



Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:46 am BST

PDB ID : 6JAF
Title : Crystal structure of Trypanosoma brucei gambiense glycerol kinase complex with PPi (pyrophosphatase reaction)
Authors : Balogun, E.O.; Chishima, T.; Ichinose, M.; Inaoka, D.K.; Kido, Y.; Ibrahim, B.; Bringaud, F.; de Koning, H.; McKerrow, J.H.; Watanabe, Y.; Nozaki, T.; Michels, P.A.M.; Harada, S.; Kita, K.; Shiba, T.
Deposited on : 2019-01-24
Resolution : 2.90 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

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 : 2.11
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) : 2.11

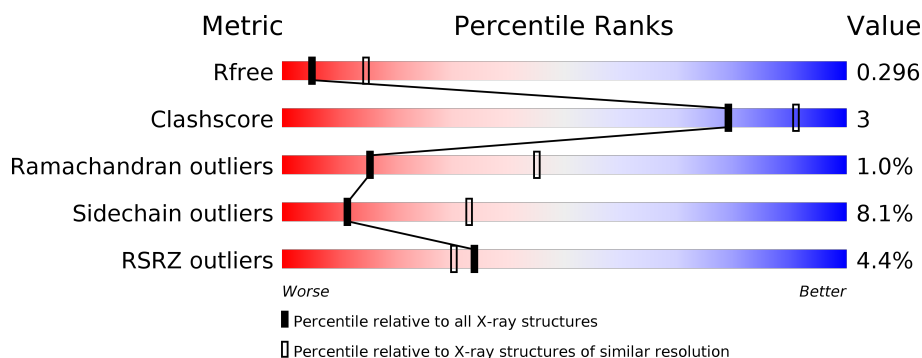
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	518	<div> <div>6%</div> <div>87%</div> <div>11%</div> <div>..</div> </div>
1	B	518	<div> <div>3%</div> <div>86%</div> <div>12%</div> <div>..</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7935 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.

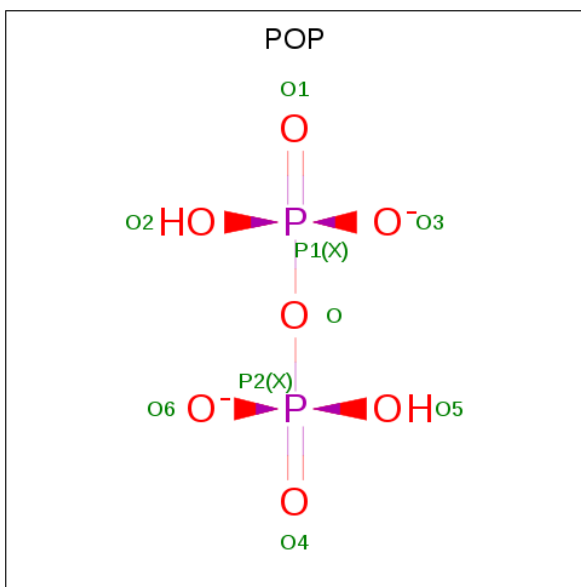
- Molecule 1 is a protein called Glycerol kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	513	Total	C	N	O	S	0	0	0
			3957	2499	694	731	33			
1	B	513	Total	C	N	O	S	0	0	0
			3957	2499	694	731	33			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP D3KVM3
A	-4	ILE	-	expression tag	UNP D3KVM3
A	-3	ASP	-	expression tag	UNP D3KVM3
A	-2	PRO	-	expression tag	UNP D3KVM3
A	-1	PHE	-	expression tag	UNP D3KVM3
A	0	THR	-	expression tag	UNP D3KVM3
B	-5	GLY	-	expression tag	UNP D3KVM3
B	-4	ILE	-	expression tag	UNP D3KVM3
B	-3	ASP	-	expression tag	UNP D3KVM3
B	-2	PRO	-	expression tag	UNP D3KVM3
B	-1	PHE	-	expression tag	UNP D3KVM3
B	0	THR	-	expression tag	UNP D3KVM3

