wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2022 – 01:12 pm BST

PDB ID	:	7PEI
Title	:	Crystal Structure of a Class D carbapenemase_E185A/R186A/R206A
Authors	:	Zhou, Q.; He, Y.; Jin, Y.
Deposited on	:	2021-08-10
Resolution	:	1.90 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

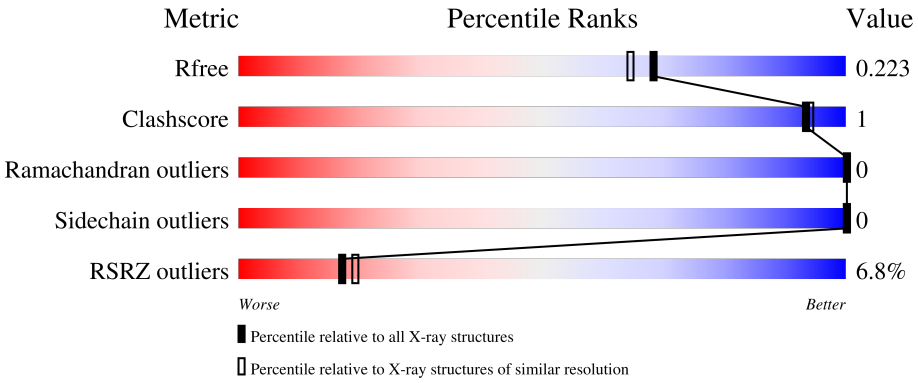
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 of 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	AAA	260	4% 90% · 7%
1	BBB	260	8% 90% · 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8196 atoms, of which 3921 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 Beta-lactamase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	242	Total	C	H	N	O	S	47	2	0
			3917	1262	1939	344	364	8			
1	BBB	242	Total	C	H	N	O	S	51	5	0
			3940	1270	1950	345	367	8			

There are 40 discrepancies between the modelled and reference sequences:

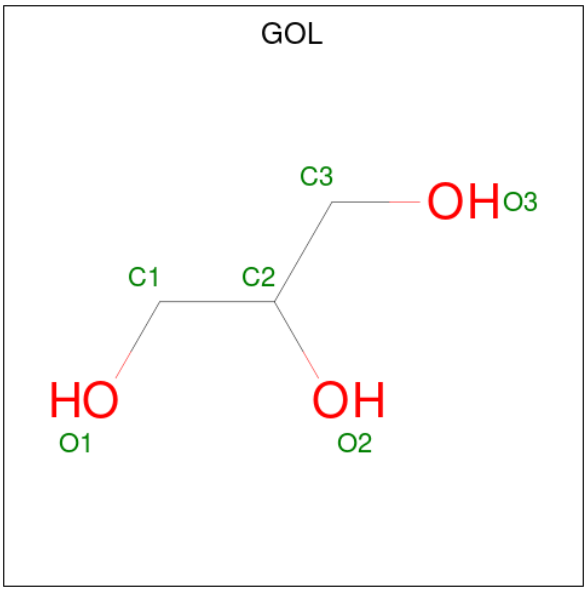
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	6	MET	-	initiating methionine	UNP A0A482LRD5
AAA	7	HIS	-	expression tag	UNP A0A482LRD5
AAA	8	HIS	-	expression tag	UNP A0A482LRD5
AAA	9	HIS	-	expression tag	UNP A0A482LRD5
AAA	10	HIS	-	expression tag	UNP A0A482LRD5
AAA	11	HIS	-	expression tag	UNP A0A482LRD5
AAA	12	HIS	-	expression tag	UNP A0A482LRD5
AAA	13	SER	-	expression tag	UNP A0A482LRD5
AAA	14	ALA	-	expression tag	UNP A0A482LRD5
AAA	15	GLY	-	expression tag	UNP A0A482LRD5
AAA	16	GLU	-	expression tag	UNP A0A482LRD5
AAA	17	ASN	-	expression tag	UNP A0A482LRD5
AAA	18	LEU	-	expression tag	UNP A0A482LRD5
AAA	19	TYR	-	expression tag	UNP A0A482LRD5
AAA	20	PHE	-	expression tag	UNP A0A482LRD5
AAA	21	GLN	-	expression tag	UNP A0A482LRD5
AAA	22	GLY	-	expression tag	UNP A0A482LRD5
AAA	185	ALA	GLU	engineered mutation	UNP A0A482LRD5
AAA	186	ALA	ARG	engineered mutation	UNP A0A482LRD5
AAA	206	ALA	ARG	engineered mutation	UNP A0A482LRD5
BBB	6	MET	-	initiating methionine	UNP A0A482LRD5
BBB	7	HIS	-	expression tag	UNP A0A482LRD5
BBB	8	HIS	-	expression tag	UNP A0A482LRD5
BBB	9	HIS	-	expression tag	UNP A0A482LRD5
BBB	10	HIS	-	expression tag	UNP A0A482LRD5

Continued on next page...

Continued from previous page...

