wwPDB X-ray Structure Validation Summary Report (i)

Jul 31, 2023 – 10:51 AM EDT

PDB ID	:	8THM
Title	:	Beta carbonic anhydrase from the carboxysome of Cyanobium PCC 7001
Authors	:	Pulsford, S.B.; Jackson, C.J.
Deposited on	:	2023-07-17
Resolution	:	2.30 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:

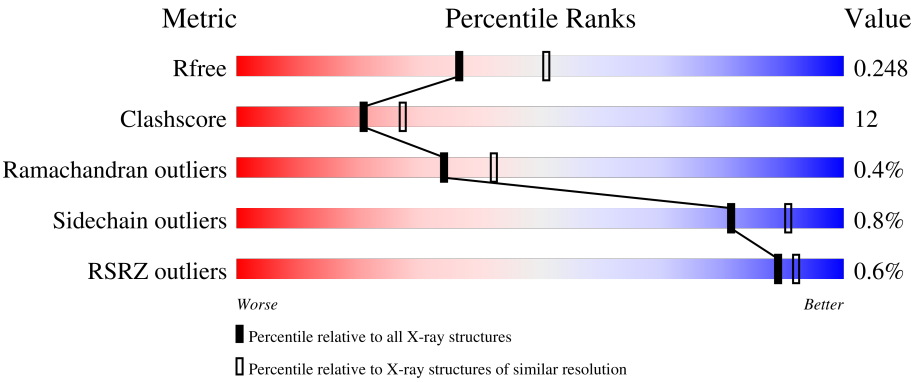
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	462	81% 19%
1	B	462	78% 22%
1	C	462	% 77% 22% .
1	D	462	79% 21%
1	E	462	% 79% 20% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	462	80% 19%

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
4	SO4	E	603	-	-	X	-
5	EDO	F	606	-	-	X	-
6	RUB	A	608	-	-	X	X
6	RUB	B	606	-	-	-	X
6	RUB	C	605	-	-	X	X
6	RUB	D	605	-	-	X	X
6	RUB	E	610	-	-	X	X
6	RUB	F	609	-	-	X	X

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 23086 atoms, of which 75 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 Carboxysome shell carbonic anhydrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	461	Total	C	N	O	S	0	7	0
			3674	2310	669	681	14			
1	B	462	Total	C	N	O	S	0	2	0
			3626	2283	653	676	14			
1	C	461	Total	C	N	O	S	0	4	0
			3635	2289	653	679	14			
1	D	462	Total	C	N	O	S	0	6	0
			3656	2301	659	682	14			
1	E	461	Total	C	N	O	S	0	7	0
			3676	2312	671	679	14			
1	F	462	Total	C	N	O	S	0	9	0
			3697	2323	674	686	14			

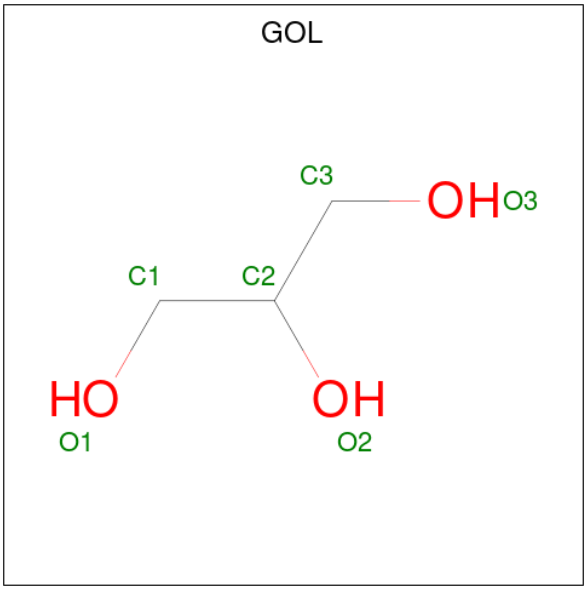
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	116	SER	-	expression tag	UNP B5ILN4
A	117	MET	-	expression tag	UNP B5ILN4
B	116	SER	-	expression tag	UNP B5ILN4
B	117	MET	-	expression tag	UNP B5ILN4
C	116	SER	-	expression tag	UNP B5ILN4
C	117	MET	-	expression tag	UNP B5ILN4
D	116	SER	-	expression tag	UNP B5ILN4
D	117	MET	-	expression tag	UNP B5ILN4
E	116	SER	-	expression tag	UNP B5ILN4
E	117	MET	-	expression tag	UNP B5ILN4
F	116	SER	-	expression tag	UNP B5ILN4
F	117	MET	-	expression tag	UNP B5ILN4

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

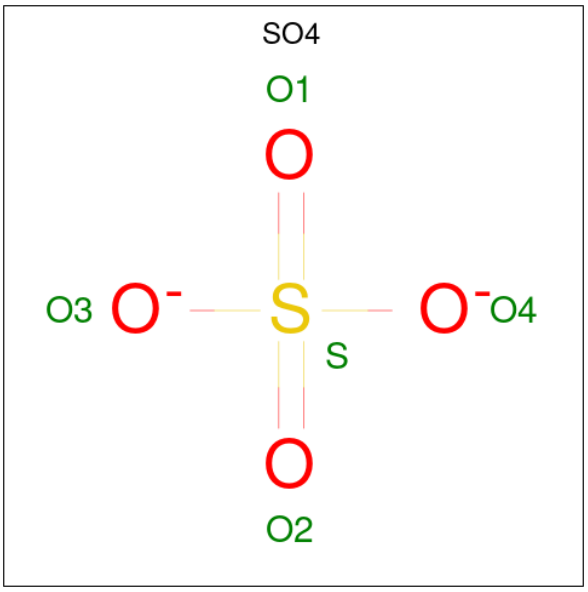
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Zn 2 2	0	0
2	B	2	Total Zn 2 2	0	0
2	C	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C H O 14 3 8 3	0	0
3	D	1	Total C O 6 3 3	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



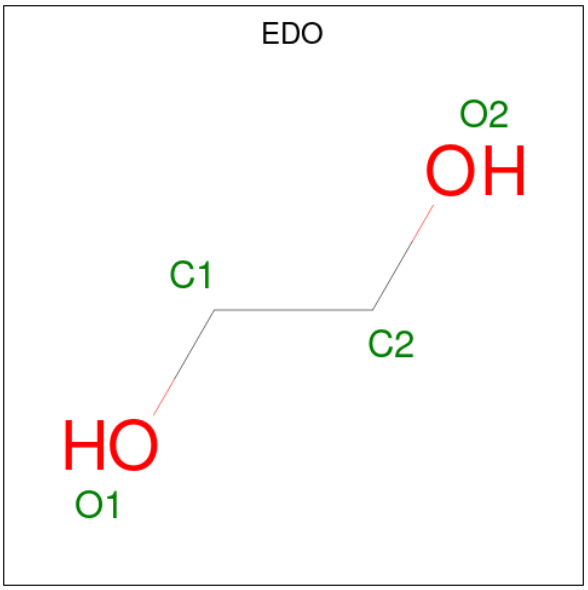
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	F	1	Total O S 5 4 1	0	0
4	F	1	Total O S 5 4 1	0	0
4	F	1	Total O S 5 4 1	0	0
4	F	1	Total O S 5 4 1	0	0

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



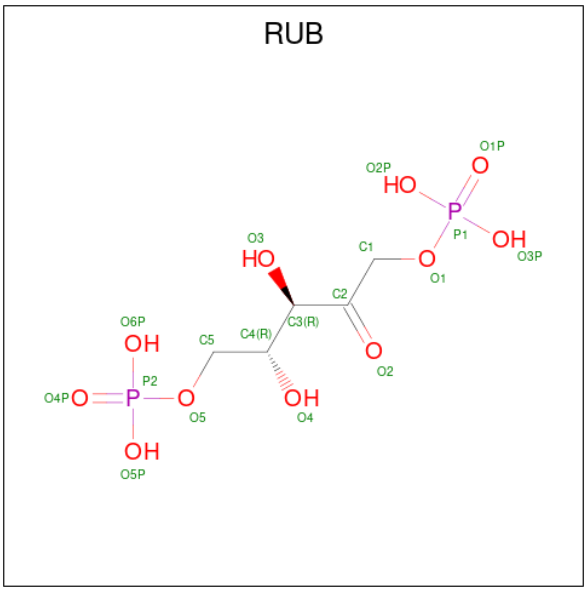
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C H O 10 2 6 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	E	1	Total C H O 10 2 6 2	0	0
5	E	1	Total C H O 10 2 6 2	0	0

Continued on next page...

Continued from previous page...

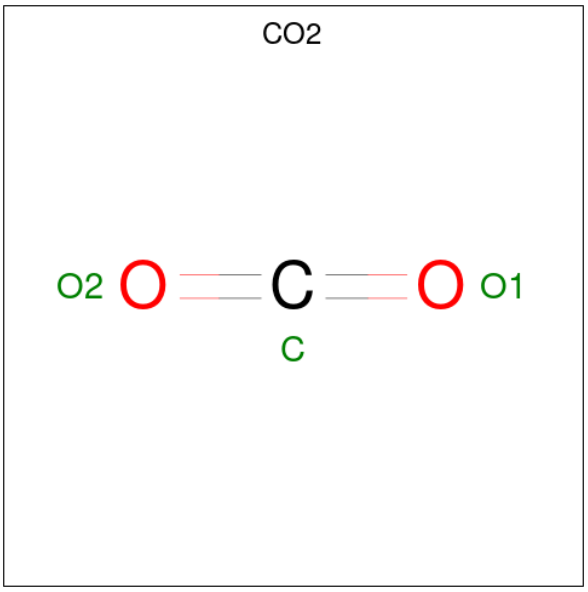
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0

- Molecule 6 is RIBULOSE-1,5-DIPHOSPHATE (three-letter code: RUB) (formula: $C_5H_{12}O_{11}P_2$) (labeled as "Ligand of Interest" by depositor).



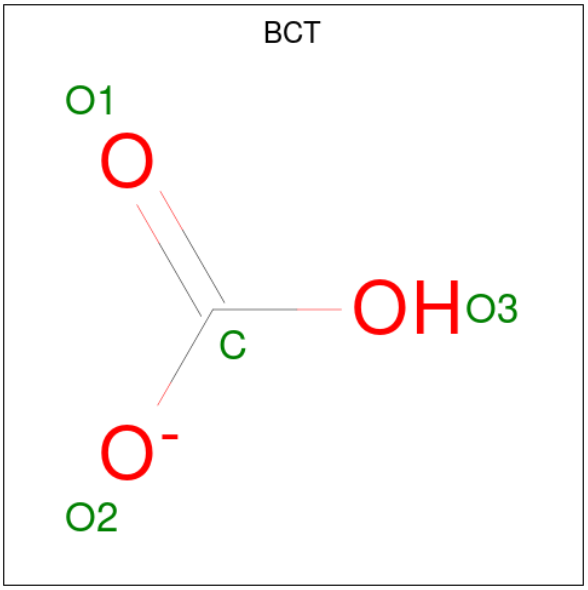
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H O P 26 5 8 11 2	0	0
6	B	1	Total C H O P 26 5 8 11 2	0	0
6	C	1	Total C H O P 26 5 8 11 2	0	0
6	D	1	Total C H O P 26 5 8 11 2	0	0
6	E	1	Total C H O P 26 5 8 11 2	0	0
6	F	1	Total C H O P 26 5 8 11 2	0	0

- Molecule 7 is CARBON DIOXIDE (three-letter code: CO2) (formula: CO_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	1	Total C O 3 1 2	0	0
7	D	1	Total C O 3 1 2	0	0

- Molecule 8 is BICARBONATE ION (three-letter code: BCT) (formula: CHO_3).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	1	Total C H O 5 1 1 3	0	0

- Molecule 9 is water.

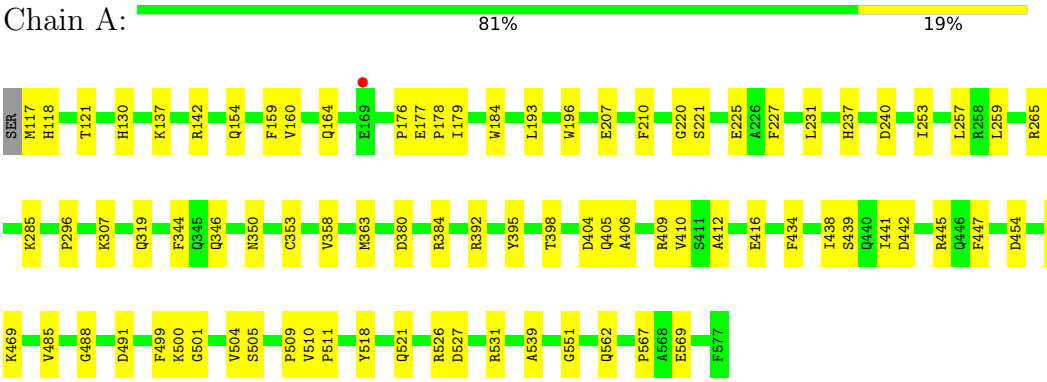
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	158	Total O 158 158	0	0
9	B	126	Total O 126 126	0	0
9	C	108	Total O 108 108	0	0
9	D	128	Total O 128 128	0	0
9	E	115	Total O 115 115	0	0
9	F	138	Total O 138 138	0	0

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.

