



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:38 AM GMT

PDB ID : 2B3Q
Title : Crystal structure of a well-folded variant of green fluorescent protein
Authors : Pedelacq, J.D.; Cabantous, S.; Tran, T.H.; Terwilliger, T.C.; Waldo, G.S.
Deposited on : 2005-09-20
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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

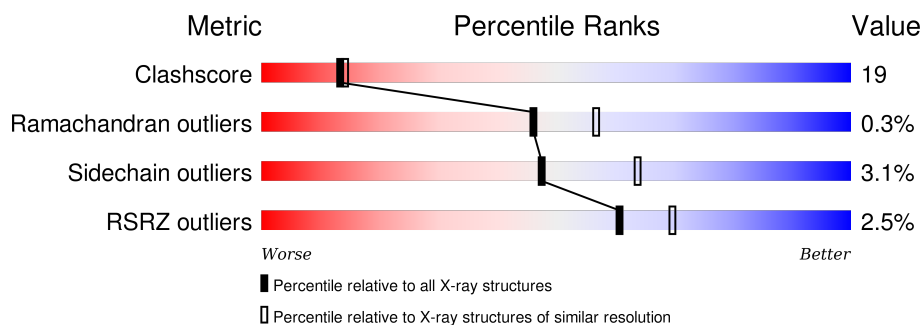
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (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. 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	244	
1	B	244	
1	C	244	
1	D	244	

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
2	MG	A	702	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7676 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	227	Total	C	N	O	S	0	0	0
			1806	1145	309	347	5			
1	B	227	Total	C	N	O	S	0	0	0
			1792	1136	304	347	5			
1	C	227	Total	C	N	O	S	0	0	0
			1806	1144	308	349	5			
1	D	227	Total	C	N	O	S	0	0	0
			1814	1150	310	349	5			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	64	LEU	PHE	ENGINEERED	UNP P42212
A	66	CRO	SER	CHROMOPHORE	UNP P42212
A	66	CRO	TYR	CHROMOPHORE	UNP P42212
A	66	CRO	GLY	CHROMOPHORE	UNP P42212
A	80	ARG	GLN	ENGINEERED	UNP P42212
A	99	SER	PHE	ENGINEERED	UNP P42212
A	153	THR	MET	ENGINEERED	UNP P42212
A	163	ALA	VAL	ENGINEERED	UNP P42212
A	239	GLY	-	EXPRESSION TAG	UNP P42212
A	240	SER	-	EXPRESSION TAG	UNP P42212
A	241	HIS	-	EXPRESSION TAG	UNP P42212
A	242	HIS	-	EXPRESSION TAG	UNP P42212
A	243	HIS	-	EXPRESSION TAG	UNP P42212
A	244	HIS	-	EXPRESSION TAG	UNP P42212
A	245	HIS	-	EXPRESSION TAG	UNP P42212
A	246	HIS	-	EXPRESSION TAG	UNP P42212
B	64	LEU	PHE	ENGINEERED	UNP P42212
B	66	CRO	SER	CHROMOPHORE	UNP P42212
B	66	CRO	TYR	CHROMOPHORE	UNP P42212
B	66	CRO	GLY	CHROMOPHORE	UNP P42212
B	80	ARG	GLN	ENGINEERED	UNP P42212

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	99	SER	PHE	ENGINEERED	UNP P42212
B	153	THR	MET	ENGINEERED	UNP P42212
B	163	ALA	VAL	ENGINEERED	UNP P42212
B	239	GLY	-	EXPRESSION TAG	UNP P42212
B	240	SER	-	EXPRESSION TAG	UNP P42212
B	241	HIS	-	EXPRESSION TAG	UNP P42212
B	242	HIS	-	EXPRESSION TAG	UNP P42212
B	243	HIS	-	EXPRESSION TAG	UNP P42212
B	244	HIS	-	EXPRESSION TAG	UNP P42212
B	245	HIS	-	EXPRESSION TAG	UNP P42212
B	246	HIS	-	EXPRESSION TAG	UNP P42212
C	64	LEU	PHE	ENGINEERED	UNP P42212
C	66	CRO	SER	CHROMOPHORE	UNP P42212
C	66	CRO	TYR	CHROMOPHORE	UNP P42212
C	66	CRO	GLY	CHROMOPHORE	UNP P42212
C	80	ARG	GLN	ENGINEERED	UNP P42212
C	99	SER	PHE	ENGINEERED	UNP P42212
C	153	THR	MET	ENGINEERED	UNP P42212
C	163	ALA	VAL	ENGINEERED	UNP P42212
C	239	GLY	-	EXPRESSION TAG	UNP P42212
C	240	SER	-	EXPRESSION TAG	UNP P42212
C	241	HIS	-	EXPRESSION TAG	UNP P42212
C	242	HIS	-	EXPRESSION TAG	UNP P42212
C	243	HIS	-	EXPRESSION TAG	UNP P42212
C	244	HIS	-	EXPRESSION TAG	UNP P42212
C	245	HIS	-	EXPRESSION TAG	UNP P42212
C	246	HIS	-	EXPRESSION TAG	UNP P42212
D	64	LEU	PHE	ENGINEERED	UNP P42212
D	66	CRO	SER	CHROMOPHORE	UNP P42212
D	66	CRO	TYR	CHROMOPHORE	UNP P42212
D	66	CRO	GLY	CHROMOPHORE	UNP P42212
D	80	ARG	GLN	ENGINEERED	UNP P42212
D	99	SER	PHE	ENGINEERED	UNP P42212
D	153	THR	MET	ENGINEERED	UNP P42212
D	163	ALA	VAL	ENGINEERED	UNP P42212
D	239	GLY	-	EXPRESSION TAG	UNP P42212
D	240	SER	-	EXPRESSION TAG	UNP P42212
D	241	HIS	-	EXPRESSION TAG	UNP P42212
D	242	HIS	-	EXPRESSION TAG	UNP P42212
D	243	HIS	-	EXPRESSION TAG	UNP P42212
D	244	HIS	-	EXPRESSION TAG	UNP P42212
D	245	HIS	-	EXPRESSION TAG	UNP P42212

