wwPDB X-ray Structure Validation Summary Report (i)

Sep 9, 2024 – 10:10 AM EDT

PDB ID	:	8U9R
Title	:	STRUCTURAL BASIS OF TRANSCRIPTION: RNA POLYMERASE II SUBSTRATE BINDING AND METAL COORDINATION USING A FREE-ELECTRON LASER
Authors	:	Arjunan, P.; Calero, G.; Kaplan, C.D.
Deposited on	:	2023-09-20
Resolution	:	3.34 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:

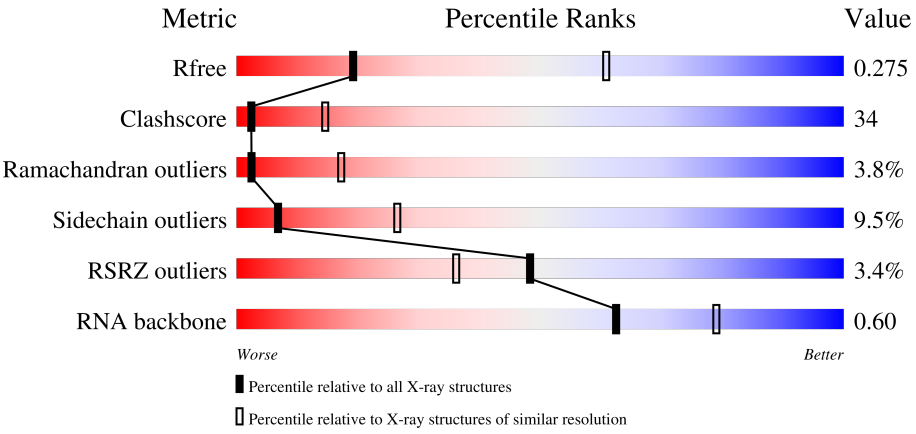
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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 3.34 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}	164625	1325 (3.38-3.30)
Clashscore	180529	1376 (3.38-3.30)
Ramachandran outliers	177936	1376 (3.38-3.30)
Sidechain outliers	177891	1375 (3.38-3.30)
RSRZ outliers	164620	1325 (3.38-3.30)
RNA backbone	3690	1003 (3.70-2.98)

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	2% 31% 44% 5% 19%
2	B	1224	4% 34% 47% 6% 13%
3	C	318	% 42% 36% 5% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	221	2% 36% 29% 7% 27%
5	E	215	5% 37% 51% 7%
6	F	155	% 19% 32% 45%
7	G	171	% 47% 46% 6%
8	H	146	8% 24% 45% 11% 20%
9	I	122	7% 41% 45% 7% 7%
10	J	70	% 40% 39% 14% 7%
11	K	120	% 49% 41% 5% 5%
12	L	70	9% 11% 27% 20% 39%
13	R	10	40% 60%
14	T	14	7% 43% 57%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
18	ATP	A	1806	-	-	X	-

2 Entry composition (i)

There are 19 unique types of molecules in this entry. The entry contains 30677 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 subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1396	10967	6917	1917	2072	61	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1060	8420	5341	1476	1548	55	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	265	2083	1310	346	414	13	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	161	1274	790	223	259	2	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	206	1693	1078	298	310	7	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	85	684	437	116	128	3	0	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	117	Total	C	N	O	S	0	0	0
			951	605	158	184	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	113	Total	C	N	O	S	0	0	0
			910	557	166	177	10			

- Molecule 10 is a protein called DNA-directed RNA polymerases II subunit RPABC5.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 12 is a protein called DNA-directed RNA polymerases II subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	43	Total	C	N	O	S	0	0	0
			343	211	69	59	4			

- Molecule 13 is a RNA chain called RNA (5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	R	10	Total	C	N	O	P	0	0	0
			217	98	45	65	9			

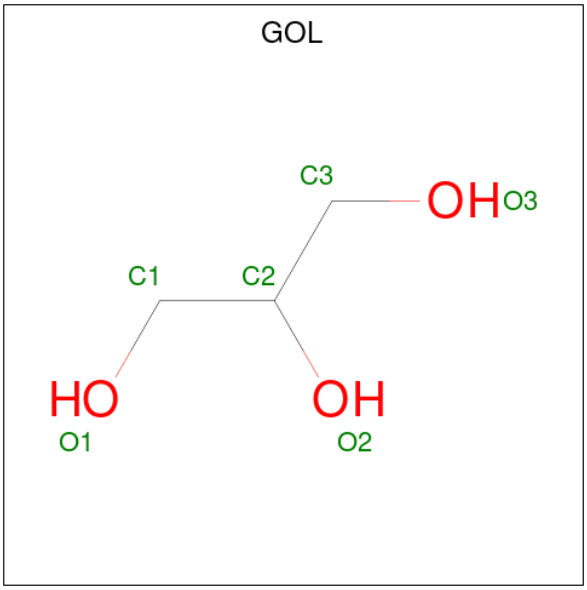
- Molecule 14 is a DNA chain called DNA (5'-D(P*CP*AP*CP*GP*TP*CP*CP*CP*TP*CP*TP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	T	14	Total	C	N	O	P	0	0	0
			279	133	47	85	14			

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 16 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

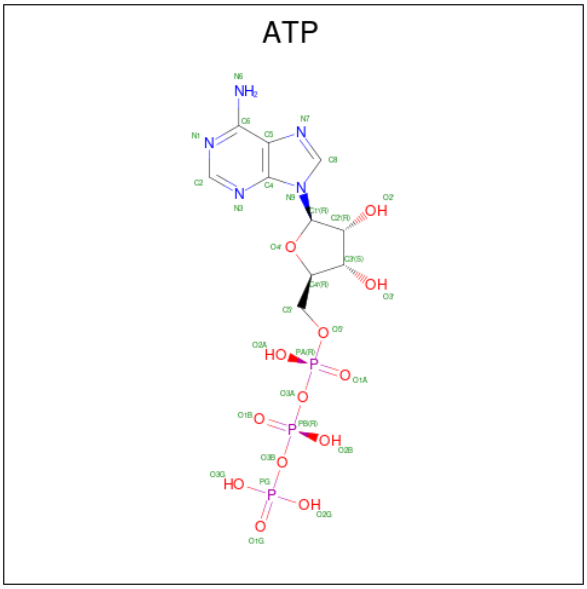


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
16	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	A	2	Total Mg 2 2	0	0
17	B	1	Total Mg 1 1	0	0

- Molecule 18 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	A	1	Total C N O P 31 10 5 13 3	0	0

- Molecule 19 is water.

