



# Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 04:59 pm BST

PDB ID : 2YX2  
Title : Crystal structure of cloned trimeric hyaluronidase from streptococcus pyogenes at 2.8 Å resolution  
Authors : Mishra, P.; Prem Kumar, R.; Bhakuni, V.; Singh, N.; Sharma, S.; Kaur, P.; Perbandt, M.; Betzel, C.; Singh, T.P.  
Deposited on : 2007-04-23  
Resolution : 2.80 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

<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  
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) : 2.11

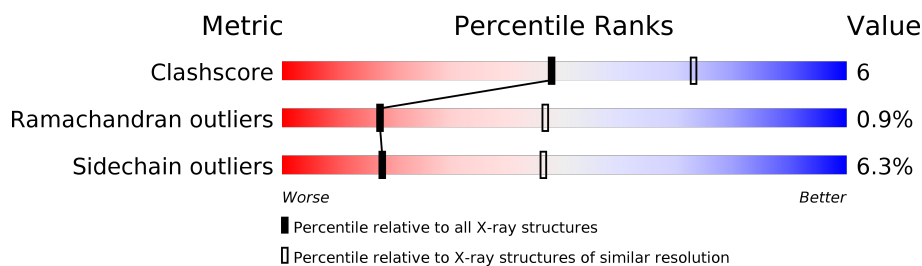
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	338	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2658 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 Hyaluronidase, phage associated.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	332	Total	C	N	O	S	0	4	0
			2552	1589	447	511	5			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	338	HIS	-	EXPRESSION TAG	UNP Q9A0M7

- Molecule 2 is water.

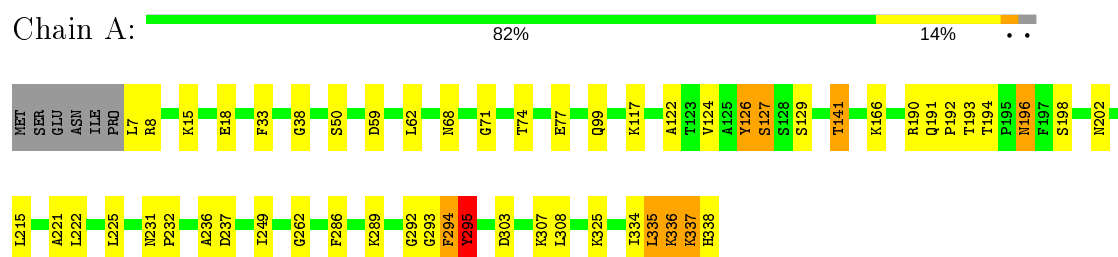
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	106	Total	O	0	0
			106	106		

### 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. 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: Hyaluronidase, phage associated



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.77Å 58.77Å 586.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	99.6 (20.00-2.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.192 , 0.224	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

## 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	0.83	7/2586 (0.3%)	0.93	11/3479 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	293	GLY	N-CA	9.77	1.60	1.46
1	A	293	GLY	C-O	8.70	1.37	1.23
1	A	294	PHE	CG-CD1	7.29	1.49	1.38
1	A	295	TYR	C-O	6.17	1.35	1.23
1	A	294	PHE	CE1-CZ	5.95	1.48	1.37
1	A	294	PHE	CD1-CE1	5.59	1.50	1.39
1	A	295	TYR	CA-CB	-5.28	1.42	1.53

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	295	TYR	CB-CG-CD1	-10.55	114.67	121.00
1	A	127	SER	CB-CA-C	10.20	129.48	110.10
1	A	294	PHE	CB-CG-CD2	-9.77	113.97	120.80
1	A	294	PHE	CB-CG-CD1	8.30	126.61	120.80
1	A	295	TYR	CB-CG-CD2	8.08	125.84	121.00
1	A	127	SER	N-CA-C	-6.55	93.30	111.00
1	A	295	TYR	CG-CD2-CE2	-6.42	116.16	121.30
1	A	122	ALA	N-CA-C	5.54	125.97	111.00
1	A	294	PHE	O-C-N	5.45	131.41	122.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	303	ASP	CB-CG-OD1	-5.37	113.47	118.30
1	A	292	GLY	C-N-CA	-5.04	111.72	122.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	295	TYR	Mainchain

## 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	2552	0	2575	31	0
2	A	106	0	0	7	0
All	All	2658	0	2575	31	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 6.

