wwPDB X-ray Structure Validation Summary Report (i)

Jun 27, 2024 – 02:10 PM EDT

PDB ID	:	5D4C
Title	:	Crystal structure of Thermus thermophilus product complex for transcription initiation with ATP and CTP
Authors	:	Zhang, Y.; Ebright, R.H.
Deposited on	:	2015-08-07
Resolution	:	3.28 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:

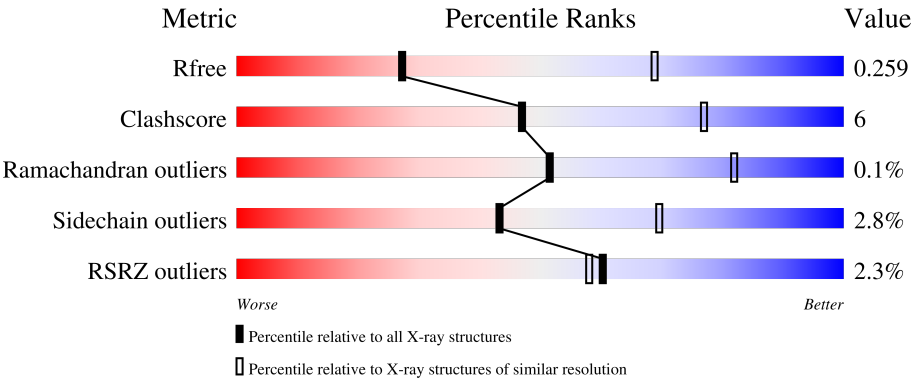
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.28 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	1177 (3.32-3.24)
Clashscore	141614	1044 (3.30-3.26)
Ramachandran outliers	138981	1026 (3.30-3.26)
Sidechain outliers	138945	1025 (3.30-3.26)
RSRZ outliers	127900	1141 (3.32-3.24)

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	315	60% 13% 27%
1	B	315	59% 11% 30%
1	K	315	61% 12% 27%
1	L	315	58% 12% 30%
2	C	1119	82% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	1119	5% 76% 19%
3	D	1524	% 81% 16%
3	N	1524	2% 80% 17%
4	E	99	83% 12% 5%
4	O	99	82% 12% 5%
5	F	443	2% 66% 11% 22%
5	P	443	6% 59% 18% 22%
6	G	19	47% 21% 11% 21%
6	R	19	58% 16% 5% 21%
7	H	27	30% 44% 26%
7	S	27	19% 44% 33%

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
8	MG	B	2001	-	-	-	X

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 56600 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	231	Total	C	N	O	S	0	0	0
			1809	1155	315	337	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1120	303	325	2			
1	K	231	Total	C	N	O	S	0	0	0
			1809	1155	315	337	2			
1	L	222	Total	C	N	O	S	0	0	0
			1750	1120	303	325	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1112	Total	C	N	O	S	0	0	0
			8774	5550	1565	1635	24			
2	M	1080	Total	C	N	O	S	0	0	0
			8508	5375	1522	1587	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			
3	N	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	94	Total	C	N	O	S	0	0	0
			758	483	132	139	4			
4	O	94	Total	C	N	O	S	0	0	0
			758	483	132	139	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			
5	P	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	initiating methionine	UNP Q5SKW1
F	-18	GLY	-	expression tag	UNP Q5SKW1
F	-17	SER	-	expression tag	UNP Q5SKW1
F	-16	SER	-	expression tag	UNP Q5SKW1
F	-15	HIS	-	expression tag	UNP Q5SKW1
F	-14	HIS	-	expression tag	UNP Q5SKW1
F	-13	HIS	-	expression tag	UNP Q5SKW1
F	-12	HIS	-	expression tag	UNP Q5SKW1
F	-11	HIS	-	expression tag	UNP Q5SKW1
F	-10	HIS	-	expression tag	UNP Q5SKW1
F	-9	SER	-	expression tag	UNP Q5SKW1
F	-8	SER	-	expression tag	UNP Q5SKW1
F	-7	GLY	-	expression tag	UNP Q5SKW1
F	-6	LEU	-	expression tag	UNP Q5SKW1
F	-5	VAL	-	expression tag	UNP Q5SKW1
F	-4	PRO	-	expression tag	UNP Q5SKW1
F	-3	ARG	-	expression tag	UNP Q5SKW1
F	-2	GLY	-	expression tag	UNP Q5SKW1
F	-1	SER	-	expression tag	UNP Q5SKW1
F	0	HIS	-	expression tag	UNP Q5SKW1
P	-19	MET	-	initiating methionine	UNP Q5SKW1
P	-18	GLY	-	expression tag	UNP Q5SKW1
P	-17	SER	-	expression tag	UNP Q5SKW1
P	-16	SER	-	expression tag	UNP Q5SKW1
P	-15	HIS	-	expression tag	UNP Q5SKW1
P	-14	HIS	-	expression tag	UNP Q5SKW1
P	-13	HIS	-	expression tag	UNP Q5SKW1
P	-12	HIS	-	expression tag	UNP Q5SKW1
P	-11	HIS	-	expression tag	UNP Q5SKW1
P	-10	HIS	-	expression tag	UNP Q5SKW1
P	-9	SER	-	expression tag	UNP Q5SKW1
P	-8	SER	-	expression tag	UNP Q5SKW1
P	-7	GLY	-	expression tag	UNP Q5SKW1
P	-6	LEU	-	expression tag	UNP Q5SKW1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-5	VAL	-	expression tag	UNP Q5SKW1
P	-4	PRO	-	expression tag	UNP Q5SKW1
P	-3	ARG	-	expression tag	UNP Q5SKW1
P	-2	GLY	-	expression tag	UNP Q5SKW1
P	-1	SER	-	expression tag	UNP Q5SKW1
P	0	HIS	-	expression tag	UNP Q5SKW1

- Molecule 6 is a DNA chain called DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*AP*GP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	15	Total	C	N	O	P	0	0	0
			308	147	60	87	14			
6	R	15	Total	C	N	O	P	0	0	0
			308	147	60	87	14			

- Molecule 7 is a DNA chain called DNA (27-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	20	Total	C	N	O	P	0	0	0
			414	197	82	116	19			
7	S	18	Total	C	N	O	P	0	0	0
			371	177	72	105	17			

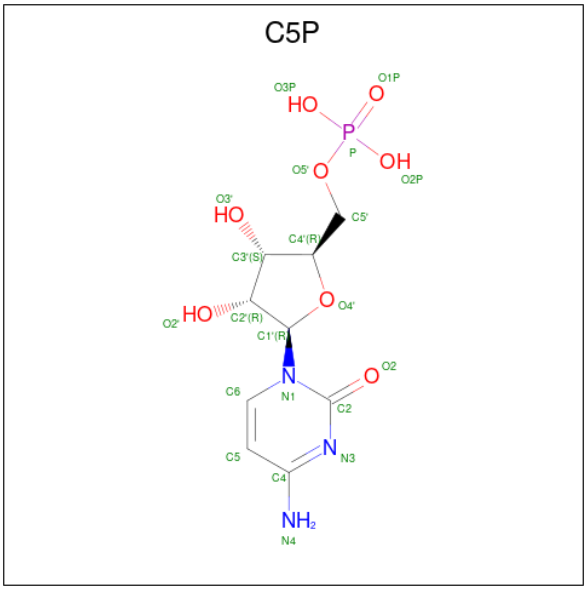
- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Mg	0	0
			1	1		
8	D	3	Total	Mg	0	0
			3	3		
8	F	1	Total	Mg	0	0
			1	1		
8	K	1	Total	Mg	0	0
			1	1		
8	L	2	Total	Mg	0	0
			2	2		
8	N	2	Total	Mg	0	0
			2	2		
8	P	1	Total	Mg	0	0
			1	1		

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

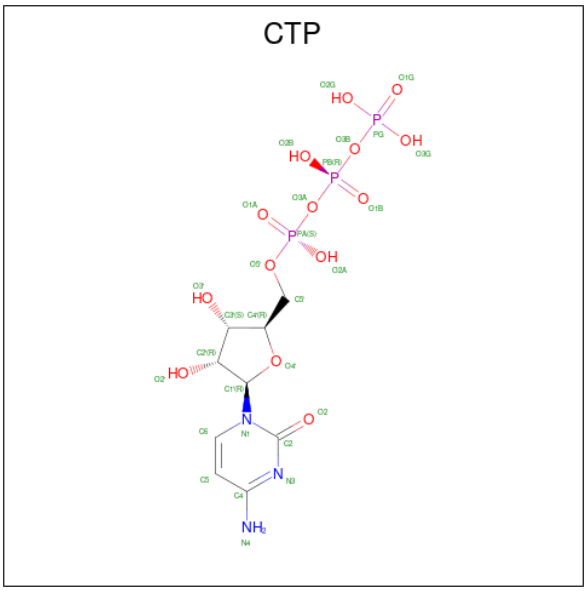
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	D	2	Total Zn 2 2	0	0
9	N	2	Total Zn 2 2	0	0

- Molecule 10 is CYTIDINE-5'-MONOPHOSPHATE (three-letter code: C5P) (formula: $C_9H_{14}N_3O_8P$).



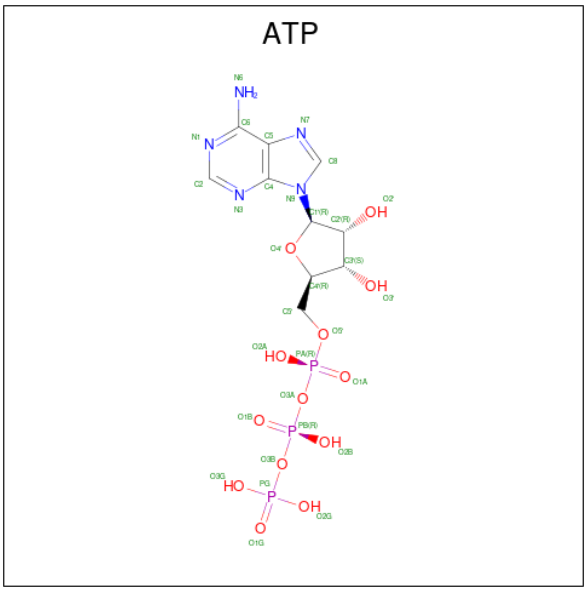
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	D	1	Total C N O P 20 9 3 7 1	0	0
10	N	1	Total C N O P 20 9 3 7 1	0	0

- Molecule 11 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	D	1	Total	O	P	0	0
			9	7	2		
11	N	1	Total	O	P	0	0
			9	7	2		

- Molecule 12 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 13 is water.

