wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2025 – 02:11 PM EST

PDB ID	:	9CQ5
Title	:	Mn-bound RuBisCO from spinach with CABP inhibitor
Authors	:	Voland, R.W.; Lancaster, K.M.
Deposited on	:	2024-07-19
Resolution	:	2.50 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

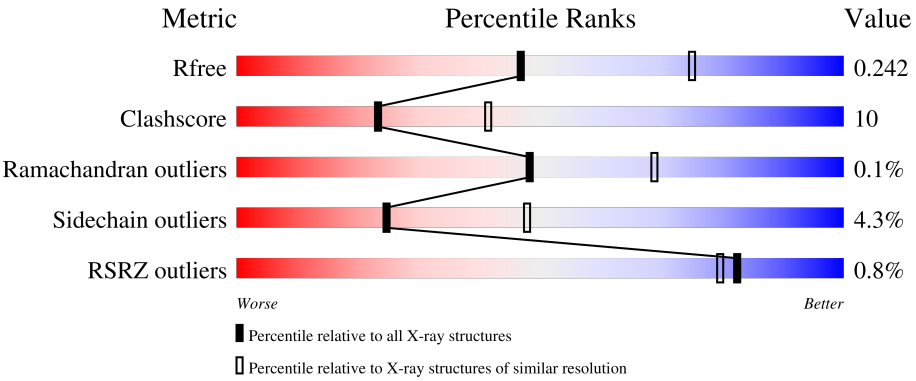
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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.



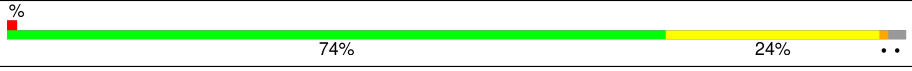
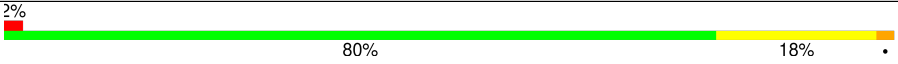
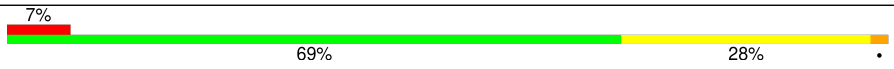
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	5504 (2.50-2.50)
Clashscore	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (2.50-2.50)

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	475	% 77% 20%
1	B	475	% 72% 25%
1	C	475	71% 27%
1	D	475	76% 22%
1	E	475	% 80% 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	475	
1	G	475	
1	H	475	
2	I	123	
2	J	123	
2	K	123	
2	L	123	
2	M	123	
2	N	123	
2	O	123	
2	P	123	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 38560 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
1	A	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	B	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	C	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	D	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	E	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	F	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	G	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			
1	H	467	Total	C	N	O	S	0	0	0
			3657	2319	640	680	18			

- Molecule 2 is a protein called Ribulose bisphosphate carboxylase small subunit, chloroplastic 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	J	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	K	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	L	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	M	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	N	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	O	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			
2	P	123	Total	C	N	O	S	0	0	0
			1033	673	167	186	7			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	2	GLN	LYS	conflict	UNP Q43832
I	6	ILE	THR	conflict	UNP Q43832
I	7	LEU	GLN	conflict	UNP Q43832
I	9	LEU	MET	conflict	UNP Q43832
I	11	LYS	ARG	conflict	UNP Q43832
I	109	GLU	GLN	conflict	UNP Q43832
I	113	ILE	VAL	conflict	UNP Q43832
J	2	GLN	LYS	conflict	UNP Q43832
J	6	ILE	THR	conflict	UNP Q43832
J	7	LEU	GLN	conflict	UNP Q43832
J	9	LEU	MET	conflict	UNP Q43832
J	11	LYS	ARG	conflict	UNP Q43832
J	109	GLU	GLN	conflict	UNP Q43832
J	113	ILE	VAL	conflict	UNP Q43832
K	2	GLN	LYS	conflict	UNP Q43832
K	6	ILE	THR	conflict	UNP Q43832
K	7	LEU	GLN	conflict	UNP Q43832
K	9	LEU	MET	conflict	UNP Q43832
K	11	LYS	ARG	conflict	UNP Q43832
K	109	GLU	GLN	conflict	UNP Q43832
K	113	ILE	VAL	conflict	UNP Q43832
L	2	GLN	LYS	conflict	UNP Q43832
L	6	ILE	THR	conflict	UNP Q43832
L	7	LEU	GLN	conflict	UNP Q43832
L	9	LEU	MET	conflict	UNP Q43832
L	11	LYS	ARG	conflict	UNP Q43832
L	109	GLU	GLN	conflict	UNP Q43832
L	113	ILE	VAL	conflict	UNP Q43832
M	2	GLN	LYS	conflict	UNP Q43832
M	6	ILE	THR	conflict	UNP Q43832
M	7	LEU	GLN	conflict	UNP Q43832
M	9	LEU	MET	conflict	UNP Q43832
M	11	LYS	ARG	conflict	UNP Q43832
M	109	GLU	GLN	conflict	UNP Q43832

Continued on next page...

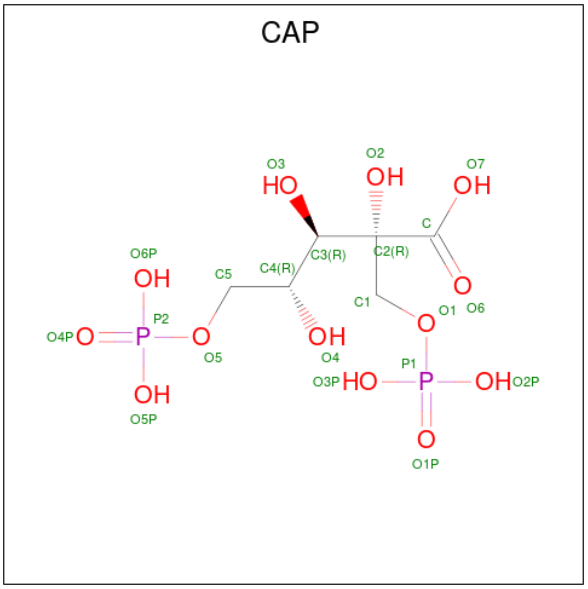
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	113	ILE	VAL	conflict	UNP Q43832
N	2	GLN	LYS	conflict	UNP Q43832
N	6	ILE	THR	conflict	UNP Q43832
N	7	LEU	GLN	conflict	UNP Q43832
N	9	LEU	MET	conflict	UNP Q43832
N	11	LYS	ARG	conflict	UNP Q43832
N	109	GLU	GLN	conflict	UNP Q43832
N	113	ILE	VAL	conflict	UNP Q43832
O	2	GLN	LYS	conflict	UNP Q43832
O	6	ILE	THR	conflict	UNP Q43832
O	7	LEU	GLN	conflict	UNP Q43832
O	9	LEU	MET	conflict	UNP Q43832
O	11	LYS	ARG	conflict	UNP Q43832
O	109	GLU	GLN	conflict	UNP Q43832
O	113	ILE	VAL	conflict	UNP Q43832
P	2	GLN	LYS	conflict	UNP Q43832
P	6	ILE	THR	conflict	UNP Q43832
P	7	LEU	GLN	conflict	UNP Q43832
P	9	LEU	MET	conflict	UNP Q43832
P	11	LYS	ARG	conflict	UNP Q43832
P	109	GLU	GLN	conflict	UNP Q43832
P	113	ILE	VAL	conflict	UNP Q43832

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mn 1 1	0	0
3	B	1	Total Mn 1 1	0	0
3	C	1	Total Mn 1 1	0	0
3	D	1	Total Mn 1 1	0	0
3	E	1	Total Mn 1 1	0	0
3	F	1	Total Mn 1 1	0	0
3	G	1	Total Mn 1 1	0	0
3	H	1	Total Mn 1 1	0	0

- Molecule 4 is 2-CARBOXYARABINITOL-1,5-DIPHOSPHATE (three-letter code: CAP) (formula: $C_6H_{14}O_{13}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O P 21 6 13 2	0	0
4	B	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	D	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	F	1	Total C O P 21 6 13 2	0	0
4	G	1	Total C O P 21 6 13 2	0	0
4	H	1	Total C O P 21 6 13 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	74	Total O 74 74	0	0
5	B	73	Total O 73 73	0	0

Continued on next page...

Continued from previous page...

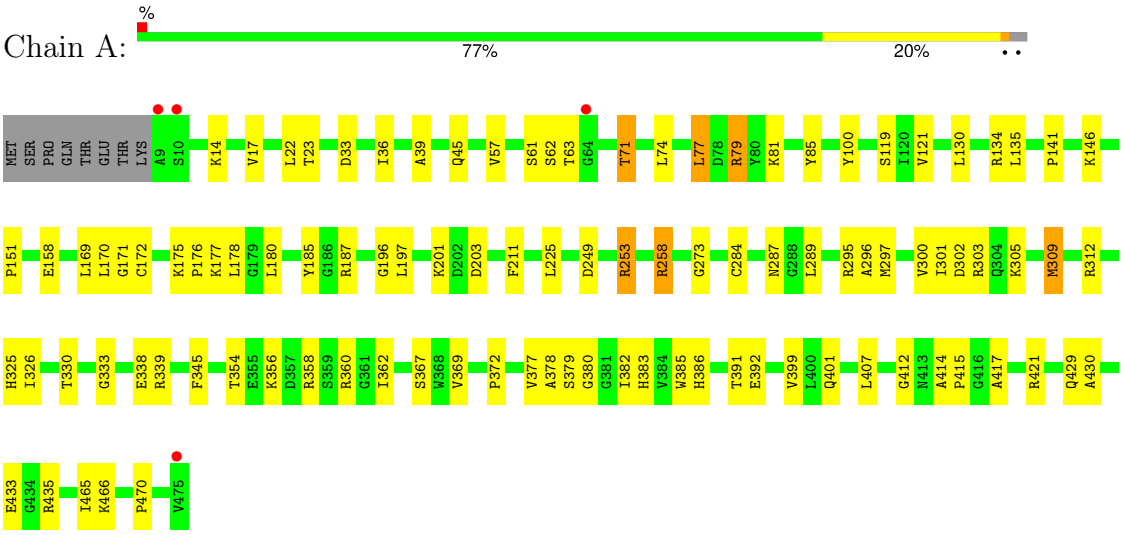
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	105	Total O 105 105	0	0
5	D	87	Total O 87 87	0	0
5	E	103	Total O 103 103	0	0
5	F	62	Total O 62 62	0	0
5	G	96	Total O 96 96	0	0
5	H	72	Total O 72 72	0	0
5	I	36	Total O 36 36	0	0
5	J	14	Total O 14 14	0	0
5	K	31	Total O 31 31	0	0
5	L	23	Total O 23 23	0	0
5	M	32	Total O 32 32	0	0
5	N	15	Total O 15 15	0	0
5	O	33	Total O 33 33	0	0
5	P	8	Total O 8 8	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.

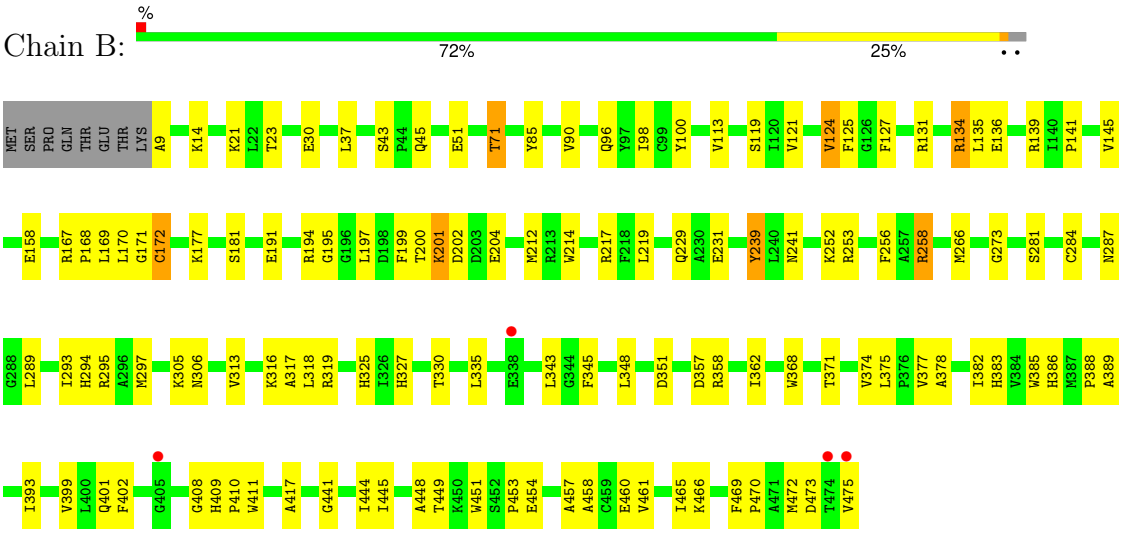
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain A:



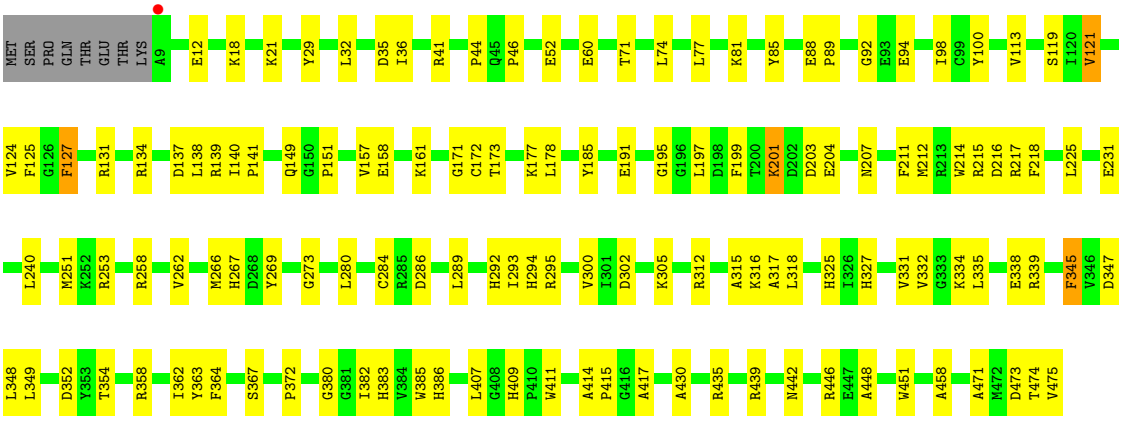
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain B:



