



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 06:28 am BST

PDB ID : 4PEG  
Title : Dbr1 in complex with guanosine-5'-monophosphate  
Authors : Montemayor, E.J.; Katolik, A.; Clark, N.E.; Taylor, A.B.; Schuermann, J.P.; Combs, D.J.; Johnsson, R.; Holloway, S.P.; Stevens, S.W.; Damha, M.J.; Hart, P.J.  
Deposited on : 2014-04-23  
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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

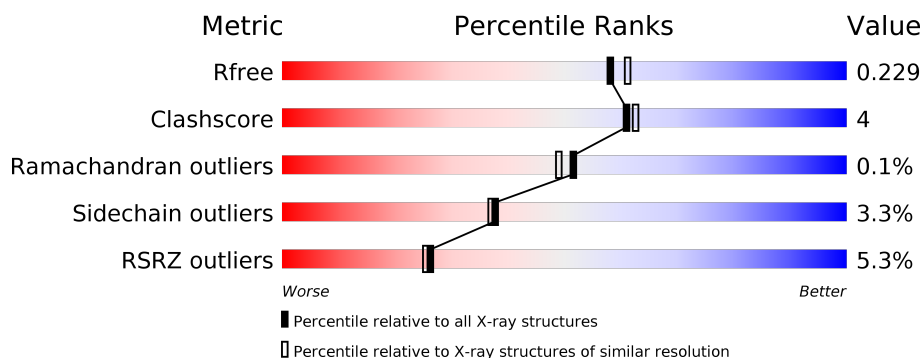
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	360	<div> <div>3%</div> <div>87%</div> <div>9%</div> <div>..</div> </div>
1	B	360	<div> <div>3%</div> <div>86%</div> <div>9%</div> <div>..</div> </div>
1	C	360	<div> <div>5%</div> <div>84%</div> <div>13%</div> <div>..</div> </div>
1	D	360	<div> <div>8%</div> <div>85%</div> <div>10%</div> <div>..</div> </div>
1	E	360	<div> <div>8%</div> <div>86%</div> <div>10%</div> <div>..</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 15560 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 RNA lariat debranching enzyme, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	349	Total	C	N	O	S	0	2	0
			2868	1863	467	522	16			
1	B	349	Total	C	N	O	S	0	1	0
			2863	1860	466	521	16			
1	C	350	Total	C	N	O	S	0	2	0
			2877	1868	468	526	15			
1	D	349	Total	C	N	O	S	0	0	0
			2857	1857	465	520	15			
1	E	349	Total	C	N	O	S	0	1	0
			2863	1860	466	521	16			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	355	HIS	-	expression tag	UNP C4M1P9
A	356	HIS	-	expression tag	UNP C4M1P9
A	357	HIS	-	expression tag	UNP C4M1P9
A	358	HIS	-	expression tag	UNP C4M1P9
A	359	HIS	-	expression tag	UNP C4M1P9
A	360	HIS	-	expression tag	UNP C4M1P9
B	355	HIS	-	expression tag	UNP C4M1P9
B	356	HIS	-	expression tag	UNP C4M1P9
B	357	HIS	-	expression tag	UNP C4M1P9
B	358	HIS	-	expression tag	UNP C4M1P9
B	359	HIS	-	expression tag	UNP C4M1P9
B	360	HIS	-	expression tag	UNP C4M1P9
C	355	HIS	-	expression tag	UNP C4M1P9
C	356	HIS	-	expression tag	UNP C4M1P9
C	357	HIS	-	expression tag	UNP C4M1P9
C	358	HIS	-	expression tag	UNP C4M1P9
C	359	HIS	-	expression tag	UNP C4M1P9
C	360	HIS	-	expression tag	UNP C4M1P9
D	355	HIS	-	expression tag	UNP C4M1P9

*Continued on next page...*

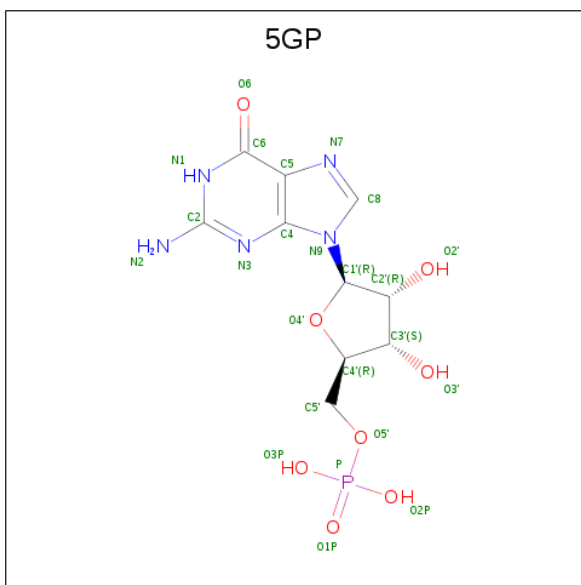
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	356	HIS	-	expression tag	UNP C4M1P9
D	357	HIS	-	expression tag	UNP C4M1P9
D	358	HIS	-	expression tag	UNP C4M1P9
D	359	HIS	-	expression tag	UNP C4M1P9
D	360	HIS	-	expression tag	UNP C4M1P9
E	355	HIS	-	expression tag	UNP C4M1P9
E	356	HIS	-	expression tag	UNP C4M1P9
E	357	HIS	-	expression tag	UNP C4M1P9
E	358	HIS	-	expression tag	UNP C4M1P9
E	359	HIS	-	expression tag	UNP C4M1P9
E	360	HIS	-	expression tag	UNP C4M1P9

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

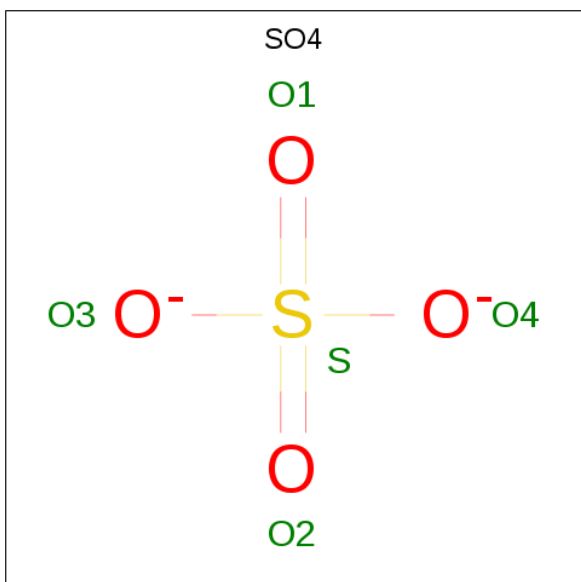
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Mn 1	0	0
2	A	1	Total 1	Mn 1	0	0
2	D	1	Total 1	Mn 1	0	0
2	C	1	Total 1	Mn 1	0	0
2	E	1	Total 1	Mn 1	0	0

- Molecule 3 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>8</sub>P).



