



# wwPDB X-ray Structure Validation Summary Report

Sep 22, 2022 – 01:19 am BST

PDB ID : 7QNG  
Title : Structure of a MHC I-Tapasin-ERp57 complex  
Authors : Mueller, I.K.; Thomas, C.; Trowitzsch, S.; Tampe, R.  
Deposited on : 2021-12-20  
Resolution : 2.70 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)  
Xtriage (Phenix) : 1.13  
EDS : 2.30  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.30

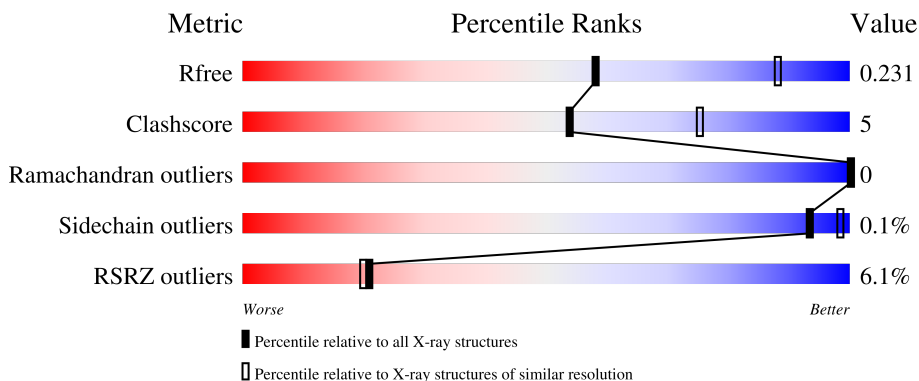
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	419	 4% 80% 10% 10%
2	B	505	 8% 80% 12% 8%
3	C	277	 4% 89% 10% .
4	D	100	 3% 91% 9%
5	E	3	 67% 33%

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 9726 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 Tapasin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	376	2847	1820	503	513	11	0	0	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	240	ARG	THR	conflict	UNP O15533
A	400	GLY	-	expression tag	UNP O15533
A	401	SER	-	expression tag	UNP O15533
A	402	GLU	-	expression tag	UNP O15533
A	403	ASN	-	expression tag	UNP O15533
A	404	LEU	-	expression tag	UNP O15533
A	405	TYR	-	expression tag	UNP O15533
A	406	PHE	-	expression tag	UNP O15533
A	407	GLN	-	expression tag	UNP O15533
A	408	GLY	-	expression tag	UNP O15533
A	409	HIS	-	expression tag	UNP O15533
A	410	HIS	-	expression tag	UNP O15533
A	411	HIS	-	expression tag	UNP O15533
A	412	HIS	-	expression tag	UNP O15533
A	413	HIS	-	expression tag	UNP O15533
A	414	HIS	-	expression tag	UNP O15533
A	415	GLY	-	expression tag	UNP O15533
A	416	SER	-	expression tag	UNP O15533
A	417	GLU	-	expression tag	UNP O15533

- Molecule 2 is a protein called Protein disulfide-isomerase A3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	464	3691	2344	620	714	13	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	36	ALA	CYS	engineered mutation	UNP P30101

- Molecule 3 is a protein called H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	274	2260	1428	398	425	9	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP P01899

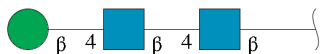
- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	100	837	533	141	159	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P61769

- Molecule 5 is an oligosaccharide called 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	N	O			
5	E	3	39	22	2	15	0	0	0

- Molecule 6 is water.

