Full wwPDB X-ray Structure Validation Report (i)

Sep 25, 2023 – 07:59 AM EDT

PDB ID	:	5WI5
Title	:	2.0 Angstrom Resolution Crystal Structure of UDP-N-acetylglucosamine 1-carboxyvinyltransferase from Streptococcus pneumoniae in Complex with Uridine-diphosphate-2(n-acetylglucosaminyl) butyric acid, (2R)-2-(phosphonooxy) propanoic acid and Magnesium.
Authors	:	Minasov, G.; Shuvalova, L.; Dubrovska, I.; Kiryukhina, O.; Grimshaw, S.; Kwon, K.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2017-07-18
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

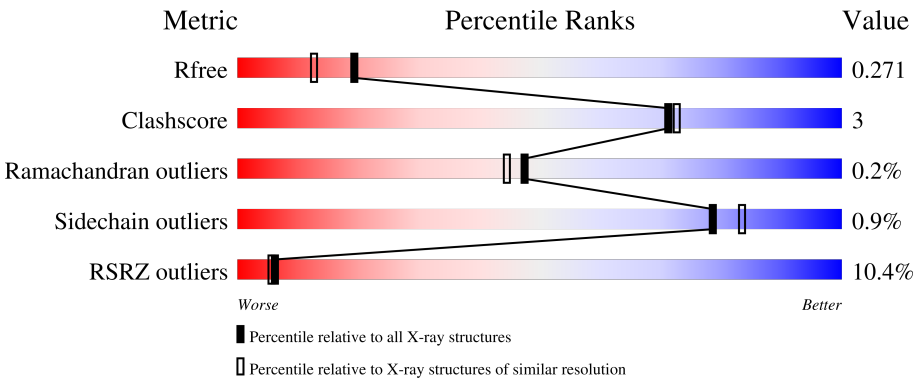
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	430	.% 91% 7% •
1	B	430	4% 90% 8% •
1	C	430	23% 87% 11% • •

Continued on next page...

Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	430	12% 91% 7%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13657 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.

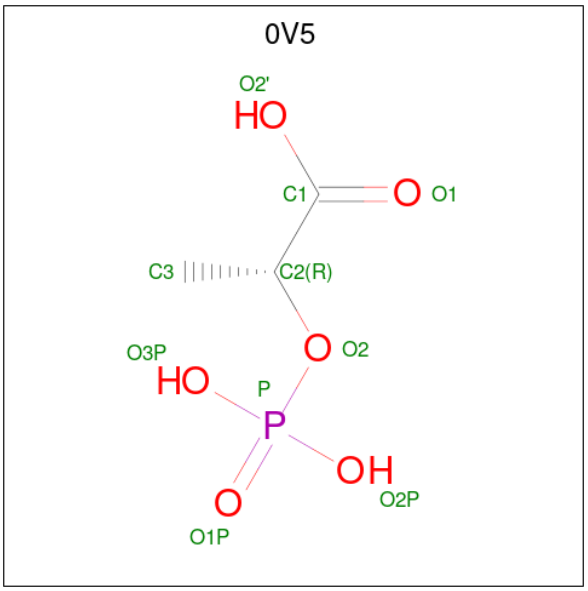
- Molecule 1 is a protein called UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	422	3242	2040	568	615	19	0	8	0
1	B	422	3192	2009	560	605	18	0	2	0
1	C	421	3178	2001	558	601	18	0	1	0
1	D	422	3173	1998	556	601	18	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q97NQ4
A	-1	ASN	-	expression tag	UNP Q97NQ4
A	0	ALA	-	expression tag	UNP Q97NQ4
B	-2	SER	-	expression tag	UNP Q97NQ4
B	-1	ASN	-	expression tag	UNP Q97NQ4
B	0	ALA	-	expression tag	UNP Q97NQ4
C	-2	SER	-	expression tag	UNP Q97NQ4
C	-1	ASN	-	expression tag	UNP Q97NQ4
C	0	ALA	-	expression tag	UNP Q97NQ4
D	-2	SER	-	expression tag	UNP Q97NQ4
D	-1	ASN	-	expression tag	UNP Q97NQ4
D	0	ALA	-	expression tag	UNP Q97NQ4

- Molecule 2 is (2R)-2-(phosphonooxy)propanoic acid (three-letter code: 0V5) (formula: $C_3H_7O_6P$) (labeled as "Ligand of Interest" by depositor).

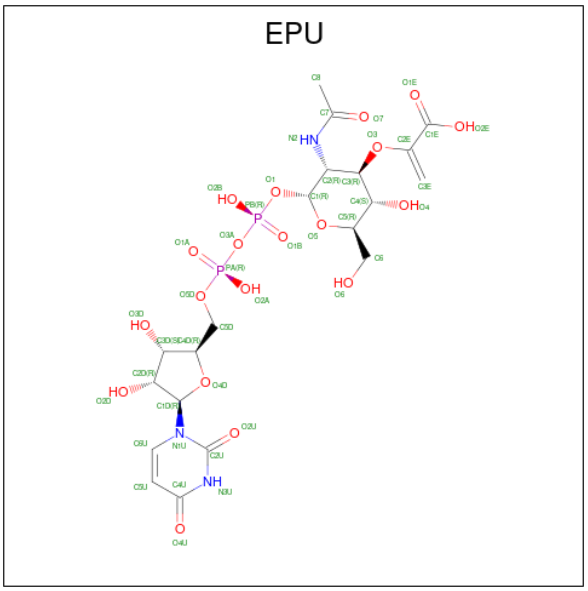


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
2	A	1	10	3	6	1	0	0
2	B	1	10	3	6	1	0	0
2	C	1	10	3	6	1	0	0
2	D	1	10	3	6	1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	A	1	1	1	0	0
3	B	1	1	1	0	0
3	C	1	1	1	0	0
3	D	1	1	1	0	0

- Molecule 4 is URIDINE-DIPHOSPHATE-2(N-ACETYLGLUCOSAMINYL) BUTYRIC ACID (three-letter code: EPU) (formula: $C_{20}H_{29}N_3O_{19}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	44	20	3	19	2	0	0
4	B	1	44	20	3	19	2	0	0
4	C	1	44	20	3	19	2	0	0
4	D	1	44	20	3	19	2	0	0

- Molecule 5 is water.

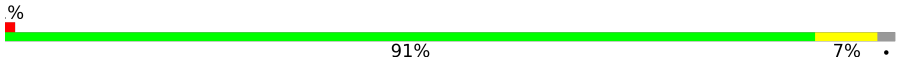
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	349	359	359	0	10
5	B	167	171	171	0	4
5	C	28	28	28	0	0
5	D	93	94	94	0	1

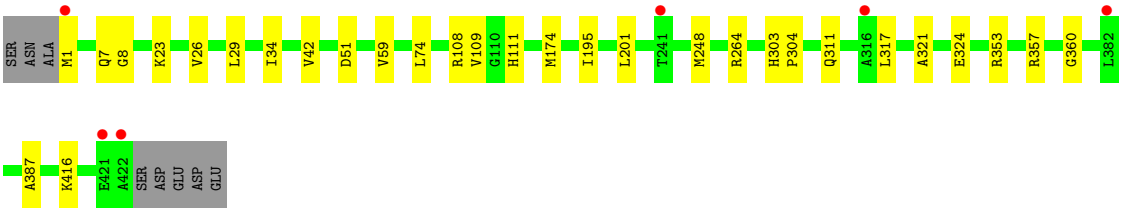
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1

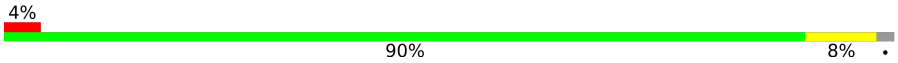
Chain A:

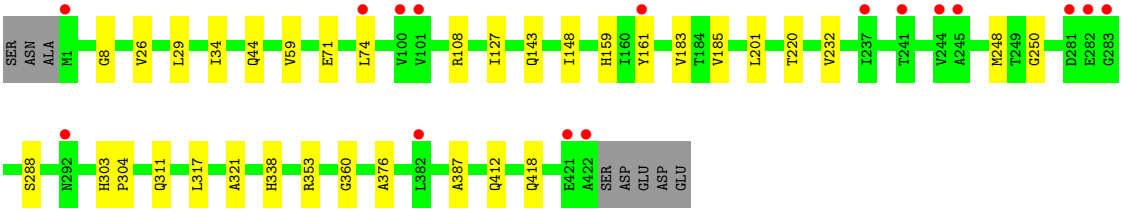




• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1

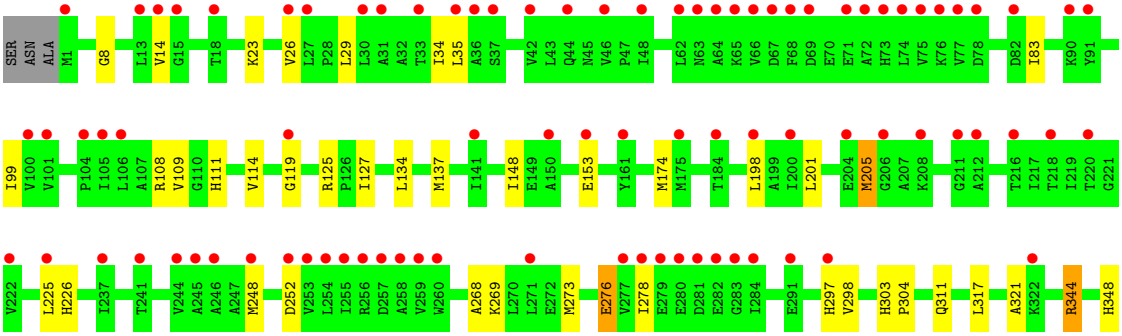
Chain B:

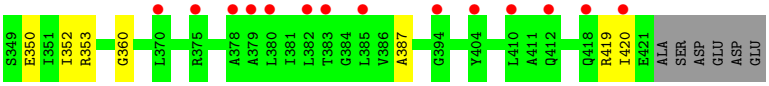




• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1

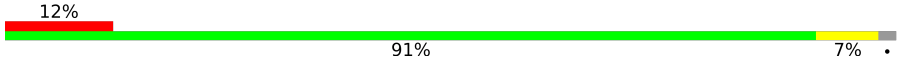
Chain C:

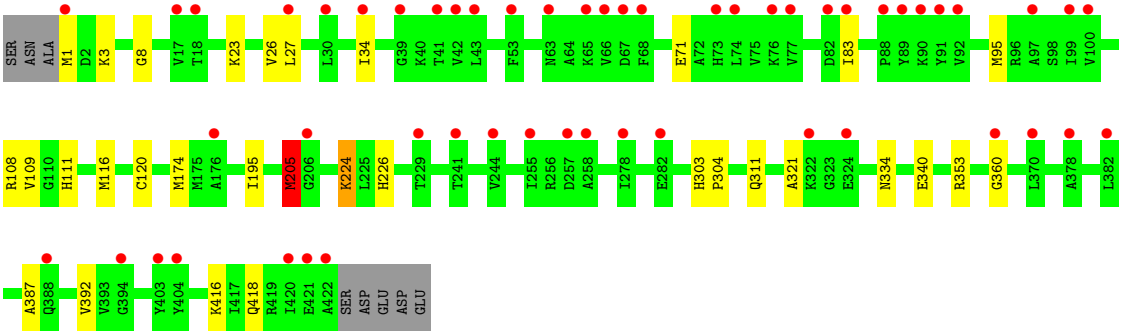




• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	84.24Å 80.13Å 126.51Å 90.00° 95.76° 90.00°	Depositor
Resolution (Å)	29.29 – 2.00 29.29 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (29.29-2.00) 99.8 (29.29-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.11 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.222 , 0.266 0.228 , 0.271	Depositor DCC
R_{free} test set	5833 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	33.0	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 51.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13657	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: EPU, MG, 0V5

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	0.59	0/3286	0.76	0/4446
1	B	0.50	0/3236	0.73	0/4379
1	C	0.45	0/3222	0.74	5/4360 (0.1%)
1	D	0.46	0/3216	0.73	2/4352 (0.0%)
All	All	0.50	0/12960	0.74	7/17537 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	205	MET	CG-SD-CE	-7.99	87.42	100.20
1	C	419	ARG	NE-CZ-NH1	6.71	123.65	120.30
1	D	205	MET	CA-CB-CG	6.26	123.94	113.30
1	C	276	GLU	CA-CB-CG	5.65	125.83	113.40
1	C	344	ARG	NE-CZ-NH2	5.60	123.10	120.30
1	C	205	MET	CA-CB-CG	5.45	122.57	113.30
1	C	273	MET	CG-SD-CE	-5.09	92.05	100.20

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	3242	0	3315	17	0
1	B	3192	0	3270	20	0
1	C	3178	0	3260	39	0
1	D	3173	0	3259	21	0
2	A	10	0	4	0	0
2	B	10	0	4	0	0
2	C	10	0	5	0	0
2	D	10	0	4	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	44	0	26	0	0
4	B	44	0	26	0	0
4	C	44	0	26	1	0
4	D	44	0	26	1	0
5	A	359	0	0	2	0
5	B	171	0	0	1	0
5	C	28	0	0	0	0
5	D	94	0	0	3	0
All	All	13657	0	13225	91	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 3.

