WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 20, 2023 – 09:37 PM EDT

PDB ID	:	5KBY
Title	:	Crystal structure of dipeptidyl peptidase IV in complex with SYR-472
Authors	:	Skene, R.J.; Jennings, A.J.
Deposited on	:	2016-06-03
Resolution	:	2.24 Å(reported)

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

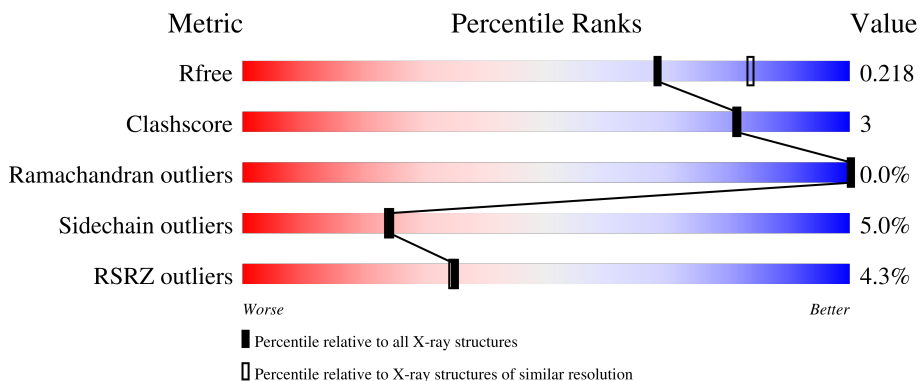|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

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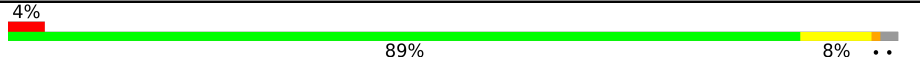*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.24 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.
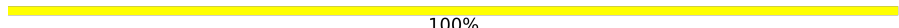
Metric	Percentile Ranks	Value
Rfree		0.218
Clashscore		3
Ramachandran outliers		0.0%
Sidechain outliers		5.0%
RSRZ outliers		4.3%

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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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% 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	4% / 89% / 8% ··
1	B	740	% / 89% / 9% ··
1	C	740	7% / 87% / 9% ··
1	D	740	4% / 87% / 10% ··
2	E	2	50% / 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	2	100%
2	G	2	100%
2	H	2	100%
2	I	2	100%
2	J	2	50% 50%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	E	2	-	-	-	X

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 25867 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	725	Total	C	N	O	S	0	1	0
			5936	3813	975	1122	26			
1	B	733	Total	C	N	O	S	0	3	0
			6024	3863	1000	1135	26			
1	C	717	Total	C	N	O	S	0	2	0
			5882	3784	963	1109	26			
1	D	725	Total	C	N	O	S	0	1	0
			5946	3819	980	1121	26			

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
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

Continued from previous page...

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

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 14	C 8	N 1	O 5	0	0
3	A	1	Total 14	C 8	N 1	O 5	0	0
3	A	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total 14	C 8	N 1	O 5	0	0
3	C	1	Total 14	C 8	N 1	O 5	0	0
3	C	1	Total 14	C 8	N 1	O 5	0	0
3	D	1	Total 14	C 8	N 1	O 5	0	0
3	D	1	Total 14	C 8	N 1	O 5	0	0
3	D	1	Total 14	C 8	N 1	O 5	0	0
3	D	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 4 is 2-[[6-[(3 {R})-3-azanylpiperidin-1-yl]-3-methyl-2,4-bis(oxidanylidene)pyrimidin-1-yl]methyl]-4-fluoranyl-benzenecarboni trile (three-letter code: 6RL) (formula: $C_{18}H_{20}FN_5O_2$).

6RL
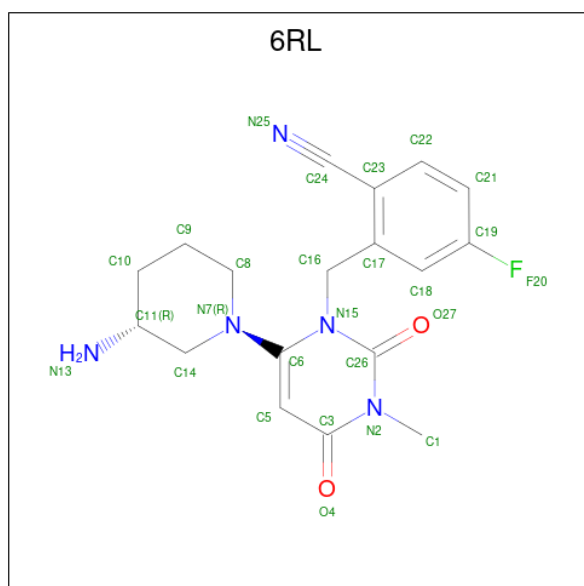
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 26	C 18	F 1	N 5	O 2	0	0
4	B	1	Total 26	C 18	F 1	N 5	O 2	0	0
4	C	1	Total 26	C 18	F 1	N 5	O 2	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	D	1	Total 26	C 18	F 1	N 5	O 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	422	Total 422	O 422	0	0
5	B	497	Total 497	O 497	0	0
5	C	300	Total 300	O 300	0	0
5	D	378	Total 378	O 378	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.
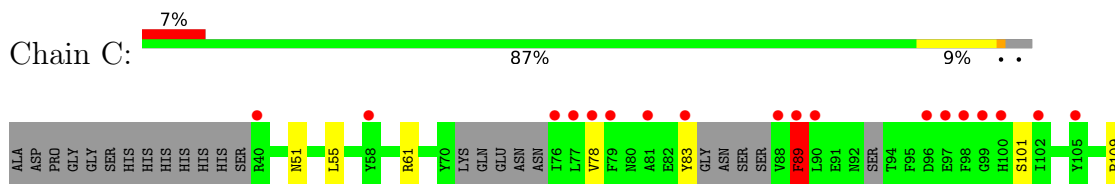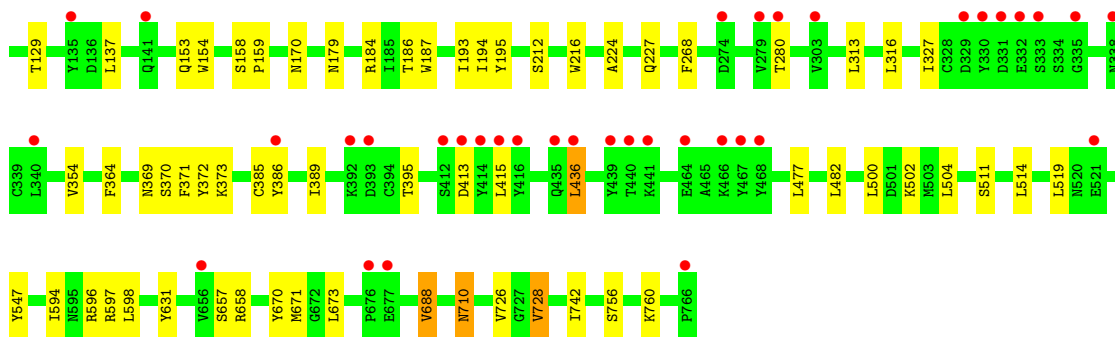
• Molecule 1: Dipeptidyl peptidase 4
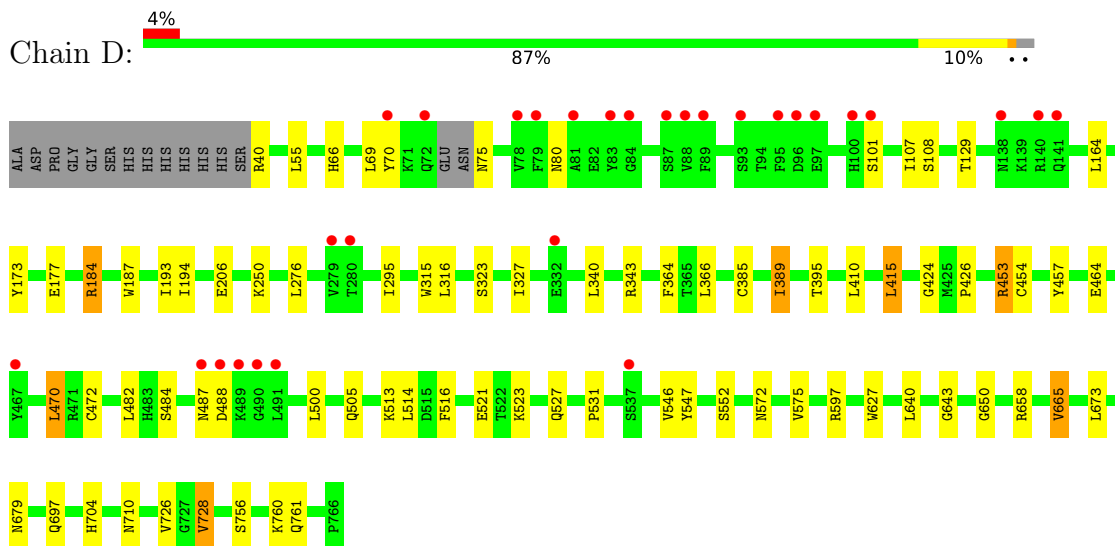
Chain A:

• Molecule 1: Dipeptidyl peptidase 4

Chain B:

• Molecule 1: Dipeptidyl peptidase 4

Chain C:

T129 Y135 D136 L137 Q141 Q153 W154 S158 P159 N170 N179 R184 I185 T186 W187 I193 I194 Y195 S212 W216 A224 Q227 F268 D274 V279 T280 V303 L313 L316 I327 C328 D329 Y330 D331 E332 S333 S334 G335 N338

