



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2023 – 12:31 PM EDT

PDB ID : 8SCQ  
Title : Bst DNA polymerase I Large Fragment wildtype D598A with 3'-amino primer, dGTP, and calcium time-resolved 2h  
Authors : Fang, Z.; Lelyveld, V.S.; Szostak, J.W.  
Deposited on : 2023-04-05  
Resolution : 2.18 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

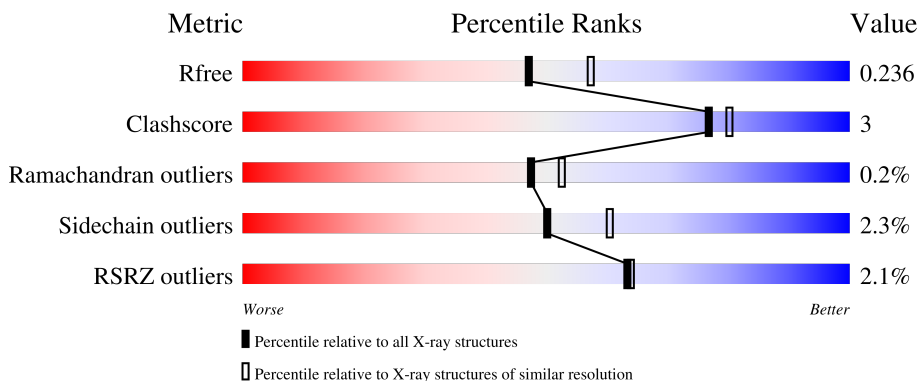
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.18 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.



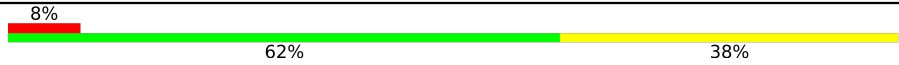
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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	579	
1	D	579	
2	B	10	
2	E	10	
3	C	13	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	13	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a red segment on the left labeled '8%', a green segment in the middle labeled '62%', and a yellow segment on the right labeled '38%'. The segments are stacked horizontally to total 100%.</p>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 10646 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 DNA polymerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	579	Total	C	N	O	S	0	3	0
			4659	2966	808	867	18			
1	D	579	Total	C	N	O	S	0	0	0
			4645	2955	807	866	17			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	713	VAL	PRO	variant	UNP D9N168
D	713	VAL	PRO	variant	UNP D9N168

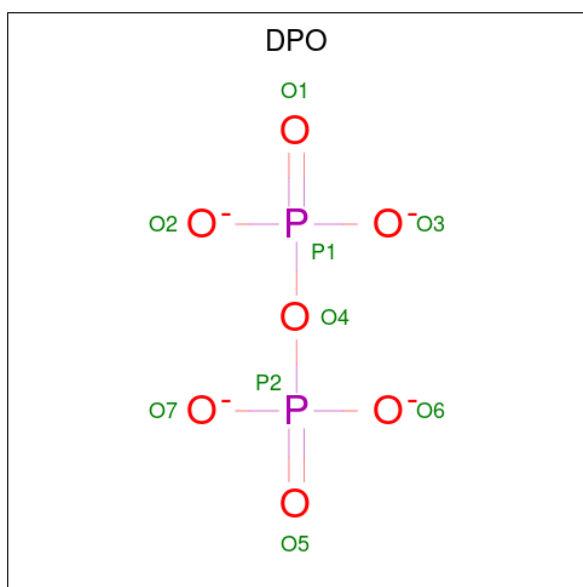
- Molecule 2 is a DNA chain called DNA primer/product.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	10	Total	C	N	O	P	0	2	0
			223	106	46	61	10			
2	E	10	Total	C	N	O	P	0	2	0
			223	106	46	61	10			

- Molecule 3 is a DNA chain called DNA template.

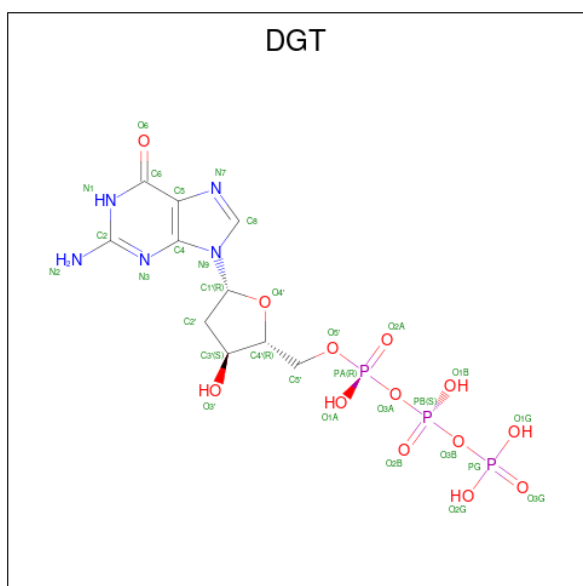
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	13	Total	C	N	O	P	0	0	1
			246	116	46	72	12			
3	F	13	Total	C	N	O	P	0	0	1
			246	116	46	72	12			

- Molecule 4 is DIPHOSPHATE (three-letter code: DPO) (formula: O<sub>7</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 9 7 2	0	1
4	D	1	Total O P 9 7 2	0	1

- Molecule 5 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O P 31 10 5 13 3	0	1

*Continued on next page...*

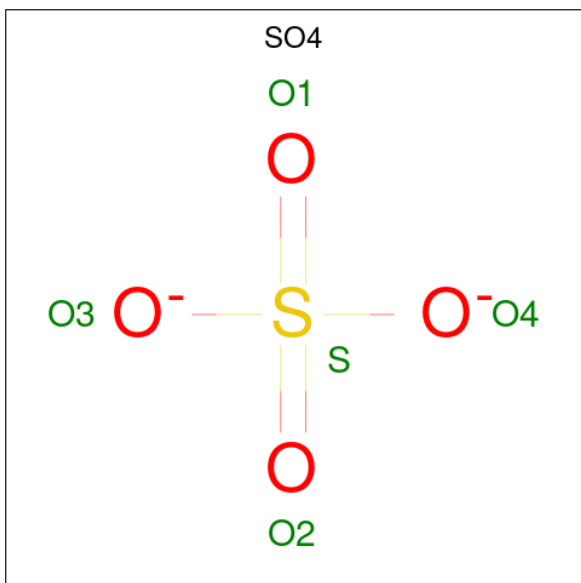
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
5	D	1	31	10	5	13	3	0	1

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
6	A	1	1	1	0	0
6	D	1	1	1	0	0

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

Continued on next page...

