wwPDB X-ray Structure Validation Summary Report (i)

Apr 19, 2022 – 01:02 pm BST

PDB ID	:	7OD9
Title	:	Crystal structure of activated CheY fused to the C-terminal domain of CheF
Authors	:	Altegoer, F.; Weiland, P.; Bange, G.
Deposited on	:	2021-04-29
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 following versions of software and data (see references (i)) were used in the production of this report:

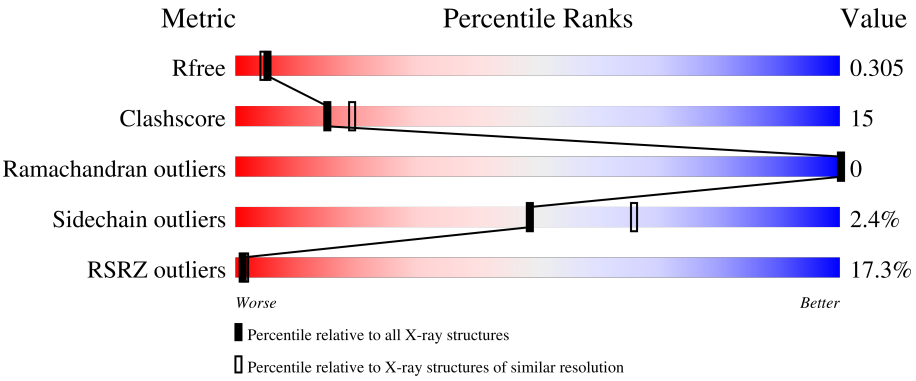
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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	142	10% 56% 27% · 14%
1	B	142	10% 73% 17% · 9%
2	C	115	17% 48% 30% · 22%
2	F	115	22% 46% 24% 30%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3374 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 Response regulator receiver protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	129	Total	C	N	O	S	0	0	0
			994	635	162	189	8			
1	A	122	Total	C	N	O	S	0	0	0
			935	596	150	181	8			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	MET	-	initiating methionine	UNP G0H061
B	-4	GLY	-	expression tag	UNP G0H061
B	-3	HIS	-	expression tag	UNP G0H061
B	-2	HIS	-	expression tag	UNP G0H061
B	-1	HIS	-	expression tag	UNP G0H061
B	0	HIS	-	expression tag	UNP G0H061
B	1	HIS	-	expression tag	UNP G0H061
B	2	HIS	-	expression tag	UNP G0H061
B	124	VAL	-	expression tag	UNP G0H061
B	125	LEU	-	expression tag	UNP G0H061
B	126	PHE	-	expression tag	UNP G0H061
B	127	GLN	-	expression tag	UNP G0H061
B	128	GLY	-	expression tag	UNP G0H061
B	129	PRO	-	expression tag	UNP G0H061
B	130	SER	-	expression tag	UNP G0H061
B	131	ALA	-	expression tag	UNP G0H061
B	132	GLY	-	expression tag	UNP G0H061
B	133	LEU	-	expression tag	UNP G0H061
B	134	VAL	-	expression tag	UNP G0H061
B	135	PRO	-	expression tag	UNP G0H061
B	136	ARG	-	expression tag	UNP G0H061
A	-5	MET	-	initiating methionine	UNP G0H061
A	-4	GLY	-	expression tag	UNP G0H061
A	-3	HIS	-	expression tag	UNP G0H061
A	-2	HIS	-	expression tag	UNP G0H061

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	HIS	-	expression tag	UNP G0H061
A	0	HIS	-	expression tag	UNP G0H061
A	1	HIS	-	expression tag	UNP G0H061
A	2	HIS	-	expression tag	UNP G0H061
A	124	VAL	-	expression tag	UNP G0H061
A	125	LEU	-	expression tag	UNP G0H061
A	126	PHE	-	expression tag	UNP G0H061
A	127	GLN	-	expression tag	UNP G0H061
A	128	GLY	-	expression tag	UNP G0H061
A	129	PRO	-	expression tag	UNP G0H061
A	130	SER	-	expression tag	UNP G0H061
A	131	ALA	-	expression tag	UNP G0H061
A	132	GLY	-	expression tag	UNP G0H061
A	133	LEU	-	expression tag	UNP G0H061
A	134	VAL	-	expression tag	UNP G0H061
A	135	PRO	-	expression tag	UNP G0H061
A	136	ARG	-	expression tag	UNP G0H061

- Molecule 2 is a protein called C-terminal domain of CheF from Methanococcus maripaludis.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	90	Total	C	N	O	S	0	0	0
			711	449	112	146	4			
2	F	81	Total	C	N	O	S	0	0	0
			641	406	102	130	3			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	234	GLY	-	expression tag	UNP G0H062
C	235	SER	-	expression tag	UNP G0H062
C	236	GLY	-	expression tag	UNP G0H062
C	237	GLY	-	expression tag	UNP G0H062
C	238	ILE	-	expression tag	UNP G0H062
C	239	GLU	-	expression tag	UNP G0H062
C	240	GLY	-	expression tag	UNP G0H062
C	241	GLY	-	expression tag	UNP G0H062
C	242	SER	-	expression tag	UNP G0H062
C	243	MET	-	expression tag	UNP G0H062
C	244	GLY	-	expression tag	UNP G0H062
F	234	GLY	-	expression tag	UNP G0H062
F	235	SER	-	expression tag	UNP G0H062

Continued on next page...

