



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 08:09 PM UTC

PDB ID : 9PKU / pdb\_00009pku  
Title : Crystal Structure of YEATS domain of human YEATS2 in complex with LS-131 peptide  
Authors : xinyi, Y.; sha, L.  
Deposited on : 2025-07-14  
Resolution : 2.88 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 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  
Mogul : **NOT EXECUTED**  
Xtrriage (Phenix) : 2.0  
EDS : **NOT EXECUTED**  
Buster-report : **NOT EXECUTED**  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

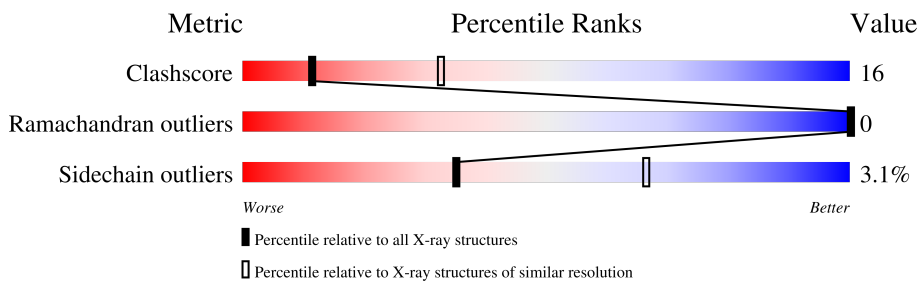
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.88 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))
Clashscore	190562	3801 (2.90-2.86)
Ramachandran outliers	187476	3699 (2.90-2.86)
Sidechain outliers	187428	3702 (2.90-2.86)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	135	
1	C	135	
2	B	5	
2	D	5	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2435 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 YEATS domain-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	135	1137	727	213	195	2	0	0	0
1	C	130	1096	703	204	188	1	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	198	SER	-	expression tag	UNP Q9ULM3
A	199	HIS	-	expression tag	UNP Q9ULM3
A	200	MET	-	expression tag	UNP Q9ULM3
C	198	SER	-	expression tag	UNP Q9ULM3
C	199	HIS	-	expression tag	UNP Q9ULM3
C	200	MET	-	expression tag	UNP Q9ULM3

- Molecule 2 is a protein called LS-131 peptide.

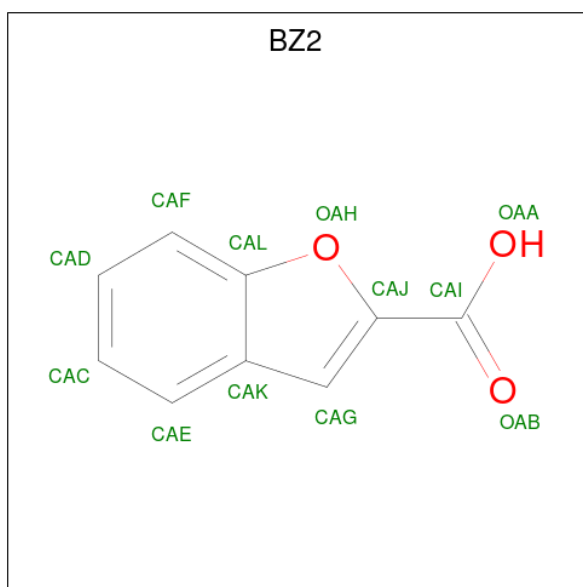
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	5	32	18	6	7	1	0	0	1
2	D	5	32	18	6	7	1	0	0	1

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0

- Molecule 4 is 1-benzofuran-2-carboxylic acid (CCD ID: BZ2) (formula: C<sub>9</sub>H<sub>6</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	C O	0	0
			11	9 2		
4	D	1	Total	C O	0	0
			11	9 2		

- Molecule 5 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	3	Total	Cl	0	0
			3	3		

- Molecule 6 is water.

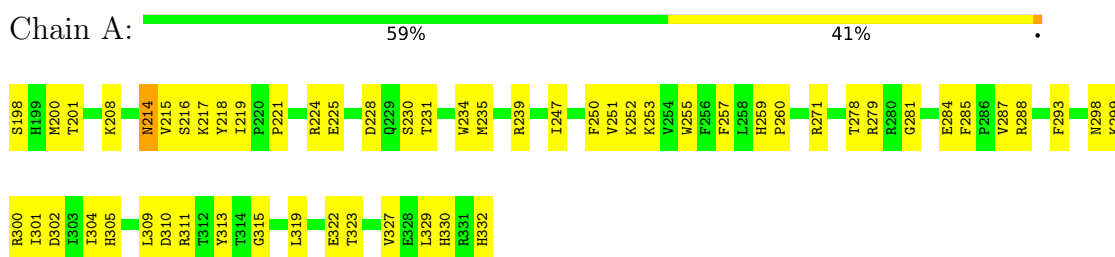
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	33	Total	O	0	0
			33	33		
6	B	4	Total	O	0	0
			4	4		
6	C	27	Total	O	0	0
			27	27		
6	D	1	Total	O	0	0
			1	1		