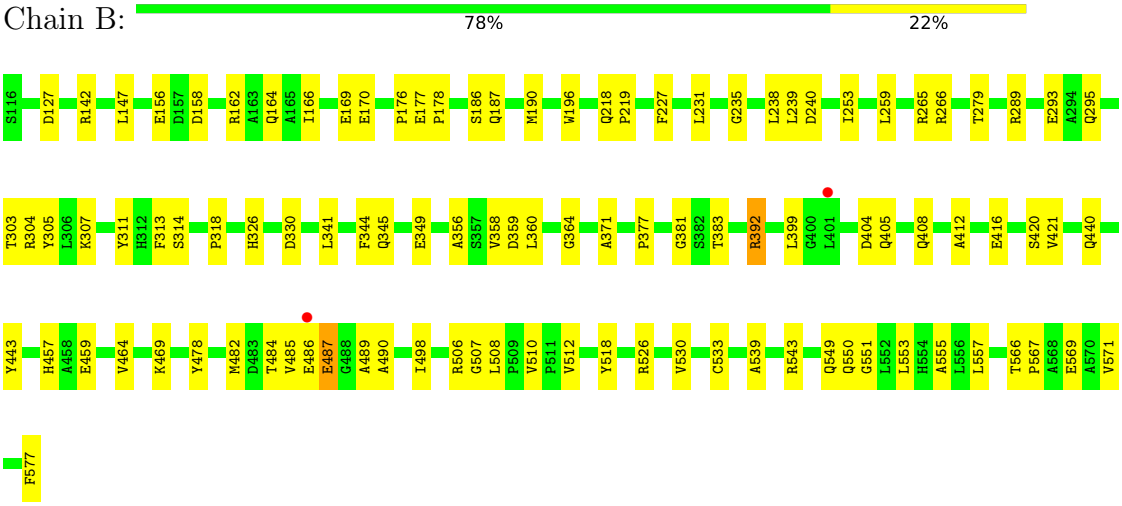
• Molecule 1: Carboxysome shell carbonic anhydrase

Chain A:



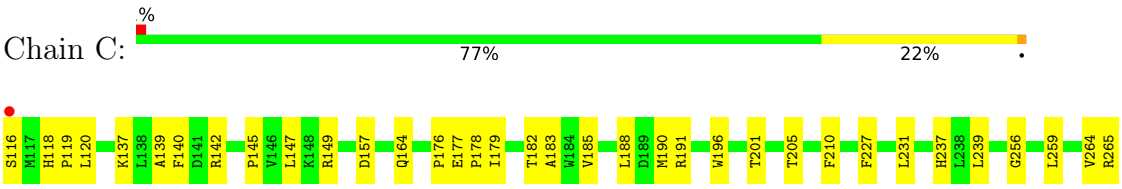
• Molecule 1: Carboxysome shell carbonic anhydrase

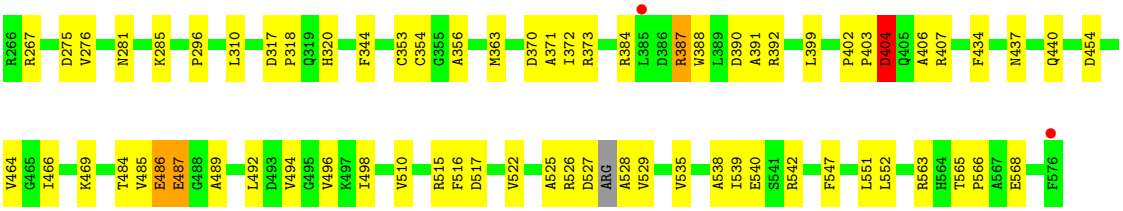
Chain B:



• Molecule 1: Carboxysome shell carbonic anhydrase

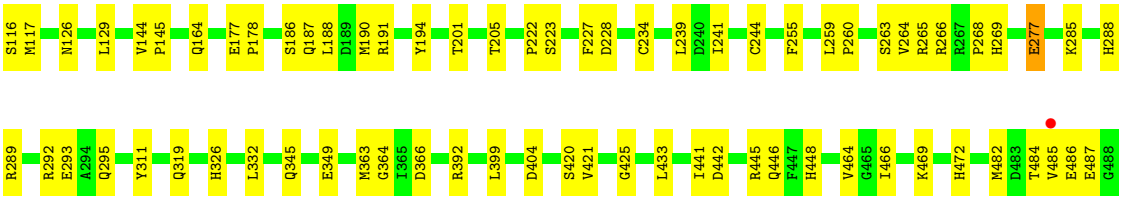
Chain C:





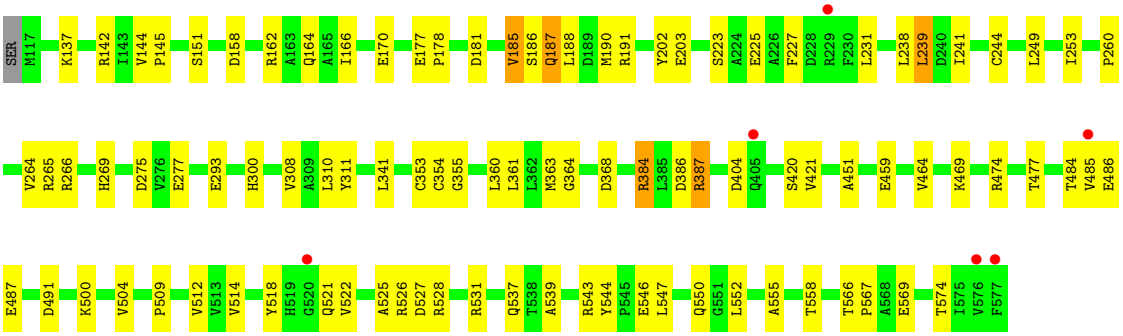
• Molecule 1: Carboxysome shell carbonic anhydrase

Chain D:



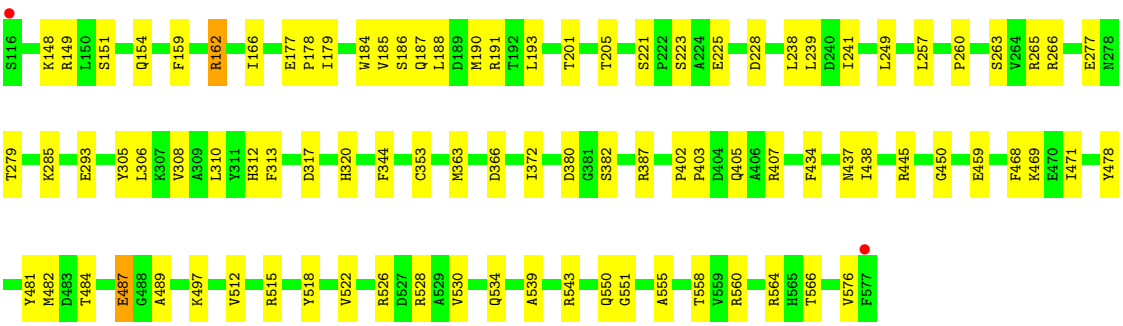
• Molecule 1: Carboxysome shell carbonic anhydrase

Chain E:



• Molecule 1: Carboxysome shell carbonic anhydrase

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	104.31Å 181.85Å 190.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.51 – 2.30 48.48 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.51-2.30) 100.0 (48.48-2.30)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 2.29Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.193 , 0.247 0.194 , 0.248	Depositor DCC
R_{free} test set	7994 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	45.3	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 35.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.003 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	23086	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.39% 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: SO4, BCT, EDO, RUB, GOL, ZN, CO2

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	0.54	0/3768	0.62	0/5120
1	B	0.49	0/3717	0.62	0/5055
1	C	0.49	0/3725	0.62	0/5065
1	D	0.53	1/3751 (0.0%)	0.63	1/5100 (0.0%)
1	E	0.49	0/3770	0.62	1/5121 (0.0%)
1	F	0.48	0/3795	0.60	0/5156
All	All	0.50	1/22526 (0.0%)	0.62	2/30617 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	194	TYR	CE1-CZ	-6.17	1.30	1.38

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	366	ASP	CB-CG-OD1	5.87	123.59	118.30
1	E	239	LEU	CB-CG-CD2	5.09	119.65	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3674	0	3553	74	0
1	B	3626	0	3498	79	0
1	C	3635	0	3507	108	0
1	D	3656	0	3526	98	0
1	E	3676	0	3563	90	0
1	F	3697	0	3571	96	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	6	0	8	0	0
3	B	6	8	8	2	0
3	D	6	0	8	0	0
4	A	15	0	0	1	0
4	B	10	0	0	0	0
4	C	10	0	0	1	0
4	D	15	0	0	1	0
4	E	20	0	0	2	0
4	F	20	0	0	1	0
5	A	4	0	6	0	0
5	B	4	0	6	0	0
5	C	4	6	6	0	0
5	E	16	12	24	1	0
5	F	12	0	18	6	0
6	A	18	8	8	11	0
6	B	18	8	8	6	0
6	C	18	8	8	8	0
6	D	18	8	8	11	0
6	E	18	8	8	8	0
6	F	18	8	8	16	0
7	C	3	0	0	0	0
7	D	3	0	0	0	0
8	D	4	1	1	1	0
9	A	158	0	0	12	0
9	B	126	0	0	10	0
9	C	108	0	0	13	0
9	D	128	0	0	15	0
9	E	115	0	0	8	0
9	F	138	0	0	7	0
All	All	23011	75	21351	533	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 12.

The worst 5 of 533 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:E:341:LEU:HB3	1:E:360:LEU:HD22	1.30	1.10
1:E:249:LEU:HD22	9:E:778:HOH:O	1.57	1.03
1:C:406:ALA:HB3	9:C:701:HOH:O	1.57	1.01
1:A:265[A]:ARG:HA	6:A:608:RUB:H52	1.43	1.01
1:A:265[B]:ARG:HA	6:A:608:RUB:H52	1.43	1.00

There are no symmetry-related clashes.

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	466/462 (101%)	451 (97%)	15 (3%)	0	100	100
1	B	462/462 (100%)	443 (96%)	17 (4%)	2 (0%)	34	42
1	C	460/462 (100%)	441 (96%)	14 (3%)	5 (1%)	14	15
1	D	466/462 (101%)	448 (96%)	17 (4%)	1 (0%)	47	58
1	E	466/462 (101%)	446 (96%)	17 (4%)	3 (1%)	25	31
1	F	469/462 (102%)	459 (98%)	10 (2%)	0	100	100
All	All	2789/2772 (101%)	2688 (96%)	90 (3%)	11 (0%)	34	42

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	487	GLU
1	C	157	ASP
1	D	404	ASP
1	C	404	ASP
1	E	187	GLN

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	384/378 (102%)	380 (99%)	4 (1%)	76	87
1	B	379/378 (100%)	378 (100%)	1 (0%)	92	97
1	C	381/378 (101%)	378 (99%)	3 (1%)	81	91
1	D	383/378 (101%)	381 (100%)	2 (0%)	88	95
1	E	384/378 (102%)	378 (98%)	6 (2%)	62	78
1	F	387/378 (102%)	382 (99%)	5 (1%)	69	82
All	All	2298/2268 (101%)	2277 (99%)	21 (1%)	81	89

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	387	ARG
1	F	162[B]	ARG
1	F	487[B]	GLU
1	F	387	ARG
1	F	162[A]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	312	HIS
1	D	446	GLN
1	E	300	HIS
1	A	346	GLN
1	A	281	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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 48 ligands modelled in this entry, 8 are monoatomic - leaving 40 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	SO4	D	604	-	4,4,4	0.11	0	6,6,6	0.20	0
4	SO4	A	605	-	4,4,4	0.14	0	6,6,6	0.46	0
3	GOL	A	603	-	5,5,5	1.16	0	5,5,5	0.85	0
5	EDO	B	605	-	3,3,3	0.55	0	2,2,2	0.23	0
6	RUB	E	610	-	15,17,17	1.56	3 (20%)	16,25,25	1.02	1 (6%)
4	SO4	C	602	-	4,4,4	1.12	0	6,6,6	0.49	0
6	RUB	A	608	-	15,17,17	1.48	1 (6%)	16,25,25	1.27	1 (6%)
4	SO4	D	603	-	4,4,4	0.20	0	6,6,6	0.47	0
6	RUB	F	609	-	15,17,17	1.59	3 (20%)	16,25,25	1.34	3 (18%)
3	GOL	D	606	-	5,5,5	1.38	1 (20%)	5,5,5	0.70	0
4	SO4	E	605	-	4,4,4	0.09	0	6,6,6	0.46	0
5	EDO	F	606	-	3,3,3	0.56	0	2,2,2	0.16	0
6	RUB	B	606	-	15,17,17	1.11	1 (6%)	16,25,25	0.80	0
8	BCT	D	608	-	2,3,3	1.11	0	2,3,3	3.76	1 (50%)
5	EDO	C	604	-	3,3,3	0.58	0	2,2,2	0.29	0
6	RUB	D	605	-	15,17,17	1.52	2 (13%)	16,25,25	1.02	2 (12%)
5	EDO	E	609	-	3,3,3	0.60	0	2,2,2	0.33	0
5	EDO	A	607	-	3,3,3	0.54	0	2,2,2	0.43	0
5	EDO	F	607	-	3,3,3	0.50	0	2,2,2	0.35	0
5	EDO	E	606	-	3,3,3	0.52	0	2,2,2	0.28	0
4	SO4	E	602	-	4,4,4	0.14	0	6,6,6	0.30	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	SO4	F	605	-	4,4,4	0.22	0	6,6,6	0.22	0
4	SO4	B	603	-	4,4,4	0.14	0	6,6,6	0.36	0
7	CO2	C	606	-	2,2,2	1.27	0	1,1,1	0.50	0
4	SO4	A	604	-	4,4,4	0.10	0	6,6,6	0.41	0
4	SO4	D	602	-	4,4,4	0.19	0	6,6,6	0.33	0
4	SO4	C	603	-	4,4,4	0.17	0	6,6,6	0.25	0
4	SO4	A	606	-	4,4,4	0.18	0	6,6,6	0.49	0
5	EDO	E	607	-	3,3,3	0.55	0	2,2,2	0.21	0
6	RUB	C	605	-	15,17,17	1.44	2 (13%)	16,25,25	1.34	1 (6%)
4	SO4	F	604	-	4,4,4	0.19	0	6,6,6	0.22	0
4	SO4	E	603	-	4,4,4	0.09	0	6,6,6	0.30	0
3	GOL	B	607	-	5,5,5	1.16	0	5,5,5	1.14	0
4	SO4	B	604	-	4,4,4	0.23	0	6,6,6	0.24	0
5	EDO	F	608	-	3,3,3	0.57	0	2,2,2	0.18	0
4	SO4	E	604	-	4,4,4	0.17	0	6,6,6	0.23	0
4	SO4	F	602	-	4,4,4	0.15	0	6,6,6	0.28	0
4	SO4	F	603	-	4,4,4	0.10	0	6,6,6	0.34	0
5	EDO	E	608	-	3,3,3	0.62	0	2,2,2	0.02	0
7	CO2	D	607	-	2,2,2	1.15	0	1,1,1	0.49	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
3	GOL	A	603	-	-	4/4/4/4	-
5	EDO	B	605	-	-	0/1/1/1	-
6	RUB	E	610	-	-	12/20/20/20	-
6	RUB	A	608	-	-	13/20/20/20	-
6	RUB	F	609	-	-	14/20/20/20	-
3	GOL	D	606	-	-	0/4/4/4	-
5	EDO	F	606	-	-	1/1/1/1	-
6	RUB	B	606	-	-	8/20/20/20	-
5	EDO	C	604	-	-	0/1/1/1	-
6	RUB	D	605	-	-	7/20/20/20	-
5	EDO	E	609	-	-	0/1/1/1	-
5	EDO	F	607	-	-	0/1/1/1	-
5	EDO	A	607	-	-	1/1/1/1	-
5	EDO	E	606	-	-	0/1/1/1	-
5	EDO	E	607	-	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	RUB	C	605	-	-	11/20/20/20	-
3	GOL	B	607	-	-	2/4/4/4	-
5	EDO	F	608	-	-	1/1/1/1	-
5	EDO	E	608	-	-	0/1/1/1	-

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	605	RUB	O1-C1	-4.10	1.40	1.43
6	A	608	RUB	O1-C1	-3.98	1.40	1.43
6	E	610	RUB	O1-C1	-3.81	1.40	1.43
6	C	605	RUB	P2-O5	3.23	1.70	1.60
6	F	609	RUB	O1-C1	-3.10	1.41	1.43

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	D	608	BCT	O2-C-O1	5.14	132.87	119.55
6	A	608	RUB	O5P-P2-O5	3.65	116.45	106.73
6	C	605	RUB	O5P-P2-O5	2.93	114.53	106.73
6	F	609	RUB	O5-C5-C4	2.60	116.29	109.36
6	D	605	RUB	O5-P2-O4P	2.41	113.23	106.47

There are no chirality outliers.

