wwPDB X-ray Structure Validation Summary Report (i)

Oct 26, 2023 – 02:09 AM EDT

PDB ID	:	3AGQ
Title	:	Structure of viral polymerase form II
Authors	:	Takeshita, D.; Tomita, K.
Deposited on	:	2010-04-06
Resolution	:	3.22 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 (i)) were used in the production of this report:

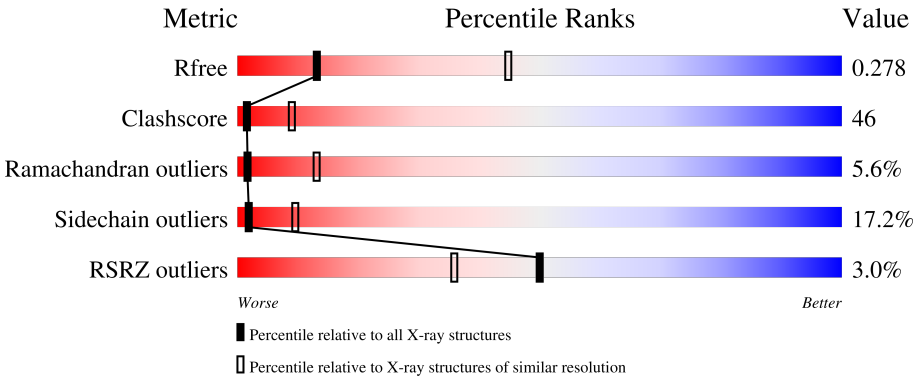
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.22 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	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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	1289	3% 30% 50% 12% 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9257 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 Elongation factor Ts, Elongation factor Tu 1, LINKER, Q beta replicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1199	9252	5843	1600	1764	45	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	284	HIS	-	linker	UNP P0A6P3
A	1284	HIS	-	expression tag	UNP Q8LTE0
A	1285	HIS	-	expression tag	UNP Q8LTE0
A	1286	HIS	-	expression tag	UNP Q8LTE0
A	1287	HIS	-	expression tag	UNP Q8LTE0
A	1288	HIS	-	expression tag	UNP Q8LTE0
A	1289	HIS	-	expression tag	UNP Q8LTE0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is water.

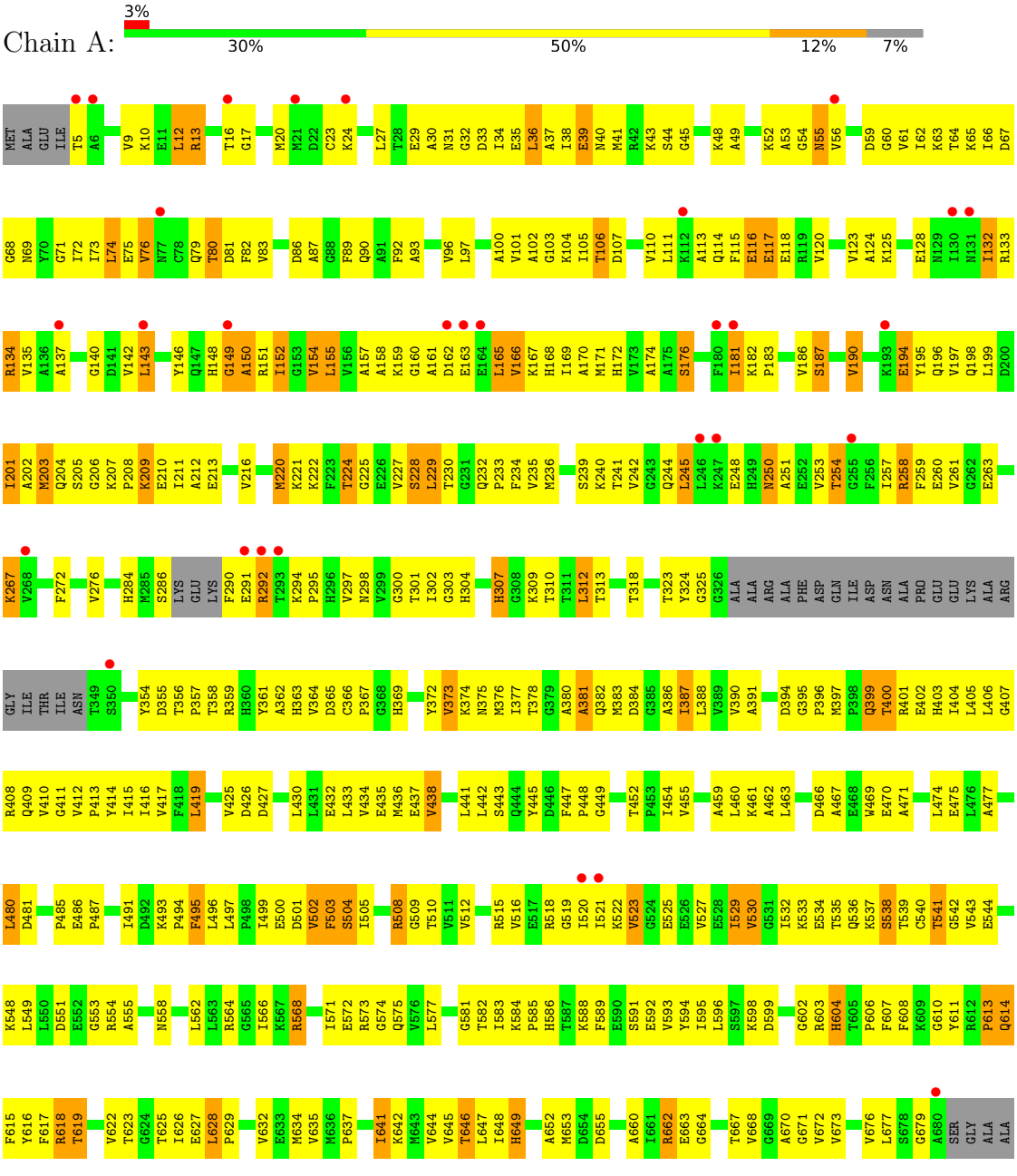
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	4	4	4	0	0

