



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 9, 2025 – 02:11 PM JST

PDB ID : 8ZYB / pdb\_00008zyb  
Title : a-KG dependent oxygenase  
Authors : Fan, A.; Fan, S.; Wu, M.; Sun, Z.  
Deposited on : 2024-06-17  
Resolution : 2.58 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](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:

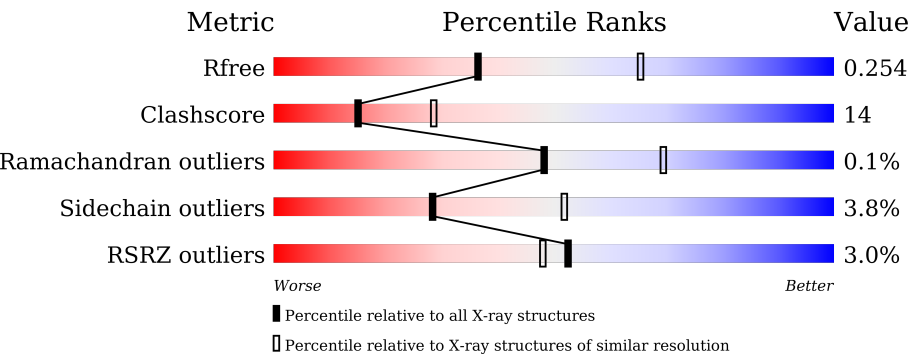
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

# 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 2.58 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	4456 (2.60-2.56)
Clashscore	180529	4905 (2.60-2.56)
Ramachandran outliers	177936	4847 (2.60-2.56)
Sidechain outliers	177891	4847 (2.60-2.56)
RSRZ outliers	164620	4456 (2.60-2.56)

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	282	
1	B	282	
1	C	282	
1	D	282	
1	E	282	
1	F	282	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	282	<div><div></div><div>5%</div><div>67%</div><div>27%</div><div></div><div></div></div>
1	H	282	<div><div></div><div>6%</div><div>64%</div><div>27%</div><div>8%</div><div></div></div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 17100 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 DhiD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	276	Total	C	N	O	S	0	0	0
			2159	1367	379	404	9			
1	A	270	Total	C	N	O	S	0	0	0
			2106	1333	372	392	9			
1	B	269	Total	C	N	O	S	0	0	0
			2101	1331	369	392	9			
1	C	278	Total	C	N	O	S	0	0	0
			2169	1374	380	406	9			
1	D	269	Total	C	N	O	S	0	0	0
			2105	1335	371	390	9			
1	F	278	Total	C	N	O	S	0	0	0
			2169	1374	380	406	9			
1	G	270	Total	C	N	O	S	0	0	0
			2120	1345	372	394	9			
1	H	259	Total	C	N	O	S	0	0	0
			2026	1283	357	377	9			

- Molecule 2 is FE (III) ION (CCD ID: FE) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total 1	Fe 1	0	0
2	H	1	Total 1	Fe 1	0	0

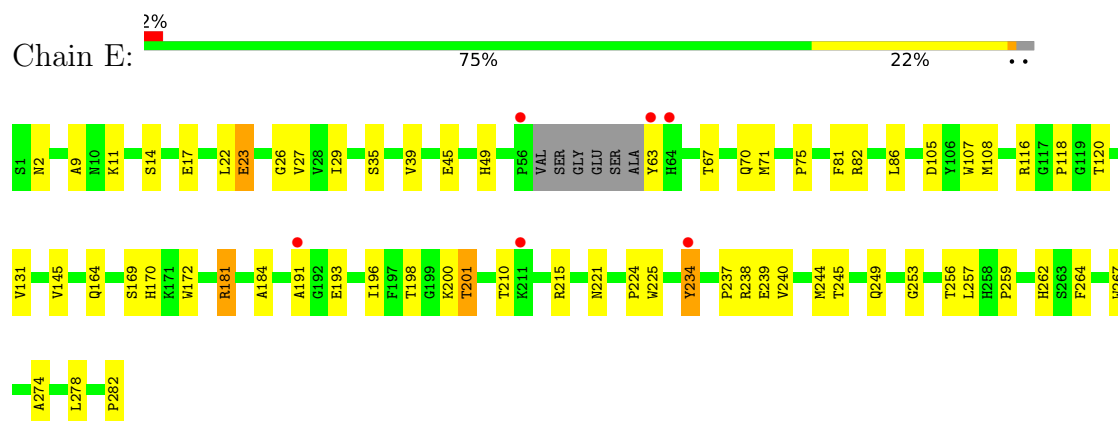
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	29	Total 29	O 29	0	0
3	A	27	Total 27	O 27	0	0
3	B	24	Total 24	O 24	0	0
3	C	11	Total 11	O 11	0	0
3	D	15	Total 15	O 15	0	0
3	F	14	Total 14	O 14	0	0
3	G	14	Total 14	O 14	0	0
3	H	3	Total 3	O 3	0	0

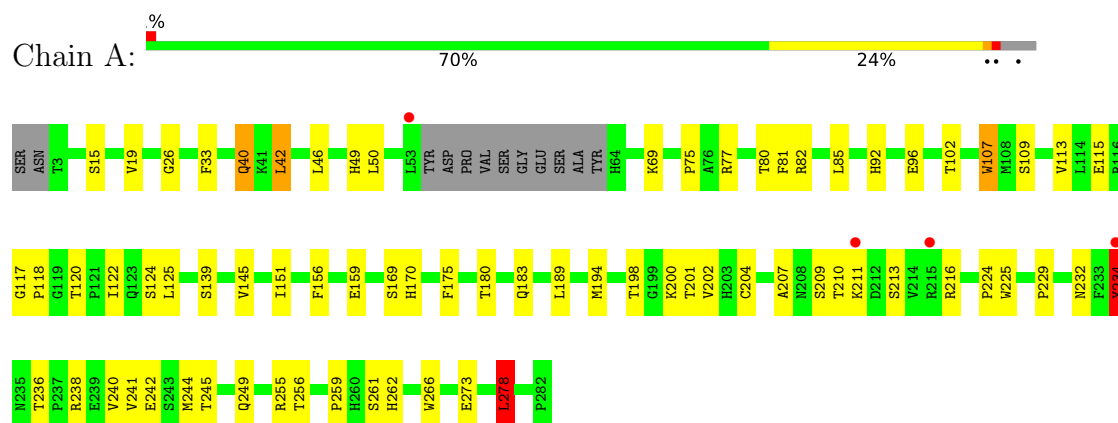
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

