



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2022 – 04:13 PM EDT

PDB ID : 5SB3  
Title : Tubulin-todalam-4-complex  
Authors : Muehlethaler, T.; Milanos, L.; Ortega, J.A.; Blum, T.B.; Gioia, D.; Prota, A.E.; Cavalli, A.; Steinmetz, M.O.  
Deposited on : 2021-07-08  
Resolution : 2.20 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 : 2.27  
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) : 2.27

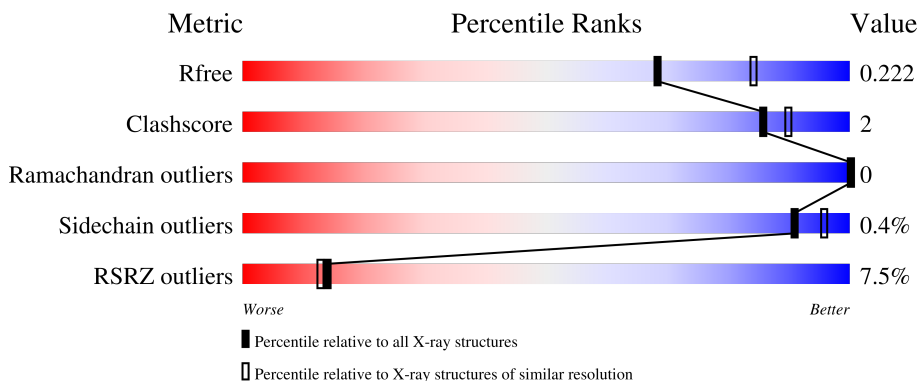
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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 of 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	451	 4% 91% 6% .
1	C	451	 % 92% 6% .
2	B	445	 2% 89% 7% 5%
2	D	445	 6% 90% 6% .
3	E	143	 5% 83% . 15%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	384	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a red segment on the left labeled '25%', a large green segment labeled '83%', a small yellow segment labeled '6%', and a grey segment on the far right labeled '12%'.</p>

## 2 Entry composition i

There are 12 unique types of molecules in this entry. The entry contains 18179 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 Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	437	Total	C	N	O	S	0	0	0
			3416	2163	581	650	22			
1	C	440	Total	C	N	O	S	0	0	0
			3437	2175	584	656	22			

- Molecule 2 is a protein called Tubulin beta-2B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	423	Total	C	N	O	S	0	1	0
			3345	2102	572	644	27			
2	D	425	Total	C	N	O	S	1	0	0
			3337	2095	569	646	27			

- Molecule 3 is a protein called Stathmin-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	121	Total	C	N	O	S	0	0	0
			1000	617	181	197	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	3	MET	-	initiating methionine	UNP P63043
E	4	ALA	-	expression tag	UNP P63043

- Molecule 4 is a protein called Tubulin-Tyrosine Ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	F	339	Total	C	N	O	S	0	0	0
			2773	1779	476	504	14			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	379	HIS	-	expression tag	UNP E1BQ43
F	380	HIS	-	expression tag	UNP E1BQ43
F	381	HIS	-	expression tag	UNP E1BQ43
F	382	HIS	-	expression tag	UNP E1BQ43
F	383	HIS	-	expression tag	UNP E1BQ43
F	384	HIS	-	expression tag	UNP E1BQ43

- Molecule 5 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
5	C	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
6	A	1	Total	Mg	0	0
			1	1		
6	B	1	Total	Mg	0	0
			1	1		
6	C	1	Total	Mg	0	0
			1	1		
6	D	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	1	Total	Mg	0	0
			1	1		

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Ca	0	0
			1	1		
7	B	1	Total	Ca	0	0
			1	1		
7	C	1	Total	Ca	0	0
			1	1		
7	E	1	Total	Ca	0	0
			1	1		

- Molecule 8 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>11</sub>P<sub>2</sub>).



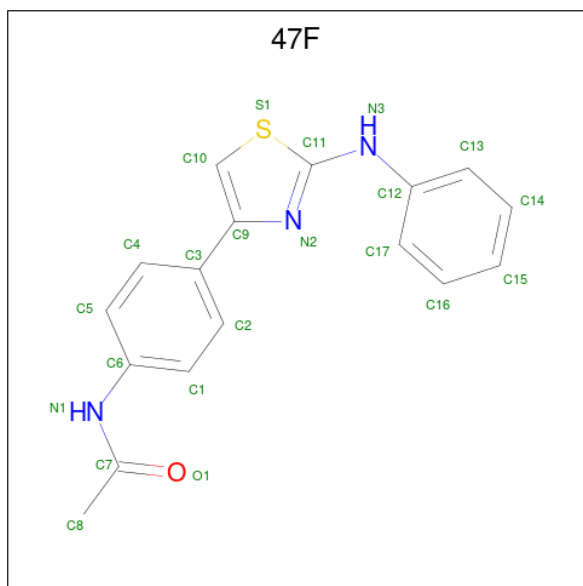
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
8	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
8	D	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 9 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



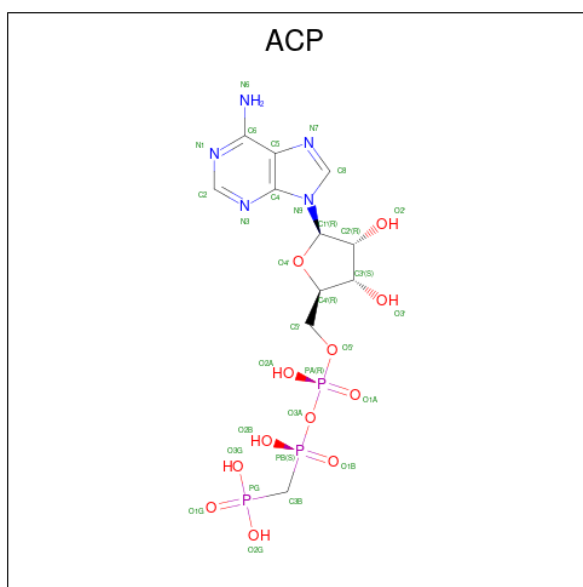
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
9	B	1	12	6	1	4	1	0	0

- Molecule 10 is N-[4-(2-anilino-1,3-thiazol-4-yl)phenyl]acetamide (three-letter code: 47F) (formula:  $C_{17}H_{15}N_3OS$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
10	C	1	22	17	3	1	1	0	0

- Molecule 11 is PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER (three-letter code: ACP) (formula:  $C_{11}H_{18}N_5O_{12}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
11	F	1	31	11	5	12	3	0	0

- Molecule 12 is water.

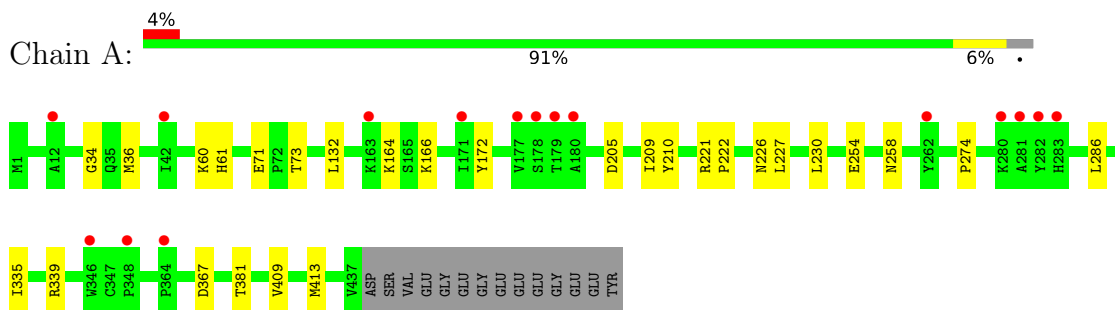
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	117	Total	O	0	0
			117	117		
12	B	135	Total	O	0	0
			135	135		
12	C	243	Total	O	0	0
			243	243		
12	D	91	Total	O	0	0
			91	91		
12	E	40	Total	O	0	0
			40	40		
12	F	51	Total	O	0	0
			51	51		



