

# Integrative Structure Validation Report

July 22, 2024 - 03:37 PM PDT

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

*Python-IHM Version 1.3*

*ATSAS Version 3.2.1 (r14885)*

*Integrative Modeling Validation Version 1.2*

PDB ID	8ZZA
PDB-Dev ID	PDBDEV_00000010
Structure Title	Integrative structure and functional anatomy of a single spoke 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

*This is a PDB-Dev IM Structure Validation Report for a publicly released PDB-Dev entry.*

*We welcome your comments at [pdb-dev@mail.wwpdb.org](mailto:pdb-dev@mail.wwpdb.org)*

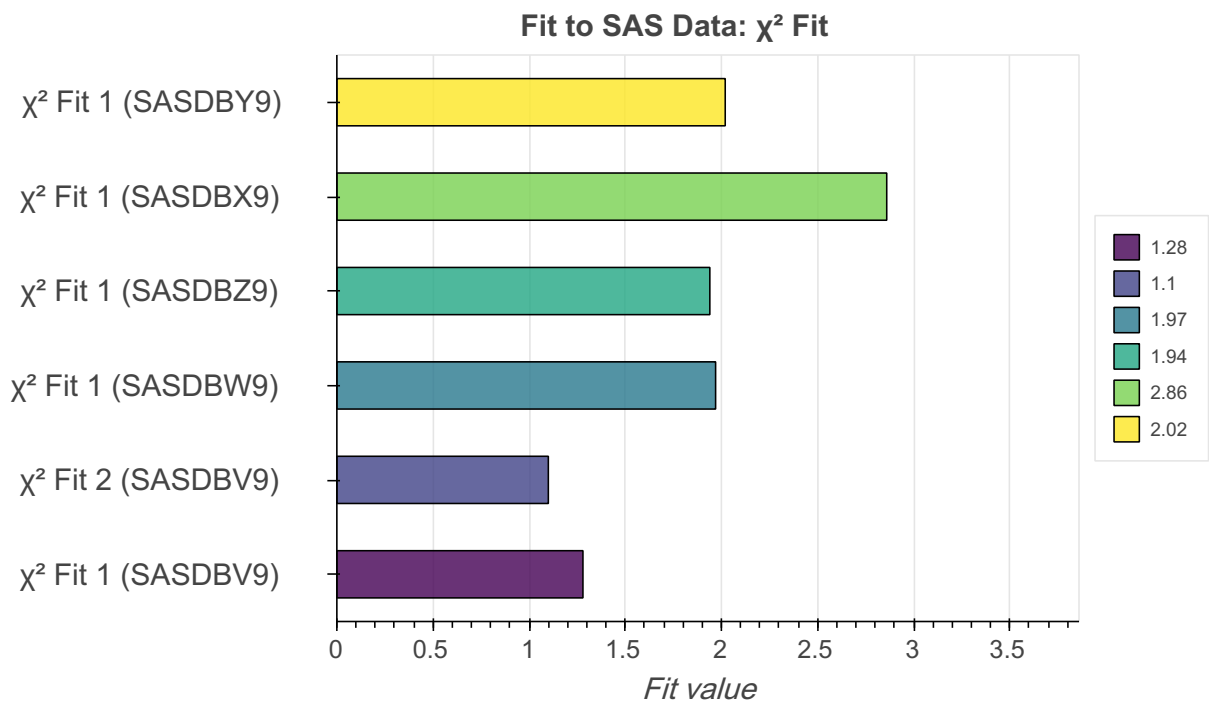
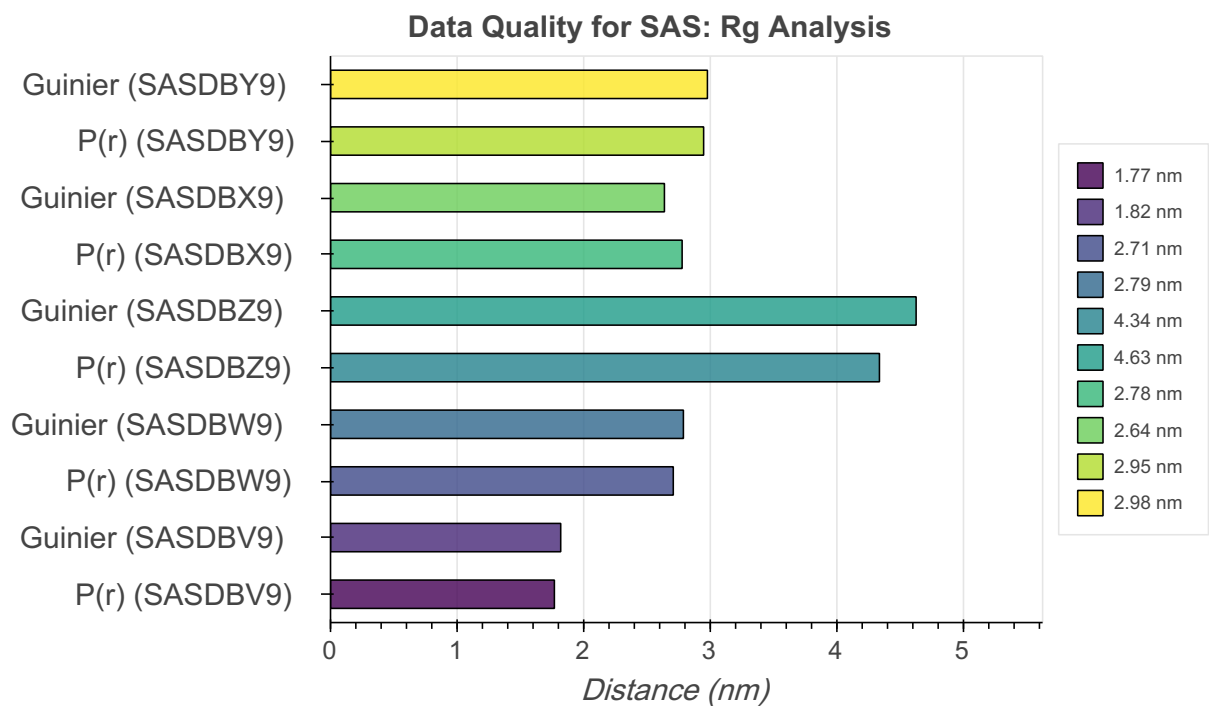
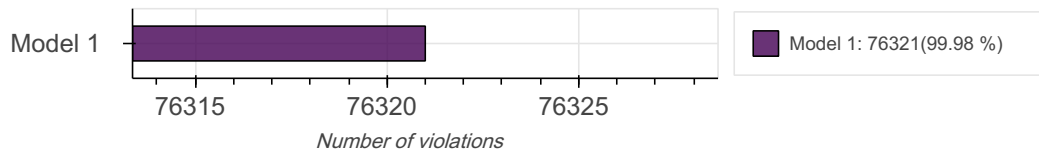
*A user guide is available at [https://pdb-dev.wwpdb.org/validation\\_help.html](https://pdb-dev.wwpdb.org/validation_help.html) with specific help available everywhere you see the  symbol.*

*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 assessment for SAS datasets and fit to model assessments for SAS 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 unique models, with 69 subunits in each model. A total of 65 datasets or restraints were used to build this entry. Each model is represented by 24 rigid bodies and 816 flexible or non-rigid units.*

## Entry composition ?

*There is 1 unique type of models in this entry. This model is titled Scaffold cluster 1/None.*

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	1	1	Nup84	A	A	726
1	2	1	Nup84	H	H	726
1	3	2	Nup85	B	B	744
1	4	2	Nup85	I	I	744
1	5	3	Nup120	C	C	1037
1	6	3	Nup120	J	J	1037
1	7	4	Nup133	D	D	1157
1	8	4	Nup133	K	K	1157
1	9	5	Nup145c	E	E	712
1	10	5	Nup145c	L	L	712
1	11	6	Seh1	F	F	349
1	12	6	Seh1	M	M	349
1	13	7	Sec13	G	G	297
1	14	7	Sec13	N	N	297

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	15	8	Dyn2	O	O	92
1	16	8	Dyn2	P	P	92
1	17	9	Nup82	Q	Q	713
1	18	9	Nup82	R	R	713
1	19	10	Nup159	S	S	1460
1	20	10	Nup159	T	T	1460
1	21	11	Nsp1	U	U	823
1	22	11	Nsp1	V	V	823
1	23	11	Nsp1	X	X	823
1	24	11	Nsp1	AB	AB	823
1	25	11	Nsp1	AF	AF	823
1	26	11	Nsp1	AJ	AJ	823
1	27	12	Nic96	W	W	839
1	28	12	Nic96	AA	AA	839
1	29	12	Nic96	AE	AE	839
1	30	12	Nic96	AI	AI	839
1	31	13	Nup49	Y	Y	472
1	32	13	Nup49	AC	AC	472
1	33	13	Nup49	AG	AG	472
1	34	13	Nup49	AK	AK	472
1	35	14	Nup57	Z	Z	541
1	36	14	Nup57	AD	AD	541
1	37	14	Nup57	AH	AH	541

