wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 04:58 AM EDT

PDB ID	:	3S5K
Title	:	Crystal structures of falcilysin, a M16 metalloprotease from the malaria parasite Plasmodium falciparum
Authors	:	Morgunova, E.; Ponpuak, M.; Istvan, E.; Popov, A.; Goldberg, D.; Eneqvist, T.
Deposited on	:	2011-05-23
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*
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:

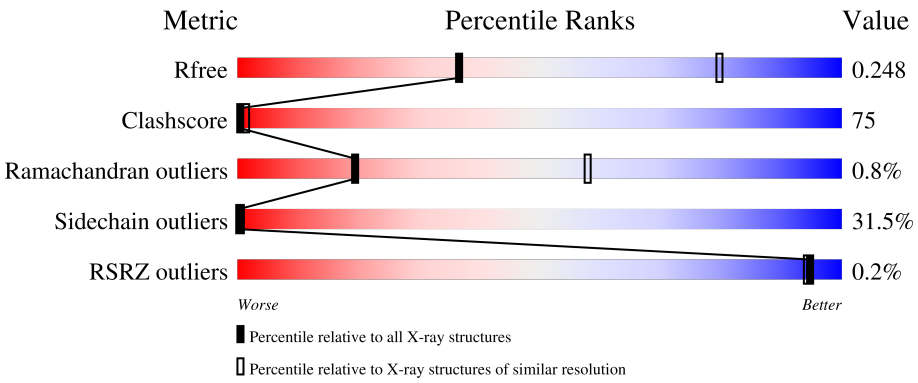
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 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)

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	A	1193	20% 44% 23% • 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8741 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 Falcilysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1053	8674	5577	1418	1653	26	0	0	0

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

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

- Molecule 3 is water.

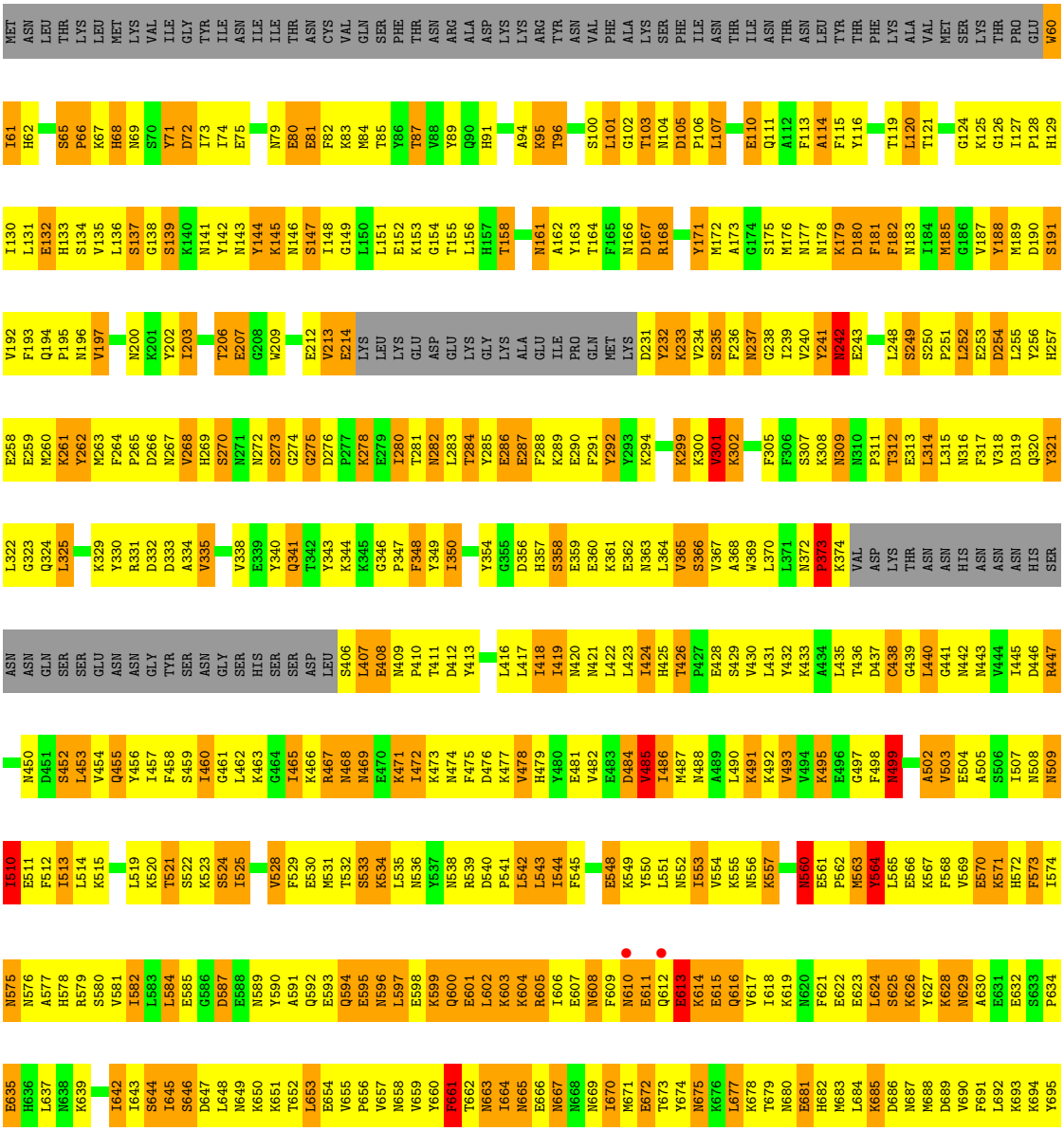
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	66	66	66	0	0

