Integrative Structure Validation Report

October 09, 2025 - 04:37 PM PDT

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

IHMValidation Version 3.0
Python-IHM Version 2.5
ATSAS Version 3.2.1 (r14885)
EMDB validation analysis Version 0.0.1.dev127
ChimeraX Version 1.9
Chimera Version 1.19
MapQ Version 1.8.1

PDB ID	8ZZH | pdb_00008zzh
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, S.J.; Sampathkumar, P.; Dutta, K.; Cahill, S.M.; Chemmama, I.E.; Williams, R.; Bonanno, J.B.; Rice, W.J.; Stokes, D.L.; Cowburn, D.; Almo, S.C.; Sali, A.; Rout, M.P.; Fernandez-Martinez, J.
Deposited on	2018-04-04

This is a PDB-IHM Structure Validation Report.

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.

1. Overview

1.1. Summary

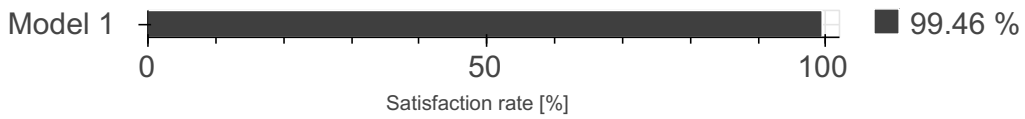
This entry consists of 1 model(s). A total of 22 dataset(s) were used to build this entry.

Name	Type	Count
2DEM class average	Experimental data	8
3DEM volume	Experimental data	2
SAS data	Experimental data	5
Experimental model	Starting model	1
Comparative model	Starting model	6

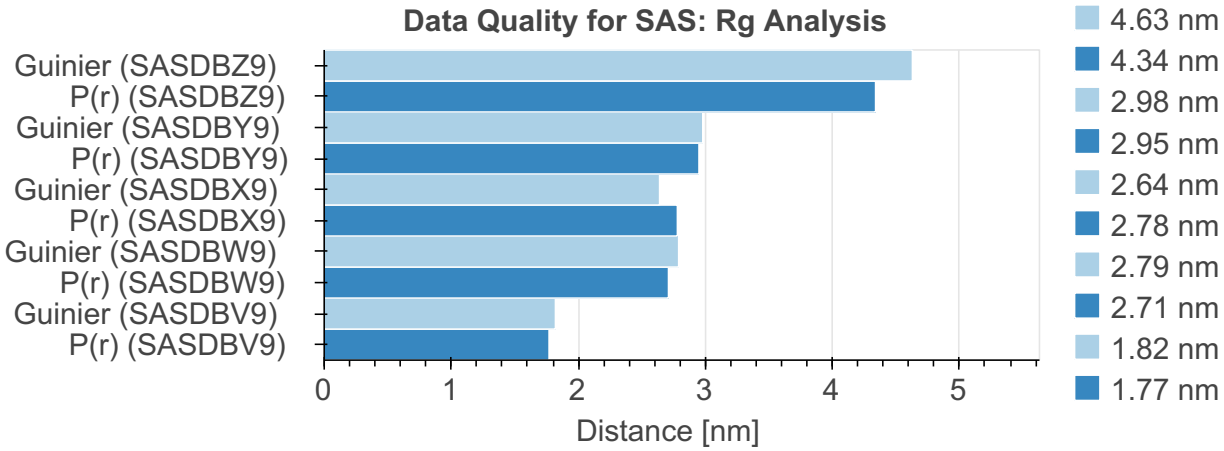
1.2. Overall quality

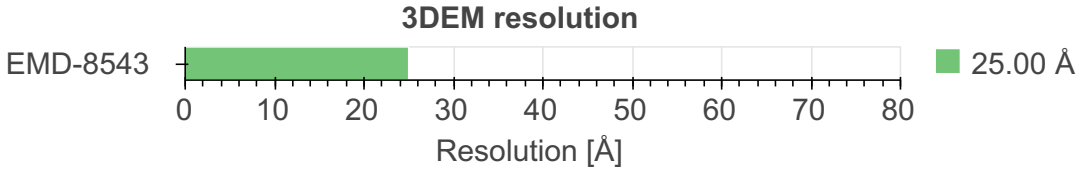
This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: Excluded Volume Analysis

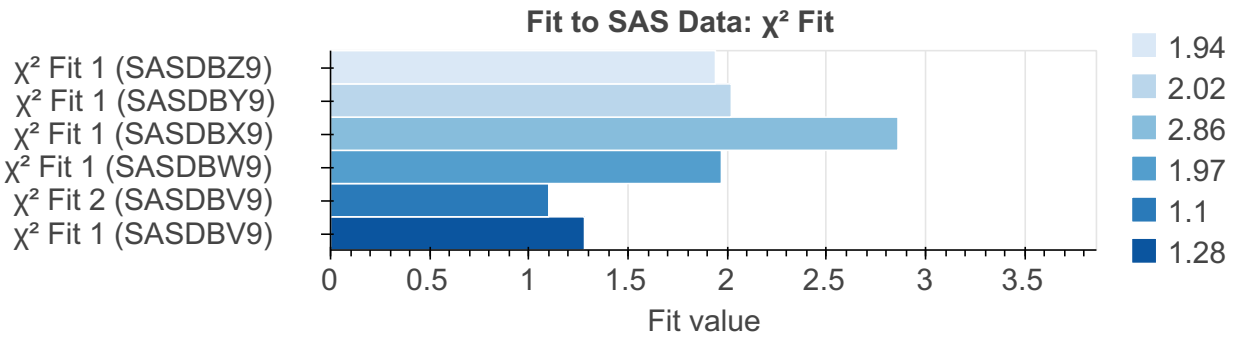


Data Quality





Fit to Data Used for Modeling



2. Model Details

2.1. Ensemble information

This entry consists of 1 distinct ensemble(s).

2.2. 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	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

2.3. 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	pdb_00005tvz
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

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

2.4. 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	Not available	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/