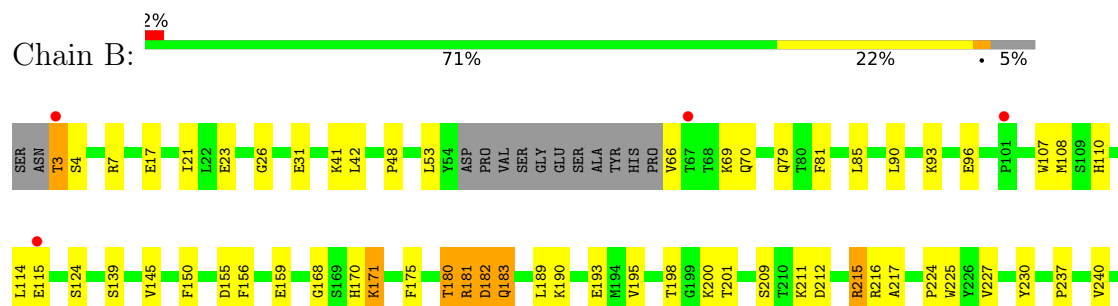
#### • Molecule 1: DhiD

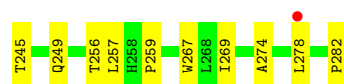


#### • Molecule 1: DhiD

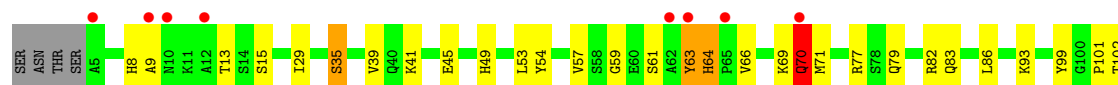


#### • Molecule 1: DhiD

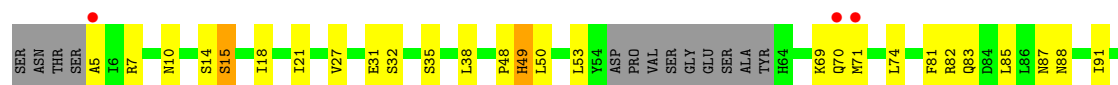




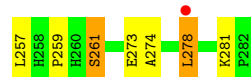
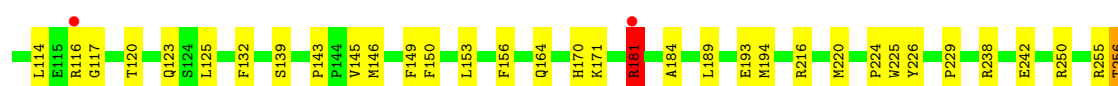
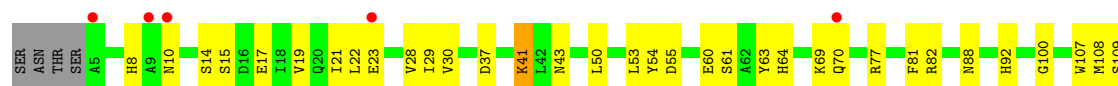
• Molecule 1: DhiD



• Molecule 1: DhiD

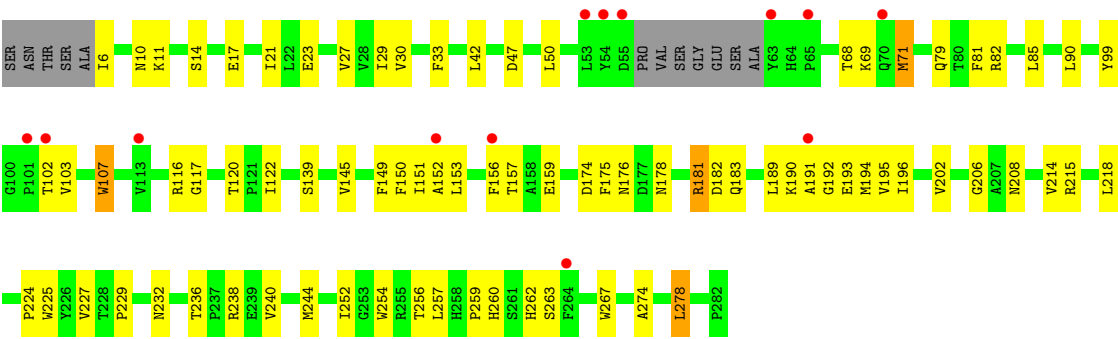


• Molecule 1: DhiD

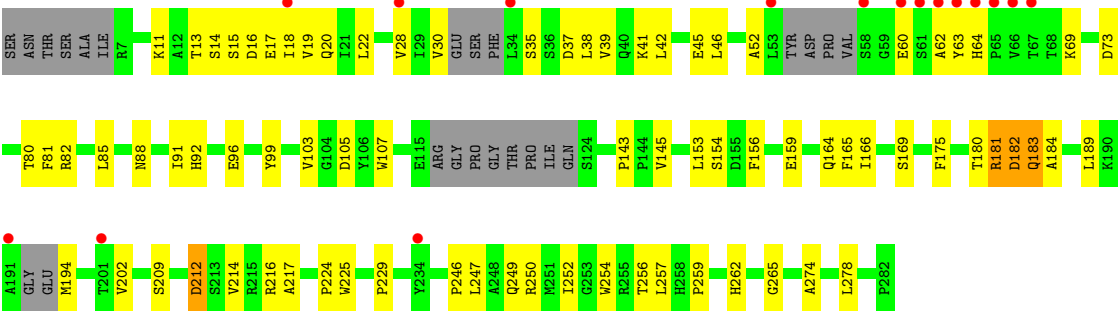


• Molecule 1: DhiD





• Molecule 1: DhiD





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	368.94Å 78.25Å 81.91Å 90.00° 98.02° 90.00°	Depositor
Resolution (Å)	30.57 – 2.58 30.57 – 2.58	Depositor EDS
% Data completeness (in resolution range)	97.5 (30.57-2.58) 97.5 (30.57-2.58)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 2.57Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.203 , 0.257 0.203 , 0.254	Depositor DCC
$R_{free}$ test set	70771 reflections (2.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtriage
Anisotropy	0.126	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 49.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17100	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.33% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: FE

