



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 24, 2021 – 05:08 pm GMT

PDB ID : 7OZF  
Title : FGFR1 kinase domain (residues 458-765) with mutations C488A, C584S in complex with 19.  
Authors : Trinh, C.H.; Turner, L.D.; Fishwick, C.W.G.  
Deposited on : 2021-06-28  
Resolution : 1.82 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 : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

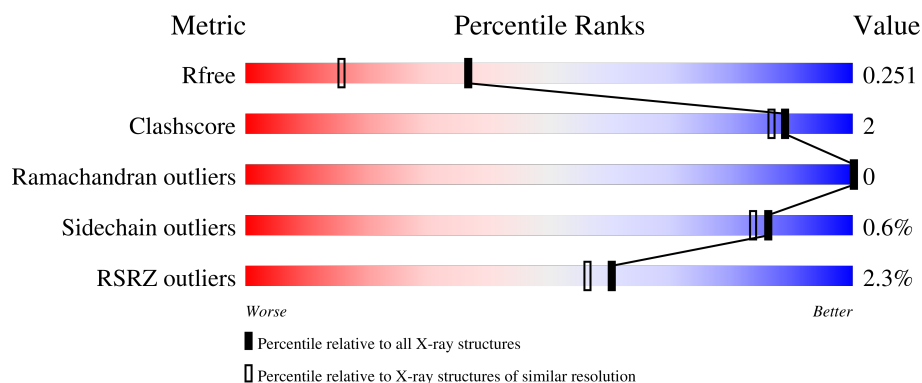
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.82 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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	309	<div> <div>3%</div> <div>89%</div> <div>7%</div> </div>
1	BBB	309	<div> <div>%</div> <div>85%</div> <div>6%</div> <div>9%</div> </div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 9295 atoms, of which 4548 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.

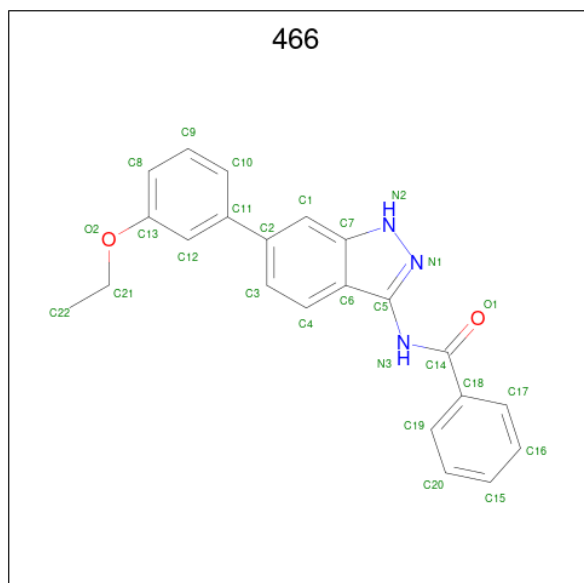
- Molecule 1 is a protein called Fibroblast growth factor receptor 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	288	Total	C	H	N	O	S	109	1	0
			4447	1416	2219	383	411	18			
1	BBB	282	Total	C	H	N	O	S	97	9	0
			4526	1442	2267	385	414	18			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	457	GLY	-	expression tag	UNP P11362
AAA	488	ALA	CYS	conflict	UNP P11362
AAA	584	SER	CYS	conflict	UNP P11362
BBB	457	GLY	-	expression tag	UNP P11362
BBB	488	ALA	CYS	conflict	UNP P11362
BBB	584	SER	CYS	conflict	UNP P11362

- Molecule 2 is N-[6-(3-ethoxyphenyl)-1H-indazol-3-yl]benzamide (three-letter code: 466) (formula: C<sub>22</sub>H<sub>19</sub>N<sub>3</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	0	0
			46	22	19	3	2		
2	BBB	1	Total	C	H	N	O	0	0
			46	22	19	3	2		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0
4	BBB	1	Total C H O 10 2 6 2	1	0
4	BBB	1	Total C H O 10 2 6 2	1	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	1	Total Cl 1 1	0	0
5	BBB	1	Total Cl 1 1	0	0

