



# Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 09:20 pm BST

PDB ID : 3ZXW  
Title : STRUCTURE OF ACTIVATED RUBISCO FROM THERMOSYNE-  
CHOCOCCUS ELONGATUS COMPLEXED WITH 2-CARBOXYARABI  
NITOL-1,5-DIPHOSPHATE  
Authors : Terlecka, B.; Wilhelmi, V.; Bialek, W.; Gubernator, B.; Szczepaniak, A.; Hof-  
mann, E.  
Deposited on : 2011-08-16  
Resolution : 2.10 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

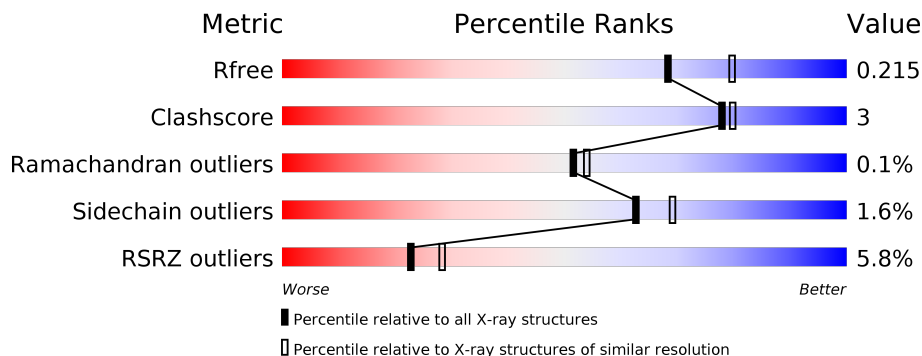
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

# 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 2.10 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.



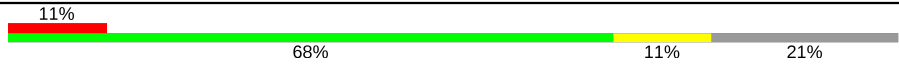

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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  $\geq 3$ , 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	475	<p>3% 91% 7%</p>
1	C	475	<p>2% 89% 8%</p>
1	E	475	<p>4% 91% 6%</p>
1	G	475	<p>3% 91% 7%</p>
2	B	118	<p>14% 69% 9% 21%</p>
2	D	118	<p>12% 69% 9% 21%</p>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	118	 <p>11% 68% 11% 21%</p>
2	H	118	 <p>22% 69% 9% 21%</p>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 18787 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 BISPHOSPHATE CARBOXYLASE LARGE CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	464	3664	2329	637	677	21	0	2	0
1	C	464	3664	2329	637	677	21	0	2	0
1	E	464	3664	2329	637	677	21	0	2	0
1	G	464	3664	2329	637	677	21	0	2	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	93	768	492	128	141	7	0	0	0
2	D	93	768	492	128	141	7	0	0	0
2	F	93	768	492	128	141	7	0	0	0
2	H	93	768	492	128	141	7	0	0	0

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

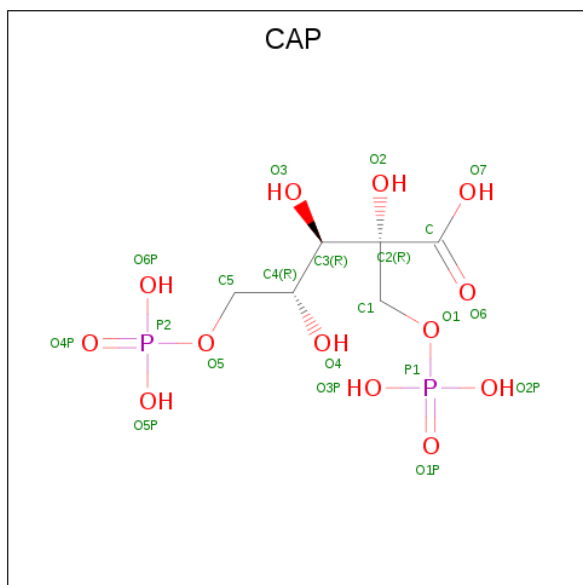
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

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

- Molecule 4 is 2-CARBOXYARABINITOL-1,5-DIPHOSPHATE (three-letter code: CAP) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>13</sub>P<sub>2</sub>).



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 GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	E	1	Total C O 6 3 3	0	0
5	E	1	Total C O 6 3 3	0	0
5	G	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	235	Total O 235 235	0	0

*Continued on next page...*

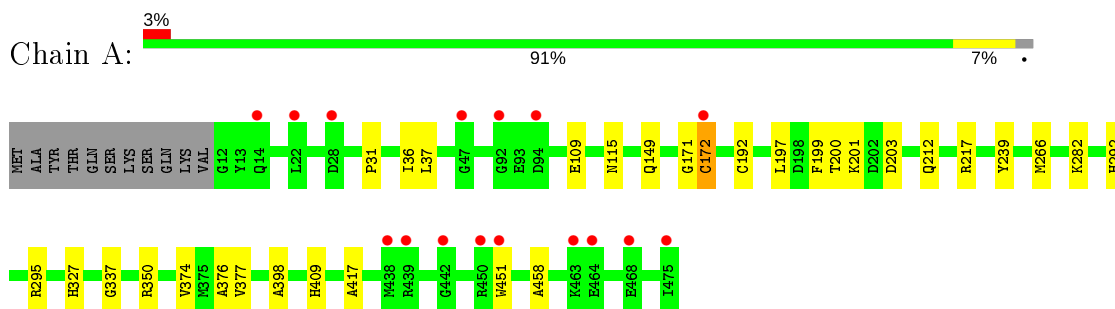
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	B	14	Total 14	O 14	0	0
6	C	226	Total 226	O 226	0	0
6	D	9	Total 9	O 9	0	0
6	E	201	Total 201	O 201	0	0
6	F	9	Total 9	O 9	0	0
6	G	204	Total 204	O 204	0	0
6	H	7	Total 7	O 7	0	0

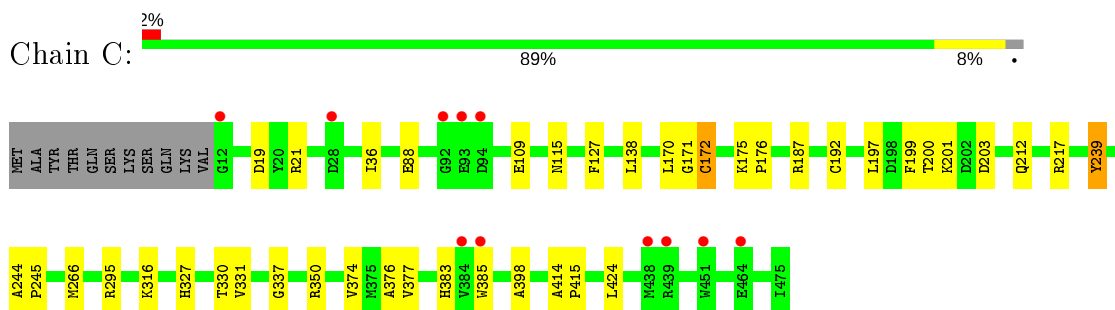
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

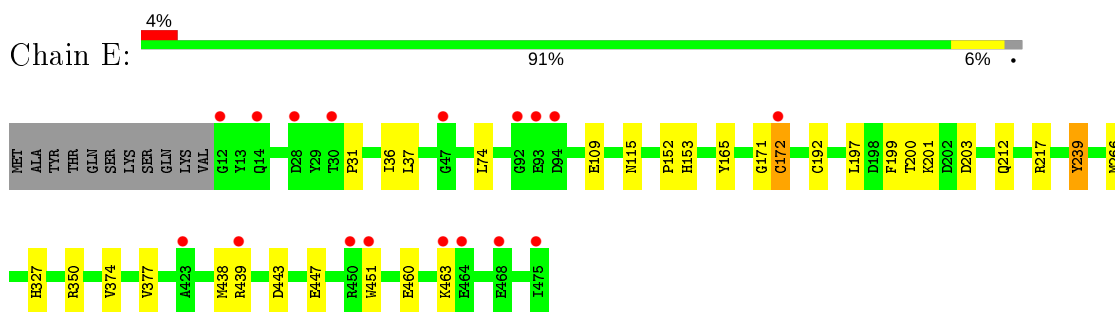
- Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN



