Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 12:06 pm BST

PDB ID	:	1OKT
Title	:	X-ray Structure of Glutathione S-Transferase from the Malarial Parasite Plasmodium falciparum
Authors	:	Fritz-Wolf, K.; Becker, A.; Rahlfs, s.; Harwaldt, P.; Schirmer, R.H.; Kabsch, W.; Becker, K.
Deposited on	:	2003-07-29
Resolution	:	1.90 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:

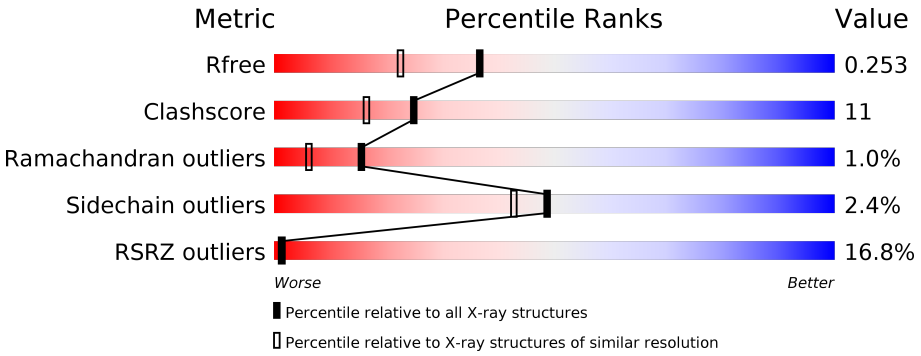
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (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 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 <=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	211	11% 81% 16%
1	B	211	22% 79% 19%

2 Entry composition (i)

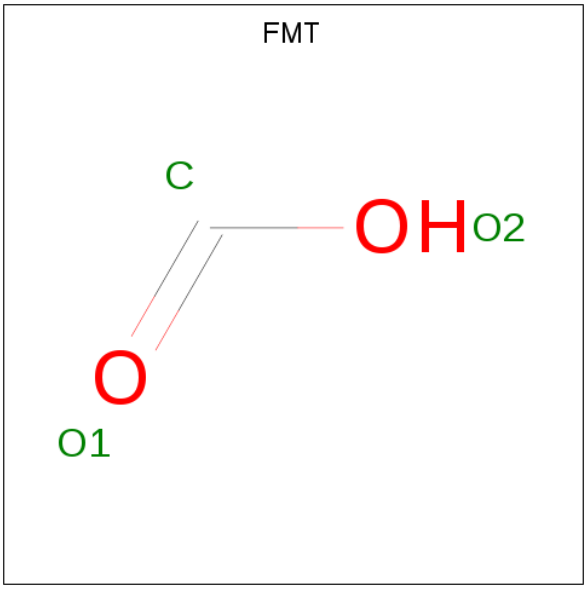
There are 3 unique types of molecules in this entry. The entry contains 3840 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 GLUTATHIONE S-TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	211	Total	C	N	O	S	0	0	0
			1757	1141	283	329	4			
1	B	211	Total	C	N	O	S	0	0	0
			1757	1141	283	329	4			

- Molecule 2 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			3	1	2		
2	A	1	Total	C	O	0	0
			3	1	2		
2	B	1	Total	C	O	0	0
			3	1	2		
2	B	1	Total	C	O	0	0
			3	1	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 3 1 2	0	0

- Molecule 3 is water.

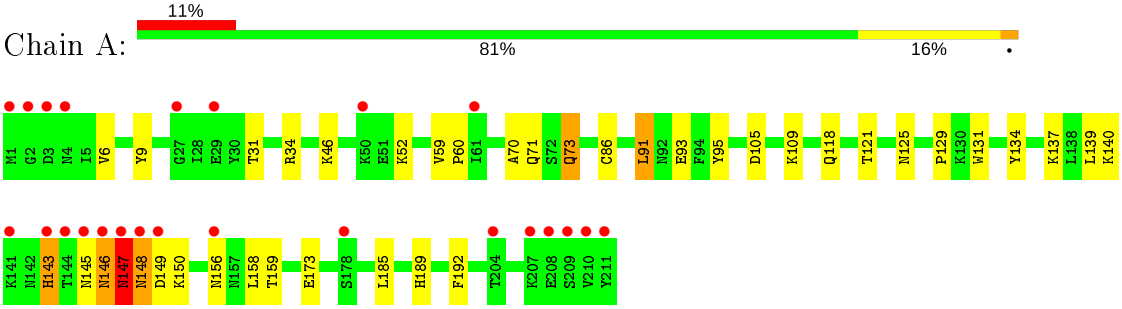
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	183	Total O 183 183	0	0
3	B	128	Total O 128 128	0	0

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 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.

