



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 16, 2022 – 12:23 PM EDT

PDB ID : 6BXC  
Title : Crystal structure of N-terminal fragment of Zebrafish Toll-Like Receptor 5 (TLR5) with Lamprey Variable Lymphocyte Receptor 9 (VLR9) bound  
Authors : Gunn, R.J.; Wilson, I.A.; Cooper, M.D.; Herrin, B.R.  
Deposited on : 2017-12-18  
Resolution : 2.50 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 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	:	2.27
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)	:	2.27

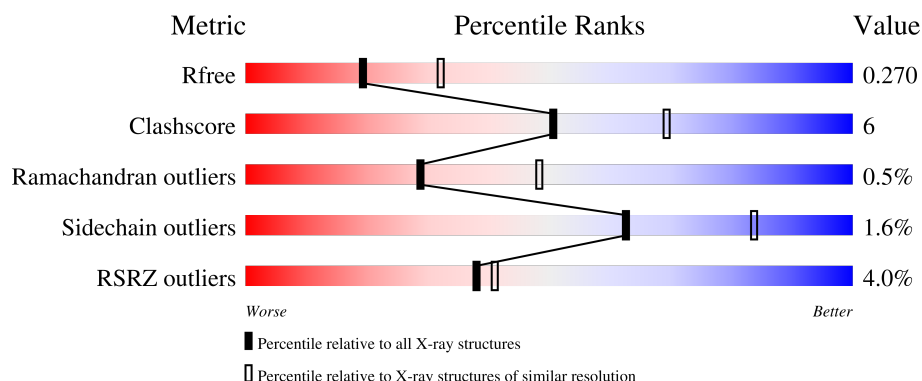
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	455	
1	B	455	
2	C	178	
2	D	178	
3	E	2	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9515 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 Toll-like receptor 5b, Variable lymphocyte receptor B chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	441	Total	C	N	O	S	0	0	0
			3421	2192	572	641	16			
1	B	441	Total	C	N	O	S	0	0	0
			3446	2206	576	648	16			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	ALA	-	expression tag	UNP F8W3J5
A	19	ASP	-	expression tag	UNP F8W3J5
A	20	PRO	-	expression tag	UNP F8W3J5
A	21	GLY	-	expression tag	UNP F8W3J5
A	466	SER	-	expression tag	UNP Q4G1L2
A	467	ALA	-	expression tag	UNP Q4G1L2
A	468	SER	-	expression tag	UNP Q4G1L2
A	469	LEU	-	expression tag	UNP Q4G1L2
A	470	VAL	-	expression tag	UNP Q4G1L2
A	471	PRO	-	expression tag	UNP Q4G1L2
A	472	ARG	-	expression tag	UNP Q4G1L2
B	18	ALA	-	expression tag	UNP F8W3J5
B	19	ASP	-	expression tag	UNP F8W3J5
B	20	PRO	-	expression tag	UNP F8W3J5
B	21	GLY	-	expression tag	UNP F8W3J5
B	466	SER	-	expression tag	UNP Q4G1L2
B	467	ALA	-	expression tag	UNP Q4G1L2
B	468	SER	-	expression tag	UNP Q4G1L2
B	469	LEU	-	expression tag	UNP Q4G1L2
B	470	VAL	-	expression tag	UNP Q4G1L2
B	471	PRO	-	expression tag	UNP Q4G1L2
B	472	ARG	-	expression tag	UNP Q4G1L2

- Molecule 2 is a protein called Variable Lymphocyte Receptor 9 (VLR9).

