



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 10:37 am BST

PDB ID : 3MBU  
Title : Structure of a bipyridine-modified PNA duplex  
Authors : Yeh, J.I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G.M.; Du, S.; Achim, C.  
Deposited on : 2010-03-26  
Resolution : 1.05 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

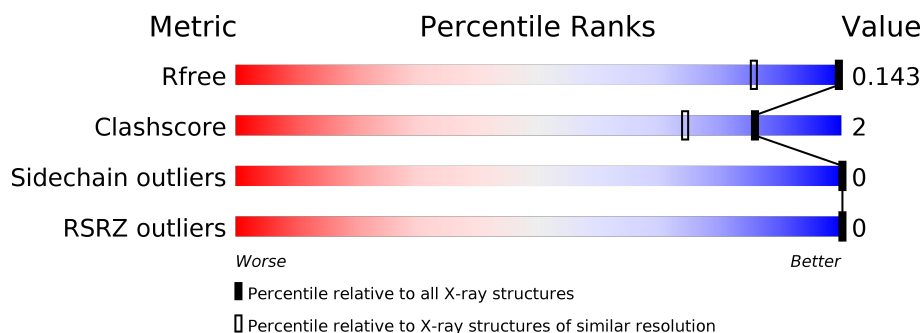
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.05 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	1202 (1.10-1.02)
Clashscore	141614	1252 (1.10-1.02)
Sidechain outliers	138945	1202 (1.10-1.02)
RSRZ outliers	127900	1178 (1.10-1.02)

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	A	10	<div><div></div>100%</div>
1	B	10	<div><div></div>100%</div>
1	C	10	<div><div></div>100%</div>
1	D	10	<div><div></div>90%<div></div>10%</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
5	GOL	D	5010	-	X	-	-

## 2 Entry composition [i](#)

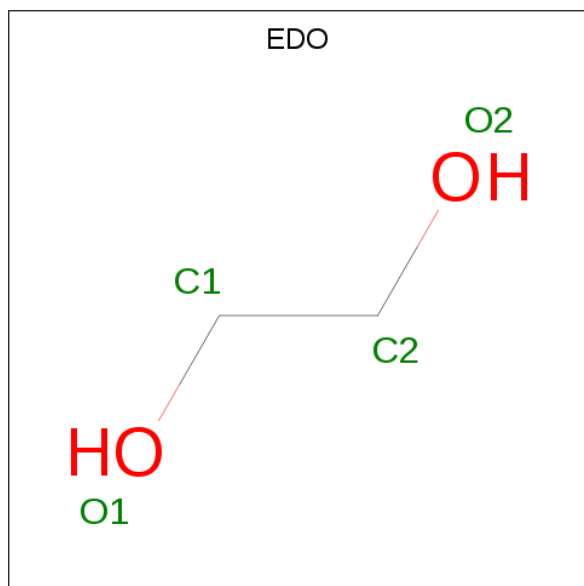
There are 7 unique types of molecules in this entry. The entry contains 1021 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bipyridine-PNA.

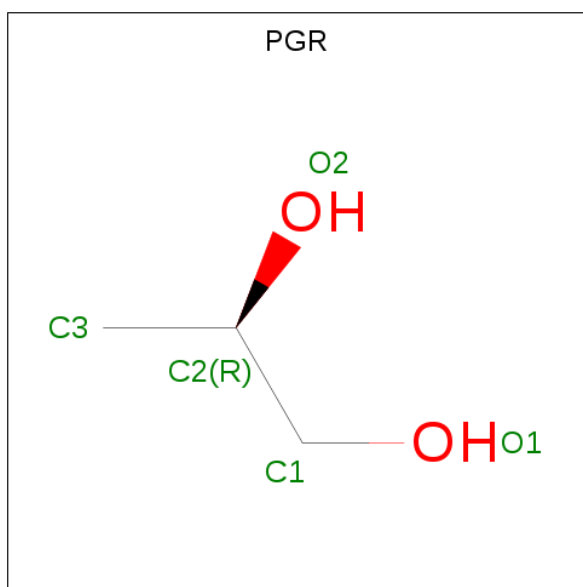
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	10	Total	C	N	O	0	0	0
			183	104	52	27			
1	B	10	Total	C	N	O	0	0	0
			183	104	52	27			
1	C	10	Total	C	N	O	0	0	0
			187	107	53	27			
1	D	10	Total	C	N	O	0	0	0
			187	107	53	27			

