WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2020 – 09:12 PM BST

PDB ID	:	1E8U
Title	:	Structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase
Authors	:	Crennell, S.; Takimoto, T.; Portner, A.; Taylor, G.
Deposited on	:	2000-10-01
Resolution	:	2.00 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
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)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.00 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.
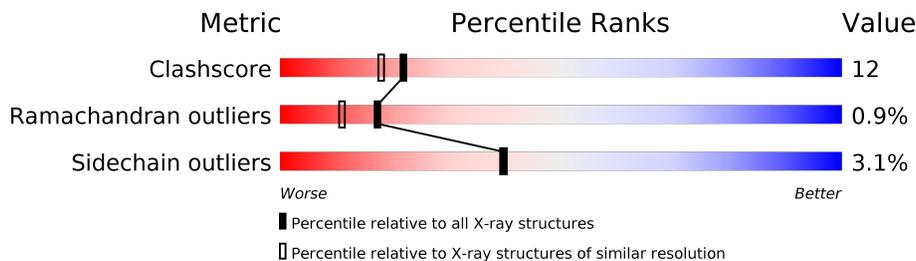
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	454	77% 20% • •
1	B	454	75% 23% ••

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	A	1573	X	-	-	-
4	NAG	B	1575	X	-	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7401 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-NEURAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	447	Total 3446	C 2172	N 588	O 667	S 19	0	0	1
1	B	450	Total 3468	C 2185	N 594	O 670	S 19	0	0	1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Ca 1	0	0
2	A	1	Total 1	Ca 1	0	0

- Molecule 3 is N-acetyl-beta-neuraminic acid (three-letter code: SLB) (formula: $C_{11}H_{19}NO_9$).
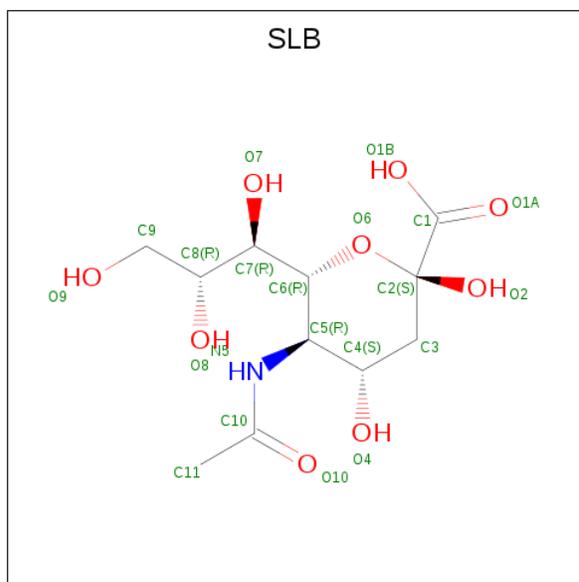
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 21	C 11	N 1	O 9	0	0
3	B	1	Total 21	C 11	N 1	O 9	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
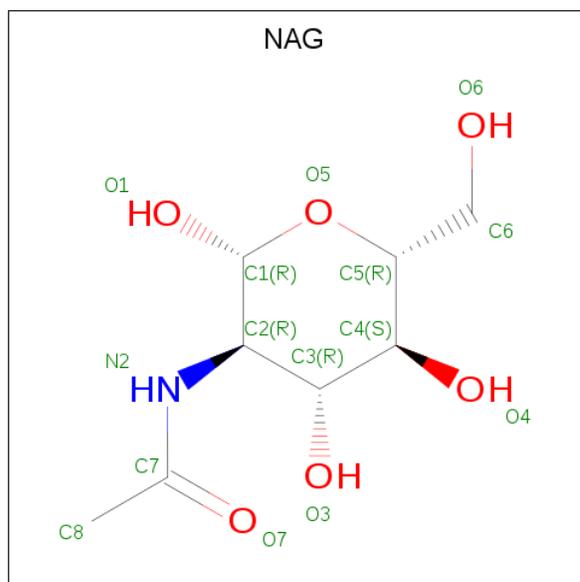
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 15	C 8	N 1	O 6	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 15	C 8	N 1	O 6	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 15	C 8	N 1	O 6	0	0

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	232	Total 232	O 232	0	0
5	B	138	Total 138	O 138	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first
graphic for a chain summarises the proportions of the various outlier classes displayed in the second
graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.
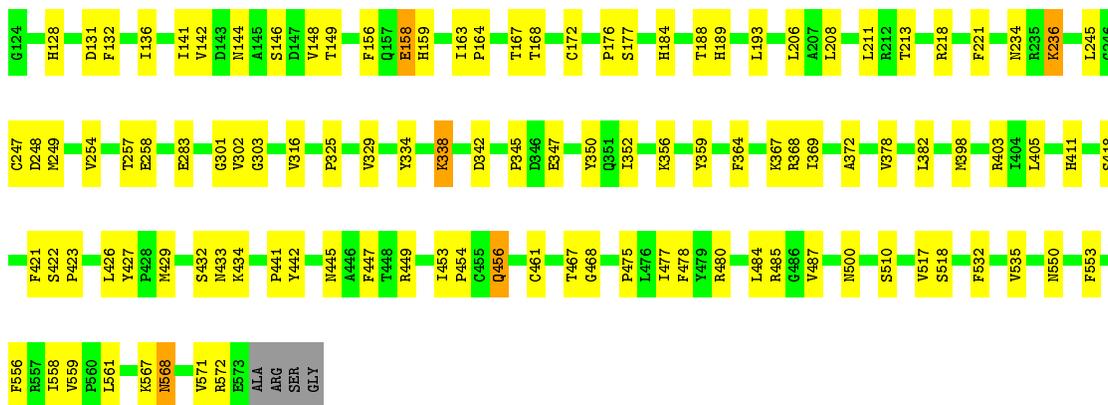
Note EDS was not executed.