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	0.29	0/2164	0.64	6/2945 (0.2%)
1	B	0.28	0/2158	0.61	4/2936 (0.1%)
1	C	0.40	2/2231 (0.1%)	0.75	12/3039 (0.4%)
1	D	0.31	0/2164	0.61	5/2945 (0.2%)
1	E	0.25	0/2220	0.63	4/3023 (0.1%)
1	F	0.31	1/2231 (0.0%)	0.61	7/3039 (0.2%)
1	G	0.33	0/2180	0.60	0/2967
1	H	0.51	1/2079 (0.0%)	0.74	5/2824 (0.2%)
All	All	0.35	4/17427 (0.0%)	0.65	43/23718 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	3
1	C	0	2
1	D	0	1
1	E	0	1
1	G	0	2
1	H	0	1
All	All	0	10

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	143	PRO	CA-C	-12.74	1.43	1.52
1	F	143	PRO	CA-C	-6.87	1.47	1.52
1	C	215	ARG	CD-NE	6.43	1.55	1.46
1	C	181	ARG	CB-CG	-5.02	1.37	1.52

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	70	GLN	CB-CG-CD	12.48	133.82	112.60
1	E	22	LEU	CA-C-N	-12.28	103.45	122.60
1	E	22	LEU	C-N-CA	-12.28	103.45	122.60
1	H	143	PRO	O-C-N	-10.99	115.98	121.15
1	A	216	ARG	CB-CG-CD	-10.39	87.40	111.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	181	ARG	Sidechain
1	B	215	ARG	Sidechain
1	B	216	ARG	Sidechain
1	C	181	ARG	Sidechain
1	E	181	ARG	Sidechain

## 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	2106	0	2031	58	1
1	B	2101	0	2026	48	1
1	C	2169	0	2082	61	0
1	D	2105	0	2028	60	1
1	E	2159	0	2074	57	1
1	F	2169	0	2082	64	0
1	G	2120	0	2036	65	0
1	H	2026	0	1945	59	0
2	A	1	0	0	0	0
2	B	1	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
2	G	1	0	0	0	0
2	H	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	27	0	0	3	0
3	B	24	0	0	1	0
3	C	11	0	0	0	0
3	D	15	0	0	1	0
3	E	29	0	0	6	0
3	F	14	0	0	7	0
3	G	14	0	0	2	0
3	H	3	0	0	1	0
All	All	17100	0	16304	449	2

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 14.

The worst 5 of 449 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:189:LEU:HD23	1:G:193:GLU:CB	1.74	1.18
1:G:189:LEU:HD23	1:G:193:GLU:HB2	1.32	1.08
1:A:118:PRO:HB3	1:A:211:LYS:HA	1.40	1.03
1:C:163:THR:H	1:C:215:ARG:HH22	1.00	0.98
1:F:123:GLN:O	1:F:181:ARG:NH2	1.98	0.96

All (2) 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:A:40:GLN:OE1	1:B:79:GLN:NE2[1_545]	1.74	0.46
1:E:63:TYR:O	1:D:270:ARG:NH1[1_565]	2.18	0.02

## 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	266/282 (94%)	254 (96%)	12 (4%)	0	100	100
1	B	265/282 (94%)	246 (93%)	19 (7%)	0	100	100
1	C	276/282 (98%)	253 (92%)	23 (8%)	0	100	100
1	D	265/282 (94%)	248 (94%)	17 (6%)	0	100	100
1	E	272/282 (96%)	259 (95%)	13 (5%)	0	100	100
1	F	276/282 (98%)	260 (94%)	16 (6%)	0	100	100
1	G	266/282 (94%)	253 (95%)	13 (5%)	0	100	100
1	H	249/282 (88%)	228 (92%)	19 (8%)	2 (1%)	16	33
All	All	2135/2256 (95%)	2001 (94%)	132 (6%)	2 (0%)	48	69

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	183	GLN
1	H	212	ASP

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	227/237 (96%)	220 (97%)	7 (3%)	35	59
1	B	226/237 (95%)	216 (96%)	10 (4%)	24	46
1	C	233/237 (98%)	226 (97%)	7 (3%)	36	60
1	D	226/237 (95%)	215 (95%)	11 (5%)	21	41
1	E	233/237 (98%)	228 (98%)	5 (2%)	48	71
1	F	233/237 (98%)	226 (97%)	7 (3%)	36	60
1	G	228/237 (96%)	219 (96%)	9 (4%)	27	51
1	H	218/237 (92%)	205 (94%)	13 (6%)	16	33
All	All	1824/1896 (96%)	1755 (96%)	69 (4%)	28	52

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	35	SER
1	H	38	LEU
1	H	182	ASP
1	C	64	HIS
1	C	63	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	79	GLN
1	H	208	ASN
1	D	20	GLN
1	G	70	GLN
1	C	276	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	270/282 (95%)	-0.12	4 (1%) 71 68	31, 56, 86, 104	0
1	B	269/282 (95%)	-0.07	5 (1%) 66 62	35, 59, 92, 120	0
1	C	278/282 (98%)	0.06	10 (3%) 46 42	41, 65, 95, 125	0
1	D	269/282 (95%)	0.08	3 (1%) 77 75	38, 69, 103, 125	0
1	E	276/282 (97%)	-0.28	6 (2%) 62 58	28, 47, 76, 118	0
1	F	278/282 (98%)	0.10	8 (2%) 54 49	37, 69, 99, 123	0
1	G	270/282 (95%)	0.28	13 (4%) 36 32	37, 77, 116, 139	0
1	H	259/282 (91%)	0.63	16 (6%) 28 24	30, 93, 139, 165	0
All	All	2169/2256 (96%)	0.08	65 (2%) 52 48	28, 65, 112, 165	0

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	63	TYR	6.4
1	E	234	TYR	4.7
1	D	70	GLN	4.5
1	H	62	ALA	4.4
1	A	234	TYR	4.2

