Full wwPDB X-ray Structure Validation Report ⓘ

Aug 19, 2023 – 06:13 PM EDT

PDB ID	:	2G2S
Title	:	Structure of S65G Y66S GFP variant after spontaneous peptide hydrolysis
Authors	:	Barondeau, D.P.
Deposited on	:	2006-02-16
Resolution	:	1.20 Å(reported)

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

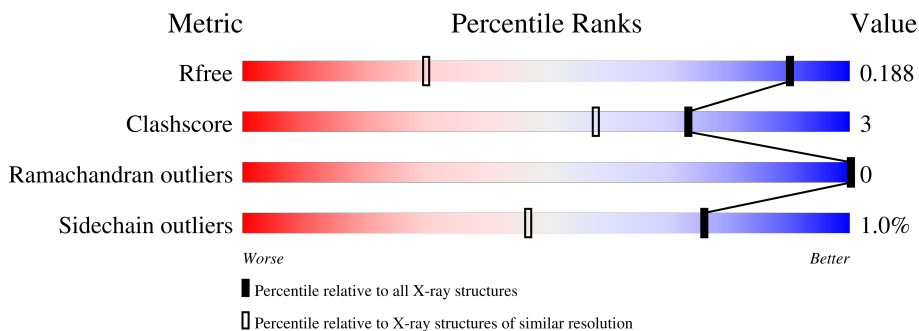|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.20 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	Percentile Ranks	Value
Rfree		0.188
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	66	86% 8% · ·
2	B	173	87% 7% · 5%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2156 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	64	Total	C	N	O	S	0	1	0
			476	307	73	95	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P42212
A	1	ALA	-	cloning artifact	UNP P42212
A	64	LEU	PHE	engineered mutation	UNP P42212
A	65	GLY	SER	engineered mutation	UNP P42212

- Molecule 2 is a protein called Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	165	Total	C	N	O	S	0	7	0
			1354	849	238	263	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	66	DHA	TYR	SEE REMARK 999	UNP P42212
B	99	SER	PHE	engineered mutation	UNP P42212
B	153	THR	MET	engineered mutation	UNP P42212
B	163	ALA	VAL	engineered mutation	UNP P42212

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	88	Total 88	O 88	0	0
4	B	237	Total 237	O 237	0	0

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. 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. 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: Green fluorescent protein

Chain A:

86%   8%   · ·

MET ALA S2 K3 G4 E5 E6 D19 K52 L64 G65

• Molecule 2: Green fluorescent protein

Chain B:

87%   7% · 5%

S66 G67 R96 Y106 R109 D117 R122 I136 Y145 Q157 K162 R168 Y182 D190 T230 HIS GLY MET ASP GLU LEU TYR LYS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.00Å 62.30Å 68.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.80 – 1.20 46.09 – 1.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (17.80-1.20) 97.2 (46.09-1.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	4.63 (at 1.20Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	(Not available) , 0.174 0.172 , 0.188	Depositor DCC
R_{free} test set	3423 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	11.6	Xtriage
Anisotropy	0.555	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 58.4	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2156	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.46% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, DHA

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	0.71	0/490	1.17	4/665 (0.6%)
2	B	0.70	0/1407	1.26	12/1899 (0.6%)
All	All	0.70	0/1897	1.23	16/2564 (0.6%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	109	ARG	NE-CZ-NH2	-10.24	115.18	120.30
2	B	168[A]	ARG	NE-CZ-NH2	8.60	124.60	120.30
2	B	168[B]	ARG	NE-CZ-NH2	8.60	124.60	120.30
2	B	109	ARG	NE-CZ-NH1	8.48	124.54	120.30
2	B	168[A]	ARG	CD-NE-CZ	7.03	133.44	123.60
2	B	168[B]	ARG	CD-NE-CZ	7.03	133.44	123.60
2	B	168[A]	ARG	NE-CZ-NH1	-6.75	116.93	120.30
2	B	168[B]	ARG	NE-CZ-NH1	-6.75	116.93	120.30
2	B	106	TYR	CB-CG-CD1	6.70	125.02	121.00
1	A	19	ASP	CB-CG-OD2	6.57	124.21	118.30
2	B	145	TYR	CB-CG-CD2	-6.03	117.38	121.00
1	A	64	LEU	CB-CA-C	-5.95	98.89	110.20
2	B	96	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	A	19	ASP	CB-CG-OD1	-5.71	113.16	118.30
2	B	122	ARG	CD-NE-CZ	5.05	130.67	123.60
1	A	6	GLU	OE1-CD-OE2	-5.02	117.28	123.30

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	476	0	478	6	0
2	B	1354	0	1300	8	0
3	B	1	0	0	0	0
4	A	88	0	0	1	0
4	B	237	0	0	5	0
All	All	2156	0	1778	12	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 3.

All (12) 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)
2:B:66:DHA:HB1	4:B:1001:HOH:O	1.76	0.85
1:A:3:LYS:HG2	4:B:1266:HOH:O	1.87	0.75
2:B:117:ASP:HB2	4:B:1281:HOH:O	1.93	0.67
2:B:162:LYS:HE2	2:B:182:TYR:OH	2.01	0.61
1:A:52:LYS:HE2	4:B:1205:HOH:O	2.00	0.59
1:A:6:GLU:HG2	4:A:1300:HOH:O	2.05	0.57
1:A:5:GLU:HG3	1:A:5:GLU:O	2.06	0.55
2:B:66:DHA:HB2	2:B:96:ARG:HH22	1.76	0.50
1:A:65:GLY:C	2:B:66:DHA:H2	2.17	0.47
2:B:157:GLN:HG3	4:B:1233:HOH:O	2.15	0.45
2:B:136:ILE:HD12	2:B:136:ILE:N	2.35	0.41
1:A:65:GLY:O	2:B:67:GLY:N	2.52	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	63/66 (96%)	63 (100%)	0	0	100	100
2	B	170/173 (98%)	167 (98%)	3 (2%)	0	100	100
All	All	233/239 (98%)	230 (99%)	3 (1%)	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	54/54 (100%)	53 (98%)	1 (2%)	57	19
2	B	151/151 (100%)	150 (99%)	1 (1%)	84	59
All	All	205/205 (100%)	203 (99%)	2 (1%)	76	47

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	LYS
2	B	190	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3　RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	DHA	B	66	2	4,4,5	1.90	1 (25%)	2,4,6	2.08	1 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DHA	B	66	2	-	0/0/2/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	66	DHA	CA-N	3.29	1.43	1.35

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	66	DHA	CB-CA-N	2.91	132.72	125.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	66	DHA	3	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.