



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 17, 2020 – 01:10 PM EST

PDB ID : 6WEX  
Title : Crystal Structure of Broadly Neutralizing Antibody 3I14-D93N Mutant Bound to the Influenza A H6 Hemagglutinin  
Authors : Harshbarger, W.D.; Lockbaum, G.J.; Deming, D.T.; Attatippaholkun, N.; Schiffer, C.A.; Marasco, W.A.  
Deposited on : 2020-04-03  
Resolution : 3.49 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 : 2.14.6  
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) : 2.14.6

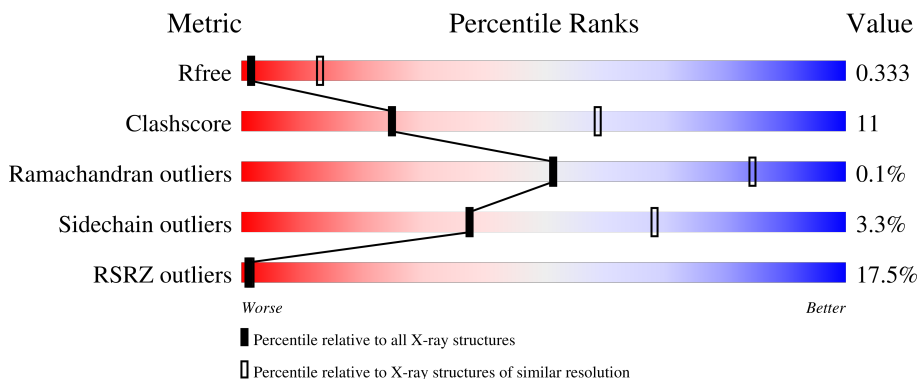
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.49 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 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	E	325	<div> <div>6%</div> <div> <div></div> <div>76%</div> <div>23%</div> <div>.</div> </div> </div>
2	F	220	<div> <div>6%</div> <div> <div></div> <div>68%</div> <div>10%</div> <div>22%</div> <div></div> </div> </div>
3	H	235	<div> <div>33%</div> <div> <div></div> <div>65%</div> <div>29%</div> <div>5%</div> <div>.</div> </div> </div>
4	L	213	<div> <div>24%</div> <div> <div></div> <div>73%</div> <div>24%</div> <div>.</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7123 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	E	325	Total	C	N	O	S	0	0	0
			2565	1627	437	488	13			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	9	PRO	-	expression tag	UNP A0A0M3KL64
E	10	GLY	-	expression tag	UNP A0A0M3KL64
E	?	-	THR	deletion	UNP A0A0M3KL64
E	?	-	ASN	deletion	UNP A0A0M3KL64

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	172	Total	C	N	O	S	0	0	0
			1385	862	243	273	7			

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	182	GLY	-	expression tag	UNP A0A0J9X245
F	183	SER	-	expression tag	UNP A0A0J9X245
F	184	PRO	-	expression tag	UNP A0A0J9X245
F	185	GLY	-	expression tag	UNP A0A0J9X245
F	186	SER	-	expression tag	UNP A0A0J9X245
F	187	GLY	-	expression tag	UNP A0A0J9X245
F	188	TYR	-	expression tag	UNP A0A0J9X245
F	189	ILE	-	expression tag	UNP A0A0J9X245
F	190	PRO	-	expression tag	UNP A0A0J9X245
F	191	GLU	-	expression tag	UNP A0A0J9X245
F	192	ALA	-	expression tag	UNP A0A0J9X245
F	193	PRO	-	expression tag	UNP A0A0J9X245

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	194	ARG	-	expression tag	UNP A0A0J9X245
F	195	ASP	-	expression tag	UNP A0A0J9X245
F	196	GLY	-	expression tag	UNP A0A0J9X245
F	197	GLN	-	expression tag	UNP A0A0J9X245
F	198	ALA	-	expression tag	UNP A0A0J9X245
F	199	TYR	-	expression tag	UNP A0A0J9X245
F	200	VAL	-	expression tag	UNP A0A0J9X245
F	201	ARG	-	expression tag	UNP A0A0J9X245
F	202	LYS	-	expression tag	UNP A0A0J9X245
F	203	ASP	-	expression tag	UNP A0A0J9X245
F	204	GLY	-	expression tag	UNP A0A0J9X245
F	205	GLU	-	expression tag	UNP A0A0J9X245
F	206	TRP	-	expression tag	UNP A0A0J9X245
F	207	VAL	-	expression tag	UNP A0A0J9X245
F	208	LEU	-	expression tag	UNP A0A0J9X245
F	209	LEU	-	expression tag	UNP A0A0J9X245
F	210	SER	-	expression tag	UNP A0A0J9X245
F	211	THR	-	expression tag	UNP A0A0J9X245
F	212	PHE	-	expression tag	UNP A0A0J9X245
F	213	LEU	-	expression tag	UNP A0A0J9X245
F	214	GLY	-	expression tag	UNP A0A0J9X245
F	215	HIS	-	expression tag	UNP A0A0J9X245
F	216	HIS	-	expression tag	UNP A0A0J9X245
F	217	HIS	-	expression tag	UNP A0A0J9X245
F	218	HIS	-	expression tag	UNP A0A0J9X245
F	219	HIS	-	expression tag	UNP A0A0J9X245
F	220	HIS	-	expression tag	UNP A0A0J9X245

