Full wwPDB X-ray Structure Validation Report ⓘ

Dec 2, 2020 – 06:02 PM EST

PDB ID	:	6USM
Title	:	Structure of nuclease domain of human parvovirus B19 non-structural protein 1 in complex with zinc
Authors	:	Tewary, S.K.
Deposited on	:	2019-10-28
Resolution	:	3.37 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

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

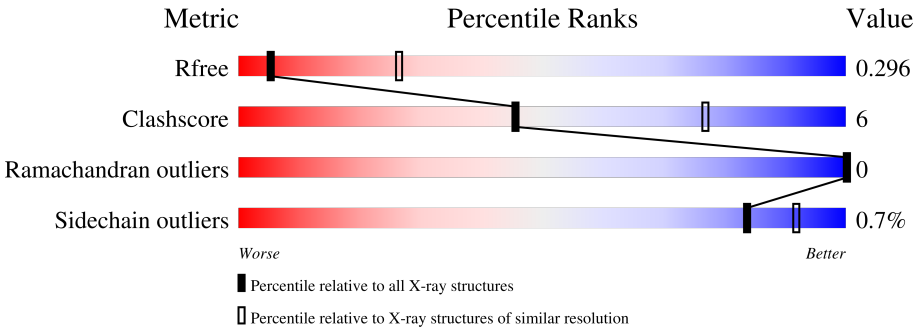
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 3.37 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	1691 (3.46-3.30)
Clashscore	141614	1762 (3.46-3.30)
Ramachandran outliers	138981	1732 (3.46-3.30)
Sidechain outliers	138945	1731 (3.46-3.30)

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	371	85% 14% .
2	B	171	77% 19% .
3	C	2	50% 50%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4190 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 Maltodextrin-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	368	2858	1844	464	543	7	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP C3SHQ8
A	139	GLY	GLU	conflict	UNP C3SHQ8
A	313	VAL	ALA	conflict	UNP C3SHQ8
A	368	VAL	-	expression tag	UNP C3SHQ8
A	369	PRO	-	expression tag	UNP C3SHQ8
A	370	HIS	-	expression tag	UNP C3SHQ8
A	371	MET	-	expression tag	UNP C3SHQ8

- Molecule 2 is a protein called Non-structural protein NS1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	164	1308	840	218	240	10	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	71	VAL	ALA	conflict	UNP Q6TV13

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	C	2	Total C O 23 12 11	0	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by author).

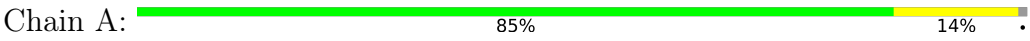
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Zn 1 1	0	0

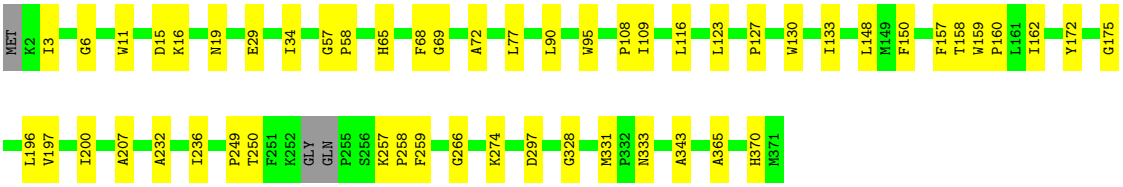
3 Residue-property plots (i)

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: Maltodextrin-binding protein

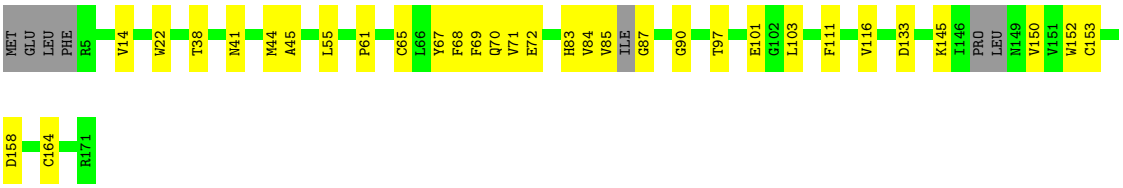
Chain A:





• Molecule 2: Non-structural protein NS1

Chain B:



• Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	102.69Å 102.69Å 134.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.00 – 3.37 47.96 – 3.37	Depositor EDS
% Data completeness (in resolution range)	96.8 (47.00-3.37) 95.6 (47.96-3.37)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.40 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	(Not available) , (Not available) 0.248 , 0.296	Depositor DCC
R_{free} test set	1148 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	81.2	Xtriage
Anisotropy	0.754	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 26.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.067 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4190	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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: ZN, GLC

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	0.48	0/2928	0.68	0/3974
2	B	0.51	0/1338	0.68	0/1815
All	All	0.49	0/4266	0.68	0/5789

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	A	2858	0	2834	34	0
2	B	1308	0	1276	19	0
3	C	23	0	21	1	0
4	B	1	0	0	0	0
All	All	4190	0	4131	53	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.

All (53) 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:90:LEU:HD23	1:A:108:PRO:HG2	1.70	0.73
1:A:116:LEU:HD22	1:A:249:PRO:HD3	1.75	0.69
2:B:55:LEU:HD23	2:B:67:TYR:CD2	2.33	0.63
1:A:157:PHE:O	1:A:160:PRO:HD2	1.99	0.63
1:A:130:TRP:HD1	1:A:250:THR:O	1.83	0.62
1:A:123:LEU:HD21	1:A:127:PRO:HD3	1.81	0.61
1:A:160:PRO:HA	1:A:257:LYS:O	2.03	0.59
2:B:14:VAL:HG22	2:B:116:VAL:HG11	1.85	0.58
2:B:70:GLN:O	2:B:83:HIS:N	2.37	0.55
1:A:15:ASP:OD1	1:A:16:LYS:HD2	2.09	0.53
1:A:90:LEU:HD22	1:A:95:TRP:CZ2	2.44	0.53
1:A:197:VAL:HA	1:A:200:ILE:HD12	1.90	0.53
1:A:130:TRP:HA	1:A:133:ILE:HD12	1.93	0.50
2:B:133:ASP:OD1	2:B:133:ASP:N	2.45	0.50
2:B:55:LEU:HD13	2:B:103:LEU:HD22	1.95	0.49
1:A:232:ALA:O	1:A:236:ILE:HG13	2.12	0.49
1:A:72:ALA:HB2	1:A:77:LEU:HD12	1.96	0.48
2:B:69:PHE:CD1	2:B:84:VAL:HG22	2.49	0.47
1:A:343:ALA:HB1	1:A:365:ALA:HB1	1.95	0.47
1:A:3:ILE:HG13	1:A:57:GLY:O	2.13	0.47
1:A:19:ASN:HB2	1:A:297:ASP:OD2	2.14	0.47
1:A:130:TRP:CD1	1:A:249:PRO:HB2	2.50	0.47
1:A:6:GLY:O	1:A:34:ILE:HG23	2.15	0.47
1:A:160:PRO:HG3	1:A:258:PRO:HA	1.97	0.46
2:B:22:TRP:CE3	2:B:111:PHE:HB3	2.50	0.46
1:A:200:ILE:HG21	1:A:207:ALA:HB2	1.98	0.46
2:B:72:GLU:OE2	2:B:145:LYS:HD2	2.16	0.46
1:A:68:PHE:HE2	1:A:266:GLY:HA3	1.82	0.45
2:B:44:MET:HB3	2:B:71:VAL:HG11	1.98	0.45
1:A:108:PRO:O	1:A:109:ILE:HD13	2.17	0.45
1:A:11:TRP:CD2	1:A:58:PRO:HG3	2.52	0.45
1:A:148:LEU:HD21	1:A:150:PHE:CE1	2.52	0.45
2:B:68:PHE:HB3	2:B:85:VAL:HB	1.99	0.45
2:B:38:THR:O	2:B:41:ASN:N	2.50	0.45
1:A:159:TRP:O	1:A:162:ILE:N	2.50	0.44
2:B:22:TRP:HD1	2:B:22:TRP:O	2.01	0.44
1:A:16:LYS:NZ	3:C:1:GLC:O2	2.51	0.43
1:A:130:TRP:CD1	1:A:250:THR:O	2.67	0.43
2:B:97:THR:O	2:B:101:GLU:HG3	2.19	0.43
1:A:65:HIS:CE1	1:A:331:MET:HB2	2.54	0.43
2:B:45:ALA:HA	2:B:152:TRP:CD1	2.55	0.42
2:B:153:CYS:SG	2:B:164:CYS:HB2	2.59	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:61:PRO:HB3	2:B:90:GLY:O	2.20	0.42
1:A:258:PRO:O	1:A:328:GLY:HA3	2.19	0.42
1:A:158:THR:HG23	1:A:196:LEU:HD22	2.01	0.42
1:A:172:TYR:HE1	1:A:175:GLY:C	2.24	0.41
1:A:257:LYS:HB3	1:A:328:GLY:HA2	2.03	0.41
1:A:29:GLU:HA	1:A:34:ILE:O	2.20	0.41
2:B:65:CYS:HB2	2:B:87:GLY:O	2.20	0.41
1:A:274:LYS:HB2	1:A:274:LYS:HE2	1.85	0.41
2:B:38:THR:H	2:B:41:ASN:HB2	1.85	0.40
1:A:69:GLY:HA3	1:A:333:ASN:O	2.21	0.40
2:B:150:VAL:HG13	2:B:150:VAL:O	2.21	0.40