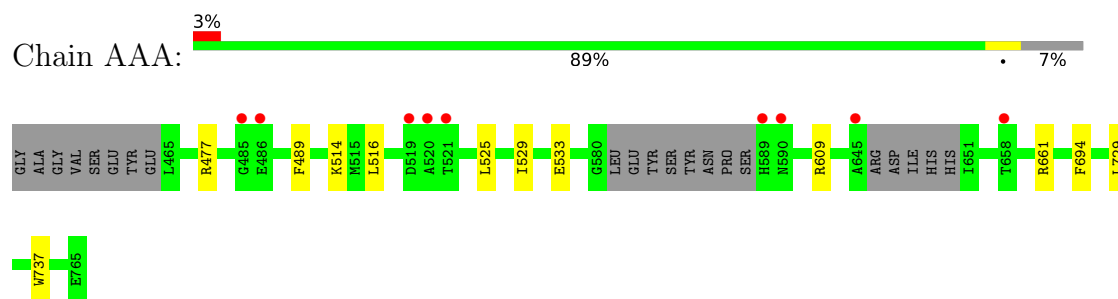
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	100	Total O 100 100	0	0
6	BBB	83	Total O 83 83	0	0

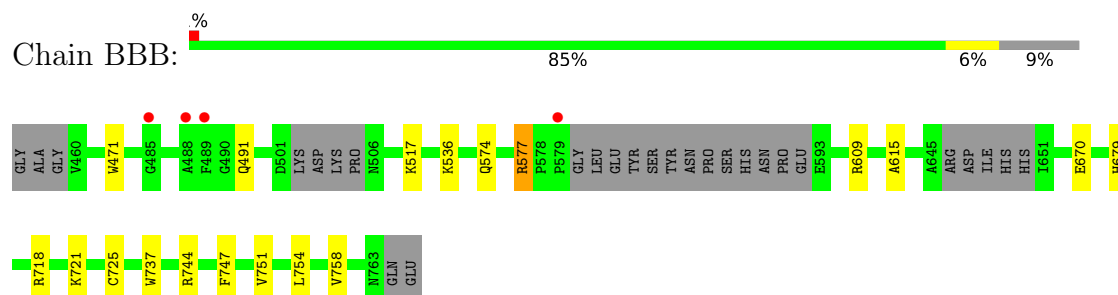
### 3 Residue-property plots [i](#)

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: Fibroblast growth factor receptor 1



- Molecule 1: Fibroblast growth factor receptor 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	208.32Å 57.77Å 65.94Å 90.00° 107.17° 90.00°	Depositor
Resolution (Å)	63.00 – 1.82 63.00 – 1.82	Depositor EDS
% Data completeness (in resolution range)	98.7 (63.00-1.82) 98.7 (63.00-1.82)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 1.82Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.204 , 0.242 0.211 , 0.251	Depositor DCC
$R_{free}$ test set	3411 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.8	Xtriage
Anisotropy	0.485	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.019 for -h-2*1,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9295	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.09% 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, EDO, 466, SO4

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	0.66	0/2275	0.74	0/3084
1	BBB	0.66	0/2304	0.74	0/3116
All	All	0.66	0/4579	0.74	0/6200

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

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	2228	2219	2172	7	0
1	BBB	2259	2267	2230	9	0
2	AAA	27	19	0	0	0
2	BBB	27	19	0	1	0
3	AAA	5	0	0	0	0
4	AAA	8	12	12	1	0
4	BBB	8	12	12	0	0
5	AAA	1	0	0	1	0
5	BBB	1	0	0	0	0
6	AAA	100	0	0	0	0
6	BBB	83	0	0	0	0
All	All	4747	4548	4426	17	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 2.

