wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 12:11 PM EDT

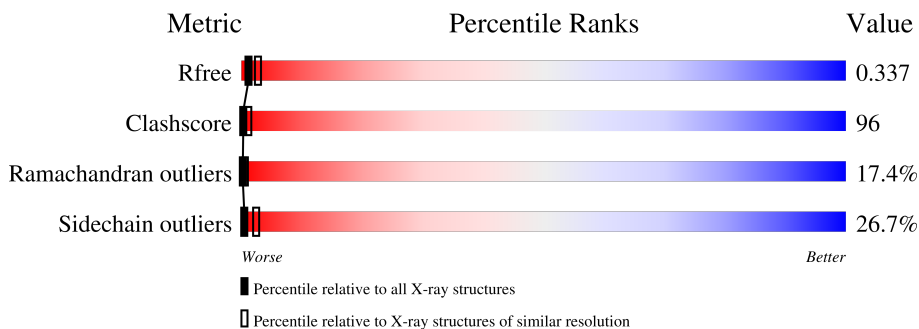|		
PDB ID	:	5C6P
Title	:	protein C
Authors	:	Lu, M.
Deposited on	:	2015-06-23
Resolution	:	3.00 Å(reported)

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 3.00 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	Percentile Ranks	Value
Rfree		0.337
Clashscore		96
Ramachandran outliers		17.4%
Sidechain outliers		26.7%

Worse — Better

▮ Percentile relative to all X-ray structures
◻ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	459	15% 55% 25% 6%
2	B	99	15% 46% 26% • 8%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4248 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 protein C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	459	Total 3508	C 2338	N 558	O 589	S 23	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	460	SER	-	expression tag	UNP Q5F9J8
A	461	SER	-	expression tag	UNP Q5F9J8
A	462	GLY	-	expression tag	UNP Q5F9J8
A	463	LEU	-	expression tag	UNP Q5F9J8

- Molecule 2 is a protein called protein D.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	91	Total 707	C 457	N 110	O 140	0	0	0

- Molecule 3 is (2S)-2-(3,4-dimethoxyphenyl)-5-{[2-(3,4-dimethoxyphenyl)ethyl](methyl)amino}-2-(propan-2-yl)pentanenitrile (three-letter code: 4YH) (formula: $C_{27}H_{38}N_2O_4$).
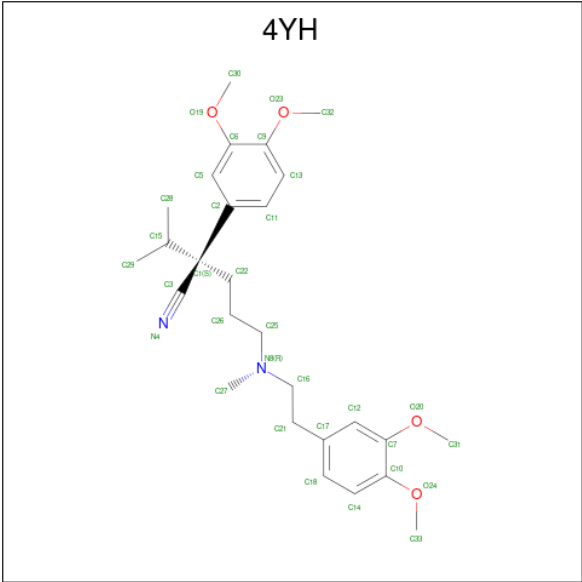
| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	33	27	2	4	0	0

3 Residue-property plots ⓘ

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.

