

# wwPDB X-ray Structure Validation Summary Report (i)

#### Apr 15, 2024 – 04:54 PM EDT

PDB ID : 8SGM

Title: Crystal Structure of CD1d-lipid complexed with Beta-2-Microglobulin, TCR

Alpha-Chain and TCR Beta-Chain

Authors : Chan Yew Poa, K.T.O.; Le Nours, J.; Rossjohn, J.

Deposited on : 2023-04-12

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
A user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (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 : 2.36.1buster-report : 1.1.7 (2018)

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

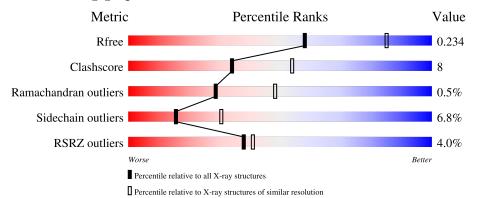
Validation Pipeline (wwPDB-VP) : 2.36.1

# 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 2.50 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}({\rm \AA})) \end{array}$
$R_{free}$	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 >=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% 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 chair	1		
1	A	274	9% 74%		22%	<del></del>
2	В	100	83%		13%	
3	С	207	85%		12%	
4	D	243	74%		22%	
5	Е	2	50%	50%		

Continued on next page...



Continued from previous page...

$\mathbf{M}$	ol	Chain	Length		Quality of chain	
6	5	F	3	33%	67%	_



# 2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 6803 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 Antigen-presenting glycoprotein CD1d.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Λ	271	Total	С	N	О	S	0	0	0
1	Α	211	2104	1346	361	390	7	U	0	U

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Α	4	MET	-	initiating methionine	UNP P15813
A	277	GLN	-	expression tag	UNP P15813

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	В	98	Total 813	C 517	N 138	O 156	S 2	0	0	0

There is a discrepancy between the modelled and reference sequences:

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

• Molecule 3 is a protein called Natural killer T cell receptor TRAV26-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	203	Total 1540	C 967	N 262	O 302	S 9	0	0	0

• Molecule 4 is a protein called Natural killer T cell receptor TRBV19 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	236	Total 1847	C 1173	N 318	O 350	S 6	0	0	0



• Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.



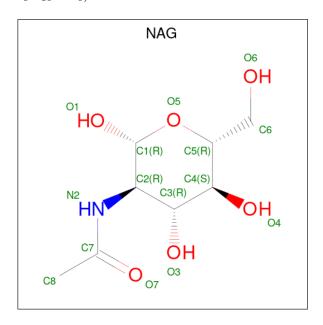
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	Е	2	Total 28	C 16	N 2	O 10	0	0	0

• Molecule 6 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-b eta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	F	3	Total 39	C 22		O 15	0	0	0

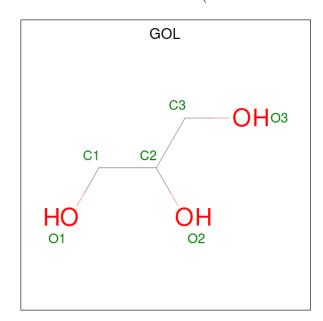
 $\bullet$  Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $\rm C_8H_{15}NO_6).$ 



I	Mol	Chain	Residues	A	ton	ns		ZeroOcc	AltConf
	7	A	1	Total 14	C 8	N 1	O 5	0	0

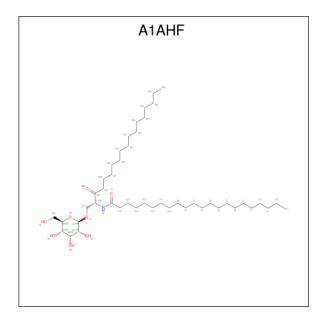


• Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mo	ol	Chain	Residues	Atoms		ZeroOcc	AltConf	
8		A	1	Total 6	C 3	O 3	0	0

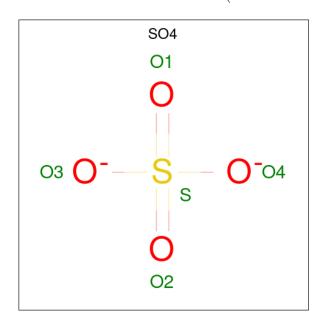
• Molecule 9 is N-[(2S)-3-oxo-1-(beta-L-talopyranosyloxy)octadecan-2-yl]docosanami de (three-letter code: A1AHF) (formula:  $C_{46}H_{89}NO_8$ ) (labeled as "Ligand of Interest" by depositor).