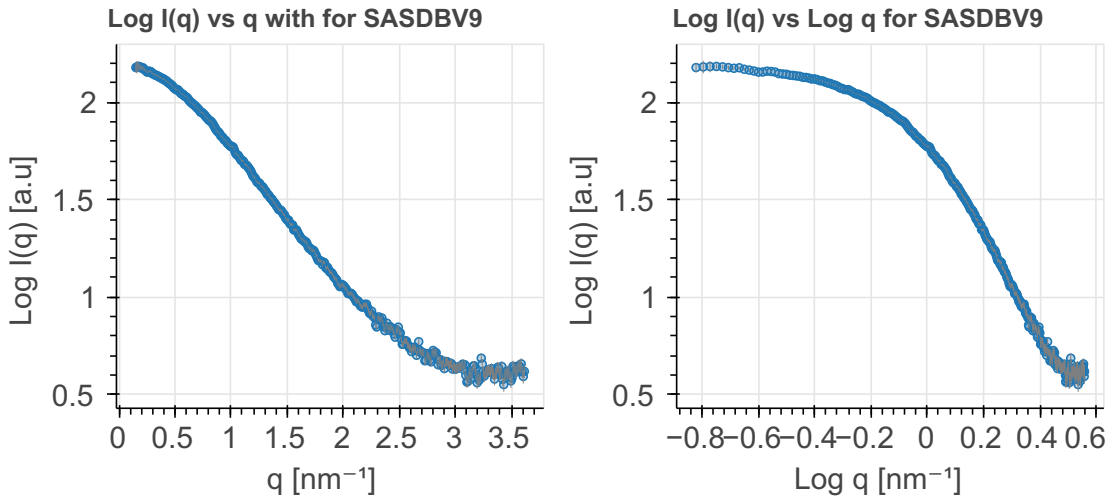
3. Data quality

3.1. SAS

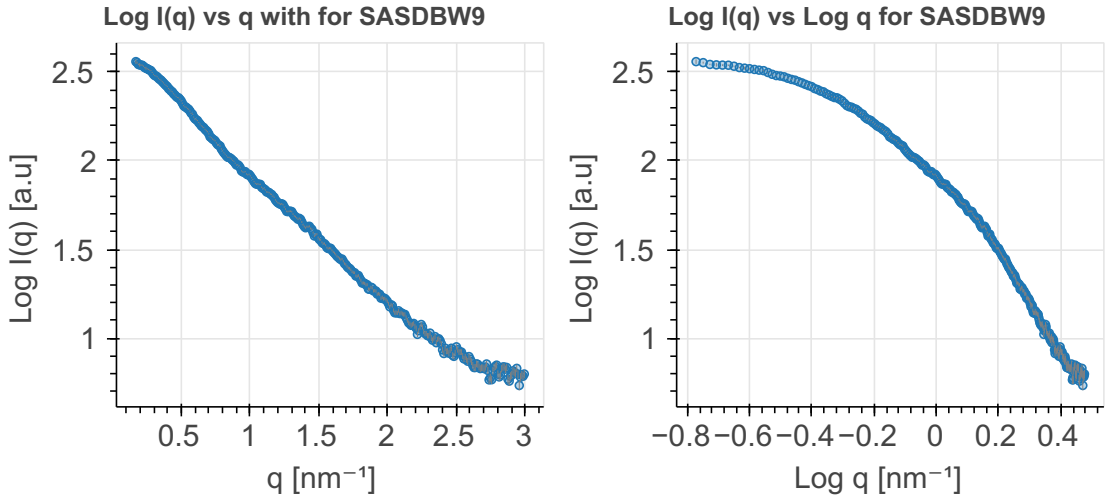
3.1.1. Scattering profile

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

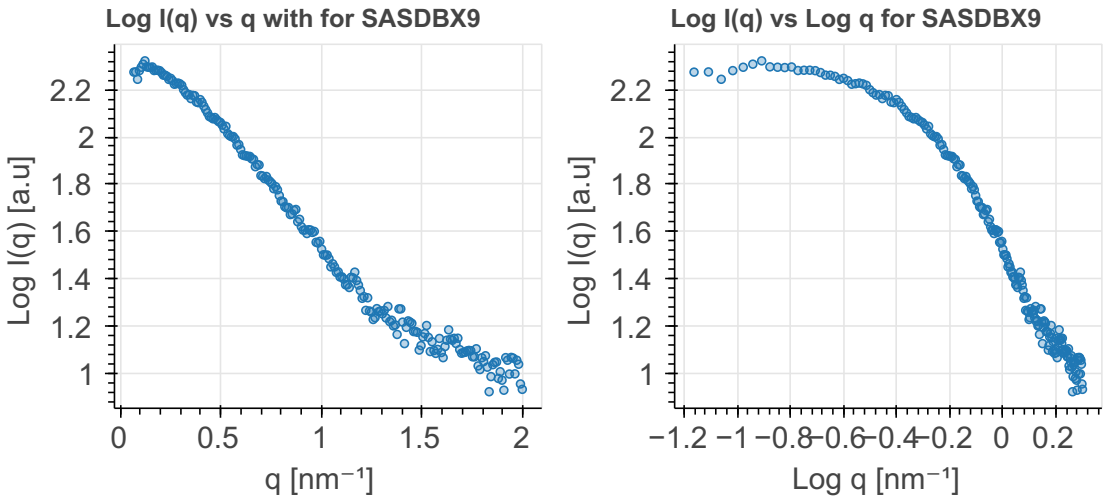
Scattering profile for SASDBV9: data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on SAS validation task force (SASvtf) recommendations. I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



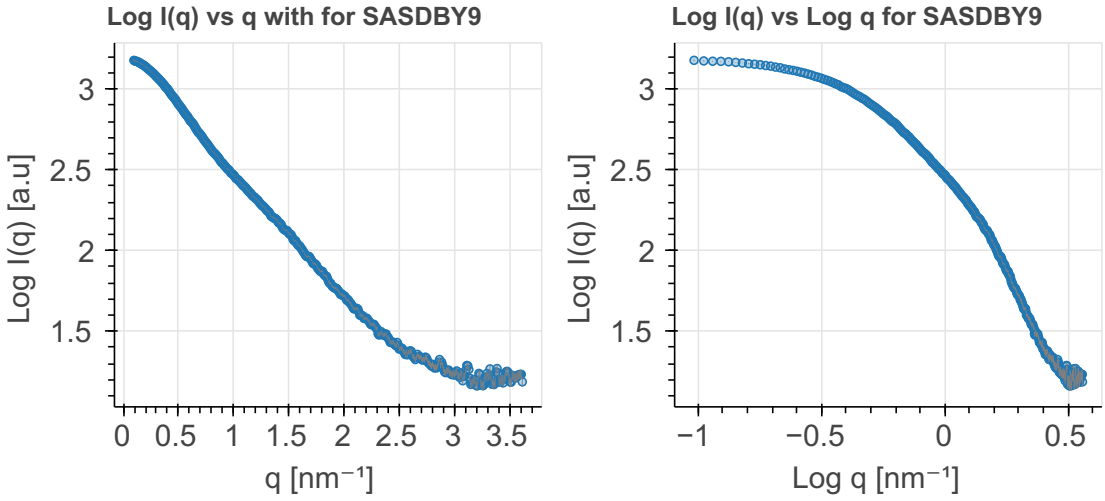
Scattering profile for SASDBW9: data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on SAS validation task force (SASvtf) recommendations. I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



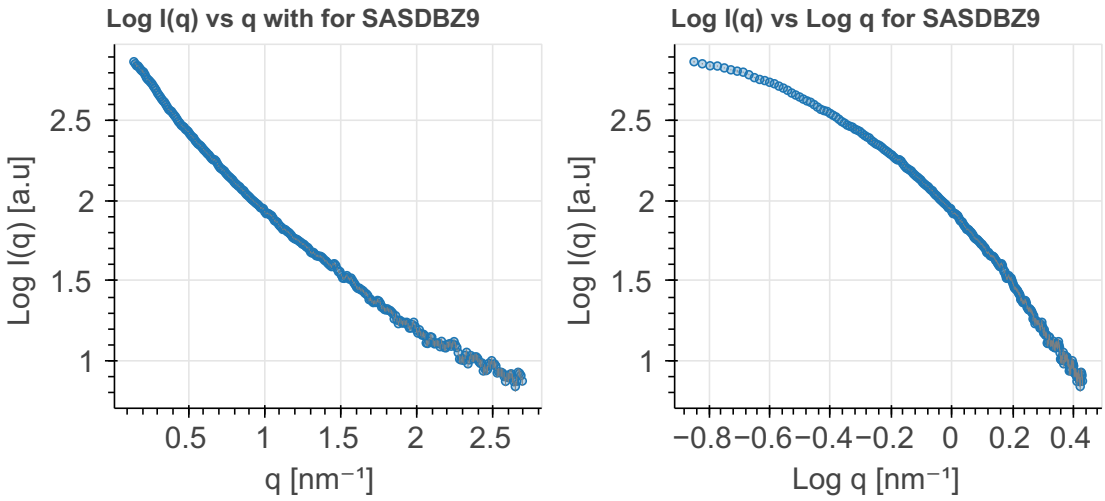
Scattering profile for SASDBX9: data from solutions of biological macromolecules are presented as both log I(q) vs q and log I(q) vs log (q) based on SAS validation task force (SASvtf) recommendations. I(q) is the intensity (in arbitrary units) and q is the modulus of the scattering vector.



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.



3.1.2. Key experimental estimates

Molecular weight (MW) estimates from experiments and analysis: Theoretical MW can be compared to SAS-derived values using the forward scatter (*I(0)*) and the known concentration and partial specific volume of the scattering particle, or as estimated from the Porod volume and partial specific volume (Trewhella et al., 2017, Trewhella et al., 2023).

