wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 03:32 am BST

PDB ID	:	6R2E
Title	:	Crystal structure of the human thymidylate synthase (hTS) interface variant Q62R
Authors	:	Pozzi, C.; Mangani, M.
Deposited on	:	2019-03-16
Resolution	:	2.55 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 following versions of software and data (see references (1)) were used in the production of this report:

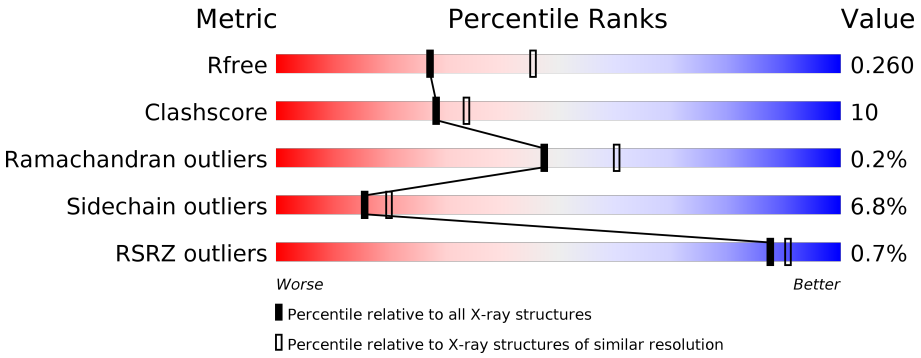
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.55 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.



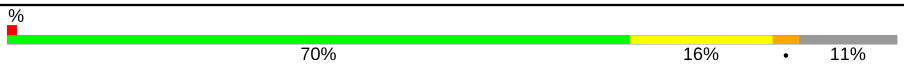
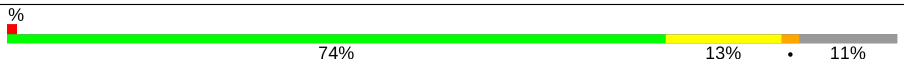
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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 on 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	325	% 69% 18% • 11%
1	B	325	% 68% 17% • 11%
1	C	325	74% 13% • 11%
1	D	325	69% 18% • 11%
1	F	325	% 71% 15% • 11%
1	H	325	71% 16% • 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	325	
2	G	325	

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
3	SO4	A	401	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 21357 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 Thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	0	0
			2312	1478	403	416	15			
1	C	288	Total	C	N	O	S	0	0	0
			2313	1478	401	419	15			
1	D	288	Total	C	N	O	S	0	1	0
			2329	1488	404	422	15			
1	B	289	Total	C	N	O	S	0	0	0
			2323	1486	406	416	15			
1	F	288	Total	C	N	O	S	0	0	0
			2323	1485	403	420	15			
1	H	288	Total	C	N	O	S	0	0	0
			2320	1484	403	418	15			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	initiating methionine	UNP P04818
A	-10	ARG	-	expression tag	UNP P04818
A	-9	GLY	-	expression tag	UNP P04818
A	-8	SER	-	expression tag	UNP P04818
A	-7	HIS	-	expression tag	UNP P04818
A	-6	HIS	-	expression tag	UNP P04818
A	-5	HIS	-	expression tag	UNP P04818
A	-4	HIS	-	expression tag	UNP P04818
A	-3	HIS	-	expression tag	UNP P04818
A	-2	HIS	-	expression tag	UNP P04818
A	-1	GLY	-	expression tag	UNP P04818
A	0	SER	-	expression tag	UNP P04818
A	62	ARG	GLN	engineered mutation	UNP P04818
C	-11	MET	-	initiating methionine	UNP P04818
C	-10	ARG	-	expression tag	UNP P04818
C	-9	GLY	-	expression tag	UNP P04818
C	-8	SER	-	expression tag	UNP P04818

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	HIS	-	expression tag	UNP P04818
C	-6	HIS	-	expression tag	UNP P04818
C	-5	HIS	-	expression tag	UNP P04818
C	-4	HIS	-	expression tag	UNP P04818
C	-3	HIS	-	expression tag	UNP P04818
C	-2	HIS	-	expression tag	UNP P04818
C	-1	GLY	-	expression tag	UNP P04818
C	0	SER	-	expression tag	UNP P04818
C	62	ARG	GLN	engineered mutation	UNP P04818
D	-11	MET	-	initiating methionine	UNP P04818
D	-10	ARG	-	expression tag	UNP P04818
D	-9	GLY	-	expression tag	UNP P04818
D	-8	SER	-	expression tag	UNP P04818
D	-7	HIS	-	expression tag	UNP P04818
D	-6	HIS	-	expression tag	UNP P04818
D	-5	HIS	-	expression tag	UNP P04818
D	-4	HIS	-	expression tag	UNP P04818
D	-3	HIS	-	expression tag	UNP P04818
D	-2	HIS	-	expression tag	UNP P04818
D	-1	GLY	-	expression tag	UNP P04818
D	0	SER	-	expression tag	UNP P04818
D	62	ARG	GLN	engineered mutation	UNP P04818
B	-11	MET	-	initiating methionine	UNP P04818
B	-10	ARG	-	expression tag	UNP P04818
B	-9	GLY	-	expression tag	UNP P04818
B	-8	SER	-	expression tag	UNP P04818
B	-7	HIS	-	expression tag	UNP P04818
B	-6	HIS	-	expression tag	UNP P04818
B	-5	HIS	-	expression tag	UNP P04818
B	-4	HIS	-	expression tag	UNP P04818
B	-3	HIS	-	expression tag	UNP P04818
B	-2	HIS	-	expression tag	UNP P04818
B	-1	GLY	-	expression tag	UNP P04818
B	0	SER	-	expression tag	UNP P04818
B	62	ARG	GLN	engineered mutation	UNP P04818
F	-11	MET	-	initiating methionine	UNP P04818
F	-10	ARG	-	expression tag	UNP P04818
F	-9	GLY	-	expression tag	UNP P04818
F	-8	SER	-	expression tag	UNP P04818
F	-7	HIS	-	expression tag	UNP P04818
F	-6	HIS	-	expression tag	UNP P04818
F	-5	HIS	-	expression tag	UNP P04818

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-4	HIS	-	expression tag	UNP P04818
F	-3	HIS	-	expression tag	UNP P04818
F	-2	HIS	-	expression tag	UNP P04818
F	-1	GLY	-	expression tag	UNP P04818
F	0	SER	-	expression tag	UNP P04818
F	62	ARG	GLN	engineered mutation	UNP P04818
H	-11	MET	-	initiating methionine	UNP P04818
H	-10	ARG	-	expression tag	UNP P04818
H	-9	GLY	-	expression tag	UNP P04818
H	-8	SER	-	expression tag	UNP P04818
H	-7	HIS	-	expression tag	UNP P04818
H	-6	HIS	-	expression tag	UNP P04818
H	-5	HIS	-	expression tag	UNP P04818
H	-4	HIS	-	expression tag	UNP P04818
H	-3	HIS	-	expression tag	UNP P04818
H	-2	HIS	-	expression tag	UNP P04818
H	-1	GLY	-	expression tag	UNP P04818
H	0	SER	-	expression tag	UNP P04818
H	62	ARG	GLN	engineered mutation	UNP P04818

- Molecule 2 is a protein called Thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	288	Total	C	N	O	S	0	0	0
			2321	1486	403	417	15			
2	G	288	Total	C	N	O	S	0	0	0
			2309	1478	399	417	15			

There are 26 discrepancies between the modelled and reference sequences:

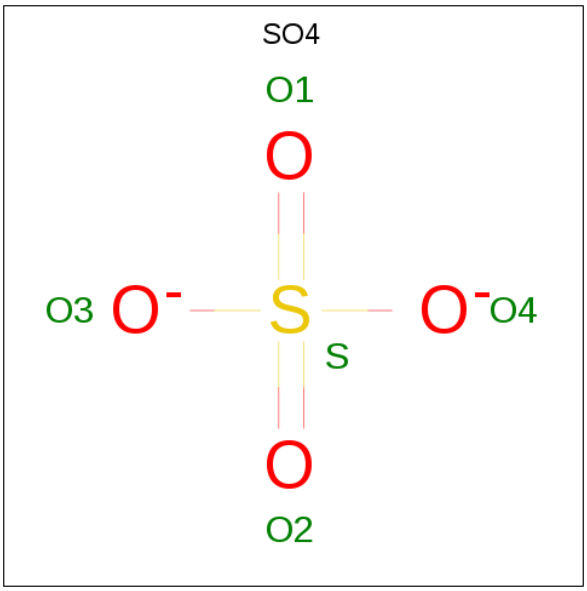
Chain	Residue	Modelled	Actual	Comment	Reference
E	-11	MET	-	initiating methionine	UNP P04818
E	-10	ARG	-	expression tag	UNP P04818
E	-9	GLY	-	expression tag	UNP P04818
E	-8	SER	-	expression tag	UNP P04818
E	-7	HIS	-	expression tag	UNP P04818
E	-6	HIS	-	expression tag	UNP P04818
E	-5	HIS	-	expression tag	UNP P04818
E	-4	HIS	-	expression tag	UNP P04818
E	-3	HIS	-	expression tag	UNP P04818
E	-2	HIS	-	expression tag	UNP P04818
E	-1	GLY	-	expression tag	UNP P04818

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	SER	-	expression tag	UNP P04818
E	62	ARG	GLN	engineered mutation	UNP P04818
G	-11	MET	-	initiating methionine	UNP P04818
G	-10	ARG	-	expression tag	UNP P04818
G	-9	GLY	-	expression tag	UNP P04818
G	-8	SER	-	expression tag	UNP P04818
G	-7	HIS	-	expression tag	UNP P04818
G	-6	HIS	-	expression tag	UNP P04818
G	-5	HIS	-	expression tag	UNP P04818
G	-4	HIS	-	expression tag	UNP P04818
G	-3	HIS	-	expression tag	UNP P04818
G	-2	HIS	-	expression tag	UNP P04818
G	-1	GLY	-	expression tag	UNP P04818
G	0	SER	-	expression tag	UNP P04818
G	62	ARG	GLN	engineered mutation	UNP P04818

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



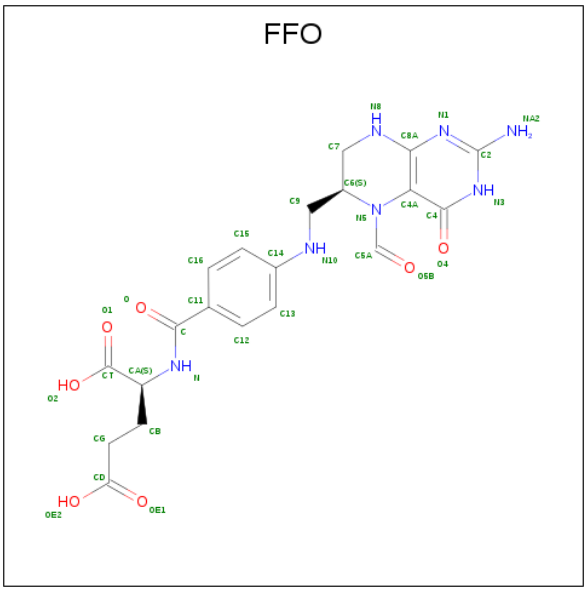
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	F	1	Total O S 5 4 1	0	0
3	F	1	Total O S 5 4 1	0	0
3	H	1	Total O S 5 4 1	0	0
3	H	1	Total O S 5 4 1	0	0
3	E	1	Total O S 5 4 1	0	0
3	E	1	Total O S 5 4 1	0	0
3	G	1	Total O S 5 4 1	0	0
3	G	1	Total O S 5 4 1	0	0

