



Full wwPDB X-ray Structure Validation Report

Mar 1, 2014 – 12:49 AM GMT

PDB ID : 2IBK
Title : Bypass of Major Benzopyrene-dG Adduct by Y-Family DNA Polymerase with Unique Structural Gap
Authors : Bauer, J.; Ling, H.; Sayer, J.M.; Xing, G.; Yagi, H.; Jerina, D.M.
Deposited on : 2006-09-11
Resolution : 2.25 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

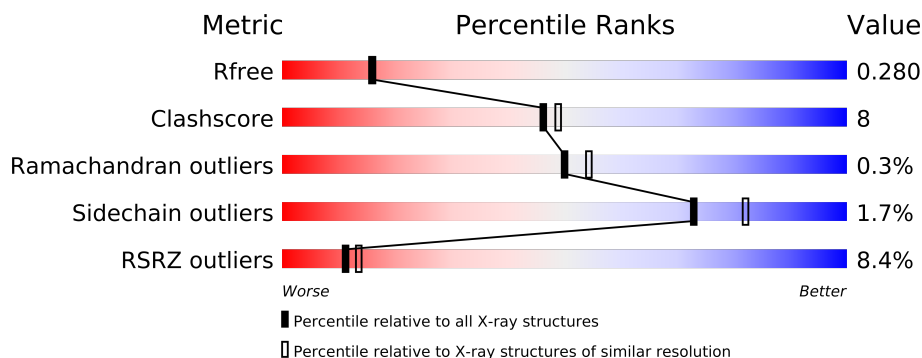
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.25 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}	66092	1108 (2.28-2.24)
Clashscore	79885	1326 (2.28-2.24)
Ramachandran outliers	78287	1291 (2.28-2.24)
Sidechain outliers	78261	1291 (2.28-2.24)
RSRZ outliers	66119	1110 (2.28-2.24)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	D	14	
2	E	17	
3	A	352	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	EDO	A	407	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 3624 atoms, of which 0 are hydrogen and 0 are deuterium.

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 5'-D(*GP*GP*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	14	Total	C	N	O	P	0	0	0
			295	140	61	81	13			

- Molecule 2 is a DNA chain called 5'-D(*TP*CP*AP*TP*GP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	17	Total	C	N	O	P	0	0	0
			333	162	54	101	16			

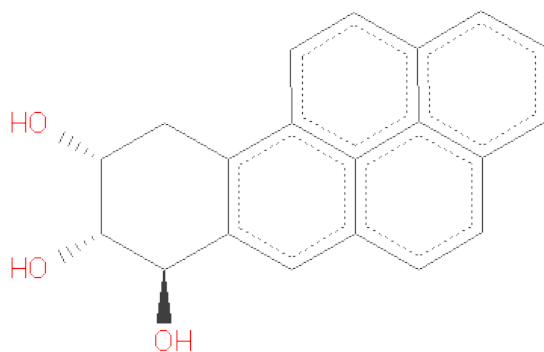
- Molecule 3 is a protein called DNA polymerase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	341	Total	C	N	O	S	0	0	0
			2743	1760	472	504	7			

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Ca	0	0
			2	2		

- Molecule 5 is 1,2,3-TRIHYDROXY-1,2,3,4-TETRAHYDROBENZO[A]PYRENE (three-letter code: BAP) (formula: C₂₀H₁₆O₃).