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



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

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

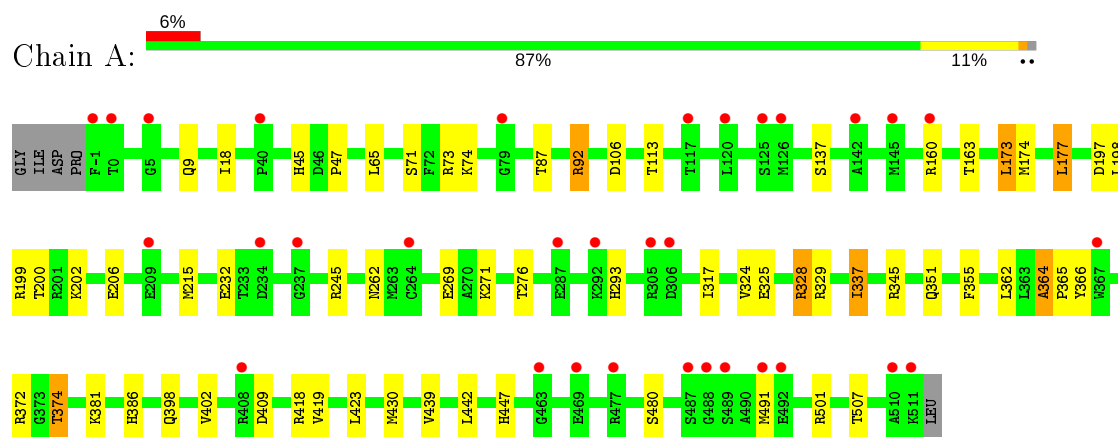


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

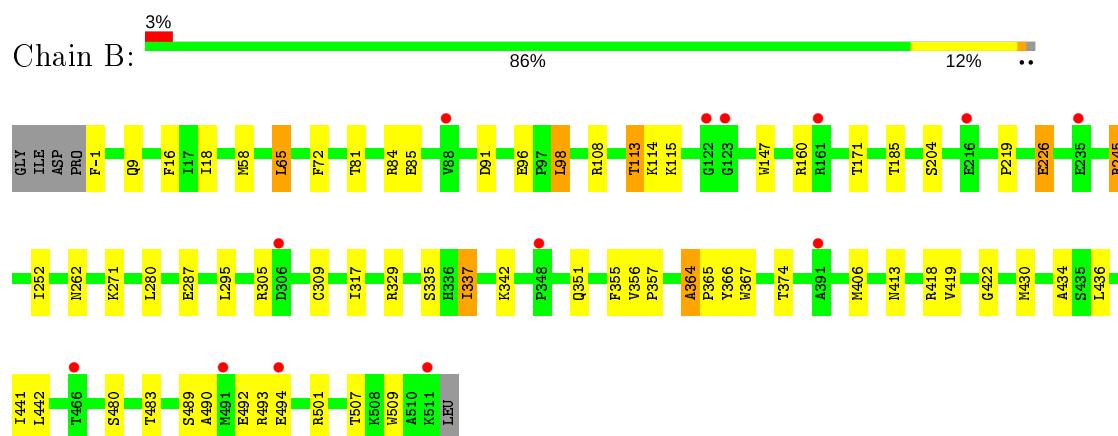
3 Residue-property plots [i](#)

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



• Molecule 1: Glycerol kinase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	63.05Å 120.46Å 154.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 20.00 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.5 (20.00-2.90) 99.9 (20.00-2.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.42 (at 2.88Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.214 , 0.298 0.220 , 0.296	Depositor DCC
R_{free} test set	1347 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	90.4	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 61.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7935	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 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	A	0.42	0/4039	0.68	0/5465
1	B	0.43	0/4039	0.70	1/5465 (0.0%)
All	All	0.42	0/8078	0.69	1/10930 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	245	ARG	NE-CZ-NH1	5.00	122.80	120.30

