



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2020 – 07:40 PM BST

PDB ID : 1GK8
Title : Rubisco from Chlamydomonas reinhardtii
Authors : Taylor, T.C.
Deposited on : 2001-08-09
Resolution : 1.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 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	:	2.13
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)	:	2.13

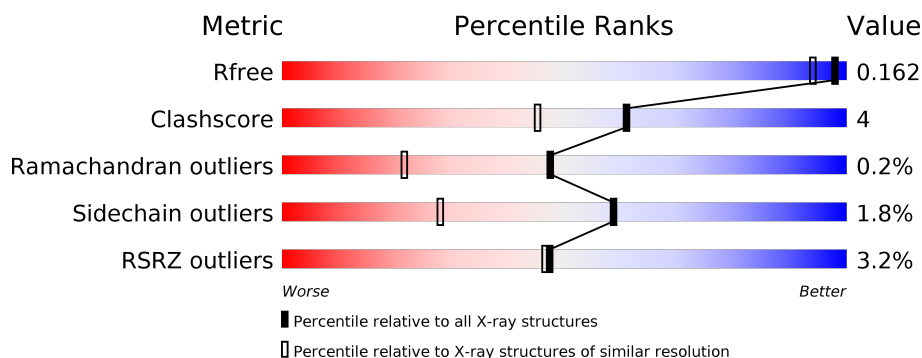
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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

Mol	Chain	Length	Quality of chain
1	A	475	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>9%</div> <div>.</div> </div> </div>
1	C	475	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div>.</div> </div> </div>
1	E	475	<div> <div>%</div> <div> <div></div> <div>87%</div> <div>10%</div> <div>..</div> </div> </div>
1	G	475	<div> <div>%</div> <div> <div></div> <div>87%</div> <div>11%</div> <div>.</div> </div> </div>
2	I	140	<div> <div>6%</div> <div> <div></div> <div>77%</div> <div>12%</div> <div>.</div> <div>10%</div> </div> </div>
2	K	140	<div> <div>5%</div> <div> <div></div> <div>74%</div> <div>15%</div> <div>..</div> <div>10%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	140	<div> <div>9%</div> <div>69%</div> <div>18%</div> <div>•</div> <div>10%</div> </div>
2	O	140	<div> <div>6%</div> <div>69%</div> <div>19%</div> <div>••</div> <div>10%</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 21642 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 RIBULOSE-1,5 BISPHOSPHATE CARBOXYLASE LARGE CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	469	Total	C	N	O	S	0	6	0
			3673	2320	646	680	27			
1	C	469	Total	C	N	O	S	0	8	0
			3679	2324	646	682	27			
1	E	465	Total	C	N	O	S	0	7	0
			3649	2306	641	675	27			
1	G	467	Total	C	N	O	S	0	8	0
			3660	2310	643	679	28			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	PRO	LEU	conflict	UNP P00877
C	46	PRO	LEU	conflict	UNP P00877
E	46	PRO	LEU	conflict	UNP P00877
G	46	PRO	LEU	conflict	UNP P00877

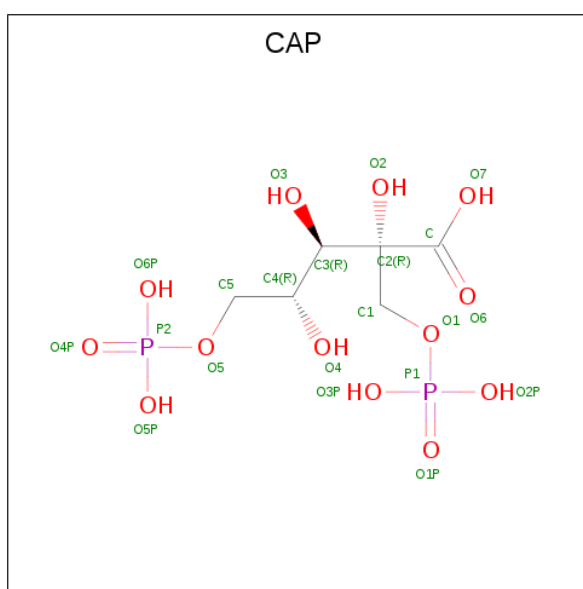
- Molecule 2 is a protein called RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	126	Total	C	N	O	S	0	5	0
			1044	679	167	186	12			
2	K	126	Total	C	N	O	S	0	7	0
			1053	686	167	187	13			
2	M	126	Total	C	N	O	S	0	8	0
			1055	687	167	187	14			
2	O	126	Total	C	N	O	S	0	7	0
			1053	686	167	187	13			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0
3	E	1	Total Mg 1 1	0	0

- Molecule 4 is 2-CARBOXYARABINITOL-1,5-DIPHOSPHATE (three-letter code: CAP) (formula: $C_6H_{14}O_{13}P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O P 21 6 13 2	0	0
4	C	1	Total C O P 21 6 13 2	0	0
4	E	1	Total C O P 21 6 13 2	0	0
4	G	1	Total C O P 21 6 13 2	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total 4	C 2	O 2	0	0
5	E	1	Total 4	C 2	O 2	0	0
5	E	1	Total 4	C 2	O 2	0	0
5	E	1	Total 4	C 2	O 2	0	0
5	E	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	G	1	Total 4	C 2	O 2	0	0
5	I	1	Total 4	C 2	O 2	0	0
5	I	1	Total 4	C 2	O 2	0	0
5	K	1	Total 4	C 2	O 2	0	0
5	K	1	Total 4	C 2	O 2	0	0
5	M	1	Total 4	C 2	O 2	0	0
5	M	1	Total 4	C 2	O 2	0	0
5	O	1	Total 4	C 2	O 2	0	0
5	O	1	Total 4	C 2	O 2	0	0

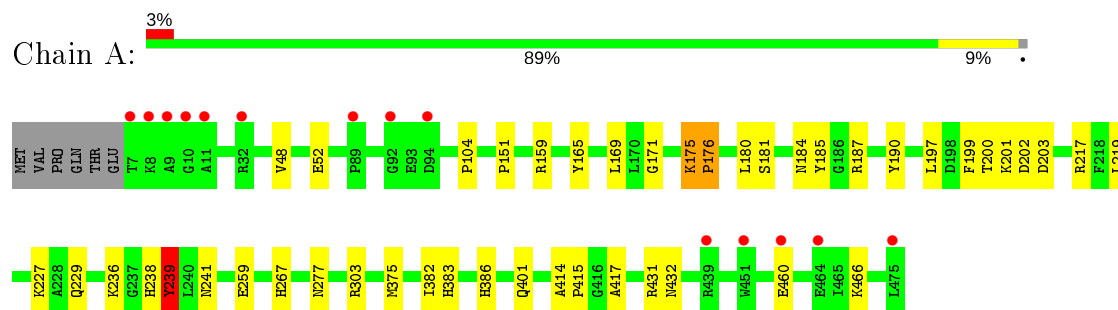
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	461	Total 461	O 461	0	4
6	C	466	Total 466	O 466	0	0
6	E	487	Total 487	O 487	0	4
6	G	474	Total 474	O 474	0	0
6	I	157	Total 157	O 157	0	0
6	K	168	Total 168	O 168	0	0
6	M	170	Total 170	O 170	0	0
6	O	173	Total 173	O 173	0	0

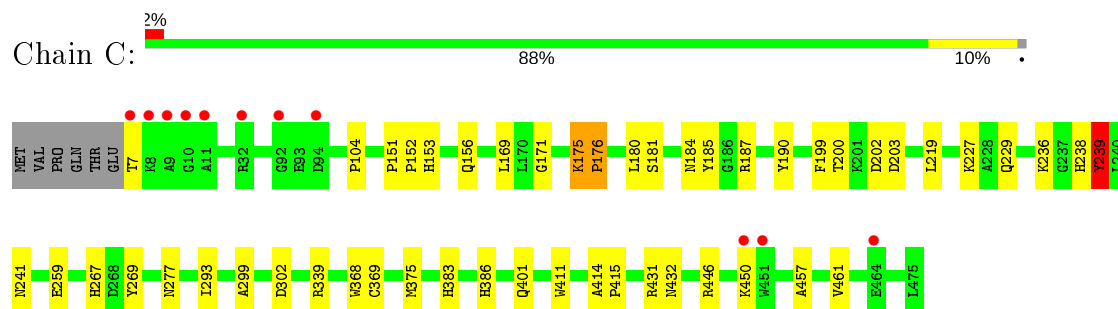
3 Residue-property plots [i](#)

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.

