



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 11:36 am BST

PDB ID : 6MU4  
Title : Bst DNA polymerase I FANA/DNA binary complex  
Authors : Jackson, L.N.; Chim, N.; Chaput, J.C.  
Deposited on : 2018-10-22  
Resolution : 1.62 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

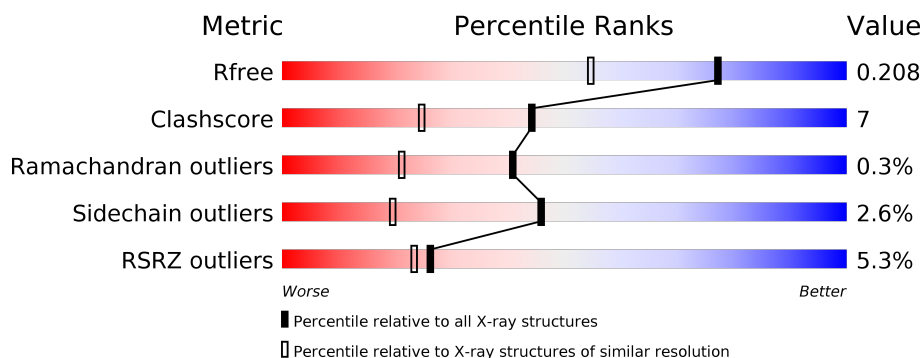
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.62 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))
$R_{free}$	130704	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (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 respectively. 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	P	11	<div> <div>36%</div> <div> <div>55%</div> <div>36%</div> <div>9%</div> </div> </div>
2	A	577	<div> <div>5%</div> <div>87%</div> <div>11%</div> </div>
3	T	13	<div> <div>38%</div> <div>62%</div> </div>

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
3	A5L	T	16	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5756 atoms, of which 42 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 DNA chain called DNA (5'-D(P\*GP\*CP\*GP\*AP\*TP\*CP\*AP\*CP\*GP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	11	Total	C	N	O	P	0	0	0
			225	107	40	67	11			

- Molecule 2 is a protein called DNA polymerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	577	Total	C	N	O	S	0	0	0
			4632	2944	807	865	16			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	550	THR	SER	conflict	UNP E1C9K5

- Molecule 3 is a DNA chain called FANA (5'-D(P\*(UF2)P\*(A5L)P\*(CFL)P\*(GFL)P\*(UF2)P\*(GFL)P\*(A5L)P\*(UF2)P\*(CFL)P\*(GFL)P\*(CFL))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	T	13	Total	C	F	N	O	P	0	0
			278	124	13	50	78	13		

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			22	6	14	2		
4	A	1	Total	C	H	O	0	0
			22	6	14	2		
4	A	1	Total	C	H	O	0	0
			22	6	14	2		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	0
			1	1		

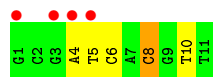
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	P	7	Total	O	0	0
			7	7		
7	A	513	Total	O	0	0
			513	513		
7	T	24	Total	O	0	0
			24	24		

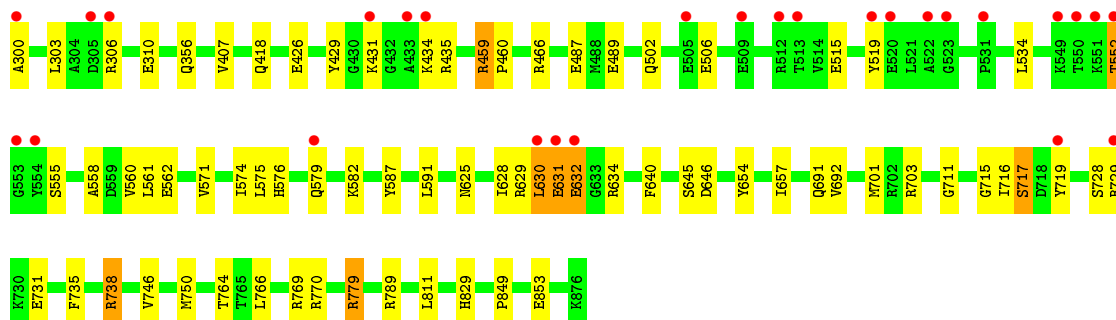
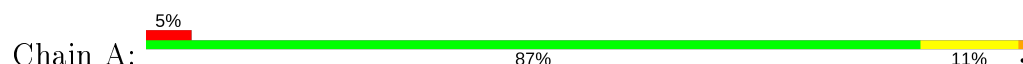
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: DNA (5'-D(P\*GP\*CP\*GP\*AP\*TP\*CP\*AP\*CP\*GP\*T)-3')



