



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 12:31 PM UTC

PDB ID : 5D98 / pdb_00005d98
Title : Influenza C Virus RNA-dependent RNA Polymerase - Space group P43212
Authors : Hengrung, N.; El Omari, K.; Serna Martin, I.; Vreede, F.T.; Cusack, S.;
Rambo, R.P.; Vonrhein, C.; Bricogne, G.; Stuart, D.I.; Grimes, J.M.; Fodor, E.
Deposited on : 2015-08-18
Resolution : 3.90 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:

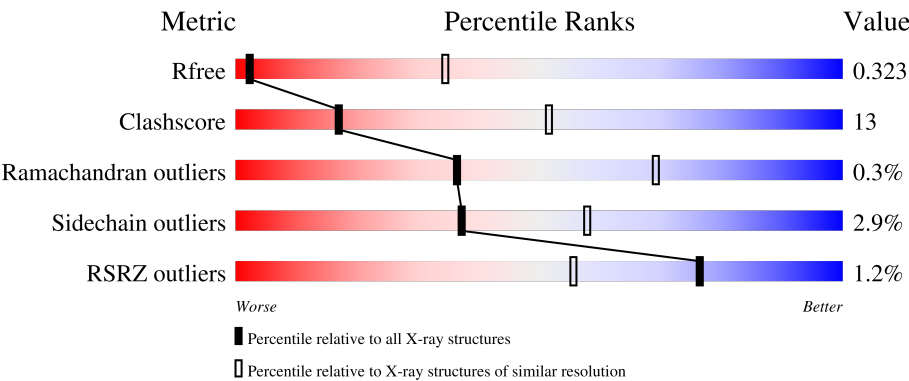
MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)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

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.90 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	1270 (4.10-3.70)
Clashscore	190562	1034 (4.08-3.72)
Ramachandran outliers	187476	1251 (4.10-3.70)
Sidechain outliers	187428	1243 (4.10-3.70)
RSRZ outliers	180081	1269 (4.10-3.70)

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	709	66%31%..
1	D	709	% 66%30%..
2	B	754	% 71%22%• 6%
2	E	754	% 70%22%• 6%
3	C	782	% 69%27%• •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	782	% 67% 29%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 34720 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	693	Total	C	N	O	S	0	0	0
			5630	3589	954	1043	44			
1	D	693	Total	C	N	O	S	0	0	0
			5630	3589	954	1043	44			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	711	Total	C	N	O	S	0	0	0
			5652	3587	956	1056	53			
2	E	711	Total	C	N	O	S	0	0	0
			5652	3587	956	1056	53			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	762	Total	C	N	O	S	0	0	0
			6076	3845	1066	1128	37			
3	F	762	Total	C	N	O	S	0	0	0
			6076	3845	1066	1128	37			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	775	ALA	-	expression tag	UNP Q9IMP3
C	776	ARG	-	expression tag	UNP Q9IMP3
C	777	GLU	-	expression tag	UNP Q9IMP3
C	778	ASN	-	expression tag	UNP Q9IMP3
C	779	LEU	-	expression tag	UNP Q9IMP3
C	780	TYR	-	expression tag	UNP Q9IMP3
C	781	PHE	-	expression tag	UNP Q9IMP3
C	782	GLN	-	expression tag	UNP Q9IMP3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	775	ALA	-	expression tag	UNP Q9IMP3
F	776	ARG	-	expression tag	UNP Q9IMP3
F	777	GLU	-	expression tag	UNP Q9IMP3
F	778	ASN	-	expression tag	UNP Q9IMP3
F	779	LEU	-	expression tag	UNP Q9IMP3
F	780	TYR	-	expression tag	UNP Q9IMP3
F	781	PHE	-	expression tag	UNP Q9IMP3
F	782	GLN	-	expression tag	UNP Q9IMP3

