wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 22, 2023 – 08:55 PM JST

PDB ID : 7BYI
Title : Structure of SHMT2 in complex with CBX
Authors : Li, L.; Chen, Y.; Lu, D.; Zhang, C.; Su, D.
Deposited on : 2020-04-23
Resolution : 2.76 Å(reported)

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

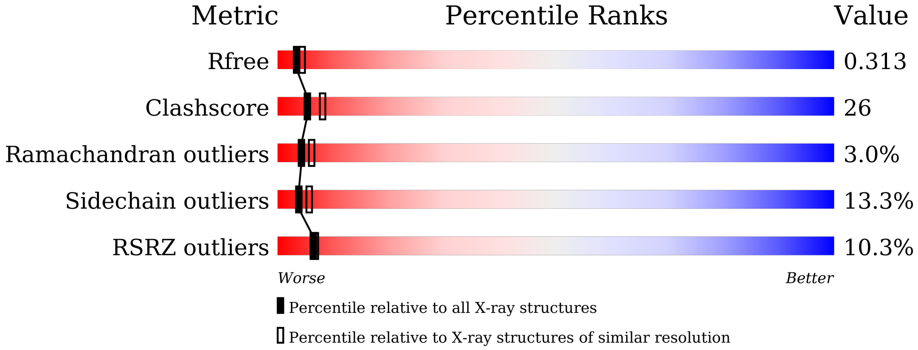
MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

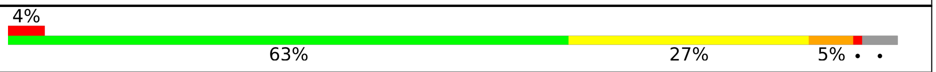
The reported resolution of this entry is 2.76 Å.

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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	483	
2	B	483	

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
5	CBO	B	601	X	-	-	-
5	CBO	B	602	X	-	-	X

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 7363 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 Serine hydroxymethyltransferase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	463	3609	2273	649	672	15	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	261	GLY	ALA	conflict	UNP P34897
A	343	ALA	CYS	conflict	UNP P34897

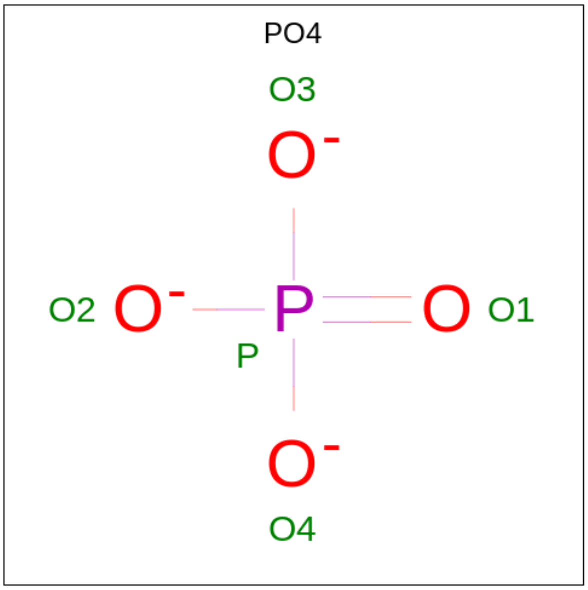
- Molecule 2 is a protein called Serine hydroxymethyltransferase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	462	3577	2251	644	666	16	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

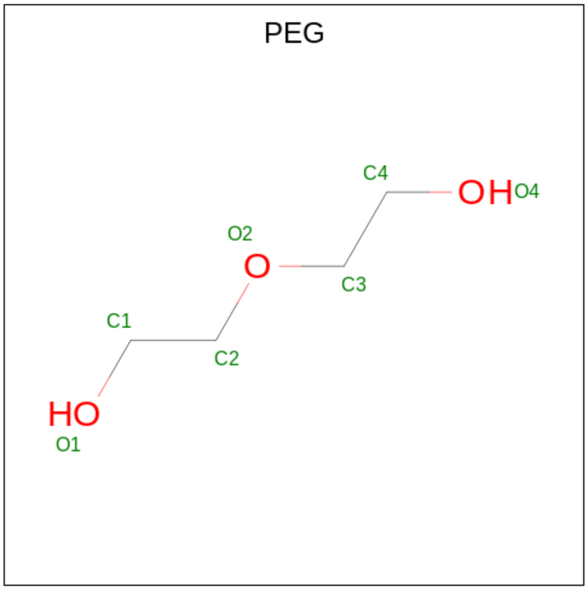
Chain	Residue	Modelled	Actual	Comment	Reference
B	196	GLY	LYS	conflict	UNP P34897
B	198	ALA	ASN	conflict	UNP P34897
B	307	ALA	ILE	conflict	UNP P34897
B	370	GLY	TYR	conflict	UNP P34897
B	371	TYR	SER	conflict	UNP P34897
B	391	GLY	LEU	conflict	UNP P34897
B	392	GLY	ASP	conflict	UNP P34897

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P) (labeled as "Ligand of Interest" by depositor).



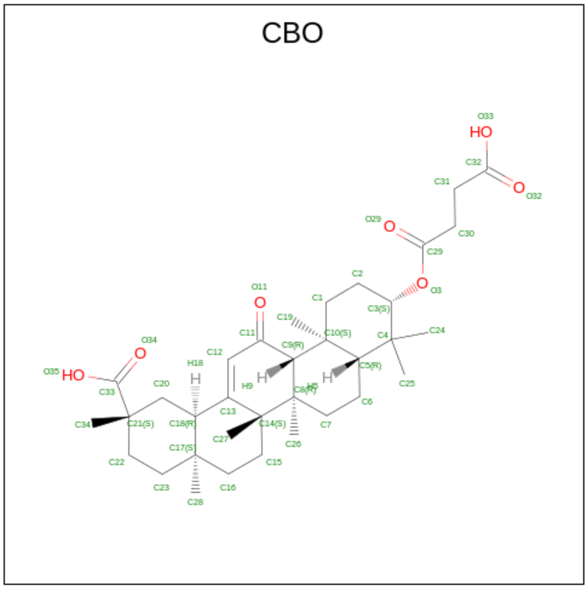
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 7 4 3	0	0

- Molecule 5 is CARBENOXOLONE (three-letter code: CBO) (formula: $C_{34}H_{50}O_7$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 41 34 7	0	0
5	B	1	Total C O 41 34 7	0	0

