



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 2, 2020 – 06:13 PM JST

PDB ID : 7C7D  
Title : Crystal structure of the catalytic unit of thermostable GH87 alpha-1,3-glucanase from *Streptomyces thermodiastaticus* strain HF3-3  
Authors : Itoh, T.; Panti, N.; Toyotake, Y.; Hayashi, J.; Suyotha, W.; Yano, S.; Wakayama, M.; Hibi, T.  
Deposited on : 2020-05-25  
Resolution : 1.16 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
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)	:	2.14.6

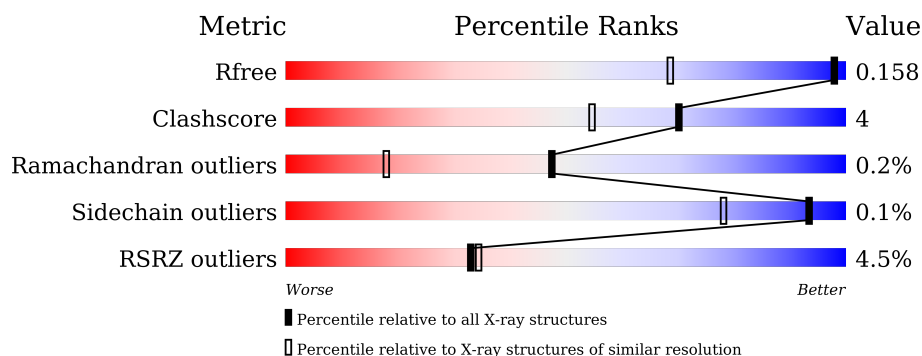
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.16 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	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)
RSRZ outliers	127900	1724 (1.20-1.12)

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 on 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	610	<div> <div>5%</div> <div>82%</div> <div>6%</div> <div>12%</div> </div>
1	B	610	<div> <div>3%</div> <div>84%</div> <div>5%</div> <div>12%</div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9253 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.

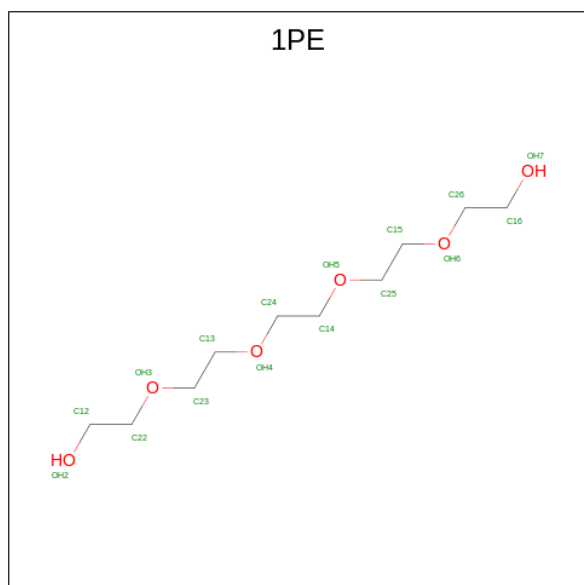
- Molecule 1 is a protein called alpha-1,3-glucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total	C	N	O	S	0	16	0
			4068	2528	703	829	8			
1	B	539	Total	C	N	O	S	0	20	0
			4078	2539	699	831	9			

