



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:02 PM GMT

PDB ID : 4UO5  
Title : Structure of the A\_Canine\_Colorado\_17864\_06 H3 haemagglutinin in complex with 3SLN  
Authors : Vachieri, S.G.; Collins, P.J.; Haire, L.F.; Ogrodowicz, R.W.; Martin, S.R.; Walker, P.A.; Xiong, X.; Gamblin, S.J.; Skehel, J.J.  
Deposited on : 2014-05-31  
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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

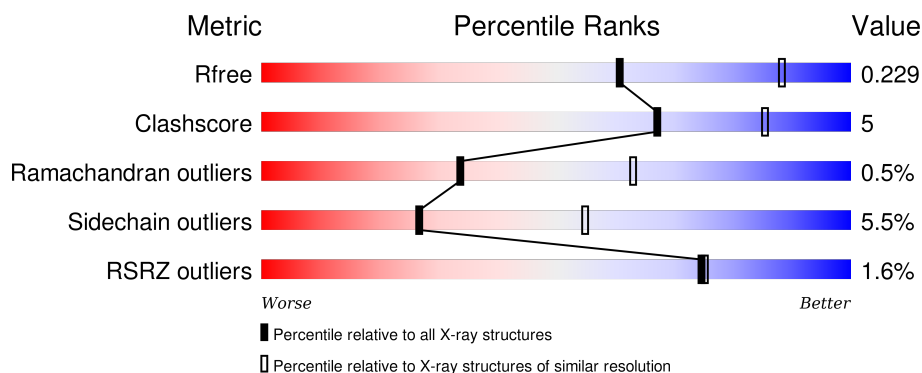
# 1 Overall quality at a glance ⓘ

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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	328	<div> <div>2%</div> <div>84%</div> <div>12%</div> <div>..</div> </div>
1	C	328	<div> <div>2%</div> <div>82%</div> <div>13%</div> <div>..</div> </div>
1	E	328	<div> <div>2%</div> <div>80%</div> <div>16%</div> <div>..</div> </div>
2	B	175	<div> <div>2%</div> <div>87%</div> <div>11%</div> <div>..</div> </div>
2	D	175	<div> <div>2%</div> <div>86%</div> <div>13%</div> <div>.</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	175	

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	E	641	-	-	-	X
5	NAG	A	641	-	-	-	X
5	NAG	C	621	X	-	-	-
9	NAG	C	641	-	-	-	X
9	NAG	C	645	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 12482 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 H3 HAEMAGGLUTININ HA1 CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2474	1544	437	477	16			
1	C	319	Total	C	N	O	S	0	0	0
			2475	1544	438	477	16			
1	E	319	Total	C	N	O	S	0	0	0
			2476	1545	438	477	16			

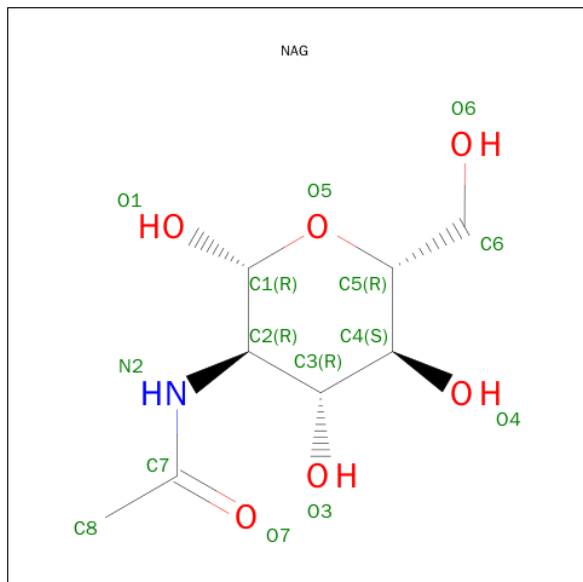
- Molecule 2 is a protein called H3 HAEMAGGLUTININ HA2 CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1425	886	250	283	6			
2	D	175	Total	C	N	O	S	0	0	0
			1419	883	247	283	6			
2	F	175	Total	C	N	O	S	0	0	0
			1406	876	243	281	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	173	SER	-	EXPRESSION TAG	UNP E0UVR5
B	174	GLY	-	EXPRESSION TAG	UNP E0UVR5
B	175	ARG	-	EXPRESSION TAG	UNP E0UVR5
B	131	GLU	ASP	CONFLICT	UNP E0UVR5
D	173	SER	-	EXPRESSION TAG	UNP E0UVR5
D	174	GLY	-	EXPRESSION TAG	UNP E0UVR5
D	175	ARG	-	EXPRESSION TAG	UNP E0UVR5
D	131	GLU	ASP	CONFLICT	UNP E0UVR5
F	173	SER	-	EXPRESSION TAG	UNP E0UVR5
F	174	GLY	-	EXPRESSION TAG	UNP E0UVR5
F	175	ARG	-	EXPRESSION TAG	UNP E0UVR5
F	131	GLU	ASP	CONFLICT	UNP E0UVR5

