



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 02:15 am GMT

PDB ID : 1FCQ  
Title : CRYSTAL STRUCTURE (MONOCLINIC) OF BEE VENOM  
HYALURONIDASE  
Authors : Markovic-Housley, Z.; Miglierini, G.; Soldatova, L.; Mueller, U.; Schirmer, T.  
Deposited on : 2000-07-19  
Resolution : 1.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<http://wwpdb.org/validation/2016/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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

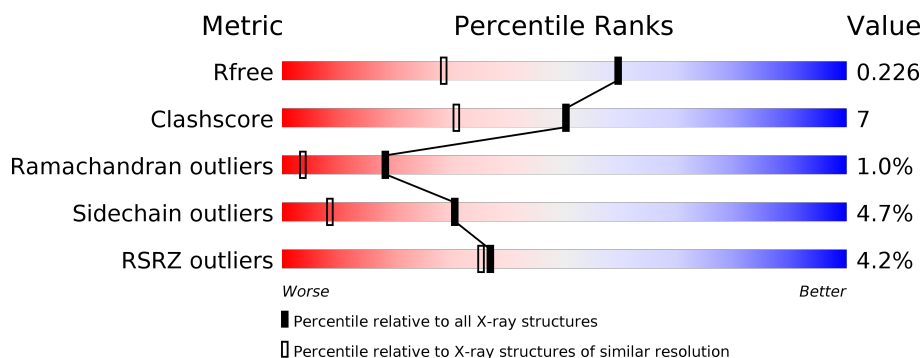
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	100719	2696 (1.60-1.60)
Clashscore	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)
RSRZ outliers	101464	2714 (1.60-1.60)

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	350	<div> <div>4%</div> <div> <div></div> <div>74%</div> <div>14%</div> <div>••</div> <div>10%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2963 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 HYALURONOGLUCOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	314	Total	C	N	O	S	68	0	0
			2609	1672	457	468	12			

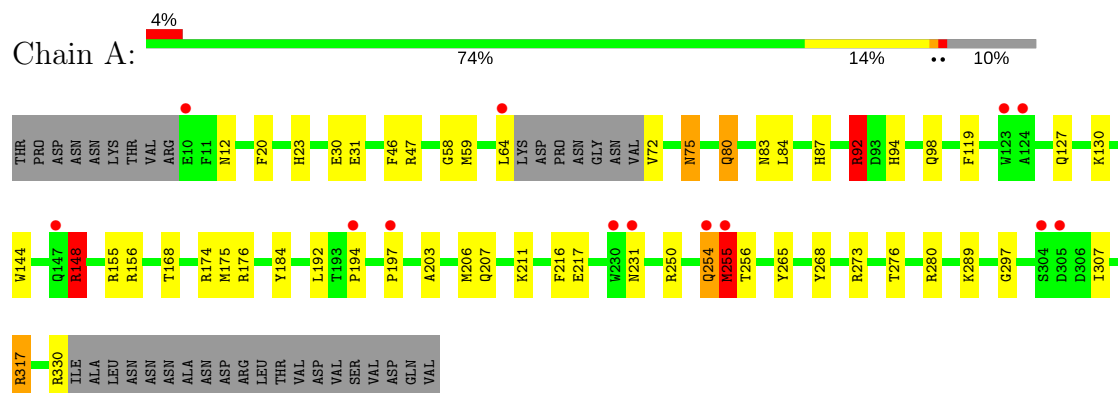
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	354	Total	O	0	0
			354	354		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

#### • Molecule 1: HYALURONOGLUCOSAMINIDASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.89Å 90.23Å 49.06Å 90.00° 92.36° 90.00°	Depositor
Resolution (Å)	8.00 – 1.60 25.64 – 1.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-1.60) 96.1 (25.64-1.60)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.86 (at 1.60Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.185 , 0.227 0.190 , 0.226	Depositor DCC
$R_{free}$ test set	4415 reflections (11.18%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.6	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 62.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.037 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2963	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.32% 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

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	0.47	2/2679 (0.1%)	1.11	15/3623 (0.4%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	330	ARG	CA-C	-6.76	1.35	1.52
1	A	330	ARG	CA-CB	5.57	1.66	1.53

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	176	ARG	NE-CZ-NH2	10.21	125.41	120.30
1	A	92	ARG	NE-CZ-NH2	-10.10	115.25	120.30
1	A	330	ARG	CA-C-O	-9.73	99.67	120.10
1	A	317	ARG	NE-CZ-NH1	8.85	124.73	120.30
1	A	194	PRO	CA-N-CD	-7.83	100.54	111.50
1	A	280	ARG	NE-CZ-NH2	7.66	124.13	120.30
1	A	155	ARG	NE-CZ-NH2	-7.45	116.57	120.30
1	A	47	ARG	NE-CZ-NH1	-7.29	116.65	120.30
1	A	148	ARG	CD-NE-CZ	7.09	133.53	123.60
1	A	156	ARG	NE-CZ-NH1	7.09	123.84	120.30
1	A	155	ARG	NE-CZ-NH1	6.33	123.47	120.30
1	A	250	ARG	NE-CZ-NH1	-6.26	117.17	120.30
1	A	184	TYR	CB-CG-CD2	-5.30	117.82	121.00
1	A	268	TYR	CB-CG-CD1	-5.08	117.95	121.00
1	A	268	TYR	CB-CG-CD2	5.04	124.02	121.00

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	2609	0	2550	33	0
2	A	354	0	0	3	0
All	All	2963	0	2550	33	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 7.