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

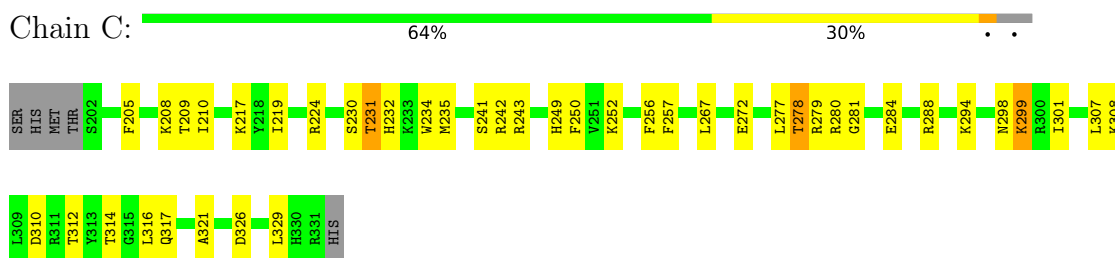
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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.

Note EDS was not executed.

- Molecule 1: YEATS domain-containing protein 2



- Molecule 1: YEATS domain-containing protein 2



- Molecule 2: LS-131 peptide



There are no outlier residues recorded for this chain.

- Molecule 2: LS-131 peptide



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

EDS was not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.40Å 51.56Å 54.38Å 90.00° 93.93° 90.00°	Depositor
Resolution (Å)	49.43 – 2.88	Depositor
% Data completeness (in resolution range)	98.7 (49.43-2.88)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.04 (at 2.91Å)	Xtrriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.208 , 0.251	Depositor
Wilson B-factor (Å <sup>2</sup> )	75.4	Xtrriage
Anisotropy	0.733	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2435	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.04% 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: CL, BZ2, A1EDD, GOL, NH2

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	0.41	0/1170	0.59	0/1583
1	C	0.42	0/1127	0.60	0/1525
2	B	0.83	0/18	1.25	0/24
2	D	1.08	0/18	1.00	0/24
All	All	0.43	0/2333	0.61	0/3156

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	A	1137	0	1136	39	0
1	C	1096	0	1101	39	0
2	B	32	0	17	0	0
2	D	32	0	17	0	0
3	A	18	0	24	0	0
3	C	30	0	40	1	0
4	B	11	0	5	3	0
4	D	11	0	5	2	0
5	C	3	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	33	0	0	1	0
6	B	4	0	0	0	0
6	C	27	0	0	1	0
6	D	1	0	0	0	0
All	All	2435	0	2345	76	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 16.