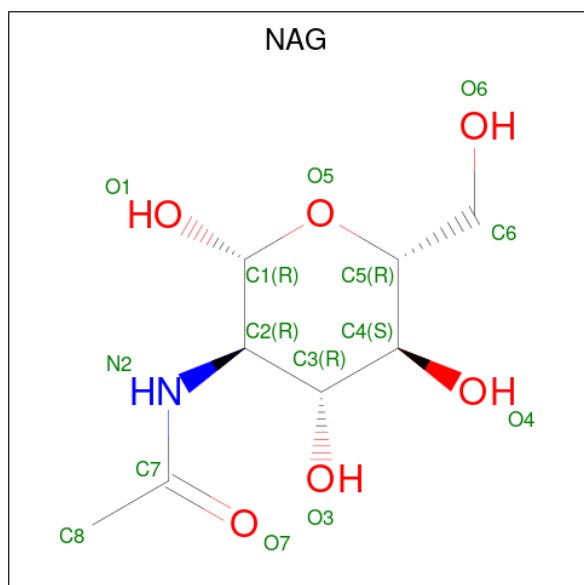
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	166	Total	C	N	O	S	0	0	0
			1225	776	201	239	9			
2	C	165	Total	C	N	O	S	0	0	0
			1228	774	205	240	9			

- Molecule 3 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
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	C	O	0	0
			4	2	2		

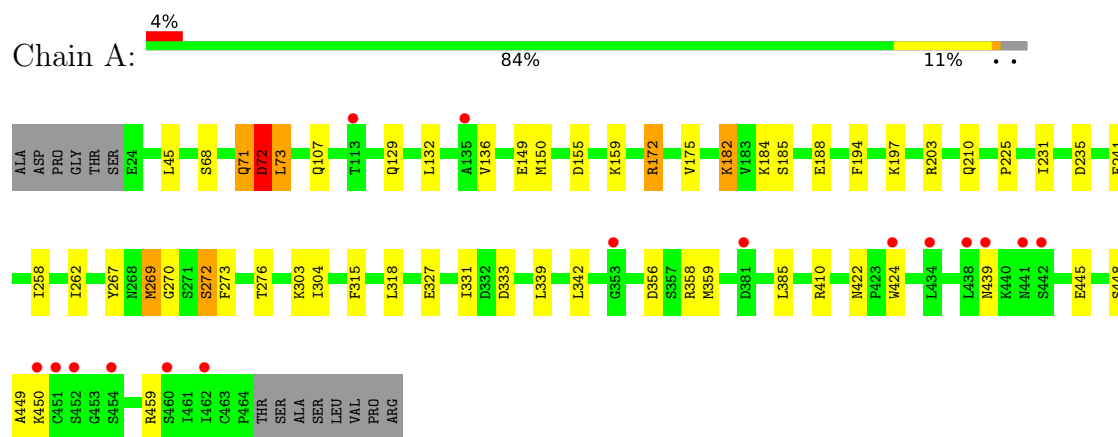
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	52	Total	O	0	0
			52	52		
6	B	44	Total	O	0	0
			44	44		
6	D	12	Total	O	0	0
			12	12		
6	C	13	Total	O	0	0
			13	13		

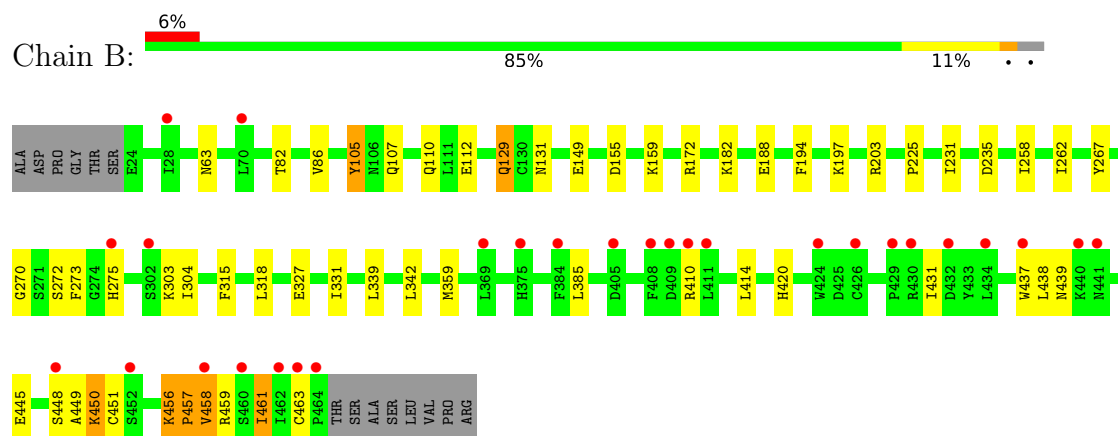
### 3 Residue-property plots [i](#)

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.

