



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 04:15 pm GMT

PDB ID : 4HVF  
Title : Crystal structure of green fluorescent protein lanGFP(Branchiostoma Lanceolatum)  
Authors : Pletnev, V.Z.; Pletneva, N.V.; Pletnev, S.V.  
Deposited on : 2012-11-06  
Resolution : 1.70 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

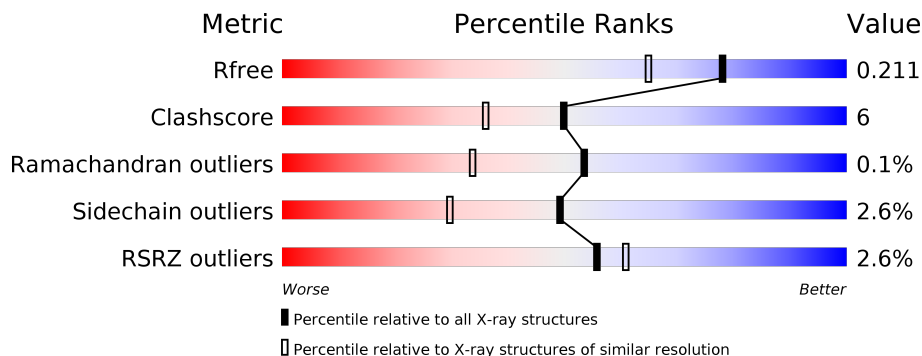
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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))
$R_{free}$	111664	3793 (1.70-1.70)
Clashscore	122126	4167 (1.70-1.70)
Ramachandran outliers	120053	4100 (1.70-1.70)
Sidechain outliers	120020	4100 (1.70-1.70)
RSRZ outliers	108989	3718 (1.70-1.70)

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. 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	226	 2% (red), 81% (green), 15% (yellow), 2% (grey)
1	B	226	 4% (red), 81% (green), 15% (yellow), 2% (grey)
1	C	226	 2% (red), 85% (green), 11% (yellow), 2% (grey)
1	D	226	 3% (red), 85% (green), 12% (yellow), 2% (grey)

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7969 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 blFP-Y6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	218	Total 1845	C 1189	N 306	O 343	S 7	0	11	0
1	B	218	Total 1811	C 1167	N 297	O 340	S 7	0	7	0
1	C	218	Total 1804	C 1164	N 295	O 338	S 7	0	6	0
1	D	225	Total 1885	C 1211	N 319	O 348	S 7	0	9	0

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	EXPRESSION TAG	UNP B1PNB8
A	-6	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	-5	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	-4	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	-3	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	-2	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	-1	HIS	-	EXPRESSION TAG	UNP B1PNB8
A	0	GLY	-	EXPRESSION TAG	UNP B1PNB8
A	1	SER	-	EXPRESSION TAG	UNP B1PNB8
A	2	LEU	PRO	ENGINEERED MUTATION	UNP B1PNB8
A	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
A	58	CR2	TYR	CHROMOPHORE	UNP B1PNB8
A	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
B	-7	MET	-	EXPRESSION TAG	UNP B1PNB8
B	-6	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	-5	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	-4	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	-3	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	-2	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	-1	HIS	-	EXPRESSION TAG	UNP B1PNB8
B	0	GLY	-	EXPRESSION TAG	UNP B1PNB8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	SER	-	EXPRESSION TAG	UNP B1PNB8
B	2	LEU	PRO	ENGINEERED MUTATION	UNP B1PNB8
B	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
B	58	CR2	TYR	CHROMOPHORE	UNP B1PNB8
B	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
C	-7	MET	-	EXPRESSION TAG	UNP B1PNB8
C	-6	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	-5	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	-4	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	-3	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	-2	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	-1	HIS	-	EXPRESSION TAG	UNP B1PNB8
C	0	GLY	-	EXPRESSION TAG	UNP B1PNB8
C	1	SER	-	EXPRESSION TAG	UNP B1PNB8
C	2	LEU	PRO	ENGINEERED MUTATION	UNP B1PNB8
C	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
C	58	CR2	TYR	CHROMOPHORE	UNP B1PNB8
C	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
D	-7	MET	-	EXPRESSION TAG	UNP B1PNB8
D	-6	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	-5	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	-4	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	-3	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	-2	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	-1	HIS	-	EXPRESSION TAG	UNP B1PNB8
D	0	GLY	-	EXPRESSION TAG	UNP B1PNB8
D	1	SER	-	EXPRESSION TAG	UNP B1PNB8
D	2	LEU	PRO	ENGINEERED MUTATION	UNP B1PNB8
D	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8
D	58	CR2	TYR	CHROMOPHORE	UNP B1PNB8
D	58	CR2	GLY	CHROMOPHORE	UNP B1PNB8