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



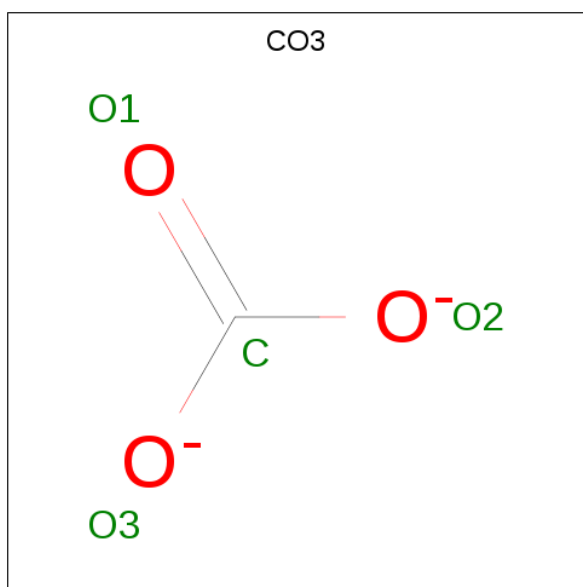
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is R-1,2-PROPANEDIOL (three-letter code: PGR) (formula:  $C_3H_8O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			5	3	2		
3	B	1	Total	C	O	0	0
			5	3	2		
3	B	1	Total	C	O	0	0
			5	3	2		

- Molecule 4 is CARBONATE ION (three-letter code: CO3) (formula:  $\text{CO}_3$ ).



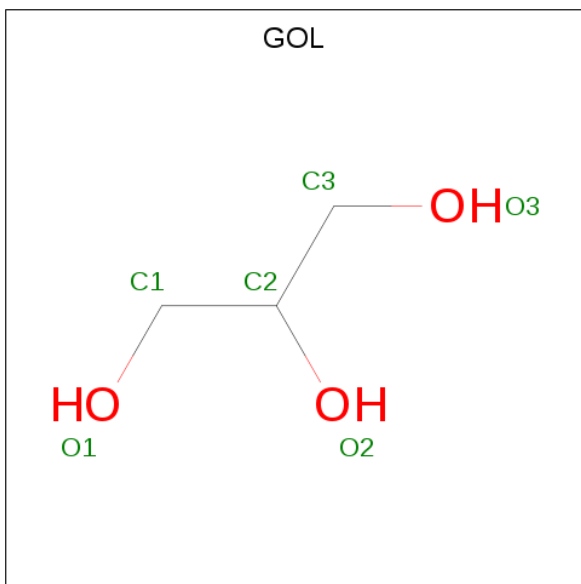
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	1	3		

*Continued on next page...*

*Continued from previous page...*

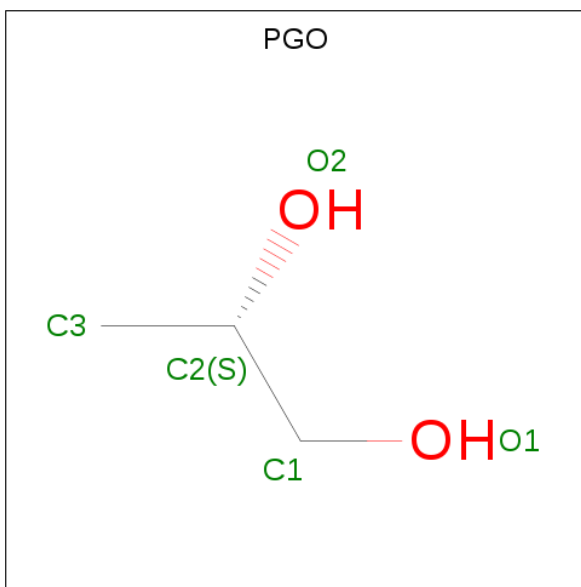
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			4	1	3		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is S-1,2-PROPANEDIOL (three-letter code: PGO) (formula:  $C_3H_8O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			5	3	2		
6	D	1	Total	C	O	0	0
			5	3	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	62	Total	O	0	0
			62	62		
7	B	49	Total	O	0	0
			49	49		
7	C	62	Total	O	0	0
			62	62		
7	D	59	Total	O	0	0
			59	59		

### 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: Bipyridine-PNA

Chain A:  100%

There are no outlier residues recorded for this chain.

- Molecule 1: Bipyridine-PNA

Chain B:  100%

There are no outlier residues recorded for this chain.

- Molecule 1: Bipyridine-PNA

Chain C:  100%

There are no outlier residues recorded for this chain.

- Molecule 1: Bipyridine-PNA

Chain D:  90% 10%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.20Å 25.73Å 53.05Å 90.00° 105.33° 90.00°	Depositor
Resolution (Å)	26.73 – 1.05 36.84 – 1.10	Depositor EDS
% Data completeness (in resolution range)	(Not available) (26.73-1.05) 90.9 (36.84-1.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.16 (at 1.10Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.127 , 0.169 0.136 , 0.143	Depositor DCC
$R_{free}$ test set	1930 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	8.3	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 57.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.38$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1021	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.35 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.2253e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, GPN, CPN, TPN, EDO, APN, PGO, PGR, KPN, CO3

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	0.40	0/4	1.46	0/4
1	B	0.74	0/4	1.16	0/4
1	C	0.79	0/8	0.89	0/8
1	D	0.59	0/8	1.18	0/8
All	All	0.66	0/24	1.14	0/24

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	183	0	123	0	0
1	B	183	0	123	0	0
1	C	187	0	134	0	0
1	D	187	0	133	3	0
2	A	4	0	6	0	0
3	A	5	0	8	0	0
3	B	10	0	16	0	0
4	C	4	0	0	0	0
4	D	4	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	6	0	8	0	0
5	D	6	0	6	3	0
6	C	5	0	8	0	0
6	D	5	0	8	0	0
7	A	62	0	0	0	0
7	B	49	0	0	0	0
7	C	62	0	0	0	0
7	D	59	0	0	0	0
All	All	1021	0	573	3	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 2.

