wwPDB X-ray Structure Validation Summary Report (i)

Apr 2, 2024 – 02:12 PM JST

PDB ID	:	8XKF
Title	:	Crystal structure of Helicobacter pylori IspDF with substrate CTP
Authors	:	Chen, X.; Wu, D.
Deposited on	:	2023-12-23
Resolution	:	2.50 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:

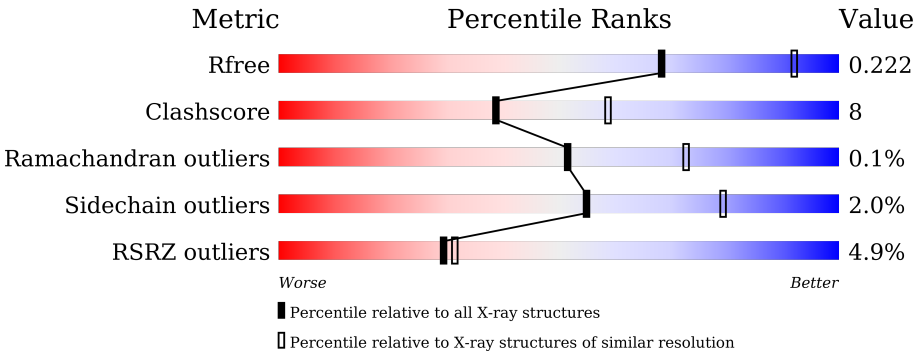
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	381	4% 81% 16% ·
1	B	381	5% 79% 15% · 5%
1	C	381	4% 86% 13% ··
1	D	381	5% 84% 14% ·
1	E	381	4% 81% 17% ··
1	F	381	6% 77% 17% · 5%

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
5	GOL	A	506	-	-	X	-

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 18605 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 Bifunctional enzyme IspD/IspF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	370	Total	C	N	O	S	0	0	0
			2922	1885	483	545	9			
1	B	362	Total	C	N	O	S	0	0	0
			2864	1850	472	533	9			
1	C	377	Total	C	N	O	S	0	0	0
			2985	1925	496	555	9			
1	D	375	Total	C	N	O	S	0	0	0
			2966	1913	492	552	9			
1	E	377	Total	C	N	O	S	0	0	0
			2985	1925	496	555	9			
1	F	361	Total	C	N	O	S	0	0	0
			2856	1843	474	530	9			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	initiating methionine	UNP O25664
A	27	HIS	-	expression tag	UNP O25664
A	28	HIS	-	expression tag	UNP O25664
A	29	HIS	-	expression tag	UNP O25664
A	30	HIS	-	expression tag	UNP O25664
A	31	HIS	-	expression tag	UNP O25664
A	32	HIS	-	expression tag	UNP O25664
B	26	MET	-	initiating methionine	UNP O25664
B	27	HIS	-	expression tag	UNP O25664
B	28	HIS	-	expression tag	UNP O25664
B	29	HIS	-	expression tag	UNP O25664
B	30	HIS	-	expression tag	UNP O25664
B	31	HIS	-	expression tag	UNP O25664
B	32	HIS	-	expression tag	UNP O25664
C	26	MET	-	initiating methionine	UNP O25664
C	27	HIS	-	expression tag	UNP O25664
C	28	HIS	-	expression tag	UNP O25664

Continued on next page...

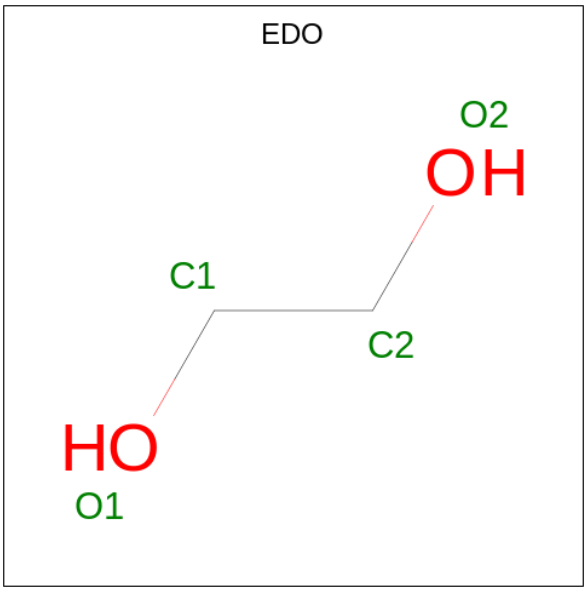
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	29	HIS	-	expression tag	UNP O25664
C	30	HIS	-	expression tag	UNP O25664
C	31	HIS	-	expression tag	UNP O25664
C	32	HIS	-	expression tag	UNP O25664
D	26	MET	-	initiating methionine	UNP O25664
D	27	HIS	-	expression tag	UNP O25664
D	28	HIS	-	expression tag	UNP O25664
D	29	HIS	-	expression tag	UNP O25664
D	30	HIS	-	expression tag	UNP O25664
D	31	HIS	-	expression tag	UNP O25664
D	32	HIS	-	expression tag	UNP O25664
E	26	MET	-	initiating methionine	UNP O25664
E	27	HIS	-	expression tag	UNP O25664
E	28	HIS	-	expression tag	UNP O25664
E	29	HIS	-	expression tag	UNP O25664
E	30	HIS	-	expression tag	UNP O25664
E	31	HIS	-	expression tag	UNP O25664
E	32	HIS	-	expression tag	UNP O25664
F	26	MET	-	initiating methionine	UNP O25664
F	27	HIS	-	expression tag	UNP O25664
F	28	HIS	-	expression tag	UNP O25664
F	29	HIS	-	expression tag	UNP O25664
F	30	HIS	-	expression tag	UNP O25664
F	31	HIS	-	expression tag	UNP O25664
F	32	HIS	-	expression tag	UNP O25664

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

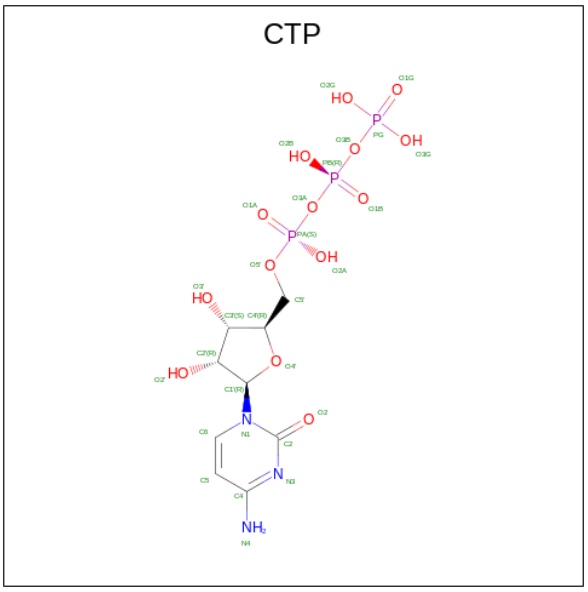
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



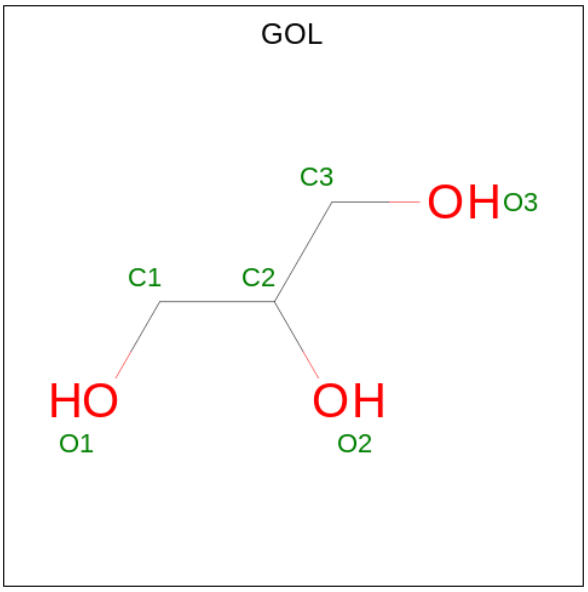
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
4	B	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
4	C	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
4	D	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
4	E	1	Total	C	N	O	P	0	0
			29	9	3	14	3		
4	F	1	Total	C	N	O	P	0	0
			29	9	3	14	3		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

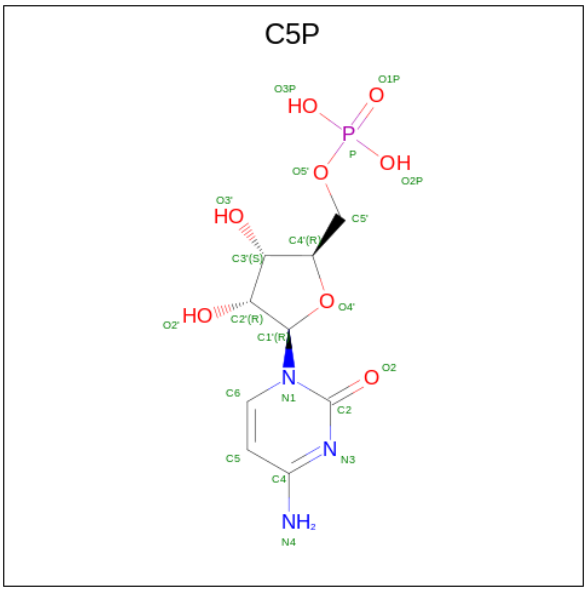


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	F	1	Total C O 6 3 3	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

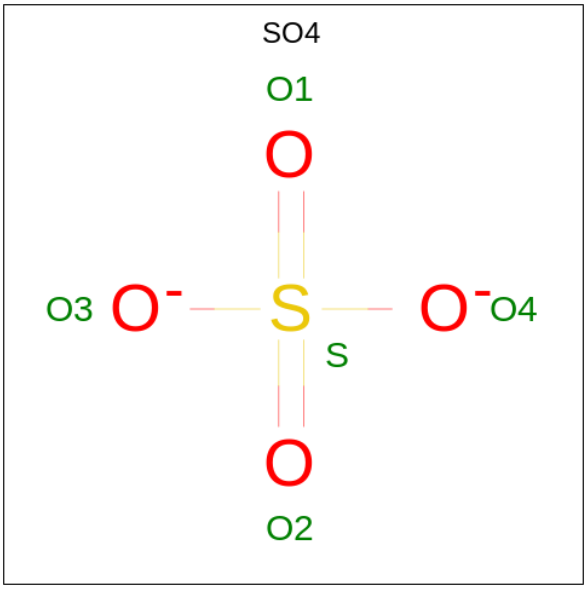
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Zn 1 1	0	0
6	D	1	Total Zn 1 1	0	0

- Molecule 7 is CYTIDINE-5'-MONOPHOSPHATE (three-letter code: C5P) (formula: $C_9H_{14}N_3O_8P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	B	1	Total	C	N	O	P	0	0
			21	9	3	8	1		
7	D	1	Total	C	N	O	P	0	0
			21	9	3	8	1		
7	F	1	Total	C	N	O	P	0	0
			21	9	3	8	1		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



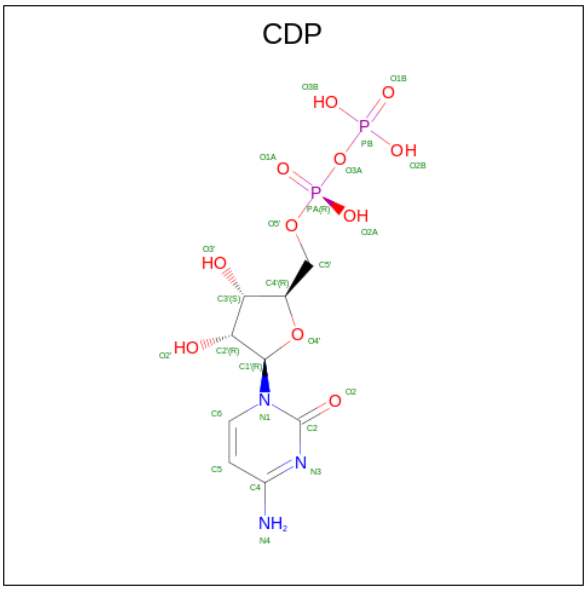
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	E	1	Total O S 5 4 1	0	0

- Molecule 9 is CYTIDINE-5'-DIPHOSPHATE (three-letter code: CDP) (formula: $C_9H_{15}N_3O_{11}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	1	Total C N O P 25 9 3 11 2	0	0
9	E	1	Total C N O P 25 9 3 11 2	0	0

- Molecule 10 is water.

