



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 6, 2025 – 12:09 PM JST

PDB ID : 9L0N / pdb\_00009l0n  
Title : The crystal structure of the beta-glucosidase Ks5A7  
Authors : Xie, W.; Yang, N.  
Deposited on : 2024-12-12  
Resolution : 2.50 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 Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.46

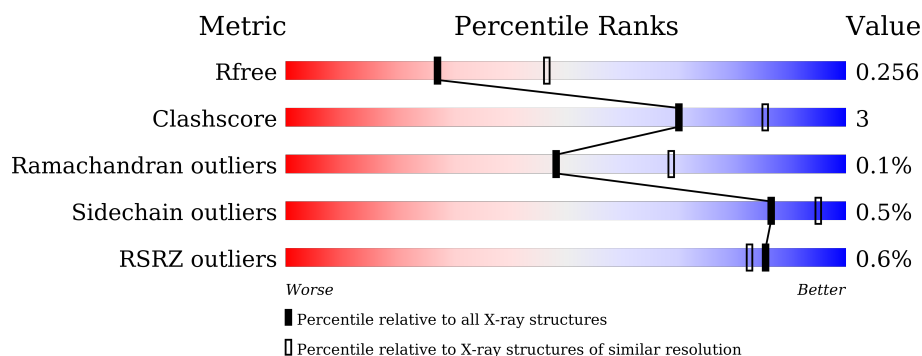
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 2.50 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	5504 (2.50-2.50)
Clashscore	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (2.50-2.50)

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	460	 90% 8% .
1	B	460	 87% 12% .
1	C	460	 92% 7% .
1	D	460	 91% 8% .

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15722 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 Beta-glucosidase Ks5A7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	455	Total	C	N	O	S	0	0	0
			3723	2411	614	684	14			
1	B	453	Total	C	N	O	S	0	0	0
			3705	2401	611	680	13			
1	C	454	Total	C	N	O	S	0	0	0
			3713	2405	611	683	14			
1	D	453	Total	C	N	O	S	0	0	0
			3705	2401	611	680	13			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	D	1	Total	Ca	0	0
			1	1		

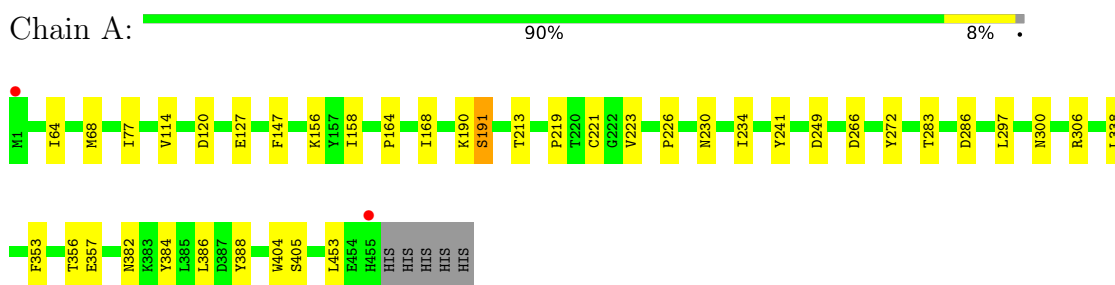
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	255	Total	O	0	0
			255	255		
3	B	219	Total	O	0	0
			219	219		
3	C	206	Total	O	0	0
			206	206		
3	D	194	Total	O	0	0
			194	194		

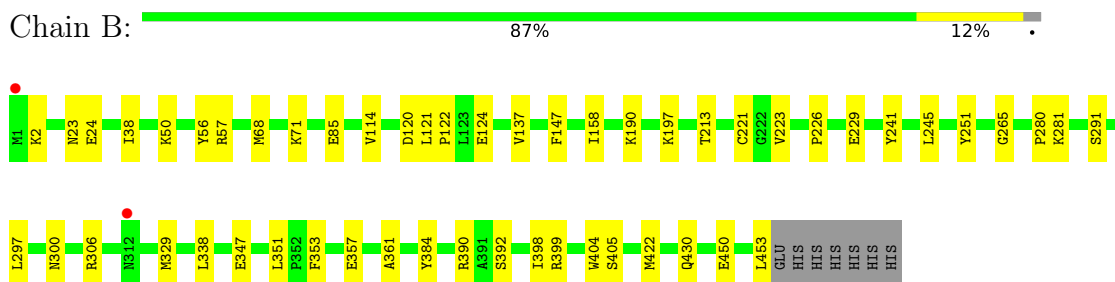
### 3 Residue-property plots

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.