### 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

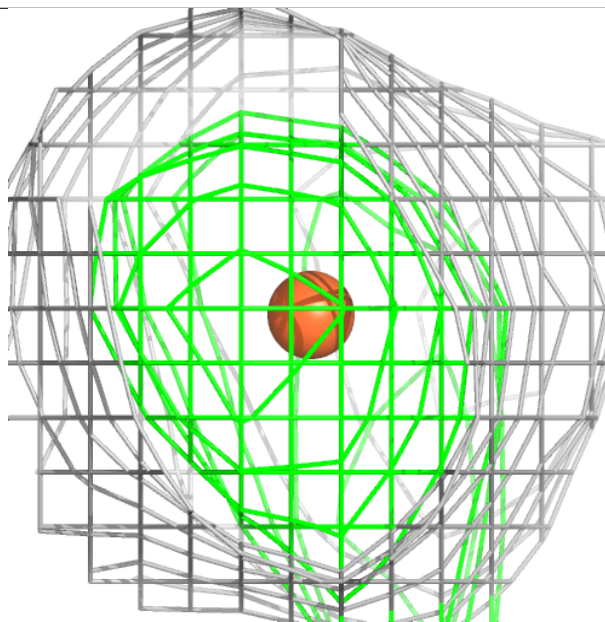
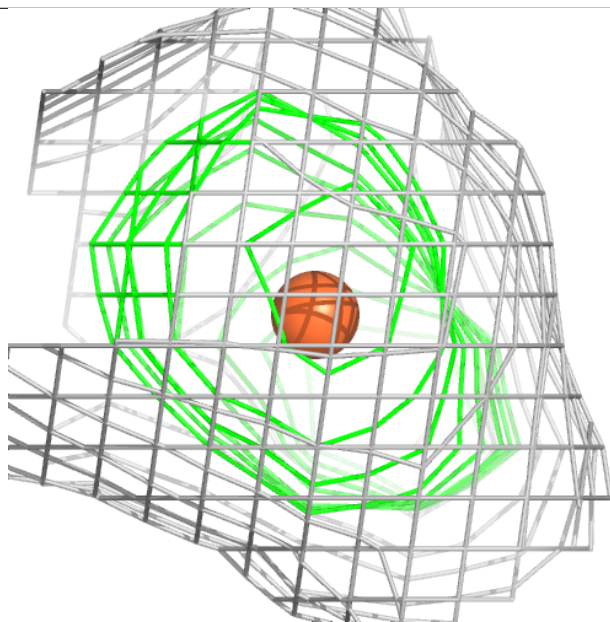
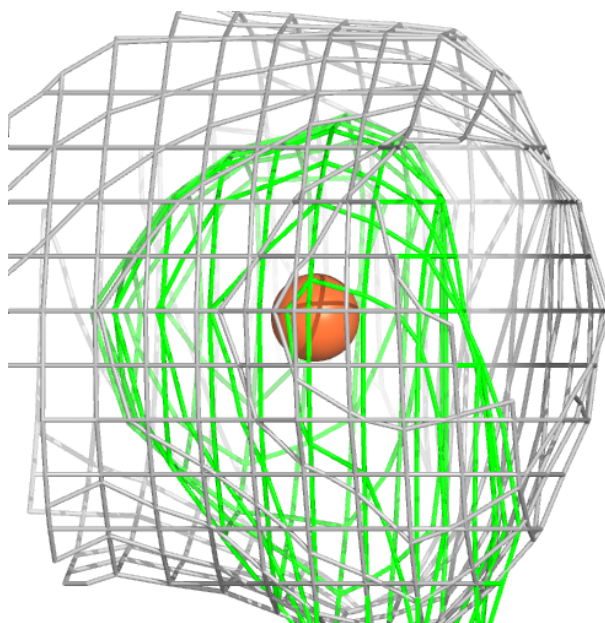
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
2	FE	H	301	1/1	0.83	0.24	144,144,144,144	0
2	FE	G	301	1/1	0.95	0.15	66,66,66,66	0
2	FE	D	301	1/1	0.97	0.09	53,53,53,53	0
2	FE	F	301	1/1	0.98	0.11	55,55,55,55	0
2	FE	C	301	1/1	0.98	0.11	54,54,54,54	0
2	FE	A	301	1/1	0.98	0.11	49,49,49,49	0
2	FE	B	301	1/1	0.99	0.08	48,48,48,48	0
2	FE	E	301	1/1	0.99	0.12	44,44,44,44	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.