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

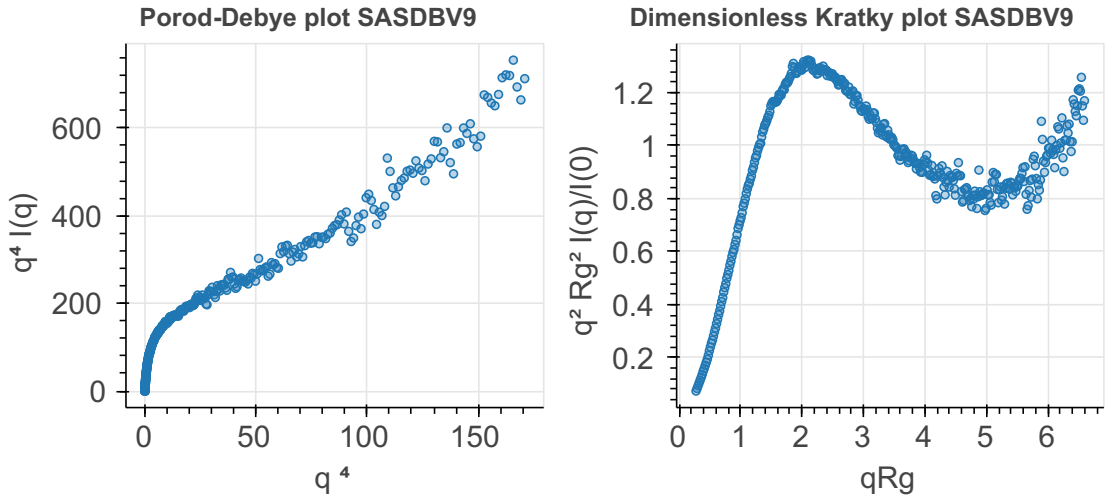
Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

SASDB ID	Estimated Volume	Porod Volume	Specific Volume	Sample Contrast	Sample Concentration
SASDBV9	Not available	17.94 nm³	Not available	Not available	Not available
SASDBW9	Not available	22.50 nm³	Not available	Not available	Not available
SASDBX9	Not available	56.68 nm³	Not available	Not available	Not available
SASDBY9	Not available	27.97 nm³	Not available	Not available	Not available
SASDBZ9	Not available	66.59 nm³	Not available	Not available	Not available

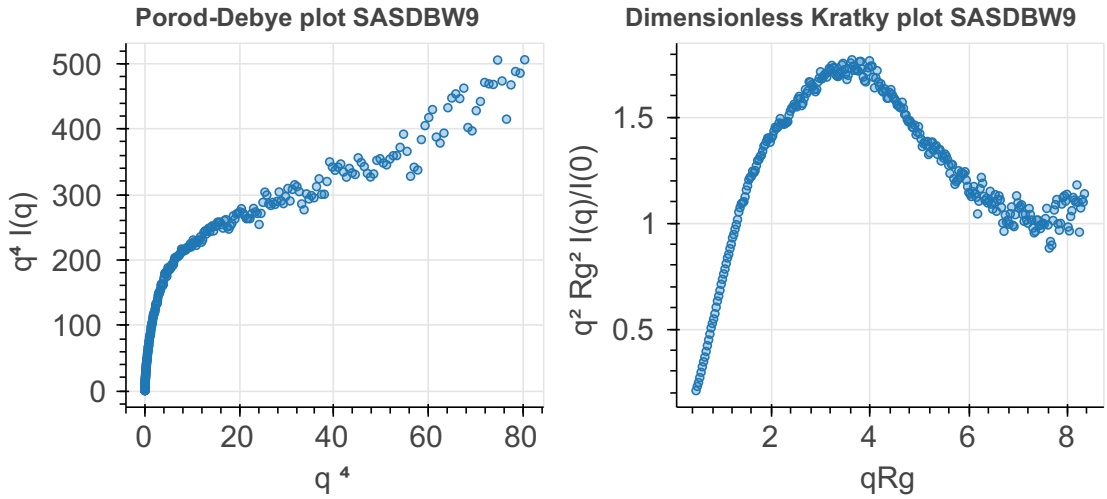
3.1.3. Flexibility analysis

In a Porod-Debye plot, a clear plateau is observed for globular (partial or fully folded) domains, whereas flexible-modular, fully unfolded domains or extended/stiff rodshaped domains lack a discernible plateau (Rambo and Tainer 2013). A bell-shaped Kratky plot ($q^2I(q)$ *vs.* q) with a well-defined maximum is observed for compact/folded structures. For partially flexible/modular or extended structures the Kratky plot can show multiple maxima and/or an increase in intensity at higher q-values depending on the degree of flexibility and extension. Fully intrinsically disordered structures yield a Kratky plot that systematically increases with increasing q values and will be near linear for highly extended molecules. The dimensionless Kratky plot ($(qR_g)^2I(q)$ *vs.* qR_g) is useful for quantifying differences in shape and foldedness among scattering objects of different sizes (Trewhella et. al., 2023).

