W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 11:05 PM BST

PDB ID	:	3OPM
Title	:	Crystal Structure of Human DPP4 Bound to TAK-294
Authors	:	Yano, J.K.; Aertgeerts, K.
Deposited on	:	2010-09-01
Resolution	:	2.72 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

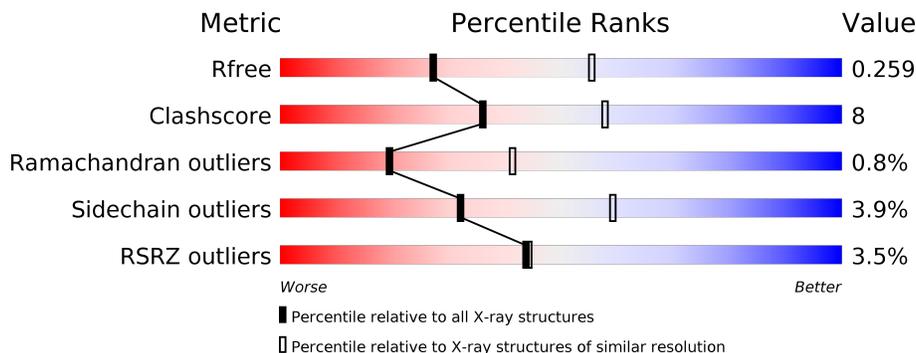|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.72 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		0.259
Clashscore		8
Ramachandran outliers		0.8%
Sidechain outliers		3.9%
RSRZ outliers		3.5%

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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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 on 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% 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	A	740	3% / 79% 18% ··
1	B	740	2% / 79% 19% ··
1	C	740	5% / 76% 20% · ·
1	D	740	4% / 82% 16% ·
2	E	2	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 24713 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	727	Total 5950	C 3821	N 977	O 1126	S 26	0	0	0
1	B	733	Total 6013	C 3857	N 997	O 1133	S 26	0	0	0
1	C	725	Total 5937	C 3810	N 979	O 1122	S 26	0	0	0
1	D	727	Total 5957	C 3824	N 981	O 1126	S 26	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	ALA	-	expression tag	UNP P27487
A	28	ASP	-	expression tag	UNP P27487
A	29	PRO	-	expression tag	UNP P27487
A	30	GLY	-	expression tag	UNP P27487
A	31	GLY	-	expression tag	UNP P27487
A	32	SER	-	expression tag	UNP P27487
A	33	HIS	-	expression tag	UNP P27487
A	34	HIS	-	expression tag	UNP P27487
A	35	HIS	-	expression tag	UNP P27487
A	36	HIS	-	expression tag	UNP P27487
A	37	HIS	-	expression tag	UNP P27487
A	38	HIS	-	expression tag	UNP P27487
B	27	ALA	-	expression tag	UNP P27487
B	28	ASP	-	expression tag	UNP P27487
B	29	PRO	-	expression tag	UNP P27487
B	30	GLY	-	expression tag	UNP P27487
B	31	GLY	-	expression tag	UNP P27487
B	32	SER	-	expression tag	UNP P27487
B	33	HIS	-	expression tag	UNP P27487
B	34	HIS	-	expression tag	UNP P27487
B	35	HIS	-	expression tag	UNP P27487

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	36	HIS	-	expression tag	UNP P27487
B	37	HIS	-	expression tag	UNP P27487
B	38	HIS	-	expression tag	UNP P27487
C	27	ALA	-	expression tag	UNP P27487
C	28	ASP	-	expression tag	UNP P27487
C	29	PRO	-	expression tag	UNP P27487
C	30	GLY	-	expression tag	UNP P27487
C	31	GLY	-	expression tag	UNP P27487
C	32	SER	-	expression tag	UNP P27487
C	33	HIS	-	expression tag	UNP P27487
C	34	HIS	-	expression tag	UNP P27487
C	35	HIS	-	expression tag	UNP P27487
C	36	HIS	-	expression tag	UNP P27487
C	37	HIS	-	expression tag	UNP P27487
C	38	HIS	-	expression tag	UNP P27487
D	27	ALA	-	expression tag	UNP P27487
D	28	ASP	-	expression tag	UNP P27487
D	29	PRO	-	expression tag	UNP P27487
D	30	GLY	-	expression tag	UNP P27487
D	31	GLY	-	expression tag	UNP P27487
D	32	SER	-	expression tag	UNP P27487
D	33	HIS	-	expression tag	UNP P27487
D	34	HIS	-	expression tag	UNP P27487
D	35	HIS	-	expression tag	UNP P27487
D	36	HIS	-	expression tag	UNP P27487
D	37	HIS	-	expression tag	UNP P27487
D	38	HIS	-	expression tag	UNP P27487

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	2	28	16	2	10	0	0	0

- Molecule 3 is 2-{[3-(aminomethyl)-2-(2-methylpropyl)-1-oxo-4-phenyl-1,2-dihydroisoquinolin-6-yl]oxy}acetamide (three-letter code: LUI) (formula: $C_{22}H_{25}N_3O_3$).
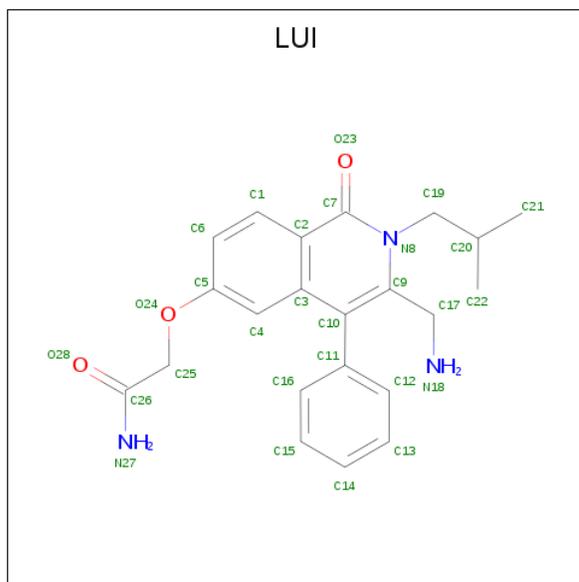
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			28	22	3	3		
3	B	1	Total	C	N	O	0	0
			28	22	3	3		
3	C	1	Total	C	N	O	0	0
			28	22	3	3		
3	D	1	Total	C	N	O	0	0
			28	22	3	3		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
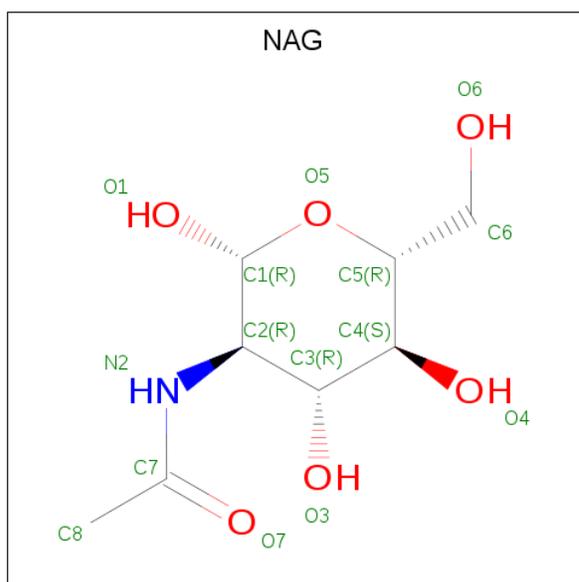
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	D	1	Total 14	C 8	N 1	O 5	0	0
4	D	1	Total 14	C 8	N 1	O 5	0	0
4	D	1	Total 14	C 8	N 1	O 5	0	0
4	D	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 5 is water.