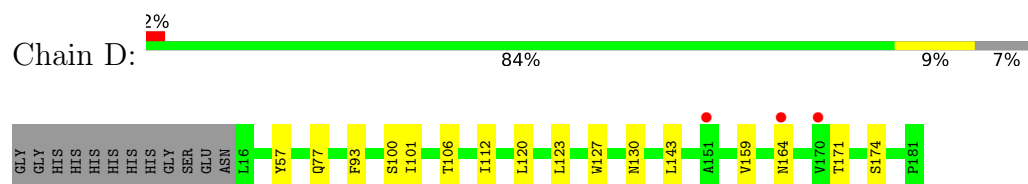
- Molecule 1: Toll-like receptor 5b, Variable lymphocyte receptor B chimera



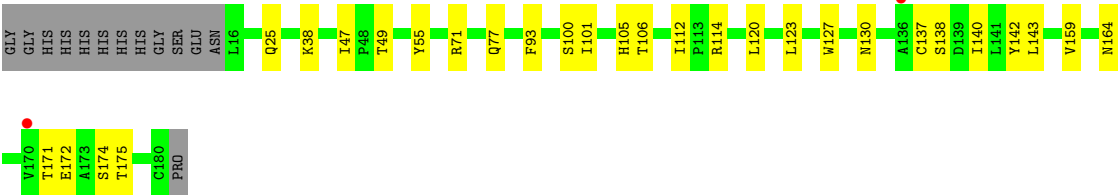
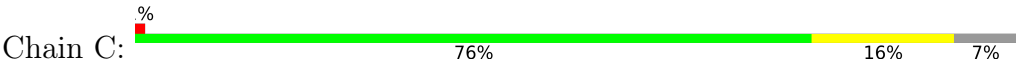
- Molecule 1: Toll-like receptor 5b, Variable lymphocyte receptor B chimera



- Molecule 2: Variable Lymphocyte Receptor 9 (VLR9)



- Molecule 2: Variable Lymphocyte Receptor 9 (VLR9)



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.38Å 69.41Å 84.56Å 112.14° 104.12° 100.81°	Depositor
Resolution (Å)	45.27 – 2.50 45.27 – 2.22	Depositor EDS
% Data completeness (in resolution range)	89.8 (45.27-2.50) 65.2 (45.27-2.22)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.35 (at 2.22Å)	Xtriage
Refinement program	PHENIX (1.12_2829)	Depositor
R, $R_{free}$	0.230 , 0.270 0.229 , 0.270	Depositor DCC
$R_{free}$ test set	2286 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.5	Xtriage
Anisotropy	0.411	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 37.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.179 for k,h,-h-k-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9515	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.85% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	A	0.31	0/3490	0.63	5/4731 (0.1%)
1	B	0.33	0/3515	0.63	2/4761 (0.0%)
2	C	0.35	0/1255	0.64	0/1719
2	D	0.37	0/1253	0.62	0/1718
All	All	0.33	0/9513	0.63	7/12929 (0.1%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	172	ARG	NE-CZ-NH2	-7.49	116.55	120.30
1	A	45	LEU	CB-CG-CD1	-7.10	98.93	111.00
1	B	105	TYR	CB-CG-CD2	6.82	125.09	121.00
1	B	105	TYR	CB-CG-CD1	-6.43	117.14	121.00
1	A	45	LEU	CA-CB-CG	6.02	129.15	115.30

There are no chirality outliers.