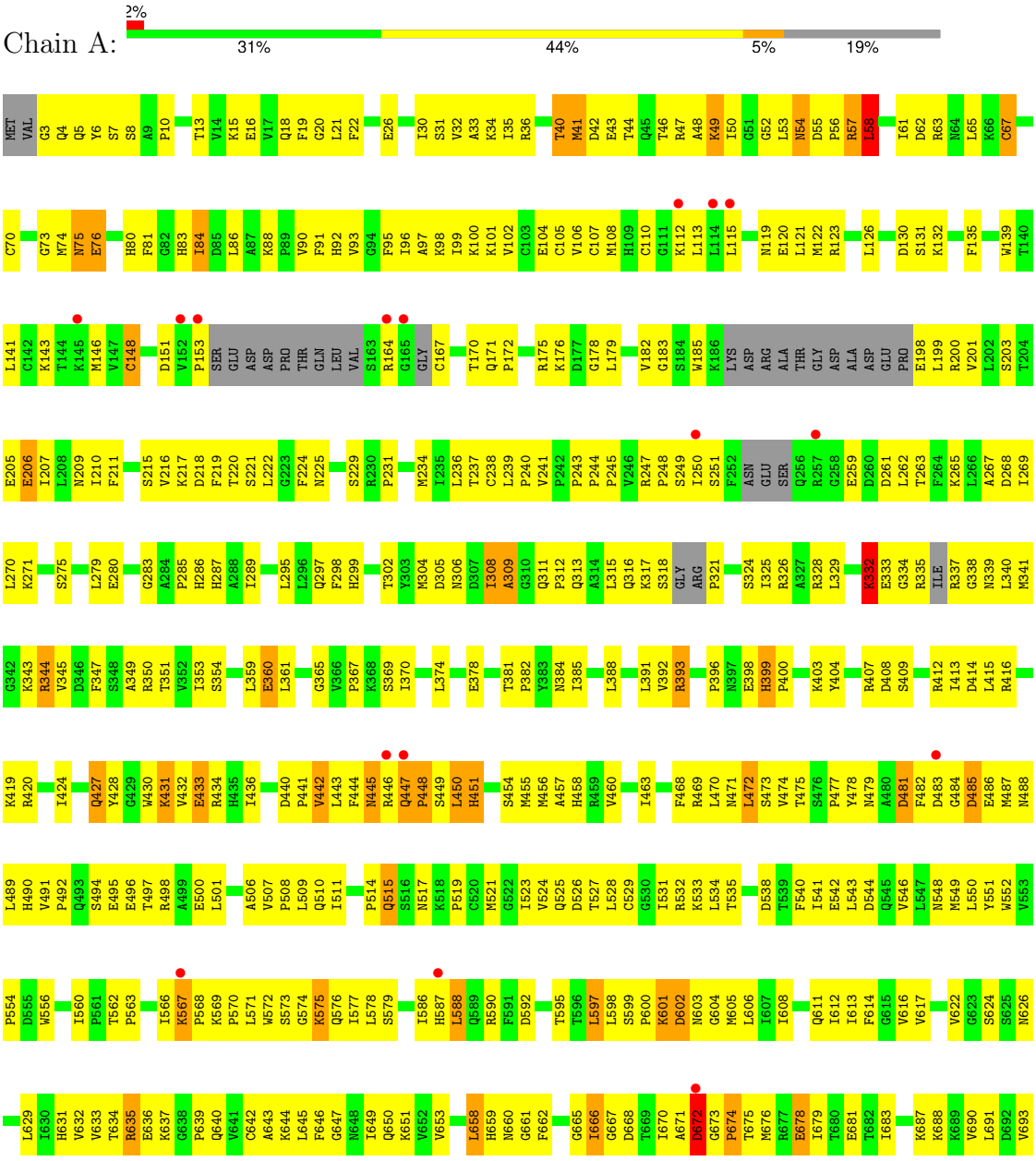
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
19	A	11	Total O 11 11	0	0
19	B	3	Total O 3 3	0	0
19	D	1	Total O 1 1	0	0
19	G	1	Total O 1 1	0	0
19	K	1	Total O 1 1	0	0

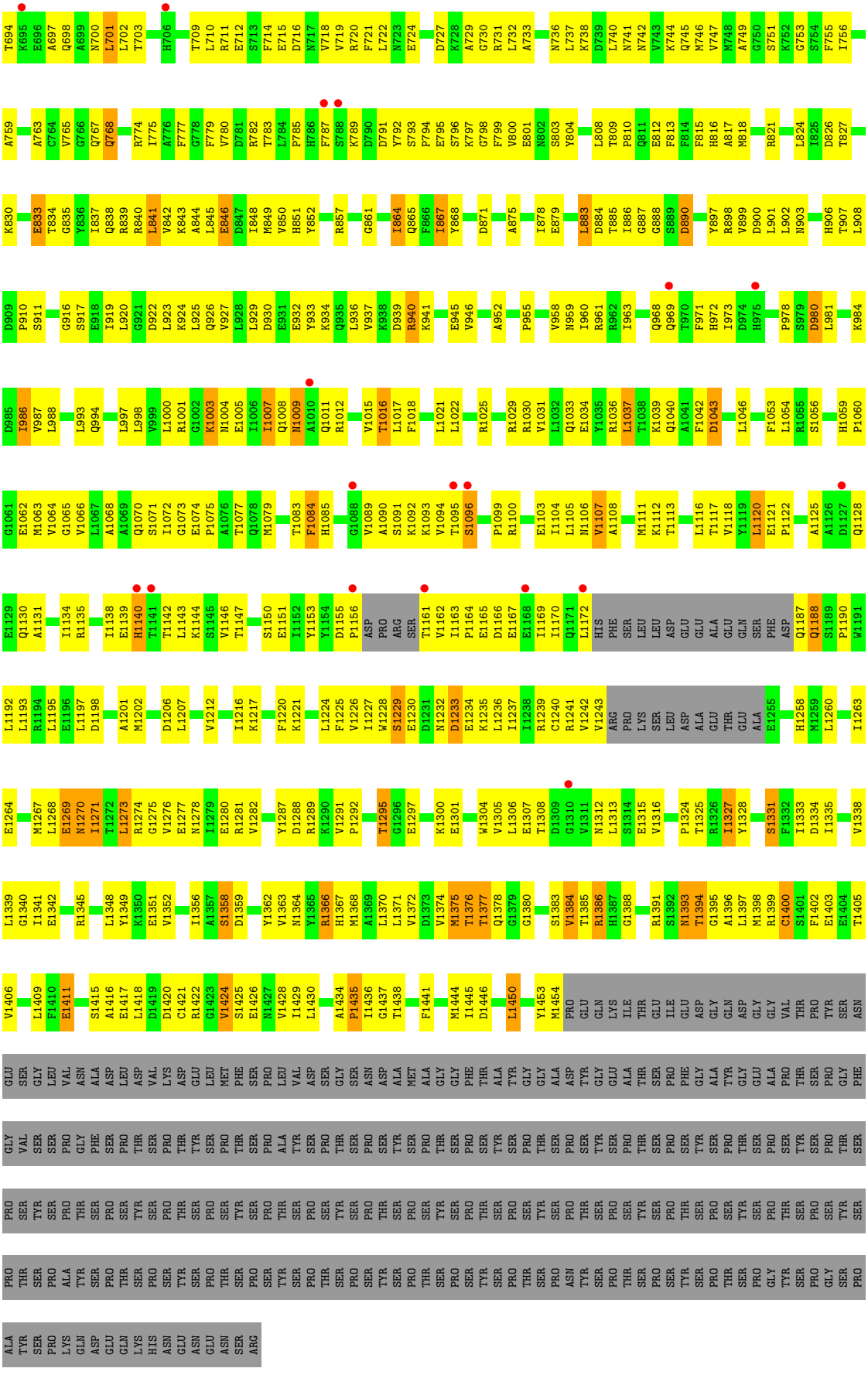
3 Residue-property plots (i)

