Full wwPDB X-ray Structure Validation Report ⓘ

Sep 2, 2023 – 11:10 PM EDT

PDB ID	:	3PX0
Title	:	Crystal Structure of Bacillus DNA Polymerase I Large Fragment Bound to DNA and dCTP-dA Mismatch (tautomer) in Closed Conformation
Authors	:	Wang, W.; Beese, L.S.
Deposited on	:	2010-12-09
Resolution	:	1.73 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:

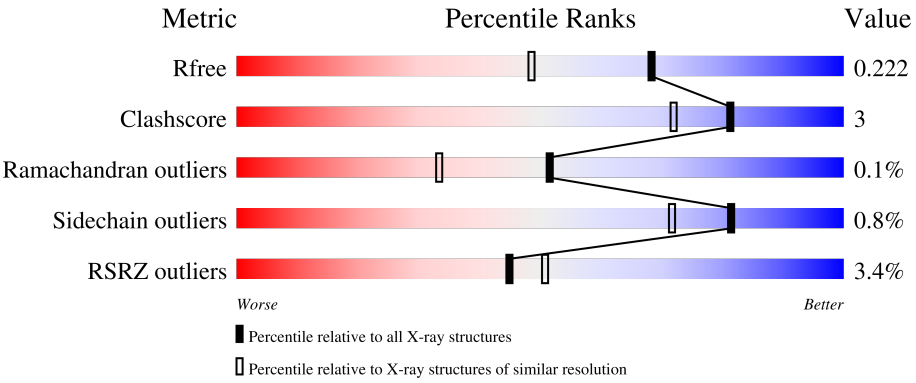
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.73 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	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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	592	6% 88% 7% 6%
1	D	592	% 93% 5% •
2	B	9	56% 44%
2	E	9	44% 56%
3	C	13	8% 46% 46% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	13	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 21278 atoms, of which 9802 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	557	Total	C	H	N	O	S	0	3	0
			9048	2855	4563	782	833	15			
1	D	579	Total	C	H	N	O	S	0	5	0
			9408	2968	4739	809	875	17			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	598	ALA	ASP	engineered mutation	UNP Q5KWC1
A	710	TYR	PHE	engineered mutation	UNP Q5KWC1
A	823	HIS	ARG	SEE REMARK 999	UNP Q5KWC1
D	598	ALA	ASP	engineered mutation	UNP Q5KWC1
D	710	TYR	PHE	engineered mutation	UNP Q5KWC1
D	823	HIS	ARG	SEE REMARK 999	UNP Q5KWC1

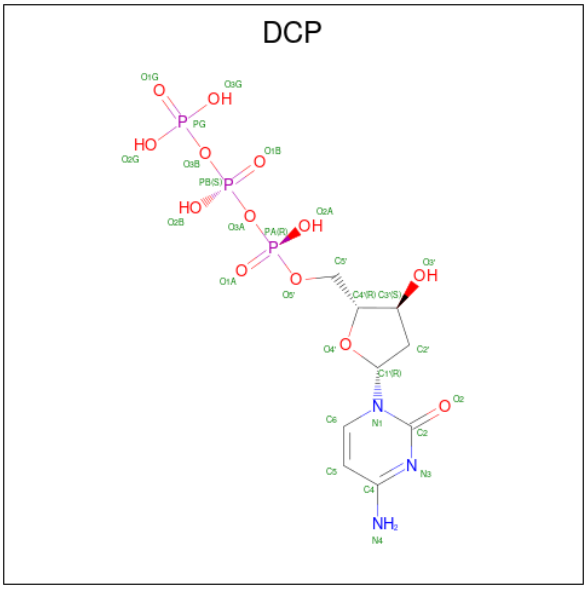
- Molecule 2 is a DNA chain called DNA (5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DOC))-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	9	Total	C	H	N	O	P	0	0	0
			275	85	101	29	52	8			
2	E	9	Total	C	H	N	O	P	0	0	0
			275	85	101	29	52	8			

- Molecule 3 is a DNA chain called DNA (5'-D(*C*AP*TP*AP*GP*GP*AP*GP*TP*CP*AP*GP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	12	Total	C	H	N	O	P	0	0	0
			383	119	133	52	68	11			
3	F	12	Total	C	H	N	O	P	0	0	0
			383	119	133	52	68	11			

- Molecule 4 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: $C_9H_{16}N_3O_{13}P_3$).

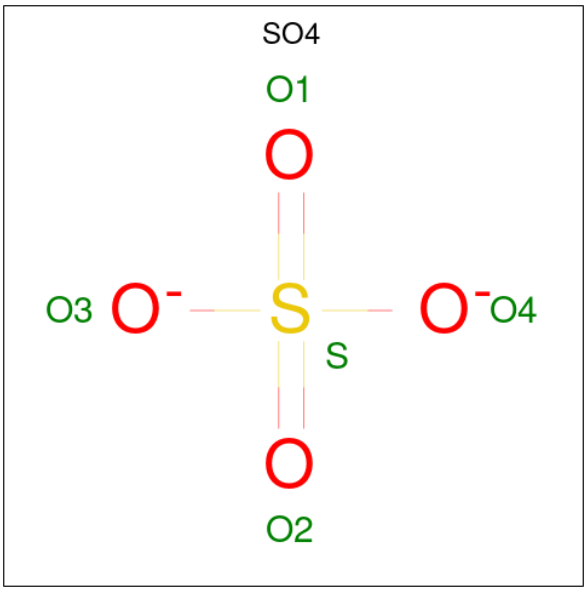


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	P	0	0
			38	9	10	3	13	3		
4	A	1	Total	C	H	N	O	P	0	0
			40	9	12	3	13	3		
4	D	1	Total	C	H	N	O	P	0	0
			38	9	10	3	13	3		

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 7 is water.

