



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:20 AM GMT

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
<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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

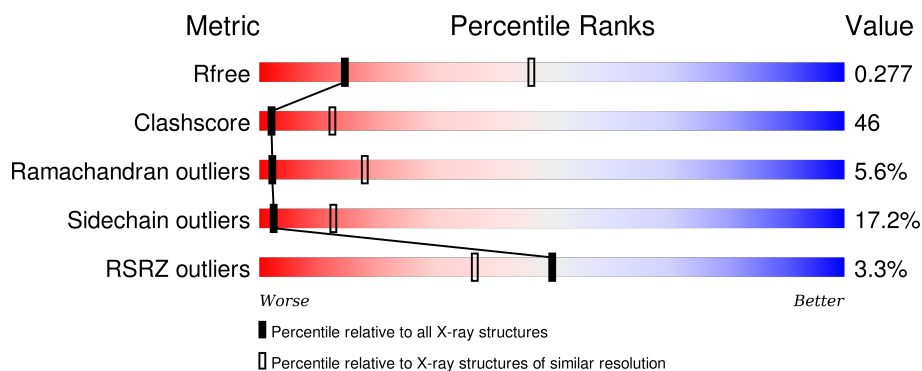
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	91344	1095 (3.26-3.18)
Clashscore	102246	1046 (3.24-3.20)
Ramachandran outliers	100387	1026 (3.24-3.20)
Sidechain outliers	100360	1025 (3.24-3.20)
RSRZ outliers	91569	1100 (3.26-3.18)

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. 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	1289	

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
1	A	1199	Total	C	N	O	S	0	0	0
			9252	5843	1600	1764	45			

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
2	A	1	Total	Mg	0	0
			1	1		

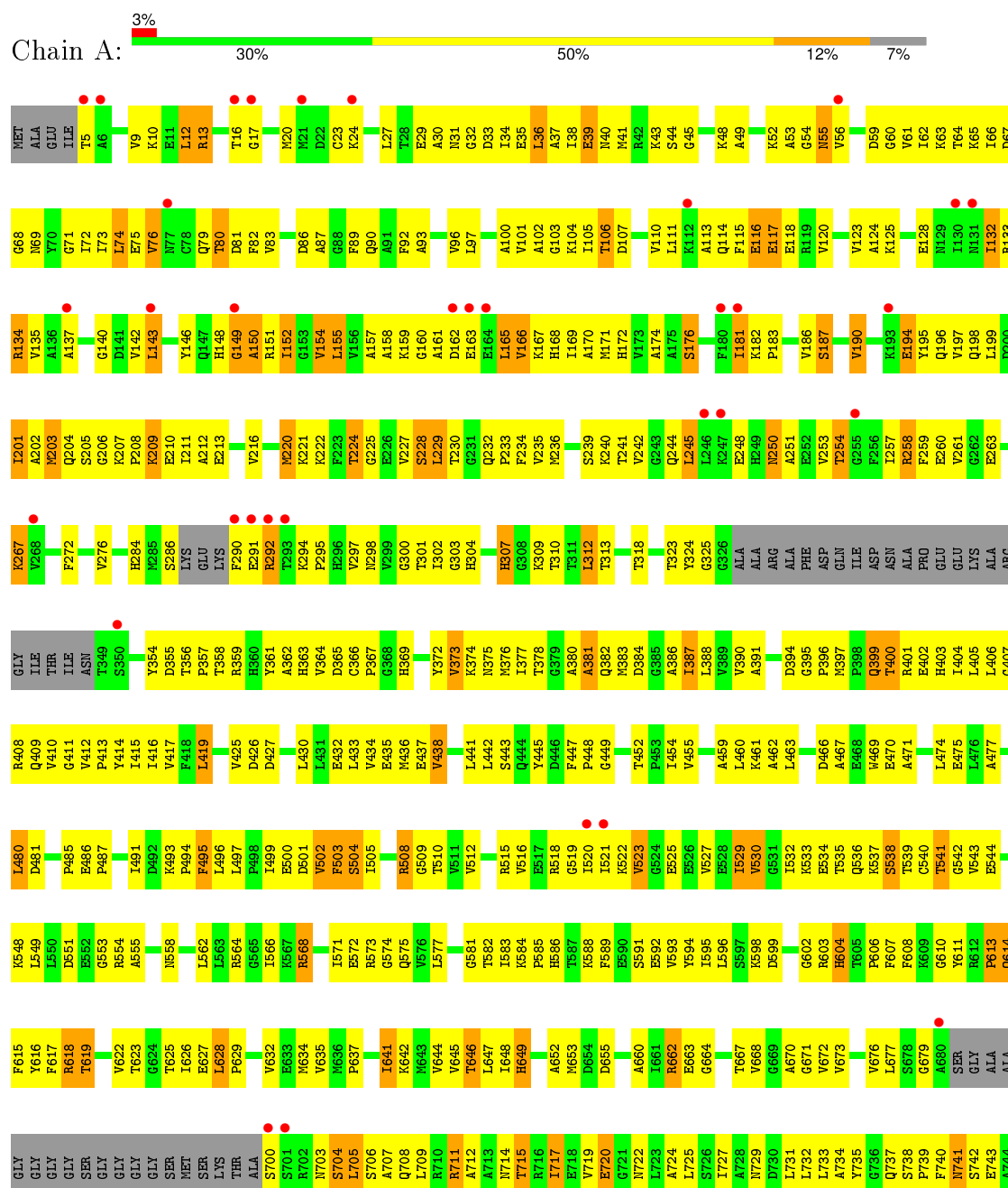
- Molecule 3 is water.

