wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2020 – 11:45 AM BST

PDB ID	:	3CU1
Title	:	Crystal Structure of 2:2:2 FGFR2D2:FGF1:SOS complex
Authors	:	Guo, F.; Dakshinamurthy, R.; Thallapuranam, S.K.K.; Sakon, J.
Deposited on	:	2008-04-15
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:

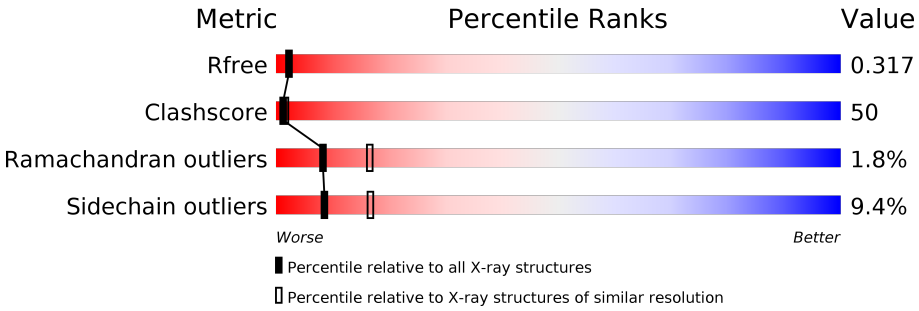
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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)

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%

Mol	Chain	Length	Quality of chain
1	A	100	32% 44% 22% •
1	C	100	33% 42% 23% •
2	B	131	40% 31% 23% • •
2	D	131	30% 43% 20% 8%
3	E	2	100%
3	F	2	100%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GU4	F	1	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4183 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.

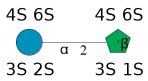
- Molecule 1 is a protein called Fibroblast growth factor receptor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	100	Total	C	N	O	S	0	3	0
			832	524	157	145	6			
1	C	100	Total	C	N	O	S	0	2	0
			825	520	154	145	6			

- Molecule 2 is a protein called Heparin-binding growth factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	128	Total	C	N	O	S	0	0	0
			1022	646	179	193	4			
2	D	131	Total	C	N	O	S	0	0	0
			1051	665	184	198	4			

- Molecule 3 is an oligosaccharide called 1,3,4,6-tetra-O-sulfo-beta-D-fructofuranose-(2-1)-2,3,4,6-tetra-O-sulfonato-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	2	Total	C	O	S	0	0	0
			55	12	35	8			
3	F	2	Total	C	O	S	0	0	0
			55	12	35	8			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	80	Total	O	0	0
			80	80		

Continued on next page...

Continued from previous page...

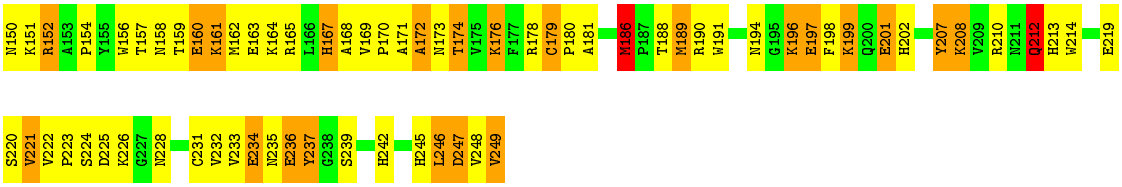
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	68	Total O 68 68	0	0
4	C	98	Total O 98 98	0	0
4	D	97	Total O 97 97	0	0

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. 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.

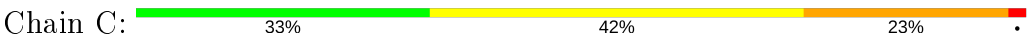
- Molecule 1: Fibroblast growth factor receptor 2

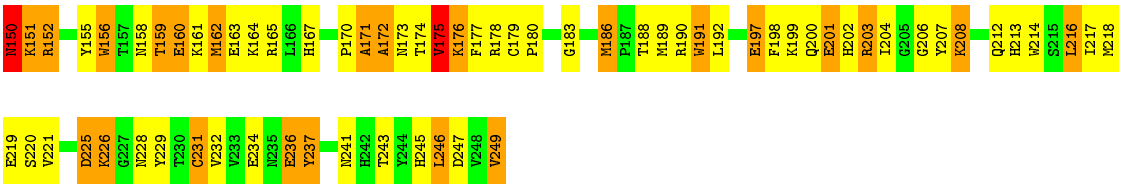
Chain A:



- Molecule 1: Fibroblast growth factor receptor 2

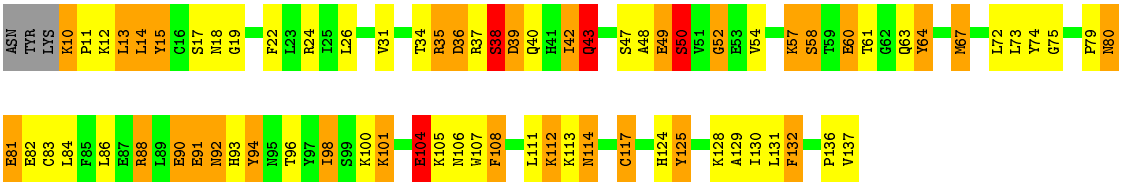
Chain C:





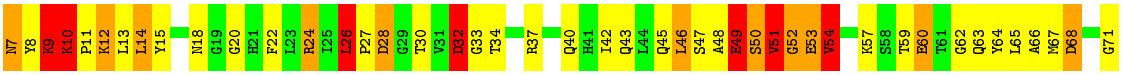
- Molecule 2: Heparin-binding growth factor 1

Chain B:



- Molecule 2: Heparin-binding growth factor 1

Chain D:



• Molecule 3: 1,3,4,6-tetra-O-sulfo-beta-D-fructofuranose-(2-1)-2,3,4,6-tetra-O-sulfonato-alpha-D-glucopyranose

Chain E:

• Molecule 3: 1,3,4,6-tetra-O-sulfo-beta-D-fructofuranose-(2-1)-2,3,4,6-tetra-O-sulfonato-alpha-D-glucopyranose

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	85.93Å 110.38Å 74.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.95 – 2.60 35.38 – 2.60	Depositor EDS
% Data completeness (in resolution range)	90.0 (42.95-2.60) 90.0 (35.38-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.24 (at 2.61Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.180 , 0.277 0.261 , 0.317	Depositor DCC
R_{free} test set	1044 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å^2)	52.0	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 57.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	4183	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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: GU4, YYJ

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	2.74	63/868 (7.3%)	1.51	9/1173 (0.8%)
1	C	2.71	59/857 (6.9%)	1.52	11/1159 (0.9%)
2	B	2.23	40/1045 (3.8%)	1.41	15/1410 (1.1%)
2	D	2.46	45/1075 (4.2%)	1.56	21/1450 (1.4%)
All	All	2.52	207/3845 (5.4%)	1.50	56/5192 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	2	0

The worst 5 of 207 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	117	CYS	CB-SG	-14.52	1.57	1.82
1	C	197	GLU	CG-CD	14.45	1.73	1.51
1	A	212[A]	GLN	CB-CG	12.71	1.86	1.52
1	A	212[B]	GLN	CB-CG	12.71	1.86	1.52
2	D	118	LYS	CE-NZ	12.13	1.79	1.49

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	247	ASP	CB-CG-OD1	-10.22	109.10	118.30
2	D	68	ASP	CB-CA-C	-9.83	90.73	110.40
2	D	91	GLU	C-N-CA	8.80	143.71	121.70
2	D	51	VAL	N-CA-C	-8.67	87.59	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	32	ASP	N-CA-CB	-8.47	95.36	110.60

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	91	GLU	CA
2	D	92	ASN	CA

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	832	0	810	57	0
1	C	825	0	801	86	0
2	B	1022	0	1007	90	0
2	D	1051	0	1035	137	0
3	E	55	0	6	4	0
3	F	55	0	6	16	0
4	A	80	0	0	26	0
4	B	68	0	0	32	0
4	C	98	0	0	32	0
4	D	97	0	0	47	0
All	All	4183	0	3665	376	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 50.

