



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 8, 2021 – 02:06 pm GMT

PDB ID : 7B74
Title : Chimeric Streptavidin With A Dimerization Domain For Artificial Transfer Hydrogenation
Authors : Igareta, N.V.; Ward, T.R.
Deposited on : 2020-12-09
Resolution : 1.85 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

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

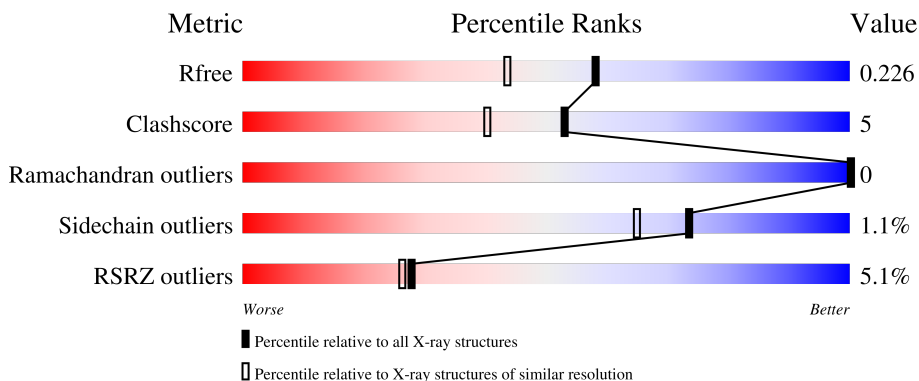
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 ≥ 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	194	<div> <div>0%</div> <div>59%</div> <div>37%</div> </div>
1	BBB	194	<div> <div>2%</div> <div>58%</div> <div>6%</div> <div>36%</div> </div>
1	CCC	194	<div> <div>4%</div> <div>53%</div> <div>10%</div> <div>37%</div> </div>
1	DDD	194	<div> <div>6%</div> <div>56%</div> <div>7%</div> <div>37%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7478 atoms, of which 3467 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 Streptavidin, Superoxide dismutase [Cu-Zn], Streptavidin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	122	Total	C	H	N	O	37	0	0
			1764	568	855	158	183			
1	BBB	125	Total	C	H	N	O	38	2	0
			1811	582	879	161	189			
1	CCC	123	Total	C	H	N	O	39	3	0
			1784	573	866	159	186			
1	DDD	122	Total	C	H	N	O	38	2	0
			1782	573	867	158	184			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P22629
AAA	2	ALA	-	expression tag	UNP P22629
AAA	3	SER	-	expression tag	UNP P22629
AAA	4	MET	-	expression tag	UNP P22629
AAA	5	THR	-	expression tag	UNP P22629
AAA	6	GLY	-	expression tag	UNP P22629
AAA	7	GLY	-	expression tag	UNP P22629
AAA	8	GLN	-	expression tag	UNP P22629
AAA	9	GLN	-	expression tag	UNP P22629
AAA	10	MET	-	expression tag	UNP P22629
AAA	11	GLY	-	expression tag	UNP P22629
AAA	12	ARG	-	expression tag	UNP P22629
AAA	13	ASP	-	expression tag	UNP P22629
AAA	14	GLN	-	expression tag	UNP P22629
AAA	146	ALA	SER	engineered mutation	UNP P22629
BBB	1	MET	-	initiating methionine	UNP P22629
BBB	2	ALA	-	expression tag	UNP P22629
BBB	3	SER	-	expression tag	UNP P22629
BBB	4	MET	-	expression tag	UNP P22629
BBB	5	THR	-	expression tag	UNP P22629
BBB	6	GLY	-	expression tag	UNP P22629