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	A	364/371 (98%)	346 (95%)	18 (5%)	0	100	100
2	B	158/171 (92%)	149 (94%)	9 (6%)	0	100	100
All	All	522/542 (96%)	495 (95%)	27 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	A	296/298 (99%)	294 (99%)	2 (1%)	84	92
2	B	146/153 (95%)	145 (99%)	1 (1%)	84	92
All	All	442/451 (98%)	439 (99%)	3 (1%)	84	92

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

Mol	Chain	Res	Type
1	A	259	PHE
1	A	370	HIS
2	B	158	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	$\#\|Z\| > 2$
3	GLC	C	1	3	12,12,12	1.65	1 (8%)	17,17,17	1.95	5 (29%)
3	GLC	C	2	3	11,11,12	1.84	2 (18%)	15,15,17	1.52	3 (20%)

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	GLC	C	1	3	-	2/2/22/22	0/1/1/1
3	GLC	C	2	3	-	2/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	GLC	C2-C3	5.17	1.60	1.52
3	C	1	GLC	C1-C2	4.58	1.63	1.52
3	C	2	GLC	C4-C5	2.07	1.57	1.53

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	1	GLC	O5-C1-C2	3.92	117.27	110.28
3	C	2	GLC	O5-C5-C6	3.81	113.17	107.20
3	C	1	GLC	C1-C2-C3	3.01	116.57	110.31
3	C	1	GLC	C1-O5-C5	2.53	118.44	113.66
3	C	1	GLC	C6-C5-C4	-2.47	107.23	113.00
3	C	1	GLC	O5-C5-C4	2.29	113.85	109.69
3	C	2	GLC	O5-C1-C2	-2.17	107.42	110.77
3	C	2	GLC	O3-C3-C4	-2.10	105.50	110.35

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2	GLC	C4-C5-C6-O6
3	C	1	GLC	C4-C5-C6-O6
3	C	2	GLC	O5-C5-C6-O6
3	C	1	GLC	O5-C5-C6-O6

There are no ring outliers.

