

Integrative Structure Validation Report ?

February 18, 2025 - 08:27 AM PST

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

Integrative Modeling Validation Version 2.0

Python-IHM Version 1.8

ATSAS Version 3.2.1 (r14885)

PDB ID	8ZZH
PDB-Dev ID	PDBDEV_00000017
Structure Title	Molecular architecture of the major membrane ring component, Pom152, of the yeast nuclear pore complex
Structure Authors	Upla P; Kim SJ; Sampathkumar P; Dutta K; Cahill SM; Chemmama IE; Williams R; Bonanno JB; Rice WJ; Stokes DL; Cowburn D; Almo SC; Sali A; Rout MP; Fernandez-Martinez J
Deposited on	2018-04-04

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

We welcome your comments at helpdesk@pdb-ihm.org

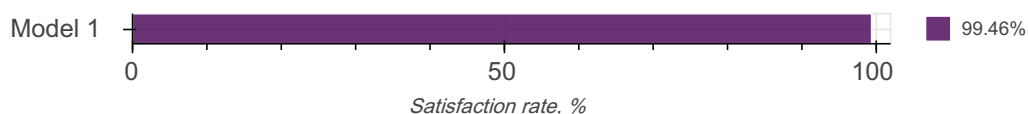
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

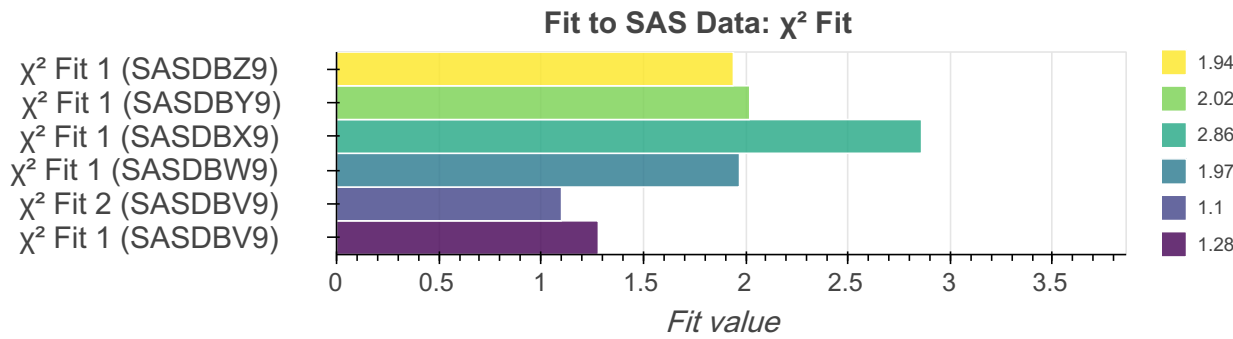
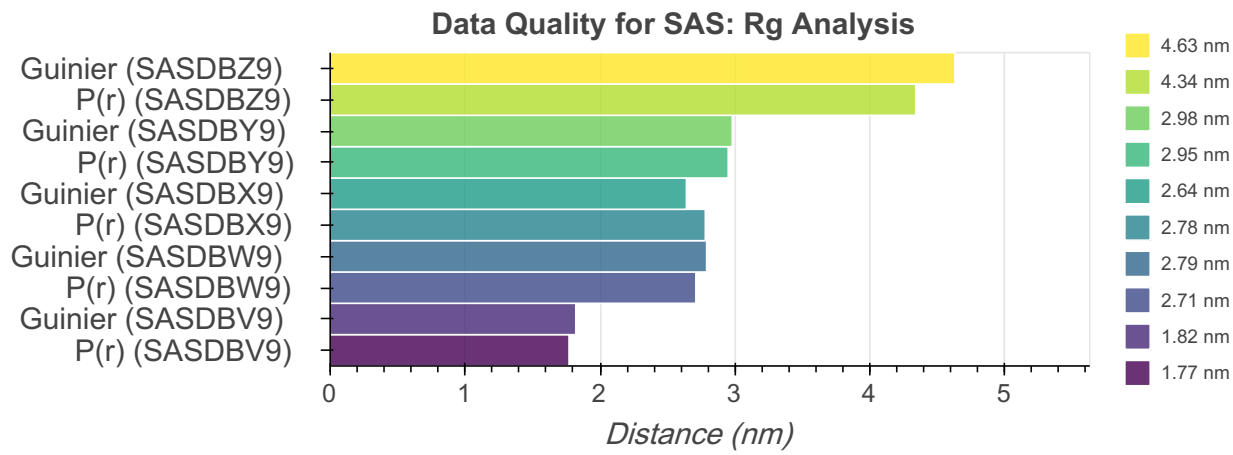
List of references used to build this report is available [here](#).

Overall quality ?

This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: Excluded Volume Analysis





Ensemble information ?

This entry consists of 1 distinct ensemble(s).

Summary ?

This entry consists of 1 model(s). A total of 22 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

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	pom152	A	1337	379-472, 520-611, 616-714, 722-818, 824-918, 931-1026, 1036-1141, 1150-1229, 1244-1337	1-378, 473-519, 612-615, 715-721, 819-823, 919-930, 1027-1035, 1142-1149, 1230-1243	100.00 / 63.80	Multiscale: Coarse-grained: 1 - 100 residue(s) per bead

Datasets used for modeling

There are 22 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	5TVZ
2	Comparative model	Zenodo	10.5281/zenodo.1231518
3	Comparative model	Zenodo	10.5281/zenodo.1231518
4	Comparative model	Zenodo	10.5281/zenodo.1231518
5	Comparative model	Zenodo	10.5281/zenodo.1231518
6	Comparative model	Zenodo	10.5281/zenodo.1231518
7	Comparative model	Zenodo	10.5281/zenodo.1231518
8	3DEM volume	EMDB	EMD-8543
9	3DEM volume	Zenodo	10.5281/zenodo.1231518
10	2DEM class average	Zenodo	10.5281/zenodo.1231518
11	2DEM class average	Zenodo	10.5281/zenodo.1231518
12	2DEM class average	Zenodo	10.5281/zenodo.1231518
13	2DEM class average	Zenodo	10.5281/zenodo.1231518
14	2DEM class average	Zenodo	10.5281/zenodo.1231518
15	2DEM class average	Zenodo	10.5281/zenodo.1231518
16	2DEM class average	Zenodo	10.5281/zenodo.1231518
17	2DEM class average	Zenodo	10.5281/zenodo.1231518
18	SAS data	SASBDB	SASDBV9
19	SAS data	SASBDB	SASDBW9
20	SAS data	SASBDB	SASDBX9
21	SAS data	SASBDB	SASDBY9

ID	Dataset type	Database name	Data access code
22	SAS data	SASBDB	SASDBZ9

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	100000	False	True

There are 3 software packages reported in this entry.

