wwPDB X-ray Structure Validation Summary Report (i)

Sep 18, 2023 – 01:02 AM EDT

PDB ID	:	4YLH
Title	:	Crystal structure of DpgC with bound substrate analog and Xe on oxygen diffusion pathway
Authors	:	Li, K.; Di Russo, N.V.; Condurso, H.L.; Roitberg, A.E.; Bruner, S.D.
Deposited on	:	2015-03-05
Resolution	:	2.58 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

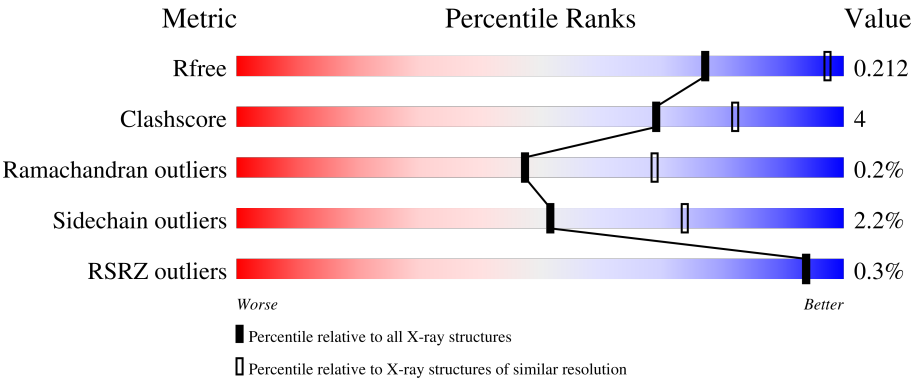
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh &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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.58 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	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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	440	86% 9%
1	B	440	88% 7%
1	C	440	86% 8%
1	D	440	86% 9%
1	E	440	86% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	440	83% 11%
1	G	440	87% 8%
1	H	440	86% 9%
1	I	440	85% 10%
1	J	440	2% 86% 9%
1	K	440	88% 7%
1	L	440	86% 8%

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
2	XE	A	501[A]	-	-	X	-
2	XE	B	501[A]	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 40478 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 DpgC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	421	Total	C	N	O	S	0	1	0
			3216	2031	595	580	10			
1	B	422	Total	C	N	O	S	0	1	0
			3268	2057	605	596	10			
1	C	421	Total	C	N	O	S	0	1	0
			3265	2054	607	594	10			
1	D	423	Total	C	N	O	S	0	1	0
			3258	2051	604	593	10			
1	E	421	Total	C	N	O	S	0	1	0
			3256	2049	604	593	10			
1	F	421	Total	C	N	O	S	0	1	0
			3258	2051	607	590	10			
1	G	422	Total	C	N	O	S	0	1	0
			3251	2048	602	591	10			
1	H	422	Total	C	N	O	S	0	1	0
			3256	2051	603	592	10			
1	I	421	Total	C	N	O	S	0	1	0
			3261	2052	607	592	10			
1	J	422	Total	C	N	O	S	0	1	0
			3240	2042	602	586	10			
1	K	422	Total	C	N	O	S	0	1	0
			3265	2055	604	596	10			
1	L	421	Total	C	N	O	S	0	1	0
			3258	2049	606	593	10			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ALA	-	expression tag	UNP Q8KLK7
A	0	MET	-	expression tag	UNP Q8KLK7
A	1	GLY	-	expression tag	UNP Q8KLK7
B	-1	ALA	-	expression tag	UNP Q8KLK7
B	0	MET	-	expression tag	UNP Q8KLK7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	expression tag	UNP Q8KLK7
C	-1	ALA	-	expression tag	UNP Q8KLK7
C	0	MET	-	expression tag	UNP Q8KLK7
C	1	GLY	-	expression tag	UNP Q8KLK7
D	-1	ALA	-	expression tag	UNP Q8KLK7
D	0	MET	-	expression tag	UNP Q8KLK7
D	1	GLY	-	expression tag	UNP Q8KLK7
E	-1	ALA	-	expression tag	UNP Q8KLK7
E	0	MET	-	expression tag	UNP Q8KLK7
E	1	GLY	-	expression tag	UNP Q8KLK7
F	-1	ALA	-	expression tag	UNP Q8KLK7
F	0	MET	-	expression tag	UNP Q8KLK7
F	1	GLY	-	expression tag	UNP Q8KLK7
G	-1	ALA	-	expression tag	UNP Q8KLK7
G	0	MET	-	expression tag	UNP Q8KLK7
G	1	GLY	-	expression tag	UNP Q8KLK7
H	-1	ALA	-	expression tag	UNP Q8KLK7
H	0	MET	-	expression tag	UNP Q8KLK7
H	1	GLY	-	expression tag	UNP Q8KLK7
I	-1	ALA	-	expression tag	UNP Q8KLK7
I	0	MET	-	expression tag	UNP Q8KLK7
I	1	GLY	-	expression tag	UNP Q8KLK7
J	-1	ALA	-	expression tag	UNP Q8KLK7
J	0	MET	-	expression tag	UNP Q8KLK7
J	1	GLY	-	expression tag	UNP Q8KLK7
K	-1	ALA	-	expression tag	UNP Q8KLK7
K	0	MET	-	expression tag	UNP Q8KLK7
K	1	GLY	-	expression tag	UNP Q8KLK7
L	-1	ALA	-	expression tag	UNP Q8KLK7
L	0	MET	-	expression tag	UNP Q8KLK7
L	1	GLY	-	expression tag	UNP Q8KLK7

- Molecule 2 is XENON (three-letter code: XE) (formula: Xe).

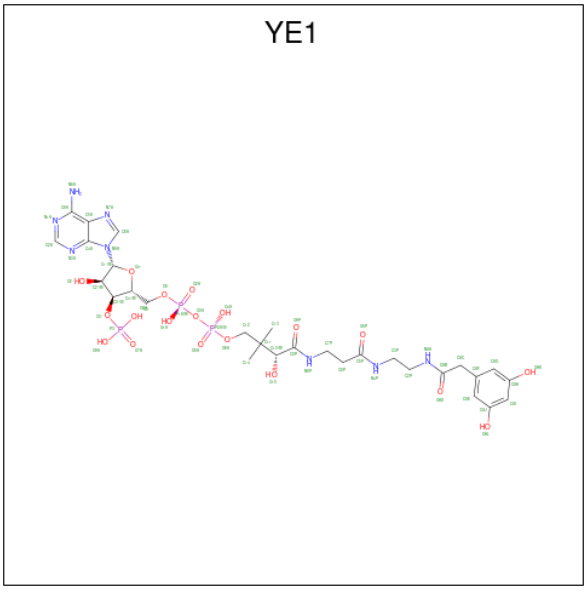
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Xe 2 2	0	1
2	B	1	Total Xe 2 2	0	1
2	C	1	Total Xe 1 1	0	1
2	D	1	Total Xe 1 1	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	E	1	Total Xe 1 1	0	1
2	F	1	Total Xe 1 1	0	1
2	G	1	Total Xe 1 1	0	1
2	H	1	Total Xe 1 1	0	1
2	I	1	Total Xe 1 1	0	1
2	J	1	Total Xe 1 1	0	1
2	K	1	Total Xe 1 1	0	1
2	L	1	Total Xe 1 1	0	1

- Molecule 3 is [(2R,3S,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-4-HYDROXY-3-(PHOSPHONOOXY)TETRAHYDROFURAN-2-YL]METHYL (3R)-4-({3-[(2-{[(3,5-DIHYDROXYPHENYL)ACETYL]AMINO}ETHYL)AMINO]-3-OXOPROPYL}AMINO)-3-HYDROXY-2,2-DIMETHYL-4-OXOBUTYL DIHYDROGEN DIPHOSPHATE (three-letter code: YE1) (formula: $C_{29}H_{43}N_8O_{19}P_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O P 59 29 8 19 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	C	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	D	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	E	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	F	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	G	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	H	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	I	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	J	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	K	1	Total	C	N	O	P	0	0
			59	29	8	19	3		
3	L	1	Total	C	N	O	P	0	0
			59	29	8	19	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	48	Total	O	0	0
			48	48		
4	B	62	Total	O	0	0
			62	62		
4	C	42	Total	O	0	0
			42	42		
4	D	56	Total	O	0	0
			56	56		
4	E	74	Total	O	0	0
			74	74		
4	F	74	Total	O	0	0
			74	74		
4	G	54	Total	O	0	0
			54	54		
4	H	59	Total	O	0	0
			59	59		

Continued on next page...

Continued from previous page...

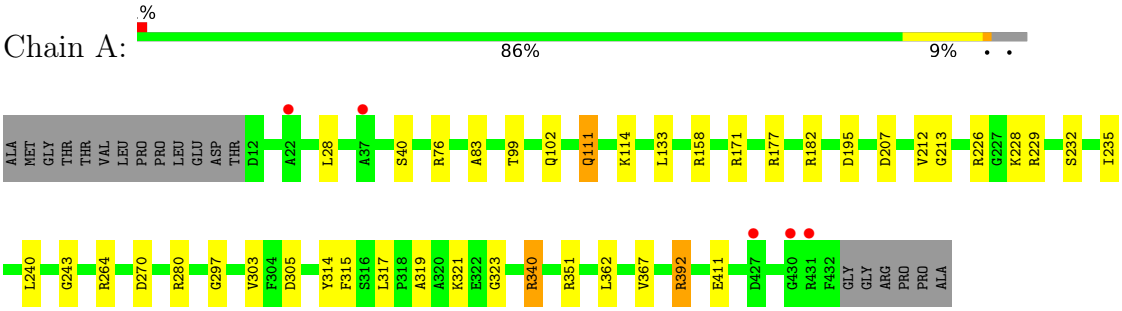
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	I	65	Total O 65 65	0	0
4	J	37	Total O 37 37	0	0
4	K	76	Total O 76 76	0	0
4	L	57	Total O 57 57	0	0

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.

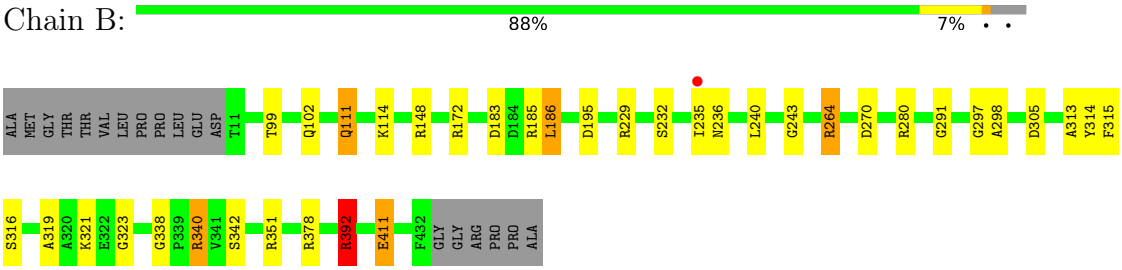
• Molecule 1: DpgC

Chain A:



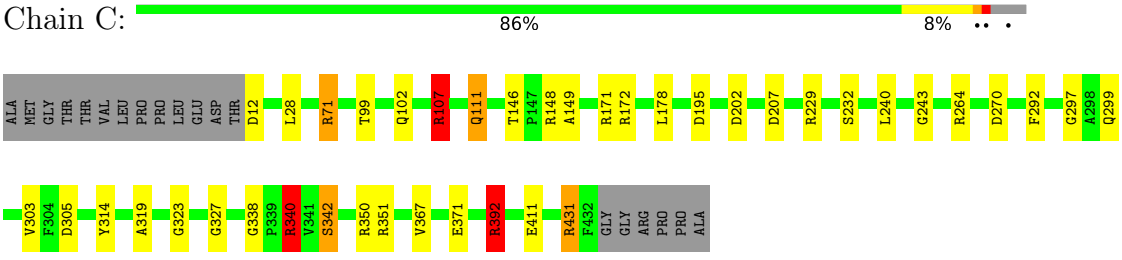
• Molecule 1: DpgC

Chain B:



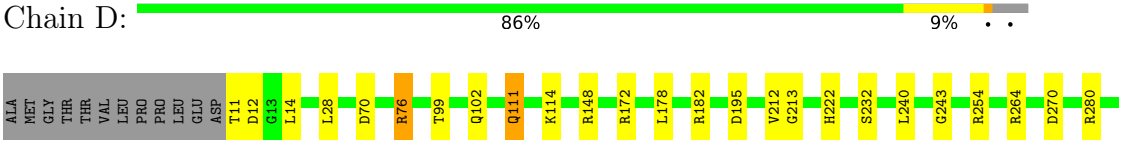
• Molecule 1: DpgC

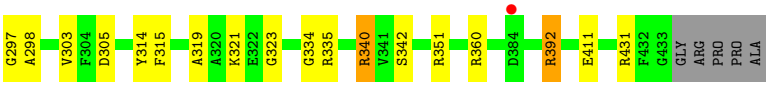
Chain C:



• Molecule 1: DpgC

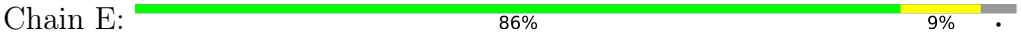
Chain D:

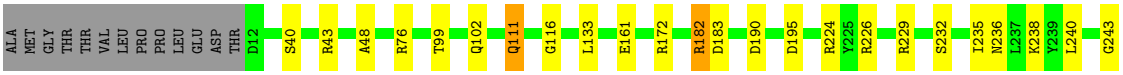


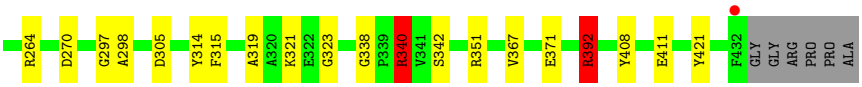


• Molecule 1: DpgC

Chain E:

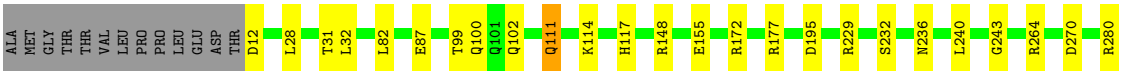






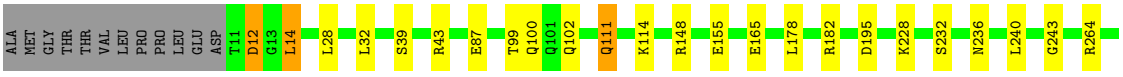
• Molecule 1: DpgC

Chain F:



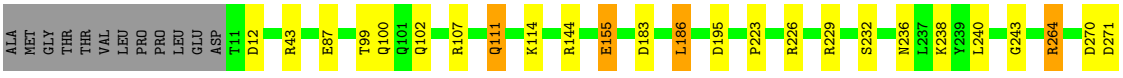
• Molecule 1: DpgC

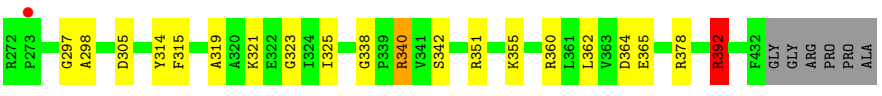
Chain G:



• Molecule 1: DpgC

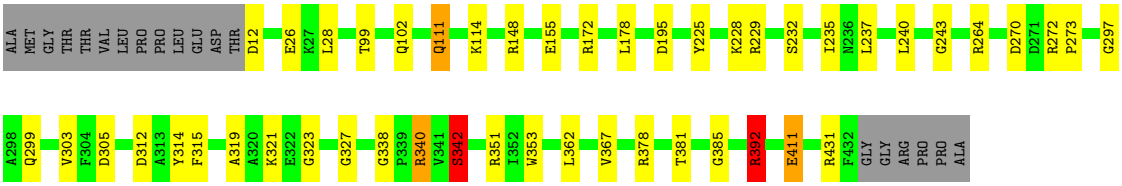
Chain H:





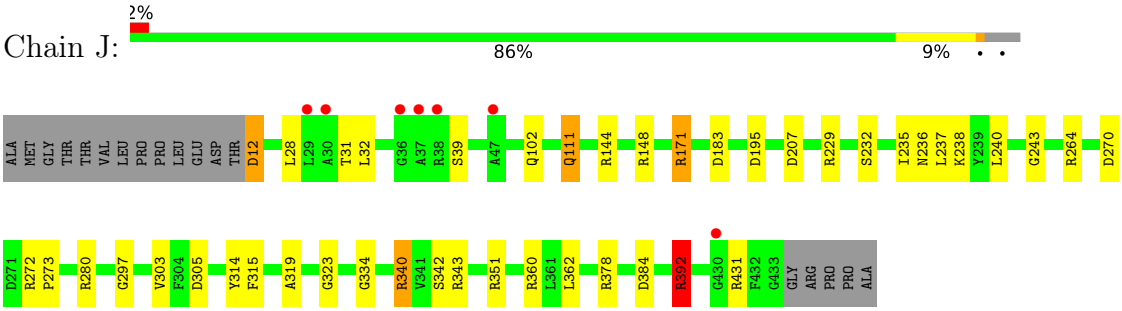
• Molecule 1: DpgC

Chain I:



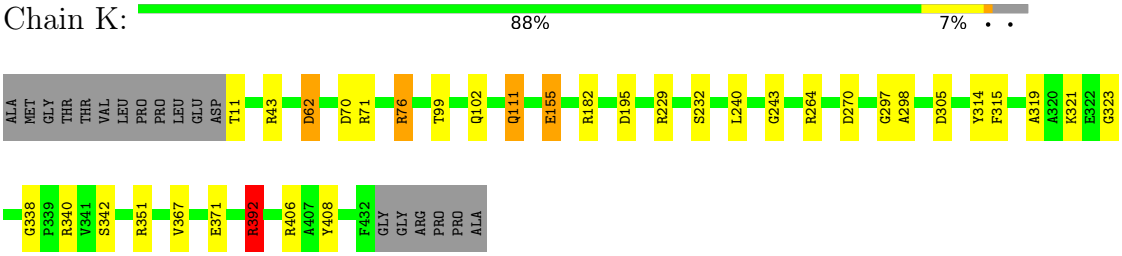
• Molecule 1: DpgC

Chain J:



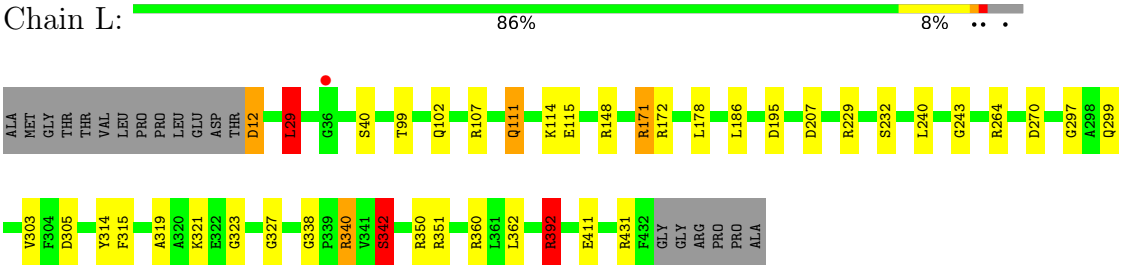
• Molecule 1: DpgC

Chain K:



• Molecule 1: DpgC

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	139.13Å 170.94Å 156.00Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	39.52 – 2.58 39.52 – 2.58	Depositor EDS
% Data completeness (in resolution range)	98.7 (39.52-2.58) 98.7 (39.52-2.58)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.08 (at 2.58Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.171 , 0.208 0.176 , 0.212	Depositor DCC
R_{free} test set	11219 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.648	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 6.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.447 for h,-k,-l	Xtriage
Reported twinning fraction	0.515 for H, K, L 0.485 for -h,-k,l	Depositor
Outliers	1 of 225931 reflections (0.000%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	40478	wwPDB-VP
Average B, all atoms (Å^2)	39.0	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 22.76 % 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 5.3374e-03. 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: XE, YE1

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	0.83	1/3277 (0.0%)	1.00	14/4449 (0.3%)
1	B	0.86	1/3330 (0.0%)	1.02	17/4516 (0.4%)
1	C	0.88	0/3327	0.99	11/4511 (0.2%)
1	D	0.91	0/3319	1.00	11/4502 (0.2%)
1	E	0.86	0/3317	1.00	16/4498 (0.4%)
1	F	0.87	1/3320 (0.0%)	1.04	19/4502 (0.4%)
1	G	0.87	0/3312	1.00	11/4494 (0.2%)
1	H	0.83	0/3317	1.02	19/4499 (0.4%)
1	I	0.93	3/3323 (0.1%)	1.00	12/4506 (0.3%)
1	J	0.78	1/3301 (0.0%)	0.98	13/4478 (0.3%)
1	K	0.93	3/3327 (0.1%)	1.11	20/4513 (0.4%)
1	L	0.87	2/3320 (0.1%)	0.97	13/4502 (0.3%)
All	All	0.87	12/39790 (0.0%)	1.01	176/53970 (0.3%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	62	ASP	CB-CG	6.15	1.64	1.51
1	I	225	TYR	CD1-CE1	-6.03	1.30	1.39
1	F	342	SER	CB-OG	-5.97	1.34	1.42
1	J	102	GLN	CG-CD	5.89	1.64	1.51
1	I	342	SER	CB-OG	-5.88	1.34	1.42

The worst 5 of 176 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	406	ARG	NE-CZ-NH1	19.54	130.07	120.30
1	K	406	ARG	NE-CZ-NH2	-13.35	113.62	120.30
1	K	62	ASP	CB-CG-OD2	12.65	129.68	118.30
1	G	392	ARG	NE-CZ-NH2	-11.59	114.51	120.30
1	K	340	ARG	NE-CZ-NH2	-11.55	114.53	120.30

There are no chirality outliers.

There are no planarity outliers.

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	3216	0	3201	23	0
1	B	3268	0	3266	21	0
1	C	3265	0	3266	26	0
1	D	3258	0	3254	19	0
1	E	3256	0	3255	29	0
1	F	3258	0	3260	27	0
1	G	3251	0	3241	21	0
1	H	3256	0	3255	28	0
1	I	3261	0	3262	33	0
1	J	3240	0	3235	27	0
1	K	3265	0	3257	16	0
1	L	3258	0	3251	25	0
2	A	2	0	0	2	0
2	B	2	0	0	3	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	A	59	0	38	1	0
3	B	59	0	39	3	0
3	C	59	0	38	2	0
3	D	59	0	38	2	0
3	E	59	0	39	4	0
3	F	59	0	36	2	0
3	G	59	0	38	2	0
3	H	59	0	38	4	0
3	I	59	0	38	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	59	0	39	9	0
3	K	59	0	38	0	0
3	L	59	0	38	1	0
4	A	48	0	0	1	0
4	B	62	0	0	0	0
4	C	42	0	0	1	0
4	D	56	0	0	1	0
4	E	74	0	0	2	0
4	F	74	0	0	2	0
4	G	54	0	0	1	0
4	H	59	0	0	4	0
4	I	65	0	0	1	0
4	J	37	0	0	0	0
4	K	76	0	0	0	0
4	L	57	0	0	1	0
All	All	40478	0	39460	292	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 4.