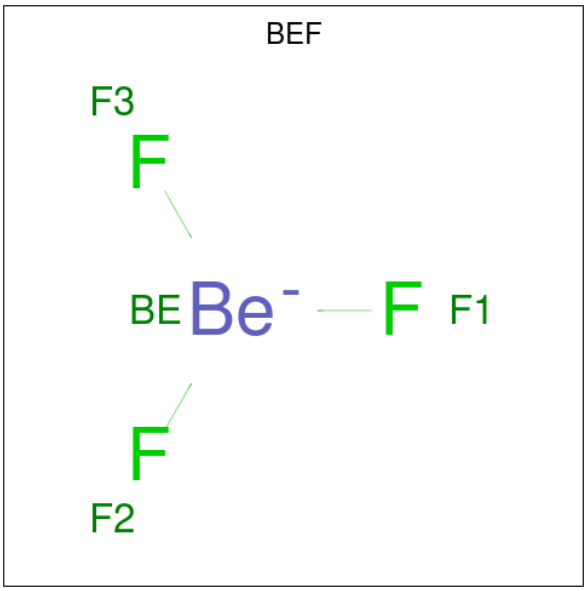
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	236	GLY	-	expression tag	UNP G0H062
F	237	GLY	-	expression tag	UNP G0H062
F	238	ILE	-	expression tag	UNP G0H062
F	239	GLU	-	expression tag	UNP G0H062
F	240	GLY	-	expression tag	UNP G0H062
F	241	GLY	-	expression tag	UNP G0H062
F	242	SER	-	expression tag	UNP G0H062
F	243	MET	-	expression tag	UNP G0H062
F	244	GLY	-	expression tag	UNP G0H062

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 4 is BERYLLIUM TRIFLUORIDE ION (three-letter code: BEF) (formula: BeF_3) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Be F 4 1 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	Be	F	0	0
			4	1	3		

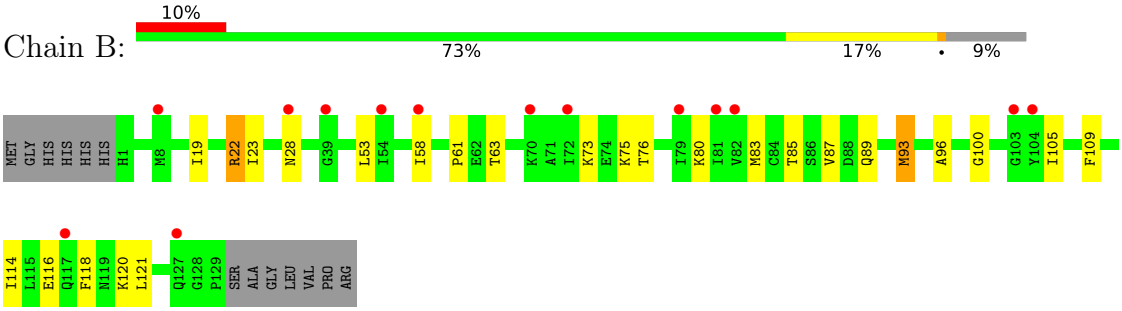
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	43	Total	O	0	0
			43	43		
5	A	27	Total	O	0	0
			27	27		
5	C	5	Total	O	0	0
			5	5		
5	F	8	Total	O	0	0
			8	8		

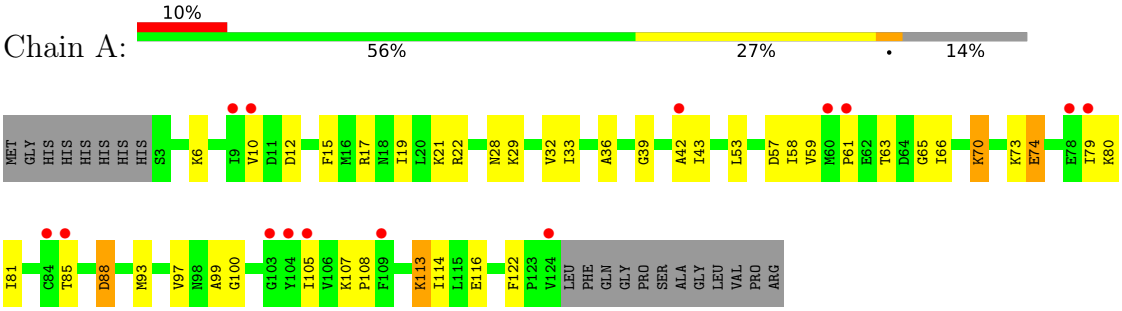
3 Residue-property plots (i)

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.