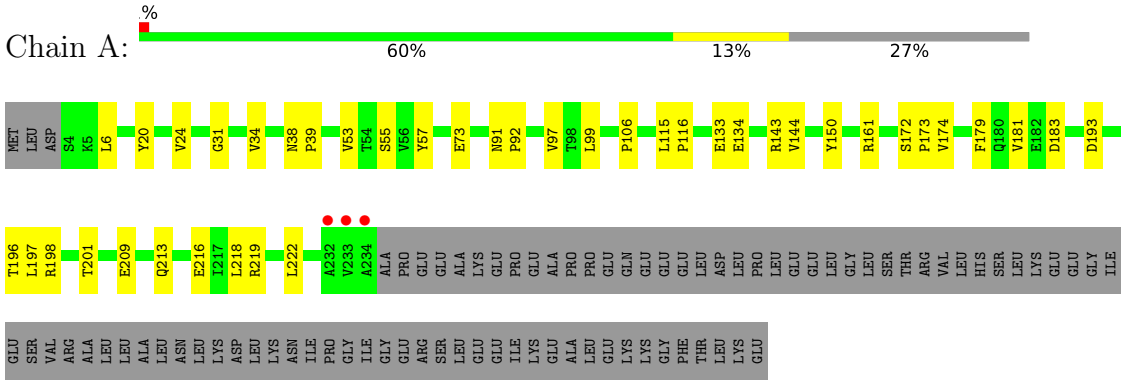
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	3	Total	O	0	0
			3	3		
13	B	2	Total	O	0	0
			2	2		
13	C	11	Total	O	0	0
			11	11		
13	D	18	Total	O	0	0
			18	18		
13	E	1	Total	O	0	0
			1	1		
13	G	2	Total	O	0	0
			2	2		
13	K	3	Total	O	0	0
			3	3		
13	L	1	Total	O	0	0
			1	1		
13	M	1	Total	O	0	0
			1	1		
13	N	14	Total	O	0	0
			14	14		
13	O	1	Total	O	0	0
			1	1		
13	P	1	Total	O	0	0
			1	1		

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.

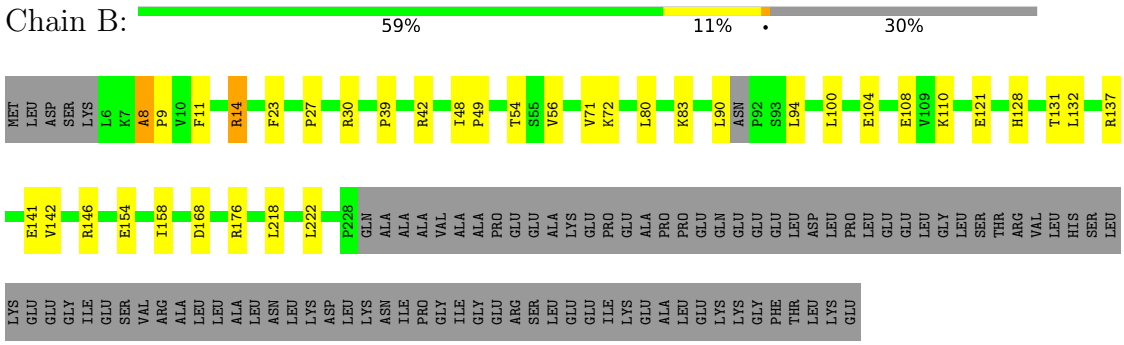
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



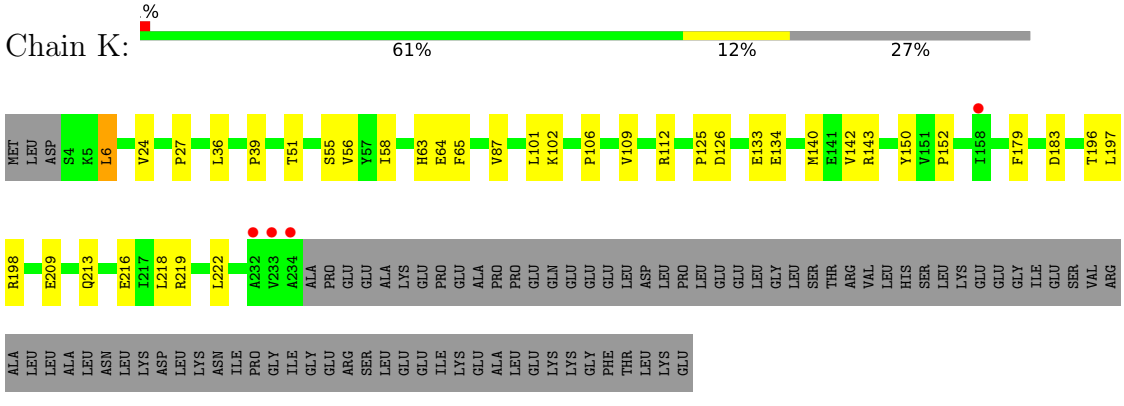
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:



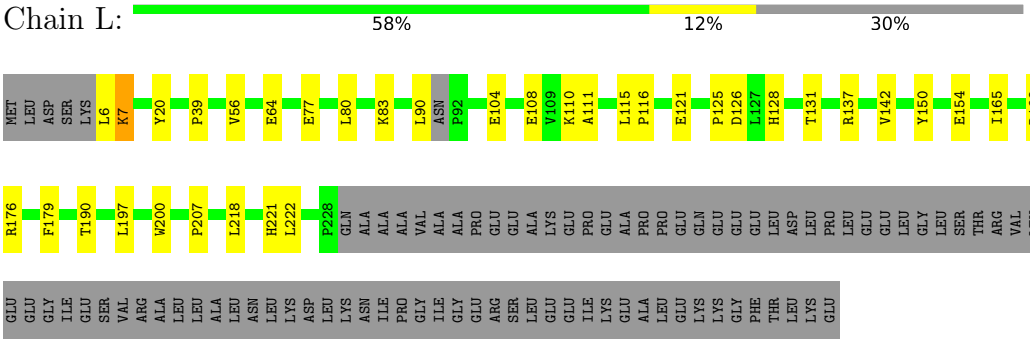
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain K:



• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain L:

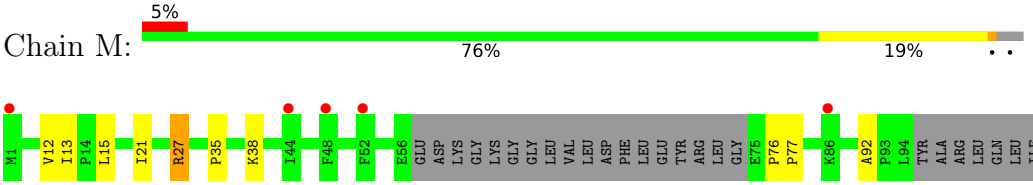


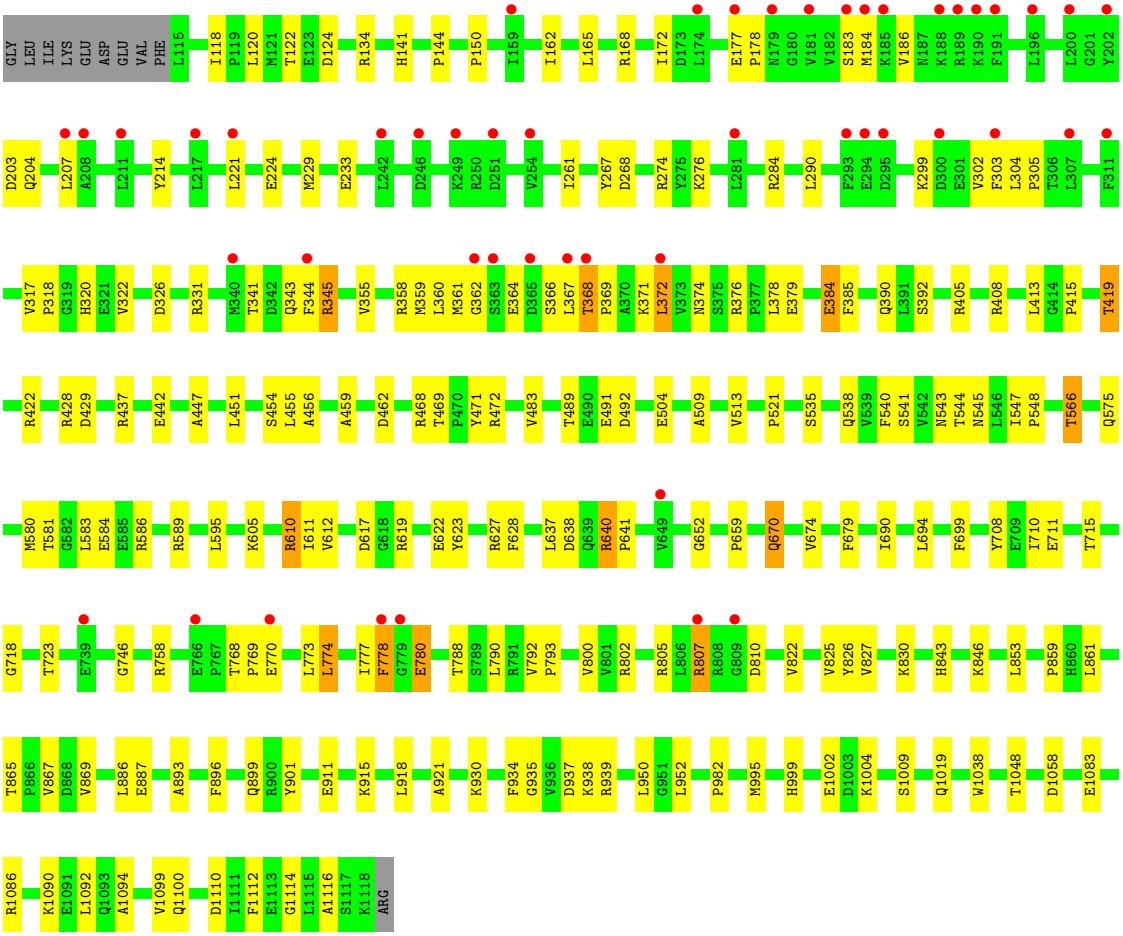
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:

• Molecule 2: DNA-directed RNA polymerase subunit beta

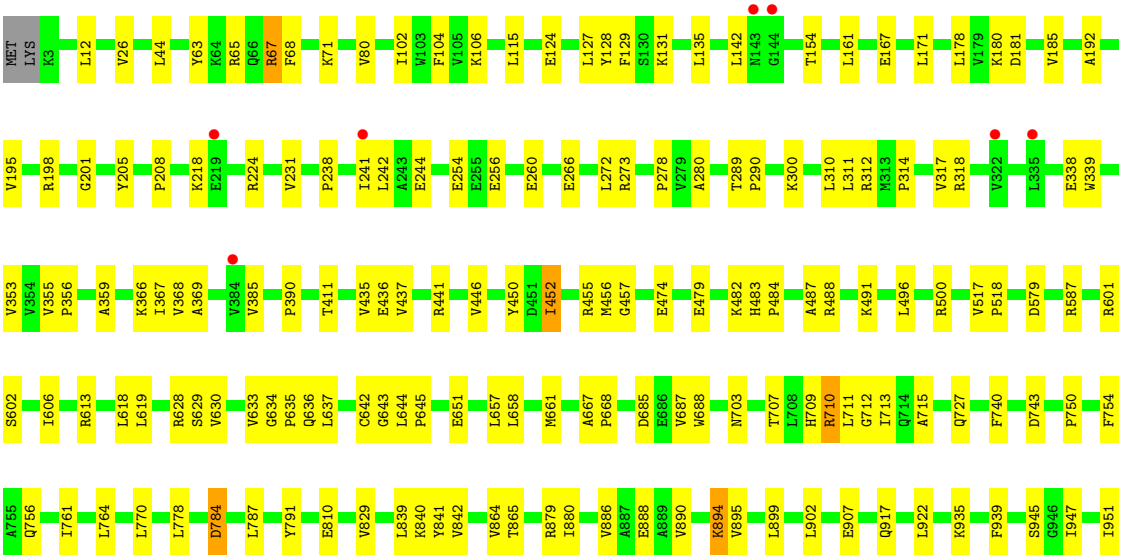
Chain M:

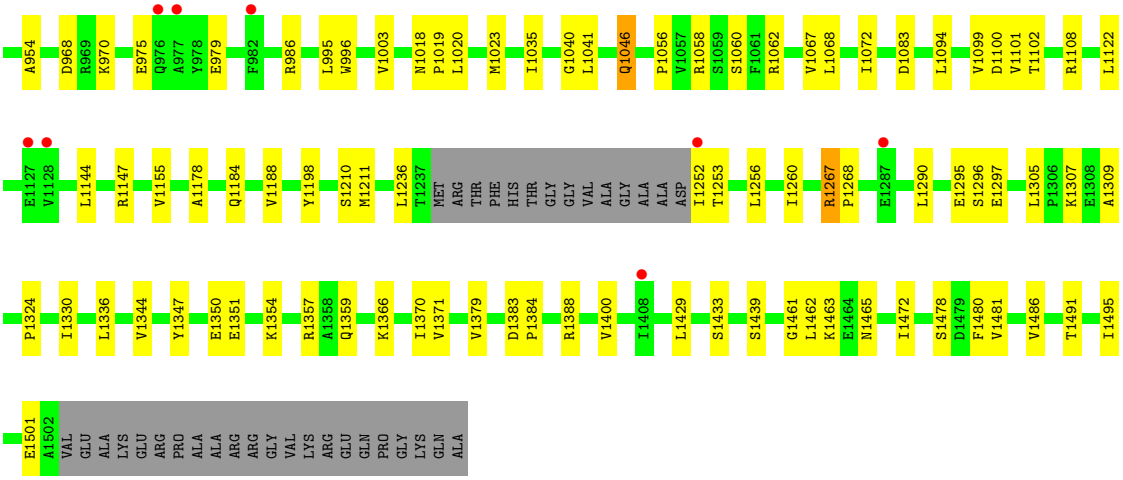




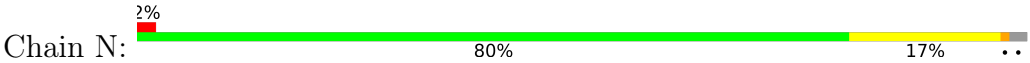
• Molecule 3: DNA-directed RNA polymerase subunit beta'

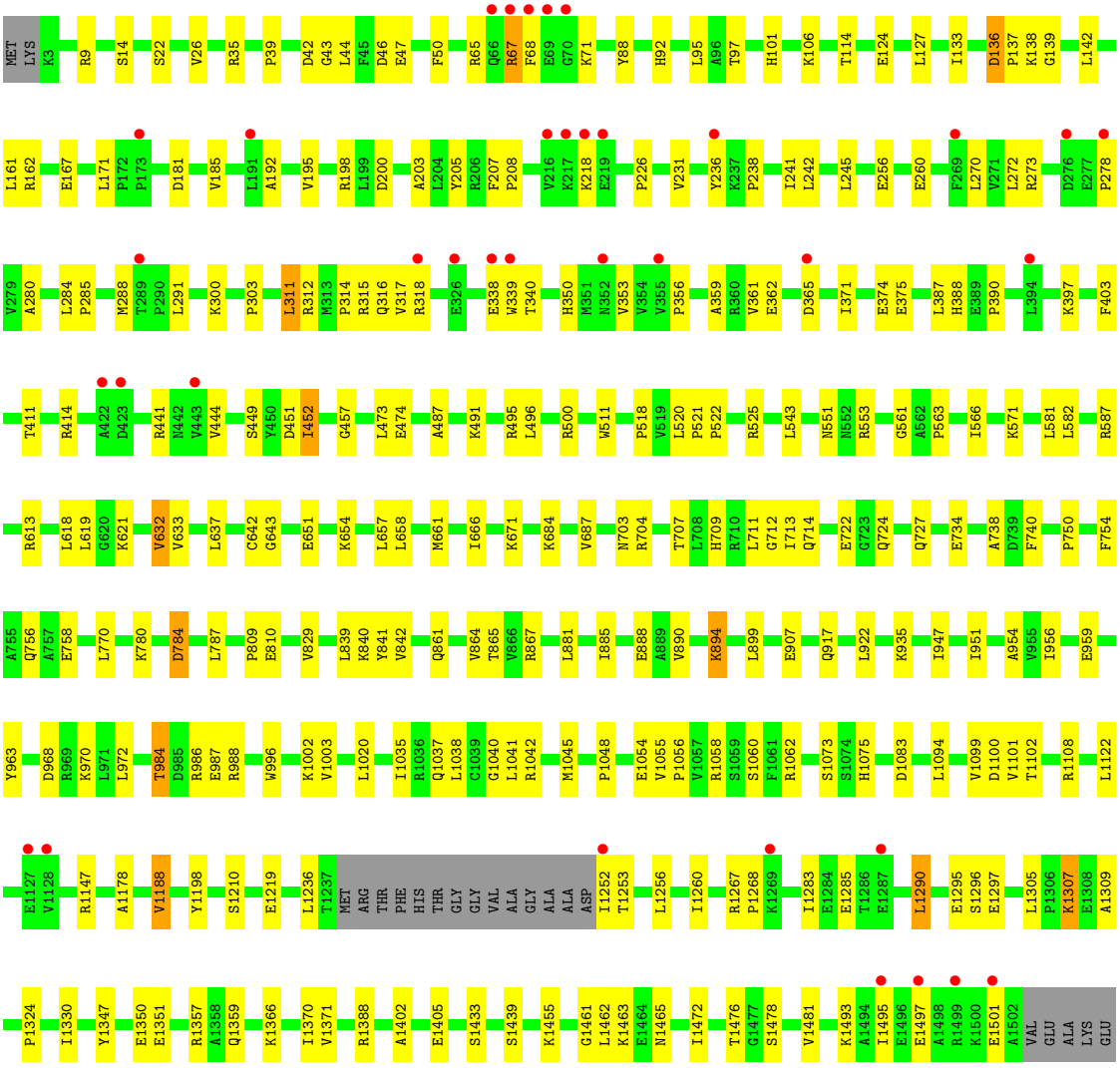
Chain D:





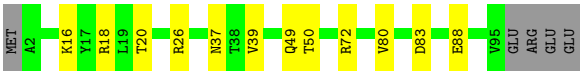
• Molecule 3: DNA-directed RNA polymerase subunit beta'





• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:

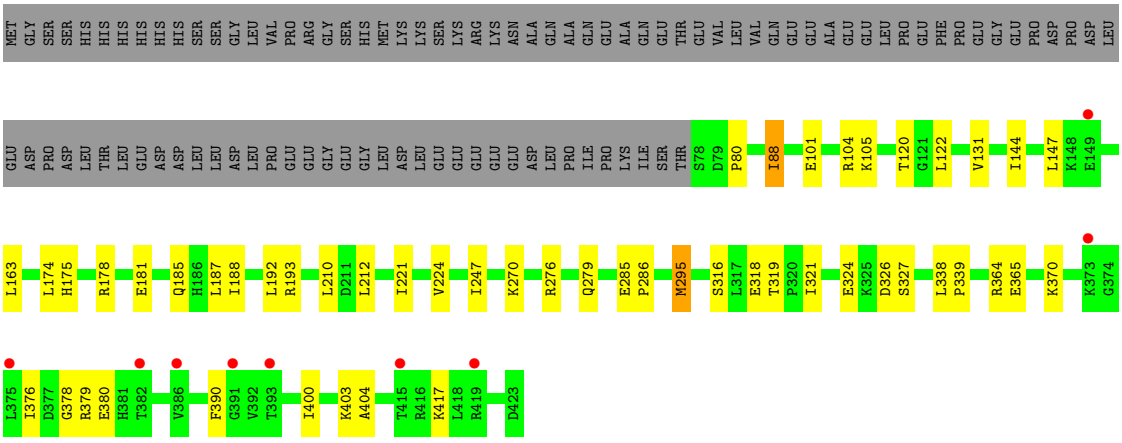


• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain O:

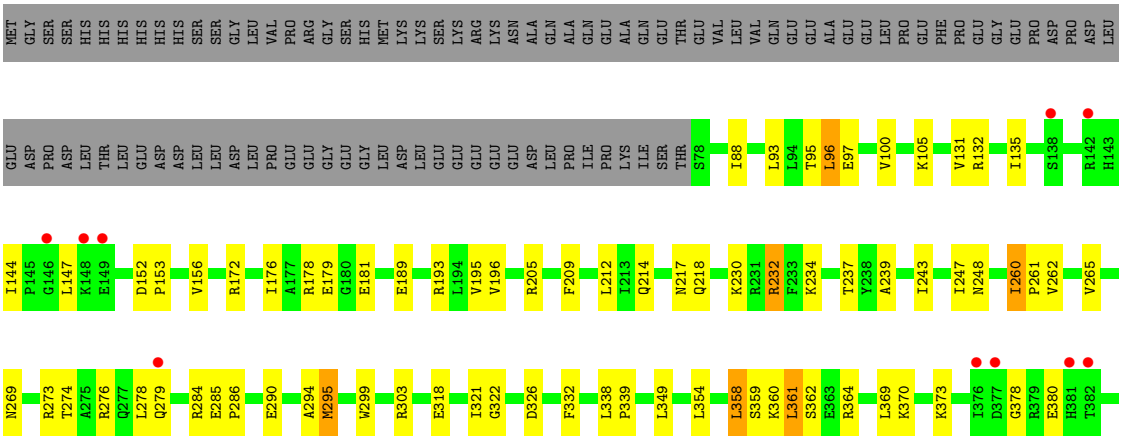
• Molecule 5: RNA polymerase sigma factor SigA

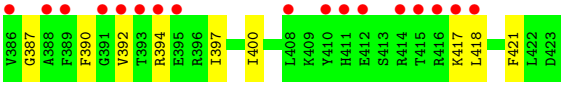
Chain F:



• Molecule 5: RNA polymerase sigma factor SigA

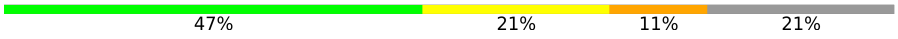
Chain P:

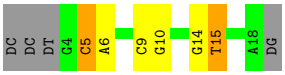




• Molecule 6: DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*AP*GP*AP*G)-3')

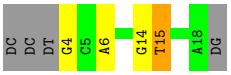
Chain G:





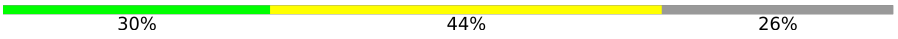
• Molecule 6: DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*AP*GP*AP*G)-3')

Chain R:



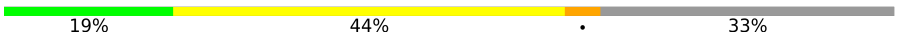
• Molecule 7: DNA (27-MER)

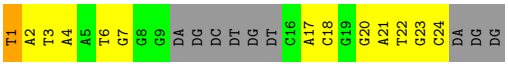
Chain H:



• Molecule 7: DNA (27-MER)

Chain S:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	185.96Å 103.64Å 297.42Å 90.00° 98.30° 90.00°	Depositor
Resolution (Å)	49.20 – 3.28 49.50 – 3.28	Depositor EDS
% Data completeness (in resolution range)	89.5 (49.20-3.28) 89.7 (49.50-3.28)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	0.18	Depositor
$< I/\sigma(I) >$ [1]	3.91 (at 3.25Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.211 , 0.257 0.213 , 0.259	Depositor DCC
R_{free} test set	7741 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	57.7	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 31.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	56600	wwPDB-VP
Average B, all atoms (Å^2)	48.0	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 52.73 % 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 4.6669e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: MG, ATP, ZN, CTP, C5P

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	0.24	0/1841	0.45	0/2504
1	B	0.23	0/1781	0.47	0/2420
1	K	0.24	0/1841	0.44	0/2504
1	L	0.24	0/1781	0.45	0/2420
2	C	0.24	0/8941	0.44	0/12092
2	M	0.24	0/8669	0.45	0/11724
3	D	0.24	0/11944	0.44	0/16149
3	N	0.24	0/11944	0.44	0/16149
4	E	0.23	0/772	0.42	0/1040
4	O	0.22	0/772	0.42	0/1040
5	F	0.24	0/2852	0.40	0/3837
5	P	0.24	0/2852	0.43	0/3837
6	G	0.51	0/346	1.14	2/533 (0.4%)
6	R	0.51	0/346	1.09	1/533 (0.2%)
7	H	0.59	0/465	1.06	0/715
7	S	0.51	0/416	1.06	1/639 (0.2%)
All	All	0.25	0/57563	0.47	4/78136 (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
3	D	0	1
3	N	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	S	1	DT	O4'-C1'-N1	6.12	112.28	108.00
6	G	15	DT	O4'-C4'-C3'	-5.81	102.17	104.50
6	G	5	DC	O4'-C1'-N1	5.27	111.69	108.00
6	R	15	DT	O4'-C4'-C3'	-5.10	102.46	104.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	633	VAL	Peptide
3	N	138	LYS	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	1809	0	1863	26	0
1	B	1750	0	1802	25	0
1	K	1809	0	1863	25	0
1	L	1750	0	1802	25	0
2	C	8774	0	8877	109	0
2	M	8508	0	8605	136	0
3	D	11738	0	11972	146	0
3	N	11738	0	11971	163	0
4	E	758	0	770	9	0
4	O	758	0	770	5	0
5	F	2807	0	2882	30	0
5	P	2807	0	2882	54	0
6	G	308	0	170	5	0
6	R	308	0	170	3	0
7	H	414	0	227	12	0
7	S	371	0	205	13	0
8	B	1	0	0	0	0
8	D	3	0	0	0	0
8	F	1	0	0	0	0
8	K	1	0	0	0	0
8	L	2	0	0	0	0
8	N	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	P	1	0	0	0	0
9	D	2	0	0	0	0
9	N	2	0	0	0	0
10	D	20	0	11	0	0
10	N	20	0	11	1	0
11	D	9	0	0	0	0
11	N	9	0	0	0	0
12	G	31	0	11	0	0
12	R	31	0	11	0	0
13	A	3	0	0	0	0
13	B	2	0	0	0	0
13	C	11	0	0	0	0
13	D	18	0	0	0	0
13	E	1	0	0	0	0
13	G	2	0	0	0	0
13	K	3	0	0	0	0
13	L	1	0	0	0	0
13	M	1	0	0	0	0
13	N	14	0	0	0	0
13	O	1	0	0	0	0
13	P	1	0	0	0	0
All	All	56600	0	56875	707	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 6.

The worst 5 of 707 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:C:243:ARG:NH1	7:H:9:DG:O6	1.99	0.93
2:M:165:LEU:HB2	2:M:168:ARG:HG3	1.59	0.84
2:M:758:ARG:HH21	2:M:788:THR:HB	1.45	0.80
3:D:1495:ILE:HG12	4:E:88:GLU:HG3	1.67	0.77
2:M:770:GLU:HB3	5:P:354:LEU:HG	1.65	0.77

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	229/315 (73%)	226 (99%)	3 (1%)	0	100	100
1	B	218/315 (69%)	212 (97%)	5 (2%)	1 (0%)	29	62
1	K	229/315 (73%)	226 (99%)	3 (1%)	0	100	100
1	L	218/315 (69%)	212 (97%)	6 (3%)	0	100	100
2	C	1108/1119 (99%)	1085 (98%)	22 (2%)	1 (0%)	51	82
2	M	1074/1119 (96%)	1049 (98%)	25 (2%)	0	100	100
3	D	1482/1524 (97%)	1451 (98%)	31 (2%)	0	100	100
3	N	1482/1524 (97%)	1446 (98%)	35 (2%)	1 (0%)	51	82
4	E	92/99 (93%)	89 (97%)	3 (3%)	0	100	100
4	O	92/99 (93%)	89 (97%)	3 (3%)	0	100	100
5	F	344/443 (78%)	340 (99%)	4 (1%)	0	100	100
5	P	344/443 (78%)	338 (98%)	5 (2%)	1 (0%)	41	72
All	All	6912/7630 (91%)	6763 (98%)	145 (2%)	4 (0%)	51	82

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	P	361	LEU
3	N	207	PHE
1	B	8	ALA
2	C	23	VAL

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	200/273 (73%)	199 (100%)	1 (0%)	88 93
1	B	195/273 (71%)	193 (99%)	2 (1%)	76 85
1	K	200/273 (73%)	197 (98%)	3 (2%)	65 81
1	L	195/273 (71%)	191 (98%)	4 (2%)	53 75
2	C	936/941 (100%)	910 (97%)	26 (3%)	43 69
2	M	908/941 (96%)	876 (96%)	32 (4%)	36 64
3	D	1253/1279 (98%)	1225 (98%)	28 (2%)	52 74
3	N	1253/1279 (98%)	1221 (97%)	32 (3%)	46 71
4	E	82/88 (93%)	80 (98%)	2 (2%)	49 73
4	O	82/88 (93%)	76 (93%)	6 (7%)	14 40
5	F	301/388 (78%)	294 (98%)	7 (2%)	50 73
5	P	301/388 (78%)	280 (93%)	21 (7%)	15 43
All	All	5906/6484 (91%)	5742 (97%)	164 (3%)	43 69

5 of 164 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	387	LEU
5	P	88	ILE
3	N	709	HIS
3	N	1252	ILE
5	P	230	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
3	N	1172	HIS
5	P	83	GLN
5	P	269	ASN
1	K	63	HIS
5	F	83	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 21 ligands modelled in this entry, 15 are monoatomic - leaving 6 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	C5P	N	1604	12,8	18,21,22	0.33	0	26,30,33	0.38	0
12	ATP	R	101	10	26,33,33	1.70	5 (19%)	31,52,52	1.55	5 (16%)
10	C5P	D	2005	12,8	18,21,22	0.32	0	26,30,33	0.49	0
11	CTP	N	1605	8	6,8,30	0.96	0	13,13,47	1.21	2 (15%)
12	ATP	G	101	10	26,33,33	1.68	5 (19%)	31,52,52	1.59	4 (12%)
11	CTP	D	2006	8	6,8,30	1.58	1 (16%)	13,13,47	1.33	1 (7%)

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
10	C5P	N	1604	12,8	-	1/7/25/26	0/2/2/2
12	ATP	R	101	10	-	3/18/38/38	0/3/3/3
10	C5P	D	2005	12,8	-	1/7/25/26	0/2/2/2
11	CTP	N	1605	8	-	3/6/6/38	-
12	ATP	G	101	10	-	4/18/38/38	0/3/3/3
11	CTP	D	2006	8	-	0/6/6/38	-

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	G	101	ATP	C2-N3	4.90	1.40	1.32
12	R	101	ATP	C2-N3	4.75	1.39	1.32
11	D	2006	CTP	PA-O1A	3.47	1.61	1.50
12	R	101	ATP	C2'-C1'	-3.44	1.48	1.53
12	G	101	ATP	C2'-C1'	-3.41	1.48	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	R	101	ATP	N3-C2-N1	-5.00	120.86	128.68
12	G	101	ATP	N3-C2-N1	-4.83	121.12	128.68
12	G	101	ATP	PA-O3A-PB	-3.68	120.19	132.83
12	R	101	ATP	PA-O3A-PB	-3.53	120.70	132.83
12	R	101	ATP	PB-O3B-PG	-3.49	120.85	132.83

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	N	1605	CTP	PA-O3A-PB-O3B
12	G	101	ATP	C5'-O5'-PA-O3A
12	R	101	ATP	PG-O3B-PB-O2B
12	G	101	ATP	C5'-O5'-PA-O1A
12	G	101	ATP	C5'-O5'-PA-O2A

There are no ring outliers.

