Full wwPDB X-ray Structure Validation Report (i)

May 19, 2022 – 02:48 pm BST

PDB ID	:	7O00
Title	:	Crystal structure of HLA-DR4 in complex with a HSP70 peptide
Authors	:	Ge, C.; Holmdahl, R.
Deposited on	:	2021-03-25
Resolution	:	2.24 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 (i) symbol.

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

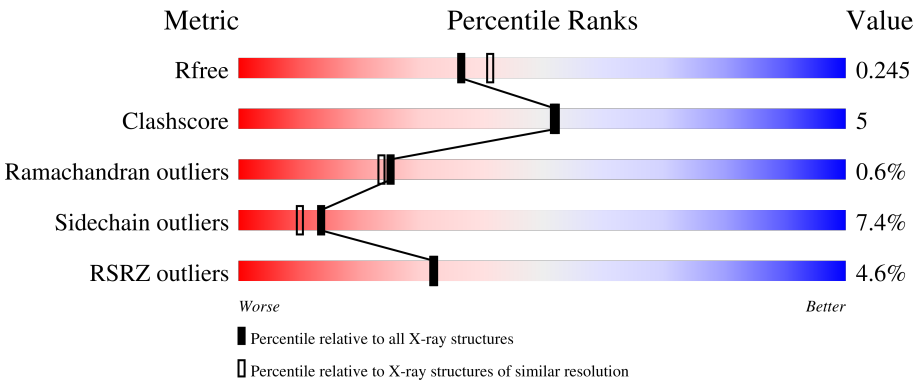
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

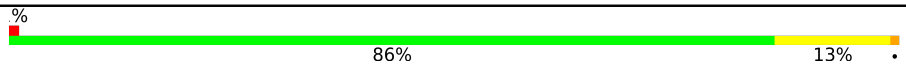
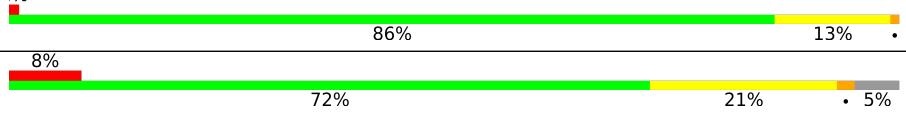
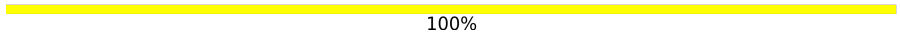
The reported resolution of this entry is 2.24 Å.

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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 chain
1	AAA	177	
2	BBB	189	
3	CCC	14	
4	AaA	2	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	AaA	2	X	-	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6154 atoms, of which 2975 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	177	2861	945	1405	237	269	5	35	0	0

- Molecule 2 is a protein called HLA class II histocompatibility antigen DR beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	BBB	180	2911	948	1416	260	282	5	45	0	0

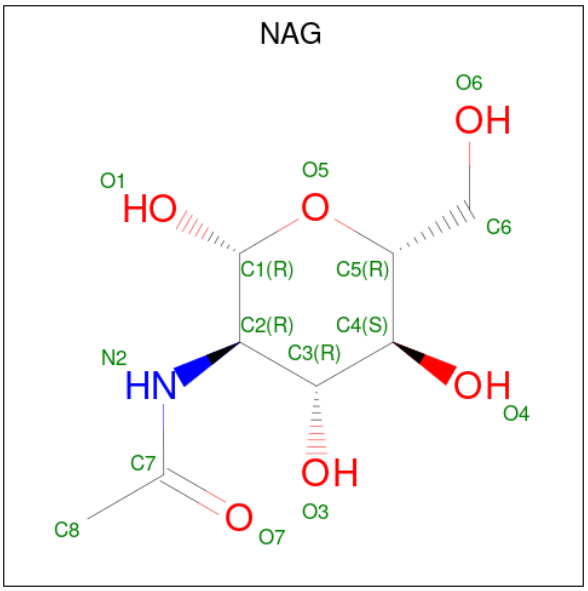
- Molecule 3 is a protein called Chaperone protein DnaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
3	CCC	14	218	67	112	19	20	3	0	0

- Molecule 4 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
			Total	C	H	N	O			
4	AaA	2	56	16	28	2	10	5	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
5	BBB	1	28	8	14	1	5	3	0

- Molecule 6 is water.