*Continued from previous page...*

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

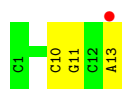
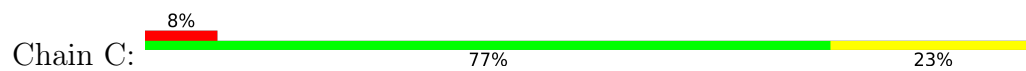
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	121	Total	O	0	0
			121	121		
8	B	10	Total	O	0	0
			10	10		
8	C	15	Total	O	0	0
			15	15		
8	D	72	Total	O	0	0
			72	72		
8	E	11	Total	O	0	0
			11	11		
8	F	13	Total	O	0	0
			13	13		





## ● Molecule 3: DNA template



## ● Molecule 3: DNA template



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.66Å 108.55Å 148.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.83 – 2.18 46.83 – 2.18	Depositor EDS
% Data completeness (in resolution range)	98.2 (46.83-2.18) 98.2 (46.83-2.18)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.76 (at 2.18Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.188 , 0.236 0.188 , 0.236	Depositor DCC
$R_{free}$ test set	3953 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.9	Xtrriage
Anisotropy	0.040	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10646	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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: DPO, C42, CA, DGT, SO4

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	0.42	0/4752	0.55	0/6420
1	D	0.38	0/4729	0.52	0/6391
2	B	0.90	0/207	0.94	0/316
2	E	0.94	0/207	1.03	0/316
3	C	0.80	0/275	0.91	0/423
3	F	0.99	0/275	0.94	0/423
All	All	0.47	0/10445	0.59	0/14289

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	4659	0	4732	21	0
1	D	4645	0	4705	36	0
2	B	223	0	122	1	0
2	E	223	0	123	2	0
3	C	246	0	135	3	0
3	F	246	0	135	10	0
4	A	9	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	9	0	0	0	0
5	A	31	0	9	1	0
5	D	31	0	10	1	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	A	40	0	0	0	0
7	B	5	0	0	0	0
7	D	30	0	0	0	0
7	E	5	0	0	0	0
8	A	121	0	0	2	0
8	B	10	0	0	0	0
8	C	15	0	0	0	0
8	D	72	0	0	0	0
8	E	11	0	0	1	0
8	F	13	0	0	0	0
All	All	10646	0	9971	61	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 (61) 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)
2:E:3:DG:N7	8:E:201:HOH:O	2.15	0.80
1:D:532:LYS:HD3	3:F:12:DC:H5''	1.77	0.66
1:A:298:LYS:NZ	8:A:1002:HOH:O	2.27	0.66
1:D:532:LYS:HE2	3:F:13:DA:H5'	1.78	0.65
1:D:677:ARG:HB3	1:D:679:LEU:HD22	1.81	0.62
1:D:532:LYS:HE3	3:F:13:DA:N3	2.15	0.61
1:A:323:VAL:O	1:A:435:ARG:NH2	2.32	0.61
1:A:648:LEU:HD12	1:A:841:MET:HG3	1.85	0.57
1:A:500:LEU:HD11	1:A:591:LEU:HD23	1.85	0.57
1:D:532:LYS:N	1:D:532:LYS:HD2	2.21	0.55
1:D:677:ARG:HD3	1:D:679:LEU:HD21	1.88	0.54
1:D:500:LEU:HD22	1:D:588:ILE:HD13	1.90	0.54
1:D:309:GLU:H	1:D:309:GLU:CD	2.11	0.53
3:C:10:DC:H2''	3:C:11:DG:C8	2.44	0.52
1:D:300:ALA:O	1:D:343:ARG:NH1	2.42	0.52
1:D:836:ALA:HB3	1:D:841:MET:HE2	1.92	0.52
1:A:836:ALA:HB3	1:A:841:MET:CE	2.40	0.51
1:D:718:ASP:OD1	1:D:719:TYR:N	2.43	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:532:LYS:HG3	3:F:13:DA:C4	2.48	0.49
1:A:658:GLU:CD	5:A:902[B]:DGT:H2'	2.33	0.49
1:D:334:GLY:HA2	1:D:348:PRO:HD3	1.94	0.49
1:D:510:GLN:O	1:D:514:VAL:HG12	2.13	0.48
1:D:414:ALA:HB1	1:D:419:TYR:HB3	1.95	0.48
1:D:672:MET:HE2	1:D:859:ARG:HB2	1.95	0.48
1:D:698:THR:OG1	1:D:701:MET:HG3	2.14	0.47
1:A:532:LYS:HG2	3:C:13:DA:C8	2.49	0.47
3:F:13:DA:N3	3:F:13:DA:H2'	2.29	0.47
3:F:4:DG:H2'	3:F:5:DC:C6	2.50	0.47
1:A:537:ILE:HA	1:A:541:LYS:HB3	1.97	0.45
1:A:718:ASP:OD1	1:A:719:TYR:N	2.50	0.45
1:A:471:ASP:N	1:A:471:ASP:OD1	2.49	0.45
1:D:505:LYS:HB3	1:D:505:LYS:HE2	1.67	0.45
1:A:532:LYS:HE2	3:C:13:DA:C8	2.51	0.45
1:D:683:THR:HG21	1:D:694:GLU:HG3	1.99	0.45
1:A:695:ASP:OD1	1:A:695:ASP:N	2.47	0.45
1:D:776:ILE:HG23	1:D:777:THR:HG23	1.98	0.44
1:D:619:THR:HG22	3:F:7:DG:H3'	1.99	0.44
1:D:767:LEU:HD12	1:D:802:ASP:HB3	2.00	0.44
1:A:408:ASP:OD1	1:A:408:ASP:N	2.51	0.44
1:D:534:LEU:HD11	1:D:574:ILE:HD13	1.99	0.44
1:A:712:ILE:HA	1:A:716:ILE:HG22	2.00	0.43
1:D:452:ALA:O	1:D:456:GLU:HG2	2.18	0.43
1:A:830:ASP:OD1	2:B:9[B]:C42:N	2.51	0.43
1:D:844:LEU:O	1:D:848:VAL:HG13	2.18	0.43
1:D:839:GLU:OE2	1:D:839:GLU:N	2.49	0.43
1:A:848:VAL:HB	1:A:849:PRO:HD3	2.01	0.43
1:D:669:ASP:O	1:D:673:GLU:HG2	2.19	0.42
1:D:677:ARG:HB3	1:D:679:LEU:CD2	2.48	0.42
1:D:618:SER:O	3:F:7:DG:H4'	2.20	0.42
1:D:830:ASP:OD1	2:E:9[B]:C42:N	2.50	0.41
1:D:843:ARG:HD2	1:D:843:ARG:HA	1.64	0.41
1:A:776:ILE:HG23	1:A:777:THR:HG23	2.02	0.41
1:D:532:LYS:HE3	3:F:13:DA:H2'	2.01	0.41
1:D:532:LYS:CD	3:F:12:DC:H5''	2.48	0.41
1:A:559:ASP:OD1	1:A:559:ASP:N	2.52	0.41
1:A:876:LYS:NZ	8:A:1008:HOH:O	2.48	0.41
1:A:346:LEU:HB3	1:A:351:ALA:HB2	2.03	0.41
1:D:408:ASP:OD1	1:D:408:ASP:N	2.54	0.41
1:D:591:LEU:HD21	1:D:640:PHE:CZ	2.56	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:764:THR:HA	1:A:769:ARG:O	2.21	0.40
1:D:658:GLU:CD	5:D:902[B]:DGT:H2'	2.42	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	580/579 (100%)	564 (97%)	15 (3%)	1 (0%)	47 52
1	D	577/579 (100%)	565 (98%)	11 (2%)	1 (0%)	47 52
All	All	1157/1158 (100%)	1129 (98%)	26 (2%)	2 (0%)	47 52

