



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:37 PM GMT

PDB ID : 4CRY  
Title : Direct visualisation of strain-induced protein post-translational modification  
Authors : Monteiro, D.C.F.; Patel, V.; Bartlett, C.P.; Grant, T.D.; Nozaki, S.; Gowdy, J.A.; Snell, E.H.; Niki, H.; Pearson, A.R.; Webb, M.E.  
Deposited on : 2014-03-02  
Resolution : 1.61 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<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

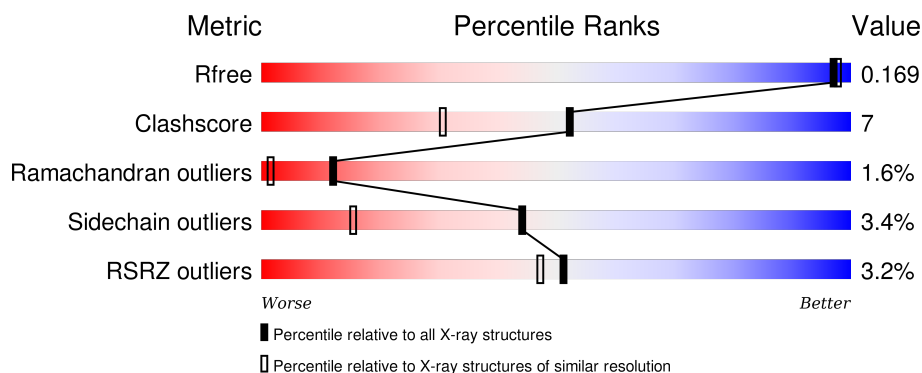
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.61 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	3202 (1.64-1.60)
Clashscore	102246	3500 (1.64-1.60)
Ramachandran outliers	100387	3411 (1.64-1.60)
Sidechain outliers	100360	3410 (1.64-1.60)
RSRZ outliers	91569	3207 (1.64-1.60)

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	41	
2	B	137	
3	G	102	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 2271 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 ASPARTATE 1-DECARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	25	Total	C	N	O	S	0	1	0
			209	129	41	37	2			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	MET	-	EXPRESSION TAG	UNP P0A790
A	-15	ARG	-	EXPRESSION TAG	UNP P0A790
A	-14	GLY	-	EXPRESSION TAG	UNP P0A790
A	-13	SER	-	EXPRESSION TAG	UNP P0A790
A	-12	HIS	-	EXPRESSION TAG	UNP P0A790
A	-11	HIS	-	EXPRESSION TAG	UNP P0A790
A	-10	HIS	-	EXPRESSION TAG	UNP P0A790
A	-9	HIS	-	EXPRESSION TAG	UNP P0A790
A	-8	HIS	-	EXPRESSION TAG	UNP P0A790
A	-7	HIS	-	EXPRESSION TAG	UNP P0A790
A	-6	GLY	-	EXPRESSION TAG	UNP P0A790
A	-5	LEU	-	EXPRESSION TAG	UNP P0A790
A	-4	VAL	-	EXPRESSION TAG	UNP P0A790
A	-3	PRO	-	EXPRESSION TAG	UNP P0A790
A	-2	ARG	-	EXPRESSION TAG	UNP P0A790
A	-1	GLY	-	EXPRESSION TAG	UNP P0A790
A	0	SER	-	EXPRESSION TAG	UNP P0A790

- Molecule 2 is a protein called PANZ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	127	Total	C	N	O	S	0	9	0
			1079	678	204	190	7			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	128	SER	-	EXPRESSION TAG	UNP P37613
B	129	GLY	-	EXPRESSION TAG	UNP P37613
B	130	LEU	-	EXPRESSION TAG	UNP P37613
B	131	GLU	-	EXPRESSION TAG	UNP P37613
B	132	HIS	-	EXPRESSION TAG	UNP P37613
B	133	HIS	-	EXPRESSION TAG	UNP P37613
B	134	HIS	-	EXPRESSION TAG	UNP P37613
B	135	HIS	-	EXPRESSION TAG	UNP P37613
B	136	HIS	-	EXPRESSION TAG	UNP P37613
B	137	HIS	-	EXPRESSION TAG	UNP P37613