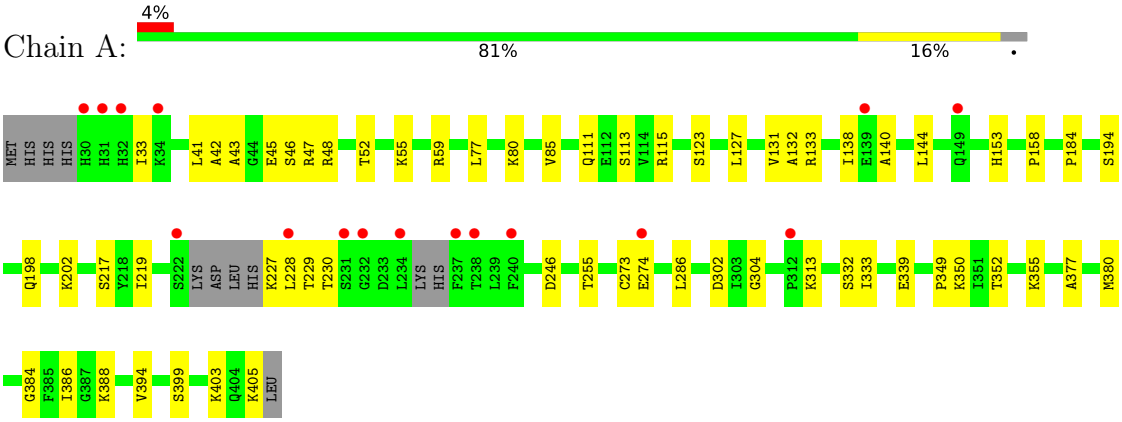
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	126	Total O 126 126	0	0
10	B	100	Total O 100 100	0	0
10	C	113	Total O 113 113	0	0
10	D	128	Total O 128 128	0	0
10	E	118	Total O 118 118	0	0
10	F	88	Total O 88 88	0	0

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.

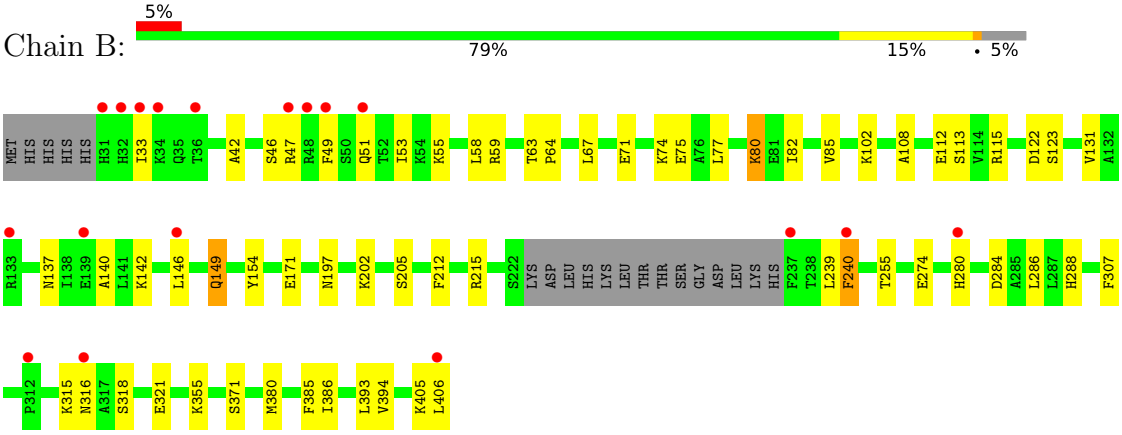
• Molecule 1: Bifunctional enzyme IspD/IspF

Chain A:



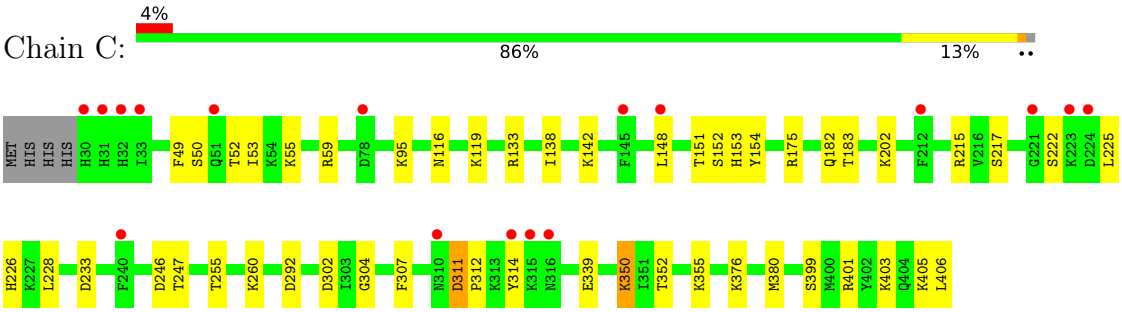
• Molecule 1: Bifunctional enzyme IspD/IspF

Chain B:



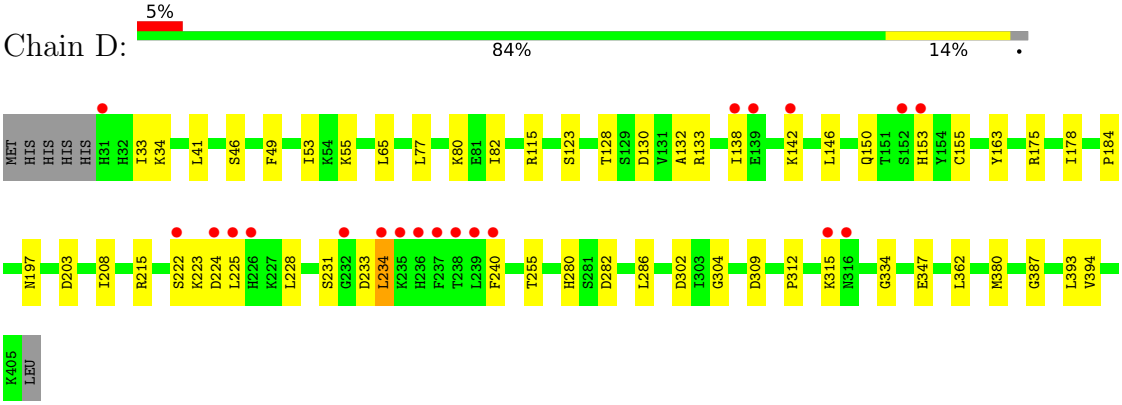
• Molecule 1: Bifunctional enzyme IspD/IspF

Chain C:



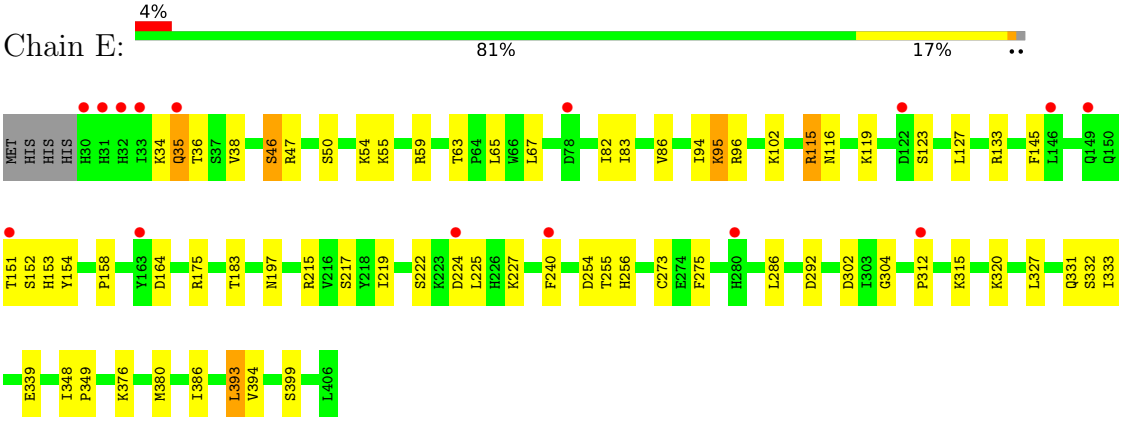
• Molecule 1: Bifunctional enzyme IspD/IspF

Chain D:



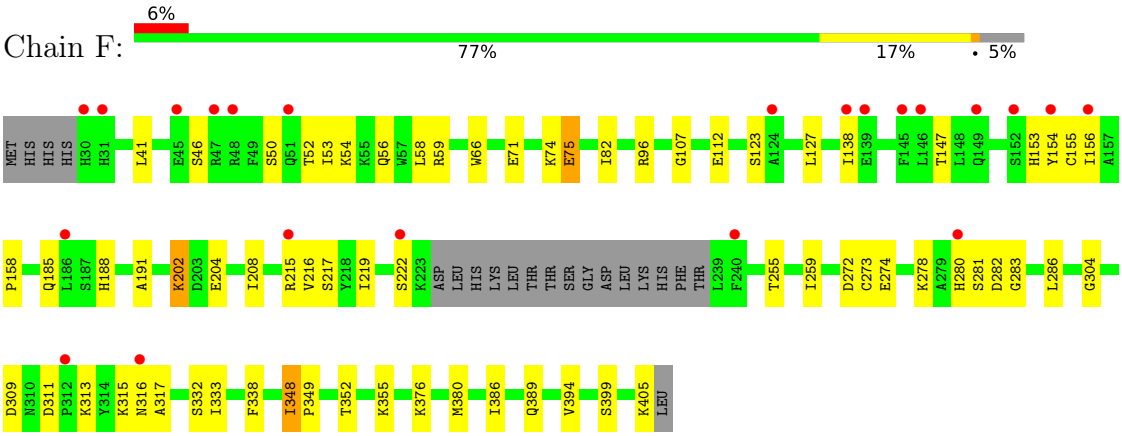
• Molecule 1: Bifunctional enzyme IspD/IspF

Chain E:



• Molecule 1: Bifunctional enzyme IspD/IspF

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	146.62Å 146.62Å 256.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.56 – 2.50 47.56 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.3 (47.56-2.50) 99.3 (47.56-2.50)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.63 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, R_{free}	0.192 , 0.227 0.196 , 0.222	Depositor DCC
R_{free} test set	5654 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	36.6	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 41.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.016 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	18605	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.61% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CTP, SO4, C5P, CDP, GOL, EDO, CL

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	0.49	0/2982	0.62	0/4025
1	B	0.48	0/2924	0.62	0/3947
1	C	0.44	0/3049	0.65	0/4116
1	D	0.48	0/3029	0.60	0/4090
1	E	0.41	0/3049	0.62	0/4116
1	F	0.45	0/2916	0.61	0/3936
All	All	0.46	0/17949	0.62	0/24230

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	2922	0	2956	42	0
1	B	2864	0	2900	48	0
1	C	2985	0	3024	39	0
1	D	2966	0	3006	53	0
1	E	2985	0	3024	45	0
1	F	2856	0	2893	45	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	12	0	17	1	0
3	B	4	0	6	1	0
3	D	4	0	6	2	0
3	E	4	0	6	0	0
4	A	29	0	12	3	0
4	B	29	0	12	2	0
4	C	29	0	12	2	0
4	D	29	0	12	1	0
4	E	29	0	12	2	0
4	F	29	0	12	2	0
5	A	6	0	8	4	0
5	C	12	0	16	4	0
5	D	6	0	8	2	0
5	F	6	0	8	0	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	B	21	0	11	2	0
7	D	21	0	11	2	0
7	F	21	0	11	2	0
8	B	5	0	0	0	0
8	E	5	0	0	0	0
9	C	25	0	10	1	0
9	E	25	0	10	4	0
10	A	126	0	0	1	0
10	B	100	0	0	1	0
10	C	113	0	0	0	0
10	D	128	0	0	0	0
10	E	118	0	0	2	0
10	F	88	0	0	2	0
All	All	18605	0	18003	270	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 8.

