



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 12, 2020 – 11:59 pm BST

PDB ID : 6QNW  
Title : Influenza A Polymerase Heterotrimer Human H3N2 Northern Territory 1968  
Authors : Keown, J.R.; Fan, H.; Grimes, J.M.  
Deposited on : 2019-02-12  
Resolution : 3.31 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

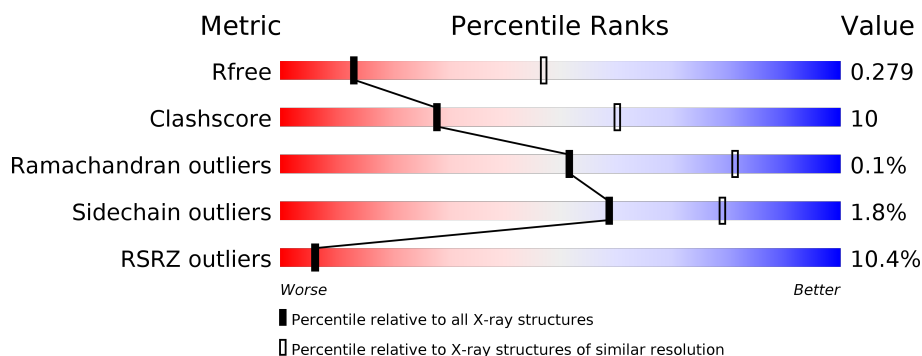
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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	716	<div> <div>3%</div> <div>72%</div> <div>25%</div> <div>..</div> </div>
1	D	716	<div> <div>17%</div> <div>73%</div> <div>21%</div> <div>..</div> </div>
1	G	716	<div> <div>2%</div> <div>72%</div> <div>25%</div> <div>..</div> </div>
1	J	716	<div> <div>10%</div> <div>75%</div> <div>22%</div> <div>..</div> </div>
2	B	757	<div> <div>4%</div> <div>70%</div> <div>20%</div> <div>10%</div> </div>
2	E	757	<div> <div>23%</div> <div>66%</div> <div>21%</div> <div>12%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	757	<p>4% 67% 21% • 11%</p>
2	K	757	<p>11% 65% 23% • 11%</p>
3	C	765	<p>10% 69% 24% • 5%</p>
3	F	765	<p>15% 70% 23% • 6%</p>
3	I	765	<p>6% 69% 26% 5%</p>
3	L	765	<p>11% 71% 22% 7%</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 67232 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	704	Total	C	N	O	S	0	0	0
			5733	3638	969	1084	42			
1	D	686	Total	C	N	O	S	0	0	0
			5604	3556	949	1058	41			
1	G	701	Total	C	N	O	S	0	0	0
			5707	3617	967	1081	42			
1	J	702	Total	C	N	O	S	0	0	0
			5713	3622	966	1083	42			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	679	Total	C	N	O	S	0	0	0
			5426	3417	932	1034	43			
2	E	665	Total	C	N	O	S	0	0	0
			5315	3351	910	1011	43			
2	H	676	Total	C	N	O	S	0	0	0
			5410	3410	929	1028	43			
2	K	673	Total	C	N	O	S	0	0	0
			5382	3390	925	1024	43			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	577	LYS	GLU	conflict	UNP P03432
E	577	LYS	GLU	conflict	UNP P03432
H	577	LYS	GLU	conflict	UNP P03432
K	577	LYS	GLU	conflict	UNP P03432

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	724	Total	C	N	O	S	0	0	0
			5772	3627	1044	1060	41			
3	F	720	Total	C	N	O	S	0	0	0
			5734	3604	1034	1056	40			
3	I	726	Total	C	N	O	S	0	0	0
			5789	3638	1047	1063	41			
3	L	709	Total	C	N	O	S	0	0	0
			5647	3552	1015	1039	41			

