Full wwPDB Integrative Structure Validation Report

June 26, 2020 -- 11:59 AM

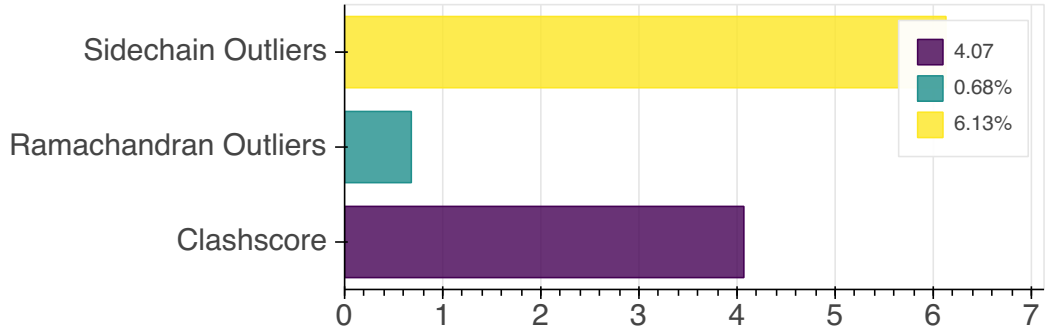
The following software were used in the production of this report:

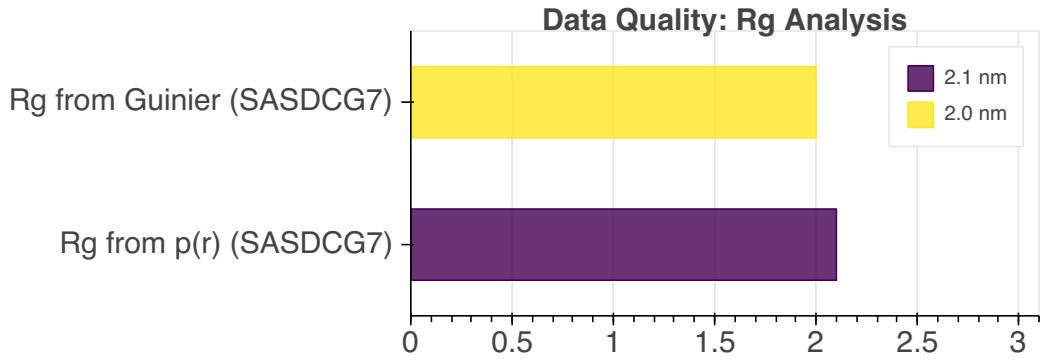
Molprobity : Version 4.4
Integrative Modeling Validation Package : Version 1.0

PDB ID	PDBDEV_00000004
Molecule Name	Structure of K63-linked Diubiquitin
Title	Characterizing Protein Dynamics with Integrative Use of Bulk and Single-Molecule Techniques
Authors	Liu Z;Gong Z;Cao Y;Ding YH;Dong MQ;Lu YB;Zhang WP;Tang C

Overall quality

Model Quality: Molprobity Analysis





Ensemble information

This entry consists of 0 distinct ensemble.

Summary

This entry consists of 3 unique models, with 2 subunits in each model. A total of 5 datasets or restraints was used to build this entry. Each model is represented by 0 rigid bodies and 2 flexible or non-rigid units.

Entry composition

There are 3 unique types of models in this entry.These models are titled None/Model 1, None/Model 2, None/Model 3 respectively.

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Total residues
1	1	1	Ubiquitin	A	76
1	2	1	Ubiquitin	B	76
2	1	1	Ubiquitin	A	76
2	2	1	Ubiquitin	B	76
3	1	1	Ubiquitin	A	76
3	2	1	Ubiquitin	B	76

Methodology and software

Step number	Protocol ID	Method name	Method type	Number of computed models	Multi state modeling	Multi scale modeling
1	1	None	None	None	True	None