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	628	ILE
1	A	628	ILE

### 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	498/495 (101%)	489 (98%)	9 (2%)	59 70
1	D	495/495 (100%)	481 (97%)	14 (3%)	43 53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	993/990 (100%)	970 (98%)	23 (2%)	50 60

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	500	LEU
1	A	584	GLN
1	A	634	ARG
1	A	653	ASP
1	A	655	SER
1	A	695	ASP
1	A	779	ARG
1	A	804	ILE
1	A	846	ARG
1	D	325	GLU
1	D	350	THR
1	D	505	LYS
1	D	507	LEU
1	D	552	THR
1	D	572	GLU
1	D	631	GLU
1	D	679	LEU
1	D	730	LYS
1	D	748	ARG
1	D	776	ILE
1	D	779	ARG
1	D	846	ARG
1	D	857	THR

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](#)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	C42	E	9[B]	3,2	18,20,21	3.28	8 (44%)	18,28,31	0.81	0
2	C42	B	9[A]	3,2	18,20,21	3.38	8 (44%)	18,28,31	1.04	1 (5%)
2	C42	E	9[A]	3,2	18,20,21	3.40	8 (44%)	18,28,31	1.06	1 (5%)
2	C42	B	9[B]	3,2	18,20,21	3.14	8 (44%)	18,28,31	1.21	2 (11%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	C42	E	9[B]	3,2	-	0/7/21/22	0/2/2/2
2	C42	B	9[A]	3,2	-	0/7/21/22	0/2/2/2
2	C42	E	9[A]	3,2	-	0/7/21/22	0/2/2/2
2	C42	B	9[B]	3,2	-	0/7/21/22	0/2/2/2

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	9[A]	C42	O4'-C1'	7.08	1.58	1.42
2	E	9[A]	C42	O4'-C1'	7.07	1.58	1.42
2	B	9[B]	C42	O4'-C1'	7.02	1.58	1.42
2	E	9[B]	C42	O4'-C1'	6.65	1.57	1.42
2	E	9[A]	C42	O4'-C4'	-6.60	1.30	1.45
2	B	9[A]	C42	O4'-C4'	-6.53	1.30	1.45
2	E	9[B]	C42	O4'-C4'	-6.51	1.30	1.45
2	B	9[B]	C42	O4'-C4'	-6.29	1.30	1.45
2	E	9[A]	C42	C1'-N1	-5.35	1.34	1.48
2	B	9[A]	C42	C1'-N1	-5.29	1.34	1.48
2	E	9[B]	C42	C4-N4	5.20	1.46	1.33
2	E	9[A]	C42	C4-N4	4.86	1.45	1.33
2	B	9[A]	C42	C4-N4	4.82	1.45	1.33
2	E	9[B]	C42	C1'-N1	-4.68	1.36	1.48
2	E	9[B]	C42	C4'-C3'	4.63	1.63	1.52

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	9[A]	C42	C4'-C3'	4.55	1.62	1.52
2	B	9[B]	C42	C4-N4	4.54	1.44	1.33
2	E	9[A]	C42	C4'-C3'	4.51	1.62	1.52
2	B	9[B]	C42	C1'-N1	-4.45	1.36	1.48
2	B	9[B]	C42	C4'-C3'	4.09	1.61	1.52
2	E	9[A]	C42	C2-N1	-3.38	1.32	1.40
2	B	9[A]	C42	C2-N1	-3.32	1.32	1.40
2	E	9[B]	C42	C2-N1	-3.18	1.33	1.40
2	E	9[A]	C42	O2-C2	-2.77	1.18	1.23
2	B	9[A]	C42	O2-C2	-2.70	1.18	1.23
2	E	9[B]	C42	O2-C2	-2.61	1.18	1.23
2	B	9[B]	C42	C2-N1	-2.40	1.34	1.40
2	E	9[A]	C42	C6-N1	-2.38	1.32	1.38
2	B	9[A]	C42	C6-N1	-2.34	1.32	1.38
2	B	9[B]	C42	C6-C5	2.17	1.40	1.35
2	E	9[B]	C42	C6-C5	2.16	1.40	1.35
2	B	9[B]	C42	C2-N3	-2.13	1.32	1.36

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	9[B]	C42	O2-C2-N3	-3.04	117.38	122.33
2	E	9[A]	C42	C2'-C1'-N1	-2.23	108.63	113.77
2	B	9[B]	C42	O4'-C1'-N1	2.20	111.80	107.86
2	B	9[A]	C42	C2'-C1'-N1	-2.11	108.90	113.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	9[B]	C42	1	0
2	B	9[B]	C42	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 22 ligands modelled in this entry, 2 are monoatomic - leaving 20 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
5	DGT	D	902[B]	6	26,33,33	3.37	9 (34%)	32,52,52	1.63	5 (15%)
7	SO4	E	101	-	4,4,4	0.15	0	6,6,6	0.11	0
7	SO4	A	905	-	4,4,4	0.20	0	6,6,6	0.75	0
7	SO4	A	909	-	4,4,4	0.22	0	6,6,6	0.15	0
7	SO4	A	906	-	4,4,4	0.16	0	6,6,6	0.10	0
5	DGT	A	902[B]	6	26,33,33	3.30	7 (26%)	32,52,52	1.69	6 (18%)
7	SO4	A	910	-	4,4,4	0.16	0	6,6,6	0.14	0
7	SO4	B	101	-	4,4,4	0.18	0	6,6,6	0.16	0
7	SO4	D	908	-	4,4,4	0.17	0	6,6,6	0.17	0
7	SO4	D	905	-	4,4,4	0.15	0	6,6,6	0.11	0
7	SO4	D	909	-	4,4,4	0.18	0	6,6,6	0.18	0
4	DPO	D	901[A]	6	6,8,8	0.73	0	13,13,13	1.12	1 (7%)
7	SO4	D	907	-	4,4,4	0.18	0	6,6,6	0.12	0
7	SO4	A	911	-	4,4,4	0.21	0	6,6,6	0.15	0
7	SO4	A	904	-	4,4,4	0.24	0	6,6,6	0.20	0
7	SO4	D	904	-	4,4,4	0.30	0	6,6,6	0.39	0
7	SO4	A	907	-	4,4,4	0.20	0	6,6,6	0.30	0
4	DPO	A	901[A]	6	6,8,8	0.73	0	13,13,13	1.06	1 (7%)
7	SO4	D	906	-	4,4,4	0.14	0	6,6,6	0.10	0
7	SO4	A	908	-	4,4,4	0.16	0	6,6,6	0.13	0

