



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 01:50 AM EDT

PDB ID : 7UQU  
Title : Crystal Structure of Mouse Cadherin-23 EC9  
Authors : Sandhu, J.S.; De La Torre, P.; Sotomayor, M.  
Deposited on : 2022-04-20  
Resolution : 1.98 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  
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

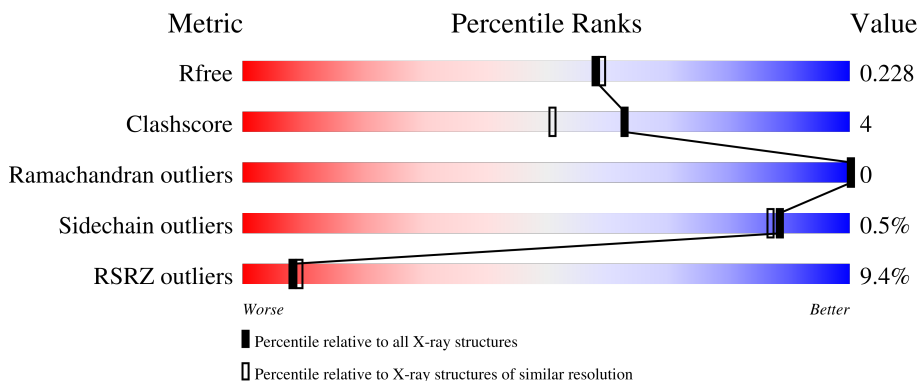
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

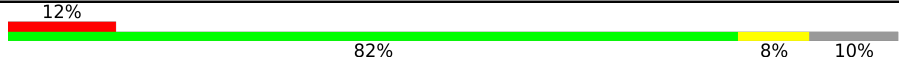
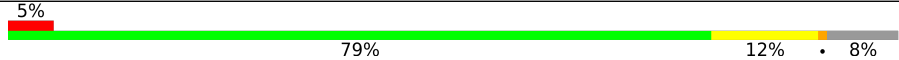
The reported resolution of this entry is 1.98 Å.

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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	AAA	117	
1	BBB	117	

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 1692 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 Cadherin-23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	105	797	496	135	163	3	0	0	0
1	BBB	108	820	509	139	169	3	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	969	LEU	-	expression tag	UNP Q99PF4
AAA	970	GLU	-	expression tag	UNP Q99PF4
AAA	971	HIS	-	expression tag	UNP Q99PF4
AAA	972	HIS	-	expression tag	UNP Q99PF4
AAA	973	HIS	-	expression tag	UNP Q99PF4
AAA	974	HIS	-	expression tag	UNP Q99PF4
AAA	975	HIS	-	expression tag	UNP Q99PF4
AAA	976	HIS	-	expression tag	UNP Q99PF4
BBB	969	LEU	-	expression tag	UNP Q99PF4
BBB	970	GLU	-	expression tag	UNP Q99PF4
BBB	971	HIS	-	expression tag	UNP Q99PF4
BBB	972	HIS	-	expression tag	UNP Q99PF4
BBB	973	HIS	-	expression tag	UNP Q99PF4
BBB	974	HIS	-	expression tag	UNP Q99PF4
BBB	975	HIS	-	expression tag	UNP Q99PF4
BBB	976	HIS	-	expression tag	UNP Q99PF4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	1	Total	Ca	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	BBB	1	Total	Ca	0	0
			1	1		

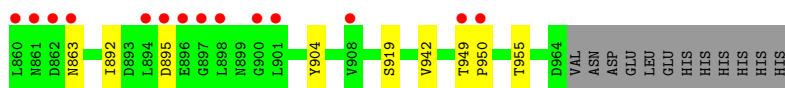
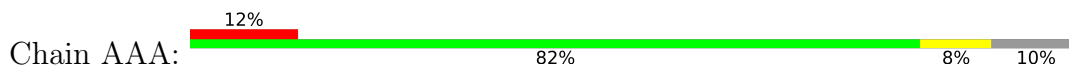
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	43	Total	O	0	0
			43	43		
3	BBB	30	Total	O	0	0
			30	30		