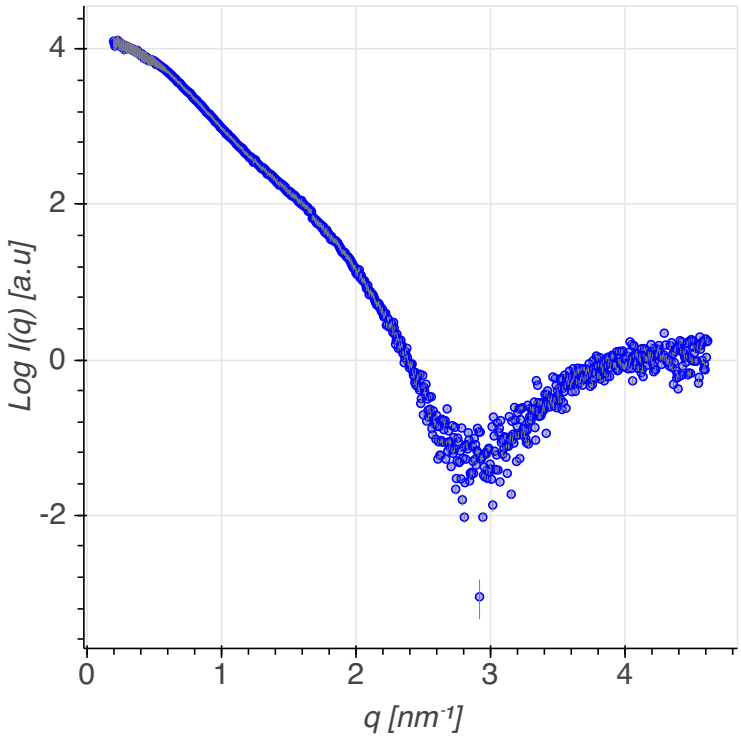
Data quality

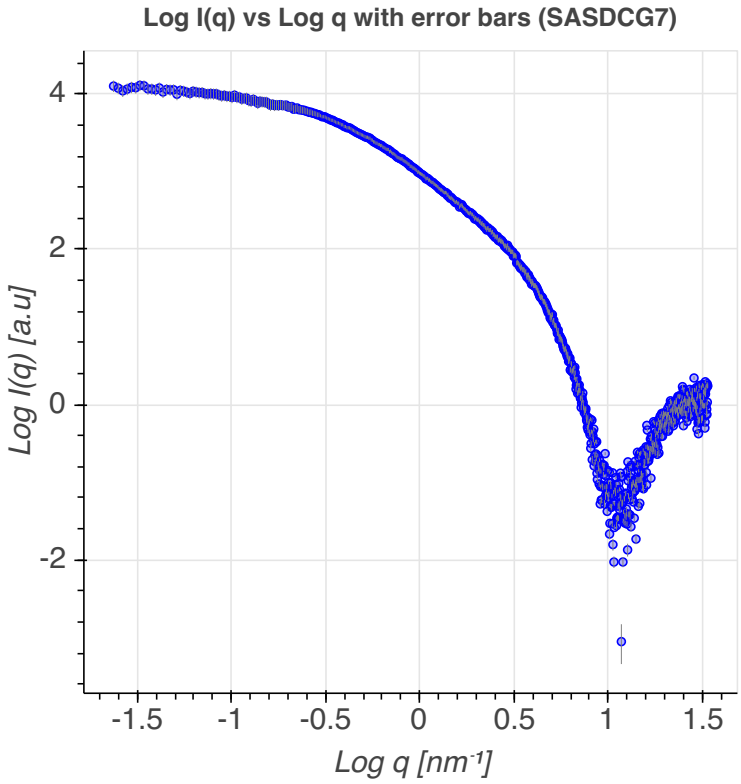
Scattering profile

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

Scattering profile for SASDCG7: 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.