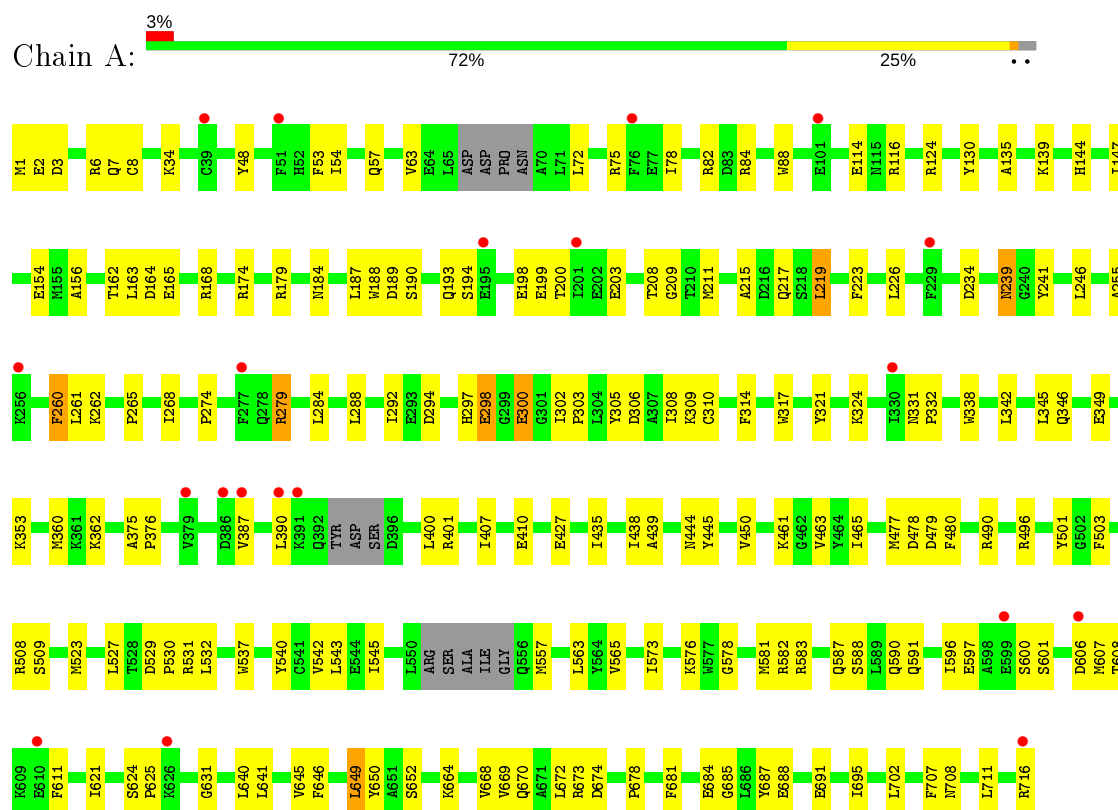
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	760	GLU	-	expression tag	UNP P03429
C	761	ASN	-	expression tag	UNP P03429
C	762	LEU	-	expression tag	UNP P03429
C	763	TYR	-	expression tag	UNP P03429
C	764	PHE	-	expression tag	UNP P03429
C	765	GLN	-	expression tag	UNP P03429
F	760	GLU	-	expression tag	UNP P03429
F	761	ASN	-	expression tag	UNP P03429
F	762	LEU	-	expression tag	UNP P03429
F	763	TYR	-	expression tag	UNP P03429
F	764	PHE	-	expression tag	UNP P03429
F	765	GLN	-	expression tag	UNP P03429
I	760	GLU	-	expression tag	UNP P03429
I	761	ASN	-	expression tag	UNP P03429
I	762	LEU	-	expression tag	UNP P03429
I	763	TYR	-	expression tag	UNP P03429
I	764	PHE	-	expression tag	UNP P03429
I	765	GLN	-	expression tag	UNP P03429
L	760	GLU	-	expression tag	UNP P03429
L	761	ASN	-	expression tag	UNP P03429
L	762	LEU	-	expression tag	UNP P03429
L	763	TYR	-	expression tag	UNP P03429
L	764	PHE	-	expression tag	UNP P03429
L	765	GLN	-	expression tag	UNP P03429

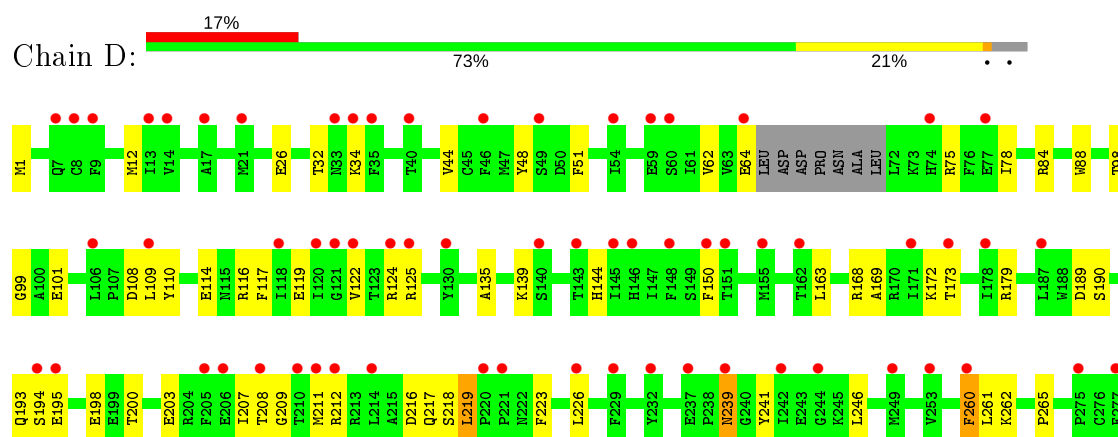
### 3 Residue-property plots

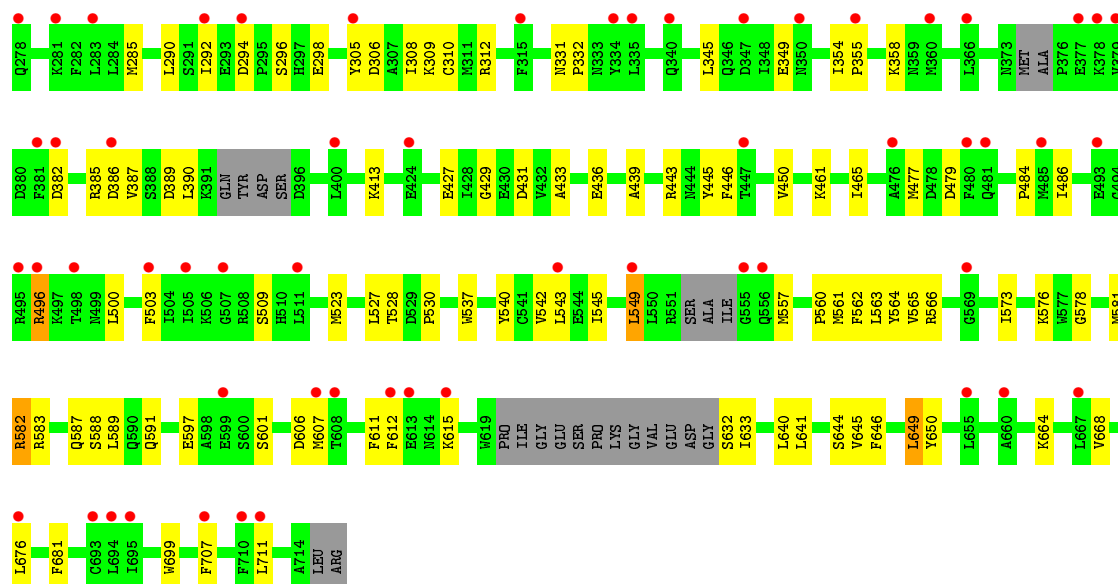
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

