



# Full wwPDB X-ray Structure Validation Report ⓘ

May 20, 2024 – 03:17 AM JST

PDB ID : 5WRN  
Title : Human thymidylate synthase complexed with dCMP  
Authors : Chen, D.; Nordlund, P.  
Deposited on : 2016-12-02  
Resolution : 2.39 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36.2  
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) : 2.36.2

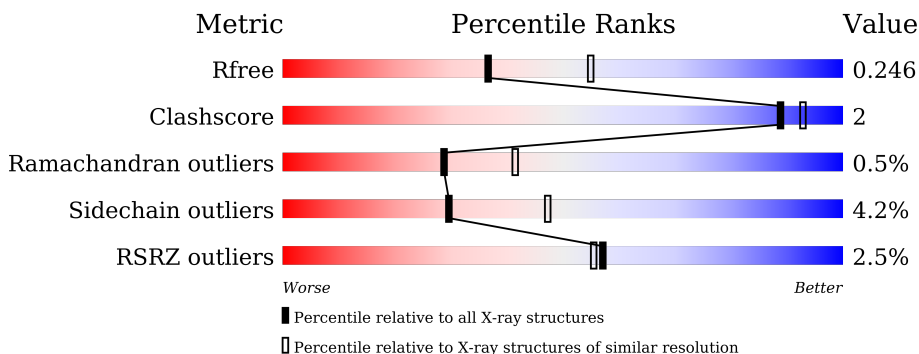
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.39 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	289	
1	B	289	
1	C	289	
1	D	289	
1	E	289	
1	F	289	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 14115 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.

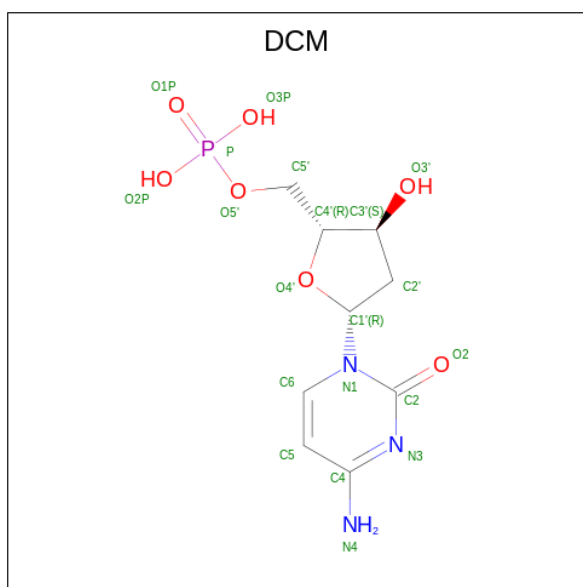
- Molecule 1 is a protein called Thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	282	2275	1458	396	409	12	0	0	0
1	A	284	2288	1463	400	414	11	0	0	0
1	C	282	2265	1449	396	409	11	0	0	0
1	D	286	2289	1466	399	412	12	0	0	0
1	E	286	2301	1471	402	416	12	0	0	0
1	F	277	2249	1440	394	404	11	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	25	MET	-	initiating methionine	UNP P04818
A	25	MET	-	initiating methionine	UNP P04818
C	25	MET	-	initiating methionine	UNP P04818
D	25	MET	-	initiating methionine	UNP P04818
E	25	MET	-	initiating methionine	UNP P04818
F	25	MET	-	initiating methionine	UNP P04818

- Molecule 2 is 2'-DEOXYCYTIDINE-5'-MONOPHOSPHATE (three-letter code: DCM) (formula: C<sub>9</sub>H<sub>14</sub>N<sub>3</sub>O<sub>7</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	B	1	20	9	3	7	1	0	0
2	A	1	20	9	3	7	1	0	0
2	C	1	20	9	3	7	1	0	0
2	D	1	20	9	3	7	1	0	0
2	E	1	20	9	3	7	1	0	0
2	F	1	20	9	3	7	1	0	0

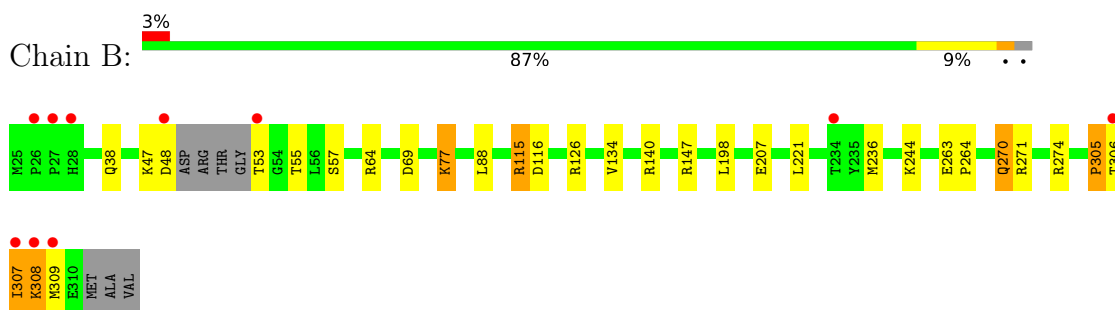
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	50	50	50	0	0
3	A	47	47	47	0	0
3	C	45	45	45	0	0
3	D	54	54	54	0	0
3	E	64	64	64	0	0
3	F	68	68	68	0	0

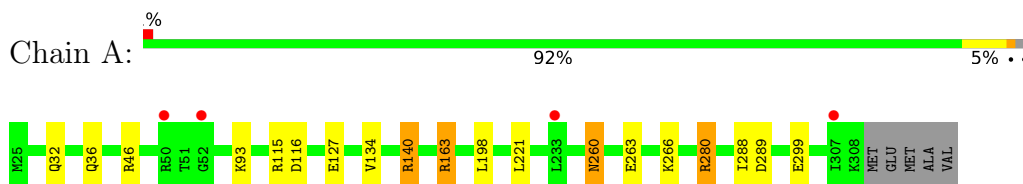
### 3 Residue-property plots [i](#)

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.

