Full wwPDB X-ray Structure Validation Report (i)

Feb 20, 2024 – 07:55 AM EST

PDB ID	:	4MX2
Title	:	Crystal Structure of adenylosuccinate lyase from Leishmania donovani
Authors	:	Wernimont, A.K.; Loppnau, P.; Dong, A.; Krojer, T.; Bradley, A.; Bushell, S.; von Delft, F.; Robinson, D.; Gilbert, I.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Hui, R.; Mottaghi, K.; Structural Genomics Consortium (SGC)
Deposited on	:	2013-09-25
Resolution	:	1.90 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:

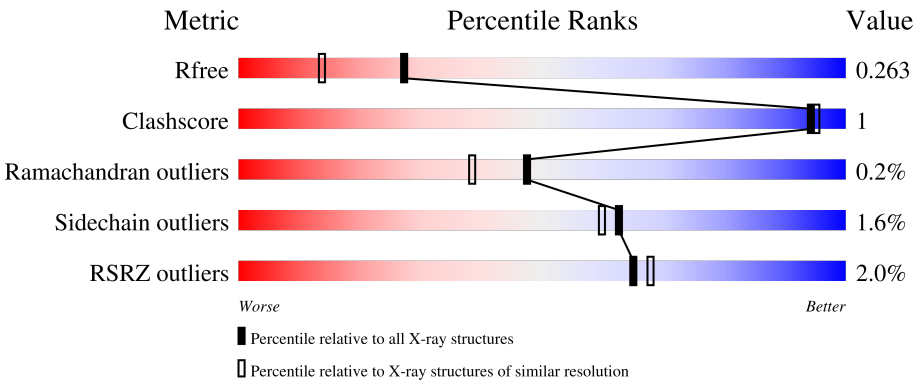
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	480	2%, 85%, •, 12%
1	B	480	%, 87%, 5%, 9%
1	C	480	3%, 85%, •, 11%
1	D	480	%, 88%, •, 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	480	3% 85% • 11%
1	F	480	% 85% 5% 10%
1	G	480	2% 85% • 11%
1	H	480	2% 86% 5% 9%

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
4	GOL	E	501	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 29460 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.

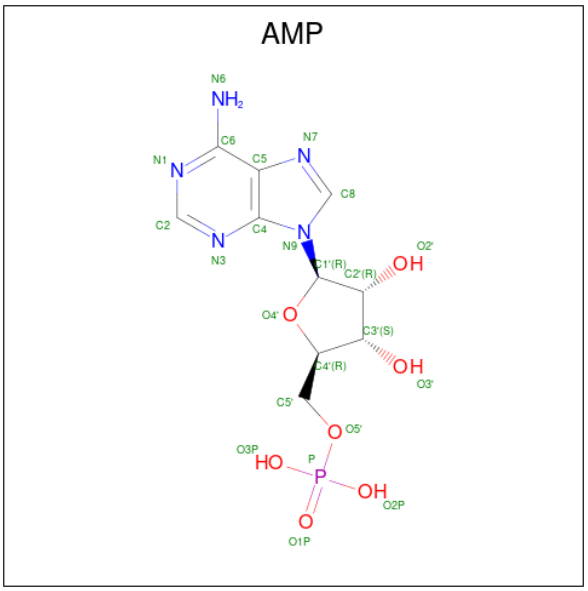
- Molecule 1 is a protein called Adenylosuccinate lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	424	3333	2145	557	617	14	0	8	0
1	B	439	3522	2261	595	652	14	0	15	0
1	C	426	3367	2167	568	618	14	0	8	0
1	D	439	3487	2239	588	647	13	0	9	0
1	E	425	3347	2154	560	619	14	0	9	0
1	F	431	3450	2222	577	637	14	0	12	0
1	G	427	3375	2174	567	620	14	0	9	0
1	H	438	3490	2247	584	644	15	0	14	0

There are 8 discrepancies between the modelled and reference sequences:

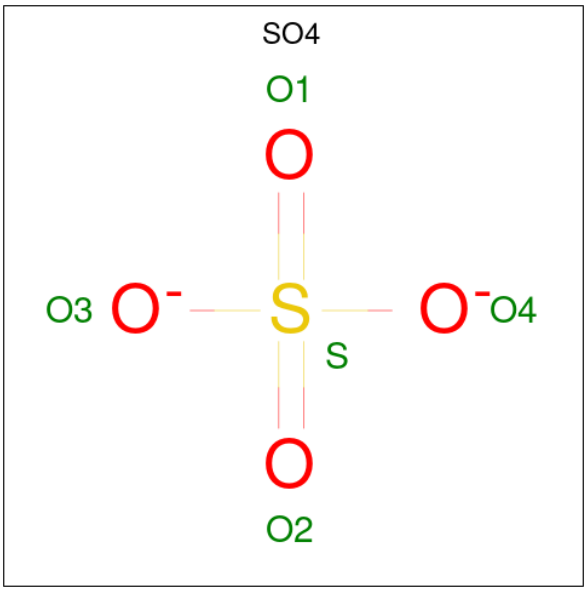
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A7LBL3
B	0	GLY	-	expression tag	UNP A7LBL3
C	0	GLY	-	expression tag	UNP A7LBL3
D	0	GLY	-	expression tag	UNP A7LBL3
E	0	GLY	-	expression tag	UNP A7LBL3
F	0	GLY	-	expression tag	UNP A7LBL3
G	0	GLY	-	expression tag	UNP A7LBL3
H	0	GLY	-	expression tag	UNP A7LBL3

- Molecule 2 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$).



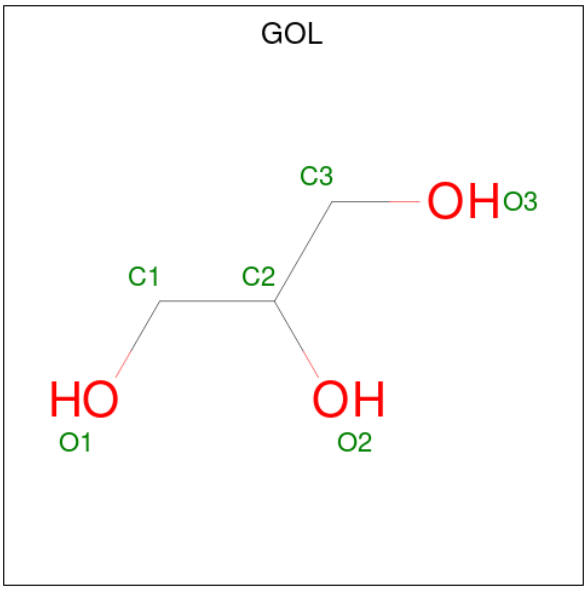
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	23	10	5	7	1	0	0
2	B	1	23	10	5	7	1	0	0
2	C	1	23	10	5	7	1	0	0
2	D	1	23	10	5	7	1	0	0
2	E	1	23	10	5	7	1	0	0
2	F	1	23	10	5	7	1	0	0
2	G	1	23	10	5	7	1	0	0
2	H	1	23	10	5	7	1	0	0

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



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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	E	1	6	3	3	0	0

- Molecule 5 is water.

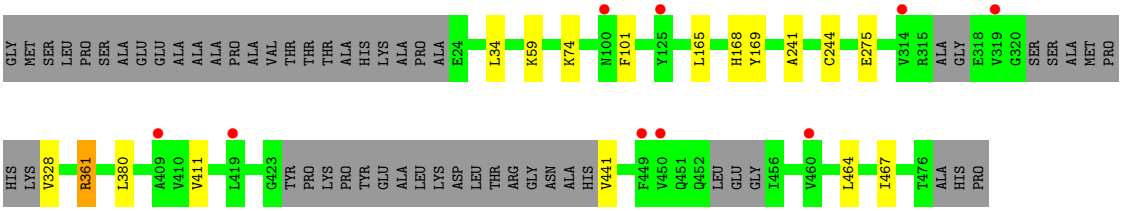
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	232	233	233	0	1
			Total	O		
5	B	244	244	244	0	0
			Total	O		
5	C	212	212	212	0	0
			Total	O		
5	D	264	265	265	0	1
			Total	O		
5	E	215	215	215	0	0
			Total	O		
5	F	239	239	239	0	0
			Total	O		
5	G	217	217	217	0	0
			Total	O		
5	H	259	259	259	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.

