



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 2, 2023 – 08:35 AM EDT

PDB ID : 6N5A  
Title : Crystal structure of an equine H7 hemagglutinin from A/equine/NY/49/73 (H7N7)  
Authors : Zhu, X.; Wilson, I.A.  
Deposited on : 2018-11-21  
Resolution : 3.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : **FAILED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

## 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 3.30 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 3940 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 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	315	2448	1531	435	467	15	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP A0A348FV55
A	8	ASP	-	expression tag	UNP A0A348FV55
A	9	PRO	-	expression tag	UNP A0A348FV55
A	10	GLY	-	expression tag	UNP A0A348FV55

- Molecule 2 is a protein called HEMAGGLUTININ HA2 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	170	1390	862	247	276	5	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	179	GLU	-	expression tag	UNP A0A348FV55
B	180	SER	-	expression tag	UNP A0A348FV55
B	181	GLY	-	expression tag	UNP A0A348FV55
B	182	ARG	-	expression tag	UNP A0A348FV55
B	183	LEU	-	expression tag	UNP A0A348FV55
B	184	VAL	-	expression tag	UNP A0A348FV55
B	185	PRO	-	expression tag	UNP A0A348FV55
B	186	ARG	-	expression tag	UNP A0A348FV55

- Molecule 3 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	
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



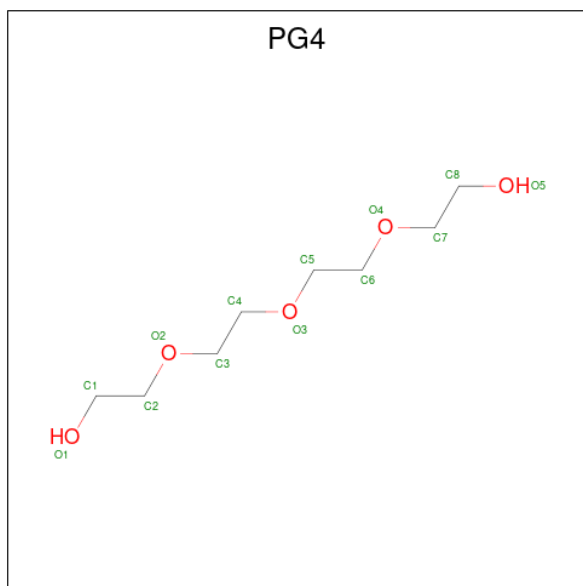
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



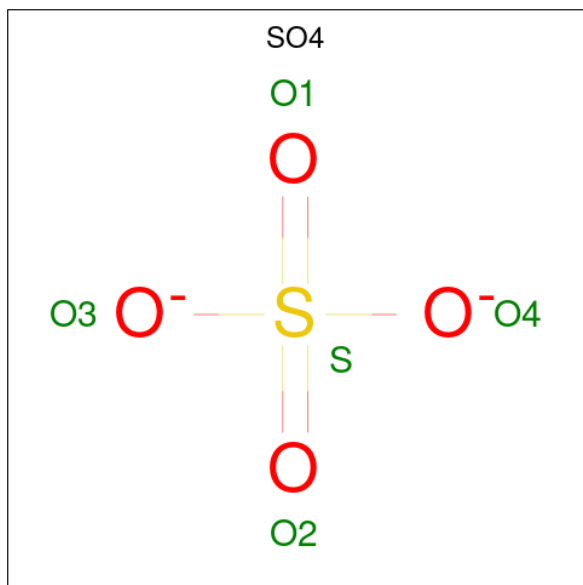
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			13	8	5		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 9 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
9	A	21	Total 21	O 21	0	0
9	B	1	Total 1	O 1	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.40Å 87.40Å 524.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.24 – 3.30	Depositor
% Data completeness (in resolution range)	99.7 (46.24-3.30)	Depositor
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.52 (at 3.25Å)	Xtrriage
Refinement program	REFMAC 5.8.0232	Depositor
R, $R_{free}$	0.200 , 0.237	Depositor
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtrriage
Anisotropy	0.051	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	3940	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.92% 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.



## 4 Model quality [i](#)

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

MolProbity failed to run properly - this section is therefore empty.

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

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

10 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	B	201	2	14,14,15	0.96	1 (7%)	17,19,21	2.11	7 (41%)
4	PGE	A	402	-	9,9,9	0.46	0	8,8,8	0.40	0
5	PEG	A	403	-	6,6,6	0.42	0	5,5,5	0.45	0
7	PG4	A	407	-	12,12,12	0.59	0	11,11,11	0.45	0
6	EDO	A	404	-	3,3,3	0.20	0	2,2,2	0.45	0
8	SO4	A	408	-	4,4,4	0.36	0	6,6,6	0.25	0
8	SO4	A	409	-	4,4,4	0.23	0	6,6,6	0.14	0
3	NAG	A	401	1	14,14,15	0.78	0	17,19,21	1.97	6 (35%)
6	EDO	A	406	-	3,3,3	0.30	0	2,2,2	0.35	0
6	EDO	A	405	-	3,3,3	0.18	0	2,2,2	0.10	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	B	201	2	-	0/6/23/26	0/1/1/1
4	PGE	A	402	-	-	4/7/7/7	-
5	PEG	A	403	-	-	3/4/4/4	-
7	PG4	A	407	-	-	4/10/10/10	-
6	EDO	A	404	-	-	1/1/1/1	-
3	NAG	A	401	1	-	2/6/23/26	0/1/1/1
6	EDO	A	406	-	-	1/1/1/1	-
6	EDO	A	405	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	201	NAG	O4-C4	2.15	1.48	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C1-O5-C5	4.69	118.55	112.19
3	B	201	NAG	C4-C3-C2	4.24	117.23	111.02
3	B	201	NAG	C3-C4-C5	-3.66	103.70	110.24

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C1-C2-N2	-3.23	104.96	110.49
3	B	201	NAG	O5-C5-C6	3.11	112.07	107.20

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	402	PGE	C3-C4-O3-C5
4	A	402	PGE	O2-C3-C4-O3
3	A	401	NAG	O5-C5-C6-O6
3	A	401	NAG	C4-C5-C6-O6
4	A	402	PGE	O3-C5-C6-O4

There are no ring outliers.

No monomer is involved in short contacts.

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

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data [i](#)

### 5.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

### 5.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.