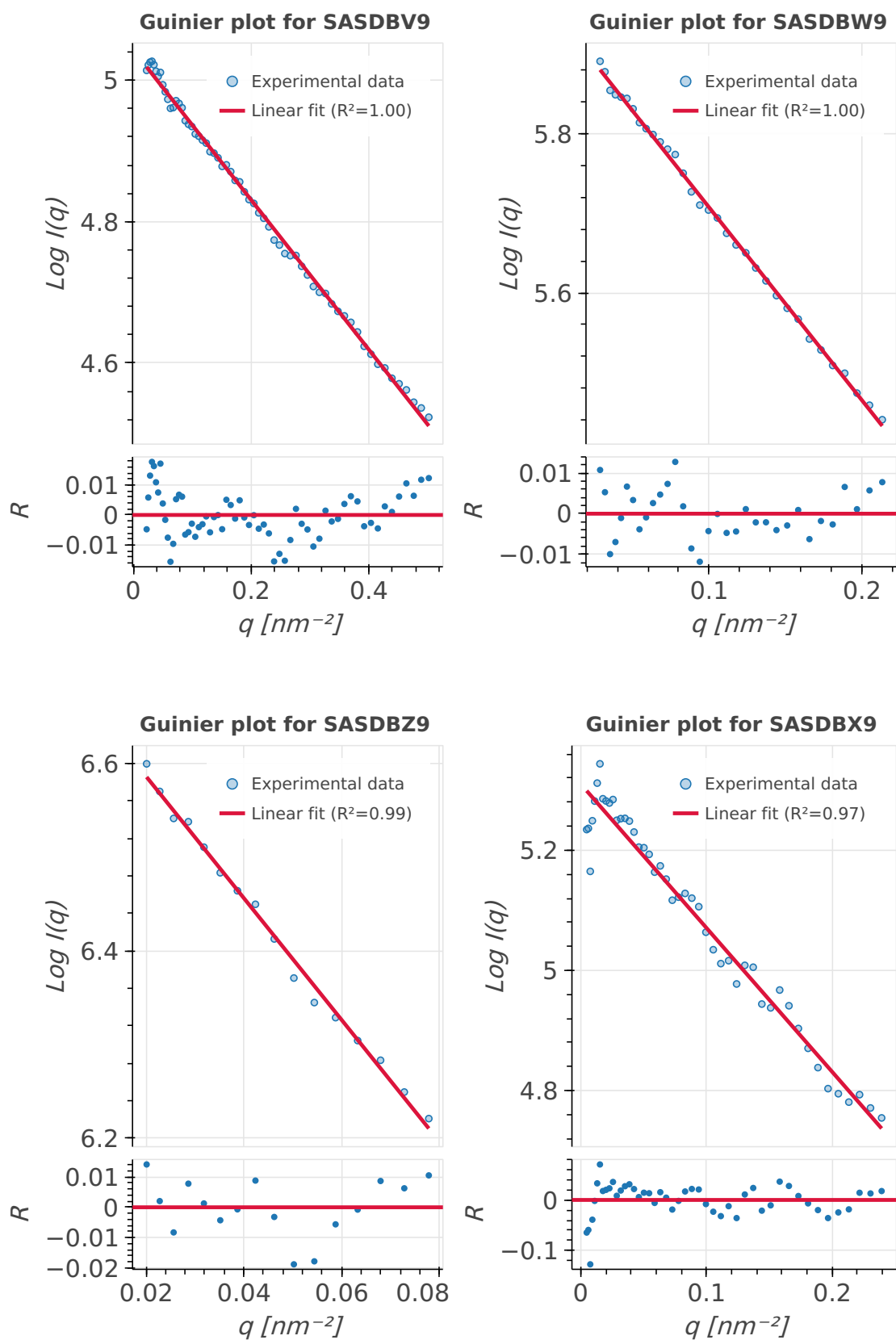


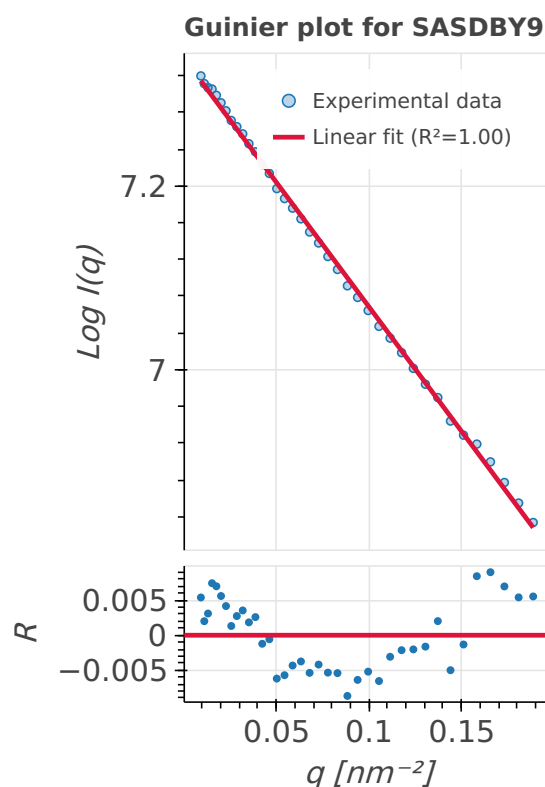
### Guinier analysis ?

Guinier analysis: agreement between the  $P(r)$  and Guinier-determined  $R_g$  (table below) is a good measure of the self-consistency of the SAS profile. Molecular weight estimates can also be compared to Porod and sample molecular weights for consistency.

SASDB ID	$R_g$	$R_g$ error	MW	MW error
SASDBV9	1.77 nm	0.05 nm	12.2 kDa	Not available
SASDBW9	2.71 nm	0.06 nm	25.2 kDa	Not available
SASDBZ9	4.34 nm	0.17 nm	48.3 kDa	Not available
SASDBX9	2.78 nm	0.18 nm	14.7 kDa	Not available
SASDBY9	2.95 nm	0.11 nm	25.2 kDa	Not available

Guinier analysis: the linearity of the Guinier plot is a sensitive indicator of the quality of the experimental SAS data; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity usually point to strong interference effects, polydispersity of the samples or improper background subtraction. Residual value plot and coefficient of determination ( $R^2$ ) are measures to assess linear fit to the data. A perfect fit has an  $R^2$  value of 1. Residual values should be equally and randomly spaced around the horizontal axis.





### 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.

### 3DEM volume

Validation for this section is under development.

### Mass Spectrometry

Validation for this section is under development.

### EM raw micrographs

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-particle or particle-atom pairs for which excluded volume was analysed.*