All (31) 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:15:LYS:HB2	1:A:18:GLU:HG3	1.61	0.81
1:A:74:THR:OG1	1:A:77:GLU:HB2	1.84	0.78
1:A:126:TYR:HD2	1:A:127:SER:N	1.86	0.74
1:A:289:LYS:NZ	2:A:428:HOH:O	2.29	0.66
1:A:126:TYR:HD2	1:A:127:SER:H	1.43	0.65
1:A:289:LYS:HE2	2:A:428:HOH:O	1.95	0.65
1:A:126:TYR:CD2	1:A:127:SER:N	2.68	0.62
1:A:249:ILE:HD11	1:A:262:GLY:HA2	1.82	0.62
1:A:196:ASN:C	1:A:196:ASN:OD1	2.41	0.60
1:A:289:LYS:CE	2:A:428:HOH:O	2.51	0.59
1:A:59:ASP:HB3	1:A:62:LEU:HG	1.85	0.58
1:A:196:ASN:ND2	1:A:198:SER:HB2	2.20	0.57

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:286:PHE:HA	1:A:295:TYR:O	2.05	0.56
1:A:231:ASN:OD1	1:A:232:PRO:HD2	2.07	0.55
1:A:335[B]:LEU:O	1:A:336[B]:LYS:HB2	2.07	0.53
1:A:337[A]:LYS:C	1:A:338[A]:HIS:CD2	2.81	0.53
1:A:68:ASN:ND2	1:A:71:GLY:HA3	2.27	0.49
1:A:190:ARG:O	1:A:192:PRO:HD3	2.12	0.49
1:A:129:SER:HB3	2:A:433:HOH:O	2.12	0.49
1:A:294:PHE:C	1:A:294:PHE:CD2	2.87	0.47
1:A:141:THR:HG22	2:A:393:HOH:O	2.13	0.47
1:A:15:LYS:HB2	1:A:18:GLU:CG	2.39	0.47
1:A:202:ASN:ND2	2:A:348:HOH:O	2.49	0.45
1:A:337[A]:LYS:HD3	1:A:337[A]:LYS:HA	1.77	0.44
1:A:126:TYR:CD2	1:A:127:SER:O	2.70	0.44
1:A:221:ALA:O	1:A:222:LEU:HD23	2.19	0.42
1:A:117:LYS:HB3	2:A:392:HOH:O	2.21	0.41
1:A:222:LEU:HD23	1:A:222:LEU:HA	1.83	0.41
1:A:33:PHE:CZ	1:A:38:GLY:HA2	2.56	0.41
1:A:215:LEU:HD21	1:A:225:LEU:HB2	2.03	0.40
1:A:307:LYS:HB3	1:A:307:LYS:HE2	1.86	0.40

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	333/338 (98%)	305 (92%)	24 (7%)	4 (1%)	13 39

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	ALA
1	A	124	VAL

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	337[A]	LYS
1	A	337[B]	LYS

### 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	274/276 (99%)	255 (93%)	19 (7%)	15	41

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	LEU
1	A	8	ARG
1	A	50	SER
1	A	99	GLN
1	A	126	TYR
1	A	141	THR
1	A	166	LYS
1	A	191	GLN
1	A	193	THR
1	A	194	THR
1	A	196	ASN
1	A	237	ASP
1	A	308	LEU
1	A	325	LYS
1	A	334	ILE
1	A	335[A]	LEU
1	A	335[B]	LEU
1	A	336[A]	LYS
1	A	336[B]	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	ASN
1	A	202	ASN
1	A	214	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 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.