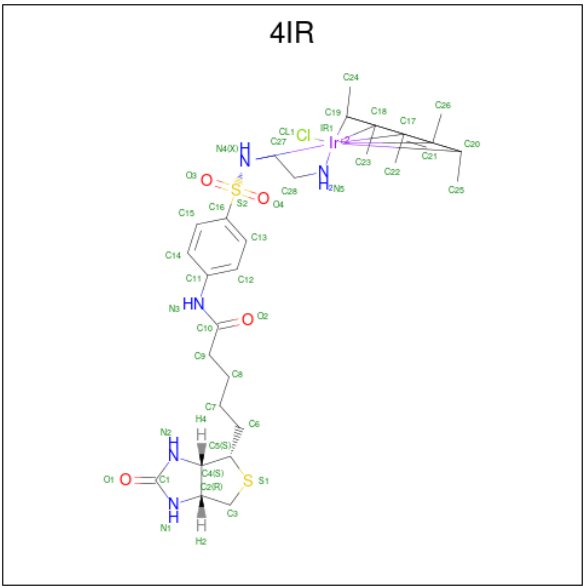
Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	7	GLY	-	expression tag	UNP P22629
BBB	8	GLN	-	expression tag	UNP P22629
BBB	9	GLN	-	expression tag	UNP P22629
BBB	10	MET	-	expression tag	UNP P22629
BBB	11	GLY	-	expression tag	UNP P22629
BBB	12	ARG	-	expression tag	UNP P22629
BBB	13	ASP	-	expression tag	UNP P22629
BBB	14	GLN	-	expression tag	UNP P22629
BBB	146	ALA	SER	engineered mutation	UNP P22629
CCC	1	MET	-	initiating methionine	UNP P22629
CCC	2	ALA	-	expression tag	UNP P22629
CCC	3	SER	-	expression tag	UNP P22629
CCC	4	MET	-	expression tag	UNP P22629
CCC	5	THR	-	expression tag	UNP P22629
CCC	6	GLY	-	expression tag	UNP P22629
CCC	7	GLY	-	expression tag	UNP P22629
CCC	8	GLN	-	expression tag	UNP P22629
CCC	9	GLN	-	expression tag	UNP P22629
CCC	10	MET	-	expression tag	UNP P22629
CCC	11	GLY	-	expression tag	UNP P22629
CCC	12	ARG	-	expression tag	UNP P22629
CCC	13	ASP	-	expression tag	UNP P22629
CCC	14	GLN	-	expression tag	UNP P22629
CCC	146	ALA	SER	engineered mutation	UNP P22629
DDD	1	MET	-	initiating methionine	UNP P22629
DDD	2	ALA	-	expression tag	UNP P22629
DDD	3	SER	-	expression tag	UNP P22629
DDD	4	MET	-	expression tag	UNP P22629
DDD	5	THR	-	expression tag	UNP P22629
DDD	6	GLY	-	expression tag	UNP P22629
DDD	7	GLY	-	expression tag	UNP P22629
DDD	8	GLN	-	expression tag	UNP P22629
DDD	9	GLN	-	expression tag	UNP P22629
DDD	10	MET	-	expression tag	UNP P22629
DDD	11	GLY	-	expression tag	UNP P22629
DDD	12	ARG	-	expression tag	UNP P22629
DDD	13	ASP	-	expression tag	UNP P22629
DDD	14	GLN	-	expression tag	UNP P22629
DDD	146	ALA	SER	engineered mutation	UNP P22629

- Molecule 2 is {N-(4-{[2-(amino-kappaN)ethyl]sulfamoyl-kappaN}phenyl)-5-[(3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl]pentanamide}(chloro)[(1,2,3,4,5-eta)-1,2,3,4,5-pentamethylcyclopentadieny]iridium(III) (three-letter code: 4IR) (formula:

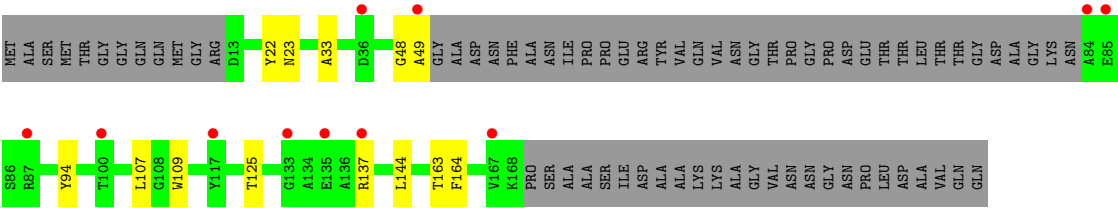
C₂₈H₄₅ClIrN₅O₄S₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	AAA	1	Total	C	Ir	N	O	S	0	0
			40	28	1	5	4	2		
2	BBB	1	Total	C	Ir	N	O	S	0	0
			40	28	1	5	4	2		
2	CCC	1	Total	C	Ir	N	O	S	0	0
			40	28	1	5	4	2		
2	DDD	1	Total	C	Ir	N	O	S	0	0
			40	28	1	5	4	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	38	Total	O	0	0
			38	38		
3	BBB	47	Total	O	0	0
			47	47		
3	CCC	50	Total	O	0	0
			50	50		
3	DDD	42	Total	O	0	0
			42	42		



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	57.28Å 57.31Å 88.03Å 90.00° 94.62° 90.00°	Depositor
Resolution (Å)	47.98 – 1.85 47.98 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.7 (47.98-1.85) 98.3 (47.98-1.85)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.56 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC 5.8.0267	Depositor
R, R_{free}	0.191 , 0.218 0.202 , 0.226	Depositor DCC
R_{free} test set	2495 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å ²)	19.7	Xtriage
Anisotropy	0.692	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7478	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 4IR

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	0.71	0/930	0.88	0/1271
1	BBB	0.75	0/962	0.89	0/1315
1	CCC	0.74	0/947	0.88	0/1294
1	DDD	0.71	0/945	0.90	0/1292
All	All	0.73	0/3784	0.89	0/5172

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	909	855	851	8	0
1	BBB	932	879	873	8	0
1	CCC	918	866	851	13	0
1	DDD	915	867	861	9	0
2	AAA	40	0	24	1	0
2	BBB	40	0	24	3	0
2	CCC	40	0	24	2	0
2	DDD	40	0	24	0	0
3	AAA	38	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	BBB	47	0	0	0	0
3	CCC	50	0	0	0	0
3	DDD	42	0	0	0	0
All	All	4011	3467	3532	40	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 5.

