wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2020 – 06:24 PM BST

PDB ID	:	4FQX
Title	:	Crystal structure of HLA-DM bound to HLA-DR1
Authors	:	Sethi, D.K.; Pos, W.; Wucherpfennig, K.W.
Deposited on	:	2012-06-25
Resolution	:	2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

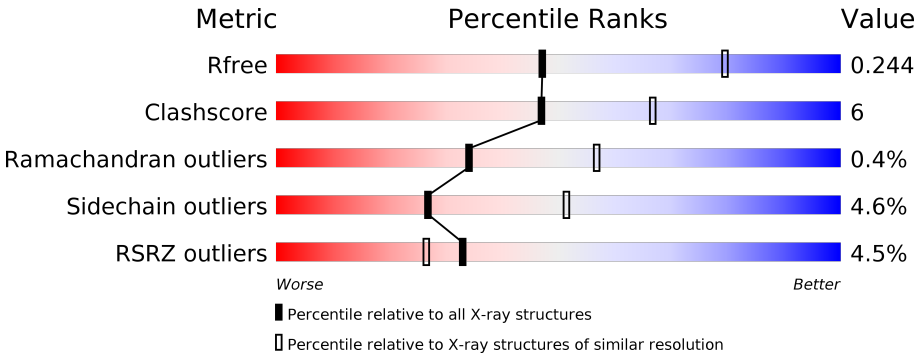
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.60 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	C	203	2% 76% 13% • 8%
2	D	199	12% 81% 13% • •
3	E	11	64% 36%
4	A	191	73% 19% • 6%
5	B	208	3% 75% 13% • 10%
6	F	2	50% 50%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6196 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 HLA class II histocompatibility antigen, DM alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	186	Total	C	N	O	S	0	0	0
			1485	969	235	275	6			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	136	GLN	HIS	variant	UNP P28067
C	137	HIS	ASP	variant	UNP P28067
C	165	ASP	ASN	engineered mutation	UNP P28067
C	200	LEU	-	expression tag	UNP P28067
C	201	VAL	-	expression tag	UNP P28067
C	202	PRO	-	expression tag	UNP P28067
C	203	ARG	-	expression tag	UNP P28067

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DM beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	191	Total	C	N	O	S	0	0	0
			1510	959	258	283	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	46	SER	CYS	engineered mutation	UNP P28068
D	92	ASP	ASN	engineered mutation	UNP P28068
D	194	GLY	-	expression tag	UNP P28068
D	195	CYS	-	expression tag	UNP P28068
D	196	LEU	-	expression tag	UNP P28068
D	197	VAL	-	expression tag	UNP P28068
D	198	PRO	-	expression tag	UNP P28068
D	199	ARG	-	expression tag	UNP P28068

- Molecule 3 is a protein called Synthetic peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	7	Total	C	N	O	S	0	0	0
			51	32	9	9	1			

- Molecule 4 is a protein called HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	180	Total	C	N	O	S	0	0	0
			1477	956	241	274	6			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	65	CYS	VAL	engineered mutation	UNP P01903

- Molecule 5 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	187	Total	C	N	O	S	0	0	0
			1521	961	268	287	5			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	VAL	-	expression tag	UNP P04229
B	-4	LEU	-	expression tag	UNP P04229
B	-3	PHE	-	expression tag	UNP P04229
B	-2	GLN	-	expression tag	UNP P04229
B	-1	GLY	-	expression tag	UNP P04229
B	0	PRO	-	expression tag	UNP P04229
B	30	SER	CYS	engineered mutation	UNP P04229
B	193	SER	-	expression tag	UNP P04229
B	194	GLY	-	expression tag	UNP P04229
B	195	GLY	-	expression tag	UNP P04229
B	196	GLY	-	expression tag	UNP P04229
B	197	SER	-	expression tag	UNP P04229
B	198	LEU	-	expression tag	UNP P04229
B	199	PRO	-	expression tag	UNP P04229
B	200	ALA	-	expression tag	UNP P04229
B	201	THR	-	expression tag	UNP P04229
B	202	GLY	-	expression tag	UNP P04229

- 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	C	N	O	0	0	0
			28	16	2	10			

- Molecule 7 is water.