- Molecule 1: protein C

Chain A:
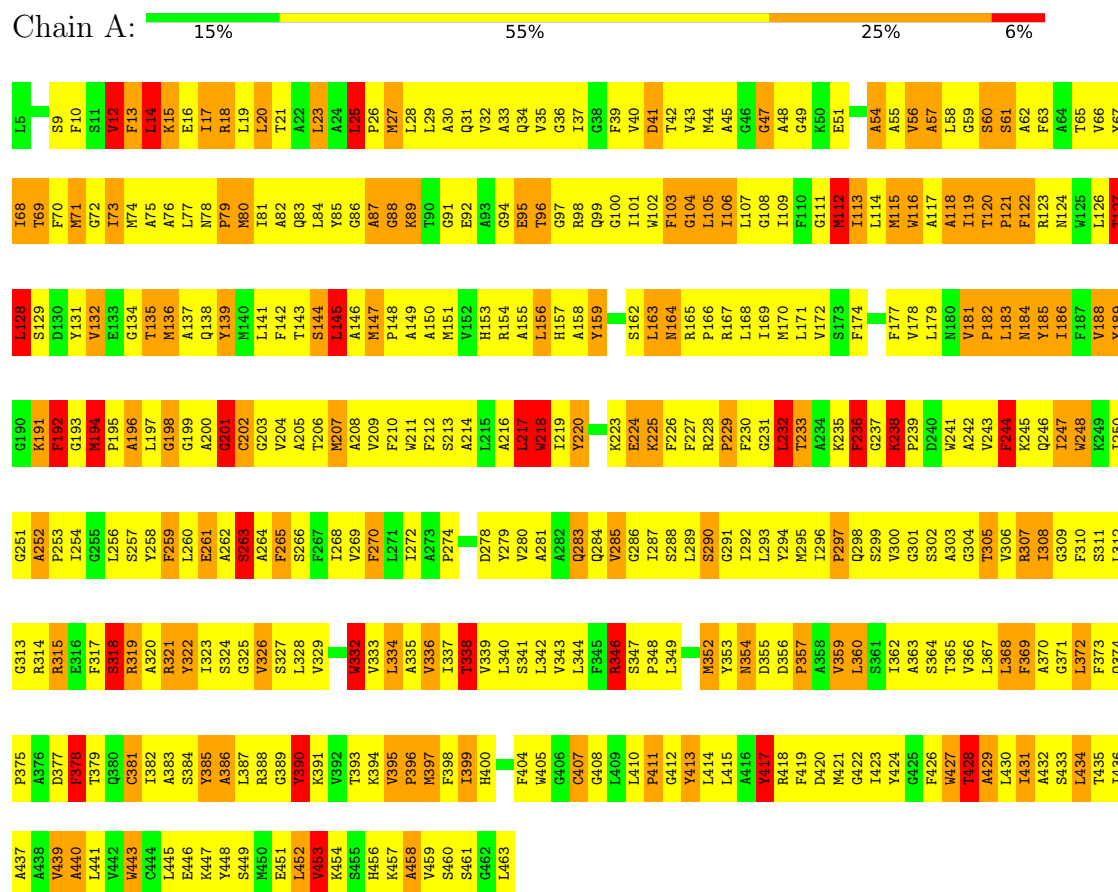
- Molecule 2: protein D

Chain B:
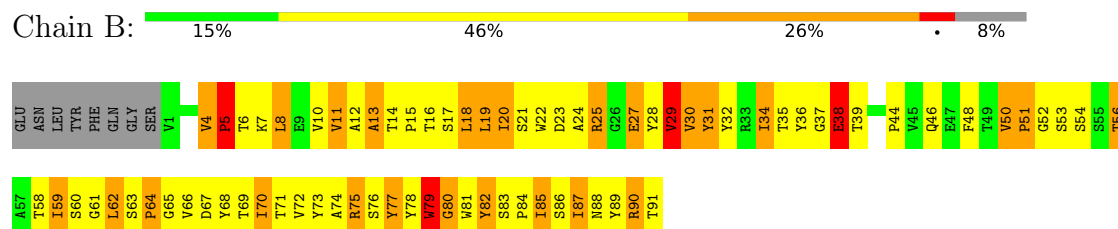
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	117.64Å 117.64Å 225.99Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 20.00 – 3.01	Depositor EDS
% Data completeness (in resolution range)	91.0 (20.00-3.00) 94.7 (20.00-3.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.29 (at 3.04Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.280 , 0.290 0.344 , 0.337	Depositor DCC
R_{free} test set	1719 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	75.5	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.12 , 65.6	EDS
L-test for twinning[2]	$< |L| > = 0.34, < L^2 > = 0.17$	Xtriage
Estimated twinning fraction	0.159 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4248	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: 4YH