5 of 74 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	603	GOL	C1-C2-C3-O3
3	B	607	GOL	C1-C2-C3-O3
6	A	608	RUB	O1-C1-C2-C3
6	A	608	RUB	O1-C1-C2-O2
6	A	608	RUB	O2-C2-C3-O3

There are no ring outliers.

15 monomers are involved in 76 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	604	SO4	1	0
4	A	605	SO4	1	0
6	E	610	RUB	8	0

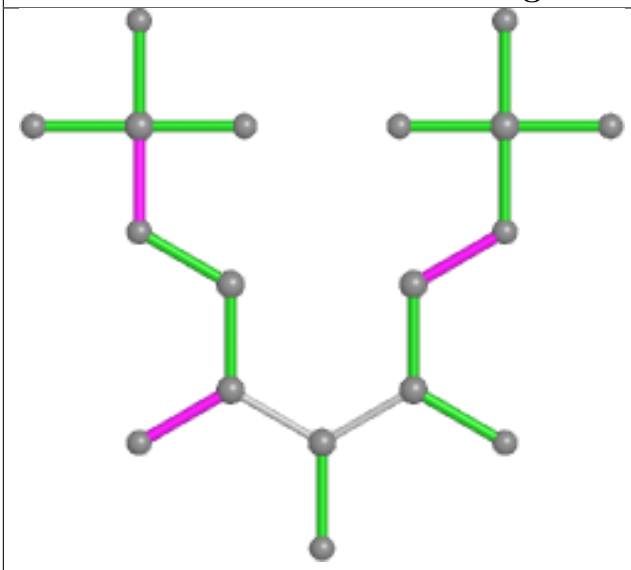
Continued on next page...

Continued from previous page...

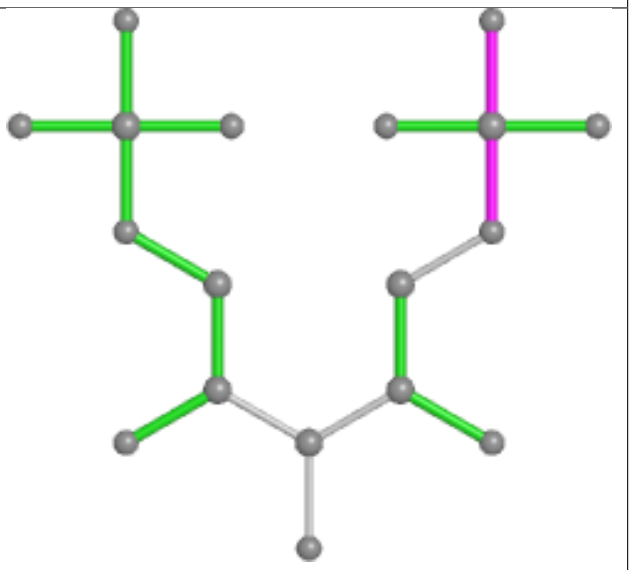
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	602	SO4	1	0
6	A	608	RUB	11	0
6	F	609	RUB	16	0
5	F	606	EDO	6	0
6	B	606	RUB	6	0
8	D	608	BCT	1	0
6	D	605	RUB	11	0
5	E	607	EDO	1	0
6	C	605	RUB	8	0
4	E	603	SO4	2	0
3	B	607	GOL	2	0
4	F	603	SO4	1	0

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.

