



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 03:14 am BST

PDB ID : 5XIQ  
Title : Crystal Structure of Toxoplasma gondii Prolyl-tRNA Synthetase (TgPRS) in complex with Halofuginone  
Authors : Jain, V.; Manickam, Y.; Sharma, A.  
Deposited on : 2017-04-26  
Resolution : 2.19 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

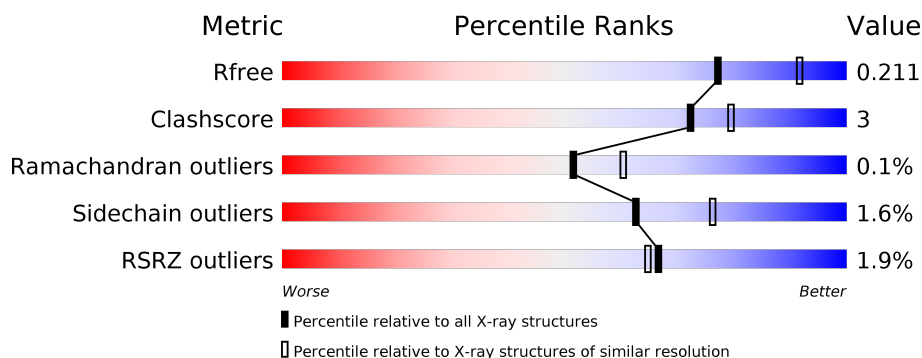
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.19 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	500	
1	B	500	
1	C	500	
1	D	500	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16320 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.

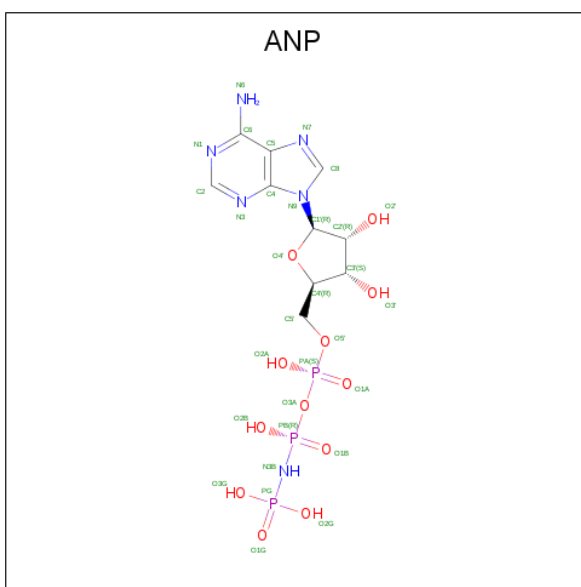
- Molecule 1 is a protein called Prolyl-tRNA synthetase (ProRS).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	478	Total	C	N	O	S	0	5	0
			3917	2521	671	703	22			
1	B	480	Total	C	N	O	S	0	1	0
			3905	2509	667	707	22			
1	C	481	Total	C	N	O	S	0	6	0
			3937	2533	681	701	22			
1	D	479	Total	C	N	O	S	0	6	0
			3941	2532	680	707	22			

There are 12 discrepancies between the modelled and reference sequences:

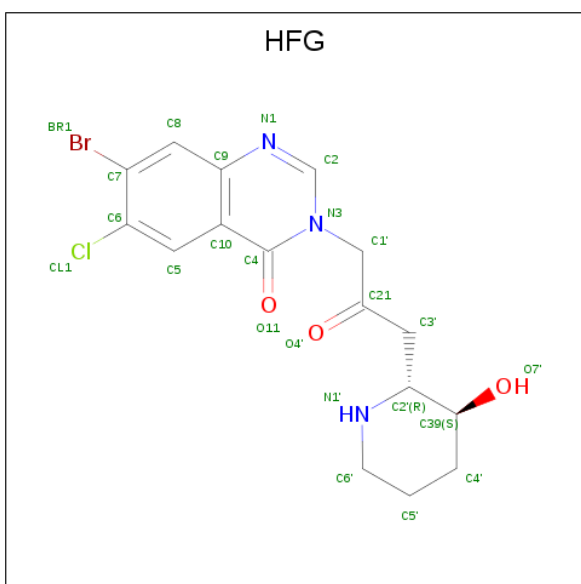
Chain	Residue	Modelled	Actual	Comment	Reference
A	331	GLY	-	expression tag	UNP S8G8I1
A	332	ALA	-	expression tag	UNP S8G8I1
A	333	MET	-	expression tag	UNP S8G8I1
B	331	GLY	-	expression tag	UNP S8G8I1
B	332	ALA	-	expression tag	UNP S8G8I1
B	333	MET	-	expression tag	UNP S8G8I1
C	331	GLY	-	expression tag	UNP S8G8I1
C	332	ALA	-	expression tag	UNP S8G8I1
C	333	MET	-	expression tag	UNP S8G8I1
D	331	GLY	-	expression tag	UNP S8G8I1
D	332	ALA	-	expression tag	UNP S8G8I1
D	333	MET	-	expression tag	UNP S8G8I1

- Molecule 2 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 3 is 7-bromo-6-chloro-3-{3-[(2R,3S)-3-hydroxypiperidin-2-yl]-2-oxopropyl}quinazolin-4(3H)-one (three-letter code: HFG) (formula:  $C_{16}H_{17}BrClN_3O_3$ ).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	Br	C	Cl	N	O	0	0
			24	1	16	1	3	3		
3	B	1	Total	Br	C	Cl	N	O	0	0
			24	1	16	1	3	3		
3	C	1	Total	Br	C	Cl	N	O	0	0
			24	1	16	1	3	3		
3	D	1	Total	Br	C	Cl	N	O	0	0
			24	1	16	1	3	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Mg	0	0
			2	2		
4	A	2	Total	Mg	0	0
			2	2		
4	D	2	Total	Mg	0	0
			2	2		
4	C	2	Total	Mg	0	0
			2	2		

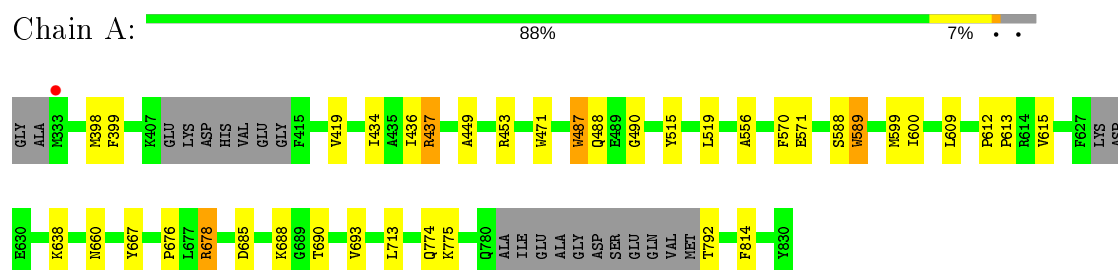
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	130	Total	O	0	2
			132	132		
5	B	73	Total	O	0	0
			73	73		
5	C	109	Total	O	0	2
			111	111		
5	D	76	Total	O	0	0
			76	76		

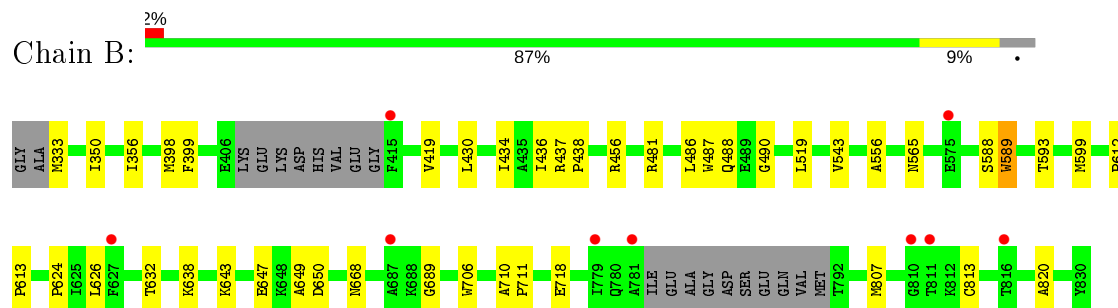
### 3 Residue-property plots [i](#)

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 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.

