



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 26, 2023 – 02:03 PM EDT

PDB ID : 3EFM  
Title : Structure of the alcaligin outer membrane recepteur FauA from Bordetella pertussis  
Authors : Brillet, K.; Lauber, E.; Reimann, C.; Armstrong, S.K.; Cobessi, D.  
Deposited on : 2008-09-09  
Resolution : 2.33 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](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

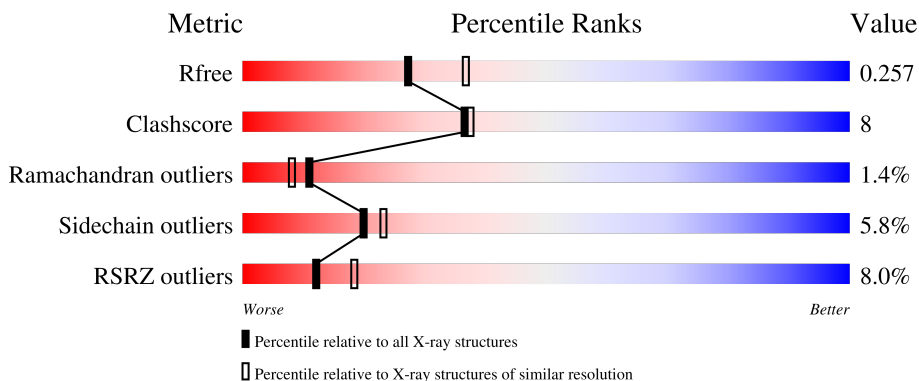
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.33 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	A	707	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4530 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 Ferric alcaligin siderophore receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	572	4384	2738	778	860	8	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	348Q	GLY	-	expression tag	UNP Q9X6A5
A	348R	SER	-	expression tag	UNP Q9X6A5
A	348S	HIS	-	expression tag	UNP Q9X6A5
A	348T	HIS	-	expression tag	UNP Q9X6A5
A	348U	HIS	-	expression tag	UNP Q9X6A5
A	348V	HIS	-	expression tag	UNP Q9X6A5
A	348W	HIS	-	expression tag	UNP Q9X6A5
A	348X	HIS	-	expression tag	UNP Q9X6A5

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

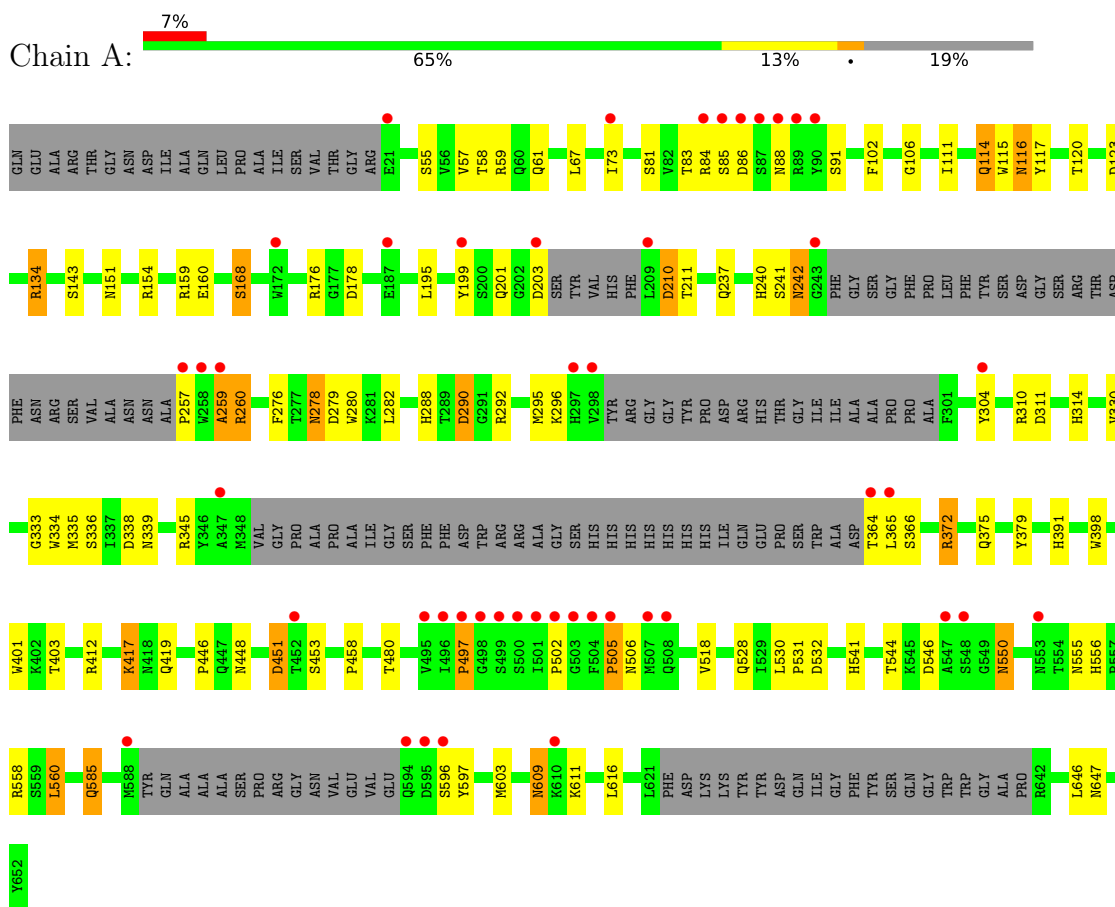
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	136	Total	O	0	0
			136	136		

