



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 10:49 PM GMT

PDB ID : 5DPI  
Title : sfGFP double mutant - 133/149 p-cyano-L-phenylalanine  
Authors : Dippel, A.B.; Olenginski, G.M.; Maurici, N.; Liskov, M.T.; Brewer, S.H.;  
Phillips-Piro, C.M.  
Deposited on : 2015-09-12  
Resolution : 2.54 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  
<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20026982

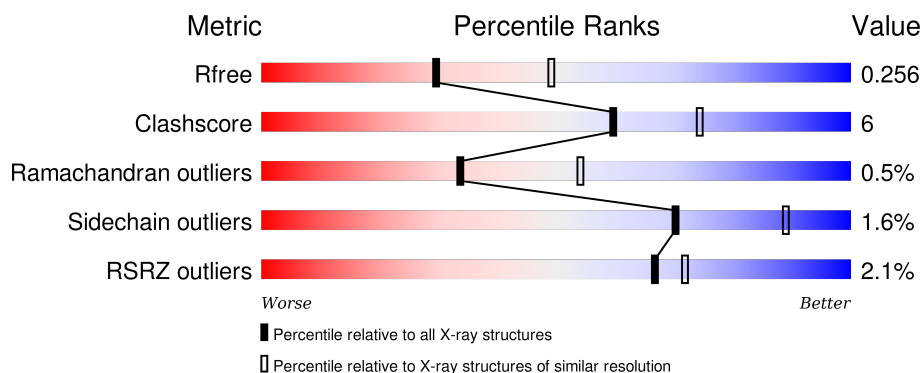
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.54 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))
$R_{free}$	91344	4549 (2.58-2.50)
Clashscore	102246	5292 (2.58-2.50)
Ramachandran outliers	100387	5194 (2.58-2.50)
Sidechain outliers	100360	5196 (2.58-2.50)
RSRZ outliers	91569	4561 (2.58-2.50)

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	237	<div> <div>86%</div> <div>9% 5%</div> </div>
1	B	237	<div> <div>2%</div> <div>82%</div> <div>12% 5%</div> </div>
1	C	237	<div> <div>%</div> <div>83%</div> <div>10% 5%</div> </div>
1	D	237	<div> <div>%</div> <div>80%</div> <div>15% 5%</div> </div>
1	E	237	<div> <div>3%</div> <div>79%</div> <div>16% 5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	237	<div><div></div><div>4%</div><div>72%</div><div>21%</div><div>• 5%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11017 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	226	Total	C	N	O	S	0	1	0
			1818	1158	311	344	5			
1	B	226	Total	C	N	O	S	0	2	0
			1823	1162	313	343	5			
1	C	224	Total	C	N	O	S	0	0	0
			1794	1142	307	340	5			
1	D	226	Total	C	N	O	S	0	0	0
			1807	1149	310	343	5			
1	E	226	Total	C	N	O	S	0	0	0
			1807	1149	310	343	5			
1	F	225	Total	C	N	O	S	0	1	0
			1800	1147	306	342	5			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP A0A059PIQ0
A	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
A	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
A	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
A	66	CRO	THR	chromophore	UNP A0A059PIQ0
A	66	CRO	TYR	chromophore	UNP A0A059PIQ0
A	66	CRO	GLY	chromophore	UNP A0A059PIQ0
A	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
A	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
A	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
A	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0
A	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0
B	0	MET	-	initiating methionine	UNP A0A059PIQ0
B	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
B	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
B	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
B	?	CRO	THR	chromophore	UNP A0A059PIQ0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	CRO	TYR	chromophore	UNP A0A059PIQ0
B	66	CRO	GLY	chromophore	UNP A0A059PIQ0
B	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
B	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
B	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
B	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0
B	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0
C	0	MET	-	initiating methionine	UNP A0A059PIQ0
C	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
C	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
C	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
C	66	CRO	THR	chromophore	UNP A0A059PIQ0
C	66	CRO	TYR	chromophore	UNP A0A059PIQ0
C	66	CRO	GLY	chromophore	UNP A0A059PIQ0
C	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
C	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
C	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
C	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0
C	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0
D	0	MET	-	initiating methionine	UNP A0A059PIQ0
D	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
D	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
D	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
D	?	CRO	THR	chromophore	UNP A0A059PIQ0
D	?	CRO	TYR	chromophore	UNP A0A059PIQ0
D	66	CRO	GLY	chromophore	UNP A0A059PIQ0
D	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
D	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
D	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
D	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0
D	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0
E	0	MET	-	initiating methionine	UNP A0A059PIQ0
E	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
E	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
E	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
E	66	CRO	THR	chromophore	UNP A0A059PIQ0
E	66	CRO	TYR	chromophore	UNP A0A059PIQ0
E	66	CRO	GLY	chromophore	UNP A0A059PIQ0
E	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
E	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
E	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
E	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0
F	0	MET	-	initiating methionine	UNP A0A059PIQ0
F	1	VAL	MET	engineered mutation	UNP A0A059PIQ0
F	2	SER	ARG	engineered mutation	UNP A0A059PIQ0
F	30	ARG	SER	engineered mutation	UNP A0A059PIQ0
F	66	CRO	THR	chromophore	UNP A0A059PIQ0
F	66	CRO	TYR	chromophore	UNP A0A059PIQ0
F	66	CRO	GLY	chromophore	UNP A0A059PIQ0
F	72	SER	ALA	engineered mutation	UNP A0A059PIQ0
F	80	ARG	GLN	engineered mutation	UNP A0A059PIQ0
F	133	4CF	ASP	engineered mutation	UNP A0A059PIQ0
F	149	4CF	ASN	engineered mutation	UNP A0A059PIQ0
F	206	VAL	ALA	engineered mutation	UNP A0A059PIQ0


- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	39	Total O 39 39	0	0
2	B	46	Total O 46 46	0	0
2	C	38	Total O 38 38	0	0
2	D	21	Total O 21 21	0	0
2	E	18	Total O 18 18	0	0
2	F	6	Total O 6 6	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 errors 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.

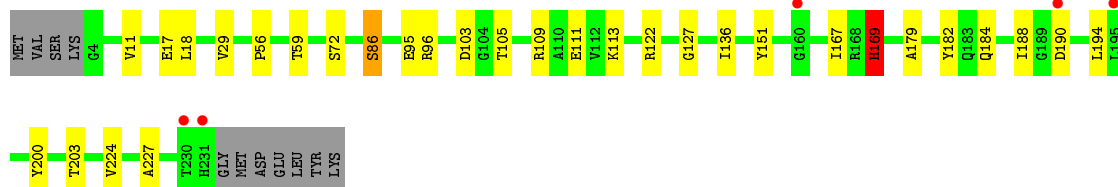
- Molecule 1: Green fluorescent protein

Chain A: 




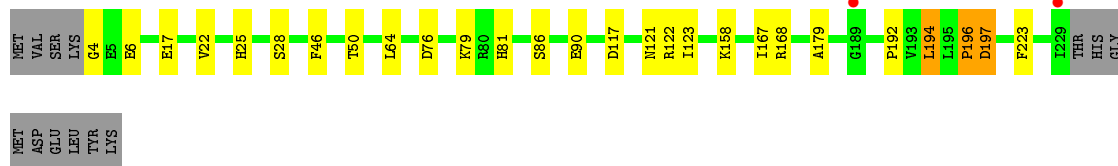
- Molecule 1: Green fluorescent protein

Chain B: 




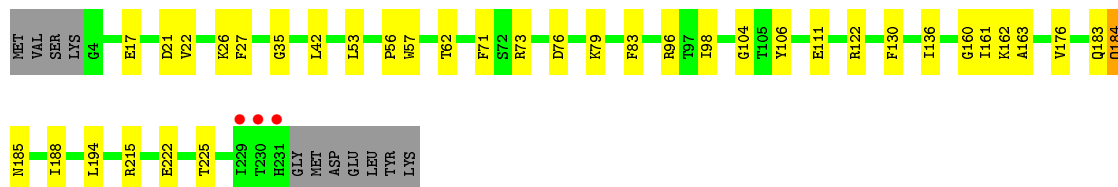
- Molecule 1: Green fluorescent protein

Chain C: 




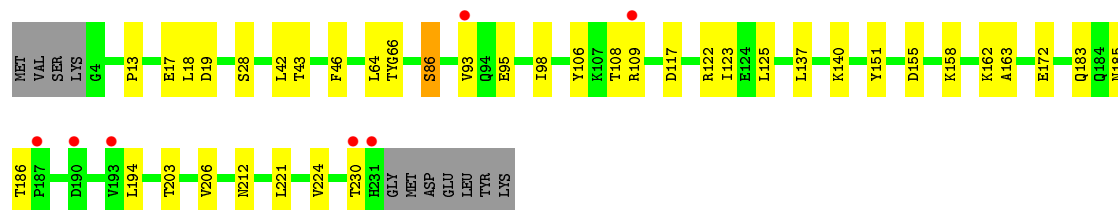
- Molecule 1: Green fluorescent protein

Chain D: 



