wwPDB X-ray Structure Validation Summary Report (i)

Sep 26, 2023 – 05:20 PM EDT

PDB ID	:	6CCV
Title	:	Crystal structure of a Mycobacterium smegmatis RNA polymerase transcription initiation complex with inhibitor Rifampicin
Authors	:	Lilic, M.; Darst, S.A.; Campbell, E.A.
Deposited on	:	2018-02-07
Resolution	:	3.05 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:

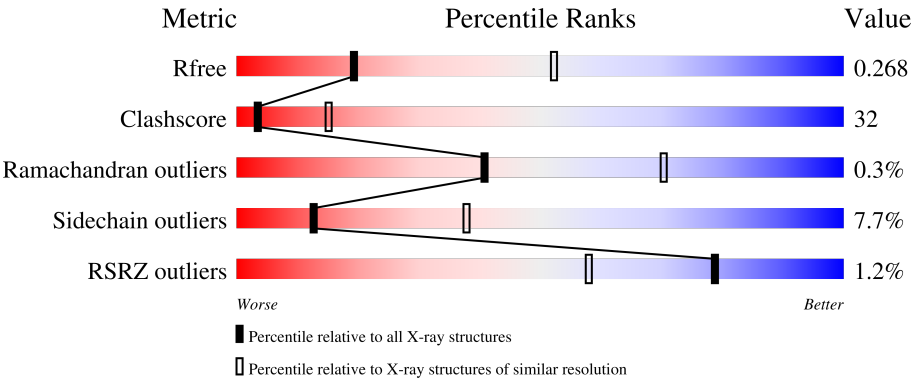
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.05 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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	350	28% 33% · 38%
1	B	350	% 27% 37% · 33%
1	T	350	3% 12% · 85%
2	C	1169	% 46% 44% · 6%
3	D	1317	50% 40% · 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	107	37% 26% 7% 29%
5	F	466	36% 27% 35%
6	G	17	100%
7	J	114	2% 37% 32% 27%
8	O	31	32% 68%
9	P	26	62% 38%

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
10	SO4	D	2004	-	-	X	-
11	EDO	F	506	-	-	-	X
14	GLU	D	2012	-	-	X	-

2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 26582 atoms, of which 48 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	0	0
			1605	1015	276	311	3			
1	B	233	Total	C	N	O	S	0	0	0
			1667	1054	289	322	2			
1	T	53	Total	C	N	O	S	0	0	0
			342	208	65	68	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1099	Total	C	N	O	S	0	0	0
			8262	5174	1450	1603	35			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1246	Total	C	N	O	S	0	0	0
			9555	5995	1720	1800	40			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	76	Total	C	N	O	0	0	0
			592	378	100	114			

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	305	Total	C	N	O	S	0	0	0
			2414	1512	436	459	7			

- Molecule 6 is a protein called Unknown Peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	G	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 7 is a protein called RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	J	83	Total	C	N	O	S	0	0	0
			671	422	119	128	2			

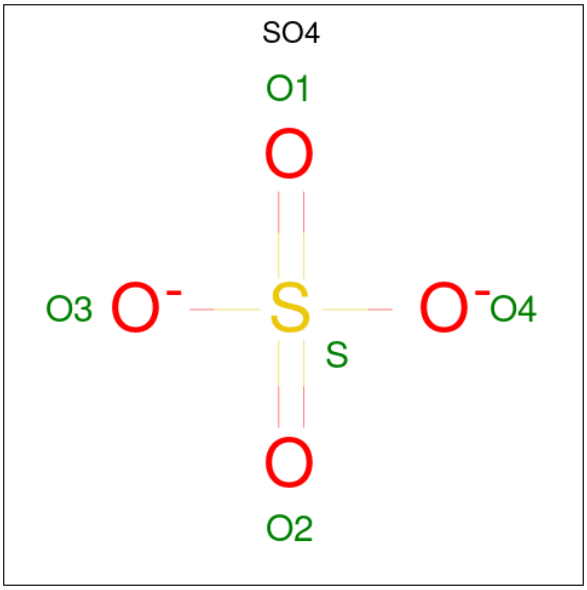
- Molecule 8 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	O	31	Total	C	N	O	P	0	0	0
			634	305	114	185	30			

- Molecule 9 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	P	26	Total	C	N	O	P	0	0	0
			526	254	94	153	25			

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



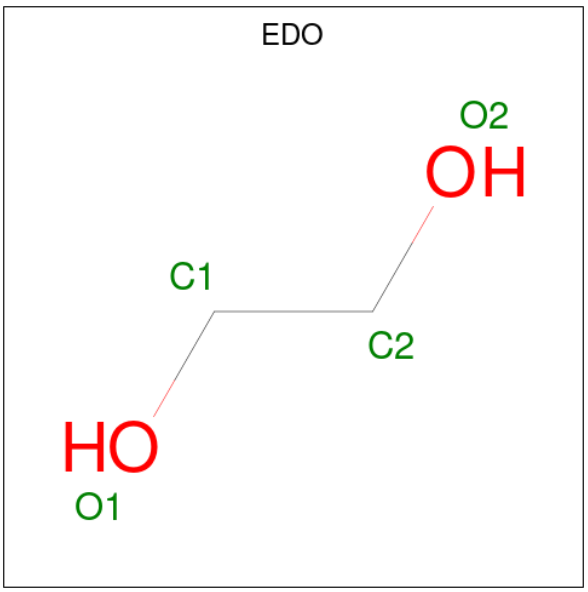
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

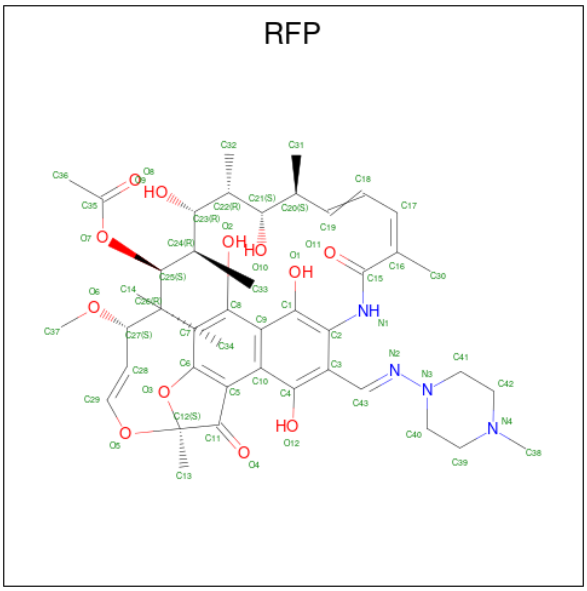
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	C	1	Total O S 5 4 1	0	0
10	C	1	Total O S 5 4 1	0	0
10	C	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0
10	F	1	Total O S 5 4 1	0	0
10	F	1	Total O S 5 4 1	0	0
10	F	1	Total O S 5 4 1	0	0
10	F	1	Total O S 5 4 1	0	0

- Molecule 11 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	C	1	Total C H O 10 2 6 2	0	0
11	C	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	F	1	Total C H O 10 2 6 2	0	0
11	F	1	Total C H O 10 2 6 2	0	0

- Molecule 12 is RIFAMPICIN (three-letter code: RFP) (formula: $C_{43}H_{58}N_4O_{12}$).

