



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 15, 2026 – 01:11 AM UTC

PDB ID : 6QPF / pdb_00006qpf
Title : Influenza A virus Polymerase Heterotrimer A/duck/Fujian/01/2002(H5N1)
Authors : Fan, H.T.; Keown, J.R.; Fodor, E.; Grimes, J.M.
Deposited on : 2019-02-13
Resolution : 3.63 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

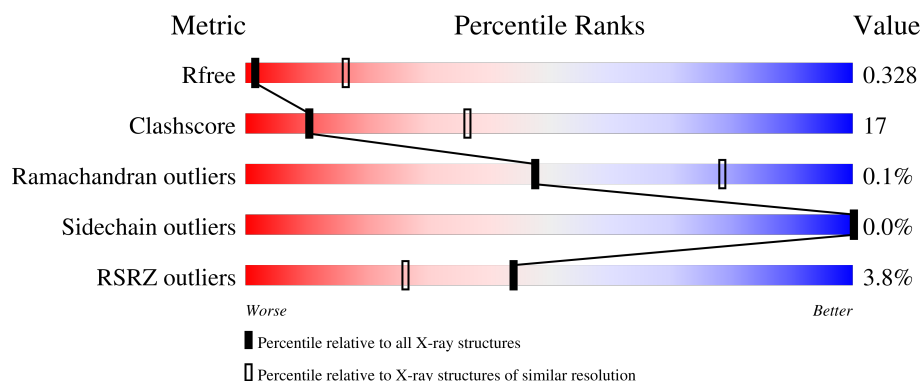
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.63 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	1132 (3.76-3.52)
Clashscore	190562	1014 (3.74-3.54)
Ramachandran outliers	187476	1123 (3.76-3.52)
Sidechain outliers	187428	1121 (3.76-3.52)
RSRZ outliers	180081	1130 (3.76-3.52)

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	716	3% 62% 38%
1	D	716	3% 68% 32%
1	G	716	2% 62% 38%
1	J	716	4% 62% 37% .
2	B	757	4% 56% 34% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	757	 4% 52% 36% 13%
2	H	757	 3% 56% 32% 11%
2	K	757	 4% 52% 35% 13%
3	C	765	 4% 64% 31% 5%
3	F	765	 4% 63% 32% 1%
3	I	765	 4% 62% 33% 6%
3	L	765	 4% 60% 35% 5%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 67760 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	716	Total	C	N	O	S	0	0	0
			5796	3671	985	1099	41			
1	D	716	Total	C	N	O	S	0	0	0
			5796	3671	985	1099	41			
1	G	716	Total	C	N	O	S	0	0	0
			5796	3671	985	1099	41			
1	J	710	Total	C	N	O	S	0	0	0
			5751	3643	978	1089	41			

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	685	Total	C	N	O	S	0	0	0
			5474	3446	950	1037	41			
2	E	661	Total	C	N	O	S	0	0	0
			5283	3328	909	1005	41			
2	H	673	Total	C	N	O	S	0	0	0
			5389	3393	936	1020	40			
2	K	662	Total	C	N	O	S	0	0	0
			5290	3339	913	998	40			

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	727	Total	C	N	O	S	0	0	0
			5790	3643	1042	1066	39			
3	F	731	Total	C	N	O	S	0	0	0
			5824	3661	1050	1074	39			
3	I	722	Total	C	N	O	S	0	0	0
			5755	3620	1038	1058	39			
3	L	730	Total	C	N	O	S	0	0	0
			5816	3657	1048	1072	39			