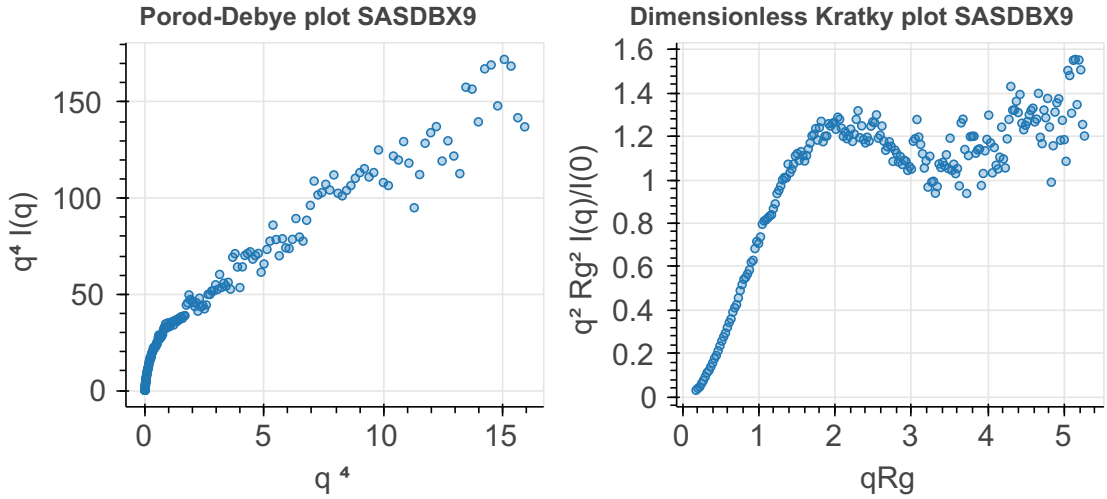
Flexibility analysis for SASDBV9



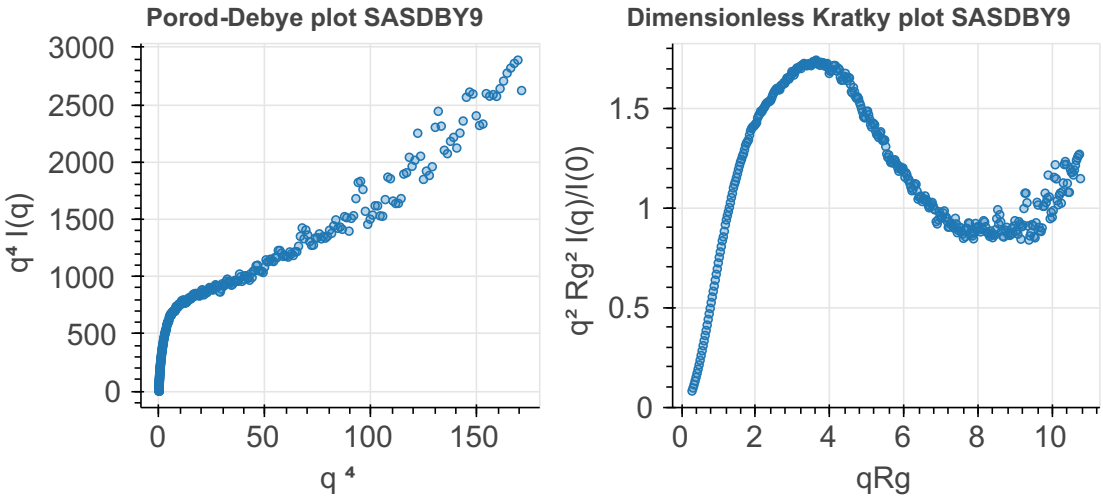
Flexibility analysis for SASDBW9



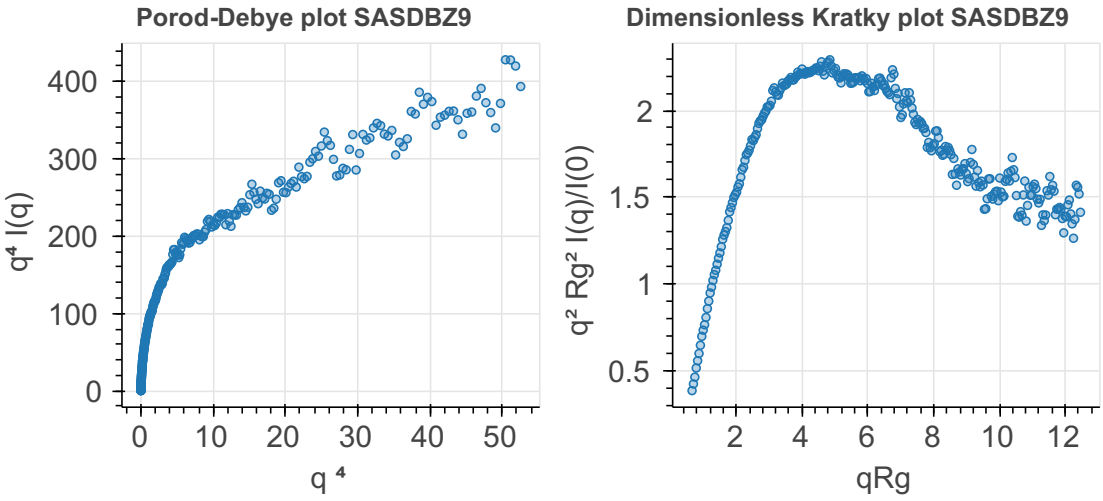
Flexibility analysis for SASDBX9



Flexibility analysis for SASDBY9



Flexibility analysis for SASDBZ9

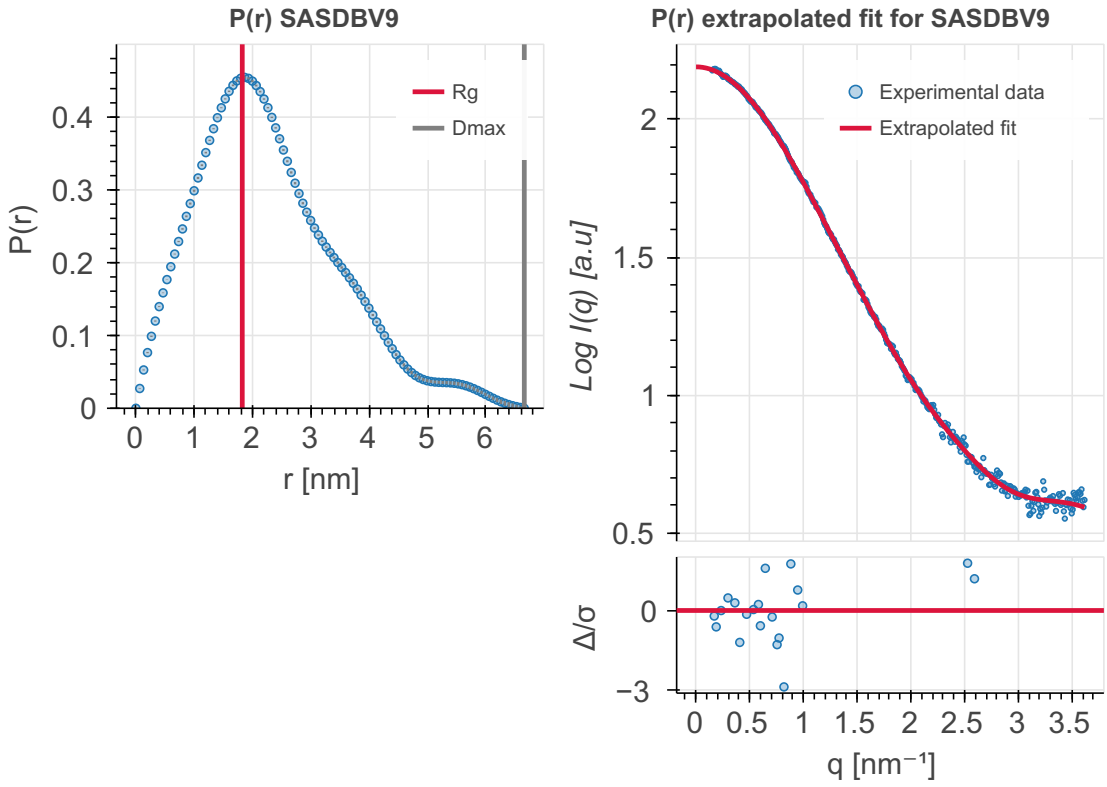


3.1.4. Pair-distance distribution analysis

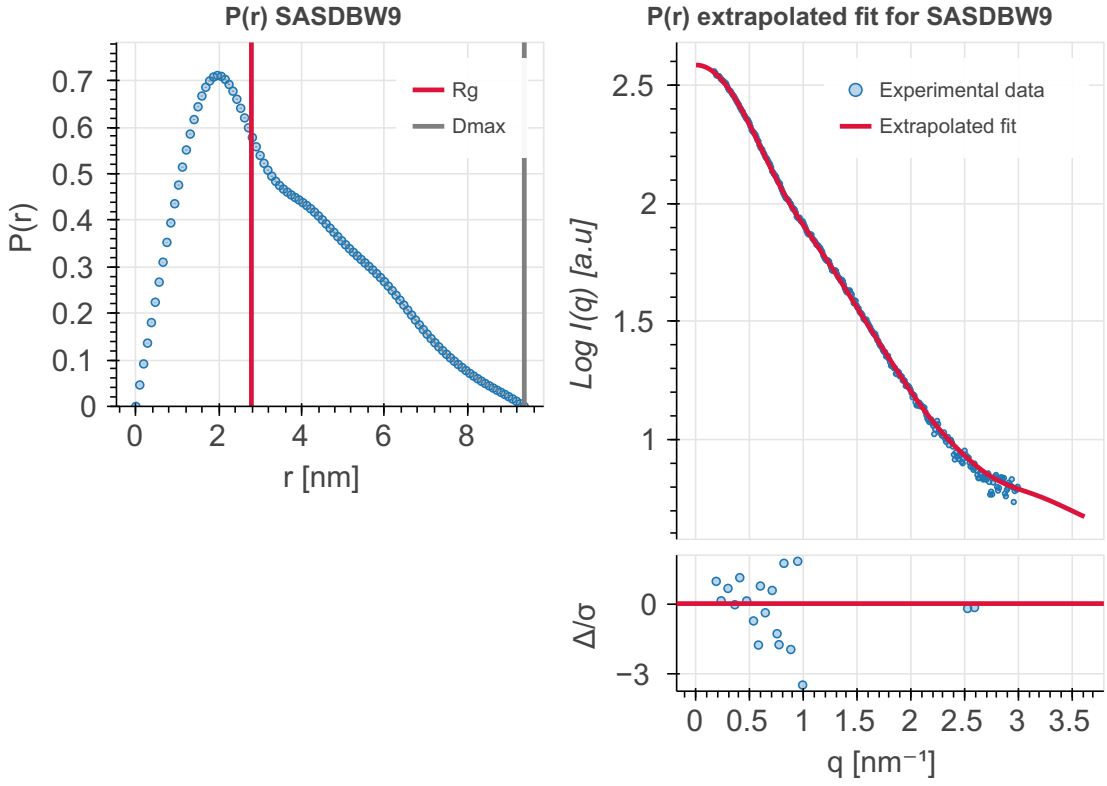
The the atom-pair distance distribution function (PDDF) or *P(r)* represents the distribution of distances between all pairs of atoms within the particle weighted by the respective scattering contrasts (Moore, 1980). The second moment of *P(r)* yields the radius of gyration (Rg), which is a measure of the overall size and shape of a macromolecule (i.e. the spatial distribution of volume elements). 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.