M	ol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	)	A	1	Total 55	C 46	N 1	O 8	0	0



 $\bullet$  Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	1	Total O S 5 4 1	0	0
10	В	1	Total O S 5 4 1	0	0

• Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	93	Total O 93 93	0	0
11	В	71	Total O 71 71	0	0
11	С	82	Total O 82 82	0	0
11	D	101	Total O 101 101	0	0

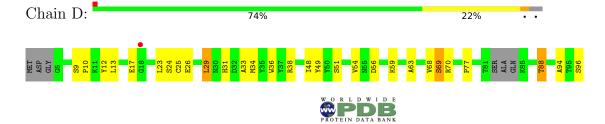


#### 3 Residue-property plots (i)

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: Antigen-presenting glycoprotein CD1d







 $\bullet$  Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 50% 50%



 $\bullet$  Molecule 6: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F: 33% 67%





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants	133.79Å 133.79Å 68.57Å	Domositon
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	47.30 - 2.50	Depositor
Resolution (A)	47.30 - 2.50	EDS
% Data completeness	100.0 (47.30-2.50)	Depositor
(in resolution range)	100.0 (47.30-2.50)	EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.29 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
D.D.	0.184 , 0.234	Depositor
$R, R_{free}$	0.184 , $0.234$	DCC
$R_{free}$ test set	2145  reflections  (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.3	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.34, 50.9	EDS
L-test for twinning <sup>2</sup>	$< L > = 0.48, < L^2> = 0.31$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6803	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.02% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: SO4, BMA, A1AHF, 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).

Mal	Mol Chain		lengths	Bond angles	
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	A	0.40	0/2165	0.56	0/2956
2	В	0.50	0/836	0.61	0/1134
3	С	0.45	0/1577	0.59	0/2155
4	D	0.44	0/1898	0.58	0/2594
All	All	0.44	0/6476	0.58	0/8839

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)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
1	A	2104	0	1970	39	0
2	В	813	0	765	8	0
3	С	1540	0	1412	22	0
4	D	1847	0	1701	42	0
5	Е	28	0	25	0	0
6	F	39	0	34	1	0
7	A	14	0	13	0	0
8	A	6	0	8	0	0
9	A	55	0	0	2	0

Continued on next page...



$\alpha \cdots$	, r	•	
Continued	trom	mromonie	maaa
-	110116	DICULUUS	Duuc
	J	1	1

Mol	Chain	Non-H	H(model)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
10	A	5	0	0	0	0
10	В	5	0	0	0	0
11	A	93	0	0	4	0
11	В	71	0	0	1	0
11	С	82	0	0	4	0
11	D	101	0	0	3	0
All	All	6803	0	5928	104	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 8.

The worst 5 of 104 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
1:A:18:PHE:HB2	1:A:94:LEU:HB2	1.63	0.81
1:A:231:GLN:HB3	1:A:248:THR:HG22	1.64	0.80
1:A:110:SER:HB3	11:A:467:HOH:O	1.87	0.74
4:D:88:THR:HB	4:D:112:THR:HA	1.72	0.72
4:D:132:GLU:OE2	4:D:132:GLU:CG	2.38	0.72

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	Perce	entiles
1	A	267/274~(97%)	252 (94%)	12 (4%)	3 (1%)	14	26
2	В	96/100 (96%)	92 (96%)	4 (4%)	0	100	100
3	С	201/207 (97%)	189 (94%)	12 (6%)	0	100	100
4	D	232/243 (96%)	214 (92%)	17 (7%)	1 (0%)	34	54
All	All	796/824 (97%)	747 (94%)	45 (6%)	4 (0%)	29	48



All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	GLY
1	A	199	PRO
1	A	253	ALA
4	D	17	GLU

#### 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	Perce	entiles
1	A	$220/240 \ (92\%)$	201 (91%)	19 (9%)	10	20
2	В	91/95 (96%)	85 (93%)	6 (7%)	16	32
3	$\mathbf{C}$	163/185 (88%)	159 (98%)	4 (2%)	47	73
4	D	191/212 (90%)	175 (92%)	16 (8%)	11	21
All	All	$665/732 \ (91\%)$	620 (93%)	45 (7%)	16	30