C339 L340 V354 F364 N369 S370 F371 Y372 K373 C385 Y386 I389 K392 D393 C394 T395 S412 D413 Y414 L415 Y416 Q435 L436 Y439 T440 K441 E464 A465 K466 Y467 Y468 L477 L482 L500 D501 K502 M503 L504 S511 L514 L519 N520 E521

Y547 I594 N595 R596 R597 L598 Y631 V656 S657 R658 Y670 M671 G672 L673 P676 E677 V688 N710 V726 G727 V728 I742 S756 K760 P766

• Molecule 1: Dipeptidyl peptidase 4

Chain D:

4%

87%   10%

ALA ASP PRO GLY SER HIS HIS HIS HIS HIS HIS SER R40 L55 H66 L69 Y70 K71 Q72 GLU ASN N75 V78 F79 N80 A81 E82 Y83 G84 S87 V88 F89 S93 T94 F95 D96 E97 H100 S101 I107 S108 T129 N138 K139 R140 Q141 L164

Y173 E177 R184 W187 I193 I194 E206 K250 L276 V279 T280 I295 W315 L316 S323 I327 E332 L340 R343 F364 T365 L366 C385 I389 T395 L410 L415 G424 M425 P426 R453 C454 Y457 E464

Y467 L470 R471 C472 L482 H483 S484 N487 D488 K489 G490 L491 L500 Q505 K513 L514 D515 F516 E521 T522 K523 Q527 P531 S537 V546 Y547 S552 N572 V575 R597 W627 L640 G643 G650 R658 V665 L673

N679 Q697 H704 N710 V726 G727 V728 S756 K760 Q761 P766

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

Chain E:

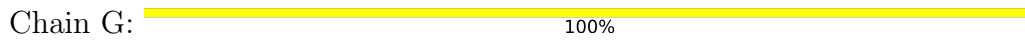50%   50%

NAG1 NAG2

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

Chain F: 100%

NAG1 NAG2

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

Chain G: 100%

NAG1 NAG2

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

Chain H:
100%

MAG1
MAG2

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

Chain I:
100%

MAG1
MAG2

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

Chain J:
50%   50%

MAG1
MAG2

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	121.57Å 122.17Å 143.70Å 90.00° 114.57° 90.00°	Depositor
Resolution (Å)	34.57 – 2.24 34.33 – 2.24	Depositor EDS
% Data completeness (in resolution range)	97.8 (34.57-2.24) 97.8 (34.33-2.24)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.78 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.7.0025	Depositor
R, R_{free}	0.174 , 0.215 0.180 , 0.218	Depositor DCC
R_{free} test set	8950 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	39.5	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 40.1	EDS
L-test for twinning[2]	$<\|L\|> = 0.50, <L^2> = 0.34$	Xtriage
Estimated twinning fraction	0.008 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	25867	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 3.19% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, 6RL

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	0.42	0/6111	0.62	1/8311 (0.0%)
1	B	0.43	0/6213	0.64	1/8449 (0.0%)
1	C	0.39	0/6059	0.59	0/8239
1	D	0.40	0/6120	0.61	1/8319 (0.0%)
All	All	0.41	0/24503	0.62	3/33318 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	453	ARG	NE-CZ-NH2	5.22	122.91	120.30
1	D	415	LEU	CA-CB-CG	5.20	127.26	115.30
1	A	453	ARG	NE-CZ-NH2	5.09	122.84	120.30

There are no chirality outliers.

There are no planarity outliers.

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	5936	0	5651	28	0
1	B	6024	0	5725	37	0
1	C	5882	0	5600	27	0
1	D	5946	0	5669	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	28	0	25	0	0
2	F	28	0	25	0	0
2	G	28	0	25	0	0
2	H	28	0	25	0	0
2	I	28	0	25	2	0
2	J	28	0	25	1	0
3	A	42	0	39	1	0
3	B	70	0	65	0	0
3	C	42	0	38	0	0
3	D	56	0	52	0	0
4	A	26	0	0	0	0
4	B	26	0	0	0	0
4	C	26	0	0	0	0
4	D	26	0	0	0	0
5	A	422	0	0	1	0
5	B	497	0	0	4	0
5	C	300	0	0	1	0
5	D	378	0	0	1	0
All	All	25867	0	22989	123	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 3.

The worst 5 of 123 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:153:GLN:HE22	1:A:170:ASN:H	1.30	0.79
1:C:153:GLN:HE22	1:C:170:ASN:H	1.31	0.77
1:C:369:ASN:O	1:C:389:ILE:HG12	1.86	0.75
1:D:552:SER:OG	5:D:1601:HOH:O	2.03	0.75
1:B:157:TRP:CE3	1:B:164:LEU:HD13	2.24	0.72

There are no symmetry-related clashes.

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	722/740 (98%)	691 (96%)	31 (4%)	0	100	100
1	B	734/740 (99%)	711 (97%)	23 (3%)	0	100	100
1	C	711/740 (96%)	682 (96%)	28 (4%)	1 (0%)	51	58
1	D	720/740 (97%)	691 (96%)	29 (4%)	0	100	100
All	All	2887/2960 (98%)	2775 (96%)	111 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	89	PHE

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	650/662 (98%)	620 (95%)	30 (5%)	27	28
1	B	661/662 (100%)	630 (95%)	31 (5%)	26	27
1	C	644/662 (97%)	611 (95%)	33 (5%)	24	23
1	D	650/662 (98%)	613 (94%)	37 (6%)	20	19
All	All	2605/2648 (98%)	2474 (95%)	131 (5%)	24	24

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	470	LEU
1	D	505	GLN
1	D	761	GLN
1	B	472	CYS
1	B	450	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35

such sidechains are listed below:

Mol	Chain	Res	Type
1	D	227	GLN
1	D	344	GLN
1	D	685	ASN
1	B	196	ASN
1	B	169	ASN

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 ⓘ

12 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	0.67	0	17,19,21	1.65	4 (23%)
2	NAG	E	2	2	14,14,15	0.49	0	17,19,21	0.97	0
2	NAG	F	1	1,2	14,14,15	0.67	0	17,19,21	1.01	1 (5%)
2	NAG	F	2	2	14,14,15	0.39	0	17,19,21	1.54	3 (17%)
2	NAG	G	1	1,2	14,14,15	0.48	0	17,19,21	1.05	2 (11%)
2	NAG	G	2	2	14,14,15	0.47	0	17,19,21	1.46	3 (17%)
2	NAG	H	1	1,2	14,14,15	0.59	0	17,19,21	1.24	2 (11%)
2	NAG	H	2	2	14,14,15	0.51	0	17,19,21	1.41	2 (11%)
2	NAG	I	1	1,2	14,14,15	0.70	0	17,19,21	1.24	3 (17%)
2	NAG	I	2	2	14,14,15	0.58	0	17,19,21	1.60	4 (23%)
2	NAG	J	1	1,2	14,14,15	0.72	0	17,19,21	1.46	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	NAG	J	2	2	14,14,15	0.47	0	17,19,21	1.40	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	-	3/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	4/6/23/26	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	3/6/23/26	0/1/1/1
2	NAG	H	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	H	2	2	-	4/6/23/26	0/1/1/1
2	NAG	I	1	1,2	-	1/6/23/26	0/1/1/1
2	NAG	I	2	2	-	2/6/23/26	0/1/1/1
2	NAG	J	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	J	2	2	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	1	NAG	O5-C1-C2	-4.22	104.63	111.29
2	I	2	NAG	C8-C7-N2	3.87	122.65	116.10
2	F	2	NAG	C8-C7-N2	3.81	122.55	116.10
2	H	2	NAG	C8-C7-N2	3.64	122.25	116.10
2	J	2	NAG	C8-C7-N2	3.41	121.86	116.10

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	1	NAG	C3-C2-N2-C7
2	J	2	NAG	O5-C5-C6-O6
2	J	2	NAG	C4-C5-C6-O6
2	G	2	NAG	O5-C5-C6-O6
2	F	2	NAG	C8-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

