Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2023 – 12:18 PM EDT

PDB ID	:	8SCL
Title	:	Bst DNA polymerase I Large Fragment mutant F710Y/D598A with 3'-amino primer, dGTP, and calcium time-resolved 6h
Authors	:	Fang, Z.; Lelyveld, V.S.; Szostak, J.W.
Deposited on	:	2023-04-05
Resolution	:	2.44 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:

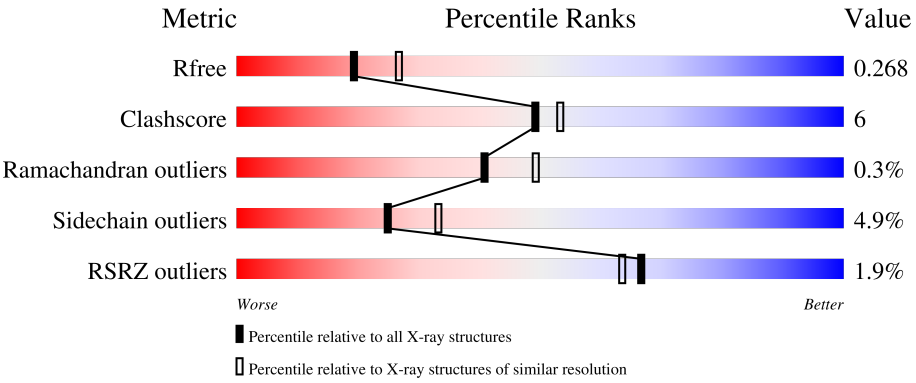
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.44 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.



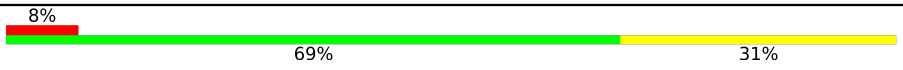
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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 <=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	87% 12% .
1	D	579	4% 82% 17% .
2	B	10	70% 20% 10%
2	E	10	60% 30% 10%
3	C	13	8% 77% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	13	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 10554 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

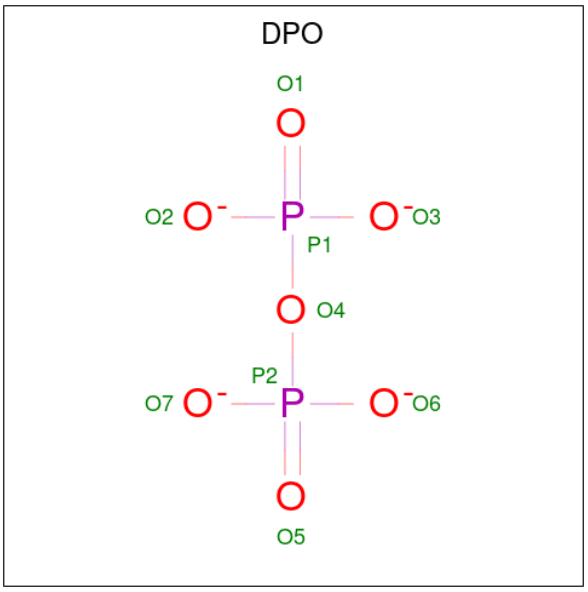
- Molecule 2 is a DNA chain called DNA primer/product.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
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.

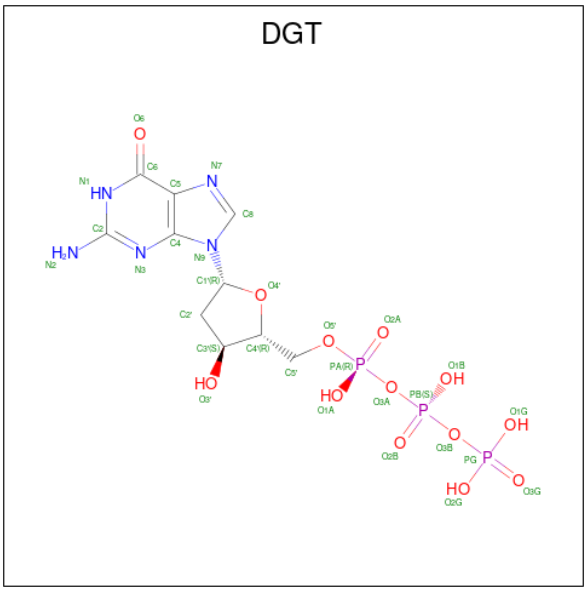
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	P	0	0	1
			246	116	46	72	12			
3	F	13	Total	C	N	O	P	0	0	1
			246	116	46	72	12			

- Molecule 4 is DIPHOSPHATE (three-letter code: DPO) (formula: O_7P_2) (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...

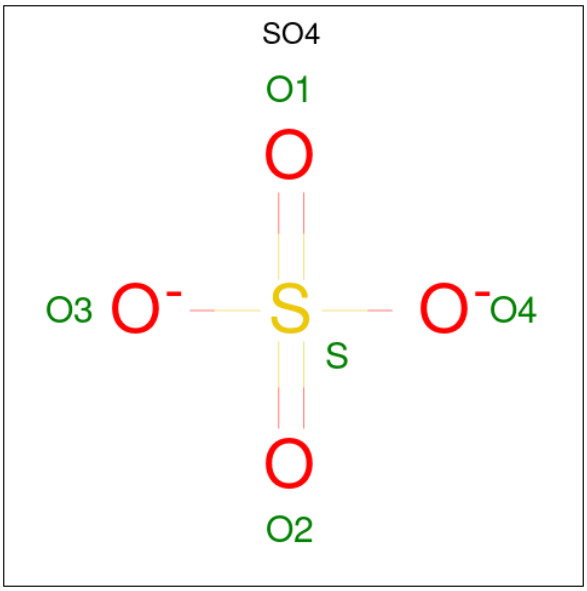
Continued from previous page...

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

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

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

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

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

- Molecule 8 is water.

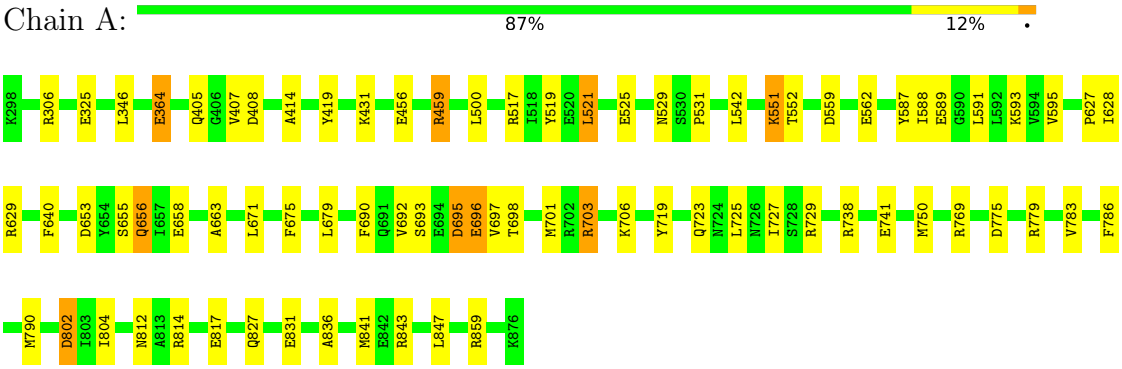
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	129	Total O 129 129	0	0
8	B	4	Total O 4 4	0	0
8	C	9	Total O 9 9	0	0
8	D	53	Total O 53 53	0	0
8	E	4	Total O 4 4	0	0
8	F	8	Total O 8 8	0	0

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.