- Molecule 2: DNA polymerase I



- Molecule 3: FANA (5'-D(P\*(UF2)P\*(A5L)P\*(CFL)P\*(GFL)P\*(UF2)P\*(GFL)P\*(A5L)P\*(UF2)P\*(CFL)P\*(GFL)P\*(CFL))-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.83Å 93.90Å 105.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.09 – 1.62 46.09 – 1.62	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.09-1.62) 94.6 (46.09-1.62)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 1.62Å)	Xtriage
Refinement program	PHENIX (1.13 _2998: ???)	Depositor
R, $R_{free}$	0.187 , 0.210 0.190 , 0.208	Depositor DCC
$R_{free}$ test set	2004 reflections (1.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.4	Xtriage
Anisotropy	0.480	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5756	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.10% 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: MG, MPD, A5L, CFL, SO4, GFL, UF2

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	P	0.96	1/251 (0.4%)	0.80	0/385
2	A	0.46	3/4715 (0.1%)	0.57	3/6373 (0.0%)
All	All	0.50	4/4966 (0.1%)	0.59	3/6758 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	489	GLU	CD-OE2	-7.73	1.17	1.25
2	A	310	GLU	CD-OE2	-7.45	1.17	1.25
1	P	8	DC	O3'-P	-6.17	1.53	1.61
2	A	489	GLU	CD-OE1	-5.18	1.20	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	738	ARG	NE-CZ-NH2	5.25	122.92	120.30
2	A	738	ARG	NE-CZ-NH1	-5.17	117.72	120.30
2	A	717	SER	CB-CA-C	-5.13	100.35	110.10

There are no chirality outliers.

There are no planarity outliers.

### 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	P	225	0	125	7	0
2	A	4632	0	4684	56	0
3	T	278	0	127	7	0
4	A	24	42	42	7	0
5	A	10	0	0	0	0
6	A	1	0	0	0	0
7	A	513	0	0	7	0
7	P	7	0	0	0	0
7	T	24	0	0	1	0
All	All	5714	42	4978	66	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 7.