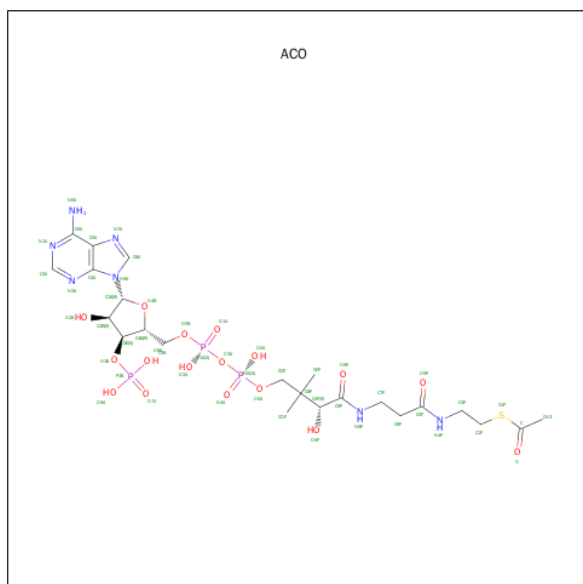
- Molecule 3 is a protein called ASPARTATE 1-DECARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	102	Total	C	N	O	S	0	5	0
			806	504	142	156	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	57	VAL	THR	ENGINEERED MUTATION	UNP P0A790

- Molecule 4 is ACETYL COENZYME \*A (three-letter code: ACO) (formula:  $C_{23}H_{38}N_7O_{17}P_3S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	S	0
			51	23	7	17	3	1	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mg	0	0
			1	1		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Cl	0	0
			1	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	19	Total	O	0	0
			19	19		
7	B	56	Total	O	0	0
			56	56		
7	G	49	Total	O	0	0
			49	49		



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.40 Å 86.40 Å 81.00 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.19 – 1.61 43.20 – 1.61	Depositor EDS
% Data completeness (in resolution range)	97.3 (43.19-1.61) 98.8 (43.20-1.61)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.40 (at 1.61 Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.129 , 0.150 0.151 , 0.169	Depositor DCC
$R_{free}$ test set	1942 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.4	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 50.2	EDS
Estimated twinning fraction	0.841 for H, K, L 0.159 for K, H, -L 0.014 for -k,-h,-l	Xtriage
Reported twinning fraction	0.841 for H, K, L 0.159 for K, H, -L	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 38084 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2271	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.78% 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: CSO, ACO, MG, CL

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	1.83	6/211 (2.8%)	1.80	6/276 (2.2%)
2	B	1.30	2/1124 (0.2%)	1.47	15/1512 (1.0%)
3	G	1.41	1/826 (0.1%)	1.53	10/1118 (0.9%)
All	All	1.40	9/2161 (0.4%)	1.53	31/2906 (1.1%)

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	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	24[A]	GLY	N-CA	9.64	1.60	1.46
1	A	24[B]	GLY	N-CA	9.64	1.60	1.46
2	B	57	SER	CB-OG	-7.25	1.32	1.42
1	A	9	LYS	CB-CG	-6.17	1.35	1.52
1	A	3	ARG	NE-CZ	-5.96	1.25	1.33

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	38	ARG	NE-CZ-NH1	16.47	128.54	120.30
3	G	34	ASP	CB-CG-OD2	-13.98	105.72	118.30
2	B	38	ARG	NE-CZ-NH2	-11.60	114.50	120.30
1	A	12	ARG	NE-CZ-NH2	-11.59	114.50	120.30
2	B	47	ARG	NE-CZ-NH2	-10.59	115.01	120.30



There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	23	GLU	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	209	0	215	2	0
2	B	1079	0	1108	24	1
3	G	806	0	796	5	1
4	B	51	0	34	0	0
5	B	1	0	0	0	0
6	B	1	0	0	0	0
7	A	19	0	0	0	0
7	B	56	0	0	2	0
7	G	49	0	0	0	0
All	All	2271	0	2153	29	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 7.