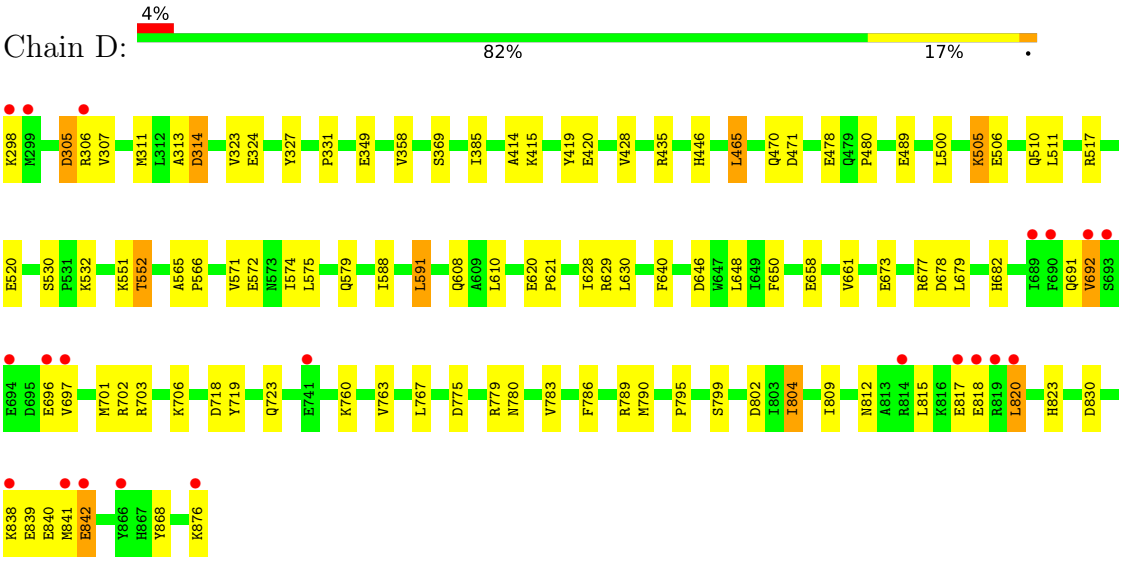
• Molecule 1: DNA polymerase I

Chain A:



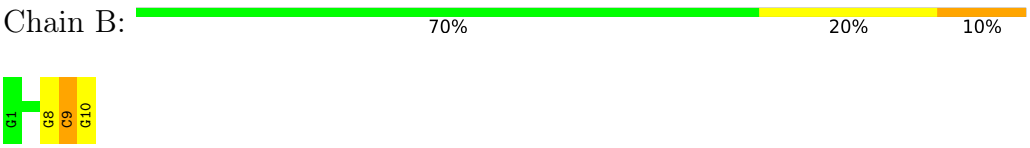
• Molecule 1: DNA polymerase I

Chain D:



• Molecule 2: DNA primer/product

Chain B:

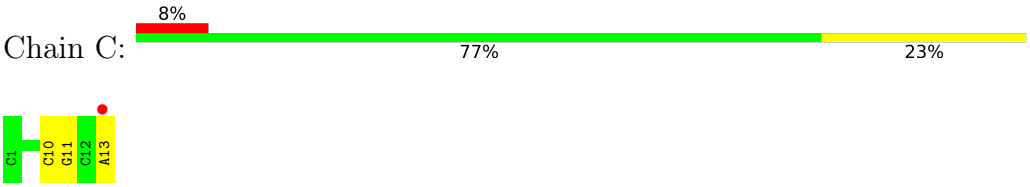


• Molecule 2: DNA primer/product

Chain E:

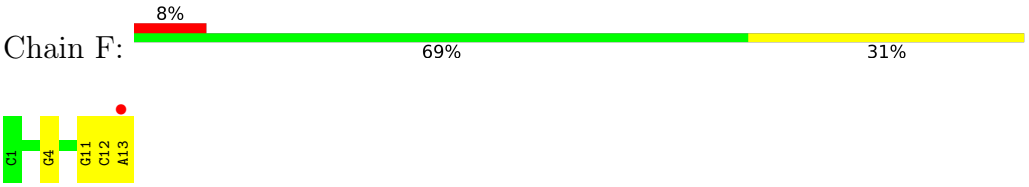
• Molecule 3: DNA template

Chain C:



• Molecule 3: DNA template

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	93.60Å 109.74Å 150.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.98 – 2.44 40.98 – 2.44	Depositor EDS
% Data completeness (in resolution range)	92.0 (40.98-2.44) 92.0 (40.98-2.44)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.19 (at 2.45Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.196 , 0.268 0.196 , 0.268	Depositor DCC
R_{free} test set	2717 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	35.0	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10554	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DGT, DPO, SO4, 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	0.47	0/4730	0.61	1/6393 (0.0%)
1	D	0.38	0/4730	0.54	0/6393
2	B	1.00	0/207	1.10	0/316
2	E	0.92	0/207	0.99	0/316
3	C	1.25	0/275	1.05	0/423
3	F	1.05	0/275	0.97	0/423
All	All	0.52	0/10424	0.64	1/14264 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	802	ASP	CB-CG-OD1	5.64	123.38	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4646	0	4705	45	0
1	D	4646	0	4705	66	0
2	B	223	0	124	5	0
2	E	223	0	125	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	246	0	135	3	0
3	F	246	0	135	8	0
4	A	9	0	0	0	0
4	D	9	0	0	0	0
5	A	31	0	7	3	0
5	D	31	0	3	3	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	A	35	0	0	0	0
8	A	129	0	0	6	1
8	B	4	0	0	0	0
8	C	9	0	0	0	0
8	D	53	0	0	1	0
8	E	4	0	0	0	0
8	F	8	0	0	0	0
All	All	10554	0	9939	119	1

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 6.