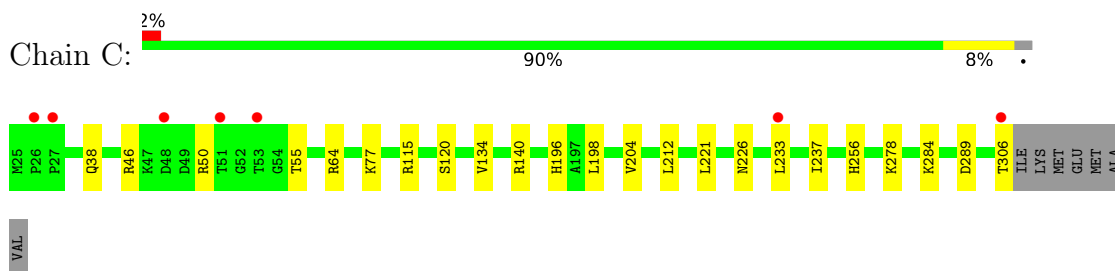
- Molecule 1: Thymidylate synthase



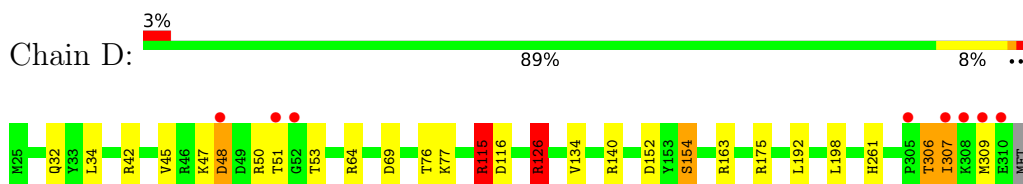
- Molecule 1: Thymidylate synthase



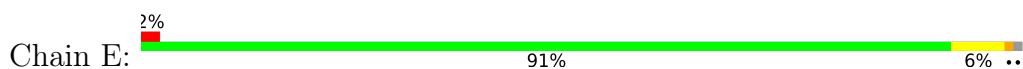
- Molecule 1: Thymidylate synthase

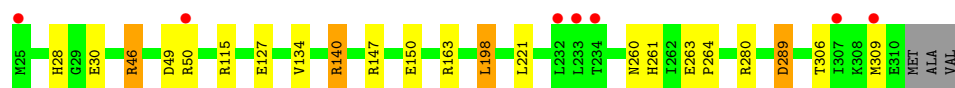


- Molecule 1: Thymidylate synthase

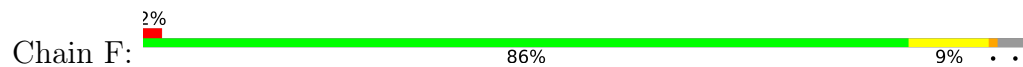


- Molecule 1: Thymidylate synthase





- Molecule 1: Thymidylate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.10Å 110.10Å 317.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.39 27.05 – 2.39	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.00-2.39) 98.1 (27.05-2.39)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.63 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.188 , 0.247 0.195 , 0.246	Depositor DCC
$R_{free}$ test set	3757 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.4	Xtrriage
Anisotropy	0.048	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 28.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14115	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 55.02 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.3588e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DCM

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	0.83	0/2348	0.94	8/3179 (0.3%)
1	B	0.83	0/2334	0.97	8/3157 (0.3%)
1	C	0.76	0/2324	0.88	4/3146 (0.1%)
1	D	0.79	0/2349	0.93	7/3181 (0.2%)
1	E	0.79	0/2361	0.93	5/3196 (0.2%)
1	F	0.76	0/2308	0.90	4/3122 (0.1%)
All	All	0.79	0/14024	0.92	36/18981 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	E	0	2
All	All	0	3

There are no bond length outliers.

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	140	ARG	NE-CZ-NH2	-10.97	114.82	120.30
1	A	163	ARG	NE-CZ-NH2	9.11	124.86	120.30
1	A	140	ARG	NE-CZ-NH2	-8.26	116.17	120.30
1	E	46	ARG	NE-CZ-NH2	7.99	124.29	120.30
1	B	140	ARG	NE-CZ-NH2	-7.37	116.61	120.30
1	D	115	ARG	NE-CZ-NH1	7.36	123.98	120.30
1	E	140	ARG	NE-CZ-NH2	-7.01	116.80	120.30
1	D	42	ARG	NE-CZ-NH1	6.79	123.69	120.30

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	175	ARG	NE-CZ-NH1	-6.75	116.92	120.30
1	D	140	ARG	NE-CZ-NH2	-6.74	116.93	120.30
1	E	115	ARG	NE-CZ-NH1	6.69	123.64	120.30
1	A	115	ARG	NE-CZ-NH1	6.63	123.62	120.30
1	E	115	ARG	NE-CZ-NH2	-6.43	117.08	120.30
1	F	140	ARG	NE-CZ-NH1	6.41	123.51	120.30
1	E	198	LEU	CA-CB-CG	6.13	129.40	115.30
1	B	64	ARG	NE-CZ-NH2	-6.11	117.24	120.30
1	B	64	ARG	NE-CZ-NH1	5.95	123.27	120.30
1	D	126	ARG	NE-CZ-NH2	-5.93	117.34	120.30
1	B	48	ASP	CB-CG-OD1	5.76	123.48	118.30
1	A	115	ARG	NE-CZ-NH2	-5.70	117.45	120.30
1	C	64	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	198	LEU	CA-CB-CG	5.59	128.16	115.30
1	B	147	ARG	NE-CZ-NH1	5.59	123.09	120.30
1	C	115	ARG	NE-CZ-NH1	5.55	123.08	120.30
1	B	115	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	B	198	LEU	CA-CB-CG	5.48	127.91	115.30
1	A	163	ARG	CD-NE-CZ	5.46	131.24	123.60
1	A	140	ARG	NE-CZ-NH1	5.42	123.01	120.30
1	C	115	ARG	NE-CZ-NH2	-5.40	117.60	120.30
1	F	68	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	A	163	ARG	CG-CD-NE	5.38	123.09	111.80
1	D	115	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	B	274	ARG	NE-CZ-NH1	5.24	122.92	120.30
1	F	115	ARG	NE-CZ-NH1	5.19	122.89	120.30
1	C	64	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	D	175	ARG	NE-CZ-NH2	5.08	122.84	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	53	THR	Peptide
1	E	309	MET	Peptide
1	E	49	ASP	Peptide

## 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	2288	0	2249	3	1
1	B	2275	0	2243	16	1
1	C	2265	0	2224	8	0
1	D	2289	0	2245	10	1
1	E	2301	0	2260	6	0
1	F	2249	0	2213	10	0
2	A	20	0	12	0	0
2	B	20	0	11	0	0
2	C	20	0	12	2	0
2	D	20	0	11	0	0
2	E	20	0	12	0	0
2	F	20	0	11	2	0
3	A	47	0	0	0	0
3	B	50	0	0	0	0
3	C	45	0	0	0	0
3	D	54	0	0	0	0
3	E	64	0	0	2	0
3	F	68	0	0	1	0
All	All	14115	0	13503	51	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (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:B:271:ARG:HH22	1:B:307:ILE:HD13	1.40	0.87
1:B:271:ARG:NH2	1:B:307:ILE:HD13	2.02	0.74
1:D:115:ARG:NH2	1:D:126:ARG:O	2.21	0.71
1:B:305:PRO:HB2	1:B:307:ILE:HD12	1.76	0.66
1:D:261:HIS:CE1	1:D:309:MET:HB3	2.32	0.65
1:E:140:ARG:NH2	1:E:289:ASP:OD2	2.31	0.63
1:F:297:GLN:HG3	3:F:564:HOH:O	1.99	0.62
1:C:226:ASN:HD21	2:C:400:DCM:HN41	1.47	0.62
1:D:51:THR:HG23	1:D:53:THR:HG22	1.83	0.60
1:B:77:LYS:HD2	1:B:308:LYS:HG3	1.84	0.59
1:C:226:ASN:ND2	2:C:400:DCM:HN41	2.01	0.59
1:A:32:GLN:O	1:A:36:GLN:HG3	2.04	0.57
1:B:271:ARG:HH22	1:B:307:ILE:CD1	2.18	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:280[B]:ARG:NH1	1:F:297:GLN:OE1	2.39	0.54
1:F:102:SER:HB2	1:F:110:ASP:OD1	2.06	0.54
1:D:198:LEU:C	1:D:198:LEU:HD12	2.29	0.53
1:B:88:LEU:HD23	1:B:236:MET:HE1	1.91	0.52
1:D:152:ASP:OD1	1:D:154:SER:OG	2.26	0.50
1:D:32:GLN:NE2	1:D:64:ARG:O	2.32	0.49
1:B:270:GLN:NE2	1:B:270:GLN:HA	2.26	0.49
1:B:263:GLU:N	1:B:264:PRO:CD	2.75	0.49
1:B:115:ARG:NH2	1:B:126:ARG:O	2.34	0.48
1:B:116:ASP:OD1	1:C:306:THR:HG23	2.13	0.48
1:F:226:ASN:HD21	2:F:400:DCM:HN41	1.62	0.48
1:B:47:LYS:O	1:B:55:THR:OG1	2.26	0.47
1:B:77:LYS:HD2	1:B:308:LYS:CG	2.45	0.47
1:A:140:ARG:NH2	1:A:289:ASP:OD2	2.46	0.46
1:E:127:GLU:HG3	3:E:550:HOH:O	2.15	0.46
1:D:77:LYS:HB2	1:D:307:ILE:HG13	1.98	0.46
1:B:207:GLU:HA	1:B:244:LYS:O	2.16	0.46
1:A:280:ARG:HH21	1:A:299:GLU:CD	2.20	0.46
1:D:47:LYS:O	1:D:48:ASP:CB	2.64	0.46
1:F:196:HIS:HB3	1:F:212:LEU:HD11	1.97	0.45
1:B:77:LYS:NZ	1:B:308:LYS:HD3	2.31	0.45
1:E:263:GLU:N	1:E:264:PRO:CD	2.80	0.44
1:D:34:LEU:HD11	1:D:76:THR:HG21	2.00	0.44
1:E:28:HIS:CD2	1:E:30:GLU:H	2.36	0.43
1:E:261:HIS:HD2	3:E:501:HOH:O	2.01	0.42
1:B:307:ILE:HG22	1:B:307:ILE:O	2.19	0.42
1:E:28:HIS:HD2	1:E:30:GLU:H	1.68	0.42
1:C:196:HIS:HB3	1:C:212:LEU:HD11	2.00	0.42
1:F:226:ASN:ND2	2:F:400:DCM:HN41	2.19	0.41
1:F:207:GLU:HA	1:F:244:LYS:O	2.20	0.41
1:B:77:LYS:HZ2	1:B:308:LYS:HD3	1.86	0.41
1:C:55:THR:HB	1:C:256:HIS:HB2	2.02	0.41
1:F:151:SER:HB2	1:F:153:TYR:CZ	2.56	0.41
1:C:198:LEU:C	1:C:198:LEU:HD12	2.42	0.40
1:C:204:VAL:HG21	1:D:45:VAL:HG11	2.04	0.40
1:C:233:LEU:HD11	1:C:237:ILE:HD11	2.03	0.40
1:F:93:LYS:HE2	1:F:100:GLU:OE1	2.21	0.40
1:F:140:ARG:NH2	1:F:289:ASP:OD2	2.55	0.40