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	3421	0	3352	33	0
1	B	3446	0	3387	45	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1228	0	1177	23	0
2	D	1225	0	1173	14	0
3	E	28	0	25	0	0
4	A	28	0	26	0	0
4	B	14	0	13	2	0
5	D	4	0	6	0	0
6	A	52	0	0	2	0
6	B	44	0	0	4	0
6	C	13	0	0	2	0
6	D	12	0	0	1	0
All	All	9515	0	9159	106	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 106 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:B:458:VAL:HG12	1:B:459:ARG:H	1.24	1.02
1:B:438:LEU:HD13	1:B:458:VAL:HG13	1.57	0.84
1:B:270:GLY:HA3	1:B:304:ILE:HA	1.66	0.78
1:A:68:SER:O	1:A:71:GLN:NE2	2.16	0.77
1:B:458:VAL:HG12	1:B:459:ARG:N	1.99	0.75

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	439/455 (96%)	418 (95%)	18 (4%)	3 (1%)	22 39
1	B	439/455 (96%)	417 (95%)	19 (4%)	3 (1%)	22 39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	163/178 (92%)	151 (93%)	12 (7%)	0	100	100
2	D	164/178 (92%)	158 (96%)	6 (4%)	0	100	100
All	All	1205/1266 (95%)	1144 (95%)	55 (5%)	6 (0%)	29	48

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	LEU
1	A	448	SER
1	B	456	LYS
1	B	457	PRO
1	A	72	ASP

### 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	377/409 (92%)	366 (97%)	11 (3%)	42	69
1	B	382/409 (93%)	377 (99%)	5 (1%)	69	87
2	C	137/153 (90%)	136 (99%)	1 (1%)	84	94
2	D	136/153 (89%)	136 (100%)	0	100	100
All	All	1032/1124 (92%)	1015 (98%)	17 (2%)	62	84

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	450	LYS
2	C	172	GLU
1	A	272	SER
1	A	276	THR
1	A	356	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	105	HIS

### 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	$\# Z  > 2$
3	NAG	E	1	1,3	14,14,15	0.53	0	17,19,21	0.68	0
3	NAG	E	2	3	14,14,15	0.41	0	17,19,21	0.39	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
3	NAG	E	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	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.

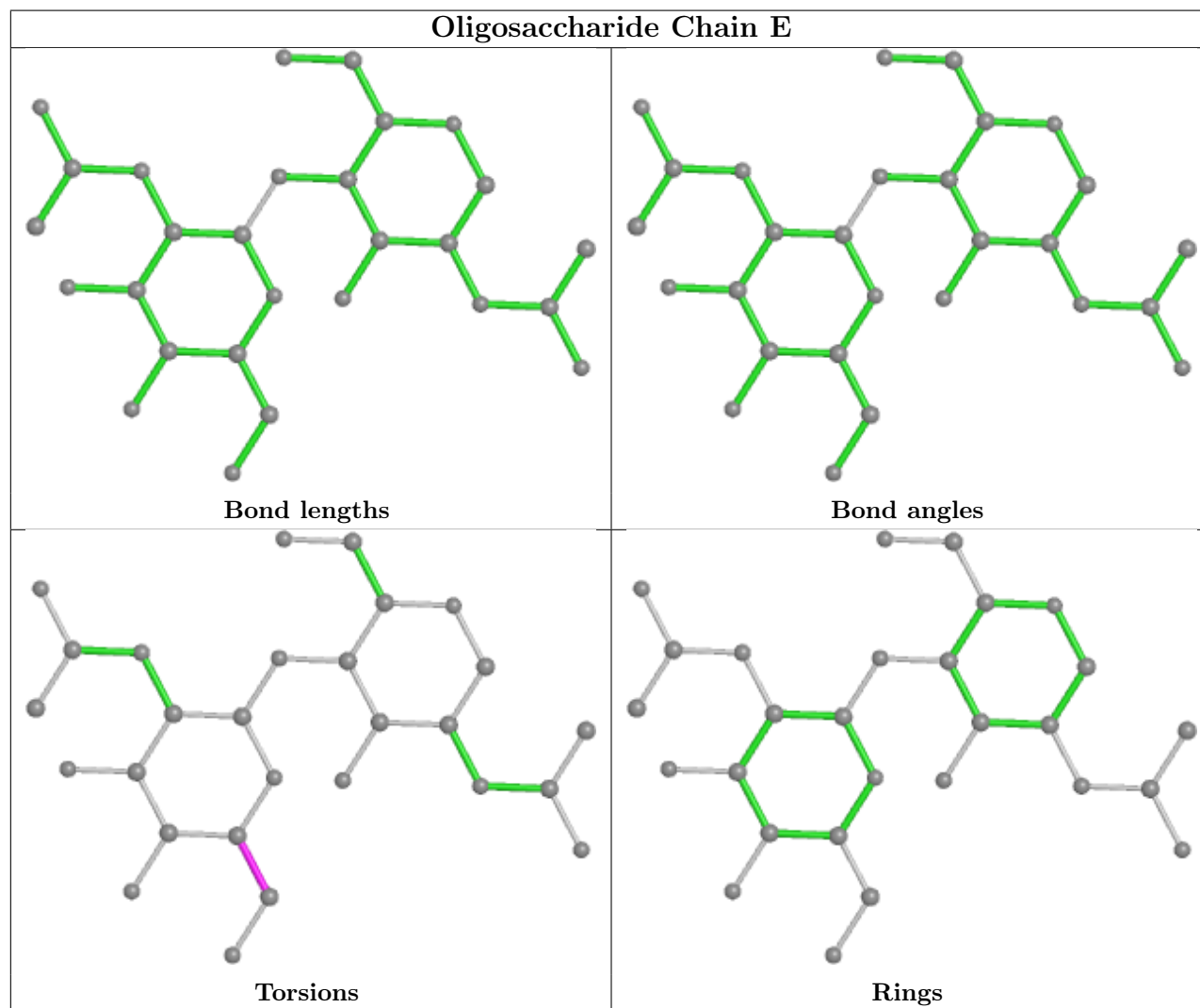
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-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 [i](#)