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 subunit RPB1

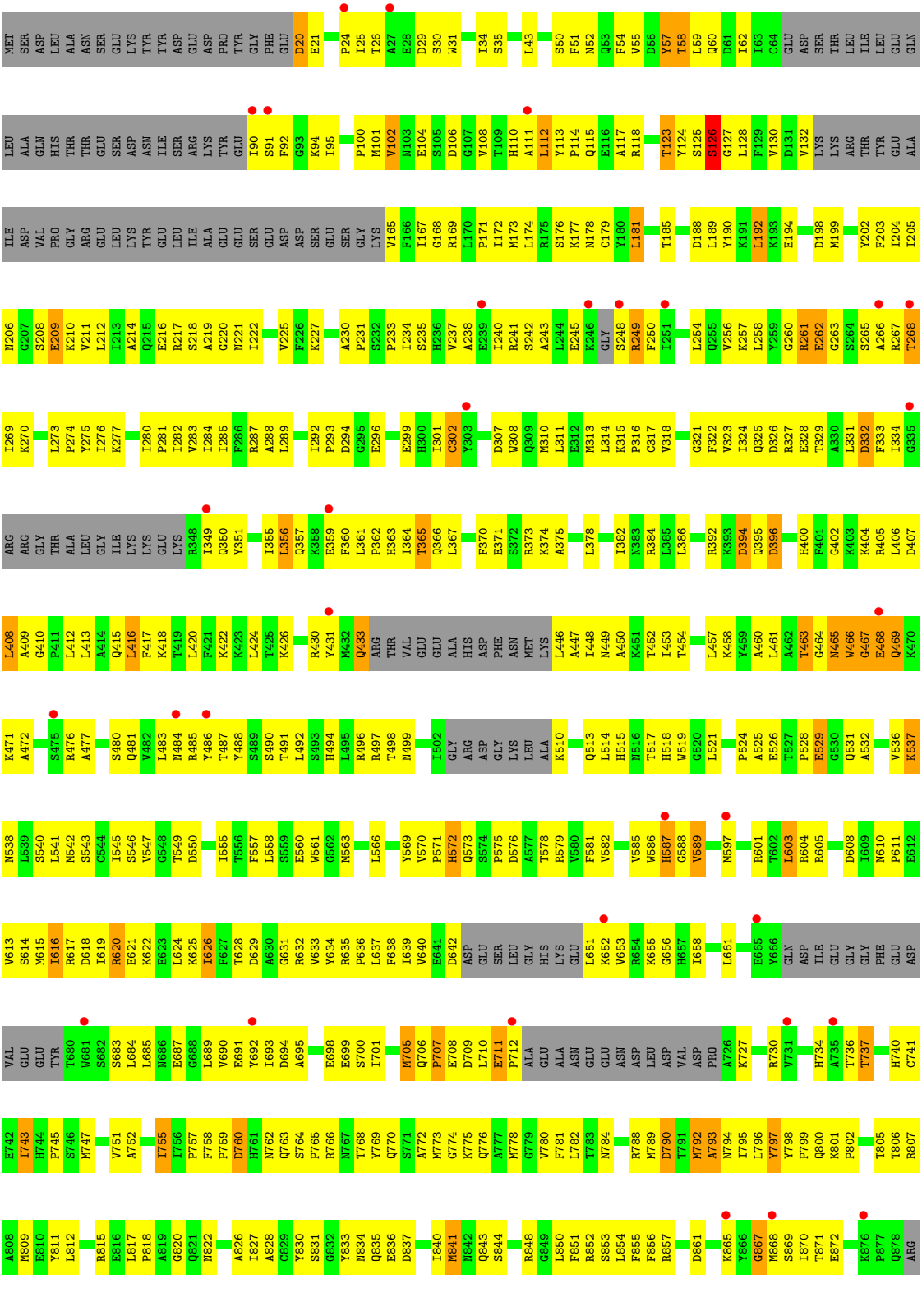
Chain A:

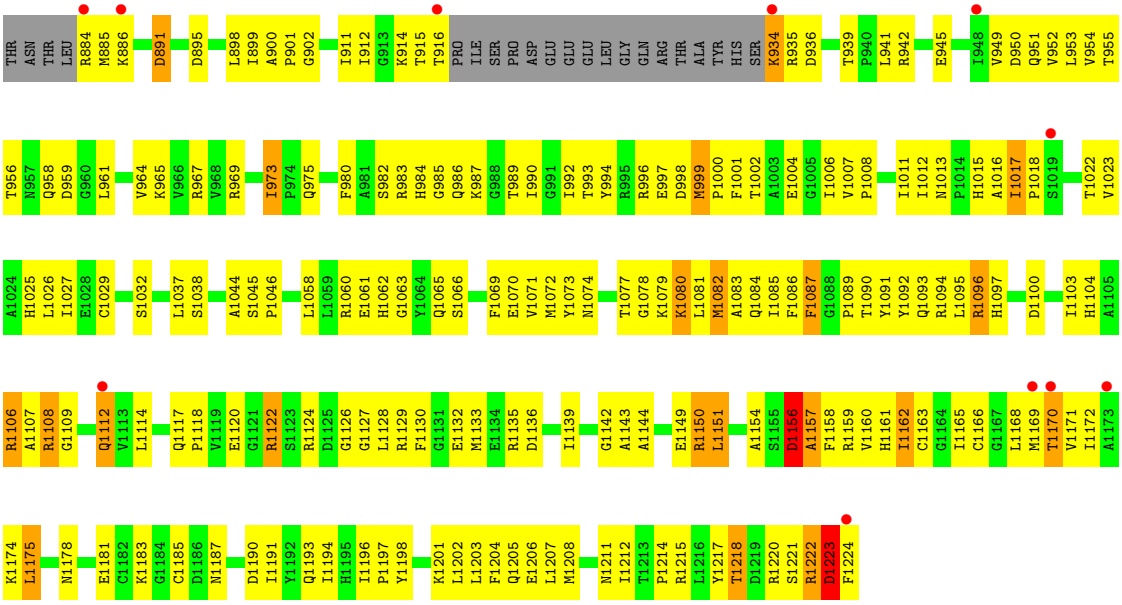




• Molecule 2: DNA-directed RNA polymerase subunit beta

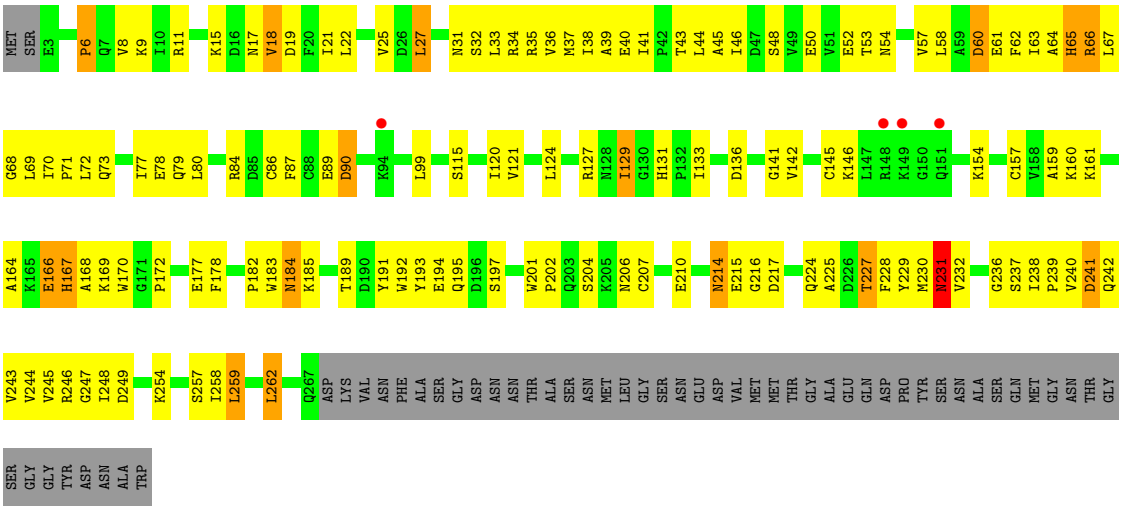
Chain B:





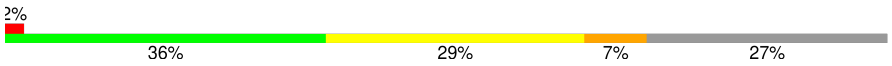
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

