



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2023 – 01:23 pm BST

PDB ID : 8CKP  
Title : X-ray structure of the crystallization-prone form of subfamily III haloalkane dehalogenase DhmeA from *Haloferax mediterranei*  
Authors : Marek, M.; Chmelova, K.; Schenkmyerova, A.; Croll, T.; Read, R.J.; Diederichs, K.  
Deposited on : 2023-02-16  
Resolution : 3.31 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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

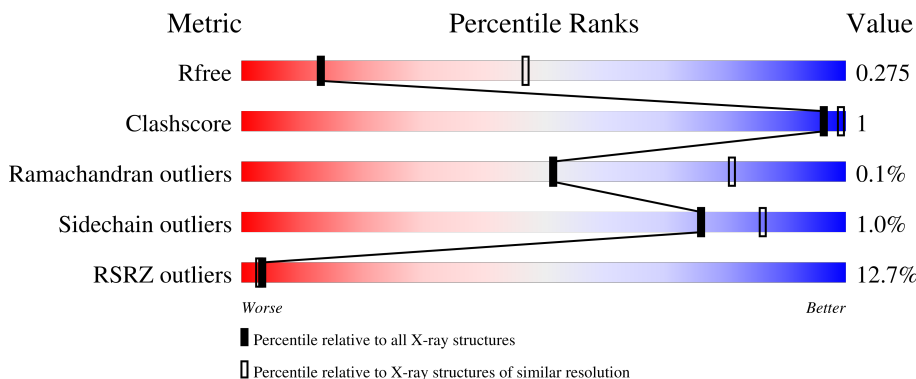
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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	311	
1	B	311	
1	C	311	
1	D	311	
1	E	311	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	311	
1	G	311	
1	H	311	
1	I	311	
1	J	311	

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
2	CL	F	1001	-	-	-	X

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 23884 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 Alpha/beta fold hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	302	Total 2442	C 1548	N 424	O 461	S 9	0	0	0
1	F	292	Total 2349	C 1482	N 410	O 450	S 7	0	0	0
1	G	300	Total 2433	C 1544	N 423	O 457	S 9	0	0	0
1	H	291	Total 2348	C 1485	N 409	O 447	S 7	0	0	0
1	I	300	Total 2428	C 1541	N 421	O 457	S 9	0	0	0
1	J	286	Total 2314	C 1465	N 405	O 437	S 7	0	0	0
1	A	287	Total 2320	C 1467	N 407	O 439	S 7	0	0	0
1	B	288	Total 2328	C 1472	N 408	O 441	S 7	0	0	0
1	C	305	Total 2468	C 1563	N 431	O 465	S 9	0	0	0
1	E	303	Total 2452	C 1554	N 427	O 462	S 9	0	0	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	GLY	deletion	UNP I3R766
D	?	-	GLY	deletion	UNP I3R766
D	306	HIS	-	expression tag	UNP I3R766
D	307	HIS	-	expression tag	UNP I3R766
D	308	HIS	-	expression tag	UNP I3R766
D	309	HIS	-	expression tag	UNP I3R766
D	310	HIS	-	expression tag	UNP I3R766
D	311	HIS	-	expression tag	UNP I3R766
F	?	-	GLY	deletion	UNP I3R766