All (119) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:414:ALA:HB1	1:D:419:TYR:HB3	1.60	0.83
1:A:859:ARG:HD3	8:A:1081:HOH:O	1.86	0.75
1:A:698:THR:HG23	1:A:701:MET:H	1.51	0.74
5:A:902[B]:DGT:O3'	8:A:1001:HOH:O	2.02	0.73
1:A:698:THR:H	1:A:701:MET:HE3	1.53	0.72
1:D:323:VAL:O	1:D:435:ARG:NH2	2.23	0.72
1:A:431:LYS:HD3	1:D:551:LYS:HE3	1.74	0.68
1:D:465:LEU:HD12	1:D:470:GLN:HB2	1.76	0.67
1:D:517:ARG:NH1	1:D:520:GLU:OE1	2.26	0.67
1:D:532:LYS:HE2	3:F:12:DC:H5"	1.77	0.66
1:D:677:ARG:HB2	1:D:679:LEU:HD13	1.79	0.65
1:D:478:GLU:OE1	1:D:608:GLN:NE2	2.31	0.64
1:A:459:ARG:NH2	8:A:1007:HOH:O	2.30	0.63
1:D:673:GLU:HG3	1:D:677:ARG:HD2	1.82	0.61
1:D:532:LYS:HA	3:F:13:DA:C6	2.37	0.59
1:D:692:VAL:HG11	1:D:701:MET:HE1	1.84	0.59
1:D:789:ARG:NH1	3:F:4:DG:OP1	2.37	0.58
1:D:532:LYS:HG2	3:F:13:DA:C4	2.39	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:629:ARG:NH2	2:E:9[B]:C42:O1P	2.38	0.57
1:D:830:ASP:OD1	2:E:9[B]:C42:N	2.38	0.57
1:A:658:GLU:CD	2:B:10[A]:DG:H2"	2.25	0.57
5:D:902[B]:DGT:O4'	2:E:9[B]:C42:H2'1	2.05	0.56
1:D:532:LYS:HD3	3:F:13:DA:H2'	1.86	0.56
1:A:663:ALA:HB2	1:A:671:LEU:HG	1.88	0.55
1:D:552:THR:OG1	3:F:13:DA:N3	2.39	0.55
1:D:358:VAL:HG13	1:D:385:ILE:HG23	1.89	0.55
1:D:719:TYR:O	1:D:723:GLN:HG2	2.07	0.55
1:D:706:LYS:NZ	5:D:902[B]:DGT:O1A	2.40	0.54
1:D:775:ASP:HB3	1:D:783:VAL:HG12	1.90	0.53
1:D:571:VAL:HA	1:D:574:ILE:HD12	1.91	0.53
1:D:314:ASP:OD1	1:D:314:ASP:N	2.38	0.53
1:A:552:THR:HB	3:C:13:DA:N3	2.24	0.53
1:D:648:LEU:HD23	1:D:838:LYS:HA	1.91	0.52
1:A:408:ASP:OD1	1:A:408:ASP:N	2.38	0.52
1:A:653:ASP:HB3	1:A:831:GLU:HG3	1.92	0.52
1:D:305:ASP:N	1:D:305:ASP:OD1	2.40	0.51
3:C:10:DC:H2"	3:C:11:DG:C8	2.45	0.51
1:D:629:ARG:HH22	2:E:9[A]:C42:H5	1.74	0.51
1:D:780:ASN:OD1	1:D:783:VAL:HG23	2.11	0.51
1:A:725:LEU:HB2	1:A:727:ILE:HG12	1.92	0.51
1:D:307:VAL:HG13	1:D:311:MET:HE2	1.92	0.51
1:D:804:ILE:HD13	1:D:830:ASP:HA	1.92	0.50
1:A:750:MET:HG2	8:A:1118:HOH:O	2.09	0.50
1:A:658:GLU:OE2	5:A:902[B]:DGT:O3'	2.30	0.50
2:E:9[A]:C42:H2'1	2:E:10[A]:DG:C8	2.47	0.50
1:A:693:SER:OG	1:A:696:GLU:OE2	2.29	0.49
2:E:9[A]:C42:H2'1	2:E:10[A]:DG:H8	1.77	0.49
1:A:786:PHE:O	1:A:790:MET:HG3	2.11	0.49
1:D:658:GLU:CD	2:E:10[A]:DG:H2"	2.33	0.49
1:A:459:ARG:HA	1:A:459:ARG:HE	1.78	0.48
1:A:629:ARG:NH1	2:B:8:DG:OP2	2.46	0.48
1:A:719:TYR:O	1:A:723:GLN:HG2	2.13	0.48
1:A:364:GLU:HG3	8:A:1092:HOH:O	2.13	0.48
1:D:575:LEU:O	1:D:579:GLN:HG3	2.13	0.48
1:D:530:SER:HA	3:F:11:DG:H4'	1.96	0.47
1:A:629:ARG:HD2	1:A:703:ARG:HD2	1.96	0.47
1:A:843:ARG:NH1	8:A:1004:HOH:O	2.22	0.47
1:A:500:LEU:HD22	1:A:588:ILE:HD13	1.96	0.47
1:D:818:GLU:HB2	1:D:820:LEU:HD22	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:695:ASP:OD1	1:A:695:ASP:N	2.39	0.47
1:D:629:ARG:NH2	2:E:9[A]:C42:H5	2.30	0.47
1:D:697:VAL:HA	1:D:701:MET:HE2	1.97	0.47
1:D:506:GLU:O	1:D:510:GLN:HG3	2.14	0.47
1:D:786:PHE:O	1:D:790:MET:HG3	2.14	0.46
1:A:814:ARG:HH11	1:A:847:LEU:HD11	1.81	0.46
1:D:650:PHE:HD1	1:D:868:TYR:HB3	1.81	0.46
2:B:8:DG:H2'	2:B:9[B]:C42:H6	1.97	0.45
1:D:313:ALA:N	8:D:1009:HOH:O	2.48	0.45
1:A:405:GLN:HB3	1:A:407:VAL:HG23	1.98	0.45
1:A:656:GLN:HG3	1:A:675:PHE:CE1	2.52	0.45
1:A:703:ARG:HH21	1:A:706:LYS:NZ	2.14	0.45
5:A:902[B]:DGT:O4'	2:B:9[B]:C42:H2'1	2.15	0.45
1:D:815:LEU:HD23	1:D:820:LEU:HD23	1.98	0.45
1:A:589:GLU:O	1:A:593:LYS:HG3	2.17	0.45
1:A:414:ALA:HB1	1:A:419:TYR:HB3	1.99	0.44
1:D:500:LEU:HD22	1:D:588:ILE:HD13	2.00	0.44
2:B:8:DG:H2'	2:B:9[A]:C42:H6	2.00	0.44
1:A:690:PHE:O	1:A:692:VAL:HG13	2.18	0.44
1:D:620:GLU:HA	1:D:621:PRO:HA	1.75	0.44
1:D:428:VAL:HG21	1:D:446:HIS:CG	2.53	0.44
1:D:682:HIS:CD2	1:D:706:LYS:HG2	2.52	0.43
1:D:480:PRO:HB2	1:D:809:ILE:HD13	2.00	0.43
1:D:511:LEU:HA	1:D:511:LEU:HD23	1.75	0.43
1:D:767:LEU:HD12	1:D:802:ASP:HB3	2.00	0.43
1:D:505:LYS:HA	1:D:505:LYS:HD3	1.85	0.43
1:A:551:LYS:H	1:A:551:LYS:HG2	1.53	0.43
1:A:559:ASP:OD1	1:A:559:ASP:N	2.52	0.43
1:A:519:TYR:CD1	1:A:525:GLU:HA	2.53	0.43
1:D:500:LEU:HD23	1:D:500:LEU:HA	1.76	0.43
2:E:1:DG:H1'	2:E:2:DC:H5'	2.01	0.43
1:A:655:SER:O	1:A:656:GLN:HG2	2.19	0.42
1:A:836:ALA:HB3	1:A:841:MET:CE	2.49	0.42
1:A:591:LEU:HD21	1:A:640:PHE:CZ	2.54	0.42
1:D:415:LYS:HD3	1:D:415:LYS:HA	1.74	0.42
1:D:591:LEU:HD21	1:D:640:PHE:HZ	1.83	0.42
1:A:769:ARG:NH2	1:A:802:ASP:OD1	2.49	0.42
1:D:349:GLU:H	1:D:349:GLU:CD	2.23	0.42
1:D:840:GLU:O	1:D:842:GLU:N	2.53	0.42
1:D:532:LYS:NZ	3:F:13:DA:OP1	2.53	0.42
1:A:591:LEU:O	1:A:595:VAL:HG23	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:775:ASP:HB3	1:A:783:VAL:HG12	2.02	0.42
1:D:591:LEU:HD21	1:D:640:PHE:CZ	2.55	0.42
1:D:415:LYS:NZ	1:D:420:GLU:HG2	2.34	0.41
1:A:738:ARG:O	1:A:741:GLU:HB2	2.20	0.41
1:D:565:ALA:N	1:D:566:PRO:HD2	2.35	0.41
1:A:517:ARG:O	1:A:521:LEU:HD23	2.20	0.41
1:D:428:VAL:HG21	1:D:446:HIS:ND1	2.35	0.41
1:D:840:GLU:C	1:D:842:GLU:H	2.23	0.41
1:A:671:LEU:HD12	1:A:671:LEU:HA	1.83	0.41
1:D:324:GLU:OE1	1:D:331:PRO:HD2	2.20	0.41
1:D:327:TYR:OH	1:D:489:GLU:OE1	2.21	0.41
1:A:531:PRO:O	3:C:13:DA:N6	2.54	0.41
1:A:697:VAL:HA	1:A:701:MET:HE3	2.01	0.41
1:D:718:ASP:OD1	1:D:719:TYR:N	2.54	0.41
1:D:763:VAL:HB	1:D:795:PRO:HG3	2.02	0.41
1:A:587:TYR:CE1	1:A:627:PRO:HD3	2.56	0.40
1:D:702:ARG:NH1	5:D:902[B]:DGT:O1G	2.46	0.40
1:D:692:VAL:HG22	1:D:696:GLU:CD	2.42	0.40
1:D:661:VAL:HG11	1:D:799:SER:HB2	2.04	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:A:1093:HOH:O	8:A:1113:HOH:O[4_545]	2.13	0.07

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%)	556 (96%)	20 (4%)	1 (0%)	47 57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	577/579 (100%)	551 (96%)	23 (4%)	3 (0%)	29 34
All	All	1154/1158 (100%)	1107 (96%)	43 (4%)	4 (0%)	41 49

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	841	MET
1	D	628	ILE
1	D	691	GLN
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	495/495 (100%)	473 (96%)	22 (4%)	28 37
1	D	495/495 (100%)	468 (94%)	27 (6%)	21 28
All	All	990/990 (100%)	941 (95%)	49 (5%)	25 33

