



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 7, 2024 – 06:49 pm GMT

PDB ID : 6HHS  
Title : MamM CTD E289D - Cadmium form  
Authors : Barber-Zucker, S.; Zarivach, R.  
Deposited on : 2018-08-29  
Resolution : 2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

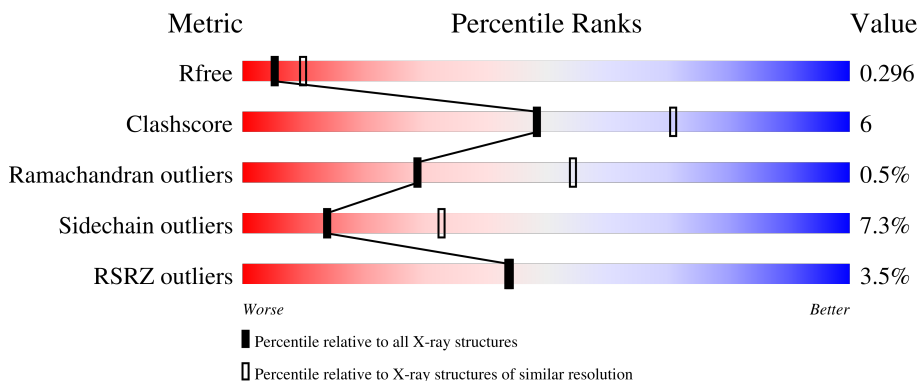
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.





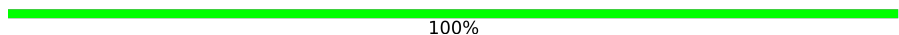
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	108	 3% 69% 6% . 23%
1	B	108	 3% 65% 11% . . 19%
1	C	108	 2% 66% 6% . 25%
1	D	108	 % 63% 10% . 23%
1	E	108	 % 65% 7% . 24%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	108	 68% 6% • 24% 9%
1	G	108	 66% 5% • • 26% 9%
2	M	8	 100%

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4710 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 Magnetosome protein MamM, Cation efflux protein family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	83	Total 641	C 393	N 123	O 121	S 4	0	0	0
1	B	88	Total 678	C 418	N 127	O 129	S 4	0	0	0
1	C	81	Total 629	C 387	N 121	O 117	S 4	0	0	0
1	D	83	Total 639	C 392	N 123	O 120	S 4	0	0	0
1	E	82	Total 635	C 390	N 122	O 119	S 4	0	0	0
1	F	82	Total 633	C 389	N 122	O 118	S 4	0	0	0
1	G	80	Total 629	C 386	N 123	O 116	S 4	0	1	0

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	GLY	-	expression tag	UNP Q6NE57
A	212	SER	-	expression tag	UNP Q6NE57
A	213	HIS	-	expression tag	UNP Q6NE57
A	214	MET	-	expression tag	UNP Q6NE57
A	289	ASP	GLU	engineered mutation	UNP Q6NE57
B	211	GLY	-	expression tag	UNP Q6NE57
B	212	SER	-	expression tag	UNP Q6NE57
B	213	HIS	-	expression tag	UNP Q6NE57
B	214	MET	-	expression tag	UNP Q6NE57
B	289	ASP	GLU	engineered mutation	UNP Q6NE57
C	211	GLY	-	expression tag	UNP Q6NE57
C	212	SER	-	expression tag	UNP Q6NE57
C	213	HIS	-	expression tag	UNP Q6NE57
C	214	MET	-	expression tag	UNP Q6NE57
C	289	ASP	GLU	engineered mutation	UNP Q6NE57

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	211	GLY	-	expression tag	UNP Q6NE57
D	212	SER	-	expression tag	UNP Q6NE57
D	213	HIS	-	expression tag	UNP Q6NE57
D	214	MET	-	expression tag	UNP Q6NE57
D	289	ASP	GLU	engineered mutation	UNP Q6NE57
E	211	GLY	-	expression tag	UNP Q6NE57
E	212	SER	-	expression tag	UNP Q6NE57
E	213	HIS	-	expression tag	UNP Q6NE57
E	214	MET	-	expression tag	UNP Q6NE57
E	289	ASP	GLU	engineered mutation	UNP Q6NE57
F	211	GLY	-	expression tag	UNP Q6NE57
F	212	SER	-	expression tag	UNP Q6NE57
F	213	HIS	-	expression tag	UNP Q6NE57
F	214	MET	-	expression tag	UNP Q6NE57
F	289	ASP	GLU	engineered mutation	UNP Q6NE57
G	211	GLY	-	expression tag	UNP Q6NE57
G	212	SER	-	expression tag	UNP Q6NE57
G	213	HIS	-	expression tag	UNP Q6NE57
G	214	MET	-	expression tag	UNP Q6NE57
G	289	ASP	GLU	engineered mutation	UNP Q6NE57

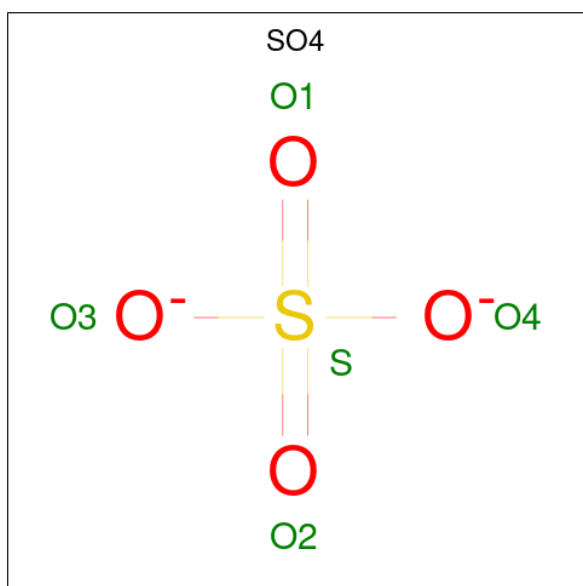
- Molecule 2 is a protein called Probable C-terminal region of MamM CTD E289D - Cadmium form.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	M	8	40	24	8	8	0	0	0

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd) (labeled as "Ligand of Interest" by depositor).

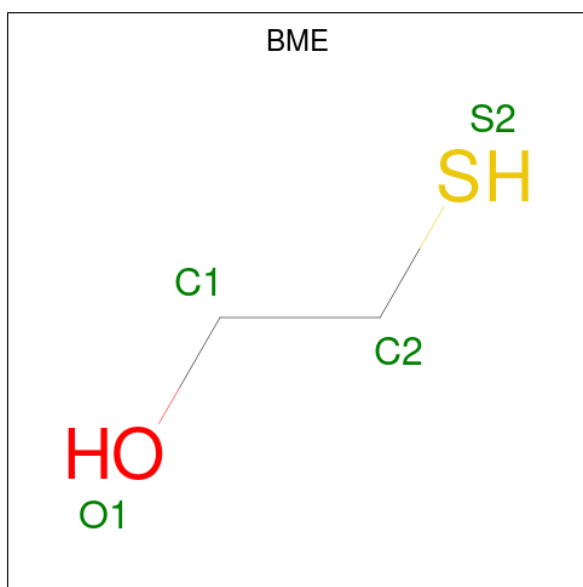
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total	Cd	0	0
			5	5		
3	B	1	Total	Cd	0	0
			1	1		
3	C	3	Total	Cd	0	0
			3	3		
3	D	1	Total	Cd	0	0
			1	1		
3	E	2	Total	Cd	0	0
			2	2		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 5 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C<sub>2</sub>H<sub>6</sub>OS).



