



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 05:10 pm GMT

PDB ID : 4CFU  
Title : Structure-based design of C8-substituted O6-cyclohexylmethoxyguanine CDK1 and 2 inhibitors.  
Authors : Carbain, B.; Paterson, D.J.; Anscombe, E.; Campbell, A.; Cano, C.; Echaliier, A.; Endicott, J.; Golding, B.T.; Haggerty, K.; Hardcastle, I.R.; Jewsbury, P.; Newell, D.R.; Noble, M.E.M.; Roche, C.; Wang, L.Z.; Griffin, R.  
Deposited on : 2013-11-19  
Resolution : 2.20 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

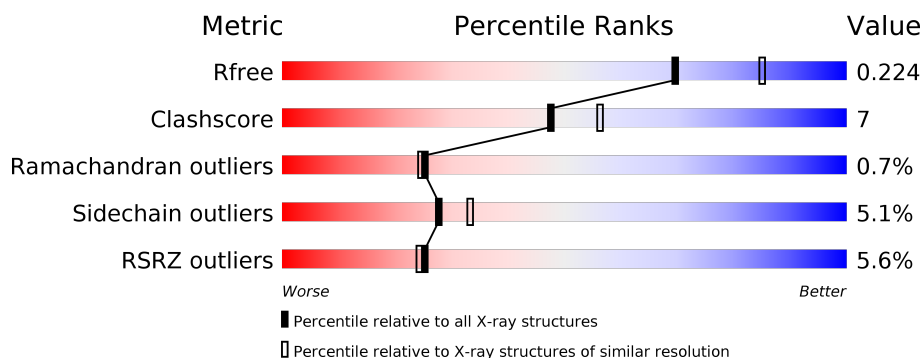
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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	303	<div> <div>6%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>• •</div> </div> </div>
1	C	303	<div> <div>7%</div> <div> <div></div> <div>76%</div> <div>12%</div> <div>• 10%</div> </div> </div>
2	B	262	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>14%</div> </div> </div>
3	D	262	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>•</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9459 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 CYCLIN-DEPENDENT KINASE 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	300	Total	C	N	O	P	S	0	6	0
			2444	1585	417	432	1	9			
1	C	273	Total	C	N	O	P	S	0	4	0
			2216	1436	381	391	1	7			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	EXPRESSION TAG	UNP P24941
A	-3	PRO	-	EXPRESSION TAG	UNP P24941
A	-2	PRO	-	EXPRESSION TAG	UNP P24941
A	-1	GLY	-	EXPRESSION TAG	UNP P24941
A	0	SER	-	EXPRESSION TAG	UNP P24941
C	-4	GLY	-	EXPRESSION TAG	UNP P24941
C	-3	PRO	-	EXPRESSION TAG	UNP P24941
C	-2	PRO	-	EXPRESSION TAG	UNP P24941
C	-1	GLY	-	EXPRESSION TAG	UNP P24941
C	0	SER	-	EXPRESSION TAG	UNP P24941

- Molecule 2 is a protein called CYCLIN-A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	262	Total	C	N	O	S	0	3	0
			2129	1379	346	394	10			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	171	GLY	-	EXPRESSION TAG	UNP P20248
B	303	ALA	THR	CONFLICT	UNP P20248
B	311	ILE	VAL	CONFLICT	UNP P20248
B	357	ALA	GLY	CONFLICT	UNP P20248

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	377	VAL	ILE	CONFLICT	UNP P20248
B	378	GLN	ARG	CONFLICT	UNP P20248
B	386	THR	SER	CONFLICT	UNP P20248
B	392	LEU	MET	CONFLICT	UNP P20248
B	400	ARG	LYS	CONFLICT	UNP P20248