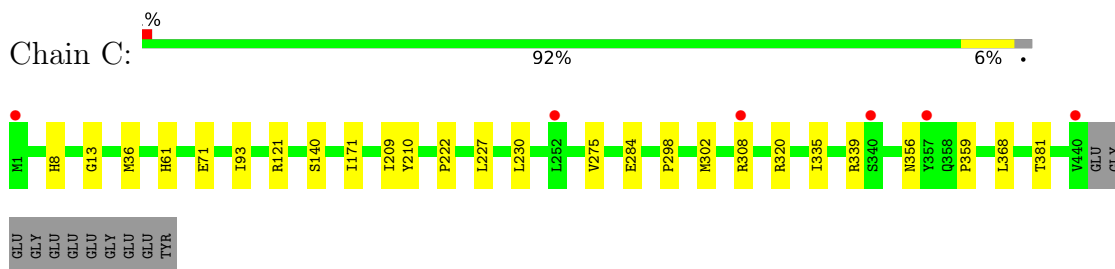
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

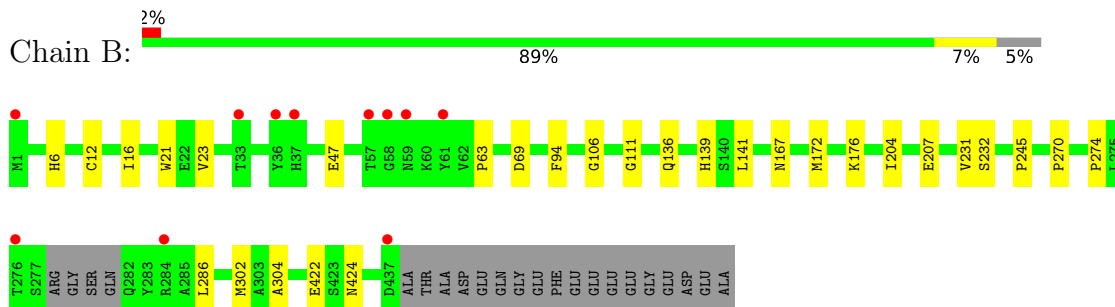
- Molecule 1: Tubulin alpha-1B chain



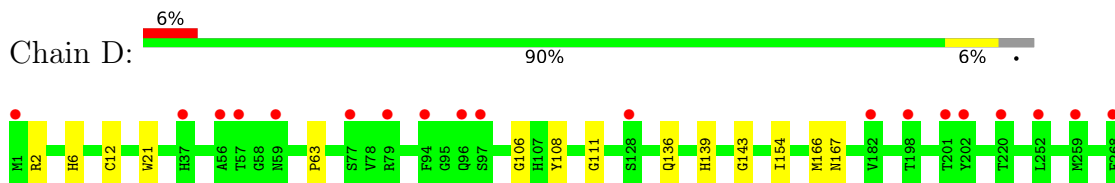
- Molecule 1: Tubulin alpha-1B chain

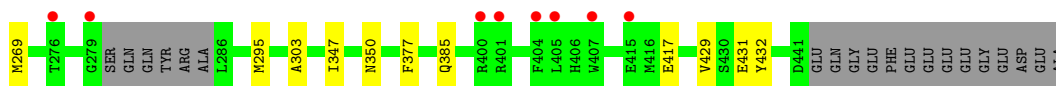


- Molecule 2: Tubulin beta-2B chain

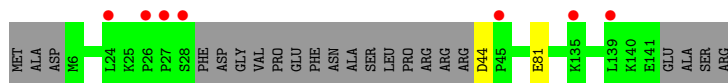
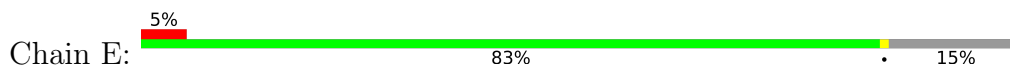


- Molecule 2: Tubulin beta-2B chain

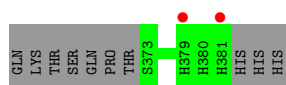
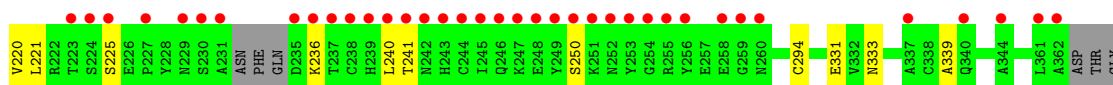
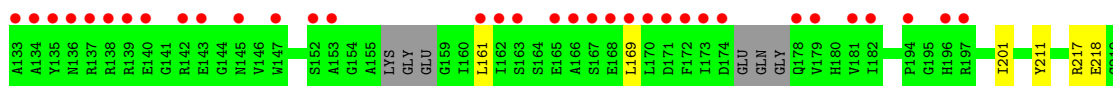
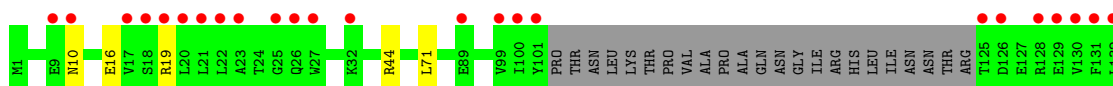
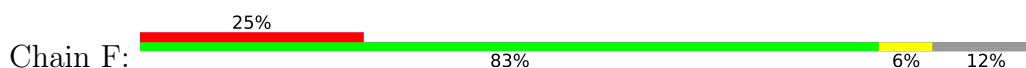




- Molecule 3: Stathmin-4



- Molecule 4: Tubulin-Tyrosine Ligase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.65Å 157.71Å 180.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.46 – 2.20 49.46 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.46-2.20) 99.9 (49.46-2.20)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.20Å)	Xtrriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.195 , 0.225 0.191 , 0.222	Depositor DCC
$R_{free}$ test set	7544 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.8	Xtrriage
Anisotropy	0.132	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 39.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	18179	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.37% 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: ACP, GTP, CA, GDP, MG, MES, 47F

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	0.25	0/3494	0.48	0/4743
1	C	0.27	0/3515	0.49	0/4772
2	B	0.26	0/3419	0.49	0/4629
2	D	0.25	0/3410	0.47	0/4618
3	E	0.24	0/1008	0.41	0/1337
4	F	0.24	0/2834	0.47	0/3824
All	All	0.25	0/17680	0.48	0/23923

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	3416	0	3330	14	0
1	C	3437	0	3348	13	0
2	B	3345	0	3227	17	0
2	D	3337	0	3217	15	0
3	E	1000	0	1018	2	0
4	F	2773	0	2744	12	0
5	A	32	0	12	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	32	0	12	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
8	B	28	0	12	1	0
8	D	28	0	12	3	0
9	B	12	0	12	1	0
10	C	22	0	0	0	0
11	F	31	0	14	2	0
12	A	117	0	0	1	0
12	B	135	0	0	3	0
12	C	243	0	0	1	0
12	D	91	0	0	4	0
12	E	40	0	0	2	0
12	F	51	0	0	0	0
All	All	18179	0	16958	75	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 2.