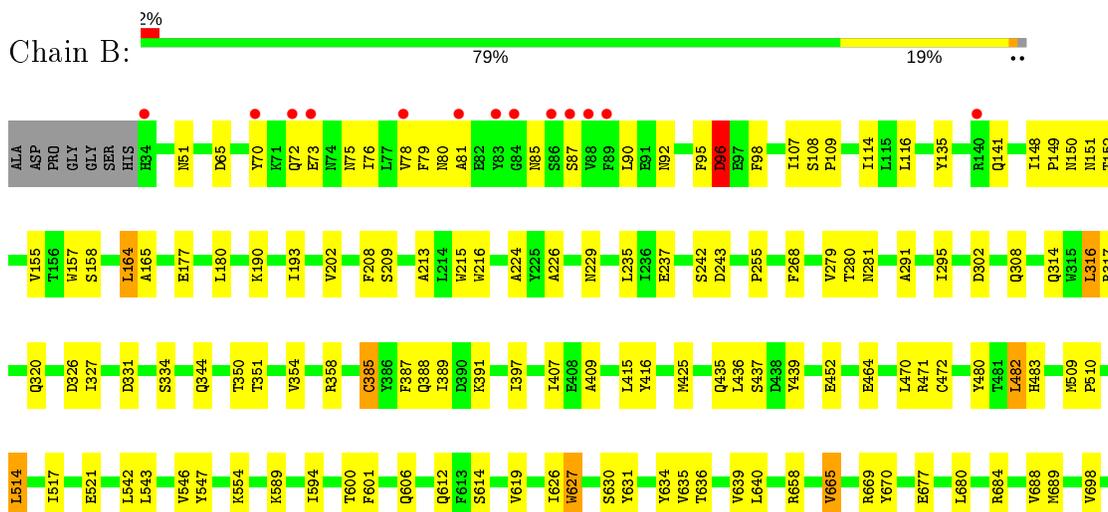| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	152	Total 152	O 152	0	0
5	B	142	Total 142	O 142	0	0
5	C	84	Total 84	O 84	0	0
5	D	128	Total 128	O 128	0	0

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 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.
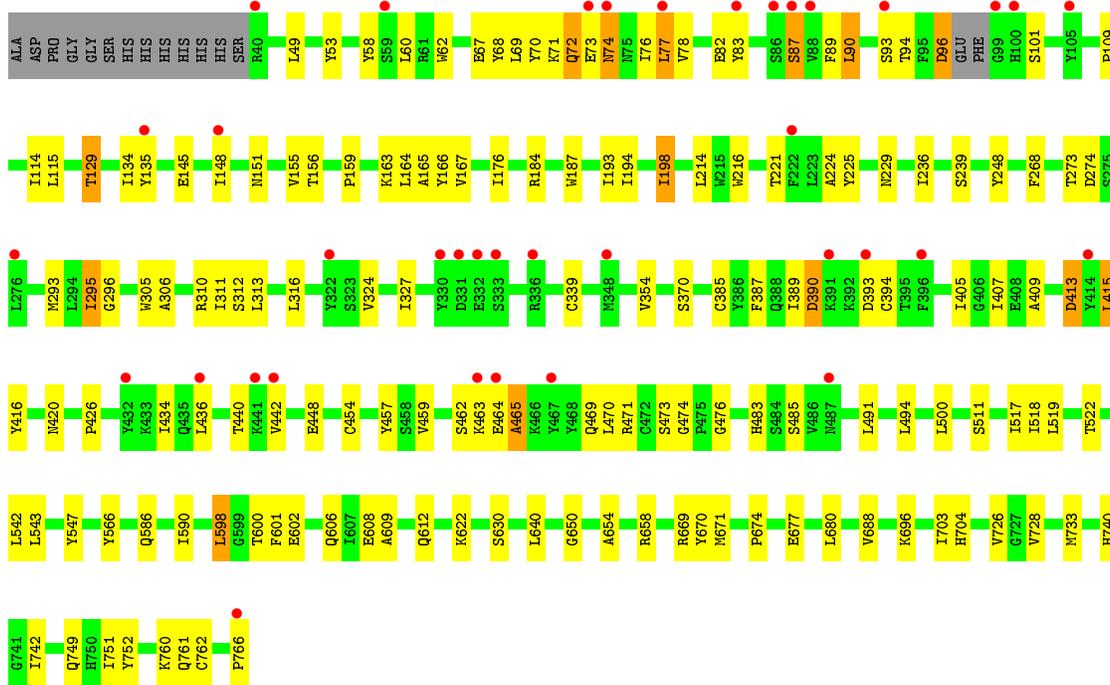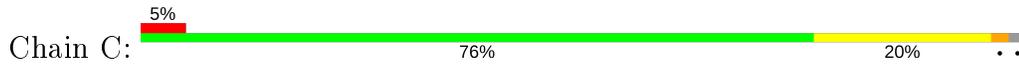
- Molecule 1: Dipeptidyl peptidase 4

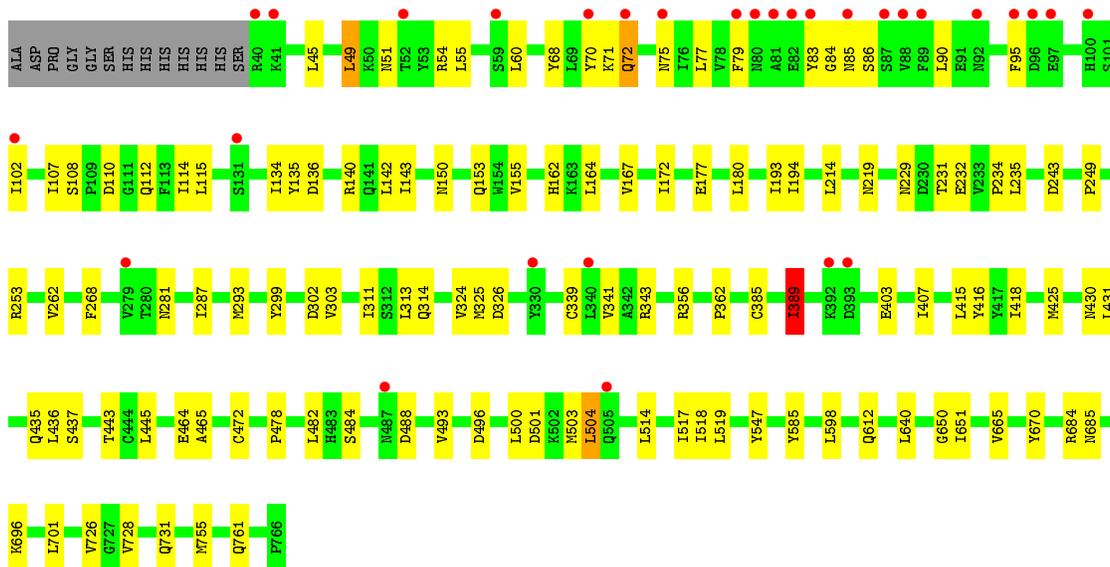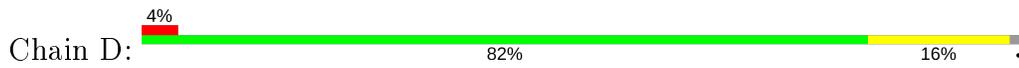
- Molecule 1: Dipeptidyl peptidase 4

H704 D708 L723 V728 H740 G741 I742 A743 S744 A747 I751 Y752 M755 L765 P766

● Molecule 1: Dipeptidyl peptidase 4

Chain C:

5%

76%   20%

ALA ASP PRO GLY SER HIS HIS HIS HIS HIS SER R40 L49 Y53 Y58 S59 L60 R61 W62 E67 Y68 L69 Y70 K71 Q72 E73 N74 N75 I76 L77 V78 E82 Y83 S86 S87 V88 F89 L90 S93 T94 F95 D96 GLU PHE G99 H100 S101 Y105 P109

I114 L115 T129 I134 Y135 E145 I148 N151 V155 T156 P159 K163 L164 A165 Y166 V167 I176 R184 W187 I193 I194 I198 L214 W215 W216 T221 F222 L223 A224 Y225 N229 I236 S239 Y248 F268 T273 D274 S275

