wwPDB X-ray Structure Validation Summary Report (i)

Oct 19, 2023 – 09:50 AM EDT

PDB ID	:	2NVY
Title	:	RNA Polymerase II form II in 150 mM Mn+2
Authors	:	Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.
Deposited on	:	2006-11-13
Resolution	:	3.40 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 (i) 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 (i)) were used in the production of this report:

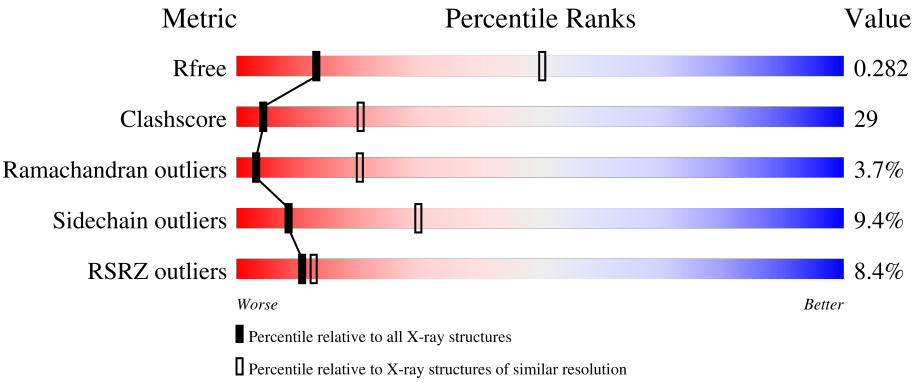
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.40 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.



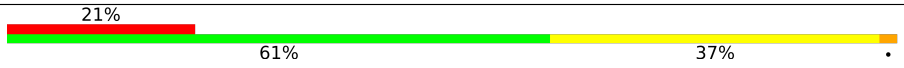
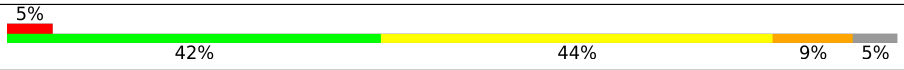
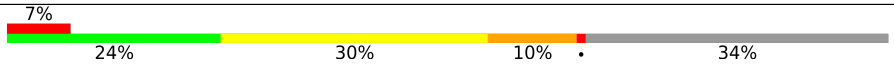
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	1733	7% 39% 37% 6% 18%
2	B	1224	8% 44% 41% · 11%
3	C	318	2% 35% 42% 6% 16%
4	E	215	4% 58% 39% ·
5	F	155	3% 21% 30% · 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	H	146	
7	I	122	
8	J	70	
9	K	120	
10	L	70	

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 28289 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 DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1419	Total	C	N	O	S	0	0	0
			11154	7023	1952	2118	61			

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1094	Total	C	N	O	S	0	0	0
			8711	5525	1519	1614	53			

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	122	Total	C	N	O	S	0	0	0
			997	613	182	191	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total	Zn	0	0
			2	2		
11	B	1	Total	Zn	0	0
			1	1		
11	C	1	Total	Zn	0	0
			1	1		
11	I	2	Total	Zn	0	0
			2	2		
11	J	1	Total	Zn	0	0
			1	1		
11	L	1	Total	Zn	0	0
			1	1		