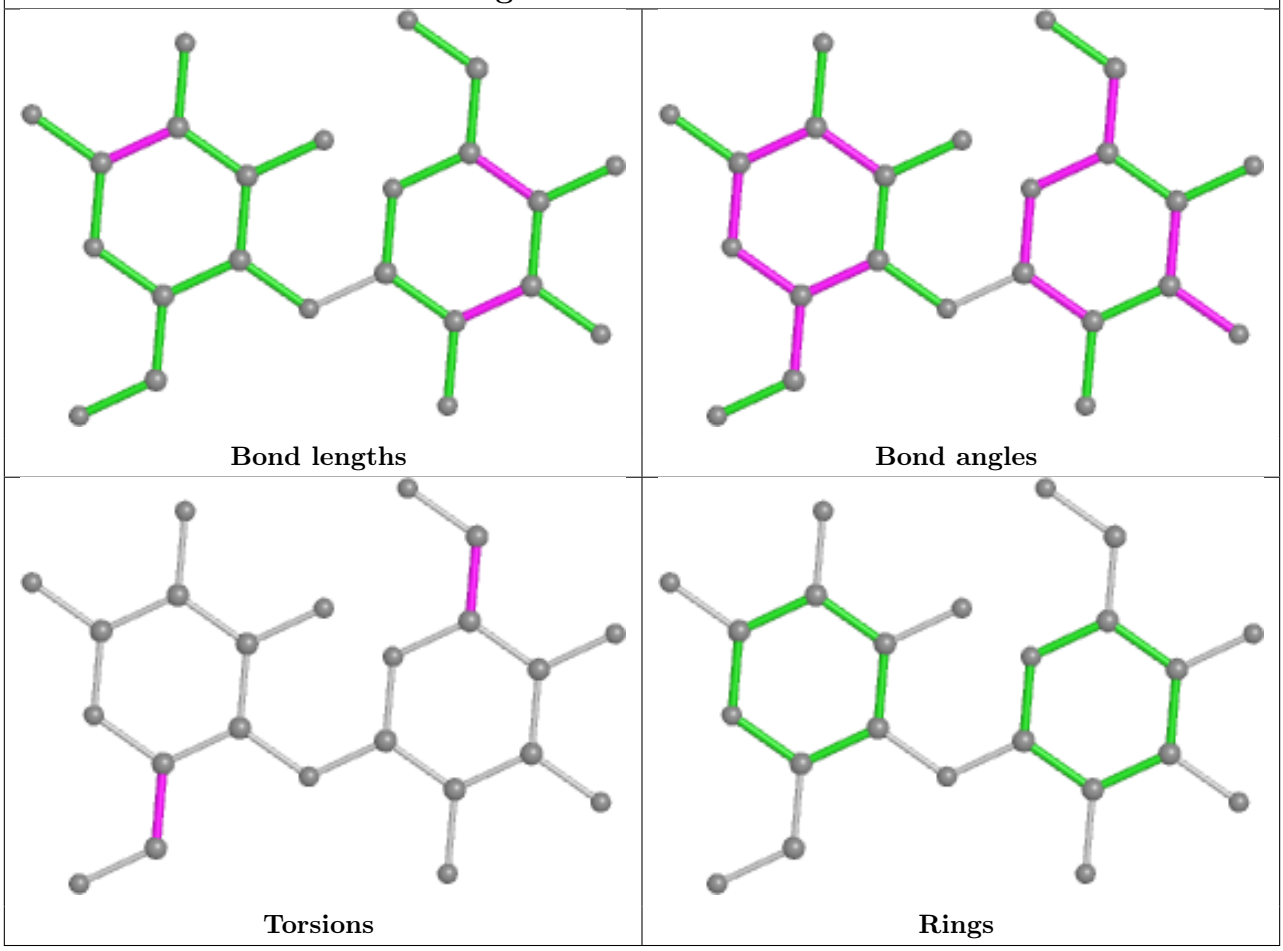
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	GLC	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 C



5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

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.