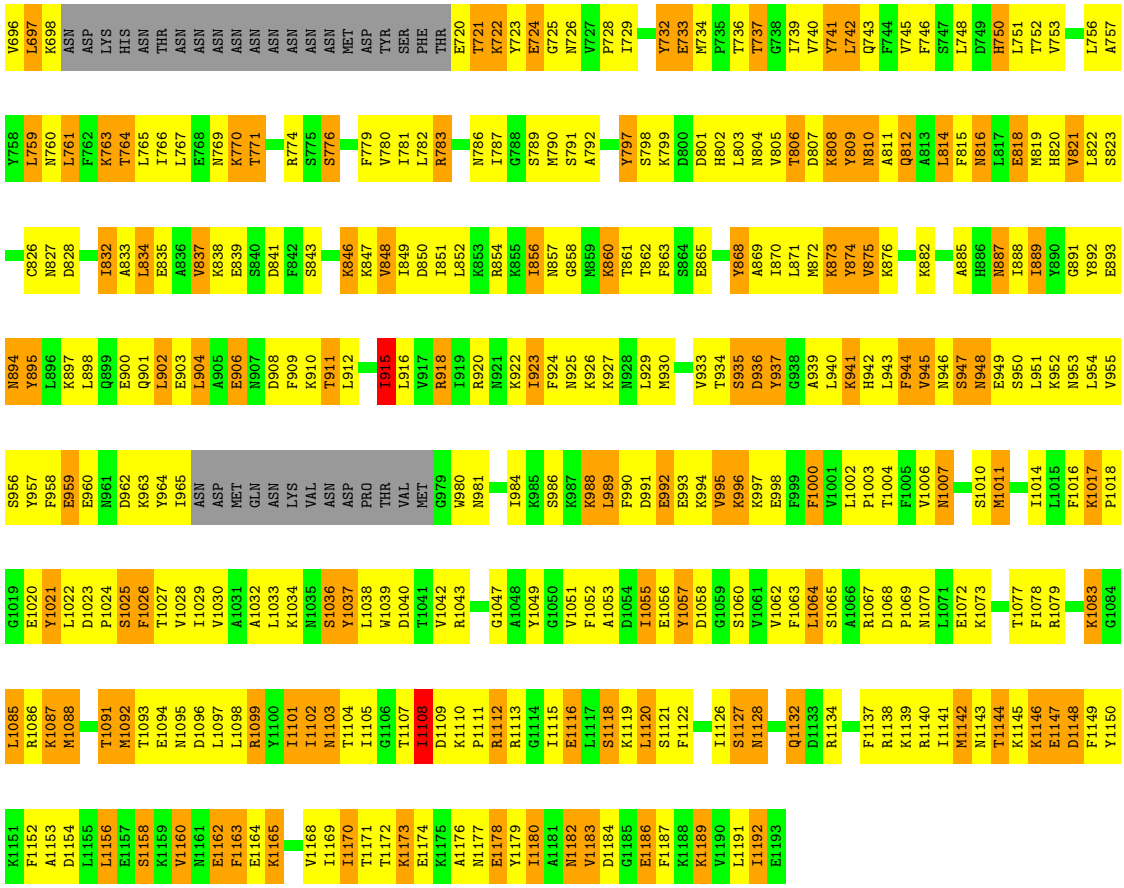
3 Residue-property plots (i)