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

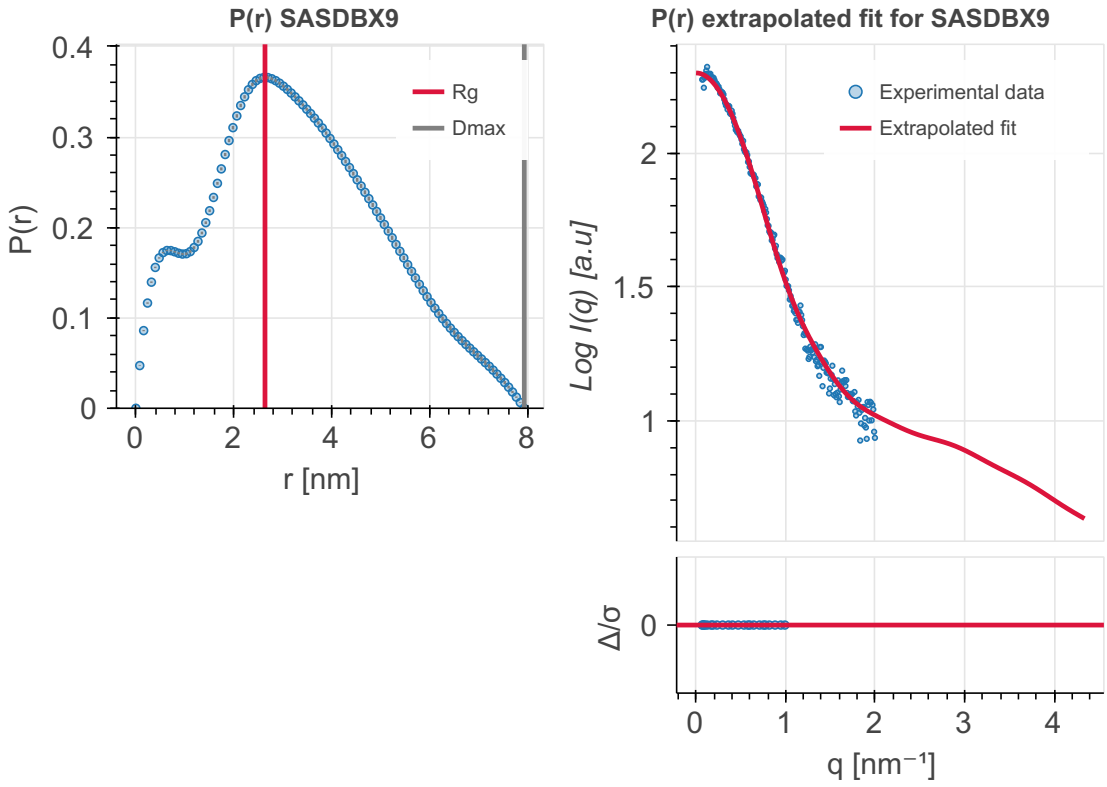
P(r) for SASDBV9: The value of P(r) should be zero beyond r=D_{max}.



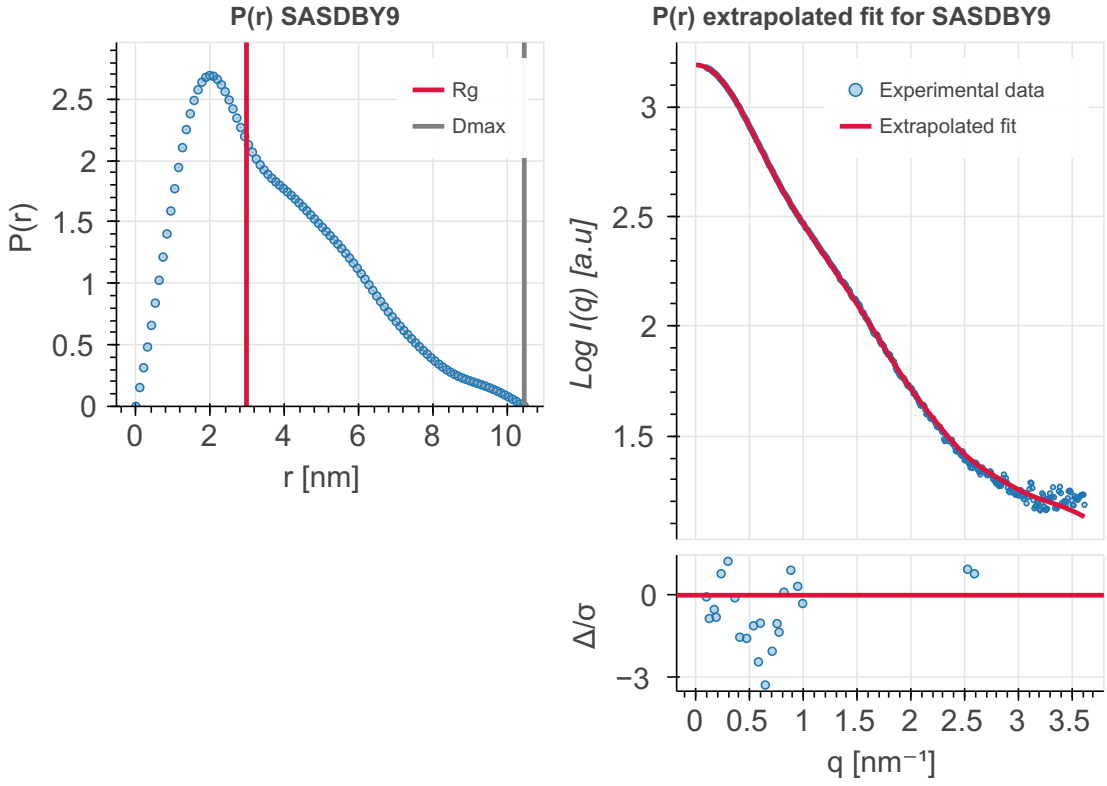
P(r) for SASDBW9: The value of P(r) should be zero beyond r=D_{max}.



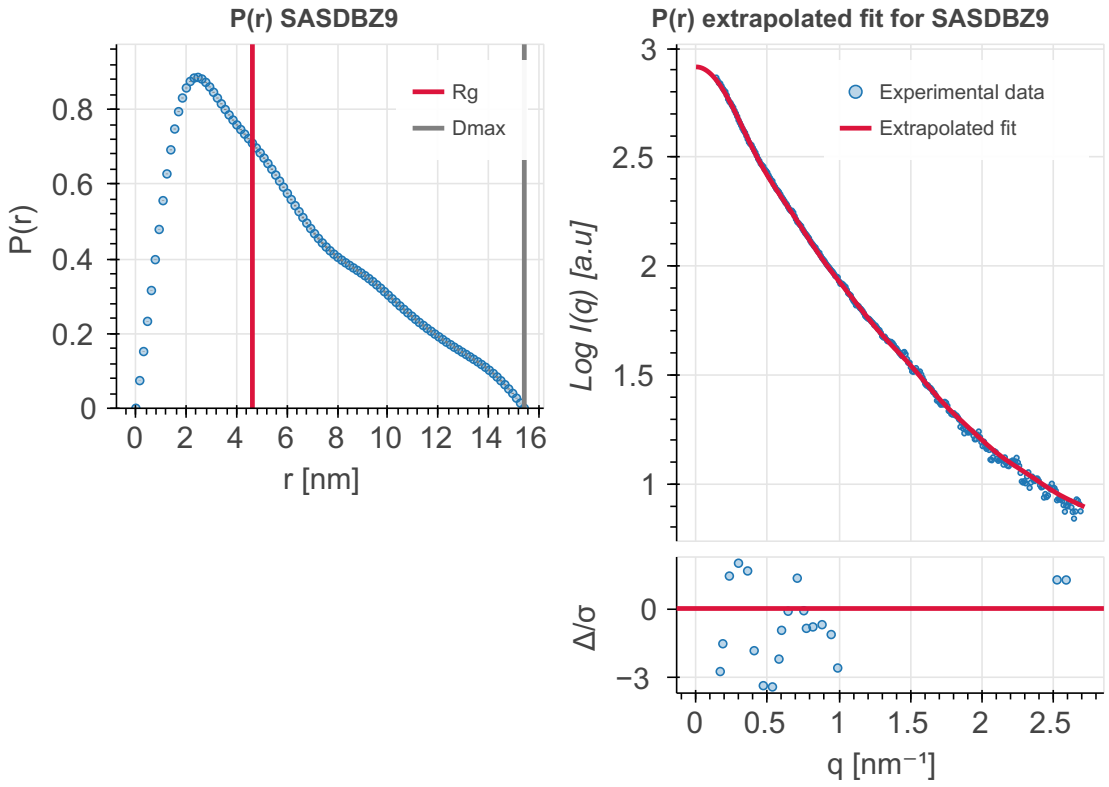
P(r) for SASDBX9: The value of P(r) should be zero beyond r=D_{max}.