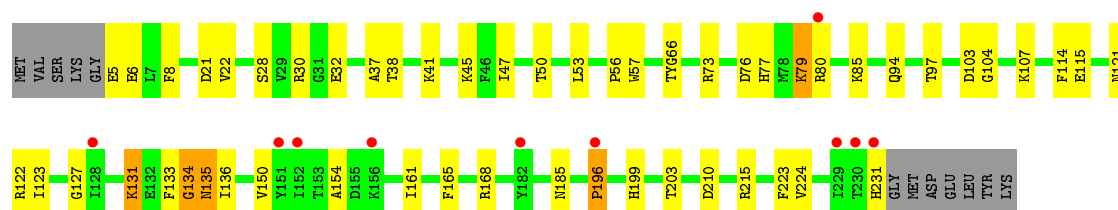
- Molecule 1: Green fluorescent protein

Chain E: 



- Molecule 1: Green fluorescent protein

Chain F: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.94Å 59.07Å 131.76Å 90.00° 108.78° 90.00°	Depositor
Resolution (Å)	43.87 – 2.54 43.87 – 2.52	Depositor EDS
% Data completeness (in resolution range)	99.2 (43.87-2.54) 93.3 (43.87-2.52)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.44 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.189 , 0.243 0.207 , 0.256	Depositor DCC
$R_{free}$ test set	1963 reflections (3.73%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.8	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 35.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 58912 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11017	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 4CF, CRO

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	0/1808	0.64	0/2443
1	B	0.51	0/1819	0.65	0/2457
1	C	0.46	0/1782	0.61	0/2406
1	D	0.44	0/1796	0.59	0/2427
1	E	0.42	0/1796	0.61	0/2427
1	F	0.42	0/1793	0.60	1/2427 (0.0%)
All	All	0.46	0/10794	0.62	1/14587 (0.0%)

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	B	0	1
1	C	0	1
1	E	0	1
1	F	0	1
All	All	0	4

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	F	134	GLY	N-CA-C	6.22	128.64	113.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	169	HIS	Sidechain
1	C	196	PRO	Peptide
1	E	230	THR	Peptide
1	F	133	4CF	Peptide

## 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	1818	0	1746	12	1
1	B	1823	0	1760	20	1
1	C	1794	0	1735	17	0
1	D	1807	0	1738	20	0
1	E	1807	0	1739	21	0
1	F	1800	0	1716	33	0
2	A	39	0	0	2	0
2	B	46	0	0	3	0
2	C	38	0	0	3	0
2	D	21	0	0	0	0
2	E	18	0	0	0	0
2	F	6	0	0	1	0
All	All	11017	0	10434	121	1