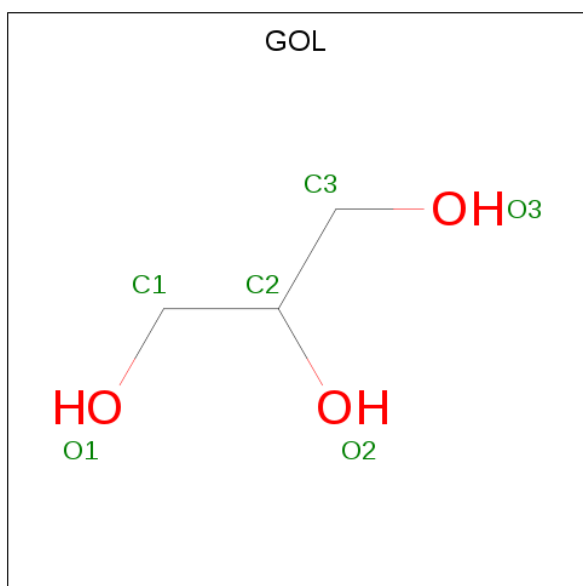
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
3	B	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
3	C	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
3	D	1	Total	C	N	O	P	0	0
			24	10	5	8	1		
3	E	1	Total	C	N	O	P	0	0
			24	10	5	8	1		

- Molecule 4 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	227	Total O 227 227	0	0
6	B	241	Total O 241 241	0	0
6	C	213	Total O 213 213	0	0

*Continued on next page...*

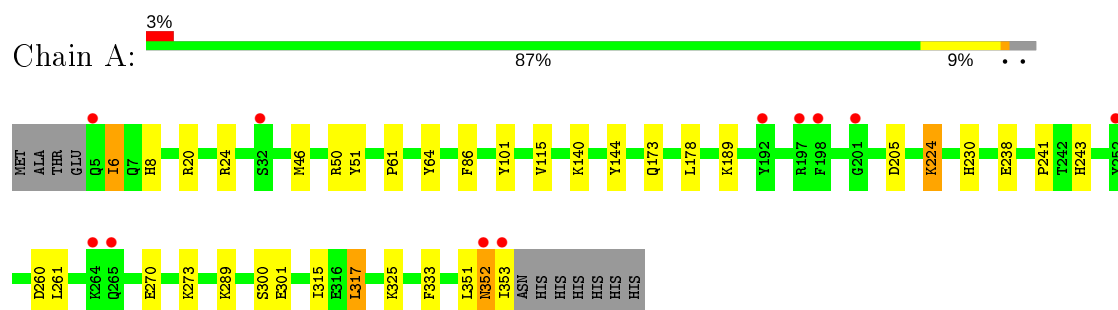
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	212	Total 212	O 212	0	0
6	E	177	Total 177	O 177	0	0

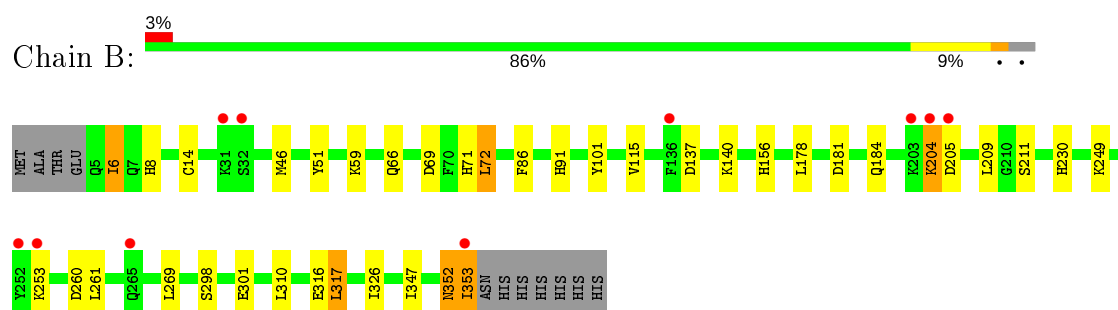
### 3 Residue-property plots [i](#)

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.

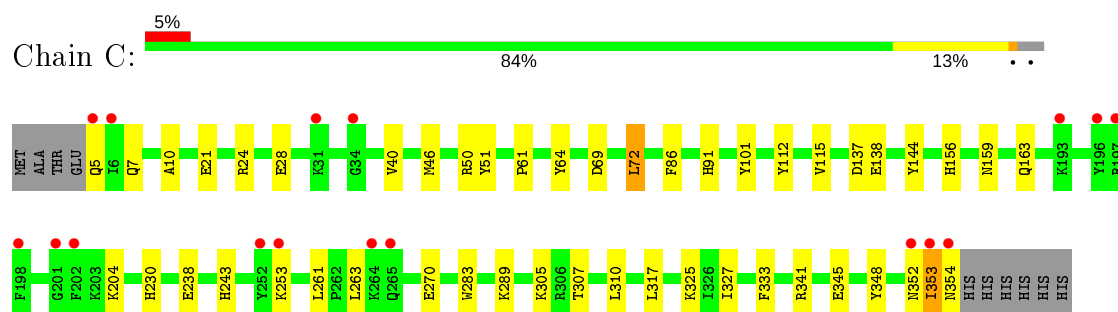
- Molecule 1: RNA lariat debranching enzyme, putative



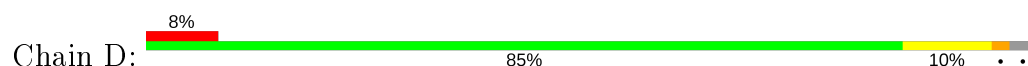
- Molecule 1: RNA lariat debranching enzyme, putative



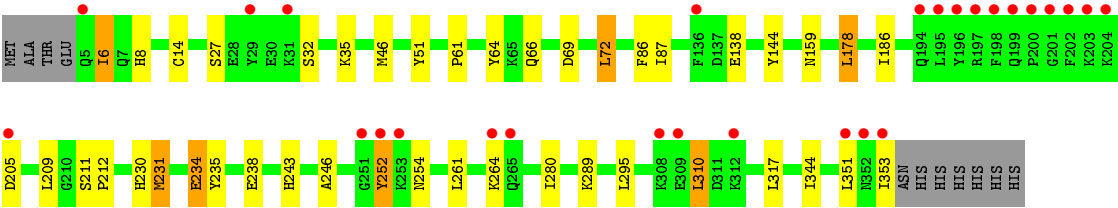
- Molecule 1: RNA lariat debranching enzyme, putative



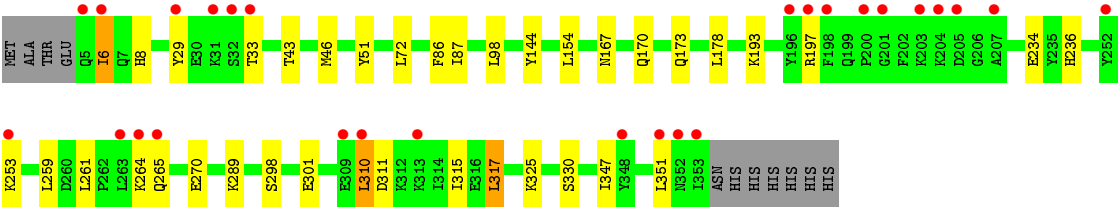
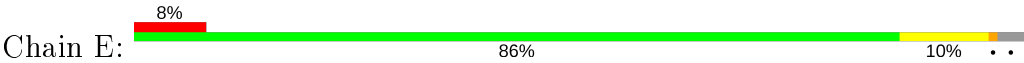
- Molecule 1: RNA lariat debranching enzyme, putative







● Molecule 1: RNA lariat debranching enzyme, putative



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.83Å 141.50Å 212.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.54 – 2.00 42.54 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.4 (42.54-2.00) 96.4 (42.54-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.73 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.197 , 0.229 0.197 , 0.229	Depositor DCC
$R_{free}$ test set	7198 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.3	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15560	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 5GP, SO4, MN

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	0.28	0/2949	0.44	0/3986
1	B	0.30	0/2941	0.45	0/3975
1	C	0.29	0/2958	0.47	0/3998
1	D	0.28	0/2935	0.47	0/3967
1	E	0.30	0/2941	0.46	0/3975
All	All	0.29	0/14724	0.46	0/19901

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	2868	0	2842	24	0
1	B	2863	0	2836	23	0
1	C	2877	0	2848	34	0
1	D	2857	0	2832	26	0
1	E	2863	0	2836	20	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1	0	0	0	0
2	E	1	0	0	0	0
3	A	24	0	12	1	0
3	B	24	0	12	1	0
3	C	24	0	12	1	0
3	D	24	0	12	1	0
3	E	24	0	12	0	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
4	E	5	0	0	0	0
5	A	6	0	8	2	0
5	B	6	0	8	0	0
6	A	227	0	0	6	0
6	B	241	0	0	7	0
6	C	213	0	0	8	0
6	D	212	0	0	3	0
6	E	177	0	0	1	0
All	All	15560	0	14270	129	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 4.