All (91) 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:B:250:GLY:HA2	1:B:288:SER:OG	1.78	0.84
1:C:276:GLU:HG2	1:C:278:ILE:HD11	1.61	0.82
1:C:276:GLU:CG	1:C:278:ILE:HD11	2.12	0.80
1:C:14:VAL:HG22	1:C:252:ASP:OD1	1.91	0.70
1:C:269:LYS:HE3	1:C:297:HIS:O	1.94	0.67
1:C:119:GLY:HA2	1:D:340:GLU:HG2	1.79	0.65
1:D:205:MET:HG3	1:D:226:HIS:O	1.98	0.64
1:C:278:ILE:HD12	1:C:278:ILE:N	2.13	0.63
1:B:26:VAL:HG23	1:B:232:VAL:HG11	1.79	0.63
1:C:269:LYS:HG2	1:C:298:VAL:HG22	1.83	0.61
1:C:34:ILE:HG23	1:C:83:ILE:HD13	1.84	0.59
1:C:99:ILE:HG22	1:C:99:ILE:O	2.01	0.59
1:B:161:TYR:CE1	1:C:268:ALA:HB1	2.37	0.58
1:C:99:ILE:HD12	1:C:114:VAL:HG12	1.86	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:269:LYS:CE	1:C:297:HIS:O	2.51	0.58
1:C:174:MET:CE	1:C:198:LEU:HD23	2.34	0.57
1:B:26:VAL:HG22	1:B:232:VAL:HG21	1.85	0.57
1:D:3:LYS:HE3	1:D:392:VAL:HG12	1.85	0.57
1:B:321:ALA:O	1:B:360:GLY:HA3	2.05	0.56
1:B:248:MET:HG2	1:B:317:LEU:HA	1.88	0.56
1:D:34:ILE:HG23	1:D:83:ILE:HD13	1.87	0.56
1:C:420:ILE:HD12	1:C:420:ILE:N	2.21	0.56
1:C:276:GLU:HG3	1:C:278:ILE:HD11	1.89	0.55
1:C:321:ALA:O	1:C:360:GLY:HA3	2.06	0.55
1:B:159:HIS:HA	1:B:185:VAL:HG23	1.89	0.55
1:D:418:GLN:HA	5:D:629:HOH:O	2.07	0.55
1:A:321:ALA:O	1:A:360:GLY:HA3	2.06	0.55
1:C:134:LEU:HD23	1:C:137:MET:CE	2.37	0.55
1:A:248:MET:HG2	1:A:317:LEU:HA	1.89	0.54
1:D:3:LYS:HE3	1:D:392:VAL:CG1	2.38	0.54
1:B:183:VAL:HG22	1:B:220:THR:HG22	1.89	0.53
1:C:174:MET:HE1	1:C:198:LEU:HD23	1.90	0.52
1:D:321:ALA:O	1:D:360:GLY:HA3	2.09	0.52
1:C:248:MET:HG2	1:C:317:LEU:HA	1.92	0.52
1:C:109:VAL:HG12	1:C:111:HIS:CD2	2.45	0.52
1:B:412:GLN:NE2	5:B:611:HOH:O	2.43	0.51
1:D:109:VAL:HG12	1:D:111:HIS:CD2	2.45	0.51
1:A:109:VAL:HG12	1:A:111:HIS:CD2	2.46	0.50
1:D:174:MET:HE1	1:D:195:ILE:HG23	1.93	0.50
1:C:344:ARG:NH2	1:C:344:ARG:HB3	2.27	0.50
1:C:276:GLU:HG2	1:C:278:ILE:CD1	2.37	0.50
1:C:205:MET:HG3	1:C:226:HIS:O	2.12	0.49
1:B:159:HIS:HA	1:B:185:VAL:CG2	2.44	0.48
1:D:8:GLY:HA3	1:D:387:ALA:O	2.15	0.47
1:C:8:GLY:HA3	1:C:387:ALA:O	2.15	0.47
1:C:99:ILE:HD12	1:C:114:VAL:CG1	2.44	0.47
1:A:8:GLY:HA3	1:A:387:ALA:O	2.15	0.46
1:A:34:ILE:HD11	1:A:59:VAL:HG13	1.96	0.46
1:B:8:GLY:HA3	1:B:387:ALA:O	2.15	0.46
1:C:303:HIS:CG	1:C:304:PRO:HA	2.51	0.46
1:B:71:GLU:H	1:B:71:GLU:CD	2.18	0.46
1:D:27:LEU:HD13	1:D:95:MET:CE	2.46	0.46
1:B:34:ILE:HD11	1:B:59:VAL:HG13	1.98	0.45
1:D:23:LYS:HE2	4:D:503:EPU:O1E	2.16	0.45
1:B:303:HIS:CG	1:B:304:PRO:HA	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:303:HIS:CG	1:A:304:PRO:HA	2.52	0.45
1:A:357:ARG:HE	1:C:352:ILE:HD11	1.81	0.45
1:C:134:LEU:HD23	1:C:137:MET:HE3	1.99	0.45
1:D:120:CYS:HA	2:D:501:0V5:C1	2.46	0.45
1:D:71:GLU:OE2	5:D:601:HOH:O	2.21	0.45
1:C:127:ILE:HD11	1:C:148:ILE:HD11	1.99	0.44
1:C:344:ARG:NE	1:D:116:MET:O	2.49	0.44
1:D:303:HIS:CG	1:D:304:PRO:HA	2.52	0.44
1:A:23:LYS:HE2	1:A:51:ASP:OD2	2.18	0.44
1:C:119:GLY:HA2	1:D:340:GLU:CG	2.48	0.44
1:D:34:ILE:O	1:D:108:ARG:HG3	2.18	0.44
1:A:174[B]:MET:HE1	1:A:195:ILE:HG23	1.99	0.43
1:D:224:LYS:HD2	1:D:224:LYS:O	2.18	0.43
1:C:34:ILE:O	1:C:108:ARG:HG3	2.17	0.43
1:D:334:ASN:ND2	5:D:609:HOH:O	2.52	0.43
1:A:34:ILE:O	1:A:108:ARG:HG3	2.19	0.43
1:B:127:ILE:HD11	1:B:148:ILE:HD11	2.00	0.42
1:C:29:LEU:HD22	1:C:201:LEU:HD22	2.00	0.42
1:A:42:VAL:CG1	1:A:74[B]:LEU:HD11	2.49	0.42
1:C:348[B]:HIS:NE2	1:C:350:GLU:OE2	2.46	0.42
1:B:44:GLN:HG2	1:B:74:LEU:CD2	2.49	0.42
1:C:23:LYS:O	1:C:26:VAL:HG12	2.20	0.42
1:C:35:LEU:HB3	1:C:225:LEU:HD12	2.00	0.42
1:C:125:ARG:HA	4:C:503:EPU:O2D	2.20	0.42
1:A:1:MET:N	5:A:611:HOH:O	2.45	0.41
1:A:357:ARG:NE	1:C:352:ILE:HD11	2.34	0.41
1:A:29:LEU:HD22	1:A:201:LEU:HD22	2.02	0.41
1:A:23:LYS:O	1:A:26:VAL:HG12	2.21	0.41
1:B:34:ILE:O	1:B:108:ARG:HG3	2.21	0.41
1:B:338:HIS:HB3	1:B:376:ALA:HB1	2.03	0.41
1:A:7[A]:GLN:HB2	1:A:416:LYS:HD2	2.02	0.41
1:C:14:VAL:HG13	1:C:252:ASP:OD1	2.21	0.41
1:B:250:GLY:CA	1:B:288:SER:OG	2.60	0.41
1:A:324:GLU:OE2	5:A:601:HOH:O	2.22	0.40
1:D:23:LYS:O	1:D:26:VAL:HG12	2.21	0.40
1:B:29:LEU:HD22	1:B:201:LEU:HD22	2.02	0.40

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	428/430 (100%)	415 (97%)	12 (3%)	1 (0%)	47	44
1	B	422/430 (98%)	408 (97%)	13 (3%)	1 (0%)	47	44
1	C	420/430 (98%)	408 (97%)	11 (3%)	1 (0%)	47	44
1	D	420/430 (98%)	408 (97%)	11 (3%)	1 (0%)	47	44
All	All	1690/1720 (98%)	1639 (97%)	47 (3%)	4 (0%)	47	44

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	353	ARG
1	B	353	ARG
1	A	353	ARG
1	D	353	ARG

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	344/343 (100%)	342 (99%)	2 (1%)	86	90
1	B	338/343 (98%)	335 (99%)	3 (1%)	78	83
1	C	337/343 (98%)	335 (99%)	2 (1%)	86	90
1	D	336/343 (98%)	331 (98%)	5 (2%)	65	69
All	All	1355/1372 (99%)	1343 (99%)	12 (1%)	78	83

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	264	ARG
1	A	311	GLN
1	B	143	GLN
1	B	311	GLN
1	B	418	GLN
1	C	153	GLU
1	C	311	GLN
1	D	1	MET
1	D	205	MET
1	D	224	LYS
1	D	311	GLN
1	D	416	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	418	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 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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	$\#\|Z\| > 2$
4	EPU	A	503	-	43,46,46	1.18	4 (9%)	61,69,69	1.72	10 (16%)
4	EPU	B	503	-	43,46,46	1.22	5 (11%)	61,69,69	1.51	8 (13%)
2	0V5	A	501	1,3	8,9,9	1.00	0	11,13,13	1.26	1 (9%)
2	0V5	B	501	1,3	8,9,9	1.11	1 (12%)	11,13,13	0.92	0
4	EPU	C	503	-	43,46,46	1.09	4 (9%)	61,69,69	1.48	9 (14%)
2	0V5	C	501	1,3	8,9,9	1.14	1 (12%)	11,13,13	1.00	0
2	0V5	D	501	1,3	8,9,9	0.87	0	11,13,13	1.13	1 (9%)
4	EPU	D	503	-	43,46,46	1.11	3 (6%)	61,69,69	1.68	15 (24%)

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	EPU	A	503	-	-	8/32/71/71	0/3/3/3
4	EPU	B	503	-	-	8/32/71/71	0/3/3/3
2	0V5	A	501	1,3	-	5/9/9/9	-
2	0V5	B	501	1,3	-	1/9/9/9	-
4	EPU	C	503	-	-	5/32/71/71	0/3/3/3
2	0V5	C	501	1,3	-	5/9/9/9	-
2	0V5	D	501	1,3	-	0/9/9/9	-
4	EPU	D	503	-	-	7/32/71/71	0/3/3/3