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	C	2	Total	C	N	O	0	0
			28	16	2	10		
4	E	2	Total	C	N	O	0	0
			28	16	2	10		
4	E	2	Total	C	N	O	0	0
			28	16	2	10		
4	E	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	3	Total	C	N	O	0	0
			39	22	2	15		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	3	Total	C	N	O	0	0
			39	22	2	15		
5	C	3	Total	C	N	O	0	0
			39	22	2	15		

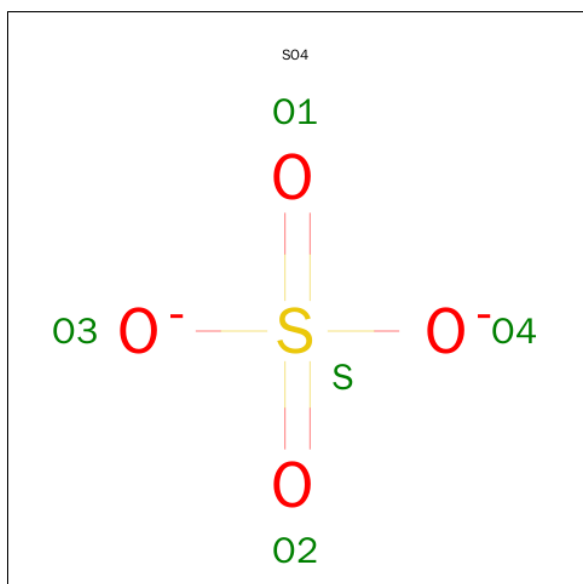
- Molecule 6 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	5	Total	C	N	O	0	0
			61	34	2	25		
6	C	5	Total	C	N	O	0	0
			61	34	2	25		
6	E	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 7 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	3	Total	C	N	O	0	0
			46	25	2	19		
7	C	3	Total	C	N	O	0	0
			46	25	2	19		
7	E	3	Total	C	N	O	0	0
			46	25	2	19		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	O	S	0	0
			5	4	1		
8	D	1	Total	O	S	0	0
			5	4	1		
8	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 9 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	C	5	Total	C	N	O	0	0
			64	36	3	25		

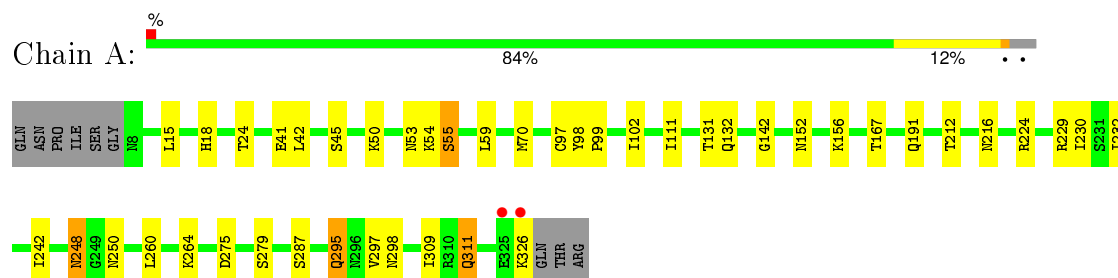
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	21	Total	O	0	0
			21	21		
10	B	13	Total	O	0	0
			13	13		
10	C	36	Total	O	0	0
			36	36		
10	D	11	Total	O	0	0
			11	11		
10	E	35	Total	O	0	0
			35	35		
10	F	6	Total	O	0	0
			6	6		

