



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 8, 2021 – 02:05 pm BST

PDB ID : 7P31  
Title : Plasmodium falciparum Hsp70-x chaperone nucleotide binding domain in complex with NCL-00023818  
Authors : Mohamad, N.; O'Donoghue, A.; Kantsadi, A.L.; Vakonakis, I.  
Deposited on : 2021-07-06  
Resolution : 2.36 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 : 2.22  
buster-report : 1.1.7 (2018)  
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) : 2.22

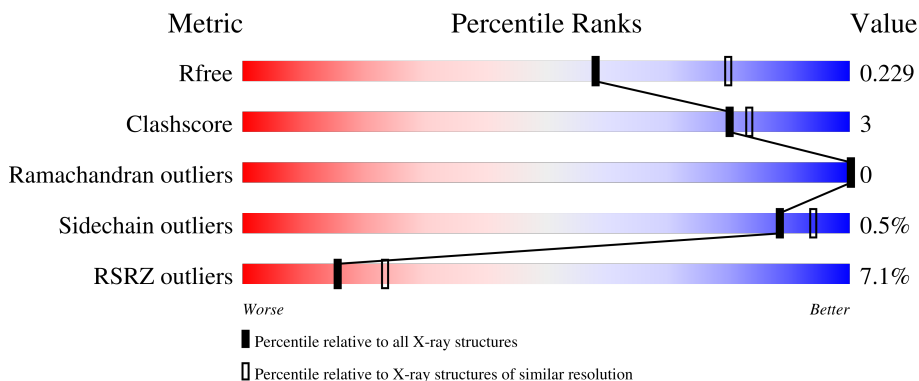
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

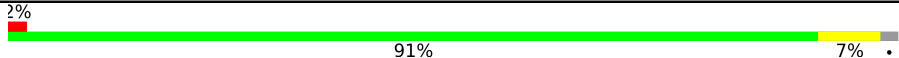
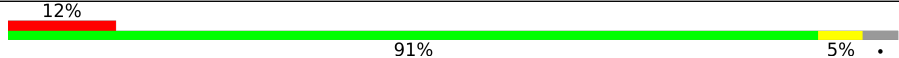
The reported resolution of this entry is 2.36 Å.

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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 of 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	393	
1	B	393	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	PYZ	A	504	-	-	X	-
5	PYZ	B	504	-	-	X	-

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 6102 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.

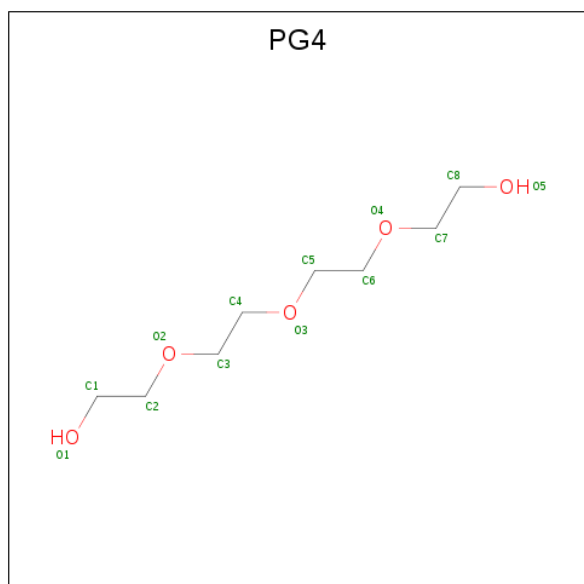
- Molecule 1 is a protein called Heat shock protein 70.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	2960	1853	512	582	13	0	1	0
1	B	376	2892	1815	499	566	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

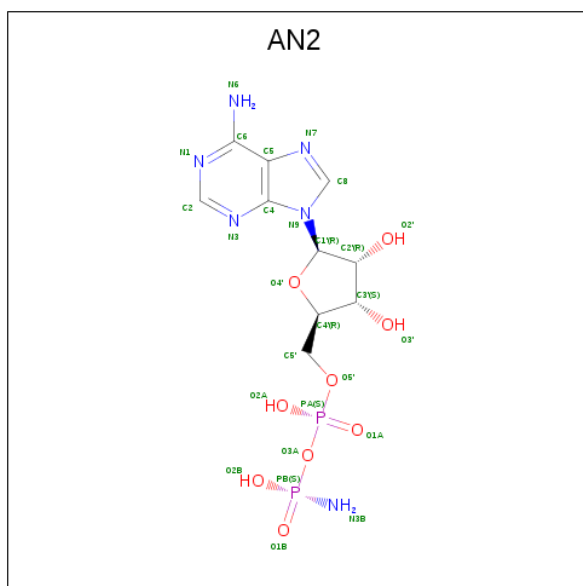
Chain	Residue	Modelled	Actual	Comment	Reference
A	27	GLY	-	expression tag	UNP K7NTP5
A	28	PRO	-	expression tag	UNP K7NTP5
B	27	GLY	-	expression tag	UNP K7NTP5
B	28	PRO	-	expression tag	UNP K7NTP5