- Molecule 4 is N-[4-({[(6S)-2-amino-5-formyl-4-oxo-3,4,5,6,7,8-hexahydropteridin-6-yl]methyl }amino)benzoyl]-L-glutamic acid (three-letter code: FFO) (formula: $C_{20}H_{23}N_7O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 34 20 7 7	0	0
4	C	1	Total C N O 34 20 7 7	0	0
4	D	1	Total C N O 34 20 7 7	0	0
4	B	1	Total C N O 34 20 7 7	0	0
4	F	1	Total C N O 34 20 7 7	0	0
4	H	1	Total C N O 34 20 7 7	0	0
4	E	1	Total C N O 34 20 7 7	0	0
4	G	1	Total C N O 34 20 7 7	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

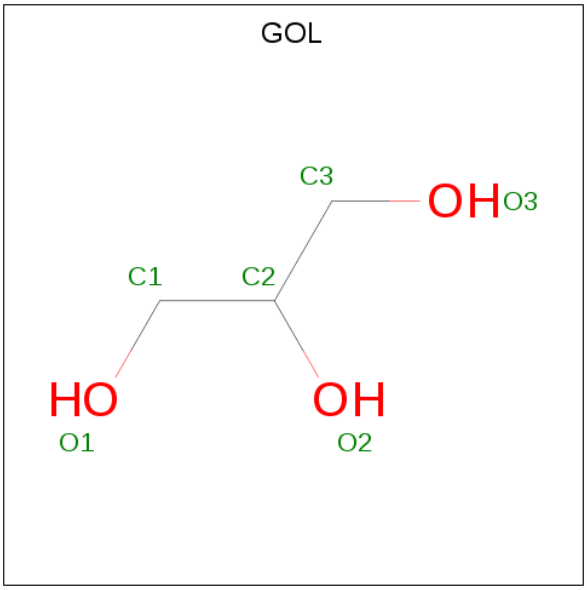
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	G	1	Total Cl 1 1	0	0
5	D	2	Total Cl 2 2	0	0
5	E	1	Total Cl 1 1	0	0
5	B	1	Total Cl 1 1	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total C O 6 3 3	0	0
6	H	1	Total C O 6 3 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	334	Total O 334 334	0	0
7	C	318	Total O 318 318	0	0
7	D	319	Total O 319 319	0	0
7	B	302	Total O 302 302	0	0
7	F	330	Total O 330 330	0	0

Continued on next page...

Continued from previous page...

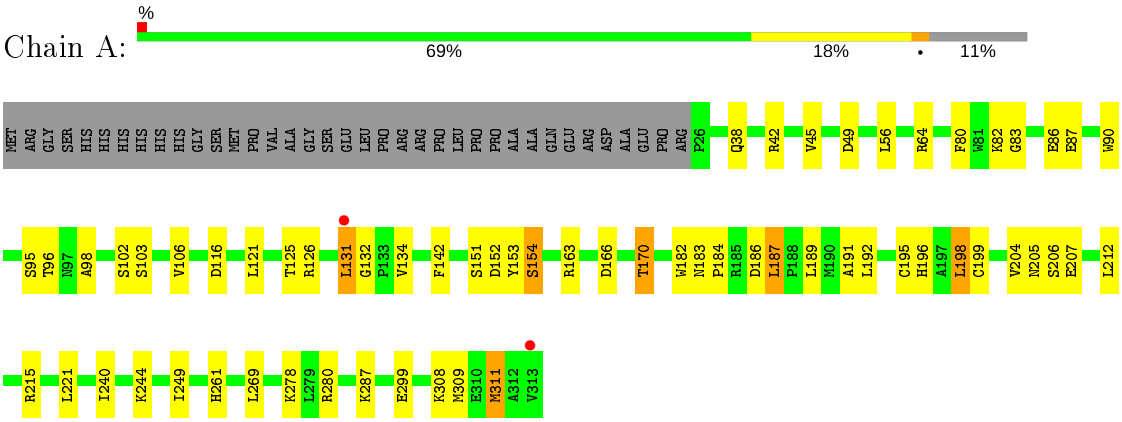
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	H	277	Total O 277 277	0	0
7	E	311	Total O 311 311	0	0
7	G	239	Total O 239 239	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

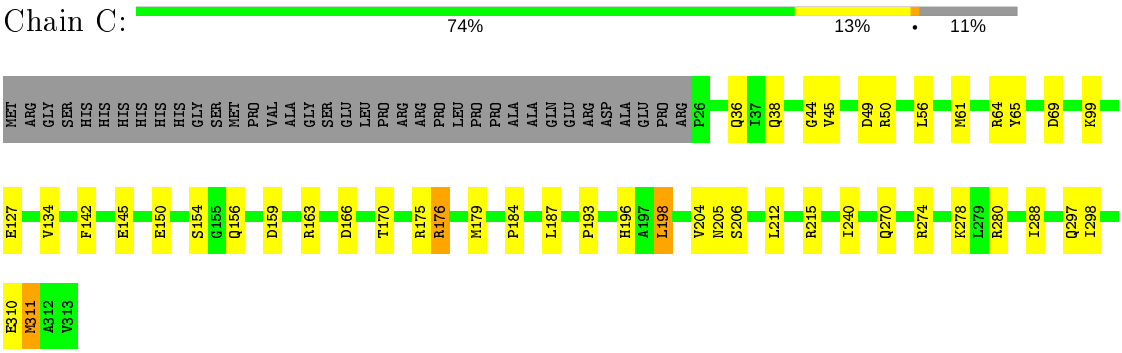
- Molecule 1: Thymidylate synthase

Chain A:



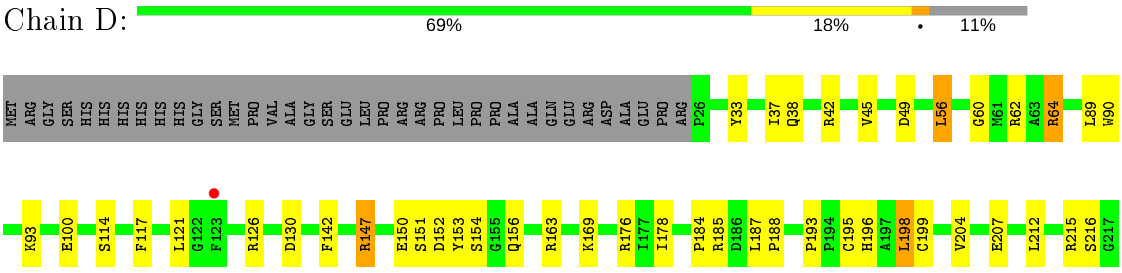
- Molecule 1: Thymidylate synthase

Chain C:



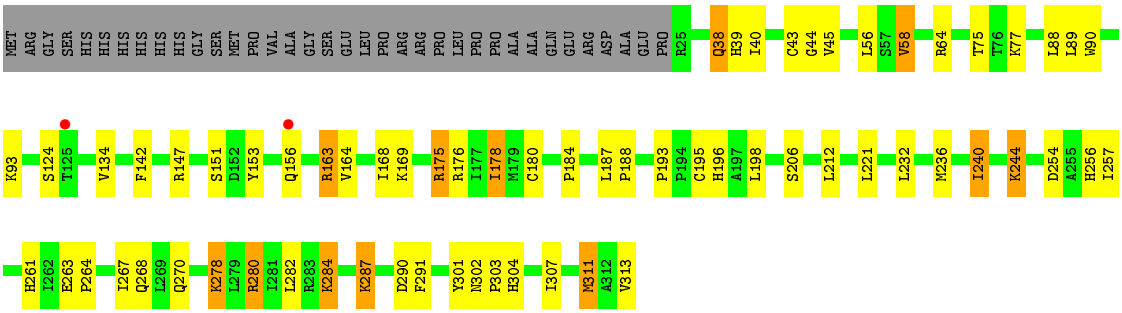
- Molecule 1: Thymidylate synthase

Chain D:



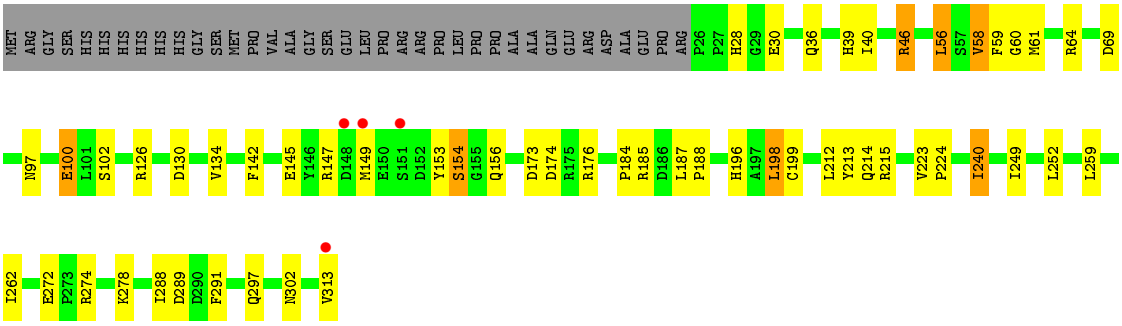
• Molecule 1: Thymidylate synthase

Chain B:



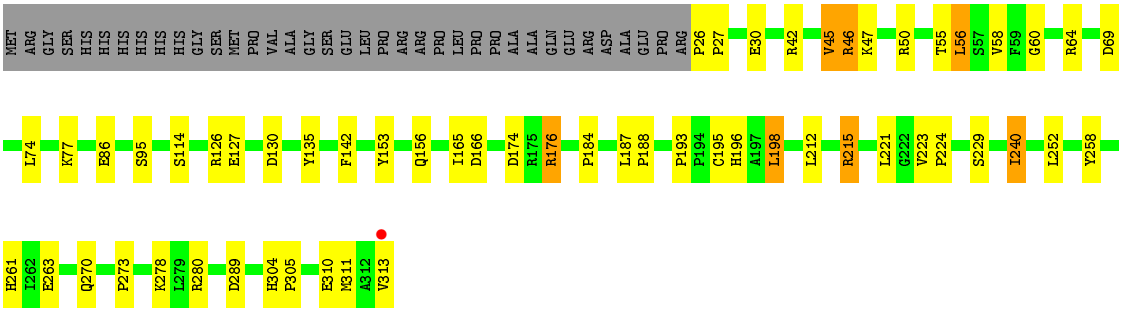
• Molecule 1: Thymidylate synthase

Chain F:



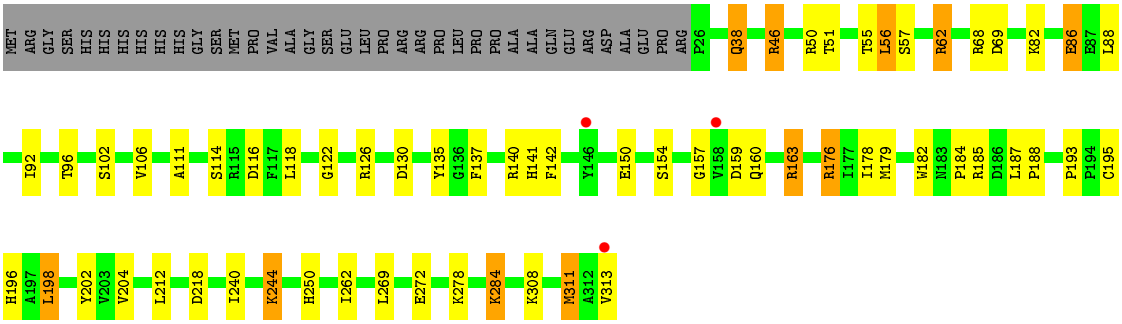
• Molecule 1: Thymidylate synthase

Chain H:

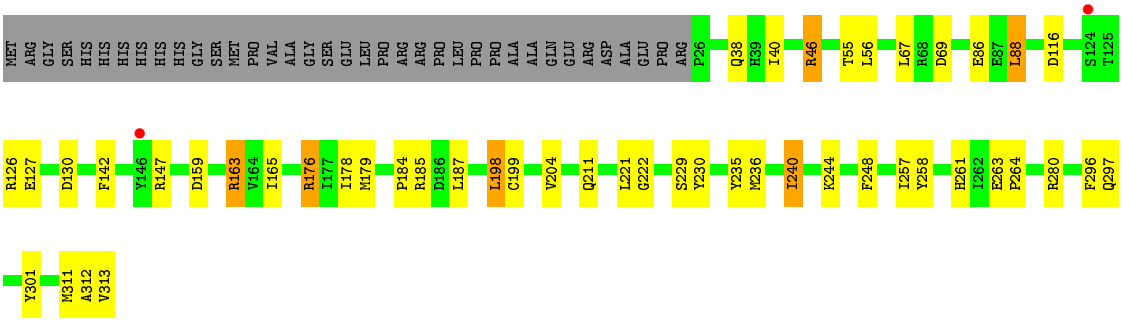


• Molecule 2: Thymidylate synthase

Chain E:



• Molecule 2: Thymidylate synthase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, α, β, γ	139.94Å 167.07Å 189.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	94.99 – 2.55 94.98 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.8 (94.99-2.55) 99.8 (94.98-2.55)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.40 (at 2.55Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.196 , 0.258 0.200 , 0.260	Depositor DCC
R_{free} test set	7182 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	29.1	Xtriage
Anisotropy	0.194	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 42.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	21357	wwPDB-VP
Average B, all atoms (Å^2)	44.0	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 59.98 % 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 1.6289e-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: GOL, SCH, CL, CME, FFO, SO4

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	0.62	0/2350	1.05	0/3178
1	B	0.63	0/2361	1.06	0/3192
1	C	0.66	0/2351	1.02	0/3180
1	D	0.62	0/2370	1.04	0/3203
1	F	0.67	0/2361	1.03	0/3191
1	H	0.62	0/2358	1.05	0/3187
2	E	0.63	0/2361	1.02	0/3189
2	G	0.59	0/2349	1.00	0/3177
All	All	0.63	0/18861	1.04	0/25497

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	2312	0	2263	47	0
1	B	2323	0	2281	64	0
1	C	2313	0	2256	35	0
1	D	2329	0	2286	53	0
1	F	2323	0	2276	50	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	2320	0	2275	56	0
2	E	2321	0	2285	54	0
2	G	2309	0	2252	45	0
3	A	10	0	0	2	0
3	B	10	0	0	1	0
3	C	15	0	0	0	0
3	D	10	0	0	0	0
3	E	10	0	0	0	0
3	F	10	0	0	0	0
3	G	10	0	0	0	0
3	H	10	0	0	1	0
4	A	34	0	21	4	0
4	B	34	0	21	7	0
4	C	34	0	21	3	0
4	D	34	0	21	2	0
4	E	34	0	21	6	0
4	F	34	0	21	0	0
4	G	34	0	21	2	0
4	H	34	0	21	2	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	2	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
6	D	6	0	8	0	0
6	H	6	0	8	0	0
7	A	334	0	0	5	0
7	B	302	0	0	9	0
7	C	318	0	0	6	0
7	D	319	0	0	6	0
7	E	311	0	0	8	0
7	F	330	0	0	4	0
7	G	239	0	0	8	0
7	H	277	0	0	12	0
All	All	21357	0	18358	367	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 10.

The worst 5 of 367 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)
1:H:26:PRO:HB2	1:H:27:PRO:CD	1.50	1.31
1:H:215:ARG:NH1	2:G:176:ARG:HG3	1.62	1.14
1:H:26:PRO:CB	1:H:27:PRO:CD	2.26	1.12
1:H:26:PRO:CB	1:H:27:PRO:HD3	1.82	1.10
1:A:215:ARG:NH1	1:B:176:ARG:HG3	1.69	1.08

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	284/325 (87%)	271 (95%)	12 (4%)	1 (0%)	34	46
1	B	285/325 (88%)	271 (95%)	13 (5%)	1 (0%)	34	46
1	C	284/325 (87%)	274 (96%)	9 (3%)	1 (0%)	34	46
1	D	285/325 (88%)	270 (95%)	15 (5%)	0	100	100
1	F	284/325 (87%)	276 (97%)	7 (2%)	1 (0%)	34	46
1	H	284/325 (87%)	268 (94%)	16 (6%)	0	100	100
2	E	284/325 (87%)	270 (95%)	14 (5%)	0	100	100
2	G	284/325 (87%)	270 (95%)	14 (5%)	0	100	100
All	All	2274/2600 (88%)	2170 (95%)	100 (4%)	4 (0%)	47	60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	134	VAL
1	C	134	VAL
1	B	134	VAL
1	A	134	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	242/279 (87%)	224 (93%)	18 (7%)	13	18
1	B	243/279 (87%)	225 (93%)	18 (7%)	13	18
1	C	242/279 (87%)	224 (93%)	18 (7%)	13	18
1	D	246/279 (88%)	232 (94%)	14 (6%)	20	27
1	F	244/279 (88%)	230 (94%)	14 (6%)	20	27
1	H	243/279 (87%)	226 (93%)	17 (7%)	15	19
2	E	244/279 (88%)	226 (93%)	18 (7%)	13	18
2	G	241/279 (86%)	226 (94%)	15 (6%)	18	24
All	All	1945/2232 (87%)	1813 (93%)	132 (7%)	16	20

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	240	ILE
1	F	102	SER
2	G	88	LEU
1	B	278	LYS
1	B	313	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	302	ASN
1	B	171	ASN
1	H	171	ASN
1	D	171	ASN
1	F	39	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

16 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\#\|Z\|>2$
1	CME	F	43	1	8,9,10	0.52	0	5,9,11	0.27	0
1	CME	C	43	1	8,9,10	0.71	0	5,9,11	0.69	0
1	CME	B	43	1	8,9,10	0.83	0	5,9,11	1.06	1 (20%)
1	CME	A	195	1	8,9,10	0.59	0	5,9,11	1.17	0
1	CME	C	195	1	8,9,10	0.51	0	5,9,11	0.66	0
2	CME	E	195	2	8,9,10	0.73	0	5,9,11	1.80	1 (20%)
2	CME	G	195	2	8,9,10	0.58	0	5,9,11	0.98	0
1	CME	D	195	1	8,9,10	0.40	0	5,9,11	1.70	1 (20%)
1	CME	H	195	1	8,9,10	0.58	0	5,9,11	1.06	0
2	SCH	G	43	2	6,7,8	0.72	0	3,7,9	1.84	1 (33%)
1	CME	B	195	1	8,9,10	0.64	0	5,9,11	1.02	1 (20%)
1	CME	A	43	1	8,9,10	0.48	0	5,9,11	1.39	1 (20%)
1	CME	D	43	1	8,9,10	0.83	0	5,9,11	1.19	0
1	CME	F	195	1	8,9,10	0.46	0	5,9,11	0.95	0
1	CME	H	43	1	8,9,10	0.85	0	5,9,11	1.69	1 (20%)
2	SCH	E	43	2	6,7,8	0.52	0	3,7,9	1.02	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
1	CME	F	43	1	-	1/5/8/10	-
1	CME	C	43	1	-	3/5/8/10	-
1	CME	B	43	1	-	1/5/8/10	-
1	CME	A	195	1	-	2/5/8/10	-
1	CME	C	195	1	-	2/5/8/10	-
2	CME	E	195	2	-	2/5/8/10	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CME	G	195	2	-	3/5/8/10	-
1	CME	D	195	1	-	1/5/8/10	-
1	CME	H	195	1	-	1/5/8/10	-
2	SCH	G	43	2	-	0/2/6/8	-
1	CME	B	195	1	-	2/5/8/10	-
1	CME	A	43	1	-	2/5/8/10	-
1	CME	D	43	1	-	3/5/8/10	-
1	CME	F	195	1	-	2/5/8/10	-
1	CME	H	43	1	-	1/5/8/10	-
2	SCH	E	43	2	-	0/2/6/8	-

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	195	CME	OH-CZ-CE	3.63	125.16	110.83
2	E	195	CME	OH-CZ-CE	3.43	124.37	110.83
2	G	43	SCH	CE-SD-SG	3.12	113.36	102.58
1	H	43	CME	CZ-CE-SD	2.75	122.93	113.37
1	A	43	CME	CZ-CE-SD	-2.70	104.00	113.37

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	43	CME	CE-SD-SG-CB
1	C	43	CME	CZ-CE-SD-SG
1	A	195	CME	CE-SD-SG-CB
1	C	195	CME	SD-CE-CZ-OH
2	E	195	CME	SD-CE-CZ-OH