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

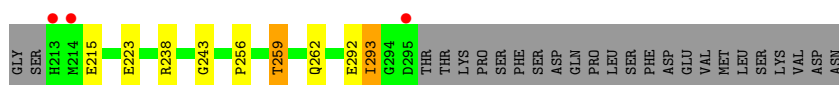
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	22	Total	O	0	0
			22	22		
6	B	19	Total	O	0	0
			19	19		
6	C	19	Total	O	0	0
			19	19		
6	D	20	Total	O	0	0
			20	20		
6	E	15	Total	O	0	0
			15	15		
6	F	11	Total	O	0	0
			11	11		
6	G	7	Total	O	0	0
			7	7		

### 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.

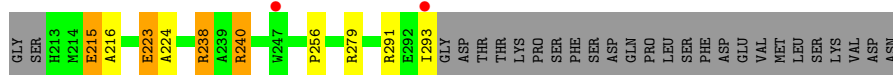
- Molecule 1: Magnetosome protein MamM, Cation efflux protein family



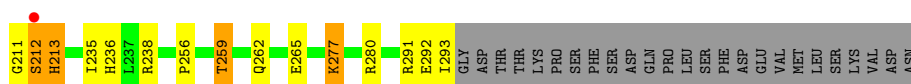
- Molecule 1: Magnetosome protein MamM, Cation efflux protein family



- Molecule 1: Magnetosome protein MamM, Cation efflux protein family



- Molecule 1: Magnetosome protein MamM, Cation efflux protein family



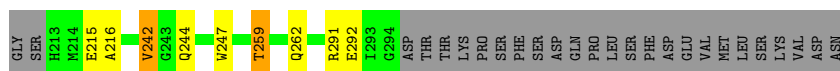
- Molecule 1: Magnetosome protein MamM, Cation efflux protein family



- Molecule 1: Magnetosome protein MamM, Cation efflux protein family

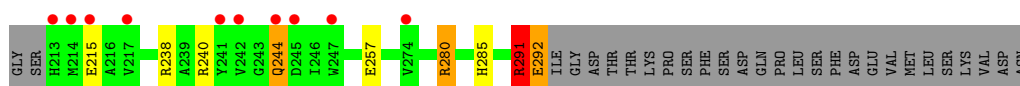


Chain F:  68% 6% 24%



- Molecule 1: Magnetosome protein MamM, Cation efflux protein family

Chain G:  9% 66% 5% 26%



- Molecule 2: Probable C-terminal region of MamM CTD E289D - Cadmium form

Chain M:  100%

There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.60Å 94.55Å 107.49Å 90.00° 91.76° 90.00°	Depositor
Resolution (Å)	47.27 – 2.70 47.27 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.3 (47.27-2.70) 98.3 (47.27-2.70)	Depositor EDS
$R_{merge}$	0.27	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0230, PHENIX	Depositor
R, $R_{free}$	0.250 , 0.290 0.252 , 0.296	Depositor DCC
$R_{free}$ test set	1140 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.3	Xtrriage
Anisotropy	0.543	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 43.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.017 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	4710	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.67% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: CD, BME, SO4

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	0.52	0/649	0.72	0/878
1	B	0.59	0/687	0.77	0/930
1	C	0.54	0/637	0.73	0/862
1	D	0.53	0/647	0.68	0/875
1	E	0.52	0/643	0.74	0/870
1	F	0.53	0/641	0.74	0/867
1	G	0.55	0/640	0.74	0/865
All	All	0.54	0/4544	0.73	0/6147

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	B	0	3
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
1	G	0	5
All	All	0	12

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (12) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	227	ARG	Sidechain
1	B	291	ARG	Sidechain
1	B	294	GLY	Peptide
1	C	240	ARG	Sidechain
1	D	291	ARG	Sidechain
1	E	238	ARG	Sidechain
1	F	291	ARG	Sidechain
1	G	238	ARG	Sidechain
1	G	240	ARG	Sidechain
1	G	280[A]	ARG	Sidechain
1	G	280[B]	ARG	Sidechain
1	G	291	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	641	0	632	10	0
1	B	678	0	673	15	0
1	C	629	0	625	7	1
1	D	639	0	634	15	0
1	E	635	0	631	9	0
1	F	633	0	628	6	1
1	G	629	0	628	4	0
2	M	40	0	10	0	0
3	A	5	0	0	0	0
3	B	1	0	0	0	0
3	C	3	0	0	1	0
3	D	1	0	0	0	0
3	E	2	0	0	0	0
4	A	10	0	0	0	0
4	B	5	0	0	0	0
4	C	10	0	0	0	0
4	D	10	0	0	0	0
4	E	5	0	0	0	0
4	F	5	0	0	0	0
5	A	4	0	5	1	0
5	B	4	0	5	0	0
5	C	4	0	5	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	F	4	0	5	0	0
6	A	22	0	0	0	0
6	B	19	0	0	2	0
6	C	19	0	0	2	0
6	D	20	0	0	1	0
6	E	15	0	0	0	0
6	F	11	0	0	0	0
6	G	7	0	0	1	0
All	All	4710	0	4481	54	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 6.

All (54) 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:D:212:SER:O	1:D:213:HIS:O	1.95	0.85
1:G:244:GLN:NE2	1:G:244:GLN:H	1.84	0.75
1:B:244:GLN:OE1	1:D:292:GLU:HB2	1.89	0.71
1:E:244:GLN:HE22	1:F:292:GLU:HB2	1.55	0.71
1:D:211:GLY:HA3	1:D:280:ARG:HH22	1.56	0.70
1:B:277:LYS:N	1:B:277:LYS:HD3	2.12	0.65
1:D:277:LYS:N	1:D:277:LYS:HD3	2.13	0.64
1:F:215:GLU:HG2	1:F:216:ALA:N	2.11	0.64
1:C:215:GLU:HG2	1:C:216:ALA:N	2.13	0.62
1:E:222:VAL:HG22	1:E:234:VAL:HG21	1.81	0.62
1:B:234:VAL:HG22	1:B:250:MET:HE1	1.81	0.61
6:C:515:HOH:O	1:D:265:GLU:HG2	2.01	0.61
1:A:238:ARG:NH1	1:D:238:ARG:HD2	2.16	0.60
3:C:403:CD:CD	6:C:501:HOH:O	1.68	0.60
1:A:238:ARG:CZ	1:D:238:ARG:HD2	2.33	0.59
1:A:243:GLY:C	1:G:291:ARG:HH21	2.06	0.59
1:E:213:HIS:O	1:E:217:VAL:HG22	2.02	0.59
1:E:214:MET:HA	1:E:217:VAL:CG2	2.32	0.59
1:A:259:THR:HG23	1:D:256:PRO:O	2.03	0.58
1:C:238:ARG:NH2	1:C:240:ARG:HG2	2.18	0.58
1:B:291:ARG:HH11	1:B:291:ARG:CG	2.18	0.57
1:B:291:ARG:NH2	1:B:297:THR:O	2.39	0.56
1:C:256:PRO:O	1:E:259:THR:HG23	2.07	0.55
1:C:279:ARG:HA	6:D:513:HOH:O	2.06	0.55
1:A:256:PRO:O	1:D:259:THR:HG23	2.06	0.54

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:ARG:NH2	1:D:238:ARG:HD2	2.22	0.54
1:C:238:ARG:HH21	1:C:240:ARG:HG2	1.73	0.54
1:E:214:MET:O	1:E:217:VAL:HG23	2.09	0.52
1:B:223:GLU:HA	6:B:507:HOH:O	2.09	0.52
1:E:214:MET:HA	1:E:217:VAL:HG23	1.93	0.50
5:A:408:BME:H12	1:B:271:GLN:HG2	1.92	0.50
1:E:259:THR:HG22	1:E:262:GLN:H	1.77	0.49
1:A:259:THR:HG22	1:A:262:GLN:H	1.77	0.48
1:F:259:THR:HG22	1:F:262:GLN:H	1.78	0.48
1:B:232:ARG:O	1:B:291:ARG:NH1	2.46	0.48
1:B:244:GLN:OE1	1:D:292:GLU:N	2.46	0.48
1:F:242:VAL:HG21	1:F:247:TRP:CD1	2.47	0.47
1:A:293:ILE:H	1:A:293:ILE:CD1	2.28	0.47
1:D:292:GLU:O	1:D:293:ILE:HB	2.14	0.47
1:G:292:GLU:HG2	6:G:402:HOH:O	2.14	0.47
1:D:259:THR:HG22	1:D:262:GLN:H	1.79	0.46
1:E:244:GLN:NE2	1:F:292:GLU:HB2	2.29	0.46
1:C:223:GLU:HG2	1:C:224:ALA:N	2.30	0.45
1:B:236:HIS:CE1	1:G:285:HIS:CE1	3.04	0.45
1:F:242:VAL:CG2	1:F:247:TRP:CD1	3.00	0.44
1:D:235:ILE:HG22	1:D:236:HIS:CD2	2.53	0.44
1:A:293:ILE:HD13	1:A:293:ILE:N	2.32	0.44
1:B:234:VAL:HG22	1:B:250:MET:CE	2.47	0.43
1:B:244:GLN:OE1	1:D:292:GLU:CB	2.61	0.43
1:A:293:ILE:H	1:A:293:ILE:HD13	1.85	0.41
1:B:291:ARG:HH11	1:B:291:ARG:HG2	1.83	0.41
1:B:227:ARG:NH1	6:B:504:HOH:O	2.54	0.41
1:B:295:ASP:O	1:B:296:THR:CB	2.68	0.41
1:C:293:ILE:O	1:C:293:ILE:HG22	2.20	0.40

