



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:48 AM GMT

PDB ID : 2BRW
Title : CRYSTAL STRUCTURE OF STREPTOCOCCUS PNEUMONIAE
HYALURONATE LYASE FROM 30PERCENT PEGMME.
Authors : Rigden, D.J.; Littlejohn, J.E.; Jedrzejewski, 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
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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

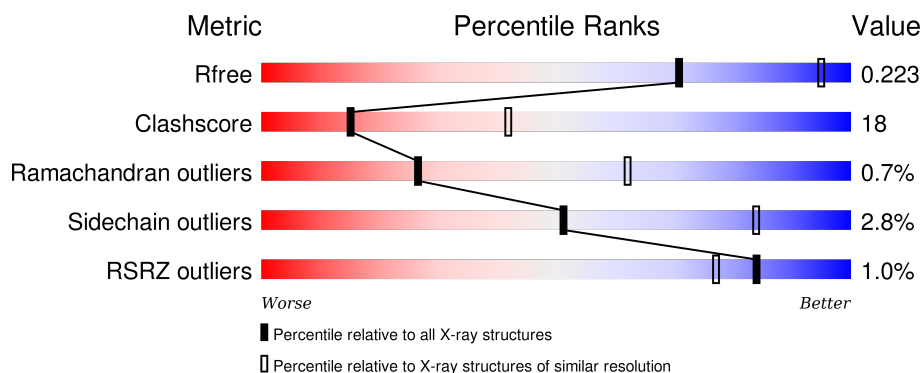
1 Overall quality at a glance ⓘ

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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (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 ≥ 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	731	<div> <div>%</div> <div> <div></div> <div>64%</div> <div>33%</div> <div>..</div> </div> </div>
1	B	731	<div> <div>%</div> <div> <div></div> <div>65%</div> <div>32%</div> <div>..</div> </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 SEE REMARK 9	UNP Q54873
A	196	ASP	GLU	CONFLICT SEE REMARK 9	UNP Q54873
A	223	ILE	THR	CONFLICT SEE REMARK 9	UNP Q54873
A	496	ARG	CYS	CONFLICT SEE REMARK 9	UNP Q54873
A	541	THR	PRO	CONFLICT SEE REMARK 9	UNP Q54873
A	704	SER	GLY	CONFLICT SEE REMARK 9	UNP Q54873
A	736	SER	PHE	CONFLICT SEE REMARK 9	UNP Q54873
A	790	GLY	ARG	CONFLICT SEE REMARK 9	UNP Q54873
B	173	THR	ALA	CONFLICT SEE REMARK 9	UNP Q54873
B	196	ASP	GLU	CONFLICT SEE REMARK 9	UNP Q54873
B	223	ILE	THR	CONFLICT SEE REMARK 9	UNP Q54873
B	496	ARG	CYS	CONFLICT SEE REMARK 9	UNP Q54873
B	541	THR	PRO	CONFLICT SEE REMARK 9	UNP Q54873
B	704	SER	GLY	CONFLICT SEE REMARK 9	UNP Q54873
B	736	SER	PHE	CONFLICT SEE REMARK 9	UNP Q54873
B	790	GLY	ARG	CONFLICT SEE REMARK 9	UNP Q54873

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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		

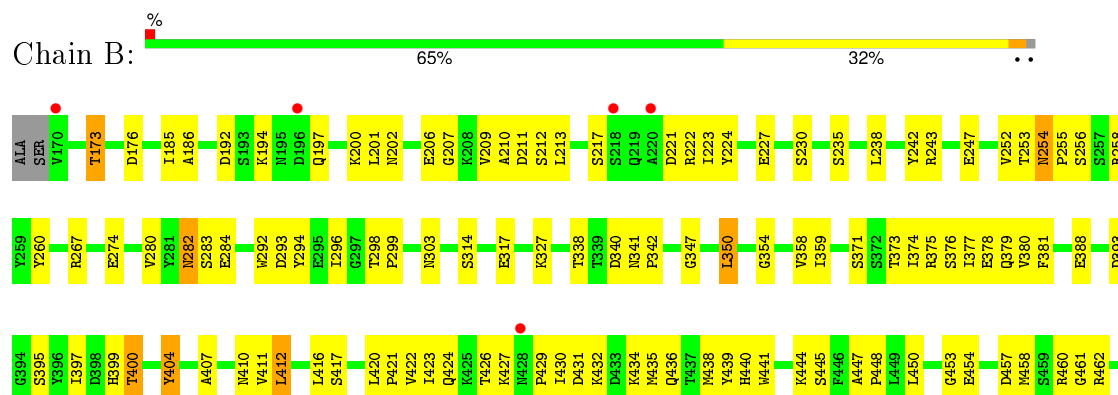
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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: HYALURONATE LYASE



