W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2023 – 03:10 PM EDT

PDB ID	:	3HUS
Title	:	Crystal structure of recombinant gamma N308K fibrinogen fragment D with the peptide ligand Gly-Pro-Arg-Pro-amide
Authors	:	Lord, S.T.; Bowley, S.R.; Okumura, N.
Deposited on	:	2009-06-15
Resolution	:	3.04 Å(reported)

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

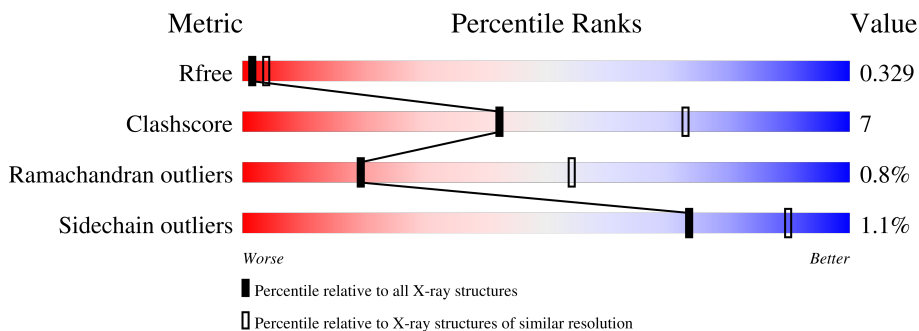|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 3.04 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		0.329
Clashscore		7
Ramachandran outliers		0.8%
Sidechain outliers		1.1%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
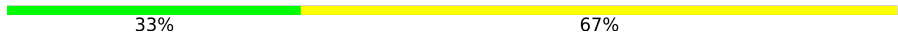
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2752 (3.08-3.00)
Clashscore	141614	3096 (3.08-3.00)
Ramachandran outliers	138981	2986 (3.08-3.00)
Sidechain outliers	138945	2988 (3.08-3.00)

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	66	79% 11% · 8%
1	D	66	82% 11% 8%
2	B	313	79% 16% · ·
2	E	313	84% 12% · ·
3	C	311	82% 11% 6%
3	F	311	79% 14% · 6%
4	G	4	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	4	50% 50%
4	I	4	75% 25%
4	J	4	75% 25%
5	K	3	67% 33%
5	L	3	33% 67%

2   Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 10837 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 Fibrinogen alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	61	Total 497	C 305	N 95	O 94	S 3	0	0	0
1	D	61	Total 497	C 305	N 95	O 94	S 3	0	0	0

- Molecule 2 is a protein called Fibrinogen beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	303	Total 2428	C 1514	N 429	O 463	S 22	0	0	0
2	E	303	Total 2428	C 1514	N 429	O 463	S 22	0	0	0

- Molecule 3 is a protein called Fibrinogen gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	291	Total 2335	C 1483	N 393	O 448	S 11	0	0	0
3	F	293	Total 2352	C 1493	N 397	O 451	S 11	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	308	LYS	ASN	engineered mutation	UNP P02679
F	308	LYS	ASN	engineered mutation	UNP P02679

- Molecule 4 is a protein called Peptide Ligand Gly-Pro-Arg-Pro-amide.
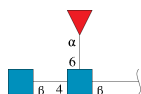
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	G	4	Total 30	C 18	N 7	O 5	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	H	4	Total 30	C 18	N 7	O 5	0	0	0
4	I	4	Total 30	C 18	N 7	O 5	0	0	0
4	J	4	Total 30	C 18	N 7	O 5	0	0	0

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	K	3	Total 38	C 22	N 2	O 14	0	0	0
5	L	3	Total 38	C 22	N 2	O 14	0	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Ca 1	0	0
6	C	1	Total 1	Ca 1	0	0
6	E	1	Total 1	Ca 1	0	0
6	F	1	Total 1	Ca 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	6	Total 6	O 6	0	0
7	B	26	Total 26	O 26	0	0
7	C	14	Total 14	O 14	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	4	Total 4	O 4	0	0
7	E	33	Total 33	O 33	0	0
7	F	17	Total 17	O 17	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.
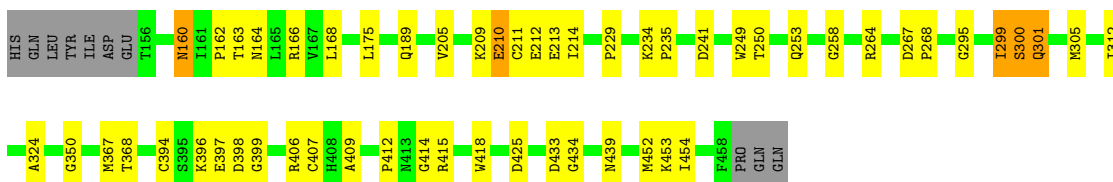
- Molecule 1: Fibrinogen alpha chain