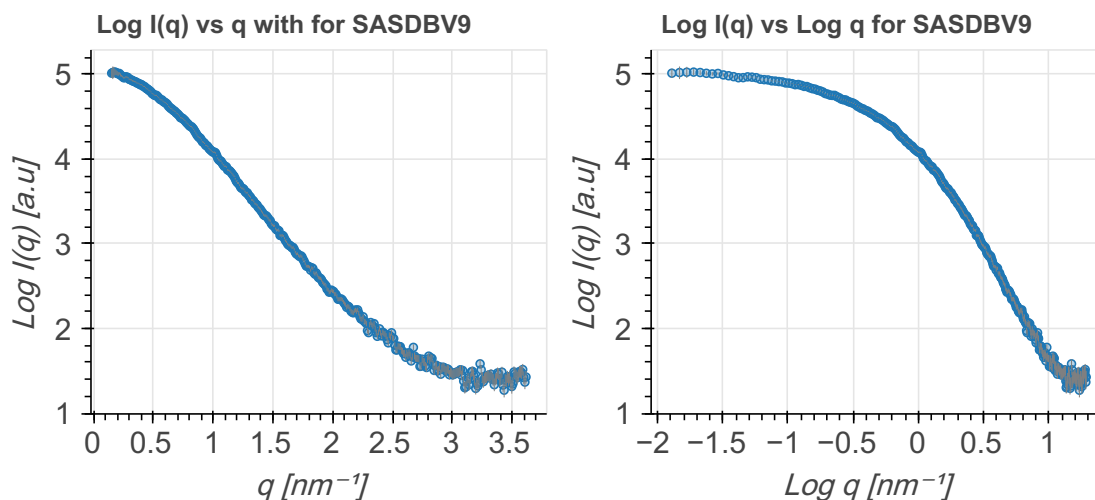
ID	Software name	Software version	Software classification	Software location
1	Integrative Modeling Platform (IMP)	develop-0a5706e202	integrative model building	https://integrativemodeling.org
2	IMP PMI module	67456c0	integrative model building	https://integrativemodeling.org
3	MODELLER	9.13	comparative modeling	https://salilab.org/modeller/

Data quality ?

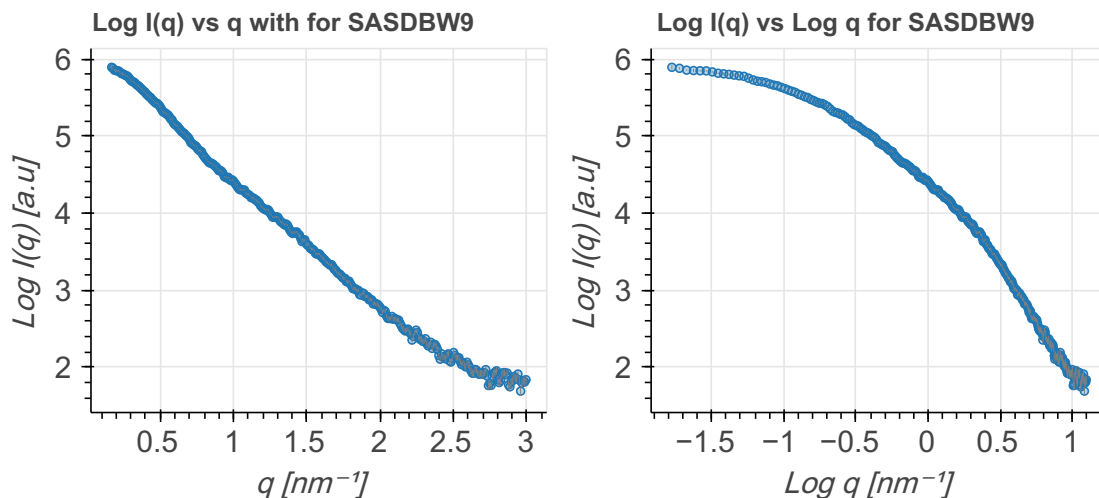
Scattering profile ?

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

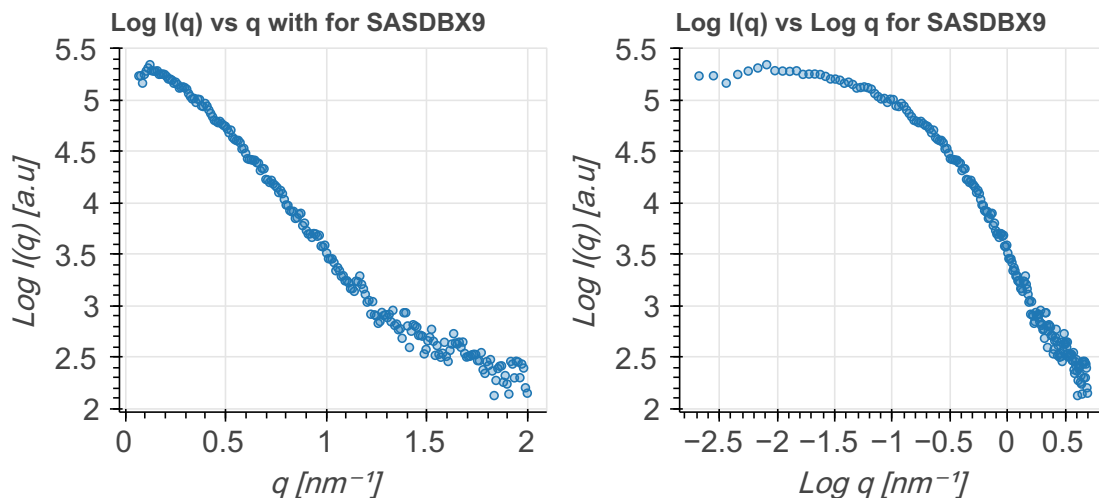
Scattering profile for [SASDBV9](#): data from solutions of biological macromolecules are presented as both $\log I(q)$ vs q and $\log I(q)$ vs $\log(q)$ based on [SAS validation task force \(SASvtf\) recommendations](#). $I(q)$ is the intensity (in arbitrary units) and q is the modulus of the scattering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