All (17) 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:AAA:516:LEU:HD11	1:AAA:525:LEU:HD13	1.77	0.67
1:BBB:721:LYS:NZ	1:BBB:725:CYS:O	2.29	0.64
1:AAA:489:PHE:O	1:AAA:514:LYS:HE3	2.05	0.57
1:AAA:661:ARG:HD2	5:AAA:805:CL:CL	2.52	0.47
1:BBB:471:TRP:CD1	1:BBB:536:LYS:HE2	2.52	0.45
1:BBB:574:GLN:O	1:BBB:577:ARG:HD2	2.16	0.44
1:AAA:477:ARG:HA	1:AAA:477:ARG:NE	2.32	0.44
1:BBB:615:ALA:HA	1:BBB:679:HIS:CE1	2.53	0.44
1:BBB:754:LEU:O	1:BBB:758:VAL:HG23	2.18	0.44
1:BBB:670:GLU:OE2	1:BBB:744:ARG:NH1	2.44	0.42
1:AAA:737:TRP:CD2	4:AAA:804:EDO:H11	2.55	0.42
1:BBB:491[B]:GLN:HB2	1:BBB:517:LYS:HG2	2.03	0.41
1:BBB:718:ARG:NH2	1:BBB:737:TRP:O	2.52	0.41
2:BBB:801:466:C4	2:BBB:801:466:O1	2.68	0.41
1:AAA:529:ILE:O	1:AAA:533:GLU:HG2	2.20	0.41
1:AAA:694:PHE:CZ	1:AAA:729:LEU:HD13	2.56	0.40
1:BBB:747:PHE:O	1:BBB:751:VAL:HG23	2.21	0.40

There are no symmetry-related clashes.

## 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	AAA	283/309 (92%)	277 (98%)	6 (2%)	0	100	100
1	BBB	283/309 (92%)	278 (98%)	5 (2%)	0	100	100
All	All	566/618 (92%)	555 (98%)	11 (2%)	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	233/269 (87%)	232 (100%)	1 (0%)	91	89
1	BBB	237/269 (88%)	235 (99%)	2 (1%)	81	77
All	All	470/538 (87%)	467 (99%)	3 (1%)	86	83

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	609	ARG
1	BBB	577	ARG
1	BBB	609	ARG

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 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

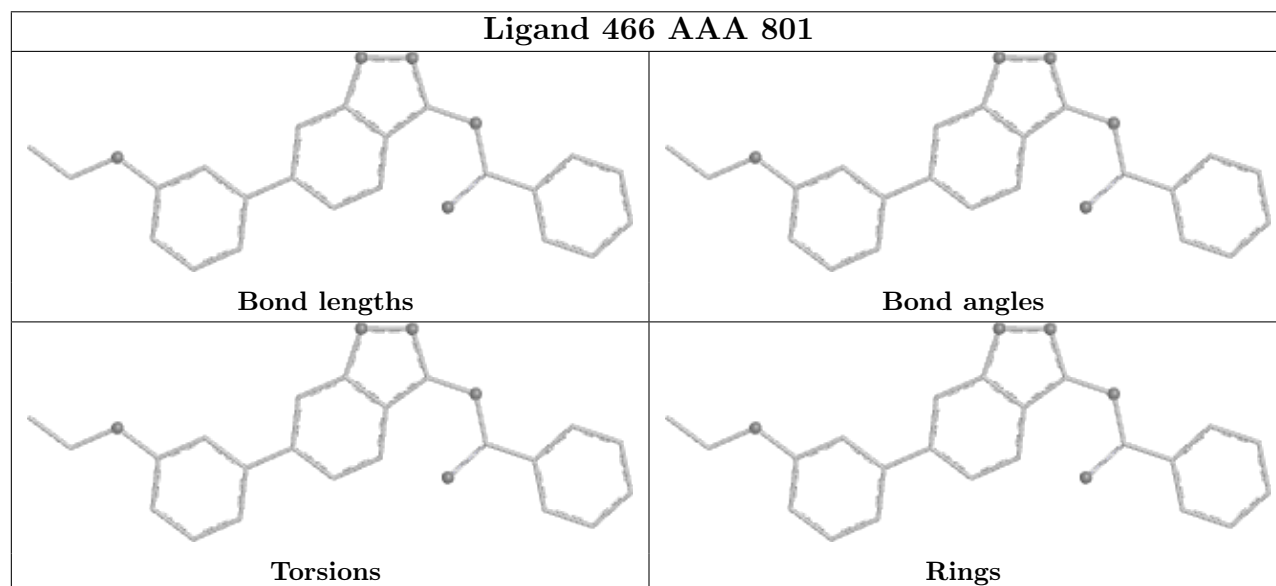
There are no chirality outliers.