The worst 5 of 376 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:C:197:GLU:CB	1:C:197:GLU:CG	1.74	1.63
2:D:128:LYS:NZ	2:D:128:LYS:CE	1.68	1.57
2:D:128:LYS:CD	2:D:128:LYS:CE	1.77	1.57
2:D:26:LEU:CD2	2:D:26:LEU:CG	1.82	1.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:10:LYS:CG	2:D:11:PRO:HD3	1.07	1.51

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	101/100 (101%)	98 (97%)	3 (3%)	0	100 100
1	C	100/100 (100%)	93 (93%)	7 (7%)	0	100 100
2	B	126/131 (96%)	113 (90%)	11 (9%)	2 (2%)	9 19
2	D	129/131 (98%)	114 (88%)	9 (7%)	6 (5%)	2 2
All	All	456/462 (99%)	418 (92%)	30 (7%)	8 (2%)	8 16

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	10	LYS
2	D	92	ASN
2	B	52	GLY
2	D	50	SER
2	D	80	ASN

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	91/88 (103%)	83 (91%)	8 (9%)	10 19
1	C	90/88 (102%)	84 (93%)	6 (7%)	16 33
2	B	112/115 (97%)	99 (88%)	13 (12%)	5 10
2	D	115/115 (100%)	103 (90%)	12 (10%)	7 13
All	All	408/406 (100%)	369 (90%)	39 (10%)	8 16

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	81	GLU
1	C	150	ASN
2	D	80	ASN
2	B	101	LYS
2	B	104	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	150	ASN
2	D	114	ASN
1	C	202	HIS
2	B	106	ASN
1	C	173	ASN

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)

4 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	$\#\|Z\| > 2$
3	GU4	E	1	3	27,27,28	1.59	6 (22%)	29,43,45	2.50	9 (31%)
3	YYJ	E	2	3	27,28,28	2.81	7 (25%)	28,46,46	2.01	9 (32%)
3	GU4	F	1	3	27,27,28	2.39	9 (33%)	29,43,45	3.75	12 (41%)
3	YYJ	F	2	3	27,28,28	2.24	7 (25%)	28,46,46	1.72	3 (10%)

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	GU4	E	1	3	-	11/21/38/41	0/1/1/1
3	YYJ	E	2	3	-	11/23/42/42	0/1/1/1
3	GU4	F	1	3	-	16/21/38/41	0/1/1/1
3	YYJ	F	2	3	-	13/23/42/42	0/1/1/1

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	2	YYJ	O3-S3	6.55	1.76	1.57
3	E	2	YYJ	O6-S6	6.09	1.73	1.56
3	E	2	YYJ	O1-S1	6.08	1.73	1.56
3	F	1	GU4	O2-S2	5.86	1.74	1.57
3	E	2	YYJ	O4-S4	5.38	1.73	1.57

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	1	GU4	O2-C2-C3	14.70	122.93	106.65
3	F	1	GU4	C2-O2-S2	7.95	128.27	117.91
3	E	1	GU4	O2-C2-C3	7.81	115.29	106.65
3	E	1	GU4	C2-O2-S2	6.07	125.83	117.91
3	E	2	YYJ	C4-O4-S4	5.86	130.19	118.88

There are no chirality outliers.

5 of 51 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	2	YYJ	O1-C1-C2-C3
3	E	2	YYJ	O1-C1-C2-O2
3	E	2	YYJ	O1-C1-C2-O5
3	E	2	YYJ	C2-C1-O1-S1
3	E	2	YYJ	C4-C3-O3-S3

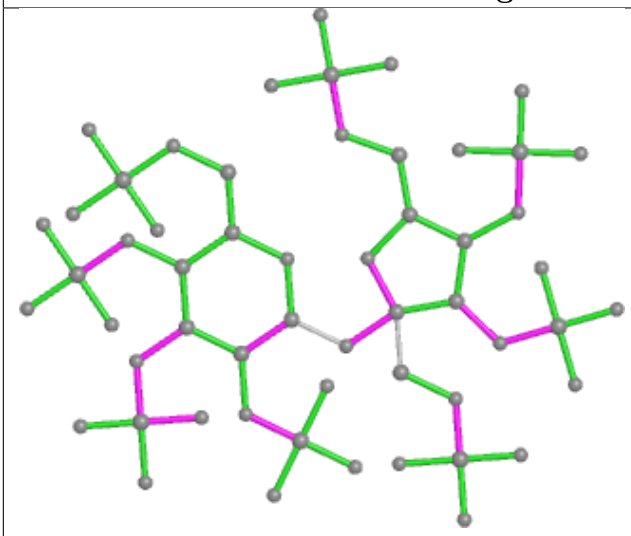
There are no ring outliers.

