wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2021 – 12:07 pm BST

PDB ID	:	7A0H
Title	:	Structure of homodimeric actin capping protein alpha subunit from Plasmodium berghei
Authors	:	Bendes, A.A.; Kursula, P.; Kursula, I.
Deposited on	:	2020-08-09
Resolution	:	2.22 Å(reported)

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

We welcome your comments at *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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

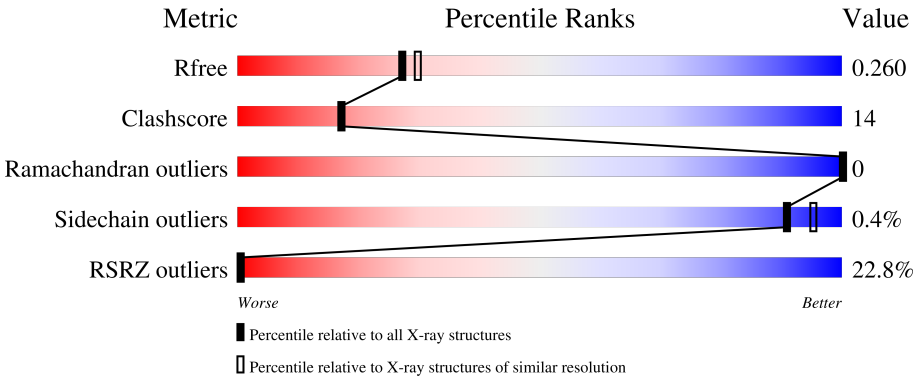
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.22 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	294	4% 72% 24% .
1	B	294	38% 62% 24% . 13%

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
2	BR	A	302	-	-	X	-
2	BR	A	321	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9136 atoms, of which 4494 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 F-actin-capping protein subunit alpha.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	285	Total	C	H	N	O	S	0	6	0
			4689	1500	2345	392	443	9			
1	B	255	Total	C	H	N	O	S	0	0	0
			4195	1346	2107	349	387	6			

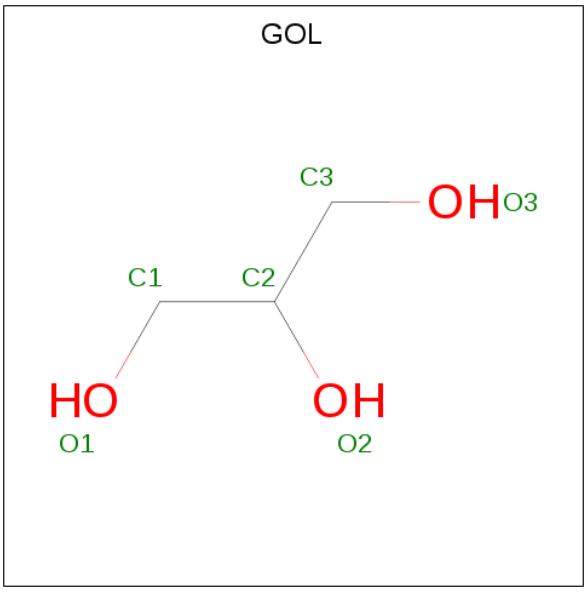
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	288	ALA	-	expression tag	UNP A0A509AR49
A	289	HIS	-	expression tag	UNP A0A509AR49
A	290	HIS	-	expression tag	UNP A0A509AR49
A	291	HIS	-	expression tag	UNP A0A509AR49
A	292	HIS	-	expression tag	UNP A0A509AR49
A	293	HIS	-	expression tag	UNP A0A509AR49
A	294	HIS	-	expression tag	UNP A0A509AR49
B	288	ALA	-	expression tag	UNP A0A509AR49
B	289	HIS	-	expression tag	UNP A0A509AR49
B	290	HIS	-	expression tag	UNP A0A509AR49
B	291	HIS	-	expression tag	UNP A0A509AR49
B	292	HIS	-	expression tag	UNP A0A509AR49
B	293	HIS	-	expression tag	UNP A0A509AR49
B	294	HIS	-	expression tag	UNP A0A509AR49

- Molecule 2 is BROMIDE ION (three-letter code: BR) (formula: Br).