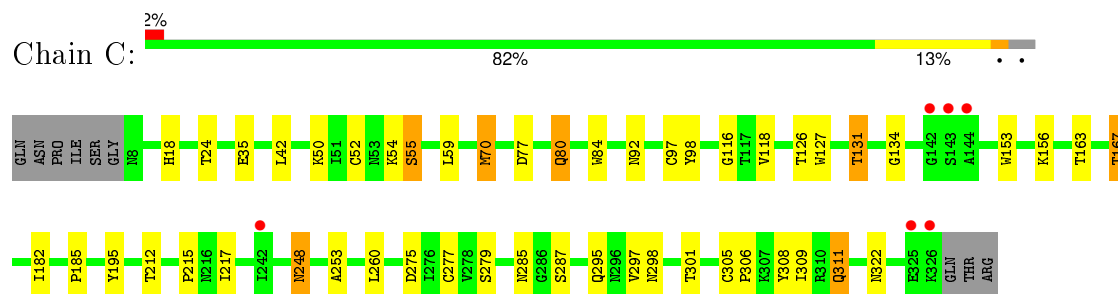
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

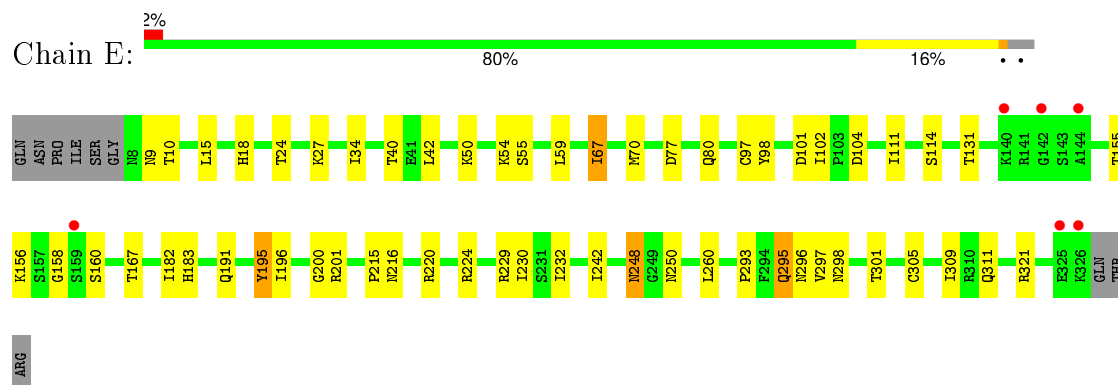
- Molecule 1: H3 HAEMAGGLUTININ HA1 CHAIN



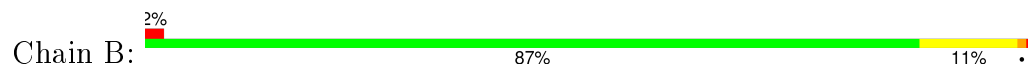
- Molecule 1: H3 HAEMAGGLUTININ HA1 CHAIN



