Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2021 – 07:17 PM EDT

PDB ID	:	1RO7
Title	:	Structural analysis of the sialyltransferase CstII from Campylobacter jejuni in complex with a substrate analogue, CMP-3FNeuAc.
Authors	:	Chiu, C.P.; Watts, A.G.; Lairson, L.L.; Gilbert, M.; Lim, D.; Wakarchuk, W.W.; Withers, S.G.; Strynadka, N.C.
Deposited on	:	2003-12-01
Resolution	:	1.80 Å(reported)

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	:	2.23.2
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)	:	2.23.2

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 1.80 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.
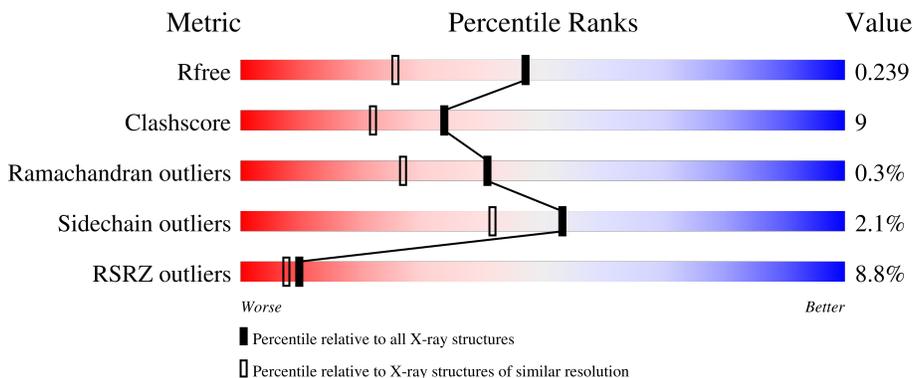
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	259	13% / 82% / 17% ·
1	B	259	6% / 72% / 17% · 8%
1	C	259	8% / 78% / 13% · 6%
1	D	259	6% / 81% / 10% · 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CSF	C	1001	X	-	-	-
3	CSF	C	2001	X	-	-	-
3	CSF	C	3001	X	-	-	-
3	CSF	C	4001	X	-	-	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 8668 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 alpha-2,3/8-sialyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	258	Total	C	N	O	S	Se	0	0	0
			2155	1410	340	396	6	3			
1	B	237	Total	C	N	O	S	Se	0	0	0
			1985	1305	311	360	6	3			
1	C	243	Total	C	N	O	S	Se	0	0	0
			2032	1335	317	371	6	3			
1	D	238	Total	C	N	O	S	Se	0	0	0
			1990	1309	311	361	6	3			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP Q9LAK3
A	53	SER	ILE	engineered mutation	UNP Q9LAK3
A	77	MSE	MET	modified residue	UNP Q9LAK3
A	136	MSE	MET	modified residue	UNP Q9LAK3
B	1	MSE	MET	modified residue	UNP Q9LAK3
B	53	SER	ILE	engineered mutation	UNP Q9LAK3
B	77	MSE	MET	modified residue	UNP Q9LAK3
B	136	MSE	MET	modified residue	UNP Q9LAK3
C	1	MSE	MET	modified residue	UNP Q9LAK3
C	53	SER	ILE	engineered mutation	UNP Q9LAK3
C	77	MSE	MET	modified residue	UNP Q9LAK3
C	136	MSE	MET	modified residue	UNP Q9LAK3
D	1	MSE	MET	modified residue	UNP Q9LAK3
D	53	SER	ILE	engineered mutation	UNP Q9LAK3
D	77	MSE	MET	modified residue	UNP Q9LAK3
D	136	MSE	MET	modified residue	UNP Q9LAK3

- Molecule 2 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).

MPD
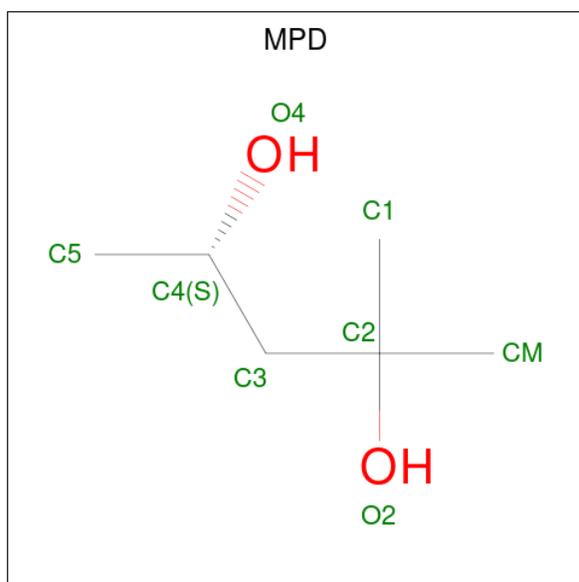
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	C	O	0	0
			8	6	2		

- Molecule 3 is CYTIDINE-5'-MONOPHOSPHATE-3-FLUORO-N-ACETYL-NEURAMINIC ACID (three-letter code: CSF) (formula: $C_{20}H_{30}FN_4O_{16}P$).

CSF
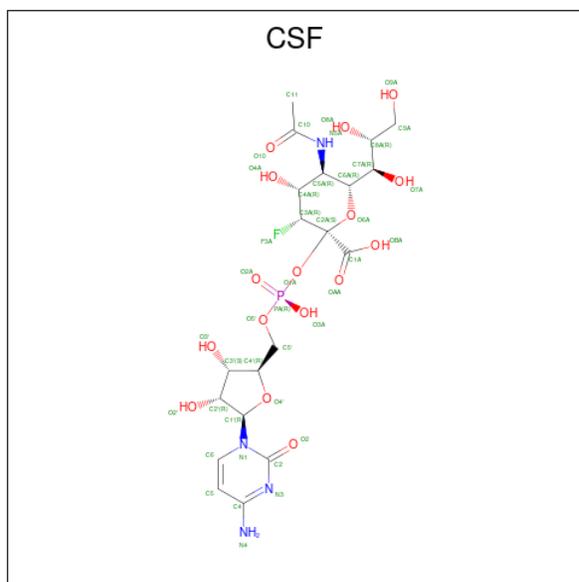
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	C	1	Total	C	F	N	O	P	0	0
			42	20	1	4	16	1		
3	C	1	Total	C	F	N	O	P	0	0
			42	20	1	4	16	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	C	1	Total 42	C 20	F 1	N 4	O 16	P 1	0	0
3	C	1	Total 42	C 20	F 1	N 4	O 16	P 1	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	89	Total 89	O 89	0	0
4	B	85	Total 85	O 85	0	0
4	C	83	Total 83	O 83	0	0
4	D	73	Total 73	O 73	0	0