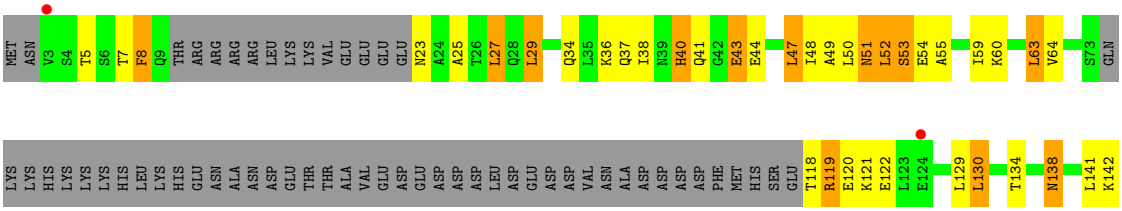
Chain C:

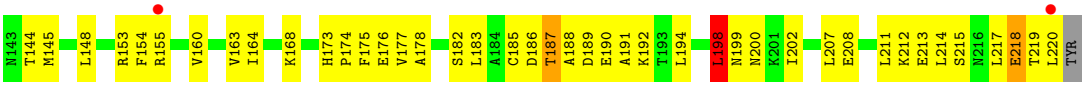


• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:

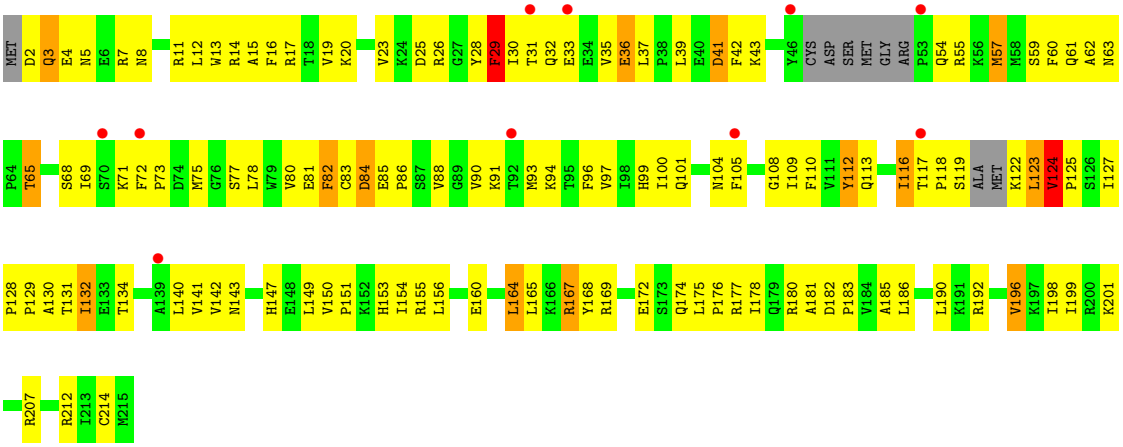






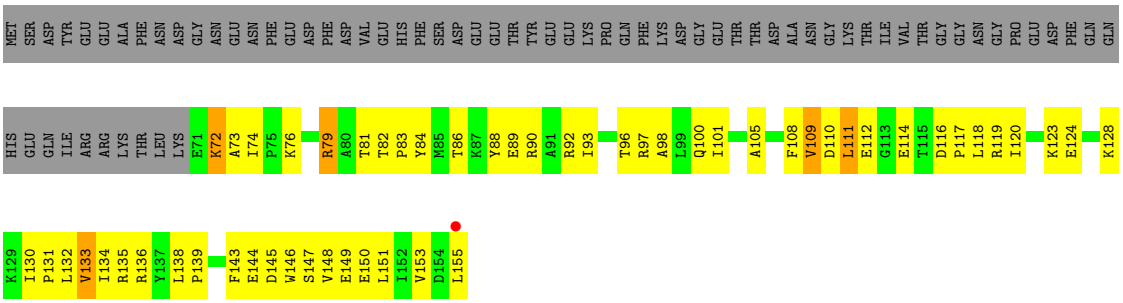
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



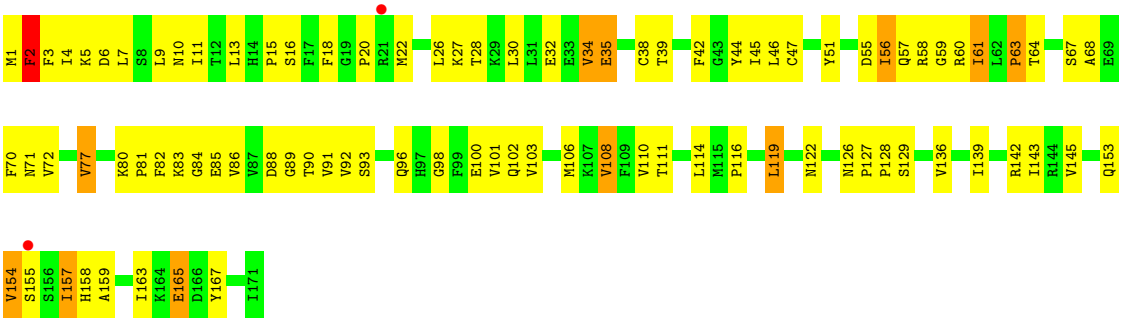
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



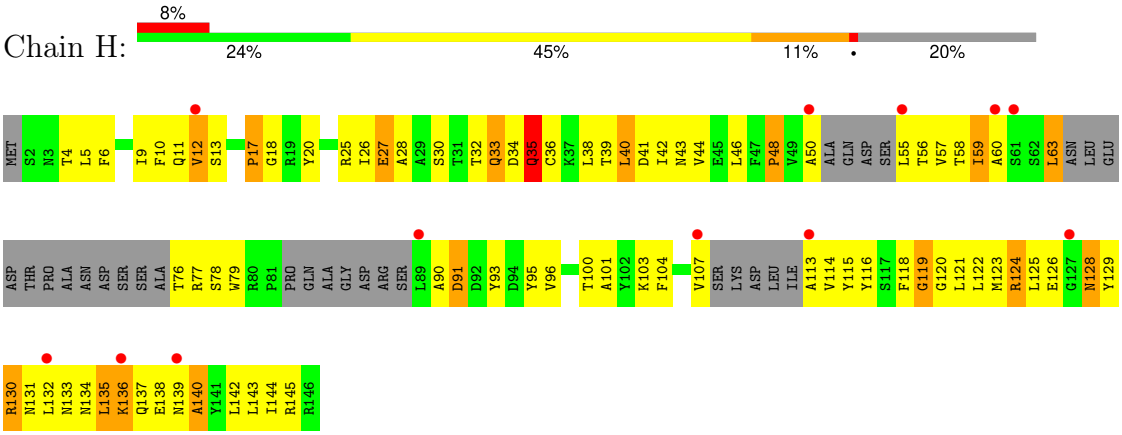
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:



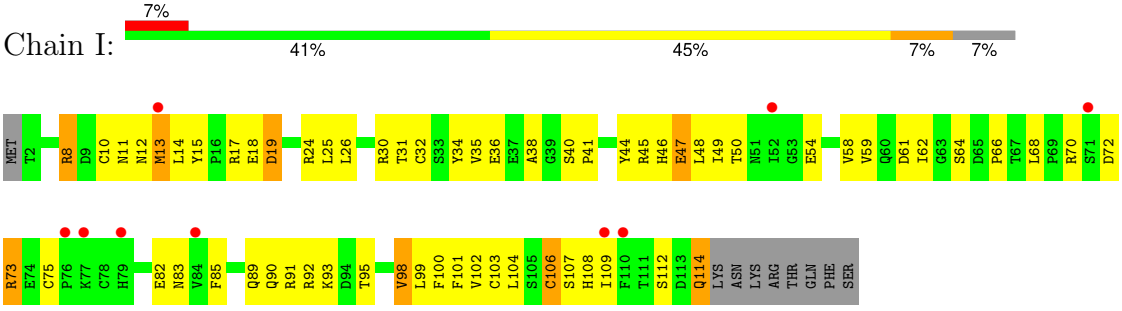
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



