



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:36 am BST

PDB ID : 2Z6X
Title : Crystal structure of 22G, the wild-type protein of the photoswitchable GFP-like protein Dronpa
Authors : Kikuchi, A.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Mizuno, H.; Miyawaki, A.
Deposited on : 2007-08-09
Resolution : 2.30 Å(reported)

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

We welcome your comments at 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 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 : 2.11
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) : 2.11

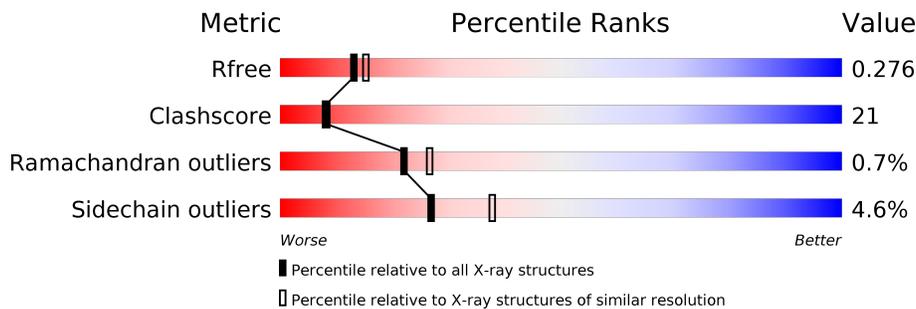
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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 on 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	255	51% 30% 15%
1	B	255	51% 30% 15%
1	C	255	54% 29% 16%
1	D	255	53% 30% 16%
1	E	255	49% 32% 15%
1	F	255	54% 29% 15%
1	G	255	58% 23% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	255	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	GYC	A	63	X	-	-	-
1	GYC	B	63	X	-	-	-
1	GYC	C	63	X	-	-	-
1	GYC	D	63	X	-	-	-
1	GYC	E	63	X	-	-	-
1	GYC	F	63	X	-	-	-
1	GYC	G	63	X	-	-	-
1	GYC	H	63	X	-	X	-

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15008 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 photochromic protein Dronpa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	217	Total 1750	C 1122	N 295	O 323	S 10	0	0	0
1	B	216	Total 1743	C 1117	N 294	O 322	S 10	0	0	0
1	C	215	Total 1737	C 1114	N 293	O 320	S 10	0	0	0
1	D	215	Total 1737	C 1114	N 293	O 320	S 10	0	0	0
1	E	216	Total 1743	C 1117	N 294	O 322	S 10	0	0	0
1	F	216	Total 1743	C 1117	N 294	O 322	S 10	0	0	0
1	G	215	Total 1737	C 1114	N 293	O 320	S 10	0	0	0
1	H	215	Total 1737	C 1114	N 293	O 320	S 10	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	113	Total 113	O 113	0	0
2	B	140	Total 140	O 140	0	0
2	C	157	Total 157	O 157	0	0
2	D	139	Total 139	O 139	0	0
2	E	148	Total 148	O 148	0	0
2	F	145	Total 145	O 145	0	0

Continued on next page...

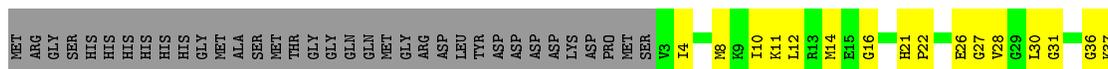
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	128	Total 128	O 128	0	0
2	H	111	Total 111	O 111	0	0



- Molecule 1: photochromic protein Dronpa

Chain H: 50% 33% 16%



4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	111.68Å 111.68Å 156.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 2.30 42.10 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.96-2.30) 99.1 (42.10-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.98 (at 2.29Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.219 , 0.278 0.216 , 0.276	Depositor DCC
R_{free} test set	4314 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	24.4	Xtrriage
Anisotropy	0.013	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 9.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.310 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15008	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0/1774	0.67	0/2394
1	B	0.40	0/1766	0.68	0/2382
1	C	0.43	0/1760	0.71	0/2374
1	D	0.43	0/1760	0.71	0/2374
1	E	0.42	0/1766	0.68	0/2382
1	F	0.45	0/1766	0.70	0/2382
1	G	0.41	0/1760	0.71	0/2374
1	H	0.40	0/1760	0.66	0/2374
All	All	0.42	0/14112	0.69	0/19036

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0
1	B	1	0
1	C	1	0
1	D	1	0
1	E	1	0
1	F	1	0
1	G	1	0
1	H	1	0
All	All	8	0

There are no bond length outliers.

There are no bond angle outliers.

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	63	GYC	CA1
1	B	63	GYC	CA1
1	C	63	GYC	CA1
1	D	63	GYC	CA1
1	E	63	GYC	CA1

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	1750	0	1693	85	0
1	B	1743	0	1686	81	0
1	C	1737	0	1681	77	0
1	D	1737	0	1681	78	0
1	E	1743	0	1686	82	0
1	F	1743	0	1686	71	0
1	G	1737	0	1681	61	0
1	H	1737	0	1682	80	0
2	A	113	0	0	9	0
2	B	140	0	0	10	0
2	C	157	0	0	10	0
2	D	139	0	0	9	0
2	E	148	0	0	8	0
2	F	145	0	0	6	0
2	G	128	0	0	9	0
2	H	111	0	0	6	0
All	All	15008	0	13476	584	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 21.