3　Residue-property plots ⓘ

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: alpha-2,3/8-sialyltransferase
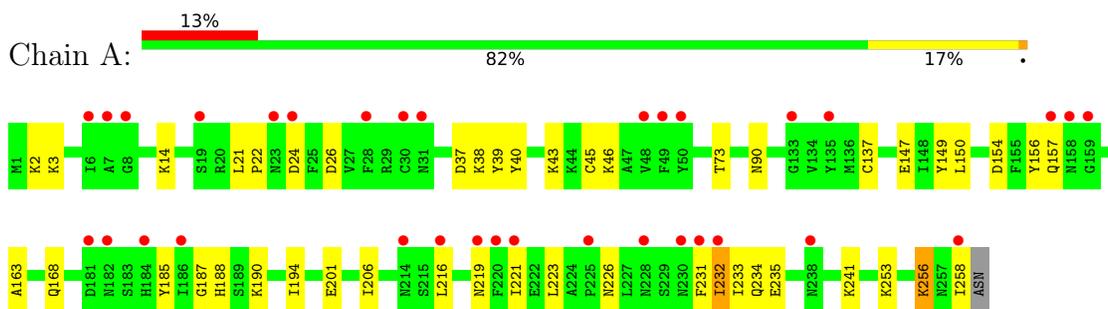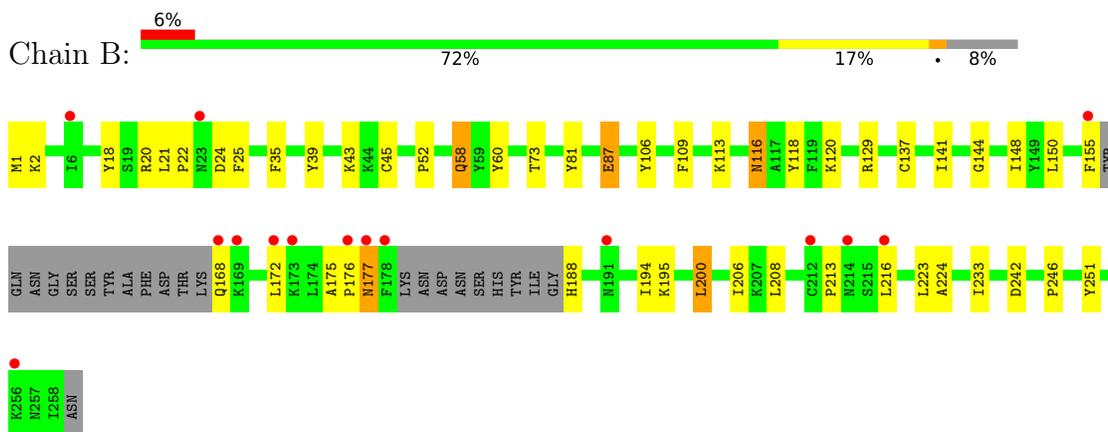
- Molecule 1: alpha-2,3/8-sialyltransferase

- Molecule 1: alpha-2,3/8-sialyltransferase

- Molecule 1: alpha-2,3/8-sialyltransferase

Chain D:
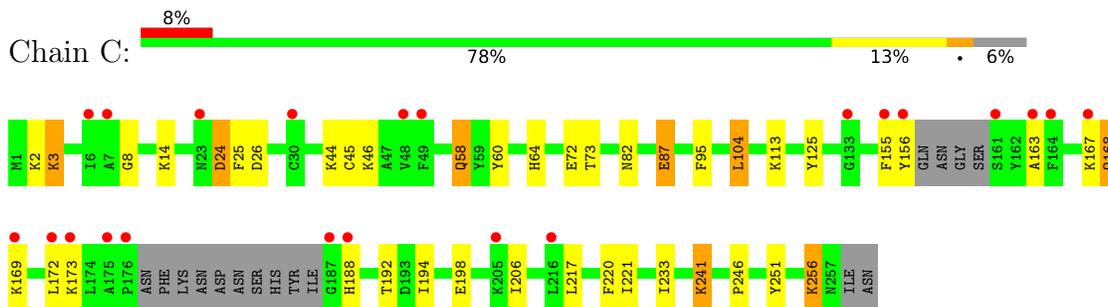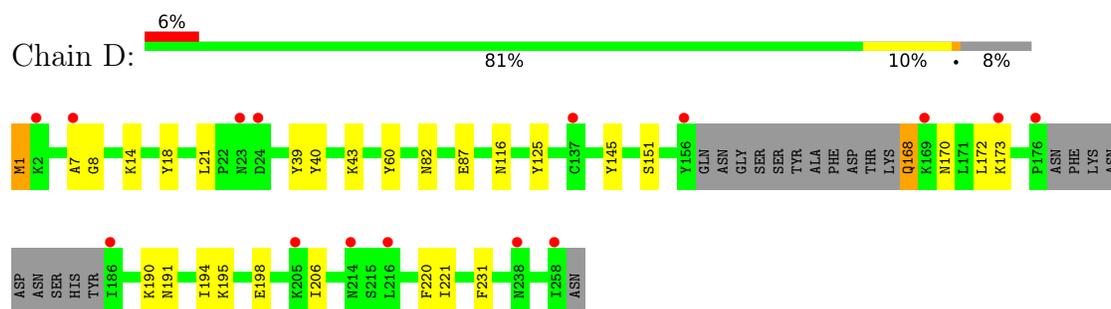
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	83.66Å 66.03Å 99.17Å 90.00° 94.42° 90.00°	Depositor
Resolution (Å)	29.64 – 1.80 29.63 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.4 (29.64-1.80) 99.6 (29.63-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.77 (at 1.80Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.217 , 0.248 0.211 , 0.239	Depositor DCC
R_{free} test set	5034 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	26.1	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 45.6	EDS
L-test for twinning[2]	$< |L| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8668	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.33% 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: CSF, MPD

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	0.35	0/2213	0.54	0/2982
1	B	0.36	0/2036	0.55	0/2740
1	C	0.36	0/2085	0.54	1/2806 (0.0%)
1	D	0.36	0/2041	0.54	1/2747 (0.0%)
All	All	0.36	0/8375	0.54	2/11275 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	8	GLY	N-CA-C	-5.21	100.07	113.10
1	C	8	GLY	N-CA-C	-5.18	100.16	113.10

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	2155	0	2087	41	0
1	B	1985	0	1938	40	0
1	C	2032	0	1976	38	0
1	D	1990	0	1946	22	0
2	B	8	0	14	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	168	0	112	5	0
4	A	89	0	0	0	0
4	B	85	0	0	0	0
4	C	83	0	0	3	0
4	D	73	0	0	0	0
All	All	8668	0	8073	140	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