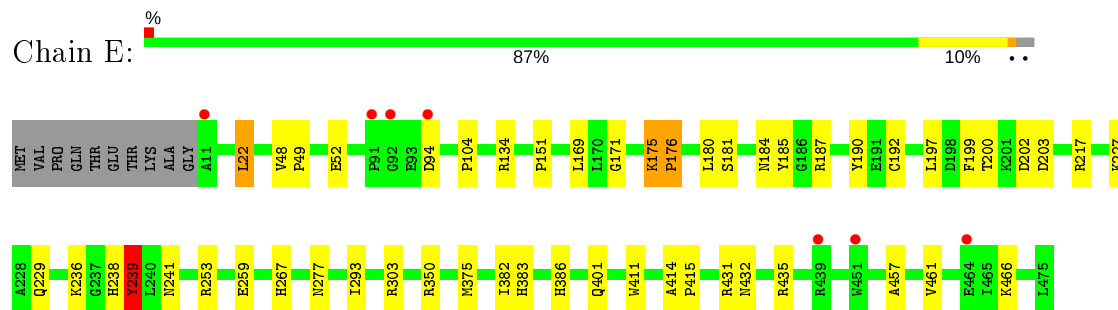
• Molecule 1: RIBULOSE-1,5 BISPHOSPHATE CARBOXYLASE LARGE CHAIN



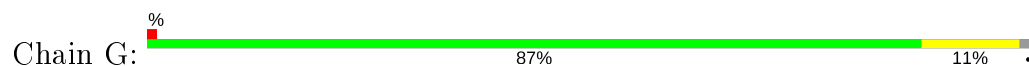
• Molecule 1: RIBULOSE-1,5 BISPHOSPHATE CARBOXYLASE LARGE CHAIN

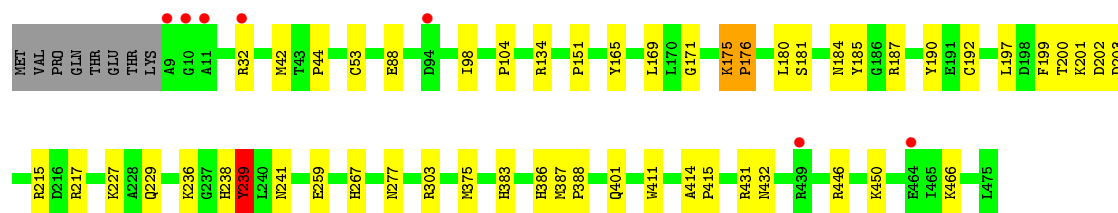


• Molecule 1: RIBULOSE-1,5 BISPHOSPHATE CARBOXYLASE LARGE CHAIN



• Molecule 1: RIBULOSE-1,5 BISPHOSPHATE CARBOXYLASE LARGE CHAIN

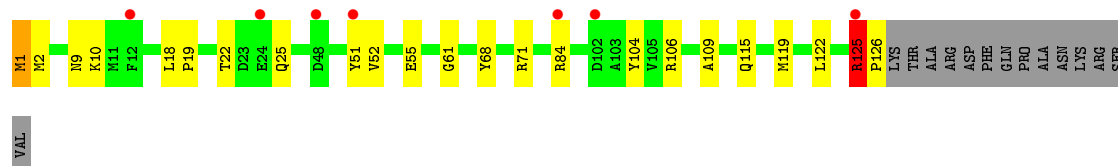




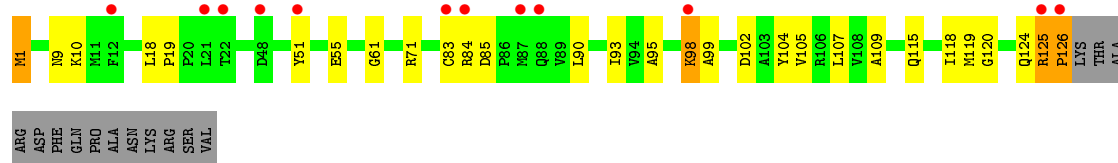
• Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1



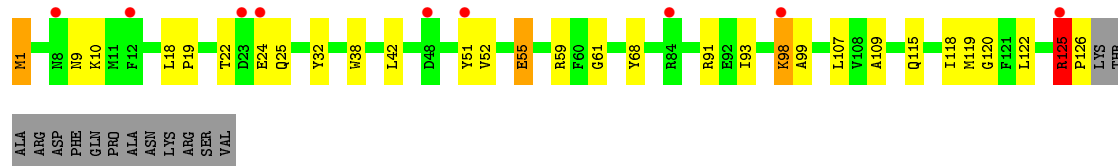
• Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1



• Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1



• Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	171.38Å 142.64Å 124.72Å 90.00° 124.13° 90.00°	Depositor
Resolution (Å)	20.00 – 1.40 84.71 – 1.40	Depositor EDS
% Data completeness (in resolution range)	97.0 (20.00-1.40) 96.1 (84.71-1.40)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 1.39Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.149 , 0.162 0.152 , 0.162	Depositor DCC
R_{free} test set	23556 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	10.3	Xtriage
Anisotropy	0.192	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	21642	wwPDB-VP
Average B, all atoms (Å ²)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MG, CAP, EDO, HYP, MME, SMC, KCX

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	0.37	0/3748	0.86	9/5064 (0.2%)
1	C	0.36	0/3763	0.87	8/5085 (0.2%)
1	E	0.38	0/3730	0.93	14/5041 (0.3%)
1	G	0.36	0/3747	0.89	13/5063 (0.3%)
2	I	0.39	1/1089 (0.1%)	0.85	2/1483 (0.1%)
2	K	0.40	1/1110 (0.1%)	0.86	5/1511 (0.3%)
2	M	0.41	1/1118 (0.1%)	0.87	5/1521 (0.3%)
2	O	0.41	1/1110 (0.1%)	1.09	9/1511 (0.6%)
All	All	0.37	4/19415 (0.0%)	0.89	65/26279 (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	A	0	2
1	C	0	2
1	E	0	2
1	G	0	2
2	I	0	1
2	M	0	1
All	All	0	10

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	M	126	PRO	N-CD	5.71	1.55	1.47
2	I	126	PRO	N-CD	5.54	1.55	1.47
2	K	126	PRO	N-CD	5.35	1.55	1.47
2	O	126	PRO	N-CD	5.18	1.55	1.47