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
4	DPO	A	901[A]	6	-	0/6/6/6	-
5	DGT	D	902[B]	6	-	2/18/34/34	0/3/3/3
4	DPO	D	901[A]	6	-	0/6/6/6	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DGT	A	902[B]	6	-	1/18/34/34	0/3/3/3

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	902[B]	DGT	O4'-C1'	8.20	1.60	1.42
5	A	902[B]	DGT	O6-C6	8.07	1.39	1.23
5	D	902[B]	DGT	O6-C6	7.95	1.39	1.23
5	A	902[B]	DGT	O4'-C4'	-7.88	1.27	1.45
5	A	902[B]	DGT	O4'-C1'	7.58	1.59	1.42
5	D	902[B]	DGT	O4'-C4'	-7.26	1.28	1.45
5	A	902[B]	DGT	C2'-C1'	-6.02	1.35	1.52
5	D	902[B]	DGT	C2'-C1'	-5.99	1.35	1.52
5	D	902[B]	DGT	C2-N2	4.87	1.45	1.34
5	A	902[B]	DGT	C2-N2	4.70	1.45	1.34
5	D	902[B]	DGT	C6-N1	-3.42	1.32	1.37
5	D	902[B]	DGT	C5-C6	-2.87	1.41	1.47
5	A	902[B]	DGT	C6-N1	-2.84	1.33	1.37
5	D	902[B]	DGT	O3'-C3'	-2.70	1.37	1.43
5	A	902[B]	DGT	C5-C6	-2.48	1.42	1.47
5	D	902[B]	DGT	C1'-N9	-2.04	1.43	1.49

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	902[B]	DGT	C2'-C1'-N9	-4.29	104.38	114.27
5	A	902[B]	DGT	C8-N7-C5	3.99	110.60	102.99
5	D	902[B]	DGT	C5-C6-N1	3.58	120.27	113.95
5	A	902[B]	DGT	C2'-C1'-N9	-3.56	106.07	114.27
5	A	902[B]	DGT	C5-C6-N1	3.44	120.03	113.95
5	D	902[B]	DGT	C8-N7-C5	3.19	109.07	102.99
5	A	902[B]	DGT	C2-N1-C6	-2.76	120.01	125.10
4	D	901[A]	DPO	P2-O4-P1	-2.66	123.71	132.83
4	A	901[A]	DPO	P2-O4-P1	-2.34	124.79	132.83
5	A	902[B]	DGT	O3'-C3'-C2'	2.31	119.17	110.90
5	D	902[B]	DGT	PA-O3A-PB	-2.30	124.93	132.83
5	D	902[B]	DGT	C2-N1-C6	-2.28	120.90	125.10
5	A	902[B]	DGT	N2-C2-N1	2.04	121.06	116.71

There are no chirality outliers.

All (3) torsion outliers are listed below:

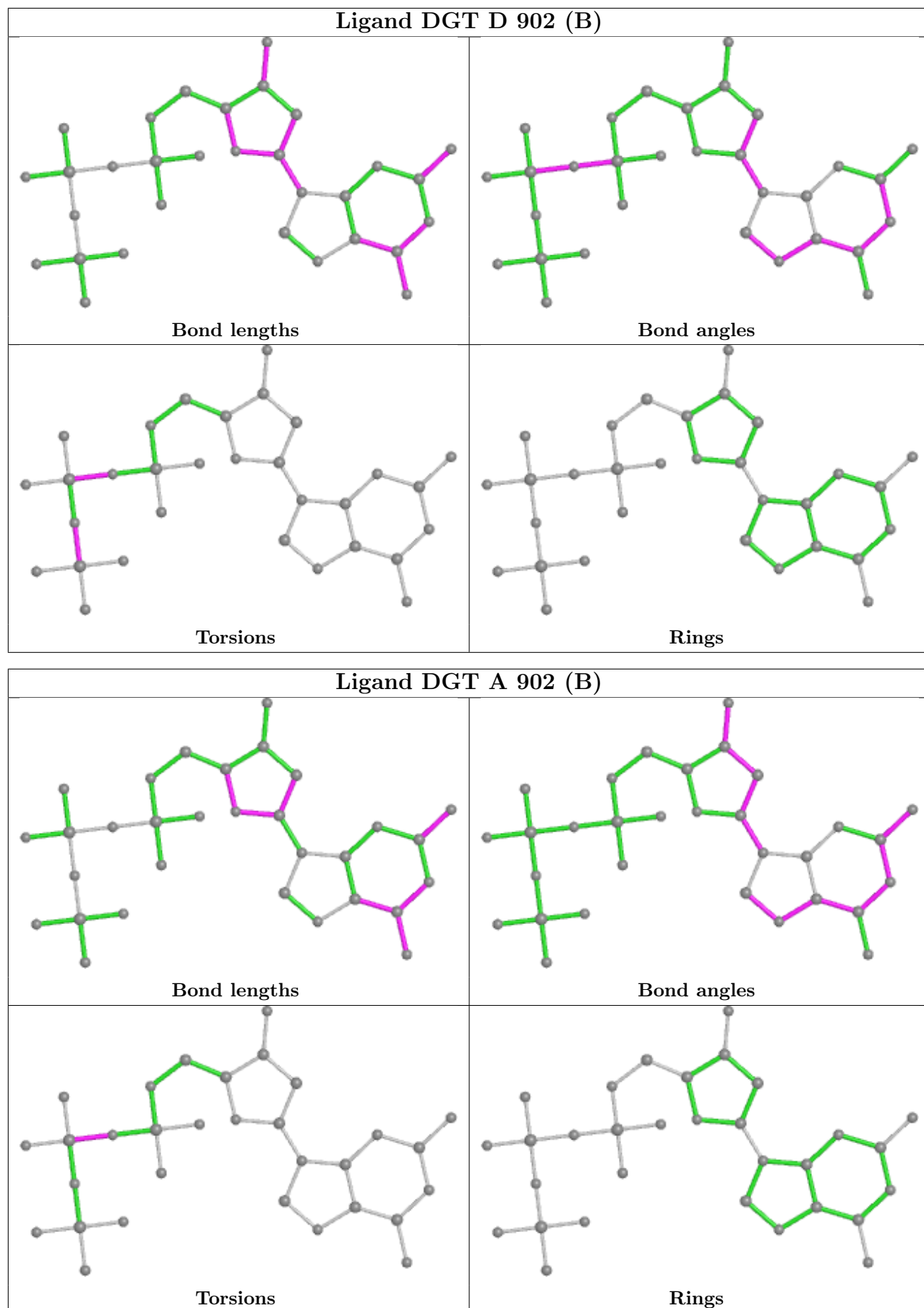
Mol	Chain	Res	Type	Atoms
5	D	902[B]	DGT	PB-O3B-PG-O3G
5	A	902[B]	DGT	PA-O3A-PB-O2B
5	D	902[B]	DGT	PA-O3A-PB-O2B

