wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2023 – 02:35 pm GMT

PDB ID	:	1H74
Title	:	CRYSTAL STRUCTURE OF HOMOSERINE KINASE COMPLEXED WITH ILE
Authors	:	Krishna, S.S.; Zhou, T.; Daugherty, M.; Osterman, A.L.; Zhang, H.
Deposited on	:	2001-07-02
Resolution	:	1.90 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:

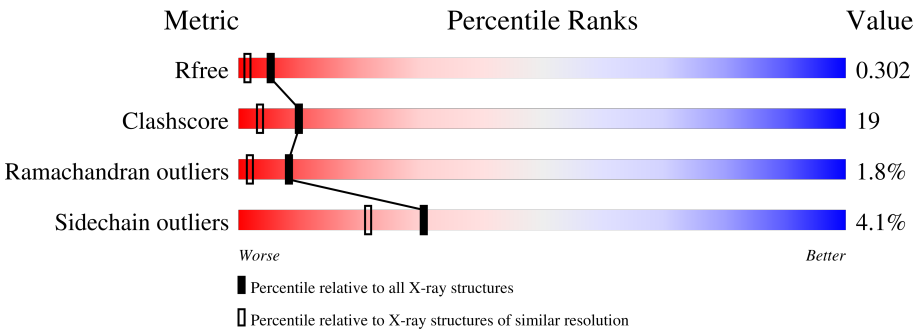
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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)

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%

Mol	Chain	Length	Quality of chain
1	A	296	72% 25% •
1	B	296	60% 34% 6%
1	C	296	80% 19% •
1	D	296	76% 22% •

2 Entry composition (i)

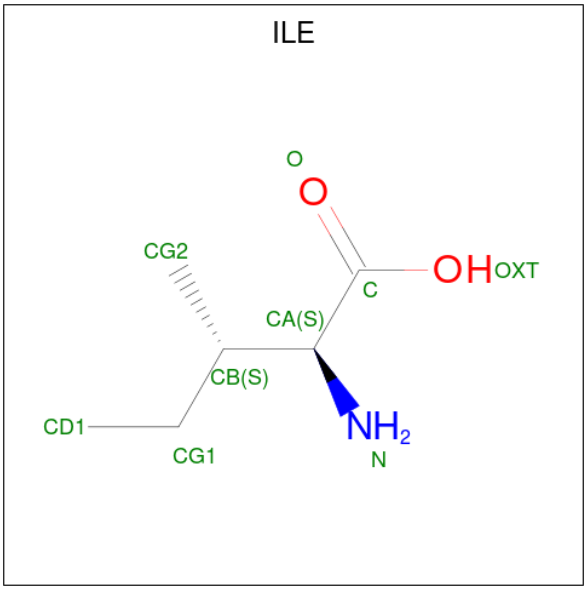
There are 6 unique types of molecules in this entry. The entry contains 9718 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HOMOSERINE KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	296	2293	1472	370	442	9	0	4	0
1	B	296	2281	1466	368	438	9	0	2	0
1	C	296	2275	1463	367	436	9	0	1	0
1	D	296	2275	1463	367	436	9	0	1	0

- Molecule 2 is ISOLEUCINE (three-letter code: ILE) (formula: $C_6H_{13}NO_2$).



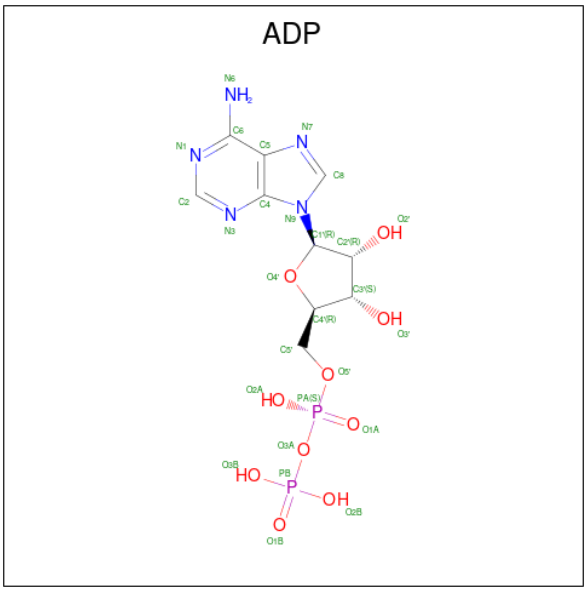
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	9	6	1	2	0	0
2	B	1	9	6	1	2	0	0

Continued on next page...

Continued from previous page...

