WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 02:12 pm BST

PDB ID : 1KI3

Title : CRYSTAL STRUCTURE OF THYMIDINE KINASE FROM HERPES SIM-
PLEX VIRUS TYPE I COMPLEXED WITH PENCICLOVIR

Authors : Champness, J.N.; Bennett, M.S.; Wien, F.; Visse, R.; Jarvest, R.L.; Summers,
W.C.; Sanderson, M.R.

Deposited on : 1998-05-15

Resolution : 2.37 Å(reported)

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

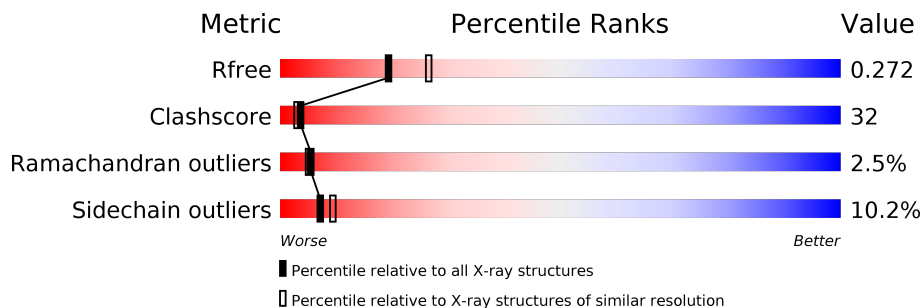|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.37 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	Percentile Ranks	Value
Rfree		0.272
Clashscore		32
Ramachandran outliers		2.5%
Sidechain outliers		10.2%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	331	48% 40% 5% 7%
1	B	331	45% 41% 7% · 6%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4869 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 THYMIDINE KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	309	Total	C	N	O	S	0	0	3
			2339	1490	411	423	15			
1	B	311	Total	C	N	O	S	0	0	3
			2352	1500	413	424	15			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is 9-(4-HYDROXY-3-(HYDROXYMETHYL)BUT-1-YL)GUANINE (three-letter code: PE2) (formula: $C_{10}H_{15}N_5O_3$).
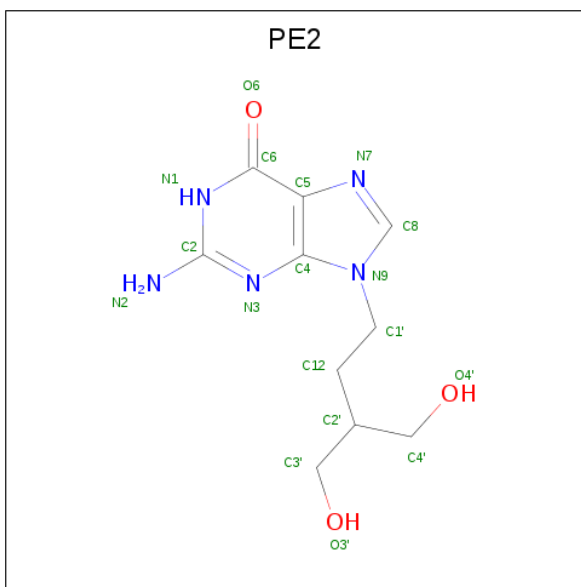
PE2

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			18	10	5	3		
3	B	1	Total	C	N	O	0	0
			18	10	5	3		