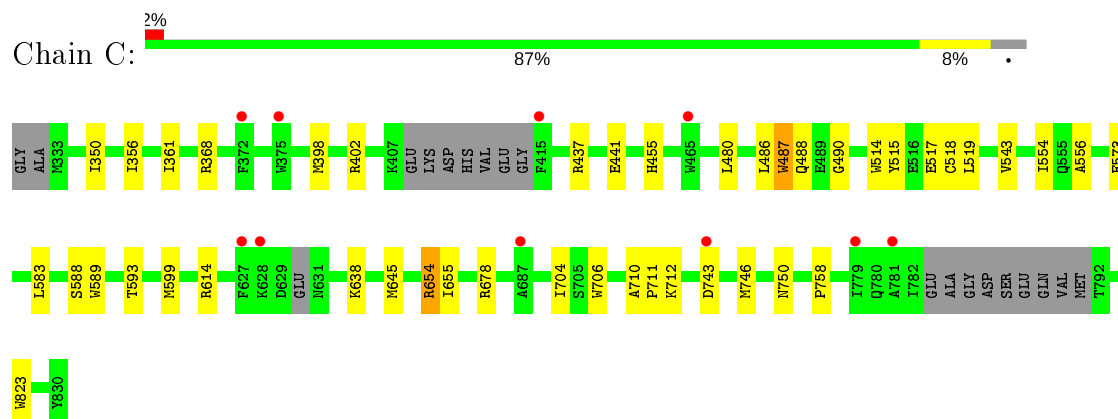
- Molecule 1: Prolyl-tRNA synthetase (ProRS)



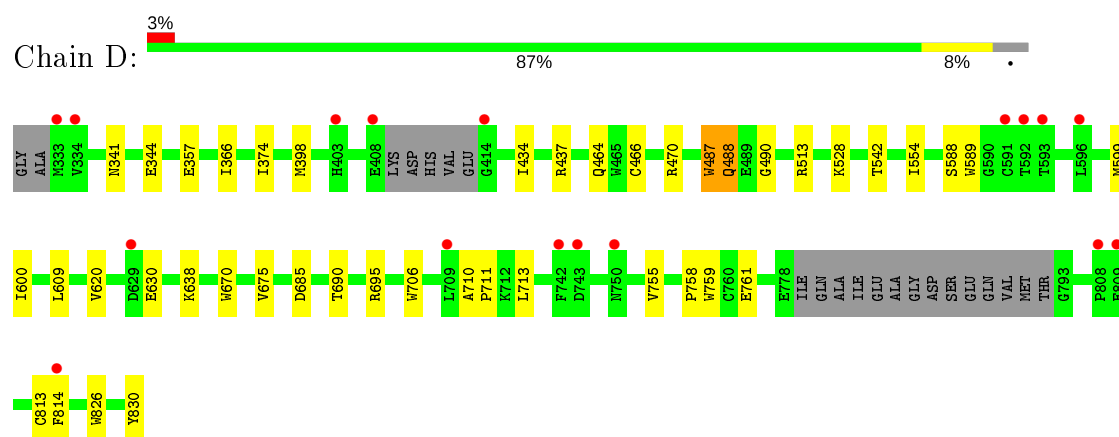
- Molecule 1: Prolyl-tRNA synthetase (ProRS)



- Molecule 1: Prolyl-tRNA synthetase (ProRS)



- Molecule 1: Prolyl-tRNA synthetase (ProRS)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.85Å 90.90Å 92.87Å 89.86° 80.92° 75.82°	Depositor
Resolution (Å)	40.32 – 2.19 40.29 – 2.19	Depositor EDS
% Data completeness (in resolution range)	97.4 (40.32-2.19) 97.5 (40.29-2.19)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.165 , 0.207 0.173 , 0.211	Depositor DCC
$R_{free}$ test set	6057 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.3	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16320	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.71% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ANP, HFG

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	0.95	0/4037	0.91	1/5461 (0.0%)
1	B	0.86	0/4014	0.89	1/5433 (0.0%)
1	C	0.88	0/4061	0.92	2/5495 (0.0%)
1	D	0.88	0/4067	0.88	3/5499 (0.1%)
All	All	0.89	0/16179	0.90	7/21888 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	513	ARG	NE-CZ-NH1	6.17	123.39	120.30
1	D	513	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	D	470	ARG	NE-CZ-NH2	-6.00	117.30	120.30
1	C	437	ARG	NE-CZ-NH2	-5.49	117.56	120.30
1	A	678	ARG	NE-CZ-NH1	5.41	123.01	120.30
1	B	481	ARG	NE-CZ-NH1	5.25	122.93	120.30
1	C	368	ARG	NE-CZ-NH2	-5.18	117.71	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	3917	0	3872	26	0
1	B	3905	0	3838	24	0
1	C	3937	0	3895	30	0
1	D	3941	0	3890	33	0
2	A	31	0	13	0	0
2	B	31	0	13	0	0
2	C	31	0	13	1	0
2	D	31	0	13	0	0
3	A	24	0	17	2	0
3	B	24	0	17	1	0
3	C	24	0	17	2	0
3	D	24	0	17	1	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
5	A	132	0	0	1	0
5	B	73	0	0	0	0
5	C	111	0	0	3	0
5	D	76	0	0	0	0
All	All	16320	0	15615	106	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 (106) 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:685:ASP:O	1:A:688:LYS:O	1.93	0.86
1:C:398:MET:HE1	1:D:398:MET:HE1	1.57	0.85
1:B:807:MET:CE	1:B:820:ALA:HB3	2.17	0.75
1:D:437:ARG:NH2	1:D:466:CYS:HB2	2.02	0.74
1:B:543:VAL:HG23	1:B:556:ALA:HB3	1.70	0.73
1:D:398:MET:HA	1:D:437:ARG:HD3	1.72	0.72
1:C:398:MET:HE1	1:D:398:MET:CE	2.20	0.70
1:C:398:MET:CE	1:D:398:MET:HE1	2.22	0.69
1:A:519:LEU:HD22	1:A:599:MET:CE	2.24	0.68
2:C:1001:ANP:O2G	5:C:1101:HOH:O	2.13	0.66
1:C:398:MET:CE	1:D:398:MET:CE	2.74	0.65
1:A:519:LEU:HD22	1:A:599:MET:HE3	1.77	0.65
1:C:487:TRP:CH2	3:C:1002:HFG:H8	2.32	0.64
1:B:710:ALA:HB3	1:B:711:PRO:HD3	1.81	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:437:ARG:HG2	1:A:437:ARG:O	2.01	0.61
1:B:807:MET:HE1	1:B:820:ALA:HB3	1.83	0.59
1:D:600:ILE:HG23	1:D:609:LEU:HD22	1.84	0.59
1:B:626:LEU:HG	1:B:632:THR:HG22	1.86	0.58
1:C:543:VAL:HG23	1:C:556:ALA:HB3	1.86	0.57
1:C:710:ALA:HB3	1:C:711:PRO:HD3	1.89	0.55
1:D:758:PRO:HD2	1:D:814:PHE:HB3	1.88	0.55
1:A:688:LYS:O	1:A:690:THR:N	2.36	0.55
1:C:638:LYS:HE3	1:C:706:TRP:CZ3	2.42	0.54
1:D:528:LYS:HD3	1:D:542:THR:HG21	1.89	0.54
1:A:398:MET:CE	1:A:436:ILE:HG23	2.38	0.54
1:C:361:ILE:HD13	1:D:434:ILE:HD12	1.89	0.53
1:A:399:PHE:HB3	1:A:434[A]:ILE:HG21	1.92	0.52
1:C:480:LEU:HD22	1:C:554:ILE:CD1	2.39	0.52
1:D:341:ASN:ND2	1:D:344:GLU:HG3	2.26	0.51
1:A:399:PHE:HB3	1:A:434[A]:ILE:CG2	2.41	0.51
1:D:710:ALA:HB3	1:D:711:PRO:CD	2.42	0.50
1:B:638:LYS:HE3	1:B:706:TRP:CE3	2.46	0.50
1:D:487:TRP:CH2	3:D:1002:HFG:H8	2.47	0.50
1:A:419:VAL:HG22	3:A:1002:HFG:BR1	2.67	0.50
1:D:670:TRP:CE3	1:D:675:VAL:HG21	2.46	0.49
1:C:455:HIS:ND1	5:C:1103:HOH:O	2.34	0.49
1:A:398:MET:HE2	1:A:436:ILE:HG23	1.94	0.49
1:D:600:ILE:HG23	1:D:609:LEU:CD2	2.43	0.49
1:A:599:MET:HE1	1:A:615:VAL:HG21	1.95	0.48
1:C:398:MET:HE3	1:D:398:MET:CE	2.44	0.48
1:B:430:LEU:HD13	1:B:434:ILE:HG12	1.96	0.48
1:C:517:GLU:O	1:C:614[A]:ARG:HD2	2.13	0.47
1:C:519:LEU:HD22	1:C:599:MET:CE	2.44	0.47
1:D:755:VAL:HG22	1:D:826:TRP:HB2	1.97	0.47
1:C:514:TRP:O	1:C:518:CYS:HB2	2.15	0.47
1:D:437:ARG:HH21	1:D:466:CYS:HB2	1.79	0.47
1:B:399:PHE:HB3	1:B:434:ILE:CG2	2.45	0.47
1:D:490:GLY:O	1:D:588:SER:HA	2.14	0.46
1:A:487:TRP:CH2	3:A:1002:HFG:H8	2.50	0.46
1:D:374:ILE:HG21	1:D:599:MET:HE3	1.97	0.46
1:C:758:PRO:HD3	1:C:823:TRP:CZ3	2.51	0.46
1:C:361:ILE:CD1	1:D:434:ILE:HD12	2.46	0.46
1:A:437:ARG:HD3	5:A:1169:HOH:O	2.16	0.46
1:C:490:GLY:O	1:C:588:SER:HA	2.15	0.46
1:C:350:ILE:HD12	1:C:356:ILE:HG12	1.98	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:486:LEU:HB2	1:C:593:THR:CG2	2.46	0.45
1:B:624:PRO:HB2	1:B:626:LEU:HD13	1.99	0.45
1:C:486:LEU:HB2	1:C:593:THR:HG23	1.98	0.45
1:D:638:LYS:HE2	1:D:706:TRP:CZ3	2.51	0.45
1:B:519:LEU:HD22	1:B:599:MET:CE	2.46	0.45
1:C:519:LEU:HD22	1:C:599:MET:HE2	1.98	0.45
1:B:490:GLY:O	1:B:588:SER:HA	2.16	0.45
1:D:357:GLU:HB3	1:D:366:ILE:HB	1.99	0.45
1:A:660:ASN:ND2	1:B:456:ARG:CZ	2.80	0.45
1:B:519:LEU:HD22	1:B:599:MET:HE2	1.99	0.45
1:A:774:GLN:OE1	1:A:792:THR:OG1	2.35	0.44
1:B:612:PRO:HA	1:B:613:PRO:HD3	1.91	0.44
1:D:630:GLU:HG3	1:D:630:GLU:O	2.17	0.44
1:B:689:GLY:HA2	1:B:706:TRP:CH2	2.53	0.44
1:D:620:VAL:HG21	1:D:713:LEU:HD13	1.99	0.43
1:B:398:MET:HE2	1:B:436:ILE:HG23	2.00	0.43
1:C:704:ILE:HD13	1:C:712:LYS:HG2	2.00	0.43
1:C:515:TYR:CZ	1:C:556:ALA:HB1	2.53	0.43
1:D:685:ASP:HB3	1:D:690:THR:O	2.18	0.43
1:C:441:GLU:OE2	3:C:1002:HFG:N1'	2.52	0.43
1:B:649:ALA:O	1:B:650:ASP:HB2	2.18	0.43
1:D:759:TRP:CE3	1:D:761:GLU:HA	2.54	0.43
1:C:746:MET:O	1:C:750:ASN:ND2	2.51	0.43
1:B:398:MET:CE	1:B:436:ILE:HG23	2.49	0.43
1:D:710:ALA:HB3	1:D:711:PRO:HD3	2.00	0.42
1:B:350:ILE:HD12	1:B:356:ILE:HG12	2.00	0.42
1:C:402[A]:ARG:HD3	5:C:1204:HOH:O	2.19	0.42
1:A:490:GLY:O	1:A:588:SER:HA	2.19	0.42
1:B:333:MET:HG2	1:B:668:ASN:OD1	2.19	0.42
1:B:490:GLY:HA3	1:B:589:TRP:CZ3	2.55	0.42
1:B:643:LYS:O	1:B:647:GLU:HG2	2.20	0.42
1:A:570:PHE:O	1:A:571:GLU:HB2	2.20	0.42
1:D:374:ILE:HD13	1:D:599:MET:HE1	2.02	0.42
1:C:654[B]:ARG:HG2	1:C:655:ILE:N	2.35	0.42
1:A:515:TYR:CZ	1:A:556:ALA:HB1	2.55	0.41
1:D:695:ARG:HH11	1:D:695:ARG:HD2	1.72	0.41
1:A:685:ASP:HB3	1:A:690:THR:O	2.20	0.41
1:A:588:SER:C	1:A:589:TRP:CD1	2.94	0.41
1:C:573:GLU:HG2	1:C:583:LEU:HD23	2.02	0.41
1:A:519:LEU:HD22	1:A:599:MET:HE2	2.01	0.41
1:B:419:VAL:HG22	3:B:1002:HFG:BR1	2.75	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:612:PRO:HA	1:A:613:PRO:HD3	1.97	0.41
1:D:554:ILE:HG13	1:D:830:TYR:HB2	2.03	0.41
1:A:667:TYR:CD1	1:A:678:ARG:HD2	2.55	0.41
1:C:519:LEU:CD2	1:C:599:MET:HE2	2.51	0.41
1:B:486:LEU:HB2	1:B:593:THR:HG23	2.03	0.41
1:A:600:ILE:HG23	1:A:609:LEU:CD2	2.51	0.41
1:D:464:GLN:O	1:D:488:GLN:HA	2.21	0.40
1:A:449:ALA:O	1:A:453:ARG:NH1	2.54	0.40
1:A:693:VAL:HG21	1:A:713:LEU:HD21	2.04	0.40
1:D:638:LYS:HE2	1:D:706:TRP:CH2	2.56	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	475/500 (95%)	464 (98%)	11 (2%)	0	100	100
1	B	475/500 (95%)	460 (97%)	14 (3%)	1 (0%)	47	55
1	C	479/500 (96%)	463 (97%)	16 (3%)	0	100	100
1	D	479/500 (96%)	468 (98%)	11 (2%)	0	100	100
All	All	1908/2000 (95%)	1855 (97%)	52 (3%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	438	PRO