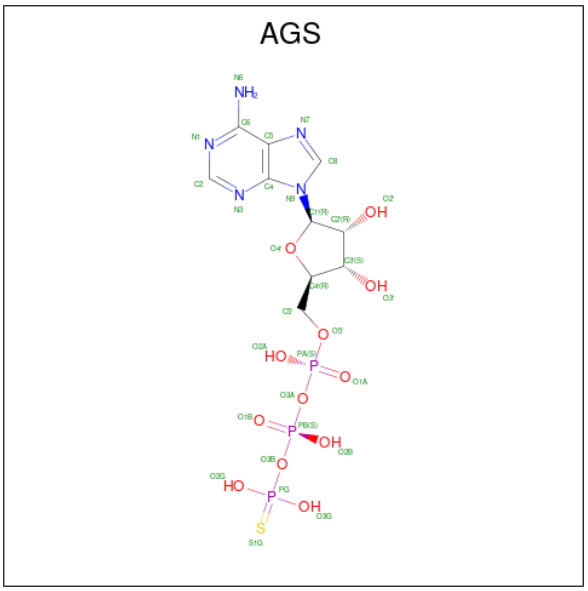
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	C	1	9	6	1	2	0	0
			Total	C	N	O		
2	D	1	9	6	1	2	0	0

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	27	10	5	10	2	0	0
			Total	C	N	O	P		
3	C	1	27	10	5	10	2	0	0

- Molecule 4 is PHOSPHOTHIOPHOSPHORIC ACID-ADENYLATE ESTER (three-letter code: AGS) (formula: $C_{10}H_{16}N_5O_{12}P_3S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
4	B	1	31	10	5	12	3	1	0	0
4	D	1	31	10	5	12	3	1	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	B	1	1	1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	105	105	105	0	0
6	B	73	73	73	0	0
6	C	157	157	157	0	0
6	D	106	106	106	0	0

3 Residue-property plots (i)

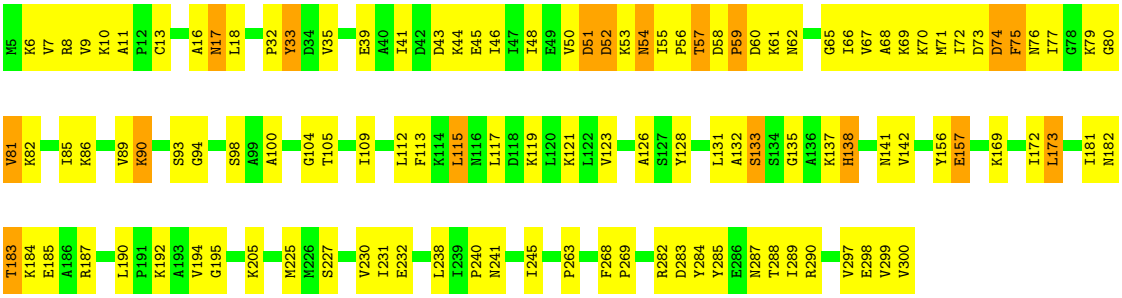
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: HOMOSERINE KINASE

