



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 09:23 PM UTC

PDB ID : 8SCG / pdb_00008scg
Title : Bst DNA polymerase I Large Fragment mutant F710Y/D598A with 3'-amino primer, dGTP, and calcium time-resolved 0h (Ground State)
Authors : Fang, Z.; Lelyveld, V.S.; Szostak, J.W.
Deposited on : 2023-04-05
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

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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

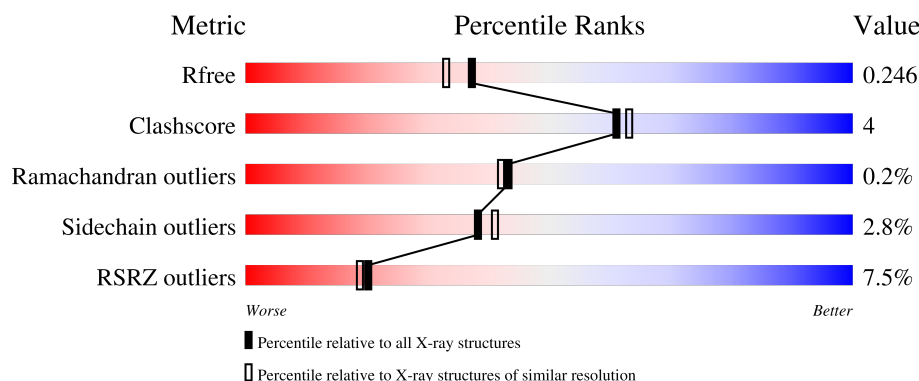
1 Overall quality at a glance

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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	3% 90% 10%
1	D	579	12% 86% 13%
2	B	9	89% 11%
2	E	9	78% 11% 11%
3	C	13	8% 54% 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	13	8% 77% 23%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10699 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
1	A	579	Total	C	N	O	S	0	0	0
			4646	2955	807	867	17			
1	D	579	Total	C	N	O	S	0	0	0
			4646	2955	807	867	17			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	710	TYR	PHE	engineered mutation	UNP D9N168
A	713	VAL	PRO	variant	UNP D9N168
D	710	TYR	PHE	engineered mutation	UNP D9N168
D	713	VAL	PRO	variant	UNP D9N168

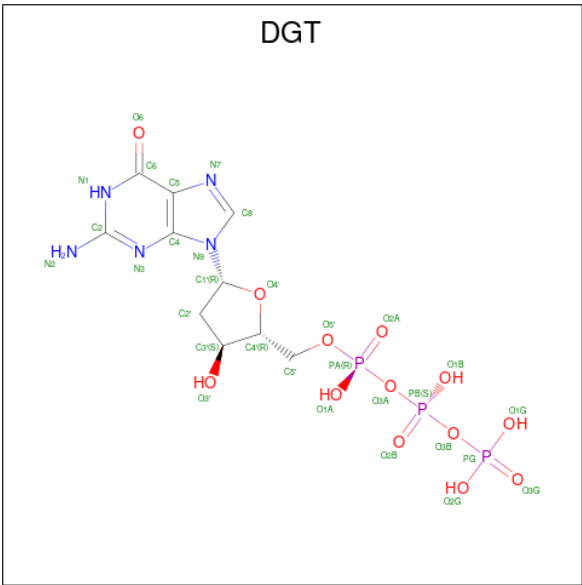
- Molecule 2 is a DNA chain called 3'-Amino DNA Primer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	9	Total	C	N	O	P	0	0	0
			182	87	37	50	8			
2	E	9	Total	C	N	O	P	0	0	0
			182	87	37	50	8			

- Molecule 3 is a DNA chain called DNA Tempelate.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
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 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (CCD ID: DGT) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).

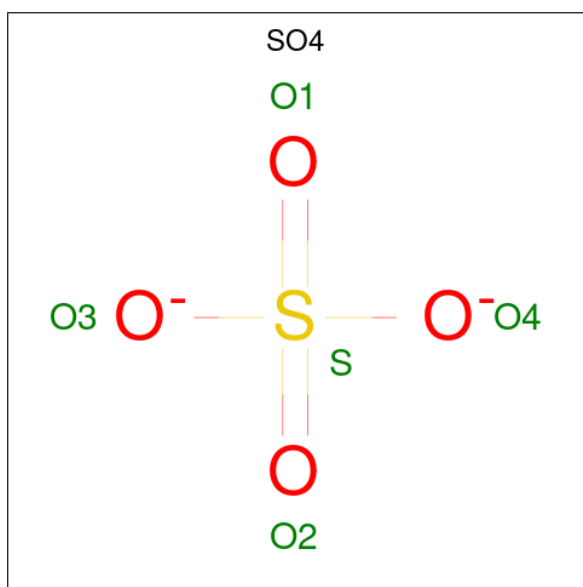


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
4	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 5 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 6 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

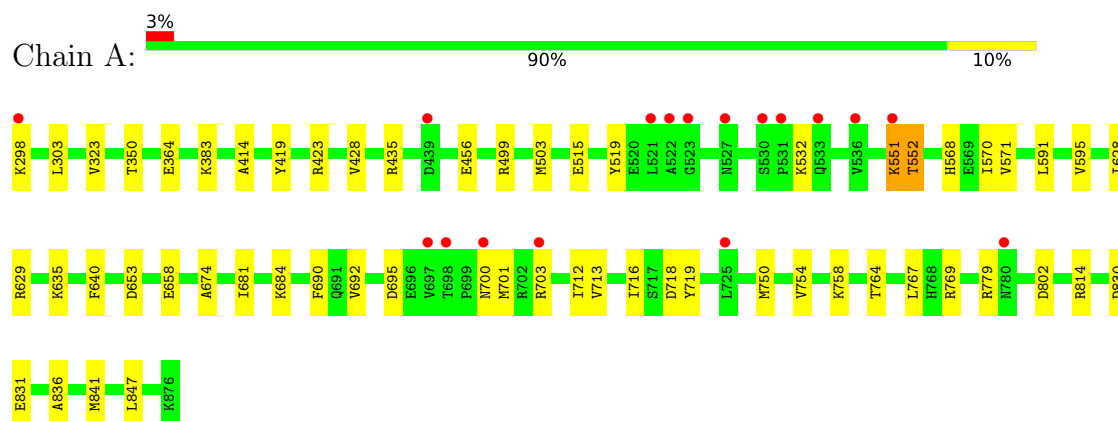
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	250	Total	O	0	0
			250	250		
7	B	19	Total	O	0	0
			19	19		
7	C	31	Total	O	0	0
			31	31		
7	D	123	Total	O	0	0
			123	123		
7	E	17	Total	O	0	0
			17	17		
7	F	27	Total	O	0	0
			27	27		