4 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	$\# Z  > 2$
4	NAG	B	901	1	14,14,15	1.04	1 (7%)	17,19,21	0.87	1 (5%)
4	NAG	A	901	1	14,14,15	0.49	0	17,19,21	0.64	0
4	NAG	A	902	1	14,14,15	0.45	0	17,19,21	0.56	0
5	EDO	D	201	-	3,3,3	0.38	0	2,2,2	0.62	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
4	NAG	B	901	1	-	1/6/23/26	0/1/1/1
4	NAG	A	901	1	-	2/6/23/26	0/1/1/1
4	NAG	A	902	1	-	0/6/23/26	0/1/1/1
5	EDO	D	201	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	901	NAG	C1-C2	3.34	1.57	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	901	NAG	C1-O5-C5	2.80	115.99	112.19

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	901	NAG	O5-C5-C6-O6
4	A	901	NAG	C3-C2-N2-C7
4	B	901	NAG	C3-C2-N2-C7
5	D	201	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	901	NAG	2	0

## 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 [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	441/455 (96%)	-0.07	16 (3%) 42 46	34, 63, 125, 161	0
1	B	441/455 (96%)	0.04	28 (6%) 20 21	33, 64, 139, 186	0
2	C	165/178 (92%)	-0.25	2 (1%) 79 80	33, 54, 99, 121	0
2	D	166/178 (93%)	-0.37	3 (1%) 68 71	35, 53, 98, 119	0
All	All	1213/1266 (95%)	-0.09	49 (4%) 38 41	33, 60, 126, 186	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	462	ILE	7.2
1	B	452	SER	6.9
1	B	434	LEU	6.1
1	A	462	ILE	6.1
1	A	441	ASN	5.0