Chain A: 72% 25% •

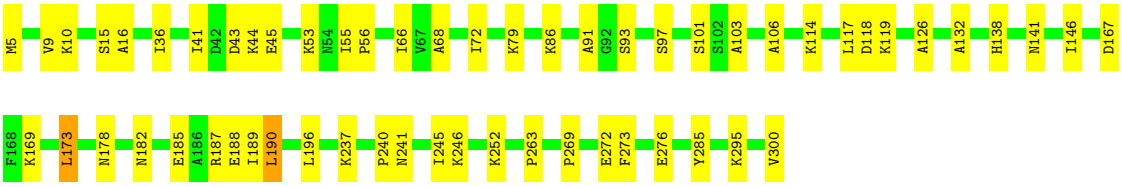
• Molecule 1: HOMOSERINE KINASE

Chain B: 60% 34% 6%



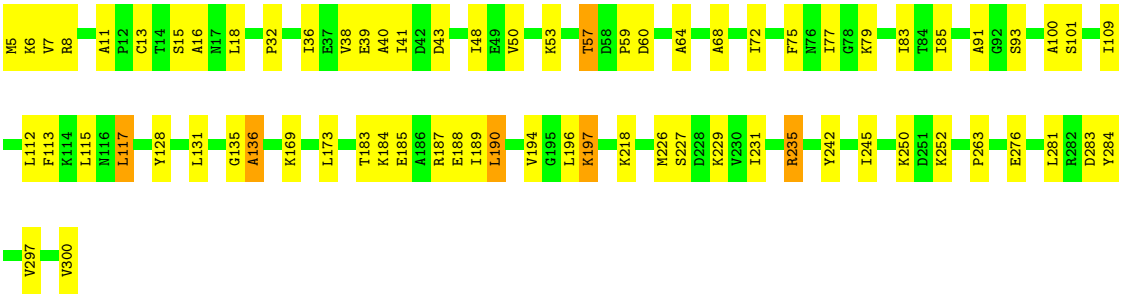
• Molecule 1: HOMOSERINE KINASE

Chain C: 80% 19% •



• Molecule 1: HOMOSERINE KINASE

Chain D: 76% 22% •


M5
K6
V7
R8
A11
P12
C13
T14
S15
A16
N17
L18
P32
I36
E37
V38
E39
A40
I41
D42
D43
I48
E49
V50
K53
T57
D58
P59
D60
A64
A68
I72
F75
N76
I77
G78
K79
I83
T84
I85
A91
G92
S93
A100
S101
I109
L112
F113
K114
L115
N116
L117
Y128
L131
G135
A136
K169
L173
T183
K184
E185
A186
R187
E188
I189
L190
V194
G195
L196
K197
K218
M226
S227
D228
K229
V230
I231
R235
Y242
I245
K250
D251
K252
P263
E276
L281
R282
D283
Y284
V297
V300

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	61.99Å 128.72Å 109.42Å 90.00° 105.57° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 48.77 – 1.89	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.00-1.90) 98.8 (48.77-1.89)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 1.88Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.210 , 0.238 0.283 , 0.302	Depositor DCC
R_{free} test set	6521 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	34.0	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 45.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.033 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9718	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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: AGS, MG, ADP

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	0.68	2/2330 (0.1%)	0.79	2/3144 (0.1%)
1	B	0.53	0/2318	0.71	0/3128
1	C	0.79	3/2312 (0.1%)	0.80	1/3120 (0.0%)
1	D	0.71	2/2312 (0.1%)	0.76	0/3120
All	All	0.68	7/9272 (0.1%)	0.77	3/12512 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	15[A]	SER	CB-OG	9.13	1.54	1.42
1	C	15[B]	SER	CB-OG	9.13	1.54	1.42
1	D	15[A]	SER	CB-OG	6.49	1.50	1.42
1	D	15[B]	SER	CB-OG	6.49	1.50	1.42
1	A	15[A]	SER	CB-OG	6.41	1.50	1.42

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	80	GLY	N-CA-C	-6.26	97.46	113.10
1	C	118	ASP	CB-CG-OD2	5.63	123.37	118.30
1	A	235	ARG	NE-CZ-NH1	-5.17	117.72	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	128	TYR	Sidechain

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	2293	0	2355	69	0
1	B	2281	0	2347	151	1
1	C	2275	0	2343	55	1
1	D	2275	0	2343	79	0
2	A	9	0	10	2	0
2	B	9	0	10	4	0
2	C	9	0	10	5	0
2	D	9	0	10	5	0
3	A	27	0	12	3	0
3	C	27	0	12	2	0
4	B	31	0	12	3	0
4	D	31	0	12	2	0
5	B	1	0	0	0	0
6	A	105	0	0	19	0
6	B	73	0	0	31	1
6	C	157	0	0	26	1
6	D	106	0	0	38	0
All	All	9718	0	9476	359	2

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 19.

The worst 5 of 359 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:46:ILE:HA	6:B:2011:HOH:O	1.48	1.12
1:C:169:LYS:HD2	6:C:2097:HOH:O	1.54	1.06
1:B:288:THR:HB	6:B:2067:HOH:O	1.55	1.04

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:141:ASN:HD21	2:B:500:ILE:HD13	1.19	1.04
1:C:246:LYS:HB2	6:C:2129:HOH:O	1.58	1.03

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:B:2069:HOH:O	6:C:2127:HOH:O[2_747]	1.90	0.30
1:B:283:ASP:O	1:C:237:LYS:NZ[2_747]	2.09	0.11

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	298/296 (101%)	283 (95%)	9 (3%)	6 (2%)	7	1
1	B	296/296 (100%)	252 (85%)	31 (10%)	13 (4%)	2	0
1	C	295/296 (100%)	287 (97%)	8 (3%)	0	100	100
1	D	295/296 (100%)	279 (95%)	14 (5%)	2 (1%)	22	12
All	All	1184/1184 (100%)	1101 (93%)	62 (5%)	21 (2%)	8	2

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	115	LEU
1	B	133	SER
1	B	138	HIS
1	A	183	THR
1	B	44	LYS

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	252/248 (102%)	237 (94%)	15 (6%)	19	9
1	B	250/248 (101%)	240 (96%)	10 (4%)	31	22
1	C	249/248 (100%)	243 (98%)	6 (2%)	49	43
1	D	249/248 (100%)	239 (96%)	10 (4%)	31	22
All	All	1000/992 (101%)	959 (96%)	41 (4%)	30	21

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	173	LEU
1	D	190	LEU
1	C	190	LEU
1	D	57	THR
1	D	196	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	163	HIS
1	C	141	ASN
1	D	138	HIS
1	B	138	HIS
1	A	178	ASN