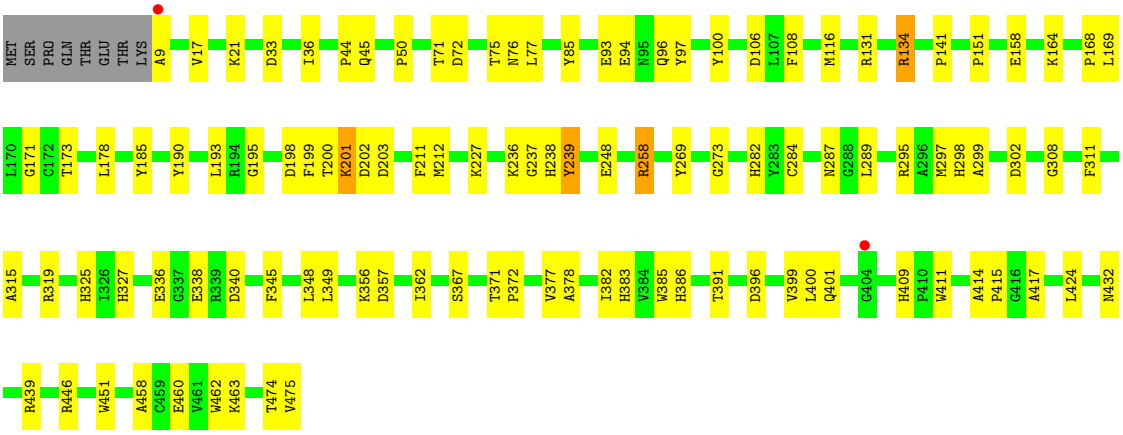
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain C:



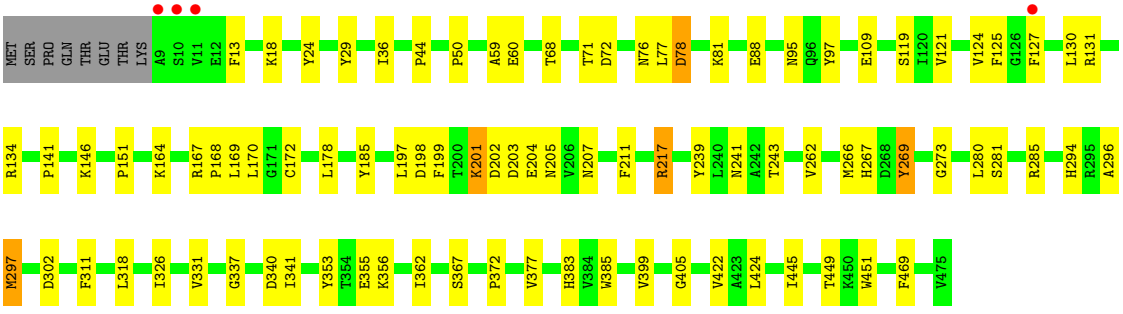
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain D:



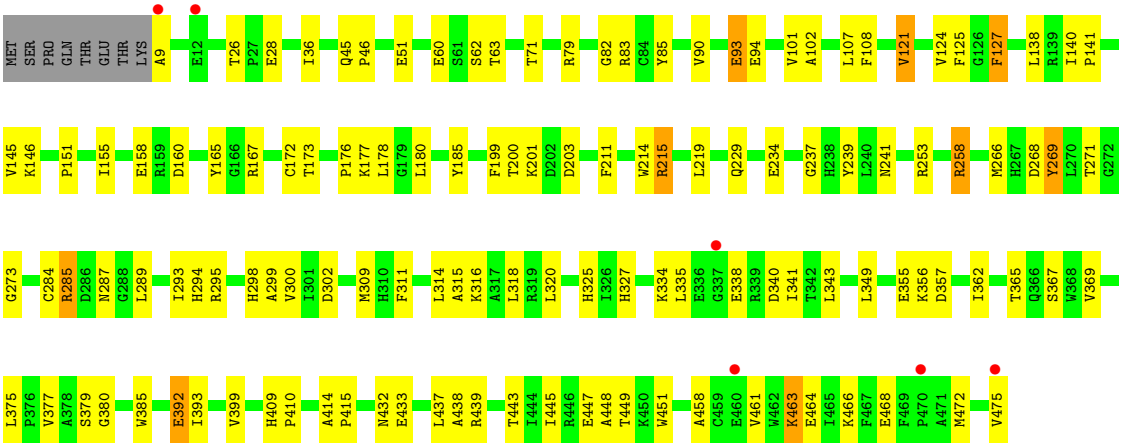
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain E:



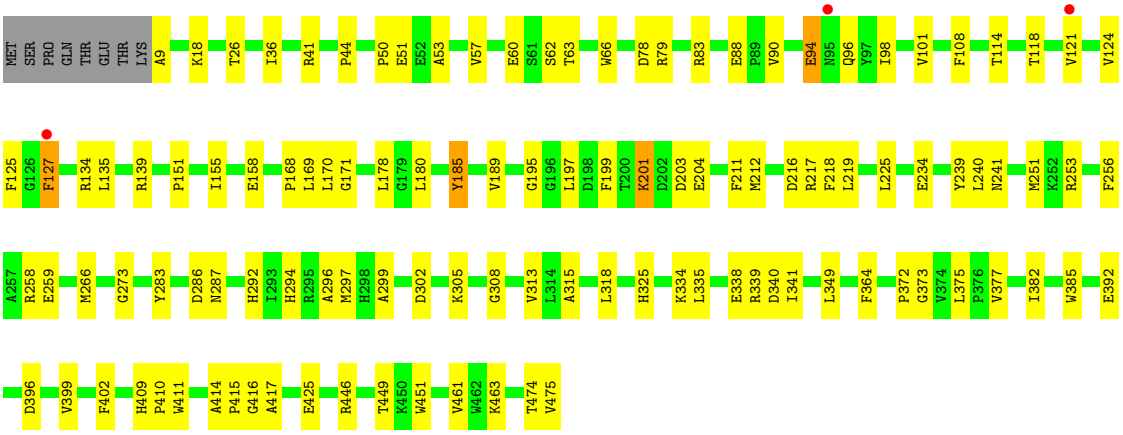
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain F:



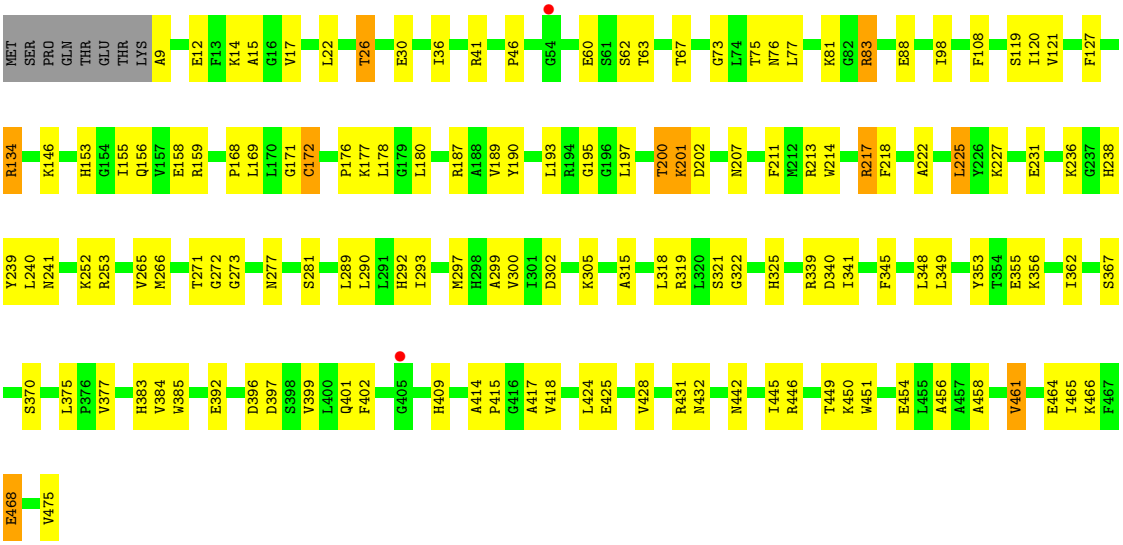
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain G:



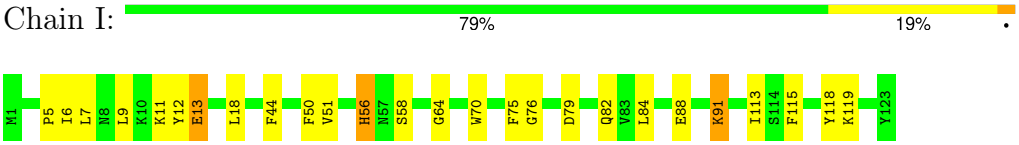
• Molecule 1: Ribulose bisphosphate carboxylase large chain

Chain H:



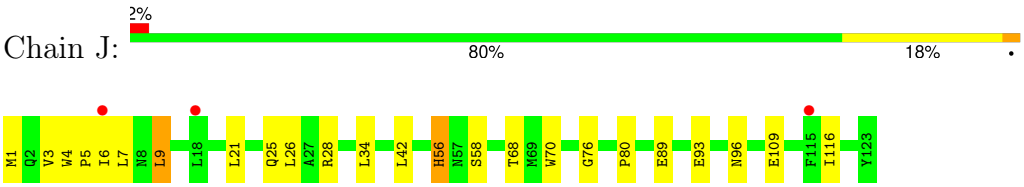
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain I:



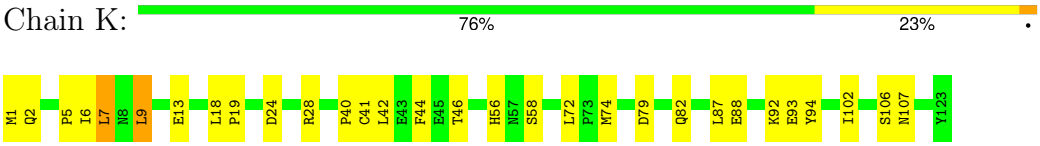
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain J:



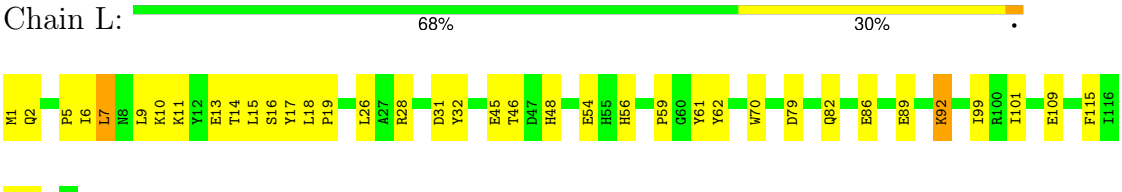
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain K:



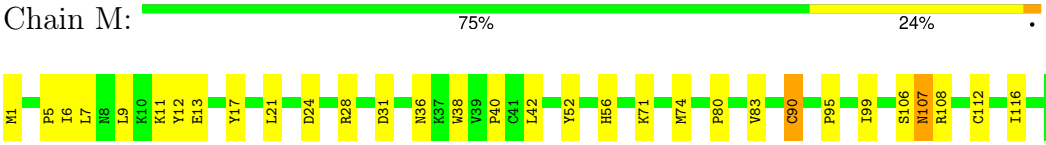
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain L:



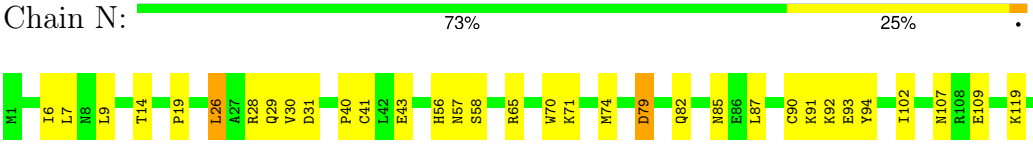
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain M:



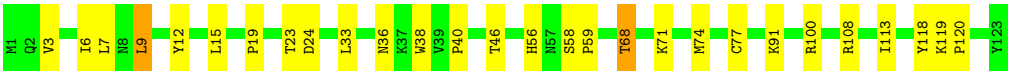
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain N:



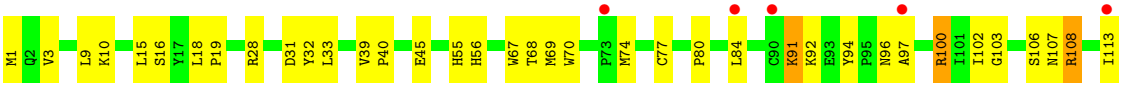
• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

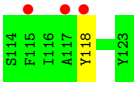
Chain O:



• Molecule 2: Ribulose bisphosphate carboxylase small subunit, chloroplastic 2

Chain P:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	218.03Å 218.89Å 111.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	99.51 – 2.50 99.51 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (99.51-2.50) 99.9 (99.51-2.50)	Depositor EDS
R_{merge}	0.56	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.49 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.197 , 0.242 0.198 , 0.242	Depositor DCC
R_{free} test set	176016 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	45.4	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 42.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.025 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	38560	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.33% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CAP, KCX, MN

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	0.43	0/3734	0.64	1/5064 (0.0%)
1	B	0.40	0/3734	0.62	0/5064
1	C	0.43	0/3734	0.64	0/5064
1	D	0.41	0/3734	0.63	0/5064
1	E	0.42	0/3734	0.65	0/5064
1	F	0.41	0/3734	0.65	0/5064
1	G	0.44	0/3734	0.64	0/5064
1	H	0.43	0/3734	0.64	0/5064
2	I	0.47	0/1068	0.60	0/1453
2	J	0.40	0/1068	0.60	1/1453 (0.1%)
2	K	0.41	0/1068	0.60	0/1453
2	L	0.40	0/1068	0.58	0/1453
2	M	0.44	0/1068	0.61	0/1453
2	N	0.39	0/1068	0.56	0/1453
2	O	0.46	0/1068	0.64	0/1453
2	P	0.37	0/1068	0.60	0/1453
All	All	0.42	0/38416	0.63	2/52136 (0.0%)

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	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
1	H	0	1
2	P	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	7

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	9	LEU	CB-CG-CD2	-7.01	99.09	111.00
1	A	225	LEU	CA-CB-CG	5.48	127.90	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	253	ARG	Sidechain
1	C	217	ARG	Sidechain
1	D	134	ARG	Sidechain
1	E	217	ARG	Sidechain
1	F	285	ARG	Sidechain

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	3657	0	3565	67	0
1	B	3657	0	3565	92	0
1	C	3657	0	3564	90	0
1	D	3657	0	3564	72	1
1	E	3657	0	3564	52	0
1	F	3657	0	3564	87	0
1	G	3657	0	3564	81	0
1	H	3657	0	3564	95	1
2	I	1033	0	990	25	0
2	J	1033	0	990	24	0
2	K	1033	0	990	21	0
2	L	1033	0	990	33	0
2	M	1033	0	990	22	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	N	1033	0	990	24	0
2	O	1033	0	990	18	0
2	P	1033	0	990	26	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
4	A	21	0	8	0	0
4	B	21	0	9	0	0
4	C	21	0	7	0	0
4	D	21	0	8	0	0
4	E	21	0	9	1	0
4	F	21	0	7	2	0
4	G	21	0	7	1	0
4	H	21	0	7	1	0
5	A	74	0	0	10	0
5	B	73	0	0	14	0
5	C	105	0	0	25	0
5	D	87	0	0	13	0
5	E	103	0	0	10	0
5	F	62	0	0	8	0
5	G	96	0	0	10	0
5	H	72	0	0	12	0
5	I	36	0	0	5	0
5	J	14	0	0	5	0
5	K	31	0	0	2	0
5	L	23	0	0	2	0
5	M	32	0	0	5	0
5	N	15	0	0	3	0
5	O	33	0	0	1	0
5	P	8	0	0	3	0
All	All	38560	0	36496	751	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 10.