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

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

- Molecule 3 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			16	10	6		

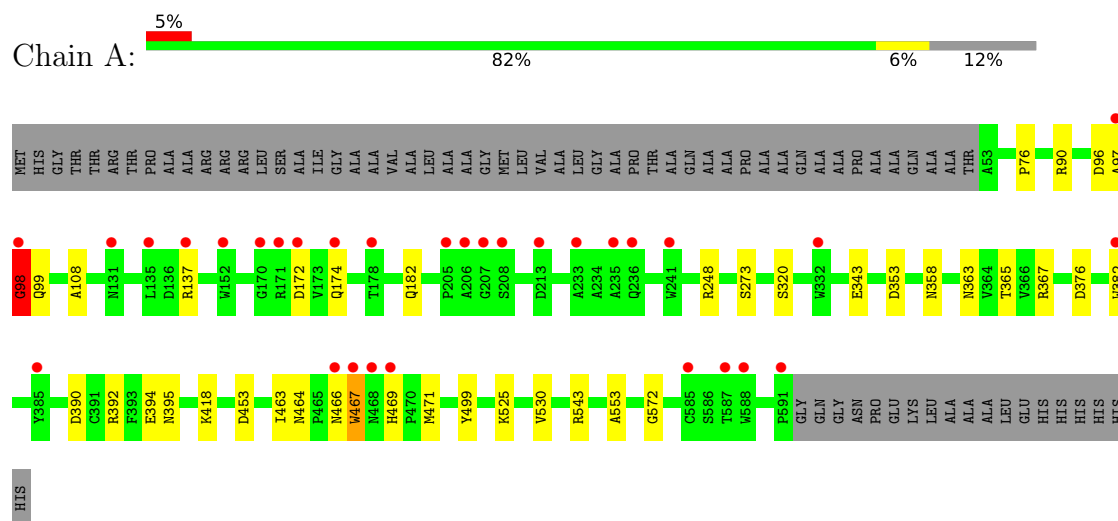
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	499	Total	O	0	0
			499	499		
4	B	590	Total	O	0	0
			590	590		

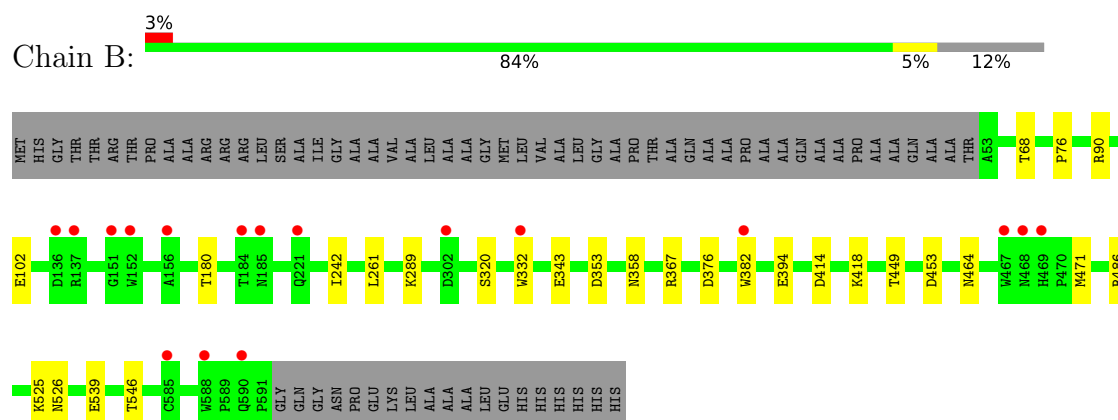
### 3 Residue-property plots [i](#)

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-1,3-glucanase



- Molecule 1: alpha-1,3-glucanase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.51Å 91.33Å 164.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.21 – 1.16 41.21 – 1.16	Depositor EDS
% Data completeness (in resolution range)	99.7 (41.21-1.16) 99.7 (41.21-1.16)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.16Å)	Xtriage
Refinement program	PHENIX 1.17.1 _3660	Depositor
R, $R_{free}$	0.135 , 0.159 0.135 , 0.158	Depositor DCC
$R_{free}$ test set	16329 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.0	Xtriage
Anisotropy	0.169	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 61.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	9253	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.22% 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, 1PE

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	0.37	1/4205 (0.0%)	0.60	0/5750
1	B	0.30	0/4227	0.59	0/5784
All	All	0.34	1/8432 (0.0%)	0.60	0/11534

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	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	467	TRP	CB-CG	-8.38	1.35	1.50

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	98	GLY	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	4068	0	3930	34	1
1	B	4078	0	3952	23	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	B	16	0	22	1	0
4	A	499	0	0	2	0
4	B	590	0	0	5	0
All	All	9253	0	7904	57	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 4.