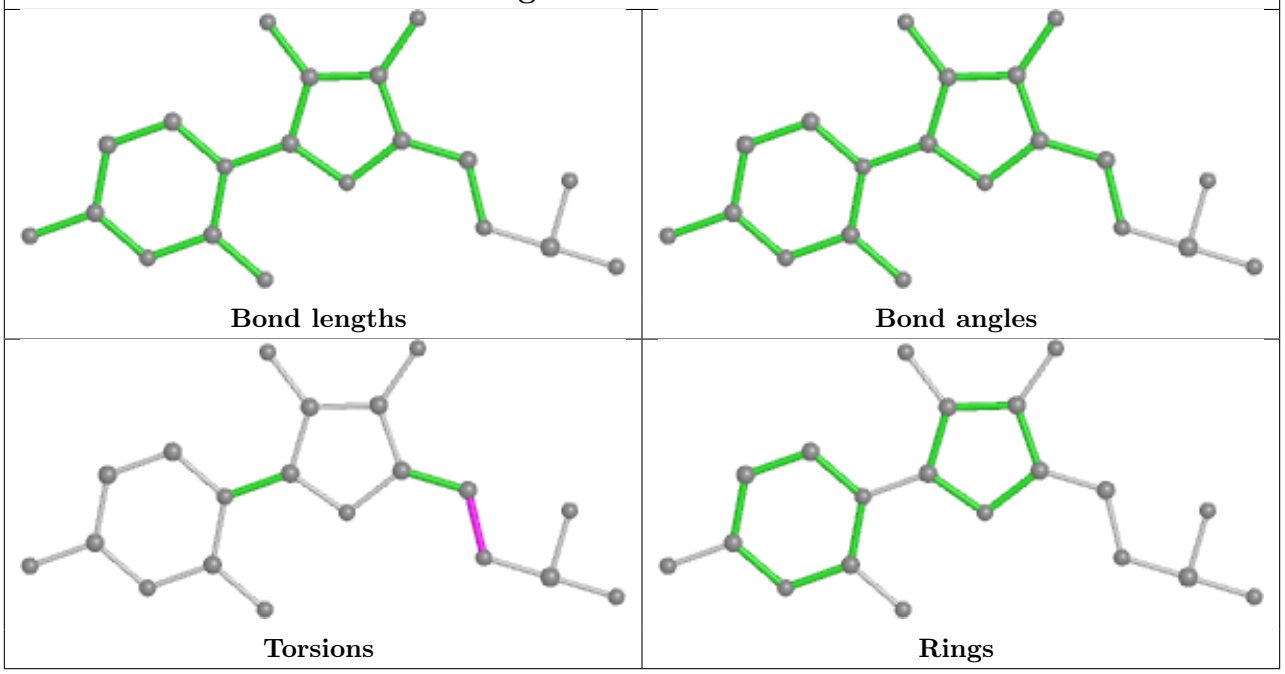
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	N	1604	C5P	1	0

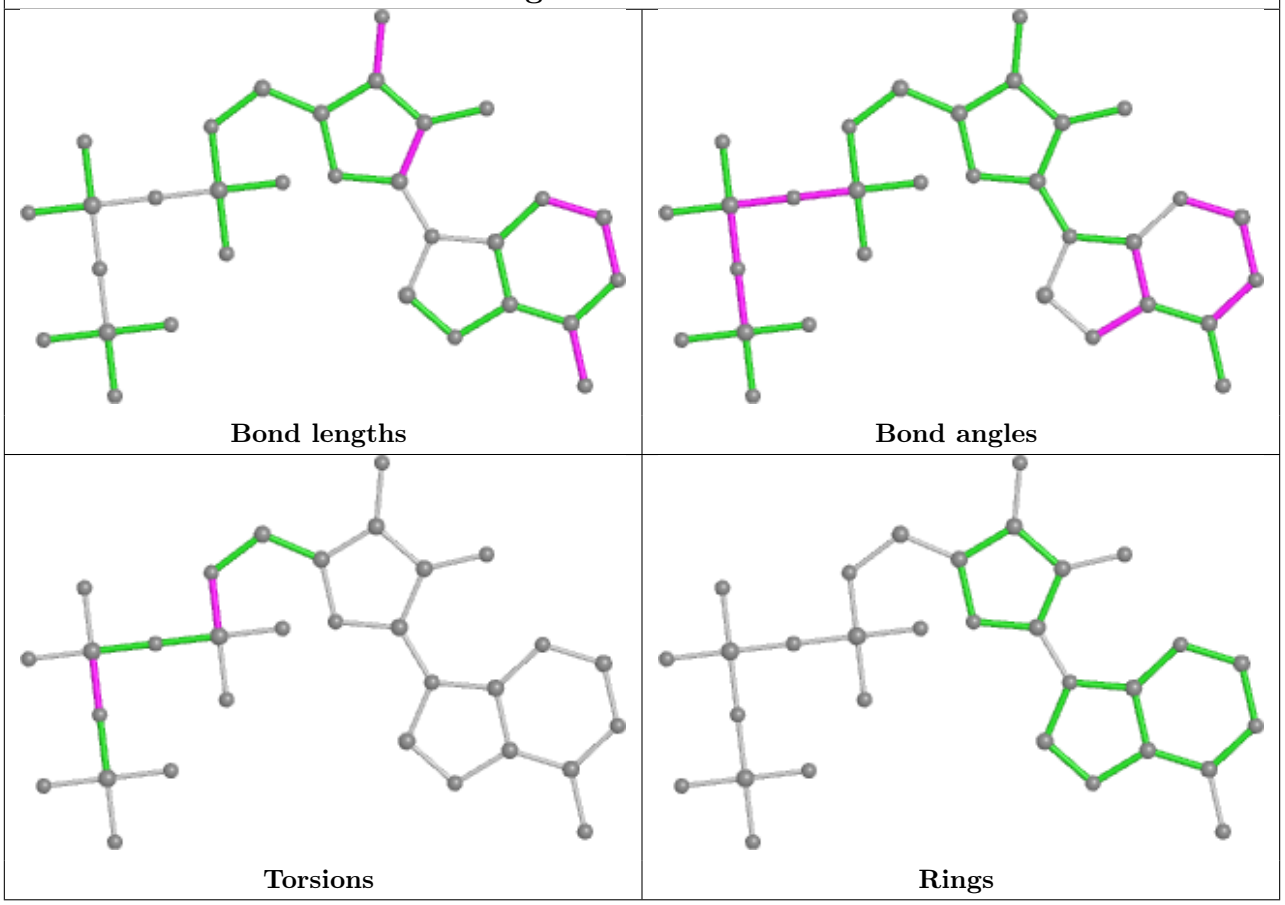
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