All (3) 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:4002:GPN:HN22	5:D:5010:GOL:H32	1.28	0.97
1:D:4002:GPN:N2	5:D:5010:GOL:H32	2.07	0.66
1:D:4002:GPN:HN22	5:D:5010:GOL:C3	2.14	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

There are no protein backbone outliers to report in this entry.

### 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	C	1/1 (100%)	1 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	1/1 (100%)	1 (100%)	0	100	100
All	All	2/2 (100%)	2 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 ⓘ

36 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	CPN	A	1003	1	15,18,19	0.71	0	16,23,25	1.63	2 (12%)
1	APN	A	1004	1	18,21,22	0.97	1 (5%)	15,28,30	1.71	3 (20%)
1	CPN	B	2009	1	15,18,19	1.03	1 (6%)	16,23,25	1.61	2 (12%)
1	CPN	C	3003	1	15,18,19	0.92	1 (6%)	16,23,25	1.60	2 (12%)
1	GPN	D	4007	1	18,22,23	1.14	1 (5%)	17,30,32	2.86	6 (35%)
1	KPN	D	4005	1	23,23,24	1.02	1 (4%)	27,29,31	1.96	8 (29%)
1	APN	C	3004	1	18,21,22	0.70	0	15,28,30	1.61	4 (26%)
1	CPN	B	2003	1	15,18,19	0.79	0	16,23,25	1.34	2 (12%)
1	TPN	C	3006	1	16,19,20	1.04	2 (12%)	15,25,27	4.13	2 (13%)
1	KPN	A	1005	1	23,23,24	1.13	1 (4%)	27,29,31	1.98	9 (33%)
1	APN	D	4004	1	18,21,22	0.84	0	15,28,30	1.61	3 (20%)
1	CPN	C	3008	1	15,18,19	0.95	1 (6%)	16,23,25	1.57	3 (18%)
1	GPN	A	1001	1	18,22,23	0.97	1 (5%)	17,30,32	2.84	6 (35%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	GPN	D	4001	1	18,22,23	1.34	1 (5%)	17,30,32	2.72	5 (29%)
1	CPN	D	4008	1	15,18,19	0.73	0	16,23,25	1.54	3 (18%)
1	GPN	C	3001	1	18,22,23	1.16	1 (5%)	17,30,32	2.73	5 (29%)
1	CPN	B	2008	1	15,18,19	0.68	0	16,23,25	1.62	2 (12%)
1	GPN	D	4002	1	18,22,23	1.16	1 (5%)	17,30,32	2.72	5 (29%)
1	CPN	A	1008	1	15,18,19	1.10	2 (13%)	16,23,25	1.30	2 (12%)
1	GPN	A	1007	1	18,22,23	1.49	2 (11%)	17,30,32	2.64	6 (35%)
1	GPN	C	3002	1	18,22,23	1.13	1 (5%)	17,30,32	2.35	4 (23%)
1	GPN	C	3007	1	18,22,23	1.12	2 (11%)	17,30,32	2.66	6 (35%)
1	CPN	D	4003	1	15,18,19	1.13	1 (6%)	16,23,25	1.24	2 (12%)
1	KPN	B	2005	1	23,23,24	1.11	2 (8%)	27,29,31	2.36	12 (44%)
1	KPN	C	3005	1	23,23,24	1.25	1 (4%)	27,29,31	1.98	9 (33%)
1	GPN	B	2001	1	18,22,23	1.33	1 (5%)	17,30,32	2.95	6 (35%)
1	APN	B	2004	1	18,21,22	0.99	0	15,28,30	1.91	3 (20%)
1	TPN	B	2006	1	16,19,20	1.19	1 (6%)	15,25,27	2.98	2 (13%)
1	GPN	A	1002	1	18,22,23	1.30	1 (5%)	17,30,32	2.70	5 (29%)
1	TPN	A	1006	1	16,19,20	1.30	1 (6%)	15,25,27	3.16	3 (20%)
1	CPN	D	4009	1	15,18,19	0.75	0	16,23,25	1.31	1 (6%)
1	TPN	D	4006	1	16,19,20	1.27	2 (12%)	15,25,27	3.76	1 (6%)
1	GPN	B	2007	1	18,22,23	1.18	1 (5%)	17,30,32	2.78	6 (35%)
1	GPN	B	2002	1	18,22,23	1.26	2 (11%)	17,30,32	2.88	6 (35%)
1	CPN	A	1009	1	15,18,19	1.15	1 (6%)	16,23,25	1.55	3 (18%)
1	CPN	C	3009	1	15,18,19	0.77	0	16,23,25	1.65	2 (12%)