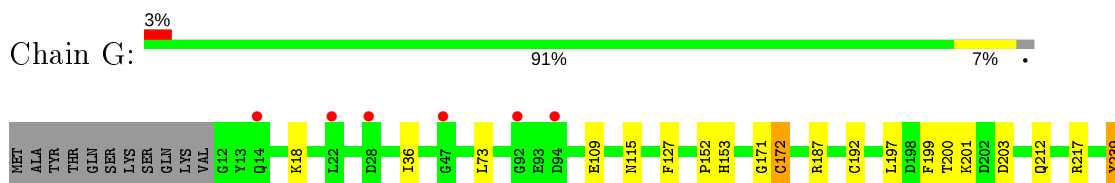
- Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN



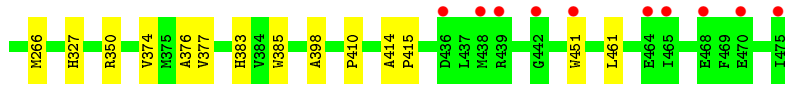
- Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN



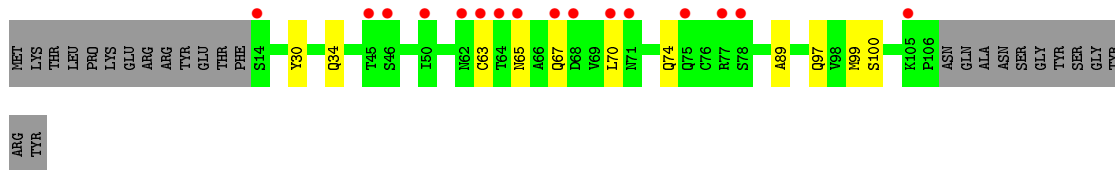
- Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN



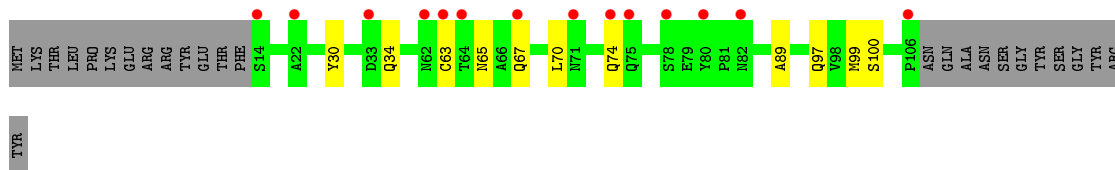




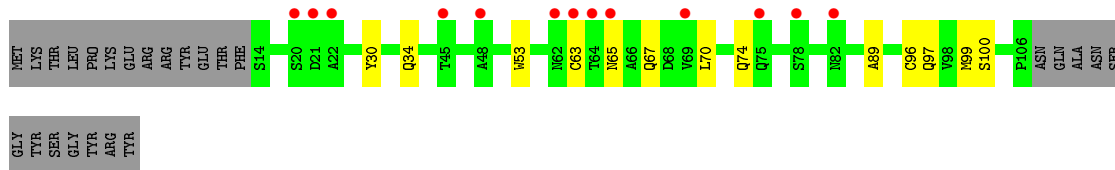
● Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN



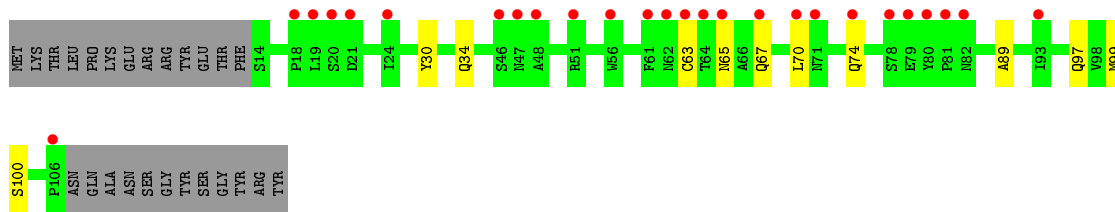
● Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN



● Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN



● Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.50Å 111.50Å 397.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.92 – 2.10 19.92 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.92-2.10) 98.8 (19.92-2.10)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 2.09Å)	Xtrriage
Refinement program	REFMAC 5.5.0088	Depositor
R, $R_{free}$	0.181 , 0.213 0.183 , 0.215	Depositor DCC
$R_{free}$ test set	7450 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.4	Xtrriage
Anisotropy	0.104	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 58.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18787	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, CAP, 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	0.39	0/3744	0.51	0/5074
1	C	0.39	0/3744	0.52	0/5074
1	E	0.39	0/3744	0.52	0/5074
1	G	0.38	0/3744	0.51	0/5074
2	B	0.34	0/789	0.46	0/1075
2	D	0.34	0/789	0.46	0/1075
2	F	0.35	0/789	0.46	0/1075
2	H	0.35	0/789	0.46	0/1075
All	All	0.38	0/18132	0.51	0/24596

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3664	0	3576	19	0
1	C	3664	0	3576	24	0
1	E	3664	0	3576	18	0
1	G	3664	0	3576	20	0
2	B	768	0	735	5	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	768	0	735	5	0
2	F	768	0	735	8	0
2	H	768	0	735	5	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	9	1	0
4	C	21	0	9	1	0
4	E	21	0	8	1	0
4	G	21	0	9	1	0
5	A	24	0	32	0	0
5	C	24	0	32	1	0
5	E	12	0	16	0	0
5	G	6	0	8	1	0
6	A	235	0	0	0	0
6	B	14	0	0	0	0
6	C	226	0	0	0	1
6	D	9	0	0	0	0
6	E	201	0	0	0	0
6	F	9	0	0	1	0
6	G	204	0	0	0	1
6	H	7	0	0	0	0
All	All	18787	0	17367	102	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 3.

All (102) 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:A:212:GLN:OE1	1:A:217:ARG:HD3	1.86	0.76
1:C:201:KCX:OQ2	4:C:477:CAP:O3	2.05	0.74
1:E:201:KCX:OQ2	4:E:477:CAP:O3	2.09	0.70
1:G:109:GLU:H	1:G:115:ASN:ND2	1.92	0.68
1:G:18:LYS:H	5:G:1476:GOL:H12	1.59	0.67
1:A:149:GLN:HE22	1:A:282:LYS:HA	1.61	0.65
1:E:212:GLN:OE1	1:E:217:ARG:HD3	1.96	0.65
1:G:201:KCX:OQ2	4:G:477:CAP:O3	2.10	0.64
1:A:36[B]:ILE:HD12	1:A:36[B]:ILE:N	2.13	0.63
1:A:201:KCX:HB2	1:A:239:TYR:CD2	2.34	0.62