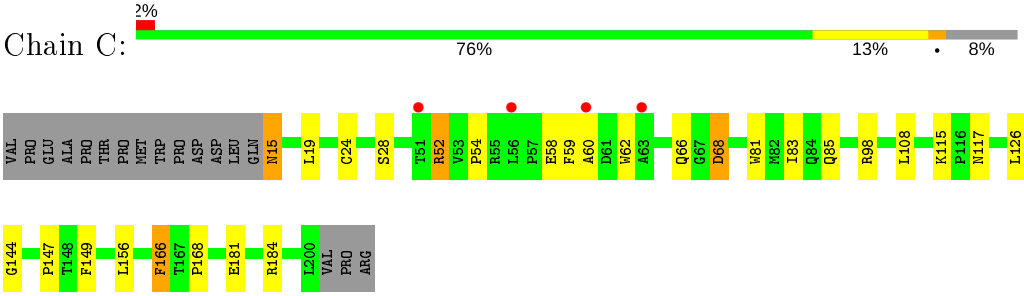
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	19	Total	O	0	0
			19	19		
7	D	13	Total	O	0	0
			13	13		
7	E	1	Total	O	0	0
			1	1		
7	A	59	Total	O	0	0
			59	59		
7	B	32	Total	O	0	0
			32	32		

3 Residue-property plots (i)

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.

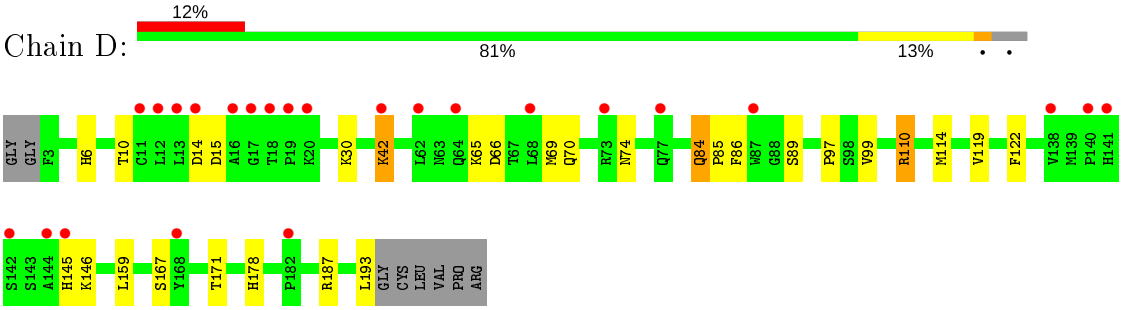
- Molecule 1: HLA class II histocompatibility antigen, DM alpha chain

Chain C:



- Molecule 2: HLA class II histocompatibility antigen, DM beta chain

Chain D:

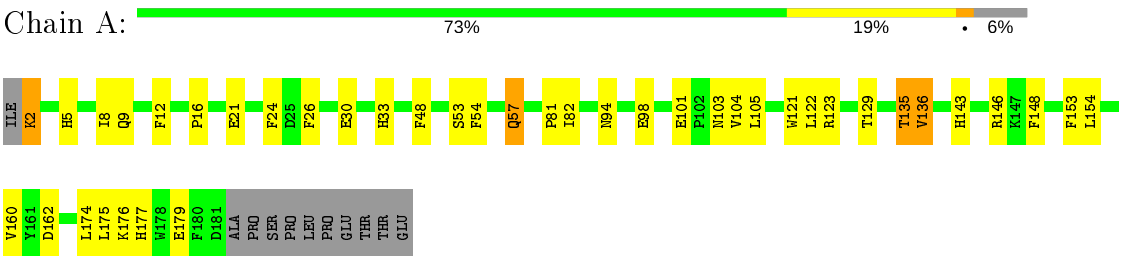


- Molecule 3: Synthetic peptide

Chain E:

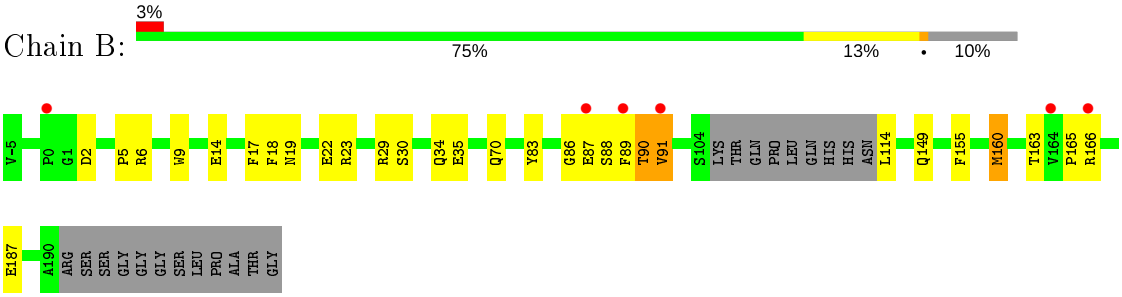
- Molecule 4: HLA class II histocompatibility antigen, DR alpha chain

Chain A:



• Molecule 5: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain B:



• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.19Å 121.65Å 138.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.14 – 2.60 36.14 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.1 (36.14-2.60) 96.8 (36.14-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.194 , 0.240 0.197 , 0.244	Depositor DCC
R_{free} test set	1742 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	54.0	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 41.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6196	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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

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	C	0.42	0/1537	0.59	0/2103
2	D	0.38	0/1556	0.56	0/2131
3	E	0.40	0/50	0.62	0/66
4	A	0.53	0/1522	0.61	0/2074
5	B	0.50	0/1559	0.65	0/2118
All	All	0.46	0/6224	0.60	0/8492

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	C	1485	0	1398	21	0
2	D	1510	0	1450	18	0
3	E	51	0	56	0	0
4	A	1477	0	1410	27	0
5	B	1521	0	1449	23	0
6	F	28	0	25	0	0
7	A	59	0	0	1	0
7	B	32	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	C	19	0	0	1	0
7	D	13	0	0	0	0
7	E	1	0	0	0	0
All	All	6196	0	5788	74	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 6.

The worst 5 of 74 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)
5:B:87:GLU:H	5:B:90:THR:HG23	1.41	0.84
2:D:110:ARG:NH1	5:B:187:GLU:OE2	2.10	0.84
1:C:181:GLU:OE2	1:C:184:ARG:NH1	2.19	0.76
2:D:70:GLN:O	2:D:74:ASN:ND2	2.18	0.75
1:C:59:PHE:HA	1:C:62:TRP:HZ3	1.52	0.73

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	C	184/203 (91%)	175 (95%)	9 (5%)	0	100 100
2	D	189/199 (95%)	181 (96%)	7 (4%)	1 (0%)	29 52
3	E	5/11 (46%)	5 (100%)	0	0	100 100
4	A	178/191 (93%)	177 (99%)	0	1 (1%)	25 47
5	B	183/208 (88%)	176 (96%)	6 (3%)	1 (0%)	29 52
All	All	739/812 (91%)	714 (97%)	22 (3%)	3 (0%)	34 57

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	84	GLN
5	B	90	THR
4	A	136	VAL

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	C	164/181 (91%)	155 (94%)	9 (6%)	21	43
2	D	169/174 (97%)	159 (94%)	10 (6%)	19	39
3	E	6/9 (67%)	6 (100%)	0	100	100
4	A	164/175 (94%)	156 (95%)	8 (5%)	25	48
5	B	166/183 (91%)	162 (98%)	4 (2%)	49	74
All	All	669/722 (93%)	638 (95%)	31 (5%)	27	51

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	146	LYS
2	D	171	THR
5	B	70	GLN
2	D	159	LEU
2	D	193	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	20	HIS
5	B	34	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	$\#\|Z\| > 2$
6	NAG	F	1	4,6	14,14,15	0.52	0	17,19,21	0.94	0
6	NAG	F	2	6	14,14,15	0.49	0	17,19,21	1.04	1 (5%)

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	4,6	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	F	2	NAG	C1-O5-C5	3.42	116.83	112.19

There are no chirality outliers.