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



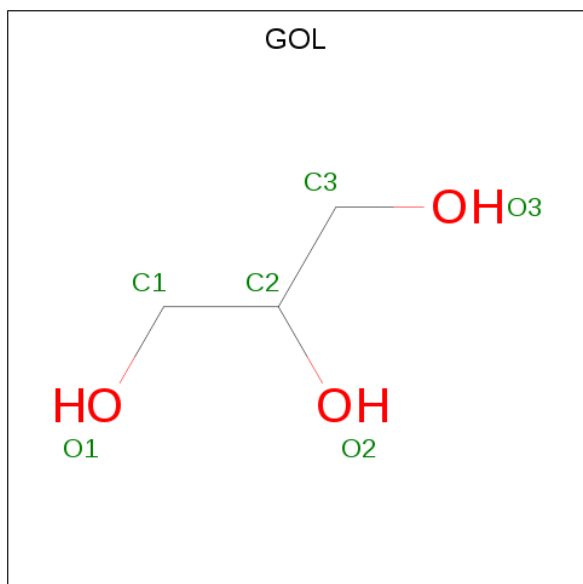
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			7	5	2		

- Molecule 3 is AMP PHOSPHORAMIDATE (three-letter code: AN2) (formula:  $C_{10}H_{16}N_6O_9P_2$ ).



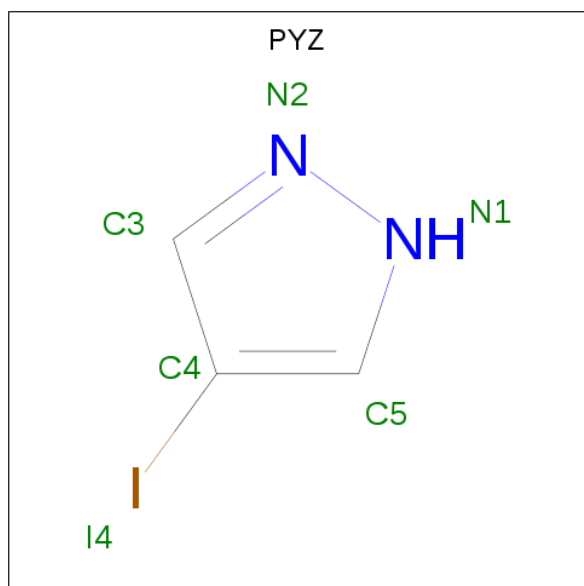
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			27	10	6	9	2		
3	B	1	Total	C	N	O	P	0	0
			27	10	6	9	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is 4-IODOPYRAZOLE (three-letter code: PYZ) (formula: C<sub>3</sub>H<sub>3</sub>IN<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	I	N	0	0
			6	3	1	2		
5	B	1	Total	C	I	N	0	0
			6	3	1	2		

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O P 5 4 1	0	0
6	B	1	Total O P 5 4 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total Cl 2 2	0	0

- Molecule 8 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total I 1 1	0	0

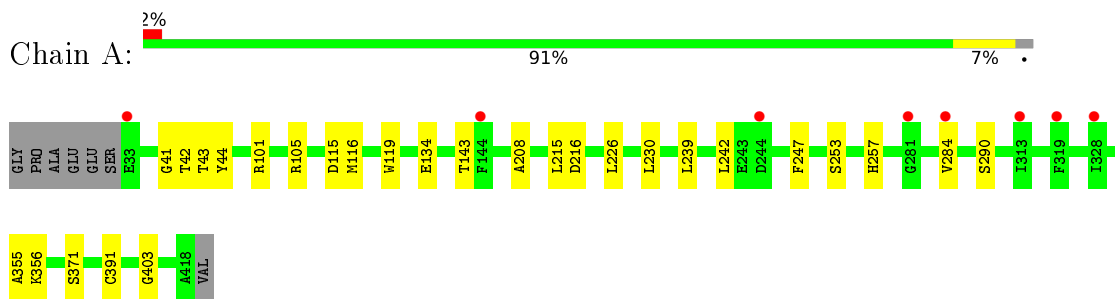
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	86	Total O 86 86	0	0
9	B	42	Total O 42 42	0	0