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



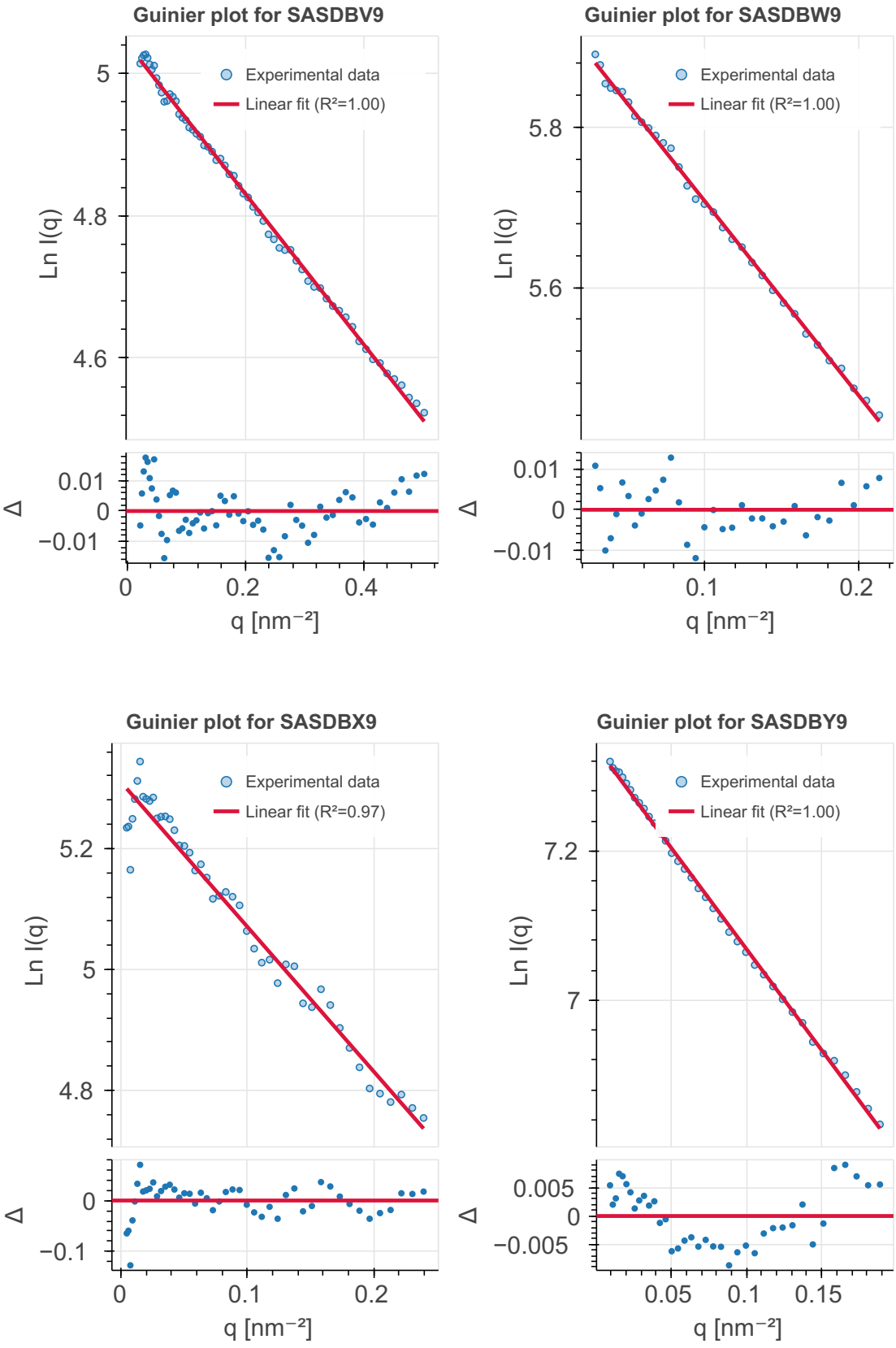
P(r) for SASDBZ9: The value of P(r) should be zero beyond r=D_{max}.

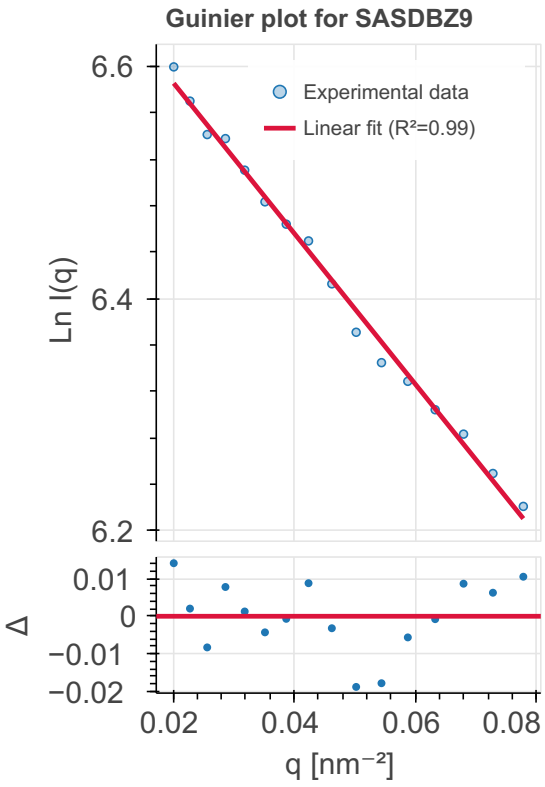


3.1.5. Guinier analysis

The linearity of the Guinier plot (ln(q) *vs.* q^2) at very-low angle (qRg < 1.3) is a sensitive indicator of the quality of the sample in relation to its homogeneity; a linear Guinier plot is a necessary but not sufficient demonstration that a solution contains monodisperse particles of the same size. Deviations from linearity can point to strong interference effects from particle attraction or repulsion, polydispersity of the samples, or improper background subtraction (Feigin et al., 2013). Residual difference plots and Pearson correlation coefficient determination (R^2) are measures to assess quality of the linear fit to the Guinier region. A perfect fit has an R^2 value of 1. Residual values should be equally and randomly spaced around the horizontal axis with no evident systematic upward or downward curvature.

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 plot for SASDBV9
Experimental data
Linear fit (R²=1.00)
Ln I(q)
Δ
q [nm⁻²]
Guinier plot for SASDBW9
Experimental data
Linear fit (R²=1.00)
Ln I(q)
Δ
q [nm⁻²]
Guinier plot for SASDBX9
Experimental data
Linear fit (R²=0.97)
Ln I(q)
Δ
q [nm⁻²]
Guinier plot for SASDBY9
Experimental data
Linear fit (R²=1.00)
Ln I(q)
Δ
q [nm⁻²]



3.3. 3DEM

This section describes quality of the 3DEM datasets

EMD-8543

3.3.1. Experimental information

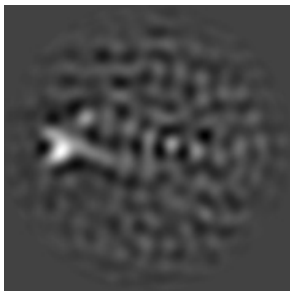
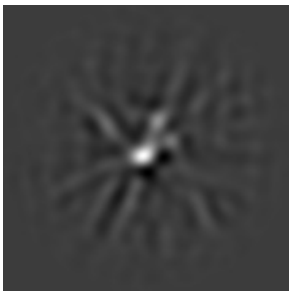
EM reconstruction method:	SINGLE PARTICLE
Resolution:	25.00 Å
Recommended level:	0.053
Estimated volume:	293.44 nm³
Specimen preparation:	Preparation ID 1 Staining: NEGATIVE; Material: Uranyl formate
Map-only validation report:	wwPDB validation report

3.3.2. Map visualisation

This section contains visualisations of the EMDB entry EMD-8543. These allow visual inspection of the internal detail of the map and identification of artifacts. Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

3.3.2.1. Orthogonal projections

Primary map

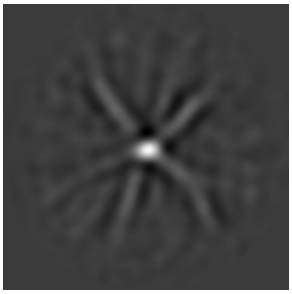




X **Y** **Z**

The images above show the map projected in three orthogonal directions.

3.3.2.2. Central slices

Primary map

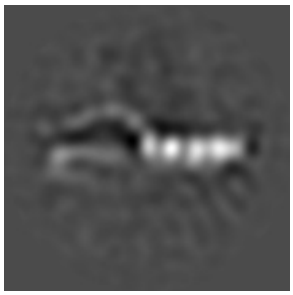




X Index: 90 **Y Index: 90** **Z Index: 90**

The images above show central slices of the map in three orthogonal directions.