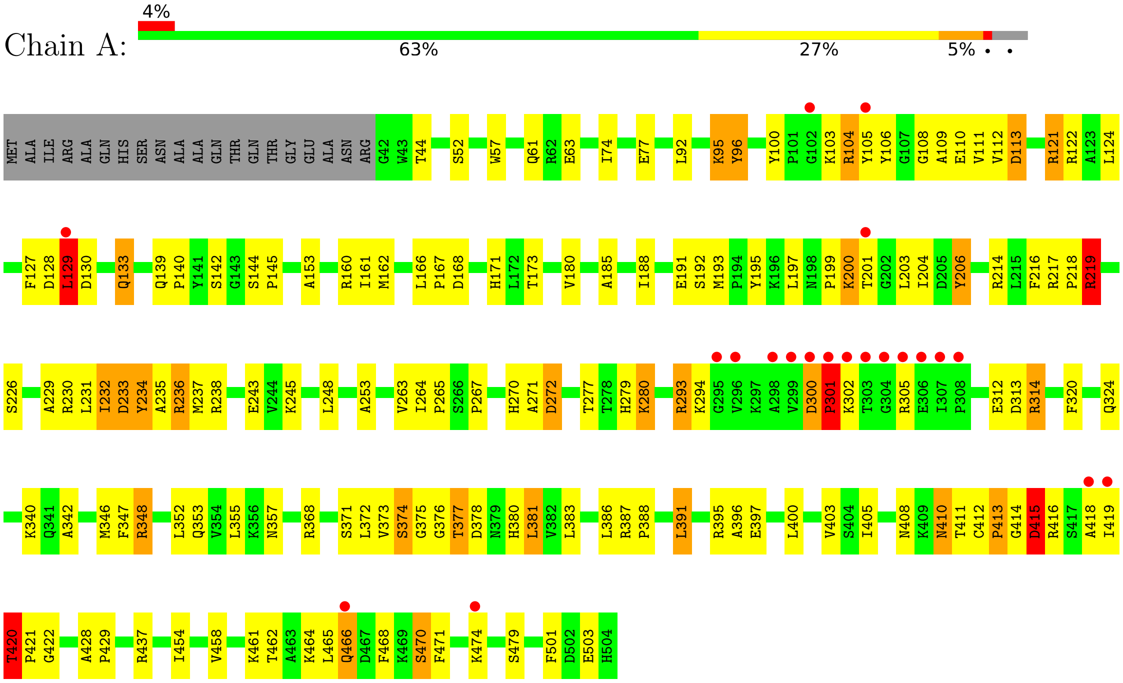
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	35	Total O 35 35	0	0
6	B	43	Total O 43 43	0	0

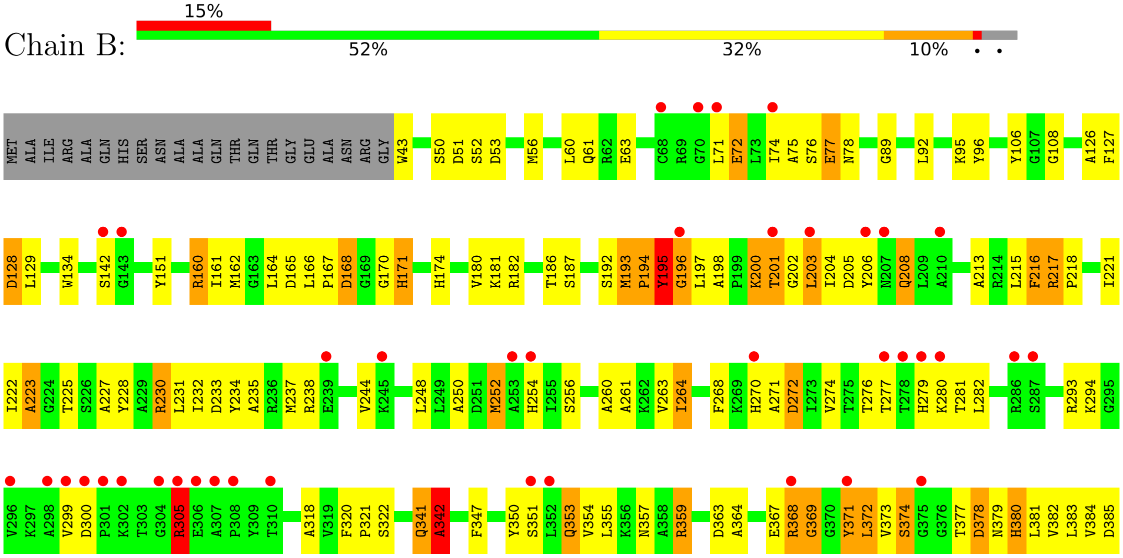
3 Residue-property plots ⓘ

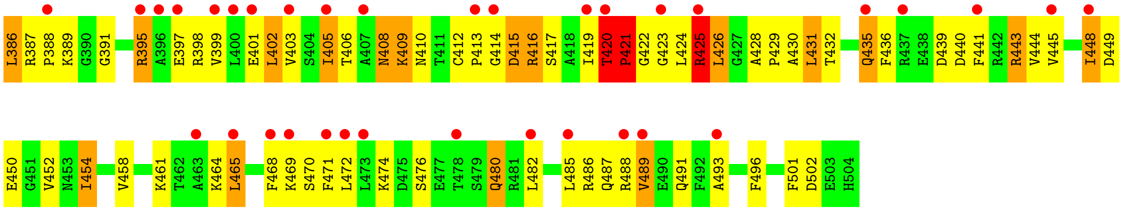
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Serine hydroxymethyltransferase, mitochondrial



- Molecule 2: Serine hydroxymethyltransferase, mitochondrial




L386
R387
P388
K389
G390
G391
R395
A396
E397
R398
V399
L400
E401
L402
V403
S404
I405
T406
A407
N408
K409
N410
T411
C412
P413
G414
D415
R416
S417
A418
I419
T420
P421
G422
G423
L424
R425
L426
G427
A428
P429
A430
L431
T432
Q435
F436
R437
E438
D439
D440
F441
R442
R443
V444
V445
I448
D449
E450
G451
V452
N453
I454
V458
K461
T462
A463
K464
L465
F468
K469
S470
F471
L472
L473
K474
D475
S476
E477
T478
S479
Q480
R481
L482
L485
R486
Q487
R488
V489
E490
Q491
F492
A493
F496
F501
D502
E503
H504