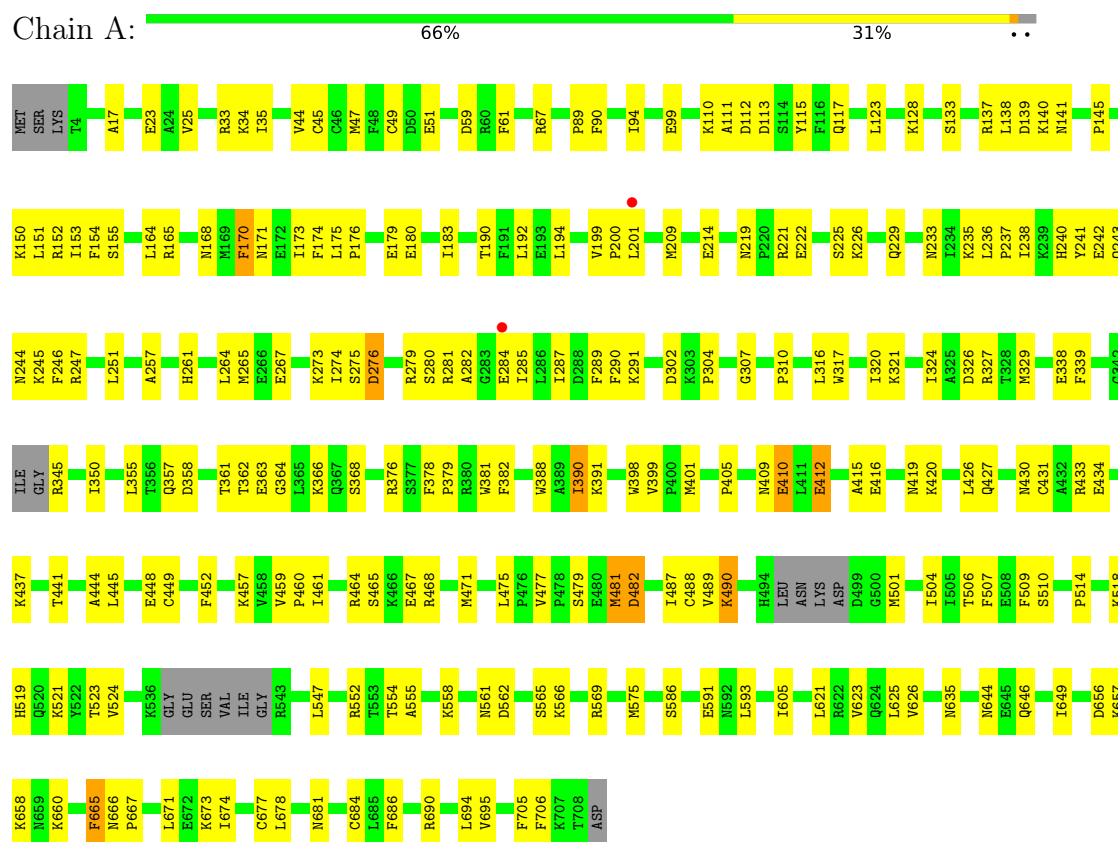
- Molecule 4 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Mg 2 2	0	0
4	D	2	Total Mg 2 2	0	0