Chain A:

- Molecule 1: Fibrinogen alpha chain

Chain D:

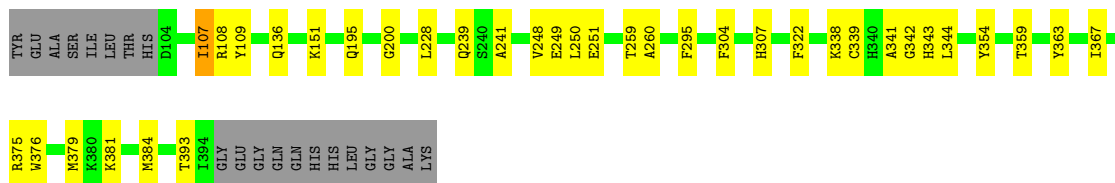- Molecule 2: Fibrinogen beta chain

Chain B:

- Molecule 2: Fibrinogen beta chain

Chain E:

- Molecule 3: Fibrinogen gamma chain

Chain C:

TYR GLU ALA SER ILE LEU THR HIS D104 I107 R108 Y109 Q136 K151 Q195 G200 L228 Q239 S240 A241 V248 E249 L250 E251 T259 A260 F295 F304 H307 F322 K338 C339 H340 A341 G342 H343 L344 Y354 T359 Y363 I367

R375 W376 M379 K380 K381 M384 T393 I394 GLY GLU GLY GLN GLN HIS HIS LEU GLY GLY ALA LYS

• Molecule 3: Fibrinogen gamma chain

Chain F:

79%   14%   • 6%

TYR GLU ALA SER ILE LEU T102 Q111 E112 I113 Y114 L124 K151 L166 L179 E183 W191 Q195 K196 R197 L198 F226 Q239 S240 A241 A245 V248 E249 L250 E251 A260 A286 T305 S306 H307 K321 N325 C326 Q329 D330 G331 S332

G333 W334 W335 M336 N337 K338 C339 H340 A341 G342 H343 L344 I367 M379 K380 K381 T382 T383 M384 F389 I394 GLY GLU GLY GLN GLN HIS HIS LEU GLY GLY ALA LYS
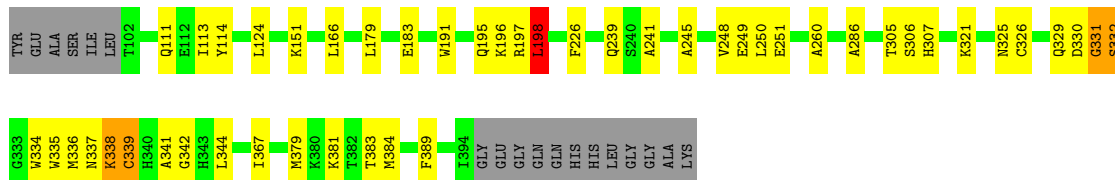
• Molecule 4: Peptide Ligand Gly-Pro-Arg-Pro-amide

Chain G:

100%

There are no outlier residues recorded for this chain.