• Molecule 1: HEMAGGLUTININ-NEURAMINIDASE
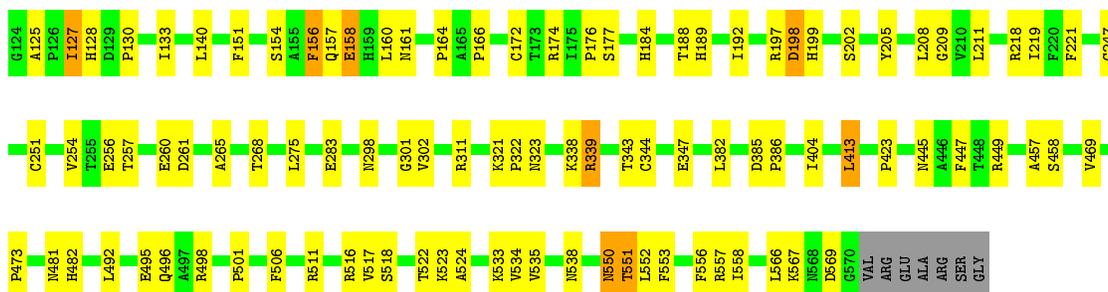
• Molecule 1: HEMAGGLUTININ-NEURAMINIDASE

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.56Å 78.49Å 199.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.00	Depositor
% Data completeness (in resolution range)	97.0 (6.00-2.00)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.232 , 0.267	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7401	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.35	0/3531	0.68	1/4804 (0.0%)
1	B	0.32	0/3553	0.61	0/4833
All	All	0.34	0/7084	0.65	1/9637 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	251	CYS	CA-CB-SG	5.33	123.59	114.00

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	3446	0	3341	74	0
1	B	3468	0	3365	86	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	21	0	18	1	0
3	B	21	0	18	2	0
4	A	44	0	43	1	0
4	B	29	0	28	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	232	0	0	1	0
5	B	138	0	0	2	0
All	All	7401	0	6813	163	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 12.

The worst 5 of 163 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:B:156:PHE:HE2	1:B:517:VAL:HA	1.25	1.00
1:A:511:ARG:HH11	1:A:511:ARG:HG2	1.29	0.93
1:B:128:HIS:HD2	1:B:211:LEU:H	1.19	0.87
1:B:156:PHE:CE2	1:B:517:VAL:HA	2.13	0.81
1:B:131:ASP:O	1:B:535:VAL:HG22	1.83	0.79

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	445/454 (98%)	417 (94%)	23 (5%)	5 (1%)	14	8
1	B	448/454 (99%)	415 (93%)	30 (7%)	3 (1%)	22	16
All	All	893/908 (98%)	832 (93%)	53 (6%)	8 (1%)	17	11

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	127	ILE
1	B	158	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	198	ASP
1	A	156	PHE
1	A	551	THR

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	387/392 (99%)	374 (97%)	13 (3%)	37	36
1	B	389/392 (99%)	378 (97%)	11 (3%)	43	44
All	All	776/784 (99%)	752 (97%)	24 (3%)	40	40

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	447	PHE
1	B	189	HIS
1	B	500	ASN
1	A	449	ARG
1	A	550	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	496	GLN
1	A	538	ASN
1	B	263	ASN
1	A	415	GLN
1	B	351	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	#$|Z| > 2$
3	SLB	B	1573	-	18,21,21	0.63	0	21,31,31	0.75	0
4	NAG	A	1573	-	15,15,15	0.40	0	21,21,21	0.53	0
4	NAG	A	1572	1	14,14,15	0.50	0	17,19,21	0.69	1 (5%)
4	NAG	B	1575	-	15,15,15	0.39	0	21,21,21	0.55	0
4	NAG	A	1571	-	15,15,15	0.40	0	21,21,21	0.56	0
4	NAG	B	1574	1	14,14,15	0.49	0	17,19,21	0.60	0
3	SLB	A	1570	-	18,21,21	0.68	0	21,31,31	0.62	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	SLB	B	1573	-	-	0/14/38/38	0/1/1/1
4	NAG	A	1573	-	1/1/6/7	3/6/26/26	0/1/1/1
4	NAG	A	1572	1	-	2/6/23/26	0/1/1/1
4	NAG	B	1575	-	1/1/6/7	2/6/26/26	0/1/1/1
4	NAG	A	1571	-	-	0/6/26/26	0/1/1/1
4	NAG	B	1574	1	-	1/6/23/26	0/1/1/1
3	SLB	A	1570	-	-	0/14/38/38	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1572	NAG	C2-N2-C7	-2.05	119.98	122.90

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1573	NAG	C1
4	B	1575	NAG	C1

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1573	NAG	C4-C5-C6-O6
4	B	1574	NAG	O5-C5-C6-O6
4	B	1575	NAG	O5-C5-C6-O6
4	B	1575	NAG	C4-C5-C6-O6
4	A	1573	NAG	O5-C5-C6-O6

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1573	SLB	2	0
4	A	1573	NAG	1	0
4	A	1572	NAG	1	0
4	B	1575	NAG	2	0
4	B	1574	NAG	2	0
3	A	1570	SLB	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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.