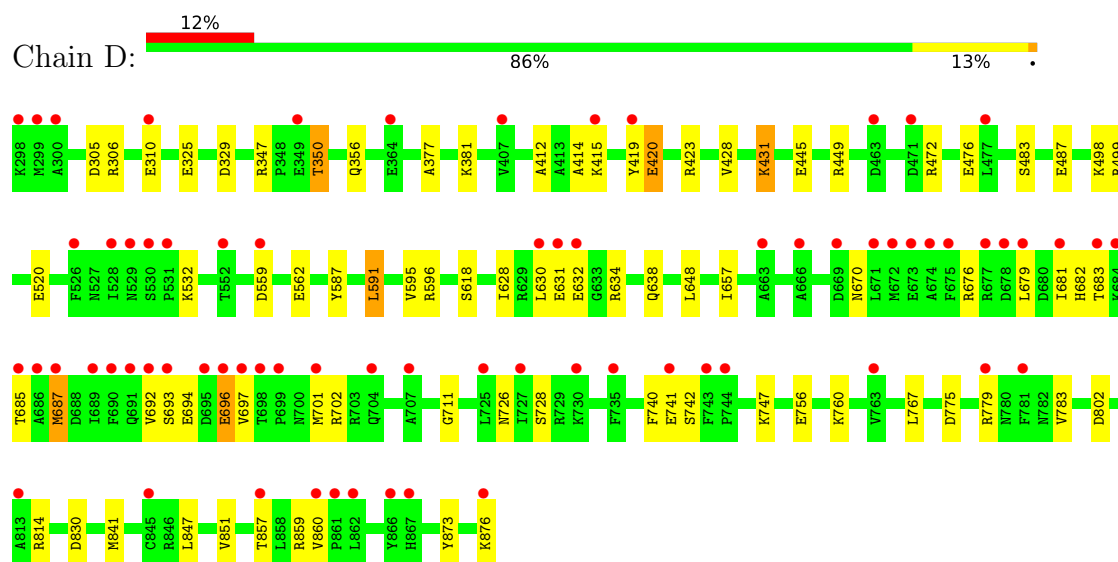
3 Residue-property plots [i](#)

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.

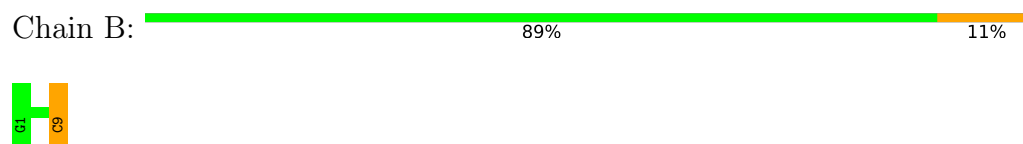
• Molecule 1: DNA polymerase I



• Molecule 1: DNA polymerase I



• Molecule 2: 3'-Amino DNA Primer



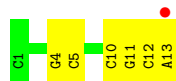
- Molecule 2: 3'-Amino DNA Primer

Chain E:  78% 11% 11%



- Molecule 3: DNA Tempelate

Chain C:  8% 54% 46%



- Molecule 3: DNA Tempelate

Chain F:  8% 77% 23%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.68Å 108.81Å 149.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.32 – 2.00 45.32 – 2.00	Depositor EDS
% Data completeness (in resolution range)	93.4 (45.32-2.00) 93.5 (45.32-2.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.50 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.202 , 0.245 0.204 , 0.246	Depositor DCC
R_{free} test set	4888 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å ²)	26.4	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 37.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10699	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: SO4, DGT, C42, CA

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	0.36	0/4730	0.55	0/6393
1	D	0.25	0/4730	0.45	0/6393
2	B	0.40	0/183	0.58	0/281
2	E	0.37	0/183	0.64	0/281
3	C	0.50	0/275	0.61	0/423
3	F	0.43	0/275	0.58	0/423
All	All	0.32	0/10376	0.51	0/14194