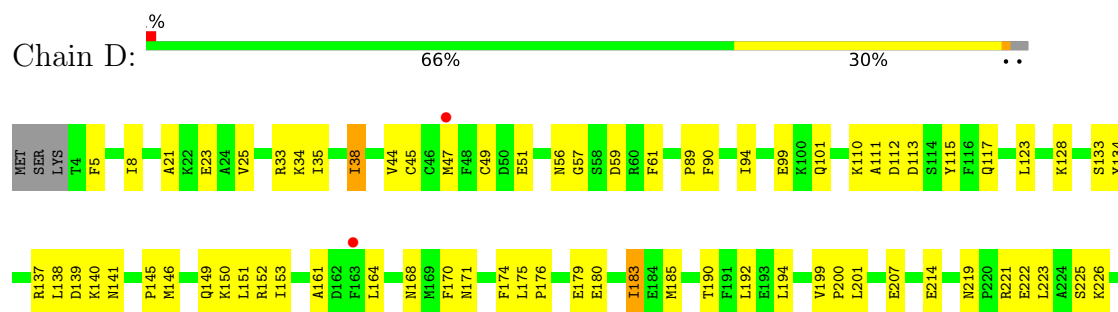
3 Residue-property plots

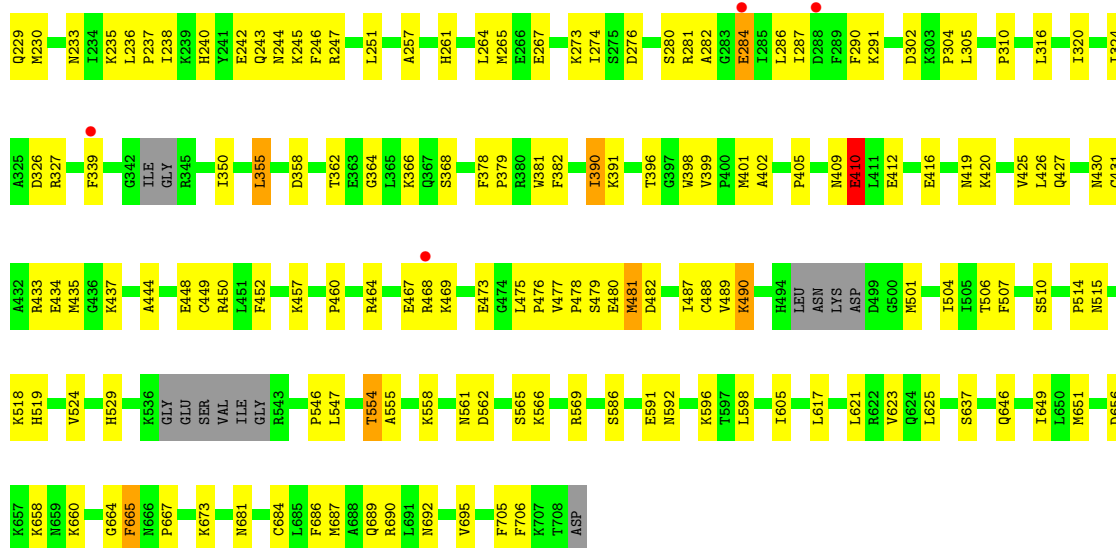
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

