



# Preliminary Full wwPDB X-ray Structure Validation Report ⓘ

Aug 30, 2021 – 06:28 PM JST

Deposition ID : D\_1300024350

This is a Preliminary Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB Deposition System during initial deposition but before annotation of the structure.

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 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	:	2.23.1
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)	:	2.23.1

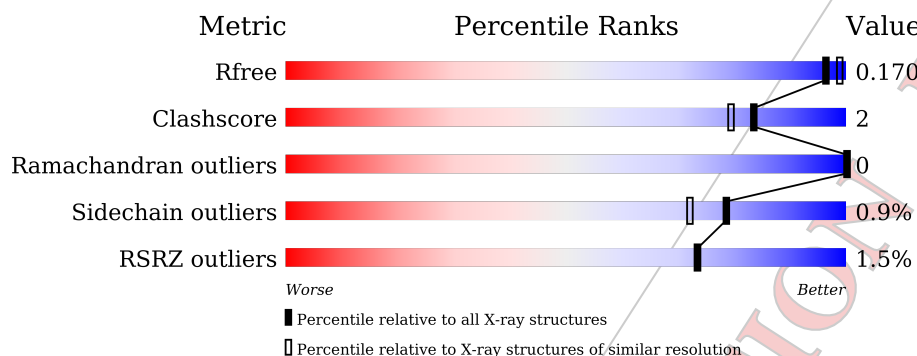
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.77 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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	884	<div> <div></div> <div>89%</div> <div>8%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8071 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 Alanine aminopeptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	860	6806	4287	1193	1307	19	0	8	0

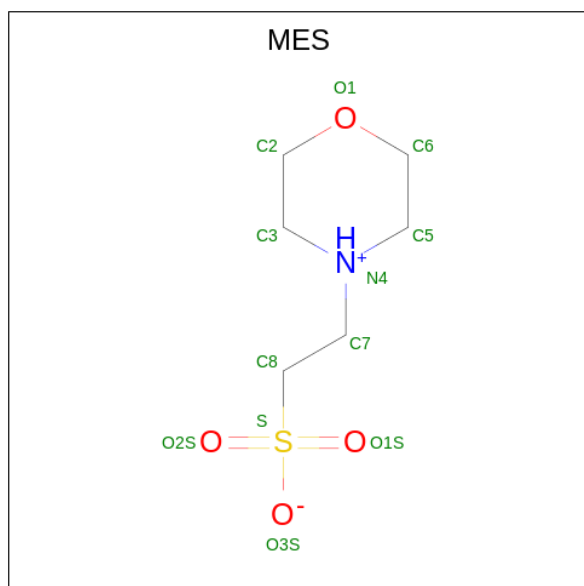
There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	initiating methionine	UNP A4F9D7
A	-22	GLY	-	expression tag	UNP A4F9D7
A	-21	SER	-	expression tag	UNP A4F9D7
A	-20	SER	-	expression tag	UNP A4F9D7
A	-19	HIS	-	expression tag	UNP A4F9D7
A	-18	HIS	-	expression tag	UNP A4F9D7
A	-17	HIS	-	expression tag	UNP A4F9D7
A	-16	HIS	-	expression tag	UNP A4F9D7
A	-15	HIS	-	expression tag	UNP A4F9D7
A	-14	HIS	-	expression tag	UNP A4F9D7
A	-13	HIS	-	expression tag	UNP A4F9D7
A	-12	HIS	-	expression tag	UNP A4F9D7
A	-11	HIS	-	expression tag	UNP A4F9D7
A	-10	HIS	-	expression tag	UNP A4F9D7
A	-9	SER	-	expression tag	UNP A4F9D7
A	-8	SER	-	expression tag	UNP A4F9D7
A	-7	GLY	-	expression tag	UNP A4F9D7
A	-6	LEU	-	expression tag	UNP A4F9D7
A	-5	VAL	-	expression tag	UNP A4F9D7
A	-4	PRO	-	expression tag	UNP A4F9D7
A	-3	ARG	-	expression tag	UNP A4F9D7
A	-2	GLY	-	expression tag	UNP A4F9D7
A	-1	SER	-	expression tag	UNP A4F9D7
A	0	HIS	-	expression tag	UNP A4F9D7
A	1	VAL	-	expression tag	UNP A4F9D7
A	802	ARG	GLU	engineered mutation	UNP A4F9D7