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O 6 3 3	0	0
2	C	1	Total C O 6 3 3	0	0

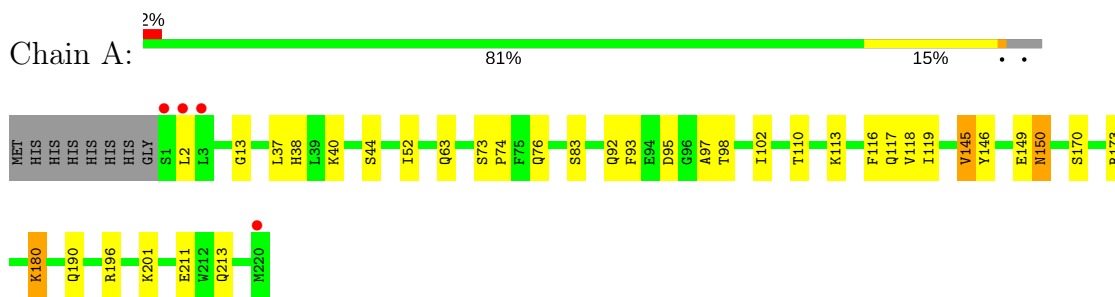
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	158	Total O 158 158	0	0
3	B	164	Total O 164 164	0	0
3	C	148	Total O 148 148	0	0
3	D	142	Total O 142 142	0	0

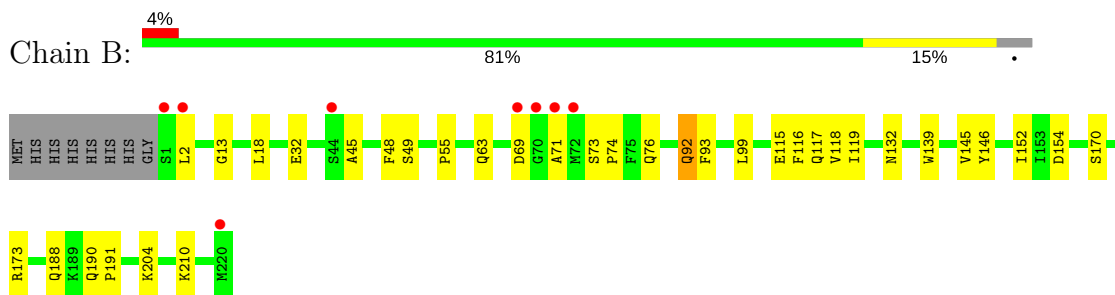
### 3 Residue-property plots [i](#)

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 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.

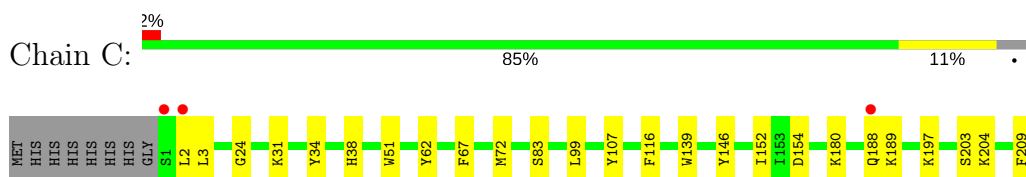
- Molecule 1: Green fluorescent protein bFP-Y6



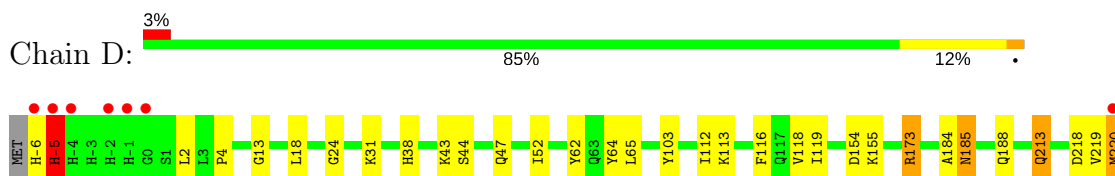
- Molecule 1: Green fluorescent protein bFP-Y6