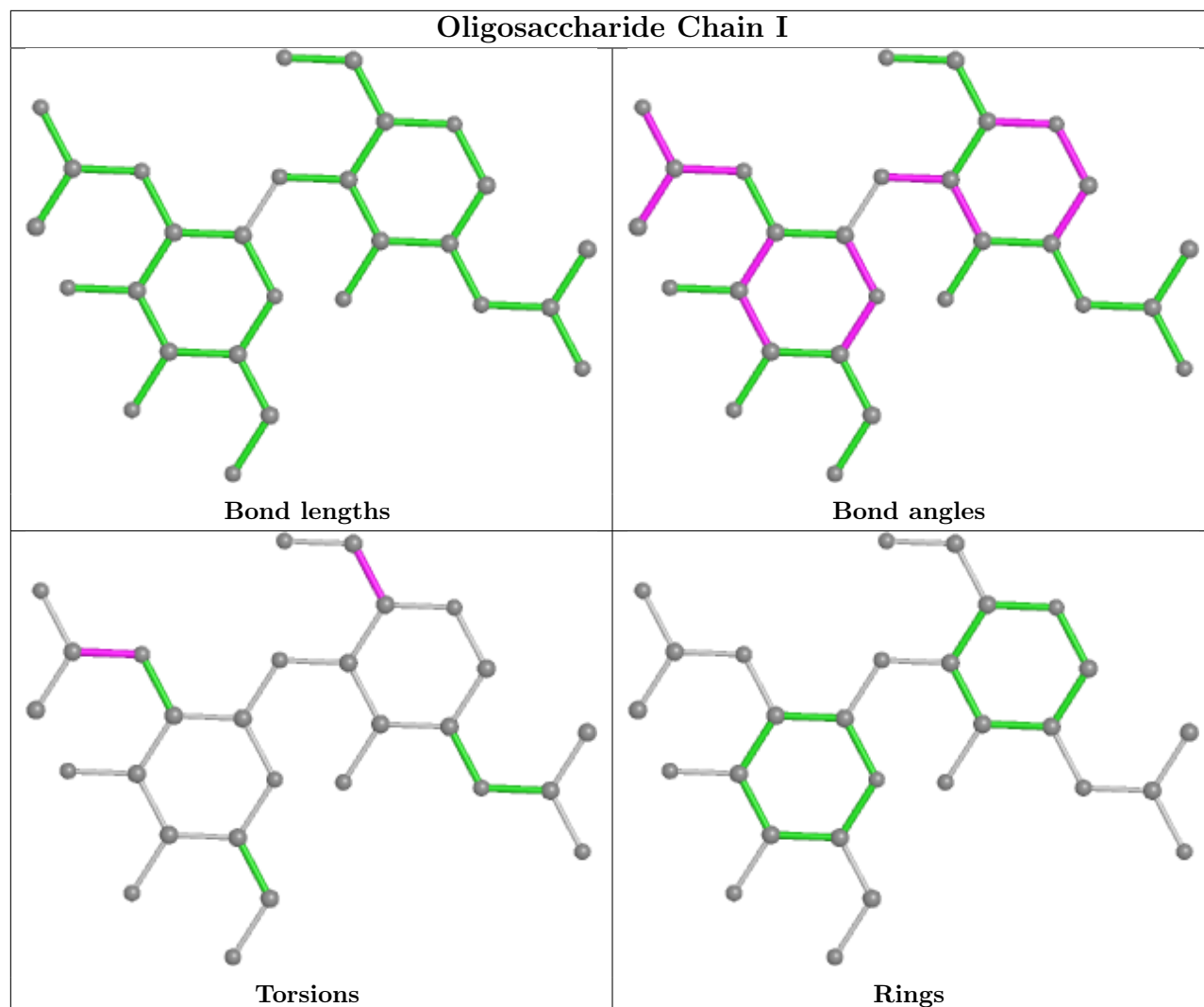| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	I	2	NAG	1	0
2	I	1	NAG	1	0
2	J	1	NAG	1	0

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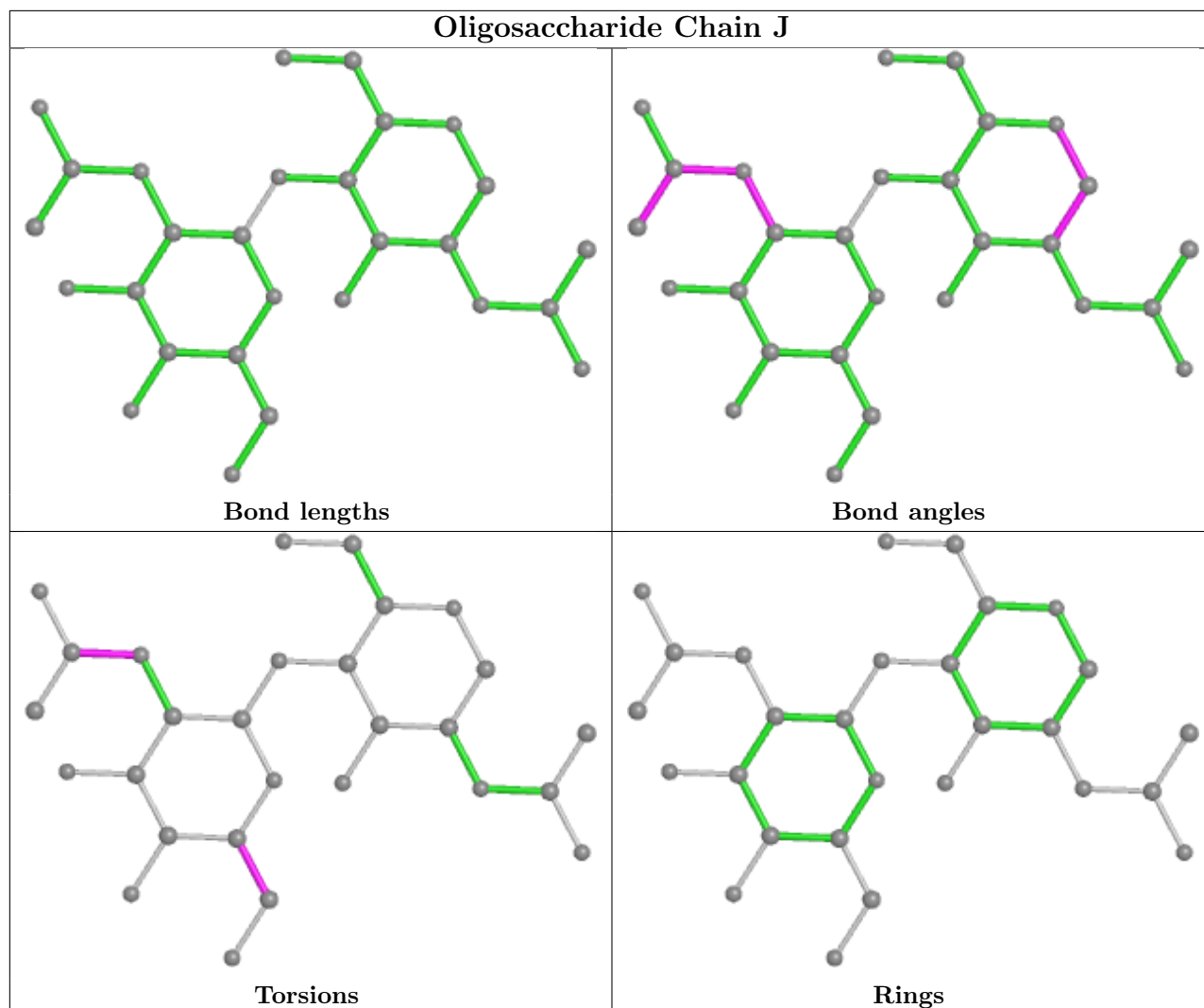

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain H

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain I

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain J

Bond lengths | Bond angles | Torsions | Rings

5.6 Ligand geometry ⓘ

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	NAG	C	1505	1	14,14,15	0.81	0	17,19,21	1.72	2 (11%)
3	NAG	A	1503	1	14,14,15	0.68	0	17,19,21	2.07	6 (35%)
3	NAG	D	1501	1	14,14,15	0.58	0	17,19,21	1.80	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
3	NAG	B	806	1	14,14,15	0.54	0	17,19,21	1.36	1 (5%)
3	NAG	A	1508	1	14,14,15	0.48	0	17,19,21	0.92	1 (5%)
3	NAG	A	1509	1	14,14,15	0.60	0	17,19,21	1.15	1 (5%)
4	6RL	B	808	-	28,28,28	0.93	1 (3%)	35,40,40	1.78	9 (25%)
4	6RL	C	1506	-	28,28,28	0.92	1 (3%)	35,40,40	1.80	6 (17%)
3	NAG	B	802	1	14,14,15	0.47	0	17,19,21	2.73	6 (35%)
3	NAG	B	801	1	14,14,15	0.52	0	17,19,21	2.03	4 (23%)
4	6RL	A	1510	-	28,28,28	0.94	0	35,40,40	1.80	9 (25%)
3	NAG	D	1502	1	14,14,15	0.64	0	17,19,21	1.07	1 (5%)
4	6RL	D	1507	-	28,28,28	0.89	1 (3%)	35,40,40	1.86	10 (28%)
3	NAG	B	805	1	14,14,15	0.36	0	17,19,21	1.47	2 (11%)
3	NAG	B	807	1	14,14,15	0.54	0	17,19,21	0.92	0
3	NAG	C	1501	1	14,14,15	0.56	0	17,19,21	1.66	3 (17%)
3	NAG	D	1506	1	14,14,15	0.63	0	17,19,21	1.34	2 (11%)
3	NAG	D	1505	1	14,14,15	0.56	0	17,19,21	0.98	1 (5%)
3	NAG	C	1504	1	14,14,15	0.49	0	17,19,21	1.31	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	NAG	C	1505	1	-	2/6/23/26	0/1/1/1
3	NAG	A	1503	1	-	0/6/23/26	0/1/1/1
3	NAG	D	1501	1	-	1/6/23/26	0/1/1/1
3	NAG	B	806	1	-	1/6/23/26	0/1/1/1
3	NAG	A	1508	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1509	1	-	2/6/23/26	0/1/1/1
4	6RL	B	808	-	-	4/10/20/20	0/3/3/3
4	6RL	C	1506	-	-	5/10/20/20	0/3/3/3
3	NAG	B	802	1	-	1/6/23/26	0/1/1/1
3	NAG	B	801	1	-	4/6/23/26	0/1/1/1
4	6RL	A	1510	-	-	3/10/20/20	0/3/3/3
3	NAG	D	1502	1	-	0/6/23/26	0/1/1/1
4	6RL	D	1507	-	-	3/10/20/20	0/3/3/3
3	NAG	B	805	1	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	B	807	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1501	1	-	4/6/23/26	0/1/1/1
3	NAG	D	1506	1	-	0/6/23/26	0/1/1/1
3	NAG	D	1505	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1504	1	-	2/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	808	6RL	C26-N15	-2.15	1.35	1.38
4	C	1506	6RL	C3-N2	-2.07	1.36	1.40
4	D	1507	6RL	C3-N2	-2.01	1.36	1.40

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	802	NAG	C1-O5-C5	8.43	123.62	112.19
3	B	801	NAG	C1-O5-C5	5.78	120.03	112.19
3	C	1505	NAG	C4-C3-C2	5.65	119.30	111.02
3	D	1501	NAG	C1-O5-C5	5.59	119.77	112.19
4	D	1507	6RL	C5-C3-N2	5.15	120.79	115.40

There are no chirality outliers.