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	С	193	ILE
4	D	68	VAL
4	D	23	LEU
4	D	29	LEU
4	D	88	THR

Sometimes 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)

5 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	Trino	Chain	Dag	Link	Bo	ond leng	ths	Bond angles		
MIOI	Type	Chain	Res	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	Е	1	5,1	14,14,15	0.68	1 (7%)	17,19,21	0.62	0
5	NAG	Е	2	5	14,14,15	0.41	0	17,19,21	0.64	0
6	NAG	F	1	6,1	14,14,15	0.27	0	17,19,21	0.76	0
6	NAG	F	2	6	14,14,15	0.81	1 (7%)	17,19,21	0.91	1 (5%)
6	BMA	F	3	6	11,11,12	1.26	1 (9%)	15,15,17	1.02	1 (6%)

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	Е	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	Е	2	5	-	4/6/23/26	0/1/1/1
6	NAG	F	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	0/6/23/26	0/1/1/1
6	BMA	F	3	6	-	1/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

$\mathbf{Mol}$	Chain	$\operatorname{Res}$	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\operatorname{\AA})$	$\operatorname{Ideal}( ext{\AA})$
6	F	2	NAG	O5-C1	-2.85	1.39	1.43
5	Е	1	NAG	C1-C2	2.47	1.56	1.52
6	F	3	BMA	C4-C3	2.32	1.58	1.52



All (2) bond angle outliers are listed below:

$\mathbf{M}$	ol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^{o})$	$\operatorname{Ideal}(^{o})$
(	3	F	2	NAG	O4-C4-C5	-2.28	103.64	109.30
(	3	F	3	BMA	C1-C2-C3	-2.02	107.19	109.67

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	Е	2	NAG	O5-C5-C6-O6
5	Е	2	NAG	C4-C5-C6-O6
6	F	3	BMA	O5-C5-C6-O6
6	F	1	NAG	C4-C5-C6-O6
6	F	1	NAG	O5-C5-C6-O6

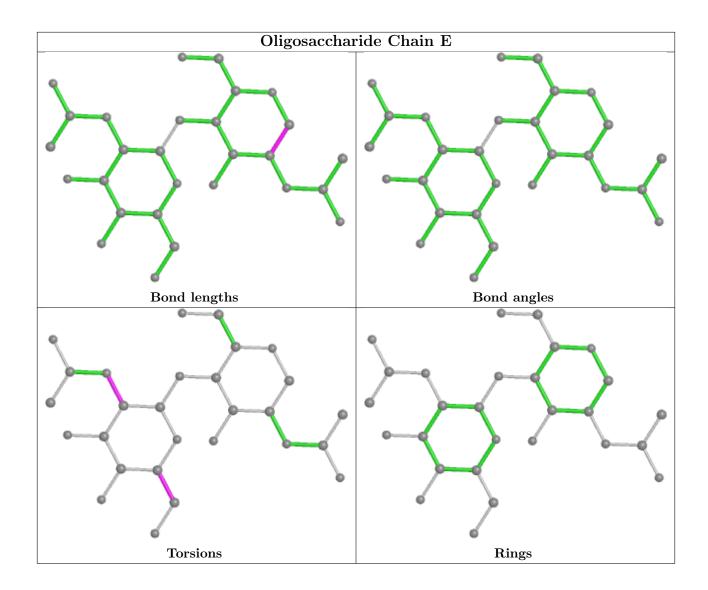
There are no ring outliers.

2 monomers are involved in 1 short contact:

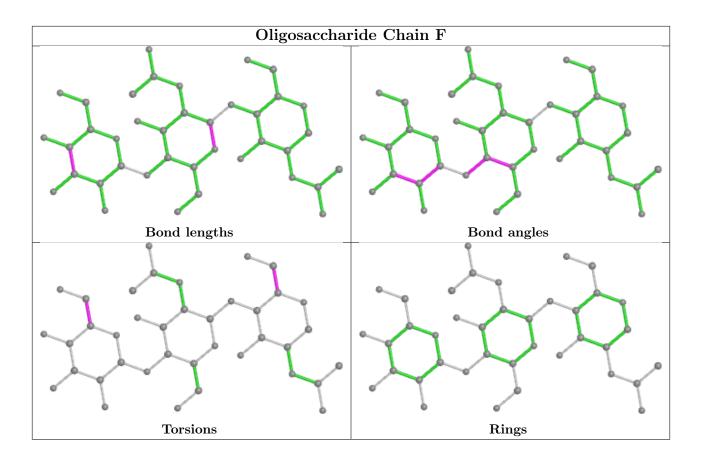
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	F	3	BMA	1	0
6	F	2	NAG	1	0