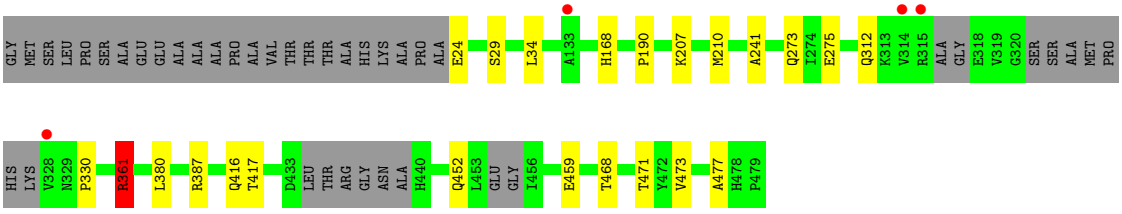
• Molecule 1: Adenylosuccinate lyase

Chain A:



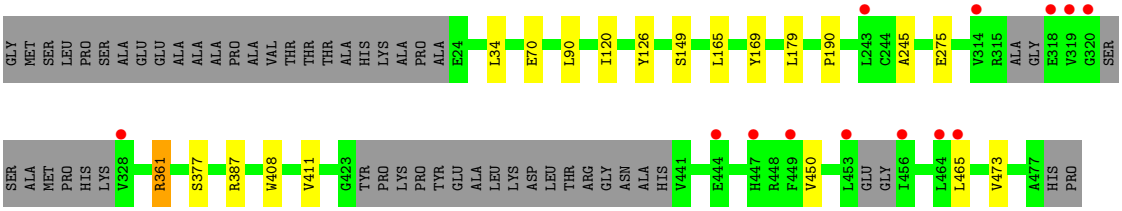
• Molecule 1: Adenylosuccinate lyase

Chain B:



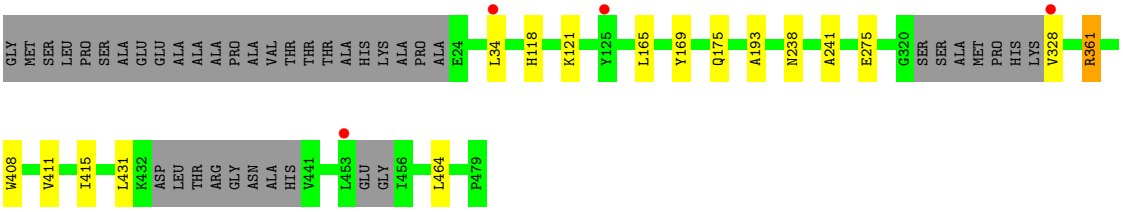
• Molecule 1: Adenylosuccinate lyase

Chain C:

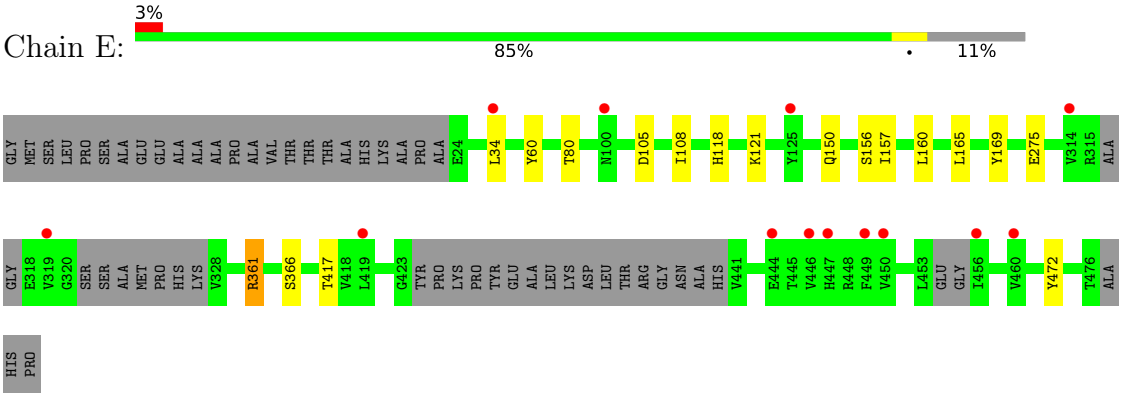


• Molecule 1: Adenylosuccinate lyase

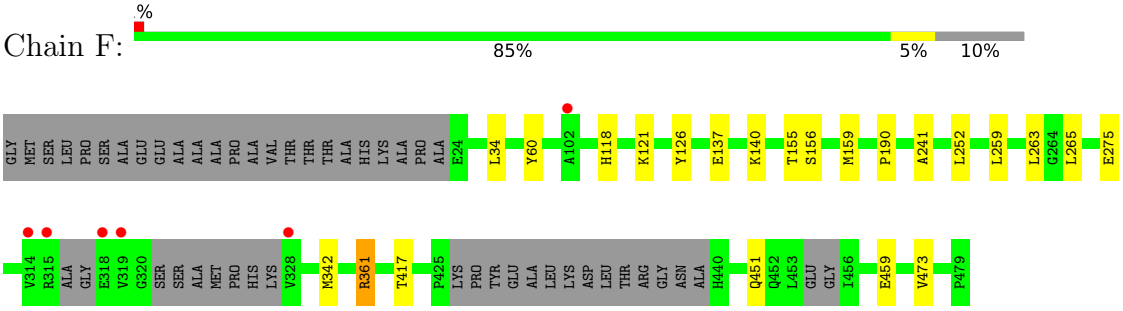
Chain D:



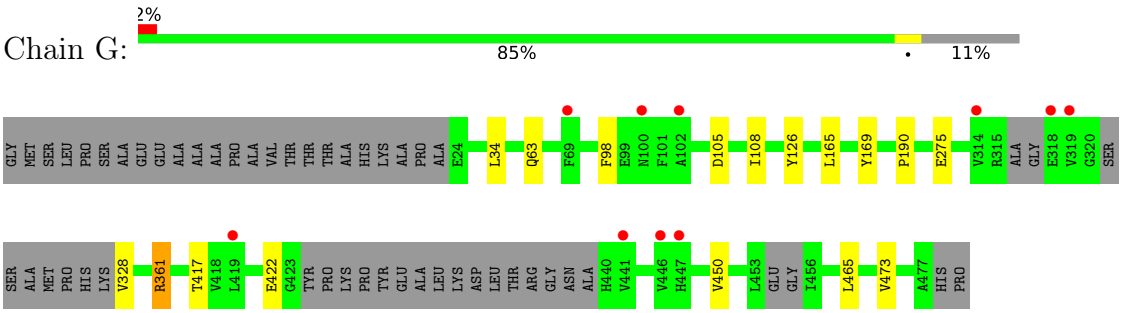
- Molecule 1: Adenylosuccinate lyase



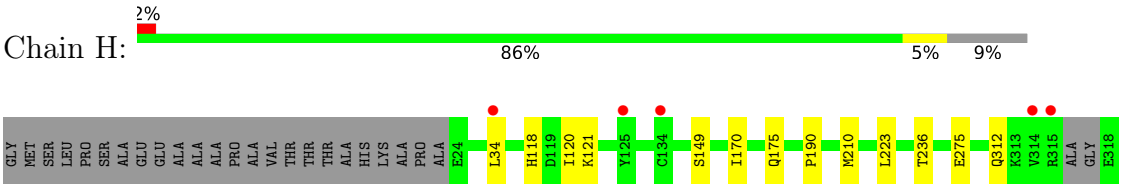
- Molecule 1: Adenylosuccinate lyase

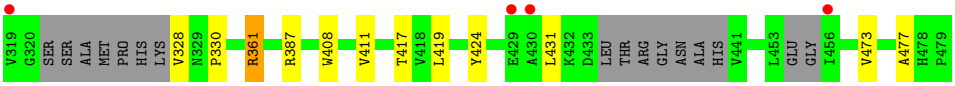


- Molecule 1: Adenylosuccinate lyase



- Molecule 1: Adenylosuccinate lyase




V319
G320
SER
SER
ALA
MET
PRO
HIS
LYS
V328
N329
P330
R361
R387
W408
V411
T417
V418
L419
Y424
E429
A430
L431
K432
D433
LEU
THR
ARG
GLY
ASN
ALA
HIS
V441
L453
GLU
GLY
I456
V473
A477
H478
P479

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	124.02Å 194.70Å 147.51Å 90.00° 91.31° 90.00°	Depositor
Resolution (Å)	34.98 – 1.90 34.86 – 1.90	Depositor EDS
% Data completeness (in resolution range)	95.7 (34.98-1.90) 89.4 (34.86-1.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.94 (at 1.89Å)	Xtriage
Refinement program	BUSTER-TNT BUSTER 2.10.0, BUSTER 2.10.0	Depositor
R, R_{free}	0.218 , 0.253 0.226 , 0.263	Depositor DCC
R_{free} test set	13161 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	16.5	Xtriage
Anisotropy	0.309	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 32.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.199 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	29460	wwPDB-VP
Average B, all atoms (Å^2)	20.0	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 44.58 % 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 1.4942e-04. 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: AMP, GOL, 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	0.50	0/3424	0.59	0/4649
1	B	0.51	0/3626	0.59	0/4921
1	C	0.50	0/3460	0.59	0/4692
1	D	0.50	0/3585	0.59	0/4870
1	E	0.48	0/3442	0.59	0/4673
1	F	0.50	0/3556	0.59	0/4827
1	G	0.49	0/3471	0.59	0/4710
1	H	0.50	0/3607	0.59	0/4900
All	All	0.50	0/28171	0.59	0/38242

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 (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	3333	0	3288	7	0
1	B	3522	0	3471	9	0
1	C	3367	0	3322	9	0
1	D	3487	0	3454	10	0
1	E	3347	0	3299	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	3450	0	3429	13	0
1	G	3375	0	3335	6	0
1	H	3490	0	3460	13	0
2	A	23	0	12	0	0
2	B	23	0	12	0	0
2	C	23	0	12	0	0
2	D	23	0	12	0	0
2	E	23	0	12	0	0
2	F	23	0	12	0	0
2	G	23	0	12	0	0
2	H	23	0	12	0	0
3	C	5	0	0	0	0
3	G	5	0	0	0	0
3	H	5	0	0	0	0
4	E	6	0	8	4	0
5	A	233	0	0	0	0
5	B	244	0	0	0	0
5	C	212	0	0	0	0
5	D	265	0	0	1	0
5	E	215	0	0	0	0
5	F	239	0	0	2	0
5	G	217	0	0	0	0
5	H	259	0	0	0	0
All	All	29460	0	27162	70	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 1.