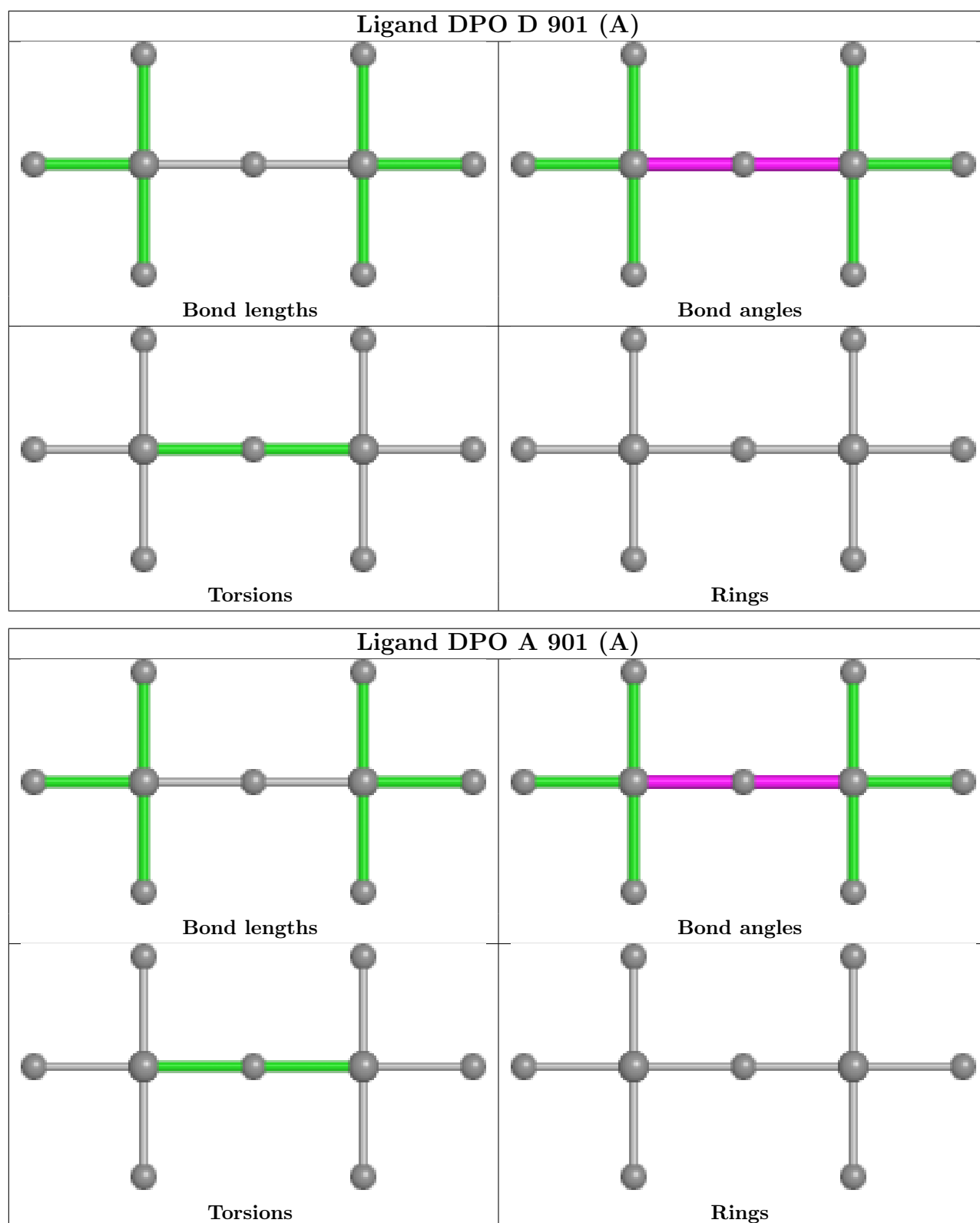
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	902[B]	DGT	1	0
5	A	902[B]	DGT	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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	579/579 (100%)	-0.05	7 (1%) 79 79	15, 27, 52, 69	0
1	D	579/579 (100%)	0.00	16 (2%) 53 54	19, 36, 58, 86	0
2	B	9/10 (90%)	-0.45	0 100 100	21, 26, 44, 49	1 (11%)
2	E	9/10 (90%)	-0.45	0 100 100	23, 26, 48, 48	1 (11%)
3	C	13/13 (100%)	-0.08	1 (7%) 13 14	18, 27, 65, 77	0
3	F	13/13 (100%)	-0.03	1 (7%) 13 14	21, 27, 72, 79	0
All	All	1202/1204 (99%)	-0.03	25 (2%) 63 64	15, 32, 57, 86	2 (0%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	695	ASP	4.0
3	F	13	DA	3.6
3	C	13	DA	3.4
1	A	554	TYR	3.0
1	D	696	GLU	2.9
1	A	523	GLY	2.9
1	D	876	LYS	2.8
1	D	298	LYS	2.7
1	D	631	GLU	2.5
1	D	678	ASP	2.5
1	D	342	GLY	2.5
1	D	523	GLY	2.4
1	A	526	PHE	2.4
1	D	305	ASP	2.4
1	A	525	GLU	2.4
1	D	694	GLU	2.3
1	A	779	ARG	2.2
1	D	299	MET	2.2
1	D	677	ARG	2.2

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	646	ASP	2.1
1	A	520	GLU	2.1
1	A	537	ILE	2.1
1	D	866	TYR	2.1
1	D	520	GLU	2.0
1	D	689	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	C42	B	9[A]	19/20	0.97	0.14	16,20,24,26	19
2	C42	B	9[B]	19/20	0.97	0.14	16,19,25,27	0
2	C42	E	9[A]	19/20	0.97	0.13	23,25,28,28	19
2	C42	E	9[B]	19/20	0.97	0.13	21,25,28,28	0