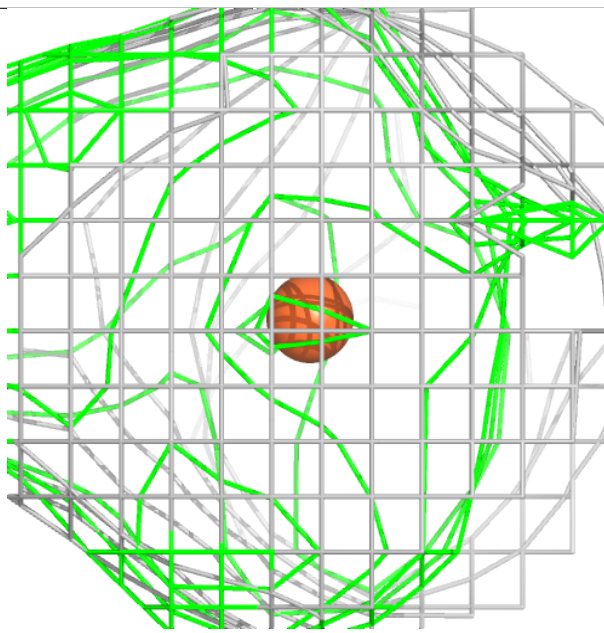
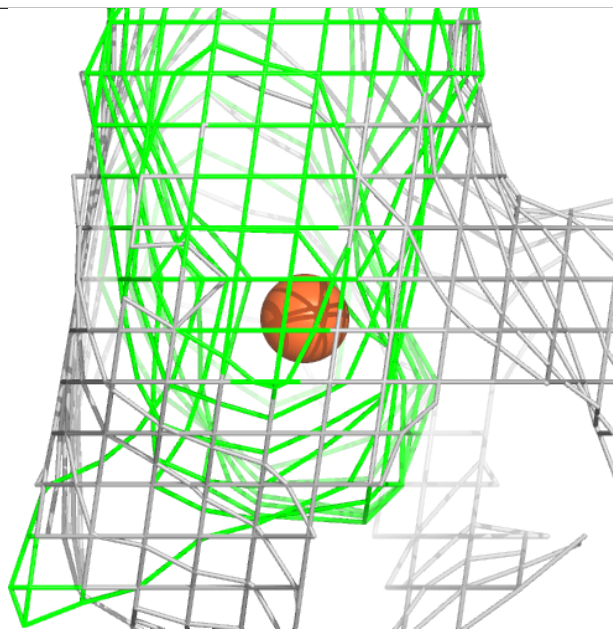
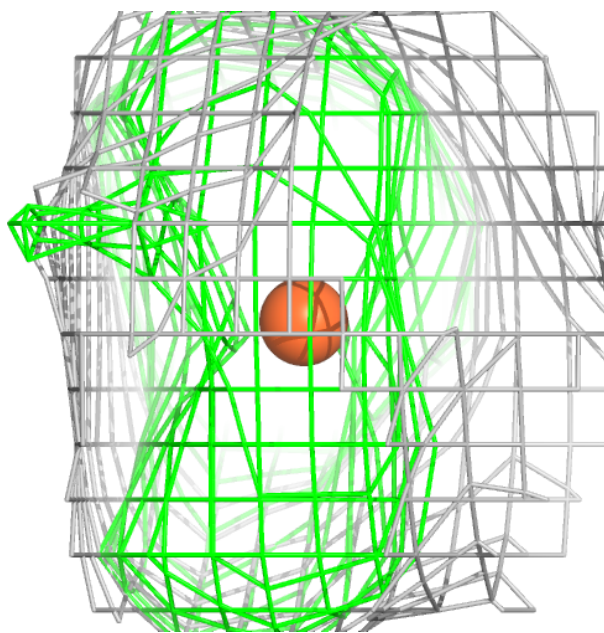
**Electron density around FE H 301:**

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



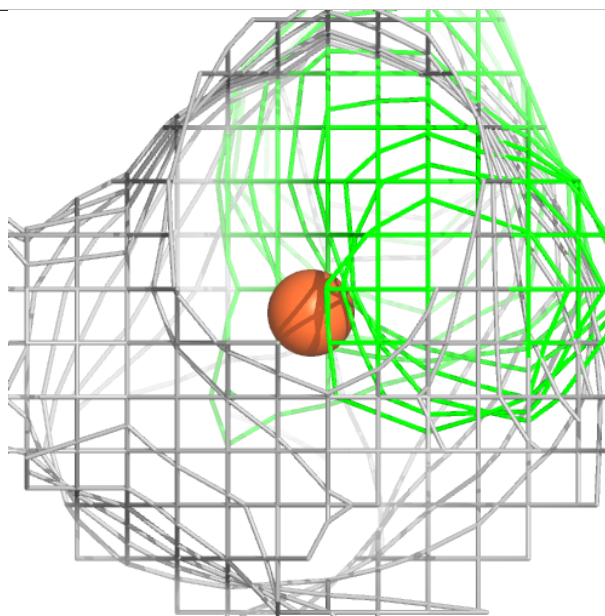
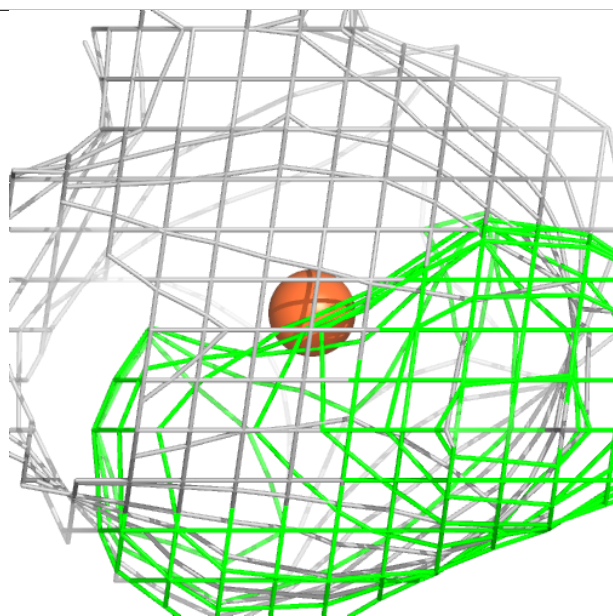
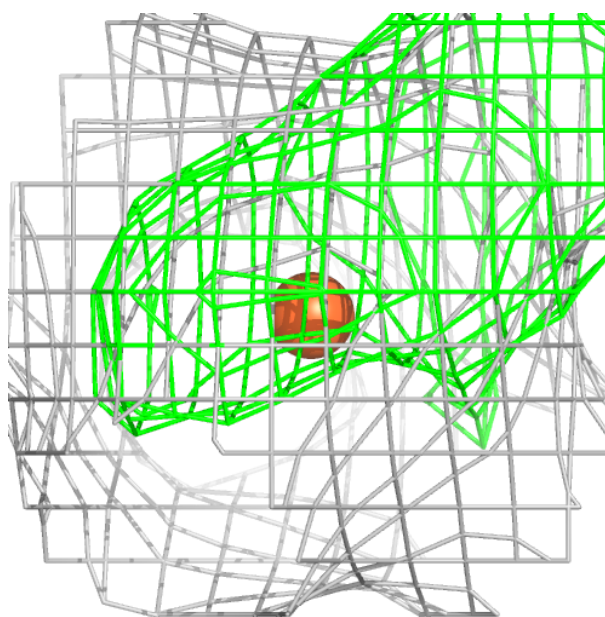
**Electron density around FE G 301:**