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	246	HIS	-	EXPRESSION TAG	UNP P42212

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

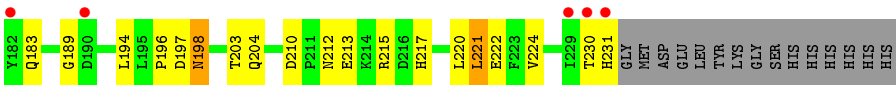
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total Mg 3 3	0	0

- Molecule 3 is water.

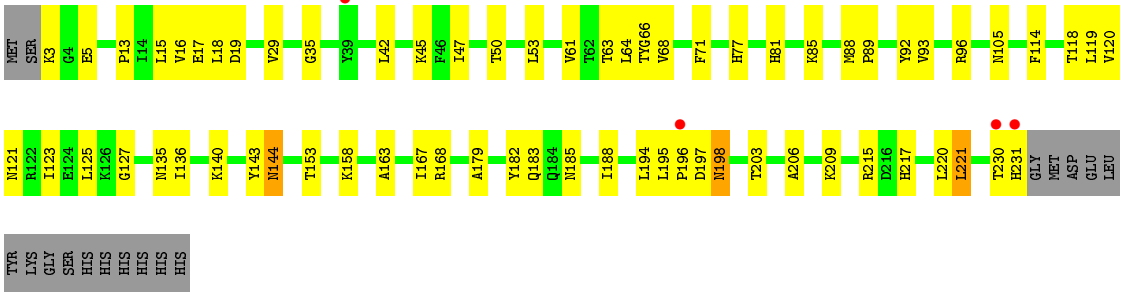
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	152	Total O 152 152	0	0
3	B	120	Total O 120 120	0	0
3	C	67	Total O 67 67	0	0
3	D	116	Total O 116 116	0	0

- Molecule 1: green fluorescent protein





● Molecule 1: green fluorescent protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	85.41Å 87.16Å 145.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 24.93 – 2.20	Depositor EDS
% Data completeness (in resolution range)	81.8 (20.00-2.30) 80.3 (24.93-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.91 (at 2.19Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.206 , 0.259 0.223 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	25.9	Xtriage
Anisotropy	0.927	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 49.7	EDS
Estimated twinning fraction	0.023 for k,h,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 44852 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7676	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.73% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, 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	0.58	0/1824	0.87	4/2467 (0.2%)
1	B	0.53	1/1810 (0.1%)	0.76	2/2452 (0.1%)
1	C	0.62	3/1824 (0.2%)	0.91	10/2468 (0.4%)
1	D	0.54	0/1832	0.79	1/2476 (0.0%)
All	All	0.57	4/7290 (0.1%)	0.83	17/9863 (0.2%)

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	0	2
1	C	0	2
1	D	0	2
All	All	0	6

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	92	TYR	C-N	9.12	1.55	1.34
1	C	204	GLN	C-N	-5.64	1.21	1.34
1	C	198	ASN	C-N	-5.19	1.22	1.34
1	B	40	GLY	C-N	5.02	1.45	1.34

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	88	MET	O-C-N	-13.29	95.85	121.10
1	C	84	PHE	O-C-N	-9.54	107.44	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	88	MET	CA-C-N	9.49	143.67	117.10
1	B	122	ARG	NE-CZ-NH2	7.61	124.11	120.30
1	C	109	ARG	NE-CZ-NH2	7.57	124.08	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	198	ASN	Mainchain
1	A	221	LEU	Mainchain
1	C	84	PHE	Mainchain
1	C	88	MET	Mainchain
1	D	50	THR	Mainchain

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	1806	0	1747	59	0
1	B	1792	0	1714	67	0
1	C	1806	0	1738	85	0
1	D	1814	0	1762	61	0
2	A	3	0	0	0	0
3	A	152	0	0	4	0
3	B	120	0	0	1	0
3	C	67	0	0	4	0
3	D	116	0	0	1	0
All	All	7676	0	6961	267	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 19.

