



# wwPDB X-ray Structure Validation Summary Report

Jan 3, 2024 – 01:19 pm GMT

PDB ID : 5K33  
Title : Crystal structure of extracellular domain of HER2 in complex with Fcab STAB19  
Authors : Humm, A.; Lobner, E.; Goritzer, K.; Mlynek, G.; Obinger, C.; Djinovic-Carugo, K.  
Deposited on : 2016-05-19  
Resolution : 3.30 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (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

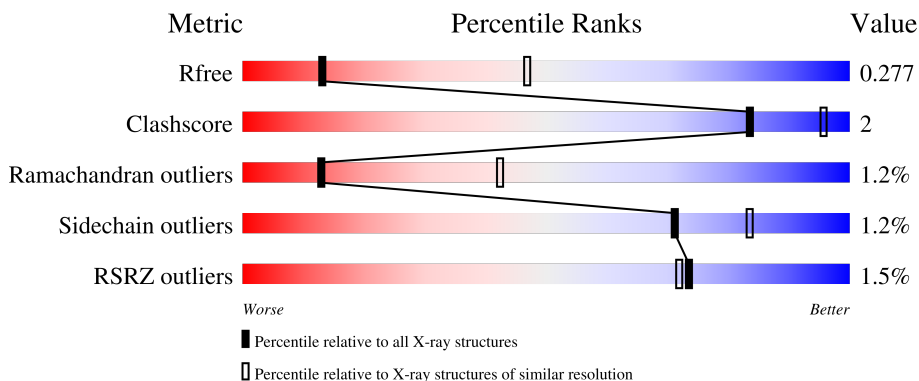
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	A	227	
2	C	607	
3	B	5	
4	D	4	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12176 atoms, of which 5992 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 Ig gamma-1 chain C region.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	213	3375	1085	1669	289	325	7	0	0	0

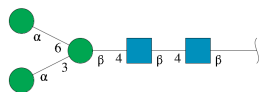
There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	358	TYR	LEU	engineered mutation	UNP P01857
A	359	LEU	THR	engineered mutation	UNP P01857
A	360	SER	LYS	engineered mutation	UNP P01857
A	361	ASP	ASN	engineered mutation	UNP P01857
A	362	SER	GLN	engineered mutation	UNP P01857
A	413	PRO	-	insertion	UNP P01857
A	414	ARG	-	insertion	UNP P01857
A	415	HIS	-	insertion	UNP P01857
A	415A	SER	-	insertion	UNP P01857
A	415B	GLU	-	insertion	UNP P01857
A	415C	THR	ASP	engineered mutation	UNP P01857
A	415D	MET	LYS	engineered mutation	UNP P01857
A	415E	ARG	SER	engineered mutation	UNP P01857
A	418	ALA	GLN	engineered mutation	UNP P01857
A	419	HIS	GLN	engineered mutation	UNP P01857

- Molecule 2 is a protein called Receptor tyrosine-protein kinase erbB-2.

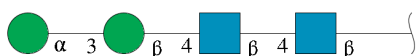
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	C	563	8516	2694	4183	779	811	49	0	0	0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



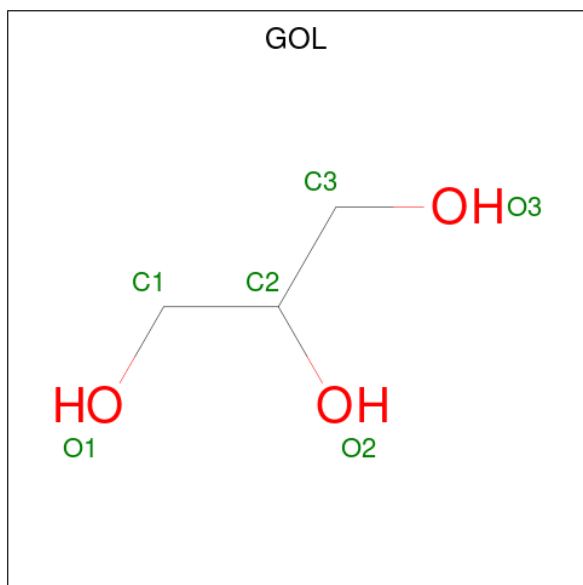
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
3	B	5	118	34	57	2	25	0	0	0