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 9 ligands modelled in this entry, 1 is 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	$\#\|Z\| > 2$
2	ILE	C	500	-	7,8,8	0.82	0	7,10,10	0.78	0
2	ILE	B	500	-	7,8,8	0.68	0	7,10,10	0.68	0
4	AGS	D	400	-	26,33,33	1.45	5 (19%)	26,52,52	1.43	4 (15%)
2	ILE	D	500	-	7,8,8	0.79	0	7,10,10	0.63	0
4	AGS	B	400	5	26,33,33	1.38	3 (11%)	26,52,52	1.50	4 (15%)
2	ILE	A	500	-	7,8,8	0.85	0	7,10,10	0.66	0
3	ADP	A	400	-	24,29,29	1.58	4 (16%)	29,45,45	1.63	4 (13%)
3	ADP	C	400	-	24,29,29	1.46	2 (8%)	29,45,45	1.47	4 (13%)

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	ILE	C	500	-	-	7/10/10/10	-
2	ILE	B	500	-	-	0/10/10/10	-
4	AGS	D	400	-	-	3/17/38/38	0/3/3/3
2	ILE	D	500	-	-	0/10/10/10	-
4	AGS	B	400	5	-	4/17/38/38	0/3/3/3
2	ILE	A	500	-	-	0/10/10/10	-
3	ADP	A	400	-	-	4/12/32/32	0/3/3/3
3	ADP	C	400	-	-	5/12/32/32	0/3/3/3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	400	ADP	C4-N3	4.88	1.42	1.35
4	D	400	AGS	PG-S1G	4.13	1.99	1.90
3	A	400	ADP	C4-N3	4.02	1.41	1.35
3	A	400	ADP	C2-N3	3.67	1.38	1.32
4	B	400	AGS	C2-N3	3.52	1.37	1.32

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	400	ADP	C1'-N9-C4	5.71	136.67	126.64
4	B	400	AGS	C1'-N9-C4	5.08	135.56	126.64
4	D	400	AGS	C1'-N9-C4	4.92	135.28	126.64
3	C	400	ADP	C1'-N9-C4	4.48	134.52	126.64
4	B	400	AGS	O2G-PG-O3B	2.55	113.14	104.64

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	500	ILE	N-CA-CB-CG1
2	C	500	ILE	N-CA-CB-CG2
2	C	500	ILE	C-CA-CB-CG1
2	C	500	ILE	C-CA-CB-CG2
3	A	400	ADP	PA-O3A-PB-O2B

There are no ring outliers.

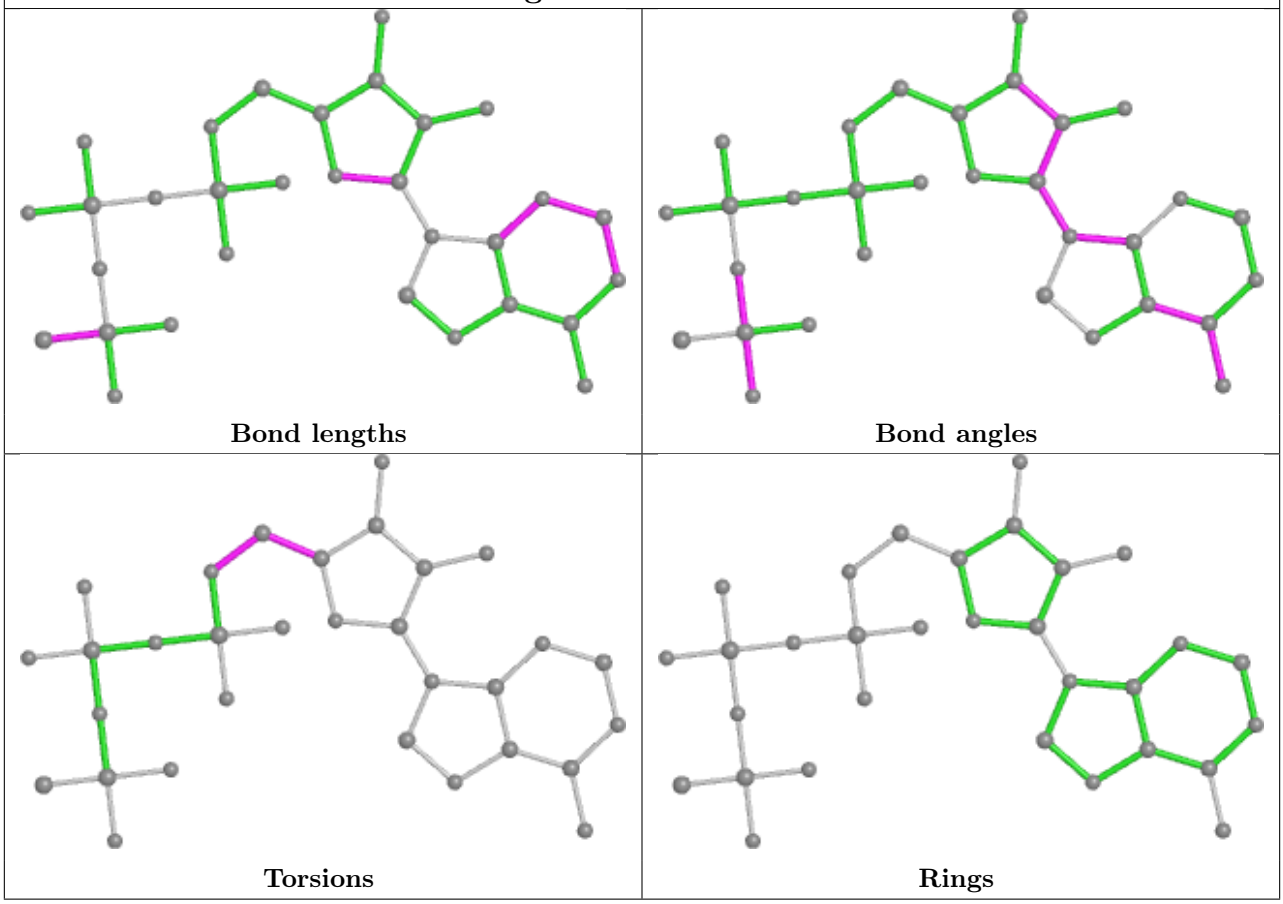
8 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	500	ILE	5	0
2	B	500	ILE	4	0
4	D	400	AGS	2	0
2	D	500	ILE	5	0
4	B	400	AGS	3	0
2	A	500	ILE	2	0
3	A	400	ADP	3	0
3	C	400	ADP	2	0