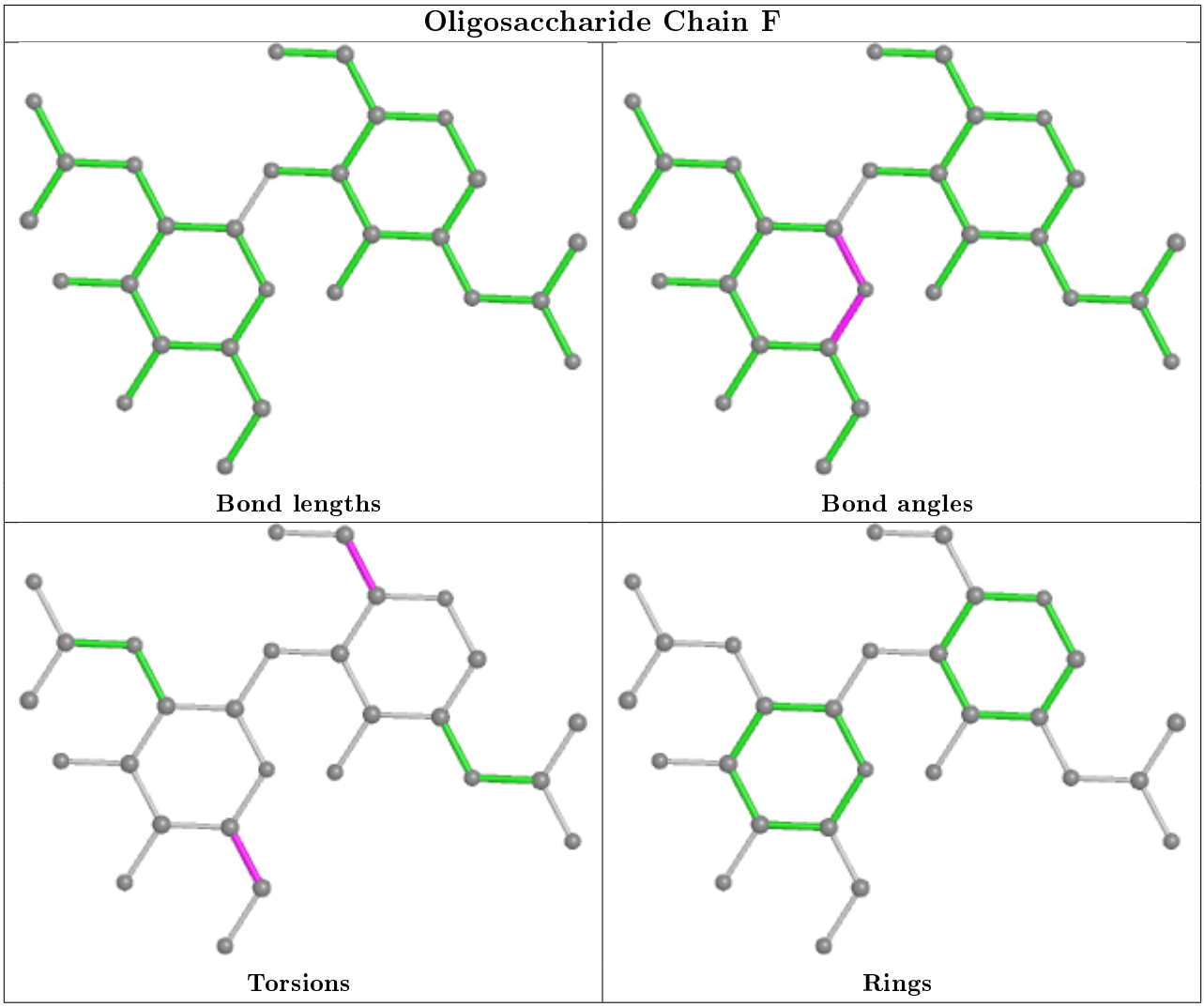
All (3) torsion outliers are listed below:

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	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.



5.6 Ligand geometry

There are no ligands in this entry.