## 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.

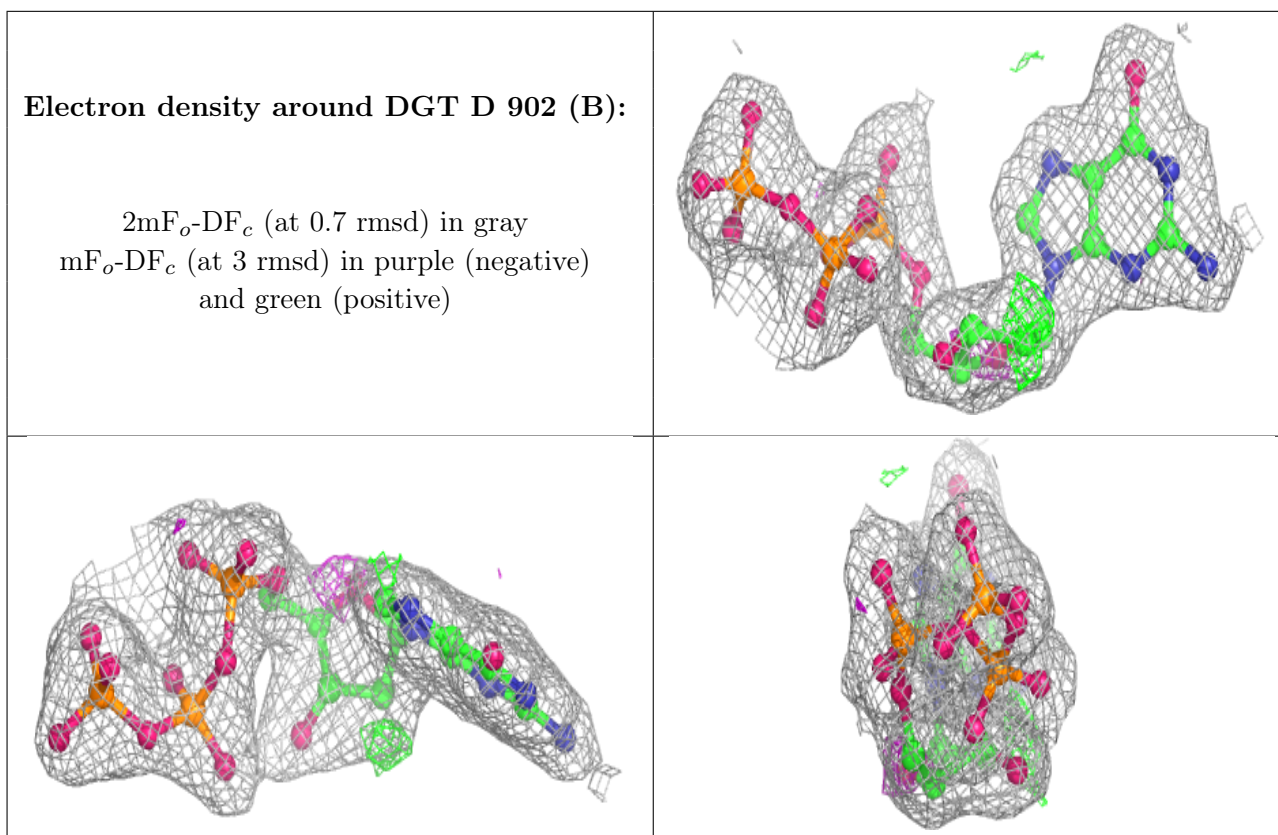
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	SO4	B	101	5/5	0.67	0.33	66,70,79,96	0
7	SO4	A	909	5/5	0.77	0.27	48,53,79,92	0
7	SO4	A	907	5/5	0.79	0.21	50,51,55,80	0
7	SO4	D	904	5/5	0.79	0.17	42,45,48,76	0
7	SO4	D	909	5/5	0.85	0.34	48,52,70,86	0
7	SO4	A	904	5/5	0.86	0.23	53,54,59,75	0
7	SO4	D	908	5/5	0.89	0.26	55,59,71,77	0
7	SO4	E	101	5/5	0.89	0.39	68,75,87,89	0

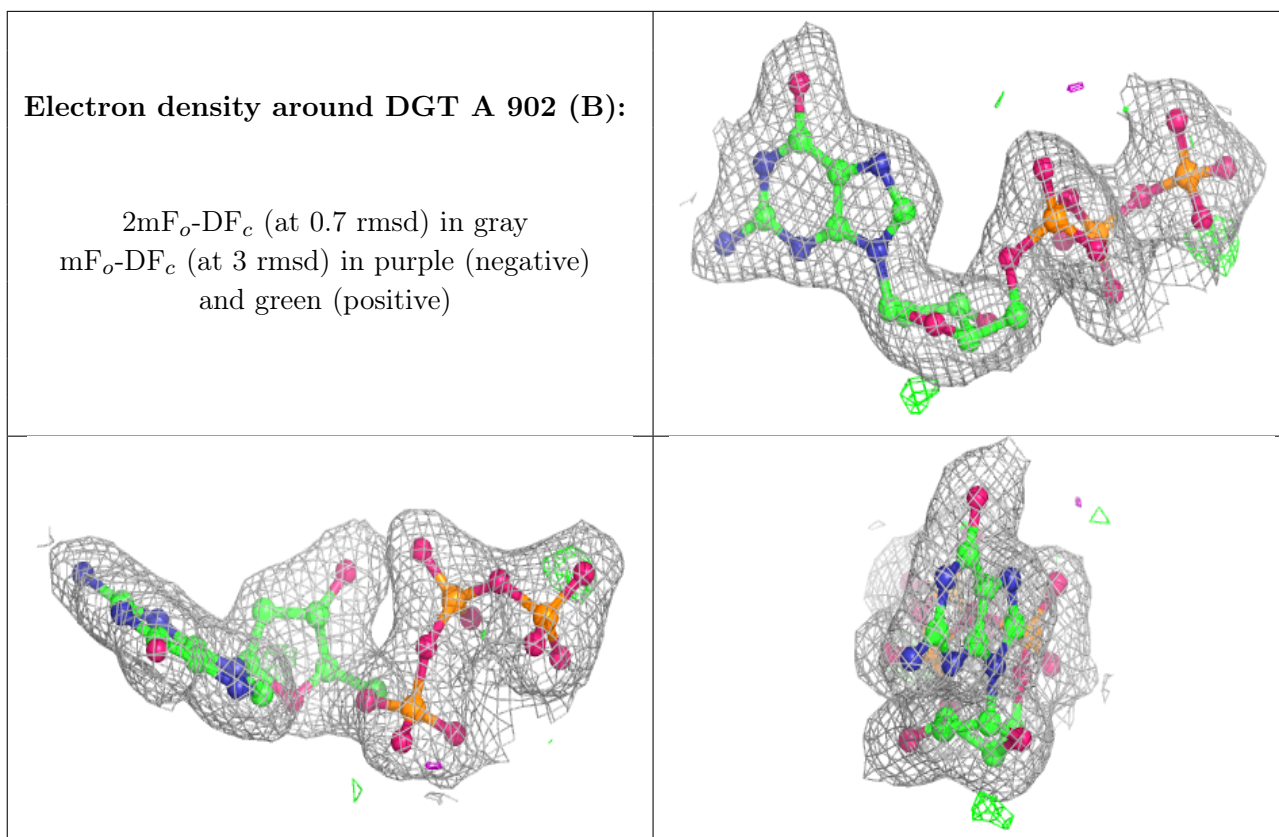
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	SO4	D	905	5/5	0.91	0.17	58,64,66,81	0
7	SO4	A	910	5/5	0.91	0.26	60,65,74,79	0
7	SO4	D	907	5/5	0.93	0.22	66,67,82,93	0
7	SO4	A	911	5/5	0.93	0.18	43,52,67,70	0
7	SO4	A	906	5/5	0.93	0.27	50,62,75,84	0
7	SO4	D	906	5/5	0.93	0.27	63,65,74,87	0
7	SO4	A	908	5/5	0.96	0.20	42,44,63,69	0
7	SO4	A	905	5/5	0.96	0.16	32,39,52,59	0
5	DGT	D	902[B]	31/31	0.97	0.10	23,28,34,39	0
5	DGT	A	902[B]	31/31	0.98	0.12	18,21,27,28	0
4	DPO	D	901[A]	9/9	0.98	0.10	28,32,35,37	9
4	DPO	A	901[A]	9/9	0.99	0.13	23,25,26,26	9
6	CA	D	903	1/1	0.99	0.07	33,33,33,33	0
6	CA	A	903	1/1	1.00	0.11	29,29,29,29	0