All (70) 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:E:157:ILE:HD11	4:E:501:GOL:H32	1.39	1.05
1:A:244[B]:CYS:HG	1:E:472:TYR:HD2	1.05	0.98
1:H:170:ILE:HD11	1:H:223:LEU:HD23	1.67	0.76
1:E:366:SER:HA	4:E:501:GOL:H12	1.67	0.75
1:E:157:ILE:CD1	4:E:501:GOL:H32	2.18	0.72
1:B:417[B]:THR:HG21	1:F:241:ALA:HB1	1.72	0.71
1:F:259[B]:LEU:HD22	1:F:265:LEU:HB2	1.77	0.66
1:B:241:ALA:HB1	1:F:417[B]:THR:HG21	1.77	0.65
1:F:259[B]:LEU:HD23	1:F:263:LEU:HD12	1.78	0.65
1:D:241:ALA:HB1	1:H:417[B]:THR:HG21	1.79	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:155:THR:HG22	1:F:259[B]:LEU:HD21	1.81	0.62
1:E:118:HIS:HD2	1:E:121:LYS:H	1.48	0.60
1:B:468:THR:H	1:B:471[A]:THR:HG22	1.65	0.60
1:A:59:LYS:HD3	1:A:101:PHE:HB3	1.87	0.57
1:A:241:ALA:HB1	1:E:417[B]:THR:HG21	1.88	0.56
1:H:170:ILE:HD11	1:H:223:LEU:CD2	2.37	0.54
1:D:415:ILE:HG23	1:D:464:LEU:HD22	1.89	0.54
1:G:190:PRO:HB3	1:G:473:VAL:HG11	1.89	0.54
1:H:120:ILE:HG13	1:H:149:SER:HB3	1.90	0.54
1:D:118:HIS:HD2	1:D:121:LYS:H	1.55	0.52
1:E:165:LEU:HA	1:E:169:TYR:HB3	1.92	0.50
1:H:118[A]:HIS:HD2	1:H:121:LYS:H	1.59	0.49
1:F:259[B]:LEU:CD2	1:F:263:LEU:HD12	2.43	0.49
1:E:105:ASP:HA	1:E:108:ILE:HD12	1.95	0.49
1:E:60:TYR:HB3	1:E:156:SER:HB2	1.94	0.49
1:B:190:PRO:HB3	1:B:473:VAL:HG11	1.94	0.49
1:F:342:MET:HE3	1:F:342:MET:HA	1.94	0.48
1:B:210[B]:MET:SD	1:B:477:ALA:HA	2.54	0.48
1:F:190:PRO:HB3	1:F:473:VAL:HG11	1.96	0.48
1:C:377:SER:OG	1:D:34:LEU:HD11	2.13	0.47
1:D:165:LEU:HA	1:D:169:TYR:HB3	1.96	0.47
1:B:312:GLN:HE22	1:B:330:PRO:HA	1.79	0.47
1:H:190:PRO:HB3	1:H:473:VAL:HG11	1.97	0.47
1:H:210[B]:MET:SD	1:H:477:ALA:HA	2.54	0.47
1:H:312:GLN:HE22	1:H:330:PRO:HA	1.80	0.47
1:B:273:GLN:HB3	1:B:361:ARG:HG2	1.97	0.47
1:C:245:ALA:HB2	1:G:417:THR:HG23	1.98	0.46
1:A:168[B]:HIS:HD2	1:A:380:LEU:HD11	1.81	0.45
1:H:175:GLN:HB3	1:H:387:ARG:NH2	2.31	0.45
1:C:165:LEU:HA	1:C:169:TYR:HB3	1.98	0.45
1:E:366:SER:HA	4:E:501:GOL:C1	2.42	0.44
1:A:411:VAL:HB	1:A:467:ILE:HG22	1.99	0.44
1:C:450:VAL:HG11	1:C:465:LEU:HG	1.99	0.44
1:F:252:LEU:HD21	5:F:605:HOH:O	2.18	0.44
1:H:419:LEU:HB3	1:H:424:TYR:HB3	2.00	0.44
1:D:241:ALA:O	1:H:417[A]:THR:HG21	2.17	0.44
1:G:105:ASP:HA	1:G:108:ILE:HD12	2.00	0.43
1:D:118:HIS:CD2	1:D:121:LYS:H	2.36	0.43
1:G:165:LEU:HA	1:G:169:TYR:HB3	2.00	0.43
1:A:165:LEU:HA	1:A:169:TYR:HB3	2.00	0.43
1:E:118:HIS:CD2	1:E:121:LYS:H	2.33	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:408:TRP:O	1:H:411:VAL:HG22	2.18	0.43
1:E:157:ILE:HA	1:E:160:LEU:HG	2.00	0.42
1:C:179:LEU:HD22	1:C:387[B]:ARG:HG3	2.02	0.42
1:F:60:TYR:HB3	1:F:156[B]:SER:HB2	2.00	0.42
1:B:207:LYS:NZ	5:F:766:HOH:O	2.53	0.42
1:C:120:ILE:HG21	1:C:149:SER:HA	2.02	0.42
1:G:63:GLN:HA	1:G:98:PHE:CD2	2.54	0.42
1:C:408:TRP:O	1:C:411:VAL:HG22	2.20	0.41
1:G:450:VAL:HG11	1:G:465:LEU:HG	2.02	0.41
1:A:464:LEU:HA	1:A:467:ILE:HD12	2.03	0.41
1:B:168[B]:HIS:HD2	1:B:380:LEU:HD11	1.85	0.41
1:D:238:ASN:ND2	5:D:852:HOH:O	2.53	0.41
1:F:137:GLU:O	1:F:140:LYS:HG2	2.20	0.41
1:C:190:PRO:HB3	1:C:473:VAL:HG11	2.02	0.40
1:F:159:MET:HG3	1:F:265:LEU:HD11	2.03	0.40
1:C:70:GLU:HG3	1:C:90:LEU:HD13	2.04	0.40
1:F:118:HIS:CD2	1:F:121:LYS:H	2.39	0.40
1:D:193:ALA:CB	1:H:236:THR:HG21	2.51	0.40
1:D:408:TRP:O	1:D:411:VAL:HG22	2.20	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	422/480 (88%)	409 (97%)	12 (3%)	1 (0%)	47 38
1	B	444/480 (92%)	422 (95%)	21 (5%)	1 (0%)	47 38
1	C	424/480 (88%)	408 (96%)	15 (4%)	1 (0%)	47 38
1	D	441/480 (92%)	421 (96%)	19 (4%)	1 (0%)	47 38
1	E	424/480 (88%)	408 (96%)	15 (4%)	1 (0%)	47 38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	433/480 (90%)	415 (96%)	17 (4%)	1 (0%)	47	38
1	G	427/480 (89%)	411 (96%)	15 (4%)	1 (0%)	47	38
1	H	442/480 (92%)	422 (96%)	19 (4%)	1 (0%)	47	38
All	All	3457/3840 (90%)	3316 (96%)	133 (4%)	8 (0%)	47	38

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	361	ARG
1	B	361	ARG
1	C	361	ARG
1	D	361	ARG
1	E	361	ARG
1	F	361	ARG
1	G	361	ARG
1	H	361	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	350/403 (87%)	344 (98%)	6 (2%)	60	57
1	B	371/403 (92%)	361 (97%)	10 (3%)	44	38
1	C	350/403 (87%)	346 (99%)	4 (1%)	73	73
1	D	370/403 (92%)	365 (99%)	5 (1%)	67	65
1	E	351/403 (87%)	346 (99%)	5 (1%)	67	65
1	F	368/403 (91%)	362 (98%)	6 (2%)	62	60
1	G	353/403 (88%)	347 (98%)	6 (2%)	60	57
1	H	372/403 (92%)	367 (99%)	5 (1%)	69	68
All	All	2885/3224 (90%)	2838 (98%)	47 (2%)	62	60

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	74	LYS
1	A	275	GLU
1	A	328	VAL
1	A	361	ARG
1	A	441	VAL
1	B	24	GLU
1	B	29	SER
1	B	34	LEU
1	B	275	GLU
1	B	361	ARG
1	B	387[A]	ARG
1	B	387[B]	ARG
1	B	416	GLN
1	B	452	GLN
1	B	459	GLU
1	C	34	LEU
1	C	126	TYR
1	C	275	GLU
1	C	361	ARG
1	D	175	GLN
1	D	275	GLU
1	D	328	VAL
1	D	361	ARG
1	D	431	LEU
1	E	34	LEU
1	E	80	THR
1	E	150	GLN
1	E	275	GLU
1	E	361	ARG
1	F	34	LEU
1	F	126	TYR
1	F	275	GLU
1	F	361	ARG
1	F	451	GLN
1	F	459	GLU
1	G	34	LEU
1	G	126	TYR
1	G	275	GLU
1	G	328	VAL
1	G	361	ARG
1	G	422	GLU
1	H	34	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	275	GLU
1	H	328	VAL
1	H	361	ARG
1	H	431	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	167	HIS
1	B	312	GLN
1	B	452	GLN
1	D	166	HIS
1	D	167	HIS
1	D	312	GLN
1	F	452	GLN
1	H	312	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)