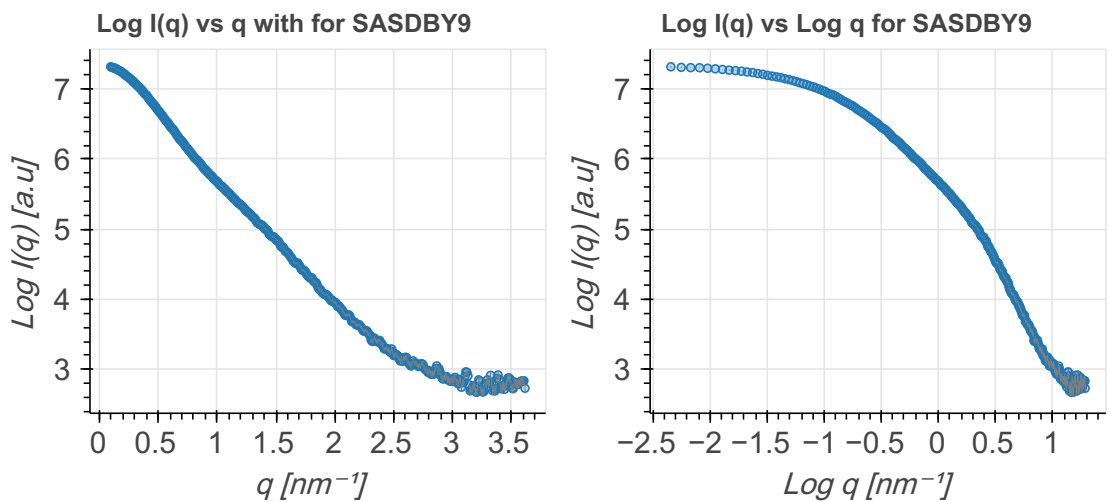
Scattering profile for [SASDBW9](#): data from solutions of biological macromolecules are presented as both $\log I(q)$ vs q and $\log I(q)$ vs $\log(q)$ based on [SAS validation task force \(SASvtf\) recommendations](#). $I(q)$ is the intensity (in arbitrary units) and q is the modulus of the scattering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



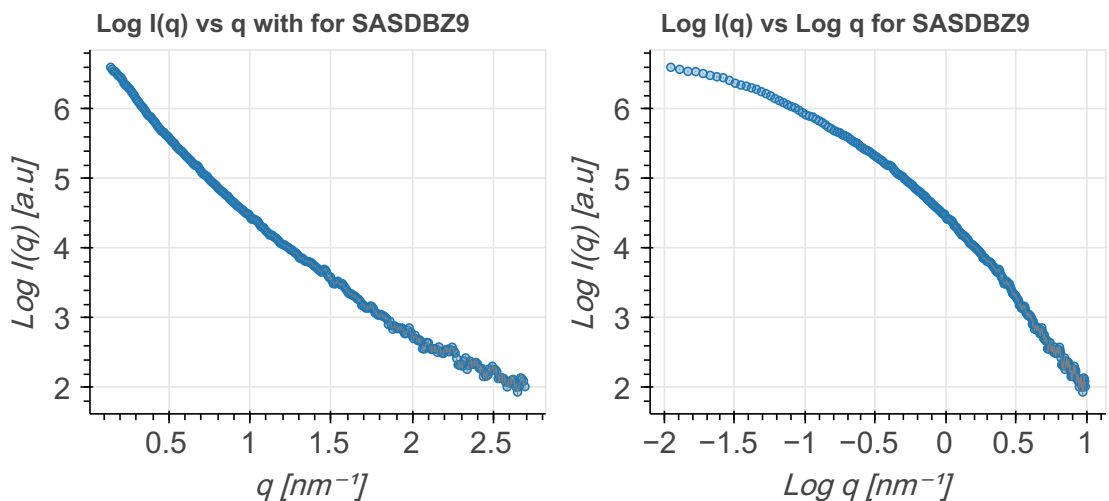
Scattering profile for [SASDBX9](#): data from solutions of biological macromolecules are presented as both $\log I(q)$ vs q and $\log I(q)$ vs $\log(q)$ based on [SAS validation task force \(SASvtf\) recommendations](#). $I(q)$ is the intensity (in arbitrary units) and q is the modulus of the scattering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



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.



Key experimental estimates ?

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

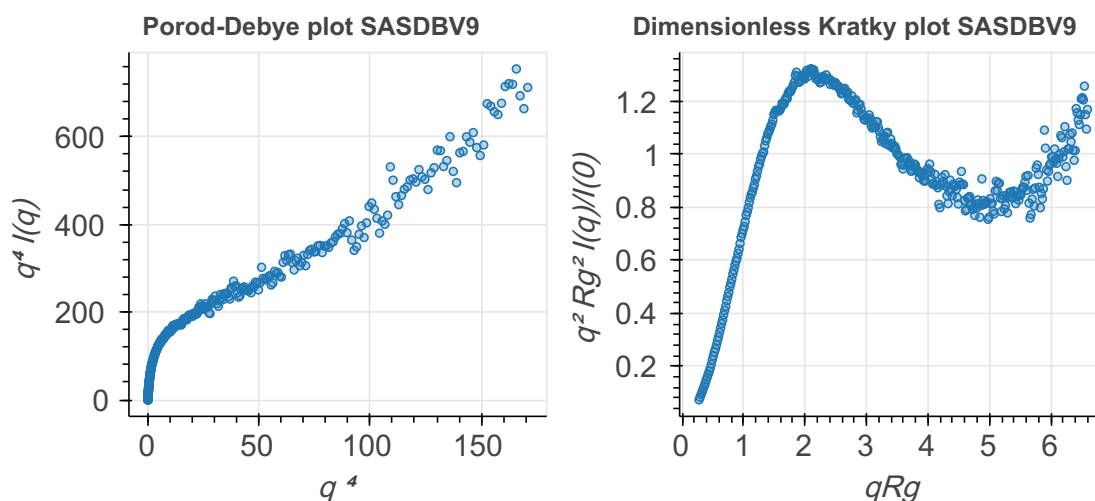
SASDB ID	Chemical composition MW	Standard MW	Porod Volume/MW
SASDBV9	12.6 kDa	12.2 kDa	Not available
SASDBW9	24.1 kDa	25.2 kDa	Not available
SASDBX9	12.5 kDa	14.7 kDa	Not available
SASDBY9	25.9 kDa	25.2 kDa	Not available
SASDBZ9	49.4 kDa	48.3 kDa	Not available

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

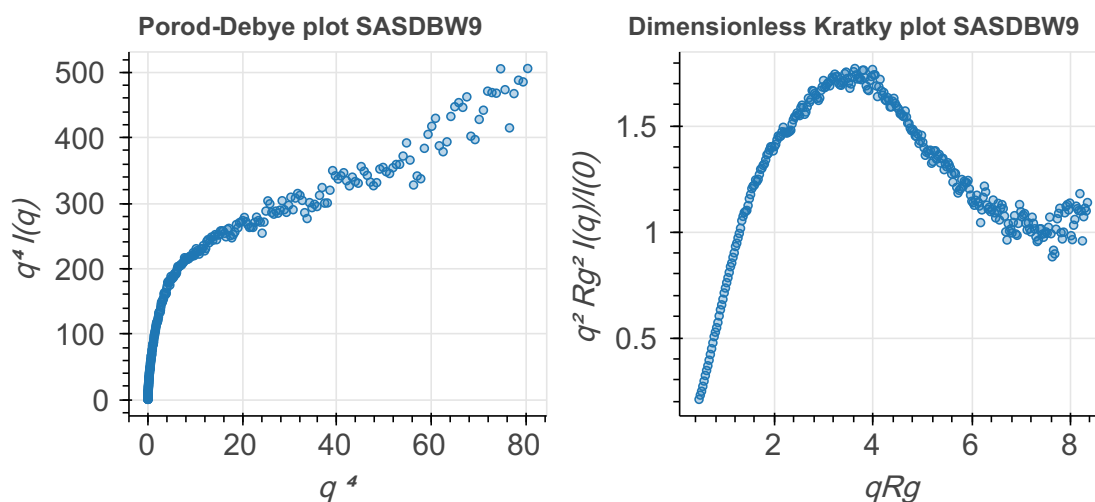
SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDBV9	Not available	17.94 nm ³	Not available	Not available	Not available
SASDBW9	Not available	22.50 nm ³	Not available	Not available	Not available
SASDBX9	Not available	56.68 nm ³	Not available	Not available	Not available
SASDBY9	Not available	27.97 nm ³	Not available	Not available	Not available
SASDBZ9	Not available	66.59 nm ³	Not available	Not available	Not available

Flexibility analysis ?

