



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 10:15 am BST

PDB ID : 2Z6Y  
Title : Crystal structure of a photoswitchable GFP-like protein Dronpa in the bright-state  
Authors : Kikuchi, A.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Mizuno, H.; Miyawaki, A.  
Deposited on : 2007-08-09  
Resolution : 2.00 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](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 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

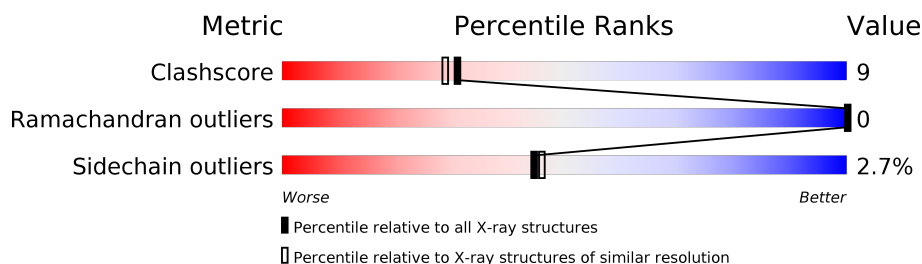
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	225	
1	B	225	
1	C	225	
1	D	225	

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	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	GYC	B	63	X	-	-	-
1	GYC	C	63	X	-	-	-
1	GYC	D	63	X	-	-	-

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7528 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 Fluorescent protein Dronpa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	0	0
			1736	1108	292	326	10			
1	B	215	Total	C	N	O	S	0	0	0
			1736	1108	292	326	10			
1	C	214	Total	C	N	O	S	0	0	0
			1730	1105	291	324	10			
1	D	214	Total	C	N	O	S	0	0	0
			1730	1105	291	324	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q5TLG6
A	-1	SER	-	EXPRESSION TAG	UNP Q5TLG6
A	0	HIS	-	EXPRESSION TAG	UNP Q5TLG6
B	-2	GLY	-	EXPRESSION TAG	UNP Q5TLG6
B	-1	SER	-	EXPRESSION TAG	UNP Q5TLG6
B	0	HIS	-	EXPRESSION TAG	UNP Q5TLG6
C	-2	GLY	-	EXPRESSION TAG	UNP Q5TLG6
C	-1	SER	-	EXPRESSION TAG	UNP Q5TLG6
C	0	HIS	-	EXPRESSION TAG	UNP Q5TLG6
D	-2	GLY	-	EXPRESSION TAG	UNP Q5TLG6
D	-1	SER	-	EXPRESSION TAG	UNP Q5TLG6
D	0	HIS	-	EXPRESSION TAG	UNP Q5TLG6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	151	Total	O	0	0
			151	151		
2	B	137	Total	O	0	0
			137	137		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	160	Total 160	O 160	0	0
2	D	148	Total 148	O 148	0	0

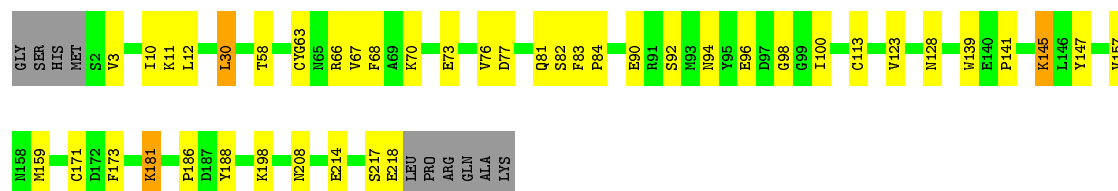
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

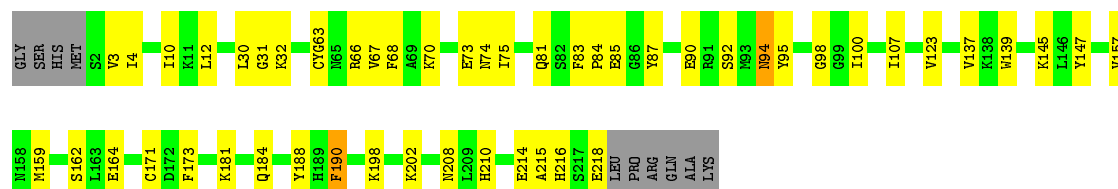
- Molecule 1: Fluorescent protein Dronpa

Chain A: 




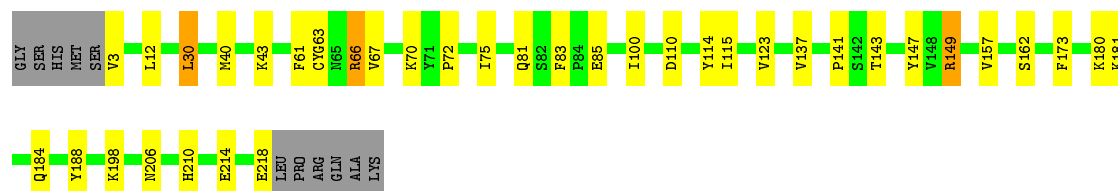
- Molecule 1: Fluorescent protein Dronpa