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



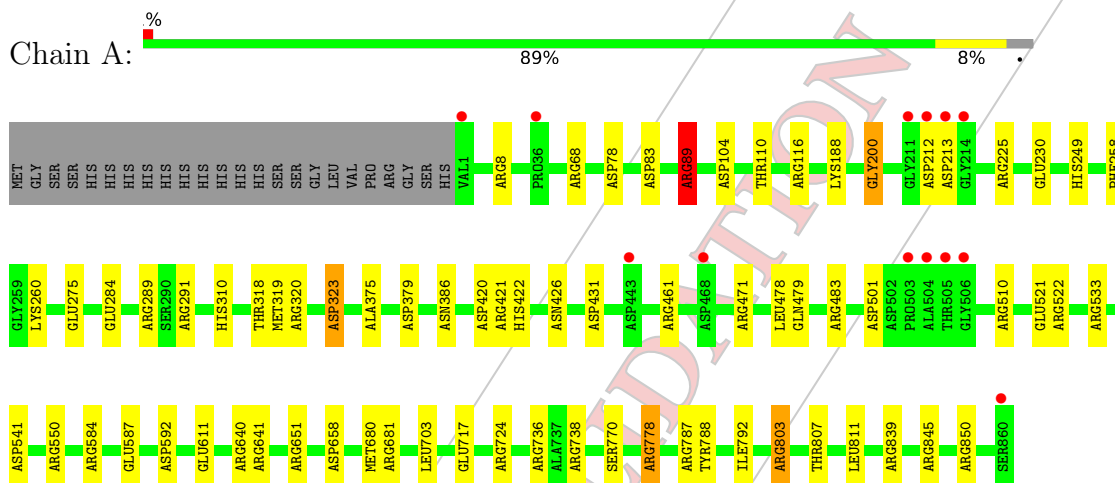
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	S	1252	Total	O	0	0
			1252	1252		



- Molecule 1: Alanine aminopeptidase



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.91Å 143.58Å 60.58Å 90.00° 114.76° 90.00°	Depositor
Resolution (Å)	19.57 – 1.77 19.57 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.57-1.77) 99.8 (19.57-1.77)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.87 (at 1.77Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.139 , 0.171 0.139 , 0.170	Depositor DCC
$R_{free}$ test set	2003 reflections (2.27%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.3	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 46.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.023 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8071	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.51% 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: ZN, MES

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	1.05	6/6966 (0.1%)	1.07	45/9491 (0.5%)

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	1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	200	GLY	C-N	8.97	1.51	1.34
1	A	611	GLU	CD-OE1	6.46	1.32	1.25
1	A	770	SER	CB-OG	5.94	1.50	1.42
1	A	483	ARG	CZ-NH2	5.72	1.40	1.33
1	A	587	GLU	CD-OE1	-5.29	1.19	1.25
1	A	284	GLU	CB-CG	-5.14	1.42	1.52

All (45) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	89	ARG	NE-CZ-NH1	11.29	125.94	120.30
1	A	116	ARG	NE-CZ-NH2	9.41	125.00	120.30
1	A	89	ARG	NE-CZ-NH2	-9.38	115.61	120.30
1	A	641	ARG	NE-CZ-NH1	-9.18	115.71	120.30
1	A	736	ARG	NE-CZ-NH1	8.17	124.38	120.30
1	A	584	ARG	NE-CZ-NH1	-8.08	116.26	120.30
1	A	289	ARG	NE-CZ-NH2	-7.59	116.51	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	533	ARG	NE-CZ-NH2	-7.39	116.61	120.30
1	A	641	ARG	CG-CD-NE	-7.17	96.74	111.80
1	A	258	PHE	CB-CA-C	7.11	124.61	110.40
1	A	681	ARG	NE-CZ-NH1	7.08	123.84	120.30
1	A	225	ARG	NE-CZ-NH1	-7.07	116.77	120.30
1	A	651	ARG	NE-CZ-NH1	-6.95	116.82	120.30
1	A	658	ASP	CB-CG-OD2	-6.91	112.08	118.30
1	A	78	ASP	CB-CG-OD1	6.85	124.47	118.30
1	A	778	ARG	NE-CZ-NH2	6.82	123.71	120.30
1	A	584	ARG	NE-CZ-NH2	6.74	123.67	120.30
1	A	483	ARG	NE-CZ-NH1	-6.70	116.95	120.30
1	A	724	ARG	NE-CZ-NH1	6.70	123.65	120.30
1	A	550	ARG	NE-CZ-NH1	6.64	123.62	120.30
1	A	78	ASP	CB-CG-OD2	-6.61	112.35	118.30
1	A	116	ARG	NE-CZ-NH1	-6.26	117.17	120.30
1	A	379	ASP	CB-CG-OD1	6.21	123.89	118.30
1	A	803	ARG	NE-CZ-NH1	6.16	123.38	120.30
1	A	681	ARG	NE-CZ-NH2	-6.03	117.28	120.30
1	A	850	ARG	NE-CZ-NH1	5.85	123.22	120.30
1	A	323	ASP	CB-CG-OD2	-5.70	113.17	118.30
1	A	640	ARG	NE-CZ-NH2	5.68	123.14	120.30
1	A	787	ARG	NE-CZ-NH1	5.67	123.14	120.30
1	A	471	ARG	NE-CZ-NH2	5.53	123.07	120.30
1	A	323	ASP	CB-CG-OD1	5.49	123.24	118.30
1	A	510	ARG	NE-CZ-NH2	-5.42	117.59	120.30
1	A	592	ASP	CB-CG-OD1	5.41	123.17	118.30
1	A	541	ASP	CB-CG-OD2	-5.39	113.44	118.30
1	A	420	ASP	CB-CG-OD1	5.35	123.11	118.30
1	A	803	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	A	420	ASP	CB-CG-OD2	-5.32	113.51	118.30
1	A	738	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	A	8	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	A	89	ARG	CD-NE-CZ	5.17	130.84	123.60
1	A	230	GLU	OE1-CD-OE2	5.11	129.43	123.30
1	A	845	ARG	NE-CZ-NH1	5.06	122.83	120.30
1	A	320	ARG	NE-CZ-NH2	5.05	122.83	120.30
1	A	736	ARG	NE-CZ-NH2	-5.03	117.79	120.30
1	A	421	ARG	NE-CZ-NH1	5.02	122.81	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	200	GLY	Mainchain

## 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	6806	0	6557	32	0
2	B	1	0	0	0	0
3	C	12	0	13	0	0
4	S	1252	0	0	11	0
All	All	8071	0	6570	32	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 2.