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	306	ARG
1	A	325	GLU
1	A	346	LEU
1	A	364	GLU
1	A	456	GLU
1	A	459	ARG
1	A	521	LEU
1	A	529	ASN
1	A	542	LEU
1	A	551	LYS
1	A	562	GLU
1	A	656	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	679	LEU
1	A	695	ASP
1	A	696	GLU
1	A	703	ARG
1	A	729	ARG
1	A	779	ARG
1	A	804	ILE
1	A	812	ASN
1	A	817	GLU
1	A	827	GLN
1	D	298	LYS
1	D	305	ASP
1	D	306	ARG
1	D	314	ASP
1	D	369	SER
1	D	465	LEU
1	D	471	ASP
1	D	505	LYS
1	D	552	THR
1	D	572	GLU
1	D	591	LEU
1	D	610	LEU
1	D	630	LEU
1	D	646	ASP
1	D	678	ASP
1	D	692	VAL
1	D	703	ARG
1	D	760	LYS
1	D	779	ARG
1	D	804	ILE
1	D	812	ASN
1	D	817	GLU
1	D	820	LEU
1	D	823	HIS
1	D	839	GLU
1	D	842	GLU
1	D	876	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	759	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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[A]	3,2	18,20,21	3.37	9 (50%)	18,28,31	1.36	3 (16%)
2	C42	B	9[B]	3,2	18,20,21	3.33	8 (44%)	18,28,31	1.08	2 (11%)
2	C42	E	9[B]	3,2	18,20,21	3.31	8 (44%)	18,28,31	1.53	3 (16%)
2	C42	B	9[A]	3,2	18,20,21	3.23	7 (38%)	18,28,31	1.01	1 (5%)

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[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
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

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	9[B]	C42	O4'-C1'	7.79	1.59	1.42
2	E	9[A]	C42	O4'-C1'	7.34	1.58	1.42
2	B	9[A]	C42	O4'-C1'	7.16	1.58	1.42
2	B	9[B]	C42	O4'-C1'	7.13	1.58	1.42
2	E	9[A]	C42	O4'-C4'	-6.38	1.30	1.45
2	B	9[B]	C42	O4'-C4'	-6.36	1.30	1.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	9[A]	C42	O4'-C4'	-6.20	1.31	1.45
2	E	9[B]	C42	O4'-C4'	-6.13	1.31	1.45
2	E	9[B]	C42	C4-N4	5.10	1.45	1.33
2	B	9[A]	C42	C4-N4	5.06	1.45	1.33
2	E	9[A]	C42	C1'-N1	-5.03	1.35	1.48
2	B	9[B]	C42	C1'-N1	-5.02	1.35	1.48
2	E	9[A]	C42	C4-N4	4.92	1.45	1.33
2	B	9[B]	C42	C4-N4	4.85	1.45	1.33
2	B	9[A]	C42	C1'-N1	-4.67	1.36	1.48
2	E	9[B]	C42	C1'-N1	-4.47	1.36	1.48
2	E	9[A]	C42	C4'-C3'	4.47	1.62	1.52
2	B	9[B]	C42	C4'-C3'	4.42	1.62	1.52
2	E	9[B]	C42	C4'-C3'	4.24	1.62	1.52
2	B	9[A]	C42	C4'-C3'	3.99	1.61	1.52
2	B	9[B]	C42	C2-N1	-3.16	1.33	1.40
2	B	9[A]	C42	C2-N1	-3.12	1.33	1.40
2	E	9[A]	C42	C2-N1	-2.89	1.33	1.40
2	E	9[A]	C42	O2-C2	-2.83	1.18	1.23
2	E	9[B]	C42	O2-C2	-2.76	1.18	1.23
2	B	9[B]	C42	O2-C2	-2.67	1.18	1.23
2	B	9[A]	C42	O2-C2	-2.50	1.19	1.23
2	E	9[B]	C42	C2-N1	-2.47	1.34	1.40
2	E	9[A]	C42	C6-N1	-2.39	1.32	1.38
2	B	9[B]	C42	C6-N1	-2.25	1.32	1.38
2	E	9[B]	C42	C6-N1	-2.22	1.32	1.38
2	E	9[A]	C42	C2-N3	-2.04	1.32	1.36

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	9[B]	C42	O4'-C1'-N1	3.54	114.19	107.86
2	E	9[B]	C42	O2-C2-N3	-2.95	117.53	122.33
2	E	9[A]	C42	O4'-C1'-N1	2.88	113.02	107.86
2	E	9[A]	C42	O2-C2-N3	-2.51	118.25	122.33
2	E	9[A]	C42	C2'-C1'-N1	-2.46	108.10	113.77
2	B	9[B]	C42	O2-C2-N3	-2.22	118.72	122.33
2	B	9[A]	C42	O4'-C1'-N1	2.20	111.79	107.86
2	E	9[B]	C42	C1'-N1-C2	2.02	121.27	117.74
2	B	9[B]	C42	O4'-C1'-N1	2.01	111.46	107.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	9[A]	C42	4	0
2	B	9[B]	C42	2	0
2	E	9[B]	C42	3	0
2	B	9[A]	C42	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 2 are monoatomic - leaving 11 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	DPO	A	901[A]	6	6,8,8	0.68	0	13,13,13	1.46	2 (15%)
7	SO4	A	909	-	4,4,4	0.21	0	6,6,6	0.08	0
7	SO4	A	910	-	4,4,4	0.17	0	6,6,6	0.14	0
5	DGT	D	902[B]	6	26,33,33	3.49	8 (30%)	32,52,52	1.35	4 (12%)
7	SO4	A	907	-	4,4,4	0.13	0	6,6,6	0.40	0
7	SO4	A	906	-	4,4,4	0.21	0	6,6,6	0.28	0
7	SO4	A	908	-	4,4,4	0.12	0	6,6,6	0.28	0
5	DGT	A	902[B]	6	26,33,33	3.57	10 (38%)	32,52,52	1.35	5 (15%)
7	SO4	A	905	-	4,4,4	0.17	0	6,6,6	0.12	0
7	SO4	A	904	-	4,4,4	0.41	0	6,6,6	0.38	0
4	DPO	D	901[A]	6	6,8,8	0.66	0	13,13,13	1.24	1 (7%)

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

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	902[B]	DGT	O6-C6	8.40	1.40	1.23
5	D	902[B]	DGT	O4'-C1'	8.19	1.60	1.42
5	A	902[B]	DGT	O4'-C4'	-8.12	1.26	1.45
5	A	902[B]	DGT	O4'-C1'	7.98	1.60	1.42
5	D	902[B]	DGT	O6-C6	7.97	1.39	1.23
5	D	902[B]	DGT	O4'-C4'	-7.67	1.27	1.45
5	A	902[B]	DGT	C2'-C1'	-6.96	1.32	1.52
5	D	902[B]	DGT	C2'-C1'	-6.94	1.33	1.52
5	D	902[B]	DGT	C2-N2	4.85	1.45	1.34
5	A	902[B]	DGT	C2-N2	4.77	1.45	1.34
5	A	902[B]	DGT	C6-N1	-3.68	1.32	1.37
5	D	902[B]	DGT	C5-C6	-3.38	1.40	1.47
5	D	902[B]	DGT	C6-N1	-3.34	1.32	1.37
5	A	902[B]	DGT	C5-C6	-3.18	1.40	1.47
5	A	902[B]	DGT	O3'-C3'	-2.83	1.37	1.43
5	D	902[B]	DGT	O3'-C3'	-2.42	1.38	1.43
5	A	902[B]	DGT	C2-N1	-2.02	1.32	1.37
5	A	902[B]	DGT	C1'-N9	-2.02	1.43	1.49

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	901[A]	DPO	P2-O4-P1	-3.83	119.68	132.83
4	D	901[A]	DPO	P2-O4-P1	-3.15	122.00	132.83
5	D	902[B]	DGT	C5-C6-N1	3.12	119.46	113.95
5	A	902[B]	DGT	PB-O3B-PG	-3.03	122.43	132.83
5	A	902[B]	DGT	PA-O3A-PB	-2.93	122.77	132.83
5	A	902[B]	DGT	C8-N7-C5	2.90	108.52	102.99
5	D	902[B]	DGT	PB-O3B-PG	-2.90	122.89	132.83
5	A	902[B]	DGT	C5-C6-N1	2.77	118.84	113.95
5	D	902[B]	DGT	C8-N7-C5	2.76	108.24	102.99
5	D	902[B]	DGT	PA-O3A-PB	-2.75	123.40	132.83
4	A	901[A]	DPO	O7-P2-O4	2.45	112.86	104.64
5	A	902[B]	DGT	C2'-C1'-N9	-2.14	109.34	114.27