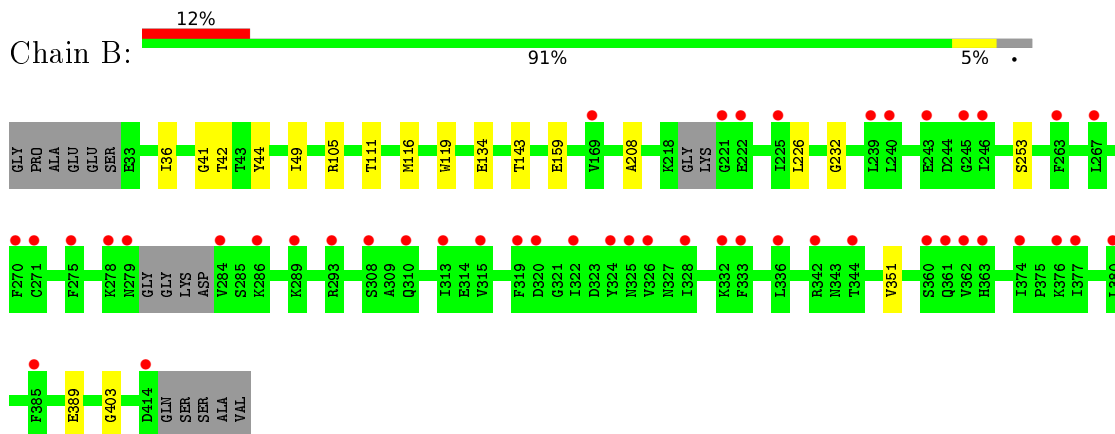
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Heat shock protein 70



- Molecule 1: Heat shock protein 70





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.98Å 102.58Å 103.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.97 – 2.36 72.97 – 2.36	Depositor EDS
% Data completeness (in resolution range)	99.9 (72.97-2.36) 100.0 (72.97-2.36)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 2.37Å)	Xtrriage
Refinement program	BUSTER 2.10.4 (24-FEB-2021)	Depositor
R, $R_{free}$	0.220 , 0.235 0.212 , 0.229	Depositor DCC
$R_{free}$ test set	1773 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.0	Xtrriage
Anisotropy	0.553	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 57.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.011 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6102	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	85.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.28% 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: AN2, GOL, CL, IOD, CSX, PG4, PYZ, PO4

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	0.45	0/2990	0.56	0/4027
1	B	0.37	0/2924	0.53	0/3938
All	All	0.41	0/5914	0.55	0/7965

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

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	2960	0	2982	20	0
1	B	2892	0	2913	12	0
2	A	7	0	6	0	0
3	A	27	0	14	1	0
3	B	27	0	14	1	0
4	A	18	0	24	2	0
4	B	18	0	24	1	0
5	A	6	0	3	4	0
5	B	6	0	3	3	0
6	A	5	0	0	1	0
6	B	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	2	0	0	0	0
8	B	1	0	0	0	0
9	A	86	0	0	1	0
9	B	42	0	0	0	0
All	All	6102	0	5983	32	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 3.