All (66) 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:P:6:DC:OP1	2:A:552:THR:OG1	1.53	1.25
3:T:7:CFL:H2'	3:T:8:GFL:H5'2	1.36	1.03
2:A:459:ARG:HH21	2:A:466:ARG:HH22	1.14	0.95
2:A:431:LYS:H	2:A:434:LYS:HE3	1.32	0.91
2:A:789:ARG:HD2	4:A:903:MPD:H52	1.54	0.88
2:A:426:GLU:HG3	2:A:431:LYS:HE2	1.59	0.82
1:P:4:DA:H2"	1:P:5:DT:H5"	1.66	0.75
2:A:719:TYR:CE2	2:A:729:ARG:HD3	2.24	0.73
2:A:692:VAL:HG21	2:A:701:MET:HE1	1.71	0.73
2:A:431:LYS:N	2:A:434:LYS:HE3	2.06	0.71
2:A:534:LEU:HD11	2:A:574:ILE:HD13	1.75	0.69
2:A:426:GLU:CG	2:A:431:LYS:HE2	2.24	0.67
2:A:426:GLU:HG3	2:A:431:LYS:HG2	1.81	0.62
2:A:628:ILE:HD11	2:A:829:HIS:CE1	2.36	0.61
1:P:4:DA:C2'	1:P:5:DT:H5"	2.31	0.61
1:P:8:DC:H4'	2:A:582:LYS:HD2	1.81	0.60
3:T:4:GFL:N3	3:T:4:GFL:H3'	2.17	0.60
2:A:692:VAL:HG21	2:A:701:MET:CE	2.32	0.60
2:A:738:ARG:CZ	7:A:1008:HOH:O	2.49	0.59
2:A:630:LEU:O	2:A:631:GLU:C	2.41	0.58
2:A:630:LEU:O	2:A:632:GLU:N	2.36	0.58
2:A:558:ALA:O	2:A:562:GLU:HG3	2.04	0.57
2:A:716:ILE:O	4:A:903:MPD:H31	2.05	0.56
3:T:14:GFL:H2'	3:T:15:CFL:O4'	2.06	0.55
2:A:502:GLN:O	2:A:506:GLU:HG3	2.05	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:555:SER:OG	2:A:560:VAL:HG21	2.07	0.55
2:A:300:ALA:N	4:A:902:MPD:HO4	2.06	0.54
1:P:4:DA:H2''	1:P:5:DT:C5'	2.35	0.53
2:A:629:ARG:CD	2:A:629:ARG:N	2.72	0.52
2:A:591:LEU:HD21	2:A:640:PHE:CZ	2.44	0.51
3:T:16:A5L:OP1	7:T:101:HOH:O	2.20	0.49
2:A:628:ILE:N	2:A:628:ILE:HD12	2.29	0.48
2:A:487:GLU:HG3	7:A:1221:HOH:O	2.13	0.47
2:A:628:ILE:HD11	2:A:829:HIS:NE2	2.29	0.47
2:A:728:SER:OG	2:A:731:GLU:HG3	2.14	0.47
2:A:561:LEU:HB2	2:A:575:LEU:HD21	1.95	0.47
2:A:716:ILE:O	4:A:903:MPD:C3	2.63	0.47
1:P:10:DT:H5''	2:A:628:ILE:HD13	1.97	0.46
2:A:789:ARG:HD2	4:A:903:MPD:C5	2.38	0.45
2:A:418:GLN:HA	7:A:1055:HOH:O	2.17	0.45
2:A:849:PRO:O	2:A:853:GLU:HG3	2.17	0.45
2:A:426:GLU:CG	2:A:431:LYS:HG2	2.46	0.45
2:A:587:TYR:O	2:A:591:LEU:HB2	2.16	0.44
2:A:561:LEU:O	2:A:571:VAL:HG11	2.18	0.44
4:A:901:MPD:H52	3:T:5:UF2:C6	2.48	0.44
2:A:459:ARG:HB3	2:A:460:PRO:HD3	1.98	0.44
2:A:429:TYR:O	2:A:435:ARG:HA	2.18	0.43
2:A:654:TYR:HB3	2:A:657:ILE:HB	2.00	0.43
3:T:14:GFL:H4'	3:T:15:CFL:O1P	2.19	0.43
2:A:426:GLU:OE2	2:A:431:LYS:HE2	2.19	0.43
2:A:576:HIS:O	2:A:579:GLN:HG2	2.19	0.43
2:A:738:ARG:NH2	7:A:1008:HOH:O	2.52	0.43
2:A:645:SER:O	2:A:646:ASP:HB2	2.18	0.42
2:A:764:THR:HA	2:A:769:ARG:O	2.19	0.42
2:A:770:ARG:HG3	7:A:1201:HOH:O	2.19	0.42
2:A:691:GLN:HG2	2:A:738:ARG:NH1	2.35	0.41
1:P:6:DC:P	2:A:552:THR:OG1	2.71	0.41
2:A:711:GLY:HA2	2:A:716:ILE:HG12	2.02	0.41
2:A:426:GLU:CD	2:A:431:LYS:HE2	2.41	0.41
2:A:703:ARG:HD3	7:A:1473:HOH:O	2.20	0.41
3:T:6:A5L:H2'	3:T:7:CFL:H6	2.02	0.41
2:A:779:ARG:NH1	7:A:1031:HOH:O	2.54	0.41
2:A:715:GLY:HA2	4:A:903:MPD:H53	2.02	0.41
2:A:515:GLU:HG3	2:A:519:TYR:CZ	2.56	0.40
2:A:746:VAL:O	2:A:750:MET:HG2	2.22	0.40
2:A:766:LEU:HA	2:A:766:LEU:HD23	1.89	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	A	575/577 (100%)	562 (98%)	11 (2%)	2 (0%)	41 21

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	631	GLU
2	A	407	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
2	A	493/494 (100%)	480 (97%)	13 (3%)	46 19

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	303	LEU
2	A	306	ARG
2	A	356	GLN
2	A	459	ARG
2	A	552	THR
2	A	625	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	A	630	LEU
2	A	632	GLU
2	A	634	ARG
2	A	717	SER
2	A	735	PHE
2	A	779	ARG
2	A	811	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

13 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	$\# Z  > 2$
3	CFL	T	13	1,3	15,21,22	5.14	11 (73%)	17,30,33	1.57	3 (17%)
3	A5L	T	16	3	18,24,25	6.48	8 (44%)	18,35,38	4.16	7 (38%)
3	UF2	T	5	1,3	14,21,22	8.10	12 (85%)	14,30,33	1.07	1 (7%)
3	GFL	T	4	3	18,25,26	6.93	14 (77%)	20,37,40	2.47	7 (35%)
3	A5L	T	11	1,3	18,24,25	6.29	9 (50%)	18,35,38	3.83	6 (33%)
3	CFL	T	7	1,3	15,21,22	5.18	12 (80%)	17,30,33	1.76	3 (17%)
3	CFL	T	15	1,3	15,21,22	5.15	11 (73%)	17,30,33	1.75	3 (17%)
3	UF2	T	9	1,3	14,21,22	7.83	12 (85%)	14,30,33	1.50	3 (21%)
3	GFL	T	10	1,3	18,25,26	6.73	14 (77%)	20,37,40	2.77	7 (35%)
3	A5L	T	6	1,3	18,24,25	6.39	8 (44%)	18,35,38	3.94	7 (38%)
3	UF2	T	12	1,3	14,21,22	7.82	12 (85%)	14,30,33	1.52	3 (21%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GFL	T	14	1,3	18,25,26	6.92	14 (77%)	20,37,40	2.99	10 (50%)
3	GFL	T	8	1,3	18,25,26	6.85	14 (77%)	20,37,40	2.79	7 (35%)