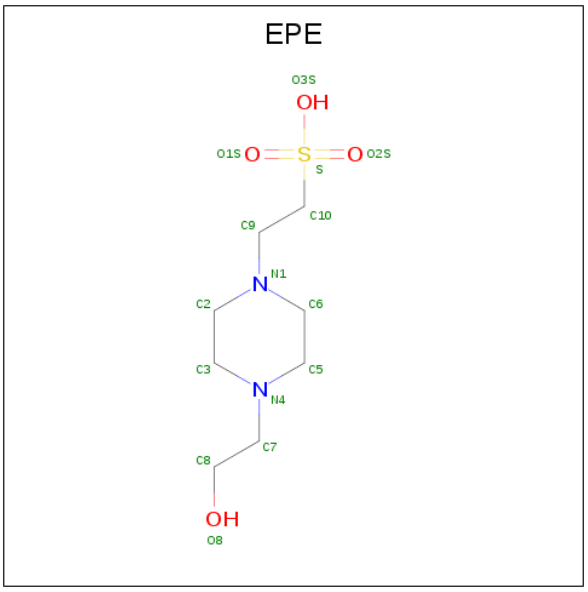
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total	Br	0	0
			21	21		
2	B	4	Total	Br	0	0
			4	4		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	B	1	Total	C	H	N	O	S	0	1
			64	16	34	4	8	2		

- Molecule 5 is water.

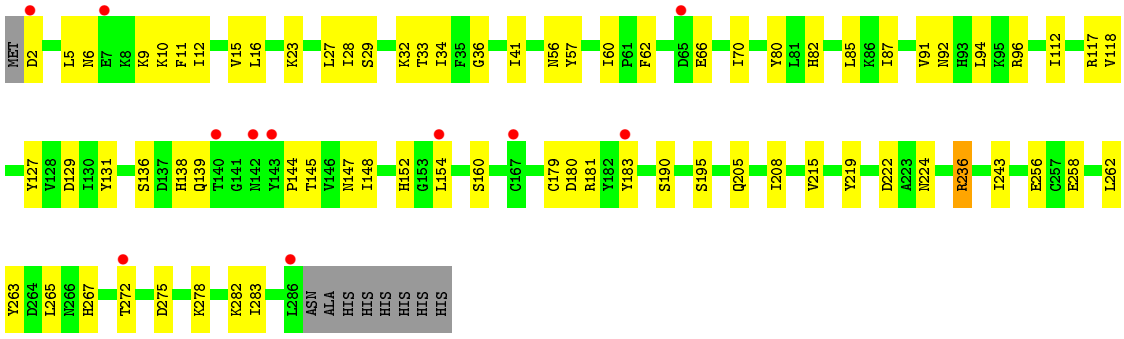
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	141	Total O 141 141	0	0
5	B	8	Total O 8 8	0	0

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.

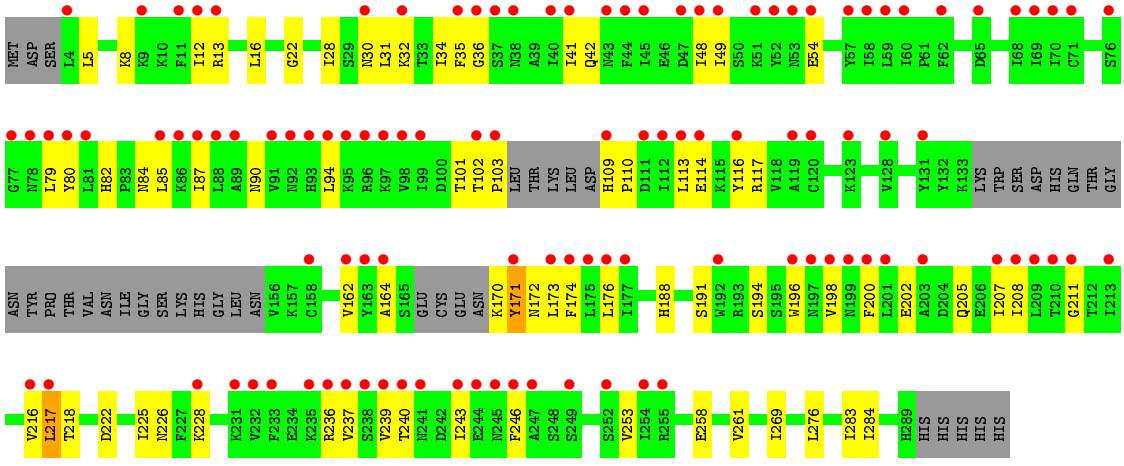
- Molecule 1: F-actin-capping protein subunit alpha