- Molecule 1: Green fluorescent protein bFP-Y6



- Molecule 1: Green fluorescent protein bFP-Y6



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.19Å 105.80Å 121.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.56 – 1.70 26.56 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.4 (26.56-1.70) 99.3 (26.56-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 1.71Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.177 , 0.213 0.176 , 0.211	Depositor DCC
$R_{free}$ test set	1013 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.2	Xtrriage
Anisotropy	0.013	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 42.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7969	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: CR2, GOL

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	1.40	5/1886 (0.3%)	1.05	4/2556 (0.2%)
1	B	1.43	7/1845 (0.4%)	1.07	2/2501 (0.1%)
1	C	1.42	5/1838 (0.3%)	1.09	3/2492 (0.1%)
1	D	1.39	5/1937 (0.3%)	1.05	2/2627 (0.1%)
All	All	1.41	22/7506 (0.3%)	1.06	11/10176 (0.1%)

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	62	TYR	CD1-CE1	5.91	1.48	1.39
1	D	62	TYR	CD2-CE2	-5.82	1.30	1.39
1	C	34	TYR	CG-CD1	-5.75	1.31	1.39
1	B	32	GLU	CD-OE1	-5.67	1.19	1.25
1	D	62	TYR	CE2-CZ	-5.59	1.31	1.38
1	B	210	LYS	CE-NZ	5.53	1.62	1.49
1	C	209	PHE	CD2-CE2	5.51	1.50	1.39
1	B	115	GLU	CD-OE1	-5.49	1.19	1.25
1	A	145	VAL	CB-CG2	-5.48	1.41	1.52
1	A	146	TYR	CD1-CE1	-5.41	1.31	1.39
1	B	188	GLN	CB-CG	-5.35	1.38	1.52
1	B	139	TRP	CE3-CZ3	5.31	1.47	1.38
1	A	211	GLU	CB-CG	5.26	1.62	1.52
1	D	103	TYR	CD2-CE2	5.24	1.47	1.39
1	B	145	VAL	CB-CG1	5.17	1.63	1.52
1	A	97	ALA	CA-CB	5.16	1.63	1.52
1	C	51	TRP	CE3-CZ3	-5.08	1.29	1.38
1	B	69	ASP	CB-CG	5.05	1.62	1.51
1	D	64	TYR	CG-CD1	5.05	1.45	1.39
1	D	116	PHE	CE2-CZ	5.04	1.47	1.37
1	A	146	TYR	CD2-CE2	-5.01	1.31	1.39
1	C	107	TYR	CD2-CE2	5.00	1.46	1.39



All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	69	ASP	CB-CG-OD1	5.87	123.58	118.30
1	C	154	ASP	CB-CG-OD1	-5.81	113.07	118.30
1	D	154	ASP	CB-CG-OD1	-5.57	113.29	118.30
1	A	116	PHE	CB-CG-CD2	-5.40	117.02	120.80
1	A	116	PHE	CB-CG-CD1	5.37	124.56	120.80
1	D	113	LYS	CD-CE-NZ	-5.36	99.37	111.70
1	B	116	PHE	CB-CG-CD2	-5.24	117.13	120.80
1	A	37	LEU	CB-CG-CD2	-5.16	102.22	111.00
1	C	116	PHE	CB-CG-CD2	-5.11	117.22	120.80
1	A	95	ASP	CB-CG-OD2	-5.01	113.79	118.30
1	C	217	SER	CB-CA-C	-5.01	100.59	110.10

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	1845	0	1785	34	0
1	B	1811	0	1741	19	0
1	C	1804	0	1738	16	0
1	D	1885	0	1804	25	0
2	C	12	0	15	1	0
3	A	158	0	0	1	0
3	B	164	0	0	2	0
3	C	148	0	0	1	0
3	D	142	0	0	1	0
All	All	7969	0	7083	91	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 6.