All (2) 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)
1:A:116:ASP:OD1	1:D:306:THR:CG2[1_455]	1.90	0.30
1:B:270:GLN:OE1	1:B:270:GLN:OE1[7_555]	2.00	0.20

## 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	282/289 (98%)	272 (96%)	8 (3%)	2 (1%)	22	32
1	B	278/289 (96%)	271 (98%)	6 (2%)	1 (0%)	34	48
1	C	280/289 (97%)	271 (97%)	8 (3%)	1 (0%)	34	48
1	D	284/289 (98%)	275 (97%)	7 (2%)	2 (1%)	22	32
1	E	284/289 (98%)	274 (96%)	8 (3%)	2 (1%)	22	32
1	F	274/289 (95%)	264 (96%)	9 (3%)	1 (0%)	34	48
All	All	1682/1734 (97%)	1627 (97%)	46 (3%)	9 (0%)	29	41

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	48	ASP
1	E	50	ARG
1	A	260	ASN
1	F	134	VAL
1	B	134	VAL
1	A	134	VAL
1	C	134	VAL
1	D	134	VAL
1	E	134	VAL

### 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	246/252 (98%)	236 (96%)	10 (4%)	30	48
1	B	245/252 (97%)	234 (96%)	11 (4%)	27	44
1	C	242/252 (96%)	232 (96%)	10 (4%)	30	48
1	D	244/252 (97%)	234 (96%)	10 (4%)	30	48
1	E	247/252 (98%)	237 (96%)	10 (4%)	31	49
1	F	242/252 (96%)	230 (95%)	12 (5%)	24	40
All	All	1466/1512 (97%)	1403 (96%)	63 (4%)	30	46

All (63) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	38	GLN
1	B	57	SER
1	B	69	ASP
1	B	77	LYS
1	B	221	LEU
1	B	270	GLN
1	B	305	PRO
1	B	306	THR
1	B	307	ILE
1	B	308	LYS
1	B	309	MET
1	A	46	ARG
1	A	93	LYS
1	A	127	GLU
1	A	163	ARG
1	A	221	LEU
1	A	260	ASN
1	A	263	GLU
1	A	266	LYS
1	A	280	ARG
1	A	288	ILE
1	C	38	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	46	ARG
1	C	50	ARG
1	C	77	LYS
1	C	120	SER
1	C	140	ARG
1	C	221	LEU
1	C	278	LYS
1	C	284	LYS
1	C	289	ASP
1	D	50	ARG
1	D	69	ASP
1	D	115	ARG
1	D	116	ASP
1	D	126	ARG
1	D	154	SER
1	D	163	ARG
1	D	192	LEU
1	D	306	THR
1	D	307	ILE
1	E	46	ARG
1	E	147	ARG
1	E	150	GLU
1	E	163	ARG
1	E	198	LEU
1	E	221	LEU
1	E	260	ASN
1	E	280	ARG
1	E	289	ASP
1	E	306	THR
1	F	57	SER
1	F	69	ASP
1	F	82	LYS
1	F	107	LYS
1	F	176	ARG
1	F	185	ARG
1	F	221	LEU
1	F	272	GLU
1	F	280[A]	ARG
1	F	280[B]	ARG
1	F	288	ILE
1	F	306	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26)

such sidechains are listed below:

Mol	Chain	Res	Type
1	B	62	GLN
1	B	112	ASN
1	B	211	GLN
1	B	270	GLN
1	A	62	GLN
1	A	112	ASN
1	A	211	GLN
1	A	260	ASN
1	C	62	GLN
1	C	211	GLN
1	D	62	GLN
1	D	112	ASN
1	D	162	GLN
1	D	211	GLN
1	D	261	HIS
1	D	270	GLN
1	E	28	HIS
1	E	62	GLN
1	E	211	GLN
1	E	260	ASN
1	E	261	HIS
1	F	62	GLN
1	F	112	ASN
1	F	205	ASN
1	F	211	GLN
1	F	261	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

6 ligands 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	DCM	A	400	-	21,21,21	2.94	7 (33%)	31,31,31	1.94	10 (32%)
2	DCM	F	400	-	21,21,21	3.00	8 (38%)	31,31,31	1.34	6 (19%)
2	DCM	C	400	-	21,21,21	2.88	7 (33%)	31,31,31	1.60	7 (22%)
2	DCM	E	400	-	21,21,21	2.95	7 (33%)	31,31,31	1.81	9 (29%)
2	DCM	D	400	-	21,21,21	2.96	6 (28%)	31,31,31	1.76	6 (19%)
2	DCM	B	400	-	21,21,21	3.00	8 (38%)	31,31,31	1.66	9 (29%)

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	DCM	A	400	-	-	1/10/22/22	0/2/2/2
2	DCM	F	400	-	-	1/10/22/22	0/2/2/2
2	DCM	C	400	-	-	1/10/22/22	0/2/2/2
2	DCM	E	400	-	-	2/10/22/22	0/2/2/2
2	DCM	D	400	-	-	6/10/22/22	0/2/2/2
2	DCM	B	400	-	-	1/10/22/22	0/2/2/2

All (43) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	400	DCM	C2'-C3'	-8.04	1.31	1.52
2	D	400	DCM	C2'-C3'	-7.83	1.32	1.52
2	F	400	DCM	C2'-C3'	-7.81	1.32	1.52
2	A	400	DCM	C2'-C3'	-7.45	1.33	1.52
2	B	400	DCM	C2'-C3'	-7.29	1.33	1.52
2	E	400	DCM	C2'-C3'	-6.99	1.34	1.52