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	CFL	T	13	1,3	-	0/5/25/26	0/2/2/2
3	A5L	T	16	3	-	2/3/25/26	0/3/3/3
3	UF2	T	5	1,3	-	1/5/25/26	0/2/2/2
3	GFL	T	4	3	-	2/3/25/26	0/3/3/3
3	A5L	T	11	1,3	-	0/3/25/26	0/3/3/3
3	CFL	T	7	1,3	-	0/5/25/26	0/2/2/2
3	CFL	T	15	1,3	-	2/5/25/26	0/2/2/2
3	UF2	T	9	1,3	-	0/5/25/26	0/2/2/2
3	GFL	T	10	1,3	-	0/3/25/26	0/3/3/3
3	A5L	T	6	1,3	-	2/3/25/26	0/3/3/3
3	UF2	T	12	1,3	-	0/5/25/26	0/2/2/2
3	GFL	T	14	1,3	-	0/3/25/26	0/3/3/3
3	GFL	T	8	1,3	-	2/3/25/26	0/3/3/3

All (151) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	14	GFL	C2'-C3'	-19.26	1.26	1.52
3	T	4	GFL	C2'-C3'	-19.25	1.26	1.52
3	T	8	GFL	C2'-C3'	-19.07	1.26	1.52
3	T	10	GFL	C2'-C3'	-18.65	1.27	1.52
3	T	16	A5L	C2'-C1'	-18.04	1.30	1.53
3	T	5	UF2	C2'-C3'	-17.94	1.28	1.52
3	T	6	A5L	C2'-C1'	-17.07	1.31	1.53
3	T	12	UF2	C2'-C3'	-16.81	1.29	1.52
3	T	11	A5L	C2'-C1'	-16.75	1.31	1.53
3	T	9	UF2	C2'-C3'	-16.65	1.29	1.52
3	T	5	UF2	F2'-C2'	-14.59	1.08	1.40
3	T	9	UF2	F2'-C2'	-14.10	1.09	1.40
3	T	12	UF2	F2'-C2'	-14.01	1.10	1.40
3	T	11	A5L	C2'-C3'	12.48	1.69	1.52
3	T	6	A5L	C2'-C3'	12.18	1.69	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	16	A5L	C2'-C3'	12.11	1.69	1.52
3	T	4	GFL	C2'-C1'	11.04	1.67	1.53
3	T	6	A5L	C3'-C4'	-10.53	1.26	1.53
3	T	14	GFL	C2'-C1'	10.48	1.66	1.53
3	T	11	A5L	C3'-C4'	-10.36	1.26	1.53
3	T	16	A5L	C3'-C4'	-10.28	1.26	1.53
3	T	8	GFL	C2'-C1'	10.20	1.66	1.53
3	T	12	UF2	C2'-C1'	9.94	1.65	1.53
3	T	9	UF2	C2'-C1'	9.91	1.65	1.53
3	T	14	GFL	C4-N3	9.85	1.51	1.35
3	T	10	GFL	C4-N3	9.59	1.50	1.35
3	T	4	GFL	C4-N3	9.53	1.50	1.35
3	T	10	GFL	C2'-C1'	9.52	1.65	1.53
3	T	8	GFL	C4-N3	9.52	1.50	1.35
3	T	6	A5L	O4'-C1'	9.44	1.54	1.41
3	T	5	UF2	C2'-C1'	9.26	1.64	1.53
3	T	16	A5L	O4'-C1'	9.00	1.53	1.41
3	T	11	A5L	O4'-C1'	8.88	1.53	1.41
3	T	13	CFL	C3'-C4'	-8.72	1.30	1.53
3	T	15	CFL	C3'-C4'	-8.48	1.31	1.53
3	T	5	UF2	O4'-C1'	-8.45	1.29	1.41
3	T	7	CFL	C6-N1	8.40	1.46	1.35
3	T	7	CFL	C3'-C4'	-8.38	1.31	1.53
3	T	15	CFL	C6-N1	8.22	1.46	1.35
3	T	9	UF2	O4'-C1'	-8.09	1.29	1.41
