



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:46 am BST

PDB ID : 2BRW  
Title : Crystal structure of Streptococcus Pneumoniae Hyaluronate Lyase from 30percent PEGMME.  
Authors : Rigden, D.J.; Littlejohn, J.E.; Jedrzejask, M.J.  
Deposited on : 2005-05-11  
Resolution : 2.80 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

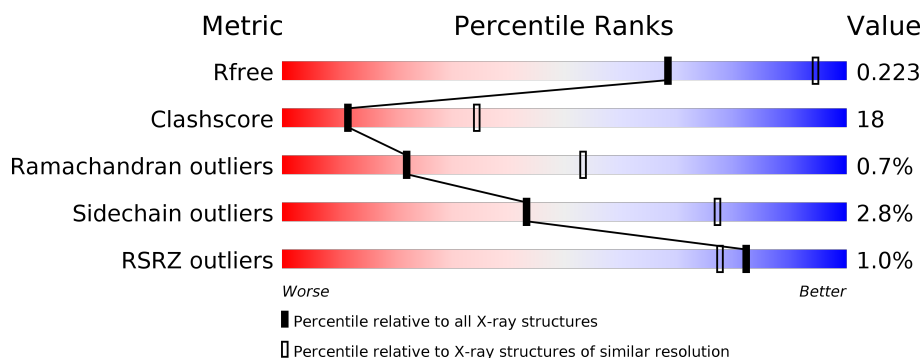
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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  $\leq 5\%$ . 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	731	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>%</span> <span>64%</span> <span>33%</span> <span>..</span> </div> </div>
1	B	731	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>%</span> <span>65%</span> <span>32%</span> <span>..</span> </div> </div>

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
2	SO4	A	1891	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11704 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 HYALURONATE LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	722	Total	C	N	O	S	0	0	1
			5791	3643	969	1157	22			
1	B	722	Total	C	N	O	S	0	0	1
			5791	3643	969	1157	22			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	173	THR	ALA	conflict	UNP Q54873
A	196	ASP	GLU	conflict	UNP Q54873
A	223	ILE	THR	conflict	UNP Q54873
A	496	ARG	CYS	conflict	UNP Q54873
A	541	THR	PRO	conflict	UNP Q54873
A	704	SER	GLY	conflict	UNP Q54873
A	736	SER	PHE	conflict	UNP Q54873
A	790	GLY	ARG	conflict	UNP Q54873
B	173	THR	ALA	conflict	UNP Q54873
B	196	ASP	GLU	conflict	UNP Q54873
B	223	ILE	THR	conflict	UNP Q54873
B	496	ARG	CYS	conflict	UNP Q54873
B	541	THR	PRO	conflict	UNP Q54873
B	704	SER	GLY	conflict	UNP Q54873
B	736	SER	PHE	conflict	UNP Q54873
B	790	GLY	ARG	conflict	UNP Q54873

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	65	Total	O	0	0
			65	65		
3	B	47	Total	O	0	0
			47	47		



E819	N820	N821	E822	Q825	V835	G836	K839	Y840	D841	S845	T846	I847	Q850	F851	Q852	V853	L854	K855	K856	E867	N874	P875	E876	T877	Q878	E879	S880	A881	L890	E891	GLN	HIS	HIS	HIS	HIS	HIS														
A693	T694	T695	I696	D697	Q698	R699	K700	L701	E715	A716	T719	E720	Q721	E722	K723	D724	T728	K749	K756	A757	L758	Q759	K760	I767	E768	Q769	Q770	S771	D772	K773	E774	V775	E776	N777	E778	F779	I782	A785	S792	F806	N807	Q808	N809	I810	K811	E812	I818			
R585	G586	W587	S590	D591	F594	H603	Y604	S605	Y608	V612	N613	P614	Y615	K616	T620	T623	D631	K634	V635	L636	G642	D647	N650	N659	W660	N661	Q662	T663	L664	T665	A666	H667	F671	K674	A678	G681	S682	N683	N686	D690										
H472	V476	R480	H483	R484	M488	S489	E490	G491	E492	T493	K494	Q498	S499	I500	V501	K502	T503	I504	V505	D508	D512	N516	S524	L530	L544	F547	N548	K549	M550	D551	K552	G565	L566	S567	L568	F569	S570	N575	H578	K581	E582	N583	K584							
G394	S395	Y396	I397	D398	H399	T400	Y404	A407	N410	V411	L412	L416	S417	L420	P421	V422	I423	Q424	K425	T426	K427	N428	P429	I430	D431	K432	D433	K434	M435	Q436	T437	M438	Y439	H440	W441	K444	S445	F446	A447	P448	L449	L450	G453	E454	D457	M458	S459	R460	G461	R462