All (91) 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:204[B]:LYS:CD	1:C:204[B]:LYS:H	1.50	1.23
1:D:-6:HIS:O	1:D:-5:HIS:HB2	1.45	1.17
1:A:201[A]:LYS:HE2	3:A:411:HOH:O	1.63	0.98
1:A:2:LEU:CD1	1:A:110:THR:HG22	1.95	0.97
1:D:18:LEU:HD13	1:D:47[B]:GLN:HG3	1.45	0.96
1:A:2:LEU:HD11	1:A:110:THR:CG2	1.94	0.96
1:B:71:ALA:HB1	3:B:396:HOH:O	1.69	0.91
1:D:173[A]:ARG:HG3	1:D:173[A]:ARG:HH11	1.42	0.84
1:D:173[A]:ARG:CG	1:D:173[A]:ARG:HH11	1.91	0.83
1:A:2:LEU:CD1	1:A:110:THR:CG2	2.55	0.82
1:A:2:LEU:HD11	1:A:110:THR:HG22	1.59	0.81
1:A:2:LEU:HD11	1:A:110:THR:HG23	1.61	0.80
1:A:2:LEU:HD12	1:A:110:THR:HG22	1.62	0.79
1:D:-6:HIS:O	1:D:-5:HIS:CB	2.30	0.79
1:A:83:SER:HB3	1:A:180[B]:LYS:HD2	1.67	0.76
1:A:117:GLN:HE22	1:B:117:GLN:HE22	1.33	0.75
1:C:204[B]:LYS:H	1:C:204[B]:LYS:HD2	0.64	0.74
1:A:102[A]:ILE:HD12	1:B:119:ILE:HD13	1.71	0.70
1:D:31:LYS:HE2	1:D:31:LYS:HA	1.73	0.69
1:A:83:SER:CB	1:A:180[B]:LYS:HD2	2.25	0.66
1:C:3:LEU:HD21	1:C:72:MET:HE3	1.77	0.65
1:A:38[A]:HIS:HE1	1:A:40:LYS:HD3	1.62	0.64
1:A:92:GLN:NE2	1:A:98:THR:OG1	2.30	0.63
1:C:203:SER:HB2	1:C:204[B]:LYS:HD2	1.84	0.59
1:A:38[A]:HIS:CE1	1:A:40:LYS:CG	2.84	0.59
1:A:38[A]:HIS:CE1	1:A:40:LYS:HD3	2.37	0.59
1:D:155:LYS:HE2	1:D:173[B]:ARG:NH1	2.19	0.58
1:D:173[A]:ARG:HG3	1:D:173[A]:ARG:NH1	2.17	0.57
1:D:185:ASN:HA	1:D:188[B]:GLN:HG2	1.87	0.56
1:D:173[A]:ARG:HH11	1:D:173[A]:ARG:CB	2.17	0.56
1:C:83:SER:HB2	1:C:180:LYS:HG3	1.89	0.54
1:A:40:LYS:HB3	1:A:40:LYS:NZ	2.23	0.54
1:B:49:SER:HA	1:B:132[B]:ASN:HD22	1.73	0.54
1:A:2:LEU:O	1:A:2:LEU:HG	2.08	0.52
1:D:173[A]:ARG:NH1	1:D:173[A]:ARG:HB2	2.25	0.52
1:A:40:LYS:CB	1:A:40:LYS:NZ	2.73	0.51
1:A:40:LYS:HB3	1:A:40:LYS:HZ3	1.76	0.51
1:A:149:GLU:HG3	1:A:150[B]:ASN:HD22	1.76	0.51
1:C:99[B]:LEU:HD23	1:C:99[B]:LEU:C	2.31	0.51
1:C:203:SER:CB	1:C:204[B]:LYS:HD2	2.43	0.49
1:D:173[A]:ARG:HB2	1:D:173[A]:ARG:HH11	1.78	0.49
1:B:92:GLN:HA	1:B:92:GLN:HE21	1.77	0.49

