



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 02:03 am BST

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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

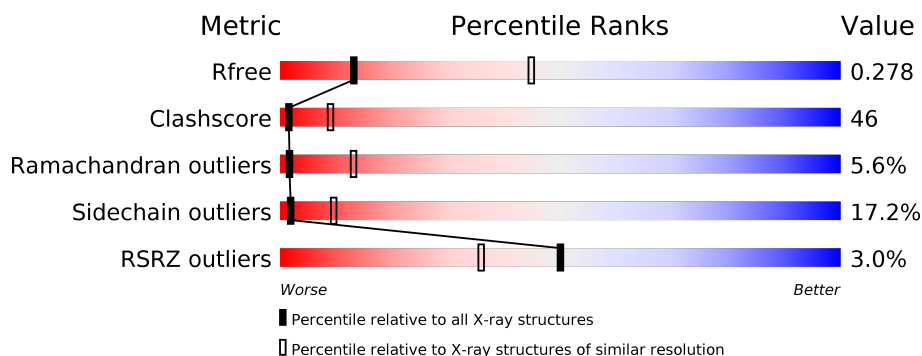
# 1 Overall quality at a glance

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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

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		

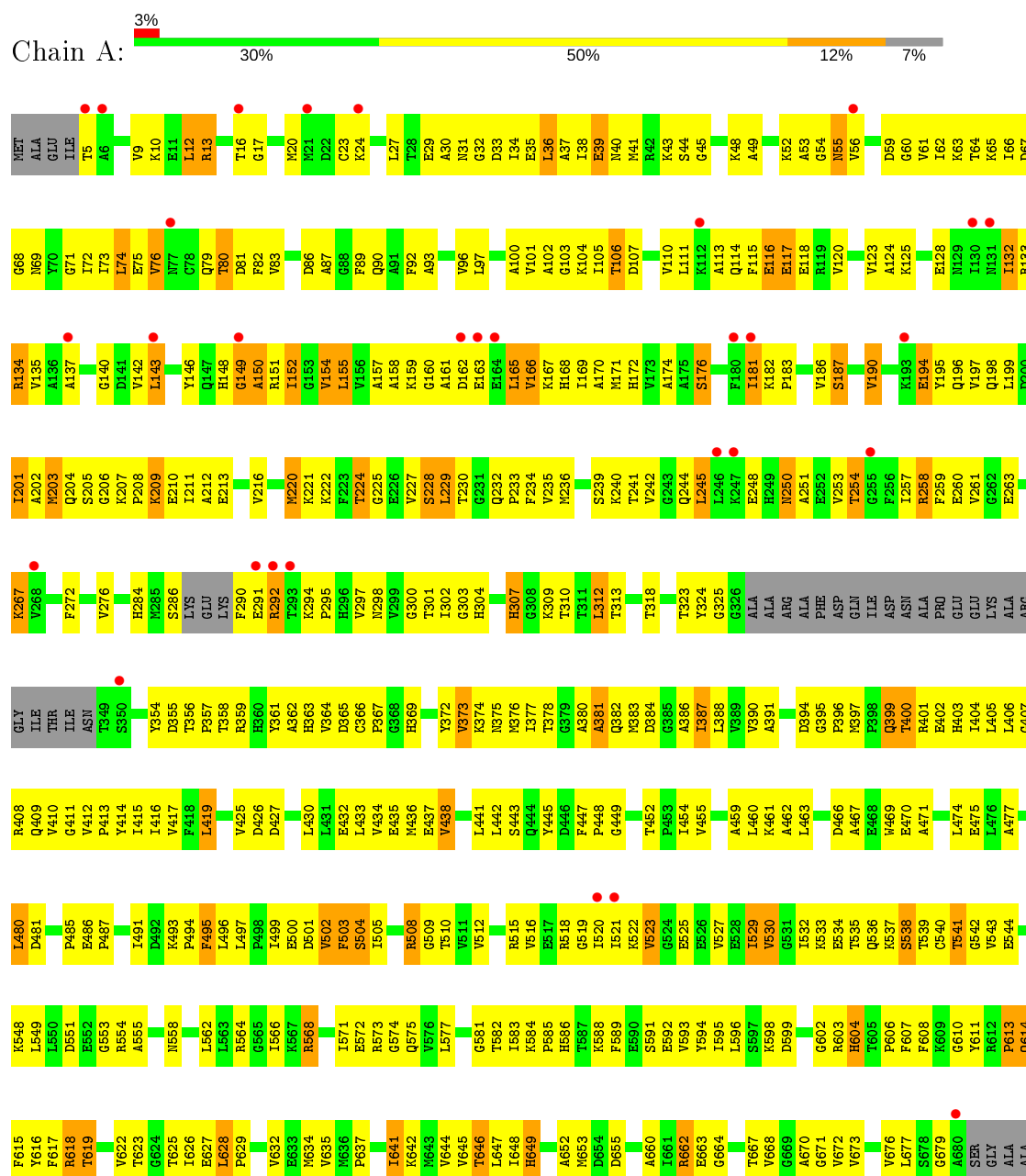
- Molecule 3 is water.