All (76) 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)
4:B:101:BZ2:HAG	1:C:281:GLY:HA2	1.38	1.04
1:A:281:GLY:HA2	4:D:101:BZ2:HAG	1.41	1.00
1:A:284:GLU:HG3	1:A:309:LEU:HD12	1.54	0.88
1:C:252:LYS:HD3	1:C:294:LYS:HG2	1.68	0.74
1:C:257:PHE:HB2	1:C:288:ARG:HB3	1.70	0.74
1:C:299:LYS:HD3	1:C:299:LYS:H	1.53	0.73
4:B:101:BZ2:HAG	1:C:281:GLY:CA	2.18	0.72
1:A:208:LYS:NZ	6:A:501:HOH:O	2.23	0.71
1:A:198:SER:HB3	1:A:201:THR:OG1	1.95	0.67
1:A:281:GLY:CA	4:D:101:BZ2:HAG	2.23	0.65
1:A:301:ILE:HD13	1:A:327:VAL:HG11	1.80	0.63
1:A:311:ARG:NH1	1:A:313:TYR:OH	2.34	0.60
1:C:232:HIS:CD2	1:C:310:ASP:H	2.19	0.60
1:A:225:GLU:N	1:A:228:ASP:OD1	2.28	0.59
1:C:235:MET:HG2	1:C:278:THR:HB	1.84	0.59
1:C:277:LEU:HD11	1:C:279:ARG:NH1	2.18	0.58
1:C:312:THR:HB	1:C:314:THR:HG23	1.86	0.58
1:C:279:ARG:HG3	1:C:279:ARG:HH11	1.69	0.57
1:A:218:TYR:O	1:A:315:GLY:HA2	2.05	0.57
1:A:257:PHE:HB2	1:A:288:ARG:HB3	1.87	0.56
1:C:210:ILE:HD11	1:C:250:PHE:HE2	1.70	0.56
1:A:253:LYS:HD2	1:A:255:TRP:NE1	2.21	0.56
4:B:101:BZ2:CAG	1:C:281:GLY:HA2	2.23	0.55
1:C:208:LYS:HD2	5:C:401:CL:CL	2.44	0.54
1:C:241:SER:OG	1:C:243:ARG:HG3	2.09	0.53
1:C:277:LEU:HD11	1:C:279:ARG:HH11	1.74	0.53
1:C:314:THR:HB	1:C:316:LEU:HD13	1.91	0.52
1:A:221:PRO:HG3	1:A:224:ARG:HH21	1.76	0.51
1:A:239:ARG:NH2	1:A:322:GLU:OE2	2.44	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:284:GLU:OE2	1:C:308:LYS:HA	2.12	0.50
1:C:231:THR:OG1	1:C:232:HIS:ND1	2.42	0.50
1:A:200:MET:SD	1:A:200:MET:N	2.78	0.49
1:A:313:TYR:CZ	1:C:242:ARG:HG3	2.47	0.48
1:C:232:HIS:HD2	1:C:310:ASP:H	1.61	0.48
1:A:284:GLU:CG	1:A:309:LEU:HD12	2.34	0.48
1:C:210:ILE:HD11	1:C:250:PHE:CE2	2.48	0.48
1:A:288:ARG:HD2	1:A:302:ASP:OD2	2.15	0.47
1:A:214:ASN:HA	1:A:235:MET:O	2.15	0.47
1:A:247:ILE:HD12	1:A:251:VAL:HG21	1.95	0.47
1:C:235:MET:CG	1:C:278:THR:HB	2.45	0.46
1:A:216:SER:OG	1:A:310:ASP:N	2.49	0.46
1:A:288:ARG:HD3	1:A:304:ILE:HD11	1.97	0.46
1:C:256:PHE:O	1:C:267:LEU:HA	2.16	0.45
1:C:301:ILE:HD12	6:C:516:HOH:O	2.15	0.45
1:C:219:ILE:HG22	1:C:224:ARG:HG3	1.98	0.45
1:C:308:LYS:HD3	1:C:321:ALA:HB2	1.97	0.45
1:A:293:PHE:CZ	1:A:301:ILE:HG12	2.51	0.44
1:A:219:ILE:HG22	1:A:224:ARG:HD3	1.99	0.44
1:A:224:ARG:NH1	1:A:230:SER:O	2.49	0.44
1:A:285:PHE:CE2	1:A:287:VAL:HG22	2.52	0.44
1:C:272:GLU:OE2	3:C:407:GOL:H12	2.18	0.44
1:A:214:ASN:HB3	1:A:305:HIS:NE2	2.33	0.44
1:A:234:TRP:CZ2	1:A:279:ARG:HB2	2.53	0.44
1:A:252:LYS:HZ3	1:A:271:ARG:HH12	1.67	0.43
1:C:280:ARG:H	1:C:280:ARG:HG2	1.74	0.43
1:C:299:LYS:HD3	1:C:299:LYS:N	2.25	0.43
1:C:205:PHE:HB2	5:C:402:CL:CL	2.55	0.43
1:C:249:HIS:H	1:C:249:HIS:CD2	2.35	0.43
1:C:257:PHE:CE2	1:C:267:LEU:HD12	2.54	0.43
1:C:217:LYS:HB3	1:C:317:GLN:HG3	2.01	0.42
1:A:259:HIS:CD2	1:A:260:PRO:HD2	2.54	0.42
1:C:234:TRP:CG	1:C:307:LEU:HD21	2.54	0.42
1:A:224:ARG:NH1	1:A:231:THR:HA	2.34	0.42
1:C:209:THR:HG23	1:C:326:ASP:OD2	2.20	0.42
1:C:279:ARG:HH11	1:C:279:ARG:CG	2.32	0.42
1:A:215:VAL:HG12	1:A:319:LEU:HA	2.01	0.42
1:C:299:LYS:H	1:C:299:LYS:CD	2.26	0.41
1:A:247:ILE:HA	1:A:250:PHE:CE1	2.55	0.41
1:A:253:LYS:HD2	1:A:255:TRP:HE1	1.84	0.41
1:A:235:MET:HG3	1:A:278:THR:OG1	2.21	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:298:ASN:ND2	1:A:329:LEU:HD12	2.35	0.41
1:A:313:TYR:CE2	1:C:242:ARG:HG3	2.56	0.41
1:A:299:LYS:HG2	1:A:300:ARG:N	2.35	0.41
1:A:214:ASN:ND2	1:A:214:ASN:C	2.80	0.40
1:A:329:LEU:HB3	1:A:330:HIS:H	1.72	0.40
1:C:298:ASN:ND2	1:C:329:LEU:HD23	2.37	0.40