Chain B: 




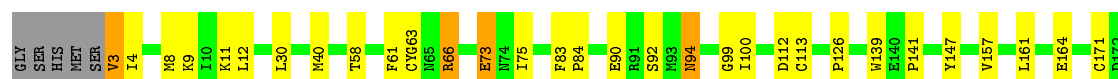
- Molecule 1: Fluorescent protein Dronpa

Chain C: 



- Molecule 1: Fluorescent protein Dronpa

Chain D: 





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.89 Å   95.83 Å   74.01 Å 90.00°   115.62°   90.00°	Depositor
Resolution (Å)	19.51 – 2.00	Depositor
% Data completeness (in resolution range)	95.2 (19.51-2.00)	Depositor
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.208 , 0.249	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7528	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP



## 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	0.36	0/1760	0.66	0/2375
1	B	0.36	0/1760	0.67	0/2375
1	C	0.38	0/1754	0.68	0/2367
1	D	0.38	0/1754	0.68	0/2367
All	All	0.37	0/7028	0.67	0/9484

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
All	All	4	0

There are no bond length outliers.

There are no bond angle outliers.

All (4) 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

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	1736	0	1660	36	0
1	B	1736	0	1660	38	0
1	C	1730	0	1655	31	0
1	D	1730	0	1655	29	0
2	A	151	0	0	4	0
2	B	137	0	0	4	0
2	C	160	0	0	3	0
2	D	148	0	0	4	0
All	All	7528	0	6630	123	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 9.

The worst 5 of 123 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:C:198:LYS:HG3	1:C:210:HIS:CD2	2.08	0.87
1:D:63:GYC:N	1:D:63:GYC:HA31	1.91	0.85
1:B:63:GYC:N	1:B:63:GYC:HA31	1.93	0.84
1:C:63:GYC:HA31	1:C:63:GYC:N	1.93	0.83
1:B:73:GLU:CD	1:B:73:GLU:H	1.81	0.82

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	212/225 (94%)	210 (99%)	2 (1%)	0	100	100
1	B	212/225 (94%)	211 (100%)	1 (0%)	0	100	100
1	C	211/225 (94%)	208 (99%)	3 (1%)	0	100	100
1	D	211/225 (94%)	209 (99%)	2 (1%)	0	100	100
All	All	846/900 (94%)	838 (99%)	8 (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	185/193 (96%)	180 (97%)	5 (3%)	44	46
1	B	185/193 (96%)	180 (97%)	5 (3%)	44	46
1	C	184/193 (95%)	181 (98%)	3 (2%)	62	67
1	D	184/193 (95%)	177 (96%)	7 (4%)	33	31
All	All	738/772 (96%)	718 (97%)	20 (3%)	44	46

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	190	PHE
1	C	30	LEU
1	D	73	GLU
1	B	145	LYS
1	B	164	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	208	ASN
1	C	81	GLN
1	D	21	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	102	ASN
1	B	206	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	A	63	1	22,22,23	3.87	11 (50%)	26,30,32	2.19	7 (26%)
1	GYC	B	63	1	22,22,23	3.86	11 (50%)	26,30,32	2.11	6 (23%)
1	GYC	C	63	1	22,22,23	3.73	13 (59%)	26,30,32	2.14	6 (23%)
1	GYC	D	63	1	22,22,23	3.72	11 (50%)	26,30,32	2.12	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. '-' means no outliers of that kind were identified.

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

The worst 5 of 46 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	63	GYC	CA1-C1	-12.26	1.27	1.51
1	A	63	GYC	CA1-C1	-12.01	1.27	1.51
1	B	63	GYC	CA1-C1	-11.85	1.27	1.51
1	C	63	GYC	CA1-C1	-11.43	1.28	1.51
1	B	63	GYC	CB2-CA2	7.29	1.41	1.35

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	63	GYC	CA3-N3-C1	7.29	135.91	127.16
1	A	63	GYC	CA3-N3-C1	7.06	135.64	127.16
1	B	63	GYC	CA3-N3-C1	7.02	135.58	127.16
1	C	63	GYC	CA3-N3-C1	6.83	135.37	127.16
1	C	63	GYC	C2-CA2-N2	4.93	112.39	108.93

All (4) 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

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	63	GYC	N-CA1-CB1-SG1
1	A	63	GYC	C1-CA1-CB1-SG1
1	B	63	GYC	C1-CA1-CB1-SG1
1	C	63	GYC	C1-CA1-CB1-SG1
1	D	63	GYC	C1-CA1-CB1-SG1

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	63	GYC	2	0
1	B	63	GYC	2	0
1	C	63	GYC	1	0
1	D	63	GYC	2	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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.