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	?	-	GLY	deletion	UNP I3R766
F	306	HIS	-	expression tag	UNP I3R766
F	307	HIS	-	expression tag	UNP I3R766
F	308	HIS	-	expression tag	UNP I3R766
F	309	HIS	-	expression tag	UNP I3R766
F	310	HIS	-	expression tag	UNP I3R766
F	311	HIS	-	expression tag	UNP I3R766
G	?	-	GLY	deletion	UNP I3R766
G	?	-	GLY	deletion	UNP I3R766
G	306	HIS	-	expression tag	UNP I3R766
G	307	HIS	-	expression tag	UNP I3R766
G	308	HIS	-	expression tag	UNP I3R766
G	309	HIS	-	expression tag	UNP I3R766
G	310	HIS	-	expression tag	UNP I3R766
G	311	HIS	-	expression tag	UNP I3R766
H	?	-	GLY	deletion	UNP I3R766
H	?	-	GLY	deletion	UNP I3R766
H	306	HIS	-	expression tag	UNP I3R766
H	307	HIS	-	expression tag	UNP I3R766
H	308	HIS	-	expression tag	UNP I3R766
H	309	HIS	-	expression tag	UNP I3R766
H	310	HIS	-	expression tag	UNP I3R766
H	311	HIS	-	expression tag	UNP I3R766
I	?	-	GLY	deletion	UNP I3R766
I	?	-	GLY	deletion	UNP I3R766
I	306	HIS	-	expression tag	UNP I3R766
I	307	HIS	-	expression tag	UNP I3R766
I	308	HIS	-	expression tag	UNP I3R766
I	309	HIS	-	expression tag	UNP I3R766
I	310	HIS	-	expression tag	UNP I3R766
I	311	HIS	-	expression tag	UNP I3R766
J	?	-	GLY	deletion	UNP I3R766
J	?	-	GLY	deletion	UNP I3R766
J	306	HIS	-	expression tag	UNP I3R766
J	307	HIS	-	expression tag	UNP I3R766
J	308	HIS	-	expression tag	UNP I3R766
J	309	HIS	-	expression tag	UNP I3R766
J	310	HIS	-	expression tag	UNP I3R766
J	311	HIS	-	expression tag	UNP I3R766
A	?	-	GLY	deletion	UNP I3R766
A	?	-	GLY	deletion	UNP I3R766
A	306	HIS	-	expression tag	UNP I3R766

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	HIS	-	expression tag	UNP I3R766
A	308	HIS	-	expression tag	UNP I3R766
A	309	HIS	-	expression tag	UNP I3R766
A	310	HIS	-	expression tag	UNP I3R766
A	311	HIS	-	expression tag	UNP I3R766
B	?	-	GLY	deletion	UNP I3R766
B	?	-	GLY	deletion	UNP I3R766
B	306	HIS	-	expression tag	UNP I3R766
B	307	HIS	-	expression tag	UNP I3R766
B	308	HIS	-	expression tag	UNP I3R766
B	309	HIS	-	expression tag	UNP I3R766
B	310	HIS	-	expression tag	UNP I3R766
B	311	HIS	-	expression tag	UNP I3R766
C	?	-	GLY	deletion	UNP I3R766
C	?	-	GLY	deletion	UNP I3R766
C	306	HIS	-	expression tag	UNP I3R766
C	307	HIS	-	expression tag	UNP I3R766
C	308	HIS	-	expression tag	UNP I3R766
C	309	HIS	-	expression tag	UNP I3R766
C	310	HIS	-	expression tag	UNP I3R766
C	311	HIS	-	expression tag	UNP I3R766
E	?	-	GLY	deletion	UNP I3R766
E	?	-	GLY	deletion	UNP I3R766
E	306	HIS	-	expression tag	UNP I3R766
E	307	HIS	-	expression tag	UNP I3R766
E	308	HIS	-	expression tag	UNP I3R766
E	309	HIS	-	expression tag	UNP I3R766
E	310	HIS	-	expression tag	UNP I3R766
E	311	HIS	-	expression tag	UNP I3R766