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	400	DCM	O4'-C4'	6.40	1.59	1.45
2	D	400	DCM	O4'-C4'	6.31	1.59	1.45
2	E	400	DCM	O4'-C4'	6.15	1.58	1.45
2	F	400	DCM	O4'-C4'	6.06	1.58	1.45
2	E	400	DCM	C2'-C1'	5.87	1.68	1.52
2	A	400	DCM	O4'-C4'	5.79	1.57	1.45
2	C	400	DCM	O4'-C4'	5.69	1.57	1.45
2	D	400	DCM	C2'-C1'	5.35	1.67	1.52
2	A	400	DCM	C2'-C1'	5.23	1.67	1.52
2	B	400	DCM	C2'-C1'	5.15	1.66	1.52
2	F	400	DCM	C2'-C1'	5.06	1.66	1.52
2	C	400	DCM	C2'-C1'	4.83	1.65	1.52
2	F	400	DCM	O4'-C1'	-4.63	1.32	1.42
2	B	400	DCM	O4'-C1'	-4.26	1.32	1.42
2	D	400	DCM	O4'-C1'	-4.24	1.32	1.42
2	C	400	DCM	O4'-C1'	-4.11	1.33	1.42
2	A	400	DCM	C5'-C4'	-4.02	1.39	1.51
2	B	400	DCM	C5'-C4'	-3.97	1.39	1.51
2	D	400	DCM	C5'-C4'	-3.96	1.39	1.51
2	E	400	DCM	C5'-C4'	-3.92	1.39	1.51
2	F	400	DCM	C5'-C4'	-3.89	1.39	1.51
2	C	400	DCM	C5'-C4'	-3.77	1.39	1.51
2	A	400	DCM	O4'-C1'	-3.76	1.33	1.42
2	E	400	DCM	O4'-C1'	-3.71	1.34	1.42
2	E	400	DCM	C4-N4	3.55	1.42	1.33
2	B	400	DCM	O3'-C3'	3.53	1.50	1.43
2	A	400	DCM	C4-N4	3.39	1.41	1.33
2	B	400	DCM	C4-N4	3.14	1.41	1.33
2	E	400	DCM	O3'-C3'	3.03	1.49	1.43
2	F	400	DCM	C4-N4	2.93	1.40	1.33
2	F	400	DCM	O3'-C3'	2.77	1.49	1.43
2	D	400	DCM	C4-N4	2.61	1.40	1.33
2	A	400	DCM	O3'-C3'	2.42	1.48	1.43
2	C	400	DCM	C4-N4	2.33	1.39	1.33
2	C	400	DCM	O3'-C3'	2.20	1.48	1.43
2	B	400	DCM	C6-C5	2.14	1.40	1.35
2	F	400	DCM	C1'-N1	-2.01	1.42	1.48

All (47) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	400	DCM	O2-C2-N3	-5.17	113.92	122.33

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	400	DCM	P-O5'-C5'	4.44	130.53	118.30
2	C	400	DCM	O2-C2-N3	-4.39	115.19	122.33
2	E	400	DCM	O2-C2-N3	-3.93	115.93	122.33
2	D	400	DCM	O4'-C1'-N1	3.93	114.89	107.86
2	A	400	DCM	O4'-C1'-N1	3.86	114.76	107.86
2	A	400	DCM	O2-C2-N3	-3.82	116.11	122.33
2	E	400	DCM	C1'-N1-C2	3.70	124.22	117.74
2	E	400	DCM	O3P-P-O5'	3.61	116.33	106.73
2	B	400	DCM	C1'-N1-C2	3.48	123.84	117.74
2	B	400	DCM	O2-C2-N3	-3.48	116.67	122.33
2	A	400	DCM	C1'-N1-C2	3.46	123.81	117.74
2	A	400	DCM	O2-C2-N1	3.44	126.00	118.89
2	C	400	DCM	C1'-N1-C2	3.37	123.64	117.74
2	C	400	DCM	O4'-C1'-N1	3.33	113.82	107.86
2	F	400	DCM	O2-C2-N3	-3.32	116.94	122.33
2	D	400	DCM	C1'-N1-C2	3.26	123.45	117.74
2	D	400	DCM	O2-C2-N1	3.13	125.36	118.89
2	E	400	DCM	O4'-C1'-N1	3.00	113.22	107.86
2	E	400	DCM	C2'-C1'-N1	2.77	120.14	113.77
2	E	400	DCM	O2-C2-N1	2.68	124.43	118.89
2	D	400	DCM	P-O5'-C5'	2.58	125.41	118.30
2	E	400	DCM	P-O5'-C5'	2.56	125.35	118.30
2	C	400	DCM	O2-C2-N1	2.54	124.15	118.89
2	B	400	DCM	N4-C4-N3	2.53	122.40	117.97
2	F	400	DCM	O2-C2-N1	2.47	124.00	118.89
2	E	400	DCM	N4-C4-N3	2.44	122.25	117.97
2	D	400	DCM	O4'-C4'-C3'	-2.40	100.07	105.67
2	A	400	DCM	C1'-N1-C6	-2.39	116.84	121.55
2	F	400	DCM	O4'-C1'-N1	2.38	112.11	107.86
2	B	400	DCM	C2'-C1'-N1	2.37	119.22	113.77
2	A	400	DCM	O3P-P-O5'	2.32	112.91	106.73
2	B	400	DCM	O4'-C1'-N1	2.30	111.97	107.86
2	F	400	DCM	P-O5'-C5'	2.28	124.56	118.30
2	B	400	DCM	P-O5'-C5'	2.25	124.49	118.30
2	A	400	DCM	N4-C4-N3	2.23	121.88	117.97
2	B	400	DCM	O4'-C4'-C5'	2.21	116.64	109.37
2	F	400	DCM	C1'-N1-C2	2.20	121.60	117.74
2	C	400	DCM	P-O5'-C5'	2.20	124.35	118.30
2	C	400	DCM	C5-C4-N4	-2.19	117.12	120.57
2	A	400	DCM	O5'-P-O1P	2.19	112.62	106.47
2	F	400	DCM	O4'-C4'-C3'	-2.14	100.68	105.67
2	B	400	DCM	O4'-C4'-C3'	-2.13	100.69	105.67

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	400	DCM	C1'-N1-C6	-2.04	117.52	121.55
2	E	400	DCM	C1'-N1-C6	-2.04	117.52	121.55
2	A	400	DCM	O4'-C4'-C5'	2.02	116.03	109.37
2	B	400	DCM	O3P-P-O5'	2.01	112.07	106.73

There are no chirality outliers.

All (12) torsion outliers are listed below:

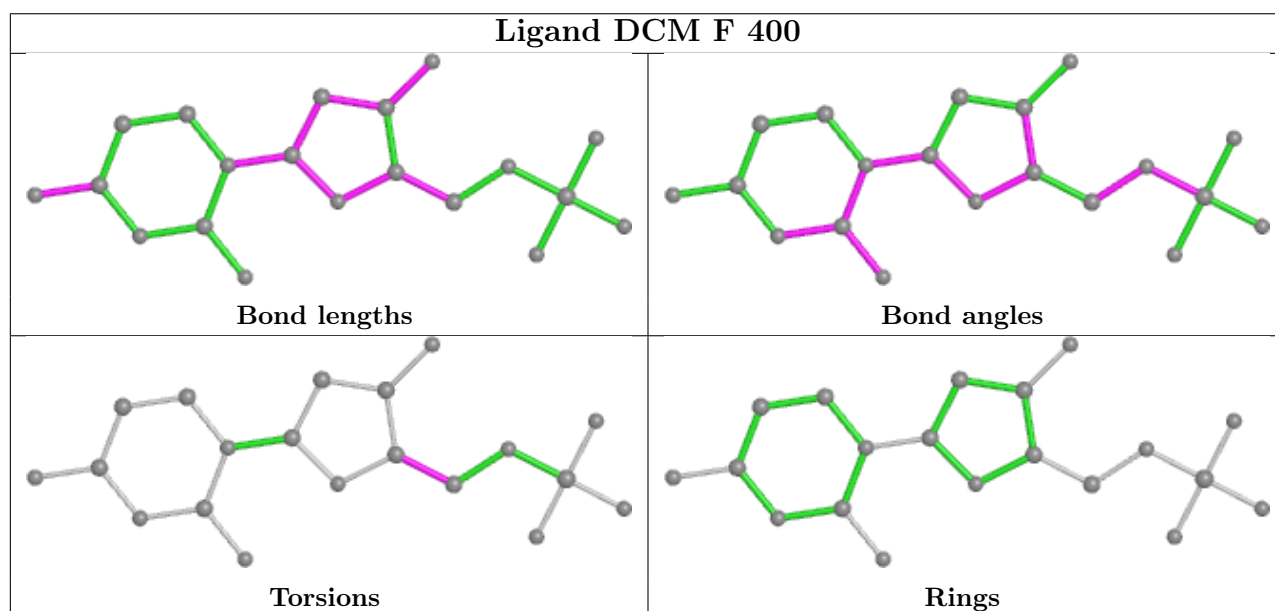
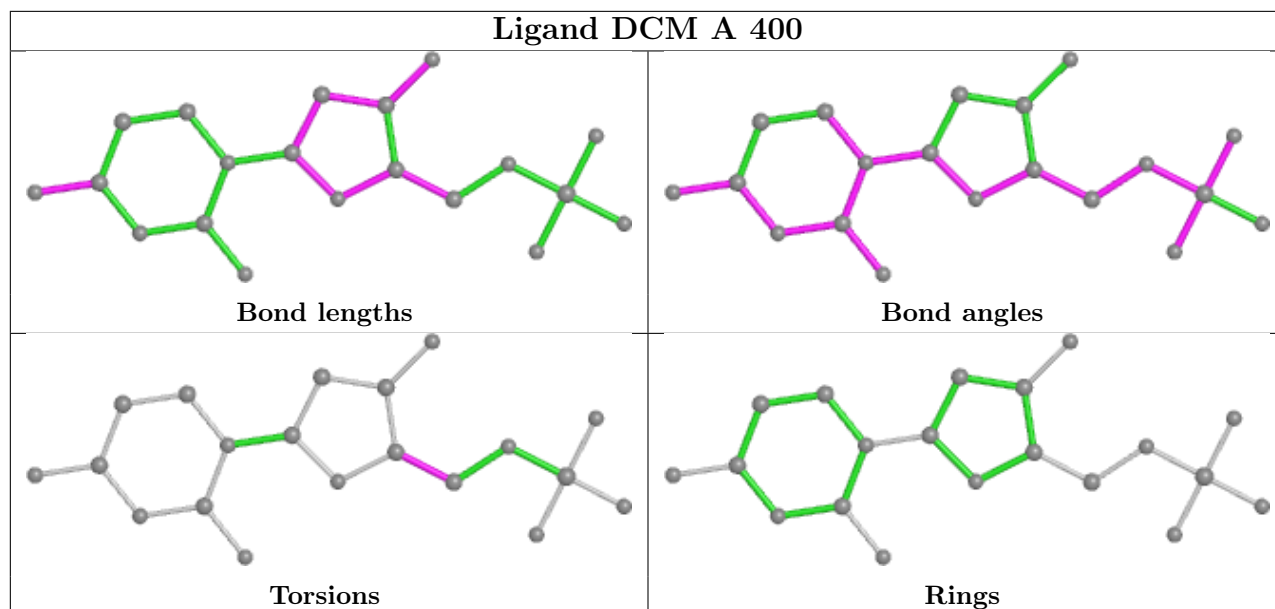
Mol	Chain	Res	Type	Atoms
2	D	400	DCM	O4'-C4'-C5'-O5'
2	D	400	DCM	C5'-O5'-P-O1P
2	D	400	DCM	C5'-O5'-P-O2P
2	D	400	DCM	C3'-C4'-C5'-O5'
2	D	400	DCM	C5'-O5'-P-O3P
2	D	400	DCM	C4'-C5'-O5'-P
2	E	400	DCM	O4'-C4'-C5'-O5'
2	B	400	DCM	O4'-C4'-C5'-O5'
2	A	400	DCM	O4'-C4'-C5'-O5'
2	E	400	DCM	C3'-C4'-C5'-O5'
2	C	400	DCM	O4'-C4'-C5'-O5'
2	F	400	DCM	O4'-C4'-C5'-O5'

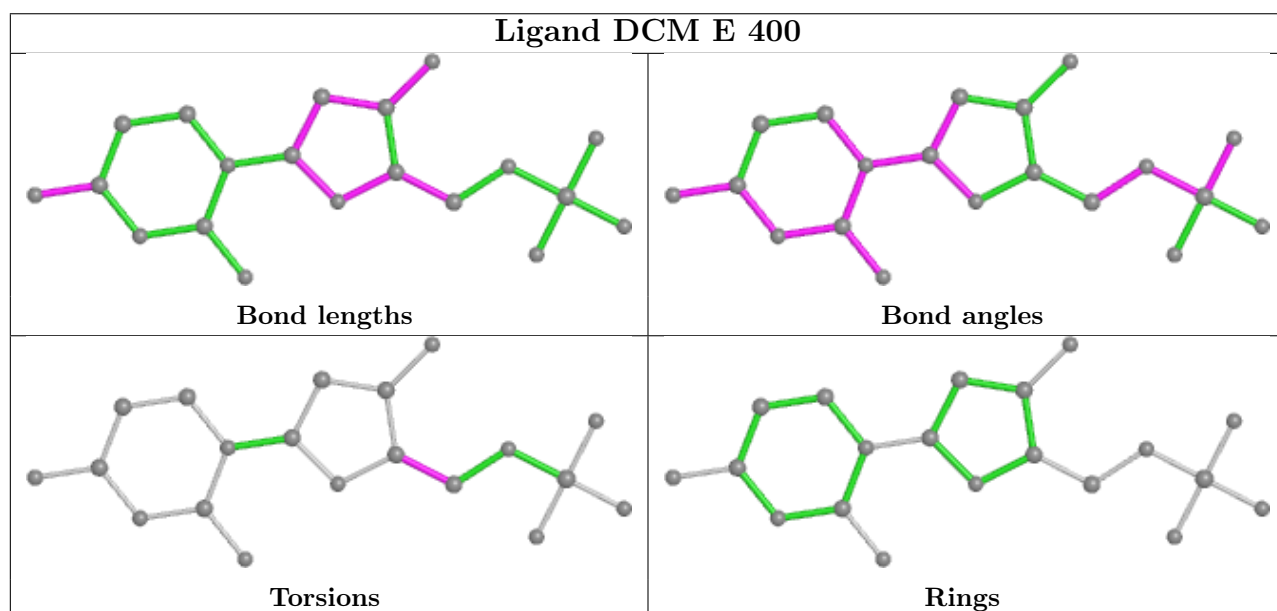
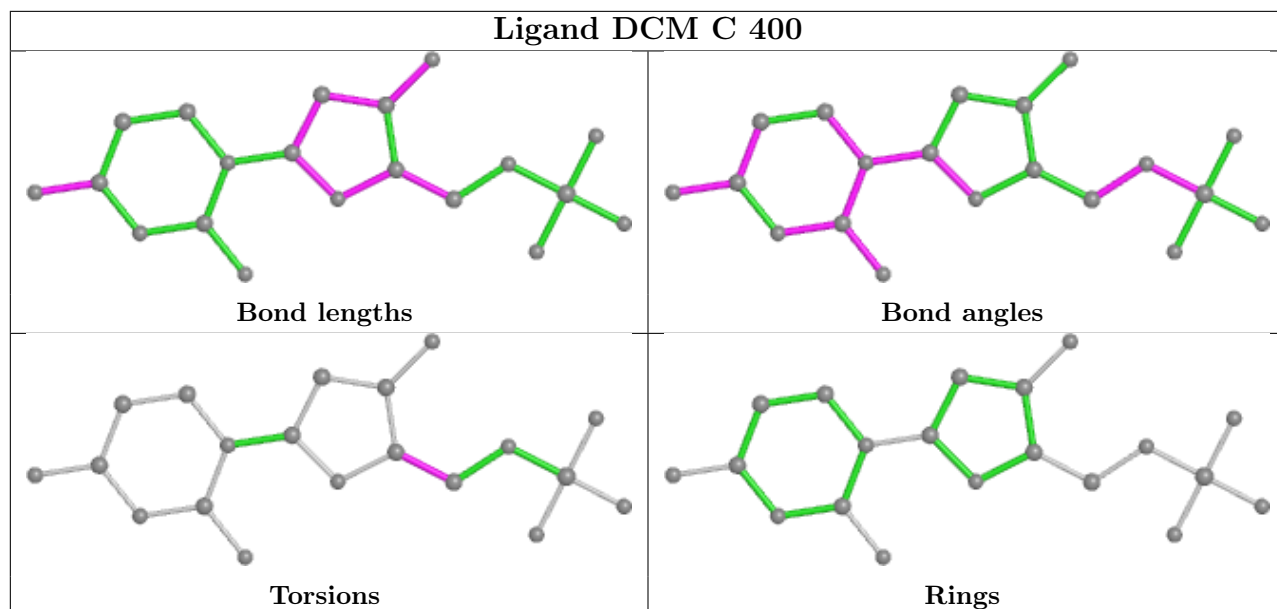
There are no ring outliers.

