



Full wwPDB Integrative Structure Validation Report

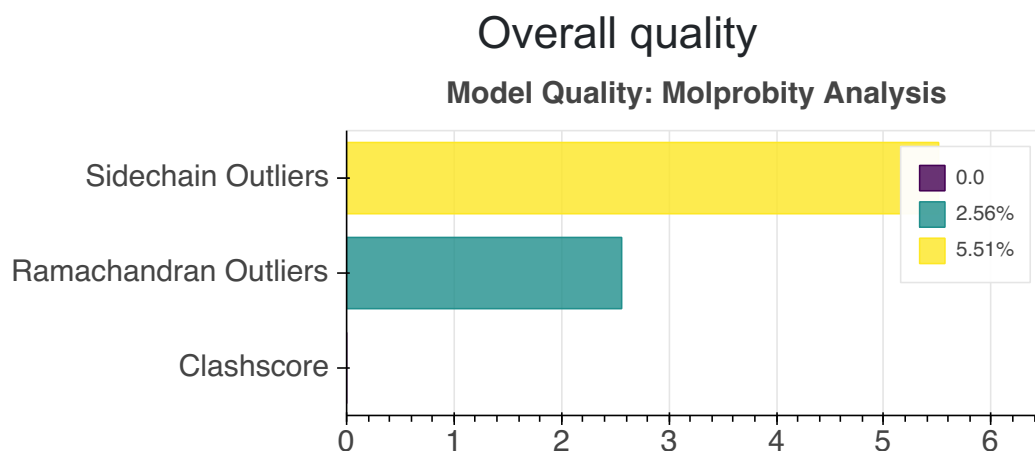
June 15, 2020 -- 09:11 PM

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

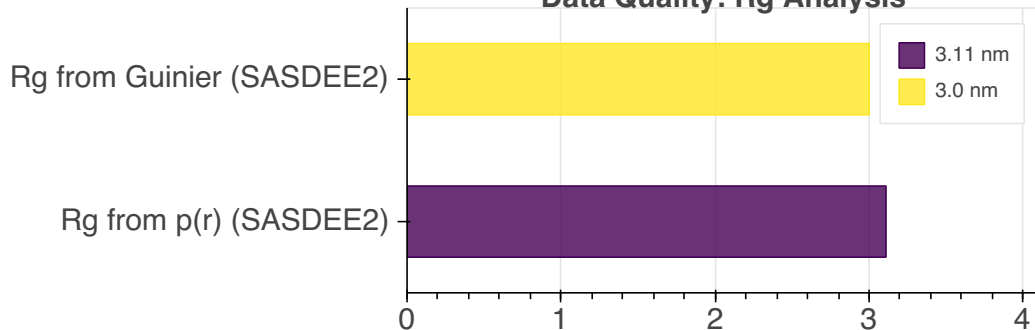
Molprobity : Version 4.4

Integrative Modeling Validation Package : Version 1.0

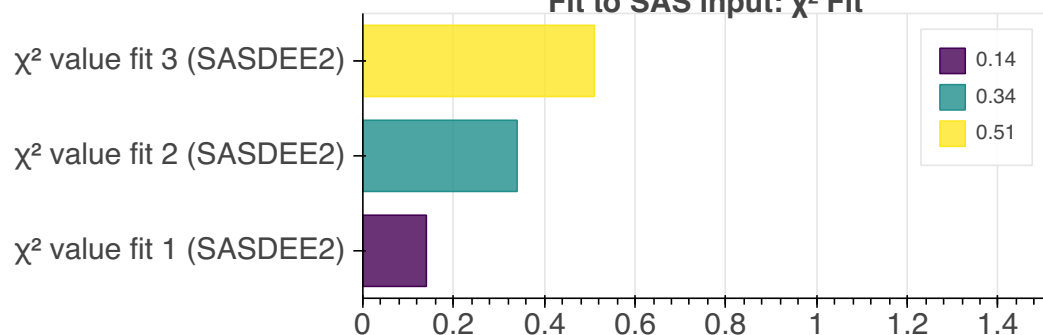
PDB ID	PDBDEV_00000027
Molecule Name	A metastable contact and structural disorder in the estrogen receptor transactivation domain
Title	A metastable contact and structural disorder in the estrogen receptor transactivation domain
Authors	Peng Y;Cao S;Kiselar J;Xiao X;Du Z;Hsien A;Ko S;Chen Y;Agrawal P;Zheng W;Shi W;Jiang W;Yang L;Chance MR;Surewicz WK;Buck M;Yang S



Data Quality: Rg Analysis



Fit to SAS input: χ^2 Fit



Ensemble information

This entry consists of 0 distinct ensemble.

Summary

This entry consists of 10 unique models, with 1 subunits in each model. A total of 2 datasets or restraints was used to build this entry. Each model is represented by 1 rigid bodies and 0 flexible or non-rigid units.

Entry composition

There are 10 unique types of models in this entry. These models are titled None/?, None/?, None/?, None/?, None/?, None/?, None/?, None/?, None/?, None/? respectively.

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Total residues
----------	----------------	------------	--------------	----------	----------------

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Total residues
1	1	1	Estrogen receptor	A	184
2	1	1	Estrogen receptor	A	184
3	1	1	Estrogen receptor	A	184
4	1	1	Estrogen receptor	A	184
5	1	1	Estrogen receptor	A	184
6	1	1	Estrogen receptor	A	184
7	1	1	Estrogen receptor	A	184
8	1	1	Estrogen receptor	A	184
9	1	1	Estrogen receptor	A	184
10	1	1	Estrogen receptor	A	184

