



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2016 – 03:42 PM EDT

PDB ID : 5L5N  
Title : Plexin A4 full extracellular region, domains 1 to 7 modeled, data to 8.5 angstrom, spacegroup P4(3)22  
Authors : Janssen, B.J.C.; Kong, Y.; Malinauskas, T.; Vangoor, V.R.; Coles, C.H.; Kaufmann, R.; Ni, T.; Gilbert, R.J.C.; Padilla-Parra, S.; Pasterkamp, R.J.; Jones, E.Y.  
Deposited on : 2016-05-28  
Resolution : 8.50 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

<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

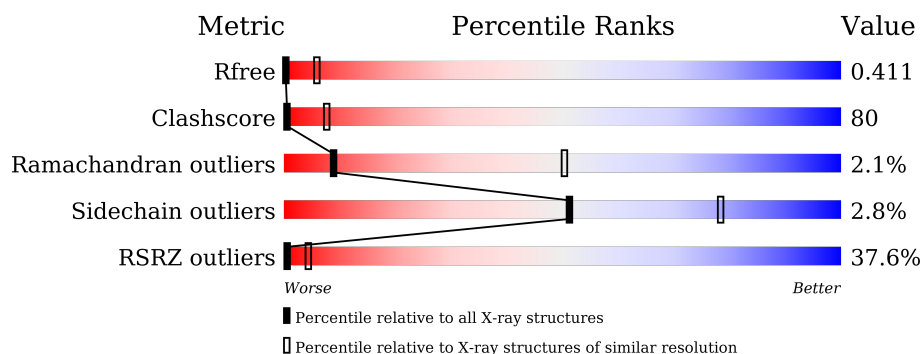
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 8.50 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	1015 (11.50-3.66)
Clashscore	102246	1064 (11.50-3.70)
Ramachandran outliers	100387	1036 (11.50-3.66)
Sidechain outliers	100360	1006 (11.50-3.66)
RSRZ outliers	91569	1014 (11.50-3.66)

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  $\geq 3$ , 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	1207	<div> <div>29%</div> <div>24%</div> <div>48%</div> <div>•</div> <div>24%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7189 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 Plexin-A4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	915	Total	C	N	O	S	0	0	0
			7189	4533	1239	1357	60			