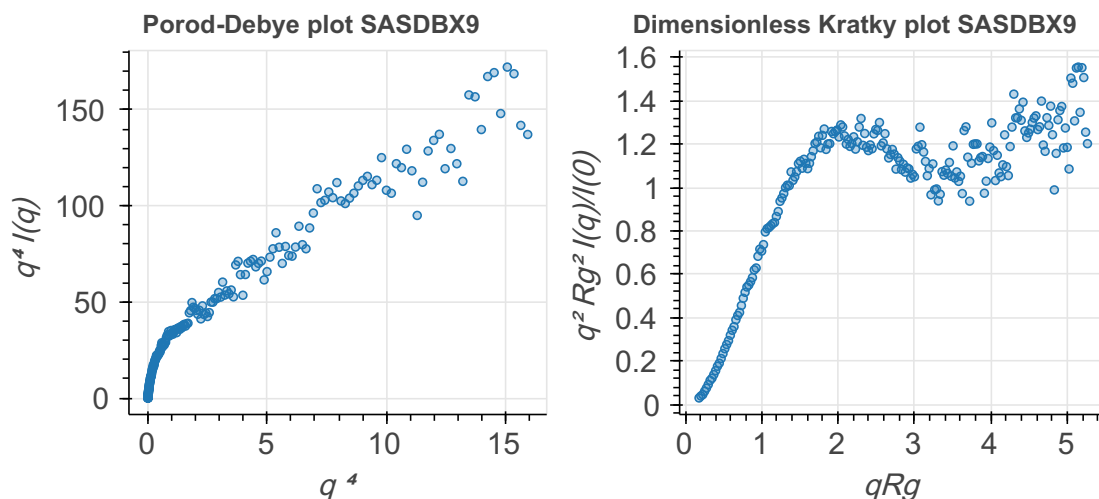
Flexibility analysis for SASDBV9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape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



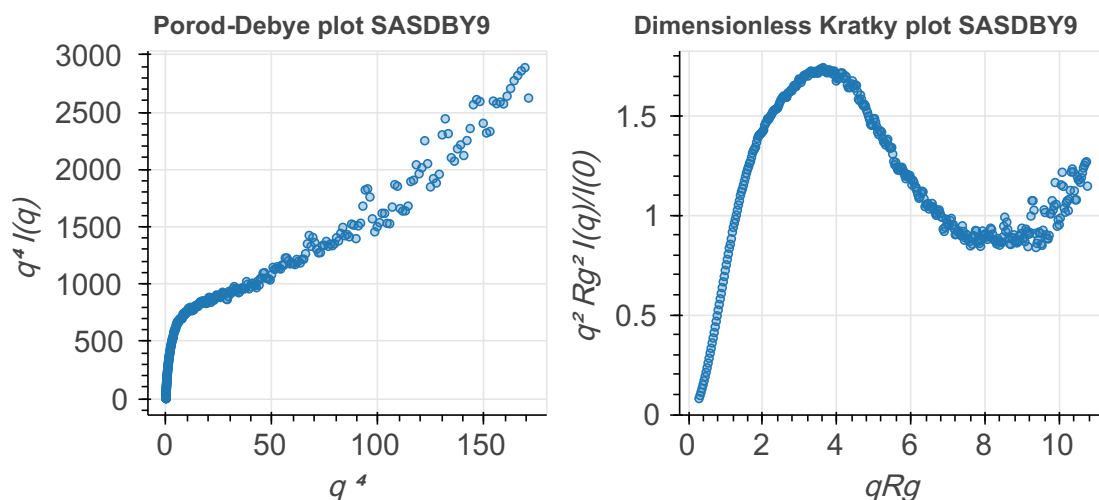
Flexibility analysis for SASDBW9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape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



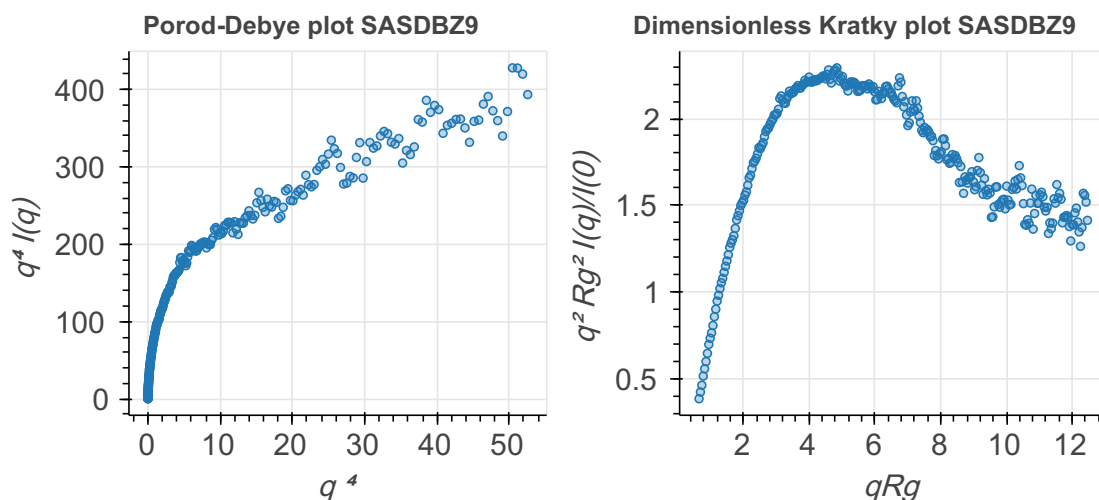
Flexibility analysis for SASDBX9: In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas, fully unfolded domains are devoid of any discernable plateau. For details, refer to Figure 5 in [Rambo and Tainer, 2011](#). In a Kratky plot, a parabolic shape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



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.