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 (121) 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:162:LYS:NZ	1:A:182:TYR:OH	2.14	0.80
1:B:111:GLU:HG2	1:B:188:ILE:HD11	1.69	0.75
1:B:151:TYR:HB2	2:B:305:HOH:O	1.87	0.73
1:B:111:GLU:OE2	2:B:301:HOH:O	2.09	0.69
1:C:6:GLU:OE2	2:C:301:HOH:O	2.10	0.69
1:E:86:SER:HB2	1:E:194:LEU:HD23	1.75	0.68
1:F:76:ASP:O	1:F:80:ARG:NH1	2.22	0.66
1:C:25:HIS:N	2:C:305:HOH:O	2.28	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:95:GLU:HG2	1:B:109:ARG:HG2	1.77	0.66
1:C:17:GLU:OE2	1:C:122:ARG:NH1	2.29	0.66
1:A:213:GLU:O	2:A:301:HOH:O	2.14	0.65
1:F:38:THR:O	1:F:73:ARG:NH2	2.30	0.65
1:F:134:GLY:O	1:F:135:ASN:HB2	1.97	0.64
1:C:81:HIS:ND1	1:C:196:PRO:O	2.25	0.64
1:A:32:GLU:OE2	2:A:302:HOH:O	2.15	0.64
1:C:196:PRO:O	1:C:197:ASP:HB2	1.96	0.64
1:A:74:TYR:O	1:A:79:LYS:NZ	2.31	0.64
1:F:66:CRO:O2	1:F:94:GLN:NE2	2.30	0.63
1:B:203:THR:HG23	1:B:224:VAL:HG22	1.81	0.63
1:E:155:ASP:OD2	1:E:158:LYS:HD2	1.99	0.63
1:A:30:ARG:NH1	1:A:32:GLU:OE1	2.23	0.62
1:D:176:VAL:HG22	1:F:168:ARG:NH1	2.14	0.62
1:A:17:GLU:OE2	1:A:122:ARG:NH1	2.33	0.61
1:B:111:GLU:OE1	1:B:113:LYS:NZ	2.26	0.61
1:B:190:ASP:OD1	1:B:190:ASP:N	2.35	0.60
1:F:21:ASP:O	1:F:127:GLY:N	2.28	0.59
1:B:59:THR:HA	1:B:169:HIS:HE1	1.67	0.59
1:D:21:ASP:OD2	1:D:26:LYS:NZ	2.34	0.59
1:F:30:ARG:NH1	1:F:32:GLU:OE2	2.27	0.58
1:F:41:LYS:HG3	1:F:223:PHE:CE1	2.40	0.57
1:D:98:ILE:HB	1:D:106:TYR:HB2	1.87	0.57
1:F:5:GLU:HA	1:F:85:LYS:HB3	1.87	0.56
1:F:22:VAL:HG22	1:F:127:GLY:HA3	1.87	0.55
1:D:17:GLU:OE2	1:D:122:ARG:NH1	2.37	0.55
1:E:203:THR:HG22	1:E:224:VAL:HG13	1.89	0.55
1:F:115:GLU:OE2	1:F:122:ARG:NH2	2.39	0.54
1:D:111:GLU:HG2	1:D:188:ILE:HD11	1.89	0.54
1:C:167:ILE:HB	1:C:179:ALA:HB3	1.90	0.54
1:B:17:GLU:OE2	1:B:122:ARG:NH1	2.40	0.54
1:E:46:PHE:CD1	1:E:64:LEU:HD13	2.44	0.52
1:C:76:ASP:HA	1:C:79:LYS:HG2	1.90	0.52
1:C:168:ARG:NH2	2:C:307:HOH:O	2.39	0.52
1:F:121:ASN:ND2	1:F:123:ILE:HD11	2.25	0.51
1:B:86:SER:HB2	1:B:194:LEU:HD12	1.92	0.51
1:E:140:LYS:O	1:E:172:GLU:HG2	2.11	0.51
1:D:76:ASP:HA	1:D:79:LYS:HG2	1.92	0.50
1:E:155:ASP:OD2	1:E:162:LYS:HE3	2.11	0.50
1:F:80:ARG:N	1:F:80:ARG:HD3	2.27	0.50
1:E:17:GLU:OE1	1:E:122:ARG:NH1	2.45	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:161:ILE:HG12	1:F:185:ASN:HB2	1.94	0.49
1:F:77:HIS:CG	1:F:231:HIS:HB2	2.47	0.49
1:C:6:GLU:H	1:C:6:GLU:CD	2.16	0.49
1:D:42:LEU:HB2	1:D:222:GLU:HB3	1.94	0.49
1:A:143:TYR:CZ	1:A:209:LYS:HE2	2.48	0.49
1:E:43:THR:HG22	1:E:221:LEU:HD13	1.94	0.49
1:E:13:PRO:HD2	1:E:117:ASP:O	2.12	0.48
1:D:104:GLY:HA3	1:D:130:PHE:CD2	2.48	0.48
1:C:223:PHE:CD2	1:E:206:VAL:HG11	2.49	0.48
1:E:163:ALA:HB3	1:E:183:GLN:HB3	1.96	0.48
1:F:103:ASP:OD1	1:F:104:GLY:N	2.45	0.48
1:E:19:ASP:OD1	1:E:28:SER:OG	2.28	0.47
1:C:4:GLY:N	1:C:6:GLU:OE2	2.47	0.47
1:D:163:ALA:HB3	1:D:183:GLN:HB3	1.95	0.47
1:C:28:SER:HB2	1:C:50:THR:HG23	1.96	0.47
1:D:162:LYS:HE2	1:D:184:GLN:HE21	1.78	0.47
1:F:8:PHE:HB3	1:F:37:ALA:HB3	1.96	0.47
1:B:167:ILE:HB	1:B:179:ALA:HB3	1.97	0.47
1:F:154:ALA:HB2	1:F:196:PRO:HD2	1.97	0.46
1:A:73:ARG:HG2	1:A:73:ARG:O	2.14	0.46
1:F:76:ASP:HA	1:F:79:LYS:HG3	1.96	0.46
1:B:103:ASP:OD2	2:B:302:HOH:O	2.20	0.46
1:D:161:ILE:CG1	1:D:185:ASN:HB2	2.46	0.46
1:D:161:ILE:HG12	1:D:185:ASN:HB2	1.98	0.46
1:B:18:LEU:HB3	1:B:29:VAL:HB	1.98	0.46
1:E:64:LEU:O	1:E:66:CRO:HA31	2.17	0.45
1:F:150:VAL:HG13	1:F:165:PHE:CD1	2.51	0.45
1:E:98:ILE:HB	1:E:106:TYR:HB2	1.99	0.45
1:C:121:ASN:ND2	1:C:123:ILE:HD11	2.32	0.45
1:C:46:PHE:CD1	1:C:64:LEU:HD13	2.52	0.45
1:D:35:GLY:HA3	1:D:71:PHE:CE1	2.52	0.45
1:F:6:GLU:OE2	2:F:301:HOH:O	2.20	0.45
1:E:93:VAL:O	1:E:185:ASN:HA	2.16	0.44
1:F:97:THR:HG23	1:F:107:LYS:HE3	1.99	0.44
1:F:47:ILE:HG12	1:F:215:ARG:CZ	2.47	0.44
1:B:56:PRO:HD3	1:B:136:ILE:O	2.18	0.44
1:D:62:THR:O	1:D:96:ARG:NH1	2.46	0.44
1:F:131:LYS:O	1:F:134:GLY:HA2	2.17	0.44
1:E:95:GLU:HG2	1:E:109:ARG:HG3	1.99	0.44
1:D:53:LEU:HD22	1:D:57:TRP:CE2	2.53	0.43
1:E:151:TYR:O	1:E:163:ALA:HA	2.18	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:128:ILE:O	1:A:129:ASP:HB2	2.18	0.43
1:A:53:LEU:HD22	1:A:57:TRP:CD2	2.53	0.43
1:B:200:TYR:CZ	1:B:227:ALA:HB3	2.53	0.43
1:F:28:SER:HB2	1:F:50:THR:HG23	1.99	0.43
1:E:108:THR:HG22	1:E:125:LEU:HG	2.00	0.43
1:F:94:GLN:HG3	1:F:185:ASN:OD1	2.19	0.43
1:F:6:GLU:CD	1:F:6:GLU:H	2.21	0.43
1:B:95:GLU:OE2	1:B:109:ARG:HD2	2.19	0.42
1:D:22:VAL:HG23	1:D:27:PHE:HE1	1.83	0.42
1:F:76:ASP:HB3	1:F:79:LYS:NZ	2.34	0.42
1:B:96:ARG:HA	1:B:182:TYR:O	2.19	0.42
1:D:73:ARG:HB3	1:D:225:THR:HG22	2.02	0.42
1:D:83:PHE:CE1	1:D:160:GLY:HA2	2.54	0.42
1:E:42:LEU:HD12	1:E:42:LEU:HA	1.83	0.42
1:F:203:THR:HG22	1:F:224:VAL:HG13	2.01	0.42
1:E:137:LEU:HD23	1:E:137:LEU:HA	1.87	0.42
1:B:111:GLU:HB3	1:B:113:LYS:CE	2.49	0.42
1:D:104:GLY:HA3	1:D:130:PHE:CG	2.55	0.42
1:F:76:ASP:O	1:F:79:LYS:HE2	2.20	0.41
1:C:22:VAL:O	1:C:25:HIS:HB2	2.20	0.41
1:B:95:GLU:HB2	1:B:184:GLN:HG3	2.02	0.41
1:F:56:PRO:HD3	1:F:136:ILE:O	2.20	0.41
1:A:89:PRO:HD2	1:A:90:GLU:OE2	2.20	0.41
1:B:105:THR:O	1:B:127:GLY:HA2	2.19	0.41
1:C:86:SER:OG	1:C:194:LEU:HD22	2.21	0.41
1:D:56:PRO:HD3	1:D:136:ILE:O	2.21	0.41
1:A:182:TYR:HE2	1:A:184:GLN:HE21	1.69	0.41
1:E:46:PHE:CE1	1:E:64:LEU:HB3	2.56	0.41
1:C:90:GLU:HG3	1:C:90:GLU:H	1.72	0.41
1:F:53:LEU:HD22	1:F:57:TRP:CD2	2.56	0.40
1:F:45:LYS:NZ	1:F:210:ASP:OD2	2.43	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:157:GLN:NE2	1:B:11:VAL:O[1_565]	2.18	0.02