• Molecule 4: Peptide Ligand Gly-Pro-Arg-Pro-amide

Chain H:

50%   50%

G1 P2 R3 P4

• Molecule 4: Peptide Ligand Gly-Pro-Arg-Pro-amide

Chain I:

75%   25%

G1 P2 R3 P4

• Molecule 4: Peptide Ligand Gly-Pro-Arg-Pro-amide

Chain J:

75%   25%

G1 P2 R3 P4

• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-ace
tamido-2-deoxy-beta-D-glucopyranose

Chain K:

67%   33%

NAG1 NAG2 FUC3

• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-ace
tamido-2-deoxy-beta-D-glucopyranose

Chain L:

33%

67%

NAG1 NAG2 FUC3

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	95.01Å 95.01Å 448.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.25 – 3.04 47.24 – 3.04	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.25-3.04) 99.4 (47.24-3.04)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.14 (at 3.07Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.220 , 0.288 0.292 , 0.329	Depositor DCC
R_{free} test set	2030 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	65.0	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 35.2	EDS
L-test for twinning[2]	$< |L| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	10837	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 42.63 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9848e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: FUC, CA, 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	0.38	0/498	0.50	0/665
1	D	0.26	0/498	0.45	0/665
2	B	0.37	0/2489	0.46	0/3362
2	E	0.36	0/2489	0.45	0/3362
3	C	0.32	0/2399	0.42	0/3243
3	F	0.38	0/2417	0.44	1/3268 (0.0%)
4	G	0.49	0/31	0.46	0/40
4	H	0.49	0/31	0.46	0/40
4	I	0.50	0/31	0.52	0/40
4	J	0.50	0/31	0.65	0/40
All	All	0.36	0/10914	0.45	1/14725 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	198	LEU	C-N-CA	-5.29	108.49	121.70

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	497	0	517	12	0
1	D	497	0	517	6	0
2	B	2428	0	2295	42	0
2	E	2428	0	2296	31	0
3	C	2335	0	2192	23	0
3	F	2352	0	2208	44	0
4	G	30	0	32	0	0
4	H	30	0	32	1	0
4	I	30	0	32	1	0
4	J	30	0	32	2	0
5	K	38	0	34	0	0
5	L	38	0	34	1	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	E	1	0	0	0	0
6	F	1	0	0	0	0
7	A	6	0	0	0	0
7	B	26	0	0	1	0
7	C	14	0	0	1	0
7	D	4	0	0	0	0
7	E	33	0	0	2	0
7	F	17	0	0	0	0
All	All	10837	0	10221	144	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 7.

The worst 5 of 144 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:159:ARG:NH1	2:B:418:TRP:CE3	2.08	1.22
3:F:198:LEU:HD12	3:F:198:LEU:N	1.50	1.17
3:F:326:CYS:HB3	3:F:339:CYS:SG	1.85	1.16
3:F:326:CYS:CB	3:F:339:CYS:SG	2.41	1.08
1:A:159:ARG:O	2:B:258:GLY:O	1.70	1.07

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	59/66 (89%)	55 (93%)	3 (5%)	1 (2%)	9	35
1	D	59/66 (89%)	58 (98%)	1 (2%)	0	100	100
2	B	301/313 (96%)	271 (90%)	29 (10%)	1 (0%)	41	74
2	E	301/313 (96%)	275 (91%)	24 (8%)	2 (1%)	22	57
3	C	289/311 (93%)	267 (92%)	21 (7%)	1 (0%)	41	74
3	F	291/311 (94%)	265 (91%)	21 (7%)	5 (2%)	9	35
4	G	2/4 (50%)	2 (100%)	0	0	100	100
4	H	2/4 (50%)	2 (100%)	0	0	100	100
4	I	2/4 (50%)	2 (100%)	0	0	100	100
4	J	2/4 (50%)	2 (100%)	0	0	100	100
All	All	1308/1396 (94%)	1199 (92%)	99 (8%)	10 (1%)	19	54

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	157	VAL
3	F	335	TRP
3	F	338	LYS
1	A	160	SER
2	E	396	LYS

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	56/61 (92%)	54 (96%)	2 (4%)	35	68
1	D	56/61 (92%)	56 (100%)	0	100	100
2	B	261/271 (96%)	257 (98%)	4 (2%)	65	86
2	E	261/271 (96%)	259 (99%)	2 (1%)	81	92
3	C	245/259 (95%)	243 (99%)	2 (1%)	81	92
3	F	247/259 (95%)	245 (99%)	2 (1%)	81	92
4	G	3/3 (100%)	3 (100%)	0	100	100
4	H	3/3 (100%)	3 (100%)	0	100	100
4	I	3/3 (100%)	3 (100%)	0	100	100
4	J	3/3 (100%)	3 (100%)	0	100	100
All	All	1138/1194 (95%)	1126 (99%)	12 (1%)	73	90