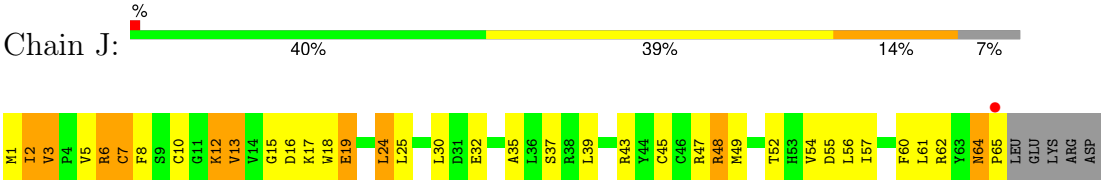
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



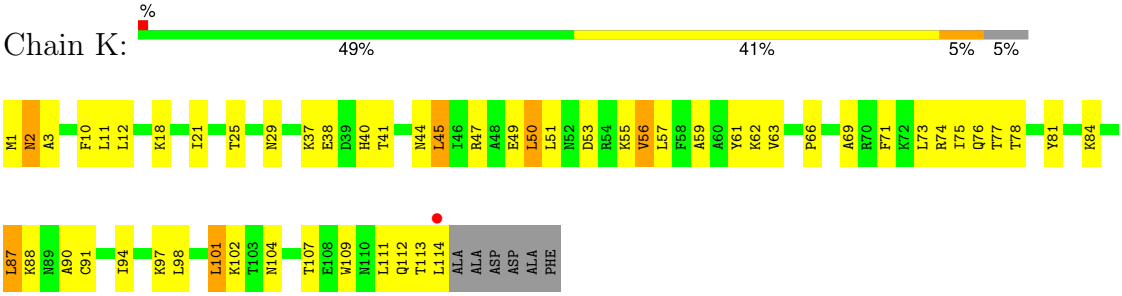
• Molecule 10: DNA-directed RNA polymerases II subunit RPABC5

Chain J:



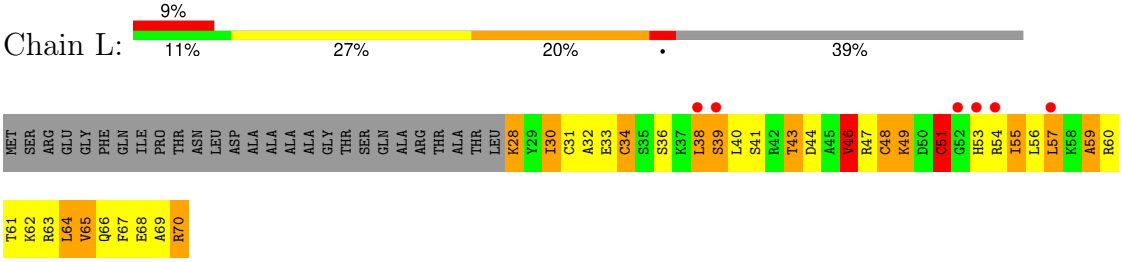
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:



• Molecule 12: DNA-directed RNA polymerases II subunit RPABC4

Chain L:

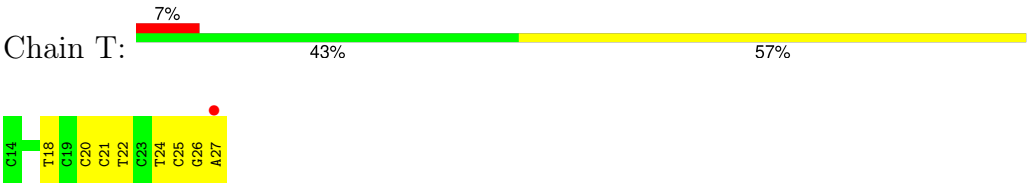


• Molecule 13: RNA (5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*G)-3')

Chain R:

• Molecule 14: DNA (5'-D(P*CP*AP*CP*GP*TP*CP*CP*CP*TP*CP*TP*CP*GP*A)-3')

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	217.19Å 387.44Å 276.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.97 – 3.34 19.97 – 3.34	Depositor EDS
% Data completeness (in resolution range)	98.7 (19.97-3.34) 99.4 (19.97-3.34)	Depositor EDS
R_{merge}	0.60	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, R_{free}	0.255 , 0.261 0.267 , 0.275	Depositor DCC
R_{free} test set	8262 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	80.0	Xtriage
Anisotropy	0.341	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 53.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.31$, $< L^2 > = 0.16$	Xtriage
Estimated twinning fraction	0.146 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.160 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	30677	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.05% 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: GOL, ZN, MG, ATP

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	0.26	0/11158	0.49	0/15080
2	B	0.27	0/8583	0.51	0/11570
3	C	0.27	0/2121	0.46	0/2876
4	D	0.24	0/1282	0.46	0/1723
5	E	0.25	0/1727	0.51	0/2323
6	F	0.26	0/696	0.50	0/940
7	G	0.25	0/1368	0.46	0/1844
8	H	0.24	0/965	0.50	0/1302
9	I	0.24	0/927	0.53	0/1250
10	J	0.27	0/541	0.52	0/727
11	K	0.25	0/937	0.44	0/1265
12	L	0.27	0/345	0.59	0/457
13	R	0.24	0/244	0.77	0/380
14	T	0.65	0/310	0.92	0/474
All	All	0.27	0/31204	0.50	0/42211

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10967	0	11024	910	0
2	B	8420	0	8456	633	0
3	C	2083	0	2036	144	0
4	D	1274	0	1287	82	0
5	E	1693	0	1715	115	0
6	F	684	0	703	60	0
7	G	1340	0	1357	97	0
8	H	951	0	926	96	0
9	I	910	0	857	53	0
10	J	532	0	542	44	0
11	K	919	0	929	48	0
12	L	343	0	363	42	0
13	R	217	0	109	4	0
14	T	279	0	158	11	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
16	A	6	0	8	0	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
18	A	31	0	12	9	0
19	A	11	0	0	0	0
19	B	3	0	0	0	0
19	D	1	0	0	0	0
19	G	1	0	0	2	0
19	K	1	0	0	0	0
All	All	30677	0	30482	2097	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 34.

The worst 5 of 2097 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:1160:VAL:HG11	2:B:1169:MET:SD	1.66	1.34
1:A:393:ARG:HG2	1:A:393:ARG:HH11	1.14	1.08
2:B:1163:CYS:HB3	2:B:1166:CYS:SG	1.96	1.04
1:A:393:ARG:HH11	1:A:393:ARG:CG	1.72	1.02
2:B:510:LYS:HB2	2:B:513:GLN:HG2	1.42	1.02

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	1376/1733 (79%)	1168 (85%)	166 (12%)	42 (3%)	3 21
2	B	1036/1224 (85%)	854 (82%)	141 (14%)	41 (4%)	2 16
3	C	263/318 (83%)	223 (85%)	29 (11%)	11 (4%)	2 15
4	D	155/221 (70%)	134 (86%)	14 (9%)	7 (4%)	2 14
5	E	200/215 (93%)	178 (89%)	17 (8%)	5 (2%)	4 25
6	F	83/155 (54%)	76 (92%)	5 (6%)	2 (2%)	5 25
7	G	169/171 (99%)	137 (81%)	24 (14%)	8 (5%)	2 13
8	H	107/146 (73%)	80 (75%)	18 (17%)	9 (8%)	0 5
9	I	111/122 (91%)	85 (77%)	24 (22%)	2 (2%)	7 31
10	J	63/70 (90%)	53 (84%)	7 (11%)	3 (5%)	2 13
11	K	112/120 (93%)	96 (86%)	13 (12%)	3 (3%)	4 23
12	L	41/70 (59%)	19 (46%)	15 (37%)	7 (17%)	0 1
All	All	3716/4565 (81%)	3103 (84%)	473 (13%)	140 (4%)	2 17