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	43	CME	1	0
1	A	195	CME	1	0
2	E	195	CME	2	0
1	D	195	CME	2	0
1	H	195	CME	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	195	CME	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 35 ligands modelled in this entry, 8 are monoatomic - leaving 27 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	SO4	B	402	-	4,4,4	0.26	0	6,6,6	0.15	0
6	GOL	D	404	-	5,5,5	0.20	0	5,5,5	0.69	0
4	FFO	B	403	-	28,36,36	1.22	3 (10%)	28,50,50	1.55	6 (21%)
3	SO4	C	403	-	4,4,4	0.31	0	6,6,6	0.28	0
3	SO4	E	402	-	4,4,4	0.26	0	6,6,6	0.12	0
3	SO4	C	402	-	4,4,4	0.31	0	6,6,6	0.18	0
4	FFO	A	403	-	28,36,36	0.96	1 (3%)	28,50,50	1.81	7 (25%)
3	SO4	F	401	-	4,4,4	0.25	0	6,6,6	0.09	0
3	SO4	C	401	-	4,4,4	0.26	0	6,6,6	0.26	0
3	SO4	A	402	-	4,4,4	0.28	0	6,6,6	0.11	0
4	FFO	H	403	-	28,36,36	0.87	1 (3%)	28,50,50	1.55	5 (17%)
4	FFO	F	403	-	28,36,36	1.12	2 (7%)	28,50,50	1.39	5 (17%)
3	SO4	F	402	-	4,4,4	0.29	0	6,6,6	0.10	0
4	FFO	E	403	-	28,36,36	1.00	2 (7%)	28,50,50	1.60	5 (17%)
3	SO4	A	401	-	4,4,4	0.05	0	6,6,6	0.37	0
3	SO4	E	401	-	4,4,4	0.29	0	6,6,6	0.21	0
3	SO4	G	401	-	4,4,4	0.28	0	6,6,6	0.37	0
4	FFO	D	403	-	28,36,36	1.41	3 (10%)	28,50,50	1.61	6 (21%)
3	SO4	B	401	-	4,4,4	0.29	0	6,6,6	0.20	0
3	SO4	H	401	-	4,4,4	0.17	0	6,6,6	0.22	0
3	SO4	H	402	-	4,4,4	0.28	0	6,6,6	0.08	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	FFO	C	404	-	28,36,36	1.03	2 (7%)	28,50,50	1.68	5 (17%)
3	SO4	D	401	-	4,4,4	0.21	0	6,6,6	0.17	0
3	SO4	D	402	-	4,4,4	0.29	0	6,6,6	0.17	0
4	FFO	G	403	-	28,36,36	1.15	2 (7%)	28,50,50	1.49	6 (21%)
3	SO4	G	402	-	4,4,4	0.28	0	6,6,6	0.24	0
6	GOL	H	404	-	5,5,5	0.32	0	5,5,5	1.15	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
4	FFO	C	404	-	-	2/18/37/37	0/2/3/3
4	FFO	H	403	-	-	3/18/37/37	0/2/3/3
4	FFO	F	403	-	-	7/18/37/37	0/2/3/3
6	GOL	D	404	-	-	2/4/4/4	-
4	FFO	B	403	-	-	6/18/37/37	0/2/3/3
4	FFO	A	403	-	-	6/18/37/37	0/2/3/3
4	FFO	G	403	-	-	9/18/37/37	0/2/3/3
4	FFO	E	403	-	-	3/18/37/37	0/2/3/3
4	FFO	D	403	-	-	9/18/37/37	0/2/3/3
6	GOL	H	404	-	-	4/4/4/4	-

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	403	FFO	C4-N3	4.57	1.41	1.33
4	B	403	FFO	C4-N3	4.50	1.40	1.33
4	G	403	FFO	C4-N3	4.09	1.40	1.33
4	D	403	FFO	C5A-N5	3.95	1.40	1.35
4	E	403	FFO	C4-N3	3.38	1.38	1.33

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	403	FFO	C4-C4A-C8A	5.34	118.61	114.44
4	D	403	FFO	C4A-N5-C6	-4.75	110.88	119.31
4	E	403	FFO	C4-C4A-C8A	4.52	117.97	114.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	404	FFO	C4-C4A-C8A	4.52	117.97	114.44
4	A	403	FFO	C4A-N5-C6	-4.43	111.45	119.31

There are no chirality outliers.

5 of 51 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	D	404	GOL	O1-C1-C2-C3
4	B	403	FFO	O5B-C5A-N5-C4A
4	B	403	FFO	O5B-C5A-N5-C6
4	B	403	FFO	N5-C6-C9-N10
4	A	403	FFO	O5B-C5A-N5-C4A

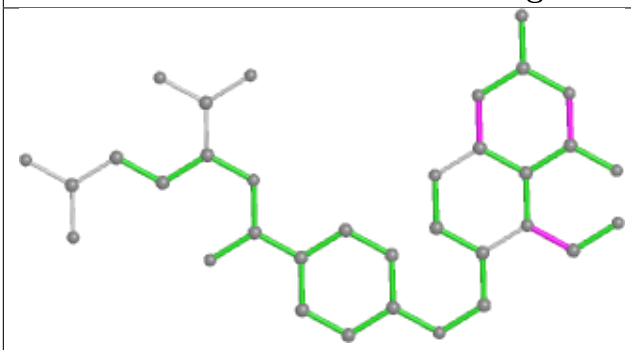
There are no ring outliers.

10 monomers are involved in 30 short contacts:

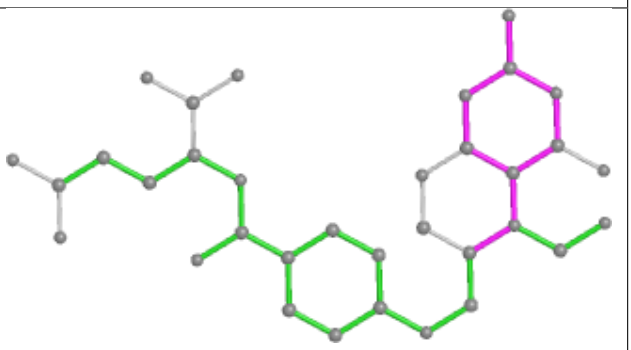
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	402	SO4	1	0
4	B	403	FFO	7	0
4	A	403	FFO	4	0
4	H	403	FFO	2	0
4	E	403	FFO	6	0
3	A	401	SO4	2	0
4	D	403	FFO	2	0
3	H	401	SO4	1	0
4	C	404	FFO	3	0
4	G	403	FFO	2	0

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.

