Full wwPDB X-ray Structure Validation Report ⓘ

Nov 17, 2024 – 05:37 PM EST

PDB ID	:	4M7W
Title	:	Crystal structure of purine nucleoside phosphorylase from Leptotrichia buccalis C-1013-b, NYSGRC Target 029767.
Authors	:	Malashkevich, V.N.; Bhosle, R.; Toro, R.; Hillerich, B.; Gizzi, A.; Garforth, S.; Kar, A.; Chan, M.K.; Lafluer, J.; Patel, H.; Matikainen, B.; Chamala, S.; Lim, S.; Celikgil, A.; Villegas, G.; Evans, B.; Love, J.; Fiser, A.; Khafizov, K.; Seidel, R.; Bonanno, J.B.; Almo, S.C.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on	:	2013-08-12
Resolution	:	1.90 Å(reported)

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

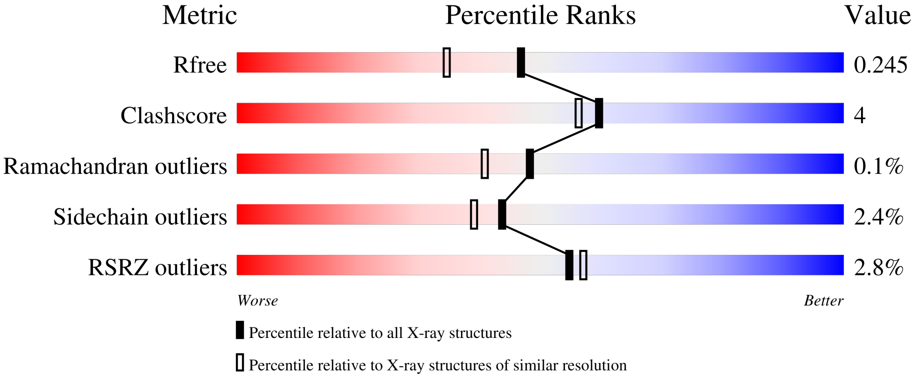
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)

1 Overall quality at a glance

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

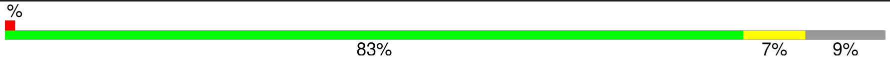
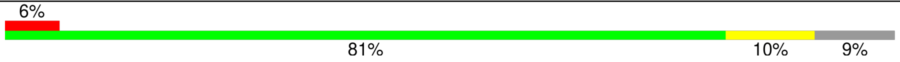
The reported resolution of this entry is 1.90 Å.

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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 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	258	
1	B	258	
1	C	258	

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

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	PO4	C	400	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5758 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 Purine nucleoside phosphorylase DeoD-type.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	236	Total 1802	C 1143	N 297	O 355	S 1	Se 6	0	1	0
1	B	234	Total 1790	C 1136	N 295	O 353	S 1	Se 5	0	1	0
1	C	236	Total 1804	C 1146	N 296	O 355	S 1	Se 6	0	2	0

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MSE	-	expression tag	UNP C7N9Q9
A	-20	HIS	-	expression tag	UNP C7N9Q9
A	-19	HIS	-	expression tag	UNP C7N9Q9
A	-18	HIS	-	expression tag	UNP C7N9Q9
A	-17	HIS	-	expression tag	UNP C7N9Q9
A	-16	HIS	-	expression tag	UNP C7N9Q9
A	-15	HIS	-	expression tag	UNP C7N9Q9
A	-14	SER	-	expression tag	UNP C7N9Q9
A	-13	SER	-	expression tag	UNP C7N9Q9
A	-12	GLY	-	expression tag	UNP C7N9Q9
A	-11	VAL	-	expression tag	UNP C7N9Q9
A	-10	ASP	-	expression tag	UNP C7N9Q9
A	-9	LEU	-	expression tag	UNP C7N9Q9
A	-8	GLY	-	expression tag	UNP C7N9Q9
A	-7	THR	-	expression tag	UNP C7N9Q9
A	-6	GLU	-	expression tag	UNP C7N9Q9
A	-5	ASN	-	expression tag	UNP C7N9Q9
A	-4	LEU	-	expression tag	UNP C7N9Q9
A	-3	TYR	-	expression tag	UNP C7N9Q9
A	-2	PHE	-	expression tag	UNP C7N9Q9
A	-1	GLN	-	expression tag	UNP C7N9Q9
A	0	SER	-	expression tag	UNP C7N9Q9
B	-21	MSE	-	expression tag	UNP C7N9Q9

Continued on next page...