All (75) 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)
3:E:44:ASP:N	12:E:304:HOH:O	2.33	0.61
1:A:71:GLU:OE2	1:A:73:THR:OG1	2.19	0.60
4:F:331:GLU:OE2	11:F:401:ACP:O3G	2.19	0.59
2:D:295:MET:HE2	2:D:377:PHE:HB2	1.85	0.58
1:A:335:ILE:HG23	1:A:339:ARG:HG3	1.84	0.58
4:F:236:LYS:HB3	4:F:240:LEU:HD13	1.84	0.58
2:B:274:PRO:HB3	2:B:286:LEU:HD22	1.86	0.58
2:D:269:MET:HG3	2:D:303:ALA:HB3	1.87	0.56
2:B:16:ILE:HD13	2:B:231:VAL:HG11	1.87	0.56
2:D:6:HIS:CD2	2:D:21:TRP:HE1	2.24	0.55
4:F:10:ASN:HB2	4:F:44:ARG:HH22	1.70	0.55
2:B:47:GLU:HG2	2:B:245:PRO:HG3	1.89	0.54

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:270:PRO:HG2	2:B:302:MET:HB2	1.89	0.53
1:A:274:PRO:HB3	1:A:286:LEU:HD12	1.90	0.53
8:D:501:GDP:O3B	12:D:601:HOH:O	2.19	0.52
2:B:176:LYS:HD2	2:B:207:GLU:HG3	1.93	0.50
2:D:431:GLU:OE1	12:D:602:HOH:O	2.20	0.50
1:C:335:ILE:HG23	1:C:339:ARG:HG3	1.93	0.50
1:C:275:VAL:HG13	1:C:368:LEU:HD21	1.94	0.49
4:F:217:ARG:HG3	4:F:218:GLU:HG2	1.95	0.49
2:B:136:GLN:HA	2:B:167:ASN:O	2.13	0.49
2:B:23:VAL:HG21	2:B:232:SER:HB3	1.95	0.48
1:A:172:TYR:HB3	1:A:205:ASP:HA	1.95	0.48
4:F:225:SER:HG	4:F:250:SER:HG	1.61	0.48
3:E:81:GLU:OE1	12:E:301:HOH:O	2.20	0.48
2:D:12:CYS:HB2	8:D:501:GDP:C8	2.47	0.48
1:A:132:LEU:O	1:A:164:LYS:NZ	2.45	0.48
2:D:136:GLN:HA	2:D:167:ASN:O	2.14	0.48
2:D:347:ILE:HG22	2:D:350:ASN:HB3	1.96	0.48
4:F:16:GLU:OE2	4:F:19:ARG:NH2	2.47	0.48
9:B:504:MES:H51	9:B:504:MES:H81	1.59	0.47
1:C:210:TYR:CZ	1:C:222:PRO:HD2	2.49	0.47
2:D:432:TYR:OH	12:D:603:HOH:O	2.20	0.47
1:C:209:ILE:HG22	1:C:227:LEU:HD22	1.97	0.46
2:B:304:ALA:N	12:B:614:HOH:O	2.47	0.46
1:C:320:ARG:HA	1:C:356:ASN:O	2.15	0.46
2:D:108:TYR:OH	2:D:417:GLU:OE2	2.27	0.46
2:B:6:HIS:CD2	2:B:21:TRP:HE1	2.34	0.45
1:C:140:SER:HA	1:C:171:ILE:HB	1.99	0.45
4:F:331:GLU:OE2	4:F:333:ASN:ND2	2.44	0.45
1:C:298:PRO:HG2	1:C:308:ARG:NH2	2.32	0.45
2:D:2:ARG:NH1	12:D:606:HOH:O	2.34	0.45
1:A:209:ILE:HG23	1:A:230:LEU:HD23	1.97	0.45
1:C:359:PRO:HB2	12:C:650:HOH:O	2.16	0.45
4:F:220:VAL:HG11	4:F:339:ALA:HB2	1.97	0.45
1:A:166:LYS:NZ	12:A:605:HOH:O	2.49	0.44
1:A:226:ASN:ND2	1:A:367:ASP:OD2	2.50	0.44
1:A:36:MET:HB3	1:A:61:HIS:CE1	2.53	0.44
2:B:106:GLY:O	2:B:111:GLY:HA3	2.18	0.44
1:C:209:ILE:HD11	1:C:302:MET:HE1	2.00	0.44
1:A:210:TYR:CZ	1:A:222:PRO:HD2	2.53	0.44
4:F:71:LEU:HD11	4:F:294:CYS:HB3	1.99	0.44
2:D:154:ILE:HG23	2:D:166:MET:HG2	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:93:ILE:HD11	1:C:121:ARG:HG3	1.99	0.43
2:D:385:GLN:HB2	2:D:429:VAL:HG13	2.00	0.43
2:B:69:ASP:O	2:B:94:PHE:HA	2.19	0.42
4:F:201:ILE:HG12	4:F:221:LEU:HG	2.02	0.42
2:B:424:ASN:HB3	12:B:626:HOH:O	2.19	0.42
1:C:8:HIS:HB3	1:C:13:GLY:O	2.19	0.42
2:B:422:GLU:HG3	12:B:629:HOH:O	2.18	0.42
2:B:204:ILE:HD13	2:B:231:VAL:HG13	2.01	0.42
4:F:241:THR:OG1	11:F:401:ACP:O3'	2.35	0.42
1:C:209:ILE:HG23	1:C:230:LEU:HD23	2.00	0.41
4:F:161:LEU:HD23	4:F:169:LEU:HA	2.01	0.41
1:A:409:VAL:HA	1:A:413:MET:O	2.20	0.41
1:C:36:MET:HB3	1:C:61:HIS:CE1	2.55	0.41
2:B:12:CYS:HB2	8:B:501:GDP:C8	2.56	0.41
1:A:34:GLY:HA3	1:A:60:LYS:HG3	2.01	0.41
1:A:254:GLU:HG2	1:A:258:ASN:ND2	2.36	0.41
2:B:21:TRP:CZ3	2:B:63:PRO:HB3	2.55	0.41
2:D:21:TRP:CZ3	2:D:63:PRO:HB3	2.55	0.41
2:D:143:GLY:HA3	8:D:501:GDP:O3A	2.19	0.41
2:B:141:LEU:HD12	2:B:172:MET:SD	2.61	0.41
1:A:209:ILE:HG22	1:A:227:LEU:HD22	2.02	0.40
2:D:106:GLY:O	2:D:111:GLY:HA3	2.22	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	435/451 (96%)	425 (98%)	10 (2%)	0	100	100
1	C	438/451 (97%)	428 (98%)	10 (2%)	0	100	100
2	B	420/445 (94%)	413 (98%)	7 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	421/445 (95%)	414 (98%)	7 (2%)	0	100	100
3	E	117/143 (82%)	116 (99%)	1 (1%)	0	100	100
4	F	327/384 (85%)	319 (98%)	8 (2%)	0	100	100
All	All	2158/2319 (93%)	2115 (98%)	43 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	368/379 (97%)	366 (100%)	2 (0%)	88	94
1	C	371/379 (98%)	368 (99%)	3 (1%)	81	90
2	B	368/383 (96%)	367 (100%)	1 (0%)	92	97
2	D	367/383 (96%)	366 (100%)	1 (0%)	92	97
3	E	109/127 (86%)	109 (100%)	0	100	100
4	F	303/342 (89%)	302 (100%)	1 (0%)	92	97
All	All	1886/1993 (95%)	1878 (100%)	8 (0%)	91	96

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	221	ARG
1	A	381	THR
2	B	139	HIS
1	C	71	GLU
1	C	284	GLU
1	C	381	THR
2	D	139	HIS
4	F	211	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.



