



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 12, 2021 – 04:25 PM EDT

PDB ID : 4FP8  
Title : Crystal structure of broadly neutralizing antibody C05 bound to H3 influenza hemagglutinin, HA1 subunit  
Authors : Ekiert, D.C.; Wilson, I.A.  
Deposited on : 2012-06-21  
Resolution : 2.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.18  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.18

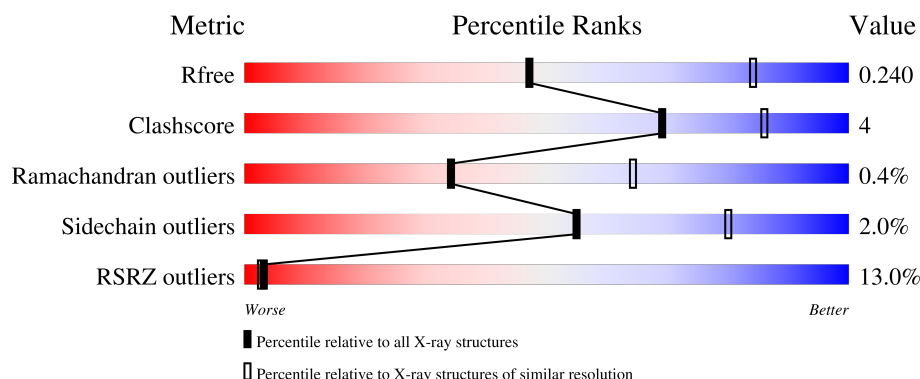
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.95 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.



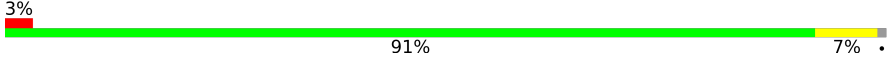

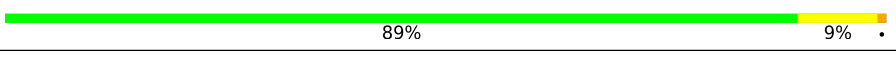
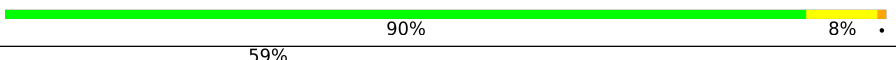

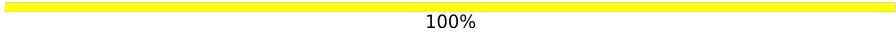
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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	278	<div> <div>7%</div> <div>88%</div> <div>8%</div> <div>.</div> </div>
1	B	278	<div> <div>11%</div> <div>92%</div> <div>.</div> <div>.</div> </div>
1	C	278	<div> <div>15%</div> <div>86%</div> <div>9%</div> <div>5%</div> </div>
1	D	278	<div> <div>19%</div> <div>88%</div> <div>8%</div> <div>.</div> </div>
2	H	241	<div> <div>3%</div> <div>88%</div> <div>10%</div> <div>..</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	I	241	 3% 91% 7% •
2	J	241	 5% 91% 8% •
2	K	241	 30% 88% 10% ••
3	L	214	 88% 10% •
3	M	214	 89% 9% •
3	N	214	 90% 8% •
3	O	214	 59% 87% 11% •
4	E	2	 100%