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	O	125	ARG	NE-CZ-NH1	13.46	127.03	120.30
2	O	55	GLU	CA-CB-CG	10.19	135.82	113.40
2	O	125	ARG	CD-NE-CZ	9.50	136.90	123.60
2	O	91	ARG	NE-CZ-NH1	8.57	124.58	120.30
1	G	217	ARG	NE-CZ-NH2	-8.55	116.03	120.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	LYS	Mainchain,Peptide
1	C	175	LYS	Mainchain,Peptide
1	E	175	LYS	Mainchain,Peptide
1	G	175	LYS	Mainchain,Peptide
2	I	125	ARG	Mainchain

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	3673	0	3583	27	0
1	C	3679	0	3592	26	0
1	E	3649	0	3562	28	1
1	G	3660	0	3562	27	0
2	I	1044	0	1015	18	0
2	K	1053	0	1021	20	0
2	M	1055	0	1022	19	0
2	O	1053	0	1021	24	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
4	A	21	0	7	0	0
4	C	21	0	8	0	0
4	E	21	0	7	0	0
4	G	21	0	7	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	28	0	42	2	0
5	C	20	0	30	0	0
5	E	28	0	42	3	0
5	G	24	0	36	0	0
5	I	8	0	12	1	0
5	K	8	0	12	0	0
5	M	8	0	12	0	0
5	O	8	0	12	0	0
6	A	461	0	0	2	0
6	C	466	0	0	2	0
6	E	487	0	0	2	0
6	G	474	0	0	1	1
6	I	157	0	0	2	0
6	K	168	0	0	0	0
6	M	170	0	0	0	0
6	O	173	0	0	3	0
All	All	21642	0	18605	169	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

The worst 5 of 169 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)
2:O:125:ARG:HH11	2:O:125:ARG:HB3	1.22	1.03
1:C:267:HIS:HD2	1:C:277:ASN:HD22	1.09	1.00
1:G:267:HIS:HD2	1:G:277:ASN:HD22	1.09	0.93
1:A:267:HIS:HD2	1:A:277:ASN:HD22	1.07	0.92
1:E:267:HIS:HD2	1:E:277:ASN:HD22	1.09	0.89

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:22:LEU:CD1	6:G:2087:HOH:O[4_455]	1.90	0.30

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	469/475 (99%)	454 (97%)	14 (3%)	1 (0%)	47	21
1	C	471/475 (99%)	456 (97%)	14 (3%)	1 (0%)	47	21
1	E	466/475 (98%)	451 (97%)	14 (3%)	1 (0%)	47	21
1	G	469/475 (99%)	455 (97%)	13 (3%)	1 (0%)	47	21
2	I	129/140 (92%)	124 (96%)	5 (4%)	0	100	100
2	K	131/140 (94%)	127 (97%)	4 (3%)	0	100	100
2	M	132/140 (94%)	126 (96%)	6 (4%)	0	100	100
2	O	131/140 (94%)	127 (97%)	4 (3%)	0	100	100
All	All	2398/2460 (98%)	2320 (97%)	74 (3%)	4 (0%)	47	21

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	176	PRO
1	E	176	PRO
1	G	176	PRO
1	A	176	PRO

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	377/376 (100%)	371 (98%)	6 (2%)	62	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	379/376 (101%)	373 (98%)	6 (2%)	62	33
1	E	376/376 (100%)	370 (98%)	6 (2%)	62	33
1	G	377/376 (100%)	372 (99%)	5 (1%)	69	42
2	I	115/122 (94%)	113 (98%)	2 (2%)	60	31
2	K	117/122 (96%)	114 (97%)	3 (3%)	46	13
2	M	118/122 (97%)	115 (98%)	3 (2%)	47	14
2	O	117/122 (96%)	113 (97%)	4 (3%)	37	8
All	All	1976/1992 (99%)	1941 (98%)	35 (2%)	59	28

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	239	TYR
1	G	203	ASP
2	O	55	GLU
1	E	241	ASN
1	E	466	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 61 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	267	HIS
1	G	163	ASN
2	O	9	ASN
1	E	277	ASN
1	E	386	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

24 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
1	KCX	C	201	1,3	7,11,12	0.51	0	4,12,14	0.12	0
2	MME	O	1	2	7,8,9	1.85	1 (14%)	5,8,10	1.03	0
1	HYP	A	151	1	6,8,9	1.81	1 (16%)	5,10,12	1.06	0
1	SMC	E	369	1	5,6,7	0.49	0	2,6,8	0.52	0
1	KCX	A	201	1,3	7,11,12	0.53	0	4,12,14	0.44	0
1	SMC	G	369	1	5,6,7	0.55	0	2,6,8	0.33	0
1	SMC	C	256	1	5,6,7	0.53	0	2,6,8	0.10	0
1	HYP	E	151	1	6,8,9	1.73	1 (16%)	5,10,12	1.42	1 (20%)
1	HYP	G	151	1	6,8,9	1.75	1 (16%)	5,10,12	1.08	0
1	SMC	A	369	1	5,6,7	0.50	0	2,6,8	0.36	0
1	SMC	E	256	1	5,6,7	0.64	0	2,6,8	0.31	0
1	SMC	C	369	1	5,6,7	0.55	0	2,6,8	0.20	0
1	HYP	A	104	1	6,8,9	1.97	1 (16%)	5,10,12	1.13	0
2	MME	I	1	2	7,8,9	1.86	1 (14%)	5,8,10	1.62	1 (20%)
2	MME	K	1	2	7,8,9	1.82	1 (14%)	5,8,10	1.18	0
1	SMC	G	256	1	5,6,7	0.55	0	2,6,8	0.34	0
1	KCX	E	201	1,3	7,11,12	0.55	0	4,12,14	0.77	0
1	SMC	A	256	1	5,6,7	0.58	0	2,6,8	0.49	0
1	KCX	G	201	1,3	7,11,12	0.51	0	4,12,14	0.88	0
1	HYP	C	104	1	6,8,9	1.89	1 (16%)	5,10,12	1.18	1 (20%)
1	HYP	G	104	1	6,8,9	1.95	1 (16%)	5,10,12	1.23	1 (20%)
1	HYP	E	104	1	6,8,9	1.77	1 (16%)	5,10,12	1.46	1 (20%)
2	MME	M	1	2	7,8,9	1.76	1 (14%)	5,8,10	1.18	0
1	HYP	C	151	1	6,8,9	1.76	1 (16%)	5,10,12	1.27	1 (20%)

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
1	KCX	C	201	1,3	-	0/7/10/12	-
2	MME	O	1	2	-	0/5/8/10	-
1	HYP	A	151	1	-	0/0/11/13	0/1/1/1
1	SMC	E	369	1	-	1/3/5/7	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	KCX	A	201	1,3	-	0/7/10/12	-
1	SMC	G	369	1	-	1/3/5/7	-
1	SMC	C	256	1	-	0/3/5/7	-
1	HYP	E	151	1	-	0/0/11/13	0/1/1/1
1	HYP	G	151	1	-	0/0/11/13	0/1/1/1
1	SMC	A	369	1	-	1/3/5/7	-
1	SMC	E	256	1	-	0/3/5/7	-
1	SMC	C	369	1	-	1/3/5/7	-
1	HYP	A	104	1	-	0/0/11/13	0/1/1/1
2	MME	I	1	2	-	0/5/8/10	-
2	MME	K	1	2	-	2/5/8/10	-
1	SMC	G	256	1	-	0/3/5/7	-
1	KCX	E	201	1,3	-	0/7/10/12	-
1	SMC	A	256	1	-	0/3/5/7	-
1	KCX	G	201	1,3	-	0/7/10/12	-
1	HYP	C	104	1	-	0/0/11/13	0/1/1/1
1	HYP	G	104	1	-	0/0/11/13	0/1/1/1
1	HYP	E	104	1	-	0/0/11/13	0/1/1/1
2	MME	M	1	2	-	1/5/8/10	-
1	HYP	C	151	1	-	0/0/11/13	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	1	MME	CM-N	-4.55	1.34	1.46
2	I	1	MME	CM-N	-4.53	1.34	1.46
2	K	1	MME	CM-N	-4.44	1.35	1.46
2	M	1	MME	CM-N	-4.33	1.35	1.46
1	A	104	HYP	CD-N	3.88	1.61	1.47