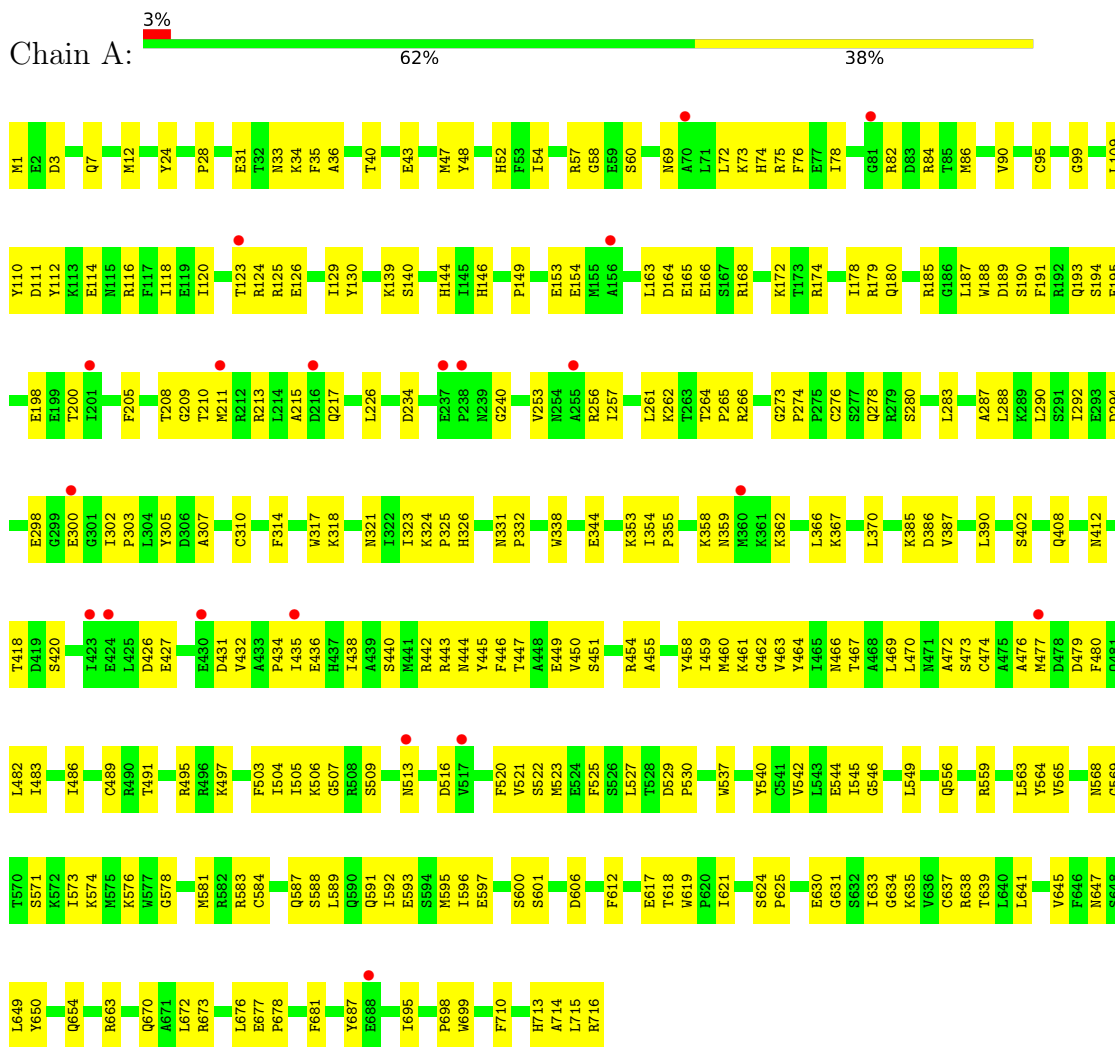
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	760	GLU	-	expression tag	UNP D3YNB5
C	761	ASN	-	expression tag	UNP D3YNB5
C	762	LEU	-	expression tag	UNP D3YNB5
C	763	TYR	-	expression tag	UNP D3YNB5
C	764	PHE	-	expression tag	UNP D3YNB5
C	765	GLN	-	expression tag	UNP D3YNB5
F	760	GLU	-	expression tag	UNP D3YNB5
F	761	ASN	-	expression tag	UNP D3YNB5
F	762	LEU	-	expression tag	UNP D3YNB5
F	763	TYR	-	expression tag	UNP D3YNB5
F	764	PHE	-	expression tag	UNP D3YNB5
F	765	GLN	-	expression tag	UNP D3YNB5
I	760	GLU	-	expression tag	UNP D3YNB5
I	761	ASN	-	expression tag	UNP D3YNB5
I	762	LEU	-	expression tag	UNP D3YNB5
I	763	TYR	-	expression tag	UNP D3YNB5
I	764	PHE	-	expression tag	UNP D3YNB5
I	765	GLN	-	expression tag	UNP D3YNB5
L	760	GLU	-	expression tag	UNP D3YNB5
L	761	ASN	-	expression tag	UNP D3YNB5
L	762	LEU	-	expression tag	UNP D3YNB5
L	763	TYR	-	expression tag	UNP D3YNB5
L	764	PHE	-	expression tag	UNP D3YNB5
L	765	GLN	-	expression tag	UNP D3YNB5

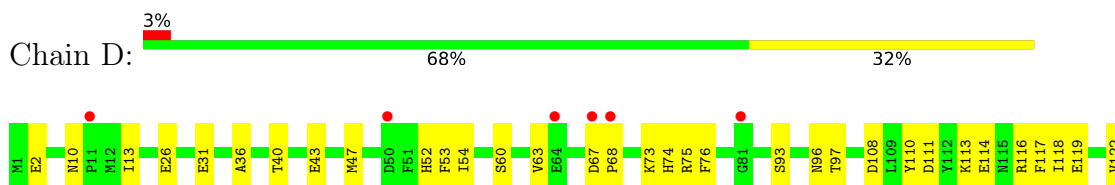
3 Residue-property plots

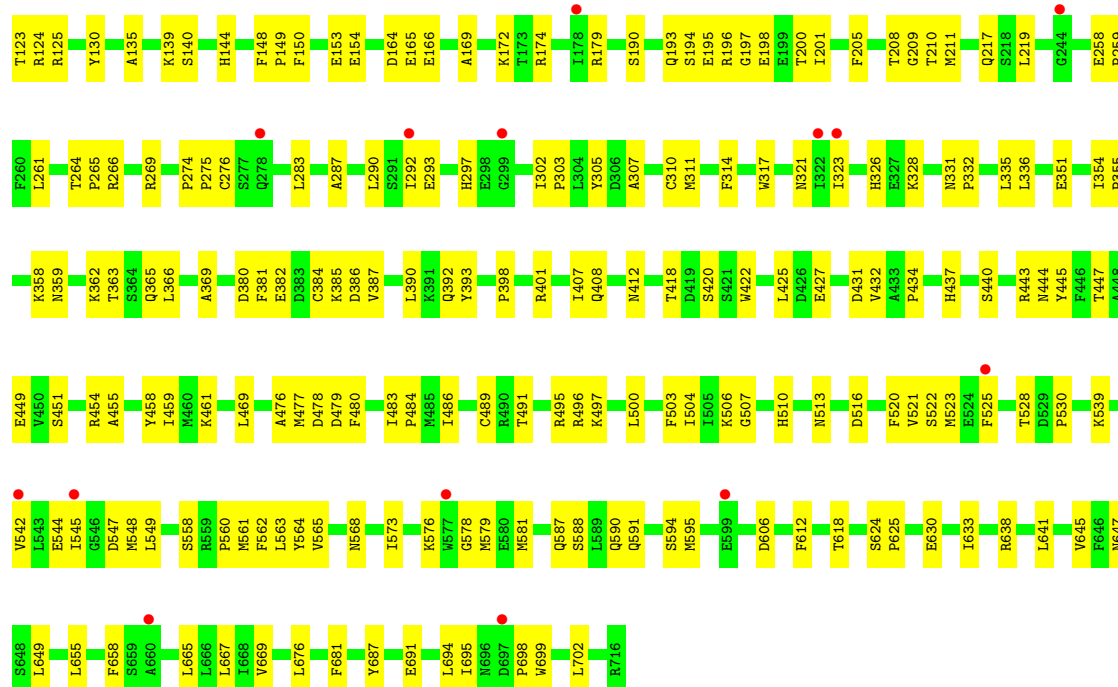
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Polymerase acidic protein