Continued from previous page...

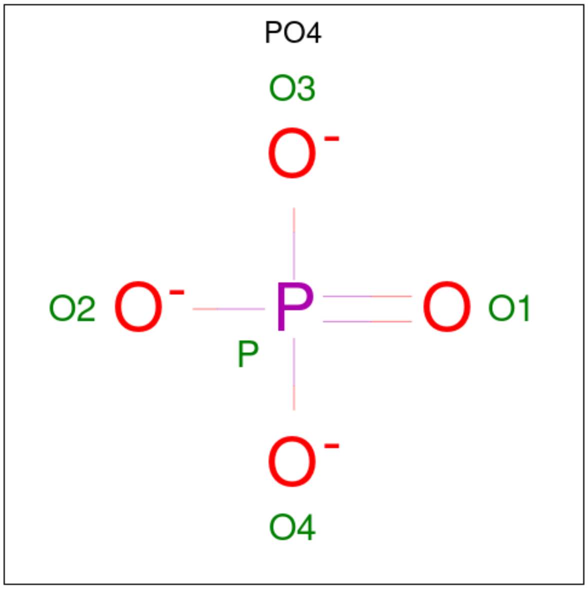
Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	HIS	-	expression tag	UNP C7N9Q9
B	-19	HIS	-	expression tag	UNP C7N9Q9
B	-18	HIS	-	expression tag	UNP C7N9Q9
B	-17	HIS	-	expression tag	UNP C7N9Q9
B	-16	HIS	-	expression tag	UNP C7N9Q9
B	-15	HIS	-	expression tag	UNP C7N9Q9
B	-14	SER	-	expression tag	UNP C7N9Q9
B	-13	SER	-	expression tag	UNP C7N9Q9
B	-12	GLY	-	expression tag	UNP C7N9Q9
B	-11	VAL	-	expression tag	UNP C7N9Q9
B	-10	ASP	-	expression tag	UNP C7N9Q9
B	-9	LEU	-	expression tag	UNP C7N9Q9
B	-8	GLY	-	expression tag	UNP C7N9Q9
B	-7	THR	-	expression tag	UNP C7N9Q9
B	-6	GLU	-	expression tag	UNP C7N9Q9
B	-5	ASN	-	expression tag	UNP C7N9Q9
B	-4	LEU	-	expression tag	UNP C7N9Q9
B	-3	TYR	-	expression tag	UNP C7N9Q9
B	-2	PHE	-	expression tag	UNP C7N9Q9
B	-1	GLN	-	expression tag	UNP C7N9Q9
B	0	SER	-	expression tag	UNP C7N9Q9
C	-21	MSE	-	expression tag	UNP C7N9Q9
C	-20	HIS	-	expression tag	UNP C7N9Q9
C	-19	HIS	-	expression tag	UNP C7N9Q9
C	-18	HIS	-	expression tag	UNP C7N9Q9
C	-17	HIS	-	expression tag	UNP C7N9Q9
C	-16	HIS	-	expression tag	UNP C7N9Q9
C	-15	HIS	-	expression tag	UNP C7N9Q9
C	-14	SER	-	expression tag	UNP C7N9Q9
C	-13	SER	-	expression tag	UNP C7N9Q9
C	-12	GLY	-	expression tag	UNP C7N9Q9
C	-11	VAL	-	expression tag	UNP C7N9Q9
C	-10	ASP	-	expression tag	UNP C7N9Q9
C	-9	LEU	-	expression tag	UNP C7N9Q9
C	-8	GLY	-	expression tag	UNP C7N9Q9
C	-7	THR	-	expression tag	UNP C7N9Q9
C	-6	GLU	-	expression tag	UNP C7N9Q9
C	-5	ASN	-	expression tag	UNP C7N9Q9
C	-4	LEU	-	expression tag	UNP C7N9Q9
C	-3	TYR	-	expression tag	UNP C7N9Q9
C	-2	PHE	-	expression tag	UNP C7N9Q9
C	-1	GLN	-	expression tag	UNP C7N9Q9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	SER	-	expression tag	UNP C7N9Q9