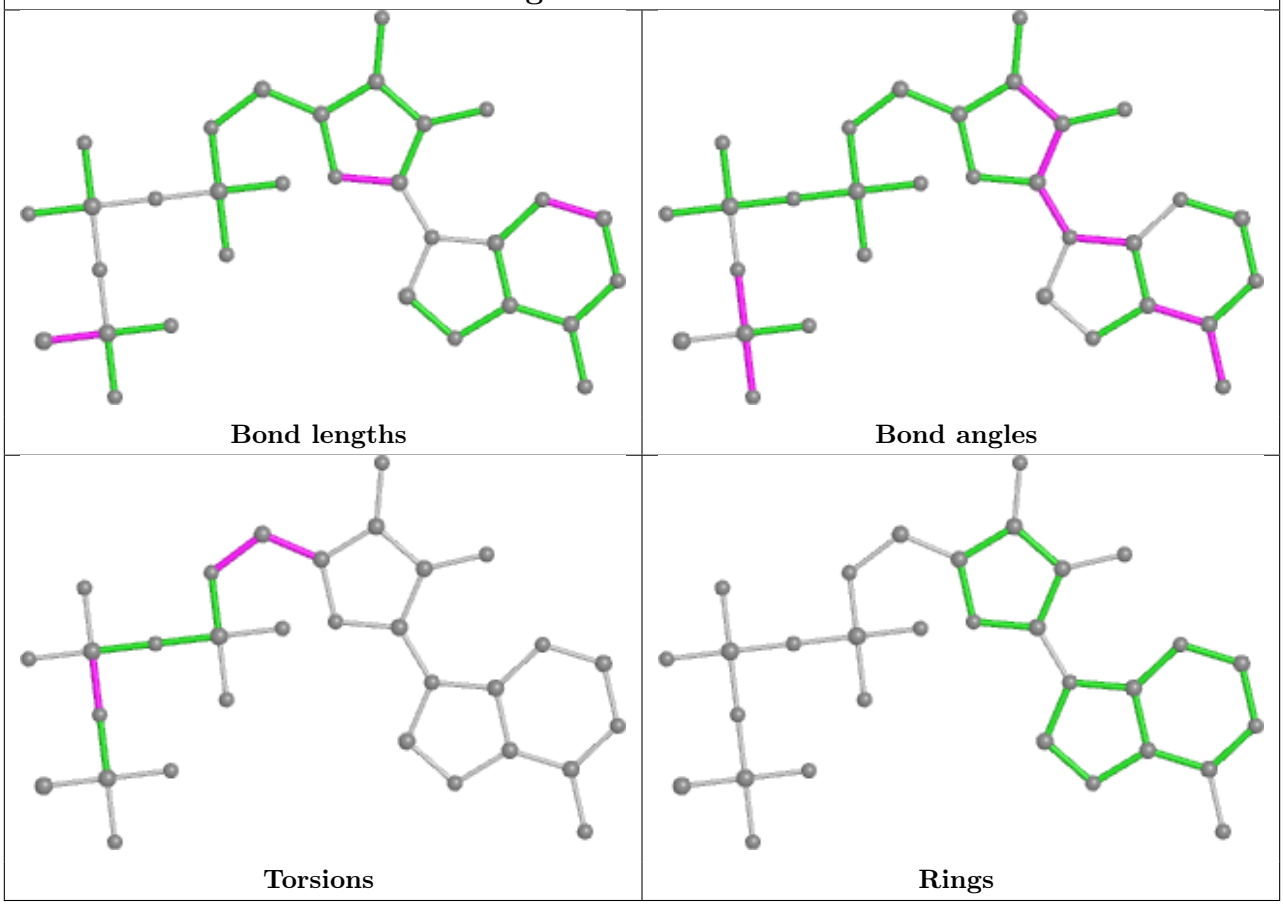
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