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22406 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 Hemagglutinin HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	8	0
			2141	1346	375	408	12			
1	B	267	Total	C	N	O	S	0	8	0
			2141	1346	375	408	12			
1	C	265	Total	C	N	O	S	0	7	0
			2116	1330	369	405	12			
1	D	266	Total	C	N	O	S	0	8	0
			2128	1335	374	407	12			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	ALA	-	expression tag	UNP Q91MA7
A	40	ASP	-	expression tag	UNP Q91MA7
A	41	PRO	-	expression tag	UNP Q91MA7
A	42	GLY	-	expression tag	UNP Q91MA7
A	310	GLY	-	expression tag	UNP Q91MA7
A	311	HIS	-	expression tag	UNP Q91MA7
A	312	HIS	-	expression tag	UNP Q91MA7
A	313	HIS	-	expression tag	UNP Q91MA7
A	314	HIS	-	expression tag	UNP Q91MA7
A	315	HIS	-	expression tag	UNP Q91MA7
A	316	HIS	-	expression tag	UNP Q91MA7
B	39	ALA	-	expression tag	UNP Q91MA7
B	40	ASP	-	expression tag	UNP Q91MA7
B	41	PRO	-	expression tag	UNP Q91MA7
B	42	GLY	-	expression tag	UNP Q91MA7
B	310	GLY	-	expression tag	UNP Q91MA7
B	311	HIS	-	expression tag	UNP Q91MA7
B	312	HIS	-	expression tag	UNP Q91MA7
B	313	HIS	-	expression tag	UNP Q91MA7
B	314	HIS	-	expression tag	UNP Q91MA7
B	315	HIS	-	expression tag	UNP Q91MA7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	316	HIS	-	expression tag	UNP Q91MA7
C	39	ALA	-	expression tag	UNP Q91MA7
C	40	ASP	-	expression tag	UNP Q91MA7
C	41	PRO	-	expression tag	UNP Q91MA7
C	42	GLY	-	expression tag	UNP Q91MA7
C	310	GLY	-	expression tag	UNP Q91MA7
C	311	HIS	-	expression tag	UNP Q91MA7
C	312	HIS	-	expression tag	UNP Q91MA7
C	313	HIS	-	expression tag	UNP Q91MA7
C	314	HIS	-	expression tag	UNP Q91MA7
C	315	HIS	-	expression tag	UNP Q91MA7
C	316	HIS	-	expression tag	UNP Q91MA7
D	39	ALA	-	expression tag	UNP Q91MA7
D	40	ASP	-	expression tag	UNP Q91MA7
D	41	PRO	-	expression tag	UNP Q91MA7
D	42	GLY	-	expression tag	UNP Q91MA7
D	310	GLY	-	expression tag	UNP Q91MA7
D	311	HIS	-	expression tag	UNP Q91MA7
D	312	HIS	-	expression tag	UNP Q91MA7
D	313	HIS	-	expression tag	UNP Q91MA7
D	314	HIS	-	expression tag	UNP Q91MA7
D	315	HIS	-	expression tag	UNP Q91MA7
D	316	HIS	-	expression tag	UNP Q91MA7

- Molecule 2 is a protein called Antibody C05, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	239	Total	C	N	O	S	0	4	0
			1806	1127	306	365	8			
2	I	239	Total	C	N	O	S	0	4	0
			1806	1127	306	365	8			
2	J	239	Total	C	N	O	S	0	4	0
			1806	1127	306	365	8			
2	K	239	Total	C	N	O	S	0	4	0
			1806	1127	306	365	8			

- Molecule 3 is a protein called Antibody C05, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	213	Total	C	N	O	S	0	1	0
			1648	1033	281	330	4			

*Continued on next page...*

Continued from previous page...

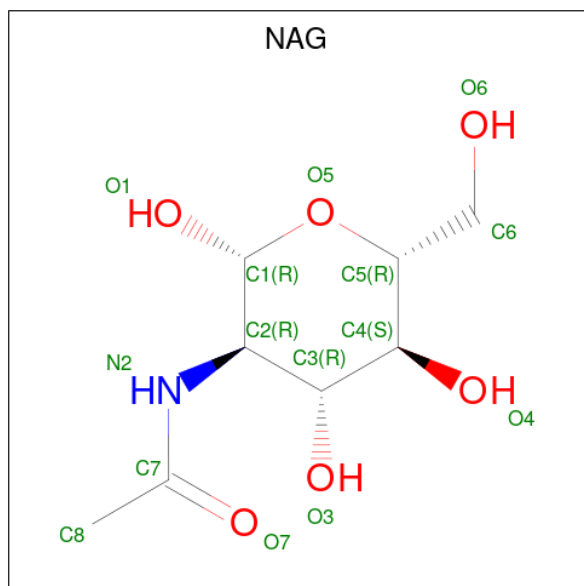
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	M	213	Total	C	N	O	S	0	1	0
			1648	1033	281	330	4			
3	N	213	Total	C	N	O	S	0	1	0
			1648	1033	281	330	4			
3	O	213	Total	C	N	O	S	0	1	0
			1648	1033	281	330	4			

- 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
4	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	1	Total 1	Zn 1	0	0
6	L	2	Total 2	Zn 2	0	0
6	M	1	Total 1	Zn 1	0	0
6	J	1	Total 1	Zn 1	0	0
6	N	2	Total 2	Zn 2	0	0
6	O	1	Total 1	Zn 1	0	0




- Molecule 1: Hemagglutinin HA1 chain








Chain M:  89% 9% .