Chain A:



- Molecule 1: F-actin-capping protein subunit alpha

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	164.49Å 35.17Å 115.44Å 90.00° 104.44° 90.00°	Depositor
Resolution (Å)	74.20 – 2.22 79.65 – 2.22	Depositor EDS
% Data completeness (in resolution range)	99.2 (74.20-2.22) 99.4 (79.65-2.22)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.06 (at 2.22Å)	Xtriage
Refinement program	PHENIX dev-3928	Depositor
R, R_{free}	0.207 , 0.257 0.209 , 0.260	Depositor DCC
R_{free} test set	1594 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	49.5	Xtriage
Anisotropy	0.608	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 80.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9136	wwPDB-VP
Average B, all atoms (Å^2)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.44% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GOL, BR, EPE

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	0.93	9/2412 (0.4%)	0.90	4/3269 (0.1%)
1	B	0.51	0/2128	0.75	3/2878 (0.1%)
All	All	0.76	9/4540 (0.2%)	0.84	7/6147 (0.1%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	258	GLU	CB-CG	7.77	1.67	1.52
1	A	195	SER	CA-CB	7.14	1.63	1.52
1	A	131	TYR	CD2-CE2	-6.27	1.29	1.39
1	A	91	VAL	CB-CG1	-5.69	1.41	1.52
1	A	179[A]	CYS	CB-SG	5.56	1.91	1.82

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	217	LEU	CB-CG-CD1	9.37	126.93	111.00
1	B	217	LEU	CB-CG-CD2	-8.25	96.97	111.00
1	A	236	ARG	NE-CZ-NH1	7.12	123.86	120.30
1	A	180[A]	ASP	CB-CG-OD1	6.58	124.22	118.30
1	A	180[B]	ASP	CB-CG-OD1	6.58	124.22	118.30

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	2344	2345	2324	49	1
1	B	2088	2107	2107	81	0
2	A	21	0	0	10	0
2	B	4	0	0	2	0
3	A	6	8	8	0	0
4	B	30	34	34	1	0
5	A	141	0	0	4	0
5	B	8	0	0	2	0
All	All	4642	4494	4473	125	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

The worst 5 of 125 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)
2:A:302:BR:BR	5:A:508:HOH:O	2.07	1.27
1:A:256:GLU:OE1	2:A:318:BR:BR	2.12	1.23
1:A:265:LEU:HD12	1:B:284:ILE:HD11	1.44	1.00
1:A:112:ILE:HD11	1:A:243:ILE:HD11	1.55	0.87
1:B:172:ASN:HB2	1:B:198:VAL:O	1.73	0.87

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:129:ASP:OD2	1:A:236:ARG:NH2[1_545]	2.18	0.02

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	289/294 (98%)	278 (96%)	11 (4%)	0	100	100
1	B	247/294 (84%)	239 (97%)	8 (3%)	0	100	100
All	All	536/588 (91%)	517 (96%)	19 (4%)	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	272/275 (99%)	270 (99%)	2 (1%)	84	91
1	B	239/275 (87%)	239 (100%)	0	100	100
All	All	511/550 (93%)	509 (100%)	2 (0%)	91	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	ILE
1	A	87	ILE