## 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	220/237 (93%)	215 (98%)	5 (2%)	0	100	100
1	B	221/237 (93%)	217 (98%)	4 (2%)	0	100	100
1	C	217/237 (92%)	208 (96%)	7 (3%)	2 (1%)	21	36
1	D	219/237 (92%)	212 (97%)	7 (3%)	0	100	100
1	E	219/237 (92%)	210 (96%)	8 (4%)	1 (0%)	34	54
1	F	219/237 (92%)	207 (94%)	9 (4%)	3 (1%)	14	23
All	All	1315/1422 (92%)	1269 (96%)	40 (3%)	6 (0%)	34	54

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	212	ASN
1	F	79	LYS
1	F	135	ASN
1	C	197	ASP
1	C	192	PRO
1	F	196	PRO

### 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	194/205 (95%)	192 (99%)	2 (1%)	82	94
1	B	195/205 (95%)	192 (98%)	3 (2%)	72	90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	192/205 (94%)	189 (98%)	3 (2%)	70	89
1	D	193/205 (94%)	190 (98%)	3 (2%)	70	89
1	E	193/205 (94%)	189 (98%)	4 (2%)	61	84
1	F	191/205 (93%)	188 (98%)	3 (2%)	70	89
All	All	1158/1230 (94%)	1140 (98%)	18 (2%)	70	89