The worst 5 of 40 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)
2:BBB:201:4IR:O2	2:BBB:201:4IR:C14	2.49	0.56
1:BBB:144:LEU:C	1:BBB:144:LEU:HD23	2.27	0.55
2:CCC:201:4IR:O2	2:CCC:201:4IR:C12	2.56	0.52
1:AAA:125:THR:HB	1:CCC:125:THR:HB	1.92	0.51
1:AAA:107:LEU:C	1:AAA:107:LEU:HD12	2.32	0.50

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	118/194 (61%)	114 (97%)	4 (3%)	0	100	100
1	BBB	123/194 (63%)	118 (96%)	5 (4%)	0	100	100
1	CCC	121/194 (62%)	119 (98%)	2 (2%)	0	100	100
1	DDD	120/194 (62%)	118 (98%)	2 (2%)	0	100	100
All	All	482/776 (62%)	469 (97%)	13 (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	88/140 (63%)	87 (99%)	1 (1%)	73	65
1	BBB	91/140 (65%)	91 (100%)	0	100	100
1	CCC	90/140 (64%)	87 (97%)	3 (3%)	38	21
1	DDD	90/140 (64%)	90 (100%)	0	100	100
All	All	359/560 (64%)	355 (99%)	4 (1%)	73	65

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	107	LEU
1	CCC	107	LEU
1	CCC	115	ASN
1	CCC	163[A]	THR

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

4 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

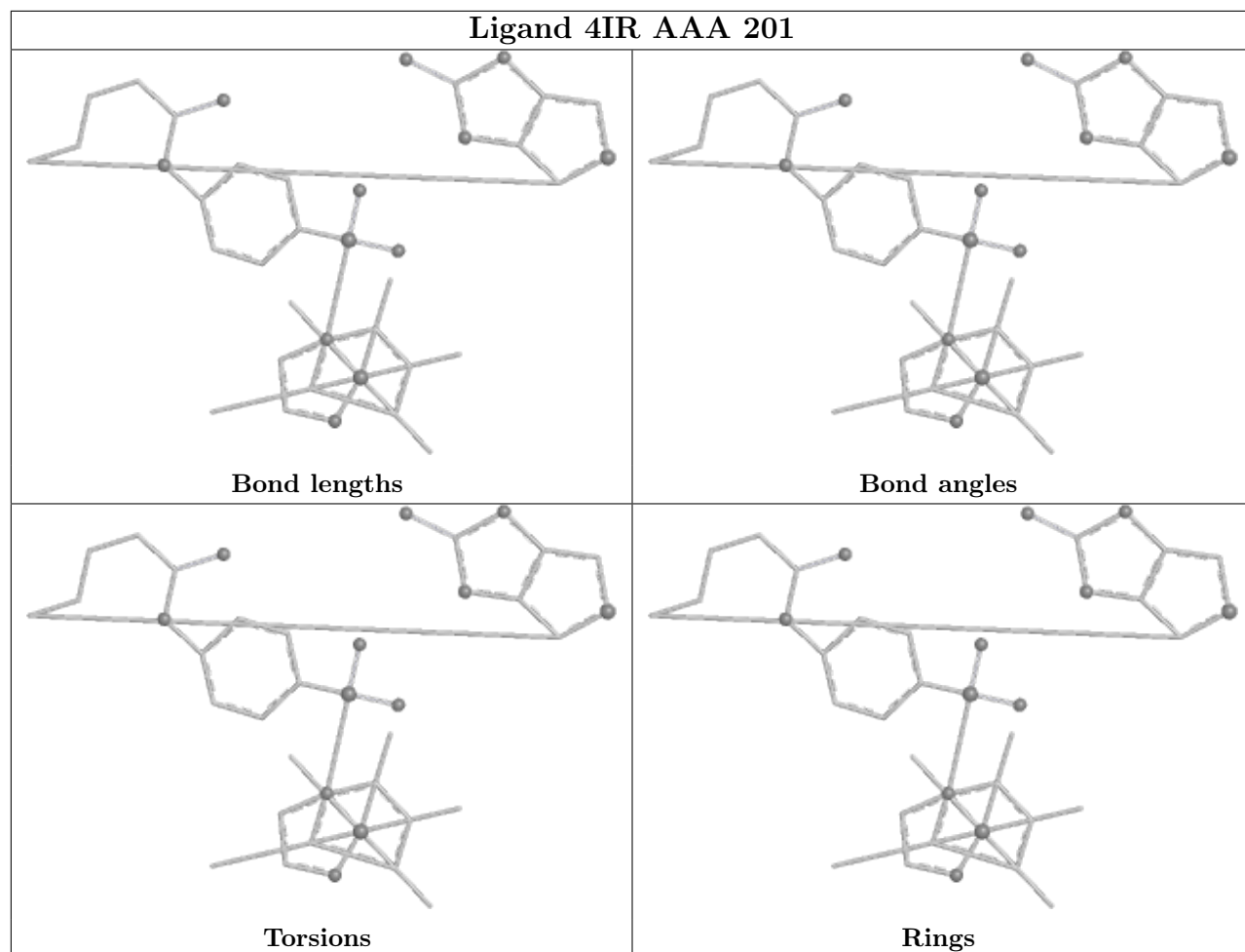
There are no torsion outliers.