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	158.62Å 158.62Å 209.43Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.50 – 2.76 34.48 – 2.76	Depositor EDS
% Data completeness (in resolution range)	99.8 (34.50-2.76) 99.9 (34.48-2.76)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.05 (at 2.76Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, R_{free}	0.242 , 0.311 0.247 , 0.313	Depositor DCC
R_{free} test set	2127 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å^2)	56.0	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 53.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	7363	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.92% 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: PO4, CBO, PEG

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	0.81	0/3682	1.08	7/4978 (0.1%)
2	B	0.85	0/3650	1.14	14/4936 (0.3%)
All	All	0.83	0/7332	1.11	21/9914 (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

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	234	TYR	CB-CG-CD1	9.42	126.65	121.00
2	B	419	ILE	CB-CA-C	9.16	129.93	111.60
1	A	234	TYR	CB-CG-CD2	-9.12	115.53	121.00
2	B	341	GLN	N-CA-C	-8.91	86.94	111.00
1	A	103	LYS	N-CA-C	-7.66	90.31	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	ARG	Peptide
1	A	129	LEU	Peptide

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	3609	0	3602	138	0
2	B	3577	0	3555	244	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	7	0	10	0	0
5	B	82	0	96	17	0
6	A	35	0	0	3	0
6	B	43	0	0	4	0
All	All	7363	0	7263	384	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 26.

The worst 5 of 384 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:B:205:ASP:OD2	2:B:208:GLN:HB3	1.31	1.27
2:B:165:ASP:CB	2:B:195:TYR:HD2	1.56	1.16
2:B:165:ASP:HB2	2:B:195:TYR:HD2	1.09	1.14
2:B:165:ASP:HB2	2:B:195:TYR:CD2	1.91	1.05
1:A:420:THR:H	1:A:421:PRO:CD	1.67	1.03

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	461/483 (95%)	391 (85%)	56 (12%)	14 (3%)	4	6
2	B	460/483 (95%)	389 (85%)	57 (12%)	14 (3%)	4	6
All	All	921/966 (95%)	780 (85%)	113 (12%)	28 (3%)	4	6

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	109	ALA
1	A	129	LEU
1	A	201	THR
1	A	302	LYS
1	A	415	ASP

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	380/394 (96%)	340 (90%)	40 (10%)	7	11
2	B	373/389 (96%)	313 (84%)	60 (16%)	2	3
All	All	753/783 (96%)	653 (87%)	100 (13%)	4	6

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	264	ILE
2	B	386	LEU
2	B	489	VAL
2	B	281	THR
2	B	359	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	270	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	353	GLN
2	B	491	GLN
2	B	435	GLN
2	B	480	GLN

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	$\#\|Z\| > 2$
3	PO4	B	603	-	4,4,4	0.89	0	6,6,6	0.41	0
5	CBO	B	602	-	45,45,45	2.63	17 (37%)	76,76,76	2.17	25 (32%)
5	CBO	B	601	-	45,45,45	3.16	15 (33%)	76,76,76	2.74	35 (46%)
3	PO4	A	701	-	4,4,4	1.08	0	6,6,6	0.56	0
4	PEG	A	702	-	6,6,6	0.23	0	5,5,5	0.12	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	CBO	B	601	-	1/1/15/16	5/15/109/109	0/5/5/5
4	PEG	A	702	-	-	0/4/4/4	-
5	CBO	B	602	-	1/1/15/16	6/15/109/109	0/5/5/5

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	602	CBO	C9-C11	-9.34	1.41	1.52
5	B	601	CBO	C10-C9	8.81	1.69	1.56
5	B	601	CBO	C20-C21	8.16	1.64	1.54
5	B	601	CBO	C14-C13	-7.81	1.40	1.53
5	B	602	CBO	C14-C13	-7.70	1.40	1.53

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	601	CBO	C8-C9-C10	-7.59	112.56	118.10
5	B	602	CBO	C17-C18-C13	-6.92	106.07	112.70
5	B	601	CBO	C17-C18-C13	-6.33	106.63	112.70
5	B	601	CBO	C14-C8-C9	5.80	112.62	107.93
5	B	601	CBO	C9-C11-C12	5.45	122.24	116.81

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	601	CBO	C5
5	B	602	CBO	C5

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	602	CBO	C2-C3-O3-C29
5	B	602	CBO	C30-C29-O3-C3
5	B	602	CBO	O29-C29-O3-C3
5	B	602	CBO	C29-C30-C31-C32
5	B	601	CBO	C30-C29-O3-C3

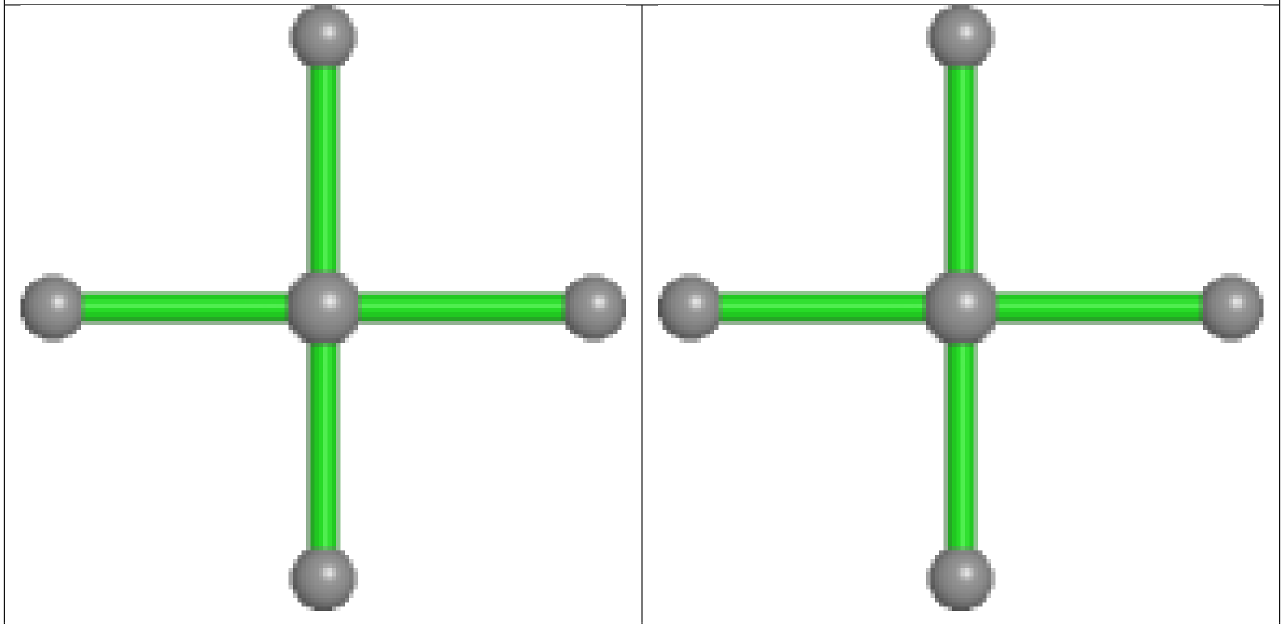
There are no ring outliers.