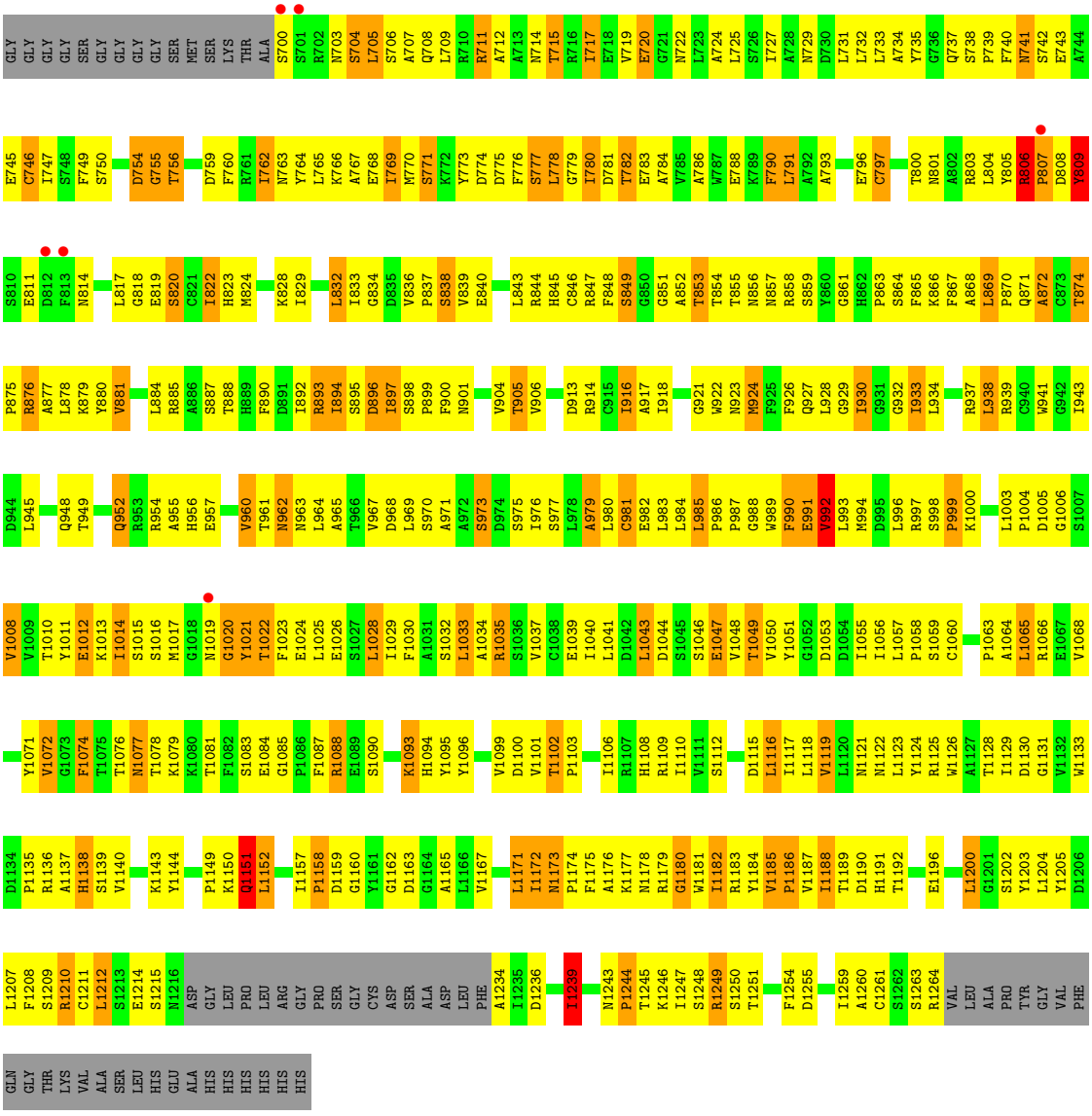
3 Residue-property plots (i)

