



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 8, 2019 – 08:38 AM EDT

PDB ID : 6OSR  
Title : Crystal structure of Influenza hemagglutinin from strain A/Melbourne/1/194  
6(H1N1)  
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2019-05-02  
Resolution : 2.55 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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : rb-20031633  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

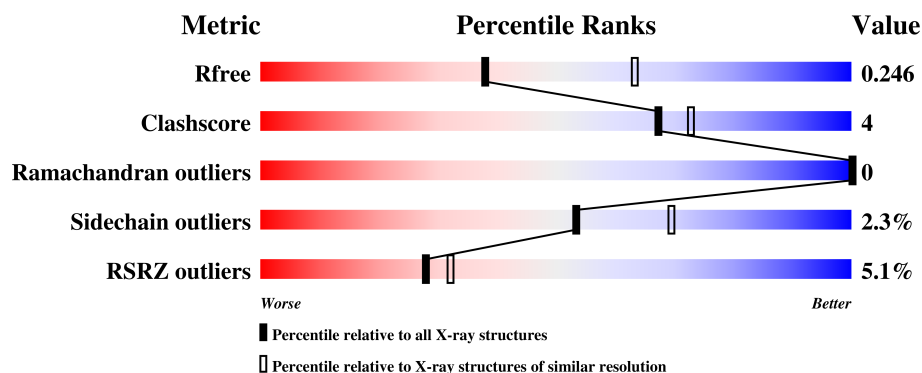
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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))
$R_{free}$	111664	1053 (2.56-2.52)
Clashscore	122126	1098 (2.56-2.52)
Ramachandran outliers	120053	1088 (2.56-2.52)
Sidechain outliers	120020	1088 (2.56-2.52)
RSRZ outliers	108989	1043 (2.56-2.52)

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	538	<div> <div>5%</div> <div>82% 7% 11%</div> </div>
1	B	538	<div> <div>4%</div> <div>77% 11% 11%</div> </div>
1	C	538	<div> <div>4%</div> <div>79% 9% 12%</div> </div>
1	D	538	<div> <div>4%</div> <div>79% 10% 10%</div> </div>
1	E	538	<div> <div>5%</div> <div>80% 8% 11%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	538	

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
2	NAG	B	604	-	-	-	X
2	NAG	C	601	-	-	-	X
2	NAG	F	602	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 23091 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	481	Total	C	N	O	S	0	0	0
			3695	2336	632	709	18			
1	B	477	Total	C	N	O	S	0	0	0
			3675	2318	632	707	18			
1	C	476	Total	C	N	O	S	0	0	0
			3650	2304	631	697	18			
1	D	483	Total	C	N	O	S	0	0	0
			3717	2349	639	711	18			
1	E	477	Total	C	N	O	S	0	0	0
			3706	2342	635	711	18			
1	F	476	Total	C	N	O	S	0	0	0
			3679	2327	631	703	18			