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	1	MME	O-C-CA	-2.61	117.94	124.78
1	G	104	HYP	O-C-CA	-2.46	118.34	124.78
1	C	104	HYP	O-C-CA	-2.40	118.48	124.78
1	E	151	HYP	CB-CG-CD	-2.37	100.36	103.27
1	E	104	HYP	O-C-CA	-2.17	119.10	124.78

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	K	1	MME	C-CA-CB-CG
2	K	1	MME	N-CA-CB-CG
1	E	369	SMC	N-CA-CB-SG
1	G	369	SMC	N-CA-CB-SG
1	A	369	SMC	N-CA-CB-SG

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	O	1	MME	1	0
1	A	201	KCX	1	0
1	C	369	SMC	1	0
2	K	1	MME	3	0
1	G	201	KCX	1	0
2	M	1	MME	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 41 ligands modelled in this entry, 4 are monoatomic - leaving 37 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	$\# Z > 2$
5	EDO	C	1479	-	3,3,3	0.55	0	2,2,2	0.49	0
5	EDO	G	1479	-	3,3,3	0.51	0	2,2,2	0.71	0
5	EDO	E	1481	-	3,3,3	0.52	0	2,2,2	0.44	0
5	EDO	C	1480	-	3,3,3	0.63	0	2,2,2	0.65	0
5	EDO	G	1478	-	3,3,3	0.57	0	2,2,2	0.20	0
4	CAP	C	1477	3	15,20,20	1.36	2 (13%)	20,31,31	0.95	2 (10%)
5	EDO	E	1479	-	3,3,3	0.54	0	2,2,2	0.64	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	EDO	E	1483	-	3,3,3	0.56	0	2,2,2	0.57	0
5	EDO	C	1482	-	3,3,3	0.56	0	2,2,2	0.45	0
5	EDO	G	1483	-	3,3,3	0.53	0	2,2,2	0.55	0
5	EDO	G	1482	-	3,3,3	0.57	0	2,2,2	0.47	0
5	EDO	M	1128	-	3,3,3	0.58	0	2,2,2	0.68	0
5	EDO	E	1482	-	3,3,3	0.53	0	2,2,2	0.41	0
5	EDO	C	1481	-	3,3,3	0.49	0	2,2,2	0.36	0
5	EDO	M	1127	-	3,3,3	0.54	0	2,2,2	0.49	0
5	EDO	C	1478	-	3,3,3	0.60	0	2,2,2	0.33	0
5	EDO	A	1484	-	3,3,3	0.59	0	2,2,2	0.67	0
5	EDO	A	1478	-	3,3,3	0.56	0	2,2,2	0.24	0
4	CAP	E	1477	3	15,20,20	1.21	1 (6%)	20,31,31	1.07	1 (5%)
5	EDO	G	1480	-	3,3,3	0.66	0	2,2,2	0.64	0
5	EDO	I	1128	-	3,3,3	0.58	0	2,2,2	0.81	0
5	EDO	K	1128	-	3,3,3	0.59	0	2,2,2	0.62	0
5	EDO	A	1482	-	3,3,3	0.55	0	2,2,2	0.50	0
5	EDO	I	1127	-	3,3,3	0.53	0	2,2,2	0.43	0
5	EDO	A	1480	-	3,3,3	0.64	0	2,2,2	0.66	0
5	EDO	E	1480	-	3,3,3	0.62	0	2,2,2	0.53	0
5	EDO	O	1127	-	3,3,3	0.55	0	2,2,2	0.53	0
5	EDO	E	1478	-	3,3,3	0.53	0	2,2,2	0.15	0
5	EDO	A	1483	-	3,3,3	0.58	0	2,2,2	0.53	0
5	EDO	O	1128	-	3,3,3	0.58	0	2,2,2	0.58	0
5	EDO	K	1127	-	3,3,3	0.57	0	2,2,2	0.52	0
5	EDO	A	1481	-	3,3,3	0.53	0	2,2,2	0.46	0
4	CAP	G	1477	3	15,20,20	1.42	2 (13%)	20,31,31	1.07	1 (5%)
5	EDO	G	1481	-	3,3,3	0.56	0	2,2,2	0.47	0
5	EDO	E	1484	-	3,3,3	0.58	0	2,2,2	0.63	0
4	CAP	A	1477	3	15,20,20	1.45	2 (13%)	20,31,31	1.07	2 (10%)
5	EDO	A	1479	-	3,3,3	0.57	0	2,2,2	0.64	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
5	EDO	C	1479	-	-	1/1/1/1	-
5	EDO	G	1479	-	-	1/1/1/1	-
5	EDO	E	1481	-	-	1/1/1/1	-
5	EDO	C	1480	-	-	0/1/1/1	-
5	EDO	G	1478	-	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CAP	C	1477	3	-	5/23/29/29	-
5	EDO	E	1479	-	-	1/1/1/1	-
5	EDO	E	1483	-	-	0/1/1/1	-
5	EDO	C	1482	-	-	0/1/1/1	-
5	EDO	G	1483	-	-	0/1/1/1	-
5	EDO	G	1482	-	-	0/1/1/1	-
5	EDO	M	1128	-	-	0/1/1/1	-
5	EDO	E	1482	-	-	0/1/1/1	-
5	EDO	C	1481	-	-	0/1/1/1	-
5	EDO	M	1127	-	-	0/1/1/1	-
5	EDO	C	1478	-	-	0/1/1/1	-
5	EDO	A	1484	-	-	0/1/1/1	-
5	EDO	A	1478	-	-	0/1/1/1	-
4	CAP	E	1477	3	-	5/23/29/29	-
5	EDO	G	1480	-	-	0/1/1/1	-
5	EDO	I	1128	-	-	1/1/1/1	-
5	EDO	K	1128	-	-	0/1/1/1	-
5	EDO	A	1482	-	-	0/1/1/1	-
5	EDO	I	1127	-	-	0/1/1/1	-
5	EDO	A	1480	-	-	0/1/1/1	-
5	EDO	E	1480	-	-	0/1/1/1	-
5	EDO	O	1127	-	-	0/1/1/1	-
5	EDO	E	1478	-	-	0/1/1/1	-
5	EDO	A	1483	-	-	0/1/1/1	-
5	EDO	O	1128	-	-	0/1/1/1	-
5	EDO	K	1127	-	-	0/1/1/1	-
5	EDO	A	1481	-	-	1/1/1/1	-
4	CAP	G	1477	3	-	5/23/29/29	-
5	EDO	G	1481	-	-	0/1/1/1	-
5	EDO	E	1484	-	-	0/1/1/1	-
4	CAP	A	1477	3	-	4/23/29/29	-
5	EDO	A	1479	-	-	1/1/1/1	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1477	CAP	O2-C2	-3.26	1.38	1.43
4	C	1477	CAP	O2-C2	-3.24	1.38	1.43
4	G	1477	CAP	O2-C2	-2.95	1.38	1.43
4	G	1477	CAP	O4-C4	-2.84	1.37	1.43
4	E	1477	CAP	O2-C2	-2.73	1.38	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	1477	CAP	O4-C4-C3	-2.83	103.12	108.78
4	G	1477	CAP	O4-C4-C3	-2.48	103.82	108.78
4	A	1477	CAP	C5-C4-C3	2.39	116.82	111.94
4	C	1477	CAP	O4-C4-C3	-2.37	104.05	108.78
4	A	1477	CAP	O4-C4-C3	-2.26	104.27	108.78