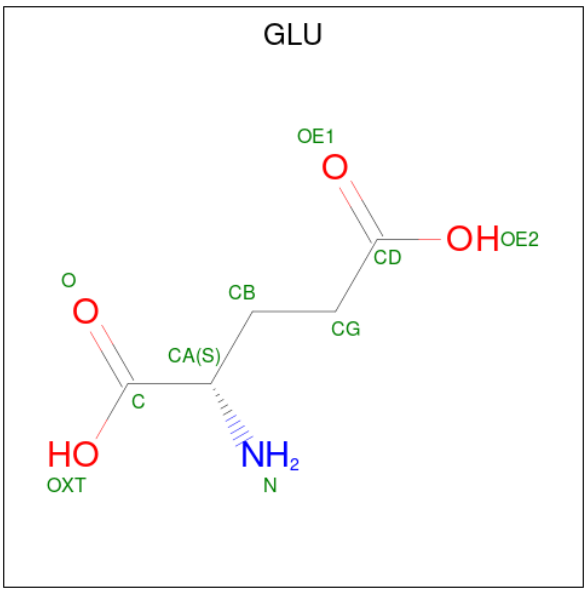


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	C	1	Total C N O 59 43 4 12	0	0

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	D	2	Total Zn 2 2	0	0

- Molecule 14 is GLUTAMIC ACID (three-letter code: GLU) (formula: $C_5H_9NO_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	D	1	Total C N O 9 5 1 3	0	0

- Molecule 15 is water.

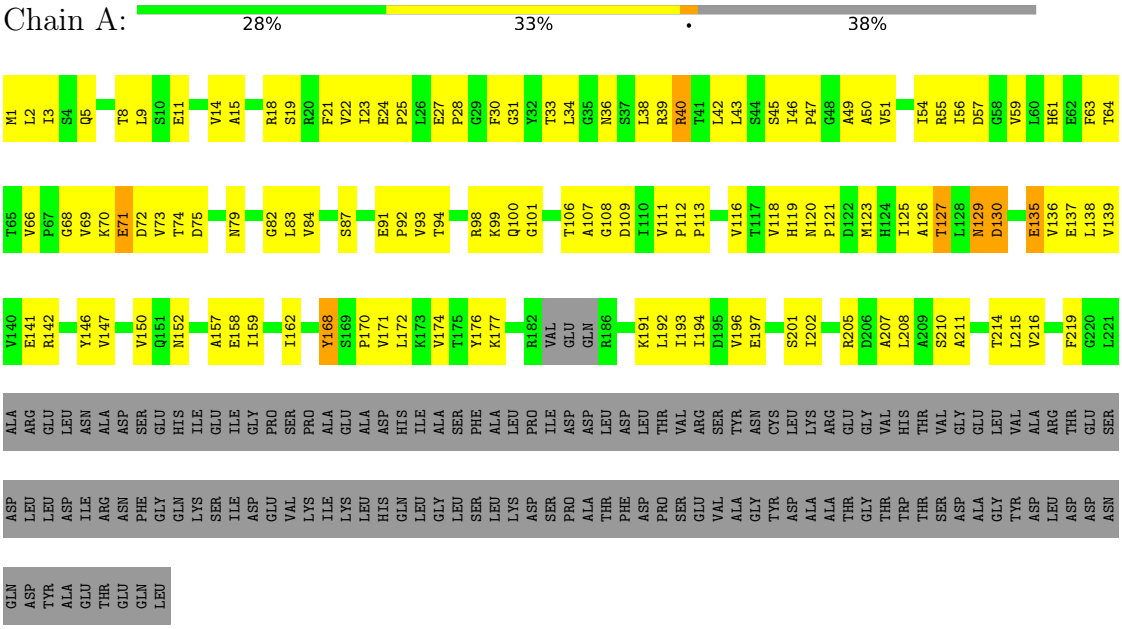
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	C	5	Total O 5 5	0	0
15	D	8	Total O 8 8	0	0
15	F	1	Total O 1 1	0	0

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.

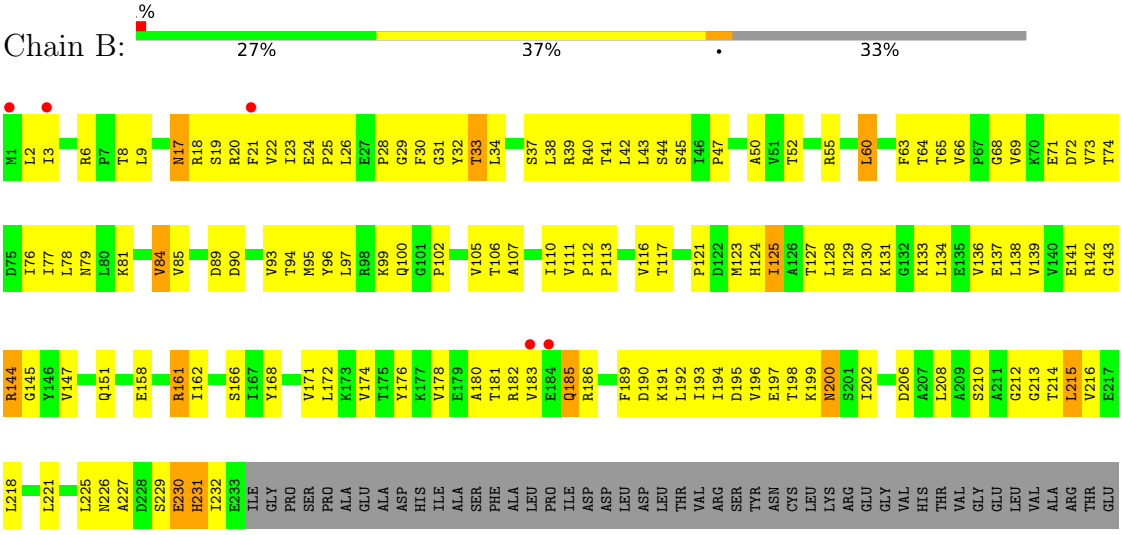
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:

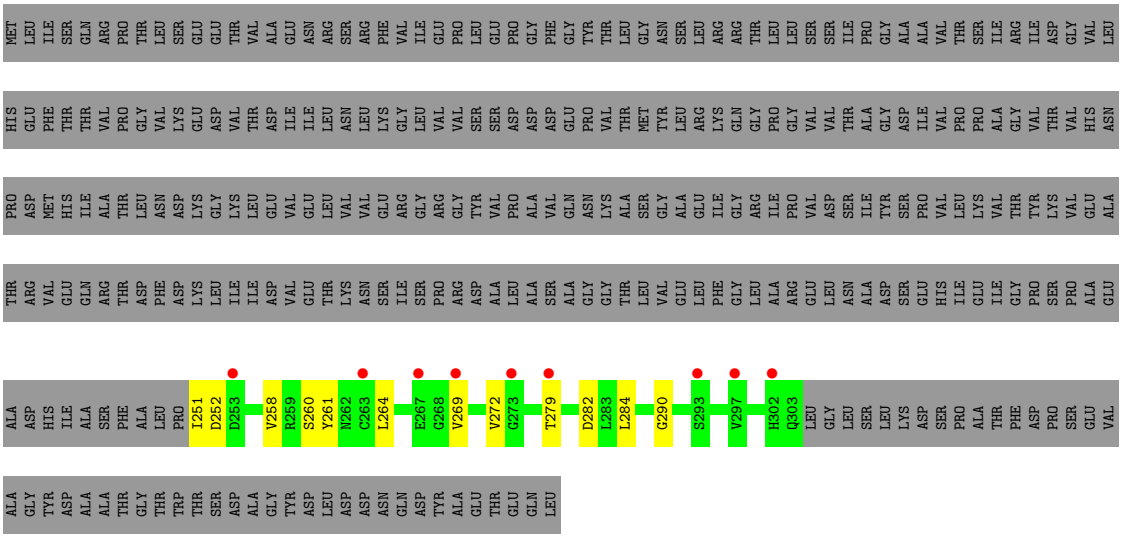


• Molecule 1: DNA-directed RNA polymerase subunit alpha

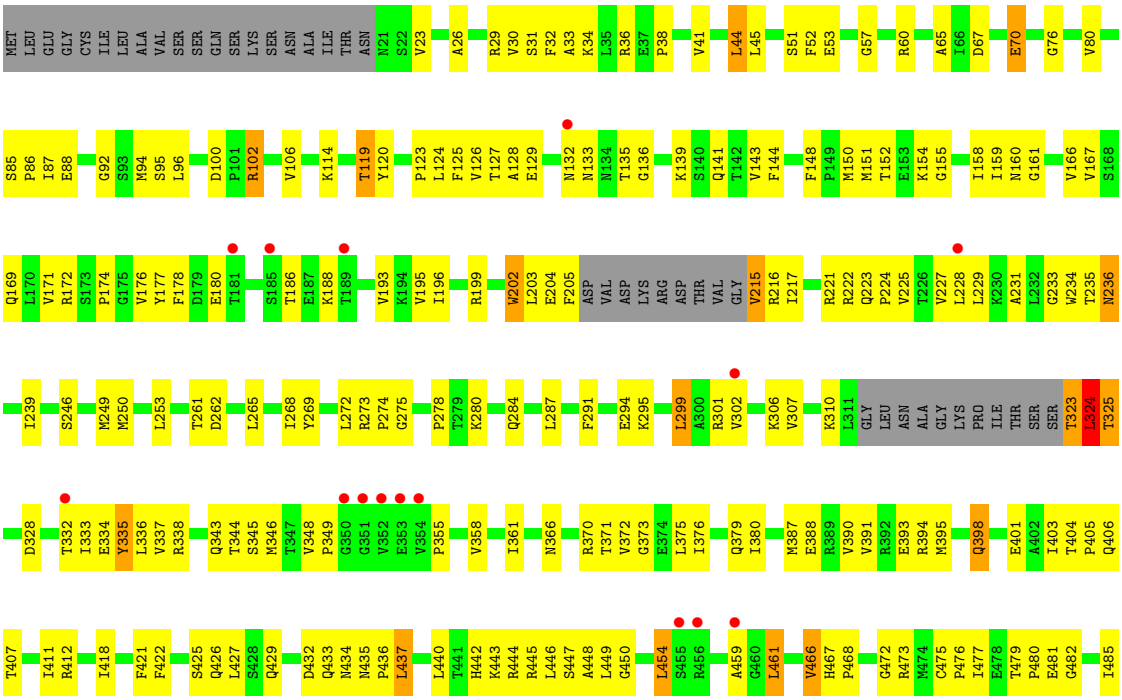
Chain B:

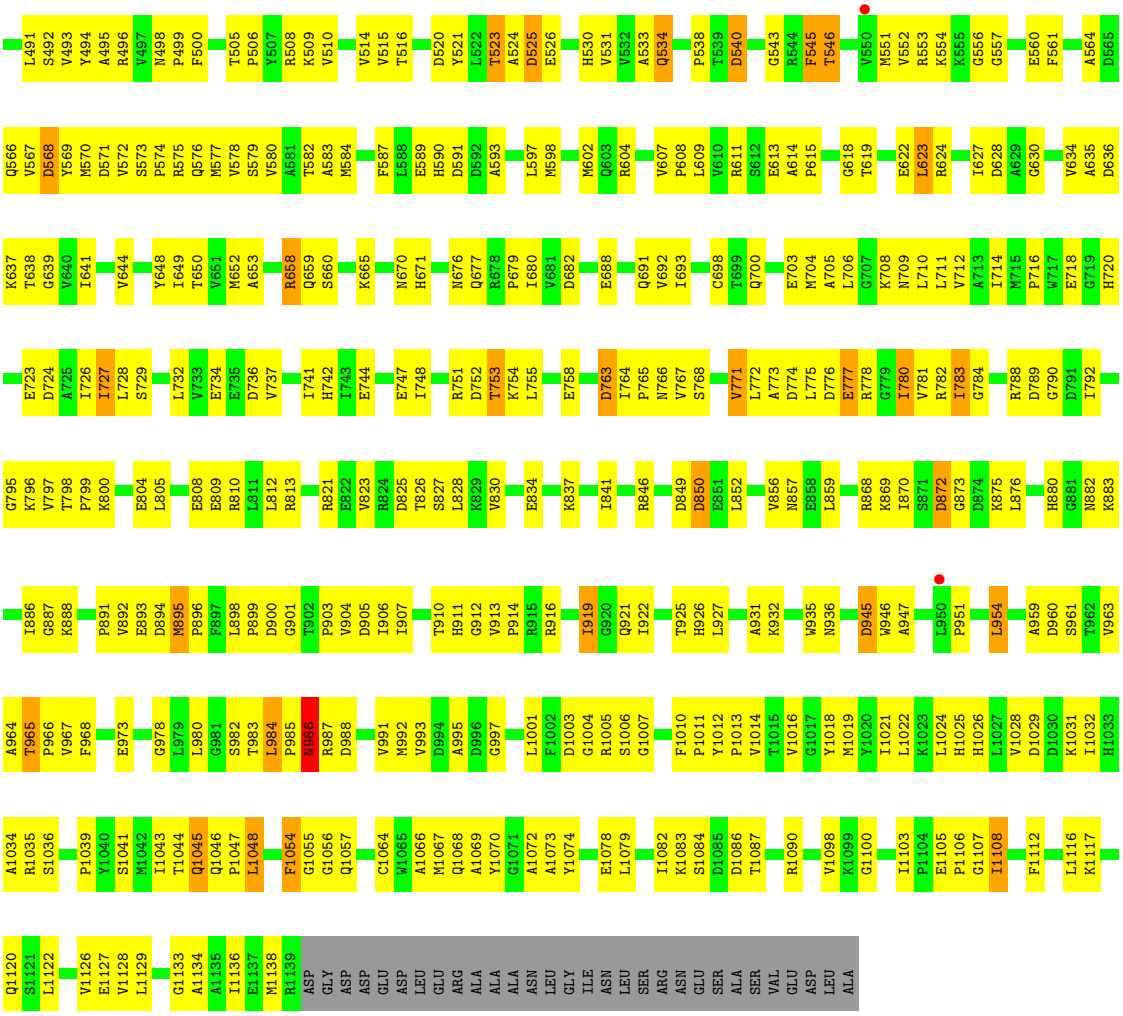


• Molecule 1: DNA-directed RNA polymerase subunit alpha



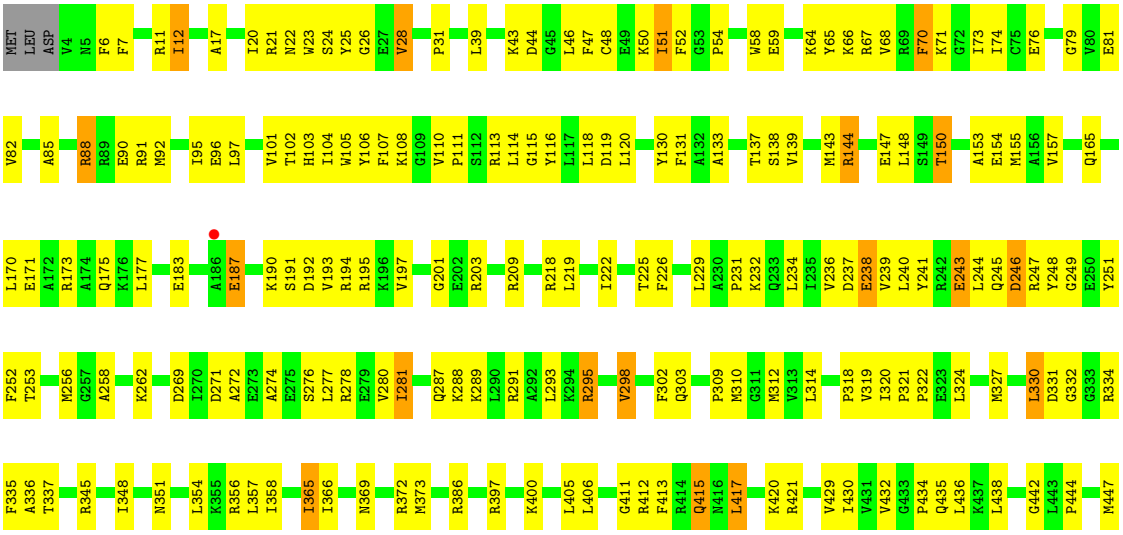
• Molecule 2: DNA-directed RNA polymerase subunit beta