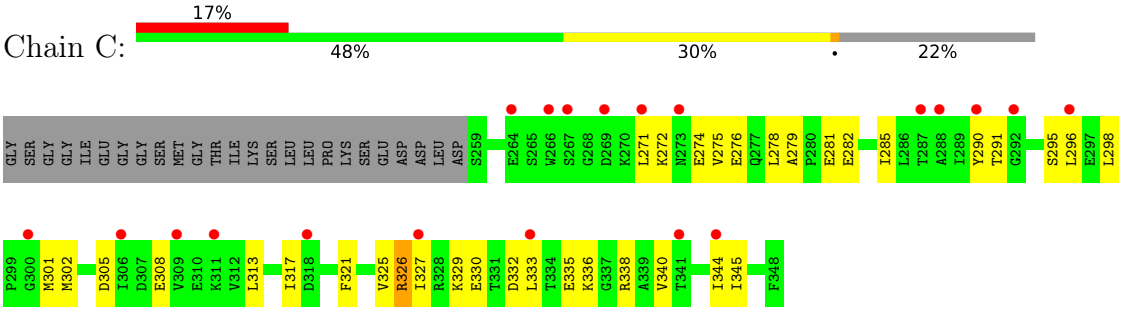
• Molecule 1: Response regulator receiver protein



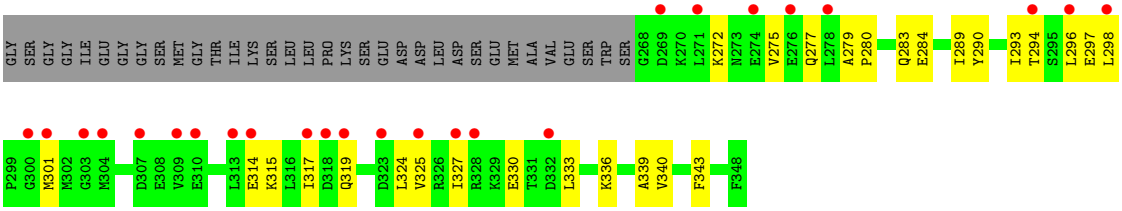
• Molecule 1: Response regulator receiver protein



• Molecule 2: C-terminal domain of CheF from Methanococcus maripaludis



• Molecule 2: C-terminal domain of CheF from Methanococcus maripaludis


GLY SER GLY GLY ILE GLU GLY GLY SER MET GLY THR ILE LYS SER LEU LEU PRO LYS SER GLU ASP ASP LEU ASP SER GLU MET ALA VAL GLU SER TRP SER G268 D269 K270 L271 K272 N273 E274 V275 E276 Q277 L278 A279 P280 Q283 E284 I289 Y290 I293 T294 S295 L296 E297 L298
P299 G300 M301 M302 G303 M304 D307 E308 V309 E310 L313 E314 K315 L316 I317 D318 Q319 D323 L324 V325 R326 I327 R328 K329 E330 T331 D332 L333 K336 A339 V340 F343 F348

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	56.29Å 70.87Å 59.82Å 90.00° 91.60° 90.00°	Depositor
Resolution (Å)	44.07 – 2.30 45.70 – 2.27	Depositor EDS
% Data completeness (in resolution range)	99.6 (44.07-2.30) 99.6 (45.70-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.98 (at 2.27Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.249 , 0.307 0.249 , 0.305	Depositor DCC
R_{free} test set	1086 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	48.6	Xtriage
Anisotropy	0.712	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3374	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.48% 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: MG, BEF

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	0.36	0/946	0.64	2/1277 (0.2%)
1	B	0.28	0/1009	0.50	0/1363
2	C	0.30	0/718	0.53	0/967
2	F	0.31	0/646	0.54	0/869
All	All	0.31	0/3319	0.56	2/4476 (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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	88	ASP	CB-CG-OD2	6.62	124.26	118.30
1	A	70	LYS	CD-CE-NZ	-6.06	97.77	111.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	74	GLU	Peptide

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	935	0	976	37	0
1	B	994	0	1031	17	0
2	C	711	0	715	39	0
2	F	641	0	655	31	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	4	0	0	1	0
4	B	4	0	0	0	0
5	A	27	0	0	5	0
5	B	43	0	0	1	0
5	C	5	0	0	0	0
5	F	8	0	0	1	0
All	All	3374	0	3377	100	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 15.

The worst 5 of 100 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:296:LEU:HB3	2:F:296:LEU:HD13	1.61	0.82
1:A:105:ILE:HD13	1:A:114:ILE:HD13	1.63	0.79
2:F:343:PHE:O	5:F:401:HOH:O	2.01	0.78
1:A:28:ASN:ND2	5:A:301:HOH:O	2.18	0.75
1:A:85:THR:O	1:A:107:LYS:N	2.22	0.72

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	120/142 (84%)	112 (93%)	8 (7%)	0	100	100
1	B	127/142 (89%)	125 (98%)	2 (2%)	0	100	100
2	C	88/115 (76%)	86 (98%)	2 (2%)	0	100	100
2	F	79/115 (69%)	77 (98%)	2 (2%)	0	100	100
All	All	414/514 (80%)	400 (97%)	14 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	105/121 (87%)	102 (97%)	3 (3%)	42	58
1	B	111/121 (92%)	108 (97%)	3 (3%)	44	61
2	C	80/99 (81%)	78 (98%)	2 (2%)	47	65
2	F	72/99 (73%)	71 (99%)	1 (1%)	67	81
All	All	368/440 (84%)	359 (98%)	9 (2%)	49	66