- Molecule 1: H3 HAEMAGGLUTININ HA1 CHAIN

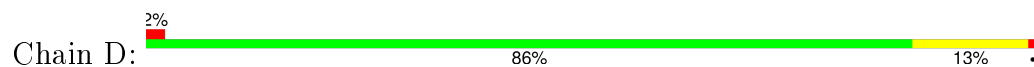


- Molecule 2: H3 HAEMAGGLUTININ HA2 CHAIN

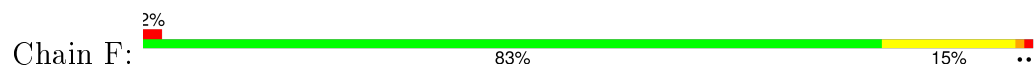




- Molecule 2: H3 HAEMAGGLUTININ HA2 CHAIN



- Molecule 2: H3 HAEMAGGLUTININ HA2 CHAIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.17Å 349.15Å 96.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	174.58 – 2.70 47.80 – 2.70	Depositor EDS
% Data completeness (in resolution range)	96.5 (174.58-2.70) 96.5 (47.80-2.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.87 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.197 , 0.229 0.197 , 0.229	Depositor DCC
$R_{free}$ test set	3789 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	67.7	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 40.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 75241 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12482	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage'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, 0.375 respectively for untwinned datasets, and 0.333, 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: BMA, NAG, SIA, GAL, SO4, MAN

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	0/2525	0.52	0/3425
1	C	0.33	0/2522	0.54	0/3413
1	E	0.32	0/2525	0.53	0/3421
2	B	0.32	0/1450	0.52	1/1949 (0.1%)
2	D	0.31	0/1444	0.53	1/1942 (0.1%)
2	F	0.33	0/1431	0.54	1/1927 (0.1%)
All	All	0.32	0/11897	0.53	3/16077 (0.0%)

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
5	C	1	0

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	110	LEU	CA-CB-CG	7.36	132.23	115.30
2	D	110	LEU	CA-CB-CG	6.86	131.08	115.30
2	B	110	LEU	CA-CB-CG	6.27	129.72	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	C	621	NAG	C1

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	2474	0	2424	21	0
1	C	2475	0	2422	32	0
1	E	2476	0	2430	27	0
2	B	1425	0	1344	15	0
2	D	1419	0	1333	16	0
2	F	1406	0	1309	22	0
3	A	14	0	13	0	0
3	C	14	0	13	1	0
4	A	28	0	25	0	0
4	C	28	0	25	0	0
4	E	84	0	75	0	0
5	A	78	0	68	0	0
5	C	39	0	34	0	0
6	A	61	0	52	3	0
6	C	61	0	52	0	0
6	E	61	0	52	0	0
7	A	46	0	40	0	0
7	C	46	0	40	0	0
7	E	46	0	40	0	0
8	B	5	0	0	0	0
8	D	5	0	0	0	0
8	F	5	0	0	0	0
9	C	64	0	55	2	0
10	A	21	0	0	0	0
10	B	13	0	0	1	0
10	C	36	0	0	1	0
10	D	11	0	0	0	0
10	E	35	0	0	0	0
10	F	6	0	0	0	0
All	All	12482	0	11846	117	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 117 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)
6:A:633:BMA:H5	6:A:637:MAN:H2	1.58	0.84
1:C:285:ASN:HD21	9:C:641:NAG:C1	1.93	0.81
2:B:134:GLY:HA2	2:D:124:ARG:HD3	1.65	0.79
2:B:25:ARG:HE	2:B:34:GLN:HE21	1.33	0.75
1:C:295:GLN:NE2	1:C:298:ASN:H	1.87	0.72

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	317/328 (97%)	301 (95%)	14 (4%)	2 (1%)	30	59
1	C	315/328 (96%)	301 (96%)	12 (4%)	2 (1%)	30	59
1	E	316/328 (96%)	300 (95%)	13 (4%)	3 (1%)	21	49
2	B	173/175 (99%)	164 (95%)	9 (5%)	0	100	100
2	D	173/175 (99%)	163 (94%)	10 (6%)	0	100	100
2	F	173/175 (99%)	163 (94%)	10 (6%)	0	100	100
All	All	1467/1509 (97%)	1392 (95%)	68 (5%)	7 (0%)	34	63

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	55	SER
1	E	196	ILE
1	A	55	SER
1	C	55	SER
1	E	158	GLY

