



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 12:16 PM EDT

PDB ID : 2DU5  
Title : Crystal structure of Archaeoglobus fulgidus O-phosphoseryl-tRNA synthetase E418N/E420N mutant complexed with tRNA<sup>Opal</sup> and O-phosphoserine ("opal complex")  
Authors : Fukunaga, R.  
Deposited on : 2006-07-20  
Resolution : 3.20 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

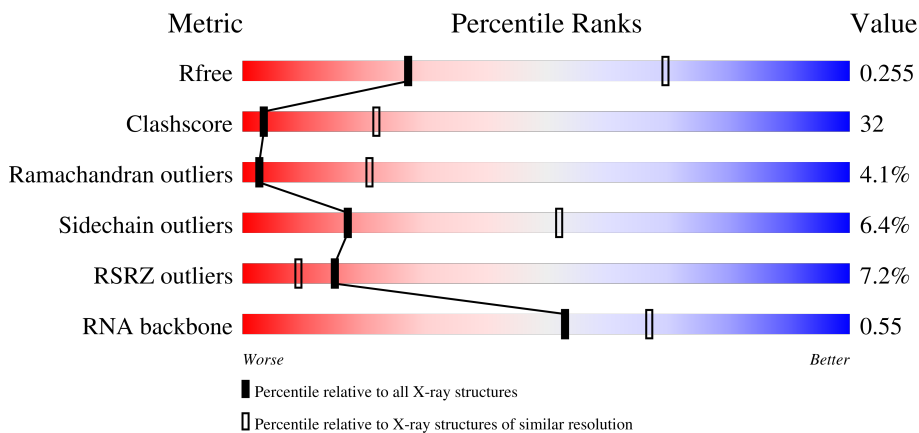
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	D	71	
2	A	534	
2	B	534	

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
3	SEP	A	1001	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9619 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 RNA chain called tRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	D	71	1517	674	269	503	71	0	0	0

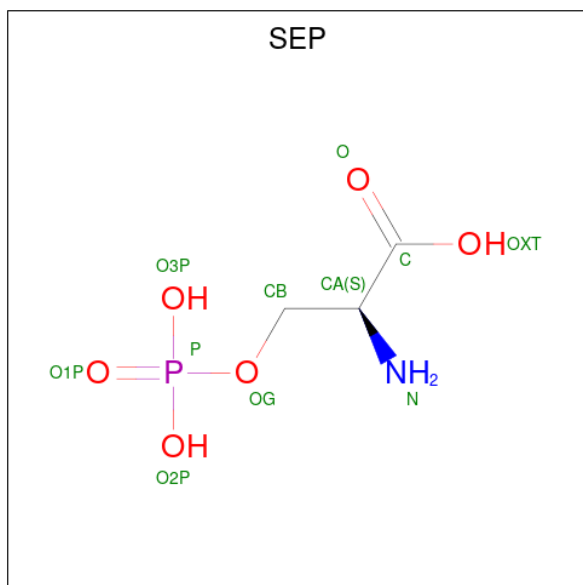
- Molecule 2 is a protein called O-phosphoseryl-tRNA synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	534	4319	2784	721	801	13	0	0	0
2	B	465	3761	2427	632	689	13	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	418	ASN	GLU	engineered mutation	UNP O30126
A	420	ASN	GLU	engineered mutation	UNP O30126
B	418	ASN	GLU	engineered mutation	UNP O30126
B	420	ASN	GLU	engineered mutation	UNP O30126

- Molecule 3 is PHOSPHOSERINE (three-letter code: SEP) (formula: C<sub>3</sub>H<sub>8</sub>NO<sub>6</sub>P).

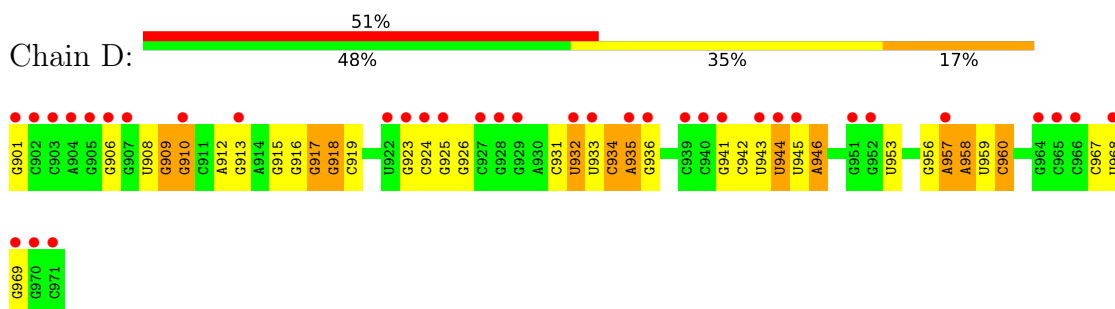


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	11	3	1	6	1	0	0
3	A	1	11	3	1	6	1	0	0