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	902[B]	DGT	PB-O3B-PG-O2G
5	A	902[B]	DGT	O4'-C4'-C5'-O5'
5	A	902[B]	DGT	C3'-C4'-C5'-O5'
5	D	902[B]	DGT	PB-O3B-PG-O2G
5	D	902[B]	DGT	O4'-C4'-C5'-O5'
5	D	902[B]	DGT	C3'-C4'-C5'-O5'
5	A	902[B]	DGT	PB-O3A-PA-O5'
5	D	902[B]	DGT	PB-O3A-PA-O5'
5	D	902[B]	DGT	PB-O3B-PG-O3G
5	A	902[B]	DGT	PB-O3B-PG-O1G

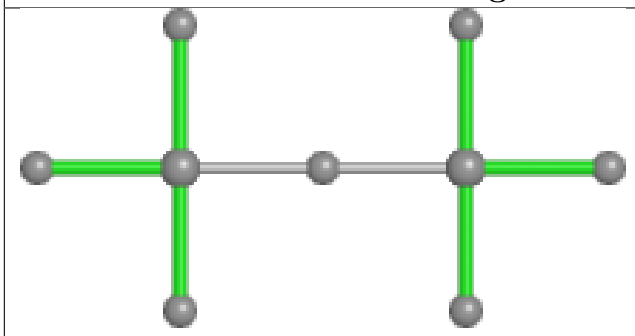
There are no ring outliers.

2 monomers are involved in 6 short contacts:

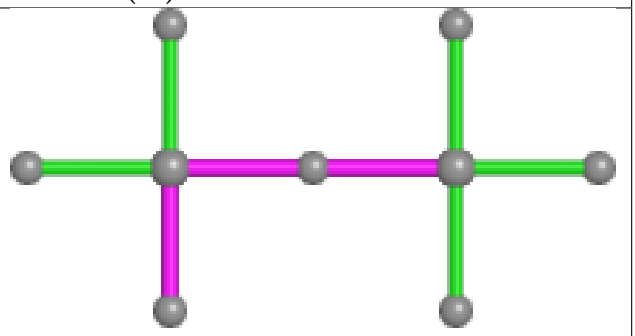
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	902[B]	DGT	3	0
5	A	902[B]	DGT	3	0

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 then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

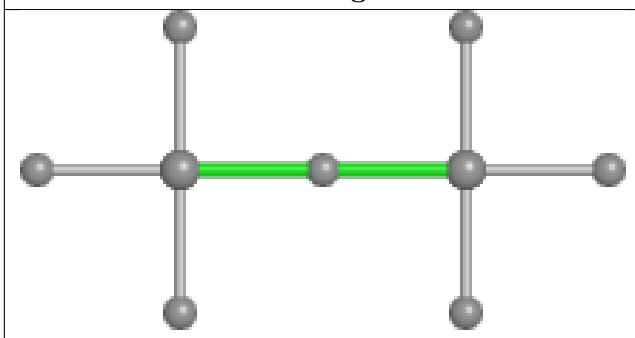
Ligand DPO A 901 (A)



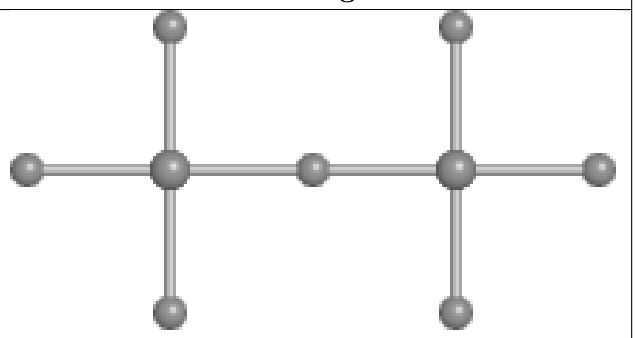
Bond lengths



Bond angles

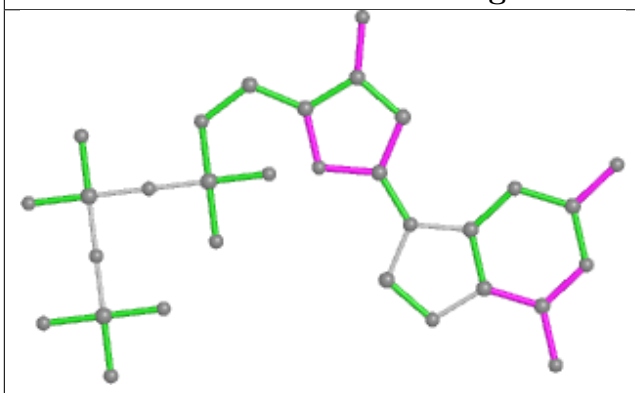


Torsions

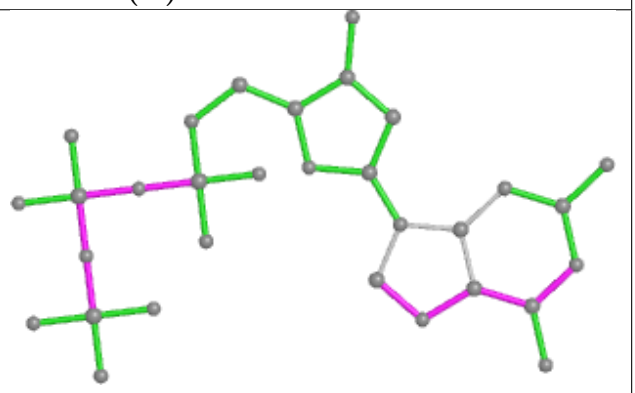


Rings

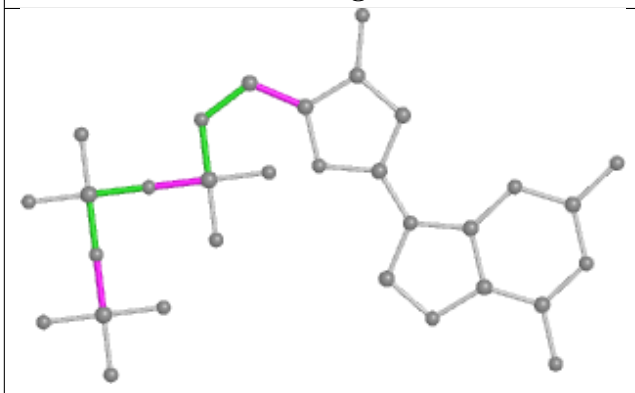
Ligand DGT D 902 (B)