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



**Electron density around FE D 301:**

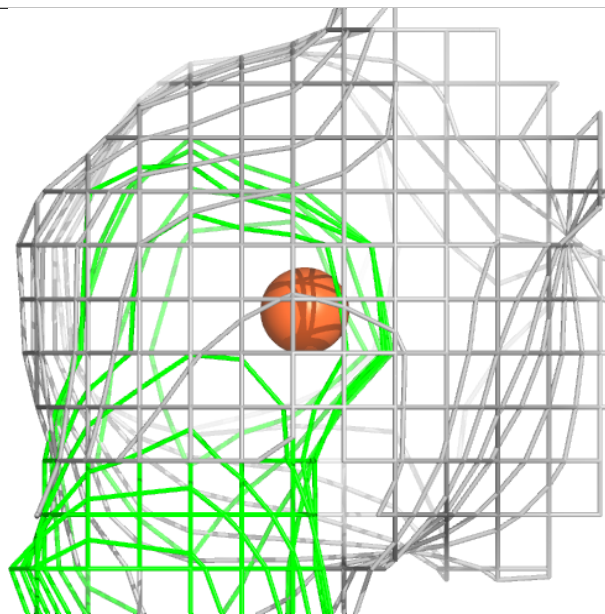
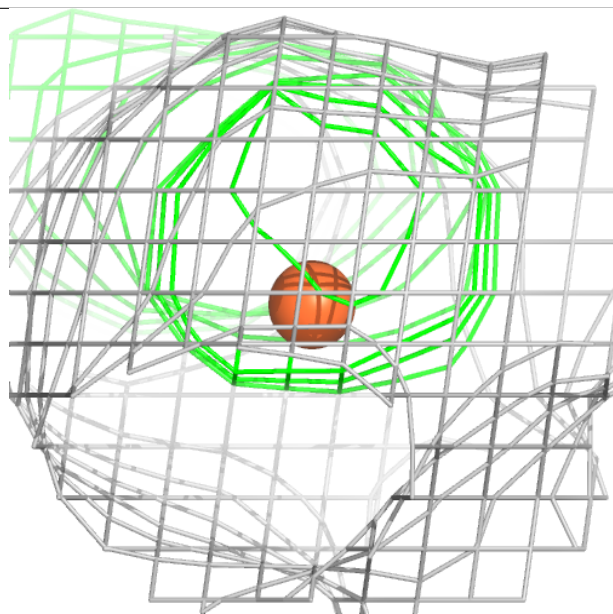
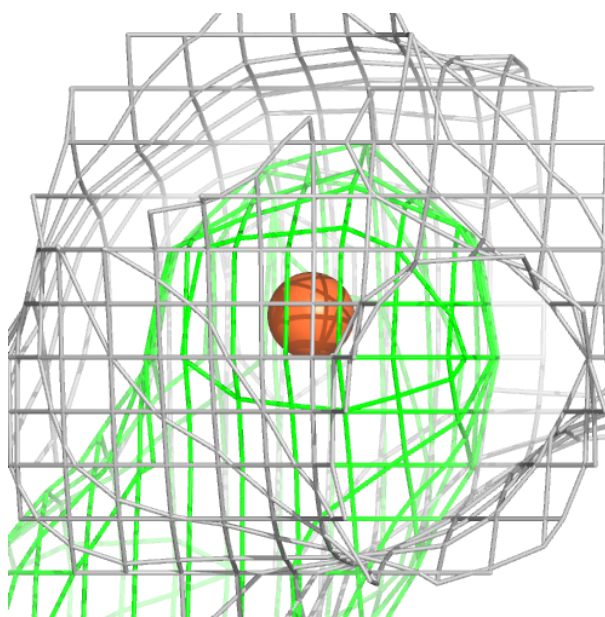
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





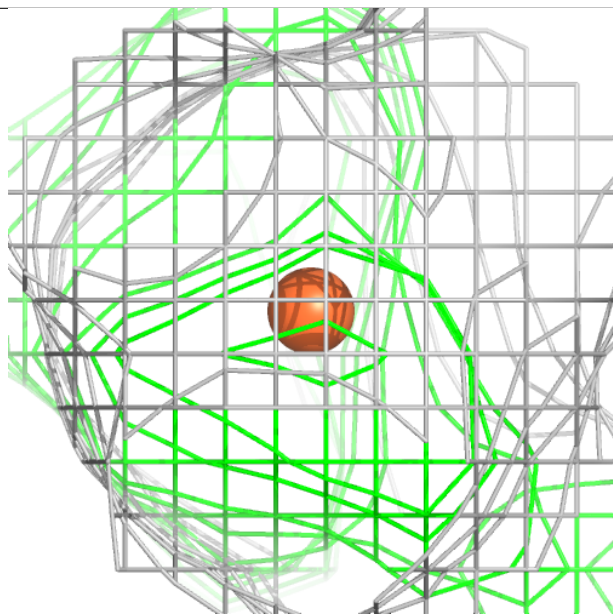
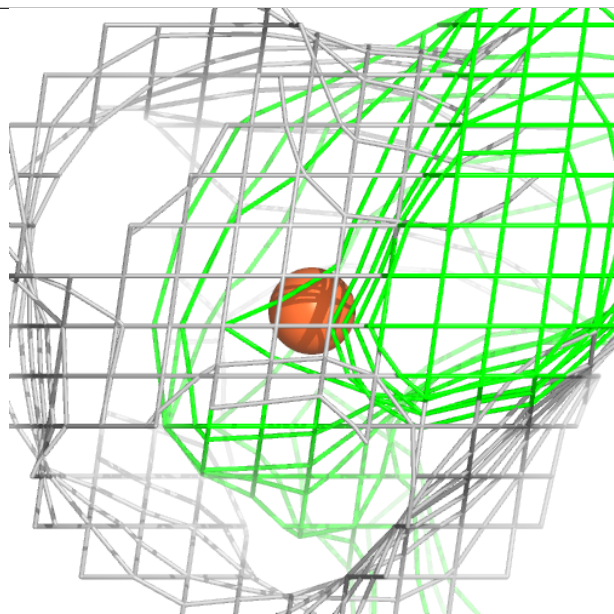
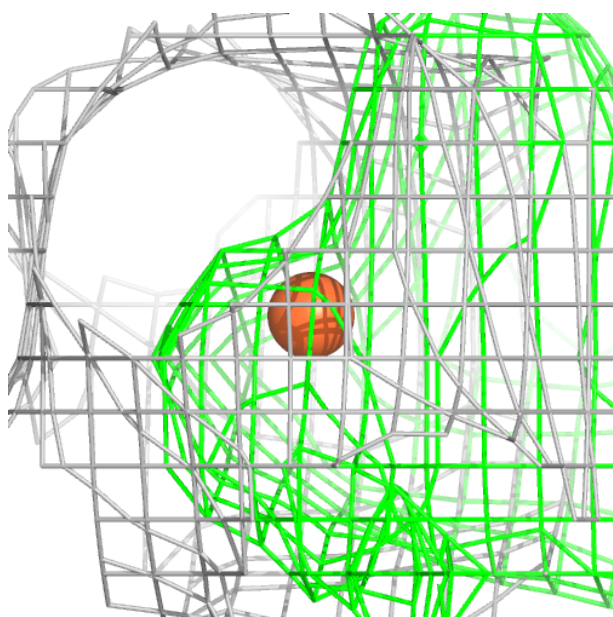
**Electron density around FE F 301:**