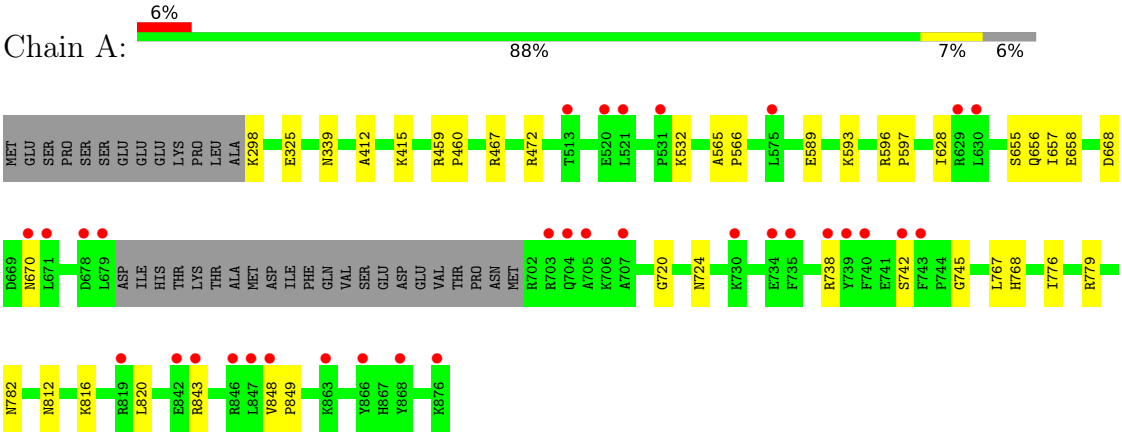
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	474	Total O 474 474	0	0
7	D	714	Total O 714 714	0	0
7	B	25	Total O 25 25	0	0
7	C	56	Total O 56 56	0	0
7	E	31	Total O 31 31	0	0
7	F	68	Total O 68 68	0	0

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.

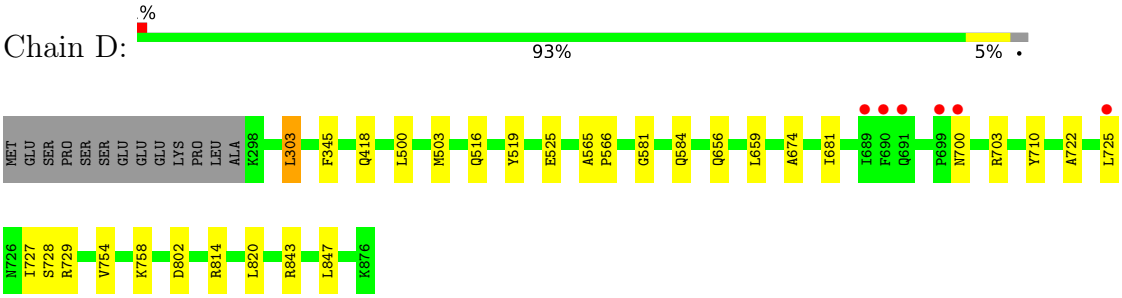
• Molecule 1: DNA polymerase I

Chain A:



• Molecule 1: DNA polymerase I

Chain D:



• Molecule 2: DNA (5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DOC))-3')

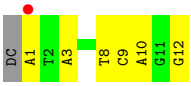
Chain B:

• Molecule 2: DNA (5'-D(*CP*CP*TP*GP*AP*CP*TP*CP*(DOC))-3')

Chain E:

• Molecule 3: DNA (5’-D(*C*AP*TP*AP*GP*GP*AP*GP*TP*CP*AP*GP*G)-3’)

Chain C:



• Molecule 3: DNA (5’-D(*C*AP*TP*AP*GP*GP*AP*GP*TP*CP*AP*GP*G)-3’)

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	93.91Å 108.96Å 150.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.24 – 1.73 64.28 – 1.73	Depositor EDS
% Data completeness (in resolution range)	93.2 (34.24-1.73) 93.2 (64.28-1.73)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 1.73Å)	Xtriage
Refinement program	PHENIX 1.7.1_743, REFMAC 5.5.0109	Depositor
R, R_{free}	0.195 , 0.224 0.194 , 0.222	Depositor DCC
R_{free} test set	6687 reflections (4.47%)	wwPDB-VP
Wilson B-factor (Å^2)	19.6	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 50.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	21278	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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: DOC, DCP, MN, 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	0.38	0/4581	0.50	0/6189
1	D	0.49	0/4771	0.60	1/6447 (0.0%)
2	B	0.81	0/173	1.66	5/264 (1.9%)
2	E	0.90	0/173	1.57	4/264 (1.5%)
3	C	0.88	0/282	1.43	1/435 (0.2%)
3	F	0.91	0/282	1.39	5/435 (1.1%)
All	All	0.49	0/10262	0.70	16/14034 (0.1%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	26	DC	O4'-C1'-N1	8.42	113.89	108.00
3	C	8	DT	O4'-C1'-N1	-8.35	102.15	108.00
2	E	23	DT	O4'-C1'-N1	-7.33	102.87	108.00
3	F	8	DT	N3-C4-O4	6.71	123.93	119.90
2	B	26	DC	O4'-C1'-N1	6.69	112.68	108.00
3	F	8	DT	C5-C4-O4	-6.17	120.58	124.90
2	B	27	DT	O4'-C4'-C3'	-6.10	102.06	104.50
2	B	23	DT	O4'-C1'-N1	-6.09	103.74	108.00
2	B	27	DT	N3-C4-O4	5.88	123.43	119.90
3	F	8	DT	O4'-C1'-N1	-5.87	103.89	108.00
2	E	27	DT	C4'-C3'-C2'	-5.63	98.03	103.10
3	F	2	DT	O4'-C1'-N1	5.62	111.94	108.00
3	F	5	DG	O4'-C4'-C3'	-5.60	102.26	104.50
2	E	27	DT	O4'-C4'-C3'	-5.32	102.37	104.50
1	D	802	ASP	CB-CG-OD2	5.22	123.00	118.30
2	B	28	DC	O4'-C4'-C3'	-5.20	102.42	104.50

