



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 05:53 PM GMT

PDB ID : 2Y0S
Title : Crystal structure of Sulfolobus shibatae RNA polymerase in P21 space group
Authors : Wojtas, M.; Peralta, B.; Ondiviela, M.; Mogni, M.; Bell, S.D.; Abrescia, N.G.A.
Deposited on : 2010-12-07
Resolution : 3.80 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

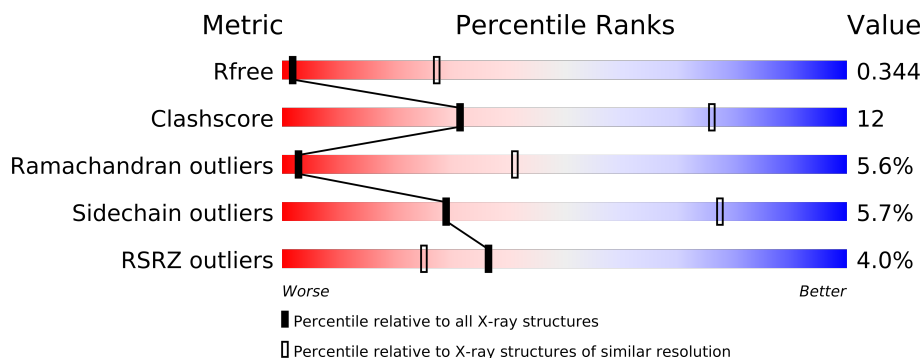
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.80 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.











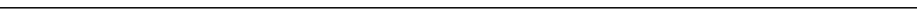



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1162 (4.20-3.40)
Clashscore	79885	1100 (4.10-3.50)
Ramachandran outliers	78287	1050 (4.10-3.50)
Sidechain outliers	78261	1042 (4.10-3.50)
RSRZ outliers	66119	1163 (4.20-3.40)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	880	
1	W	880	
2	B	1131	
2	R	1131	
3	C	395	
3	Y	395	
4	D	265	
4	S	265	
5	E	180	
5	T	180	
6	F	113	
6	U	113	
7	G	132	
7	V	132	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	84	
8	Z	84	
9	I	95	
9	K	95	
10	J	104	
10	Q	104	
11	L	92	
11	M	92	
12	N	66	
12	O	66	
13	P	48	
13	X	48	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
14	ZN	B	2124	-	X
14	ZN	B	2125	-	X
14	ZN	R	2124	-	X
14	ZN	R	2125	-	X
15	MG	W	1882	-	X

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 52472 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	831	Total	C	N	O	S	0	0	0
			6620	4211	1170	1212	27			
1	W	831	Total	C	N	O	S	0	0	0
			6620	4211	1170	1212	27			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1085	Total	C	N	O	S	0	0	0
			8615	5461	1526	1599	29			
2	R	1085	Total	C	N	O	S	0	0	0
			8615	5461	1526	1599	29			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT A”.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	368	Total	C	N	O	S	0	0	0
			2845	1803	486	548	8			
3	Y	368	Total	C	N	O	S	0	0	0
			2845	1803	486	548	8			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	262	Total	C	N	O	S	0	0	0
			2081	1338	336	394	13			
4	S	262	Total	C	N	O	S	0	0	0
			2081	1338	336	394	13			

- Molecule 5 is a protein called RNA POLYMERASE SUBUNIT 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	164	Total	C	N	O	S	0	0	0
			1297	833	219	240	5			
5	T	164	Total	C	N	O	S	0	0	0
			1297	833	219	240	5			

- Molecule 6 is a protein called RNA POLYMERASE SUBUNIT 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	90	Total	C	N	O	S	0	0	0
			701	439	114	145	3			
6	U	90	Total	C	N	O	S	0	0	0
			701	439	114	145	3			

- Molecule 7 is a protein called RNA POLYMERASE SUBUNIT 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	110	Total	C	N	O	S	0	0	0
			878	558	147	169	4			
7	V	110	Total	C	N	O	S	0	0	0
			878	558	147	169	4			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT H.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	H	74	Total	C	N	O	0	0	0
			609	396	108	105			
8	Z	74	Total	C	N	O	0	0	0
			609	396	108	105			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	82	Total	C	N	O	S	0	0	0
			653	419	117	116	1			
9	K	82	Total	C	N	O	S	0	0	0
			653	419	117	116	1			

- Molecule 10 is a protein called RNA POLYMERASE SUBUNIT 13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	39	Total	C	N	O	S	0	0	0
			332	210	54	67	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	Q	39	Total	C	N	O	S	0	0	0
			332	210	54	67	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	40	ASN	ASP	CONFLICT	UNP B8YB65
Q	40	ASN	ASP	CONFLICT	UNP B8YB65

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	L	91	Total	C	N	O	S	0	0	0
			707	454	114	137	2			
11	M	91	Total	C	N	O	S	0	0	0
			707	454	114	137	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	N	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			
12	O	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			

- Molecule 13 is a protein called RNA POLYMERASE SUBUNIT 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	P	45	Total	C	N	O	S	0	0	0
			369	245	63	56	5			
13	X	45	Total	C	N	O	S	0	0	0
			369	245	63	56	5			

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	P	1	Total	Zn	0	0
			1	1		
14	B	3	Total	Zn	0	0
			3	3		
14	W	2	Total	Zn	0	0
			2	2		

Continued on next page...