- Molecule 12 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

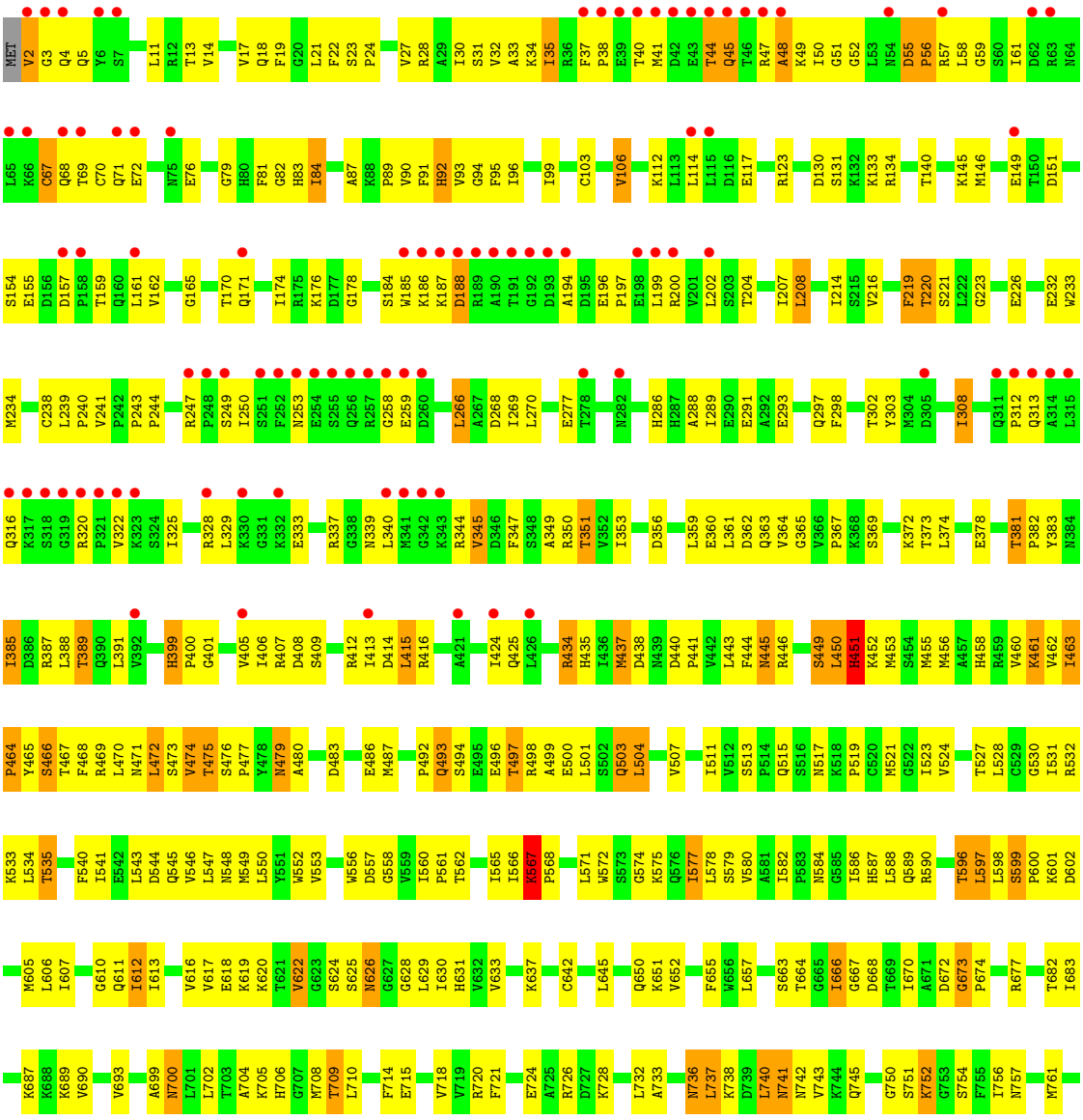
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	2	Total Mn 2 2	0	0