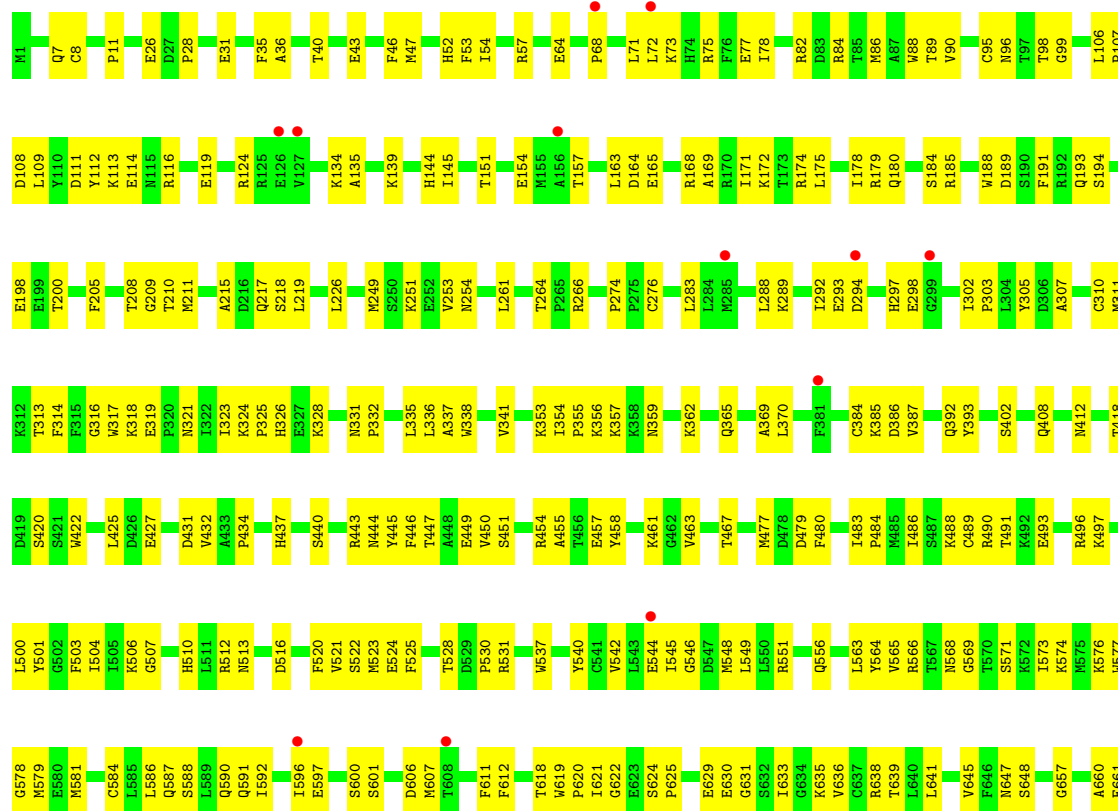


• Molecule 1: Polymerase acidic protein

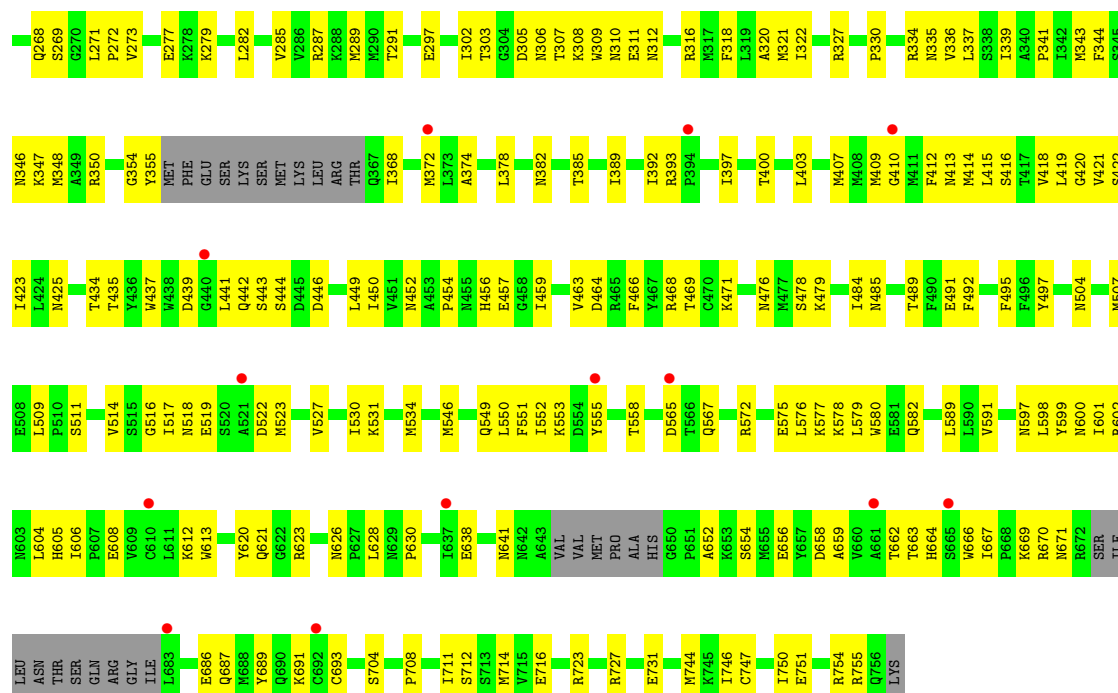




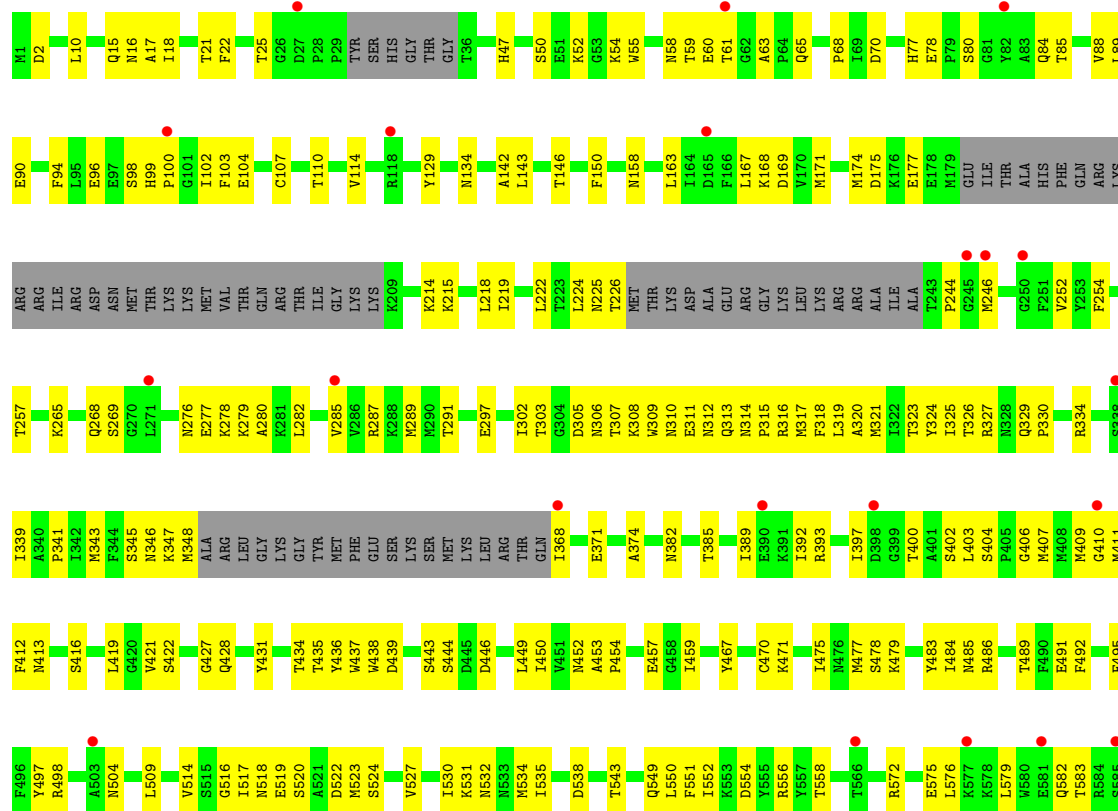