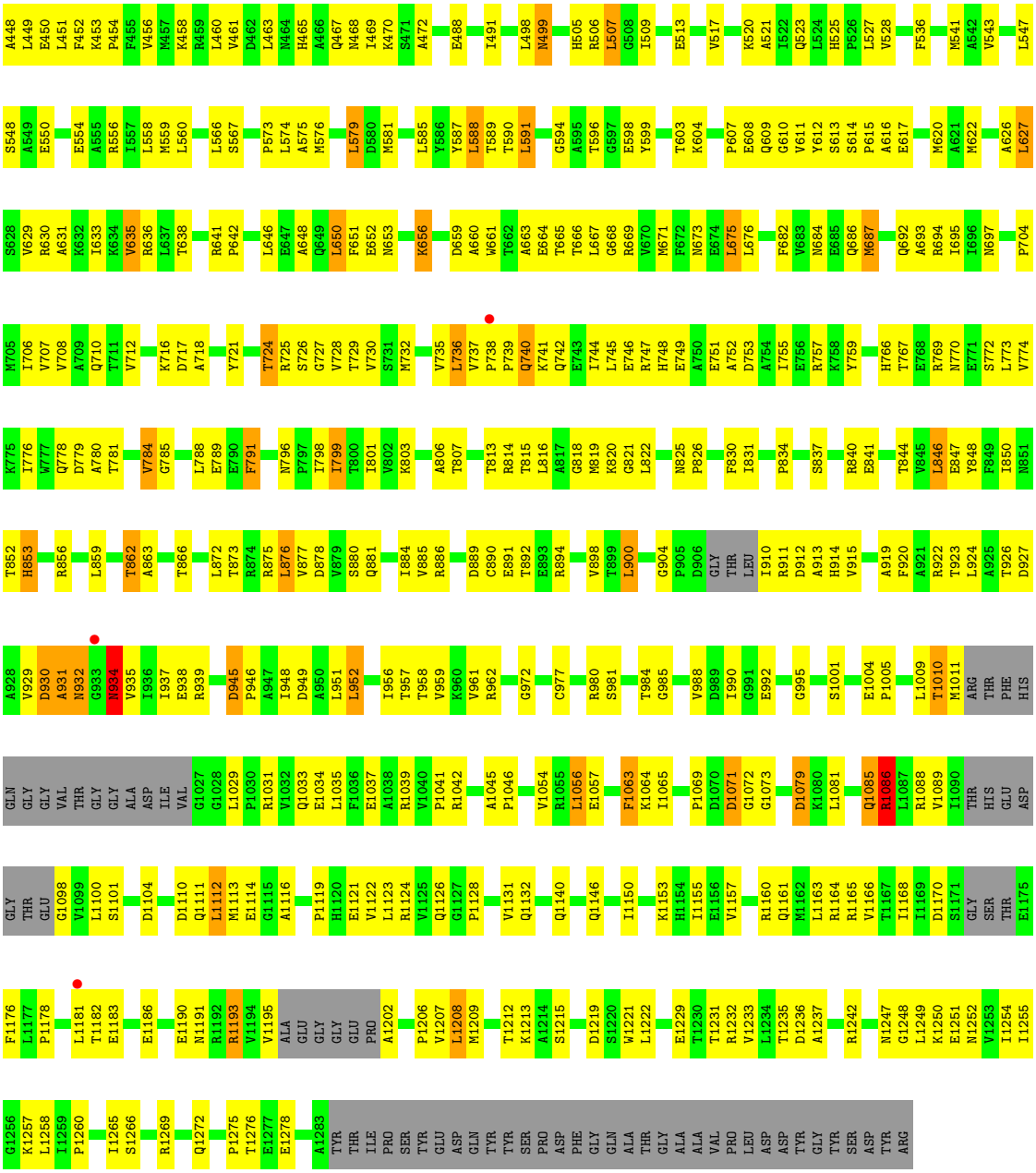




• Molecule 3: DNA-directed RNA polymerase subunit beta'

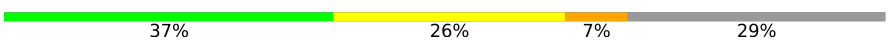
Chain D:

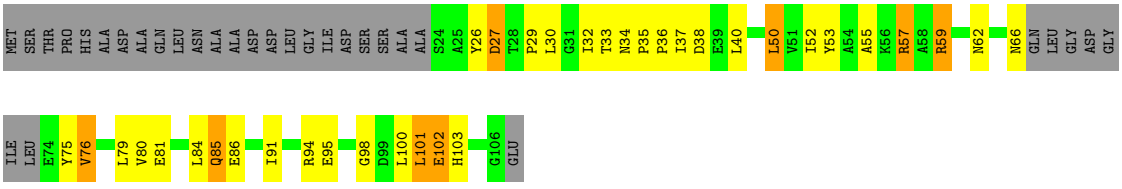




• Molecule 4: DNA-directed RNA polymerase subunit omega

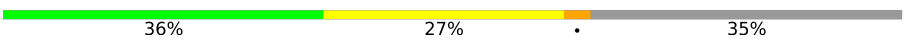
Chain E:

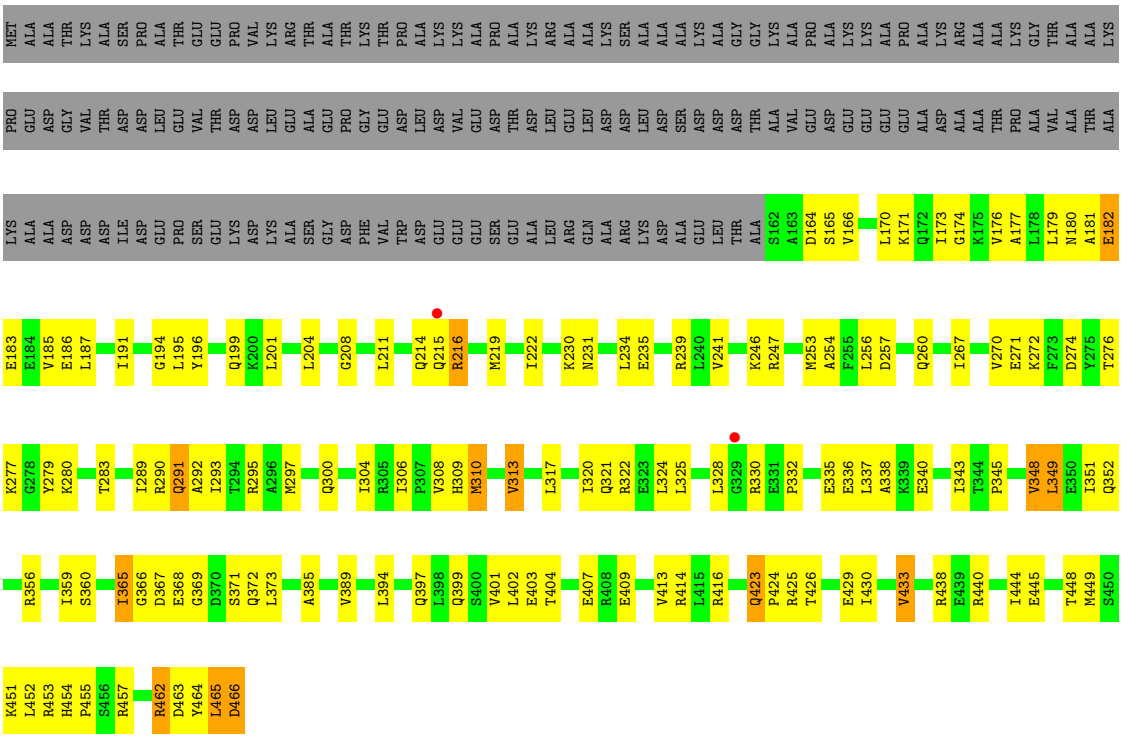




• Molecule 5: RNA polymerase sigma factor SigA

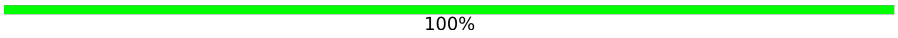
Chain F:





• Molecule 6: Unknown Peptide

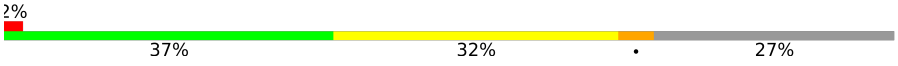
Chain G:

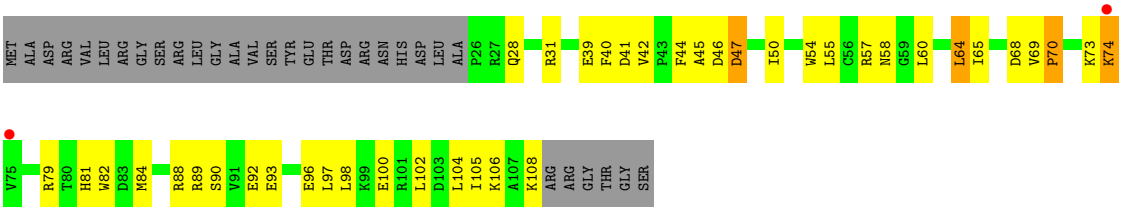


There are no outlier residues recorded for this chain.