2 monomers are involved in 17 short contacts:

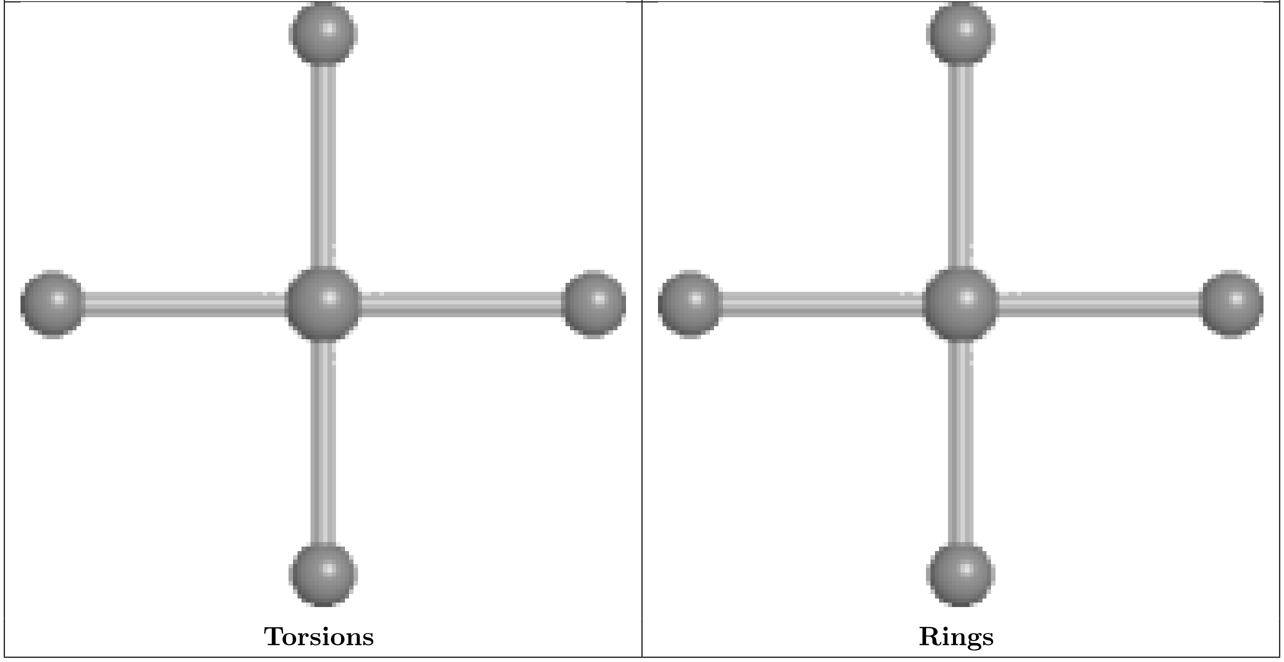
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	602	CBO	10	0
5	B	601	CBO	7	0