All (32) 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:A:101:ARG:HA	5:A:504:PYZ:H5	1.35	1.05
1:A:216:ASP:HA	1:A:242:LEU:HD21	1.51	0.92
1:A:105:ARG:HH21	5:A:504:PYZ:HN1	1.16	0.89
1:A:43:THR:H	3:A:1002:AN2:H3B1	1.18	0.89
1:B:105:ARG:HH21	5:B:504:PYZ:HN1	1.27	0.82
1:A:101:ARG:HA	5:A:504:PYZ:C5	2.12	0.77
4:A:1006:GOL:H31	9:A:1124:HOH:O	1.87	0.73
1:B:105:ARG:NH2	5:B:504:PYZ:HN1	2.01	0.54
1:B:159:GLU:HB3	4:B:503:GOL:H2	1.89	0.54
1:A:105:ARG:NH2	5:A:504:PYZ:HN1	1.94	0.53
1:B:42:THR:OG1	5:B:504:PYZ:I4	2.92	0.53
1:B:253:SER:OG	1:B:351:VAL:HG13	2.10	0.51
1:A:134:GLU:HG2	1:A:143:THR:HG22	1.92	0.51
1:B:134:GLU:HG2	1:B:143:THR:HG22	1.93	0.51
1:A:42:THR:OG1	6:A:1007:PO4:O3	2.29	0.50
1:B:116:MET:HA	1:B:119:TRP:CE3	2.46	0.50
1:A:239:LEU:HD23	1:A:355:ALA:CB	2.42	0.49
1:B:36:ILE:HG22	1:B:49:ILE:HG22	1.94	0.49
1:A:208:ALA:O	1:A:403:GLY:HA3	2.13	0.47
1:A:41:GLY:HA3	1:A:44:TYR:O	2.15	0.47
1:A:215:LEU:O	1:A:242:LEU:CD2	2.63	0.46
1:B:232:GLY:HA2	3:B:502:AN2:H4'	1.98	0.46
1:A:356:LYS:NZ	1:B:389:GLU:CD	2.69	0.45
1:A:216:ASP:HA	1:A:242:LEU:CD2	2.34	0.44
1:B:208:ALA:O	1:B:403:GLY:HA3	2.17	0.44
1:A:253:SER:HA	4:A:1006:GOL:H12	2.00	0.43
1:B:41:GLY:HA3	1:B:44:TYR:O	2.19	0.42
1:A:230:LEU:HG	1:A:371:SER:HB2	2.01	0.41
1:A:116:MET:HA	1:A:119:TRP:CE3	2.56	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:115:ASP:OD1	1:A:257:HIS:NE2	2.54	0.41
1:A:284:VAL:HG13	1:A:290:SER:HB3	2.03	0.40
1:A:216:ASP:HB3	1:A:247:PHE:CE2	2.57	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	
1	A	383/393 (98%)	380 (99%)	3 (1%)	0	100	100
1	B	369/393 (94%)	367 (100%)	2 (0%)	0	100	100
All	All	752/786 (96%)	747 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	321/326 (98%)	320 (100%)	1 (0%)	92	96
1	B	315/326 (97%)	313 (99%)	2 (1%)	86	93
All	All	636/652 (98%)	633 (100%)	3 (0%)	88	94

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	226	LEU
1	B	111	THR
1	B	226	LEU

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

3 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	CSX	A	391[B]	-	3,6,7	0.87	0	1,6,8	3.35	1 (100%)
1	CSX	A	391[A]	-	3,6,7	0.82	0	1,6,8	0.63	0
1	CSX	B	391	1	3,6,7	0.72	0	1,6,8	1.79	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
1	CSX	A	391[B]	-	-	0/1/5/7	-
1	CSX	A	391[A]	-	-	0/1/5/7	-
1	CSX	B	391	1	-	0/1/5/7	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	391[B]	CSX	CA-CB-SG	3.35	120.67	113.36

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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 3 are monoatomic - leaving 13 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	GOL	B	505	-	5,5,5	0.06	0	5,5,5	0.22	0
6	PO4	B	506	-	4,4,4	2.53	1 (25%)	6,6,6	0.43	0
5	PYZ	B	504	-	6,6,6	0.84	0	2,7,7	0.32	0
3	AN2	B	502	-	24,29,29	2.38	4 (16%)	25,45,45	1.42	4 (16%)
4	GOL	A	1005	-	5,5,5	0.17	0	5,5,5	0.29	0
4	GOL	A	1006	-	5,5,5	0.16	0	5,5,5	0.33	0
4	GOL	B	501	-	5,5,5	0.12	0	5,5,5	0.18	0
4	GOL	A	1003	-	5,5,5	0.13	0	5,5,5	0.30	0
6	PO4	A	1007	-	4,4,4	2.52	1 (25%)	6,6,6	0.60	0
2	PG4	A	1001	-	6,6,12	0.24	0	5,5,11	0.19	0
3	AN2	A	1002	-	24,29,29	1.91	3 (12%)	25,45,45	1.43	5 (20%)
5	PYZ	A	504	-	6,6,6	1.01	0	2,7,7	0.56	0
4	GOL	B	503	-	5,5,5	0.10	0	5,5,5	0.26	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	GOL	B	505	-	-	2/4/4/4	-
5	PYZ	B	504	-	-	-	0/1/1/1
3	AN2	B	502	-	-	0/9/32/32	0/3/3/3
4	GOL	A	1005	-	-	2/4/4/4	-
4	GOL	A	1006	-	-	1/4/4/4	-
4	GOL	B	501	-	-	2/4/4/4	-
4	GOL	A	1003	-	-	1/4/4/4	-
2	PG4	A	1001	-	-	2/4/4/10	-
3	AN2	A	1002	-	-	0/9/32/32	0/3/3/3
5	PYZ	A	504	-	-	-	0/1/1/1
4	GOL	B	503	-	-	0/4/4/4	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	AN2	PB-O1B	9.58	1.61	1.46
3	A	1002	AN2	PB-O1B	7.25	1.57	1.46
6	B	506	PO4	P-O1	4.15	1.60	1.50
3	A	1002	AN2	PB-O2B	-4.10	1.45	1.56
6	A	1007	PO4	P-O1	4.09	1.60	1.50
3	B	502	AN2	PB-O2B	-3.68	1.46	1.56
3	B	502	AN2	PB-O3A	3.59	1.63	1.59
3	B	502	AN2	C5-C4	2.30	1.47	1.40
3	A	1002	AN2	C5-C4	2.23	1.46	1.40