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	176	VAL
1	A	194	LEU
1	B	72	SER
1	B	86	SER
1	B	169	HIS
1	C	117	ASP
1	C	158	LYS
1	C	194	LEU
1	D	184	GLN
1	D	194	LEU
1	D	215	ARG
1	E	18	LEU
1	E	86	SER
1	E	123	ILE
1	E	186	THR
1	F	114	PHE
1	F	131	LYS
1	F	199	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	169	HIS
1	C	139	HIS
1	D	184	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



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

18 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	4CF	A	133	1	11,13,14	0.90	0	14,16,18	0.59	0
1	4CF	A	149	1	11,13,14	0.87	0	14,16,18	1.01	0
1	CRO	A	66	1	23,23,24	2.74	6 (26%)	30,32,34	2.66	9 (30%)
1	4CF	B	133	1	11,13,14	0.84	0	14,16,18	0.81	1 (7%)
1	4CF	B	149	1	11,13,14	0.73	0	14,16,18	0.95	0
1	CRO	B	66	1	23,23,24	2.53	5 (21%)	30,32,34	2.70	9 (30%)
1	4CF	C	133	1	11,13,14	0.90	0	14,16,18	0.97	1 (7%)
1	4CF	C	149	1	11,13,14	0.81	0	14,16,18	0.62	0
1	CRO	C	66	1	23,23,24	2.65	5 (21%)	30,32,34	2.86	9 (30%)
1	4CF	D	133	1	11,13,14	0.80	0	14,16,18	0.88	0
1	4CF	D	149	1	11,13,14	0.82	0	14,16,18	0.75	0
1	CRO	D	66	1	23,23,24	2.70	5 (21%)	30,32,34	3.01	11 (36%)
1	4CF	E	133	1	11,13,14	0.84	0	14,16,18	0.81	1 (7%)
1	4CF	E	149	1	11,13,14	0.67	0	14,16,18	0.84	0
1	CRO	E	66	1	23,23,24	2.65	5 (21%)	30,32,34	2.66	9 (30%)
1	4CF	F	133	1	11,13,14	0.84	0	14,16,18	0.63	1 (7%)
1	4CF	F	149	1	11,13,14	0.68	0	14,16,18	0.65	0
1	CRO	F	66	1	23,23,24	2.65	5 (21%)	30,32,34	2.81	9 (30%)

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	4CF	A	133	1	-	0/6/8/10	0/1/1/1
1	4CF	A	149	1	-	0/6/8/10	0/1/1/1
1	CRO	A	66	1	-	0/12/31/32	0/2/2/2
1	4CF	B	133	1	-	0/6/8/10	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	4CF	B	149	1	-	0/6/8/10	0/1/1/1
1	CRO	B	66	1	-	0/12/31/32	0/2/2/2
1	4CF	C	133	1	-	0/6/8/10	0/1/1/1
1	4CF	C	149	1	-	0/6/8/10	0/1/1/1
1	CRO	C	66	1	-	0/12/31/32	0/2/2/2
1	4CF	D	133	1	-	0/6/8/10	0/1/1/1
1	4CF	D	149	1	-	0/6/8/10	0/1/1/1
1	CRO	D	66	1	-	0/12/31/32	0/2/2/2
1	4CF	E	133	1	-	0/6/8/10	0/1/1/1
1	4CF	E	149	1	-	0/6/8/10	0/1/1/1
1	CRO	E	66	1	-	0/12/31/32	0/2/2/2
1	4CF	F	133	1	-	0/6/8/10	0/1/1/1
1	4CF	F	149	1	-	0/6/8/10	0/1/1/1
1	CRO	F	66	1	-	0/12/31/32	0/2/2/2