All (32) 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:717:GLU:HG2	4:S:1238:HOH:O	1.65	0.96
1:A:478[B]:LEU:C	1:A:478[B]:LEU:HD23	2.04	0.79
1:A:461[A]:ARG:HD3	4:S:1092:HOH:O	1.93	0.68
1:A:422:HIS:HD2	4:S:1357:HOH:O	1.77	0.67
1:A:68:ARG:NH1	4:S:1111:HOH:O	2.25	0.67
1:A:386:ASN:HD22	1:A:839:ARG:HH12	1.44	0.65
1:A:213:ASP:HB3	1:A:249:HIS:HE1	1.66	0.59
1:A:478[A]:LEU:HG	1:A:522:ARG:NE	2.22	0.54
1:A:478[A]:LEU:HD12	1:A:522:ARG:HD3	1.91	0.53
1:A:422:HIS:HE1	4:S:233:HOH:O	1.90	0.52
1:A:461[A]:ARG:HG2	4:S:1092:HOH:O	2.10	0.52
1:A:68:ARG:NH2	1:A:104:ASP:OD1	2.43	0.50
1:A:680[A]:MET:HG3	1:A:703:LEU:HD22	1.93	0.49
1:A:778:ARG:NH1	4:S:792:HOH:O	2.42	0.49
1:A:478[B]:LEU:HD23	1:A:479:GLN:N	2.28	0.47
1:A:807:THR:HG23	1:A:811:LEU:HD12	1.96	0.47
1:A:323:ASP:HB2	1:A:375:ALA:O	2.15	0.46
1:A:478[B]:LEU:C	1:A:478[B]:LEU:CD2	2.82	0.46
1:A:461[A]:ARG:CD	4:S:1092:HOH:O	2.59	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:521:GLU:OE1	1:A:522:ARG:HB2	2.16	0.46
1:A:89:ARG:HH11	1:A:89:ARG:HG3	1.81	0.45
1:A:461[A]:ARG:CG	4:S:1092:HOH:O	2.63	0.45
1:A:478[A]:LEU:HG	1:A:522:ARG:CZ	2.47	0.45
1:A:478[A]:LEU:HD12	1:A:478[A]:LEU:N	2.34	0.43
1:A:275:GLU:HG3	1:A:310:HIS:HB3	2.02	0.42
1:A:188:LYS:NZ	4:S:716:HOH:O	2.53	0.41
1:A:422:HIS:CE1	1:A:431:ASP:HB3	2.55	0.41
1:A:212:ASP:HB2	4:S:866:HOH:O	2.19	0.41
1:A:110:THR:O	1:A:803:ARG:HD3	2.20	0.41
1:A:478[A]:LEU:N	1:A:478[A]:LEU:CD1	2.84	0.41
1:A:788:TYR:CE2	1:A:792[A]:ILE:HD13	2.56	0.40
1:A:318:THR:O	1:A:426:ASN:HA	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	866/884 (98%)	848 (98%)	18 (2%)	0	<b>100</b> <b>100</b>

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	A	704/717 (98%)	698 (99%)	6 (1%)	78 72

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	ASP
1	A	89	ARG
1	A	260	LYS
1	A	291	ARG
1	A	319	MET
1	A	501	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	106	GLN
1	A	130	GLN
1	A	276	ASN
1	A	381	GLN
1	A	386	ASN
1	A	422	HIS
1	A	426	ASN
1	A	453	GLN
1	A	512	HIS
1	A	532	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	MES	C	901	-	12,12,12	0.75	0	14,16,16	0.48	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
3	MES	C	901	-	-	1/6/14/14	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	901	MES	C8-C7-N4-C5

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 [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	A	860/884 (97%)	-0.33	13 (1%) 73 73	9, 13, 26, 59	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	213	ASP	6.4
1	A	212	ASP	6.3
1	A	504	ALA	5.8
1	A	503	PRO	5.8
1	A	1	VAL	5.0
1	A	214	GLY	4.5
1	A	468	ASP	3.2
1	A	443	ASP	3.2
1	A	211	GLY	2.8
1	A	506	GLY	2.7
1	A	860	SER	2.2
1	A	505	THR	2.1
1	A	36	PRO	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( $\text{\AA}^2$ )	Q<0.9
3	MES	C	901	12/?	0.95	0.15	15,21,28,34	0
2	ZN	B	1	1/?	0.99	0.04	15,15,15,15	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.