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

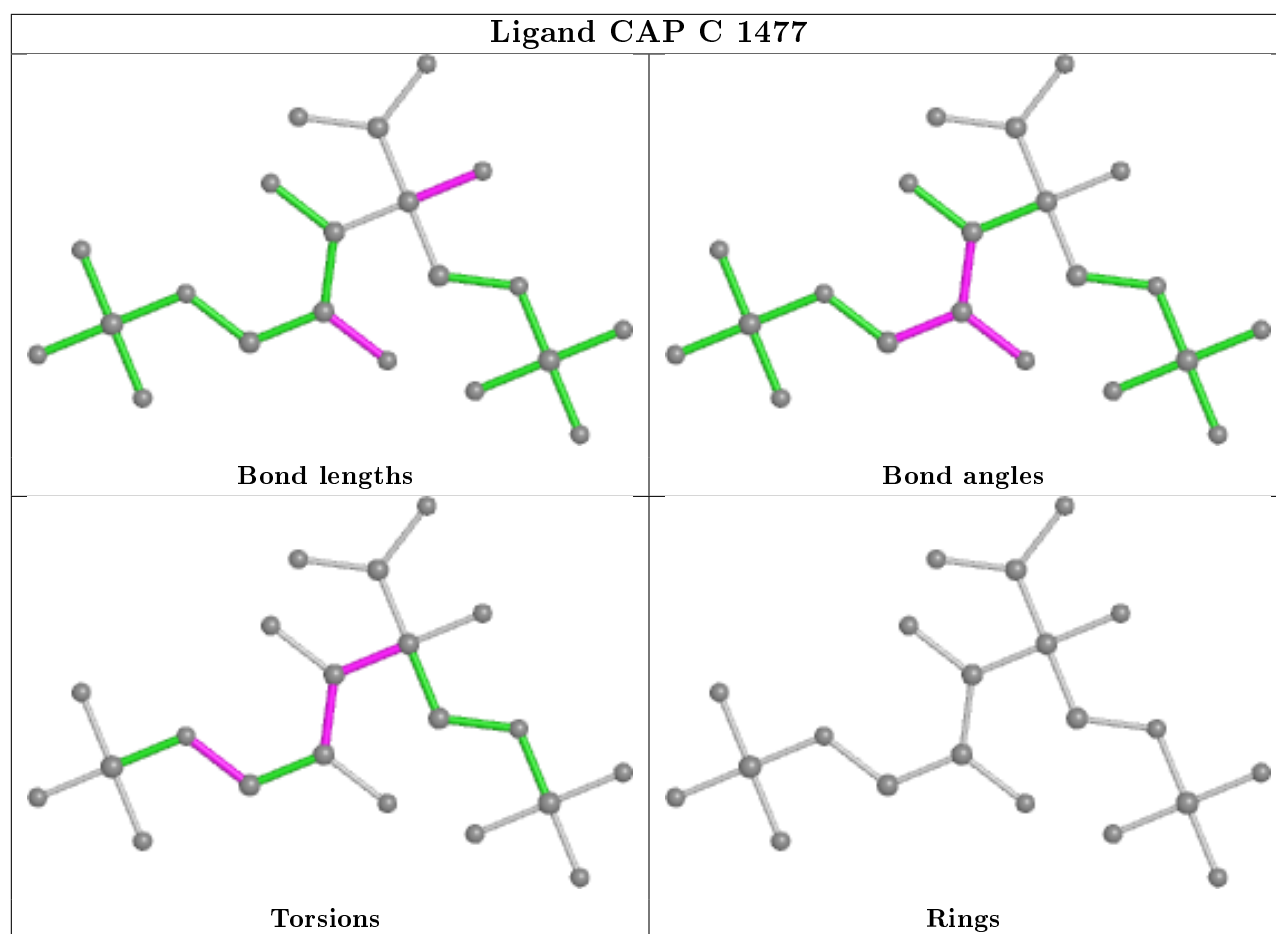
Mol	Chain	Res	Type	Atoms
4	C	1477	CAP	C2-C3-C4-O4
4	C	1477	CAP	O3-C3-C4-O4
4	E	1477	CAP	C2-C3-C4-O4
4	E	1477	CAP	O3-C3-C4-O4
4	G	1477	CAP	C2-C3-C4-O4

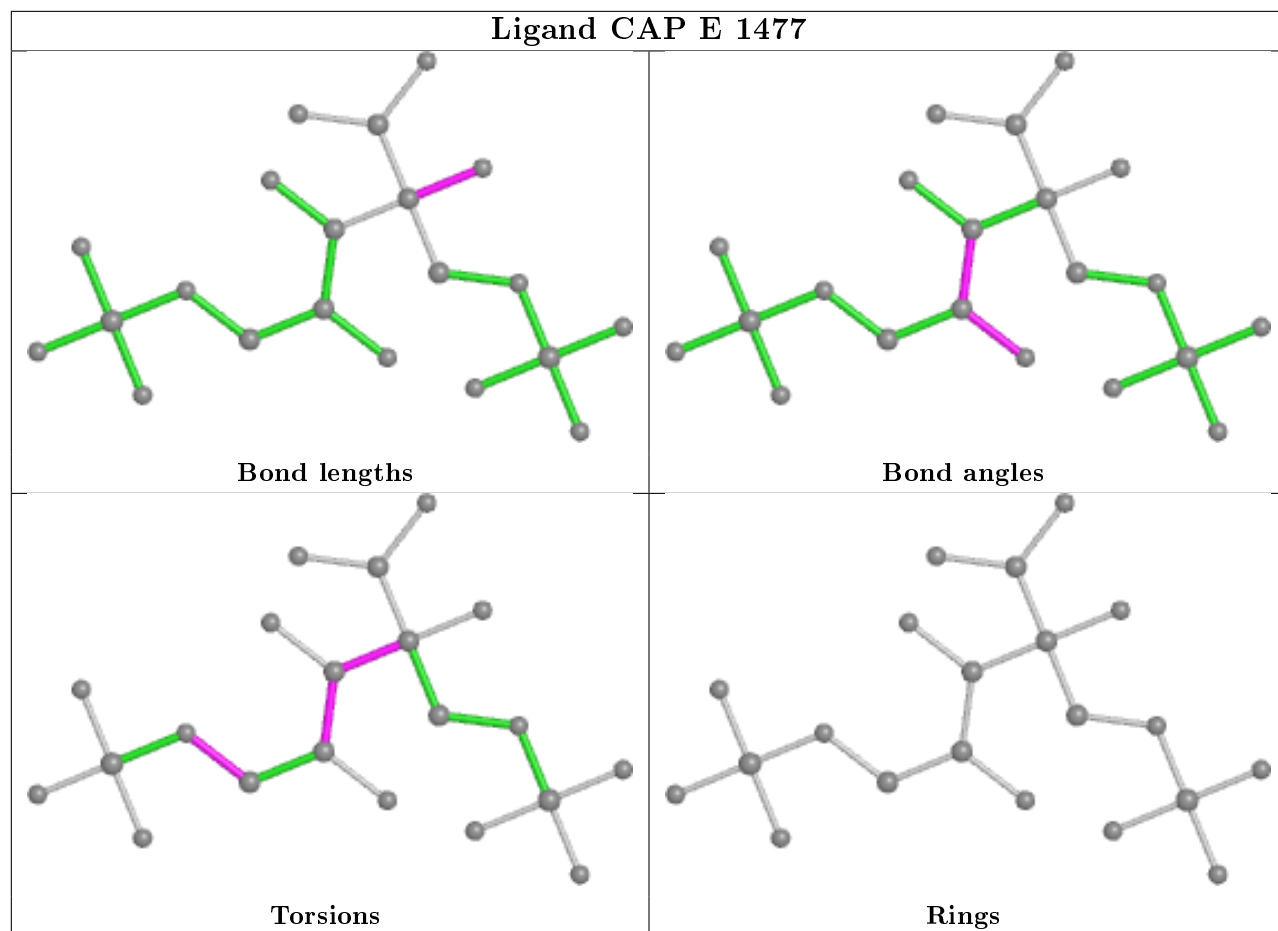
There are no ring outliers.