The worst 5 of 751 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:G:79:ARG:NH1	5:G:601:HOH:O	1.84	1.09
1:E:29:TYR:O	5:E:601:HOH:O	1.82	0.98
1:C:216:ASP:OD2	5:C:601:HOH:O	1.84	0.94
1:C:21:LYS:NZ	5:C:605:HOH:O	2.01	0.93
4:E:502:CAP:O2P	5:E:602:HOH:O	1.87	0.92

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:D:439:ARG:NH2	1:H:468:GLU:O[4_454]	1.64	0.56

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	464/475 (98%)	452 (97%)	12 (3%)	0	100	100
1	B	464/475 (98%)	452 (97%)	12 (3%)	0	100	100
1	C	464/475 (98%)	454 (98%)	10 (2%)	0	100	100
1	D	464/475 (98%)	453 (98%)	11 (2%)	0	100	100
1	E	464/475 (98%)	454 (98%)	9 (2%)	1 (0%)	44	64
1	F	464/475 (98%)	453 (98%)	11 (2%)	0	100	100
1	G	464/475 (98%)	454 (98%)	10 (2%)	0	100	100
1	H	464/475 (98%)	452 (97%)	12 (3%)	0	100	100
2	I	121/123 (98%)	117 (97%)	3 (2%)	1 (1%)	16	31
2	J	121/123 (98%)	118 (98%)	3 (2%)	0	100	100
2	K	121/123 (98%)	117 (97%)	3 (2%)	1 (1%)	16	31
2	L	121/123 (98%)	116 (96%)	5 (4%)	0	100	100
2	M	121/123 (98%)	115 (95%)	5 (4%)	1 (1%)	16	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	N	121/123 (98%)	116 (96%)	5 (4%)	0	100	100
2	O	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
2	P	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
All	All	4680/4784 (98%)	4553 (97%)	123 (3%)	4 (0%)	48	69

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	13	GLU
2	K	13	GLU
2	M	13	GLU
1	E	297	MET

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	378/386 (98%)	363 (96%)	15 (4%)	27	51
1	B	378/386 (98%)	359 (95%)	19 (5%)	20	41
1	C	378/386 (98%)	361 (96%)	17 (4%)	23	46
1	D	378/386 (98%)	364 (96%)	14 (4%)	29	55
1	E	378/386 (98%)	363 (96%)	15 (4%)	27	51
1	F	378/386 (98%)	359 (95%)	19 (5%)	20	41
1	G	378/386 (98%)	367 (97%)	11 (3%)	37	64
1	H	378/386 (98%)	363 (96%)	15 (4%)	27	51
2	I	112/112 (100%)	110 (98%)	2 (2%)	54	78
2	J	112/112 (100%)	107 (96%)	5 (4%)	23	46
2	K	112/112 (100%)	109 (97%)	3 (3%)	40	67
2	L	112/112 (100%)	107 (96%)	5 (4%)	23	46
2	M	112/112 (100%)	106 (95%)	6 (5%)	18	37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	N	112/112 (100%)	104 (93%)	8 (7%)	12 25
2	O	112/112 (100%)	106 (95%)	6 (5%)	18 37
2	P	112/112 (100%)	103 (92%)	9 (8%)	10 20
All	All	3920/3984 (98%)	3751 (96%)	169 (4%)	25 48

5 of 169 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	134	ARG
2	M	90	CYS
1	H	252	LYS
2	J	56	HIS
2	N	79	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	O	82	GLN
2	O	56	HIS
2	L	82	GLN
2	O	2	GLN
2	K	107	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

8 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	A	201	1,3	10,11,12	1.01	1 (10%)	6,12,14	0.88	0
1	KCX	F	201	1,3	10,11,12	0.97	0	6,12,14	0.93	0
1	KCX	B	201	1,3	10,11,12	0.95	0	6,12,14	1.73	2 (33%)
1	KCX	G	201	1,3	10,11,12	0.78	0	6,12,14	1.53	2 (33%)
1	KCX	H	201	1,3	10,11,12	1.09	1 (10%)	6,12,14	1.25	1 (16%)
1	KCX	D	201	1,3	10,11,12	0.88	0	6,12,14	1.33	1 (16%)
1	KCX	C	201	1,3	10,11,12	0.98	0	6,12,14	1.53	1 (16%)
1	KCX	E	201	1,3	10,11,12	0.99	0	6,12,14	1.09	1 (16%)

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	A	201	1,3	-	2/9/10/12	-
1	KCX	F	201	1,3	-	2/9/10/12	-
1	KCX	B	201	1,3	-	5/9/10/12	-
1	KCX	G	201	1,3	-	4/9/10/12	-
1	KCX	H	201	1,3	-	2/9/10/12	-
1	KCX	D	201	1,3	-	3/9/10/12	-
1	KCX	C	201	1,3	-	1/9/10/12	-
1	KCX	E	201	1,3	-	0/9/10/12	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	201	KCX	OQ1-CX	2.43	1.26	1.21
1	A	201	KCX	OQ1-CX	2.01	1.25	1.21

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	201	KCX	OQ1-CX-NZ	-3.68	119.33	124.92
1	B	201	KCX	OQ1-CX-NZ	-3.36	119.81	124.92
1	G	201	KCX	CD-CE-NZ	-2.93	103.97	112.20
1	H	201	KCX	OQ1-CX-NZ	-2.54	121.06	124.92
1	B	201	KCX	CE-NZ-CX	-2.39	117.92	121.98

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	201	KCX	N-CA-CB-CG
1	B	201	KCX	C-CA-CB-CG
1	D	201	KCX	C-CA-CB-CG
1	G	201	KCX	N-CA-CB-CG
1	G	201	KCX	C-CA-CB-CG

There are no ring outliers.

7 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	F	201	KCX	2	0
1	B	201	KCX	2	0
1	G	201	KCX	1	0
1	H	201	KCX	1	0
1	D	201	KCX	2	0
1	C	201	KCX	1	0
1	E	201	KCX	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	$\#\|Z\| > 2$
4	CAP	G	502	3	18,20,20	0.87	0	23,31,31	1.24	2 (8%)
4	CAP	F	502	3	18,20,20	0.98	0	23,31,31	1.06	2 (8%)
4	CAP	E	502	3	18,20,20	1.19	2 (11%)	23,31,31	1.34	3 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	CAP	H	502	3	18,20,20	1.07	1 (5%)	23,31,31	1.28	4 (17%)
4	CAP	D	502	3	18,20,20	0.96	0	23,31,31	1.43	2 (8%)
4	CAP	A	502	3	18,20,20	1.05	0	23,31,31	1.66	5 (21%)
4	CAP	C	502	3	18,20,20	0.99	0	23,31,31	1.56	3 (13%)
4	CAP	B	502	3	18,20,20	1.12	1 (5%)	23,31,31	1.29	3 (13%)

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
4	CAP	G	502	3	-	8/29/29/29	-
4	CAP	F	502	3	-	7/29/29/29	-
4	CAP	E	502	3	-	8/29/29/29	-
4	CAP	H	502	3	-	7/29/29/29	-
4	CAP	D	502	3	-	6/29/29/29	-
4	CAP	A	502	3	-	12/29/29/29	-
4	CAP	C	502	3	-	6/29/29/29	-
4	CAP	B	502	3	-	12/29/29/29	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	502	CAP	C2-C	-2.89	1.50	1.53
4	E	502	CAP	C2-C	-2.63	1.51	1.53
4	E	502	CAP	C4-C3	-2.60	1.51	1.54
4	H	502	CAP	C4-C3	-2.07	1.52	1.54

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	502	CAP	O7-C-C2	5.18	122.74	114.06
4	C	502	CAP	O7-C-C2	4.70	121.93	114.06
4	A	502	CAP	O7-C-C2	3.81	120.44	114.06
4	E	502	CAP	O7-C-C2	3.57	120.04	114.06
4	B	502	CAP	O7-C-C2	3.56	120.03	114.06

There are no chirality outliers.

5 of 66 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	502	CAP	O6-C-C2-C1
4	A	502	CAP	O7-C-C2-C1
4	A	502	CAP	O6-C-C2-O2
4	A	502	CAP	O7-C-C2-O2
4	A	502	CAP	C2-C3-C4-O4

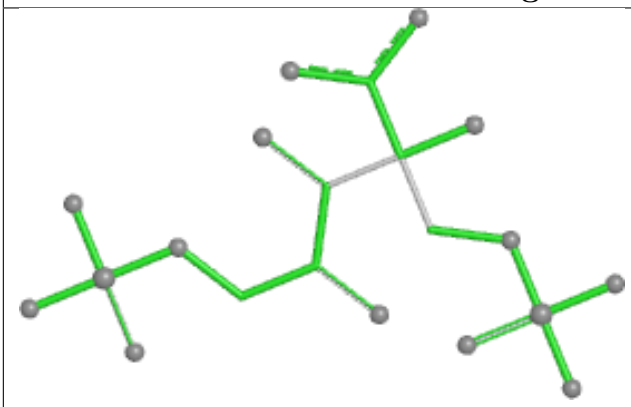
There are no ring outliers.

4 monomers are involved in 5 short contacts:

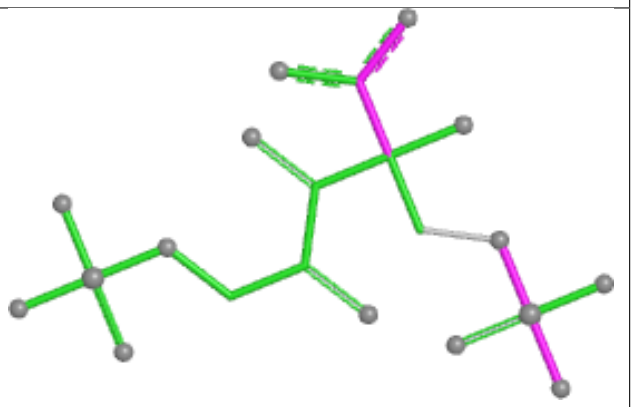
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	502	CAP	1	0
4	F	502	CAP	2	0
4	E	502	CAP	1	0
4	H	502	CAP	1	0

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 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.