All (57) 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:A:467:TRP:HA	1:A:467:TRP:CE3	2.17	0.79
1:A:137:ARG:NH1	1:A:137:ARG:HB2	2.01	0.76
1:A:467:TRP:HE3	1:A:467:TRP:HA	1.50	0.73
1:B:486:ARG:HG3	3:B:801:1PE:H221	1.69	0.73
1:A:96:ASP:O	1:A:98:GLY:N	2.27	0.68
1:A:248[B]:ARG:HH21	1:A:273:SER:HA	1.64	0.63
1:A:467:TRP:CA	1:A:467:TRP:CE3	2.83	0.61
1:A:471[C]:MET:HE3	4:A:1012:HOH:O	2.02	0.59
1:A:137:ARG:HB2	1:A:137:ARG:HH11	1.67	0.58
1:B:353:ASP:OD1	1:B:376:ASP:HB3	2.03	0.58
1:B:546[B]:THR:HG22	4:B:955:HOH:O	2.04	0.57
1:A:467:TRP:HB2	1:A:469:HIS:H	1.71	0.55
1:A:108:ALA:HB2	1:A:174:GLN:NE2	2.23	0.53
1:A:464:ASN:HD22	1:A:471[C]:MET:CE	2.22	0.52
1:A:365:THR:OG1	1:A:392[B]:ARG:HD3	2.09	0.52
1:B:539:GLU:OE2	4:B:901:HOH:O	2.19	0.51
1:A:353:ASP:OD1	1:A:376:ASP:HB3	2.12	0.49
1:B:464:ASN:HD22	1:B:471[C]:MET:CE	2.25	0.49
1:A:98:GLY:HA3	1:A:99:GLN:HG3	1.94	0.49
1:A:98:GLY:CA	1:A:99:GLN:HG3	2.43	0.48
1:B:343:GLU:HA	1:B:367:ARG:O	2.14	0.47
1:B:367:ARG:HA	1:B:394:GLU:O	2.15	0.47
1:A:343:GLU:HA	1:A:367:ARG:O	2.15	0.47
1:A:543[B]:ARG:HG2	1:A:572:GLY:HA3	1.98	0.46
1:A:464:ASN:HD22	1:A:471[C]:MET:HE1	1.80	0.46

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:394[C]:GLU:HA	1:A:418:LYS:O	2.16	0.45
1:A:395:ASN:ND2	4:A:811:HOH:O	2.48	0.45
1:A:137:ARG:HB2	1:A:137:ARG:CZ	2.45	0.45
1:B:418:LYS:HA	1:B:453:ASP:O	2.15	0.45
1:A:394[A]:GLU:HA	1:A:418:LYS:O	2.17	0.45
1:B:394:GLU:HA	1:B:418:LYS:O	2.17	0.45
1:A:394[B]:GLU:HA	1:A:418:LYS:O	2.17	0.45
1:B:102:GLU:HG2	1:B:180:THR:OG1	2.16	0.44
1:B:320:SER:HA	1:B:343:GLU:O	2.17	0.44
1:B:526[A]:ASN:ND2	4:B:919:HOH:O	2.50	0.44
1:A:463:ILE:HG12	1:A:499:TYR:CZ	2.53	0.44
1:A:525[A]:LYS:HB3	1:A:525[A]:LYS:HE2	1.82	0.44
1:B:358[B]:ASN:OD1	1:B:382:TRP:HB3	2.18	0.44
1:B:471[C]:MET:HE3	4:B:908:HOH:O	2.18	0.44
1:B:289:LYS:HA	1:B:320:SER:O	2.18	0.43
1:B:242:ILE:HD11	1:B:261[A]:LEU:HG	1.99	0.43
1:A:358[B]:ASN:OD1	1:A:382:TRP:HB3	2.19	0.43
1:A:418:LYS:HA	1:A:453:ASP:O	2.19	0.43
1:B:525[B]:LYS:NZ	1:B:526[B]:ASN:OD1	2.47	0.43
1:A:76:PRO:HA	1:A:90:ARG:O	2.19	0.43
1:B:68[B]:THR:OG1	1:B:102:GLU:HB2	2.18	0.43
1:B:76:PRO:HA	1:B:90:ARG:O	2.19	0.43
1:A:320:SER:HA	1:A:343:GLU:O	2.19	0.42
1:B:526[B]:ASN:ND2	4:B:921:HOH:O	2.52	0.42
1:B:414:ASP:HA	1:B:449:THR:O	2.20	0.42
1:B:464:ASN:HD22	1:B:471[C]:MET:HE2	1.84	0.42
1:A:530:VAL:HA	1:A:553:ALA:O	2.21	0.41
1:A:464:ASN:HB2	1:A:471[C]:MET:HE2	2.03	0.41
1:B:332:TRP:HE1	1:B:358[B]:ASN:ND2	2.19	0.41
1:A:363:ASN:HA	1:A:390:ASP:O	2.21	0.40
1:A:394[A]:GLU:HG2	1:A:418:LYS:HB3	2.02	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:A:182:GLN:NE2	1:A:466:ASN:O[3_655]	2.19	0.01