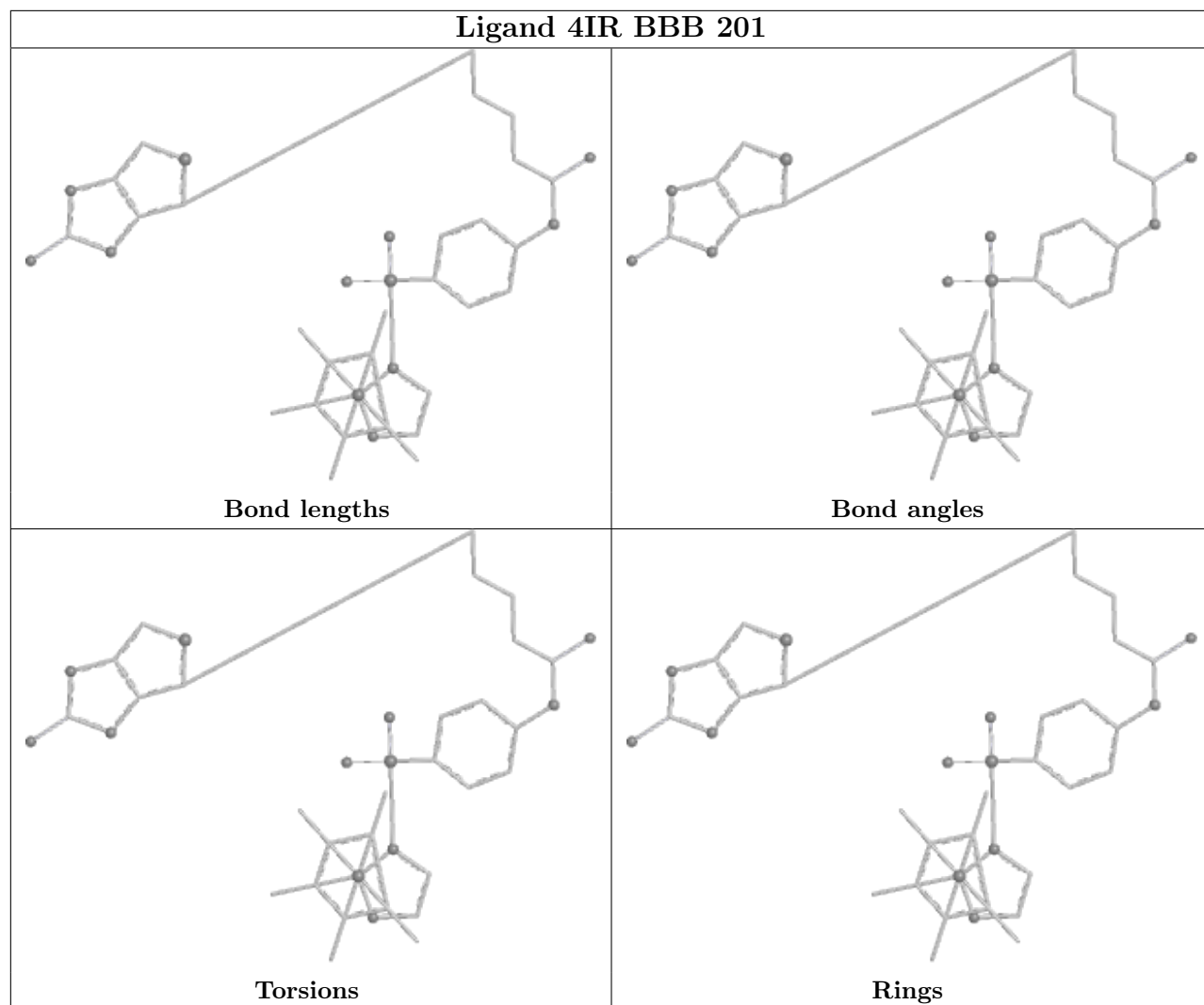
There are no ring outliers.

No monomer is involved in short contacts.

