



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 02:20 pm BST

PDB ID : 3Q8F  
Title : Crystal structure of 2-Fluorohistidine labeled Protective Antigen (pH 5.8)  
Authors : Lovell, S.; Battaile, K.P.; Rajapaksha, M.; Janowiak, B.E.; Andra, K.K.;  
Bann, J.G.  
Deposited on : 2011-01-06  
Resolution : 2.10 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	:	2.11
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)	:	2.11

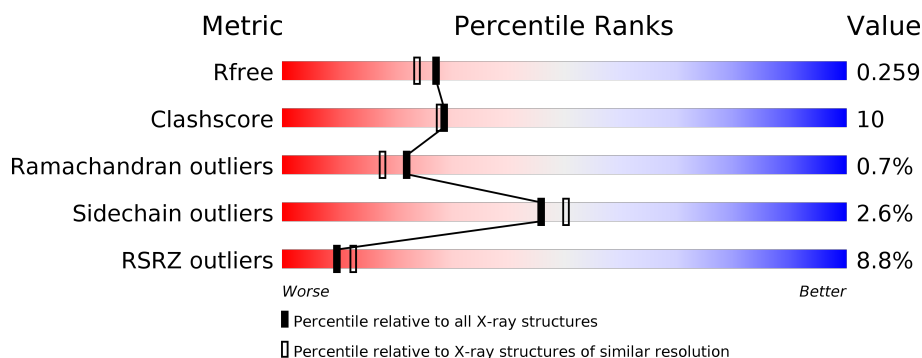
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	A	735	<div> <div>8%</div> <div> <div></div> <div>70%</div> <div>17%</div> <div>•</div> <div>12%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5204 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.

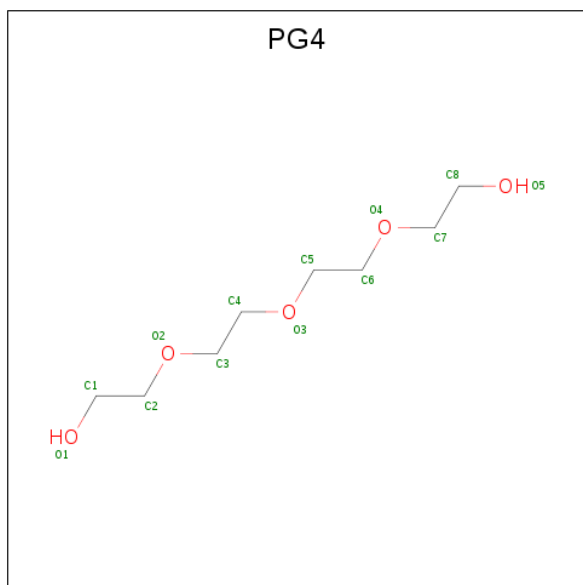
- Molecule 1 is a protein called Protective antigen.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	F	N	O	S			
1	A	645	5052	3181	7	838	1018	8	0	0	0

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

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

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	124	Total 124	O 124	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.29 Å 93.66 Å 115.56 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.40 – 2.10 43.40 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.0 (43.40-2.10) 97.9 (43.40-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.79 (at 2.10 Å)	Xtriage
Refinement program	PHENIX dev_601	Depositor
R, $R_{free}$	0.222 , 0.267 0.215 , 0.259	Depositor DCC
$R_{free}$ test set	2341 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.0	Xtriage
Anisotropy	0.462	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5204	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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: CA, PG4, 2HF

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	0.51	0/5045	0.62	0/6839

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

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	5052	0	4852	97	0
2	A	2	0	0	0	0
3	A	26	0	36	3	0
4	A	124	0	0	5	0
All	All	5204	0	4888	97	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 10.