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



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

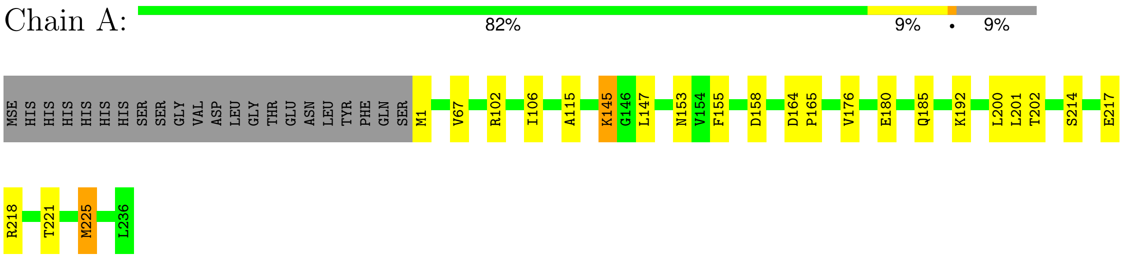
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	127	Total O 127 127	0	0
3	B	123	Total O 123 123	0	0
3	C	97	Total O 97 97	0	0

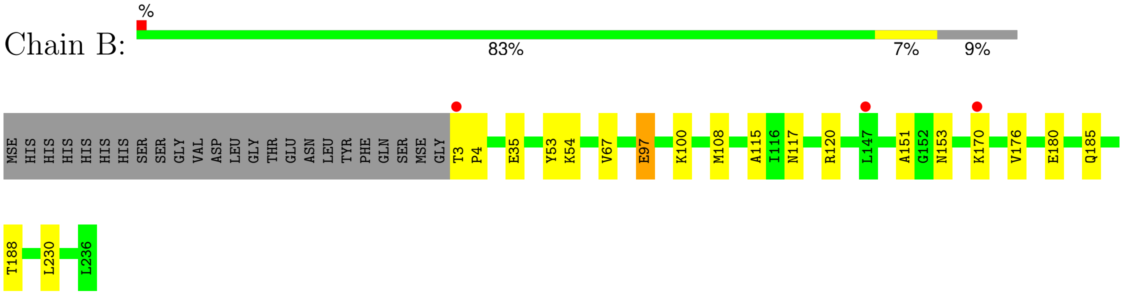
3 Residue-property plots ⓘ

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.

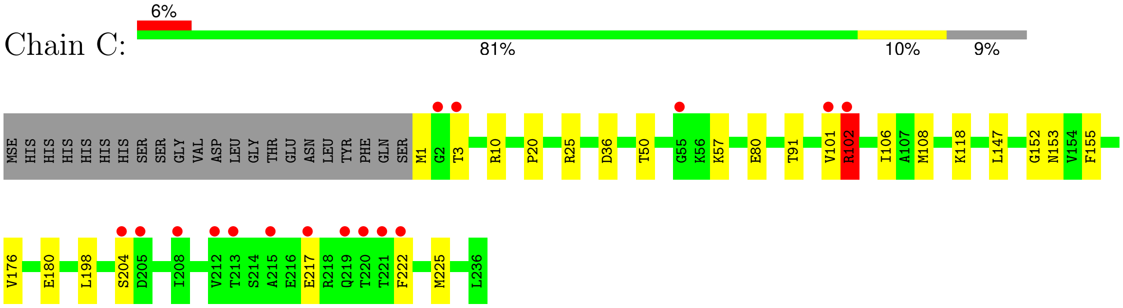
- Molecule 1: Purine nucleoside phosphorylase DeoD-type