There are 276 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP C7S1Y2
A	2	SER	-	expression tag	UNP C7S1Y2
A	495	PHE	-	expression tag	UNP C7S1Y2
A	496	LEU	-	expression tag	UNP C7S1Y2
A	497	VAL	-	expression tag	UNP C7S1Y2
A	498	PRO	-	expression tag	UNP C7S1Y2
A	499	ARG	-	expression tag	UNP C7S1Y2
A	500	GLY	-	expression tag	UNP C7S1Y2
A	501	SER	-	expression tag	UNP C7S1Y2
A	502	PRO	-	expression tag	UNP C7S1Y2
A	503	GLY	-	expression tag	UNP C7S1Y2
A	504	SER	-	expression tag	UNP C7S1Y2
A	505	GLY	-	expression tag	UNP C7S1Y2
A	506	TYR	-	expression tag	UNP C7S1Y2
A	507	ILE	-	expression tag	UNP C7S1Y2
A	508	PRO	-	expression tag	UNP C7S1Y2
A	509	GLU	-	expression tag	UNP C7S1Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	510	ALA	-	expression tag	UNP C7S1Y2
A	511	PRO	-	expression tag	UNP C7S1Y2
A	512	ARG	-	expression tag	UNP C7S1Y2
A	513	ASP	-	expression tag	UNP C7S1Y2
A	514	GLY	-	expression tag	UNP C7S1Y2
A	515	GLN	-	expression tag	UNP C7S1Y2
A	516	ALA	-	expression tag	UNP C7S1Y2
A	517	TYR	-	expression tag	UNP C7S1Y2
A	518	VAL	-	expression tag	UNP C7S1Y2
A	519	ARG	-	expression tag	UNP C7S1Y2
A	520	LYS	-	expression tag	UNP C7S1Y2
A	521	ASP	-	expression tag	UNP C7S1Y2
A	522	GLY	-	expression tag	UNP C7S1Y2
A	523	GLU	-	expression tag	UNP C7S1Y2
A	524	TRP	-	expression tag	UNP C7S1Y2
A	525	VAL	-	expression tag	UNP C7S1Y2
A	526	LEU	-	expression tag	UNP C7S1Y2
A	527	LEU	-	expression tag	UNP C7S1Y2
A	528	SER	-	expression tag	UNP C7S1Y2
A	529	THR	-	expression tag	UNP C7S1Y2
A	530	PHE	-	expression tag	UNP C7S1Y2
A	531	LEU	-	expression tag	UNP C7S1Y2
A	532	GLY	-	expression tag	UNP C7S1Y2
A	533	HIS	-	expression tag	UNP C7S1Y2
A	534	HIS	-	expression tag	UNP C7S1Y2
A	535	HIS	-	expression tag	UNP C7S1Y2
A	536	HIS	-	expression tag	UNP C7S1Y2
A	537	HIS	-	expression tag	UNP C7S1Y2
A	538	HIS	-	expression tag	UNP C7S1Y2
B	1	GLY	-	expression tag	UNP C7S1Y2
B	2	SER	-	expression tag	UNP C7S1Y2
B	495	PHE	-	expression tag	UNP C7S1Y2
B	496	LEU	-	expression tag	UNP C7S1Y2
B	497	VAL	-	expression tag	UNP C7S1Y2
B	498	PRO	-	expression tag	UNP C7S1Y2
B	499	ARG	-	expression tag	UNP C7S1Y2
B	500	GLY	-	expression tag	UNP C7S1Y2
B	501	SER	-	expression tag	UNP C7S1Y2
B	502	PRO	-	expression tag	UNP C7S1Y2
B	503	GLY	-	expression tag	UNP C7S1Y2
B	504	SER	-	expression tag	UNP C7S1Y2
B	505	GLY	-	expression tag	UNP C7S1Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	506	TYR	-	expression tag	UNP C7S1Y2
B	507	ILE	-	expression tag	UNP C7S1Y2
B	508	PRO	-	expression tag	UNP C7S1Y2
B	509	GLU	-	expression tag	UNP C7S1Y2
B	510	ALA	-	expression tag	UNP C7S1Y2
B	511	PRO	-	expression tag	UNP C7S1Y2
B	512	ARG	-	expression tag	UNP C7S1Y2
B	513	ASP	-	expression tag	UNP C7S1Y2
B	514	GLY	-	expression tag	UNP C7S1Y2
B	515	GLN	-	expression tag	UNP C7S1Y2
B	516	ALA	-	expression tag	UNP C7S1Y2
B	517	TYR	-	expression tag	UNP C7S1Y2
B	518	VAL	-	expression tag	UNP C7S1Y2
B	519	ARG	-	expression tag	UNP C7S1Y2
B	520	LYS	-	expression tag	UNP C7S1Y2
B	521	ASP	-	expression tag	UNP C7S1Y2
B	522	GLY	-	expression tag	UNP C7S1Y2
B	523	GLU	-	expression tag	UNP C7S1Y2
B	524	TRP	-	expression tag	UNP C7S1Y2
B	525	VAL	-	expression tag	UNP C7S1Y2
B	526	LEU	-	expression tag	UNP C7S1Y2
B	527	LEU	-	expression tag	UNP C7S1Y2
B	528	SER	-	expression tag	UNP C7S1Y2
B	529	THR	-	expression tag	UNP C7S1Y2
B	530	PHE	-	expression tag	UNP C7S1Y2
B	531	LEU	-	expression tag	UNP C7S1Y2
B	532	GLY	-	expression tag	UNP C7S1Y2
B	533	HIS	-	expression tag	UNP C7S1Y2
B	534	HIS	-	expression tag	UNP C7S1Y2
B	535	HIS	-	expression tag	UNP C7S1Y2
B	536	HIS	-	expression tag	UNP C7S1Y2
B	537	HIS	-	expression tag	UNP C7S1Y2
B	538	HIS	-	expression tag	UNP C7S1Y2
C	1	GLY	-	expression tag	UNP C7S1Y2
C	2	SER	-	expression tag	UNP C7S1Y2
C	495	PHE	-	expression tag	UNP C7S1Y2
C	496	LEU	-	expression tag	UNP C7S1Y2
C	497	VAL	-	expression tag	UNP C7S1Y2
C	498	PRO	-	expression tag	UNP C7S1Y2
C	499	ARG	-	expression tag	UNP C7S1Y2
C	500	GLY	-	expression tag	UNP C7S1Y2
C	501	SER	-	expression tag	UNP C7S1Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	502	PRO	-	expression tag	UNP C7S1Y2
C	503	GLY	-	expression tag	UNP C7S1Y2
C	504	SER	-	expression tag	UNP C7S1Y2
C	505	GLY	-	expression tag	UNP C7S1Y2
C	506	TYR	-	expression tag	UNP C7S1Y2
C	507	ILE	-	expression tag	UNP C7S1Y2
C	508	PRO	-	expression tag	UNP C7S1Y2
C	509	GLU	-	expression tag	UNP C7S1Y2
C	510	ALA	-	expression tag	UNP C7S1Y2
C	511	PRO	-	expression tag	UNP C7S1Y2
C	512	ARG	-	expression tag	UNP C7S1Y2
C	513	ASP	-	expression tag	UNP C7S1Y2
C	514	GLY	-	expression tag	UNP C7S1Y2
C	515	GLN	-	expression tag	UNP C7S1Y2
C	516	ALA	-	expression tag	UNP C7S1Y2
C	517	TYR	-	expression tag	UNP C7S1Y2
C	518	VAL	-	expression tag	UNP C7S1Y2
C	519	ARG	-	expression tag	UNP C7S1Y2
C	520	LYS	-	expression tag	UNP C7S1Y2
C	521	ASP	-	expression tag	UNP C7S1Y2
C	522	GLY	-	expression tag	UNP C7S1Y2
C	523	GLU	-	expression tag	UNP C7S1Y2
C	524	TRP	-	expression tag	UNP C7S1Y2
C	525	VAL	-	expression tag	UNP C7S1Y2
C	526	LEU	-	expression tag	UNP C7S1Y2
C	527	LEU	-	expression tag	UNP C7S1Y2
C	528	SER	-	expression tag	UNP C7S1Y2
C	529	THR	-	expression tag	UNP C7S1Y2
C	530	PHE	-	expression tag	UNP C7S1Y2
C	531	LEU	-	expression tag	UNP C7S1Y2
C	532	GLY	-	expression tag	UNP C7S1Y2
C	533	HIS	-	expression tag	UNP C7S1Y2
C	534	HIS	-	expression tag	UNP C7S1Y2
C	535	HIS	-	expression tag	UNP C7S1Y2
C	536	HIS	-	expression tag	UNP C7S1Y2
C	537	HIS	-	expression tag	UNP C7S1Y2
C	538	HIS	-	expression tag	UNP C7S1Y2
D	1	GLY	-	expression tag	UNP C7S1Y2
D	2	SER	-	expression tag	UNP C7S1Y2
D	495	PHE	-	expression tag	UNP C7S1Y2
D	496	LEU	-	expression tag	UNP C7S1Y2
D	497	VAL	-	expression tag	UNP C7S1Y2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	498	PRO	-	expression tag	UNP C7S1Y2
D	499	ARG	-	expression tag	UNP C7S1Y2
D	500	GLY	-	expression tag	UNP C7S1Y2
D	501	SER	-	expression tag	UNP C7S1Y2
D	502	PRO	-	expression tag	UNP C7S1Y2
D	503	GLY	-	expression tag	UNP C7S1Y2
D	504	SER	-	expression tag	UNP C7S1Y2
D	505	GLY	-	expression tag	UNP C7S1Y2
D	506	TYR	-	expression tag	UNP C7S1Y2
D	507	ILE	-	expression tag	UNP C7S1Y2
D	508	PRO	-	expression tag	UNP C7S1Y2
D	509	GLU	-	expression tag	UNP C7S1Y2
D	510	ALA	-	expression tag	UNP C7S1Y2
D	511	PRO	-	expression tag	UNP C7S1Y2
D	512	ARG	-	expression tag	UNP C7S1Y2
D	513	ASP	-	expression tag	UNP C7S1Y2
D	514	GLY	-	expression tag	UNP C7S1Y2
D	515	GLN	-	expression tag	UNP C7S1Y2
D	516	ALA	-	expression tag	UNP C7S1Y2
D	517	TYR	-	expression tag	UNP C7S1Y2
D	518	VAL	-	expression tag	UNP C7S1Y2
D	519	ARG	-	expression tag	UNP C7S1Y2
D	520	LYS	-	expression tag	UNP C7S1Y2
D	521	ASP	-	expression tag	UNP C7S1Y2
D	522	GLY	-	expression tag	UNP C7S1Y2
D	523	GLU	-	expression tag	UNP C7S1Y2
D	524	TRP	-	expression tag	UNP C7S1Y2
D	525	VAL	-	expression tag	UNP C7S1Y2
D	526	LEU	-	expression tag	UNP C7S1Y2
D	527	LEU	-	expression tag	UNP C7S1Y2
D	528	SER	-	expression tag	UNP C7S1Y2
D	529	THR	-	expression tag	UNP C7S1Y2
D	530	PHE	-	expression tag	UNP C7S1Y2
D	531	LEU	-	expression tag	UNP C7S1Y2
D	532	GLY	-	expression tag	UNP C7S1Y2
D	533	HIS	-	expression tag	UNP C7S1Y2
D	534	HIS	-	expression tag	UNP C7S1Y2
D	535	HIS	-	expression tag	UNP C7S1Y2
D	536	HIS	-	expression tag	UNP C7S1Y2
D	537	HIS	-	expression tag	UNP C7S1Y2
D	538	HIS	-	expression tag	UNP C7S1Y2
E	1	GLY	-	expression tag	UNP C7S1Y2

