



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 06:51 PM UTC

PDB ID : 3AGP / pdb_00003agp
Title : Structure of viral polymerase form I
Authors : Takeshita, D.; Tomita, K.
Deposited on : 2010-04-06
Resolution : 2.80 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

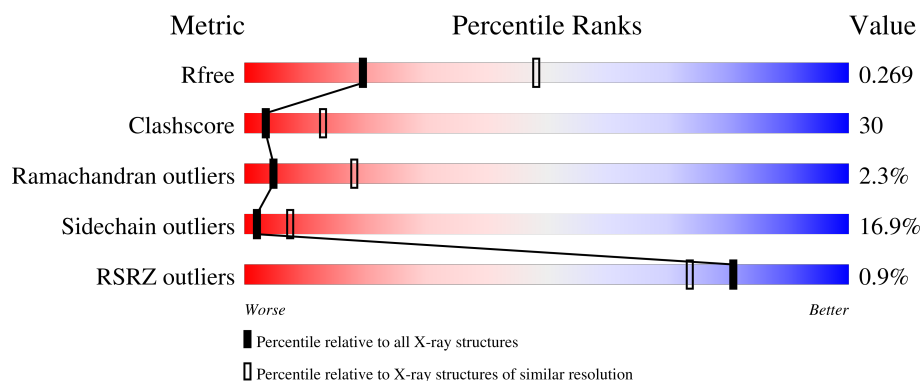
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	180053	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (2.80-2.80)

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 $\leq 5\%$. 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	46% 37% 10% 7%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9311 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, LINKER, Q beta replicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1202	Total	C	N	O	S	0	0	0
			9274	5857	1603	1769	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 CALCIUM ION (CCD ID: CA) (formula: Ca).

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

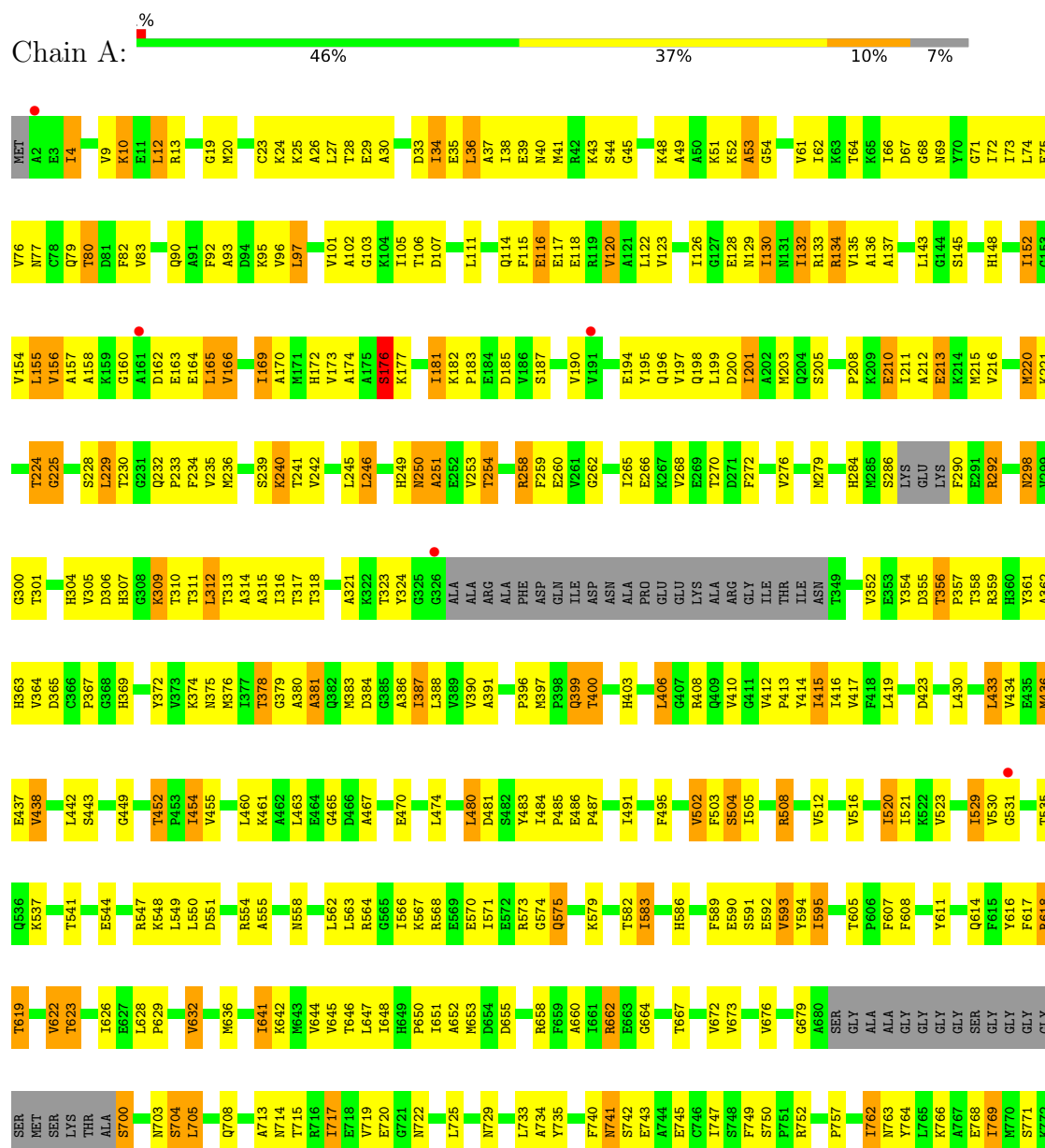
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	36	Total	O	0	0
			36	36		