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:C:291:ARG:NH1	1:F:244:GLN:OE1[2_455]	2.12	0.08

## 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	81/108 (75%)	79 (98%)	2 (2%)	0	100	100
1	B	86/108 (80%)	84 (98%)	0	2 (2%)	6	16
1	C	79/108 (73%)	78 (99%)	1 (1%)	0	100	100
1	D	81/108 (75%)	78 (96%)	2 (2%)	1 (1%)	13	32
1	E	80/108 (74%)	79 (99%)	1 (1%)	0	100	100
1	F	80/108 (74%)	79 (99%)	1 (1%)	0	100	100
1	G	79/108 (73%)	78 (99%)	1 (1%)	0	100	100
All	All	566/756 (75%)	555 (98%)	8 (1%)	3 (0%)	29	54

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	296	THR
1	D	213	HIS
1	B	295	ASP

### 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	66/90 (73%)	61 (92%)	5 (8%)	13	30
1	B	71/90 (79%)	63 (89%)	8 (11%)	6	13
1	C	65/90 (72%)	62 (95%)	3 (5%)	27	54

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	66/90 (73%)	63 (96%)	3 (4%)	27	55
1	E	66/90 (73%)	59 (89%)	7 (11%)	6	15
1	F	65/90 (72%)	63 (97%)	2 (3%)	40	69
1	G	65/90 (72%)	58 (89%)	7 (11%)	6	15
All	All	464/630 (74%)	429 (92%)	35 (8%)	14	31

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	215	GLU
1	A	223	GLU
1	A	259	THR
1	A	292	GLU
1	A	293	ILE
1	B	215	GLU
1	B	238	ARG
1	B	261	GLU
1	B	277	LYS
1	B	289	ASP
1	B	291	ARG
1	B	295	ASP
1	B	297	THR
1	C	215	GLU
1	C	223	GLU
1	C	238	ARG
1	D	212	SER
1	D	259	THR
1	D	277	LYS
1	E	213	HIS
1	E	215	GLU
1	E	217	VAL
1	E	234	VAL
1	E	259	THR
1	E	292	GLU
1	E	293	ILE
1	F	242	VAL
1	F	259	THR
1	G	215	GLU
1	G	244	GLN
1	G	257	GLU
1	G	280[A]	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	G	280[B]	ARG
1	G	291	ARG
1	G	292	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	236	HIS
1	E	244	GLN
1	G	244	GLN
1	G	285	HIS

### 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 25 ligands modelled in this entry, 12 are monoatomic - leaving 13 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	403	-	4,4,4	0.40	0	6,6,6	0.31	0
4	SO4	A	406	-	4,4,4	0.40	0	6,6,6	0.29	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	D	402	-	4,4,4	0.27	0	6,6,6	0.44	0
5	BME	C	406	1	3,3,3	0.72	0	1,2,2	1.31	0
5	BME	B	403	1	3,3,3	0.40	0	1,2,2	0.02	0
5	BME	A	408	1	3,3,3	0.50	0	1,2,2	0.02	0
4	SO4	B	402	-	4,4,4	0.38	0	6,6,6	0.21	0
4	SO4	F	401	-	4,4,4	0.26	0	6,6,6	0.48	0
4	SO4	E	403	-	4,4,4	0.37	0	6,6,6	0.36	0
4	SO4	C	404	-	4,4,4	0.28	0	6,6,6	0.23	0
4	SO4	A	407	-	4,4,4	0.56	0	6,6,6	0.54	0
4	SO4	C	405	-	4,4,4	0.32	0	6,6,6	0.44	0
5	BME	F	402	1	3,3,3	0.79	0	1,2,2	0.75	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BME	A	408	1	-	1/1/1/1	-
5	BME	C	406	1	-	1/1/1/1	-
5	BME	B	403	1	-	1/1/1/1	-
5	BME	F	402	1	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	F	402	BME	O1-C1-C2-S2
5	B	403	BME	O1-C1-C2-S2
5	C	406	BME	O1-C1-C2-S2
5	A	408	BME	O1-C1-C2-S2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	408	BME	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	83/108 (76%)	0.27	3 (3%) 42 42	24, 31, 65, 86	0
1	B	88/108 (81%)	0.42	3 (3%) 45 45	26, 38, 59, 70	1 (1%)
1	C	81/108 (75%)	0.22	2 (2%) 57 59	23, 31, 44, 73	0
1	D	83/108 (76%)	0.15	1 (1%) 79 80	26, 34, 53, 86	0
1	E	82/108 (75%)	0.22	1 (1%) 79 80	25, 35, 56, 81	0
1	F	82/108 (75%)	0.18	0 100 100	23, 32, 48, 82	0
1	G	80/108 (74%)	0.73	10 (12%) 3 3	26, 42, 82, 93	1 (1%)
2	M	0/8	-	-	-	-
All	All	579/764 (75%)	0.31	20 (3%) 44 44	23, 34, 64, 93	2 (0%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	244	GLN	4.5
1	G	247	TRP	3.6
1	G	217	VAL	3.4
1	G	214	MET	3.3
1	G	213	HIS	3.3
1	G	241	TYR	2.9
1	A	214	MET	2.9
1	G	242	VAL	2.9
1	A	295	ASP	2.6
1	D	212	SER	2.6
1	C	293	ILE	2.5
1	G	245	ASP	2.4
1	G	215	GLU	2.4
1	G	274	VAL	2.3
1	A	213	HIS	2.2
1	B	296	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	300	SER	2.1
1	C	247	TRP	2.1
1	B	217	VAL	2.0
1	E	293	ILE	2.0

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

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

## 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<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	BME	C	406	4/4	0.77	0.30	43,43,45,46	0
5	BME	F	402	4/4	0.88	0.30	40,45,46,48	0
5	BME	B	403	4/4	0.89	0.19	57,60,63,63	0
5	BME	A	408	4/4	0.92	0.29	54,63,63,66	0
4	SO4	D	403	5/5	0.94	0.11	52,54,61,62	0
4	SO4	A	407	5/5	0.96	0.20	30,31,33,34	5
4	SO4	B	402	5/5	0.97	0.15	36,36,39,40	0
4	SO4	C	404	5/5	0.98	0.15	33,34,36,36	0
4	SO4	C	405	5/5	0.98	0.14	29,30,31,36	0
4	SO4	A	406	5/5	0.98	0.12	25,26,27,28	0
4	SO4	E	403	5/5	0.98	0.10	34,37,41,41	0
3	CD	D	401	1/1	0.99	0.13	32,32,32,32	0
4	SO4	F	401	5/5	0.99	0.17	30,31,32,34	0
3	CD	E	401	1/1	0.99	0.12	27,27,27,27	0
3	CD	A	405	1/1	0.99	0.10	24,24,24,24	0
4	SO4	D	402	5/5	0.99	0.19	27,28,28,29	0
3	CD	C	402	1/1	0.99	0.12	26,26,26,26	0
3	CD	A	403	1/1	1.00	0.11	30,30,30,30	0
3	CD	A	404	1/1	1.00	0.10	34,34,34,34	0