All (129) 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:D:353:ILE:H	1:D:353:ILE:HD12	1.39	0.87
1:D:14:CYS:SG	6:D:712:HOH:O	2.34	0.85
1:C:28:GLU:HG2	6:C:695:HOH:O	1.77	0.84
1:C:353:ILE:HD13	1:C:353:ILE:H	1.42	0.82
1:A:352:ASN:O	1:A:353:ILE:HG13	1.79	0.80
1:B:14[A]:CYS:SG	6:B:626:HOH:O	2.45	0.74
1:D:353:ILE:H	1:D:353:ILE:CD1	2.00	0.73
1:C:353:ILE:N	1:C:353:ILE:HD13	2.03	0.72
1:D:209:LEU:HD21	1:D:231:MET:HE2	1.71	0.72
1:C:163:GLN:NE2	6:C:577:HOH:O	2.24	0.70
1:A:224:LYS:NZ	1:A:241:PRO:O	2.25	0.70
1:B:352:ASN:O	1:B:353:ILE:HG23	1.93	0.68
1:C:348:TYR:CZ	1:C:353:ILE:HG13	2.29	0.67
1:C:5:GLN:CG	6:C:706:HOH:O	2.41	0.67

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:353:ILE:HD12	1:D:353:ILE:N	2.09	0.67
1:C:138:GLU:HG2	1:C:159:ASN:HB2	1.75	0.66
1:C:5:GLN:HG2	6:C:706:HOH:O	1.96	0.65
1:C:238:GLU:OE1	1:C:243:HIS:ND1	2.26	0.63
1:B:230:HIS:HA	6:B:626:HOH:O	1.98	0.63
1:A:351:LEU:O	1:A:352:ASN:HB2	1.98	0.62
1:C:230:HIS:HA	6:C:663:HOH:O	1.99	0.62
1:C:307:THR:HG22	1:C:310:LEU:HB2	1.82	0.62
1:C:341:ARG:HG3	6:C:581:HOH:O	2.00	0.61
1:D:234:GLU:HB2	1:D:254:ASN:HB2	1.82	0.61
1:B:6:ILE:HD11	1:B:8:HIS:CE1	2.35	0.61
1:A:352:ASN:C	1:A:353:ILE:HG13	2.21	0.61
1:A:238:GLU:OE1	1:A:243:HIS:ND1	2.32	0.61
1:D:46:MET:HG3	1:D:86:PHE:CD2	2.38	0.59
1:A:61:PRO:HG2	1:A:64:TYR:HD2	1.68	0.58
1:C:21[B]:GLU:OE1	1:C:24:ARG:NH1	2.35	0.58
1:C:7:GLN:HB3	1:C:263:LEU:HD13	1.86	0.57
1:D:6:ILE:HD11	1:D:8:HIS:CE1	2.38	0.57
1:A:270:GLU:HG2	1:A:325:LYS:HD3	1.88	0.56
1:D:8:HIS:CE1	1:D:35:LYS:HG2	2.40	0.56
1:E:259:LEU:HB3	1:E:261:LEU:HD13	1.87	0.56
5:A:404:GOL:H31	6:A:680:HOH:O	2.06	0.55
1:C:46:MET:HG3	1:C:86:PHE:CD2	2.41	0.55
1:B:59:LYS:HE2	6:B:695:HOH:O	2.06	0.55
1:E:298:SER:HB3	1:E:301:GLU:HB2	1.89	0.55
1:A:173:GLN:HG2	6:A:511:HOH:O	2.07	0.54
1:A:6:ILE:HD11	1:A:8:HIS:CE1	2.43	0.54
1:E:46:MET:HG3	1:E:86:PHE:CD2	2.41	0.54
1:B:204:LYS:HG2	1:B:205:ASP:N	2.23	0.54
1:C:61:PRO:HG2	1:C:64:TYR:HD2	1.72	0.54
1:A:230:HIS:HA	6:A:629:HOH:O	2.07	0.53
1:E:347:ILE:O	1:E:351:LEU:HG	2.09	0.53
1:D:138:GLU:HG2	1:D:159:ASN:HB2	1.90	0.53
1:E:144:TYR:HB3	1:E:289:LYS:O	2.10	0.52
1:D:351:LEU:HB2	1:D:353:ILE:CD1	2.39	0.52
1:E:167:ASN:O	1:E:170:GLN:HG2	2.09	0.52
1:C:353:ILE:HG12	1:C:354:ASN:H	1.73	0.52
1:B:46:MET:HG3	1:B:86:PHE:CD2	2.45	0.52
1:A:224:LYS:HE3	1:A:243:HIS:CD2	2.45	0.52
1:C:270:GLU:HG2	1:C:325:LYS:HD3	1.92	0.52
1:B:269:LEU:HG	1:B:326:ILE:HD12	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:46:MET:HG3	1:A:86:PHE:CD1	2.45	0.51
1:A:325:LYS:HG2	6:A:713:HOH:O	2.10	0.51
1:C:353:ILE:HG12	1:C:354:ASN:N	2.26	0.51
1:C:10:ALA:HB3	1:C:40:VAL:HG22	1.92	0.50
1:C:5:GLN:HB3	6:C:706:HOH:O	2.11	0.50
1:E:317:LEU:HB3	1:E:347:ILE:HG12	1.94	0.50
1:E:270:GLU:HG2	1:E:325:LYS:HD3	1.94	0.49
1:B:137:ASP:OD2	1:B:156:HIS:ND1	2.38	0.49
1:C:144:TYR:HB3	1:C:289:LYS:O	2.13	0.49
1:C:50:ARG:HD3	1:C:333:PHE:CE2	2.48	0.49
1:C:353:ILE:CD1	1:C:353:ILE:N	2.73	0.48
1:D:144:TYR:HB3	1:D:289:LYS:O	2.13	0.48
1:E:43:THR:HG22	1:E:178:LEU:HD12	1.96	0.48
1:A:353:ILE:O	1:A:353:ILE:HG22	2.11	0.48
1:B:140:LYS:HE3	6:B:717:HOH:O	2.13	0.48
1:C:5:GLN:CB	6:C:706:HOH:O	2.61	0.47
1:A:140:LYS:HE3	6:A:515:HOH:O	2.14	0.47
5:A:404:GOL:C3	6:A:680:HOH:O	2.60	0.47
1:D:61:PRO:HG2	1:D:64:TYR:HD2	1.80	0.47
1:E:234:GLU:OE2	1:E:236:HIS:NE2	2.48	0.47
1:B:301:GLU:HG2	6:B:614:HOH:O	2.13	0.47
1:B:298:SER:HB3	1:B:301:GLU:HB3	1.97	0.47
1:E:310:LEU:HA	1:E:310:LEU:HD13	1.75	0.47
1:D:235:TYR:HB3	1:D:246:ALA:HB3	1.97	0.46
1:B:14[A]:CYS:HB2	1:B:249:LYS:HA	1.97	0.46
1:B:71:HIS:CE1	1:B:72:LEU:HD13	2.51	0.46
1:C:101:TYR:HB2	1:C:115:VAL:HG23	1.98	0.46
1:B:91:HIS:NE2	3:B:402:5GP:O3P	2.38	0.46
1:E:310:LEU:HD12	1:E:351:LEU:HD22	1.98	0.46
1:B:69:ASP:O	1:B:72:LEU:HB2	2.16	0.46
1:E:98:LEU:HD22	1:E:154:LEU:HD22	1.98	0.46
1:C:137:ASP:OD2	1:C:156:HIS:ND1	2.45	0.46
1:E:310:LEU:CD1	1:E:351:LEU:HD22	2.45	0.46
1:C:204:LYS:HB2	1:C:204:LYS:HE3	1.67	0.46
1:E:6:ILE:HD11	1:E:8:HIS:CE1	2.51	0.45
1:D:69:ASP:O	1:D:72:LEU:HB2	2.16	0.45
1:A:144:TYR:HB3	1:A:289:LYS:O	2.17	0.45
1:C:283:TRP:CZ2	1:C:345:GLU:HB2	2.52	0.45
1:A:6:ILE:HD12	1:A:260:ASP:HB3	1.97	0.45
1:B:317:LEU:HB3	1:B:347:ILE:HG12	1.99	0.45
1:B:6:ILE:HD13	1:B:260:ASP:HB3	1.99	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:316:GLU:OE1	6:B:654:HOH:O	2.21	0.44
1:D:211:SER:HA	1:D:212:PRO:HD2	1.89	0.44
1:C:86:PHE:CZ	1:C:112:TYR:HB2	2.53	0.44
1:D:351:LEU:HB2	1:D:353:ILE:HD12	2.00	0.44
1:C:91:HIS:NE2	3:C:402:5GP:O3P	2.42	0.43
1:D:252:TYR:N	1:D:252:TYR:CD1	2.85	0.43
1:D:252:TYR:HD1	1:D:252:TYR:N	2.15	0.43
1:B:101:TYR:HB2	1:B:115:VAL:HG23	2.01	0.43
1:C:307:THR:CG2	1:C:310:LEU:H	2.31	0.43
1:A:189:LYS:HA	1:A:189:LYS:HD2	1.78	0.43
1:B:181:ASP:OD2	1:B:211:SER:HB2	2.19	0.43
1:E:29:TYR:O	1:E:33:THR:OG1	2.24	0.43
1:C:69:ASP:O	1:C:72:LEU:HB2	2.20	0.42
1:D:87:ILE:HD13	1:D:178:LEU:HB3	2.01	0.42
1:D:280:ILE:HD13	1:D:344:ILE:HG23	2.02	0.42
1:A:205:ASP:OD2	3:A:402:5GP:N2	2.42	0.42
1:E:167:ASN:CG	6:E:663:HOH:O	2.58	0.42
1:A:273:LYS:HE3	1:A:315:ILE:HG21	2.02	0.42
1:D:230:HIS:HA	6:D:712:HOH:O	2.20	0.42
1:D:205:ASP:OD2	3:D:402:5GP:N2	2.42	0.41
1:D:238:GLU:CD	1:D:243:HIS:HD1	2.21	0.41
1:E:311:ASP:O	1:E:315:ILE:HG12	2.20	0.41
1:E:87:ILE:HD13	1:E:178:LEU:HB3	2.02	0.41
1:B:184:GLN:HA	1:B:209:LEU:O	2.20	0.41
1:E:193:LYS:O	1:E:197:ARG:HG3	2.19	0.41
1:A:50:ARG:HD3	1:A:333:PHE:CE2	2.56	0.41
1:B:310:LEU:HB3	6:B:648:HOH:O	2.20	0.41
1:D:310:LEU:HB3	6:D:646:HOH:O	2.21	0.41
1:C:348:TYR:OH	1:C:353:ILE:HG13	2.21	0.41
1:A:317:LEU:HD12	1:A:317:LEU:HA	1.89	0.41
1:A:20:ARG:O	1:A:24:ARG:HG3	2.20	0.41
1:D:295:LEU:HD23	1:D:295:LEU:HA	1.89	0.40
1:A:101:TYR:HB2	1:A:115:VAL:HG23	2.02	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	349/360 (97%)	340 (97%)	8 (2%)	1 (0%)	41	37
1	B	348/360 (97%)	340 (98%)	7 (2%)	1 (0%)	41	37
1	C	350/360 (97%)	342 (98%)	8 (2%)	0	100	100
1	D	347/360 (96%)	338 (97%)	9 (3%)	0	100	100
1	E	348/360 (97%)	342 (98%)	6 (2%)	0	100	100
All	All	1742/1800 (97%)	1702 (98%)	38 (2%)	2 (0%)	51	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	352	ASN
1	B	352	ASN