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ARG
1	A	433	GLU
1	A	567	LYS
1	A	672	ASP
1	A	674	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	1213/1520 (80%)	1113 (92%)	100 (8%)	9 32
2	B	917/1061 (86%)	842 (92%)	75 (8%)	9 32
3	C	232/274 (85%)	219 (94%)	13 (6%)	17 46
4	D	143/200 (72%)	128 (90%)	15 (10%)	5 22
5	E	189/197 (96%)	158 (84%)	31 (16%)	2 8
6	F	74/137 (54%)	64 (86%)	10 (14%)	3 14
7	G	152/152 (100%)	137 (90%)	15 (10%)	6 25
8	H	104/128 (81%)	90 (86%)	14 (14%)	3 14
9	I	105/116 (90%)	91 (87%)	14 (13%)	3 14
10	J	60/65 (92%)	51 (85%)	9 (15%)	2 11
11	K	99/102 (97%)	90 (91%)	9 (9%)	7 28
12	L	38/57 (67%)	26 (68%)	12 (32%)	0 1
All	All	3326/4009 (83%)	3009 (90%)	317 (10%)	7 26

5 of 317 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	182	ASP
9	I	114	GLN
6	F	118	LEU
8	H	33	GLN
11	K	87	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	776	GLN
2	B	1062	HIS
2	B	1025	HIS
2	B	1084	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	736	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	R	8/10 (80%)	2 (25%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	R	2	U
13	R	3	C

There are no RNA pucker outliers to report.

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 13 ligands modelled in this entry, 11 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	GOL	A	1803	-	5,5,5	1.02	0	5,5,5	1.03	0
18	ATP	A	1806	17	28,33,33	0.81	0	34,52,52	0.97	3 (8%)

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	GOL	A	1803	-	-	2/4/4/4	-
18	ATP	A	1806	17	-	6/18/38/38	0/3/3/3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	A	1806	ATP	O2A-PA-O3A	2.77	114.77	107.27
18	A	1806	ATP	O3A-PA-O1A	-2.42	103.41	110.70
18	A	1806	ATP	C5-C6-N6	2.32	123.84	120.31

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
18	A	1806	ATP	C5'-O5'-PA-O2A
18	A	1806	ATP	C5'-O5'-PA-O3A
18	A	1806	ATP	C3'-C4'-C5'-O5'
18	A	1806	ATP	O4'-C4'-C5'-O5'
16	A	1803	GOL	C1-C2-C3-O3

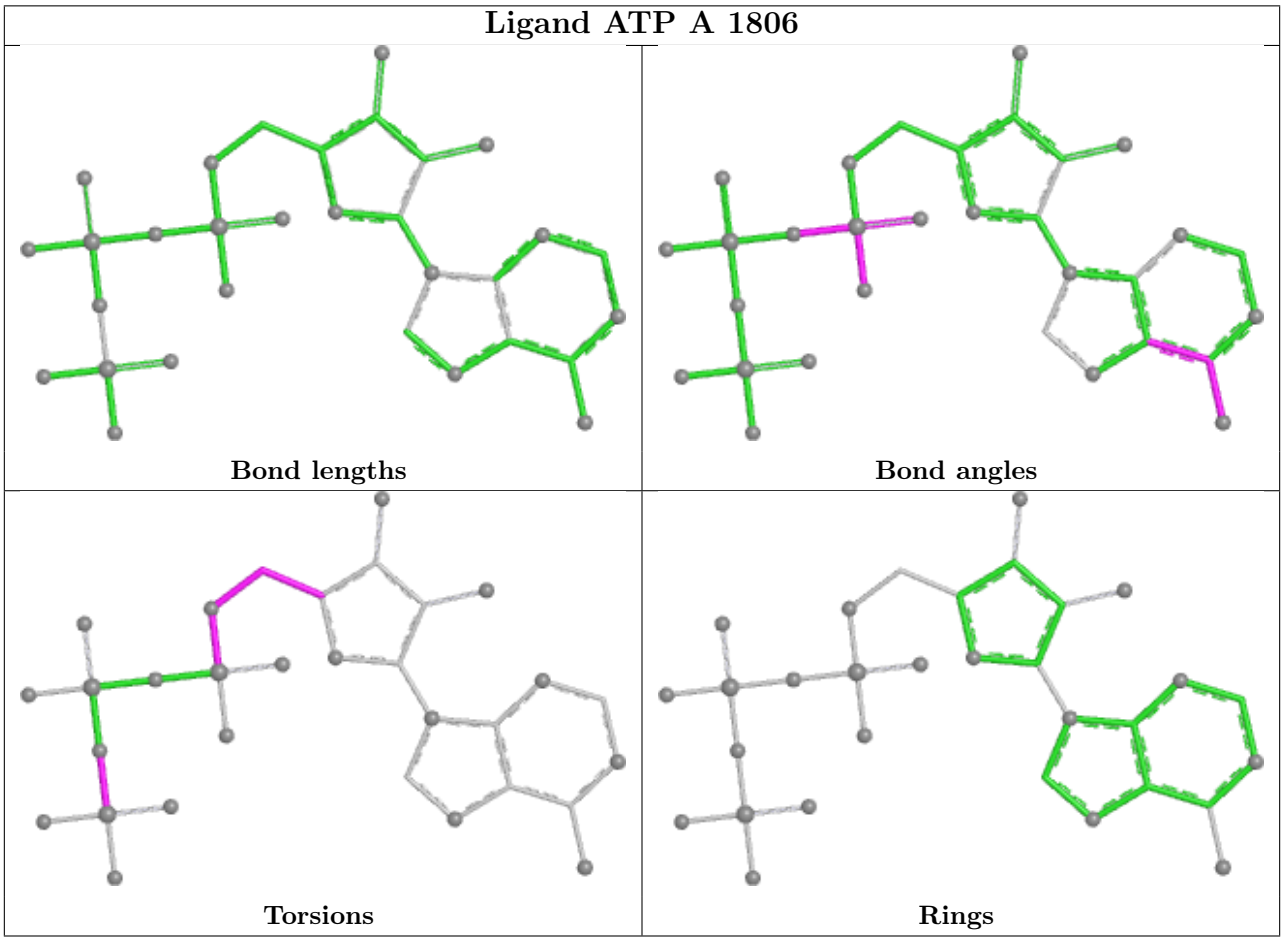
There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
18	A	1806	ATP	9	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier.