• Molecule 1: HYALURONATE LYASE



H472	R585	A693	E819
	G586	T694	I820
V476	W587	T695	I821
		I696	E822
R480	S590	D697	Q825
R483	D691	Q698	W835
R484	F594	R699	G836
		K700	
M488	H603	L701	R839
S489	Y604	E715	Y840
E490	S605	A716	D841
G491			
E492	Y608	T719	S845
T493		E720	T846
K494	V612	Q721	I847
	N613	E722	
Q498	P614	K723	Q850
S499	Y615	D724	F851
L500	K616		Q852
V501	T620	T728	W853
T503			L854
L504	T623	K749	R855
V505			R856
	D631	K756	
D508		A757	E867
	K634	L758	
D512	V635	Q759	I874
	L636	K760	P875
N516			R876
S524	G642	N767	T877
	D647	E768	Q878
L530		G769	E879
	N650	Q770	S880
L544		S771	A881
	N659	D772	
F547	M660	K773	L890
N548	N661	E774	E891
K549	Q662	V775	GLN
M550	T663	E776	HIS
D551	L664	N777	HIS
K552	T665	E778	HIS
	A666	F779	HIS
G565	H667	I782	HIS
L566			
S567	F671	A785	
L568			
F569	K674	S792	
S570			
N575	A678	F806	
		N807	
H578	G681	Q808	
	S682	W809	
K581	N683	I810	
E582		K811	
N583	N686	E812	
K584	D690	I818	

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	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
$\langle I/\sigma(I) \rangle$ ¹	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	1582 reflections (5.43%)	DCC
Wilson B-factor (Å ²)	31.5	Xtriage
Anisotropy	0.570	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 44.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	1 of 47765 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	11704	wwPDB-VP
Average B, all atoms (Å ²)	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.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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%)	24	58
1	B	720/731 (98%)	662 (92%)	54 (8%)	4 (1%)	30	65
All	All	1440/1462 (98%)	1325 (92%)	105 (7%)	10 (1%)	26	62

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%)	47	81
1	B	640/649 (99%)	624 (98%)	16 (2%)	55	86
All	All	1280/1298 (99%)	1244 (97%)	36 (3%)	51	84

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

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	A	1891	-	4,4,4	0.38	0	6,6,6	0.08	0
2	SO4	B	1891	-	4,4,4	0.16	0	6,6,6	0.10	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
2	SO4	A	1891	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1891	-	-	0/0/0/0	0/0/0/0

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	A	1891	SO4	2	0
2	B	1891	SO4	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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 ²)	Q<0.9
1	A	722/731 (98%)	-0.31	8 (1%) 82 74	12, 30, 55, 81	0
1	B	722/731 (98%)	-0.26	7 (0%) 84 77	14, 31, 54, 79	0
All	All	1444/1462 (98%)	-0.28	15 (1%) 84 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	4.5
1	A	867	GLU	4.4
1	A	227	GLU	3.8
1	B	218	SER	3.7
1	B	719	THR	3.1

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(\AA^2)	Q<0.9
2	SO4	A	1891	5/5	0.91	0.16	-0.50	67,67,68,68	0
2	SO4	B	1891	5/5	0.91	0.21	-	65,66,66,66	0

6.5 Other polymers [i](#)

There are no such residues in this entry.