5 of 34 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1509	NAG	C4-C5-C6-O6
3	C	1505	NAG	C1-C2-N2-C7
3	C	1504	NAG	O5-C5-C6-O6
3	A	1509	NAG	O5-C5-C6-O6
3	B	801	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1509	NAG	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 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 6RL B 808

Bond lengths

Bond angles

Torsions

Rings

Ligand 6RL C 1506

Bond lengths

Bond angles

Torsions

Rings

Ligand 6RL A 1510

Bond lengths

Bond angles

Torsions

Rings

Ligand 6RL D 1507

Bond lengths

Bond angles

Torsions

Rings
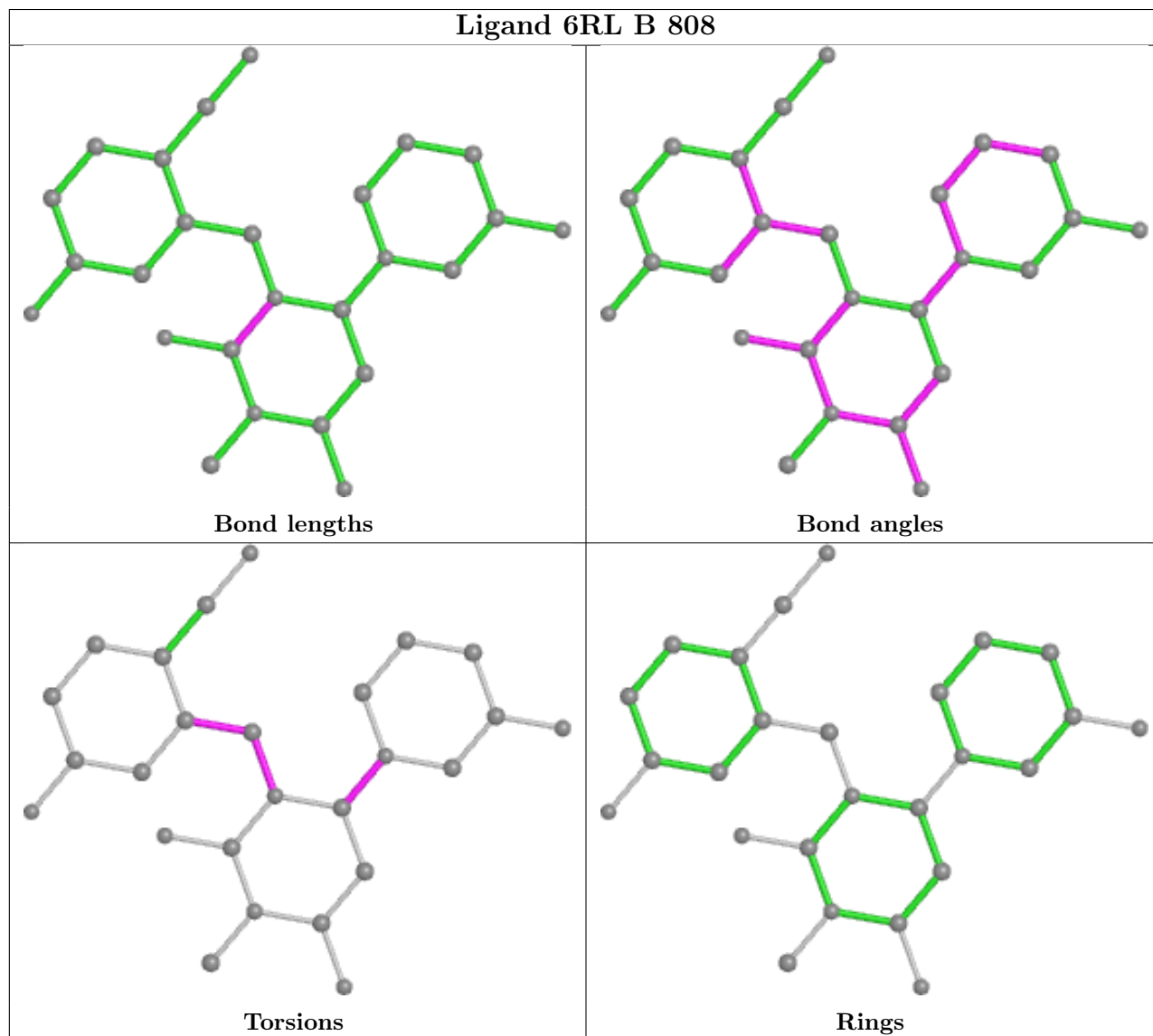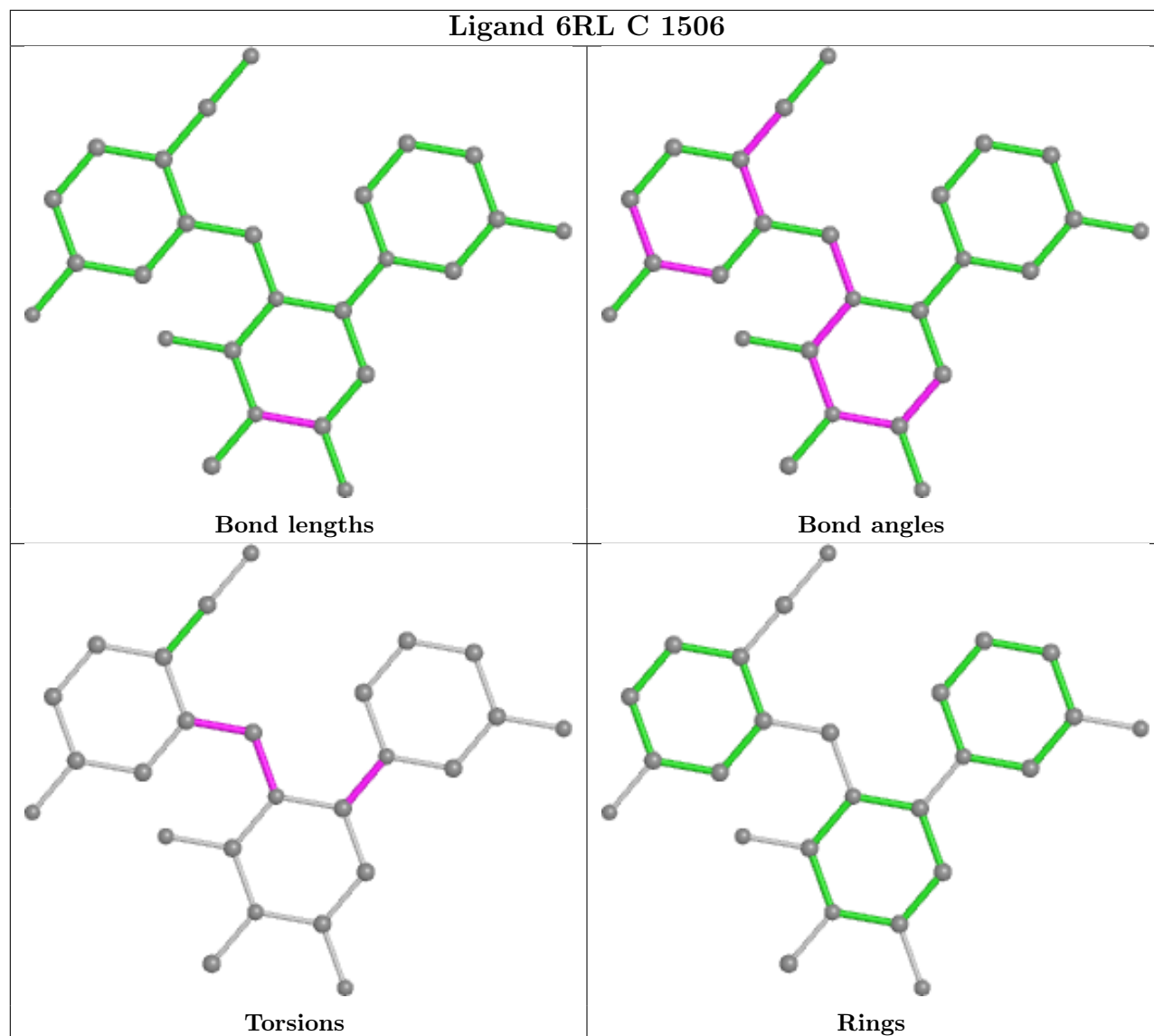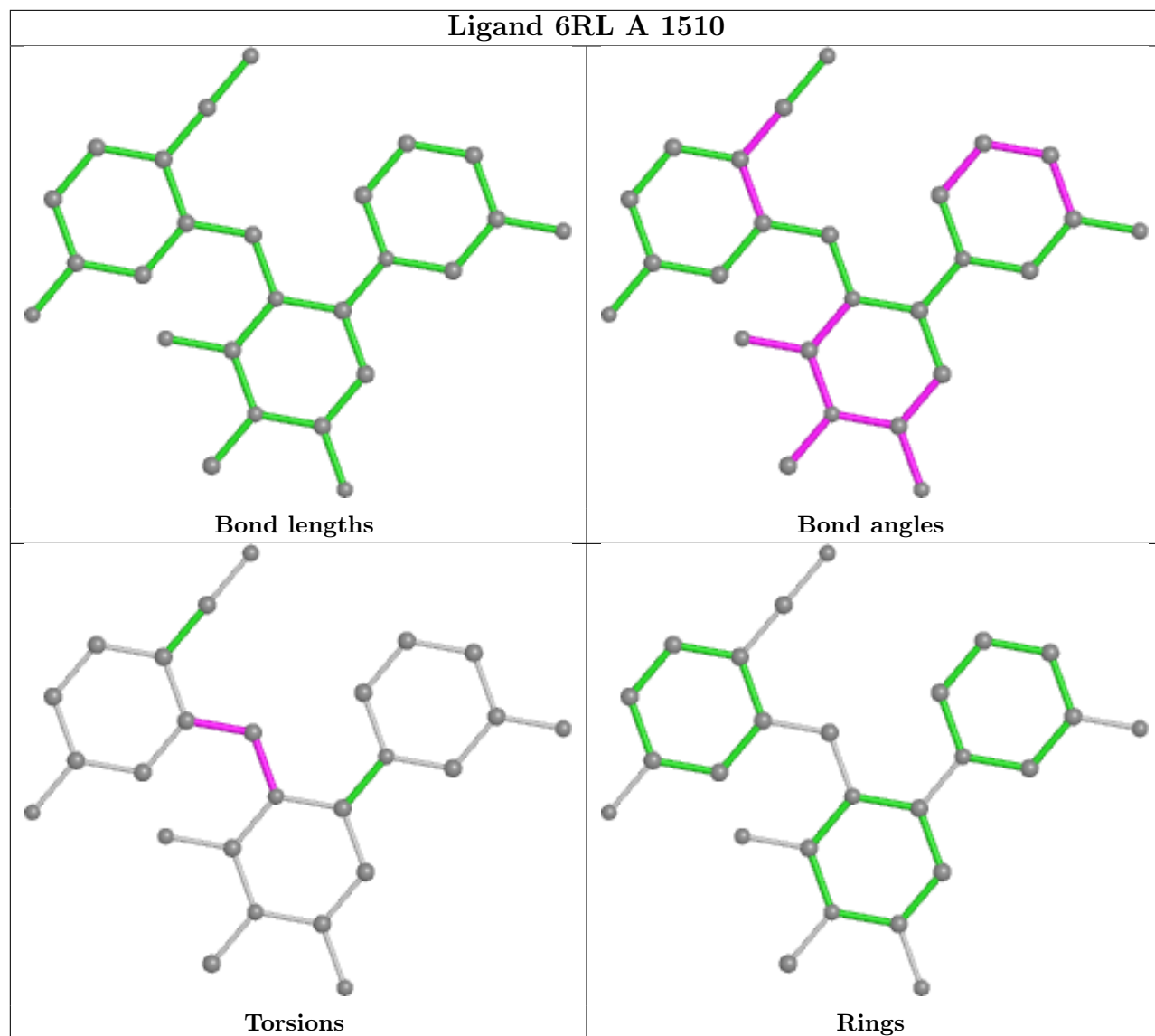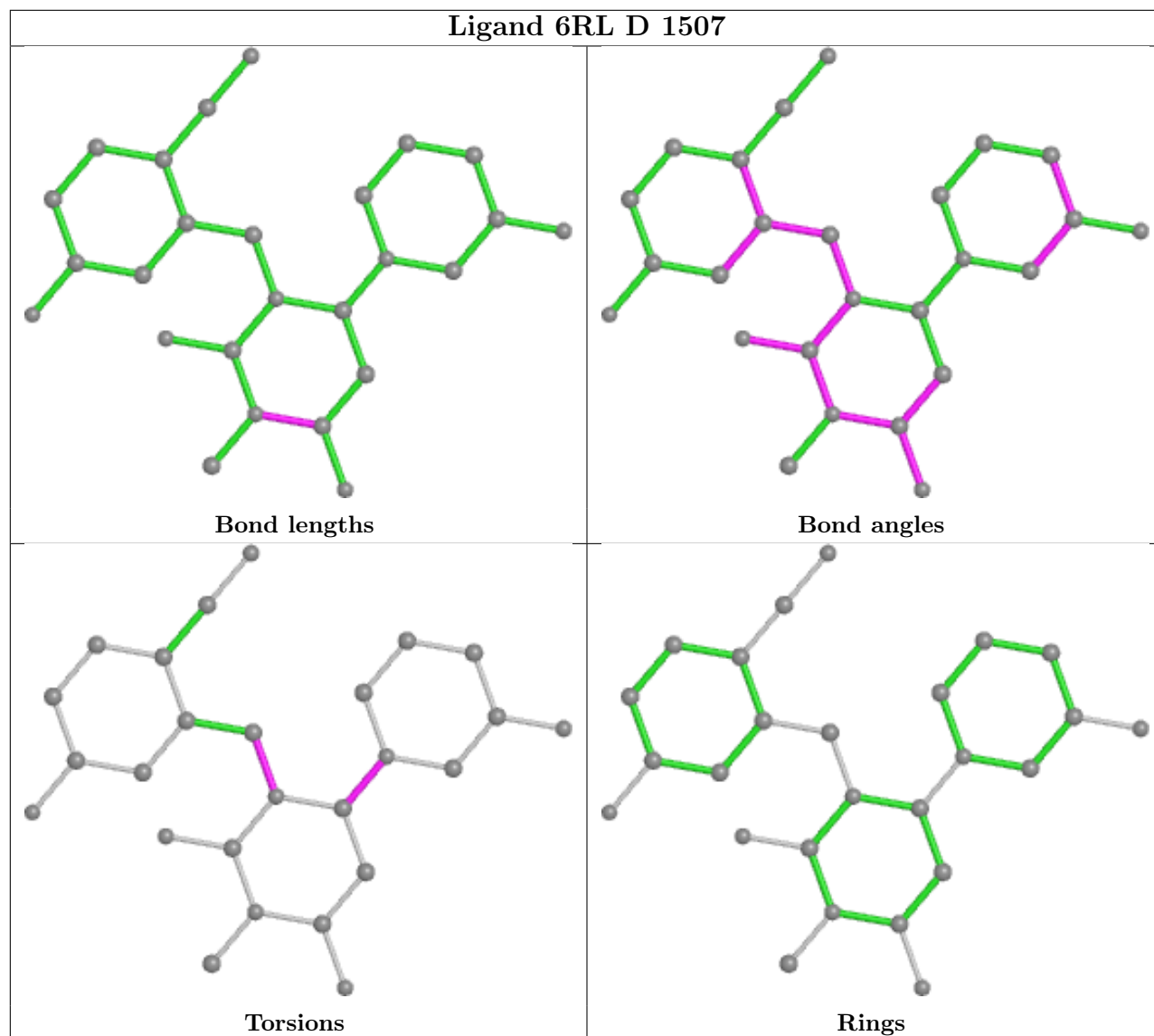
5.7 Other polymers ⓘ