### 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	280/289 (97%)	263 (94%)	17 (6%)	23	49
1	C	279/289 (96%)	265 (95%)	14 (5%)	30	60
1	E	280/289 (97%)	260 (93%)	20 (7%)	18	41
2	B	148/148 (100%)	143 (97%)	5 (3%)	44	75
2	D	147/148 (99%)	141 (96%)	6 (4%)	37	69
2	F	144/148 (97%)	136 (94%)	8 (6%)	26	54
All	All	1278/1311 (98%)	1208 (94%)	70 (6%)	27	55

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	167	THR
2	D	110	LEU
2	F	86	ASP
1	C	195	TYR
1	C	311	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	250	ASN
2	D	27	GLN
2	F	116	ASN
1	C	311	GLN
2	D	34	GLN

### 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 ⓘ

48 carbohydrates 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	A	611	1,4	14,14,15	0.49	0	15,19,21	0.95	1 (6%)
4	NAG	A	612	4	14,14,15	0.61	0	15,19,21	1.29	2 (13%)
5	NAG	A	621	1,5	14,14,15	0.56	0	15,19,21	0.74	0
5	NAG	A	622	5	14,14,15	0.64	0	15,19,21	1.16	1 (6%)
5	BMA	A	623	5	11,11,12	0.46	0	14,15,17	0.99	1 (7%)
6	NAG	A	631	1,6	14,14,15	0.56	0	15,19,21	1.09	1 (6%)
6	NAG	A	632	6	14,14,15	0.51	0	15,19,21	1.35	2 (13%)
6	BMA	A	633	6	11,11,12	0.75	0	14,15,17	1.95	4 (28%)
6	MAN	A	634	6	11,11,12	0.62	0	14,15,17	1.87	2 (14%)
6	MAN	A	637	6	11,11,12	0.54	0	14,15,17	1.18	1 (7%)
5	NAG	A	641	5	14,14,15	0.54	0	15,19,21	1.60	2 (13%)
5	NAG	A	642	5	14,14,15	0.57	0	15,19,21	1.26	1 (6%)
5	BMA	A	643	5	11,11,12	0.46	0	14,15,17	1.59	2 (14%)
7	SIA	A	701	7	16,20,21	0.35	0	18,28,31	0.69	0
7	GAL	A	702	7	11,11,12	0.57	0	14,15,17	0.91	0
7	NAG	A	703	7	15,15,15	0.51	0	17,21,21	1.30	1 (5%)
4	NAG	C	611	1,4	14,14,15	0.63	0	15,19,21	0.99	1 (6%)
4	NAG	C	612	4	14,14,15	0.59	0	15,19,21	1.12	2 (13%)
5	NAG	C	621	1,5	14,14,15	0.42	0	15,19,21	1.33	1 (6%)
5	NAG	C	622	5	14,14,15	0.56	0	15,19,21	1.20	1 (6%)
5	BMA	C	623	5	11,11,12	0.43	0	14,15,17	0.78	0
6	NAG	C	631	1,6	14,14,15	0.57	0	15,19,21	1.18	3 (20%)
6	NAG	C	632	6	14,14,15	0.57	0	15,19,21	0.64	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	BMA	C	633	6	11,11,12	0.41	0	14,15,17	1.41	1 (7%)
6	MAN	C	634	6	11,11,12	0.59	0	14,15,17	1.43	2 (14%)
6	MAN	C	637	6	11,11,12	0.65	0	14,15,17	0.90	0
9	NAG	C	641	9	14,14,15	0.60	0	15,19,21	1.36	2 (13%)
9	NAG	C	642	9	14,14,15	0.50	0	15,19,21	1.01	0
9	BMA	C	643	9	11,11,12	0.48	0	14,15,17	1.28	2 (14%)
9	MAN	C	644	9	11,11,12	0.57	0	14,15,17	0.90	1 (7%)
9	NAG	C	645	9	14,14,15	0.46	0	15,19,21	0.90	1 (6%)
7	SIA	C	701	7	16,20,21	0.36	0	18,28,31	0.62	0
7	GAL	C	702	7	11,11,12	0.62	0	14,15,17	1.30	1 (7%)
7	NAG	C	703	7	15,15,15	0.52	0	17,21,21	1.26	1 (5%)
4	NAG	E	611	1,4	14,14,15	0.55	0	15,19,21	0.79	1 (6%)
4	NAG	E	612	4	14,14,15	0.44	0	15,19,21	1.07	1 (6%)
4	NAG	E	621	1,4	14,14,15	0.54	0	15,19,21	0.87	0
4	NAG	E	622	4	14,14,15	0.64	0	15,19,21	1.05	1 (6%)