All (140) 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:C:256:LYS:HD2	1:C:256:LYS:H	1.40	0.86
1:C:44:LYS:HD2	1:C:72:GLU:HG3	1.62	0.82
1:C:14:LYS:HG3	1:C:163:ALA:HB1	1.62	0.81
1:D:168:GLN:HE21	1:D:168:GLN:N	1.82	0.76
1:A:154:ASP:HB3	1:A:157:GLN:HE21	1.51	0.75
1:C:14:LYS:HD2	1:C:163:ALA:O	1.87	0.74
1:B:1:MSE:HG3	1:B:144:GLY:O	1.89	0.73
1:C:246:PRO:HG2	1:C:251:TYR:CE1	2.25	0.71
1:A:256:LYS:NZ	1:A:256:LYS:H	1.88	0.70
1:A:14:LYS:HG3	1:A:163:ALA:HB1	1.73	0.69
1:C:194:ILE:O	1:C:198:GLU:HG3	1.93	0.69
1:B:118:TYR:OH	1:B:195:LYS:HG2	1.94	0.68
1:A:185:TYR:HB3	1:A:188:HIS:CD2	2.29	0.67
1:D:170:ASN:O	1:D:173:LYS:HG2	1.96	0.66
1:A:233:ILE:HD12	1:A:233:ILE:N	2.10	0.65
1:D:194:ILE:O	1:D:198:GLU:HG3	1.96	0.65
1:D:168:GLN:O	1:D:172:LEU:HD23	1.97	0.65
1:D:168:GLN:N	1:D:168:GLN:NE2	2.44	0.64
1:B:129:ARG:HH21	1:B:188:HIS:HE1	1.47	0.62
1:D:221:ILE:N	1:D:221:ILE:HD12	2.14	0.62
1:B:58:GLN:HA	1:B:58:GLN:HE21	1.66	0.61
1:C:188:HIS:O	1:C:192:THR:HB	2.00	0.61
1:C:155:PHE:CZ	1:C:194:ILE:HD11	2.35	0.61
1:A:2:LYS:HG2	1:A:24:ASP:OD2	2.01	0.60
1:D:14:LYS:HG2	1:D:40:TYR:CZ	2.36	0.60
1:A:256:LYS:H	1:A:256:LYS:HD2	1.64	0.60
1:A:190:LYS:HE2	1:A:216:LEU:HD21	1.82	0.60
1:C:256:LYS:H	1:C:256:LYS:CD	2.14	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:39:TYR:CE2	1:D:43:LYS:HG2	2.38	0.59
1:A:231:PHE:O	1:A:232:ILE:HB	2.03	0.58
1:C:58:GLN:HE21	1:C:58:GLN:HA	1.68	0.58
1:A:256:LYS:H	1:A:256:LYS:HZ3	1.51	0.58
1:C:241:LYS:HD3	4:C:4084:HOH:O	2.04	0.57
1:B:81:TYR:N	1:B:87:GLU:OE1	2.35	0.57
1:B:155:PHE:CZ	1:B:194:ILE:HD11	2.40	0.57
1:C:188:HIS:HA	4:C:4048:HOH:O	2.04	0.57
1:D:1:MSE:HB2	1:D:145:TYR:CE1	2.39	0.56
1:A:256:LYS:H	1:A:256:LYS:CD	2.19	0.56
1:B:18:TYR:CE2	1:B:233:ILE:HD11	2.39	0.56
1:B:35:PHE:CE2	1:B:168:GLN:HG2	2.40	0.56
1:A:216:LEU:HD12	1:A:219:ASN:HD22	1.71	0.55
1:A:37:ASP:OD2	1:A:38:LYS:HG3	2.07	0.55
1:A:258:ILE:HG13	1:A:258:ILE:O	2.06	0.54
1:A:253:LYS:O	1:A:256:LYS:HE2	2.07	0.54
1:B:52:PRO:HG2	1:B:87:GLU:HG2	1.89	0.53
1:B:206:ILE:HD12	1:B:206:ILE:C	2.29	0.53
1:C:44:LYS:CD	1:C:72:GLU:HG3	2.38	0.53
3:C:3001:CSF:H3'	3:C:3001:CSF:O3A	2.09	0.52
1:D:170:ASN:HA	1:D:173:LYS:HE3	1.91	0.52
1:B:113:LYS:HB3	1:B:113:LYS:HZ3	1.75	0.52
1:D:190:LYS:O	1:D:194:ILE:HG12	2.10	0.52
1:C:221:ILE:HD12	1:C:221:ILE:N	2.24	0.52
1:B:129:ARG:HH21	1:B:188:HIS:CE1	2.27	0.52
1:B:137:CYS:SG	1:B:150:LEU:HD21	2.50	0.52
1:C:241:LYS:O	1:C:241:LYS:HG2	2.10	0.51
1:D:191:ASN:HB3	1:D:195:LYS:NZ	2.24	0.51
1:B:2:LYS:HE2	1:B:25:PHE:C	2.30	0.51
1:A:39:TYR:CE2	1:A:43:LYS:HG2	2.45	0.51
1:B:206:ILE:HD12	1:B:206:ILE:O	2.11	0.51
1:A:206:ILE:C	1:A:206:ILE:HD12	2.31	0.51
1:C:2:LYS:HE3	1:C:24:ASP:OD1	2.10	0.51
1:A:256:LYS:HD2	1:A:256:LYS:N	2.26	0.50
1:D:220:PHE:HB2	1:D:221:ILE:HD12	1.93	0.50
1:A:14:LYS:HG2	1:A:40:TYR:CZ	2.46	0.50
1:A:21:LEU:HD12	1:A:22:PRO:HD2	1.94	0.50
1:A:233:ILE:HD12	1:A:233:ILE:H	1.74	0.50
1:A:256:LYS:NZ	1:A:256:LYS:N	2.59	0.50
1:C:2:LYS:HE2	1:C:25:PHE:C	2.32	0.50
1:A:201:GLU:HG3	1:A:206:ILE:O	2.11	0.49

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:45:CYS:O	1:A:73:THR:HA	2.12	0.49
1:A:185:TYR:CE1	1:A:187:GLY:HA3	2.47	0.49
1:B:2:LYS:HE3	1:B:24:ASP:CG	2.32	0.49
1:C:156:TYR:OH	3:C:3001:CSF:H3A	2.13	0.49
1:C:156:TYR:HA	3:C:3001:CSF:H41	1.77	0.49
1:B:60:TYR:HB2	1:D:125:TYR:CE2	2.46	0.49
1:C:256:LYS:HD2	1:C:256:LYS:N	2.19	0.48
1:B:216:LEU:HD22	1:B:216:LEU:N	2.29	0.48
1:A:156:TYR:HA	3:C:1001:CSF:H41	1.77	0.48
1:D:206:ILE:C	1:D:206:ILE:HD12	2.33	0.47
1:C:168:GLN:O	1:C:172:LEU:HG	2.13	0.47
1:B:113:LYS:HB3	1:B:113:LYS:NZ	2.29	0.47
1:A:190:LYS:O	1:A:194:ILE:HG12	2.15	0.47
1:B:216:LEU:HD22	1:B:216:LEU:H	1.80	0.46
1:A:3:LYS:HE3	1:A:147:GLU:OE1	2.16	0.46
1:B:116:ASN:HD21	1:B:120:LYS:NZ	2.14	0.46
1:B:246:PRO:HG2	1:B:251:TYR:CE2	2.51	0.45
1:A:43:LYS:HE3	1:A:234:GLN:OE1	2.16	0.45
1:D:21:LEU:HD12	1:D:231:PHE:HB2	1.99	0.45
1:B:1:MSE:HA	1:B:1:MSE:HE2	1.98	0.45
1:A:216:LEU:HD12	1:A:219:ASN:ND2	2.32	0.45
1:B:21:LEU:HD12	1:B:22:PRO:HD2	1.99	0.45
1:B:168:GLN:HB3	1:B:242:ASP:OD2	2.17	0.45
1:D:190:LYS:HD3	1:D:220:PHE:CE2	2.52	0.45
1:A:221:ILE:HD12	1:A:221:ILE:N	2.32	0.45
1:B:141:ILE:HD12	1:B:200:LEU:HG	1.98	0.44
1:A:26:ASP:HB3	1:A:46:LYS:HB2	2.00	0.44
1:A:3:LYS:HD3	1:A:149:TYR:CE2	2.53	0.44
1:B:148:ILE:HB	1:B:208:LEU:HD23	1.99	0.44
1:C:217:LEU:CD1	1:C:221:ILE:HD13	2.48	0.44
1:B:39:TYR:CZ	1:B:43:LYS:HG2	2.53	0.43
1:C:60:TYR:CE2	1:C:64:HIS:HE1	2.36	0.43
1:D:7:ALA:HA	1:D:151:SER:O	2.18	0.43
1:B:168:GLN:O	1:B:172:LEU:HG	2.17	0.43
1:A:137:CYS:SG	1:A:150:LEU:HD21	2.59	0.43
1:A:223:LEU:HD12	1:A:223:LEU:N	2.33	0.43
1:C:95:PHE:CG	1:C:104:LEU:HD13	2.53	0.43
1:C:206:ILE:C	1:C:206:ILE:HD12	2.39	0.43
1:C:95:PHE:CD2	1:C:104:LEU:HD13	2.52	0.43
1:C:168:GLN:HE21	1:C:241:LYS:NZ	2.16	0.43
1:B:106:TYR:HA	1:B:109:PHE:HB3	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:3:LYS:H	1:C:3:LYS:HD2	1.84	0.43
1:B:45:CYS:O	1:B:73:THR:HA	2.18	0.43
1:C:82:ASN:N	1:C:87:GLU:OE1	2.52	0.42
1:C:45:CYS:O	1:C:73:THR:HA	2.19	0.42
1:B:177:ASN:ND2	1:B:177:ASN:C	2.72	0.42
1:B:2:LYS:HE2	1:B:25:PHE:O	2.18	0.42
1:C:113:LYS:HD2	4:C:4028:HOH:O	2.19	0.42
1:B:113:LYS:NZ	1:B:113:LYS:CB	2.83	0.42
1:B:213:PRO:HB3	1:B:223:LEU:HD21	2.02	0.42
1:A:39:TYR:CZ	1:A:43:LYS:HG2	2.54	0.42
1:B:216:LEU:H	1:B:216:LEU:CD2	2.33	0.41
1:A:156:TYR:HA	3:C:1001:CSF:N4	2.36	0.41
1:A:190:LYS:HE2	1:A:216:LEU:CD2	2.49	0.41
1:A:168:GLN:HG3	1:A:241:LYS:O	2.20	0.41
1:A:43:LYS:CE	1:A:234:GLN:HE22	2.34	0.41
1:C:26:ASP:HB3	1:C:46:LYS:HB2	2.03	0.41
1:C:104:LEU:HD12	1:C:104:LEU:HA	1.87	0.41
1:C:125:TYR:CE2	1:D:60:TYR:HB2	2.55	0.41
1:C:220:PHE:HB2	1:C:221:ILE:HD12	2.03	0.41
1:D:82:ASN:N	1:D:87:GLU:OE2	2.48	0.41
1:B:177:ASN:C	1:B:177:ASN:HD22	2.23	0.41
1:A:201:GLU:HA	1:A:206:ILE:CD1	2.51	0.41
1:B:20:ARG:HD3	1:B:224:ALA:O	2.21	0.41
1:C:233:ILE:N	1:C:233:ILE:HD12	2.36	0.40
1:D:39:TYR:CZ	1:D:43:LYS:HG2	2.56	0.40
1:B:35:PHE:CZ	1:B:168:GLN:HG2	2.57	0.40
1:C:168:GLN:HG2	1:C:241:LYS:O	2.22	0.40
1:D:18:TYR:HD2	1:D:21:LEU:HD11	1.85	0.40
1:B:175:ALA:HA	1:B:176:PRO:HD2	1.96	0.40
1:C:169:LYS:O	1:C:173:LYS:HG3	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	256/259 (99%)	239 (93%)	15 (6%)	2 (1%)	19	7
1	B	231/259 (89%)	225 (97%)	6 (3%)	0	100	100
1	C	237/259 (92%)	229 (97%)	7 (3%)	1 (0%)	34	21
1	D	232/259 (90%)	225 (97%)	7 (3%)	0	100	100
All	All	956/1036 (92%)	918 (96%)	35 (4%)	3 (0%)	41	27