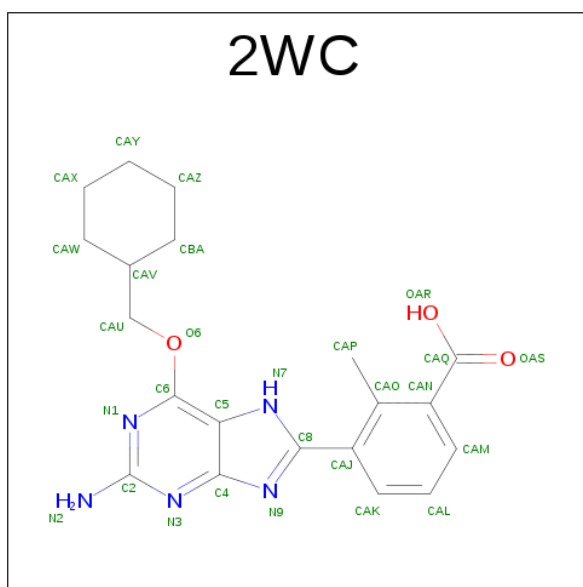
- Molecule 3 is a protein called CYCLIN-A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	262	Total	C	N	O	S	0	1	0
			2123	1373	345	395	10			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	171	ASN	-	EXPRESSION TAG	UNP P20248
D	172	GLU	-	EXPRESSION TAG	UNP P20248
D	303	ALA	THR	CONFLICT	UNP P20248
D	311	ILE	VAL	CONFLICT	UNP P20248
D	357	ALA	GLY	CONFLICT	UNP P20248
D	377	VAL	ILE	CONFLICT	UNP P20248
D	378	GLN	ARG	CONFLICT	UNP P20248
D	386	THR	SER	CONFLICT	UNP P20248
D	392	LEU	MET	CONFLICT	UNP P20248
D	400	ARG	LYS	CONFLICT	UNP P20248

- Molecule 4 is 3-[2-AZANYL-6-(CYCLOHEXYLMETHOXY)-7H-PURIN-8-YL]-2-METHYL-BENZOIC ACID (three-letter code: 2WC) (formula: C<sub>20</sub>H<sub>23</sub>N<sub>5</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 28	C 20	N 5	O 3	0	0
4	C	1	Total 28	C 20	N 5	O 3	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mg 1 1	0	0
5	D	1	Total Mg 1 1	0	0

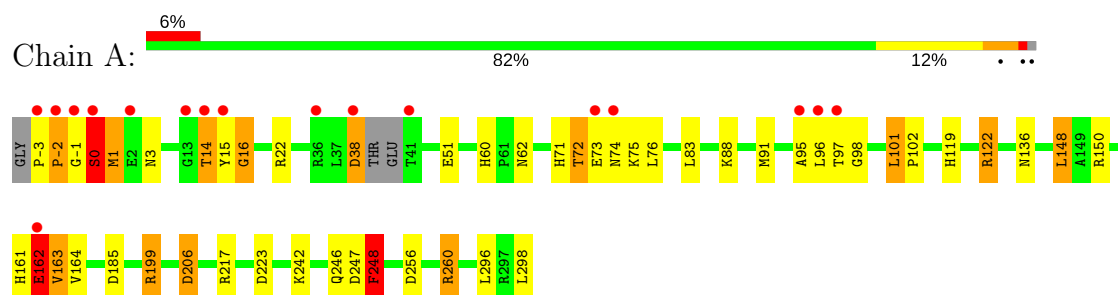
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	155	Total O 155 155	0	0
6	B	131	Total O 131 131	0	0
6	C	112	Total O 112 112	0	0
6	D	91	Total O 91 91	0	0

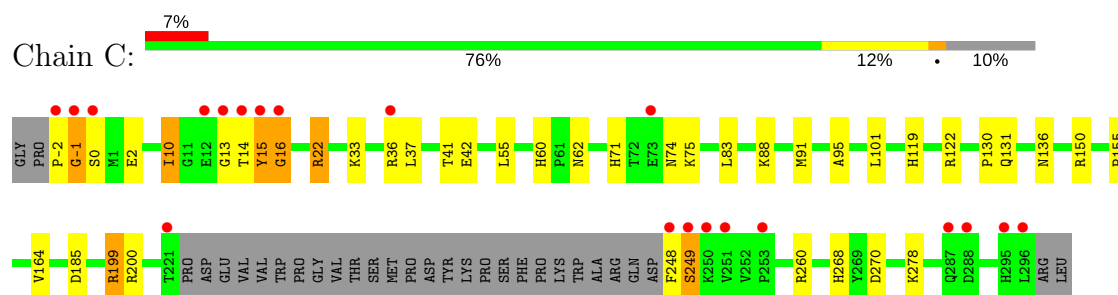
### 3 Residue-property plots

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.