*Continued on next page...*

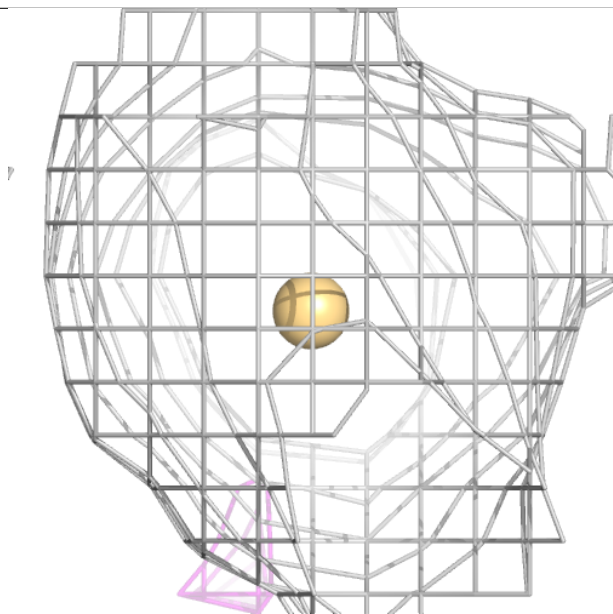
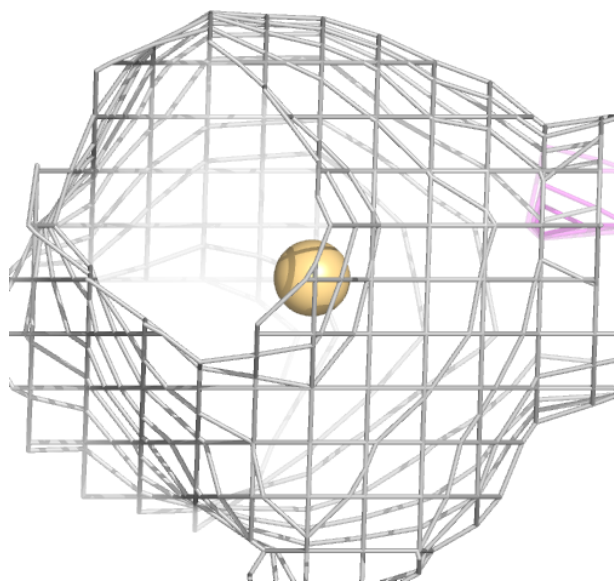
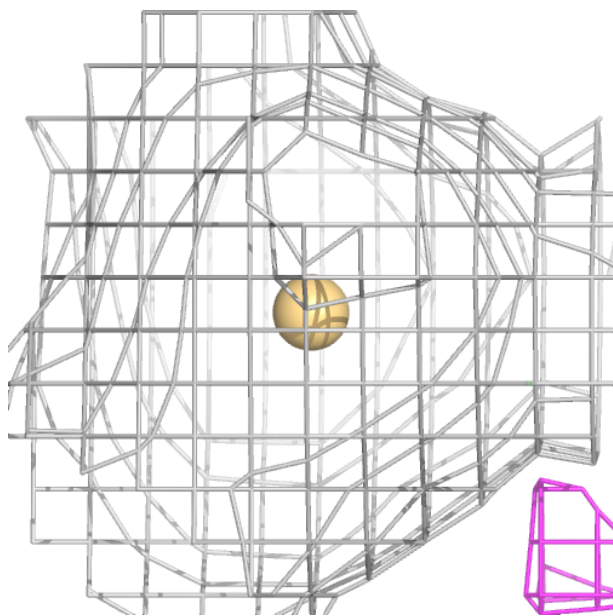
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CD	E	402	1/1	1.00	0.10	29,29,29,29	0
3	CD	A	401	1/1	1.00	0.14	25,25,25,25	0
3	CD	B	401	1/1	1.00	0.11	29,29,29,29	0
3	CD	C	401	1/1	1.00	0.12	29,29,29,29	0
3	CD	A	402	1/1	1.00	0.12	28,28,28,28	0
3	CD	C	403	1/1	1.00	0.12	27,27,27,27	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.

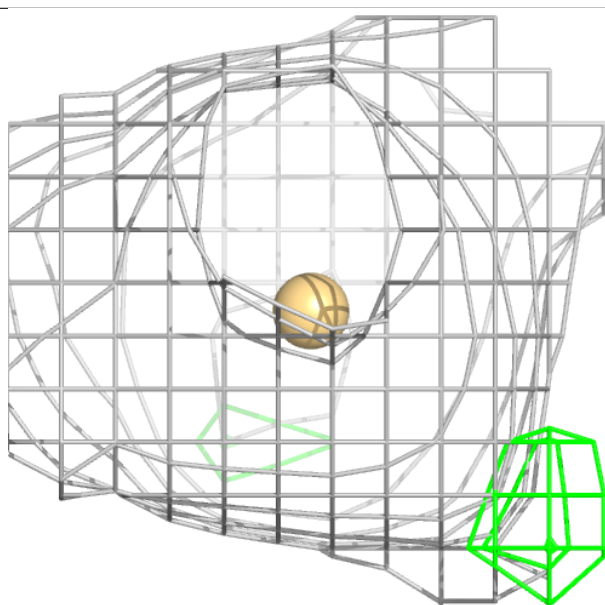
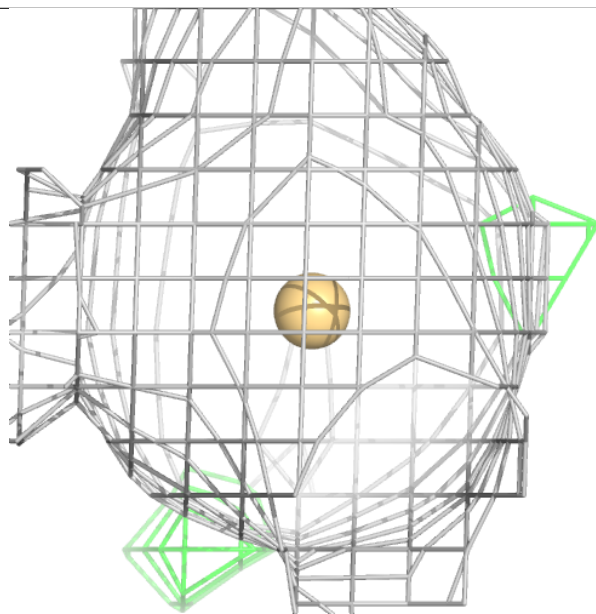
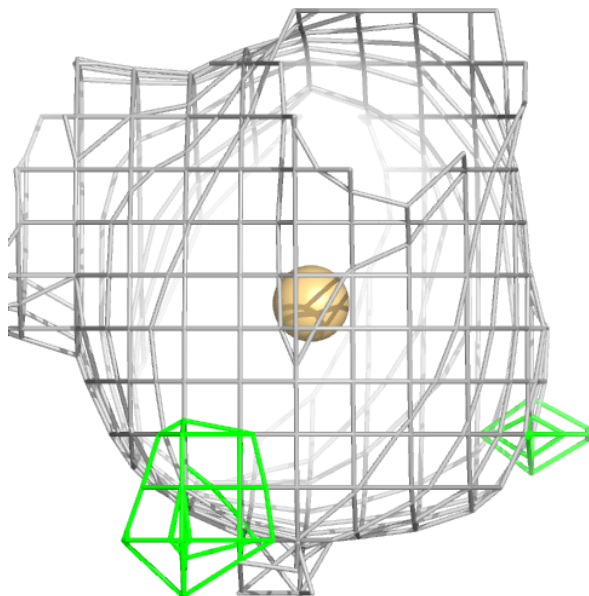
**Electron density around CD D 401:**

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



**Electron density around CD E 401:**

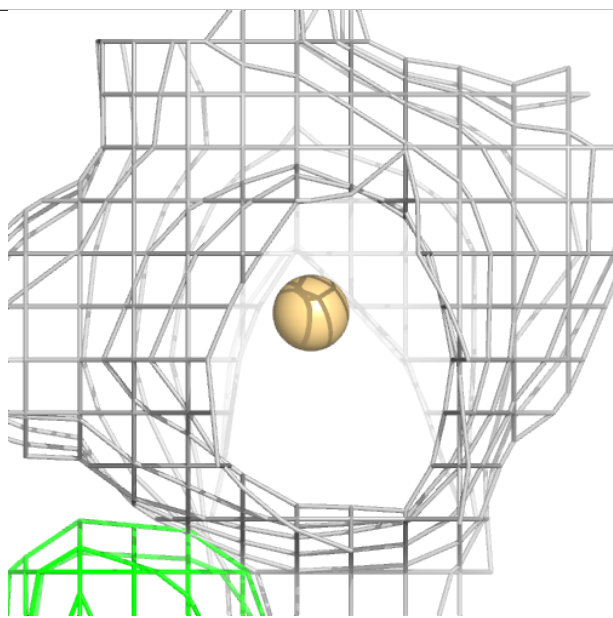
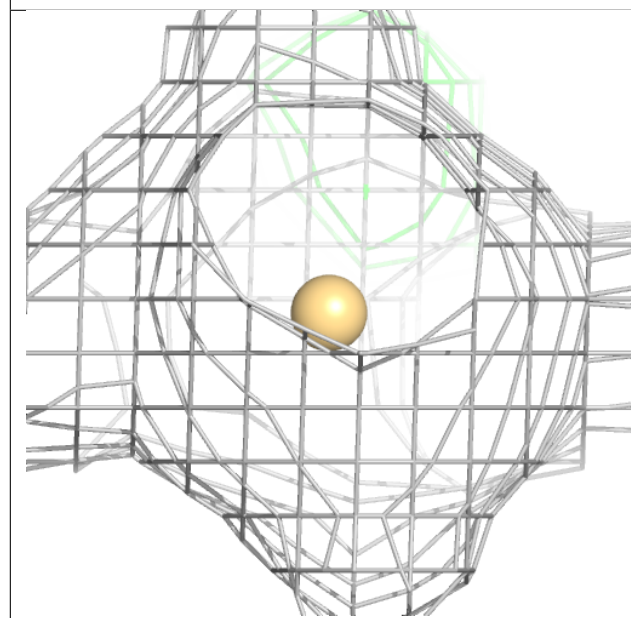
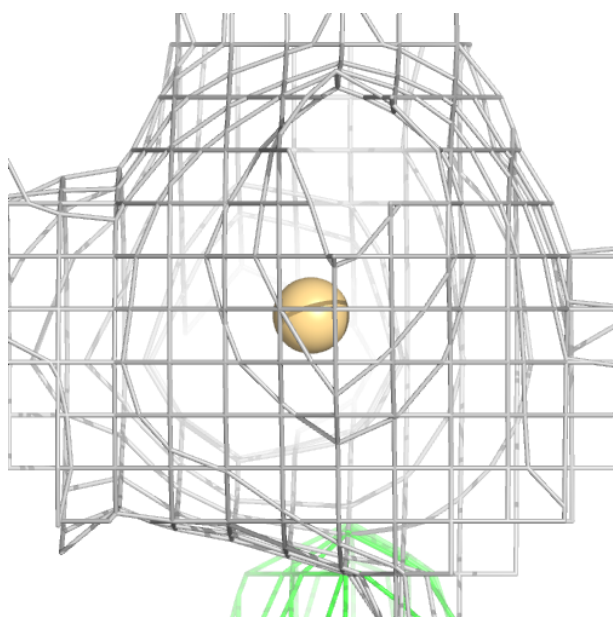
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





