WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 21, 2024 – 12:03 am GMT

PDB ID	:	7Z42
Title	:	Influenza B polymerase with Pol II pSer5 CTD peptide mimic bound in site 2B
Authors	:	Cusack, S.; Drncova, P.
Deposited on	:	2022-03-03
Resolution	:	2.42 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

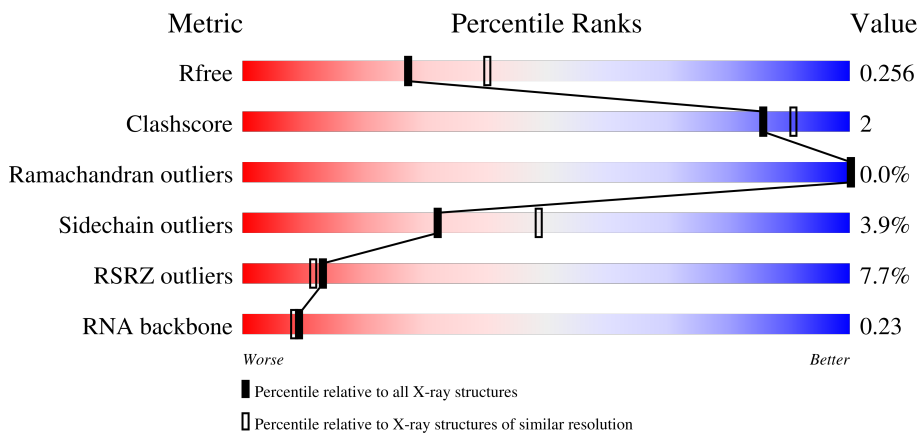|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.42 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	Percentile Ranks	Value
Rfree		0.256
Clashscore		2
Ramachandran outliers		0.0%
Sidechain outliers		3.9%
RSRZ outliers		7.7%
RNA backbone		0.23

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)
RNA backbone	3102	1080 (2.80-2.04)

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 <=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	751	5% / 88% / 7% ··
1	D	751	2% / 87% / 8% 5%
2	B	772	2% / 90% / 7% ··
2	E	772	% / 87% / 9% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	798	15% / 82% / 8% · 10%
3	F	798	18% / 82% / 8% · 9%
4	G	28	11% / 75% / 25%
4	I	28	4% / 64% / 36%
4	X	28	11% / 14% / 86%
4	Y	28	7% / 7% · 89%
5	H	13	46% / 46% / 8%
5	V	13	46% / 54%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 35949 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	721	Total	C	N	O	S	0	1	0
			5787	3675	970	1101	41			
1	D	715	Total	C	N	O	S	0	1	0
			5750	3652	962	1095	41			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	expression tag	UNP Q5V8Z9
A	-12	SER	-	expression tag	UNP Q5V8Z9
A	-11	HIS	-	expression tag	UNP Q5V8Z9
A	-10	HIS	-	expression tag	UNP Q5V8Z9
A	-9	HIS	-	expression tag	UNP Q5V8Z9
A	-8	HIS	-	expression tag	UNP Q5V8Z9
A	-7	HIS	-	expression tag	UNP Q5V8Z9
A	-6	HIS	-	expression tag	UNP Q5V8Z9
A	-5	HIS	-	expression tag	UNP Q5V8Z9
A	-4	HIS	-	expression tag	UNP Q5V8Z9
A	-3	GLY	-	expression tag	UNP Q5V8Z9
A	-2	SER	-	expression tag	UNP Q5V8Z9
A	-1	GLY	-	expression tag	UNP Q5V8Z9
A	0	SER	-	expression tag	UNP Q5V8Z9
A	727	GLY	-	expression tag	UNP Q5V8Z9
A	728	SER	-	expression tag	UNP Q5V8Z9
A	729	GLY	-	expression tag	UNP Q5V8Z9
A	730	SER	-	expression tag	UNP Q5V8Z9
A	731	GLY	-	expression tag	UNP Q5V8Z9
A	732	GLU	-	expression tag	UNP Q5V8Z9
A	733	ASN	-	expression tag	UNP Q5V8Z9
A	734	LEU	-	expression tag	UNP Q5V8Z9
A	735	TYR	-	expression tag	UNP Q5V8Z9
A	736	PHE	-	expression tag	UNP Q5V8Z9
A	737	GLN	-	expression tag	UNP Q5V8Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-13	GLY	-	expression tag	UNP Q5V8Z9
D	-12	SER	-	expression tag	UNP Q5V8Z9
D	-11	HIS	-	expression tag	UNP Q5V8Z9
D	-10	HIS	-	expression tag	UNP Q5V8Z9
D	-9	HIS	-	expression tag	UNP Q5V8Z9
D	-8	HIS	-	expression tag	UNP Q5V8Z9
D	-7	HIS	-	expression tag	UNP Q5V8Z9
D	-6	HIS	-	expression tag	UNP Q5V8Z9
D	-5	HIS	-	expression tag	UNP Q5V8Z9
D	-4	HIS	-	expression tag	UNP Q5V8Z9
D	-3	GLY	-	expression tag	UNP Q5V8Z9
D	-2	SER	-	expression tag	UNP Q5V8Z9
D	-1	GLY	-	expression tag	UNP Q5V8Z9
D	0	SER	-	expression tag	UNP Q5V8Z9
D	727	GLY	-	expression tag	UNP Q5V8Z9
D	728	SER	-	expression tag	UNP Q5V8Z9
D	729	GLY	-	expression tag	UNP Q5V8Z9
D	730	SER	-	expression tag	UNP Q5V8Z9
D	731	GLY	-	expression tag	UNP Q5V8Z9
D	732	GLU	-	expression tag	UNP Q5V8Z9
D	733	ASN	-	expression tag	UNP Q5V8Z9
D	734	LEU	-	expression tag	UNP Q5V8Z9
D	735	TYR	-	expression tag	UNP Q5V8Z9
D	736	PHE	-	expression tag	UNP Q5V8Z9
D	737	GLN	-	expression tag	UNP Q5V8Z9