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	4646	0	4705	36	0
1	D	4646	0	4705	45	0
2	B	182	0	103	2	0
2	E	182	0	103	4	0
3	C	246	0	135	5	0
3	F	246	0	135	3	0
4	A	31	0	12	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	31	0	12	1	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	15	0	0	0	0
6	C	5	0	0	0	0
7	A	250	0	0	4	0
7	B	19	0	0	0	0
7	C	31	0	0	0	0
7	D	123	0	0	4	0
7	E	17	0	0	1	0
7	F	27	0	0	0	0
All	All	10699	0	9910	85	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 (85) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:814:ARG:HG2	1:A:847:LEU:HD11	1.60	0.81
2:E:3:DG:N7	7:E:101:HOH:O	2.16	0.78
1:D:445:GLU:OE2	1:D:449:ARG:NH1	2.20	0.74
1:D:779:ARG:NH2	7:D:1001:HOH:O	2.23	0.72
1:A:703:ARG:O	1:A:703:ARG:NH1	2.25	0.69
1:A:551:LYS:HZ2	1:D:431:LYS:HG2	1.58	0.68
1:A:552:THR:OG1	3:C:13:DA:N3	2.27	0.67
1:D:697:VAL:HG13	1:D:701:MET:HE2	1.78	0.65
1:D:830:ASP:OD1	2:E:9:C42:N	2.29	0.64
1:A:383:LYS:NZ	7:A:1003:HOH:O	2.31	0.62
1:D:356:GLN:NE2	7:D:1004:HOH:O	2.33	0.61
1:D:329:ASP:OD2	1:D:596:ARG:NH1	2.33	0.61
1:D:591:LEU:O	1:D:595:VAL:HG23	2.01	0.61
4:A:901:DGT:O4'	2:B:9:C42:H2'1	2.01	0.60
1:D:472:ARG:NH1	1:D:476:GLU:OE2	2.34	0.60
1:D:687:MET:HE1	1:D:694:GLU:HA	1.82	0.59
1:D:498:LYS:HD3	7:D:1003:HOH:O	2.02	0.59
1:A:814:ARG:CG	1:A:847:LEU:HD11	2.31	0.58
1:A:629:ARG:NH1	1:A:700:ASN:OD1	2.37	0.57
1:A:591:LEU:O	1:A:595:VAL:HG23	2.03	0.57
1:D:499:ARG:NH1	1:D:638:GLN:OE1	2.39	0.55
1:D:414:ALA:HB1	1:D:419:TYR:HB3	1.88	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:10:DC:H2''	3:C:11:DG:C8	2.42	0.54
1:D:412:ALA:O	1:D:415:LYS:HG2	2.07	0.54
1:D:634:ARG:HH12	1:D:876:LYS:HD3	1.72	0.54
1:D:670:ASN:ND2	1:D:742:SER:O	2.38	0.54
1:D:676:ARG:NH1	1:D:859:ARG:O	2.22	0.54
1:A:692:VAL:HG21	1:A:701:MET:HE1	1.88	0.54
1:D:483:SER:O	1:D:487:GLU:HG3	2.09	0.53
4:D:901:DGT:O4'	2:E:9:C42:H2'1	2.08	0.52
1:A:414:ALA:HB1	1:A:419:TYR:HB3	1.91	0.52
1:A:423:ARG:NH2	1:A:428:VAL:HG22	2.25	0.51
1:A:690:PHE:HB2	1:A:701:MET:HE3	1.92	0.50
1:D:814:ARG:HG2	1:D:847:LEU:HD11	1.92	0.50
1:D:682:HIS:HB2	1:D:702:ARG:NH1	2.27	0.50
1:A:779:ARG:NH1	1:D:420:GLU:OE1	2.46	0.49
1:A:532:LYS:HG2	3:C:13:DA:C8	2.47	0.49
1:D:814:ARG:CG	1:D:847:LEU:HD11	2.42	0.49
1:D:596:ARG:NH2	7:D:1002:HOH:O	2.28	0.48
1:D:423:ARG:NH1	1:D:428:VAL:HG22	2.29	0.48
1:D:648:LEU:HD12	1:D:841:MET:HG3	1.94	0.48
1:D:847:LEU:O	1:D:851:VAL:HG23	2.13	0.48
1:D:587:TYR:O	1:D:591:LEU:HB2	2.14	0.47
1:A:515:GLU:HG2	1:A:519:TYR:CE2	2.49	0.47
1:A:591:LEU:HD21	1:A:640:PHE:CZ	2.50	0.47
1:D:687:MET:HE2	1:D:687:MET:HB3	1.61	0.47
1:D:740:PHE:HB3	1:D:747:LYS:HD2	1.96	0.47
3:C:12:DC:O3'	3:C:13:DA:H3'	2.13	0.47
1:D:767:LEU:HD12	1:D:802:ASP:HB3	1.96	0.47
1:A:674:ALA:HB1	1:A:681:ILE:HD11	1.97	0.47
1:D:618:SER:O	3:F:7:DG:H4'	2.15	0.47
1:D:681:ILE:O	1:D:685:THR:HG23	2.16	0.46
1:A:712:ILE:HA	1:A:716:ILE:HG22	1.96	0.46
1:A:713:VAL:O	1:A:750:MET:HE3	2.16	0.46
1:A:830:ASP:OD1	2:B:9:C42:N	2.46	0.46
1:A:684:LYS:HD3	1:A:684:LYS:HA	1.71	0.46
1:A:695:ASP:N	1:A:695:ASP:OD1	2.48	0.45
1:D:630:LEU:HD12	1:D:631:GLU:H	1.81	0.45
1:A:767:LEU:HD12	1:A:802:ASP:HB3	1.99	0.45
1:D:347:ARG:HB2	1:D:350:THR:HG23	1.98	0.45
1:D:687:MET:H	1:D:687:MET:HG2	1.56	0.45
1:A:690:PHE:CB	1:A:701:MET:HE3	2.47	0.44
1:A:364:GLU:HG3	7:A:1158:HOH:O	2.17	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:559:ASP:N	1:D:559:ASP:OD1	2.50	0.44
1:A:658:GLU:OE2	4:A:901:DGT:O3'	2.25	0.44
1:A:298:LYS:HE2	7:A:1212:HOH:O	2.17	0.43
1:A:718:ASP:OD1	1:A:719:TYR:N	2.50	0.43
1:A:764:THR:HA	1:A:769:ARG:O	2.17	0.43
1:D:693:SER:O	1:D:696:GLU:HG2	2.18	0.43
1:D:532:LYS:HD3	3:F:13:DA:C8	2.53	0.43
3:C:4:DG:H2'	3:C:5:DC:C6	2.54	0.42
1:A:503:MET:HE1	1:A:635:LYS:O	2.18	0.42
1:D:377:ALA:O	1:D:381:LYS:HG3	2.20	0.42
1:D:498:LYS:N	1:D:498:LYS:HE2	2.35	0.42
1:D:775:ASP:HB3	1:D:783:VAL:HG12	2.02	0.42
1:A:754:VAL:HG12	1:A:758:LYS:HE2	2.01	0.41
1:A:456:GLU:HG2	7:A:1056:HOH:O	2.19	0.41
1:D:711:GLY:HA2	3:F:3:DC:C2	2.55	0.41
1:D:756:GLU:HG2	1:D:760:LYS:HD3	2.02	0.41
1:A:568:HIS:O	1:A:571:VAL:HG22	2.21	0.41
1:A:653:ASP:HB2	1:A:831:GLU:HG2	2.03	0.40
1:D:830:ASP:CG	2:E:9:C42:HN1	2.26	0.40
1:A:323:VAL:O	1:A:435:ARG:NH2	2.47	0.40
1:A:836:ALA:HB3	1:A:841:MET:CE	2.51	0.40
1:D:638:GLN:HG3	1:D:873:TYR:CG	2.57	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	577/579 (100%)	562 (97%)	14 (2%)	1 (0%)	43	42
1	D	577/579 (100%)	557 (96%)	19 (3%)	1 (0%)	43	42
All	All	1154/1158 (100%)	1119 (97%)	33 (3%)	2 (0%)	43	42

All (2) Ramachandran outliers are listed below:

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

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	495/495 (100%)	489 (99%)	6 (1%)	63	70
1	D	495/495 (100%)	473 (96%)	22 (4%)	25	24
All	All	990/990 (100%)	962 (97%)	28 (3%)	38	41