The worst 5 of 292 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:A:76:ARG:HH12	1:A:133:LEU:HD22	1.31	0.95
1:L:29:LEU:HD21	1:L:115:GLU:HG3	1.49	0.94
1:E:76:ARG:HH12	1:E:133:LEU:HD22	1.34	0.93
1:A:317:LEU:HD11	2:A:501[A]:XE:XE	2.56	0.84
1:K:70:ASP:OD1	1:K:76:ARG:NH1	2.11	0.83

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	420/440 (96%)	411 (98%)	8 (2%)	1 (0%)	47 69
1	B	421/440 (96%)	413 (98%)	7 (2%)	1 (0%)	47 69
1	C	420/440 (96%)	411 (98%)	8 (2%)	1 (0%)	47 69
1	D	422/440 (96%)	414 (98%)	7 (2%)	1 (0%)	47 69
1	E	420/440 (96%)	411 (98%)	8 (2%)	1 (0%)	47 69
1	F	420/440 (96%)	410 (98%)	9 (2%)	1 (0%)	47 69
1	G	421/440 (96%)	412 (98%)	8 (2%)	1 (0%)	47 69
1	H	421/440 (96%)	413 (98%)	7 (2%)	1 (0%)	47 69
1	I	420/440 (96%)	411 (98%)	8 (2%)	1 (0%)	47 69
1	J	421/440 (96%)	413 (98%)	7 (2%)	1 (0%)	47 69
1	K	421/440 (96%)	411 (98%)	9 (2%)	1 (0%)	47 69
1	L	420/440 (96%)	411 (98%)	8 (2%)	1 (0%)	47 69
All	All	5047/5280 (96%)	4941 (98%)	94 (2%)	12 (0%)	47 69

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	270	ASP
1	A	270	ASP
1	B	270	ASP
1	D	270	ASP
1	E	270	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	316/345 (92%)	310 (98%)	6 (2%)	57 77
1	B	328/345 (95%)	322 (98%)	6 (2%)	59 78
1	C	328/345 (95%)	320 (98%)	8 (2%)	49 72
1	D	325/345 (94%)	317 (98%)	8 (2%)	47 70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	E	326/345 (94%)	321 (98%)	5 (2%)	65 82
1	F	326/345 (94%)	318 (98%)	8 (2%)	47 70
1	G	323/345 (94%)	313 (97%)	10 (3%)	40 64
1	H	325/345 (94%)	319 (98%)	6 (2%)	59 78
1	I	327/345 (95%)	319 (98%)	8 (2%)	49 72
1	J	321/345 (93%)	312 (97%)	9 (3%)	43 67
1	K	327/345 (95%)	321 (98%)	6 (2%)	59 78
1	L	326/345 (94%)	319 (98%)	7 (2%)	53 75
All	All	3898/4140 (94%)	3811 (98%)	87 (2%)	52 74

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	111	GLN
1	J	360	ARG
1	I	303	VAL
1	J	12	ASP
1	K	62	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	423	HIS
1	J	111	GLN
1	I	79	HIS
1	I	299	GLN
1	K	79	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)

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 26 ligands modelled in this entry, 14 are monoatomic - leaving 12 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	YE1	L	502	-	54,62,62	2.76	22 (40%)	70,92,92	1.85	17 (24%)
3	YE1	F	502	-	54,62,62	2.52	20 (37%)	70,92,92	1.86	24 (34%)
3	YE1	E	502	-	54,62,62	2.42	15 (27%)	70,92,92	1.59	12 (17%)
3	YE1	J	502	-	54,62,62	2.81	18 (33%)	70,92,92	1.54	11 (15%)
3	YE1	A	502	-	54,62,62	2.58	17 (31%)	70,92,92	1.61	15 (21%)
3	YE1	K	502	-	54,62,62	2.47	16 (29%)	70,92,92	1.61	13 (18%)
3	YE1	H	502	-	54,62,62	2.65	20 (37%)	70,92,92	1.71	15 (21%)
3	YE1	G	502	-	54,62,62	2.73	21 (38%)	70,92,92	1.59	17 (24%)
3	YE1	B	502	-	54,62,62	2.49	21 (38%)	70,92,92	1.62	16 (22%)
3	YE1	I	502	-	54,62,62	2.47	16 (29%)	70,92,92	1.61	13 (18%)
3	YE1	C	502	-	54,62,62	3.39	26 (48%)	70,92,92	1.59	14 (20%)
3	YE1	D	502	-	54,62,62	2.55	21 (38%)	70,92,92	1.68	15 (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	YE1	L	502	-	-	11/51/71/71	0/4/4/4
3	YE1	F	502	-	-	8/51/71/71	0/4/4/4
3	YE1	E	502	-	-	7/51/71/71	0/4/4/4
3	YE1	J	502	-	-	18/51/71/71	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	YE1	A	502	-	-	10/51/71/71	0/4/4/4
3	YE1	K	502	-	-	4/51/71/71	0/4/4/4
3	YE1	H	502	-	-	11/51/71/71	0/4/4/4
3	YE1	G	502	-	-	8/51/71/71	0/4/4/4
3	YE1	B	502	-	-	11/51/71/71	0/4/4/4
3	YE1	I	502	-	-	4/51/71/71	0/4/4/4
3	YE1	C	502	-	-	5/51/71/71	0/4/4/4
3	YE1	D	502	-	-	13/51/71/71	0/4/4/4

The worst 5 of 233 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	J	502	YE1	P3'-O3'	14.35	1.86	1.59
3	C	502	YE1	P3'-O3'	14.18	1.86	1.59
3	A	502	YE1	P3'-O3'	13.16	1.84	1.59
3	L	502	YE1	P3'-O3'	10.53	1.79	1.59
3	G	502	YE1	P3'-O3'	10.22	1.78	1.59

The worst 5 of 182 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	L	502	YE1	O6A-C12-C11	-7.59	98.35	110.55
3	E	502	YE1	O3'-P3'-O7A	-5.43	88.43	109.39
3	B	502	YE1	O3'-P3'-O7A	-5.22	89.26	109.39
3	F	502	YE1	CAJ-CAE-CAF	-5.04	116.62	120.35
3	C	502	YE1	CAC-CAB-NAA	-4.94	109.51	116.19

There are no chirality outliers.

5 of 110 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	YE1	C2'-C3'-O3'-P3'
3	A	502	YE1	C5'-O5'-P1A-O2A
3	B	502	YE1	P1A-O3A-P2A-O6A
3	B	502	YE1	C14-C11-C12-O6A
3	B	502	YE1	C13-C11-C12-O6A

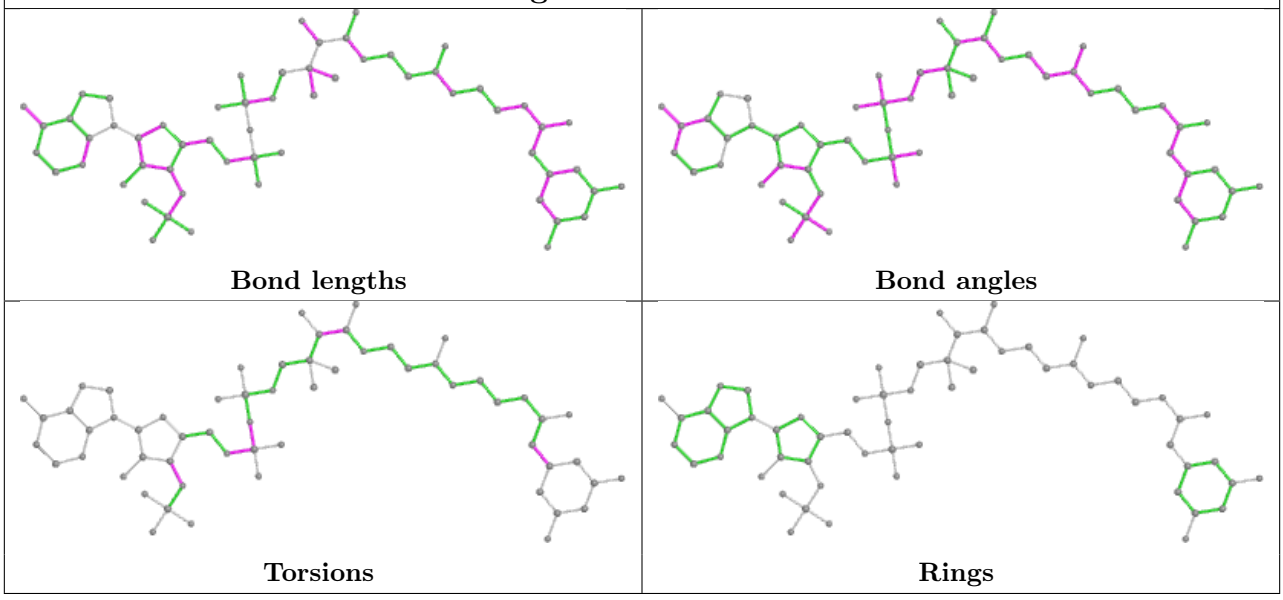
There are no ring outliers.

11 monomers are involved in 34 short contacts:

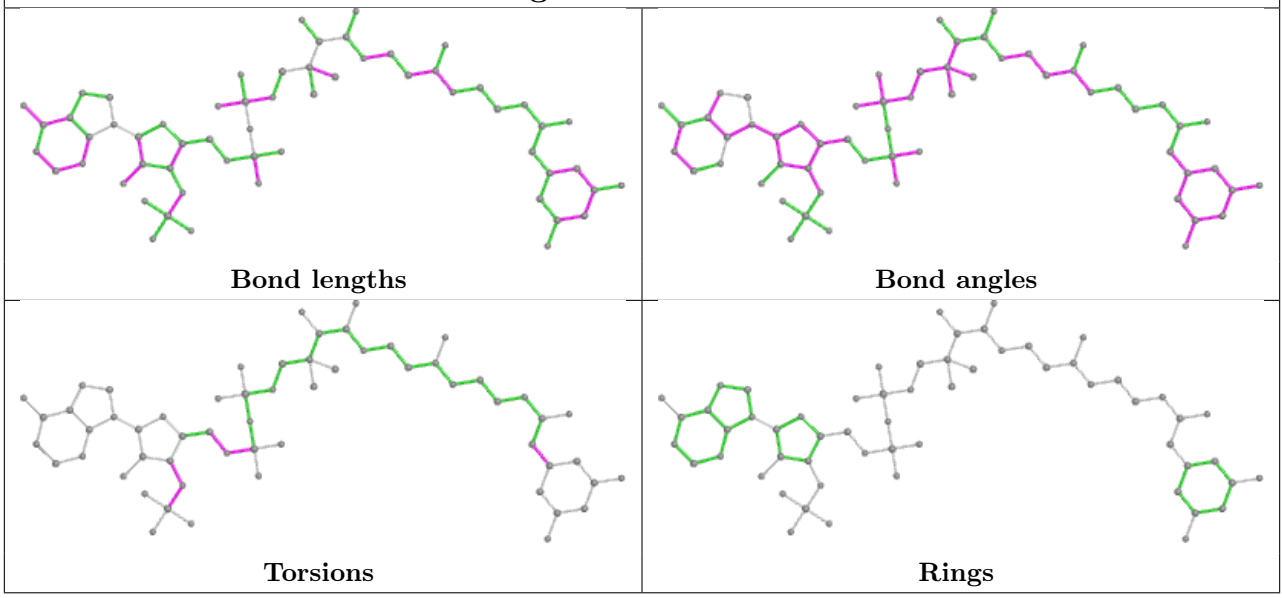
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	L	502	YE1	1	0
3	F	502	YE1	2	0
3	E	502	YE1	4	0
3	J	502	YE1	9	0
3	A	502	YE1	1	0
3	H	502	YE1	4	0
3	G	502	YE1	2	0
3	B	502	YE1	3	0
3	I	502	YE1	4	0
3	C	502	YE1	2	0
3	D	502	YE1	2	0

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.