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	3856044471	229275	99.99

## Fit of model to data used for modeling ?

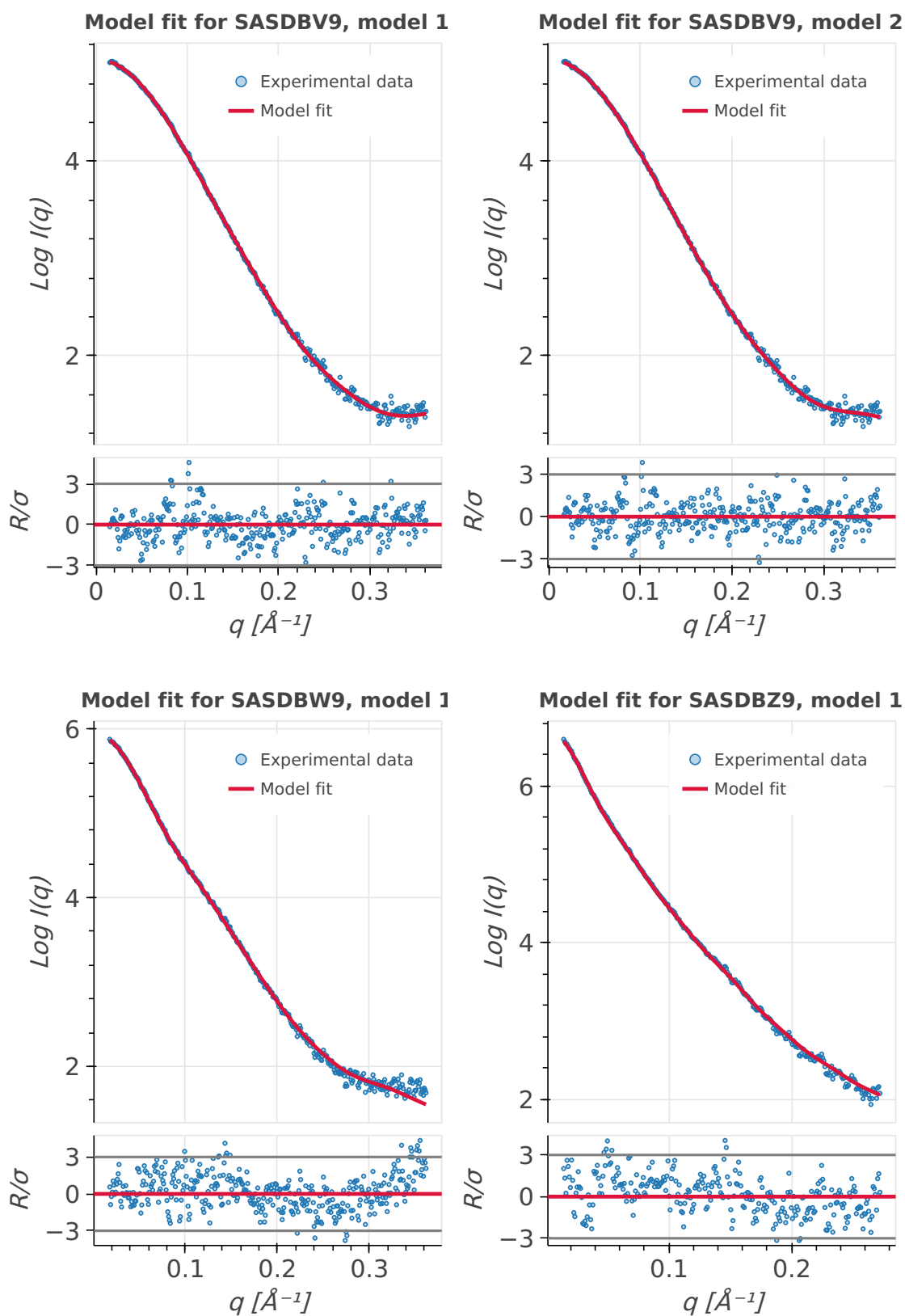
### Fit of model(s) to SAS data

#### $\chi^2$ goodness of fit and cormap analysis ?

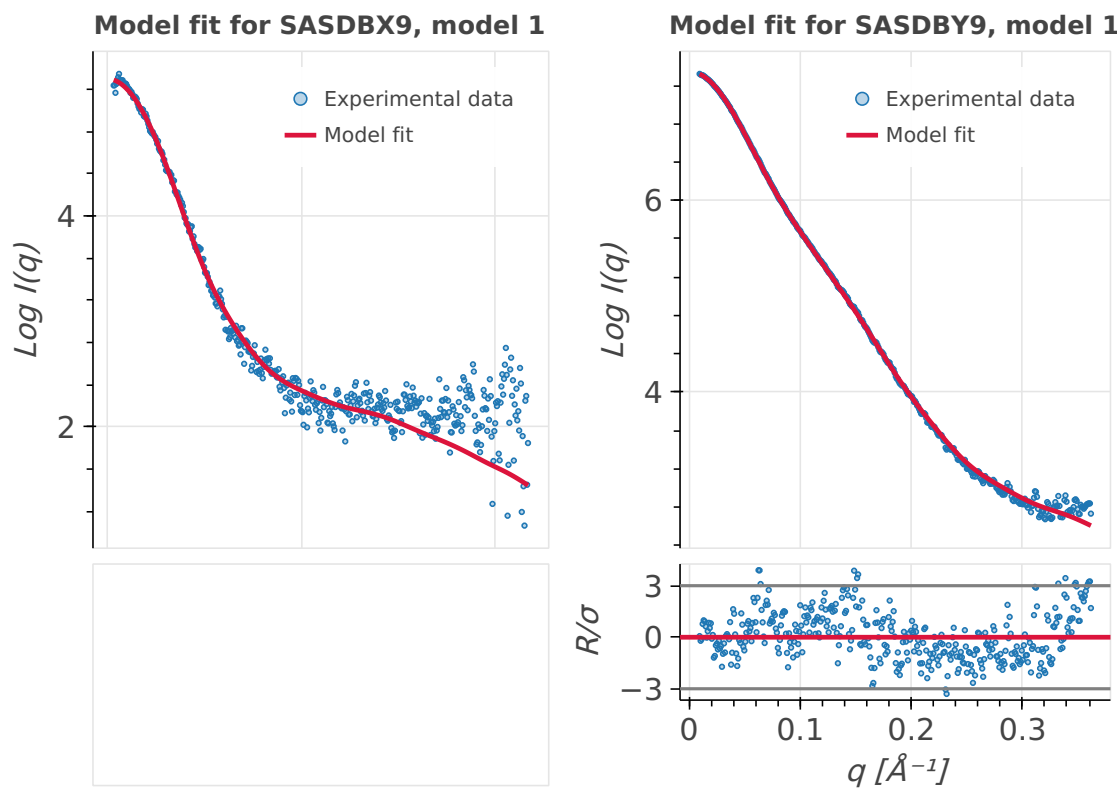
*Model and fits displayed below were obtained from SASBDB.  $\chi^2$  values are a measure of fit of the model to data. A perfect fit has a  $\chi^2$  value of 1.0. ATSAS DATCMP was used for hypothesis testing. All data sets are similar (i.e. the fit and the data collected) is the null hypothesis. p-value is a measure of evidence against the null hypothesis, smaller the value, the stronger the evidence that you should reject the null hypothesis.*

SASDB ID	Model	$\chi^2$	p-value
SASDBV9	1	1.28	0.02