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)

5 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	Во	Bond lengths			ond ang	gles
MIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	A1AHF	A	303	-	55,55,55	1.32	4 (7%)	60,63,63	1.59	13 (21%)
8	GOL	A	302	-	5,5,5	0.47	0	5,5,5	0.39	0
10	SO4	В	101	-	4,4,4	0.19	0	6,6,6	0.16	0
10	SO4	A	304	-	4,4,4	0.15	0	6,6,6	0.20	0
7	NAG	A	301	1	14,14,15	0.59	0	17,19,21	0.76	1 (5%)

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	nο	outliers	$\alpha$ f	that	kind	were	identified.
	means	$\mathbf{n}$	Outilities	OI	unat	MILLA	WCIC	identifica.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	A1AHF	A	303	-	-	23/50/71/71	0/1/1/1
8	GOL	A	302	-	-	0/4/4/4	-
7	NAG	A	301	1	-	2/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}( ext{\AA})$
9	A	303	A1AHF	C17-N	6.62	1.48	1.34
9	A	303	A1AHF	O-C17	-2.56	1.18	1.23
9	A	303	A1AHF	O26-C22	2.27	1.48	1.43
9	A	303	A1AHF	O2-C20	2.06	1.47	1.41

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	$\Gamma_{\rm ype}     \text{Atoms}    $		$\operatorname{Observed}(^{o})$	$\operatorname{Ideal}({}^{o})$
9	A	303	A1AHF	C16-C17-N	6.04	126.31	115.83
9	A	303	A1AHF	C23-C22-C21	3.43	116.81	110.82
9	A	303	A1AHF	C22-C23-C24	3.41	116.32	110.24
9	A	303	A1AHF	O-C17-N	-3.31	117.37	122.95
9	A	303	A1AHF	O-C17-C16	-3.11	116.32	122.02

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	303	A1AHF	N-C18-C19-O1
9	A	303	A1AHF	C54-C18-C19-O1
9	A	303	A1AHF	O2-C20-O1-C19
9	A	303	A1AHF	C18-C54-C55-C56
9	A	303	A1AHF	O28-C54-C55-C56

There are no ring outliers.

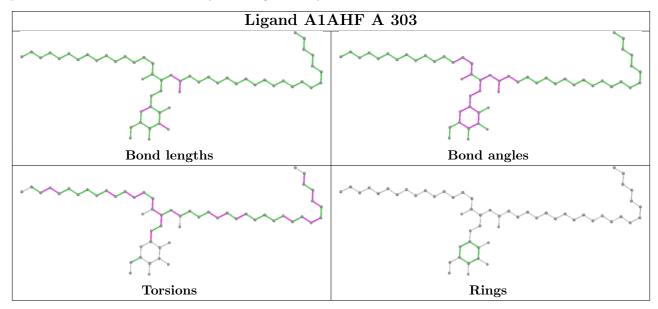
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	303	A1AHF	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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will



also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



# 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,  $95^{th}$  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>	$\#\mathrm{RSRZ}{>}2$	$OWAB(\AA^2)$	Q<0.9
1	A	271/274 (98%)	0.28	24 (8%) 9 9	35, 49, 92, 108	0
2	В	98/100 (98%)	-0.16	0 100 100	34, 42, 64, 82	0
3	С	203/207 (98%)	-0.01	5 (2%) 57 61	30, 45, 74, 87	0
4	D	236/243 (97%)	-0.10	3 (1%) 77 79	30, 53, 77, 88	0
All	All	808/824 (98%)	0.04	32 (3%) 38 41	30, 48, 79, 108	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	253	ALA	5.5
1	A	199	PRO	4.6
1	A	254	GLY	4.1
1	A	109	ALA	4.0
3	С	129	SER	3.8