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	F	1	Total Cl 1 1	0	0

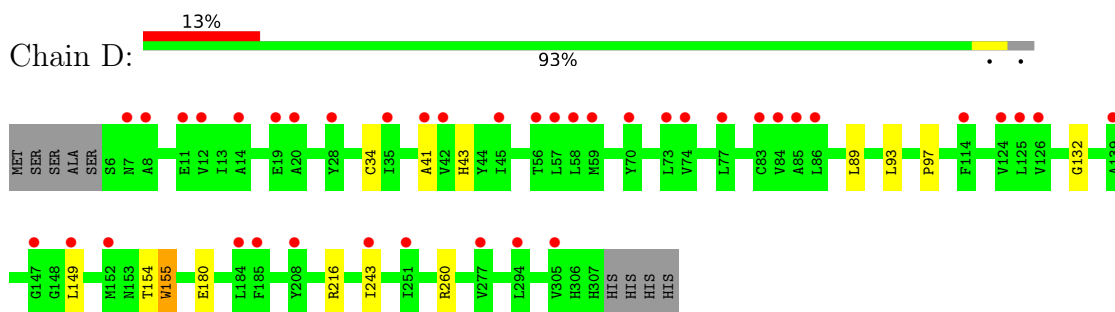
- Molecule 3 is water.

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

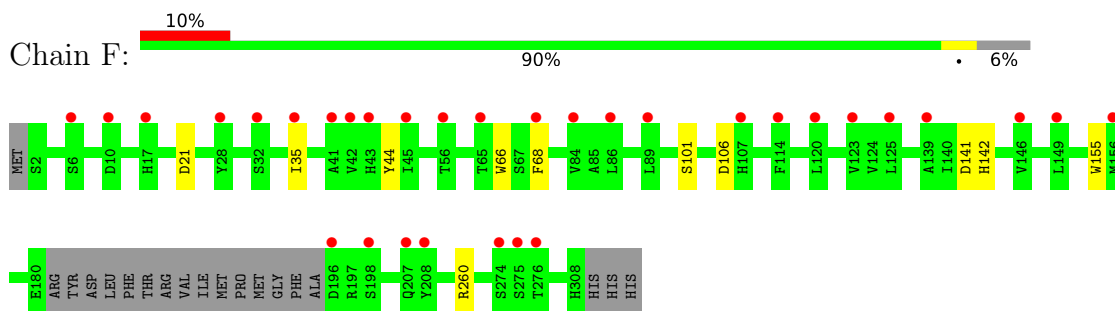
### 3 Residue-property plots [i](#)

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.