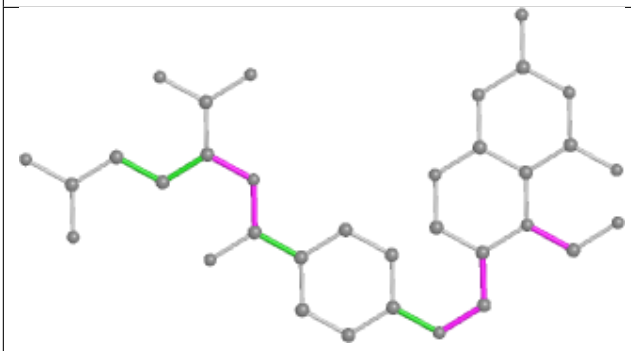
Ligand FFO B 403



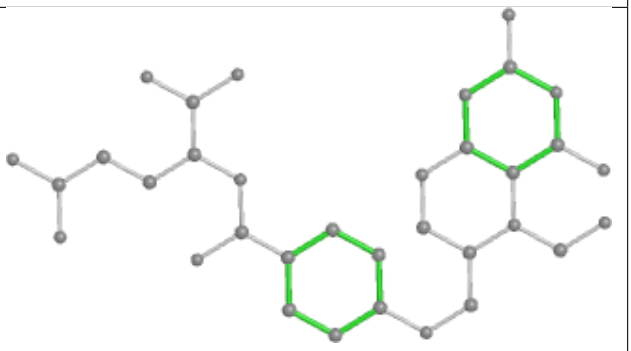
Bond lengths



Bond angles

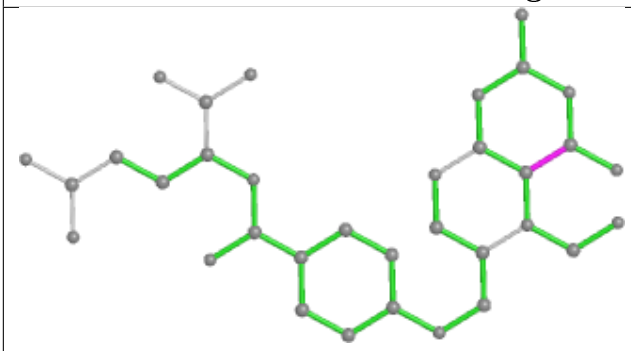


Torsions

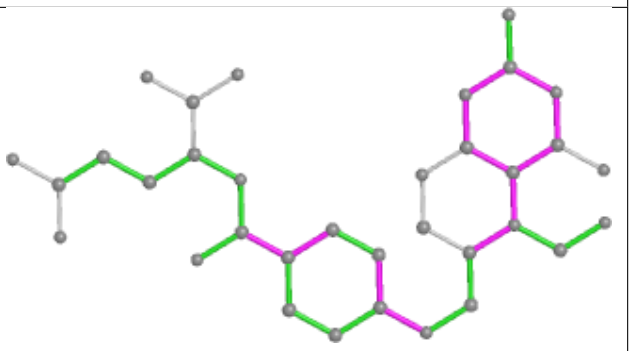


Rings

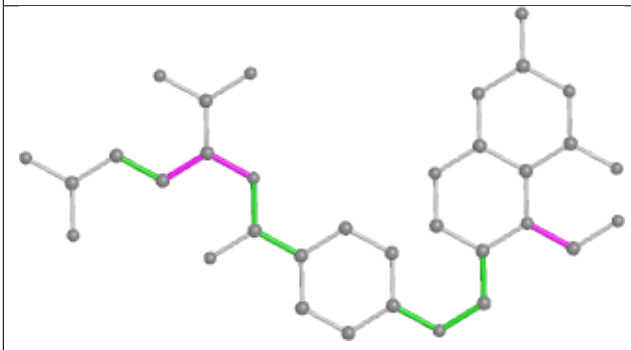
Ligand FFO A 403



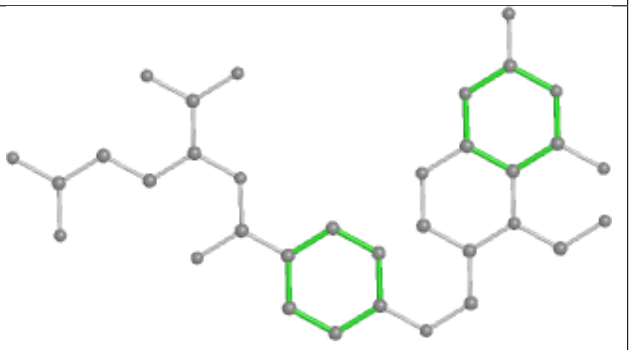
Bond lengths



Bond angles

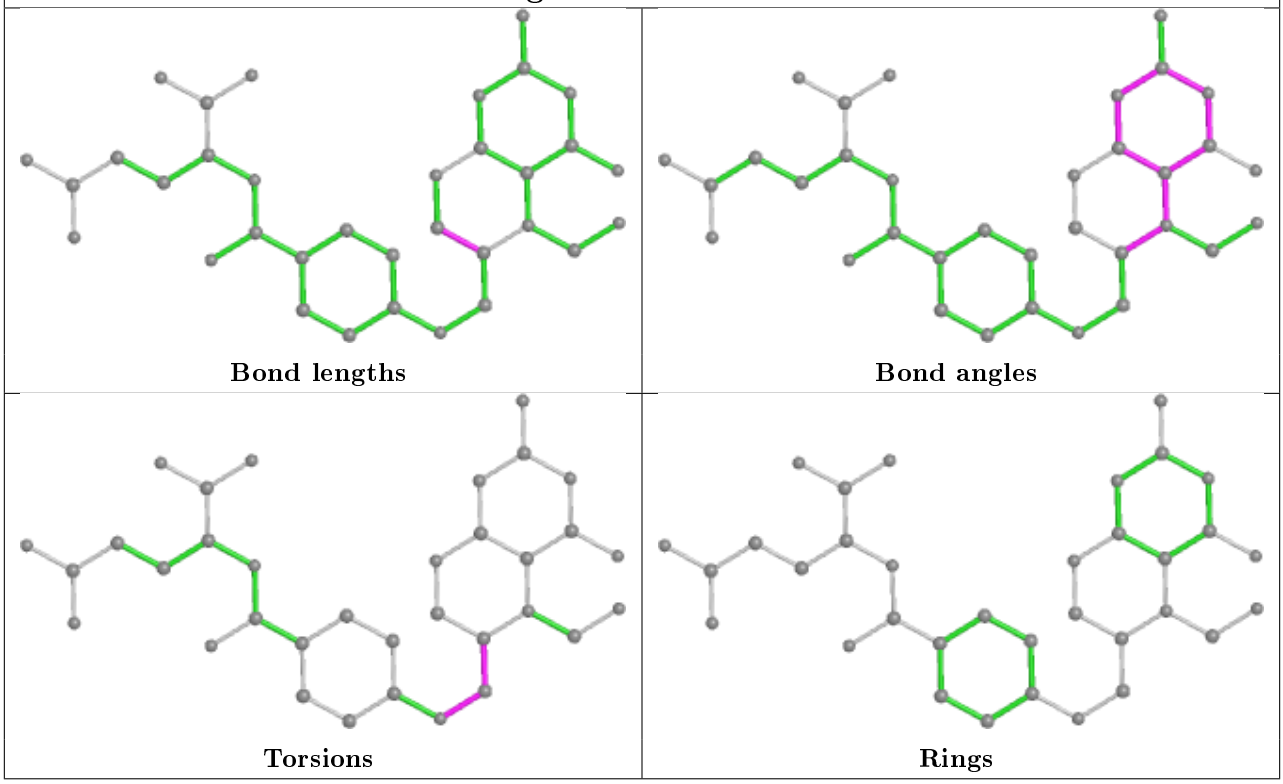


Torsions

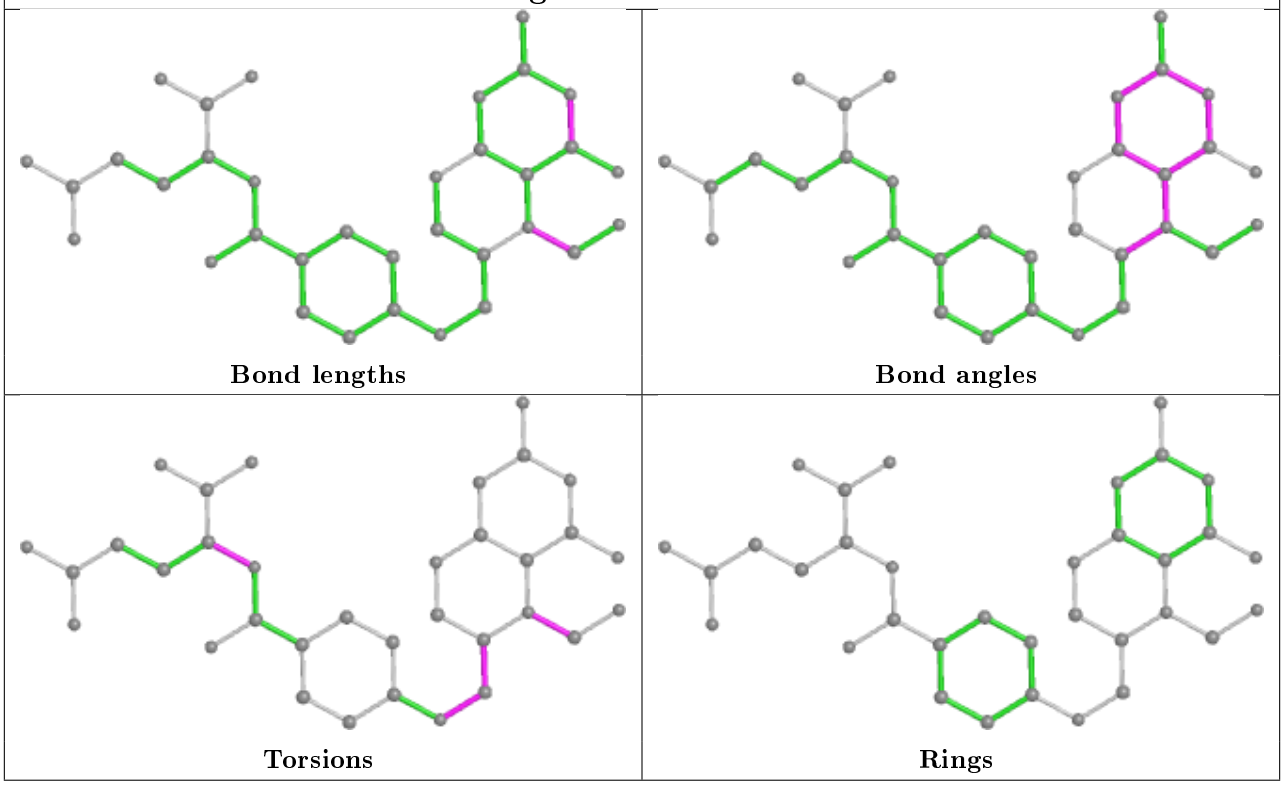


Rings

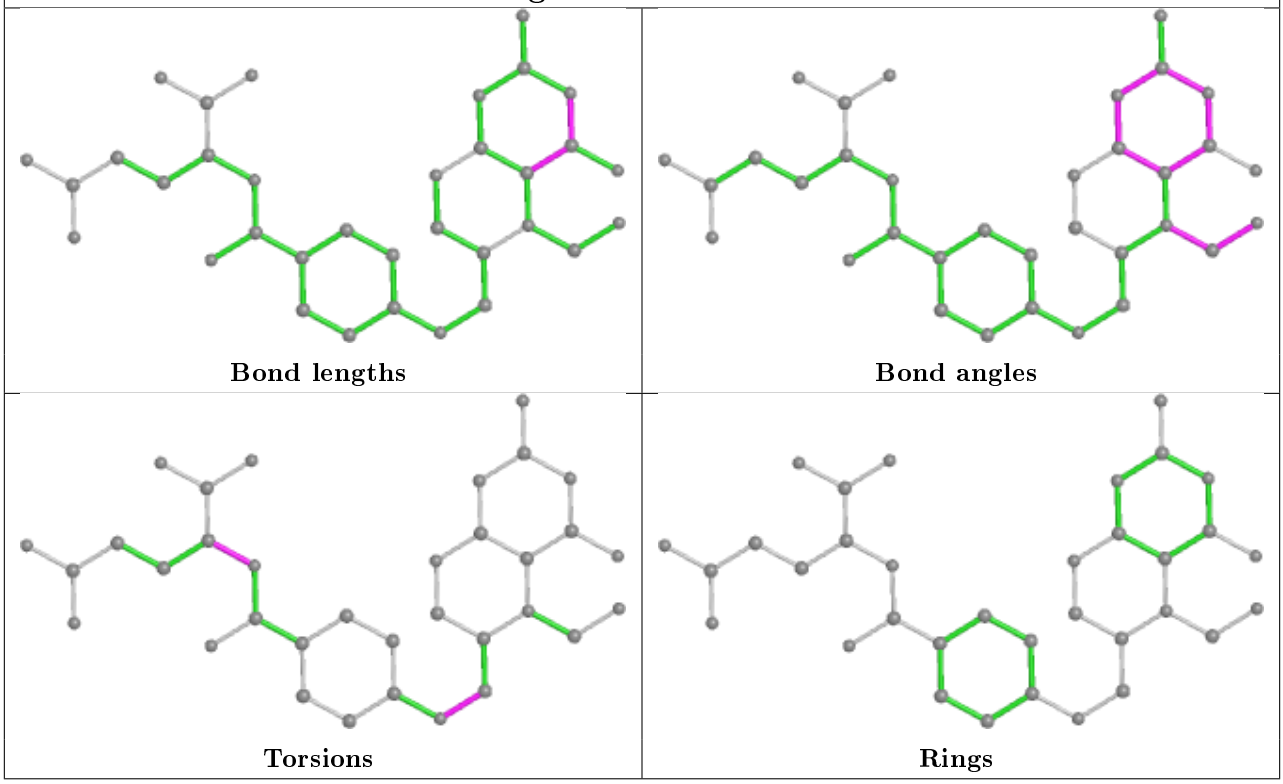
Ligand FFO H 403



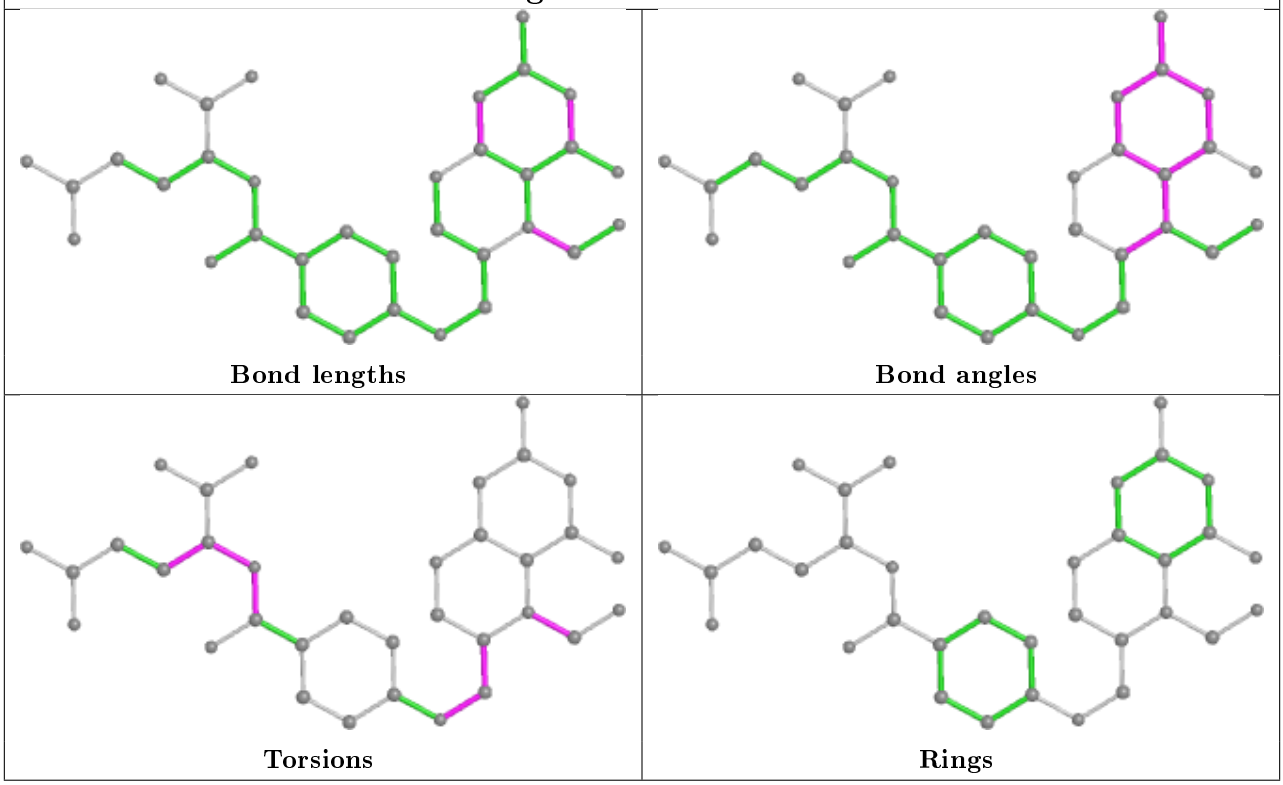
Ligand FFO F 403



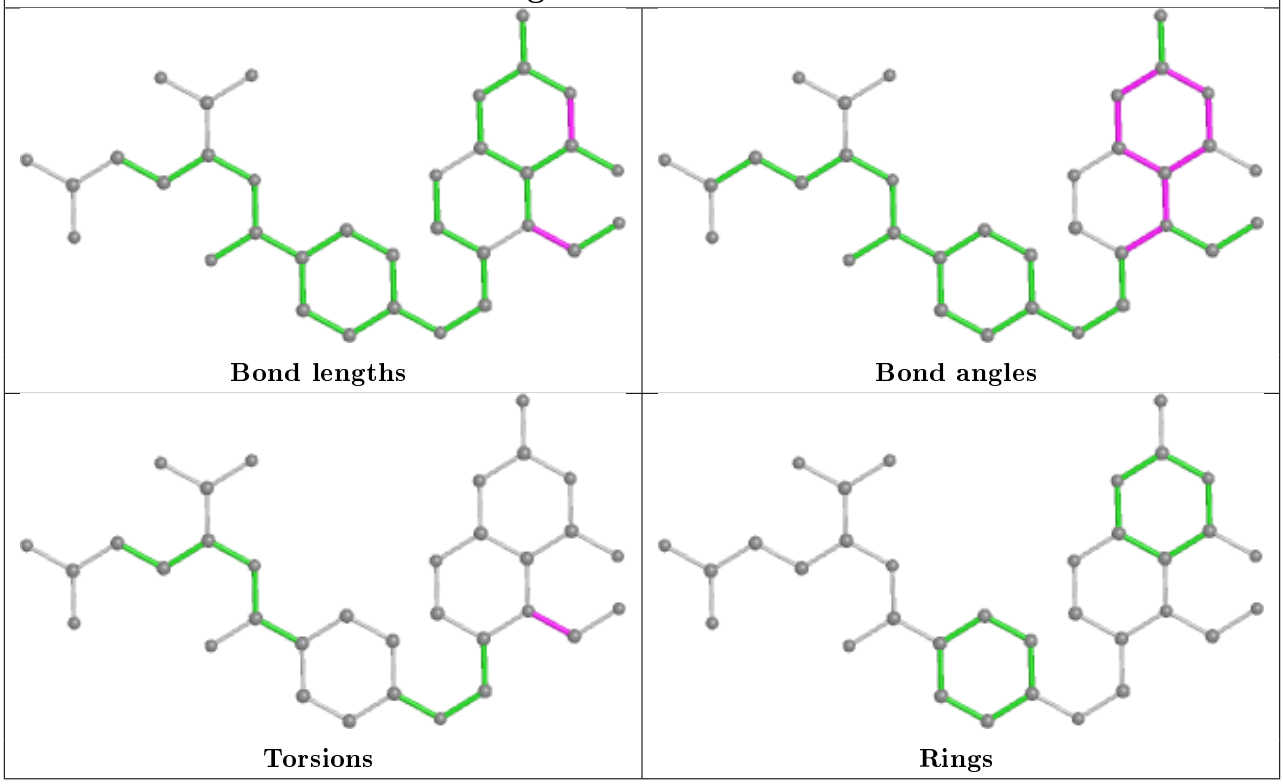
Ligand FFO E 403



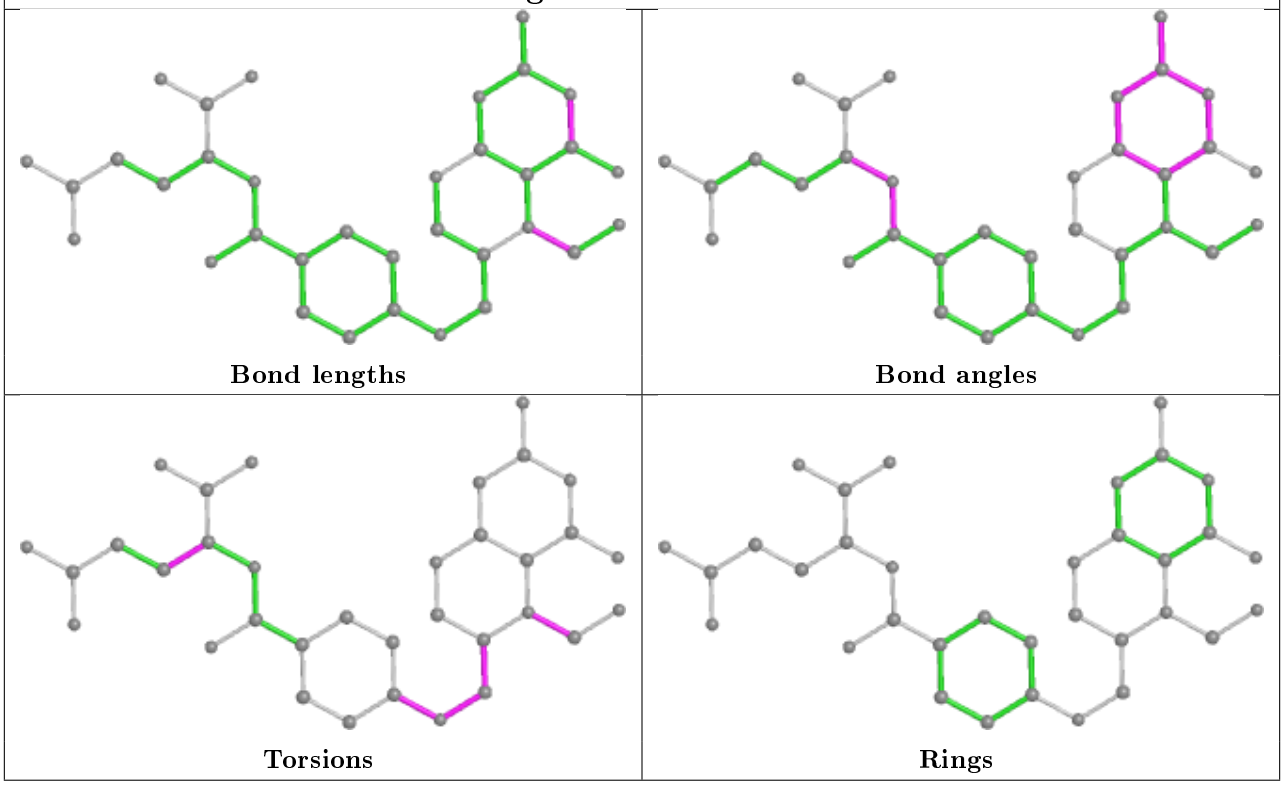
Ligand FFO D 403



Ligand FFO C 404



Ligand FFO G 403



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	286/325 (88%)	-0.04	2 (0%) 87 90	26, 42, 67, 117	0
1	B	287/325 (88%)	0.07	2 (0%) 87 90	26, 41, 67, 99	0
1	C	286/325 (88%)	-0.08	0 100 100	24, 38, 62, 94	0
1	D	286/325 (88%)	-0.02	1 (0%) 94 96	25, 40, 66, 95	0
1	F	286/325 (88%)	-0.03	4 (1%) 75 81	25, 40, 63, 108	0
1	H	286/325 (88%)	0.09	1 (0%) 94 96	27, 46, 73, 116	0
2	E	286/325 (88%)	0.01	3 (1%) 82 86	25, 41, 69, 92	0
2	G	286/325 (88%)	0.02	2 (0%) 87 90	26, 45, 71, 90	0
All	All	2289/2600 (88%)	0.00	15 (0%) 87 90	24, 42, 68, 117	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	313	VAL	3.0
1	A	313	VAL	2.9
1	F	151	SER	2.7
1	H	313	VAL	2.5
2	G	124	SER	2.4

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

