



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 5, 2016 – 03:04 PM EDT

PDB ID : 5E6M
Title : Crystal structure of human wild type GlyRS bound with tRNAGly
Authors : Xie, W.; Qin, X.; Deng, X.; Chen, L.; Liu, Y.
Deposited on : 2015-10-10
Resolution : 2.93 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
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) : rb-20027939

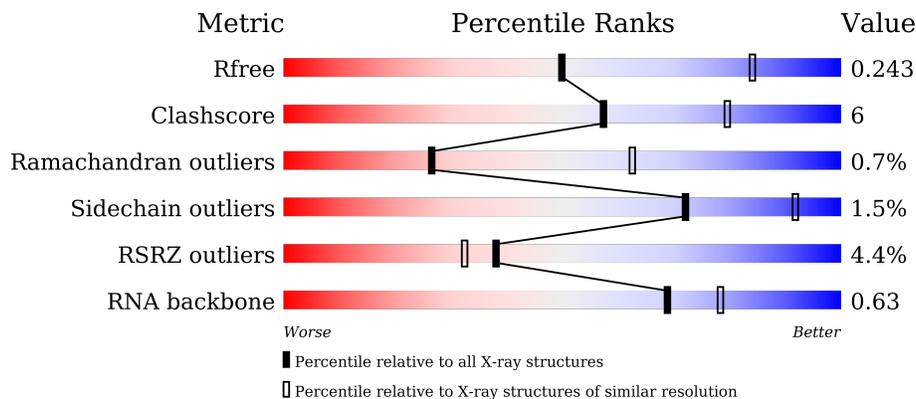
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.93 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	1643 (2.94-2.90)
Clashscore	102246	1871 (2.94-2.90)
Ramachandran outliers	100387	1824 (2.94-2.90)
Sidechain outliers	100360	1826 (2.94-2.90)
RSRZ outliers	91569	1650 (2.94-2.90)
RNA backbone	2183	1004 (3.30-2.54)

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 $\leq 5\%$. 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	693	 4% 62% 12% 25%
1	B	693	 4% 62% 11% 27%
2	C	74	 70% 27%
2	E	74	 66% 28% 5%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 11178 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 Glycine-tRNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	519	Total	C	N	O	S	0	0	0
			3993	2541	696	738	18			
1	B	506	Total	C	N	O	S	0	0	0
			3892	2478	675	722	17			

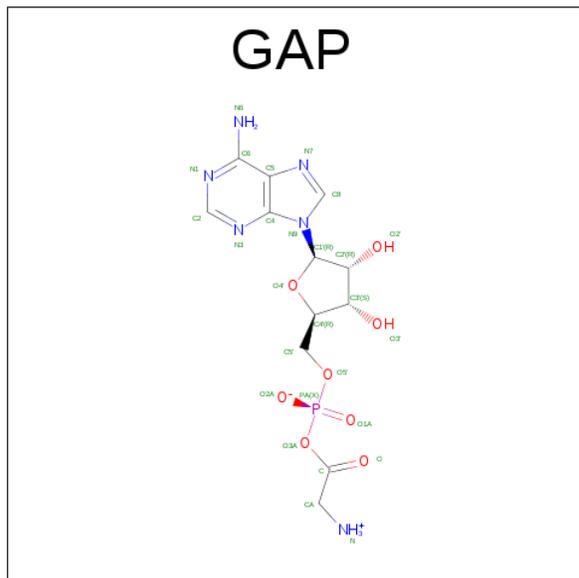
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	686	LEU	-	expression tag	UNP P41250
A	687	GLU	-	expression tag	UNP P41250
A	688	HIS	-	expression tag	UNP P41250
A	689	HIS	-	expression tag	UNP P41250
A	690	HIS	-	expression tag	UNP P41250
A	691	HIS	-	expression tag	UNP P41250
A	692	HIS	-	expression tag	UNP P41250
A	693	HIS	-	expression tag	UNP P41250
B	686	LEU	-	expression tag	UNP P41250
B	687	GLU	-	expression tag	UNP P41250
B	688	HIS	-	expression tag	UNP P41250
B	689	HIS	-	expression tag	UNP P41250
B	690	HIS	-	expression tag	UNP P41250
B	691	HIS	-	expression tag	UNP P41250
B	692	HIS	-	expression tag	UNP P41250
B	693	HIS	-	expression tag	UNP P41250

- Molecule 2 is a RNA chain called tRNA(Gly).