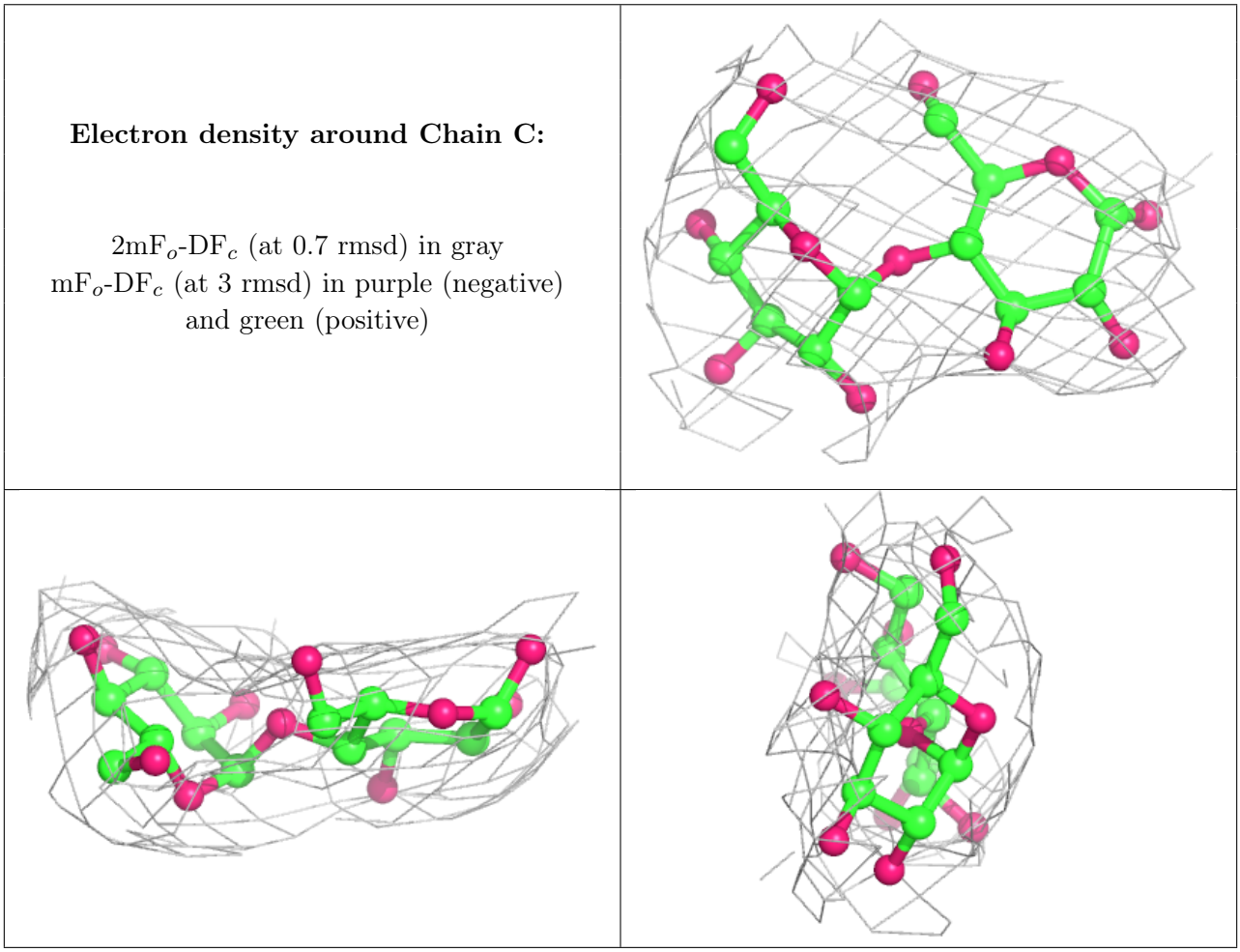
6.3 Carbohydrates

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 C:

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



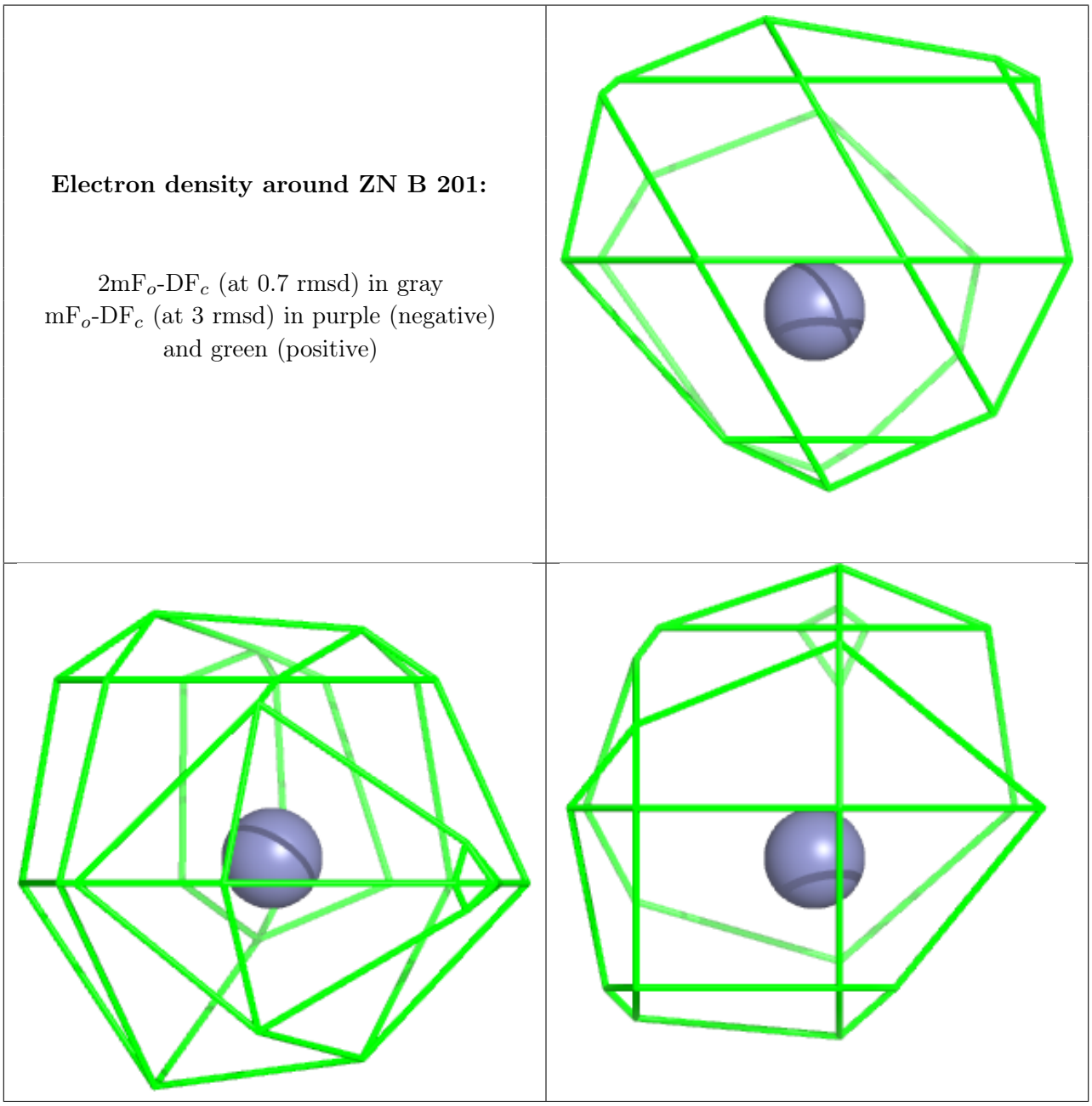
6.4 Ligands

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

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 ZN B 201:

$2mF_o$-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

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