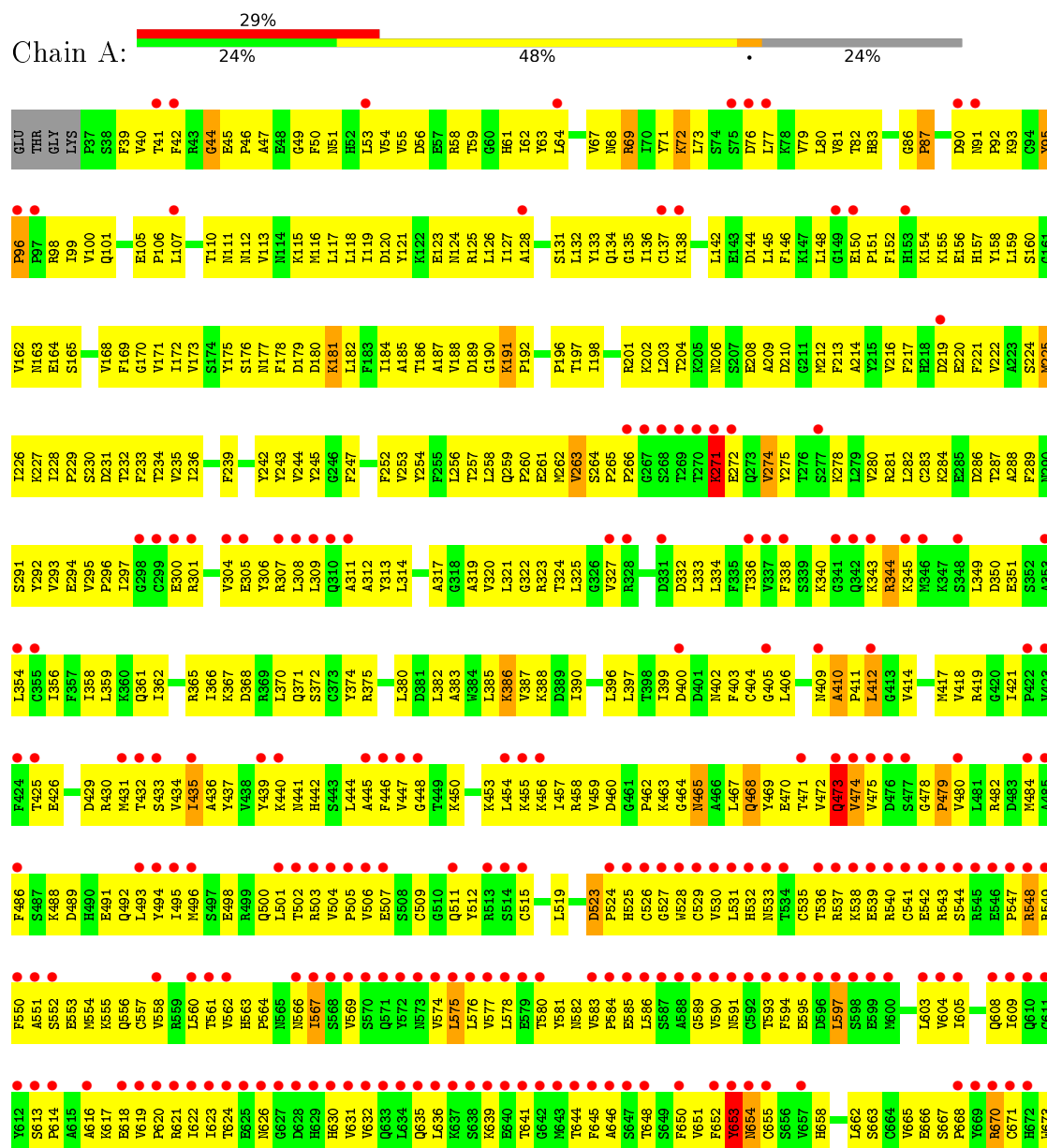
There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	GLU	-	expression tag	UNP Q80UG2
A	34	THR	-	expression tag	UNP Q80UG2
A	35	GLY	-	expression tag	UNP Q80UG2
A	1230	GLY	-	expression tag	UNP Q80UG2
A	1231	ARG	-	expression tag	UNP Q80UG2
A	1232	THR	-	expression tag	UNP Q80UG2
A	1233	LYS	-	expression tag	UNP Q80UG2
A	1234	HIS	-	expression tag	UNP Q80UG2
A	1235	HIS	-	expression tag	UNP Q80UG2
A	1236	HIS	-	expression tag	UNP Q80UG2
A	1237	HIS	-	expression tag	UNP Q80UG2
A	1238	HIS	-	expression tag	UNP Q80UG2
A	1239	HIS	-	expression tag	UNP Q80UG2