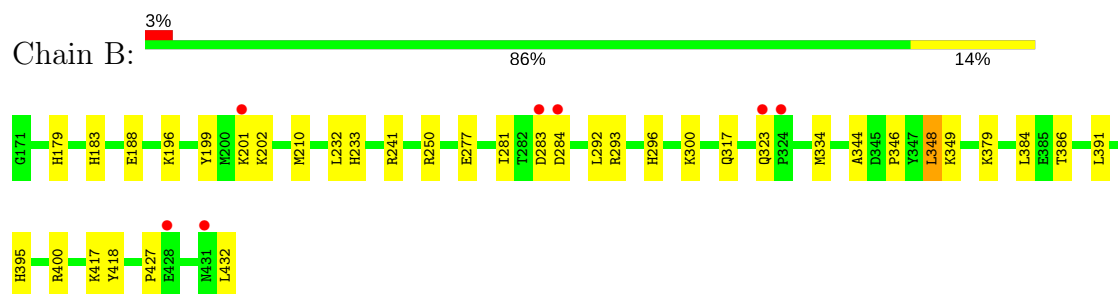
#### • Molecule 1: CYCLIN-DEPENDENT KINASE 2



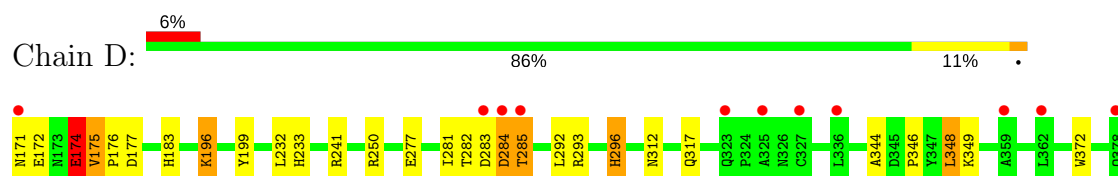
#### • Molecule 1: CYCLIN-DEPENDENT KINASE 2

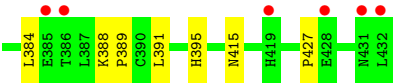


#### • Molecule 2: CYCLIN-A2



#### • Molecule 3: CYCLIN-A2





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.81Å 133.74Å 147.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.82 – 2.20 19.82 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.4 (19.82-2.20) 99.5 (19.82-2.20)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.25 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.182 , 0.221 0.189 , 0.224	Depositor DCC
$R_{free}$ test set	3750 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.0	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 38.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9459	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.95% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: TPO, MG, 2WC

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	1.05	5/2506 (0.2%)	1.08	16/3396 (0.5%)
1	C	0.87	0/2261	0.92	3/3057 (0.1%)
2	B	0.92	1/2189 (0.0%)	0.95	3/2976 (0.1%)
3	D	0.85	0/2176	0.95	5/2958 (0.2%)
All	All	0.93	6/9132 (0.1%)	0.98	27/12387 (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	1
1	C	0	1
3	D	0	2
All	All	0	4

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	51	GLU	CD-OE2	12.43	1.39	1.25
2	B	188	GLU	CD-OE1	6.50	1.32	1.25
1	A	256	ASP	CB-CG	6.27	1.65	1.51
1	A	162	GLU	C-O	6.22	1.35	1.23
1	A	0	SER	CB-OG	5.94	1.50	1.42
1	A	248	PHE	CG-CD2	-5.12	1.31	1.38

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	217	ARG	NE-CZ-NH2	-11.31	114.65	120.30
2	B	334	MET	CG-SD-CE	-10.37	83.60	100.20
1	C	199	ARG	NE-CZ-NH2	-8.06	116.27	120.30
1	A	217	ARG	NE-CZ-NH1	8.01	124.31	120.30
3	D	175	VAL	N-CA-C	-7.91	89.64	111.00
3	D	241	ARG	NE-CZ-NH2	-7.04	116.78	120.30
3	D	250	ARG	NE-CZ-NH2	-7.00	116.80	120.30
1	C	199	ARG	NE-CZ-NH1	6.94	123.77	120.30
1	A	223	ASP	CB-CG-OD1	6.82	124.44	118.30
2	B	241	ARG	NE-CZ-NH2	-6.79	116.90	120.30
1	A	247[A]	ASP	CB-CG-OD1	6.43	124.09	118.30
1	A	247[B]	ASP	CB-CG-OD1	6.43	124.09	118.30
1	A	199	ARG	NE-CZ-NH1	5.99	123.30	120.30
1	A	38	ASP	CB-CG-OD2	-5.84	113.05	118.30
3	D	241	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	A	256	ASP	CB-CG-OD1	5.53	123.28	118.30
1	C	10	ILE	CB-CA-C	-5.49	100.63	111.60
1	A	260	ARG	NE-CZ-NH1	5.45	123.03	120.30
1	A	122	ARG	NE-CZ-NH1	5.42	123.01	120.30
1	A	162	GLU	C-N-CA	5.36	135.09	121.70
1	A	162	GLU	CB-CA-C	5.30	121.01	110.40
2	B	241	ARG	NE-CZ-NH1	5.29	122.94	120.30
1	A	206[A]	ASP	CB-CG-OD2	-5.24	113.58	118.30
1	A	206[B]	ASP	CB-CG-OD2	-5.24	113.58	118.30
1	A	217	ARG	CG-CD-NE	-5.20	100.88	111.80
3	D	174	GLU	C-N-CA	5.18	134.65	121.70
1	A	223	ASP	CB-CG-OD2	-5.02	113.78	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	162	GLU	Peptide
1	C	-1	GLY	Peptide
3	D	174	GLU	Peptide
3	D	284	ASP	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	2444	0	2490	47	0
1	C	2216	0	2282	45	0
2	B	2129	0	2147	29	0
3	D	2123	0	2136	18	0
4	A	28	0	22	0	0
4	C	28	0	22	2	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	155	0	0	17	1
6	B	131	0	0	9	0
6	C	112	0	0	10	1
6	D	91	0	0	2	0
All	All	9459	0	9099	126	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.