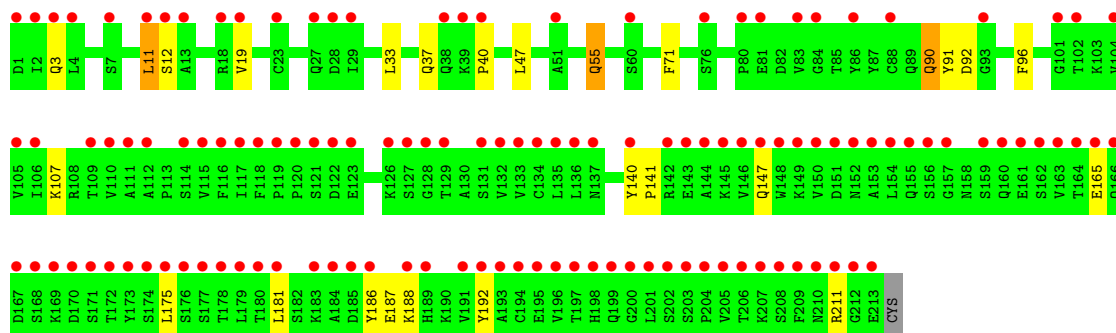
- Molecule 3: Antibody C05, light chain

Chain N:  90% 8% .



- Molecule 3: Antibody C05, light chain

Chain O:  59% 87% 11% .



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

Chain E:  100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.79Å 247.72Å 95.38Å 90.00° 91.33° 90.00°	Depositor
Resolution (Å)	44.49 – 2.95 44.49 – 2.95	Depositor EDS
% Data completeness (in resolution range)	95.1 (44.49-2.95) 95.0 (44.49-2.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.64 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.212 , 0.241 0.210 , 0.240	Depositor DCC
$R_{free}$ test set	4191 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.8	Xtriage
Anisotropy	0.240	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 36.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.008 for l,k,-h 0.028 for h,-k,-l 0.021 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	22406	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, ZN, PCA

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	0.30	0/2195	0.58	0/2990
1	B	0.29	0/2195	0.57	0/2990
1	C	0.29	0/2170	0.59	0/2956
1	D	0.31	0/2181	0.58	0/2971
2	H	0.31	0/1837	0.58	1/2500 (0.0%)
2	I	0.28	0/1837	0.54	0/2500
2	J	0.28	0/1837	0.54	0/2500
2	K	0.26	0/1837	0.51	0/2500
3	L	0.34	0/1682	0.59	0/2280
3	M	0.32	0/1682	0.57	0/2280
3	N	0.30	0/1682	0.56	0/2280
3	O	0.26	0/1682	0.52	0/2280
All	All	0.30	0/22817	0.56	1/31027 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	100(H)	ARG	NE-CZ-NH2	-5.03	117.78	120.30

There are no chirality outliers.

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	2141	0	2074	13	0
1	B	2141	0	2075	12	0
1	C	2116	0	2044	21	0
1	D	2128	0	2058	18	0
2	H	1806	0	1749	16	0
2	I	1806	0	1749	13	0
2	J	1806	0	1749	13	0
2	K	1806	0	1749	20	0
3	L	1648	0	1614	19	0
3	M	1648	0	1614	15	0
3	N	1648	0	1614	12	0
3	O	1648	0	1614	15	0
4	E	28	0	25	0	0
5	A	14	0	13	0	0
5	D	14	0	13	0	0
6	H	1	0	0	0	0
6	J	1	0	0	0	0
6	L	2	0	0	0	0
6	M	1	0	0	0	0
6	N	2	0	0	0	0
6	O	1	0	0	0	0
All	All	22406	0	21754	165	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 4.

