Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 02:17 AM GMT

PDB ID	:	2GGO
Title	:	Crystal Structure of glucose-1-phosphate thymidylyltransferase from Sulfolobus tokodaii
Authors	:	Rajakannan, V.; Mizushima, T.; Suzuki, A.; Masui, R.; Kuramitsu, S.; Yamane, T.
Deposited on	:	2006-03-24
Resolution	:	1.80 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
http://wwpdb.org/validation/2016/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:

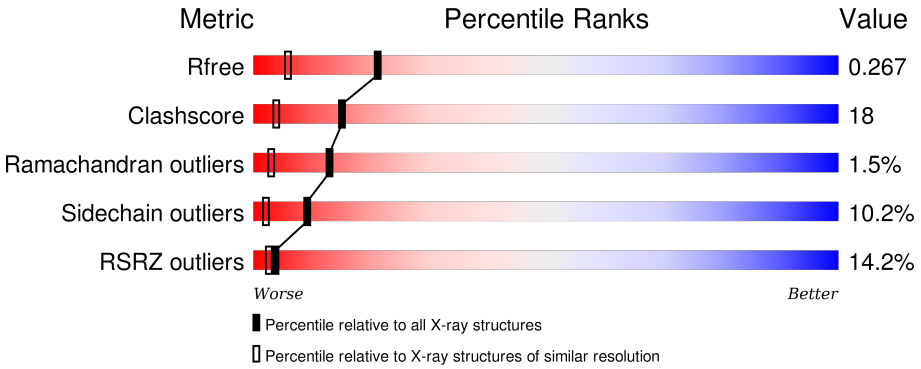
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.80 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}	91344	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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. 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	401	14%, 63%, 30%, 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3332 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 401aa long hypothetical glucose-1-phosphate thymidylyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	401	3141	2021	524	590	6	0	0	0

- Molecule 2 is water.

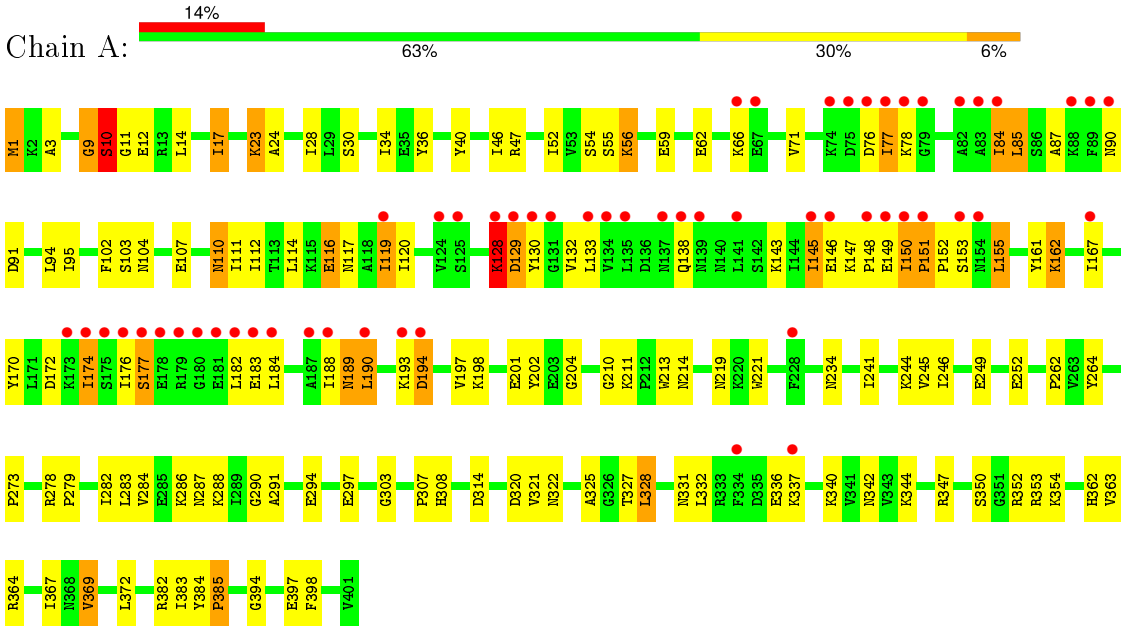
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	191	191	191	0	0

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: 401aa long hypothetical glucose-1-phosphate thymidylyltransferase

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	123.00Å 123.00Å 94.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.00 – 1.80 18.75 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.6 (19.00-1.80) 96.6 (18.75-1.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.222 , 0.267 0.221 , 0.267	Depositor DCC
R_{free} test set	2438 reflections (5.36%)	DCC
Wilson B-factor (Å^2)	19.4	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 76.2	EDS
Estimated twinning fraction	0.031 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	0 of 47908 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3332	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.11% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.96	4/3195 (0.1%)	0.96	3/4311 (0.1%)