SASDBV9	2	1.10	0.01
SASDBW9	1	1.97	0.00
SASDBZ9	1	1.94	0.00
SASDBX9	1	2.86	0.00
SASDBY9	1	2.02	0.00

Model fit(s): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.







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.

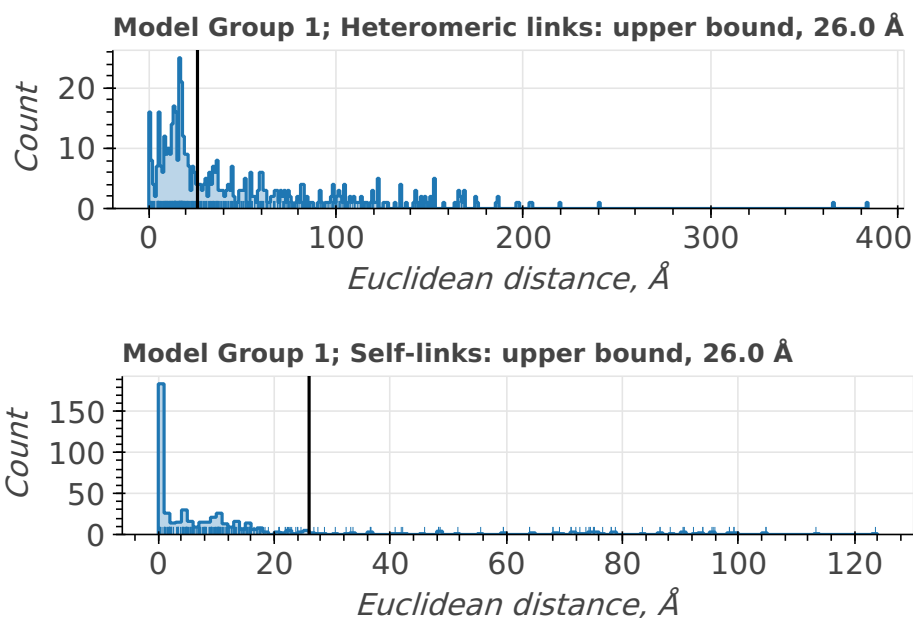
There are 1080 crosslinking restraints combined in 618 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	coarse-grained	LYS	coarse-grained	upper bound	26.0	925
DSS	LYS	CA	LYS	CA	upper bound	26.0	102
DSS	LYS	coarse-grained	THR	coarse-grained	upper bound	26.0	13
DSS	LYS	coarse-grained	MET	coarse-grained	upper bound	26.0	27
DSS	ASN	coarse-grained	LYS	coarse-grained	upper bound	26.0	2
DSS	ARG	coarse-grained	LYS	coarse-grained	upper bound	26.0	2
DSS	GLN	coarse-grained	LYS	coarse-grained	upper bound	26.0	2