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	4485	4563	4542	24	0
1	D	4669	4739	4718	20	0
2	B	174	101	103	0	0
2	E	174	101	103	1	0
3	C	250	133	136	4	0
3	F	250	133	136	1	0
4	A	56	22	24	5	0
4	D	28	10	12	2	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	10	0	0	0	0
6	D	10	0	0	0	0
7	A	474	0	0	6	4
7	B	25	0	0	0	0
7	C	56	0	0	2	0
7	D	714	0	0	6	4
7	E	31	0	0	0	0
7	F	68	0	0	0	0
All	All	11476	9802	9774	51	4

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 (51) 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:584:GLN:HG2	7:D:969:HOH:O	1.78	0.82
4:A:201:DCP:HN41	3:C:3:DA:H61	1.30	0.78
4:D:202:DCP:HN41	3:F:3:DA:H61	1.38	0.71
1:D:581:GLY:HA3	7:D:974:HOH:O	1.99	0.62
1:D:814:ARG:NH1	7:D:1352:HOH:O	2.34	0.61
1:A:532:LYS:HE3	7:A:1274:HOH:O	2.02	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:700:ASN:OD1	1:D:703:ARG:NH2	2.36	0.58
1:A:412:ALA:O	1:A:415:LYS:HG2	2.05	0.56
1:D:754:VAL:CG1	1:D:758:LYS:HE2	2.37	0.55
3:C:12:DG:H1'	7:C:1216:HOH:O	2.06	0.54
1:D:754:VAL:HG12	1:D:758:LYS:HE2	1.90	0.53
1:A:298:LYS:NZ	7:A:1273:HOH:O	2.42	0.53
1:A:656:GLN:HA	4:A:201:DCP:O2B	2.08	0.53
1:A:658:GLU:CD	4:A:201:DCP:H2'2	2.28	0.53
1:A:459:ARG:HB3	1:A:460:PRO:HD3	1.91	0.53
3:C:1:DA:N6	7:C:1086:HOH:O	2.41	0.52
1:A:820:LEU:HD21	1:A:843:ARG:HE	1.75	0.52
1:A:412:ALA:O	1:A:415:LYS:HE2	2.09	0.51
1:A:657:ILE:HG23	1:A:658:GLU:N	2.26	0.51
1:A:848:VAL:HB	1:A:849:PRO:HD3	1.94	0.49
1:D:656:GLN:HA	4:D:202:DCP:O2B	2.14	0.48
1:A:467:ARG:HD3	7:D:927:HOH:O	2.13	0.47
1:A:779:ARG:NH2	7:A:1219:HOH:O	2.46	0.47
1:A:720:GLY:O	1:A:724:ASN:HB2	2.15	0.47
1:A:767:LEU:O	1:A:768:HIS:HB2	2.14	0.47
1:A:325:GLU:H	1:A:325:GLU:CD	2.19	0.46
1:D:814:ARG:HG3	1:D:847:LEU:HD11	1.97	0.46
3:C:9:DC:H2'	3:C:10:DA:C8	2.50	0.46
1:A:655:SER:O	7:A:942:HOH:O	2.20	0.46
1:A:472:ARG:HB2	4:A:203:DCP:C2	2.46	0.46
1:D:725:LEU:O	1:D:727:ILE:HG23	2.16	0.46
1:A:668:ASP:OD1	1:A:745:GLY:N	2.48	0.46
1:D:722:ALA:HB2	1:D:729:ARG:HA	1.97	0.45
1:D:303:LEU:HD13	1:D:345:PHE:HD2	1.82	0.45
1:A:589:GLU:O	1:A:593:LYS:HG3	2.17	0.44
1:D:659:LEU:HG	1:D:710:TYR:CE1	2.53	0.44
1:A:565:ALA:N	1:A:566:PRO:CD	2.81	0.44
1:D:565:ALA:N	1:D:566:PRO:CD	2.81	0.44
1:D:847:LEU:C	1:D:847:LEU:HD23	2.38	0.43
1:D:516:GLN:HG2	7:D:991:HOH:O	2.18	0.43
1:D:500:LEU:HD12	1:D:503:MET:HE3	2.01	0.42
1:A:339:ASN:HB2	7:A:890:HOH:O	2.20	0.42
1:D:820:LEU:HD21	1:D:843:ARG:CZ	2.50	0.41
2:E:28:DC:H2'	2:E:29:DOC:H6	2.02	0.41
4:A:203:DCP:O5'	4:A:203:DCP:H6	2.20	0.41
1:A:816:LYS:HD2	7:A:1318:HOH:O	2.19	0.41
1:D:418:GLN:HA	7:D:1386:HOH:O	2.19	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:596:ARG:HA	1:A:597:PRO:HD3	1.96	0.41
1:D:674:ALA:HB1	1:D:681:ILE:HD11	2.02	0.41
1:D:519:TYR:CD2	1:D:525:GLU:HG2	2.55	0.41
1:A:412:ALA:HA	1:A:415:LYS:HD3	2.02	0.40

