



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 09:57 PM EDT

PDB ID : 5XGC  
Title : Crystal structure of SmgGDS-558  
Authors : Shimizu, H.; Toma-Fukai, S.; Shimizu, T.  
Deposited on : 2017-04-13  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

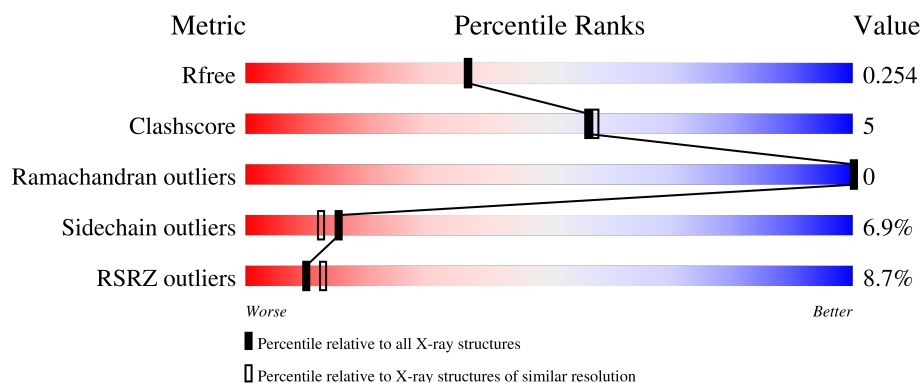
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	503	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3531 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Rap1 GTPase-GDP dissociation stimulator 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	473	Total	C	N	O	S	0	0	0
			3457	2178	602	656	21			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	56	GLY	-	expression tag	UNP P52306
A	57	PRO	-	expression tag	UNP P52306
A	58	LEU	-	expression tag	UNP P52306
A	59	GLY	-	expression tag	UNP P52306
A	60	SER	-	expression tag	UNP P52306

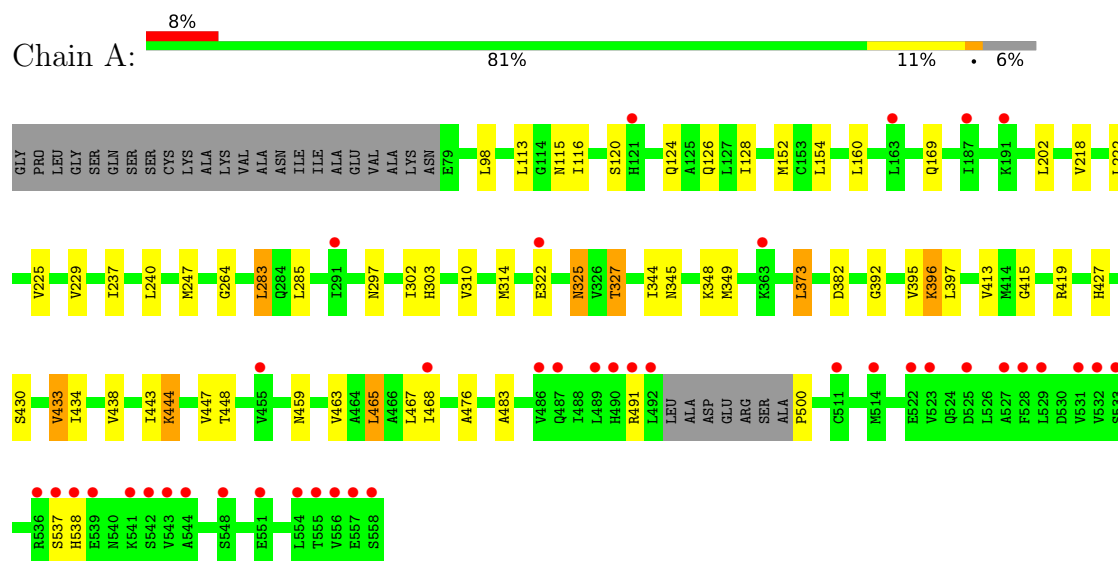
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	74	Total	O	0	0
			74	74		

### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Rap1 GTPase-GDP dissociation stimulator 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	194.33Å 51.12Å 52.44Å 90.00° 92.77° 90.00°	Depositor
Resolution (Å)	49.43 – 2.10 40.11 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.43-2.10) 99.9 (40.11-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.44 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.215 , 0.252 0.220 , 0.254	Depositor DCC
$R_{free}$ test set	1546 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.5	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 64.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3531	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.53% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.40	0/3490	0.61	1/4724 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	500	PRO	N-CA-CB	5.52	109.92	103.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3457	0	3415	34	0
2	A	74	0	0	1	0
All	All	3531	0	3415	34	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (34) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:325:ASN:HD21	1:A:327:THR:HB	1.48	0.76