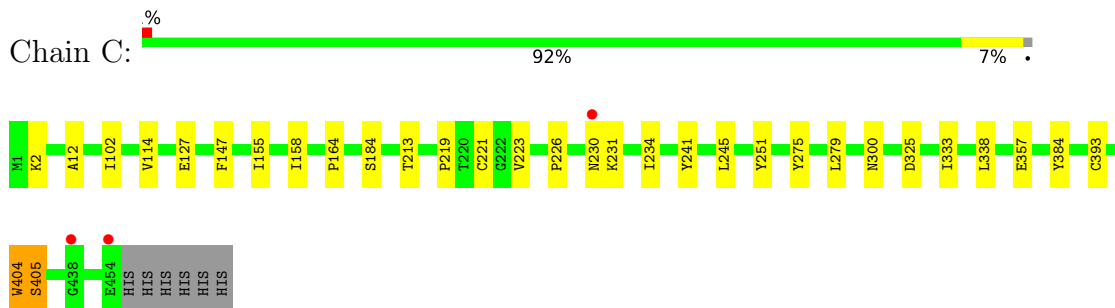
- Molecule 1: Beta-glucosidase Ks5A7



- Molecule 1: Beta-glucosidase Ks5A7

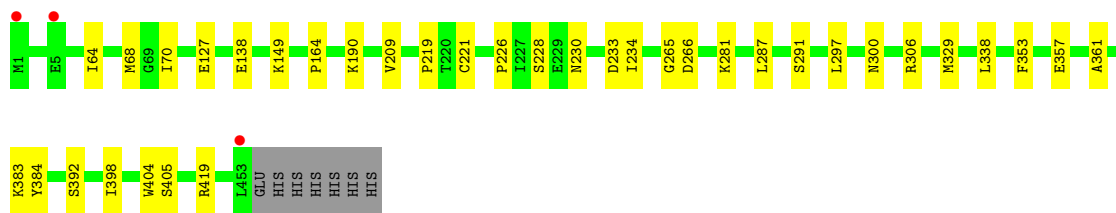


- Molecule 1: Beta-glucosidase Ks5A7



- Molecule 1: Beta-glucosidase Ks5A7





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.04Å 112.94Å 99.39Å 90.00° 101.23° 90.00°	Depositor
Resolution (Å)	27.34 – 2.50 27.34 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (27.34-2.50) 99.7 (27.34-2.50)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.27 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, $R_{free}$	0.216 , 0.256 0.215 , 0.256	Depositor DCC
$R_{free}$ test set	3338 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.3	Xtriage
Anisotropy	0.768	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 37.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	15722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.1639e-03.*

<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	A	0.13	0/3834	0.31	0/5203
1	B	0.12	0/3815	0.32	0/5177
1	C	0.12	0/3823	0.32	0/5188
1	D	0.12	0/3815	0.31	0/5177
All	All	0.12	0/15287	0.32	0/20745

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	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	404	TRP	Peptide

### 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	3723	0	3561	24	0
1	B	3705	0	3552	33	0
1	C	3713	0	3554	20	0
1	D	3705	0	3552	23	0
2	B	1	0	0	0	0
2	D	1	0	0	0	0
3	A	255	0	0	1	0
3	B	219	0	0	6	0
3	C	206	0	0	2	0
3	D	194	0	0	2	0
All	All	15722	0	14219	96	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 3.

The worst 5 of 96 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:C:393:CYS:SG	3:C:561:HOH:O	2.39	0.79
1:B:229:GLU:OE2	1:B:390:ARG:NH2	2.20	0.71
1:B:50:LYS:HE2	1:B:430:GLN:HE21	1.57	0.70
1:D:138:GLU:OE2	3:D:601:HOH:O	2.10	0.68
1:B:121:LEU:HD22	1:B:122:PRO:HD2	1.82	0.62

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	453/460 (98%)	442 (98%)	11 (2%)	0	100	100
1	B	451/460 (98%)	440 (98%)	11 (2%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	452/460 (98%)	439 (97%)	12 (3%)	1 (0%)	44	64
1	D	451/460 (98%)	439 (97%)	12 (3%)	0	100	100
All	All	1807/1840 (98%)	1760 (97%)	46 (2%)	1 (0%)	48	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	405	SER

### 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	391/402 (97%)	388 (99%)	3 (1%)	79	91
1	B	389/402 (97%)	387 (100%)	2 (0%)	86	95
1	C	390/402 (97%)	389 (100%)	1 (0%)	91	97
1	D	389/402 (97%)	387 (100%)	2 (0%)	86	95
All	All	1559/1608 (97%)	1551 (100%)	8 (0%)	86	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	281	LYS
1	D	221	CYS
1	B	453	LEU
1	B	221	CYS
1	C	221	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	92	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	212	ASN
1	D	292	GLN
1	D	23	ASN
1	B	212	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

### 5.6 Ligand geometry ⓘ

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 ⓘ

There are no such residues in this entry.