### 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 16 ligands modelled in this entry, 9 are monoatomic - leaving 7 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
8	GDP	B	501	6	24,30,30	1.19	2 (8%)	31,47,47	1.87	7 (22%)
5	GTP	C	501	6	26,34,34	0.99	1 (3%)	33,54,54	1.64	6 (18%)
5	GTP	A	501	6	26,34,34	0.94	1 (3%)	33,54,54	1.72	6 (18%)
10	47F	C	504	-	21,24,24	1.24	2 (9%)	26,32,32	1.25	2 (7%)
8	GDP	D	501	6	24,30,30	1.20	2 (8%)	31,47,47	1.93	7 (22%)
9	MES	B	504	-	12,12,12	2.23	1 (8%)	14,16,16	1.94	4 (28%)
11	ACP	F	401	6	27,33,33	1.38	5 (18%)	32,52,52	1.40	4 (12%)

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
8	GDP	B	501	6	-	3/12/32/32	0/3/3/3

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GTP	C	501	6	-	6/18/38/38	0/3/3/3
5	GTP	A	501	6	-	7/18/38/38	0/3/3/3
10	47F	C	504	-	-	2/10/12/12	0/3/3/3
8	GDP	D	501	6	-	4/12/32/32	0/3/3/3
9	MES	B	504	-	-	4/6/14/14	0/1/1/1
11	ACP	F	401	6	-	6/15/38/38	0/3/3/3

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	504	MES	C8-S	-7.48	1.66	1.77
8	D	501	GDP	C5-C6	4.21	1.48	1.41
8	B	501	GDP	C5-C6	4.07	1.48	1.41
10	C	504	47F	C11-N3	3.24	1.42	1.36
5	C	501	GTP	C6-N1	3.16	1.38	1.33
5	A	501	GTP	C6-N1	2.92	1.38	1.33
11	F	401	ACP	PG-O3G	2.90	1.61	1.54
11	F	401	ACP	PG-O2G	2.90	1.61	1.54
11	F	401	ACP	PB-O3A	2.70	1.61	1.58
11	F	401	ACP	C5-C4	2.48	1.47	1.40
8	D	501	GDP	C5-C4	2.46	1.47	1.40
8	B	501	GDP	C5-C4	2.40	1.47	1.40
11	F	401	ACP	PB-O2B	2.19	1.61	1.56
10	C	504	47F	C3-C9	-2.02	1.45	1.48

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	501	GTP	N3-C2-N1	-5.30	120.15	127.22
5	C	501	GTP	N3-C2-N1	-5.16	120.34	127.22
8	D	501	GDP	C2-N3-C4	4.75	120.78	115.36
8	B	501	GDP	C2-N3-C4	4.58	120.59	115.36
8	B	501	GDP	C4-C5-C6	-4.33	116.66	120.80
8	D	501	GDP	C2-N1-C6	4.19	122.58	115.93
8	D	501	GDP	C4-C5-C6	-4.11	116.87	120.80
5	A	501	GTP	C2-N3-C4	4.09	120.03	115.36
9	B	504	MES	O1S-S-C8	4.08	111.83	106.92
8	D	501	GDP	C5-C6-N1	-4.02	117.94	123.43
8	B	501	GDP	C2-N1-C6	3.99	122.27	115.93
8	B	501	GDP	C5-C6-N1	-3.82	118.21	123.43
9	B	504	MES	C5-N4-C3	3.80	117.38	108.83

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	501	GTP	C2-N3-C4	3.70	119.58	115.36
11	F	401	ACP	PB-O3A-PA	-3.61	121.10	132.56
8	D	501	GDP	N3-C2-N1	-3.37	122.72	127.22
8	B	501	GDP	N3-C2-N1	-3.32	122.80	127.22
11	F	401	ACP	C3'-C2'-C1'	3.28	105.91	100.98
11	F	401	ACP	N3-C2-N1	-3.20	123.68	128.68
5	A	501	GTP	C5-C6-N1	-3.19	119.07	123.43
5	C	501	GTP	C5-C6-N1	-3.07	119.23	123.43
10	C	504	47F	C10-C9-C3	-3.03	125.22	129.44
8	D	501	GDP	C4-C5-N7	-2.82	106.45	109.40
5	A	501	GTP	C2-N1-C6	2.78	120.35	115.93
8	D	501	GDP	PA-O3A-PB	-2.76	123.36	132.83
5	A	501	GTP	PA-O3A-PB	-2.73	123.46	132.83
10	C	504	47F	C12-N3-C11	-2.69	122.57	129.39
11	F	401	ACP	C4-C5-N7	-2.69	106.59	109.40
5	C	501	GTP	PA-O3A-PB	-2.69	123.60	132.83
5	C	501	GTP	PB-O3B-PG	-2.66	123.69	132.83
8	B	501	GDP	PA-O3A-PB	-2.65	123.72	132.83
5	C	501	GTP	C2-N1-C6	2.64	120.13	115.93
8	B	501	GDP	C4-C5-N7	-2.61	106.68	109.40
9	B	504	MES	C6-C5-N4	-2.29	106.63	110.10
9	B	504	MES	C7-N4-C5	2.24	116.96	111.23
5	A	501	GTP	PB-O3B-PG	-2.23	125.18	132.83

There are no chirality outliers.

All (32) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	501	GTP	C5'-O5'-PA-O1A
5	A	501	GTP	C5'-O5'-PA-O2A
5	C	501	GTP	PB-O3B-PG-O3G
5	C	501	GTP	C5'-O5'-PA-O1A
5	C	501	GTP	C5'-O5'-PA-O2A
8	B	501	GDP	C5'-O5'-PA-O1A
8	B	501	GDP	C5'-O5'-PA-O2A
8	D	501	GDP	PA-O3A-PB-O3B
8	D	501	GDP	C5'-O5'-PA-O1A
8	D	501	GDP	C5'-O5'-PA-O2A
9	B	504	MES	C8-C7-N4-C5
11	F	401	ACP	PB-C3B-PG-O1G
11	F	401	ACP	PB-C3B-PG-O2G
11	F	401	ACP	C5'-O5'-PA-O1A