12 ligands 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	$\#\|Z\| > 2$
2	AMP	H	501	-	22,25,25	0.59	0	25,38,38	0.88	2 (8%)
2	AMP	G	501	-	22,25,25	0.62	0	25,38,38	0.92	2 (8%)
2	AMP	E	502	-	22,25,25	0.58	0	25,38,38	0.86	1 (4%)
2	AMP	A	501	-	22,25,25	0.63	0	25,38,38	0.96	1 (4%)
2	AMP	F	501	-	22,25,25	0.57	0	25,38,38	0.94	2 (8%)
3	SO4	G	502	-	4,4,4	0.17	0	6,6,6	0.11	0
3	SO4	C	502	-	4,4,4	0.27	0	6,6,6	0.10	0
4	GOL	E	501	-	5,5,5	0.11	0	5,5,5	0.50	0
2	AMP	D	501	-	22,25,25	0.57	0	25,38,38	0.94	2 (8%)
2	AMP	C	501	-	22,25,25	0.65	0	25,38,38	0.85	1 (4%)
2	AMP	B	501	-	22,25,25	0.67	0	25,38,38	0.89	2 (8%)
3	SO4	H	502	-	4,4,4	0.19	0	6,6,6	0.22	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
2	AMP	H	501	-	-	1/6/26/26	0/3/3/3
2	AMP	G	501	-	-	1/6/26/26	0/3/3/3
2	AMP	E	502	-	-	1/6/26/26	0/3/3/3
2	AMP	A	501	-	-	1/6/26/26	0/3/3/3
2	AMP	F	501	-	-	1/6/26/26	0/3/3/3
4	GOL	E	501	-	-	0/4/4/4	-
2	AMP	D	501	-	-	1/6/26/26	0/3/3/3
2	AMP	C	501	-	-	1/6/26/26	0/3/3/3
2	AMP	B	501	-	-	1/6/26/26	0/3/3/3

There are no bond length outliers.

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	501	AMP	P-O5'-C5'	3.06	126.72	118.30
2	G	501	AMP	P-O5'-C5'	2.62	125.50	118.30
2	G	501	AMP	C5-C6-N6	2.46	124.10	120.35
2	F	501	AMP	C5-C6-N6	2.36	123.93	120.35
2	C	501	AMP	C5-C6-N6	2.33	123.90	120.35
2	E	502	AMP	C5-C6-N6	2.32	123.88	120.35
2	H	501	AMP	P-O5'-C5'	2.31	124.66	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	AMP	C5-C6-N6	2.26	123.78	120.35
2	D	501	AMP	P-O5'-C5'	2.25	124.50	118.30
2	D	501	AMP	C5-C6-N6	2.22	123.73	120.35
2	B	501	AMP	C5-C6-N6	2.21	123.71	120.35
2	B	501	AMP	P-O5'-C5'	2.21	124.38	118.30
2	H	501	AMP	C5-C6-N6	2.04	123.45	120.35

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	AMP	O4'-C4'-C5'-O5'
2	F	501	AMP	O4'-C4'-C5'-O5'
2	E	502	AMP	O4'-C4'-C5'-O5'
2	G	501	AMP	O4'-C4'-C5'-O5'
2	A	501	AMP	O4'-C4'-C5'-O5'
2	C	501	AMP	O4'-C4'-C5'-O5'
2	D	501	AMP	O4'-C4'-C5'-O5'
2	H	501	AMP	O4'-C4'-C5'-O5'

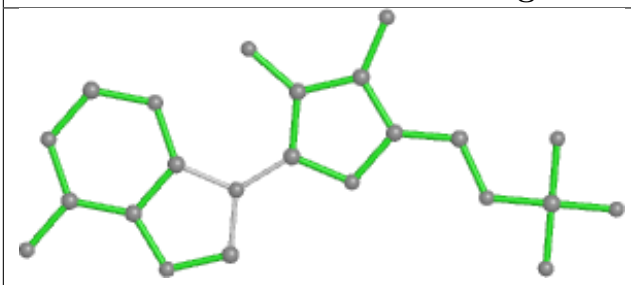
There are no ring outliers.

1 monomer is involved in 4 short contacts:

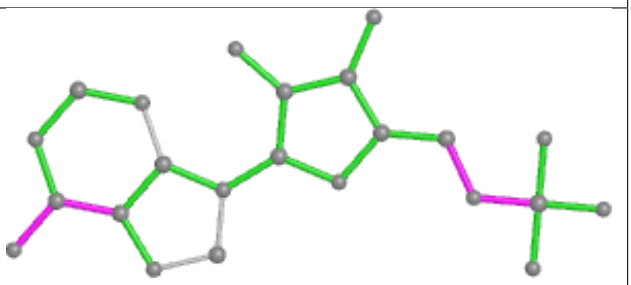
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	501	GOL	4	0

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.