Log I(q) vs q with error bars (SASDCG7)





Key experimental estimates

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

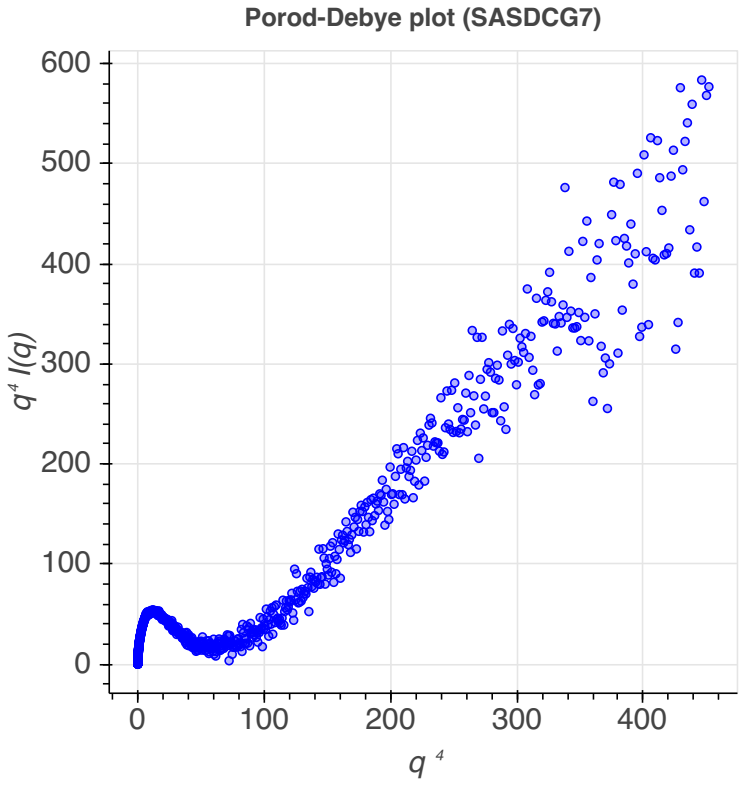
SASDB ID	Chemical composition MW	Standard MW	Porod MW
SASDCG7	13.0 kDa	N/A	13.0 kDa

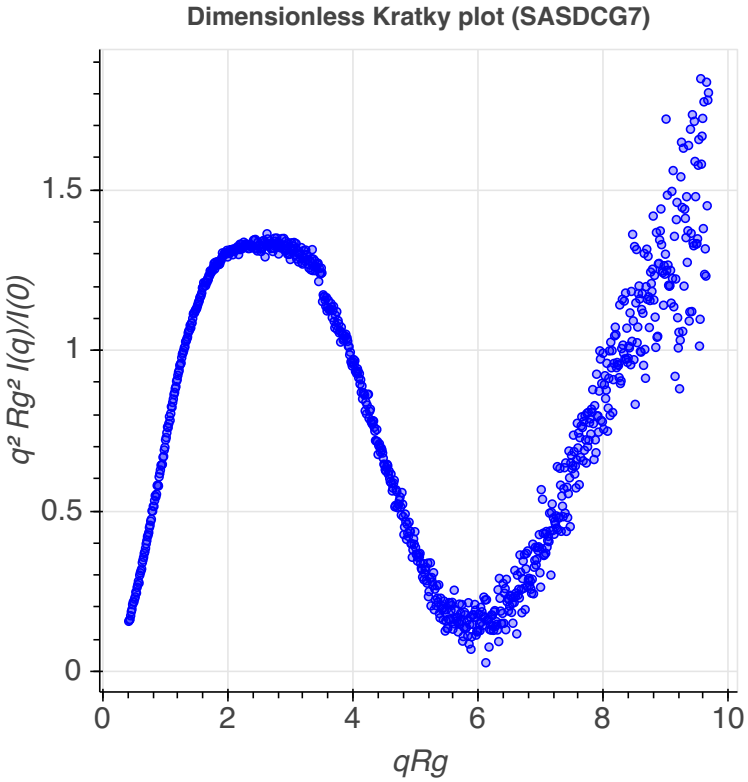
Volume estimates from experiments and analysis: estimated volume can be compared to Porod volume obtained from scattering profiles.

SASDB ID	Estimated volume	Estimated volume method	Porod volume
SASDCG7	None	None	22.00 nm³

Flexibility analysis

Flexibility analysis for SASDCG7: 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.