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	3957	0	3970	23	0
1	B	3957	0	3970	21	0
2	A	9	0	0	0	0
3	A	6	0	8	0	0
3	B	6	0	8	1	0
All	All	7935	0	7956	44	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 (44) 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:374:THR:HG21	1:A:507:THR:HA	1.73	0.69
1:A:364:ALA:HB1	1:A:365:PRO:CD	2.24	0.68
1:B:364:ALA:HB1	1:B:365:PRO:HD2	1.79	0.64
1:B:364:ALA:HB1	1:B:365:PRO:CD	2.30	0.61
1:A:197:ASP:HB3	1:A:200:THR:HG22	1.85	0.59
1:A:351:GLN:HG2	1:A:386:HIS:CD2	2.41	0.56
1:A:419:VAL:HG21	1:A:430:MET:SD	2.47	0.55
1:A:325:GLU:OE1	1:A:328:ARG:NH1	2.42	0.53
1:B:171:THR:HG21	1:B:219:PRO:HG3	1.90	0.52
1:A:269:GLU:OE1	1:A:418:ARG:NH2	2.43	0.52
1:A:374:THR:CG2	1:A:507:THR:HG22	2.40	0.51
1:A:364:ALA:HB1	1:A:365:PRO:HD2	1.93	0.51
1:B:91:ASP:HB2	1:B:98:LEU:HD11	1.93	0.51
1:B:16:PHE:CD2	1:B:58:MET:HA	2.46	0.51
1:B:84:ARG:HE	3:B:601:GOL:H31	1.77	0.50
1:A:324:VAL:HG21	1:A:423:LEU:HD21	1.93	0.50
1:A:45:HIS:O	1:A:47:PRO:HD3	2.11	0.49
1:B:262:ASN:HB3	1:B:418:ARG:HG3	1.95	0.49
1:B:171:THR:HG21	1:B:219:PRO:CG	2.43	0.48
1:B:65:LEU:HD13	1:B:72:PHE:CD2	2.49	0.48
1:A:364:ALA:CB	1:A:365:PRO:CD	2.90	0.48
1:B:364:ALA:O	1:B:367:TRP:NE1	2.48	0.47
1:B:419:VAL:HG21	1:B:430:MET:SD	2.55	0.46
1:B:434:ALA:N	1:B:441:ILE:HD12	2.31	0.45
1:B:81:THR:HA	1:B:252:ILE:O	2.16	0.45
1:A:372:ARG:O	1:A:374:THR:HG22	2.16	0.45
1:A:398:GLN:O	1:A:402:VAL:HG23	2.17	0.44
1:A:262:ASN:HB3	1:A:418:ARG:HG3	1.99	0.44
1:B:113:THR:HG23	1:B:114:LYS:HE2	1.98	0.44
1:A:293:HIS:CD2	1:A:409:ASP:OD2	2.70	0.44
1:A:18:ILE:HG21	1:A:65:LEU:CD1	2.47	0.44
1:B:489:SER:O	1:B:492:GLU:N	2.50	0.44
1:B:436:LEU:O	1:B:493:ARG:NE	2.34	0.43
1:B:18:ILE:HG21	1:B:65:LEU:HD12	2.01	0.42
1:A:374:THR:HG23	1:A:507:THR:HG22	2.02	0.42
1:B:108:ARG:HD3	1:B:147:TRP:CZ2	2.55	0.42
1:A:197:ASP:OD1	1:A:200:THR:HG22	2.20	0.42
1:A:92:ARG:CD	1:A:163:THR:HA	2.49	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:356:VAL:HG11	1:B:507:THR:HG21	2.02	0.42
1:A:173:LEU:O	1:A:177:LEU:HD22	2.20	0.41
1:A:232:GLU:OE1	1:A:245:ARG:NH1	2.53	0.41
1:B:295:LEU:HD13	1:B:406:MET:HG2	2.02	0.41
1:B:364:ALA:CB	1:B:365:PRO:CD	2.99	0.40
1:A:18:ILE:HG21	1:A:65:LEU:HD12	2.03	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	511/518 (99%)	469 (92%)	40 (8%)	2 (0%)	34	66
1	B	511/518 (99%)	478 (94%)	25 (5%)	8 (2%)	9	32
All	All	1022/1036 (99%)	947 (93%)	65 (6%)	10 (1%)	15	45

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	364	ALA
1	B	337	ILE
1	B	364	ALA
1	B	490	ALA
1	A	337	ILE
1	B	85	GLU
1	B	226	GLU
1	B	509	TRP
1	B	422	GLY
1	B	357	PRO

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	421/425 (99%)	385 (91%)	36 (9%)	10	30
1	B	421/425 (99%)	389 (92%)	32 (8%)	13	36
All	All	842/850 (99%)	774 (92%)	68 (8%)	11	33