### 5.3.2 Protein sidechains ⓘ

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	416/436 (95%)	407 (98%)	9 (2%)	52	65
1	B	414/436 (95%)	407 (98%)	7 (2%)	60	74
1	C	417/436 (96%)	409 (98%)	8 (2%)	57	71
1	D	419/436 (96%)	415 (99%)	4 (1%)	76	86
All	All	1666/1744 (96%)	1638 (98%)	28 (2%)	62	74

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	437	ARG
1	A	471	TRP
1	A	487	TRP
1	A	488	GLN
1	A	589	TRP
1	A	638	LYS
1	A	676	PRO
1	A	775	LYS
1	A	814	PHE
1	B	437	ARG
1	B	487	TRP
1	B	488	GLN
1	B	565	ASN
1	B	589	TRP
1	B	718	GLU
1	B	813	CYS
1	C	487	TRP
1	C	488	GLN
1	C	589	TRP
1	C	645	MET
1	C	654[A]	ARG
1	C	654[B]	ARG
1	C	678	ARG
1	C	743	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	487	TRP
1	D	488	GLN
1	D	589	TRP
1	D	813	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	750	ASN
1	C	780	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 8 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
3	HFG	C	1002	-	26,26,26	1.05	1 (3%)	22,37,37	0.79	1 (4%)
3	HFG	D	1002	-	26,26,26	1.15	1 (3%)	22,37,37	1.21	2 (9%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HFG	A	1002	-	26,26,26	1.28	2 (7%)	22,37,37	0.93	2 (9%)
3	HFG	B	1002	-	26,26,26	1.24	2 (7%)	22,37,37	0.96	2 (9%)
2	ANP	C	1001	4	29,33,33	0.99	1 (3%)	31,52,52	0.98	3 (9%)
2	ANP	B	1001	4	29,33,33	1.64	4 (13%)	31,52,52	1.00	2 (6%)
2	ANP	A	1001	4	29,33,33	0.99	1 (3%)	31,52,52	1.05	2 (6%)
2	ANP	D	1001	4	29,33,33	1.04	2 (6%)	31,52,52	0.99	3 (9%)

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	HFG	C	1002	-	-	0/8/19/19	0/3/3/3
3	HFG	D	1002	-	-	0/8/19/19	0/3/3/3
3	HFG	A	1002	-	-	0/8/19/19	0/3/3/3
3	HFG	B	1002	-	-	0/8/19/19	0/3/3/3
2	ANP	C	1001	4	-	4/14/38/38	0/3/3/3
2	ANP	B	1001	4	-	3/14/38/38	0/3/3/3
2	ANP	A	1001	4	-	4/14/38/38	0/3/3/3
2	ANP	D	1001	4	-	6/14/38/38	0/3/3/3

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1001	ANP	PB-O1B	4.98	1.54	1.46
2	B	1001	ANP	PB-O2B	-4.77	1.43	1.56
2	D	1001	ANP	O4'-C1'	3.60	1.46	1.41
2	C	1001	ANP	O4'-C1'	3.50	1.46	1.41
2	B	1001	ANP	O4'-C1'	3.49	1.46	1.41
2	A	1001	ANP	O4'-C1'	3.23	1.45	1.41
3	C	1002	HFG	C4-C10	3.14	1.47	1.41
3	A	1002	HFG	C4-C10	2.55	1.46	1.41
3	B	1002	HFG	C4-C10	2.54	1.46	1.41
3	D	1002	HFG	C4-C10	2.47	1.46	1.41
3	A	1002	HFG	C1'-C21	2.32	1.54	1.51
3	B	1002	HFG	C8-C9	-2.21	1.38	1.41
2	B	1001	ANP	PB-O3A	2.07	1.61	1.59
2	D	1001	ANP	PB-O3A	2.00	1.61	1.59



All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	1002	HFG	C4'-C39-C2'	-3.29	108.74	110.89
2	A	1001	ANP	O1G-PG-N3B	-2.86	107.56	111.77
3	B	1002	HFG	C4'-C39-C2'	-2.68	109.14	110.89
2	C	1001	ANP	O1G-PG-N3B	-2.59	107.96	111.77
2	B	1001	ANP	O1G-PG-N3B	-2.58	107.97	111.77
2	D	1001	ANP	O1G-PG-N3B	-2.57	107.98	111.77
3	A	1002	HFG	C4'-C39-C2'	-2.36	109.34	110.89
2	A	1001	ANP	O1B-PB-N3B	2.29	115.14	111.77
3	B	1002	HFG	C10-C4-N3	-2.25	114.61	116.15
2	B	1001	ANP	O1B-PB-N3B	2.25	115.09	111.77
3	C	1002	HFG	C10-C4-N3	-2.25	114.61	116.15
3	D	1002	HFG	C10-C4-N3	-2.24	114.62	116.15
2	D	1001	ANP	C5'-C4'-C3'	-2.18	107.00	115.18
3	A	1002	HFG	C10-C4-N3	-2.14	114.69	116.15
2	C	1001	ANP	C5'-C4'-C3'	-2.12	107.23	115.18
2	C	1001	ANP	O1B-PB-N3B	2.11	114.88	111.77
2	D	1001	ANP	O1B-PB-N3B	2.10	114.86	111.77

There are no chirality outliers.