3 Residue-property plots

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, LINKER, Q beta replicase



PHE	I1188	R1088	D995	P920	E840	Y773
GLN	T1189	E1089	L996	G921	G841	D774
THR	T1192	S1090	R997	N922	M842	D776
LYS	R1193	K1093	P999	N923	L843	F776
VAL	D1194	H1094	L1003	F926	C846	S777
ALA	R1195	H1095	P1004	Q927	R847	L778
SER	E1196	Y1096	V1008	T930	F848	G779
LEU	R1197	V1099	E1012	G931	S849	T780
HIS	L1200	D1100	K1013	G932	T853	D781
GLU	R1210	Y1101	E1014	I934	T854	T782
ALA	G1211	T1102	T1014	R935	T855	A784
HIS	L1212	Y1105	S1015	L938	M856	V785
HIS	S1215	T1106	M1017	I943	M857	A786
HIS	N1216	R1109	Y1021	D947	S859	E787
ASP	GLY	I1110	T1022	Q948	H862	F788
LEU	PRO	L1116	F1023	T949	S863	K789
PRO	LEU	L1121	E1024	N951	S864	F790
ARG	GLY	Y1124	L1028	Q952	F865	L791
PRO	GLY	W1125	T1029	R953	K866	
SER	GLY	A1127	F1030	R954	F867	C797
CYS	ASP	T1128	L1033	A955	L868	A798
SER	SER	I1129	I1040	G958	L869	L799
ALA	ALA	D1130	L1041	S959	P870	T800
ASP	ASP	R1136	D1042	V960	Q871	N801
LEU	PHE	K1143	L1043	T961	A872	A802
PHE	A1234	Y1144	D1044	G962	C873	L804
I1235	I1234	K1150	S1045	N963	T874	Y805
R1241	S1242	Q1151	E1046	L964	V881	R806
N1243	P1244	L1152	E1047	T966	S887	P807
T1245	K1246	I1157	Y1048	A965	T892	D808
K1246	I1247	L1165	T1049	V967	R893	Y809
I1247	S1248	V1167	V1050	D968	T894	S810
R1249	S1250	I1172	I1055	L969	F812	E811
F1254	F1254	P1173	L1056	S970	N814	F813
S1263	R1264	F1174	L1057	A971	S815	N814
VAL	LEU	R1175	P1058	L978	S816	F815
LEU	ALA	R1179	G1072	A979	L817	S816
ALA	PRO	T1182	F1073	L980	T822	L817
PRO	TYR	R1183	T1075	C981	H823	
TYR	VAL	Y1184	T1076	E982	M824	
VAL		V1185	M1077	L983		K828
			T1078	L984		K831
			T1081			L832
						G834
						D835
						V836
						P837
						S838
						V839

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	139.50Å 256.43Å 100.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.90 – 2.80 19.90 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.2 (19.90-2.80) 97.9 (19.90-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 2.80Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.215 , 0.283 0.210 , 0.269	Depositor DCC
R_{free} test set	2230 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	60.4	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 34.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.022 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.023 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9311	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: CA

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	0/9443	0.90	16/12771 (0.1%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	910	SER	N-CA-C	-7.27	104.54	113.41
1	A	1243	ASN	N-CA-C	7.01	119.76	110.36
1	A	1243	ASN	CA-C-N	6.93	128.50	119.84
1	A	1243	ASN	C-N-CA	6.93	128.50	119.84
1	A	251	ALA	N-CA-C	6.91	120.43	109.65

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	9274	0	9256	550	0
2	A	1	0	0	0	0
3	A	36	0	0	3	0
All	All	9311	0	9256	550	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 30.

The worst 5 of 550 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:491:ILE:HG22	1:A:555:ALA:HB3	1.36	1.07
1:A:1100:ASP:OD1	1:A:1102:THR:HG22	1.59	1.02
1:A:133:ARG:HG2	1:A:134:ARG:H	1.25	1.01
1:A:854:THR:HG21	1:A:872:ALA:HB3	1.42	1.01
1:A:806:ARG:H	1:A:807:PRO:HD2	1.31	0.96

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	1192/1289 (92%)	1045 (88%)	120 (10%)	27 (2%)	5 18

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	53	ALA
1	A	778	LEU
1	A	806	ARG
1	A	970	SER
1	A	19	GLY

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	993/1060 (94%)	825 (83%)	168 (17%)	2 7

5 of 168 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	873	CYS
1	A	1040	ILE
1	A	894	ILE
1	A	950	ILE
1	A	1093	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1121	ASN
1	A	1122	ASN
1	A	1173	ASN
1	A	558	ASN
1	A	536	GLN

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 oligosaccharides 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

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	1202/1289 (93%)	-0.31	11 (0%) 81 74	35, 75, 124, 148	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	326	GLY	4.3
1	A	1235	ILE	4.0
1	A	1216	ASN	3.9
1	A	807	PRO	2.9
1	A	2	ALA	2.5

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 oligosaccharides 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, 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	B-factors(Å ²)	Q<0.9
2	CA	A	2001	1/1	0.95	0.10	82,82,82,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.