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 ⓘ

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	725/740 (97%)	0.03	32 (4%)	34	34	23, 43, 88, 137	0
1	B	733/740 (99%)	-0.23	11 (1%)	73	74	23, 42, 74, 108	0
1	C	717/740 (96%)	0.20	54 (7%)	14	13	24, 55, 102, 135	0
1	D	725/740 (97%)	0.04	29 (4%)	38	38	22, 48, 94, 124	0
All	All	2900/2960 (97%)	0.01	126 (4%)	35	34	22, 46, 92, 137	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	97	GLU	7.4
1	D	81	ALA	6.3
1	C	77	LEU	6.0
1	A	81	ALA	5.9
1	C	335	GLY	5.6

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

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

6.3 Carbohydrates ⓘ

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	1	14/15	0.65	0.31	75,86,99,114	0
2	NAG	E	2	14/15	0.73	0.47	101,120,128,132	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	H	2	14/15	0.77	0.25	62,85,99,105	0
2	NAG	J	2	14/15	0.82	0.36	60,93,108,112	0
2	NAG	G	2	14/15	0.86	0.23	88,93,99,102	0
2	NAG	F	2	14/15	0.87	0.28	59,69,73,75	0
2	NAG	G	1	14/15	0.89	0.16	60,80,91,92	0
2	NAG	I	2	14/15	0.92	0.17	50,58,67,69	0
2	NAG	J	1	14/15	0.94	0.12	47,54,68,71	0
2	NAG	H	1	14/15	0.95	0.11	44,54,63,74	0
2	NAG	I	1	14/15	0.96	0.08	37,47,54,58	0
2	NAG	F	1	14/15	0.96	0.19	40,46,58,60	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)
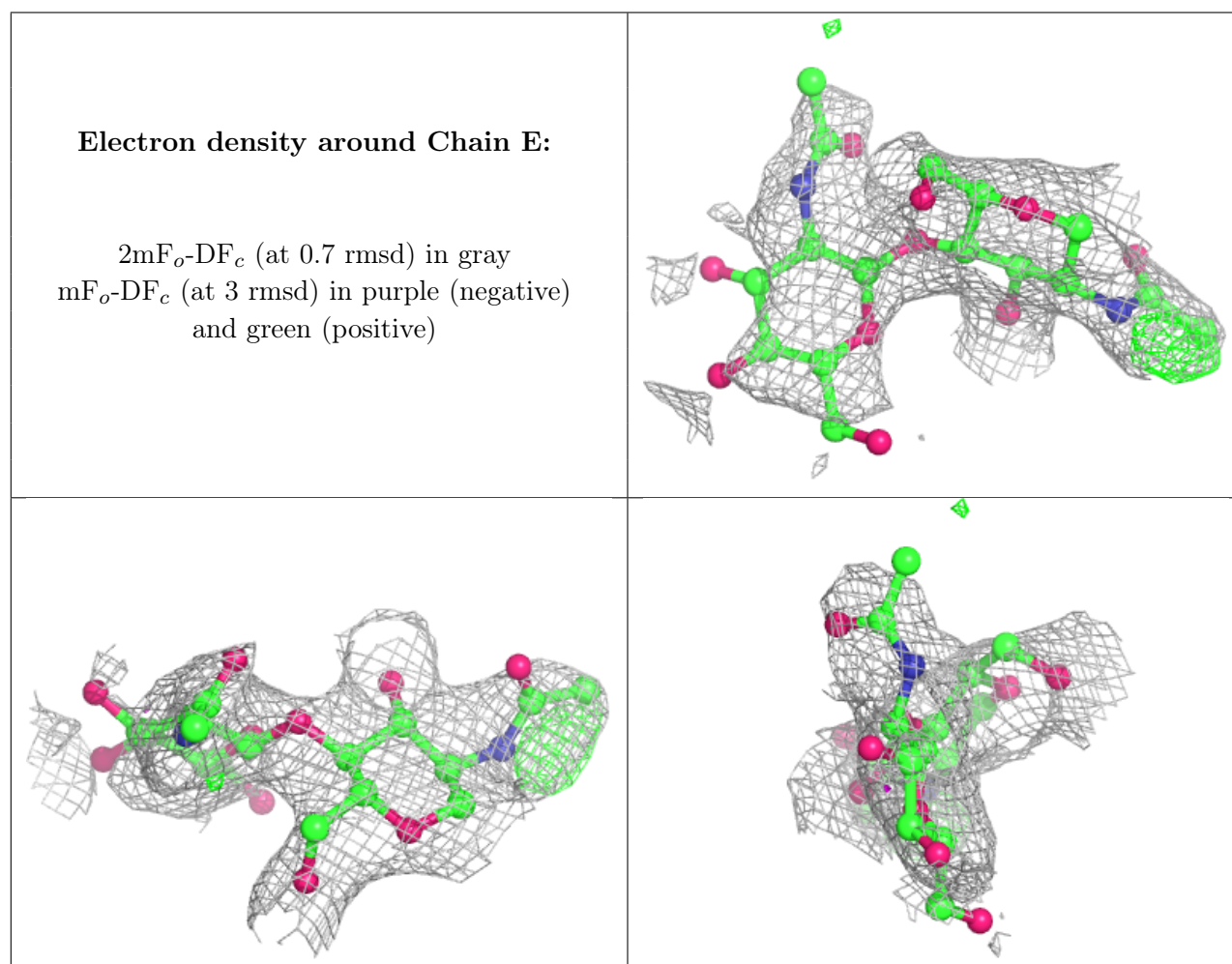
Electron density around Chain F:

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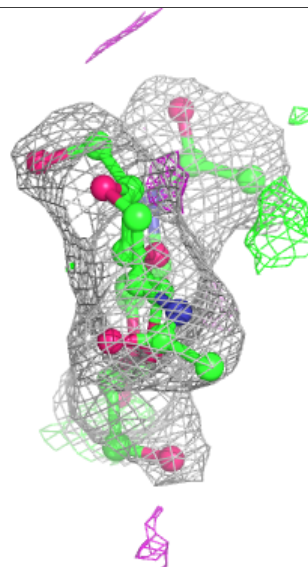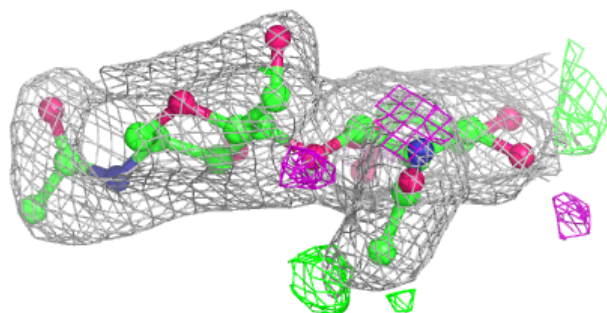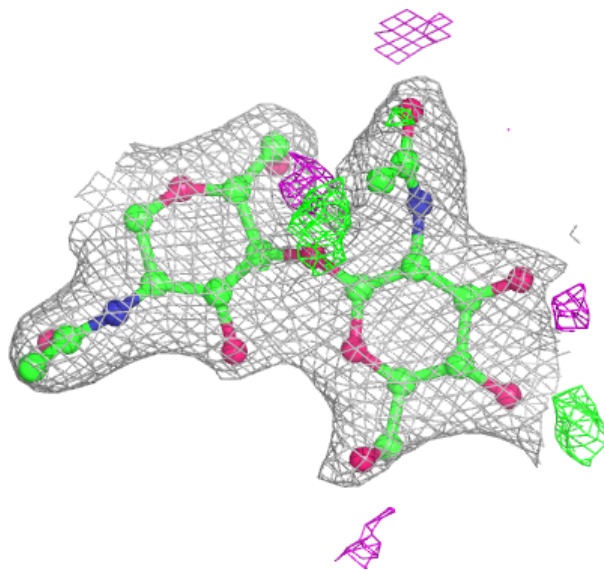
Electron density around Chain G:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
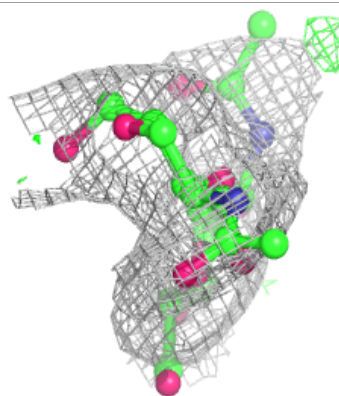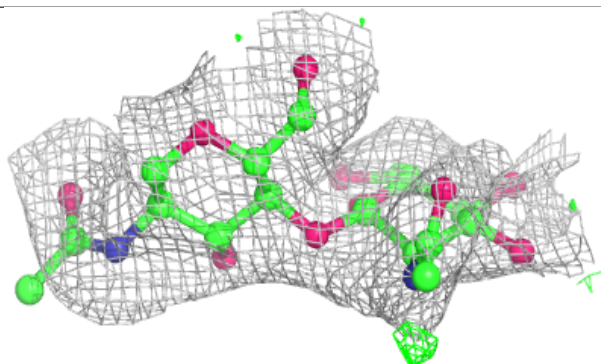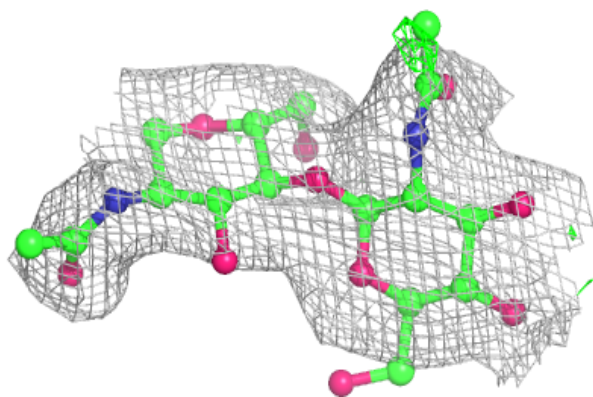
Electron density around Chain H:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
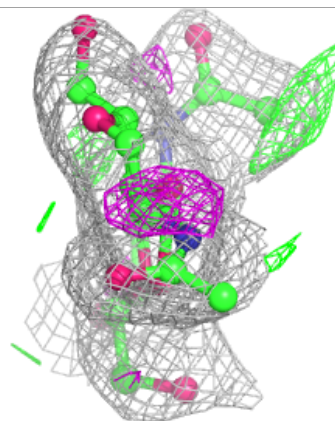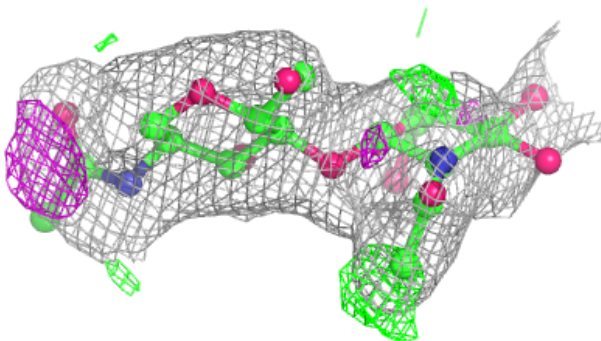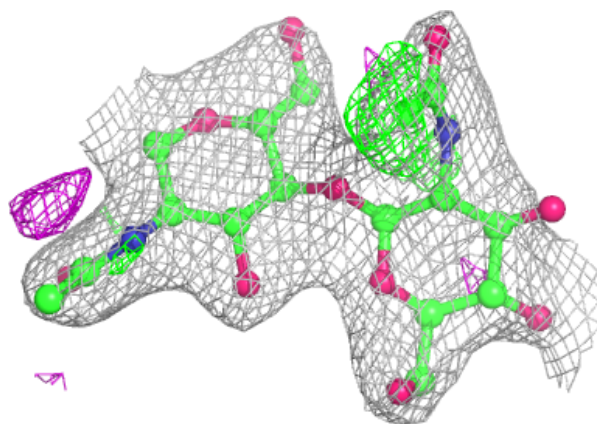
Electron density around Chain I:

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

Electron density around Chain J:

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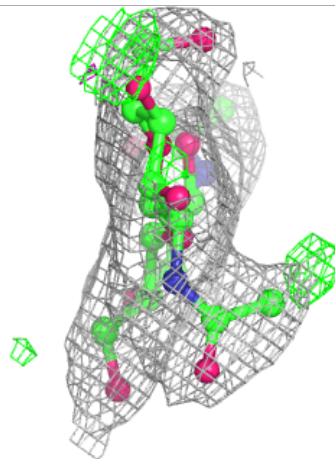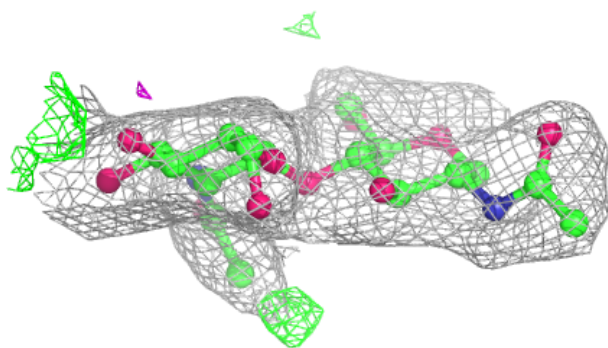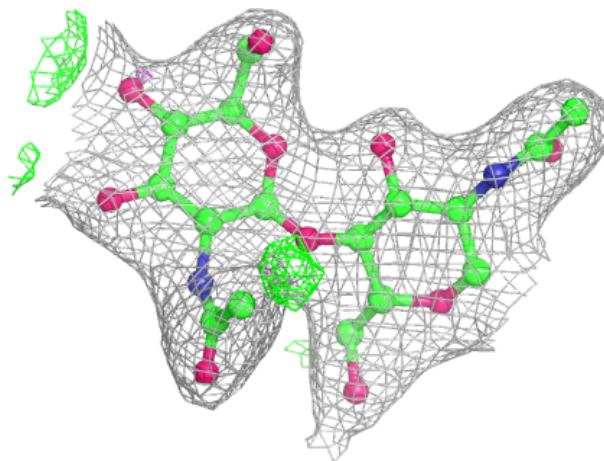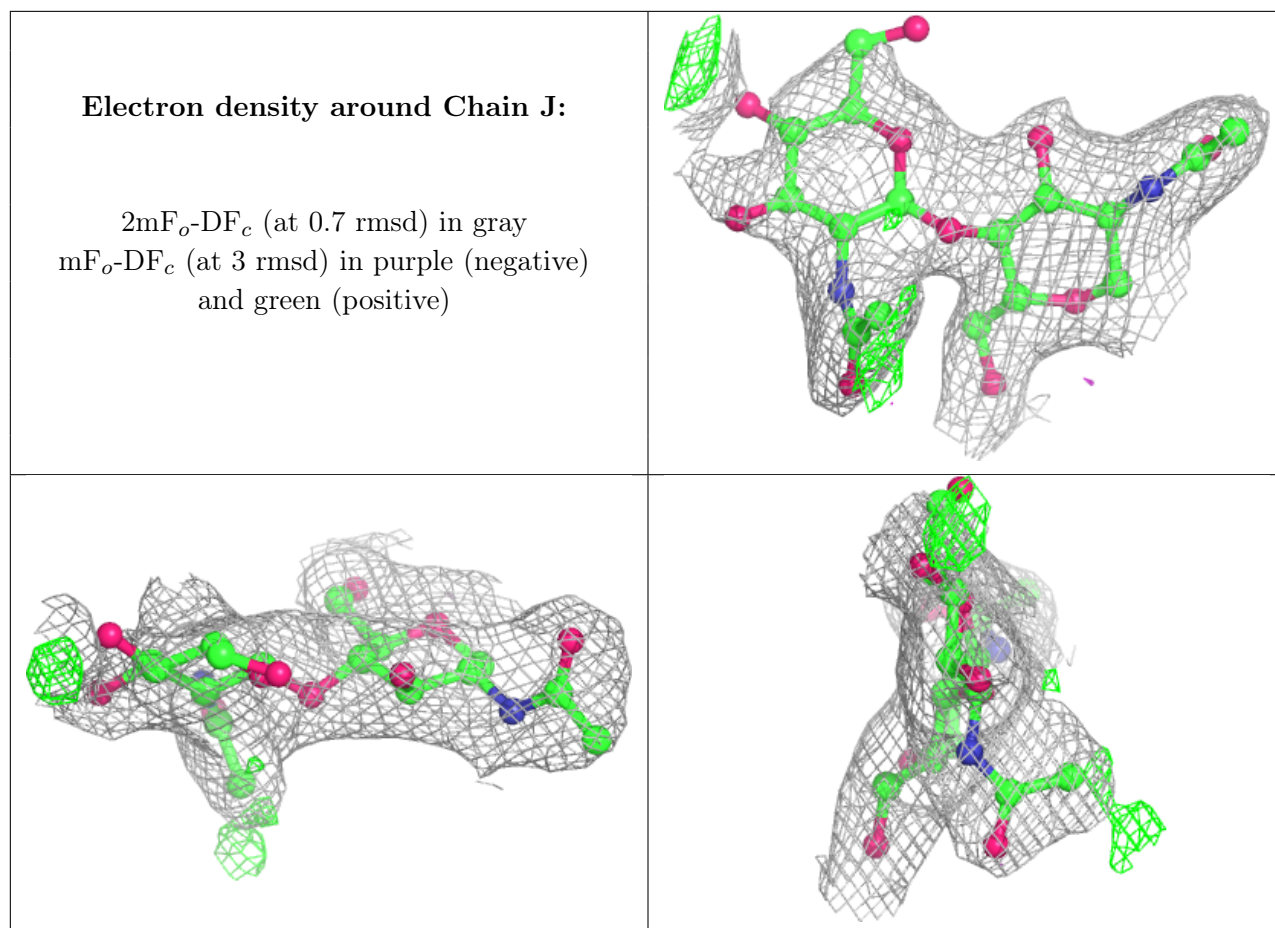
6.4 Ligands (i)

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
3	NAG	A	1509	14/15	0.66	0.23	99,115,122,124	0
3	NAG	C	1505	14/15	0.68	0.32	76,87,98,106	0
3	NAG	C	1501	14/15	0.71	0.17	64,74,83,85	0
3	NAG	B	801	14/15	0.73	0.19	69,85,94,94	0
3	NAG	D	1505	14/15	0.74	0.19	74,84,91,96	0
3	NAG	B	806	14/15	0.79	0.35	67,76,80,82	0
3	NAG	A	1503	14/15	0.79	0.38	35,40,50,51	0
3	NAG	B	805	14/15	0.81	0.16	64,71,78,79	0
3	NAG	C	1504	14/15	0.81	0.15	61,83,107,108	0
3	NAG	D	1501	14/15	0.82	0.24	66,77,88,89	0
3	NAG	B	802	14/15	0.86	0.22	56,63,67,71	0
3	NAG	B	807	14/15	0.88	0.28	54,65,72,75	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	D	1502	14/15	0.89	0.31	60,68,81,85	0
3	NAG	D	1506	14/15	0.89	0.23	64,79,86,87	0
3	NAG	A	1508	14/15	0.90	0.20	59,69,73,79	0
4	6RL	A	1510	26/26	0.97	0.15	24,27,30,33	0
4	6RL	D	1507	26/26	0.97	0.12	22,28,32,37	0
4	6RL	C	1506	26/26	0.98	0.14	26,30,34,36	0
4	6RL	B	808	26/26	0.98	0.14	19,26,31,33	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.
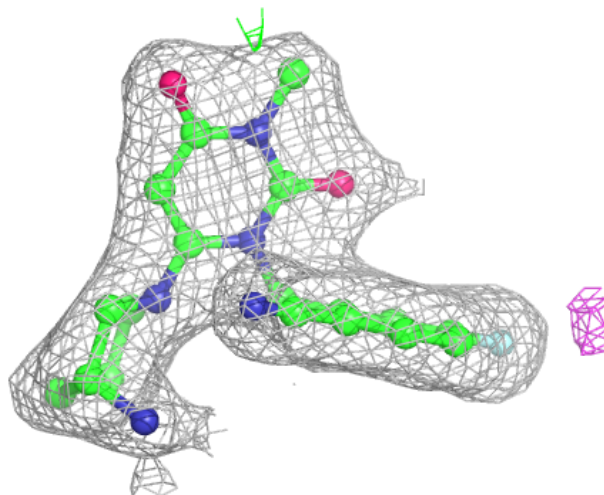
Electron density around 6RL A 1510:

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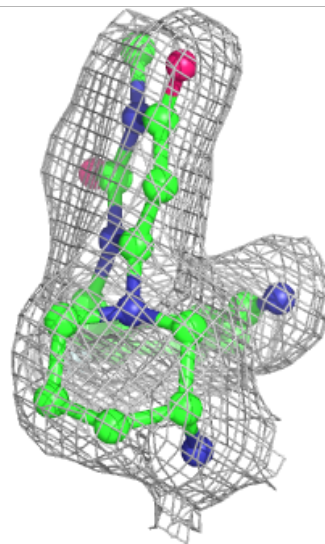mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