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: Elongation factor Ts, Elongation factor Tu 1, LINKER, Q beta replicase

Chain A:




GLY GLY GLY GLY SER GLY GLY GLY GLY SER MET SER LYS THR ALA S700 S701 R702 N703 S704 L705 S706 A707 Q708 L709 R710 R711 A712 A713 N714 T715 R716 I717 E718 V719 E720 G721 N722 L723 A724 L725 S726 I727 A728 N729 D730 L731 L732 L733 A734 Y735 G736 Q737 S738 P739 F740 N741 S742 E743 A744
E745 C746 I747 S748 F749 S750 D754 G755 T756 D759 F760 R761 I762 N763 Y764 L765 K766 A767 E768 I769 M770 S771 K772 Y773 D774 D775 F776 S777 L778 G779 I780 D781 T782 E783 A784 V785 A786 W787 E788 K789 F790 L791 A792 A793 E796 C797 T800 N801 A802 R803 L804 Y805 R806 P807 D808 V809
S810 E811 D812 F813 N814 L817 G818 E819 S820 C821 I822 H823 M824 K828 I829 L832 I833 G834 D835 V836 M770 P837 S838 V839 E840 L843 R844 H845 C846 R847 F848 S849 G850 G851 A852 T853 T854 T855 N856 N857 R858 S859 Y860 G861 H862 P863 S864 F865 K866 F867 A868 L869 P870 Q871 A872 C873 T874
P875 R876 A877 L878 K879 Y880 V881 L884 R885 A886 S887 T888 H889 F890 D891 I892 R893 I894 S895 D896 I897 S898 P899 F900 N901 V904 T905 V906 D913 R914 C915 I916 A917 I918 G921 W922 N923 M924 F925 F926 Q927 L928 G929 I930 G931 G932 I933 L934 R937 L938 R939 G940 W941 G942 I943
D944 L945 Q948 T949 Q952 R953 R954 A955 H956 E957 V960 T961 N962 N963 L964 A965 T966 V967 D968 L969 S970 A971 A972 S973 D974 S975 I976 S977 L978 A979 L980 C981 E982 L983 L984 L985 P986 P987 G988 W989 F990 E991 V992 L993 M994 D995 L996 R997 S998 P999 K1000 L1003 P1004 D1005 G1006 S1007
V1008 V1009 T1010 Y1011 E1012 K1013 I1014 S1015 S1016 M1017 G1018 N1019 G1020 Y1021 T1022 F1023 E1024 L1025 E1026 S1027 L1028 I1029 F1030 A1031 S1032 L1033 A1034 R1035 S1036 V1037 C1038 E1039 I1040 L1041 D1042 L1043 D1044 S1045 S1046 E1047 V1048 T1049 V1050 Y1051 G1052 D1053 D1054 I1055 I1056 L1057 P1058 S1059 C1060 P1063 A1064 L1065 R1066 E1067 V1068
Y1071 V1072 G1073 F1074 T1075 T1076 N1077 T1078 K1079 K1080 T1081 F1082 S1083 E1084 G1085 P1086 F1087 R1088 E1089 S1090 K1093 H1094 Y1095 Y1096 V1099 D1100 V1101 T1102 P1103 I1106 R1107 H1108 R1109 I1110 V1111 S1112 D1115 L1116 I1117 L1118 V1119 L1120 N1121 N1122 L1123 Y1124 R1125 W1126 A1127 T1128 I1129 D1130 G1131 V1132 W1133
D1134 P1135 R1136 A1137 H1138 S1139 V1140 K1143 Y1144 P1149 K1150 G1151 L1152 I1157 P1158 D1159 G1160 Y1161 G1162 D1163 G1164 A1165 L1166 V1167 L1171 I1172 N1173 P1174 F1175 A1176 K1177 N1178 R1179 G1180 W1181 I1182 R1183 Y1184 V1185 P1186 V1187 I1188 T1189 D1190 H1191 T1192 E1196 L1200 G1201 S1202 Y1203 L1204 Y1205 D1206