All (17) torsion outliers are listed below:

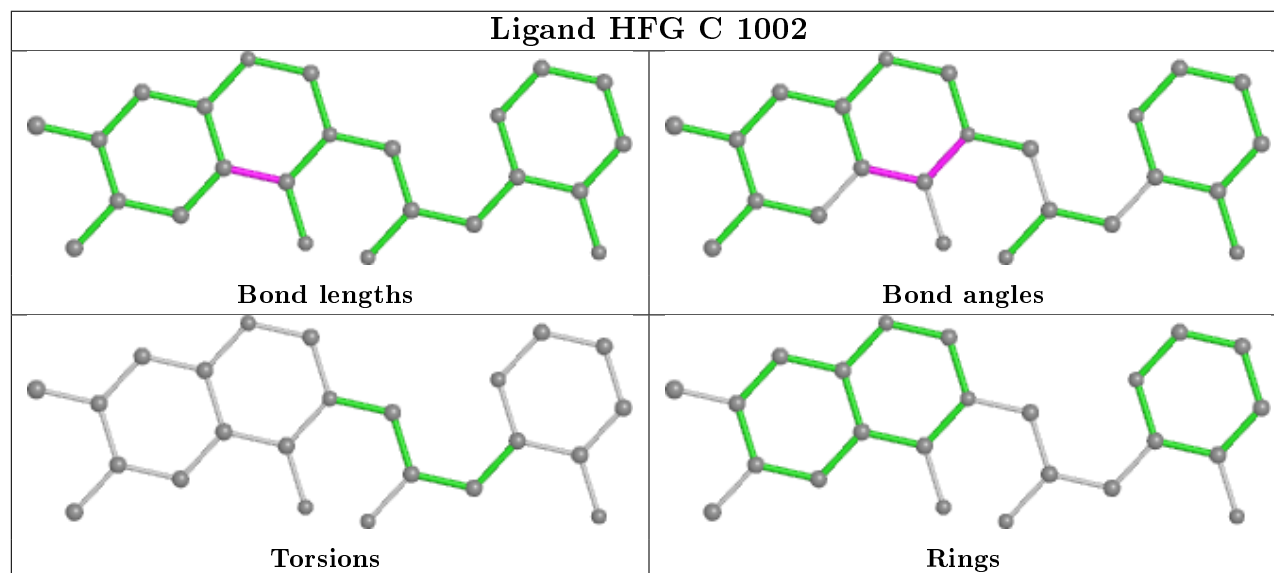
Mol	Chain	Res	Type	Atoms
2	C	1001	ANP	PG-N3B-PB-O1B
2	C	1001	ANP	C5'-O5'-PA-O1A
2	B	1001	ANP	C5'-O5'-PA-O1A
2	A	1001	ANP	PG-N3B-PB-O1B
2	A	1001	ANP	C5'-O5'-PA-O1A
2	D	1001	ANP	PG-N3B-PB-O1B
2	D	1001	ANP	PA-O3A-PB-O1B
2	D	1001	ANP	PA-O3A-PB-O2B
2	D	1001	ANP	C5'-O5'-PA-O1A
2	D	1001	ANP	C5'-O5'-PA-O2A
2	C	1001	ANP	C5'-O5'-PA-O3A
2	B	1001	ANP	C5'-O5'-PA-O3A
2	A	1001	ANP	C5'-O5'-PA-O3A
2	D	1001	ANP	C5'-O5'-PA-O3A
2	C	1001	ANP	C5'-O5'-PA-O2A
2	B	1001	ANP	C5'-O5'-PA-O2A
2	A	1001	ANP	C5'-O5'-PA-O2A

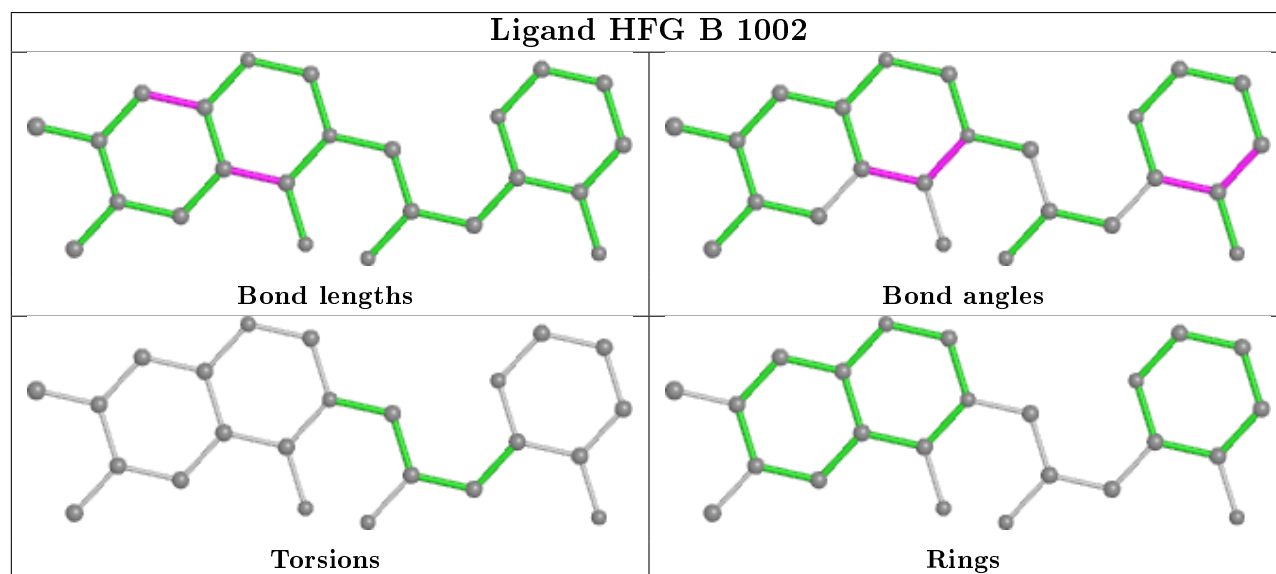
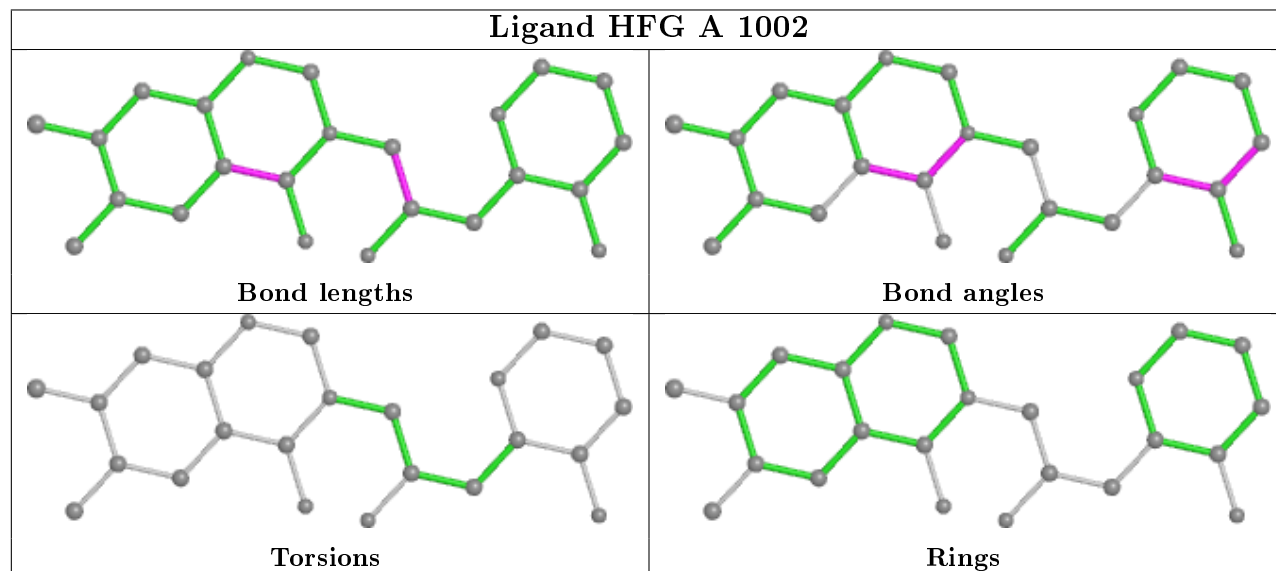
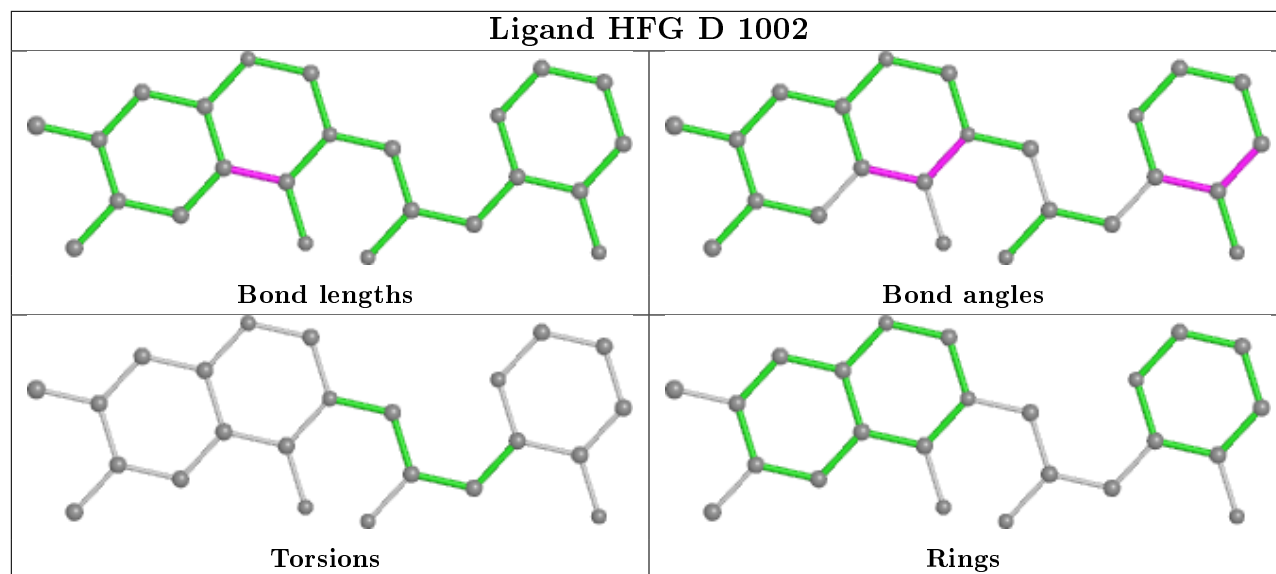
There are no ring outliers.