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

Mol	Chain	Res	Type
1	A	303	LEU
1	A	350	THR
1	A	499	ARG
1	A	551	LYS
1	A	552	THR
1	A	570	ILE
1	D	305	ASP
1	D	306	ARG
1	D	310	GLU
1	D	325	GLU
1	D	350	THR
1	D	420	GLU
1	D	431	LYS
1	D	520	GLU
1	D	562	GLU
1	D	591	LEU
1	D	632	GLU
1	D	657	ILE
1	D	679	LEU
1	D	683	THR
1	D	687	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	692	VAL
1	D	696	GLU
1	D	726	ASN
1	D	728	SER
1	D	741	GLU
1	D	857	THR
1	D	860	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	468	ASN
1	A	470	GLN
1	A	573	ASN
1	A	759	GLN
1	A	793	ASN
1	A	867	HIS
1	D	470	GLN
1	D	502	GLN
1	D	768	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	B	9	2,3	17,20,21	3.19	10 (58%)	20,28,31	1.48	3 (15%)
2	C42	E	9	2,3	17,20,21	3.31	9 (52%)	20,28,31	1.21	2 (10%)

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	B	9	2,3	-	0/7/21/22	0/2/2/2
2	C42	E	9	2,3	-	0/7/21/22	0/2/2/2

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	9	C42	O4'-C1'	6.90	1.57	1.42
2	B	9	C42	O4'-C1'	6.75	1.57	1.42
2	E	9	C42	O4'-C4'	-6.61	1.30	1.45
2	B	9	C42	O4'-C4'	-6.11	1.31	1.45
2	E	9	C42	C1'-N1	-4.67	1.36	1.48
2	E	9	C42	C4-N4	4.59	1.45	1.33
2	B	9	C42	C4-N4	4.37	1.44	1.33
2	E	9	C42	C4'-C3'	4.34	1.61	1.52
2	B	9	C42	C1'-N1	-4.32	1.37	1.48
2	B	9	C42	C4'-C3'	3.77	1.60	1.52
2	B	9	C42	C2-N1	-3.09	1.33	1.40
2	E	9	C42	C2-N1	-2.88	1.34	1.40
2	E	9	C42	C6-N1	-2.43	1.32	1.38
2	E	9	C42	O2-C2	-2.33	1.19	1.23
2	B	9	C42	O2-C2	-2.31	1.19	1.23
2	E	9	C42	C2-N3	-2.29	1.31	1.36
2	B	9	C42	C2-N3	-2.26	1.31	1.36
2	B	9	C42	C6-N1	-2.19	1.32	1.38
2	B	9	C42	C6-C5	2.04	1.39	1.35

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	9	C42	C2'-C3'-C4'	-3.99	97.81	103.06
2	B	9	C42	O2-C2-N3	-2.91	117.74	122.33
2	E	9	C42	O2-C2-N3	-2.81	117.90	122.33
2	E	9	C42	O4'-C1'-N1	2.32	111.98	107.86
2	B	9	C42	O4'-C1'-C2'	-2.28	101.98	106.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	9	C42	2	0
2	E	9	C42	3	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	$\# Z > 2$
4	DGT	A	901	5	32,33,33	3.22	12 (37%)	48,52,52	1.85	10 (20%)
6	SO4	A	904	-	4,4,4	0.30	0	6,6,6	0.23	0
4	DGT	D	901	5	32,33,33	3.47	14 (43%)	48,52,52	1.67	8 (16%)
6	SO4	A	905	-	4,4,4	0.30	0	6,6,6	0.10	0
6	SO4	C	101	-	4,4,4	0.29	0	6,6,6	0.19	0
6	SO4	A	903	-	4,4,4	0.24	0	6,6,6	0.37	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	DGT	D	901	5	-	3/22/34/34	0/3/3/3
4	DGT	A	901	5	-	5/22/34/34	0/3/3/3

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	DGT	O6-C6	8.34	1.39	1.23
4	D	901	DGT	O6-C6	8.33	1.39	1.23
4	D	901	DGT	O4'-C1'	7.97	1.60	1.42
4	D	901	DGT	O4'-C4'	-7.85	1.27	1.45
4	A	901	DGT	O4'-C1'	7.63	1.59	1.42
4	A	901	DGT	O4'-C4'	-7.49	1.28	1.45
4	D	901	DGT	C2'-C1'	-6.94	1.33	1.52
4	A	901	DGT	C2'-C1'	-6.85	1.33	1.52
4	A	901	DGT	C2-N2	4.97	1.45	1.34
4	D	901	DGT	PB-O3A	4.94	1.64	1.59
4	D	901	DGT	C2-N2	4.76	1.45	1.34
4	D	901	DGT	PB-O3B	4.73	1.64	1.59
4	D	901	DGT	PA-O3A	4.27	1.64	1.59
4	D	901	DGT	C6-N1	-3.75	1.31	1.38
4	A	901	DGT	PB-O3B	3.43	1.63	1.59
4	A	901	DGT	C6-N1	-3.21	1.32	1.38
4	A	901	DGT	PB-O3A	3.16	1.62	1.59
4	D	901	DGT	O3'-C3'	-2.70	1.37	1.43
4	A	901	DGT	C2-N1	-2.54	1.31	1.37
4	A	901	DGT	C8-N9	-2.25	1.32	1.37
4	D	901	DGT	C2-N1	-2.24	1.32	1.37
4	A	901	DGT	PA-O3A	2.22	1.61	1.59
4	A	901	DGT	O3'-C3'	-2.20	1.38	1.43
4	D	901	DGT	C5-N7	-2.04	1.35	1.39
4	D	901	DGT	C8-N9	-2.03	1.33	1.37
4	D	901	DGT	C1'-N9	-2.01	1.43	1.48

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	DGT	C5-C4-N3	-5.93	118.95	128.39
4	A	901	DGT	C2-N3-C4	5.63	122.00	112.30
4	D	901	DGT	C5-C4-N3	-5.43	119.74	128.39
4	D	901	DGT	C2-N3-C4	5.14	121.16	112.30
4	A	901	DGT	N9-C4-N3	3.75	133.45	125.95
4	D	901	DGT	N9-C4-N3	3.64	133.24	125.95
4	A	901	DGT	C6-C5-N7	3.57	136.78	130.29
4	D	901	DGT	C6-C5-N7	3.21	136.13	130.29
4	A	901	DGT	C4-C5-N7	-2.78	106.26	110.67
4	A	901	DGT	C5-C6-N1	2.58	119.81	113.25
4	D	901	DGT	C5-C6-N1	2.49	119.59	113.25
4	A	901	DGT	C8-N7-C5	2.40	108.54	104.26
4	D	901	DGT	C4-C5-N7	-2.39	106.88	110.67

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	DGT	O2G-PG-O3B	2.37	112.58	104.64
4	A	901	DGT	O3A-PB-O2B	2.28	117.58	110.70
4	A	901	DGT	O1A-PA-O3A	-2.28	101.11	107.27
4	D	901	DGT	C8-N7-C5	2.23	108.23	104.26
4	D	901	DGT	N9-C8-N7	-2.15	109.41	113.40

There are no chirality outliers.