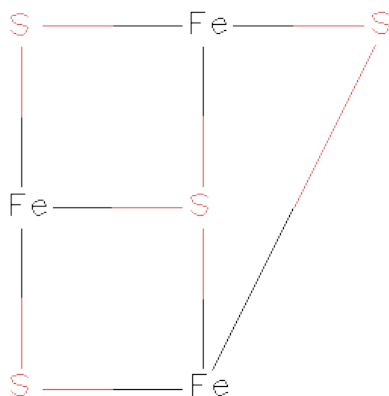
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total 2	Zn 2	0	0
14	N	1	Total 1	Zn 1	0	0
14	X	1	Total 1	Zn 1	0	0
14	O	1	Total 1	Zn 1	0	0
14	R	3	Total 3	Zn 3	0	0

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	W	1	Total 1	Mg 1	0	0
15	A	1	Total 1	Mg 1	0	0

- Molecule 16 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe₃S₄).



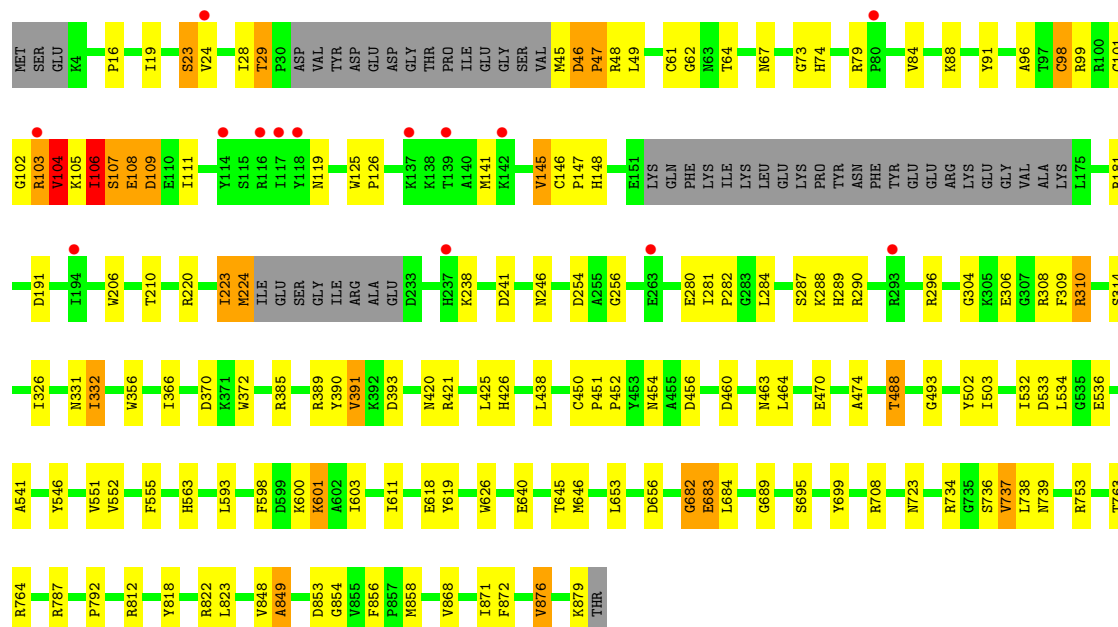
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
16	D	1	Total 7	Fe 3	S 4	0	0
16	S	1	Total 7	Fe 3	S 4	0	0

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 errors 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.

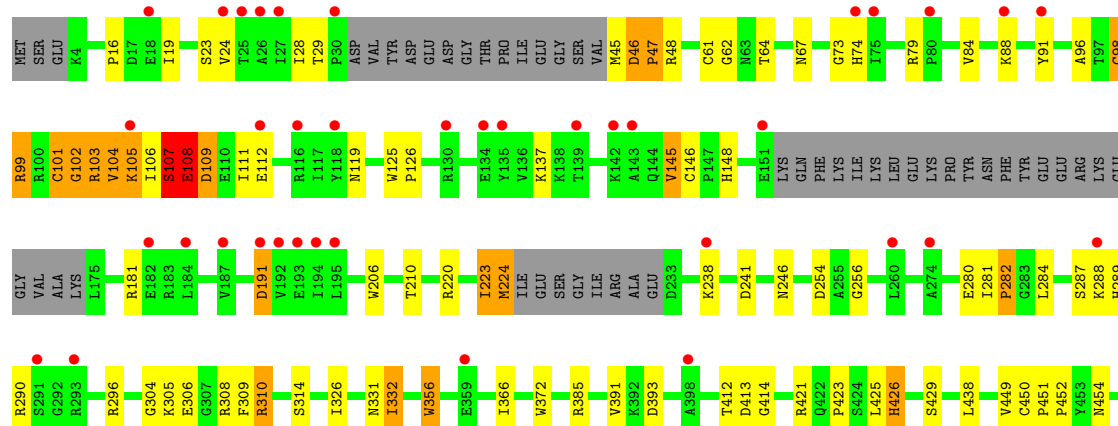
• Molecule 1: DNA-DIRECTED RNA POLYMERASE

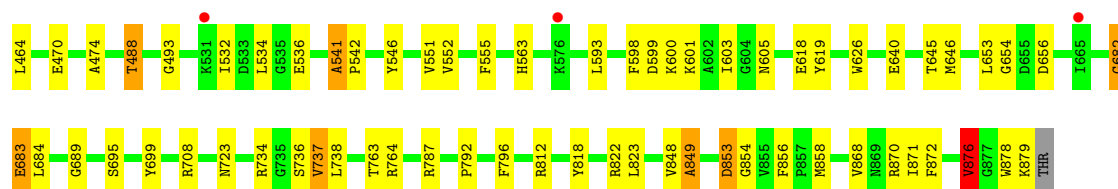
Chain A: 