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	502	AN2	N3-C2-N1	-3.76	122.79	128.68
3	A	1002	AN2	N3-C2-N1	-3.36	123.42	128.68
3	B	502	AN2	C4-C5-N7	-2.62	106.66	109.40
3	A	1002	AN2	C1'-N9-C4	-2.44	122.35	126.64
3	A	1002	AN2	N6-C6-N1	2.28	123.31	118.57
3	B	502	AN2	O4'-C1'-C2'	-2.26	103.63	106.93
3	B	502	AN2	O3'-C3'-C4'	-2.25	104.53	111.05
3	A	1002	AN2	C2-N1-C6	2.06	122.28	118.75
3	A	1002	AN2	O4'-C1'-C2'	-2.01	103.98	106.93

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	501	GOL	C1-C2-C3-O3
4	B	505	GOL	C1-C2-C3-O3
2	A	1001	PG4	O2-C3-C4-O3
4	A	1003	GOL	O1-C1-C2-C3
4	A	1005	GOL	O1-C1-C2-C3
4	B	501	GOL	O2-C2-C3-O3
4	B	505	GOL	O2-C2-C3-O3
4	A	1005	GOL	O1-C1-C2-O2
2	A	1001	PG4	C3-C4-O3-C5
4	A	1006	GOL	C1-C2-C3-O3