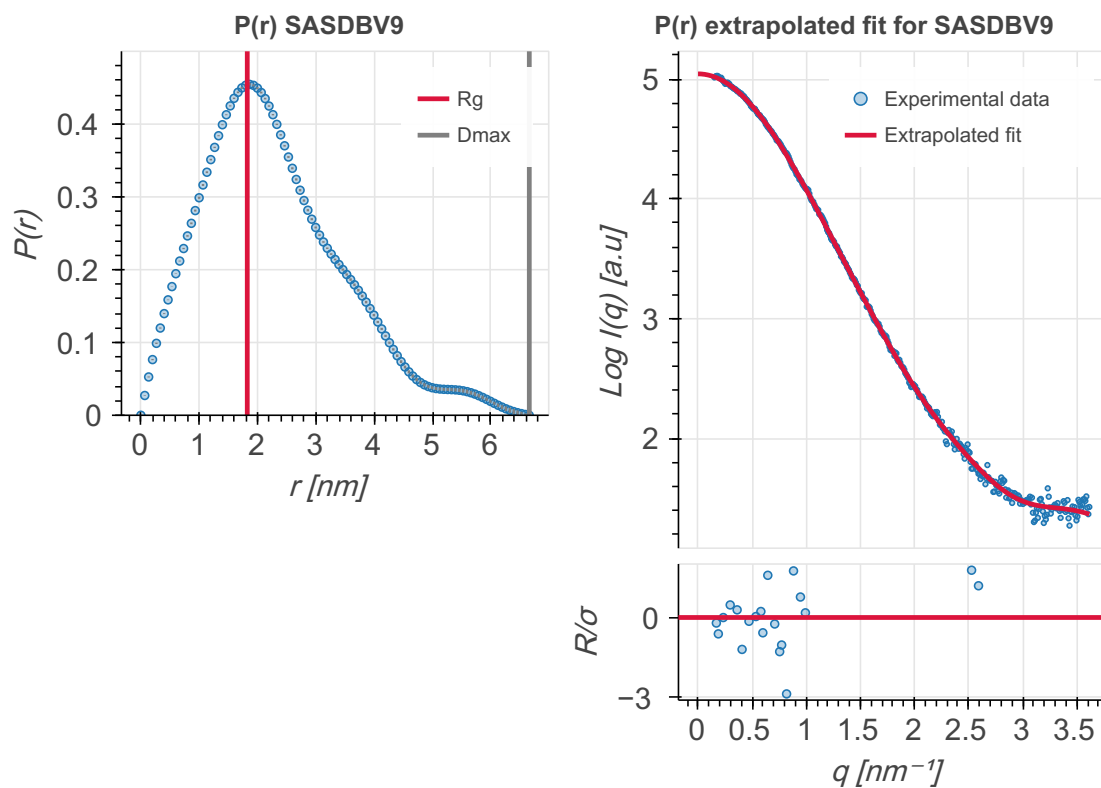


Pair-distance distribution analysis ?

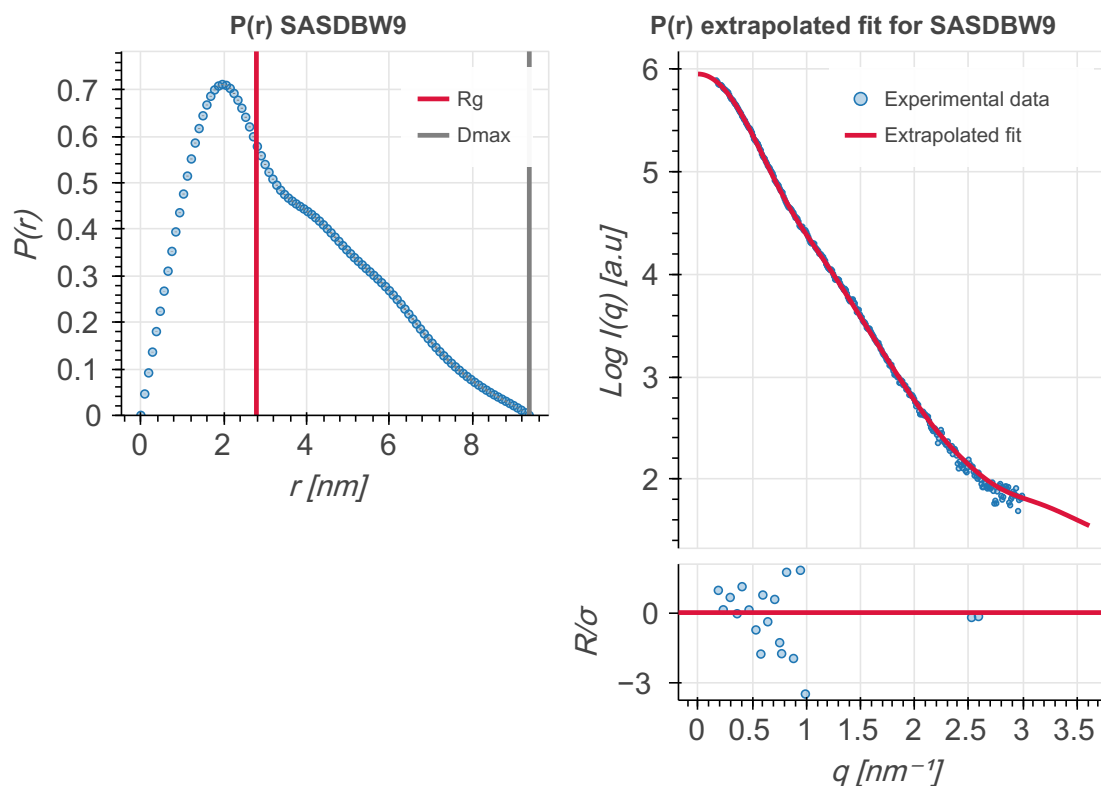
P(r) analysis: P(r) represents the distribution of distances between all pairs of atoms within the particle weighted by the respective electron densities. P(r) is the Fourier transform of I(s) (and vice versa). R_g can be estimated from integrating the P(r) function. Agreement between the P(r) and Guinier-determined 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
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
SASDBZ9	GNOM 4.5a	15.430 nm	Not available	4.629 nm	0.011 nm