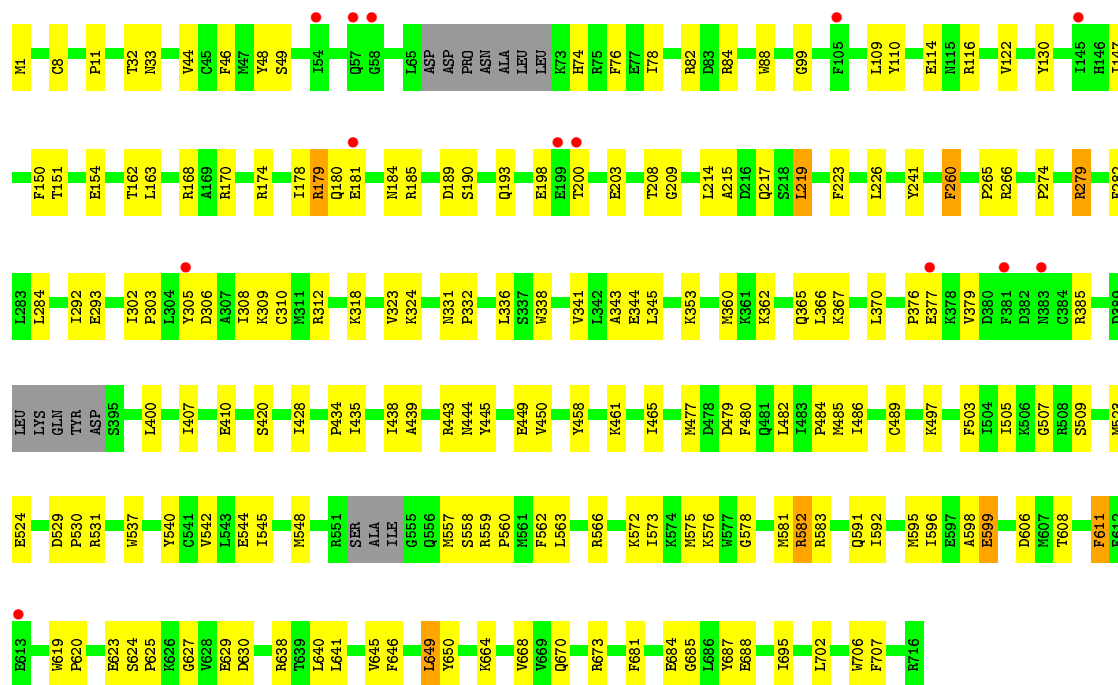


#### • Molecule 1: Polymerase acidic protein

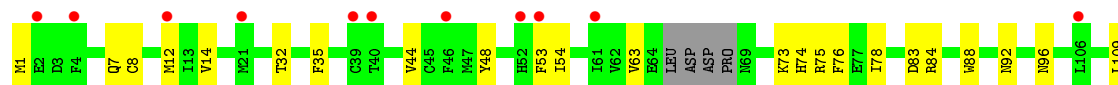
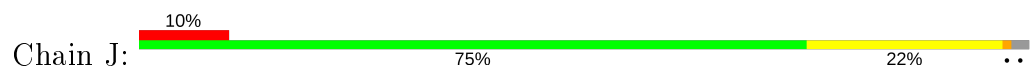