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	CPN	A	1003	1	-	0/13/14/15	0/1/1/1
1	APN	A	1004	1	-	0/13/14/15	0/2/2/2
1	CPN	B	2009	1	-	0/13/14/15	0/1/1/1
1	CPN	C	3003	1	-	0/13/14/15	0/1/1/1
1	GPN	D	4007	1	-	0/13/14/15	0/2/2/2
1	KPN	D	4005	1	-	3/17/18/19	0/2/2/2
1	APN	C	3004	1	-	0/13/14/15	0/2/2/2

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CPN	B	2003	1	-	0/13/14/15	0/1/1/1
1	TPN	C	3006	1	-	0/13/14/15	0/1/1/1
1	KPN	A	1005	1	-	3/17/18/19	0/2/2/2
1	APN	D	4004	1	-	0/13/14/15	0/2/2/2
1	CPN	C	3008	1	-	0/13/14/15	0/1/1/1
1	GPN	A	1001	1	-	1/13/14/15	0/2/2/2
1	GPN	D	4001	1	-	0/13/14/15	0/2/2/2
1	CPN	D	4008	1	-	0/13/14/15	0/1/1/1
1	GPN	C	3001	1	-	0/13/14/15	0/2/2/2
1	CPN	B	2008	1	-	0/13/14/15	0/1/1/1
1	GPN	D	4002	1	-	0/13/14/15	0/2/2/2
1	CPN	A	1008	1	-	0/13/14/15	0/1/1/1
1	GPN	A	1007	1	-	0/13/14/15	0/2/2/2
1	GPN	C	3002	1	-	0/13/14/15	0/2/2/2
1	GPN	C	3007	1	-	0/13/14/15	0/2/2/2
1	CPN	D	4003	1	-	0/13/14/15	0/1/1/1
1	KPN	B	2005	1	-	6/17/18/19	0/2/2/2
1	KPN	C	3005	1	-	1/17/18/19	0/2/2/2
1	GPN	B	2001	1	-	0/13/14/15	0/2/2/2
1	APN	B	2004	1	-	0/13/14/15	0/2/2/2
1	TPN	B	2006	1	-	0/13/14/15	0/1/1/1
1	GPN	A	1002	1	-	0/13/14/15	0/2/2/2
1	TPN	A	1006	1	-	0/13/14/15	0/1/1/1
1	CPN	D	4009	1	-	0/13/14/15	0/1/1/1
1	TPN	D	4006	1	-	0/13/14/15	0/1/1/1
1	GPN	B	2007	1	-	0/13/14/15	0/2/2/2
1	GPN	B	2002	1	-	0/13/14/15	0/2/2/2
1	CPN	A	1009	1	-	0/13/14/15	0/1/1/1
1	CPN	C	3009	1	-	0/13/14/15	0/1/1/1

All (34) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	4001	GPN	C6-N1	4.69	1.41	1.33
1	B	2001	GPN	C6-N1	4.59	1.41	1.33
1	C	3005	KPN	C6-C5	-4.56	1.38	1.49
1	A	1002	GPN	C6-N1	4.54	1.40	1.33
1	A	1007	GPN	C6-N1	4.42	1.40	1.33
1	C	3001	GPN	C6-N1	4.41	1.40	1.33
1	A	1006	TPN	C4-N3	4.19	1.40	1.33
1	C	3002	GPN	C6-N1	3.99	1.40	1.33
1	B	2007	GPN	C6-N1	3.98	1.40	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	4007	GPN	C6-N1	3.86	1.39	1.33
1	B	2002	GPN	C6-N1	3.81	1.39	1.33
1	D	4005	KPN	C6-C5	-3.74	1.40	1.49
1	A	1005	KPN	C6-C5	-3.74	1.40	1.49
1	D	4002	GPN	C6-N1	3.67	1.39	1.33
1	B	2005	KPN	C6-C5	-3.45	1.41	1.49
1	B	2006	TPN	C4-N3	3.34	1.38	1.33
1	C	3007	GPN	C6-N1	3.26	1.38	1.33
1	A	1009	CPN	C5'-C'	3.19	1.60	1.49
1	A	1007	GPN	C2-N1	3.15	1.41	1.35
1	A	1001	GPN	C6-N1	2.95	1.38	1.33
1	D	4006	TPN	C8'-N1	-2.88	1.44	1.47
1	A	1004	APN	C5'-N4'	2.87	1.51	1.47
1	A	1008	CPN	C8'-N1	-2.75	1.44	1.47
1	B	2005	KPN	C8'-C2	2.50	1.55	1.51
1	D	4003	CPN	C4-N3	2.49	1.39	1.35
1	C	3007	GPN	C2-N1	2.44	1.39	1.35
1	C	3006	TPN	C4-N3	2.44	1.37	1.33
1	C	3008	CPN	C8'-N1	-2.42	1.45	1.47
1	D	4006	TPN	C4-N3	2.36	1.37	1.33
1	B	2009	CPN	C6-N1	2.35	1.39	1.35
1	C	3003	CPN	C8'-N1	-2.20	1.45	1.47
1	A	1008	CPN	C6-C5	-2.10	1.33	1.38
1	B	2002	GPN	C8-N7	-2.04	1.31	1.34
1	C	3006	TPN	C8'-N1	-2.03	1.45	1.47