6	NAG	E	631	1,6	14,14,15	0.57	0	15,19,21	1.03	1 (6%)
6	NAG	E	632	6	14,14,15	0.47	0	15,19,21	0.92	1 (6%)
6	BMA	E	633	6	11,11,12	0.46	0	14,15,17	2.31	2 (14%)
6	MAN	E	634	6	11,11,12	0.71	0	14,15,17	1.67	3 (21%)
6	MAN	E	637	6	11,11,12	0.67	0	14,15,17	0.88	1 (7%)
4	NAG	E	641	4	14,14,15	0.52	0	15,19,21	1.02	2 (13%)
4	NAG	E	642	4	14,14,15	0.55	0	15,19,21	1.31	2 (13%)
7	SIA	E	701	7	16,20,21	0.42	0	18,28,31	0.65	0
7	GAL	E	702	7	11,11,12	0.64	0	14,15,17	0.89	0
7	NAG	E	703	7	15,15,15	0.49	0	17,21,21	1.00	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 no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	611	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	612	4	-	0/6/23/26	0/1/1/1
5	NAG	A	621	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	622	5	-	0/6/23/26	0/1/1/1
5	BMA	A	623	5	-	0/2/19/22	0/1/1/1
6	NAG	A	631	1,6	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	A	632	6	-	0/6/23/26	0/1/1/1
6	BMA	A	633	6	-	0/2/19/22	1/1/1/1
6	MAN	A	634	6	-	0/2/19/22	0/1/1/1
6	MAN	A	637	6	-	0/2/19/22	1/1/1/1
5	NAG	A	641	5	-	0/6/23/26	0/1/1/1
5	NAG	A	642	5	-	0/6/23/26	0/1/1/1
5	BMA	A	643	5	-	0/2/19/22	0/1/1/1
7	SIA	A	701	7	-	0/14/34/38	0/1/1/1
7	GAL	A	702	7	-	0/2/19/22	0/1/1/1
7	NAG	A	703	7	-	0/6/26/26	0/1/1/1
4	NAG	C	611	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	612	4	-	0/6/23/26	0/1/1/1
5	NAG	C	621	1,5	1/1/5/7	0/6/23/26	0/1/1/1
5	NAG	C	622	5	-	0/6/23/26	0/1/1/1
5	BMA	C	623	5	-	0/2/19/22	0/1/1/1
6	NAG	C	631	1,6	-	0/6/23/26	0/1/1/1
6	NAG	C	632	6	-	0/6/23/26	0/1/1/1
6	BMA	C	633	6	-	0/2/19/22	0/1/1/1
6	MAN	C	634	6	-	0/2/19/22	0/1/1/1
6	MAN	C	637	6	-	0/2/19/22	0/1/1/1
9	NAG	C	641	9	-	0/6/23/26	0/1/1/1
9	NAG	C	642	9	-	0/6/23/26	0/1/1/1
9	BMA	C	643	9	-	0/2/19/22	0/1/1/1
9	MAN	C	644	9	-	0/2/19/22	0/1/1/1
9	NAG	C	645	9	-	0/6/23/26	0/1/1/1
7	SIA	C	701	7	-	0/14/34/38	0/1/1/1
7	GAL	C	702	7	-	0/2/19/22	0/1/1/1
7	NAG	C	703	7	-	0/6/26/26	0/1/1/1
4	NAG	E	611	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	612	4	-	0/6/23/26	0/1/1/1
4	NAG	E	621	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	622	4	-	0/6/23/26	0/1/1/1
6	NAG	E	631	1,6	-	0/6/23/26	0/1/1/1
6	NAG	E	632	6	-	0/6/23/26	0/1/1/1
6	BMA	E	633	6	-	0/2/19/22	0/1/1/1
6	MAN	E	634	6	-	0/2/19/22	0/1/1/1
6	MAN	E	637	6	-	0/2/19/22	0/1/1/1
4	NAG	E	641	4	-	0/6/23/26	0/1/1/1
4	NAG	E	642	4	-	0/6/23/26	0/1/1/1
7	SIA	E	701	7	-	0/14/34/38	0/1/1/1
7	GAL	E	702	7	-	0/2/19/22	0/1/1/1
7	NAG	E	703	7	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	632	NAG	C3-C4-C5	-2.64	105.60	110.20
6	A	633	BMA	C3-C4-C5	-2.30	106.19	110.20
6	A	633	BMA	C2-C3-C4	-2.22	107.27	111.04
6	C	631	NAG	O7-C7-C8	-2.09	118.22	122.06
9	C	645	NAG	C1-O5-C5	2.02	114.81	112.25

