wwPDB X-ray Structure Validation Summary Report (i)

May 28, 2020 – 08:41 pm BST

PDB ID	:	1U6G
Title	:	Crystal Structure of The Cand1-Cul1-Roc1 Complex
Authors	:	Goldenberg, S.J.; Shumway, S.D.; Cascio, T.C.; Garbutt, K.C.; Liu, J.; Xiong, Y.; Zheng, N.
Deposited on	:	2004-07-29
Resolution	:	3.10 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 following versions of software and data (see references (1)) were used in the production of this report:

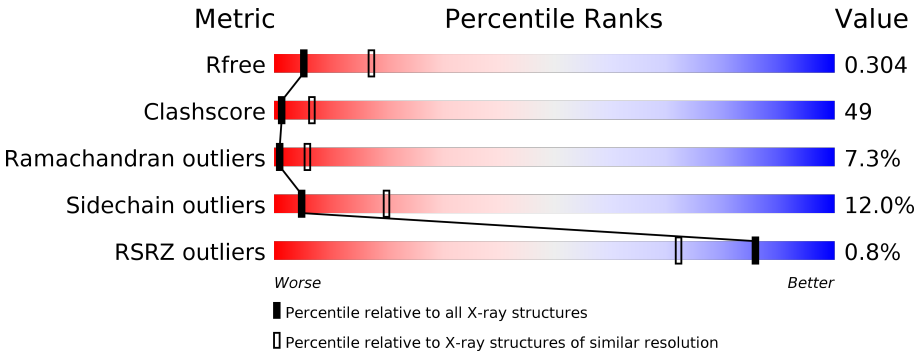
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 >=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	776	% 41% 40% 10% • 8%
2	B	108	2% 31% 38% 11% • 19%
3	C	1230	30% 50% 11% • 7%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15511 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 Cullin homolog 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	715	Total	C	N	O	S	0	0	0
			5855	3719	998	1109	29			

- Molecule 2 is a protein called RING-box protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	88	Total	C	N	O	S	0	0	0
			731	464	133	125	9			

- Molecule 3 is a protein called TIP120 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	1146	Total	C	N	O	S	0	0	0
			8904	5667	1509	1672	56			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	3	Total	Zn	0	0
			3	3		

- Molecule 5 is water.