All (31) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	O2-C2	-2.16	1.18	1.23
1	B	66	CRO	C1-N3	2.35	1.41	1.37
1	D	66	CRO	C1-N3	2.47	1.41	1.37
1	A	66	CRO	C1-N3	2.64	1.42	1.37
1	C	66	CRO	C1-N3	2.65	1.42	1.37
1	E	66	CRO	C1-N3	2.78	1.42	1.37
1	F	66	CRO	OH-CZ	2.90	1.44	1.37
1	C	66	CRO	OH-CZ	2.95	1.44	1.37
1	F	66	CRO	C1-N3	3.10	1.42	1.37
1	D	66	CRO	OH-CZ	3.16	1.44	1.37
1	A	66	CRO	OH-CZ	3.30	1.44	1.37
1	E	66	CRO	OH-CZ	3.41	1.45	1.37
1	B	66	CRO	OH-CZ	3.49	1.45	1.37
1	D	66	CRO	CG2-CB2	3.53	1.53	1.46
1	F	66	CRO	CG2-CB2	3.57	1.53	1.46
1	C	66	CRO	CG2-CB2	3.58	1.53	1.46
1	B	66	CRO	CG2-CB2	3.71	1.54	1.46
1	A	66	CRO	CG2-CB2	3.76	1.54	1.46
1	E	66	CRO	CG2-CB2	3.88	1.54	1.46
1	A	66	CRO	C2-N3	4.08	1.49	1.39
1	B	66	CRO	C2-N3	4.60	1.51	1.39
1	E	66	CRO	C2-N3	4.66	1.51	1.39
1	D	66	CRO	C2-N3	4.95	1.51	1.39
1	F	66	CRO	C2-N3	5.09	1.52	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	66	CRO	C2-N3	5.25	1.52	1.39
1	B	66	CRO	CA2-C2	9.14	1.58	1.48
1	E	66	CRO	CA2-C2	9.69	1.58	1.48
1	F	66	CRO	CA2-C2	9.69	1.58	1.48
1	C	66	CRO	CA2-C2	9.70	1.58	1.48
1	D	66	CRO	CA2-C2	10.25	1.59	1.48
1	A	66	CRO	CA2-C2	10.50	1.59	1.48

All (60) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	66	CRO	C2-CA2-N2	-7.00	103.71	109.03
1	C	66	CRO	O2-C2-CA2	-6.60	127.17	130.97
1	E	66	CRO	C2-CA2-N2	-6.38	104.17	109.03
1	C	66	CRO	C2-CA2-N2	-6.17	104.34	109.03
1	B	66	CRO	O2-C2-CA2	-6.08	127.47	130.97
1	F	66	CRO	C2-CA2-N2	-6.04	104.44	109.03
1	A	66	CRO	CA1-C1-N3	-6.01	117.70	124.98
1	F	66	CRO	O2-C2-CA2	-5.89	127.58	130.97
1	B	66	CRO	C2-CA2-N2	-5.56	104.80	109.03
1	D	66	CRO	O2-C2-CA2	-5.53	127.79	130.97
1	B	66	CRO	CA1-C1-N3	-5.26	118.62	124.98
1	A	66	CRO	C2-CA2-N2	-5.09	105.16	109.03
1	C	66	CRO	CA1-C1-N3	-4.37	119.69	124.98
1	D	66	CRO	CA1-C1-N3	-4.03	120.09	124.98
1	F	66	CRO	CG2-CB2-CA2	-3.58	125.74	130.27
1	E	66	CRO	O2-C2-CA2	-3.50	128.96	130.97
1	E	66	CRO	CA1-C1-N3	-3.32	120.96	124.98
1	A	66	CRO	CG1-CB1-CA1	-3.23	106.77	112.42
1	D	66	CRO	C2-N3-C1	-3.22	102.27	108.28
1	E	66	CRO	CG1-CB1-CA1	-2.99	107.19	112.42
1	A	66	CRO	O2-C2-CA2	-2.97	129.26	130.97
1	A	66	CRO	C2-N3-C1	-2.91	102.86	108.28
1	C	66	CRO	C2-N3-C1	-2.85	102.96	108.28
1	C	133	4CF	CG-CB-CA	-2.82	107.63	114.12
1	A	66	CRO	CG2-CB2-CA2	-2.71	126.83	130.27
1	D	66	CRO	CG1-CB1-CA1	-2.69	107.70	112.42
1	E	66	CRO	C2-N3-C1	-2.68	103.28	108.28
1	F	66	CRO	C2-N3-C1	-2.63	103.36	108.28
1	C	66	CRO	CG2-CB2-CA2	-2.57	127.02	130.27
1	F	66	CRO	CA1-C1-N3	-2.56	121.88	124.98
1	D	66	CRO	CG2-CB2-CA2	-2.51	127.09	130.27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	66	CRO	C2-N3-C1	-2.51	103.59	108.28
1	F	66	CRO	CG1-CB1-CA1	-2.47	108.09	112.42
1	B	133	4CF	CG-CB-CA	-2.16	109.15	114.12
1	E	133	4CF	CG-CB-CA	-2.01	109.49	114.12
1	C	66	CRO	CA3-N3-C2	2.01	128.38	124.21
1	B	66	CRO	CA1-C1-N2	2.06	126.65	123.71
1	F	133	4CF	CB-CA-N	2.06	121.64	111.91
1	B	66	CRO	CA2-C2-N3	2.12	104.42	103.37
1	B	66	CRO	CA3-N3-C2	2.15	128.65	124.21
1	D	66	CRO	CB2-CA2-C2	2.17	125.29	122.24
1	F	66	CRO	CB2-CA2-C2	2.21	125.35	122.24
1	D	66	CRO	CA3-N3-C2	2.24	128.85	124.21
1	A	66	CRO	CA2-C2-N3	2.63	104.67	103.37
1	E	66	CRO	CA3-N3-C1	2.69	130.35	127.38
1	E	66	CRO	CA2-C2-N3	2.97	104.84	103.37
1	C	66	CRO	CA2-C2-N3	3.16	104.93	103.37
1	D	66	CRO	CA2-C2-N3	3.57	105.14	103.37
1	F	66	CRO	N3-C1-N2	3.59	114.19	111.53
1	E	66	CRO	N3-C1-N2	3.97	114.47	111.53
1	B	66	CRO	N3-C1-N2	4.31	114.73	111.53
1	C	66	CRO	N3-C1-N2	4.53	114.88	111.53
1	D	66	CRO	N3-C1-N2	5.39	115.53	111.53
1	A	66	CRO	CA2-N2-C1	6.38	110.90	105.54
1	A	66	CRO	N3-C1-N2	6.58	116.41	111.53
1	B	66	CRO	CA2-N2-C1	8.22	112.45	105.54
1	C	66	CRO	CA2-N2-C1	8.74	112.89	105.54
1	E	66	CRO	CA2-N2-C1	9.15	113.23	105.54
1	D	66	CRO	CA2-N2-C1	9.31	113.37	105.54
1	F	66	CRO	CA2-N2-C1	9.71	113.70	105.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	66	CRO	1	0
1	F	66	CRO	1	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 ⓘ