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: Falcilysin

Chain A: 20% 44% 23% • 12%




V696 L697 K698 ASN ASP LYS HIS ASN THR ASN ASN ASN ASN ASN ASN ASN ASN ASN MET ASP TYR SER PHE THR E720 T721 K722 Y723 E724 G725 N726 V727 P728 I729 Y732 E733 M734 P735 T736 T737 G738 I739 V740 Y741 L742 Q743 F744 V745 F746 S747 L748 D749 H750 L751 T752 V753 L756 A757
Y758 L759 N760 L761 F762 K763 T764 L765 I766 L767 E768 N769 K770 T771 R774 S775 S776 F779 V780 I781 L782 R783 N786 I787 G788 S789 M790 S791 A792 Y797 S798 K799 D800 D801 H802 L803 N804 V805 T806 D807 K808 Y809 N810 A811 Q812 A813 L814 F815 N816 L817 E818 M819 H820 V821 L822 S823
C826 N827 D828 I832 A833 L834 E835 A836 V837 K838 E839 S840 D841 F842 S843 K846 K847 V848 I849 D850 I851 L852 K853 R854 K855 I856 N857 G858 M859 K860 T861 T862 F863 S864 E865 Y868 A869 I870 L871 M872 K873 Y874 V875 K876 K882 A885 H886 N887 I888 I889 Y890 G891 Y892 E893
N894 Y895 L896 K897 L898 Q899 E900 Q901 L902 E903 L904 A905 E906 N907 D908 F909 K910 T911 L912 I915 L916 V917 R918 I919 R920 N921 K922 I923 F924 N925 K926 K927 N928 L929 M930 V933 T934 S935 D936 Y937 G938 A939 L940 K941 H942 L943 F944 V945 N946 S947 N948 E949 S950 L951 K952 N953 L954 V955
S956 Y957 F958 E959 E960 N961 D962 K963 Y964 I965 ASN ASP MET GLN ASN LYS VAL ASN ASP PRO THR VAL MET G979 W980 N981 I984 K985 S986 K987 K988 L989 F990 D991 E992 E993 K994 V995 K996 K997 E998 F999 F1000 V1001 L1002 P1003 T1004 F1005 V1006 N1007 S1010 M1011 I1014 L1015 F1016 K1017 P1018
G1019 E1020 Y1021 L1022 D1023 P1024 S1025 F1026 T1027 V1028 I1029 V1030 A1031 A1032 L1033 K1034 N1035 S1036 Y1037 L1038 W1039 D1040 T1041 V1042 R1043 G1047 A1048 Y1049 G1050 V1051 F1052 A1053 D1054 I1055 E1056 Y1057 D1058 G1059 S1060 V1061 V1062 F1063 L1064 S1065 A1066 R1067 D1068 P1069 N1070 L1071 E1072 K1073 T1077 F1078 R1079 K1083 G1084
L1085 R1086 K1087 M1088 T1091 M1092 T1093 E1094 N1095 D1096 L1097 L1098 R1099 Y1100 I1101 I1102 N1103 T1104 I1105 G1106 T1107 I1108 D1109 K1110 P1111 R1112 R1113 G1114 I1115 E1116 L1117 S1118 K1119 L1120 S1121 F1122 I1126 S1127 N1128 Q1132 D1133 R1134 F1137 R1138 K1139 R1140 I1141 M1142 N1143 T1144 K1145 K1146 E1147 D1148 F1149 Y1150
K1151 F1152 A1153 D1154 L1155 L1156 E1157 S1158 K1159 V1160 N1161 E1162 F1163 E1164 K1165 V1168 I1169 I1170 T1171 T1172 K1173 E1174 K1175 A1176 N1177 E1178 Y1179 I1180 A1181 N1182 V1183 D1184 G1185 E1186 F1187 K1188 K1189 V1190 L1191 I1192 E1193