All (126) 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:260[B]:ARG:HG2	1:C:260[B]:ARG:HH11	1.38	0.88
1:A:161:HIS:O	1:A:162:GLU:C	2.12	0.87
1:C:33:LYS:HD3	6:C:2032:HOH:O	1.73	0.86
1:A:71:HIS:NE2	2:B:296:HIS:CE1	2.46	0.84
1:A:22[B]:ARG:NH1	6:A:2009:HOH:O	2.12	0.82
2:B:210:MET:HE1	2:B:250:ARG:CB	2.13	0.79
1:C:268:HIS:HE1	6:C:2107:HOH:O	1.64	0.79
1:C:60:HIS:HD2	1:C:62:ASN:H	1.32	0.78
1:A:72:THR:OG1	6:A:2037:HOH:O	2.00	0.77
1:A:60:HIS:HD2	1:A:62:ASN:H	1.32	0.77
2:B:183:HIS:HD2	6:B:2018:HOH:O	1.68	0.75
1:A:246:GLN:O	6:A:2111:HOH:O	2.05	0.75
1:A:95:ALA:O	1:A:96:LEU:HB2	1.87	0.73
1:C:15:TYR:HB2	1:C:16:GLY:HA2	1.70	0.73
6:A:2035:HOH:O	2:B:300:LYS:NZ	2.20	0.73
1:A:15:TYR:N	1:A:16:GLY:CA	2.53	0.71
1:C:260[B]:ARG:NH1	1:C:260[B]:ARG:HG2	2.07	0.70
1:A:-2:PRO:O	1:A:0:SER:HB2	1.92	0.69
1:A:-2:PRO:HD2	6:A:2001:HOH:O	1.93	0.68
1:A:60:HIS:CD2	1:A:62:ASN:H	2.11	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:71:HIS:HD2	3:D:296:HIS:HE1	1.40	0.68
1:A:161:HIS:HD2	6:A:2094:HOH:O	1.77	0.67
1:C:60:HIS:CD2	1:C:62:ASN:H	2.13	0.66
2:B:183:HIS:HB2	2:B:317:GLN:HE22	1.60	0.66
1:A:0:SER:O	1:A:1:MET:HB2	1.96	0.65
2:B:210:MET:HE1	2:B:250:ARG:HB2	1.79	0.65
1:A:97:THR:HB	6:A:2045:HOH:O	1.96	0.65
1:C:15:TYR:HB2	1:C:16:GLY:CA	2.26	0.65
1:C:74:ASN:ND2	6:C:2023:HOH:O	2.30	0.65
1:C:71:HIS:HD2	3:D:296:HIS:CE1	2.16	0.63
1:C:13:GLY:HA3	1:C:15:TYR:CE2	2.36	0.60
1:C:15:TYR:CD2	1:C:16:GLY:HA3	2.36	0.60
1:C:41:THR:HB	6:C:2020:HOH:O	2.00	0.60
1:C:15:TYR:CB	1:C:16:GLY:CA	2.79	0.60
2:B:233:HIS:HD2	6:B:2077:HOH:O	1.83	0.59
1:A:162:GLU:HG3	1:A:163:VAL:HA	1.84	0.59
2:B:210:MET:HE1	2:B:250:ARG:HB3	1.82	0.59
2:B:344:ALA:HB1	2:B:348:LEU:HD22	1.84	0.58
1:A:260:ARG:HD3	6:A:2131:HOH:O	2.03	0.58
1:A:95:ALA:HA	1:A:199:ARG:HH11	1.68	0.58
3:D:344:ALA:HB1	3:D:348:LEU:HD22	1.85	0.58