All (33) 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:80:GLN:HE21	1:A:80:GLN:H	1.11	0.97
1:A:59:MET:H	1:A:87:HIS:HD2	1.19	0.90
1:A:58:GLY:H	1:A:94:HIS:HD2	1.29	0.77
1:A:216:PHE:C	1:A:255:MET:SD	2.73	0.67
1:A:148:ARG:HG3	1:A:148:ARG:HH11	1.59	0.67
1:A:207:GLN:HG2	2:A:541:HOH:O	1.94	0.66
1:A:80:GLN:NE2	1:A:80:GLN:H	1.92	0.64
1:A:75:ASN:HD21	1:A:83:ASN:H	1.46	0.63
1:A:59:MET:H	1:A:87:HIS:CD2	2.09	0.63
1:A:174:ARG:HG3	2:A:415:HOH:O	1.98	0.61
1:A:98:GLN:HA	1:A:98:GLN:HE21	1.66	0.61
1:A:87:HIS:HE1	1:A:168:THR:OG1	1.86	0.58
1:A:80:GLN:HE21	1:A:80:GLN:N	1.93	0.56
1:A:75:ASN:ND2	1:A:83:ASN:H	2.03	0.56
1:A:203:ALA:HA	1:A:206:MET:HE2	1.91	0.53
1:A:58:GLY:H	1:A:94:HIS:CD2	2.20	0.53
1:A:59:MET:N	1:A:87:HIS:HD2	1.99	0.52
1:A:31:GLU:OE2	1:A:317:ARG:NH2	2.44	0.51
1:A:64:LEU:O	1:A:72:VAL:N	2.45	0.50
1:A:20:PHE:CD2	1:A:46:PHE:HB2	2.47	0.49
1:A:98:GLN:HA	1:A:98:GLN:NE2	2.28	0.49
1:A:119:PHE:O	1:A:130:LYS:HE2	2.13	0.49
1:A:217:GLU:HG2	1:A:255:MET:SD	2.52	0.49
1:A:12:ASN:O	1:A:297:GLY:HA3	2.14	0.48
1:A:30:GLU:HG2	2:A:533:HOH:O	2.16	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:20:PHE:HA	1:A:23:HIS:CE1	2.51	0.46
1:A:192:LEU:HD23	1:A:197:PRO:O	2.17	0.44
1:A:92:ARG:HG3	1:A:175:MET:CE	2.50	0.42
1:A:216:PHE:CB	1:A:255:MET:SD	3.08	0.42
1:A:273:ARG:NH2	1:A:276:THR:HG21	2.35	0.42
1:A:92:ARG:HB3	1:A:92:ARG:HE	1.59	0.41
1:A:144:TRP:CE3	1:A:148:ARG:HG2	2.55	0.41
1:A:216:PHE:HB3	1:A:255:MET:SD	2.60	0.41

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	310/350 (89%)	298 (96%)	9 (3%)	3 (1%)	18 4

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	254	GLN
1	A	255	MET
1	A	307	ILE

### 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	276/309 (89%)	263 (95%)	13 (5%)	30 8

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	80	GLN
1	A	84	LEU
1	A	92	ARG
1	A	127	GLN
1	A	148	ARG
1	A	211	LYS
1	A	231	ASN
1	A	254	GLN
1	A	255	MET
1	A	256	THR
1	A	265	TYR
1	A	289	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	80	GLN
1	A	87	HIS
1	A	94	HIS
1	A	97	ASN
1	A	98	GLN
1	A	121	GLN
1	A	162	GLN
1	A	180	ASN
1	A	231	ASN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	309/350 (88%)	0.31	13 (4%) 37 35	11, 19, 35, 58	5 (1%)

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	255	MET	21.5
1	A	194	PRO	5.8
1	A	254	GLN	5.3
1	A	305	ASP	4.4
1	A	230	TRP	4.0
1	A	10	GLU	3.5
1	A	124	ALA	3.3
1	A	64	LEU	3.1
1	A	304	SER	2.9
1	A	123	TRP	2.7
1	A	231	ASN	2.5
1	A	147	GLN	2.2
1	A	197	PRO	2.2

### 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 [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.