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	754	Total 5911	C 3731	N 1024	O 1104	S 52	0	0	0
2	E	743	Total 5818	C 3671	N 1004	O 1091	S 52	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP Q5V8Y6
B	-7	SER	-	expression tag	UNP Q5V8Y6
B	-6	GLY	-	expression tag	UNP Q5V8Y6
B	-5	SER	-	expression tag	UNP Q5V8Y6
B	-4	GLY	-	expression tag	UNP Q5V8Y6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	SER	-	expression tag	UNP Q5V8Y6
B	-2	GLY	-	expression tag	UNP Q5V8Y6
B	-1	SER	-	expression tag	UNP Q5V8Y6
B	0	GLY	-	expression tag	UNP Q5V8Y6
B	753	GLY	-	expression tag	UNP Q5V8Y6
B	754	SER	-	expression tag	UNP Q5V8Y6
B	755	GLY	-	expression tag	UNP Q5V8Y6
B	756	SER	-	expression tag	UNP Q5V8Y6
B	757	GLY	-	expression tag	UNP Q5V8Y6
B	758	GLU	-	expression tag	UNP Q5V8Y6
B	759	ASN	-	expression tag	UNP Q5V8Y6
B	760	LEU	-	expression tag	UNP Q5V8Y6
B	761	TYR	-	expression tag	UNP Q5V8Y6
B	762	PHE	-	expression tag	UNP Q5V8Y6
B	763	GLN	-	expression tag	UNP Q5V8Y6
E	-8	GLY	-	expression tag	UNP Q5V8Y6
E	-7	SER	-	expression tag	UNP Q5V8Y6
E	-6	GLY	-	expression tag	UNP Q5V8Y6
E	-5	SER	-	expression tag	UNP Q5V8Y6
E	-4	GLY	-	expression tag	UNP Q5V8Y6
E	-3	SER	-	expression tag	UNP Q5V8Y6
E	-2	GLY	-	expression tag	UNP Q5V8Y6
E	-1	SER	-	expression tag	UNP Q5V8Y6
E	0	GLY	-	expression tag	UNP Q5V8Y6
E	753	GLY	-	expression tag	UNP Q5V8Y6
E	754	SER	-	expression tag	UNP Q5V8Y6
E	755	GLY	-	expression tag	UNP Q5V8Y6
E	756	SER	-	expression tag	UNP Q5V8Y6
E	757	GLY	-	expression tag	UNP Q5V8Y6
E	758	GLU	-	expression tag	UNP Q5V8Y6
E	759	ASN	-	expression tag	UNP Q5V8Y6
E	760	LEU	-	expression tag	UNP Q5V8Y6
E	761	TYR	-	expression tag	UNP Q5V8Y6
E	762	PHE	-	expression tag	UNP Q5V8Y6
E	763	GLN	-	expression tag	UNP Q5V8Y6

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	718	Total 5746	C 3654	N 1001	O 1051	S 40	0	0	0
3	F	728	Total 5838	C 3712	N 1023	O 1063	S 40	0	1	0