L276 M293 L294 I295 G296 W305 A306 R310 I311 S312 L313 L316 Y322 S323 V324 I327 Y330 D331 E332 S333 R336 C339 M348 V354 S370 C385 Y386 F387 Q388 I389 D390 K391 K392 D393 C394 T395 F396 I405 G406 I407 E408 A409 D413 Y414 L415

Y416 N420 P426 Y432 K433 I434 Q435 L436 T440 K441 V442 E448 C454 Y457 S458 V459 S462 K463 E464 A465 K466 Y467 V468 Q469 L470 R471 C472 S473 G474 P475 G476 H483 S484 S485 V486 N487 L491 L494 L500 S511 I517 I518 L519 T522

L542 L543 Y547 Y566 Q586 I590 L598 G599 T600 F601 E602 Q606 I607 E608 A609 Q612 K622 S630 L640 G650 A654 R658 R669 Y670 M671 P674 E677 L680 V688 K696 I703 H704 V726 G727 V728 M733 H740

G741 I742 Q749 H750 I751 Y752 K760 Q761 C762 P766

● Molecule 1: Dipeptidyl peptidase 4

Chain D:

4%

82%   16%

ALA ASP PRO GLY SER HIS HIS HIS HIS HIS SER R40 K41 L45 L49 K50 N51 T52 Y53 R54 L55 S59 L60 Y68 L69 Y70 K71 Q72 N75 I76 L77 V78 F79 N80 A81 E82 Y83 G84 N85 S86 S87 V88 F89 L90 E91 N92 F95 D96 E97 H100 S101

I102 I107 S108 P109 D110 G111 Q112 F113 I114 L115 S131 I134 Y135 D136 R140 Q141 L142 I143 N150 Q153 W154 V155 H162 K163 L164 V167 I172 E177 L180 I193 I194 L214 N219 N229 N230 D230 T231 E232 V233 P234 L235 D243 P249

R253 V262 F268 V279 T280 N281 I287 M293 Y299 D302 V303 I311 S312 L313 Q314 V324 M325 D326 Y330 C339 L340 V341 A342 R343 R356 P362 C385 I389 K392 D393 E403 I407 L415 Y416 Y417 I418 M425 N430 L431

Q435 L436 S437 T443 C444 L445 E464 A465 C472 P478 L482 H483 S484 N487 D488 V493 D496 L500 D501 K502 M503 L504 Q505 L514 I517 I518 L519 Y547 Y585 L598 Q612 L640 G650 I651 V665 Y670 R684 N685

K696 L701 V726 G727 V728 Q731 M755 Q761 P766

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc opyranose

Chain E:

100%

MAG1
MAG2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	121.77Å 122.56Å 144.74Å 90.00° 115.00° 90.00°	Depositor
Resolution (Å)	35.00 – 2.72 34.68 – 2.72	Depositor EDS
% Data completeness (in resolution range)	97.5 (35.00-2.72) 97.5 (34.68-2.72)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.37 (at 2.72Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.195 , 0.254 0.199 , 0.259	Depositor DCC
R_{free} test set	5038 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	43.8	Xtriage
Anisotropy	0.419	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 35.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.013 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	24713	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% 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: LUI, NAG

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	0.52	0/6122	0.63	0/8327
1	B	0.50	0/6190	0.61	0/8419
1	C	0.46	0/6107	0.58	0/8305
1	D	0.49	0/6129	0.62	1/8336 (0.0%)
All	All	0.49	0/24548	0.61	1/33387 (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	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	356	ARG	NE-CZ-NH1	5.34	122.97	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	96	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	A	5950	0	5664	84	0
1	B	6013	0	5718	95	0
1	C	5937	0	5664	100	0
1	D	5957	0	5679	81	0
2	E	28	0	26	5	0
3	A	28	0	25	0	0
3	B	28	0	25	0	0
3	C	28	0	25	3	0
3	D	28	0	25	0	0
4	A	84	0	78	9	0
4	B	70	0	65	10	0
4	D	56	0	52	9	0
5	A	152	0	0	1	0
5	B	142	0	0	4	0
5	C	84	0	0	2	0
5	D	128	0	0	3	0
All	All	24713	0	23046	365	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 8.

The worst 5 of 365 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:B:150:ASN:HD21	4:B:1501:NAG:C1	1.09	1.57
1:C:229:ASN:HD21	2:E:1:NAG:C1	1.07	1.55
1:B:229:ASN:HD21	4:B:2291:NAG:C1	0.94	1.54
1:B:281:ASN:HD21	4:B:2811:NAG:C1	1.16	1.51
1:D:229:ASN:HD21	4:D:2291:NAG:C1	1.15	1.51

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	725/740 (98%)	677 (93%)	42 (6%)	6 (1%)	19	41
1	B	731/740 (99%)	691 (94%)	39 (5%)	1 (0%)	51	77
1	C	721/740 (97%)	673 (93%)	39 (5%)	9 (1%)	13	30
1	D	725/740 (98%)	674 (93%)	45 (6%)	6 (1%)	19	41
All	All	2902/2960 (98%)	2715 (94%)	165 (6%)	22 (1%)	19	41

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	GLN
1	A	73	GLU
1	C	90	LEU
1	C	93	SER
1	C	465	ALA

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	651/662 (98%)	625 (96%)	26 (4%)	31	58
1	B	658/662 (99%)	631 (96%)	27 (4%)	30	57
1	C	650/662 (98%)	618 (95%)	32 (5%)	25	50
1	D	652/662 (98%)	634 (97%)	18 (3%)	43	71
All	All	2611/2648 (99%)	2508 (96%)	103 (4%)	32	59

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	594	ILE
1	C	83	TYR
1	D	389	ILE
1	B	627	TRP
1	B	704	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	74	ASN
1	C	314	GLN
1	D	748	HIS
1	C	229	ASN
1	C	344	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)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

2 monosaccharides 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	NAG	E	1	1,2	14,14,15	0.46	0	17,19,21	1.34	2 (11%)
2	NAG	E	2	2	14,14,15	0.58	0	17,19,21	1.43	2 (11%)

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	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	2	NAG	O5-C1-C2	-3.22	106.21	111.29
2	E	1	NAG	C1-C2-N2	2.80	115.27	110.49
2	E	2	NAG	O7-C7-N2	2.27	126.13	121.95
2	E	1	NAG	O4-C4-C3	-2.21	105.25	110.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	1	NAG	5	0
2	E	2	NAG	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain E