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:218:ASP:OD1	1:D:220:MET:HG3	2.12	0.49
1:B:2:LEU:HD23	1:B:2:LEU:O	2.13	0.48
1:A:145:VAL:HA	1:A:190[B]:GLN:HE22	1.79	0.48
1:B:13:GLY:HA3	1:B:118:VAL:O	2.14	0.48
1:D:184:ALA:O	1:D:188[A]:GLN:HG2	2.13	0.48
1:A:119[B]:ILE:HG23	1:A:119[B]:ILE:O	2.14	0.47
1:A:196:ARG:HG2	1:A:213:GLN:HB3	1.95	0.47
1:B:48:PHE:C	1:B:132[B]:ASN:HD21	2.17	0.47
1:D:18:LEU:CD1	1:D:47[B]:GLN:HG3	2.29	0.47
1:C:3:LEU:HD21	1:C:72:MET:CE	2.43	0.47
1:B:173:ARG:HD3	3:B:435:HOH:O	2.15	0.47
1:C:203:SER:HB2	1:C:204[B]:LYS:CD	2.45	0.46
1:C:83:SER:CB	1:C:180:LYS:HG3	2.44	0.46
1:D:24:GLY:HA3	1:D:38:HIS:O	2.15	0.46
1:B:93:PHE:CD2	1:B:170:SER:HB3	2.50	0.46
1:A:38[A]:HIS:CE1	1:A:40:LYS:HG3	2.50	0.46
1:A:13:GLY:HA3	1:A:118:VAL:O	2.15	0.45
1:C:146:TYR:HD1	1:C:152:ILE:HG12	1.83	0.44
1:D:43:LYS:O	1:D:44:SER:OG	2.30	0.44
1:D:185:ASN:O	1:D:188[B]:GLN:HG3	2.18	0.43
1:B:45:ALA:HB1	1:B:204:LYS:HB3	1.99	0.43
1:D:13:GLY:HA3	1:D:118:VAL:O	2.18	0.43
1:B:146:TYR:CD1	1:B:152:ILE:HG12	2.54	0.43
1:B:63:GLN:O	1:B:76:GLN:HG2	2.19	0.43
1:D:4:PRO:HB3	1:D:112:ILE:HD11	2.01	0.43
1:A:38[A]:HIS:CE1	1:A:40:LYS:CD	3.02	0.42
1:A:173[B]:ARG:HH21	1:B:92:GLN:NE2	2.17	0.42
1:A:38[A]:HIS:HE1	1:A:40:LYS:CD	2.31	0.42
1:A:63:GLN:O	1:A:76:GLN:HG2	2.20	0.42
1:D:119[C]:ILE:O	1:D:119[C]:ILE:HG23	2.17	0.42
1:D:219:VAL:O	3:D:411:HOH:O	2.22	0.42
1:B:146:TYR:HD1	1:B:152:ILE:HG12	1.85	0.42
1:B:190:GLN:HA	1:B:191:PRO:C	2.39	0.42
1:B:73:SER:HB2	1:B:74:PRO:HD2	2.01	0.42
1:C:24:GLY:HA3	1:C:38:HIS:O	2.20	0.42
1:D:185:ASN:O	1:D:188[B]:GLN:CG	2.68	0.42
1:D:52:ILE:HD12	1:D:52:ILE:HA	1.91	0.42
1:A:102[A]:ILE:CD1	1:B:119:ILE:HG21	2.50	0.42
1:A:73:SER:HB2	1:A:74:PRO:HD2	2.01	0.41
1:C:203:SER:HB2	1:C:204[B]:LYS:CE	2.50	0.41
1:B:99:LEU:HD11	1:B:118:VAL:CG1	2.51	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:139:TRP:CD1	1:C:197:LYS:HG2	2.56	0.40
1:D:65:LEU:HD13	1:D:213:GLN:HB2	2.03	0.40
1:A:93:PHE:CD2	1:A:170:SER:HB3	2.56	0.40
2:C:302:GOL:H11	3:C:512:HOH:O	2.21	0.40
1:A:149:GLU:HG3	1:A:150[B]:ASN:ND2	2.36	0.40
1:A:52:ILE:HD12	1:A:52:ILE:HA	1.93	0.40

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	224/226 (99%)	222 (99%)	2 (1%)	0	100	100
1	B	220/226 (97%)	217 (99%)	3 (1%)	0	100	100
1	C	219/226 (97%)	215 (98%)	4 (2%)	0	100	100
1	D	229/226 (101%)	224 (98%)	4 (2%)	1 (0%)	36	19
All	All	892/904 (99%)	878 (98%)	13 (2%)	1 (0%)	53	34

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	-5	HIS

### 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	199/195 (102%)	193 (97%)	6 (3%)	44	24
1	B	195/195 (100%)	191 (98%)	4 (2%)	56	38
1	C	194/195 (100%)	188 (97%)	6 (3%)	43	23
1	D	203/195 (104%)	196 (97%)	7 (3%)	40	19
All	All	791/780 (101%)	768 (97%)	23 (3%)	49	25