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	503	EPU	C2E-C1E	-3.87	1.45	1.49
4	B	503	EPU	C2E-C1E	-3.55	1.45	1.49
4	D	503	EPU	O3-C2E	3.51	1.48	1.36
4	C	503	EPU	O3-C2E	3.37	1.47	1.36
4	B	503	EPU	O3-C2E	3.21	1.47	1.36
4	D	503	EPU	C2U-N1U	2.98	1.43	1.38
4	A	503	EPU	O3-C2E	2.75	1.45	1.36
4	B	503	EPU	C5U-C4U	-2.63	1.37	1.43
4	B	503	EPU	C2U-N1U	2.58	1.42	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	503	EPU	C2U-N3U	-2.34	1.33	1.38
4	B	503	EPU	O2U-C2U	2.31	1.27	1.23
4	C	503	EPU	C2U-N1U	2.16	1.41	1.38
4	C	503	EPU	C4U-N3U	-2.15	1.34	1.38
4	D	503	EPU	C2U-N3U	-2.12	1.34	1.38
2	C	501	0V5	P-O2	2.11	1.63	1.59
2	B	501	0V5	O1-C1	2.08	1.28	1.22
4	A	503	EPU	C4U-N3U	-2.07	1.34	1.38
4	C	503	EPU	C5U-C4U	-2.01	1.39	1.43

All (44) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	503	EPU	O2E-C1E-C2E	5.92	124.01	113.91
4	A	503	EPU	O1E-C1E-C2E	-5.72	113.17	121.79
4	C	503	EPU	O2E-C1E-C2E	4.55	121.67	113.91
4	D	503	EPU	C4U-N3U-C2U	-4.27	120.95	126.58
4	B	503	EPU	C5U-C4U-N3U	4.12	121.00	114.84
4	B	503	EPU	O4U-C4U-C5U	-4.06	118.02	125.16
4	D	503	EPU	O2E-C1E-C2E	4.04	120.80	113.91
4	C	503	EPU	C5U-C4U-N3U	4.02	120.86	114.84
4	B	503	EPU	O2E-C1E-C2E	3.99	120.72	113.91
4	D	503	EPU	N3U-C2U-N1U	3.93	120.11	114.89
4	C	503	EPU	C4U-N3U-C2U	-3.89	121.45	126.58
4	B	503	EPU	C4U-N3U-C2U	-3.89	121.46	126.58
4	A	503	EPU	C5U-C4U-N3U	3.64	120.28	114.84
4	A	503	EPU	C4U-N3U-C2U	-3.48	122.00	126.58
4	D	503	EPU	O4U-C4U-C5U	-3.35	119.28	125.16
4	D	503	EPU	C5U-C4U-N3U	3.34	119.83	114.84
4	A	503	EPU	N3U-C2U-N1U	3.33	119.31	114.89
4	B	503	EPU	O1E-C1E-C2E	-3.32	116.79	121.79
4	C	503	EPU	N3U-C2U-N1U	3.16	119.08	114.89
4	C	503	EPU	O4U-C4U-C5U	-3.10	119.71	125.16
4	A	503	EPU	C2D-C1D-N1U	-2.88	105.06	113.22
4	A	503	EPU	PB-O3A-PA	-2.83	123.10	132.83
4	D	503	EPU	O3-C3-C2	2.63	112.84	107.91
4	B	503	EPU	N3U-C2U-N1U	2.51	118.22	114.89
4	C	503	EPU	PB-O3A-PA	-2.50	124.23	132.83
4	B	503	EPU	C3E-C2E-C1E	-2.50	118.11	122.73
4	A	503	EPU	O4U-C4U-C5U	-2.49	120.79	125.16
4	D	503	EPU	O2B-PB-O1B	2.47	124.43	112.24
4	D	503	EPU	O3A-PB-O1	-2.41	97.62	102.48

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	0V5	O2'-C1-C2	2.41	122.71	113.89
4	C	503	EPU	O3-C3-C4	2.39	113.16	107.70
4	B	503	EPU	O3-C3-C2	2.38	112.36	107.91
4	D	503	EPU	C3-C2-N2	2.34	114.79	110.91
2	D	501	0V5	C3-C2-C1	2.27	113.02	109.28
4	D	503	EPU	C1D-N1U-C2U	2.22	121.60	117.57
4	C	503	EPU	O1E-C1E-C2E	-2.22	118.45	121.79
4	D	503	EPU	PB-O3A-PA	-2.21	125.25	132.83
4	D	503	EPU	O4D-C1D-N1U	2.13	113.23	108.36
4	C	503	EPU	O2B-PB-O1B	2.07	122.47	112.24
4	D	503	EPU	C3E-C2E-C1E	-2.06	118.93	122.73
4	D	503	EPU	O3-C2E-C1E	2.05	120.69	115.59
4	D	503	EPU	O2A-PA-O5D	2.03	117.17	107.75
4	A	503	EPU	O3A-PB-O1	-2.02	98.42	102.48
4	A	503	EPU	O4D-C4D-C3D	2.01	109.09	105.11

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	0V5	C2-O2-P-O1P
4	A	503	EPU	C5D-O5D-PA-O1A
4	A	503	EPU	C5D-O5D-PA-O2A
4	A	503	EPU	O1E-C1E-C2E-C3E
4	A	503	EPU	O2E-C1E-C2E-O3
4	A	503	EPU	O2E-C1E-C2E-C3E
4	B	503	EPU	C5D-O5D-PA-O1A
4	B	503	EPU	C5D-O5D-PA-O2A
4	B	503	EPU	O1E-C1E-C2E-C3E
4	B	503	EPU	O2E-C1E-C2E-O3
4	B	503	EPU	O2E-C1E-C2E-C3E
4	C	503	EPU	C5D-O5D-PA-O2A
4	D	503	EPU	C5D-O5D-PA-O2A
4	D	503	EPU	O1E-C1E-C2E-C3E
4	D	503	EPU	O2E-C1E-C2E-O3
4	D	503	EPU	O2E-C1E-C2E-C3E
4	A	503	EPU	O1E-C1E-C2E-O3
4	B	503	EPU	O1E-C1E-C2E-O3
4	D	503	EPU	O1E-C1E-C2E-O3
2	A	501	0V5	C2-O2-P-O1P
4	A	503	EPU	C5D-O5D-PA-O3A
4	C	503	EPU	C5D-O5D-PA-O3A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	D	503	EPU	C5D-O5D-PA-O3A
4	C	503	EPU	PB-O3A-PA-O1A
2	A	501	0V5	O2'-C1-C2-C3
2	C	501	0V5	O1-C1-C2-C3
2	C	501	0V5	O2'-C1-C2-C3
2	C	501	0V5	O2'-C1-C2-O2
4	C	503	EPU	C5D-O5D-PA-O1A
4	B	503	EPU	O4D-C4D-C5D-O5D
4	A	503	EPU	PB-O3A-PA-O1A
2	C	501	0V5	C2-O2-P-O1P
2	A	501	0V5	C2-O2-P-O3P
4	B	503	EPU	C5D-O5D-PA-O3A
4	C	503	EPU	PB-O3A-PA-O2A
2	A	501	0V5	O1-C1-C2-C3
2	A	501	0V5	O2'-C1-C2-O2
2	C	501	0V5	O1-C1-C2-O2
4	D	503	EPU	C5D-O5D-PA-O1A

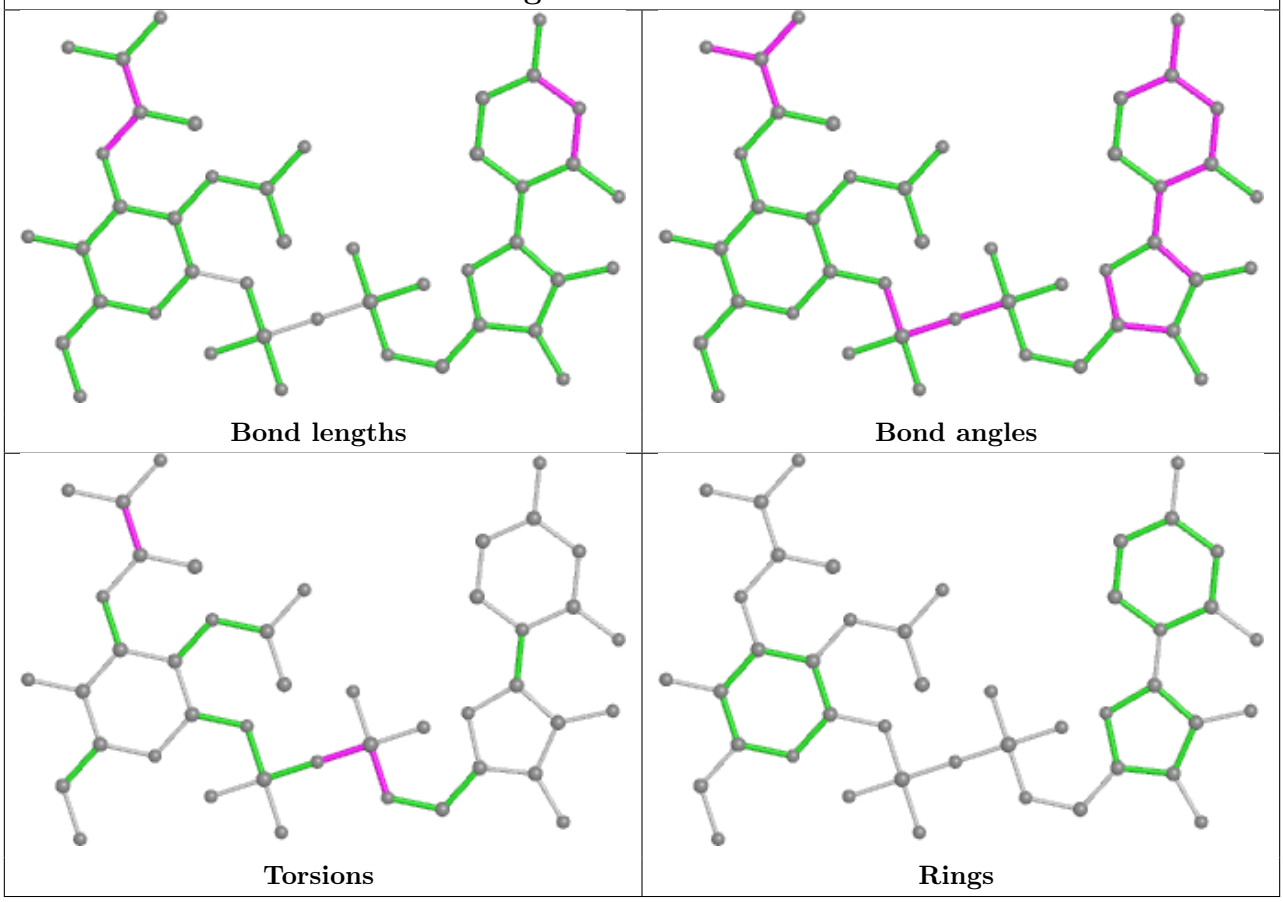
There are no ring outliers.

3 monomers are involved in 3 short contacts:

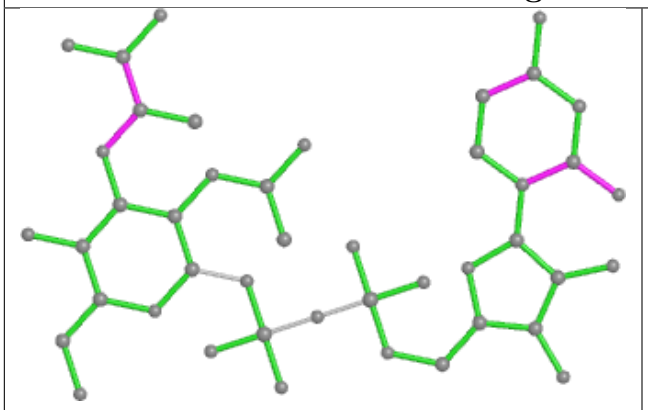
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	503	EPU	1	0
2	D	501	0V5	1	0
4	D	503	EPU	1	0

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.