The worst 5 of 584 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:H:63:GYC:HA31	1:H:63:GYC:HN2	1.15	1.08

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:138:LYS:HG2	1:C:162:LEU:HD21	1.38	1.05
1:C:187:ASP:HB3	2:C:371:HOH:O	1.59	1.01
1:H:198:LYS:HA	1:H:198:LYS:HE2	1.47	0.97
1:D:198:LYS:HB2	1:D:208:ASN:HD22	1.30	0.94

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	214/255 (84%)	199 (93%)	13 (6%)	2 (1%)	17	20
1	B	213/255 (84%)	196 (92%)	16 (8%)	1 (0%)	29	35
1	C	212/255 (83%)	198 (93%)	14 (7%)	0	100	100
1	D	212/255 (83%)	202 (95%)	10 (5%)	0	100	100
1	E	213/255 (84%)	203 (95%)	9 (4%)	1 (0%)	29	35
1	F	213/255 (84%)	204 (96%)	8 (4%)	1 (0%)	29	35
1	G	212/255 (83%)	202 (95%)	7 (3%)	3 (1%)	11	11
1	H	212/255 (83%)	195 (92%)	13 (6%)	4 (2%)	8	7
All	All	1701/2040 (83%)	1599 (94%)	90 (5%)	12 (1%)	22	26

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	167	GLY
1	H	202	LYS
1	A	200	HIS
1	G	218	GLY
1	H	167	GLY

5.3.2 Protein sidechains [i](#)

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	186/217 (86%)	177 (95%)	9 (5%)	25	36
1	B	185/217 (85%)	173 (94%)	12 (6%)	17	23
1	C	184/217 (85%)	180 (98%)	4 (2%)	52	69
1	D	184/217 (85%)	175 (95%)	9 (5%)	25	35
1	E	185/217 (85%)	173 (94%)	12 (6%)	17	23
1	F	185/217 (85%)	179 (97%)	6 (3%)	39	54
1	G	184/217 (85%)	173 (94%)	11 (6%)	19	26
1	H	184/217 (85%)	179 (97%)	5 (3%)	44	61
All	All	1477/1736 (85%)	1409 (95%)	68 (5%)	27	38

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	100	ILE
1	E	94	ASN
1	G	211	GLU
1	D	145	LYS
1	E	58	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	205	ASN
1	E	94	ASN
1	H	158	ASN
1	D	216	HIS
1	E	158	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues are 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
1	GYC	D	63	1	22,22,23	3.69	10 (45%)	26,30,32	2.25	6 (23%)
1	GYC	G	63	1	22,22,23	3.84	11 (50%)	26,30,32	2.03	6 (23%)
1	GYC	A	63	1	22,22,23	3.79	11 (50%)	26,30,32	2.04	5 (19%)
1	GYC	C	63	1	22,22,23	3.74	11 (50%)	26,30,32	2.18	8 (30%)
1	GYC	F	63	1	22,22,23	3.81	13 (59%)	26,30,32	2.14	6 (23%)
1	GYC	H	63	1	22,22,23	3.81	10 (45%)	26,30,32	2.04	7 (26%)
1	GYC	B	63	1	22,22,23	3.66	12 (54%)	26,30,32	2.03	6 (23%)
1	GYC	E	63	1	22,22,23	3.83	10 (45%)	26,30,32	1.93	6 (23%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	GYC	D	63	1	1/1/5/7	3/9/29/30	0/2/2/2
1	GYC	G	63	1	1/1/5/7	1/9/29/30	0/2/2/2
1	GYC	A	63	1	1/1/5/7	2/9/29/30	0/2/2/2
1	GYC	C	63	1	1/1/5/7	2/9/29/30	0/2/2/2
1	GYC	F	63	1	1/1/5/7	2/9/29/30	0/2/2/2
1	GYC	H	63	1	1/1/5/7	2/9/29/30	0/2/2/2
1	GYC	B	63	1	1/1/5/7	2/9/29/30	0/2/2/2
1	GYC	E	63	1	1/1/5/7	3/9/29/30	0/2/2/2

The worst 5 of 88 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	63	GYC	CA1-C1	-12.40	1.26	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	63	GYC	CA1-C1	-11.82	1.28	1.51
1	G	63	GYC	CA1-C1	-11.72	1.28	1.51
1	D	63	GYC	CA1-C1	-11.67	1.28	1.51
1	A	63	GYC	CA1-C1	-11.36	1.28	1.51

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	GYC	CA3-N3-C1	7.01	135.58	127.16
1	D	63	GYC	CA3-N3-C1	6.83	135.36	127.16
1	G	63	GYC	CA3-N3-C1	6.63	135.12	127.16
1	C	63	GYC	CA3-N3-C1	6.60	135.08	127.16
1	F	63	GYC	CA3-N3-C1	6.46	134.91	127.16

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	63	GYC	CA1
1	G	63	GYC	CA1
1	A	63	GYC	CA1
1	C	63	GYC	CA1
1	F	63	GYC	CA1

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	63	GYC	C1-CA1-CB1-SG1
1	D	63	GYC	C-CA3-N3-C2
1	G	63	GYC	C1-CA1-CB1-SG1
1	A	63	GYC	C1-CA1-CB1-SG1
1	C	63	GYC	C1-CA1-CB1-SG1

There are no ring outliers.

8 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	63	GYC	2	0
1	G	63	GYC	2	0
1	A	63	GYC	2	0
1	C	63	GYC	2	0
1	F	63	GYC	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	H	63	GYC	7	0
1	B	63	GYC	3	0
1	E	63	GYC	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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

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.