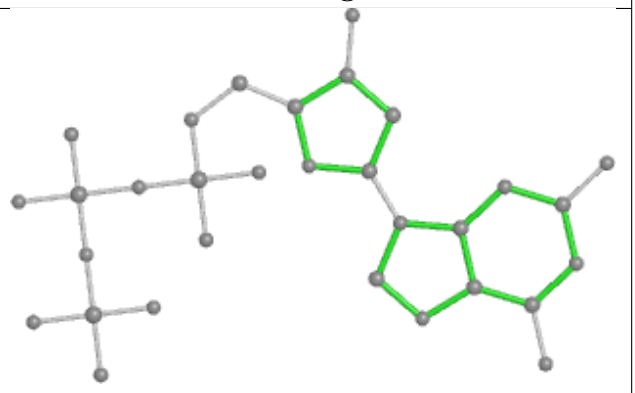
Bond lengths



Bond angles

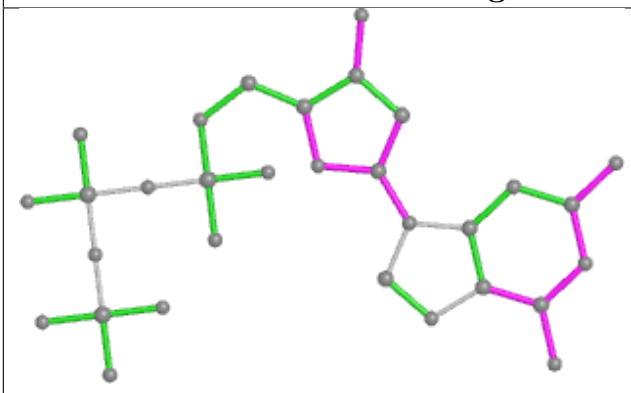


Torsions

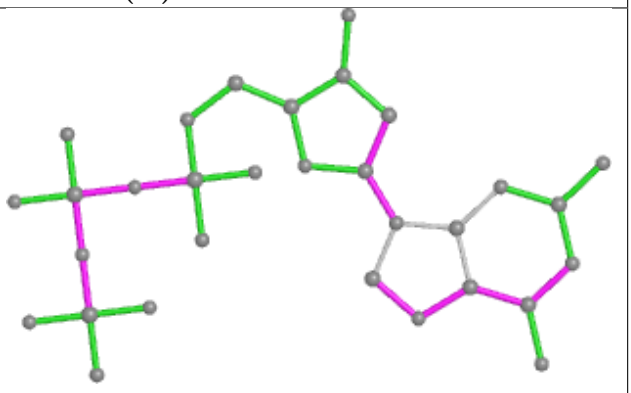


Rings

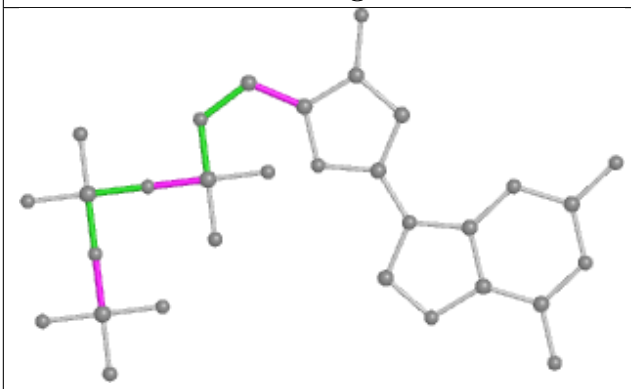
Ligand DGT A 902 (B)



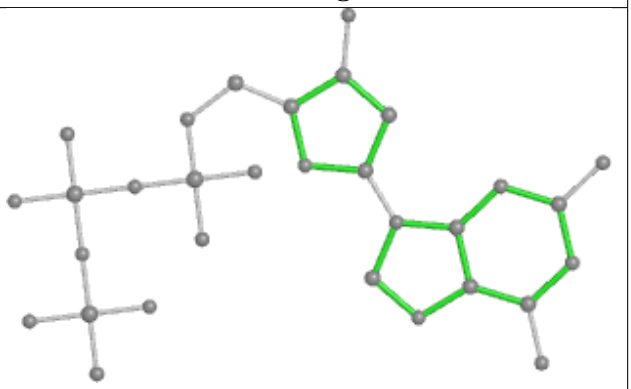
Bond lengths



Bond angles

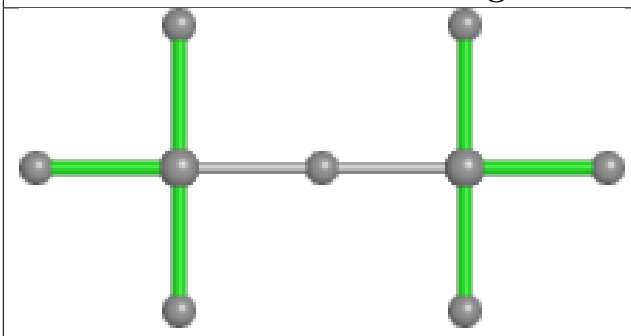


Torsions

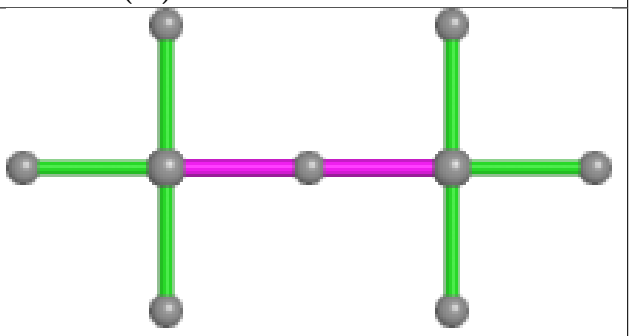


Rings

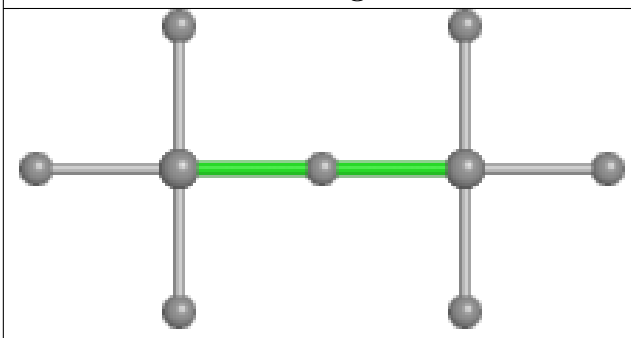
Ligand DPO D 901 (A)



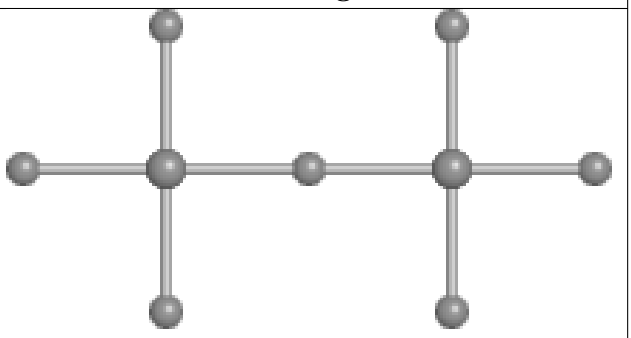
Bond lengths



Bond angles



Torsions



Rings

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^{th} 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%)	-0.32	0 100 100	13, 29, 56, 76	0
1	D	579/579 (100%)	0.07	21 (3%) 42 39	26, 48, 75, 105	0
2	B	9/10 (90%)	-0.29	0 100 100	23, 25, 43, 44	1 (11%)
2	E	9/10 (90%)	-0.27	0 100 100	32, 33, 46, 48	1 (11%)
3	C	13/13 (100%)	-0.29	1 (7%) 13 10	18, 26, 57, 85	0
3	F	13/13 (100%)	-0.36	1 (7%) 13 10	25, 33, 77, 84	0
All	All	1202/1204 (99%)	-0.14	23 (1%) 66 63	13, 39, 69, 105	2 (0%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	692	VAL	4.5
1	D	819	ARG	4.5
1	D	689	ILE	3.5
1	D	842	GLU	3.3
3	C	13	DA	3.3
1	D	838	LYS	3.0
1	D	693	SER	3.0
3	F	13	DA	2.9
1	D	817	GLU	2.9
1	D	694	GLU	2.9
1	D	306	ARG	2.6
1	D	876	LYS	2.5
1	D	741	GLU	2.5
1	D	298	LYS	2.4
1	D	697	VAL	2.4
1	D	820	LEU	2.4
1	D	299	MET	2.4
1	D	814	ARG	2.3
1	D	690	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	818	GLU	2.2
1	D	696	GLU	2.1
1	D	841	MET	2.0
1	D	866	TYR	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^{th} 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^2)	Q<0.9
2	C42	E	9[A]	19/20	0.96	0.17	28,34,43,44	19
2	C42	E	9[B]	19/20	0.96	0.17	27,34,43,44	19
2	C42	B	9[A]	19/20	0.98	0.14	19,21,25,25	19
2	C42	B	9[B]	19/20	0.98	0.14	19,21,25,25	19

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^{th} 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^2)	Q<0.9
7	SO4	A	904	5/5	0.70	0.26	40,55,70,97	0
7	SO4	A	910	5/5	0.83	0.24	65,67,92,104	0
7	SO4	A	906	5/5	0.86	0.16	49,53,70,88	0
4	DPO	D	901[A]	9/9	0.88	0.17	50,55,61,62	9
7	SO4	A	909	5/5	0.89	0.22	57,72,88,92	0
7	SO4	A	908	5/5	0.92	0.26	56,60,74,88	0
4	DPO	A	901[A]	9/9	0.92	0.17	35,42,47,50	9
5	DGT	D	902[B]	31/31	0.92	0.17	30,39,60,62	31
6	CA	D	903	1/1	0.93	0.06	57,57,57,57	0
7	SO4	A	905	5/5	0.93	0.17	60,64,77,94	0

Continued on next page...