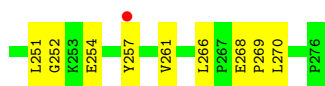
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	25	Total O 25 25	0	0

*Continued on next page...*

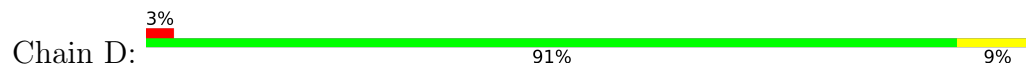
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
6	B	10	Total O 10 10	0	0
6	C	7	Total O 7 7	0	0
6	D	10	Total O 10 10	0	0





- Molecule 4: Beta-2-microglobulin



- Molecule 5: beta-D-mannopyranose-(1-4)-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 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.76Å 168.53Å 187.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.00 – 2.70 48.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.00-2.70) 99.8 (48.00-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.34 (at 2.69Å)	Xtrriage
Refinement program	PDB-REDO, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.198 , 0.229 0.206 , 0.231	Depositor DCC
$R_{free}$ test set	3284 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	87.6	Xtrriage
Anisotropy	0.190	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9726	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	106.0	wwPDB-VP

Xtrriage'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.*

<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: BMA, 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	A	0.45	0/2943	0.66	0/4037
2	B	0.46	0/3774	0.60	0/5095
3	C	0.46	0/2329	0.64	0/3162
4	D	0.41	0/860	0.62	0/1162
All	All	0.45	0/9906	0.63	0/13456

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 [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	2847	0	2802	29	0
2	B	3691	0	3597	49	0
3	C	2260	0	2131	15	0
4	D	837	0	803	7	0
5	E	39	0	34	6	0
6	A	25	0	0	0	0
6	B	10	0	0	0	0
6	C	7	0	0	1	0
6	D	10	0	0	1	0
All	All	9726	0	9367	100	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 5.

The worst 5 of 100 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:E:2:NAG:H83	5:E:2:NAG:H3	1.69	0.75
2:B:402:ASP:HB2	2:B:465:VAL:HB	1.70	0.72
2:B:395:LEU:HD11	2:B:453:ILE:HG13	1.79	0.65
3:C:215:LEU:HD22	3:C:261:VAL:HG22	1.80	0.63
2:B:34:GLY:O	2:B:38:ARG:HG2	1.99	0.62

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	372/419 (89%)	356 (96%)	16 (4%)	0	100	100
2	B	460/505 (91%)	432 (94%)	28 (6%)	0	100	100
3	C	271/277 (98%)	257 (95%)	14 (5%)	0	100	100
4	D	98/100 (98%)	95 (97%)	3 (3%)	0	100	100
All	All	1201/1301 (92%)	1140 (95%)	61 (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	299/331 (90%)	299 (100%)	0	100	100
2	B	399/428 (93%)	398 (100%)	1 (0%)	92	98
3	C	235/235 (100%)	235 (100%)	0	100	100
4	D	95/95 (100%)	95 (100%)	0	100	100
All	All	1028/1089 (94%)	1027 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	382	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	368	ASN
3	C	72	GLN
3	C	80	ASN
1	A	205	ASN
1	A	93	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](#)

3 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
5	NAG	E	1	5,1	14,14,15	0.39	0	17,19,21	0.53	0
5	NAG	E	2	5	14,14,15	0.38	0	17,19,21	1.20	1 (5%)
5	BMA	E	3	5	11,11,12	0.28	0	15,15,17	0.51	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
5	NAG	E	1	5,1	-	2/6/23/26	0/1/1/1
5	NAG	E	2	5	-	3/6/23/26	0/1/1/1
5	BMA	E	3	5	-	0/2/19/22	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(°)	Ideal(°)
5	E	2	NAG	C2-N2-C7	4.34	129.08	122.90

There are no chirality outliers.

All (5) torsion outliers are listed below:

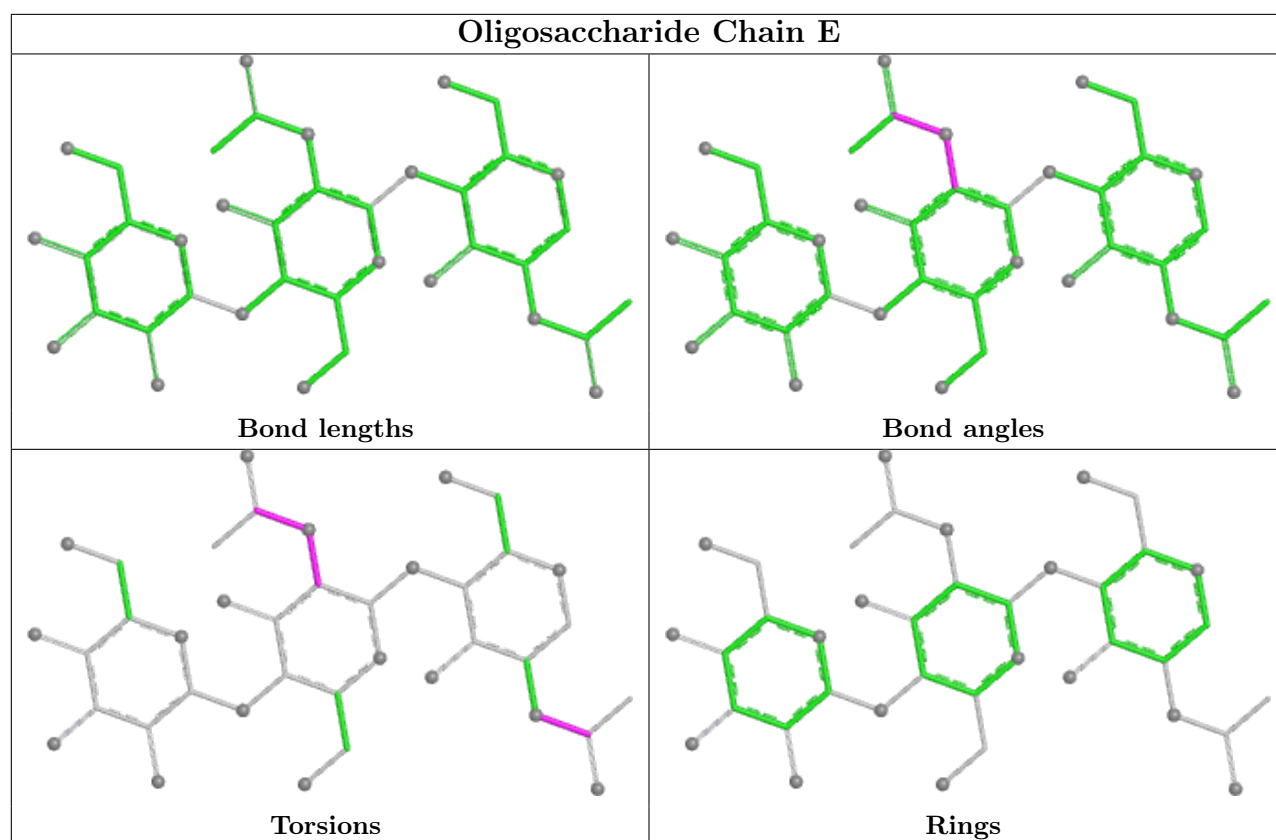
Mol	Chain	Res	Type	Atoms
5	E	1	NAG	C8-C7-N2-C2
5	E	1	NAG	O7-C7-N2-C2
5	E	2	NAG	C8-C7-N2-C2
5	E	2	NAG	O7-C7-N2-C2
5	E	2	NAG	C3-C2-N2-C7

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	1	NAG	3	0
5	E	2	NAG	3	0
5	E	3	BMA	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.



## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	376/419 (89%)	0.68	18 (4%) 30 28	55, 80, 174, 204	0
2	B	464/505 (91%)	0.60	42 (9%) 9 7	63, 112, 189, 215	0
3	C	274/277 (98%)	0.46	11 (4%) 38 37	59, 106, 155, 178	0
4	D	100/100 (100%)	0.57	3 (3%) 50 51	63, 92, 129, 149	0
All	All	1214/1301 (93%)	0.59	74 (6%) 21 20	55, 99, 176, 215	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	405	ILE	13.4
2	B	406	VAL	11.2
1	A	281	LEU	6.7
2	B	456	LEU	6.1
2	B	453	ILE	5.3