The worst 5 of 97 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:A:327:SER:HA	4:A:862:HOH:O	1.87	0.75
1:A:376:ASN:HA	1:A:401:LEU:HD13	1.74	0.70
1:A:71:VAL:HG11	1:A:74:SER:N	2.08	0.68
1:A:69:ILE:HG23	1:A:116:ILE:HG13	1.77	0.67
1:A:253:2HF:HND1	1:A:255:LEU:H	1.47	0.62

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	614/735 (84%)	581 (95%)	29 (5%)	4 (1%)	22	18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	141	ASN
1	A	354	THR
1	A	93	ASP
1	A	124	ASN

### 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	547/651 (84%)	533 (97%)	14 (3%)	46	50

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	224	GLU
1	A	338	LEU
1	A	551	ASP
1	A	141	ASN
1	A	539	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	422	ASN
1	A	437	ASN
1	A	601	ASN
1	A	408	ASN
1	A	553	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

7 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	$\# Z  > 2$
1	2HF	A	263	1	8,11,12	1.20	1 (12%)	2,14,16	0.69	0
1	2HF	A	616	1	8,11,12	1.14	1 (12%)	2,14,16	1.90	1 (50%)
1	2HF	A	299	1	8,11,12	1.11	1 (12%)	2,14,16	0.47	0
1	2HF	A	253	1	8,11,12	1.24	1 (12%)	2,14,16	0.60	0
1	2HF	A	336	1	8,11,12	1.25	1 (12%)	2,14,16	2.58	1 (50%)
1	2HF	A	86	1	8,11,12	1.10	1 (12%)	2,14,16	1.30	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	2HF	A	597	1	8,11,12	1.14	1 (12%)	2,14,16	1.70	1 (50%)

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
1	2HF	A	263	1	-	0/5/6/8	0/1/1/1
1	2HF	A	616	1	-	1/5/6/8	0/1/1/1
1	2HF	A	299	1	-	0/5/6/8	0/1/1/1
1	2HF	A	253	1	-	0/5/6/8	0/1/1/1
1	2HF	A	336	1	-	0/5/6/8	0/1/1/1
1	2HF	A	86	1	-	1/5/6/8	0/1/1/1
1	2HF	A	597	1	-	1/5/6/8	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	336	2HF	F1-CE1	2.78	1.40	1.33
1	A	597	2HF	F1-CE1	2.73	1.40	1.33
1	A	616	2HF	F1-CE1	2.55	1.39	1.33
1	A	263	2HF	F1-CE1	2.49	1.39	1.33
1	A	86	2HF	F1-CE1	2.46	1.39	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	336	2HF	CB-CA-C	-3.55	104.81	111.47
1	A	616	2HF	CB-CA-C	-2.68	106.44	111.47
1	A	597	2HF	CB-CA-C	-2.33	107.09	111.47

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	616	2HF	O-C-CA-CB
1	A	597	2HF	O-C-CA-CB
1	A	86	2HF	CA-CB-CG-ND1

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	299	2HF	1	0
1	A	253	2HF	1	0
1	A	336	2HF	1	0
1	A	86	2HF	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 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	$\# Z  > 2$
3	PG4	A	739	-	12,12,12	0.45	0	11,11,11	0.92	0
3	PG4	A	738	-	12,12,12	0.68	0	11,11,11	0.67	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	PG4	A	739	-	-	5/10/10/10	-
3	PG4	A	738	-	-	2/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	739	PG4	O3-C5-C6-O4
3	A	738	PG4	O4-C7-C8-O5
3	A	739	PG4	O2-C3-C4-O3
3	A	739	PG4	C5-C6-O4-C7
3	A	738	PG4	C5-C6-O4-C7

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	739	PG4	3	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	638/735 (86%)	0.62	56 (8%) 10 12	30, 50, 80, 91	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	402	SER	5.3
1	A	137	THR	5.2
1	A	88	THR	4.4
1	A	74	SER	4.3
1	A	96	VAL	4.3

### 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
1	2HF	A	86	11/12	0.76	0.25	76,81,85,86	0
1	2HF	A	299	11/12	0.95	0.13	36,44,52,57	0
1	2HF	A	336	11/12	0.96	0.12	38,41,46,60	0
1	2HF	A	616	11/12	0.96	0.13	35,37,43,44	0
1	2HF	A	253	11/12	0.97	0.16	31,37,41,41	0
1	2HF	A	597	11/12	0.97	0.10	33,40,48,50	0
1	2HF	A	263	11/12	0.98	0.15	28,31,34,35	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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
3	PG4	A	739	13/13	0.78	0.25	50,59,65,75	0
3	PG4	A	738	13/13	0.93	0.14	43,45,53,53	0
2	CA	A	736	1/1	0.99	0.12	33,33,33,33	0
2	CA	A	737	1/1	0.99	0.12	38,38,38,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.