



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 06:15 pm BST

PDB ID : 1RMP  
Title : Probing the Role of Tryptophans in Aequorea Victoria Green Fluorescent Proteins with an Expanded Genetic Code  
Authors : Budisa, N.; Pal, P.P.; Alefelder, S.; Birle, P.; Krywcun, T.; Rubini, M.; Wenger, W.; Bae, J.H.; Steiner, T.  
Deposited on : 2003-11-28  
Resolution : 3.00 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 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

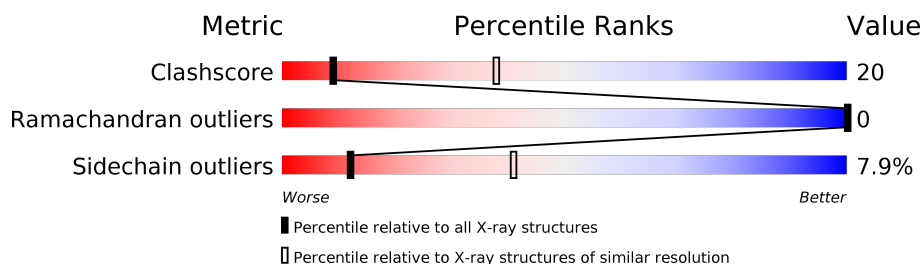
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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	226	

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	23S	A	57	-	-	X	-

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1787 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 SIGF1-GFP fusion protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	O	S	Se	0	0	0
			1775	1127	301	340	6	1			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	66	CRO	THR	CHROMOPHORE	UNP P42212
A	66	CRO	TYR	CHROMOPHORE	UNP P42212
A	66	CRO	GLY	CHROMOPHORE	UNP P42212

- Molecule 2 is water.

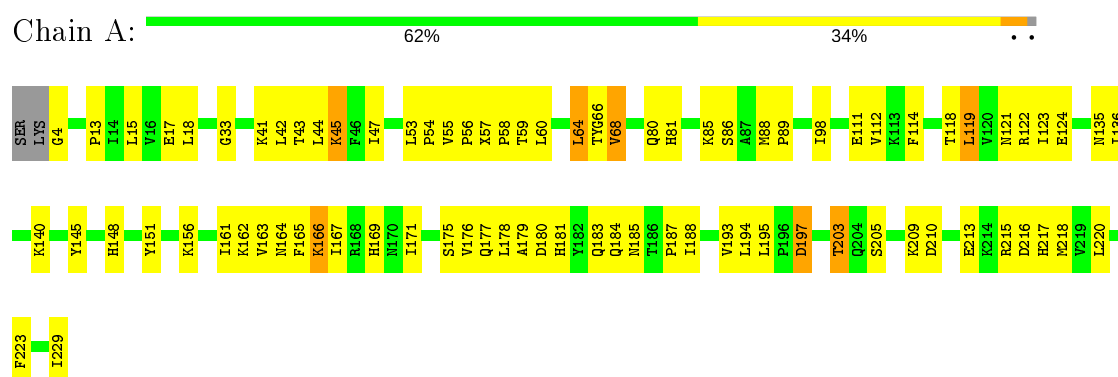
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	12	Total	O	0	0
			12	12		

### 3 Residue-property plots [i](#)

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: SIGF1-GFP fusion protein



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.77Å 62.77Å 70.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.67 – 3.00	Depositor
% Data completeness (in resolution range)	76.0 (28.67-3.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.225 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1787	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	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: CRO, 23S

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	0.51	1/1776 (0.1%)	0.87	1/2397 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	LEU	C-O	9.69	1.41	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	68	VAL	N-CA-CB	-6.96	96.18	111.50

There are no chirality outliers.

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	1775	0	1708	70	0
2	A	12	0	0	1	0
All	All	1787	0	1708	70	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 20.