All (8) torsion outliers are listed below:

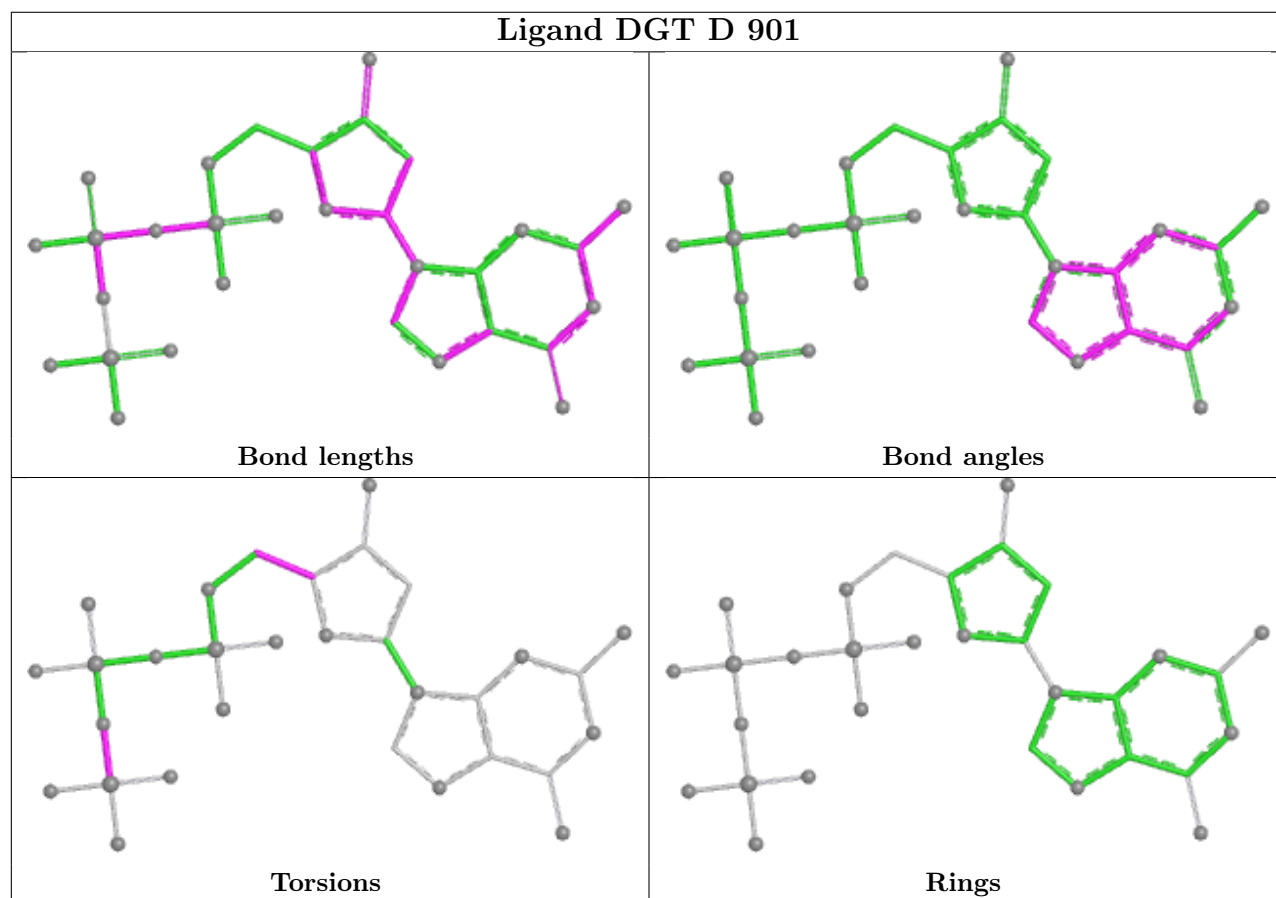
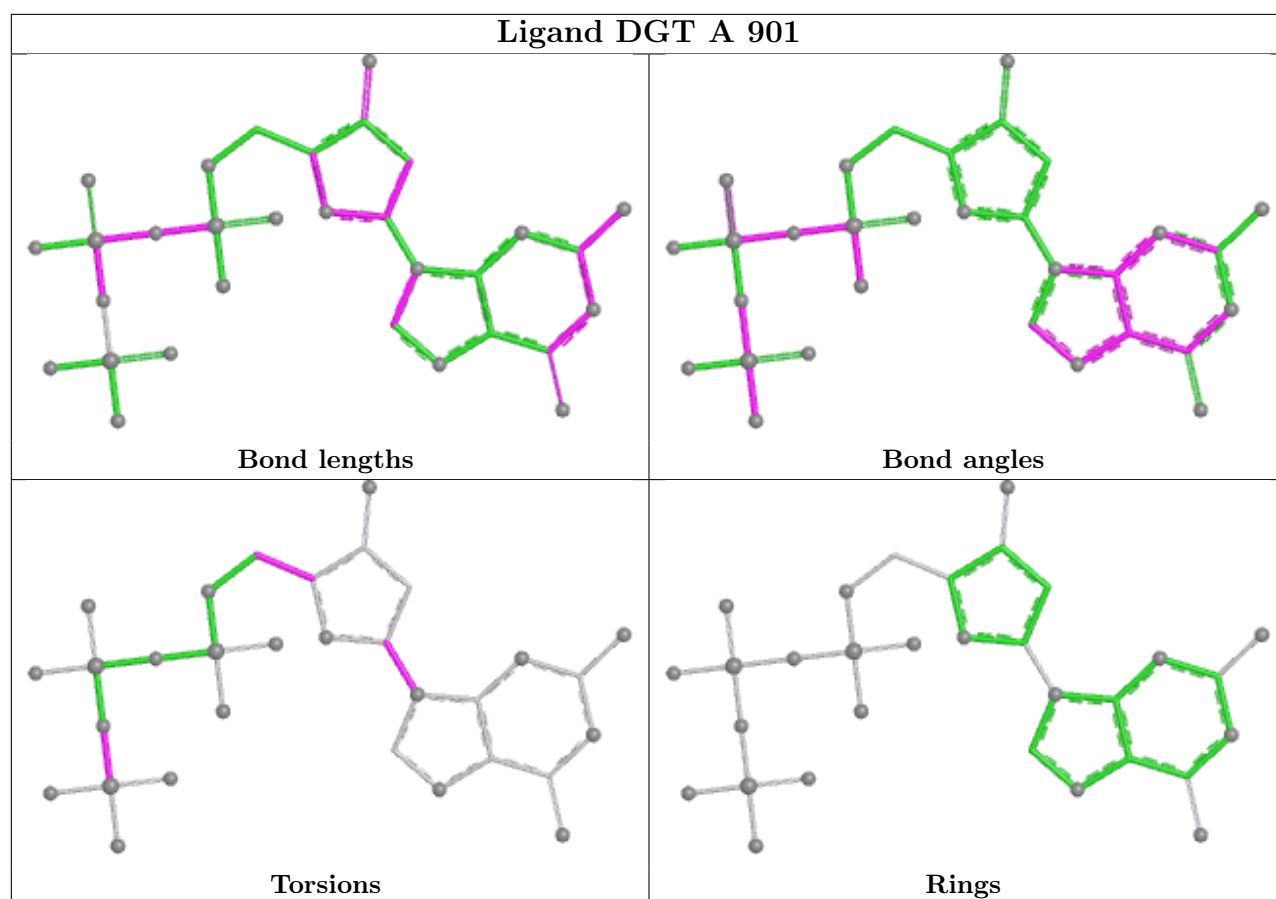
Mol	Chain	Res	Type	Atoms
4	A	901	DGT	PB-O3B-PG-O2G
4	A	901	DGT	O4'-C4'-C5'-O5'
4	A	901	DGT	C3'-C4'-C5'-O5'
4	D	901	DGT	O4'-C4'-C5'-O5'
4	D	901	DGT	C3'-C4'-C5'-O5'
4	A	901	DGT	PB-O3B-PG-O3G
4	D	901	DGT	PB-O3B-PG-O3G
4	A	901	DGT	O4'-C1'-N9-C4