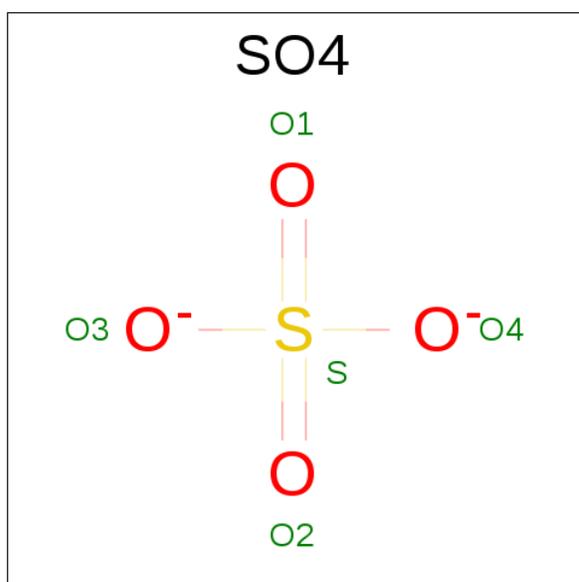
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	74	Total	C	N	O	P	0	0	0
			1580	700	273	531	76			
2	E	74	Total	C	N	O	P	0	0	0
			1580	700	273	531	76			

- Molecule 3 is GLYCYL-ADENOSINE-5'-PHOSPHATE (three-letter code: GAP) (formula: $C_{12}H_{17}N_6O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	27	12	6	8	1	0	0
3	B	1	27	12	6	8	1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



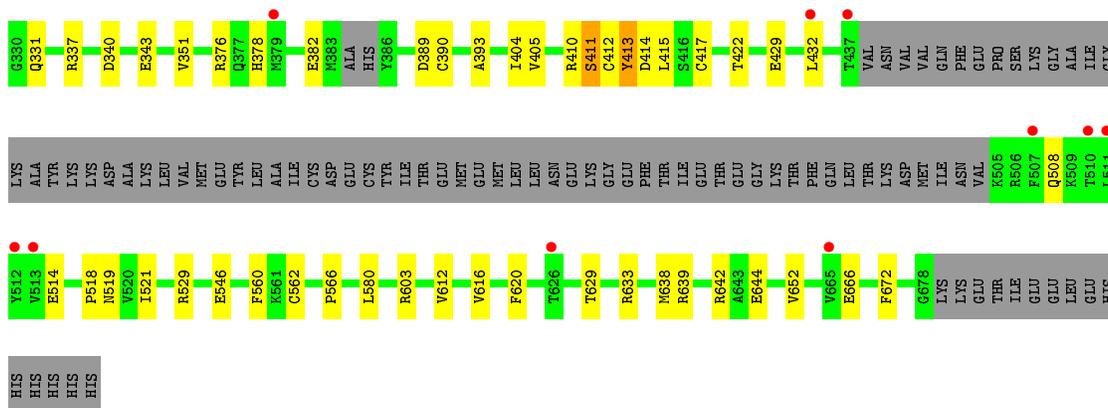
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0

- Molecule 5 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total Ni 1 1	0	0
5	E	1	Total Ni 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	22	Total O 22 22	0	0
6	B	24	Total O 24 24	0	0
6	C	17	Total O 17 17	0	0
6	E	9	Total O 9 9	0	0



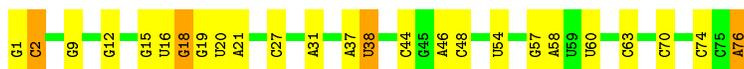
- Molecule 2: tRNA(Gly)

Chain C: 70% 27%



- Molecule 2: tRNA(Gly)

Chain E: 66% 28% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	206.86 Å 122.20 Å 84.73 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.56 – 2.93 49.56 – 2.93	Depositor EDS
% Data completeness (in resolution range)	96.0 (49.56-2.93) 96.0 (49.56-2.93)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.22 (at 2.91 Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.194 , 0.243 0.194 , 0.243	Depositor DCC
R_{free} test set	2281 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	68.5	Xtrriage
Anisotropy	0.041	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 55.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11178	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: GTP, NI, SO4, GAP

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	0.27	0/4078	0.46	0/5524
1	B	0.30	0/3981	0.45	0/5400
2	C	0.25	0/1726	0.76	0/2687
2	E	0.23	0/1726	0.74	1/2687 (0.0%)
All	All	0.27	0/11511	0.57	1/16298 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	C	O5'-P-OP1	5.42	117.20	110.70

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	3993	0	3826	59	0
1	B	3892	0	3703	45	0
2	C	1580	0	798	14	0
2	E	1580	0	798	17	0
3	A	27	0	17	3	0
3	B	27	0	17	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	5	0	0	1	0
5	C	1	0	0	0	0
5	E	1	0	0	0	0
6	A	22	0	0	2	0
6	B	24	0	0	4	0
6	C	17	0	0	0	0
6	E	9	0	0	1	0
All	All	11178	0	9159	121	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 6.