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	2	SER	-	expression tag	UNP C7S1Y2
E	495	PHE	-	expression tag	UNP C7S1Y2
E	496	LEU	-	expression tag	UNP C7S1Y2
E	497	VAL	-	expression tag	UNP C7S1Y2
E	498	PRO	-	expression tag	UNP C7S1Y2
E	499	ARG	-	expression tag	UNP C7S1Y2
E	500	GLY	-	expression tag	UNP C7S1Y2
E	501	SER	-	expression tag	UNP C7S1Y2
E	502	PRO	-	expression tag	UNP C7S1Y2
E	503	GLY	-	expression tag	UNP C7S1Y2
E	504	SER	-	expression tag	UNP C7S1Y2
E	505	GLY	-	expression tag	UNP C7S1Y2
E	506	TYR	-	expression tag	UNP C7S1Y2
E	507	ILE	-	expression tag	UNP C7S1Y2
E	508	PRO	-	expression tag	UNP C7S1Y2
E	509	GLU	-	expression tag	UNP C7S1Y2
E	510	ALA	-	expression tag	UNP C7S1Y2
E	511	PRO	-	expression tag	UNP C7S1Y2
E	512	ARG	-	expression tag	UNP C7S1Y2
E	513	ASP	-	expression tag	UNP C7S1Y2
E	514	GLY	-	expression tag	UNP C7S1Y2
E	515	GLN	-	expression tag	UNP C7S1Y2
E	516	ALA	-	expression tag	UNP C7S1Y2
E	517	TYR	-	expression tag	UNP C7S1Y2
E	518	VAL	-	expression tag	UNP C7S1Y2
E	519	ARG	-	expression tag	UNP C7S1Y2
E	520	LYS	-	expression tag	UNP C7S1Y2
E	521	ASP	-	expression tag	UNP C7S1Y2
E	522	GLY	-	expression tag	UNP C7S1Y2
E	523	GLU	-	expression tag	UNP C7S1Y2
E	524	TRP	-	expression tag	UNP C7S1Y2
E	525	VAL	-	expression tag	UNP C7S1Y2
E	526	LEU	-	expression tag	UNP C7S1Y2
E	527	LEU	-	expression tag	UNP C7S1Y2
E	528	SER	-	expression tag	UNP C7S1Y2
E	529	THR	-	expression tag	UNP C7S1Y2
E	530	PHE	-	expression tag	UNP C7S1Y2
E	531	LEU	-	expression tag	UNP C7S1Y2
E	532	GLY	-	expression tag	UNP C7S1Y2
E	533	HIS	-	expression tag	UNP C7S1Y2
E	534	HIS	-	expression tag	UNP C7S1Y2
E	535	HIS	-	expression tag	UNP C7S1Y2