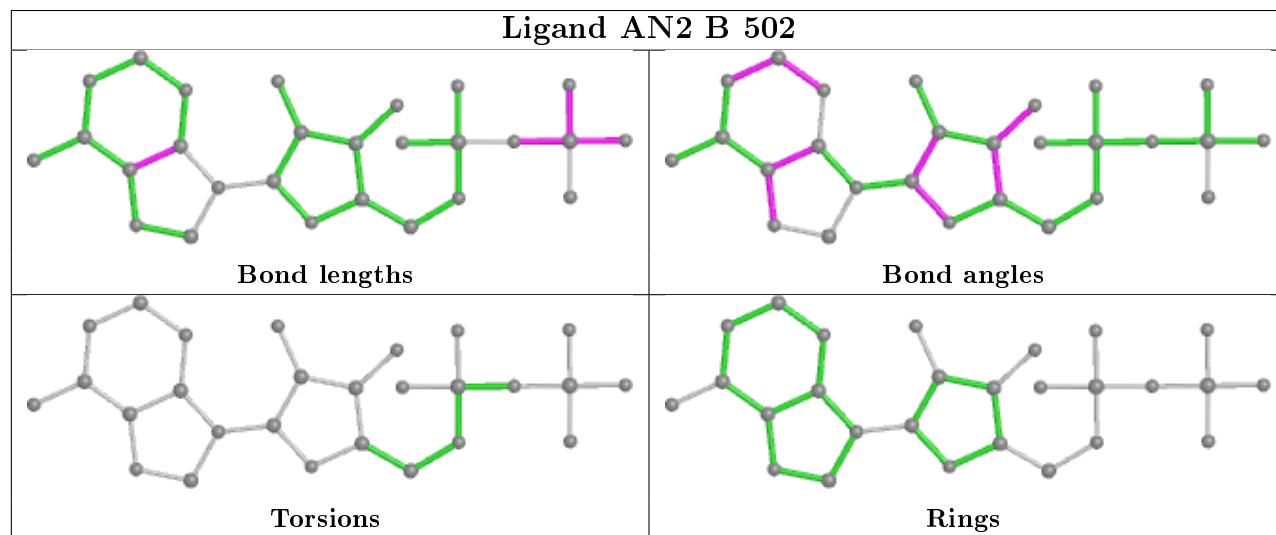
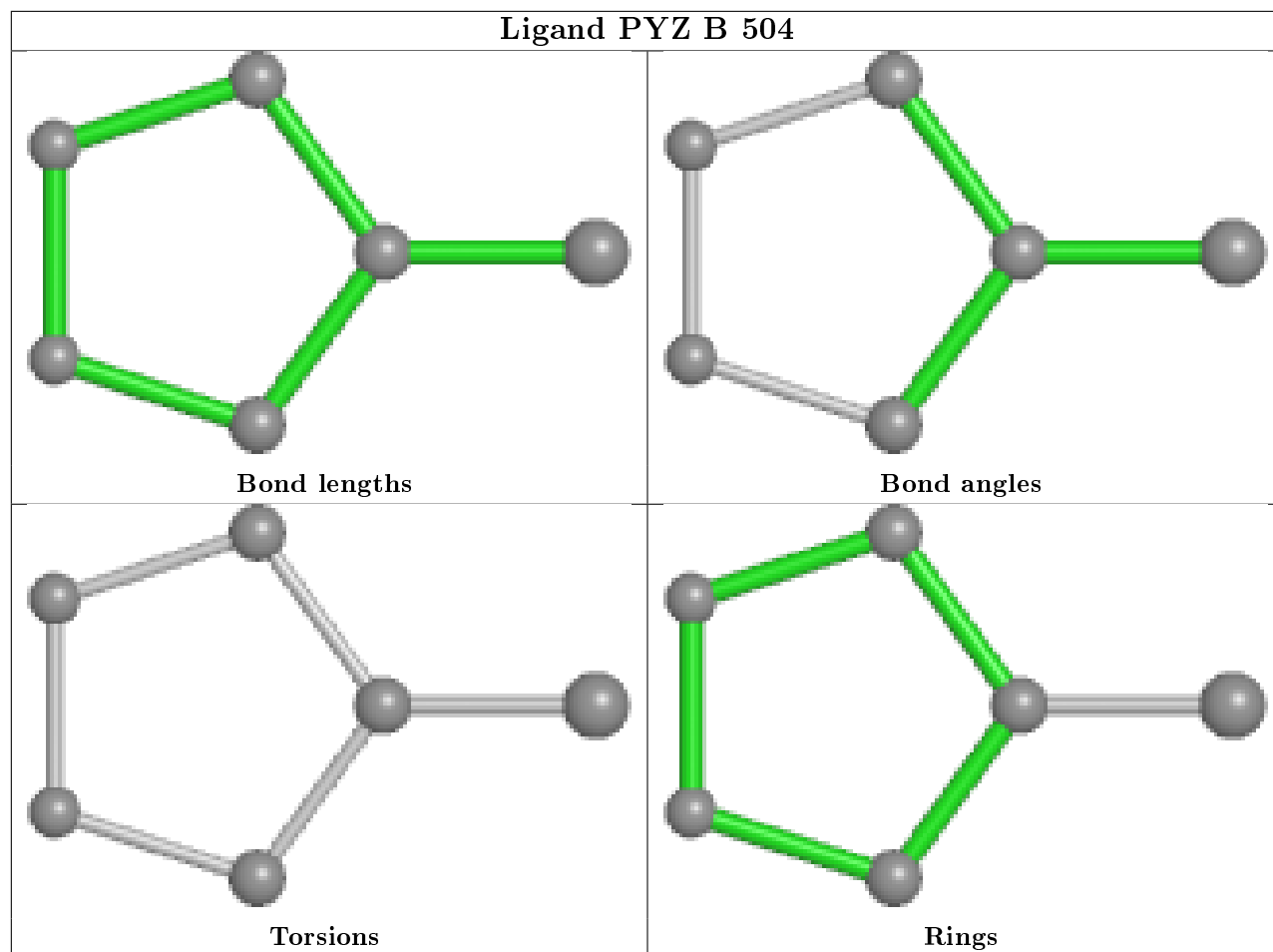
There are no ring outliers.