• Molecule 1: Polymerase acidic protein

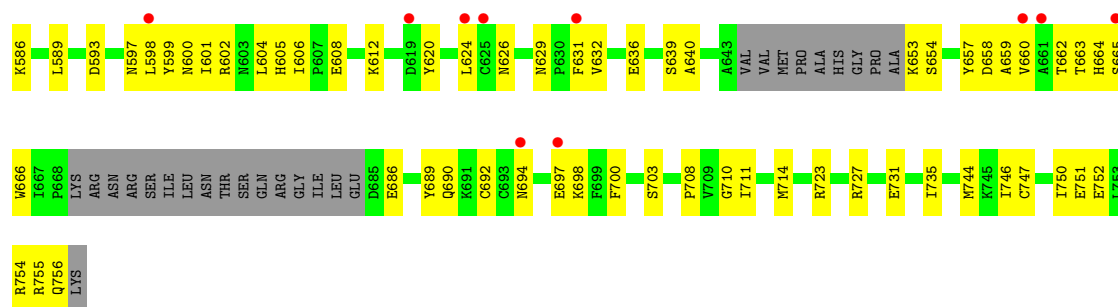




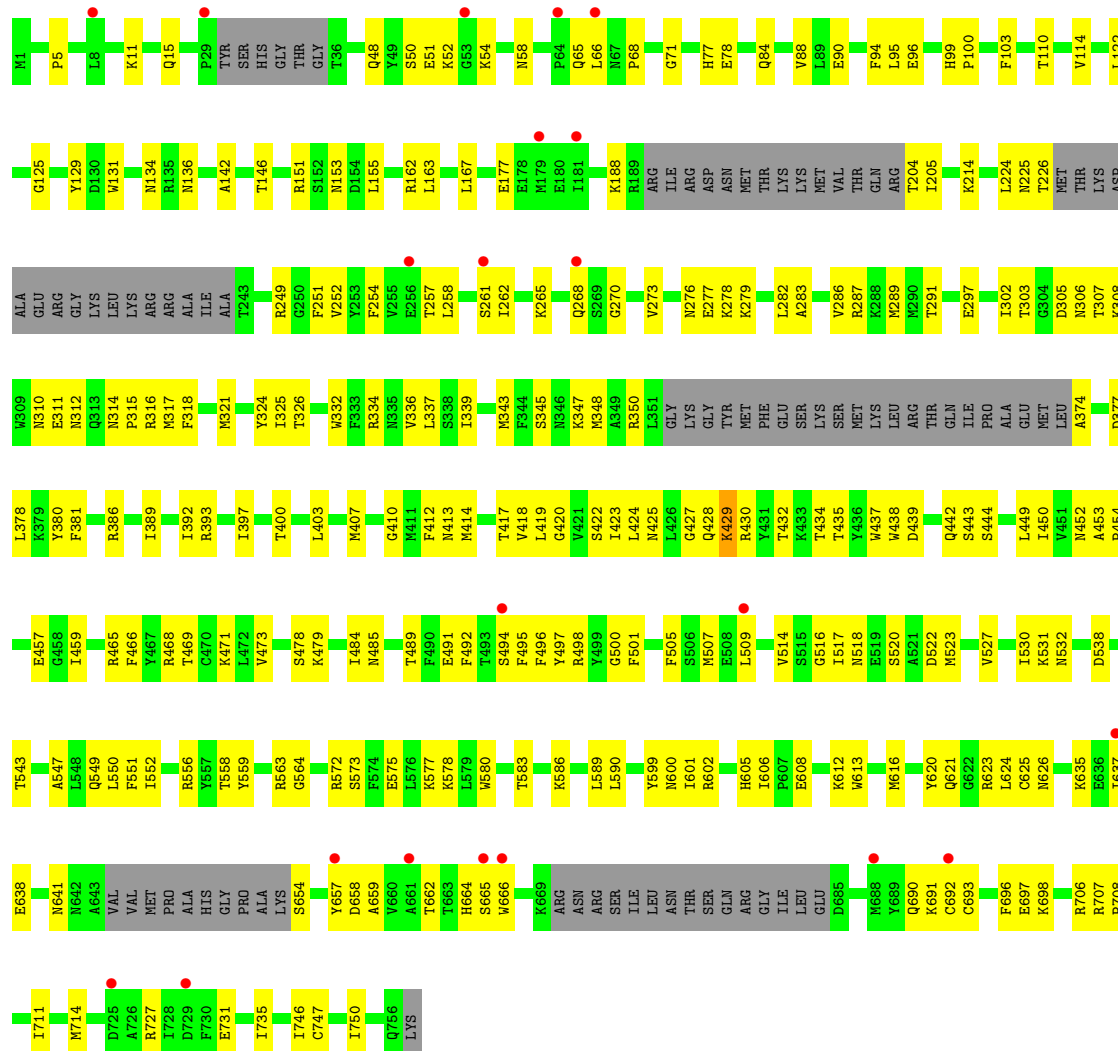


• Molecule 2: RNA-directed RNA polymerase catalytic subunit