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLY	-	expression tag	UNP Q5V8X3
C	-7	SER	-	expression tag	UNP Q5V8X3
C	-6	GLY	-	expression tag	UNP Q5V8X3
C	-5	SER	-	expression tag	UNP Q5V8X3
C	-4	GLY	-	expression tag	UNP Q5V8X3
C	-3	SER	-	expression tag	UNP Q5V8X3
C	-2	GLY	-	expression tag	UNP Q5V8X3
C	-1	SER	-	expression tag	UNP Q5V8X3
C	0	GLY	-	expression tag	UNP Q5V8X3
C	771	GLY	-	expression tag	UNP Q5V8X3
C	772	TRP	-	expression tag	UNP Q5V8X3
C	773	SER	-	expression tag	UNP Q5V8X3
C	774	HIS	-	expression tag	UNP Q5V8X3
C	775	PRO	-	expression tag	UNP Q5V8X3
C	776	GLN	-	expression tag	UNP Q5V8X3
C	777	PHE	-	expression tag	UNP Q5V8X3
C	778	GLU	-	expression tag	UNP Q5V8X3
C	779	LYS	-	expression tag	UNP Q5V8X3
C	780	GLY	-	expression tag	UNP Q5V8X3
C	781	SER	-	expression tag	UNP Q5V8X3
C	782	GLY	-	expression tag	UNP Q5V8X3
C	783	SER	-	expression tag	UNP Q5V8X3
C	784	GLU	-	expression tag	UNP Q5V8X3
C	785	ASN	-	expression tag	UNP Q5V8X3
C	786	LEU	-	expression tag	UNP Q5V8X3
C	787	TYR	-	expression tag	UNP Q5V8X3
C	788	PHE	-	expression tag	UNP Q5V8X3
C	789	GLN	-	expression tag	UNP Q5V8X3
F	-8	GLY	-	expression tag	UNP Q5V8X3
F	-7	SER	-	expression tag	UNP Q5V8X3
F	-6	GLY	-	expression tag	UNP Q5V8X3
F	-5	SER	-	expression tag	UNP Q5V8X3
F	-4	GLY	-	expression tag	UNP Q5V8X3
F	-3	SER	-	expression tag	UNP Q5V8X3
F	-2	GLY	-	expression tag	UNP Q5V8X3
F	-1	SER	-	expression tag	UNP Q5V8X3
F	0	GLY	-	expression tag	UNP Q5V8X3
F	771	GLY	-	expression tag	UNP Q5V8X3
F	772	TRP	-	expression tag	UNP Q5V8X3
F	773	SER	-	expression tag	UNP Q5V8X3
F	774	HIS	-	expression tag	UNP Q5V8X3
F	775	PRO	-	expression tag	UNP Q5V8X3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	776	GLN	-	expression tag	UNP Q5V8X3
F	777	PHE	-	expression tag	UNP Q5V8X3
F	778	GLU	-	expression tag	UNP Q5V8X3
F	779	LYS	-	expression tag	UNP Q5V8X3
F	780	GLY	-	expression tag	UNP Q5V8X3
F	781	SER	-	expression tag	UNP Q5V8X3
F	782	GLY	-	expression tag	UNP Q5V8X3
F	783	SER	-	expression tag	UNP Q5V8X3
F	784	GLU	-	expression tag	UNP Q5V8X3
F	785	ASN	-	expression tag	UNP Q5V8X3
F	786	LEU	-	expression tag	UNP Q5V8X3
F	787	TYR	-	expression tag	UNP Q5V8X3
F	788	PHE	-	expression tag	UNP Q5V8X3
F	789	GLN	-	expression tag	UNP Q5V8X3

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	Y	3	Total	C	N	O	P	0	0	0
			24	12	3	8	1			
4	G	21	Total	C	N	O	P	0	0	0
			165	96	21	45	3			
4	X	4	Total	C	N	O	P	0	0	0
			30	15	4	10	1			
4	I	18	Total	C	N	O	P	0	0	0
			142	85	18	37	2			

- Molecule 5 is a RNA chain called RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP *AP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	13	Total	C	N	O	P	0	0	0
			266	117	52	84	13			
5	V	13	Total	C	N	O	P	0	0	0
			266	117	52	84	13			

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	43	Total	O	0	0
			43	43		

Continued from previous page...

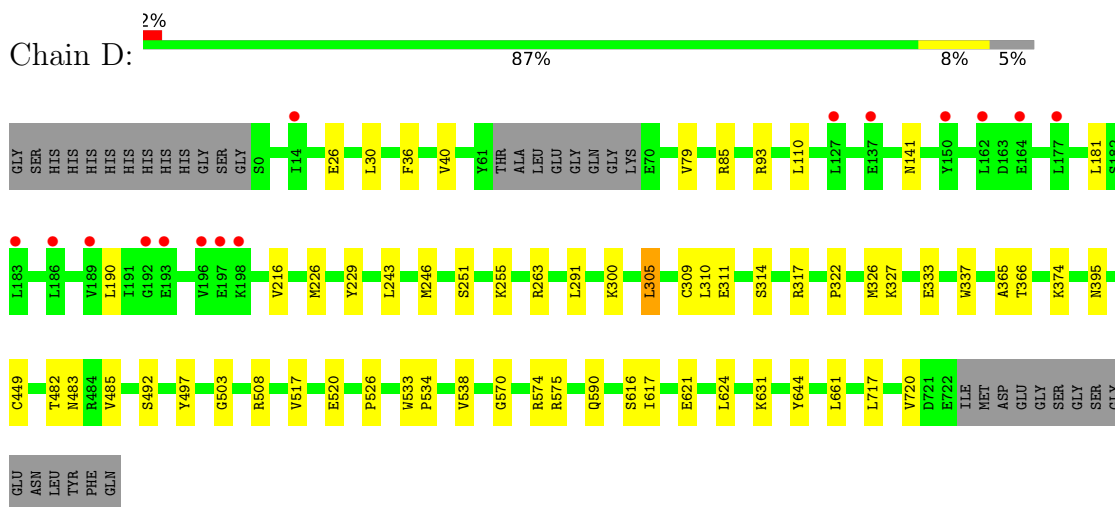| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	30	Total 30	O 30	0	0
6	C	20	Total 20	O 20	0	0
6	D	39	Total 39	O 39	0	0
6	E	43	Total 43	O 43	0	0
6	F	16	Total 16	O 16	0	0
6	G	1	Total 1	O 1	0	0
6	H	6	Total 6	O 6	0	0
6	I	1	Total 1	O 1	0	0
6	V	7	Total 7	O 7	0	0