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	330	GLU
2	F	272	LYS
1	A	22	ARG
1	A	74	GLU
1	A	113	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	117	GLN
1	A	28	ASN
1	A	89	GLN
1	A	117	GLN
2	C	283	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	BEF	B	202	1	0,3,3	-	-	-		
4	BEF	A	202	-	0,3,3	-	-	-		

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

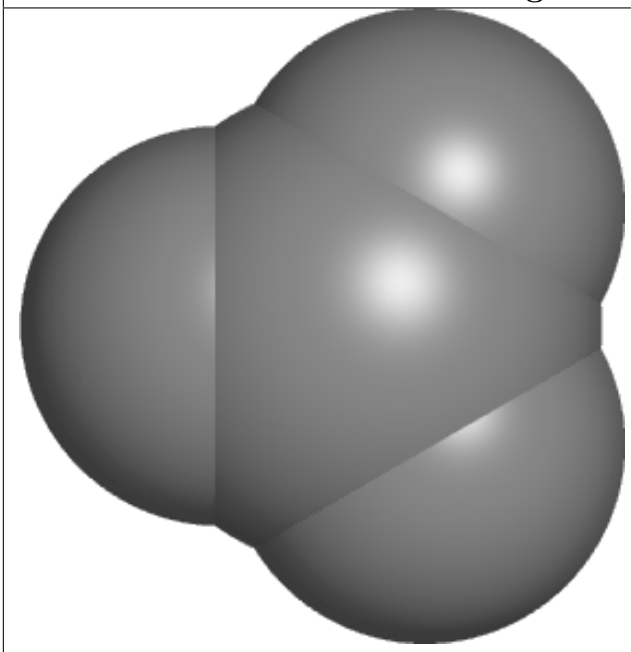
There are no ring outliers.

1 monomer is involved in 1 short contact:

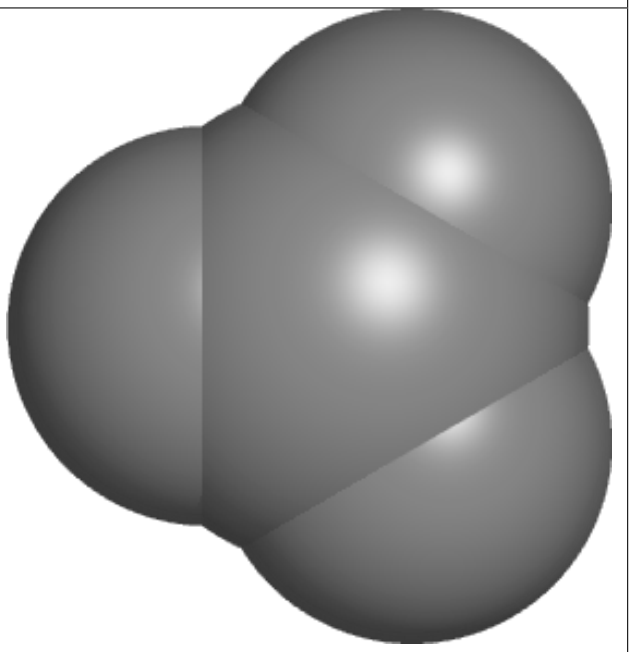
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	202	BEF	1	0

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.