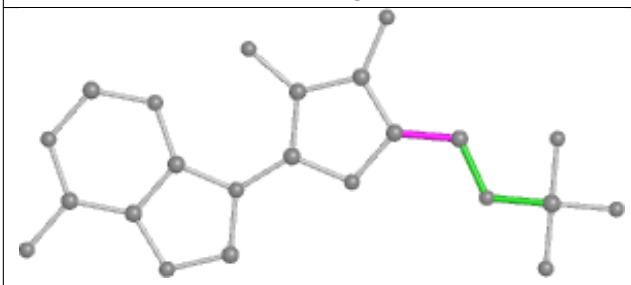
Ligand AMP H 501



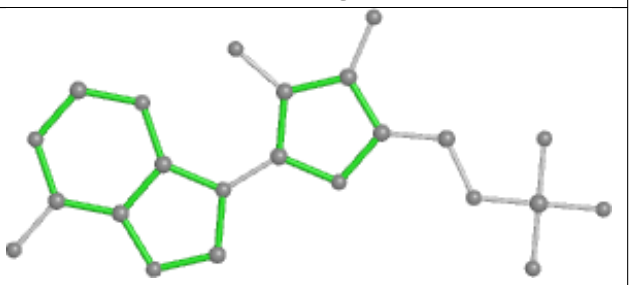
Bond lengths



Bond angles

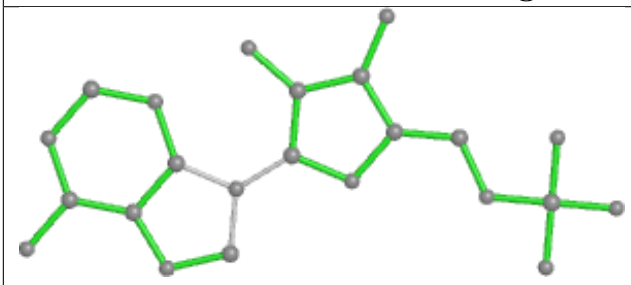


Torsions

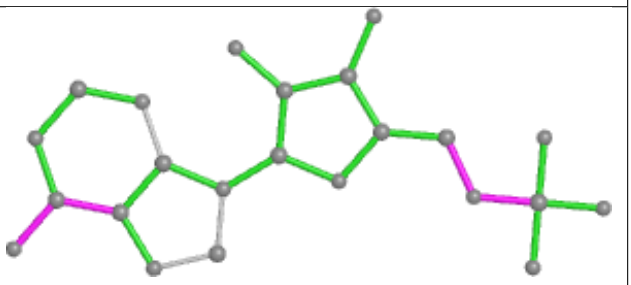


Rings

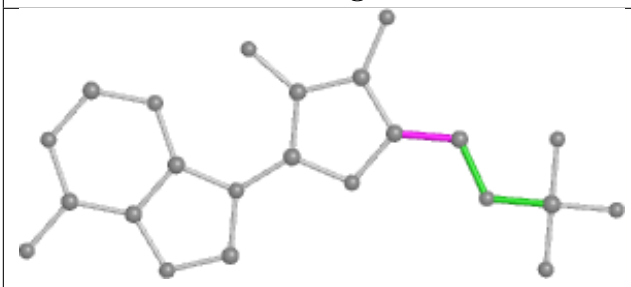
Ligand AMP G 501



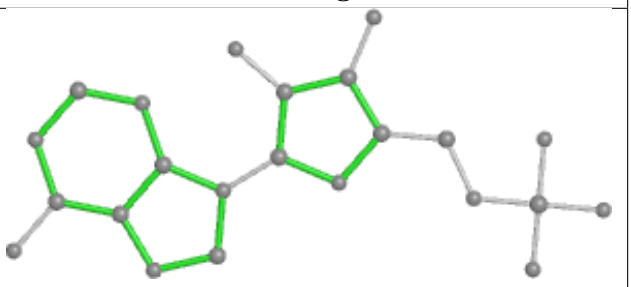
Bond lengths



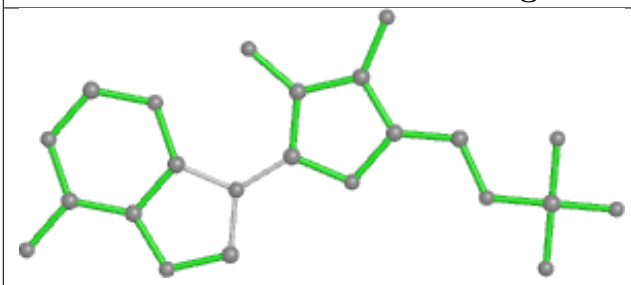
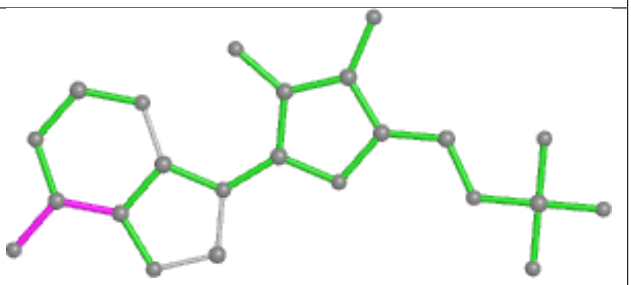
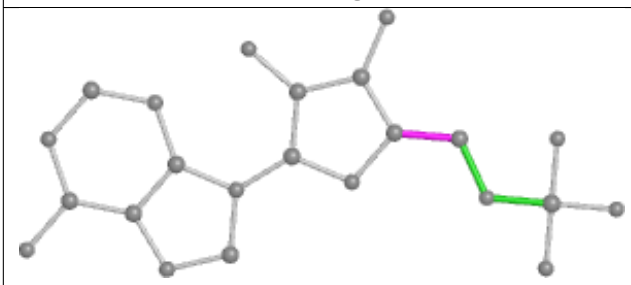
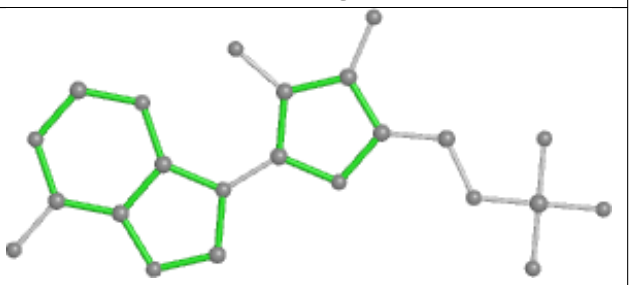
Bond angles

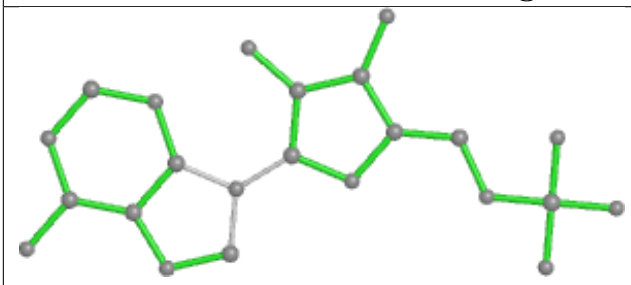
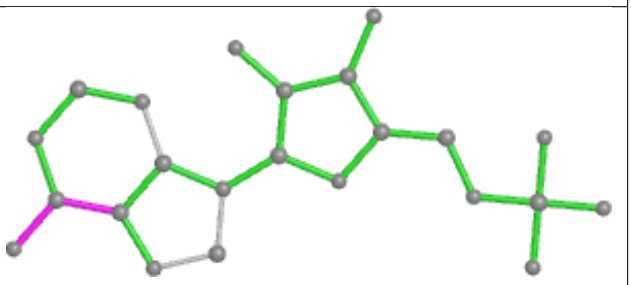
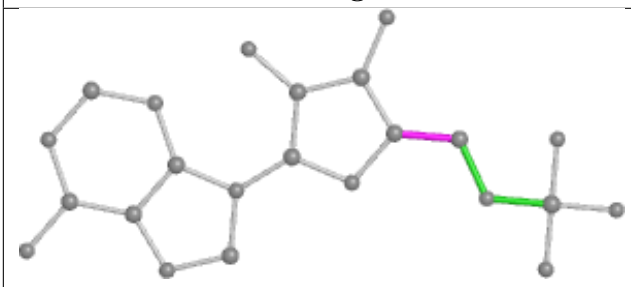
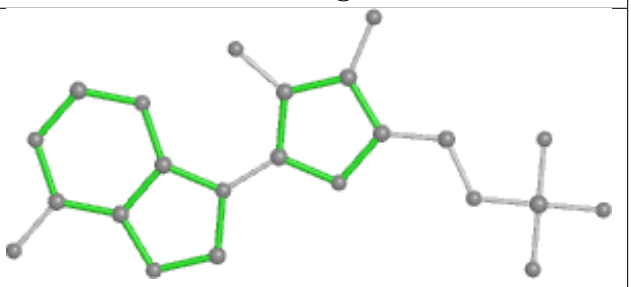


Torsions

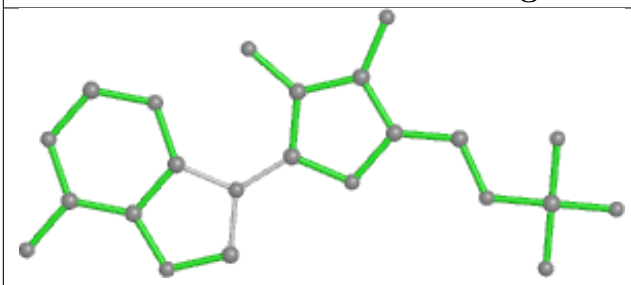


Rings

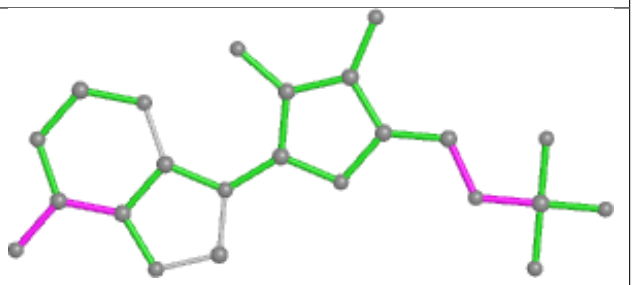
Ligand AMP E 502	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand AMP A 501	
 Bond lengths	 Bond angles
 Torsions	 Rings