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: Plexin-A4



LVS	GLY	ARG	GLY	CYS	V933	T873	L812	L740	C874
HIS	GLY	PRO	ASN	LEU	A934	K874	C813	L740	C874
HIS	ASN	GLU	THR	PHE	V935	T875	L814	N741	K675
HIS	VAL	GLU	PRO	HIS	C936	T876	K815	Q743	Y676
HIS	LVS	PHE	ILE	ARG	R937	I877	A816	G744	R677
HIS	LEU	GLY	ALA	ARG	P938	R878	D817	I745	H678
ASN	ASN	PHE	VAL	SER	E939	G879	P818	E746	V679
THR	TYR	ILE	TRP	PRO	F940	E880	D819	Q747	C880
THR	LEU	GLY	LEU	SER	N941	N881	F820	R748	T681
VAL	VAL	ASP	THR	TYR	A942	L882	E821		H682
LEU	LEU	ASP	HIS	ASN	R943	G883	C822		D683
VAL	VAL	VAL	ILE	ILE	S944	L884	G823	R753	R684
GLY	GLY	GLN	ASP	CYS	E945	E885	H824	V759	N685
GLU	GLU	SER	LEU	ASN	Q946	F886	C825	Q760	T686
LVS	LVS	LEU	ILE	THR	L947	R887	Q826		F689
PRO	PRO	LEU	GLN	THR	Y948	D888	S827	T763	E690
CYS	CYS	ILE	ASN	SER	Y949	I889	P828	T764	E691
THR	THR	LEU	PRO	SER	F950	A890	S829	S765	
VAL	VAL	ASN	GLN	GLU	THR	E891	Q830	V766	V694
THR	THR	LVS	ILE	GLY	LEU	H932	C831	S767	K695
VAL	VAL	ASN	ARG	VAL	THR	K934	T832	V768	L696
SER	SER	THR	ALA	LEU	THR	R894	L833	E769	P697
ASP	ASP	PHE	LVS	ASP	LEU	V895			E698
VAL	VAL	THR	HIS	MET	ALA	A896	H836	I773	D699
THR	THR	TYR	GLY	LVS	ASP	G897	C837	T774	C700
GLN	GLN	PRO	GLY	VAL	LEU	E898	P838	N775	P701
LEU	LEU	TYR	PRO	THR	LVS	E999	A839	L776	Q702
LEU	LEU	PRO	LVS	THR	ASN	C900	H840	P777	L703
CYS	CYS	ASN	VAL	VAL	PRO	S901	E841	V778	L704
GLU	GLU	PRO	GLN	GLN	ASN	P902	S842	E779	R705
SER	SER	VAL	ILE	VAL	ARG	L903	R843	L780	V706
PRO	PRO	PHE	ASN	ASP	GLY	V904	H844	T781	D707
ASN	ASN	GLU	ILE	ARG	ALA	P904	L845	V782	K708
LEU	LEU	ALA	CYS	ALA	MET	D905	S846	V783	
ILE	ILE	PHE	GLU	ARG	SER	G906	E846	V784	I709
GLY	GLY	SER	VAL	ILE	GLY	Y907	L847	L710	
ARG	ARG	PRO	LEU	ARG	GLY	I908	S848	N785	V711
HIS	HIS	SER	ASN	GLN	THR	P909	G849	G786	P712
LVS	LVS	GLY	ALA	ASP	GLN	A910	A850	H787	V713
VAL	VAL	ILE	THR	LEU	VAL	E911	N851	F788	E714
MET	MET	LEU	GLU	THR	THR	Q912	S852	N789	V715
ALA	ALA	GLU	MET	PHE	ILE	I913	K853	I790	I716
ARG	ARG	LEU	THR	GLN	THR	V914	C854	D791	K717
VAL	VAL	LVS	CYS	TYR	THR	E915	T855	N792	
GLY	GLY	PRO	GLU	VAL	GLY	E916	N856		P718
GLY	GLY	GLY	ALA	GLU	THR	M917	P857		I719
MET	MET	THR	ASP	ASP	LEU	G918	R858	N796	T720
TYR	TYR	ALA	PRO	THR	ASN	E919	T859	K797	L721
SER	SER	ILE	LEU	THR	ALA	A920	T860	V798	A723
PRO	PRO	ILE	ILE	ILE	GLY	K921	E861	L800	K724
PRO	PRO	LEU	VAL	VAL	SER	E922	T862	N725	
GLY	GLY	LVS	ARG	ARG	ASN	S923	I863	R802	L726
MET	MET	GLY	GLY	ILE	ASN	Q924	P864	C803	P727
VAL	VAL	LVS	ASP	GLU	VAL	H925	V865	G804	Q728
THR	THR	ASN	ASP	PRO	VAL	A926	T866	A805	P729
ILE	ILE	LEU	GLN	GLU	VAL	G927	G867	P806	
ALA	ALA	ILE	SER	THR	PHE	F928	P868	R807	P730
PRO	PRO	PRO	ASP	SER	GLY	V929	R869	E808	S731
GLY	GLY	PRO	LEU	ILE	ILE	E930	E870	S809	
ARG	ARG	VAL	GLN	VAL	GLN	I931	C810	C510	Y736
THR	THR	ALA	GLU	SER	THR	C932	T872	C511	T739