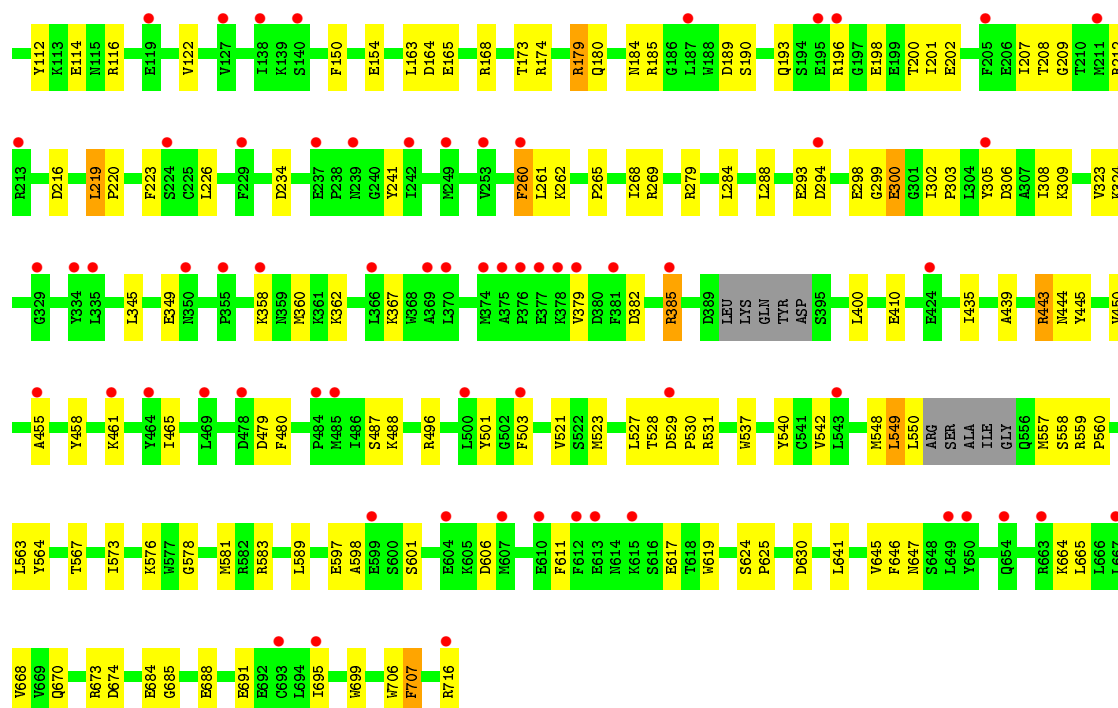


- Molecule 1: Polymerase acidic protein

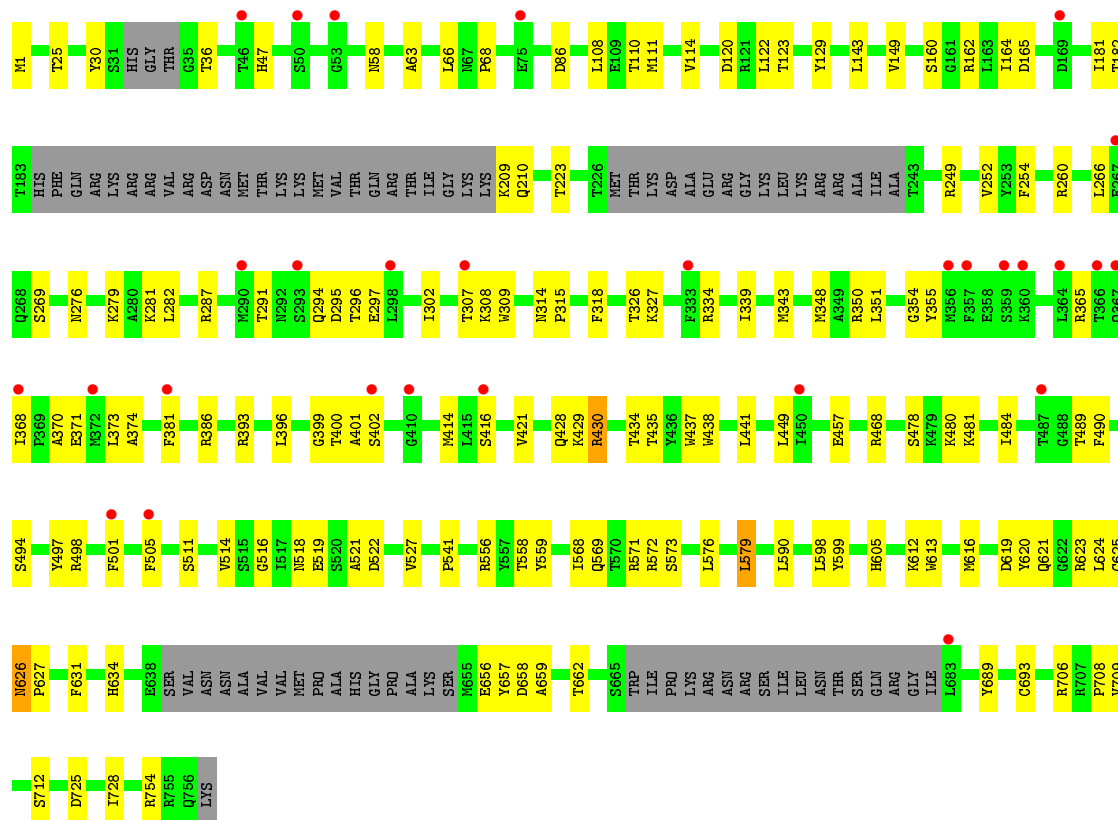


- Molecule 1: Polymerase acidic protein



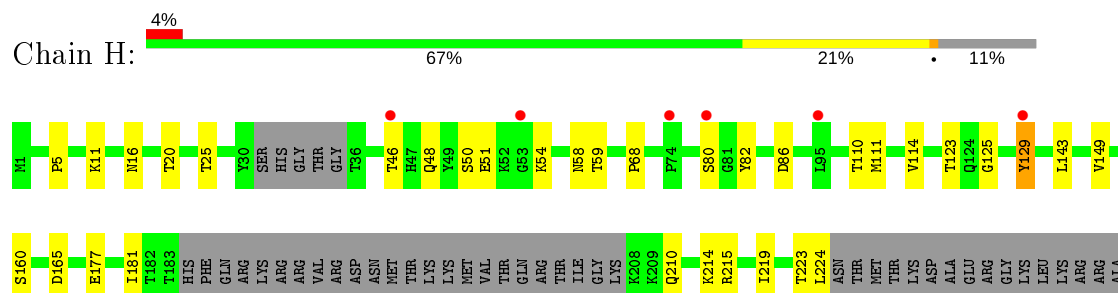