*Continued on next page...*

*Continued from previous page...*

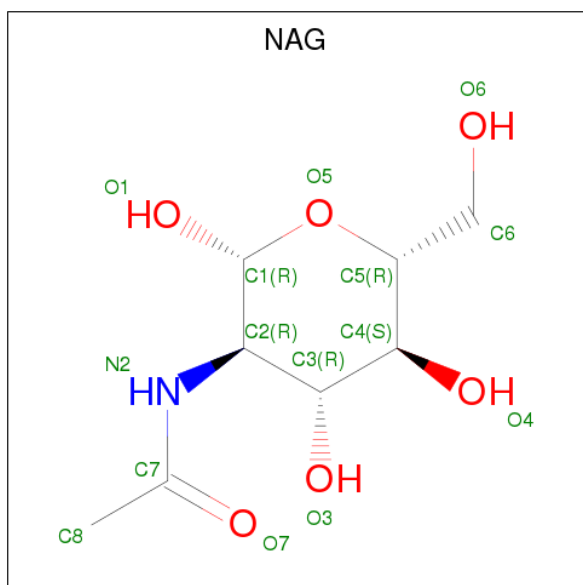
Chain	Residue	Modelled	Actual	Comment	Reference
E	536	HIS	-	expression tag	UNP C7S1Y2
E	537	HIS	-	expression tag	UNP C7S1Y2
E	538	HIS	-	expression tag	UNP C7S1Y2
F	1	GLY	-	expression tag	UNP C7S1Y2
F	2	SER	-	expression tag	UNP C7S1Y2
F	495	PHE	-	expression tag	UNP C7S1Y2
F	496	LEU	-	expression tag	UNP C7S1Y2
F	497	VAL	-	expression tag	UNP C7S1Y2
F	498	PRO	-	expression tag	UNP C7S1Y2
F	499	ARG	-	expression tag	UNP C7S1Y2
F	500	GLY	-	expression tag	UNP C7S1Y2
F	501	SER	-	expression tag	UNP C7S1Y2
F	502	PRO	-	expression tag	UNP C7S1Y2
F	503	GLY	-	expression tag	UNP C7S1Y2
F	504	SER	-	expression tag	UNP C7S1Y2
F	505	GLY	-	expression tag	UNP C7S1Y2
F	506	TYR	-	expression tag	UNP C7S1Y2
F	507	ILE	-	expression tag	UNP C7S1Y2
F	508	PRO	-	expression tag	UNP C7S1Y2
F	509	GLU	-	expression tag	UNP C7S1Y2
F	510	ALA	-	expression tag	UNP C7S1Y2
F	511	PRO	-	expression tag	UNP C7S1Y2
F	512	ARG	-	expression tag	UNP C7S1Y2
F	513	ASP	-	expression tag	UNP C7S1Y2
F	514	GLY	-	expression tag	UNP C7S1Y2
F	515	GLN	-	expression tag	UNP C7S1Y2
F	516	ALA	-	expression tag	UNP C7S1Y2
F	517	TYR	-	expression tag	UNP C7S1Y2
F	518	VAL	-	expression tag	UNP C7S1Y2
F	519	ARG	-	expression tag	UNP C7S1Y2
F	520	LYS	-	expression tag	UNP C7S1Y2
F	521	ASP	-	expression tag	UNP C7S1Y2
F	522	GLY	-	expression tag	UNP C7S1Y2
F	523	GLU	-	expression tag	UNP C7S1Y2
F	524	TRP	-	expression tag	UNP C7S1Y2
F	525	VAL	-	expression tag	UNP C7S1Y2
F	526	LEU	-	expression tag	UNP C7S1Y2
F	527	LEU	-	expression tag	UNP C7S1Y2
F	528	SER	-	expression tag	UNP C7S1Y2
F	529	THR	-	expression tag	UNP C7S1Y2
F	530	PHE	-	expression tag	UNP C7S1Y2
F	531	LEU	-	expression tag	UNP C7S1Y2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	532	GLY	-	expression tag	UNP C7S1Y2
F	533	HIS	-	expression tag	UNP C7S1Y2
F	534	HIS	-	expression tag	UNP C7S1Y2
F	535	HIS	-	expression tag	UNP C7S1Y2
F	536	HIS	-	expression tag	UNP C7S1Y2
F	537	HIS	-	expression tag	UNP C7S1Y2
F	538	HIS	-	expression tag	UNP C7S1Y2

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



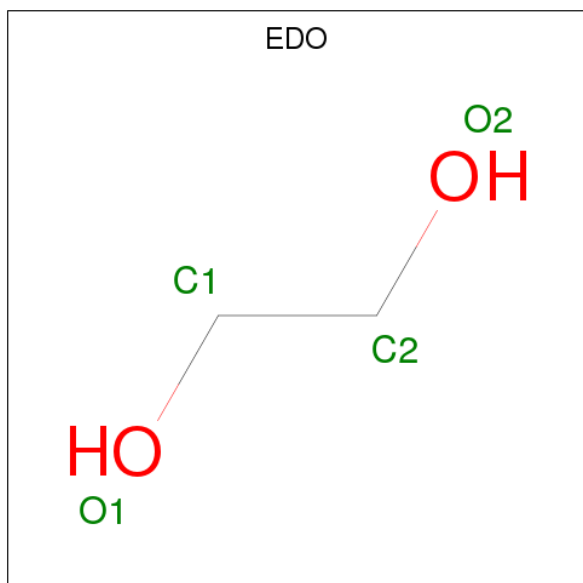
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