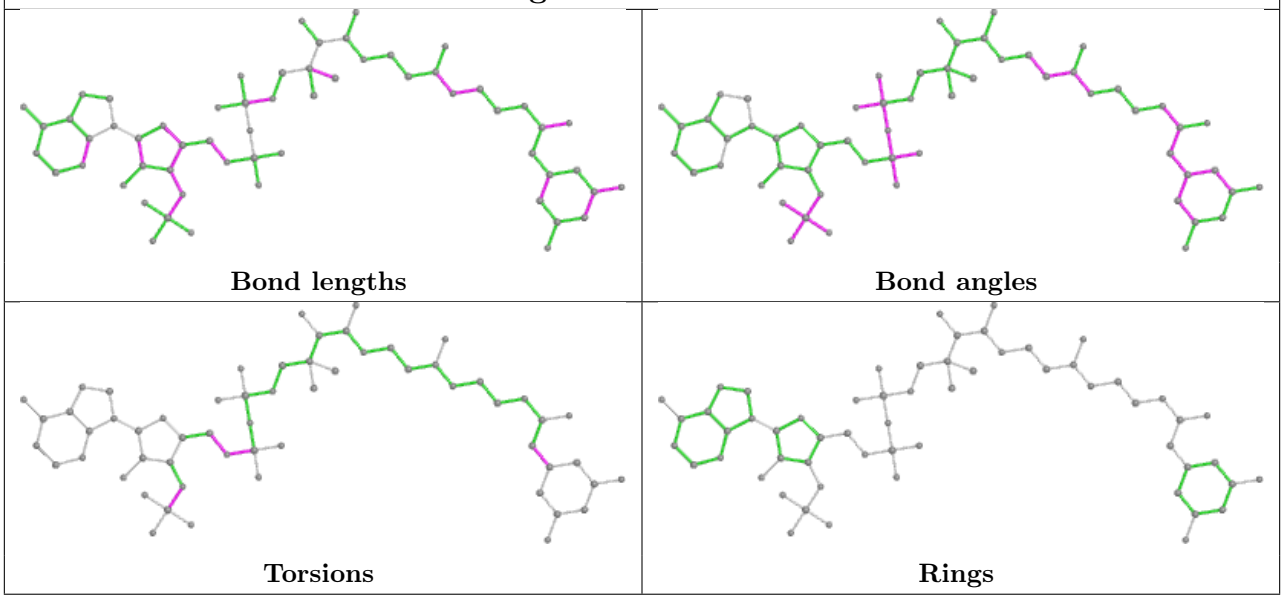
Ligand YE1 L 502



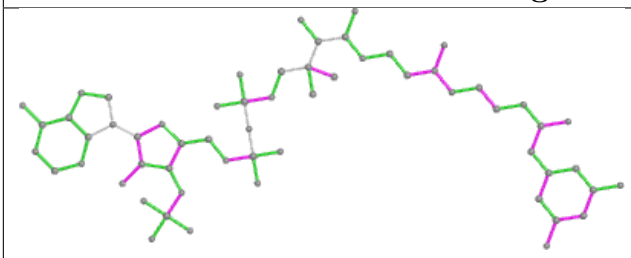
Ligand YE1 F 502



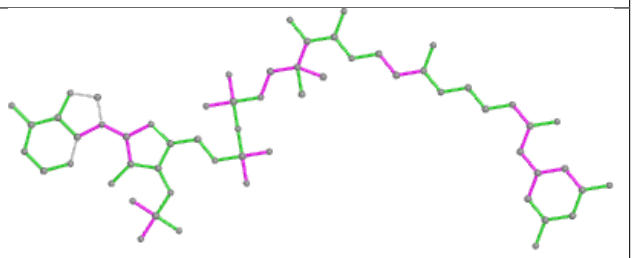
Ligand YE1 E 502



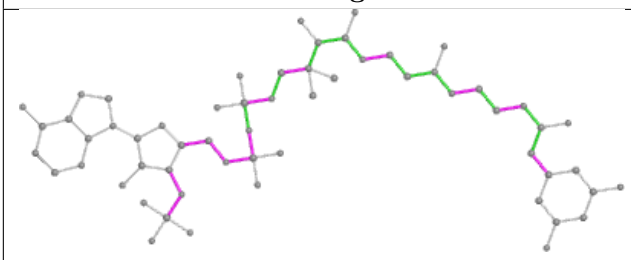
Ligand YE1 J 502



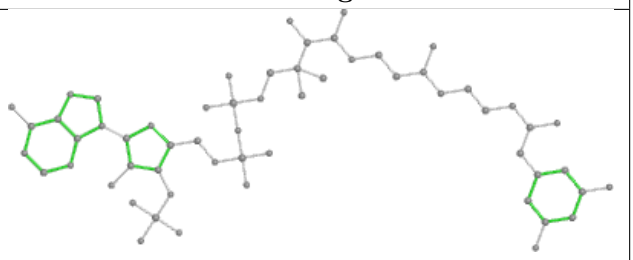
Bond lengths



Bond angles

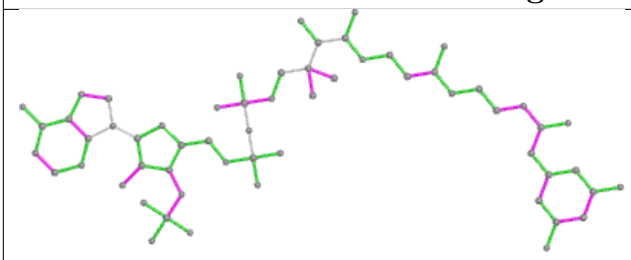


Torsions

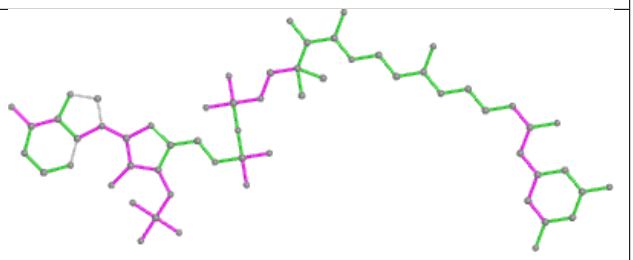


Rings

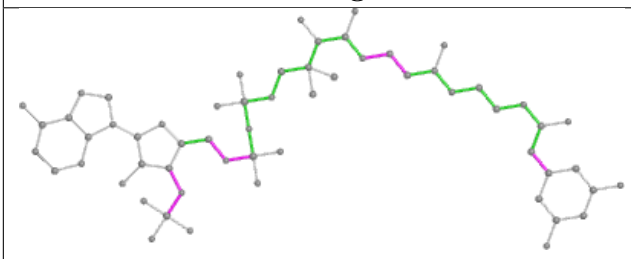
Ligand YE1 A 502



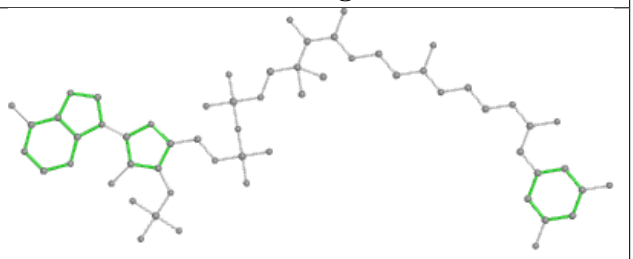
Bond lengths



Bond angles

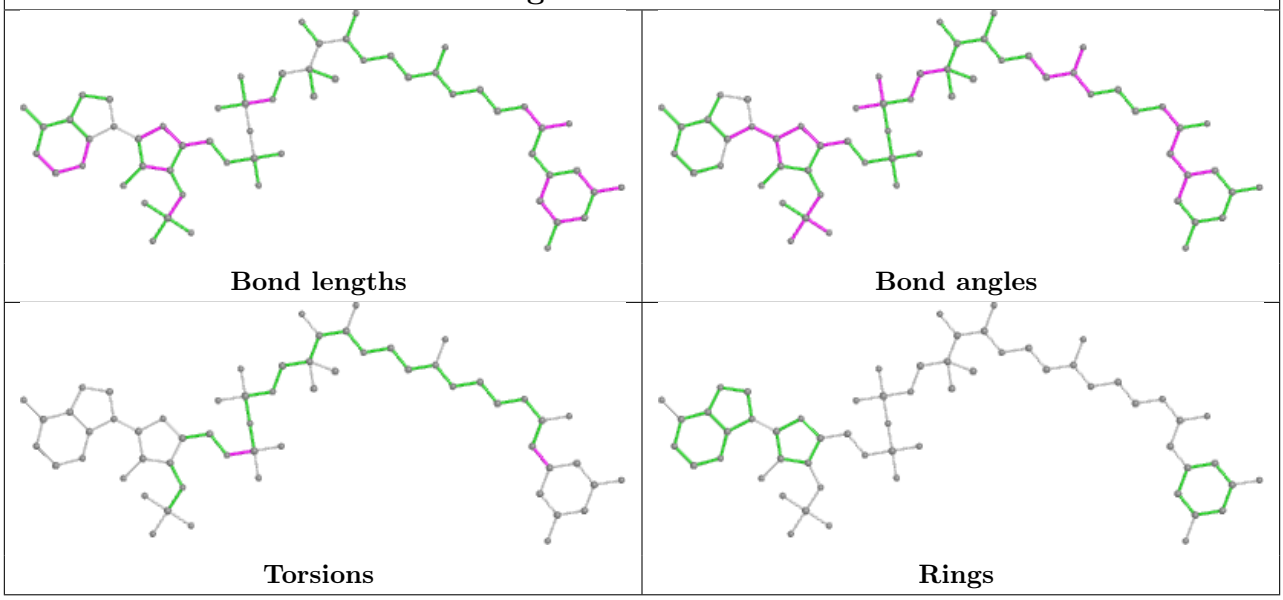


Torsions

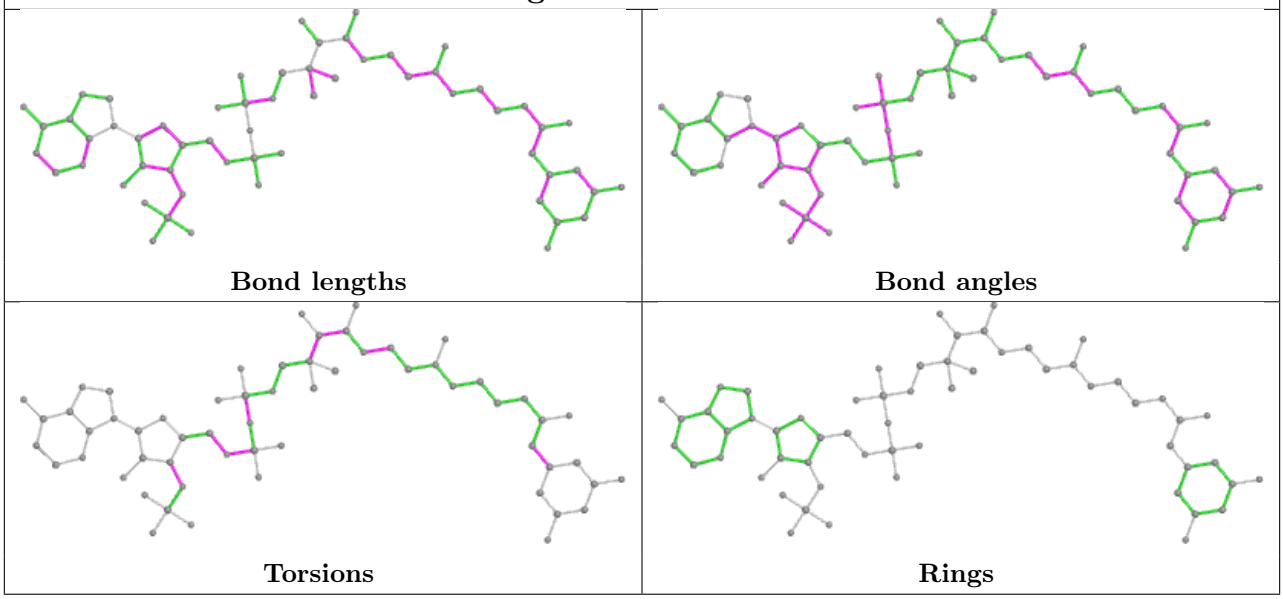


Rings

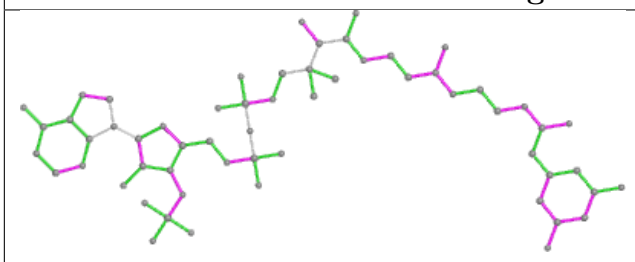
Ligand YE1 K 502



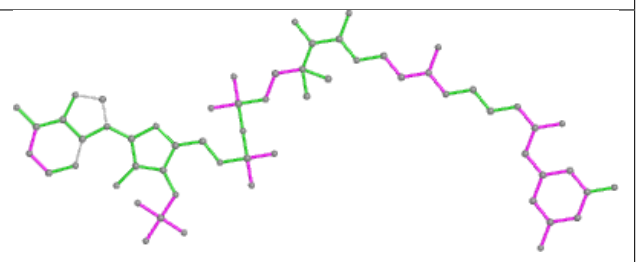
Ligand YE1 H 502



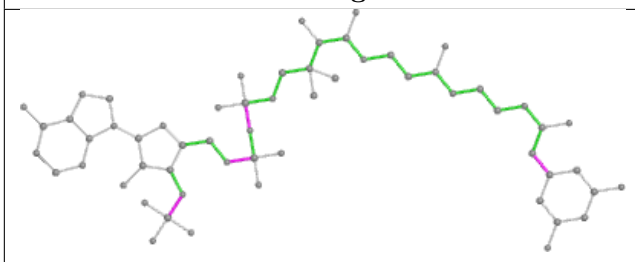
Ligand YE1 G 502



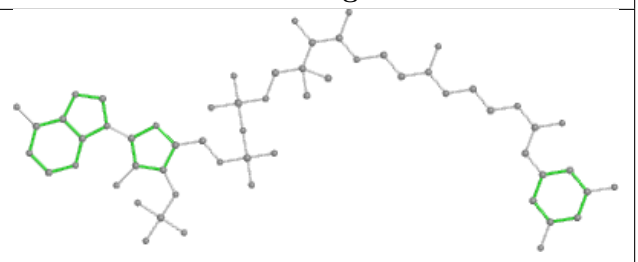
Bond lengths



Bond angles

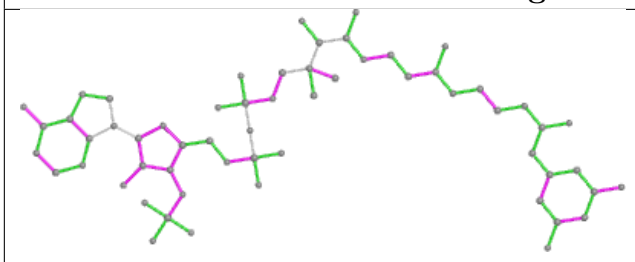


Torsions

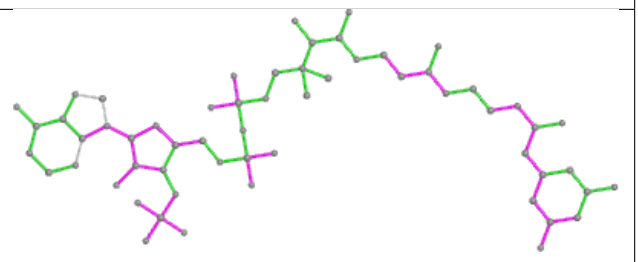


Rings

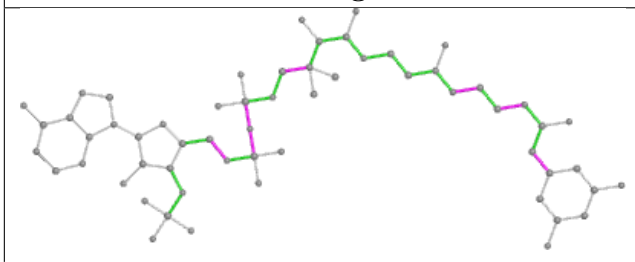
Ligand YE1 B 502



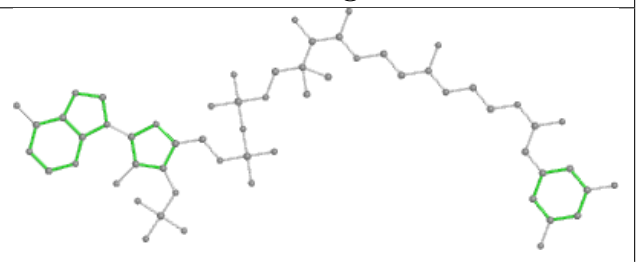
Bond lengths



Bond angles