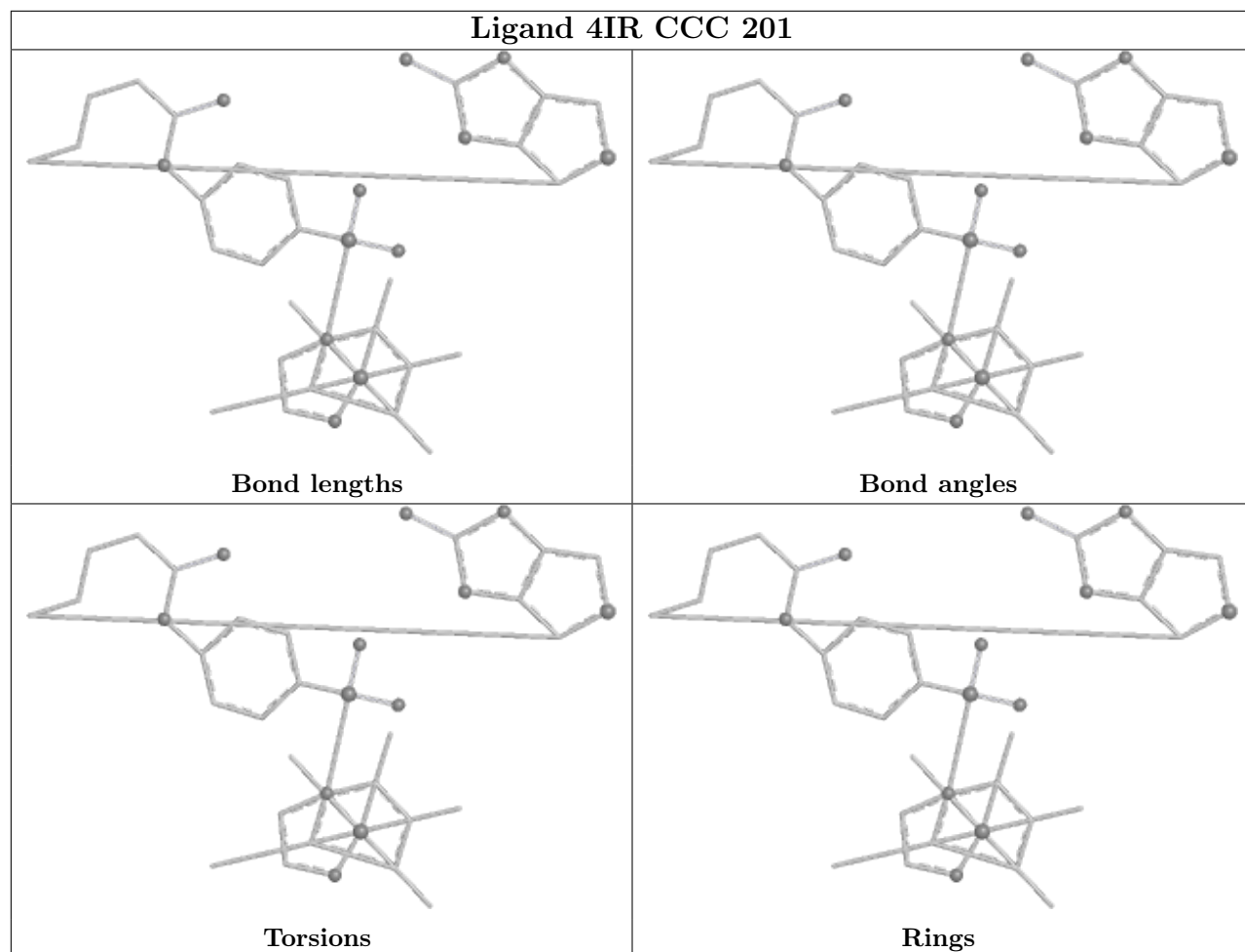
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