### 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	318/326 (98%)	310 (98%)	8 (2%)	47	49
1	B	317/326 (97%)	307 (97%)	10 (3%)	39	38
1	C	319/326 (98%)	310 (97%)	9 (3%)	43	44
1	D	316/326 (97%)	301 (95%)	15 (5%)	26	22
1	E	317/326 (97%)	307 (97%)	10 (3%)	39	38
All	All	1587/1630 (97%)	1535 (97%)	52 (3%)	38	37



All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	ILE
1	A	51	TYR
1	A	178	LEU
1	A	224	LYS
1	A	261	LEU
1	A	300	SER
1	A	301	GLU
1	A	317	LEU
1	B	6	ILE
1	B	51	TYR
1	B	66	GLN
1	B	72	LEU
1	B	178	LEU
1	B	204	LYS
1	B	253	LYS
1	B	261	LEU
1	B	317	LEU
1	B	353	ILE
1	C	51	TYR
1	C	72	LEU
1	C	253	LYS
1	C	261	LEU
1	C	305	LYS
1	C	317	LEU
1	C	327	ILE
1	C	352	ASN
1	C	353	ILE
1	D	6	ILE
1	D	27	SER
1	D	32	SER
1	D	51	TYR
1	D	66	GLN
1	D	72	LEU
1	D	178	LEU
1	D	186	ILE
1	D	231	MET
1	D	234	GLU
1	D	252	TYR
1	D	261	LEU
1	D	264	LYS
1	D	310	LEU
1	D	317	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	6	ILE
1	E	51	TYR
1	E	72	LEU
1	E	173	GLN
1	E	253	LYS
1	E	264	LYS
1	E	265	GLN
1	E	310	LEU
1	E	317	LEU
1	E	330	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	47	GLN
1	B	243	HIS
1	D	16	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](#)

Of 17 ligands modelled in this entry, 5 are monoatomic - leaving 12 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	5GP	E	402	2	22,26,26	0.54	0	27,40,40	1.66	5 (18%)
4	SO4	A	403	-	4,4,4	0.17	0	6,6,6	0.09	0
3	5GP	C	402	2	22,26,26	0.55	0	27,40,40	1.67	5 (18%)
3	5GP	A	402	2	22,26,26	0.52	0	27,40,40	1.62	5 (18%)
4	SO4	E	403	-	4,4,4	0.15	0	6,6,6	0.08	0
4	SO4	B	403	-	4,4,4	0.15	0	6,6,6	0.06	0
5	GOL	B	404	-	5,5,5	0.22	0	5,5,5	0.23	0
4	SO4	C	403	-	4,4,4	0.14	0	6,6,6	0.10	0
4	SO4	D	403	-	4,4,4	0.15	0	6,6,6	0.11	0
5	GOL	A	404	-	5,5,5	0.34	0	5,5,5	0.37	0
3	5GP	B	402	2	22,26,26	0.56	0	27,40,40	1.65	4 (14%)
3	5GP	D	402	2	22,26,26	0.56	0	27,40,40	1.73	6 (22%)

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	5GP	E	402	2	-	5/6/26/26	0/3/3/3
3	5GP	C	402	2	-	0/6/26/26	0/3/3/3
3	5GP	A	402	2	-	5/6/26/26	0/3/3/3
5	GOL	B	404	-	-	0/4/4/4	-
5	GOL	A	404	-	-	2/4/4/4	-
3	5GP	B	402	2	-	0/6/26/26	0/3/3/3
3	5GP	D	402	2	-	1/6/26/26	0/3/3/3

There are no bond length outliers.

