Full wwPDB X-ray Structure Validation Report (i)

Oct 17, 2021 – 12:24 AM EDT

PDB ID	:	1N7N
Title	:	Streptococcus pneumoniae Hyaluronate Lyase W292A Mutant
Authors	:	Nukui, M.; Taylor, K.B.; McPherson, D.T.; Shigenaga, M.; Jedrzejas, M.J.
Deposited on	:	2002-11-16
Resolution	:	1.55 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

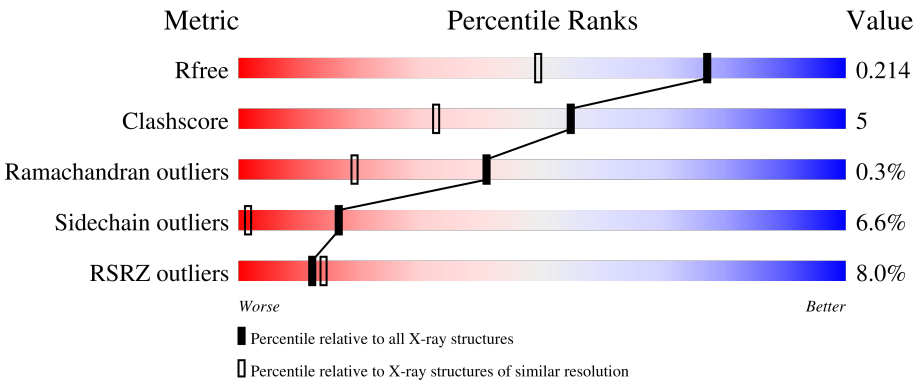
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.55 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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	721	8% 86% 11% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6507 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	721	5775	3630	967	1156	22	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	292	ALA	TRP	engineered mutation	UNP Q54873
A	731	VAL	GLY	SEE REMARK 999	UNP Q54873

- Molecule 2 is water.

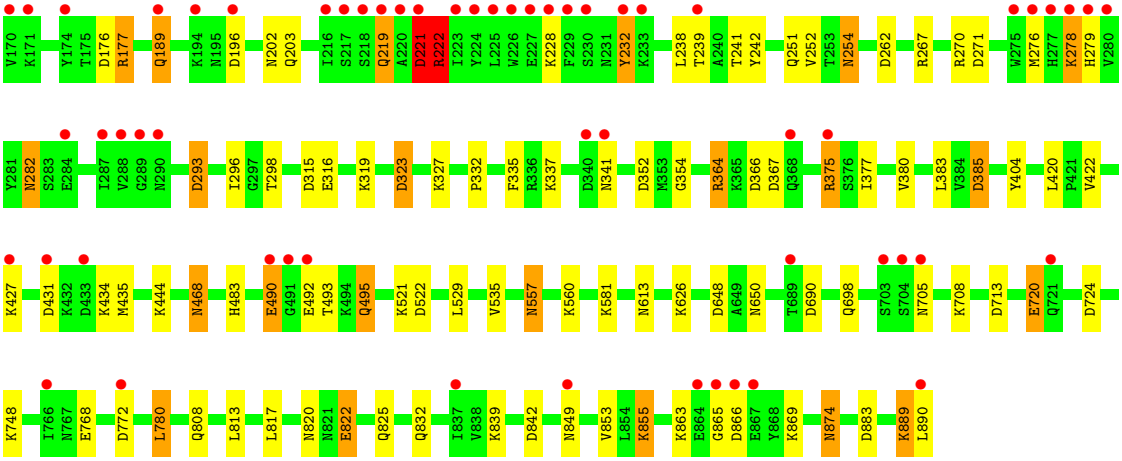
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	732	732	732	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: HYALURONIDASE

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	84.07Å 103.58Å 101.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.55 19.99 – 1.55	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-1.55) 96.9 (19.99-1.55)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 1.55Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.180 , 0.209 0.191 , 0.214	Depositor DCC
R_{free} test set	1266 reflections (1.01%)	wwPDB-VP
Wilson B-factor (Å^2)	18.6	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 58.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.014 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6507	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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)

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	0.58	0/5892	0.82	22/7957 (0.3%)

There are no bond length outliers.

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	271	ASP	CB-CG-OD2	9.33	126.69	118.30
1	A	323	ASP	CB-CG-OD2	8.80	126.22	118.30
1	A	385	ASP	CB-CG-OD2	7.32	124.88	118.30
1	A	842	ASP	CB-CG-OD2	6.86	124.48	118.30
1	A	176	ASP	CB-CG-OD2	6.74	124.36	118.30
1	A	648	ASP	CB-CG-OD1	6.48	124.13	118.30
1	A	772	ASP	CB-CG-OD2	6.36	124.02	118.30
1	A	315	ASP	CB-CG-OD2	6.31	123.98	118.30
1	A	724	ASP	CB-CG-OD2	6.22	123.90	118.30
1	A	262	ASP	CB-CG-OD2	5.86	123.57	118.30
1	A	883	ASP	CB-CG-OD2	5.86	123.57	118.30
1	A	522	ASP	CB-CG-OD2	5.81	123.53	118.30
1	A	780	LEU	CA-CB-CG	5.79	128.61	115.30
1	A	713	ASP	CB-CG-OD2	5.65	123.39	118.30
1	A	293	ASP	CB-CG-OD2	5.63	123.37	118.30
1	A	431	ASP	CB-CG-OD2	5.54	123.29	118.30
1	A	366	ASP	CB-CG-OD2	5.42	123.17	118.30
1	A	352	ASP	CB-CG-OD2	5.28	123.06	118.30
1	A	866	ASP	CB-CG-OD2	5.25	123.03	118.30
1	A	690	ASP	CB-CG-OD2	5.21	122.99	118.30
1	A	367	ASP	CB-CG-OD2	5.06	122.85	118.30
1	A	196	ASP	CB-CG-OD2	5.05	122.85	118.30

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	5775	0	5590	59	0
2	A	732	0	0	28	0
All	All	6507	0	5590	59	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 5.