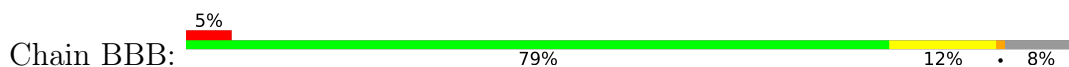
### 3 Residue-property plots [i](#)

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: Cadherin-23



- Molecule 1: Cadherin-23



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.31Å 41.66Å 56.91Å 90.00° 114.20° 90.00°	Depositor
Resolution (Å)	51.91 – 1.98 44.09 – 1.98	Depositor EDS
% Data completeness (in resolution range)	95.8 (51.91-1.98) 95.8 (44.09-1.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.03 (at 1.98Å)	Xtrriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.180 , 0.224 0.191 , 0.228	Depositor DCC
$R_{free}$ test set	681 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.0	Xtrriage
Anisotropy	0.325	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.034 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1692	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.29% 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: CA

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	AAA	0.69	0/807	0.85	0/1101
1	BBB	0.67	0/830	0.81	0/1133
All	All	0.68	0/1637	0.83	0/2234

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	AAA	797	0	794	7	0
1	BBB	820	0	813	9	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	43	0	0	0	0
3	BBB	30	0	0	0	0
All	All	1692	0	1607	14	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 4.

All (14) 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:BBB:937:GLU:OE1	1:BBB:961:ARG:NH2	2.37	0.56
1:AAA:949:THR:OG1	1:AAA:950:PRO:HD3	2.09	0.53
1:AAA:904:TYR:O	1:AAA:919:SER:HA	2.09	0.51
1:BBB:910:MET:SD	1:BBB:935:ILE:HD13	2.50	0.51
1:AAA:949:THR:OG1	1:AAA:950:PRO:CD	2.62	0.47
1:AAA:892:ILE:CD1	1:BBB:892:ILE:HG12	2.47	0.45
1:AAA:892:ILE:HD13	1:BBB:892:ILE:HG12	1.99	0.45
1:BBB:901:LEU:O	1:BBB:946:ASP:HA	2.17	0.44
1:BBB:949:THR:HB	1:BBB:950:PRO:CD	2.47	0.44
1:BBB:908:VAL:CG2	1:BBB:913:MET:HG3	2.49	0.43
1:BBB:861:ASN:ND2	1:BBB:948:GLY:HA3	2.34	0.43
1:AAA:863:ASN:ND2	1:AAA:895:ASP:OD1	2.52	0.42
1:AAA:942:VAL:O	1:AAA:955:THR:HA	2.19	0.42
1:BBB:874:ALA:O	1:BBB:960:VAL:HA	2.20	0.41

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	AAA	103/117 (88%)	101 (98%)	2 (2%)	0	100	100
1	BBB	106/117 (91%)	101 (95%)	5 (5%)	0	100	100
All	All	209/234 (89%)	202 (97%)	7 (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	AAA	91/103 (88%)	91 (100%)	0	100	100
1	BBB	94/103 (91%)	93 (99%)	1 (1%)	73	70
All	All	185/206 (90%)	184 (100%)	1 (0%)	88	87

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	908	VAL

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 2 ligands modelled in this entry, 2 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

### 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	AAA	105/117 (89%)	0.61	14 (13%) <b>3</b> <b>3</b>	12, 20, 59, 74	0
1	BBB	108/117 (92%)	0.06	6 (5%) <b>24</b> <b>26</b>	12, 23, 46, 67	0
All	All	213/234 (91%)	0.33	20 (9%) <b>8</b> <b>9</b>	12, 21, 56, 74	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	949	THR	8.7
1	AAA	898	LEU	6.9
1	AAA	860	LEU	6.2
1	AAA	896	GLU	4.8
1	BBB	949	THR	4.1
1	AAA	901	LEU	3.7
1	AAA	897	GLY	3.5
1	AAA	862	ASP	3.5
1	AAA	950	PRO	3.3
1	BBB	898	LEU	3.2
1	AAA	894	LEU	2.9
1	AAA	863	ASN	2.8
1	AAA	861	ASN	2.7
1	BBB	909	GLY	2.7
1	BBB	908	VAL	2.6
1	AAA	908	VAL	2.5
1	BBB	894	LEU	2.3
1	BBB	860	LEU	2.3
1	AAA	900	GLY	2.2
1	AAA	895	ASP	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](#)