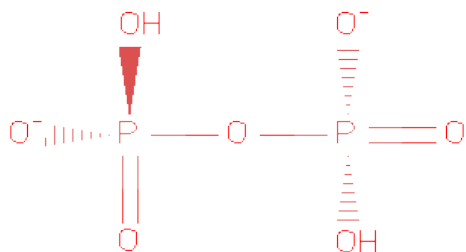
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			23	20	3		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



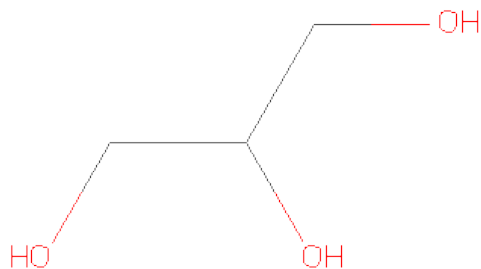
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is PYROPHOSPHATE 2- (three-letter code: POP) (formula: $\text{H}_2\text{O}_7\text{P}_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	P	0	0
			9	7	2		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $\text{C}_3\text{H}_8\text{O}_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	136	Total 136	O 136	0	0
9	D	31	Total 31	O 31	0	0
9	E	34	Total 34	O 34	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	96.71Å 100.47Å 52.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.20 – 2.25 48.36 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.2 (17.20-2.25) 99.3 (48.36-2.25)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.48 (at 2.24Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.250 , 0.272 0.263 , 0.280	Depositor DCC
R_{free} test set	728 reflections (2.95%)	DCC
Wilson B-factor (Å ²)	56.6	Xtriage
Anisotropy	0.338	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 52.1	EDS
Estimated twinning fraction	0.014 for k,h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 24724 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3624	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CA, EDO, BAP, POP

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	D	0.63	0/333	0.86	0/515
2	E	0.73	0/370	1.12	4/566 (0.7%)
3	A	0.45	0/2782	0.67	0/3736
All	All	0.50	0/3485	0.76	4/4817 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	105	DT	O3'-P-O5'	6.34	116.05	104.00
2	E	106	DG	O5'-P-OP1	-5.61	100.65	105.70
2	E	104	DA	C4-N9-C1'	5.38	135.98	126.30
2	E	104	DA	C8-N9-C1'	-5.12	118.48	127.70

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	295	0	159	0	0
2	E	333	0	191	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	2743	0	2889	48	0
4	A	2	0	0	0	0
5	E	23	0	15	1	0
6	A	8	0	12	0	0
6	E	4	0	6	0	0
7	A	9	0	0	0	0
8	E	6	0	8	0	0
9	A	136	0	0	0	0
9	D	31	0	0	0	0
9	E	34	0	0	0	0
All	All	3624	0	3280	55	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