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	38	14	Nup57	AL	AL	541
1	39	15	Nup157	AM	AM	1391
1	40	15	Nup157	AQ	AQ	1391
1	41	16	Nup170	AN	AN	1502
1	42	16	Nup170	AR	AR	1502
1	43	17	Nup188	AO	AO	1655
1	44	17	Nup188	AS	AS	1655
1	45	18	Nup192	AP	AP	1683
1	46	18	Nup192	AT	AT	1683
1	47	19	Nup53	AU	AU	475
1	48	19	Nup53	AZ	AZ	475
1	49	20	Nup59	AV	AV	528
1	50	20	Nup59	BA	BA	528
1	51	21	Ndc1	AW	AW	655
1	52	21	Ndc1	BB	BB	655
1	53	22	Pom34	AX	AX	299
1	54	22	Pom34	BC	BC	299
1	55	23	Pom152	AY	AY	1337
1	56	23	Pom152	BD	BD	1337
1	57	24	Nup100	BE	BE	959
1	58	24	Nup100	BF	BF	959
1	59	25	Nup116	BG	BG	1113
1	60	25	Nup116	BH	BH	1113

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	61	26	Nup42	BI	BI	430
1	62	27	Gle1	BJ	BJ	538
1	63	28	Nup145	BK	BK	1317
1	64	28	Nup145	BL	BL	1317
1	65	29	Nup1	BM	BM	1076
1	66	30	Nup60	BN	BN	539
1	67	30	Nup60	BO	BO	539
1	68	31	Mlp1	BP	BP	1875
1	69	32	Mlp2	BQ	BQ	1679

### 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	File	10.1016/j.cell.2016.10.028
2	Integrative model	File	10.5281/zenodo.1194547
3	Experimental model	PDB	5CWS
4	Comparative model	File	10.5281/zenodo.1194547
5	Experimental model	PDB	2QX5
6	Experimental model	File	10.5281/zenodo.1194547
7	Experimental model	File	10.5281/zenodo.1194547
8	Comparative model	File	10.5281/zenodo.1194547
9	Comparative model	File	10.5281/zenodo.1194547
10	Comparative model	File	10.5281/zenodo.1194547

<b>ID</b>	<b>Dataset type</b>	<b>Database name</b>	<b>Data access code</b>
11	Comparative model	File	10.5281/zenodo.1194547
12	Comparative model	File	10.5281/zenodo.1194547
13	Comparative model	File	10.5281/zenodo.1194547
14	Integrative model	File	10.1016/j.str.2017.01.006
15	Integrative model	File	10.5281/zenodo.1194547
16	Experimental model	PDB	3NF5
17	Comparative model	File	10.5281/zenodo.1194547
18	Comparative model	File	10.5281/zenodo.1194547
19	Experimental model	PDB	3KEP
20	Experimental model	File	10.5281/zenodo.1194547
21	Mass Spectrometry data	File	10.5281/zenodo.1149746
22	Crosslinking-MS data	File	10.5281/zenodo.1194547
23	Crosslinking-MS data	File	10.5281/zenodo.1194547
24	EM raw micrographs	EMPIAR	EMPIAR-10155
25	3DEM volume	EMDB	EMD-7321
26	3DEM volume	File	10.5281/zenodo.1194547
27	SAS data	File	10.5281/zenodo.1194547
28	SAS data	File	10.5281/zenodo.1194547
29	SAS data	File	10.5281/zenodo.1194547
30	SAS data	File	10.5281/zenodo.1194547
31	SAS data	File	10.5281/zenodo.1194547
32	SAS data	File	10.5281/zenodo.1194547
33	SAS data	File	10.5281/zenodo.1194547

ID	Dataset type	Database name	Data access code
34	SAS data	File	10.5281/zenodo.1194547
35	SAS data	File	10.5281/zenodo.1194547
36	SAS data	File	10.5281/zenodo.1194547
37	SAS data	File	10.5281/zenodo.1194547
38	SAS data	File	10.5281/zenodo.1194547
39	SAS data	File	10.5281/zenodo.1194547
40	SAS data	File	10.5281/zenodo.1194547
41	SAS data	File	10.5281/zenodo.1194547
42	SAS data	File	10.5281/zenodo.1194547
43	SAS data	File	10.5281/zenodo.1194547
44	SAS data	File	10.5281/zenodo.1194547
45	SAS data	File	10.5281/zenodo.1194547
46	SAS data	File	10.5281/zenodo.1194547
47	SAS data	File	10.5281/zenodo.1194547
48	SAS data	SASBDB	SASDBV9
49	SAS data	SASBDB	SASDBW9
50	SAS data	SASBDB	SASDBZ9
51	SAS data	SASBDB	SASDBX9
52	SAS data	SASBDB	SASDBY9
53	SAS data	File	10.5281/zenodo.1194547
54	SAS data	File	10.5281/zenodo.1194547
55	SAS data	File	10.5281/zenodo.1194547
56	SAS data	File	10.5281/zenodo.1194547



ID	Dataset type	Database name	Data access code
57	SAS data	File	10.5281/zenodo.1194547
58	SAS data	File	10.5281/zenodo.1194547
59	SAS data	File	10.5281/zenodo.1194547
60	SAS data	File	10.5281/zenodo.1194547
61	SAS data	File	10.5281/zenodo.1194547
62	SAS data	File	10.5281/zenodo.1194547
63	EM raw micrographs	EMPIAR	EMPIAR-10162
64	2DEM class average	File	10.5281/zenodo.1194547
65	2DEM class average	File	10.5281/zenodo.1194547

## Representation

*This entry has only one representation and includes 24 rigid bodies and 816 flexible units*

Chain ID	Rigid bodies	Non-rigid segments
A	-	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
B	-	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
C	-	1-1, 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, 1037-1037
D	-	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
E	-	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

Chain ID	Rigid bodies	Non-rigid segments
F	-	1-248, 249-287, 288-346, 347-349
G	-	1-9, 10-158, 159-165, 166-296, 297-297
H	-	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
I	-	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
J	-	1-1, 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, 1037-1037
K	-	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
L	-	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
M	-	1-248, 249-287, 288-346, 347-349
N	-	1-9, 10-158, 159-165, 166-296, 297-297
O	-	1-6, 7-92
P	-	1-6, 7-92
Q	-	1-6, 7-16, 17-22, 23-120, 121-122, 123-452, 453-521, 522-612, 613-624, 625-669, 670-677, 678-713
R	-	1-6, 7-16, 17-22, 23-120, 121-122, 123-452, 453-521, 522-612, 613-624, 625-669, 670-677, 678-713
S	-	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
T	-	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
U	-	601-636, 637-727, 728-741, 742-778, 779-787, 788-823

Chain ID	Rigid bodies	Non-rigid segments
V	-	601-636, 637-727, 728-741, 742-778, 779-787, 788-823
W	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
X	637-727, 742-778, 788-823	601-636, 728-741, 779-787
Y	270-359, 369-407, 433-472	201-269, 360-368, 408-432
Z	287-423, 433-476, 505-540	201-286, 424-432, 477-504, 541-541
AA	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
AB	637-727, 742-778, 788-823	601-636, 728-741, 779-787
AC	270-359, 369-407, 433-472	201-269, 360-368, 408-432

Chain ID	Rigid bodies	Non-rigid segments
AD	287-423, 433-476, 505-540	201-286, 424-432, 477-504, 541-541
AE	-	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
AF	-	601-636, 637-727, 728-741, 742-778, 779-787, 788-823
AG	-	201-269, 270-359, 360-368, 369-407, 408-432, 433-472
AH	-	201-286, 287-423, 424-432, 433-476, 477-504, 505-540, 541-541
AI	-	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
AJ	-	601-636, 637-727, 728-741, 742-778, 779-787, 788-823
AK	-	201-269, 270-359, 360-368, 369-407, 408-432, 433-472
AL	-	201-286, 287-423, 424-432, 433-476, 477-504, 505-540, 541-541
AM	-	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, 1391-1391
AN	-	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
AO	-	1-11, 12-34, 35-39, 40-91, 92-100, 101-123, 124-130, 131-166, 167-173, 174-224, 225-255, 256-282, 283-287, 288-304, 305-317, 318-434, 435-438, 439-479, 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

Chain ID	Rigid bodies	Non-rigid segments
AP	-	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
AQ	-	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, 1391-1391
AR	-	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
AS	-	1-11, 12-34, 35-39, 40-91, 92-100, 101-123, 124-130, 131-166, 167-173, 174-224, 225-255, 256-282, 283-287, 288-304, 305-317, 318-434, 435-438, 439-479, 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
AT	-	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
AU	-	1-247, 248-284, 285-303, 304-360, 361-475
AV	-	1-265, 266-302, 303-345, 346-402, 403-528
AW	-	1-655
AX	-	1-299
AY	-	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
AZ	-	1-247, 248-284, 285-303, 304-360, 361-475
BA	-	1-265, 266-302, 303-345, 346-402, 403-528
BB	-	1-655
BC	-	1-299

