wwPDB X-ray Structure Validation Summary Report (i)

Oct 14, 2023 – 02:33 PM EDT

PDB ID : 7T2C
Title : Crystal structure of the B5 TCR in complex with HLA-DP4-Ply
Authors : Ciacchi, L.; Farenc, C.; Petersen, J.; Reid, H.H.; Rossjohn, J.
Deposited on : 2021-12-03
Resolution : 3.10 Å(reported)

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

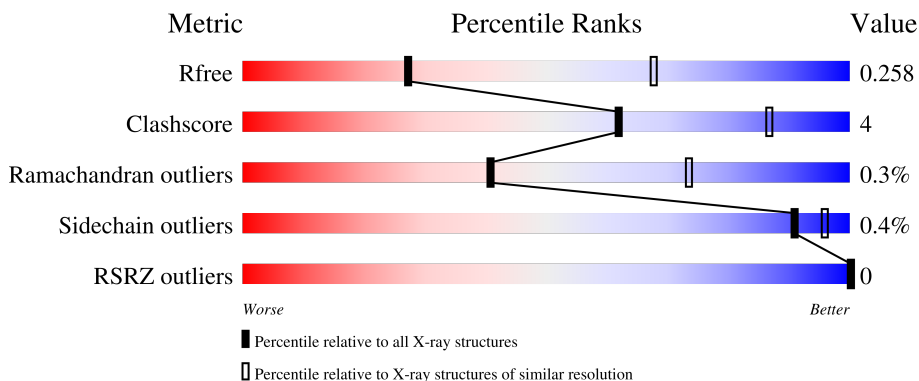|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.10 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		0.258
Clashscore		4
Ramachandran outliers		0.3%
Sidechain outliers		0.4%
RSRZ outliers		0

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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	D	204	84% 7% 9%
2	E	249	86% 14%
3	A	181	89% 10% ·
4	B	188	79% 14% · 6%
5	C	15	53% 27% 20%

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	2	100%

2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 6477 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 T cell receptor, B5, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	185	Total	C	N	O	S	0	0	0
			1440	900	240	292	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	CYS	THR	engineered mutation	UNP P01848

- Molecule 2 is a protein called T cell receptor, B5, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	248	Total	C	N	O	S	0	0	0
			1957	1234	335	379	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	184	CYS	SER	engineered mutation	UNP P01850
E	202	ALA	CYS	engineered mutation	UNP P01850

- Molecule 3 is a protein called HLA class II histocompatibility antigen, DP alpha 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	180	Total	C	N	O	S	0	0	0
			1473	953	238	277	5			

- Molecule 4 is a protein called HLA class II histocompatibility antigen, DP beta 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	176	Total	C	N	O	S	0	0	0
			1446	909	255	274	8			

- Molecule 5 is a protein called Pneumolysin-derived peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	C	12	Total 112	C 77	N 18	O 17	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	cloning artifact	UNP Q04IN8
C	-2	ALA	-	cloning artifact	UNP Q04IN8

- Molecule 6 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
6	F	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
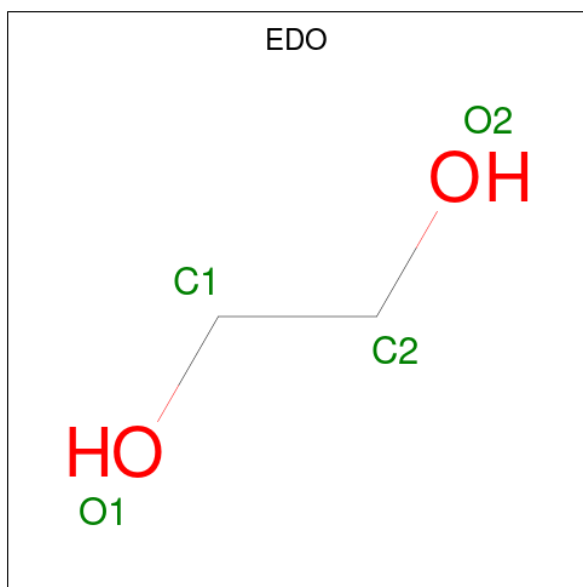
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	E	1	Total 4	C 2	O 2	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