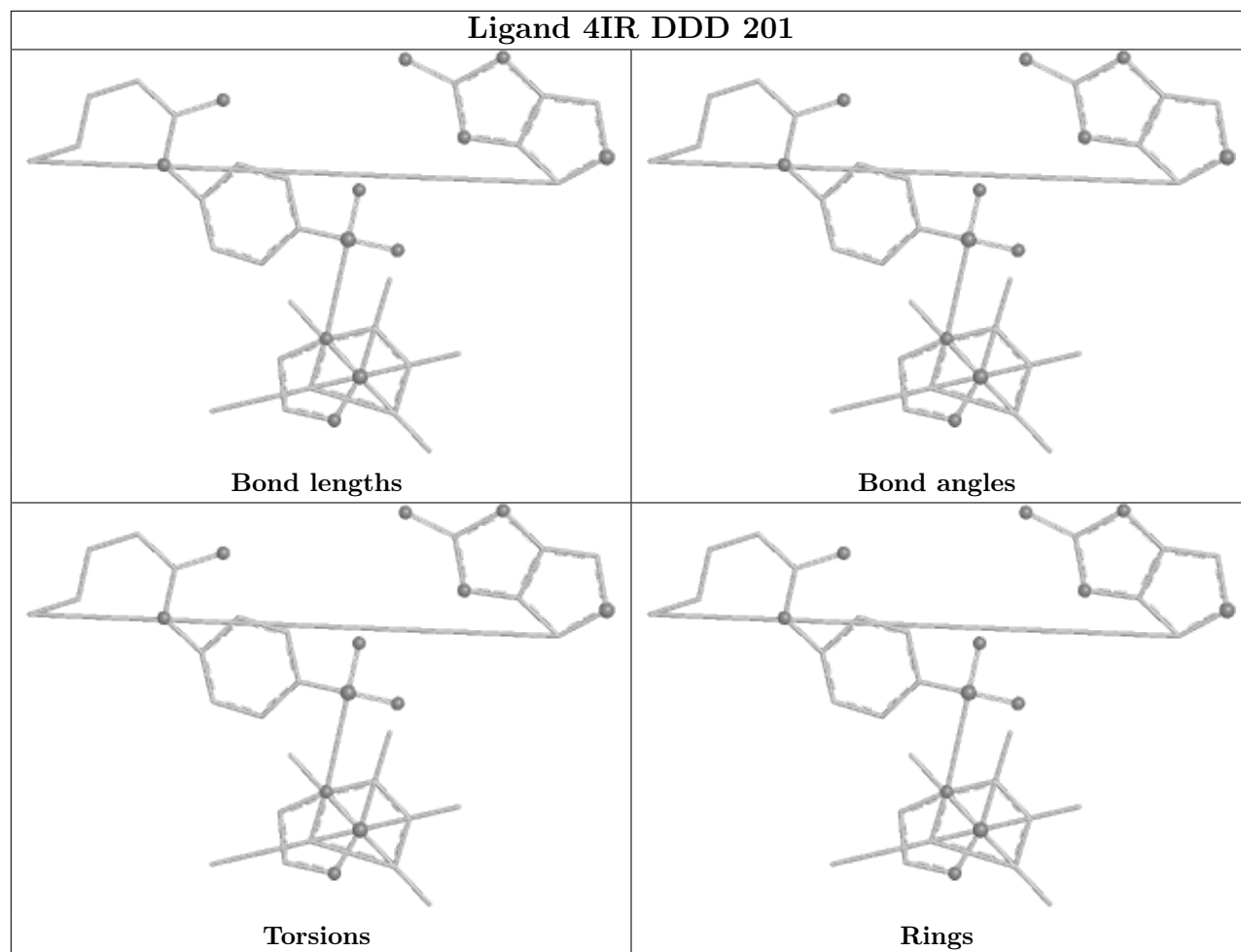
Ligand 4IR AAA 201





Ligand 4IR CCC 201





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, 95th 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 ²)	Q<0.9
1	AAA	122/194 (62%)	0.28	2 (1%) 72 72	14, 23, 46, 56	0
1	BBB	125/194 (64%)	0.30	4 (3%) 47 45	14, 23, 51, 65	0
1	CCC	123/194 (63%)	0.40	8 (6%) 18 18	13, 23, 47, 64	0
1	DDD	122/194 (62%)	0.50	11 (9%) 9 9	14, 23, 49, 63	0
All	All	492/776 (63%)	0.37	25 (5%) 28 26	13, 23, 50, 65	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	117	TYR	4.1
1	DDD	167	VAL	3.7
1	DDD	100	THR	3.4
1	DDD	133	GLY	3.3
1	CCC	13	ASP	3.1

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](#)