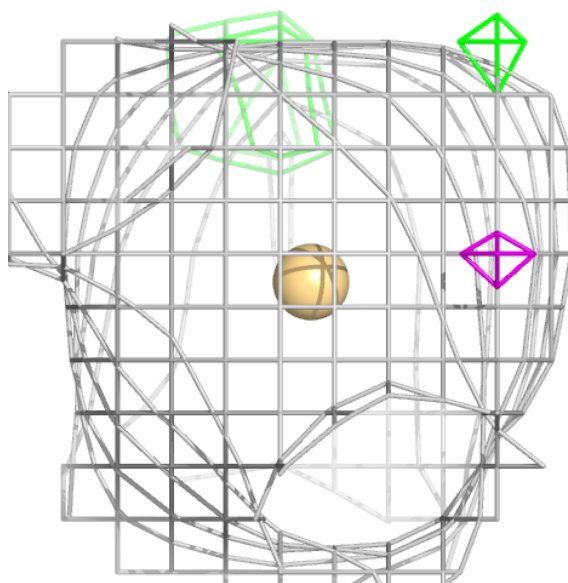
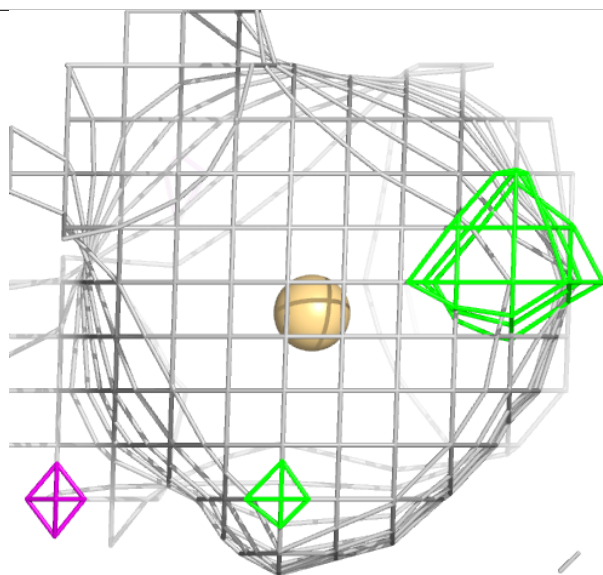
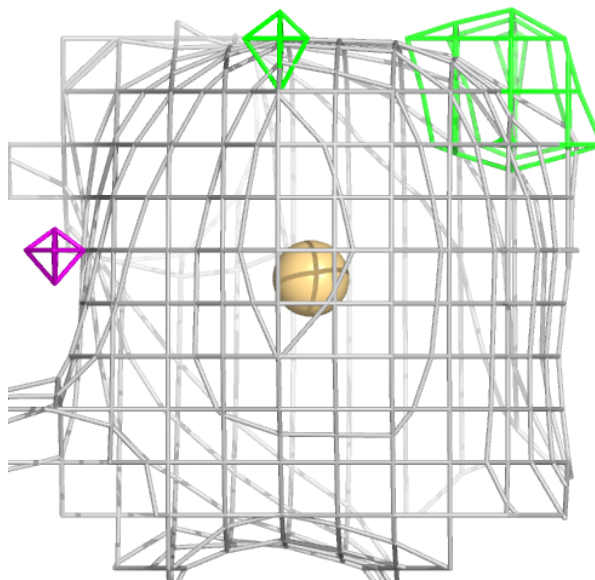
**Electron density around CD A 405:**

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



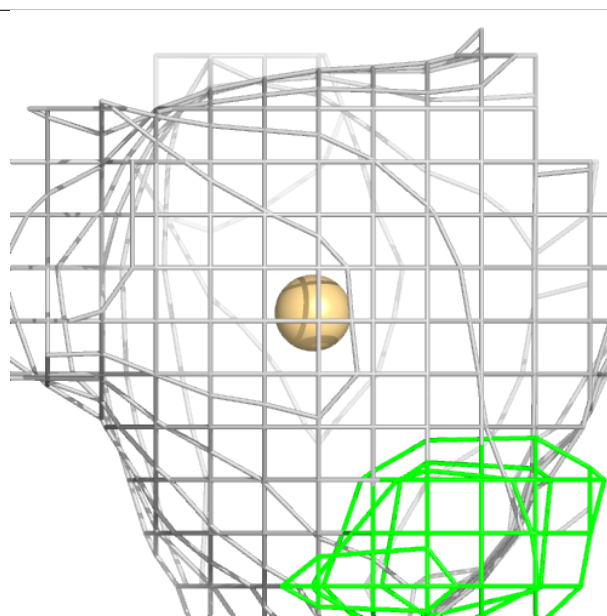
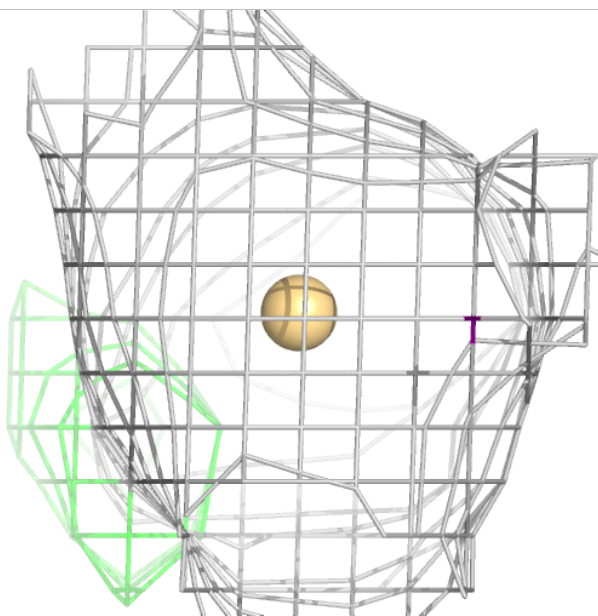
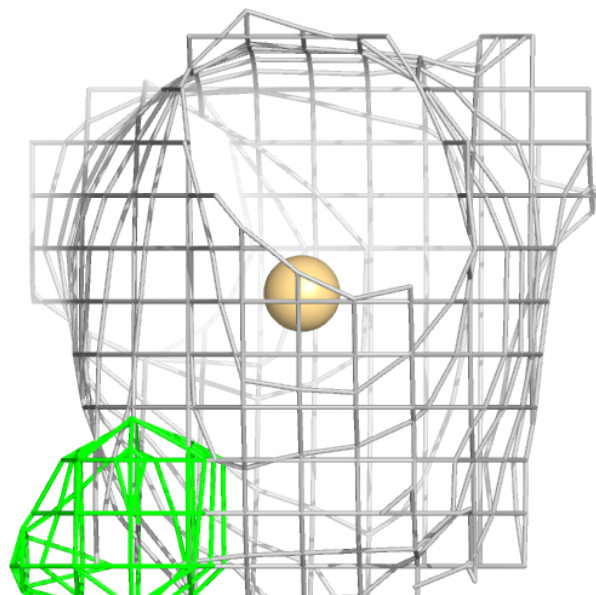
**Electron density around CD C 402:**