3:D:183:HIS:HB2	3:D:317:GLN:HE22	1.68	0.57
1:A:162:GLU:CG	1:A:163:VAL:HA	2.34	0.57
1:C:119:HIS:HD2	6:C:2061:HOH:O	1.86	0.56
1:A:15:TYR:N	1:A:16:GLY:HA3	2.18	0.56
2:B:284:ASP:HB2	6:B:2068:HOH:O	2.06	0.56
1:A:15:TYR:N	1:A:16:GLY:HA2	2.22	0.54
3:D:196:LYS:HE2	6:D:2041:HOH:O	2.08	0.54
2:B:179[A]:HIS:CE1	6:B:2013:HOH:O	2.62	0.53
6:C:2044:HOH:O	3:D:296:HIS:CD2	2.63	0.52
1:A:0:SER:O	1:A:1:MET:CB	2.57	0.52
1:A:74:ASN:HB2	6:A:2038:HOH:O	2.09	0.52
2:B:432:LEU:O	6:B:2131:HOH:O	2.19	0.52
1:A:71:HIS:CD2	2:B:296:HIS:CE1	2.98	0.51
1:A:14:THR:C	1:A:16:GLY:HA3	2.30	0.51
1:A:248:PHE:O	6:A:2130:HOH:O	2.18	0.51
2:B:395:HIS:HE1	2:B:427:PRO:O	1.94	0.51
1:A:73:GLU:OE1	1:C:2:GLU:OE1	2.29	0.51
3:D:233:HIS:HD2	6:D:2062:HOH:O	1.94	0.50
1:C:15:TYR:HD2	1:C:16:GLY:CA	2.24	0.50
2:B:346:PRO:O	2:B:349:LYS:HG2	2.11	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:95:ALA:O	1:C:199:ARG:NH1	2.44	0.49
1:C:36:ARG:HD3	6:C:2017:HOH:O	2.11	0.49
3:D:346:PRO:O	3:D:349:LYS:HG3	2.12	0.49
1:A:119:HIS:HD2	6:A:2056:HOH:O	1.95	0.48
1:A:38:ASP:CB	6:A:2019:HOH:O	2.61	0.48
1:A:83:LEU:HD23	1:A:136:ASN:HB3	1.95	0.48
1:C:15:TYR:CD2	1:C:16:GLY:CA	2.96	0.48
1:C:83:LEU:HD23	1:C:136:ASN:HB3	1.95	0.48
1:C:15:TYR:HE2	6:C:2008:HOH:O	1.97	0.48
1:A:161:HIS:O	1:A:162:GLU:O	2.32	0.48
1:A:88:LYS:HA	1:A:91:MET:HE2	1.96	0.48
1:C:278:LYS:HE3	3:D:177:ASP:O	2.13	0.48
1:A:-3:PRO:O	1:A:-1:GLY:N	2.44	0.47
2:B:400:ARG:NH1	6:B:2114:HOH:O	2.47	0.47
4:C:1297:2WC:OAS	4:C:1297:2WC:HAP1	2.13	0.47
3:D:282:THR:O	3:D:285:THR:OG1	2.29	0.47
1:C:268:HIS:HD2	1:C:270:ASP:H	1.61	0.47
1:C:260[B]:ARG:CG	1:C:260[B]:ARG:NH1	2.73	0.47
1:A:60:HIS:HE1	6:A:2018:HOH:O	1.98	0.46
1:A:71:HIS:CD2	1:A:76:LEU:HD13	2.50	0.46
1:C:95:ALA:HA	1:C:199:ARG:HD2	1.97	0.46
1:C:88:LYS:HA	1:C:91:MET:HE2	1.97	0.46
3:D:388:LYS:HB3	3:D:389:PRO:HD3	1.98	0.46
1:A:119:HIS:HE1	1:A:185:ASP:OD2	1.99	0.45