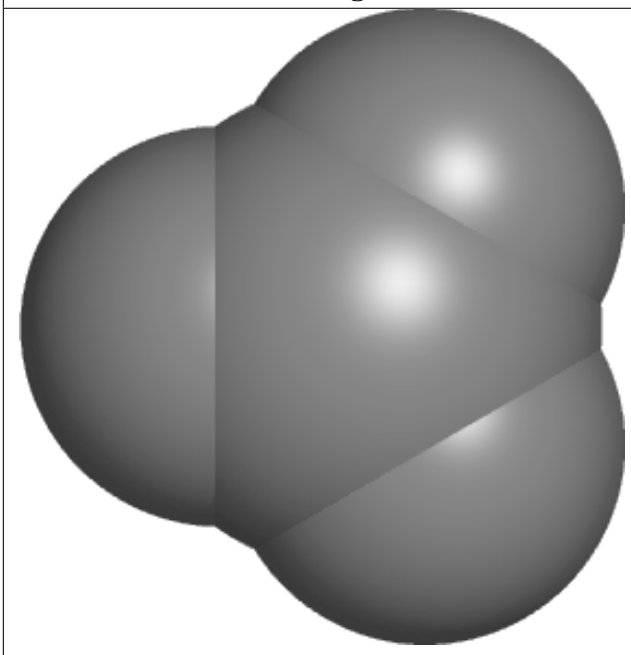
Ligand BEF B 202



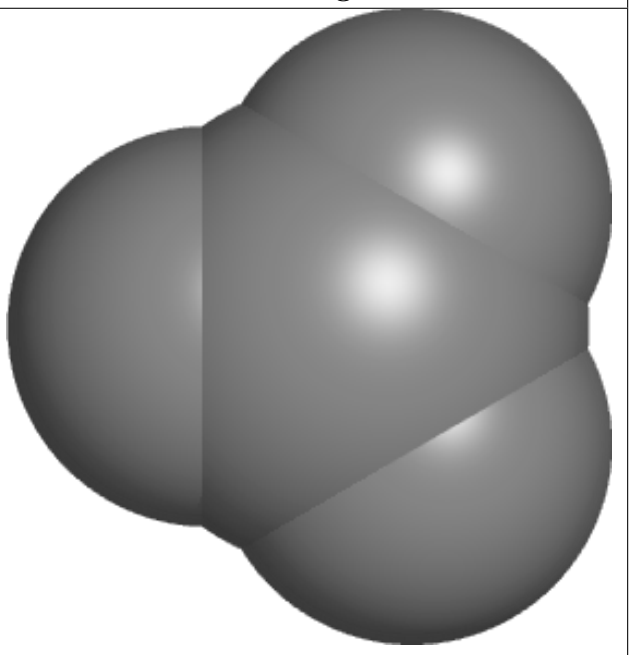
Bond lengths



Bond angles

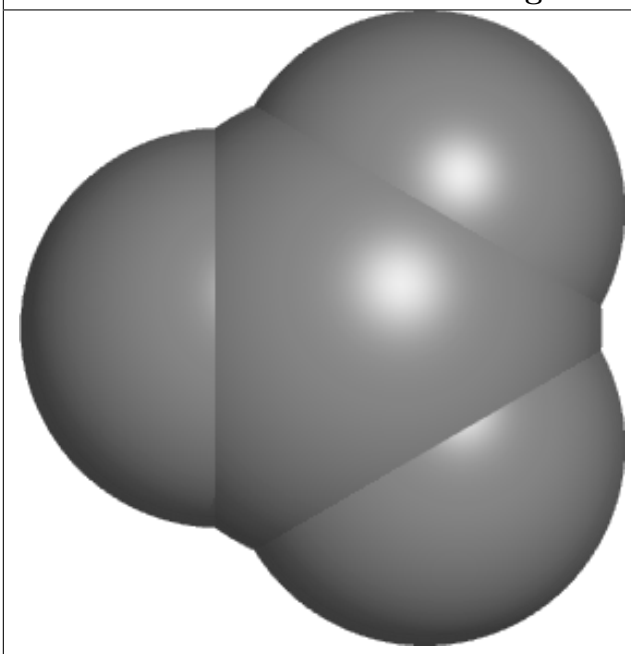


Torsions

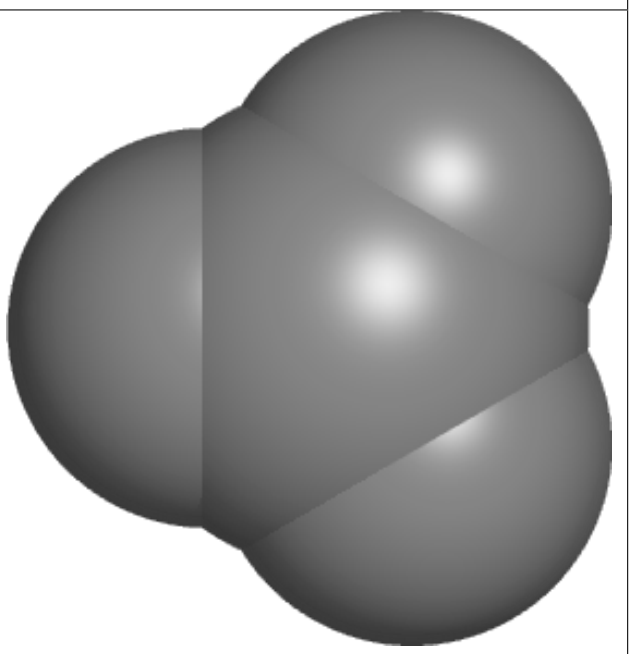


Rings

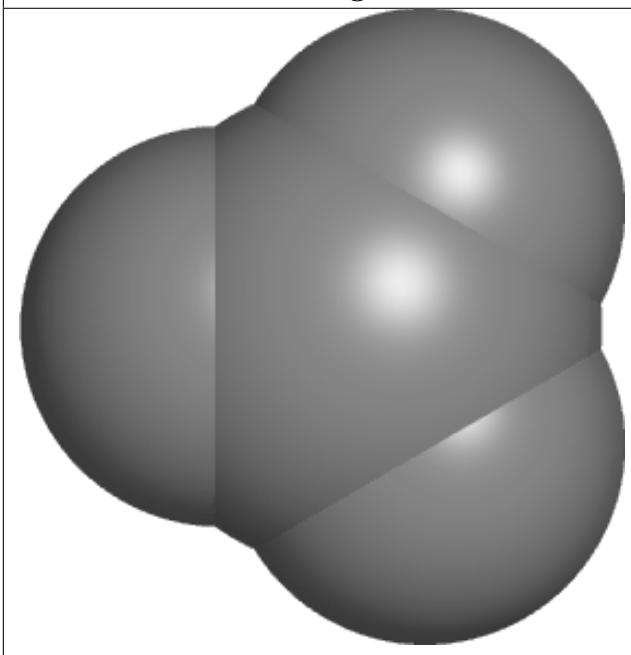
Ligand BEF A 202



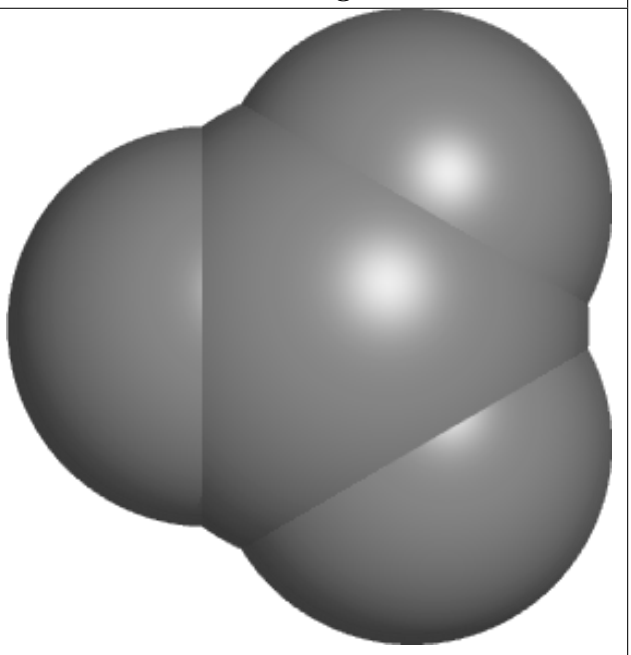
Bond lengths



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	122/142 (85%)	1.01	14 (11%)	4	7	51, 72, 98, 131	0
1	B	129/142 (90%)	1.00	14 (10%)	5	8	42, 59, 85, 112	0
2	C	90/115 (78%)	1.24	20 (22%)	0	1	60, 97, 122, 131	0
2	F	81/115 (70%)	1.41	25 (30%)	0	0	59, 98, 119, 129	0