The worst 5 of 121 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)
2:C:54:U:H3	2:C:58:A:H2	1.13	0.97
1:B:140:HIS:HD2	2:C:76:A:H8	1.35	0.73
2:E:12:G:N7	6:E:201:HOH:O	2.22	0.72
1:B:629:THR:HA	1:B:644:GLU:HA	1.72	0.71
2:E:54:U:H3	2:E:58:A:H2	1.37	0.71

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	509/693 (73%)	490 (96%)	16 (3%)	3 (1%)	30	66
1	B	498/693 (72%)	472 (95%)	22 (4%)	4 (1%)	24	58
All	All	1007/1386 (73%)	962 (96%)	38 (4%)	7 (1%)	26	62

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	411	SER
1	A	413	TYR
1	B	150	LYS
1	B	411	SER
1	B	413	TYR

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	407/600 (68%)	400 (98%)	7 (2%)	68	90
1	B	399/600 (66%)	394 (99%)	5 (1%)	76	93
All	All	806/1200 (67%)	794 (98%)	12 (2%)	72	92

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	246	THR
1	A	621	ASP
1	B	414	ASP
1	A	86	ILE
1	B	343	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	72/74 (97%)	7 (9%)	0
2	E	72/74 (97%)	11 (15%)	0
All	All	144/148 (97%)	18 (12%)	0

5 of 18 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	9	G
2	C	12	G
2	C	18	G
2	C	37	A
2	C	38	U

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues 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	GTP	C	1	2	26,34,34	2.62	9 (34%)	29,54,54	1.79	7 (24%)
2	GTP	E	1	2	26,34,34	2.63	9 (34%)	29,54,54	1.71	6 (20%)

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
2	GTP	C	1	2	-	0/18/38/38	0/3/3/3
2	GTP	E	1	2	-	0/18/38/38	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	GTP	C8-N7	-5.40	1.24	1.34
2	E	1	GTP	C8-N7	-5.30	1.24	1.34
2	C	1	GTP	C2'-C3'	-4.70	1.40	1.53
2	E	1	GTP	C2'-C3'	-3.92	1.42	1.53
2	E	1	GTP	C2'-C1'	-3.50	1.48	1.53

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1	GTP	C1'-N9-C4	-5.44	120.73	126.81
2	C	1	GTP	C1'-N9-C4	-4.87	121.37	126.81
2	C	1	GTP	N3-C2-N1	-3.68	122.56	127.56
2	E	1	GTP	N3-C2-N1	-3.53	122.75	127.56
2	C	1	GTP	C5-C6-N1	-3.12	119.45	123.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	GTP	1	0
2	E	1	GTP	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

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	GAP	A	701	-	26,29,29	1.63	2 (7%)	25,43,43	1.60	2 (8%)
4	SO4	A	702	-	4,4,4	0.87	0	6,6,6	0.14	0
3	GAP	B	701	-	26,29,29	1.58	2 (7%)	25,43,43	1.56	1 (4%)

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	GAP	A	701	-	-	0/11/33/33	0/3/3/3
4	SO4	A	702	-	-	0/0/0/0	0/0/0/0
3	GAP	B	701	-	-	0/11/33/33	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	701	GAP	C5-C4	3.03	1.47	1.40
3	A	701	GAP	C5-C4	3.28	1.47	1.40
3	B	701	GAP	O3A-C	6.55	1.46	1.37
3	A	701	GAP	O3A-C	6.63	1.46	1.37

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	701	GAP	N3-C2-N1	-6.64	123.65	128.87
3	A	701	GAP	N3-C2-N1	-6.32	123.91	128.87
3	A	701	GAP	O2A-PA-O3A	2.05	110.50	104.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	701	GAP	3	0
4	A	702	SO4	1	0
3	B	701	GAP	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	519/693 (74%)	0.17	25 (4%) 34 29	32, 56, 108, 137	0
1	B	506/693 (73%)	0.29	27 (5%) 30 24	36, 58, 107, 136	0
2	C	73/74 (98%)	-0.18	0 100 100	44, 57, 91, 96	0
2	E	73/74 (98%)	-0.07	0 100 100	45, 69, 99, 111	0
All	All	1171/1534 (76%)	0.18	52 (4%) 38 32	32, 58, 107, 137	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	LEU	5.4
1	B	626	THR	5.2
1	B	510	THR	5.2
1	A	42	ALA	4.5
1	A	44	LEU	4.1

6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	GTP	C	1	32/32	0.82	0.19	0.17	98,111,156,183	0
2	GTP	E	1	32/32	0.89	0.16	-0.74	84,111,147,159	0

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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	SO4	A	702	5/5	0.94	0.30	1.55	89,91,105,145	0
3	GAP	B	701	27/27	0.96	0.20	0.12	40,50,87,109	0
3	GAP	A	701	27/27	0.96	0.17	-0.85	34,53,82,95	0
5	NI	C	101	1/1	0.93	0.13	-	89,89,89,89	0
5	NI	E	101	1/1	0.87	0.12	-	115,115,115,115	0

6.5 Other polymers [i](#)

There are no such residues in this entry.