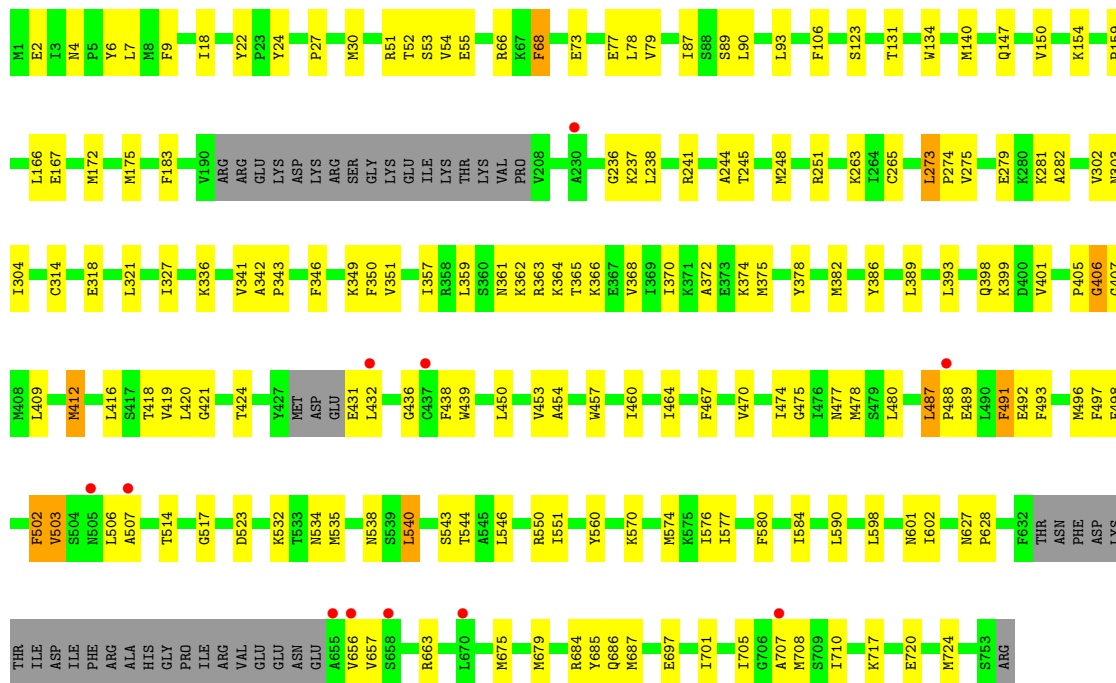


• Molecule 1: Polymerase acidic protein

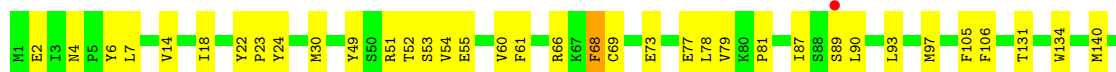


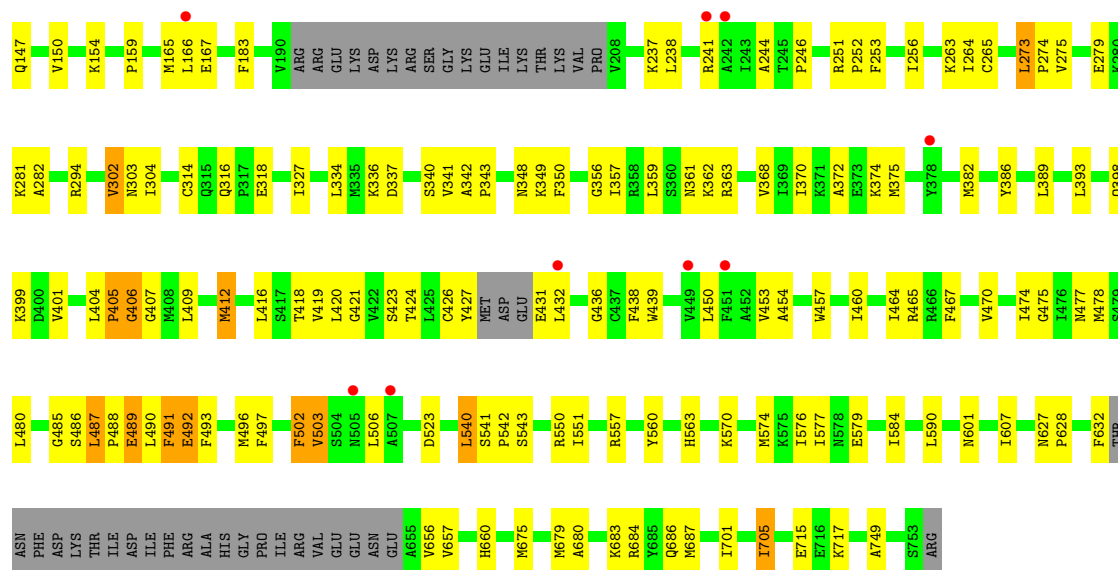


• Molecule 2: RNA-directed RNA polymerase catalytic subunit

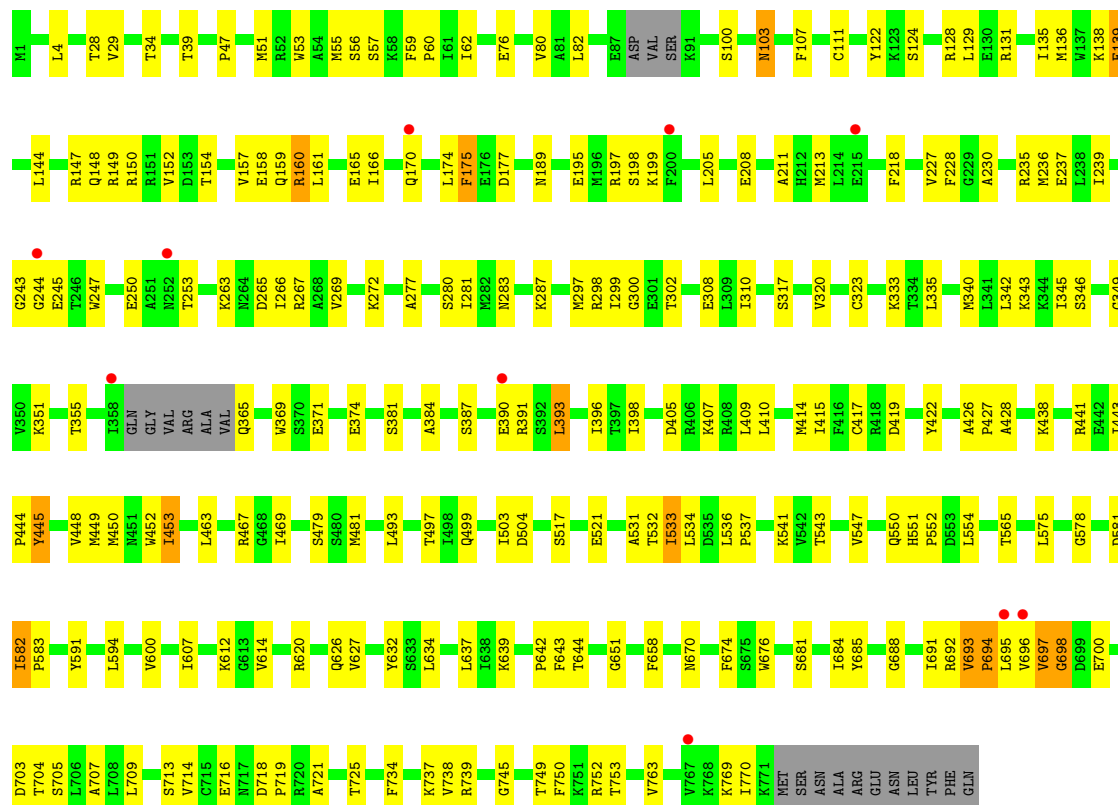


• Molecule 2: RNA-directed RNA polymerase catalytic subunit





• Molecule 3: Polymerase basic protein 2



• Molecule 3: Polymerase basic protein 2



LEU	I684	T543	D419	L331	R235	H1
TTR	Y685	V547	Y422	G332	M236	L4
PHE	G688	Q550	A426	K333	E237	Y11
GLN	I691	H551	A427	P339	I239	Q16
	R692	P552	A428	M340	E245	Q16
	V693	D553	T429	K343	T246	T28
	P694	L554	K438	S346	W247	V29
	V696	T565	R441	G349	I248	
		L575	P444	V350	Q249	K37
	D703	G578	Y445	K351	E250	W38
	T704	I582	V448	T355	T253	T39
	S713	P583	M449	G355	A254	L49
	V714	M450	M450	I358	K263	L50
	E716	M451	M451	GLN	W264	M51
	W717	I453	I453	GLY	D265	R52
	D718	I586	L463	VAL	I266	W53
	P719	Y591	L463	ARG	R267	A54
	R720	L594	S466	ALA	A268	M55
	A721	Y595	R467	VAL	V269	S56
		V600	I469	Q365	C270	S57
	T725	I607	T472	S381	R271	K58
	F734	K612	S479	E374	K272	F59
	K737	G613	M481	S370	C273	P60
	V738	V614	S492	E371	C274	E76
	R739	R620	L493	A384	I275	H77
	L740	Q626	T497	E390	A276	V80
	F741	V627	Q499	R391	A277	A81
	Q744	Y632	I503	S392	S280	L82
	V747	L634	D504	L393	M282	E87
	R748	L637	I512	E395	W283	ASP
	T749	P642	H513	I396	K287	VAL
	F750	F643	S517	T397	L286	SER
	K751	T644	E521	I398	I292	R91
	R752	Q651	A531	D405	M297	I102