Methodology and software

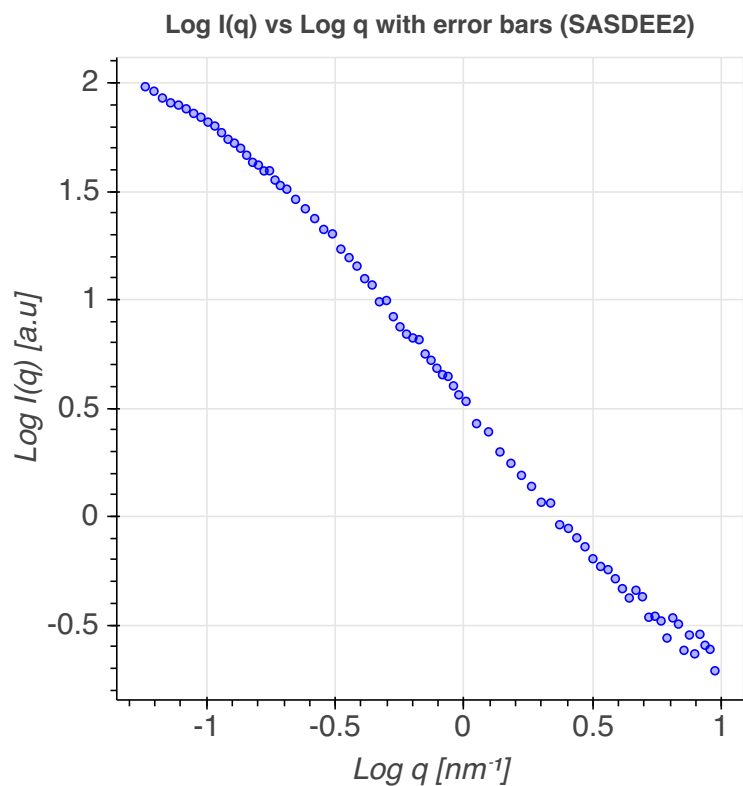
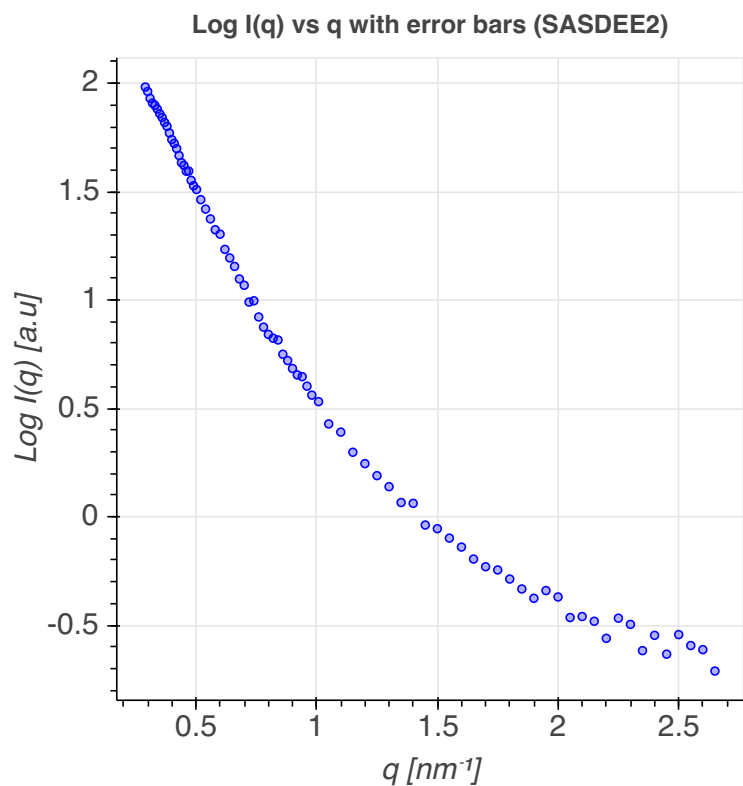
Step number	Protocol ID	Method name	Method type	Number of computed models	Multi state modeling	Multi scale modeling
1	1	?	Modeling estrogen receptor N-terminal domain	None	False	False

Data quality

Scattering profile

SAS data used in this integrative model was obtained from 1 deposited SASBDB entry (entries).

Scattering profile for SASDEE2: 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 MW
SASDEE2	20.18 kDa	N/A	N/A

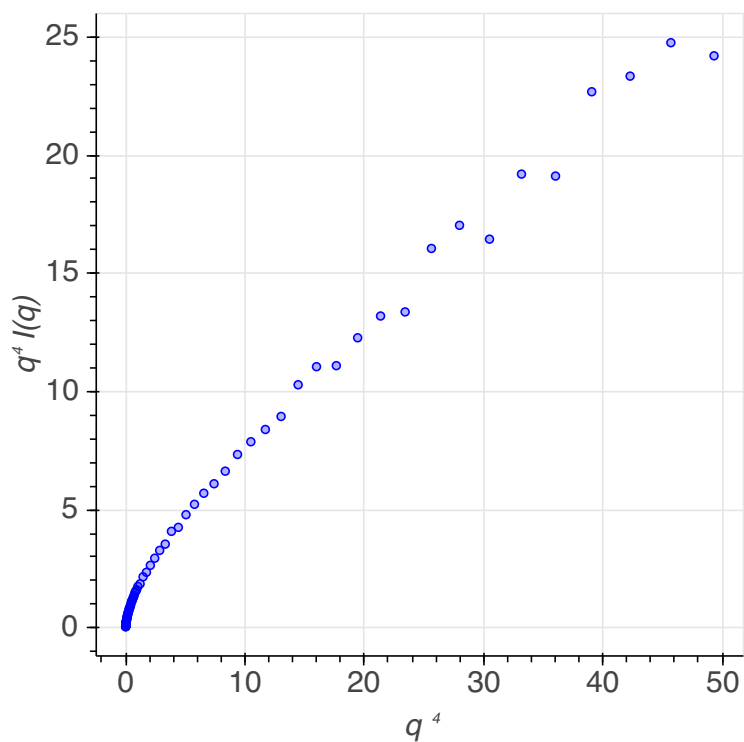
Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

SASDB ID	Estimated volume	Estimated volume method	Porod volume
SASDEE2	None	None	N/A

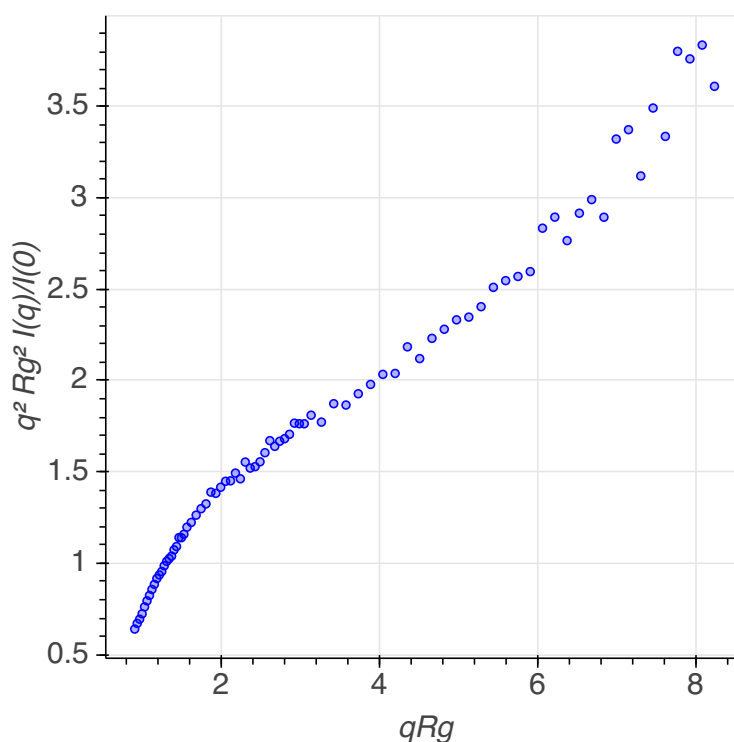
Flexibility analysis

Flexibility analysis for SASDEE2: 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.