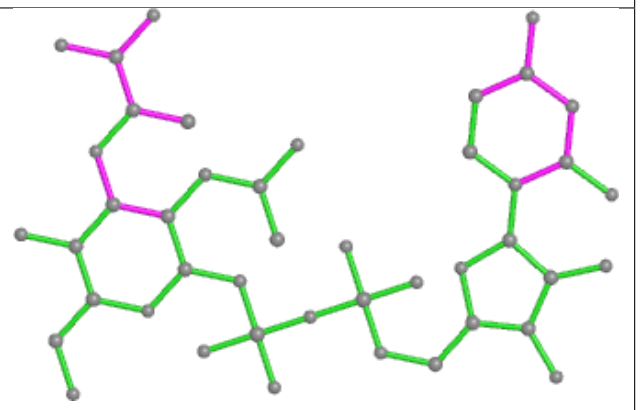
Ligand EPU A 503



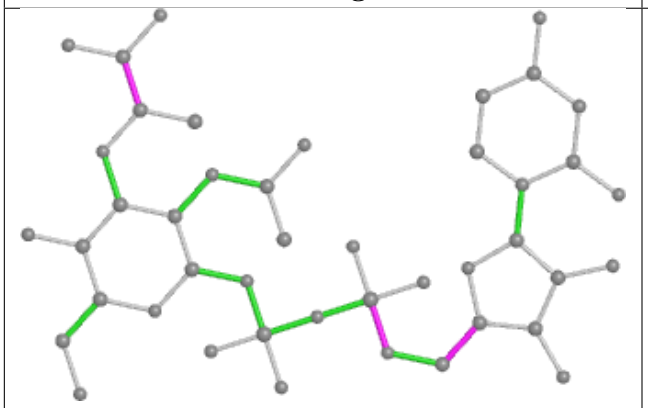
Ligand EPU B 503



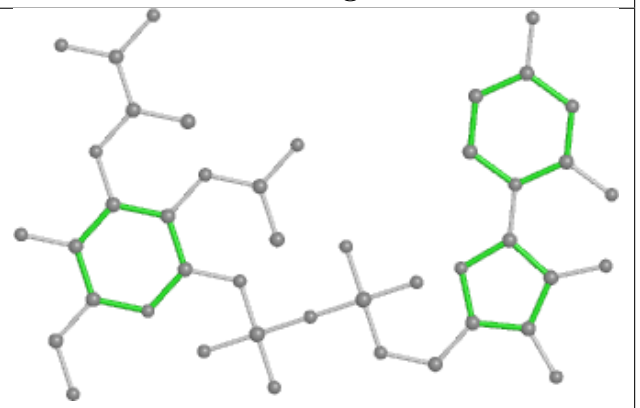
Bond lengths



Bond angles

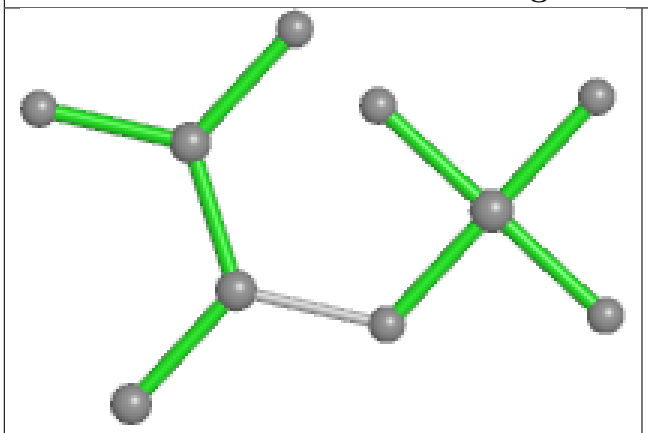


Torsions

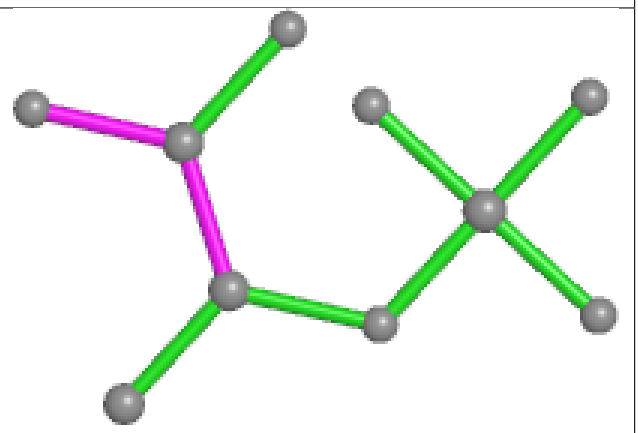


Rings

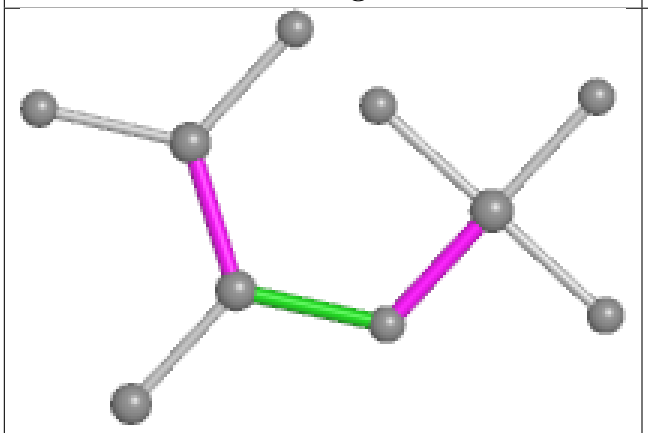
Ligand 0V5 A 501



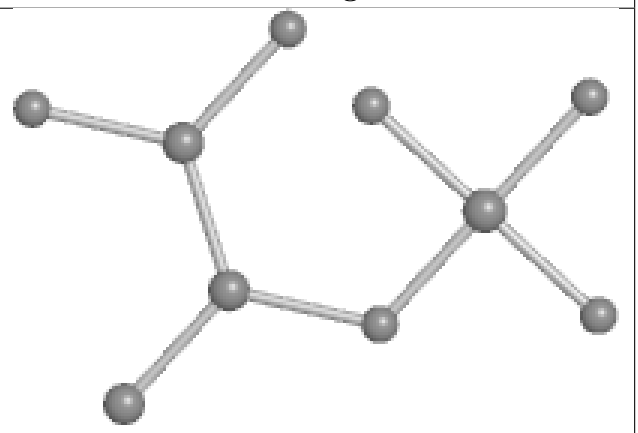
Bond lengths



Bond angles

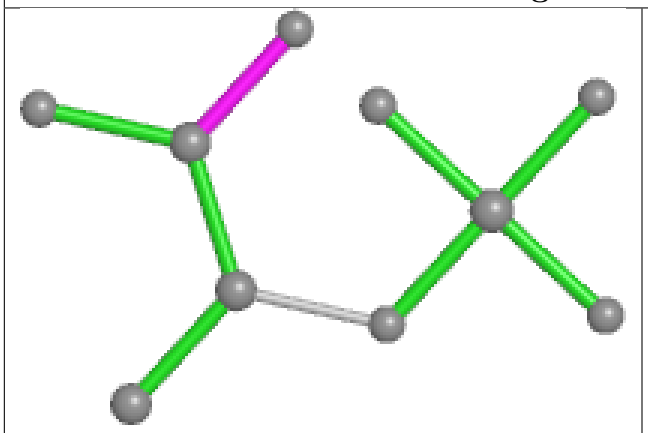


Torsions

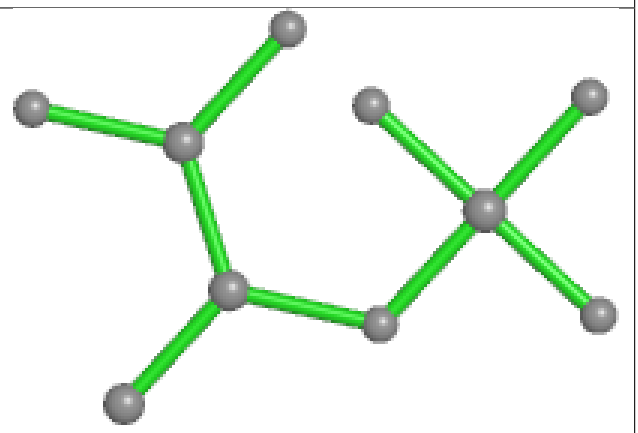


Rings

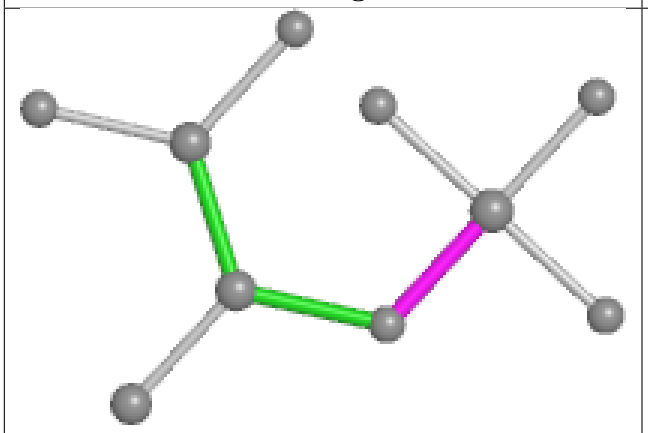
Ligand 0V5 B 501



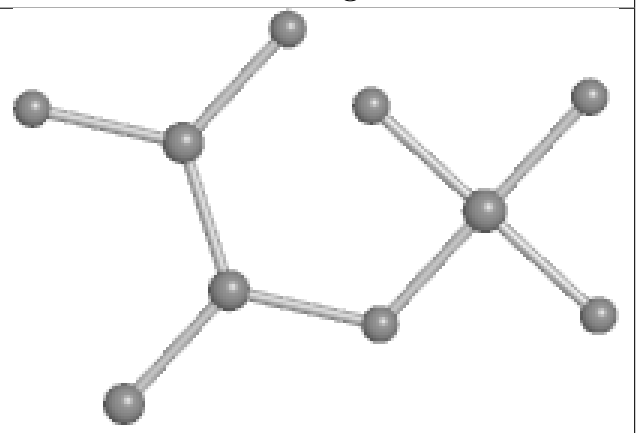
Bond lengths



Bond angles

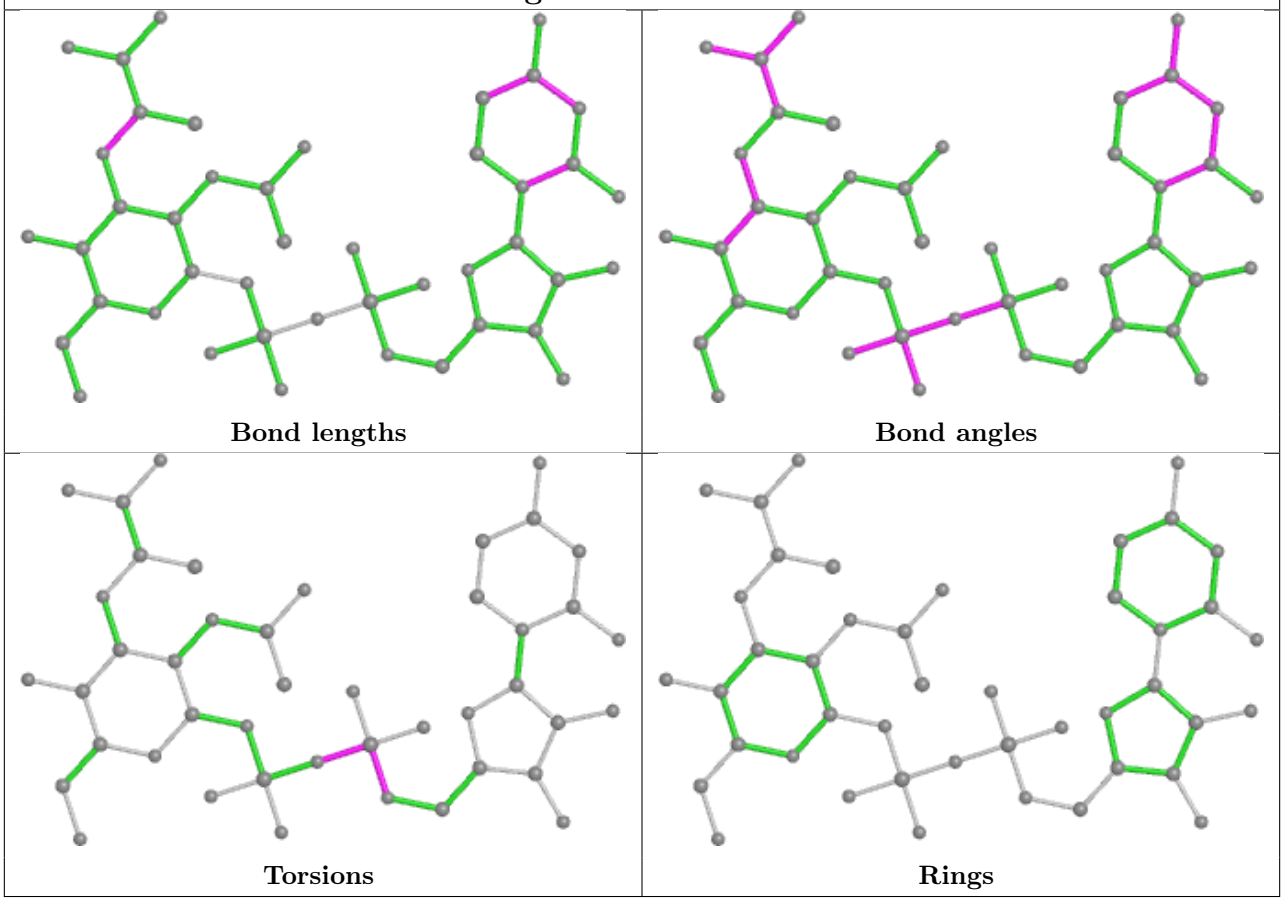


Torsions

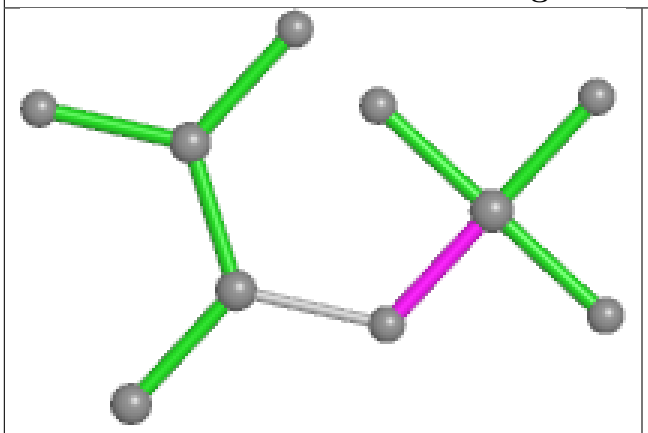


Rings

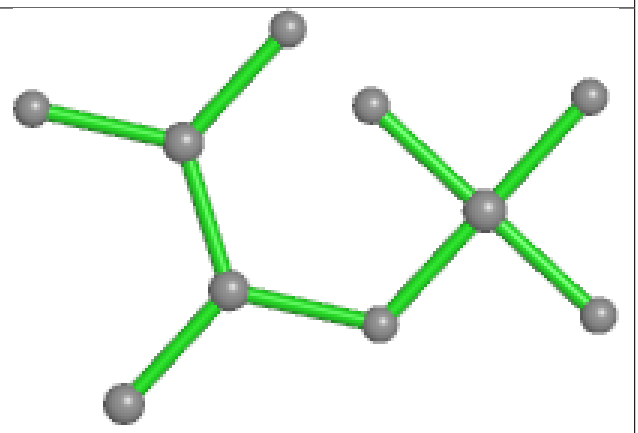
Ligand EPU C 503