### 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data [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	455/460 (98%)	-0.19	2 (0%) 89 86	5, 13, 24, 51	0
1	B	453/460 (98%)	-0.12	2 (0%) 89 86	6, 13, 25, 45	0
1	C	454/460 (98%)	-0.08	3 (0%) 84 81	6, 15, 29, 48	0
1	D	453/460 (98%)	0.07	3 (0%) 84 81	7, 19, 30, 45	0
All	All	1815/1840 (98%)	-0.08	10 (0%) 85 83	5, 15, 28, 51	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	230	ASN	3.6
1	A	455	HIS	3.4
1	C	454	GLU	2.5
1	D	1	MET	2.4
1	A	1	MET	2.3

### 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](#)

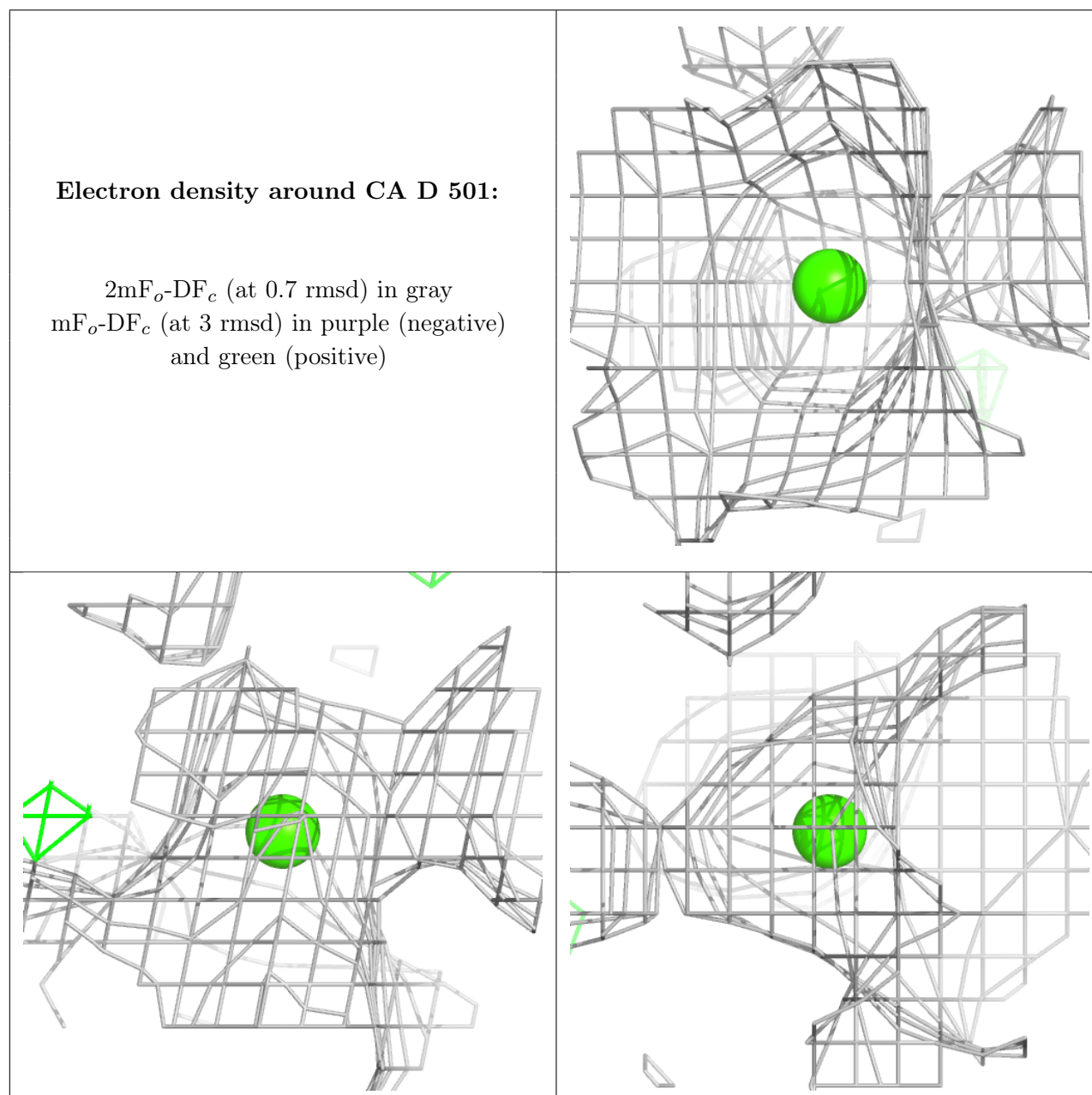
There are no oligosaccharides in this entry.