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:36[B]:ILE:N	1:E:36[B]:ILE:HD12	2.14	0.62
1:G:212:GLN:OE1	1:G:217:ARG:HD3	1.99	0.62
1:G:109:GLU:H	1:G:115:ASN:HD22	1.49	0.60
2:F:96:CYS:HB2	6:F:2008:HOH:O	2.01	0.60
1:C:109:GLU:H	1:C:115:ASN:ND2	2.01	0.59
1:A:171:GLY:HA2	1:A:199:PHE:O	2.02	0.59
1:G:171:GLY:HA2	1:G:199:PHE:O	2.04	0.58
1:C:36[B]:ILE:N	1:C:36[B]:ILE:HD12	2.18	0.58
1:G:172:CYS:SG	1:G:200:THR:HG22	2.43	0.58
1:G:239:TYR:HB3	1:G:266:MET:HB3	1.87	0.57
1:G:36[B]:ILE:N	1:G:36[B]:ILE:HD12	2.20	0.57
1:A:109:GLU:H	1:A:115:ASN:ND2	2.03	0.56
1:A:239:TYR:HB3	1:A:266:MET:HB3	1.86	0.56
1:G:201:KCX:HB2	1:G:239:TYR:CD2	2.41	0.56
1:E:172:CYS:SG	1:E:200:THR:HG22	2.46	0.55
1:E:172:CYS:HB3	1:E:197:LEU:CD1	2.38	0.54
1:C:201:KCX:HB2	1:C:239:TYR:CD2	2.42	0.54
1:A:172:CYS:HB3	1:A:197:LEU:CD1	2.37	0.54
1:C:212:GLN:OE1	1:C:217:ARG:HD3	2.09	0.52
1:E:109:GLU:H	1:E:115:ASN:ND2	2.08	0.52
1:A:201:KCX:OQ2	4:A:477:CAP:O3	2.15	0.52
1:C:172:CYS:HB3	1:C:197:LEU:CD1	2.40	0.52
1:E:31:PRO:HB3	1:E:37:LEU:HD21	1.92	0.51
1:C:171:GLY:HA2	1:C:199:PHE:O	2.10	0.51
1:C:172:CYS:SG	1:C:200:THR:HG22	2.51	0.51
1:E:171:GLY:HA2	1:E:199:PHE:O	2.09	0.51
1:C:414:ALA:HB3	1:C:415:PRO:HD3	1.94	0.50
1:C:383:HIS:CE1	1:C:385:TRP:HB2	2.47	0.50
2:F:97:GLN:HE21	2:F:100:SER:HB2	1.77	0.49
1:C:109:GLU:H	1:C:115:ASN:HD22	1.61	0.48
2:D:97:GLN:HE21	2:D:100:SER:HB2	1.79	0.48
2:B:97:GLN:HE21	2:B:100:SER:HB2	1.79	0.48
1:E:201:KCX:HB2	1:E:239:TYR:CD2	2.48	0.48
1:G:414:ALA:HB3	1:G:415:PRO:HD3	1.95	0.48
1:E:239:TYR:HB3	1:E:266:MET:HB3	1.96	0.48
1:G:410:PRO:HD3	1:G:461:LEU:HD22	1.96	0.47
1:G:172:CYS:HB3	1:G:197:LEU:CD1	2.44	0.47
2:F:70:LEU:O	2:F:74:GLN:HG2	2.15	0.47
2:H:97:GLN:HE21	2:H:100:SER:HB2	1.79	0.47
1:A:31:PRO:HB3	1:A:37:LEU:HD21	1.97	0.47
1:E:327:HIS:HA	1:E:377:VAL:HB	1.97	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:70:LEU:O	2:D:74:GLN:HG2	2.15	0.47
1:G:383:HIS:CE1	1:G:385:TRP:HB2	2.49	0.47
1:C:327:HIS:HA	1:C:377:VAL:HB	1.96	0.46
2:H:70:LEU:O	2:H:74:GLN:HG2	2.15	0.46
1:E:36[B]:ILE:N	1:E:36[B]:ILE:CD1	2.78	0.46
1:A:36[B]:ILE:CD1	1:A:36[B]:ILE:N	2.78	0.46
2:B:70:LEU:O	2:B:74:GLN:HG2	2.16	0.46
1:E:443:ASP:O	1:E:447:GLU:HB2	2.15	0.46
1:A:172:CYS:SG	1:A:200:THR:HG22	2.57	0.45
1:A:295:ARG:HG2	1:A:327:HIS:HB2	1.99	0.45
1:C:244:ALA:HB1	1:C:245:PRO:HD2	1.99	0.45
1:E:109:GLU:H	1:E:115:ASN:HD22	1.63	0.45
1:E:152:PRO:HB2	1:E:153:HIS:CD2	2.52	0.45
1:C:19:ASP:HB3	1:C:21:ARG:HG2	1.98	0.45
1:A:327:HIS:HA	1:A:377:VAL:HB	2.00	0.44
1:E:350:ARG:HG2	1:E:374:VAL:O	2.17	0.43
1:G:376:ALA:O	1:G:398:ALA:HA	2.19	0.43
1:C:239:TYR:HB3	1:C:266:MET:HB3	2.00	0.43
1:C:36[B]:ILE:CD1	1:C:36[B]:ILE:N	2.81	0.43
2:D:30:TYR:O	2:D:34:GLN:HG2	2.18	0.43
2:F:65:ASN:HD21	2:F:67:GLN:HB3	1.84	0.43
2:H:30:TYR:O	2:H:34:GLN:HG2	2.19	0.43
1:C:175:LYS:HA	1:C:176:PRO:C	2.40	0.43
2:F:30:TYR:O	2:F:34:GLN:HG2	2.19	0.43
1:C:330:THR:O	1:C:331:VAL:HB	2.18	0.42
1:C:170:LEU:HG	1:C:424:LEU:HD22	2.00	0.42
2:B:65:ASN:HD21	2:B:67:GLN:HB3	1.84	0.42
1:G:350:ARG:HG2	1:G:374:VAL:O	2.19	0.42
2:F:89:ALA:HB3	2:F:99:MET:HG2	2.02	0.42
1:E:460:GLU:HA	1:E:463:LYS:HE3	2.01	0.42
1:G:327:HIS:HA	1:G:377:VAL:HB	2.02	0.42
1:C:295:ARG:HG2	1:C:327:HIS:HB2	2.02	0.42
2:H:65:ASN:HD21	2:H:67:GLN:HB3	1.85	0.42
2:B:30:TYR:O	2:B:34:GLN:HG2	2.19	0.41
2:B:89:ALA:HB3	2:B:99:MET:HG2	2.01	0.41
1:C:350:ARG:HG2	1:C:374:VAL:O	2.21	0.41
1:C:88:GLU:HG3	5:C:1477:GOL:H12	2.01	0.41
1:E:165:TYR:CD1	2:F:97:GLN:HB3	2.55	0.41
2:F:53:TRP:O	1:G:187:ARG:NH1	2.50	0.41
1:G:152:PRO:HB2	1:G:153:HIS:CD2	2.56	0.41
2:D:89:ALA:HB3	2:D:99:MET:HG2	2.02	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:36[B]:ILE:N	1:G:36[B]:ILE:CD1	2.83	0.41
1:A:197:LEU:HG	1:A:417:ALA:HB1	2.03	0.41
1:A:266:MET:HA	1:A:292:HIS:O	2.21	0.41
1:C:376:ALA:O	1:C:398:ALA:HA	2.21	0.41
1:C:138:LEU:O	1:C:316:LYS:NZ	2.36	0.41
1:A:376:ALA:O	1:A:398:ALA:HA	2.20	0.40
1:A:409:HIS:CD2	1:A:458:ALA:HB2	2.57	0.40
2:H:89:ALA:HB3	2:H:99:MET:HG2	2.03	0.40
1:A:350:ARG:HG2	1:A:374:VAL:O	2.21	0.40
2:D:65:ASN:HD21	2:D:67:GLN:HB3	1.85	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)
6:G:2134:HOH:O	6:G:2134:HOH:O[7_555]	1.70	0.50
6:C:2169:HOH:O	6:C:2169:HOH:O[7_555]	2.10	0.10