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
4	D	4	97	28	47	2	20	0	0	0

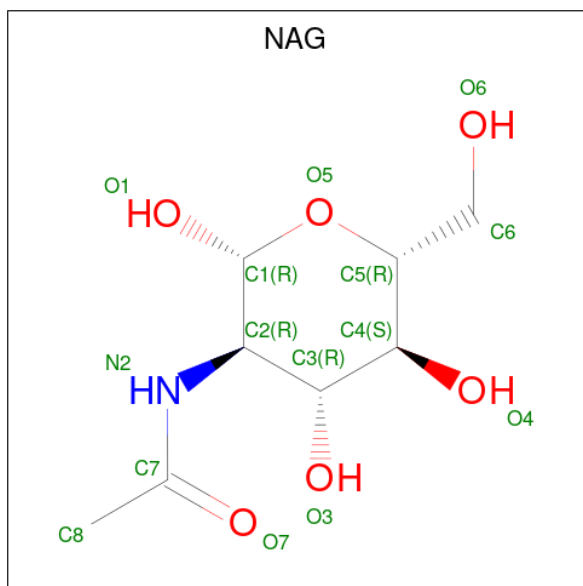
- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
5	A	1	14	3	8	3	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:

C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).




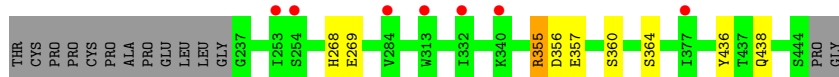
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			
6	C	1	Total	28	8	14	1	5	0	0
6	C	1	Total	28	8	14	1	5	0	0

### 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 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.

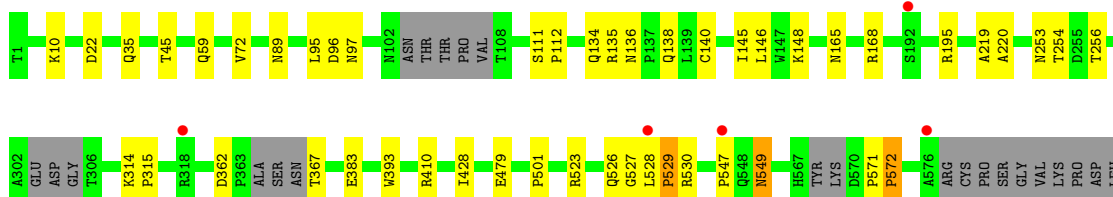
- Molecule 1: Ig gamma-1 chain C region

Chain A: 



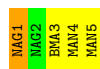
- Molecule 2: Receptor tyrosine-protein kinase erbB-2

Chain C: 



- Molecule 3: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 



- Molecule 4: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D: 



## 4 Data and refinement statistics

Property	Value
Space group	P 32 2 1
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.62Å 108.62Å 174.34Å 90.00° 90.00° 120.00°
Resolution (Å)	47.03 – 3.30 47.03 – 3.00
% Data completeness (in resolution range)	99.9 (47.03-3.30) 99.9 (47.03-3.00)
$R_{merge}$	0.14
$R_{sym}$	(Not available)
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 3.01Å)
Refinement program	BUSTER Snapshot v.20150616, REFMAC CCP4 v7.0.0, PHENIX 1.10-
R, $R_{free}$	0.217 , 0.271 0.228 , 0.277
$R_{free}$ test set	2017 reflections (8.26%)
Wilson B-factor (Å <sup>2</sup> )	107.7
Anisotropy	0.114
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 69.7
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$
Estimated twinning fraction	0.046 for -h,-k,l
$F_o, F_c$ correlation	0.94
Total number of atoms	12176
Average B, all atoms (Å <sup>2</sup> )	132.0

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: MAN, BMA, NAG, GOL

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	0.25	0/1756	0.42	0/2394
2	C	0.26	0/4430	0.47	0/6023
All	All	0.26	0/6186	0.45	0/8417

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	1706	1669	1660	4	0
2	C	4333	4183	4152	22	0
3	B	61	57	52	2	0
4	D	50	47	43	1	0
5	A	6	8	8	0	0
6	C	28	28	26	0	0
All	All	6184	5992	5941	29	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 2.