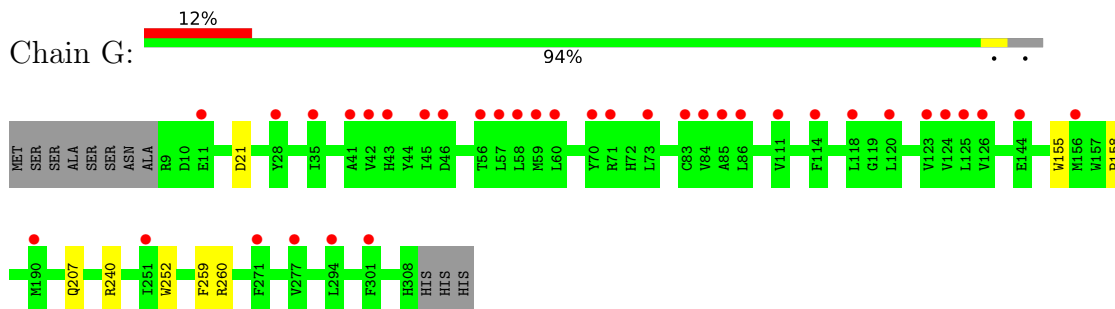
- Molecule 1: Alpha/beta fold hydrolase



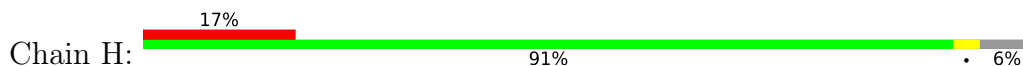
- Molecule 1: Alpha/beta fold hydrolase

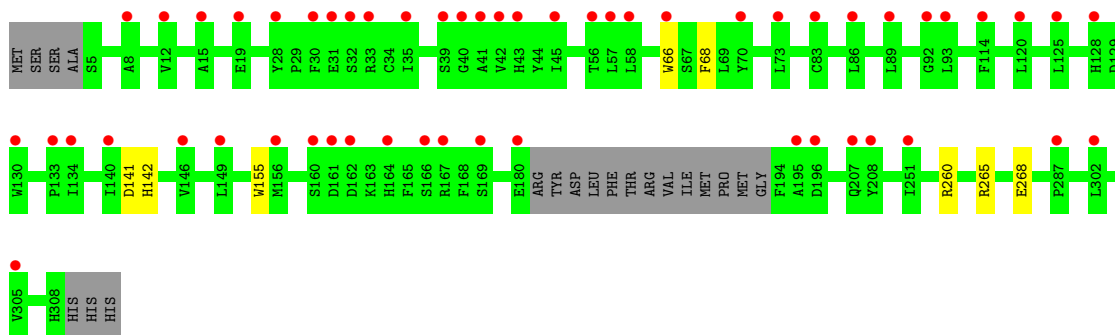


- Molecule 1: Alpha/beta fold hydrolase

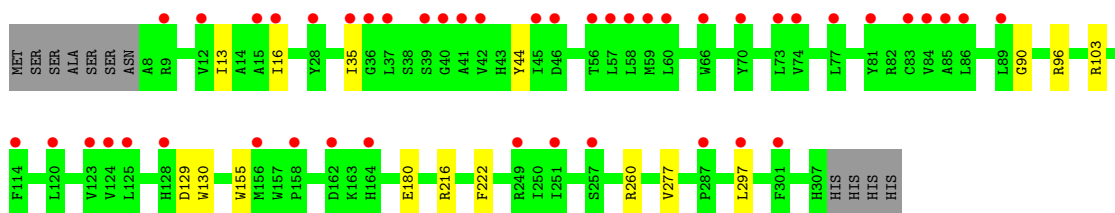
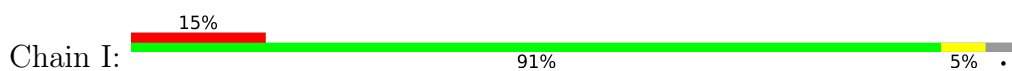


- Molecule 1: Alpha/beta fold hydrolase

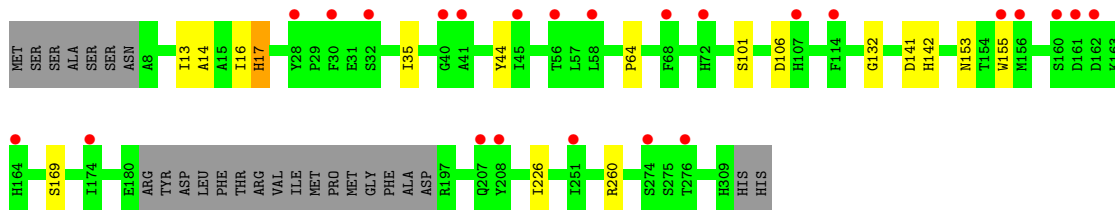
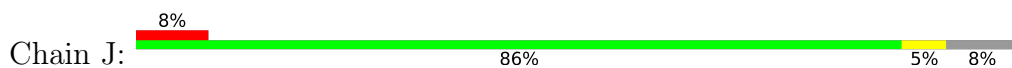




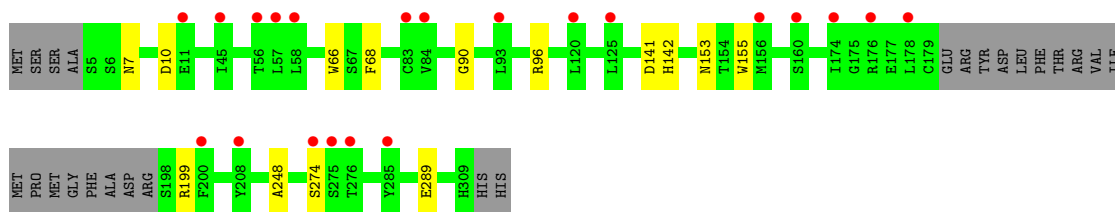
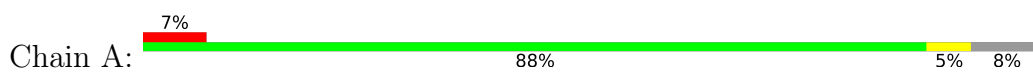
• Molecule 1: Alpha/beta fold hydrolase