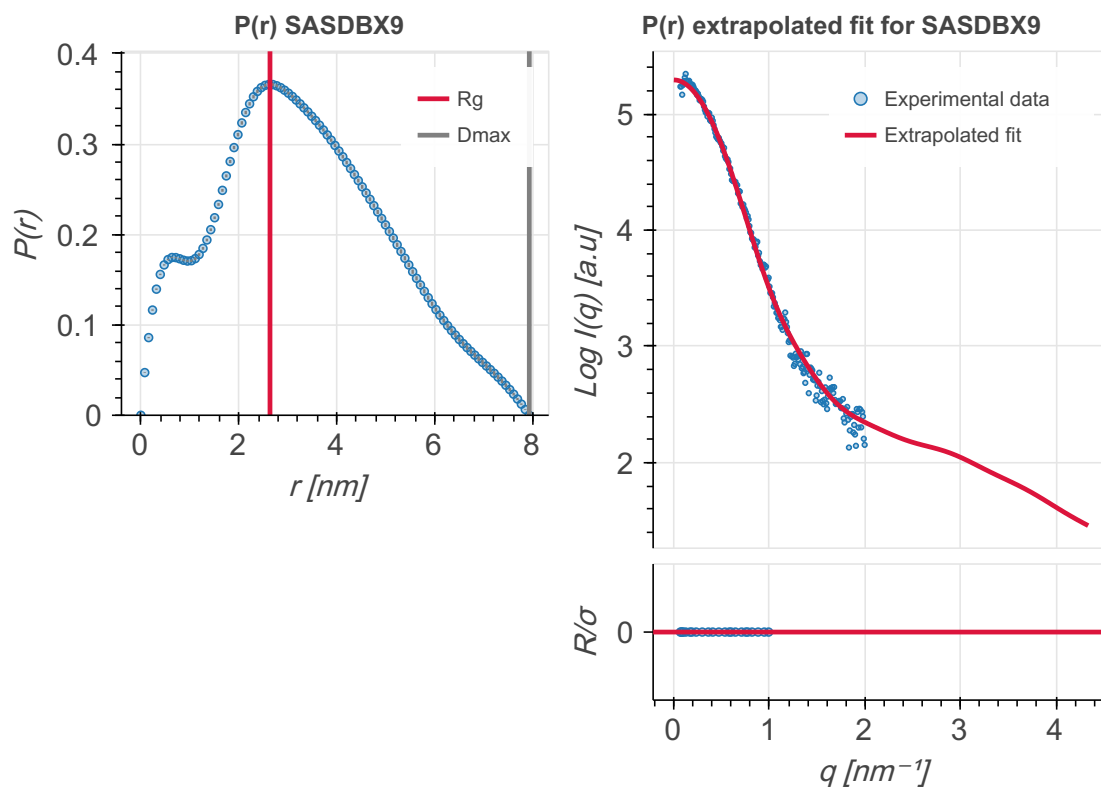
P(r) for SASDBV9: The value of P(r) should be zero beyond $r=D_{max}$.



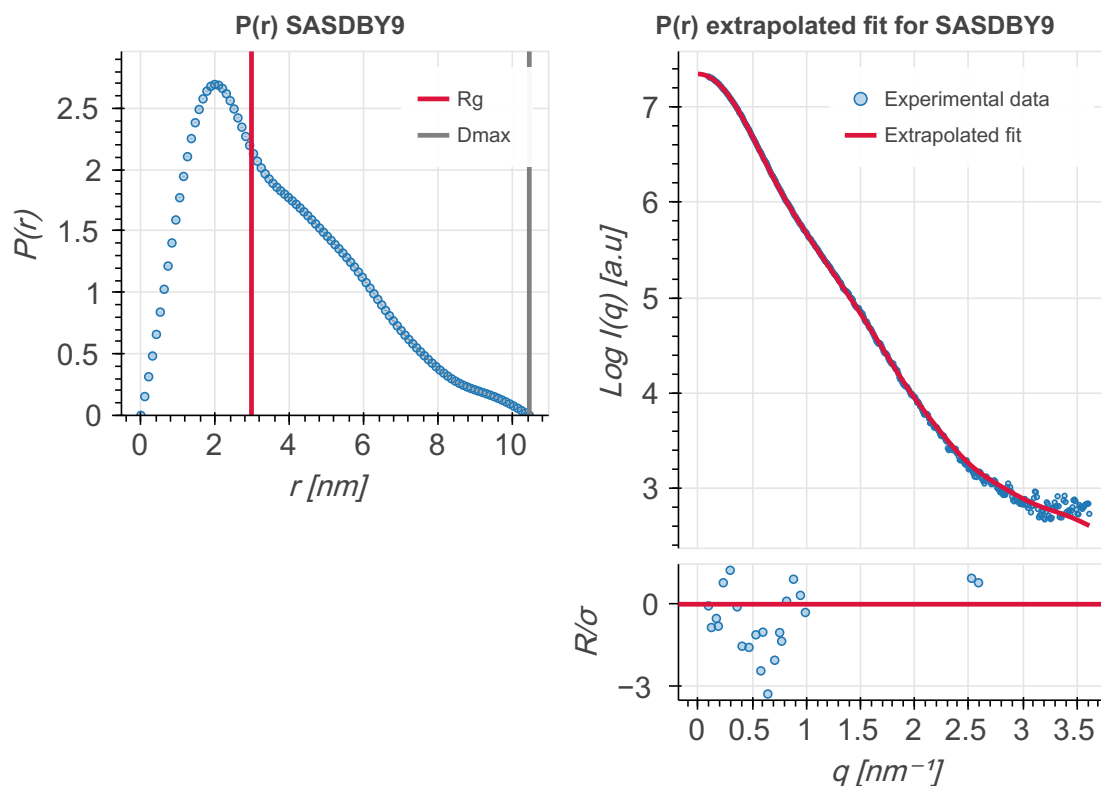
P(r) for SASDBV9: The value of $P(r)$ should be zero beyond $r=D_{max}$.



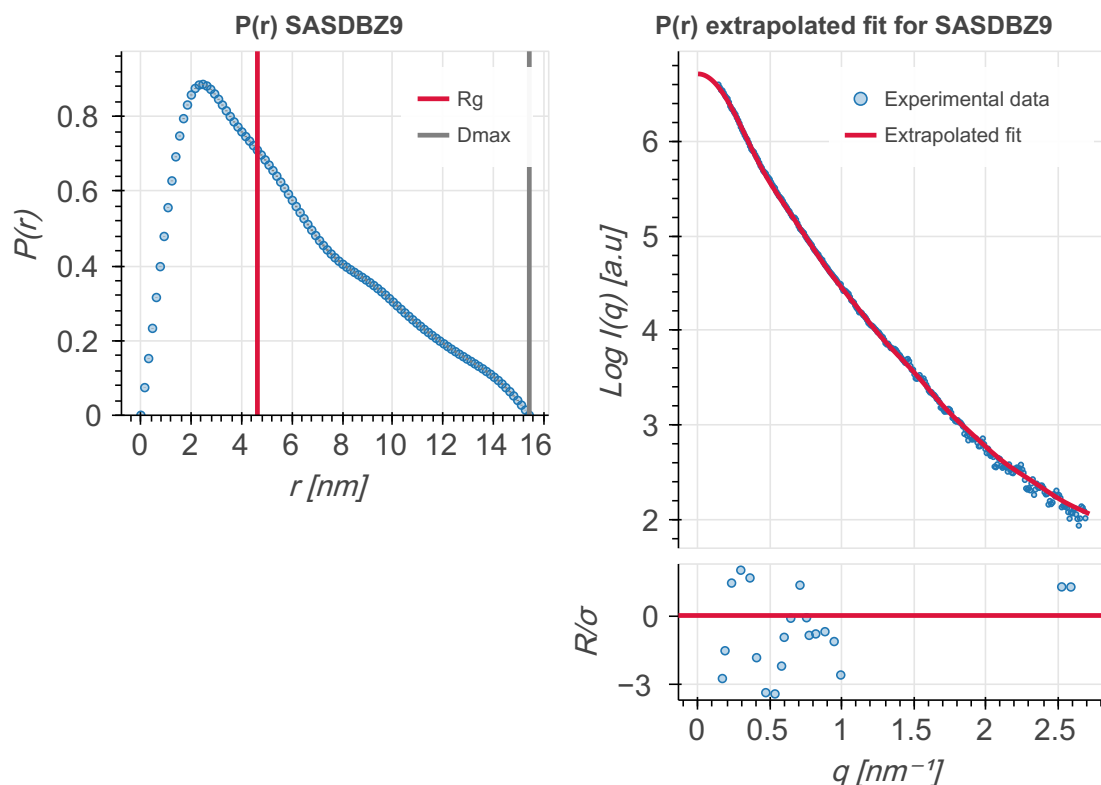
P(r) for SASDBX9: The value of $P(r)$ should be zero beyond $r=D_{max}$.