All (55) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:157:MET:CE	3:A:166:ILE:HD11	1.99	0.91
3:A:300:ARG:HH11	3:A:300:ARG:HG2	1.42	0.83
3:A:153:ILE:O	3:A:157:MET:HG3	1.87	0.75
3:A:256:ARG:HB2	3:A:323:LEU:HD22	1.69	0.75
3:A:115:VAL:HB	3:A:120:GLU:HB3	1.70	0.73
3:A:273:TYR:HA	3:A:276:LEU:HD12	1.74	0.70
3:A:267:ARG:O	3:A:271:GLU:HG3	1.91	0.69
3:A:300:ARG:CG	3:A:300:ARG:HH11	2.05	0.69
2:E:108:DA:H2'	3:A:336:ARG:HH21	1.59	0.67
2:E:113:DT:OP1	2:E:113:DT:H4'	1.97	0.65
3:A:245:ILE:HG23	3:A:275:LYS:HE2	1.79	0.63
2:E:105:DT:H5'	2:E:106:DG:H4'	1.80	0.62
3:A:242:ARG:HH11	3:A:242:ARG:HG3	1.64	0.62
3:A:283:ALA:HB2	3:A:339:LYS:HD2	1.82	0.62
3:A:277:ASP:O	3:A:278:LYS:HB2	2.00	0.61
3:A:157:MET:HE2	3:A:166:ILE:HD11	1.81	0.61
3:A:87:ARG:HD2	3:A:136:GLU:OE2	2.01	0.60
2:E:108:DA:H2'	3:A:336:ARG:NH2	2.18	0.59
2:E:103:DC:H2''	2:E:104:DA:OP2	2.03	0.59
3:A:256:ARG:HB2	3:A:323:LEU:CD2	2.33	0.58
3:A:274:TYR:HD2	3:A:274:TYR:O	1.86	0.57
3:A:36:ARG:HH12	3:A:253:ARG:HA	1.70	0.56
3:A:251:MET:HE1	3:A:261:ILE:HG12	1.88	0.55
3:A:157:MET:HE1	3:A:166:ILE:HD11	1.86	0.54
3:A:49:GLU:HA	3:A:52:LYS:HE2	1.91	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:46:ALA:HB1	3:A:50:ALA:HB3	1.92	0.52
3:A:9:ASP:O	3:A:10:TYR:C	2.49	0.51
3:A:36:ARG:HG3	3:A:37:PHE:HD2	1.75	0.51
3:A:153:ILE:HG22	3:A:157:MET:HE1	1.92	0.50
3:A:8:PHE:N	3:A:8:PHE:CD1	2.79	0.50
3:A:270:GLU:OE1	3:A:308:LYS:HD3	2.11	0.50
3:A:9:ASP:O	3:A:14:GLN:NE2	2.44	0.49
3:A:166:ILE:HG23	3:A:170:GLU:HB3	1.93	0.49
2:E:106:DG:N3	2:E:106:DG:H2'	2.28	0.48
3:A:199:ILE:HG23	3:A:204:ASP:HB2	1.96	0.48
5:E:1106:BAP:H10	3:A:274:TYR:CG	2.49	0.47
2:E:102:DT:H2''	2:E:103:DC:OP2	2.15	0.47
2:E:111:DC:H2''	2:E:112:DT:H71	1.97	0.46
3:A:289:VAL:HB	3:A:332:ARG:HB2	1.98	0.46
3:A:217:ILE:HD11	3:A:222:ALA:HA	1.98	0.45
2:E:103:DC:O2	3:A:42:ALA:HB2	2.17	0.45
3:A:309:GLU:H	3:A:309:GLU:CD	2.20	0.44
3:A:300:ARG:NH1	3:A:300:ARG:HG2	2.21	0.44
3:A:242:ARG:NH1	3:A:242:ARG:HG3	2.31	0.43
3:A:36:ARG:HG2	3:A:331:ARG:NH2	2.32	0.43
3:A:27:PRO:HG2	3:A:71:ALA:HA	2.00	0.43
3:A:10:TYR:HD1	3:A:13:ALA:HB3	1.83	0.43
3:A:300:ARG:NH1	3:A:301:THR:O	2.52	0.43
2:E:106:DG:H5''	3:A:76:MET:CE	2.49	0.43
3:A:319:LEU:HD13	3:A:333:ILE:HD11	2.01	0.43
3:A:12:TYR:HB2	3:A:45:THR:HG22	2.02	0.42
2:E:106:DG:H5''	3:A:76:MET:HE1	2.03	0.41
3:A:9:ASP:O	3:A:11:PHE:N	2.53	0.41
3:A:59:ILE:HD12	3:A:63:GLU:HG2	2.01	0.41
2:E:102:DT:C2'	2:E:102:DT:O2	2.69	0.40

There are no symmetry-related clashes.

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
3	A	339/352 (96%)	326 (96%)	12 (4%)	1 (0%)	50 54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	10	TYR

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
3	A	300/309 (97%)	295 (98%)	5 (2%)	73 82

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	120	GLU
3	A	238	ARG
3	A	248	ILE
3	A	274	TYR
3	A	300	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	188	ASN
3	A	285	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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$
7	POP	A	405	4	8,8,8	2.94	5 (62%)	13,13,13	2.00	3 (23%)
6	EDO	A	407	-	3,3,3	0.63	0	2,2,2	0.37	0
6	EDO	A	408	-	3,3,3	0.59	0	2,2,2	0.37	0
5	BAP	E	1106	2	27,27,27	1.53	7 (25%)	42,42,42	1.03	2 (4%)
6	EDO	E	406	-	3,3,3	0.74	0	2,2,2	0.19	0
8	GOL	E	409	-	5,5,5	0.17	0	5,5,5	0.38	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
7	POP	A	405	4	-	0/6/6/6	0/0/0/0
6	EDO	A	407	-	-	0/1/1/1	0/0/0/0
6	EDO	A	408	-	-	0/1/1/1	0/0/0/0
5	BAP	E	1106	2	-	0/0/16/16	0/0/5/5
6	EDO	E	406	-	-	0/1/1/1	0/0/0/0
8	GOL	E	409	-	-	0/4/4/4	0/0/0/0