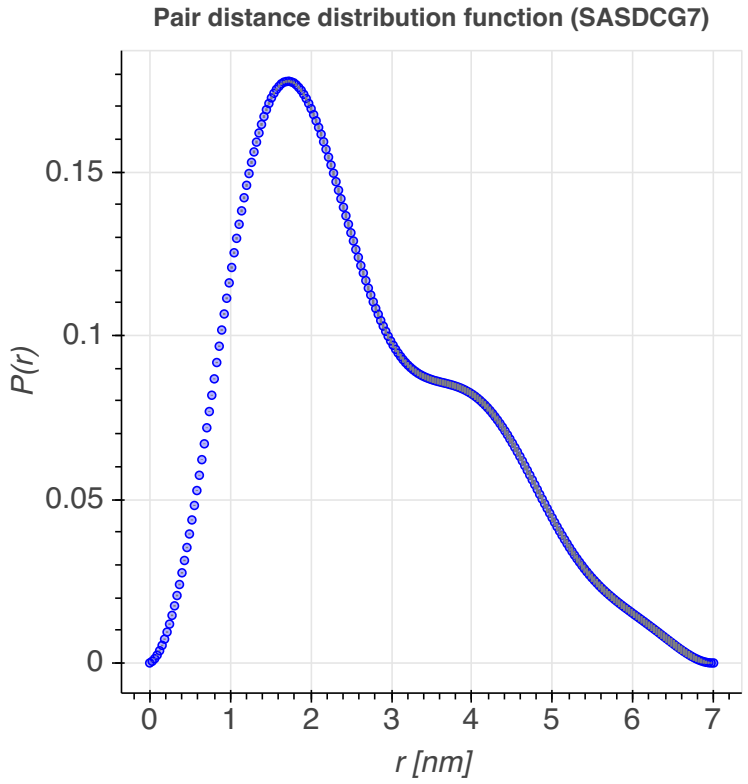
Pair-distance distribution analysis

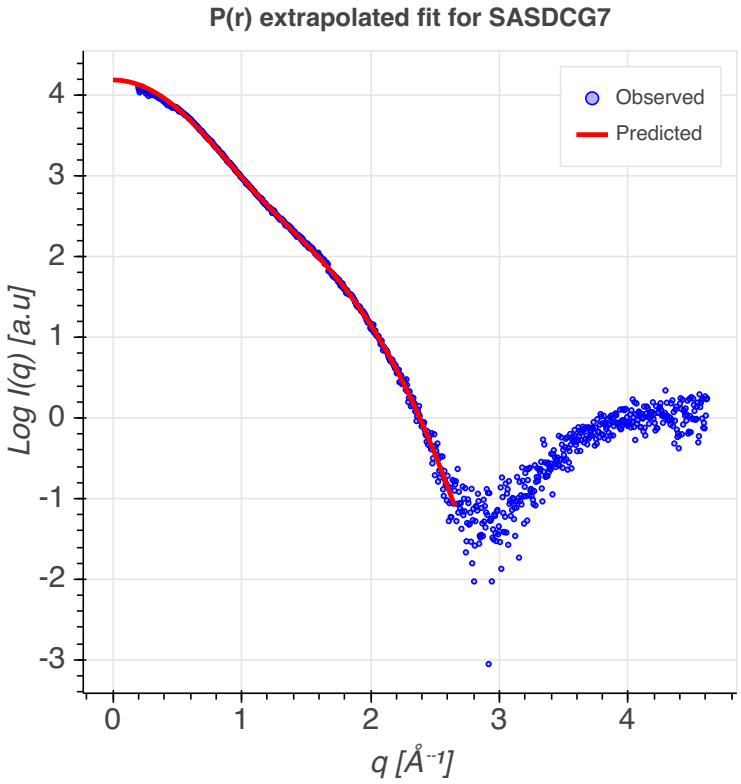
P(r) analysis: p(r) represents the distribution of distances between all pairs of atoms within the particle weighted by the respective electron densities. p(r) is the Fourier transform of I(s) (and vice versa). Rg can be estimated from integrating the p(r) function. Agreement between the p(r) and Guinier-

determined Rg (table below) is a good measure of the self-consistency of the SAS profile. Rg is a measure for the overall size of a macromolecule; e.g. a protein with a smaller Rg is more compact than a protein with a larger Rg, provided both have the same molecular weight (MW). The point where p(r) is decaying to zero is called Dmax and represents the maximum size of the particle.