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, 95^{th} 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	C	186/203 (91%)	0.12	4 (2%) 62 56	31, 66, 116, 162	0
2	D	191/199 (95%)	0.64	24 (12%) 3 2	36, 78, 137, 178	0
3	E	7/11 (63%)	0.08	0 100 100	44, 49, 85, 113	0
4	A	180/191 (94%)	-0.15	0 100 100	27, 40, 70, 93	0
5	B	187/208 (89%)	0.05	6 (3%) 47 40	28, 49, 108, 144	0
All	All	751/812 (92%)	0.17	34 (4%) 33 26	27, 56, 118, 178	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	91	VAL	6.8
2	D	16	ALA	5.6
2	D	17	GLY	5.6
2	D	20	LYS	4.4
2	D	12	LEU	4.3

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, 95^{th} 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
6	NAG	F	2	14/15	0.78	0.23	89,104,115,118	0

Continued on next page...

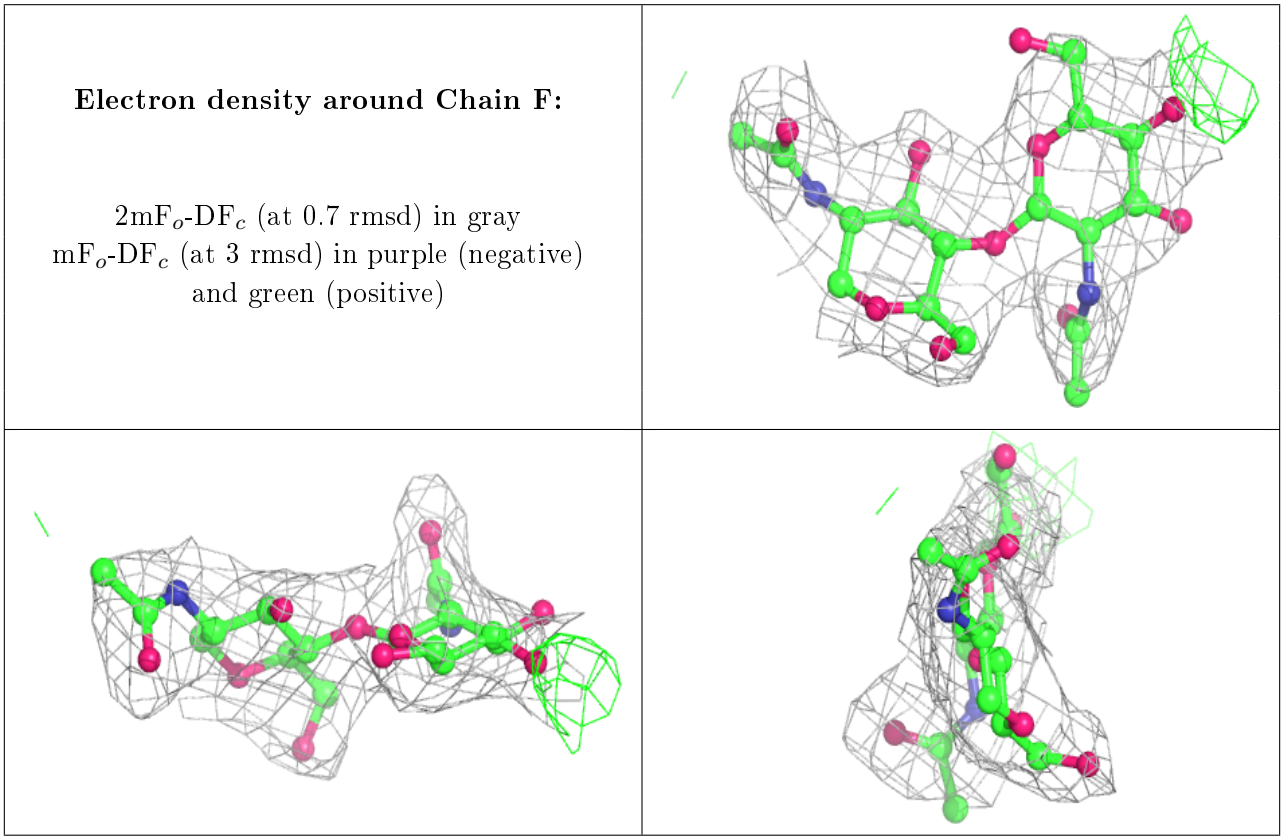
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	NAG	F	1	14/15	0.95	0.14	48,57,61,69	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:

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



6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.