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 PO4 B 603

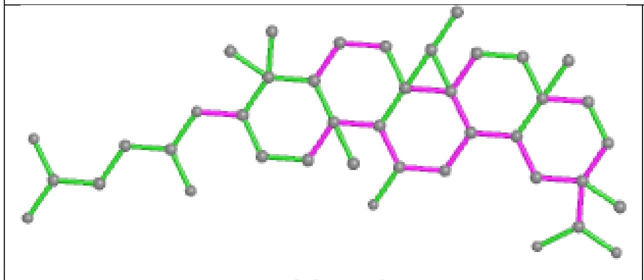


Bond lengths **Bond angles**

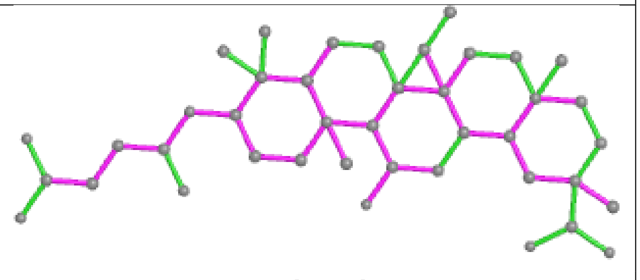


Torsions **Rings**

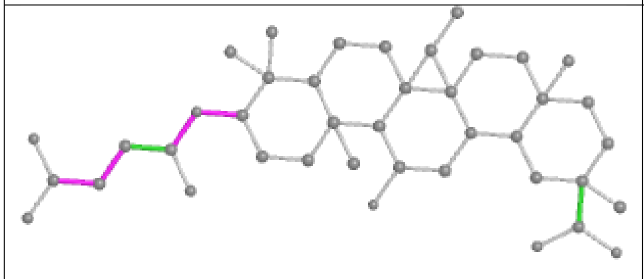
Ligand CBO B 602



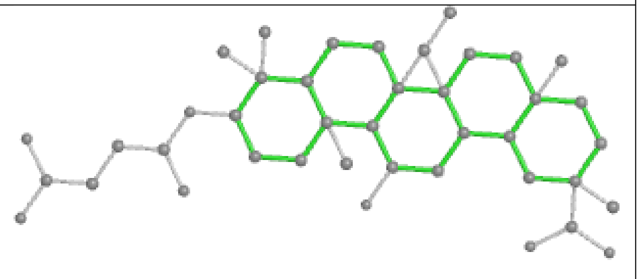
Bond lengths



Bond angles

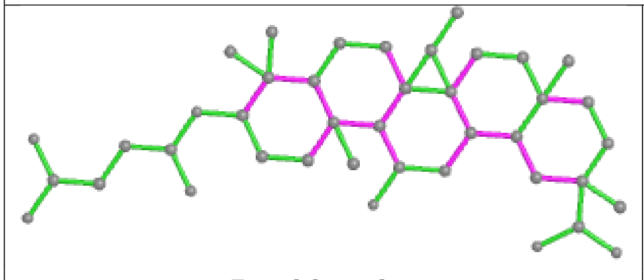


Torsions

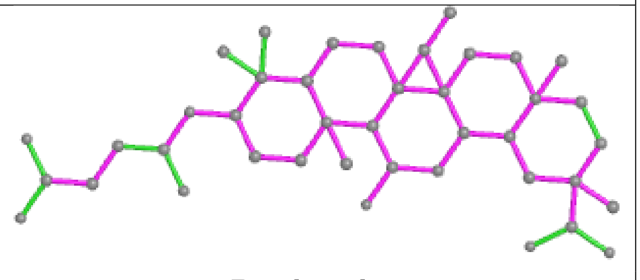


Rings

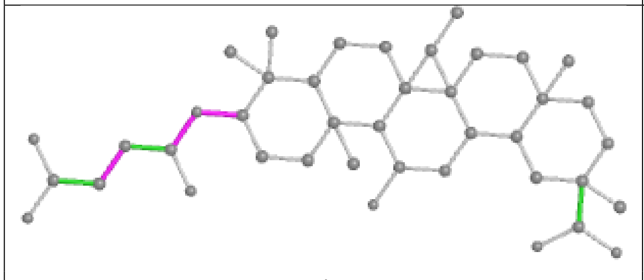
Ligand CBO B 601



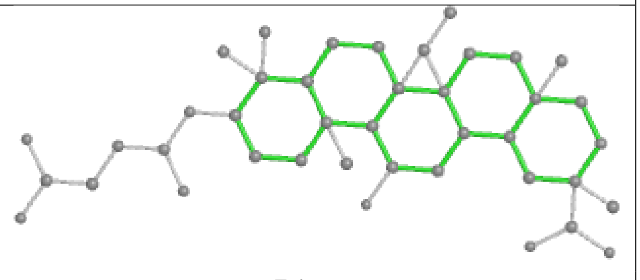
Bond lengths



Bond angles

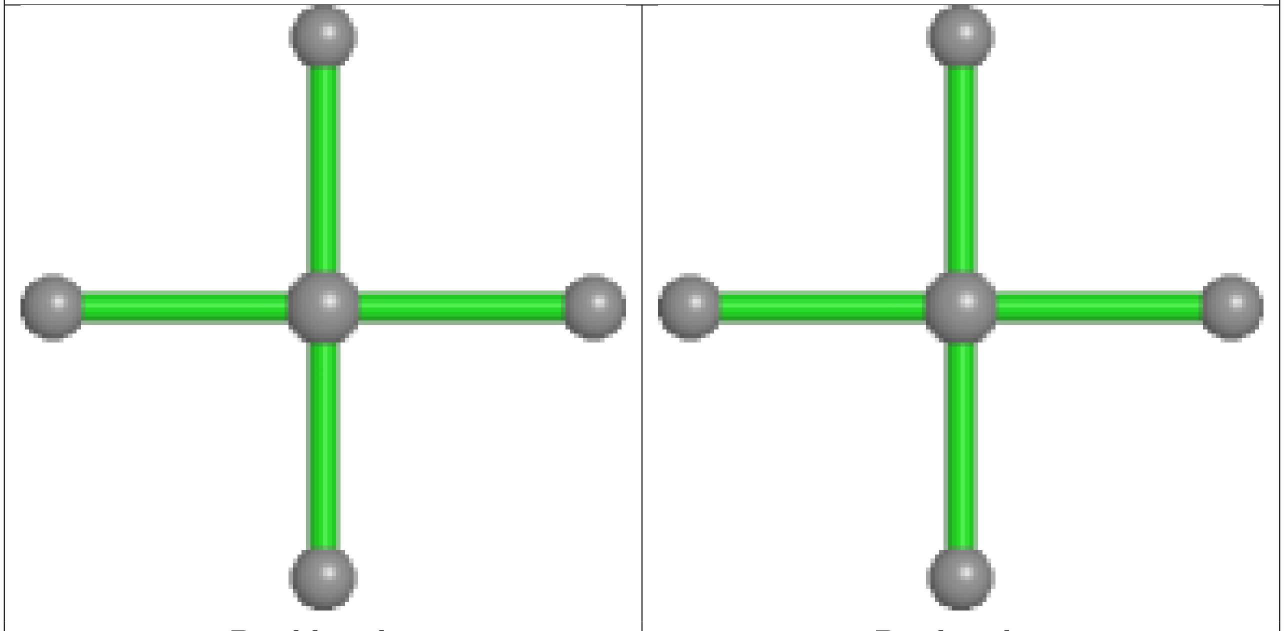


Torsions



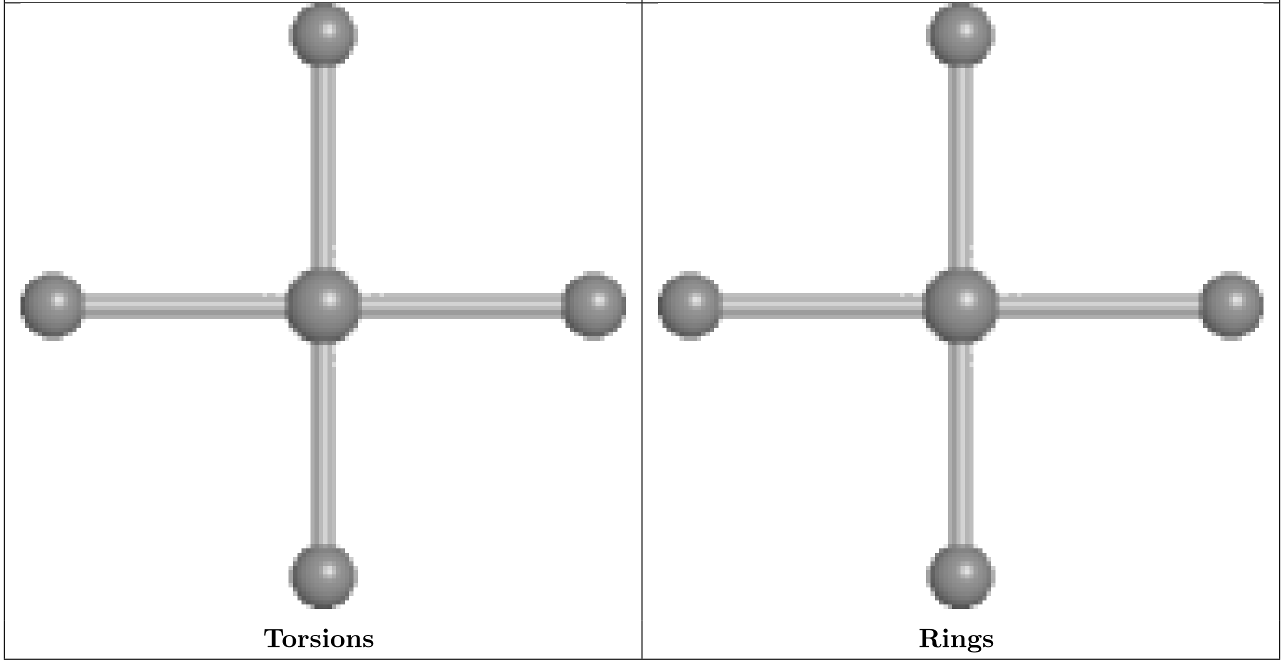
Rings

Ligand PO4 A 701



Bond lengths

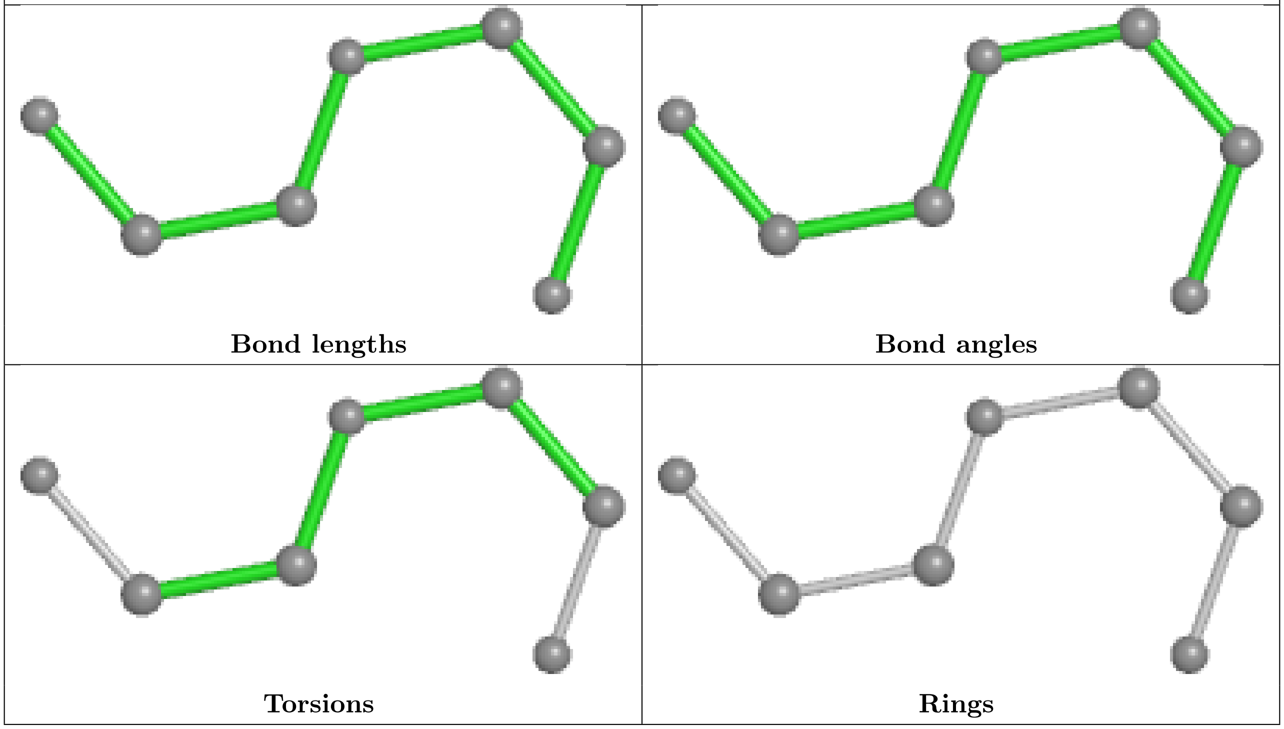
Bond angles



Torsions

Rings

Ligand PEG A 702