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	C	621	NAG	C1

There are no torsion outliers.

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	633	BMA	C1-C2-C3-C4-C5-O5
6	A	637	MAN	C1-C2-C3-C4-C5-O5

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	633	BMA	3	0
6	A	634	MAN	1	0
6	A	637	MAN	1	0
9	C	641	NAG	2	0

## 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	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	A	601	1	14,14,15	0.61	0	15,19,21	1.64	3 (20%)
8	SO4	B	1176	-	4,4,4	0.38	0	6,6,6	0.06	0
3	NAG	C	601	1	14,14,15	0.64	0	15,19,21	3.05	4 (26%)
8	SO4	D	1176	-	4,4,4	0.37	0	6,6,6	0.10	0
8	SO4	F	1176	-	4,4,4	0.39	0	6,6,6	0.09	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	A	601	1	-	0/6/23/26	0/1/1/1
8	SO4	B	1176	-	-	0/0/0/0	0/0/0/0
3	NAG	C	601	1	-	0/6/23/26	0/1/1/1
8	SO4	D	1176	-	-	0/0/0/0	0/0/0/0
8	SO4	F	1176	-	-	0/0/0/0	0/0/0/0

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(°)
3	C	601	NAG	O7-C7-C8	-2.45	117.56	122.06
3	A	601	NAG	C8-C7-N2	2.18	120.28	116.11
3	A	601	NAG	C2-N2-C7	2.57	126.33	123.04
3	C	601	NAG	C8-C7-N2	3.01	121.87	116.11
3	A	601	NAG	C1-O5-C5	4.25	117.64	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	601	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	319/328 (97%)	0.03	2 (0%) 90 91	46, 66, 93, 130	0
1	C	319/328 (97%)	-0.00	6 (1%) 70 70	45, 64, 96, 132	0
1	E	319/328 (97%)	0.06	6 (1%) 70 70	46, 64, 95, 127	0
2	B	175/175 (100%)	0.26	4 (2%) 64 64	44, 75, 97, 152	0
2	D	175/175 (100%)	0.19	3 (1%) 73 74	43, 73, 98, 137	0
2	F	175/175 (100%)	0.14	3 (1%) 73 74	42, 71, 95, 128	0
All	All	1482/1509 (98%)	0.09	24 (1%) 74 75	42, 68, 96, 152	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	174	GLY	9.1
2	B	175	ARG	8.1
2	D	173	SER	6.8
2	D	175	ARG	5.2
2	D	174	GLY	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
9	NAG	C	645	14/15	0.66	0.41	5.70	130,150,156,158	0
9	NAG	C	641	14/15	0.84	0.31	3.73	111,124,129,132	0
4	NAG	E	641	14/15	0.79	0.33	3.54	109,125,128,131	0
5	NAG	A	641	14/15	0.84	0.20	3.35	103,107,114,116	0
7	SIA	C	701	20/21	0.96	0.23	0.78	72,79,86,86	0
7	SIA	A	701	20/21	0.97	0.20	0.31	70,75,81,82	0
6	NAG	C	632	14/15	0.96	0.21	-0.04	76,89,98,106	0
7	SIA	E	701	20/21	0.94	0.16	-0.35	71,73,82,85	0
6	NAG	A	632	14/15	0.96	0.18	-0.48	87,95,107,114	0