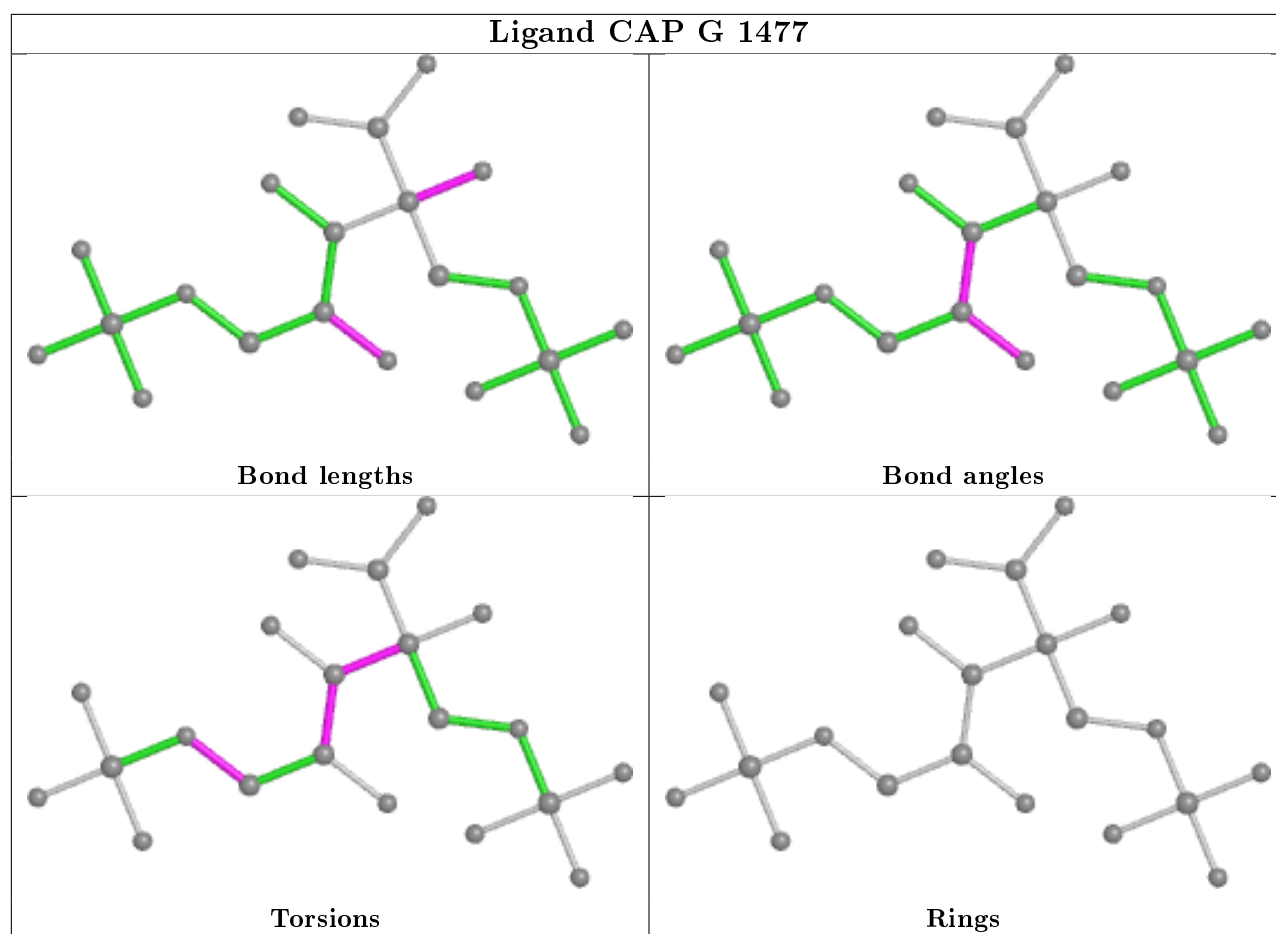
3 monomers are involved in 6 short contacts:

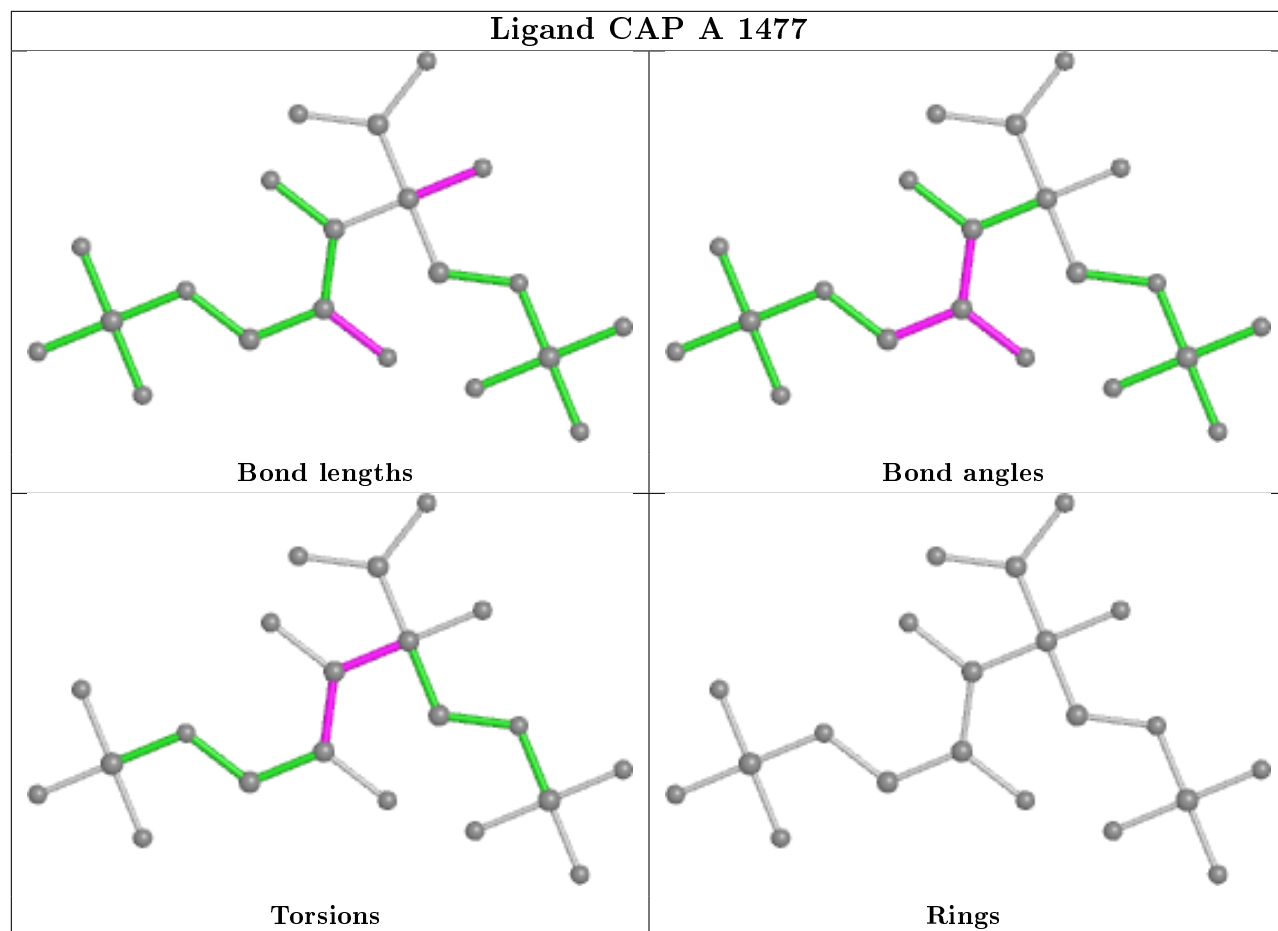
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	1481	EDO	3	0
5	I	1128	EDO	1	0
5	A	1481	EDO	2	0