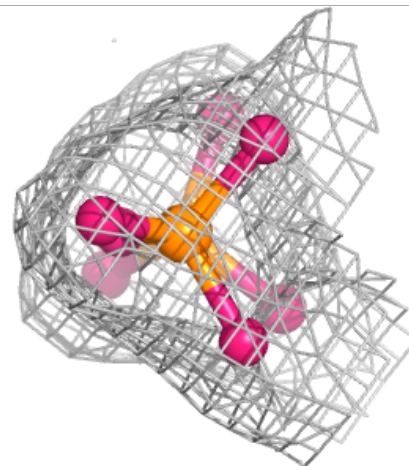
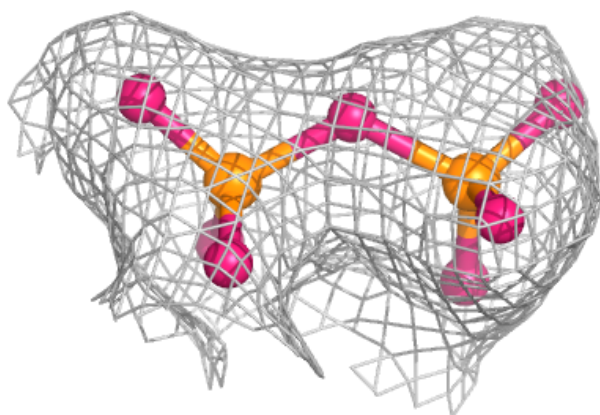
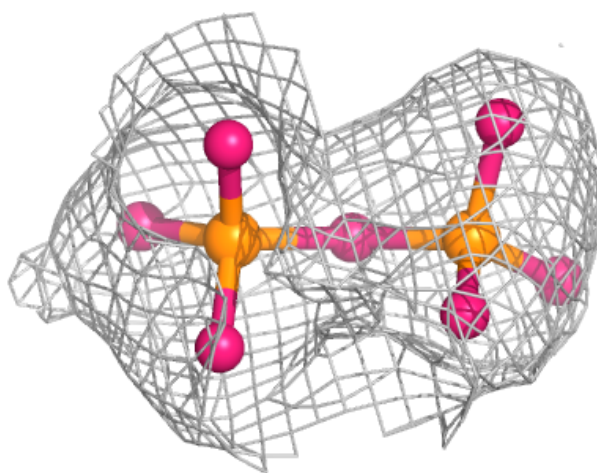
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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





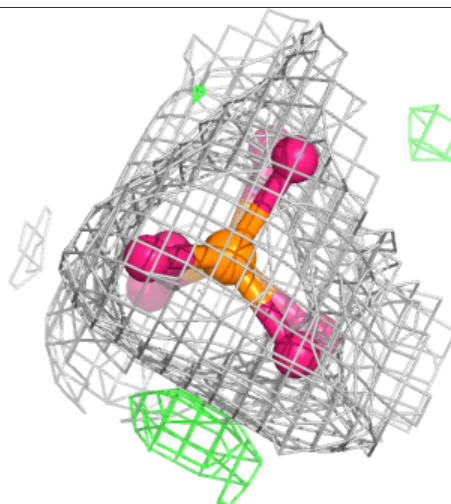
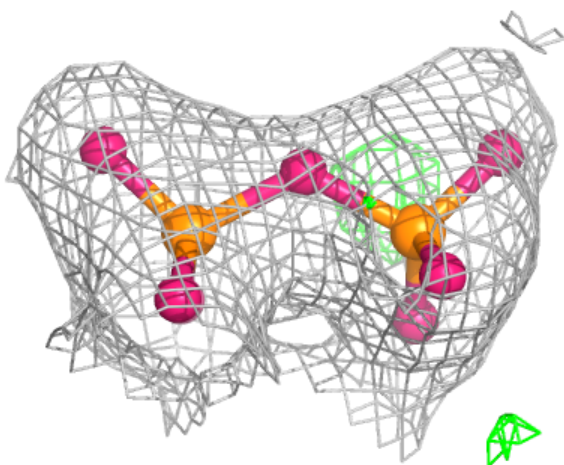
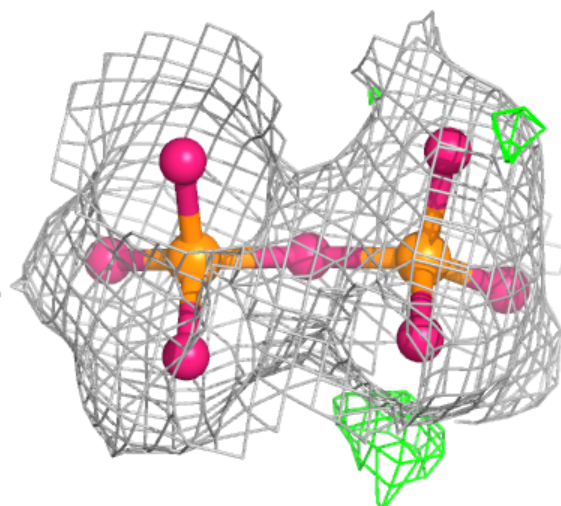
**Electron density around DPO D 901 (A):**