Porod-Debye plot (SASDEE2)



Dimensionless Kratky plot (SASDEE2)

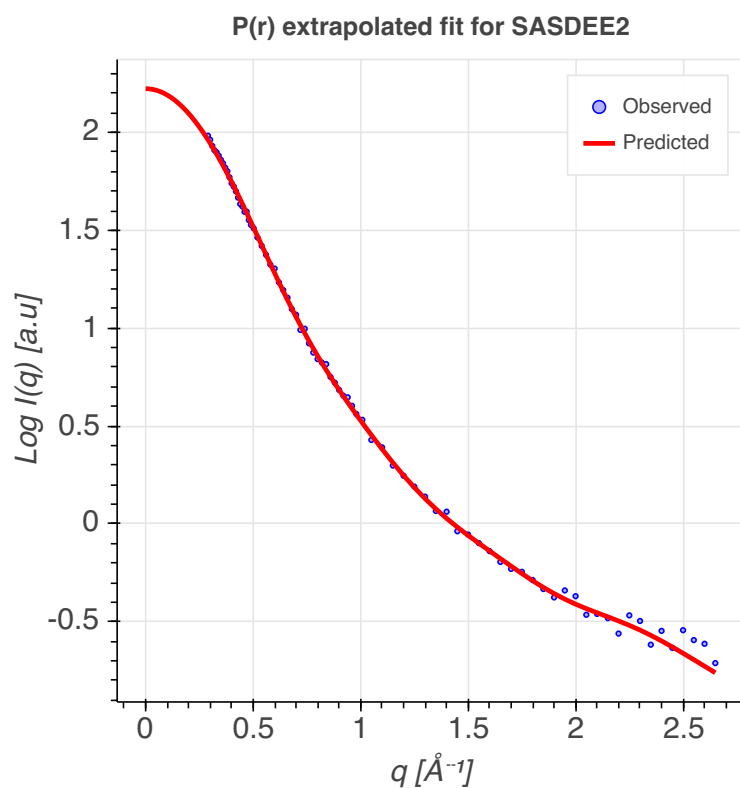
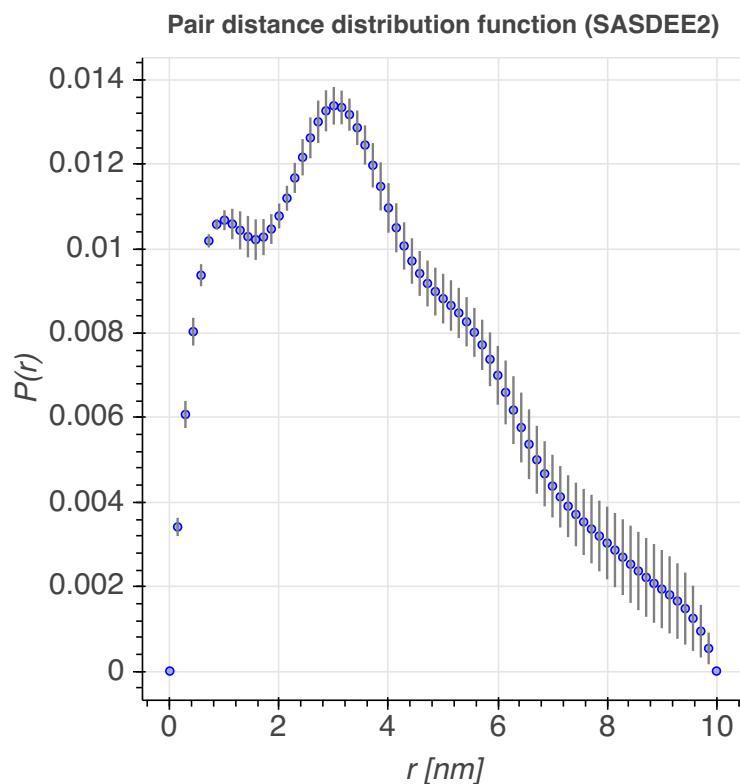


Pair-distance distribution analysis

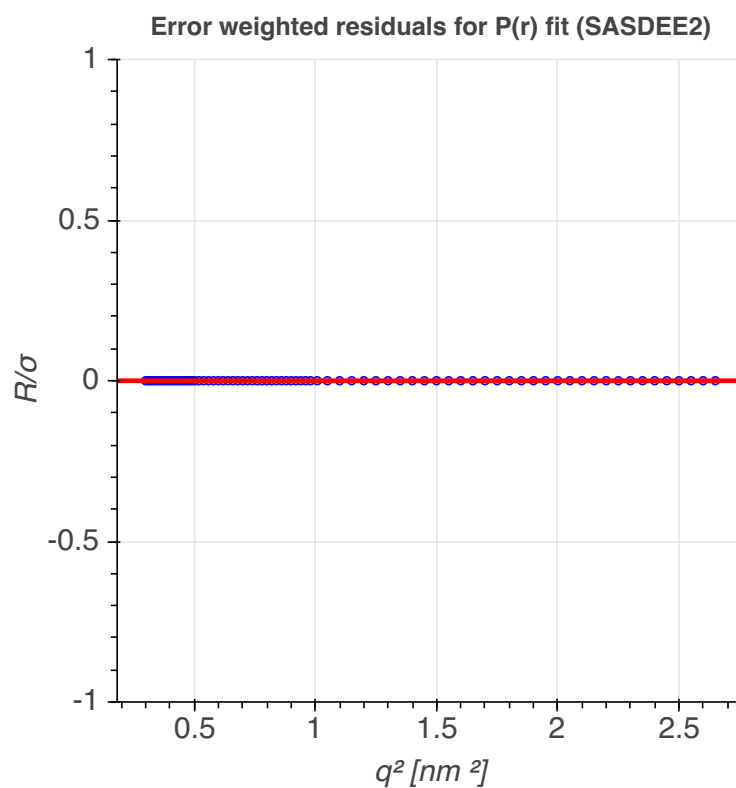
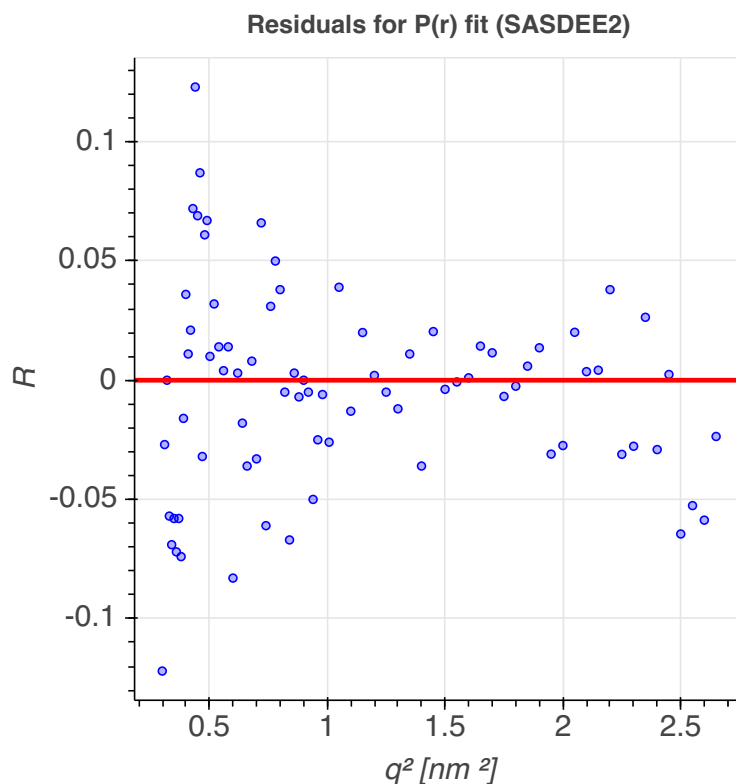
P(r) analysis: $p(r)$ represents the distribution of distances between all pairs of atoms within the particle weighted by the respective electron densities. $p(r)$ is the Fourier transform of $I(s)$ (and vice versa). R_g can be estimated from integrating the $p(r)$ function. Agreement between the $p(r)$ 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
SASDEE2	ATSAS GNOM	10.00 nm	N/A	3.110 nm	None nm