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.
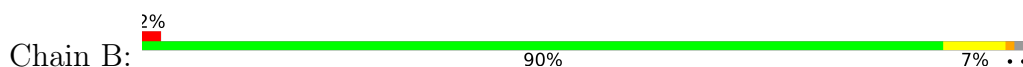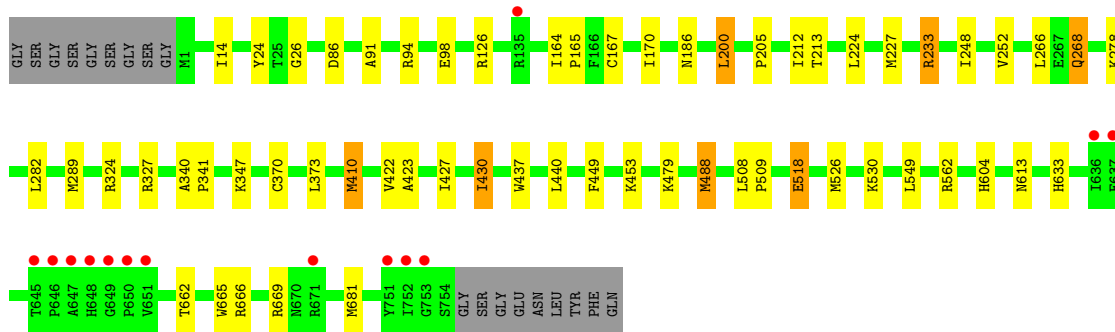
- Molecule 1: Polymerase acidic protein
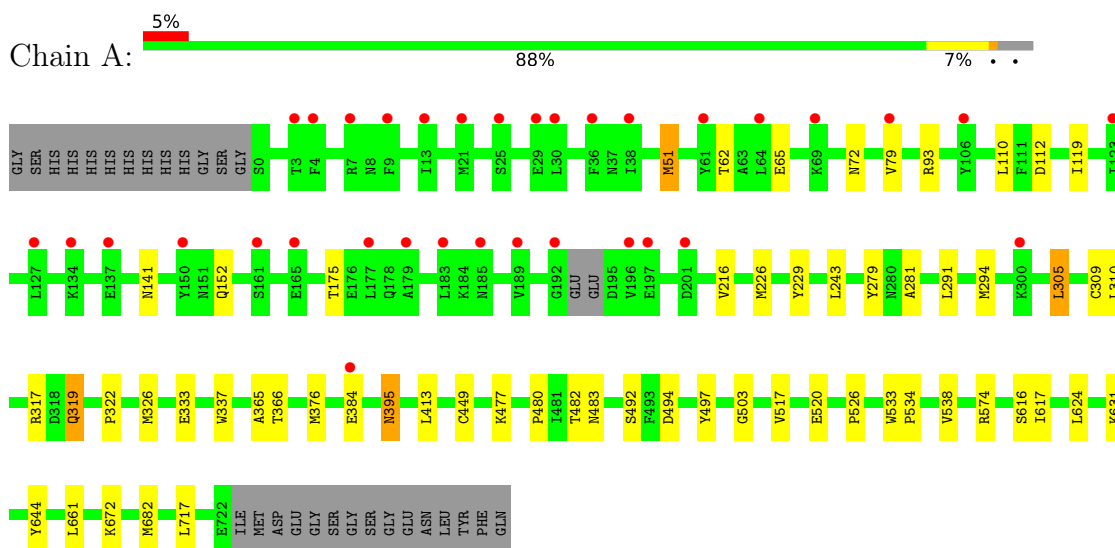
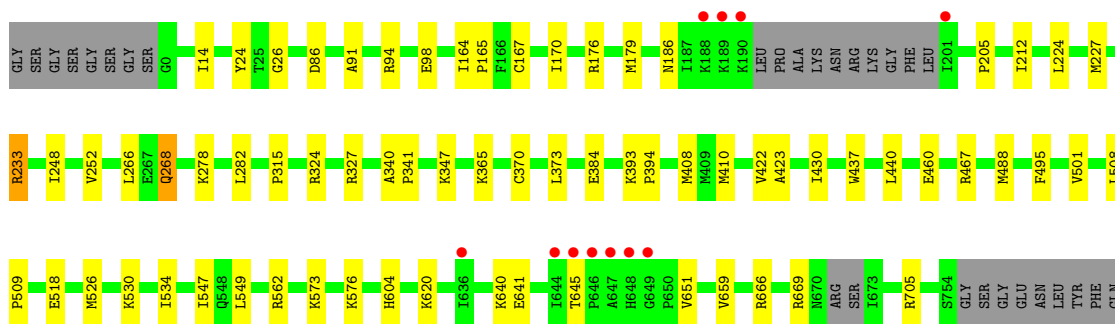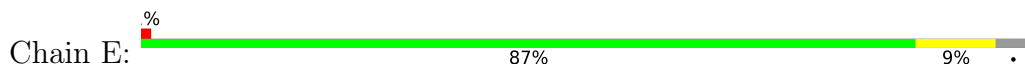- Molecule 1: Polymerase acidic protein