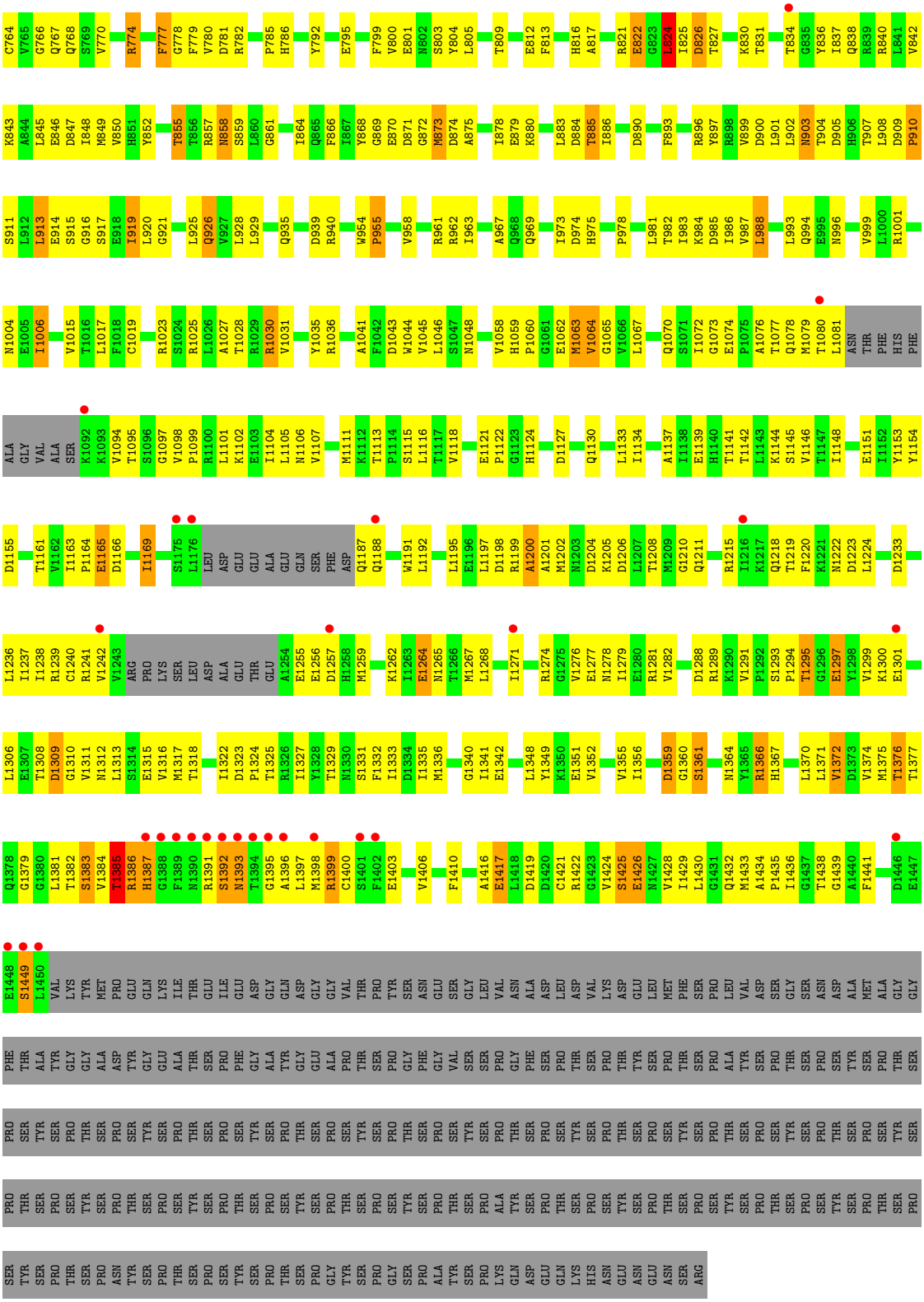
3 Residue-property plots (i)

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: DNA-directed RNA polymerase II largest subunit