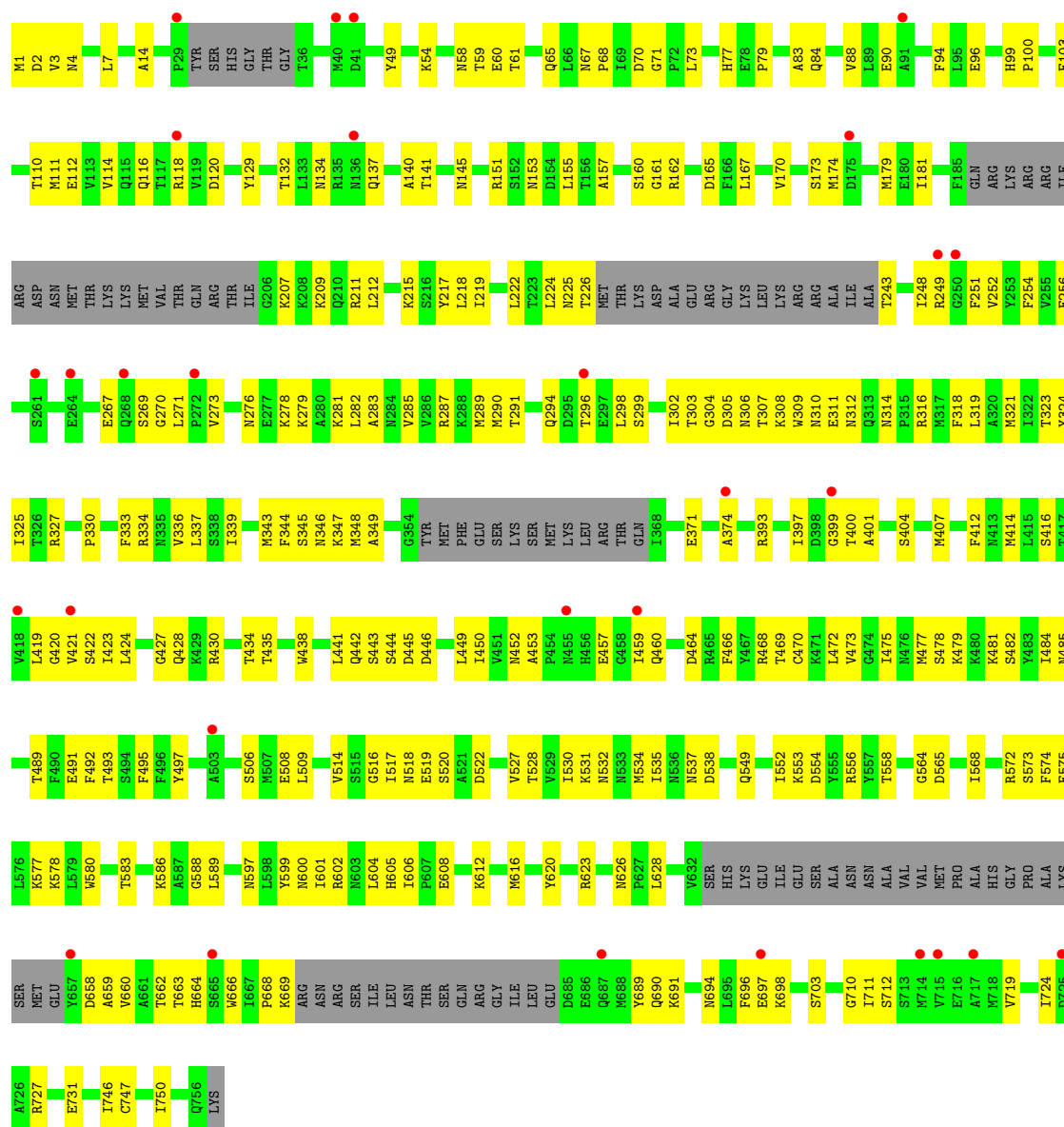


• Molecule 2: RNA-directed RNA polymerase catalytic subunit

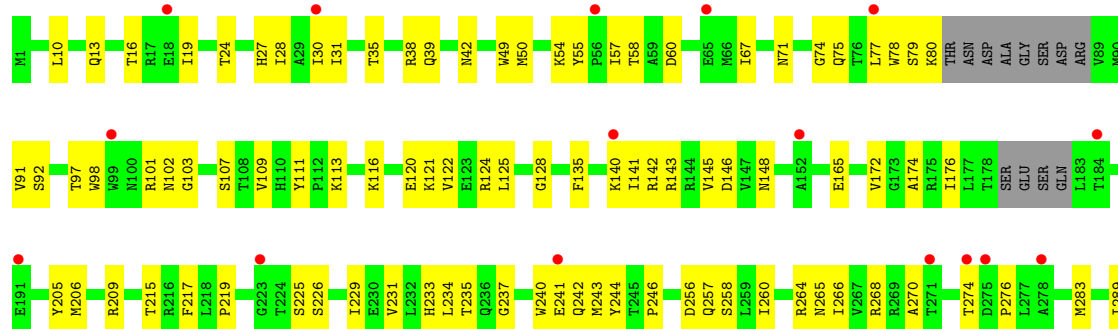