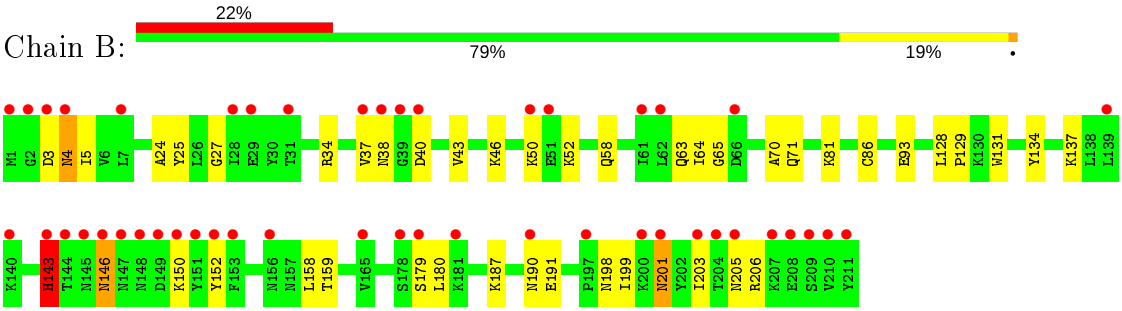
- Molecule 1: GLUTATHIONE S-TRANSFERASE

Chain A:



- Molecule 1: GLUTATHIONE S-TRANSFERASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	89.36Å 63.00Å 75.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.22 – 1.90 19.21 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.9 (19.22-1.90) 99.1 (19.21-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.23 (at 1.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.222 , 0.259 0.216 , 0.253	Depositor DCC
R_{free} test set	1017 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å^2)	27.9	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 52.0	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	3840	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.59% 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: FMT

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	0.35	0/1800	0.56	0/2430
1	B	0.32	0/1800	0.51	0/2430
All	All	0.34	0/3600	0.54	0/4860

There are no bond length outliers.

There are no bond angle outliers.

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	1757	0	1713	38	0
1	B	1757	0	1713	38	1
2	A	6	0	2	0	0
2	B	9	0	3	0	0
3	A	183	0	0	4	2
3	B	128	0	0	4	1
All	All	3840	0	3431	75	4

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 11.