3.3.2.3. Largest variance slices

Primary map

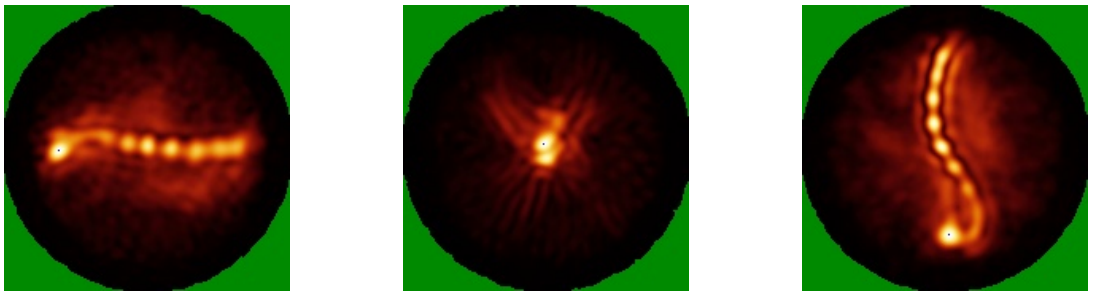


X Index: 84 **Y Index: 34** **Z Index: 89**

The images above show the largest variance slices of the map in three orthogonal directions.

3.3.2.4 Orthogonal standard-deviation projections (false-color)

Primary map

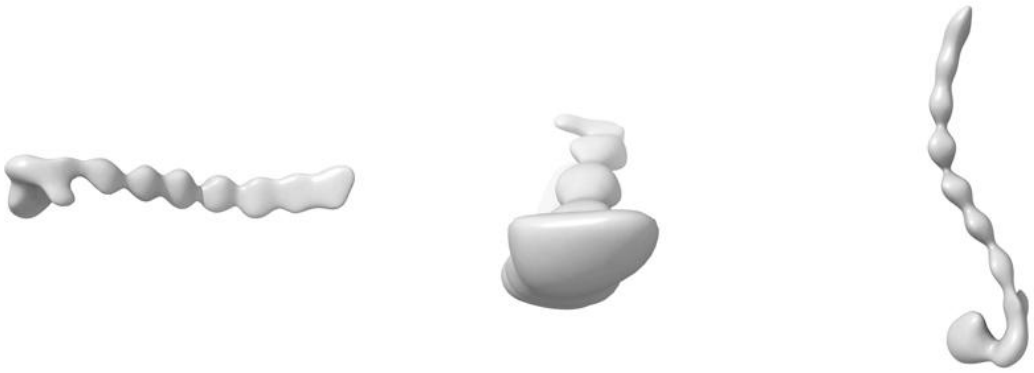


X **Y** **Z**

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

3.3.2.5. Orthogonal surface views

Primary map



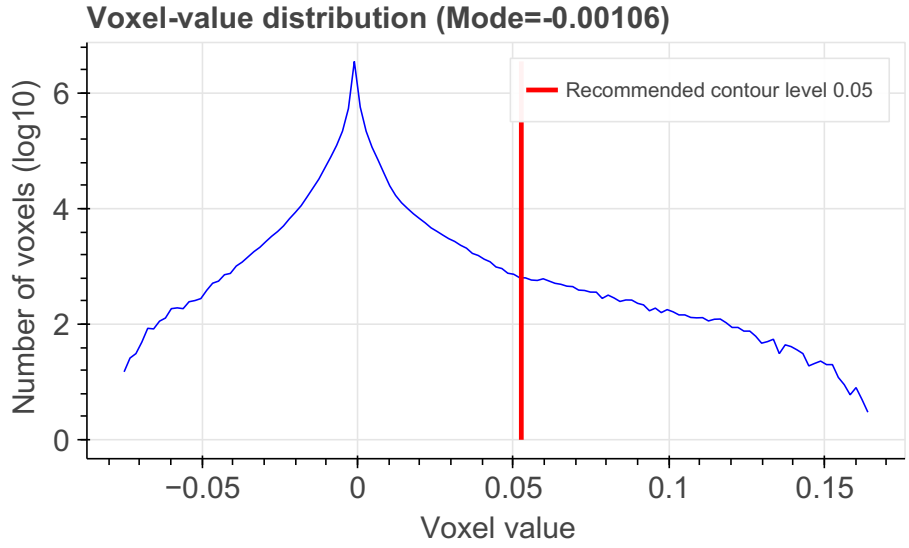
X **Y** **Z**

The images above show the 3D surface view of the map at the recommended contour level 0.053 . These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

3.3.3. Map analysis

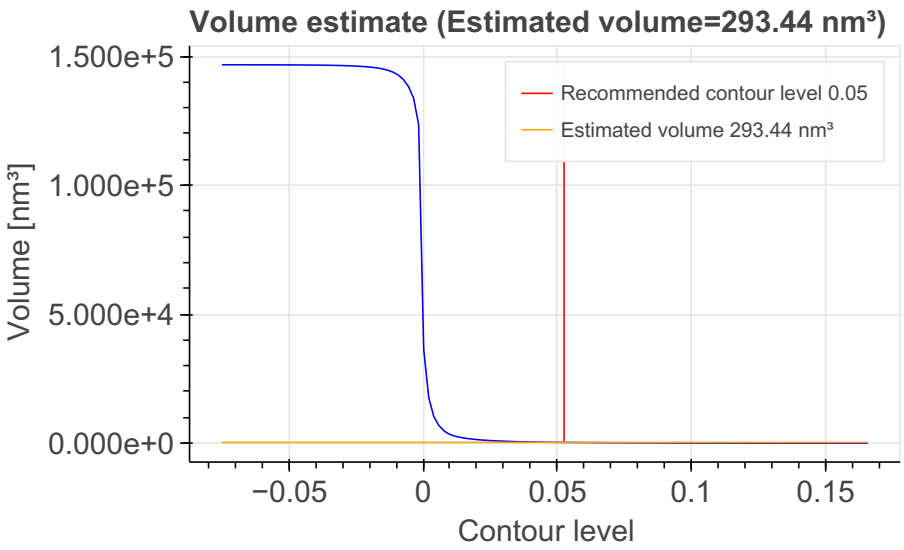
This section contains the results of statistical analysis of the map.

3.3.3.1. Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

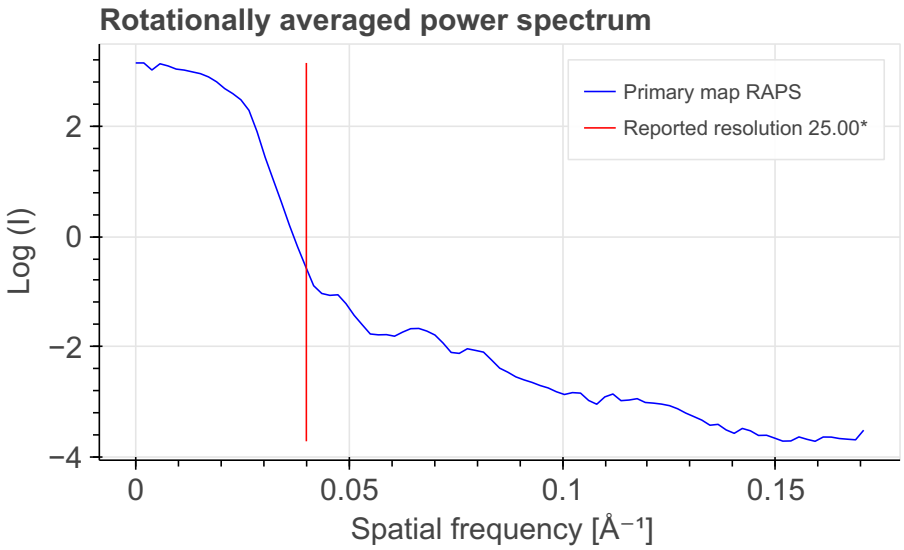
3.3.3.2. Volume estimate



The volume at the recommended contour level is 293.44 nm³.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

3.3.3.3. Rotationally averaged power spectrum



*Reported resolution corresponds to spatial frequency of 0.040 Å⁻¹

3.3.4. Fourier-Shell correlation

3.3.4.2. Resolution estimates

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	25.00	-	-
Author-provided FSC curve	25.00	33.22	25.84

3.4. 2DEM class average

Validation for this section is under development.

4. 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.

4.1a. Excluded Volume Analysis

Excluded volume satisfaction for the models in the entry are listed below. The Analysed column shows the number of particle-partice or particle-atom pairs for which excluded volume was analysed.