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



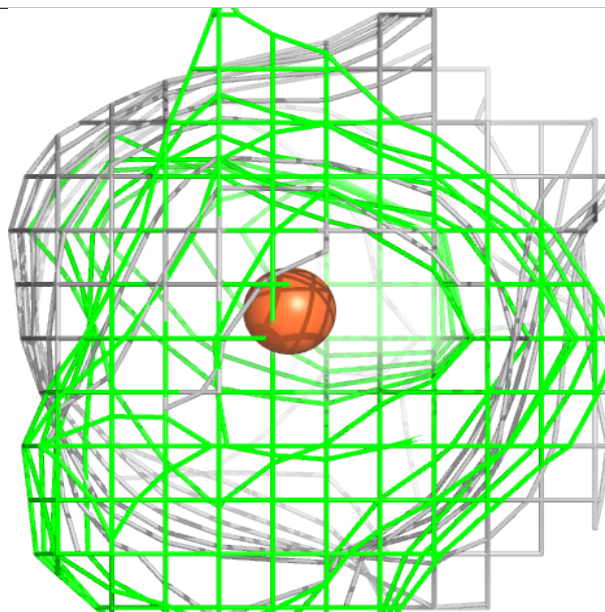
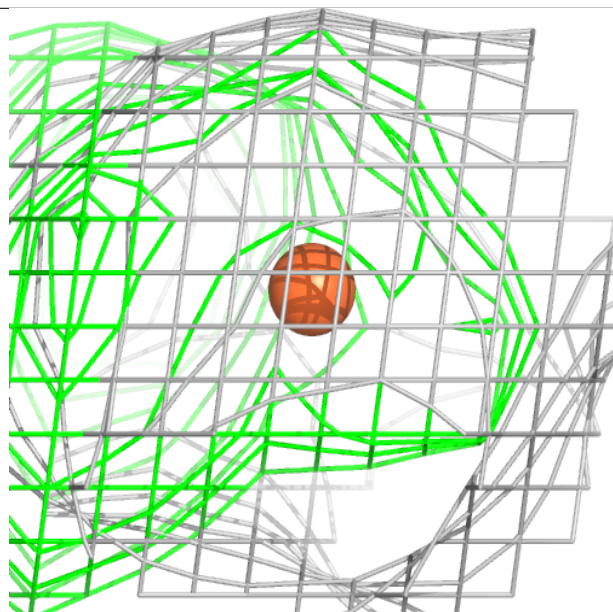
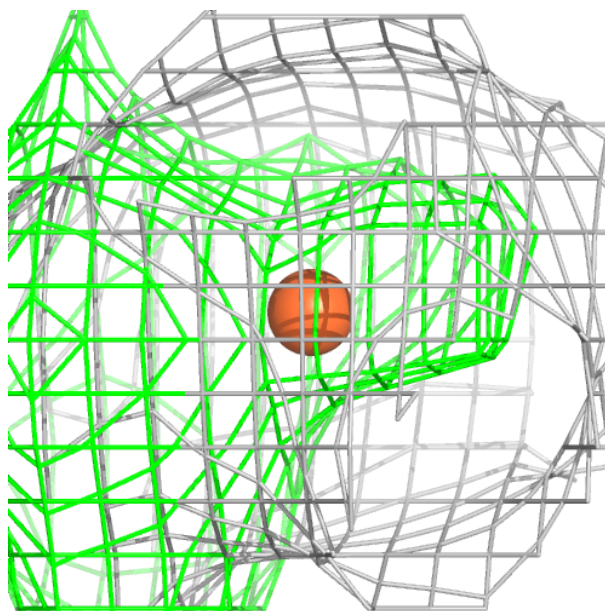
**Electron density around FE C 301:**