The worst 5 of 29 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:355:ARG:NH1	1:A:356:ASP:OD1	2.17	0.78
2:C:134:GLN:O	2:C:136:ASN:ND2	2.17	0.77
2:C:95:LEU:O	2:C:97:ASN:ND2	2.25	0.70
3:B:1:NAG:O7	3:B:1:NAG:O3	2.15	0.64
2:C:383:GLU:OE2	2:C:410:ARG:NE	2.33	0.61

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	211/227 (93%)	202 (96%)	8 (4%)	1 (0%)	29	61
2	C	552/607 (91%)	497 (90%)	47 (8%)	8 (1%)	11	38
All	All	763/834 (92%)	699 (92%)	55 (7%)	9 (1%)	13	42

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	549	ASN
2	C	10	LYS
1	A	360	SER
2	C	572	PRO
2	C	89	ASN

### 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	198/209 (95%)	196 (99%)	2 (1%)	76	86
2	C	484/523 (92%)	478 (99%)	6 (1%)	71	83
All	All	682/732 (93%)	674 (99%)	8 (1%)	71	83

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	479	GLU
2	C	165	ASN
2	C	72	VAL
2	C	45	THR
2	C	135	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	66	HIS
2	C	296	HIS
2	C	511	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](#)

9 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	NAG	B	1	1,3	14,14,15	0.20	0	17,19,21	0.76	1 (5%)
3	NAG	B	2	3	14,14,15	0.22	0	17,19,21	0.39	0
3	BMA	B	3	3	11,11,12	1.02	1 (9%)	15,15,17	1.05	1 (6%)
3	MAN	B	4	3	11,11,12	0.95	1 (9%)	15,15,17	1.35	3 (20%)
3	MAN	B	5	3	11,11,12	1.09	2 (18%)	15,15,17	1.63	2 (13%)
4	NAG	D	1	2,4	14,14,15	0.36	0	17,19,21	1.14	2 (11%)
4	NAG	D	2	4	14,14,15	0.36	0	17,19,21	0.53	0
4	BMA	D	3	4	11,11,12	1.00	0	15,15,17	0.90	1 (6%)
4	MAN	D	4	4	11,11,12	0.69	1 (9%)	15,15,17	1.03	2 (13%)

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	NAG	B	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	B	2	3	-	2/6/23/26	0/1/1/1
3	BMA	B	3	3	-	2/2/19/22	0/1/1/1
3	MAN	B	4	3	-	0/2/19/22	0/1/1/1
3	MAN	B	5	3	-	0/2/19/22	0/1/1/1
4	NAG	D	1	2,4	-	5/6/23/26	0/1/1/1
4	NAG	D	2	4	-	3/6/23/26	0/1/1/1
4	BMA	D	3	4	-	2/2/19/22	0/1/1/1
4	MAN	D	4	4	-	0/2/19/22	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	4	MAN	C1-C2	2.98	1.59	1.52
3	B	5	MAN	C1-C2	2.60	1.58	1.52
3	B	3	BMA	C1-C2	2.14	1.57	1.52
3	B	5	MAN	O5-C1	2.01	1.46	1.43
4	D	4	MAN	C1-C2	2.00	1.56	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	5	MAN	C1-O5-C5	5.05	119.03	112.19
4	D	1	NAG	C2-N2-C7	3.18	127.42	122.90
3	B	4	MAN	C1-C2-C3	2.74	113.04	109.67
3	B	4	MAN	O2-C2-C3	-2.58	104.97	110.14
4	D	4	MAN	C1-O5-C5	2.57	115.67	112.19