The worst 5 of 270 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)
7:F:502:C5P:C4'	7:F:502:C5P:O4'	1.68	1.17
1:D:312:PRO:CA	1:D:315:LYS:HE3	1.78	1.14
1:D:312:PRO:HA	1:D:315:LYS:HE3	1.14	1.12
1:F:71:GLU:O	1:F:75:GLU:HG3	1.50	1.11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:B:503:C5P:C4'	7:B:503:C5P:O4'	1.68	1.10

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	364/381 (96%)	359 (99%)	5 (1%)	0	100	100
1	B	358/381 (94%)	351 (98%)	7 (2%)	0	100	100
1	C	375/381 (98%)	360 (96%)	15 (4%)	0	100	100
1	D	373/381 (98%)	365 (98%)	7 (2%)	1 (0%)	41	61
1	E	375/381 (98%)	367 (98%)	7 (2%)	1 (0%)	41	61
1	F	357/381 (94%)	351 (98%)	5 (1%)	1 (0%)	41	61
All	All	2202/2286 (96%)	2153 (98%)	46 (2%)	3 (0%)	51	73

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	152	SER
1	F	282	ASP
1	D	282	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	318/329 (97%)	318 (100%)	0	100	100
1	B	311/329 (94%)	300 (96%)	11 (4%)	36	62
1	C	325/329 (99%)	319 (98%)	6 (2%)	59	81
1	D	323/329 (98%)	320 (99%)	3 (1%)	78	92
1	E	325/329 (99%)	317 (98%)	8 (2%)	47	73
1	F	310/329 (94%)	299 (96%)	11 (4%)	36	62
All	All	1912/1974 (97%)	1873 (98%)	39 (2%)	55	79

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	50	SER
1	F	272	ASP
1	F	52	THR
1	F	202	LYS
1	F	348	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	193	GLN
1	E	331	GLN
1	D	197	ASN
1	F	31	HIS
1	E	35	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 27 ligands modelled in this entry, 3 are monoatomic - leaving 24 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	$\#\|Z\| > 2$
3	EDO	A	502	-	3,3,3	0.73	0	2,2,2	0.30	0
8	SO4	E	504	-	4,4,4	0.14	0	6,6,6	0.25	0
5	GOL	C	504	-	5,5,5	0.43	0	5,5,5	0.47	0
8	SO4	B	504	-	4,4,4	0.15	0	6,6,6	0.23	0
5	GOL	A	506	-	5,5,5	0.68	0	5,5,5	0.58	0
4	CTP	F	501	-	26,30,30	3.82	15 (57%)	39,47,47	1.05	2 (5%)
3	EDO	A	503	-	3,3,3	0.47	0	2,2,2	0.49	0
5	GOL	F	503	-	5,5,5	0.46	0	5,5,5	0.41	0
4	CTP	B	502	-	26,30,30	3.78	15 (57%)	39,47,47	0.93	2 (5%)
3	EDO	A	504	-	3,3,3	0.77	0	2,2,2	0.51	0
5	GOL	D	504	-	5,5,5	0.48	0	5,5,5	0.29	0
7	C5P	D	503	-	22,22,22	4.32	14 (63%)	33,33,33	0.95	1 (3%)
5	GOL	C	503	-	5,5,5	0.43	0	5,5,5	0.31	0
4	CTP	C	501	-	26,30,30	3.82	15 (57%)	39,47,47	1.23	4 (10%)
9	CDP	E	503	-	24,26,26	3.97	13 (54%)	37,40,40	1.79	11 (29%)
3	EDO	E	501	-	3,3,3	0.60	0	2,2,2	0.72	0
7	C5P	B	503	-	22,22,22	4.31	14 (63%)	33,33,33	0.80	0
3	EDO	B	501	-	3,3,3	0.61	0	2,2,2	0.17	0
4	CTP	A	505	-	26,30,30	3.79	15 (57%)	39,47,47	0.96	2 (5%)
3	EDO	D	501	-	3,3,3	0.68	0	2,2,2	0.34	0
4	CTP	E	502	-	26,30,30	3.75	15 (57%)	39,47,47	1.05	3 (7%)
7	C5P	F	502	-	22,22,22	4.36	14 (63%)	33,33,33	0.96	0
4	CTP	D	502	-	26,30,30	3.79	15 (57%)	39,47,47	1.04	2 (5%)
9	CDP	C	502	6	24,26,26	4.14	13 (54%)	37,40,40	1.62	8 (21%)

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
3	EDO	A	502	-	-	1/1/1/1	-
5	GOL	C	504	-	-	2/4/4/4	-
5	GOL	A	506	-	-	2/4/4/4	-
4	CTP	F	501	-	-	4/22/38/38	0/2/2/2
3	EDO	A	503	-	-	1/1/1/1	-
5	GOL	F	503	-	-	0/4/4/4	-
4	CTP	B	502	-	-	1/22/38/38	0/2/2/2
3	EDO	A	504	-	-	1/1/1/1	-
5	GOL	D	504	-	-	4/4/4/4	-
7	C5P	D	503	-	-	3/10/26/26	0/2/2/2
5	GOL	C	503	-	-	3/4/4/4	-
4	CTP	C	501	-	-	2/22/38/38	0/2/2/2
9	CDP	E	503	-	-	3/16/32/32	0/2/2/2
3	EDO	E	501	-	-	1/1/1/1	-
7	C5P	B	503	-	-	1/10/26/26	0/2/2/2
3	EDO	B	501	-	-	1/1/1/1	-
4	CTP	A	505	-	-	6/22/38/38	0/2/2/2
3	EDO	D	501	-	-	1/1/1/1	-
4	CTP	E	502	-	-	6/22/38/38	0/2/2/2
7	C5P	F	502	-	-	1/10/26/26	0/2/2/2
4	CTP	D	502	-	-	4/22/38/38	0/2/2/2
9	CDP	C	502	6	-	2/16/32/32	0/2/2/2

The worst 5 of 158 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	C	502	CDP	C2'-C3'	-11.52	1.21	1.53
9	E	503	CDP	C2'-C3'	-10.45	1.24	1.53
7	F	502	C5P	O4'-C4'	10.39	1.68	1.45
7	D	503	C5P	O4'-C4'	10.38	1.68	1.45
7	B	503	C5P	O4'-C4'	10.34	1.68	1.45

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	E	503	CDP	O2-C2-N3	-3.91	115.97	122.33
4	C	501	CTP	PB-O3B-PG	-3.57	120.58	132.83
4	C	501	CTP	O4'-C1'-N1	3.51	116.38	108.36
9	C	502	CDP	C5-C6-N1	-3.24	116.39	121.81
4	F	501	CTP	C3'-C2'-C1'	3.22	107.54	101.43

There are no chirality outliers.

5 of 50 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	502	CTP	C5'-O5'-PA-O1A
4	F	501	CTP	C5'-O5'-PA-O1A
4	F	501	CTP	C5'-O5'-PA-O2A
4	F	501	CTP	PB-O3A-PA-O5'
5	A	506	GOL	C1-C2-C3-O3

There are no ring outliers.

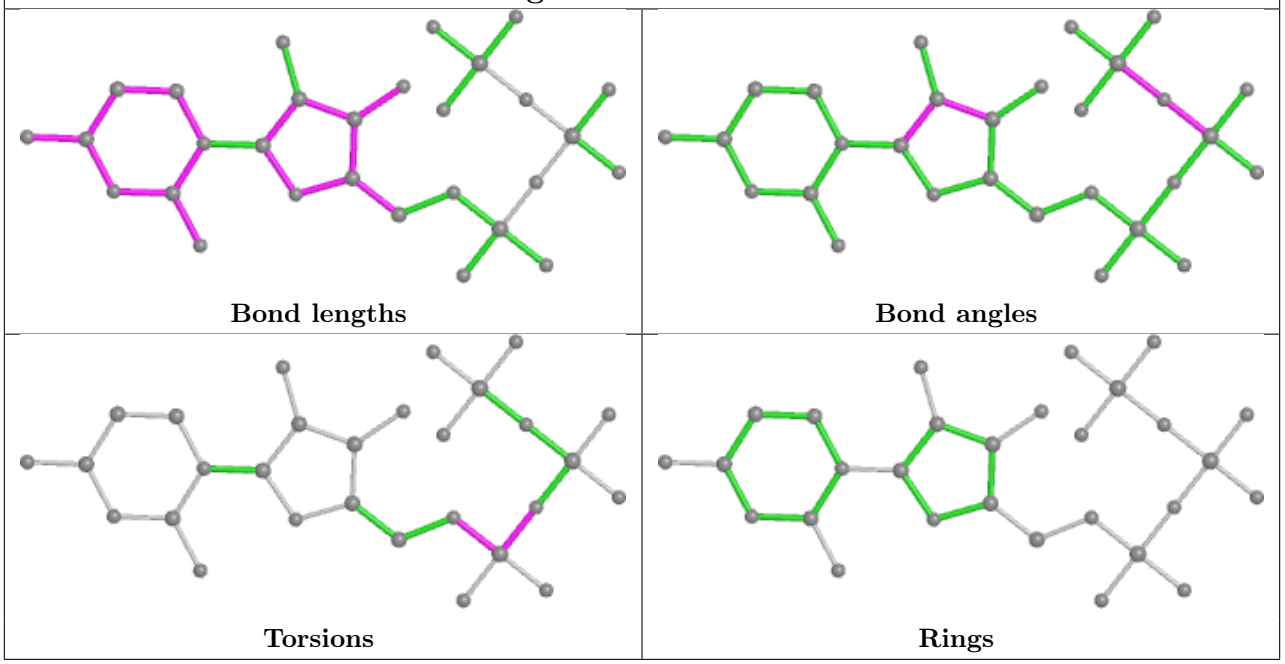
18 monomers are involved in 33 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	EDO	1	0
5	C	504	GOL	1	0
5	A	506	GOL	4	0
4	F	501	CTP	2	0
4	B	502	CTP	2	0
5	D	504	GOL	2	0
7	D	503	C5P	2	0
5	C	503	GOL	3	0
4	C	501	CTP	2	0
9	E	503	CDP	4	0
7	B	503	C5P	2	0
3	B	501	EDO	1	0
4	A	505	CTP	3	0
3	D	501	EDO	2	0
4	E	502	CTP	2	0
7	F	502	C5P	2	0
4	D	502	CTP	1	0
9	C	502	CDP	1	0

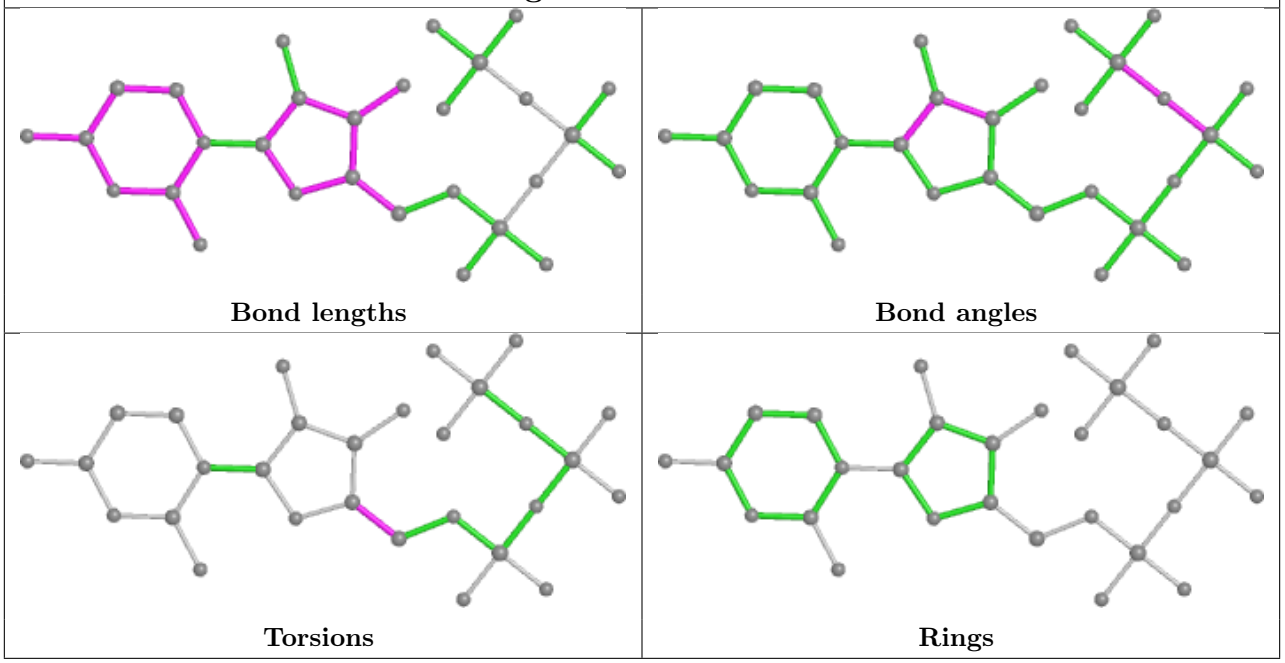
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