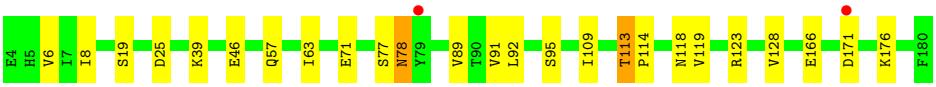
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	AAA	35	35	35	0	0
			Total	O		
6	BBB	41	41	41	0	0
			Total	O		
6	CCC	4	4	4	0	0

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.

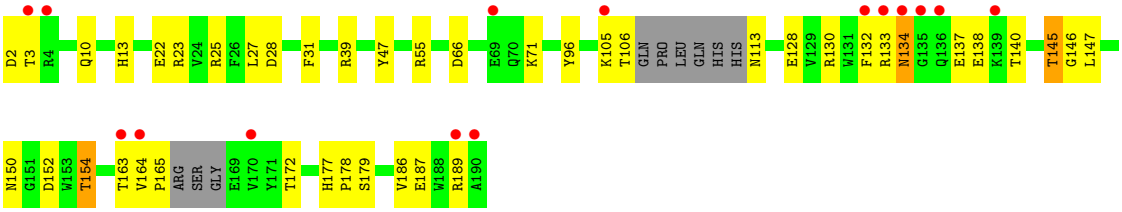
• Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain AAA:



• Molecule 2: HLA class II histocompatibility antigen DR beta chain

Chain BBB:

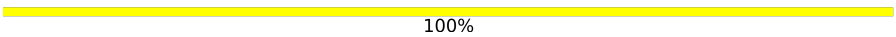


• Molecule 3: Chaperone protein DnaK

Chain CCC:

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

Chain AaA:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	76.23Å 76.23Å 170.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.45 – 2.24 51.40 – 2.24	Depositor EDS
% Data completeness (in resolution range)	89.7 (51.45-2.24) 89.7 (51.40-2.24)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.85 (at 2.25Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC 5.8.0267	Depositor
R, R_{free}	0.199 , 0.242 0.203 , 0.245	Depositor DCC
R_{free} test set	1105 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	46.1	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6154	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	AAA	0.68	0/1501	0.88	0/2047
2	BBB	0.66	0/1534	0.88	0/2083
3	CCC	0.75	0/107	0.83	0/143
All	All	0.67	0/3142	0.88	0/4273

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	BBB	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	BBB	3	THR	Peptide

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	AAA	1456	1405	1396	8	0
2	BBB	1495	1416	1406	24	0
3	CCC	106	112	112	2	0
4	AaA	28	28	25	0	0
5	BBB	14	14	13	0	0
6	AAA	35	0	0	0	0
6	BBB	41	0	0	3	0
6	CCC	4	0	0	1	0
All	All	3179	2975	2952	33	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.

