



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 6, 2022 – 10:12 am BST

PDB ID : 7OU6
Title : Human O-GlcNAc hydrolase in complex with DNJNAc-thiazolidines
Authors : Males, A.; Davies, G.J.; Gonzalez-Cuesta, M.; Mellet, C.O.; Fernandez, J.M.G.; Sidhu, P.; Ashmus, R.; Busmann, J.; Vocadlo, D.J.; Foster, L.
Deposited on : 2021-06-11
Resolution : 2.41 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, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
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)	:	2.27

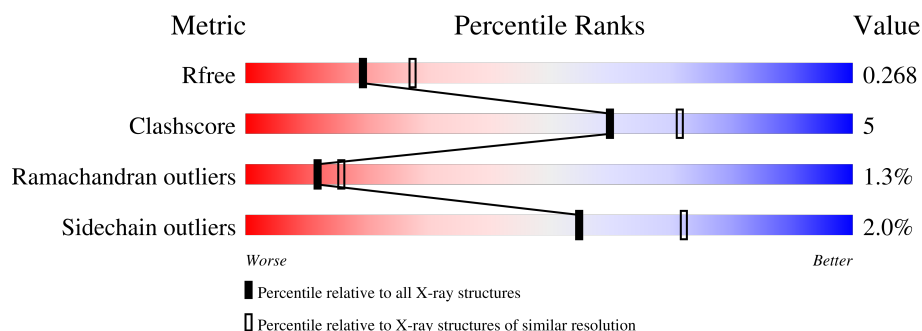
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.41 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	AAA	916	
1	BBB	916	

2 Entry composition [i](#)

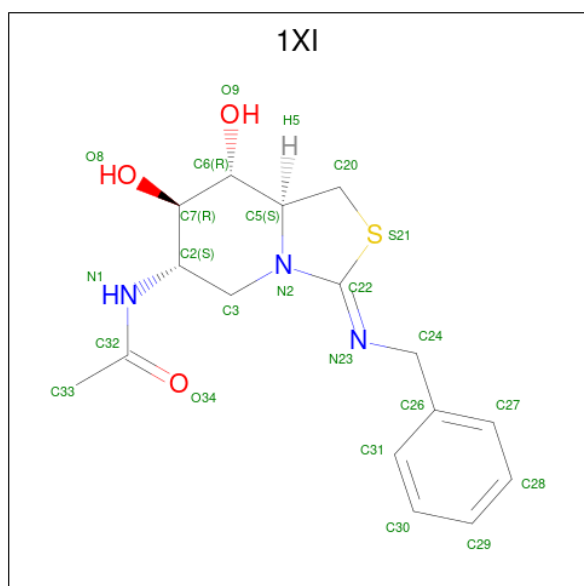
There are 3 unique types of molecules in this entry. The entry contains 13220 atoms, of which 6452 are hydrogens and 0 are deuteriums.

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 Protein O-GlcNAcase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	430	Total	C	H	N	O	S	152	0	0
			6372	2131	3098	531	592	20			
1	BBB	436	Total	C	H	N	O	S	127	0	0
			6742	2236	3312	549	621	24			

- Molecule 2 is {N}-[(3 {Z},6 {S},7 {R},8 {R},8 {a} {S})-7,8-bis(oxidanyl)-3-(phenylmethyl)imino-1,5,6,7,8,8 {a}-hexahydro-[1,3]thiazolo[3,4-a]pyridin-6-yl]ethanamide (three-letter code: 1XI) (formula: C₁₆H₂₁N₃O₃S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	S	2	0
			44	16	21	3	3	1		
2	BBB	1	Total	C	H	N	O	S	2	0
			44	16	21	3	3	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	8	Total 8	O 8	0	0
3	BBB	10	Total 10	O 10	0	0

[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	101.29Å 101.29Å 284.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.28 – 2.41 58.21 – 2.41	Depositor EDS
% Data completeness (in resolution range)	99.9 (58.28-2.41) 100.0 (58.21-2.41)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.221 , 0.271 0.221 , 0.268	Depositor DCC
R_{free} test set	1735 reflections (2.98%)	wwPDB-VP
Wilson B-factor (Å ²)	78.3	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	13220	wwPDB-VP
Average B, all atoms (Å ²)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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: 1XI

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	0/3365	0.83	1/4589 (0.0%)
1	BBB	0.70	0/3525	0.84	0/4795
All	All	0.70	0/6890	0.83	1/9384 (0.0%)