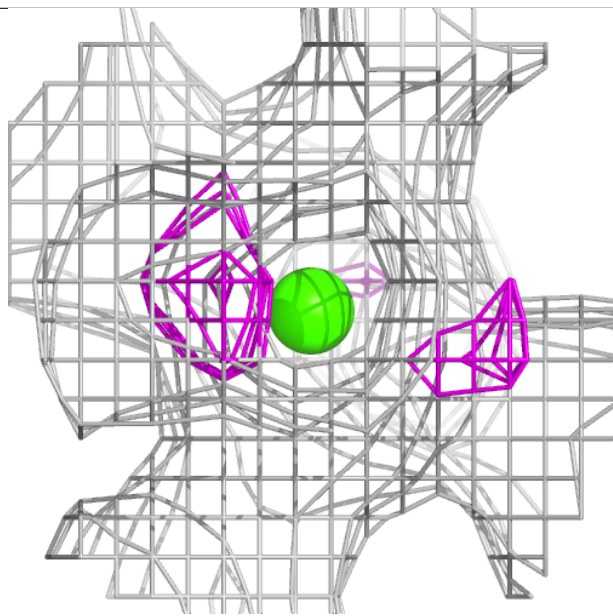
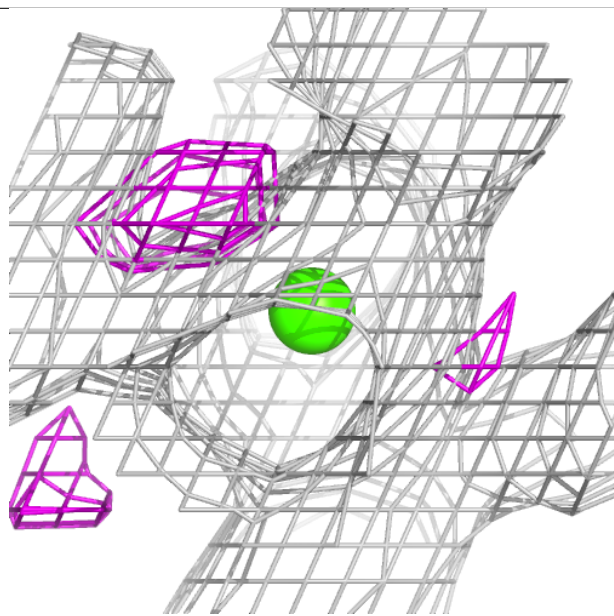
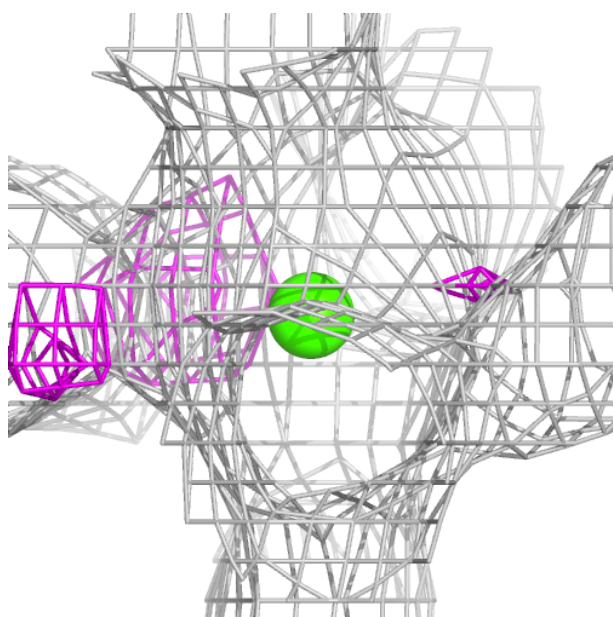
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