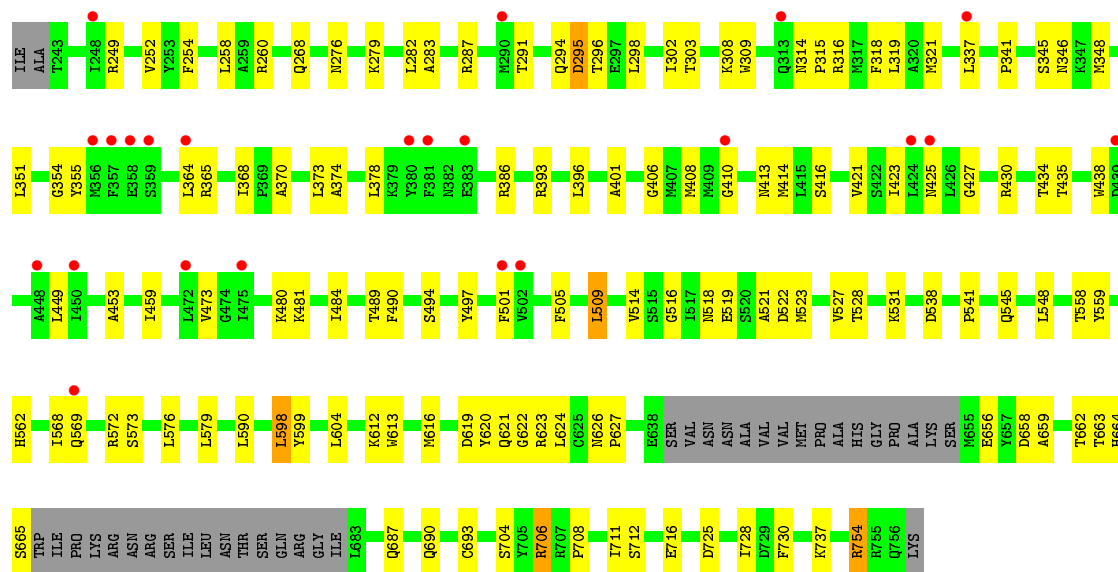
• Molecule 2: RNA-directed RNA polymerase catalytic subunit



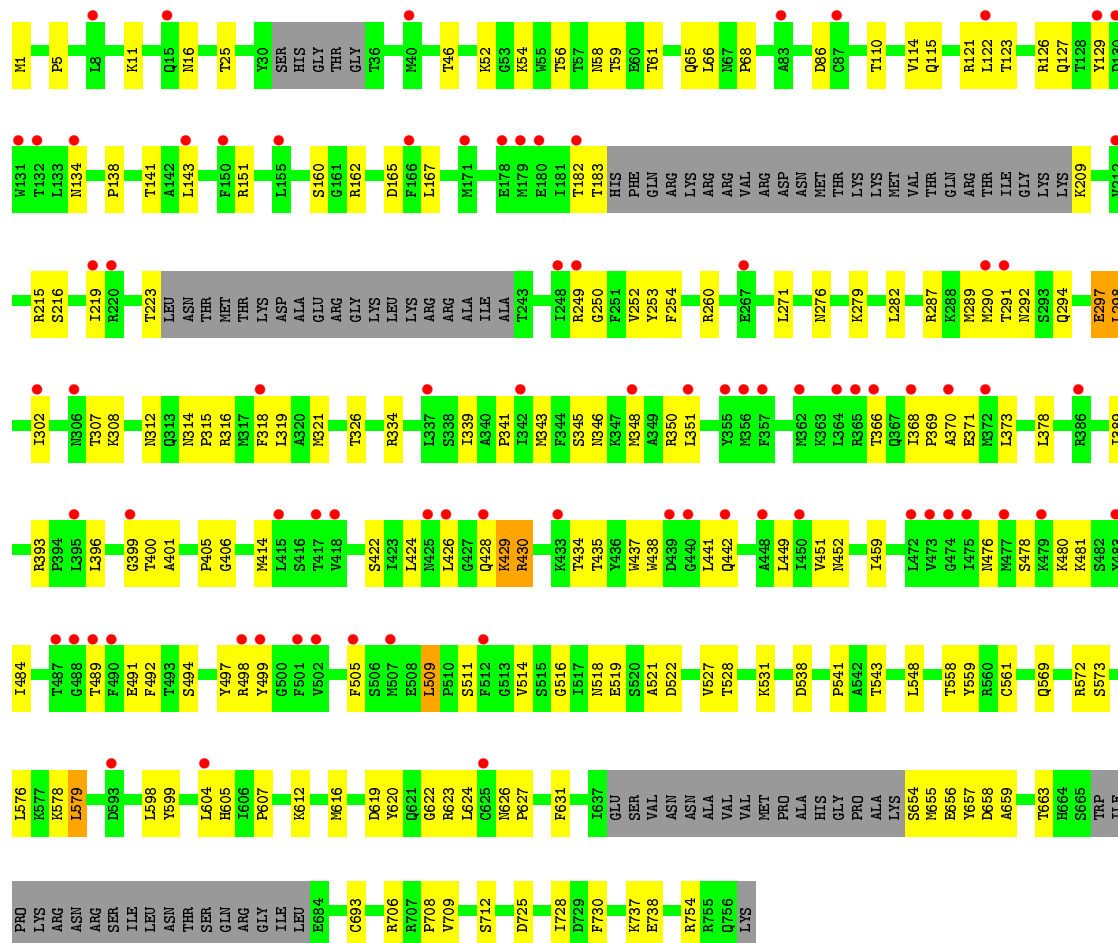
• Molecule 2: RNA-directed RNA polymerase catalytic subunit