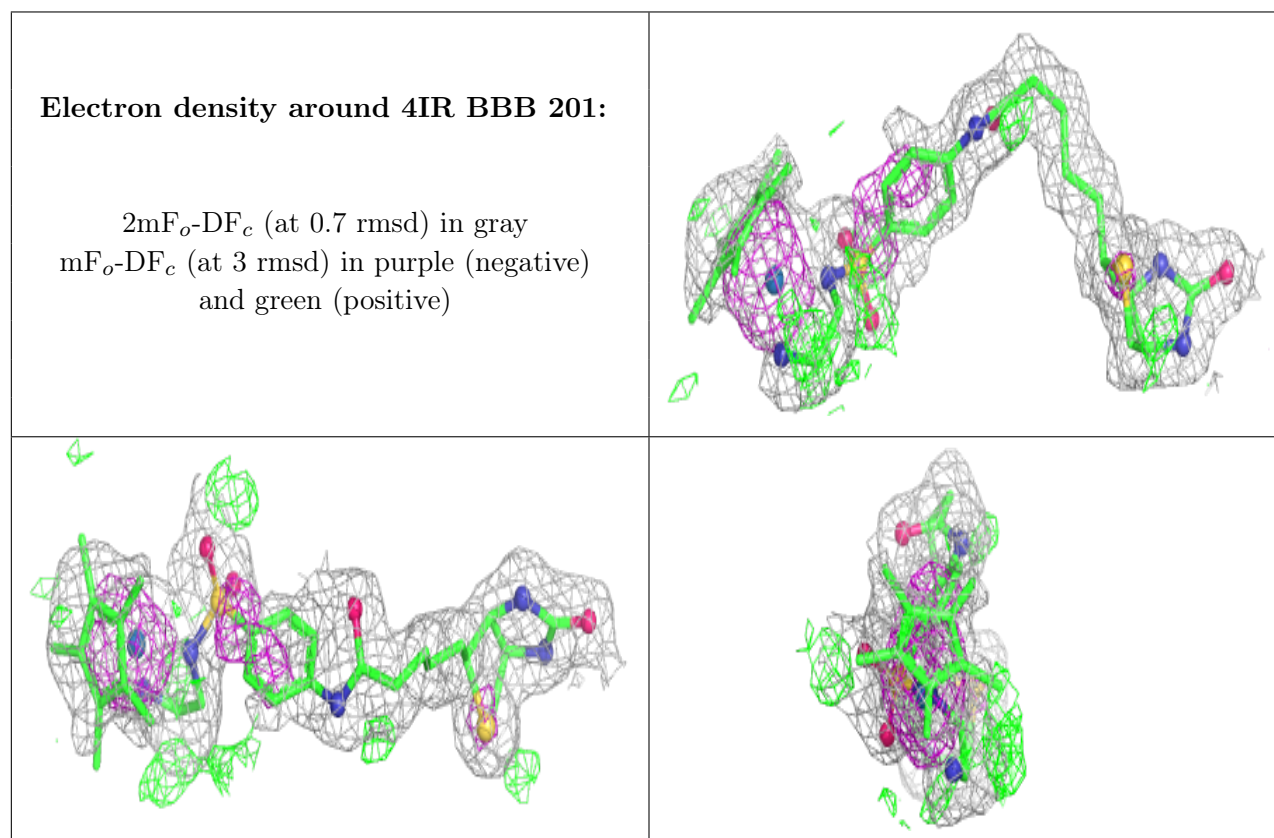
There are no monosaccharides in this entry.

6.4 Ligands [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, 95th 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.

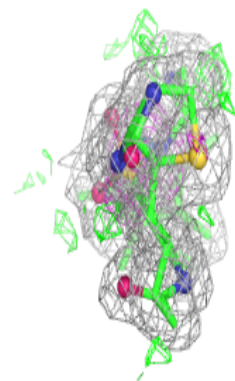
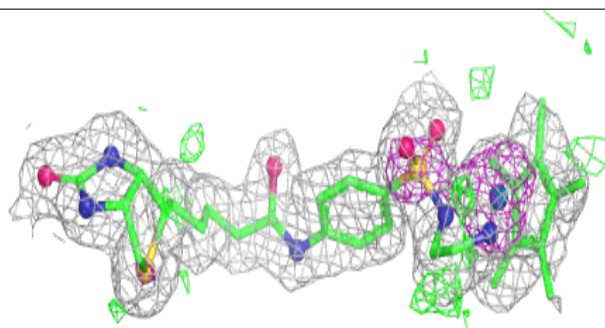
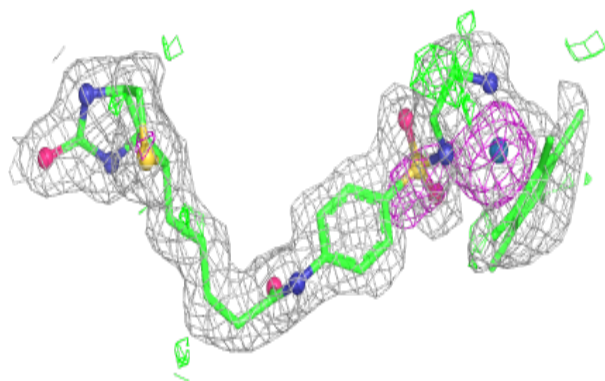
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	4IR	BBB	201	40/41	0.94	0.16	14,38,52,54	1
2	4IR	AAA	201	40/41	0.97	0.14	15,31,43,47	1
2	4IR	CCC	201	40/41	0.97	0.14	14,32,42,43	1
2	4IR	DDD	201	40/41	0.97	0.14	15,30,42,44	1

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.