The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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, 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	1396/1733 (80%)	0.04	34 (2%) 59 46	40, 92, 158, 219	0
2	B	1060/1224 (86%)	0.16	43 (4%) 42 31	30, 99, 165, 276	0
3	C	265/318 (83%)	-0.01	4 (1%) 71 60	60, 98, 143, 188	0
4	D	161/221 (72%)	0.10	4 (2%) 58 45	74, 107, 159, 212	0
5	E	206/215 (95%)	0.38	10 (4%) 36 28	61, 124, 184, 234	0
6	F	85/155 (54%)	-0.29	1 (1%) 76 65	46, 73, 106, 132	0
7	G	171/171 (100%)	-0.12	2 (1%) 76 65	70, 96, 150, 175	0
8	H	117/146 (80%)	0.61	12 (10%) 13 13	95, 135, 180, 221	0
9	I	113/122 (92%)	0.73	9 (7%) 20 18	99, 139, 200, 222	0
10	J	65/70 (92%)	-0.08	1 (1%) 71 60	70, 95, 138, 181	0
11	K	114/120 (95%)	-0.23	1 (0%) 81 71	57, 97, 143, 170	0
12	L	43/70 (61%)	1.03	6 (13%) 7 8	80, 128, 180, 197	0
13	R	10/10 (100%)	0.40	0 100 100	98, 113, 170, 205	0
14	T	14/14 (100%)	0.24	1 (7%) 23 20	86, 106, 203, 205	0
All	All	3820/4589 (83%)	0.11	128 (3%) 48 36	30, 99, 166, 276	0

The worst 5 of 128 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	114	LEU	7.2
9	I	110	PHE	6.1
9	I	76	PRO	5.5
2	B	916	THR	5.3
2	B	246	LYS	5.3

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

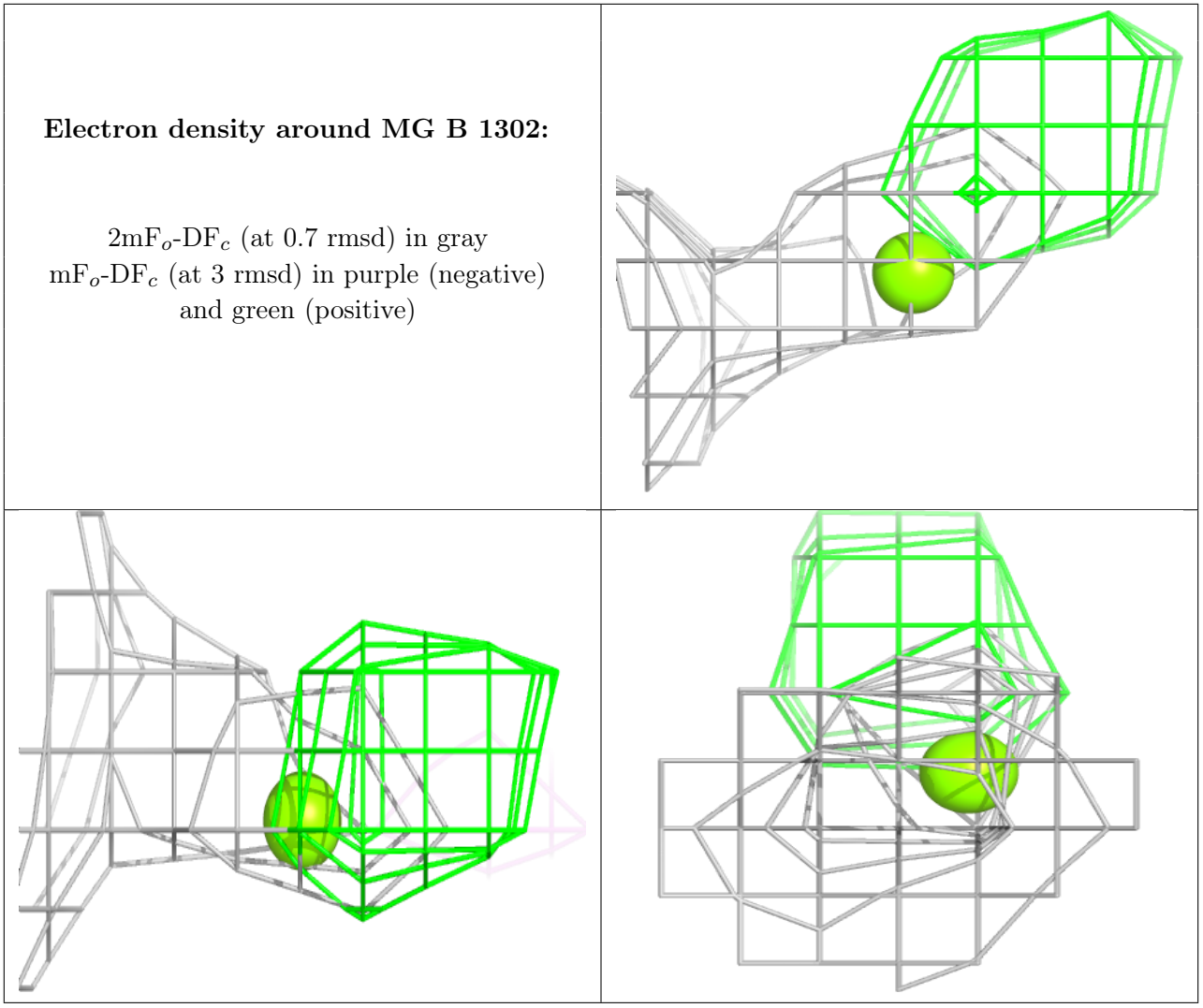
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
15	ZN	A	1801	1/1	0.79	0.11	213,213,213,213	0
17	MG	B	1302	1/1	0.81	0.76	92,92,92,92	0
17	MG	A	1805	1/1	0.83	0.15	105,105,105,105	0
15	ZN	L	101	1/1	0.90	0.08	173,173,173,173	0
18	ATP	A	1806	31/31	0.92	0.12	112,118,133,155	0
17	MG	A	1804	1/1	0.94	0.10	94,94,94,94	0
16	GOL	A	1803	6/6	0.96	0.14	103,110,116,118	0
15	ZN	A	1802	1/1	0.97	0.03	56,56,56,56	0
15	ZN	I	201	1/1	0.99	0.12	103,103,103,103	0
15	ZN	I	202	1/1	0.99	0.04	181,181,181,181	0
15	ZN	B	1301	1/1	0.99	0.03	61,61,61,61	0
15	ZN	C	401	1/1	0.99	0.04	69,69,69,69	0
15	ZN	J	101	1/1	1.00	0.02	79,79,79,79	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

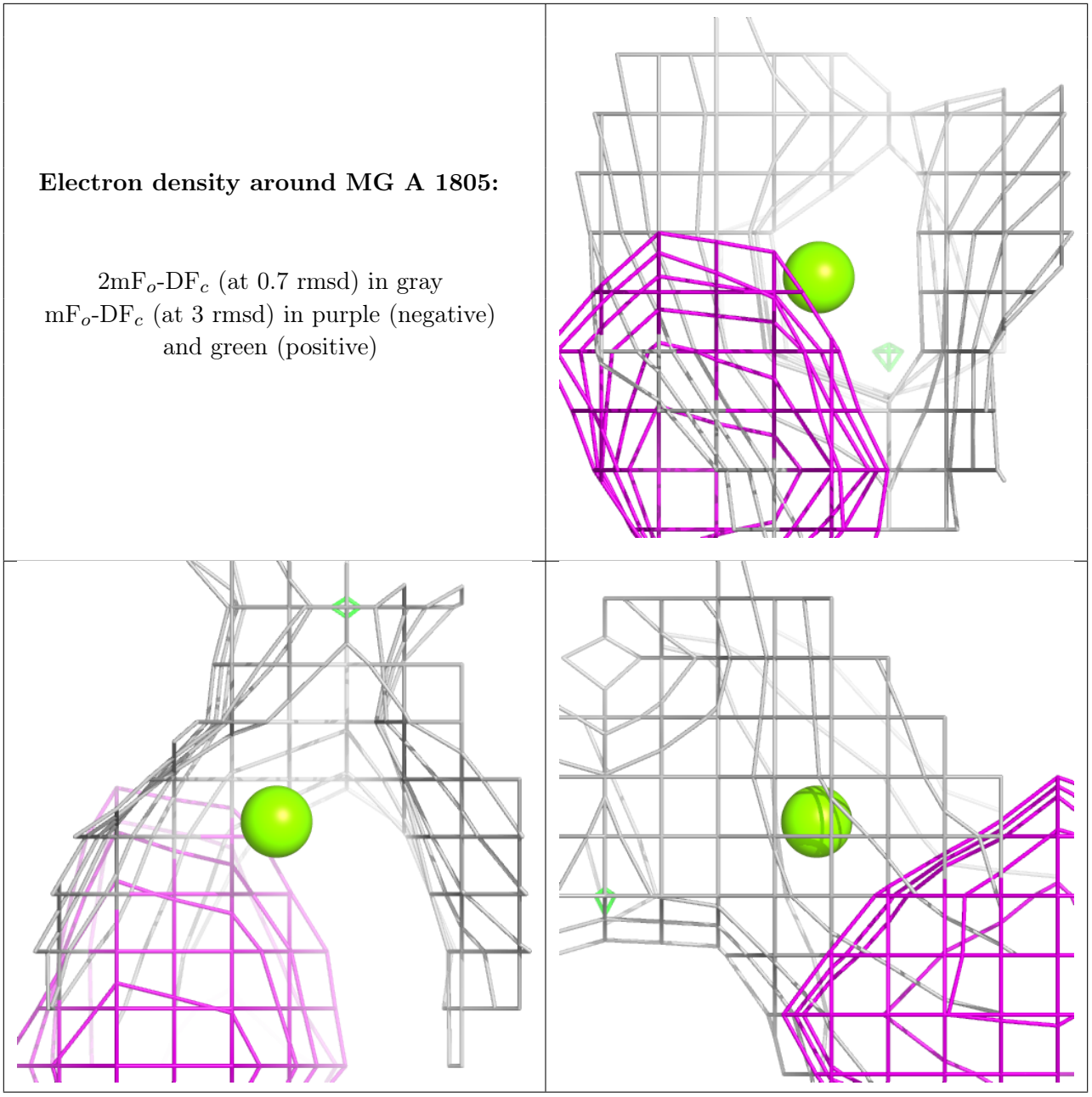
Electron density around MG B 1302:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