• Molecule 7: RNA polymerase-binding protein RbpA

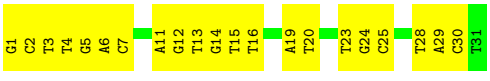
Chain J:





• Molecule 8: DNA (31-MER)

Chain O:



• Molecule 9: DNA (26-MER)

Chain P:


A1
G2
C3
A4
C5
A11
A12
T17
G20
T21
C22
A23
A24
G25
C26

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	132.35Å 162.32Å 139.40Å 90.00° 107.37° 90.00°	Depositor
Resolution (Å)	57.12 – 3.05 57.12 – 2.88	Depositor EDS
% Data completeness (in resolution range)	98.0 (57.12-3.05) 88.8 (57.12-2.88)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.22 (at 2.86Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.223 , 0.269 0.223 , 0.268	Depositor DCC
R_{free} test set	1877 reflections (1.67%)	wwPDB-VP
Wilson B-factor (Å^2)	78.8	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 61.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	26582	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% 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: RFP, EDO, SO4, ZN

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	0.23	0/1629	0.45	0/2220
1	B	0.23	0/1693	0.43	0/2316
1	T	0.22	0/343	0.37	0/468
2	C	0.24	0/8408	0.43	0/11428
3	D	0.24	0/9706	0.43	1/13140 (0.0%)
4	E	0.24	0/604	0.43	0/822
5	F	0.22	0/2445	0.39	0/3300
7	J	0.23	0/685	0.42	0/927
8	O	0.51	0/710	0.94	0/1095
9	P	0.57	0/589	0.93	0/906
All	All	0.26	0/26812	0.47	1/36622 (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
2	C	0	5
3	D	0	3
All	All	0	8

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	579	LEU	CA-CB-CG	5.98	129.06	115.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	1133	GLY	Peptide
2	C	324	LEU	Peptide
2	C	540	ASP	Peptide
2	C	985	PRO	Peptide
2	C	986	ASN	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	1605	0	1623	141	0
1	B	1667	0	1636	166	0
1	T	342	0	275	15	0
2	C	8262	0	8009	627	1
3	D	9555	0	9509	600	1
4	E	592	0	583	40	0
5	F	2414	0	2434	151	0
6	G	85	0	19	0	0
7	J	671	0	660	48	0
8	O	634	0	350	40	0
9	P	526	0	296	15	0
10	C	20	0	0	3	0
10	D	25	0	0	3	0
10	F	20	0	0	1	0
11	C	8	12	12	0	0
11	D	16	24	24	3	0
11	F	8	12	12	3	0
12	C	59	0	58	9	0
13	D	2	0	0	0	0
14	D	9	0	5	10	0
15	C	5	0	0	0	0
15	D	8	0	0	1	0
15	F	1	0	0	0	0
All	All	26534	48	25505	1669	1

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 32.

The worst 5 of 1669 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:603:THR:HG22	3:D:604:LYS:HG3	1.19	1.15
1:A:197:GLU:OE1	2:C:987:ARG:NH1	1.87	1.07
2:C:203:LEU:HG	2:C:217:ILE:HG22	1.35	1.07
2:C:771:VAL:HG22	2:C:772:LEU:HD12	1.35	1.07
2:C:53:GLU:OE2	2:C:60:ARG:NH1	1.89	1.06

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:658:ARG:NH1	3:D:147:GLU:OE1[2_356]	1.99	0.21

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	214/350 (61%)	204 (95%)	9 (4%)	1 (0%)	29	60
1	B	231/350 (66%)	211 (91%)	20 (9%)	0	100	100
1	T	51/350 (15%)	51 (100%)	0	0	100	100
2	C	1093/1169 (94%)	1040 (95%)	51 (5%)	2 (0%)	47	77
3	D	1234/1317 (94%)	1185 (96%)	45 (4%)	4 (0%)	41	70
4	E	72/107 (67%)	66 (92%)	5 (7%)	1 (1%)	11	36
5	F	303/466 (65%)	299 (99%)	4 (1%)	0	100	100
7	J	81/114 (71%)	77 (95%)	3 (4%)	1 (1%)	13	40
All	All	3279/4223 (78%)	3133 (96%)	137 (4%)	9 (0%)	41	70

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	931	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	1010	THR
2	C	325	THR
3	D	1086	ARG
2	C	850	ASP

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	174/297 (59%)	161 (92%)	13 (8%)	13 39
1	B	171/297 (58%)	158 (92%)	13 (8%)	13 38
1	T	26/297 (9%)	25 (96%)	1 (4%)	33 63
2	C	860/984 (87%)	798 (93%)	62 (7%)	14 40
3	D	989/1095 (90%)	913 (92%)	76 (8%)	13 38
4	E	62/86 (72%)	51 (82%)	11 (18%)	2 7
5	F	253/379 (67%)	236 (93%)	17 (7%)	16 43
7	J	72/98 (74%)	65 (90%)	7 (10%)	8 27
All	All	2607/3533 (74%)	2407 (92%)	200 (8%)	13 38