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	24	ASP
1	A	226	ASN
1	A	232	ILE

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	234/232 (101%)	231 (99%)	3 (1%)	69	62
1	B	216/232 (93%)	211 (98%)	5 (2%)	50	37
1	C	220/232 (95%)	212 (96%)	8 (4%)	35	20
1	D	216/232 (93%)	213 (99%)	3 (1%)	67	59
All	All	886/928 (96%)	867 (98%)	19 (2%)	53	42

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	90	ASN
1	A	235	GLU
1	A	256	LYS
1	B	58	GLN
1	B	87	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	116	ASN
1	B	177	ASN
1	B	200	LEU
1	C	3	LYS
1	C	58	GLN
1	C	87	GLU
1	C	104	LEU
1	C	167	LYS
1	C	168	GLN
1	C	241	LYS
1	C	256	LYS
1	D	1	MSE
1	D	116	ASN
1	D	168	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (31) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	67	GLN
1	A	69	GLN
1	A	128	GLN
1	A	157	GLN
1	A	158	ASN
1	A	177	ASN
1	A	180	ASN
1	A	219	ASN
1	A	228	ASN
1	A	257	ASN
1	B	51	ASN
1	B	58	GLN
1	B	111	GLN
1	B	116	ASN
1	B	128	GLN
1	B	177	ASN
1	B	188	HIS
1	C	32	GLN
1	C	58	GLN
1	C	67	GLN
1	C	90	ASN
1	C	128	GLN
1	C	168	GLN
1	C	228	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	234	GLN
1	D	83	GLN
1	D	116	ASN
1	D	128	GLN
1	D	188	HIS
1	D	219	ASN
1	D	238	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

5 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	MPD	B	401	-	7,7,7	0.51	0	9,10,10	0.49	0
3	CSF	C	3001	-	36,44,44	2.62	10 (27%)	42,67,67	2.18	15 (35%)
3	CSF	C	4001	-	36,44,44	2.68	10 (27%)	42,67,67	2.04	13 (30%)
3	CSF	C	2001	-	36,44,44	2.66	11 (30%)	42,67,67	2.10	11 (26%)
3	CSF	C	1001	-	36,44,44	2.56	10 (27%)	42,67,67	2.02	12 (28%)

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	MPD	B	401	-	-	0/5/5/5	-
3	CSF	C	3001	-	1/1/14/15	2/25/75/75	0/3/3/3
3	CSF	C	4001	-	1/1/14/15	2/25/75/75	0/3/3/3
3	CSF	C	2001	-	1/1/14/15	2/25/75/75	0/3/3/3
3	CSF	C	1001	-	1/1/14/15	0/25/75/75	0/3/3/3

All (41) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	4001	CSF	C6-N1	9.18	1.47	1.35
3	C	2001	CSF	C6-N1	8.86	1.46	1.35
3	C	3001	CSF	C6-N1	8.71	1.46	1.35
3	C	1001	CSF	C6-N1	8.70	1.46	1.35
3	C	2001	CSF	C3A-C4A	6.58	1.58	1.52
3	C	3001	CSF	C3A-C4A	6.36	1.58	1.52
3	C	4001	CSF	C3A-C4A	6.27	1.58	1.52
3	C	1001	CSF	C3A-C4A	5.72	1.58	1.52
3	C	4001	CSF	C5-C4	5.12	1.53	1.41
3	C	2001	CSF	C2-N3	5.02	1.48	1.38
3	C	3001	CSF	C2-N3	4.81	1.47	1.38
3	C	2001	CSF	C5-C4	4.80	1.52	1.41
3	C	1001	CSF	C5-C4	4.78	1.52	1.41
3	C	3001	CSF	C5-C4	4.64	1.52	1.41
3	C	4001	CSF	C2-N3	4.63	1.47	1.38
3	C	1001	CSF	C2-N3	4.55	1.47	1.38
3	C	4001	CSF	O4'-C1'	4.01	1.46	1.41
3	C	1001	CSF	O1A-C2A	3.93	1.51	1.43
3	C	2001	CSF	C4-N3	3.88	1.41	1.35
3	C	3001	CSF	C4-N3	3.83	1.41	1.35
3	C	3001	CSF	O4'-C1'	3.71	1.46	1.41
3	C	3001	CSF	PA-O3A	-3.69	1.38	1.55
3	C	1001	CSF	C4-N3	3.64	1.41	1.35
3	C	4001	CSF	PA-O3A	-3.57	1.38	1.55
3	C	2001	CSF	PA-O3A	-3.52	1.38	1.55
3	C	4001	CSF	C4-N3	3.50	1.41	1.35
3	C	3001	CSF	O1A-C2A	3.30	1.50	1.43
3	C	1001	CSF	PA-O3A	-3.10	1.40	1.55
3	C	1001	CSF	O4'-C1'	3.06	1.45	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2001	CSF	O4'-C1'	3.05	1.45	1.41
3	C	4001	CSF	O1A-C2A	2.91	1.49	1.43
3	C	2001	CSF	O1A-C2A	2.84	1.49	1.43
3	C	2001	CSF	C6A-C5A	2.73	1.57	1.53
3	C	4001	CSF	C6A-C5A	2.46	1.57	1.53
3	C	3001	CSF	PA-O1A	2.29	1.66	1.60
3	C	2001	CSF	F3A-C3A	-2.26	1.36	1.40
3	C	3001	CSF	C6A-C5A	2.23	1.56	1.53
3	C	1001	CSF	C5A-N5A	-2.21	1.42	1.45
3	C	4001	CSF	C5A-N5A	-2.19	1.42	1.45
3	C	1001	CSF	PA-O1A	2.17	1.66	1.60
3	C	2001	CSF	PA-O1A	2.05	1.65	1.60

