wwPDB X-ray Structure Validation Summary Report (i)

Sep 5, 2023 – 03:30 AM EDT

PDB ID	:	3V0C
Title	:	4.3 angstrom crystal structure of an inactive BoNT/A (E224Q/R363A/Y366F)
Authors	:	Gu, S.; Rumpel, S.; Zhou, J.; Strotmeier, J.; Bigalke, H.; Perry, K.; Shoemaker, C.B.; Rummel, A.; Jin, R.
Deposited on	:	2011-12-07
Resolution	:	4.30 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:

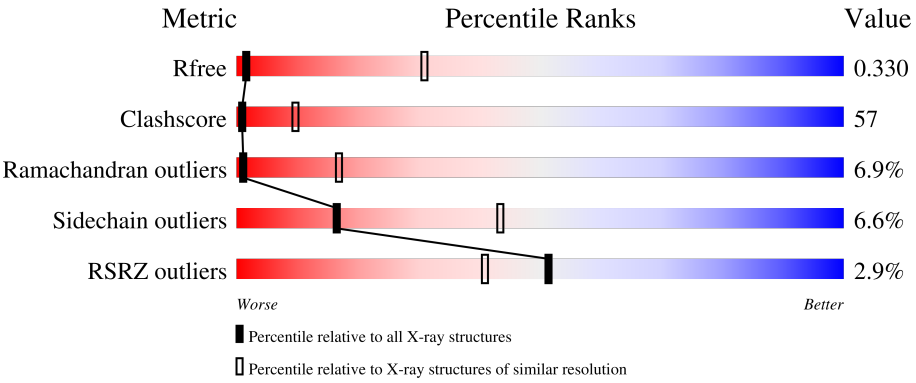
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 4.30 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	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)
RSRZ outliers	127900	1075 (4.90-3.70)

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	1312	3% 28% 59% 9% .

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10390 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 BoNT/A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1277	10389	6664	1714	1979	32	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	224	GLN	GLU	engineered mutation	UNP Q7B8V4
A	363	ALA	ARG	engineered mutation	UNP Q7B8V4
A	366	PHE	TYR	engineered mutation	UNP Q7B8V4
A	1158	ALA	THR	conflict	UNP Q7B8V4
A	1297	VAL	-	expression tag	UNP Q7B8V4
A	1298	PRO	-	expression tag	UNP Q7B8V4
A	1299	PRO	-	expression tag	UNP Q7B8V4
A	1300	THR	-	expression tag	UNP Q7B8V4
A	1301	PRO	-	expression tag	UNP Q7B8V4
A	1302	GLY	-	expression tag	UNP Q7B8V4
A	1303	SER	-	expression tag	UNP Q7B8V4
A	1304	ALA	-	expression tag	UNP Q7B8V4
A	1305	TRP	-	expression tag	UNP Q7B8V4
A	1306	SER	-	expression tag	UNP Q7B8V4
A	1307	HIS	-	expression tag	UNP Q7B8V4
A	1308	PRO	-	expression tag	UNP Q7B8V4
A	1309	GLN	-	expression tag	UNP Q7B8V4
A	1310	PHE	-	expression tag	UNP Q7B8V4
A	1311	GLU	-	expression tag	UNP Q7B8V4
A	1312	LYS	-	expression tag	UNP Q7B8V4