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	5GP	N3-C2-N1	-4.49	121.23	127.22
3	D	402	5GP	N3-C2-N1	-4.49	121.24	127.22
3	C	402	5GP	N3-C2-N1	-4.45	121.28	127.22
3	E	402	5GP	N3-C2-N1	-4.33	121.45	127.22
3	A	402	5GP	N3-C2-N1	-4.19	121.63	127.22
3	B	402	5GP	C2-N3-C4	4.14	120.08	115.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	402	5GP	C2-N3-C4	3.91	119.82	115.36
3	C	402	5GP	C2-N3-C4	3.90	119.81	115.36
3	D	402	5GP	C2-N3-C4	3.75	119.64	115.36
3	A	402	5GP	C2-N3-C4	3.68	119.56	115.36
3	D	402	5GP	C6-N1-C2	3.16	120.95	115.93
3	C	402	5GP	C6-N1-C2	3.10	120.85	115.93
3	A	402	5GP	C5-C6-N1	-3.06	119.24	123.43
3	B	402	5GP	C6-N1-C2	3.02	120.73	115.93
3	D	402	5GP	C5-C6-N1	-3.00	119.33	123.43
3	E	402	5GP	C6-N1-C2	3.00	120.70	115.93
3	A	402	5GP	C6-N1-C2	3.00	120.69	115.93
3	E	402	5GP	C5-C6-N1	-2.99	119.34	123.43
3	C	402	5GP	C5-C6-N1	-2.97	119.36	123.43
3	B	402	5GP	C5-C6-N1	-2.89	119.48	123.43
3	D	402	5GP	C6-C5-C4	-2.22	118.67	120.80
3	A	402	5GP	C4-C5-N7	-2.10	107.21	109.40
3	D	402	5GP	C4-C5-N7	-2.08	107.23	109.40
3	E	402	5GP	C4-C5-N7	-2.07	107.24	109.40
3	C	402	5GP	C4-C5-N7	-2.04	107.27	109.40

There are no chirality outliers.

All (13) torsion outliers are listed below:

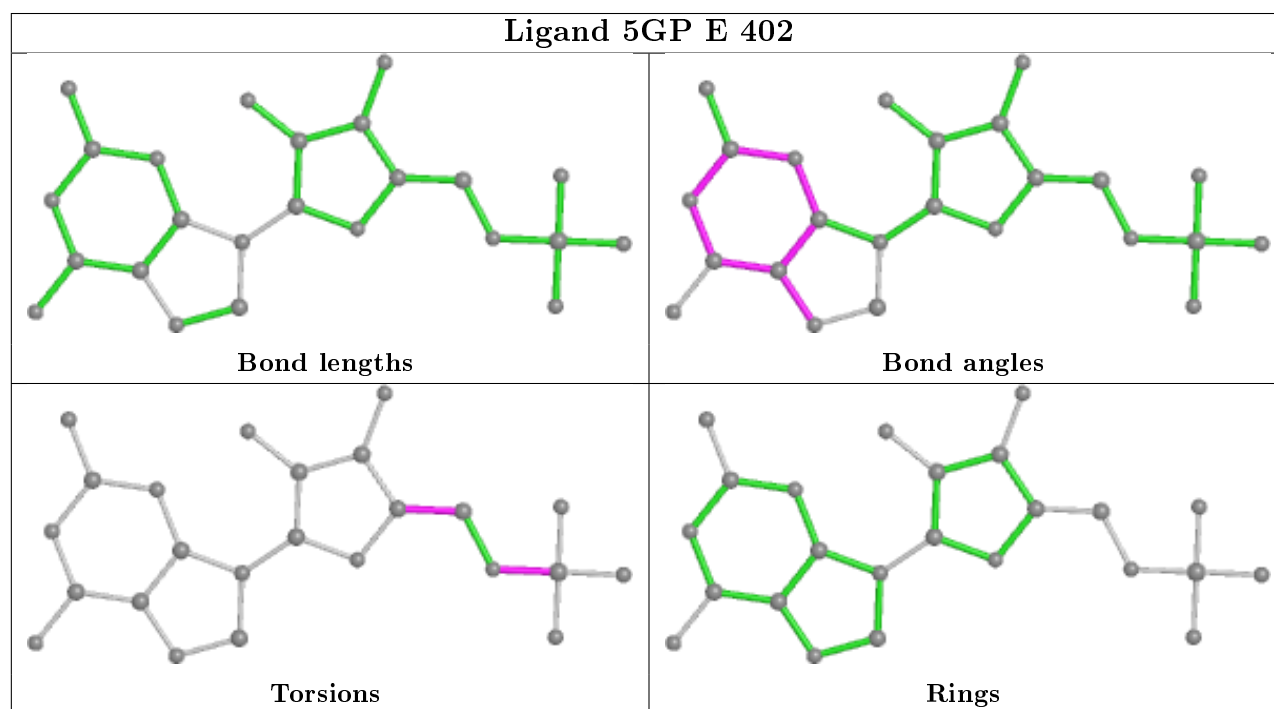
Mol	Chain	Res	Type	Atoms
3	E	402	5GP	C5'-O5'-P-O1P
3	E	402	5GP	C5'-O5'-P-O2P
3	E	402	5GP	C5'-O5'-P-O3P
3	E	402	5GP	C3'-C4'-C5'-O5'
3	A	402	5GP	C5'-O5'-P-O3P
5	A	404	GOL	O1-C1-C2-O2
3	E	402	5GP	O4'-C4'-C5'-O5'
3	A	402	5GP	O4'-C4'-C5'-O5'
3	A	402	5GP	C3'-C4'-C5'-O5'
5	A	404	GOL	O1-C1-C2-C3
3	A	402	5GP	C5'-O5'-P-O1P
3	A	402	5GP	C5'-O5'-P-O2P
3	D	402	5GP	O4'-C4'-C5'-O5'

There are no ring outliers.