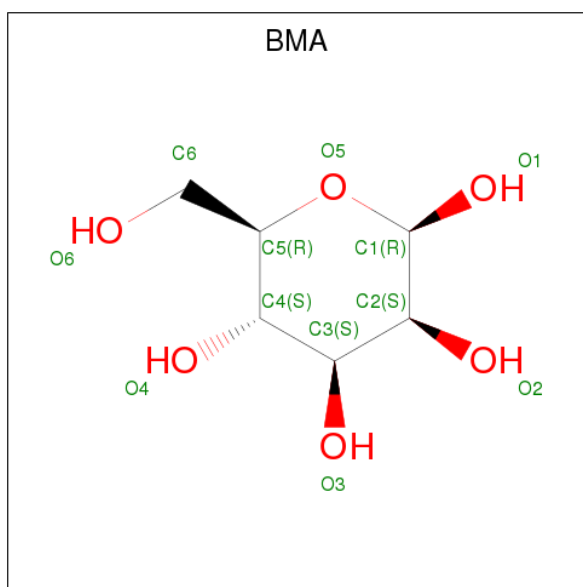
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	E	1	Total C O 4 2 2	0	0
3	E	1	Total C O 4 2 2	0	0
3	F	1	Total C O 4 2 2	0	0

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

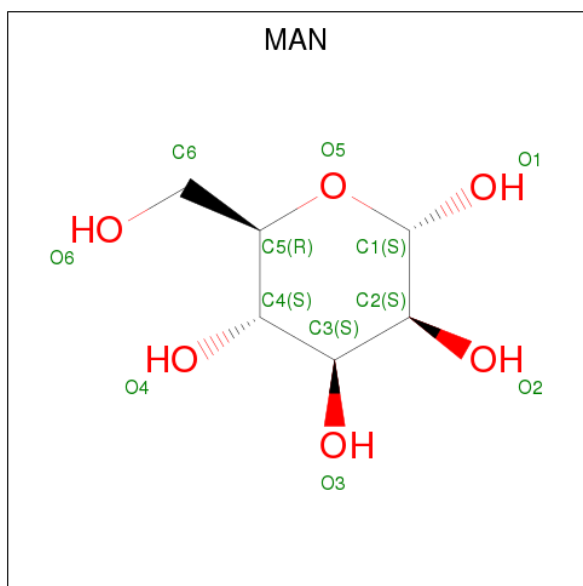
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total K 1 1	0	0
4	D	1	Total K 1 1	0	0

- Molecule 5 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

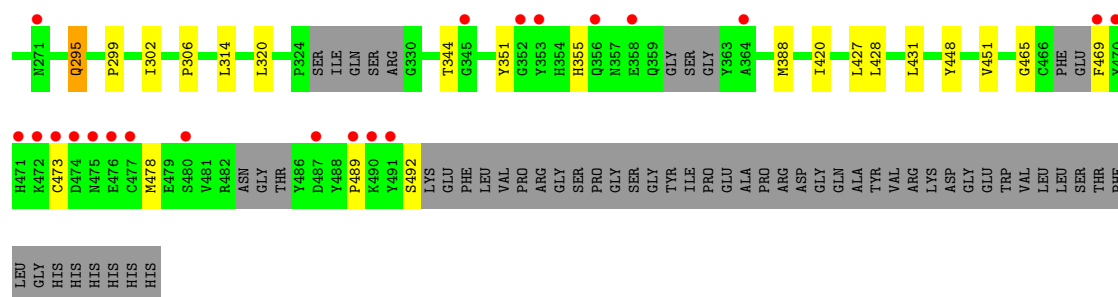
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	1	Total Na 1 1	0	0
7	F	1	Total Na 1 1	0	0
7	E	1	Total Na 1 1	0	0

- Molecule 8 is water.

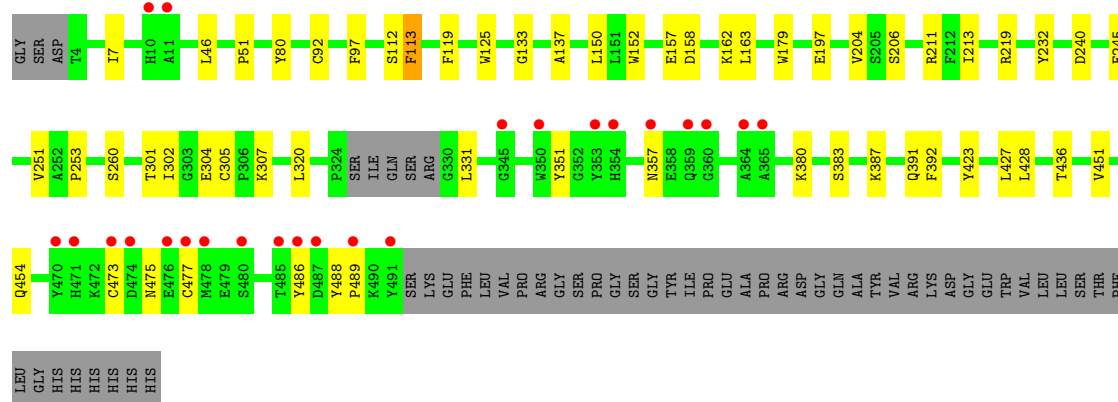
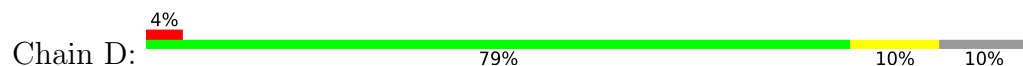
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	100	Total O 101 101	0	1
8	B	84	Total O 84 84	0	0
8	C	112	Total O 112 112	0	0
8	D	131	Total O 131 131	0	0
8	E	94	Total O 95 95	0	1
8	F	123	Total O 123 123	0	0