All (51) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	2001	CSF	O6A-C6A-C7A	-6.09	97.89	107.29
3	C	3001	CSF	O6A-C6A-C7A	-6.01	98.03	107.29
3	C	4001	CSF	C2-N3-C4	5.94	122.36	116.34
3	C	1001	CSF	C2-N3-C4	5.65	122.07	116.34
3	C	3001	CSF	C2-N3-C4	5.64	122.06	116.34
3	C	1001	CSF	O6A-C6A-C7A	-5.58	98.68	107.29
3	C	2001	CSF	C2-N3-C4	5.55	121.97	116.34
3	C	4001	CSF	O6A-C6A-C7A	-5.35	99.03	107.29
3	C	3001	CSF	O5'-C5'-C4'	4.49	124.43	108.99
3	C	4001	CSF	C8A-C7A-C6A	4.37	121.32	113.03
3	C	1001	CSF	C8A-C7A-C6A	4.36	121.30	113.03
3	C	2001	CSF	C8A-C7A-C6A	4.15	120.91	113.03
3	C	3001	CSF	C3A-C4A-C5A	3.98	114.96	109.83
3	C	4001	CSF	C3A-C4A-C5A	3.93	114.91	109.83
3	C	3001	CSF	C8A-C7A-C6A	3.81	120.25	113.03
3	C	2001	CSF	C3A-C4A-C5A	3.63	114.51	109.83
3	C	2001	CSF	C3'-C2'-C1'	3.48	106.21	100.98
3	C	1001	CSF	C3'-C2'-C1'	3.02	105.52	100.98
3	C	1001	CSF	C4A-C5A-N5A	3.00	116.28	110.62
3	C	1001	CSF	C3A-C4A-C5A	2.90	113.57	109.83
3	C	2001	CSF	C4A-C5A-N5A	2.90	116.09	110.62
3	C	2001	CSF	O8A-C8A-C9A	-2.76	102.68	109.14
3	C	3001	CSF	C4A-C5A-N5A	2.71	115.72	110.62
3	C	4001	CSF	O8A-C8A-C9A	-2.67	102.88	109.14
3	C	1001	CSF	O6A-C6A-C5A	2.67	112.38	109.78
3	C	3001	CSF	O6A-C6A-C5A	2.67	112.38	109.78

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	3001	CSF	O8A-C8A-C9A	-2.62	103.00	109.14
3	C	3001	CSF	O1A-PA-O2A	2.57	118.70	109.29
3	C	3001	CSF	C5A-N5A-C10	2.55	129.39	123.18
3	C	1001	CSF	O8A-C8A-C9A	-2.55	103.16	109.14
3	C	3001	CSF	C3'-C2'-C1'	2.55	104.82	100.98
3	C	4001	CSF	C4A-C5A-N5A	2.53	115.39	110.62
3	C	4001	CSF	C3'-C2'-C1'	2.51	104.75	100.98
3	C	4001	CSF	C5A-N5A-C10	2.47	129.18	123.18
3	C	1001	CSF	C6A-C5A-N5A	-2.41	106.91	110.91
3	C	1001	CSF	C5A-N5A-C10	2.39	129.00	123.18
3	C	2001	CSF	O5'-PA-O2A	-2.32	100.01	109.07
3	C	4001	CSF	O5'-PA-O2A	-2.25	100.27	109.07
3	C	2001	CSF	C5A-N5A-C10	2.24	128.62	123.18
3	C	4001	CSF	O4A-C4A-C5A	-2.18	105.25	109.66
3	C	3001	CSF	O3A-PA-O2A	-2.18	101.47	112.24
3	C	1001	CSF	O5'-PA-O2A	-2.17	100.58	109.07
3	C	1001	CSF	O4A-C4A-C5A	-2.12	105.37	109.66
3	C	2001	CSF	O6A-C6A-C5A	2.12	111.85	109.78
3	C	2001	CSF	C6A-C5A-N5A	-2.10	107.42	110.91
3	C	4001	CSF	C6A-C5A-N5A	-2.05	107.50	110.91
3	C	4001	CSF	O5'-C5'-C4'	2.04	116.03	108.99
3	C	3001	CSF	C6A-C5A-N5A	-2.03	107.54	110.91
3	C	3001	CSF	O4A-C4A-C5A	-2.02	105.57	109.66
3	C	4001	CSF	O6A-C6A-C5A	2.01	111.74	109.78
3	C	3001	CSF	O5'-PA-O2A	-2.01	101.23	109.07

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	1001	CSF	C2A
3	C	2001	CSF	C2A
3	C	3001	CSF	C2A
3	C	4001	CSF	C2A

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	3001	CSF	C4'-C5'-O5'-PA
3	C	4001	CSF	C7A-C8A-C9A-O9A
3	C	4001	CSF	O8A-C8A-C9A-O9A
3	C	2001	CSF	C2A-O1A-PA-O2A
3	C	2001	CSF	C4'-C5'-O5'-PA

Continued on next page...

Continued from previous page...

| Mol	Chain	Res	Type	Atoms
3	C	3001	CSF	C5'-O5'-PA-O1A

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	3001	CSF	3	0
3	C	1001	CSF	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 CSF C 3001