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	69	Total	O	0	0
			69	69		
4	B	63	Total	O	0	0
			63	63		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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: THYMIDINE KINASE
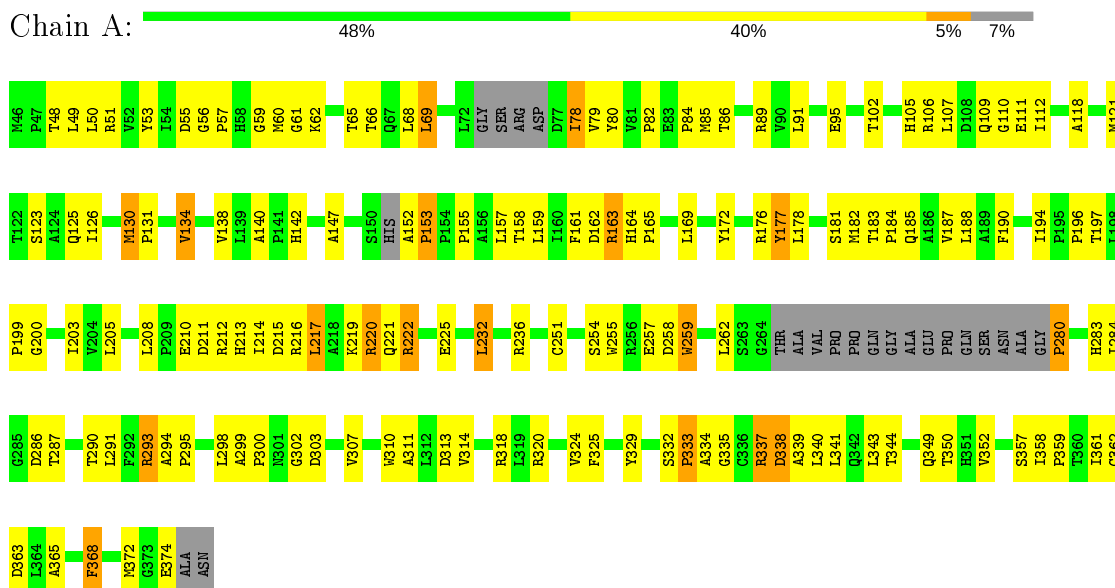
- Molecule 1: THYMIDINE KINASE
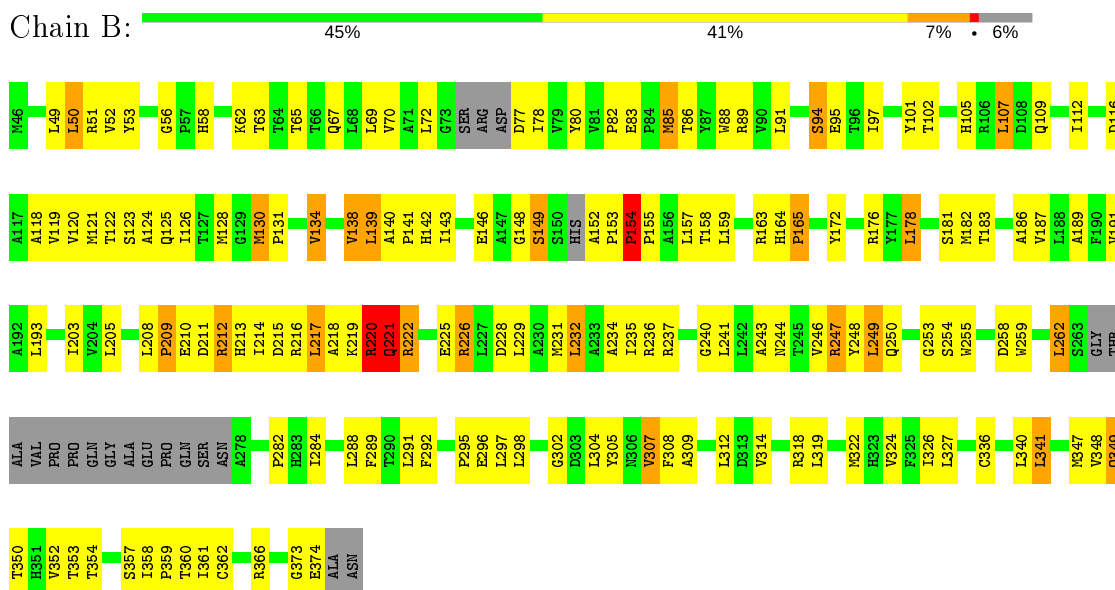
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	113.80Å 117.70Å 108.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 2.37 11.99 – 2.36	Depositor EDS
% Data completeness (in resolution range)	70.4 (12.00-2.37) 69.7 (11.99-2.36)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$<I/\sigma(I)>$ [1]	1.63 (at 2.35Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.200 , 0.314 0.235 , 0.272	Depositor DCC
R_{free} test set	2096 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	25.7	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 37.6	EDS
L-test for twinning[2]	$<|L|> = 0.44, <L^2> = 0.27$	Xtriage
Estimated twinning fraction	0.057 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	4869	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% 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: SO4, PE2

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	0.67	0/2392	0.85	3/3267 (0.1%)
1	B	0.70	0/2406	0.86	2/3287 (0.1%)
All	All	0.69	0/4798	0.85	5/6554 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	280	PRO	N-CA-CB	6.45	111.03	103.30
1	A	69	LEU	CA-CB-CG	6.31	129.81	115.30
1	B	304	LEU	CA-CB-CG	5.64	128.27	115.30
1	A	291	LEU	CA-CB-CG	5.48	127.91	115.30
1	B	139	LEU	CA-CB-CG	5.45	127.83	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2339	0	2349	149	0
1	B	2352	0	2368	152	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	18	0	15	0	0
3	B	18	0	15	5	0
4	A	69	0	0	5	0
4	B	63	0	0	4	0
All	All	4869	0	4747	299	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 32.