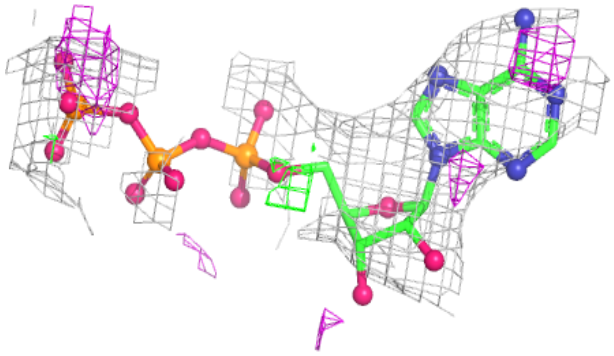
Electron density around MG A 1805:

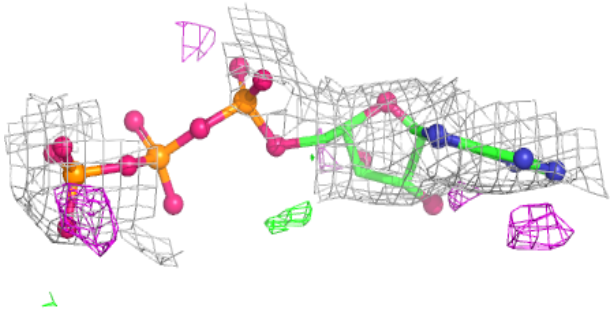
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

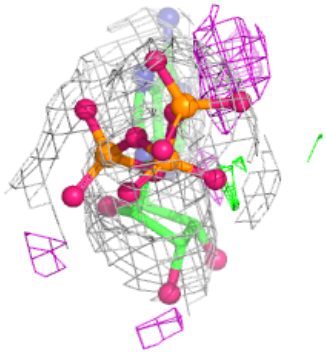


Electron density around ATP A 1806:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

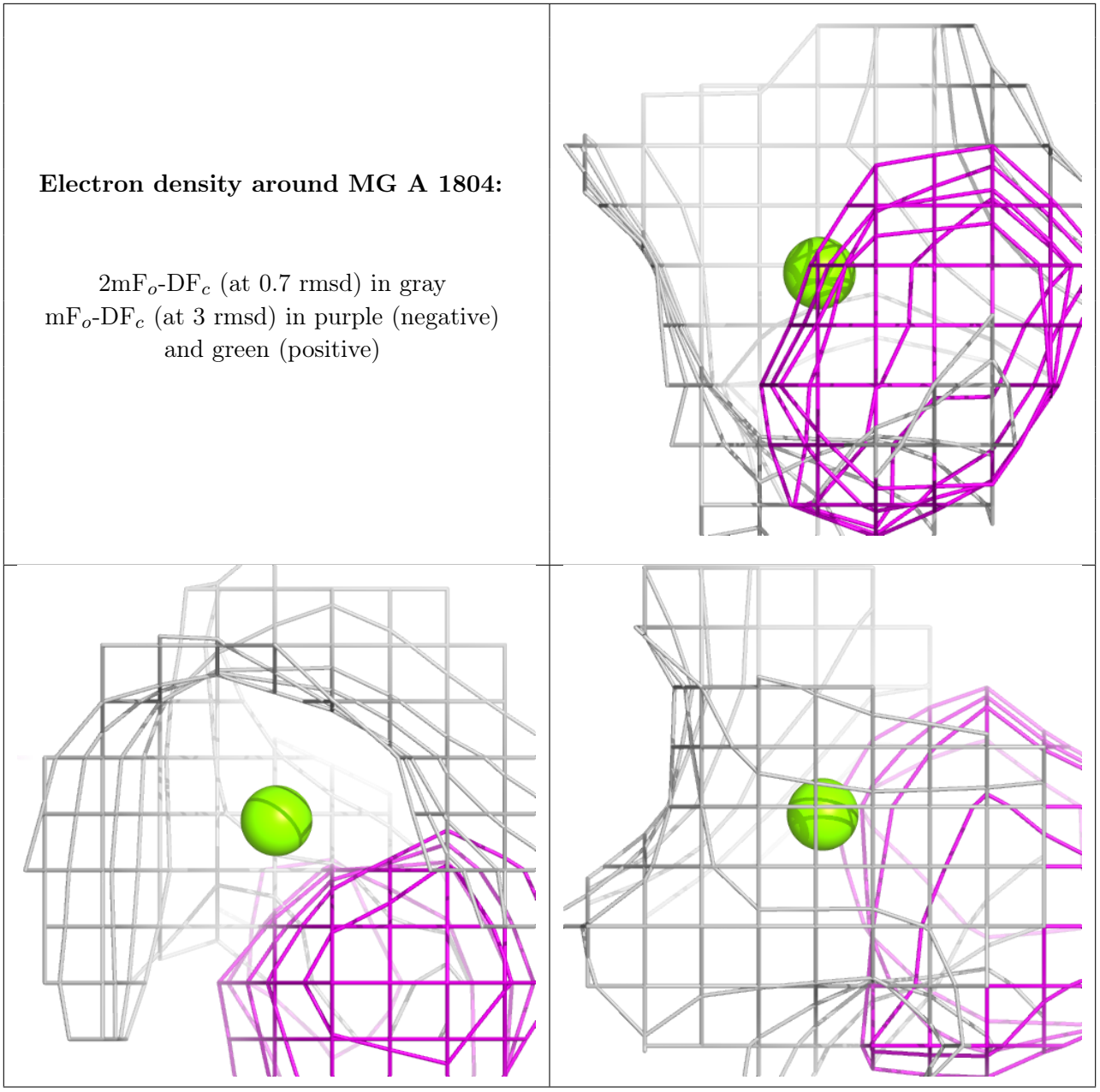






Electron density around MG A 1804:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.