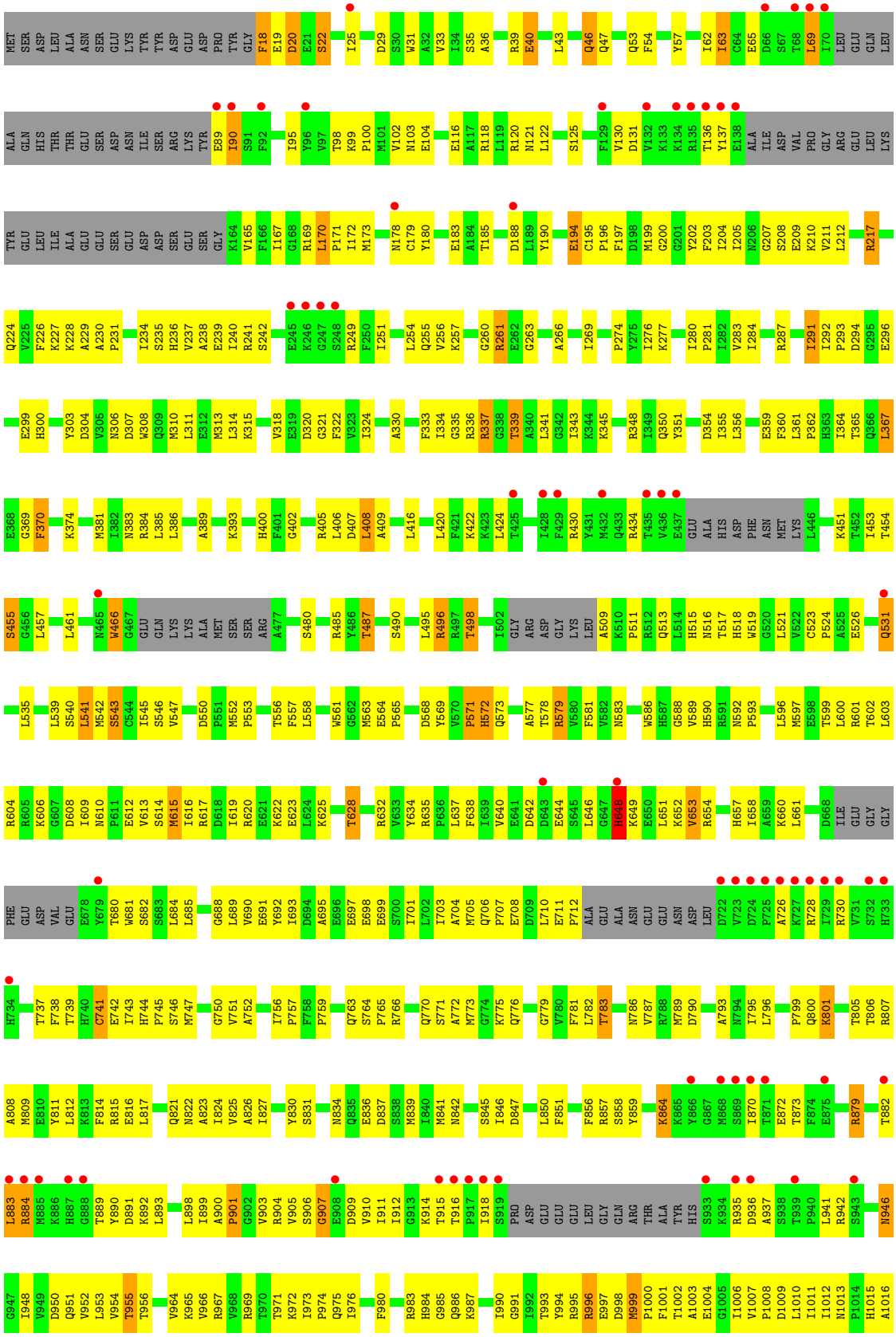
Chain A:

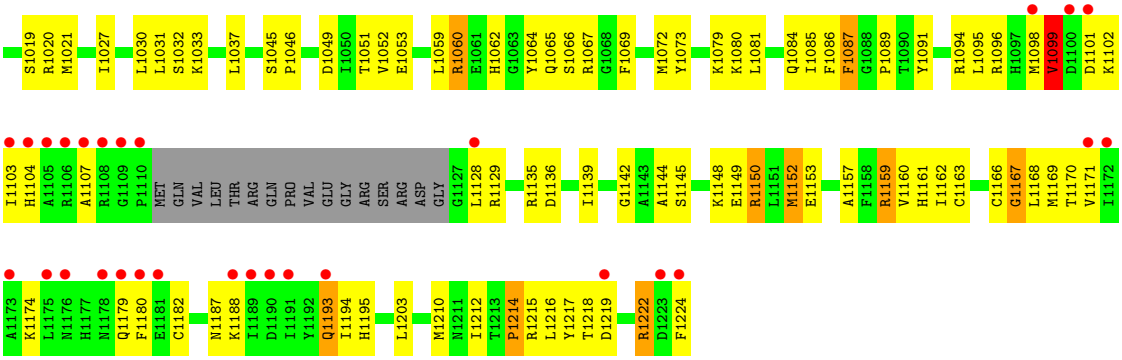




• Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide

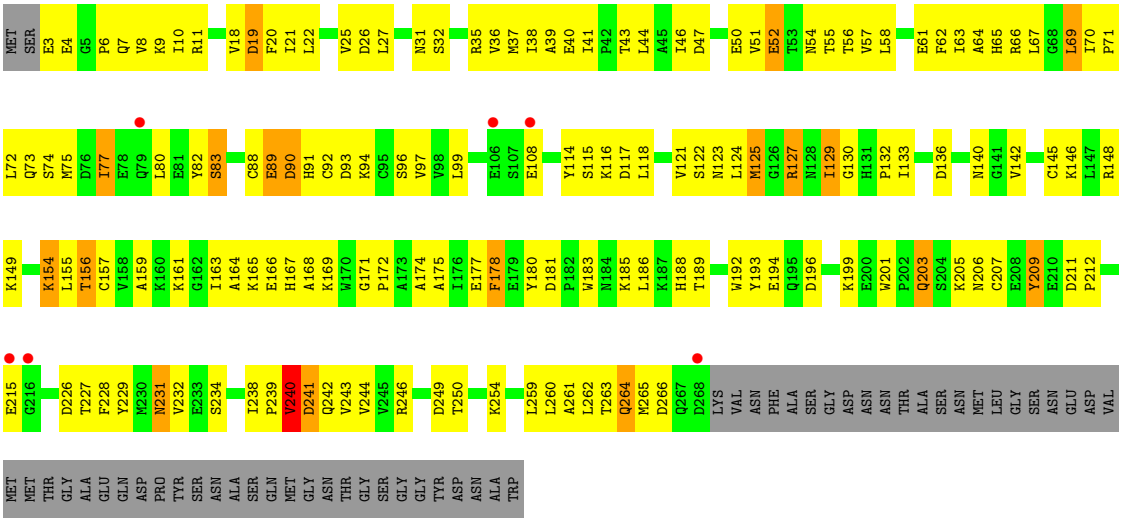
Chain B:





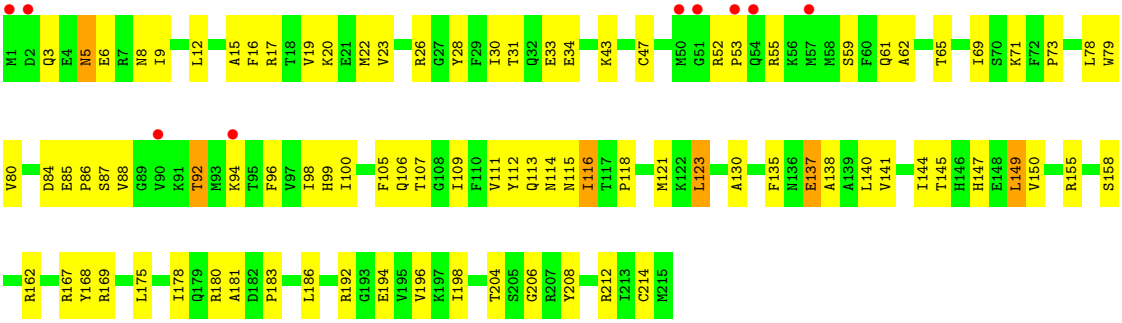
• Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C:



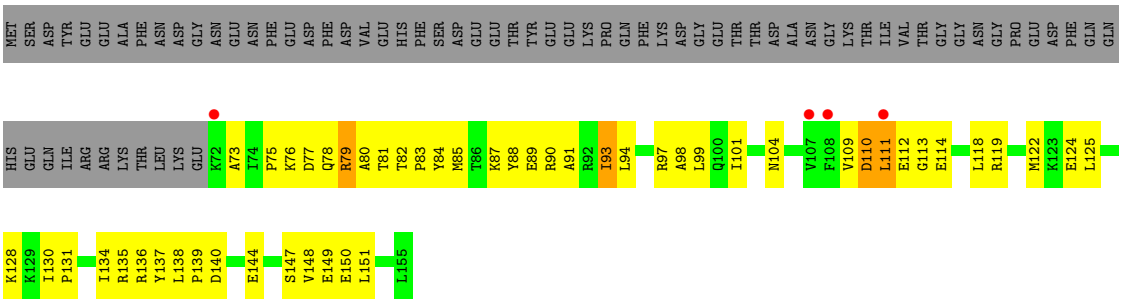
• Molecule 4: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E:



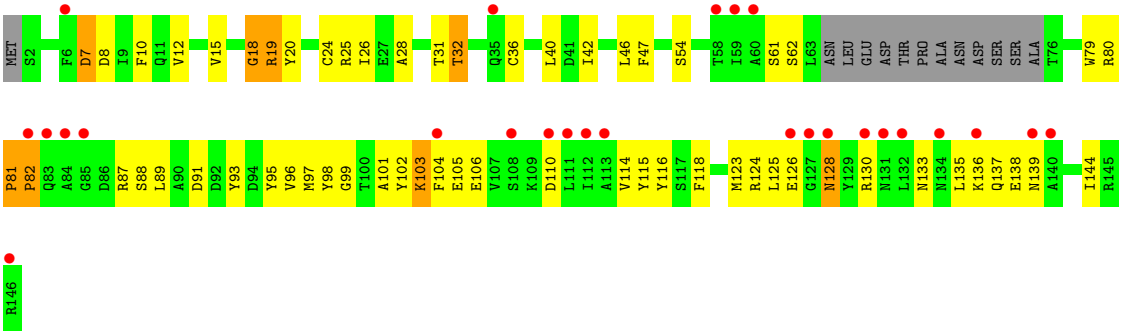
• Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F:



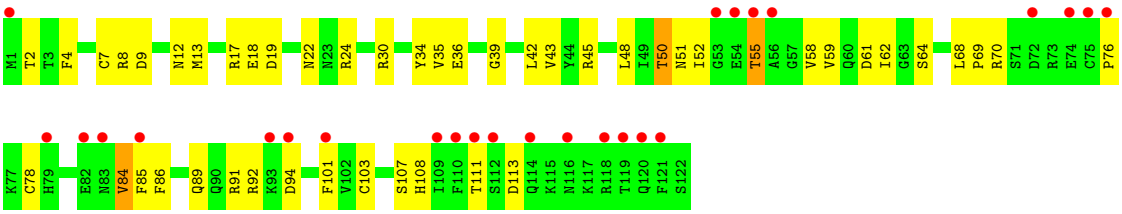
• Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H:



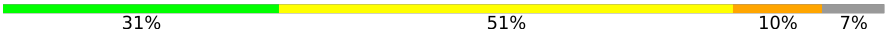
• Molecule 7: DNA-directed RNA polymerase II subunit 9

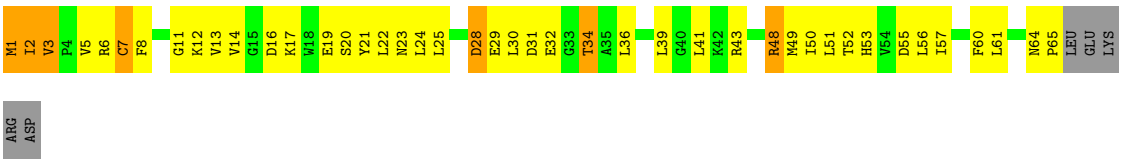
Chain I:



• Molecule 8: DNA-directed RNA polymerases I/II/III subunit 10

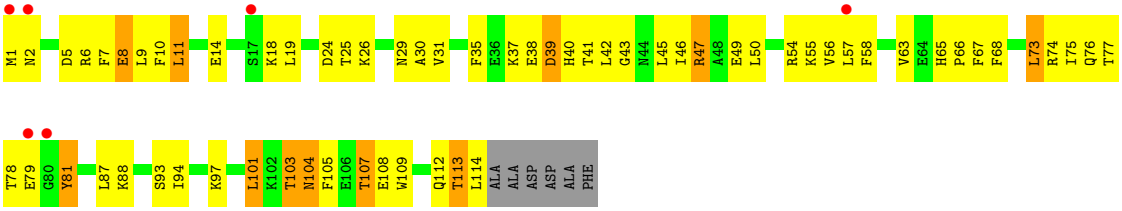
Chain J:



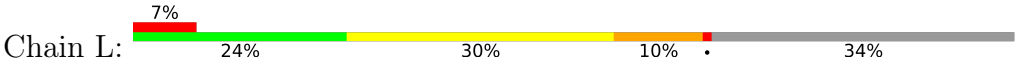


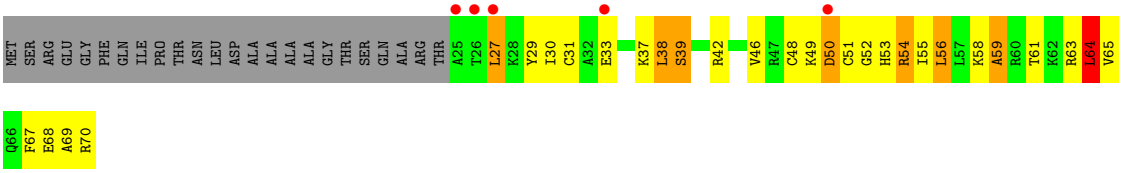
• Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:



• Molecule 10: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide





4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	118.04Å 218.91Å 369.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.40 39.71 – 3.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.40) 97.8 (39.71-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	4.00 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.249 , 0.323 0.229 , 0.282	Depositor DCC
R_{free} test set	1970 reflections (3.04%)	wwPDB-VP
Wilson B-factor (Å^2)	89.6	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 48.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	28289	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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: ZN, MN

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	0.60	2/11352 (0.0%)	0.71	1/15352 (0.0%)
2	B	0.58	0/8882	0.69	0/11976
3	C	0.58	0/2133	0.70	0/2891
4	E	0.57	0/1796	0.70	0/2416
5	F	0.61	0/691	0.71	0/933
6	H	0.47	0/1086	0.65	0/1470
7	I	0.52	0/1016	0.64	0/1365
8	J	0.62	0/541	0.76	0/727
9	K	0.56	0/937	0.71	0/1265
10	L	0.54	0/366	0.73	0/485
All	All	0.58	2/28800 (0.0%)	0.70	1/38880 (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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1449	SER	CB-OG	12.62	1.58	1.42
1	A	1421	CYS	CB-SG	-5.25	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	824	LEU	CA-CB-CG	5.03	126.86	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	450	LEU	Peptide

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	11154	0	11225	712	0
2	B	8711	0	8737	561	0
3	C	2095	0	2051	161	0
4	E	1760	0	1788	74	0
5	F	679	0	701	54	0
6	H	1068	0	1040	74	0
7	I	997	0	955	40	0
8	J	532	0	542	59	0
9	K	919	0	929	81	0
10	L	364	0	388	35	0
11	A	2	0	0	0	0
11	B	1	0	0	0	0
11	C	1	0	0	0	0
11	I	2	0	0	0	0
11	J	1	0	0	0	0
11	L	1	0	0	0	0
12	A	2	0	0	0	0
All	All	28289	0	28356	1637	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 29.

The worst 5 of 1637 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)
2:B:800:GLN:HB3	8:J:52:THR:CG2	1.63	1.25
1:A:567:LYS:HB2	1:A:568:PRO:CD	1.66	1.21
1:A:672:ASP:HB2	1:A:736:ASN:ND2	1.60	1.15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:855:THR:HG21	1:A:857:ARG:HE	0.98	1.14
1:A:672:ASP:CB	1:A:736:ASN:HD21	1.59	1.14