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
1	CME	F	43	10/11	0.90	0.23	30,35,53,53	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	CME	F	195	10/11	0.91	0.15	41,52,68,80	0
1	CME	C	43	10/11	0.93	0.17	39,52,84,85	0
2	CME	E	195	10/11	0.94	0.15	42,53,64,73	0
1	CME	H	43	10/11	0.94	0.18	34,36,48,51	0
2	SCH	E	43	8/9	0.94	0.15	37,40,81,90	0
1	CME	D	195	10/11	0.95	0.14	36,44,55,69	0
1	CME	H	195	10/11	0.95	0.16	45,50,71,76	0
1	CME	D	43	10/11	0.95	0.15	33,41,76,76	0
2	CME	G	195	10/11	0.96	0.16	42,51,72,89	0
2	SCH	G	43	8/9	0.96	0.15	35,38,63,65	0
1	CME	B	195	10/11	0.96	0.15	36,47,80,82	0
1	CME	A	43	10/11	0.96	0.15	40,50,75,76	0
1	CME	C	195	10/11	0.97	0.15	35,48,59,74	0
1	CME	B	43	10/11	0.97	0.15	38,47,83,90	0
1	CME	A	195	10/11	0.97	0.15	34,44,63,69	0

6.3 Carbohydrates (i)

There are no carbohydrates 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, 95th 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
6	GOL	H	404	6/6	0.72	0.20	64,71,86,87	0
6	GOL	D	404	6/6	0.84	0.16	52,63,66,67	0
5	CL	A	404	1/1	0.88	0.07	55,55,55,55	0
5	CL	G	404	1/1	0.91	0.06	52,52,52,52	0
5	CL	B	404	1/1	0.92	0.23	51,51,51,51	0
3	SO4	A	402	5/5	0.93	0.19	91,92,98,108	0
3	SO4	H	402	5/5	0.93	0.27	96,101,106,119	0
3	SO4	F	402	5/5	0.93	0.19	87,93,103,108	0
3	SO4	C	403	5/5	0.93	0.19	78,85,89,96	0
4	FFO	G	403	34/34	0.94	0.12	35,48,54,64	0
5	CL	D	406	1/1	0.94	0.07	53,53,53,53	0
4	FFO	B	403	34/34	0.95	0.13	33,41,46,46	0
4	FFO	A	403	34/34	0.95	0.15	36,40,48,53	0

Continued on next page...