P(r) for SASDEE2: The value of $P(r)$ should be zero beyond $r=D_{max}$.



Residuals and error weighted residuals for P(r) analysis for SASDEE2: Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.

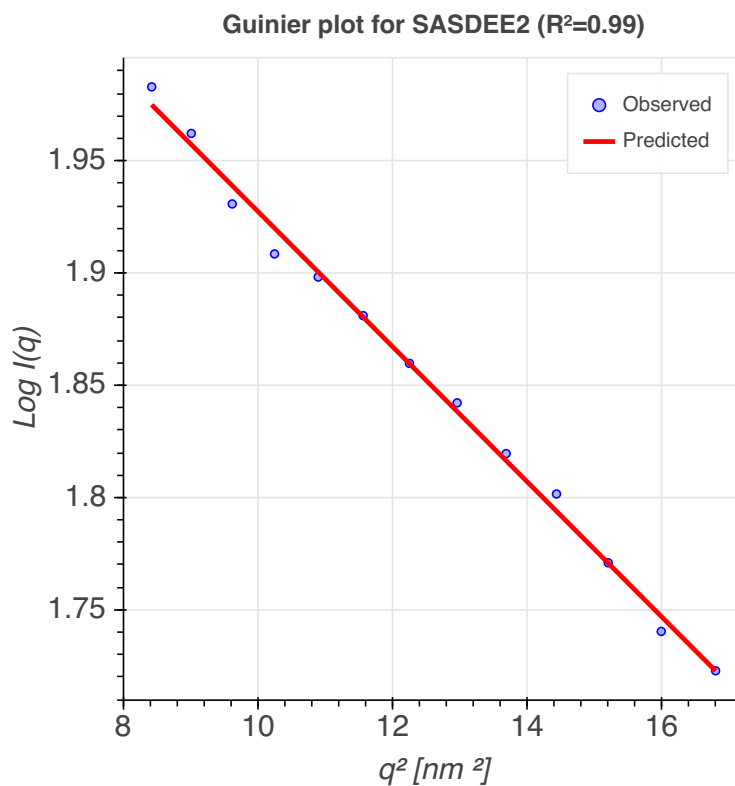


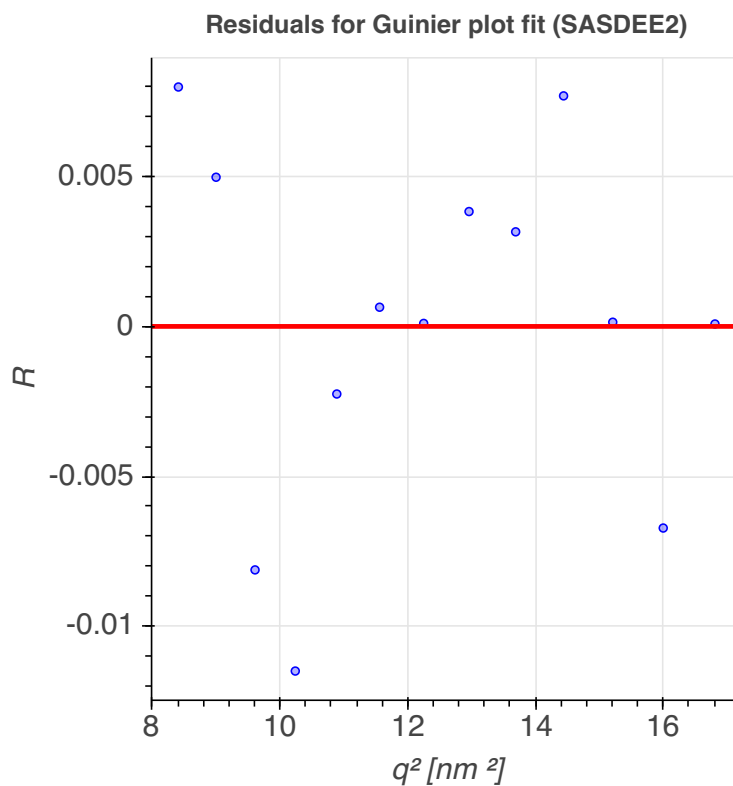
Guinier analysis

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	Rg	Rg error	MW	MW error
SASDEE2	3.0 nm	0.160 nm	N/A	N/A

Guinier analysis for SASDEE2: 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.





Model quality

Standard geometry

There are 13530 bond outliers in this entry.

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
ND2--HD22	1.008	0.86	109
N--H	1.008	0.86	1609
NE2--HE2	1.008	0.86	39
NH1--HH12	1.008	0.86	89
CD2--HD2	1.078	0.93	199
CE1--HE1	1.078	0.93	199
CE2--HE2	1.078	0.93	159
NH1--HH11	1.009	0.86	89
ND2--HD21	1.009	0.86	109

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NE2--HE22	1.009	0.86	99
CZ--HZ	1.079	0.93	49
NH2--HH22	1.009	0.86	89
NE2--HE21	1.008	0.86	99
NE--HE	1.009	0.86	89
CD1--HD1	1.078	0.93	159
NH2--HH21	1.009	0.86	89
CG--HG3	1.089	0.97	649
CA--HA	1.088	0.97	1659
CB--HB	1.088	0.97	149
CD--HD2	1.089	0.97	359
CB--HB2	1.088	0.97	1509
CD2--HD22	1.088	0.97	189
CE--HE1	1.088	0.97	59
CB--HB3	1.088	0.97	1509
CB--HB1	1.088	0.97	189
CG2--HG21	1.089	0.97	149
CA--HA2	1.089	0.97	179
CD2--HD21	1.088	0.97	189
CG--HG2	1.088	0.97	649
CA--HA3	1.089	0.97	179
CD--HD3	1.089	0.97	359
CD1--HD11	1.089	0.97	209

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CD1--HD13	1.088	0.97	209
NZ--HZ3	1.009	0.89	49
CG1--HG12	1.089	0.97	79
CG2--HG22	1.088	0.97	149
N--H3	1.009	0.89	9
CG2--HG23	1.089	0.97	149
OG--HG	0.959	0.84	139
CE--HE2	1.089	0.97	109
CD1--HD12	1.089	0.97	209
CG1--HG11	1.089	0.97	59
NZ--HZ2	1.009	0.89	49
CG1--HG13	1.088	0.97	79
CD2--HD23	1.089	0.97	189
OG1--HG1	0.959	0.84	69
CE--HE3	1.089	0.97	109
NZ--HZ1	1.009	0.89	49
OH--HH	0.959	0.84	109
CG--HG	1.089	0.97	189
N--H2	1.009	0.89	9
N--H1	1.009	0.89	9

There are 316 angle outliers in this entry.