Chain ID	Rigid bodies	Non-rigid segments
BD	-	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
BE	816-958	551-815, 959-959
BF	816-958	551-815, 959-959
BG	-	751-965, 966-1111, 1112-1113
BH	-	751-965, 966-1111, 1112-1113
BI	-	-
BJ	-	1-120
BK	459-605	201-458
BL	459-605	201-458
BM	-	1-351
BN	-	1-398
BO	-	1-398
BP	-	238-716
BQ	-	215-690

## 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
-------------	-------------	-------------	-------------	--------------------	---------------------------	----------------------	----------------------

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
1	1	Replica exchange monte carlo	Sampling	None	500	False	True
2	1	Replica exchange monte carlo	Sampling	None	3000	False	True
3	1	Replica exchange monte carlo	Sampling	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>	4.0	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>	2.2	image processing	<a href="http://blake.bcm.edu/emanwiki/EMAN2">http://blake.bcm.edu/emanwiki/EMAN2</a>
9	<a href="#">RELION</a>	1.4	image processing	<a href="https://www2.mrc-lmb.cam.ac.uk/relion/">https://www2.mrc-lmb.cam.ac.uk/relion/</a>

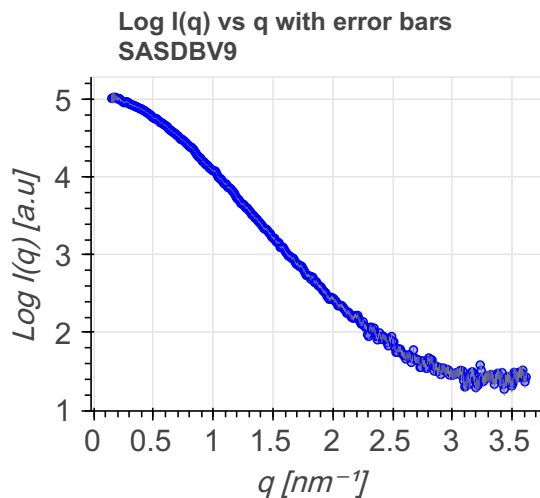
ID	Software name	Software version	Software classification	Software location
10	SGD	Not available	database	<a href="https://www.yeastgenome.org/">https://www.yeastgenome.org/</a>
11	HeliQuest	Not available	helix prediction	<a href="http://heliquest.ipmc.cnrs.fr/">http://heliquest.ipmc.cnrs.fr/</a>
12	MODELLER	9.15	comparative modeling	<a href="https://salilab.org/modeller/">https://salilab.org/modeller/</a>
13	MODELLER	9.13	comparative modeling	<a href="https://salilab.org/modeller/">https://salilab.org/modeller/</a>

### Data quality ?

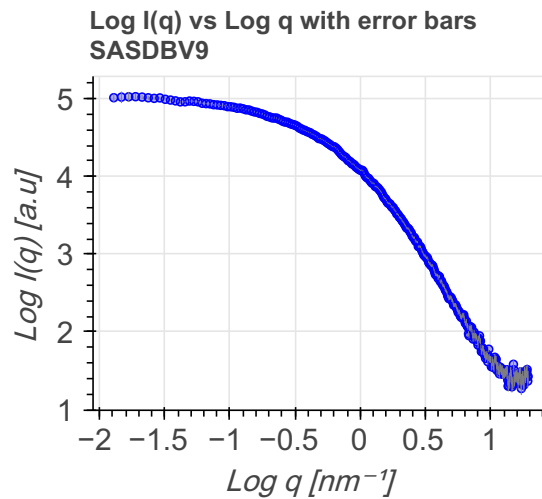
### Scattering profile ?

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

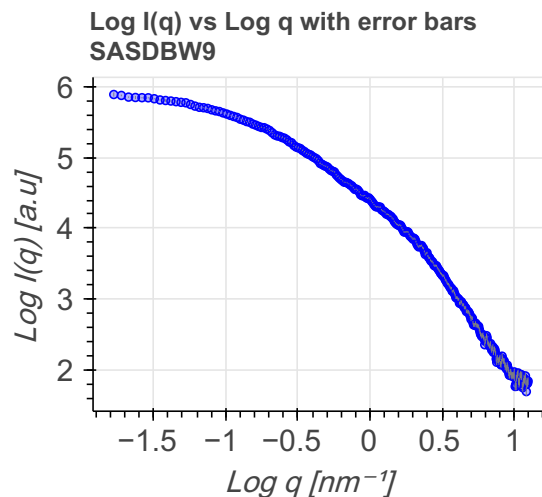
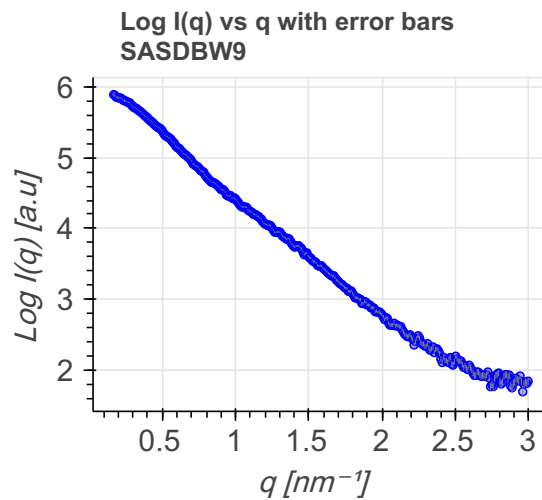
Scattering profile for [SASDBV9](#): data from solutions of biological macromolecules are presented as both  $\log I(q)$  vs  $q$  and  $\log I(q)$  vs  $\log(q)$  based on [SAS validation task force \(SASvtf\) recommendations](#).  $I(q)$  is the intensity (in arbitrary units) and  $q$  is the modulus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



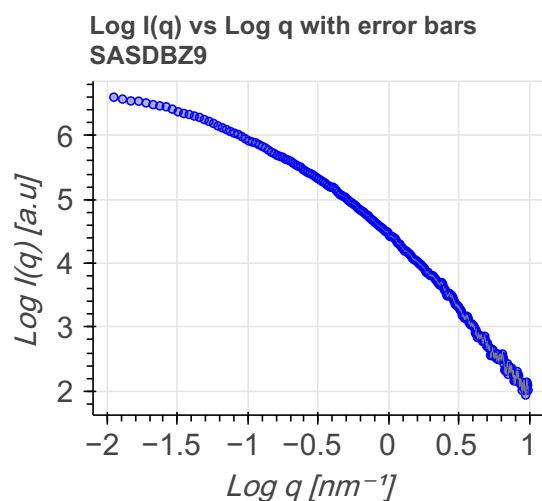
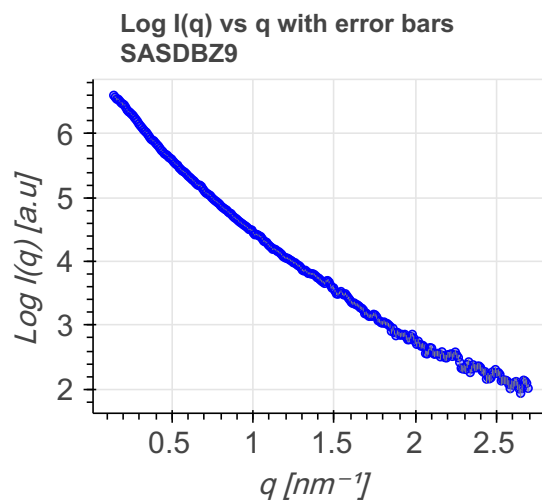




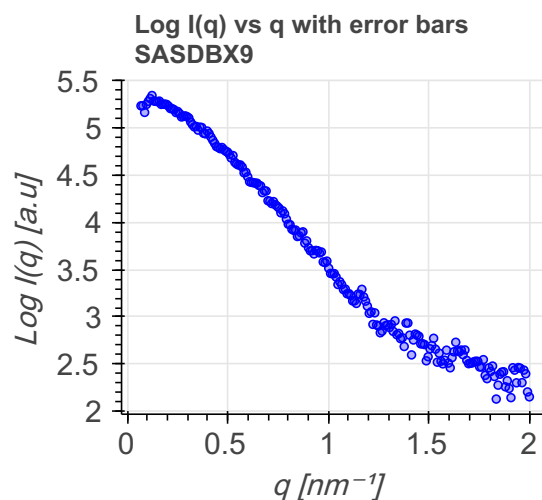
Scattering profile for [SASDBW9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.