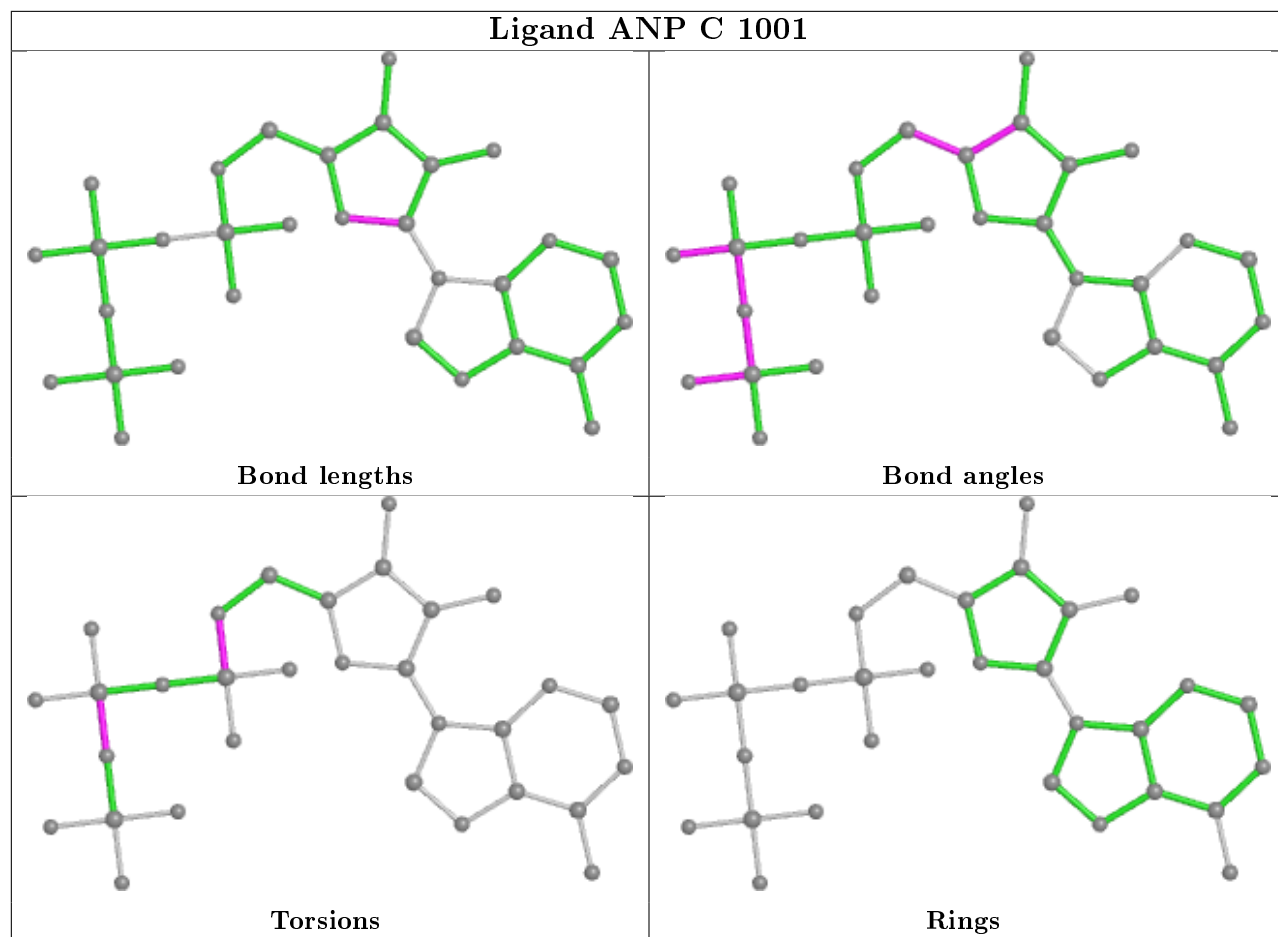
5 monomers are involved in 7 short contacts:

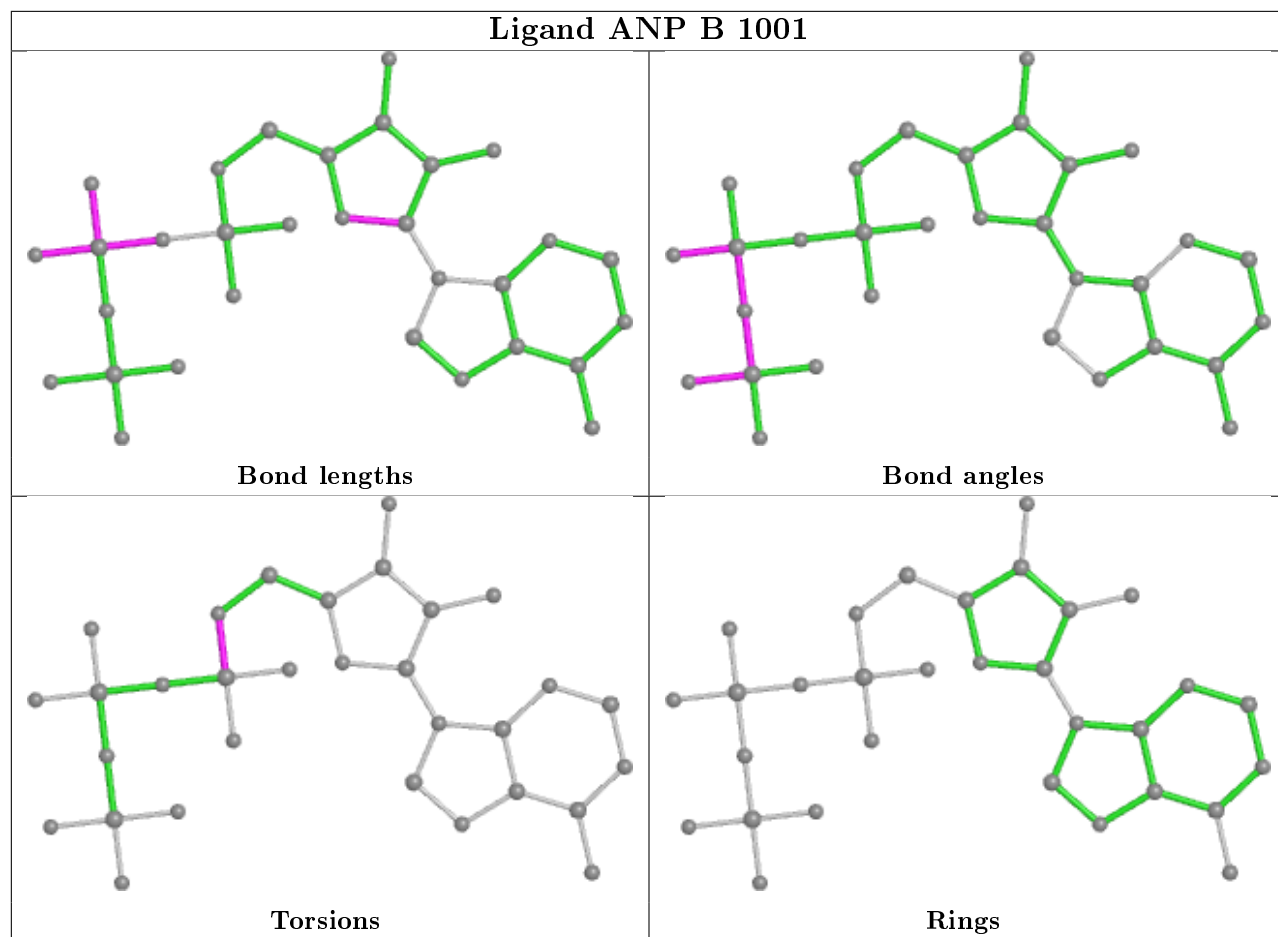
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1002	HFG	2	0
3	D	1002	HFG	1	0
3	A	1002	HFG	2	0
3	B	1002	HFG	1	0
2	C	1001	ANP	1	0