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



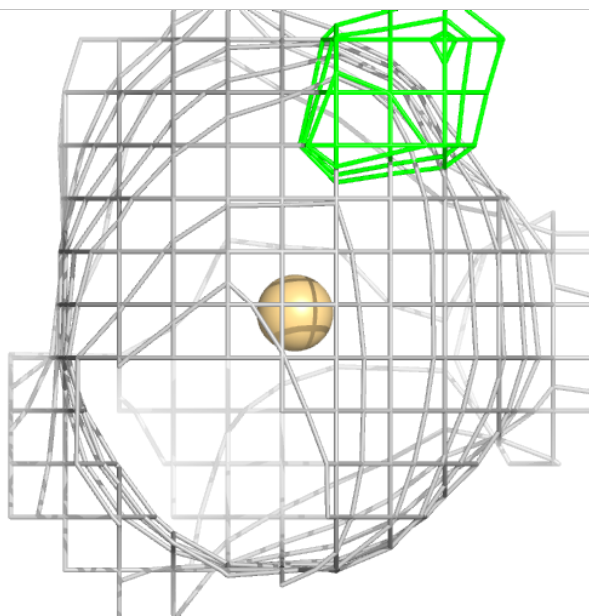
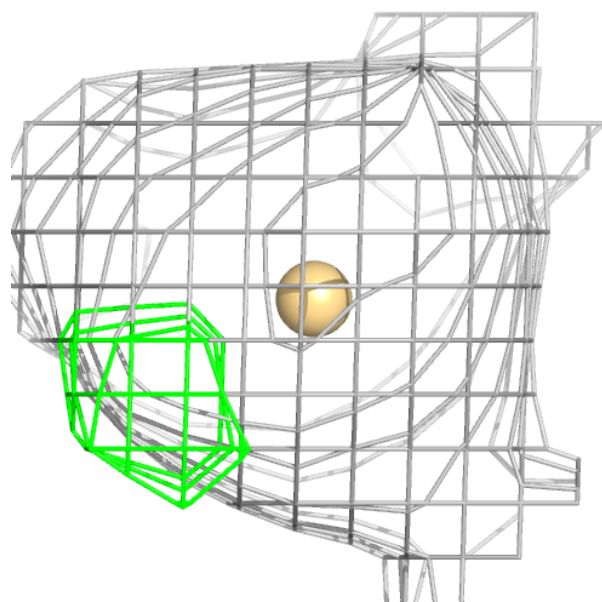
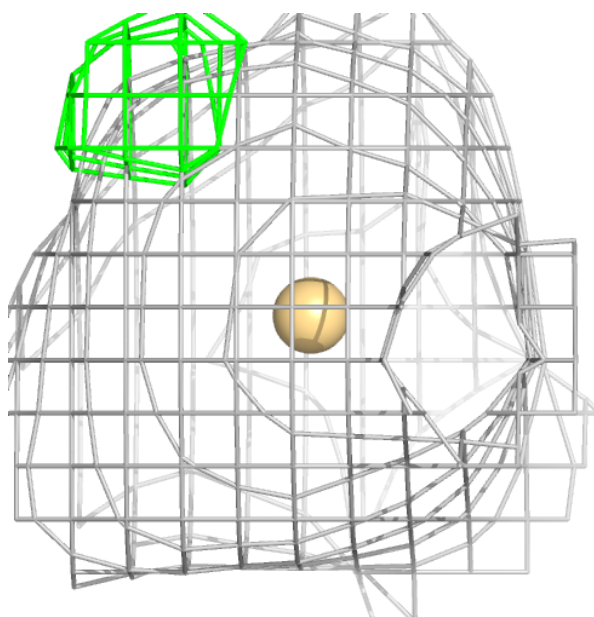
**Electron density around CD A 403:**

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



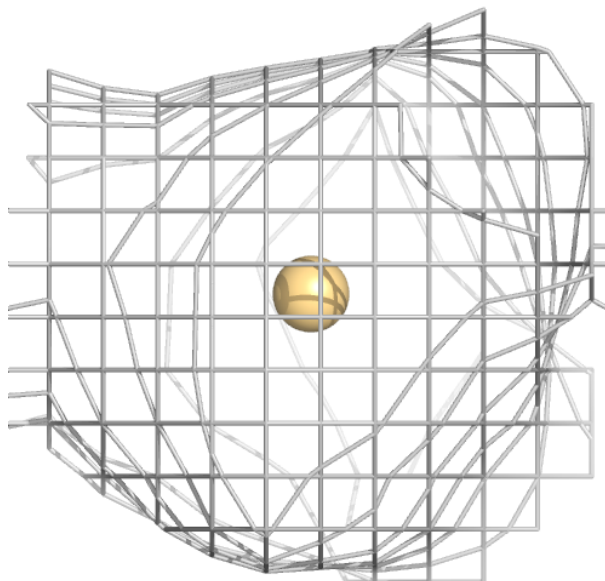
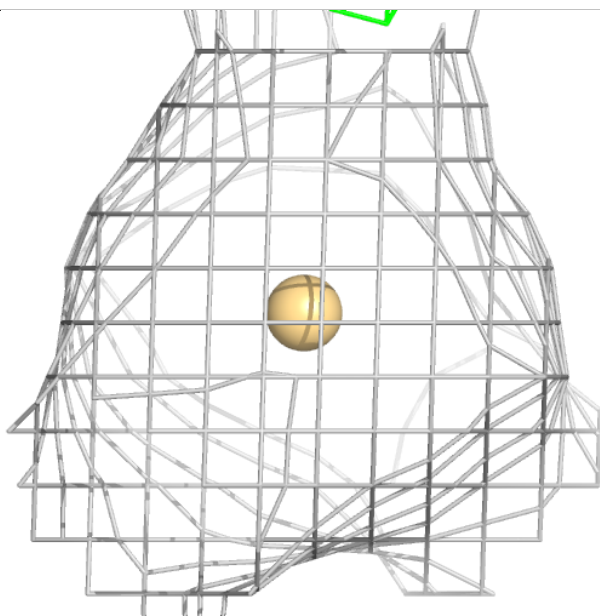
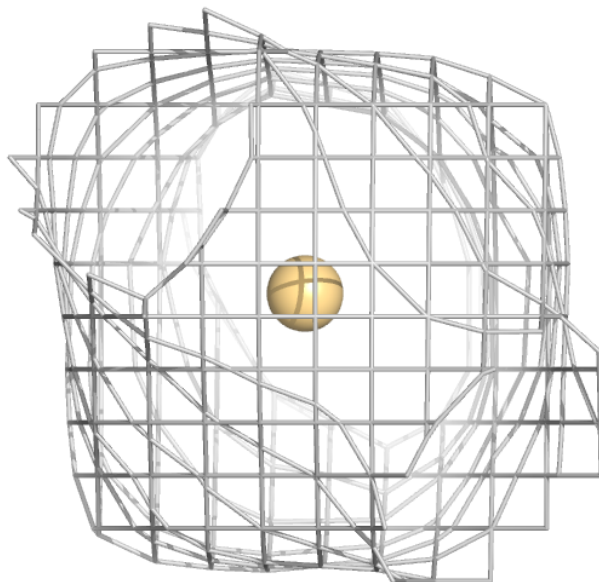
**Electron density around CD A 404:**