- Molecule 2: RNA-directed RNA polymerase catalytic subunit

GLY SER GLY SER GLY SER GLY M1 I14 Y24 T25 G26 D86 A91 R94 E98 R126 R135 I164 P165 F166 C167 I170 N186 L200 P205 I212 T213 L224 M227 R233 I248 V252 L266 E267 Q268 K278

L282 M289 R324 R327 A340 P341 K347 C370 L373 M410 V422 A423 I427 I430 W437 L440 F449 K453 K479 M488 L508 P509 E518 M526 K530 L549 R562 H604 N613 H633 I636 E637

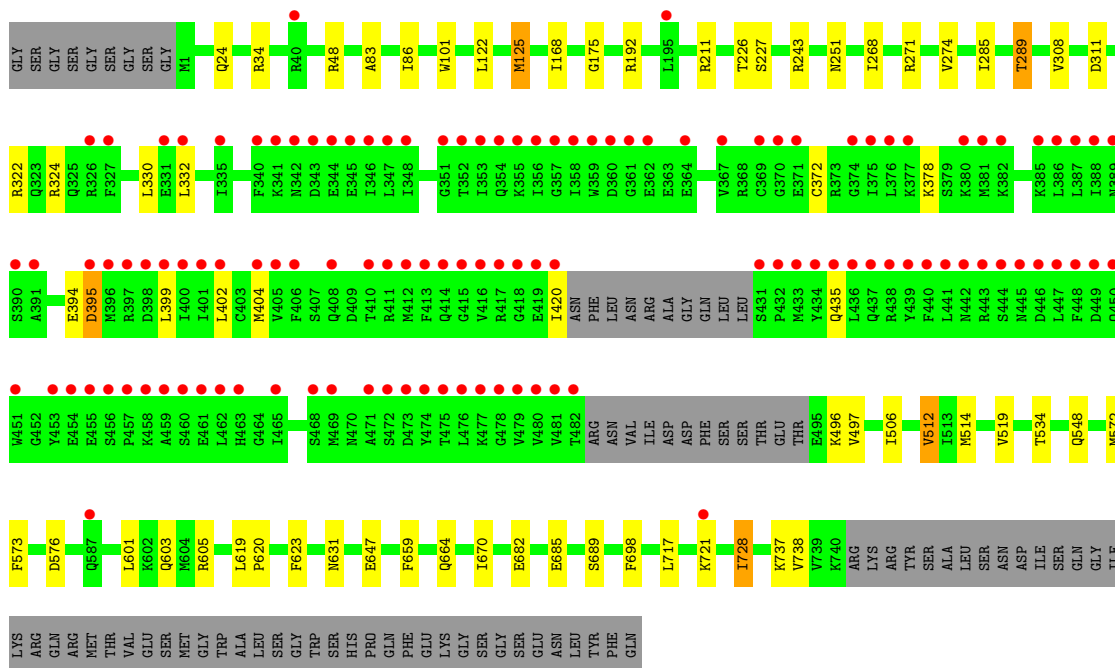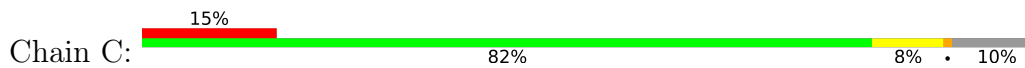T645 P646 A647 H648 G649 P650 V651 T662 W665 R666 R669 N670 R671 M681 Y751 I752 G753 S754 GLY SER GLY GLU ASN LEU TYR PHE GLN
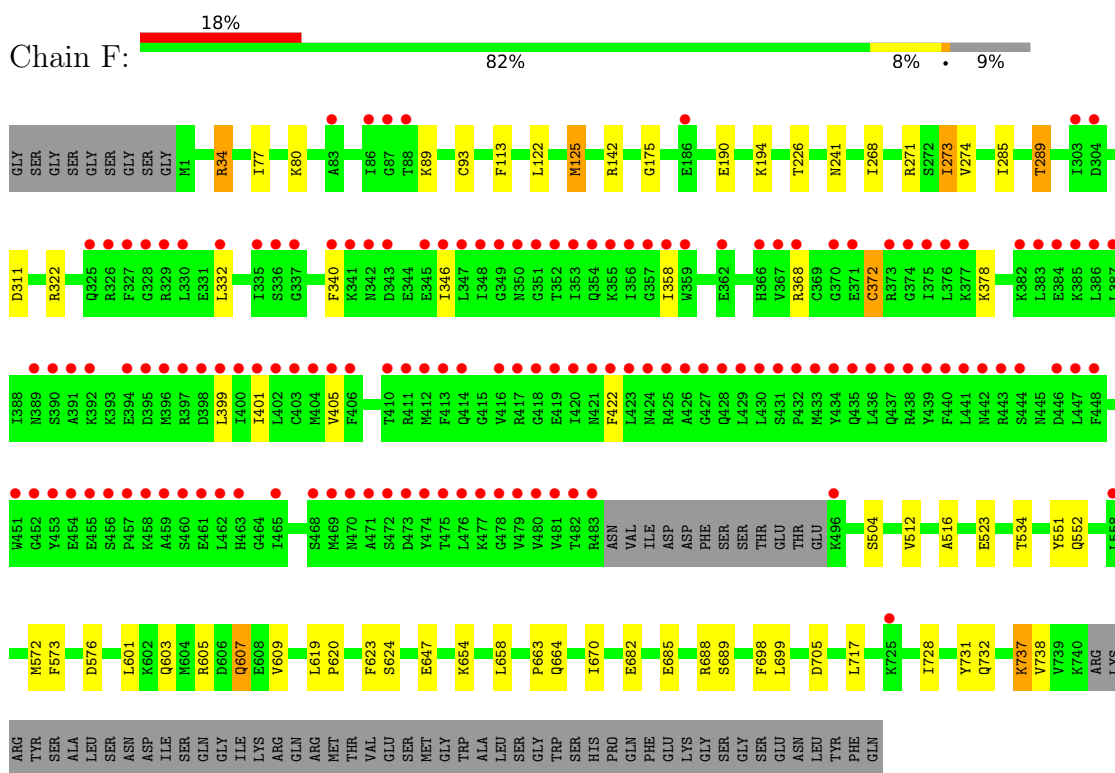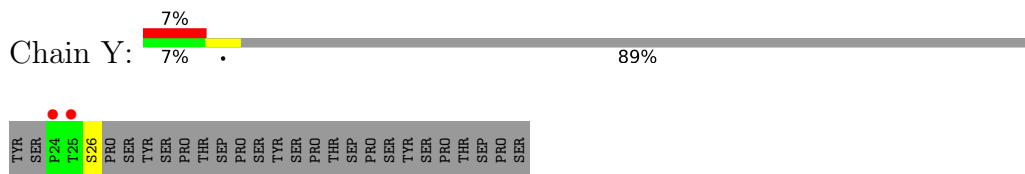
• Molecule 2: RNA-directed RNA polymerase catalytic subunit