### 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
6	BMA	F	3	11/12	0.78	0.17	63,69,75,76	0
5	NAG	Е	2	14/15	0.90	0.30	82,89,91,92	0

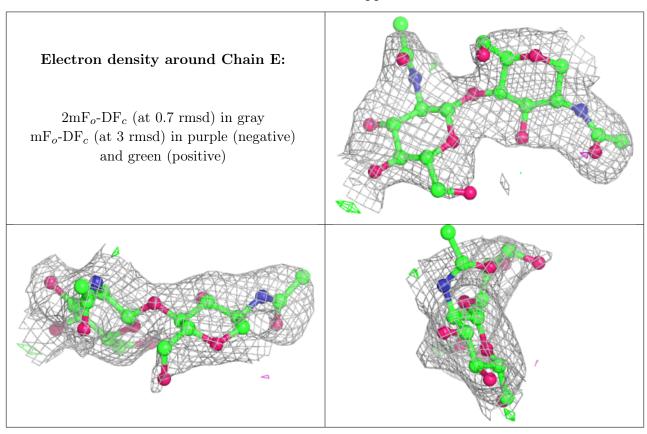
Continued on next page...



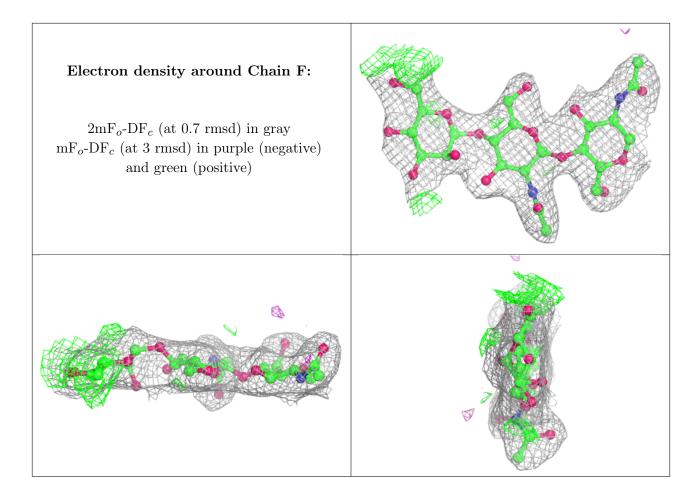
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
6	NAG	F	2	14/15	0.93	0.11	41,56,64,68	0
5	NAG	E	1	14/15	0.93	0.19	53,59,73,80	0
6	NAG	F	1	14/15	0.95	0.12	39,50,55,57	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.







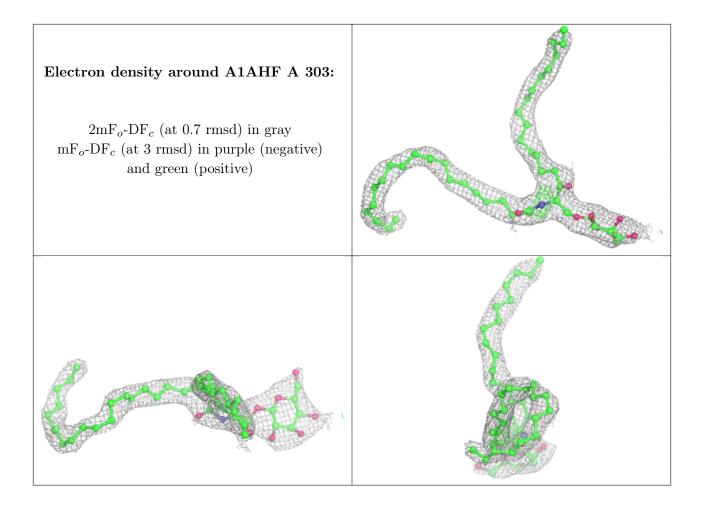
# 6.4 Ligands (i)

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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
7	NAG	A	301	14/15	0.89	0.24	65,73,78,80	0
9	A1AHF	A	303	55/55	0.91	0.22	40,49,73,77	0
8	GOL	A	302	6/6	0.93	0.21	52,54,59,60	0
10	SO4	A	304	5/5	0.94	0.32	87,87,87,94	0
10	SO4	В	101	5/5	0.97	0.16	91,91,92,93	0

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.





# 6.5 Other polymers (i)

There are no such residues in this entry.