## 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	463/475 (98%)	450 (97%)	12 (3%)	1 (0%)	47	49
1	C	463/475 (98%)	451 (97%)	11 (2%)	1 (0%)	47	49
1	E	463/475 (98%)	449 (97%)	14 (3%)	0	100	100
1	G	463/475 (98%)	451 (97%)	12 (3%)	0	100	100
2	B	91/118 (77%)	90 (99%)	1 (1%)	0	100	100
2	D	91/118 (77%)	90 (99%)	1 (1%)	0	100	100
2	F	91/118 (77%)	90 (99%)	1 (1%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	91/118 (77%)	90 (99%)	1 (1%)	0	100	100
All	All	2216/2372 (93%)	2161 (98%)	53 (2%)	2 (0%)	51	54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	337	GLY
1	C	337	GLY

### 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	377/385 (98%)	373 (99%)	4 (1%)	73	79
1	C	377/385 (98%)	371 (98%)	6 (2%)	62	69
1	E	377/385 (98%)	369 (98%)	8 (2%)	53	59
1	G	377/385 (98%)	370 (98%)	7 (2%)	57	63
2	B	86/108 (80%)	85 (99%)	1 (1%)	71	77
2	D	86/108 (80%)	85 (99%)	1 (1%)	71	77
2	F	86/108 (80%)	85 (99%)	1 (1%)	71	77
2	H	86/108 (80%)	85 (99%)	1 (1%)	71	77
All	All	1852/1972 (94%)	1823 (98%)	29 (2%)	62	69

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	172	CYS
1	A	192	CYS
1	A	203	ASP
1	A	451	TRP
2	B	63	CYS
1	C	127	PHE

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	C	172	CYS
1	C	187	ARG
1	C	192	CYS
1	C	203	ASP
1	C	239	TYR
2	D	63	CYS
1	E	74	LEU
1	E	172	CYS
1	E	192	CYS
1	E	203	ASP
1	E	239	TYR
1	E	438	MET
1	E	439	ARG
1	E	451	TRP
2	F	63	CYS
1	G	73	LEU
1	G	127	PHE
1	G	172	CYS
1	G	192	CYS
1	G	203	ASP
1	G	239	TYR
1	G	451	TRP
2	H	63	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	115	ASN
1	A	149	GLN
1	A	153	HIS
1	A	401	GLN
2	B	65	ASN
2	B	97	GLN
1	C	115	ASN
1	C	153	HIS
1	C	401	GLN
2	D	65	ASN
2	D	97	GLN
1	E	115	ASN
1	E	153	HIS
1	E	401	GLN
2	F	65	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	F	97	GLN
1	G	115	ASN
1	G	153	HIS
1	G	401	GLN
2	H	65	ASN
2	H	97	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	KCX	E	201	1,3	7,11,12	0.54	0	4,12,14	1.69	1 (25%)
1	KCX	G	201	1,3	7,11,12	0.46	0	4,12,14	1.73	1 (25%)
1	KCX	A	201	1,3	7,11,12	0.50	0	4,12,14	1.76	1 (25%)
1	KCX	C	201	1,3	7,11,12	0.53	0	4,12,14	1.79	1 (25%)

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	E	201	1,3	-	2/7/10/12	-
1	KCX	G	201	1,3	-	2/7/10/12	-
1	KCX	A	201	1,3	-	1/7/10/12	-
1	KCX	C	201	1,3	-	1/7/10/12	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	201	KCX	CD-CE-NZ	-3.40	102.36	111.49
1	G	201	KCX	CD-CE-NZ	-3.33	102.53	111.49
1	E	201	KCX	CD-CE-NZ	-3.28	102.68	111.49
1	C	201	KCX	CD-CE-NZ	-3.08	103.22	111.49

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	201	KCX	CG-CD-CE-NZ
1	C	201	KCX	CG-CD-CE-NZ
1	G	201	KCX	CG-CD-CE-NZ
1	E	201	KCX	CG-CD-CE-NZ
1	G	201	KCX	CD-CE-NZ-CX
1	E	201	KCX	CD-CE-NZ-CX

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	201	KCX	2	0
1	G	201	KCX	2	0
1	A	201	KCX	2	0
1	C	201	KCX	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 4 are monoatomic - leaving 15 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	GOL	A	1479	-	5,5,5	0.36	0	5,5,5	0.33	0
5	GOL	E	1477	-	5,5,5	0.37	0	5,5,5	0.16	0
5	GOL	A	1477	-	5,5,5	0.35	0	5,5,5	0.23	0
5	GOL	C	1477	-	5,5,5	0.36	0	5,5,5	0.18	0
4	CAP	G	477	3	15,20,20	0.83	0	20,31,31	0.67	0
4	CAP	E	477	3	15,20,20	0.85	0	20,31,31	0.57	0
5	GOL	E	1476	-	5,5,5	0.35	0	5,5,5	0.31	0
4	CAP	C	477	3	15,20,20	0.84	0	20,31,31	0.64	0
5	GOL	G	1476	-	5,5,5	0.38	0	5,5,5	0.25	0
4	CAP	A	477	3	15,20,20	0.83	0	20,31,31	0.59	0
5	GOL	C	1479	-	5,5,5	0.35	0	5,5,5	0.21	0
5	GOL	C	1476	-	5,5,5	0.36	0	5,5,5	0.21	0
5	GOL	C	1478	-	5,5,5	0.33	0	5,5,5	0.39	0
5	GOL	A	1478	-	5,5,5	0.37	0	5,5,5	0.30	0
5	GOL	A	1476	-	5,5,5	0.35	0	5,5,5	0.21	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	GOL	A	1479	-	-	2/4/4/4	-
5	GOL	E	1477	-	-	1/4/4/4	-
5	GOL	A	1477	-	-	3/4/4/4	-
5	GOL	C	1477	-	-	2/4/4/4	-
4	CAP	G	477	3	-	3/23/29/29	-
4	CAP	E	477	3	-	3/23/29/29	-
5	GOL	E	1476	-	-	0/4/4/4	-
4	CAP	C	477	3	-	3/23/29/29	-
5	GOL	G	1476	-	-	2/4/4/4	-
4	CAP	A	477	3	-	5/23/29/29	-
5	GOL	C	1479	-	-	2/4/4/4	-
5	GOL	C	1476	-	-	0/4/4/4	-
5	GOL	C	1478	-	-	2/4/4/4	-
5	GOL	A	1478	-	-	2/4/4/4	-
5	GOL	A	1476	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (34) torsion outliers are listed below:

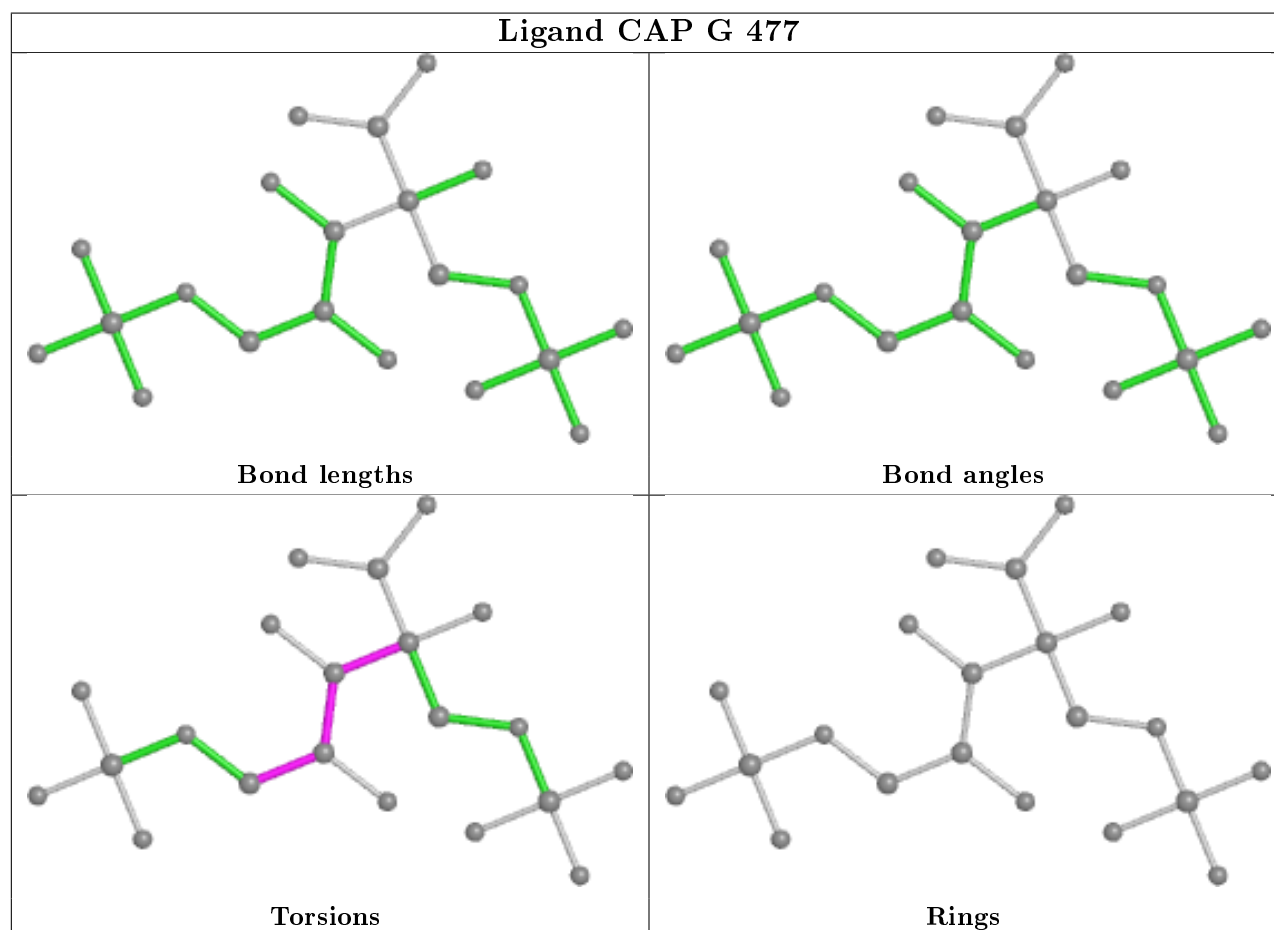
Mol	Chain	Res	Type	Atoms
5	C	1477	GOL	C1-C2-C3-O3
5	C	1477	GOL	O2-C2-C3-O3
4	G	477	CAP	O3-C3-C4-O4
4	E	477	CAP	C2-C3-C4-O4
4	E	477	CAP	O3-C3-C4-O4
4	C	477	CAP	C2-C3-C4-O4
4	C	477	CAP	O3-C3-C4-O4
4	A	477	CAP	O1-C1-C2-O2
4	A	477	CAP	C2-C3-C4-O4
4	A	477	CAP	O3-C3-C4-O4
5	C	1478	GOL	O1-C1-C2-C3
5	C	1478	GOL	O1-C1-C2-O2
5	A	1479	GOL	C1-C2-C3-O3
5	G	1476	GOL	O1-C1-C2-C3
5	C	1479	GOL	C1-C2-C3-O3
5	A	1476	GOL	C1-C2-C3-O3
5	A	1478	GOL	O1-C1-C2-C3
4	G	477	CAP	O2-C2-C3-C4
4	E	477	CAP	O2-C2-C3-C4
4	C	477	CAP	O2-C2-C3-C4
4	A	477	CAP	O2-C2-C3-C4
5	A	1479	GOL	O2-C2-C3-O3
5	A	1476	GOL	O2-C2-C3-O3
5	G	1476	GOL	O1-C1-C2-O2
5	A	1476	GOL	O1-C1-C2-O2
5	E	1477	GOL	C1-C2-C3-O3
5	C	1479	GOL	O2-C2-C3-O3
5	A	1478	GOL	O1-C1-C2-O2
5	A	1477	GOL	C1-C2-C3-O3
5	A	1476	GOL	O1-C1-C2-C3
5	A	1477	GOL	O2-C2-C3-O3
4	G	477	CAP	O4-C4-C5-O5
5	A	1477	GOL	O1-C1-C2-C3
4	A	477	CAP	O1-C1-C2-C3

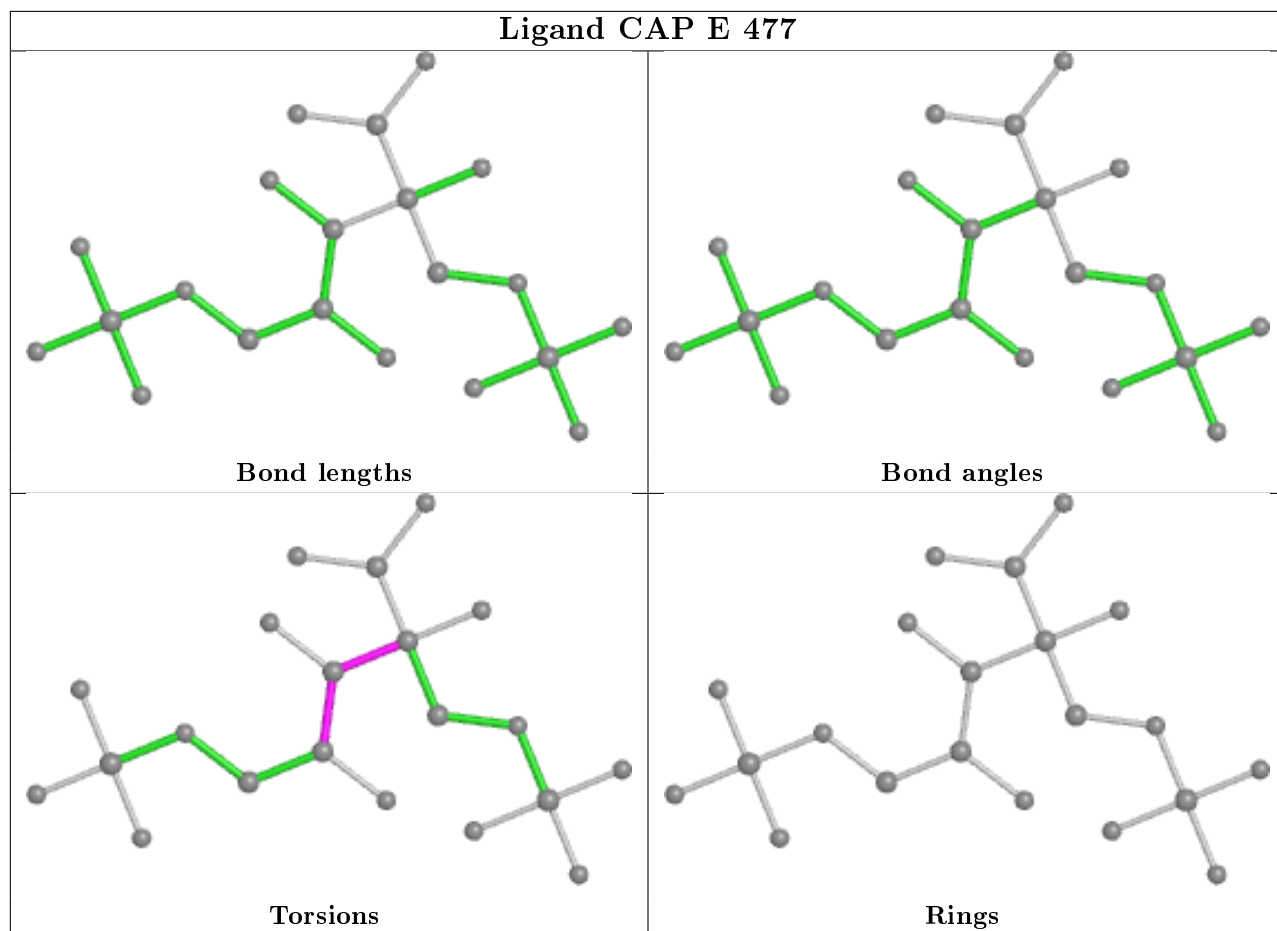
There are no ring outliers.