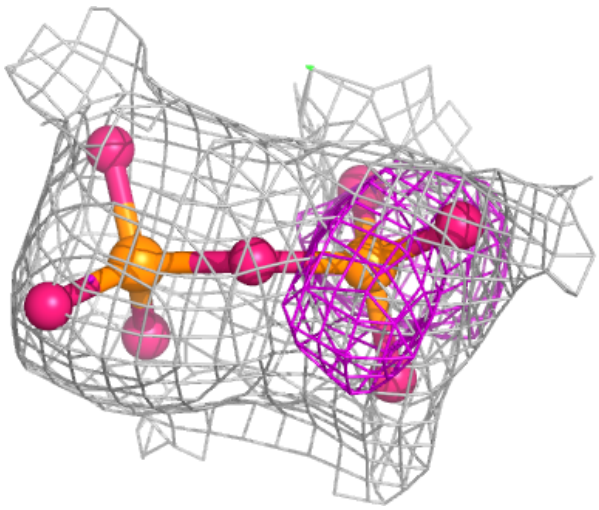
Continued from previous page...

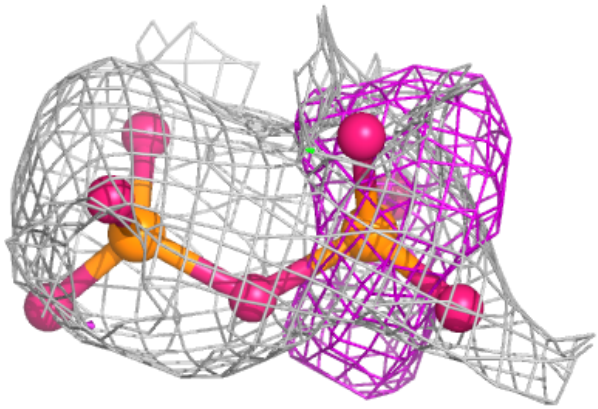
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	DGT	A	902[B]	31/31	0.94	0.17	20,23,44,46	31
7	SO4	A	907	5/5	0.99	0.09	24,37,42,46	0
6	CA	A	903	1/1	0.99	0.05	42,42,42,42	0

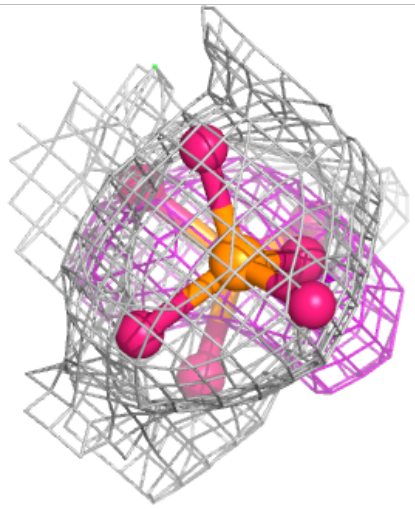
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

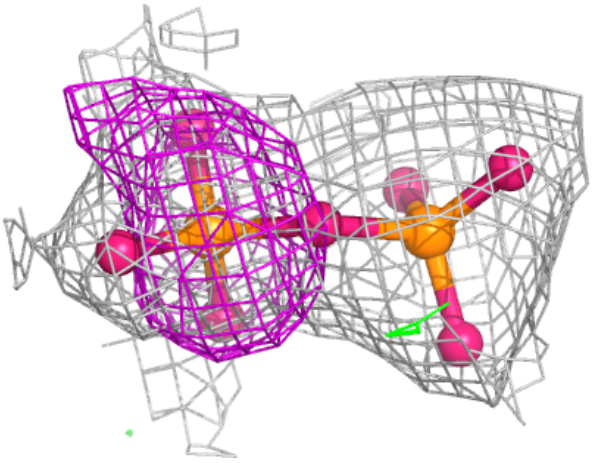


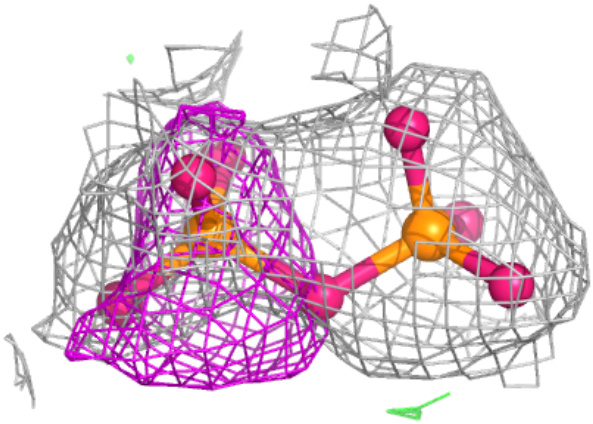


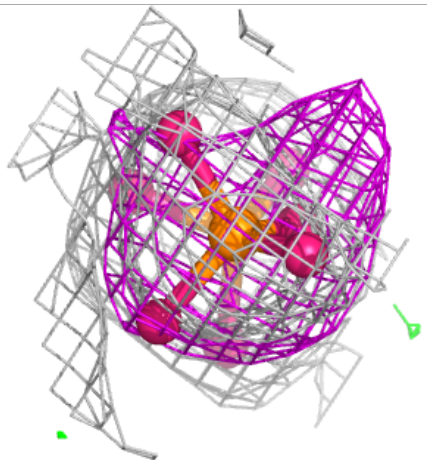


Electron density around DPO A 901 (A):

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

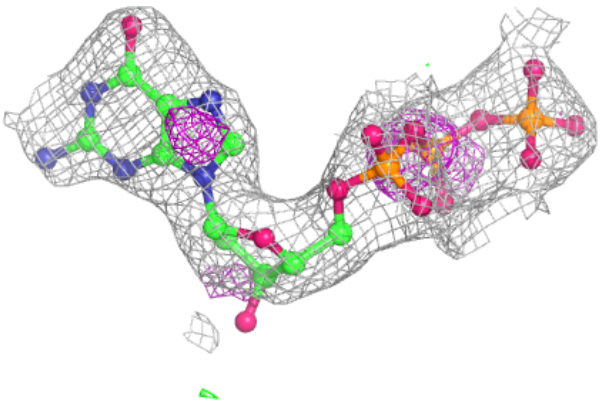


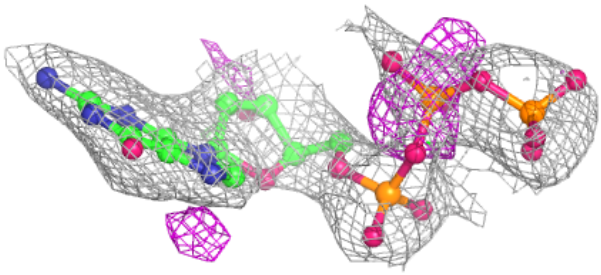


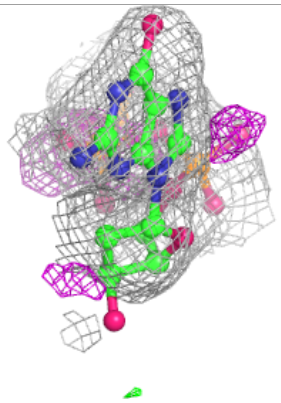


Electron density around DGT D 902 (B):