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

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total 1	O 1	0	0
9	B	1	Total 1	O 1	0	0
9	C	1	Total 1	O 1	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.
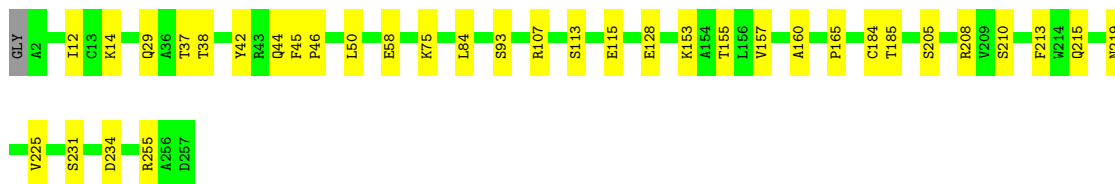
- Molecule 1: T cell receptor, B5, alpha chain

Chain D:

84%   7%   9%

LEU A2 T30 K58 A76 L91 P92 G112 L140 R141 D142 S143 V150 V162 K175 C176 W193 A204 PHE ASN ASN SER ILE ILE PRO GLU ASP THR PHE PHE PRO SER PRO GLU SER SER

- Molecule 2: T cell receptor, B5, beta chain

Chain E:

86%   14%

GLY A2 I12 C13 K14 Q29 A36 T37 T38 Y42 R43 Q44 F45 P46 L50 E58 K75 L84 S93 R107 S113 E115 E128 K153 A154 T155 L156 V157 A160 P165 C184 T185 S205 R208 V209 S210 F213 W214 Q215 N219

V225 S231 D234 R255 A256 D257

- Molecule 3: HLA class II histocompatibility antigen, DP alpha 1 chain

Chain A:

89%   10%   •

I1 K2 A3 Y9 F12 D27 K39 F52 A61 L73 D85 P86 P87 F112 P115 N118 S144 F145 E166 Q172 P173 A180 GLN

- Molecule 4: HLA class II histocompatibility antigen, DP beta 1 chain

Chain B:

79%   14%   •   6%

ARG ALA T3 L8 F9 Q10 Q13 F18 I31 R34 E52 R55 P56 A57 Q64 I67 R72 R77 L85 G86 P97 R98 V99 S104 LYS LYS GLY PRO LEU GLN HIS HIS N113 V116 C117 H118 V119 F122 R149 D152 W153 T154

F155 Q156 M160 L180 V184 W188 LYS ALA
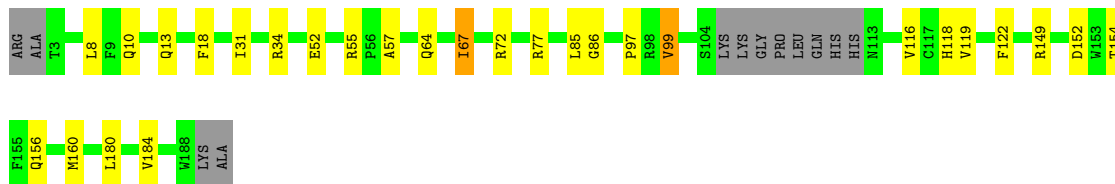
- Molecule 5: Pneumolysin-derived peptide

Chain C:

53%   27%   20%

GLY ALA T-1 G0 L1 E4 R7 T8 V9 Y10 GLU
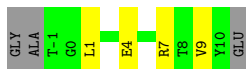
• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc
opyranose

Chain F:

100%

NAG1 NAG2

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	91.93Å 77.71Å 153.68Å 90.00° 96.71° 90.00°	Depositor
Resolution (Å)	46.80 – 3.10	Depositor
	48.58 – 3.10	EDS
% Data completeness	99.6 (46.80-3.10)	Depositor
(in resolution range)	99.6 (48.58-3.10)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.211 , 0.257	Depositor
	0.214 , 0.258	DCC
R_{free} test set	983 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	61.5	Xtriage
Anisotropy	0.537	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 26.9	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6477	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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, EDO