*Continued on next page...*

*Continued from previous page...*

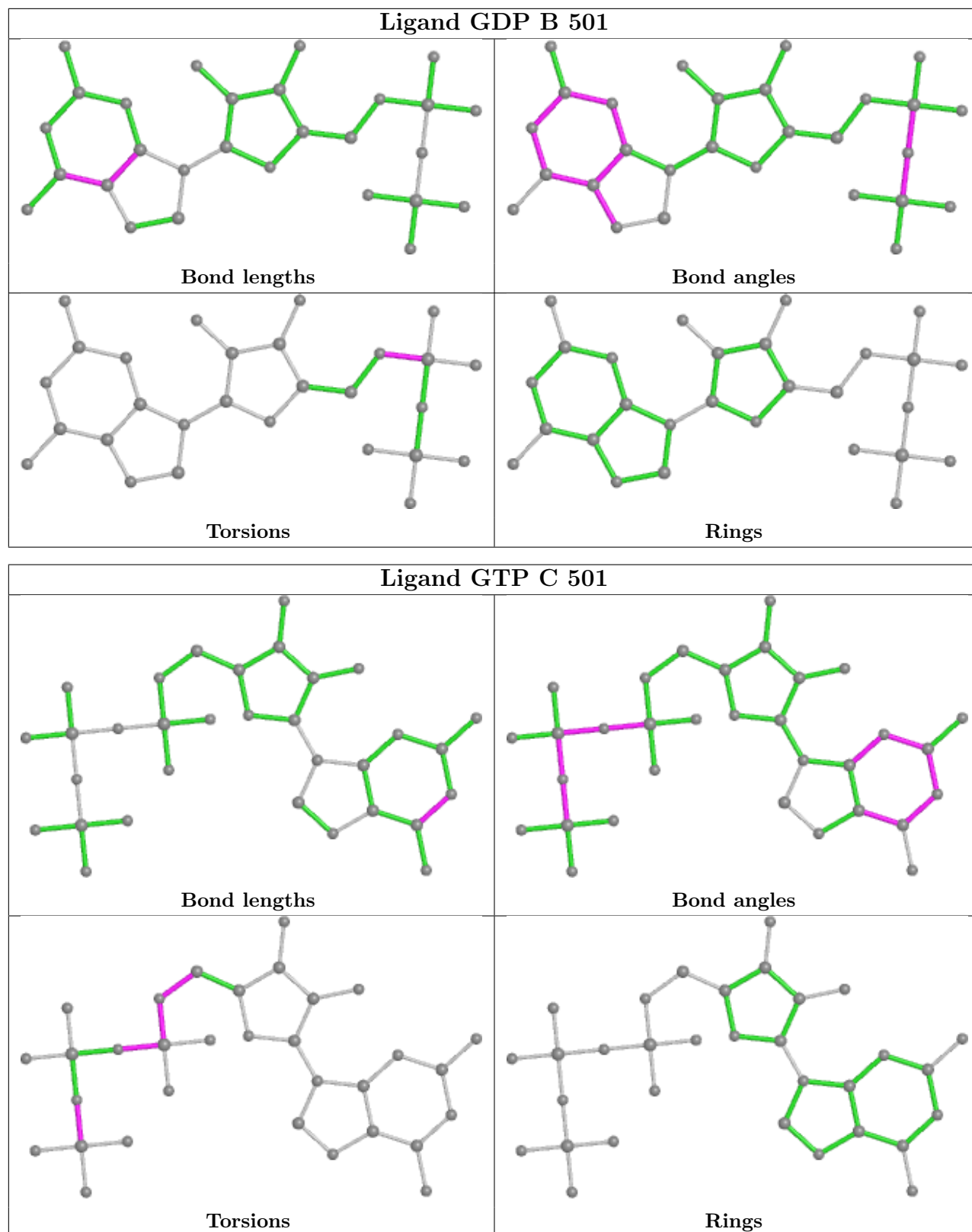
Mol	Chain	Res	Type	Atoms
11	F	401	ACP	C5'-O5'-PA-O3A
9	B	504	MES	C7-C8-S-O3S
5	A	501	GTP	PB-O3B-PG-O1G
5	A	501	GTP	C5'-O5'-PA-O3A
5	C	501	GTP	PB-O3A-PA-O2A
11	F	401	ACP	C5'-O5'-PA-O2A
9	B	504	MES	C7-C8-S-O1S
9	B	504	MES	C7-C8-S-O2S
11	F	401	ACP	PB-C3B-PG-O3G
10	C	504	47F	C13-C12-N3-C11
5	A	501	GTP	PB-O3A-PA-O2A
10	C	504	47F	C17-C12-N3-C11
5	C	501	GTP	C4'-C5'-O5'-PA
5	A	501	GTP	PB-O3B-PG-O2G
5	A	501	GTP	PB-O3B-PG-O3G
5	C	501	GTP	C5'-O5'-PA-O3A
8	B	501	GDP	C5'-O5'-PA-O3A
8	D	501	GDP	C5'-O5'-PA-O3A

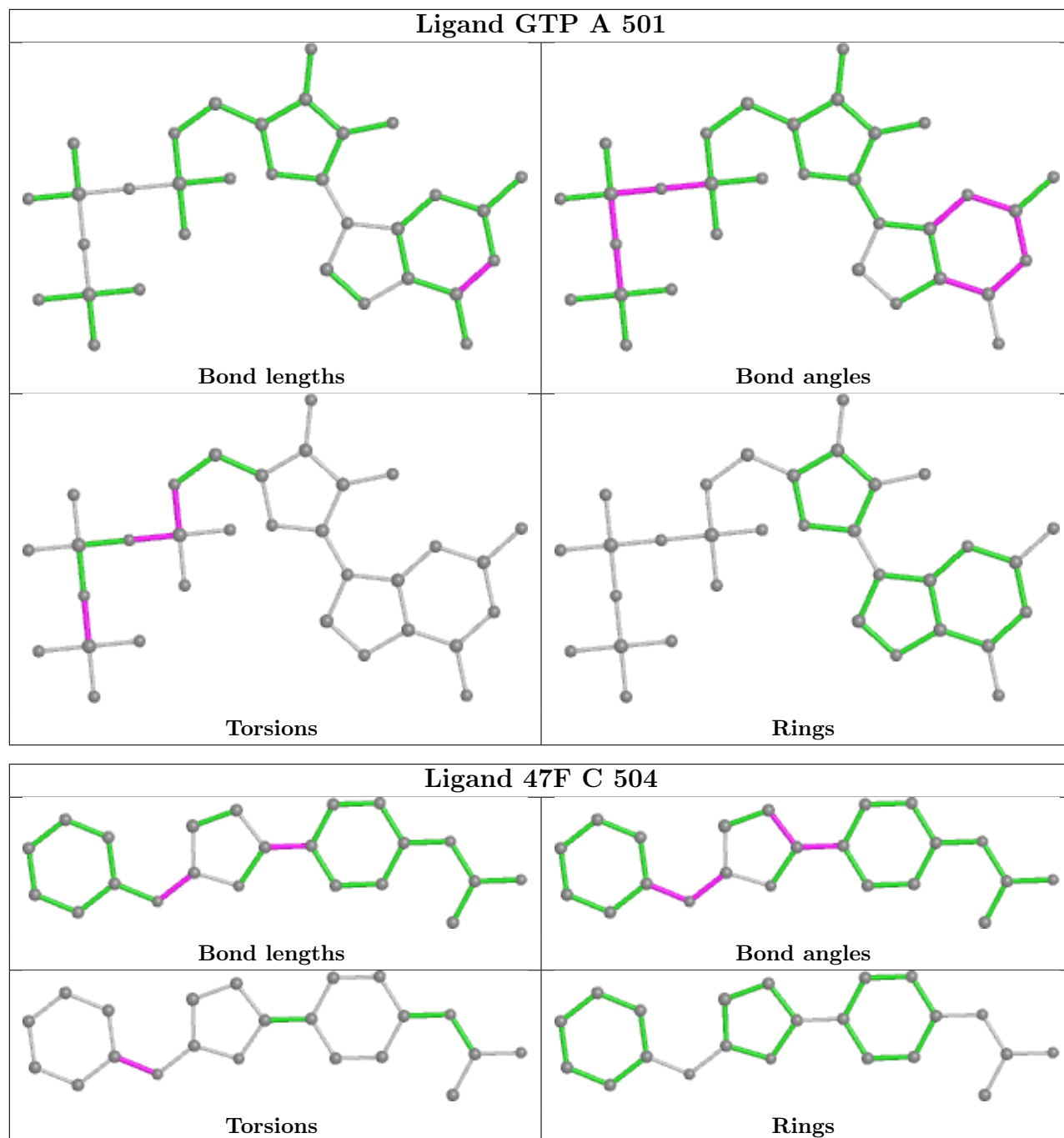
There are no ring outliers.