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	223/237 (94%)	-0.16	1 (0%) 93 94	33, 45, 65, 97	0
1	B	223/237 (94%)	-0.11	5 (2%) 65 70	32, 49, 72, 109	0
1	C	221/237 (93%)	-0.02	2 (0%) 85 88	36, 51, 72, 97	0
1	D	223/237 (94%)	-0.01	3 (1%) 79 83	45, 59, 79, 108	0
1	E	223/237 (94%)	-0.04	7 (3%) 52 58	44, 63, 80, 109	0
1	F	222/237 (93%)	0.36	10 (4%) 37 43	48, 74, 103, 123	0
All	All	1335/1422 (93%)	0.00	28 (2%) 67 72	32, 56, 89, 123	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	229	ILE	5.0
1	A	231	HIS	4.3
1	B	231	HIS	4.3
1	F	231	HIS	3.7
1	F	230	THR	3.6
1	E	231	HIS	3.6
1	F	182[A]	TYR	3.6
1	D	230	THR	3.4
1	D	229	ILE	3.4
1	C	189	GLY	3.2
1	F	80	ARG	3.0
1	F	151	TYR	3.0
1	B	160	GLY	2.7
1	B	230	THR	2.5
1	E	93	VAL	2.4
1	E	190	ASP	2.4
1	F	128	ILE	2.4
1	E	230	THR	2.4
1	F	152	ILE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	231	HIS	2.3
1	B	190	ASP	2.2
1	E	187	PRO	2.2
1	F	156	LYS	2.2
1	B	195	LEU	2.2
1	C	229	ILE	2.1
1	E	109	ARG	2.1
1	E	193	VAL	2.1
1	F	196	PRO	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	4CF	D	133	13/14	0.84	0.22	-	60,68,74,78	0
1	4CF	E	133	13/14	0.94	0.22	-	70,75,83,85	0
1	4CF	B	133	13/14	0.94	0.14	-	46,50,54,55	0
1	4CF	C	133	13/14	0.94	0.14	-	48,57,62,64	0
1	4CF	F	149	13/14	0.94	0.19	-	66,69,72,72	0
1	CRO	D	66	22/23	0.96	0.16	-	43,48,54,59	0
1	CRO	F	66	22/23	0.95	0.20	-	59,65,70,74	0
1	CRO	B	66	22/23	0.97	0.15	-	30,36,46,50	0
1	4CF	A	149	13/14	0.96	0.16	-	32,37,42,43	0
1	4CF	C	149	13/14	0.94	0.22	-	40,48,57,57	0
1	4CF	B	149	13/14	0.96	0.17	-	36,41,46,46	0
1	4CF	E	149	13/14	0.98	0.13	-	44,45,51,51	0
1	4CF	D	149	13/14	0.97	0.12	-	44,52,59,61	0
1	4CF	A	133	13/14	0.84	0.21	-	54,62,67,71	0
1	CRO	C	66	22/23	0.97	0.19	-	39,43,45,47	0
1	CRO	E	66	22/23	0.98	0.12	-	41,52,59,65	0
1	CRO	A	66	22/23	0.97	0.20	-	30,35,39,41	0
1	4CF	F	133	13/14	0.79	0.38	-	87,98,104,104	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.