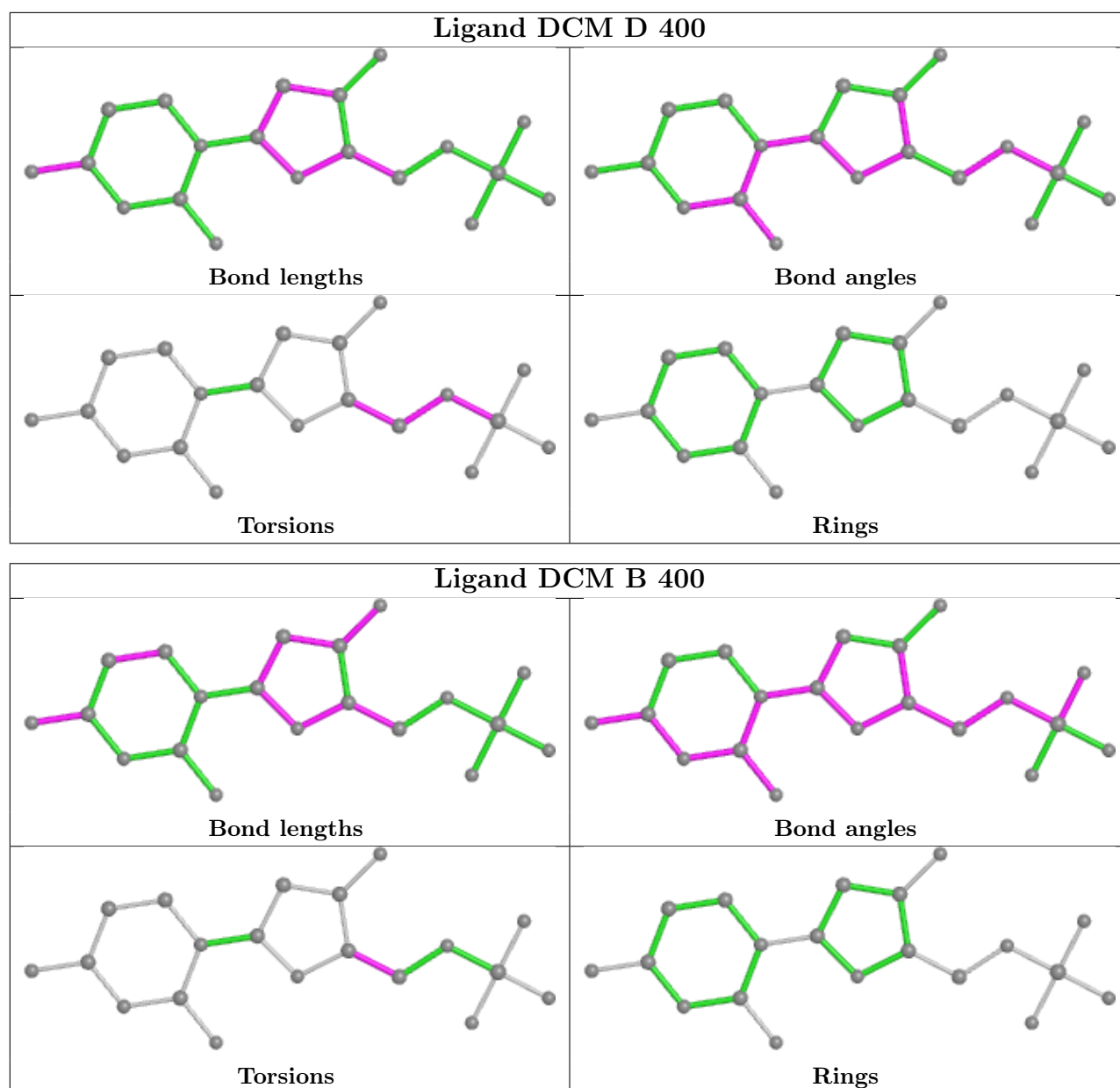
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	400	DCM	2	0
2	C	400	DCM	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







## 5.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [\(i\)](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	284/289 (98%)	-0.30	4 (1%) 75 73	18, 31, 56, 99	0
1	B	282/289 (97%)	-0.21	10 (3%) 44 43	18, 32, 61, 111	0
1	C	282/289 (97%)	-0.16	7 (2%) 57 55	21, 36, 62, 82	0
1	D	286/289 (98%)	-0.15	8 (2%) 53 51	23, 37, 65, 115	0
1	E	286/289 (98%)	-0.24	7 (2%) 59 57	22, 35, 62, 96	0
1	F	277/289 (95%)	-0.27	7 (2%) 57 55	20, 33, 57, 92	0
All	All	1697/1734 (97%)	-0.22	43 (2%) 57 55	18, 34, 62, 115	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	309	MET	6.3
1	B	307	ILE	6.1
1	D	52	GLY	4.7
1	B	53	THR	4.4
1	D	307	ILE	4.2
1	A	52	GLY	4.1
1	D	309	MET	4.0
1	C	51	THR	4.0
1	A	307	ILE	3.8
1	E	50	ARG	3.8
1	B	308	LYS	3.8
1	E	309	MET	3.6
1	F	26	PRO	3.3
1	E	307	ILE	3.2
1	B	26	PRO	3.0
1	E	233	LEU	3.0
1	B	28	HIS	2.9
1	C	233	LEU	2.8
1	F	28	HIS	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	51	THR	2.8
1	F	48	ASP	2.6
1	B	27	PRO	2.6
1	D	308	LYS	2.6
1	D	48	ASP	2.5
1	C	26	PRO	2.5
1	E	25	MET	2.5
1	E	234	THR	2.5
1	E	232	LEU	2.4
1	C	306	THR	2.4
1	D	310	GLU	2.4
1	F	230	TYR	2.2
1	C	48	ASP	2.2
1	B	48	ASP	2.2
1	F	25	MET	2.2
1	F	46	ARG	2.1
1	F	210	CYS	2.1
1	D	305	PRO	2.1
1	A	233	LEU	2.1
1	B	306	THR	2.1
1	A	50	ARG	2.1
1	B	234	THR	2.1
1	C	27	PRO	2.1
1	C	53	THR	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