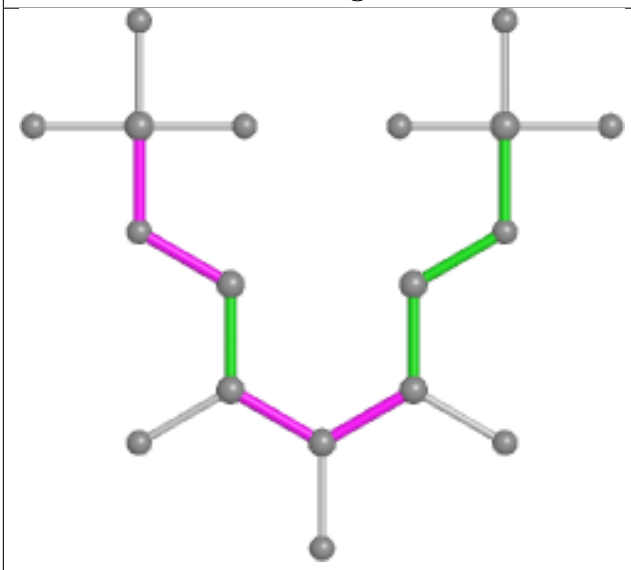
Ligand RUB E 610



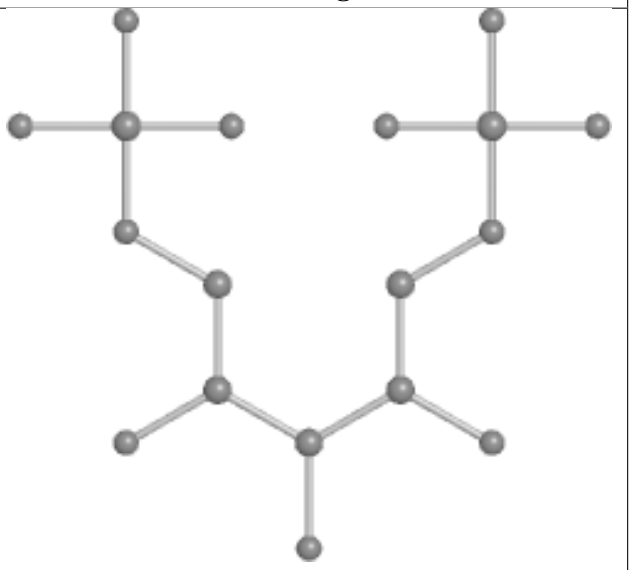
Bond lengths



Bond angles

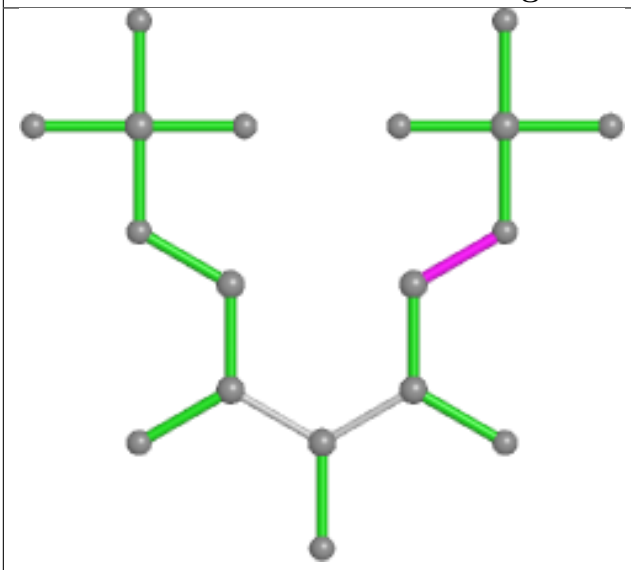


Torsions

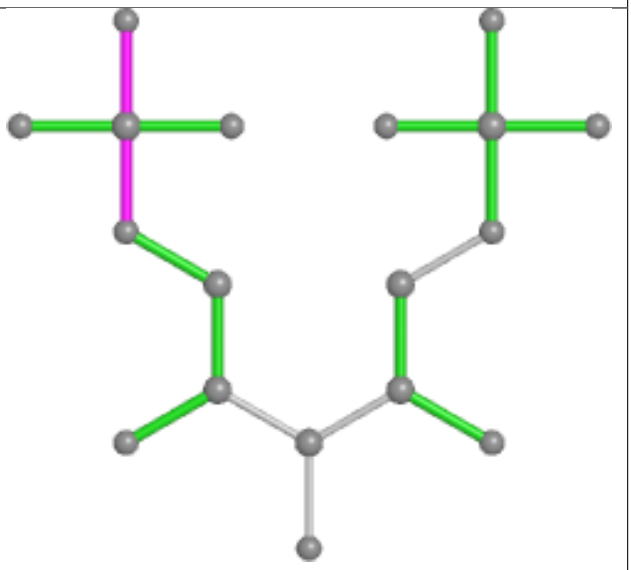


Rings

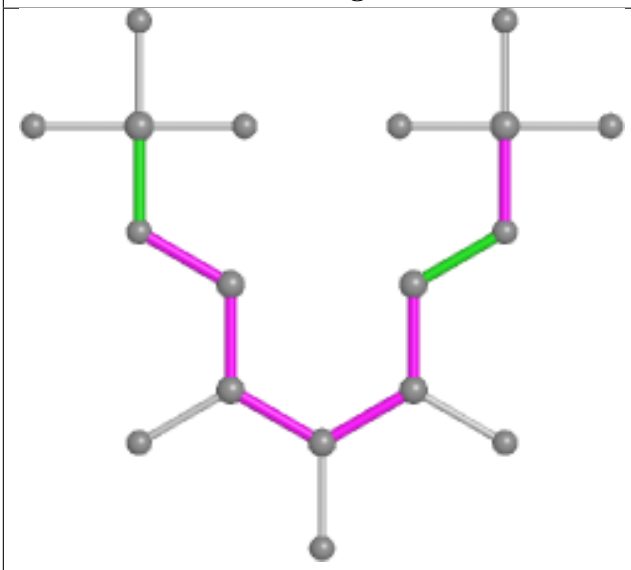
Ligand RUB A 608



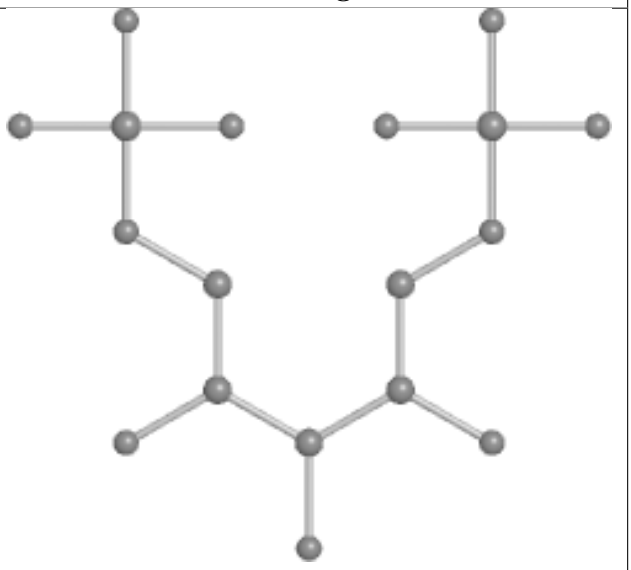
Bond lengths



Bond angles

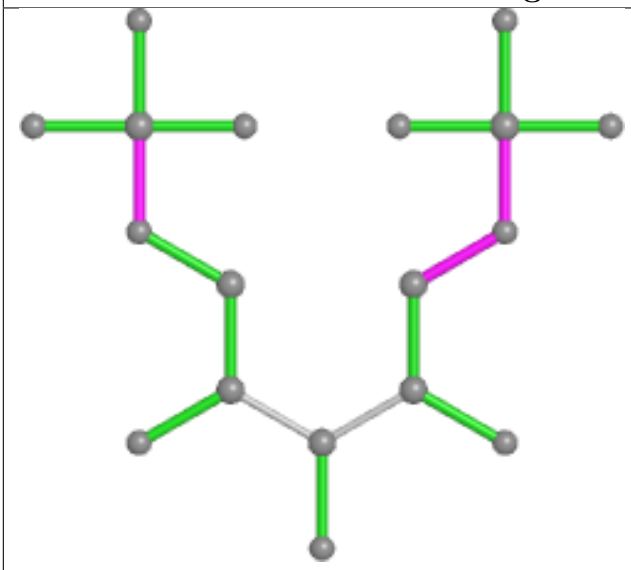


Torsions

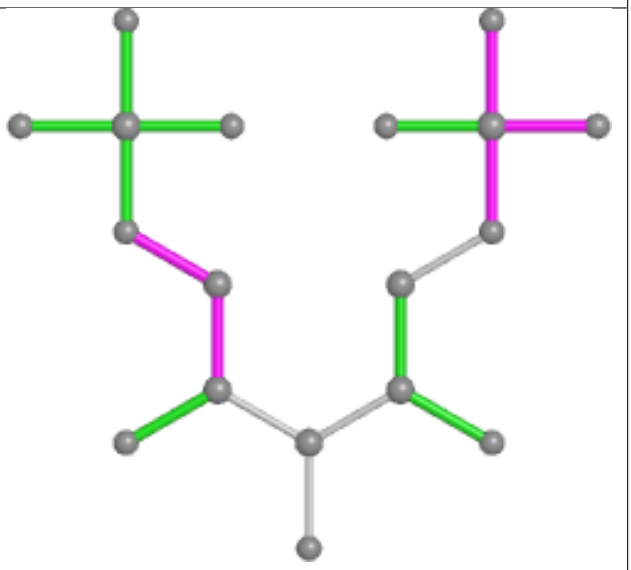


Rings

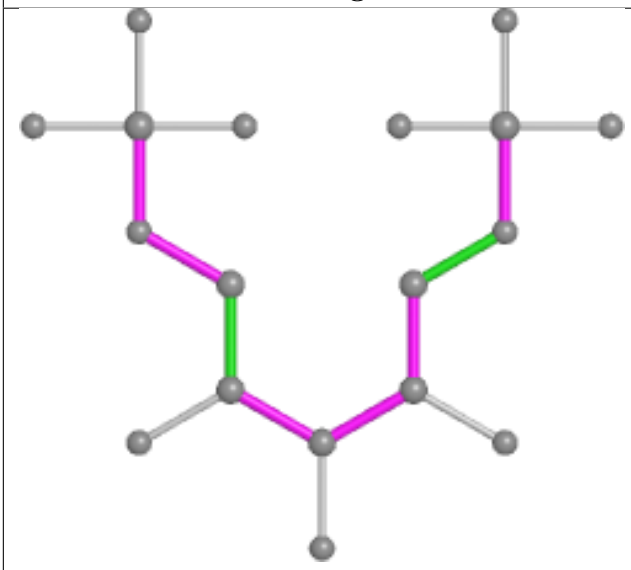
Ligand RUB F 609



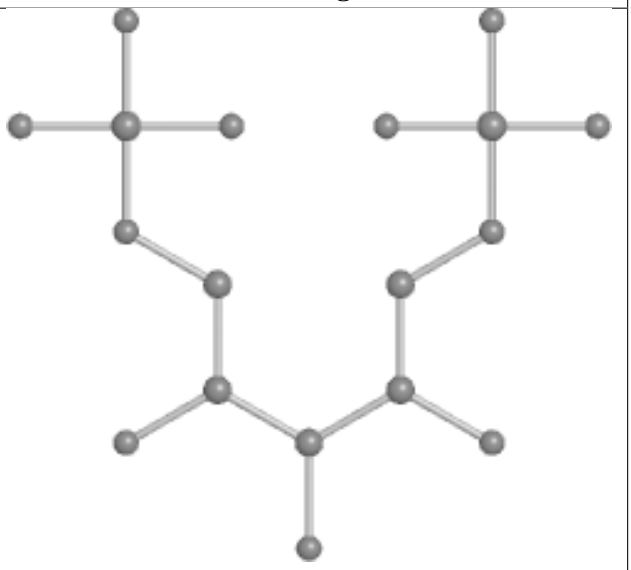
Bond lengths



Bond angles

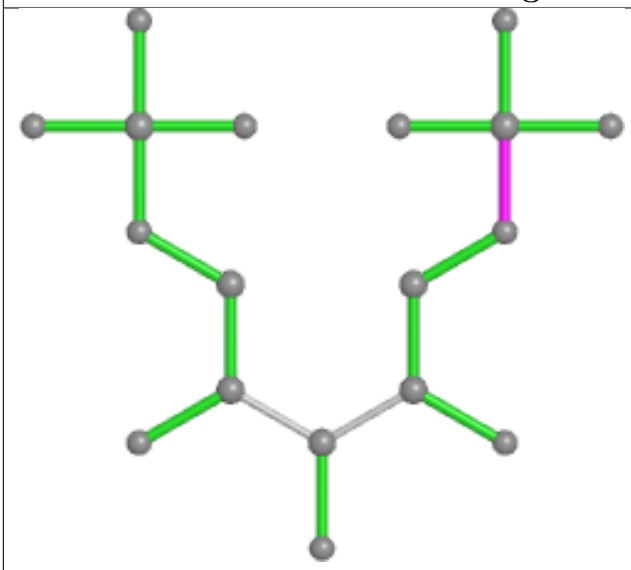


Torsions

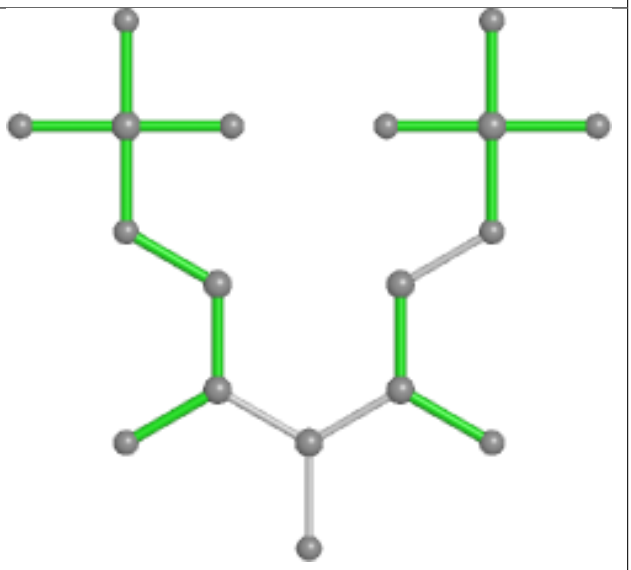


Rings

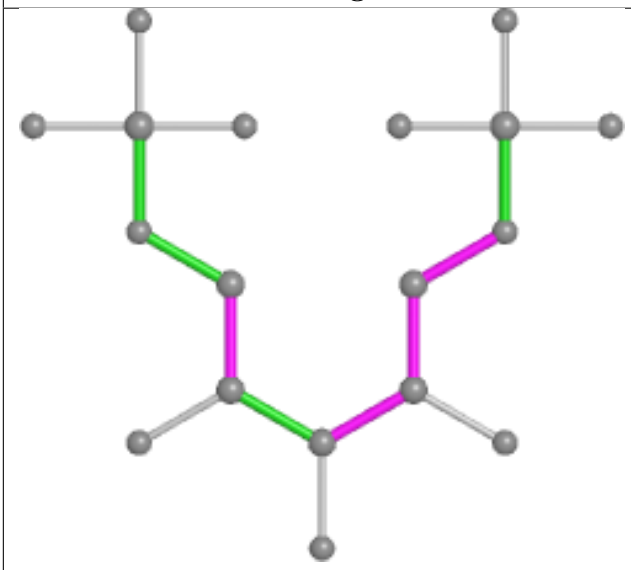
Ligand RUB B 606



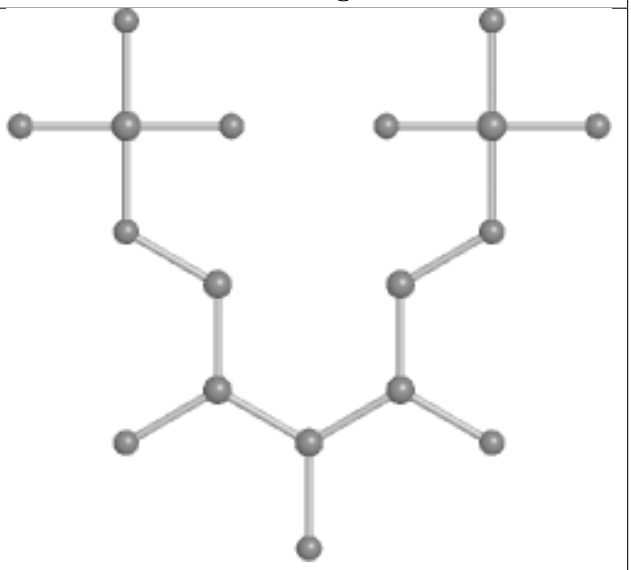
Bond lengths



Bond angles

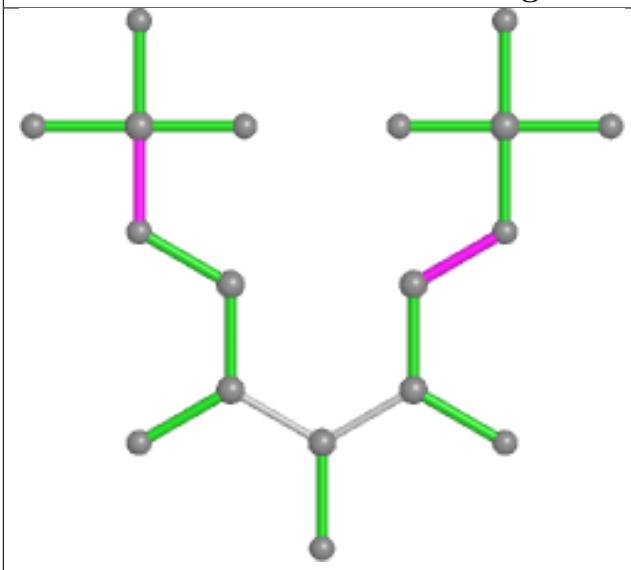


Torsions

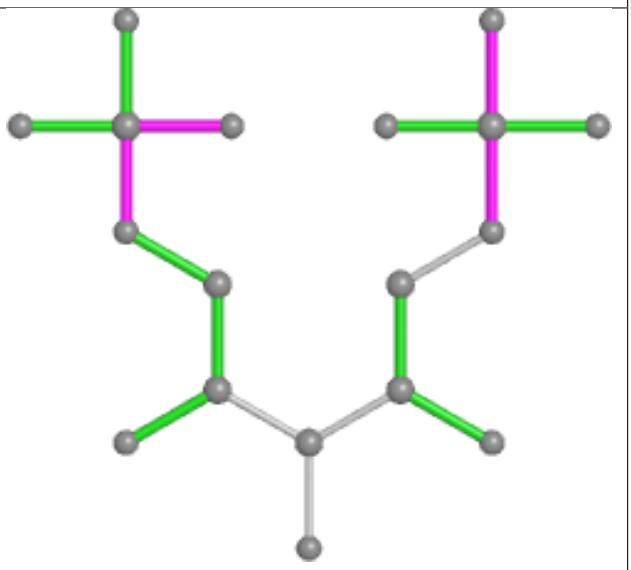


Rings

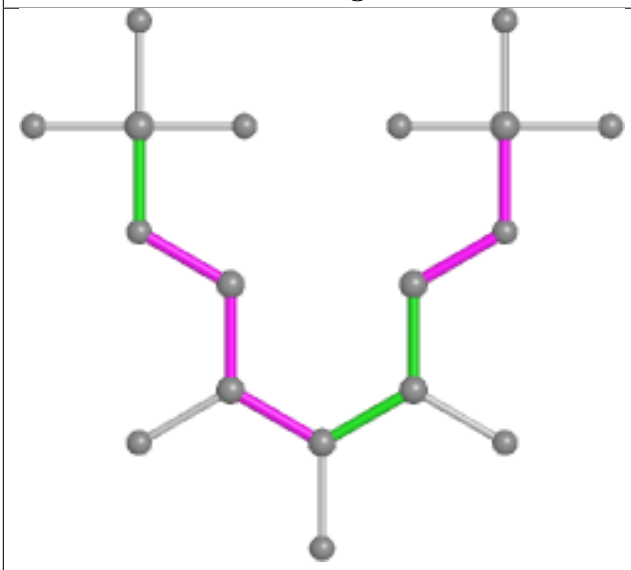
Ligand RUB D 605



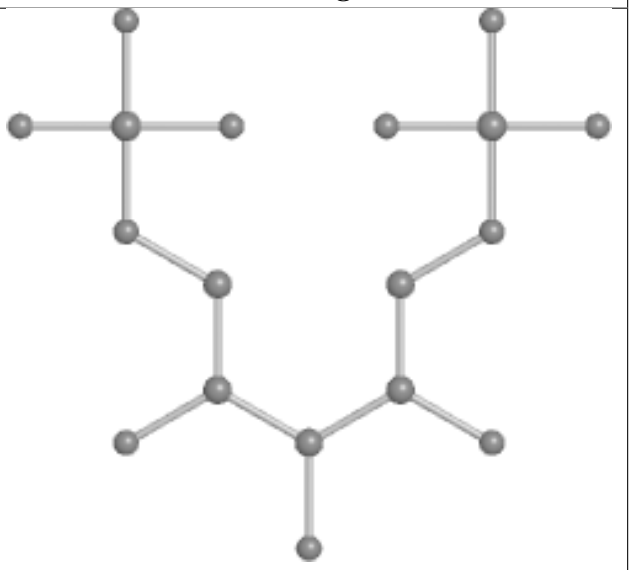
Bond lengths



Bond angles

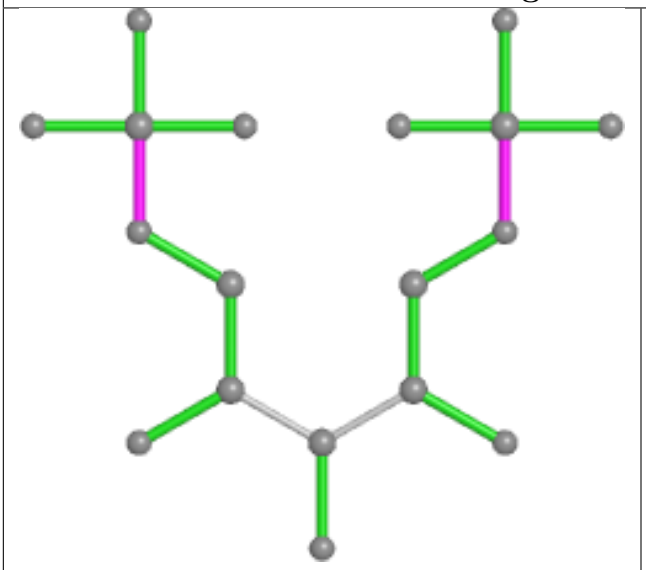


Torsions

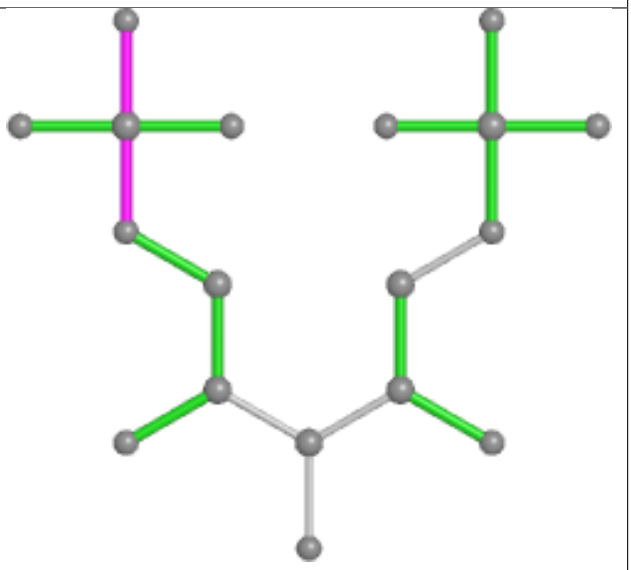


Rings

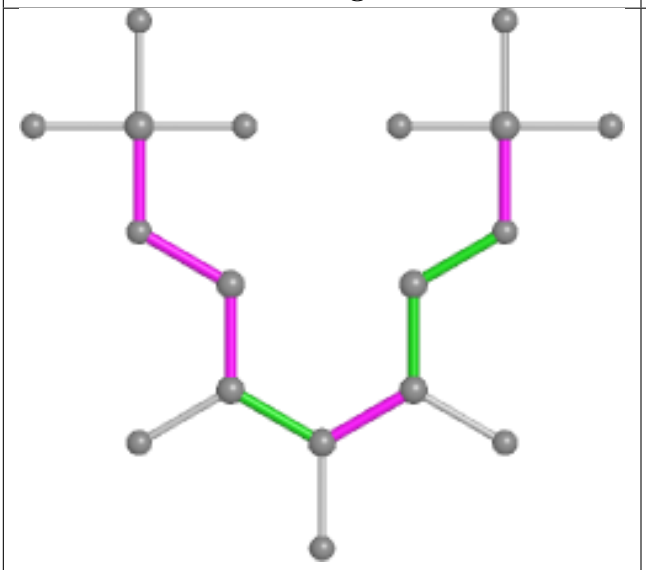
Ligand RUB C 605



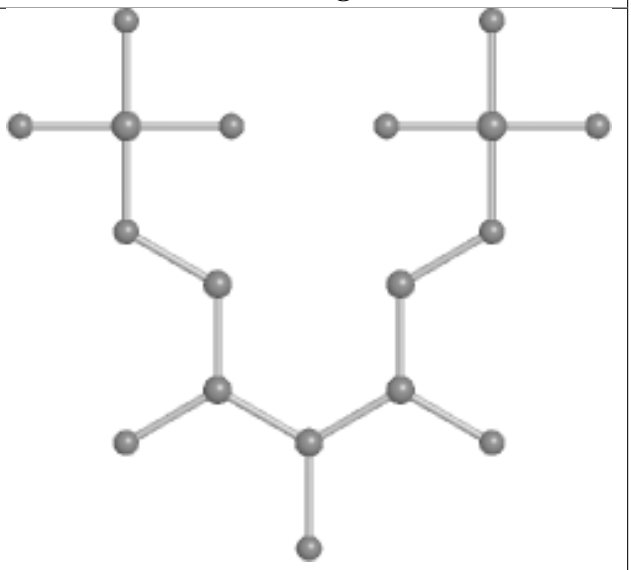
Bond lengths



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	461/462 (99%)	-0.29	1 (0%)	95	96	36, 47, 67, 96	1 (0%)
1	B	462/462 (100%)	-0.28	2 (0%)	92	95	36, 49, 69, 103	0
1	C	461/462 (99%)	-0.13	3 (0%)	87	91	40, 55, 74, 94	0
1	D	462/462 (100%)	-0.25	2 (0%)	92	95	36, 51, 74, 92	0
1	E	461/462 (99%)	-0.21	6 (1%)	77	81	38, 52, 73, 93	2 (0%)
1	F	462/462 (100%)	-0.31	2 (0%)	92	95	37, 48, 63, 86	0
All	All	2769/2772 (99%)	-0.24	16 (0%)	89	92	36, 50, 71, 103	3 (0%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	485	VAL	3.3
1	D	546	GLU	3.0
1	E	577	PHE	2.6
1	C	576	PHE	2.6
1	E	520	GLY	2.6

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 (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
6	RUB	C	605	18/18	0.54	0.49	47,63,75,89	26
6	RUB	F	609	18/18	0.57	0.49	39,54,70,76	26
