W O R L D W I D E

wwPDB

P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2014 – 11:04 AM GMT

PDB ID	:	4BEY
Title	:	Night blindness causing G90D rhodopsin in complex with GaCT2 peptide
Authors	:	Singhal, A.; Ostermaier, M.K.; Vishnivetskiy, S.A.; Panneels, V.; Homan, K.T.; Tesmer, J.J.G.; Veprintsev, D.; Deupi, X.; Gurevich, V.V.; Schertler, G.F.X.; Standfuss, J.
Deposited on	:	2013-03-12
Resolution	:	2.90 Å(reported)

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.90 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	Percentile Ranks	Value
Rfree		0.261
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		5.2%
RSRZ outliers		3.5%

Worse .... *Better*

▌ Percentile relative to all X-ray structures

❙ Percentile relative to X-ray structures of similar resolution
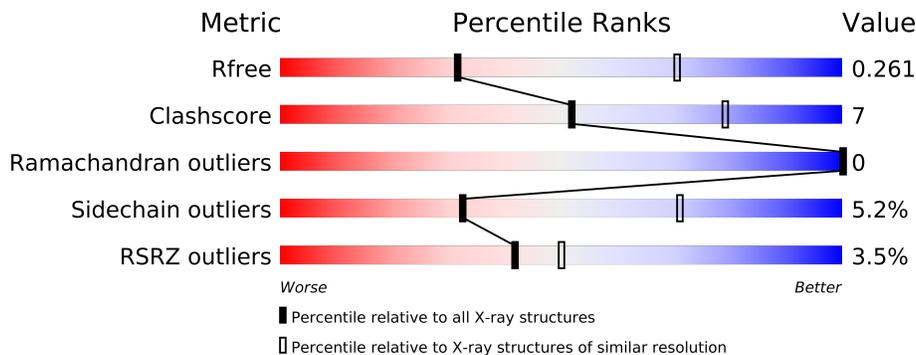
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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	A	349	
2	B	11	

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

Mol	Type	Chain	Res	Geometry	Electron density
3	SO4	A	1327	X	X
5	PLM	A	1329	-	X
7	BOG	A	1334	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 2810 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 protein called RHODOPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	327	2606	1737	400	441	28	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ACE	-	EXPRESSION TAG	UNP P02699
A	2	CYS	ASN	ENGINEERED MUTATION	UNP P02699
A	90	ASP	GLY	ENGINEERED MUTATION	UNP P02699
A	282	CYS	ASP	ENGINEERED MUTATION	UNP P02699

- Molecule 2 is a protein called GUANINE NUCLEOTIDE-BINDING PROTEIN G(T) SUB-UNIT ALPHA-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	11	88	57	13	17	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	341	LEU	LYS	ENGINEERED MUTATION	UNP P04695

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
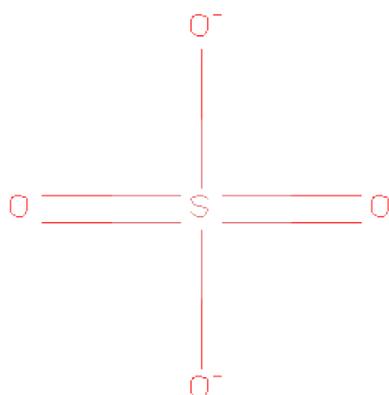
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
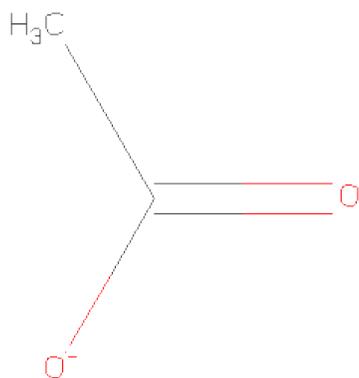
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is PALMITIC ACID (three-letter code: PLM) (formula: $C_{16}H_{32}O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 17	C 16	O 1	0	0

- Molecule 6 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	4	Total 50	C 28	N 2	O 20	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ACE	-	EXPRESSION TAG	UNP P02699
A	2	CYS	ASN	ENGINEERED MUTATION	UNP P02699
A	90	ASP	GLY	ENGINEERED MUTATION	UNP P02699
A	282	CYS	ASP	ENGINEERED MUTATION	UNP P02699