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
C-N-H	112.269	124.3	122

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
HG2-CG-HG3	97.708	110.0	8
CG-CB-HB2	120.202	108.0	7
HD22-CD2-HD23	97.918	110.0	2
HG12-CG1-HG13	92.282	110.0	0
HB2-CB-HB3	97.993	110.0	31
HZ1-NZ-HZ3	122.371	109.0	2
CA-N-H	126.273	114.0	7
CB-OG1-HG1	93.037	110.0	0
CB-CA-HA	96.938	109.0	3
N-CA-HA	97.858	110.0	6
HD21-CD2-HD22	97.003	110.0	1
HB1-CB-HB2	96.99	110.0	5
CG-CB-HB3	95.827	108.0	7
HD2-CD-HD3	122.233	110.0	7
C-CA-HA	96.951	109.0	10
CE-NZ-HZ2	94.864	110.0	0
CA-CB-HB3	96.927	109.0	9
CA-CB-HB2	121.002	109.0	4
HD12-CD1-HD13	97.874	110.0	1
HD11-CD1-HD12	95.703	110.0	0
HG22-CG2-HG23	96.722	109.0	3
SD-CE-HE1	95.744	109.0	1
HG21-CG2-HG23	95.829	110.0	0

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CB-CG-HG3	121.179	109.0	2
CB-CG1-HG11	122.984	109.0	0
CG-CD2-HD23	121.251	109.0	1
NZ-CE-HE2	121.714	108.0	0
HD21-CD2-HD23	96.708	110.0	1
N-CA-HA2	96.384	110.0	0
HE2-CE-HE3	96.468	110.0	0
CB-CG-HG2	121.139	109.0	2
CD-CG-HG2	121.262	108.0	0
CB-CG2-HG22	97.308	110.0	1
HG11-CG1-HG13	97.438	110.0	3
SD-CE-HE2	121.941	109.0	0
HE1-CE-HE3	97.073	110.0	0
CZ-CE2-HE2	107.364	120.2	0
HG11-CG1-HG12	97.195	110.0	0
CZ-OH-HH	122.334	110.0	1
HH21-NH2-HH22	107.333	120.0	0
HB1-CB-HB3	97.473	110.0	1
OG-CB-HB2	121.604	109.0	0
CD-CG-HG3	97.925	110.0	2
N-CD-HD3	121.548	109.0	0
CG-CD1-HD12	121.475	109.0	0
N-CD-HD2	121.456	109.0	0

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CG-CD2-HD22	121.003	109.0	1
CD2-CG-HG	95.7	108.0	0
CZ-NH2-HH22	107.711	120.0	0
CG2-CB-HB	95.715	108.0	0
HZ2-NZ-HZ3	121.179	109.0	1
OG-CB-HB3	96.789	109.0	0
CB-CG2-HG23	121.082	109.0	1
CG-CD-HD3	121.18	109.0	0
HD11-CD1-HD13	97.867	110.0	0
C-CA-HA3	96.869	109.0	0
HA2-CA-HA3	96.936	109.0	1
CZ-NH1-HH11	107.88	120.0	0
CG-ND2-HD21	107.932	120.0	0
CA-CB-HB1	121.049	109.0	0
NZ-CE-HE3	120.023	108.0	0

Too-close contacts

The following all-atom clashscore is based on a MolProbity analysis. All-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The table below contains clashscores for all the models in this entry.

Model ID	Clash score	Number of clashes
Model 1	0.00	0
Model 10	0.00	0
Model 2	0.00	0
Model 3	0.00	0

Model ID	Clash score	Number of clashes
Model 4	0.00	0
Model 5	0.00	0
Model 6	0.00	0
Model 7	0.00	0
Model 8	0.00	0
Model 9	0.00	0

All 0 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Torsion angles: Protein backbone

In the following table, Ramachandran outliers are listed. The Analysed column shows the number of residues for which the backbone conformation was analysed.

Model ID	Analyzed	Favored	Allowed	Outliers
1	182	147	32	3
2	182	150	28	4
3	182	169	10	3
4	182	154	26	2
5	182	161	14	7
6	182	153	23	6
7	182	159	19	4
8	182	153	23	6
9	182	156	19	7
10	182	155	23	4

Detailed list of outliers are tabulated below.

Model ID	Chain and res ID	Residue type
----------	------------------	--------------

Model ID	Chain and res ID	Residue type
1	A:55	PRO
1	A:74	GLY
1	A:136	PRO
2	A:21	1ASN
2	A:35	1LEU
2	A:55	1PRO
2	A:129	1PRO
3	A:55	PRO
3	A:72	VAL
3	A:124	HIS
4	A:55	PRO
4	A:99	PRO
5	A:10	SER
5	A:11	GLY
5	A:14	LEU
5	A:55	PRO
5	A:76	THR
5	A:117	LEU
5	A:144	ALA
6	A:55	PRO
6	A:70	ALA
6	A:103	VAL
6	A:126	GLN

Model ID	Chain and res ID	Residue type
6	A:146	PRO
6	A:163	GLU
7	A:44	LEU
7	A:47	SER
7	A:55	PRO
7	A:110	LEU
8	A:32	LYS
8	A:55	PRO
8	A:79	PRO
8	A:86	ALA
8	A:90	GLY
8	A:144	ALA
9	A:55	PRO
9	A:57	GLY
9	A:69	ASN
9	A:70	ALA
9	A:88	ALA
9	A:96	GLY
9	A:126	GLN
10	A:42	VAL
10	A:55	PRO
10	A:57	GLY
10	A:141	VAL

Torsion angles: Protein sidechains

In the following table, sidechain outliers are listed. The Analysed column shows the number of residues for which the sidechain conformation was analysed.

Model ID	Analyzed	Favored	Allowed	Outliers
1	147	123	14	10
2	147	133	12	2
3	147	132	9	6
4	147	125	16	6
5	147	108	24	15
6	147	126	13	8
7	147	116	19	12
8	147	119	21	7
9	147	124	15	8
10	147	119	21	7

Detailed list of outliers are tabulated below.