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	133/135 (98%)	129 (97%)	4 (3%)	0	100	100
1	C	128/135 (95%)	124 (97%)	4 (3%)	0	100	100
2	B	3/5 (60%)	3 (100%)	0	0	100	100
2	D	3/5 (60%)	2 (67%)	1 (33%)	0	100	100
All	All	267/280 (95%)	258 (97%)	9 (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	A	128/128 (100%)	124 (97%)	4 (3%)	35	66
1	C	123/128 (96%)	119 (97%)	4 (3%)	33	65
2	B	2/2 (100%)	2 (100%)	0	100	100
2	D	2/2 (100%)	2 (100%)	0	100	100
All	All	255/260 (98%)	247 (97%)	8 (3%)	35	66

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	214	ASN
1	A	217	LYS
1	A	323	THR
1	A	332	HIS
1	C	230	SER
1	C	231	THR
1	C	278	THR
1	C	299	LYS

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

Mol	Chain	Res	Type
1	A	290	GLN
1	C	306	ASN
1	C	317	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

Mogul was not executed - this section is therefore empty.

## 5.5 Carbohydrates [i](#)

Mogul was not executed - this section is therefore empty.

## 5.6 Ligand geometry [i](#)

Mogul was not executed - this section is therefore empty.

## 5.7 Other polymers [i](#)

Mogul was not executed - this section is therefore empty.

## 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

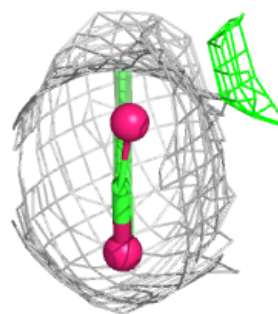
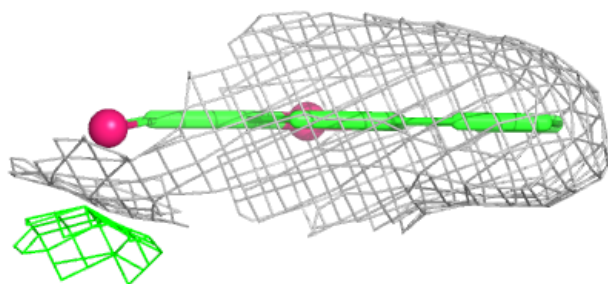
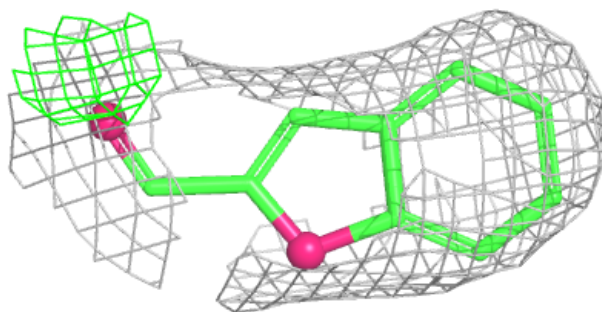
### 6.4 Ligands

EDS was not executed - this section is therefore empty.

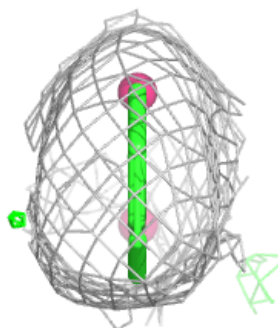
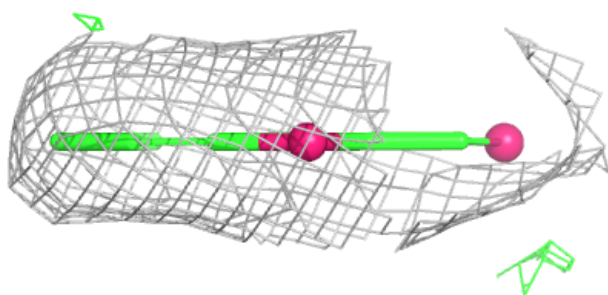
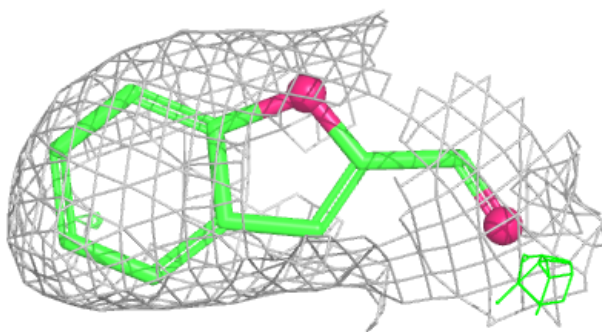
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around BZ2 B 101:**

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

**Electron density around BZ2 D 101:**

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



## 6.5 Other polymers

EDS was not executed - this section is therefore empty.