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	405	POP	P2-O6	4.50	1.65	1.50
7	A	405	POP	P1-O1	3.62	1.62	1.50
7	A	405	POP	P1-O3	3.57	1.62	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1106	BAP	C6-C7	3.35	1.44	1.37
7	A	405	POP	P1-O	3.22	1.67	1.60
5	E	1106	BAP	C7-C1'	-2.66	1.47	1.51
5	E	1106	BAP	C2-C3	2.57	1.42	1.36
5	E	1106	BAP	C2-C1	2.39	1.42	1.36
7	A	405	POP	P2-O	2.30	1.65	1.60
5	E	1106	BAP	C8-C7	2.24	1.39	1.36
5	E	1106	BAP	C10-C9	2.07	1.40	1.34
5	E	1106	BAP	C5-C4	2.02	1.40	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	405	POP	P2-O-P1	-4.06	119.53	131.74
7	A	405	POP	O2-P1-O	3.96	116.89	104.29
5	E	1106	BAP	C3'-C2'-C1'	3.15	115.28	110.64
7	A	405	POP	O-P1-O1	-2.56	99.96	107.57
5	E	1106	BAP	C8-C7-C1'	-2.45	116.53	120.33

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 95th 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(Å²)	Q<0.9	
1	D	14/14 (100%)	0.30	0	100	100	48, 58, 66, 67	0
2	E	17/17 (100%)	0.17	0	100	100	46, 66, 89, 90	0
3	A	341/352 (96%)	0.86	30 (8%)	10	12	33, 59, 79, 85	0
All	All	372/383 (97%)	0.81	30 (8%)	11	14	33, 59, 79, 90	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	274	TYR	4.6
3	A	331	ARG	4.4
3	A	2	ILE	4.0
3	A	177	GLU	3.6
3	A	252	LYS	3.6
3	A	253	ARG	3.3
3	A	10	TYR	3.2
3	A	199	ILE	3.1
3	A	263	PRO	2.9
3	A	169	GLU	2.9
3	A	55	VAL	2.8
3	A	188	ASN	2.7
3	A	48	TYR	2.7
3	A	116	ARG	2.6
3	A	293	LEU	2.6
3	A	312	TYR	2.6
3	A	123	ASN	2.5
3	A	258	LEU	2.3
3	A	250	THR	2.3
3	A	175	ILE	2.3
3	A	112	SER	2.2
3	A	270	GLU	2.2
3	A	67	ILE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	118	TYR	2.2
3	A	159	LYS	2.1
3	A	277	ASP	2.1
3	A	62	VAL	2.1
3	A	37	PHE	2.1
3	A	327	GLU	2.0
3	A	68	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

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. 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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	EDO	A	407	4/4	0.86	33.44	86,86,86,86	0
6	EDO	E	406	4/4	0.18	1.98	68,71,72,73	0
7	POP	A	405	9/9	0.31	1.87	65,67,68,69	0
6	EDO	A	408	4/4	0.27	0.03	86,87,88,88	0
8	GOL	E	409	6/6	0.16	-0.29	79,81,81,81	0
5	BAP	E	1106	23/23	0.16	-0.45	51,53,55,55	0
4	CA	A	401	1/1	0.05	-1.83	54,54,54,54	0
4	CA	A	402	1/1	0.05	-6.18	66,66,66,66	0

6.5 Other polymers ⓘ

There are no such residues in this entry.