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.67Å 84.06Å 98.67Å 90.00° 98.87° 90.00°	Depositor
Resolution (Å)	95.35 – 2.80 40.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	88.5 (95.35-2.80) 87.3 (40.00-2.40)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, $R_{free}$	0.229 , 0.277 0.234 , 0.223	Depositor DCC
$R_{free}$ test set	2429 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.5	Xtriage
Anisotropy	0.570	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 44.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	11704	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.90% 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: SO4

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	0.43	0/5910	0.66	0/7983
1	B	0.43	0/5910	0.66	0/7983
All	All	0.43	0/11820	0.66	0/15966

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	A	5791	0	5611	209	0
1	B	5791	0	5611	193	0
2	A	5	0	0	2	0
2	B	5	0	0	1	0
3	A	65	0	0	7	0
3	B	47	0	0	6	0
All	All	11704	0	11222	400	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 18.

The worst 5 of 400 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:424:GLN:HE22	1:B:430:ILE:HG22	1.05	1.21
1:A:196:ASP:O	1:A:200:LYS:HB2	1.63	0.97
1:B:424:GLN:HE22	1:B:430:ILE:CG2	1.80	0.95
1:B:424:GLN:NE2	1:B:430:ILE:HG22	1.84	0.93
1:B:719:THR:HG22	1:B:721:GLN:H	1.31	0.93

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	720/731 (98%)	663 (92%)	51 (7%)	6 (1%)	19	49
1	B	720/731 (98%)	662 (92%)	54 (8%)	4 (1%)	25	56
All	All	1440/1462 (98%)	1325 (92%)	105 (7%)	10 (1%)	22	53

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	431	ASP
1	A	193	SER
1	A	721	GLN
1	A	674	LYS
1	B	721	GLN

### 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	640/649 (99%)	620 (97%)	20 (3%)	40	74
1	B	640/649 (99%)	624 (98%)	16 (2%)	47	80
All	All	1280/1298 (99%)	1244 (97%)	36 (3%)	43	77

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	822	GLU
1	B	173	THR
1	B	822	GLU
1	A	854	LEU
1	B	212	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	832	GLN
1	B	254	ASN
1	B	832	GLN
1	A	850	GLN
1	B	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands 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	$\# Z  > 2$
2	SO4	B	1891	-	4,4,4	0.18	0	6,6,6	0.09	0
2	SO4	A	1891	-	4,4,4	0.36	0	6,6,6	0.07	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1891	SO4	1	0
2	A	1891	SO4	2	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	722/731 (98%)	-0.32	8 (1%) 80 75	12, 30, 55, 81	0
1	B	722/731 (98%)	-0.27	7 (0%) 82 77	14, 31, 54, 79	0
All	All	1444/1462 (98%)	-0.29	15 (1%) 82 77	12, 31, 55, 81	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	218	SER	4.4
1	A	867	GLU	4.3
1	A	227	GLU	3.7
1	B	218	SER	3.6
1	B	719	THR	3.0

### 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](#)

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
2	SO4	B	1891	5/5	0.91	0.21	65,66,66,66	0
2	SO4	A	1891	5/5	0.91	0.16	67,67,68,68	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.