All (75) 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:147:ASN:HB3	1:A:150:LYS:HG2	1.35	1.05
1:A:73:GLN:H	1:A:73:GLN:HE21	1.19	0.90
1:B:37:VAL:HG23	1:B:38:ASN:ND2	2.02	0.75
1:A:147:ASN:HB3	1:A:150:LYS:CG	2.18	0.71
1:B:46:LYS:HG2	1:B:50:LYS:NZ	2.08	0.68
1:A:145:ASN:C	1:A:147:ASN:H	1.99	0.66
1:B:143:HIS:CD2	1:B:146:ASN:HD21	2.14	0.65
1:A:147:ASN:HD22	1:A:147:ASN:C	2.01	0.63
1:B:25:TYR:HB2	1:B:199:ILE:HD11	1.80	0.62
1:B:52:LYS:NZ	1:B:63:GLN:HE22	1.98	0.62
1:A:146:ASN:N	1:A:146:ASN:HD22	2.00	0.60
1:B:34:ARG:O	1:B:37:VAL:HG22	2.02	0.59
1:B:25:TYR:CE2	1:B:158:LEU:HD13	2.38	0.57
1:A:145:ASN:ND2	1:A:156:ASN:HB3	2.19	0.57
1:A:105:ASP:OD1	1:A:109:LYS:HE2	2.05	0.56
1:A:147:ASN:HD22	1:A:148:ASN:N	2.04	0.56
1:B:158:LEU:C	1:B:158:LEU:HD23	2.27	0.56
1:B:46:LYS:HG2	1:B:50:LYS:HZ3	1.69	0.56
1:B:40:ASP:OD1	1:B:43:VAL:HG23	2.08	0.54
1:B:143:HIS:HD2	1:B:146:ASN:HD21	1.53	0.54
1:B:179:SER:HB3	3:B:2118:HOH:O	2.07	0.53
1:A:91:LEU:HD22	1:A:95:TYR:CE1	2.45	0.52
1:A:192:PHE:HB2	3:A:2143:HOH:O	2.09	0.51
1:B:150:LYS:HE2	1:B:152:TYR:HE1	1.75	0.51
1:B:158:LEU:HD23	1:B:159:THR:N	2.26	0.51
1:A:158:LEU:HD23	1:A:158:LEU:C	2.31	0.50
1:A:158:LEU:HD23	1:A:159:THR:N	2.27	0.49
1:A:46:LYS:HB2	1:A:46:LYS:NZ	2.27	0.49
1:B:38:ASN:N	1:B:38:ASN:HD22	2.10	0.49
1:B:81:LYS:HA	1:B:86:CYS:SG	2.53	0.49
1:B:143:HIS:HB2	1:B:146:ASN:HD21	1.78	0.49
1:A:145:ASN:C	1:A:147:ASN:N	2.66	0.48
1:B:86:CYS:O	1:B:93:GLU:HA	2.12	0.48
1:A:147:ASN:HD22	1:A:149:ASP:H	1.61	0.48
1:B:70:ALA:O	1:B:71:GLN:HB2	2.14	0.47
1:A:140:LYS:O	1:A:143:HIS:HB2	2.15	0.47
1:A:146:ASN:N	1:A:146:ASN:ND2	2.63	0.47
1:B:24:ALA:O	1:B:198:ASN:HB3	2.14	0.47
1:A:147:ASN:C	1:A:147:ASN:ND2	2.67	0.46
1:A:147:ASN:ND2	1:A:148:ASN:N	2.62	0.46
1:B:128:LEU:HB3	1:B:129:PRO:HD3	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:118:GLN:NE2	1:B:58:GLN:HB3	2.31	0.46
1:B:187:LYS:O	1:B:191:GLU:HG3	2.15	0.46
1:B:4:ASN:HD22	1:B:5:ILE:N	2.13	0.46
1:A:147:ASN:O	1:A:149:ASP:N	2.49	0.46
1:A:129:PRO:HG3	3:A:2131:HOH:O	2.15	0.45
1:B:143:HIS:HB2	1:B:146:ASN:ND2	2.30	0.45
1:B:180:LEU:HD13	1:B:190:ASN:HD22	1.82	0.45
1:A:121:THR:HG23	1:A:125:ASN:ND2	2.32	0.45
1:B:131:TRP:HA	1:B:134:TYR:CD2	2.52	0.45
1:A:131:TRP:HA	1:A:134:TYR:CD2	2.51	0.44
1:A:146:ASN:O	1:A:147:ASN:C	2.56	0.44
1:B:40:ASP:HB3	1:B:43:VAL:CG2	2.48	0.44
1:A:59:VAL:HB	1:A:60:PRO:HA	2.00	0.44
1:A:6:VAL:HG22	1:A:31:THR:HB	2.00	0.44
1:A:137:LYS:HD3	3:A:2135:HOH:O	2.17	0.43
1:A:137:LYS:NZ	3:A:2137:HOH:O	2.51	0.43
1:B:3:ASP:HB3	1:B:65:GLY:CA	2.48	0.43
1:A:70:ALA:O	1:A:71:GLN:HB2	2.19	0.43
1:B:5:ILE:HD12	1:B:64:ILE:HG12	2.01	0.43
1:A:143:HIS:ND1	1:A:146:ASN:HB3	2.34	0.43
1:B:143:HIS:HD2	1:B:146:ASN:ND2	2.17	0.42
1:A:146:ASN:O	1:A:147:ASN:O	2.38	0.42
1:A:139:LEU:HD23	1:A:185:LEU:HB3	2.01	0.42
1:B:5:ILE:CD1	1:B:64:ILE:HG23	2.49	0.42
1:B:137:LYS:HG3	3:B:2099:HOH:O	2.21	0.41
1:B:180:LEU:HD11	3:B:2113:HOH:O	2.19	0.41
1:B:201:ASN:ND2	1:B:205:ASN:HD21	2.17	0.41
1:A:9:TYR:O	1:A:34:ARG:HA	2.21	0.41
1:A:139:LEU:HD21	1:A:189:HIS:HB2	2.02	0.41
1:A:86:CYS:O	1:A:93:GLU:HA	2.21	0.41
1:A:52:LYS:HB2	1:A:52:LYS:HE2	1.91	0.40
1:B:27:GLY:HA2	3:B:2012:HOH:O	2.21	0.40
1:B:38:ASN:N	1:B:38:ASN:ND2	2.70	0.40
1:B:203:ILE:O	1:B:206:ARG:HG2	2.22	0.40

All (4) 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)
3:B:2064:HOH:O	3:B:2064:HOH:O[2_665]	1.63	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:37:VAL:O	1:B:37:VAL:O[2_675]	1.71	0.49
3:A:2099:HOH:O	3:A:2099:HOH:O[2_665]	1.81	0.39
3:A:2106:HOH:O	3:A:2106:HOH:O[2_665]	1.86	0.34

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	209/211 (99%)	198 (95%)	9 (4%)	2 (1%)	15 6
1	B	209/211 (99%)	196 (94%)	11 (5%)	2 (1%)	15 6
All	All	418/422 (99%)	394 (94%)	20 (5%)	4 (1%)	15 6

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	147	ASN
1	A	148	ASN
1	B	143	HIS
1	B	146	ASN