All (151) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	3006	TPN	C4-N3-C2	15.39	128.13	115.14
1	D	4006	TPN	C4-N3-C2	13.92	126.89	115.14
1	A	1006	TPN	C4-N3-C2	11.44	124.80	115.14
1	B	2006	TPN	C4-N3-C2	10.45	123.96	115.14
1	B	2001	GPN	C5-C6-N1	-8.40	111.95	123.43
1	D	4001	GPN	C5-C6-N1	-8.03	112.45	123.43
1	B	2002	GPN	C5-C6-N1	-7.73	112.85	123.43
1	C	3001	GPN	C5-C6-N1	-7.67	112.94	123.43
1	A	1001	GPN	C5-C6-N1	-7.57	113.08	123.43
1	A	1002	GPN	C5-C6-N1	-7.40	113.31	123.43
1	D	4002	GPN	C5-C6-N1	-7.37	113.36	123.43
1	B	2007	GPN	C5-C6-N1	-7.34	113.40	123.43
1	D	4007	GPN	C5-C6-N1	-7.32	113.41	123.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	3007	GPN	C5-C6-N1	-7.31	113.43	123.43
1	A	1007	GPN	C5-C6-N1	-7.06	113.77	123.43
1	C	3002	GPN	C5-C6-N1	-6.73	114.22	123.43
1	B	2002	GPN	C6-N1-C2	6.12	125.65	115.93
1	B	2005	KPN	C10-N2-C6	6.12	125.81	117.23
1	A	1001	GPN	C6-N1-C2	5.93	125.35	115.93
1	D	4007	GPN	C6-N1-C2	5.83	125.20	115.93
1	B	2007	GPN	C6-N1-C2	5.80	125.14	115.93
1	B	2001	GPN	C6-N1-C2	5.67	124.94	115.93
1	C	3001	GPN	C6-N1-C2	5.64	124.89	115.93
1	A	1002	GPN	C6-N1-C2	5.45	124.58	115.93
1	C	3007	GPN	C6-N1-C2	5.38	124.48	115.93
1	D	4002	GPN	C6-N1-C2	5.30	124.36	115.93
1	B	2008	CPN	C2-N3-C4	5.20	121.61	116.34
1	D	4001	GPN	C6-N1-C2	5.08	124.00	115.93
1	D	4005	KPN	C10-N2-C6	4.93	124.14	117.23
1	B	2004	APN	C7'-C8'-N9	4.90	116.59	110.76
1	C	3009	CPN	C2-N3-C4	4.80	121.20	116.34
1	A	1007	GPN	C6-N1-C2	4.69	123.38	115.93
1	B	2009	CPN	C2-N3-C4	4.51	120.92	116.34
1	C	3005	KPN	C10-N2-C6	4.49	123.54	117.23
1	C	3003	CPN	C2-N3-C4	4.49	120.89	116.34
1	A	1003	CPN	C2-N3-C4	4.48	120.88	116.34
1	C	3008	CPN	C2-N3-C4	4.47	120.87	116.34
1	C	3002	GPN	C6-N1-C2	4.32	122.79	115.93
1	D	4009	CPN	C2-N3-C4	4.31	120.71	116.34
1	B	2002	GPN	N3-C2-N1	-4.09	121.76	127.22
1	A	1002	GPN	N3-C2-N1	-4.05	121.83	127.22
1	C	3005	KPN	C8'-C7'-N4'	4.04	124.76	117.74
1	A	1009	CPN	C2-N3-C4	3.95	120.35	116.34
1	B	2007	GPN	N3-C2-N1	-3.92	121.99	127.22
1	A	1001	GPN	N3-C2-N1	-3.86	122.07	127.22
1	A	1007	GPN	N3-C2-N1	-3.73	122.25	127.22
1	A	1005	KPN	C8'-C7'-N4'	3.72	124.20	117.74
1	C	3001	GPN	N3-C2-N1	-3.69	122.31	127.22
1	C	3007	GPN	N3-C2-N1	-3.67	122.32	127.22
1	A	1004	APN	C5-C6-N6	3.65	125.89	120.35
1	D	4007	GPN	N3-C2-N1	-3.65	122.36	127.22
1	D	4004	APN	C7'-C8'-N9	3.64	115.10	110.76
1	B	2006	TPN	C7'-C8'-N1	-3.62	106.45	110.76
1	B	2001	GPN	C7'-C8'-N9	-3.60	106.47	110.76
1	B	2005	KPN	C8'-C2-C3	3.54	125.97	120.89