Bond lengths	Bond angles
Torsions	Rings
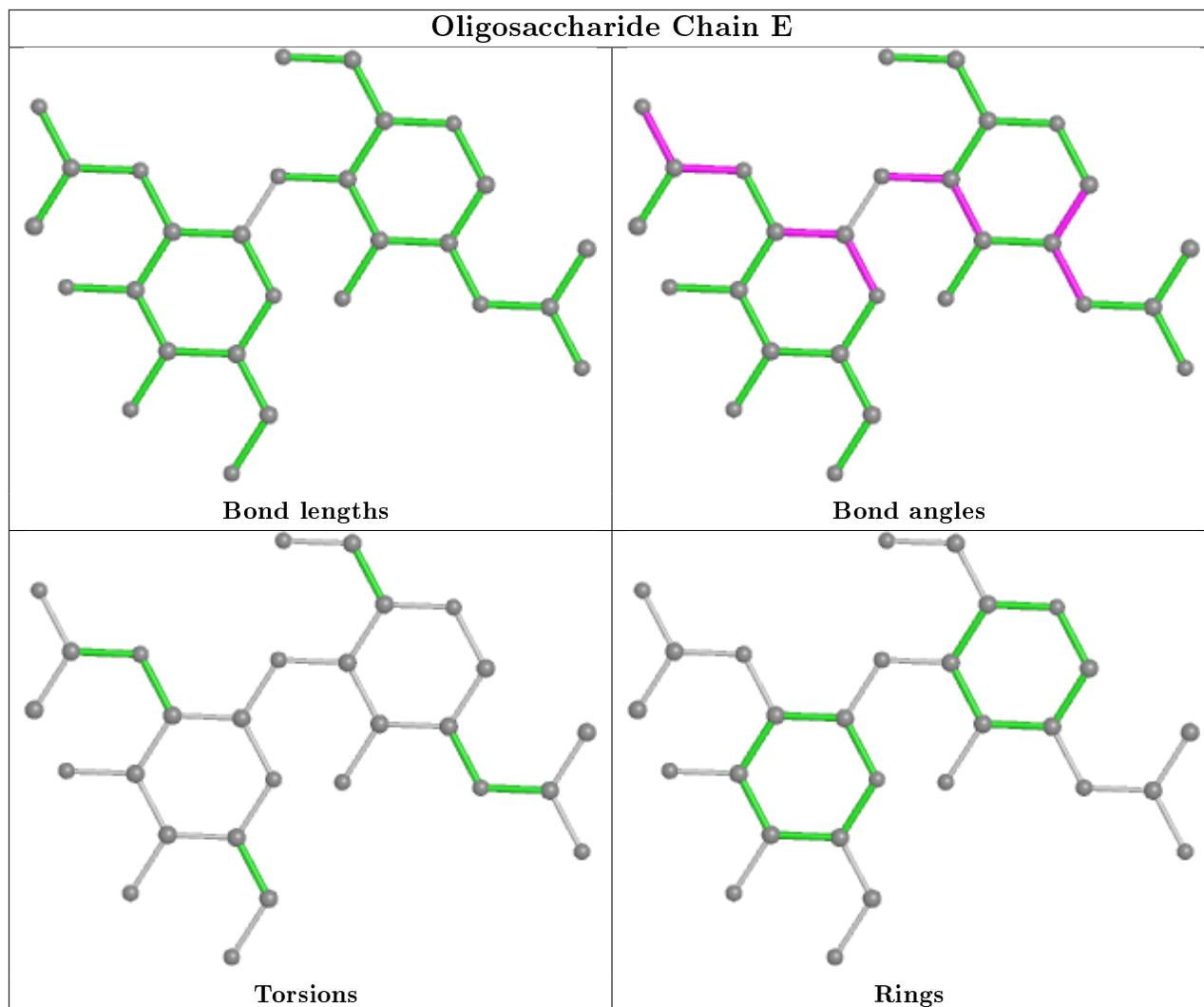
5.6 Ligand geometry (i)

19 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	#$|Z| > 2$
3	LUI	B	800	-	29,30,30	2.40	5 (17%)	32,42,42	2.05	3 (9%)
4	NAG	D	2191	1	14,14,15	0.61	0	17,19,21	1.03	1 (5%)
3	LUI	D	800	-	29,30,30	2.68	5 (17%)	32,42,42	2.09	4 (12%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	NAG	B	901	1	14,14,15	0.77	0	17,19,21	2.13	4 (23%)
4	NAG	A	1501	1	14,14,15	0.68	0	17,19,21	1.97	5 (29%)
4	NAG	D	2291	1	14,14,15	0.53	0	17,19,21	2.49	6 (35%)
4	NAG	A	3211	1	14,14,15	0.60	0	17,19,21	0.84	0
4	NAG	A	2291	1	14,14,15	0.67	0	17,19,21	1.32	1 (5%)
4	NAG	B	3211	1	14,14,15	0.72	0	17,19,21	1.33	1 (5%)
4	NAG	A	2191	1	14,14,15	0.67	0	17,19,21	0.93	1 (5%)
3	LUI	A	800	-	29,30,30	2.49	5 (17%)	32,42,42	1.94	3 (9%)
3	LUI	C	800	-	29,30,30	2.71	5 (17%)	32,42,42	1.90	3 (9%)
4	NAG	B	1501	1	14,14,15	0.59	0	17,19,21	1.31	2 (11%)
4	NAG	D	1501	1	14,14,15	0.85	0	17,19,21	1.64	4 (23%)
4	NAG	D	2811	1	14,14,15	0.86	0	17,19,21	1.92	4 (23%)
4	NAG	A	2811	1	14,14,15	0.88	0	17,19,21	1.37	5 (29%)
4	NAG	B	2811	1	14,14,15	0.52	0	17,19,21	0.91	0
4	NAG	B	2291	1	14,14,15	0.52	0	17,19,21	1.05	1 (5%)
4	NAG	A	2292	-	14,14,15	0.67	0	17,19,21	1.22	2 (11%)

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	LUI	B	800	-	-	2/13/15/15	0/3/3/3
4	NAG	D	2191	1	-	0/6/23/26	0/1/1/1
3	LUI	D	800	-	-	3/13/15/15	0/3/3/3
4	NAG	B	901	1	-	1/6/23/26	0/1/1/1
4	NAG	A	1501	1	-	3/6/23/26	0/1/1/1
4	NAG	D	2291	1	-	2/6/23/26	0/1/1/1
4	NAG	A	3211	1	-	0/6/23/26	0/1/1/1
4	NAG	A	2291	1	-	1/6/23/26	0/1/1/1
4	NAG	B	3211	1	-	2/6/23/26	0/1/1/1
4	NAG	A	2191	1	-	0/6/23/26	0/1/1/1
3	LUI	A	800	-	-	4/13/15/15	0/3/3/3
3	LUI	C	800	-	-	3/13/15/15	0/3/3/3
4	NAG	B	1501	1	-	2/6/23/26	0/1/1/1
4	NAG	D	1501	1	-	3/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	D	2811	1	-	0/6/23/26	0/1/1/1
4	NAG	A	2811	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2811	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2291	1	-	0/6/23/26	0/1/1/1
4	NAG	A	2292	-	-	2/6/23/26	0/1/1/1

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	800	LUI	C10-C9	12.92	1.50	1.36
3	D	800	LUI	C10-C9	12.70	1.49	1.36
3	A	800	LUI	C10-C9	11.49	1.48	1.36
3	B	800	LUI	C10-C9	11.03	1.48	1.36
3	D	800	LUI	C7-C2	4.84	1.50	1.41

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	800	LUI	C2-C7-N8	9.47	122.62	116.15
3	D	800	LUI	C2-C7-N8	9.45	122.60	116.15
3	A	800	LUI	C2-C7-N8	9.12	122.38	116.15
3	C	800	LUI	C2-C7-N8	8.87	122.20	116.15
4	D	2291	NAG	C1-O5-C5	7.10	121.82	112.19

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1501	NAG	C3-C2-N2-C7
3	C	800	LUI	O24-C25-C26-N27
4	B	2811	NAG	O5-C5-C6-O6
4	B	1501	NAG	O5-C5-C6-O6
4	D	1501	NAG	O5-C5-C6-O6

There are no ring outliers.