### 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.

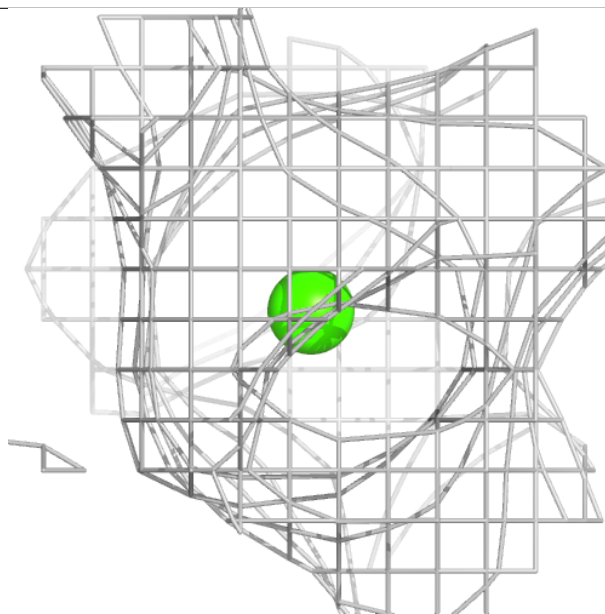
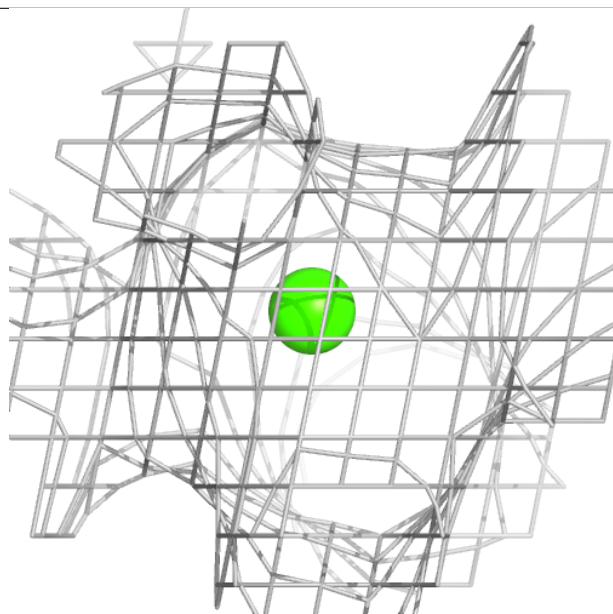
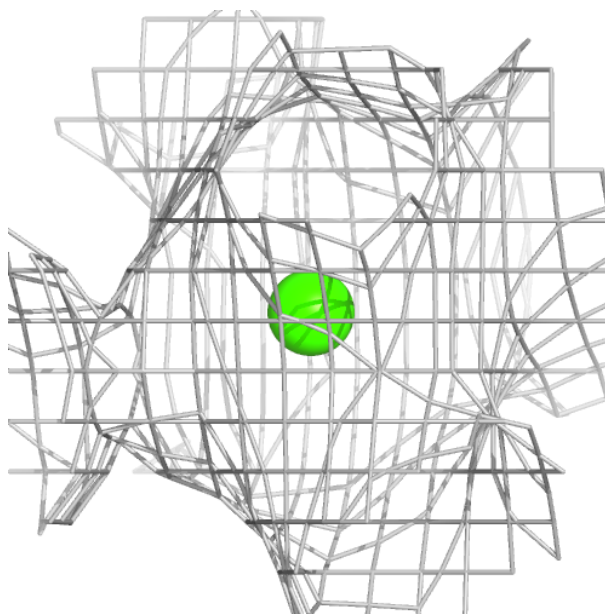
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	D	501	1/1	0.87	0.09	48,48,48,48	0
2	CA	B	501	1/1	0.95	0.04	17,17,17,17	0

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 B 501:**

$2mF_o-DF_c$  (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.