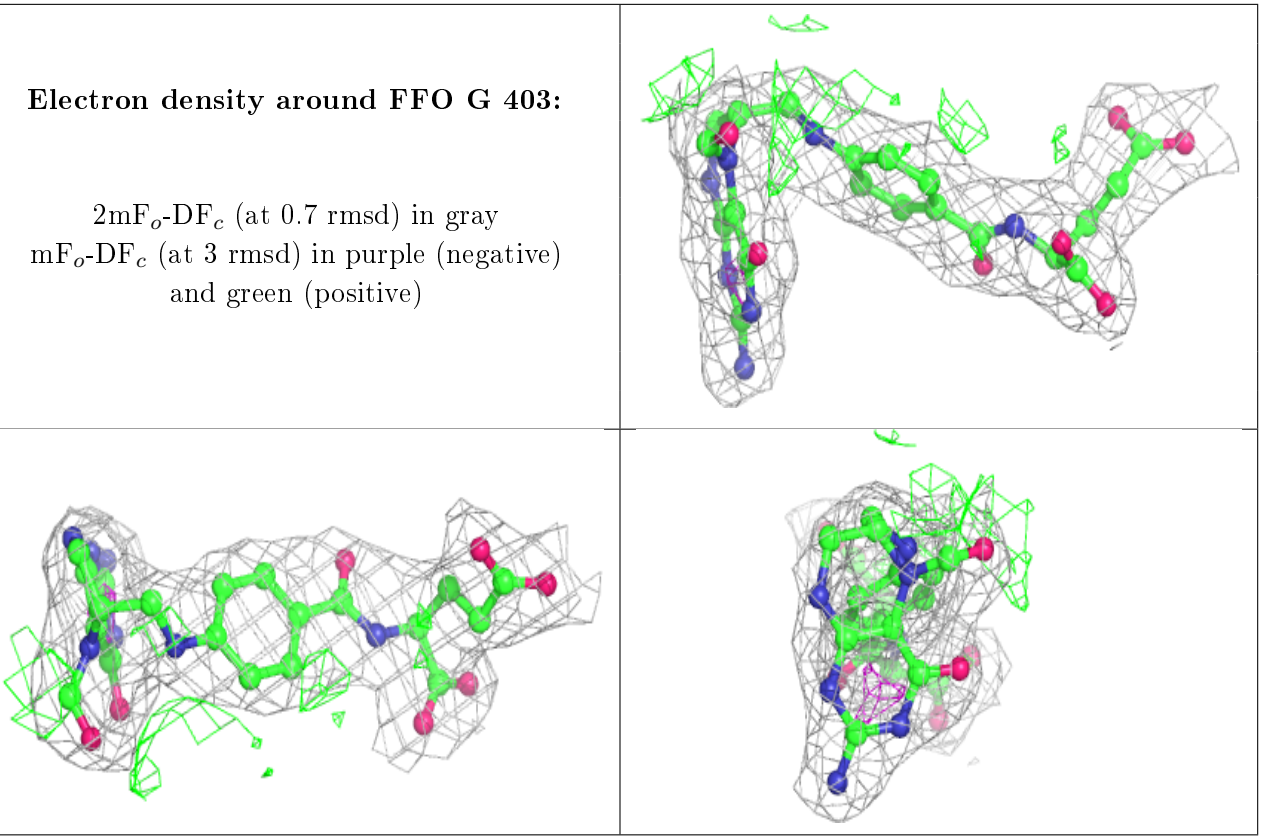
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FFO	D	403	34/34	0.95	0.14	26,39,44,47	0
4	FFO	F	403	34/34	0.95	0.17	30,37,45,58	0
5	CL	E	404	1/1	0.96	0.15	47,47,47,47	0
5	CL	A	405	1/1	0.96	0.11	49,49,49,49	0
4	FFO	C	404	34/34	0.96	0.17	30,38,44,45	0
4	FFO	E	403	34/34	0.96	0.12	34,43,55,64	0
5	CL	C	405	1/1	0.96	0.11	48,48,48,48	0
4	FFO	H	403	34/34	0.96	0.15	34,45,54,69	0
5	CL	D	405	1/1	0.97	0.20	54,54,54,54	0
3	SO4	G	402	5/5	0.97	0.14	63,63,69,73	0
3	SO4	E	402	5/5	0.97	0.15	65,68,74,81	0
3	SO4	D	401	5/5	0.98	0.11	50,59,65,66	0
3	SO4	D	402	5/5	0.98	0.15	70,72,75,80	0
3	SO4	B	402	5/5	0.98	0.07	66,70,74,79	0
3	SO4	G	401	5/5	0.99	0.21	36,45,48,50	0
3	SO4	C	401	5/5	0.99	0.18	27,32,35,36	0
3	SO4	B	401	5/5	0.99	0.15	35,37,42,43	0
3	SO4	H	401	5/5	0.99	0.14	34,38,39,41	0
3	SO4	A	401	5/5	0.99	0.18	24,27,35,35	0
3	SO4	E	401	5/5	0.99	0.17	42,44,52,53	0
3	SO4	C	402	5/5	0.99	0.20	33,34,39,40	0
3	SO4	F	401	5/5	1.00	0.15	31,32,33,34	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.

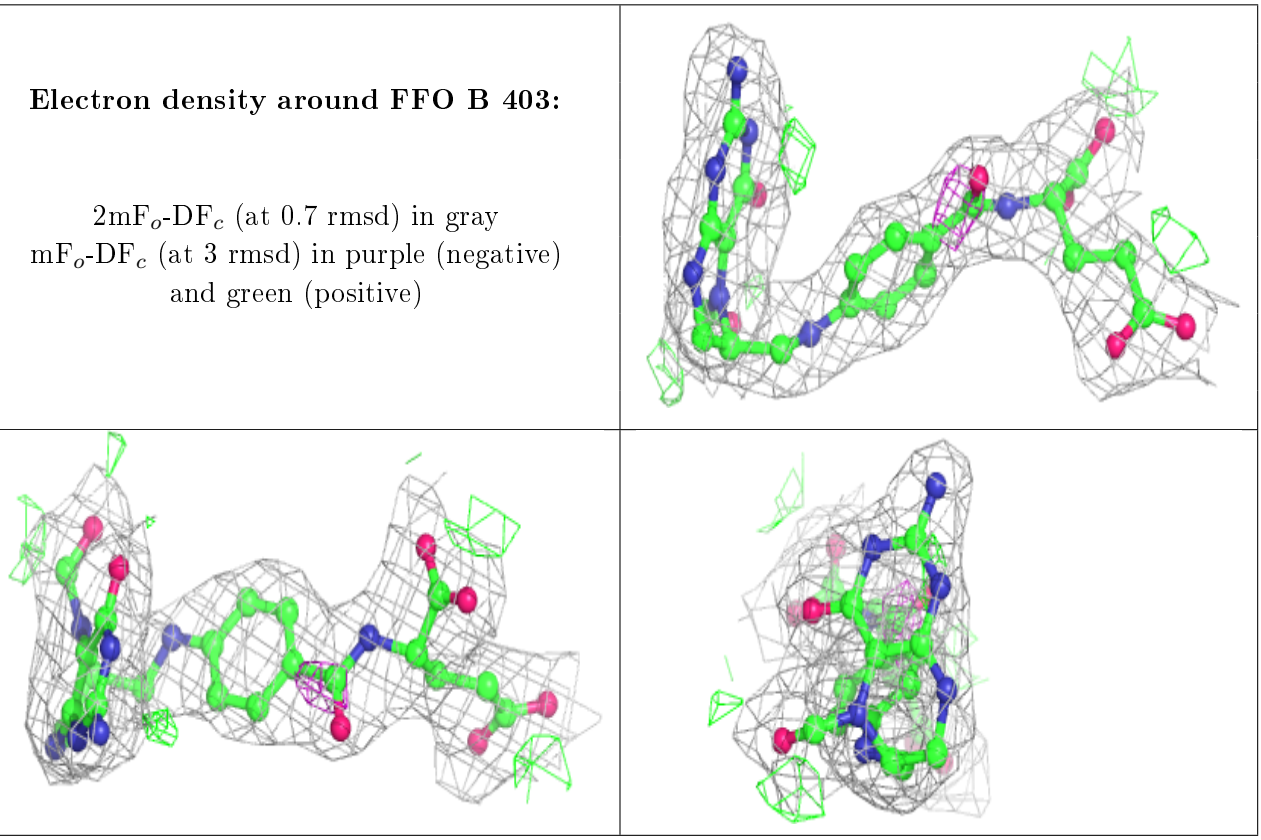
Electron density around FFO G 403:

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



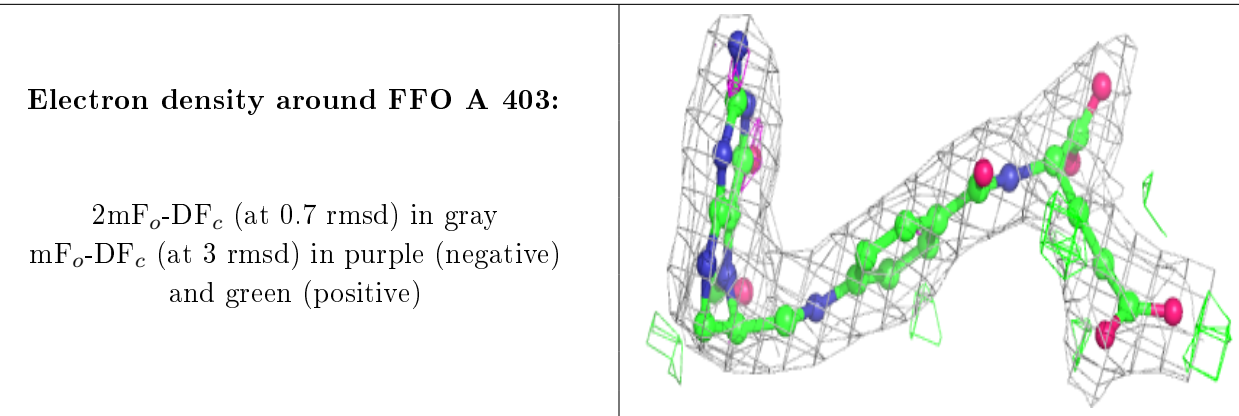
Electron density around FFO B 403:

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



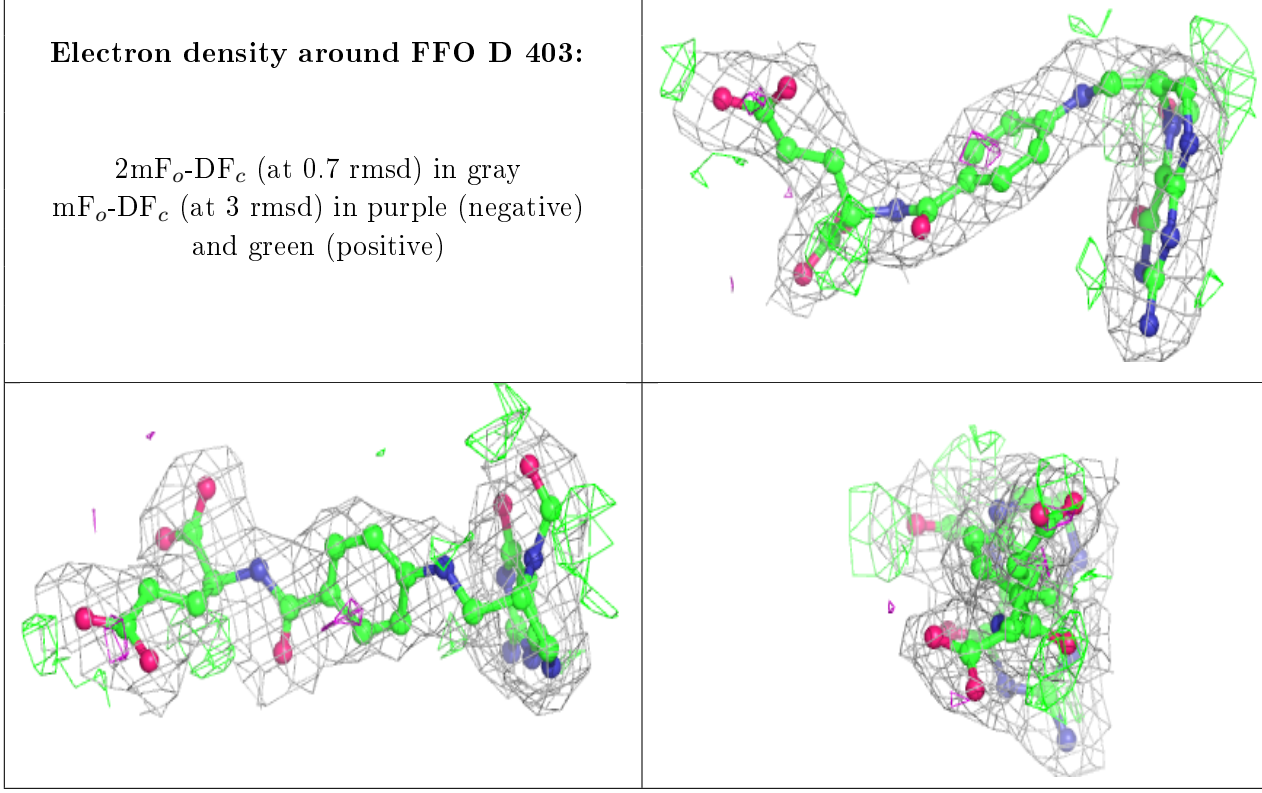
Electron density around FFO A 403:

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



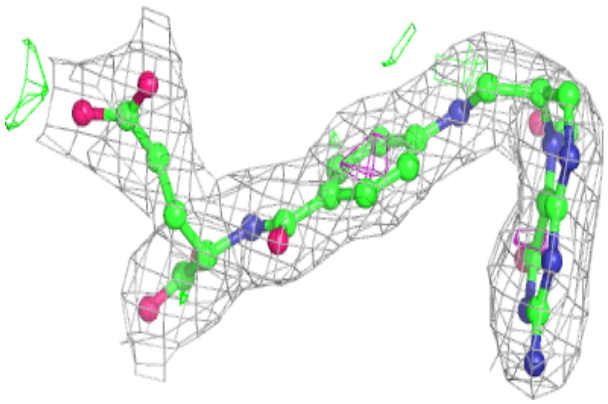
Electron density around FFO D 403:

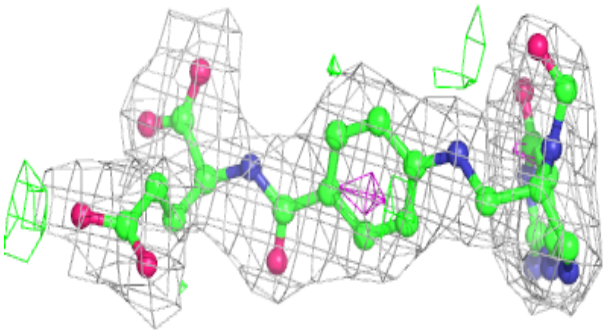
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

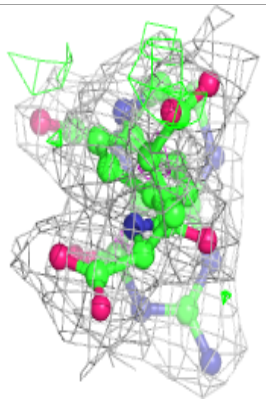


Electron density around FFO F 403:

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

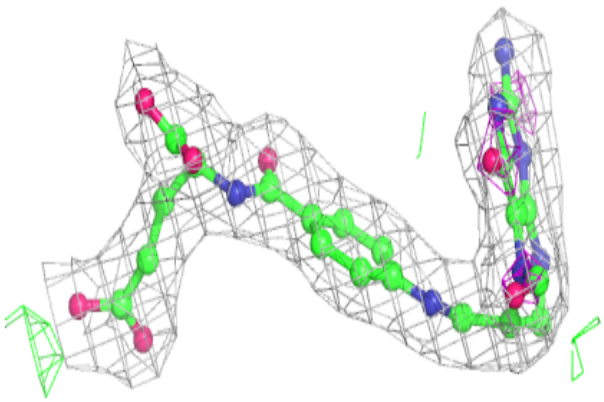


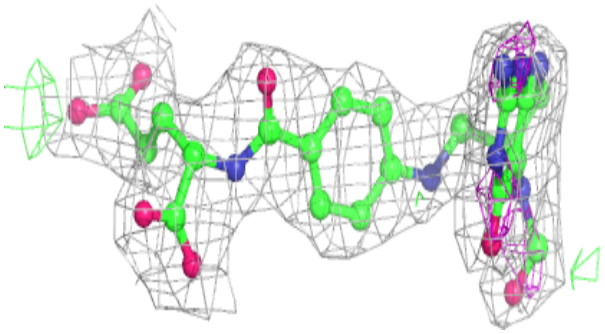


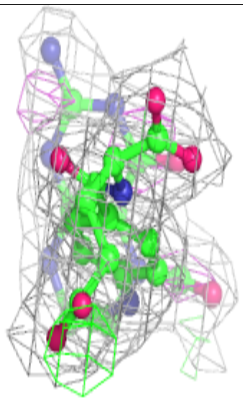


Electron density around FFO C 404:

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

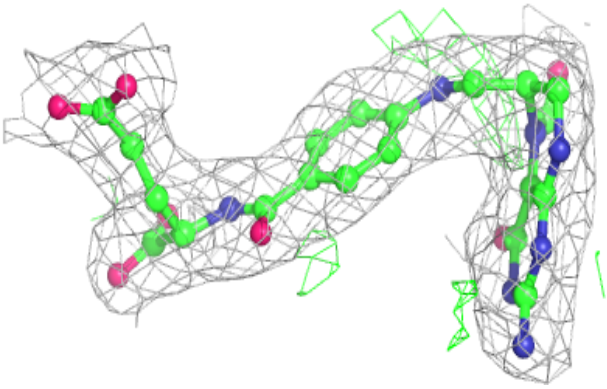


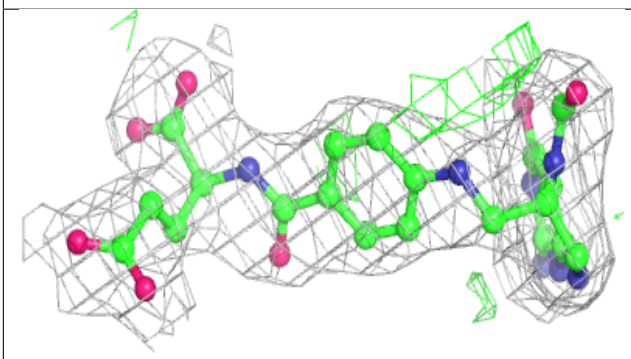


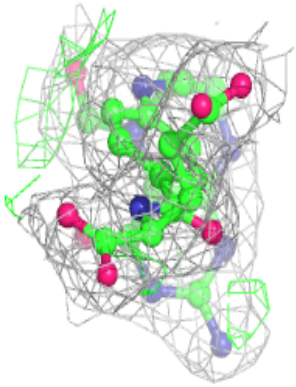


Electron density around FFO E 403:

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

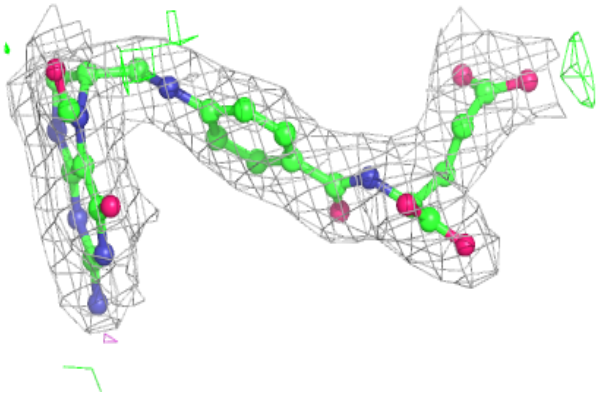


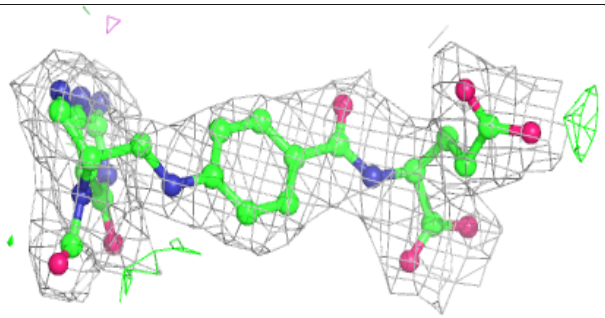


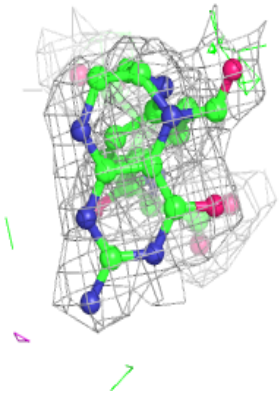


Electron density around FFO H 403:

$2mF_o$-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.