3	T	13	CFL	C6-N1	8.08	1.45	1.35
3	T	12	UF2	O4'-C1'	-7.99	1.29	1.41
3	T	16	A5L	O4'-C4'	7.89	1.62	1.45
3	T	8	GFL	C2-N2	7.82	1.49	1.33
3	T	14	GFL	C2-N2	7.78	1.49	1.33
3	T	6	A5L	O4'-C4'	7.77	1.62	1.45
3	T	4	GFL	C2-N2	7.76	1.49	1.33
3	T	15	CFL	O4'-C4'	7.71	1.62	1.45
3	T	10	GFL	C2-N2	7.69	1.49	1.33
3	T	5	UF2	C4-N3	7.67	1.46	1.33
3	T	15	CFL	O4'-C1'	-7.56	1.30	1.41
3	T	7	CFL	O4'-C4'	7.55	1.61	1.45
3	T	5	UF2	C6-N1	7.55	1.45	1.35
3	T	9	UF2	C4-N3	7.53	1.46	1.33
3	T	7	CFL	O4'-C1'	-7.40	1.30	1.41
3	T	13	CFL	O4'-C1'	-7.39	1.30	1.41
3	T	12	UF2	C4-N3	7.37	1.45	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	11	A5L	O4'-C4'	7.35	1.61	1.45
3	T	9	UF2	C6-N1	7.32	1.44	1.35
3	T	12	UF2	C6-N1	7.30	1.44	1.35
3	T	13	CFL	O4'-C4'	7.28	1.61	1.45
3	T	4	GFL	C6-N1	7.04	1.45	1.33
3	T	8	GFL	C6-N1	6.89	1.45	1.33
3	T	10	GFL	C6-N1	6.88	1.45	1.33
3	T	10	GFL	O4'-C1'	-6.82	1.31	1.41
3	T	8	GFL	O4'-C1'	-6.80	1.31	1.41
3	T	14	GFL	C6-N1	6.80	1.44	1.33
3	T	10	GFL	C6-C5	6.71	1.52	1.41
3	T	8	GFL	C6-C5	6.46	1.52	1.41
3	T	4	GFL	O4'-C1'	-6.44	1.32	1.41
3	T	14	GFL	O4'-C1'	-6.40	1.32	1.41
3	T	14	GFL	C6-C5	6.35	1.52	1.41
3	T	4	GFL	C6-C5	6.33	1.52	1.41
3	T	5	UF2	C2-N3	6.28	1.50	1.38
3	T	7	CFL	C4-N3	6.17	1.45	1.35
3	T	9	UF2	C2-N3	6.11	1.50	1.38
3	T	13	CFL	C4-N3	6.07	1.45	1.35
3	T	12	UF2	C2-N3	6.03	1.50	1.38
3	T	15	CFL	C4-N3	5.93	1.45	1.35
3	T	7	CFL	C2-N3	5.56	1.49	1.38
3	T	13	CFL	C2-N3	5.44	1.48	1.38
3	T	15	CFL	C2-N3	5.31	1.48	1.38
3	T	8	GFL	C5'-C4'	-5.31	1.35	1.51
3	T	7	CFL	C6-C5	5.24	1.49	1.38
3	T	13	CFL	C6-C5	5.20	1.49	1.38
3	T	15	CFL	C6-C5	5.19	1.49	1.38
3	T	14	GFL	C5'-C4'	-5.05	1.35	1.51
3	T	4	GFL	C5'-C4'	-4.95	1.36	1.51
3	T	4	GFL	C2-N1	4.93	1.44	1.35
3	T	10	GFL	C5'-C4'	-4.85	1.36	1.51
3	T	8	GFL	C2-N1	4.85	1.44	1.35
3	T	10	GFL	C2-N1	4.78	1.43	1.35
3	T	5	UF2	C6-C5	4.72	1.48	1.38
3	T	14	GFL	C2-N1	4.71	1.43	1.35
3	T	9	UF2	C6-C5	4.63	1.48	1.38
3	T	12	UF2	C6-C5	4.55	1.48	1.38
3	T	5	UF2	O4'-C4'	3.97	1.53	1.45
3	T	9	UF2	O4'-C4'	3.96	1.53	1.45
3	T	12	UF2	O4'-C4'	3.91	1.53	1.45