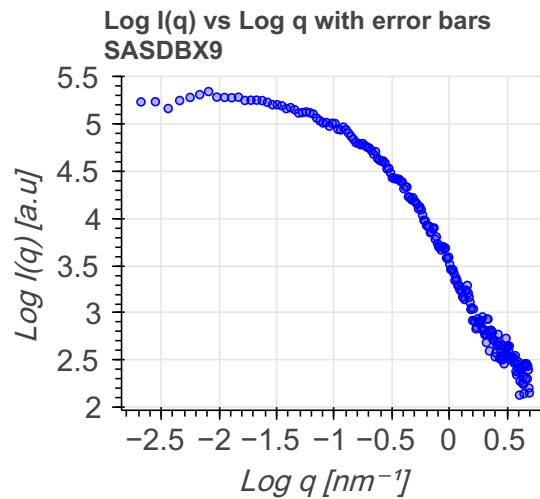


Scattering profile for [SASDBZ9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.

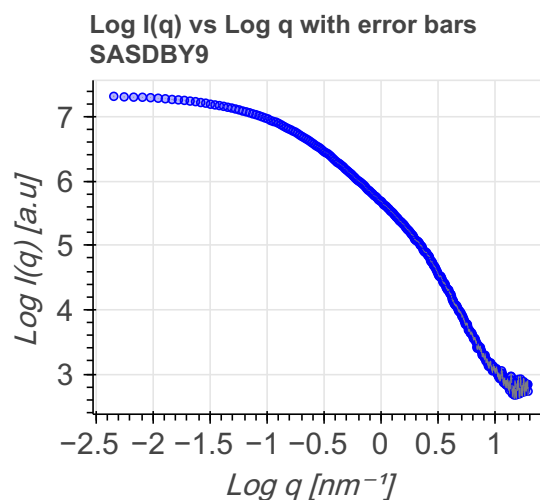
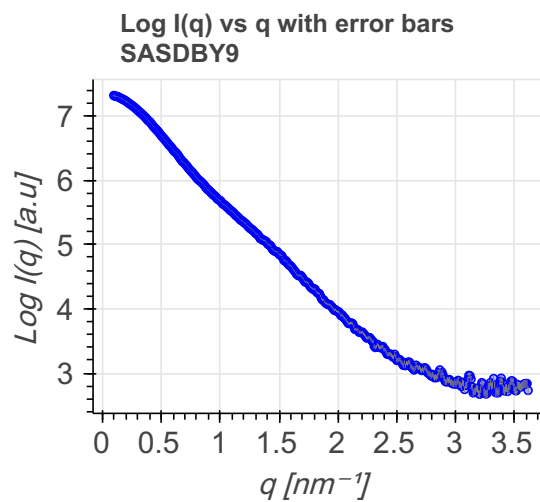


Scattering profile for [SASDBX9](#): data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on [SAS validation task force \(SASvtf\) recommendations](#). I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.





Scattering profile for [SASDBY9](#): 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.



Key experimental estimates

[Molecular weight \(MW\) estimates from experiments and analysis](#) true molecular weight can be compared to the Porod

estimate from scattering profiles.

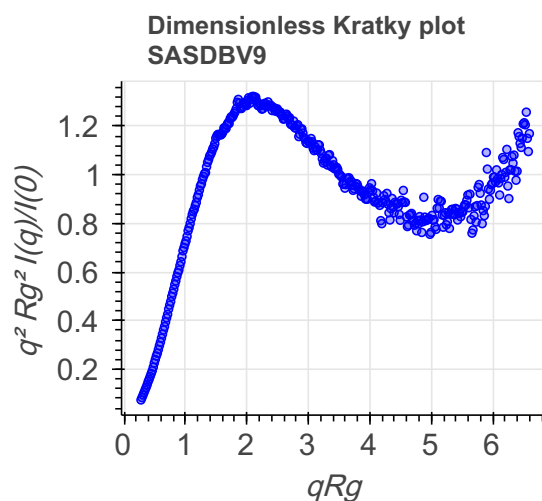
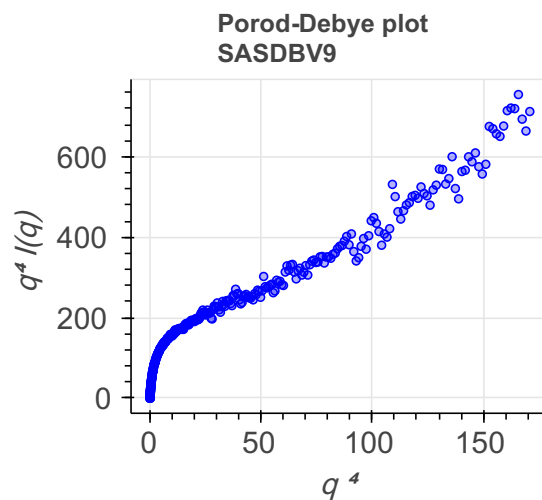
SASDB ID	Chemical composition MW	Standard MW	Porod Volume/MW
SASDBV9	12.6 kDa	12.2 kDa	N/A
SASDBW9	24.1 kDa	25.2 kDa	N/A
SASDBZ9	49.4 kDa	48.3 kDa	N/A
SASDBX9	12.5 kDa	14.7 kDa	N/A
SASDBY9	25.9 kDa	25.2 kDa	N/A

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

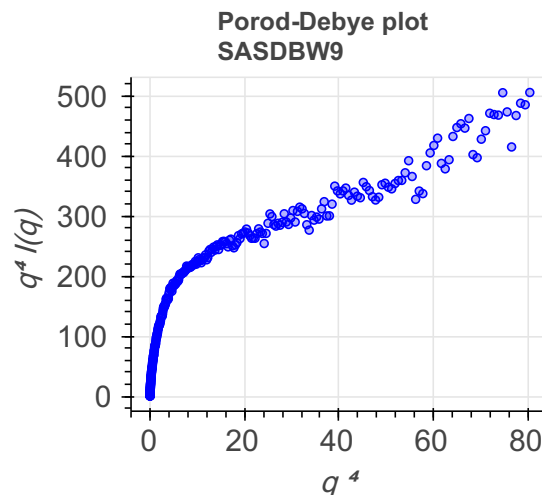
SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDBV9	N/A	17.94 nm <sup>3</sup>	N/A	N/A	N/A
SASDBW9	N/A	22.50 nm <sup>3</sup>	N/A	N/A	N/A
SASDBZ9	N/A	66.59 nm <sup>3</sup>	N/A	N/A	N/A
SASDBX9	N/A	56.68 nm <sup>3</sup>	N/A	N/A	N/A
SASDBY9	N/A	27.97 nm <sup>3</sup>	N/A	N/A	N/A

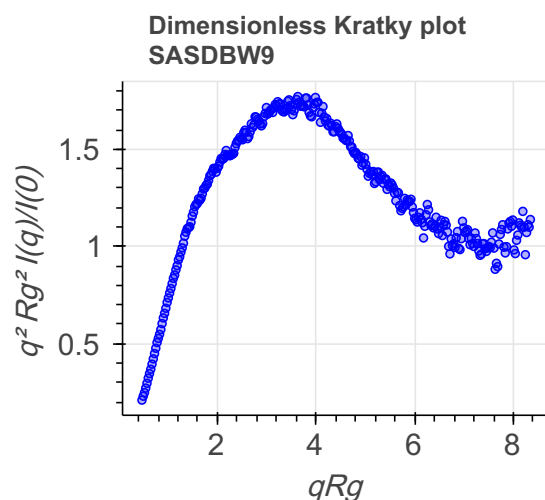
### Flexibility analysis

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.

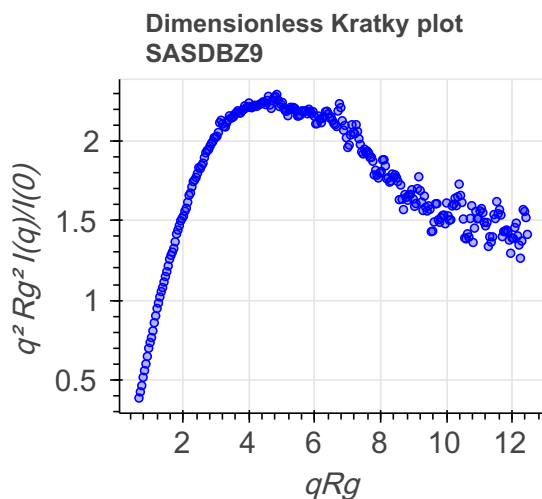
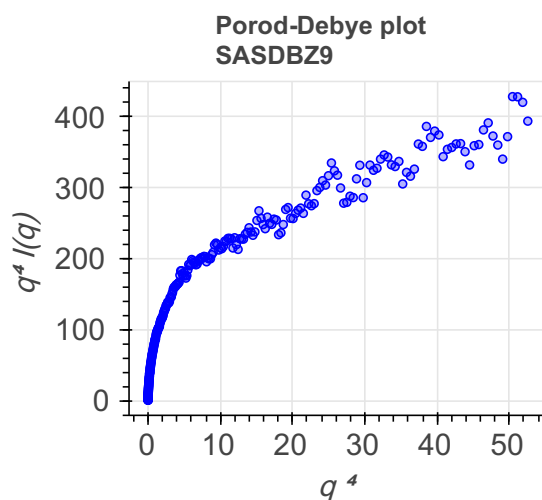


Flexibility analysis for SASDBW9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape is observed for globular (partial or fully folded) domains and a hyperbolic shape is observed for fully unfolded domains.