- Molecule 3 is a protein called heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	223	Total	C	N	O	S	0	0	0
			1693	1083	285	319	6			

- Molecule 4 is a protein called light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	213	Total	C	N	O	S	0	0	0
			1466	916	247	298	5			

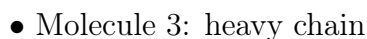
- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



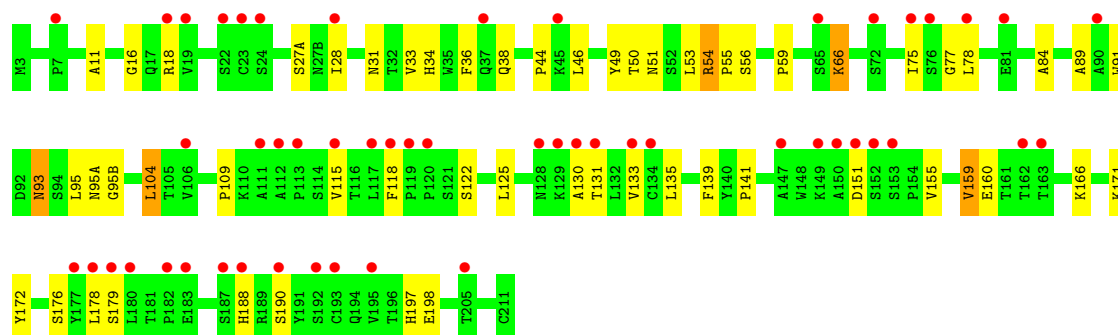
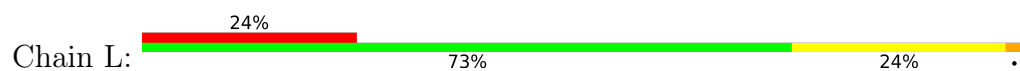
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	E	1	14	8	1	5	0	0



- Molecule 1: Hemagglutinin



• Molecule 4: light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.81Å 117.81Å 438.26Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.26 – 3.49 48.26 – 3.49	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.26-3.49) 95.4 (48.26-3.49)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.17.1 _3660	Depositor
R, $R_{free}$	0.276 , 0.333 0.276 , 0.333	Depositor DCC
$R_{free}$ test set	1530 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.9	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 91.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	7123	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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: NAG

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	E	0.24	0/2627	0.46	0/3573
2	F	0.25	0/1413	0.38	0/1902
3	H	0.26	0/1739	0.49	0/2368
4	L	0.26	0/1496	0.54	0/2036
All	All	0.25	0/7275	0.47	0/9879

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	E	2565	0	2514	49	0
2	F	1385	0	1298	13	0
3	H	1693	0	1667	59	0
4	L	1466	0	1303	44	0
5	E	14	0	13	0	0
All	All	7123	0	6795	150	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 11.

The worst 5 of 150 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:E:156:LYS:HD2	1:E:159:ALA:HA	1.67	0.75
1:E:315:ARG:HD3	1:E:315:ARG:H	1.51	0.75
3:H:35:HIS:HB2	3:H:93:ALA:O	1.87	0.74
3:H:101:HIS:ND1	3:H:101:HIS:O	2.22	0.73
4:L:160:GLU:OE1	4:L:160:GLU:N	2.21	0.72

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	E	323/325 (99%)	310 (96%)	13 (4%)	0	100	100
2	F	170/220 (77%)	166 (98%)	4 (2%)	0	100	100
3	H	219/235 (93%)	192 (88%)	27 (12%)	0	100	100
4	L	211/213 (99%)	184 (87%)	26 (12%)	1 (0%)	29	68
All	All	923/993 (93%)	852 (92%)	70 (8%)	1 (0%)	51	84

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	L	159	VAL

### 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	E	285/285 (100%)	280 (98%)	5 (2%)	59	81
2	F	146/185 (79%)	142 (97%)	4 (3%)	44	73
3	H	188/198 (95%)	180 (96%)	8 (4%)	29	62
4	L	147/177 (83%)	139 (95%)	8 (5%)	22	55
All	All	766/845 (91%)	741 (97%)	25 (3%)	38	68

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	27	PHE
3	H	82(A)	GLU
4	L	159	VAL
3	H	38	ARG
3	H	100	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	226	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 ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

1 ligand is 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	$\# Z  > 2$
5	NAG	E	401	1	14,14,15	0.31	0	17,19,21	0.59	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
5	NAG	E	401	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	401	NAG	O5-C5-C6-O6
5	E	401	NAG	C4-C5-C6-O6
5	E	401	NAG	C3-C2-N2-C7

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	E	325/325 (100%)	0.45	19 (5%)	23 20	37, 81, 120, 172	0
2	F	172/220 (78%)	0.58	14 (8%)	12 12	38, 94, 141, 166	0
3	H	223/235 (94%)	1.70	78 (34%)	0 0	85, 158, 204, 234	0
4	L	213/213 (100%)	1.15	52 (24%)	0 0	70, 127, 189, 218	0
All	All	933/993 (93%)	0.94	163 (17%)	1 1	37, 109, 187, 234	0

The worst 5 of 163 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	112	SER	14.2
4	L	180	LEU	8.2
3	H	119	PRO	7.5
3	H	117	LYS	7.4
3	H	120	SER	6.8

### 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 monosaccharides in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
5	NAG	E	401	14/15	0.83	0.27	89,108,117,118	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.