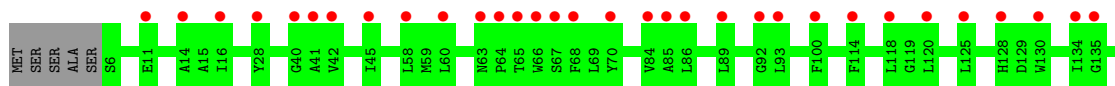
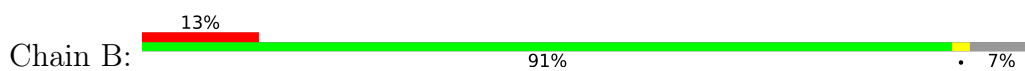
• Molecule 1: Alpha/beta fold hydrolase



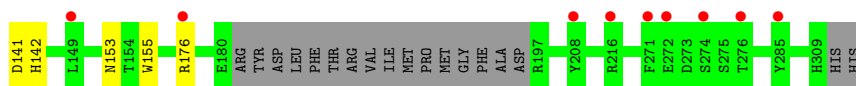
• Molecule 1: Alpha/beta fold hydrolase



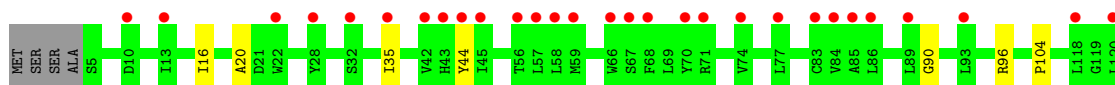
• Molecule 1: Alpha/beta fold hydrolase



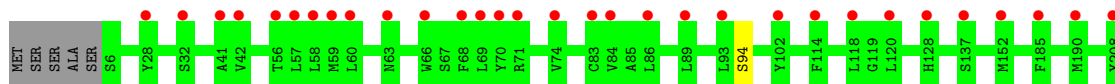




- Molecule 1: Alpha/beta fold hydrolase



- Molecule 1: Alpha/beta fold hydrolase



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	172.76Å 289.90Å 168.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.47 – 3.31 49.47 – 3.31	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.47-3.31) 99.6 (49.47-3.31)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.96 (at 3.33Å)	Xtrriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, $R_{free}$	0.236 , 0.275 0.232 , 0.275	Depositor DCC
$R_{free}$ test set	3153 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	151.5	Xtrriage
Anisotropy	0.239	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 101.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.021 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	23884	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	155.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.27% 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: CL

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	0.25	0/2385	0.55	0/3239
1	B	0.25	0/2393	0.54	0/3250
1	C	0.24	0/2538	0.53	0/3447
1	D	0.25	0/2510	0.54	0/3409
1	E	0.25	0/2521	0.54	0/3424
1	F	0.24	0/2413	0.54	0/3277
1	G	0.25	0/2502	0.54	0/3398
1	H	0.25	0/2413	0.53	0/3277
1	I	0.24	0/2496	0.54	0/3390
1	J	0.24	0/2379	0.54	0/3231
All	All	0.25	0/24550	0.54	0/33342

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	1
1	G	0	1
1	I	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	176	ARG	Sidechain
1	G	21	ASP	Peptide
1	I	103	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	2320	0	2171	6	0
1	B	2328	0	2174	1	0
1	C	2468	0	2312	6	0
1	D	2442	0	2293	9	0
1	E	2452	0	2300	3	0
1	F	2349	0	2191	4	0
1	G	2433	0	2284	3	0
1	H	2348	0	2190	4	0
1	I	2428	0	2282	6	0
1	J	2314	0	2163	6	0
2	F	1	0	0	0	0
3	D	1	0	0	0	0
All	All	23884	0	22360	46	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 1.