All (4) 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)
7:A:1322:HOH:O	7:D:1323:HOH:O[4_445]	1.74	0.46
7:A:1198:HOH:O	7:D:1412:HOH:O[2_745]	1.86	0.34
7:A:1263:HOH:O	7:D:1267:HOH:O[2_745]	2.03	0.17
7:A:1357:HOH:O	7:D:1348:HOH:O[4_445]	2.14	0.06

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	556/592 (94%)	544 (98%)	11 (2%)	1 (0%)	47 29
1	D	582/592 (98%)	569 (98%)	13 (2%)	0	100 100
All	All	1138/1184 (96%)	1113 (98%)	24 (2%)	1 (0%)	51 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
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	477/507 (94%)	471 (99%)	6 (1%)	69 52
1	D	500/507 (99%)	498 (100%)	2 (0%)	91 86
All	All	977/1014 (96%)	969 (99%)	8 (1%)	81 72

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

Mol	Chain	Res	Type
1	A	670	ASN
1	A	738	ARG
1	A	742	SER
1	A	776	ILE
1	A	782	ASN
1	A	812	ASN
1	D	303	LEU
1	D	728	SER

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)

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	DOC	B	29	2,3	16,19,20	0.44	0	20,26,29	0.68	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	DOC	E	29	2,3	16,19,20	0.51	0	20,26,29	0.46	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
2	DOC	B	29	2,3	-	0/7/18/19	0/2/2/2
2	DOC	E	29	2,3	-	0/7/18/19	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	29	DOC	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
6	SO4	D	1	-	4,4,4	0.14	0	6,6,6	0.27	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	DCP	A	201	5	25,29,29	1.61	4 (16%)	37,45,45	1.52	5 (13%)
6	SO4	A	3	-	4,4,4	0.17	0	6,6,6	0.21	0
6	SO4	D	4	-	4,4,4	0.18	0	6,6,6	0.13	0
4	DCP	D	202	5	25,29,29	1.58	5 (20%)	37,45,45	1.33	7 (18%)
4	DCP	A	203	-	25,29,29	1.77	4 (16%)	37,45,45	1.32	4 (10%)
6	SO4	A	2	-	4,4,4	0.15	0	6,6,6	0.16	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	DCP	D	202	5	-	4/22/34/34	0/2/2/2
4	DCP	A	203	-	-	5/22/34/34	0/2/2/2
4	DCP	A	201	5	-	3/22/34/34	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	203	DCP	C2'-C3'	-4.96	1.39	1.52
4	A	201	DCP	C2'-C3'	-4.61	1.40	1.52
4	D	202	DCP	C2'-C3'	-4.24	1.41	1.52
4	A	201	DCP	C4-N4	3.38	1.41	1.33
4	A	203	DCP	C4-N4	3.13	1.41	1.33
4	D	202	DCP	C4-N4	2.95	1.40	1.33
4	A	203	DCP	C2-N1	-2.80	1.33	1.40
4	A	201	DCP	C6-C5	2.49	1.40	1.35
4	D	202	DCP	C2-N1	-2.48	1.34	1.40
4	A	203	DCP	O4'-C4'	-2.20	1.40	1.45
4	D	202	DCP	C6-C5	2.06	1.39	1.35
4	A	201	DCP	C2-N1	-2.03	1.35	1.40
4	D	202	DCP	O3'-C3'	2.02	1.47	1.43

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	201	DCP	O4'-C1'-N1	5.57	117.82	107.86
4	A	203	DCP	PB-O3A-PA	-4.15	118.60	132.83
4	A	203	DCP	PB-O3B-PG	-3.90	119.45	132.83
4	D	202	DCP	O4'-C1'-N1	3.55	114.20	107.86

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	201	DCP	O2-C2-N3	-2.86	117.68	122.33
4	A	201	DCP	O5'-C5'-C4'	2.86	118.82	108.99
4	A	201	DCP	C2'-C1'-N1	-2.83	107.25	113.77
4	A	201	DCP	O2G-PG-O3B	2.63	113.45	104.64
4	D	202	DCP	O2G-PG-O3B	2.55	113.20	104.64
4	D	202	DCP	C1'-N1-C2	2.51	122.13	117.74
4	D	202	DCP	O5'-C5'-C4'	2.39	117.23	108.99
4	D	202	DCP	C1'-N1-C6	-2.27	117.06	121.55
4	D	202	DCP	O2-C2-N3	-2.27	118.64	122.33
4	A	203	DCP	O5'-C5'-C4'	2.16	116.44	108.99
4	D	202	DCP	C2'-C1'-N1	-2.09	108.95	113.77
4	A	203	DCP	O4'-C1'-N1	2.08	111.57	107.86

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	203	DCP	C5'-O5'-PA-O1A
4	D	202	DCP	PB-O3B-PG-O1G
4	A	203	DCP	PG-O3B-PB-O1B
4	D	202	DCP	PB-O3B-PG-O2G
4	A	201	DCP	PB-O3A-PA-O2A
4	D	202	DCP	PB-O3A-PA-O2A
4	A	203	DCP	C5'-O5'-PA-O2A
4	A	201	DCP	PB-O3B-PG-O1G
4	A	203	DCP	PG-O3B-PB-O2B
4	A	203	DCP	C5'-O5'-PA-O3A
4	A	201	DCP	PG-O3B-PB-O2B
4	D	202	DCP	PB-O3A-PA-O1A

There are no ring outliers.