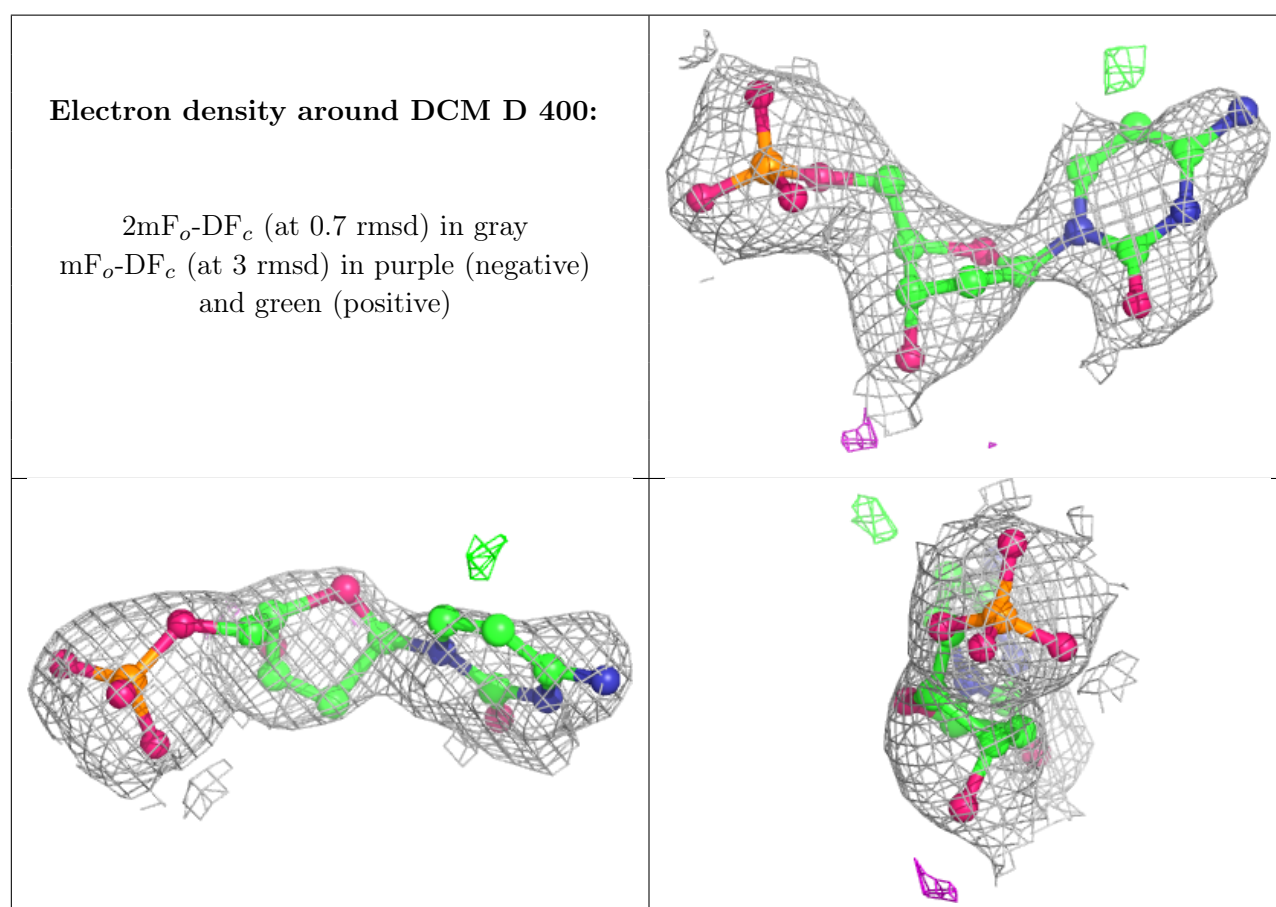
## 6.4 Ligands [i](#)

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



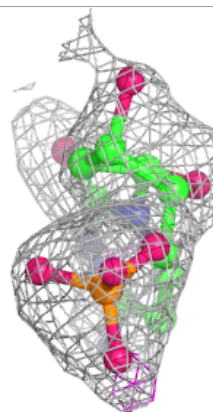
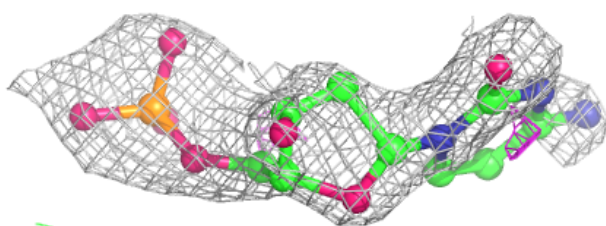
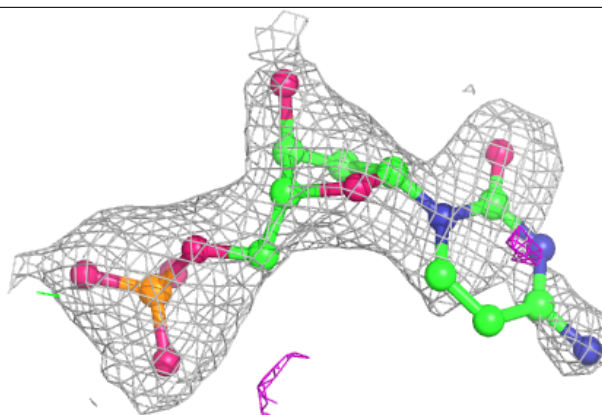
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	DCM	D	400	20/20	0.91	0.17	58,71,95,95	0
2	DCM	F	400	20/20	0.92	0.19	53,76,92,93	0
2	DCM	E	400	20/20	0.93	0.14	43,56,73,77	0
2	DCM	B	400	20/20	0.94	0.14	53,64,80,83	0
2	DCM	A	400	20/20	0.95	0.13	39,51,64,67	0
2	DCM	C	400	20/20	0.96	0.11	40,52,74,76	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.

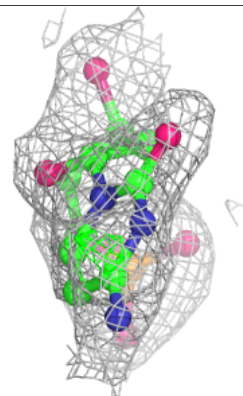
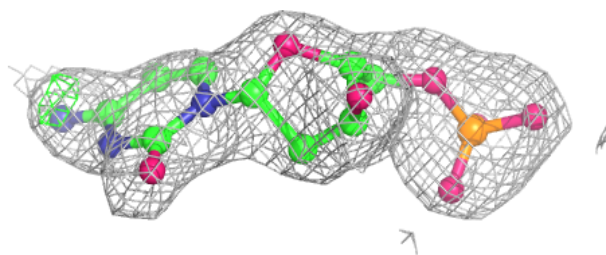
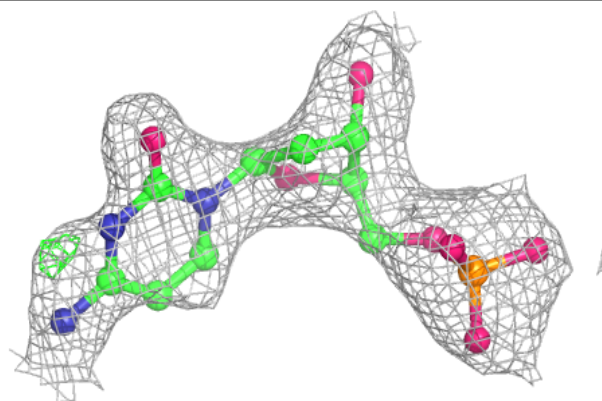