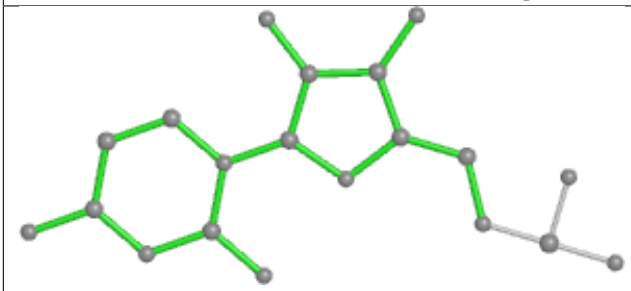
Ligand C5P N 1604



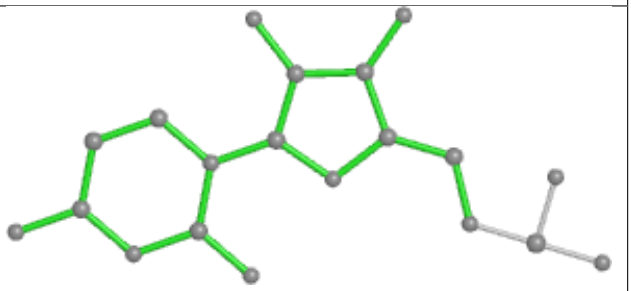
Ligand ATP R 101



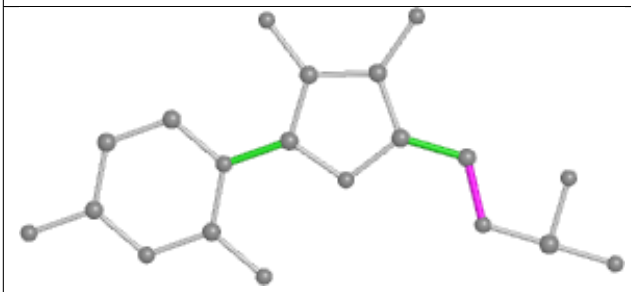
Ligand C5P D 2005



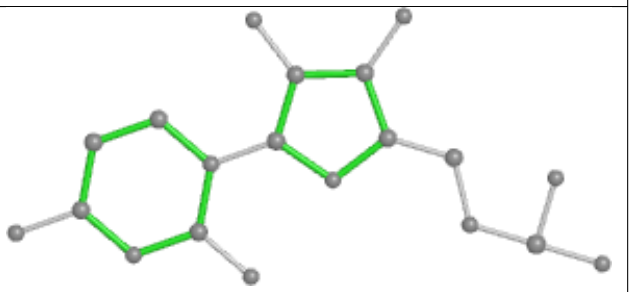
Bond lengths



Bond angles

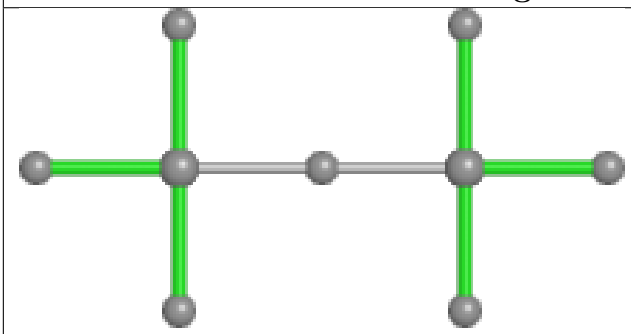


Torsions

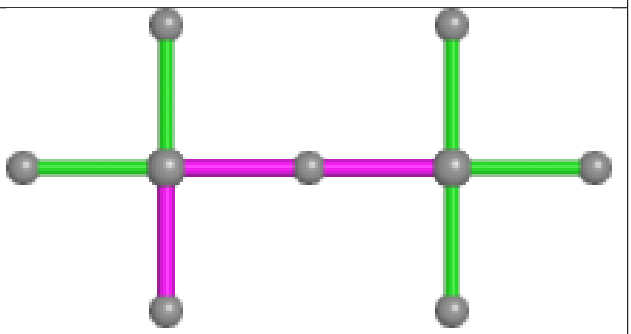


Rings

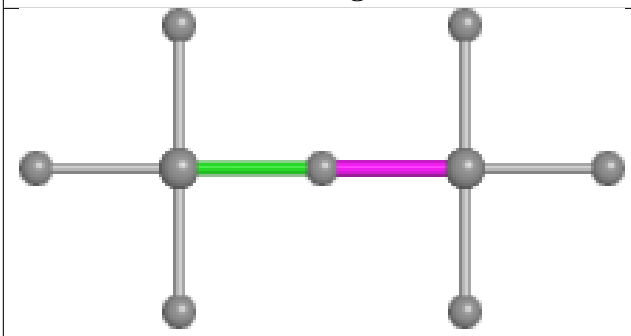
Ligand CTP N 1605



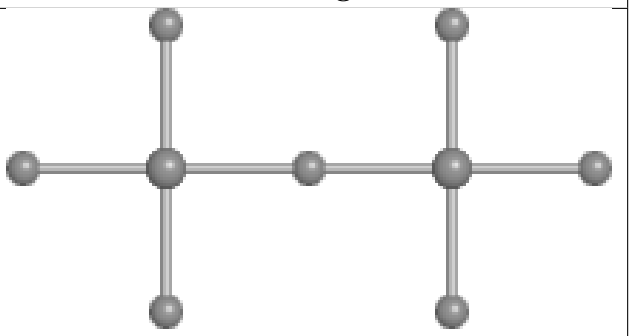
Bond lengths



Bond angles

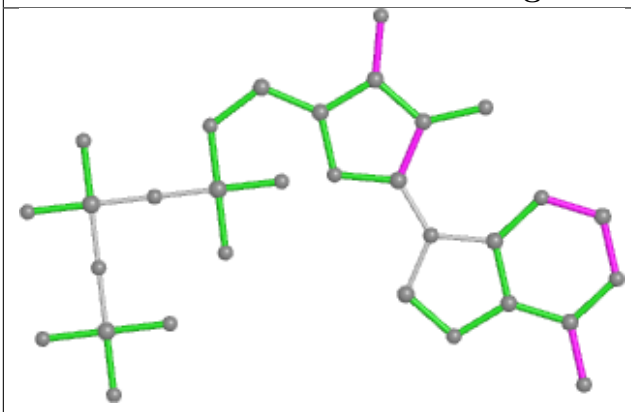


Torsions

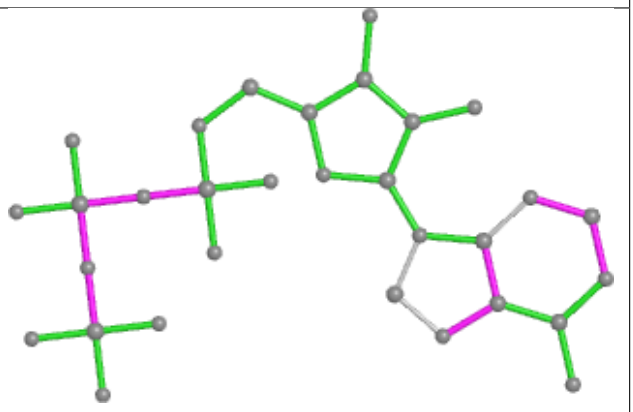


Rings

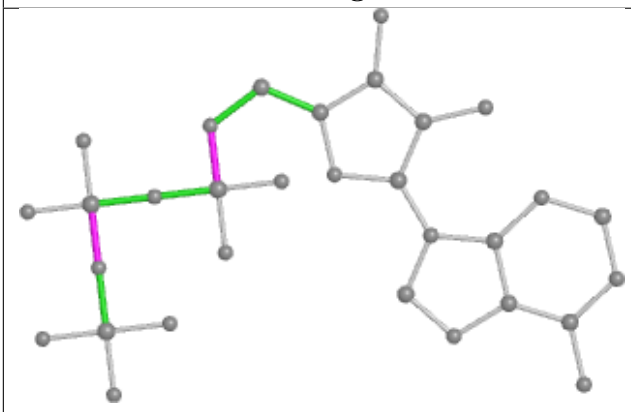
Ligand ATP G 101



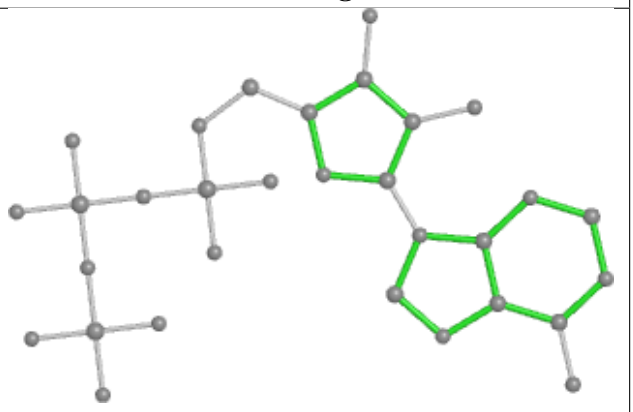
Bond lengths



Bond angles

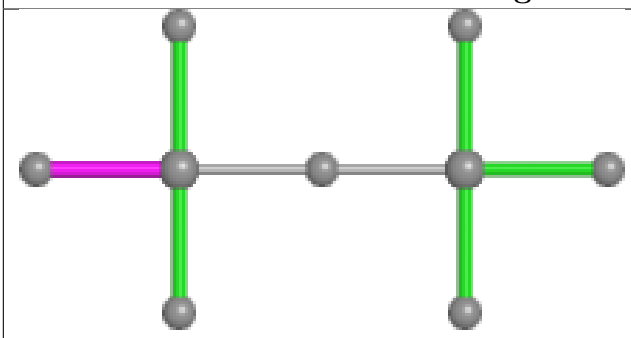


Torsions

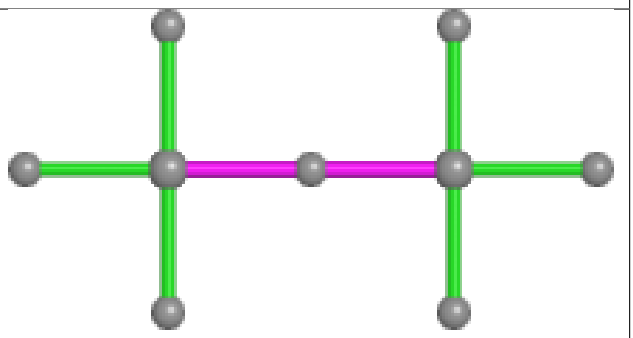


Rings

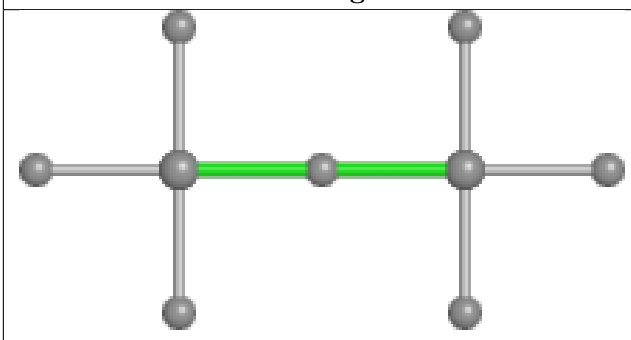
Ligand CTP D 2006



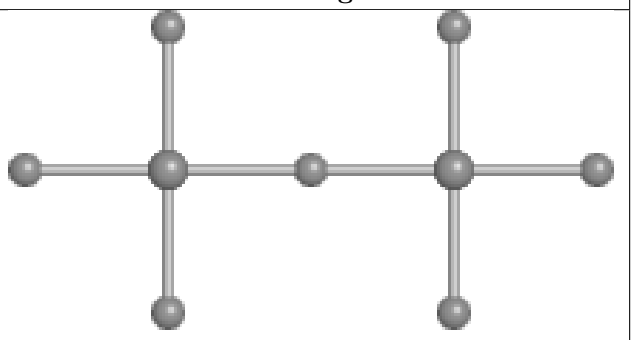
Bond lengths



Bond angles



Torsions



Rings

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	231/315 (73%)	-0.32	3 (1%) 77 76	13, 36, 63, 128	1 (0%)
1	B	222/315 (70%)	-0.25	0 100 100	16, 42, 80, 94	0
1	K	231/315 (73%)	-0.14	4 (1%) 70 67	18, 47, 73, 123	1 (0%)
1	L	222/315 (70%)	-0.27	0 100 100	17, 46, 87, 114	0
2	C	1112/1119 (99%)	-0.28	11 (0%) 82 82	4, 31, 89, 122	3 (0%)
2	M	1080/1119 (96%)	0.18	54 (5%) 28 27	5, 58, 121, 135	2 (0%)
3	D	1486/1524 (97%)	-0.23	15 (1%) 82 82	3, 32, 90, 122	4 (0%)
3	N	1486/1524 (97%)	-0.08	36 (2%) 59 55	3, 40, 103, 143	4 (0%)
4	E	94/99 (94%)	-0.43	0 100 100	11, 29, 67, 94	0
4	O	94/99 (94%)	-0.40	0 100 100	13, 40, 81, 99	0
5	F	346/443 (78%)	-0.10	9 (2%) 56 52	13, 47, 108, 124	0
5	P	346/443 (78%)	0.30	27 (7%) 13 12	29, 74, 142, 167	0
6	G	15/19 (78%)	-0.33	0 100 100	15, 34, 132, 135	0
6	R	15/19 (78%)	-0.41	0 100 100	24, 43, 131, 132	0
7	H	20/27 (74%)	-0.30	0 100 100	36, 66, 117, 125	0
7	S	18/27 (66%)	-0.21	0 100 100	63, 80, 125, 140	0
All	All	7018/7722 (90%)	-0.12	159 (2%) 60 58	3, 42, 105, 167	15 (0%)