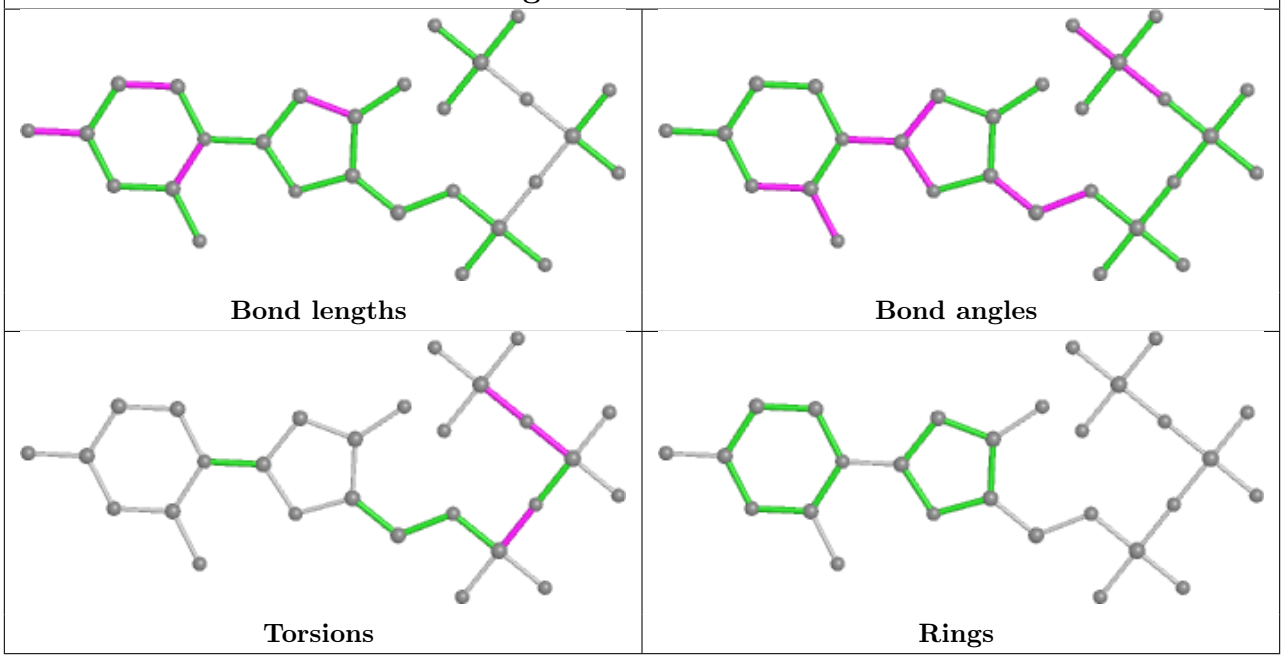
3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	201	DCP	3	0
4	D	202	DCP	2	0
4	A	203	DCP	2	0

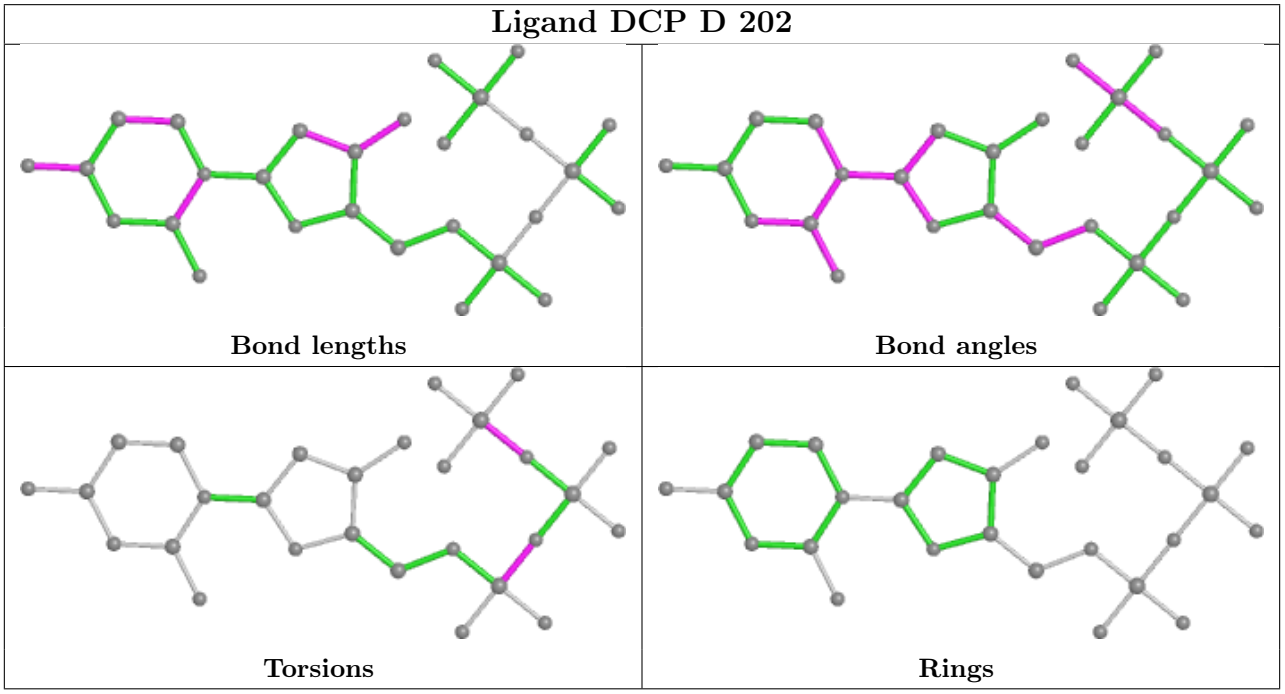
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