All (70) 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:57:23S:HE1	1:A:218:MET:HB2	0.96	1.08
1:A:57:23S:NE1	1:A:218:MET:HB2	1.74	1.01
1:A:53:LEU:HD22	1:A:57:23S:HZ3	1.38	1.01
1:A:86:SER:HB2	1:A:194:LEU:HD13	1.68	0.75
1:A:156:LYS:HG2	1:A:195:LEU:HD13	1.70	0.74
1:A:111:GLU:HG3	1:A:188:ILE:HD11	1.72	0.70
1:A:53:LEU:CD2	1:A:57:23S:HZ3	2.20	0.68
1:A:53:LEU:HD13	1:A:57:23S:CZ2	2.24	0.68
1:A:47:ILE:HD13	1:A:215:ARG:CZ	2.24	0.67
1:A:57:23S:HE1	1:A:218:MET:CB	1.91	0.65
1:A:166:LYS:HB2	1:A:166:LYS:NZ	2.12	0.63
1:A:53:LEU:HD22	1:A:57:23S:CZ3	2.22	0.62
1:A:56:PRO:HB2	1:A:59:THR:HG23	1.81	0.60
1:A:56:PRO:HD3	1:A:136:ILE:O	2.01	0.60
1:A:176:VAL:HG12	1:A:177:GLN:N	2.17	0.60
1:A:13:PRO:HG2	1:A:118:THR:HA	1.85	0.58
1:A:171:ILE:HD12	1:A:175:SER:HB3	1.84	0.58
1:A:18:LEU:HD23	1:A:123:ILE:HB	1.86	0.57
1:A:47:ILE:HD13	1:A:215:ARG:NH1	2.20	0.56
1:A:56:PRO:HD2	1:A:136:ILE:HG23	1.87	0.55
1:A:53:LEU:HD23	1:A:55:VAL:O	2.05	0.55
1:A:156:LYS:CG	1:A:195:LEU:HD13	2.38	0.54
1:A:156:LYS:HA	1:A:195:LEU:CD1	2.39	0.53
1:A:121:ASN:ND2	1:A:123:ILE:HD11	2.24	0.53
1:A:169:HIS:HB2	1:A:177:GLN:HB3	1.91	0.53
1:A:33:GLY:HA3	1:A:44:LEU:HD13	1.90	0.53
1:A:156:LYS:HA	1:A:195:LEU:HD11	1.92	0.52
1:A:187:PRO:HG3	1:A:193:VAL:HG11	1.93	0.51
1:A:81:HIS:CE1	1:A:197:ASP:HB2	2.45	0.51
1:A:66:CRO:HG13	1:A:68:VAL:HG21	1.92	0.50
1:A:60:LEU:O	1:A:64:LEU:HG	2.10	0.49
1:A:209:LYS:HG3	1:A:217:HIS:NE2	2.26	0.49
1:A:53:LEU:HD21	1:A:60:LEU:HD11	1.93	0.49
1:A:15:LEU:HG	1:A:118:THR:HG21	1.93	0.49
1:A:162:LYS:CE	1:A:184:GLN:HE21	2.25	0.49
1:A:98:ILE:HG23	1:A:181:HIS:CE1	2.49	0.48
1:A:57:23S:N	1:A:58:PRO:CD	2.76	0.47
1:A:88:MET:HB3	1:A:89:PRO:HA	1.96	0.47
1:A:151:TYR:O	1:A:163:VAL:HG13	2.15	0.47
1:A:161:ILE:HG12	1:A:185:ASN:HB2	1.97	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:183:GLN:HG2	1:A:184:GLN:N	2.31	0.46
1:A:135:ASN:OD1	1:A:140:LYS:HD2	2.16	0.46
1:A:166:LYS:HD2	1:A:178:LEU:HD13	1.97	0.46
1:A:162:LYS:HE2	1:A:184:GLN:HE21	1.80	0.46
1:A:53:LEU:HG	1:A:54:PRO:HD2	1.98	0.46
1:A:156:LYS:HG2	1:A:195:LEU:CD1	2.44	0.45
1:A:209:LYS:HG3	1:A:217:HIS:CE1	2.52	0.45
1:A:56:PRO:HD2	1:A:136:ILE:CG2	2.46	0.45
1:A:41:LYS:HG3	1:A:223:PHE:CE1	2.51	0.45
1:A:86:SER:CB	1:A:194:LEU:HD13	2.44	0.45
1:A:68:VAL:HG23	2:A:236:HOH:O	2.16	0.44
1:A:53:LEU:HD13	1:A:57:23S:CZ3	2.48	0.43
1:A:17:GLU:CB	1:A:122:ARG:NH2	2.81	0.43
1:A:167:ILE:HB	1:A:179:ALA:HB3	2.00	0.43
1:A:166:LYS:HG3	1:A:178:LEU:HD22	1.99	0.43
1:A:148:HIS:O	1:A:203:THR:HG23	2.18	0.43
1:A:166:LYS:HB2	1:A:166:LYS:HZ2	1.80	0.43
1:A:4:GLY:HA3	1:A:85:LYS:O	2.18	0.43
1:A:176:VAL:CG1	1:A:177:GLN:N	2.82	0.42
1:A:210:ASP:HB3	1:A:213:GLU:HB2	2.01	0.42
1:A:161:ILE:CG1	1:A:185:ASN:HB2	2.50	0.42
1:A:123:ILE:HG22	1:A:124:GLU:N	2.34	0.41
1:A:165:PHE:CD1	1:A:165:PHE:N	2.87	0.41
1:A:112:VAL:HG13	1:A:119:LEU:HD11	2.02	0.41
1:A:45:LYS:HE3	1:A:47:ILE:HD11	2.03	0.41
1:A:98:ILE:HD13	1:A:181:HIS:NE2	2.36	0.40
1:A:33:GLY:HA3	1:A:43:THR:O	2.21	0.40
1:A:176:VAL:HG12	1:A:177:GLN:H	1.86	0.40
1:A:148:HIS:CG	1:A:167:ILE:HD13	2.56	0.40
1:A:178:LEU:HD23	1:A:178:LEU:HA	1.78	0.40