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	ALA	coarse-grained	LYS	coarse-grained	upper bound	26.0	2
DSS	LYS	coarse-grained	VAL	coarse-grained	upper bound	26.0	3
DSS	GLN	coarse-grained	MET	coarse-grained	upper bound	26.0	1
DSS	LYS	CA	MET	CA	upper bound	26.0	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.



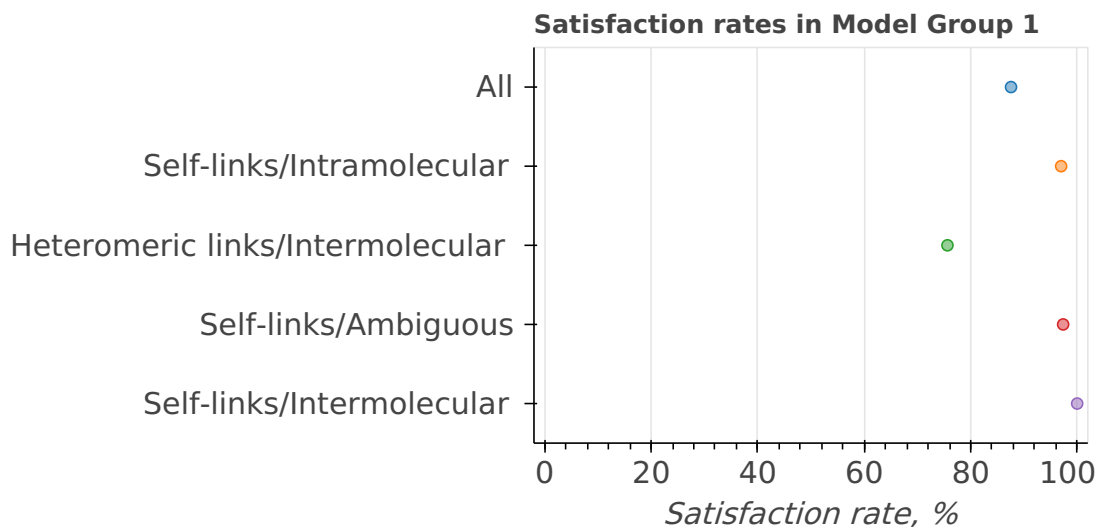
#### Satisfaction of restraints

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=618)
1	1	1	1/5	All	87.58	12.42	612
				Self-links/ Intramolecular	96.99	3.01	299
				Heteromeric links/ Intermolecular	75.65	24.35	271
				Self-links/ Ambiguous	97.37	2.63	38
				Self-links/ Intermolecular	100.00	0.00	4

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.

3DEM volume

Validation for this section is under development.

Mass Spectrometry

Validation for this section is under development.

EM raw micrographs

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.*