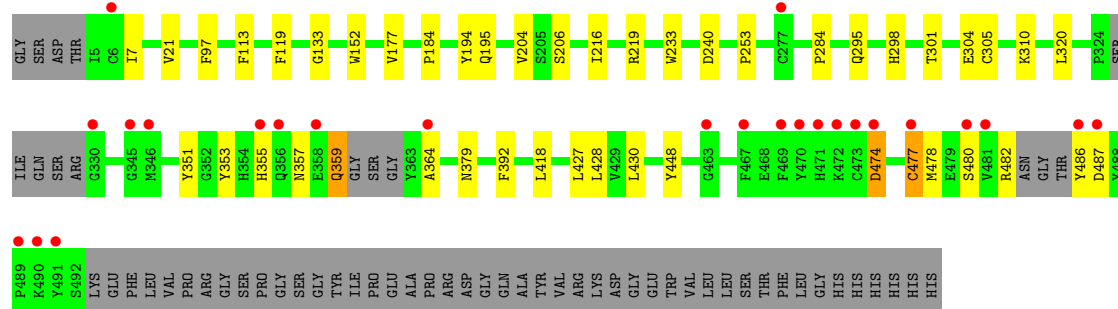
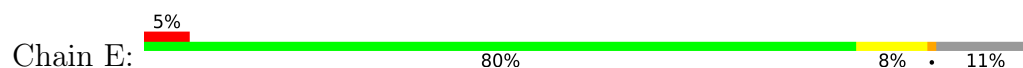




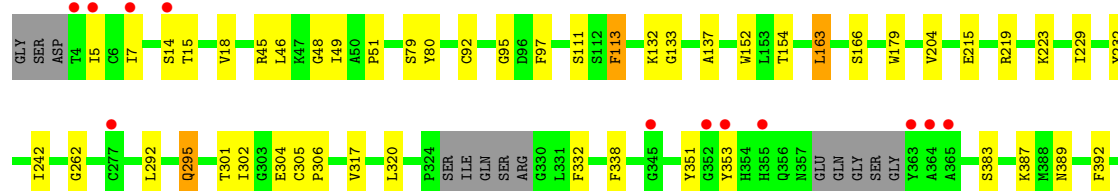
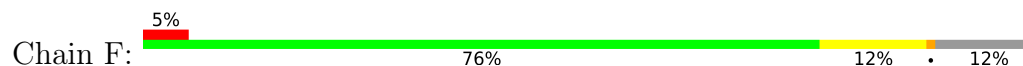
• Molecule 1: Hemagglutinin

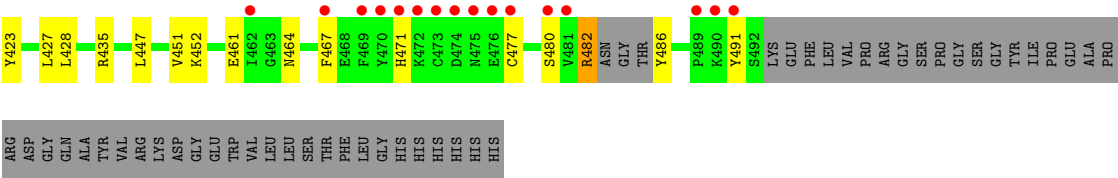


• Molecule 1: Hemagglutinin



• Molecule 1: Hemagglutinin





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.10Å 113.97Å 246.38Å 90.00° 91.44° 90.00°	Depositor
Resolution (Å)	48.74 – 2.55 48.74 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.74-2.55) 99.9 (48.74-2.55)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.54Å)	Xtriage
Refinement program	PHENIX dev_3409	Depositor
R, $R_{free}$	0.191 , 0.246 0.191 , 0.246	Depositor DCC
$R_{free}$ test set	1955 reflections (1.49%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.2	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 59.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.026 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	23091	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.95% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, NA, K, EDO, 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	0.45	0/3784	0.60	0/5151
1	B	0.47	0/3761	0.59	0/5115
1	C	0.43	0/3734	0.60	0/5077
1	D	0.45	0/3808	0.62	0/5179
1	E	0.42	0/3794	0.60	0/5154
1	F	0.41	0/3766	0.61	0/5119
All	All	0.44	0/22647	0.60	0/30795

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3695	0	3430	26	0
1	B	3675	0	3445	37	0
1	C	3650	0	3420	29	0
1	D	3717	0	3469	37	0
1	E	3706	0	3491	30	0
1	F	3679	0	3466	46	0
2	A	56	0	52	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	70	0	63	1	0
2	C	42	0	39	0	0
2	D	14	0	13	0	0
2	E	28	0	26	0	0
2	F	42	0	39	0	0
3	A	8	0	12	0	0
3	C	8	0	12	2	0
3	D	16	0	24	2	0
3	E	8	0	12	0	0
3	F	4	0	6	0	0
4	A	1	0	0	0	0
4	D	1	0	0	0	0
5	B	11	0	9	0	0
6	B	11	0	10	0	0
7	C	1	0	0	0	0
7	E	1	0	0	0	0
7	F	1	0	0	0	0
8	A	101	0	0	1	0
8	B	84	0	0	1	0
8	C	112	0	0	1	0
8	D	131	0	0	0	0
8	E	95	0	0	0	0
8	F	123	0	0	0	0
All	All	23091	0	21038	183	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 183 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:163:LEU:HA	3:C:604:EDO:H22	1.58	0.84
1:A:155:LYS:HD2	1:A:195:GLN:HG2	1.66	0.76
1:A:427:LEU:HD21	1:C:428:LEU:HD13	1.67	0.76
1:D:219:ARG:HG2	1:F:204:VAL:HG11	1.67	0.75
1:D:204:VAL:HG11	1:E:219:ARG:HG2	1.69	0.74