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



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

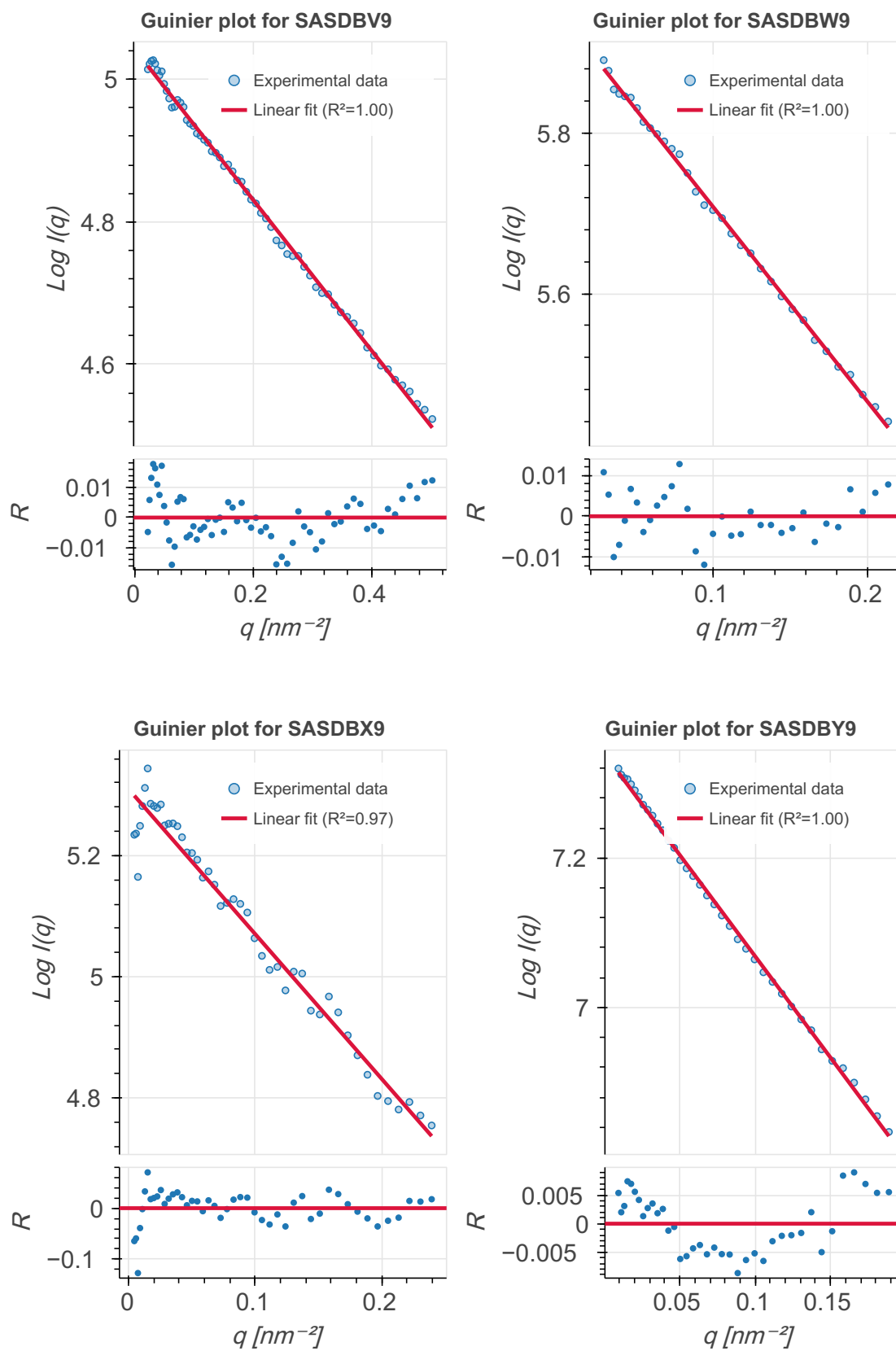


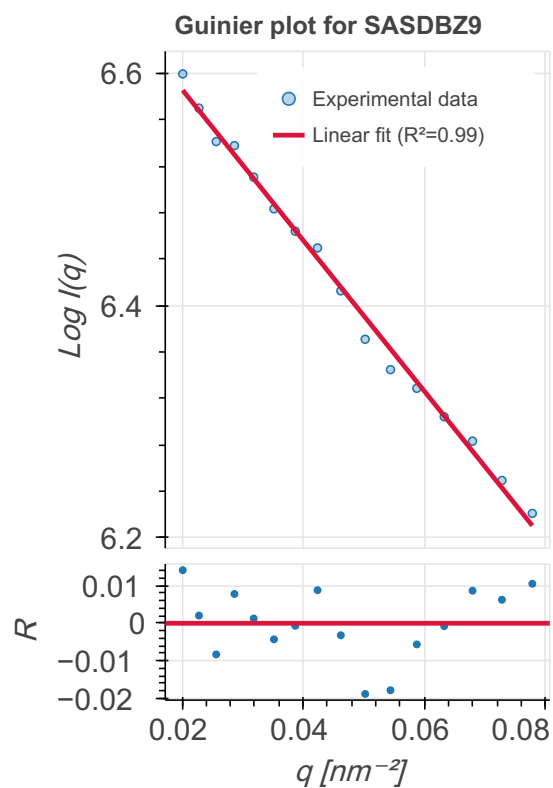
Guinier analysis ?

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
SASDBX9	2.78 nm	0.18 nm	14.7 kDa	Not available
SASDBY9	2.95 nm	0.11 nm	25.2 kDa	Not available
SASDBZ9	4.34 nm	0.17 nm	48.3 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.





2DEM class average

Validation for this section is under development.

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

Model ID	Analysed	Number of violations	Excluded Volume Satisfaction (%)
1	388521	2089	99.46

Fit of model to data used for modeling ?