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

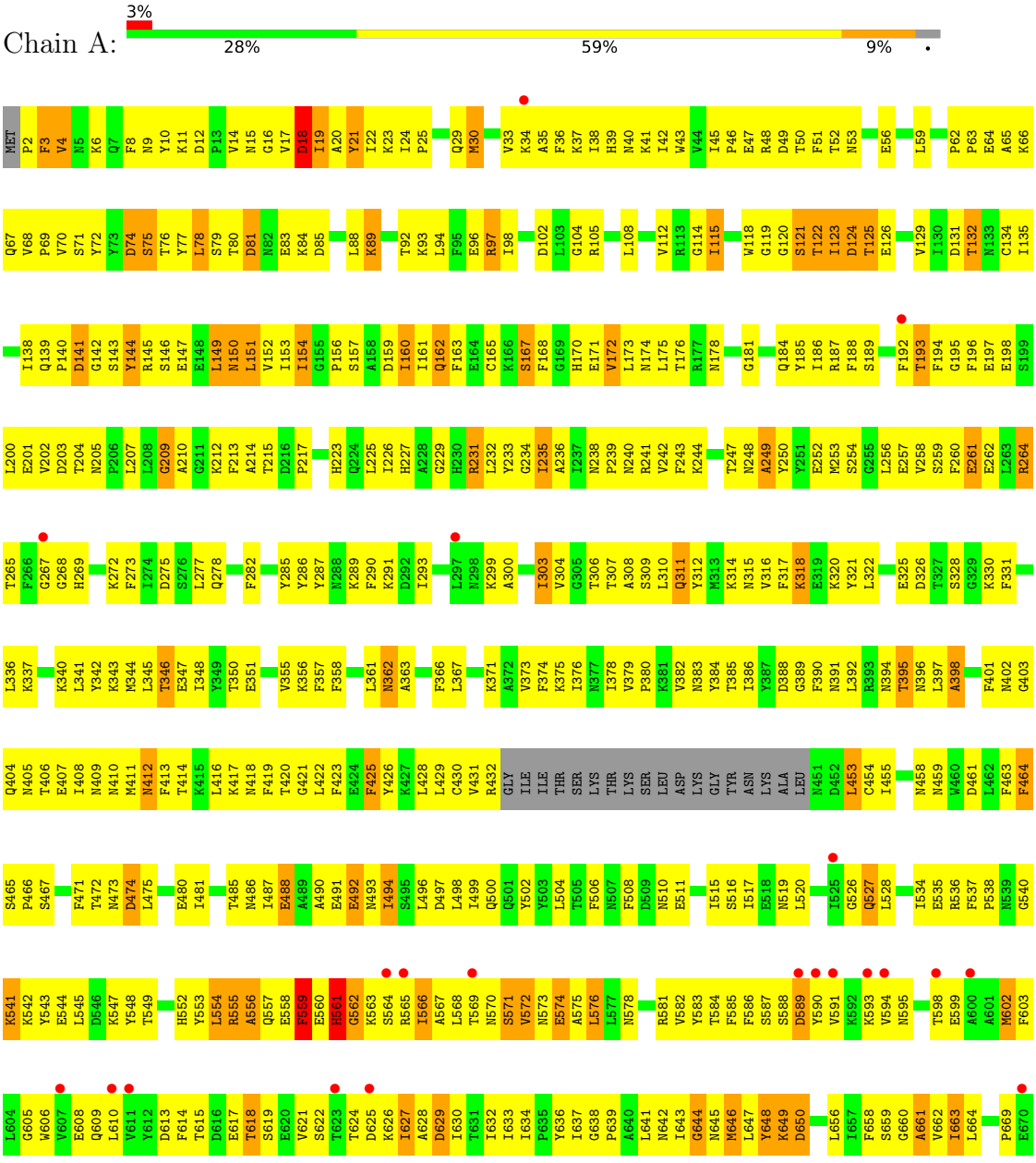
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0

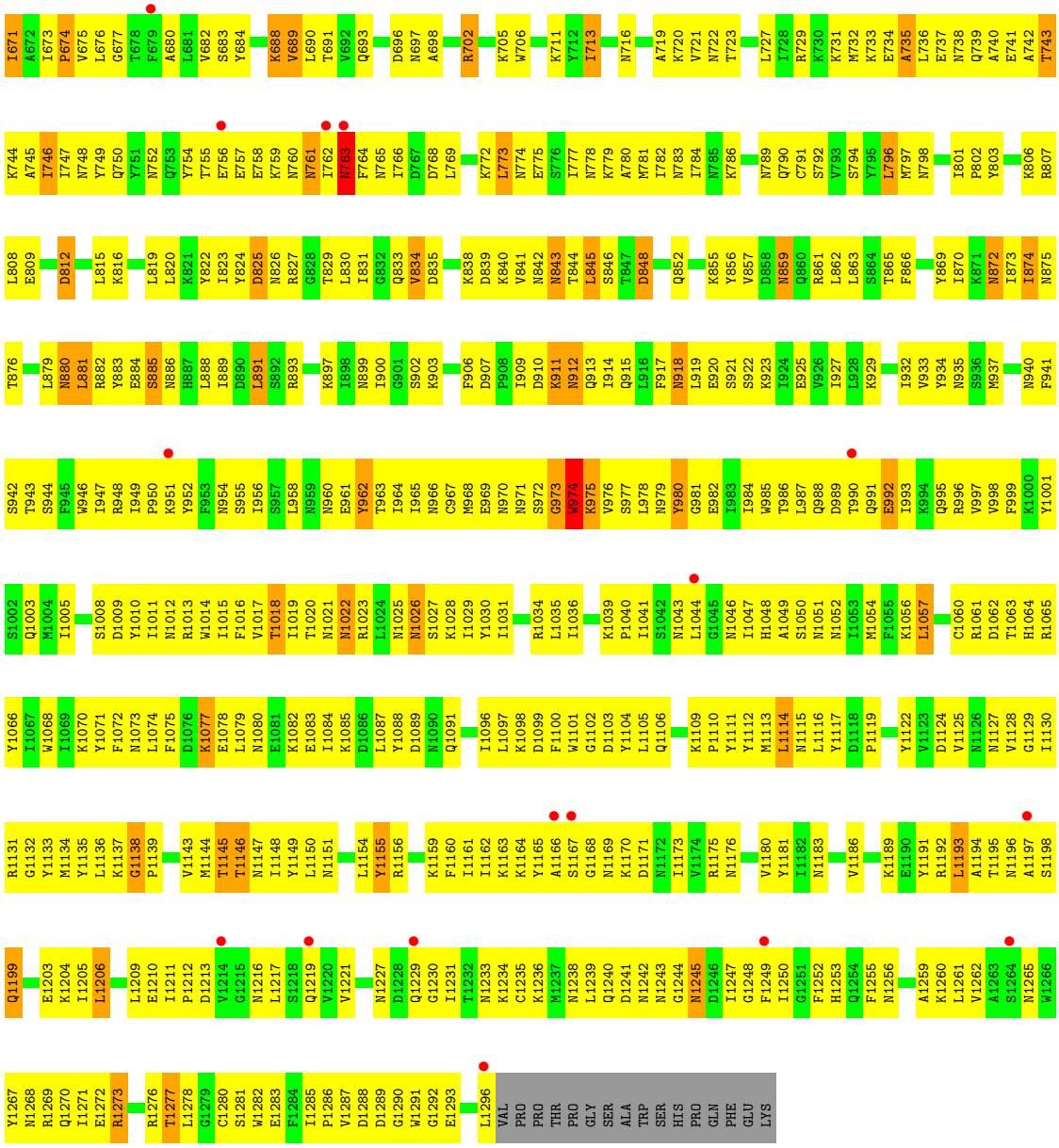
3 Residue-property plots (i)