- Molecule 1: Purine nucleoside phosphorylase DeoD-type



- Molecule 1: Purine nucleoside phosphorylase DeoD-type



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	164.88Å 164.88Å 45.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.90 – 1.90 32.90 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (32.90-1.90) 99.3 (32.90-1.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 1.89Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.201 , 0.244 0.202 , 0.245	Depositor DCC
R_{free} test set	2831 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	24.7	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 37.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5758	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.48	0/1831	0.62	1/2474 (0.0%)
1	B	0.49	0/1819	0.65	0/2459
1	C	0.46	0/1836	0.62	0/2481
All	All	0.48	0/5486	0.63	1/7414 (0.0%)

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	225	MSE	CG-SE-CE	-5.75	86.24	98.90

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	1802	0	1795	14	0
1	B	1790	0	1780	13	0
1	C	1804	0	1803	19	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	2	0
3	A	127	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	123	0	0	1	0
3	C	97	0	0	1	0
All	All	5758	0	5378	45	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 4.

All (45) 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:C:101:VAL:O	1:C:102:ARG:HB2	1.72	0.89
1:A:115:ALA:HB2	1:B:115:ALA:HB2	1.68	0.75
1:A:201:LEU:C	1:A:225:MSE:HE1	2.13	0.69
1:B:3:THR:HB	1:B:4:PRO:HD3	1.75	0.68
1:A:102:ARG:HD2	1:A:221:THR:OG1	2.01	0.61
1:A:202:THR:N	1:A:225:MSE:HE1	2.19	0.57
1:C:25:ARG:HD3	1:C:222:PHE:CE1	2.40	0.56
1:B:117:ASN:ND2	1:B:120:ARG:HH11	2.04	0.56
1:B:153:ASN:HD22	1:B:176:VAL:HG22	1.72	0.55
1:B:170:LYS:HG3	3:B:579:HOH:O	2.05	0.55
1:C:153:ASN:HD22	1:C:176:VAL:HG22	1.75	0.51
1:C:36:ASP:HB3	3:C:542:HOH:O	2.09	0.51
1:C:25:ARG:NH2	2:C:400:PO4:O4	2.43	0.51
1:B:117:ASN:HD21	1:B:188:THR:HG21	1.76	0.50
1:C:108:MSE:CE	1:C:152:GLY:HA2	2.42	0.50
1:A:102:ARG:CZ	1:A:217:GLU:HG2	2.41	0.50
1:A:145:LYS:HB3	1:A:147:LEU:HG	1.94	0.49
1:C:101:VAL:O	1:C:102:ARG:CB	2.49	0.49
1:C:108:MSE:HE2	1:C:152:GLY:HA2	1.94	0.48
1:C:20:PRO:HB3	1:C:25:ARG:HD2	1.94	0.48
1:A:192:LYS:NZ	3:A:598:HOH:O	2.42	0.48
1:B:108:MSE:HE2	1:B:151:ALA:O	2.14	0.47
1:A:225:MSE:HE3	1:A:225:MSE:HB2	1.49	0.47
1:A:164:ASP:HA	1:A:165:PRO:HD2	1.78	0.46
1:B:117:ASN:HD22	1:B:120:ARG:HH11	1.63	0.46
1:C:91:THR:HB	1:C:204[A]:SER:OG	2.15	0.46
1:C:25:ARG:NH1	2:C:400:PO4:O3	2.46	0.45
1:B:153:ASN:ND2	1:B:176:VAL:HG22	2.32	0.45
1:A:214:SER:OG	1:A:217:GLU:HB2	2.17	0.44
1:C:50:THR:CG2	1:C:57:LYS:HE3	2.47	0.44
1:A:67:VAL:HG13	1:A:185:GLN:HB2	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:25:ARG:HD3	1:C:222:PHE:HE1	1.81	0.44
1:A:153:ASN:HD22	1:A:176:VAL:HG22	1.83	0.43
1:C:50:THR:HG22	1:C:57:LYS:HE3	2.00	0.43
1:B:97:GLU:O	1:B:100:LYS:HE3	2.18	0.42
1:C:3:THR:CG2	1:C:80:GLU:HG3	2.50	0.42
1:A:106:ILE:HG12	1:A:200:LEU:CD2	2.50	0.42
1:B:67:VAL:HG13	1:B:185:GLN:HB2	2.01	0.42
1:C:118:LYS:HA	1:C:118:LYS:HE2	2.02	0.41
1:A:155:PHE:HB2	1:A:176:VAL:HG11	2.01	0.41
1:B:53:TYR:CZ	1:B:54:LYS:HE2	2.55	0.41
1:B:230:LEU:HD23	1:B:230:LEU:HA	1.85	0.41
1:C:155:PHE:HB2	1:C:176:VAL:HG11	2.03	0.41
1:C:102:ARG:HG2	1:C:217:GLU:OE2	2.20	0.40
1:C:106:ILE:HG23	1:C:198:LEU:HD11	2.02	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	
1	A	235/258 (91%)	231 (98%)	4 (2%)	0	100	100
1	B	233/258 (90%)	229 (98%)	4 (2%)	0	100	100
1	C	236/258 (92%)	232 (98%)	3 (1%)	1 (0%)	30	22
All	All	704/774 (91%)	692 (98%)	11 (2%)	1 (0%)	48	41