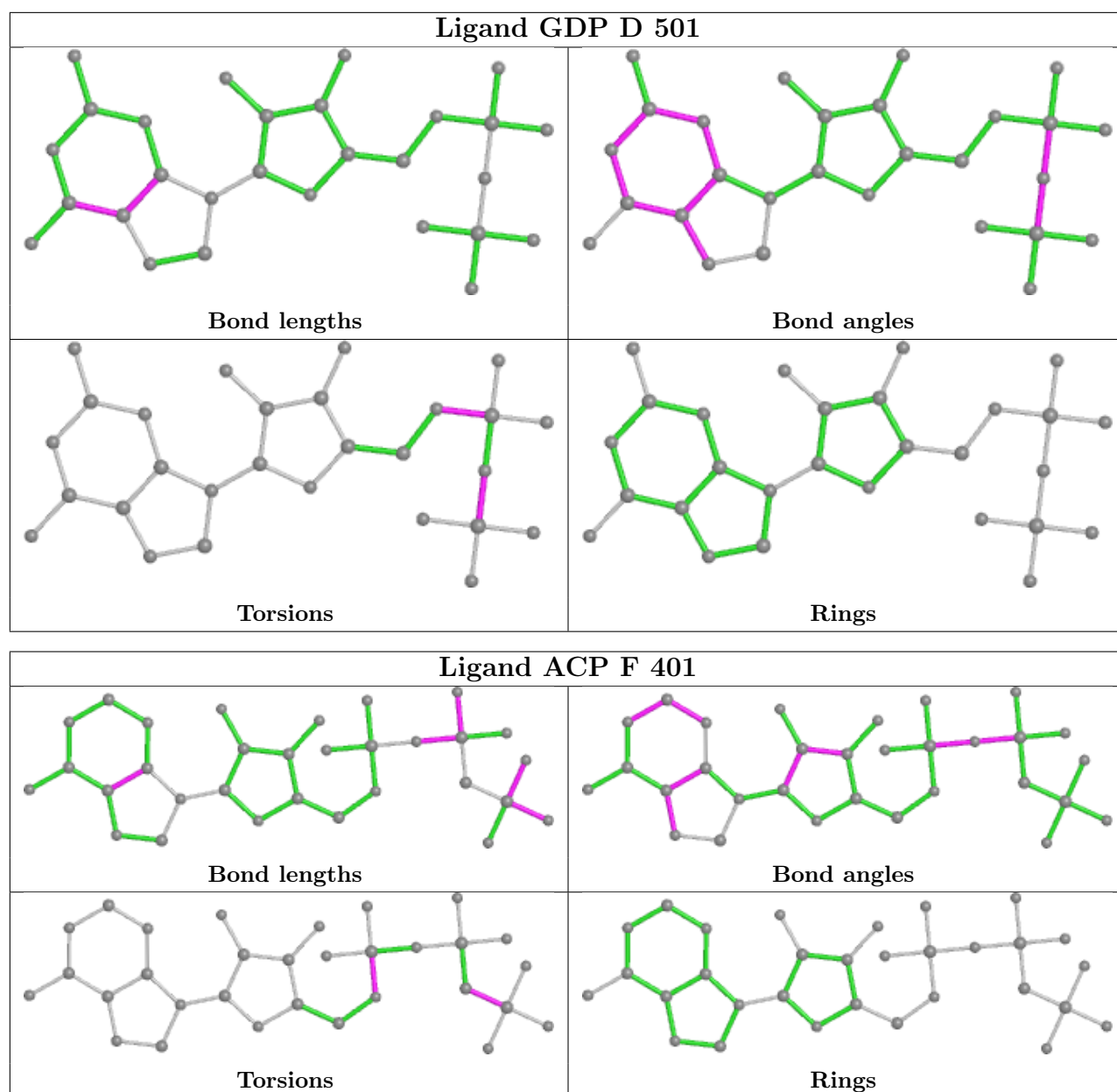
4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	501	GDP	1	0
8	D	501	GDP	3	0
9	B	504	MES	1	0
11	F	401	ACP	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 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	437/451 (96%)	0.13	16 (3%) 41 39	37, 54, 83, 133	0
1	C	440/451 (97%)	-0.07	6 (1%) 75 73	29, 42, 67, 106	0
2	B	423/445 (95%)	0.12	11 (2%) 56 53	31, 50, 83, 119	0
2	D	425/445 (95%)	0.31	27 (6%) 19 18	37, 60, 94, 125	2 (0%)
3	E	121/143 (84%)	0.28	7 (5%) 23 22	44, 64, 103, 111	0
4	F	339/384 (88%)	1.23	97 (28%) 0 0	49, 78, 144, 160	0
All	All	2185/2319 (94%)	0.30	164 (7%) 14 13	29, 56, 105, 160	2 (0%)

All (164) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	251	LYS	9.8
4	F	249	TYR	9.6
4	F	244	CYS	9.6
4	F	250	SER	8.6
4	F	130	VAL	8.1
4	F	173	ILE	7.8
4	F	253	TYR	7.5
4	F	259	GLY	7.1
4	F	166	ALA	6.9
4	F	131	PHE	6.8
4	F	240	LEU	6.8
4	F	132	LEU	6.5
2	B	1	MET	6.3
4	F	243	HIS	6.3
3	E	26	PRO	6.1
1	A	282	TYR	6.1
4	F	135	TYR	6.1
2	D	57	THR	6.1
4	F	254	GLY	5.9

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
4	F	170	LEU	5.8
4	F	161	LEU	5.7
4	F	248	GLU	5.6
4	F	133	ALA	5.6
4	F	100	ILE	5.3
4	F	252	ASN	5.3
4	F	142	ARG	5.2
4	F	134	ALA	5.2
4	F	129	GLU	5.0
4	F	225	SER	5.0
3	E	27	PRO	4.9
4	F	125	THR	4.8
2	B	276	THR	4.8
4	F	179	VAL	4.8
2	B	59	ASN	4.8
2	D	400	ARG	4.7
3	E	139	LEU	4.7
4	F	172	PHE	4.7
4	F	101	TYR	4.7
4	F	256	TYR	4.5
4	F	182	ILE	4.5
4	F	169	LEU	4.4
4	F	247	LYS	4.3
4	F	167	SER	4.3
4	F	20	LEU	4.2
4	F	361	LEU	4.1
4	F	231	ALA	4.1
4	F	362	ALA	4.1
2	B	57	THR	4.0
4	F	239	HIS	4.0
4	F	241	THR	4.0
4	F	181	VAL	4.0
4	F	163	SER	4.0
4	F	246	GLN	4.0
4	F	242	ASN	3.9
4	F	99	VAL	3.9
1	C	340	SER	3.8
4	F	236	LYS	3.8
1	A	281	ALA	3.7
1	A	283	HIS	3.7
2	B	37	HIS	3.7
4	F	255	ARG	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	42	ILE	3.7
4	F	143	GLU	3.6
2	D	401	ARG	3.6
2	D	415	GLU	3.5
4	F	230	SER	3.5
4	F	139	ARG	3.4
4	F	165	GLU	3.4
4	F	89	GLU	3.4
4	F	194	PRO	3.4
4	F	17	VAL	3.4
1	A	262	TYR	3.3
4	F	245	ILE	3.3
2	D	59	ASN	3.2
2	D	94	PHE	3.2
4	F	138	ARG	3.2
4	F	25	GLY	3.2
4	F	126	ASP	3.1
4	F	238	CYS	3.1
2	B	284	ARG	3.1
3	E	24	LEU	3.1
3	E	45	PRO	3.0
2	D	407	TRP	3.0
4	F	178	GLN	3.0
4	F	22	LEU	3.0
4	F	381	HIS	3.0
4	F	196	HIS	2.9
2	D	182	VAL	2.9
2	B	36	TYR	2.9
4	F	227	PRO	2.9
2	B	437	ASP	2.9
4	F	153	ALA	2.8
4	F	21	LEU	2.7
2	D	279	GLY	2.7
1	A	179	THR	2.7
2	D	404	PHE	2.7
4	F	147	TRP	2.7
4	F	258	GLU	2.7
2	D	276	THR	2.7
4	F	340	GLN	2.7
2	D	37	HIS	2.6
1	C	440	VAL	2.6
4	F	229	ASN	2.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	B	33	THR	2.6
2	D	201	THR	2.6
1	C	1	MET	2.6
2	D	220	THR	2.5
4	F	223	THR	2.5
4	F	337	ALA	2.5
4	F	18	SER	2.5
4	F	224	SER	2.5
3	E	135	LYS	2.5
2	D	56	ALA	2.5
1	A	177	VAL	2.5
4	F	27	TRP	2.5
4	F	171	ASP	2.5
2	D	202	TYR	2.4
4	F	26	GLN	2.4
1	C	357	TYR	2.4
3	E	28	SER	2.4
4	F	145	ASN	2.4
4	F	344	ALA	2.4
4	F	128	ARG	2.4
4	F	140	GLU	2.4
2	D	252	LEU	2.4
2	D	96	GLN	2.4
2	B	58	GLY	2.4
4	F	174	ASP	2.4
4	F	23	ALA	2.4
4	F	237	THR	2.4
4	F	197	ARG	2.3
4	F	235	ASP	2.3
4	F	168	GLU	2.3
1	A	348	PRO	2.3
1	C	252	LEU	2.3
4	F	137	ARG	2.3
4	F	9	GLU	2.3
2	D	97	SER	2.3
1	A	346	TRP	2.3
4	F	260	ASN	2.2
2	D	1	MET	2.2
4	F	136	ASN	2.2
1	A	180	ALA	2.2
1	A	178	SER	2.2
4	F	19	ARG	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	364	PRO	2.2
4	F	162	ILE	2.1
1	A	171	ILE	2.1
4	F	379	HIS	2.1
2	D	198	THR	2.1
2	D	128	SER	2.1
2	D	268	PHE	2.1
1	A	280	LYS	2.1
4	F	32	LYS	2.1
1	C	308	ARG	2.1
2	D	405	LEU	2.1
4	F	10	ASN	2.1
2	B	61	TYR	2.1
2	D	79	ARG	2.0
2	D	259	MET	2.0
1	A	12	ALA	2.0
2	D	77	SER	2.0
4	F	152	SER	2.0
1	A	163	LYS	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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.