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 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



GLN	L1207	D1134	Y1071	Y1008	D944	P875	S810	E745	GLY
GLY	F1208	P1135	Y1072	V1009	L945	R876	E811	C746	GLY
THR	S1209	R1136	G1073	T1010	A877	L878	D812	C747	GLY
LYS	R1210	A1137	F1073	Y1011	Q948	L879	F813	S748	GLY
VAL	C1211	H1138	T1074	K1012	T949	K879	N814	F749	SER
ALA	L1212	S1139	F1075	K1013	Y880	V881	N814	S750	GLY
SER	S1213	V1140	T1076	I1014	Q982	V881	L817	D754	GLY
LEU	E1214	K1143	M1077	S1015	R983	L884	G818	D755	GLY
HIS	S1215	Y1144	T1078	S1016	R984	R885	E819	G755	GLY
GLU	N1216	K1080	K1079	M1017	A955	R885	S820	T756	SER
ALA	ASP	P1149	K1080	G1018	H956	A886	C821	W759	MET
HIS	LEU	P1150	T1081	N1019	E957	S887	I822	D759	SER
HIS	LEU	K1150	F1082	G1020	V960	T888	H823	F760	LYS
HIS	PRO	Q1151	S1083	Y1021	T961	H889	H824	K761	THR
HIS	ARG	L1152	E1084	T1022	N962	P990	I829	I762	ALA
HIS	GLY	I1157	G1085	F1023	N963	L892	K828	N763	S700
PRO	SER	P1158	F1086	E1024	N964	R893	I829	Y764	S701
GLY	CYS	D1159	F1087	L1025	L964	T894	L832	L765	R702
ASP	ALA	G1160	R1088	E1026	A965	S895	I833	K766	N703
LEU	LEU	Y1161	S1089	S1027	T966	D896	G834	A767	S704
PHE	PHE	V1167	S1090	L1028	V967	I897	D835	I768	L705
ALA	ALA	L1171	K1093	F1030	D968	S898	V836	I769	S706
LEU	LEU	I1172	H1094	A1031	L969	P899	P837	M770	A707
ALA	ALA	N1173	Y1095	S1032	S970	F900	S838	S771	Q708
LEU	LEU	P1174	Y1096	L1033	A971	N901	V839	K772	L709
ALA	ALA	F1175	V1099	A1034	A972	V904	E940	Y773	K710
LEU	LEU	I1176	D1100	R1035	D974	N904	D774	D775	R711
LEU	LEU	K1177	Y1101	S1036	S975	T905	D775	A712	A712
LEU	LEU	R1178	T1102	V1037	I976	V906	R844	S776	A713
LEU	LEU	G1180	P1103	C1038	S977	D913	H845	S777	N714
LEU	LEU	Y1181	I1106	E1038	L978	R914	C846	L778	T715
LEU	LEU	I1182	R1107	I1040	A979	R915	R847	G779	T716
LEU	LEU	R1183	H1108	L1041	L980	C915	I780	I781	I717
LEU	LEU	Y1184	R1109	D1042	C981	I916	D782	V719	E718
LEU	LEU	V1185	R1179	L1043	E982	A917	G850	E720	E720
LEU	LEU	P1186	I1110	D1044	L983	I918	G851	E721	E721
LEU	LEU	V1187	T1111	S1045	L984	I918	A852	A784	G721
LEU	LEU	I1188	D1115	S1046	L985	G921	T853	V785	N722
LEU	LEU	T1189	L1116	E1047	P986	N922	T854	A786	I723
LEU	LEU	D1190	L1117	V1048	P987	N923	T855	A787	A724
LEU	LEU	H1191	L1118	T1049	G988	N924	T856	L725	L725
LEU	LEU	T1192	L1119	V1050	W989	F925	N856	K789	S726
LEU	LEU	E1196	L1120	G1052	P990	F926	N857	F790	I727
LEU	LEU	L1200	A1127	D1053	E991	Q927	R858	L791	A728
LEU	LEU	G1201	T1128	D1054	V992	L928	T860	A792	N729
LEU	LEU	S1202	I1129	I1055	L993	N929	G861	A793	I730
LEU	LEU	Y1203	D1130	I1056	N994	G930	H862	L731	L731
LEU	LEU	Y1204	L1131	I1057	D995	G931	S863	L732	L732
LEU	LEU	Y1205	V1132	L1057	L996	G932	S864	L733	L733
LEU	LEU	D1206	W1133	P1058	R997	I933	F865	A734	Y735
LEU	LEU	L1207	W1126	S1059	S998	I934	K866	G796	G736
LEU	LEU	G1208	A1127	C1060	P999	I934	F867	C797	Q737
LEU	LEU	S1209	T1128	P1063	K1000	R937	A868	A802	Q737
LEU	LEU	Y1210	I1129	A1064	L1003	L938	R803	S738	S738
LEU	LEU	Y1211	D1130	L085	P1004	R939	L804	F739	F739
LEU	LEU	Y1212	L1131	R1066	D1005	G940	Y805	F740	F740
LEU	LEU	Y1213	V1132	E1067	G1006	R941	K806	N741	N741
LEU	LEU	Y1214	W1133	V1068	S1007	G942	P807	S742	S742
LEU	LEU	Y1215	W1133	V1068	S1007	G942	D808	E743	E743
LEU	LEU	Y1216	W1133	V1068	S1007	G942	Y809	A744	A744

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	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 <sup>2</sup> )	86.5	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 58.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	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 <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: 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:HB3	1:A:893:ARG:HH11	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%)	<b>2</b> <b>13</b>

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%)	<b>2</b> <b>9</b>

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, 95<sup>th</sup> 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 <sup>2</sup> )	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 [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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	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.