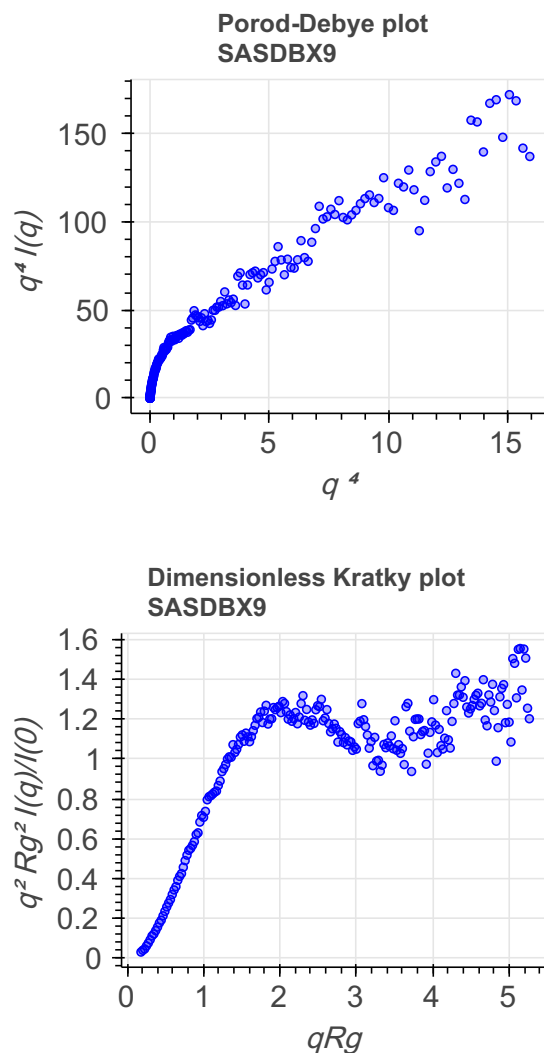


Flexibility analysis for SASDBZ9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape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

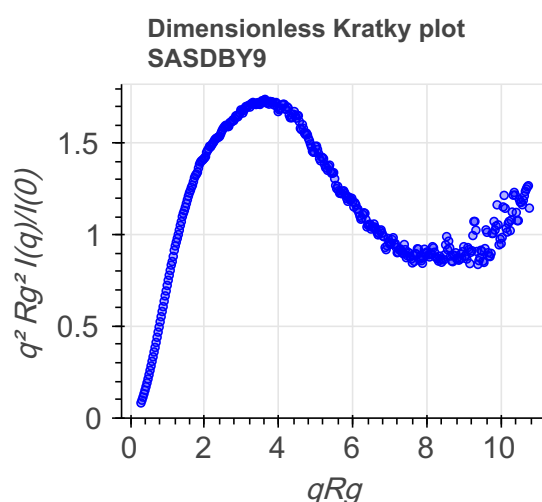
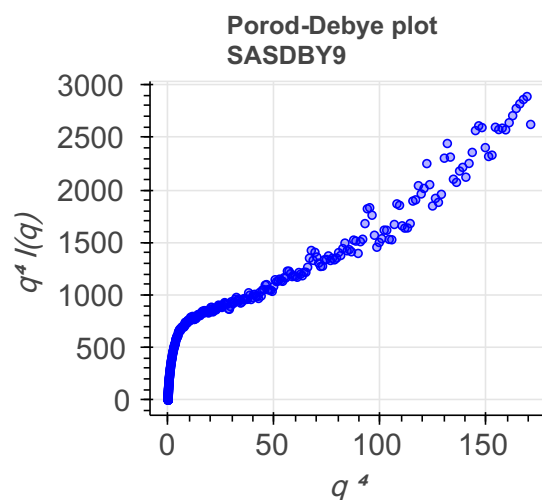


Flexibility analysis for SASDBX9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo](#)

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.



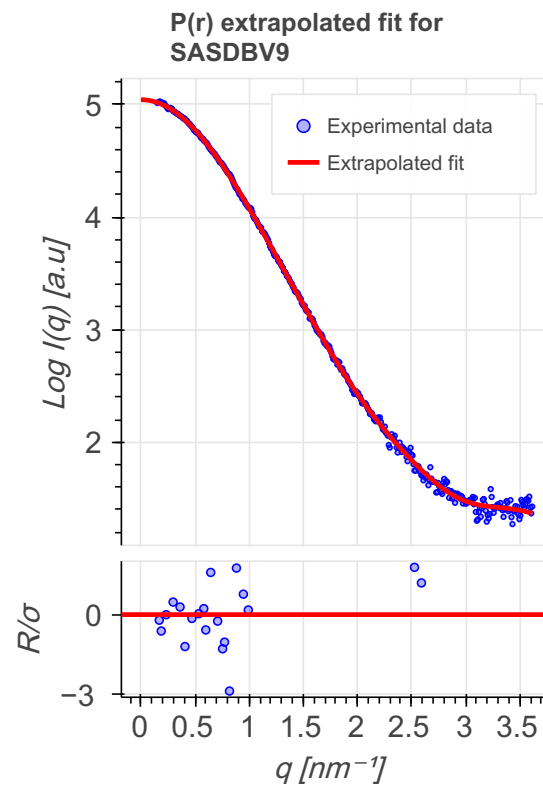
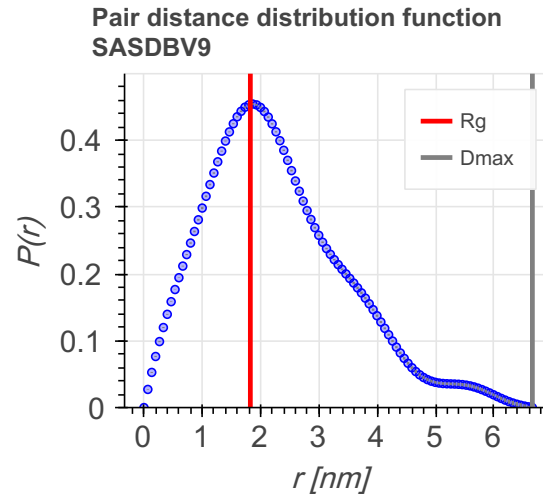
### Pair-distance distribution analysis ?

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.

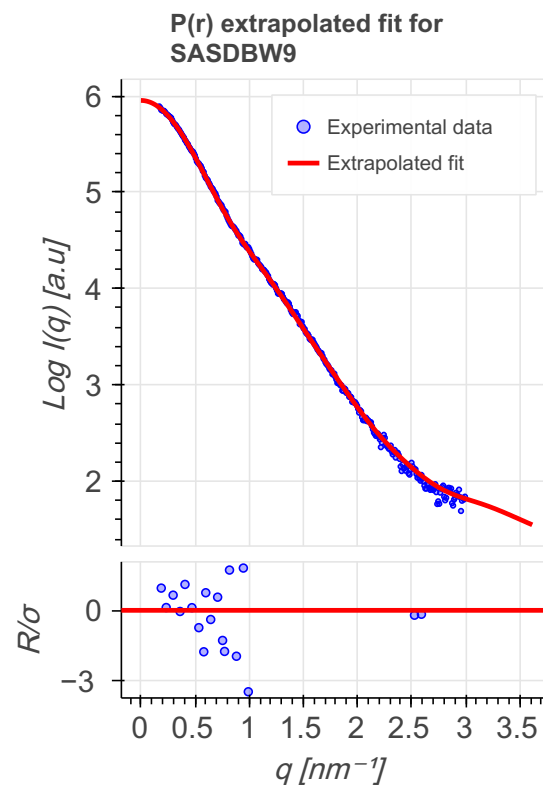
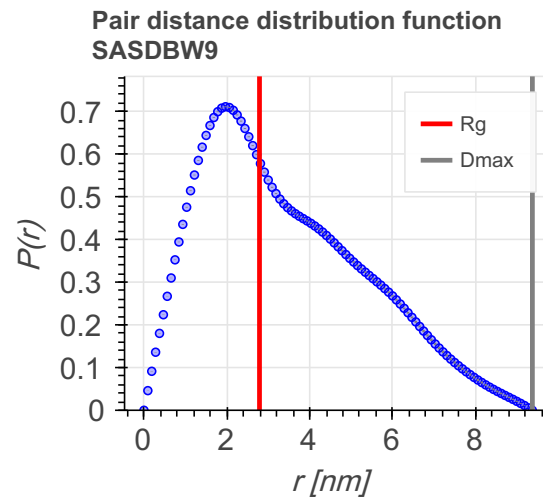
SASDB ID	Software used	Dmax	Dmax error	Rg	Rg error
SASDBV9	GNOM 4.5a	6.660 nm	N/A	1.824 nm	0.006 nm
SASDBW9	GNOM 4.5a	9.370 nm	N/A	2.787 nm	0.007 nm
SASDBZ9	GNOM 4.5a	15.430 nm	N/A	4.629 nm	0.011 nm
SASDBX9	GNOM 4.5a	7.930 nm	N/A	2.636 nm	0.008 nm
SASDBY9	GNOM 4.5a	10.450 nm	N/A	2.976 nm	0.005 nm