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1005	KPN	C9-C10-N2	-3.54	117.65	123.43
1	D	4008	CPN	C2-N3-C4	3.54	119.93	116.34
1	B	2005	KPN	C7-C6-N2	-3.49	117.06	122.26
1	B	2003	CPN	O1'-C'-C5'	-3.41	116.09	126.39
1	D	4001	GPN	N3-C2-N1	-3.40	122.69	127.22
1	D	4007	GPN	C8'-N9-C8	3.39	129.67	125.66
1	D	4002	GPN	N3-C2-N1	-3.33	122.78	127.22
1	A	1005	KPN	C10-N2-C6	3.32	121.89	117.23
1	B	2004	APN	C5-C6-N6	3.28	125.33	120.35
1	A	1005	KPN	C2-C8'-C7'	-3.26	106.86	112.61
1	A	1007	GPN	C7'-C8'-N9	-3.26	106.88	110.76
1	C	3004	APN	C5-C6-N6	3.24	125.28	120.35
1	B	2005	KPN	C9-C10-N2	-3.22	118.16	123.43
1	C	3003	CPN	O1'-C'-C5'	-3.19	116.77	126.39
1	B	2001	GPN	N3-C2-N1	-3.17	123.00	127.22
1	C	3004	APN	C7'-C8'-N9	3.15	114.51	110.76
1	B	2005	KPN	C6-C5-N1	3.15	122.35	116.77
1	D	4005	KPN	C7-C6-N2	-3.13	117.59	122.26
1	A	1009	CPN	C5'-N4'-C3'	-3.13	112.61	117.06
1	B	2002	GPN	O1'-C'-C5'	-3.13	116.95	126.39
1	D	4004	APN	C5-C6-N6	3.13	125.10	120.35
1	B	2005	KPN	C5-C6-N2	3.10	122.26	116.77
1	B	2005	KPN	O7'-C7'-C8'	-3.09	116.57	121.60
1	B	2005	KPN	C1-N1-C5	3.09	121.97	117.90
1	B	2005	KPN	C8'-C2-C1	-3.07	116.04	120.94
1	A	1005	KPN	C1-N1-C5	3.05	121.91	117.90
1	D	4005	KPN	C8'-C7'-N4'	3.00	122.96	117.74
1	A	1004	APN	C7'-C8'-N9	2.98	114.31	110.76
1	B	2003	CPN	C2-N3-C4	2.97	119.35	116.34
1	A	1002	GPN	O1'-C'-C5'	-2.96	117.46	126.39
1	D	4001	GPN	O1'-C'-C5'	-2.94	117.50	126.39
1	D	4008	CPN	C5-C6-N1	-2.89	118.97	121.28
1	D	4005	KPN	C9-C10-N2	-2.89	118.71	123.43
1	B	2004	APN	N3-C2-N1	-2.86	124.21	128.68
1	C	3005	KPN	C2-C8'-C7'	-2.86	107.57	112.61
1	A	1005	KPN	C9-C8-C7	-2.83	115.87	120.19
1	D	4008	CPN	O1'-C'-C5'	-2.79	117.96	126.39
1	D	4003	CPN	O1'-C'-C5'	-2.79	117.97	126.39
1	D	4002	GPN	O1'-C'-C5'	-2.76	118.05	126.39
1	C	3005	KPN	C9-C10-N2	-2.74	118.96	123.43
1	A	1008	CPN	O1'-C'-C5'	-2.73	118.15	126.39
1	A	1003	CPN	O1'-C'-C5'	-2.69	118.27	126.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1001	GPN	O1'-C'-C5'	-2.66	118.34	126.39
1	A	1006	TPN	C7'-C8'-N1	-2.66	107.59	110.76
1	B	2005	KPN	C8'-C7'-N4'	2.66	122.36	117.74
1	A	1004	APN	C5'-N4'-C3'	-2.64	113.31	117.06
1	A	1007	GPN	O1'-C'-C5'	-2.64	118.42	126.39
1	C	3005	KPN	C1-N1-C5	2.62	121.36	117.90
1	C	3005	KPN	C5-C6-N2	2.59	121.35	116.77
1	C	3007	GPN	O1'-C'-C5'	-2.59	118.58	126.39
1	B	2007	GPN	C6-C5-C4	-2.58	118.34	120.80
1	C	3009	CPN	C8'-N1-C6	2.57	123.21	118.63
1	C	3001	GPN	C6-C5-C4	-2.55	118.37	120.80
1	A	1007	GPN	N2-C2-N3	2.52	121.91	117.79
1	C	3002	GPN	O1'-C'-C5'	-2.49	118.86	126.39
1	B	2009	CPN	C5-C6-N1	-2.46	119.31	121.28
1	D	4005	KPN	C4-C5-N1	-2.46	118.59	122.26
1	C	3008	CPN	O1'-C'-C5'	-2.45	119.00	126.39
1	A	1005	KPN	C8-C9-C10	2.43	123.32	118.62
1	C	3005	KPN	C2'-C3'-N4'	-2.41	108.55	113.11
1	B	2001	GPN	O1'-C'-C5'	-2.41	119.10	126.39
1	B	2007	GPN	C8'-C7'-N4'	2.41	120.09	117.07
1	C	3005	KPN	O7'-C7'-N4'	-2.39	117.39	122.05
1	D	4007	GPN	C6-C5-C4	-2.39	118.52	120.80
1	D	4001	GPN	C8'-C7'-N4'	2.37	120.04	117.07
1	A	1005	KPN	O7'-C7'-C8'	-2.36	117.77	121.60
1	D	4005	KPN	C1-N1-C5	2.35	121.00	117.90
1	C	3005	KPN	O7'-C7'-C8'	-2.33	117.80	121.60
1	C	3007	GPN	C6-C5-C4	-2.32	118.58	120.80
1	C	3006	TPN	C8'-N1-C2	-2.28	115.26	117.86
1	C	3008	CPN	C8'-C7'-N4'	2.26	119.90	117.07
1	B	2008	CPN	O1'-C'-C5'	-2.26	119.58	126.39
1	D	4002	GPN	C8'-C7'-N4'	2.25	119.89	117.07
1	D	4005	KPN	C3'-N4'-C7'	2.25	128.98	121.03
1	D	4007	GPN	C8'-C7'-N4'	2.20	119.83	117.07
1	B	2007	GPN	O1'-C'-C5'	-2.20	119.75	126.39
1	A	1005	KPN	O7'-C7'-N4'	-2.20	117.77	122.05
1	C	3002	GPN	C2'-C3'-N4'	2.19	117.26	113.11
1	C	3007	GPN	N2-C2-N3	2.15	121.30	117.79
1	A	1001	GPN	C2'-C3'-N4'	2.15	117.18	113.11
1	C	3001	GPN	C8'-N9-C8	2.15	128.20	125.66
1	B	2005	KPN	C4-C5-N1	-2.14	119.07	122.26
1	B	2002	GPN	C8'-C7'-N4'	2.13	119.73	117.07
1	A	1009	CPN	O1'-C'-C5'	-2.07	120.15	126.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	3004	APN	N3-C2-N1	-2.06	125.45	128.68
1	A	1006	TPN	C3'-N4'-C7'	2.05	128.29	121.03
1	D	4003	CPN	C2-N3-C4	2.05	118.42	116.34
1	A	1001	GPN	C6-C5-C4	-2.05	118.84	120.80
1	D	4005	KPN	C8'-C2-C3	2.05	123.82	120.89
1	A	1002	GPN	C6-C5-C4	-2.04	118.84	120.80
1	A	1008	CPN	C8'-C7'-N4'	2.04	119.62	117.07
1	B	2001	GPN	C2'-C3'-N4'	2.03	116.95	113.11
1	C	3004	APN	C5-C6-N1	-2.03	115.76	120.35
1	B	2005	KPN	C3'-N4'-C7'	2.02	128.17	121.03
1	B	2002	GPN	C6-C5-C4	-2.02	118.87	120.80
1	D	4004	APN	C8'-N9-C8	2.00	128.03	125.66