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	5	UF2	C5'-C4'	-3.65	1.40	1.51
3	T	12	UF2	C5'-C4'	-3.59	1.40	1.51
3	T	9	UF2	C5'-C4'	-3.55	1.40	1.51
3	T	7	CFL	C4-N4	3.54	1.45	1.35
3	T	15	CFL	C4-N4	3.54	1.45	1.35
3	T	13	CFL	C4-N4	3.51	1.45	1.35
3	T	13	CFL	F-C2'	-3.40	1.32	1.40
3	T	7	CFL	C5-C4	3.39	1.49	1.41
3	T	13	CFL	C5-C4	3.39	1.49	1.41
3	T	14	GFL	O3'-C3'	3.38	1.50	1.43
3	T	15	CFL	F-C2'	-3.27	1.33	1.40
3	T	15	CFL	C5-C4	3.25	1.49	1.41
3	T	7	CFL	F-C2'	-3.19	1.33	1.40
3	T	16	A5L	C6-N6	3.16	1.45	1.34
3	T	5	UF2	O3'-C3'	3.15	1.50	1.43
3	T	9	UF2	O3'-C3'	3.14	1.50	1.43
3	T	12	UF2	O3'-C3'	3.13	1.50	1.43
3	T	6	A5L	C6-N6	3.13	1.45	1.34
3	T	11	A5L	C6-N6	3.13	1.45	1.34
3	T	14	GFL	O4'-C4'	3.03	1.51	1.45
3	T	4	GFL	O6-C6	-3.03	1.16	1.24
3	T	7	CFL	O3'-C3'	3.02	1.50	1.43
3	T	10	GFL	O6-C6	-3.02	1.17	1.24
3	T	14	GFL	O6-C6	-3.01	1.17	1.24
3	T	15	CFL	O3'-C3'	2.98	1.50	1.43
3	T	8	GFL	O6-C6	-2.98	1.17	1.24
3	T	4	GFL	O4'-C4'	2.94	1.51	1.45
3	T	10	GFL	O4'-C4'	2.90	1.51	1.45
3	T	13	CFL	O3'-C3'	2.87	1.49	1.43
3	T	10	GFL	C3'-C4'	2.72	1.59	1.53
3	T	16	A5L	C5-C4	-2.70	1.33	1.40
3	T	8	GFL	O4'-C4'	2.69	1.51	1.45
3	T	6	A5L	C5-C4	-2.64	1.33	1.40
3	T	11	A5L	C5-C4	-2.63	1.34	1.40
3	T	14	GFL	C2-N3	2.62	1.46	1.34
3	T	14	GFL	C3'-C4'	2.61	1.59	1.53
3	T	4	GFL	O3'-C3'	2.55	1.49	1.43
3	T	4	GFL	C3'-C4'	2.53	1.59	1.53
3	T	8	GFL	C2-N3	2.50	1.46	1.34
3	T	10	GFL	C2-N3	2.48	1.46	1.34
3	T	4	GFL	C2-N3	2.41	1.46	1.34
3	T	11	A5L	C2-N3	2.37	1.35	1.32

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	12	UF2	O4-C4	-2.29	1.18	1.24
3	T	8	GFL	O3'-C3'	2.27	1.48	1.43
3	T	10	GFL	O3'-C3'	2.26	1.48	1.43
3	T	6	A5L	C2-N3	2.24	1.35	1.32
3	T	16	A5L	C2-N3	2.21	1.35	1.32
3	T	5	UF2	O4-C4	-2.18	1.19	1.24
3	T	7	CFL	C2'-C3'	2.11	1.55	1.52
3	T	9	UF2	O4-C4	-2.08	1.19	1.24
3	T	8	GFL	C3'-C4'	2.08	1.58	1.53
3	T	11	A5L	O5'-C5'	-2.03	1.39	1.44