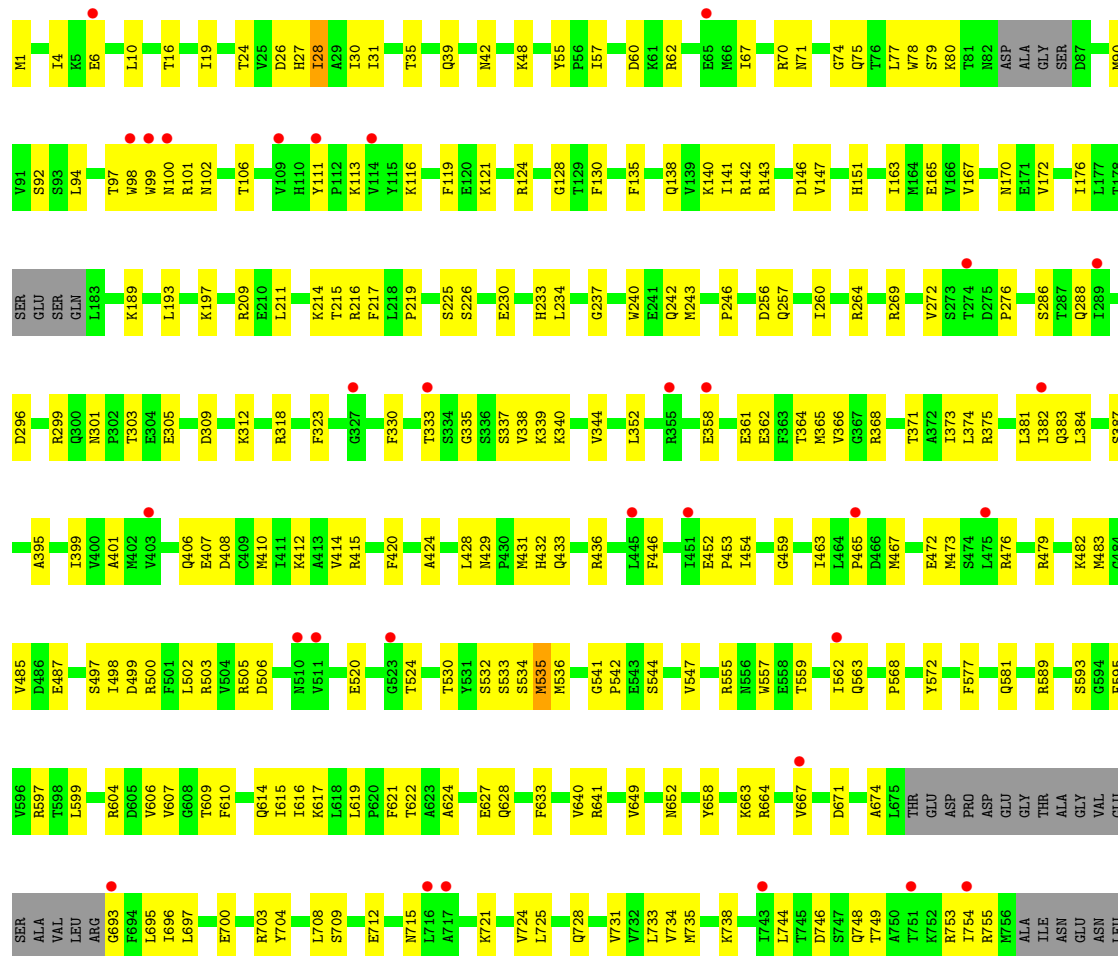
• Molecule 2: RNA-directed RNA polymerase catalytic subunit





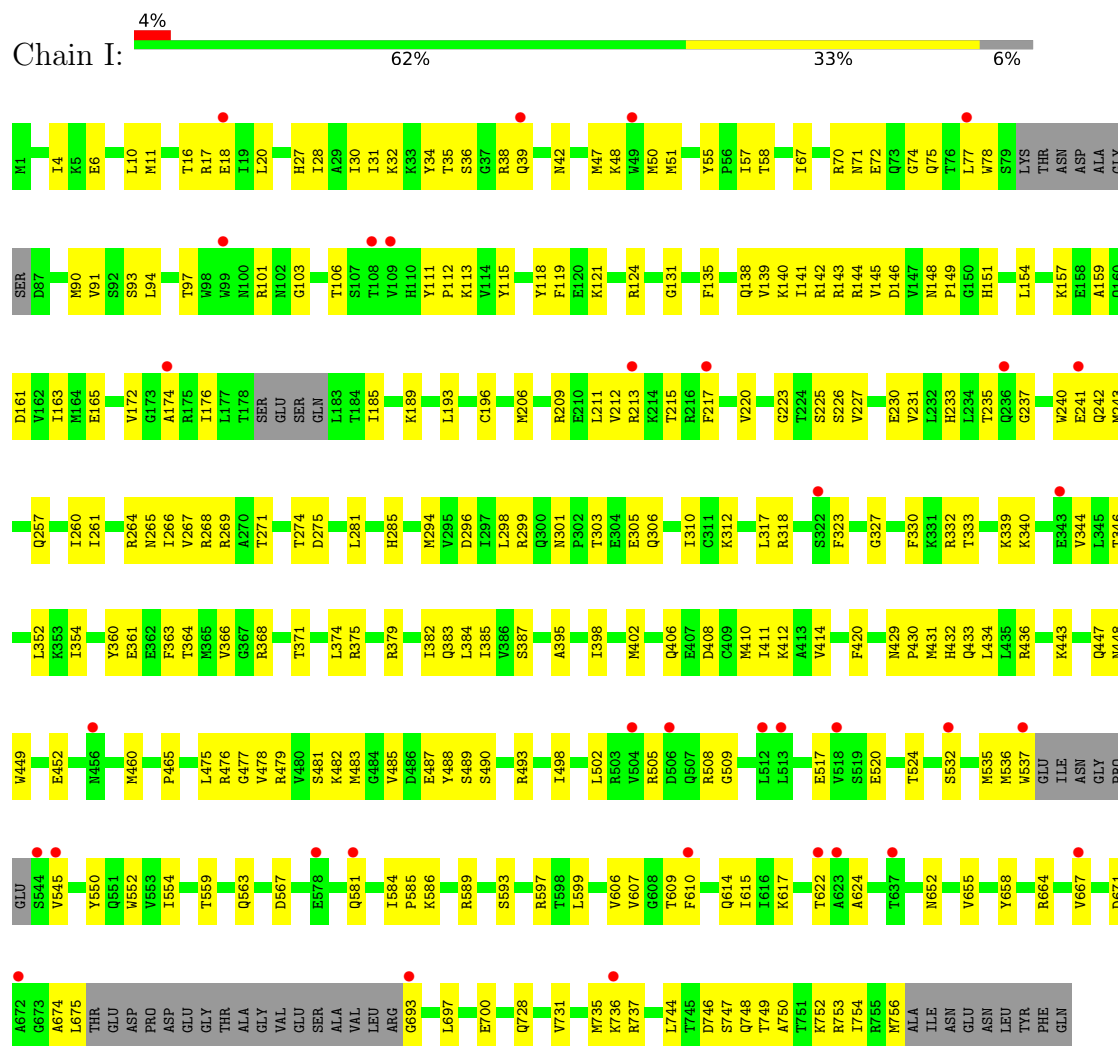
• Molecule 3: Polymerase basic protein 2