6	NAG	E	632	14/15	0.96	0.14	-1.19	77,93,100,109	0
4	NAG	E	612	14/15	0.83	0.42	-	119,125,129,132	0
6	NAG	C	631	14/15	0.97	0.13	-	75,79,81,82	0
9	BMA	C	643	11/12	0.83	0.32	-	136,143,147,147	0
6	NAG	E	631	14/15	0.97	0.13	-	82,86,92,93	0
6	MAN	C	634	11/12	0.79	0.35	-	132,135,141,141	0
6	MAN	E	634	11/12	0.75	0.49	-	145,150,157,160	0
7	NAG	C	703	15/15	0.85	0.52	-	111,120,135,136	0
5	BMA	C	623	11/12	0.69	0.36	-	149,158,160,160	0
7	NAG	A	703	15/15	0.88	0.36	-	100,107,118,118	0
5	NAG	C	621	14/15	0.84	0.29	-	101,108,119,132	0
4	NAG	E	642	14/15	0.52	0.41	-	125,138,144,146	0
5	NAG	A	621	14/15	0.85	0.35	-	104,113,122,133	0
5	NAG	A	622	14/15	0.74	0.51	-	135,148,153,154	0
7	GAL	A	702	11/12	0.95	0.22	-	84,90,95,95	0
6	MAN	C	637	11/12	0.74	0.28	-	135,141,150,152	0
6	MAN	E	637	11/12	0.70	0.27	-	142,148,153,154	0
6	BMA	C	633	11/12	0.85	0.23	-	115,125,132,137	0
6	MAN	A	634	11/12	0.76	0.30	-	146,150,154,156	0
4	NAG	A	612	14/15	0.84	0.25	-	116,122,128,129	0
4	NAG	C	612	14/15	0.78	0.50	-	120,124,134,136	0
6	NAG	A	631	14/15	0.95	0.15	-	78,82,85,86	0
9	NAG	C	642	14/15	0.81	0.39	-	127,134,141,143	0
4	NAG	E	621	14/15	0.88	0.30	-	99,108,113,122	0
6	BMA	E	633	11/12	0.79	0.27	-	120,131,141,142	0
7	GAL	C	702	11/12	0.95	0.18	-	92,99,105,105	0
9	MAN	C	644	11/12	0.84	0.43	-	144,147,149,149	0
5	BMA	A	623	11/12	0.71	0.44	-	147,153,155,155	0
4	NAG	A	611	14/15	0.94	0.13	-	84,89,98,109	0
5	BMA	A	643	11/12	0.80	0.33	-	137,143,146,146	0
5	NAG	A	642	14/15	0.88	0.32	-	116,124,132,138	0
4	NAG	E	622	14/15	0.78	0.44	-	130,135,138,139	0
4	NAG	E	611	14/15	0.92	0.25	-	92,98,106,113	0
6	BMA	A	633	11/12	0.76	0.29	-	124,139,146,155	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	GAL	E	702	11/12	0.95	0.13	-	87,93,97,97	0
5	NAG	C	622	14/15	0.67	0.41	-	141,148,155,156	0
6	MAN	A	637	11/12	0.73	0.33	-	154,157,160,161	0
4	NAG	C	611	14/15	0.87	0.24	-	91,98,106,115	0
7	NAG	E	703	15/15	0.81	0.35	-	102,110,121,123	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	SO4	B	1176	5/5	0.95	0.20	-0.24	93,99,99,104	0
8	SO4	F	1176	5/5	0.96	0.17	-0.82	91,91,93,94	0
8	SO4	D	1176	5/5	0.93	0.13	-1.69	92,94,96,96	0
3	NAG	C	601	14/15	0.78	0.40	-	107,113,124,126	0
3	NAG	A	601	14/15	0.78	0.34	-	114,121,125,126	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.