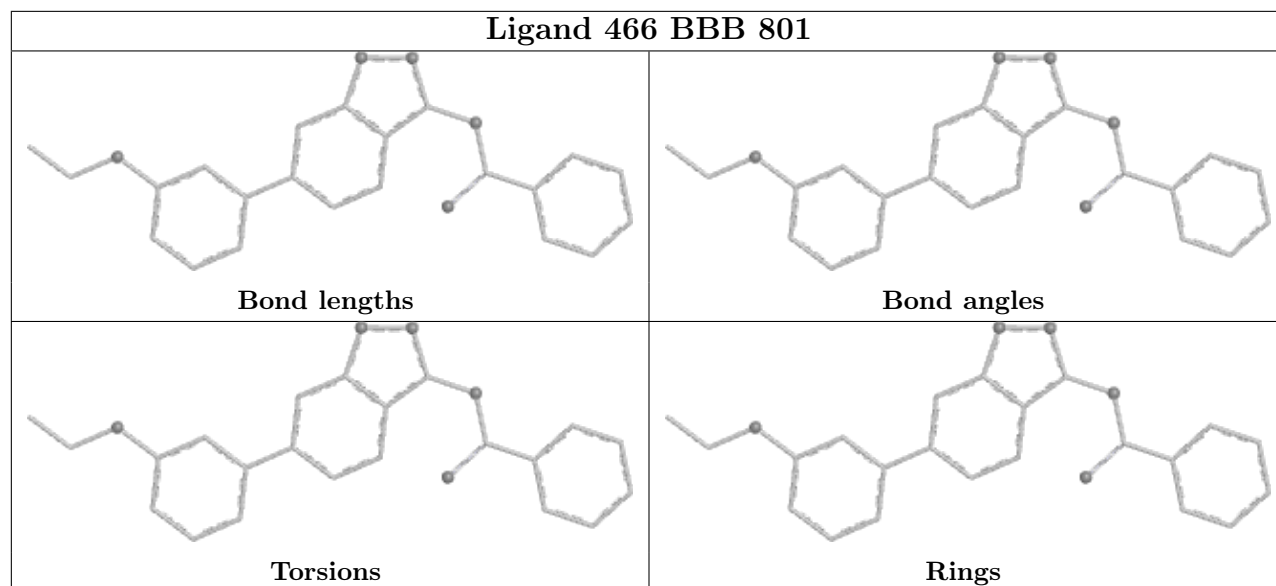
There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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	AAA	288/309 (93%)	0.07	9 (3%)	49	43	30, 47, 81, 111	0
1	BBB	282/309 (91%)	0.08	4 (1%)	75	72	35, 47, 74, 90	0
All	All	570/618 (92%)	0.08	13 (2%)	60	56	30, 47, 78, 111	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	489[A]	PHE	3.8
1	BBB	488[A]	ALA	3.5
1	AAA	520	ALA	3.3
1	AAA	521	THR	3.1
1	AAA	486	GLU	3.0
1	BBB	485[A]	GLY	2.9
1	AAA	485	GLY	2.7
1	BBB	579	PRO	2.4
1	AAA	519	ASP	2.4
1	AAA	658	THR	2.3
1	AAA	590	ASN	2.2
1	AAA	589	HIS	2.1
1	AAA	645	ALA	2.1