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	DGT	2	0
4	D	901	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

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

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, 95th 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 ²)	Q<0.9
1	A	579/579 (100%)	0.14	17 (2%) 53 52	12, 25, 47, 66	0
1	D	579/579 (100%)	0.90	71 (12%) 8 7	21, 39, 67, 104	0
2	B	8/9 (88%)	-0.08	0 100 100	19, 28, 43, 45	0
2	E	8/9 (88%)	-0.17	0 100 100	25, 28, 45, 45	0
3	C	13/13 (100%)	0.04	1 (7%) 19 18	17, 24, 51, 70	0
3	F	13/13 (100%)	0.04	1 (7%) 19 18	21, 28, 68, 68	0
All	All	1200/1202 (99%)	0.50	90 (7%) 20 19	12, 32, 60, 104	0

All (90) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	683	THR	5.2
1	D	699	PRO	4.5
1	D	861	PRO	4.2
1	D	693	SER	4.0
1	D	679	LEU	3.9
1	D	698	THR	3.9
1	D	678	ASP	3.8
1	D	695	ASP	3.8
1	D	697	VAL	3.8
1	D	687	MET	3.6
1	D	674	ALA	3.6
1	D	686	ALA	3.6
1	D	866	TYR	3.5
1	D	669	ASP	3.2
1	D	685	THR	3.2
1	D	690	PHE	3.2
1	A	298	LYS	3.1
1	D	531	PRO	3.1
1	D	677	ARG	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	691	GLN	3.1
1	D	857	THR	3.1
1	D	630	LEU	3.1
1	D	530	SER	3.1
1	D	671	LEU	3.0
1	D	689	ILE	3.0
1	A	531	PRO	2.9
1	A	703	ARG	2.9
1	A	530	SER	2.9
1	A	523	GLY	2.9
3	F	13	DA	2.9
1	D	779	ARG	2.8
1	D	845	CYS	2.8
1	D	701	MET	2.8
1	D	526	PHE	2.8
1	D	559	ASP	2.8
1	D	298	LYS	2.8
1	D	876	LYS	2.7
1	A	521	LEU	2.7
1	D	631	GLU	2.7
1	D	692	VAL	2.7
1	D	675	PHE	2.6
1	D	860	VAL	2.6
1	D	681	ILE	2.6
1	D	299	MET	2.6
1	D	663	ALA	2.6
3	C	13	DA	2.5
1	D	781	PHE	2.5
1	D	672	MET	2.5
1	D	463	ASP	2.5
1	D	741	GLU	2.5
1	D	684	LYS	2.5
1	D	364	GLU	2.4
1	D	528	ILE	2.4
1	D	744	PRO	2.4
1	D	666	ALA	2.4
1	A	697	VAL	2.4
1	D	419	TYR	2.4
1	D	696	GLU	2.4
1	A	551	LYS	2.4
1	D	529	ASN	2.3
1	D	862	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	704	GLN	2.3
1	D	300	ALA	2.3
1	A	698	THR	2.3
1	D	725	LEU	2.2
1	A	780	ASN	2.2
1	D	552	THR	2.2
1	D	310	GLU	2.2
1	A	439	ASP	2.2
1	D	813	ALA	2.2
1	D	471	ASP	2.2
1	D	743	PHE	2.1
1	D	730	LYS	2.1
1	D	673	GLU	2.1
1	A	533	GLN	2.1
1	D	727	ILE	2.1
1	A	522	ALA	2.1
1	D	707	ALA	2.1
1	D	867	HIS	2.1
1	D	349	GLU	2.1
1	A	700	ASN	2.1
1	A	536	VAL	2.1
1	D	735	PHE	2.1
1	A	725	LEU	2.0
1	D	407	VAL	2.0
1	D	632	GLU	2.0
1	D	763	VAL	2.0
1	A	527	ASN	2.0
1	D	477	LEU	2.0
1	D	415	LYS	2.0