**Electron density around DCM F 400:**

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

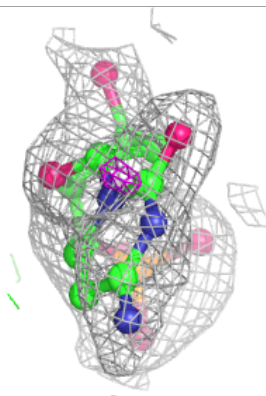
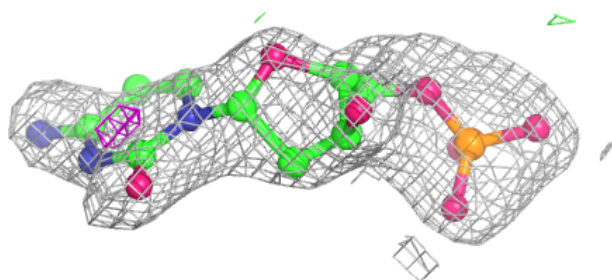
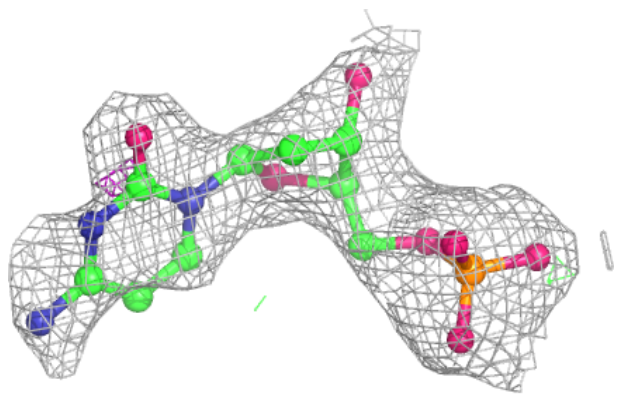
**Electron density around DCM E 400:**

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

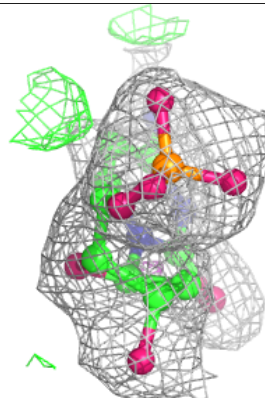
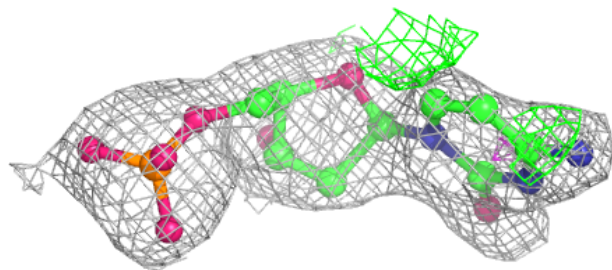
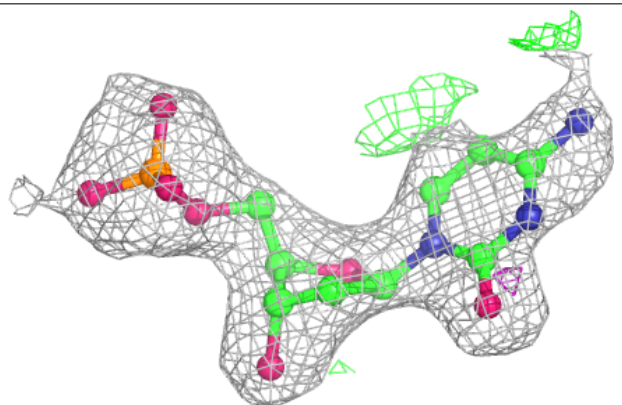


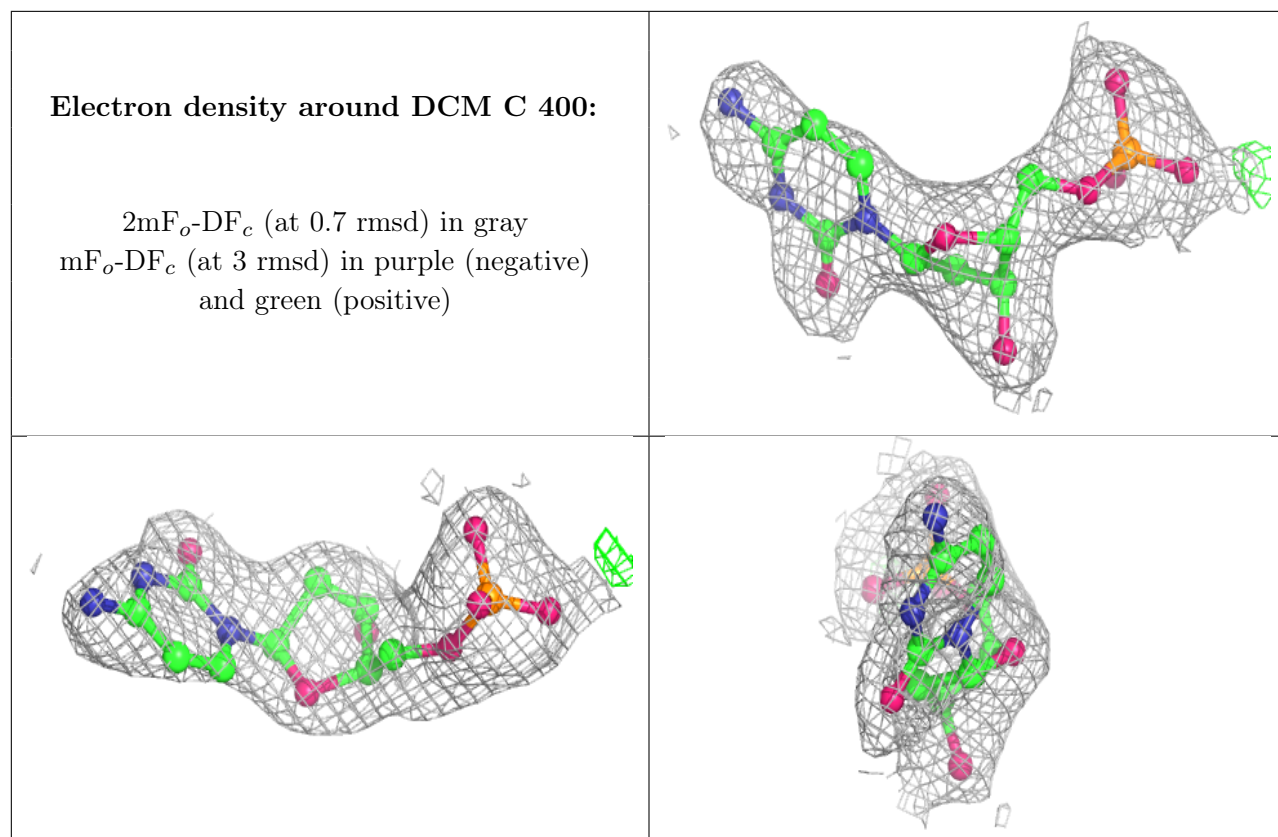
**Electron density around DCM B 400:**

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

**Electron density around DCM A 400:**

$2mF_o-DF_c$  (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.