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, 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	464/475 (97%)	-0.09	14 (3%) 50 49	6, 10, 20, 40	0
1	C	464/475 (97%)	-0.06	11 (2%) 59 58	6, 9, 19, 42	1 (0%)
1	E	460/475 (96%)	-0.14	7 (1%) 73 72	6, 8, 18, 30	0
1	G	462/475 (97%)	-0.06	7 (1%) 73 72	6, 9, 21, 42	0
2	I	125/140 (89%)	0.25	9 (7%) 15 14	9, 14, 23, 32	0
2	K	125/140 (89%)	0.20	7 (5%) 24 22	8, 14, 22, 32	0
2	M	125/140 (89%)	0.31	12 (9%) 8 7	9, 14, 22, 32	0
2	O	125/140 (89%)	0.35	9 (7%) 15 14	9, 14, 21, 32	0
All	All	2350/2460 (95%)	-0.01	76 (3%) 47 46	6, 10, 21, 42	1 (0%)

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	9	ALA	14.8
1	A	9	ALA	6.8
1	E	11	ALA	6.4
1	C	8	LYS	6.2
1	A	8	LYS	5.7

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	MME	O	1	9/10	0.71	0.18	18,21,29,29	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	MME	K	1	9/10	0.91	0.14	17,20,22,22	0
1	HYP	E	104	8/9	0.92	0.13	7,7,9,10	0
2	MME	M	1	9/10	0.93	0.13	16,18,20,22	0
2	MME	I	1	9/10	0.95	0.13	16,17,18,18	0
1	HYP	C	104	8/9	0.97	0.07	8,8,9,10	0
1	KCX	C	201	12/13	0.97	0.08	6,6,6,6	0
1	HYP	G	151	8/9	0.97	0.09	6,8,8,8	0
1	HYP	C	151	8/9	0.97	0.10	6,7,8,8	0
1	KCX	E	201	12/13	0.98	0.06	5,6,6,7	0
1	KCX	G	201	12/13	0.98	0.07	6,6,7,8	0
1	HYP	A	151	8/9	0.98	0.09	7,8,8,8	0
1	HYP	G	104	8/9	0.98	0.07	7,8,8,9	0
1	HYP	A	104	8/9	0.98	0.07	8,8,9,9	0
1	KCX	A	201	12/13	0.98	0.06	6,7,7,7	0
1	HYP	E	151	8/9	0.98	0.10	7,7,7,8	0
1	SMC	A	369	7/8	0.99	0.06	8,9,9,11	0
1	SMC	A	256	7/8	0.99	0.08	7,7,8,8	0
1	SMC	E	256	7/8	0.99	0.08	6,7,7,8	0
1	SMC	C	369	7/8	0.99	0.06	7,8,9,10	0
1	SMC	C	256	7/8	0.99	0.10	6,6,8,8	0
1	SMC	E	369	7/8	0.99	0.07	7,8,8,10	0
1	SMC	G	369	7/8	0.99	0.06	9,9,10,11	0
1	SMC	G	256	7/8	0.99	0.08	6,7,8,8	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, 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(\AA^2)	Q<0.9
5	EDO	E	1481	4/4	0.69	0.28	32,33,33,33	0
5	EDO	C	1480	4/4	0.72	0.14	29,29,29,30	0
5	EDO	E	1480	4/4	0.85	0.13	20,21,22,22	0
5	EDO	A	1481	4/4	0.85	0.35	36,36,36,36	0
5	EDO	O	1127	4/4	0.86	0.15	21,21,21,23	0

Continued on next page...