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

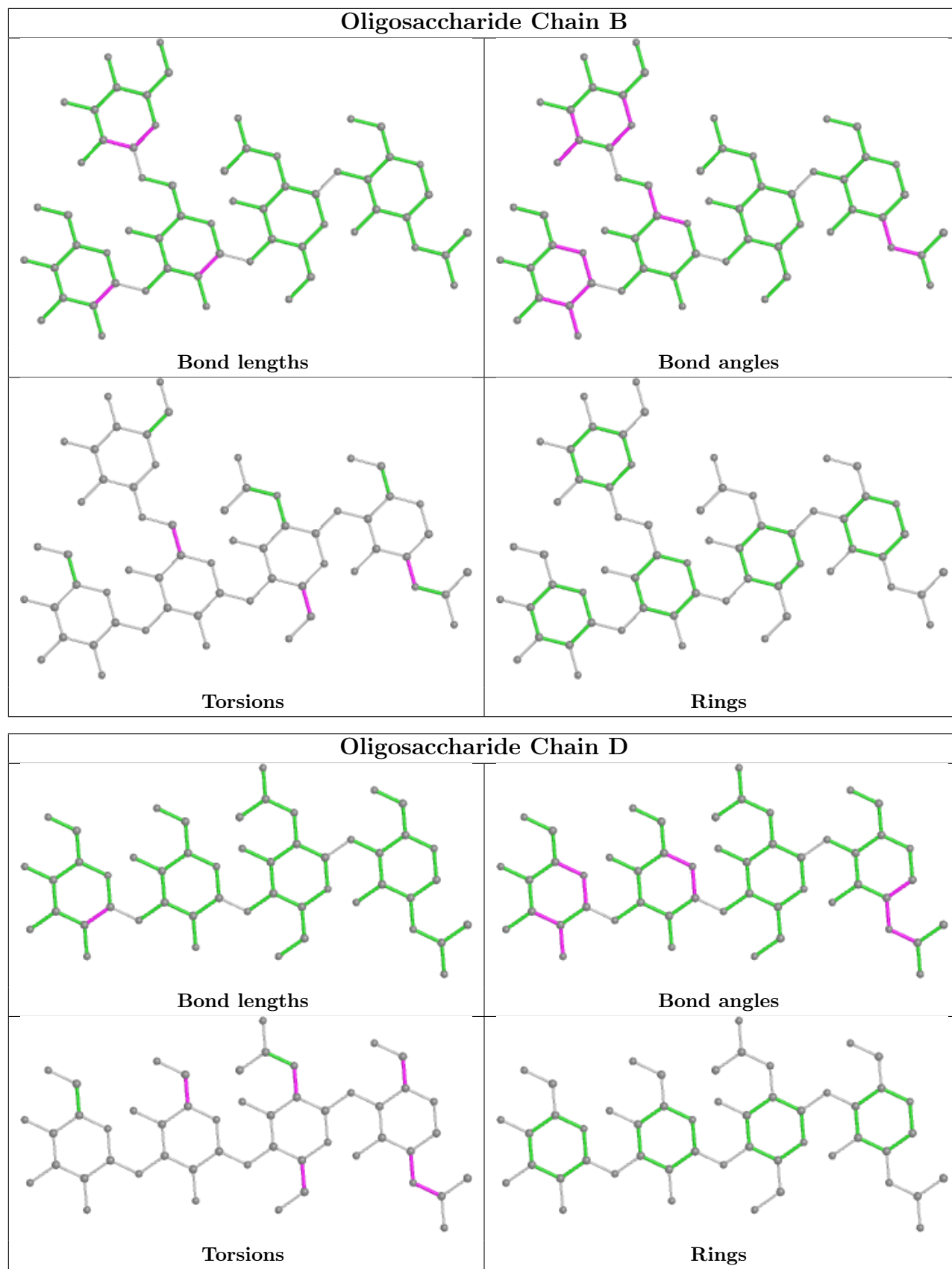
Mol	Chain	Res	Type	Atoms
3	B	1	NAG	C3-C2-N2-C7
3	B	3	BMA	O5-C5-C6-O6
3	B	3	BMA	C4-C5-C6-O6
4	D	1	NAG	C8-C7-N2-C2
4	D	1	NAG	O7-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	1	NAG	1	0
3	B	1	NAG	2	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.



## 5.6 Ligand geometry

3 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
6	NAG	C	705	2	14,14,15	0.42	0	17,19,21	0.54	0
5	GOL	A	506	-	5,5,5	0.37	0	5,5,5	0.18	0
6	NAG	C	706	2	14,14,15	0.39	0	17,19,21	0.57	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	C	705	2	-	1/6/23/26	0/1/1/1
5	GOL	A	506	-	-	0/4/4/4	-
6	NAG	C	706	2	-	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 (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	706	NAG	O5-C5-C6-O6
6	C	706	NAG	C4-C5-C6-O6
6	C	705	NAG	O5-C5-C6-O6