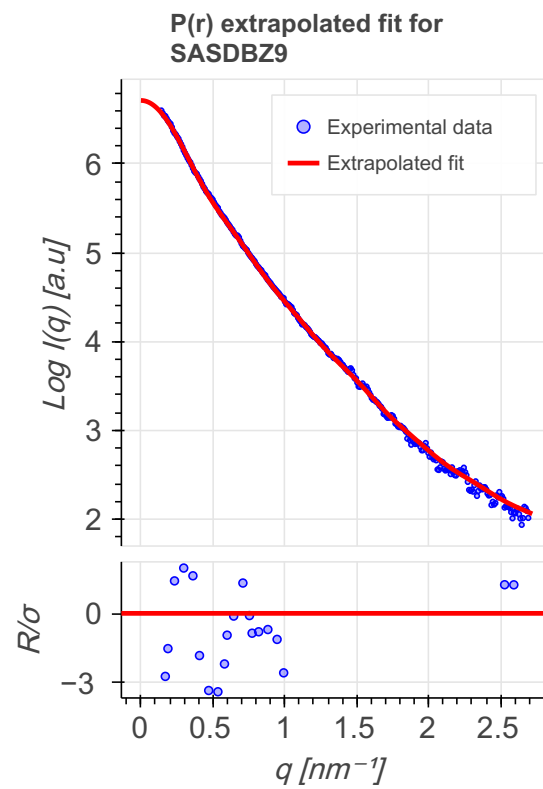
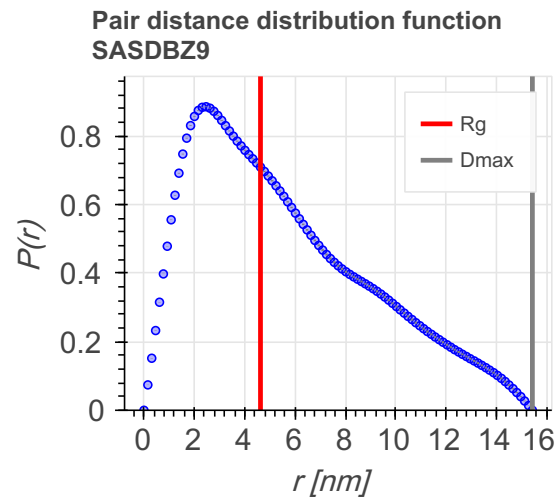
P(r) for SASDBV9: The value of P(r) should be zero beyond  $r=D_{\max}$ .



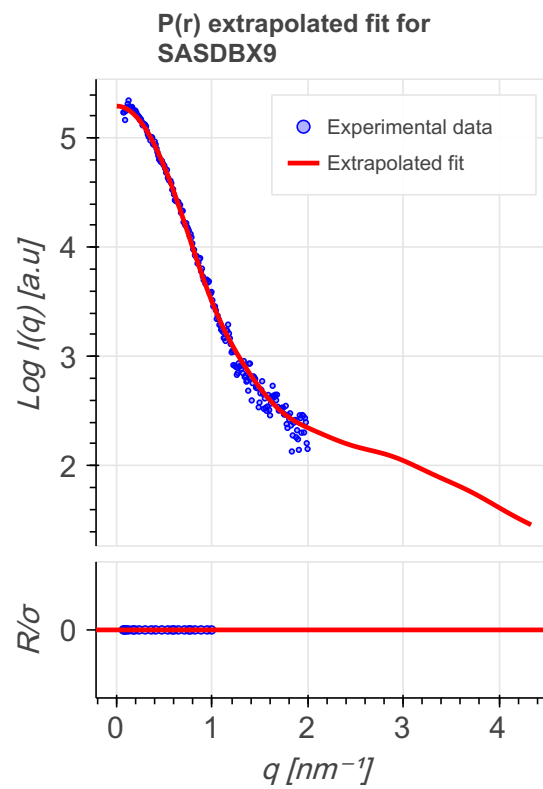
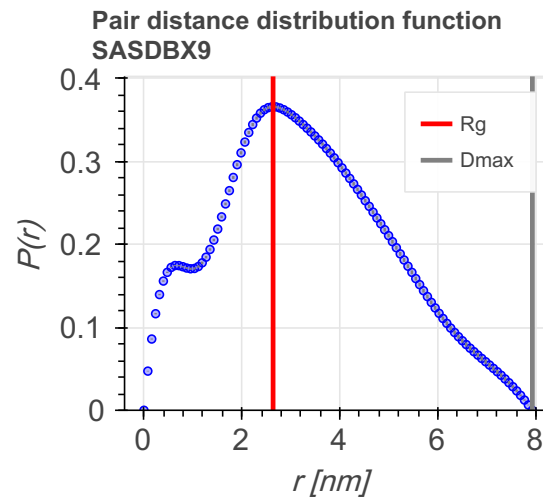
P(r) for SASDBV9: The value of P(r) should be zero beyond  $r=D_{\max}$ .



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

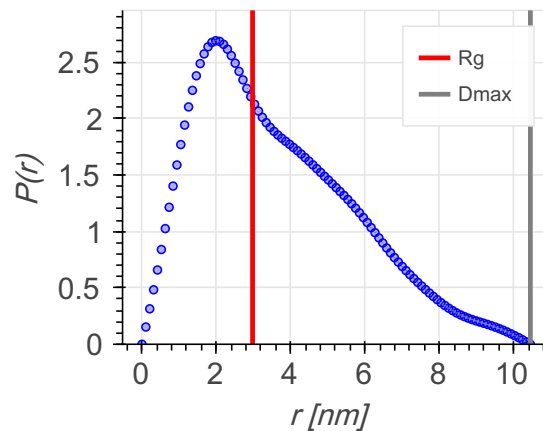


$P(r)$  for SASDBX9: The value of  $P(r)$  should be zero beyond  $r=D_{max}$ .

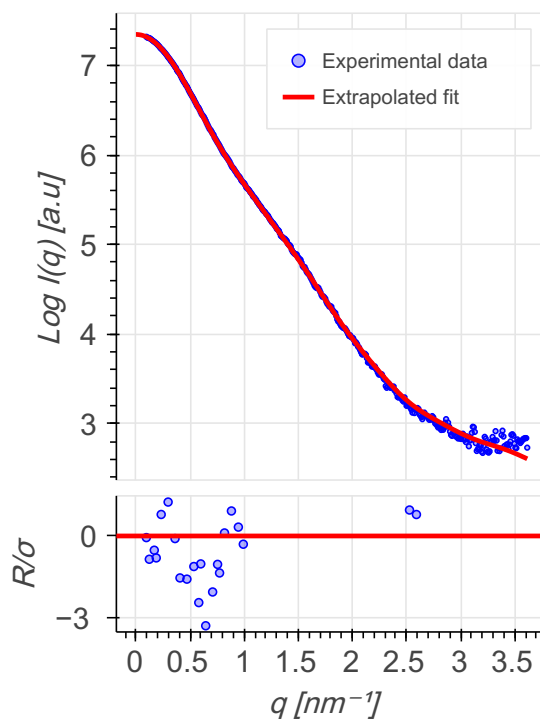


P(r) for SASDBY9: The value of  $P(r)$  should be zero beyond  $r=D_{\text{max}}$ .

Pair distance distribution function  
SASDBY9



P(r) extrapolated fit for  
SASDBY9



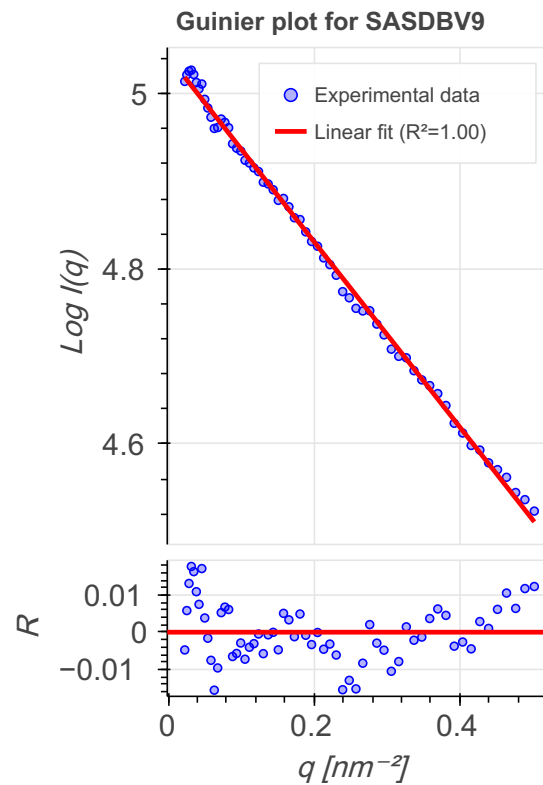
### Guinier analysis ?