The worst 5 of 165 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:205:SER:HB3	1:C:210:GLN:HG3	1.64	0.80
1:A:77:ASP:OD2	1:A:141:ARG:NH1	2.16	0.78
3:M:11:LEU:HD12	3:M:11:LEU:C	2.09	0.72
1:D:299:LYS:HA	1:D:308:TYR:CG	2.28	0.69
1:D:210:GLN:HB3	3:M:12:SER:HB2	1.77	0.65

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	273/278 (98%)	266 (97%)	6 (2%)	1 (0%)	34	64
1	B	273/278 (98%)	266 (97%)	6 (2%)	1 (0%)	34	64
1	C	270/278 (97%)	264 (98%)	5 (2%)	1 (0%)	34	64
1	D	272/278 (98%)	265 (97%)	6 (2%)	1 (0%)	34	64
2	H	241/241 (100%)	234 (97%)	5 (2%)	2 (1%)	19	49
2	I	241/241 (100%)	234 (97%)	6 (2%)	1 (0%)	34	64
2	J	241/241 (100%)	235 (98%)	5 (2%)	1 (0%)	34	64
2	K	241/241 (100%)	233 (97%)	6 (2%)	2 (1%)	19	49
3	L	212/214 (99%)	208 (98%)	4 (2%)	0	100	100
3	M	212/214 (99%)	208 (98%)	4 (2%)	0	100	100
3	N	212/214 (99%)	207 (98%)	5 (2%)	0	100	100
3	O	212/214 (99%)	207 (98%)	5 (2%)	0	100	100
All	All	2900/2932 (99%)	2827 (98%)	63 (2%)	10 (0%)	34	69

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	189	LEU
2	K	189	LEU
1	A	62	ILE
1	B	62	ILE
1	C	62	ILE

### 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	244/244 (100%)	243 (100%)	1 (0%)	91	97
1	B	244/244 (100%)	243 (100%)	1 (0%)	91	97
1	C	241/244 (99%)	240 (100%)	1 (0%)	91	97
1	D	242/244 (99%)	241 (100%)	1 (0%)	91	97
2	H	202/200 (101%)	195 (96%)	7 (4%)	36	67
2	I	202/200 (101%)	199 (98%)	3 (2%)	65	85
2	J	202/200 (101%)	199 (98%)	3 (2%)	65	85
2	K	202/200 (101%)	198 (98%)	4 (2%)	55	80
3	L	187/187 (100%)	180 (96%)	7 (4%)	34	65
3	M	187/187 (100%)	180 (96%)	7 (4%)	34	65
3	N	187/187 (100%)	180 (96%)	7 (4%)	34	65
3	O	187/187 (100%)	180 (96%)	7 (4%)	34	65
All	All	2527/2524 (100%)	2478 (98%)	49 (2%)	55	81

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	78	LEU
3	N	147	GLN
2	J	92	CYS
3	N	11	LEU
3	N	188	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	138	ASN
3	L	199	GLN
3	O	189	HIS
3	L	189	HIS
3	M	55	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues 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
2	PCA	J	1	2	7,8,9	2.16	2 (28%)	9,10,12	2.22	5 (55%)
2	PCA	K	1	2	7,8,9	2.18	2 (28%)	9,10,12	2.21	5 (55%)
2	PCA	I	1	2	7,8,9	2.16	2 (28%)	9,10,12	2.23	5 (55%)
2	PCA	H	1	2	7,8,9	2.15	2 (28%)	9,10,12	2.20	5 (55%)

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
2	PCA	J	1	2	-	0/0/11/13	0/1/1/1
2	PCA	K	1	2	-	0/0/11/13	0/1/1/1
2	PCA	I	1	2	-	0/0/11/13	0/1/1/1
2	PCA	H	1	2	-	0/0/11/13	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	1	PCA	CD-N	4.66	1.46	1.34
2	J	1	PCA	CD-N	4.65	1.46	1.34
2	I	1	PCA	CD-N	4.63	1.46	1.34
2	H	1	PCA	CD-N	4.54	1.46	1.34
2	H	1	PCA	CA-N	3.26	1.50	1.46