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



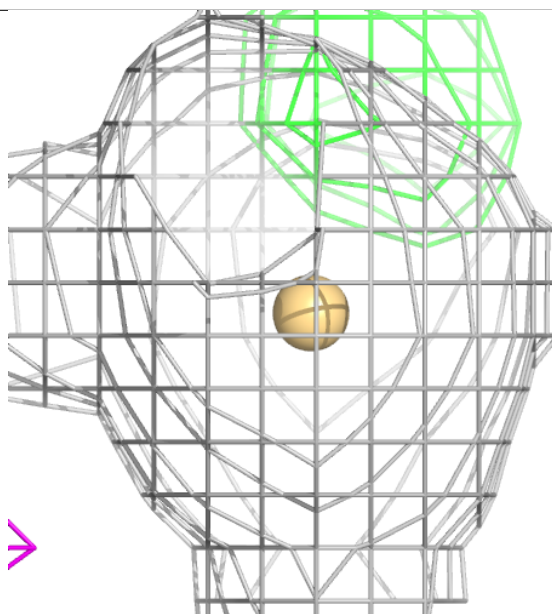
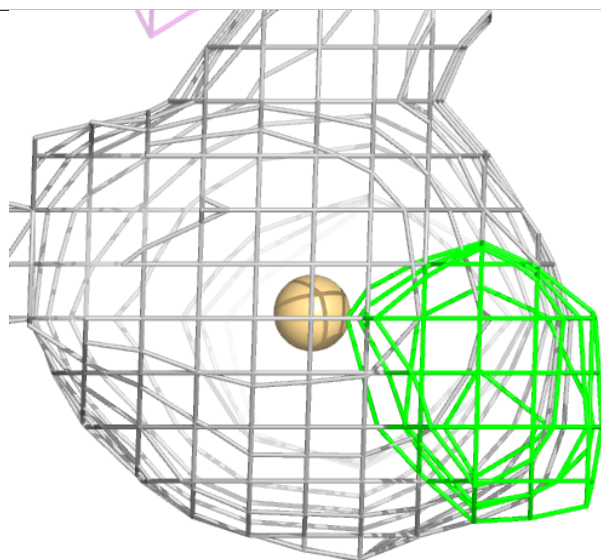
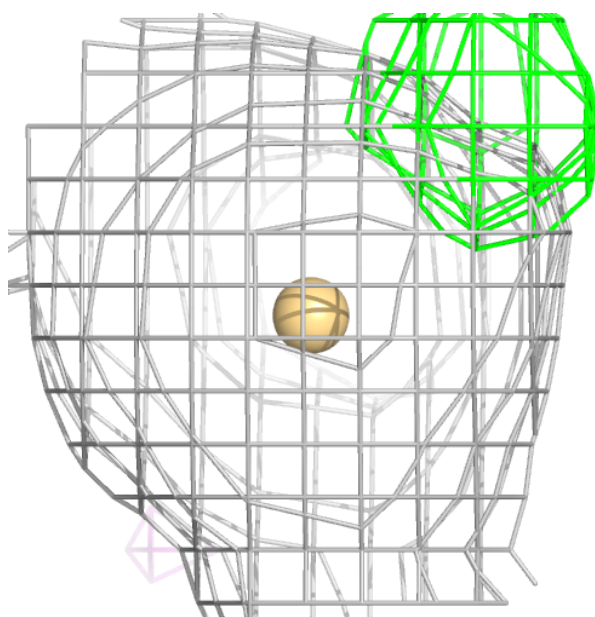
**Electron density around CD E 402:**

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



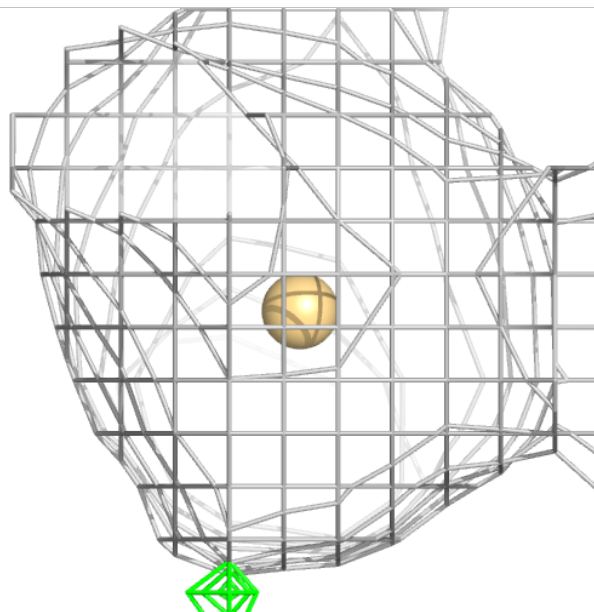
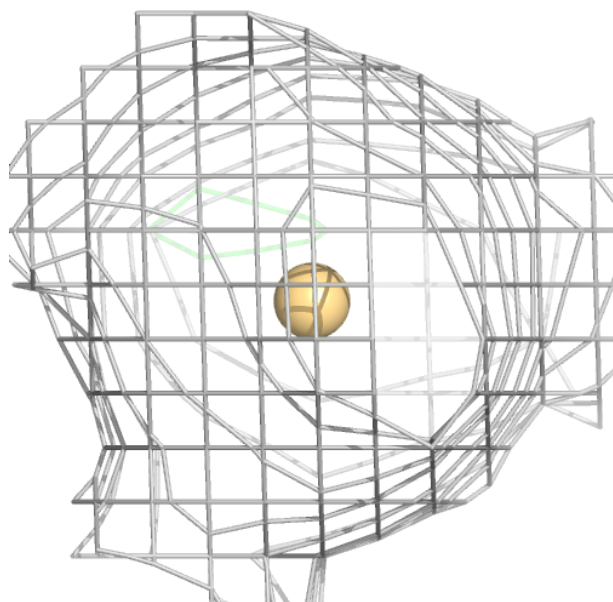
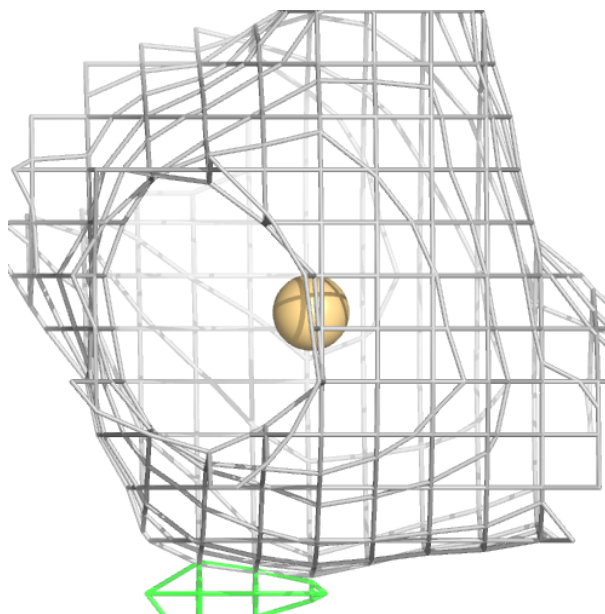
**Electron density around CD A 401:**

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



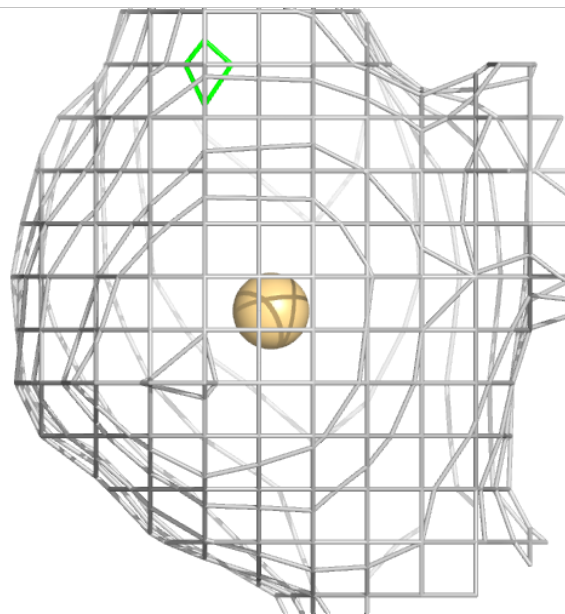
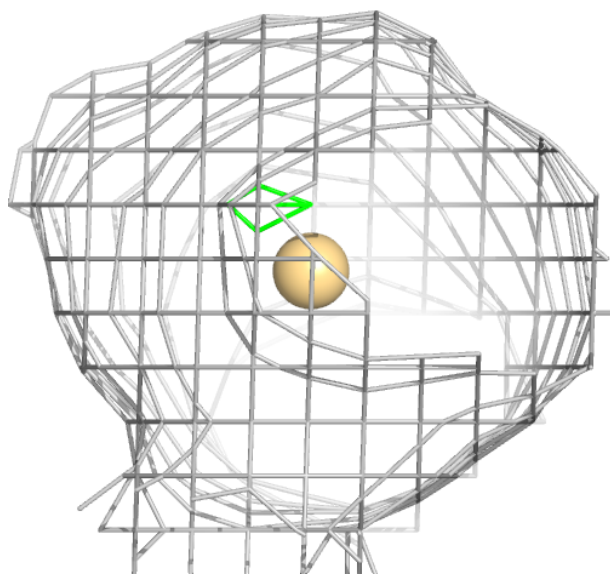
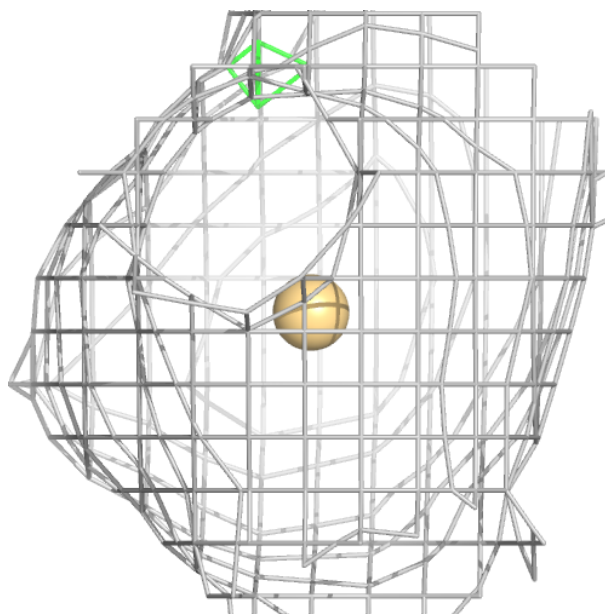
**Electron density around CD B 401:**