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



**Electron density around FE A 301:**

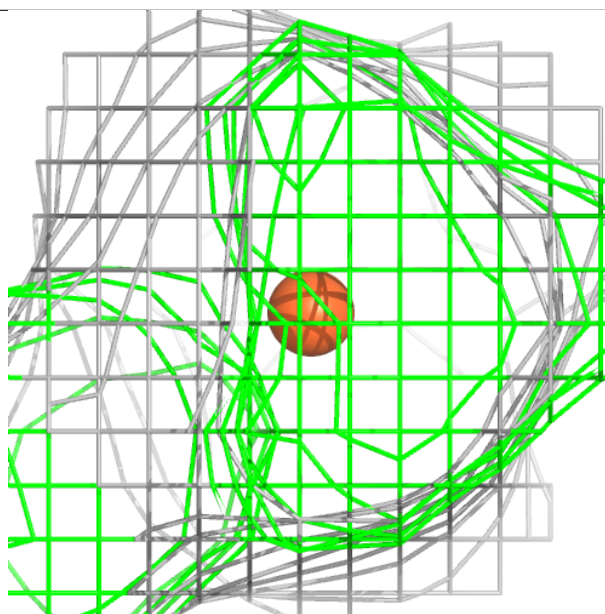
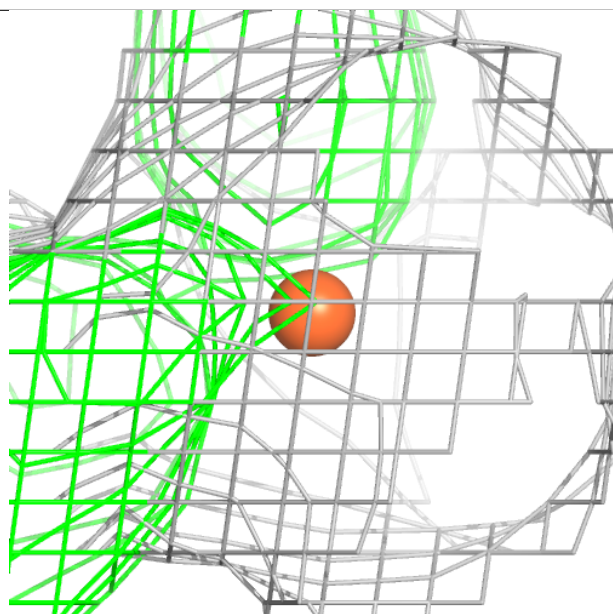
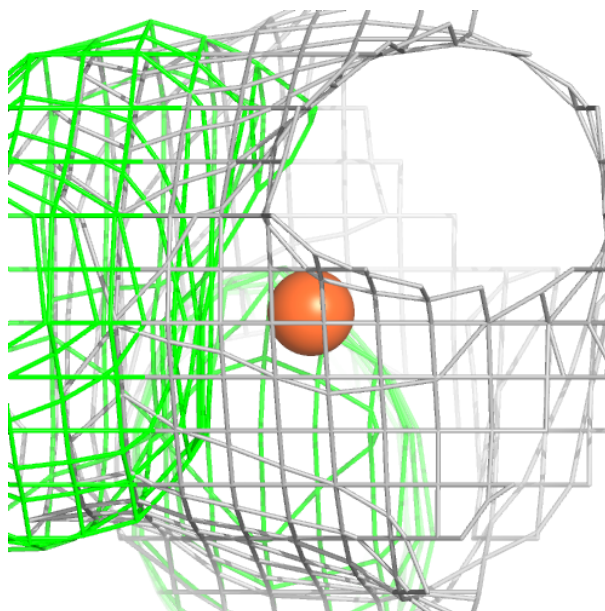
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



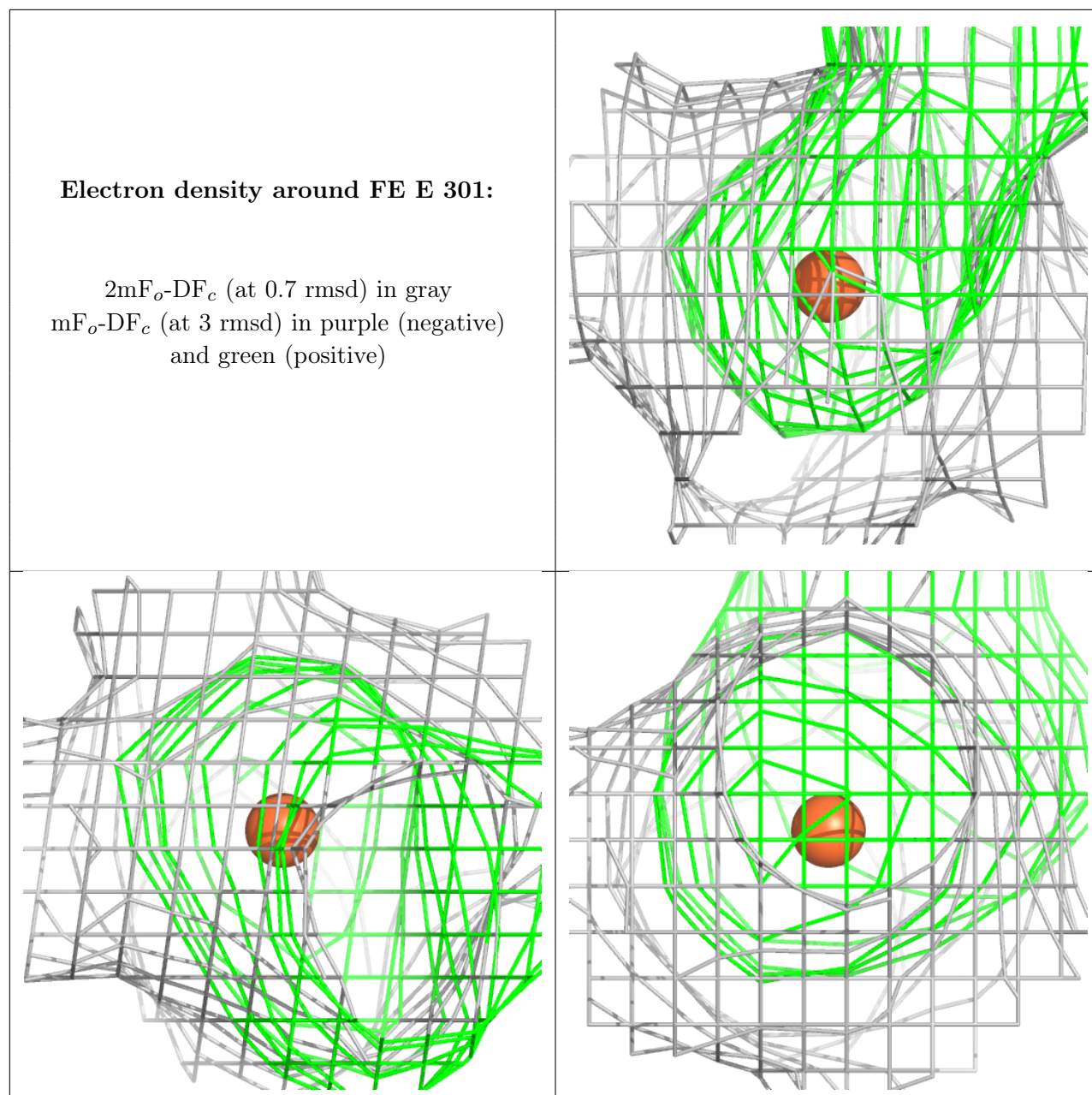


**Electron density around FE B 301:**

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







## 6.5 Other polymers ⓘ

There are no such residues in this entry.