Chain E:

%
87%   9%   •

GLY SER GLY SER GLY SER GLY G0 I14 Y24 T25 G26 D86 A91 R94 E98 I164 P165 F166 C167 I170 R176 M179 N186 I187 K188 K189 K190 LEU PRO ALA LYS ASN ARG LYS GLY PHE LEU I201 P205 I212 L224 M227

R233 I248 V252 L266 E267 Q268 K278 L282 P315 R324 R327 A340 P341 K347 K365 C370 L373 E384 K393 P394 M408 M409 M410 V422 A423 I430 W437 L440 E460 R467 M488 F495 V501 L508

P509 E518 M526 K530 I534 I547 Q548 L549 R562 K573 K576 H604 K620 I636 K640 E641 I644 T645 P646 A647 H648 G649 P650 V651 V659 R666 R669 N670 ARG SER I673 R705 S754 GLY SER GLY GLU ASN LEU TYR PHE GLN

• Molecule 3: Polymerase basic protein 2

Chain C:

15%
82%   8%   •   10%

GLY SER GLY SER GLY SER GLY M1 Q24 R34 R40 R48 A83 I86 W101 L122 M125 I168 G175 R192 L195 R211 T226 S227 R243 N251 I268 R271 V274 I285 T289 V308 D311

R322 Q323 R324 Q325 R326 F327 L330 E331 L332 I335 F340 K341 N342 D343 E344 E345 I346 L347 I348 G351 T352 I353 Q354 K355 I356 G357 I358 W359 D360 G361 E362 E363 E364 V367 R368 C369 G370 E371 C372 R373 G374 I375 L376 K377 K378 G374 L376 K377 K378 K385 L386 L387 I388 N389

S390 A391 E394 D395 M396 R397 D398 L399 I400 I401 L402 C403 M404 V405 F406 S407 Q408 D409 T410 R411 M412 F413 Q414 G415 V416 R417 G418 E419 L420 ASN PHE LEU ASN ARG ALA GLY GLN LEU LEU S431 P432 M433 Y434 Q435 L436 Q437 R438 Y439 F440 L441 N442 R443 S444 N445 D446 L447 F448 D449 Q450

W451 G452 Y453 E454 E455 S456 P457 K458 A459 S460 E461 L462 H463 G464 I465 S468 M469 N470 A471 S472 D473 Y474 T475 L476 K477 G478 V479 V480 V481 T482 ARG ASN VAL ILE ASP ASP PHE SER SER THR GLU THR E495 K496 V497 I506 V512 I513 M514 V519 T534 Q548 M572

F573 D576 Q587 L601 K602 Q603 M604 R605 L619 P620 F623 N631 E647 F659 Q664 I670 E682 E685 S689 F698 L717 K721 I728 K737 V738 V739 K740 ARG LYS TYR SER ALA LEU SER ASN ASP ILE SER GLN GLY ILE