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



**Electron density around CD C 401:**

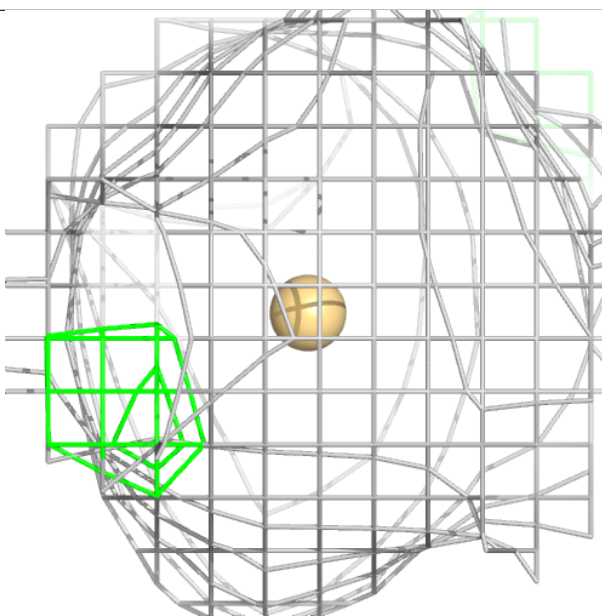
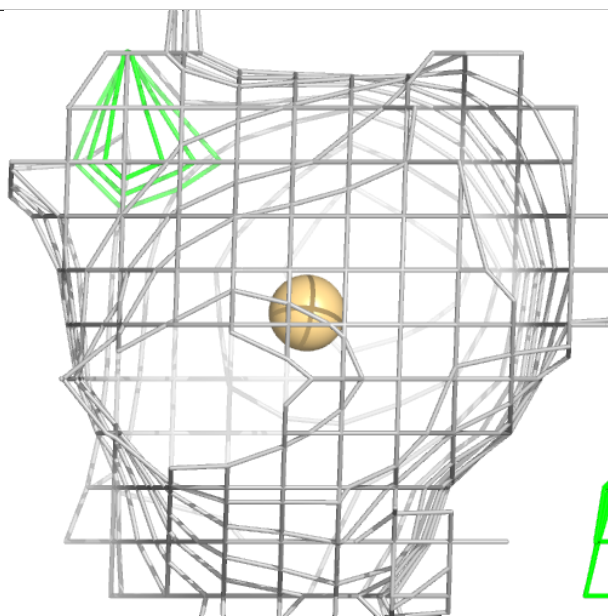
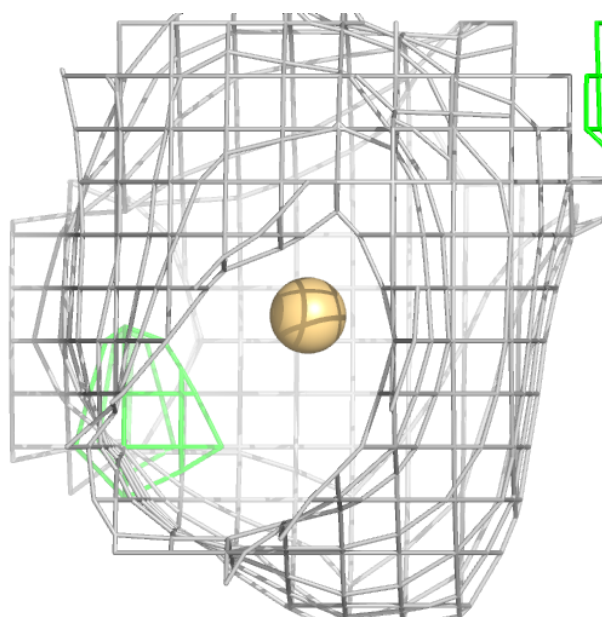
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

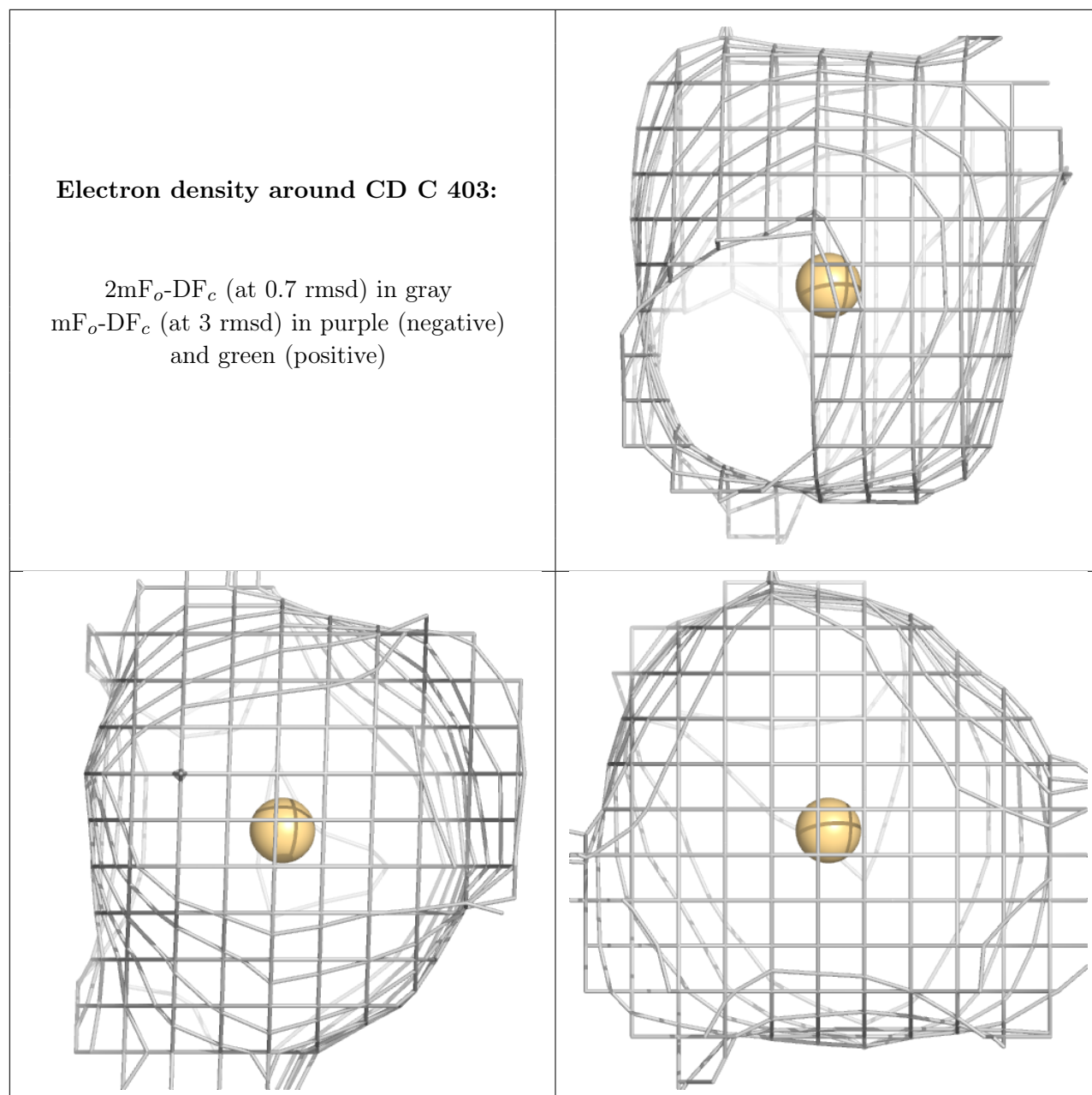




**Electron density around CD A 402:**

$2mF_o-DF_c$  (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.