### 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.

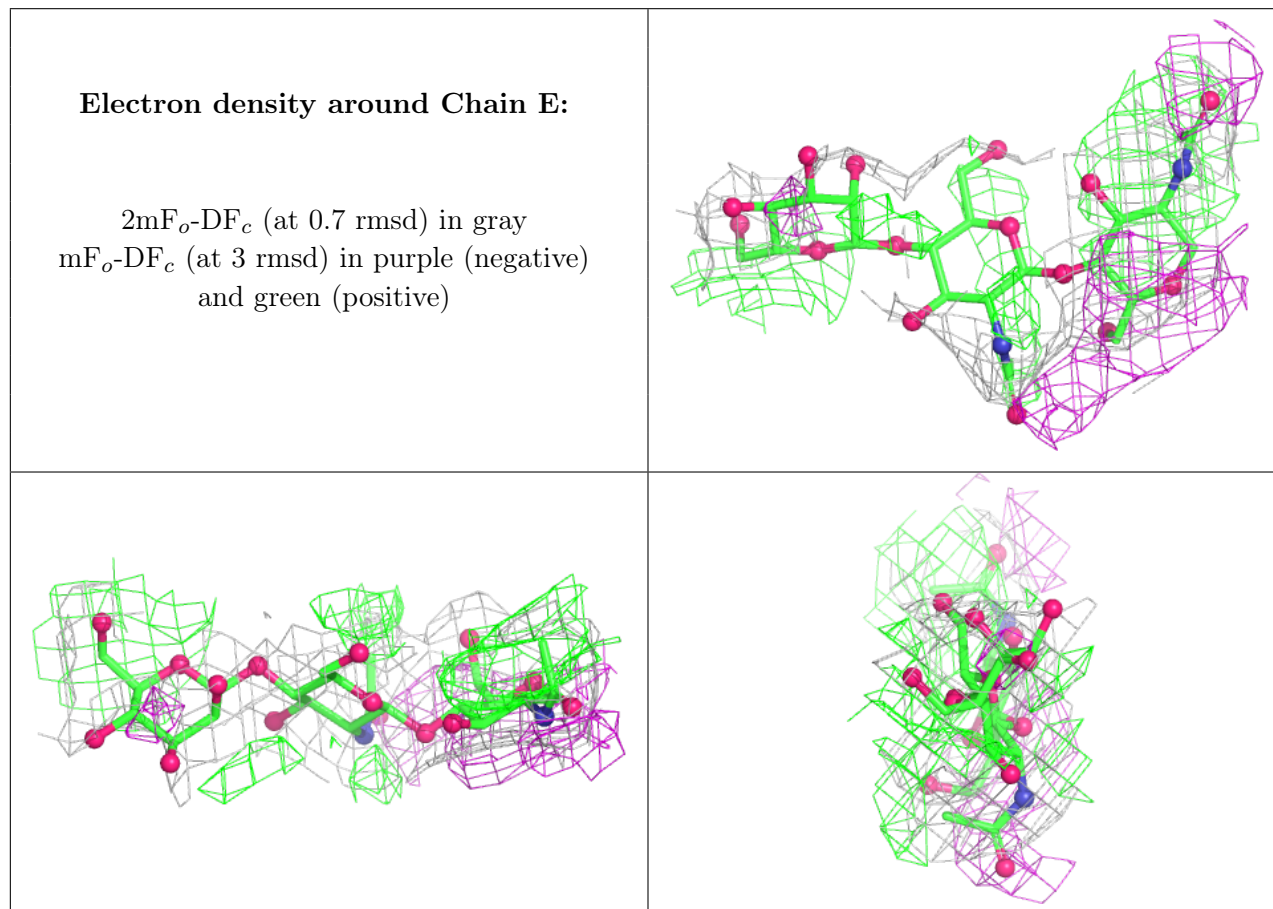
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	BMA	E	3	11/12	0.53	0.26	159,166,176,180	0
5	NAG	E	1	14/15	0.71	0.31	80,91,106,120	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	E	2	14/15	0.86	0.17	99,131,152,160	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.