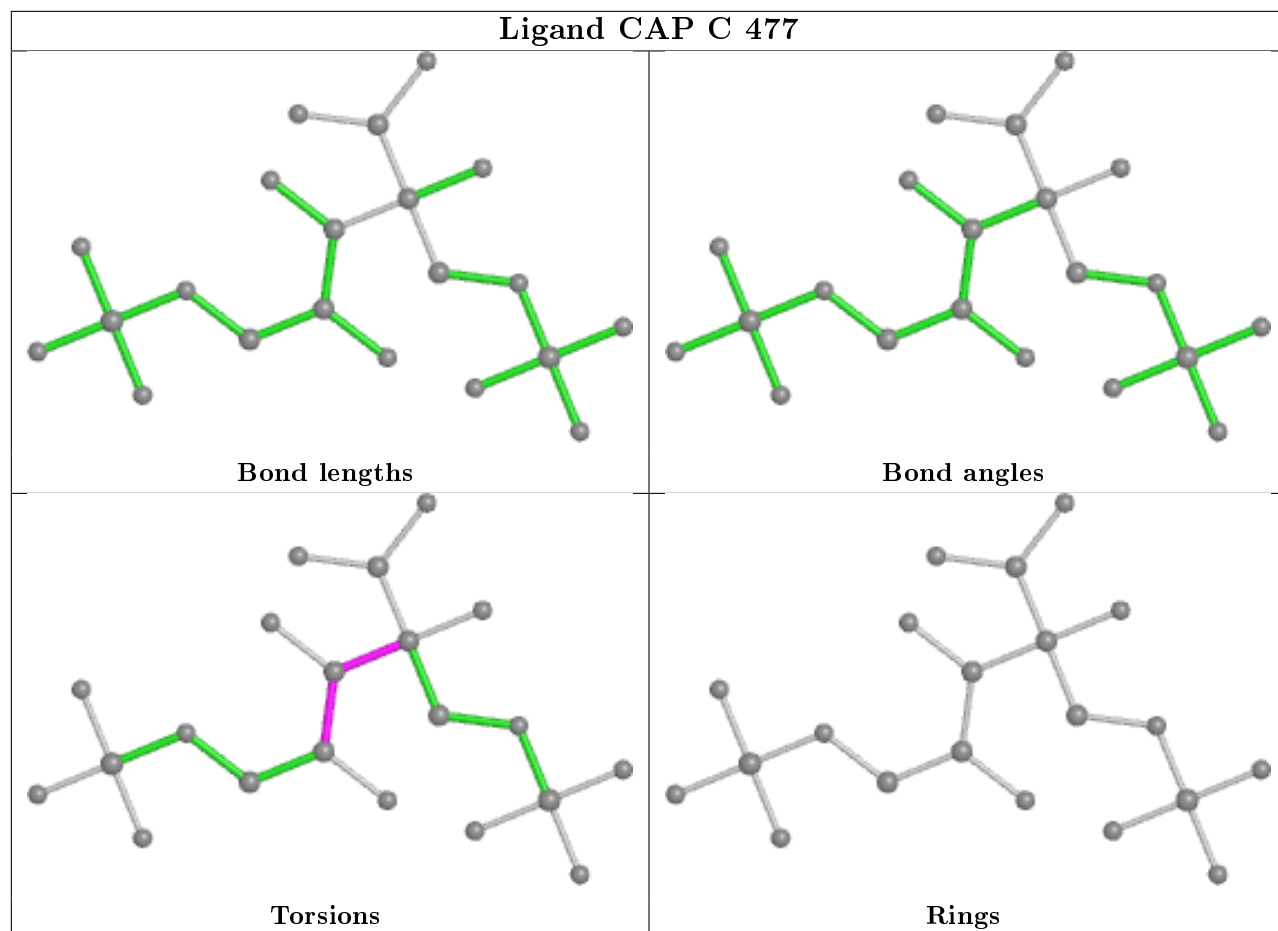
6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	1477	GOL	1	0
4	G	477	CAP	1	0
4	E	477	CAP	1	0
4	C	477	CAP	1	0
5	G	1476	GOL	1	0
4	A	477	CAP	1	0

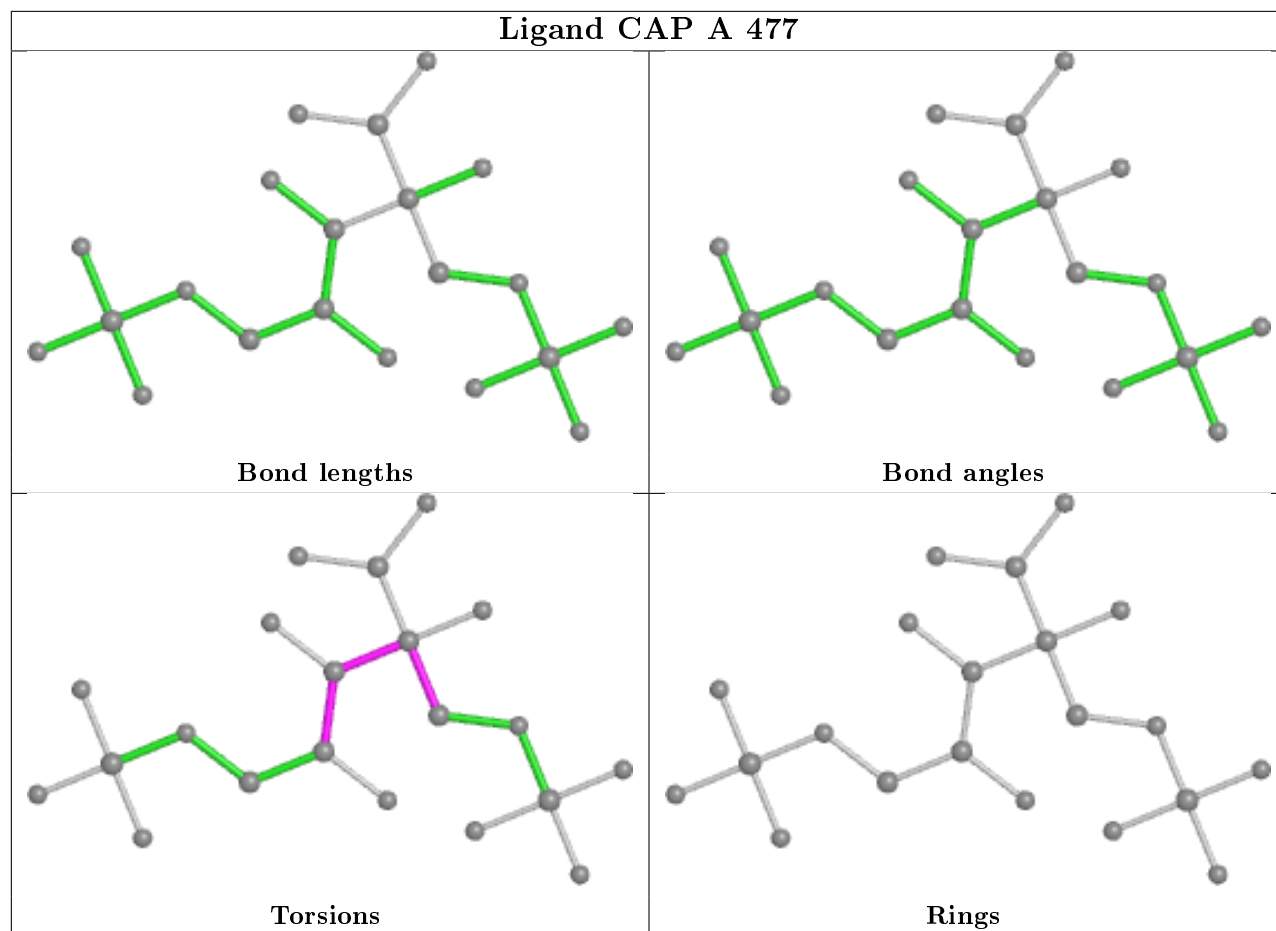
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	463/475 (97%)	-0.14	16 (3%) 44 50	14, 30, 55, 95	0
1	C	463/475 (97%)	-0.20	11 (2%) 59 64	13, 29, 55, 95	0
1	E	463/475 (97%)	-0.17	17 (3%) 41 48	14, 30, 55, 95	0
1	G	463/475 (97%)	-0.28	16 (3%) 44 50	14, 30, 55, 95	0
2	B	93/118 (78%)	0.89	16 (17%) 1 1	33, 55, 71, 76	0
2	D	93/118 (78%)	0.97	14 (15%) 2 3	33, 55, 71, 76	0
2	F	93/118 (78%)	0.99	13 (13%) 2 3	33, 55, 71, 76	0
2	H	93/118 (78%)	1.42	26 (27%) 0 0	33, 55, 71, 104	0
All	All	2224/2372 (93%)	0.02	129 (5%) 23 28	13, 33, 65, 104	0