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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 28 ligands modelled in this entry, 25 are monoatomic - leaving 3 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	EPE	B	305[B]	-	15,15,15	0.95	1 (6%)	18,20,20	2.17	6 (33%)
4	EPE	B	305[A]	-	15,15,15	0.79	1 (6%)	18,20,20	2.08	6 (33%)
3	GOL	A	322	-	5,5,5	2.31	2 (40%)	5,5,5	0.80	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	EPE	B	305[B]	-	-	3/9/19/19	0/1/1/1
4	EPE	B	305[A]	-	-	5/9/19/19	0/1/1/1
3	GOL	A	322	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	322	GOL	C1-C2	3.75	1.67	1.51
4	B	305[B]	EPE	C10-S	3.07	1.81	1.77
4	B	305[A]	EPE	C10-S	2.53	1.81	1.77
3	A	322	GOL	O2-C2	2.40	1.50	1.43

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	305[A]	EPE	C7-N4-C5	4.57	122.93	111.23
4	B	305[B]	EPE	C9-N1-C2	4.28	122.17	111.23
4	B	305[B]	EPE	C7-N4-C3	3.74	120.81	111.23
4	B	305[B]	EPE	C5-N4-C3	3.72	117.20	108.83
4	B	305[A]	EPE	C5-N4-C3	3.62	116.98	108.83

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	322	GOL	O1-C1-C2-O2
3	A	322	GOL	O1-C1-C2-C3
4	B	305[A]	EPE	C8-C7-N4-C5
4	B	305[A]	EPE	S-C10-C9-N1
4	B	305[A]	EPE	C9-C10-S-O1S

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	305[B]	EPE	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^{th} 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^2)	Q<0.9
1	A	285/294 (96%)	0.62	11 (3%) 39 37	37, 65, 136, 170	0
1	B	255/294 (86%)	2.01	112 (43%) 0 0	44, 151, 189, 224	0
All	All	540/588 (91%)	1.28	123 (22%) 0 0	37, 105, 180, 224	0

The worst 5 of 123 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	98	VAL	8.6
1	B	200	PHE	8.4
1	B	163	TYR	8.4
1	B	41	ILE	7.5
1	B	13	ARG	7.3

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 monosaccharides 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^{th} 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^2)	Q<0.9
2	BR	A	312	1/1	0.63	0.10	123,123,123,123	1
4	EPE	B	305[A]	15/15	0.76	0.26	88,106,125,130	32
4	EPE	B	305[B]	15/15	0.76	0.26	89,111,125,133	32
2	BR	A	304	1/1	0.77	0.25	91,91,91,91	1
2	BR	A	316	1/1	0.78	0.19	105,105,105,105	1
2	BR	A	319	1/1	0.78	0.15	123,123,123,123	1
2	BR	A	310	1/1	0.80	0.23	65,65,65,65	1
2	BR	A	314	1/1	0.80	0.16	90,90,90,90	1
2	BR	A	317	1/1	0.81	0.30	61,61,61,61	1
2	BR	B	303	1/1	0.82	0.18	88,88,88,88	1
2	BR	B	304	1/1	0.82	0.09	151,151,151,151	1
2	BR	A	313	1/1	0.83	0.16	79,79,79,79	1
2	BR	A	315	1/1	0.85	0.22	88,88,88,88	1
2	BR	A	320	1/1	0.85	0.16	105,105,105,105	1
3	GOL	A	322	6/6	0.87	0.23	67,81,96,96	0
2	BR	A	318	1/1	0.89	0.21	82,82,82,82	1
2	BR	A	302	1/1	0.89	0.34	58,58,58,58	1
2	BR	A	303	1/1	0.90	0.14	65,65,65,65	1
2	BR	B	301	1/1	0.92	0.09	96,96,96,96	1
2	BR	B	302	1/1	0.93	0.06	125,125,125,125	1
2	BR	A	307	1/1	0.93	0.19	87,87,87,87	1
2	BR	A	311	1/1	0.93	0.15	74,74,74,74	1
2	BR	A	321	1/1	0.94	0.15	110,110,110,110	1
2	BR	A	308	1/1	0.95	0.10	70,70,70,70	1
2	BR	A	305	1/1	0.95	0.24	69,69,69,69	1
2	BR	A	309	1/1	0.96	0.16	64,64,64,64	1
2	BR	A	301	1/1	0.96	0.19	65,65,65,65	1
2	BR	A	306	1/1	0.98	0.19	56,56,56,56	1

6.5 Other polymers (i)

There are no such residues in this entry.