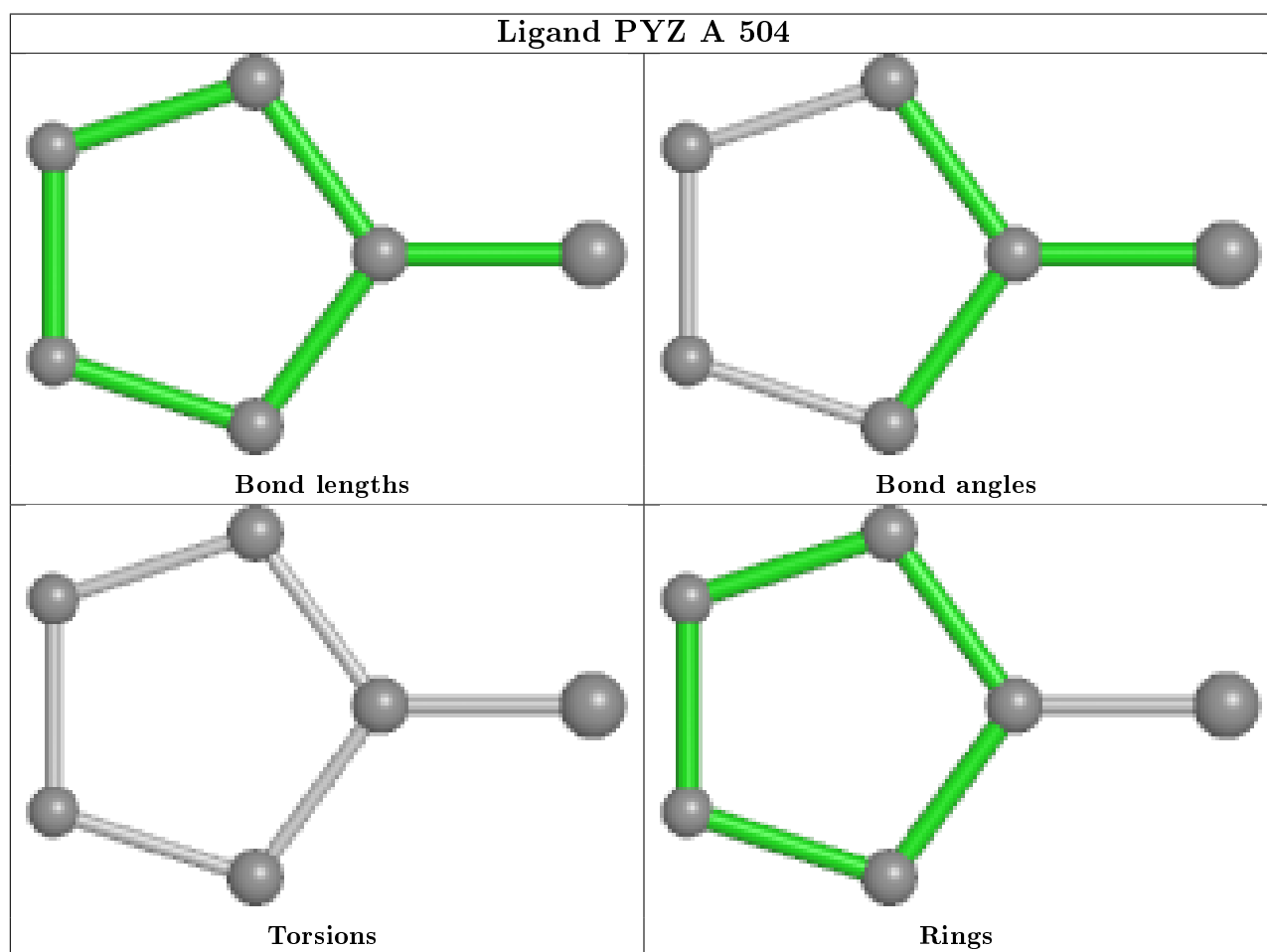
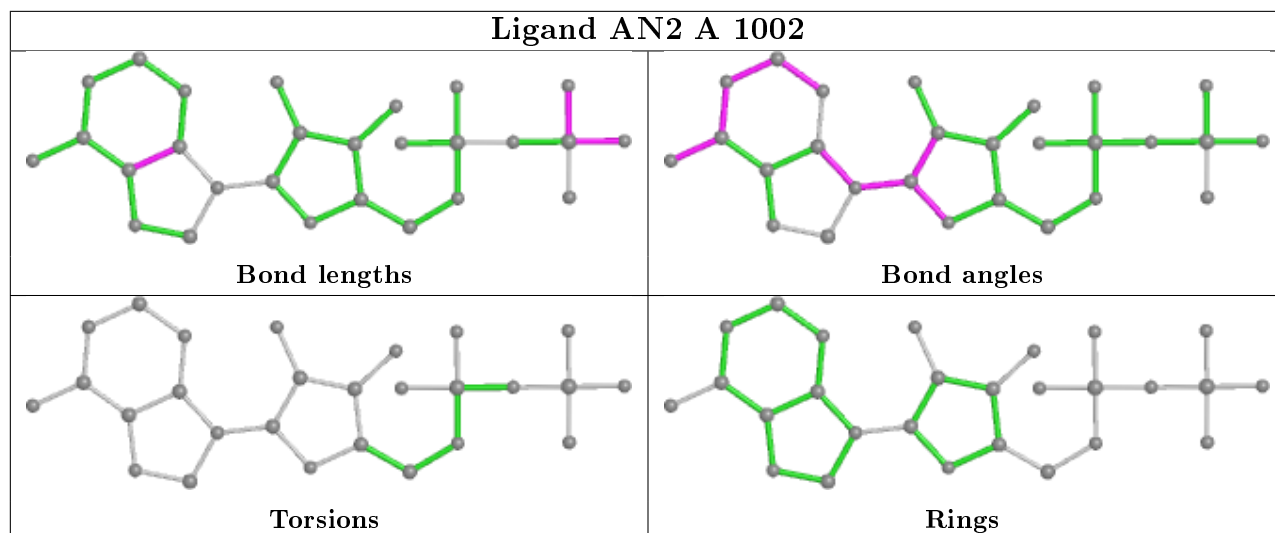
7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	504	PYZ	3	0
3	B	502	AN2	1	0
4	A	1006	GOL	2	0
6	A	1007	PO4	1	0
3	A	1002	AN2	1	0
5	A	504	PYZ	4	0
4	B	503	GOL	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