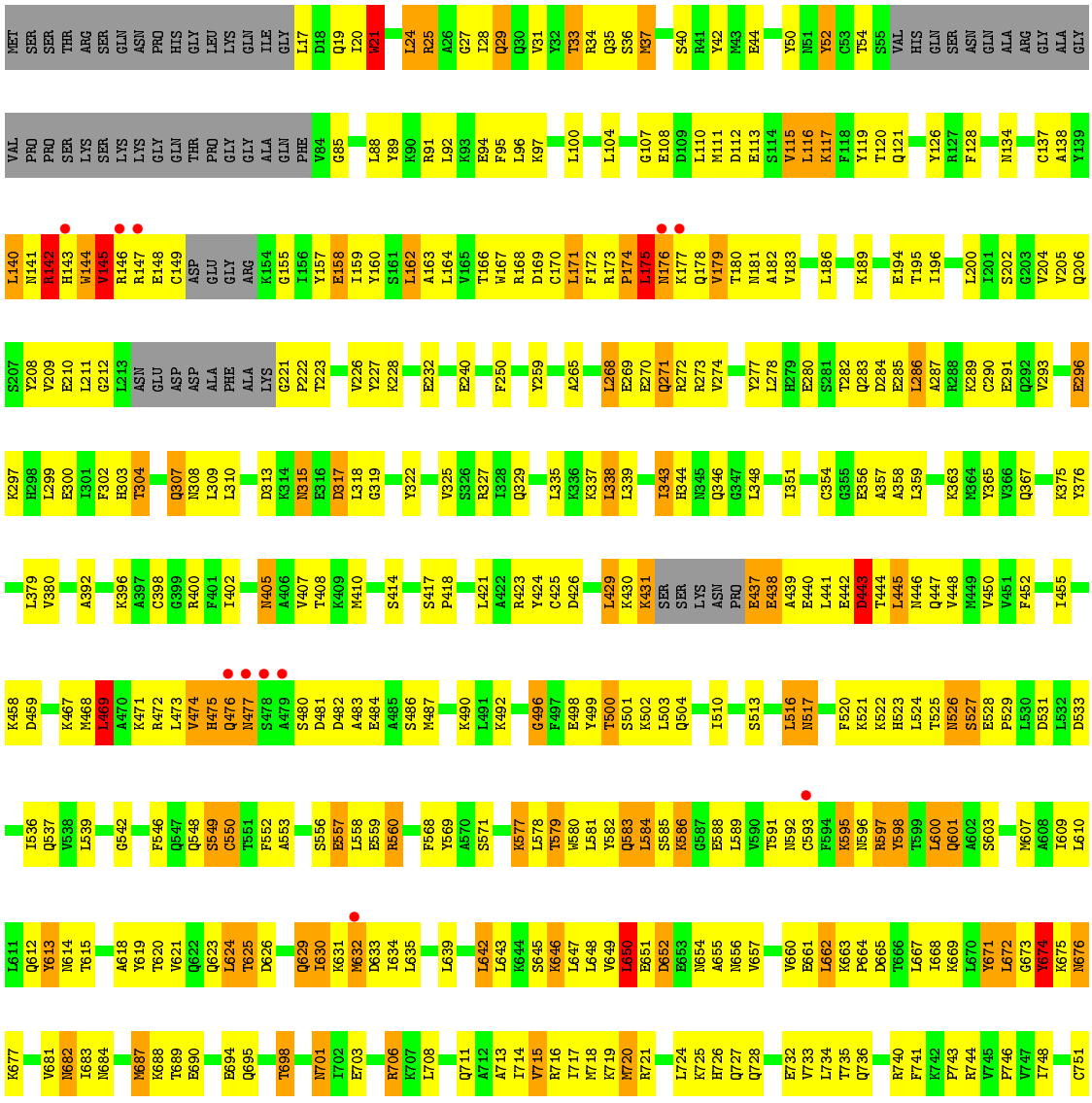
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	10	Total	O	0	0
			10	10		
5	B	3	Total	O	0	0
			3	3		
5	C	5	Total	O	0	0
			5	5		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

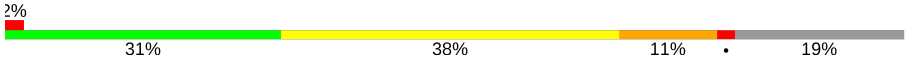
- Molecule 1: Cullin homolog 1

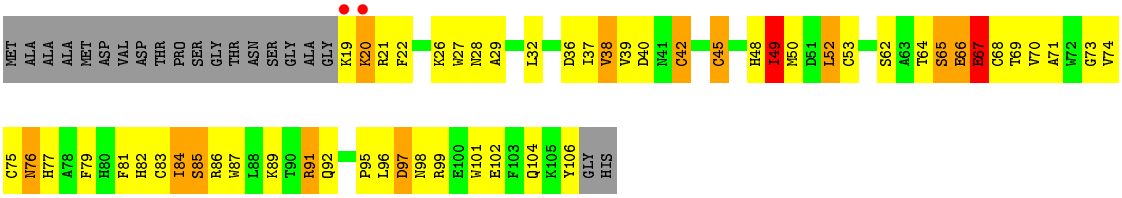
Chain A:



- Molecule 2: RING-box protein 1

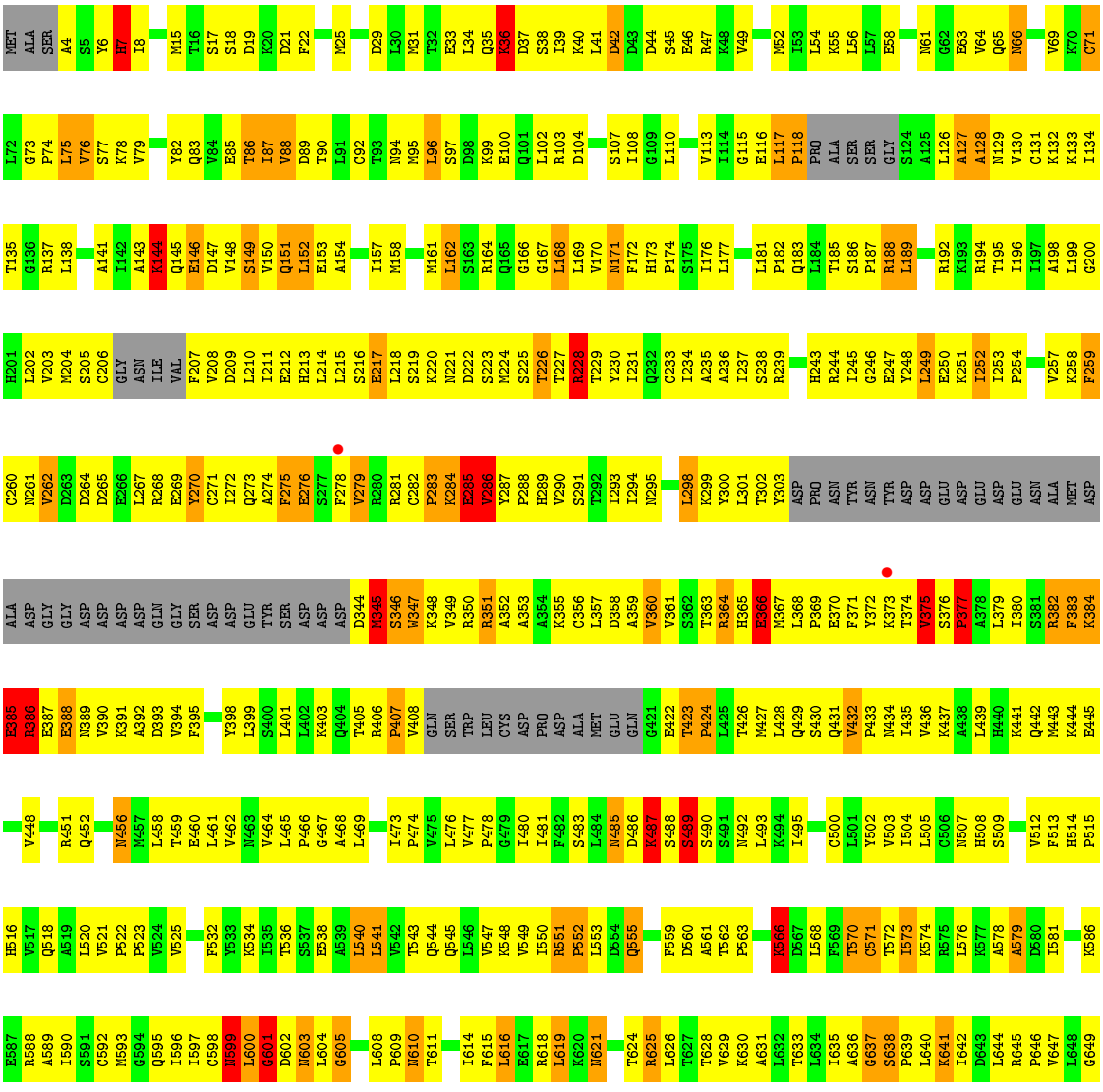
Chain B:

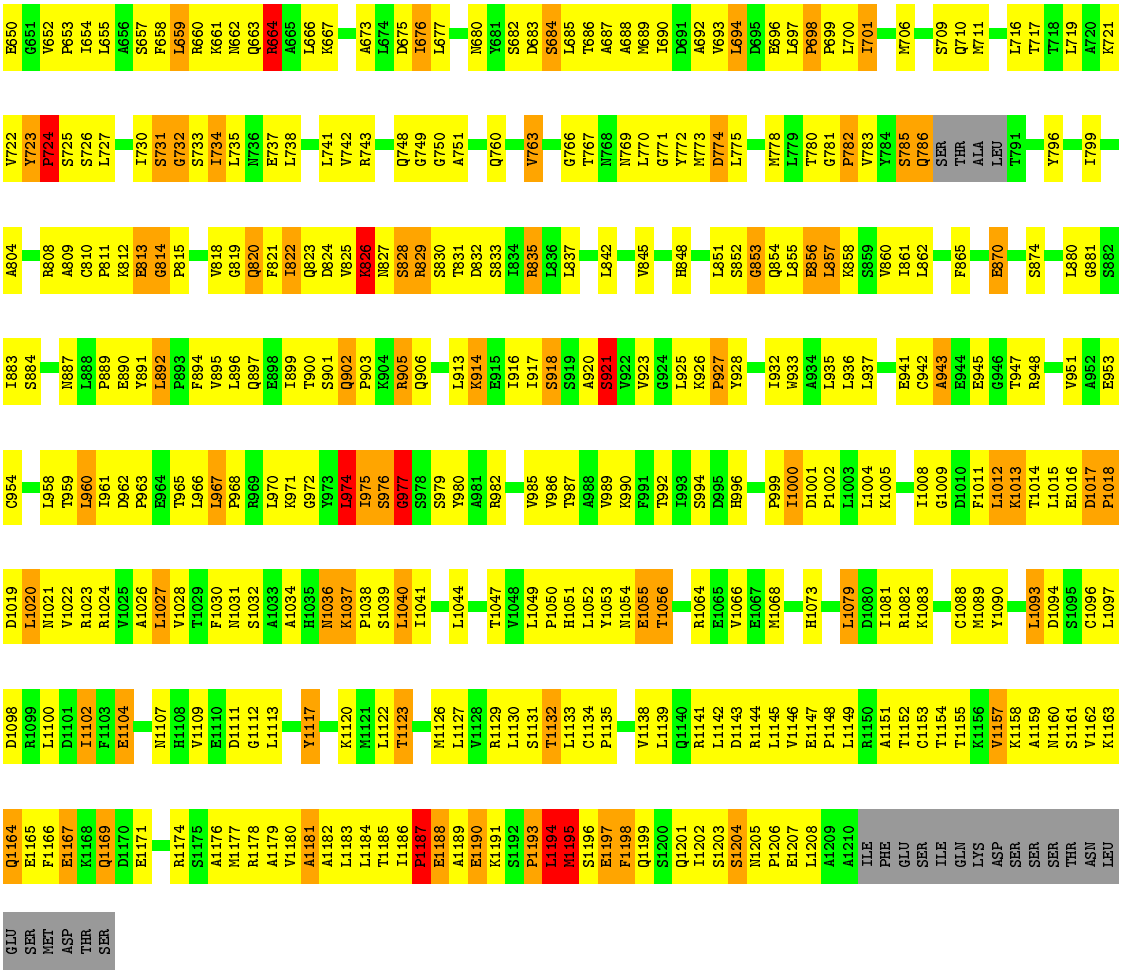




- Molecule 3: TIP120 protein

Chain C:




E650 G651 V652 P653 I654 L655 A656 S657 F658 L659 R660 K661 N662 Q663 R664 A665 L666 K667 A673 L674 D675 I676 L677 N680 Y681 S682 D683 S684 L685 T686 A687 A688 M689 I690 D691 A692 V693 L694 D695 E696 L697 P698 P699 L700 I701 M706 S709 Q710 M711 L716 T717 T718 L719 A720 K721
V722 Y723 P724 S725 S726 L727 I730 S731 G732 S733 I734 L735 N736 E737 L738 L741 V742 R743 Q748 G749 G750 A751 Q760 V763 G766 T767 N768 N769 L770 G771 Y772 M773 D774 L775 M778 L779 T780 G781 P782 V783 Y784 S785 Q786 SER THR ALA LEU T791 Y796 I799
A804 R808 A809 C810 P811 K812 E813 G814 P815 V818 G819 Q820 F821 I822 Q823 D824 V825 K826 N827 S828 R829 S830 T831 D832 S833 I834 R835 L836 L837 L842 V845 H848 L851 S852 G853 Q854 L855 E856 L857 K858 S859 V860 I861 L862 F865 E870 S874 L880 G881 S882
I883 S884 N887 L888 P889 E890 Y891 L892 P893 F894 V895 L896 Q897 E898 I899 T900 S901 Q902 P903 K904 R905 Q906 L913 K914 E915 I916 I917 S918 S919 A920 S921 V922 V923 G924 L925 K926 P927 Y928 I932 W933 A934 L935 L936 L937 E941 C942 A943 E944 E945 G946 T947 R948 V951 A952 E953
C954 L958 T959 L960 I961 D962 P963 E964 T965 I966 L967 P968 R969 L970 K971 G972 Y973 L974 I975 S976 G977 S978 S979 Y980 A981 R982 V985 V986 T987 A988 V989 R990 F991 T992 I993 S994 D995 H996 P999 I1000 D1001 P1002 L1003 L1004 K1005 I1008 G1009 D1010 F1011 L1012 K1013 T1014 L1015 E1016 D1017 P1018
D1019 L1020 N1021 V1022 R1023 R1024 V1025 A1026 L1027 V1028 T1029 F1030 N1031 S1032 A1033 A1034 H1035 N1036 K1037 P1038 S1039 L1040 I1041 L1044 T1047 V1048 L1049 P1050 H1051 L1052 Y1053 N1054 E1055 T1056 R1064 E1065 V1066 E1067 M1068 H1073 L1079 D1080 I1081 R1082 K1083 C1088 M1089 Y1090 L1093 D1094 S1095 C1096 L1097
D1098 R1099 L1100 D1101 I1102 F1103 E1104 N1107 H1108 V1109 E1110 D1111 G1112 L1113 Y1117 K1120 M1121 L1122 T1123 M1126 L1127 V1128 R1129 L1130 S1131 T1132 L1133 C1134 P1135 V1138 L1139 Q1140 R1141 L1142 D1143 R1144 L1145 V1146 E1147 P1148 L1149 R1150 A1151 T1152 C1153 T1154 T1155 K1156 V1157 K1158 A1159 N1160 S1161 V1162 K1163
Q1164 E1165 F1166 E1167 K1168 Q1169 D1170 E1171 R1174 S1175 A1176 M1177 R1178 A1179 V1180 A1181 A1182 L1183 L1184 T1185 I1186 P1187 E1188 A1189 E1190 K1191 S1192 P1193 L1194 M1195 S1196 E1197 F1198 Q1199 S1200 Q1201 I1202 S1203 S1204 N1205 P1206 E1207 L1208 A1209 A1210 ILE PHE GLU SER ILE GLN LYS ASP SER SER SER THR ASN LEU
GLU SER MET ASP THR SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	108.47Å 151.33Å 215.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 49.68 – 2.49	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.10) 73.1 (49.68-2.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.84 (at 2.48Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.243 , 0.317 0.236 , 0.304	Depositor DCC
R_{free} test set	3108 reflections (2.95%)	wwPDB-VP
Wilson B-factor (Å^2)	56.5	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 53.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	15511	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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: 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	0.49	3/5949 (0.1%)	1.14	14/8007 (0.2%)
2	B	0.51	0/752	0.86	1/1020 (0.1%)
3	C	0.45	1/9041 (0.0%)	0.86	31/12243 (0.3%)
All	All	0.47	4/15742 (0.0%)	0.97	46/21270 (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	1
3	C	0	2
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	437	GLU	C-N	-8.90	1.13	1.34
1	A	443	ASP	C-N	7.93	1.52	1.34
3	C	601	GLY	C-N	-6.35	1.19	1.34
1	A	630	ILE	CG1-CD1	5.03	1.85	1.50