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.

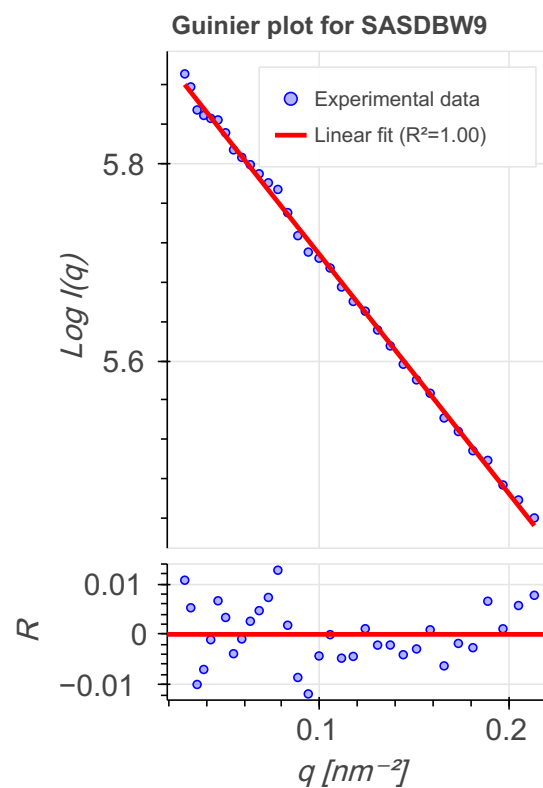
SASDB ID	$R_g$	$R_g$ error	MW	MW error
SASDBV9	1.77 nm	0.05 nm	12.2 kDa	N/A
SASDBW9	2.71 nm	0.06 nm	25.2 kDa	N/A

SASDB ID	Rg	Rg error	MW	MW error
SASDBZ9	4.34 nm	0.17 nm	48.3 kDa	N/A
SASDBX9	2.78 nm	0.18 nm	14.7 kDa	N/A
SASDBY9	2.95 nm	0.11 nm	25.2 kDa	N/A

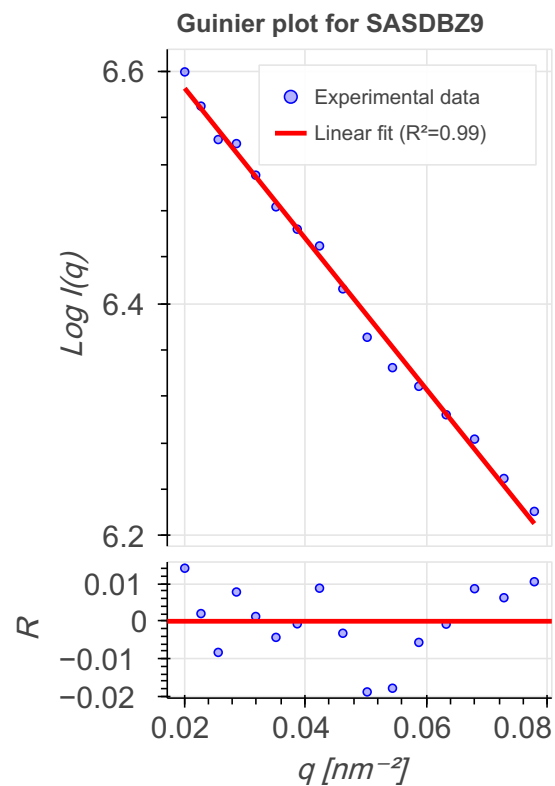
Guinier analysis for SASDBV9: the linearity of the Guinier plot is a sensitive indicator of the quality of the experimental SAS data; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity usually 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.



Guinier analysis for SASDBW9: 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.

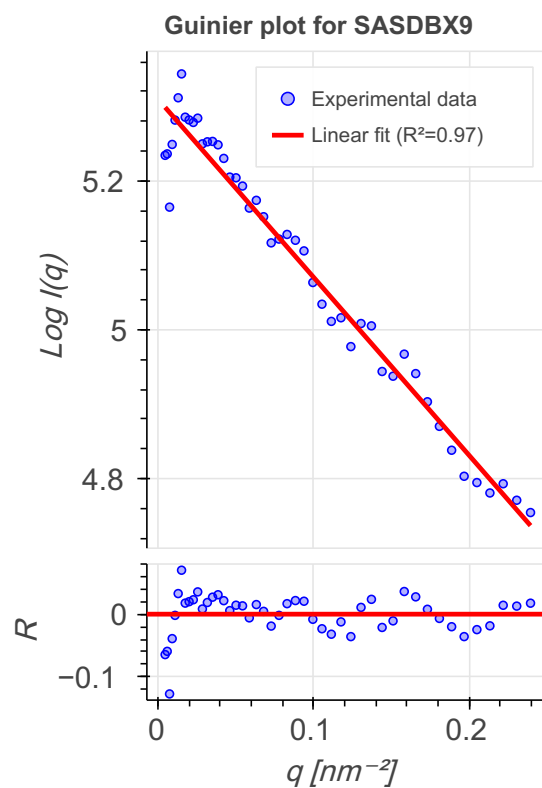


Guinier analysis for SASDBZ9: the linearity of the Guinier plot is a sensitive indicator of the quality of the experimental SAS data; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity usually 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.

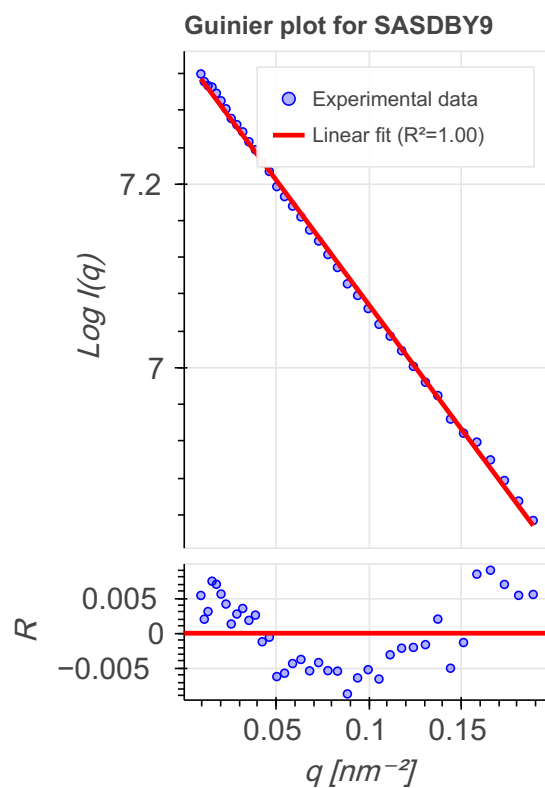


Guinier analysis for SASDBX9: the linearity of the Guinier plot is a sensitive indicator of the quality of the experimental SAS data; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity usually 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.





Guinier analysis for SASDBY9: 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.



### Mass Spectrometry

Validation for this section is under development.

### 2DEM class average

Validation for this section is under development.

### 3DEM volume

Validation for this section is under development.

### Crosslinking-MS

Validation for this section is under development.

### EM raw micrographs

Validation for this section is under development.

## Model quality ?

For models with atomic structures, molprobability 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.*