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	555/610 (91%)	531 (96%)	22 (4%)	2 (0%)	34	10
1	B	559/610 (92%)	535 (96%)	24 (4%)	0	100	100
All	All	1114/1220 (91%)	1066 (96%)	46 (4%)	2 (0%)	47	18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	ALA
1	A	98	GLY

### 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	441/464 (95%)	440 (100%)	1 (0%)	93	78
1	B	445/464 (96%)	445 (100%)	0	100	100
All	All	886/928 (96%)	885 (100%)	1 (0%)	93	78

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

Mol	Chain	Res	Type
1	A	172	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	131	ASN
1	A	197	GLN
1	A	469	HIS

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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$
3	1PE	B	801	-	15,15,15	0.53	0	14,14,14	0.24	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
3	1PE	B	801	-	-	4/13/13/13	-

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
3	B	801	1PE	OH2-C12-C22-OH3
3	B	801	1PE	C14-C24-OH4-C13
3	B	801	1PE	OH4-C13-C23-OH3
3	B	801	1PE	OH5-C14-C24-OH4

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	801	1PE	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	539/610 (88%)	0.35	31 (5%)	23 24	10, 16, 37, 68	0
1	B	539/610 (88%)	0.14	17 (3%)	47 48	10, 15, 30, 73	0
All	All	1078/1220 (88%)	0.25	48 (4%)	33 34	10, 16, 34, 73	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	467	TRP	19.4
1	B	467	TRP	13.7
1	A	98	GLY	10.7
1	A	172	ASP	6.3
1	A	171	ARG	5.5
1	B	469	HIS	5.2
1	B	588	TRP	4.6
1	A	97	ALA	4.6
1	A	469	HIS	4.6
1	A	588	TRP	4.4
1	B	590	GLN	3.9
1	B	185	ASN	3.6
1	A	174	GLN	3.6
1	A	170	GLY	3.5
1	B	136	ASP	3.4
1	A	468	ASN	3.4
1	B	152	TRP	3.2
1	A	206	ALA	3.1
1	B	221	GLN	3.0
1	B	184	THR	2.9
1	B	468	ASN	2.9
1	A	205	PRO	2.9
1	A	591	PRO	2.8
1	A	585	CYS	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	585	CYS	2.8
1	A	137	ARG	2.7
1	B	382	TRP	2.6
1	B	332	TRP	2.5
1	A	152	TRP	2.5
1	A	213	ASP	2.5
1	A	207	GLY	2.4
1	B	151	GLY	2.4
1	A	332	TRP	2.4
1	B	302	ASP	2.4
1	A	385	TYR	2.4
1	A	235	ALA	2.4
1	A	241	TRP	2.4
1	A	178	THR	2.3
1	A	233	ALA	2.2
1	A	587	THR	2.2
1	A	131	ASN	2.2
1	B	156	ALA	2.2
1	A	236	GLN	2.1
1	A	135	LEU	2.1
1	A	466	ASN	2.1
1	A	208	SER	2.1
1	B	137	ARG	2.0
1	A	382	TRP	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](#)