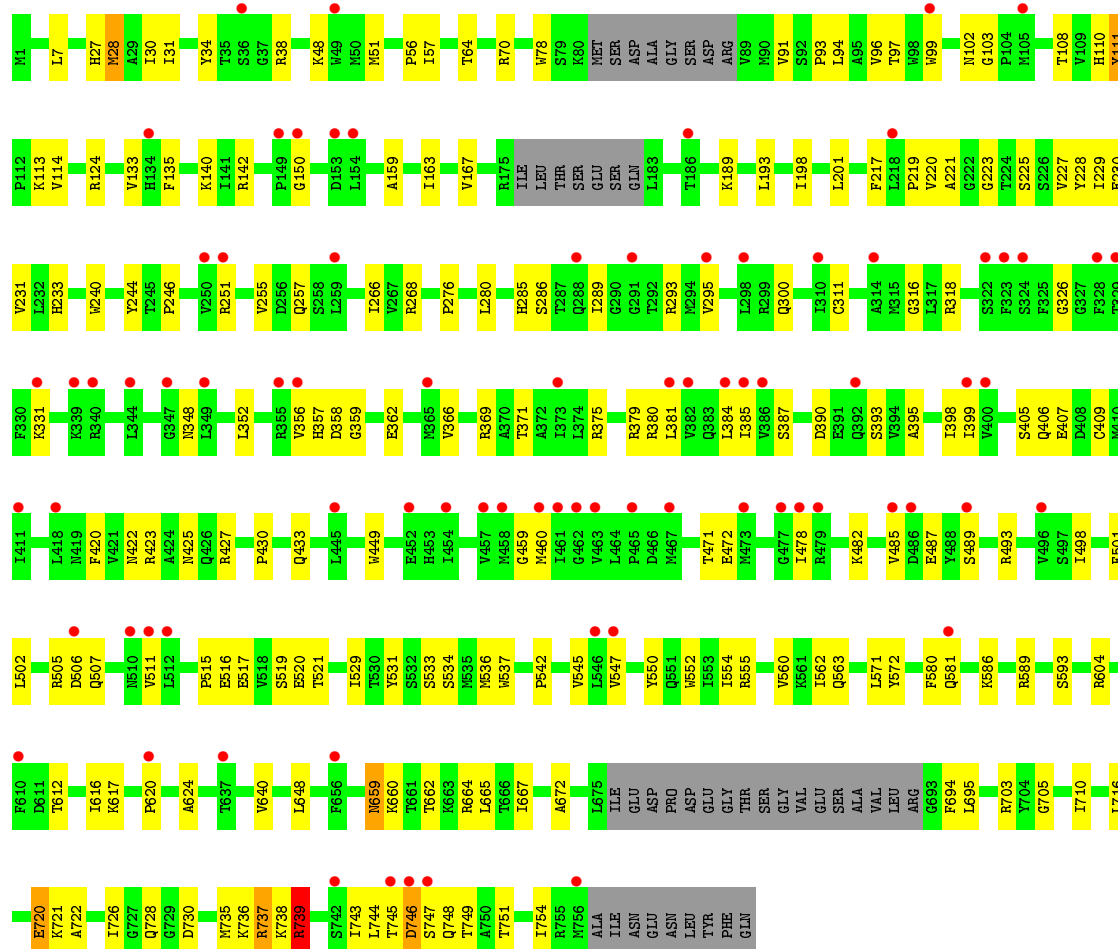


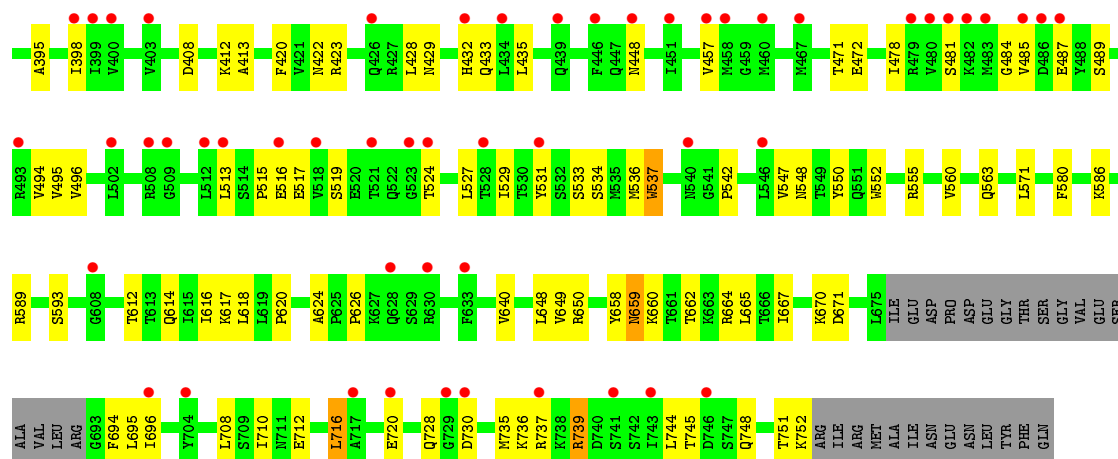


• Molecule 2: RNA-directed RNA polymerase catalytic subunit

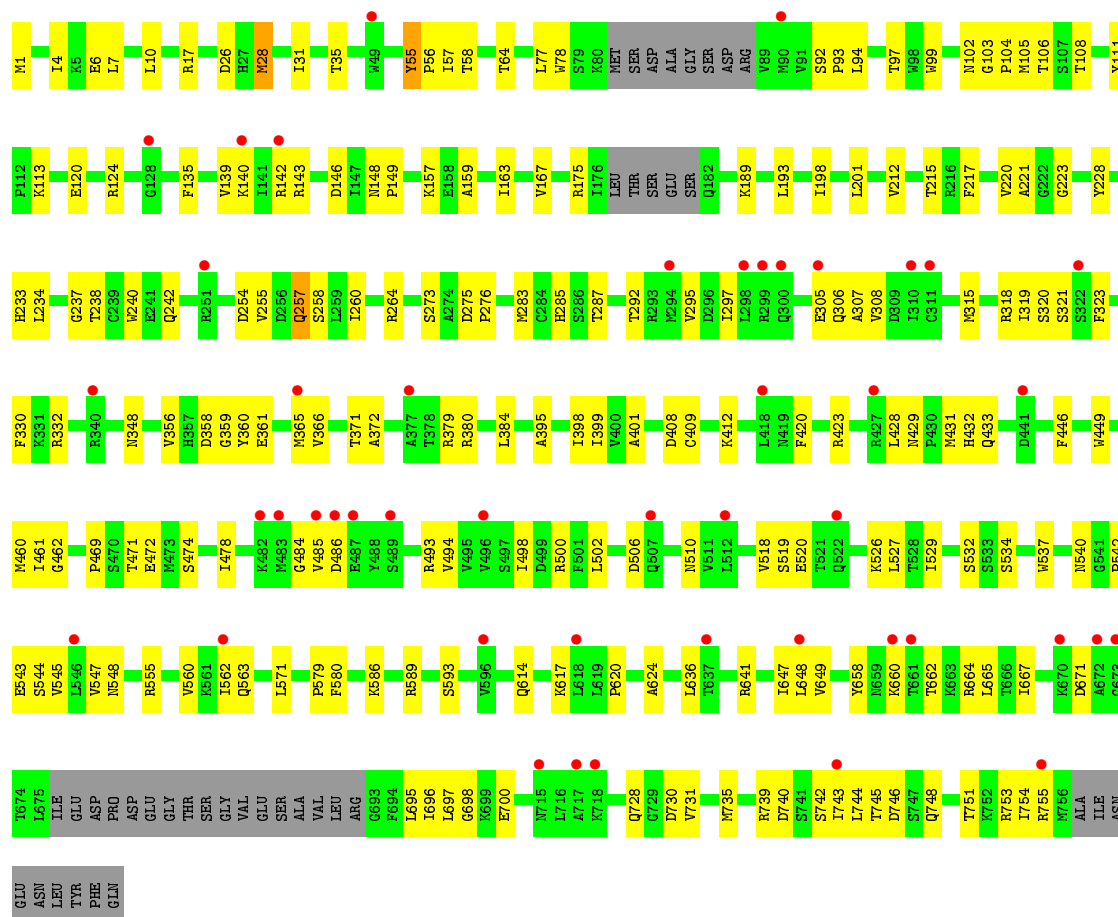


• Molecule 3: Polymerase basic protein 2





• Molecule 3: Polymerase basic protein 2



• Molecule 3: Polymerase basic protein 2



M1	K113	G223	T329	F438	Q523	R657	W756
I4	K121	Y228	F330	Q439	L527	Y658	ALA
R17	R124	Q236	K331	A442	T528	R659	ILE
E18	R134	T237	S336	F446	I529	R660	ASN
H27	F135	T238	L344	W449	S532	L665	GLU
M28	R136	C239	G347	H453	S533	A672	ASN
A29	N137	W240	N348	V457	S534	L675	LEU
I30	E241	Q242	L349	M458	M535	ILE	TTR
I31	K140	W243	L352	G459	W537	GLU	PHE
Q38	T141	G247	H357	M460	P542	ASP	GLN
N42	R142	GLY	D358	V463	E543	PRQ	
N42	R144	GLU	G359	V463	S544	ASP	
K48	G150	VAL	E362	T471	V547	GLY	
W49	D153	W252	M365	M473	Y550	SER	
A52	L154	D253	V366	S474	I554	GLY	
P56	K157	D254	V366	M475	R555	THR	
T64	E158	I260	T371	I478	V560	SER	
E72	A159	I261	A372	R479	E561	GLY	
W78	I163	R268	I373	V480	I562	VAL	
S79	V167	R269	L374	V485	Q563	GLU	
LYS	A174	W283	L381	D486	Q566	SER	
MET	R175	T287	L384	E487	L571	ALA	
SER	ILE	T292	I385	Y488	M575	VAL	
ASP	THR	T292	V386	S489	E576	LEU	
ALA	SER	T292	S387	E492	F577	ARG	
GLY	GLU	L298	I399	V494	E578	G693	
SER	SER	R299	D408	V495	P579	F694	
ASP	SER	Q300	C409	V496	Y592	L695	
ARG	Q182	R301	K412	S497	S593	L696	
V89	L183	P302	L418	I498	K617	L697	