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	D	0.24	0/1469	0.44	0/1996
2	E	0.24	0/2011	0.41	0/2736
3	A	0.24	0/1521	0.42	0/2076
4	B	0.25	0/1481	0.43	0/2011
5	C	0.20	0/118	0.36	0/163
All	All	0.24	0/6600	0.42	0/8982

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	D	1440	0	1387	7	0
2	E	1957	0	1860	24	0
3	A	1473	0	1376	15	0
4	B	1446	0	1365	21	0
5	C	112	0	102	5	0
6	F	28	0	25	0	0
7	E	4	0	6	0	0
8	B	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	B	1	0	0	1	0
9	C	1	0	0	0	0
9	D	1	0	0	0	0
All	All	6477	0	6134	55	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 4.

The worst 5 of 55 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)
3:A:9:TYR:HB2	4:B:13:GLN:HB2	1.69	0.73
1:D:176:CYS:CB	2:E:184:CYS:HG	2.01	0.68
2:E:58:GLU:O	3:A:39:LYS:NZ	2.29	0.65
1:D:150:VAL:HG23	1:D:193:TRP:HB3	1.78	0.64
3:A:12:PHE:HB2	4:B:8:LEU:HD11	1.81	0.62

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	D	183/204 (90%)	171 (93%)	10 (6%)	2 (1%)	14	46
2	E	246/249 (99%)	238 (97%)	8 (3%)	0	100	100
3	A	178/181 (98%)	173 (97%)	5 (3%)	0	100	100
4	B	172/188 (92%)	164 (95%)	8 (5%)	0	100	100
5	C	10/15 (67%)	10 (100%)	0	0	100	100
All	All	789/837 (94%)	756 (96%)	31 (4%)	2 (0%)	41	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	58	LYS
1	D	92	PRO

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	D	167/186 (90%)	166 (99%)	1 (1%)	86	94
2	E	216/216 (100%)	216 (100%)	0	100	100
3	A	161/162 (99%)	161 (100%)	0	100	100
4	B	158/167 (95%)	156 (99%)	2 (1%)	69	87
5	C	10/11 (91%)	10 (100%)	0	100	100
All	All	712/742 (96%)	709 (100%)	3 (0%)	91	96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	30	THR
4	B	67	ILE
4	B	99	VAL

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

Mol	Chain	Res	Type
1	D	159	GLN
2	E	48	GLN
3	A	68	ASN
4	B	13	GLN
4	B	156	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

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	$\#|Z| > 2$
6	NAG	F	1	6,3	14,14,15	0.24	0	17,19,21	0.48	0
6	NAG	F	2	6	14,14,15	0.24	0	17,19,21	0.40	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
6	NAG	F	1	6,3	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	F	1	NAG	O5-C5-C6-O6
6	F	1	NAG	C4-C5-C6-O6
6	F	2	NAG	C4-C5-C6-O6
6	F	2	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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 F