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	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	369	VAL	CB-CG2	10.25	1.74	1.52
1	A	321	VAL	CB-CG1	6.79	1.67	1.52
1	A	363	VAL	CB-CG2	-5.45	1.41	1.52
1	A	294	GLU	CB-CG	5.04	1.61	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	328	LEU	CB-CG-CD2	-10.56	93.04	111.00
1	A	308	HIS	CB-CA-C	-5.80	98.80	110.40
1	A	369	VAL	CG1-CB-CG2	-5.23	102.53	110.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	128	LYS	Peptide
1	A	9	GLY	Peptide

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	3141	0	3244	112	0
2	A	191	0	0	15	0
All	All	3332	0	3244	112	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 18.

All (112) 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:320:ASP:HB3	2:A:559:HOH:O	1.49	1.11
1:A:84:ILE:HG22	1:A:87:ALA:HB2	1.38	1.06
1:A:176:ILE:HG12	1:A:182:LEU:HD13	1.47	0.93
1:A:128:LYS:N	1:A:129:ASP:HB2	1.96	0.81
1:A:107:GLU:HG3	1:A:202:TYR:CD1	2.16	0.80
1:A:1:MET:HE1	1:A:162:LYS:HD3	1.67	0.77
1:A:84:ILE:HD11	1:A:184:LEU:HD22	1.65	0.76
1:A:129:ASP:OD1	1:A:129:ASP:O	2.03	0.75
1:A:107:GLU:HG3	1:A:202:TYR:HD1	1.51	0.74
1:A:394:GLY:N	1:A:397:GLU:OE1	2.20	0.73
1:A:40:TYR:CD2	1:A:102:PHE:CE1	2.76	0.73
1:A:148:PRO:HG2	1:A:150:ILE:O	1.93	0.69
1:A:147:LYS:HB3	2:A:590:HOH:O	1.93	0.68
1:A:128:LYS:O	1:A:148:PRO:HD2	1.93	0.67
1:A:40:TYR:HD2	1:A:102:PHE:CE1	2.14	0.66
1:A:327:THR:HA	1:A:369:VAL:O	1.96	0.65
1:A:353:ARG:HD3	2:A:499:HOH:O	1.97	0.64
1:A:286:LYS:O	1:A:303:GLY:HA2	1.99	0.63
1:A:119:ILE:HD11	1:A:188:ILE:HD13	1.80	0.62
1:A:1:MET:CE	1:A:162:LYS:HD3	2.31	0.59
1:A:342:ASN:OD1	1:A:347:ARG:HD2	2.02	0.59
1:A:288:LYS:HE2	2:A:439:HOH:O	2.02	0.59
1:A:249:GLU:HG2	2:A:527:HOH:O	2.02	0.58
1:A:54:SER:OG	1:A:56:LYS:HG3	2.03	0.58
1:A:40:TYR:CE2	1:A:102:PHE:O	2.57	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:143:LYS:HA	1:A:189:ASN:HD21	1.68	0.58
1:A:336:GLU:OE1	1:A:353:ARG:NH2	2.38	0.57
1:A:150:ILE:O	1:A:152:PRO:HD3	2.05	0.57
1:A:336:GLU:HB3	1:A:353:ARG:HH21	1.70	0.56
1:A:340:LYS:HB2	1:A:347:ARG:NH2	2.20	0.56
1:A:3:ALA:HB2	1:A:46:ILE:HG21	1.87	0.56
1:A:241:ILE:HG23	1:A:245:VAL:HG21	1.88	0.56
1:A:320:ASP:CB	2:A:559:HOH:O	2.28	0.56
1:A:145:ILE:CD1	1:A:152:PRO:HG2	2.35	0.56
1:A:194:ASP:OD1	1:A:194:ASP:N	2.39	0.55
1:A:119:ILE:HG12	1:A:161:TYR:CD2	2.41	0.55
1:A:40:TYR:HE2	1:A:102:PHE:O	1.89	0.55
1:A:244:LYS:O	1:A:262:PRO:HA	2.07	0.54
1:A:352:ARG:HD2	2:A:563:HOH:O	2.08	0.54
1:A:290:GLY:HA3	1:A:307:PRO:O	2.07	0.54
1:A:76:ASP:HA	2:A:461:HOH:O	2.08	0.53
1:A:325:ALA:O	1:A:367:ILE:HA	2.08	0.53
1:A:47:ARG:HD3	2:A:405:HOH:O	2.09	0.53
1:A:143:LYS:HA	1:A:189:ASN:ND2	2.25	0.52
1:A:77:ILE:CD1	1:A:85:LEU:HD12	2.40	0.52
1:A:128:LYS:CA	1:A:129:ASP:HB2	2.40	0.52
1:A:129:ASP:OD1	1:A:129:ASP:C	2.48	0.52
1:A:133:LEU:HD12	1:A:155:LEU:HD13	1.92	0.51
1:A:219:ASN:HB3	1:A:282:ILE:HD13	1.92	0.51