All (33) 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)
2:BBB:150:ASN:HD22	2:BBB:154:THR:CG2	1.89	0.85
2:BBB:150:ASN:HD22	2:BBB:154:THR:HG22	1.43	0.83
2:BBB:128:GLU:OE1	2:BBB:130:ARG:NH2	2.28	0.65
1:AAA:118:ASN:HB2	1:AAA:166:GLU:HB2	1.80	0.63
2:BBB:13:HIS:CD2	2:BBB:28:ASP:OD1	2.51	0.63
1:AAA:113:THR:HG22	1:AAA:114:PRO:HA	1.84	0.60
3:CCC:1:ARG:N	6:CCC:101:HOH:O	2.35	0.59
2:BBB:71:LYS:HE2	3:CCC:10:ASP:OD2	2.02	0.58
2:BBB:150:ASN:HB2	2:BBB:154:THR:HG22	1.92	0.52
2:BBB:13:HIS:HD2	2:BBB:28:ASP:OD1	1.93	0.50
1:AAA:77:SER:O	1:AAA:78:ASN:HB2	2.11	0.50
2:BBB:145:THR:HG22	2:BBB:146:GLY:O	2.12	0.49
2:BBB:106:THR:O	2:BBB:113:ASN:N	2.47	0.47
2:BBB:150:ASN:HD22	2:BBB:154:THR:HG23	1.77	0.47
2:BBB:71:LYS:NZ	6:BBB:301:HOH:O	2.47	0.47
2:BBB:145:THR:CG2	2:BBB:146:GLY:O	2.64	0.46
2:BBB:152:ASP:OD1	2:BBB:154:THR:HB	2.16	0.45
1:AAA:6:VAL:HG12	1:AAA:8:ILE:HG13	1.98	0.45
2:BBB:164:VAL:HA	6:BBB:329:HOH:O	2.17	0.45
2:BBB:2:ASP:N	6:BBB:302:HOH:O	2.49	0.45
2:BBB:96:TYR:CD1	2:BBB:179:SER:HB2	2.52	0.44
2:BBB:163:THR:O	2:BBB:165:PRO:HD3	2.18	0.44
2:BBB:10:GLN:HB2	2:BBB:31:PHE:HB2	2.00	0.43
2:BBB:25:ARG:NH1	2:BBB:27:LEU:HD21	2.33	0.43
1:AAA:123:ARG:NH1	1:AAA:128:VAL:HG11	2.34	0.43
1:AAA:91:VAL:HG23	1:AAA:176:LYS:HB3	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:BBB:130:ARG:HD3	2:BBB:137:GLU:OE2	2.19	0.42
2:BBB:133:ARG:O	2:BBB:134:ASN:C	2.57	0.42
2:BBB:47:TYR:OH	2:BBB:71:LYS:HE3	2.19	0.41
2:BBB:132:PHE:HB2	2:BBB:172:THR:OG1	2.20	0.41
1:AAA:89:VAL:O	1:AAA:176:LYS:HE2	2.21	0.41
2:BBB:177:HIS:CG	2:BBB:178:PRO:HD2	2.56	0.40
1:AAA:109:ILE:HD13	1:AAA:119:VAL:HG21	2.03	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	AAA	175/177 (99%)	168 (96%)	6 (3%)	1 (1%)	25	23
2	BBB	174/189 (92%)	167 (96%)	6 (3%)	1 (1%)	25	23
3	CCC	12/14 (86%)	12 (100%)	0	0	100	100
All	All	361/380 (95%)	347 (96%)	12 (3%)	2 (1%)	25	23

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	46	GLU
2	BBB	134	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	AAA	162/162 (100%)	151 (93%)	11 (7%)	16 13
2	BBB	163/171 (95%)	149 (91%)	14 (9%)	10 7
3	CCC	12/12 (100%)	12 (100%)	0	100 100
All	All	337/345 (98%)	312 (93%)	25 (7%)	13 10

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	19	SER
1	AAA	25	ASP
1	AAA	39	LYS
1	AAA	57	GLN
1	AAA	63	ILE
1	AAA	71	GLU
1	AAA	78	ASN
1	AAA	92	LEU
1	AAA	95	SER
1	AAA	113	THR
1	AAA	171	ASP
2	BBB	22	GLU
2	BBB	23	ARG
2	BBB	39	ARG
2	BBB	55	ARG
2	BBB	66	ASP
2	BBB	105	LYS
2	BBB	138	GLU
2	BBB	140	THR
2	BBB	145	THR
2	BBB	147	LEU
2	BBB	154	THR
2	BBB	186	VAL
2	BBB	187	GLU
2	BBB	189	ARG

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)

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	$\#\|Z\| > 2$
4	NAG	AaA	1	4,1	14,14,15	0.48	0	17,19,21	1.50	3 (17%)
4	NAG	AaA	2	4	14,14,15	0.50	0	17,19,21	1.21	2 (11%)

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	AaA	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	AaA	2	4	1/1/7/7	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AaA	1	NAG	C3-C4-C5	-3.51	103.98	110.24
4	AaA	2	NAG	C4-C3-C2	3.48	116.12	111.02
4	AaA	1	NAG	C4-C3-C2	-2.40	107.51	111.02
4	AaA	2	NAG	C3-C4-C5	2.22	114.19	110.24
4	AaA	1	NAG	O3-C3-C4	2.09	115.19	110.35