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	4005	KPN	C1-C2-C8'-C7'
1	B	2005	KPN	C1-C2-C8'-C7'
1	D	4005	KPN	O7'-C7'-C8'-C2
1	B	2005	KPN	O7'-C7'-C8'-C2
1	A	1001	GPN	C2'-C3'-N4'-C5'
1	A	1005	KPN	O7'-C7'-C8'-C2
1	D	4005	KPN	N4'-C7'-C8'-C2
1	A	1005	KPN	N4'-C7'-C8'-C2
1	C	3005	KPN	N4'-C7'-C8'-C2
1	B	2005	KPN	C8'-C7'-N4'-C5'
1	A	1005	KPN	C'-C5'-N4'-C3'
1	B	2005	KPN	C'-C5'-N4'-C3'
1	B	2005	KPN	N4'-C7'-C8'-C2
1	B	2005	KPN	C3-C2-C8'-C7'

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	4002	GPN	3	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

10 ligands 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
4	CO3	C	5002	-	0,3,3	0.00	-	0,3,3	0.00	-
3	PGR	B	5005	-	3,4,4	0.99	0	1,4,4	0.53	0
4	CO3	D	5001	-	0,3,3	0.00	-	0,3,3	0.00	-
6	PGO	D	5008	-	3,4,4	0.40	0	1,4,4	0.17	0
5	GOL	D	5010	-	5,5,5	1.34	0	5,5,5	3.92	4 (80%)
3	PGR	B	5004	-	3,4,4	0.43	0	1,4,4	0.02	0
6	PGO	C	5007	-	3,4,4	0.25	0	1,4,4	0.27	0
2	EDO	A	5003	-	3,3,3	0.37	0	2,2,2	0.14	0
5	GOL	C	5006	-	5,5,5	0.55	0	5,5,5	0.88	0
3	PGR	A	5009	-	3,4,4	0.38	0	1,4,4	0.58	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 Torsions and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	PGO	D	5008	-	-	0/2/2/2	-
3	PGR	B	5005	-	-	0/2/2/2	-
5	GOL	D	5010	-	-	3/4/4/4	-
3	PGR	B	5004	-	-	0/2/2/2	-
6	PGO	C	5007	-	-	0/2/2/2	-
2	EDO	A	5003	-	-	0/1/1/1	-
5	GOL	C	5006	-	-	0/4/4/4	-
3	PGR	A	5009	-	-	0/2/2/2	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	5010	GOL	O3-C3-C2	6.84	142.98	110.20
5	D	5010	GOL	O2-C2-C1	4.02	126.81	109.12
5	D	5010	GOL	O1-C1-C2	2.88	124.01	110.20
5	D	5010	GOL	O2-C2-C3	2.31	119.28	109.12