LYS ARG GLN ARG MET THR VAL GLU SER MET GLY TRP ALA LEU SER GLY TRP SER HIS PRO GLN PHE GLU LYS SER GLY SER GLY SER GLU ASN LEU TYR PHE GLN
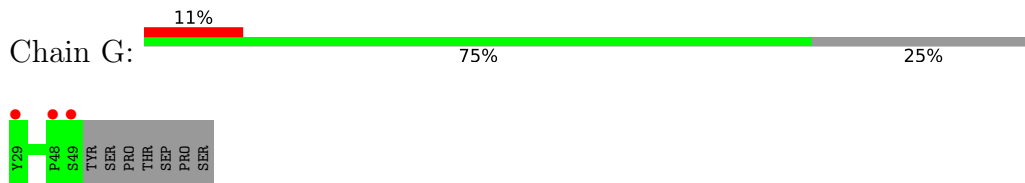
• Molecule 3: Polymerase basic protein 2
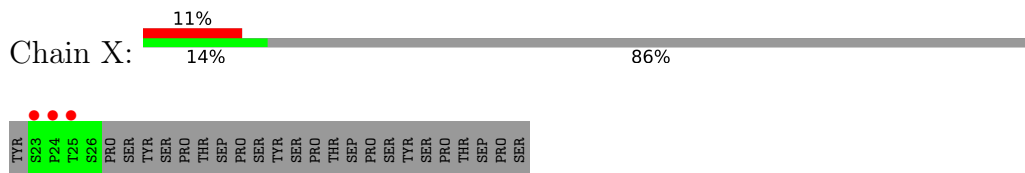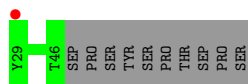
Chain F:

• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain Y:

• Molecule 4: DNA-directed RNA polymerase II subunit RPB1
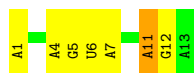
Chain G:

• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain X:

• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain I:

Y29 T46 SEP PRO SER TYR SER. PRO THR SEP PRO SER

• Molecule 5: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*GP*A)-3')

Chain H:
46%　46%　8%

A1 A4 G5 U6 A7 A11 G12 A13

• Molecule 5: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*GP*A)-3')

Chain V:
46%　54%

A1 A4 G5 U6 A7 A11 G12 A13

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	130.93Å 202.27Å 135.73Å 90.00° 110.54° 90.00°	Depositor
Resolution (Å)	127.11 – 2.42 127.11 – 2.42	Depositor EDS
% Data completeness (in resolution range)	61.4 (127.11-2.42) 61.4 (127.11-2.42)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.38 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.223 , 0.257 0.224 , 0.256	Depositor DCC
R_{free} test set	7573 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	44.1	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 39.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.001 for l,-k,h	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	35949	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SEP

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	0.67	0/5903	0.71	0/7959
1	D	0.67	0/5866	0.70	0/7910
2	B	0.67	0/6028	0.71	0/8127
2	E	0.67	0/5931	0.71	0/7995
3	C	0.67	0/5844	0.72	0/7852
3	F	0.67	0/5939	0.72	0/7982
4	G	0.60	0/140	0.63	0/189
4	I	0.62	0/127	0.63	0/173
4	X	0.71	0/20	0.59	0/27
4	Y	0.57	0/14	0.80	0/18
5	H	0.93	1/298 (0.3%)	1.26	3/462 (0.6%)
5	V	0.93	1/298 (0.3%)	1.25	2/462 (0.4%)
All	All	0.67	2/36408 (0.0%)	0.73	5/49156 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	V	1	A	OP3-P	-7.51	1.52	1.61
5	H	1	A	OP3-P	-7.43	1.52	1.61

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	V	5	G	P-O3'-C3'	-6.80	111.54	119.70
5	H	12	G	P-O3'-C3'	-6.47	111.93	119.70
5	H	5	G	P-O3'-C3'	-6.04	112.45	119.70
5	H	6	U	P-O3'-C3'	-5.63	112.95	119.70
5	V	6	U	P-O3'-C3'	-5.56	113.03	119.70

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	5787	0	5761	33	0
1	D	5750	0	5717	28	0
2	B	5911	0	5932	30	0
2	E	5818	0	5824	34	0
3	C	5746	0	5909	25	0
3	F	5838	0	6007	33	0
4	G	165	0	130	0	0
4	I	142	0	115	0	0
4	X	30	0	21	0	0
4	Y	24	0	17	1	0
5	H	266	0	130	1	0
5	V	266	0	130	1	0
6	A	43	0	0	0	0
6	B	30	0	0	0	0
6	C	20	0	0	0	0
6	D	39	0	0	0	0
6	E	43	0	0	1	0
6	F	16	0	0	0	0
6	G	1	0	0	0	0
6	H	6	0	0	0	0
6	I	1	0	0	0	0
6	V	7	0	0	0	0
All	All	35949	0	35693	160	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 2.

The worst 5 of 160 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:51:MET:O	1:A:62:THR:HA	1.92	0.70
2:B:282:LEU:HD22	2:B:440:LEU:HD13	1.76	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:717:LEU:HD21	3:F:728:ILE:HD11	1.77	0.66
2:E:282:LEU:HD22	2:E:440:LEU:HD13	1.78	0.66
3:F:273:ILE:HD11	3:F:516:ALA:CB	2.26	0.66

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	718/751 (96%)	691 (96%)	26 (4%)	1 (0%)	51	67
1	D	712/751 (95%)	688 (97%)	24 (3%)	0	100	100
2	B	752/772 (97%)	726 (96%)	26 (4%)	0	100	100
2	E	737/772 (96%)	714 (97%)	23 (3%)	0	100	100
3	C	712/798 (89%)	678 (95%)	34 (5%)	0	100	100
3	F	725/798 (91%)	693 (96%)	31 (4%)	1 (0%)	51	67
4	G	16/28 (57%)	16 (100%)	0	0	100	100
4	I	14/28 (50%)	13 (93%)	1 (7%)	0	100	100
4	X	2/28 (7%)	2 (100%)	0	0	100	100
4	Y	1/28 (4%)	1 (100%)	0	0	100	100
All	All	4389/4754 (92%)	4222 (96%)	165 (4%)	2 (0%)	100	100