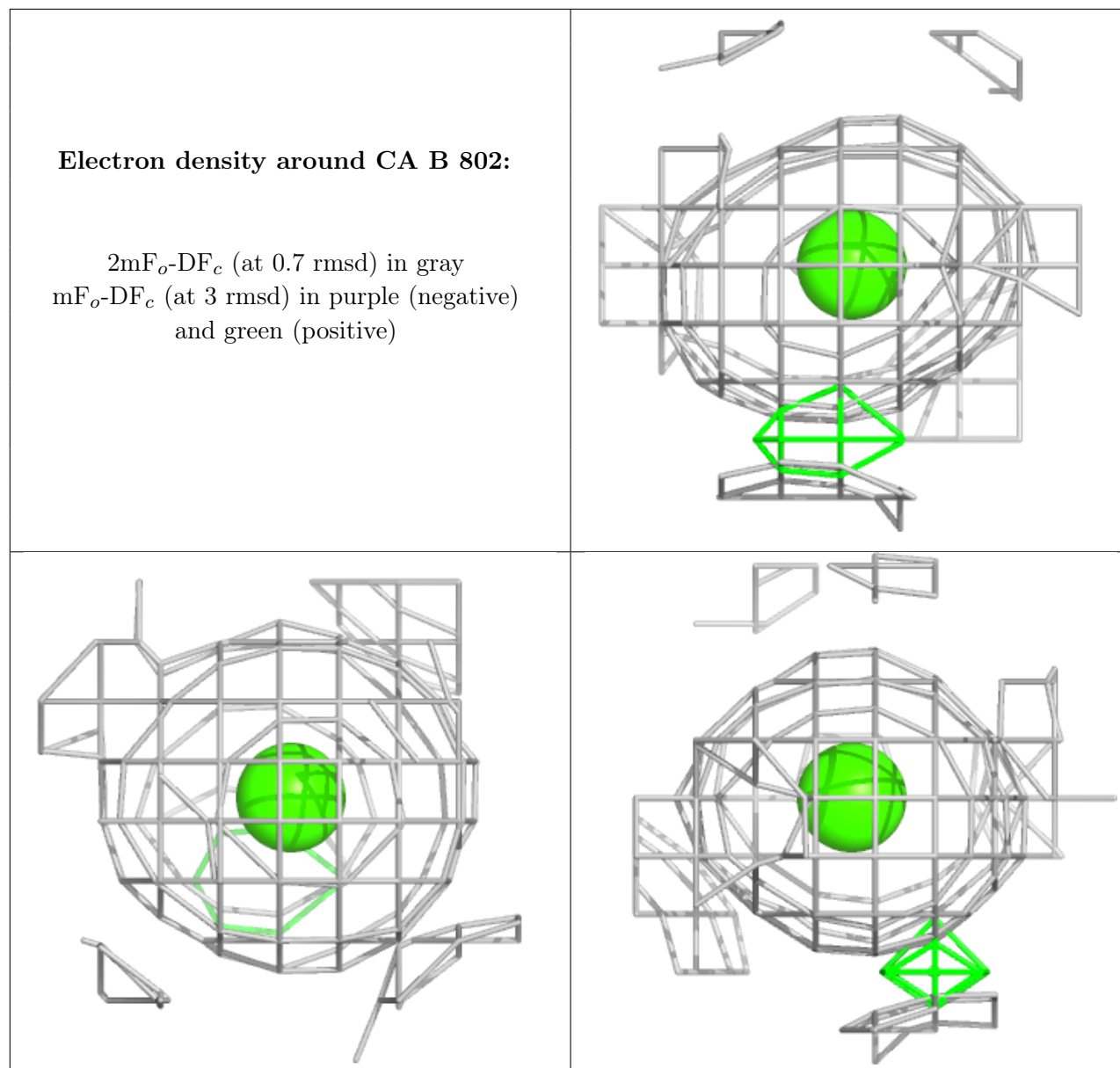
There are no monosaccharides in this entry.