### 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](#)

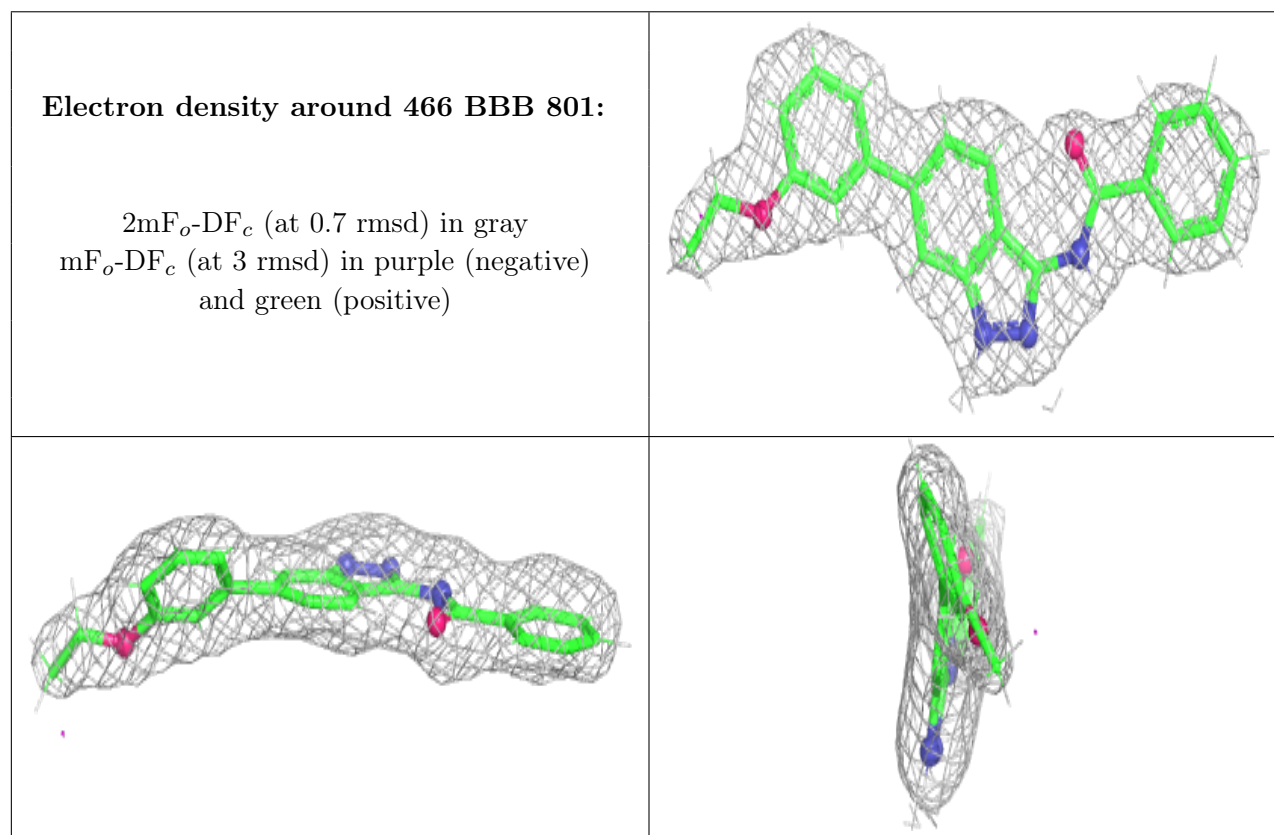
There are no monosaccharides in this entry.