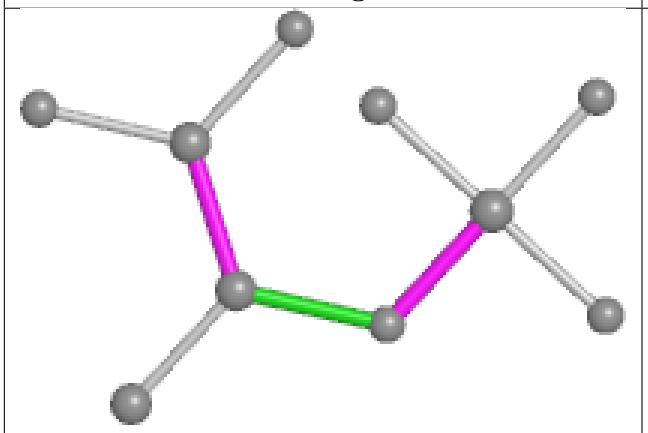
Ligand 0V5 C 501



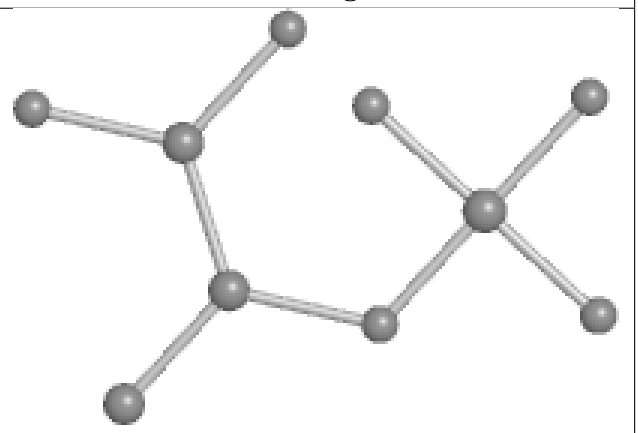
Bond lengths



Bond angles

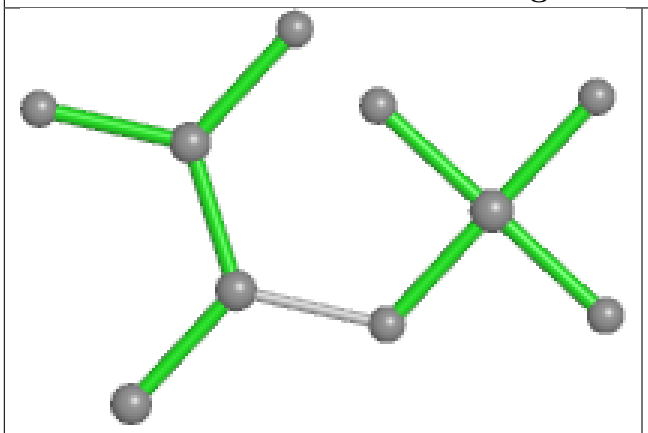


Torsions

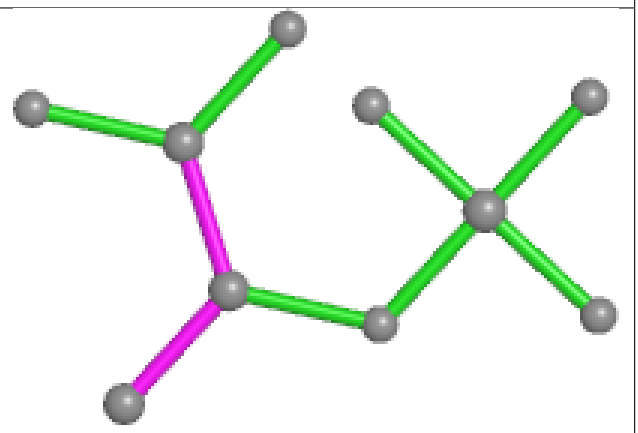


Rings

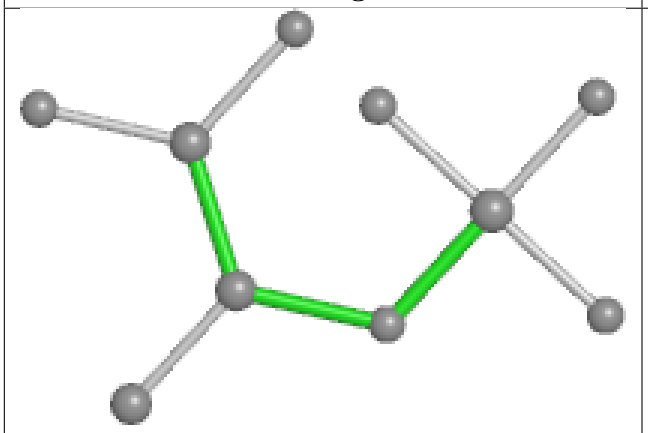
Ligand 0V5 D 501



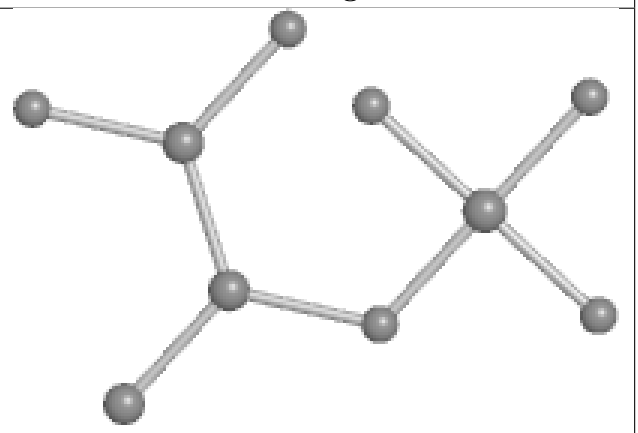
Bond lengths



Bond angles

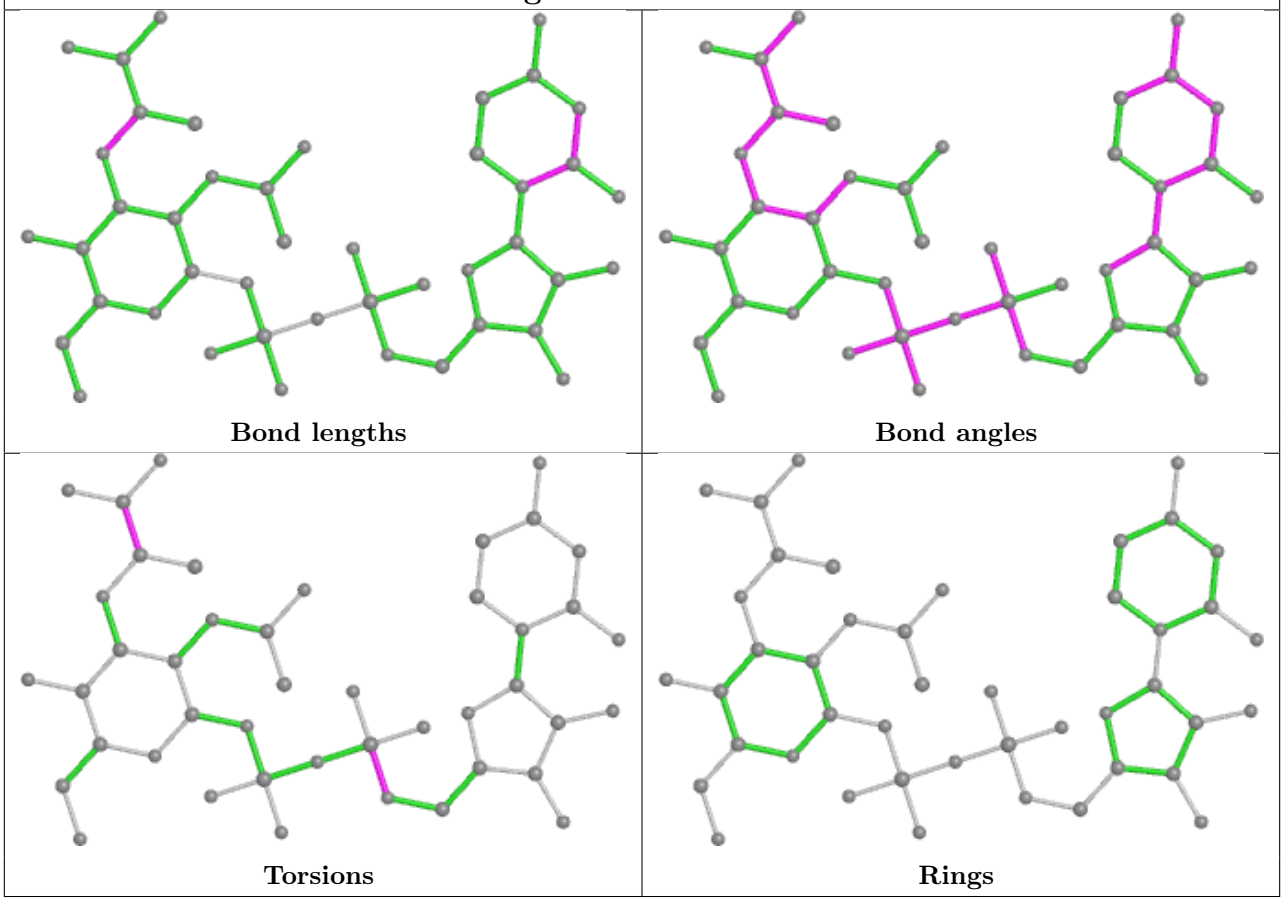


Torsions



Rings

Ligand EPU D 503



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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	422/430 (98%)	-0.22	6 (1%) 75 74	17, 26, 42, 81	0
1	B	422/430 (98%)	0.28	16 (3%) 40 39	23, 43, 65, 112	0
1	C	421/430 (97%)	1.28	101 (23%) 0 0	42, 72, 98, 121	0
1	D	422/430 (98%)	0.87	53 (12%) 3 3	34, 60, 89, 104	0
All	All	1687/1720 (98%)	0.55	176 (10%) 6 5	17, 50, 90, 121	0