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	191/191 (100%)	185 (97%)	6 (3%)	40 32

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	191/191 (100%)	188 (98%)	3 (2%)	62 60
All	All	382/382 (100%)	373 (98%)	9 (2%)	49 43

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	GLN
1	A	91	LEU
1	A	143	HIS
1	A	146	ASN
1	A	147	ASN
1	A	173	GLU
1	B	4	ASN
1	B	143	HIS
1	B	201	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	GLN
1	A	104	GLN
1	A	118	GLN
1	A	125	ASN
1	A	143	HIS
1	A	146	ASN
1	A	147	ASN
1	A	198	ASN
1	A	205	ASN
1	B	4	ASN
1	B	38	ASN
1	B	63	GLN
1	B	118	GLN
1	B	125	ASN
1	B	143	HIS
1	B	146	ASN
1	B	190	ASN
1	B	195	ASN
1	B	198	ASN
1	B	201	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 carbohydrates in this entry.

5.6 Ligand geometry

5 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	FMT	A	1212	-	0,2,2	0.00	-	0,1,1	0.00	-
2	FMT	B	1214	-	0,2,2	0.00	-	0,1,1	0.00	-
2	FMT	B	1212	-	0,2,2	0.00	-	0,1,1	0.00	-
2	FMT	A	1213	-	0,2,2	0.00	-	0,1,1	0.00	-
2	FMT	B	1213	-	0,2,2	0.00	-	0,1,1	0.00	-

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.

No monomer is involved in short contacts.

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	211/211 (100%)	0.82	24 (11%)	5	5	16, 29, 70, 89	0
1	B	211/211 (100%)	1.46	47 (22%)	0	0	21, 39, 78, 95	0
All	All	422/422 (100%)	1.14	71 (16%)	1	1	16, 33, 75, 95	0

All (71) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	145	ASN	14.4
1	B	2	GLY	13.5
1	B	148	ASN	12.4
1	A	146	ASN	11.9
1	A	148	ASN	10.7
1	B	146	ASN	10.0
1	B	1	MET	9.9
1	B	144	THR	9.8
1	A	1	MET	8.9
1	B	147	ASN	8.1
1	B	37	VAL	7.8
1	B	149	ASP	7.8
1	A	145	ASN	7.6
1	A	2	GLY	7.0
1	B	38	ASN	6.9
1	A	143	HIS	6.2
1	A	149	ASP	6.2
1	A	3	ASP	5.5
1	A	211	TYR	5.5
1	A	210	VAL	5.3
1	B	210	VAL	5.2
1	B	39	GLY	5.1
1	A	147	ASN	5.0
1	B	209	SER	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	207	LYS	4.9
1	A	209	SER	4.9
1	B	3	ASP	4.9
1	A	144	THR	4.2
1	B	151	TYR	4.2
1	B	29	GLU	4.0
1	B	208	GLU	4.0
1	B	178	SER	3.7
1	B	143	HIS	3.7
1	B	66	ASP	3.6
1	B	181	LYS	3.5
1	A	208	GLU	3.4
1	B	204	THR	3.4
1	A	204	THR	3.3
1	B	179	SER	3.2
1	B	165	VAL	3.2
1	B	197	PRO	3.1
1	A	178	SER	3.1
1	B	211	TYR	3.1
1	B	150	LYS	3.0
1	B	156	ASN	3.0
1	B	200	LYS	3.0
1	B	205	ASN	2.9
1	B	51	GLU	2.9
1	A	4	ASN	2.8
1	B	153	PHE	2.7
1	B	203	ILE	2.7
1	B	201	ASN	2.7
1	B	190	ASN	2.6
1	A	27	GLY	2.5
1	A	141	LYS	2.5
1	B	140	LYS	2.5
1	B	61	ILE	2.5
1	A	29	GLU	2.5
1	B	4	ASN	2.4
1	B	40	ASP	2.3
1	B	7	LEU	2.2
1	A	156	ASN	2.2
1	A	61	ILE	2.2
1	B	62	LEU	2.2
1	B	31	THR	2.1
1	B	139	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	152	TYR	2.1
1	A	50	LYS	2.1
1	B	50	LYS	2.1
1	B	28	ILE	2.0
1	A	207	LYS	2.0

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

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	FMT	A	1213	3/3	0.95	0.11	36,36,40,43	0
2	FMT	B	1213	3/3	0.95	0.17	44,44,45,45	0
2	FMT	B	1214	3/3	0.97	0.10	29,29,32,35	0
2	FMT	B	1212	3/3	0.97	0.08	28,28,30,31	0
2	FMT	A	1212	3/3	0.99	0.06	23,23,23,26	0

6.5 Other polymers (i)

There are no such residues in this entry.