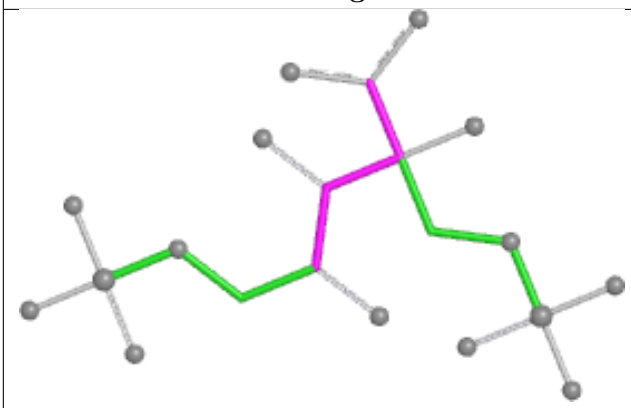
Ligand CAP G 502



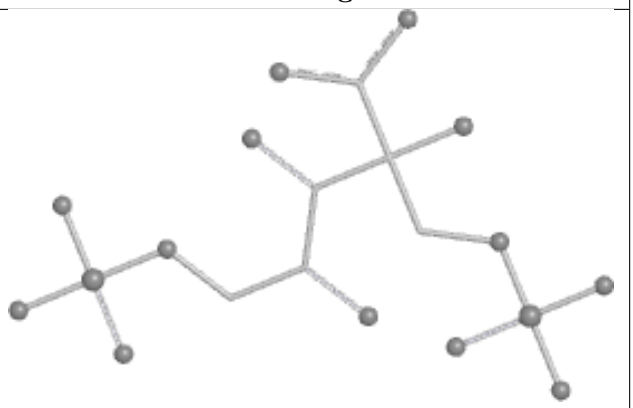
Bond lengths



Bond angles

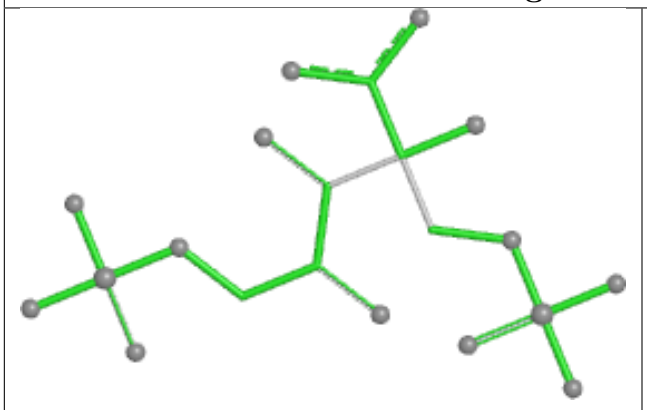


Torsions

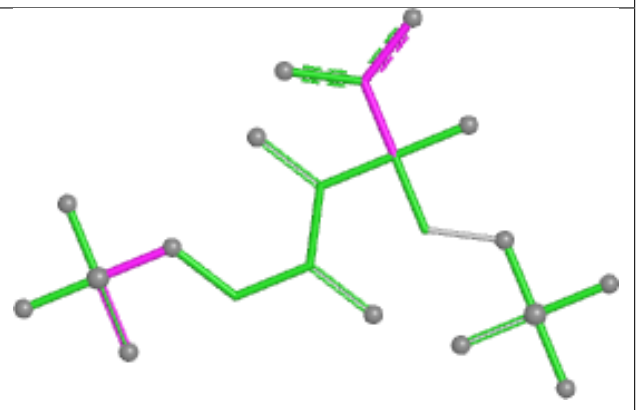


Rings

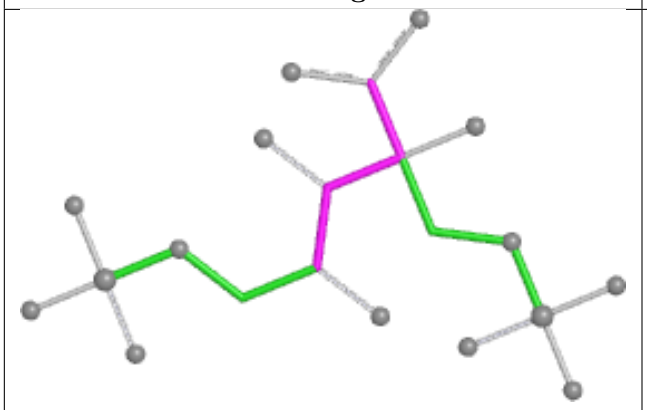
Ligand CAP F 502



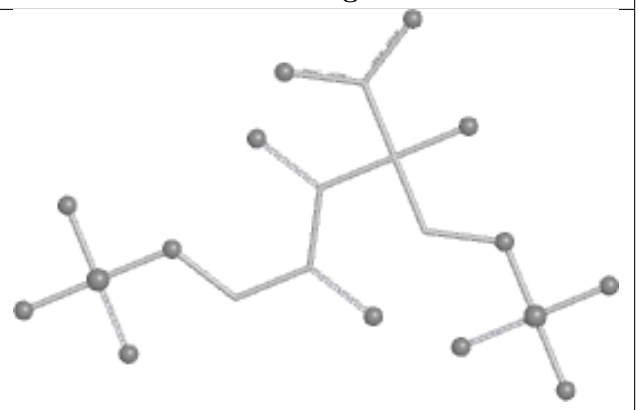
Bond lengths



Bond angles

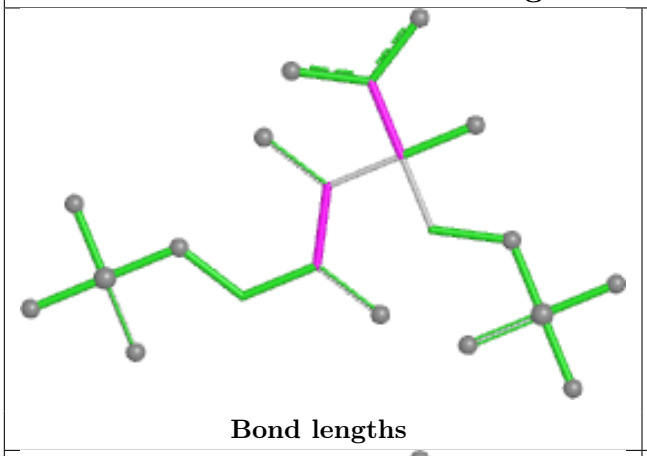


Torsions

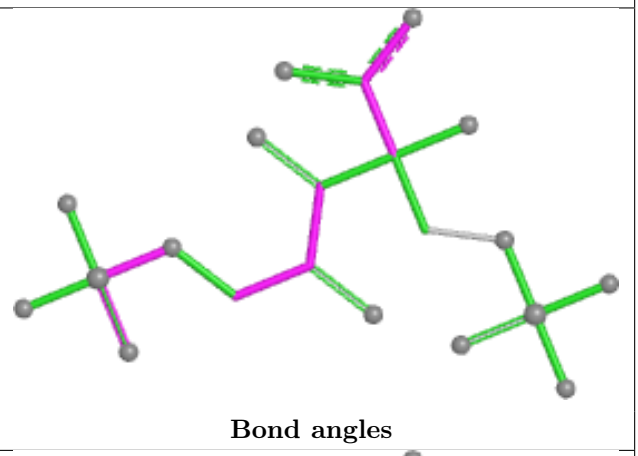


Rings

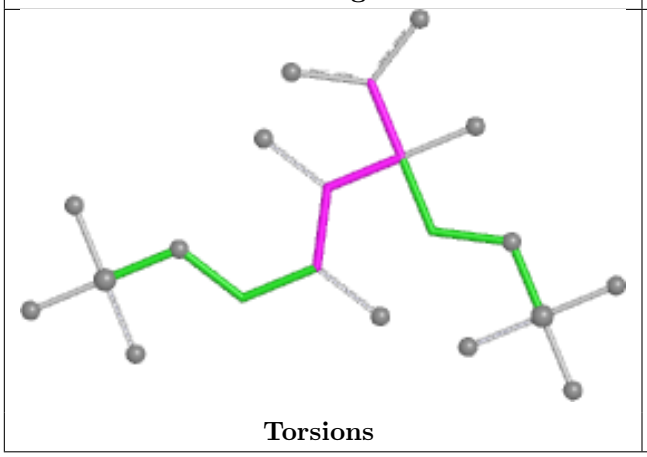
Ligand CAP E 502



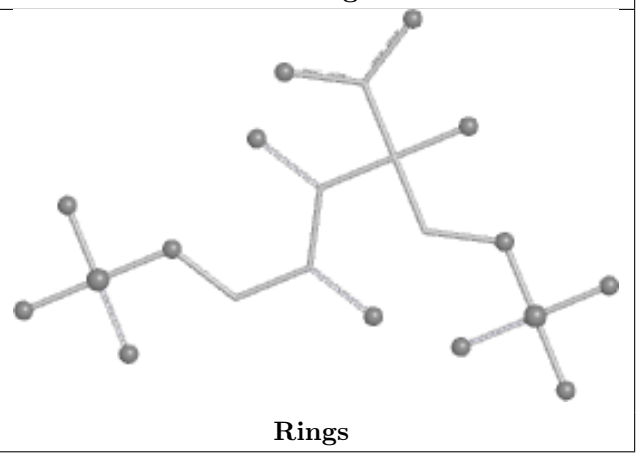
Bond lengths



Bond angles

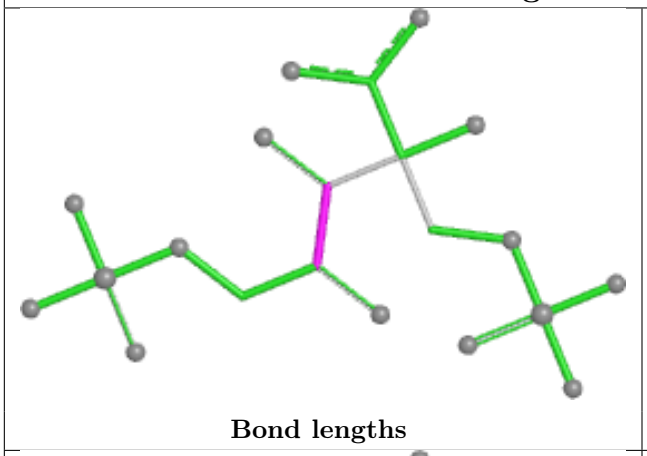


Torsions

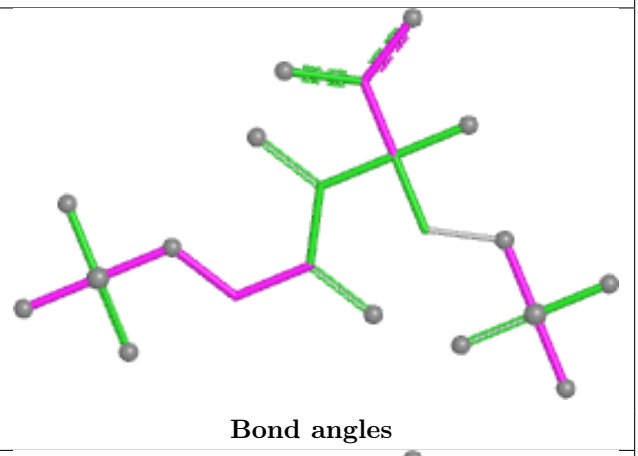


Rings

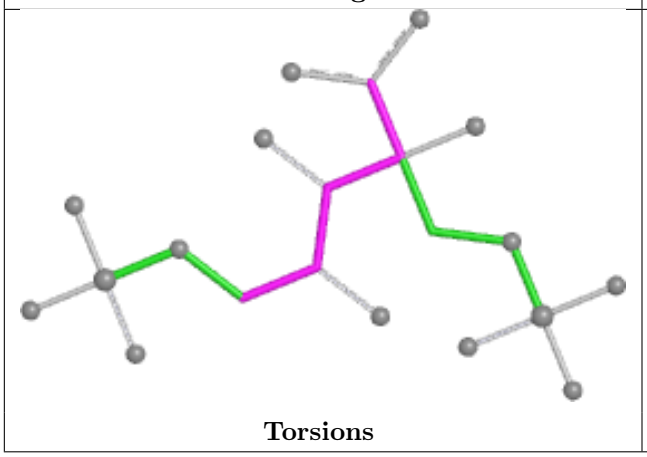
Ligand CAP H 502



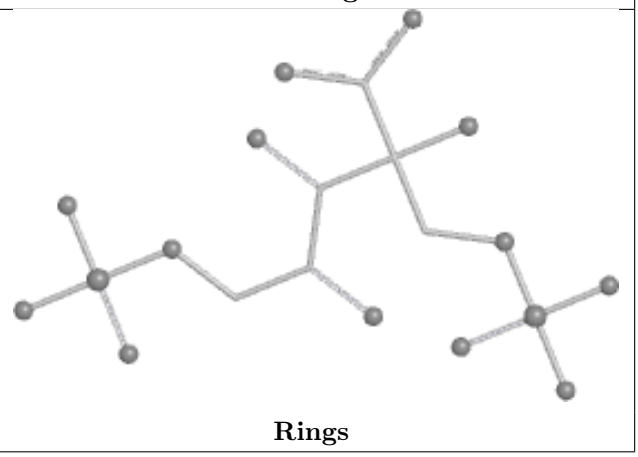
Bond lengths



Bond angles

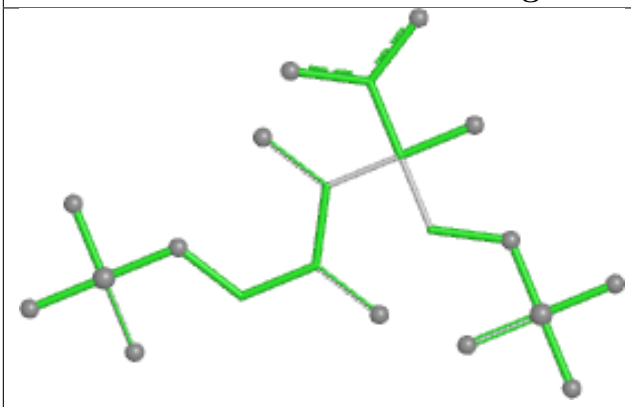


Torsions

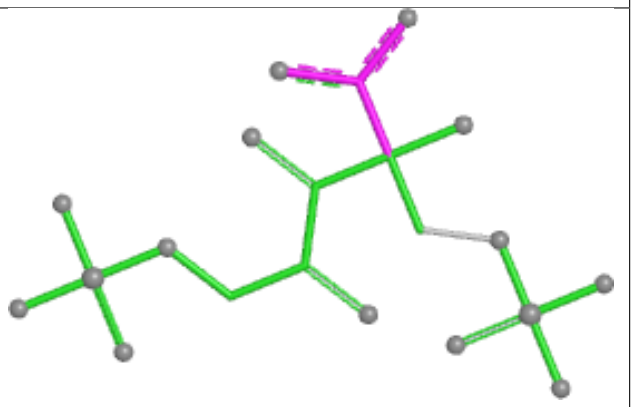


Rings

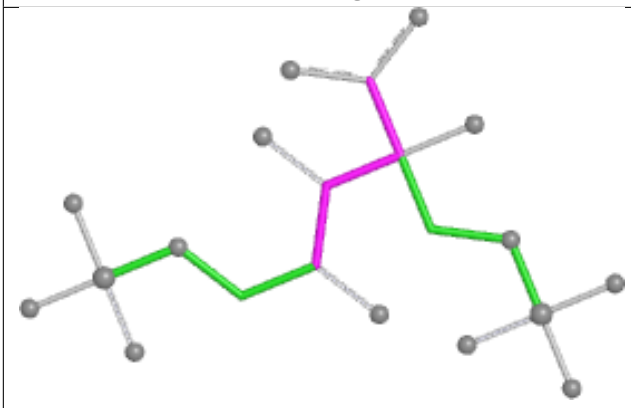
Ligand CAP D 502



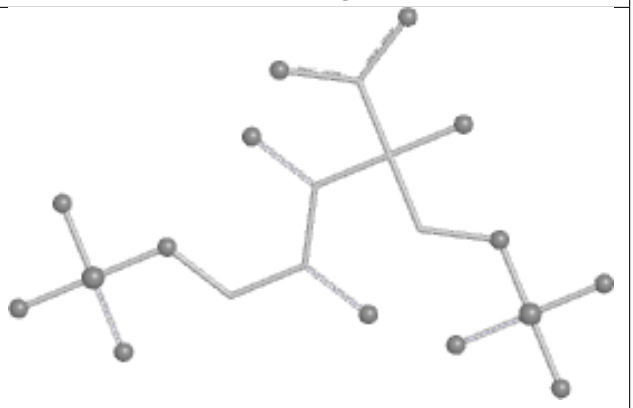
Bond lengths



Bond angles

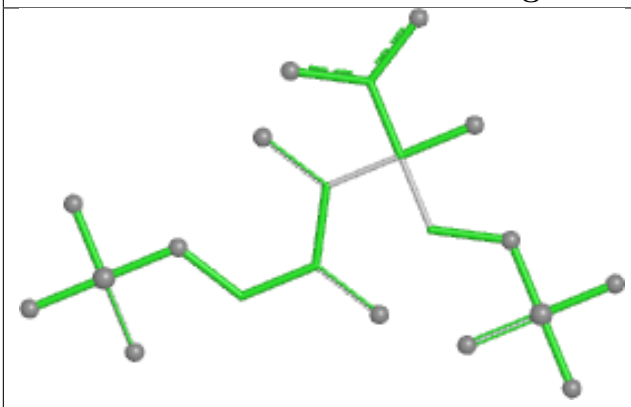


Torsions

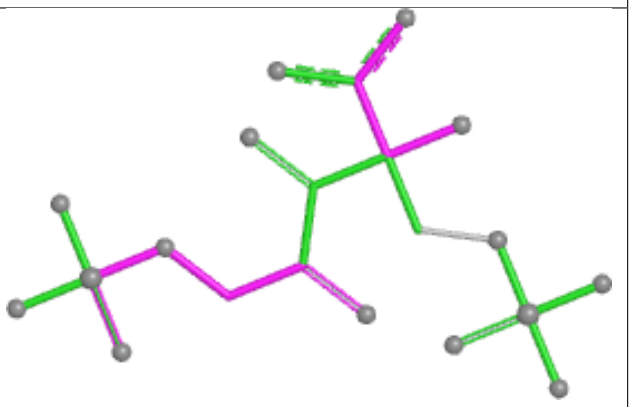


Rings

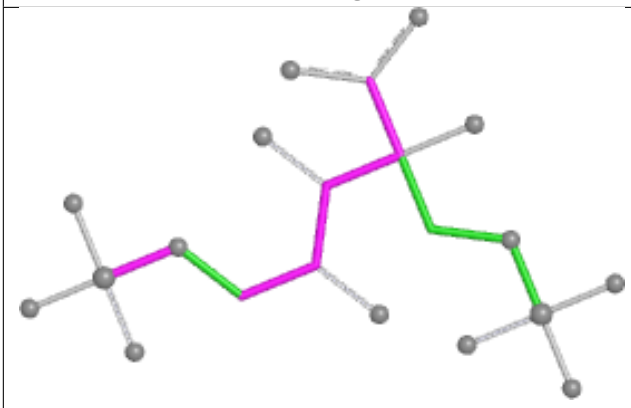
Ligand CAP A 502



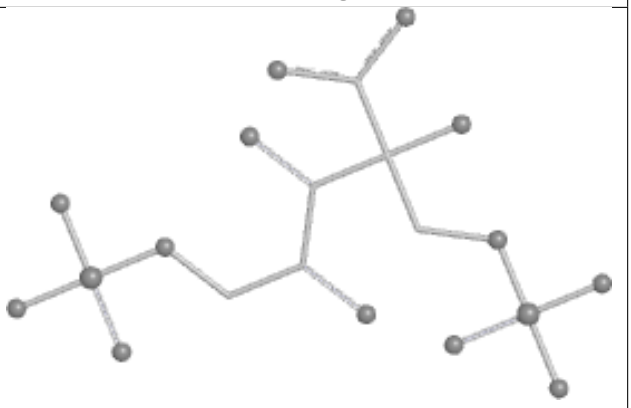
Bond lengths



Bond angles

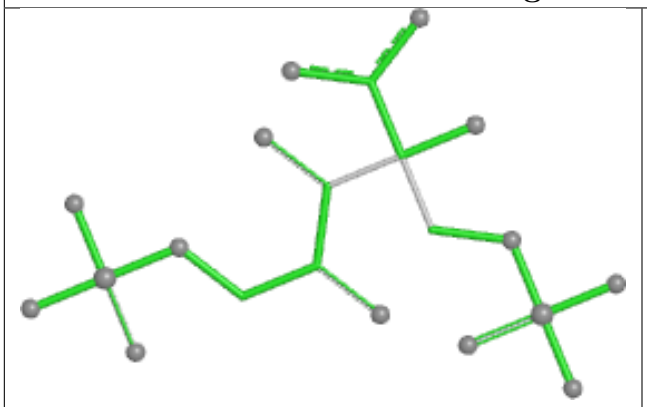


Torsions

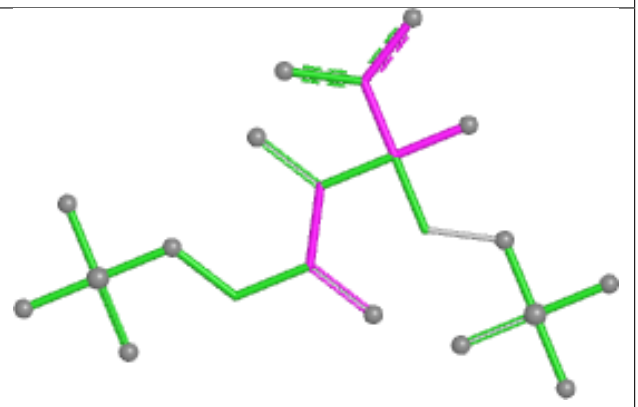


Rings

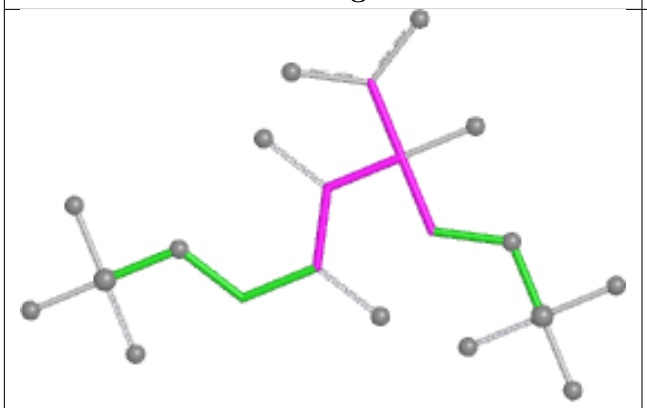
Ligand CAP C 502



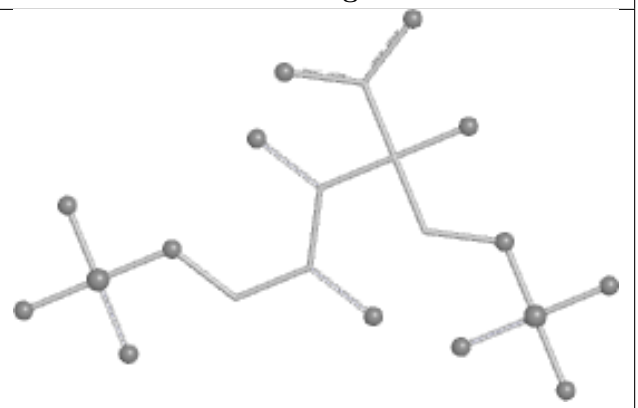
Bond lengths



Bond angles

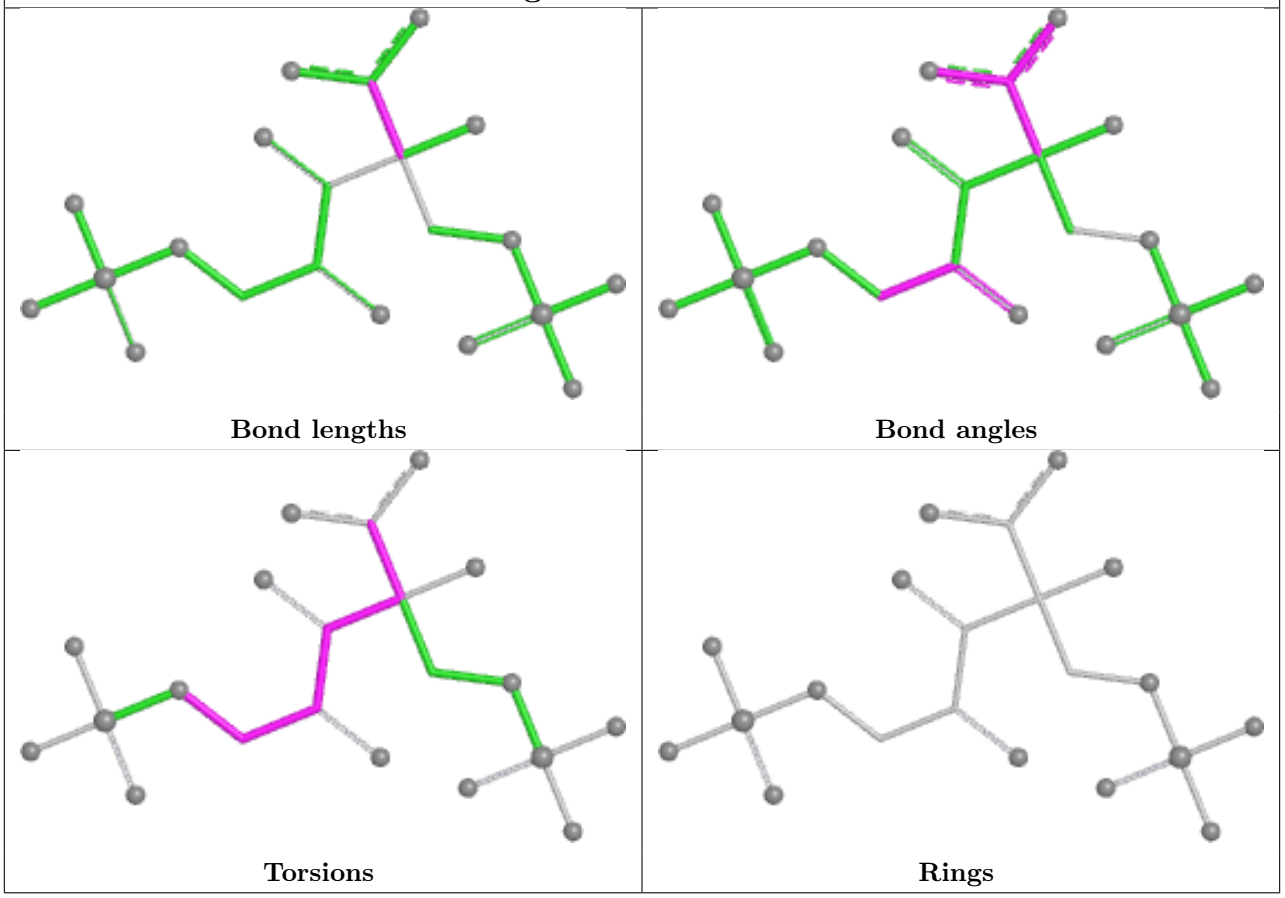


Torsions



Rings

Ligand CAP B 502



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^{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	466/475 (98%)	0.25	4 (0%) 81 78	32, 44, 63, 87	0
1	B	466/475 (98%)	0.44	4 (0%) 81 78	35, 51, 71, 84	0
1	C	466/475 (98%)	0.06	1 (0%) 92 90	30, 42, 61, 80	0
1	D	466/475 (98%)	0.10	2 (0%) 89 86	33, 45, 62, 73	0
1	E	466/475 (98%)	0.02	4 (0%) 81 78	30, 41, 61, 90	0