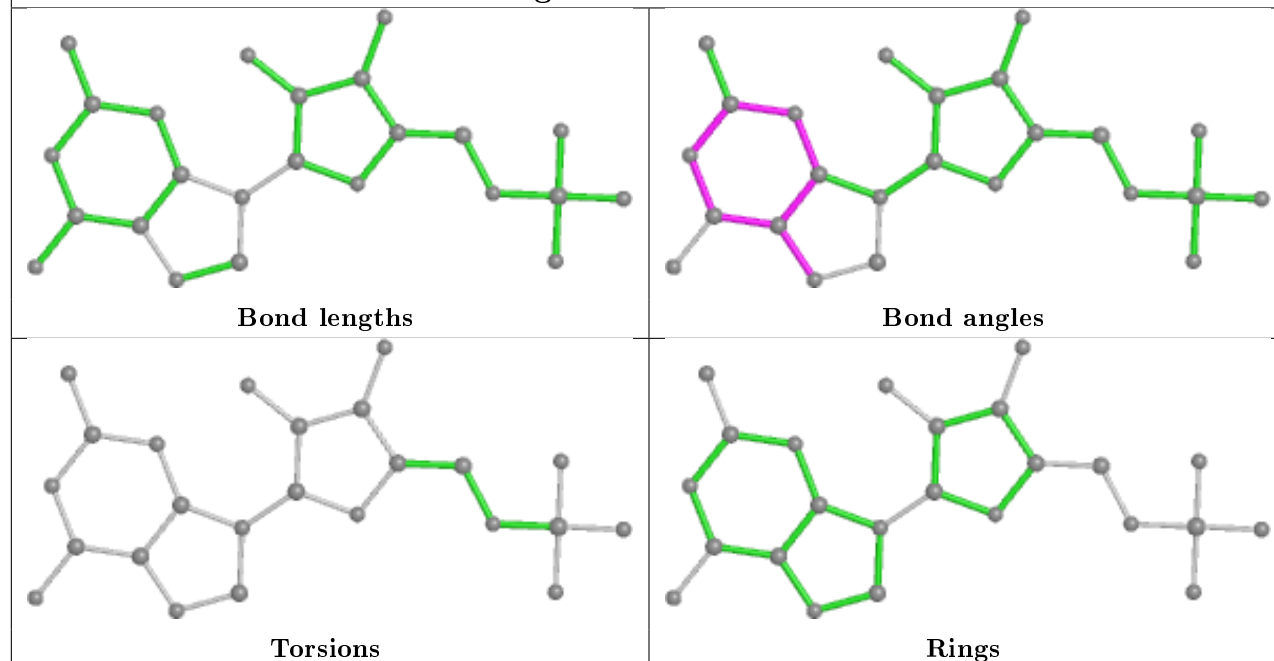
5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	402	5GP	1	0
3	A	402	5GP	1	0
5	A	404	GOL	2	0
3	B	402	5GP	1	0
3	D	402	5GP	1	0

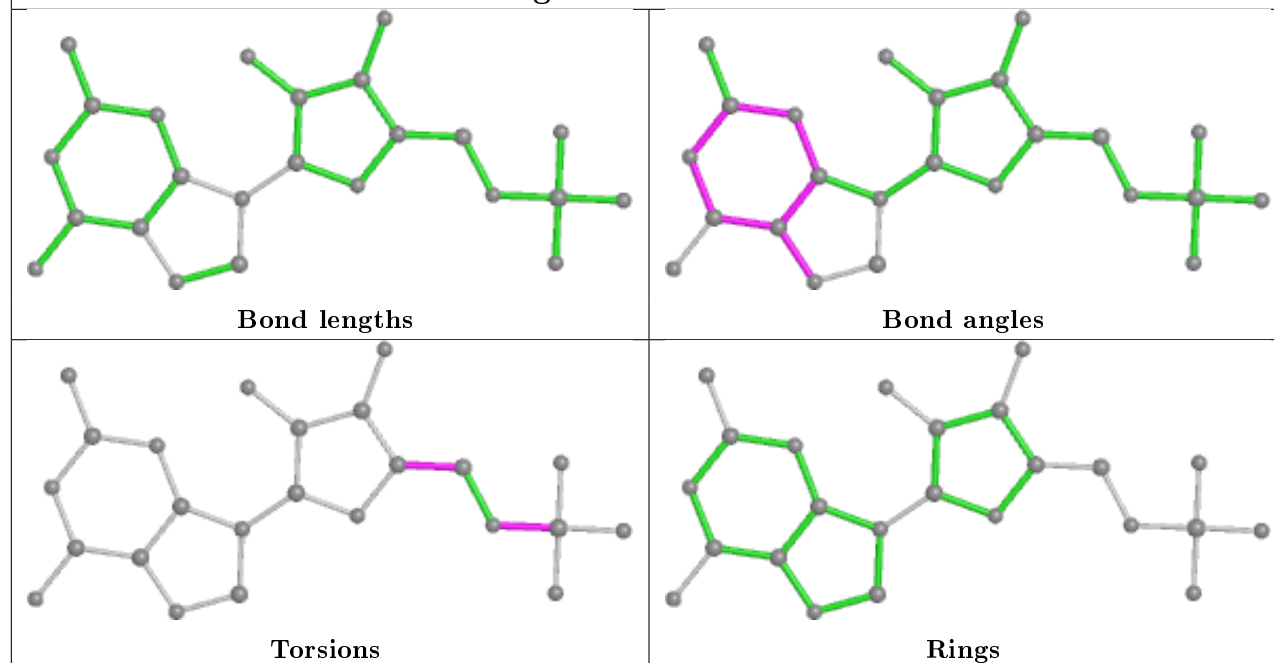
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