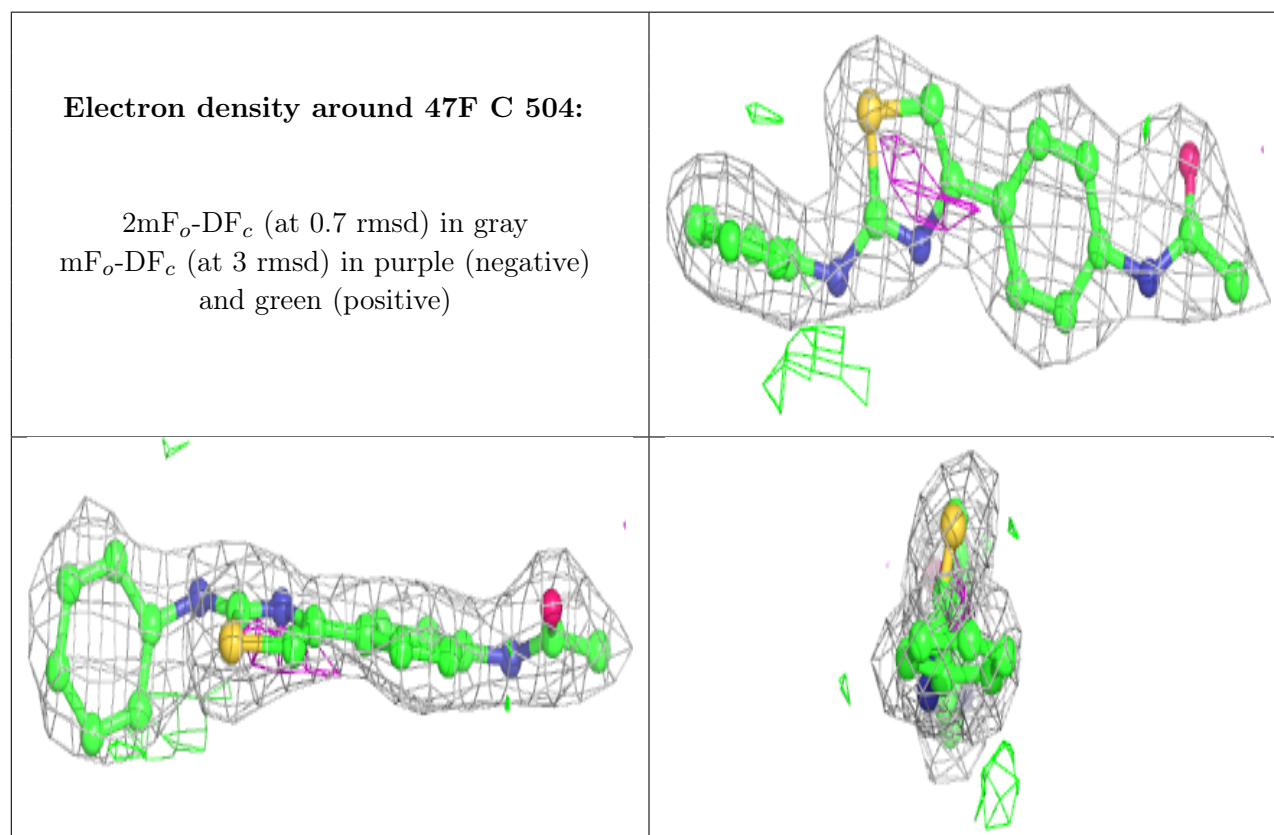
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	MG	D	502	1/1	0.79	0.20	70,70,70,70	0
10	47F	C	504	22/22	0.89	0.29	35,45,51,54	22
11	ACP	F	401	31/31	0.89	0.16	90,105,112,114	0
7	CA	E	201	1/1	0.91	0.15	105,105,105,105	0

*Continued on next page...*

Continued from previous page...

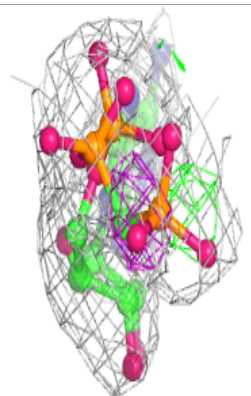
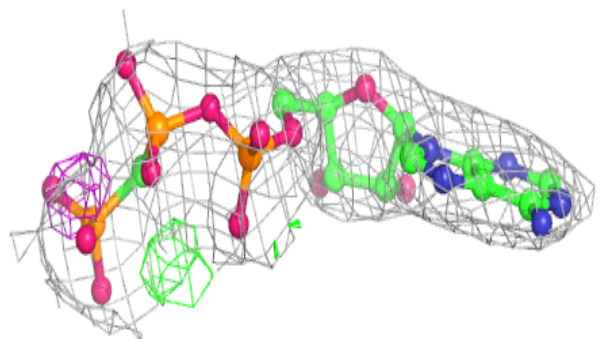
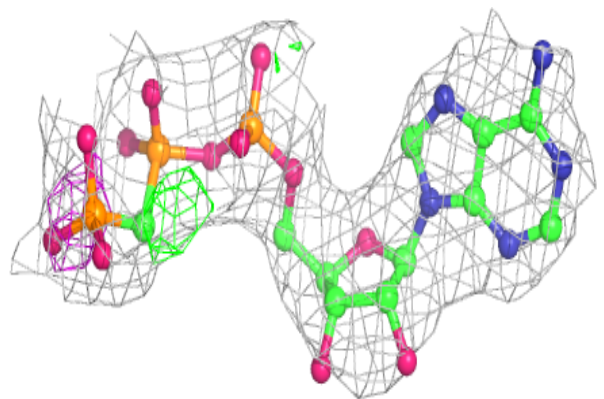
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	CA	B	503	1/1	0.95	0.05	94,94,94,94	0
6	MG	A	502	1/1	0.95	0.15	39,39,39,39	0
6	MG	C	502	1/1	0.96	0.14	32,32,32,32	0
8	GDP	D	501	28/28	0.96	0.12	55,59,67,71	0
5	GTP	A	501	32/32	0.97	0.20	34,38,41,45	0
7	CA	C	503	1/1	0.97	0.04	60,60,60,60	0
6	MG	F	402	1/1	0.97	0.06	90,90,90,90	0
6	MG	B	502	1/1	0.98	0.25	30,30,30,30	0
9	MES	B	504	12/12	0.98	0.11	42,48,55,57	0
5	GTP	C	501	32/32	0.98	0.16	29,33,37,40	0
7	CA	A	503	1/1	0.98	0.04	70,70,70,70	0
8	GDP	B	501	28/28	0.99	0.19	32,37,42,44	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.