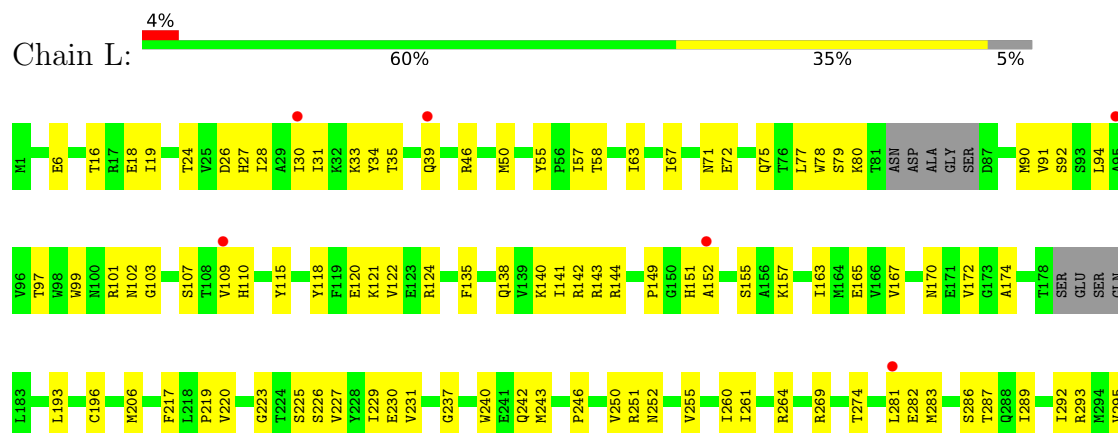


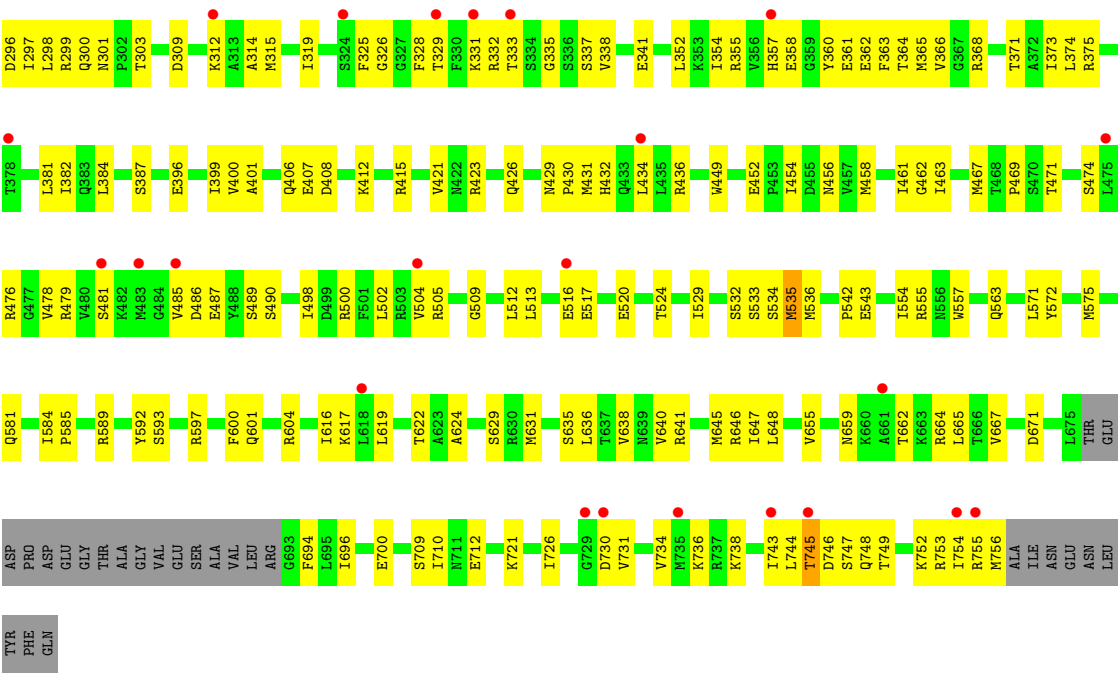
TYR
PHE
GLN

• Molecule 3: Polymerase basic protein 2



• Molecule 3: Polymerase basic protein 2





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	337.12Å 192.86Å 235.69Å 90.00° 91.46° 90.00°	Depositor
Resolution (Å)	235.61 – 3.63 235.61 – 3.63	Depositor EDS
% Data completeness (in resolution range)	48.0 (235.61-3.63) 48.0 (235.61-3.63)	Depositor EDS
R_{merge}	0.48	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 3.68Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.277 , 0.323 0.281 , 0.328	Depositor DCC
R_{free} test set	4071 reflections (2.40%)	wwPDB-VP
Wilson B-factor (Å ²)	76.8	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 114.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.023 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.027 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.026 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.035 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.035 for -h,-k,l	Xtriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	67760	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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](#)