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	575:VAL	C	576:ALA	N	4.08

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	213/227 (93%)	0.23	7 (3%) 46 44	102, 138, 173, 206	0
2	C	563/607 (92%)	0.06	5 (0%) 84 84	65, 107, 164, 200	0
All	All	776/834 (93%)	0.10	12 (1%) 73 72	65, 119, 170, 206	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	340	LYS	3.7
1	A	332	ILE	3.7
2	C	576	ALA	2.9
1	A	377	ILE	2.9
2	C	528	LEU	2.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	MAN	B	4	11/12	0.67	0.22	140,151,181,182	0
4	BMA	D	3	11/12	0.73	0.22	164,172,202,207	0
3	MAN	B	5	11/12	0.78	0.18	142,147,176,176	0
3	NAG	B	1	14/15	0.78	0.20	110,125,147,164	0
4	MAN	D	4	11/12	0.78	0.25	180,181,217,218	0

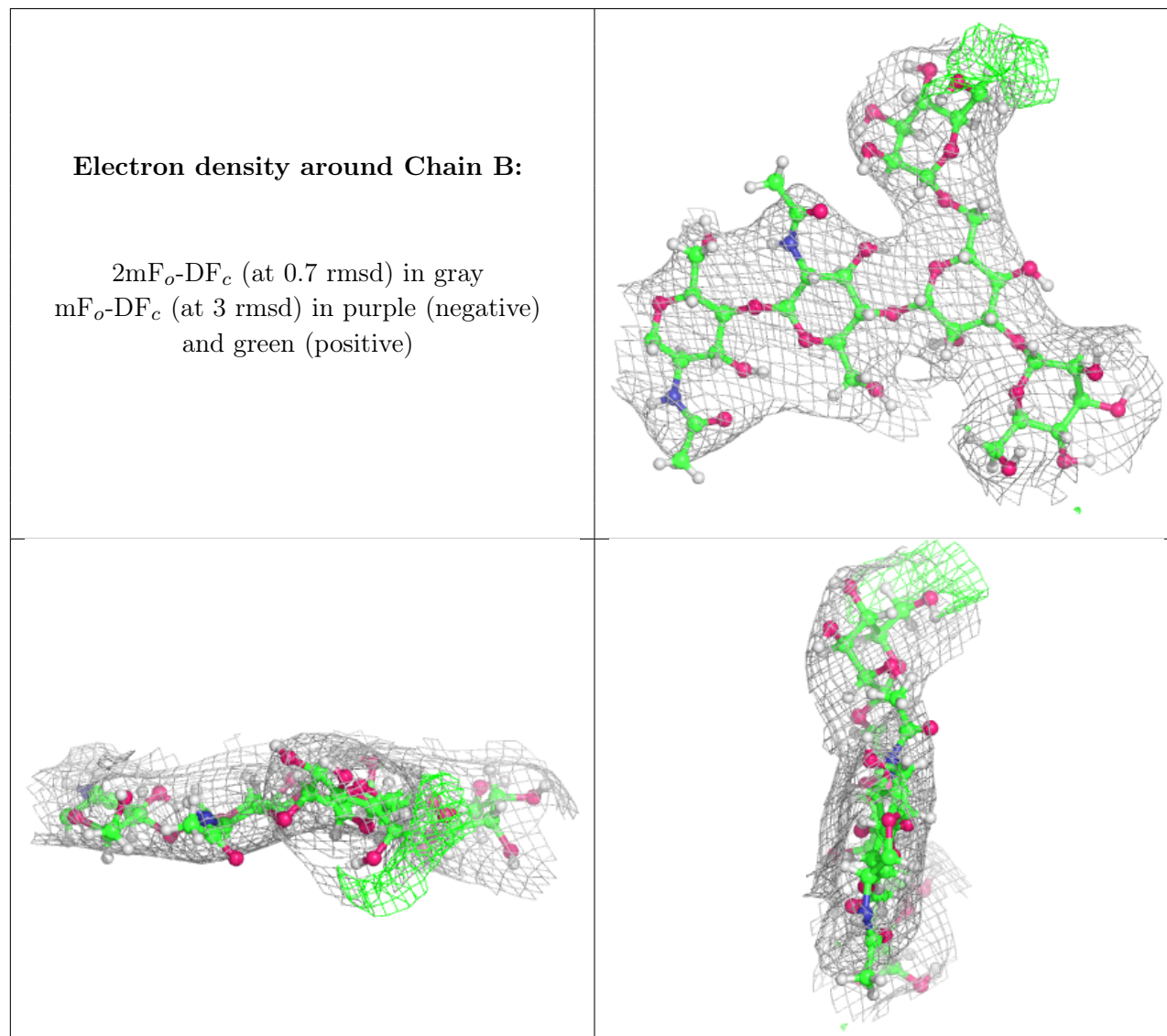
*Continued on next page...*

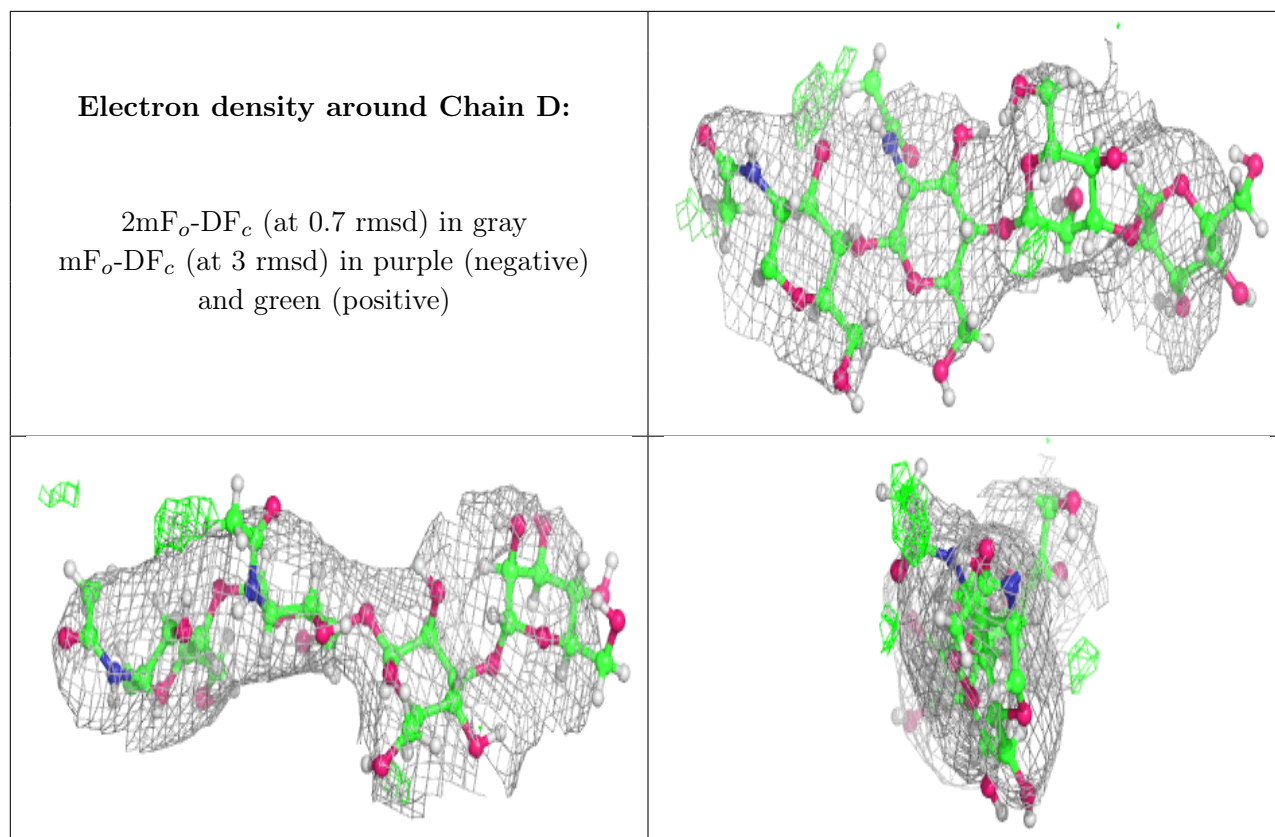


Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	2	14/15	0.83	0.23	120,124,149,149	0
3	BMA	B	3	11/12	0.83	0.18	122,132,156,156	0
4	NAG	D	2	14/15	0.91	0.21	133,153,179,179	0
4	NAG	D	1	14/15	0.94	0.15	95,109,123,131	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.





## 6.4 Ligands [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	GOL	A	506	6/6	0.61	0.23	144,172,178,180	0
6	NAG	C	705	14/15	0.83	0.17	127,150,175,175	0
6	NAG	C	706	14/15	0.88	0.22	144,171,205,205	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.