The worst 5 of 267 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:D:13:PRO:HB2	1:D:118:THR:HG22	1.41	0.97

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:64:LEU:C	1:D:66:CRO:H2	1.72	0.93
1:D:71:PHE:CE2	1:D:119:LEU:HD22	2.06	0.91
1:A:163:ALA:HB3	1:A:183:GLN:HB3	1.54	0.89
1:D:71:PHE:HE2	1:D:119:LEU:HD22	1.39	0.88

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	222/244 (91%)	217 (98%)	5 (2%)	0	100	100
1	B	222/244 (91%)	205 (92%)	17 (8%)	0	100	100
1	C	222/244 (91%)	205 (92%)	14 (6%)	3 (1%)	14	13
1	D	222/244 (91%)	217 (98%)	5 (2%)	0	100	100
All	All	888/976 (91%)	844 (95%)	41 (5%)	3 (0%)	46	57

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	157	GLN
1	C	136	ILE
1	C	61	VAL

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	195/212 (92%)	189 (97%)	6 (3%)	47	64
1	B	192/212 (91%)	186 (97%)	6 (3%)	47	64
1	C	195/212 (92%)	190 (97%)	5 (3%)	54	71
1	D	197/212 (93%)	190 (96%)	7 (4%)	42	57
All	All	779/848 (92%)	755 (97%)	24 (3%)	47	64

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	203	THR
1	C	144	ASN
1	D	203	THR
1	B	221	LEU
1	C	88	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	157	GLN
1	C	135	ASN
1	D	149	ASN
1	B	170	ASN
1	A	149	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	CRO	A	66	-	23,23,24	2.74	9 (39%)	29,32,34	1.99	8 (27%)
1	CRO	B	66	-	23,23,24	2.77	7 (30%)	29,32,34	1.90	9 (31%)
1	CRO	C	66	-	23,23,24	2.67	9 (39%)	29,32,34	1.95	9 (31%)
1	CRO	D	66	-	23,23,24	2.59	9 (39%)	29,32,34	1.77	7 (24%)

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	-	-	0/12/31/32	0/2/2/2
1	CRO	B	66	-	-	0/12/31/32	0/2/2/2
1	CRO	C	66	-	-	0/12/31/32	0/2/2/2
1	CRO	D	66	-	-	0/12/31/32	0/2/2/2

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	66	CRO	OH-CZ	-4.62	1.26	1.37
1	B	66	CRO	OH-CZ	-4.28	1.26	1.37
1	A	66	CRO	OH-CZ	-4.11	1.27	1.37
1	C	66	CRO	OH-CZ	-4.04	1.27	1.37
1	C	66	CRO	CG2-CB2	-2.42	1.42	1.46

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	CRO	CA3-N3-C1	-3.71	123.05	127.36
1	C	66	CRO	CE2-CZ-CE1	-3.49	114.82	119.79
1	A	66	CRO	CE2-CZ-CE1	-3.35	115.02	119.79
1	D	66	CRO	CE2-CZ-CE1	-3.32	115.06	119.79
1	B	66	CRO	CE2-CZ-CE1	-3.28	115.12	119.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	66	CRO	6	0
1	B	66	CRO	8	0
1	C	66	CRO	6	0
1	D	66	CRO	8	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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, 95th 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 ²)	Q<0.9
1	A	226/244 (92%)	-0.35	1 (0%) 93 95	12, 22, 35, 65	0
1	B	226/244 (92%)	0.03	7 (3%) 52 62	13, 32, 52, 68	0
1	C	226/244 (92%)	0.29	11 (4%) 33 42	20, 36, 56, 75	0
1	D	226/244 (92%)	-0.24	4 (1%) 71 78	13, 24, 39, 58	0
All	All	904/976 (92%)	-0.07	23 (2%) 61 70	12, 28, 50, 75	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	231	HIS	4.4
1	B	231	HIS	4.3
1	C	230	THR	4.3
1	C	6	GLU	3.7
1	B	128	ILE	3.6

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

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, 95th 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 ²)	Q<0.9
1	CRO	C	66	22/23	0.94	0.20	-	16,23,29,30	0
1	CRO	A	66	22/23	0.96	0.19	-	13,14,16,23	0
1	CRO	D	66	22/23	0.96	0.17	-	11,14,16,22	0
1	CRO	B	66	22/23	0.95	0.20	-	18,22,24,26	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. 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, 95th 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(\AA^2)	Q<0.9
2	MG	A	702	1/1	1.00	0.15	4.74	8,8,8,8	0
2	MG	A	703	1/1	0.91	0.08	-	45,45,45,45	0
2	MG	A	701	1/1	0.91	0.13	-	49,49,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.