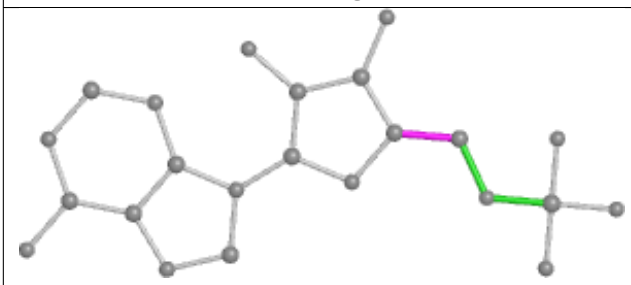
Ligand AMP F 501



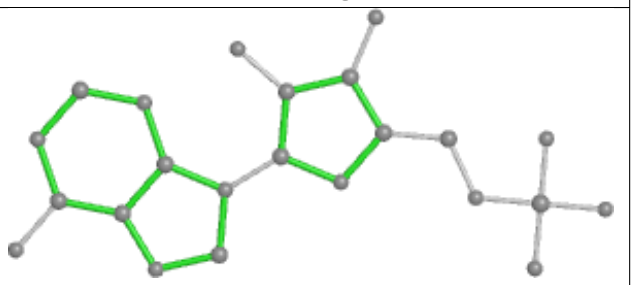
Bond lengths



Bond angles

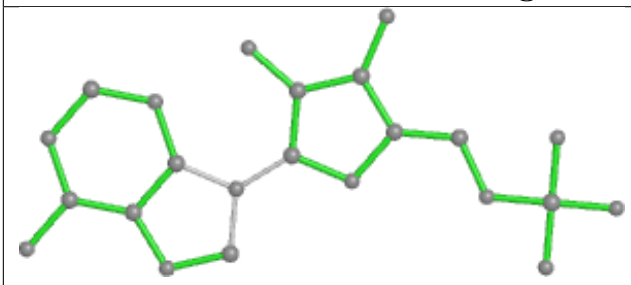


Torsions

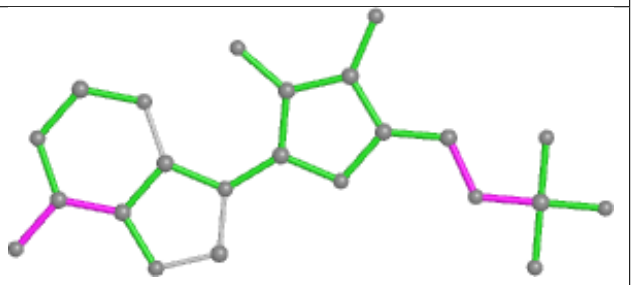


Rings

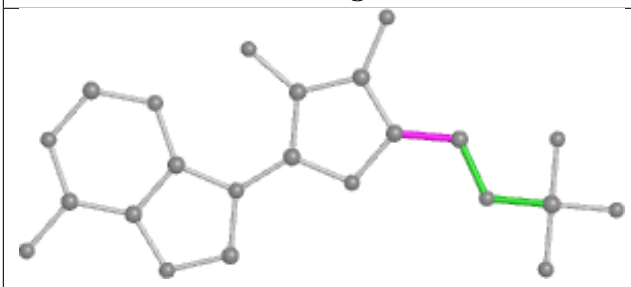
Ligand AMP D 501



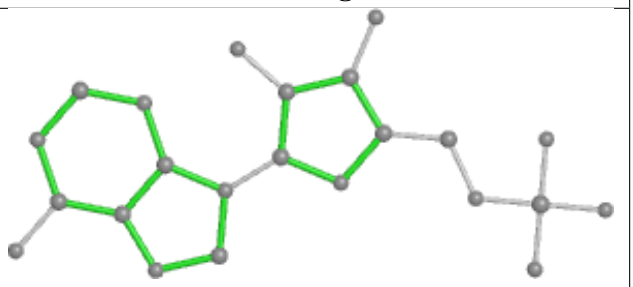
Bond lengths



Bond angles

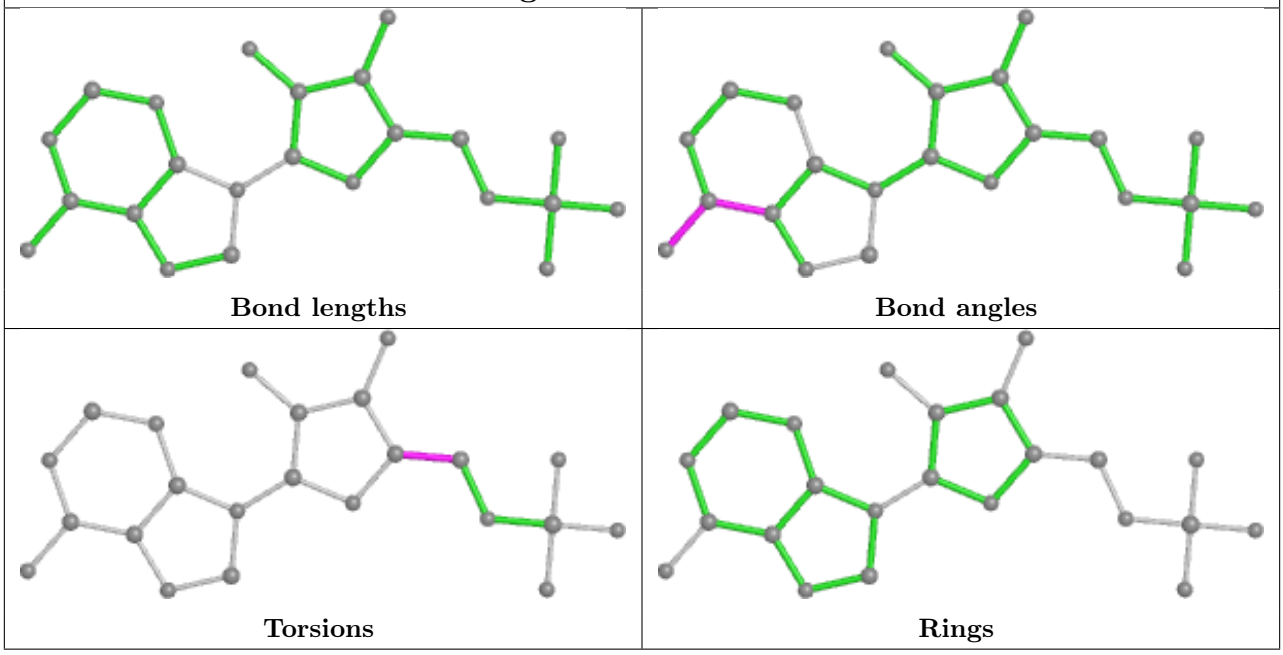


Torsions

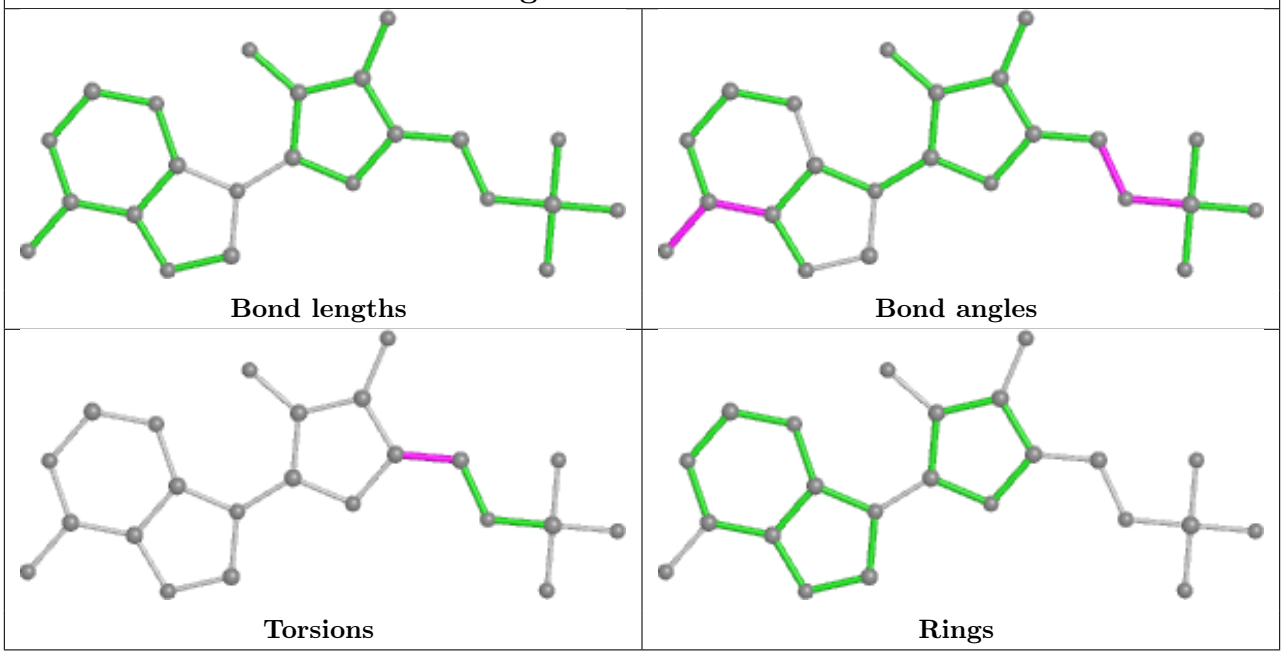


Rings

Ligand AMP C 501



Ligand AMP B 501



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	424/480 (88%)	0.04	9 (2%)	63	66	9, 17, 43, 83	0
1	B	439/480 (91%)	-0.03	4 (0%)	84	85	8, 16, 40, 75	0
1	C	426/480 (88%)	0.13	13 (3%)	49	51	9, 18, 47, 75	0
1	D	439/480 (91%)	-0.02	4 (0%)	84	85	8, 16, 38, 64	0
1	E	425/480 (88%)	0.14	13 (3%)	49	51	10, 19, 46, 74	0
1	F	431/480 (89%)	0.05	6 (1%)	75	77	9, 17, 38, 81	0
1	G	427/480 (88%)	0.13	10 (2%)	60	63	9, 19, 43, 77	0
1	H	438/480 (91%)	0.01	9 (2%)	63	66	8, 16, 39, 72	0
All	All	3449/3840 (89%)	0.06	68 (1%)	65	68	8, 17, 43, 83	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	314	VAL	6.8
1	A	314	VAL	6.6
1	B	314	VAL	6.2
1	A	319	VAL	5.7
1	G	314	VAL	5.0
1	H	314	VAL	4.9
1	C	453	LEU	4.2
1	A	419	LEU	4.0
1	A	450	VAL	3.8
1	E	319	VAL	3.8
1	E	449	PHE	3.7
1	E	450	VAL	3.7
1	C	319	VAL	3.6
1	F	328	VAL	3.6
1	A	449	PHE	3.4
1	G	102	ALA	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	460	VAL	3.1
1	C	449	PHE	3.1
1	H	34	LEU	3.0
1	C	320	GLY	3.0
1	F	318	GLU	3.0
1	A	100	ASN	3.0
1	D	34	LEU	2.9
1	E	419	LEU	2.9
1	C	464	LEU	2.9
1	E	447	HIS	2.8
1	F	315	ARG	2.7
1	G	419	LEU	2.7
1	E	100	ASN	2.6
1	E	314	VAL	2.6
1	G	318	GLU	2.6
1	H	315	ARG	2.6
1	E	444	GLU	2.6
1	E	456	ILE	2.6
1	C	444	GLU	2.5
1	C	314	VAL	2.5
1	A	125	TYR	2.4
1	A	460	VAL	2.4
1	D	328	VAL	2.4
1	H	429	GLU	2.3
1	H	430	ALA	2.3
1	B	133	ALA	2.3
1	D	453	LEU	2.3
1	H	134[A]	CYS	2.3
1	G	69	PHE	2.2
1	G	446	VAL	2.2
1	G	319	VAL	2.2
1	G	441	VAL	2.2
1	C	456	ILE	2.2
1	E	34	LEU	2.2
1	C	318	GLU	2.2
1	D	125	TYR	2.2
1	C	465	LEU	2.2
1	A	409	ALA	2.2
1	H	319	VAL	2.2
1	C	243	LEU	2.1
1	F	102	ALA	2.1
1	B	315	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	125	TYR	2.1
1	G	447	HIS	2.1
1	B	328	VAL	2.1
1	F	319	VAL	2.1
1	H	456	ILE	2.1
1	C	447	HIS	2.1
1	E	125	TYR	2.1
1	G	100	ASN	2.0
1	E	446	VAL	2.0
1	C	328	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)