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	SER
1	A	113	LYS
1	A	150[A]	ASN
1	A	150[B]	ASN
1	A	180[A]	LYS
1	A	180[B]	LYS
1	B	18	LEU
1	B	55	PRO
1	B	92	GLN
1	B	154	ASP
1	C	2	LEU
1	C	31	LYS
1	C	67	PHE
1	C	188	GLN
1	C	189	LYS
1	C	217	SER
1	D	-5	HIS
1	D	2	LEU
1	D	173[A]	ARG
1	D	173[B]	ARG
1	D	185	ASN
1	D	213	GLN
1	D	220	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	92	GLN
1	A	117	GLN
1	B	92	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	56	GLN
1	D	92	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	CR2	A	58	1	20,20,21	2.06	6 (30%)	25,27,29	2.81	12 (48%)
1	CR2	B	58	1	20,20,21	1.80	6 (30%)	25,27,29	3.42	13 (52%)
1	CR2	C	58	1	20,20,21	1.94	8 (40%)	25,27,29	3.01	12 (48%)
1	CR2	D	58	1	20,20,21	1.35	2 (10%)	25,27,29	2.86	8 (32%)

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	CR2	A	58	1	-	0/6/25/26	0/2/2/2
1	CR2	B	58	1	-	0/6/25/26	0/2/2/2
1	CR2	C	58	1	-	0/6/25/26	0/2/2/2
1	CR2	D	58	1	-	0/6/25/26	0/2/2/2

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	58	CR2	OH-CZ	-3.93	1.27	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	58	CR2	CA2-C2	-3.65	1.45	1.48
1	C	58	CR2	OH-CZ	-3.19	1.29	1.37
1	B	58	CR2	CA2-C2	-3.07	1.45	1.48
1	C	58	CR2	CA2-C2	-3.00	1.45	1.48
1	A	58	CR2	C1-N2	-2.50	1.28	1.32
1	B	58	CR2	C2-N3	-2.40	1.34	1.39
1	C	58	CR2	C1-N2	-2.35	1.28	1.32
1	B	58	CR2	OH-CZ	-2.22	1.31	1.37
1	C	58	CR2	O2-C2	-2.21	1.18	1.23
1	B	58	CR2	O2-C2	-2.20	1.18	1.23
1	D	58	CR2	O2-C2	-2.19	1.18	1.23
1	A	58	CR2	O2-C2	-2.18	1.18	1.23
1	C	58	CR2	CE2-CZ	2.19	1.43	1.38
1	C	58	CR2	CD2-CG2	2.27	1.43	1.39
1	A	58	CR2	CA1-C1	2.30	1.52	1.49
1	C	58	CR2	CB2-CA2	2.88	1.37	1.35
1	B	58	CR2	CA1-C1	2.96	1.52	1.49
1	D	58	CR2	CA1-C1	3.07	1.52	1.49
1	B	58	CR2	CE1-CD1	3.64	1.45	1.38
1	C	58	CR2	CA1-C1	4.11	1.54	1.49
1	A	58	CR2	CB2-CA2	5.19	1.39	1.35