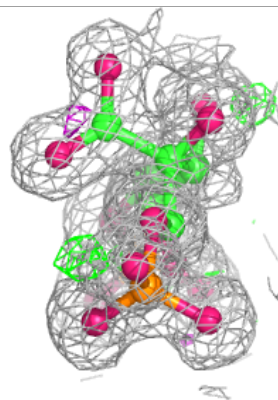
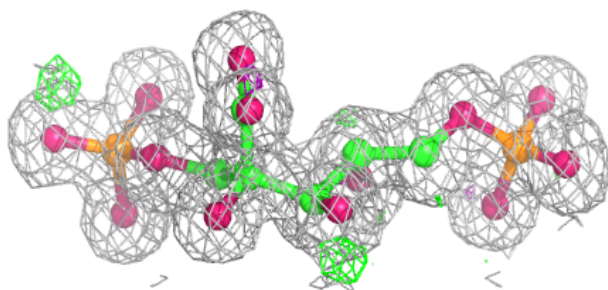
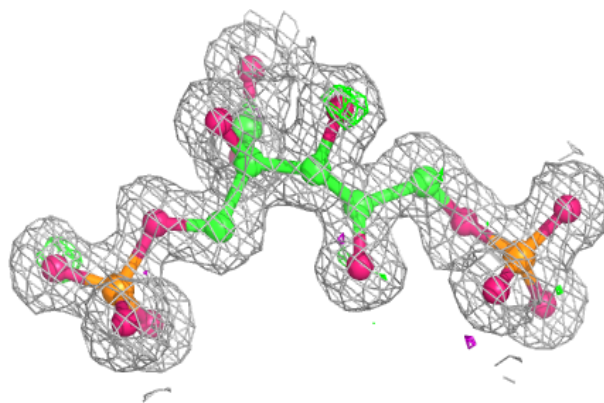
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	EDO	A	1480	4/4	0.87	0.11	23,23,23,24	0
5	EDO	G	1480	4/4	0.88	0.11	20,20,21,21	0
5	EDO	K	1128	4/4	0.88	0.12	23,23,24,26	0
5	EDO	G	1482	4/4	0.88	0.15	20,21,21,22	0
5	EDO	O	1128	4/4	0.89	0.13	19,20,20,22	0
5	EDO	I	1128	4/4	0.89	0.13	25,27,28,29	0
5	EDO	A	1483	4/4	0.91	0.14	20,21,21,22	0
5	EDO	M	1127	4/4	0.92	0.12	20,20,21,22	0
5	EDO	K	1127	4/4	0.92	0.11	19,19,19,21	0
5	EDO	M	1128	4/4	0.92	0.14	19,21,21,23	0
5	EDO	G	1481	4/4	0.92	0.12	18,18,18,18	0
5	EDO	A	1479	4/4	0.92	0.10	15,15,15,17	0
5	EDO	I	1127	4/4	0.93	0.17	20,21,21,21	0
5	EDO	E	1483	4/4	0.93	0.13	19,20,20,21	0
5	EDO	E	1479	4/4	0.94	0.11	13,13,14,15	0
5	EDO	G	1479	4/4	0.94	0.11	13,14,14,16	0
5	EDO	C	1479	4/4	0.94	0.10	15,15,16,17	0
5	EDO	C	1482	4/4	0.94	0.12	18,19,19,20	0
5	EDO	E	1482	4/4	0.96	0.10	12,12,13,13	0
5	EDO	A	1482	4/4	0.96	0.10	15,16,16,16	0
5	EDO	C	1478	4/4	0.96	0.09	10,11,11,12	0
5	EDO	G	1478	4/4	0.97	0.07	9,9,9,10	0
5	EDO	A	1478	4/4	0.97	0.09	10,11,11,11	0
5	EDO	C	1481	4/4	0.97	0.06	14,15,15,15	0
5	EDO	G	1483	4/4	0.97	0.10	13,13,13,14	0
4	CAP	G	1477	21/21	0.98	0.07	7,8,9,10	0
4	CAP	C	1477	21/21	0.98	0.07	6,7,8,9	0
5	EDO	E	1484	4/4	0.98	0.08	12,12,12,12	4
4	CAP	A	1477	21/21	0.98	0.06	7,8,9,11	0
5	EDO	E	1478	4/4	0.98	0.08	9,9,10,11	0
5	EDO	A	1484	4/4	0.99	0.06	13,13,13,13	4
4	CAP	E	1477	21/21	0.99	0.06	6,7,8,9	0
3	MG	A	1476	1/1	1.00	0.06	7,7,7,7	0
3	MG	E	1476	1/1	1.00	0.07	6,6,6,6	0
3	MG	G	1476	1/1	1.00	0.05	6,6,6,6	0
3	MG	C	1476	1/1	1.00	0.07	6,6,6,6	0

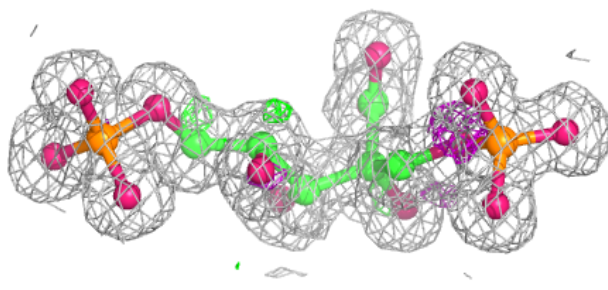
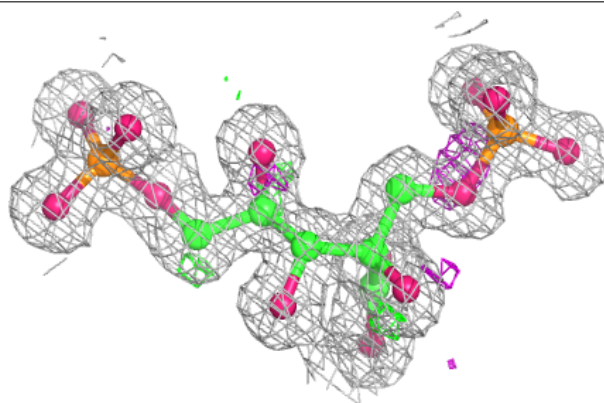
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CAP G 1477:

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

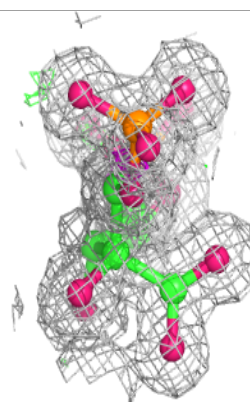
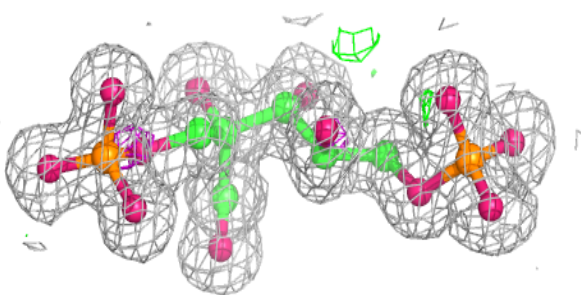
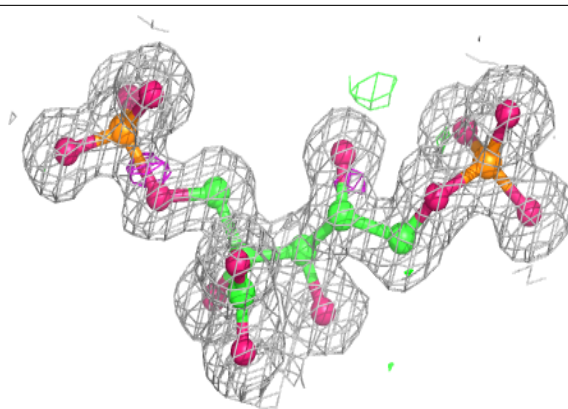
**Electron density around CAP C 1477:**

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

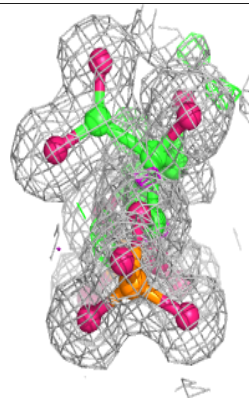
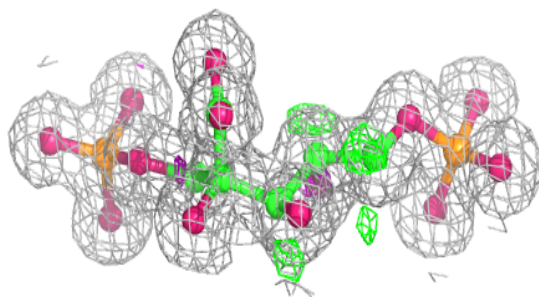
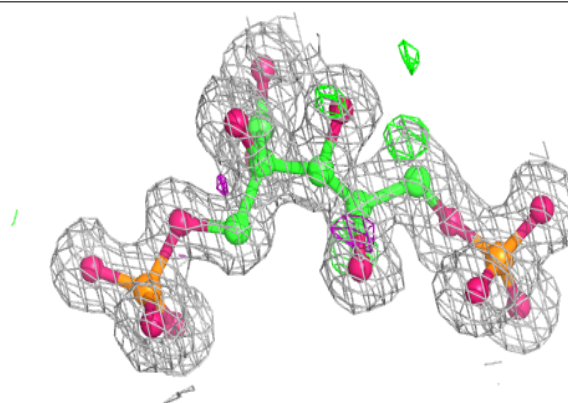


Electron density around CAP A 1477:

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

**Electron density around CAP E 1477:**

$2mF_o-DF_c$ (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.