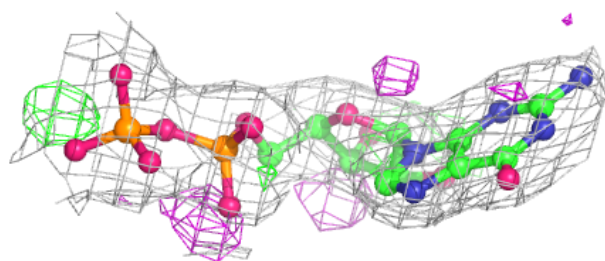
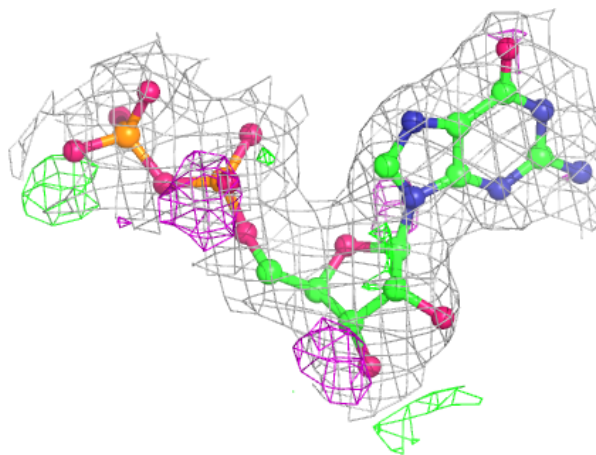
**Electron density around ACP F 401:**

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

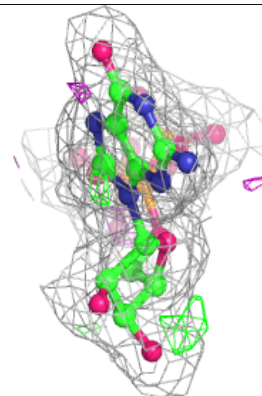
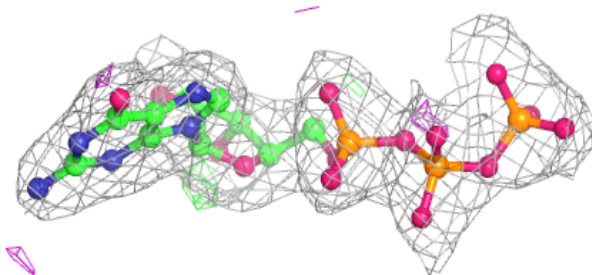
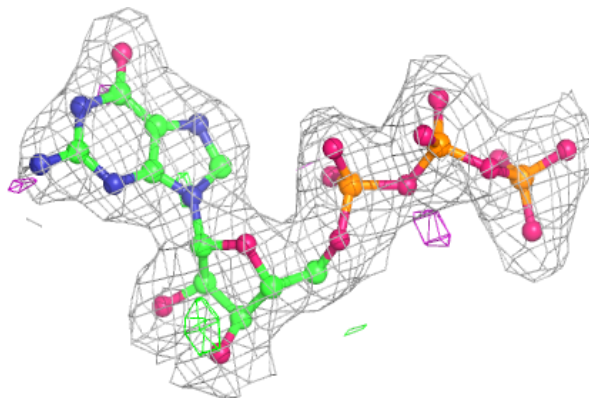


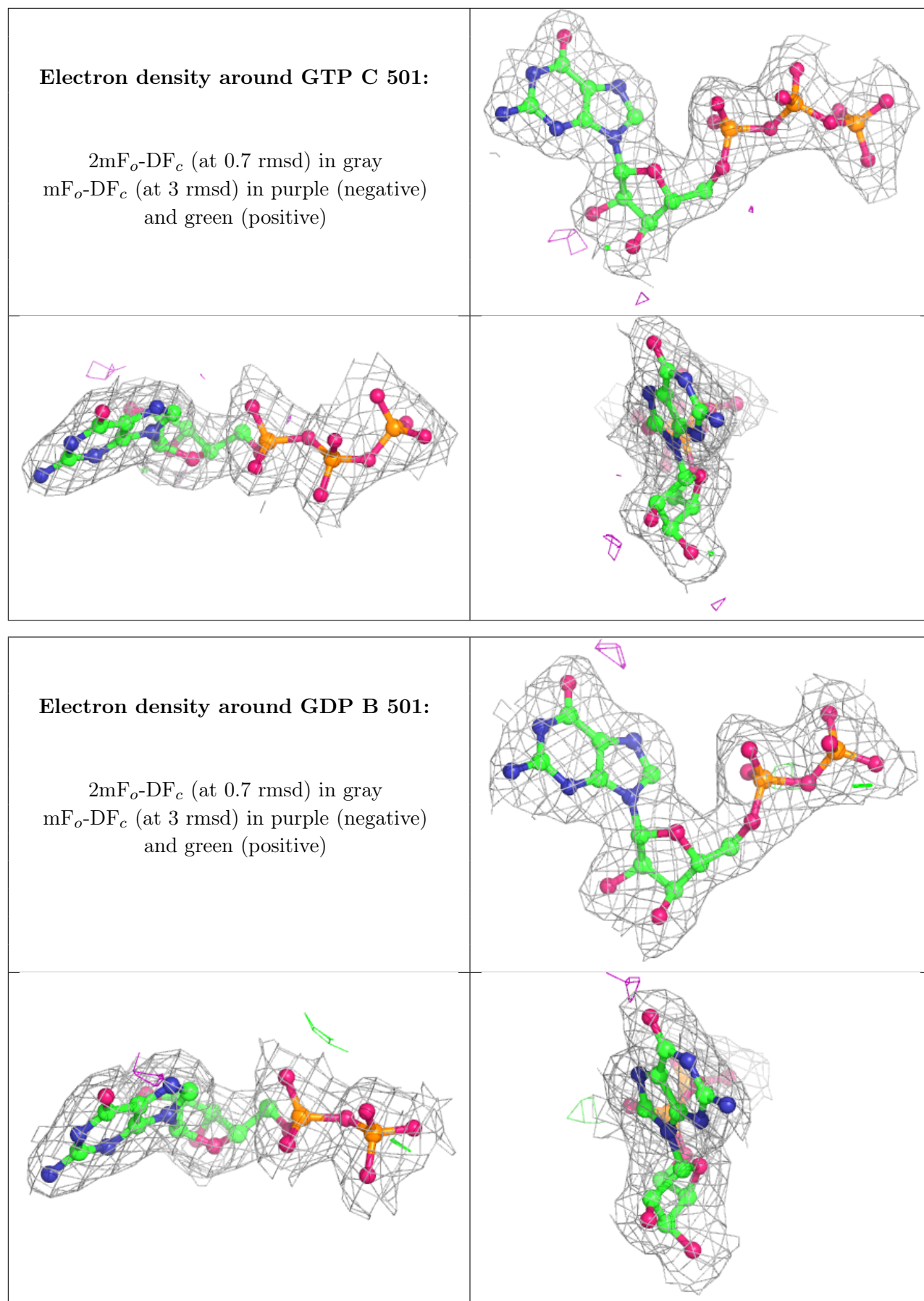
**Electron density around GDP D 501:**

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

**Electron density around GTP A 501:**

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







## 6.5 Other polymers [i](#)

There are no such residues in this entry.