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	AaA	2	NAG	C1

All (4) torsion outliers are listed below:

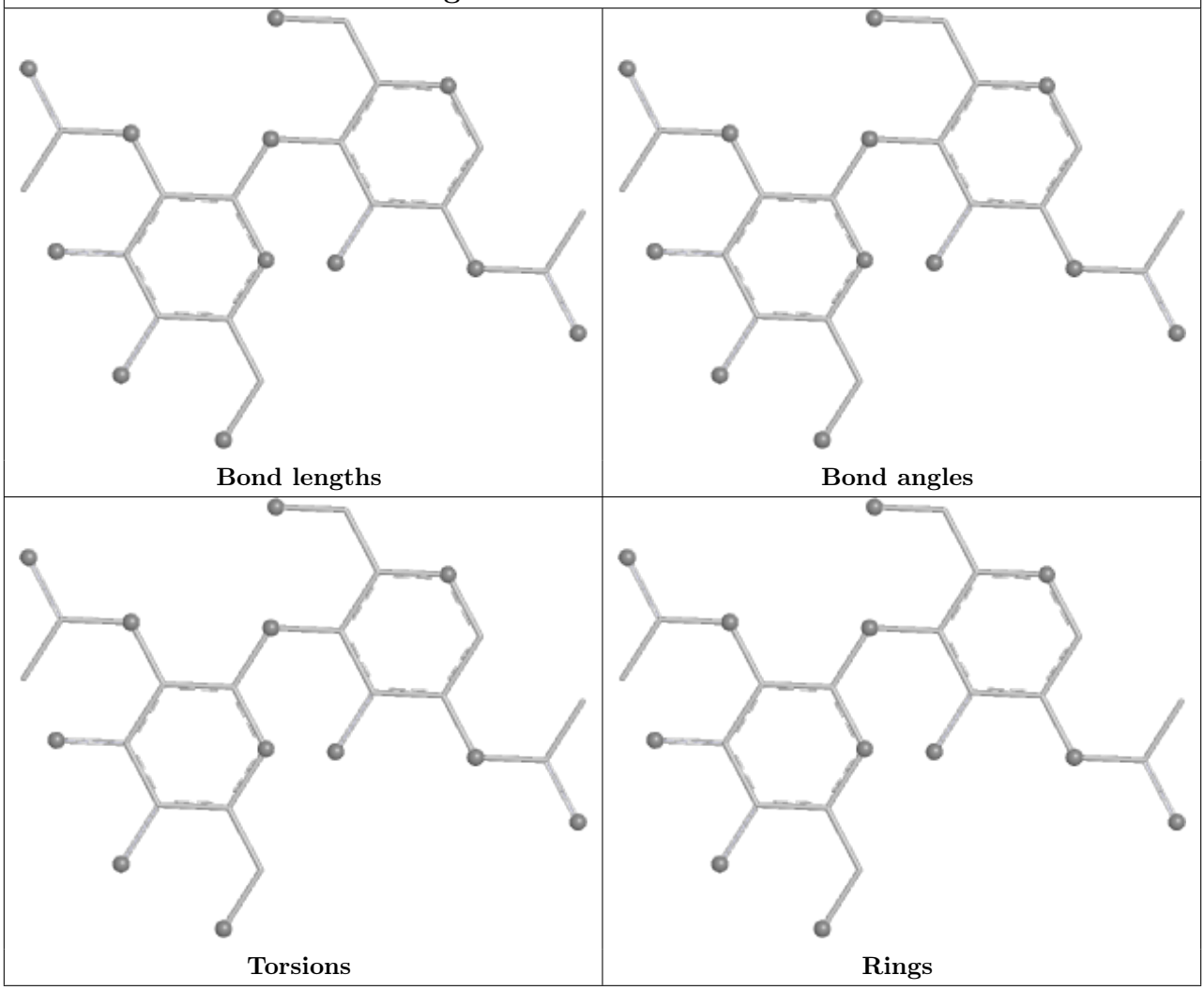
Mol	Chain	Res	Type	Atoms
4	AaA	1	NAG	O5-C5-C6-O6
4	AaA	2	NAG	O5-C5-C6-O6
4	AaA	1	NAG	C4-C5-C6-O6
4	AaA	2	NAG	C4-C5-C6-O6

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.