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



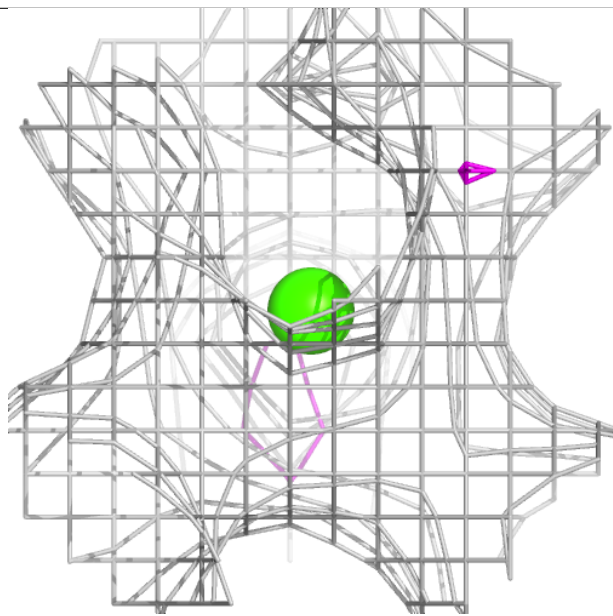
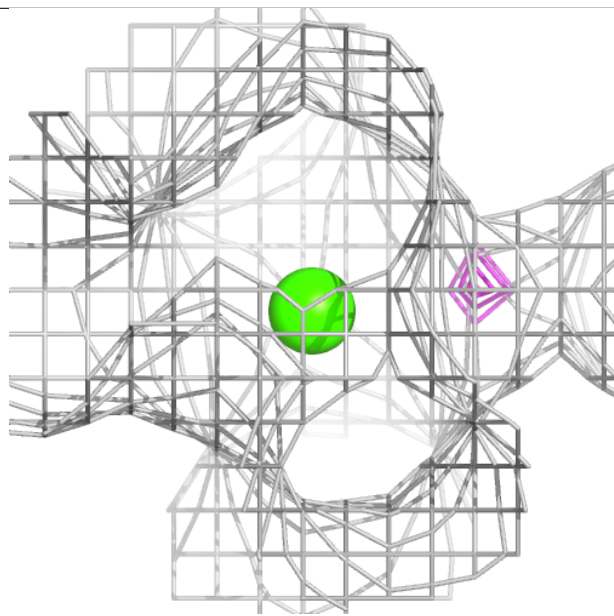
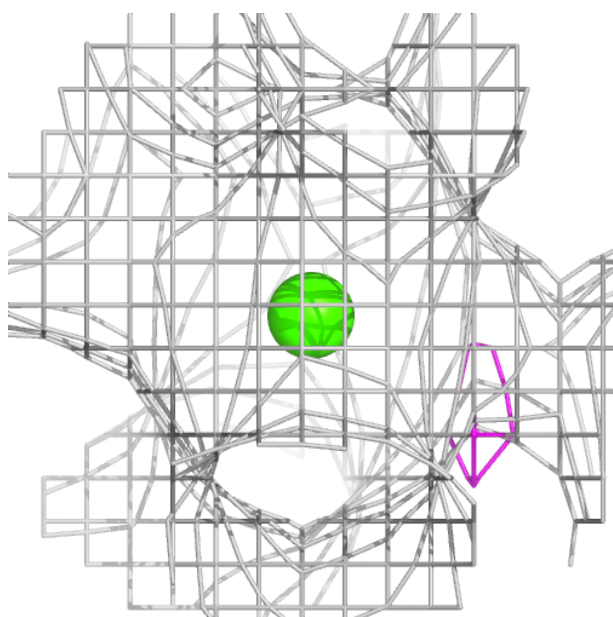
**Electron density around DPO A 901 (A):**

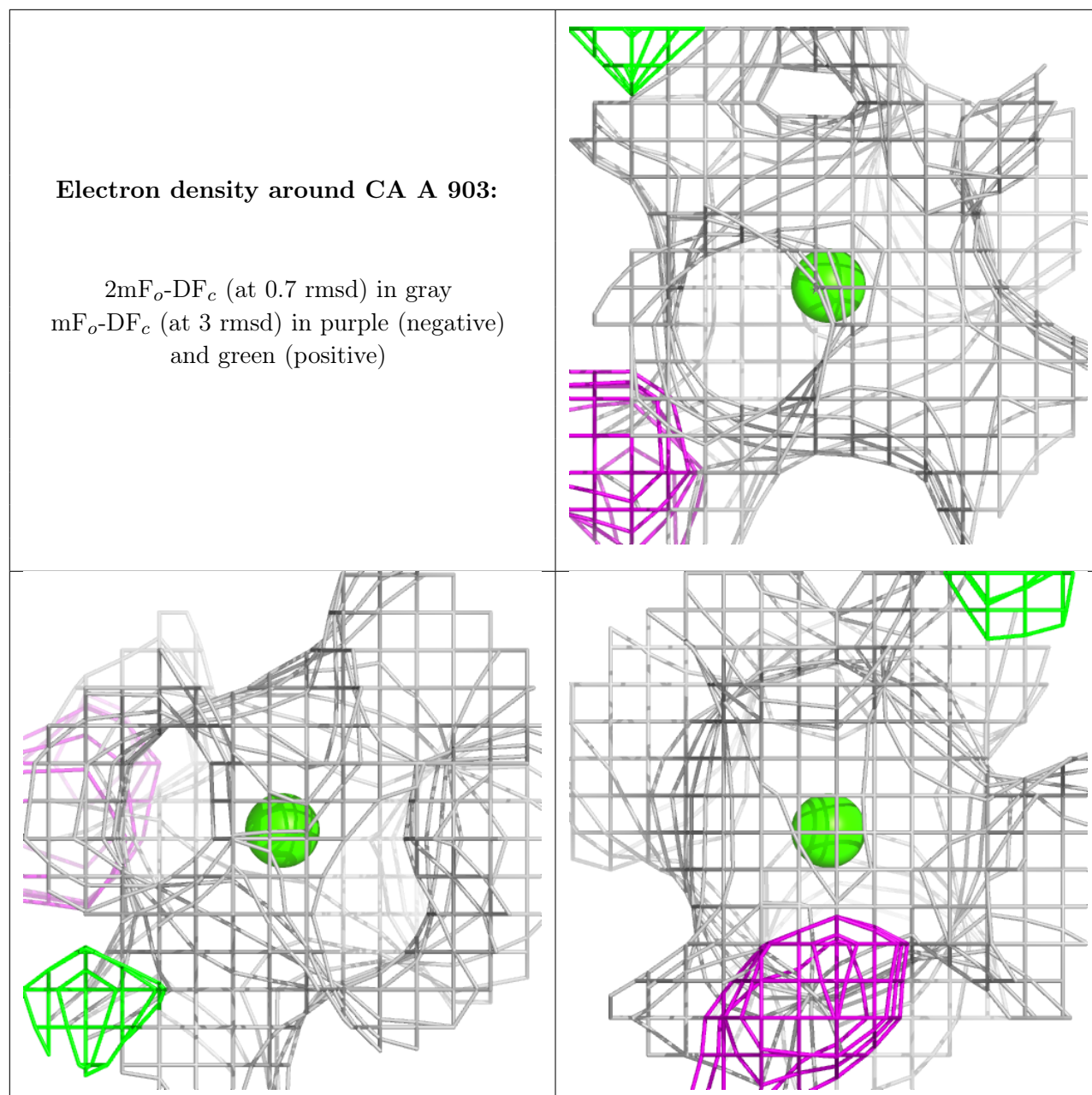
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around CA D 903:**

$2mF_o-DF_c$  (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.