4 monomers are involved in 20 short contacts:

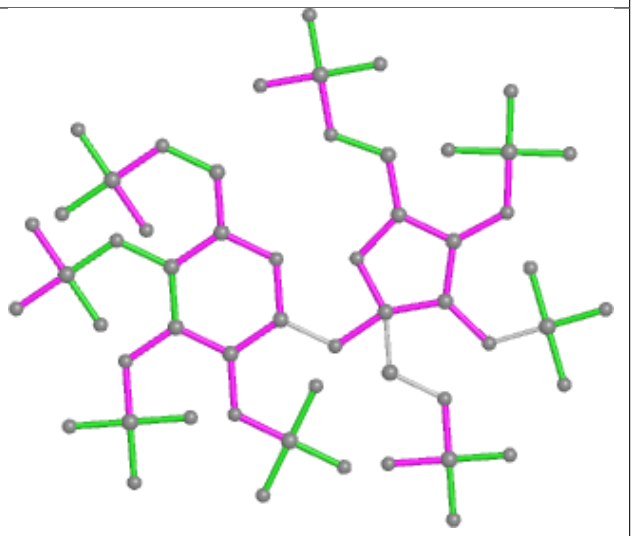
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	2	YYJ	3	0
3	E	1	GU4	3	0
3	F	1	GU4	11	0
3	F	2	YYJ	5	0