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

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	C42	E	9	19/20	0.95	0.09	23,27,34,37	0
2	C42	B	9	19/20	0.97	0.06	14,18,24,26	0

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

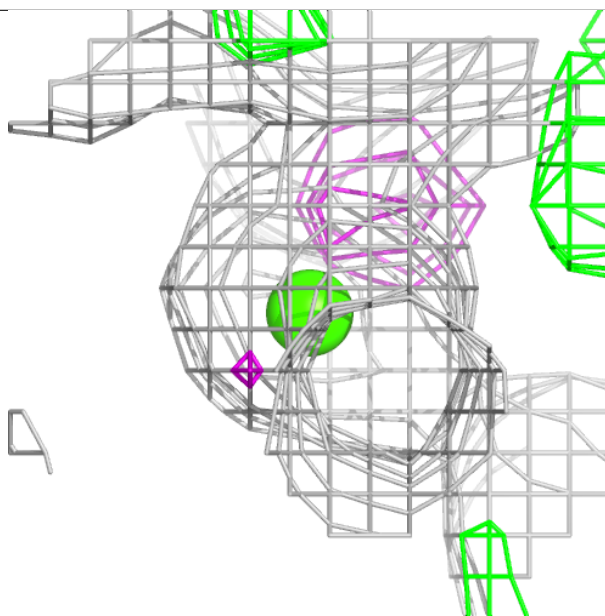
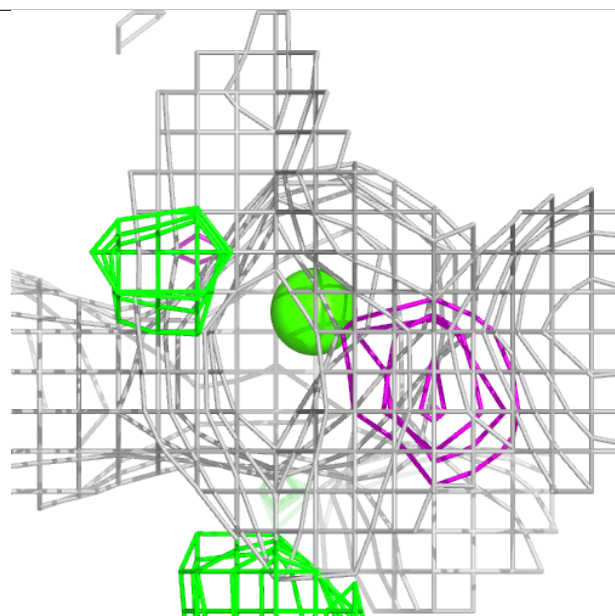
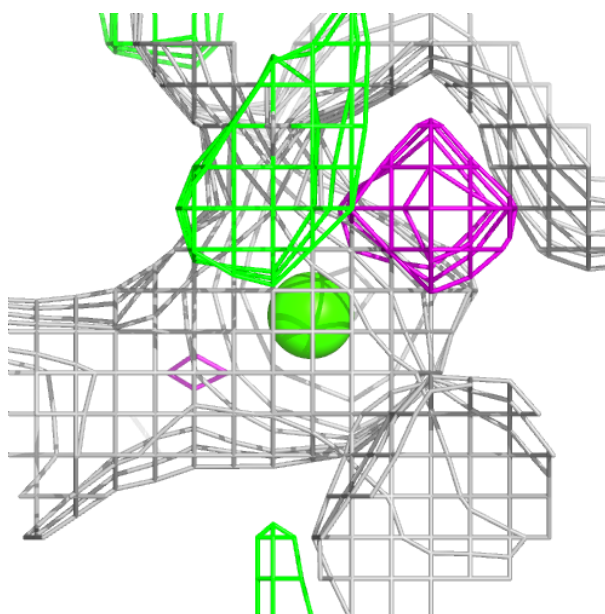
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
6	SO4	C	101	5/5	0.52	0.15	52,57,67,80	0
6	SO4	A	904	5/5	0.73	0.17	41,52,59,76	0
6	SO4	A	905	5/5	0.74	0.12	48,54,60,80	0
5	CA	D	902	1/1	0.86	0.18	62,62,62,62	0
4	DGT	D	901	31/31	0.89	0.11	26,43,56,60	0
6	SO4	A	903	5/5	0.93	0.12	30,40,42,52	0
4	DGT	A	901	31/31	0.95	0.09	16,25,36,39	0
5	CA	A	902	1/1	0.98	0.14	43,43,43,43	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.

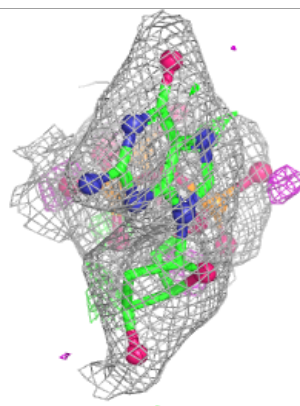
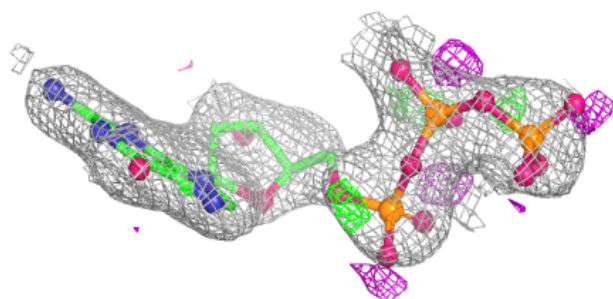
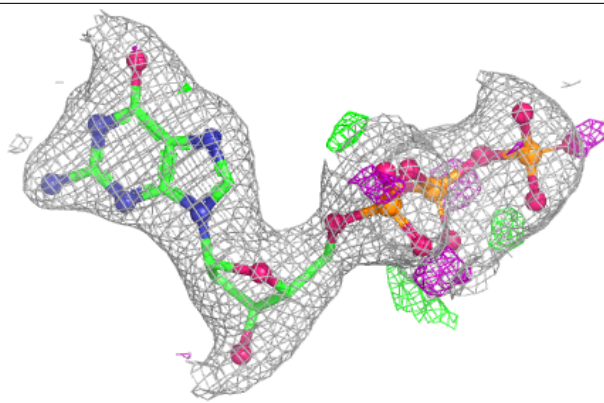
Electron density around CA D 902:

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

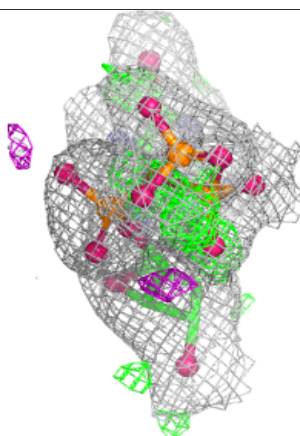
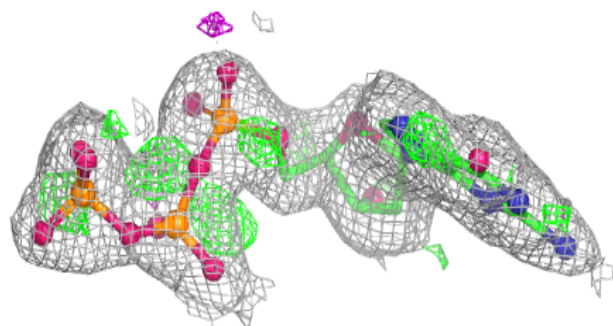
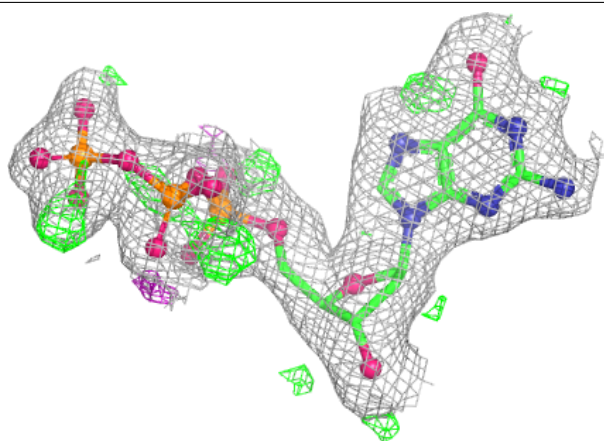


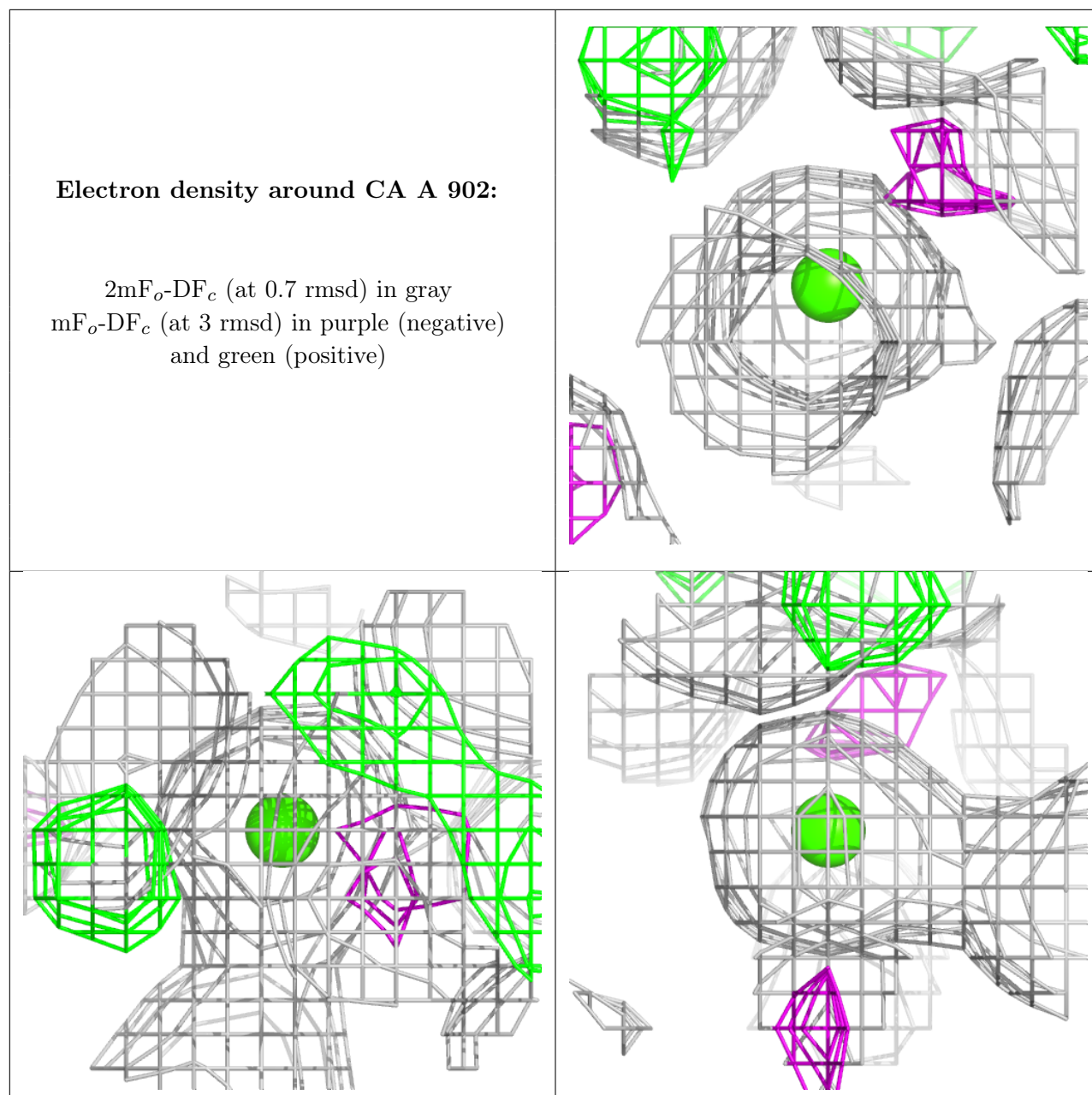
Electron density around DGT D 901:

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

**Electron density around DGT A 901:**

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





6.5 Other polymers ⓘ

There are no such residues in this entry.