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

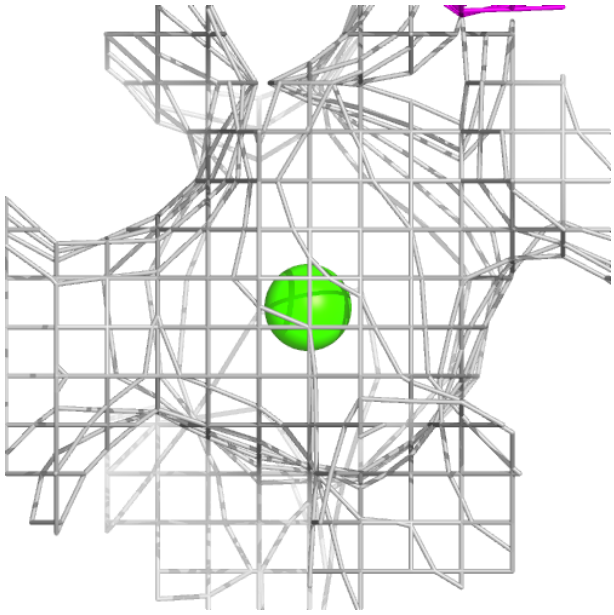


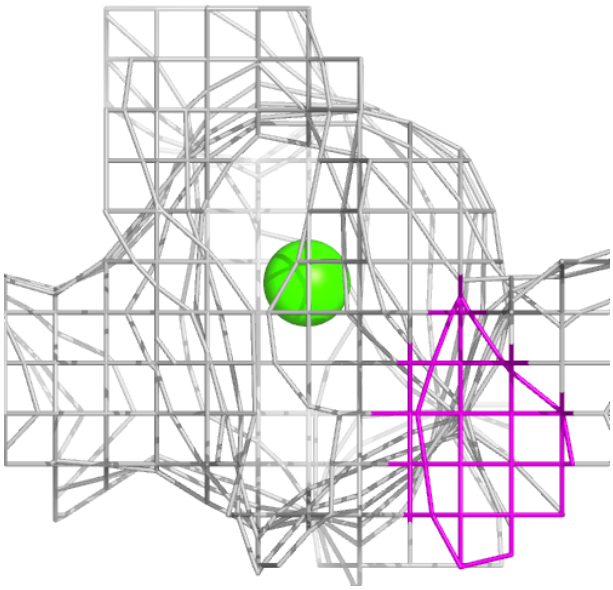


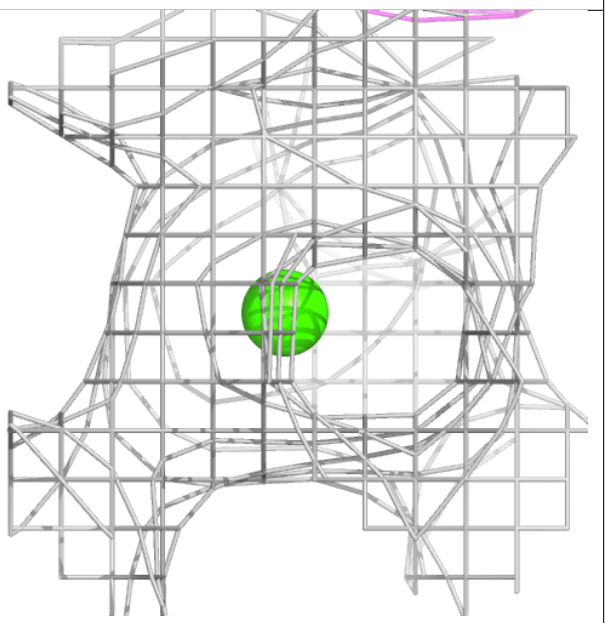


Electron density around CA D 903:

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

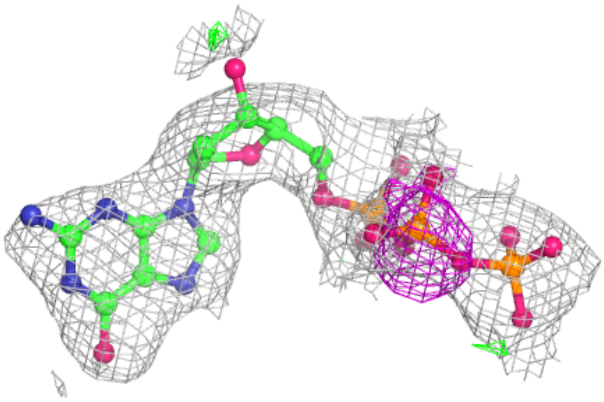


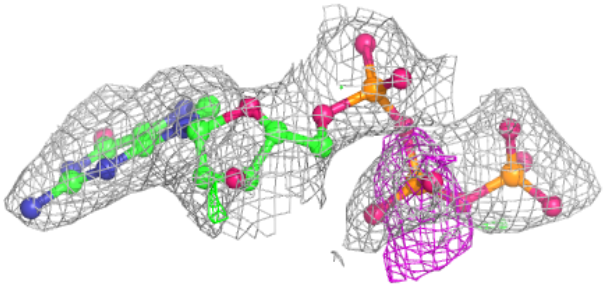


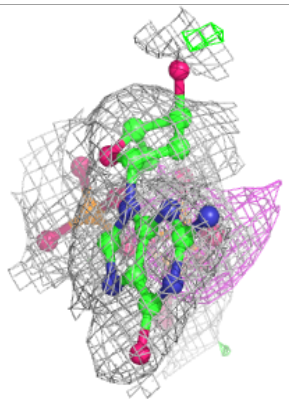


Electron density around DGT A 902 (B):

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

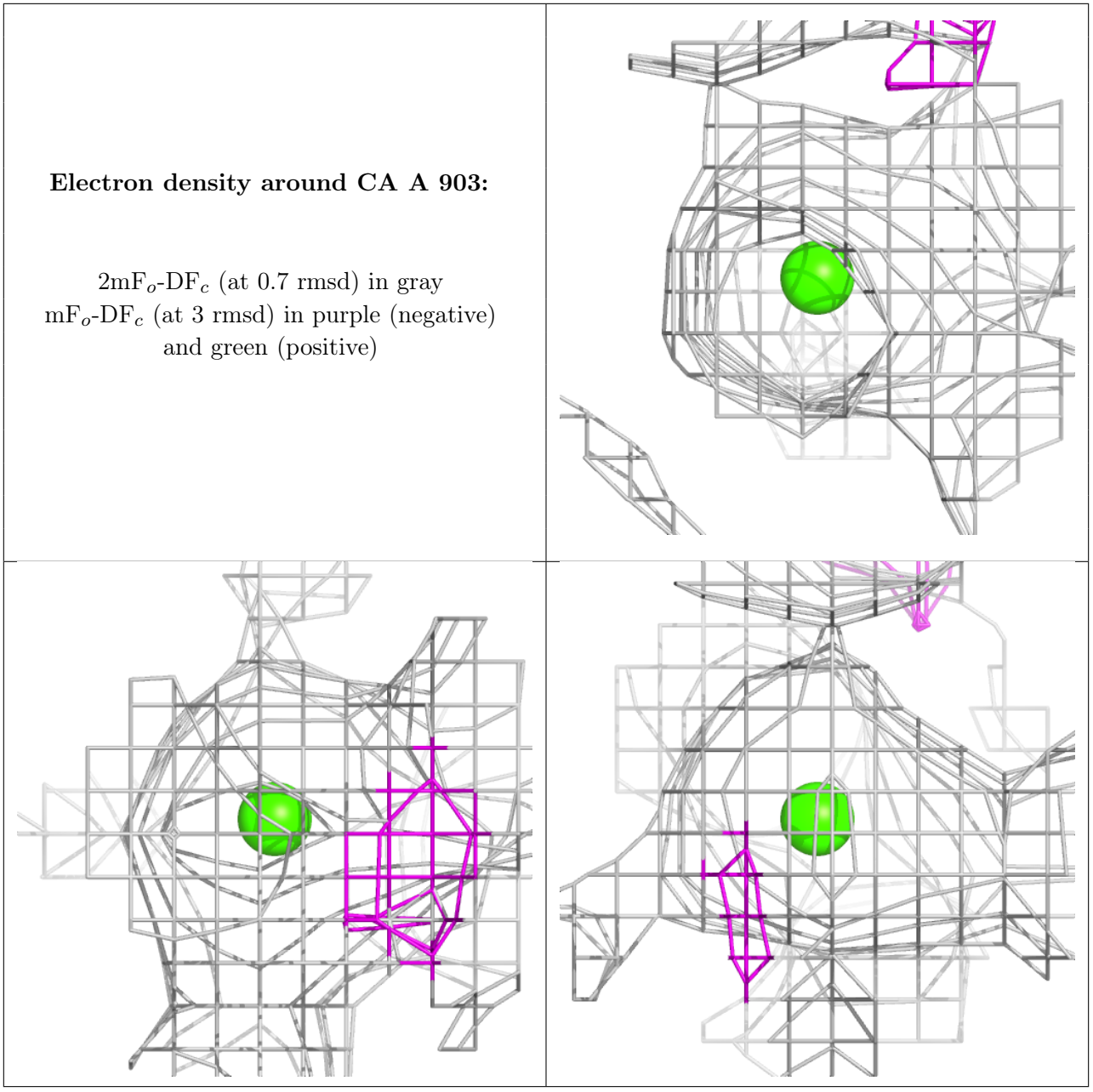






Electron density around CA A 903:

$2mF_o$-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.