All (68) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	71	SER
1	A	73	ARG
1	A	74	LYS
1	A	87	THR
1	A	92	ARG
1	A	106	ASP
1	A	113	THR
1	A	137	SER
1	A	160	ARG
1	A	173	LEU
1	A	174	MET
1	A	177	LEU
1	A	198	LEU
1	A	199	ARG
1	A	202	LYS
1	A	206	GLU
1	A	215	MET
1	A	271	LYS
1	A	276	THR
1	A	317	ILE
1	A	328	ARG
1	A	329	ARG
1	A	337	ILE
1	A	345	ARG
1	A	355	PHE
1	A	362	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	366	TYR
1	A	374	THR
1	A	381	LYS
1	A	439	VAL
1	A	442	LEU
1	A	447	HIS
1	A	480	SER
1	A	491	MET
1	A	501	ARG
1	B	-1	PHE
1	B	9	GLN
1	B	65	LEU
1	B	96	GLU
1	B	98	LEU
1	B	113	THR
1	B	115	LYS
1	B	160	ARG
1	B	185	THR
1	B	204	SER
1	B	226	GLU
1	B	245	ARG
1	B	271	LYS
1	B	280	LEU
1	B	287	GLU
1	B	305	ARG
1	B	309	CYS
1	B	317	ILE
1	B	329	ARG
1	B	335	SER
1	B	337	ILE
1	B	342	LYS
1	B	351	GLN
1	B	355	PHE
1	B	366	TYR
1	B	374	THR
1	B	413	ASN
1	B	442	LEU
1	B	480	SER
1	B	483	THR
1	B	494	GLU
1	B	501	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	GLN
1	A	224	ASN
1	A	272	ASN
1	A	293	HIS
1	A	351	GLN
1	A	447	HIS
1	B	23	GLN
1	B	224	ASN
1	B	272	ASN
1	B	293	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	$\# Z > 2$
3	GOL	A	702	-	5,5,5	0.33	0	5,5,5	0.47	0
2	POP	A	701	-	6,8,8	0.58	0	13,13,13	0.81	0
3	GOL	B	601	-	5,5,5	0.30	0	5,5,5	0.47	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
3	GOL	A	702	-	-	0/4/4/4	-
2	POP	A	701	-	-	0/6/6/6	-
3	GOL	B	601	-	-	0/4/4/4	-

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 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	601	GOL	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	513/518 (99%)	0.29	32 (6%)	20 16	63, 90, 119, 142	0
1	B	513/518 (99%)	0.14	13 (2%)	57 55	56, 80, 104, 117	0
All	All	1026/1036 (99%)	0.22	45 (4%)	34 30	56, 84, 115, 142	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	491	MET	4.6
1	A	463	GLY	4.4
1	A	489	SER	3.7
1	A	287	GLU	3.5
1	A	234	ASP	3.5
1	A	292	LYS	3.4
1	A	5	GLY	3.4
1	B	123	GLY	3.3
1	A	306	ASP	3.3
1	A	120	LEU	3.1
1	B	348	PRO	3.0
1	B	306	ASP	3.0
1	A	209	GLU	3.0
1	A	408	ARG	3.0
1	A	510	ALA	2.9
1	A	117	THR	2.8
1	A	367	TRP	2.8
1	B	216	GLU	2.8
1	A	40	PRO	2.8
1	A	125	SER	2.7
1	A	488	GLY	2.7
1	A	305	ARG	2.6
1	A	145	MET	2.6
1	A	0	THR	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	-1	PHE	2.5
1	B	466	THR	2.5
1	A	487	SER	2.4
1	A	511	LYS	2.4
1	A	126	MET	2.4
1	A	142	ALA	2.4
1	B	88	VAL	2.3
1	A	492	GLU	2.3
1	B	491	MET	2.3
1	B	511	LYS	2.2
1	A	469	GLU	2.2
1	A	477	ARG	2.2
1	B	161	ARG	2.2
1	A	160	ARG	2.1
1	B	494	GLU	2.1
1	A	237	GLY	2.1
1	B	391	ALA	2.1
1	B	122	GLY	2.1
1	A	79	GLY	2.0
1	A	264	CYS	2.0
1	B	235	GLU	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. 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	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	GOL	A	702	6/6	0.83	0.38	65,72,78,84	0
2	POP	A	701	9/9	0.83	0.20	85,96,137,138	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	GOL	B	601	6/6	0.84	0.36	61,67,71,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.