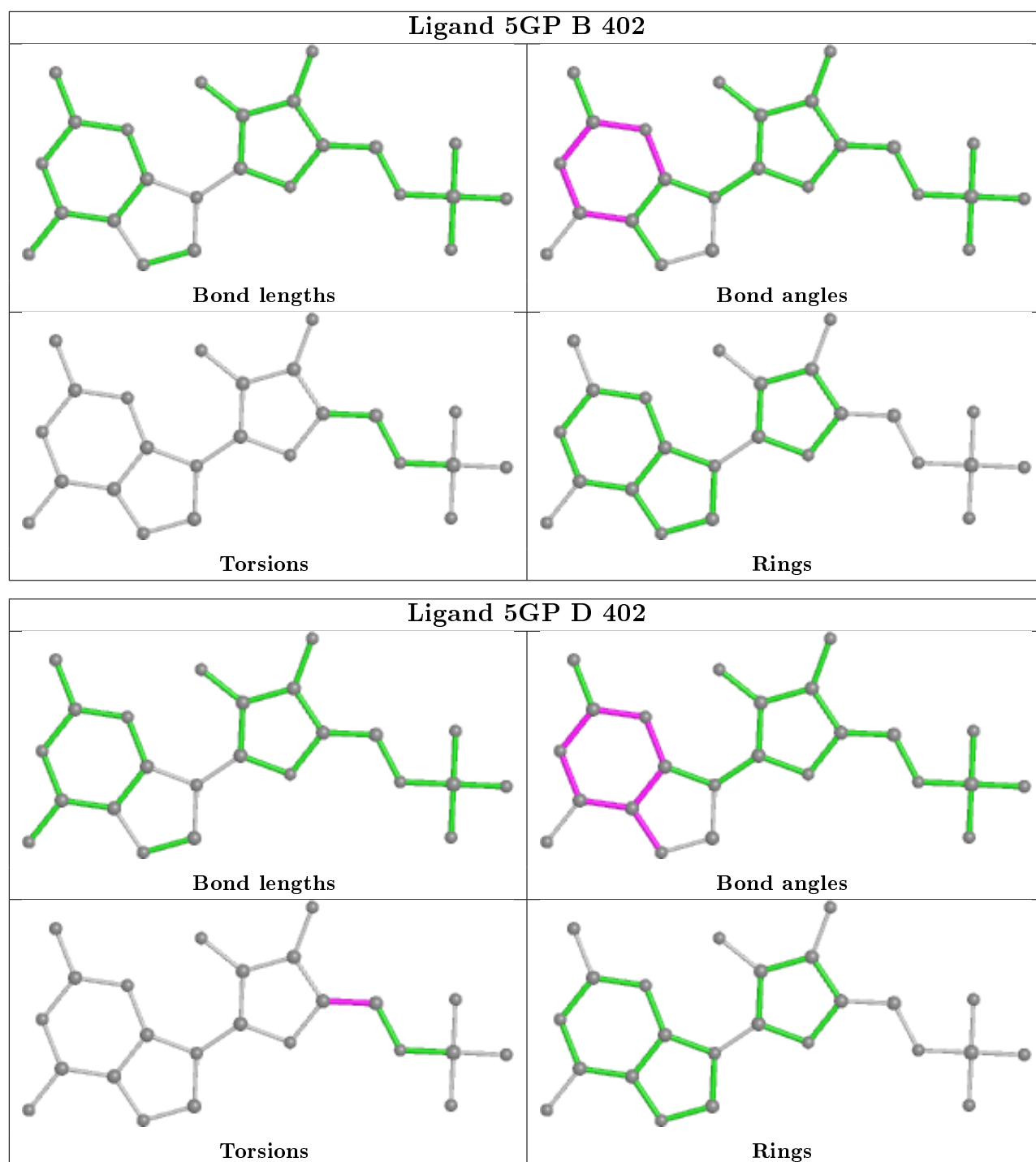


## Ligand 5GP C 402



## Ligand 5GP A 402





## 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	349/360 (96%)	0.11	11 (3%) 47 46	16, 27, 48, 69	0
1	B	349/360 (96%)	0.05	10 (2%) 51 50	16, 27, 46, 67	0
1	C	350/360 (97%)	0.37	17 (4%) 29 28	20, 30, 49, 81	0
1	D	349/360 (96%)	0.39	27 (7%) 13 12	16, 30, 62, 77	0
1	E	349/360 (96%)	0.44	27 (7%) 13 12	18, 33, 58, 80	0
All	All	1746/1800 (97%)	0.27	92 (5%) 26 25	16, 30, 55, 81	0

All (92) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	353	ILE	8.3
1	E	264	LYS	5.9
1	C	353	ILE	5.7
1	C	354	ASN	5.7
1	E	198	PHE	5.4
1	D	31	LYS	5.2
1	D	252	TYR	5.2
1	D	203	LYS	5.1
1	C	198	PHE	4.3
1	A	265	GLN	4.3
1	E	265	GLN	4.3
1	E	352	ASN	4.3
1	A	5	GLN	4.2
1	C	264	LYS	4.2
1	D	198	PHE	4.2
1	D	353	ILE	4.0
1	C	253	LYS	3.9
1	E	253	LYS	3.9
1	E	5	GLN	3.8
1	D	5	GLN	3.8

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	264	LYS	3.8
1	A	198	PHE	3.7
1	D	202	PHE	3.5
1	A	353	ILE	3.5
1	B	253	LYS	3.4
1	E	197	ARG	3.4
1	D	352	ASN	3.4
1	E	200	PRO	3.4
1	E	204	LYS	3.3
1	E	32	SER	3.2
1	D	195	LEU	3.2
1	E	351	LEU	3.2
1	D	265	GLN	3.1
1	D	29	TYR	3.1
1	D	136	PHE	3.1
1	D	197	ARG	3.1
1	A	352	ASN	3.1
1	C	31	LYS	3.0
1	E	201	GLY	2.9
1	E	29	TYR	2.9
1	C	202	PHE	2.8
1	D	196	TYR	2.8
1	D	205	ASP	2.8
1	C	352	ASN	2.8
1	B	136	PHE	2.8
1	C	265	GLN	2.7
1	D	253	LYS	2.7
1	B	252	TYR	2.7
1	E	31	LYS	2.7
1	D	309	GLU	2.7
1	A	32	SER	2.7
1	E	6	ILE	2.6
1	D	204	LYS	2.6
1	D	308	LYS	2.6
1	D	199	GLN	2.6
1	C	5	GLN	2.6
1	E	205	ASP	2.5
1	E	263	LEU	2.5
1	B	31	LYS	2.5
1	D	251	GLY	2.5
1	C	34	GLY	2.5
1	E	207	ALA	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	204	LYS	2.5
1	E	203	LYS	2.5
1	D	200	PRO	2.4
1	E	309	GLU	2.4
1	C	252	TYR	2.4
1	E	33	THR	2.4
1	D	201	GLY	2.4
1	A	252	TYR	2.4
1	C	6	ILE	2.3
1	C	193	LYS	2.3
1	C	197	ARG	2.3
1	B	32	SER	2.3
1	D	312	LYS	2.3
1	B	265	GLN	2.3
1	B	353	ILE	2.3
1	A	201	GLY	2.2
1	E	313	LYS	2.2
1	E	196	TYR	2.2
1	A	264	LYS	2.2
1	B	203	LYS	2.2
1	E	310	LEU	2.2
1	D	351	LEU	2.2
1	C	201	GLY	2.2
1	D	194	GLN	2.1
1	E	252	TYR	2.1
1	E	348	TYR	2.1
1	C	196	TYR	2.1
1	B	205	ASP	2.1
1	A	192	TYR	2.0
1	A	197	ARG	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

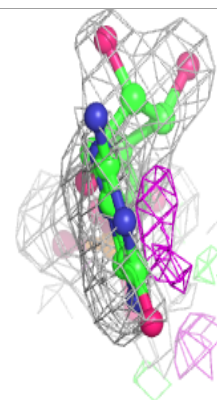
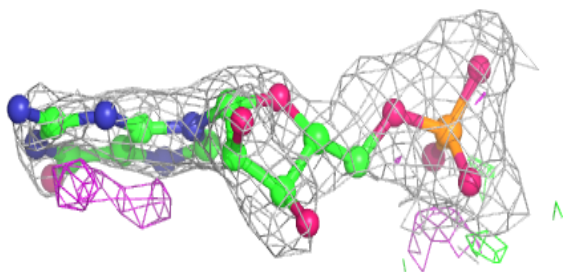
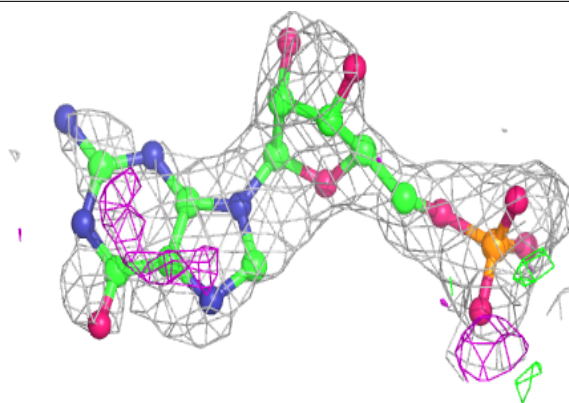
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GOL	A	404	6/6	0.84	0.20	46,53,54,55	0
5	GOL	B	404	6/6	0.86	0.20	56,60,60,60	0
3	5GP	B	402	24/24	0.88	0.24	39,66,77,77	0
3	5GP	C	402	24/24	0.89	0.27	39,61,63,63	0
3	5GP	D	402	24/24	0.89	0.23	41,70,70,71	0
3	5GP	E	402	24/24	0.92	0.20	31,63,70,72	0
3	5GP	A	402	24/24	0.93	0.17	34,52,57,57	0
4	SO4	A	403	5/5	0.94	0.17	54,55,56,57	0
4	SO4	E	403	5/5	0.97	0.14	48,48,50,52	0
4	SO4	C	403	5/5	0.97	0.13	52,52,53,54	0
4	SO4	B	403	5/5	0.98	0.12	42,44,47,49	0
4	SO4	D	403	5/5	0.99	0.13	36,37,41,42	0
2	MN	C	401	1/1	0.99	0.07	24,24,24,24	0
2	MN	A	401	1/1	0.99	0.10	25,25,25,25	0
2	MN	B	401	1/1	0.99	0.05	23,23,23,23	0
2	MN	D	401	1/1	0.99	0.07	30,30,30,30	0
2	MN	E	401	1/1	1.00	0.08	23,23,23,23	0