All (176) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	382	LEU	6.4
1	B	1	MET	6.2
1	B	421	GLU	5.6
1	D	91	TYR	5.3
1	C	278	ILE	5.2
1	D	422	ALA	5.1
1	C	281	ASP	5.1
1	C	48	ILE	4.9
1	B	281	ASP	4.6
1	D	66	VAL	4.5
1	D	229	THR	4.5
1	C	200	ILE	4.4
1	D	420	ILE	4.3
1	C	271	LEU	4.3
1	C	257	ASP	4.3
1	C	76	LYS	4.2
1	C	277	VAL	4.2
1	B	161	TYR	4.1
1	C	36	ALA	4.1
1	C	260	TRP	4.0
1	D	282	GLU	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	39	GLY	3.9
1	C	82	ASP	3.9
1	A	422	ALA	3.9
1	C	212	ALA	3.9
1	D	421	GLU	3.9
1	C	64	ALA	3.9
1	C	91	TYR	3.8
1	C	258	ALA	3.7
1	C	280	GLU	3.7
1	C	77	VAL	3.6
1	C	225	LEU	3.6
1	D	27	LEU	3.6
1	C	420	ILE	3.6
1	C	410	LEU	3.6
1	C	101	VAL	3.6
1	C	279	GLU	3.6
1	D	206	GLY	3.6
1	D	53	PHE	3.6
1	C	237	ILE	3.6
1	C	27	LEU	3.5
1	D	77	VAL	3.5
1	D	73	HIS	3.4
1	C	1	MET	3.4
1	D	382	LEU	3.4
1	B	422	ALA	3.3
1	D	370	LEU	3.3
1	D	82	ASP	3.3
1	C	394	GLY	3.3
1	D	394	GLY	3.3
1	C	13	LEU	3.2
1	C	67	ASP	3.2
1	C	100	VAL	3.2
1	D	74	LEU	3.2
1	C	252	ASP	3.2
1	C	150	ALA	3.2
1	C	245	ALA	3.2
1	C	244	VAL	3.1
1	C	72	ALA	3.1
1	C	284	ILE	3.1
1	C	246	ALA	3.1
1	D	241	THR	3.1
1	C	255	ILE	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	222	VAL	3.1
1	C	283	GLY	3.0
1	C	204	GLU	3.0
1	C	220	THR	3.0
1	C	254	LEU	3.0
1	D	100	VAL	3.0
1	C	74	LEU	3.0
1	C	216	THR	3.0
1	D	1	MET	3.0
1	B	283	GLY	2.9
1	D	89	TYR	2.9
1	C	378	ALA	2.9
1	C	31	ALA	2.8
1	C	184	THR	2.8
1	D	42	VAL	2.8
1	C	78	ASP	2.8
1	D	176	ALA	2.8
1	D	388	GLN	2.8
1	B	74	LEU	2.8
1	C	37	SER	2.8
1	D	30	LEU	2.8
1	B	245	ALA	2.8
1	C	15	GLY	2.7
1	C	69	ASP	2.7
1	C	418	GLN	2.7
1	C	35	LEU	2.7
1	A	1	MET	2.7
1	C	33	THR	2.7
1	C	75	VAL	2.6
1	C	282	GLU	2.6
1	C	385	LEU	2.6
1	C	412	GLN	2.6
1	D	278	ILE	2.6
1	C	65	LYS	2.6
1	B	292	ASN	2.6
1	D	34	ILE	2.6
1	C	141	ILE	2.6
1	D	65	LYS	2.6
1	D	99	ILE	2.6
1	C	322	LYS	2.5
1	B	237	ILE	2.5
1	A	382	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	380	LEU	2.5
1	C	208	LYS	2.5
1	D	324	GLU	2.5
1	C	259	VAL	2.4
1	D	92	VAL	2.4
1	C	71	GLU	2.4
1	C	153	GLU	2.4
1	B	382	LEU	2.4
1	D	378	ALA	2.4
1	C	206	GLY	2.4
1	C	14	VAL	2.4
1	C	26	VAL	2.4
1	D	76	LYS	2.4
1	D	67	ASP	2.4
1	C	383	THR	2.4
1	C	370	LEU	2.4
1	C	297	HIS	2.4
1	C	105	ILE	2.4
1	D	18	THR	2.4
1	C	90	LYS	2.4
1	D	90	LYS	2.4
1	D	322	LYS	2.3
1	D	17	VAL	2.3
1	D	257	ASP	2.3
1	C	161	TYR	2.3
1	C	106	LEU	2.3
1	B	101	VAL	2.3
1	C	175	MET	2.3
1	D	43	LEU	2.3
1	C	211	GLY	2.3
1	A	241	THR	2.2
1	C	68	PHE	2.2
1	C	119	GLY	2.2
1	D	258	ALA	2.2
1	C	66	VAL	2.2
1	A	316	ALA	2.2
1	C	30	LEU	2.2
1	D	403	TYR	2.2
1	C	248	MET	2.2
1	C	379	ALA	2.2
1	C	291	GLU	2.2
1	C	44	GLN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	375	ARG	2.2
1	B	282	GLU	2.1
1	C	198	LEU	2.1
1	B	244	VAL	2.1
1	D	68	PHE	2.1
1	A	421	GLU	2.1
1	C	218	THR	2.1
1	C	73	HIS	2.1
1	C	46	VAL	2.1
1	C	18	THR	2.1
1	C	104	PRO	2.1
1	D	41	THR	2.1
1	D	255	ILE	2.1
1	C	62	LEU	2.1
1	C	42	VAL	2.1
1	C	256	ARG	2.1
1	D	97	ALA	2.1
1	C	404	TYR	2.1
1	D	83	ILE	2.0
1	D	244	VAL	2.0
1	D	360	GLY	2.0
1	D	88	PRO	2.0
1	B	241	THR	2.0
1	C	241	THR	2.0
1	C	253	VAL	2.0
1	C	63	ASN	2.0
1	D	63	ASN	2.0
1	D	404	TYR	2.0
1	B	100	VAL	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

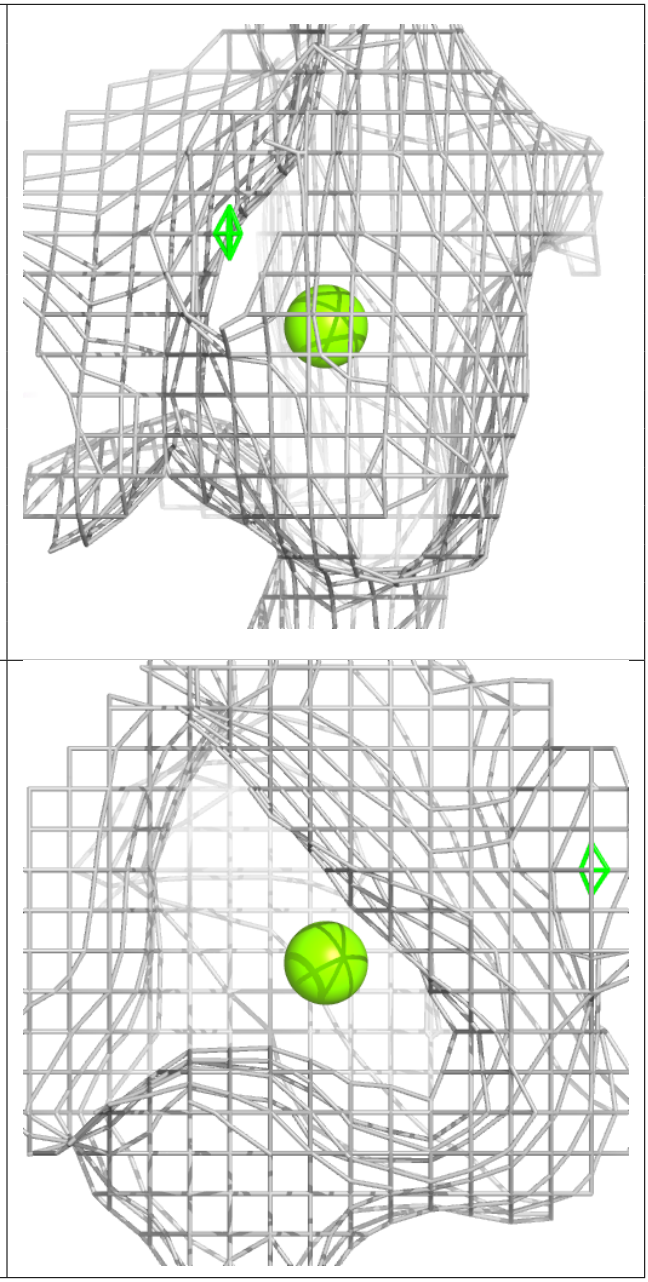
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	D	502	1/1	0.85	0.08	59,59,59,59	0
3	MG	C	502	1/1	0.90	0.05	60,60,60,60	0
4	EPU	C	503	44/44	0.91	0.14	40,52,69,76	0
2	0V5	D	501	10/10	0.94	0.11	36,53,57,58	0
4	EPU	B	503	44/44	0.95	0.12	28,37,44,46	0
2	0V5	C	501	10/10	0.95	0.10	37,54,57,64	0
4	EPU	D	503	44/44	0.95	0.13	29,43,51,60	0
3	MG	A	502	1/1	0.96	0.05	31,31,31,31	0
3	MG	B	502	1/1	0.97	0.03	35,35,35,35	0
2	0V5	B	501	10/10	0.97	0.09	28,31,41,47	0
4	EPU	A	503	44/44	0.98	0.11	18,24,33,40	0
2	0V5	A	501	10/10	0.98	0.08	25,28,31,32	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.