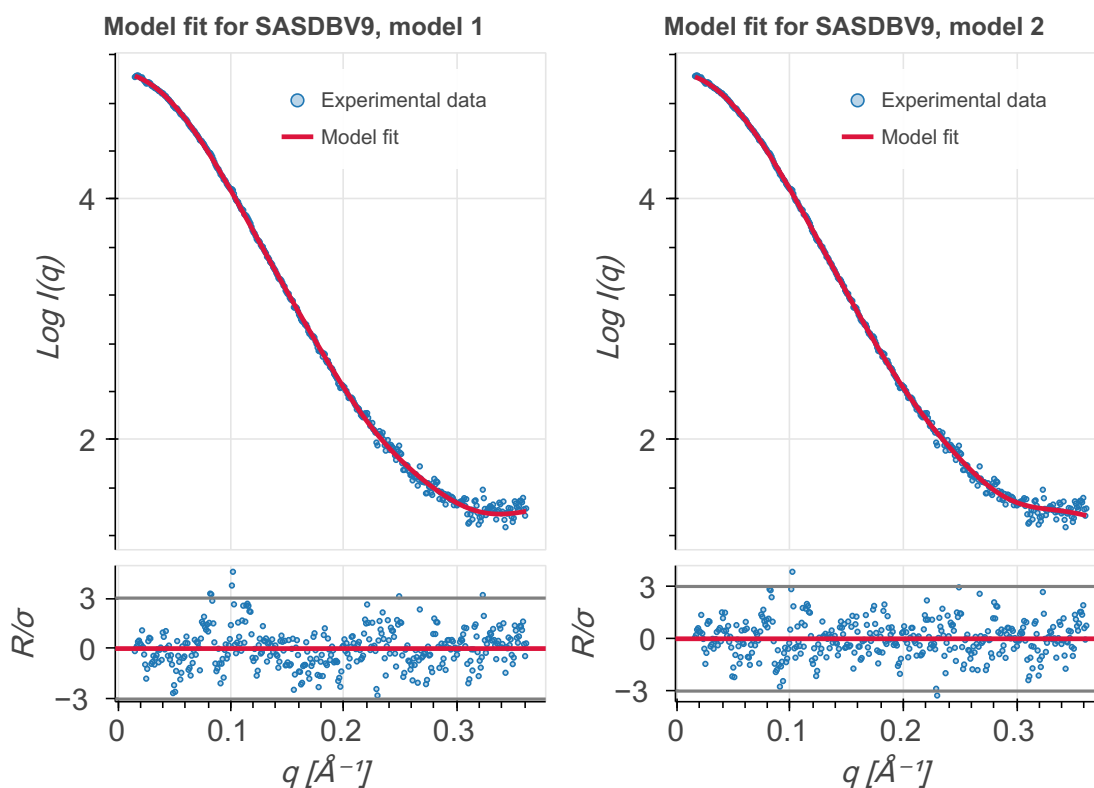
Fit of model(s) to SAS data

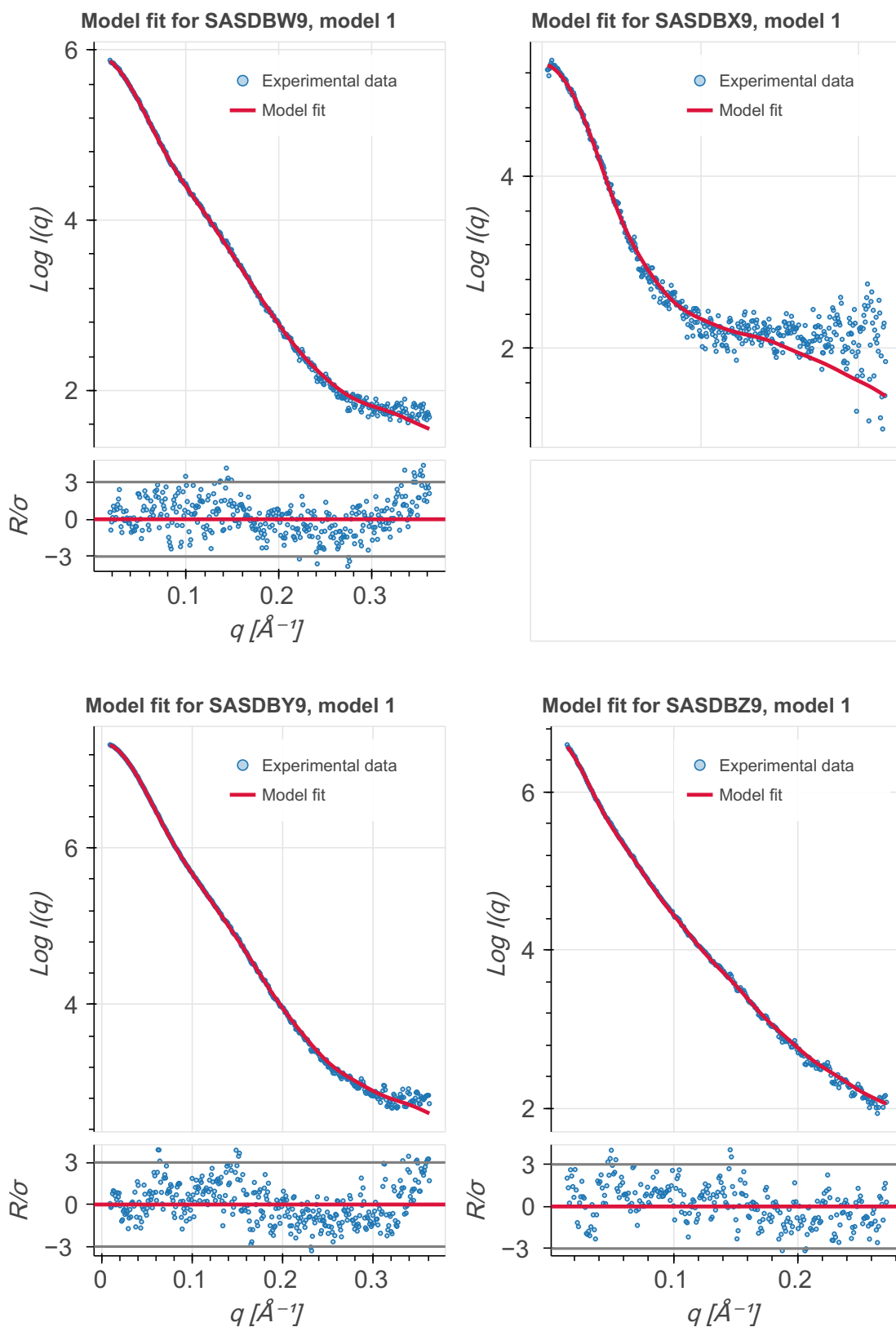
χ^2 goodness of fit and cormap analysis ?

Model and fits displayed below were obtained from SASBDB. χ^2 values are a measure of fit of the model to data. A perfect fit has a χ^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	χ^2	p-value
SASDBV9	1	1.28	0.02
SASDBV9	2	1.10	0.01
SASDBW9	1	1.97	0.00
SASDBX9	1	2.86	0.00
SASDBY9	1	2.02	0.00
SASDBZ9	1	1.94	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.





2DEM class average

Validation for this section is under development.

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 Trehwella, 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.