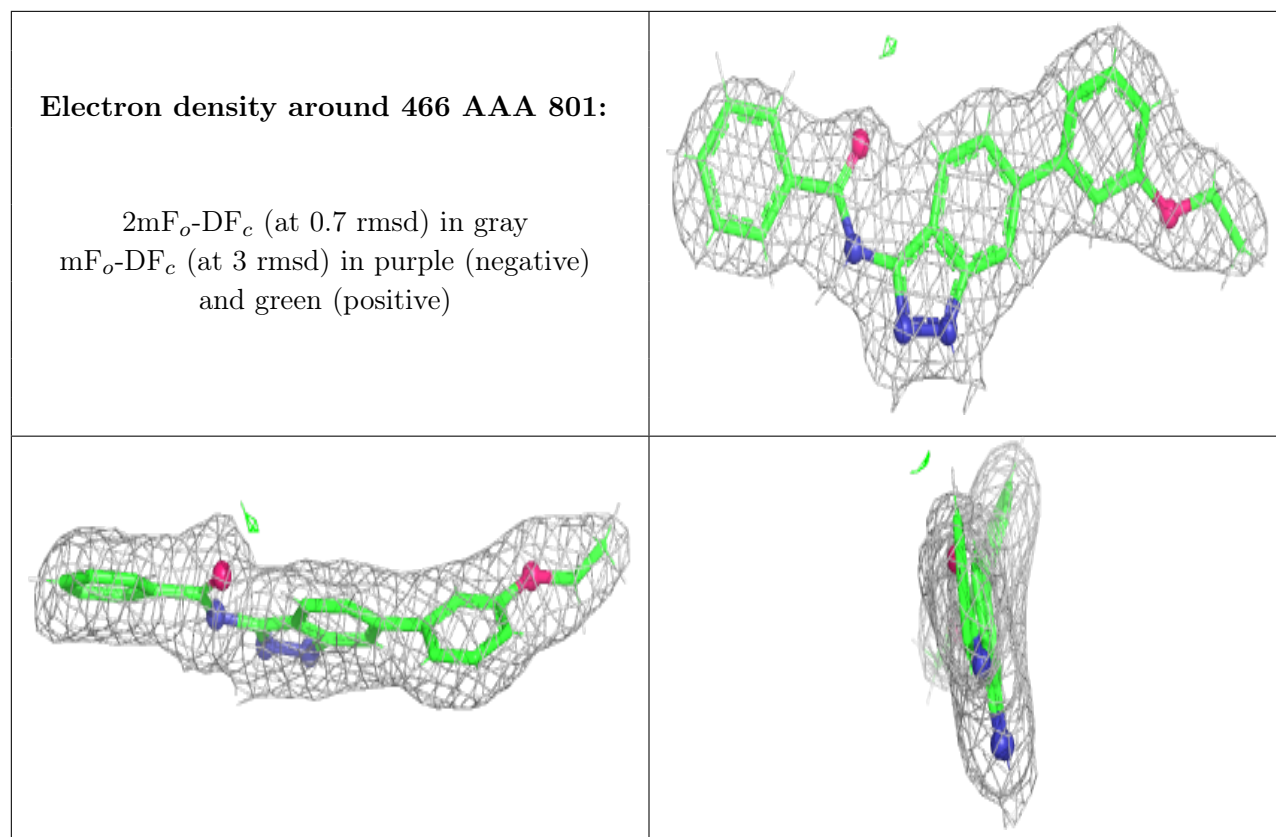
## 6.4 Ligands ⓘ

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
3	SO4	AAA	802	5/5	0.79	0.13	102,103,106,110	0
4	EDO	BBB	802	4/4	0.91	0.12	51,55,56,57	1
5	CL	BBB	804	1/1	0.92	0.27	58,58,58,58	0
4	EDO	BBB	803	4/4	0.93	0.10	53,54,58,58	1
4	EDO	AAA	803	4/4	0.94	0.12	47,50,52,54	1
2	466	BBB	801	27/27	0.94	0.12	40,46,66,67	0
5	CL	AAA	805	1/1	0.96	0.23	61,61,61,61	0
2	466	AAA	801	27/27	0.96	0.10	37,47,55,56	0
4	EDO	AAA	804	4/4	0.98	0.09	42,46,47,47	1

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.