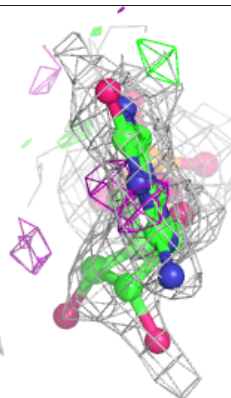
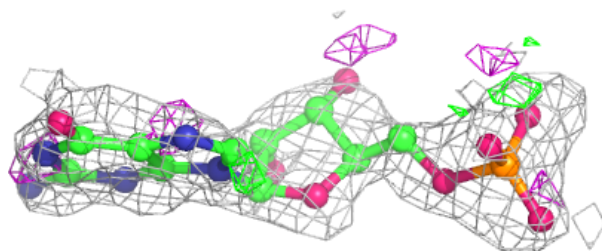
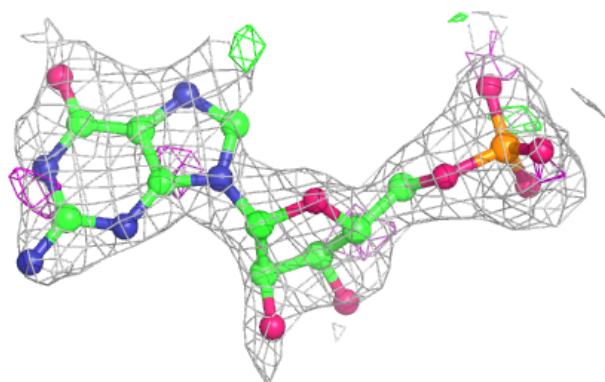
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around 5GP B 402:**

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

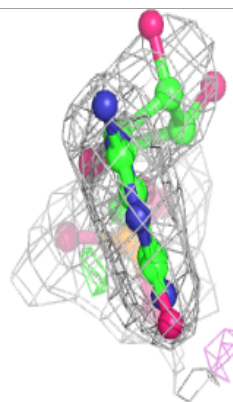
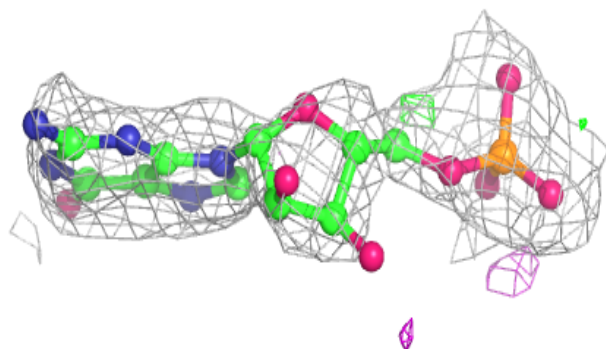
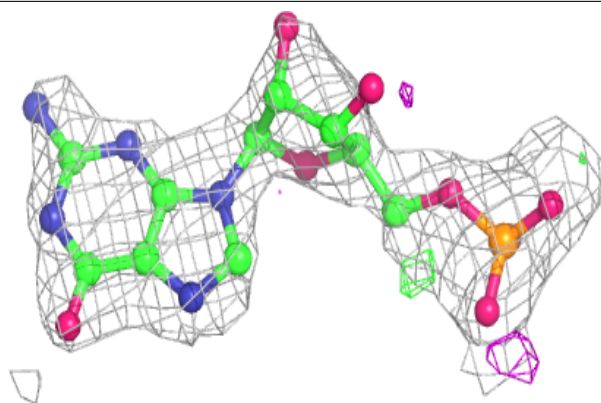
**Electron density around 5GP C 402:**

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

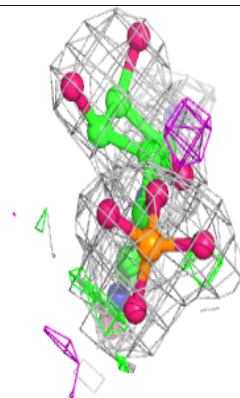
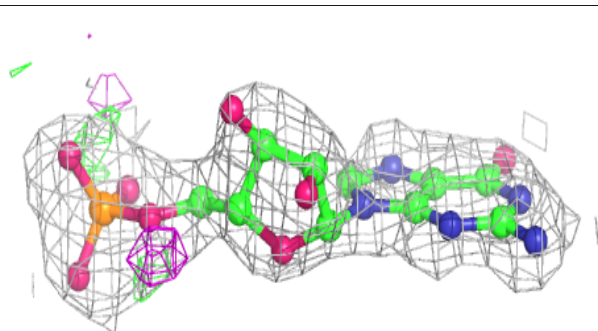
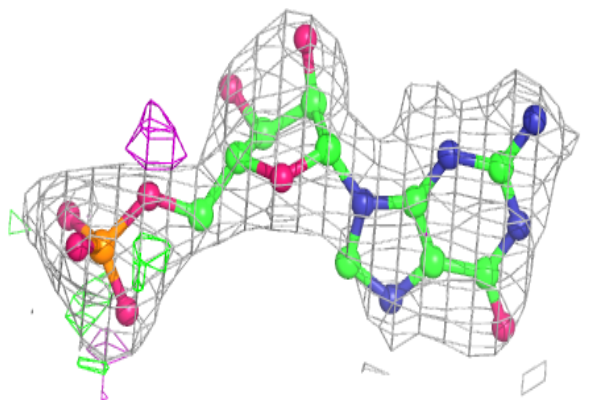


**Electron density around 5GP D 402:**

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

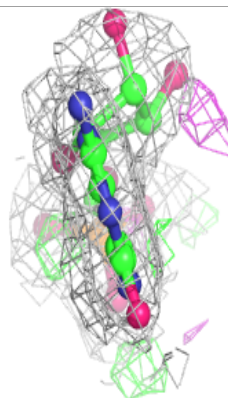
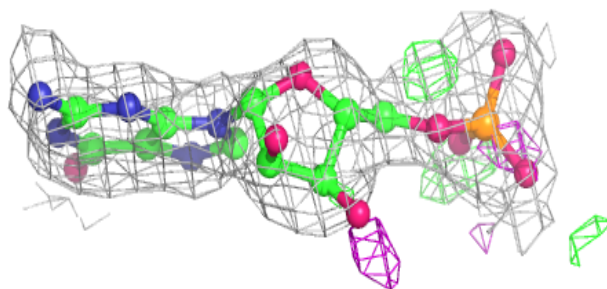
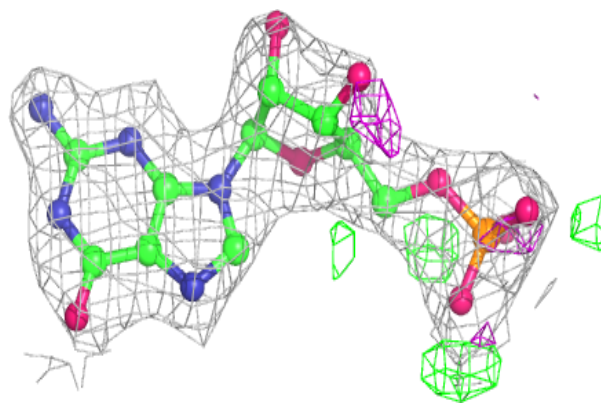
**Electron density around 5GP E 402:**

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



**Electron density around 5GP A 402:**

$2mF_o-DF_c$  (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.