Bond lengths

Bond angles

Torsions

Rings

Ligand CSF C 4001

Bond lengths

Bond angles

Torsions

Rings

Ligand CSF C 2001

Bond lengths

Bond angles

Torsions

Rings

Ligand CSF C 1001

Bond lengths

Bond angles

Torsions

Rings
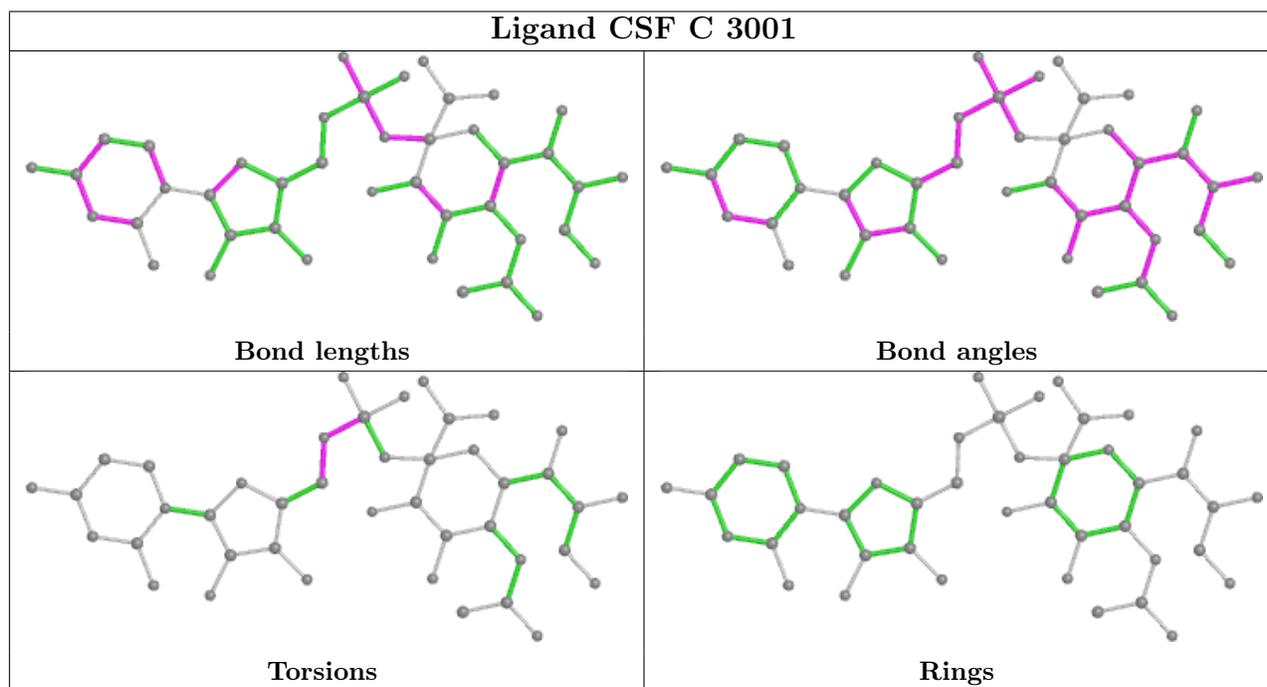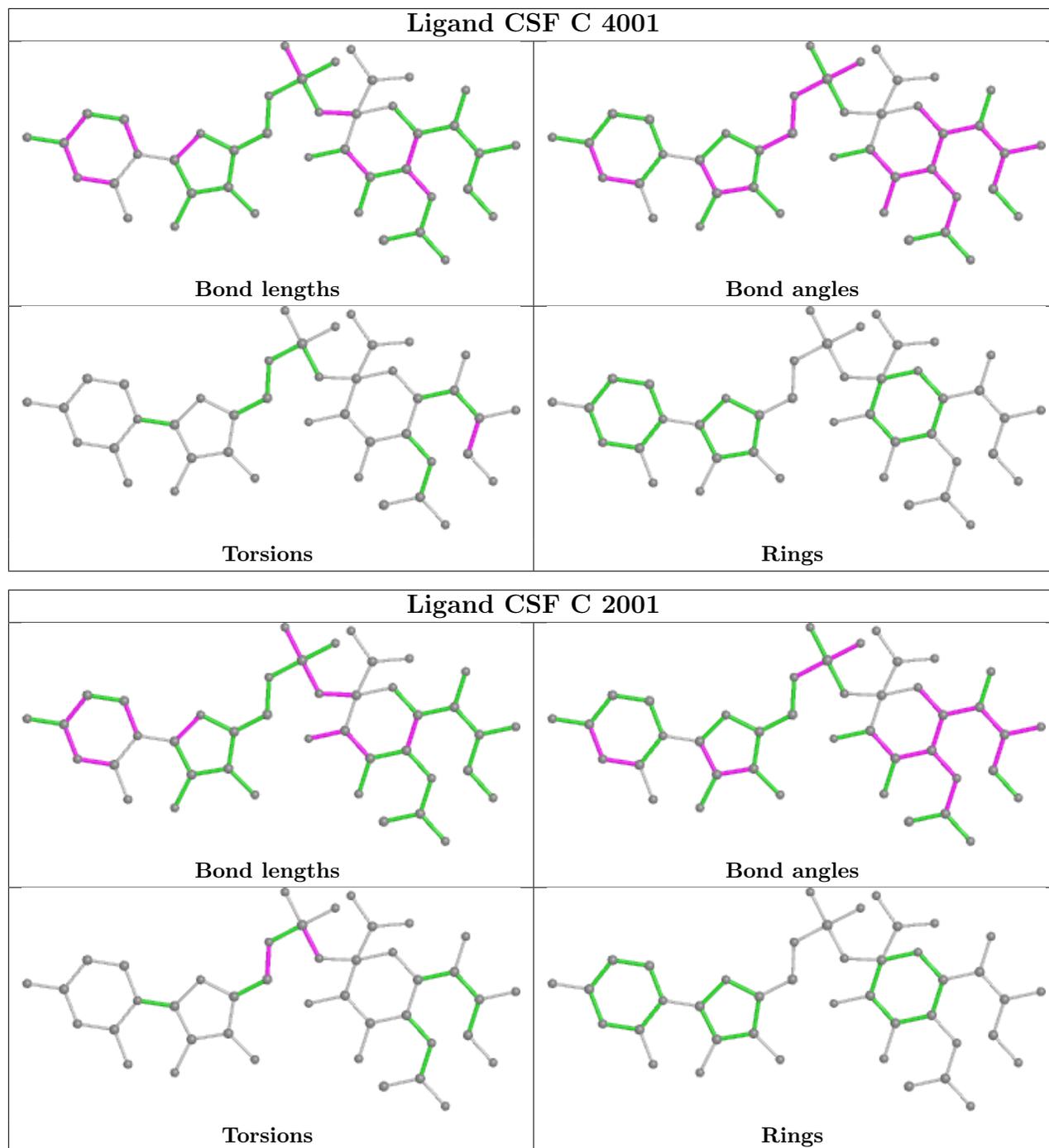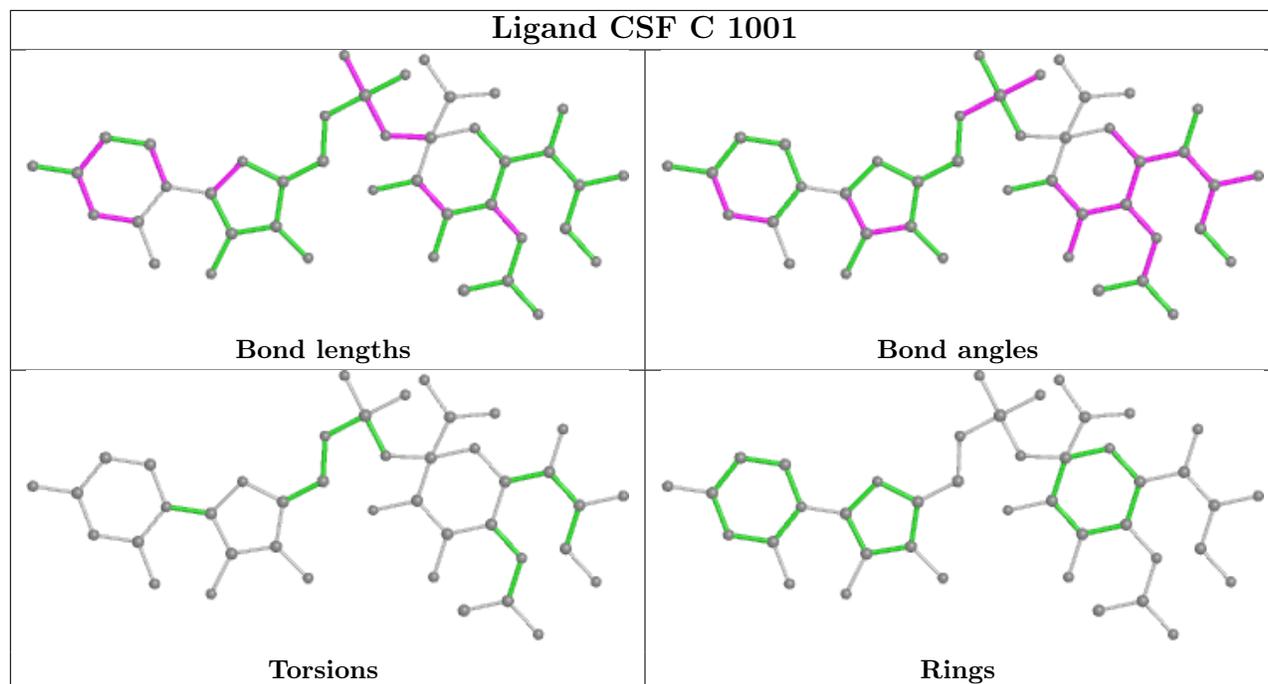
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	255/259 (98%)	0.60	33 (12%)	3	2	17, 34, 63, 77	0
1	B	234/259 (90%)	0.33	15 (6%)	19	15	18, 29, 51, 63	0
1	C	240/259 (92%)	0.38	22 (9%)	9	6	21, 30, 56, 73	0
1	D	235/259 (90%)	0.38	15 (6%)	19	15	20, 32, 53, 65	0
All	All	964/1036 (93%)	0.43	85 (8%)	10	7	17, 31, 57, 77	0