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

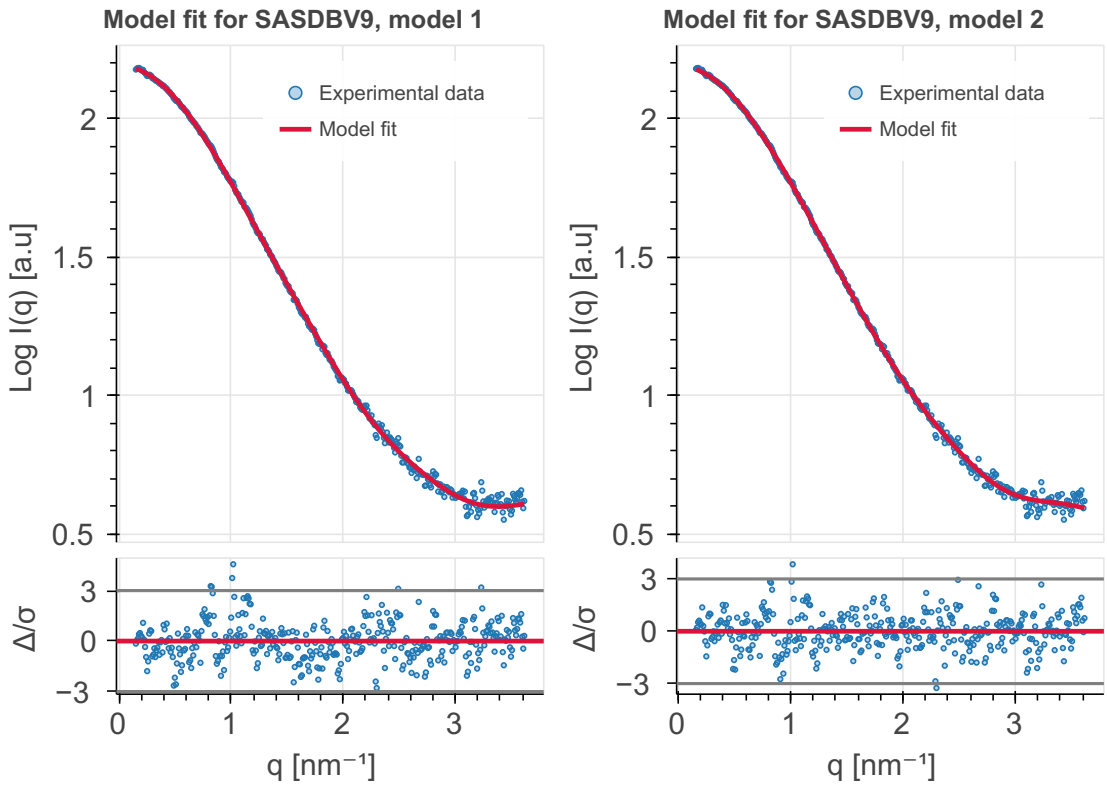
5. Fit to Data Used for Modeling Assessment

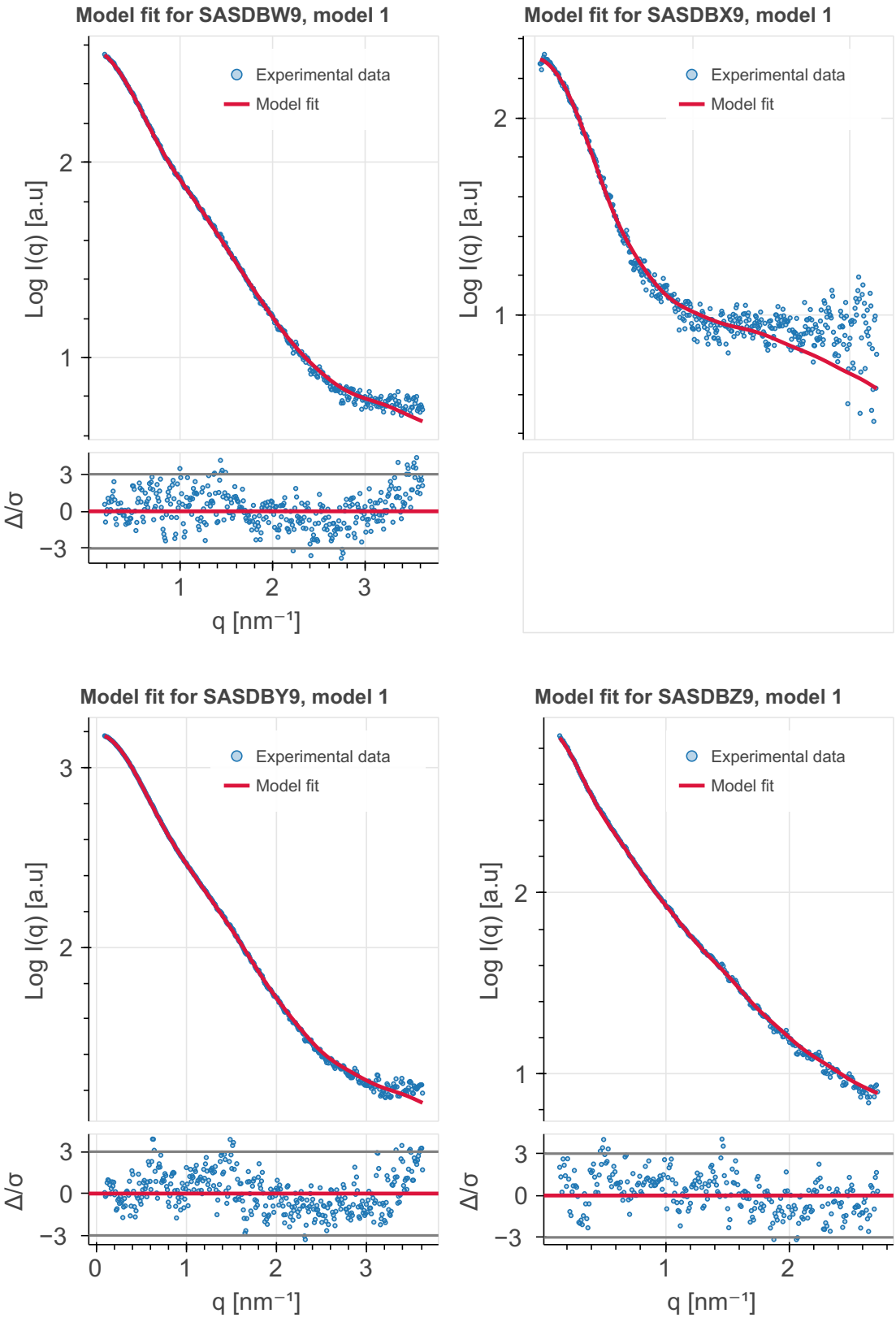
5.1. SAS

Model and fits displayed below were obtained from SASBDB.

5.1.1 Model versus Experimental Scattering Profiles

Experimental (blue) and model (red) scattering profiles are presented as log *I(q) vs. q* together with error weighted residual difference plot between the experimental and model *I(q) vs q*. The *I(q)* is the intensity (preferably in absolute units of cm-1 or arbitrary units) and *q* is the modulus of the scattering vector. For a good fit, residual values should be equally and randomly spaced around the horizontal axis largely within +/- three standard deviations.





5.1.2 χ^2 goodness of fit and 5.1.3 cormap analysis

χ^2 values are a measure of the overall fit of the model to the 1D scattering profile. A model that fits the data within its error estimates will have a χ^2 value close to one, provided that the dominant errors are the random statistical errors (i.e. no systematic errors) from the SAS measurement that are correctly propagated. Correlation Map (CorMap) test is a variance-covariance analysis on the scattering intensities comparing two (or more) scattering profiles (e.g. model versus experiment or multiple measures from the same sample). The CorMap test complements χ^2 and importantly is independent of the reported errors. The method assigns a probability (P-value based on

a 1-tailed Schilling test) for finding the longest string of experimental data points that lie systematically above (+1) or below (-1) the model profile. The P-value lies between 0 – 1 and a significance threshold is chosen below which the model fit is judged to show systematic deviation from experiment. A typical range statisticians use to indicate significant deviation is 0.01 - 0.05. As implemented in the ATSAS suite, the reported CorMap P-value is green (model fit is good) for $P > 0.05$, yellow for $0.01 < P < 0.05$, and red (model deviates significantly) for $P < 0.01$.

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

5.3. 3DEM

This section describes fit of models to the 3DEM data. Only results for the representative model, selected as a first model with the largest number of asymmetric units.

3DEM validation for coarse-grained structures is under development.

5.4. 2DEM class average

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment

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.