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.78Å 105.29Å 114.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.83 – 3.20 42.83 – 2.84	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.83-3.20) 95.4 (42.83-2.84)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 2.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6_289)	Depositor
R, R_{free}	0.210 , 0.316 0.224 , 0.248	Depositor DCC
R_{free} test set	1075 reflections (4.11%)	wwPDB-VP
Wilson B-factor (Å^2)	50.1	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 78.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	8741	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	1.79	58/8849 (0.7%)	1.11	36/11918 (0.3%)

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

The worst 5 of 58 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	438	CYS	CB-SG	-10.15	1.65	1.82
1	A	826	CYS	CB-SG	-7.60	1.69	1.82
1	A	818	GLU	CG-CD	-7.31	1.41	1.51
1	A	188	TYR	CE1-CZ	-6.19	1.30	1.38
1	A	821	VAL	CA-CB	-6.18	1.41	1.54

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	915	ILE	CB-CA-C	-7.40	96.80	111.60
1	A	365	VAL	CB-CA-C	-6.51	99.02	111.40
1	A	945	VAL	CB-CA-C	-6.49	99.07	111.40
1	A	837	VAL	CB-CA-C	-6.44	99.16	111.40
1	A	856	ILE	CB-CA-C	-6.38	98.83	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	560	ASN	Peptide

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	8674	0	8613	1304	2
2	A	1	0	0	0	0
3	A	66	0	0	6	0
All	All	8741	0	8613	1304	2

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 75.

The worst 5 of 1304 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:263:MET:CE	1:A:457:ILE:HD12	1.14	1.54
1:A:610:ASN:C	1:A:613:GLU:CG	1.75	1.50
1:A:232:TYR:HB2	1:A:614:LYS:NZ	1.30	1.45
1:A:151:LEU:O	1:A:155:THR:CG2	1.67	1.42
1:A:612:GLN:CB	1:A:616:GLN:HB3	1.49	1.42

All (2) 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:608:ASN:ND2	1:A:988:LYS:CG[2_554]	1.98	0.22
1:A:538:ASN:O	1:A:679:THR:OG1[2_654]	2.10	0.10

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	1043/1193 (87%)	998 (96%)	37 (4%)	8 (1%)	19 58

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	VAL
1	A	610	ASN
1	A	613	GLU
1	A	1128	ASN
1	A	373	PRO

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	972/1104 (88%)	666 (68%)	306 (32%)	0 0

5 of 306 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	902	LEU
1	A	1143	ASN
1	A	920	ARG
1	A	1007	ASN
1	A	1182	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	928	ASN
1	A	946	ASN
1	A	1177	ASN
1	A	443	ASN
1	A	425	HIS

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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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 (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, 95th 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	A	1053/1193 (88%)	-0.51	2 (0%)	95	94	9, 32, 70, 128	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	610	ASN	3.9
1	A	612	GLN	2.7

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, 95th 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	ZN	A	1194	1/1	0.35	0.20	38,38,38,38	0

6.5 Other polymers (i)

There are no such residues in this entry.