All (67) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	16	A5L	C5-C6-N6	12.38	139.16	120.35
3	T	6	A5L	C5-C6-N6	11.99	138.57	120.35
3	T	11	A5L	C5-C6-N6	11.55	137.90	120.35
3	T	16	A5L	N6-C6-N1	-8.64	100.63	118.57
3	T	6	A5L	N6-C6-N1	-8.24	101.46	118.57
3	T	11	A5L	N6-C6-N1	-8.08	101.81	118.57
3	T	8	GFL	C1'-N9-C4	-7.39	113.66	126.64
3	T	10	GFL	C1'-N9-C4	-7.37	113.70	126.64
3	T	4	GFL	C1'-N9-C4	-6.12	115.88	126.64
3	T	14	GFL	C1'-N9-C4	-6.05	116.01	126.64
3	T	16	A5L	N3-C2-N1	-5.68	119.79	128.68
3	T	11	A5L	N3-C2-N1	-5.67	119.81	128.68
3	T	4	GFL	N3-C2-N1	-5.58	119.78	127.22
3	T	6	A5L	N3-C2-N1	-5.52	120.05	128.68
3	T	8	GFL	N3-C2-N1	-5.50	119.88	127.22
3	T	14	GFL	N3-C2-N1	-5.41	120.01	127.22
3	T	10	GFL	N3-C2-N1	-5.39	120.03	127.22
3	T	14	GFL	C3'-C2'-C1'	4.94	109.11	103.13
3	T	15	CFL	C2-N3-C4	4.84	121.24	116.34
3	T	13	CFL	C2-N3-C4	4.42	120.82	116.34
3	T	4	GFL	C2-N3-C4	4.39	120.37	115.36
3	T	14	GFL	C2-N3-C4	4.38	120.35	115.36
3	T	7	CFL	C2-N3-C4	4.27	120.67	116.34
3	T	10	GFL	C2'-C3'-C4'	4.22	107.86	102.40
3	T	14	GFL	C2'-C3'-C4'	4.21	107.84	102.40
3	T	16	A5L	O4'-C1'-C2'	-4.18	101.49	105.79
3	T	8	GFL	C3'-C2'-C1'	4.16	108.17	103.13
3	T	10	GFL	C2-N3-C4	4.09	120.03	115.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	6	A5L	C2'-C3'-C4'	4.06	107.65	102.40
3	T	8	GFL	C2-N3-C4	4.00	119.93	115.36
3	T	11	A5L	C2'-C3'-C4'	3.70	107.18	102.40
3	T	14	GFL	C5'-C4'-C3'	-3.68	101.41	115.18
3	T	7	CFL	C2'-C3'-C4'	3.50	106.92	102.40
3	T	16	A5L	C2'-C3'-C4'	3.30	106.67	102.40
3	T	12	UF2	C3'-C2'-C1'	3.26	107.07	103.13
3	T	11	A5L	C3'-C2'-C1'	3.14	106.94	103.13
3	T	9	UF2	C2'-C3'-C4'	3.01	106.30	102.40
3	T	12	UF2	C2'-C3'-C4'	3.00	106.28	102.40
3	T	10	GFL	C3'-C2'-C1'	2.96	106.72	103.13
3	T	14	GFL	O3'-C3'-C4'	2.95	119.59	111.05
3	T	15	CFL	C2'-C3'-C4'	2.93	106.19	102.40
3	T	16	A5L	C3'-C2'-C1'	-2.85	99.68	103.13
3	T	9	UF2	C3'-C2'-C1'	2.64	106.33	103.13
3	T	5	UF2	C2'-C3'-C4'	2.60	105.76	102.40
3	T	8	GFL	C5-C6-N1	-2.56	119.92	123.43
3	T	6	A5L	C3'-C2'-C1'	2.54	106.20	103.13
3	T	8	GFL	C6-N1-C2	2.53	119.95	115.93
3	T	13	CFL	C2'-C3'-C4'	2.53	105.67	102.40
3	T	6	A5L	F-C2'-C1'	2.48	114.25	109.08
3	T	7	CFL	N4-C4-N3	2.43	120.34	116.49
3	T	4	GFL	C3'-C2'-C1'	2.42	106.07	103.13
3	T	16	A5L	C1'-N9-C4	-2.41	122.40	126.64
3	T	4	GFL	C6-N1-C2	2.38	119.71	115.93
3	T	8	GFL	C5'-C4'-C3'	-2.37	106.31	115.18
3	T	10	GFL	C6-N1-C2	2.35	119.66	115.93
3	T	4	GFL	C5-C6-N1	-2.34	120.23	123.43
3	T	14	GFL	C5-C6-N1	-2.33	120.25	123.43
3	T	14	GFL	C6-N1-C2	2.30	119.59	115.93
3	T	15	CFL	N4-C4-N3	2.29	120.10	116.49
3	T	11	A5L	F-C2'-C1'	2.26	113.79	109.08
3	T	14	GFL	O4'-C1'-C2'	2.24	108.10	105.79
3	T	10	GFL	C5-C6-N1	-2.23	120.39	123.43
3	T	13	CFL	N4-C4-N3	2.11	119.83	116.49
3	T	4	GFL	N2-C2-N1	2.08	120.48	117.25
3	T	9	UF2	F2'-C2'-C3'	2.07	113.56	109.22
3	T	12	UF2	F2'-C2'-C3'	2.06	113.54	109.22
3	T	6	A5L	C5'-C4'-C3'	-2.00	107.67	115.18

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	T	6	A5L	C3'-C4'-C5'-O5'
3	T	8	GFL	C3'-C4'-C5'-O5'
3	T	15	CFL	C3'-C4'-C5'-O5'
3	T	8	GFL	O4'-C4'-C5'-O5'
3	T	6	A5L	O4'-C4'-C5'-O5'
3	T	4	GFL	C3'-C4'-C5'-O5'
3	T	4	GFL	O4'-C4'-C5'-O5'
3	T	16	A5L	O4'-C4'-C5'-O5'
3	T	15	CFL	O4'-C4'-C5'-O5'
3	T	16	A5L	C3'-C4'-C5'-O5'
3	T	5	UF2	O4'-C4'-C5'-O5'

There are no ring outliers.