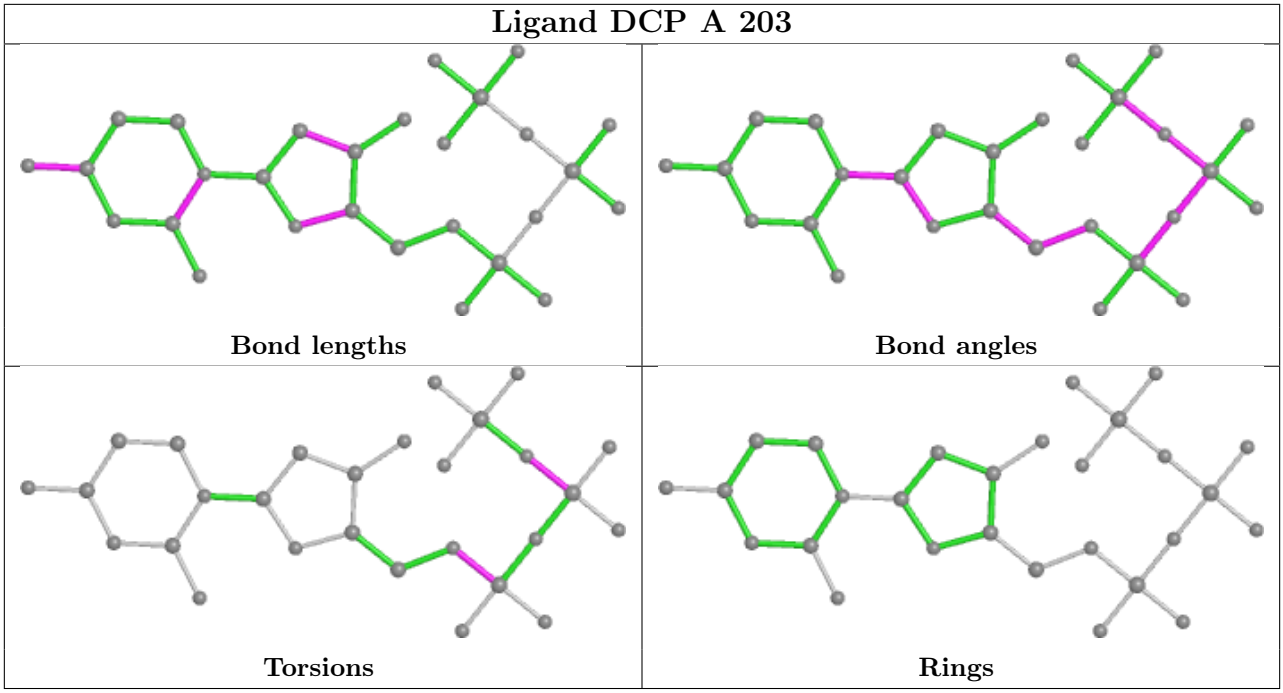
also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted 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 DCP A 201



Ligand DCP D 202





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, 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	557/592 (94%)	0.37	33 (5%) 22 26	15, 32, 54, 66	0
1	D	579/592 (97%)	0.12	6 (1%) 82 87	9, 20, 39, 52	0
2	B	8/9 (88%)	-0.33	0 100 100	19, 25, 44, 51	0
2	E	8/9 (88%)	-0.17	0 100 100	14, 24, 41, 50	0
3	C	12/13 (92%)	0.10	1 (8%) 11 14	18, 25, 60, 78	0
3	F	12/13 (92%)	0.13	0 100 100	12, 20, 46, 63	0
All	All	1176/1228 (95%)	0.24	40 (3%) 45 51	9, 26, 49, 78	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	735	PHE	5.1
1	A	630	LEU	4.7
1	A	819	ARG	4.7
1	A	743	PHE	4.1
3	C	1	DA	4.1
1	A	739	TYR	3.7
1	A	704	GLN	3.6
1	A	679	LEU	3.6
1	A	842	GLU	3.6
1	A	703	ARG	3.2
1	A	846	ARG	3.1
1	D	689	ILE	3.0
1	A	707	ALA	3.0
1	A	866	TYR	2.9
1	A	671	LEU	2.8
1	A	705	ALA	2.7
1	A	742	SER	2.7
1	A	531	PRO	2.6
1	A	863	LYS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	734	GLU	2.5
1	A	629	ARG	2.5
1	A	868	TYR	2.4
1	D	691	GLN	2.4
1	A	730	LYS	2.3
1	A	847	LEU	2.3
1	A	738	ARG	2.3
1	A	740	PHE	2.2
1	D	699	PRO	2.2
1	A	876	LYS	2.2
1	A	521	LEU	2.2
1	A	670	ASN	2.1
1	D	700	ASN	2.1
1	A	678	ASP	2.1
1	A	575	LEU	2.1
1	A	520	GLU	2.1
1	A	843	ARG	2.0
1	D	725	LEU	2.0
1	A	848	VAL	2.0
1	D	690	PHE	2.0
1	A	513	THR	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	DOC	B	29	18/19	0.96	0.08	16,25,32,35	0
2	DOC	E	29	18/19	0.98	0.11	8,14,21,22	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,