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	GLU
3	F	89	LYS

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	641/664 (96%)	624 (97%)	17 (3%)	44	63
1	D	638/664 (96%)	613 (96%)	25 (4%)	32	49
2	B	646/657 (98%)	620 (96%)	26 (4%)	31	48
2	E	636/657 (97%)	614 (96%)	22 (4%)	36	53
3	C	628/694 (90%)	597 (95%)	31 (5%)	25	39
3	F	637/694 (92%)	609 (96%)	28 (4%)	28	43
4	G	18/24 (75%)	18 (100%)	0	100	100
4	I	16/24 (67%)	16 (100%)	0	100	100
4	X	3/24 (12%)	3 (100%)	0	100	100
4	Y	2/24 (8%)	2 (100%)	0	100	100
All	All	3865/4126 (94%)	3716 (96%)	149 (4%)	32	49

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	562	ARG
3	F	607	GLN
2	E	669	ARG
3	F	368	ARG
3	C	227	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	590	GLN
2	E	310	ASN
3	F	325	GLN
2	B	613	ASN
1	A	319	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	H	11/13 (84%)	3 (27%)	0
5	V	11/13 (84%)	3 (27%)	0
All	All	22/26 (84%)	6 (27%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	H	4	A
5	H	7	A
5	H	11	A
5	V	4	A
5	V	7	A

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

7 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
4	SEP	G	47	4	8,9,10	0.54	0	8,12,14	0.64	0
4	SEP	G	33	4	8,9,10	0.56	0	8,12,14	0.64	0
4	SEP	I	33	4	8,9,10	0.56	0	8,12,14	0.62	0
4	SEP	X	26	4	8,9,10	0.55	0	8,12,14	0.68	0
4	SEP	Y	26	4	8,9,10	0.57	0	8,12,14	0.64	0
4	SEP	I	40	4	8,9,10	0.51	0	8,12,14	0.64	0
4	SEP	G	40	4	8,9,10	0.72	0	8,12,14	0.63	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	SEP	G	47	4	-	2/5/8/10	-
4	SEP	G	33	4	-	2/5/8/10	-
4	SEP	I	33	4	-	2/5/8/10	-
4	SEP	X	26	4	-	1/5/8/10	-
4	SEP	Y	26	4	-	1/5/8/10	-
4	SEP	I	40	4	-	0/5/8/10	-
4	SEP	G	40	4	-	4/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	G	40	SEP	CB-OG-P-O1P
4	G	40	SEP	CB-OG-P-O2P
4	G	40	SEP	CB-OG-P-O3P
4	Y	26	SEP	N-CA-CB-OG
4	X	26	SEP	N-CA-CB-OG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	Y	26	SEP	1	0

5.5 Carbohydrates (i)

There are no monosaccharides 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	721/751 (96%)	0.34	34 (4%)	31	29	21, 40, 126, 149	0
1	D	715/751 (95%)	0.22	15 (2%)	63	60	23, 39, 116, 139	0
2	B	754/772 (97%)	0.15	14 (1%)	66	64	23, 45, 76, 116	0
2	E	743/772 (96%)	0.16	11 (1%)	73	71	23, 42, 69, 110	0
3	C	718/798 (89%)	1.06	119 (16%)	1	1	25, 56, 185, 226	0
3	F	728/798 (91%)	1.24	140 (19%)	1	0	26, 54, 172, 193	0
4	G	18/28 (64%)	0.87	3 (16%)	1	1	46, 68, 131, 137	0
4	I	16/28 (57%)	0.69	1 (6%)	20	18	46, 63, 75, 89	0
4	X	3/28 (10%)	4.20	3 (100%)	0	0	94, 94, 96, 97	0
4	Y	2/28 (7%)	3.86	2 (100%)	0	0	98, 98, 98, 98	0
5	H	13/13 (100%)	-0.46	0	100	100	25, 29, 56, 69	0
5	V	13/13 (100%)	-0.39	0	100	100	25, 28, 40, 47	0
All	All	4444/4780 (92%)	0.53	342 (7%)	13	11	21, 46, 139, 226	0

The worst 5 of 342 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	425	ARG	25.0
3	F	423	LEU	22.2
3	C	460	SER	16.8
3	C	431	SER	16.0
3	C	419	GLU	14.7

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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	SEP	G	47	10/11	0.70	0.27	124,135,140,141	0
4	SEP	Y	26	10/11	0.82	0.22	79,84,91,92	0
4	SEP	X	26	10/11	0.87	0.17	79,83,87,89	0
4	SEP	G	40	10/11	0.93	0.11	64,67,72,72	0
4	SEP	I	40	10/11	0.93	0.12	65,71,84,84	0
4	SEP	I	33	10/11	0.96	0.09	55,58,61,62	0
4	SEP	G	33	10/11	0.96	0.11	49,53,54,55	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.