2:B:210:MET:CE	2:B:250:ARG:HB2	2.47	0.45
1:C:10:ILE:O	1:C:10:ILE:CG2	2.63	0.45
1:C:200:ARG:HG2	1:C:200:ARG:HH11	1.82	0.45
1:A:97:THR:CB	6:A:2045:HOH:O	2.59	0.45
3:D:372:TRP:HB3	3:D:384:LEU:HD11	1.98	0.45
1:A:98:GLY:HA2	1:A:199:ARG:HD3	1.99	0.45
1:C:22[A]:ARG:HD3	6:C:2012:HOH:O	2.16	0.44
1:C:83:LEU:O	4:C:1297:2WC:N9	2.51	0.44
1:C:119:HIS:HE1	1:C:185:ASP:OD2	2.01	0.44
2:B:300:LYS:HD3	1:C:-2:PRO:HG2	1.99	0.43
1:A:161:HIS:CD2	6:A:2094:HOH:O	2.62	0.43
1:A:101:LEU:N	1:A:102:PRO:CD	2.81	0.43
1:A:3:ASN:OD1	3:D:293:ARG:NH1	2.52	0.43
1:C:-2:PRO:N	1:C:-1:GLY:HA3	2.34	0.42
2:B:179[A]:HIS:CD2	2:B:379:LYS:HZ1	2.37	0.42
1:C:36:ARG:NH1	1:C:75:LYS:HE3	2.33	0.42
3:D:171:ASN:O	3:D:176:PRO:HD3	2.18	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:161:HIS:O	1:A:163:VAL:N	2.52	0.42
1:C:155:PRO:HB3	3:D:174:GLU:HB3	2.01	0.42
2:B:277:GLU:O	2:B:281:ILE:HG13	2.20	0.42
1:A:71:HIS:CE1	2:B:296:HIS:CE1	3.05	0.42
1:C:41:THR:HG22	1:C:42:GLU:N	2.35	0.42
1:A:148:LEU:HD23	6:A:2061:HOH:O	2.20	0.42
2:B:386:THR:HB	6:B:2112:HOH:O	2.20	0.42
1:C:278:LYS:CE	3:D:177:ASP:O	2.68	0.42
3:D:395:HIS:HE1	3:D:427:PRO:O	2.02	0.41
1:C:249:SER:N	6:C:2102:HOH:O	2.42	0.41
1:A:242:LYS:HE3	6:A:2119:HOH:O	2.19	0.41
3:D:277:GLU:O	3:D:281:ILE:HG13	2.21	0.41
1:C:249:SER:HA	1:C:260[B]:ARG:NH1	2.36	0.41
2:B:300:LYS:HB3	1:C:-2:PRO:HG2	2.03	0.41
2:B:183:HIS:HE1	6:B:2017:HOH:O	2.04	0.41
1:A:14:THR:HG22	1:A:15:TYR:HA	2.02	0.41
1:A:71:HIS:HE2	2:B:296:HIS:CE1	2.33	0.40
1:A:88:LYS:HA	1:A:91:MET:CE	2.51	0.40
2:B:202:LYS:HB3	2:B:202:LYS:HE3	1.89	0.40
1:C:88:LYS:HB2	1:C:130:PRO:HB2	2.04	0.40
2:B:417:LYS:HE3	2:B:418:TYR:OH	2.21	0.40
2:B:300:LYS:HB3	1:C:-2:PRO:CG	2.51	0.40
2:B:284:ASP:CB	6:B:2068:HOH:O	2.66	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)
6:A:2153:HOH:O	6:C:2099:HOH:O[3_454]	1.84	0.36