The worst 5 of 299 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:254:SER:HB3	1:A:257:GLU:HG2	1.30	1.10
1:A:84:PRO:HG3	1:A:368:PHE:HZ	1.19	1.02
1:B:62:LYS:HE2	1:B:163:ARG:HH12	1.37	0.89
1:A:232:LEU:HD13	1:A:236:ARG:HD2	1.54	0.88
1:A:84:PRO:HG3	1:A:368:PHE:CZ	2.09	0.85

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	301/331 (91%)	260 (86%)	31 (10%)	10 (3%)	4	3
1	B	303/331 (92%)	267 (88%)	31 (10%)	5 (2%)	9	10
All	All	604/662 (91%)	527 (87%)	62 (10%)	15 (2%)	5	5

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	ARG
1	B	149	SER
1	B	220	ARG
1	A	110	GLY
1	A	220	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	245/264 (93%)	224 (91%)	21 (9%)	10	14
1	B	246/264 (93%)	217 (88%)	29 (12%)	5	6
All	All	491/528 (93%)	441 (90%)	50 (10%)	7	9

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	72	LEU
1	B	130	MET
1	B	349	GLN
1	B	85	MET
1	B	109	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	67	GLN
1	A	99	ASN
1	A	142	HIS
1	A	351	HIS
1	B	142	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

4 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	SO4	A	3	-	4,4,4	0.61	0	6,6,6	1.20	1 (16%)
3	PE2	B	2	-	15,19,19	1.43	4 (26%)	16,26,26	4.26	5 (31%)
3	PE2	A	1	-	15,19,19	2.03	4 (26%)	16,26,26	4.19	7 (43%)
2	SO4	B	4	-	4,4,4	0.71	0	6,6,6	1.35	1 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PE2	B	2	-	-	1/9/9/9	0/2/2/2
3	PE2	A	1	-	-	6/9/9/9	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1	PE2	C6-N1	4.45	1.40	1.33
3	A	1	PE2	C2-N1	3.89	1.42	1.35
3	A	1	PE2	C8-N7	-3.42	1.28	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2	PE2	C4-N3	-2.91	1.31	1.35
3	B	2	PE2	C8-N7	-2.56	1.30	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	2	PE2	C5-C6-N1	-12.89	105.80	123.43
3	A	1	PE2	C12-C1'-N9	9.28	130.78	112.13
3	A	1	PE2	C5-C6-N1	-9.18	110.88	123.43
3	B	2	PE2	C6-N1-C2	7.26	127.47	115.93
3	B	2	PE2	C2-N3-C4	-6.21	108.27	115.36

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1	PE2	C12-C2'-C3'-O3'
3	A	1	PE2	C12-C1'-N9-C4
3	A	1	PE2	C12-C1'-N9-C8
3	B	2	PE2	N9-C1'-C12-C2'
3	A	1	PE2	C1'-C12-C2'-C3'

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	2	PE2	5	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 PE2 B 2

Bond lengths

Bond angles

Torsions

Rings

Ligand PE2 A 1

Bond lengths

Bond angles

Torsions

Rings
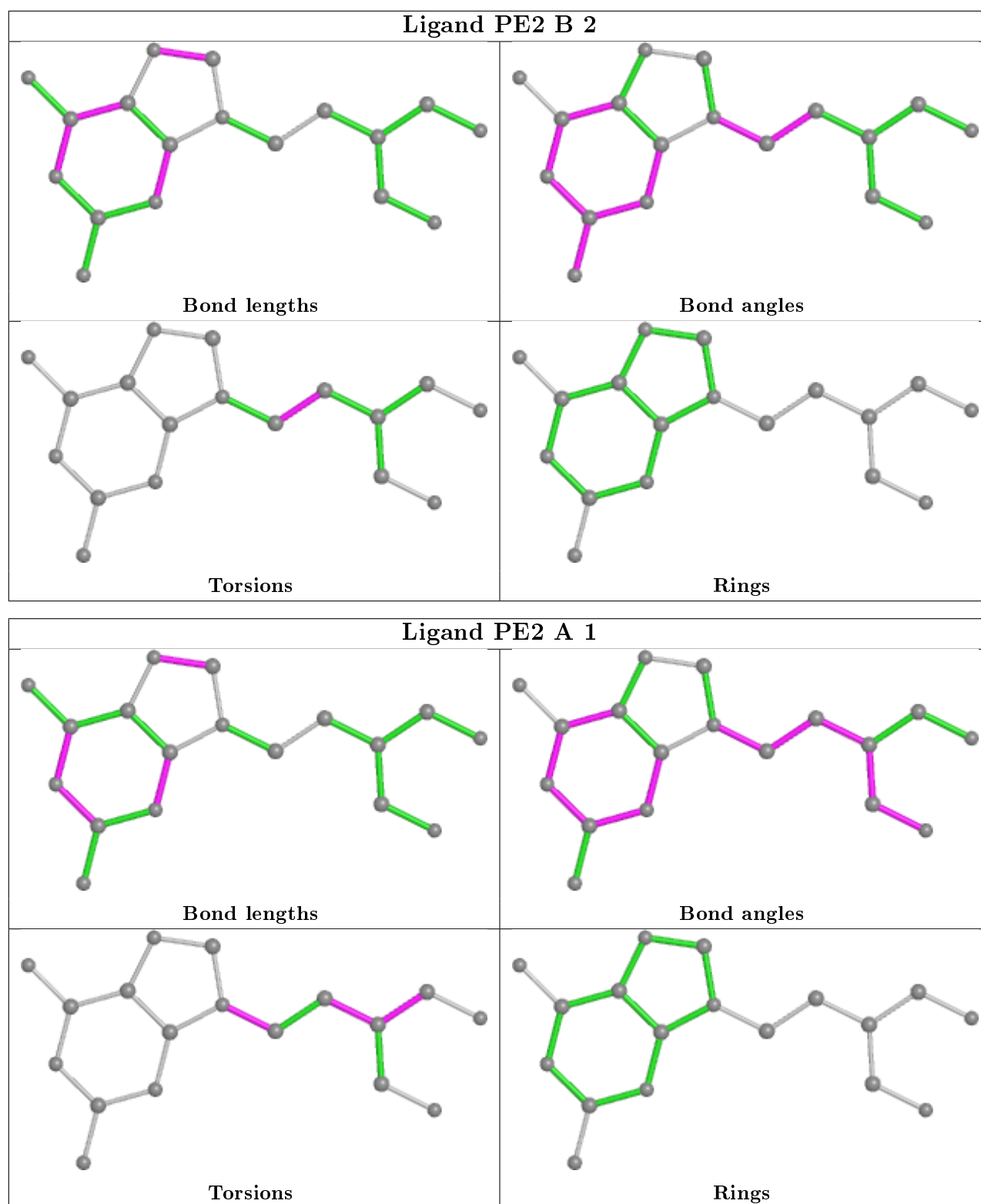
5.7   Other polymers (i)