• Molecule 1: DNA-DIRECTED RNA POLYMERASE

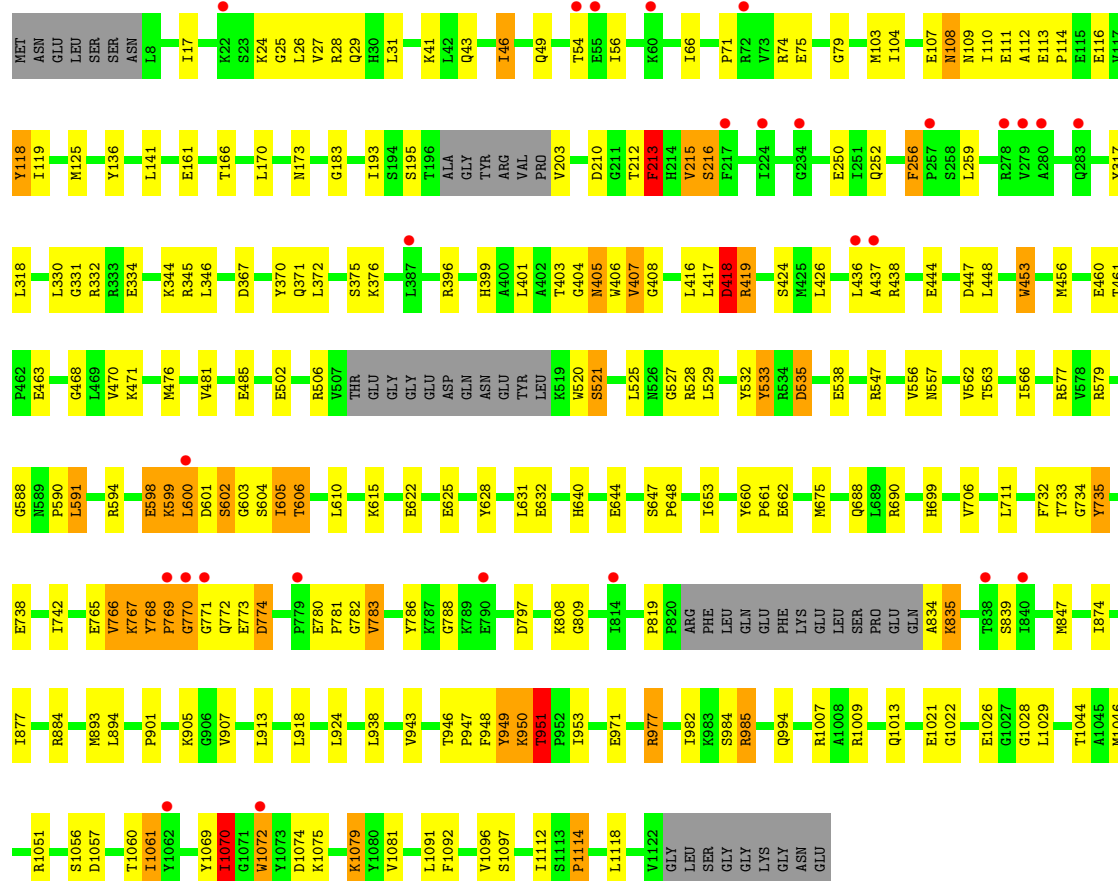
Chain W: 





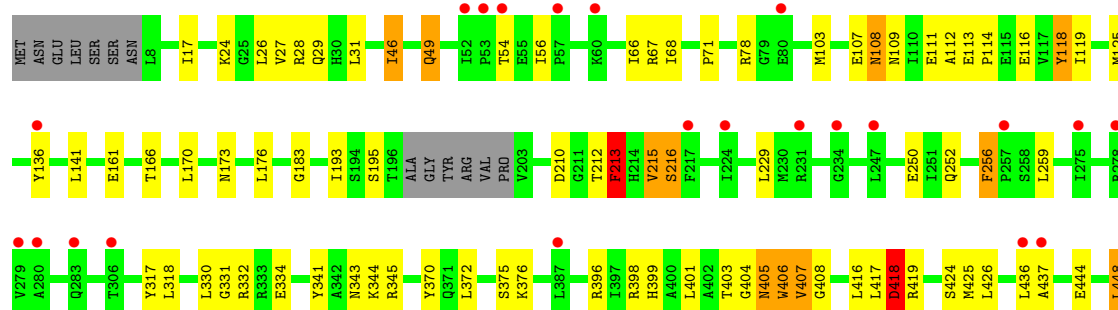
• Molecule 2: DNA-DIRECTED RNA POLYMERASE