## 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	301/303 (99%)	291 (97%)	5 (2%)	5 (2%)	11	7
1	C	272/303 (90%)	260 (96%)	8 (3%)	4 (2%)	12	9
2	B	263/262 (100%)	260 (99%)	3 (1%)	0	100	100
3	D	261/262 (100%)	255 (98%)	6 (2%)	0	100	100
All	All	1097/1130 (97%)	1066 (97%)	22 (2%)	9 (1%)	25	21

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	162	GLU
1	A	164	VAL
1	C	0[A]	SER
1	C	0[B]	SER
1	A	-2	PRO
1	C	16	GLY
1	C	164	VAL
1	A	16	GLY

### 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	268/265 (101%)	254 (95%)	14 (5%)	27	32
1	C	242/265 (91%)	230 (95%)	12 (5%)	28	34
2	B	237/234 (101%)	225 (95%)	12 (5%)	28	33
3	D	236/235 (100%)	222 (94%)	14 (6%)	23	26
All	All	983/999 (98%)	931 (95%)	52 (5%)	28	31

All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	SER
1	A	14	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	72	THR
1	A	75	LYS
1	A	101	LEU
1	A	122	ARG
1	A	148	LEU
1	A	150	ARG
1	A	163	VAL
1	A	206[A]	ASP
1	A	206[B]	ASP
1	A	248	PHE
1	A	296	LEU
1	A	298	LEU
2	B	196	LYS
2	B	199	TYR
2	B	201	LYS
2	B	232	LEU
2	B	283[A]	ASP
2	B	283[B]	ASP
2	B	292	LEU
2	B	293	ARG
2	B	323	GLN
2	B	348	LEU
2	B	384	LEU
2	B	391	LEU
1	C	14	THR
1	C	15	TYR
1	C	22[A]	ARG
1	C	22[B]	ARG
1	C	37	LEU
1	C	55	LEU
1	C	101	LEU
1	C	122	ARG
1	C	131	GLN
1	C	150	ARG
1	C	248	PHE
1	C	249	SER
3	D	172	GLU
3	D	175	VAL
3	D	196	LYS
3	D	199	TYR
3	D	232	LEU
3	D	283	ASP

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
3	D	284	ASP
3	D	285	THR
3	D	292	LEU
3	D	296	HIS
3	D	312	ASN
3	D	348	LEU
3	D	391	LEU
3	D	415	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (29) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	60	HIS
1	A	119	HIS
1	A	161	HIS
1	A	268	HIS
2	B	183	HIS
2	B	233	HIS
2	B	254	GLN
2	B	296	HIS
2	B	317	GLN
2	B	395	HIS
2	B	396	GLN
2	B	403	GLN
2	B	431	ASN
1	C	60	HIS
1	C	71	HIS
1	C	84	HIS
1	C	85	GLN
1	C	119	HIS
1	C	268	HIS
1	C	295	HIS
3	D	173	ASN
3	D	183	HIS
3	D	233	HIS
3	D	296	HIS
3	D	317	GLN
3	D	395	HIS
3	D	396	GLN
3	D	425	ASN
3	D	431	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	TPO	A	160	1	9,10,11	1.03	0	10,14,16	1.16	0
1	TPO	C	160	1	9,10,11	1.60	3 (33%)	10,14,16	1.06	1 (10%)

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	TPO	A	160	1	-	0/8/11/13	0/0/0/0
1	TPO	C	160	1	-	0/8/11/13	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	160	TPO	CG2-CB	2.04	1.56	1.51
1	C	160	TPO	P-OG1	2.34	1.63	1.59
1	C	160	TPO	CA-C	2.68	1.53	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	160	TPO	O3P-P-O2P	2.56	117.96	107.61

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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	2WC	A	1299	-	27,31,31	2.01	6 (22%)	30,44,44	2.03	7 (23%)
4	2WC	C	1297	-	27,31,31	2.21	8 (29%)	30,44,44	2.36	12 (40%)

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	2WC	A	1299	-	-	0/9/21/21	0/4/4/4
4	2WC	C	1297	-	-	0/9/21/21	0/4/4/4

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	1297	2WC	CAJ-C8	-6.21	1.36	1.48
4	A	1299	2WC	CAJ-C8	-6.20	1.36	1.48
4	C	1297	2WC	CAP-CAO	-3.80	1.43	1.51
4	A	1299	2WC	CAP-CAO	-3.69	1.44	1.51
4	A	1299	2WC	CAJ-CAO	2.22	1.44	1.40
4	C	1297	2WC	O6-CAU	2.36	1.51	1.44
4	C	1297	2WC	CAM-CAN	2.54	1.44	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	1297	2WC	CAZ-CBA	2.67	1.60	1.53
4	A	1299	2WC	C6-N1	2.73	1.36	1.31
4	C	1297	2WC	C6-N1	2.78	1.36	1.31
4	C	1297	2WC	CAN-CAO	2.82	1.45	1.40
4	A	1299	2WC	CAZ-CBA	2.91	1.60	1.53
4	A	1299	2WC	O6-C6	4.00	1.37	1.35
4	C	1297	2WC	O6-C6	5.06	1.38	1.35