Model ID	Chain and res ID	Residue type
1	A:2	THR
1	A:18	ILE
1	A:26	LEU
1	A:72	VAL
1	A:109	MET
1	A:116	GLN
1	A:151	ARG
1	A:165	LEU
1	A:168	THR

Model ID	Chain and res ID	Residue type
1	A:184	TYR
2	A:106	SER
2	A:177	GLU
3	A:17	GLN
3	A:26	LEU
3	A:33	ILE
3	A:99	PRO
3	A:102	SER
3	A:140	THR
4	A:12	MET
4	A:51	VAL
4	A:91	SER
4	A:97	PHE
4	A:112	HIS
4	A:140	THR
5	A:14	LEU
5	A:39	LEU
5	A:42	VAL
5	A:43	TYR
5	A:60	TYR
5	A:97	PHE
5	A:103	VAL
5	A:117	LEU

Model ID	Chain and res ID	Residue type
5	A:120	PHE
5	A:135	GLU
5	A:154	SER
5	A:164	ARG
5	A:168	THR
5	A:169	ASN
5	A:181	GLU
6	A:4	THR
6	A:52	TYR
6	A:85	GLU
6	A:110	LEU
6	A:134	ASN
6	A:140	THR
6	A:165	LEU
6	A:177	GLU
7	A:1	MET
7	A:4	THR
7	A:31	LEU
7	A:35	LEU
7	A:45	ASP
7	A:91	SER
7	A:126	GLN
7	A:128	VAL

Model ID	Chain and res ID	Residue type
7	A:132	LEU
7	A:165	LEU
7	A:180	LYS
7	A:183	ARG
8	A:44	LEU
8	A:72	VAL
8	A:108	LEU
8	A:140	THR
8	A:143	GLU
8	A:159	GLN
8	A:168	THR
9	A:6	HIS
9	A:15	LEU
9	A:19	GLN
9	A:104	SER
9	A:126	GLN
9	A:127	GLN
9	A:174	MET
9	A:178	SER
10	A:35	LEU
10	A:39	LEU
10	A:42	VAL
10	A:78	LEU

Model ID	Chain and res ID	Residue type
10	A:100	LEU
10	A:129	PRO
10	A:176	MET

Fit of model to data used for modeling

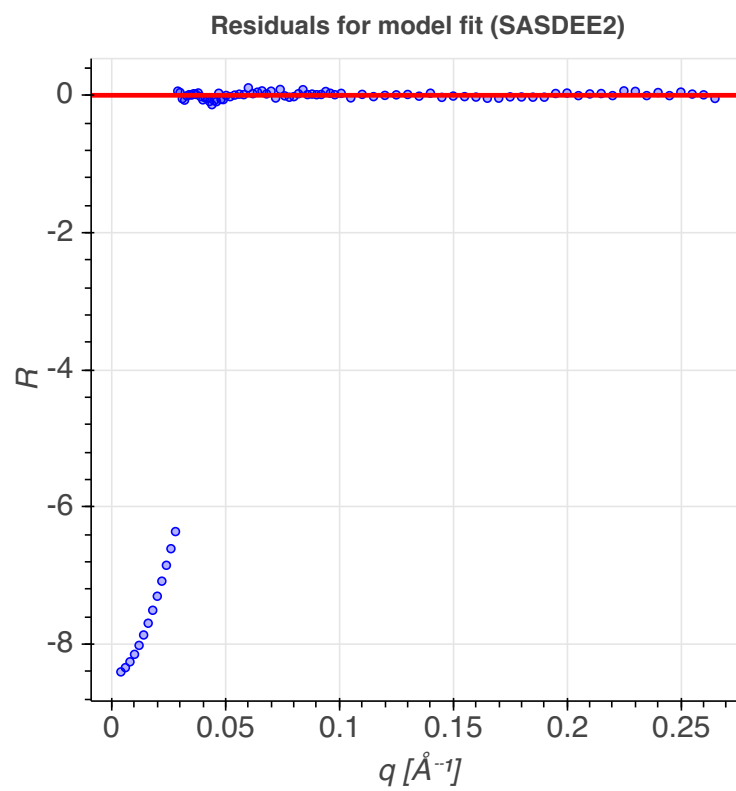
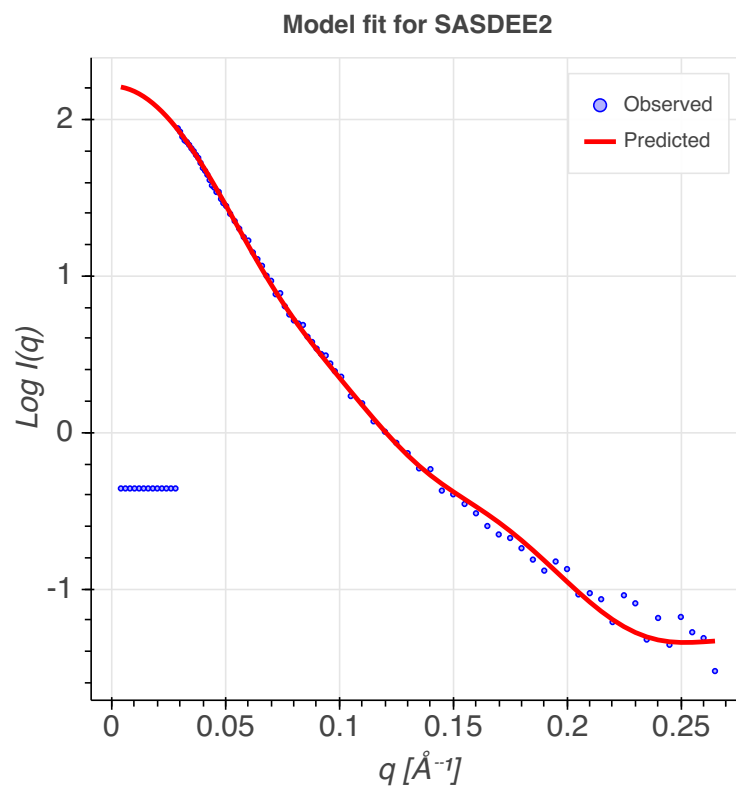
Fit of model(s) to SAS data