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	0.59	2/3605 (0.1%)	0.93	8/4898 (0.2%)
2	B	0.71	0/729	1.13	3/1004 (0.3%)
All	All	0.61	2/4334 (0.0%)	0.97	11/5902 (0.2%)

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	13
2	B	0	4
All	All	0	17

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	346	ARG	C-O	10.16	1.42	1.23
1	A	332	TRP	CB-CG	-6.61	1.38	1.50

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	145	LEU	CA-CB-CG	11.35	141.41	115.30
1	A	346	ARG	O-C-N	-11.18	104.81	122.70
2	B	8	LEU	CA-CB-CG	9.48	137.10	115.30
1	A	198	GLY	N-CA-C	-6.31	97.33	113.10
1	A	127	THR	N-CA-C	5.90	126.94	111.00

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	126	LEU	Peptide
1	A	127	THR	Peptide
1	A	14	LEU	Peptide
1	A	25	LEU	Peptide
1	A	54	ALA	Peptide

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	A	3508	0	3587	666	1
2	B	707	0	681	168	1
3	A	33	0	38	9	0
All	All	4248	0	4306	825	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 96.

The worst 5 of 825 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:458:ALA:HB3	2:B:8:LEU:O	1.31	1.26
2:B:27:GLU:HB3	2:B:52:GLY:O	1.39	1.23
1:A:158:ALA:CB	1:A:217:LEU:HD11	1.67	1.23
1:A:292:ILE:HG13	1:A:293:LEU:H	1.11	1.13
1:A:151:MET:SD	1:A:214:ALA:CB	2.39	1.10

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:196:ALA:N	2:B:31:TYR:OH[5_545]	2.13	0.07

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	457/459 (100%)	258 (56%)	117 (26%)	82 (18%)	0	0
2	B	89/99 (90%)	55 (62%)	21 (24%)	13 (15%)	0	1
All	All	546/558 (98%)	313 (57%)	138 (25%)	95 (17%)	0	0

5 of 95 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	PHE
1	A	20	LEU
1	A	80	MET
1	A	106	ILE
1	A	119	ILE

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	353/353 (100%)	258 (73%)	95 (27%)	0	2
2	B	78/85 (92%)	58 (74%)	20 (26%)	0	3
All	All	431/438 (98%)	316 (73%)	115 (27%)	0	2

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	261	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	75	ARG
1	A	338	THR
2	B	70	ILE
2	B	25	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	34	GLN
1	A	99	GLN
1	A	246	GLN
1	A	284	GLN

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 ⓘ

1 ligand is 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
3	4YH	A	601	-	32,34,34	1.71	3 (9%)	36,46,46	1.82	7 (19%)

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	4YH	A	601	-	-	20/32/35/35	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	4YH	C1-C2	-6.80	1.40	1.53
3	A	601	4YH	C3-N4	4.01	1.20	1.14
3	A	601	4YH	C21-C17	-3.64	1.41	1.51

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	601	4YH	C1-C3-N4	-6.66	166.22	177.45
3	A	601	4YH	C31-O20-C7	-5.40	109.38	117.53
3	A	601	4YH	C30-O19-C6	-3.25	112.62	117.53
3	A	601	4YH	O20-C7-C12	-2.30	120.17	124.12
3	A	601	4YH	O20-C7-C10	2.28	118.58	115.41

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	601	4YH	C3-C1-C2-C5
3	A	601	4YH	C3-C1-C2-C11
3	A	601	4YH	C2-C1-C15-C29
3	A	601	4YH	C2-C1-C15-C28
3	A	601	4YH	C3-C1-C15-C29

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	4YH	9	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 then 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 4YH A 601