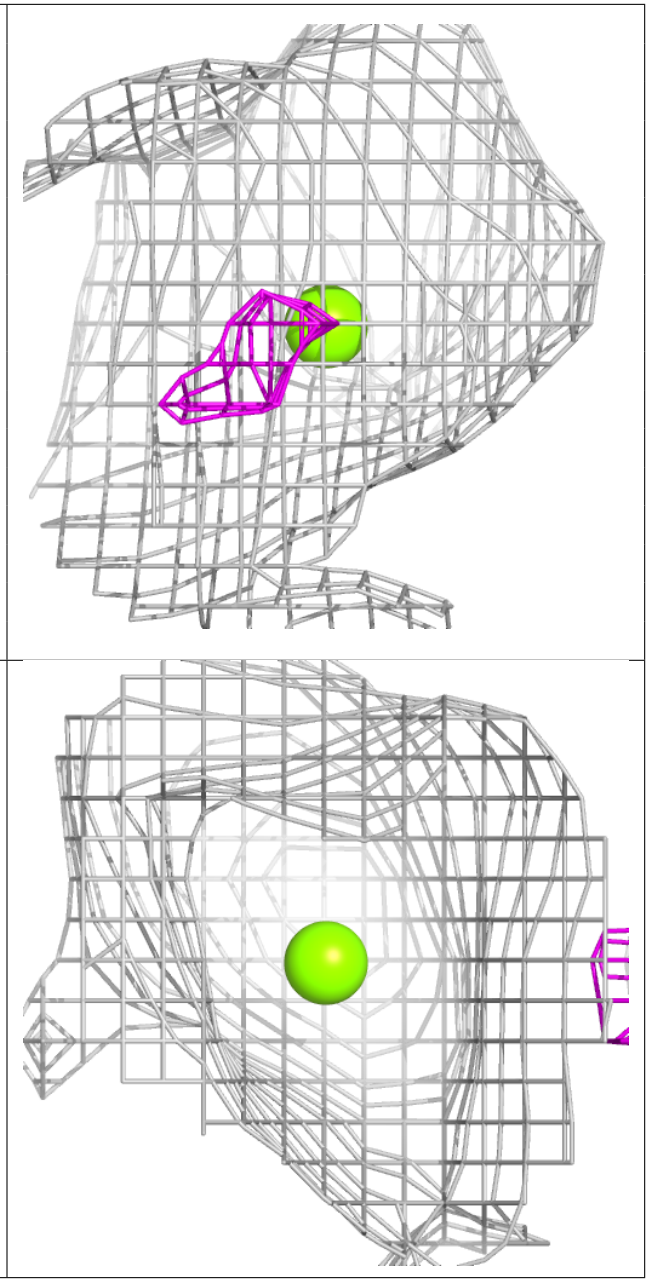
Electron density around MG D 502:

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



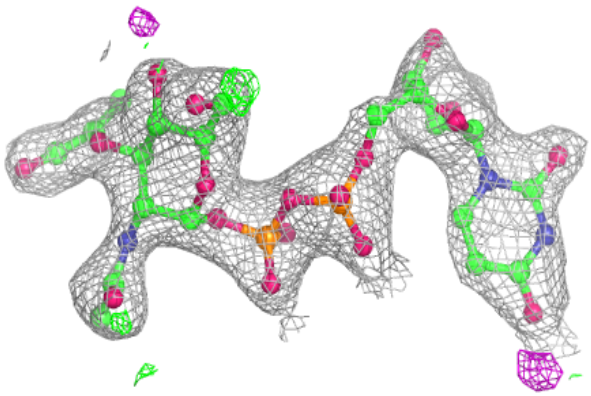
Electron density around MG C 502:

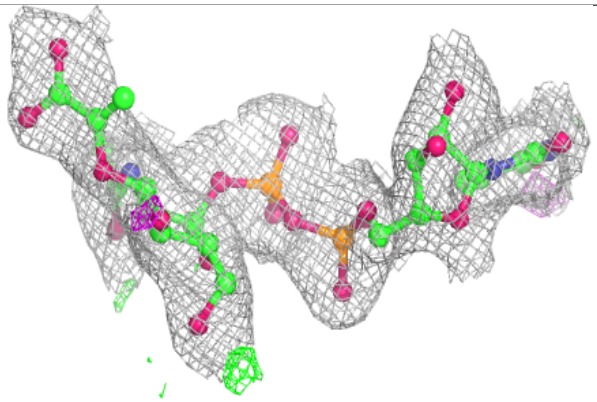
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

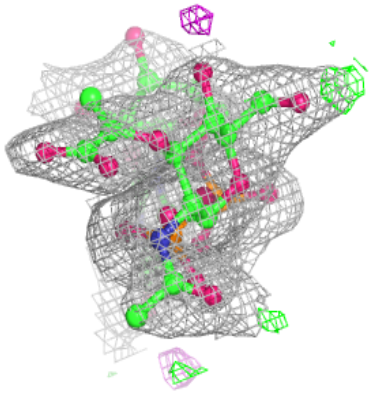


Electron density around EPU C 503:

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

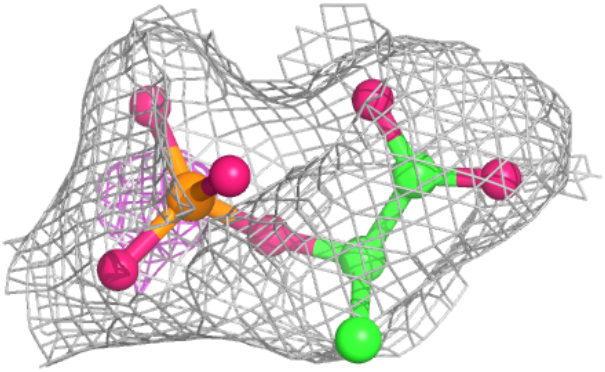


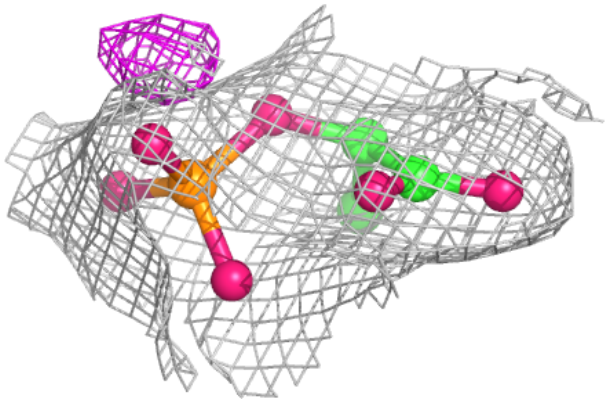


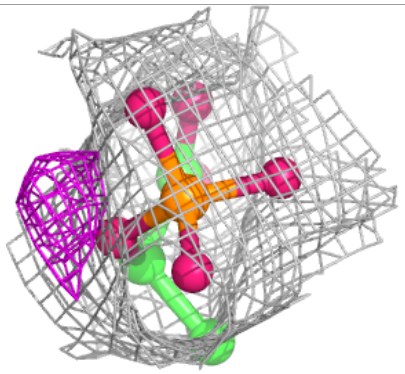


Electron density around 0V5 D 501:

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

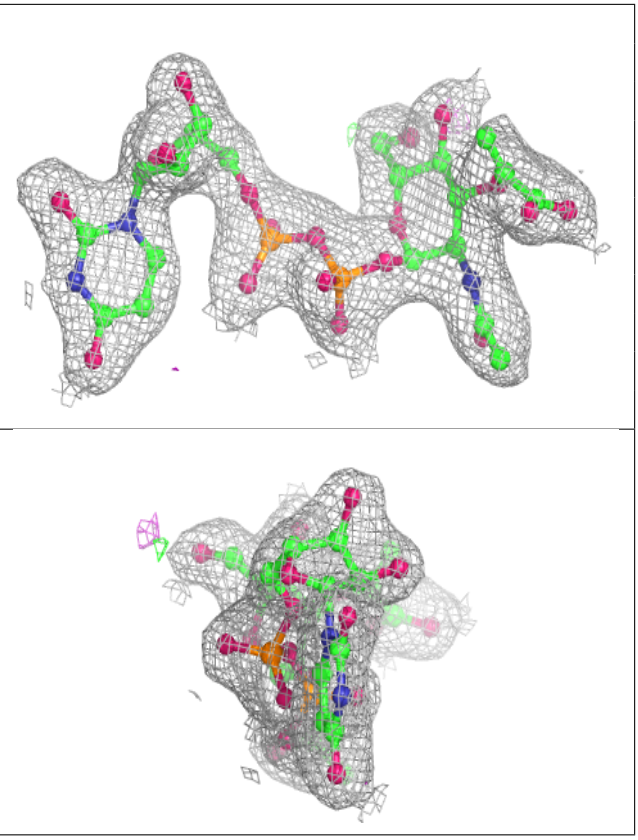






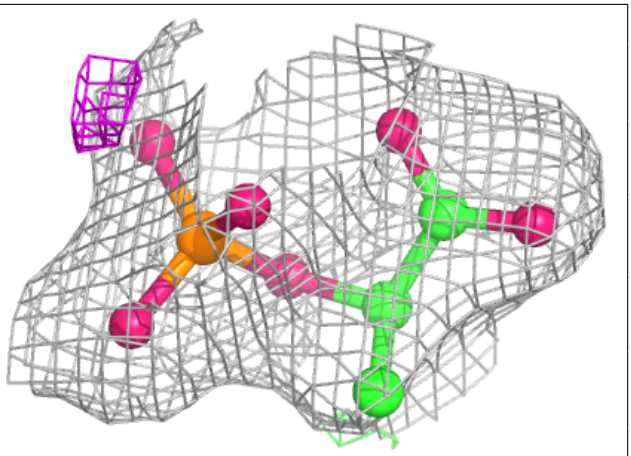
Electron density around EPU B 503:

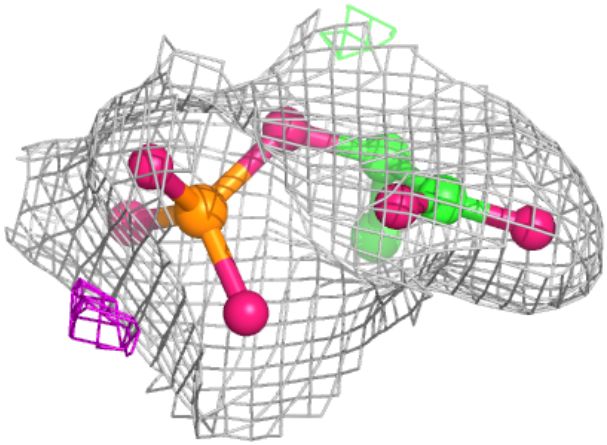
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

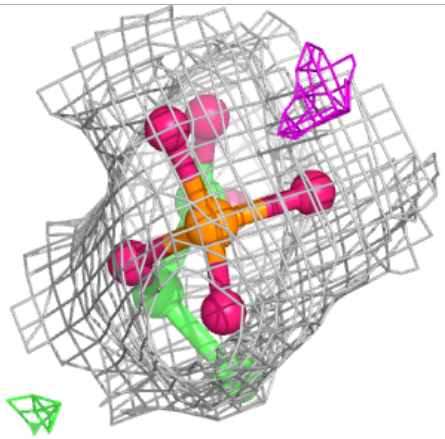


Electron density around 0V5 C 501:

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

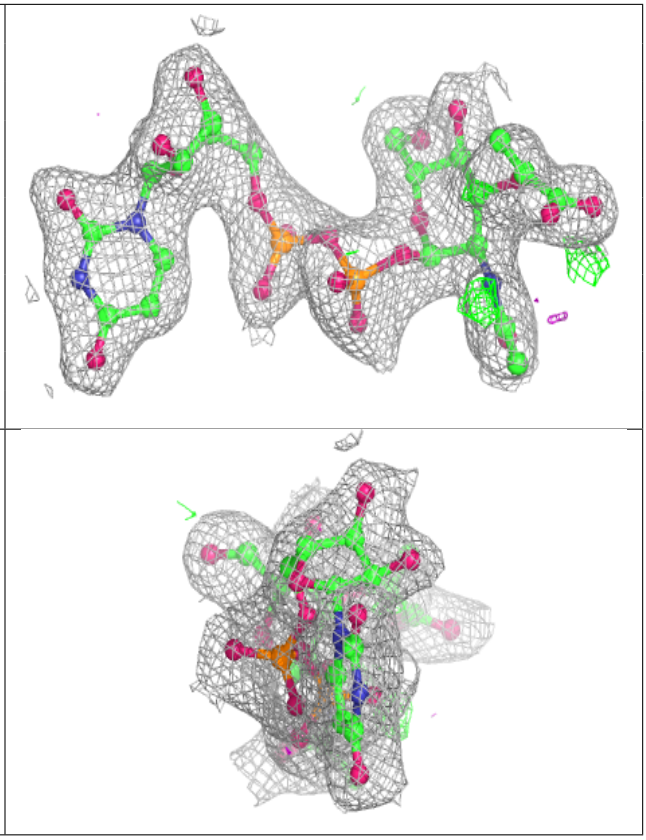






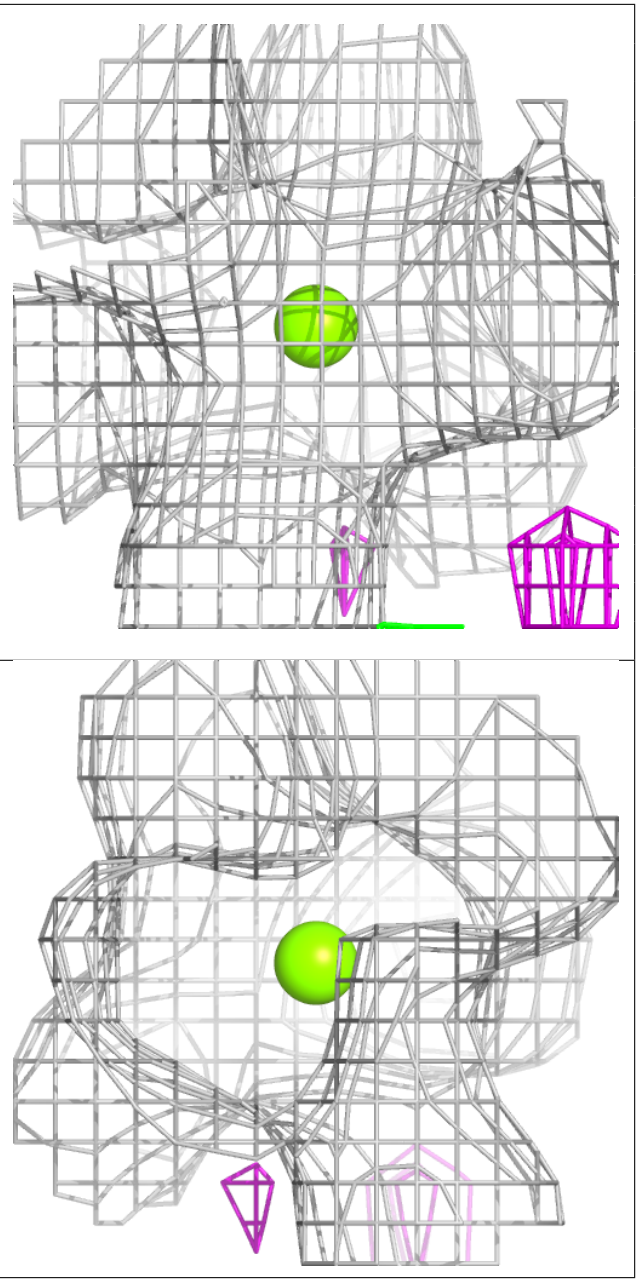
Electron density around EPU D 503:

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



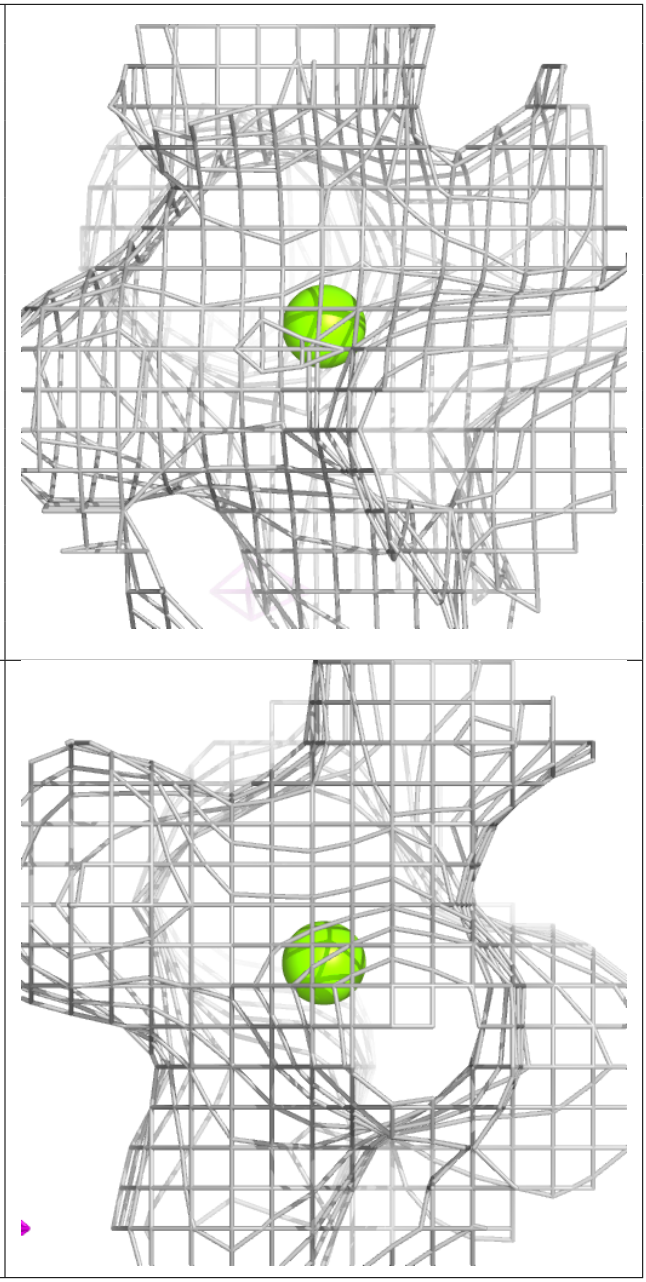
Electron density around MG A 502:

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



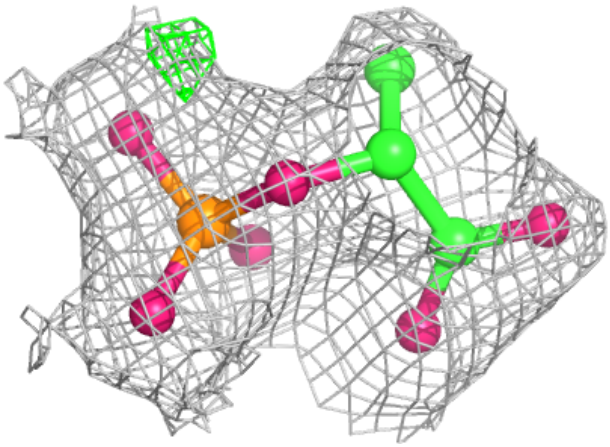
Electron density around MG B 502:

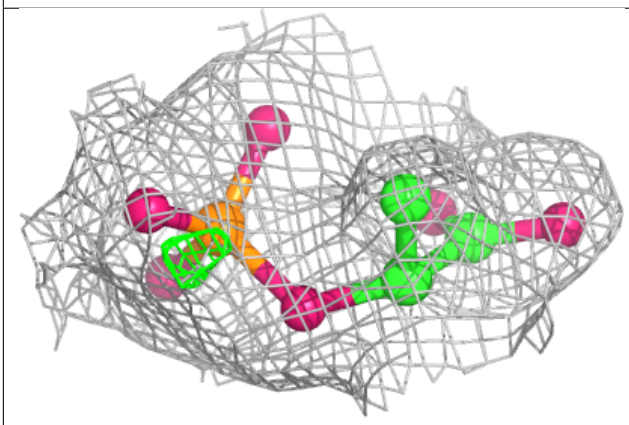
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

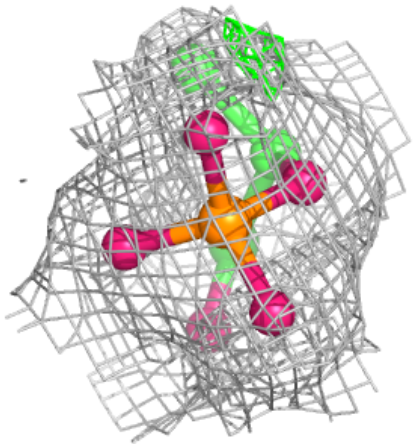


Electron density around 0V5 B 501:

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

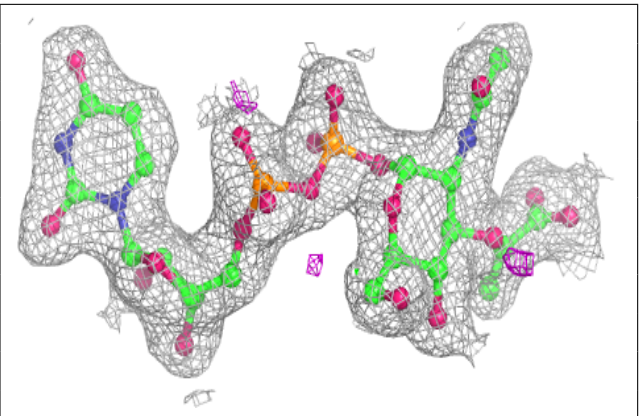


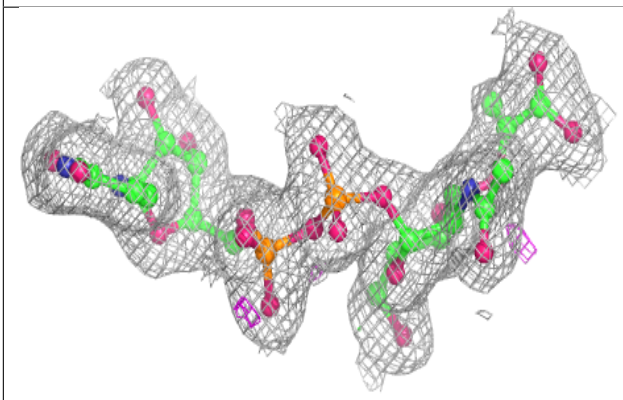


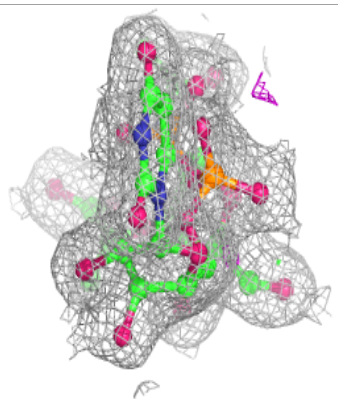


Electron density around EPU A 503:

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

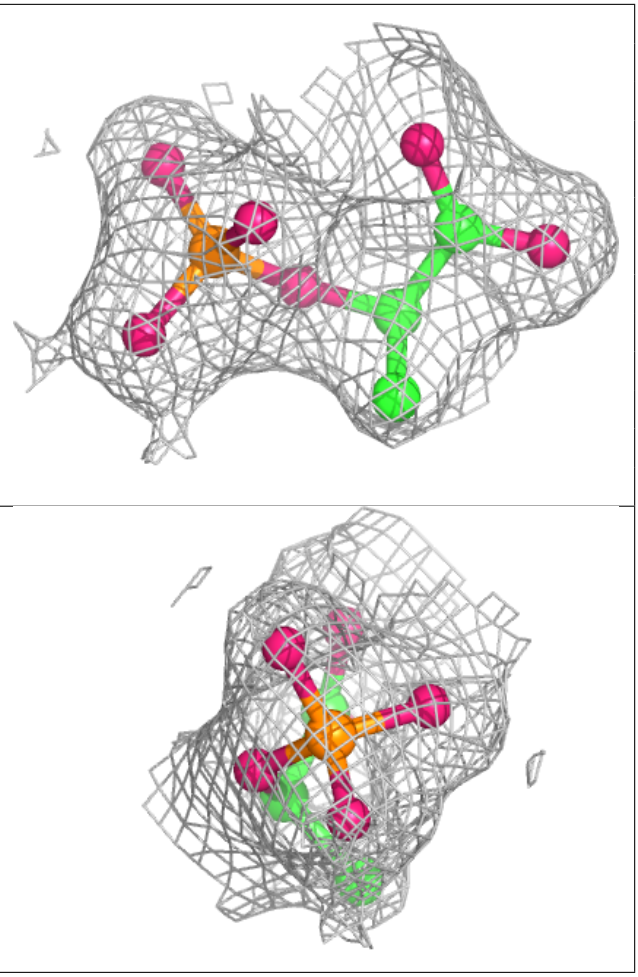






Electron density around 0V5 A 501:

$2mF_o$-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.