1	F	466/475 (98%)	0.14	6 (1%) 74 71	33, 46, 65, 90	0
1	G	466/475 (98%)	0.05	3 (0%) 85 83	30, 41, 58, 75	0
1	H	466/475 (98%)	0.29	2 (0%) 89 86	35, 46, 66, 80	0
2	I	123/123 (100%)	0.11	0 100 100	35, 46, 60, 69	0
2	J	123/123 (100%)	0.66	3 (2%) 59 56	44, 61, 73, 79	0
2	K	123/123 (100%)	0.22	0 100 100	35, 48, 61, 75	0
2	L	123/123 (100%)	0.48	0 100 100	43, 58, 75, 84	0
2	M	123/123 (100%)	0.09	0 100 100	35, 45, 61, 66	0
2	N	123/123 (100%)	0.58	0 100 100	42, 58, 72, 83	0
2	O	123/123 (100%)	0.04	0 100 100	36, 45, 59, 67	0
2	P	123/123 (100%)	0.94	8 (6%) 26 24	42, 64, 80, 89	0
All	All	4712/4784 (98%)	0.22	37 (0%) 82 79	30, 46, 68, 90	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	470	PRO	3.6
1	F	475	VAL	3.5
1	A	9	ALA	3.2
1	B	475	VAL	2.9
2	J	6	ILE	2.7

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^{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
1	KCX	B	201	12/13	0.89	0.13	51,54,57,57	0
1	KCX	F	201	12/13	0.91	0.11	43,47,55,56	0
1	KCX	G	201	12/13	0.91	0.10	35,40,44,47	0
1	KCX	A	201	12/13	0.93	0.08	36,41,46,46	0
1	KCX	E	201	12/13	0.93	0.09	36,39,42,45	0
1	KCX	C	201	12/13	0.94	0.09	34,39,43,46	0
1	KCX	H	201	12/13	0.94	0.08	36,44,53,54	0
1	KCX	D	201	12/13	0.95	0.08	42,44,50,51	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, 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	CAP	B	502	21/21	0.92	0.10	45,52,59,62	0
4	CAP	A	502	21/21	0.94	0.08	42,47,51,53	0
4	CAP	F	502	21/21	0.94	0.09	45,52,57,62	0
4	CAP	H	502	21/21	0.94	0.09	36,48,51,57	0
4	CAP	D	502	21/21	0.95	0.07	42,51,55,56	0
4	CAP	G	502	21/21	0.96	0.07	35,43,48,50	0
4	CAP	C	502	21/21	0.96	0.07	31,43,50,51	0
4	CAP	E	502	21/21	0.97	0.07	33,39,45,48	0
3	MN	E	501	1/1	0.99	0.03	50,50,50,50	0
3	MN	H	501	1/1	0.99	0.07	63,63,63,63	0
3	MN	A	501	1/1	0.99	0.02	48,48,48,48	0
3	MN	B	501	1/1	0.99	0.03	54,54,54,54	0
3	MN	D	501	1/1	0.99	0.03	53,53,53,53	0
3	MN	C	501	1/1	1.00	0.04	46,46,46,46	0

Continued on next page...

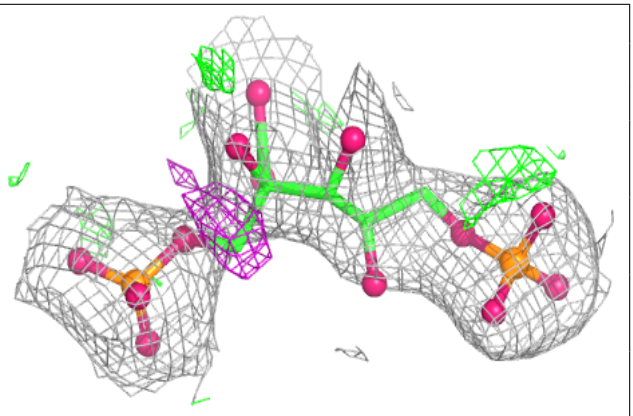
Continued from previous page...