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	5010	GOL	O1-C1-C2-C3
5	D	5010	GOL	O2-C2-C3-O3
5	D	5010	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	5010	GOL	3	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	1/10 (10%)	0.32	0 100 100	21, 21, 21, 21	0
1	B	1/10 (10%)	-0.27	0 100 100	17, 17, 17, 17	0
1	C	1/10 (10%)	0.47	0 100 100	18, 18, 18, 18	0
1	D	1/10 (10%)	0.14	0 100 100	17, 17, 17, 17	0
All	All	4/40 (10%)	0.16	0 100 100	17, 17, 18, 21	0

There are no RSRZ outliers to report.

### 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. 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
1	KPN	A	1005	22/23	0.92	0.12	16,22,26,28	0
1	KPN	B	2005	22/23	0.94	0.09	14,19,23,34	0
1	KPN	C	3005	22/23	0.94	0.11	11,14,19,20	0
1	KPN	D	4005	22/23	0.95	0.08	11,14,21,24	0
1	APN	D	4004	20/21	0.97	0.08	9,10,12,13	0
1	APN	C	3004	20/21	0.97	0.07	8,10,13,14	0
1	APN	A	1004	20/21	0.97	0.07	10,11,16,16	0
1	TPN	B	2006	19/20	0.97	0.07	11,14,24,25	0
1	TPN	A	1006	19/20	0.97	0.07	10,13,22,22	0
1	CPN	C	3003	18/19	0.98	0.07	7,7,9,10	0
1	GPN	D	4007	21/22	0.98	0.07	6,7,11,13	0
1	GPN	A	1001	21/22	0.98	0.07	6,8,10,13	0
1	GPN	D	4001	21/22	0.98	0.07	6,8,11,16	0
1	GPN	C	3001	21/22	0.98	0.07	6,8,11,23	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CPN	B	2008	18/19	0.98	0.07	5,6,8,8	0
1	GPN	D	4002	21/22	0.98	0.07	5,7,8,9	0
1	CPN	A	1008	18/19	0.98	0.07	5,6,7,8	0
1	GPN	A	1007	21/22	0.98	0.07	7,8,10,11	0
1	GPN	C	3002	21/22	0.98	0.07	5,7,8,8	0
1	CPN	D	4003	18/19	0.98	0.07	6,7,10,14	0
1	CPN	A	1003	18/19	0.98	0.07	6,8,10,11	0
1	CPN	B	2009	18/19	0.98	0.07	7,8,13,14	0
1	GPN	B	2001	21/22	0.98	0.07	6,7,9,9	0
1	APN	B	2004	20/21	0.98	0.07	10,12,17,19	0
1	CPN	B	2003	18/19	0.98	0.07	6,8,11,16	0
1	TPN	C	3006	19/20	0.98	0.06	8,8,10,11	0
1	CPN	D	4009	18/19	0.98	0.07	6,7,9,10	0
1	TPN	D	4006	19/20	0.98	0.07	9,10,13,13	0
1	GPN	B	2007	21/22	0.98	0.07	6,8,12,13	0
1	GPN	B	2002	21/22	0.98	0.07	5,7,8,9	0
1	CPN	A	1009	18/19	0.98	0.08	6,7,12,14	0
1	GPN	C	3007	21/22	0.99	0.07	6,7,8,11	0
1	GPN	A	1002	21/22	0.99	0.07	6,6,8,8	0
1	CPN	C	3008	18/19	0.99	0.07	5,6,7,8	0
1	CPN	D	4008	18/19	0.99	0.07	5,6,7,8	0
1	CPN	C	3009	18/19	0.99	0.07	6,7,9,10	0

### 6.3 Carbohydrates

There are no carbohydrates in this entry.

### 6.4 Ligands

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
5	GOL	D	5010	6/6	0.95	0.15	9,14,22,23	0
2	EDO	A	5003	4/4	0.96	0.13	9,14,15,15	0
3	PGR	B	5005	5/5	0.97	0.14	9,15,16,18	0
3	PGR	B	5004	5/5	0.97	0.09	10,13,14,15	0
6	PGO	C	5007	5/5	0.97	0.08	8,11,12,15	0
4	CO3	D	5001	4/4	0.97	0.08	11,11,14,16	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PGR	A	5009	5/5	0.97	0.11	8,14,16,24	0
6	PGO	D	5008	5/5	0.98	0.07	8,10,10,13	0
4	CO3	C	5002	4/4	0.98	0.08	9,11,11,12	0
5	GOL	C	5006	6/6	0.99	0.09	8,15,20,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.