### 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: Ferric alcaligin siderophore receptor



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	165.89Å 188.85Å 62.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.35 – 2.33 29.35 – 2.33	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.35-2.33) 99.9 (29.35-2.33)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 2.34Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.219 , 0.258 0.220 , 0.257	Depositor DCC
$R_{free}$ test set	2141 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.9	Xtrriage
Anisotropy	0.417	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4530	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.40% 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: SO4

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	0.71	0/4474	0.83	10/6081 (0.2%)

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	134	ARG	NE-CZ-NH2	-10.69	114.95	120.30
1	A	154	ARG	NE-CZ-NH2	-8.73	115.93	120.30
1	A	59	ARG	NE-CZ-NH2	-7.92	116.34	120.30
1	A	134	ARG	NE-CZ-NH1	6.97	123.79	120.30
1	A	154	ARG	NE-CZ-NH1	6.76	123.68	120.30
1	A	560	LEU	CA-CB-CG	6.54	130.35	115.30
1	A	505	PRO	N-CA-CB	6.38	110.96	103.30
1	A	257	PRO	N-CA-CB	5.99	110.49	103.30
1	A	497	PRO	N-CA-CB	5.90	110.38	103.30
1	A	502	PRO	N-CA-CB	5.28	109.63	103.30

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	4384	0	4110	67	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	10	0	0	0	0
3	A	136	0	0	3	0
All	All	4530	0	4110	67	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 8.