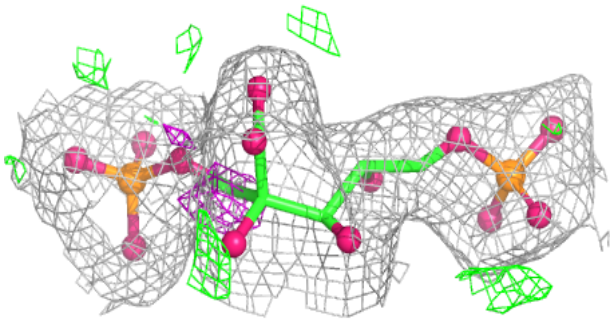
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MN	F	501	1/1	1.00	0.02	49,49,49,49	0
3	MN	G	501	1/1	1.00	0.02	47,47,47,47	0

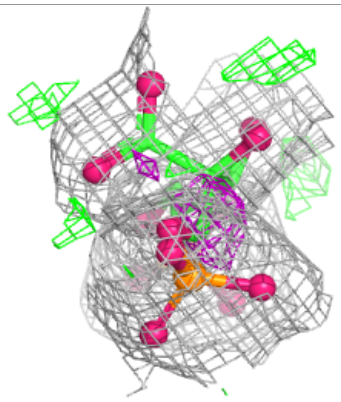
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CAP B 502:

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

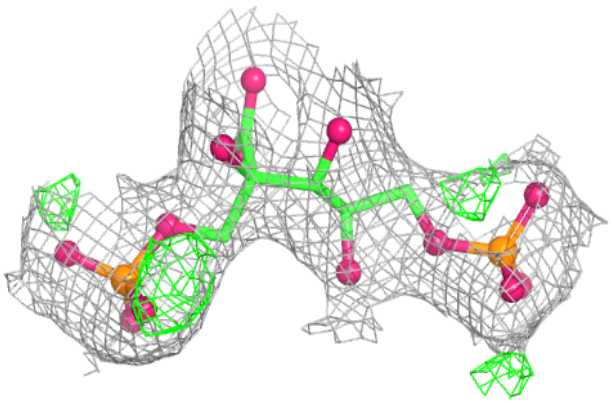


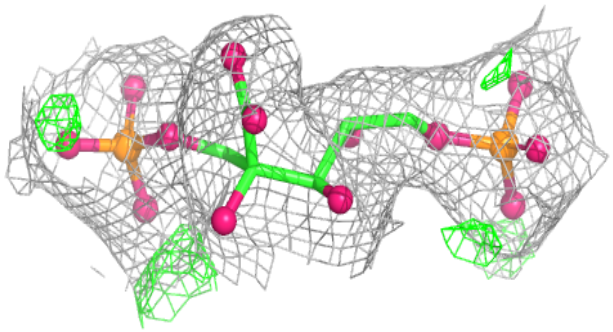


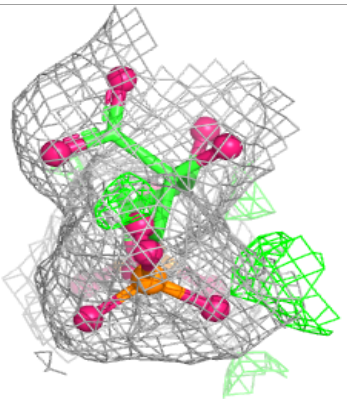


Electron density around CAP A 502:

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

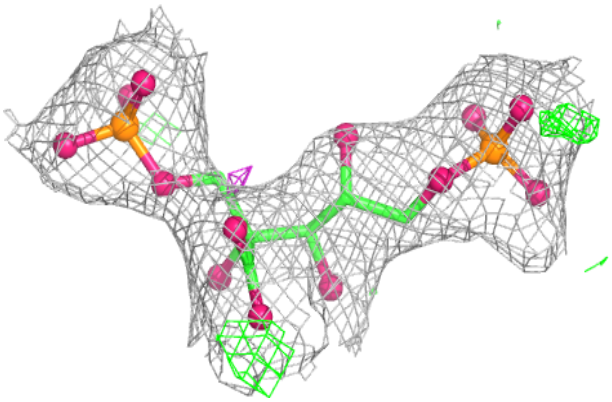


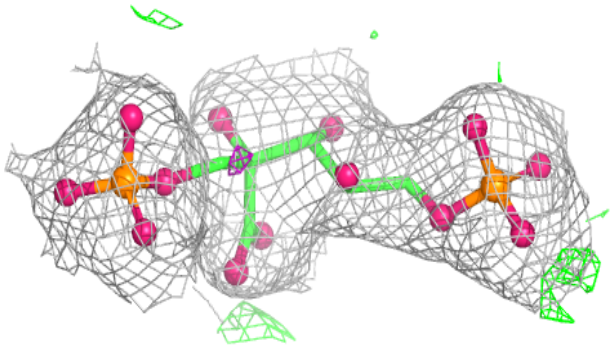


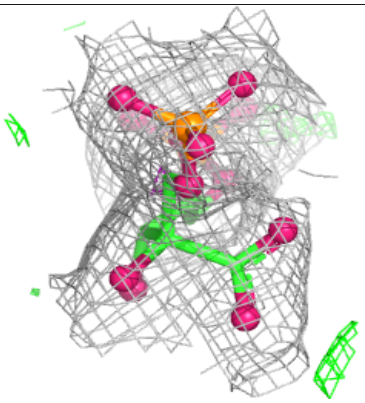


Electron density around CAP F 502:

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

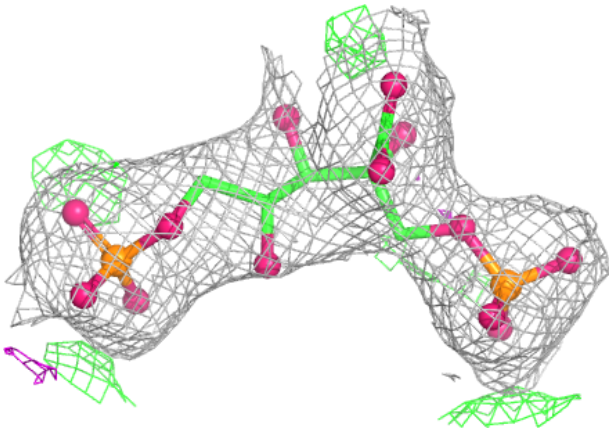


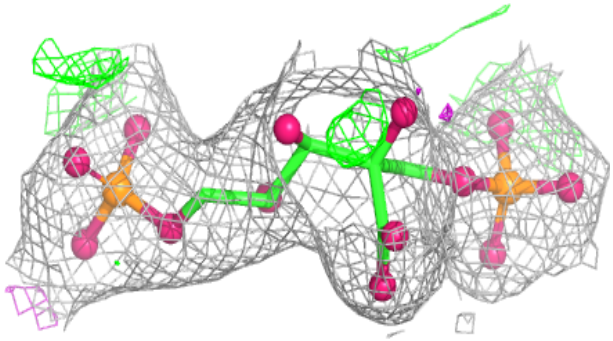


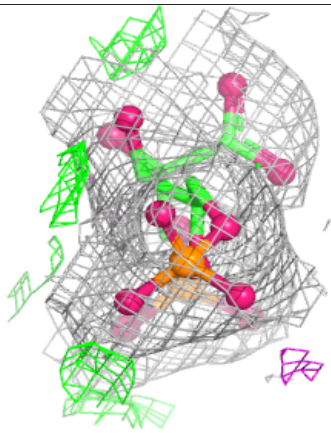


Electron density around CAP H 502:

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

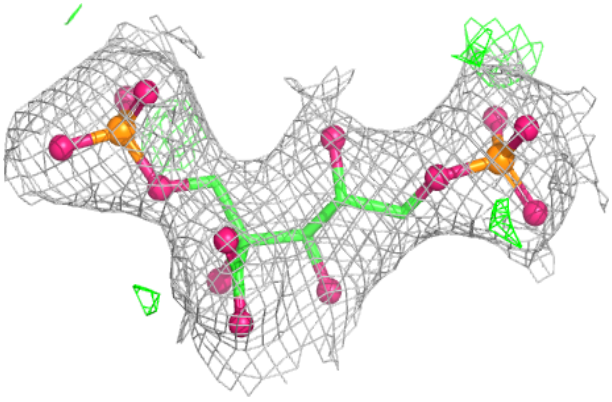


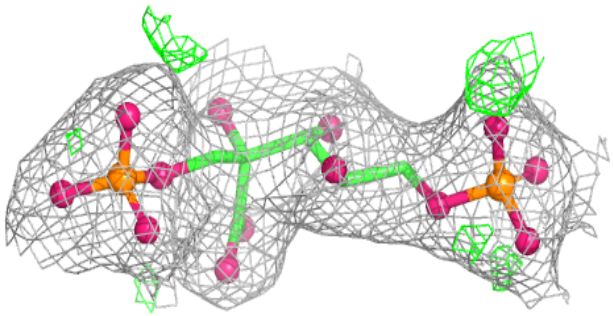


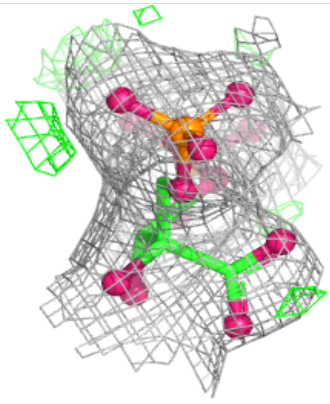


Electron density around CAP D 502:

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

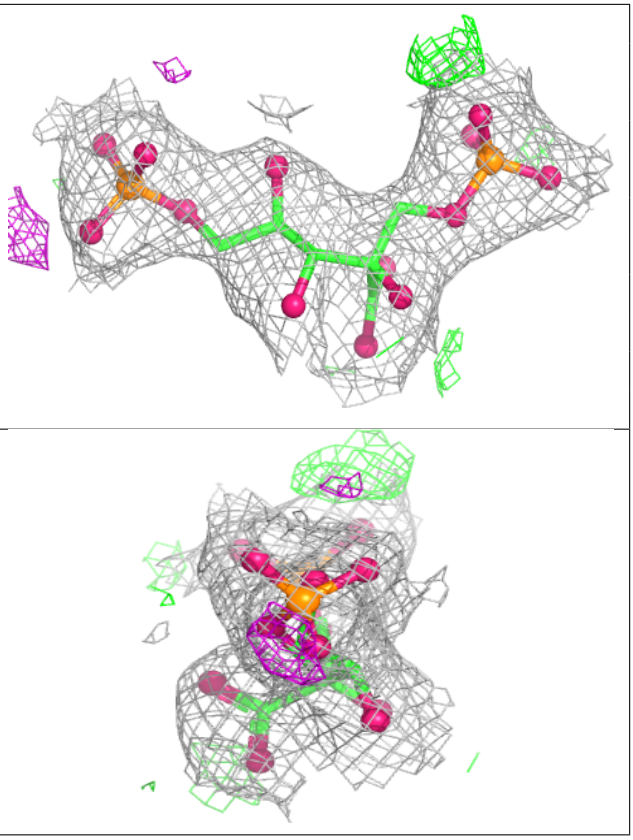






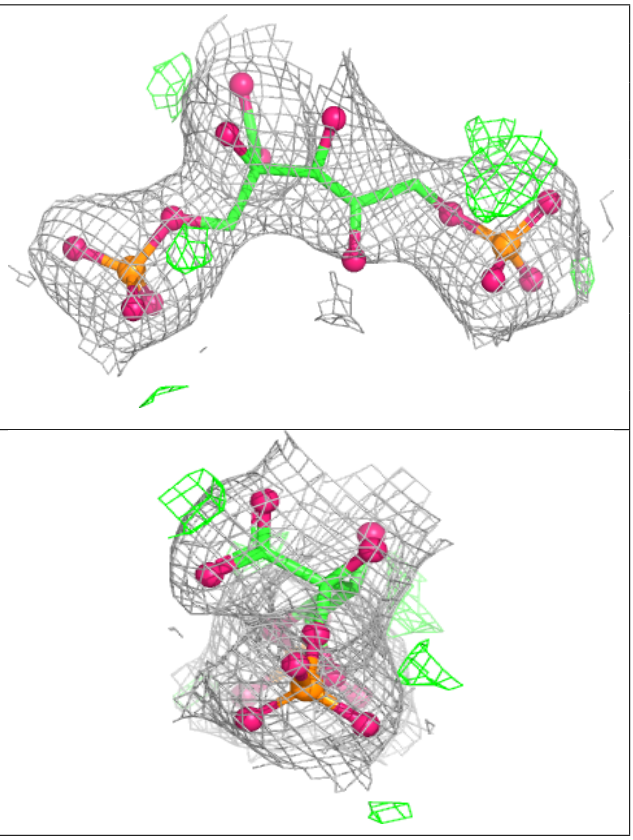
Electron density around CAP G 502:

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



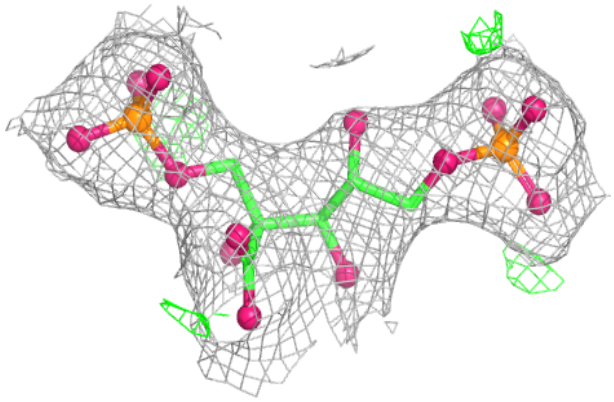
Electron density around CAP C 502:

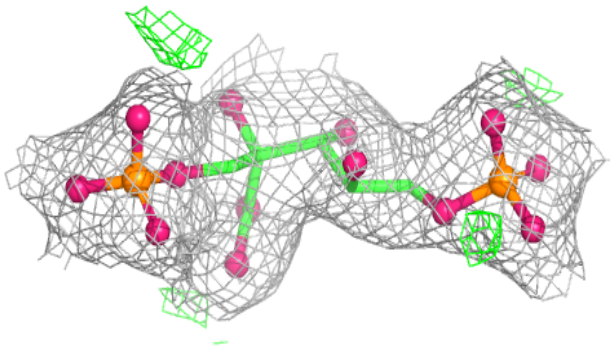
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

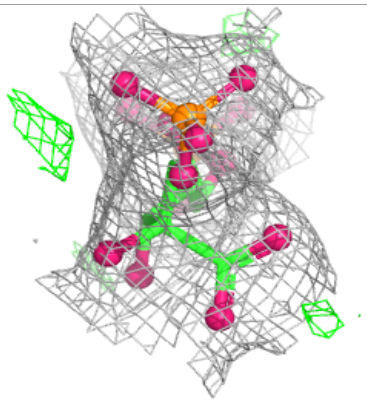


Electron density around CAP E 502:

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

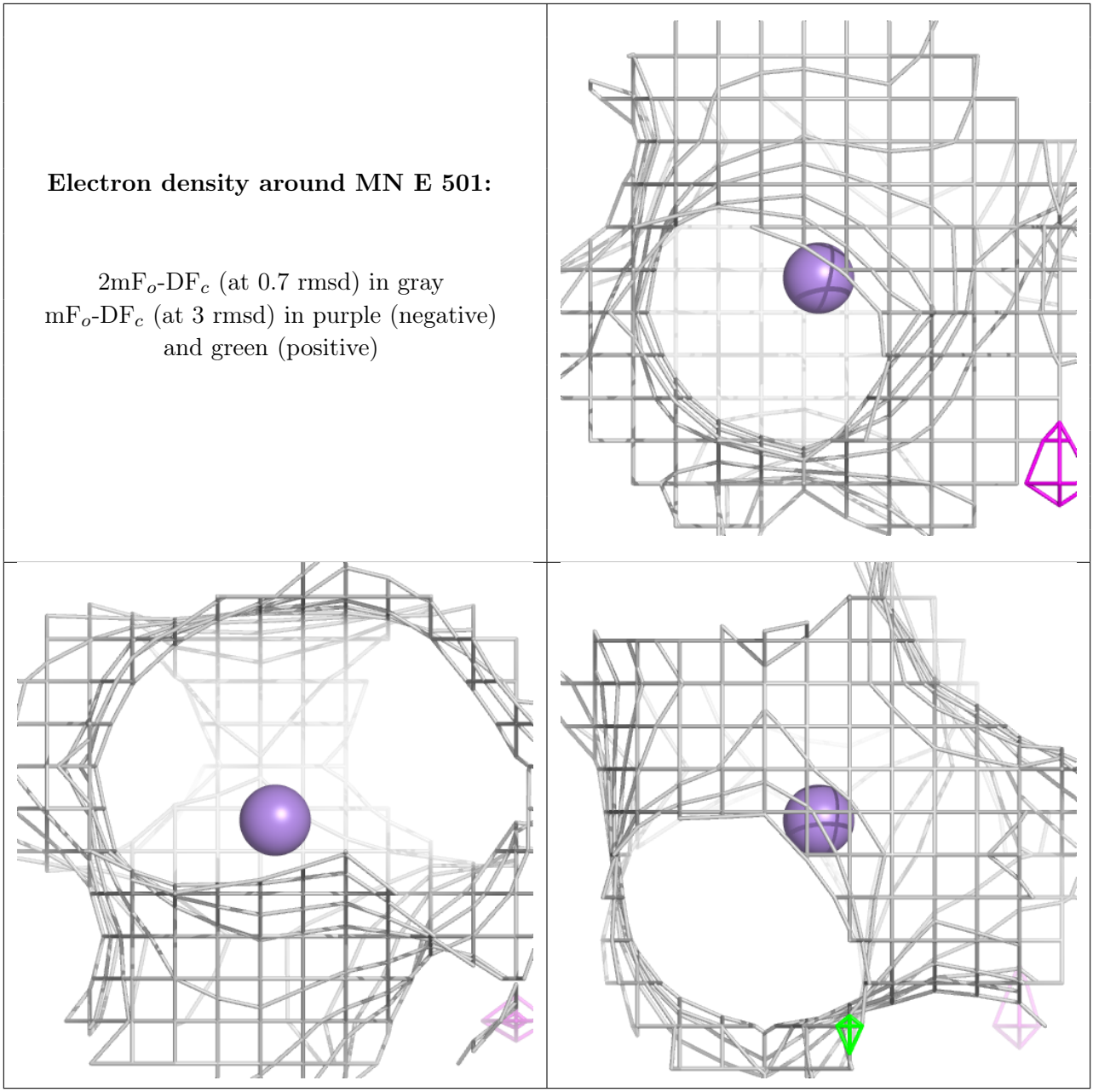






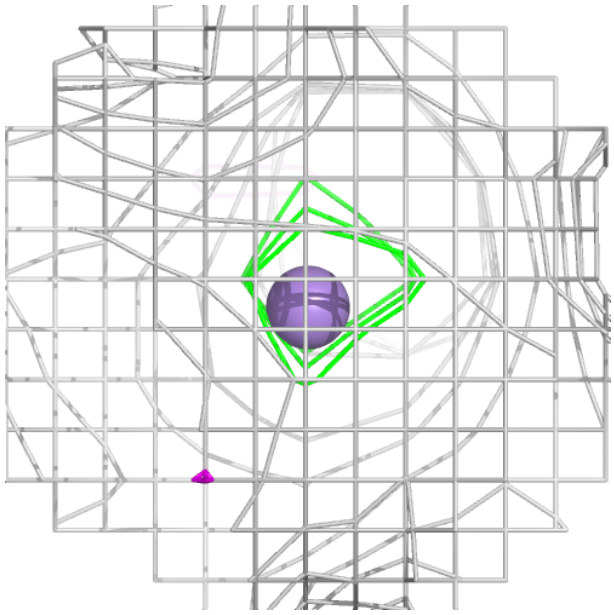
Electron density around MN E 501:

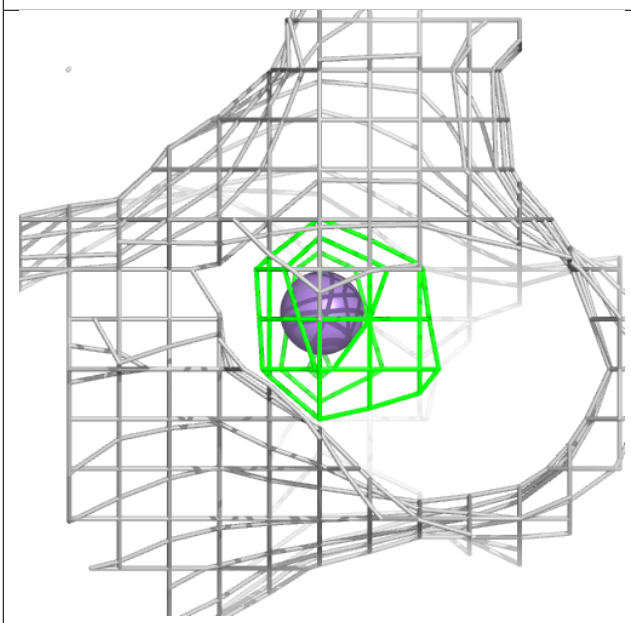
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

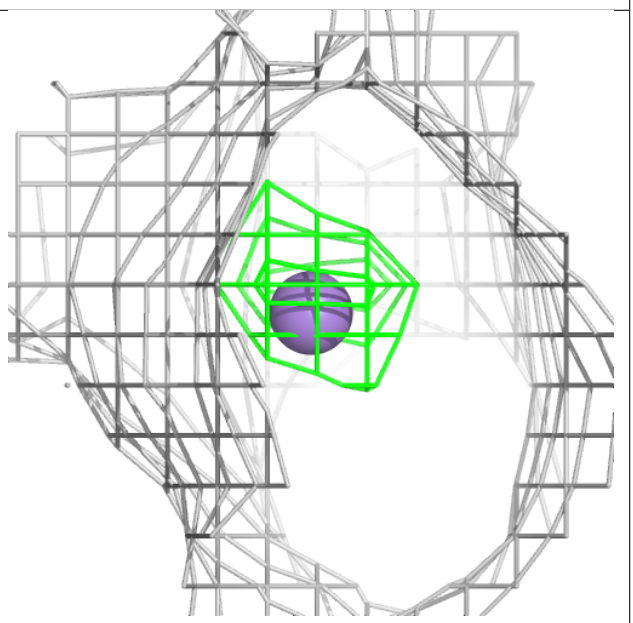


Electron density around MN H 501:

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

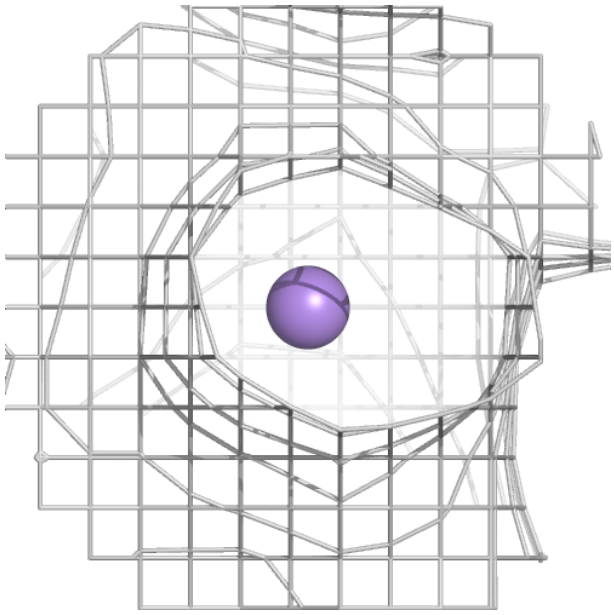


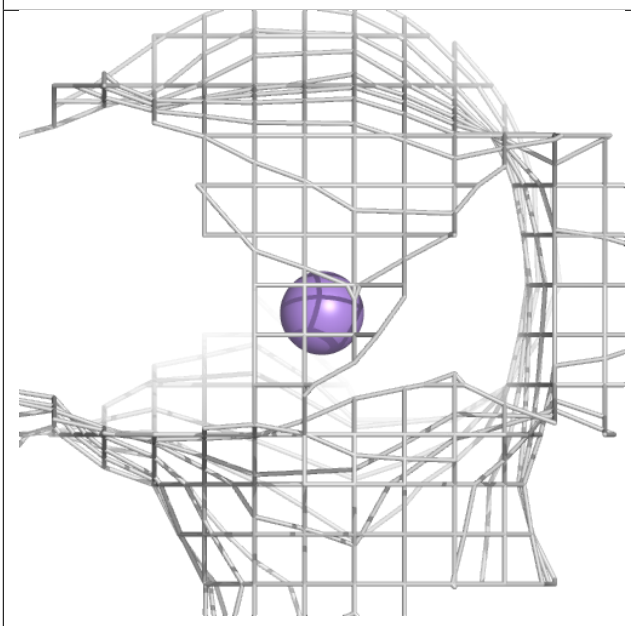


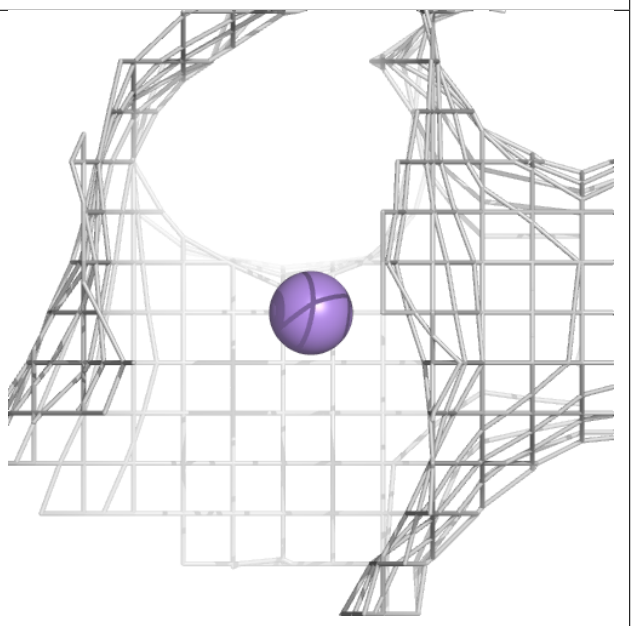


Electron density around MN A 501:

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

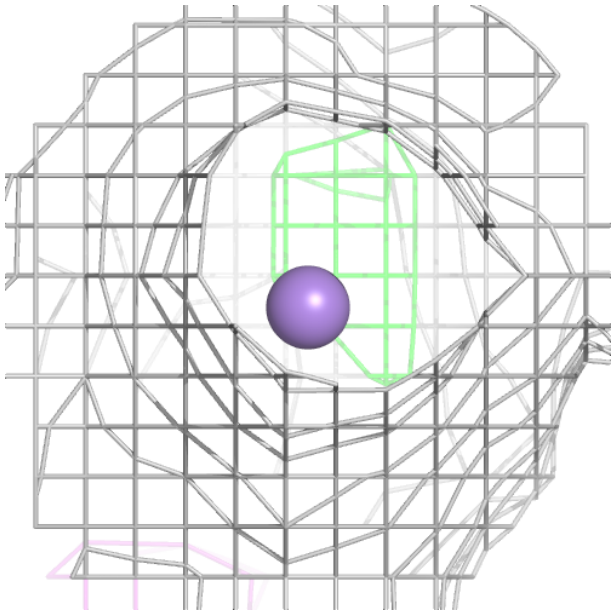


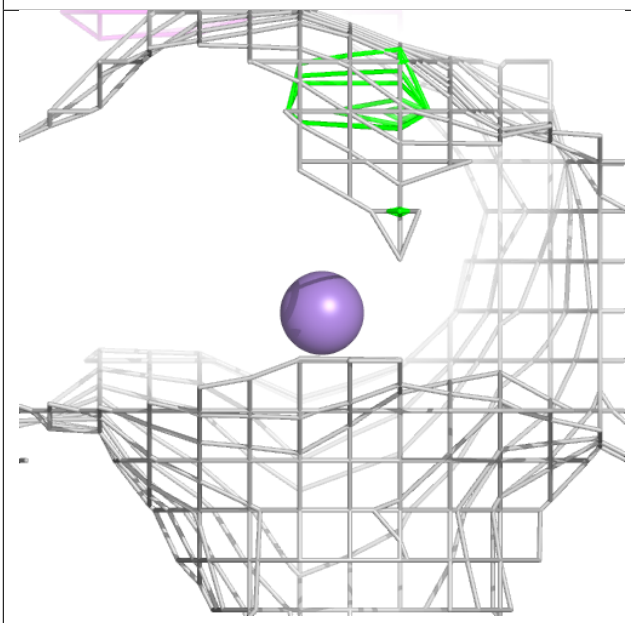


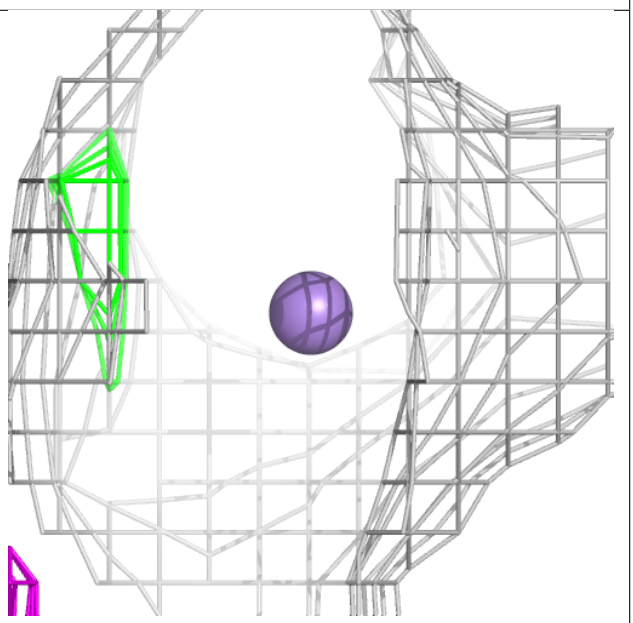


Electron density around MN B 501:

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

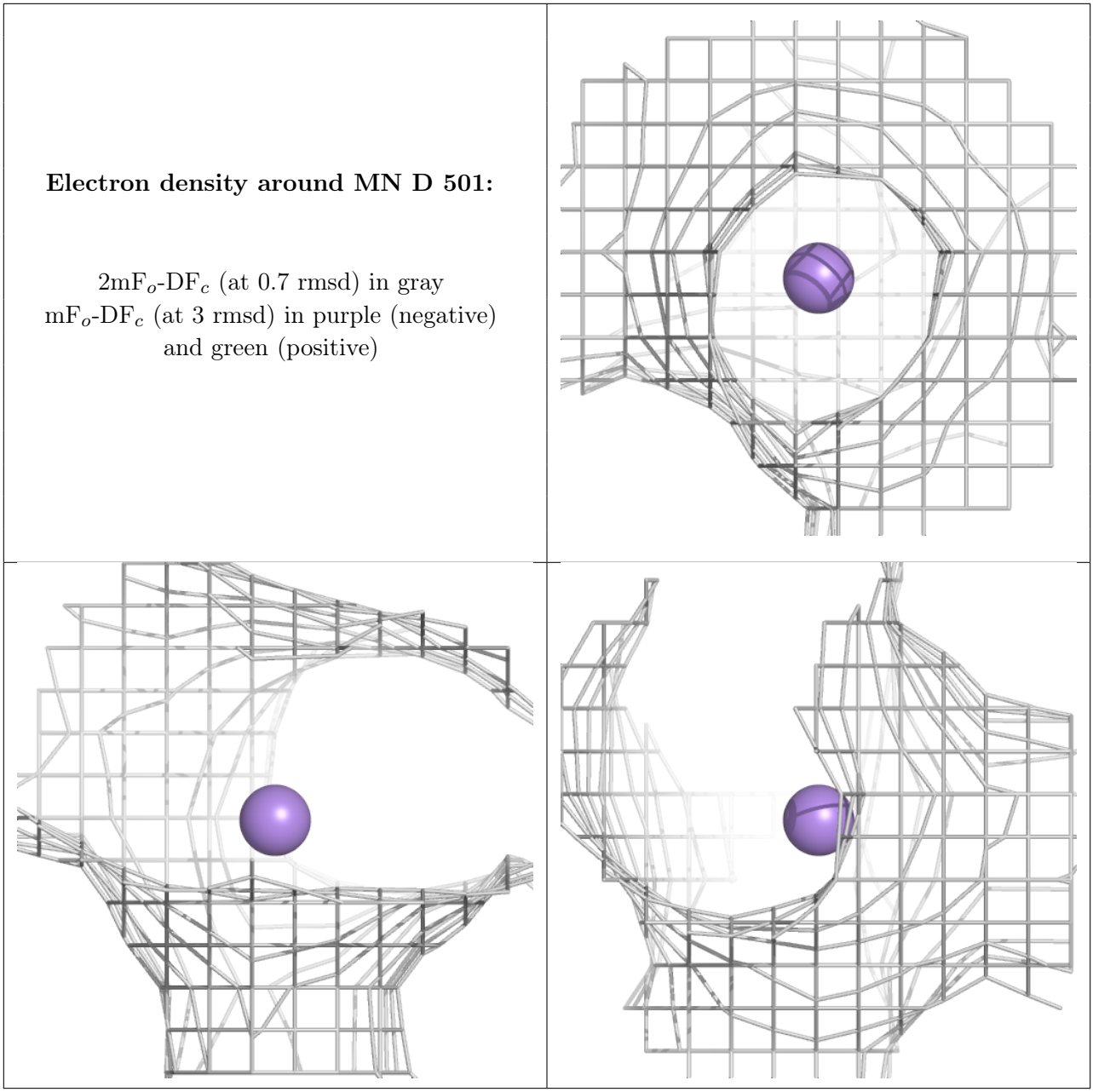






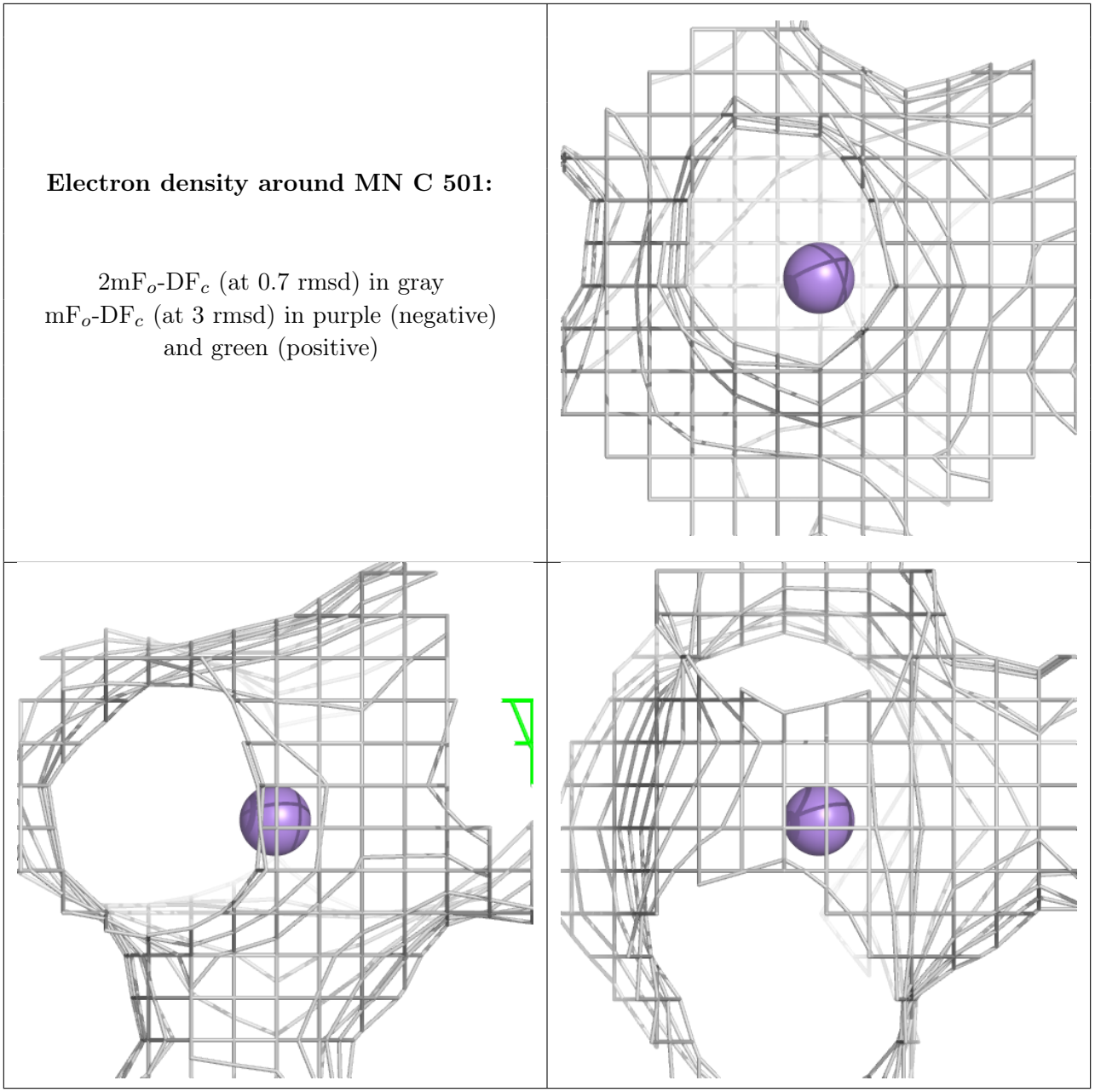
Electron density around MN D 501:

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



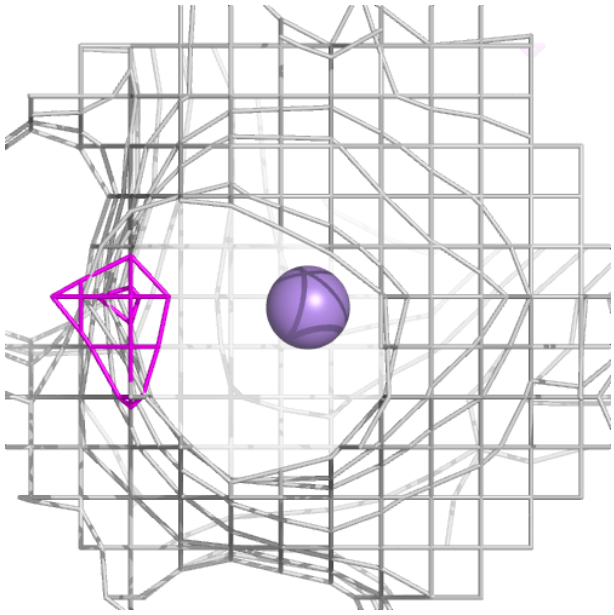
Electron density around MN C 501:

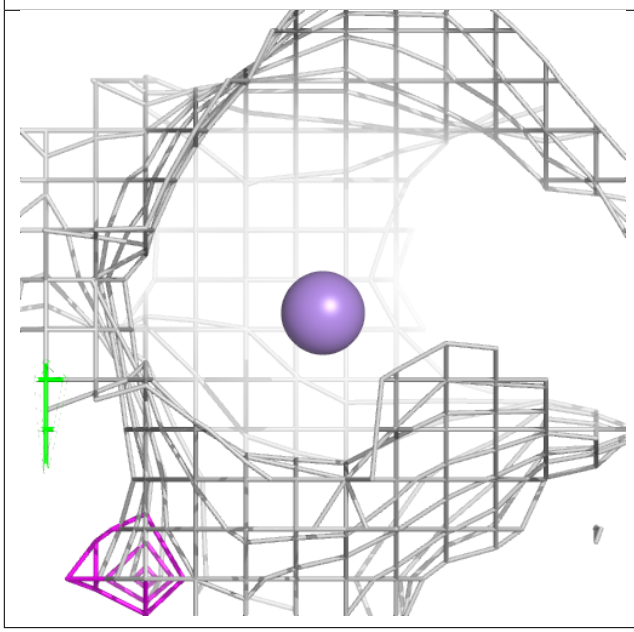
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

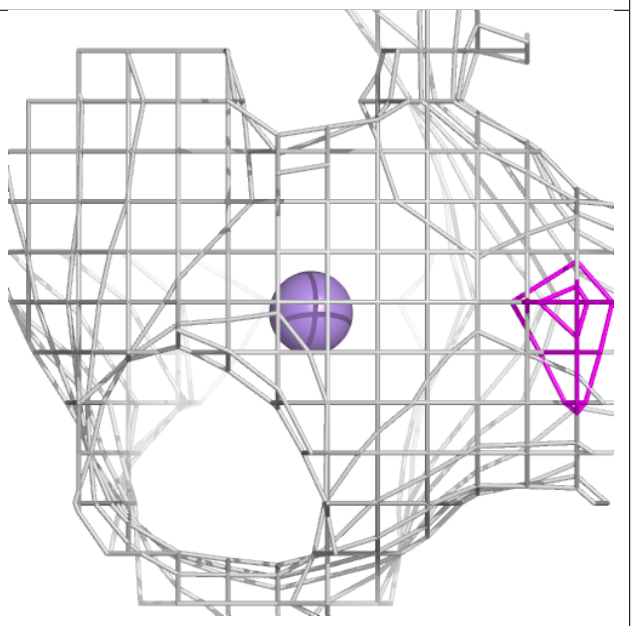


Electron density around MN F 501:

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

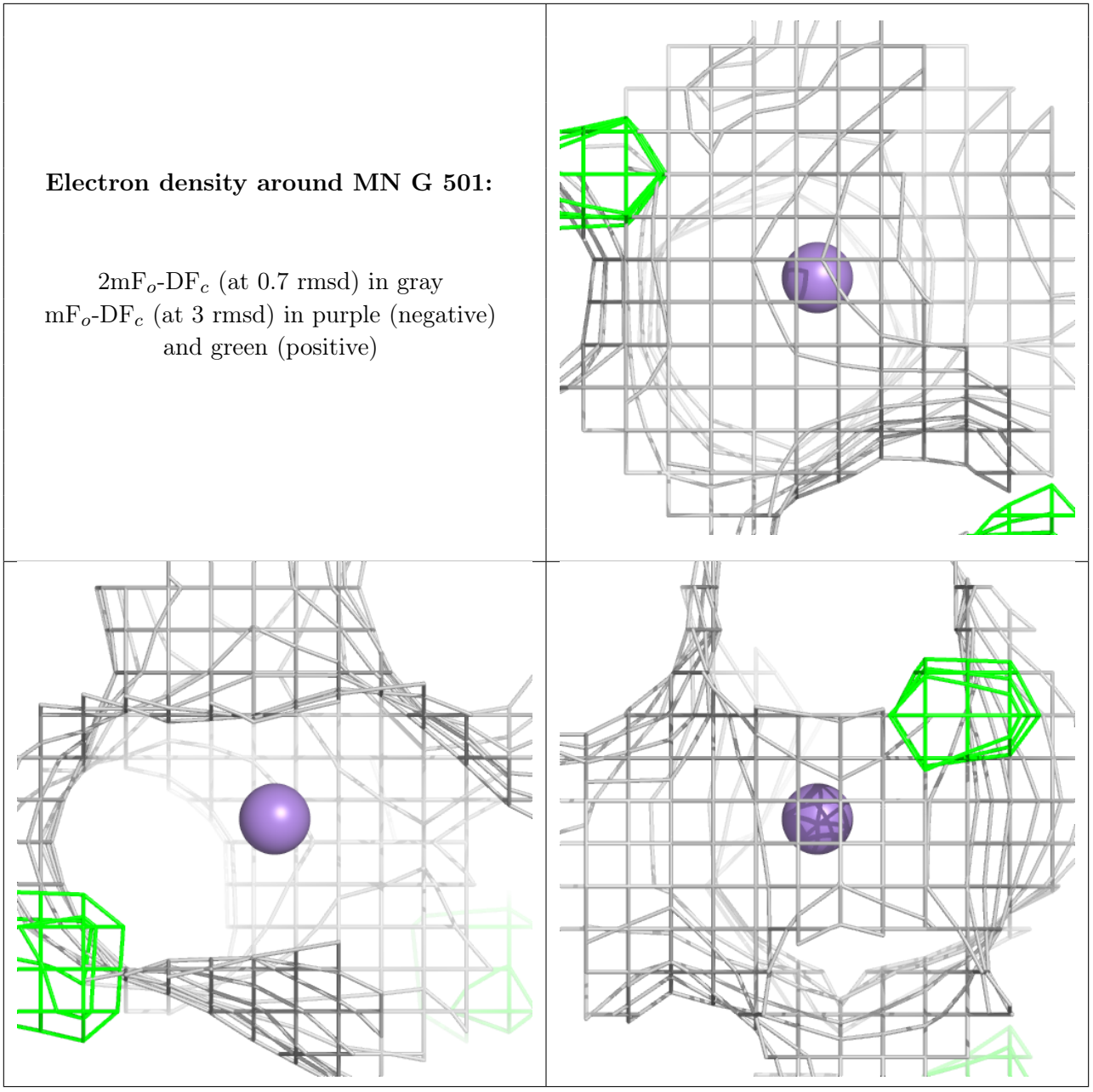






Electron density around MN G 501:

$2mF_o$-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.