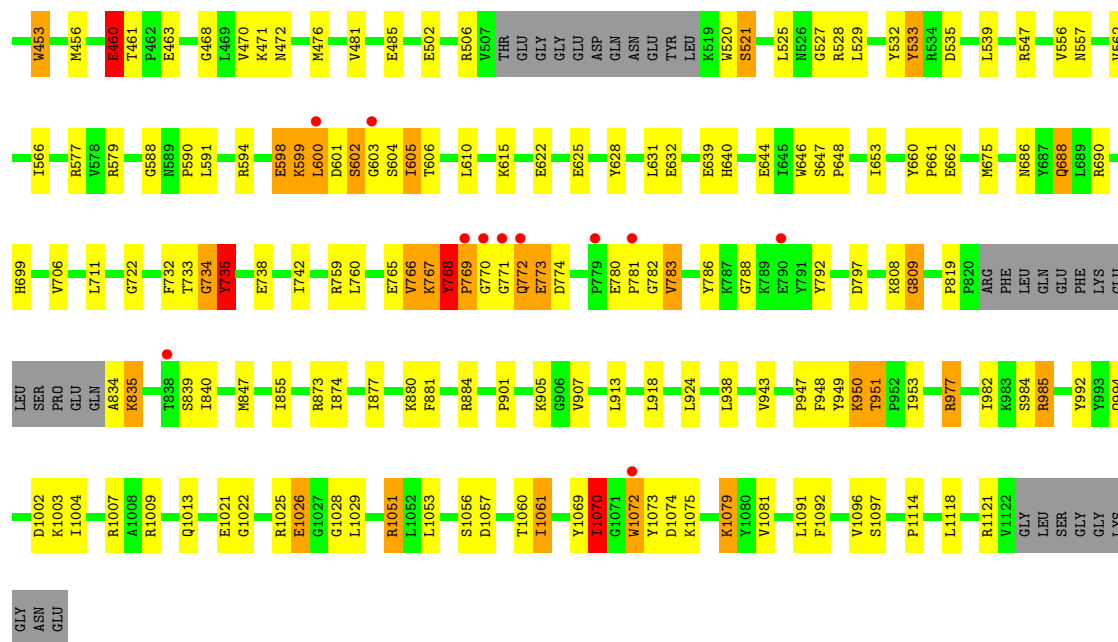
Chain B:



• Molecule 2: DNA-DIRECTED RNA POLYMERASE

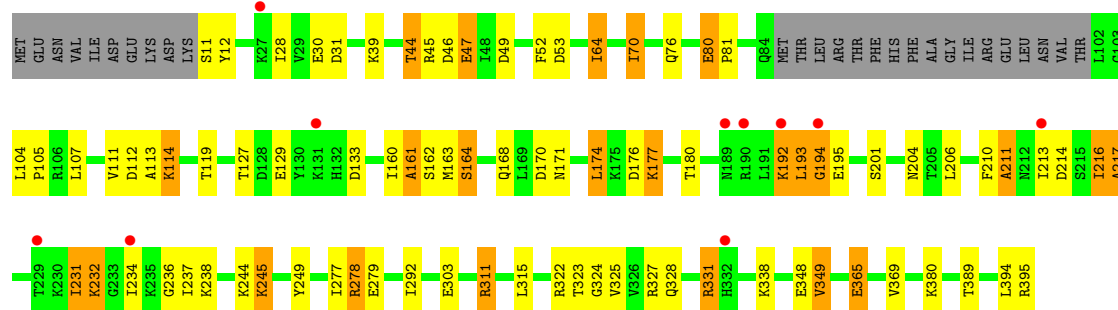
Chain R:





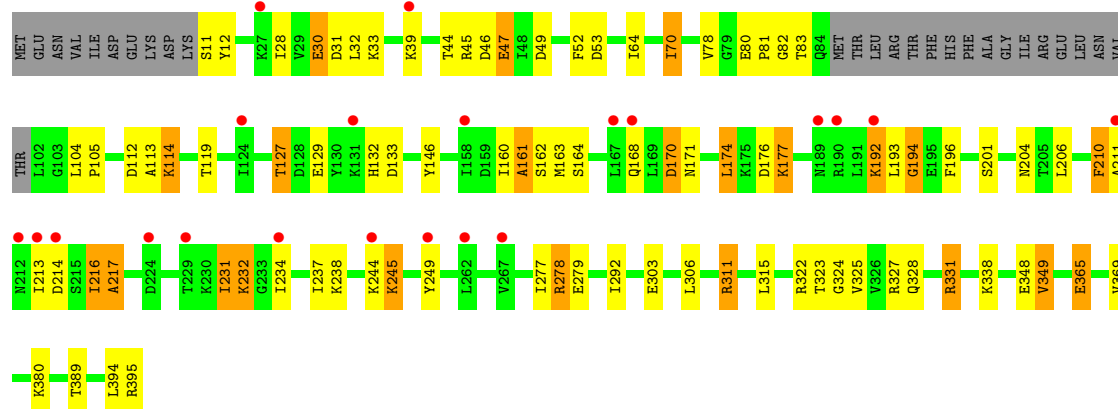
• Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT A”

Chain C:



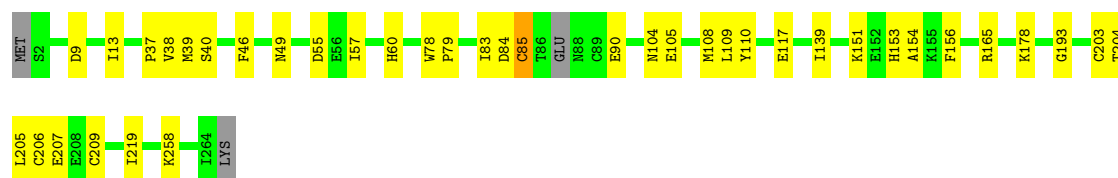
• Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT A”

Chain Y:



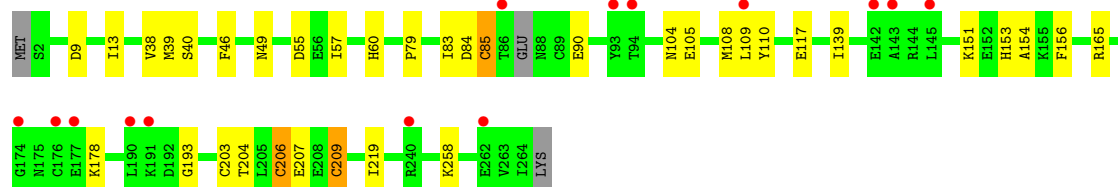
• Molecule 4: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

Chain D: 



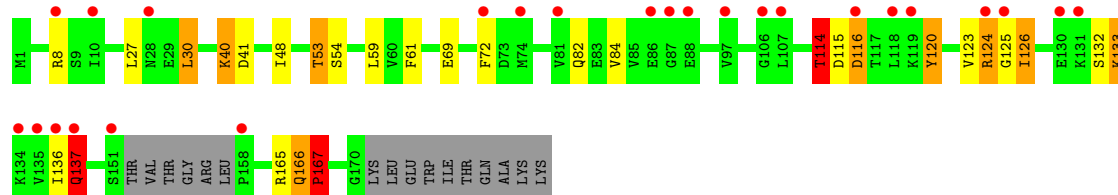
- Molecule 4: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

Chain S: 



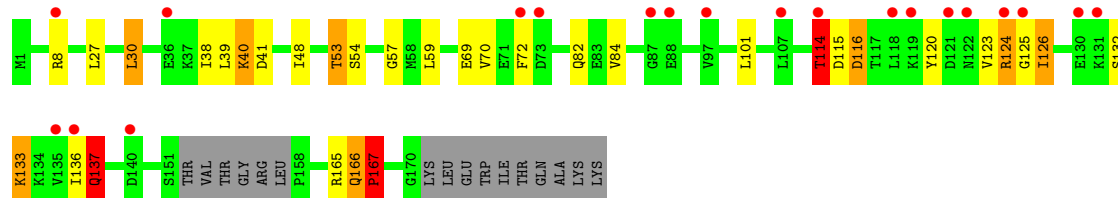
- Molecule 5: RNA POLYMERASE SUBUNIT 4

Chain E: 



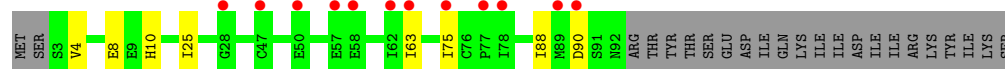
- Molecule 5: RNA POLYMERASE SUBUNIT 4

Chain T: 



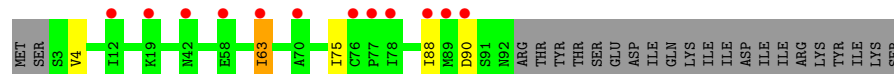
- Molecule 6: RNA POLYMERASE SUBUNIT 7

Chain F: 



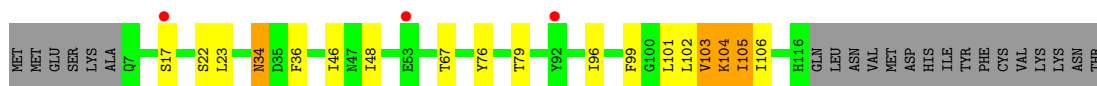
- Molecule 6: RNA POLYMERASE SUBUNIT 7

Chain U: 



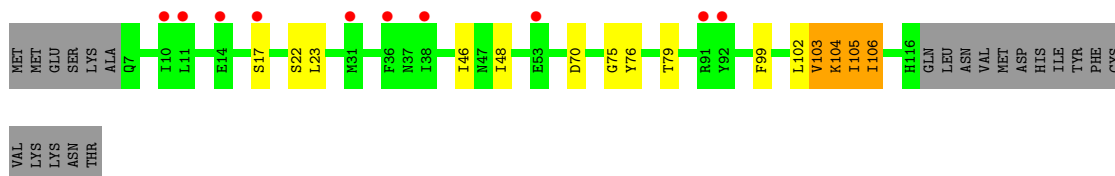
- Molecule 7: RNA POLYMERASE SUBUNIT 8

Chain G:



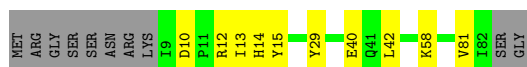
- Molecule 7: RNA POLYMERASE SUBUNIT 8

Chain V:



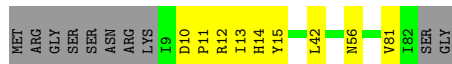
- Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT H

Chain H:



- Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT H

Chain Z:



- Molecule 9: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain I:



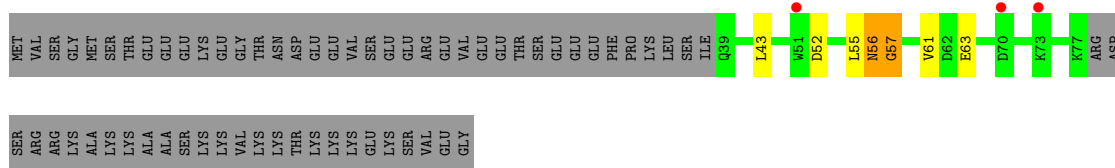
- Molecule 9: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain K:



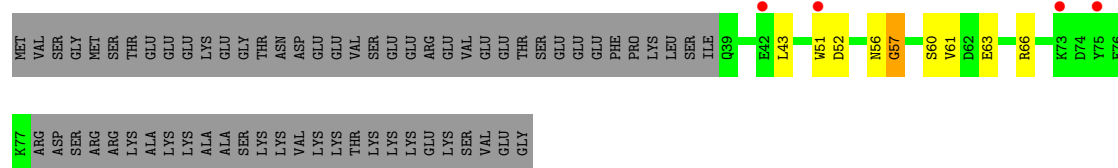
- Molecule 10: RNA POLYMERASE SUBUNIT 13

Chain J:



- Molecule 10: RNA POLYMERASE SUBUNIT 13

Chain Q:



- Molecule 11: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain L:



- Molecule 11: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain M:



- Molecule 12: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

Chain N:



- Molecule 12: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

Chain O:



- Molecule 13: RNA POLYMERASE SUBUNIT 12

Chain P:



- Molecule 13: RNA POLYMERASE SUBUNIT 12

Chain X:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	129.33Å 195.91Å 212.45Å 90.00° 105.73° 90.00°	Depositor
Resolution (Å)	29.86 – 3.80 47.12 – 3.79	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.86-3.80) 98.6 (47.12-3.79)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.94 (at 3.77Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.300 , 0.354 0.287 , 0.344	Depositor DCC
R_{free} test set	4995 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	79.6	Xtriage
Anisotropy	0.414	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 73.3	EDS
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 99994 reflections	Xtriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	52472	wwPDB-VP
Average B, all atoms (Å ²)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.01 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.6829e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, F3S, 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	3/6762 (0.0%)	0.81	7/9151 (0.1%)
1	W	0.44	2/6762 (0.0%)	0.62	4/9151 (0.0%)
2	B	0.50	0/8778	0.81	1/11873 (0.0%)
2	R	0.42	0/8778	0.60	0/11873
3	C	0.50	1/2869 (0.0%)	0.80	0/3862
3	Y	0.41	0/2869	0.61	0/3862
4	D	0.39	0/2116	0.64	0/2859
4	S	0.32	0/2116	0.48	0/2859
5	E	0.40	0/1316	0.71	0/1775
5	T	0.36	1/1316 (0.1%)	0.53	0/1775
6	F	0.28	0/709	0.62	0/961
6	U	0.25	0/709	0.41	0/961
7	G	0.40	0/890	0.67	0/1194
7	V	0.33	0/890	0.51	0/1194
8	H	0.53	0/623	0.77	0/845
8	Z	0.42	0/623	0.58	0/845
9	I	1.29	3/662 (0.5%)	0.90	5/896 (0.6%)
9	K	1.33	4/662 (0.6%)	1.01	5/896 (0.6%)
10	J	0.36	0/336	0.51	0/450
10	Q	0.43	0/336	0.71	0/450
11	L	0.40	0/717	0.68	0/968
11	M	0.31	0/717	0.52	0/968
12	N	0.43	0/524	0.68	0/706
12	O	0.35	0/524	0.52	0/706
13	P	0.49	0/378	0.80	0/507
13	X	0.44	0/378	0.66	0/507
All	All	0.49	14/53360 (0.0%)	0.69	22/72094 (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
5	E	0	1
5	T	0	1
9	I	0	1
9	K	0	1
All	All	0	5

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	56	ILE	C-N	-22.63	0.82	1.34
9	K	55	ASN	C-N	21.74	1.84	1.34
9	K	56	ILE	C-N	-21.68	0.84	1.34
9	I	55	ASN	C-N	20.09	1.80	1.34
1	A	104	VAL	CB-CG2	-9.48	1.32	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	I	55	ASN	O-C-N	9.91	138.56	122.70
9	K	55	ASN	CA-C-N	-9.88	95.47	117.20
9	I	55	ASN	CA-C-N	-9.78	95.68	117.20
9	I	56	ILE	O-C-N	-9.40	107.66	122.70
9	K	55	ASN	O-C-N	9.26	137.52	122.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	VAL	Mainchain
5	E	167	PRO	Peptide
9	I	56	ILE	Mainchain
9	K	56	ILE	Mainchain
5	T	167	PRO	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit,

and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6620	0	0	85	0
1	W	6620	0	0	86	0
2	B	8615	0	0	115	0
2	R	8615	0	0	107	0
3	C	2845	0	0	48	0
3	Y	2845	0	0	49	0
4	D	2081	0	0	19	0
4	S	2081	0	0	19	0
5	E	1297	0	0	20	0
5	T	1297	0	0	24	0
6	F	701	0	0	3	0
6	U	701	0	0	2	0
7	G	878	0	0	6	0
7	V	878	0	0	7	0
8	H	609	0	0	5	0
8	Z	609	0	0	2	0
9	I	653	0	0	21	0
9	K	653	0	0	20	0
10	J	332	0	0	3	0
10	Q	332	0	0	5	0
11	L	707	0	0	2	0
11	M	707	0	0	2	0
12	N	514	0	0	2	0
12	O	514	0	0	4	0
13	P	369	0	0	6	0
13	X	369	0	0	8	0
14	A	2	0	0	0	0
14	B	3	0	0	0	0
14	N	1	0	0	0	0
14	O	1	0	0	0	0
14	P	1	0	0	0	0
14	R	3	0	0	0	0
14	W	2	0	0	0	0
14	X	1	0	0	0	0
15	A	1	0	0	0	0
15	W	1	0	0	0	0
16	D	7	0	0	3	0
16	S	7	0	0	5	0
All	All	52472	0	0	606	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 12.