8 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	T	16	A5L	1	0
3	T	5	UF2	1	0
3	T	4	GFL	1	0
3	T	7	CFL	2	0
3	T	15	CFL	2	0
3	T	6	A5L	1	0
3	T	14	GFL	2	0
3	T	8	GFL	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\# Z  > 2$
5	SO4	A	905	-	4,4,4	0.14	0	6,6,6	0.07	0
4	MPD	A	903	-	7,7,7	0.38	0	9,10,10	0.78	0
5	SO4	A	904	-	4,4,4	0.15	0	6,6,6	0.12	0
4	MPD	A	902	-	7,7,7	0.28	0	9,10,10	0.41	0
4	MPD	A	901	-	7,7,7	0.37	0	9,10,10	0.56	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MPD	A	903	-	-	1/5/5/5	-
4	MPD	A	902	-	-	1/5/5/5	-
4	MPD	A	901	-	-	0/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	903	MPD	C1-C2-C3-C4
4	A	902	MPD	C2-C3-C4-O4

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	903	MPD	5	0
4	A	902	MPD	1	0
4	A	901	MPD	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	P	11/11 (100%)	1.77	4 (36%) 0 0	59, 73, 113, 149	0
2	A	577/577 (100%)	0.23	27 (4%) 31 28	18, 31, 60, 100	0
3	T	0/13	-	-	-	-
All	All	588/601 (97%)	0.26	31 (5%) 26 23	18, 31, 64, 149	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	433	ALA	5.8
2	A	550	THR	5.5
2	A	551	LYS	5.4
1	P	1	DG	5.1
2	A	631	GLU	4.6
2	A	300	ALA	4.5
2	A	552	THR	4.5
2	A	553	GLY	4.2
2	A	520	GLU	4.2
2	A	306	ARG	4.1
2	A	632	GLU	4.0
1	P	3	DG	3.9
2	A	522	ALA	3.9
2	A	719	TYR	3.7
2	A	434	LYS	3.7
2	A	512	ARG	3.3
2	A	305	ASP	3.2
2	A	531	PRO	3.0
1	P	4	DA	3.0
2	A	505	GLU	2.9
2	A	554	TYR	2.7
2	A	509	GLU	2.6
1	P	5	DT	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	A	549	LYS	2.4
2	A	431	LYS	2.3
2	A	513	THR	2.3
2	A	729	ARG	2.2
2	A	630	LEU	2.1
2	A	579	GLN	2.0
2	A	523	GLY	2.0
2	A	519	TYR	2.0

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	UF2	T	5	20/21	0.62	0.17	93,103,117,117	0
3	GFL	T	4	23/24	0.64	0.25	77,86,120,121	0
3	A5L	T	16	22/23	0.65	0.45	92,101,123,126	0
3	CFL	T	15	20/21	0.67	0.36	99,108,124,125	0
3	GFL	T	14	23/24	0.70	0.31	77,98,106,119	0
3	CFL	T	7	20/21	0.75	0.16	56,75,94,98	0
3	A5L	T	6	22/23	0.76	0.15	83,94,120,123	0
3	GFL	T	10	23/24	0.84	0.18	43,53,58,59	0
3	GFL	T	8	23/24	0.84	0.13	38,56,61,67	0
3	UF2	T	9	20/21	0.91	0.12	32,44,55,58	0
3	CFL	T	13	20/21	0.92	0.27	47,64,73,81	0
3	UF2	T	12	20/21	0.92	0.16	40,53,60,68	0
3	A5L	T	11	22/23	0.94	0.16	33,43,50,51	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

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



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MPD	A	901	8/8	0.56	0.25	33,55,65,66	0
4	MPD	A	903	8/8	0.79	0.21	46,61,66,79	0
4	MPD	A	902	8/8	0.86	0.21	37,53,60,67	0
5	SO4	A	905	5/5	0.87	0.15	85,91,93,93	0
6	MG	A	906	1/1	0.95	0.12	52,52,52,52	0
5	SO4	A	904	5/5	0.98	0.07	58,66,72,73	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.