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: BoNT/A

Chain A:




I671 A672 I673 P674 V675 L676 G677 T678 F679 A680 L681 V682 S683 Y684 K688 V689 L690 T691 V692 Q693 D696 N697 A698 R702 K705 W706 K711 Y712 I713 N716 A719 K720 V721 T722 T723 L727 I728 R729 K730 K731 M732 K733 E734 A735 L736 E737 N738 Q739 A740 E741 A742 T743
K744 A745 I746 I747 N748 Y749 Q750 Y751 N752 Q753 Y754 T755 E756 E757 E758 K759 N760 N761 I762 N763 F764 N765 I766 D767 D768 L769 K772 L773 N774 E775 S776 I777 N778 K779 A780 M781 I782 N783 I784 N785 K786 N789 Q790 C791 S792 V793 S794 Y796 L796 M797 N798 I801 P802 Y803 K806 R807
L808 E809 D812 L815 K816 L819 L820 K821 Y822 I823 Y824 D825 N826 R827 G828 T829 L830 I831 G832 Q833 V834 D835 K838 D839 K840 V841 N842 N843 T844 L845 S846 T847 D848 Q852 K855 Y856 V857 D858 N859 Q860 R861 L862 L863 S864 T865 F866 Y869 I870 K871 N872 I873 I874 N875
T876 L879 N880 L881 R882 Y883 E884 S885 N886 H887 L888 I889 D890 L891 S892 R893 K897 I898 N899 I900 Q901 S902 K903 F906 D907 P908 I909 D910 K911 N912 Q913 I914 Q915 L916 F917 N918 L919 E920 S921 S922 K923 I924 E925 V926 I927 L928 K929 I932 V933 Y934 N935 S936 M937 N940 F941
S942 T943 S944 F945 W946 I947 R948 I949 P950 K951 Y952 F953 N954 S955 I956 S957 L958 N959 N960 E961 Y962 T963 I964 I965 N966 C967 M968 E969 N970 N971 S972 G973 W974 K975 V976 S977 L978 N979 Y980 G981 E982 I983 I984 W985 T986 L987 Q988 D989 T990 Q991 E992 I993 K994 Q995 R996 V997 V998 F999 K1000 Y1001
S1002 Q1003 M1004 I1005 S1008 D1009 Y1010 I1011 N1012 R1013 W1014 I1015 F1016 V1017 T1018 I1019 T1020 N1021 N1022 R1023 L1024 N1025 N1026 S1027 K1028 I1029 Y1030 I1031 R1034 L1035 I1036 K1039 P1040 I1041 S1042 N1043 L1044 G1045 N1046 I1047 H1048 A1049 S1050 N1051 N1052 I1053 M1054 F1055 K1056 L1057 C1060 R1061 D1062 T1063 H1064 R1065
Y1066 I1067 W1068 I1069 K1070 Y1071 F1072 N1073 L1074 F1075 D1076 K1077 E1078 L1079 N1080 E1081 K1082 E1083 I1084 K1085 D1086 L1087 Y1088 D1089 N1090 Q1091 I1096 L1097 K1098 D1099 F1100 W1101 G1102 D1103 Y1104 L1105 Q1106 K1109 P1110 Y1111 Y1112 M1113 L1114 N1115 L1116 Y1117 D1118 P1119 Y1122 V1123 D1124 V1125 N1126 N1127 V1128 G1129 I1130
R1131 G1132 Y1133 M1134 Y1135 L1136 K1137 G1138 P1139 V1143 M1144 T1145 T1146 N1147 I1148 Y1149 L1150 N1151 L1154 Y1155 R1156 K1159 F1160 I1161 I1162 K1163 K1164 Y1165 A1166 S1167 G1168 N1169 K1170 D1171 N1172 I1173 V1174 R1175 N1176 V1180 Y1181 I1182 N1183 V1186 K1189 E1190 Y1191 R1192 L1193 A1194 T1195 N1196 A1197 S1198
Q1199 E1203 K1204 I1205 L1206 L1209 E1210 I1211 P1212 D1213 V1214 G1215 N1216 L1217 S1218 Q1219 V1220 V1221 N1227 D1228 Q1229 G1230 I1231 T1232 N1233 K1234 C1235 K1236 M1237 N1238 L1239 Q1240 D1241 N1242 N1243 G1244 N1245 D1246 I1247 G1248 F1249 I1250 G1251 F1252 H1253 Q1264 F1255 N1256 A1259 K1260 L1261 V1262 A1263 S1264 N1265 W1266
Y1267 N1268 R1269 Q1270 I1271 E1272 R1273 R1276 T1277 L1278 G1279 C1280 S1281 W1282 E1283 F1284 I1285 P1286 V1287 D1288 D1289 G1290 W1291 G1292 E1293 L1296 VAL PRO PRO THR PRO GLY SER ALA TRP SER HIS PRO GLN PHE GLU LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	167.52Å 167.52Å 158.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.11 – 4.30 45.11 – 4.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.11-4.30) 99.9 (45.11-4.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.96 (at 4.28Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, R_{free}	0.322 , 0.349 0.312 , 0.330	Depositor DCC
R_{free} test set	913 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	161.8	Xtriage
Anisotropy	0.203	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 241.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	10390	wwPDB-VP
Average B, all atoms (Å^2)	235.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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	1.16	0/10610	0.66	1/14367 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	151	LEU	CA-CB-CG	6.20	129.55	115.30