The worst 5 of 606 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
9:I:55:ASN:C	9:I:56:ILE:N	1.80	1.32
9:K:55:ASN:C	9:K:56:ILE:N	1.84	1.30
9:I:56:ILE:O	9:I:57:SER:N	1.67	1.24
9:K:56:ILE:O	9:K:57:SER:N	1.71	1.19
9:I:56:ILE:C	9:I:57:SER:CA	2.13	1.17

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	823/880 (94%)	669 (81%)	120 (15%)	34 (4%)	4	50
1	W	823/880 (94%)	672 (82%)	111 (14%)	40 (5%)	3	45
2	B	1077/1131 (95%)	887 (82%)	118 (11%)	72 (7%)	2	36
2	R	1077/1131 (95%)	881 (82%)	122 (11%)	74 (7%)	2	34
3	C	364/395 (92%)	278 (76%)	63 (17%)	23 (6%)	2	38
3	Y	364/395 (92%)	282 (78%)	54 (15%)	28 (8%)	1	30
4	D	258/265 (97%)	219 (85%)	31 (12%)	8 (3%)	7	59
4	S	258/265 (97%)	220 (85%)	30 (12%)	8 (3%)	7	59
5	E	160/180 (89%)	127 (79%)	21 (13%)	12 (8%)	2	31
5	T	160/180 (89%)	125 (78%)	24 (15%)	11 (7%)	2	34
6	F	88/113 (78%)	75 (85%)	11 (12%)	2 (2%)	10	65
6	U	88/113 (78%)	77 (88%)	9 (10%)	2 (2%)	10	65
7	G	108/132 (82%)	89 (82%)	15 (14%)	4 (4%)	5	54
7	V	108/132 (82%)	88 (82%)	16 (15%)	4 (4%)	5	54
8	H	72/84 (86%)	59 (82%)	9 (12%)	4 (6%)	3	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
8	Z	72/84 (86%)	59 (82%)	7 (10%)	6 (8%)	1	28
9	I	80/95 (84%)	64 (80%)	9 (11%)	7 (9%)	1	26
9	K	80/95 (84%)	64 (80%)	9 (11%)	7 (9%)	1	26
10	J	37/104 (36%)	27 (73%)	8 (22%)	2 (5%)	3	42
10	Q	37/104 (36%)	27 (73%)	8 (22%)	2 (5%)	3	42
11	L	89/92 (97%)	78 (88%)	9 (10%)	2 (2%)	10	66
11	M	89/92 (97%)	79 (89%)	8 (9%)	2 (2%)	10	66
12	N	62/66 (94%)	50 (81%)	10 (16%)	2 (3%)	6	58
12	O	62/66 (94%)	50 (81%)	9 (14%)	3 (5%)	4	45
13	P	43/48 (90%)	33 (77%)	7 (16%)	3 (7%)	2	34
13	X	43/48 (90%)	33 (77%)	7 (16%)	3 (7%)	2	34
All	All	6522/7170 (91%)	5312 (81%)	845 (13%)	365 (6%)	3	41

5 of 365 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	SER
1	A	103	ARG
1	A	108	GLU
1	A	683	GLU
1	A	737	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	723/766 (94%)	683 (94%)	40 (6%)	30	79
1	W	723/766 (94%)	678 (94%)	45 (6%)	26	75
2	B	937/975 (96%)	880 (94%)	57 (6%)	26	75
2	R	937/975 (96%)	871 (93%)	66 (7%)	21	70
3	C	316/341 (93%)	290 (92%)	26 (8%)	17	63
3	Y	316/341 (93%)	294 (93%)	22 (7%)	21	70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	D	233/238 (98%)	227 (97%)	6 (3%)	59	90
4	S	233/238 (98%)	226 (97%)	7 (3%)	53	89
5	E	144/158 (91%)	136 (94%)	8 (6%)	30	78
5	T	144/158 (91%)	135 (94%)	9 (6%)	25	74
6	F	84/107 (78%)	81 (96%)	3 (4%)	47	87
6	U	84/107 (78%)	81 (96%)	3 (4%)	47	87
7	G	104/125 (83%)	96 (92%)	8 (8%)	18	66
7	V	104/125 (83%)	99 (95%)	5 (5%)	35	82
8	H	67/75 (89%)	67 (100%)	0	100	100
8	Z	67/75 (89%)	67 (100%)	0	100	100
9	I	72/83 (87%)	70 (97%)	2 (3%)	56	89
9	K	72/83 (87%)	70 (97%)	2 (3%)	56	89
10	J	37/96 (38%)	35 (95%)	2 (5%)	31	79
10	Q	37/96 (38%)	35 (95%)	2 (5%)	31	79
11	L	79/80 (99%)	79 (100%)	0	100	100
11	M	79/80 (99%)	79 (100%)	0	100	100
12	N	58/60 (97%)	53 (91%)	5 (9%)	15	61
12	O	58/60 (97%)	54 (93%)	4 (7%)	22	70
13	P	41/43 (95%)	38 (93%)	3 (7%)	20	69
13	X	41/43 (95%)	38 (93%)	3 (7%)	20	69
All	All	5790/6294 (92%)	5462 (94%)	328 (6%)	29	77

5 of 328 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	25	ASN
2	R	341	TYR
13	X	18	LEU
9	K	25	ASN
10	Q	43	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 18 ligands modelled in this entry, 16 are monoatomic - leaving 2 for Mogul analysis.