13 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	2191	NAG	2	0
4	B	901	NAG	2	0

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	2291	NAG	3	0
4	A	3211	NAG	2	0
4	A	2291	NAG	4	0
3	C	800	LUI	3	0
4	B	1501	NAG	4	0
4	D	1501	NAG	1	0
4	D	2811	NAG	3	0
4	A	2811	NAG	3	0
4	B	2811	NAG	2	0
4	B	2291	NAG	2	0
4	A	2292	NAG	3	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 LUI B 800

Bond lengths

Bond angles

Torsions

Rings

Ligand LUI D 800

Bond lengths

Bond angles

Torsions

Rings

Ligand LUI A 800

Bond lengths

Bond angles

Torsions

Rings

Ligand LUI C 800

Bond lengths

Bond angles

Torsions

Rings
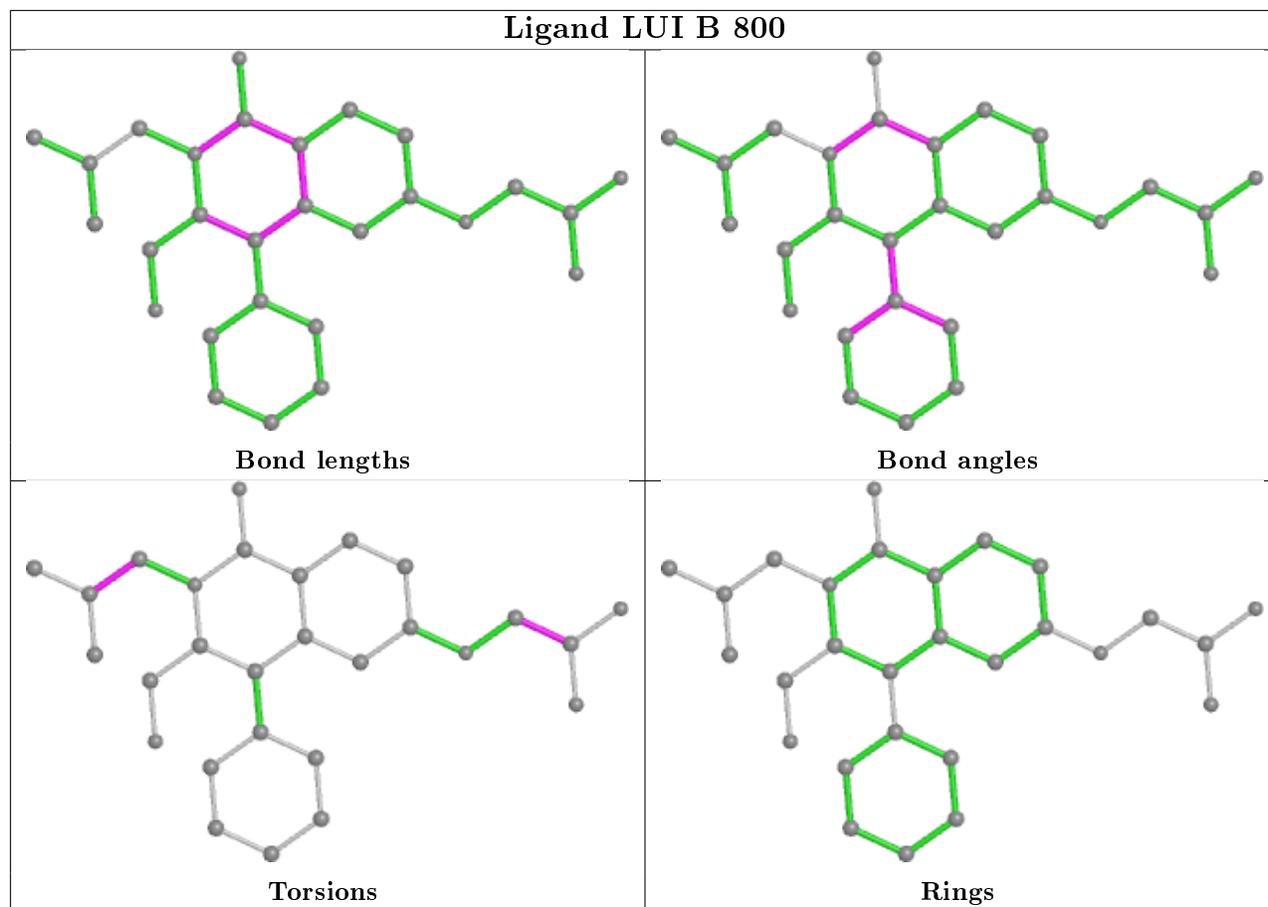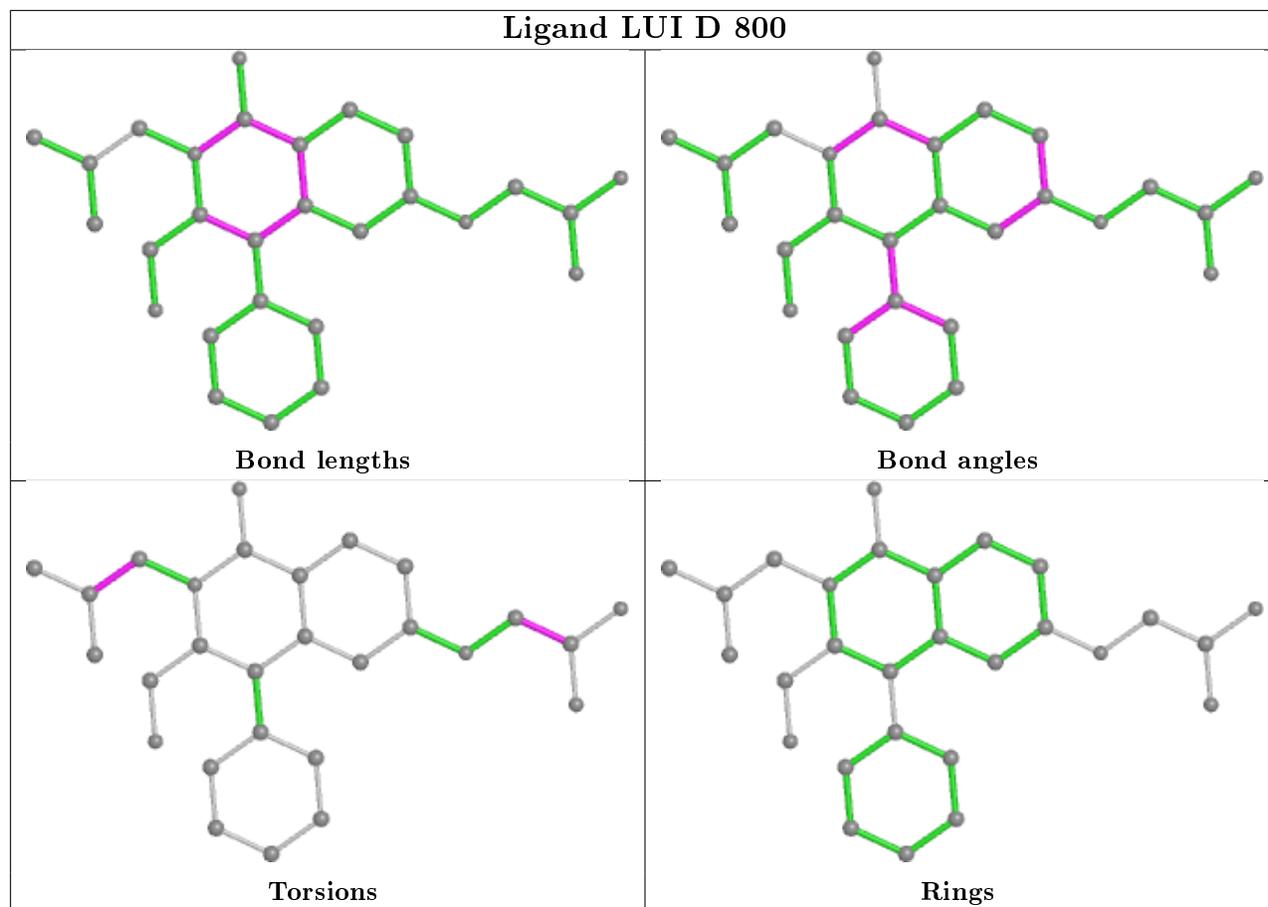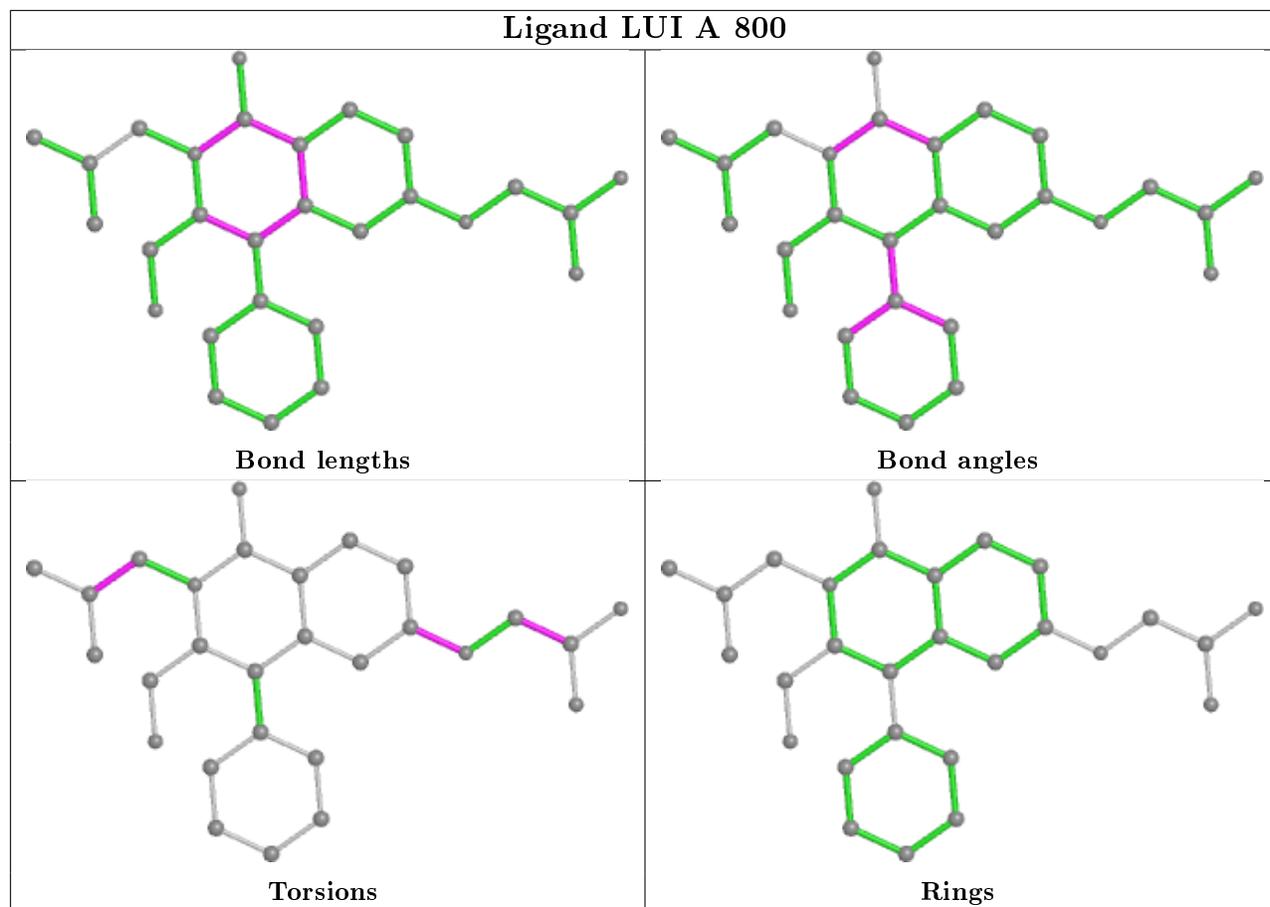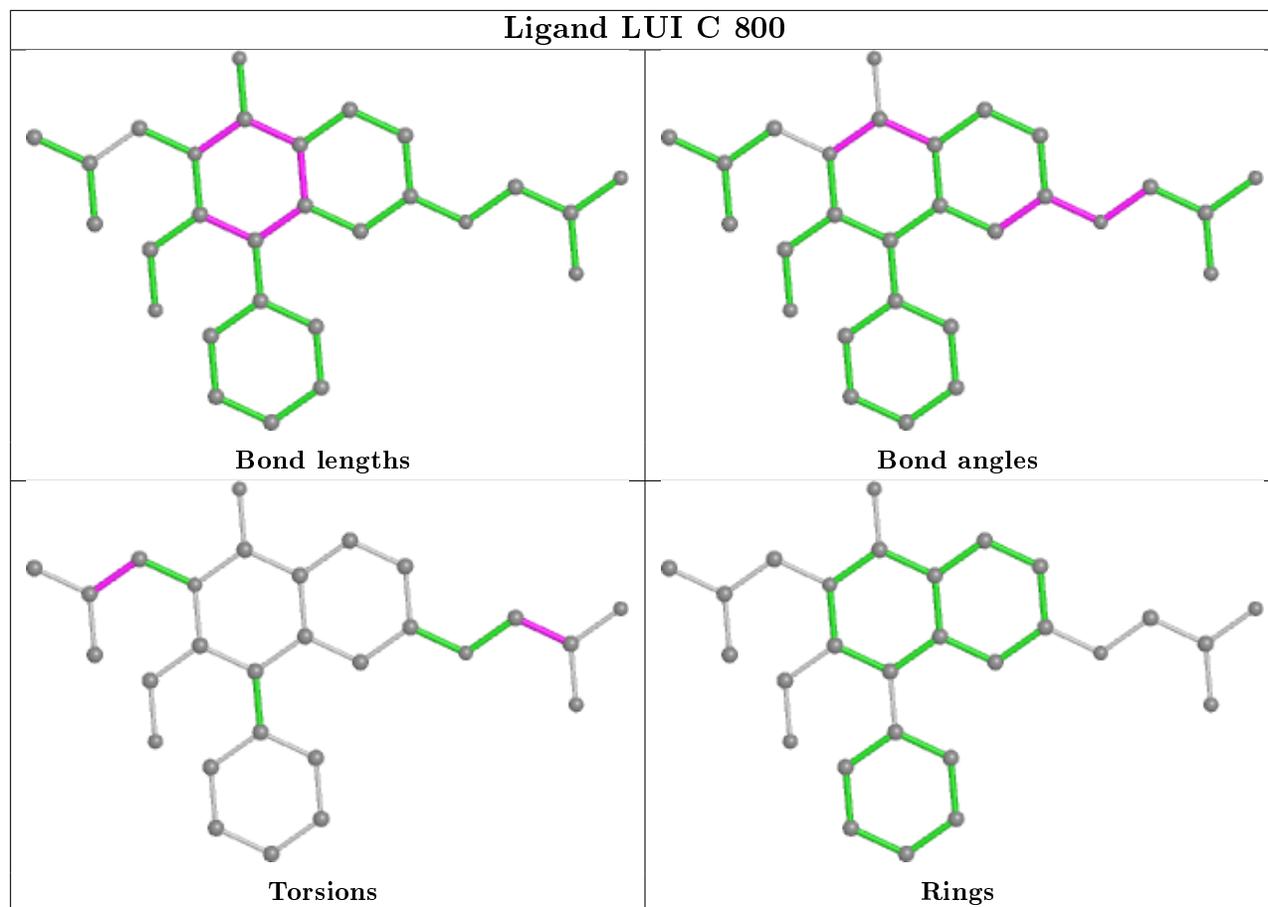
5.7 Other polymers (i)