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	393	THR
2	E	280	THR
3	F	339	CYS
2	E	428	LYS
2	B	299	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	253	GLN
2	E	364	ASN
3	F	230	ASN
2	E	339	GLN
2	E	408	HIS

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 ⓘ

6 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	$\#|Z| > 2$
5	NAG	K	1	5,2	14,14,15	0.52	0	17,19,21	0.86	0
5	NAG	K	2	5	14,14,15	0.76	0	17,19,21	2.43	5 (29%)
5	FUC	K	3	5	10,10,11	0.64	0	14,14,16	0.69	0
5	NAG	L	1	5	14,14,15	0.56	0	17,19,21	0.96	0
5	NAG	L	2	5	14,14,15	0.64	0	17,19,21	1.54	3 (17%)
5	FUC	L	3	5	10,10,11	0.62	0	14,14,16	0.74	0

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
5	NAG	K	1	5,2	-	0/6/23/26	0/1/1/1
5	NAG	K	2	5	-	5/6/23/26	0/1/1/1
5	FUC	K	3	5	-	-	0/1/1/1
5	NAG	L	1	5	-	2/6/23/26	0/1/1/1
5	NAG	L	2	5	-	4/6/23/26	0/1/1/1
5	FUC	L	3	5	-	-	0/1/1/1

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	K	2	NAG	C2-N2-C7	6.85	132.66	122.90
5	K	2	NAG	C4-C3-C2	4.28	117.29	111.02
5	L	2	NAG	C2-N2-C7	3.65	128.10	122.90
5	K	2	NAG	C8-C7-N2	3.61	122.21	116.10
5	L	2	NAG	C4-C3-C2	3.26	115.80	111.02

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	L	1	NAG	O5-C5-C6-O6
5	L	1	NAG	C4-C5-C6-O6
5	K	2	NAG	C8-C7-N2-C2
5	K	2	NAG	O7-C7-N2-C2
5	L	2	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	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 K