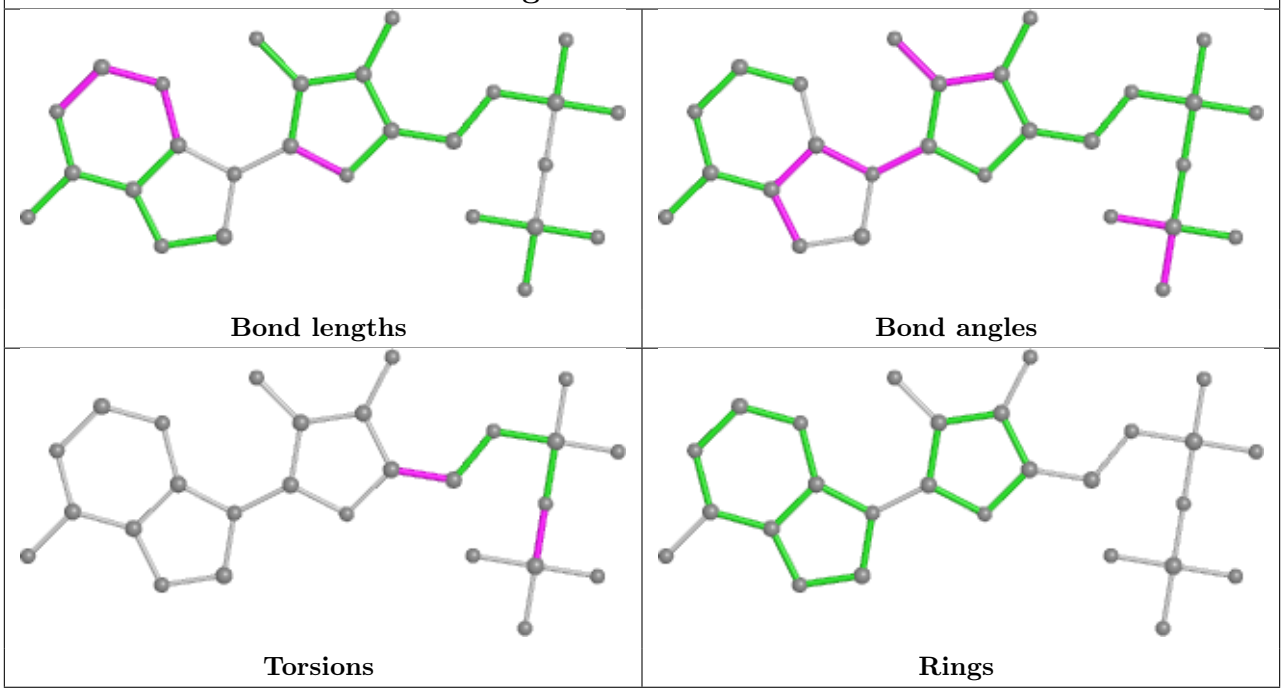
Ligand AGS D 400



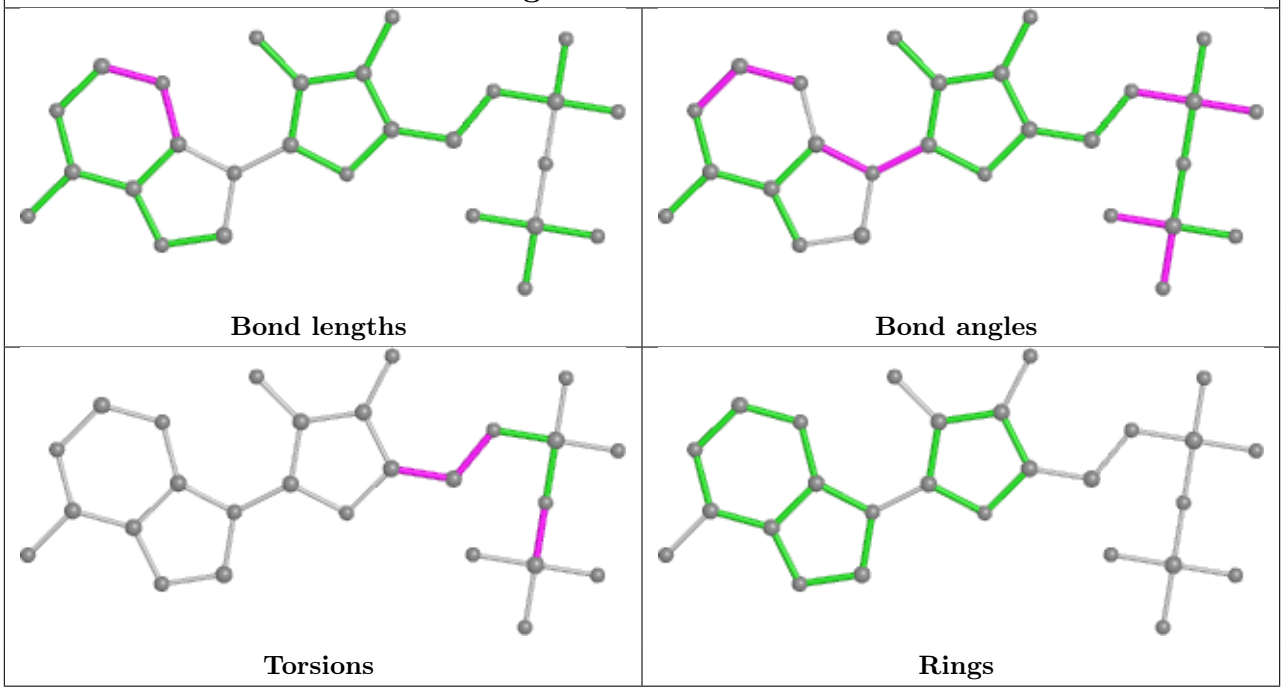
Ligand AGS B 400



Ligand ADP A 400



Ligand ADP C 400



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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