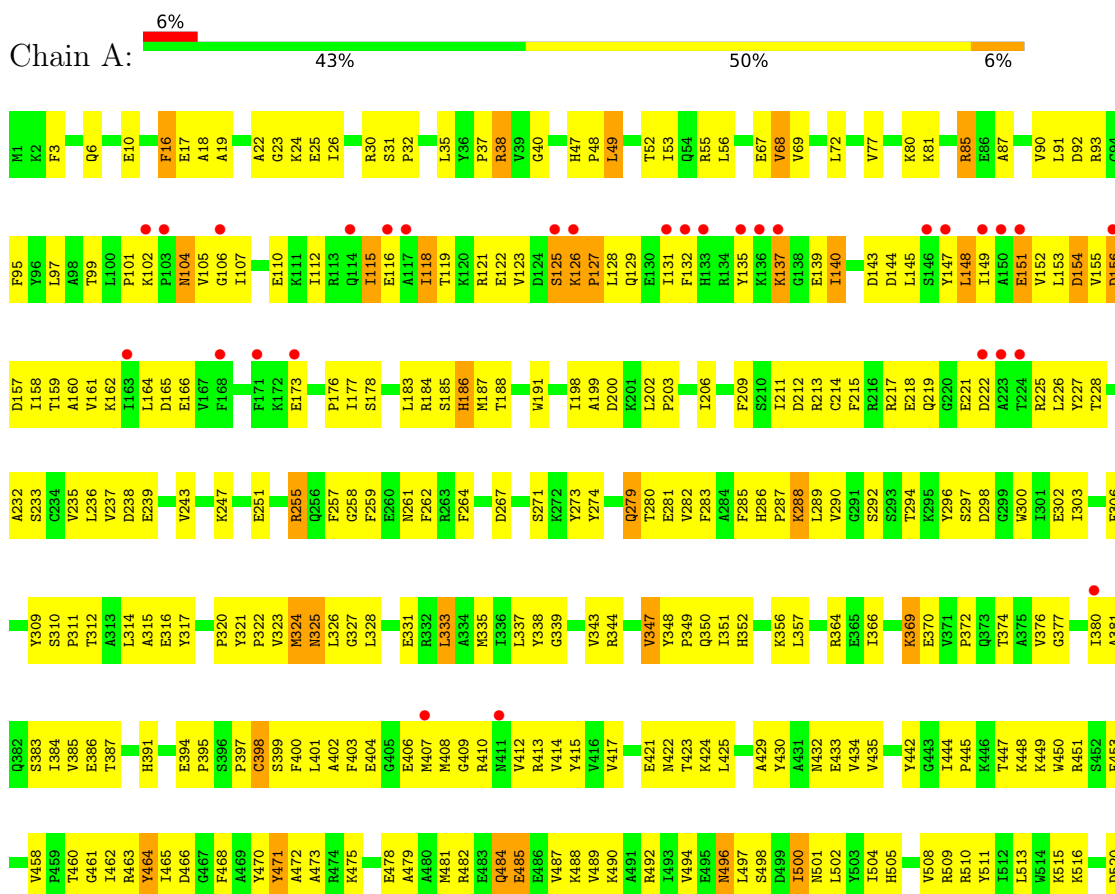
### 3 Residue-property plots [i](#)

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: tRNA

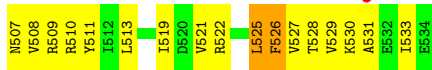
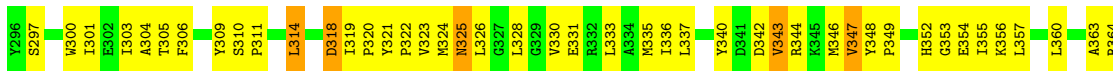
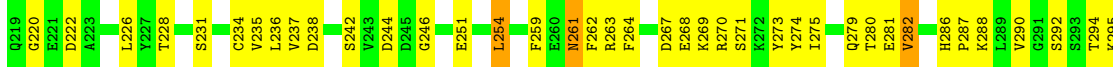
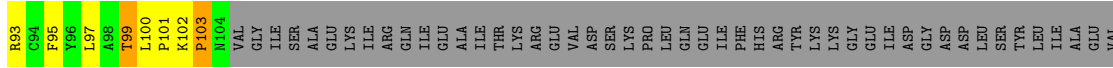
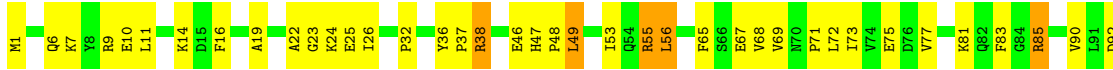