1:A:145:ILE:HD13	1:A:152:PRO:HG2	1.92	0.50
1:A:210:GLY:O	1:A:211:LYS:HG3	2.12	0.50
1:A:114:LEU:HD13	1:A:198:LYS:HD3	1.94	0.49
1:A:119:ILE:HG12	1:A:161:TYR:HD2	1.77	0.49
1:A:84:ILE:HD11	1:A:184:LEU:CD2	2.39	0.49
1:A:30:SER:HB3	1:A:264:TYR:OH	2.13	0.48
1:A:384:TYR:CD1	1:A:385:PRO:HD2	2.48	0.48
1:A:150:ILE:HD13	1:A:151:PRO:HD2	1.95	0.48
1:A:155:LEU:HD21	1:A:201:GLU:HG3	1.95	0.48
1:A:297:GLU:O	1:A:314:ASP:HA	2.13	0.48
1:A:198:LYS:NZ	2:A:522:HOH:O	2.45	0.48
1:A:246:ILE:HD12	1:A:246:ILE:N	2.28	0.48
1:A:84:ILE:CG2	1:A:87:ALA:HB2	2.27	0.48
1:A:322:ASN:ND2	2:A:410:HOH:O	2.46	0.47
1:A:52:ILE:HA	1:A:71:VAL:O	2.13	0.47
1:A:23:LYS:HG3	2:A:404:HOH:O	2.14	0.47
1:A:110:ASN:HD22	1:A:110:ASN:C	2.18	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:241:ILE:CG2	1:A:245:VAL:HG21	2.45	0.46
1:A:364:ARG:HD3	2:A:477:HOH:O	2.15	0.46
1:A:174:ILE:HG12	1:A:190:LEU:HD11	1.98	0.46
1:A:234:ASN:HD22	1:A:252:GLU:HG3	1.79	0.45
1:A:128:LYS:N	1:A:129:ASP:CB	2.75	0.45
1:A:24:ALA:HB1	1:A:34:ILE:HB	1.98	0.45
1:A:283:LEU:HB3	1:A:287:ASN:ND2	2.31	0.45
1:A:183:GLU:HB3	2:A:418:HOH:O	2.16	0.45
1:A:23:LYS:HG3	1:A:23:LYS:H	1.57	0.45
1:A:211:LYS:H	1:A:214:ASN:HD22	1.65	0.44
1:A:278:ARG:HB3	1:A:279:PRO:HD2	1.99	0.44
1:A:1:MET:HG3	1:A:112:ILE:HB	2.00	0.44
1:A:174:ILE:HG23	1:A:176:ILE:HG23	2.00	0.44
1:A:146:GLU:OE2	1:A:147:LYS:HG3	2.17	0.44
1:A:372:LEU:N	1:A:372:LEU:HD12	2.32	0.44
1:A:369:VAL:HG21	1:A:383:ILE:HG22	1.99	0.44
1:A:320:ASP:O	1:A:362:HIS:HA	2.17	0.43
1:A:10:SER:HA	1:A:11:GLY:HA3	1.66	0.43
1:A:103:SER:HB3	1:A:221:TRP:CH2	2.53	0.43
1:A:77:ILE:HD12	1:A:85:LEU:HD12	2.00	0.43
1:A:170:TYR:HB3	1:A:190:LEU:HD12	2.01	0.43
1:A:331:ASN:O	1:A:354:LYS:HA	2.18	0.43
1:A:17:ILE:N	1:A:17:ILE:CD1	2.81	0.43
1:A:129:ASP:HB3	1:A:130:TYR:HB2	1.99	0.43
1:A:193:LYS:HA	1:A:193:LYS:HD3	1.81	0.42
1:A:111:ILE:HG12	1:A:162:LYS:HB2	2.01	0.42
1:A:161:TYR:HE1	1:A:184:LEU:HD23	1.85	0.42
1:A:116:GLU:OE2	1:A:198:LYS:HG3	2.19	0.42
1:A:332:LEU:HD13	1:A:354:LYS:HE3	2.00	0.42
1:A:3:ALA:HA	1:A:94:LEU:O	2.20	0.42
1:A:120:ILE:CG2	1:A:202:TYR:HB2	2.50	0.42
1:A:167:ILE:HA	1:A:167:ILE:HD12	1.93	0.42
1:A:176:ILE:HG22	1:A:177:SER:OG	2.19	0.42
1:A:328:LEU:HA	1:A:328:LEU:HD23	1.80	0.42
1:A:14:LEU:HA	1:A:14:LEU:HD12	1.93	0.42
1:A:336:GLU:HB3	1:A:353:ARG:NH2	2.33	0.41
1:A:213:TRP:CZ2	1:A:350:SER:HA	2.55	0.41
1:A:273:PRO:O	1:A:291:ALA:HA	2.21	0.41
1:A:128:LYS:HB3	2:A:570:HOH:O	2.21	0.41
1:A:394:GLY:CA	1:A:397:GLU:OE1	2.68	0.41
1:A:12:GLU:HB3	1:A:344:LYS:HG2	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:132:VAL:HG12	1:A:153:SER:HB3	2.02	0.41
1:A:382:ARG:O	1:A:398:PHE:HA	2.20	0.41
1:A:117:ASN:O	1:A:197:VAL:HA	2.21	0.41
1:A:176:ILE:HA	1:A:177:SER:HA	1.93	0.40
1:A:28:ILE:HB	1:A:36:TYR:CE1	2.56	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	399/401 (100%)	364 (91%)	29 (7%)	6 (2%)	13 3