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, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	463/483 (95%)	0.08	21 (4%)	33	39	25, 60, 106, 174	0
2	B	462/483 (95%)	0.78	74 (16%)	1	2	38, 77, 120, 183	0
All	All	925/966 (95%)	0.43	95 (10%)	6	7	25, 68, 115, 183	0

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	489	VAL	7.0
1	A	299	VAL	6.9
1	A	419	ILE	6.5
2	B	396	ALA	6.4
2	B	472	LEU	6.2

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 ⓘ

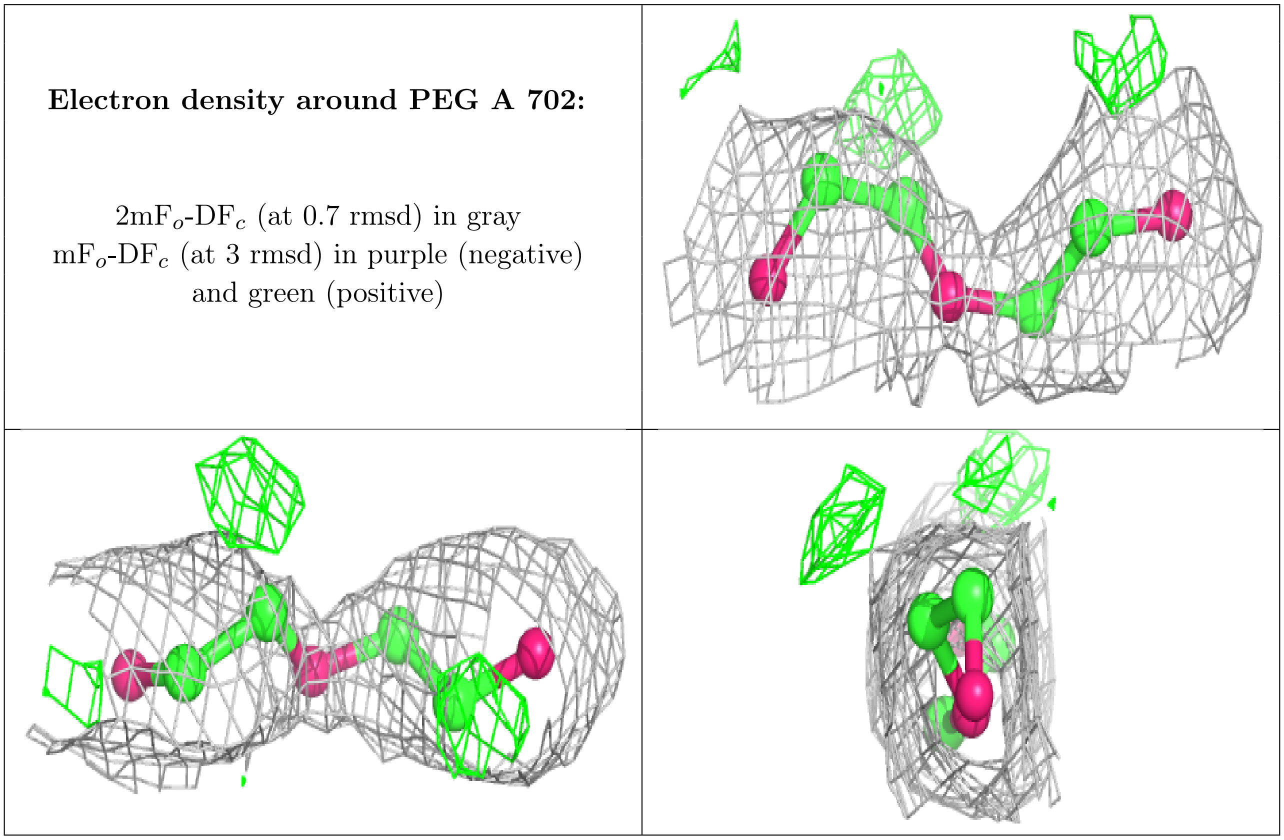
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	PEG	A	702	7/7	0.68	0.26	75,90,98,100	0
5	CBO	B	602	41/41	0.73	0.41	120,140,149,153	0
5	CBO	B	601	41/41	0.79	0.30	80,140,156,162	0
3	PO4	B	603	5/5	0.97	0.48	63,74,81,90	0
3	PO4	A	701	5/5	0.98	0.25	49,50,64,68	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.