The worst 5 of 46 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:154:THR:HG22	1:D:155:TRP:H	1.58	0.69
1:D:132:GLY:HA3	1:D:154:THR:HG21	1.77	0.64
1:J:14:ALA:HA	1:J:17:HIS:CD2	2.33	0.62
1:I:180:GLU:O	1:I:216:ARG:NH1	2.33	0.61
1:D:154:THR:HG22	1:D:155:TRP:N	2.19	0.55

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	283/311 (91%)	267 (94%)	15 (5%)	1 (0%)	34	66
1	B	284/311 (91%)	266 (94%)	17 (6%)	1 (0%)	34	66
1	C	303/311 (97%)	291 (96%)	12 (4%)	0	100	100
1	D	300/311 (96%)	287 (96%)	13 (4%)	0	100	100
1	E	301/311 (97%)	289 (96%)	12 (4%)	0	100	100
1	F	288/311 (93%)	274 (95%)	14 (5%)	0	100	100
1	G	298/311 (96%)	288 (97%)	10 (3%)	0	100	100
1	H	287/311 (92%)	278 (97%)	9 (3%)	0	100	100
1	I	298/311 (96%)	282 (95%)	16 (5%)	0	100	100
1	J	282/311 (91%)	272 (96%)	8 (3%)	2 (1%)	22	55
All	All	2924/3110 (94%)	2794 (96%)	126 (4%)	4 (0%)	51	81

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	153	ASN
1	J	153	ASN
1	A	153	ASN
1	J	132	GLY

### 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	245/266 (92%)	244 (100%)	1 (0%)	91	95
1	B	245/266 (92%)	244 (100%)	1 (0%)	91	95
1	C	260/266 (98%)	257 (99%)	3 (1%)	71	84
1	D	257/266 (97%)	255 (99%)	2 (1%)	81	89
1	E	258/266 (97%)	255 (99%)	3 (1%)	71	84
1	F	248/266 (93%)	245 (99%)	3 (1%)	71	84
1	G	256/266 (96%)	253 (99%)	3 (1%)	71	84
1	H	247/266 (93%)	245 (99%)	2 (1%)	81	89
1	I	255/266 (96%)	252 (99%)	3 (1%)	71	84
1	J	243/266 (91%)	239 (98%)	4 (2%)	62	80
All	All	2514/2660 (94%)	2489 (99%)	25 (1%)	76	86

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	64	PRO
1	A	155	TRP
1	E	260	ARG
1	J	260	ARG
1	B	155	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 1 ligands modelled in this entry, 1 is 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	287/311 (92%)	0.48	21 (7%) 15 15	123, 154, 193, 226	0
1	B	288/311 (92%)	0.72	41 (14%) 2 2	119, 146, 189, 225	0
1	C	305/311 (98%)	0.85	45 (14%) 2 1	119, 149, 194, 211	0
1	D	302/311 (97%)	0.66	40 (13%) 3 2	134, 163, 201, 228	0
1	E	303/311 (97%)	0.66	37 (12%) 4 3	126, 157, 197, 225	0
1	F	292/311 (93%)	0.63	32 (10%) 5 5	108, 141, 205, 226	0
1	G	300/311 (96%)	0.67	36 (12%) 4 3	125, 158, 190, 214	0
1	H	291/311 (93%)	0.79	54 (18%) 1 1	115, 147, 205, 230	0
1	I	300/311 (96%)	0.86	46 (15%) 2 1	121, 149, 181, 221	0
1	J	286/311 (91%)	0.45	24 (8%) 11 11	126, 152, 184, 223	0
All	All	2954/3110 (94%)	0.68	376 (12%) 3 3	108, 152, 197, 230	0

The worst 5 of 376 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	28	TYR	6.9
1	G	57	LEU	6.6
1	G	56	THR	6.1
1	I	84	VAL	6.0
1	G	84	VAL	6.0