All (67) 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:417:LYS:N	1:A:417:LYS:HD3	1.64	1.10
1:A:417:LYS:HD3	1:A:417:LYS:H	0.95	1.07
1:A:310:ARG:HE	1:A:339:ASN:HD21	0.96	0.96
1:A:58:THR:H	1:A:61:GLN:HE21	1.05	0.94
1:A:417:LYS:H	1:A:417:LYS:CD	1.80	0.94
1:A:55:SER:OG	1:A:134:ARG:NH2	2.12	0.82
1:A:310:ARG:NE	1:A:339:ASN:HD21	1.78	0.81
1:A:58:THR:H	1:A:61:GLN:NE2	1.78	0.81
1:A:123:ASP:H	1:A:237:GLN:HE22	1.30	0.79
1:A:143:SER:H	1:A:375:GLN:HE22	1.25	0.79
1:A:310:ARG:HE	1:A:339:ASN:ND2	1.80	0.79
1:A:609:ASN:HD22	1:A:611:LYS:H	1.34	0.73
1:A:609:ASN:ND2	1:A:611:LYS:H	1.89	0.70
1:A:295:MET:HB3	1:A:304:TYR:HB2	1.73	0.69
1:A:106:GLY:H	1:A:151:ASN:HD21	1.44	0.64
1:A:518:VAL:HG22	1:A:544:THR:HG22	1.78	0.64
1:A:334:TRP:CH2	1:A:336:SER:HB2	2.38	0.59
1:A:417:LYS:HE3	3:A:690:HOH:O	2.02	0.59
1:A:259:ALA:O	1:A:260:ARG:HB3	2.03	0.58
1:A:288:HIS:ND1	1:A:311:ASP:OD2	2.30	0.58
1:A:333:GLY:HA3	1:A:379:TYR:CZ	2.40	0.56
1:A:417:LYS:NZ	1:A:458:PRO:HA	2.20	0.56
1:A:84:ARG:O	1:A:86:ASP:N	2.39	0.55
1:A:345:ARG:HD2	1:A:366:SER:O	2.07	0.55
1:A:541:HIS:HA	1:A:558:ARG:O	2.05	0.55
1:A:143:SER:H	1:A:375:GLN:NE2	2.01	0.54
1:A:451:ASP:HB3	1:A:453:SER:H	1.71	0.54
1:A:178:ASP:OD1	3:A:771:HOH:O	2.19	0.53
1:A:290:ASP:OD2	1:A:290:ASP:N	2.41	0.52
1:A:116:ASN:HD21	1:A:120:THR:H	1.59	0.51
1:A:609:ASN:HD22	1:A:609:ASN:C	2.14	0.50

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:546:ASP:OD2	1:A:550:ASN:ND2	2.45	0.50
1:A:585:GLN:HB2	1:A:597:TYR:CZ	2.47	0.50
1:A:168:SER:OG	1:A:647:ASN:ND2	2.46	0.49
1:A:338:ASP:OD1	1:A:372:ARG:NH1	2.46	0.49
1:A:123:ASP:OD2	1:A:176:ARG:NH2	2.46	0.49
1:A:314:HIS:HE1	1:A:379:TYR:OH	1.95	0.49
1:A:81:SER:HB3	1:A:585:GLN:HE21	1.79	0.48
1:A:412:ARG:HD3	1:A:448:ASN:ND2	2.28	0.48
1:A:83:THR:HG21	1:A:556:HIS:CD2	2.49	0.48
1:A:102:PHE:CD1	1:A:111:ILE:HD13	2.48	0.48
1:A:242:ASN:C	1:A:242:ASN:HD22	2.17	0.48
1:A:211:THR:HG22	1:A:241:SER:CB	2.44	0.47
1:A:531:PRO:O	1:A:532:ASP:HB2	2.14	0.47
1:A:114:GLN:O	1:A:117:TYR:HE1	1.97	0.47
1:A:278:ASN:ND2	1:A:280:TRP:H	2.12	0.47
1:A:391:HIS:HD2	3:A:669:HOH:O	1.98	0.46
1:A:115:TRP:CZ3	1:A:339:ASN:HB3	2.50	0.46
1:A:401:TRP:CD2	1:A:446:PRO:HG2	2.50	0.46
1:A:278:ASN:O	1:A:279:ASP:HB2	2.16	0.46
1:A:143:SER:N	1:A:375:GLN:HE22	2.03	0.46
1:A:106:GLY:H	1:A:151:ASN:ND2	2.12	0.45
1:A:401:TRP:CE2	1:A:446:PRO:HG2	2.52	0.45
1:A:603:MET:HA	1:A:616:LEU:O	2.17	0.45
1:A:211:THR:HG22	1:A:241:SER:HB2	1.99	0.44
1:A:364:THR:HG22	1:A:365:LEU:N	2.32	0.44
1:A:276:PHE:HB2	1:A:278:ASN:ND2	2.33	0.44
1:A:203:ASP:HB3	1:A:210:ASP:HA	2.00	0.42
1:A:199:TYR:CZ	1:A:201:GLN:HG3	2.54	0.42
1:A:57:VAL:HA	1:A:61:GLN:NE2	2.34	0.42
1:A:240:HIS:HE1	1:A:242:ASN:OD1	2.03	0.42
1:A:398:TRP:HA	1:A:419:GLN:O	2.20	0.41
1:A:67:LEU:HD13	1:A:73:ILE:HA	2.02	0.41
1:A:417:LYS:HZ3	1:A:458:PRO:HA	1.83	0.41
1:A:123:ASP:N	1:A:237:GLN:HE22	2.08	0.40
1:A:401:TRP:CH2	1:A:403:THR:HB	2.57	0.40
1:A:528:GLN:HG2	1:A:530:LEU:O	2.22	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	558/707 (79%)	529 (95%)	21 (4%)	8 (1%)	11 8