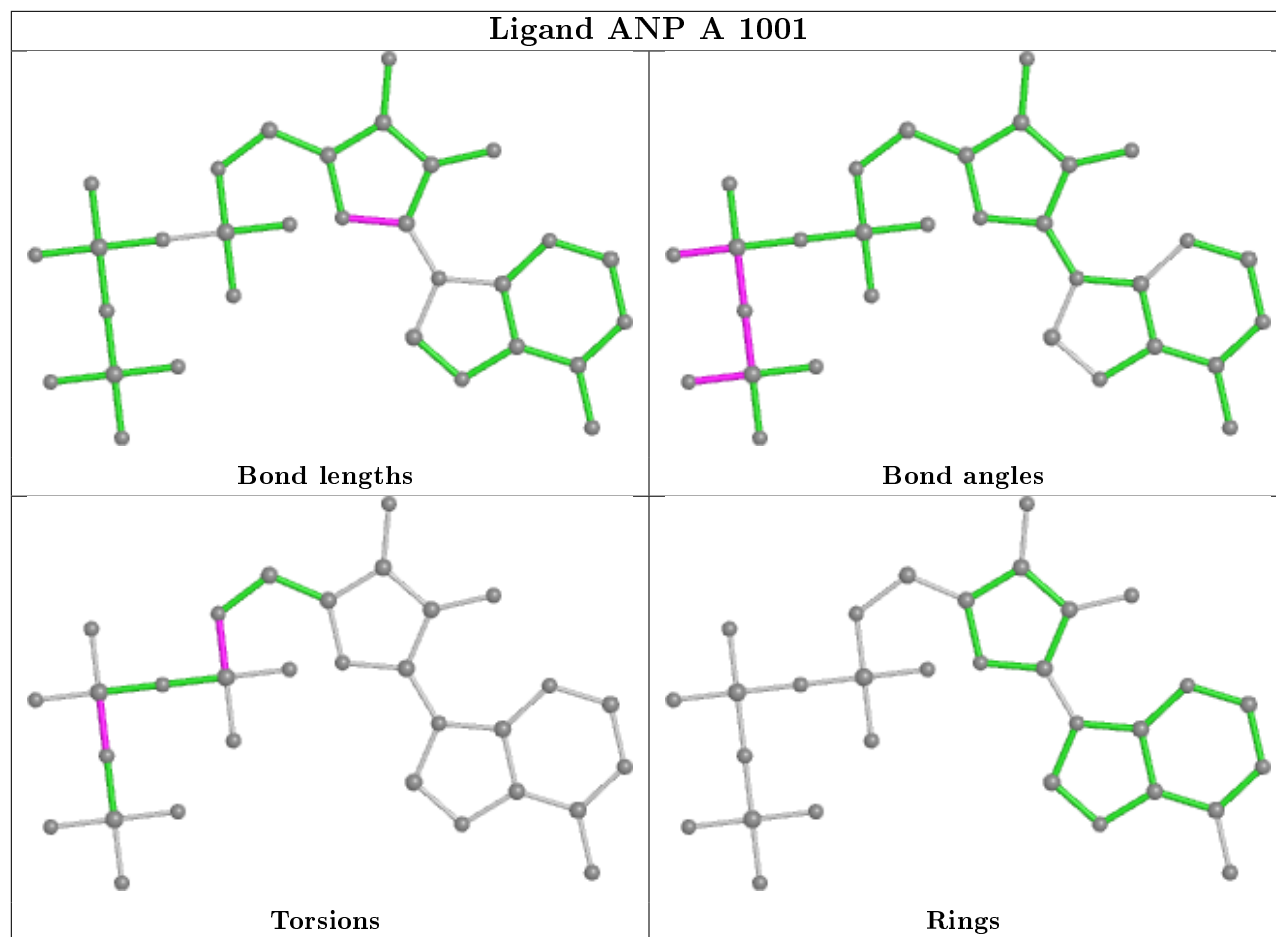
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

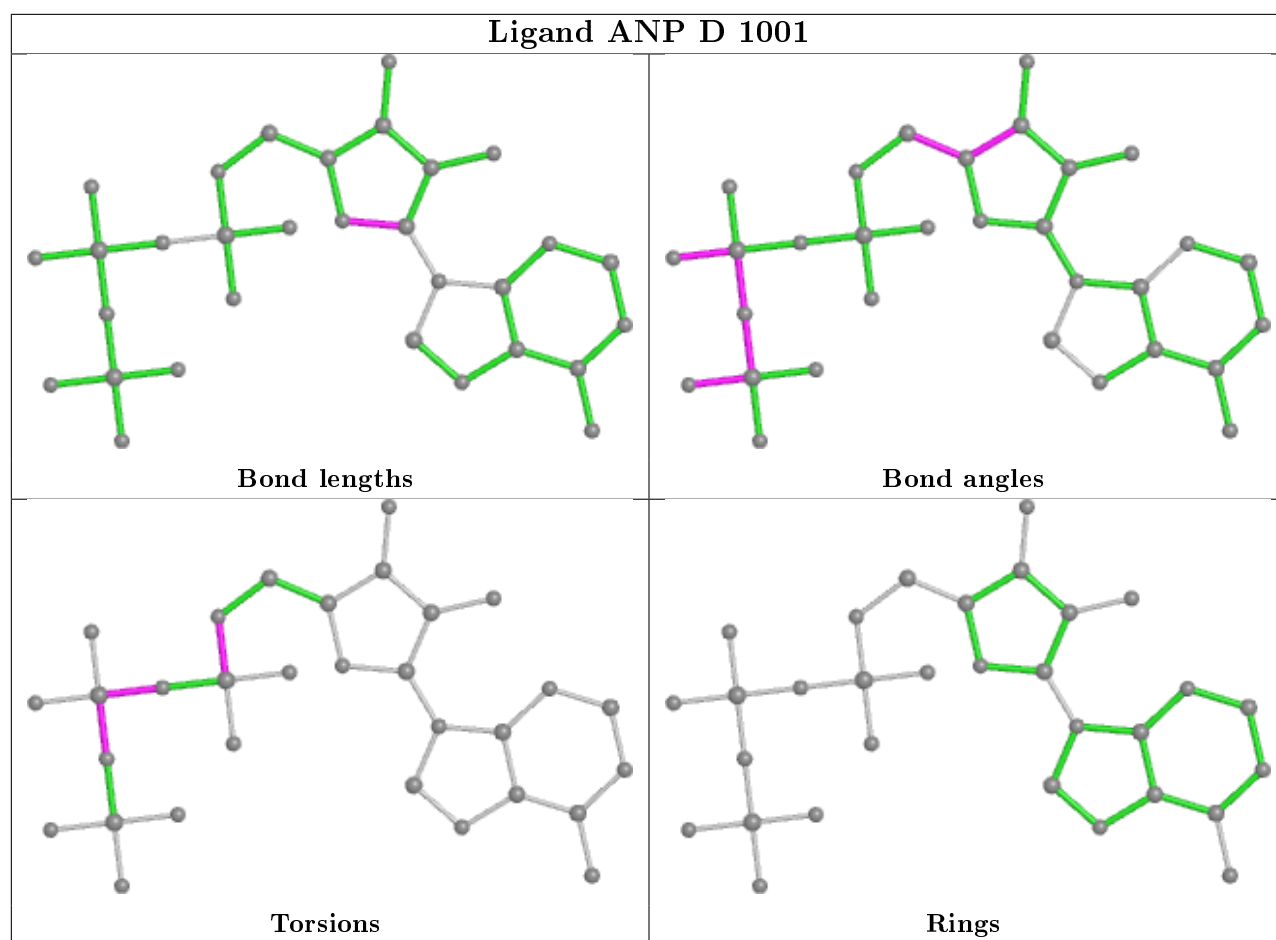












## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	478/500 (95%)	-0.28	1 (0%) 95 94	23, 37, 62, 84	0
1	B	480/500 (96%)	-0.08	9 (1%) 66 65	29, 43, 73, 94	1 (0%)
1	C	481/500 (96%)	-0.08	10 (2%) 63 61	25, 40, 73, 125	0
1	D	479/500 (95%)	0.06	17 (3%) 44 42	27, 44, 74, 111	0
All	All	1918/2000 (95%)	-0.09	37 (1%) 66 65	23, 41, 72, 125	1 (0%)

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	415	PHE	7.0
1	C	627	PHE	4.0
1	D	596	LEU	3.6
1	C	687	ALA	3.5
1	D	629	ASP	3.4
1	D	334	VAL	3.2
1	B	810	GLY	3.2
1	B	779	ILE	3.0
1	D	408	GLU	3.0
1	B	781	ALA	3.0
1	D	742	PHE	3.0
1	D	333	MET	2.9
1	D	809	GLU	2.9
1	B	575	GLU	2.8
1	B	627	PHE	2.8
1	D	743	ASP	2.8
1	D	709	LEU	2.8
1	C	779	ILE	2.7
1	C	415	PHE	2.6
1	C	781	ALA	2.5
1	A	333	MET	2.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	375	TRP	2.4
1	C	465	TRP	2.3
1	C	372	PHE	2.3
1	D	593	THR	2.3
1	D	591	CYS	2.3
1	B	816	THR	2.2
1	C	628	LYS	2.2
1	B	811	THR	2.2
1	D	414	GLY	2.2
1	D	814	PHE	2.1
1	C	743	ASP	2.1
1	D	403	HIS	2.1
1	D	592	THR	2.1
1	D	750	ASN	2.0
1	D	808	PRO	2.0
1	B	687	ALA	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	MG	C	1004	1/1	0.88	0.05	44,44,44,44	0
4	MG	B	1003	1/1	0.90	0.12	36,36,36,36	0
3	HFG	D	1002	24/24	0.92	0.12	27,31,40,67	0
3	HFG	A	1002	24/24	0.92	0.12	27,32,42,59	0
3	HFG	C	1002	24/24	0.93	0.12	26,31,43,64	0
3	HFG	B	1002	24/24	0.93	0.12	30,34,39,66	0
4	MG	A	1003	1/1	0.94	0.05	32,32,32,32	0