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	$\# Z > 2$
16	F3S	D	1001	4	3,9,9	32.58	3 (100%)	0,15,15	0.00	-
16	F3S	S	1001	4	3,9,9	32.59	3 (100%)	0,15,15	0.00	-

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
16	F3S	D	1001	4	-	0/0/24/24	0/0/3/3
16	F3S	S	1001	4	-	0/0/24/24	0/0/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	D	1001	F3S	S3-FE4	-33.25	2.10	2.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	S	1001	F3S	S3-FE4	-33.18	2.10	2.33
16	S	1001	F3S	S3-FE1	-32.34	2.11	2.33
16	D	1001	F3S	S3-FE1	-32.32	2.11	2.33
16	S	1001	F3S	S3-FE3	-32.25	2.11	2.33

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 ⓘ

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(Å ²)	Q<0.9
1	A	831/880 (94%)	0.28	14 (1%)	67	46	28, 92, 200, 308	0
1	W	831/880 (94%)	0.42	41 (4%)	28	20	46, 106, 212, 378	0
2	B	1085/1131 (95%)	0.29	27 (2%)	54	36	29, 88, 181, 293	0
2	R	1085/1131 (95%)	0.32	33 (3%)	48	31	55, 93, 190, 334	0
3	C	368/395 (93%)	0.53	10 (2%)	52	34	33, 108, 214, 329	0
3	Y	368/395 (93%)	0.58	21 (5%)	23	16	56, 115, 224, 338	0
4	D	262/265 (98%)	0.32	0	100	100	59, 116, 219, 298	0
4	S	262/265 (98%)	0.50	14 (5%)	25	18	77, 134, 209, 265	0
5	E	164/180 (91%)	1.16	25 (15%)	3	4	70, 166, 254, 302	0
5	T	164/180 (91%)	1.12	20 (12%)	5	6	72, 164, 255, 300	0
6	F	90/113 (79%)	0.95	12 (13%)	4	4	98, 188, 257, 327	0
6	U	90/113 (79%)	1.03	12 (13%)	4	4	128, 184, 276, 323	0
7	G	110/132 (83%)	0.55	3 (2%)	52	34	66, 132, 226, 301	0
7	V	110/132 (83%)	0.82	10 (9%)	9	9	75, 145, 228, 295	0
8	H	74/84 (88%)	0.19	0	100	100	53, 82, 153, 205	0
8	Z	74/84 (88%)	0.24	0	100	100	83, 99, 171, 205	0
9	I	82/95 (86%)	0.27	0	100	100	57, 91, 160, 195	0
9	K	82/95 (86%)	0.24	0	100	100	35, 84, 181, 219	0
10	J	39/104 (37%)	1.02	3 (7%)	13	11	130, 183, 243, 268	0
10	Q	39/104 (37%)	0.81	4 (10%)	7	8	112, 162, 213, 265	0
11	L	91/92 (98%)	0.22	0	100	100	65, 99, 194, 250	0
11	M	91/92 (98%)	0.42	3 (3%)	44	29	95, 121, 175, 247	0
12	N	64/66 (96%)	0.31	0	100	100	59, 103, 177, 205	0
12	O	64/66 (96%)	0.13	0	100	100	91, 112, 168, 192	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
13	P	45/48 (93%)	0.53	4 (8%) 10 10	71, 93, 189, 275	0
13	X	45/48 (93%)	0.51	3 (6%) 17 14	85, 105, 197, 249	0
All	All	6610/7170 (92%)	0.43	259 (3%) 36 26	28, 105, 217, 378	0

The worst 5 of 259 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	119	LYS	7.0
5	T	119	LYS	6.6
1	W	24	VAL	5.2
5	E	87	GLY	5.1
1	W	118	TYR	5.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
14	ZN	B	2125	1/1	0.69	20.43	246,246,246,246	0
14	ZN	R	2125	1/1	1.04	7.00	246,246,246,246	0
14	ZN	R	2124	1/1	0.51	5.18	246,246,246,246	0
14	ZN	B	2124	1/1	0.50	3.12	246,246,246,246	0
15	MG	W	1882	1/1	0.47	2.40	246,246,246,246	0
16	F3S	D	1001	7/7	0.35	1.76	232,235,241,245	0
16	F3S	S	1001	7/7	0.34	1.02	247,251,257,260	0
15	MG	A	1882	1/1	0.25	0.96	246,246,246,246	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
14	ZN	W	1881	1/1	0.19	-0.58	246,246,246,246	0
14	ZN	B	2123	1/1	0.27	-0.63	246,246,246,246	0
14	ZN	A	1880	1/1	0.14	-0.95	246,246,246,246	0
14	ZN	R	2123	1/1	0.21	-1.09	246,246,246,246	0
14	ZN	X	1049	1/1	0.12	-1.14	157,157,157,157	0
14	ZN	P	1049	1/1	0.09	-1.17	142,142,142,142	0
14	ZN	A	1881	1/1	0.07	-1.23	246,246,246,246	0
14	ZN	N	1065	1/1	0.12	-1.29	130,130,130,130	0
14	ZN	O	1065	1/1	0.10	-1.64	165,165,165,165	0
14	ZN	W	1880	1/1	0.08	-2.32	246,246,246,246	0

6.5 Other polymers ⓘ

There are no such residues in this entry.