All	All	422/514 (82%)	1.13	73 (17%)	1	1	42, 76, 113, 131	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	271	LEU	5.6
2	C	296	LEU	4.6
2	F	296	LEU	4.5
2	F	300	GLY	4.3
2	F	298	LEU	3.9

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

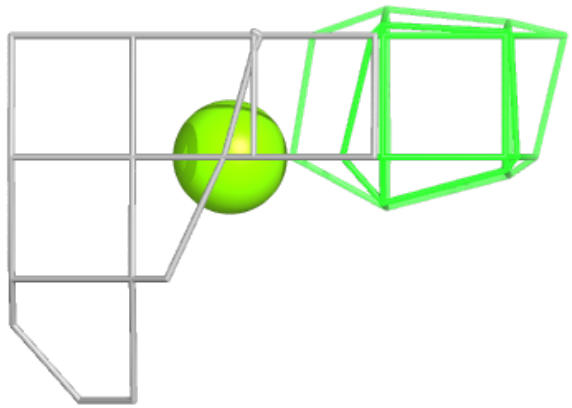
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

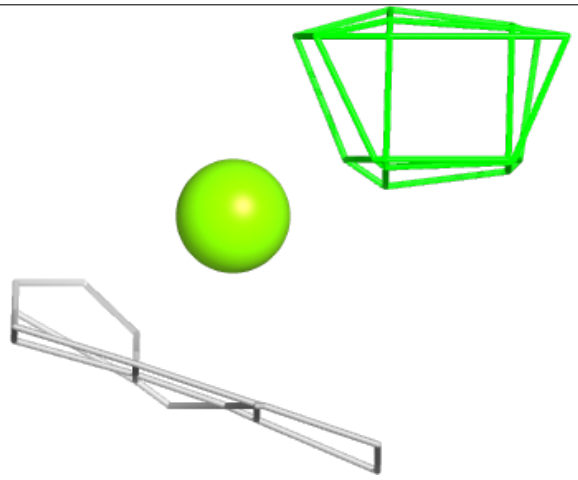
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	A	201	1/1	0.37	0.12	80,80,80,80	0
4	BEF	B	202	4/4	0.85	0.12	37,43,49,54	0
3	MG	B	201	1/1	0.92	0.09	50,50,50,50	0
4	BEF	A	202	4/4	0.94	0.08	69,73,75,88	0

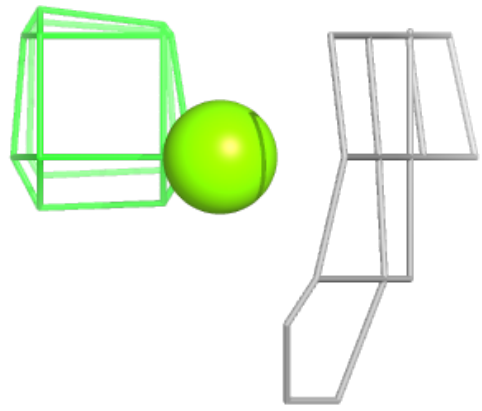
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around MG A 201:

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

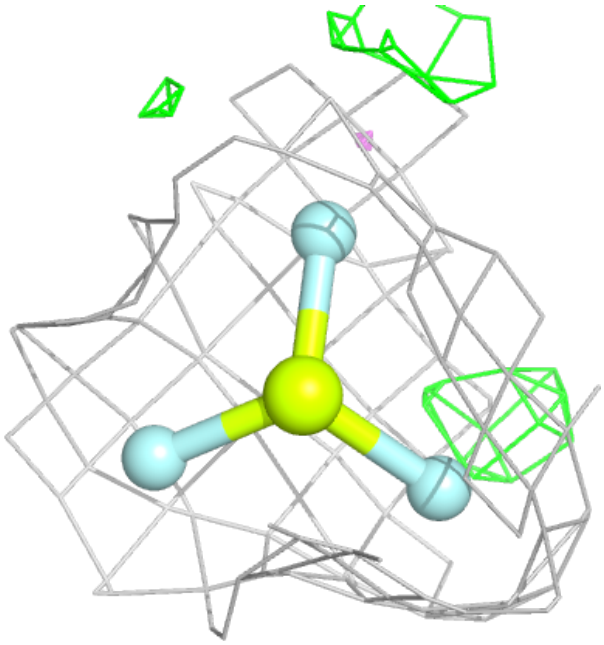


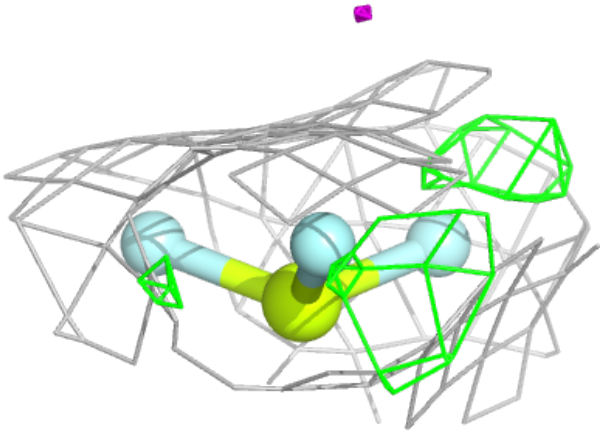


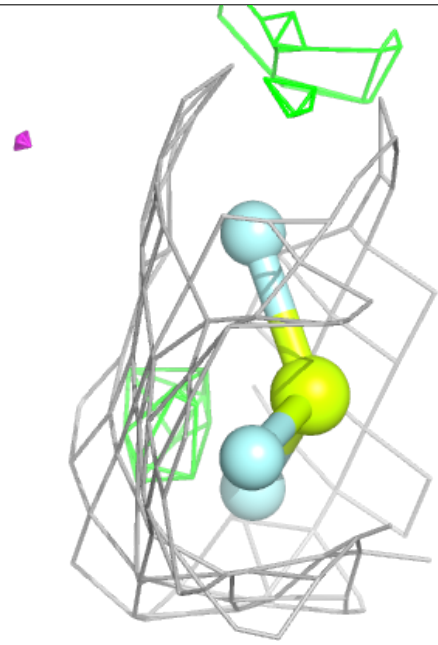


Electron density around BEF B 202:

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

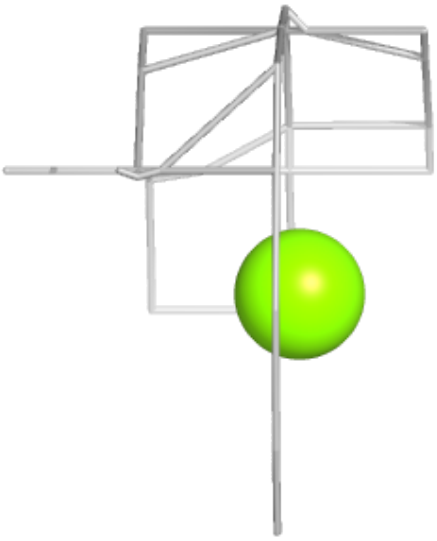


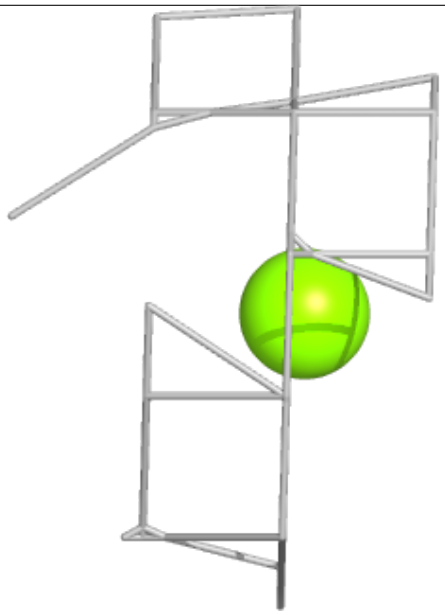


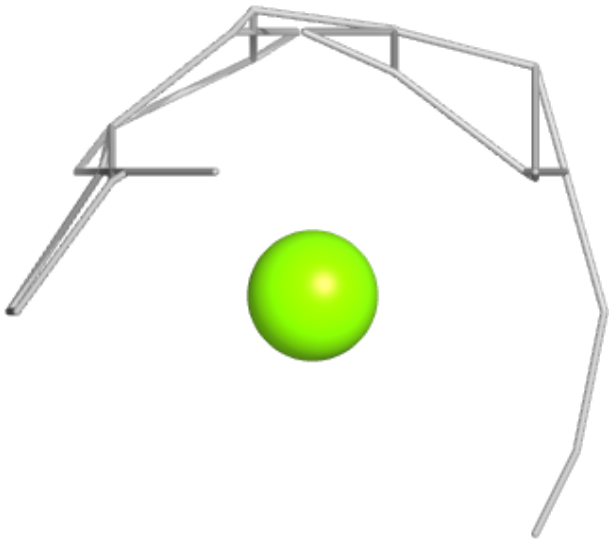


Electron density around MG B 201:

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

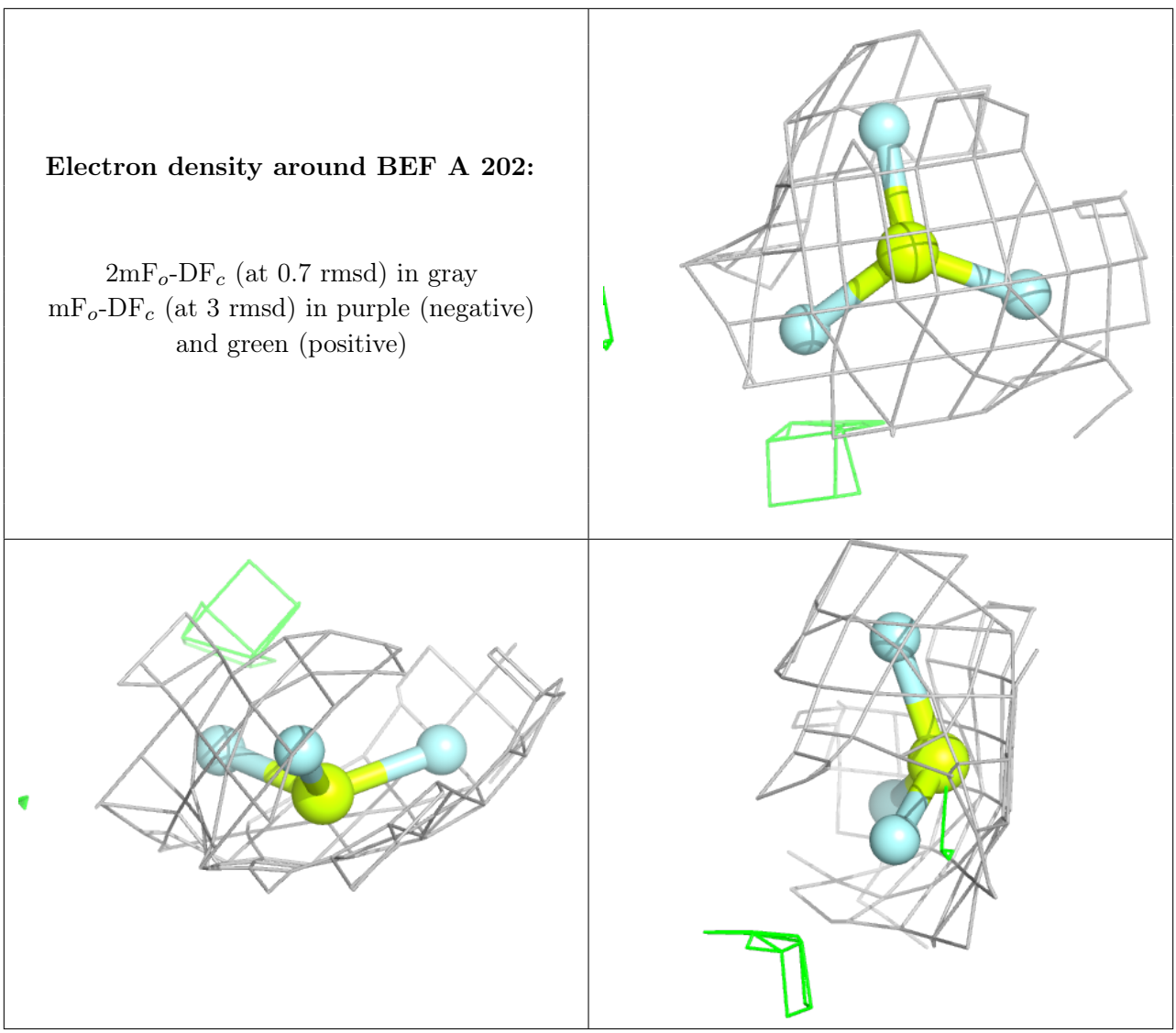






Electron density around BEF A 202:

$2mF_o$-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.