- Molecule 7 is SUGAR (B-OCTYLGLUCOSIDE) (three-letter code: BOG) (formula: $C_{14}H_{28}O_6$).
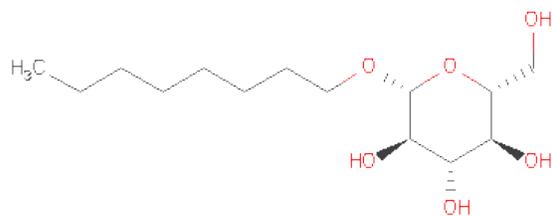
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 20	C 14	O 6	0	0

• Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	20	Total 20	O 20	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: RHODOPSIN

Chain A:
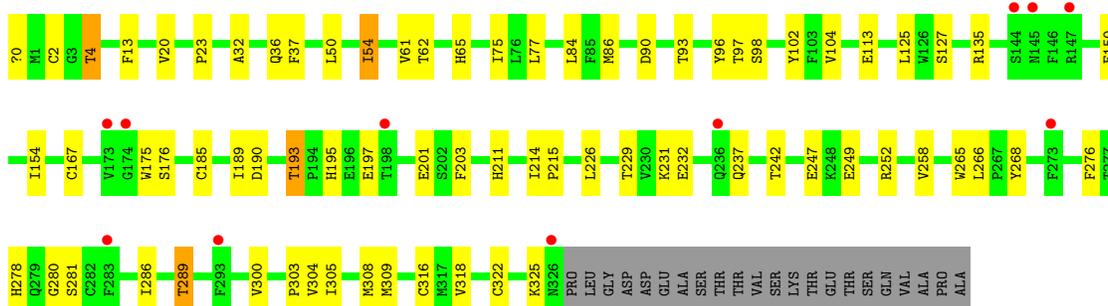
• Molecule 2: GUANINE NUCLEOTIDE-BINDING PROTEIN G(T) SUBUNIT ALPHA-1

Chain B:
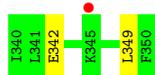
4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	242.44Å 242.44Å 110.36Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.26 – 2.90 29.26 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.26-2.90) 100.0 (29.26-2.90)	Depositor EDS
R_{merge}	0.27	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.71 (at 2.90Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8.2_1309)	Depositor
R, R_{free}	0.234 , 0.259 0.233 , 0.261	Depositor DCC
R_{free} test set	1375 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	60.7	Xtriage
Anisotropy	0.628	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 27510 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	2810	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.25% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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: BMA, NAG, ACE, SO4, ACT, PLM, BOG, MAN

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	0.53	0/2691	0.71	1/3668 (0.0%)
2	B	0.64	0/88	0.60	0/115
All	All	0.53	0/2779	0.70	1/3783 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	0	ACE	C-N-CA	9.24	144.81	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	TRP	Peptide

5.2 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 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	A	2606	0	2578	40	0
2	B	88	0	89	1	0
3	A	5	0	0	0	0
4	A	4	0	3	0	0
5	A	17	0	31	0	0
6	A	50	0	43	0	1
7	A	20	0	28	2	0
8	A	20	0	0	1	0
All	All	2810	0	2772	40	1

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 7.

The worst 5 of 40 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:54:ILE:HD11	1:A:303:PRO:HB2	1.64	0.78
1:A:304:VAL:HA	1:A:308:MET:HG3	1.74	0.69
1:A:176:SER:O	8:A:2013:HOH:O	2.11	0.68
1:A:93:THR:HG21	1:A:113:GLU:HG3	1.75	0.68
1:A:300:VAL:O	1:A:303:PRO:HD2	1.96	0.66

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	Distance(Å)	Clash(Å)
6:A:1330:NAG:O3	6:A:1330:NAG:O3[16_545]	2.11	0.09