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	475/538 (88%)	462 (97%)	13 (3%)	0	100	100
1	B	471/538 (88%)	453 (96%)	18 (4%)	0	100	100
1	C	466/538 (87%)	452 (97%)	14 (3%)	0	100	100
1	D	479/538 (89%)	463 (97%)	16 (3%)	0	100	100
1	E	469/538 (87%)	456 (97%)	13 (3%)	0	100	100
1	F	468/538 (87%)	454 (97%)	14 (3%)	0	100	100
All	All	2828/3228 (88%)	2740 (97%)	88 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	387/467 (83%)	377 (97%)	10 (3%)	49	65
1	B	391/467 (84%)	383 (98%)	8 (2%)	58	74
1	C	386/467 (83%)	376 (97%)	10 (3%)	49	65
1	D	390/467 (84%)	383 (98%)	7 (2%)	62	77
1	E	397/467 (85%)	388 (98%)	9 (2%)	53	69
1	F	392/467 (84%)	381 (97%)	11 (3%)	47	62
All	All	2343/2802 (84%)	2288 (98%)	55 (2%)	53	69

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	320	LEU
1	D	162	LYS
1	F	320	LEU
1	C	351	TYR
1	C	473	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 36 ligands modelled in this entry, 5 are monoatomic - leaving 31 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	$\# Z  > 2$
2	NAG	A	601	1	14,14,15	0.49	0	17,19,21	0.75	1 (5%)
2	NAG	A	602	1	14,14,15	0.46	0	17,19,21	0.41	0
2	NAG	A	603	1	14,14,15	0.48	0	17,19,21	0.87	1 (5%)
2	NAG	A	604	1	14,14,15	0.45	0	17,19,21	0.61	0
3	EDO	A	605	-	3,3,3	0.56	0	2,2,2	0.31	0
3	EDO	A	606	-	3,3,3	0.46	0	2,2,2	0.61	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	B	601	1	14,14,15	0.44	0	17,19,21	0.39	0
2	NAG	B	602	1	14,14,15	0.64	1 (7%)	17,19,21	1.24	1 (5%)
2	NAG	B	603	1,2	14,14,15	0.25	0	17,19,21	0.48	0
2	NAG	B	604	2,5	14,14,15	0.41	0	17,19,21	0.47	0
5	BMA	B	605	2,6	11,11,12	0.87	0	15,15,17	0.96	1 (6%)
6	MAN	B	606	5	11,11,12	1.07	1 (9%)	15,15,17	1.30	2 (13%)
2	NAG	B	607	1	14,14,15	0.60	0	17,19,21	0.63	1 (5%)
2	NAG	C	601	1	14,14,15	0.37	0	17,19,21	0.45	0
2	NAG	C	602	1	14,14,15	0.42	0	17,19,21	0.47	0
2	NAG	C	603	1	14,14,15	0.55	0	17,19,21	0.47	0
3	EDO	C	604	-	3,3,3	0.45	0	2,2,2	0.25	0
3	EDO	C	605	-	3,3,3	0.48	0	2,2,2	0.37	0
2	NAG	D	601	1	14,14,15	0.35	0	17,19,21	0.41	0
3	EDO	D	602	-	3,3,3	0.40	0	2,2,2	0.40	0
3	EDO	D	603	-	3,3,3	0.40	0	2,2,2	0.57	0
3	EDO	D	604	-	3,3,3	0.61	0	2,2,2	0.39	0
3	EDO	D	605	-	3,3,3	0.52	0	2,2,2	0.22	0
2	NAG	E	601	1	14,14,15	0.55	0	17,19,21	0.49	0
2	NAG	E	602	1	14,14,15	0.31	0	17,19,21	0.45	0
3	EDO	E	603	-	3,3,3	0.58	0	2,2,2	0.20	0
3	EDO	E	604	-	3,3,3	0.50	0	2,2,2	0.25	0
2	NAG	F	601	1	14,14,15	0.41	0	17,19,21	0.38	0
2	NAG	F	602	1	14,14,15	0.71	1 (7%)	17,19,21	0.49	0
3	EDO	F	603	-	3,3,3	0.49	0	2,2,2	0.28	0
2	NAG	F	604	1	14,14,15	0.67	0	17,19,21	1.70	4 (23%)

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	NAG	A	601	1	-	0/6/23/26	0/1/1/1
2	NAG	A	602	1	-	0/6/23/26	0/1/1/1
2	NAG	A	603	1	-	0/6/23/26	0/1/1/1
2	NAG	A	604	1	-	0/6/23/26	0/1/1/1
3	EDO	A	605	-	-	0/1/1/1	0/0/0/0
3	EDO	A	606	-	-	0/1/1/1	0/0/0/0
2	NAG	B	601	1	-	0/6/23/26	0/1/1/1
2	NAG	B	602	1	-	0/6/23/26	0/1/1/1
2	NAG	B	603	1,2	-	0/6/23/26	0/1/1/1