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

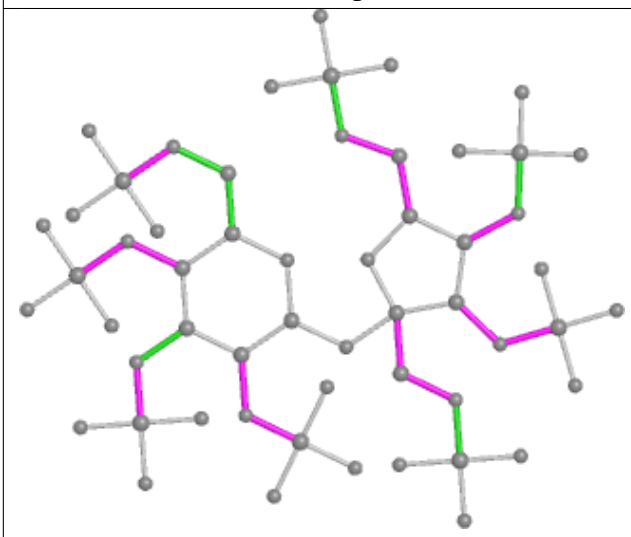
Oligosaccharide Chain E



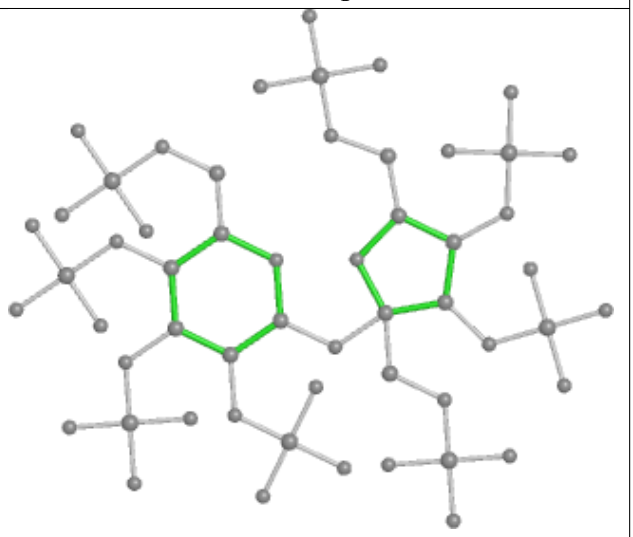
Bond lengths



Bond angles

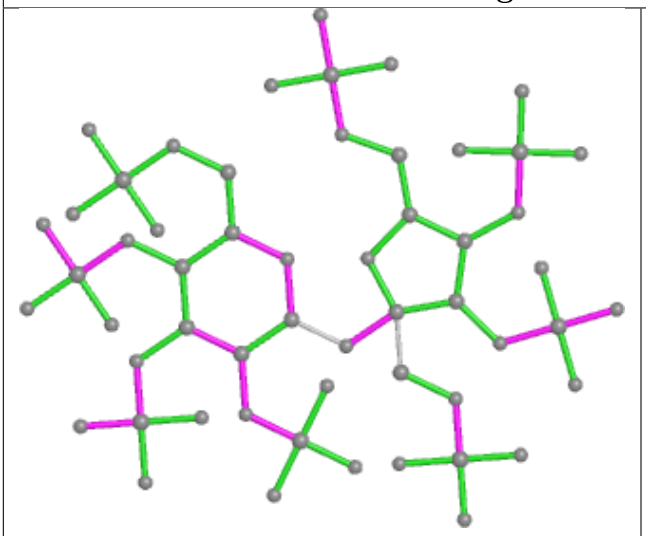


Torsions

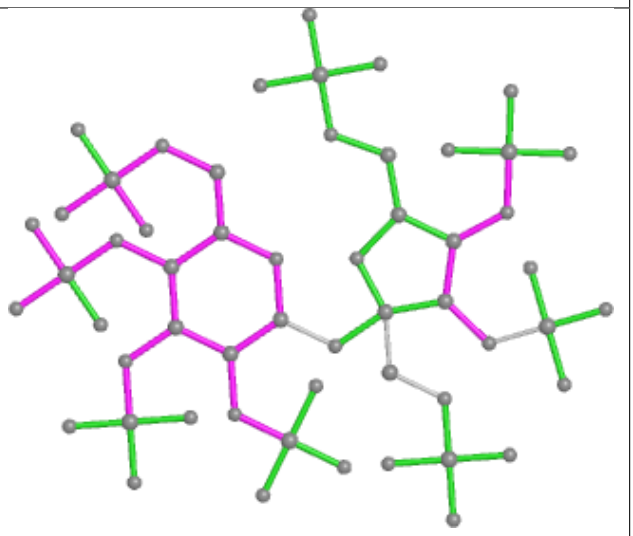


Rings

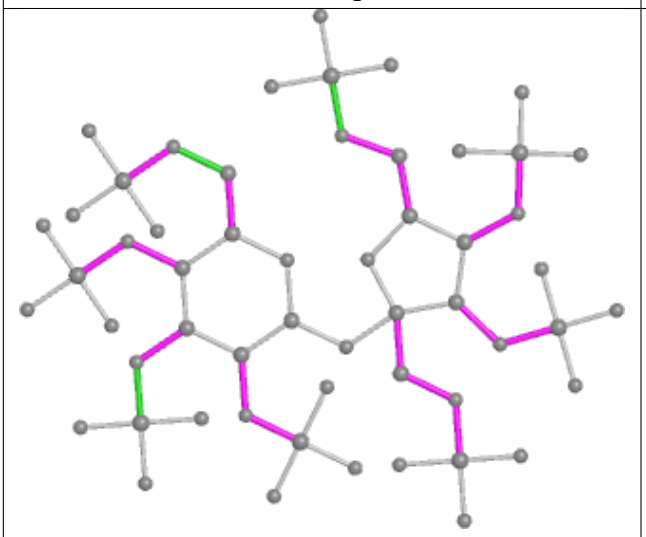
Oligosaccharide Chain F



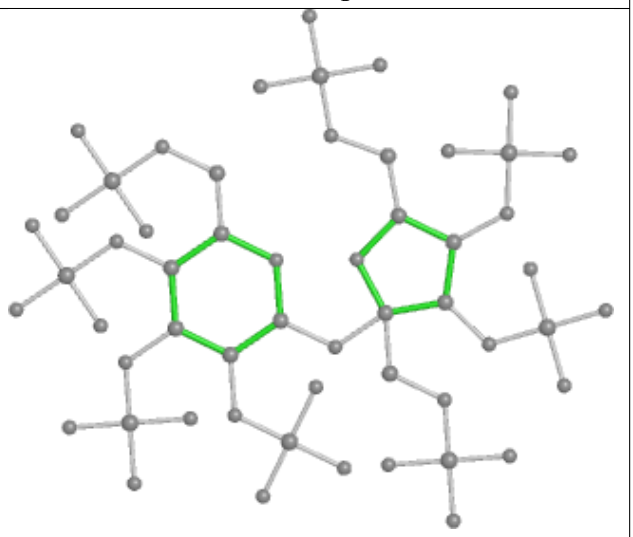
Bond lengths



Bond angles



Torsions



Rings

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

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.