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	1	PCA	CA-N-CD	-3.26	102.43	113.58
2	J	1	PCA	CA-N-CD	-3.16	102.75	113.58

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1	PCA	CA-N-CD	-3.16	102.77	113.58
2	J	1	PCA	OE-CD-CG	-3.10	121.35	126.76
2	K	1	PCA	OE-CD-CG	-3.10	121.36	126.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	E	1	4,1	14,14,15	0.55	0	17,19,21	0.91	1 (5%)
4	NAG	E	2	4	14,14,15	0.52	0	17,19,21	0.72	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	E	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	E	2	4	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	1	NAG	O5-C1-C2	-2.94	106.65	111.29
4	E	2	NAG	O5-C5-C6	2.15	110.57	107.20

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	2	NAG	C8-C7-N2-C2
4	E	2	NAG	O7-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 for Mogul analysis.

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	A	701	1	14,14,15	0.59	0	17,19,21	0.91	1 (5%)
5	NAG	D	701	1	14,14,15	0.57	0	17,19,21	0.88	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
5	NAG	A	701	1	-	0/6/23/26	0/1/1/1
5	NAG	D	701	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	A	701	NAG	O5-C1-C2	-3.12	106.37	111.29
5	D	701	NAG	O5-C1-C2	-3.03	106.50	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, 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	267/278 (96%)	0.32	19 (7%) 16 14	30, 54, 125, 181	0
1	B	267/278 (96%)	0.44	31 (11%) 4 4	29, 60, 146, 202	0
1	C	265/278 (95%)	0.81	43 (16%) 1 1	33, 59, 237, 285	0
1	D	266/278 (95%)	1.20	53 (19%) 1 1	31, 62, 281, 438	0
2	H	238/241 (98%)	0.09	7 (2%) 51 51	31, 50, 97, 265	0
2	I	238/241 (98%)	0.25	8 (3%) 45 43	31, 63, 133, 250	0
2	J	238/241 (98%)	0.27	11 (4%) 32 32	35, 68, 131, 246	0
2	K	238/241 (98%)	2.34	73 (30%) 0 0	36, 100, 312, 488	0
3	L	213/214 (99%)	-0.02	1 (0%) 91 91	30, 46, 75, 106	0
3	M	213/214 (99%)	-0.00	1 (0%) 91 91	27, 46, 86, 147	0
3	N	213/214 (99%)	0.11	1 (0%) 91 91	49, 69, 113, 140	0
3	O	213/214 (99%)	3.18	126 (59%) 0 0	89, 178, 276, 356	0
All	All	2869/2932 (97%)	0.74	374 (13%) 3 3	27, 62, 228, 488	0

The worst 5 of 374 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	182	VAL	22.9
2	K	138	LEU	22.1
2	K	139	GLY	16.0
2	K	140	CYS	15.4
2	K	158	ALA	14.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	PCA	K	1	8/9	0.82	0.20	122,139,147,148	0
2	PCA	J	1	8/9	0.87	0.19	86,98,103,106	0
2	PCA	I	1	8/9	0.92	0.20	61,81,87,88	0
2	PCA	H	1	8/9	0.95	0.22	79,96,104,105	0

### 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
4	NAG	E	1	14/15	0.80	0.32	109,120,130,136	0
4	NAG	E	2	14/15	0.82	0.31	120,125,127,128	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. 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	NAG	A	701	14/15	0.70	0.28	122,134,148,154	0
6	ZN	H	301	1/1	0.84	0.18	134,134,134,134	0
5	NAG	D	701	14/15	0.85	0.27	100,115,129,135	0
6	ZN	L	301	1/1	0.94	0.09	106,106,106,106	0
6	ZN	L	302	1/1	0.94	0.14	72,72,72,72	0
6	ZN	J	301	1/1	0.94	0.04	143,143,143,143	0
6	ZN	N	302	1/1	0.94	0.07	104,104,104,104	0
6	ZN	N	301	1/1	0.95	0.12	102,102,102,102	0
6	ZN	O	301	1/1	0.95	0.12	133,133,133,133	0
6	ZN	M	301	1/1	0.98	0.09	94,94,94,94	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.