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	1411/1733 (81%)	1182 (84%)	174 (12%)	55 (4%)	3 19
2	B	1074/1224 (88%)	924 (86%)	122 (11%)	28 (3%)	5 26
3	C	264/318 (83%)	221 (84%)	34 (13%)	9 (3%)	3 21
4	E	213/215 (99%)	179 (84%)	30 (14%)	4 (2%)	8 31
5	F	82/155 (53%)	68 (83%)	11 (13%)	3 (4%)	3 20
6	H	129/146 (88%)	95 (74%)	22 (17%)	12 (9%)	0 4
7	I	120/122 (98%)	96 (80%)	21 (18%)	3 (2%)	5 26
8	J	63/70 (90%)	53 (84%)	8 (13%)	2 (3%)	4 22
9	K	112/120 (93%)	92 (82%)	14 (12%)	6 (5%)	2 13
10	L	44/70 (63%)	25 (57%)	12 (27%)	7 (16%)	0 0
All	All	3512/4173 (84%)	2935 (84%)	448 (13%)	129 (4%)	3 20

5 of 129 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	48	ALA
1	A	55	ASP
1	A	415	LEU
1	A	464	PRO

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	1239/1520 (82%)	1108 (89%)	131 (11%)	6 24
2	B	950/1061 (90%)	867 (91%)	83 (9%)	10 34
3	C	234/274 (85%)	208 (89%)	26 (11%)	6 22
4	E	197/197 (100%)	188 (95%)	9 (5%)	27 57
5	F	74/137 (54%)	68 (92%)	6 (8%)	11 38
6	H	117/128 (91%)	115 (98%)	2 (2%)	60 80
7	I	116/116 (100%)	108 (93%)	8 (7%)	15 45
8	J	60/65 (92%)	49 (82%)	11 (18%)	1 5
9	K	99/102 (97%)	88 (89%)	11 (11%)	6 22
10	L	40/57 (70%)	34 (85%)	6 (15%)	3 12
All	All	3126/3657 (86%)	2833 (91%)	293 (9%)	8 30

5 of 293 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	189	THR
9	K	113	THR
3	C	264	GLN
7	I	35	VAL
1	A	1118	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	24	ASN
6	H	128	ASN
3	C	65	HIS
3	C	242	GLN
9	K	29	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 10 ligands modelled in this entry, 10 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

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, 95^{th} 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	1419/1733 (81%)	0.40	119 (8%) 11 13	41, 79, 162, 188	0
2	B	1094/1224 (89%)	0.39	98 (8%) 9 11	47, 80, 140, 167	0
3	C	266/318 (83%)	0.13	6 (2%) 60 59	58, 78, 111, 133	0
4	E	215/215 (100%)	0.20	9 (4%) 36 35	51, 98, 130, 145	0
5	F	84/155 (54%)	0.20	4 (4%) 30 31	52, 71, 92, 98	0
6	H	133/146 (91%)	1.01	26 (19%) 1 1	103, 119, 147, 149	0
7	I	122/122 (100%)	1.07	26 (21%) 0 1	75, 107, 129, 150	0
8	J	65/70 (92%)	-0.05	0 100 100	58, 70, 91, 94	0
9	K	114/120 (95%)	0.33	6 (5%) 26 27	58, 85, 107, 126	0
10	L	46/70 (65%)	0.76	5 (10%) 5 6	87, 137, 144, 145	0
All	All	3558/4173 (85%)	0.40	299 (8%) 11 13	41, 82, 146, 188	0

The worst 5 of 299 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1110	PRO	9.7
1	A	323	LYS	9.1
1	A	1390	ASN	9.1
1	A	188	ASP	8.2
2	B	919	SER	8.1

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 (i)

There are no monosaccharides 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, 95^{th} 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
12	MN	A	3010	1/1	0.83	0.11	50,50,50,50	0
11	ZN	L	3005	1/1	0.93	0.06	128,128,128,128	0
12	MN	A	3009	1/1	0.95	0.15	49,49,49,49	0
11	ZN	B	3007	1/1	0.97	0.04	87,87,87,87	0
11	ZN	A	3008	1/1	0.97	0.07	143,143,143,143	0
11	ZN	A	3006	1/1	0.98	0.04	95,95,95,95	0
11	ZN	C	3002	1/1	0.99	0.04	79,79,79,79	0
11	ZN	I	3003	1/1	0.99	0.08	97,97,97,97	0
11	ZN	I	3004	1/1	0.99	0.12	117,117,117,117	0
11	ZN	J	3001	1/1	1.00	0.13	68,68,68,68	0

6.5 Other polymers (i)

There are no such residues in this entry.