Bond lengths

Bond angles

Torsions

Rings
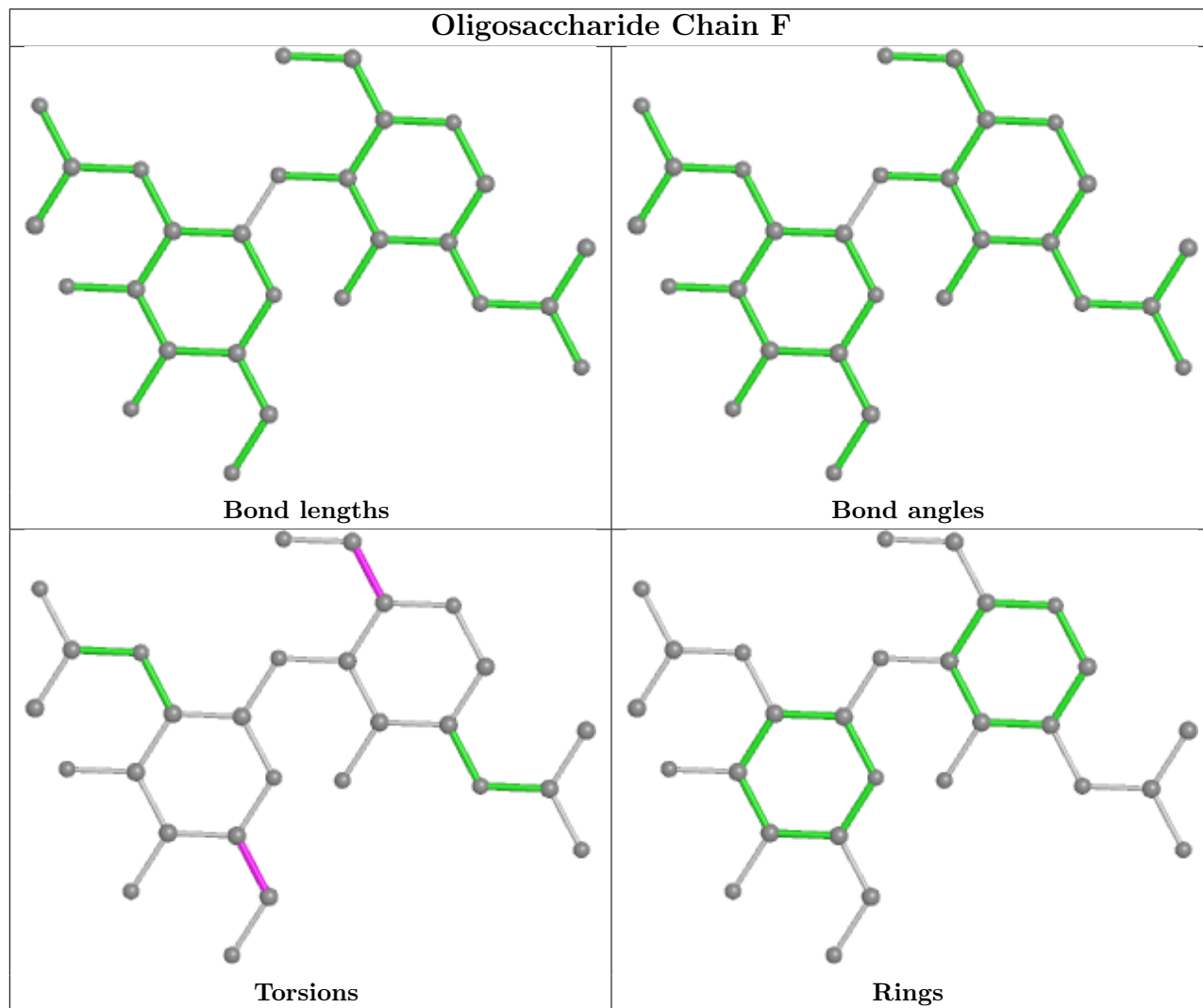
5.6   Ligand geometry (i)

2 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	$\#\|Z\|>2$
8	NAG	B	201	4	14,14,15	0.30	0	17,19,21	0.46	0
7	EDO	E	301	-	3,3,3	0.46	0	2,2,2	0.32	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
8	NAG	B	201	4	-	0/6/23/26	0/1/1/1
7	EDO	E	301	-	-	0/1/1/1	-

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 ⓘ

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(Å²)	Q<0.9
1	D	185/204 (90%)	-0.16	0	100	100	42, 65, 126, 168	0
2	E	248/249 (99%)	-0.22	0	100	100	44, 65, 103, 135	0
3	A	180/181 (99%)	-0.21	0	100	100	46, 65, 90, 116	0
4	B	176/188 (93%)	-0.15	0	100	100	45, 67, 101, 131	0
5	C	12/15 (80%)	-0.16	0	100	100	46, 50, 66, 107	0
All	All	801/837 (95%)	-0.19	0	100	100	42, 65, 106, 168	0

There are no RSRZ outliers to report.