5 of 200 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	456	VAL
3	D	876	LEU
1	T	284	LEU
3	D	558	LEU
3	D	736	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	684	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	888	HIS
7	J	58	ASN
3	D	692	GLN
3	D	778	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 25 ligands modelled in this entry, 2 are monoatomic - leaving 23 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	$\#\|Z\| > 2$
10	SO4	D	2003	-	4,4,4	0.14	0	6,6,6	0.07	0
11	EDO	F	506	-	3,3,3	0.46	0	2,2,2	0.27	0
10	SO4	D	2006	-	4,4,4	0.14	0	6,6,6	0.05	0
10	SO4	F	504	-	4,4,4	0.14	0	6,6,6	0.05	0
11	EDO	F	505	-	3,3,3	0.46	0	2,2,2	0.26	0
10	SO4	F	501	5	4,4,4	0.13	0	6,6,6	0.04	0
10	SO4	D	2010	-	4,4,4	0.14	0	6,6,6	0.04	0
11	EDO	C	1207	-	3,3,3	0.44	0	2,2,2	0.32	0
10	SO4	D	2005	-	4,4,4	0.14	0	6,6,6	0.05	0
10	SO4	C	1202	-	4,4,4	0.14	0	6,6,6	0.05	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
11	EDO	D	2009	-	3,3,3	0.47	0	2,2,2	0.26	0
12	RFP	C	1205	-	63,63,63	2.64	17 (26%)	94,94,94	1.93	19 (20%)
10	SO4	C	1203	-	4,4,4	0.14	0	6,6,6	0.05	0
10	SO4	C	1206	-	4,4,4	0.15	0	6,6,6	0.06	0
10	SO4	D	2004	-	4,4,4	0.13	0	6,6,6	0.06	0
10	SO4	F	502	-	4,4,4	0.14	0	6,6,6	0.05	0
11	EDO	D	2008	-	3,3,3	0.45	0	2,2,2	0.31	0
11	EDO	D	2011	-	3,3,3	0.45	0	2,2,2	0.26	0
11	EDO	C	1204	-	3,3,3	0.43	0	2,2,2	0.42	0
14	GLU	D	2012	-	7,8,9	0.90	0	4,9,11	1.08	0
10	SO4	F	503	-	4,4,4	0.14	0	6,6,6	0.06	0
11	EDO	D	2007	-	3,3,3	0.45	0	2,2,2	0.30	0
10	SO4	C	1201	-	4,4,4	0.14	0	6,6,6	0.05	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
11	EDO	C	1207	-	-	0/1/1/1	-
11	EDO	F	506	-	-	1/1/1/1	-
11	EDO	D	2008	-	-	0/1/1/1	-
12	RFP	C	1205	-	-	13/60/85/85	0/5/5/5
11	EDO	D	2011	-	-	1/1/1/1	-
11	EDO	F	505	-	-	1/1/1/1	-
11	EDO	C	1204	-	-	1/1/1/1	-
14	GLU	D	2012	-	-	3/6/7/9	-
11	EDO	D	2007	-	-	0/1/1/1	-
11	EDO	D	2009	-	-	1/1/1/1	-

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	C	1205	RFP	C17-C16	8.22	1.58	1.34
12	C	1205	RFP	C15-N1	7.10	1.50	1.35
12	C	1205	RFP	C1-C9	6.55	1.63	1.43
12	C	1205	RFP	C18-C19	6.42	1.59	1.33
12	C	1205	RFP	C29-C28	5.85	1.60	1.30

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	C	1205	RFP	C2-C3-C43	-6.94	117.11	124.17
12	C	1205	RFP	C2-C3-C4	6.63	123.46	119.20
12	C	1205	RFP	O3-C6-C7	6.10	131.63	121.14
12	C	1205	RFP	O7-C35-C36	5.18	120.61	111.09
12	C	1205	RFP	C20-C21-C22	-3.52	107.80	114.96

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	C	1205	RFP	C4-C3-C43-N2
12	C	1205	RFP	C13-C12-O5-C29
12	C	1205	RFP	C26-C27-C28-C29
12	C	1205	RFP	O6-C27-C28-C29
12	C	1205	RFP	C43-N2-N3-C40

There are no ring outliers.

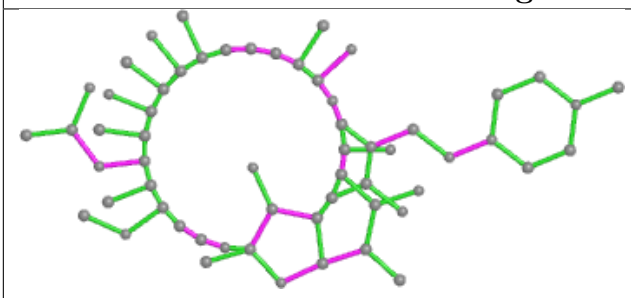
11 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	F	506	EDO	3	0
10	D	2010	SO4	1	0
12	C	1205	RFP	9	0
10	C	1203	SO4	1	0
10	C	1206	SO4	1	0
10	D	2004	SO4	2	0
10	F	502	SO4	1	0
11	D	2011	EDO	2	0
14	D	2012	GLU	10	0
11	D	2007	EDO	1	0
10	C	1201	SO4	1	0

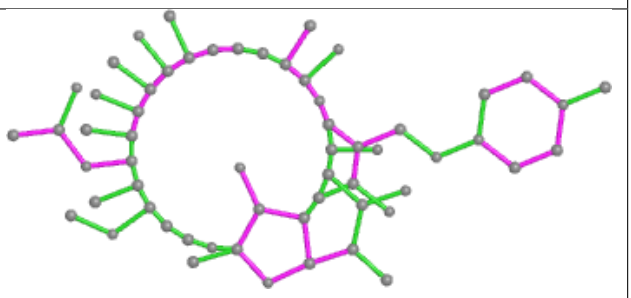
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

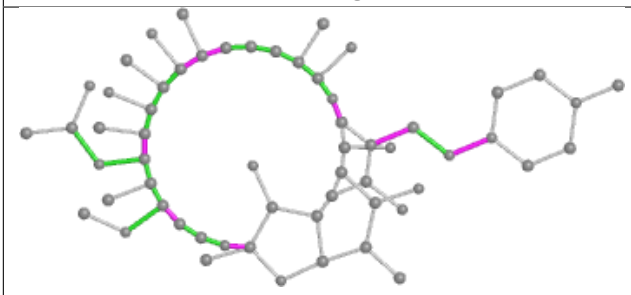
Ligand RFP C 1205



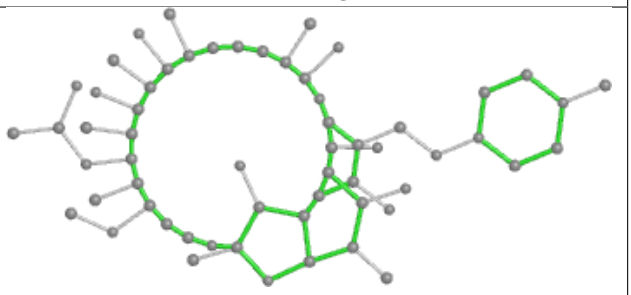
Bond lengths



Bond angles



Torsions



Rings

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	218/350 (62%)	-0.27	0 100 100	68, 97, 128, 146	0
1	B	233/350 (66%)	-0.00	5 (2%) 63 39	85, 121, 145, 161	0
1	T	53/350 (15%)	1.19	9 (16%) 1 0	123, 159, 184, 195	0
2	C	1099/1169 (94%)	-0.07	17 (1%) 73 51	47, 93, 154, 176	0
3	D	1246/1317 (94%)	-0.21	4 (0%) 94 85	41, 84, 140, 168	0
4	E	76/107 (71%)	-0.13	0 100 100	61, 89, 131, 144	0
5	F	305/466 (65%)	-0.37	2 (0%) 87 72	46, 86, 135, 170	0
6	G	0/17	-	-	-	-
7	J	83/114 (72%)	-0.11	2 (2%) 59 34	70, 111, 158, 173	0
8	O	31/31 (100%)	-0.89	0 100 100	59, 73, 96, 100	0
9	P	26/26 (100%)	-0.94	0 100 100	66, 80, 94, 104	0
All	All	3370/4297 (78%)	-0.16	39 (1%) 79 58	41, 92, 150, 195	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	455	SER	7.0
1	B	1	MET	3.8
2	C	354	VAL	3.5
2	C	352	VAL	3.4
2	C	185	SER	3.1