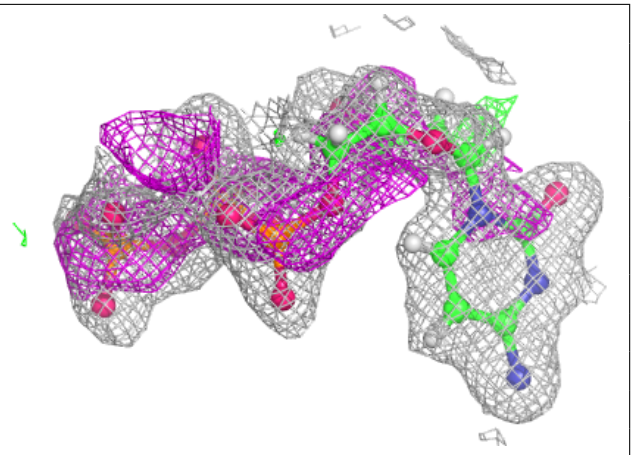
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

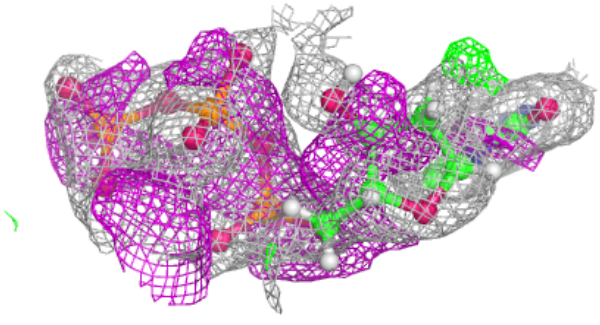
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	DCP	A	201	28/28	0.88	0.18	23,34,44,45	0
6	SO4	A	3	5/5	0.89	0.28	48,51,60,65	0
4	DCP	A	203	28/28	0.90	0.11	30,46,72,84	0
6	SO4	D	4	5/5	0.94	0.26	50,51,56,58	0
4	DCP	D	202	28/28	0.95	0.09	13,20,29,30	0
6	SO4	A	2	5/5	0.95	0.09	33,34,43,48	0
6	SO4	D	1	5/5	0.98	0.09	28,33,41,43	0
5	MN	A	1	1/1	0.98	0.08	37,37,37,37	0
5	MN	D	2	1/1	0.99	0.05	20,20,20,20	0

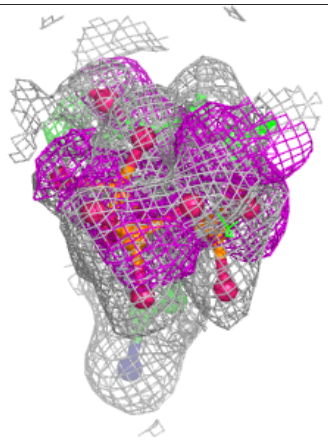
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DCP A 201:

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

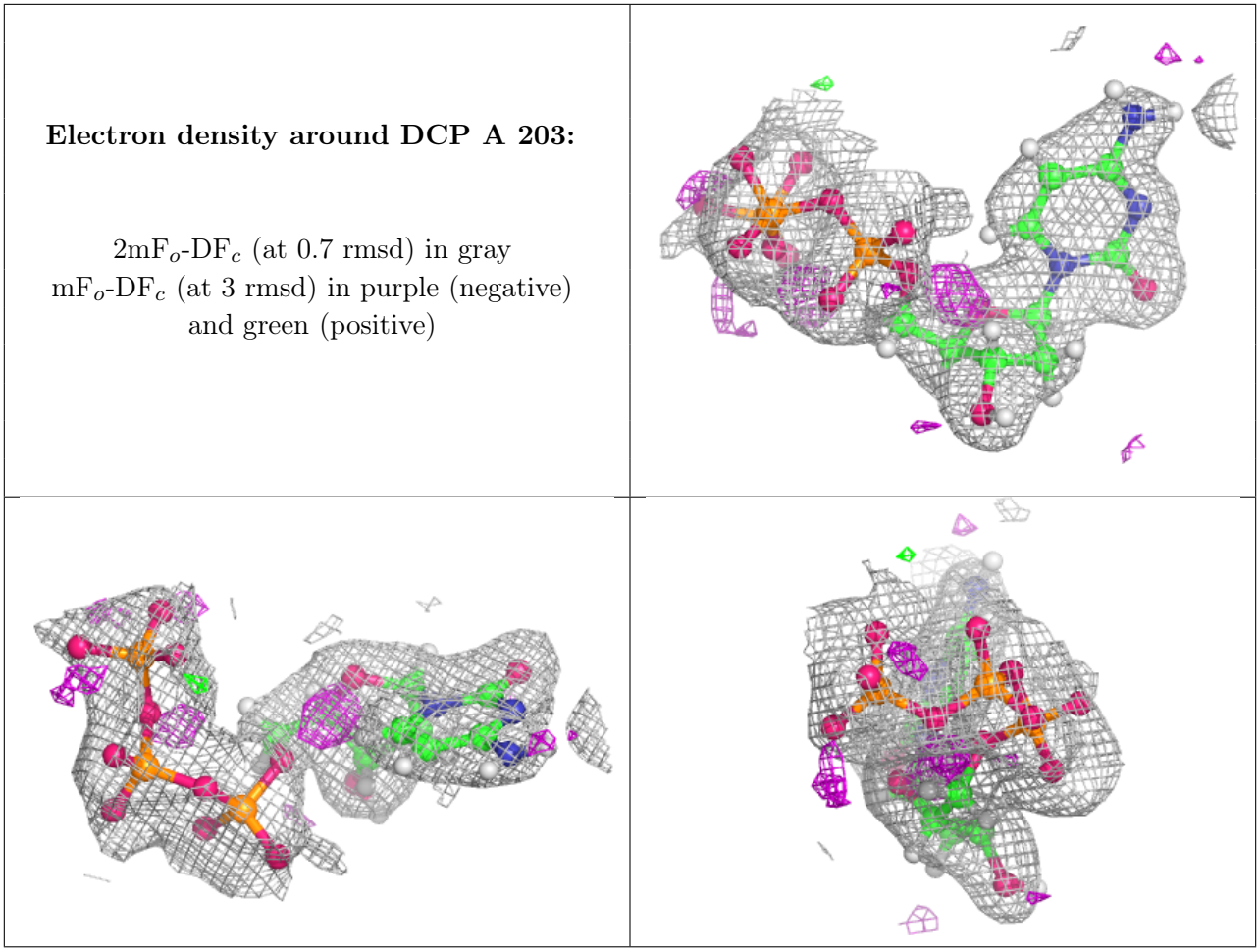






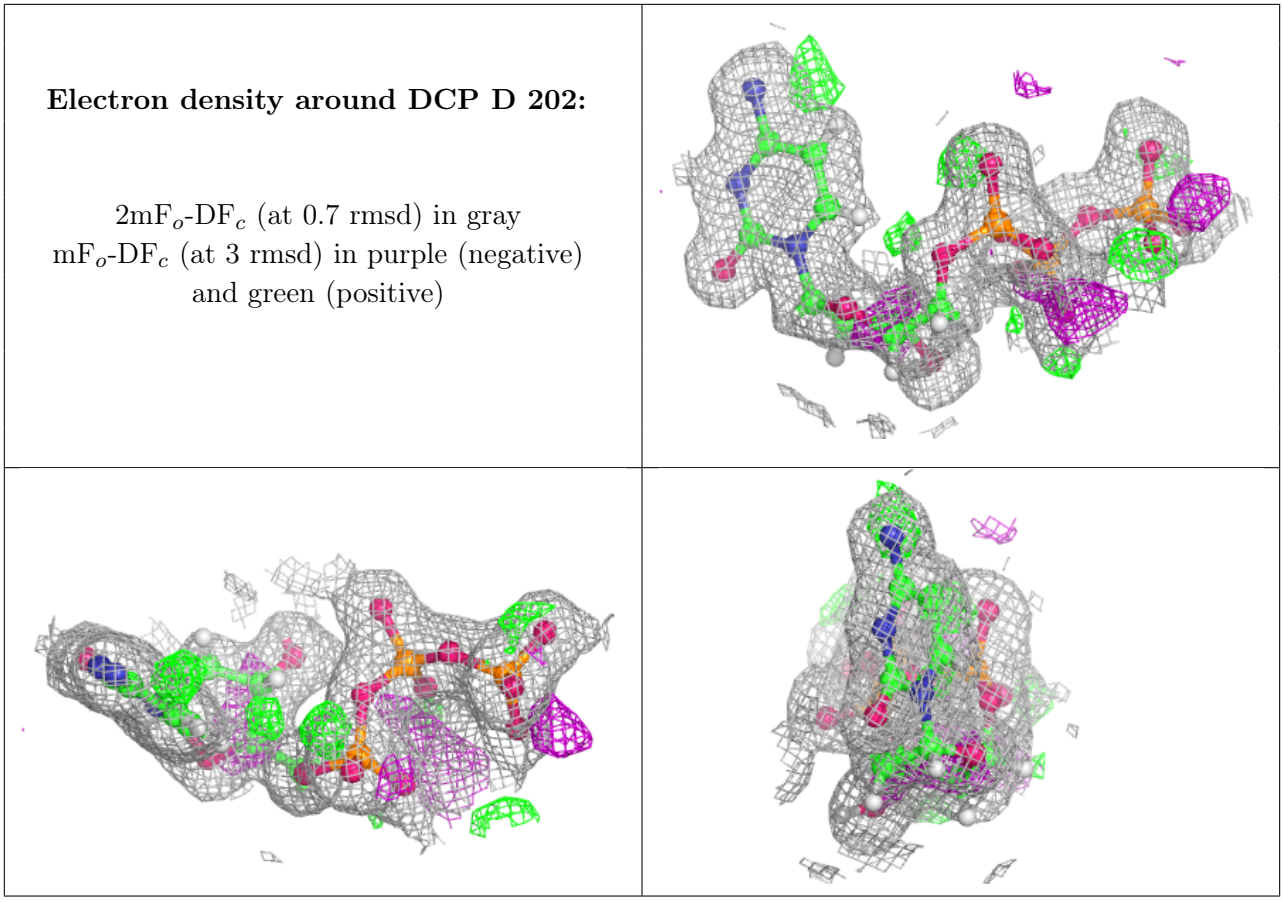
Electron density around DCP A 203:

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



Electron density around DCP D 202:

$2mF_o$-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.