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	218/226 (96%)	202 (93%)	16 (7%)	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	190/196 (97%)	175 (92%)	15 (8%)	12	41

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	LEU
1	A	45	LYS
1	A	80	GLN
1	A	114	PHE
1	A	119	LEU
1	A	145	TYR
1	A	164	ASN
1	A	166	LYS
1	A	180	ASP
1	A	197	ASP
1	A	203	THR
1	A	205	SER
1	A	216	ASP
1	A	220	LEU
1	A	229	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	80	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	144	ASN
1	A	159	ASN
1	A	184	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	CRO	A	66	1	23,23,24	2.35	10 (43%)	30,32,34	3.29	9 (30%)
1	23S	A	57	1	7,14,15	1.32	1 (14%)	6,19,21	2.82	4 (66%)

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	CRO	A	66	1	-	0/12/31/32	0/2/2/2
1	23S	A	57	1	-	0/4/6/8	0/2/2/2

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	CB2-CA2	4.98	1.39	1.35
1	A	66	CRO	OH-CZ	-4.36	1.26	1.37
1	A	66	CRO	CG2-CB2	-3.47	1.40	1.46
1	A	66	CRO	CA3-C3	3.39	1.60	1.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	CA2-C2	-3.37	1.45	1.48
1	A	66	CRO	CA2-N2	-3.16	1.31	1.38
1	A	66	CRO	CE2-CD2	2.91	1.44	1.38
1	A	57	23S	CD2-CE2	-2.81	1.35	1.42
1	A	66	CRO	O3-C3	2.50	1.34	1.19
1	A	66	CRO	CD1-CG2	2.37	1.44	1.39
1	A	66	CRO	CA1-N1	2.04	1.53	1.47

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	CRO	CA2-N2-C1	10.65	113.62	105.77
1	A	66	CRO	O2-C2-CA2	-8.07	126.43	130.96
1	A	66	CRO	C2-CA2-N2	-6.68	104.25	108.93
1	A	66	CRO	CA2-C2-N3	6.58	106.48	103.37
1	A	57	23S	CG-CB-CA	5.12	122.45	114.53
1	A	66	CRO	C2-N3-C1	-3.17	106.36	107.97
1	A	66	CRO	N3-C1-N2	-3.14	109.28	111.45
1	A	57	23S	CZ3-CD2-CE2	2.86	113.43	105.30
1	A	57	23S	CB-CA-C	2.78	116.67	111.47
1	A	66	CRO	CB1-CA1-N1	2.60	131.06	113.84
1	A	66	CRO	C1-CA1-N1	-2.53	105.85	109.96
1	A	66	CRO	CB2-CA2-C2	2.45	125.20	122.28
1	A	57	23S	CZ2-CZ3-CD2	2.04	116.07	110.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	66	CRO	1	0
1	A	57	23S	9	0

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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.