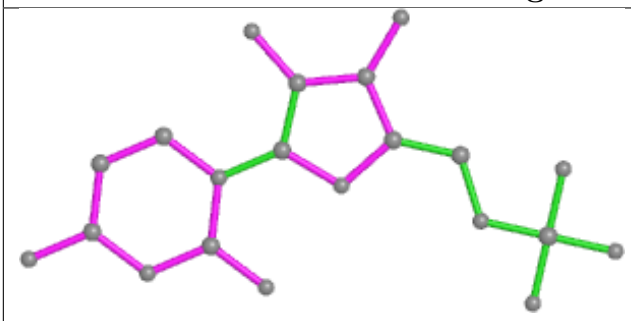
Ligand CTP F 501



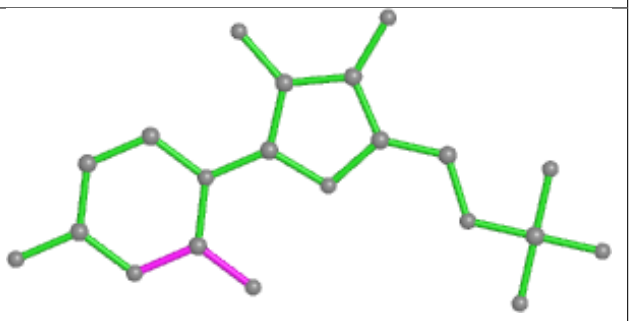
Ligand CTP B 502



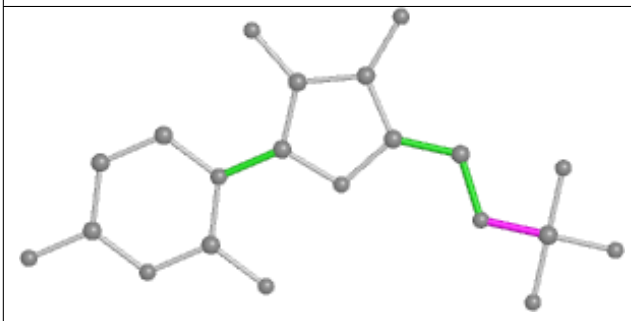
Ligand C5P D 503



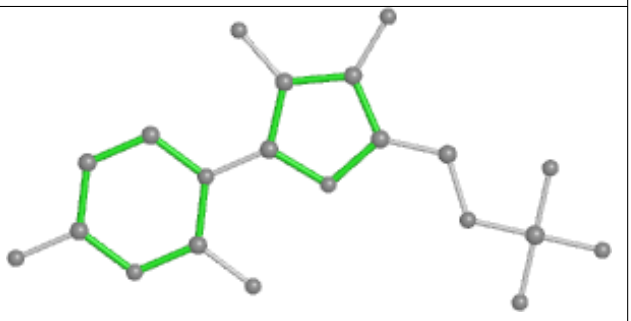
Bond lengths



Bond angles

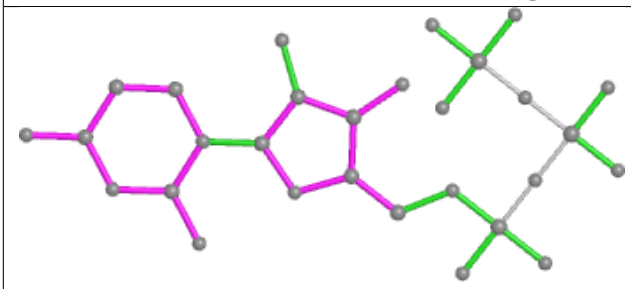


Torsions

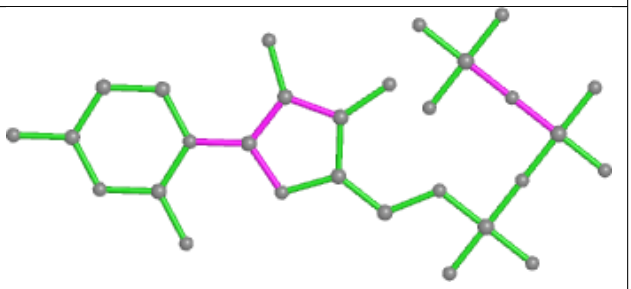


Rings

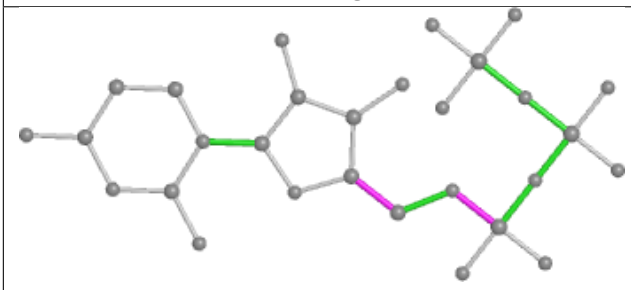
Ligand CTP C 501



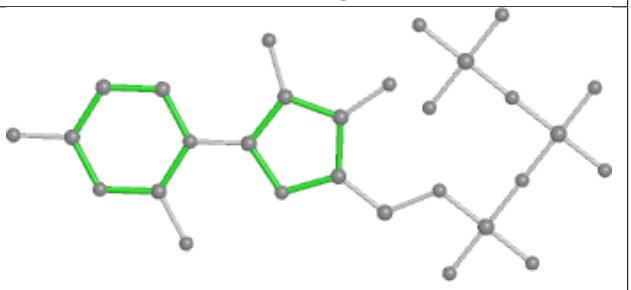
Bond lengths



Bond angles

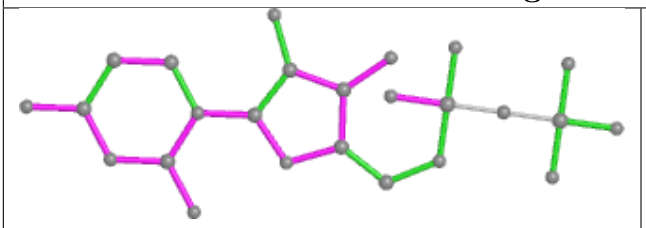


Torsions

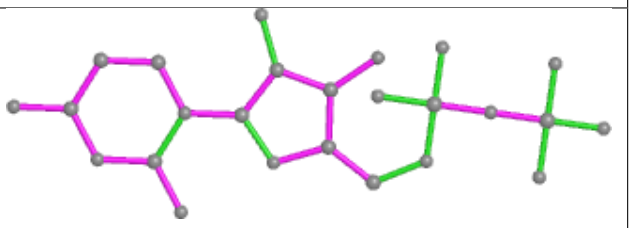


Rings

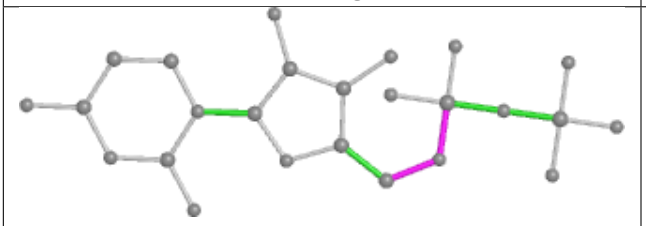
Ligand CDP E 503



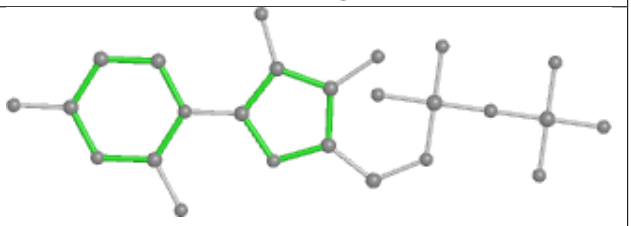
Bond lengths



Bond angles

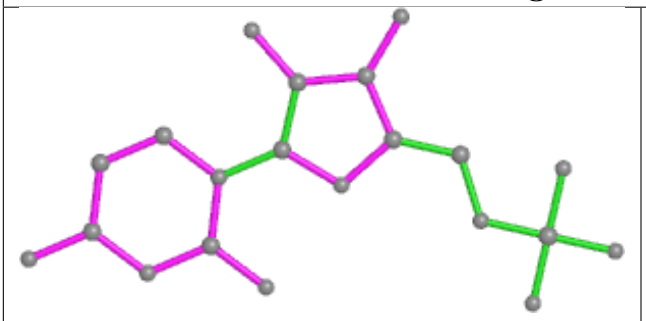


Torsions

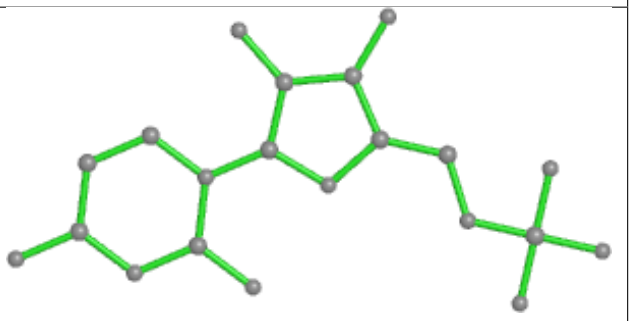


Rings

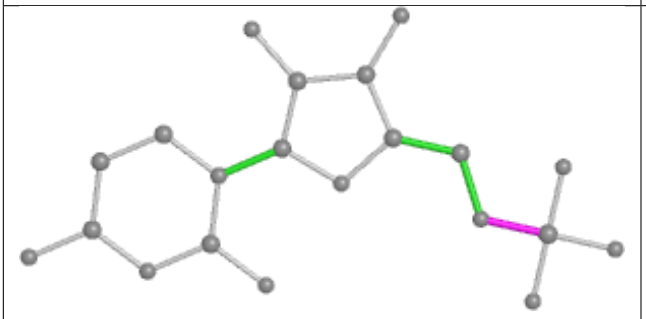
Ligand C5P B 503



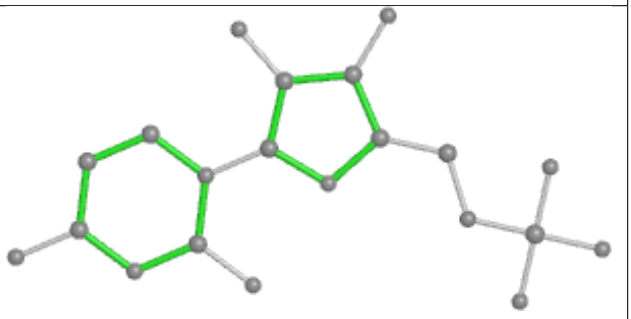
Bond lengths



Bond angles

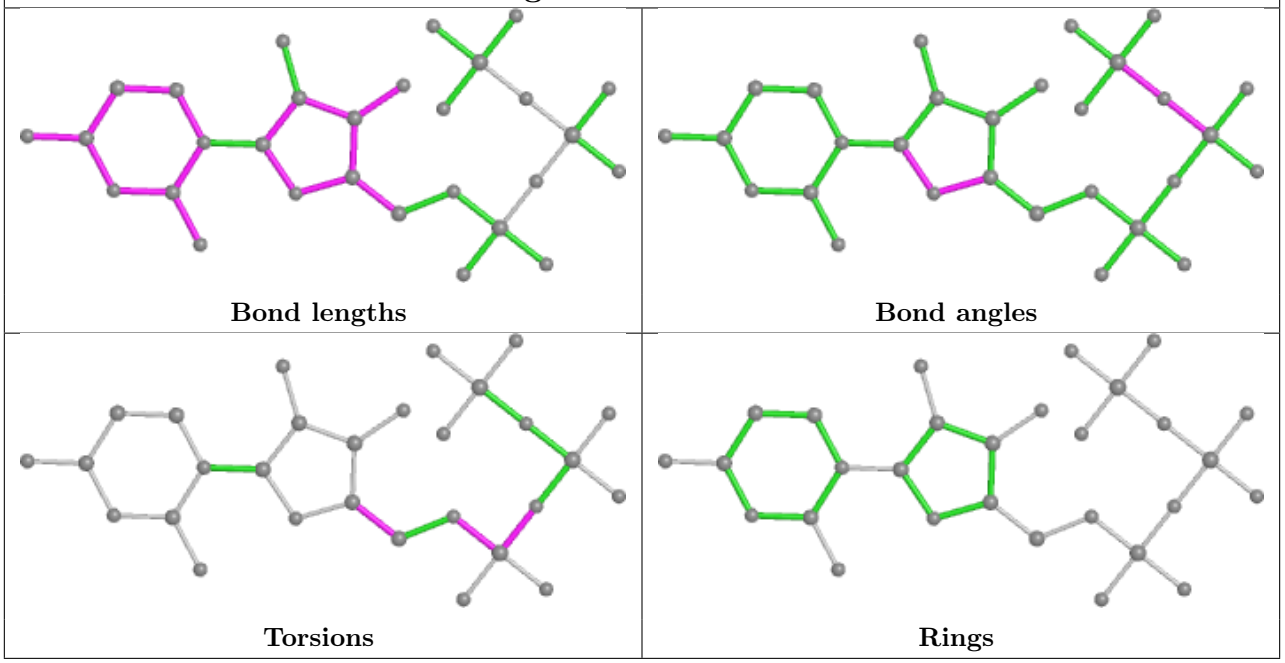


Torsions

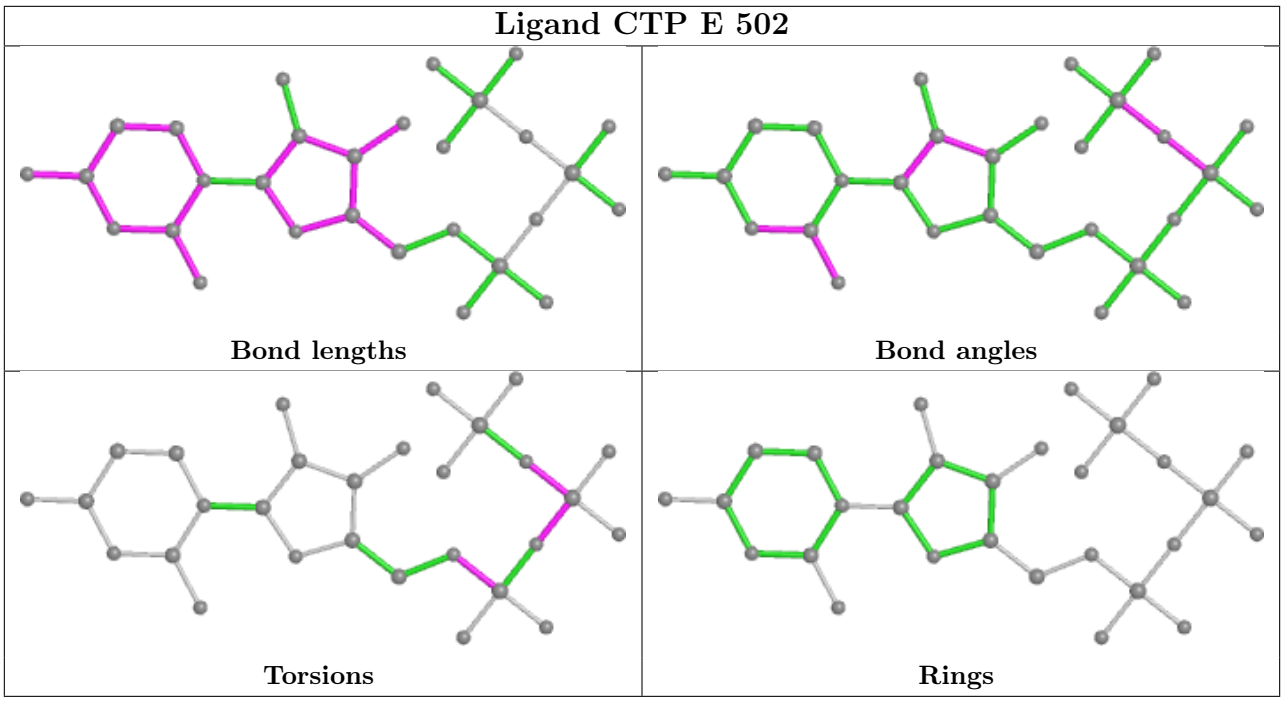


Rings

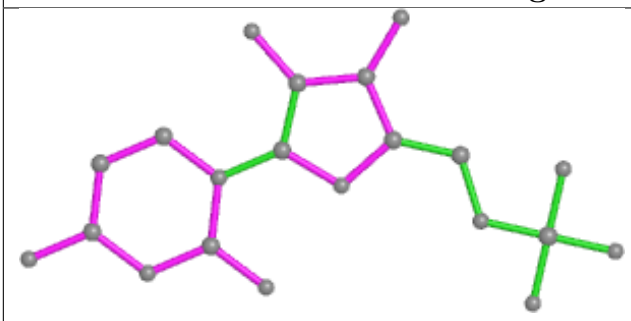
Ligand CTP A 505



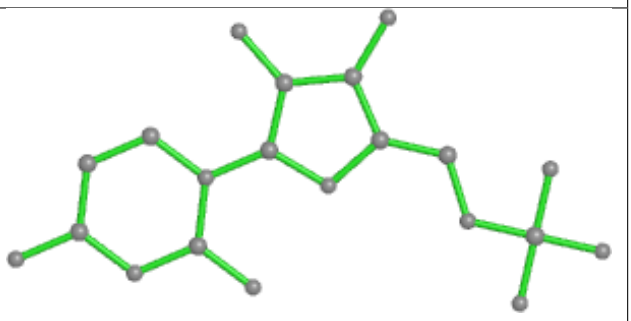
Ligand CTP E 502



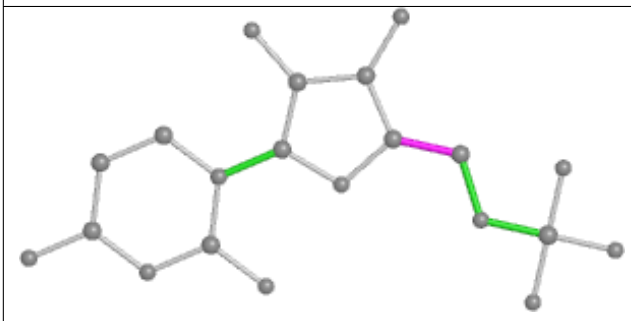
Ligand C5P F 502