Oligosaccharide Chain AaA



5.6 Ligand geometry (i)

1 ligand is 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	BBB	201	2	14,14,15	0.62	0	17,19,21	2.04	3 (17%)

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	BBB	201	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	BBB	201	NAG	C1-O5-C5	6.68	121.24	112.19
5	BBB	201	NAG	C4-C3-C2	-3.01	106.60	111.02
5	BBB	201	NAG	O5-C1-C2	2.39	115.06	111.29

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 (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, 95th 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^2)	Q<0.9
1	AAA	177/177 (100%)	0.25	2 (1%) 80 81	34, 51, 72, 84	0
2	BBB	180/189 (95%)	0.68	15 (8%) 11 11	38, 56, 95, 104	0
3	CCC	14/14 (100%)	0.32	0 100 100	44, 55, 73, 76	0
All	All	371/380 (97%)	0.47	17 (4%) 32 32	34, 54, 86, 104	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	BBB	164	VAL	8.5
2	BBB	136	GLN	6.6
2	BBB	3	THR	4.1
2	BBB	189	ARG	3.3
2	BBB	105	LYS	3.2
2	BBB	135	GLY	3.2
1	AAA	171	ASP	3.2
2	BBB	133	ARG	2.9
2	BBB	69	GLU	2.8
2	BBB	170	VAL	2.7
2	BBB	190	ALA	2.7
2	BBB	132	PHE	2.6
2	BBB	4	ARG	2.5
2	BBB	139	LYS	2.3
2	BBB	134	ASN	2.2
2	BBB	163	THR	2.1
1	AAA	79	TYR	2.1

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)

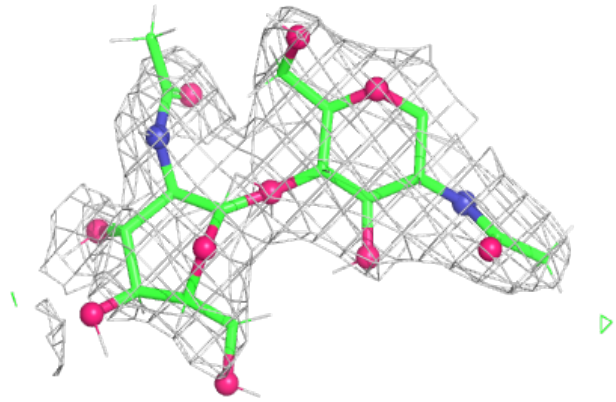
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

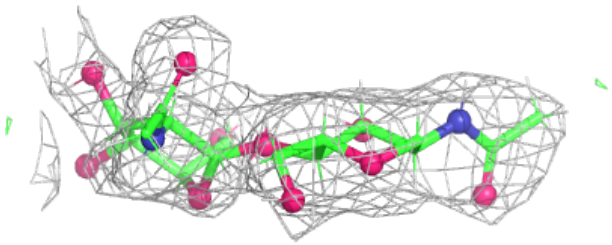
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	AaA	2	14/15	0.81	0.35	30,106,113,118	3
4	NAG	AaA	1	14/15	0.91	0.17	30,71,85,89	2

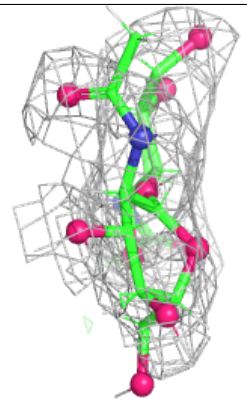
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain AaA:

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 (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, 95th 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^2)	Q<0.9
5	NAG	BBB	201	14/15	0.74	0.28	30,109,117,120	3

6.5 Other polymers (i)

There are no such residues in this entry.