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain L

Bond lengths

Bond angles

Torsions

Rings
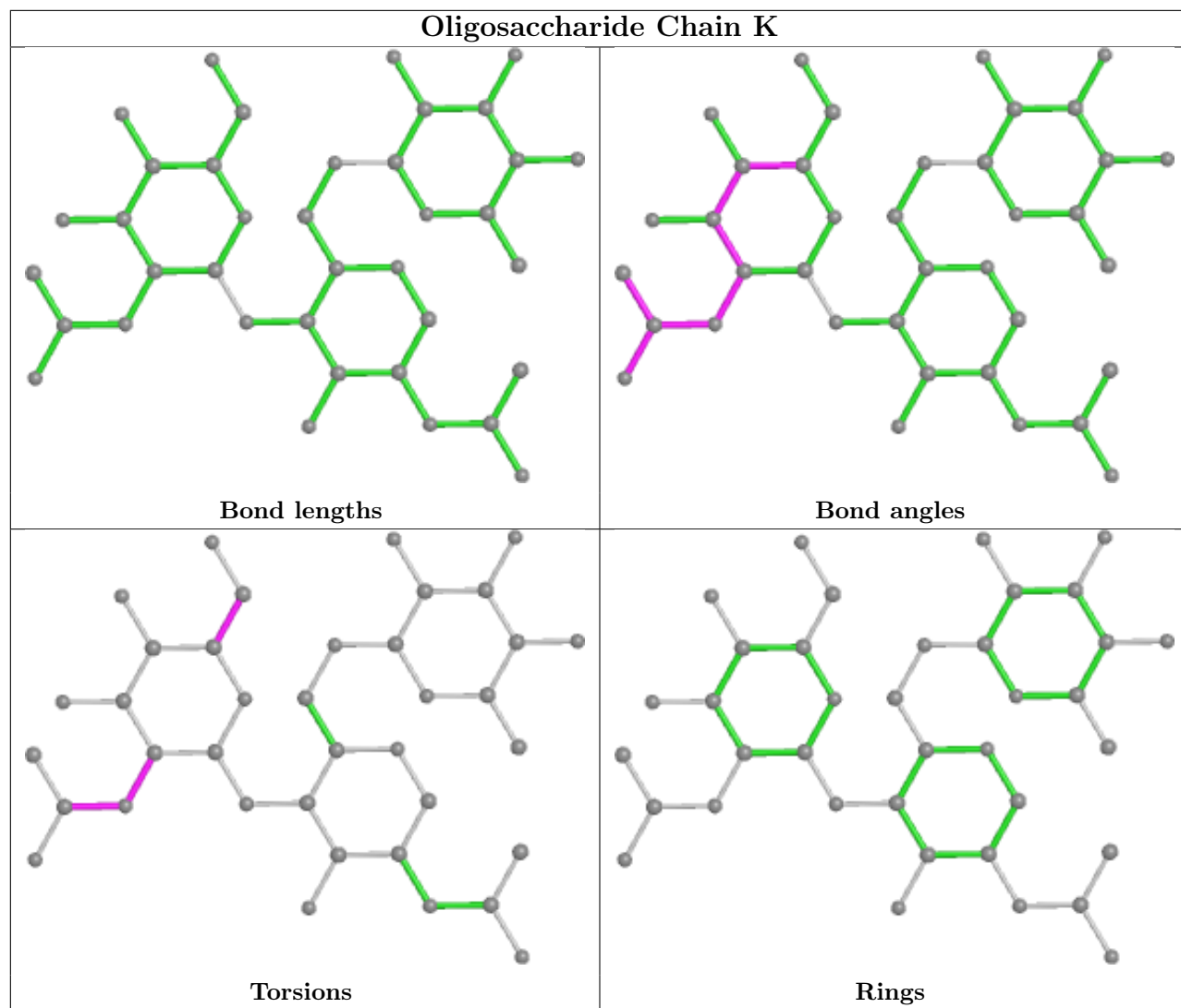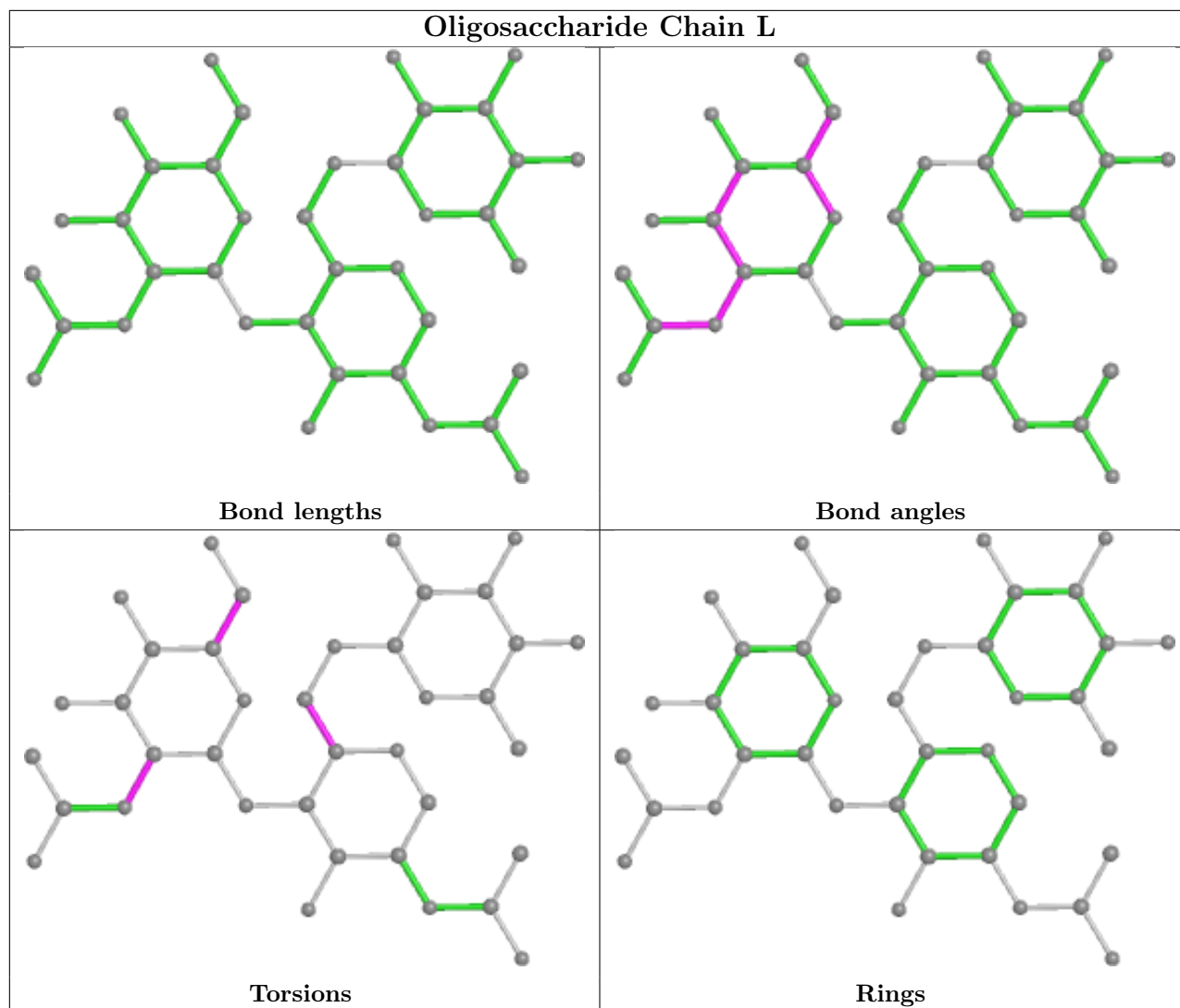
5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.