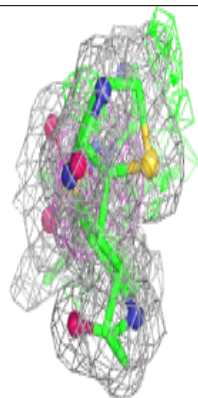
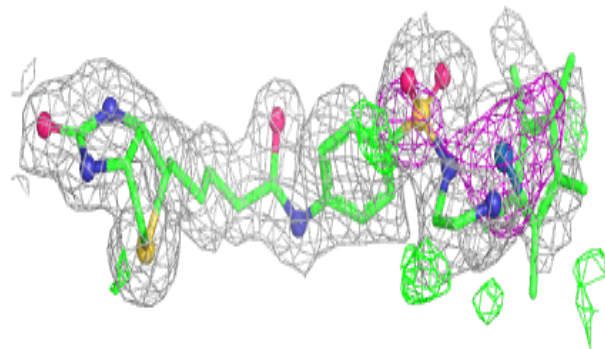
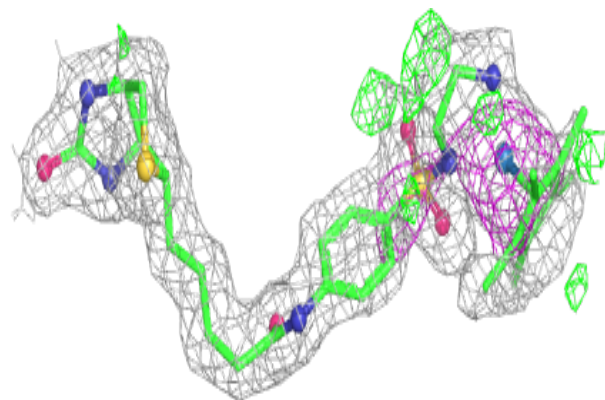


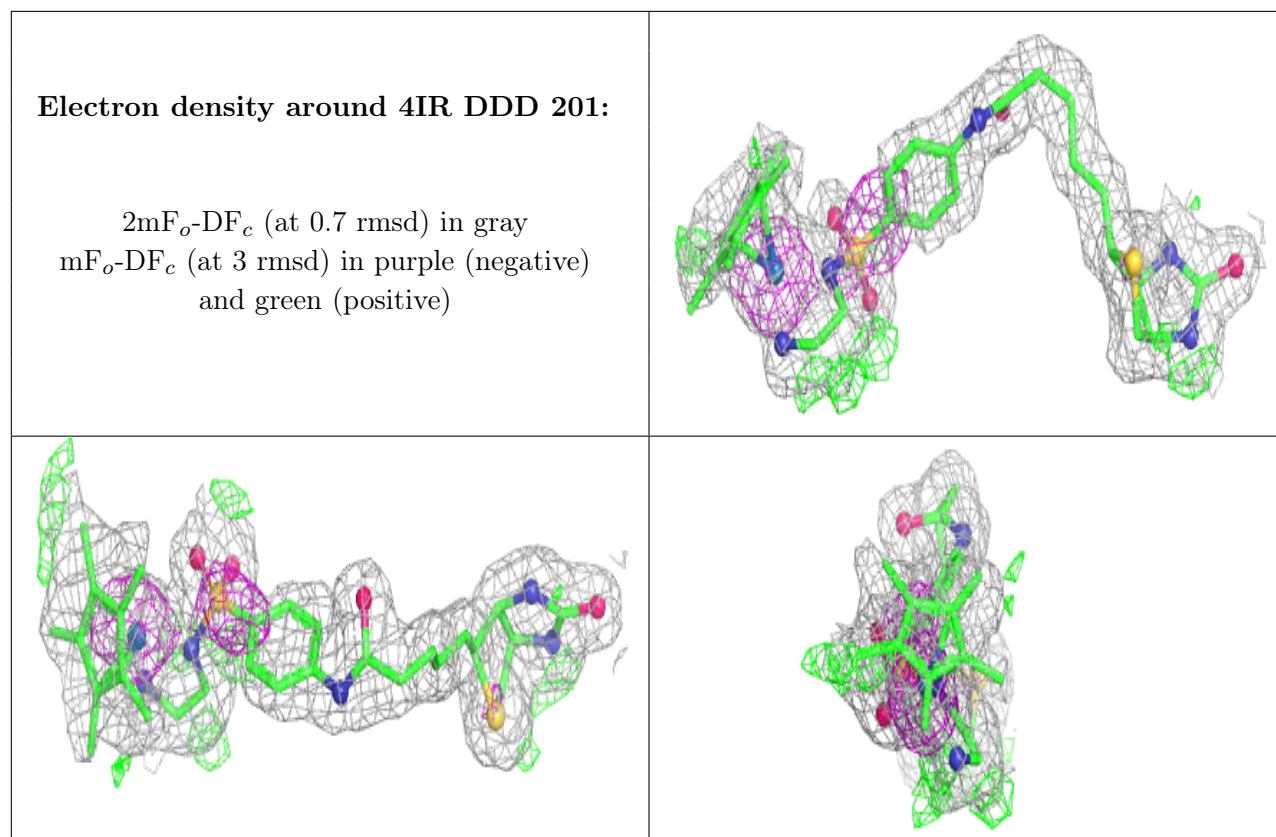
Electron density around 4IR AAA 201:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around 4IR CCC 201:**

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





6.5 Other polymers [i](#)

There are no such residues in this entry.