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	102	ARG

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	192/204 (94%)	187 (97%)	5 (3%)	41	36
1	B	191/204 (94%)	188 (98%)	3 (2%)	58	56
1	C	193/204 (95%)	187 (97%)	6 (3%)	35	29
All	All	576/612 (94%)	562 (98%)	14 (2%)	44	39

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MSE
1	A	145	LYS
1	A	158	ASP
1	A	180	GLU
1	A	218	ARG
1	B	35	GLU
1	B	97	GLU
1	B	180	GLU
1	C	1	MSE
1	C	10	ARG
1	C	102	ARG
1	C	147	LEU
1	C	180	GLU
1	C	225	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	153	ASN
1	B	117	ASN
1	B	153	ASN
1	B	185	GLN
1	C	85	ASN
1	C	153	ASN
1	C	196	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

3 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	$\#\|Z\| > 2$
2	PO4	B	400	-	4,4,4	1.05	0	6,6,6	0.68	0
2	PO4	C	400	-	4,4,4	1.17	0	6,6,6	0.84	0
2	PO4	A	400	-	4,4,4	1.10	0	6,6,6	0.73	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	**Chain**	**Res**	**Type**	**Clashes**	**Symm-Clashes**
2	C	400	PO4	2	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, 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	230/258 (89%)	-0.12	0	100	100	19, 28, 52, 75	1 (0%)
1	B	229/258 (88%)	-0.11	3 (1%)	74	76	19, 29, 49, 79	1 (0%)
1	C	230/258 (89%)	0.22	16 (6%)	24	25	16, 32, 66, 94	2 (0%)
All	All	689/774 (89%)	-0.00	19 (2%)	55	57	16, 30, 55, 94	4 (0%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	212	VAL	4.8
1	C	222	PHE	4.7
1	C	2	GLY	4.0
1	B	3	THR	3.5
1	C	215	ALA	3.4
1	C	213	THR	3.1
1	C	204[A]	SER	2.9
1	C	101	VAL	2.7
1	C	221	THR	2.7
1	C	205	ASP	2.6
1	C	219	GLN	2.5
1	C	220	THR	2.5
1	C	102	ARG	2.5
1	C	55	GLY	2.4
1	C	217	GLU	2.3
1	C	3	THR	2.2
1	C	208	ILE	2.1
1	B	170	LYS	2.1
1	B	147	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 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	PO4	A	400	5/5	0.95	0.09	44,46,47,52	0
2	PO4	C	400	5/5	0.95	0.09	45,45,46,51	0
2	PO4	B	400	5/5	0.98	0.07	30,32,34,34	0

6.5 Other polymers ⓘ

There are no such residues in this entry.