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	
1	A	326/349 (93%)	313 (96%)	13 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	9/11 (82%)	9 (100%)	0	0	100	100
All	All	335/360 (93%)	322 (96%)	13 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	280/297 (94%)	266 (95%)	14 (5%)	34	73
2	B	10/10 (100%)	9 (90%)	1 (10%)	11	32
All	All	290/307 (94%)	275 (95%)	15 (5%)	32	71

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	229	THR
1	A	232	GLU
1	A	289	THR
1	A	226	LEU
1	A	268	TYR

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

Mol	Chain	Res	Type
1	A	100	HIS
1	A	326	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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)

4 carbohydrates 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	#$\|Z\| > 2$
6	NAG	A	1330	1,6	12,14,15	0.49	0	15,19,21	0.84	0
6	NAG	A	1331	6	12,14,15	0.50	0	15,19,21	0.72	0
6	BMA	A	1332	6	10,11,12	0.60	0	11,15,17	2.29	3 (27%)
6	MAN	A	1333	6	10,11,12	1.00	1 (10%)	11,15,17	1.11	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
6	NAG	A	1330	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	1331	6	-	0/6/23/26	0/1/1/1
6	BMA	A	1332	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1333	6	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1333	MAN	C4-C3	2.47	1.59	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1332	BMA	O5-C5-C4	5.64	117.81	110.65
6	A	1332	BMA	C3-C4-C5	3.58	116.59	110.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1332	BMA	C4-C3-C2	2.95	114.47	110.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry (i)

4 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	#$|Z| > 2$
3	SO4	A	1327	-	4,4,4	2.24	4 (100%)	6,6,6	2.31	2 (33%)
4	ACT	A	1328	-	1,3,3	1.32	0	0,3,3	0.00	-
5	PLM	A	1329	1	16,16,17	5.35	2 (12%)	13,15,17	1.06	1 (7%)
7	BOG	A	1334	-	20,20,20	1.07	2 (10%)	25,25,25	1.57	3 (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
3	SO4	A	1327	-	-	0/0/0/0	0/0/0/0
4	ACT	A	1328	-	-	0/0/0/0	0/0/0/0
5	PLM	A	1329	1	-	0/13/14/15	0/0/0/0
7	BOG	A	1334	-	-	0/11/31/31	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1329	PLM	O2-C1	21.06	1.25	1.11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1334	BOG	O1-C1	3.39	1.46	1.40
5	A	1329	PLM	C2-C1	3.27	1.57	1.47
3	A	1327	SO4	O2-S	2.36	1.54	1.47
3	A	1327	SO4	O4-S	2.26	1.55	1.47

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	1334	BOG	C1'-O1-C1	5.07	123.08	113.96
3	A	1327	SO4	O2-S-O1	3.81	122.14	109.53
3	A	1327	SO4	O4-S-O3	3.79	125.09	109.08
7	A	1334	BOG	O5-C5-C6	3.13	114.04	106.34
5	A	1329	PLM	C4-C3-C2	2.41	123.06	113.36

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(Å^2)	Q<0.9
1	A	327/349 (93%)	-0.13	11 (3%)	43	51	34, 52, 85, 126	0
2	B	11/11 (100%)	0.07	1 (9%)	9	12	51, 59, 67, 73	0
All	All	338/360 (93%)	-0.12	12 (3%)	42	48	34, 52, 85, 126	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	145	ASN	3.7
1	A	326	ASN	3.6
1	A	283	PHE	3.2
1	A	147	ARG	2.7
1	A	173	VAL	2.6

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 ⓘ

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	MAN	A	1333	11/12	0.26	7.68	56,61,67,70	0
6	NAG	A	1330	14/15	0.12	-0.94	35,39,41,51	0
6	NAG	A	1331	14/15	0.12	-5.80	25,32,39,51	0

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å^2)	Q<0.9
6	BMA	A	1332	11/12	0.20	-	49,57,60,60	0

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
7	BOG	A	1334	20/20	0.30	7.25	61,63,78,92	0
5	PLM	A	1329	17/18	0.30	4.02	42,50,76,81	0
3	SO4	A	1327	5/5	0.38	2.49	124,125,135,137	0
4	ACT	A	1328	4/4	0.23	0.42	61,67,70,70	0

6.5 Other polymers ⓘ

There are no such residues in this entry.