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	A	10389	0	10255	1168	9
2	A	1	0	0	0	0
All	All	10390	0	10255	1168	9

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 57.

The worst 5 of 1168 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:277:LEU:HD22	1:A:474:ASP:HB3	1.24	1.13
1:A:872:ASN:HD21	1:A:874:ILE:HB	1.12	1.11
1:A:948:ARG:HB3	1:A:1068:TRP:HB2	1.28	1.10
1:A:310:LEU:HD11	1:A:314:LYS:HE3	1.38	1.05
1:A:969:GLU:H	1:A:972:SER:HB3	1.24	1.02

The worst 5 of 9 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:486:ASN:OD1	1:A:1273:ARG:NH2[6_555]	1.12	1.08
1:A:63:PRO:O	1:A:309:SER:N[3_564]	1.82	0.38
1:A:486:ASN:CG	1:A:1273:ARG:NH2[6_555]	1.90	0.30
1:A:693:GLN:NE2	1:A:1276:ARG:CB[6_555]	1.96	0.24
1:A:697:ASN:ND2	1:A:1276:ARG:NH2[6_555]	2.00	0.20

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	1273/1312 (97%)	1008 (79%)	177 (14%)	88 (7%)	1 17

5 of 88 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	121	SER
1	A	398	ALA
1	A	488	GLU
1	A	541	LYS
1	A	559	PHE

5.3.2 Protein sidechains

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	1158/1190 (97%)	1082 (93%)	76 (7%)	16 43

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	843	ASN
1	A	1077	LYS
1	A	874	ILE
1	A	974	TRP
1	A	1277	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	915	GLN
1	A	1012	ASN
1	A	960	ASN
1	A	979	ASN
1	A	1026	ASN

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 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

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	1277/1312 (97%)	0.14	37 (2%) 51 41	119, 222, 356, 750	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	590	TYR	5.5
1	A	569	THR	5.1
1	A	594	VAL	4.7
1	A	1044	LEU	4.1
1	A	625	ASP	4.1

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 monosaccharides 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
2	ZN	A	1313	1/1	0.93	0.47	147,147,147,147	0

6.5 Other polymers (i)

There are no such residues in this entry.