All (129) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	63	CYS	6.4
2	H	63	CYS	6.4
2	H	62	ASN	6.2
2	D	62	ASN	6.1
1	E	451	TRP	5.7
2	H	64	THR	5.7
1	G	439	ARG	5.7
1	A	439	ARG	5.6
2	F	62	ASN	5.5
2	H	78	SER	5.3
2	B	62	ASN	5.2
2	F	64	THR	4.9
2	H	19	LEU	4.5
1	C	439	ARG	4.5
2	D	63	CYS	4.3
2	H	70	LEU	4.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	F	48	ALA	4.2
1	G	451	TRP	4.1
1	G	438	MET	4.1
1	E	14	GLN	4.1
1	A	475	ILE	4.0
1	A	438	MET	4.0
1	C	92	GLY	4.0
1	A	450	ARG	4.0
1	E	439	ARG	3.9
1	C	94	ASP	3.9
2	D	82	ASN	3.7
1	G	92	GLY	3.7
1	E	12	GLY	3.7
2	D	106	PRO	3.5
2	B	63	CYS	3.5
1	C	12	GLY	3.4
2	H	82	ASN	3.4
2	B	64	THR	3.4
1	G	470	GLU	3.4
1	E	172	CYS	3.3
1	A	464	GLU	3.3
1	A	172	CYS	3.3
2	B	70	LEU	3.3
2	H	93	ILE	3.2
2	B	75	GLN	3.2
2	B	78	SER	3.2
2	D	64	THR	3.2
2	F	82	ASN	3.1
2	B	105	LYS	3.0
2	H	48	ALA	3.0
2	H	74	GLN	3.0
2	H	20	SER	3.0
1	E	92	GLY	3.0
1	C	28	ASP	3.0
1	E	28	ASP	3.0
2	B	67	GLN	3.0
1	A	451	TRP	2.9
1	G	14	GLN	2.9
1	G	464	GLU	2.9
1	A	442	GLY	2.9
1	E	93	GLU	2.9
1	G	94	ASP	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	451	TRP	2.8
2	D	74	GLN	2.8
2	H	106	PRO	2.8
1	E	468	GLU	2.8
2	B	45	THR	2.8
1	C	93	GLU	2.8
1	E	47	GLY	2.7
2	F	78	SER	2.7
1	E	94	ASP	2.7
2	D	78	SER	2.7
1	E	30	THR	2.7
2	B	14	SER	2.6
1	G	22	LEU	2.6
2	F	21	ASP	2.6
1	A	22	LEU	2.6
2	D	14	SER	2.6
1	G	442	GLY	2.6
1	G	465	ILE	2.6
1	A	468	GLU	2.5
2	H	47	ASN	2.5
1	C	384	VAL	2.5
2	F	75	GLN	2.5
2	D	75	GLN	2.5
2	H	67	GLN	2.5
2	H	81	PRO	2.5
2	B	65	ASN	2.4
1	C	464	GLU	2.4
2	H	79	GLU	2.4
1	A	47	GLY	2.4
2	H	80	TYR	2.4
2	B	71	ASN	2.4
1	A	14	GLN	2.4
2	H	65	ASN	2.4
1	G	475	ILE	2.4
2	B	46	SER	2.4
2	D	80	TYR	2.4
2	H	46	SER	2.4
2	D	71	ASN	2.4
2	H	61	PHE	2.3
1	G	468	GLU	2.3
2	D	22	ALA	2.3
2	H	51	ARG	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	65	ASN	2.3
2	H	71	ASN	2.3
1	A	28	ASP	2.3
2	F	20	SER	2.3
1	E	464	GLU	2.3
2	F	22	ALA	2.3
2	F	45	THR	2.2
1	E	463	LYS	2.2
2	F	69	VAL	2.2
1	E	423	ALA	2.2
2	H	56	TRP	2.2
2	H	18	PRO	2.2
2	D	67	GLN	2.2
1	C	385	TRP	2.2
2	B	50	ILE	2.2
1	E	475	ILE	2.1
2	B	68	ASP	2.1
1	A	94	ASP	2.1
1	G	28	ASP	2.1
1	A	92	GLY	2.1
1	G	436	ASP	2.1
1	G	47	GLY	2.1
1	C	438	MET	2.1
1	E	450	ARG	2.1
1	A	463	LYS	2.0
2	D	33	ASP	2.0
2	H	21	ASP	2.0
2	H	24	ILE	2.0
2	B	77	ARG	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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(Å <sup>2</sup> )	Q<0.9
1	KCX	A	201	12/13	0.90	0.12	24,25,34,35	0
1	KCX	G	201	12/13	0.93	0.13	22,27,37,39	0
1	KCX	E	201	12/13	0.93	0.12	19,26,38,38	0
1	KCX	C	201	12/13	0.95	0.11	21,26,37,38	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