#### • Molecule 2: O-phosphoseryl-tRNA synthetase





● Molecule 2: O-phosphoseryl-tRNA synthetase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.04Å 148.04Å 152.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.07 – 3.20 49.07 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.07-3.20) 99.9 (49.07-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.60 (at 3.19Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.223 , 0.299 0.190 , 0.255	Depositor DCC
$R_{free}$ test set	1603 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	102.1	Xtrriage
Anisotropy	0.036	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 116.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.038 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9619	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	113.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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: SEP

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	D	0.32	1/1693 (0.1%)	0.67	0/2637
2	A	0.31	0/4416	0.55	0/5966
2	B	0.33	0/3849	0.58	0/5200
All	All	0.32	1/9958 (0.0%)	0.59	0/13803

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	901	G	OP3-P	-7.04	1.52	1.61

There are no bond angle outliers.

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	D	1517	0	768	35	0
2	A	4319	0	4339	330	0
2	B	3761	0	3774	249	0
3	A	22	0	10	5	0
All	All	9619	0	8891	593	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 32.

The worst 5 of 593 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:B:464:TYR:HA	2:B:500:ILE:HD11	1.32	1.07
2:B:448:LYS:H	2:B:448:LYS:HD3	1.20	1.05
2:B:55:ARG:HG2	2:B:55:ARG:HH21	1.34	0.92
2:B:47:HIS:HD2	2:B:49:LEU:H	1.13	0.92
2:A:460:THR:HG23	2:A:462:ILE:H	1.36	0.89

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	
2	A	532/534 (100%)	417 (78%)	88 (16%)	27 (5%)	2	15
2	B	461/534 (86%)	376 (82%)	71 (15%)	14 (3%)	4	28
All	All	993/1068 (93%)	793 (80%)	159 (16%)	41 (4%)	3	21

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	118	ILE
2	A	398	CYS
2	A	526	PHE
2	B	14	LYS
2	B	484	GLN

### 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	
2	A	464/464 (100%)	437 (94%)	27 (6%)	20	55
2	B	402/464 (87%)	374 (93%)	28 (7%)	15	48
All	All	866/928 (93%)	811 (94%)	55 (6%)	17	52

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	49	LEU
2	B	236	LEU
2	B	525	LEU
2	B	485	GLU
2	B	55	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	382	GLN
2	B	496	ASN
2	B	432	ASN
2	A	422	ASN
2	B	286	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	D	70/71 (98%)	15 (21%)	3 (4%)

5 of 15 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	D	906	G
1	D	908	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	909	G
1	D	910	G
1	D	917	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	D	917	G
1	D	934	C
1	D	957	A

## 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 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	SEP	A	1002	-	9,10,10	1.53	1 (11%)	12,14,14	1.74	2 (16%)
3	SEP	A	1001	-	9,10,10	1.53	1 (11%)	12,14,14	1.65	2 (16%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SEP	A	1002	-	-	1/10/10/10	-
3	SEP	A	1001	-	-	7/10/10/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	SEP	P-O1P	3.37	1.61	1.50
3	A	1002	SEP	P-O1P	3.36	1.61	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1002	SEP	OG-CB-CA	4.78	112.23	108.06
3	A	1001	SEP	OG-CB-CA	4.48	111.97	108.06
3	A	1002	SEP	O3P-P-OG	2.28	112.80	106.73
3	A	1001	SEP	O3P-P-OG	2.20	112.60	106.73

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	SEP	CB-OG-P-O1P
3	A	1001	SEP	CB-OG-P-O2P
3	A	1001	SEP	CB-OG-P-O3P
3	A	1001	SEP	O-C-CA-CB
3	A	1001	SEP	OXT-C-CA-CB

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1002	SEP	1	0
3	A	1001	SEP	4	0

## 5.7 Other polymers [i](#)

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	D	71/71 (100%)	2.28	36 (50%) 0 0	104, 195, 200, 200	0
2	A	534/534 (100%)	0.33	33 (6%) 20 11	20, 92, 197, 202	0
2	B	465/534 (87%)	-0.03	8 (1%) 70 57	25, 80, 185, 202	0
All	All	1070/1139 (93%)	0.30	77 (7%) 15 9	20, 91, 198, 202	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	223	ALA	8.9
1	D	969	G	7.7
1	D	970	G	7.5
1	D	903	C	5.6
2	A	149	ILE	4.9

### 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 monosaccharides 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	SEP	A	1001	11/11	0.94	0.25	41,103,137,172	0
3	SEP	A	1002	11/11	0.95	0.22	41,77,116,154	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.