6	RUB	A	608	18/18	0.58	0.44	44,61,75,77	26
6	RUB	B	606	18/18	0.61	0.43	37,66,82,97	26
6	RUB	E	610	18/18	0.70	0.43	47,66,74,80	26
6	RUB	D	605	18/18	0.70	0.41	43,58,64,70	26
5	EDO	A	607	4/4	0.76	0.21	70,71,73,74	0
3	GOL	B	607	6/6	0.78	0.24	70,85,100,101	0
3	GOL	D	606	6/6	0.82	0.26	64,72,74,82	0
5	EDO	E	609	4/4	0.82	0.30	62,81,89,97	0
3	GOL	A	603	6/6	0.83	0.12	58,61,64,64	0
5	EDO	F	606	4/4	0.83	0.32	59,77,85,88	0
8	BCT	D	608	4/4	0.85	0.16	35,43,55,58	5
7	CO2	C	606	3/3	0.86	0.29	56,56,62,67	0
4	SO4	E	603	5/5	0.86	0.15	71,77,87,95	5
7	CO2	D	607	3/3	0.88	0.11	55,55,70,71	0
4	SO4	D	604	5/5	0.89	0.26	54,62,63,66	5
5	EDO	C	604	4/4	0.89	0.29	51,79,83,99	0
5	EDO	F	607	4/4	0.89	0.09	63,65,69,70	0
5	EDO	F	608	4/4	0.90	0.18	59,60,71,72	0
4	SO4	F	605	5/5	0.91	0.11	73,77,94,98	0
5	EDO	B	605	4/4	0.91	0.29	47,52,58,71	0
5	EDO	E	607	4/4	0.92	0.11	68,71,74,80	0
4	SO4	F	604	5/5	0.92	0.23	39,46,52,70	5
5	EDO	E	608	4/4	0.93	0.52	55,72,86,86	0
5	EDO	E	606	4/4	0.93	0.19	54,61,64,70	0
4	SO4	E	604	5/5	0.95	0.20	50,52,59,70	5
4	SO4	C	602	5/5	0.95	0.17	62,70,74,75	0
4	SO4	A	606	5/5	0.96	0.17	41,48,57,67	5
2	ZN	B	601	1/1	0.98	0.10	51,51,51,51	0
4	SO4	B	603	5/5	0.98	0.12	51,58,66,71	0
4	SO4	E	605	5/5	0.98	0.13	63,66,76,79	0
2	ZN	C	601	1/1	0.99	0.13	55,55,55,55	0
4	SO4	F	602	5/5	0.99	0.17	47,48,50,57	0
4	SO4	F	603	5/5	0.99	0.17	45,60,67,69	0
4	SO4	B	604	5/5	0.99	0.08	44,45,54,55	0
2	ZN	E	601	1/1	0.99	0.13	46,46,46,46	0

Continued on next page...

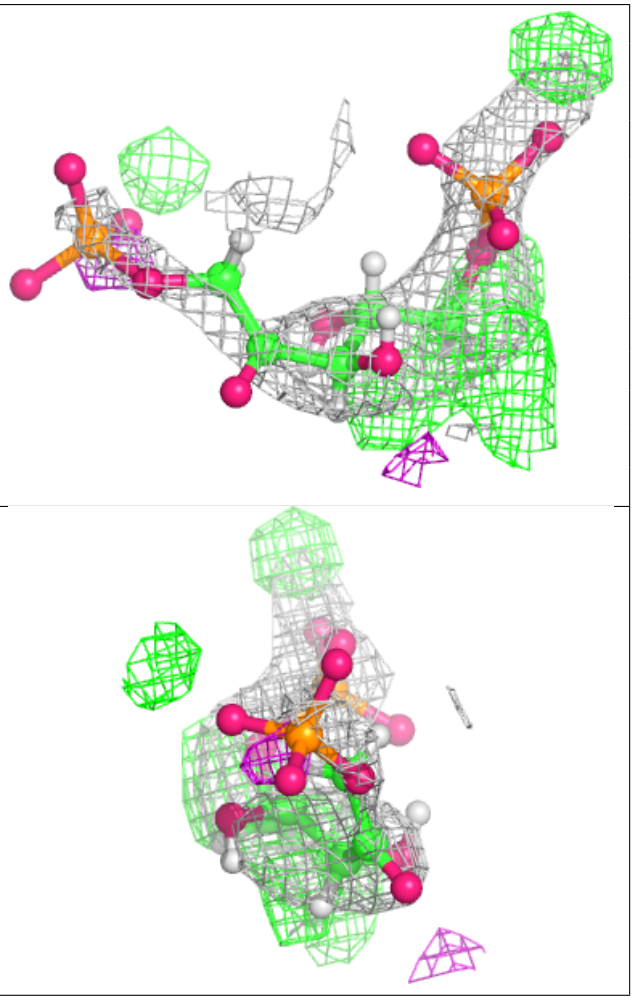
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	C	603	5/5	0.99	0.19	45,61,62,65	0
4	SO4	D	602	5/5	0.99	0.22	58,68,69,71	0
4	SO4	D	603	5/5	0.99	0.17	43,49,57,60	0
4	SO4	A	604	5/5	0.99	0.14	43,45,50,53	0
4	SO4	E	602	5/5	0.99	0.14	46,50,54,56	0
4	SO4	A	605	5/5	0.99	0.11	52,55,58,78	0
2	ZN	B	602	1/1	0.99	0.13	46,46,46,46	0
2	ZN	A	601	1/1	1.00	0.15	41,41,41,41	0
2	ZN	D	601	1/1	1.00	0.11	47,47,47,47	0
2	ZN	A	602	1/1	1.00	0.15	43,43,43,43	0
2	ZN	F	601	1/1	1.00	0.14	45,45,45,45	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.

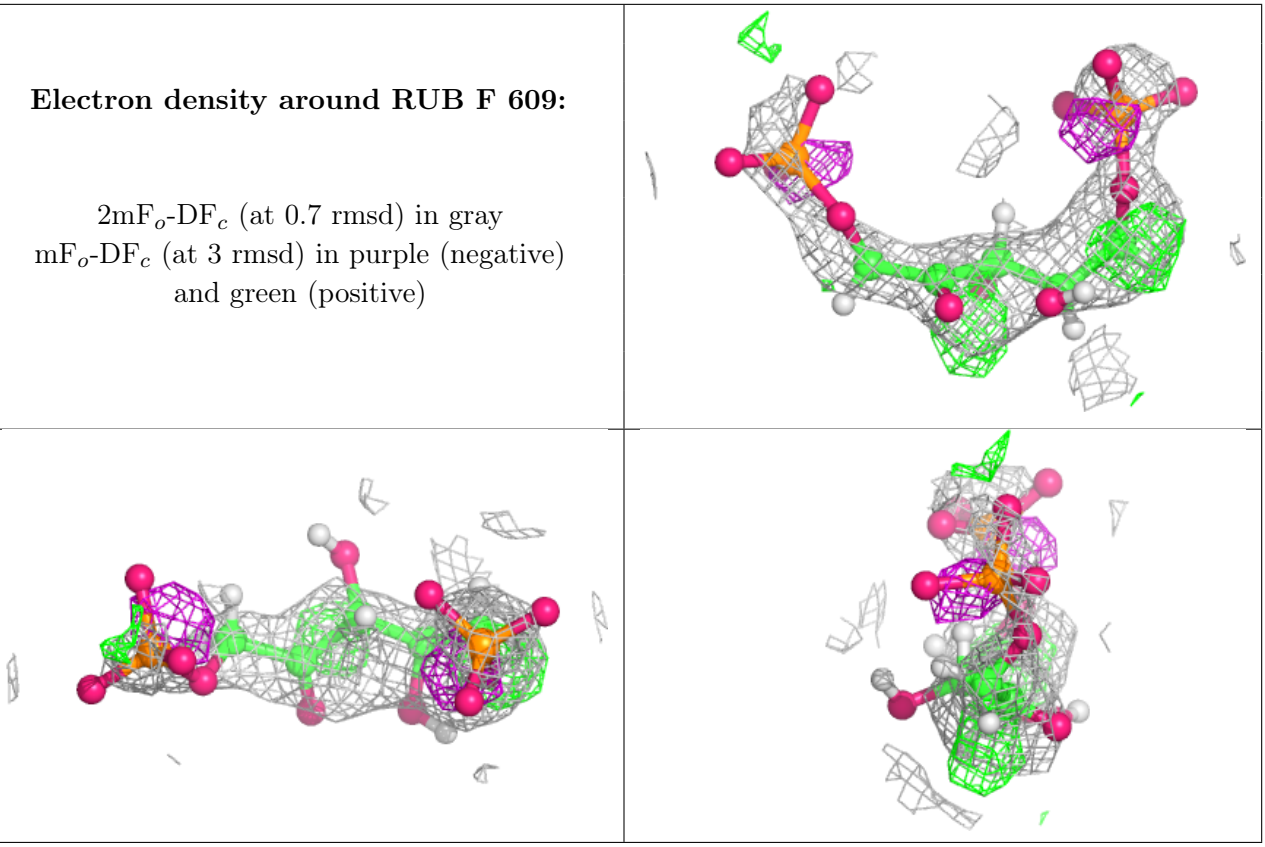
Electron density around RUB C 605:

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



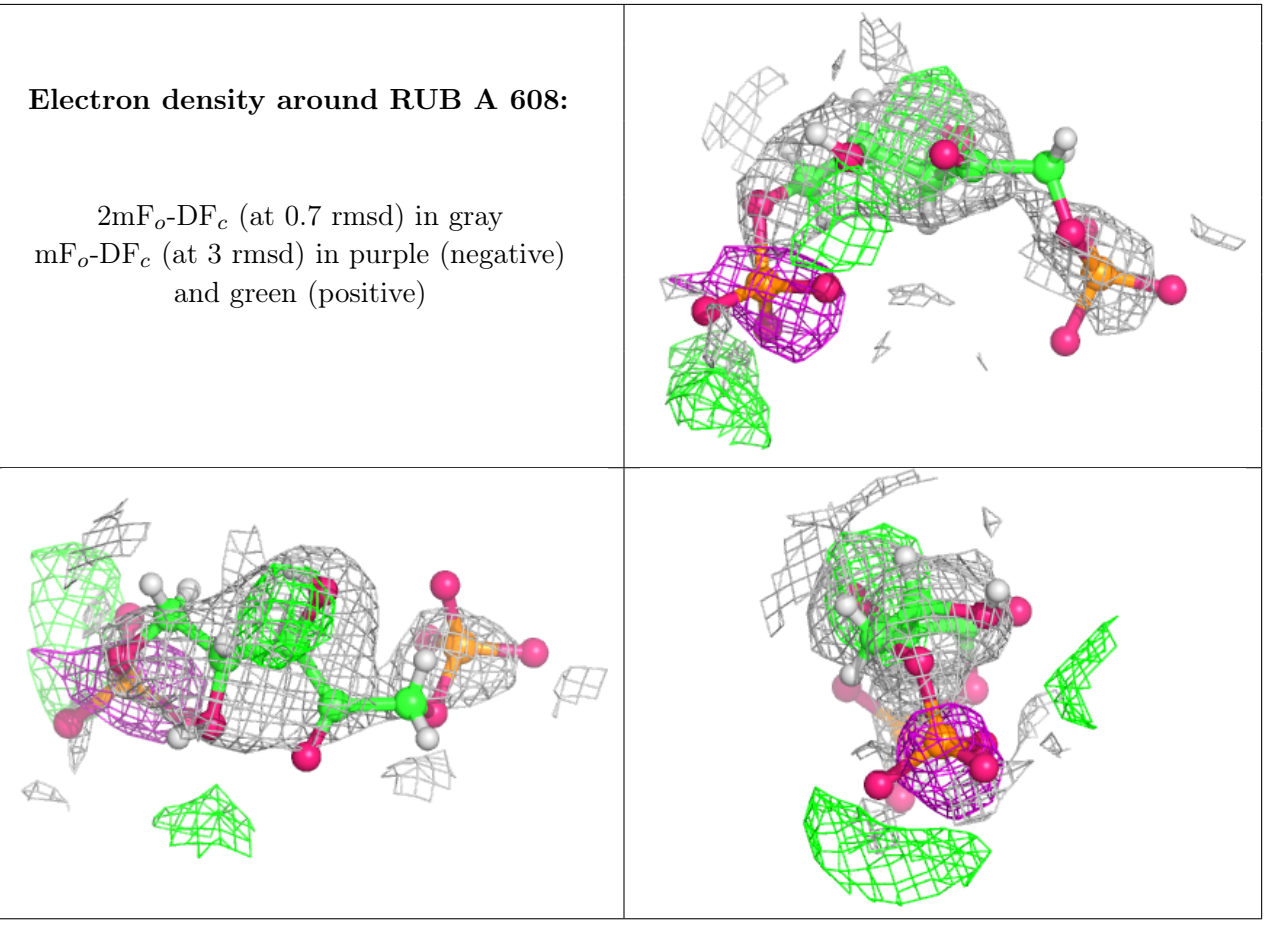
Electron density around RUB F 609:

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



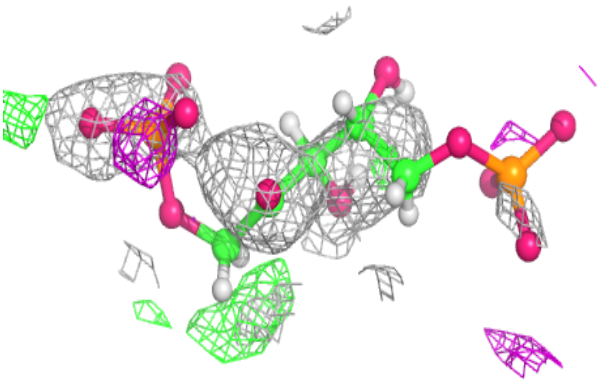
Electron density around RUB A 608:

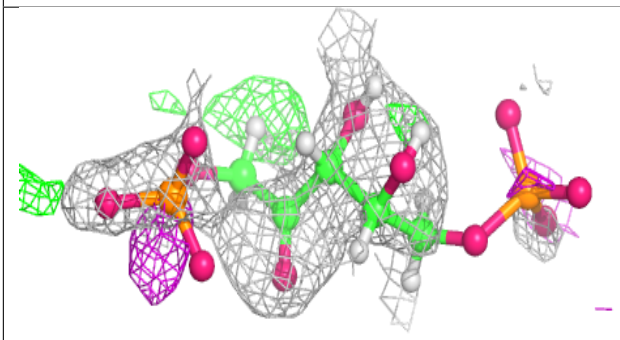
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

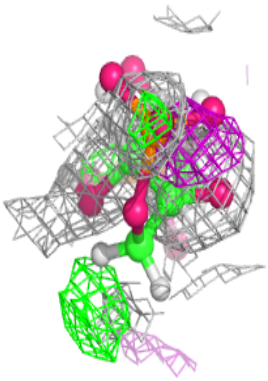


Electron density around RUB B 606:

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

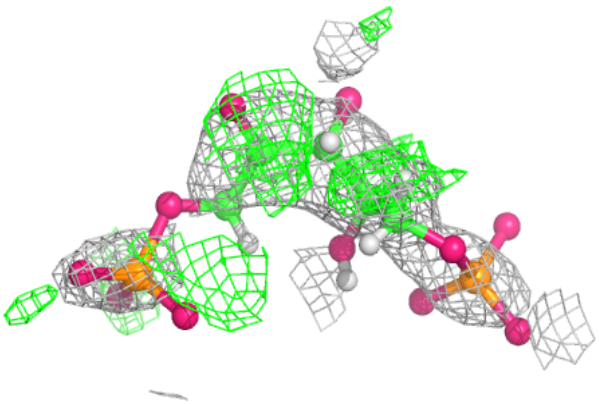


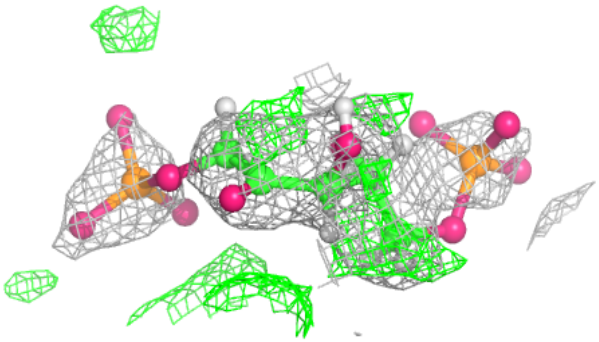


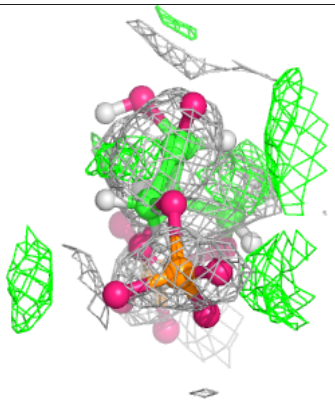


Electron density around RUB E 610:

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

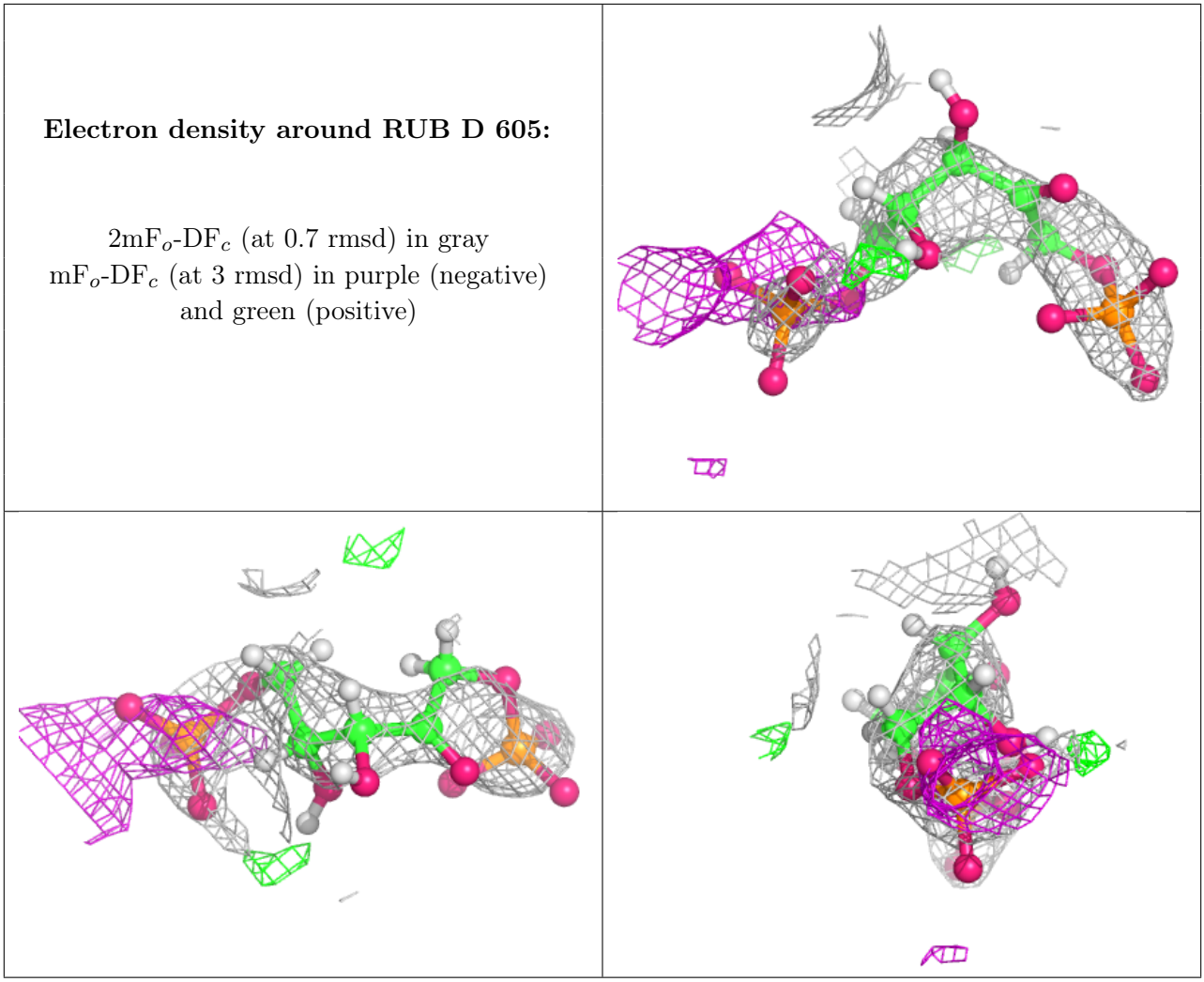






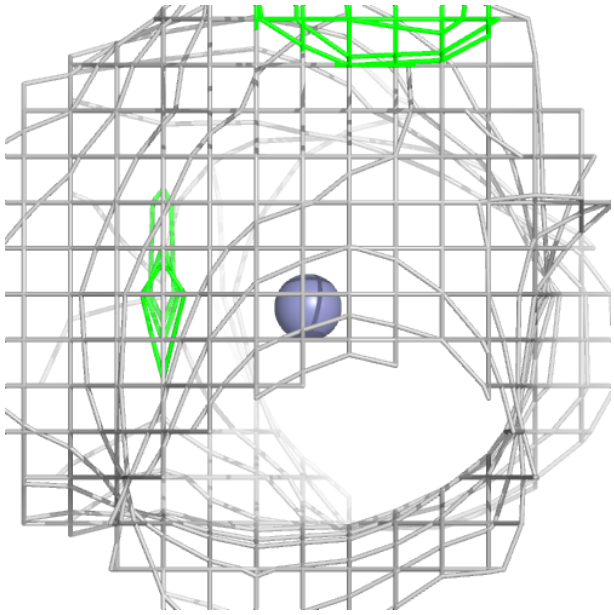
Electron density around RUB D 605:

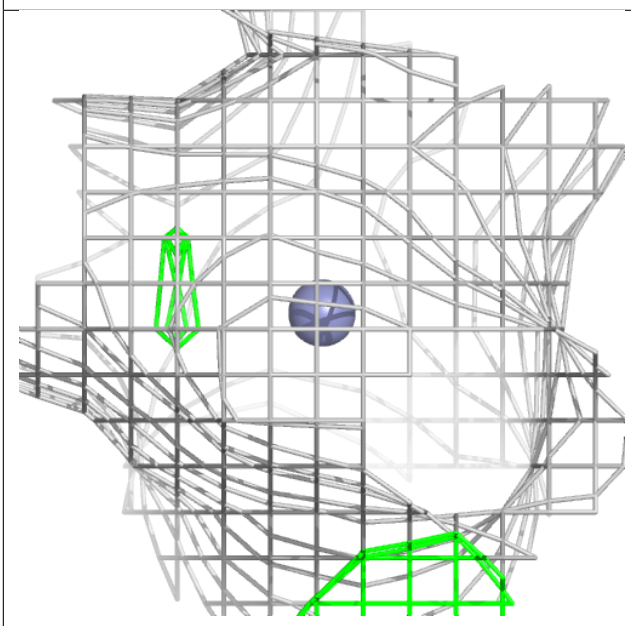
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