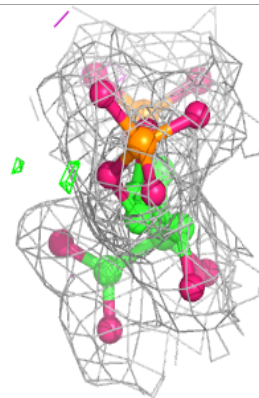
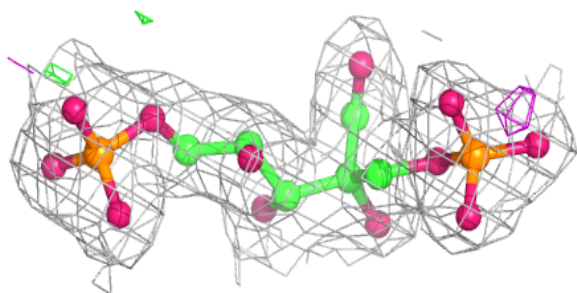
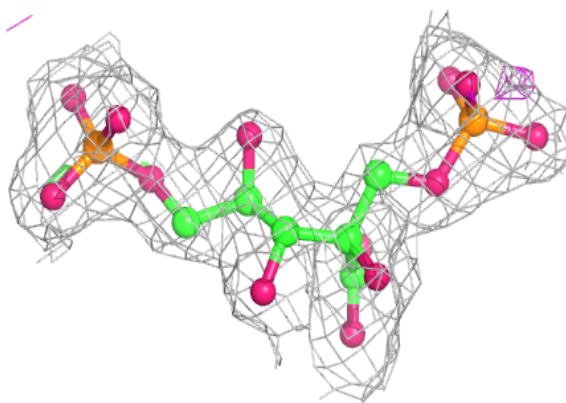
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
5	GOL	C	1479	6/6	0.65	0.35	40,48,68,94	0
5	GOL	A	1479	6/6	0.70	0.19	58,68,87,88	0
5	GOL	C	1477	6/6	0.80	0.19	41,66,74,100	0
5	GOL	E	1476	6/6	0.81	0.18	35,64,74,78	0
5	GOL	C	1478	6/6	0.82	0.17	42,64,74,101	0
5	GOL	A	1476	6/6	0.86	0.14	45,60,81,83	0
5	GOL	A	1478	6/6	0.87	0.22	53,56,86,144	0
5	GOL	C	1476	6/6	0.89	0.15	31,44,82,92	0
5	GOL	E	1477	6/6	0.89	0.40	48,56,65,73	0
5	GOL	A	1477	6/6	0.89	0.41	44,55,91,101	0
5	GOL	G	1476	6/6	0.91	0.14	46,51,59,61	0
3	MG	C	476	1/1	0.94	0.07	28,28,28,28	0
3	MG	G	476	1/1	0.96	0.11	29,29,29,29	0
4	CAP	C	477	21/21	0.97	0.08	22,28,39,57	0
3	MG	E	476	1/1	0.97	0.05	30,30,30,30	0
4	CAP	E	477	21/21	0.97	0.07	24,33,50,65	0
3	MG	A	476	1/1	0.98	0.08	35,35,35,35	0
4	CAP	G	477	21/21	0.98	0.06	22,29,40,154	0
4	CAP	A	477	21/21	0.98	0.08	25,34,46,70	0

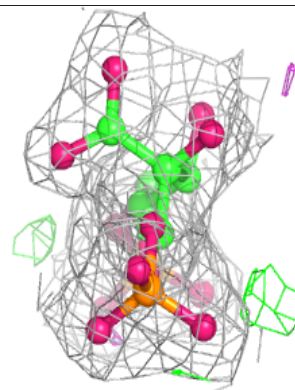
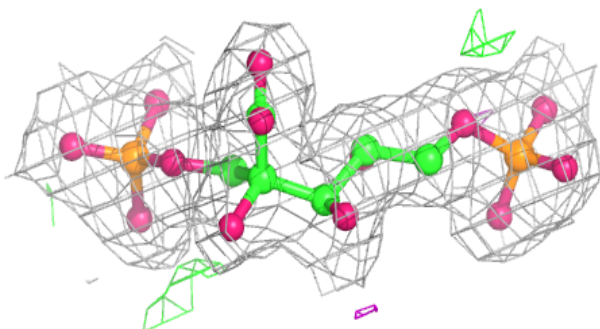
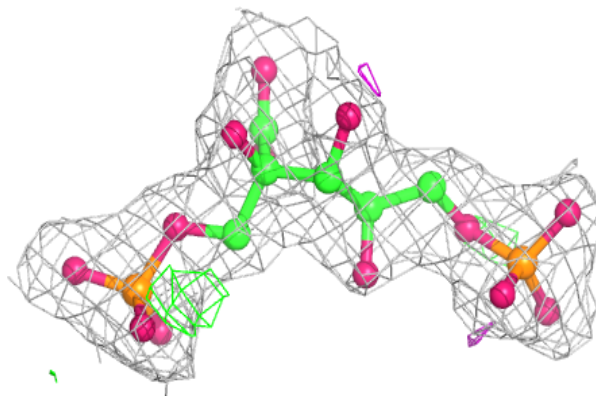
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

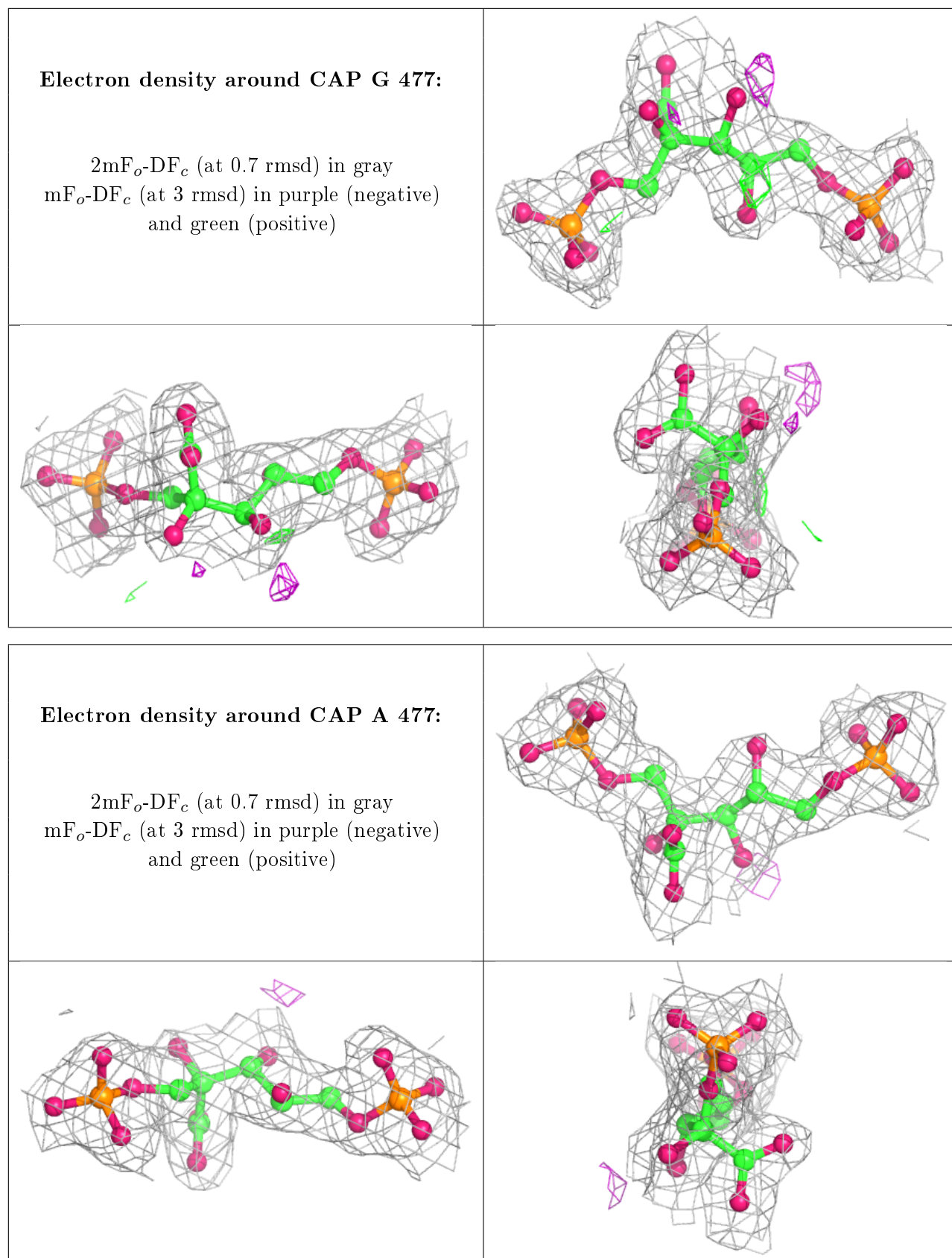
**Electron density around CAP C 477:**

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

**Electron density around CAP E 477:**

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







## 6.5 Other polymers

There are no such residues in this entry.