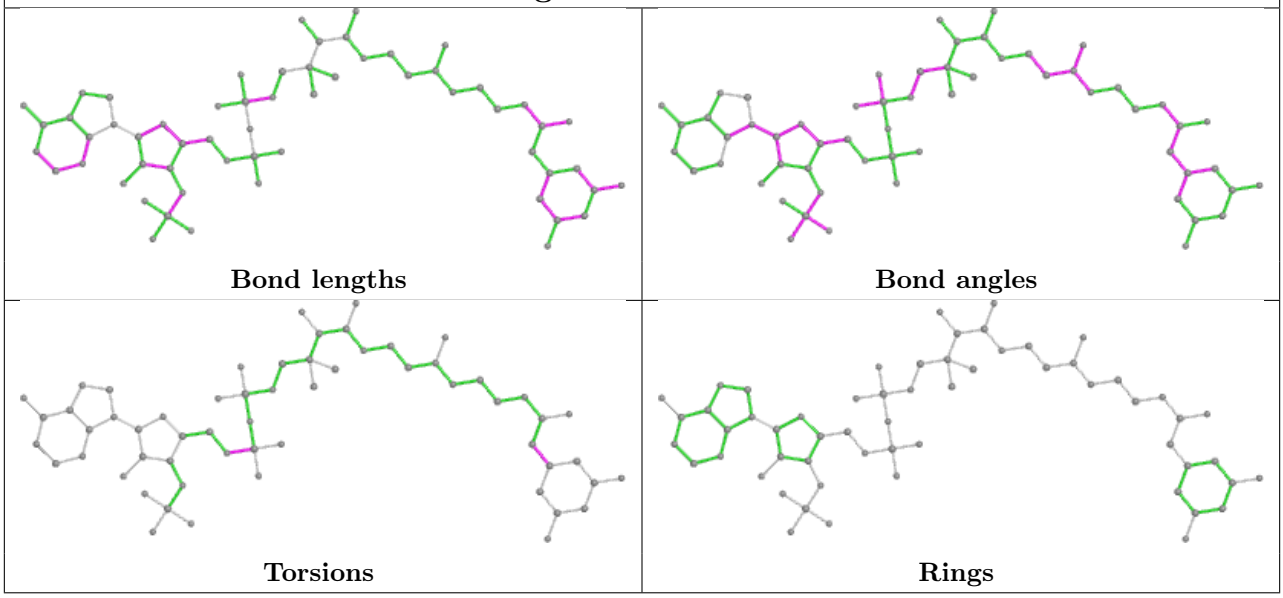


Torsions

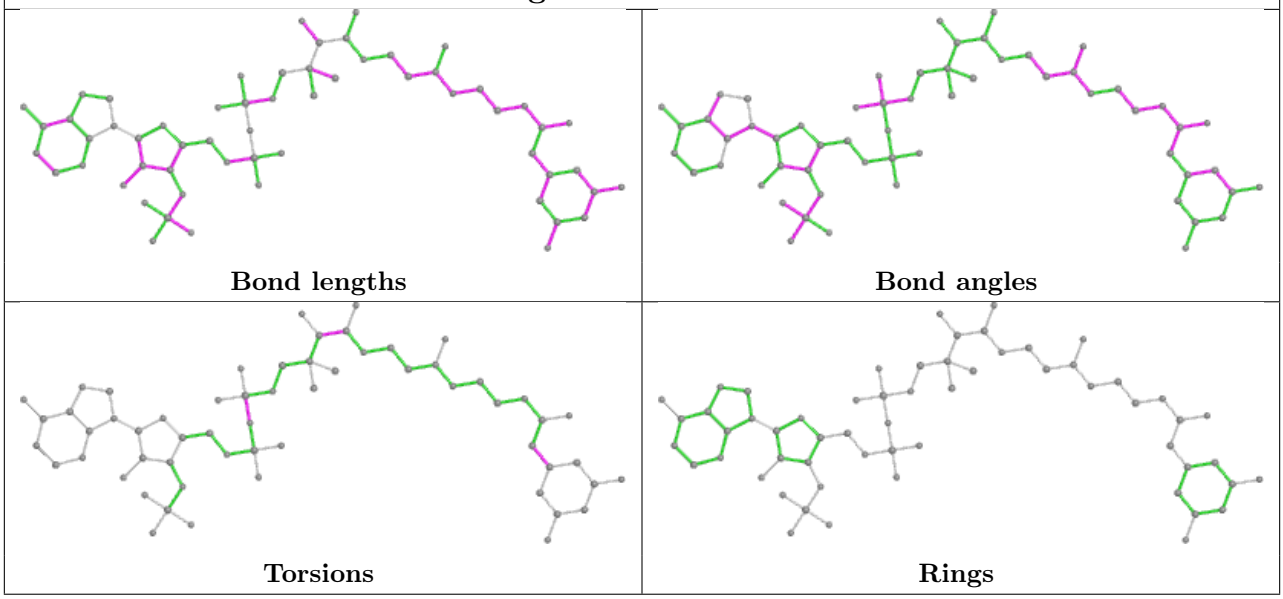


Rings

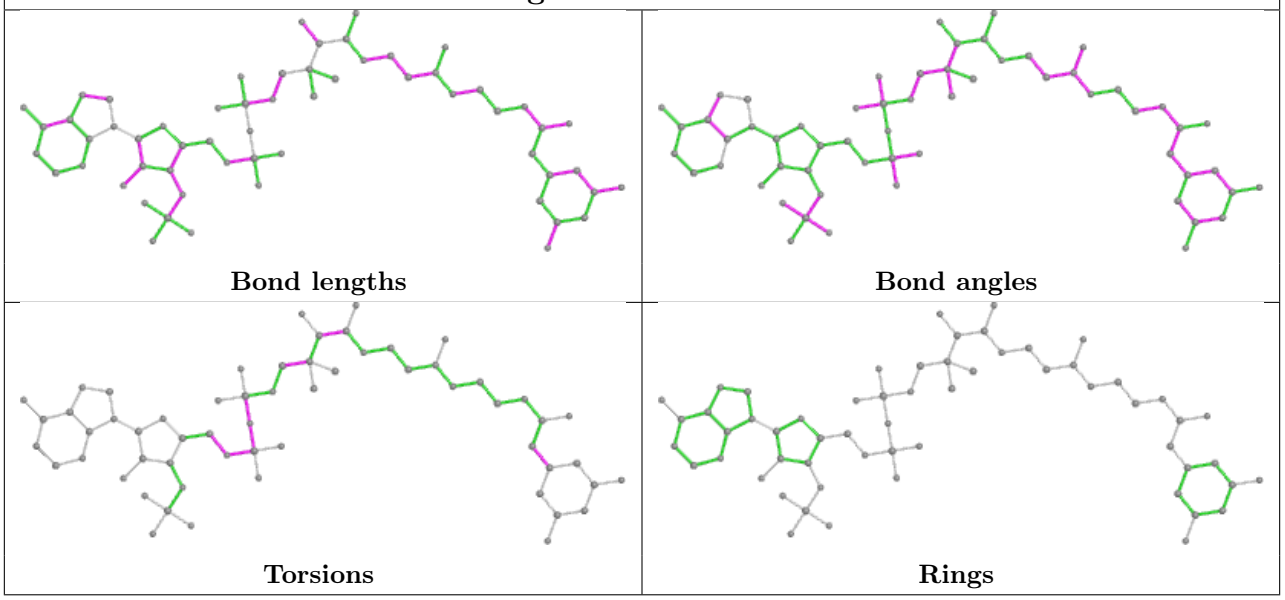
Ligand YE1 I 502



Ligand YE1 C 502



Ligand YE1 D 502



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	421/440 (95%)	-0.09	5 (1%) 79 77	16, 39, 80, 111	0
1	B	422/440 (95%)	-0.23	1 (0%) 95 95	19, 38, 62, 77	0
1	C	421/440 (95%)	-0.20	0 100 100	17, 37, 67, 90	0
1	D	423/440 (96%)	-0.28	1 (0%) 95 95	16, 32, 52, 68	0
1	E	421/440 (95%)	-0.23	1 (0%) 95 95	19, 35, 59, 76	0
1	F	421/440 (95%)	-0.24	0 100 100	16, 34, 56, 74	0
1	G	422/440 (95%)	-0.25	0 100 100	16, 32, 62, 83	0
1	H	422/440 (95%)	-0.14	1 (0%) 95 95	17, 41, 65, 86	0
1	I	421/440 (95%)	-0.24	0 100 100	15, 34, 59, 74	0
1	J	422/440 (95%)	-0.03	7 (1%) 70 67	25, 44, 78, 108	0
1	K	422/440 (95%)	-0.24	0 100 100	13, 33, 59, 77	0
1	L	421/440 (95%)	-0.15	1 (0%) 95 95	18, 38, 72, 103	0
All	All	5059/5280 (95%)	-0.19	17 (0%) 94 94	13, 36, 65, 111	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	427	ASP	4.0
1	J	37	ALA	3.7
1	A	37	ALA	3.6
1	J	430	GLY	3.0
1	J	29	LEU	3.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