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

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

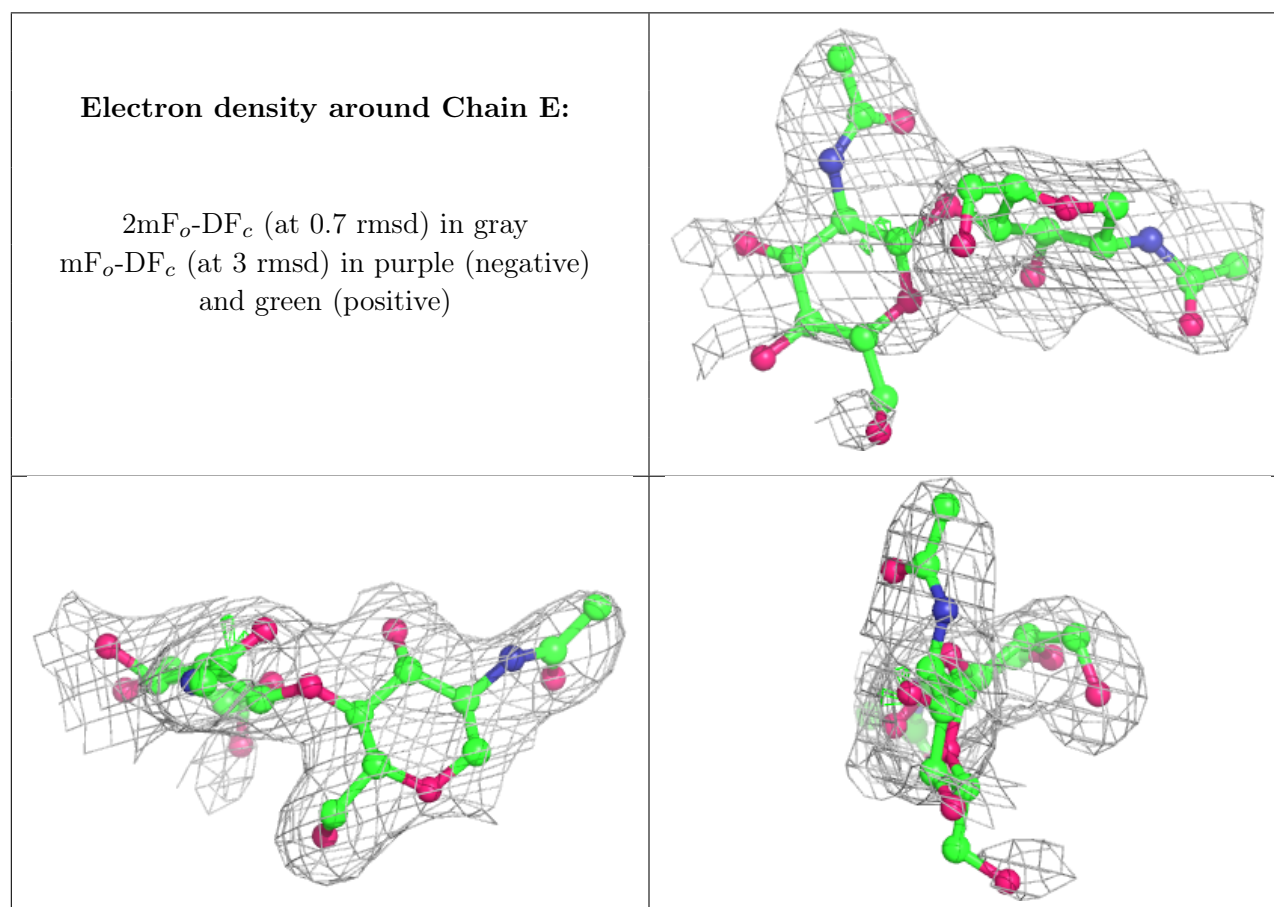
### 6.3 Carbohydrates [i](#)

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	NAG	E	2	14/15	0.73	0.24	74,81,91,96	0
3	NAG	E	1	14/15	0.89	0.11	51,59,66,69	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.



## 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	901	14/15	0.76	0.16	69,75,82,85	0
4	NAG	A	901	14/15	0.79	0.18	63,73,79,80	0
4	NAG	A	902	14/15	0.91	0.14	42,55,62,72	0
5	EDO	D	201	4/4	0.94	0.15	35,36,45,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.