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	85	SER
1	A	497	PRO
1	A	505	PRO
1	A	88	ASN
1	A	259	ALA
1	A	451	ASP
1	A	506	ASN
1	A	260	ARG

### 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	447/582 (77%)	421 (94%)	26 (6%)	20 23

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	91	SER
1	A	114	GLN
1	A	116	ASN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	159	ARG
1	A	160	GLU
1	A	168	SER
1	A	195	LEU
1	A	210	ASP
1	A	242	ASN
1	A	278	ASN
1	A	282	LEU
1	A	290	ASP
1	A	292	ARG
1	A	296	LYS
1	A	330	VAL
1	A	335	MET
1	A	372	ARG
1	A	417	LYS
1	A	480	THR
1	A	550	ASN
1	A	555	ASN
1	A	560	LEU
1	A	585	GLN
1	A	596	SER
1	A	609	ASN
1	A	646	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	60	GLN
1	A	61	GLN
1	A	101	ASN
1	A	116	ASN
1	A	151	ASN
1	A	237	GLN
1	A	240	HIS
1	A	242	ASN
1	A	278	ASN
1	A	314	HIS
1	A	339	ASN
1	A	375	GLN
1	A	391	HIS
1	A	405	GLN
1	A	445	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	448	ASN
1	A	550	ASN
1	A	556	HIS
1	A	609	ASN
1	A	619	ASN
1	A	647	ASN

### 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 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	$\# Z  > 2$
2	SO4	A	653	-	4,4,4	0.17	0	6,6,6	0.24	0
2	SO4	A	654	-	4,4,4	0.16	0	6,6,6	0.38	0

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 [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, 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	A	572/707 (80%)	0.42	46 (8%) <b>12</b> <b>18</b>	28, 44, 65, 80	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	497	PRO	8.0
1	A	87	SER	7.5
1	A	504	PHE	6.9
1	A	503	GLY	6.8
1	A	88	ASN	6.4
1	A	203	ASP	5.2
1	A	90	TYR	5.1
1	A	364	THR	5.1
1	A	85	SER	4.7
1	A	495	VAL	4.5
1	A	258	TRP	4.4
1	A	507	MET	4.4
1	A	498	GLY	4.3
1	A	588	MET	4.1
1	A	594	GLN	4.0
1	A	86	ASP	3.9
1	A	595	ASP	3.9
1	A	365	LEU	3.9
1	A	199	TYR	3.7
1	A	499	SER	3.7
1	A	259	ALA	3.6
1	A	502	PRO	3.5
1	A	172	TRP	3.4
1	A	297	HIS	3.4
1	A	452	THR	3.4
1	A	257	PRO	3.3
1	A	304	TYR	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	347	ALA	3.0
1	A	547	ALA	2.9
1	A	596	SER	2.9
1	A	500	SER	2.8
1	A	243	GLY	2.8
1	A	187	GLU	2.7
1	A	89	ARG	2.7
1	A	21	GLU	2.6
1	A	610	LYS	2.5
1	A	501	ILE	2.5
1	A	298	VAL	2.5
1	A	496	ILE	2.4
1	A	84	ARG	2.3
1	A	505	PRO	2.3
1	A	209	LEU	2.2
1	A	553	ASN	2.2
1	A	73	ILE	2.2
1	A	548	SER	2.2
1	A	508	GLN	2.1

## 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(Å <sup>2</sup> )	Q<0.9
2	SO4	A	653	5/5	0.96	0.23	83,83,83,84	0
2	SO4	A	654	5/5	0.96	0.16	80,81,81,82	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.