L1207 F1208 S1209 R1210 C1211 L1212 S1213 E1214 S1215 N1216 ASP GLY LEU PRO LEU ARG GLY PRO SER GLY CYS ASP SER ALA ASP LEU PHE A1234 I1235 D1236 I1239 N1243 P1244 T1245 K1246 I1247 S1248 R1249 S1250 T1251 F1254 D1255 I1259 A1260 C1261 S1262 S1263 R1264 VAL LEU ALA PRO TYR GLY VAL PHE
GLN GLY THR LYS VAL ALA SER LEU HIS GLU ALA HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	138.77Å 255.12Å 100.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.22 30.28 – 3.22	Depositor EDS
% Data completeness (in resolution range)	97.6 (20.00-3.22) 97.6 (30.28-3.22)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.44 (at 3.24Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.251 , 0.317 0.222 , 0.278	Depositor DCC
R_{free} test set	1452 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	86.5	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 58.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	0.077 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.022 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9257	wwPDB-VP
Average B, all atoms (Å^2)	132.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.26% 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: MG

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	0.40	0/9421	0.62	1/12741 (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	286	SER	CA-C-O	-18.36	81.55	120.10

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	9252	0	9234	845	0
2	A	1	0	0	0	0
3	A	4	0	0	0	0
All	All	9257	0	9234	845	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 46.

The worst 5 of 845 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:930:ILE:HD11	1:A:1021:TYR:HD2	1.17	1.07
1:A:893:ARG:HH11	1:A:893:ARG:HB3	1.20	1.06
1:A:133:ARG:HG2	1:A:134:ARG:H	1.20	1.04
1:A:198:GLN:HA	1:A:201:ILE:HG12	1.37	1.04
1:A:399:GLN:H	1:A:399:GLN:HE21	1.06	1.02

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	1189/1289 (92%)	944 (79%)	179 (15%)	66 (6%)	2 13

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	GLU
1	A	134	ARG
1	A	150	ALA
1	A	176	SER
1	A	209	LYS

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	991/1060 (94%)	821 (83%)	170 (17%)	2 9

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	905	THR
1	A	1049	THR
1	A	933	ILE
1	A	991	GLU
1	A	1102	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	649	HIS
1	A	1155	ASN
1	A	741	ASN
1	A	1094	HIS
1	A	729	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

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	1199/1289 (93%)	-0.02	36 (3%) 50 36	72, 137, 182, 211	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	700	SER	5.3
1	A	5	THR	4.3
1	A	701	SER	4.0
1	A	6	ALA	3.6
1	A	812	ASP	3.5

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	MG	A	2001	1/1	0.67	0.34	84,84,84,84	0

6.5 Other polymers (i)

There are no such residues in this entry.