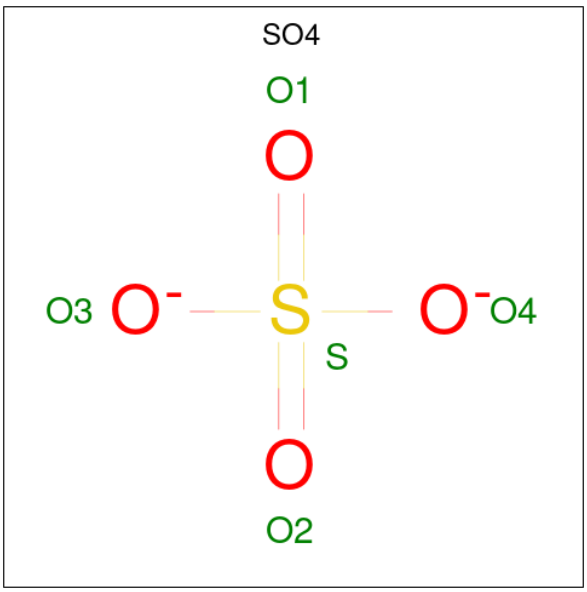
Chain	Residue	Modelled	Actual	Comment	Reference
BBB	11	HIS	-	expression tag	UNP A0A482LRD5
BBB	12	HIS	-	expression tag	UNP A0A482LRD5
BBB	13	SER	-	expression tag	UNP A0A482LRD5
BBB	14	ALA	-	expression tag	UNP A0A482LRD5
BBB	15	GLY	-	expression tag	UNP A0A482LRD5
BBB	16	GLU	-	expression tag	UNP A0A482LRD5
BBB	17	ASN	-	expression tag	UNP A0A482LRD5
BBB	18	LEU	-	expression tag	UNP A0A482LRD5
BBB	19	TYR	-	expression tag	UNP A0A482LRD5
BBB	20	PHE	-	expression tag	UNP A0A482LRD5
BBB	21	GLN	-	expression tag	UNP A0A482LRD5
BBB	22	GLY	-	expression tag	UNP A0A482LRD5
BBB	185	ALA	GLU	engineered mutation	UNP A0A482LRD5
BBB	186	ALA	ARG	engineered mutation	UNP A0A482LRD5
BBB	206	ALA	ARG	engineered mutation	UNP A0A482LRD5

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		
2	AAA	1	Total	C	H	O	2	0
			14	3	8	3		
2	BBB	1	Total	C	H	O	2	0
			14	3	8	3		
2	BBB	1	Total	C	H	O	2	0
			14	3	8	3		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total O S 5 4 1	0	0
3	AAA	1	Total O S 5 4 1	0	0
3	AAA	1	Total O S 5 4 1	0	0
3	BBB	1	Total O S 5 4 1	0	0
3	BBB	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

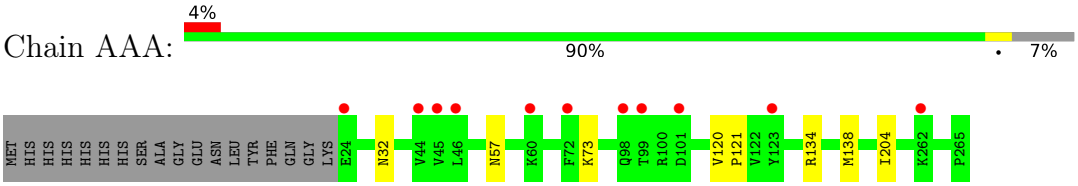
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	145	Total O 146 146	0	1
4	BBB	112	Total O 112 112	0	0

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.

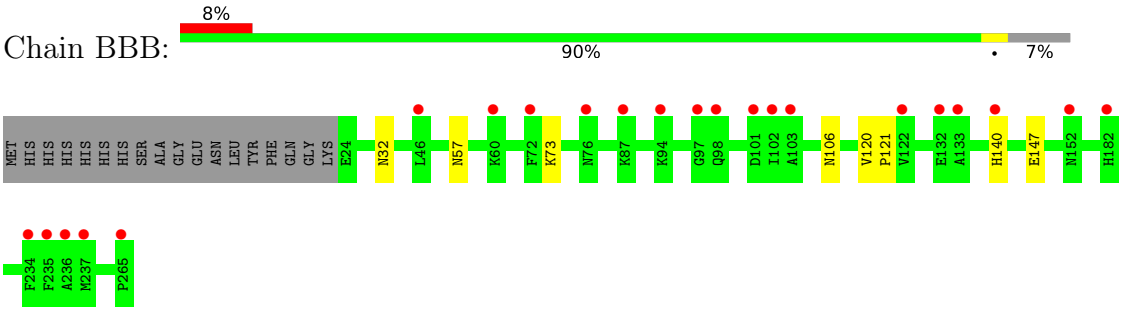
• Molecule 1: Beta-lactamase

Chain AAA:



• Molecule 1: Beta-lactamase

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62	Depositor
Cell constants a, b, c, α, β, γ	146.19Å 146.19Å 54.89Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.69 – 1.90 19.69 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.69-1.90) 100.0 (19.69-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.177 , 0.211 0.189 , 0.223	Depositor DCC
R_{free} test set	2657 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	33.6	Xtriage
Anisotropy	0.838	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.030 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8196	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% 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: KCX, GOL, 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	AAA	0.77	0/2020	0.85	0/2733
1	BBB	0.76	0/2042	0.80	0/2763
All	All	0.77	0/4062	0.82	0/5496

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	AAA	1978	1939	1932	4	0
1	BBB	1990	1950	1945	4	0
2	AAA	12	16	16	2	0
2	BBB	12	16	16	1	0
3	AAA	15	0	0	1	0
3	BBB	10	0	0	0	0
4	AAA	146	0	0	2	0
4	BBB	112	0	0	2	0
All	All	4275	3921	3909	11	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 1.