## 5.7 Other polymers [i](#)

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	A	385/393 (97%)	0.22	8 (2%) 63 74	45, 67, 100, 113	0
1	B	375/393 (95%)	0.70	46 (12%) 4 7	52, 95, 155, 194	0
All	All	760/786 (96%)	0.46	54 (7%) 16 23	45, 78, 148, 194	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	324	TYR	9.4
1	B	275	PHE	6.8
1	B	333	PHE	6.5
1	B	286	LYS	6.5
1	B	284	VAL	6.2
1	B	239	LEU	5.9
1	B	279	ASN	5.4
1	B	362	VAL	5.2
1	B	322	ILE	4.5
1	B	271	CYS	4.5
1	B	278	LYS	4.3
1	B	310	GLN	4.3
1	B	380	LEU	3.6
1	B	313	ILE	3.6
1	B	315	VAL	3.6
1	B	319	PHE	3.6
1	B	342	ARG	3.5
1	B	293	ARG	3.4
1	B	332	LYS	3.4
1	B	326	VAL	3.3
1	B	320	ASP	3.3
1	B	222	GLU	3.2
1	B	374	ILE	3.2
1	B	243	GLU	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	363	HIS	3.1
1	B	245	GLY	2.8
1	B	308	SER	2.8
1	A	319	PHE	2.7
1	B	270	PHE	2.7
1	B	263	PHE	2.6
1	B	360	SER	2.5
1	B	361	GLN	2.5
1	B	414	ASP	2.5
1	B	267	LEU	2.4
1	A	284	VAL	2.4
1	A	33	GLU	2.4
1	B	344	THR	2.4
1	B	221	GLY	2.3
1	B	169	VAL	2.3
1	A	244	ASP	2.3
1	B	377	ILE	2.3
1	B	225	ILE	2.2
1	B	328	ILE	2.2
1	A	281	GLY	2.2
1	B	336	LEU	2.2
1	B	289	LYS	2.2
1	B	246	ILE	2.1
1	A	144	PHE	2.1
1	B	385	PHE	2.1
1	A	313	ILE	2.1
1	A	328	ILE	2.1
1	B	240	LEU	2.1
1	B	376	LYS	2.1
1	B	325	ASN	2.0

## 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	CSX	B	391	7/8	0.86	0.16	102,105,108,108	0
1	CSX	A	391[B]	7/8	0.95	0.15	61,62,65,65	4
1	CSX	A	391[A]	7/8	0.95	0.15	61,62,65,65	4

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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(Å <sup>2</sup> )	Q<0.9
7	CL	A	1008	1/1	0.50	0.21	116,116,116,116	0
4	GOL	B	505	6/6	0.76	0.17	133,133,133,133	0
4	GOL	A	1003	6/6	0.76	0.33	92,93,93,93	0
4	GOL	A	1005	6/6	0.79	0.24	76,79,79,80	0
4	GOL	B	503	6/6	0.88	0.37	95,96,96,96	0
7	CL	A	1009	1/1	0.88	0.11	86,86,86,86	0
4	GOL	A	1006	6/6	0.89	0.12	85,85,85,86	0
8	IOD	B	507	1/1	0.89	0.07	167,167,167,167	0
5	PYZ	A	504	6/6	0.92	0.19	67,68,69,69	1
2	PG4	A	1001	7/13	0.92	0.20	82,83,83,84	0
4	GOL	B	501	6/6	0.93	0.35	95,96,96,96	0
6	PO4	B	506	5/5	0.94	0.19	116,116,117,117	0
5	PYZ	B	504	6/6	0.97	0.18	95,96,96,96	6
6	PO4	A	1007	5/5	0.97	0.20	78,78,79,80	0
3	AN2	B	502	27/27	0.97	0.13	76,78,81,81	0
3	AN2	A	1002	27/27	0.99	0.13	49,52,56,57	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.