	T753	F658	T532	R406	K298	M103
	A754	N670	I533	K407	I299	N106
	A758	W676	L534	R408	G300	C111
		S681	D536	L409	E308	V117
			L536	L309	I310	V121
			ASN	L410	M212	Y122
			ALA	A411	M213	K123
			ARG	M412	F218	S124
			GLU	C413	F228	R125
			ASN	M414	G229	F126
				I415	A230	G127
				F416	E234	R128
				C417		L129
				R418		E130

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	185.66Å 185.66Å 598.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 3.90 50.01 – 3.90	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.01-3.90) 98.8 (50.01-3.90)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.92 (at 3.89Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.286 , 0.326 0.286 , 0.323	Depositor DCC
R_{free} test set	4772 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	161.2	Xtriage
Anisotropy	0.644	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 156.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	34720	wwPDB-VP
Average B, all atoms (Å ²)	207.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.59% 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](#)

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	0.54	0/5746	0.88	4/7717 (0.1%)
1	D	0.56	0/5746	0.88	4/7717 (0.1%)
2	B	0.56	0/5749	0.91	5/7723 (0.1%)
2	E	0.57	0/5749	0.89	6/7723 (0.1%)
3	C	0.56	0/6185	0.89	6/8322 (0.1%)
3	F	0.57	0/6185	0.88	4/8322 (0.0%)
All	All	0.56	0/35360	0.89	29/47524 (0.1%)

There are no bond length outliers.