All (45) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	58	CR2	CA2-C2-N3	-5.68	100.84	103.30
1	B	58	CR2	CE1-CD1-CG2	-5.61	113.89	121.27
1	B	58	CR2	CA3-N3-C1	-4.84	120.98	127.85
1	A	58	CR2	CE1-CD1-CG2	-4.83	114.92	121.27
1	D	58	CR2	CE2-CD2-CG2	-4.73	115.05	121.27
1	A	58	CR2	CD2-CE2-CZ	-4.13	115.29	119.88
1	C	58	CR2	CD2-CG2-CB2	-3.09	110.64	121.23
1	C	58	CR2	CB2-CA2-N2	-3.00	124.50	128.82
1	B	58	CR2	O3-C3-CA3	-2.97	116.55	126.38
1	B	58	CR2	CD2-CE2-CZ	-2.95	116.60	119.88
1	D	58	CR2	CA3-N3-C1	-2.86	123.80	127.85
1	D	58	CR2	O3-C3-CA3	-2.53	118.02	126.38
1	D	58	CR2	CD2-CG2-CB2	-2.53	112.58	121.23
1	B	58	CR2	CA2-N2-C1	-2.49	103.90	105.76
1	A	58	CR2	O3-C3-CA3	-2.45	118.30	126.38
1	A	58	CR2	CA2-C2-N3	-2.43	102.25	103.30
1	C	58	CR2	CE2-CD2-CG2	-2.40	118.12	121.27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	58	CR2	CA3-N3-C1	-2.39	124.45	127.85
1	C	58	CR2	O3-C3-CA3	-2.33	118.69	126.38
1	B	58	CR2	CB2-CA2-N2	-2.27	125.55	128.82
1	A	58	CR2	CB2-CA2-C2	-2.15	119.65	122.28
1	B	58	CR2	O2-C2-N3	-2.03	120.15	124.38
1	A	58	CR2	CD2-CG2-CB2	-2.01	114.34	121.23
1	C	58	CR2	N3-C1-N2	2.14	114.12	111.81
1	B	58	CR2	N3-C1-N2	2.14	114.12	111.81
1	D	58	CR2	N3-C1-N2	2.14	114.12	111.81
1	A	58	CR2	N3-C1-N2	2.16	114.15	111.81
1	A	58	CR2	C1-CA1-N1	2.34	117.97	112.96
1	A	58	CR2	CD1-CE1-CZ	2.37	122.52	119.88
1	C	58	CR2	CG2-CB2-CA2	2.42	132.86	130.10
1	B	58	CR2	CB2-CA2-C2	2.72	125.61	122.28
1	C	58	CR2	CD1-CG2-CD2	2.99	122.12	117.63
1	A	58	CR2	C2-CA2-N2	3.44	111.36	108.93
1	A	58	CR2	CD1-CG2-CD2	3.49	122.88	117.63
1	B	58	CR2	CD1-CG2-CD2	3.80	123.35	117.63
1	D	58	CR2	CD1-CG2-CD2	4.62	124.58	117.63
1	C	58	CR2	C2-CA2-N2	4.68	112.23	108.93
1	B	58	CR2	CG2-CB2-CA2	4.98	135.79	130.10
1	C	58	CR2	O2-C2-CA2	6.57	134.46	130.97
1	D	58	CR2	O2-C2-CA2	7.22	134.81	130.97
1	C	58	CR2	C1-CA1-N1	7.84	129.71	112.96
1	D	58	CR2	C1-CA1-N1	7.85	129.72	112.96
1	B	58	CR2	C1-CA1-N1	7.85	129.72	112.96
1	B	58	CR2	O2-C2-CA2	8.86	135.68	130.97
1	A	58	CR2	O2-C2-CA2	9.23	135.88	130.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.



## 5.6 Ligand geometry [i](#)

2 ligands 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
2	GOL	C	301	-	5,5,5	0.71	0	5,5,5	1.06	0
2	GOL	C	302	-	5,5,5	1.06	0	5,5,5	1.06	0

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	GOL	C	301	-	-	0/4/4/4	0/0/0/0
2	GOL	C	302	-	-	0/4/4/4	0/0/0/0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	302	GOL	1	0

## 5.7 Other polymers [i](#)

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

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	217/226 (96%)	-0.11	4 (1%) 68 73	9, 15, 27, 47	0
1	B	217/226 (96%)	-0.02	8 (3%) 41 47	9, 16, 29, 51	0
1	C	217/226 (96%)	0.05	4 (1%) 68 73	9, 16, 30, 50	0
1	D	224/226 (99%)	-0.03	7 (3%) 49 54	10, 17, 29, 40	4 (1%)
All	All	875/904 (96%)	-0.03	23 (2%) 56 61	9, 16, 29, 51	4 (0%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	LEU	9.3
1	B	1	SER	6.6
1	C	1	SER	6.5
1	D	-6	HIS	5.9
1	D	220	MET	5.8
1	C	2	LEU	5.6
1	A	220	MET	4.2
1	D	-4	HIS	4.1
1	B	69	ASP	3.9
1	B	70	GLY	3.8
1	A	1	SER	3.7
1	B	220	MET	3.3
1	D	-5	HIS	3.3
1	B	2	LEU	3.0
1	B	72	MET	3.0
1	D	-1	HIS	2.9
1	A	3	LEU	2.8
1	C	220	MET	2.4
1	D	0	GLY	2.3
1	C	188	GLN	2.3
1	D	-2	HIS	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	44	SER	2.2
1	B	71	ALA	2.1

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	CR2	C	58	19/20	0.93	0.14	10,14,17,18	0
1	CR2	D	58	19/20	0.94	0.10	12,14,17,20	0
1	CR2	B	58	19/20	0.95	0.11	10,13,16,16	0
1	CR2	A	58	19/20	0.96	0.10	7,12,15,16	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GOL	C	302	6/6	0.77	0.33	20,20,20,20	6
2	GOL	C	301	6/6	0.80	0.33	20,20,20,20	6

## 6.5 Other polymers [i](#)

There are no such residues in this entry.