The worst 5 of 29 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)
3:G:67[A]:ARG:HH22	3:G:113[A]:GLU:HG2	1.19	1.00
2:B:28:PRO:HA	2:B:29:SER:C	1.88	0.94
3:G:67[A]:ARG:NH2	3:G:113[A]:GLU:HG2	1.83	0.93
2:B:28:PRO:HG2	2:B:31:LEU:HD22	1.65	0.79
2:B:27:SER:N	2:B:28:PRO:HD2	2.01	0.76

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)
2:B:75[A]:ARG:NH1	3:G:40:GLU:OE2[4_555]	1.85	0.35

## 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	23/41 (56%)	22 (96%)	1 (4%)	0	100	100
2	B	134/137 (98%)	128 (96%)	3 (2%)	3 (2%)	8	1
3	G	103/102 (101%)	102 (99%)	0	1 (1%)	19	4
All	All	260/280 (93%)	252 (97%)	4 (2%)	4 (2%)	12	2

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	28	PRO
2	B	30	SER
2	B	27	SER
3	G	57	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	22/35 (63%)	21 (96%)	1 (4%)	34	9
2	B	116/116 (100%)	111 (96%)	5 (4%)	35	10
3	G	83/78 (106%)	82 (99%)	1 (1%)	78	59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	221/229 (96%)	214 (97%)	7 (3%)	44 17

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	31	LEU
3	G	89	SER
2	B	32	GLN
2	B	25	GLU
2	B	35	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	120	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
3	CSO	G	78	3	3,6,7	1.42	1 (33%)	1,6,8	2.28	1 (100%)

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
3	CSO	G	78	3	-	0/1/5/7	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	78	CSO	CB-CA	-2.34	1.47	1.53

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	78	CSO	O-C-CA	-2.28	119.56	125.49

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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	$\# Z  > 2$
4	ACO	B	1128	5	43,53,53	1.14	2 (4%)	55,79,79	2.29	17 (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
4	ACO	B	1128	5	-	0/47/67/67	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1128	ACO	C4A-N3A	2.30	1.39	1.35
4	B	1128	ACO	P3B-O3B	2.38	1.67	1.60

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1128	ACO	N3A-C2A-N1A	-9.89	121.32	128.89
4	B	1128	ACO	O6A-CCP-CBP	-3.59	104.78	110.55
4	B	1128	ACO	O8A-P3B-O7A	-3.45	99.46	110.58
4	B	1128	ACO	O9P-C9P-N8P	-3.30	116.47	123.08
4	B	1128	ACO	C1B-N9A-C4A	-3.16	122.18	126.94

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	25/41 (60%)	-0.04	0 <span style="border: 1px solid blue; padding: 0 2px;">100</span> <span style="border: 1px solid blue; padding: 0 2px;">100</span>	14, 17, 32, 52	0
2	B	127/137 (92%)	0.04	8 (6%) <span style="border: 1px solid red; padding: 0 2px;">23</span> <span style="border: 1px solid red; padding: 0 2px;">20</span>	16, 29, 51, 100	0
3	G	101/102 (99%)	-0.18	0 <span style="border: 1px solid blue; padding: 0 2px;">100</span> <span style="border: 1px solid blue; padding: 0 2px;">100</span>	13, 17, 33, 36	0
All	All	253/280 (90%)	-0.05	8 (3%) <span style="border: 1px solid gray; padding: 0 2px;">51</span> <span style="border: 1px solid gray; padding: 0 2px;">47</span>	13, 22, 45, 100	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	29	SER	5.8
2	B	28	PRO	5.4
2	B	27	SER	5.0
2	B	32	GLN	4.9
2	B	31	LEU	3.1

### 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
3	CSO	G	78	7/8	0.99	0.08	-	15,17,24,33	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, 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( $\text{\AA}^2$ )	Q<0.9
5	MG	B	1129	1/1	1.00	0.07	0.60	26,26,26,26	0
4	ACO	B	1128	51/51	0.98	0.06	-0.79	18,28,36,37	0
6	CL	B	1130	1/1	0.92	0.48	-	30,30,30,30	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.