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

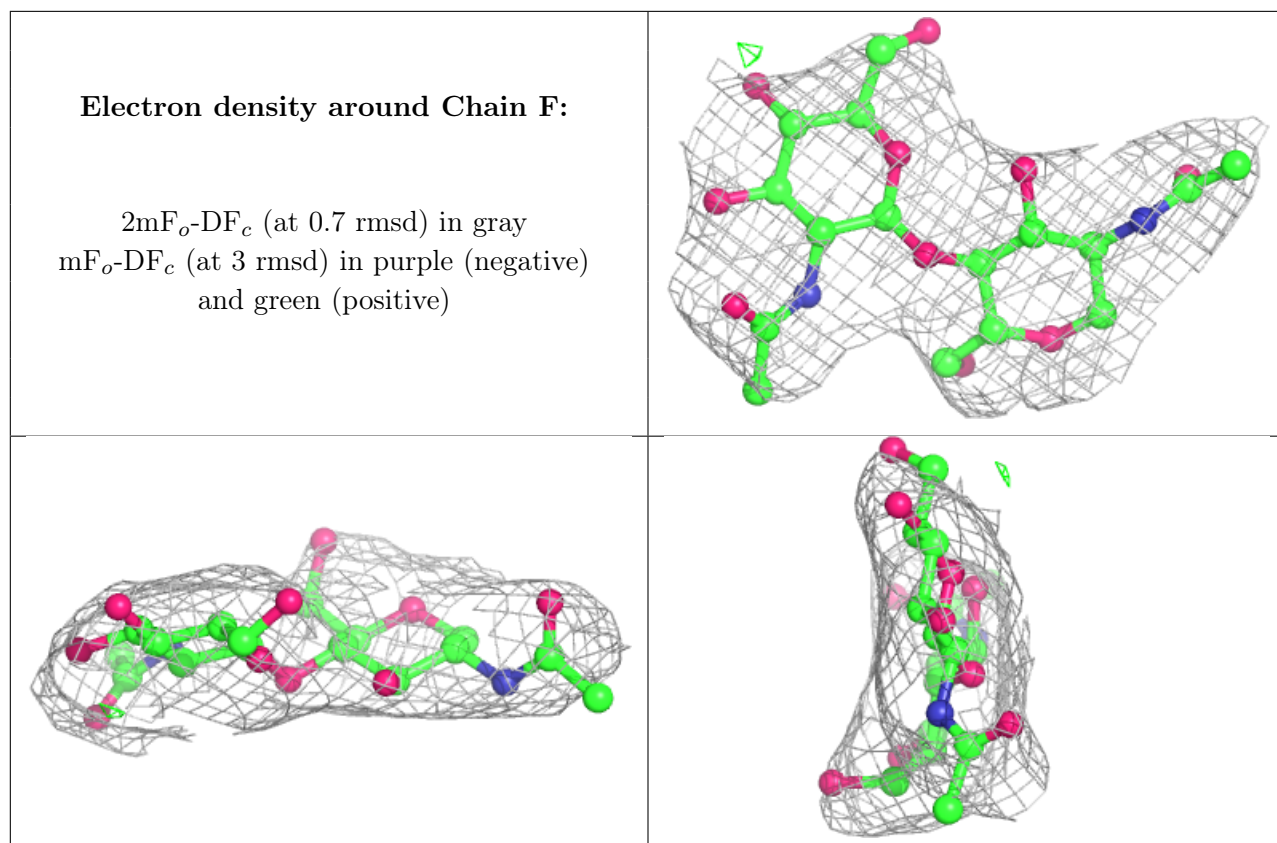There are no non-standard protein/DNA/RNA residues in this entry.

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(Å²)	Q<0.9
6	NAG	F	2	14/15	0.87	0.20	97,114,136,140	0
6	NAG	F	1	14/15	0.94	0.20	69,82,100,110	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 F:

$2mF_o\text{-}DF_c$ (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
8	NAG	B	201	14/15	0.71	0.19	79,113,126,127	0
7	EDO	E	301	4/4	0.81	0.26	50,59,64,69	0

6.5 Other polymers ⓘ

There are no such residues in this entry.