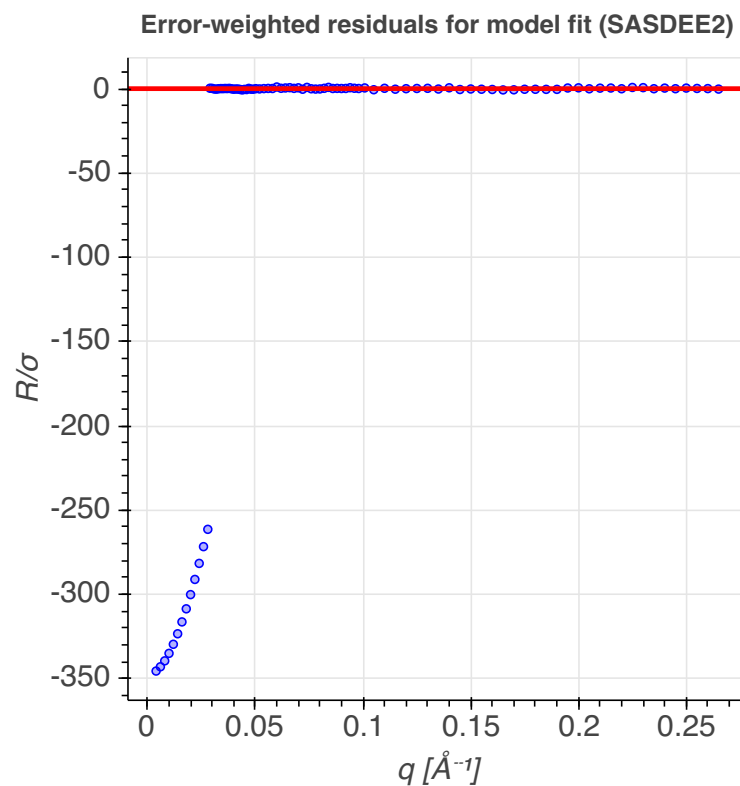
χ^2 goodness of fit analysis

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 zero.

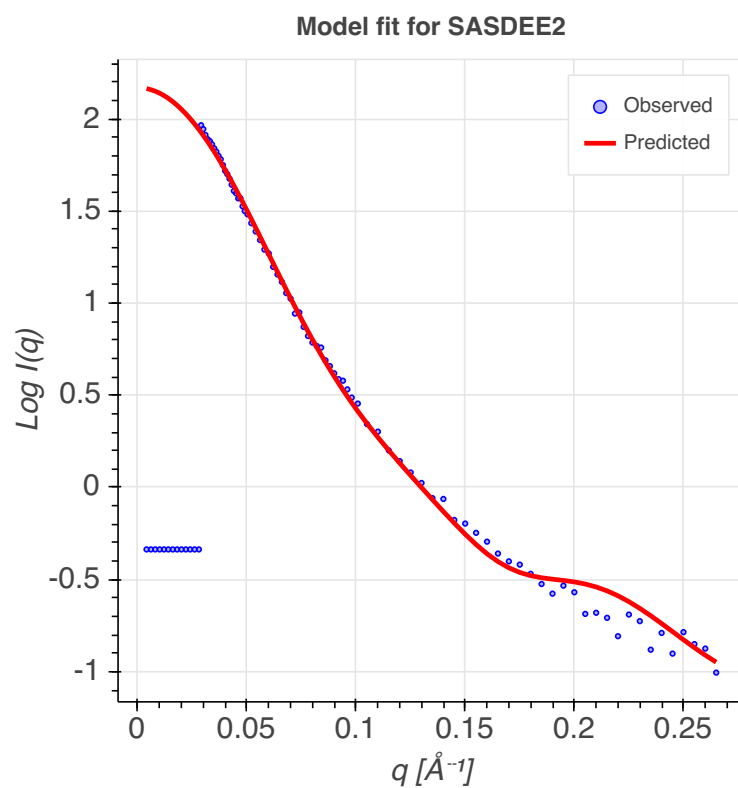
SASDB ID	Model	χ^2
SASDEE2	1	2.73
SASDEE2	2	2.32
SASDEE2	3	1.98

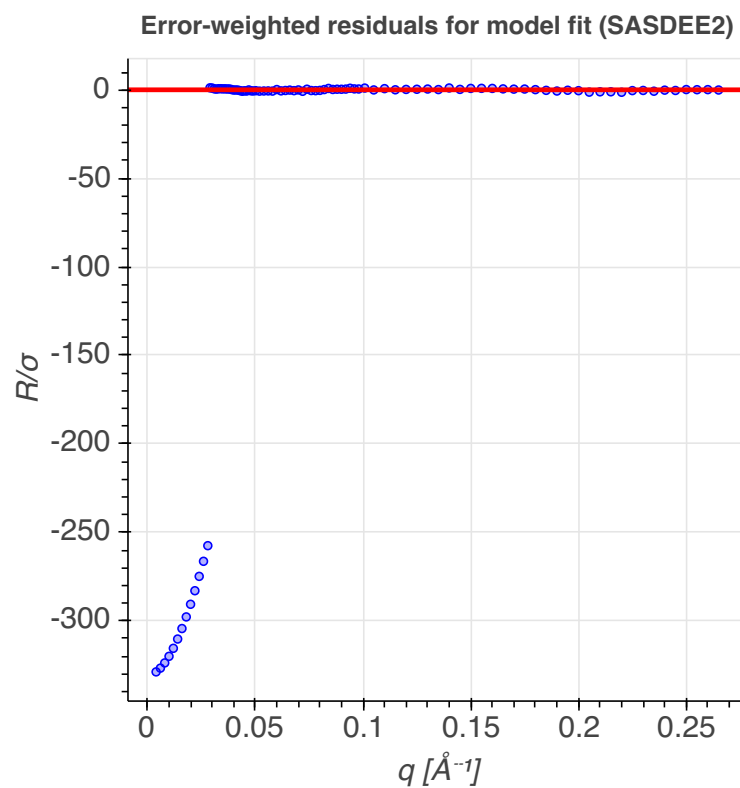
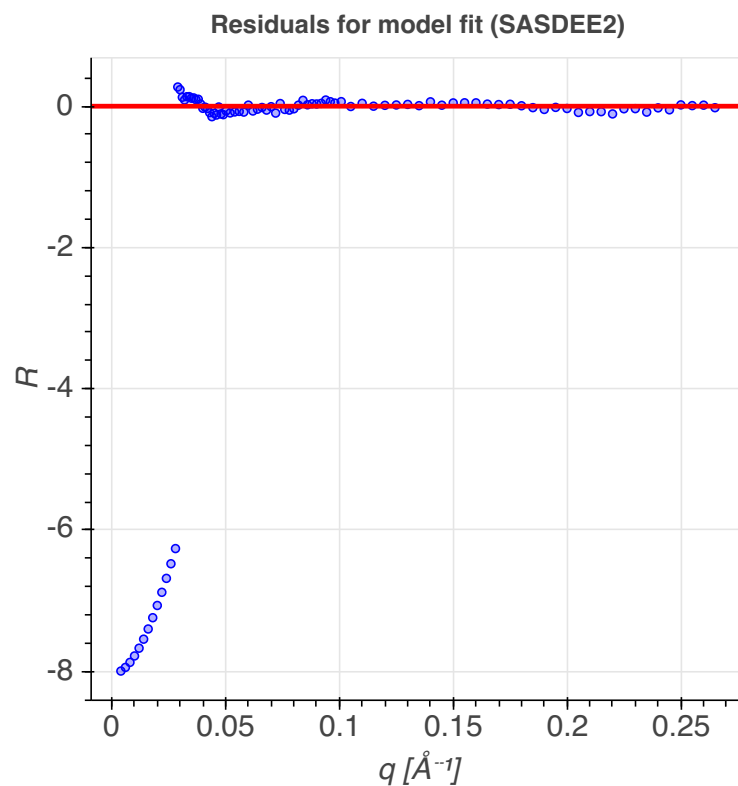
Model fit for SASDEE2 (fit/model number 1): Residual value plot is a measure to assess fit to the data. sResidual values should be equally and randomly spaced around the horizontal axis.