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	149	GLU
1	A	9	GLY
1	A	10	SER
1	A	129	ASP
1	A	204	GLY
1	A	151	PRO

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	344/344 (100%)	309 (90%)	35 (10%)	9 2

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	10	SER
1	A	17	ILE
1	A	23	LYS
1	A	55	SER
1	A	56	LYS
1	A	59	GLU
1	A	62	GLU
1	A	66	LYS
1	A	77	ILE
1	A	78	LYS
1	A	84	ILE
1	A	85	LEU
1	A	90	ASN
1	A	91	ASP
1	A	95	ILE
1	A	104	ASN
1	A	110	ASN
1	A	116	GLU
1	A	119	ILE
1	A	128	LYS
1	A	138	GLN
1	A	145	ILE
1	A	150	ILE
1	A	155	LEU
1	A	162	LYS
1	A	172	ASP
1	A	174	ILE
1	A	177	SER
1	A	189	ASN
1	A	190	LEU
1	A	194	ASP
1	A	284	VAL
1	A	337	LYS
1	A	385	PRO

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

Mol	Chain	Res	Type
1	A	104	ASN
1	A	110	ASN
1	A	189	ASN
1	A	214	ASN
1	A	234	ASN
1	A	287	ASN
1	A	308	HIS

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

There are no ligands in this entry.

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	401/401 (100%)	0.43	57 (14%) 4 2	5, 37, 97, 122	0

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	77	ILE	11.8
1	A	76	ASP	7.0
1	A	177	SER	6.7
1	A	175	SER	6.0
1	A	179	ARG	5.7
1	A	174	ILE	5.7
1	A	150	ILE	5.6
1	A	176	ILE	5.5
1	A	180	GLY	5.4
1	A	173	LYS	5.2
1	A	137	ASN	5.2
1	A	178	GLU	5.1
1	A	182	LEU	4.8
1	A	74	LYS	4.6
1	A	149	GLU	4.6
1	A	129	ASP	4.1
1	A	145	ILE	4.0
1	A	151	PRO	4.0
1	A	148	PRO	4.0
1	A	89	PHE	4.0
1	A	66	LYS	3.9
1	A	78	LYS	3.7
1	A	75	ASP	3.6
1	A	82	ALA	3.5
1	A	90	ASN	3.3
1	A	188	ILE	3.3
1	A	130	TYR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	228	PHE	3.1
1	A	194	ASP	3.1
1	A	83	ALA	3.0
1	A	141	LEU	3.0
1	A	128	LYS	3.0
1	A	133	LEU	2.9
1	A	124	VAL	2.8
1	A	88	LYS	2.7
1	A	79	GLY	2.7
1	A	146	GLU	2.7
1	A	131	GLY	2.7
1	A	125	SER	2.6
1	A	84	ILE	2.5
1	A	138	GLN	2.4
1	A	167	ILE	2.4
1	A	153	SER	2.4
1	A	181	GLU	2.4
1	A	184	LEU	2.4
1	A	193	LYS	2.3
1	A	135	LEU	2.2
1	A	139	ASN	2.2
1	A	337	LYS	2.2
1	A	187	ALA	2.2
1	A	334	PHE	2.2
1	A	190	LEU	2.2
1	A	134	VAL	2.2
1	A	183	GLU	2.1
1	A	67	GLU	2.1
1	A	119	ILE	2.0
1	A	154	ASN	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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.