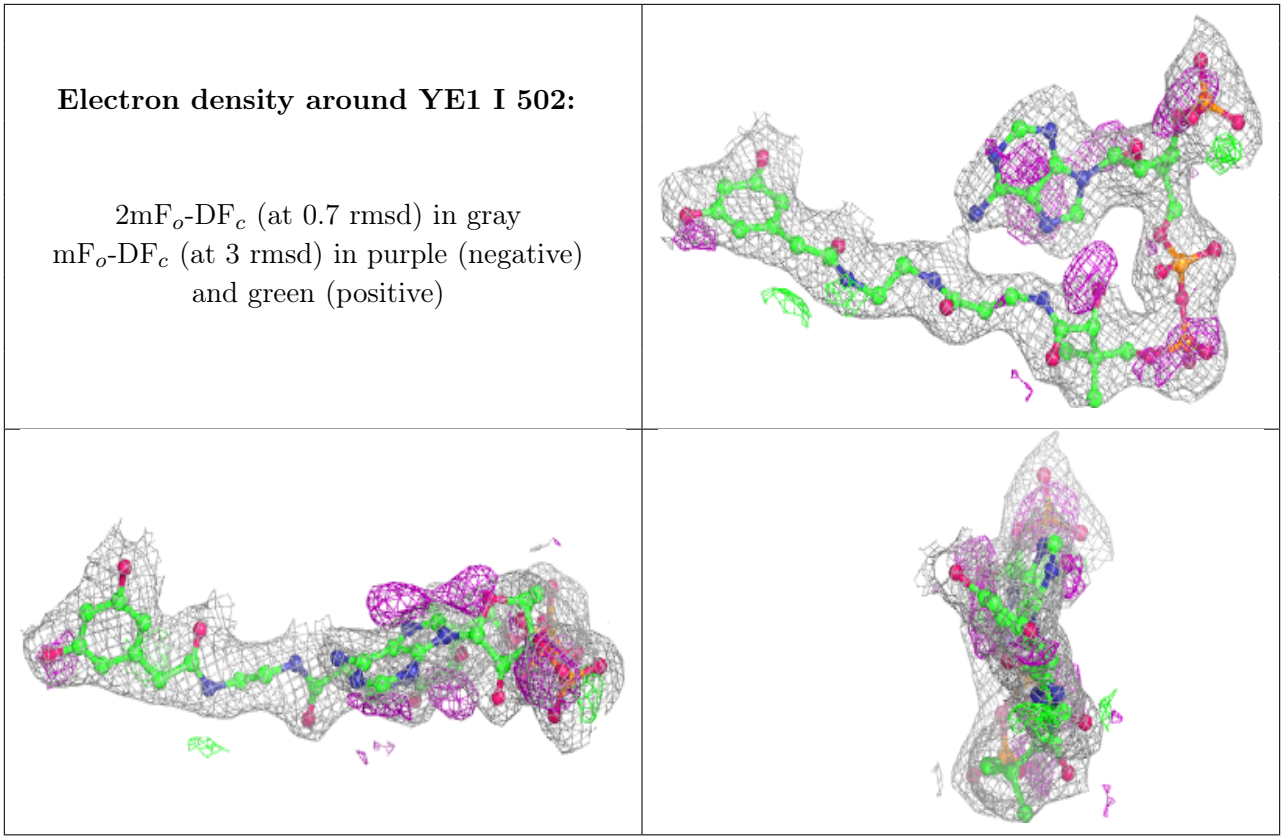
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	YE1	I	502	59/59	0.95	0.16	17,37,47,50	0
3	YE1	J	502	59/59	0.96	0.13	28,49,68,75	0
3	YE1	E	502	59/59	0.97	0.16	27,45,59,63	0
3	YE1	H	502	59/59	0.97	0.15	19,42,67,73	0
3	YE1	A	502	59/59	0.97	0.14	29,46,64,73	0
3	YE1	B	502	59/59	0.97	0.13	14,43,55,57	0
3	YE1	L	502	59/59	0.97	0.14	23,37,49,63	0
3	YE1	F	502	59/59	0.98	0.14	23,35,44,53	0
3	YE1	G	502	59/59	0.98	0.14	19,36,54,58	0
2	XE	H	501[B]	1/1	0.98	0.06	47,47,47,47	1
3	YE1	C	502	59/59	0.98	0.12	13,32,49,55	0
3	YE1	D	502	59/59	0.98	0.12	15,28,43,47	0
3	YE1	K	502	59/59	0.98	0.14	17,37,47,50	0
2	XE	D	501[B]	1/1	0.98	0.10	24,24,24,24	1
2	XE	A	501[A]	1/1	0.99	0.10	58,58,58,58	1
2	XE	E	501[B]	1/1	0.99	0.11	55,55,55,55	1
2	XE	F	501[B]	1/1	0.99	0.06	39,39,39,39	1
2	XE	A	501[B]	1/1	0.99	0.10	47,47,47,47	1
2	XE	I	501[B]	1/1	0.99	0.06	37,37,37,37	1
2	XE	J	501[B]	1/1	0.99	0.07	46,46,46,46	1
2	XE	K	501[B]	1/1	0.99	0.08	63,63,63,63	1
2	XE	L	501[B]	1/1	0.99	0.07	101,101,101,101	1
2	XE	B	501[A]	1/1	0.99	0.12	27,27,27,27	1
2	XE	B	501[B]	1/1	0.99	0.12	43,43,43,43	1
2	XE	G	501[B]	1/1	1.00	0.11	36,36,36,36	1
2	XE	C	501[B]	1/1	1.00	0.07	48,48,48,48	1

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.

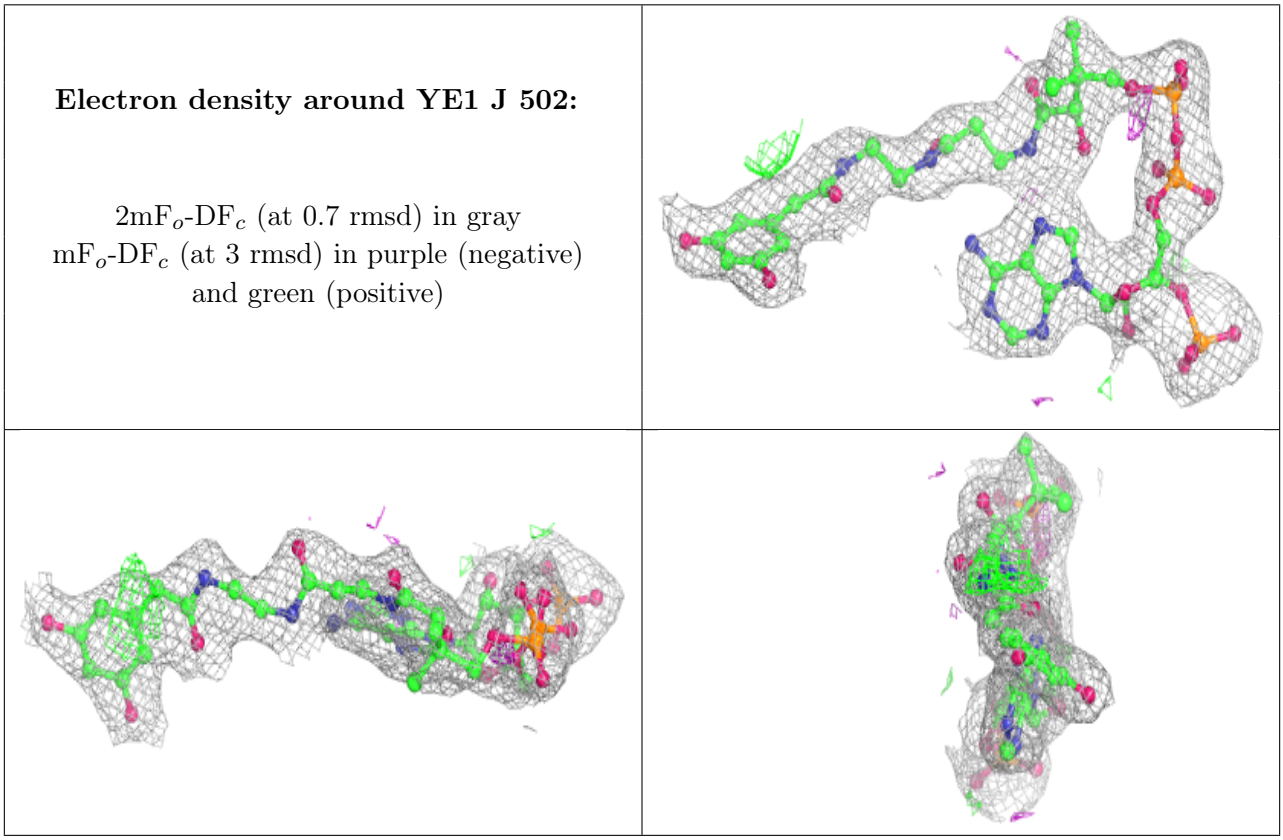
Electron density around YE1 I 502:

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



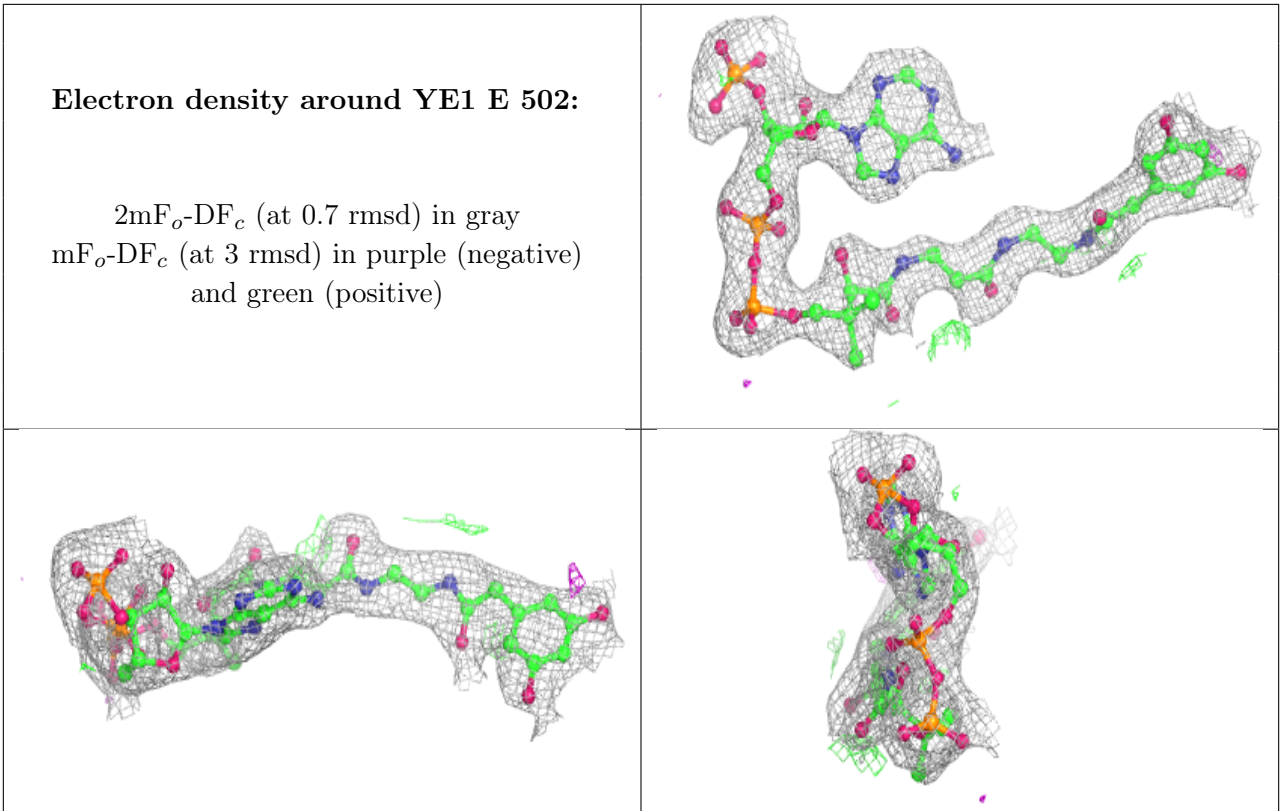
Electron density around YE1 J 502:

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



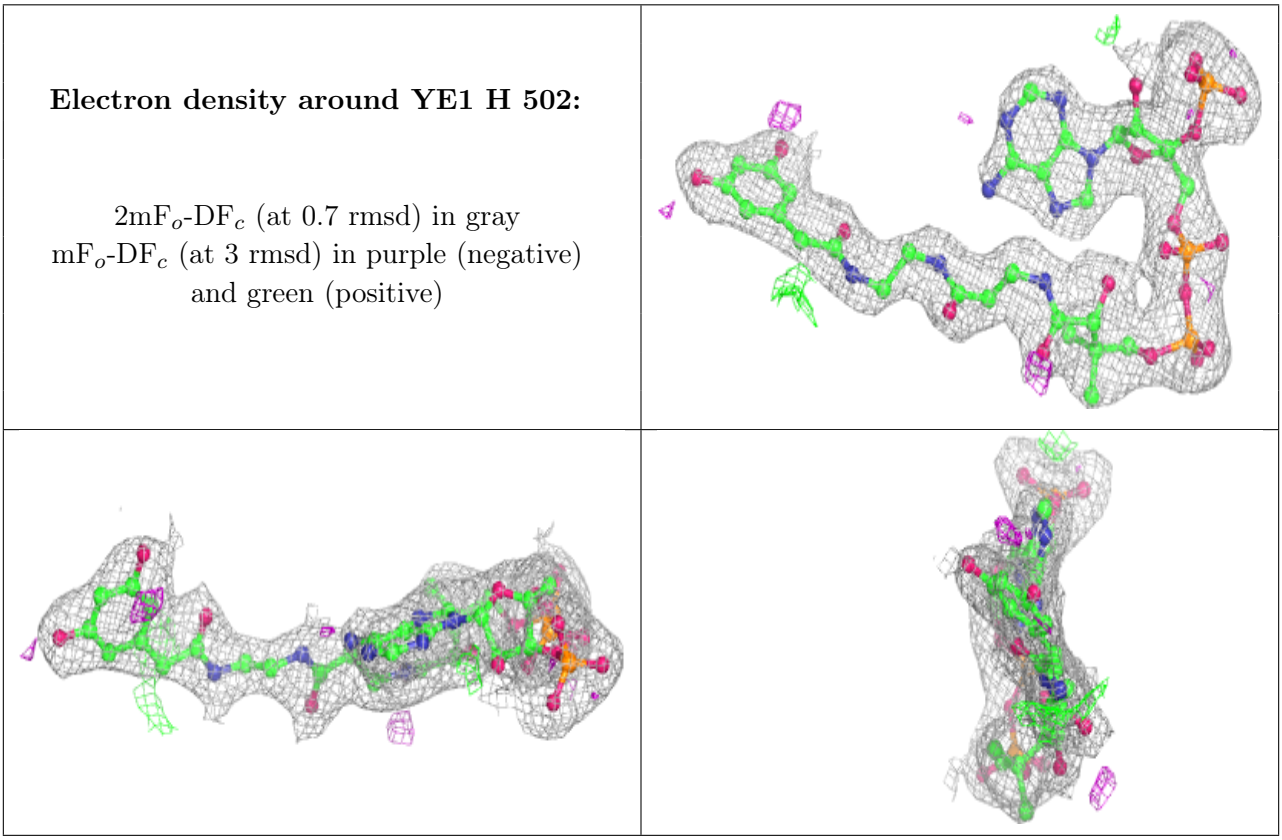
Electron density around YE1 E 502:

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



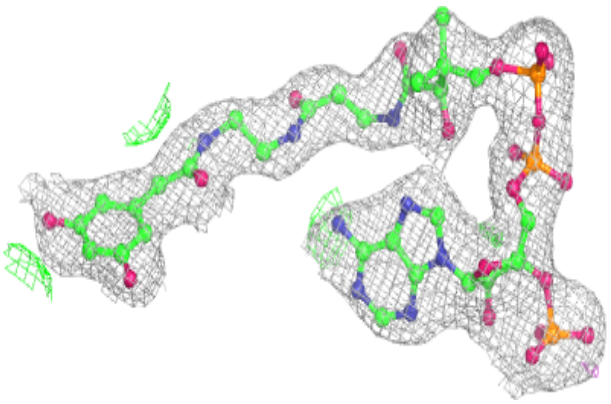
Electron density around YE1 H 502:

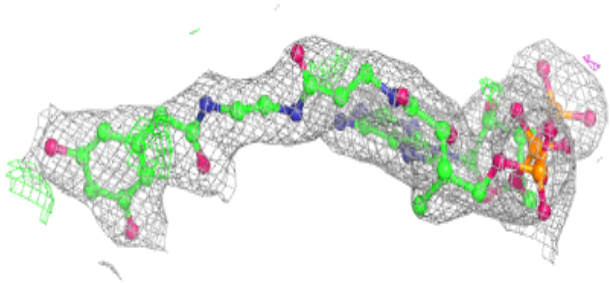
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

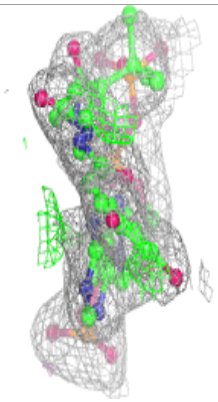


Electron density around YE1 A 502:

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

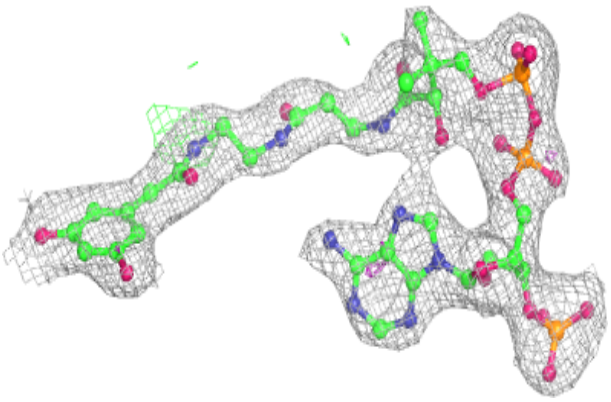


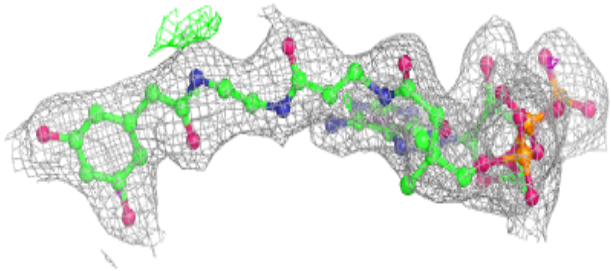


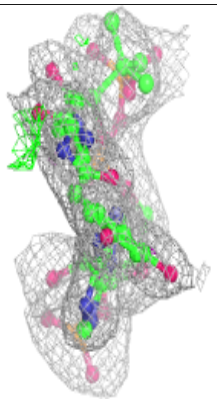


Electron density around YE1 B 502:

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

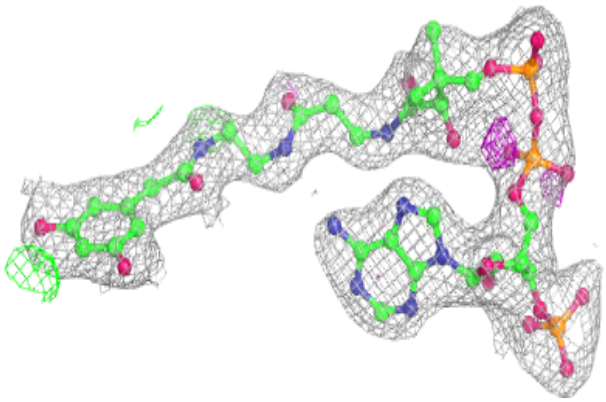


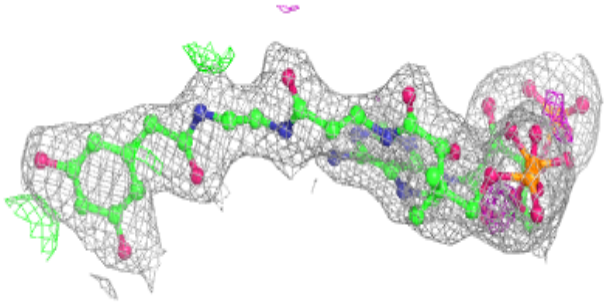


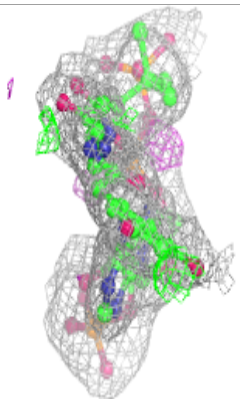


Electron density around YE1 L 502:

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

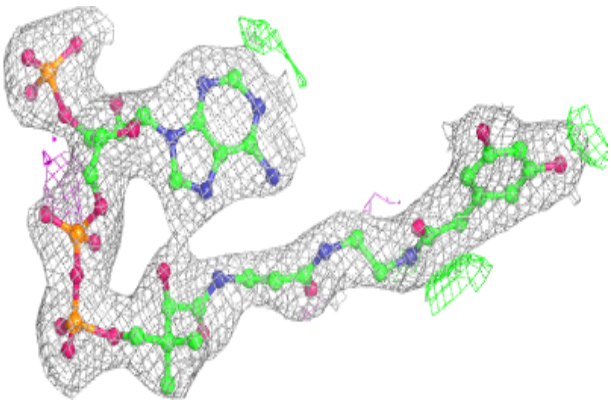


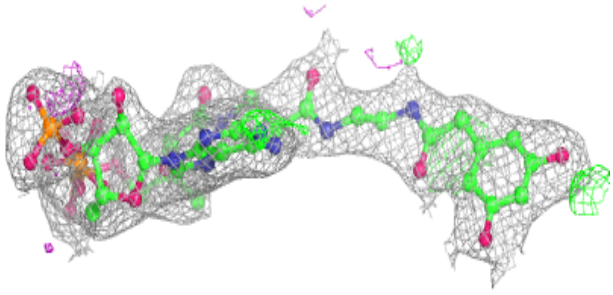


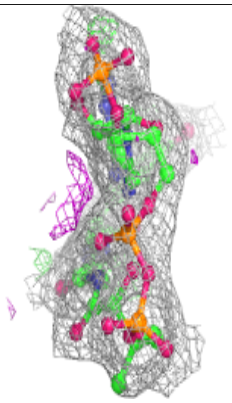


Electron density around YE1 F 502:

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

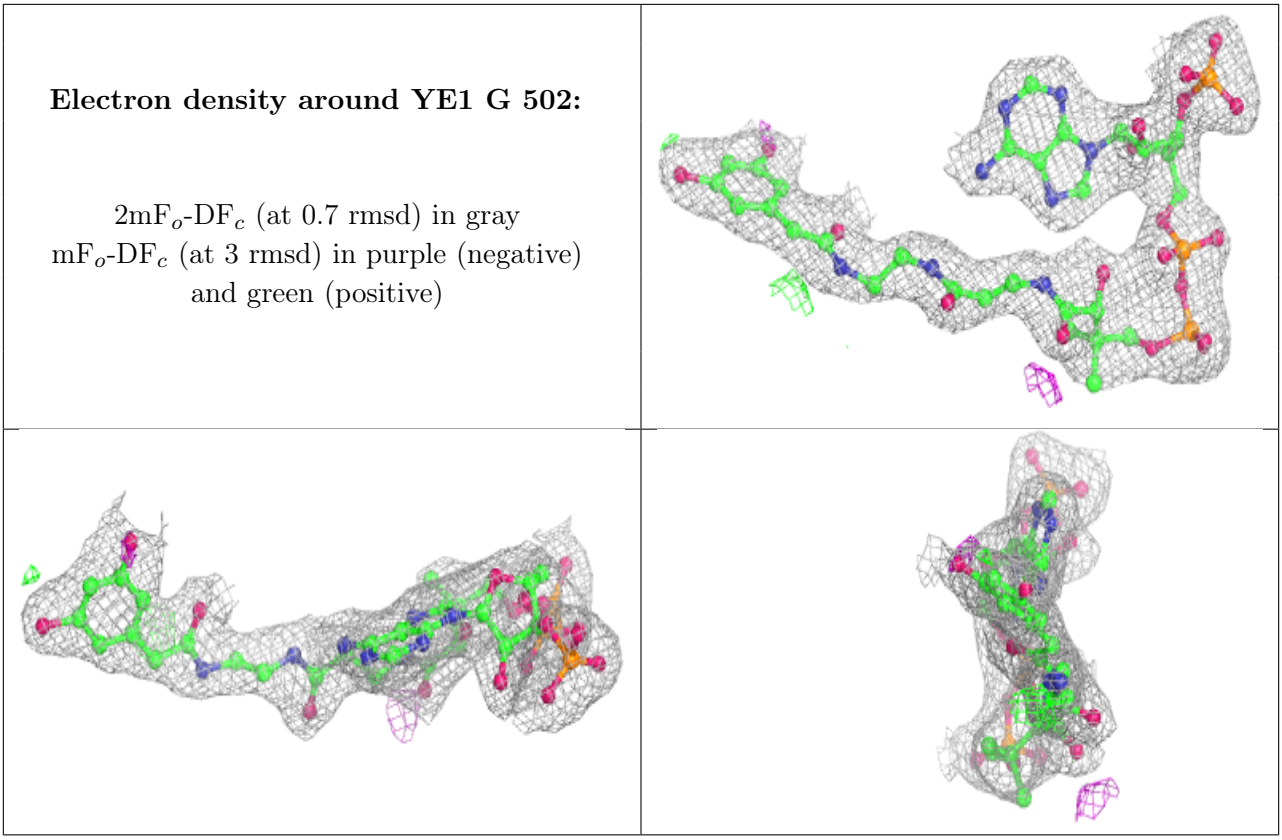






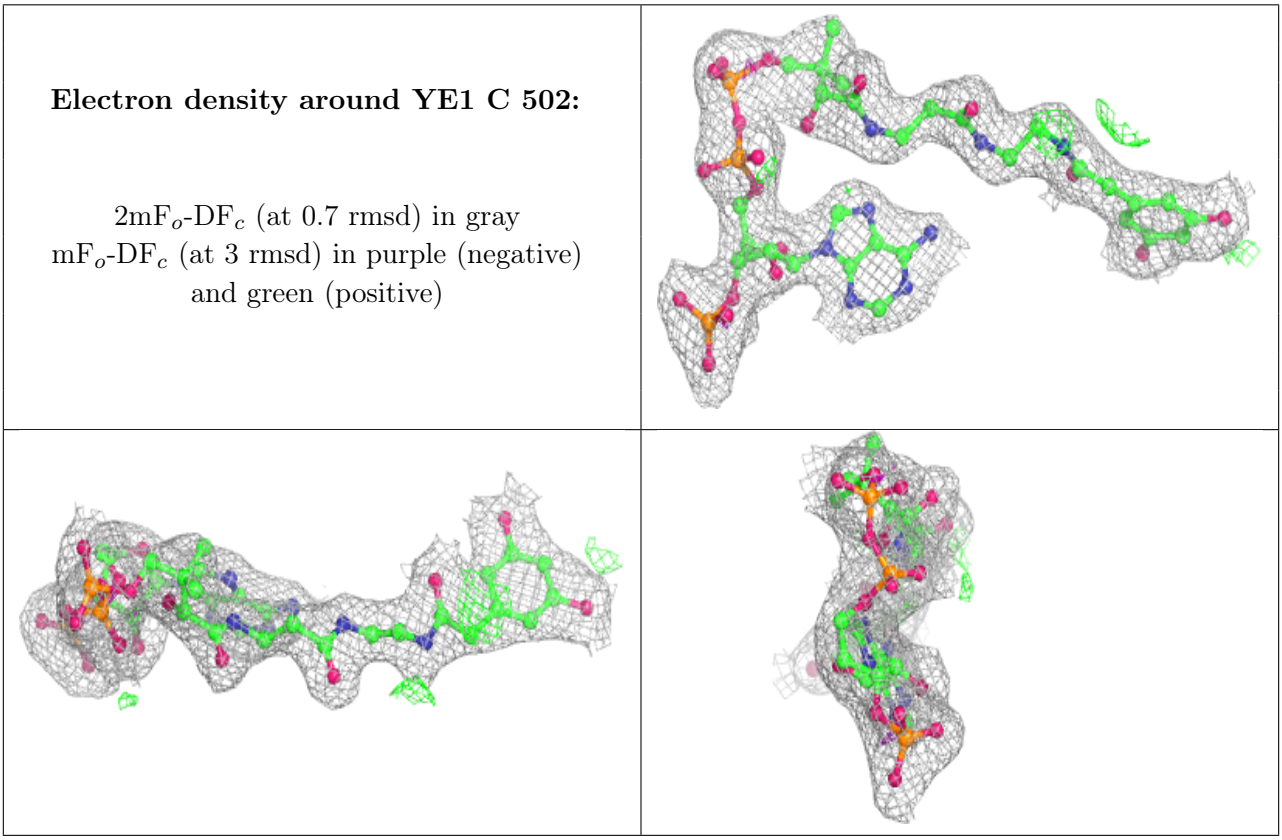
Electron density around YE1 G 502:

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



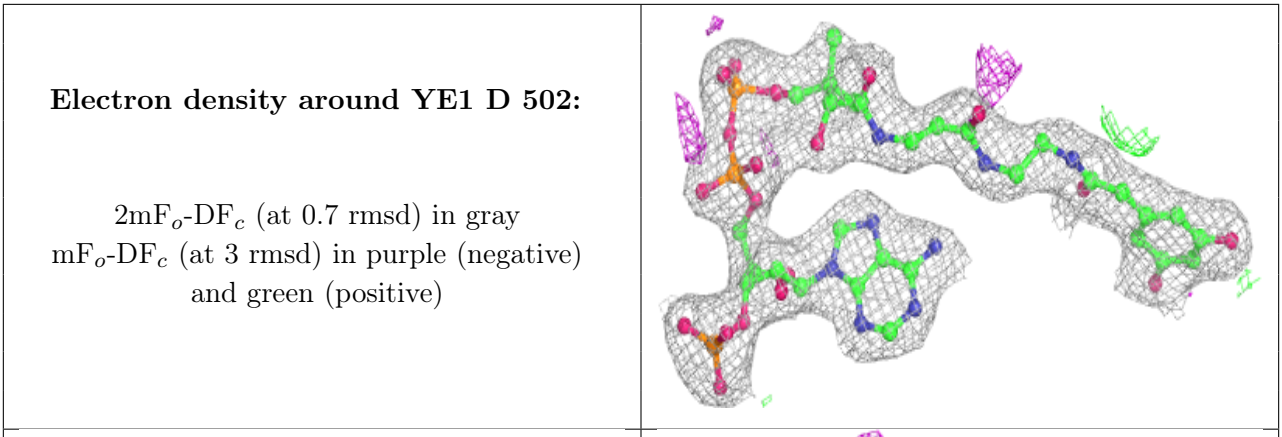
Electron density around YE1 C 502:

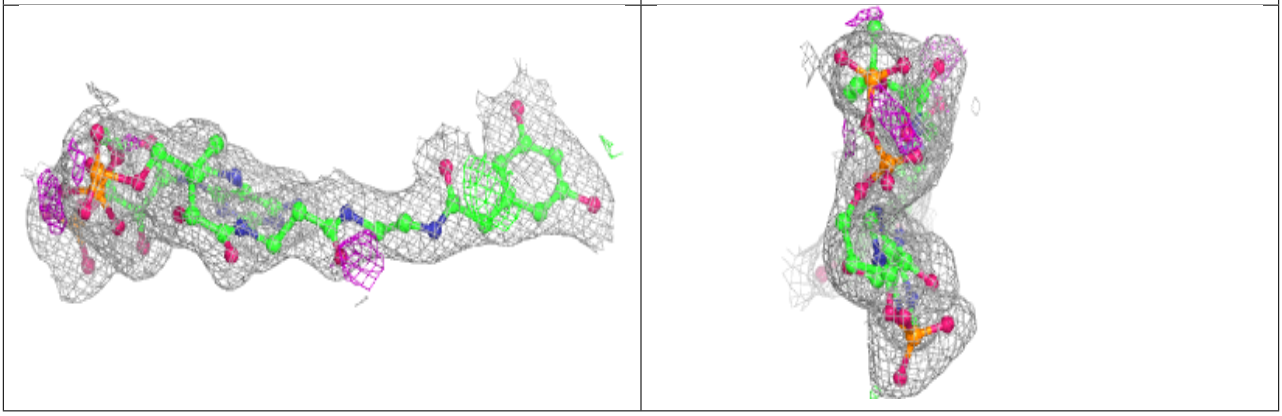
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around YE1 D 502:

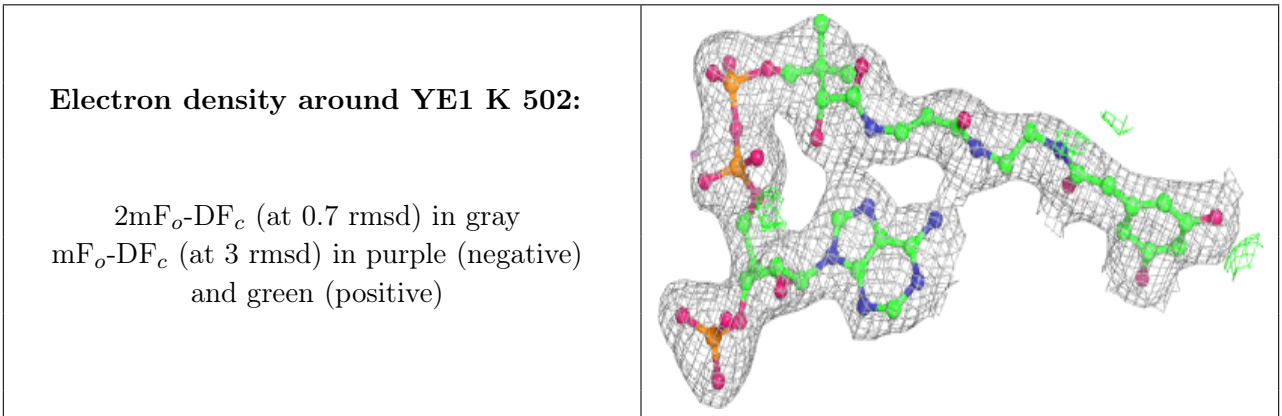
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

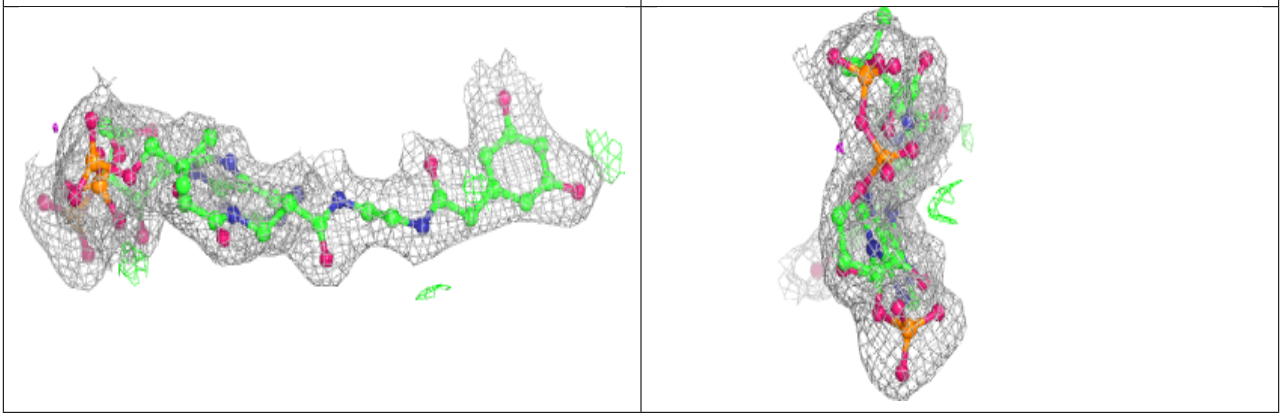




Electron density around YE1 K 502:

$2mF_o$-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 (i)

There are no such residues in this entry.