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)

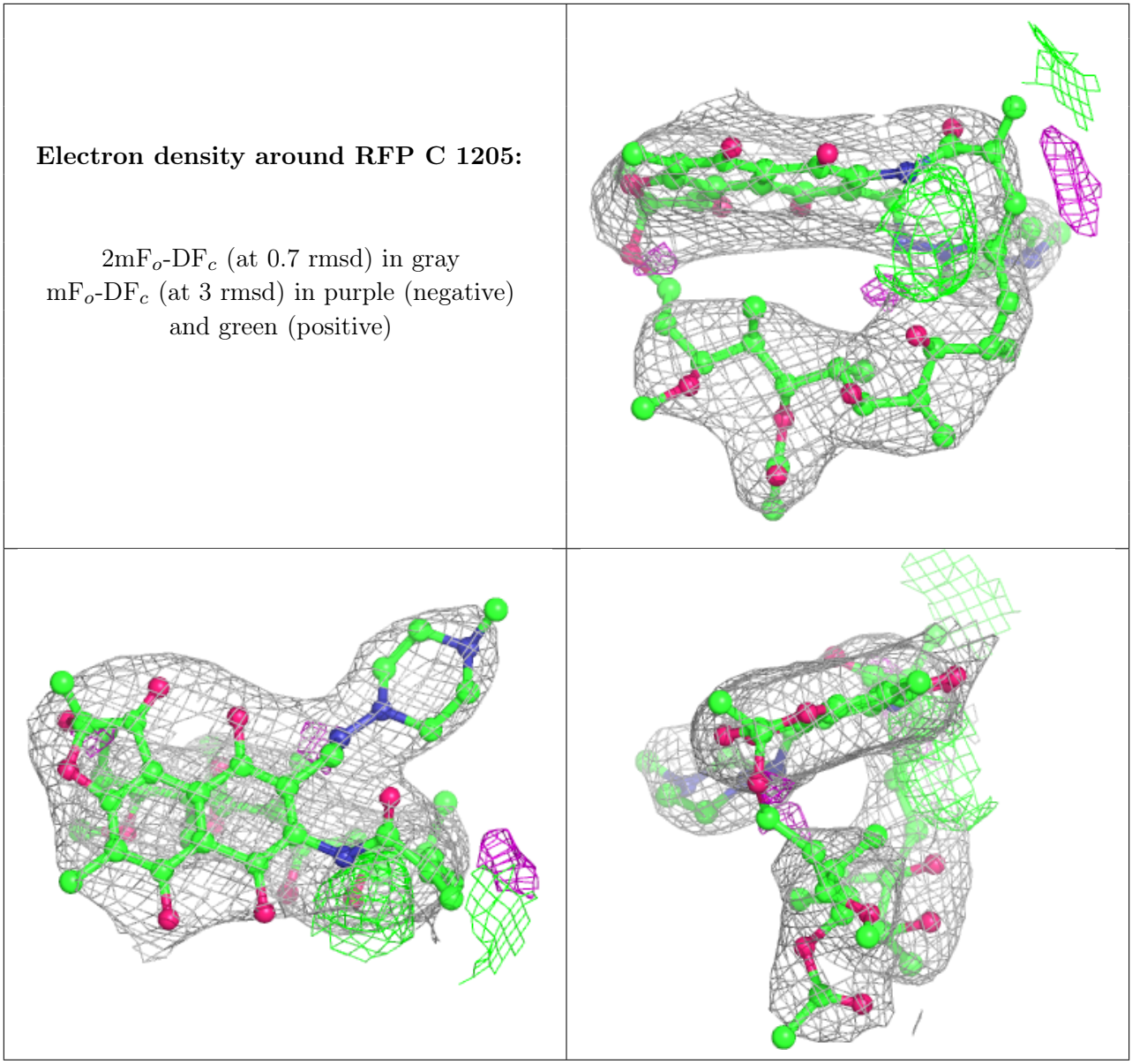
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
14	GLU	D	2012	9/10	0.71	0.30	90,97,100,110	0
10	SO4	C	1206	5/5	0.76	0.35	105,119,155,236	0
11	EDO	F	506	4/4	0.77	0.69	110,132,156,188	0
11	EDO	D	2007	4/4	0.82	0.28	75,94,113,123	0
10	SO4	D	2010	5/5	0.83	0.35	116,117,152,170	0
10	SO4	C	1203	5/5	0.86	0.29	114,127,160,168	0
11	EDO	D	2011	4/4	0.86	0.24	73,87,105,105	0
10	SO4	D	2006	5/5	0.87	0.25	116,124,133,140	0
11	EDO	F	505	4/4	0.87	0.32	81,97,121,121	0
10	SO4	D	2005	5/5	0.89	0.36	97,122,136,138	0
10	SO4	C	1202	5/5	0.89	0.16	119,126,141,149	0
10	SO4	F	504	5/5	0.90	0.11	121,127,136,142	0
11	EDO	D	2008	4/4	0.90	0.20	89,107,111,117	0
11	EDO	C	1207	4/4	0.90	0.28	94,113,134,134	0
10	SO4	F	503	5/5	0.91	0.12	104,110,129,130	0
10	SO4	D	2004	5/5	0.93	0.16	87,91,121,132	0
12	RFP	C	1205	59/59	0.94	0.24	47,69,104,119	0
11	EDO	D	2009	4/4	0.95	0.24	65,87,104,104	0
11	EDO	C	1204	4/4	0.95	0.26	61,74,92,92	0
10	SO4	C	1201	5/5	0.95	0.19	101,114,125,135	0
10	SO4	D	2003	5/5	0.96	0.22	64,76,98,102	0
13	ZN	D	2002	1/1	0.97	0.24	120,120,120,120	0
13	ZN	D	2001	1/1	0.98	0.33	129,129,129,129	0
10	SO4	F	502	5/5	0.98	0.18	79,101,117,120	0
10	SO4	F	501	5/5	0.98	0.04	104,105,128,134	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 RFP C 1205:

$2mF_o$-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.