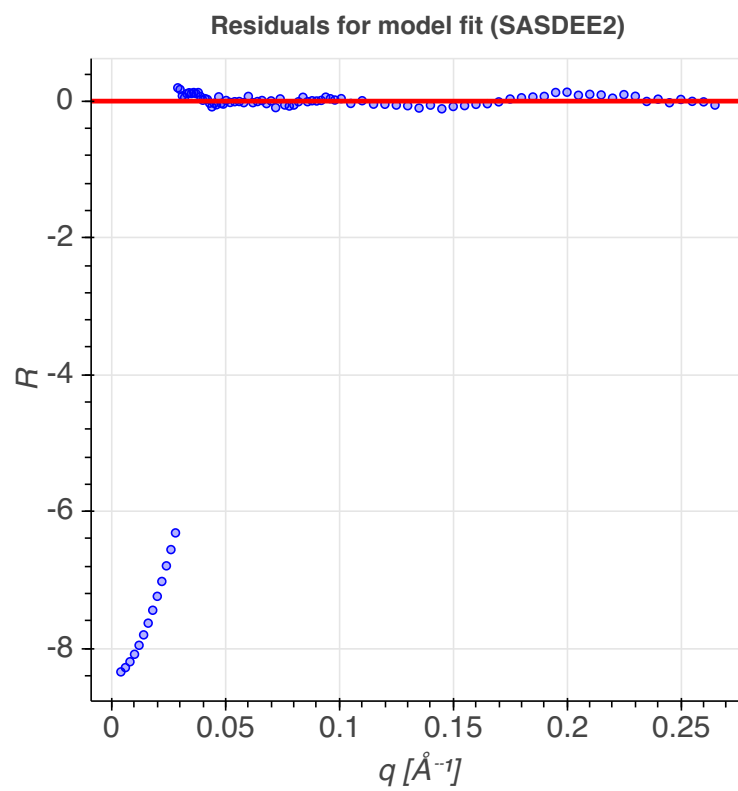
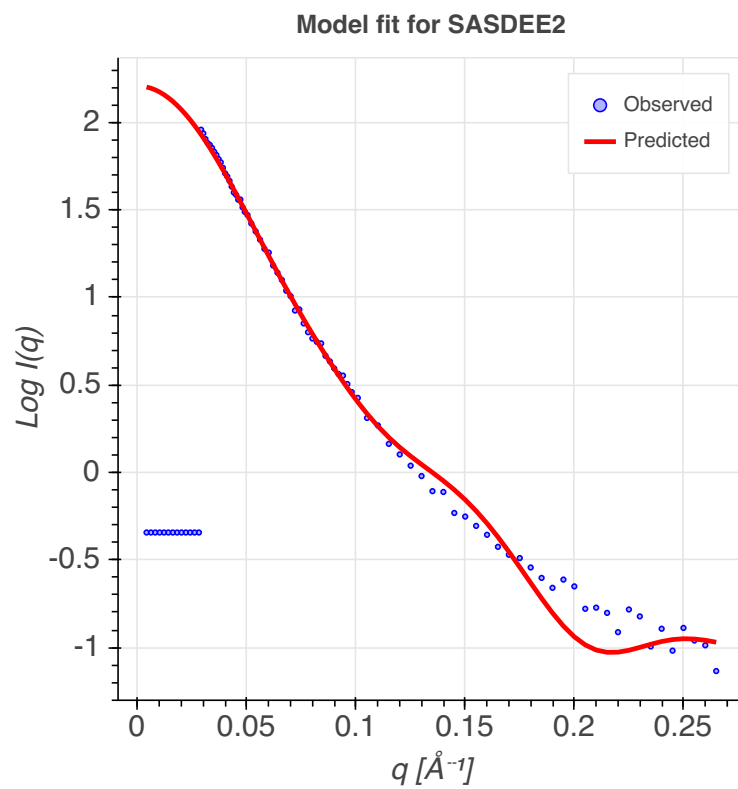


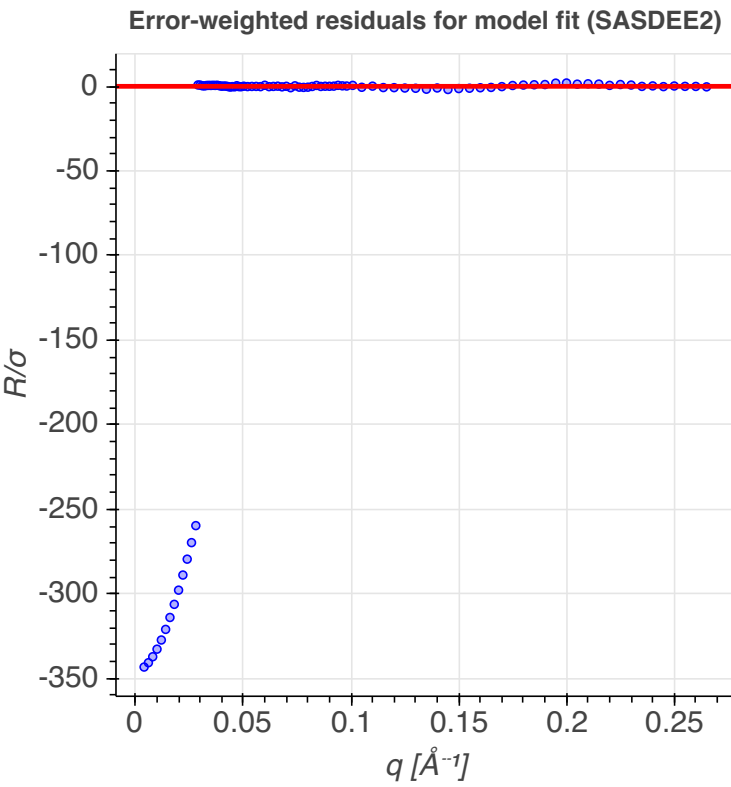
Model fit for SASDEE2 (fit/model number 2): Residual value plot is a measure to assess fit to the data. sResidual values should be equally and randomly spaced around the horizontal axis.





Model fit for SASDEE2 (fit/model number 3): Residual value plot is a measure to assess fit to the data. Residual values should be equally and randomly spaced around the horizontal axis.





Cormap p-value analysis of fits

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	p-value
SASDEE2	1	9.97E-03
SASDEE2	2	2.44E-03
SASDEE2	3	4.94E-03

Fit of model to data not used for modeling

Uncertainty of data and model