*Continued on next page...*

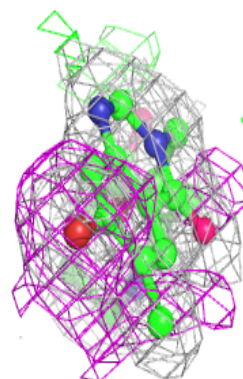
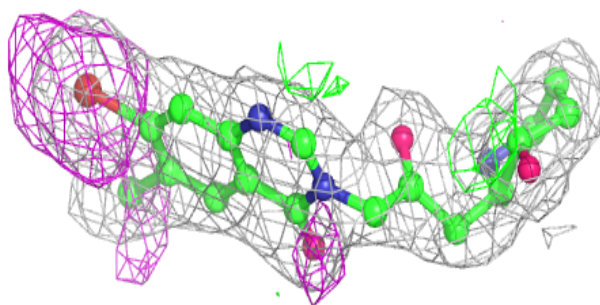
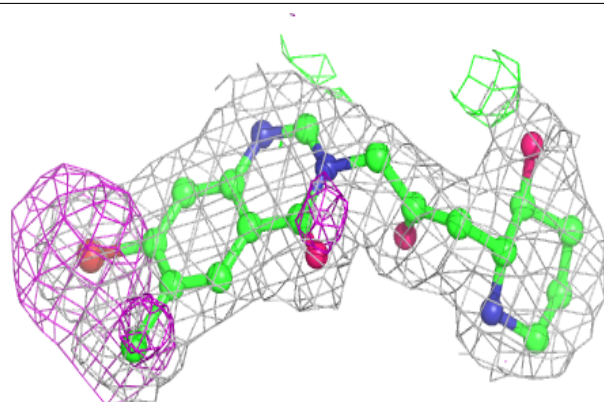
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MG	A	1004	1/1	0.94	0.03	51,51,51,51	0
4	MG	B	1004	1/1	0.95	0.06	56,56,56,56	0
4	MG	D	1004	1/1	0.96	0.03	50,50,50,50	0
4	MG	C	1003	1/1	0.96	0.10	37,37,37,37	0
2	ANP	D	1001	31/31	0.97	0.15	26,32,41,53	0
2	ANP	A	1001	31/31	0.97	0.13	27,30,35,36	0
2	ANP	B	1001	31/31	0.97	0.13	30,35,41,42	0
4	MG	D	1003	1/1	0.98	0.13	29,29,29,29	0
2	ANP	C	1001	31/31	0.98	0.13	28,32,35,39	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.

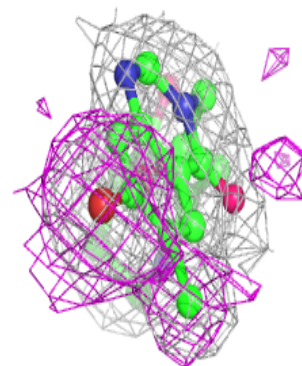
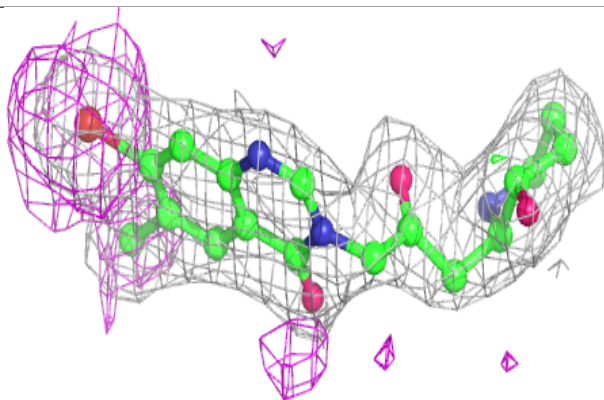
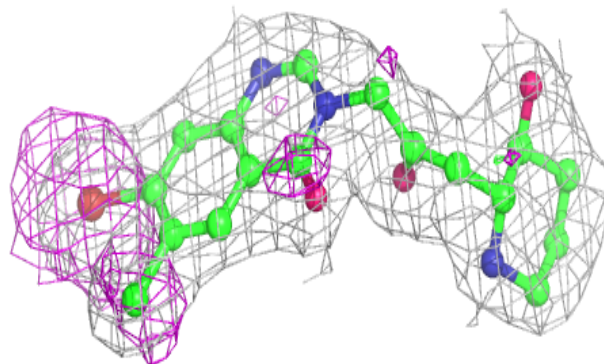
**Electron density around HFG D 1002:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

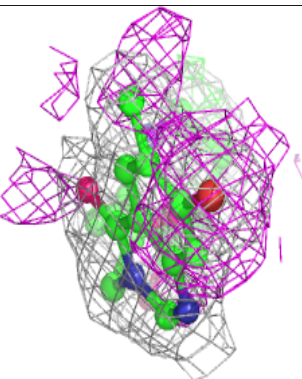
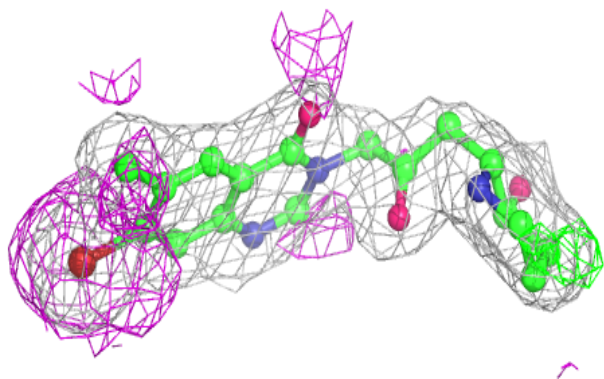
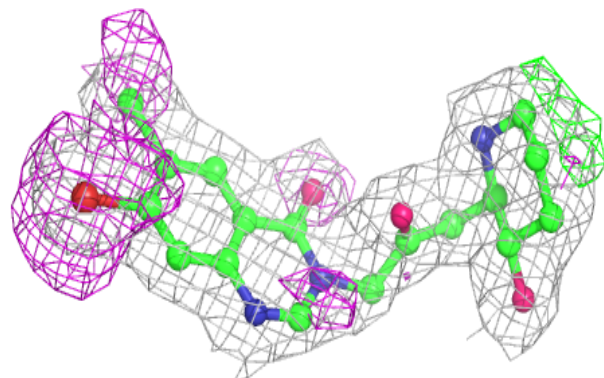


**Electron density around HFG A 1002:**

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

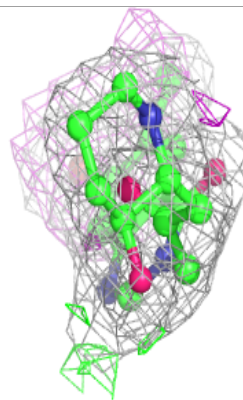
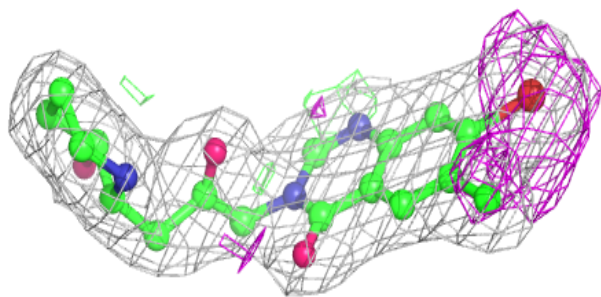
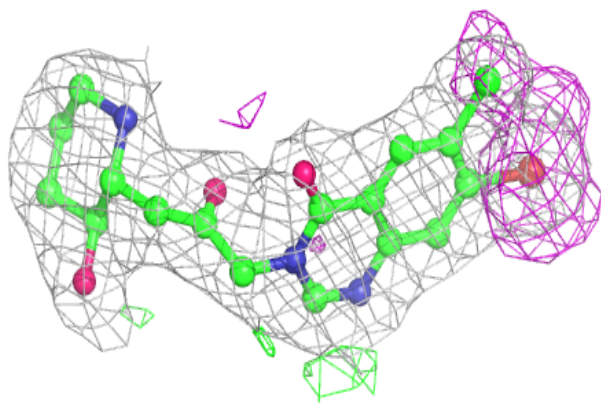
**Electron density around HFG C 1002:**