Bond lengths

Bond angles

Torsions

Rings
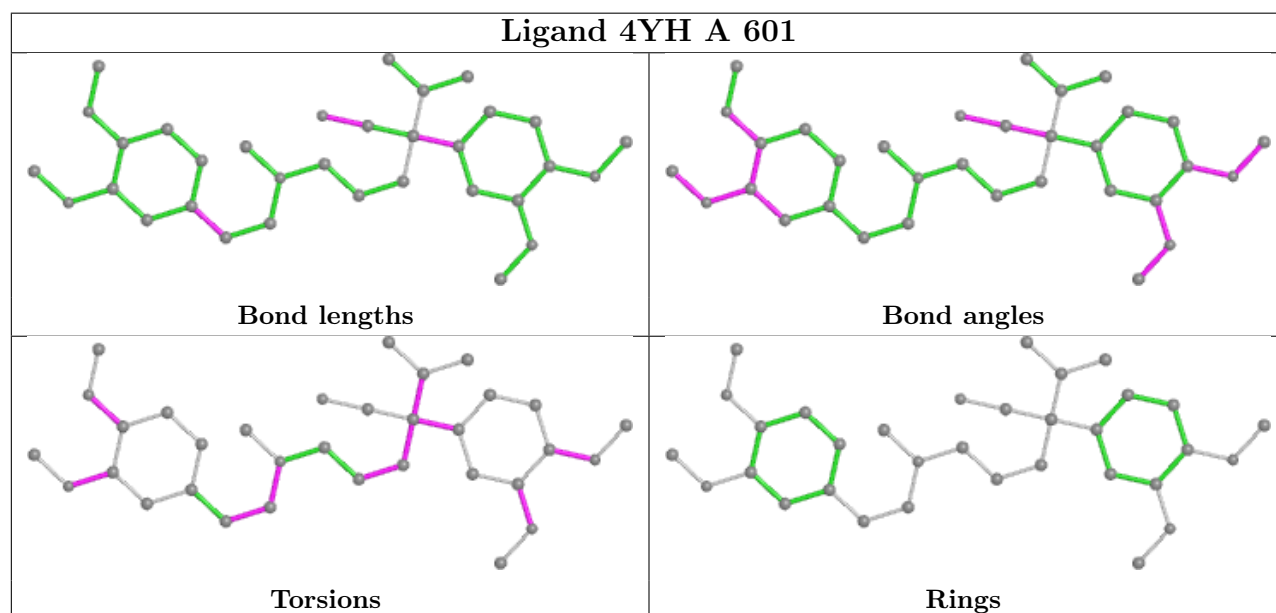
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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.
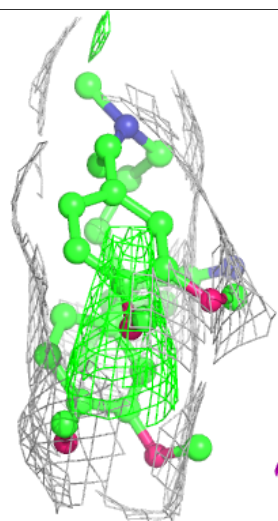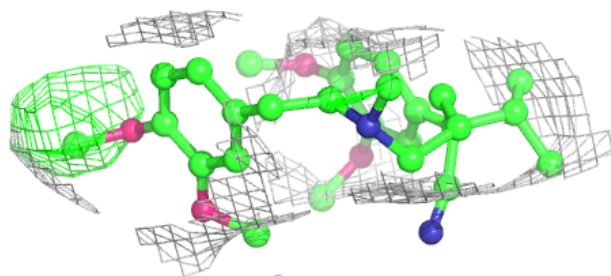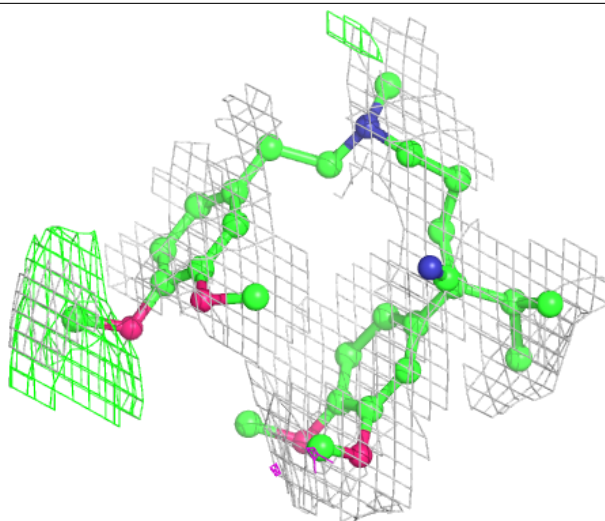
6.4 Ligands ⓘ

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 4YH A 601:

$2mF_o$-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 ⓘ

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