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 K:

$2mF_o$-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 L:

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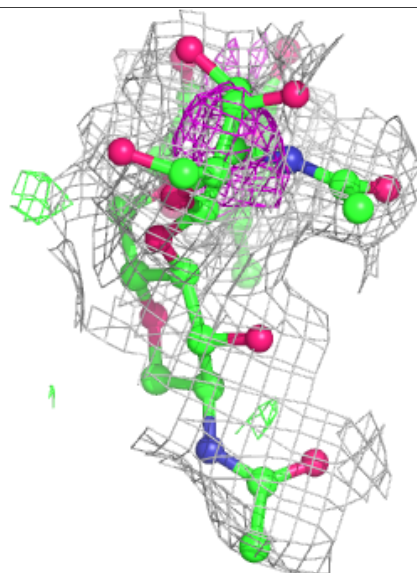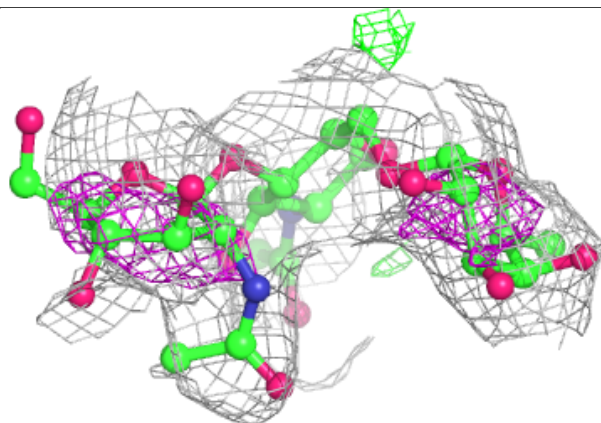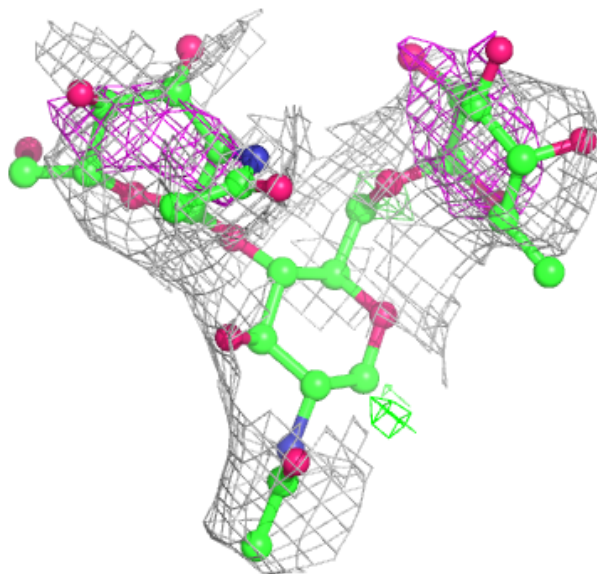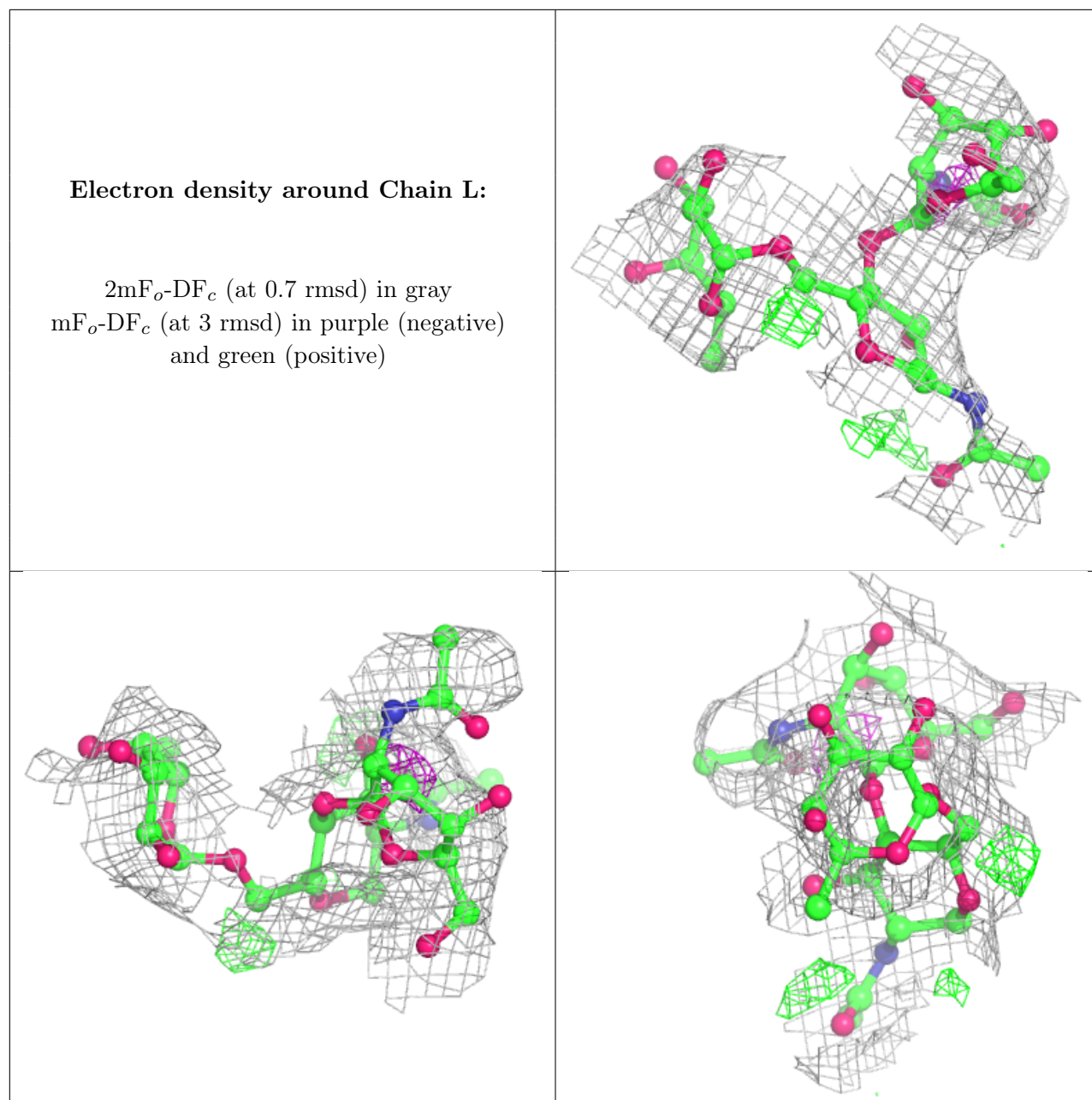
6.4 Ligands ⓘ

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

6.5 Other polymers ⓘ

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