## 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.

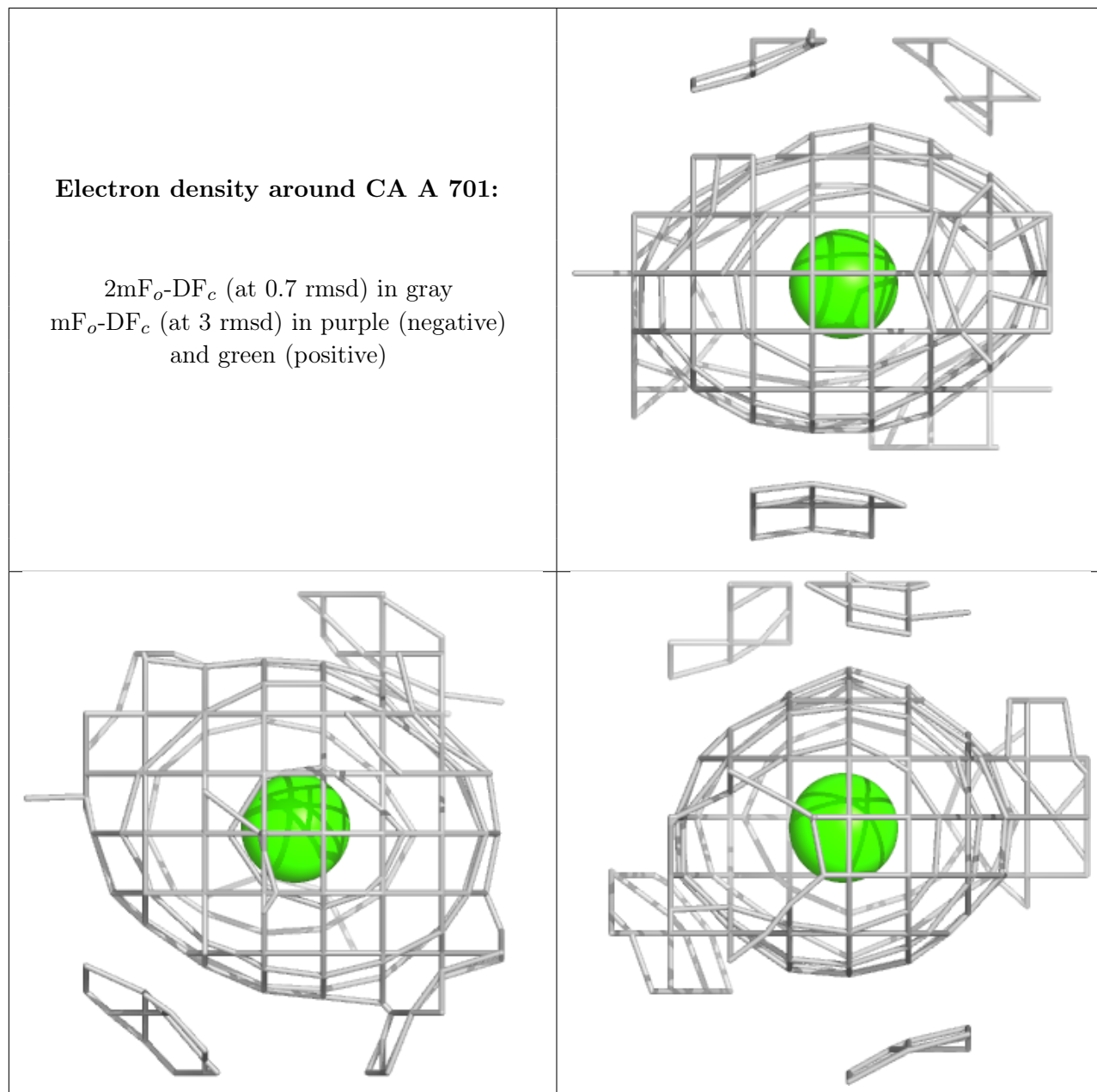
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	1PE	B	801	16/16	0.79	0.14	45,48,49,49	0
2	CA	B	802	1/1	1.00	0.05	12,12,12,12	0
2	CA	A	701	1/1	1.00	0.06	11,11,11,11	0

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 A 701:**

$2mF_o-DF_c$  (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.