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



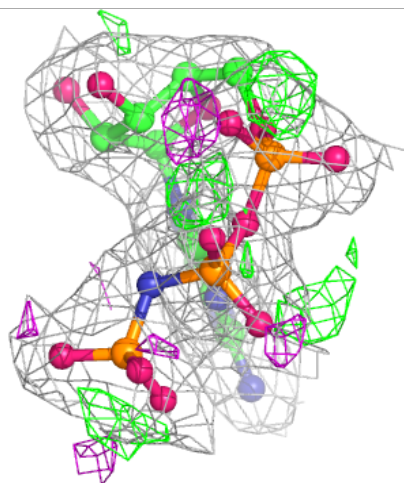
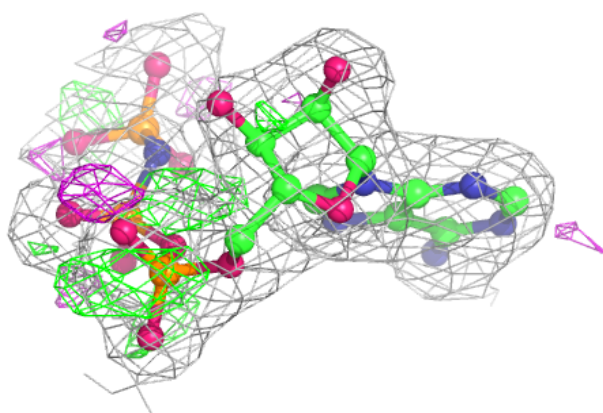
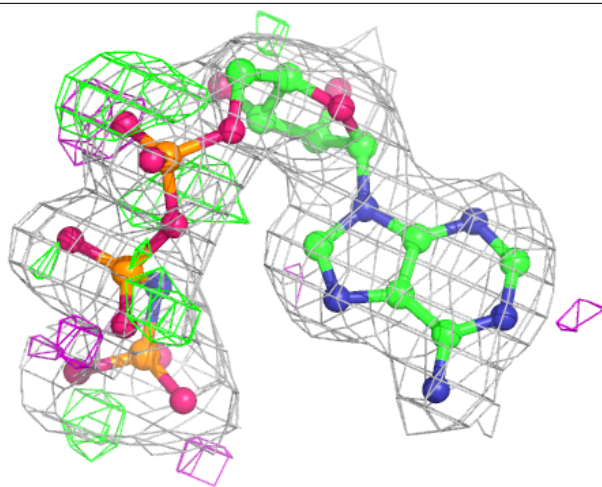
**Electron density around HFG B 1002:**

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



**Electron density around ANP D 1001:**

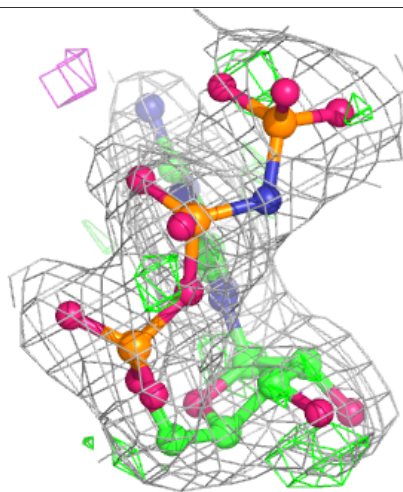
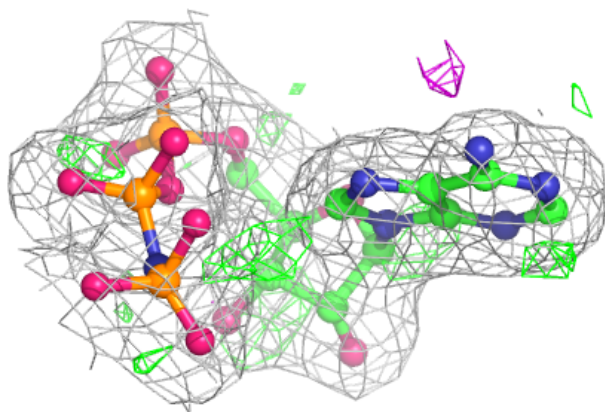
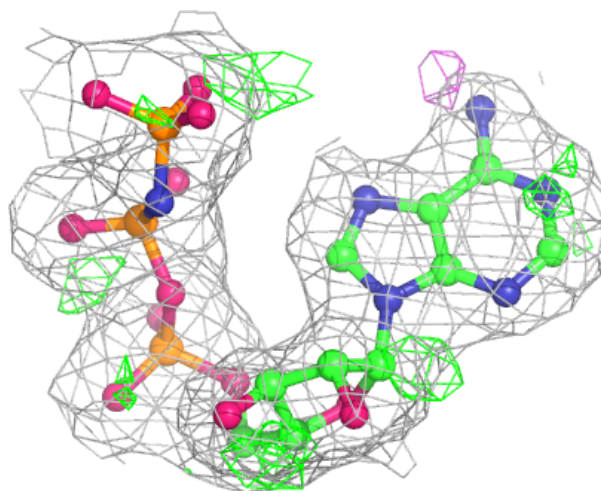
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





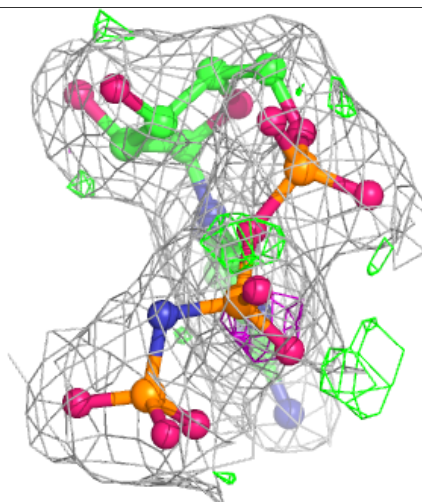
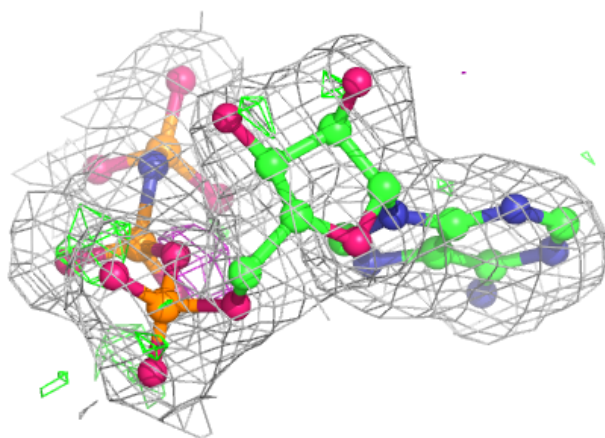
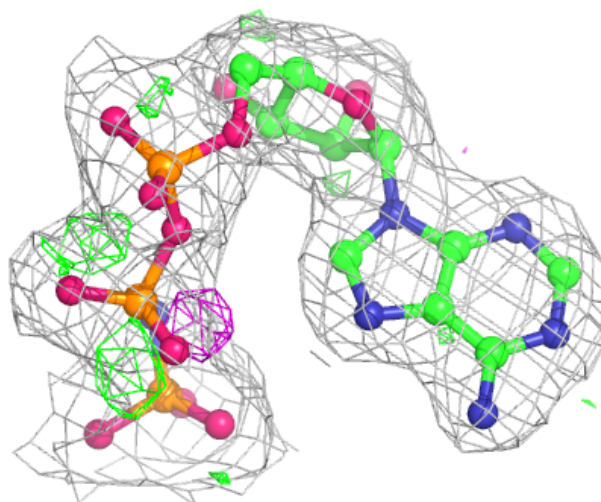
**Electron density around ANP A 1001:**

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



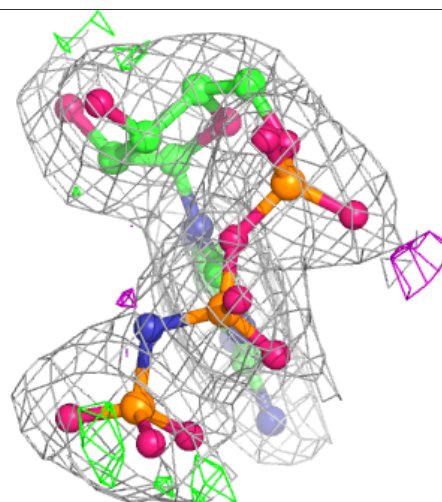
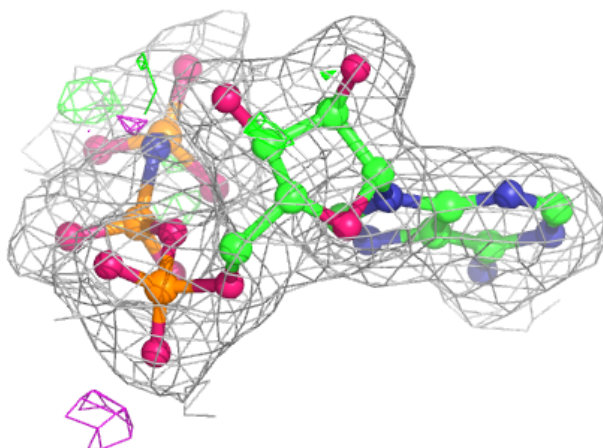
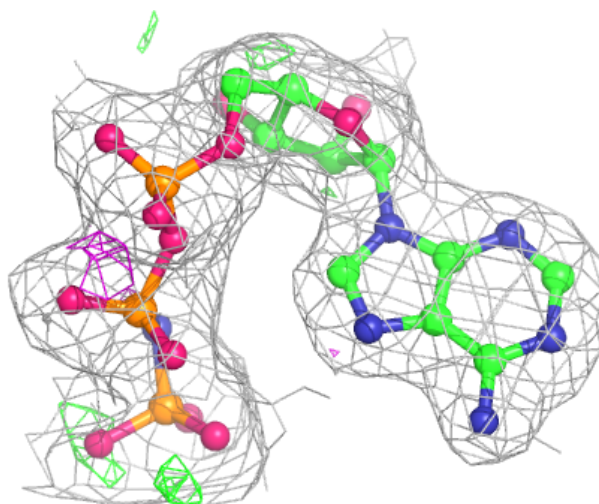
**Electron density around ANP B 1001:**

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



**Electron density around ANP C 1001:**

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



## 6.5 Other polymers ⓘ

There are no such residues in this entry.