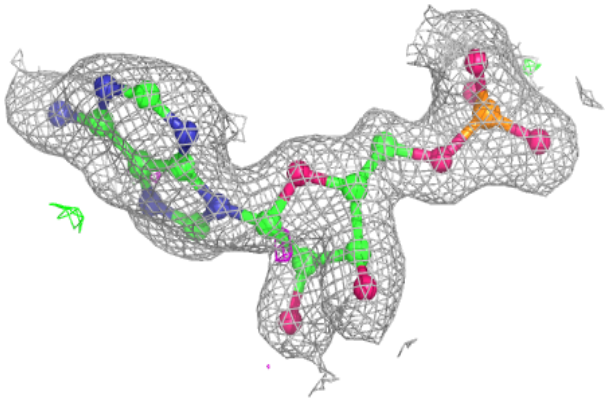
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

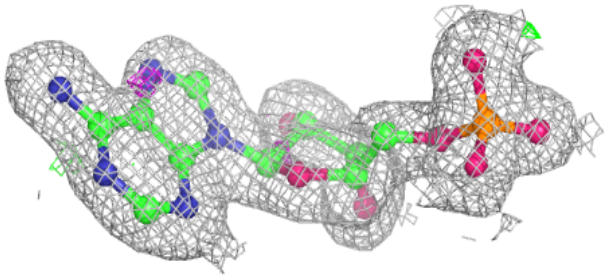
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	E	501	6/6	0.83	0.17	26,27,30,31	0
2	AMP	E	502	23/23	0.93	0.12	15,22,23,24	0
2	AMP	G	501	23/23	0.94	0.08	11,15,19,20	0
3	SO4	H	502	5/5	0.94	0.10	44,47,48,48	0
2	AMP	C	501	23/23	0.94	0.09	8,14,18,20	0
2	AMP	F	501	23/23	0.95	0.09	14,19,22,23	0
2	AMP	A	501	23/23	0.96	0.09	8,14,16,17	0
2	AMP	H	501	23/23	0.96	0.08	11,18,20,21	0
3	SO4	C	502	5/5	0.97	0.10	29,30,32,32	0
3	SO4	G	502	5/5	0.97	0.07	30,34,36,38	0
2	AMP	B	501	23/23	0.97	0.08	8,16,19,19	0
2	AMP	D	501	23/23	0.97	0.09	10,17,19,20	0

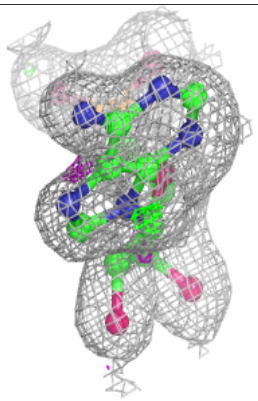
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around AMP 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)

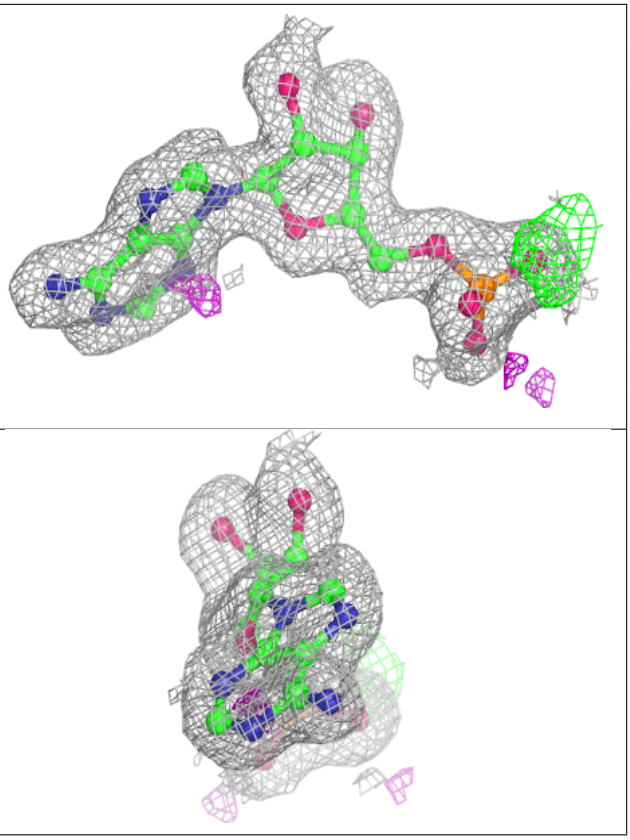






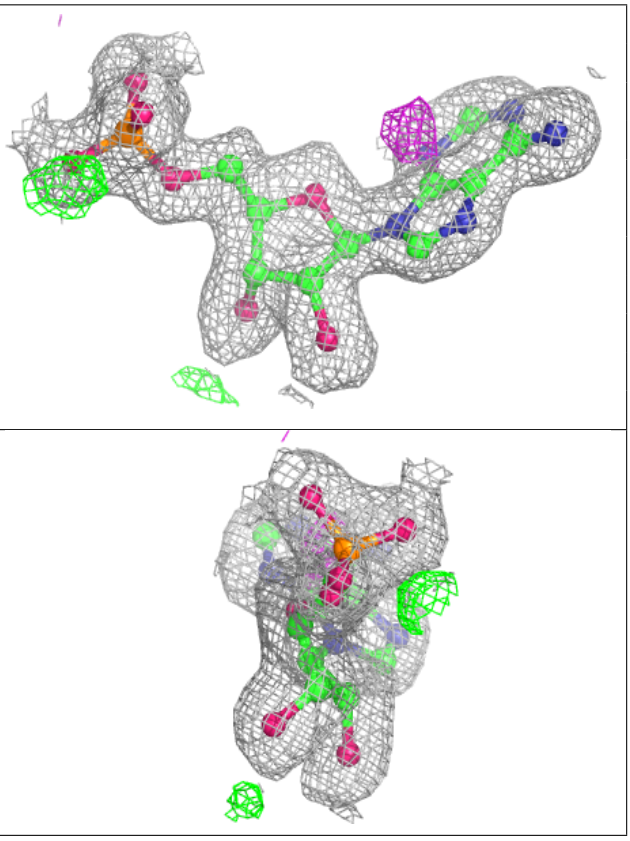
Electron density around AMP G 501:

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



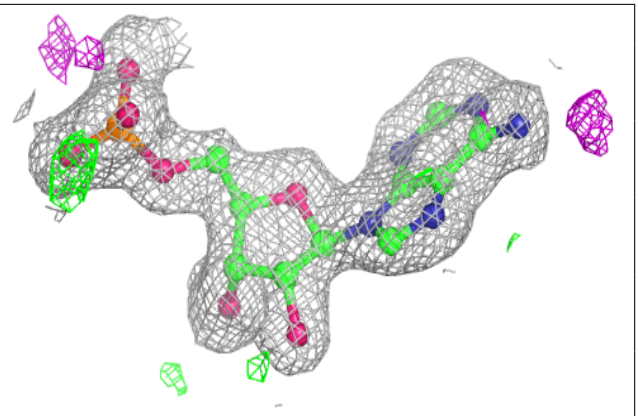
Electron density around AMP C 501:

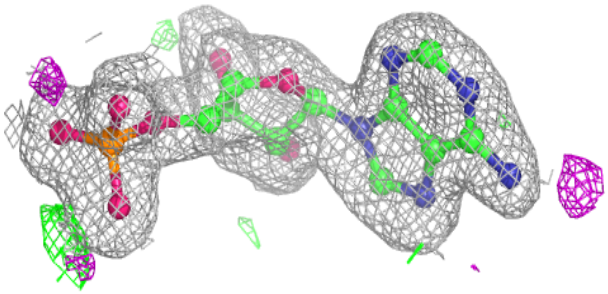
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

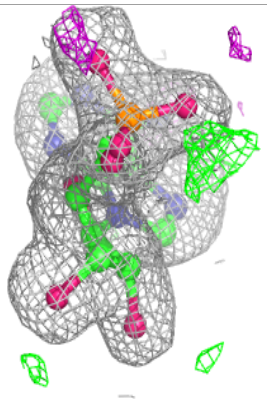


Electron density around AMP F 501:

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

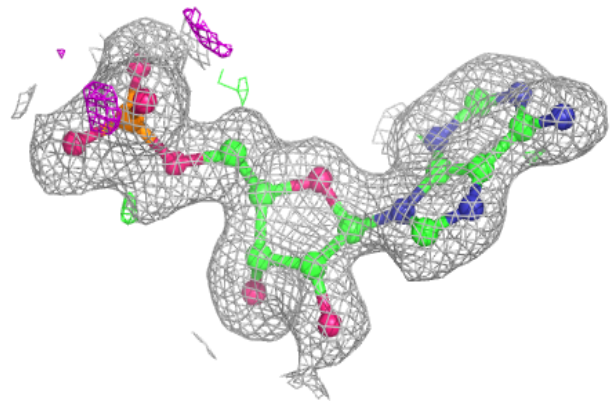


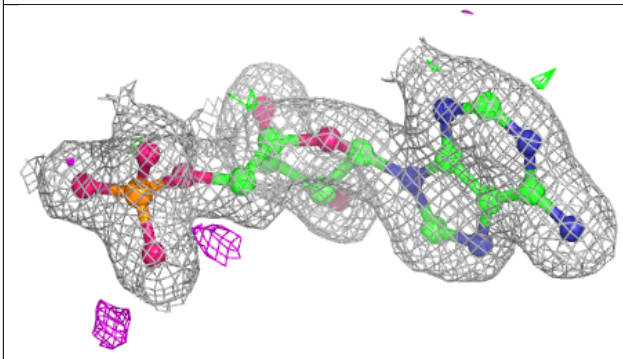


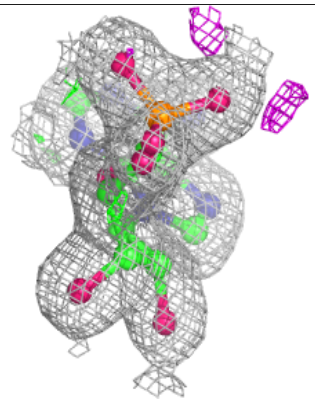


Electron density around AMP A 501:

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

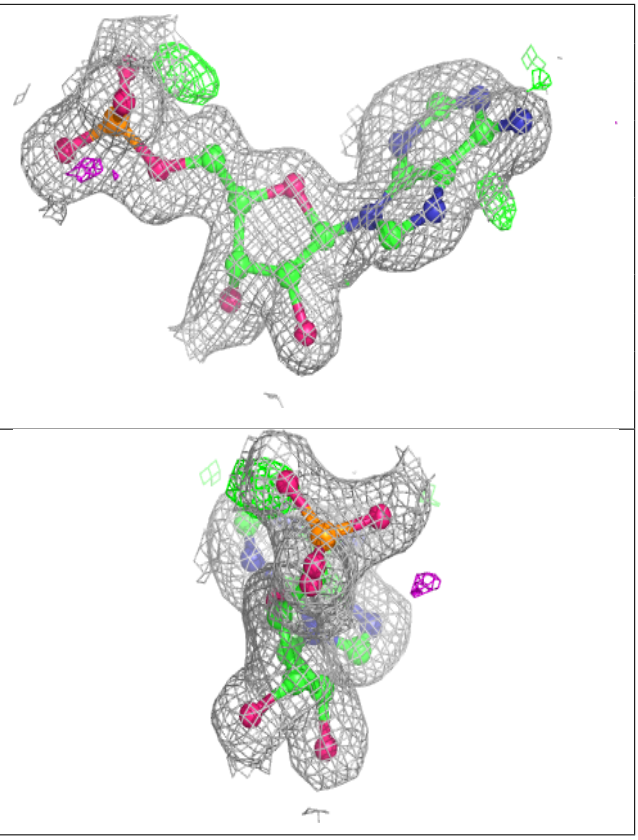


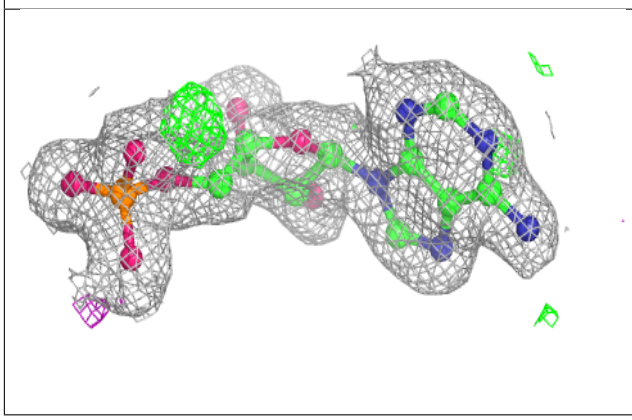




Electron density around AMP H 501:

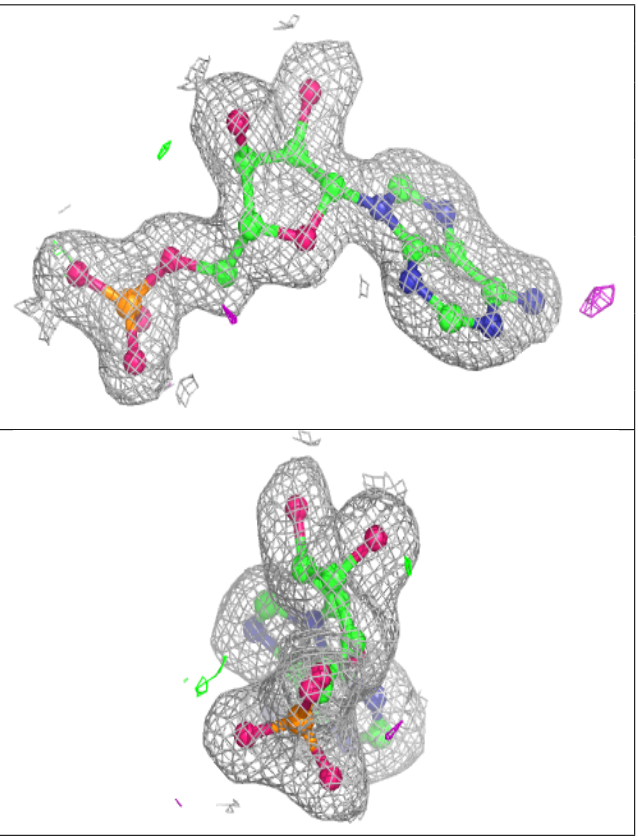
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

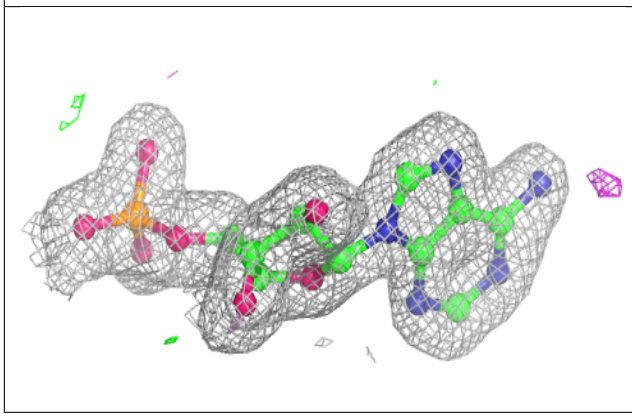




Electron density around AMP B 501:

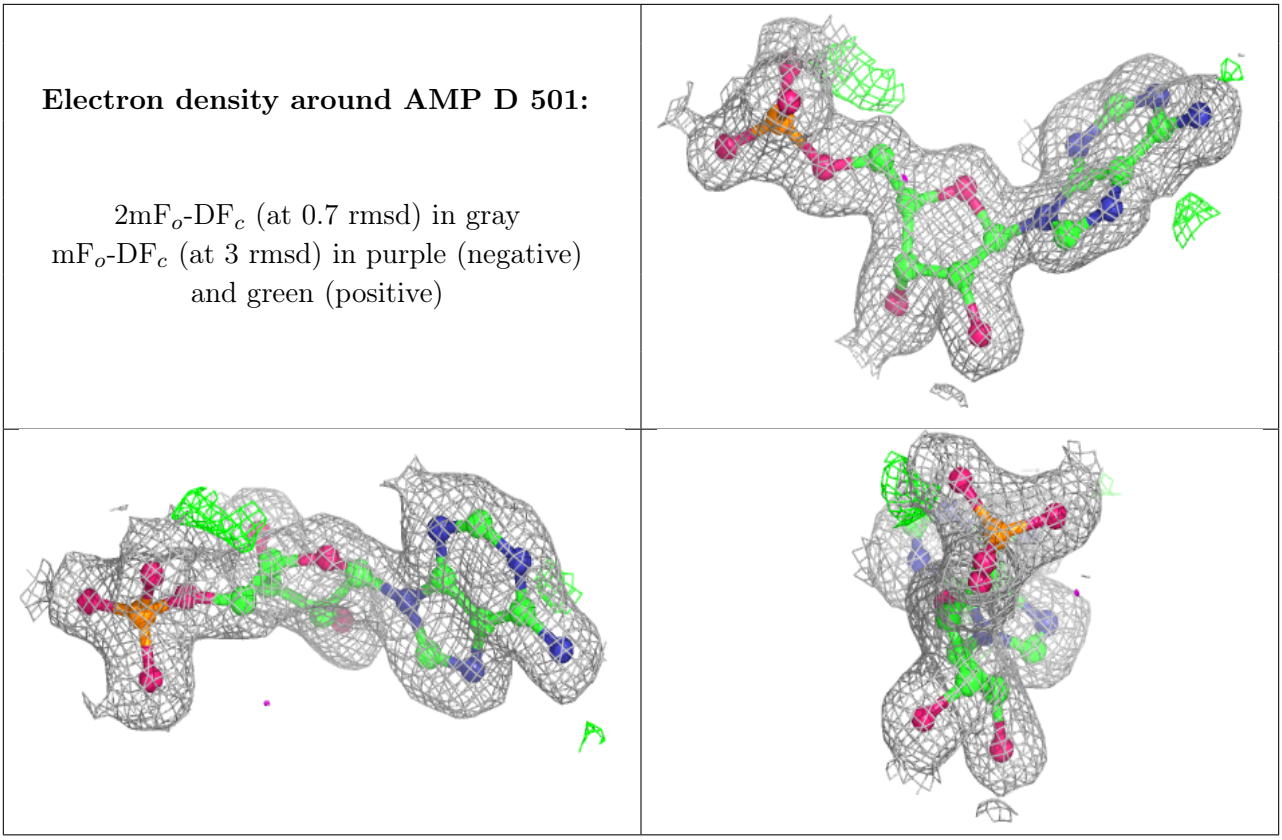
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





Electron density around AMP 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)



6.5 Other polymers (i)

There are no such residues in this entry.