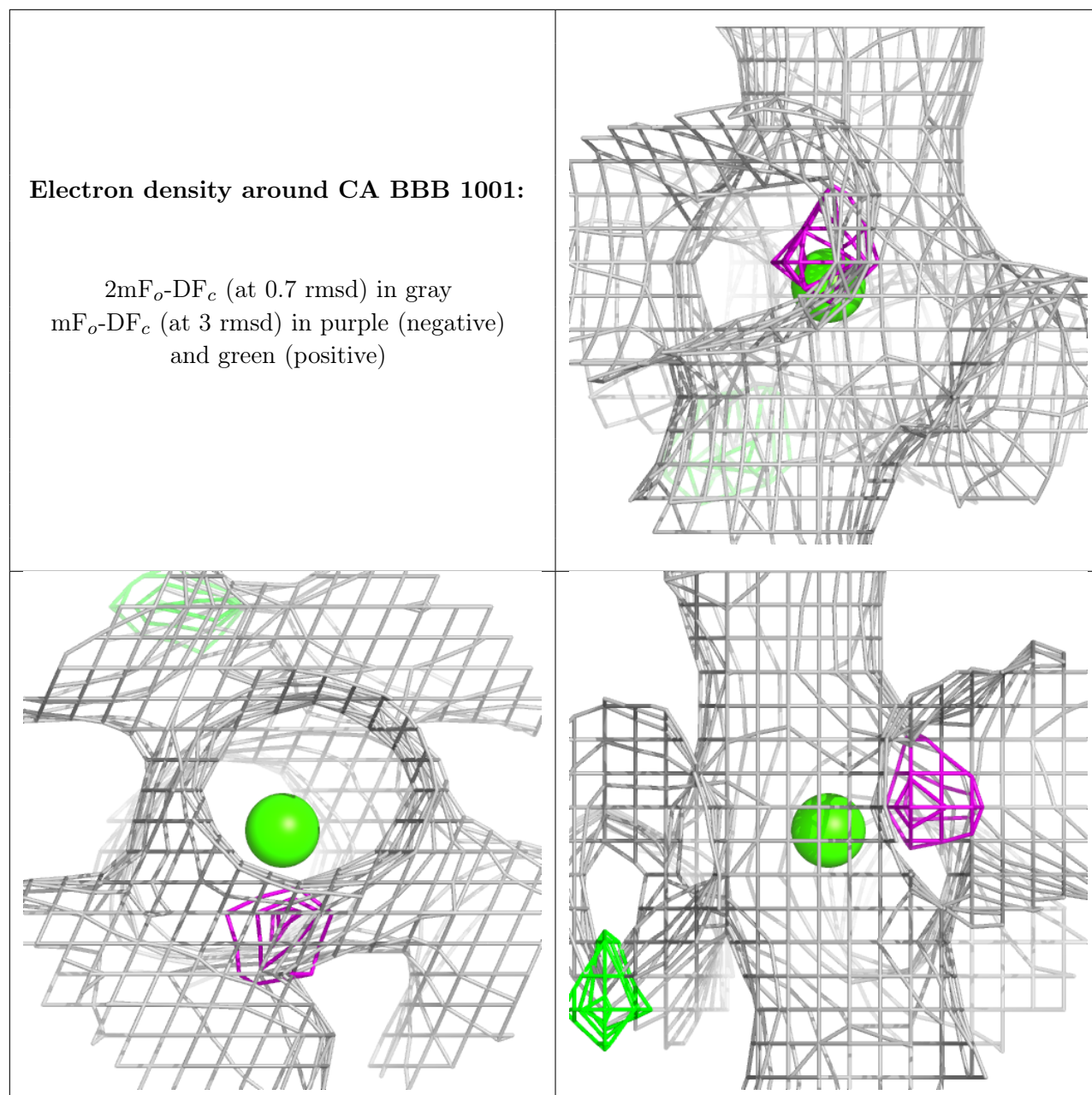
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	AAA	1001	1/1	0.94	0.12	45,45,45,45	0
2	CA	BBB	1001	1/1	0.97	0.07	33,33,33,33	0

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 CA AAA 1001:**

$2mF_o-DF_c$  (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.