All (85) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	156	TYR	7.5
1	B	178	PHE	7.1
1	C	187	GLY	6.5
1	B	23	ASN	5.6
1	A	158	ASN	5.5
1	D	186	ILE	5.5
1	B	176	PRO	5.1
1	A	230	ASN	5.1
1	A	23	ASN	4.9
1	B	172	LEU	4.8
1	A	6	ILE	4.7
1	C	188	HIS	4.7
1	C	164	PHE	4.0
1	A	184	HIS	3.9
1	A	228	ASN	3.9
1	C	163	ALA	3.9
1	A	24	ASP	3.8
1	B	177	ASN	3.8
1	A	30	CYS	3.6
1	D	23	ASN	3.5
1	C	169	LYS	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	49	PHE	3.4
1	D	214	ASN	3.3
1	A	219	ASN	3.3
1	A	258	ILE	3.3
1	A	159	GLY	3.3
1	A	232	ILE	3.3
1	D	238	ASN	3.2
1	D	2	LYS	3.2
1	A	231	PHE	3.2
1	B	169	LYS	3.1
1	A	220	PHE	3.1
1	C	173	LYS	3.1
1	D	176	PRO	3.0
1	A	28	PHE	3.0
1	C	156	TYR	2.9
1	B	214	ASN	2.9
1	A	238	ASN	2.9
1	A	7	ALA	2.8
1	A	48	VAL	2.8
1	C	161	SER	2.8
1	C	167	LYS	2.7
1	B	168	GLN	2.7
1	B	212	CYS	2.7
1	C	48	VAL	2.7
1	D	137	CYS	2.7
1	C	172	LEU	2.7
1	B	216	LEU	2.6
1	D	216	LEU	2.6
1	A	157	GLN	2.6
1	A	19	SER	2.6
1	C	6	ILE	2.6
1	C	23	ASN	2.5
1	B	256	LYS	2.5
1	D	169	LYS	2.5
1	D	24	ASP	2.4
1	A	135	TYR	2.4
1	C	155	PHE	2.4
1	A	133	GLY	2.4
1	D	258	ILE	2.4
1	C	30	CYS	2.3
1	C	49	PHE	2.3
1	A	181	ASP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	186	ILE	2.3
1	B	191	ASN	2.3
1	B	6	ILE	2.3
1	A	214	ASN	2.3
1	C	133	GLY	2.3
1	C	216	LEU	2.3
1	A	182	ASN	2.2
1	D	7	ALA	2.2
1	B	173	LYS	2.2
1	A	221	ILE	2.2
1	C	176	PRO	2.2
1	D	173	LYS	2.2
1	C	205	LYS	2.2
1	D	205	LYS	2.2
1	C	175	ALA	2.1
1	C	7	ALA	2.1
1	A	216	LEU	2.1
1	A	50	TYR	2.1
1	A	31	ASN	2.0
1	B	155	PHE	2.0
1	A	8	GLY	2.0
1	A	225	PRO	2.0

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CSF	C	2001	42/42	0.80	0.20	39,49,60,61	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CSF	C	3001	42/42	0.81	0.18	36,43,48,49	0
2	MPD	B	401	8/8	0.82	0.17	40,46,47,48	0
3	CSF	C	4001	42/42	0.87	0.15	37,41,45,49	0
3	CSF	C	1001	42/42	0.91	0.14	21,29,37,40	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 CSF C 2001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)
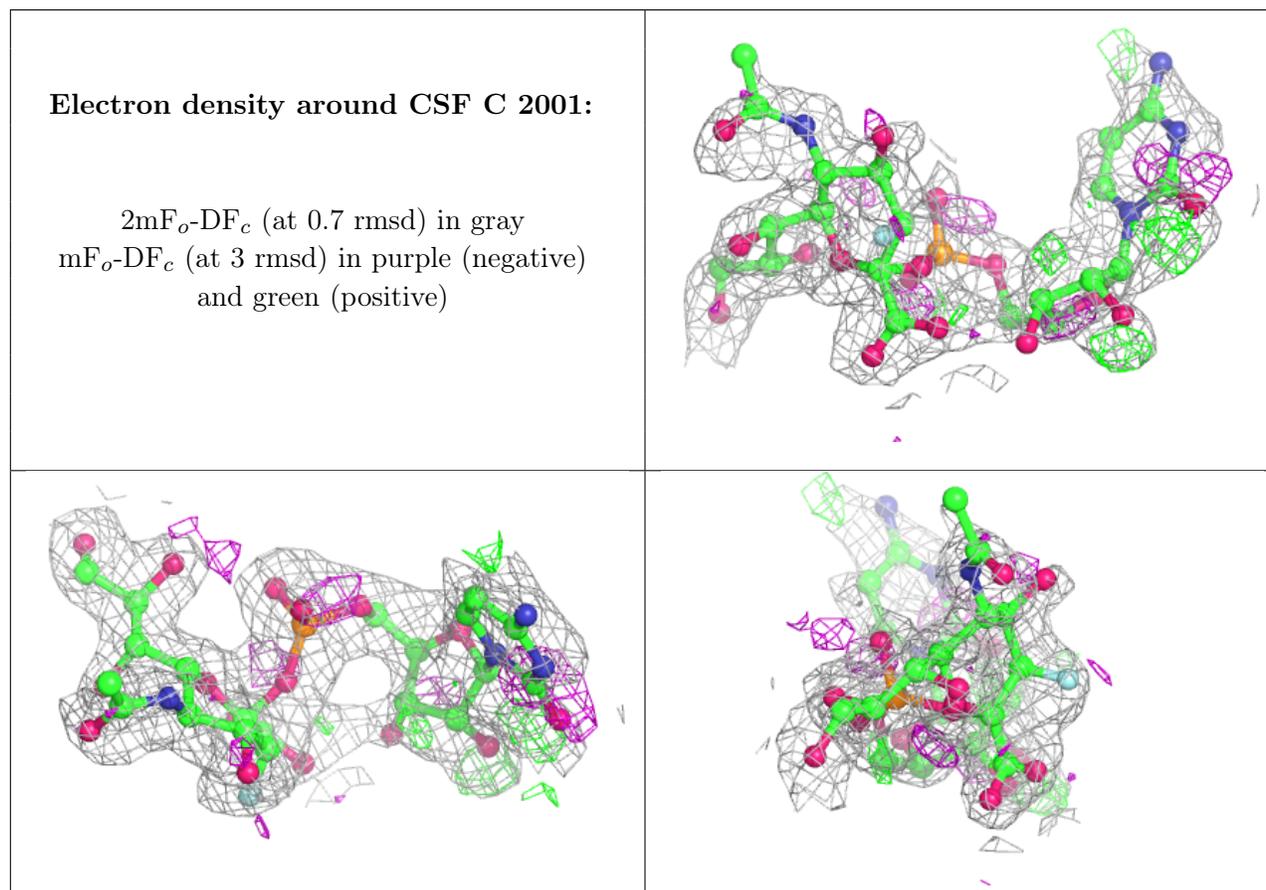
Electron density around CSF C 3001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
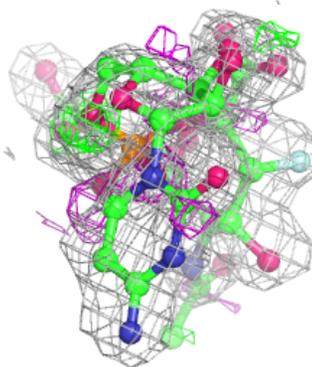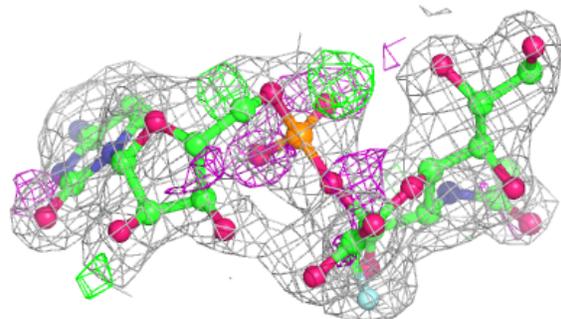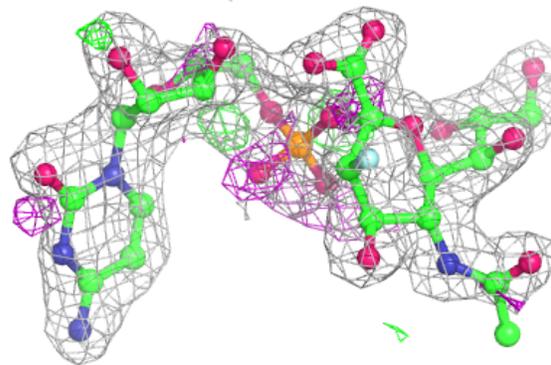
Electron density around CSF C 4001:

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

Electron density around CSF C 1001:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
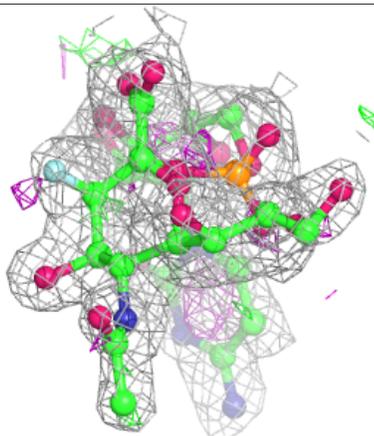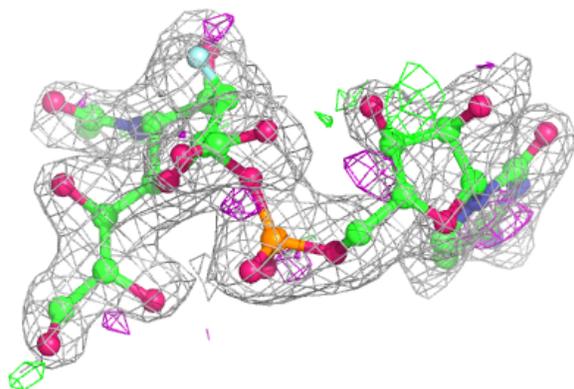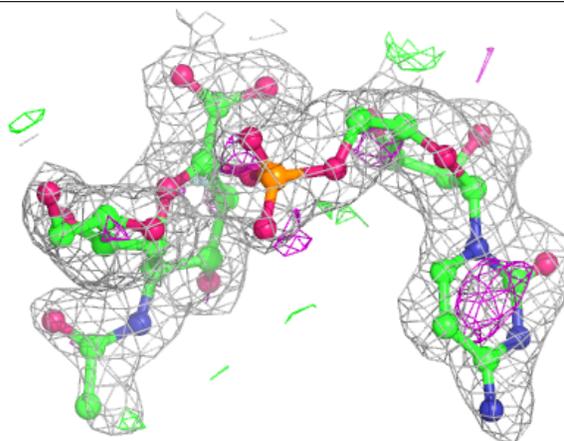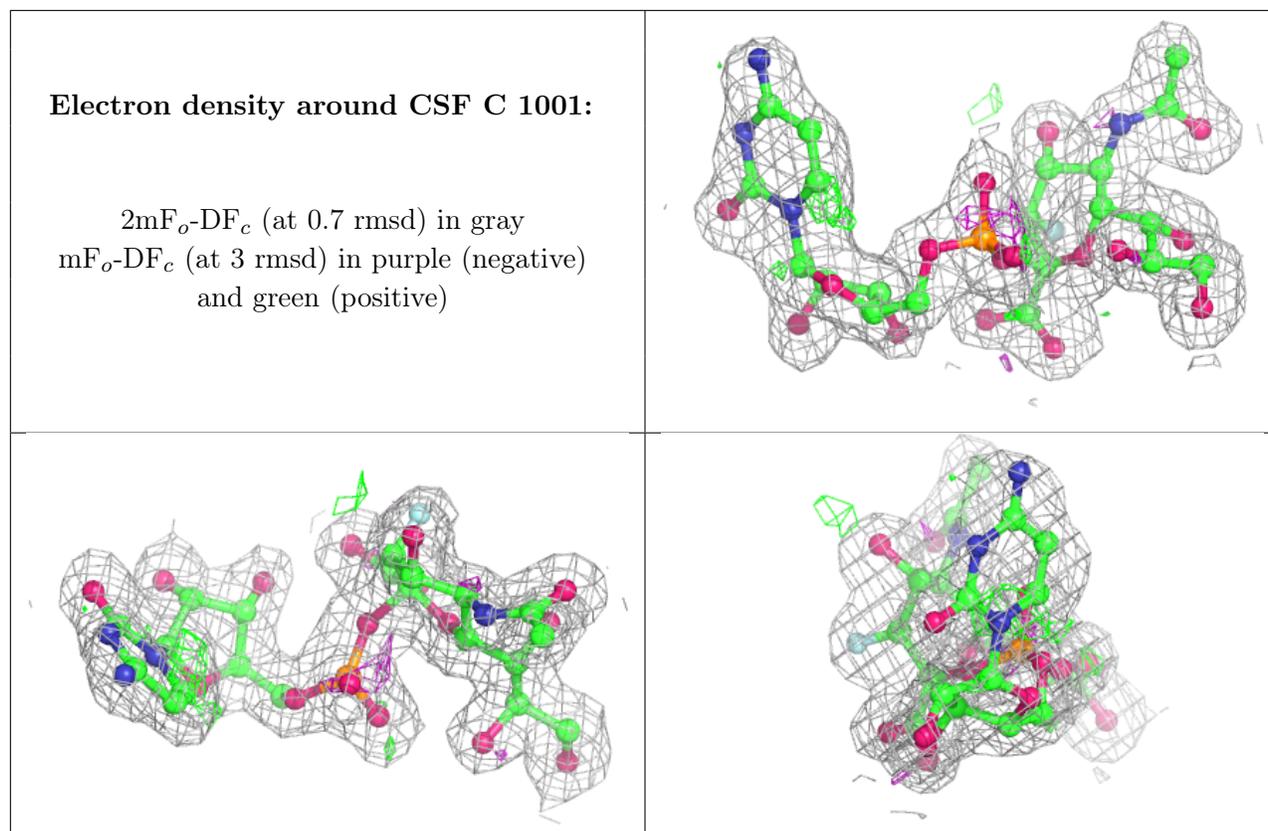
6.5 Other polymers (i)

There are no such residues in this entry.