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	189.48Å 189.48Å 252.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.37 – 8.50 70.37 – 8.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (70.37-8.50) 99.6 (70.37-8.50)	Depositor EDS
$R_{merge}$	0.33	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 8.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.401 , 0.412 0.404 , 0.411	Depositor DCC
$R_{free}$ test set	201 reflections (4.60%)	DCC
Wilson B-factor (Å <sup>2</sup> )	502.9	Xtriage
Anisotropy	0.294	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 155.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.34$ , $\langle L^2 \rangle = 0.17$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.54	EDS
Total number of atoms	7189	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	162.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.07% 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

### 5.1 Standard geometry

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	1.10	4/7343 (0.1%)	1.26	20/9940 (0.2%)

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	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	653	TYR	C-N	-36.36	0.50	1.34
1	A	700	CYS	C-N	-33.68	0.70	1.34
1	A	49	GLY	CA-C	6.43	1.62	1.51
1	A	49	GLY	C-N	5.09	1.45	1.34

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	747	GLN	CG-CD-OE1	-38.83	43.94	121.60
1	A	653	TYR	O-C-N	-20.66	89.64	122.70
1	A	700	CYS	C-N-CD	-19.18	78.41	120.60
1	A	700	CYS	O-C-N	-10.23	101.67	121.10
1	A	747	GLN	CG-CD-NE2	-9.56	93.74	116.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	653	TYR	Mainchain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	700	CYS	Mainchain
1	A	863	ILE	Peptide
1	A	95	TYR	Peptide

## 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	A	7189	0	7049	1133	26
All	All	7189	0	7049	1133	26

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 80.

The worst 5 of 1133 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:468:GLN:HG3	1:A:524:PRO:CD	1.18	1.58
1:A:530:VAL:HG11	1:A:584:PRO:CD	1.11	1.56
1:A:439:TYR:CE2	1:A:538:LYS:NZ	1.78	1.51
1:A:468:GLN:CG	1:A:524:PRO:HD3	1.39	1.51
1:A:439:TYR:HE2	1:A:538:LYS:NZ	1.11	1.44

The worst 5 of 26 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:234:THR:CA	1:A:234:THR:CA[5_455]	0.86	1.34
1:A:233:PHE:O	1:A:234:THR:OG1[5_455]	0.87	1.33
1:A:146:PHE:CE1	1:A:730:GLN:OE1[4_555]	1.03	1.17
1:A:234:THR:O	1:A:234:THR:O[5_455]	1.40	0.80
1:A:83:HIS:CE1	1:A:731:SER:OG[4_555]	1.57	0.63



## 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	905/1207 (75%)	842 (93%)	44 (5%)	19 (2%)	9	50

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	PRO
1	A	181	LYS
1	A	191	LYS
1	A	410	ALA
1	A	465	ASN

### 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	812/1067 (76%)	789 (97%)	23 (3%)	51	78

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	548	ARG
1	A	575	LEU
1	A	854	CYS
1	A	567	ILE
1	A	597	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	A	672	HIS
1	A	836	HIS
1	A	629	HIS
1	A	163	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	6

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	802:LYS	C	803:CYS	N	4.06

*Continued on next page...*

*Continued from previous page...*

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	854:CYS	C	855:THR	N	3.30
1	A	557:CYS	C	558:VAL	N	2.81
1	A	506:VAL	C	507:GLU	N	2.58
1	A	700:CYS	C	701:PRO	N	0.70

## 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	915/1207 (75%)	1.95	344 (37%) <b>0</b> <b>4</b>	100, 150, 216, 216	0

The worst 5 of 344 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	584	PRO	11.0
1	A	859	ILE	9.8
1	A	585	GLU	9.7
1	A	902	PRO	9.4
1	A	860	THR	9.3

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.