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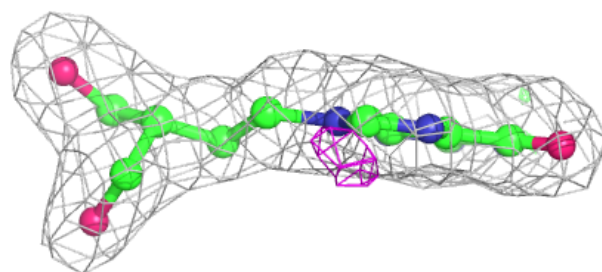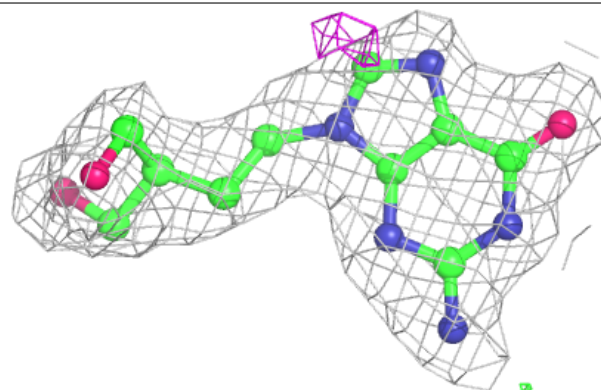
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around PE2 B 2:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
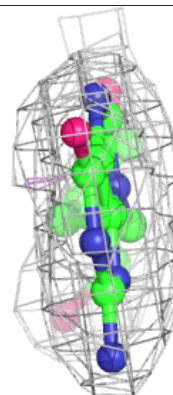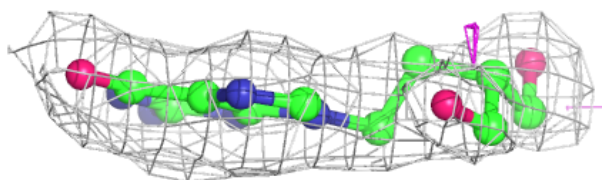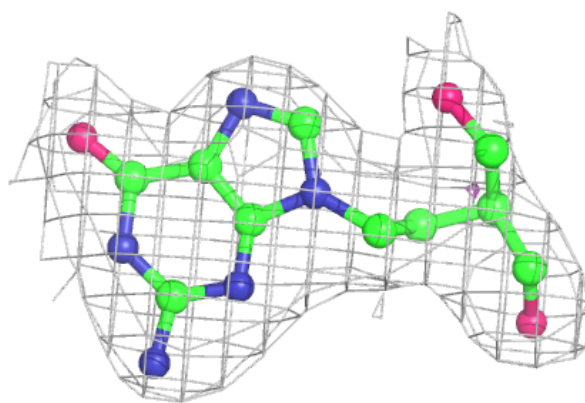
Electron density around PE2 A 1:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (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.