The worst 5 of 159 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	P	377	ASP	7.9
5	P	391	GLY	7.5
5	P	381	HIS	5.9
5	P	392	VAL	5.9
3	N	70	GLY	5.2

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)

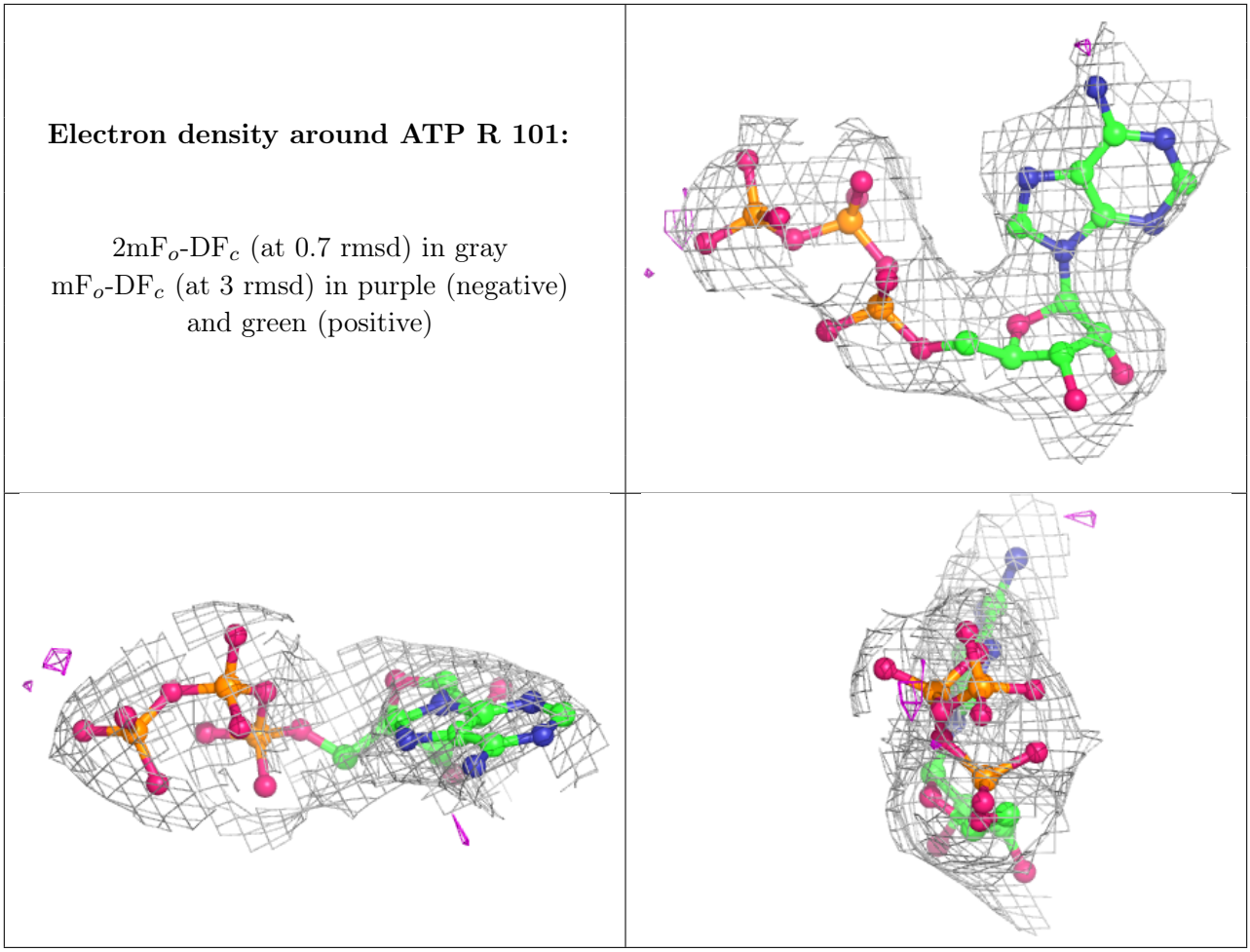
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
8	MG	B	2001	1/1	0.37	0.49	70,70,70,70	0
8	MG	L	2001	1/1	0.70	0.28	72,72,72,72	0
9	ZN	N	1602	1/1	0.91	0.06	128,128,128,128	0
12	ATP	R	101	31/31	0.92	0.17	28,39,97,99	0
10	C5P	N	1604	20/21	0.93	0.17	23,34,41,41	0
11	CTP	D	2006	9/29	0.94	0.30	27,37,51,69	0
10	C5P	D	2005	20/21	0.94	0.17	16,22,28,34	0
8	MG	F	2001	1/1	0.95	0.05	31,31,31,31	0
8	MG	D	2004	1/1	0.95	0.18	29,29,29,29	0
8	MG	N	1606	1/1	0.95	0.30	30,30,30,30	0
12	ATP	G	101	31/31	0.95	0.15	18,28,68,77	0
8	MG	D	2007	1/1	0.95	0.34	17,17,17,17	0
8	MG	K	901	1/1	0.96	0.26	36,36,36,36	0
11	CTP	N	1605	9/29	0.97	0.32	34,40,54,72	0
9	ZN	D	2002	1/1	0.98	0.03	92,92,92,92	0
8	MG	P	2001	1/1	0.98	0.06	77,77,77,77	0
9	ZN	D	2001	1/1	0.99	0.12	10,10,10,10	0
8	MG	N	1603	1/1	0.99	0.21	12,12,12,12	0
9	ZN	N	1601	1/1	0.99	0.14	8,8,8,8	0
8	MG	D	2003	1/1	0.99	0.26	6,6,6,6	0
8	MG	L	2002	1/1	0.99	0.16	25,25,25,25	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.

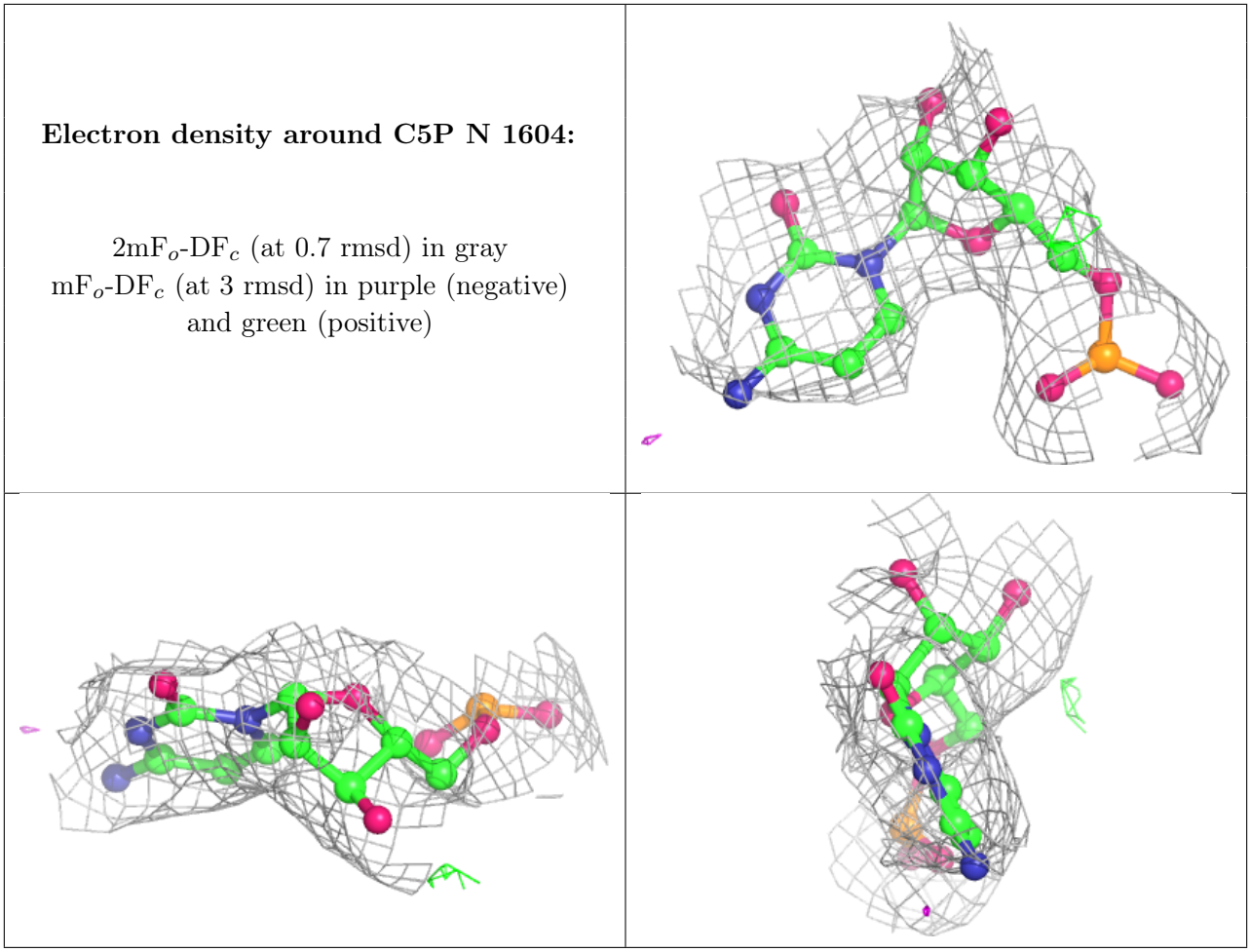
Electron density around ATP R 101:

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



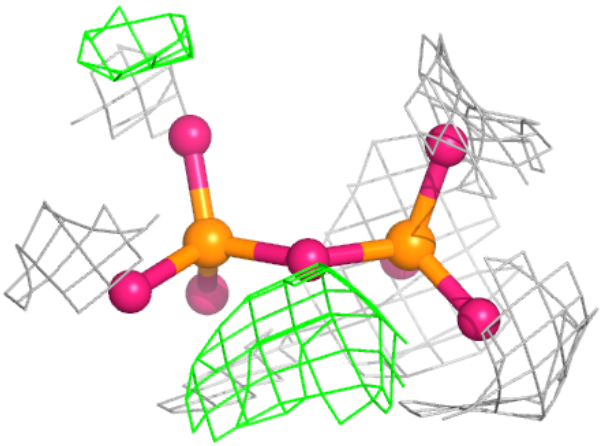
Electron density around C5P N 1604:

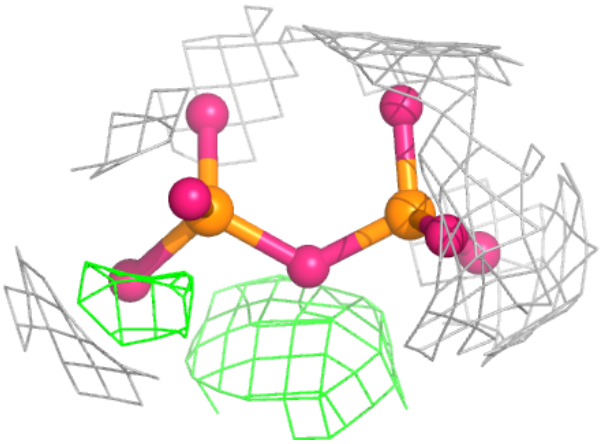
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

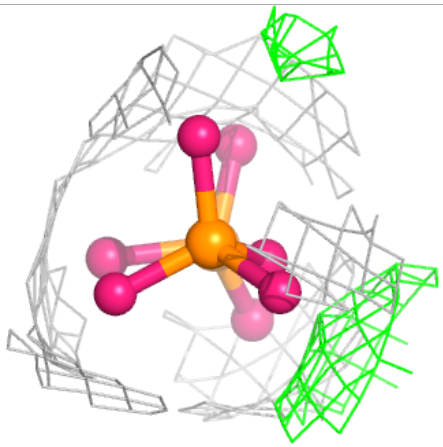


Electron density around CTP D 2006:

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

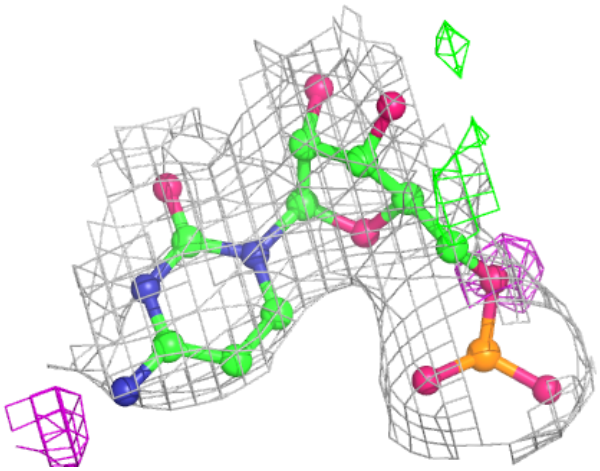


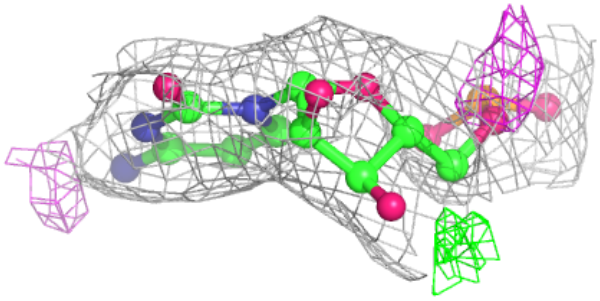




Electron density around C5P D 2005:

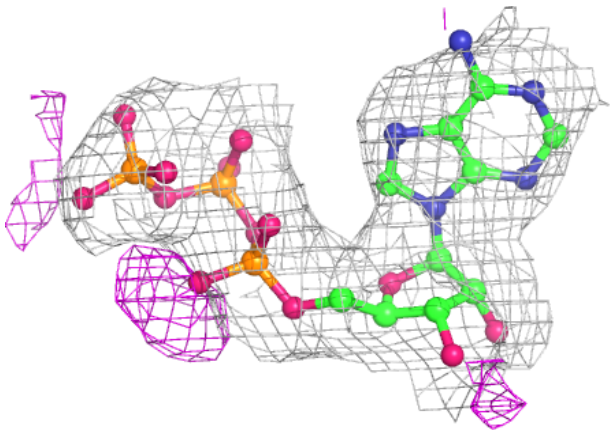
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

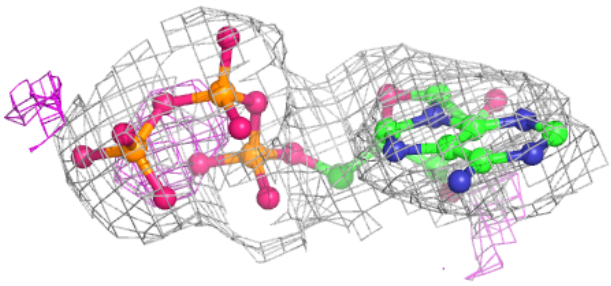


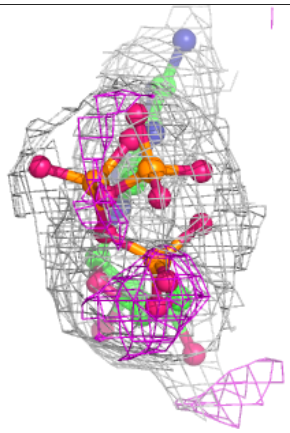


Electron density around ATP G 101:

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

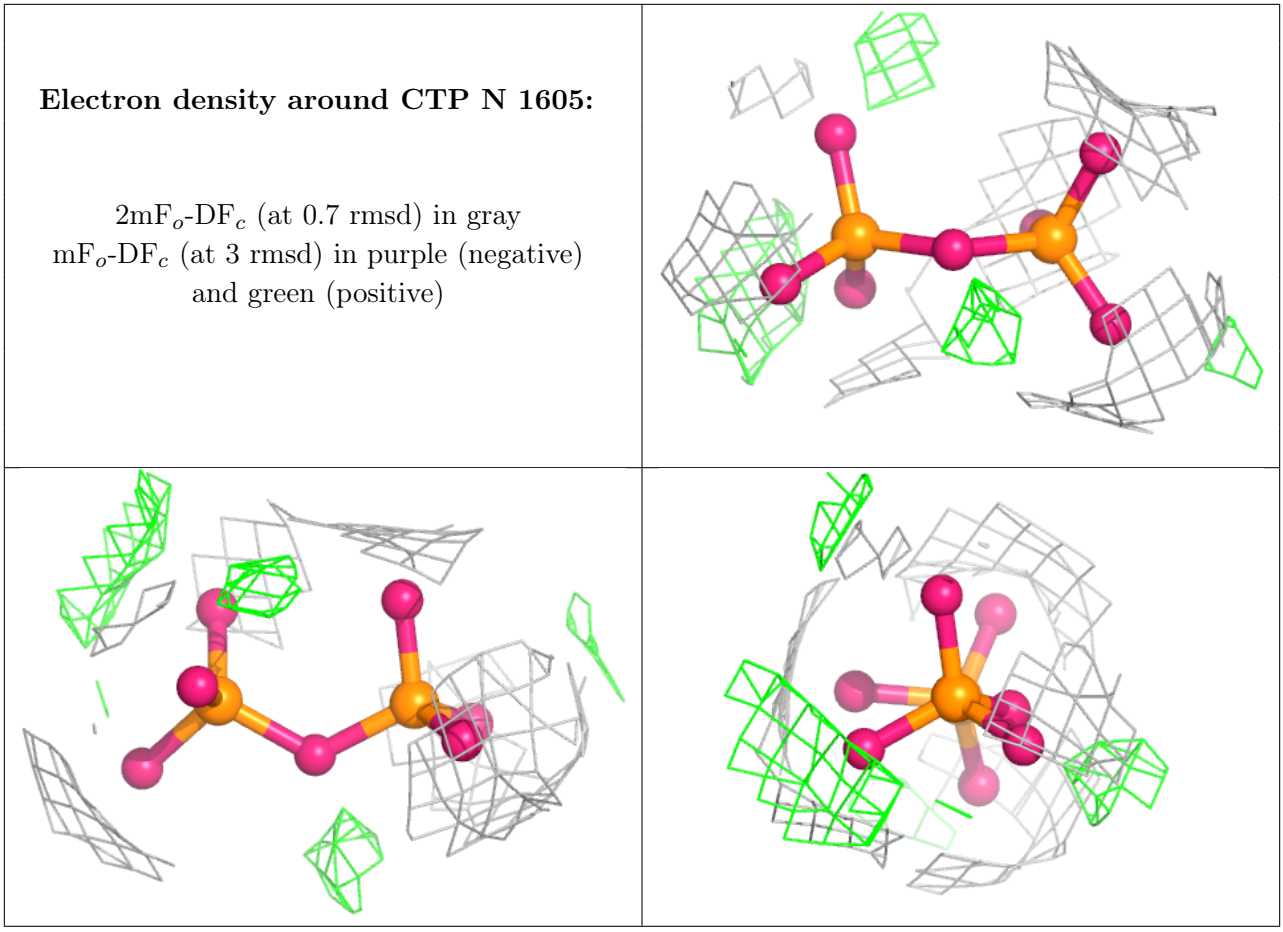






Electron density around CTP N 1605:

$2mF_o$-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

There are no such residues in this entry.