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	BBB	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	AAA	104	ARG	NE-CZ-NH2	-5.04	117.78	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	100	ASP	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	AAA	3274	3098	2955	38	0
1	BBB	3430	3312	3226	30	0
2	AAA	23	21	0	1	0
2	BBB	23	21	0	0	0
3	AAA	8	0	0	0	0
3	BBB	10	0	0	0	0
All	All	6768	6452	6181	59	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 59 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:693:LEU:HB3	1:AAA:694:PRO:CD	2.20	0.71
1:AAA:693:LEU:HB3	1:AAA:694:PRO:HD3	1.80	0.63
1:AAA:241:LEU:O	1:AAA:273:ARG:NH1	2.32	0.61
1:AAA:626:THR:O	1:AAA:630:ASN:ND2	2.38	0.57
1:AAA:142:ASP:OD1	1:BBB:547:LEU:HA	2.06	0.55

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	420/916 (46%)	372 (89%)	41 (10%)	7 (2%)	9	11
1	BBB	424/916 (46%)	396 (93%)	24 (6%)	4 (1%)	17	24
All	All	844/1832 (46%)	768 (91%)	65 (8%)	11 (1%)	12	16

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	587	ALA
1	AAA	693	LEU
1	AAA	189	SER
1	AAA	588	ASN
1	BBB	101	TYR

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	306/797 (38%)	300 (98%)	6 (2%)	55	72
1	BBB	348/797 (44%)	341 (98%)	7 (2%)	55	72
All	All	654/1594 (41%)	641 (98%)	13 (2%)	55	72

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	152	SER
1	BBB	191	PHE
1	BBB	314	PRO
1	BBB	222	THR
1	BBB	278	TRP

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](#)

2 ligands are modelled in this entry.

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
2	1XI	AAA	1001	-	23,25,25	0.72	1 (4%)	24,35,35	1.20	3 (12%)
2	1XI	BBB	1001	-	23,25,25	0.87	1 (4%)	24,35,35	1.05	2 (8%)

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
2	1XI	AAA	1001	-	-	1/9/38/38	0/3/3/3
2	1XI	BBB	1001	-	-	3/9/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BBB	1001	1XI	C22-N23	2.15	1.30	1.26
2	AAA	1001	1XI	C3-N2	-2.06	1.44	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	1001	1XI	C2-C3-N2	-3.54	105.11	110.38
2	AAA	1001	1XI	C20-C5-C6	3.08	120.38	115.72
2	BBB	1001	1XI	C7-C6-C5	-2.46	107.21	111.37
2	BBB	1001	1XI	C6-C7-C2	2.41	114.55	111.02

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	1001	1XI	C6-C7-C2	2.35	114.47	111.02

There are no chirality outliers.

All (4) torsion outliers are listed below:

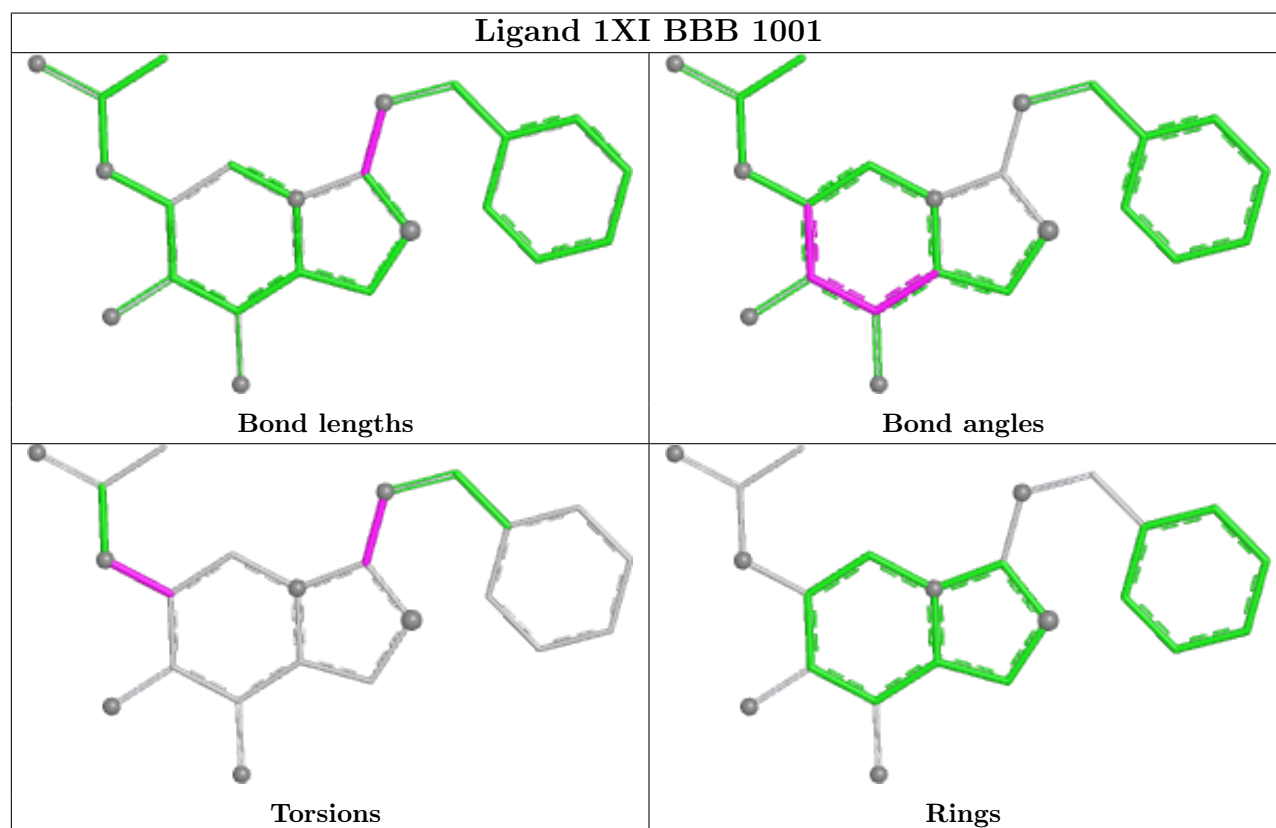
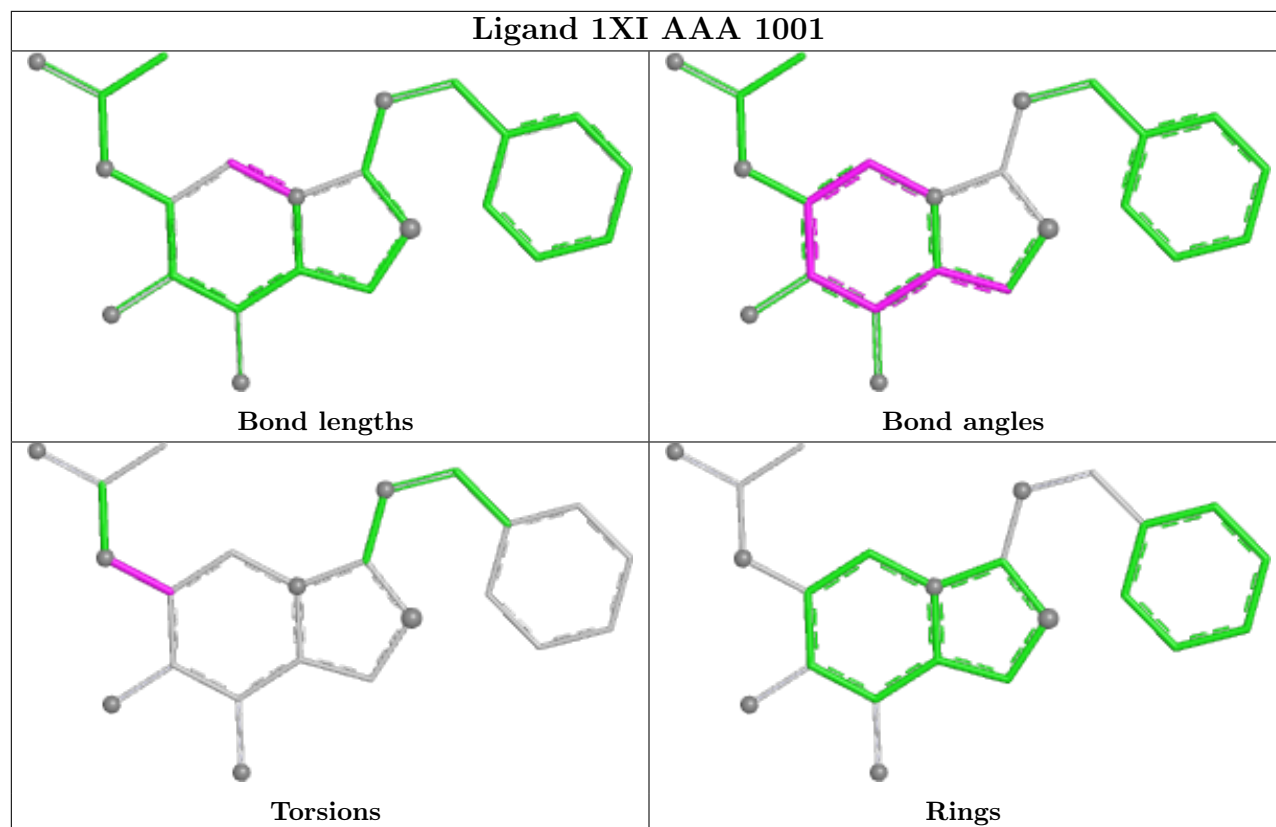
Mol	Chain	Res	Type	Atoms
2	BBB	1001	1XI	N2-C22-N23-C24
2	AAA	1001	1XI	C3-C2-N1-C32
2	BBB	1001	1XI	C3-C2-N1-C32
2	BBB	1001	1XI	C7-C2-N1-C32