Models	Excluded Volume Satisfaction (%)	Number of violations

Models	Excluded Volume Satisfaction (%)	Number of violations
1	99.98	76321.0

## Fit of model to data used for modeling ?

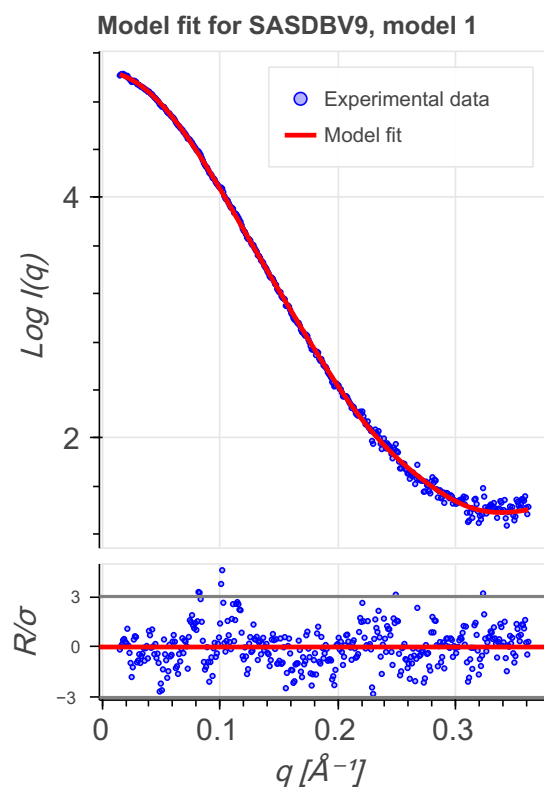
### Fit of model(s) to SAS data

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

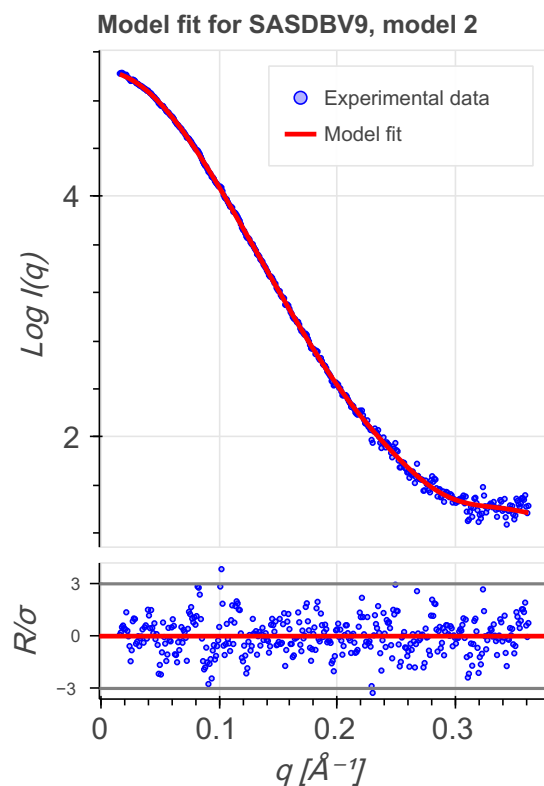
Model and fits displayed below were obtained from SASBDB.  $\chi^2$  values are a measure of fit of the model to data. A perfect fit has a  $\chi^2$  value of 1.0. ATSAS datcmp was used for hypothesis testing. All data sets are similar (i.e. the fit and the data collected) is the null hypothesis. p-value 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	2.22E-02
SASDBV9	2	1.10	5.56E-03
SASDBW9	1	1.97	1.10E-05
SASDBZ9	1	1.94	0.00E+00
SASDBX9	1	2.86	0.00E+00
SASDBY9	1	2.02	3.00E-06

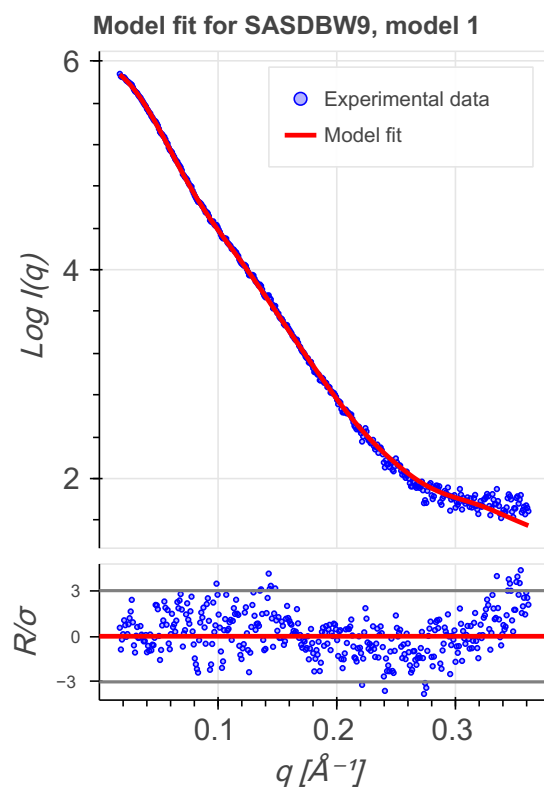
Model fit for SASDBV9 (fit/model number 1): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



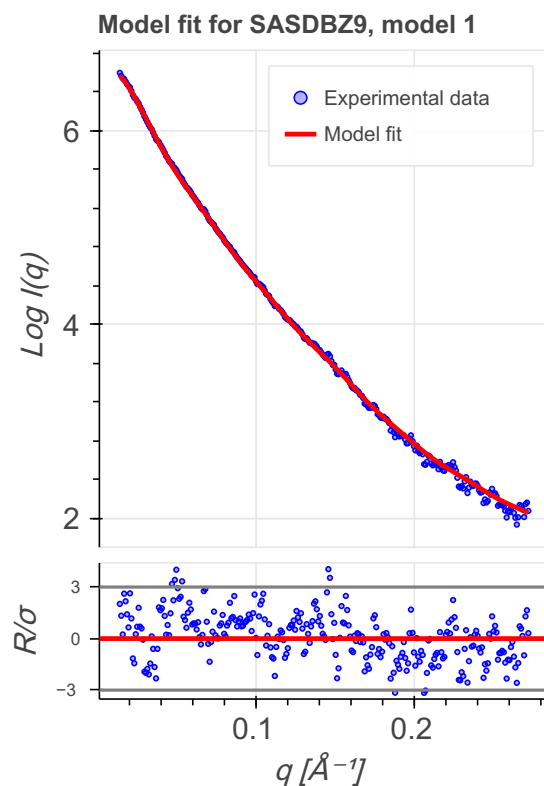
Model fit for SASDBV9 (fit/model number 2): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



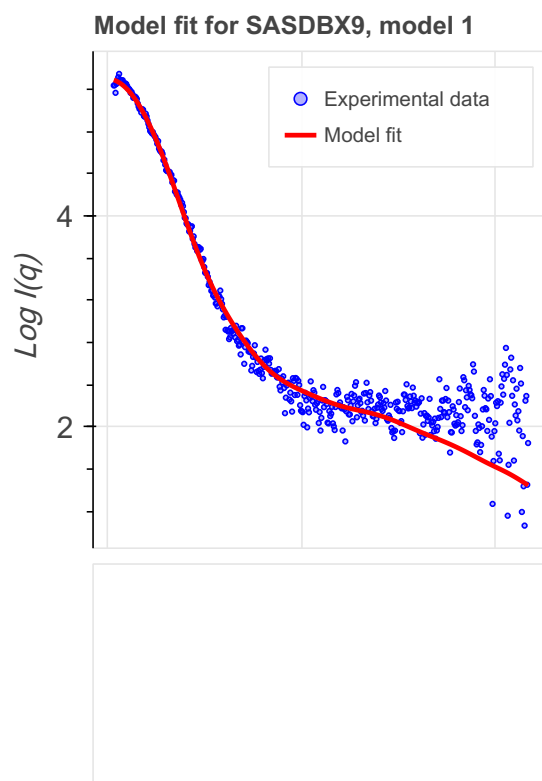
Model fit for SASDBV9 (fit/model number 1): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



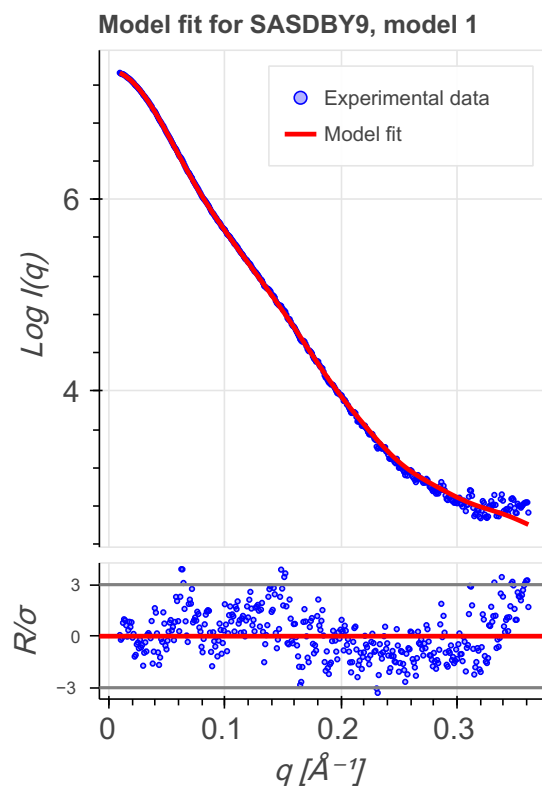
Model fit for SASDBZ9 (fit/model number 1): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



Model fit for SASDBX9 (fit/model number 1): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



Model fit for SASDBY9 (fit/model number 1): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.



### Mass Spectrometry

Validation for this section is under development.

### 2DEM class average

Validation for this section is under development.

### 3DEM volume

Validation for this section is under development.

### Crosslinking-MS

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.

---

### *Acknowledgements*

*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 ABI awards (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250). The [PDB-Dev team](#) and members of [Sali lab](#) contributed model validation metrics and software packages.*

*Implementation of validation methods for SAS data and SAS-based models are funded by [RCSB PDB](#) (grant number DBI-1832184). Dr. Stephen Burley, Dr. John Westbrook, and Dr. Jasmine Young from [RCSB PDB](#), Dr. Jill Trehwella, Dr. Dina Schneidman, and members of the [SASBDB](#) repository are acknowledged for their advice and support in implementing SAS validation methods.*

*Members of the [wwPDB Integrative/Hybrid Methods Task Force](#) provided recommendations and community support for the project.*