The worst 5 of 11 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:BBB:302:GOL:O1	4:BBB:401:HOH:O	2.03	0.76
3:AAA:305:SO4:O4	4:AAA:401:HOH:O	2.07	0.70
1:AAA:134:ARG:O	1:AAA:138:MET:HG2	2.00	0.62
2:AAA:302:GOL:H11	4:AAA:486:HOH:O	2.07	0.54
1:AAA:204:ILE:HG21	2:AAA:302:GOL:H12	1.92	0.50

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	AAA	241/260 (93%)	236 (98%)	5 (2%)	0	100	100
1	BBB	244/260 (94%)	239 (98%)	5 (2%)	0	100	100
All	All	485/520 (93%)	475 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	209/222 (94%)	209 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	BBB	212/222 (96%)	212 (100%)	0	100	100
All	All	421/444 (95%)	421 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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	$\#\|Z\| > 2$
1	KCX	AAA	73	1	9,11,12	0.95	1 (11%)	5,12,14	0.68	0
1	KCX	BBB	73	1	9,11,12	0.98	1 (11%)	5,12,14	0.85	0

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
1	KCX	AAA	73	1	-	0/9/10/12	-
1	KCX	BBB	73	1	-	0/9/10/12	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	73	KCX	CX-NZ	2.21	1.39	1.35
1	AAA	73	KCX	CX-NZ	2.17	1.38	1.35

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.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

9 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	GOL	AAA	302	-	5,5,5	0.18	0	5,5,5	0.52	0
2	GOL	BBB	301	-	5,5,5	0.18	0	5,5,5	0.26	0
2	GOL	BBB	302	-	5,5,5	0.05	0	5,5,5	0.21	0
3	SO4	BBB	303	-	4,4,4	0.43	0	6,6,6	0.20	0
3	SO4	AAA	304	-	4,4,4	0.26	0	6,6,6	0.37	0
3	SO4	AAA	303	-	4,4,4	0.27	0	6,6,6	0.19	0
3	SO4	BBB	304	-	4,4,4	0.44	0	6,6,6	0.09	0
3	SO4	AAA	305	-	4,4,4	0.38	0	6,6,6	0.11	0
2	GOL	AAA	301	-	5,5,5	0.14	0	5,5,5	0.36	0

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	GOL	AAA	302	-	-	2/4/4/4	-
2	GOL	AAA	301	-	-	2/4/4/4	-
2	GOL	BBB	301	-	-	2/4/4/4	-
2	GOL	BBB	302	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	302	GOL	C1-C2-C3-O3
2	BBB	301	GOL	C1-C2-C3-O3
2	BBB	302	GOL	C1-C2-C3-O3
2	AAA	301	GOL	C1-C2-C3-O3
2	AAA	302	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	302	GOL	2	0
2	BBB	302	GOL	1	0
3	AAA	305	SO4	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

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	AAA	241/260 (92%)	0.37	11 (4%) 32 35	22, 35, 56, 76	1 (0%)
1	BBB	241/260 (92%)	0.55	22 (9%) 9 10	24, 41, 62, 79	4 (1%)
All	All	482/520 (92%)	0.46	33 (6%) 17 19	22, 37, 61, 79	5 (1%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	101	ASP	4.0
1	AAA	44	VAL	3.8
1	BBB	97	GLY	3.5
1	BBB	76	ASN	3.3
1	BBB	102	ILE	3.3

6.2 Non-standard residues in protein, DNA, RNA chains

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
1	KCX	BBB	73	12/13	0.90	0.15	30,34,43,43	1
1	KCX	AAA	73	12/13	0.92	0.18	24,29,33,38	1

6.3 Carbohydrates

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^{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	GOL	AAA	302	6/6	0.83	0.36	38,45,49,49	14
2	GOL	BBB	301	6/6	0.84	0.29	45,48,56,56	14
3	SO4	AAA	304	5/5	0.84	0.15	42,42,51,56	5
2	GOL	BBB	302	6/6	0.88	0.27	31,44,49,49	14
2	GOL	AAA	301	6/6	0.89	0.19	41,46,52,52	14
3	SO4	AAA	305	5/5	0.90	0.25	41,49,55,57	5
3	SO4	BBB	304	5/5	0.91	0.14	41,47,56,56	5
3	SO4	AAA	303	5/5	0.93	0.12	38,42,46,52	5
3	SO4	BBB	303	5/5	0.95	0.11	38,42,46,49	5

6.5 Other polymers (i)

There are no such residues in this entry.