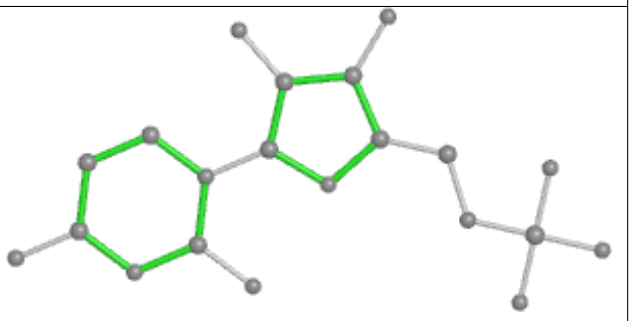
Bond lengths



Bond angles

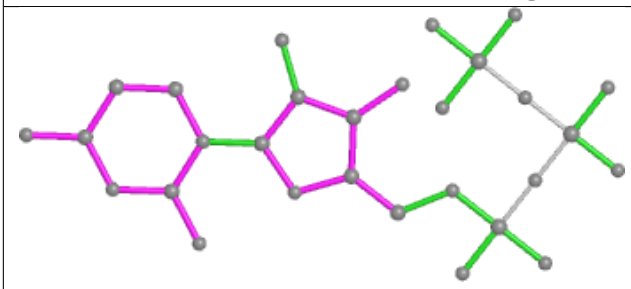


Torsions

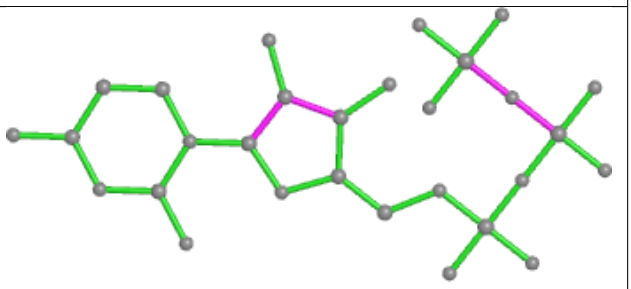


Rings

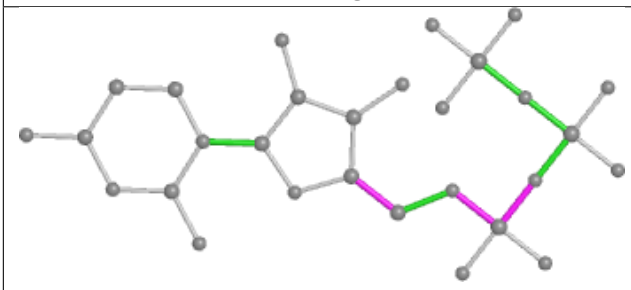
Ligand CTP D 502



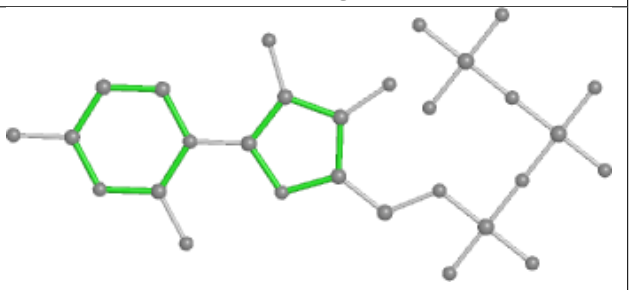
Bond lengths



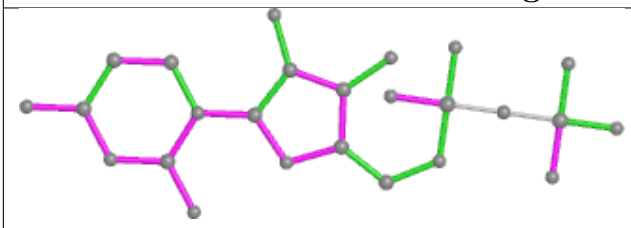
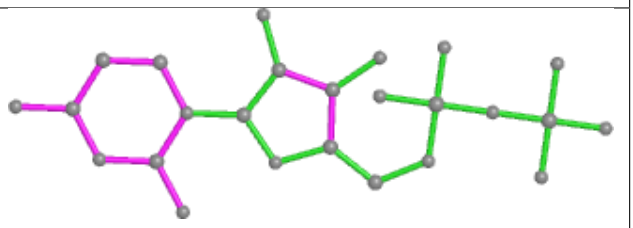
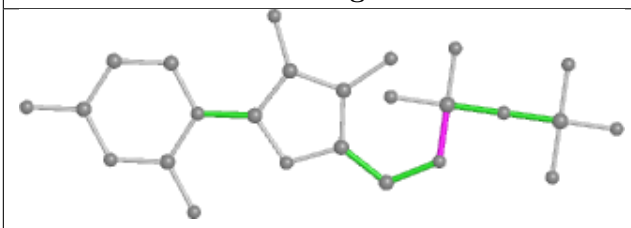
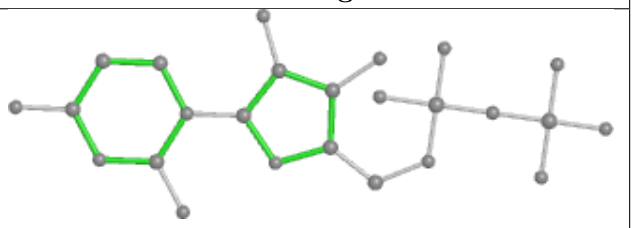
Bond angles



Torsions



Rings

Ligand CDP C 502	
 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	370/381 (97%)	0.05	16 (4%)	35	38	20, 33, 61, 76	0
1	B	362/381 (95%)	0.13	18 (4%)	28	30	21, 39, 68, 87	0
1	C	377/381 (98%)	0.09	17 (4%)	33	36	21, 36, 68, 79	0
1	D	375/381 (98%)	0.12	20 (5%)	26	28	21, 35, 64, 79	0
1	E	377/381 (98%)	0.04	15 (3%)	38	41	22, 38, 64, 76	0
1	F	361/381 (94%)	0.28	22 (6%)	21	22	24, 43, 68, 87	0
All	All	2222/2286 (97%)	0.12	108 (4%)	29	31	20, 37, 66, 87	0

The worst 5 of 108 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	280	HIS	4.8
1	D	236	HIS	4.8
1	E	30	HIS	4.6
1	F	31	HIS	4.6
1	F	312	PRO	4.5

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands (i)

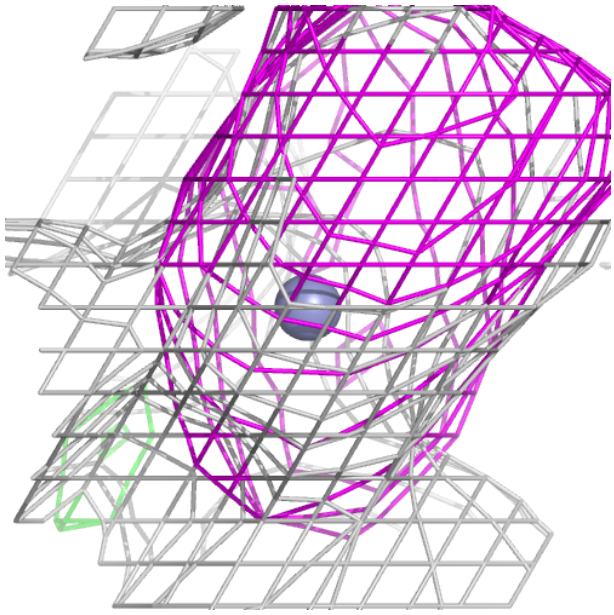
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