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

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



PHE	D1206	D1134		V1008	D944	P875	S810	E745
GLN	L1207	F1136	Y1071	V1009	L945	R876	E811	C746
GLY	F1208	A1137	V1072	T1010		A877	D812	I747
THR	S1209	A1137	G1073	Y1011	Q948	L878	F813	S748
LYS	R1210	H1138	F1074	E1012	T949	K379	N814	F749
VAL	C1211	S1139	T1075	K1013		Y880		S750
ALA	L1212	V1140	T1076	I1014	Q952	V881	L817	
SER	S1213		M1077	S1015	R953		G818	D754
LEU	E1214	K1143	T1078	S1016	R954	L884	E819	G755
HIS	S1215	Y1144	K1079	M1017	A955	R885	S820	T756
GLU	M1216		K1080	G1018	R956	A886	C821	
ALA	ASP	P1149	T1081	N1019	E957	S887	I822	F759
GLY	GLY	K1150	F1082	G1020		T888	H823	D760
HIS	LEU	G1151	S1083	Y1021	V960	H889	K824	R761
HIS	PRO	L1152	E1084	T1022	T961	F890		I762
HIS	LEU		G1085	F1023	N962	D891	K828	M763
HIS	ARG	I1157	P1086	E1024	N963	I892	I829	Y764
GLY	GLY	F1158	F1087	L1025	N964	R893	L765	L765
PRO	PRO	D1159	R1088	E1026	A965	I894	K832	K766
SER	SER	G1160	E1089	S1027	T966	S895	I833	A767
GLY	GLY	Y1161	S1090	L1028	V967	D896	E834	E768
CYS	ASP	G1162		I1029	D968	I897	G835	I769
ASP	ASP	D1163	K1093	F1030	L969	S898	V836	M770
ALA	ALA	G1164	H1094	A1031	S970	P899	P837	S771
ALA	ASP	A1165	Y1095	S1032	A971	F900	S838	K772
LEU	LEU	L1166	Y1096	L1033	A972	N901	V839	Y773
PHE	PHE	V1167		A1034	S973		E840	D774
	A1234	L1171	V1099	R1035	D974	V904		D775
I1235	I1235	L1172	D1100	S1036	S975	T905	L843	F776
D1236	D1236	M1173	V1101	V1037	I976	V906	R844	S777
		P1174	T1102	C1038	S977		H845	L778
I1239		F1175	P1103	E1039	L978	D813	C846	G779
		K1176		I1040	A979	R914	R847	I780
M1243		K1177	I1106	L1041	C915	C915	F848	D781
P1244		N1178	H1107	D1042	C981	I916	S849	T782
T1245		R1179	H1108	L1043	E982	A917	G850	E783
K1246		G1180	I1110	D1044	L983	I918	G851	A784
I1247		W1181	V1111	S1045	L984		A852	V785
S1248		R1182	S1112	S1046	L985	G921	T853	A786
R1249		R1183		E1047	P986	W922	T854	I787
S1250		Y1184		V1048	P987	N923	T855	E788
T1251		V1185	D1115	T1049	G988	N924	N856	K789
		P1186	L1116	V1050	W989	F925	N857	F790
F1254		V1187	L1117	Y1051	F990	F926	R858	L791
D1255		I1188	L1118	G1052	E991	Q927	S859	A792
		T1189	V1119	D1053	V992	L928	Y860	A793
I1259		D1190	L1120	D1054	L993	G929	G861	
A1260		H1191	N1121	I1055	M994	I930	H862	E796
C1261		T1192	L1122	L1056	D995	G931	P863	C797
S1262		R1193	L1123	L1057	L996	I932	S864	
S1263			Y1124	P1058	A997	I933	F865	T800
R1264		E1196	R1125	S1059	S998	L934	K866	N801
VAL	VAL		W1126	C1060	P999		F867	A802
LEU	LEU	L1200	A1127	P1063	K1000	R937	A868	R803
ALA	ALA	G1201	T1128	A1064		L938	L869	L804
PRO	PRO	S1202	I1129	L1065	L1003	R939	P870	Y805
TYR	TYR	Y1203	D1130	R1066	P1004	C940	Q871	R806
GLY	GLY	L1204	G1131	E1067	D1005	W941	A872	P807
VAL	VAL	Y1205	W1133	V1068	G1006	G942	C873	D808
					S1007	I943	T874	Y809

4 Data and refinement statistics

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
$\langle I/\sigma(I) \rangle$ ¹	2.44 (at 3.24Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.251 , 0.317 0.221 , 0.277	Depositor DCC
R_{free} test set	1451 reflections (5.07%)	DCC
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
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
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	2 of 28652 reflections (0.007%)	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 $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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(°)	Ideal(°)
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 [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	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:HE21	1:A:399:GLN:H	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 18

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
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	991/1060 (94%)	821 (83%)	170 (17%)	2 12

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

Mol	Chain	Res	Type
1	A	662	ARG
1	A	791	LEU
1	A	1172	ILE
1	A	705	LEU
1	A	741	ASN

Some 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	614	GLN
1	A	729	ASN
1	A	1151	GLN
1	A	649	HIS
1	A	741	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](#)

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	1199/1289 (93%)	0.01	39 (3%)	50	37	72, 137, 182, 211	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	700	SER	5.4
1	A	5	THR	4.4
1	A	701	SER	4.1
1	A	6	ALA	3.7
1	A	812	ASP	3.6

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	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.