### 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](#)

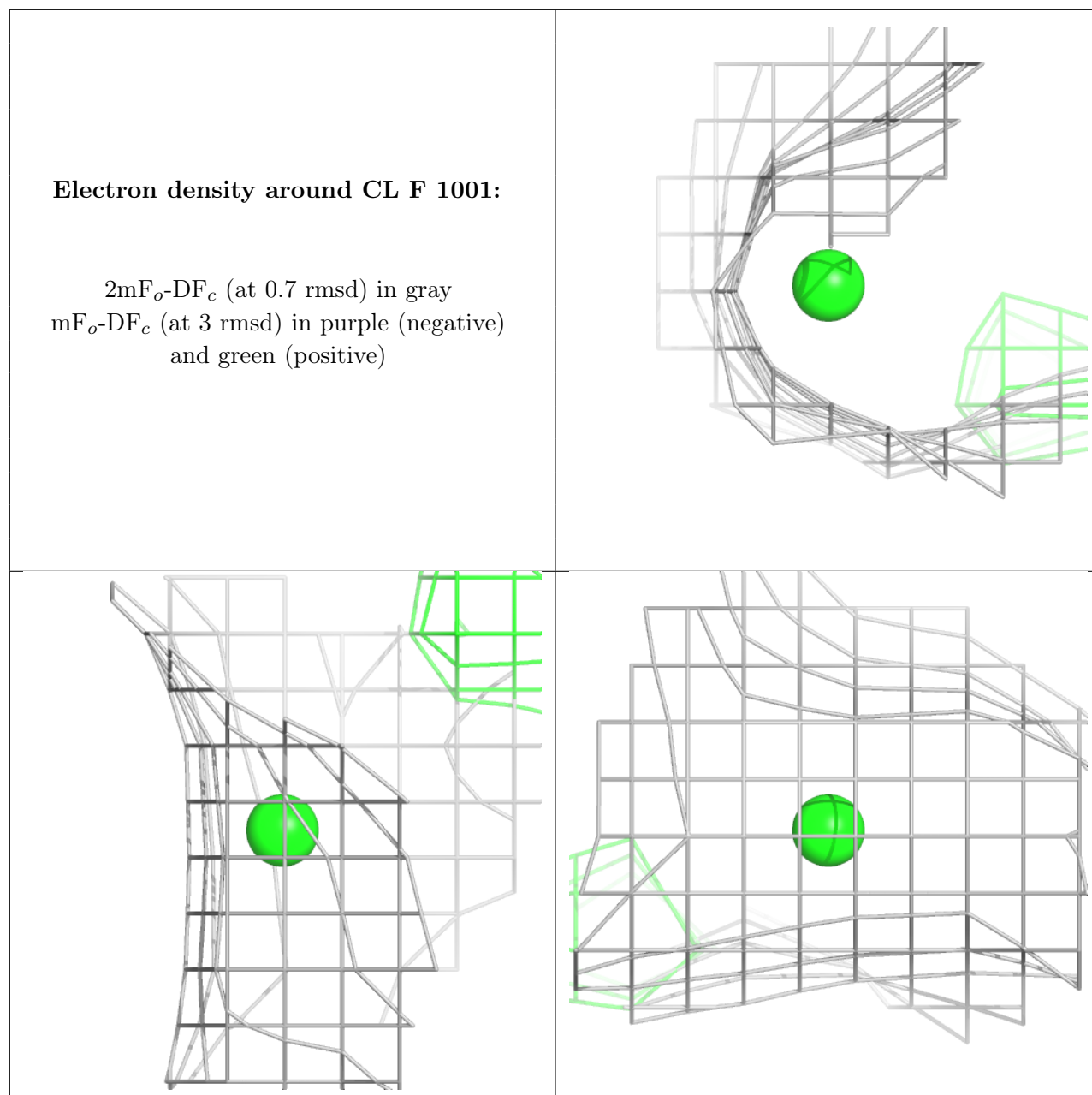
There are no monosaccharides in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
2	CL	F	1001	1/1	0.70	0.47	138,138,138,138	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.