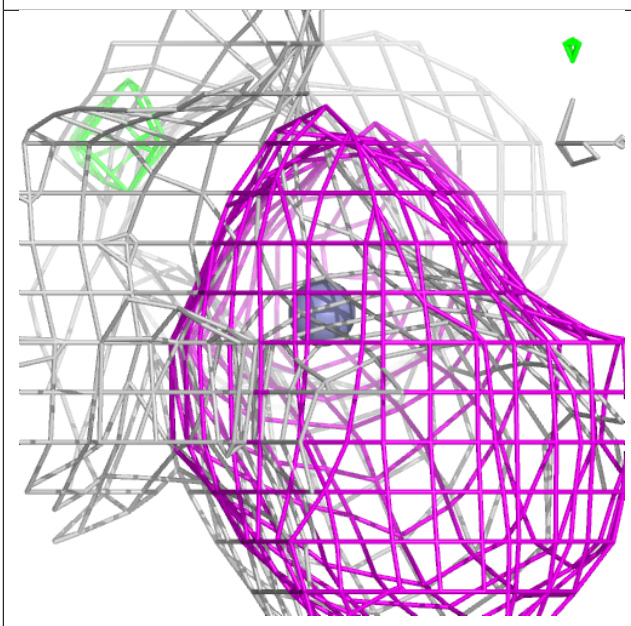
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EDO	B	501	4/4	0.60	0.24	50,53,55,55	0
6	ZN	A	507	1/1	0.75	0.29	84,84,84,84	0
3	EDO	A	503	4/4	0.81	0.18	37,40,50,54	0
3	EDO	E	501	4/4	0.85	0.16	45,46,48,51	0
7	C5P	D	503	21/21	0.86	0.27	34,40,60,75	0
5	GOL	F	503	6/6	0.87	0.13	33,34,35,35	0
9	CDP	E	503	25/25	0.87	0.22	31,41,63,79	0
3	EDO	D	501	4/4	0.88	0.15	43,44,47,49	0
7	C5P	F	502	21/21	0.88	0.18	40,46,58,70	0
5	GOL	A	506	6/6	0.88	0.21	28,34,35,40	0
3	EDO	A	502	4/4	0.89	0.15	33,34,39,39	0
2	CL	A	501	1/1	0.89	0.14	46,46,46,46	0
9	CDP	C	502	25/25	0.89	0.20	30,39,64,79	0
6	ZN	D	505	1/1	0.89	0.24	86,86,86,86	0
7	C5P	B	503	21/21	0.90	0.16	32,38,50,54	0
5	GOL	C	503	6/6	0.91	0.10	31,34,38,38	0
4	CTP	F	501	29/29	0.91	0.18	40,51,78,82	0
4	CTP	B	502	29/29	0.91	0.20	40,52,80,88	0
4	CTP	A	505	29/29	0.94	0.14	32,42,64,67	0
3	EDO	A	504	4/4	0.95	0.13	33,35,37,44	0
5	GOL	C	504	6/6	0.95	0.10	30,32,32,35	0
4	CTP	C	501	29/29	0.96	0.14	31,40,56,63	0
4	CTP	D	502	29/29	0.96	0.12	33,41,59,66	0
4	CTP	E	502	29/29	0.96	0.12	29,38,52,53	0
5	GOL	D	504	6/6	0.98	0.10	39,41,42,45	0
8	SO4	B	504	5/5	0.99	0.12	28,29,31,37	0
8	SO4	E	504	5/5	0.99	0.11	32,35,38,39	0

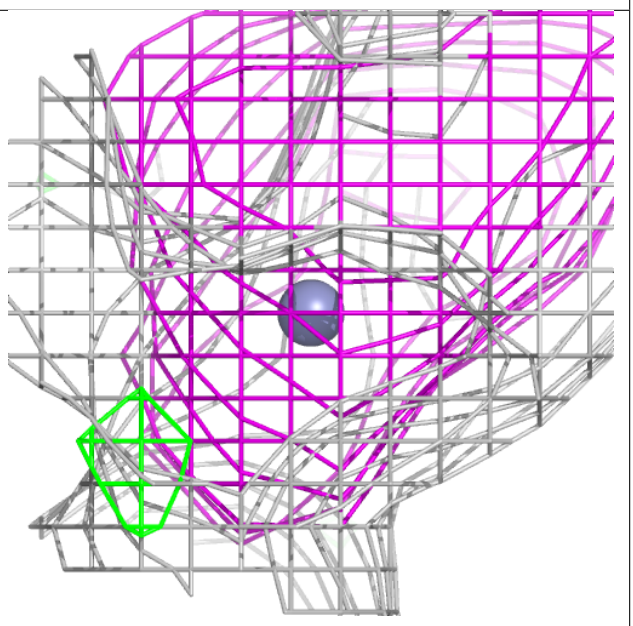
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ZN A 507:

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

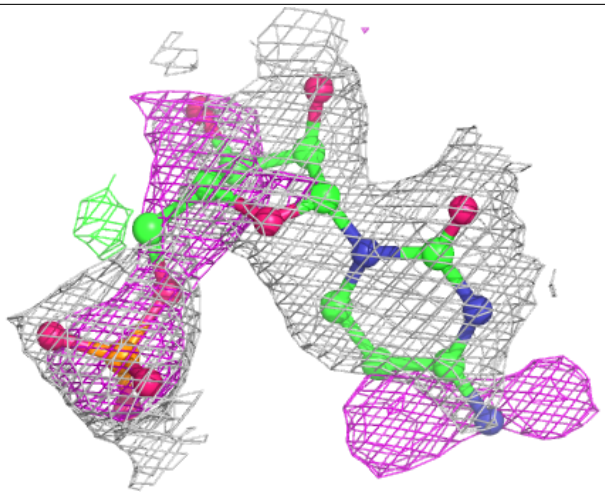


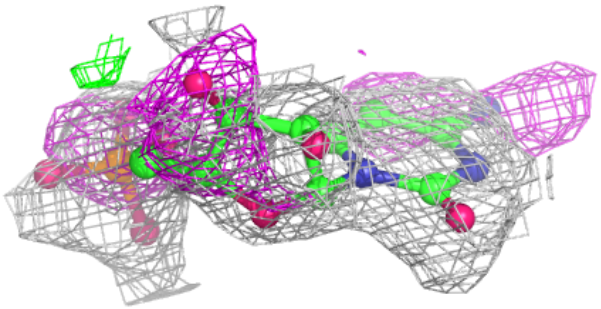


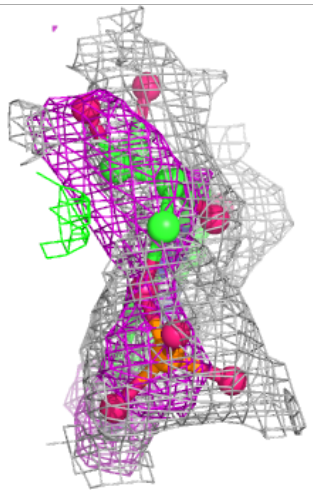


Electron density around C5P D 503:

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

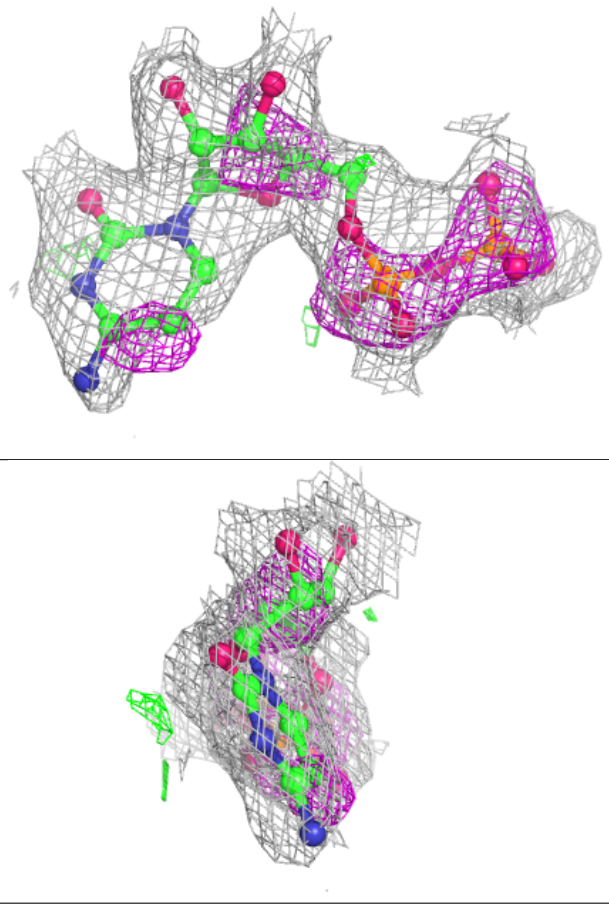






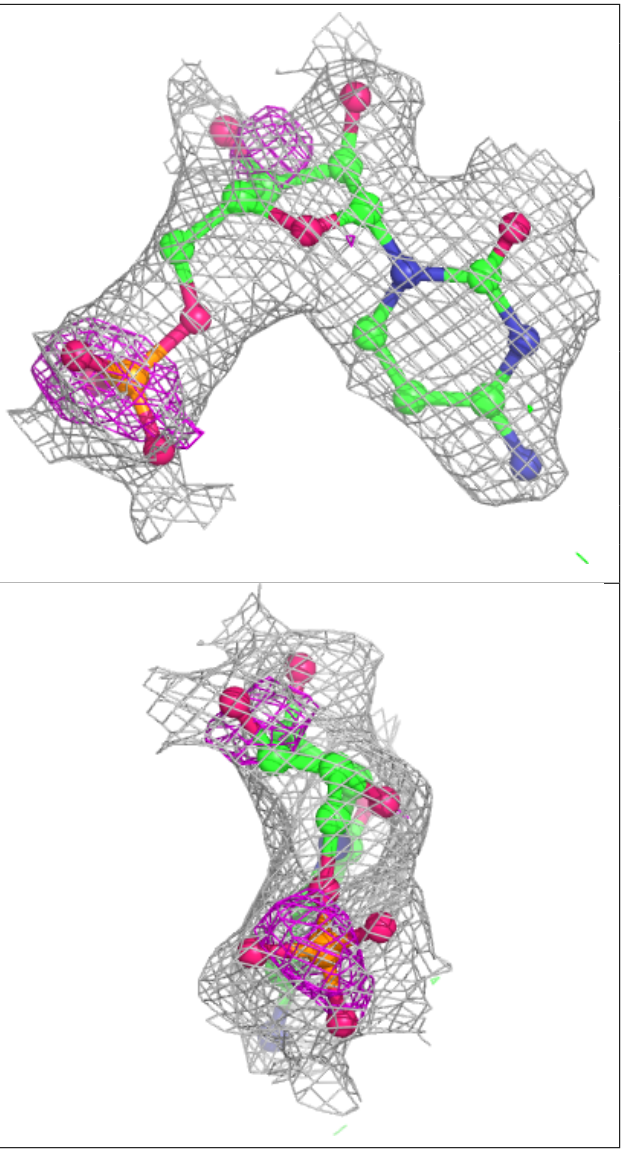
Electron density around CDP E 503:

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



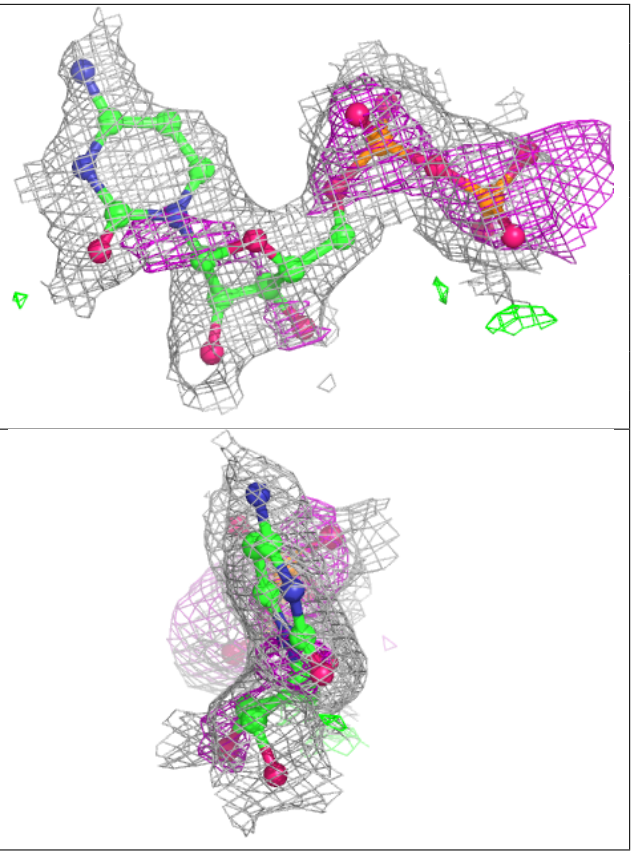
Electron density around C5P F 502:

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



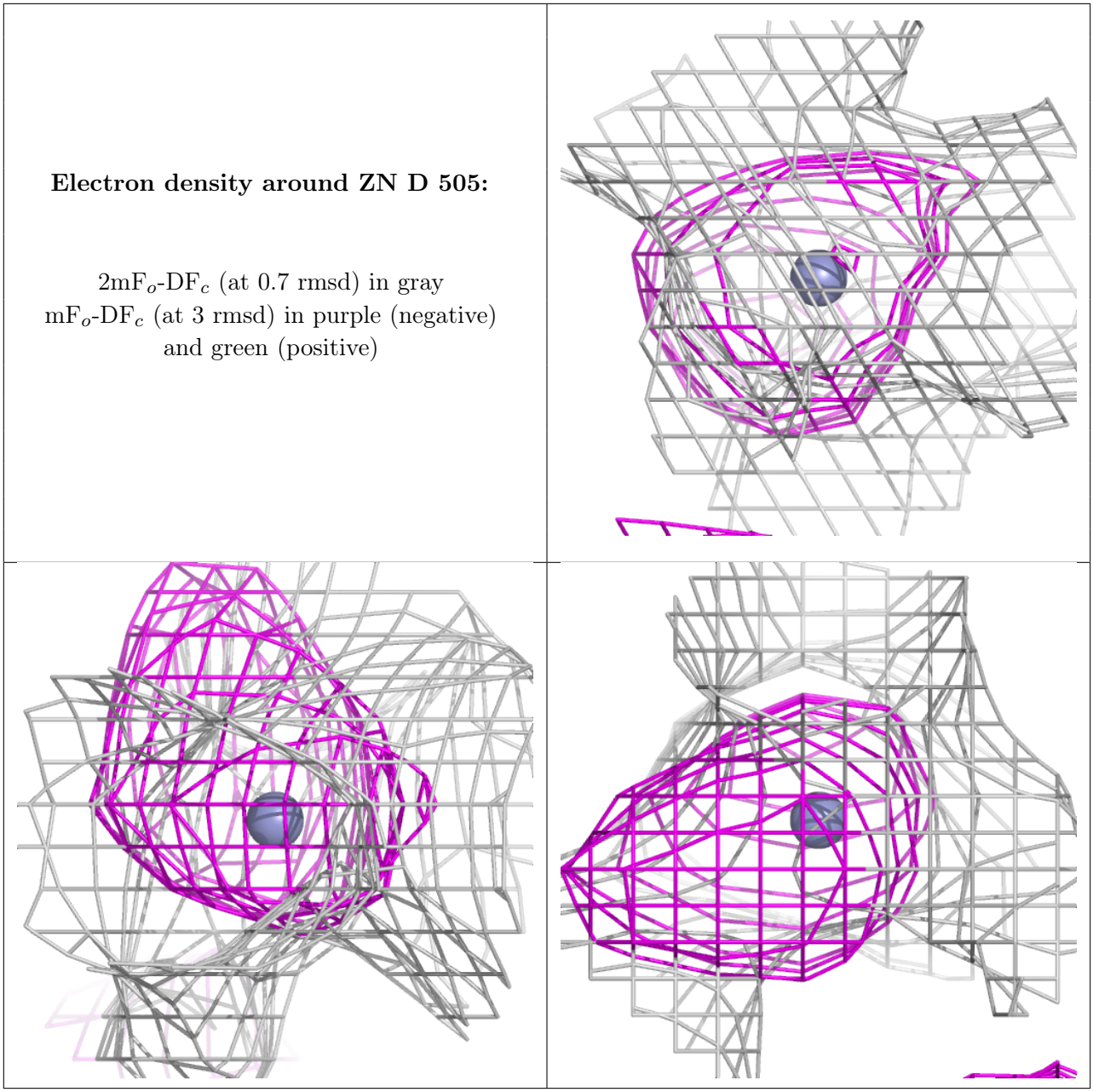
Electron density around CDP C 502:

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



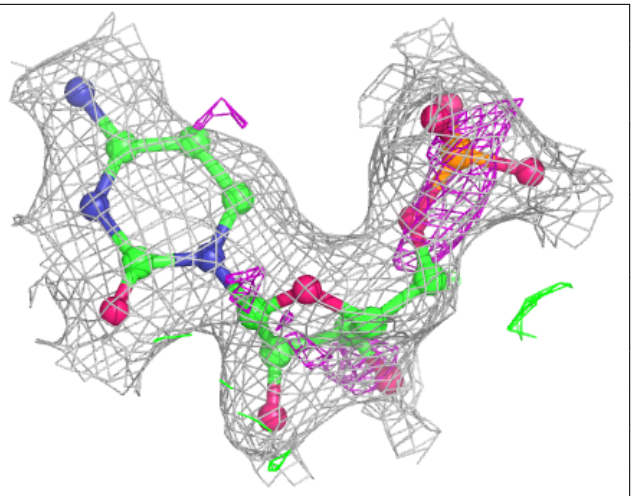
Electron density around ZN D 505:

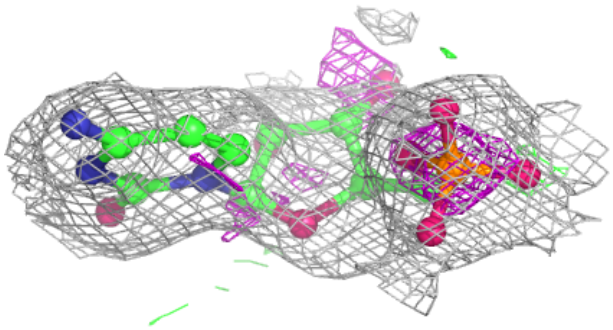
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

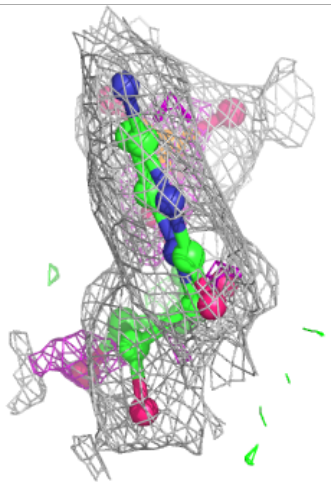


Electron density around C5P B 503:

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

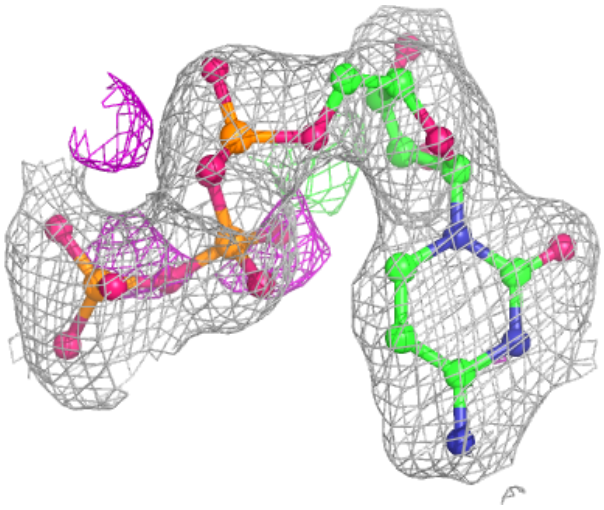


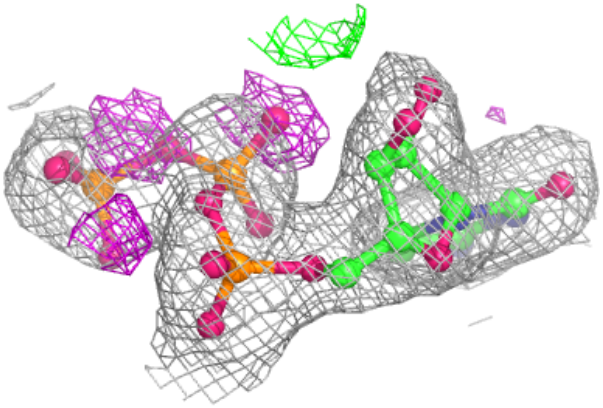


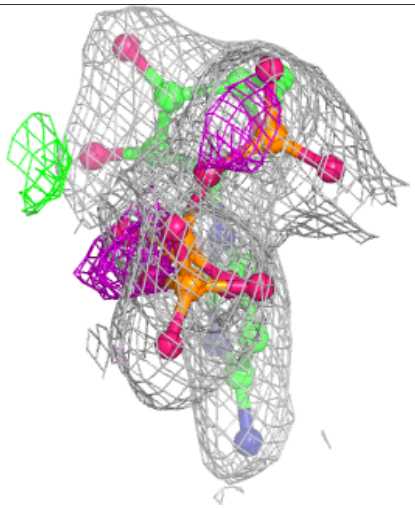


Electron density around CTP F 501:

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

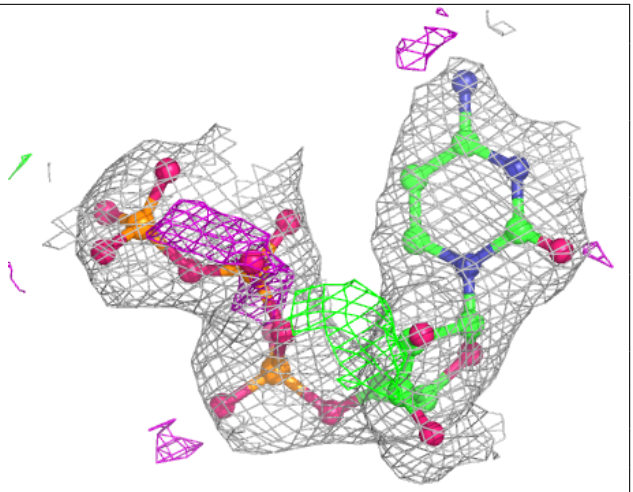


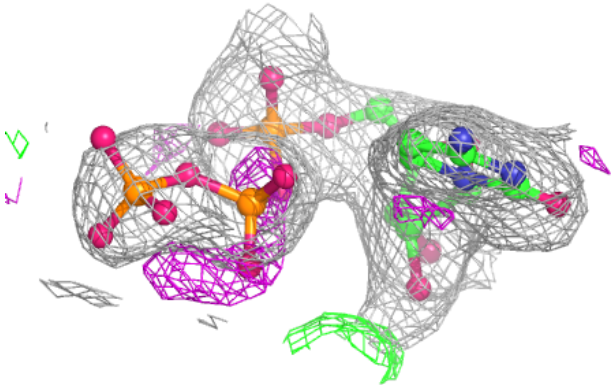


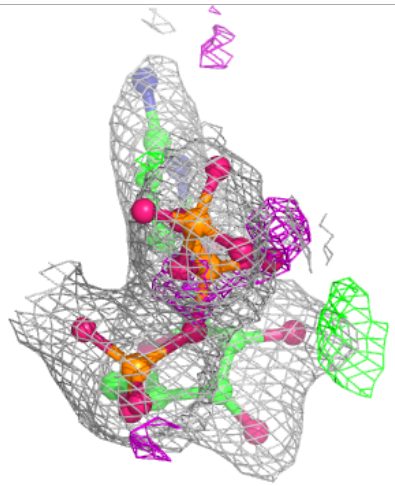


Electron density around CTP B 502:

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

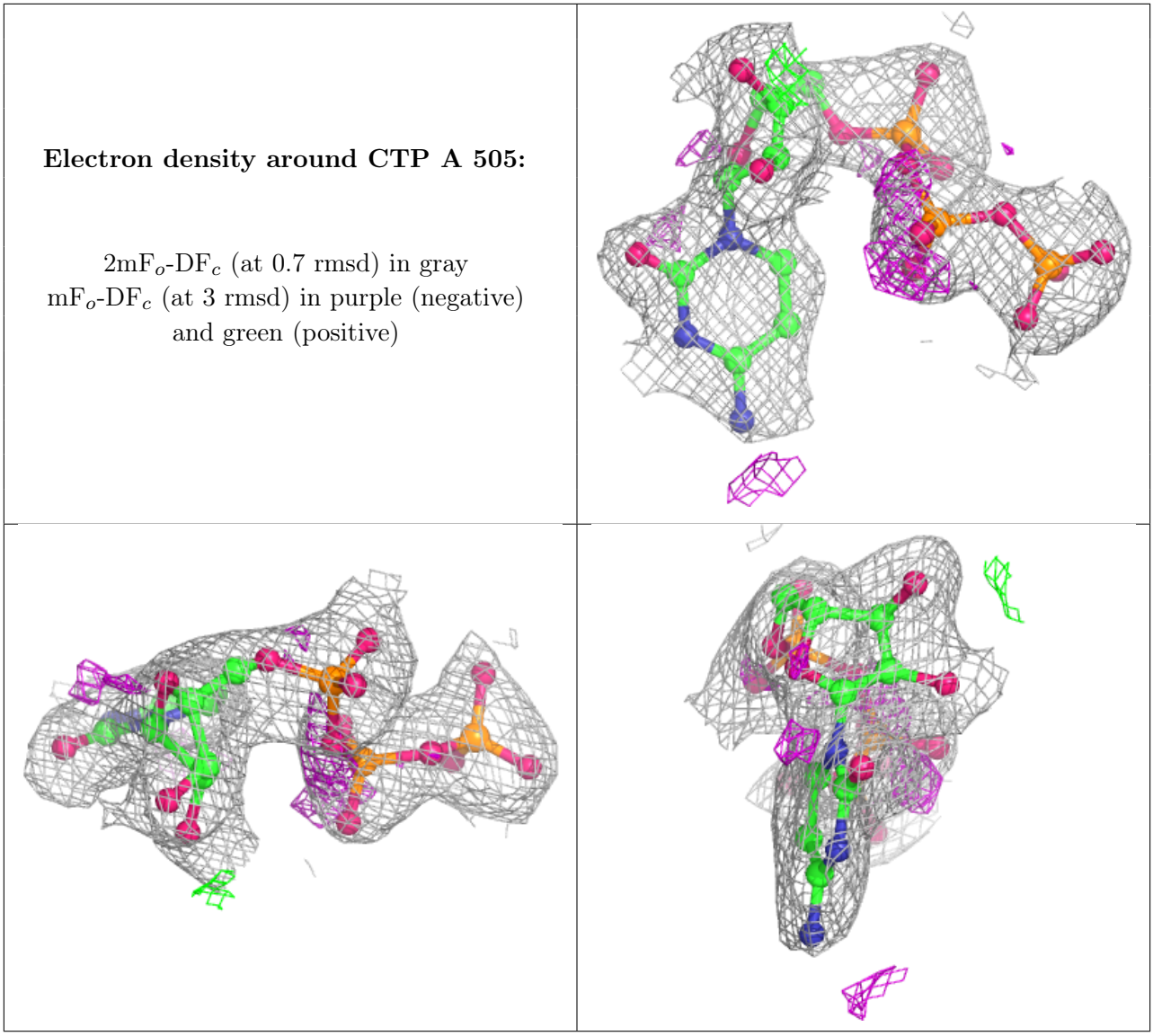






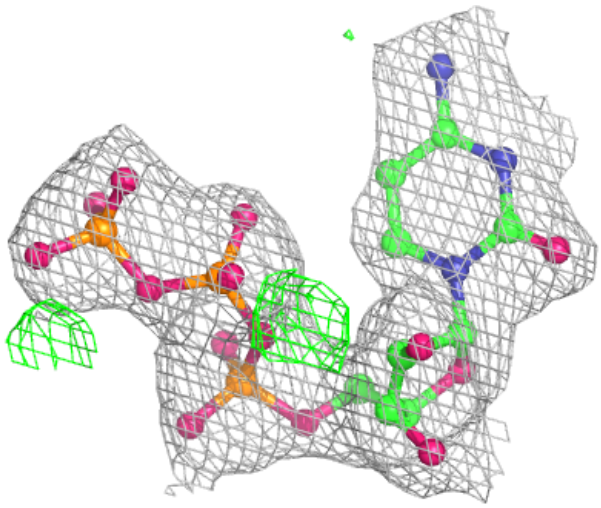
Electron density around CTP A 505:

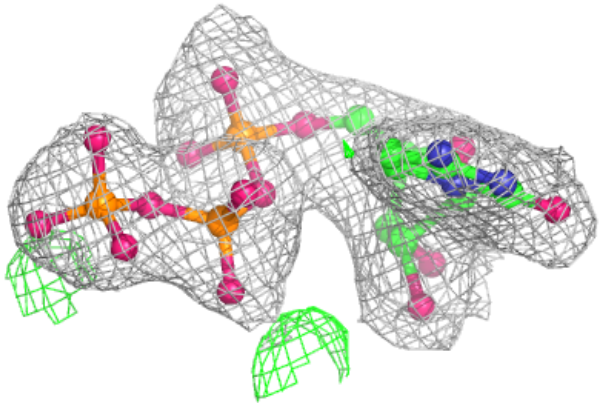
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

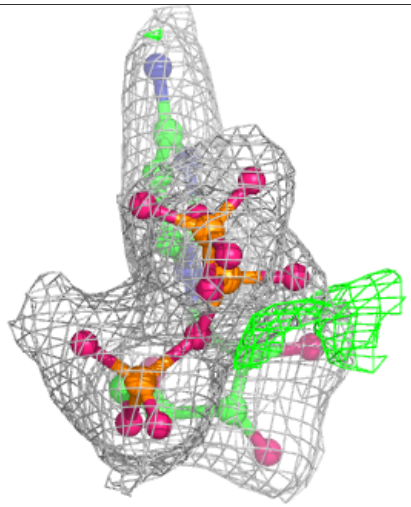


Electron density around CTP C 501:

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

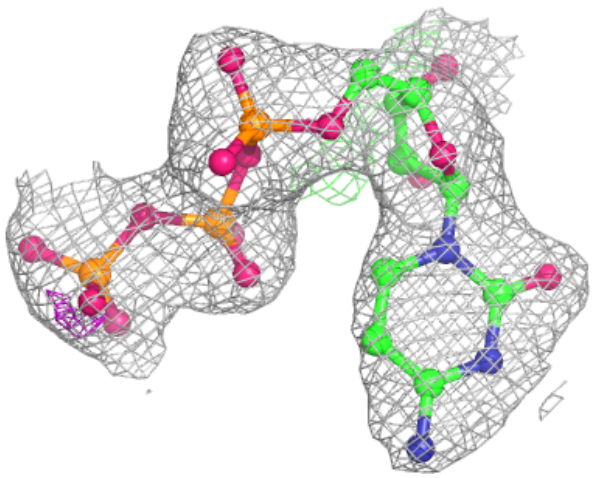


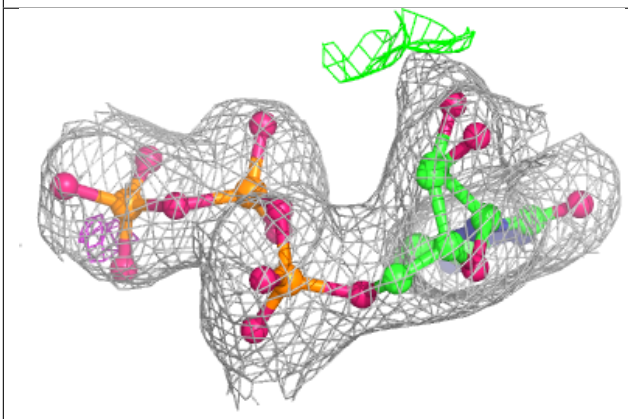


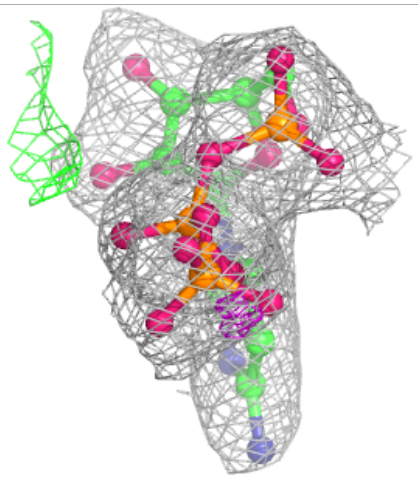


Electron density around CTP D 502:

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

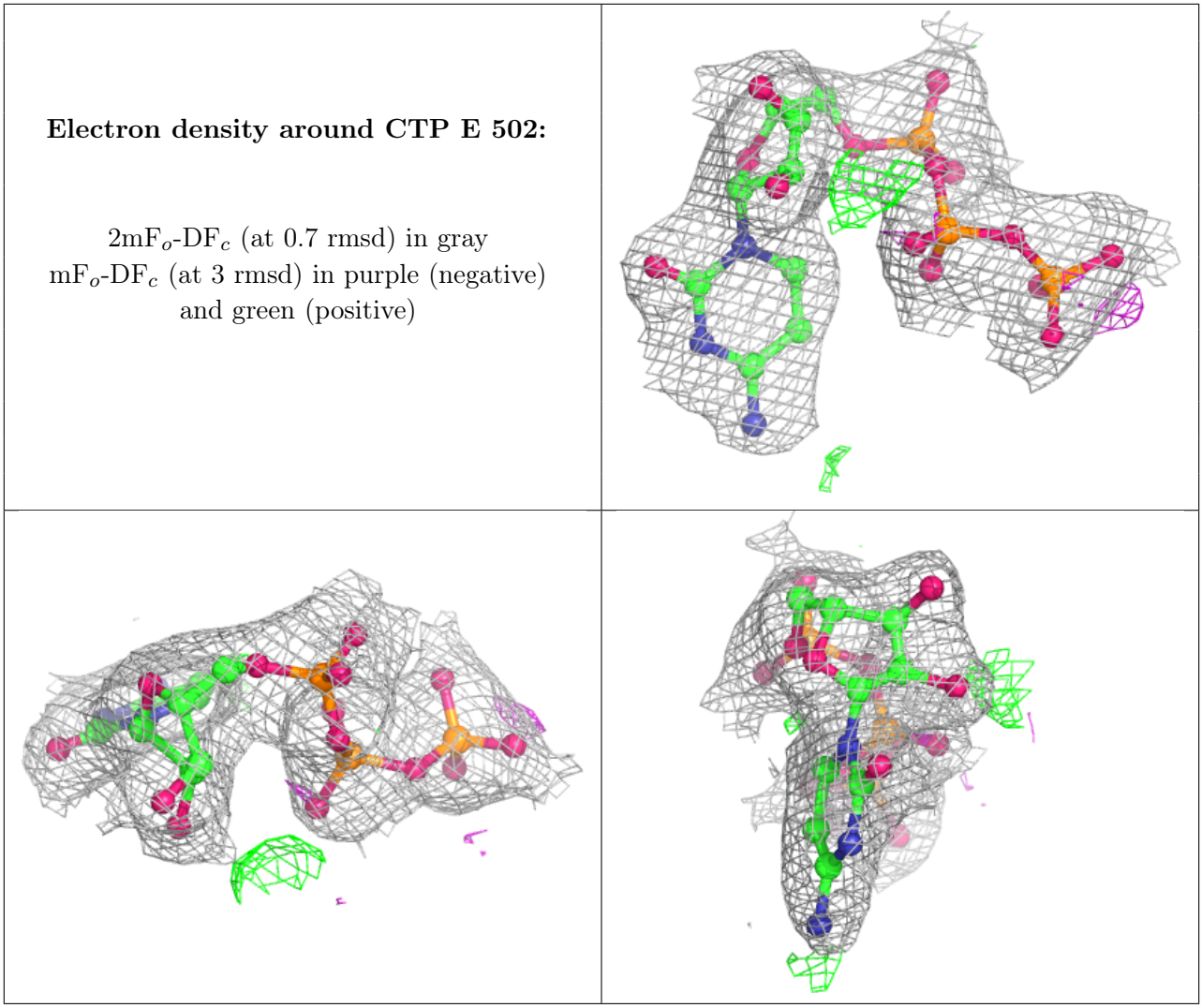






Electron density around CTP E 502:

$2mF_o$-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.