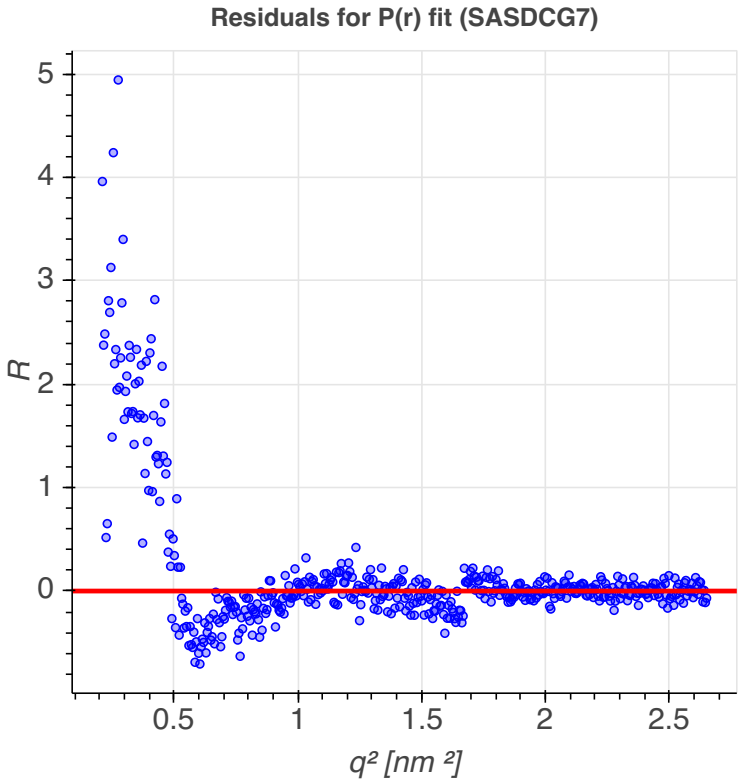
SASDB ID	Software used	Dmax	Dmax error	Rg	Rg error
SASDCG7	ATSAS GNOM	7.00 nm	N/A	2.100 nm	0.010 nm

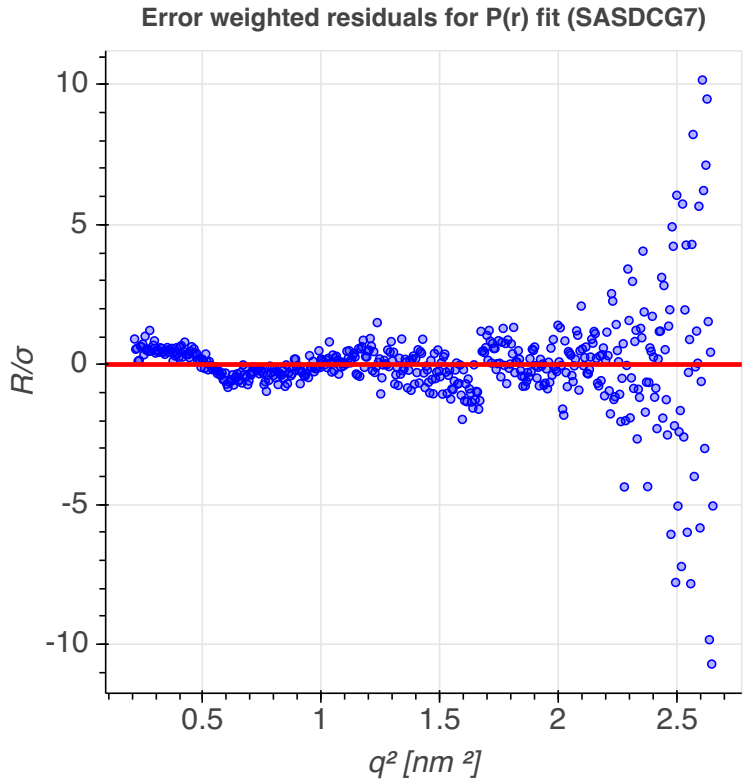
P(r) for SASDCG7: The value of P(r) should be zero beyond r=Dmax.