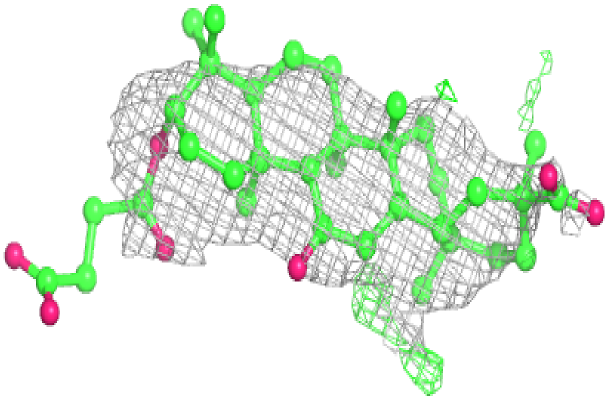
Electron density around PEG A 702:

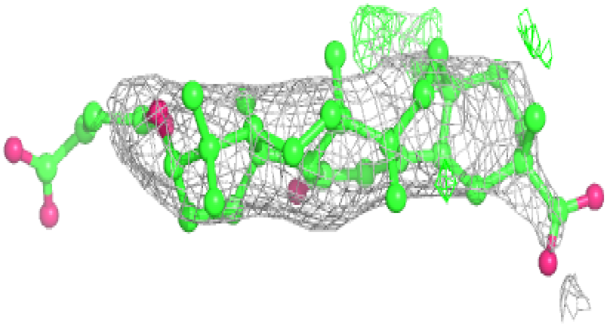
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

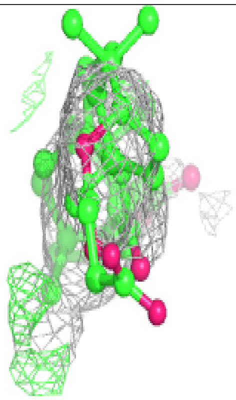


Electron density around CBO B 602:

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

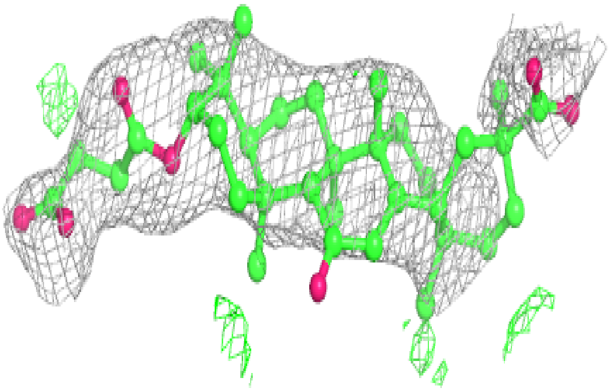


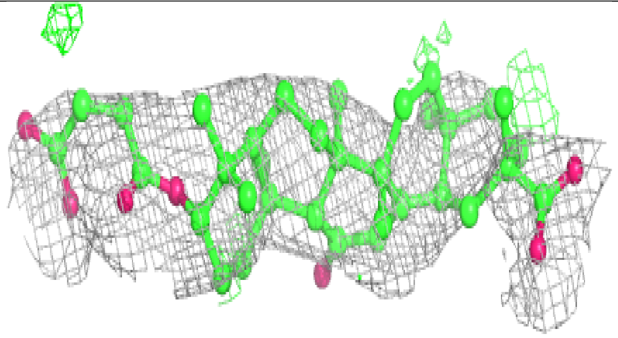


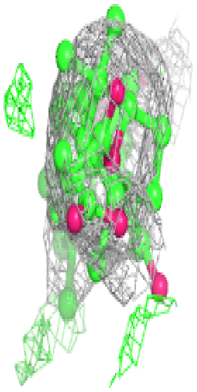


Electron density around CBO B 601:

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

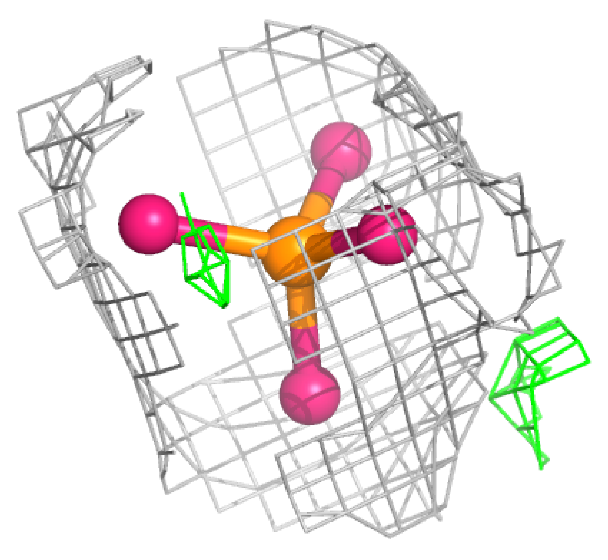


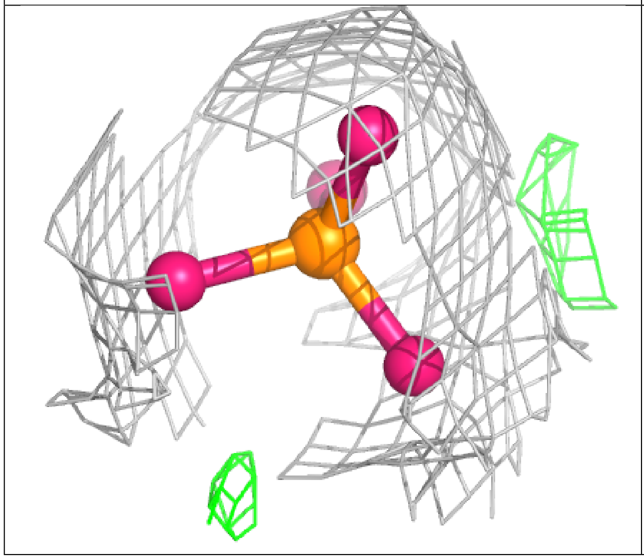


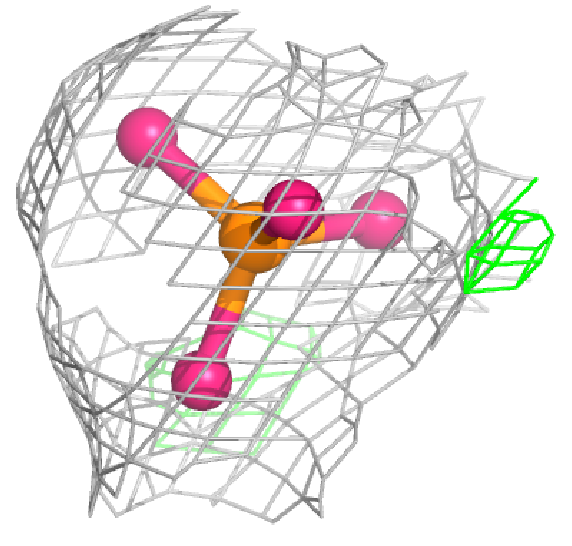


Electron density around PO4 B 603:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

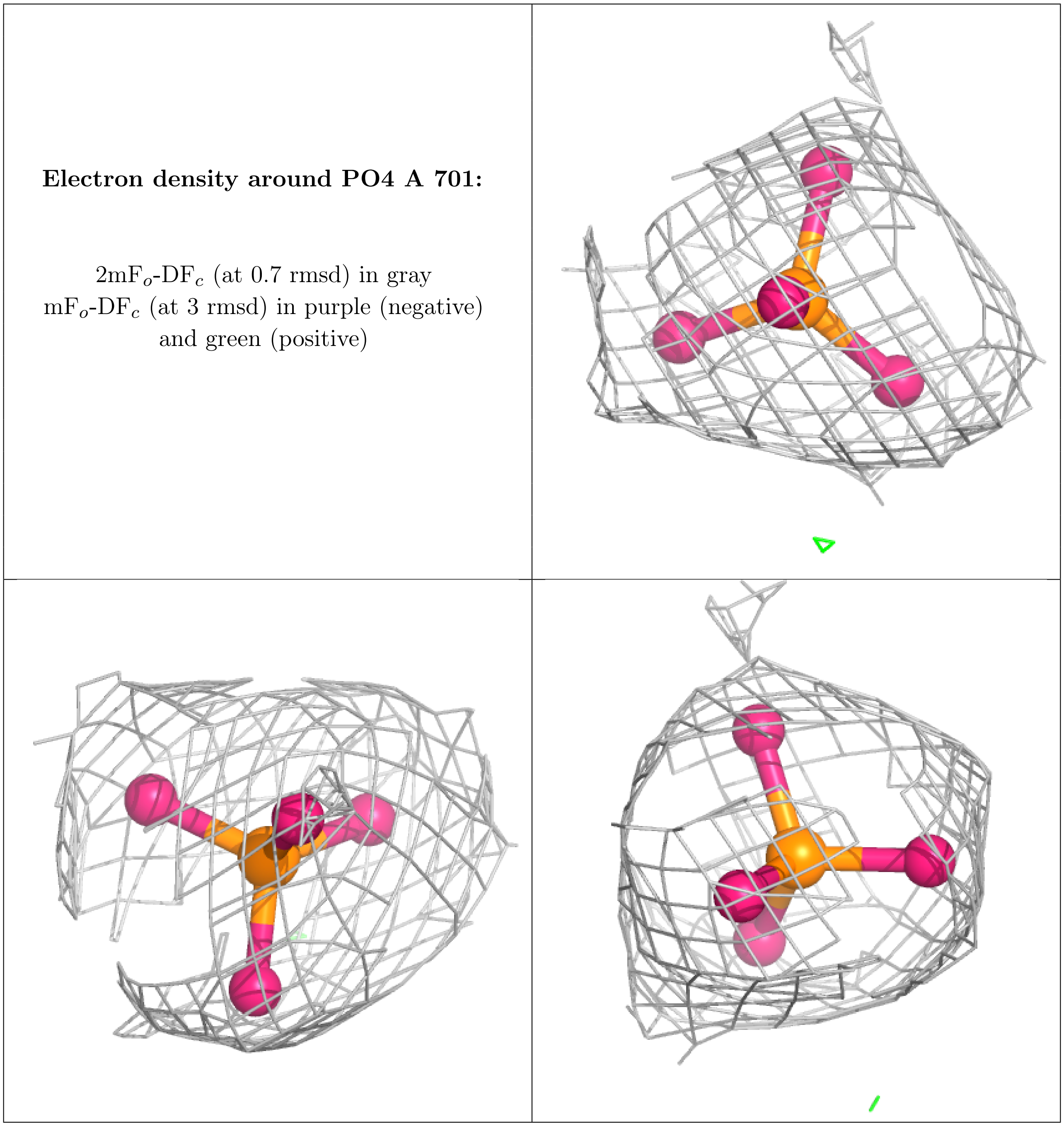






Electron density around PO4 A 701:

$2mF_o$-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.