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	437	GLU	O-C-N	-72.37	6.91	122.70
3	C	117	LEU	C-N-CD	-20.01	76.58	120.60
3	C	117	LEU	C-N-CA	13.74	179.69	122.00
3	C	487	LYS	CB-CA-C	12.45	135.31	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	377	PRO	CA-N-CD	-10.61	96.64	111.50

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	437	GLU	Mainchain
3	C	599	ASN	Mainchain
3	C	601	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	5855	0	5900	502	0
2	B	731	0	689	76	0
3	C	8904	0	9248	998	0
4	B	3	0	0	0	0
5	A	10	0	0	0	0
5	B	3	0	0	0	0
5	C	5	0	0	0	0
All	All	15511	0	15837	1525	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 49.

The worst 5 of 1525 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)
3:C:373:LYS:CE	3:C:427:MET:HE1	1.31	1.56
3:C:373:LYS:HE3	3:C:427:MET:CE	1.36	1.54
1:A:630:ILE:CD1	1:A:630:ILE:CG1	1.85	1.51
3:C:373:LYS:CE	3:C:427:MET:CE	1.84	1.48
3:C:373:LYS:CD	3:C:427:MET:HE1	1.53	1.37

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	705/776 (91%)	584 (83%)	91 (13%)	30 (4%)	2 16
2	B	86/108 (80%)	68 (79%)	13 (15%)	5 (6%)	1 10
3	C	1134/1230 (92%)	793 (70%)	236 (21%)	105 (9%)	0 3
All	All	1925/2114 (91%)	1445 (75%)	340 (18%)	140 (7%)	1 6

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	144	TRP
1	A	145	VAL
1	A	646	LYS
1	A	652	ASP
1	A	674	TYR

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	650/698 (93%)	575 (88%)	75 (12%)	5 22
2	B	78/90 (87%)	65 (83%)	13 (17%)	2 9
3	C	1022/1098 (93%)	900 (88%)	122 (12%)	5 20
All	All	1750/1886 (93%)	1540 (88%)	210 (12%)	5 20

5 of 210 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	96	LEU
3	C	298	LEU
3	C	1102	ILE
3	C	144	LYS
3	C	228	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	676	ASN
3	C	66	ASN
3	C	1036	ASN
1	A	682	ASN
1	A	736	GLN

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)

Of 3 ligands modelled in this entry, 3 are 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	601:GLY	C	602:ASP	N	1.19
1	A	437:GLU	C	438:GLU	N	1.13

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	A	715/776 (92%)	-0.57	11 (1%) 73 54	13, 46, 129, 197	0
2	B	88/108 (81%)	-0.62	2 (2%) 60 39	3, 38, 76, 151	0
3	C	1146/1230 (93%)	-0.68	2 (0%) 95 90	14, 61, 122, 190	0
All	All	1949/2114 (92%)	-0.64	15 (0%) 86 72	3, 55, 124, 197	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	478	SER	16.0
1	A	477	ASN	10.7
1	A	479	ALA	4.6
1	A	176	ASN	4.1
2	B	20	LYS	3.7

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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
4	ZN	B	1230	1/1	0.93	0.14	56,56,56,56	0
4	ZN	B	1229	1/1	0.95	0.12	42,42,42,42	0
4	ZN	B	1231	1/1	0.97	0.10	36,36,36,36	0

6.5 Other polymers (i)

There are no such residues in this entry.