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	604	2,5	-	0/6/23/26	0/1/1/1
5	BMA	B	605	2,6	-	0/2/19/22	0/1/1/1
6	MAN	B	606	5	-	0/2/19/22	0/1/1/1
2	NAG	B	607	1	-	0/6/23/26	0/1/1/1
2	NAG	C	601	1	-	0/6/23/26	0/1/1/1
2	NAG	C	602	1	-	0/6/23/26	0/1/1/1
2	NAG	C	603	1	-	0/6/23/26	0/1/1/1
3	EDO	C	604	-	-	0/1/1/1	0/0/0/0
3	EDO	C	605	-	-	0/1/1/1	0/0/0/0
2	NAG	D	601	1	-	0/6/23/26	0/1/1/1
3	EDO	D	602	-	-	0/1/1/1	0/0/0/0
3	EDO	D	603	-	-	0/1/1/1	0/0/0/0
3	EDO	D	604	-	-	0/1/1/1	0/0/0/0
3	EDO	D	605	-	-	0/1/1/1	0/0/0/0
2	NAG	E	601	1	-	0/6/23/26	0/1/1/1
2	NAG	E	602	1	-	0/6/23/26	0/1/1/1
3	EDO	E	603	-	-	0/1/1/1	0/0/0/0
3	EDO	E	604	-	-	0/1/1/1	0/0/0/0
2	NAG	F	601	1	-	0/6/23/26	0/1/1/1
2	NAG	F	602	1	-	0/6/23/26	0/1/1/1
3	EDO	F	603	-	-	0/1/1/1	0/0/0/0
2	NAG	F	604	1	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	606	MAN	O5-C1	-2.06	1.40	1.43
2	B	602	NAG	C1-C2	2.02	1.55	1.52
2	F	602	NAG	C1-C2	2.48	1.56	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	604	NAG	O5-C1-C2	-3.52	105.82	111.36
6	B	606	MAN	O2-C2-C3	-2.88	104.54	110.16
5	B	605	BMA	O2-C2-C3	-2.20	105.86	110.16
2	F	604	NAG	C2-N2-C7	2.05	125.85	122.92
2	B	607	NAG	C1-O5-C5	2.26	115.27	112.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	NAG	1	0
2	B	602	NAG	1	0
3	C	604	EDO	1	0
3	C	605	EDO	1	0
3	D	602	EDO	1	0
3	D	604	EDO	1	0

## 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	481/538 (89%)	0.18	26 (5%) 26 30	24, 47, 111, 147	0
1	B	477/538 (88%)	0.14	20 (4%) 36 42	27, 52, 99, 125	0
1	C	476/538 (88%)	0.20	22 (4%) 32 39	24, 48, 110, 190	0
1	D	483/538 (89%)	0.14	24 (4%) 29 33	25, 43, 109, 130	0
1	E	477/538 (88%)	0.16	25 (5%) 27 31	24, 49, 107, 135	0
1	F	476/538 (88%)	0.20	28 (5%) 22 26	24, 50, 115, 157	0
All	All	2870/3228 (88%)	0.17	145 (5%) 28 32	24, 49, 110, 190	0

The worst 5 of 145 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	473	CYS	6.1
1	F	473	CYS	5.8
1	B	364	ALA	5.7
1	C	473	CYS	5.7
1	C	477	CYS	5.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	NAG	A	603	14/15	0.48	0.38	91,106,115,116	0
2	NAG	F	602	14/15	0.61	0.43	78,102,111,115	0
2	NAG	B	602	14/15	0.63	0.28	80,94,109,112	0
2	NAG	E	601	14/15	0.66	0.38	91,104,114,119	0
2	NAG	A	601	14/15	0.67	0.35	77,92,97,99	0
2	NAG	A	602	14/15	0.68	0.33	87,101,107,108	0
2	NAG	B	601	14/15	0.68	0.37	81,97,107,111	0
2	NAG	F	601	14/15	0.70	0.36	74,91,99,99	0
2	NAG	C	601	14/15	0.71	0.42	86,108,121,122	0
2	NAG	C	602	14/15	0.76	0.17	78,94,97,99	0
2	NAG	F	604	14/15	0.77	0.31	89,102,111,113	0
2	NAG	D	601	14/15	0.77	0.36	74,92,109,113	0
2	NAG	B	607	14/15	0.79	0.34	74,96,99,102	0
2	NAG	A	604	14/15	0.80	0.46	79,94,108,117	0
2	NAG	B	604	14/15	0.80	0.49	94,108,114,115	0
3	EDO	D	604	4/4	0.80	0.32	58,60,64,70	0
2	NAG	C	603	14/15	0.86	0.28	75,94,98,102	0
2	NAG	E	602	14/15	0.86	0.38	75,90,101,103	0
3	EDO	A	605	4/4	0.88	0.21	51,61,66,67	0
3	EDO	E	603	4/4	0.88	0.23	43,44,46,54	0
5	BMA	B	605	11/12	0.89	0.28	72,94,103,104	0
2	NAG	B	603	14/15	0.89	0.38	66,86,95,105	0
7	NA	F	605	1/1	0.90	0.35	40,40,40,40	0
3	EDO	D	603	4/4	0.90	0.23	59,59,60,67	0
7	NA	C	606	1/1	0.91	0.34	44,44,44,44	0
3	EDO	C	604	4/4	0.92	0.17	55,57,58,59	0
3	EDO	D	605	4/4	0.94	0.23	58,59,61,62	0
3	EDO	F	603	4/4	0.94	0.19	50,51,52,52	0
6	MAN	B	606	11/12	0.94	0.13	37,54,63,64	0
3	EDO	C	605	4/4	0.95	0.20	53,55,56,64	0
3	EDO	E	604	4/4	0.95	0.22	56,56,59,60	0
3	EDO	D	602	4/4	0.96	0.14	45,47,48,57	0
7	NA	E	605	1/1	0.96	0.38	47,47,47,47	0
3	EDO	A	606	4/4	0.96	0.20	43,46,52,57	0
4	K	A	607	1/1	0.97	0.25	48,48,48,48	0
4	K	D	606	1/1	0.97	0.18	61,61,61,61	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.