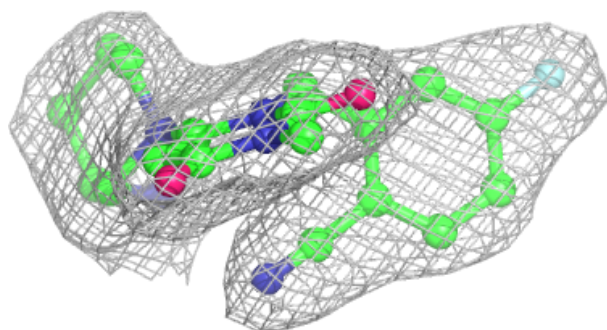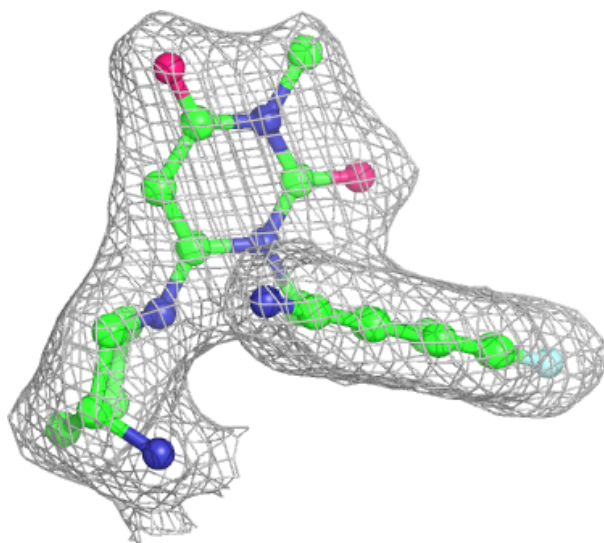wwPDB X-ray Structure Validation Summary Report
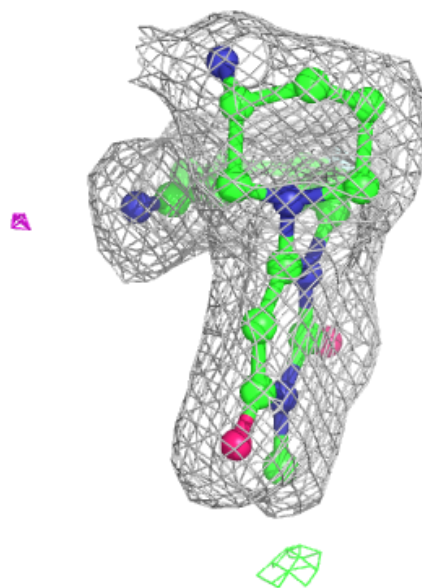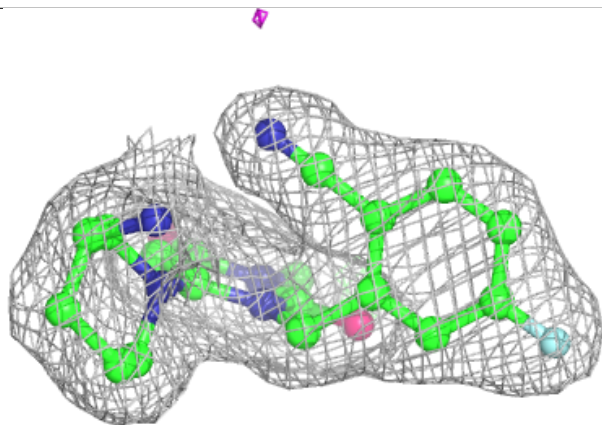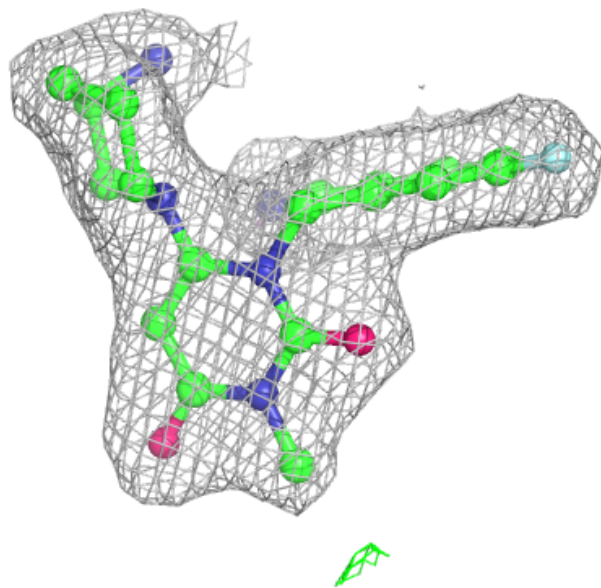
Electron density around 6RL D 1507:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
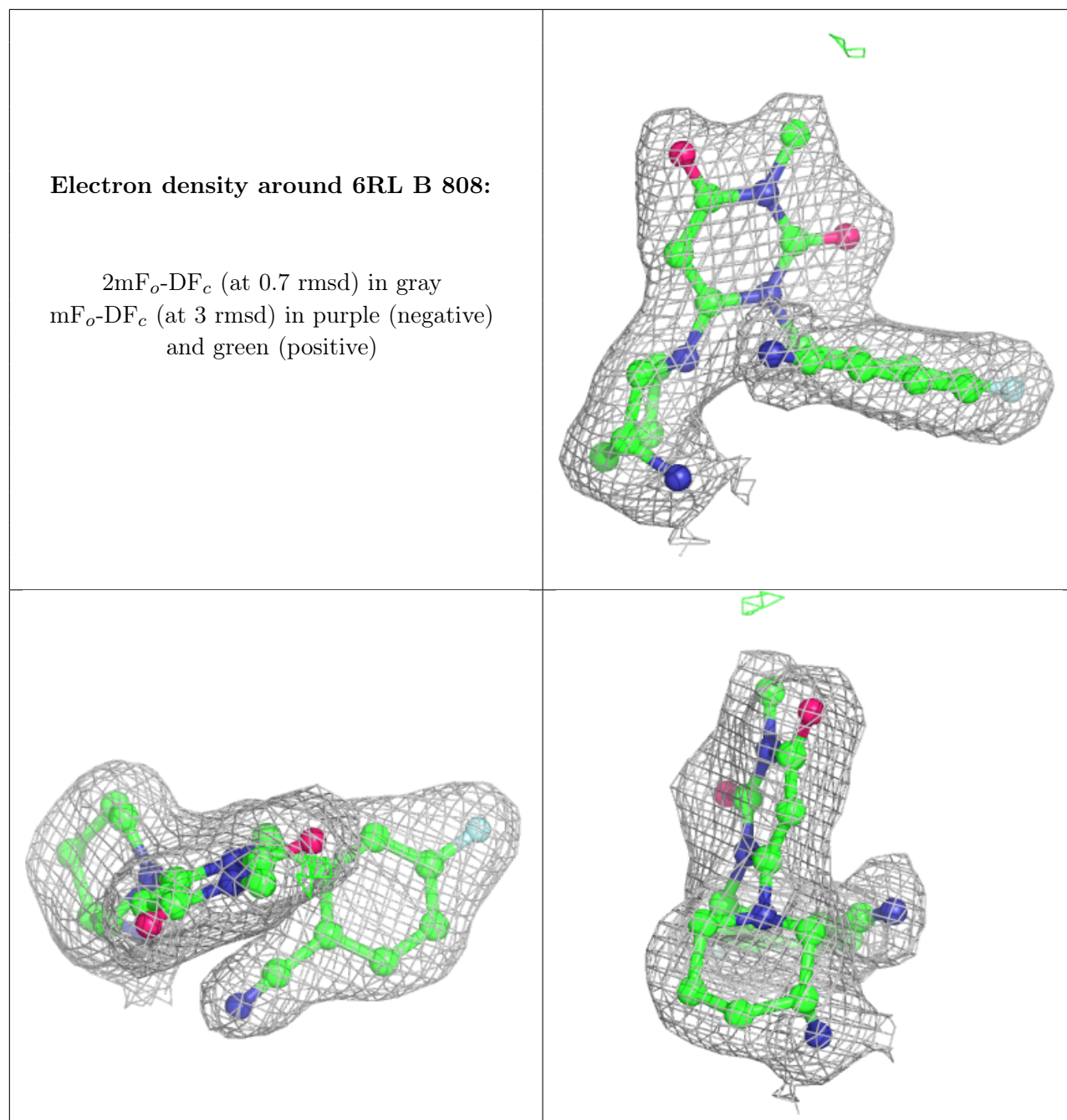and green (positive)
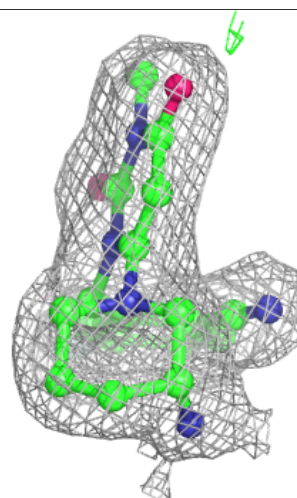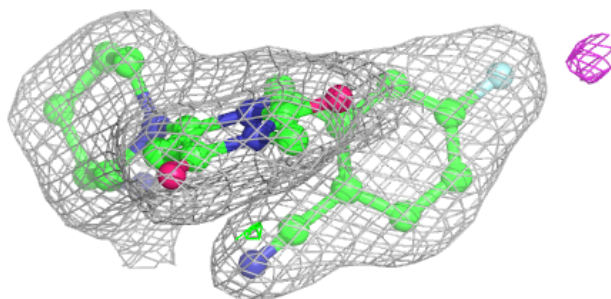
Electron density around 6RL C 1506:

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

Electron density around 6RL B 808:

$2mF_o$-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 ⓘ

There are no such residues in this entry.