Residuals and error weighted residuals for P(r) analysis for SASDCG7: Residual value plot is a measure to assess fit to the data. sResidual values should be equally and randomly spaced around the horizontal axis.



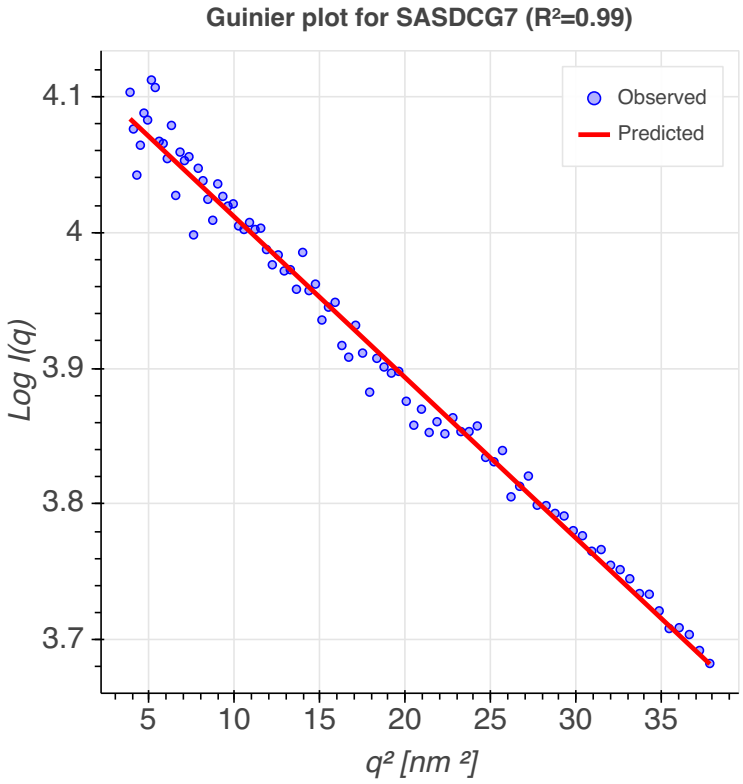


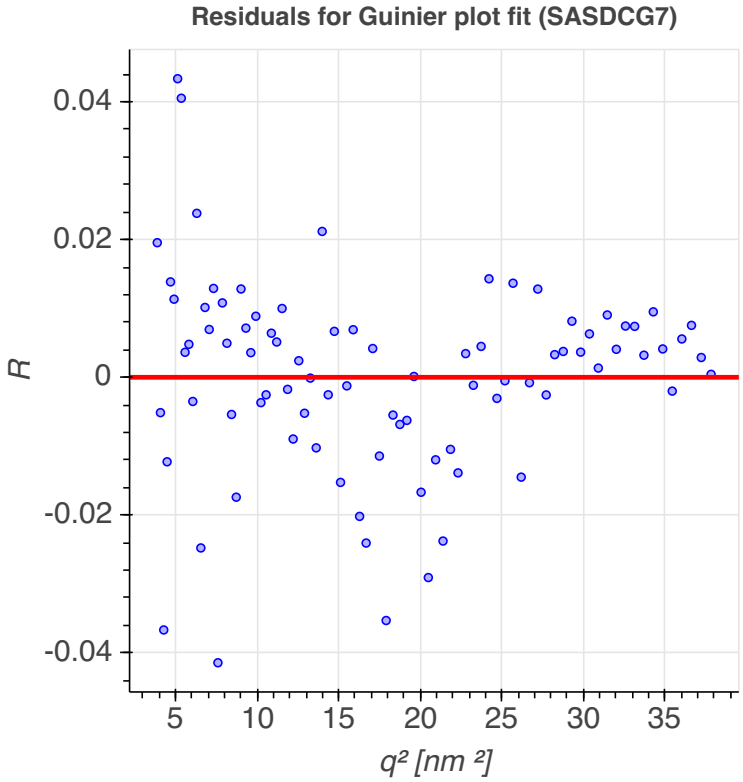
Guinier analysis

Guinier analysis: agreement between the p(r) and Guinier-determined Rg (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
SASDCG7	2.0 nm	0.150 nm	N/A	N/A

Guinier analysis for SASDCG7: 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 samesize. 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 3774 bond outliers in this entry.