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	406	GLY	N-CA-C	7.22	121.40	112.73
1	A	481	MET	N-CA-C	6.39	117.01	108.38
3	F	245	GLU	N-CA-C	6.37	119.04	111.71
1	D	284	GLU	N-CA-C	-6.17	105.79	113.38
2	E	489	GLU	N-CA-C	6.16	117.99	111.28

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	5630	0	5632	155	0
1	D	5630	0	5632	175	0
2	B	5652	0	5749	141	0
2	E	5652	0	5749	140	0
3	C	6076	0	6183	190	0
3	F	6076	0	6183	205	0
4	A	2	0	0	0	0
4	D	2	0	0	0	0
All	All	34720	0	35128	888	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 13.

The worst 5 of 888 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:138:LYS:HE2	3:F:250:GLU:CG	1.22	1.59
3:F:138:LYS:CE	3:F:250:GLU:HG2	1.34	1.48
3:F:138:LYS:CE	3:F:250:GLU:CD	1.86	1.46
3:F:138:LYS:HE3	3:F:250:GLU:CD	1.42	1.38
3:F:138:LYS:CE	3:F:250:GLU:CG	1.83	1.36

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	685/709 (97%)	643 (94%)	42 (6%)	0	100	100
1	D	685/709 (97%)	645 (94%)	40 (6%)	0	100	100
2	B	703/754 (93%)	666 (95%)	35 (5%)	2 (0%)	36	69
2	E	703/754 (93%)	673 (96%)	27 (4%)	3 (0%)	30	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	756/782 (97%)	688 (91%)	63 (8%)	5 (1%)	18	53
3	F	756/782 (97%)	686 (91%)	67 (9%)	3 (0%)	30	64
All	All	4288/4490 (96%)	4001 (93%)	274 (6%)	13 (0%)	36	69

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	405	PRO
3	C	698	GLY
2	E	405	PRO
2	B	503	VAL
2	E	503	VAL

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	618/631 (98%)	599 (97%)	19 (3%)	35	57
1	D	618/631 (98%)	600 (97%)	18 (3%)	37	58
2	B	629/669 (94%)	610 (97%)	19 (3%)	36	57
2	E	629/669 (94%)	609 (97%)	20 (3%)	34	56
3	C	669/686 (98%)	653 (98%)	16 (2%)	43	63
3	F	669/686 (98%)	650 (97%)	19 (3%)	38	59
All	All	3832/3972 (96%)	3721 (97%)	111 (3%)	37	58

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	183	ILE
3	F	738	VAL
1	D	673	LYS
3	F	725	THR
3	F	270	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	141	ASN
3	F	403	ASN
1	D	519	HIS
3	F	212	HIS
2	E	660	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are 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 [i](#)

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	693/709 (97%)	-0.28	2 (0%) 90 78	132, 195, 269, 299	0
1	D	693/709 (97%)	-0.19	6 (0%) 81 62	129, 191, 258, 295	0
2	B	711/754 (94%)	-0.28	11 (1%) 72 50	131, 192, 256, 317	0
2	E	711/754 (94%)	-0.20	10 (1%) 73 52	130, 184, 266, 315	0
3	C	762/782 (97%)	-0.06	10 (1%) 75 54	149, 222, 285, 358	0
3	F	762/782 (97%)	0.03	11 (1%) 73 52	134, 224, 287, 343	0
All	All	4332/4490 (96%)	-0.16	50 (1%) 76 55	129, 202, 275, 358	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	695	LEU	4.9
3	C	696	VAL	4.2
2	E	242	ALA	3.9
2	E	166	LEU	3.4
3	F	586	ILE	3.2

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
4	MG	D	800	1/1	0.69	0.15	200,200,200,200	0
4	MG	A	800	1/1	0.71	0.11	204,204,204,204	0
4	MG	D	801	1/1	0.72	0.14	182,182,182,182	0
4	MG	A	801	1/1	0.92	0.06	185,185,185,185	0

6.5 Other polymers [i](#)

There are no such residues in this entry.