P93	K189	E304	ASN	L502	L618	I710	
L94	L193	A307	PHE	R505	L619	R715	
A95	I198	D309	VAL	D506	P620	L716	
V96	W98	I310	ASN	Q507	A623	L717	
W98	W99	R318	ARG	R508	A624	E720	
N100	R101	I319	ALA	V511	S629	I726	
N101	N102	S320	ASN	L512	V640	G727	
G103	P104	W212	GLN	L513	L648	Q728	
M105	M105	S321	ARG	P515	E516	G729	
T108	T108	S322	LEU	E517	V649	D730	
V109	V109	F323	ASN	V518	R650	R735	
H110	Y111	F325	P430	S519	G651	R736	
P112	G222	G326	L435	H437	F656	R737	

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	336.48Å 191.90Å 235.74Å 90.00° 91.50° 90.00°	Depositor
Resolution (Å)	138.62 – 3.31 138.62 – 3.31	Depositor EDS
% Data completeness (in resolution range)	55.1 (138.62-3.31) 55.1 (138.62-3.31)	Depositor EDS
$R_{merge}$	0.46	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.235 , 0.279 0.235 , 0.279	Depositor DCC
$R_{free}$ test set	6172 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	88.5	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 76.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.025 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.024 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 0.024 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.029 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	67232	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.07% 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 ⓘ