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NH1--HH12	0.999	0.86	23
NZ--HZ2	1.039	0.89	38
NZ--HZ1	1.038	0.89	41
NE--HE	1.008	0.86	23
N--H2	1.04	0.89	5
CE1--HE1	1.078	0.93	23
CD1--HD1	1.079	0.93	17
CZ--HZ	1.078	0.93	11
CD2--HD2	1.078	0.93	23
CE2--HE2	1.078	0.93	17
NZ--HZ3	1.038	0.89	38
N--H3	1.038	0.89	5
NH2--HH22	0.999	0.86	23
N--H1	1.039	0.89	5
NH2--HH21	0.985	0.86	23
NH1--HH11	0.977	0.86	23
ND2--HD22	0.979	0.86	11
NE2--HE22	0.979	0.86	35
CD--HD2	1.077	0.97	83
OG1--HG1	0.959	0.84	41
N--H	0.978	0.86	431
NE2--HE21	0.958	0.86	35
OG--HG	0.959	0.84	17

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
ND1--HD1	0.979	0.86	5
ND2--HD21	0.956	0.86	11
OH--HH	0.958	0.84	5
NE2--HE2	0.978	0.86	5
CB--HB2	1.073	0.97	311
CA--HA	1.076	0.97	419
CB--HB	1.079	0.97	107
CB--HB3	1.069	0.97	311
CG--HG3	1.076	0.97	161
CG--HG2	1.074	0.97	161
CG--HG	1.077	0.97	53
CA--HA2	1.076	0.97	35
CA--HA3	1.076	0.97	35
CG1--HG13	1.078	0.97	65
CE--HE2	1.077	0.97	47
CD--HD3	1.057	0.97	83
CD1--HD13	1.078	0.97	95
CG1--HG12	1.077	0.97	65
CE--HE3	1.078	0.97	47
CG2--HG23	1.078	0.97	107
CD2--HD21	1.077	0.97	53
CD1--HD12	1.077	0.97	95
CD2--HD22	1.079	0.97	53

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CG2--HG21	1.077	0.97	107
CG2--HG22	1.078	0.97	107
CD1--HD11	1.078	0.97	95
CD2--HD23	1.078	0.97	53
CE--HE1	1.081	0.97	5
CG1--HG11	1.079	0.97	23
CB--HB1	1.079	0.97	11

There are 0 angle outliers in this entry.

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	4.07	10
Model 2	0.00	0
Model 3	0.00	0

All 10 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	148	144	3	1
2	148	146	1	1
3	148	145	2	1

Detailed list of outliers are tabulated below.

Model ID	Chain and res ID	Residue type
1	B:73	LEU
2	B:73	LEU
3	B:75	GLY

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	136	116	12	8
2	136	116	12	8
3	136	114	13	9

Detailed list of outliers are tabulated below.

Model ID	Chain and res ID	Residue type
1	A:13	ILE
1	A:15	LEU
1	A:39	ASP
1	A:71	LEU
1	B:13	ILE
1	B:15	LEU
1	B:39	ASP
1	B:71	LEU
2	A:13	ILE
2	A:15	LEU
2	A:39	ASP
2	A:71	LEU

Model ID	Chain and res ID	Residue type
2	B:13	ILE
2	B:15	LEU
2	B:39	ASP
2	B:71	LEU
3	A:13	ILE
3	A:15	LEU
3	A:39	ASP
3	A:71	LEU
3	A:74	ARG
3	B:13	ILE
3	B:15	LEU
3	B:39	ASP
3	B:71	LEU

Fit of model to data used for modeling

Fit of model(s) to SAS data

χ^2 goodness of fit analysis

Model(s) and/or fit for this entry has not been deposited.

Cormap p-value analysis of fits

Model(s) and/or fit for this entry has not been deposited.

Fit of model to data not used for modeling

Uncertainty of data and model