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1297	2WC	CAX-CAW-CAV	-5.53	102.04	112.19
4	C	1297	2WC	N3-C2-N1	-5.33	119.67	127.46
4	A	1299	2WC	CAZ-CBA-CAV	-5.08	102.87	112.19
4	A	1299	2WC	O6-CAU-CAV	-4.42	97.70	107.81
4	A	1299	2WC	N3-C2-N1	-4.33	121.13	127.46
4	C	1297	2WC	CAY-CAX-CAW	-3.12	104.97	111.42
4	A	1299	2WC	CAX-CAW-CAV	-3.05	106.60	112.19
4	C	1297	2WC	O6-CAU-CAV	-3.00	100.95	107.81
4	C	1297	2WC	CAP-CAO-CAJ	-2.75	118.37	121.78
4	C	1297	2WC	CAP-CAO-CAN	2.09	124.38	121.78
4	C	1297	2WC	CAJ-C8-N7	2.20	126.44	123.71
4	C	1297	2WC	CAL-CAK-CAJ	2.35	124.36	120.37
4	A	1299	2WC	O6-C6-N1	2.59	122.90	120.23
4	A	1299	2WC	C2-N3-C4	2.72	118.34	115.16
4	C	1297	2WC	C2-N1-C6	2.96	120.45	116.18
4	C	1297	2WC	N2-C2-N3	3.45	124.12	117.75
4	C	1297	2WC	C2-N3-C4	3.67	119.44	115.16
4	C	1297	2WC	CAM-CAN-CAO	3.71	123.57	119.94
4	A	1299	2WC	CAJ-C8-N7	4.05	128.74	123.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1297	2WC	2	0

## 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	299/303 (98%)	0.00	17 (5%) 24 24	18, 28, 50, 75	0
1	C	272/303 (89%)	0.29	20 (7%) 15 14	25, 37, 55, 73	0
2	B	262/262 (100%)	-0.14	7 (2%) 55 52	21, 31, 45, 58	0
3	D	262/262 (100%)	0.20	17 (6%) 20 18	21, 39, 72, 90	0
All	All	1095/1130 (96%)	0.09	61 (5%) 25 24	18, 33, 60, 90	0

All (61) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	13	GLY	8.0
1	C	14	THR	7.4
3	D	171	ASN	7.3
1	A	-2	PRO	6.5
1	C	15	TYR	6.2
1	A	15	TYR	6.1
1	C	251	VAL	5.9
1	C	250	LYS	5.5
1	A	96	LEU	5.5
1	A	41	THR	5.4
1	C	248	PHE	5.3
1	C	221	THR	5.1
3	D	283	ASP	5.0
3	D	284	ASP	4.9
1	C	12	GLU	4.9
3	D	323	GLN	4.8
1	C	249	SER	4.6
1	A	36	ARG	4.6
2	B	284	ASP	4.5
3	D	428	GLU	4.5
3	D	385	GLU	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	-2	PRO	4.2
2	B	323	GLN	4.2
1	A	13	GLY	4.2
3	D	431	ASN	4.1
2	B	283[A]	ASP	4.1
1	C	295	HIS	4.0
1	A	38	ASP	3.9
1	C	253	PRO	3.8
1	A	14	THR	3.7
1	A	73	GLU	3.7
3	D	336	LEU	3.4
2	B	324	PRO	3.3
3	D	419	HIS	3.3
1	A	97	THR	3.3
3	D	325	ALA	3.2
2	B	201	LYS	3.1
1	A	162	GLU	3.1
1	A	74	ASN	3.0
1	C	296	LEU	2.9
1	A	-1	GLY	2.8
2	B	431	ASN	2.8
1	C	287	GLN	2.7
1	A	2	GLU	2.7
1	A	95	ALA	2.7
3	D	359	ALA	2.6
2	B	428	GLU	2.5
1	C	-1	GLY	2.5
3	D	432	LEU	2.5
1	C	288	ASP	2.4
1	C	16	GLY	2.4
3	D	285	THR	2.4
3	D	386	THR	2.4
3	D	378	GLN	2.3
1	A	-3	PRO	2.3
1	C	36	ARG	2.3
3	D	362	LEU	2.3
1	A	0	SER	2.3
1	C	0[A]	SER	2.3
3	D	327	CYS	2.3
1	C	73	GLU	2.2

## 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	TPO	A	160	11/12	0.99	0.07	-	25,27,28,29	0
1	TPO	C	160	11/12	0.98	0.06	-	29,30,33,34	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(Å <sup>2</sup> )	Q<0.9
4	2WC	A	1299	28/28	0.91	0.14	0.94	27,36,41,42	0
5	MG	A	1300	1/1	0.96	0.14	0.19	34,34,34,34	0
5	MG	D	1433	1/1	0.97	0.15	0.17	32,32,32,32	0
4	2WC	C	1297	28/28	0.81	0.19	0.15	34,39,45,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.