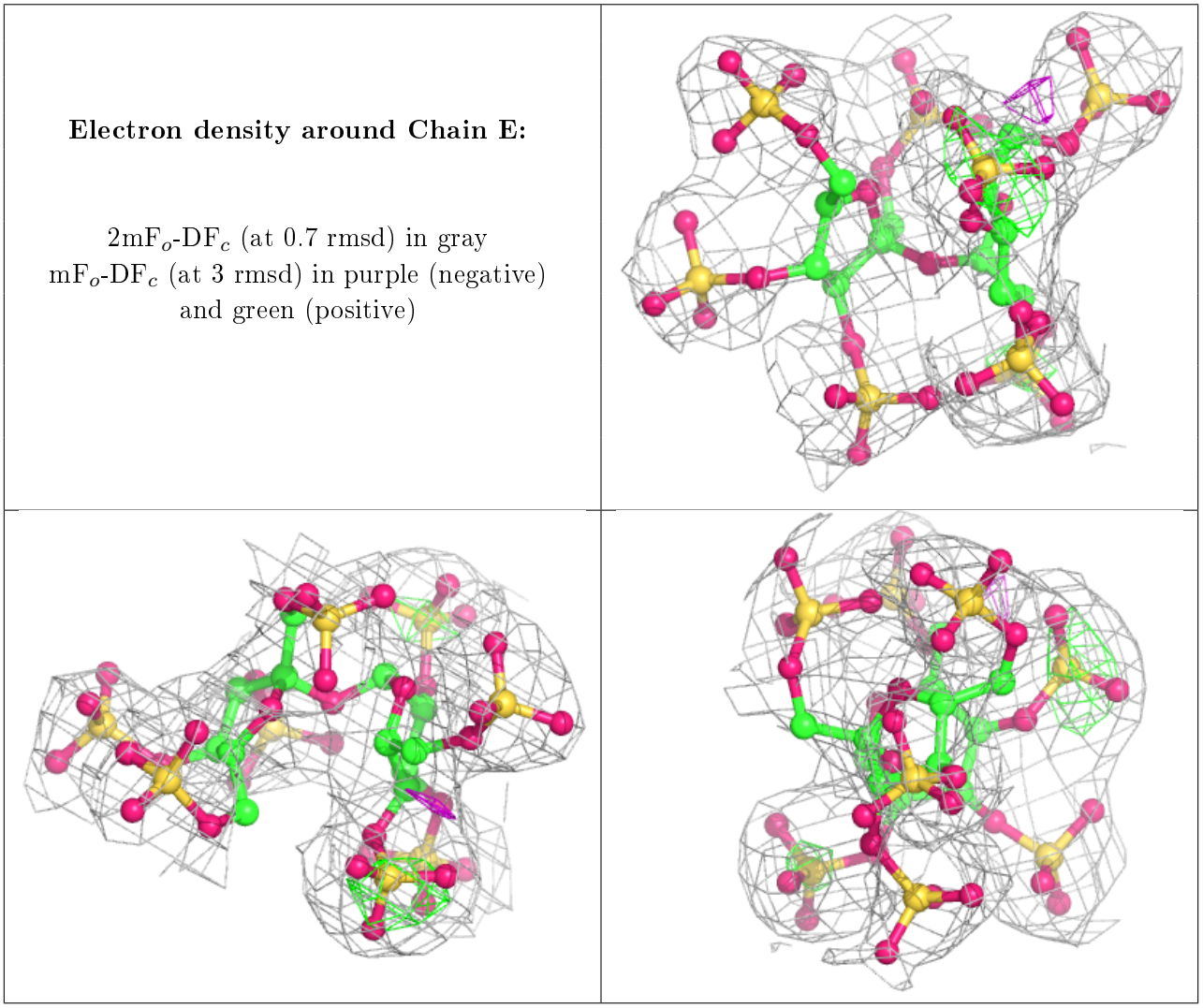
6.3 Carbohydrates

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.