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	A	727/740 (98%)	-0.07	23 (3%)	47	48	21, 42, 79, 100	0
1	B	733/740 (99%)	-0.19	13 (1%)	68	70	24, 46, 69, 85	0
1	C	725/740 (97%)	0.16	37 (5%)	28	27	25, 61, 101, 122	0
1	D	727/740 (98%)	-0.01	30 (4%)	37	36	24, 47, 91, 117	0
All	All	2912/2960 (98%)	-0.03	103 (3%)	44	44	21, 48, 91, 122	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	93	SER	4.6
1	D	92	ASN	4.6
1	C	73	GLU	4.5
1	D	81	ALA	4.4
1	D	97	GLU	4.4

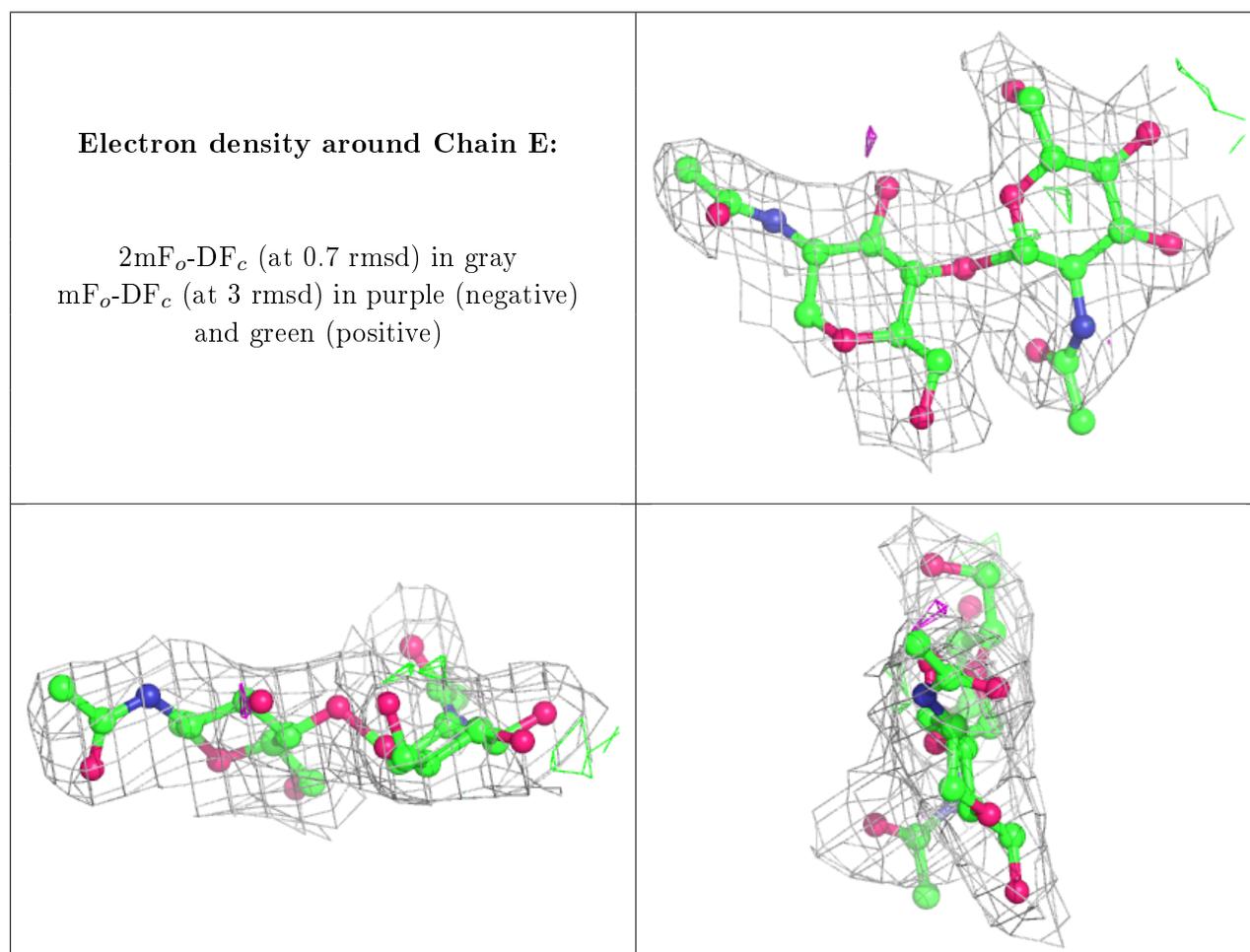## 6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

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.
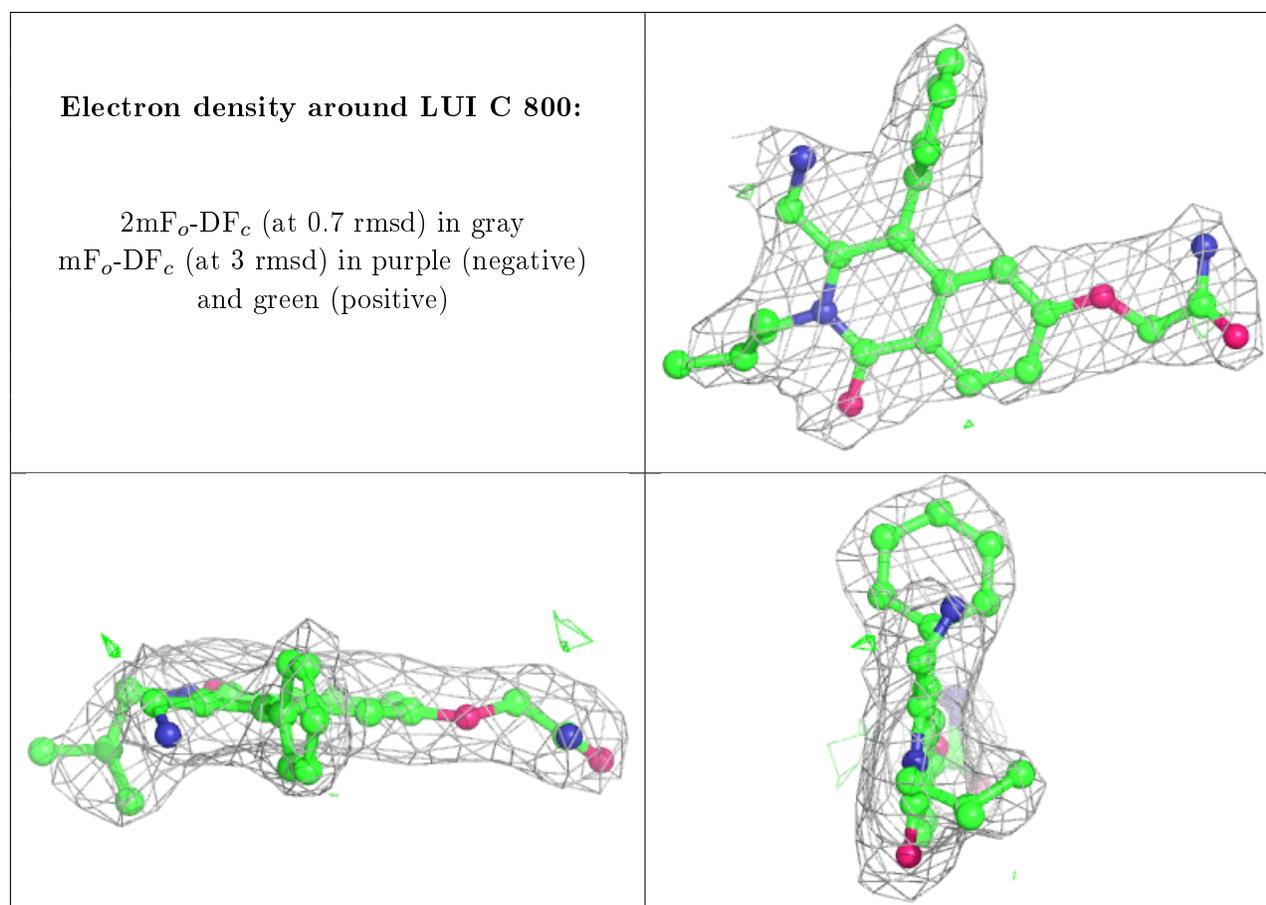
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	E	2	14/15	0.82	0.24	55,55,56,56	0
2	NAG	E	1	14/15	0.94	0.13	42,44,46,46	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain E:

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

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, 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(Å^2)	Q<0.9
4	NAG	B	901	14/15	0.75	0.15	53,54,55,55	0
4	NAG	A	2811	14/15	0.76	0.17	48,51,54,54	0
4	NAG	A	1501	14/15	0.80	0.32	53,55,57,59	0
4	NAG	A	2191	14/15	0.81	0.34	53,56,59,60	0
4	NAG	D	1501	14/15	0.82	0.28	51,53,54,54	0
4	NAG	D	2811	14/15	0.83	0.23	52,53,54,54	0
4	NAG	B	3211	14/15	0.84	0.37	55,57,57,58	0
4	NAG	A	2292	14/15	0.86	0.24	50,51,52,52	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
4	NAG	D	2191	14/15	0.87	0.27	51,53,57,58	0
4	NAG	B	2811	14/15	0.87	0.13	51,52,55,55	0
4	NAG	B	1501	14/15	0.89	0.27	50,51,52,52	0
4	NAG	D	2291	14/15	0.89	0.21	50,51,53,54	0
4	NAG	B	2291	14/15	0.90	0.15	45,48,49,49	0
4	NAG	A	2291	14/15	0.93	0.19	45,48,50,51	0
4	NAG	A	3211	14/15	0.93	0.26	49,52,54,55	0
3	LUI	C	800	28/28	0.95	0.16	39,41,46,47	0
3	LUI	A	800	28/28	0.95	0.21	36,40,47,50	0
3	LUI	D	800	28/28	0.96	0.15	35,38,42,43	0
3	LUI	B	800	28/28	0.97	0.17	37,39,40,40	0

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.
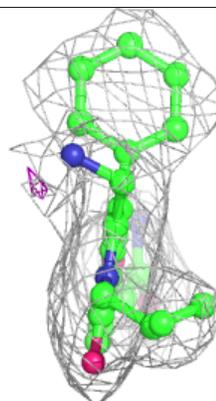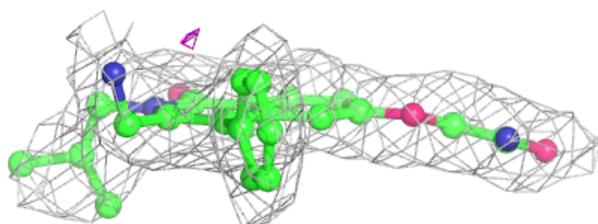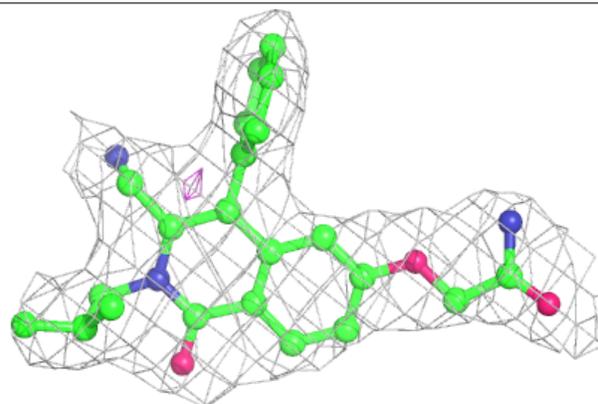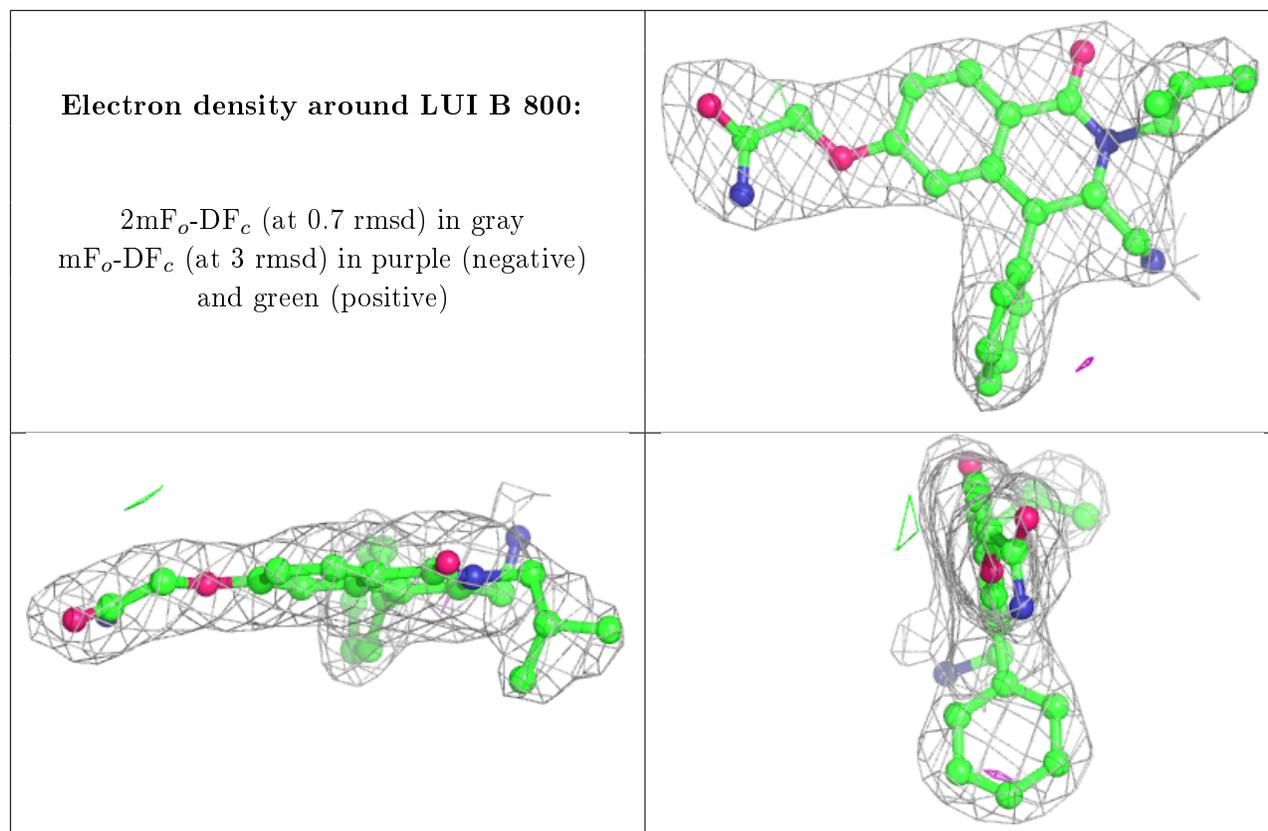

Electron density around LUI C 800:

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

Electron density around LUI A 800:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
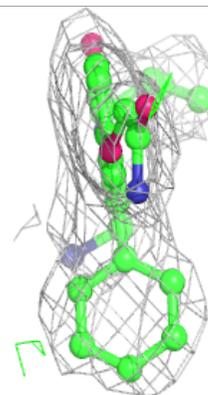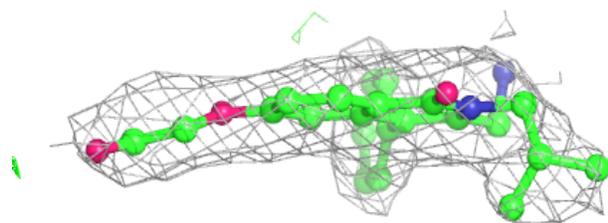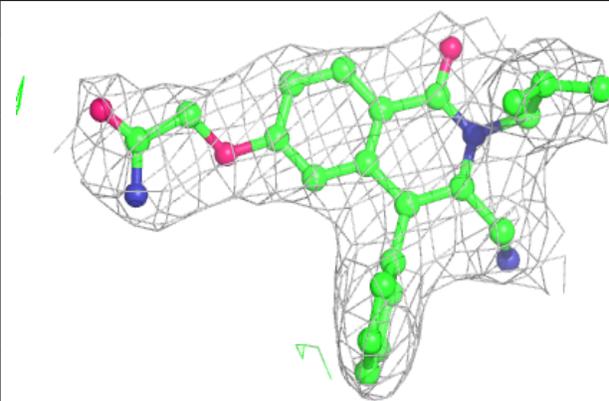
Electron density around LUI D 800:

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

Electron density around LUI B 800:

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

6.5 Other polymers (i)

There are no such residues in this entry.