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	1001	1XI	1	0

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

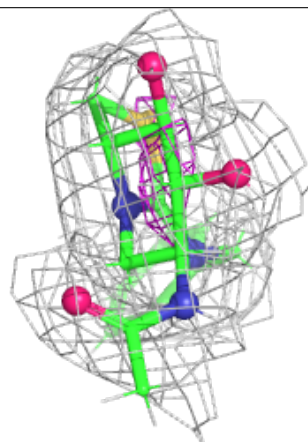
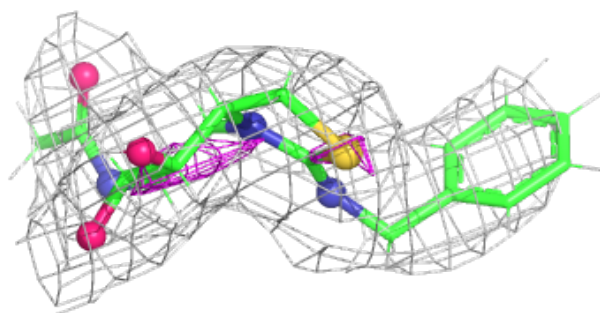
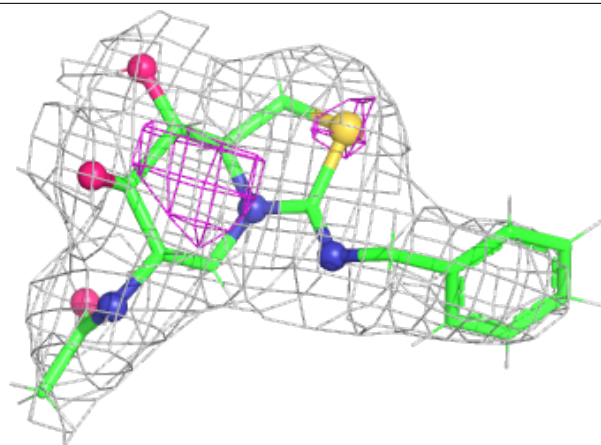
6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

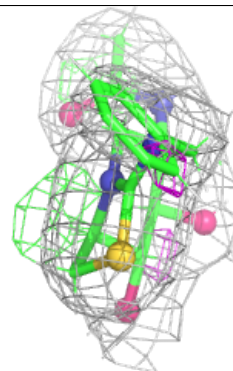
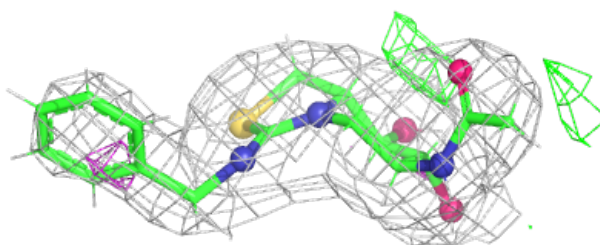
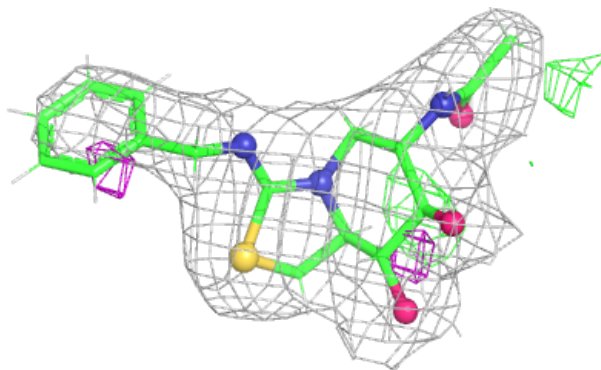
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 1XI 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)

**Electron density around 1XI BBB 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 [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.