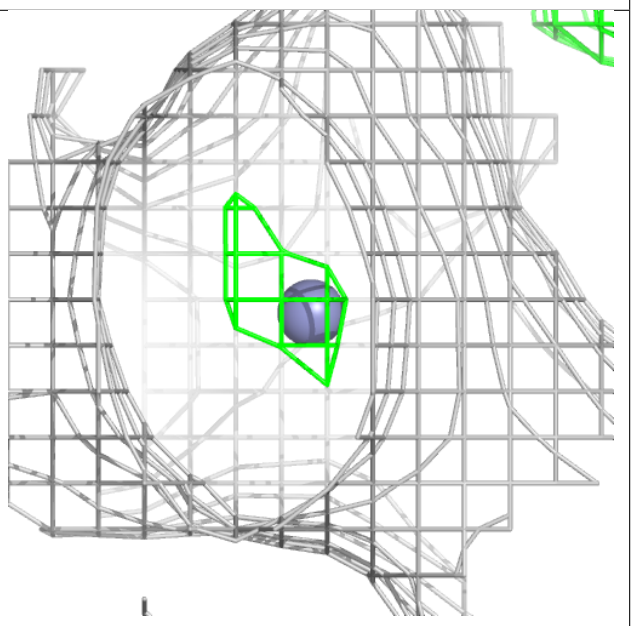


Electron density around ZN B 601:

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

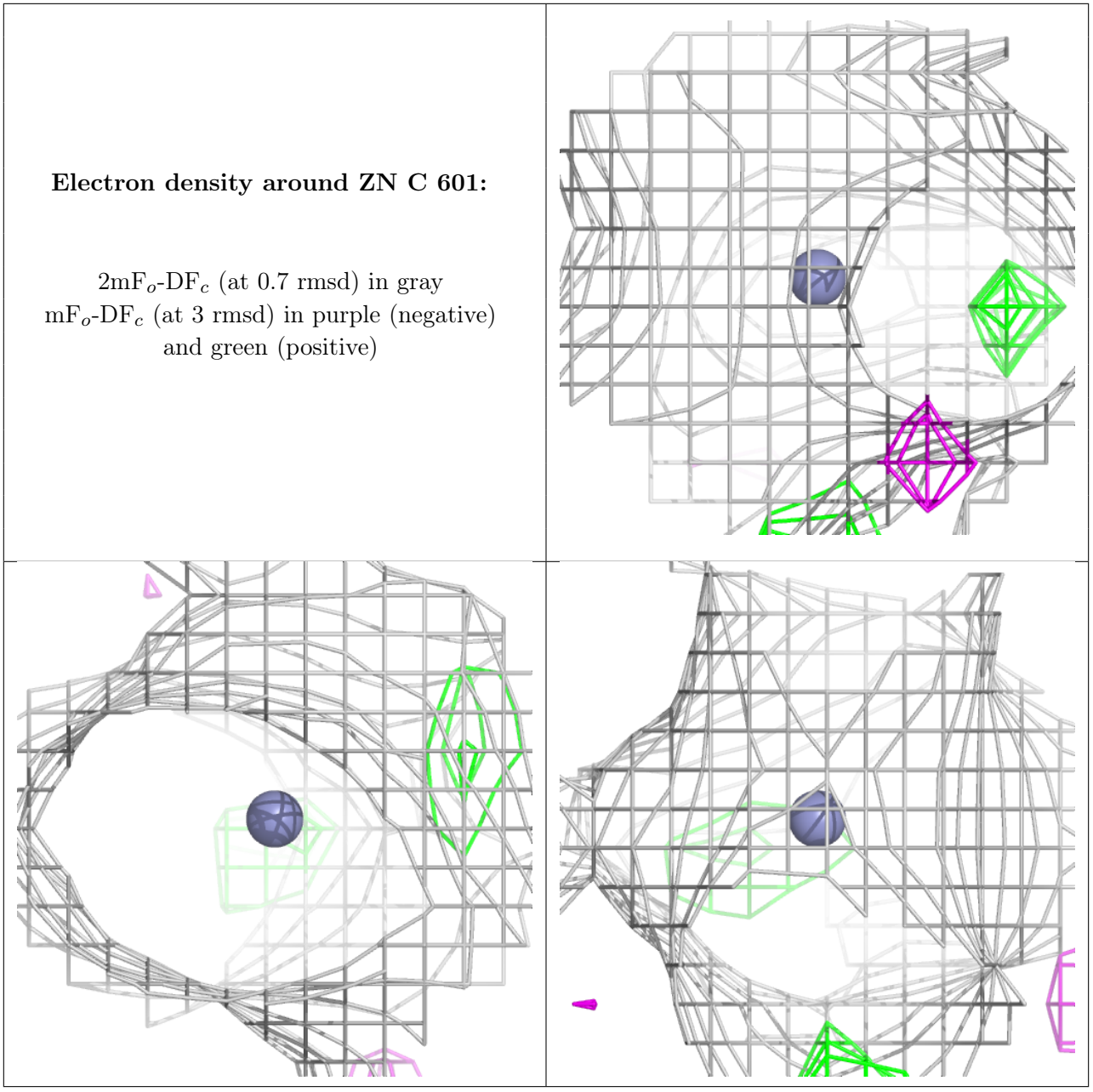






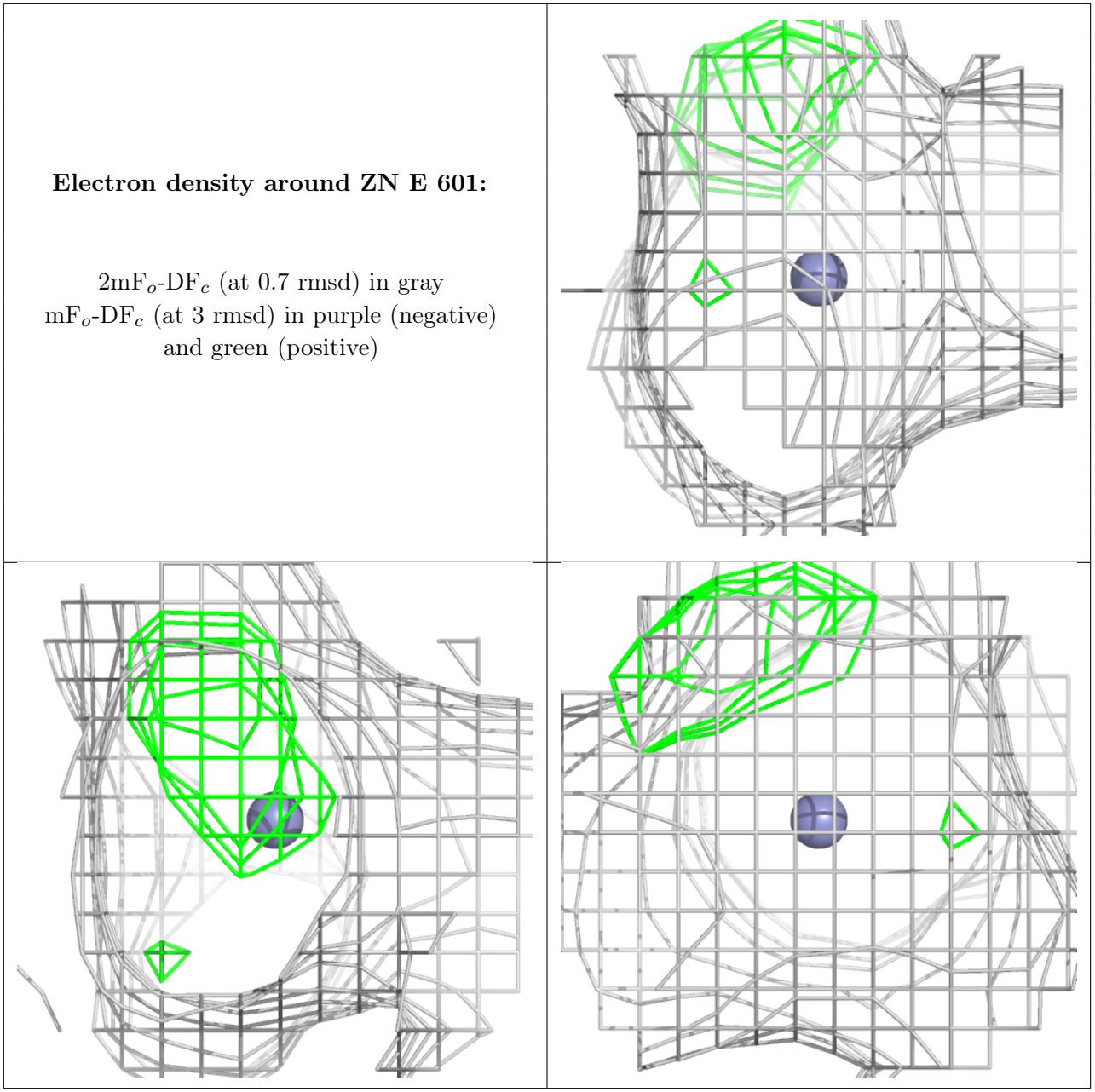
Electron density around ZN C 601:

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



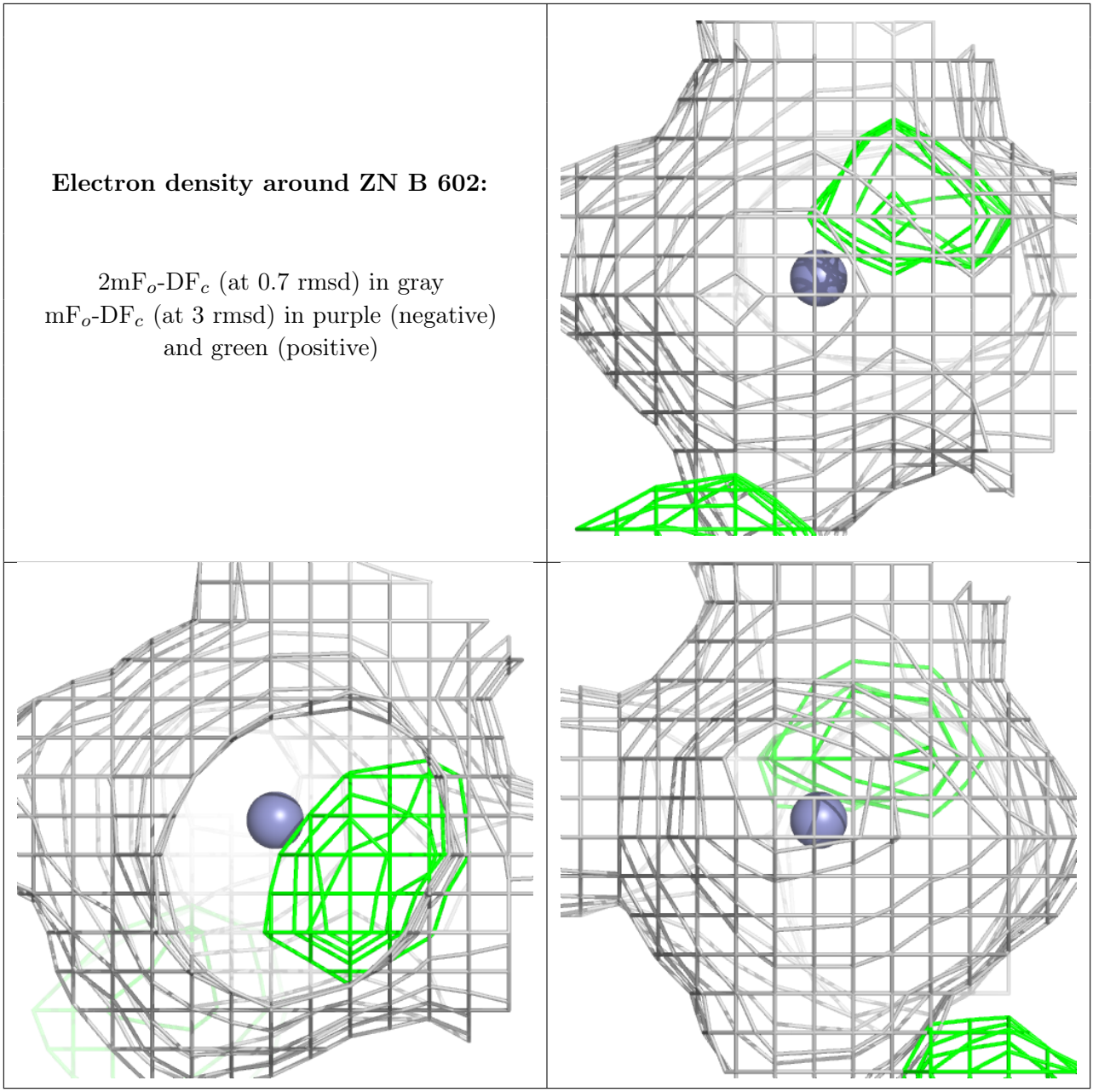
Electron density around ZN E 601:

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



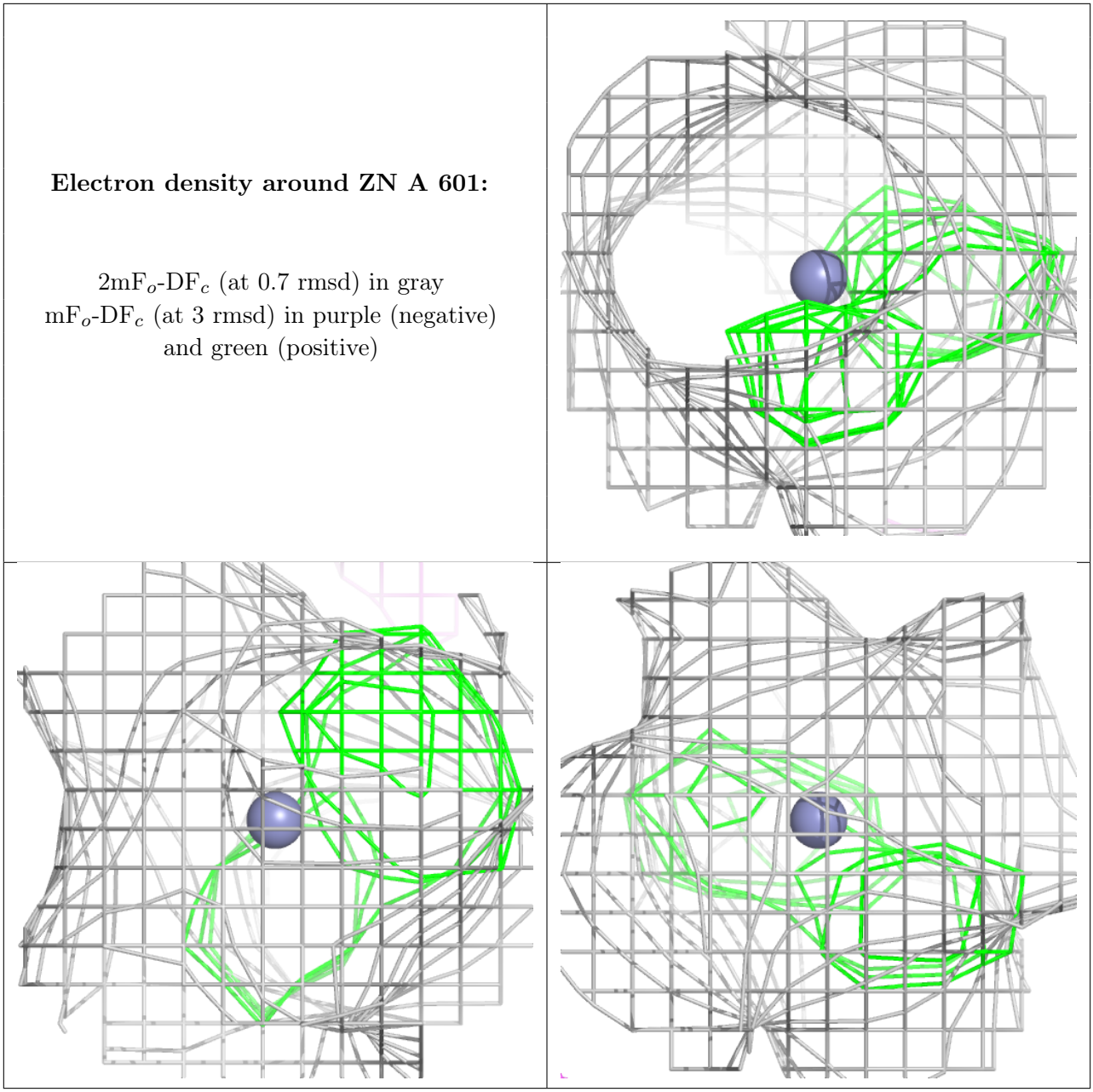
Electron density around ZN B 602:

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



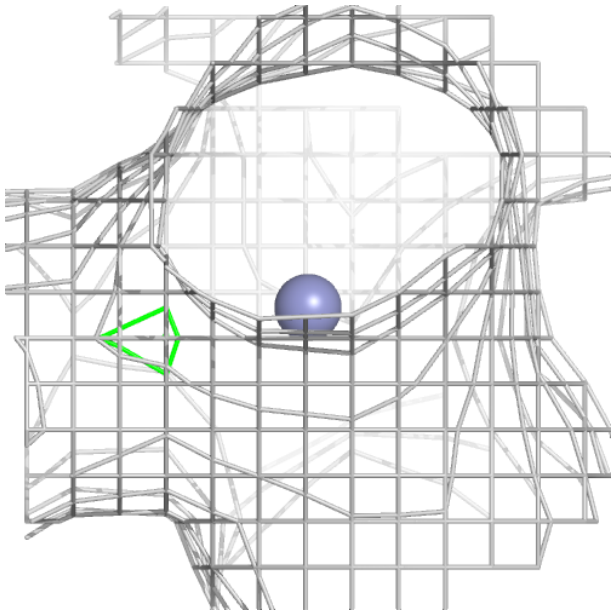
Electron density around ZN A 601:

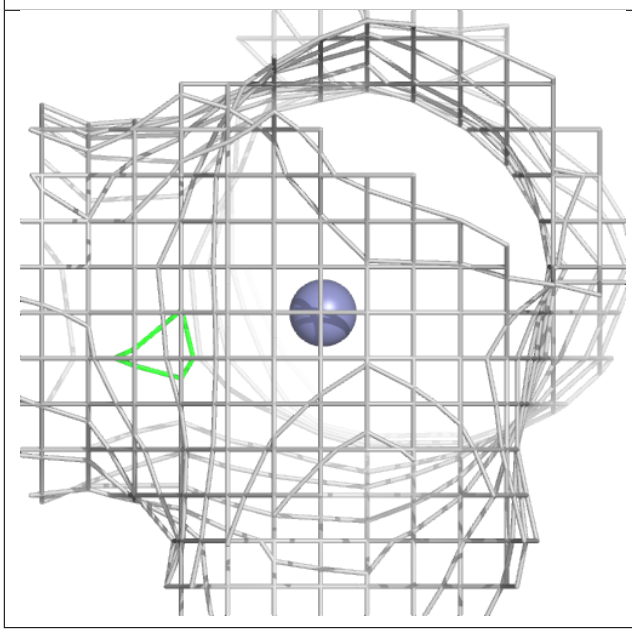
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

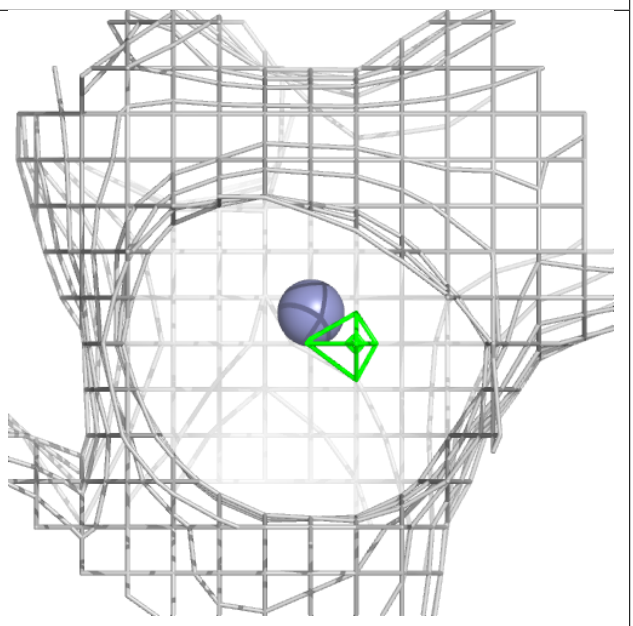


Electron density around ZN D 601:

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

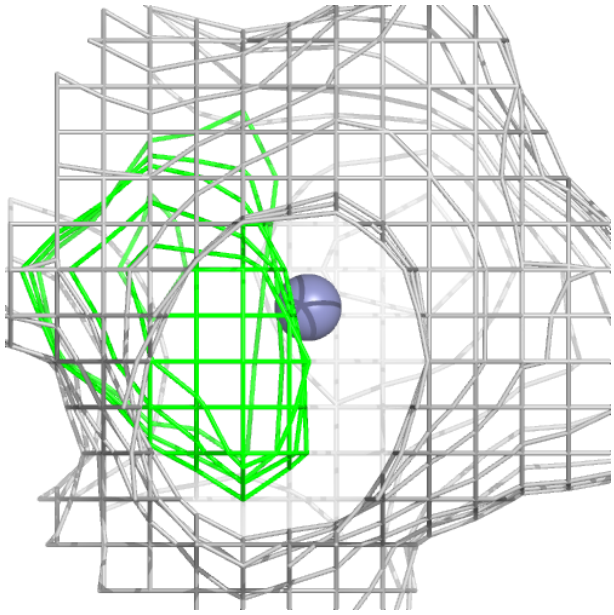


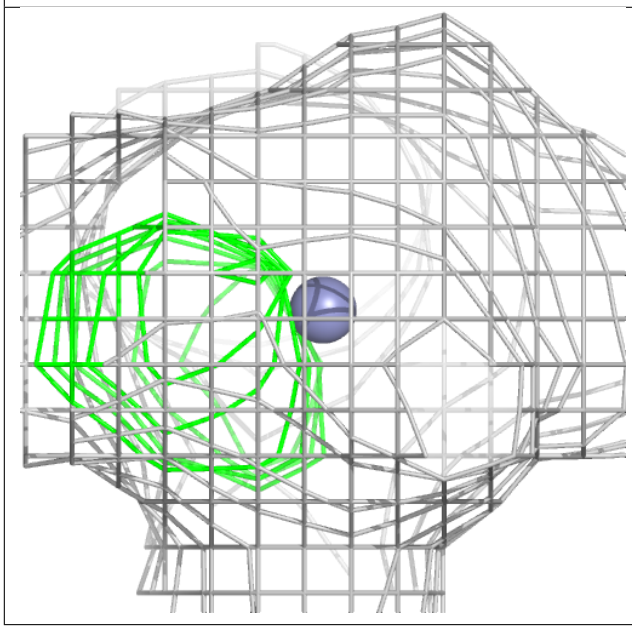


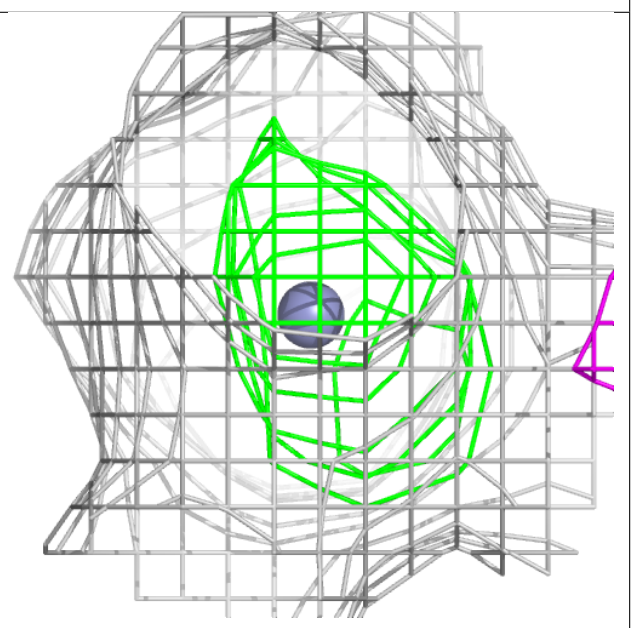


Electron density around ZN A 602:

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

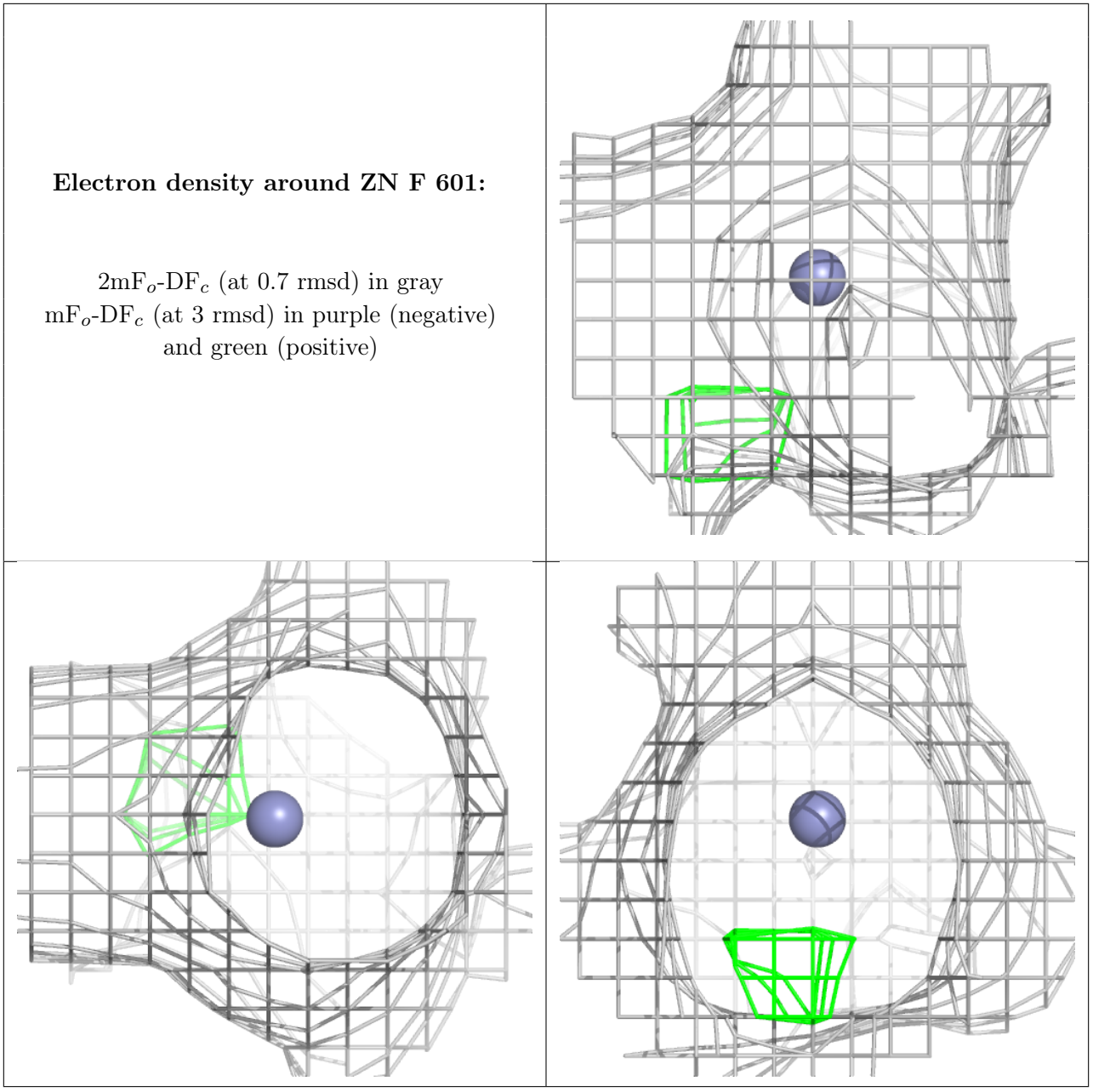






Electron density around ZN F 601:

$2mF_o$-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.