### 5.1 Standard geometry ⓘ

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	0.25	0/5853	0.46	0/7886
1	D	0.25	0/5720	0.44	0/7701
1	G	0.25	0/5827	0.45	0/7850
1	J	0.25	0/5833	0.45	0/7860
2	B	0.25	0/5529	0.45	1/7456 (0.0%)
2	E	0.25	0/5417	0.43	0/7306
2	H	0.25	0/5513	0.44	0/7433
2	K	0.25	0/5485	0.44	0/7396
3	C	0.24	0/5866	0.47	0/7900
3	F	0.24	0/5828	0.47	1/7851 (0.0%)
3	I	0.24	0/5883	0.46	0/7923
3	L	0.24	0/5738	0.46	0/7725
All	All	0.25	0/68492	0.45	2/92287 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
3	C	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	579	LEU	CB-CG-CD2	-9.30	95.18	111.00
3	F	716	LEU	CA-CB-CG	6.75	130.82	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	297	HIS	Peptide
3	C	739	ARG	Peptide

## 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	5733	0	5662	123	0
1	D	5604	0	5527	120	0
1	G	5707	0	5624	134	0
1	J	5713	0	5630	116	0
2	B	5426	0	5374	109	0
2	E	5315	0	5262	129	0
2	H	5410	0	5366	138	0
2	K	5382	0	5330	136	0
3	C	5772	0	5915	141	0
3	F	5734	0	5869	138	0
3	I	5789	0	5934	152	0
3	L	5647	0	5786	120	0
All	All	67232	0	67279	1334	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 10.

The worst 5 of 1334 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:F:317:LEU:HD12	3:F:496:VAL:H	1.41	0.86
3:F:289:ILE:HG22	3:F:529:ILE:HG13	1.63	0.81
2:E:708:PRO:HA	3:F:728:GLN:HE21	1.46	0.79
3:F:149:PRO:HD2	3:F:209:ARG:HH12	1.49	0.77
2:B:309:TRP:HZ2	2:B:416:SER:HB3	1.49	0.77

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	696/716 (97%)	646 (93%)	49 (7%)	1 (0%)	51	81
1	D	674/716 (94%)	632 (94%)	42 (6%)	0	100	100
1	G	693/716 (97%)	644 (93%)	49 (7%)	0	100	100
1	J	694/716 (97%)	646 (93%)	48 (7%)	0	100	100
2	B	667/757 (88%)	628 (94%)	38 (6%)	1 (0%)	51	81
2	E	653/757 (86%)	621 (95%)	32 (5%)	0	100	100
2	H	664/757 (88%)	629 (95%)	34 (5%)	1 (0%)	47	76
2	K	661/757 (87%)	623 (94%)	37 (6%)	1 (0%)	47	76
3	C	716/765 (94%)	658 (92%)	56 (8%)	2 (0%)	41	71
3	F	712/765 (93%)	660 (93%)	52 (7%)	0	100	100
3	I	718/765 (94%)	658 (92%)	59 (8%)	1 (0%)	51	81
3	L	697/765 (91%)	647 (93%)	49 (7%)	1 (0%)	51	81
All	All	8245/8952 (92%)	7692 (93%)	545 (7%)	8 (0%)	51	81

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	298	GLU
2	B	295	ASP
2	K	429	LYS
3	I	745	THR
3	L	516	GLU

### 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	634/644 (98%)	623 (98%)	11 (2%)	60	79
1	D	620/644 (96%)	608 (98%)	12 (2%)	57	78
1	G	631/644 (98%)	622 (99%)	9 (1%)	67	82
1	J	632/644 (98%)	620 (98%)	12 (2%)	57	78
2	B	602/669 (90%)	593 (98%)	9 (2%)	65	81
2	E	589/669 (88%)	576 (98%)	13 (2%)	52	76
2	H	600/669 (90%)	590 (98%)	10 (2%)	60	79
2	K	597/669 (89%)	588 (98%)	9 (2%)	65	81
3	C	641/676 (95%)	631 (98%)	10 (2%)	62	80
3	F	637/676 (94%)	619 (97%)	18 (3%)	43	71
3	I	643/676 (95%)	635 (99%)	8 (1%)	71	84
3	L	628/676 (93%)	617 (98%)	11 (2%)	59	79
All	All	7454/7956 (94%)	7322 (98%)	132 (2%)	59	79

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	111	TYR
1	G	179	ARG
3	L	17	ARG
3	F	209	ARG
3	F	650	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	217	GLN
2	H	65	GLN
1	J	217	GLN
1	G	146	HIS
2	K	292	ASN

### 5.3.3 RNA ⓘ

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](#)

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, 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	704/716 (98%)	0.71	20 (2%) 53 51	36, 86, 152, 183	0
1	D	686/716 (95%)	1.11	121 (17%) 1 1	89, 142, 214, 261	0
1	G	701/716 (97%)	0.59	13 (1%) 66 65	39, 86, 147, 214	0
1	J	702/716 (98%)	0.87	75 (10%) 6 5	67, 121, 184, 240	0
2	B	679/757 (89%)	0.77	29 (4%) 35 35	35, 82, 137, 210	0
2	E	665/757 (87%)	1.40	177 (26%) 0 0	93, 164, 215, 239	0
2	H	676/757 (89%)	0.73	29 (4%) 35 35	35, 89, 157, 214	0
2	K	673/757 (88%)	0.92	81 (12%) 4 3	65, 124, 184, 221	0
3	C	724/765 (94%)	0.89	80 (11%) 5 5	46, 110, 167, 217	0
3	F	720/765 (94%)	0.98	116 (16%) 1 1	55, 134, 185, 220	0
3	I	726/765 (94%)	0.70	46 (6%) 20 21	42, 107, 163, 200	0
3	L	709/765 (92%)	0.89	83 (11%) 4 3	70, 123, 173, 215	0
All	All	8365/8952 (93%)	0.88	870 (10%) 6 6	35, 114, 186, 261	0

The worst 5 of 870 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	440	GLY	12.8
3	L	717	ALA	11.4
3	C	250	VAL	9.8
3	F	140	LYS	9.7
2	E	439	ASP	8.3

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.