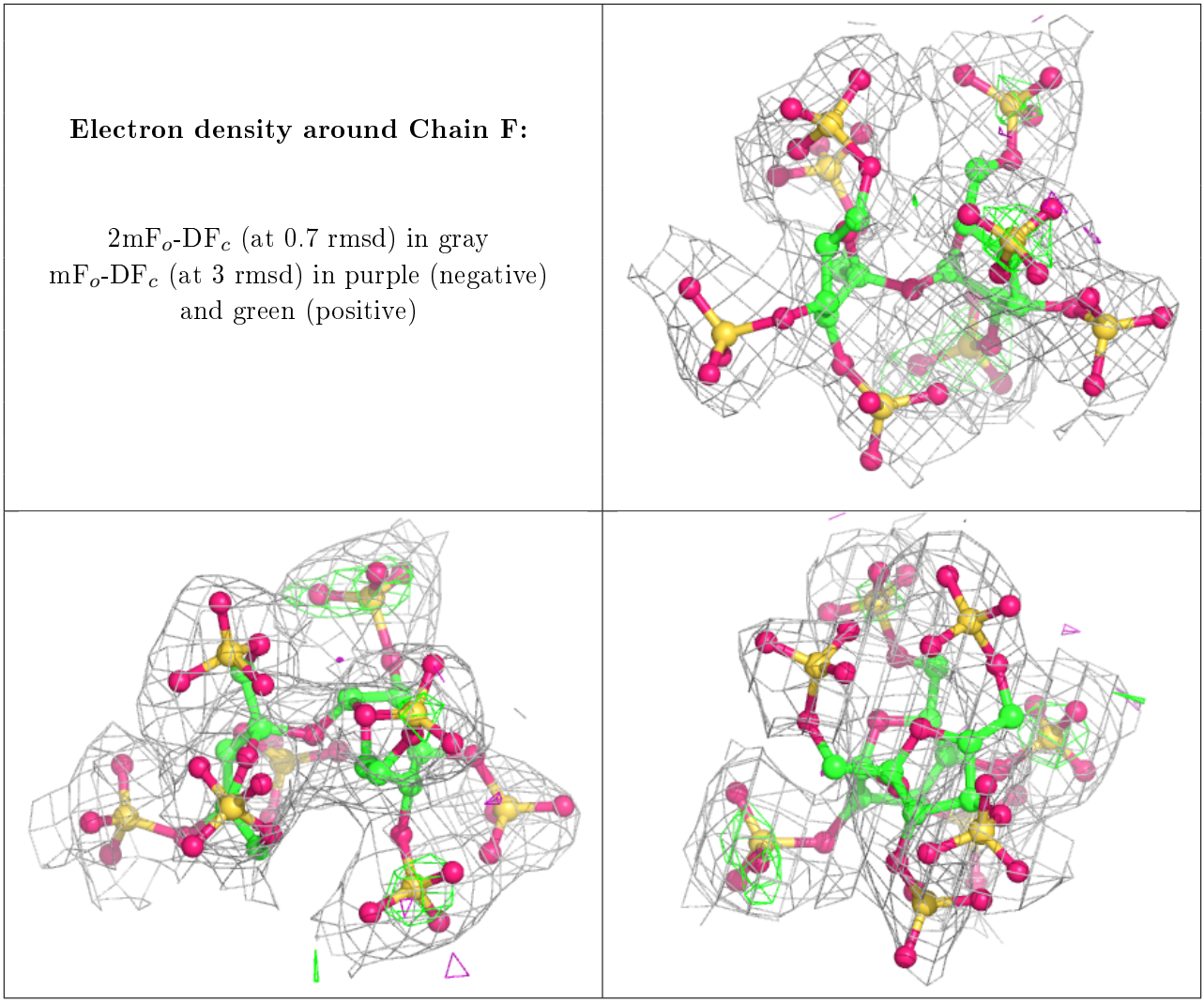
Electron density around Chain E:

$2mF_o$-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 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

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.