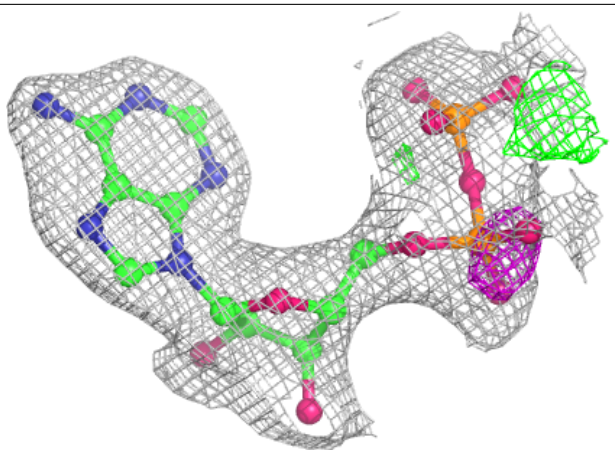
6.4 Ligands (i)

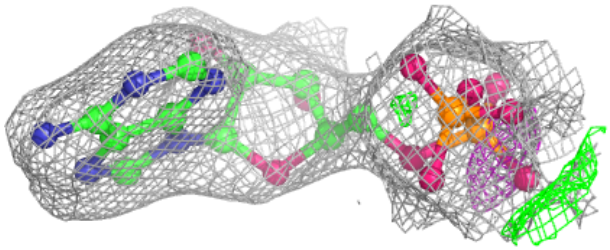
Unable to reproduce the depositors R factor - this section is therefore empty.

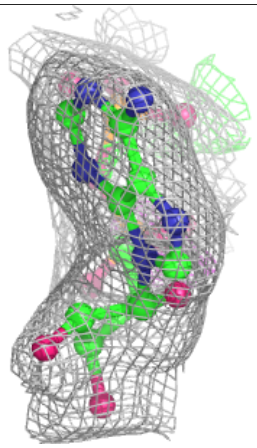
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ADP A 400:

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

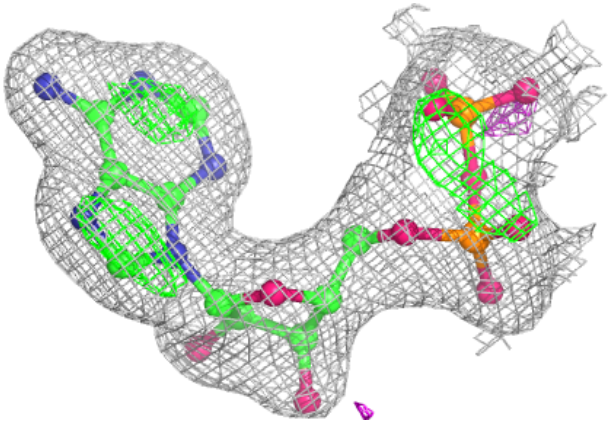


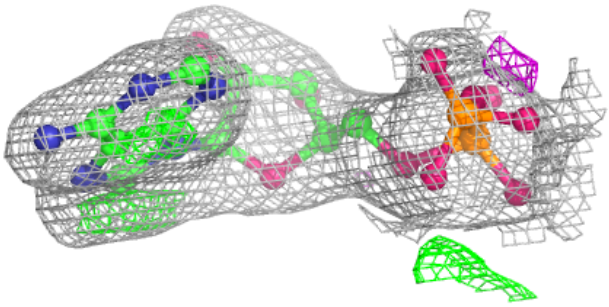


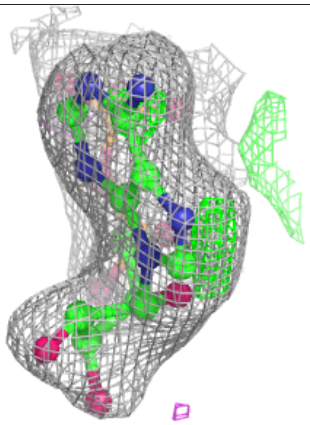


Electron density around ADP C 400:

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

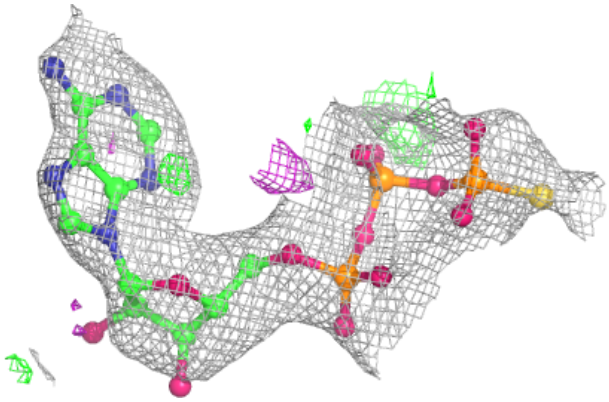


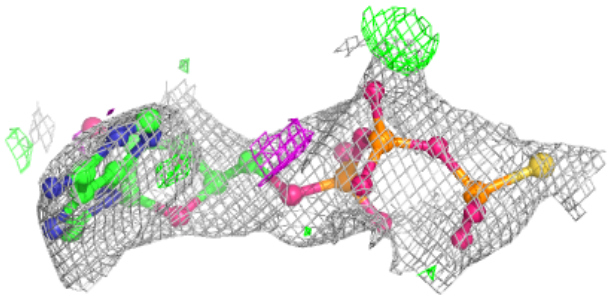


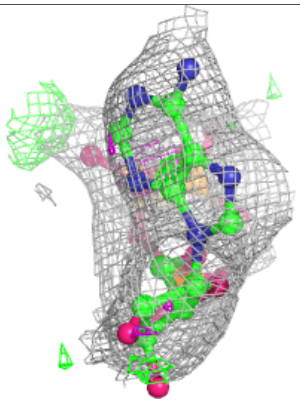


Electron density around AGS B 400:

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

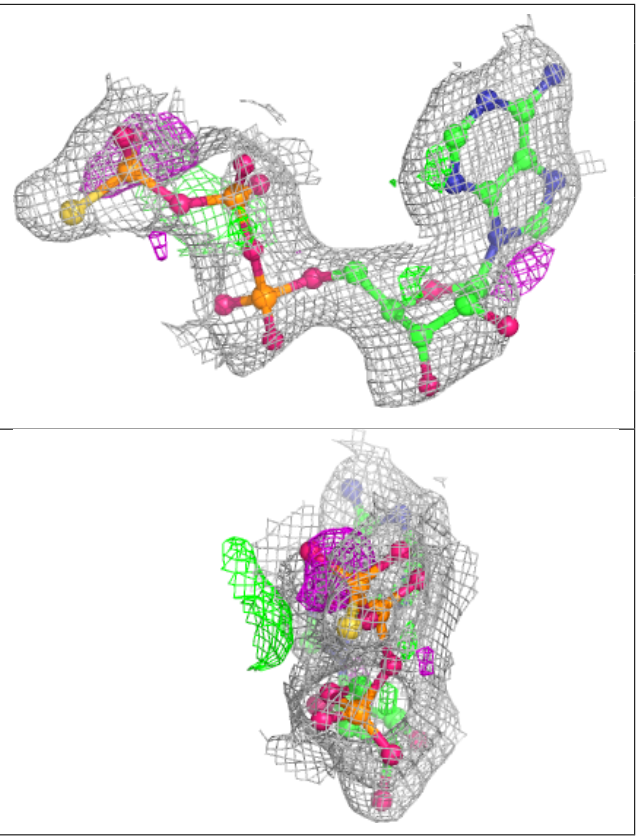






Electron density around AGS D 400:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.