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:382:ASP:O	1:A:427:HIS:CE1	2.48	0.67
1:A:325:ASN:ND2	1:A:327:THR:HB	2.12	0.64
1:A:222:LEU:HD23	1:A:247:MET:HE1	1.78	0.64
1:A:225:VAL:O	1:A:229:VAL:HG23	1.97	0.63
1:A:537:SER:HA	1:A:538:HIS:C	2.18	0.63
1:A:310:VAL:HG12	1:A:314:MET:HE3	1.82	0.62
1:A:264:GLY:HA3	1:A:303:HIS:CE1	2.36	0.60
1:A:438:VAL:HG12	1:A:443:ILE:HD11	1.84	0.60
1:A:373:LEU:C	1:A:373:LEU:HD12	2.24	0.57
1:A:438:VAL:HG12	1:A:443:ILE:CD1	2.36	0.56
1:A:128:ILE:HD12	1:A:169:GLN:HG3	1.88	0.56
1:A:240:LEU:HD21	1:A:283:LEU:HD13	1.89	0.55
1:A:285:LEU:HD13	1:A:327:THR:CG2	2.37	0.54
1:A:218:VAL:HG22	2:A:604:HOH:O	2.08	0.52
1:A:285:LEU:HD13	1:A:327:THR:HG23	1.91	0.51
1:A:310:VAL:HG12	1:A:314:MET:CE	2.41	0.51
1:A:468:ILE:HD11	1:A:476:ALA:HB2	1.93	0.50
1:A:444:LYS:O	1:A:448:THR:HG23	2.12	0.50
1:A:124:GLN:O	1:A:128:ILE:HG12	2.13	0.49
1:A:397:LEU:C	1:A:397:LEU:HD23	2.33	0.49
1:A:113:LEU:HD23	1:A:116:ILE:HD11	1.95	0.48
1:A:344:ILE:HG22	1:A:348:LYS:HE3	1.96	0.48
1:A:434:ILE:O	1:A:438:VAL:HG13	2.14	0.47
1:A:310:VAL:HG21	1:A:349:MET:HG2	1.97	0.46
1:A:415:GLY:O	1:A:419:ARG:HG3	2.16	0.46
1:A:392:GLY:HA3	1:A:433:VAL:HB	1.98	0.45
1:A:302:ILE:HA	1:A:345:ASN:HD21	1.81	0.44
1:A:325:ASN:C	1:A:325:ASN:HD22	2.22	0.43
1:A:459:ASN:O	1:A:463:VAL:HG23	2.19	0.43
1:A:447:VAL:HG11	1:A:483:ALA:HB1	2.02	0.42
1:A:222:LEU:CD2	1:A:247:MET:HE1	2.47	0.41
1:A:396:LYS:HE3	1:A:396:LYS:O	2.21	0.41
1:A:465:LEU:HA	1:A:468:ILE:HG22	2.03	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	469/503 (93%)	451 (96%)	18 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	347/426 (82%)	323 (93%)	24 (7%)	15	12

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	98	LEU
1	A	115	ASN
1	A	120	SER
1	A	126	GLN
1	A	152	MET
1	A	154	LEU
1	A	160	LEU
1	A	202	LEU
1	A	237	ILE
1	A	283	LEU
1	A	297	ASN
1	A	322	GLU
1	A	325	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	327	THR
1	A	373	LEU
1	A	395	VAL
1	A	396	LYS
1	A	413	VAL
1	A	430	SER
1	A	433	VAL
1	A	444	LYS
1	A	465	LEU
1	A	467	LEU
1	A	491	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	108	GLN
1	A	258	GLN
1	A	297	ASN
1	A	303	HIS
1	A	325	ASN
1	A	345	ASN
1	A	458	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

### 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	473/503 (94%)	0.70	41 (8%) <b>10</b> <b>13</b>	39, 59, 124, 157	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	537	SER	11.7
1	A	532	VAL	7.5
1	A	187	ILE	6.6
1	A	528	PHE	6.4
1	A	555	THR	6.1
1	A	558	SER	5.9
1	A	531	VAL	5.8
1	A	554	LEU	5.6
1	A	533	SER	4.9
1	A	551	GLU	4.7
1	A	527	ALA	4.7
1	A	541	LYS	4.7
1	A	538	HIS	4.7
1	A	542	SER	4.6
1	A	557	GLU	4.6
1	A	536	ARG	4.5
1	A	556	VAL	4.4
1	A	548	SER	4.3
1	A	492	LEU	3.8
1	A	486	VAL	3.4
1	A	322	GLU	3.2
1	A	543	VAL	3.2
1	A	121	HIS	3.0
1	A	523	VAL	3.0
1	A	455	VAL	2.8
1	A	511	CYS	2.8
1	A	525	ASP	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	489	LEU	2.7
1	A	191	LYS	2.7
1	A	490	HIS	2.6
1	A	491	ARG	2.6
1	A	363	LYS	2.4
1	A	487	GLN	2.3
1	A	163	LEU	2.3
1	A	544	ALA	2.3
1	A	514	MET	2.1
1	A	529	LEU	2.1
1	A	539	GLU	2.1
1	A	291	ILE	2.1
1	A	522	GLU	2.0
1	A	468	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.