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	0.18	0/5921	0.42	0/7981
1	D	0.17	0/5921	0.41	0/7981
1	G	0.17	0/5921	0.43	0/7981
1	J	0.16	0/5874	0.42	0/7915
2	B	0.18	0/5580	0.43	0/7527
2	E	0.17	0/5386	0.41	0/7270
2	H	0.18	0/5493	0.43	0/7409
2	K	0.17	0/5394	0.42	0/7278
3	C	0.17	0/5883	0.44	0/7928
3	F	0.16	0/5917	0.44	0/7974
3	I	0.19	1/5846 (0.0%)	0.46	0/7876
3	L	0.17	0/5909	0.44	0/7963
All	All	0.17	1/69045 (0.0%)	0.43	0/93083

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	567	ASP	C-O	5.64	1.26	1.23

There are no bond angle outliers.

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	5796	0	5734	216	0
1	D	5796	0	5734	177	0
1	G	5796	0	5734	210	0
1	J	5751	0	5689	223	0
2	B	5474	0	5423	216	0
2	E	5283	0	5219	225	0
2	H	5389	0	5338	226	0
2	K	5290	0	5246	236	0
3	C	5790	0	5939	210	0
3	F	5824	0	5969	187	1
3	I	5755	0	5903	218	1
3	L	5816	0	5963	221	0
All	All	67760	0	67891	2293	1

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 17.

The worst 5 of 2293 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)
3:I:318:ARG:HD3	3:I:493:ARG:HB3	1.54	0.90
1:G:321:ASN:HB3	1:G:544:GLU:HB3	1.52	0.89
1:G:68:PRO:HB3	1:G:72:LEU:HG	1.56	0.87
1:J:7:GLN:NE2	1:J:8:CYS:SG	2.48	0.87
2:H:485:ASN:ND2	2:H:489:THR:O	2.08	0.86

All (1) 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)
3:F:415:ARG:NH2	3:I:448:ASN:OD1[4_545]	2.19	0.01

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	714/716 (100%)	654 (92%)	60 (8%)	0	100	100
1	D	714/716 (100%)	652 (91%)	62 (9%)	0	100	100
1	G	714/716 (100%)	648 (91%)	66 (9%)	0	100	100
1	J	706/716 (99%)	647 (92%)	57 (8%)	2 (0%)	36	64
2	B	671/757 (89%)	611 (91%)	60 (9%)	0	100	100
2	E	647/757 (86%)	594 (92%)	53 (8%)	0	100	100
2	H	659/757 (87%)	599 (91%)	59 (9%)	1 (0%)	43	71
2	K	648/757 (86%)	600 (93%)	48 (7%)	0	100	100
3	C	719/765 (94%)	649 (90%)	69 (10%)	1 (0%)	48	78
3	F	723/765 (94%)	645 (89%)	77 (11%)	1 (0%)	48	78
3	I	712/765 (93%)	643 (90%)	69 (10%)	0	100	100
3	L	722/765 (94%)	652 (90%)	68 (9%)	2 (0%)	36	64
All	All	8349/8952 (93%)	7594 (91%)	748 (9%)	7 (0%)	48	78

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	536	MET
3	F	535	MET
2	H	429	LYS
3	L	535	MET
3	L	745	THR

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	640/640 (100%)	640 (100%)	0	100	100
1	D	640/640 (100%)	640 (100%)	0	100	100
1	G	640/640 (100%)	640 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	635/640 (99%)	635 (100%)	0	100	100
2	B	605/668 (91%)	604 (100%)	1 (0%)	87	84
2	E	587/668 (88%)	587 (100%)	0	100	100
2	H	597/668 (89%)	597 (100%)	0	100	100
2	K	585/668 (88%)	585 (100%)	0	100	100
3	C	643/674 (95%)	642 (100%)	1 (0%)	87	84
3	F	647/674 (96%)	646 (100%)	1 (0%)	87	84
3	I	639/674 (95%)	639 (100%)	0	100	100
3	L	646/674 (96%)	646 (100%)	0	100	100
All	All	7504/7928 (95%)	7501 (100%)	3 (0%)	100	100

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	153	ASN
3	C	266	ILE
3	F	28	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 76 such sidechains are listed below:

Mol	Chain	Res	Type
3	I	236	GLN
2	K	562	HIS
3	I	300	GLN
1	J	69	ASN
3	L	632	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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [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	716/716 (100%)	0.39	20 (2%)	55 32	39, 83, 142, 203	0
1	D	716/716 (100%)	0.47	20 (2%)	55 32	56, 106, 163, 212	0
1	G	716/716 (100%)	0.33	13 (1%)	67 41	42, 82, 147, 205	0
1	J	710/716 (99%)	0.51	31 (4%)	39 23	59, 101, 145, 200	0
2	B	685/757 (90%)	0.51	31 (4%)	38 23	49, 85, 137, 215	0
2	E	661/757 (87%)	0.60	31 (4%)	36 22	64, 108, 152, 240	0
2	H	673/757 (88%)	0.44	21 (3%)	51 30	39, 88, 138, 185	0
2	K	662/757 (87%)	0.50	29 (4%)	39 23	44, 98, 142, 202	0
3	C	727/765 (95%)	0.51	28 (3%)	43 26	55, 103, 163, 227	0
3	F	731/765 (95%)	0.66	31 (4%)	40 24	64, 119, 176, 210	0
3	I	722/765 (94%)	0.56	34 (4%)	36 22	46, 99, 151, 179	0
3	L	730/765 (95%)	0.49	29 (3%)	42 25	55, 100, 160, 232	0
All	All	8449/8952 (94%)	0.50	318 (3%)	44 26	39, 97, 155, 240	0

The worst 5 of 318 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	456	ASN	5.3
1	G	126	GLU	5.2
2	E	665	SER	4.9
2	K	261	SER	4.8
2	E	410	GLY	4.4

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.