All (59) 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:613:ASN:H	1:A:698:GLN:HE22	1.08	0.96
1:A:650:ASN:HD21	1:A:832:GLN:HE22	1.35	0.74
1:A:420:LEU:HD22	1:A:435:MET:CE	2.17	0.74
1:A:535:VAL:HG22	2:A:1612:HOH:O	1.89	0.72
1:A:483:HIS:HE1	2:A:1215:HOH:O	1.74	0.70
1:A:296:ILE:HD11	1:A:335:PHE:CE1	2.29	0.69
1:A:354:GLY:HA3	1:A:377:ILE:HD11	1.76	0.67
1:A:705:ASN:HB2	2:A:1335:HOH:O	1.94	0.66
1:A:863:LYS:NZ	1:A:865:GLY:O	2.25	0.66
1:A:557:ASN:HD22	1:A:557:ASN:C	1.99	0.65
1:A:375:ARG:CD	2:A:1579:HOH:O	2.46	0.63
1:A:375:ARG:HD2	2:A:1579:HOH:O	1.98	0.62
1:A:364:ARG:HD3	2:A:1141:HOH:O	1.99	0.61
1:A:874:ASN:C	1:A:874:ASN:HD22	2.04	0.60
1:A:296:ILE:HD11	1:A:335:PHE:CZ	2.38	0.59
1:A:267:ARG:HD3	2:A:1177:HOH:O	2.05	0.57
1:A:720:GLU:HG3	2:A:1318:HOH:O	2.05	0.56
1:A:822:GLU:HG3	2:A:1298:HOH:O	2.05	0.55
1:A:354:GLY:CA	1:A:377:ILE:HD11	2.35	0.55
1:A:232:TYR:HB3	2:A:1212:HOH:O	2.07	0.55
1:A:708:LYS:HE2	2:A:1386:HOH:O	2.07	0.54
1:A:221:ASP:O	1:A:222:ARG:O	2.26	0.54
1:A:495:GLN:HG2	2:A:1612:HOH:O	2.08	0.54
1:A:177:ARG:HG3	1:A:422:VAL:HG13	1.88	0.53
1:A:557:ASN:ND2	1:A:560:LYS:H	2.07	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:820:ASN:HD22	1:A:825:GLN:HG2	1.74	0.52
1:A:189:GLN:HG3	2:A:1367:HOH:O	2.10	0.51
1:A:202:ASN:ND2	1:A:251:GLN:HE22	2.09	0.51
1:A:279:HIS:CE1	2:A:1618:HOH:O	2.63	0.50
1:A:267:ARG:CD	2:A:1177:HOH:O	2.59	0.50
1:A:889:LYS:NZ	1:A:889:LYS:H	2.10	0.50
1:A:521:LYS:HD3	2:A:1332:HOH:O	2.12	0.50
1:A:332:PRO:HG2	2:A:1566:HOH:O	2.12	0.49
1:A:839:LYS:HD2	1:A:853:VAL:HG23	1.95	0.49
1:A:239:THR:HG21	1:A:293:ASP:OD2	2.13	0.49
1:A:420:LEU:HD22	1:A:435:MET:HE1	1.94	0.48
1:A:278:LYS:HG3	1:A:279:HIS:CD2	2.50	0.46
1:A:817:LEU:C	1:A:817:LEU:HD13	2.35	0.46
1:A:282:ASN:HB2	2:A:1292:HOH:O	2.15	0.46
1:A:890:LEU:CB	2:A:1455:HOH:O	2.64	0.46
1:A:254:ASN:C	1:A:254:ASN:HD22	2.18	0.46
1:A:341:ASN:ND2	2:A:1515:HOH:O	2.36	0.45
1:A:341:ASN:HB2	2:A:1459:HOH:O	2.17	0.44
1:A:720:GLU:CG	2:A:1318:HOH:O	2.64	0.44
1:A:581:LYS:HB3	1:A:768:GLU:HB2	2.00	0.44
1:A:855:LYS:HE3	1:A:855:LYS:HB3	1.80	0.44
1:A:468:ASN:C	1:A:468:ASN:HD22	2.21	0.43
1:A:323:ASP:HA	2:A:1596:HOH:O	2.18	0.43
1:A:316:GLU:CD	1:A:316:GLU:N	2.72	0.42
1:A:493:THR:HG21	2:A:1344:HOH:O	2.19	0.42
1:A:493:THR:HG22	2:A:1493:HOH:O	2.18	0.42
1:A:296:ILE:CD1	1:A:335:PHE:CE1	3.01	0.42
1:A:375:ARG:NH2	2:A:1611:HOH:O	2.53	0.42
1:A:242:TYR:CD2	1:A:298:THR:HG23	2.55	0.41
1:A:490:GLU:O	1:A:493:THR:HG22	2.20	0.41
1:A:238:LEU:HB2	2:A:1621:HOH:O	2.20	0.41
1:A:385:ASP:HB2	2:A:1588:HOH:O	2.20	0.40
1:A:219:GLN:CA	1:A:219:GLN:HE21	2.35	0.40
1:A:241:THR:HG21	1:A:276:MET:HE3	2.03	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	719/721 (100%)	694 (96%)	23 (3%)	2 (0%)	41 19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	ARG
1	A	221	ASP

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	637/639 (100%)	595 (93%)	42 (7%)	16 1

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	177	ARG
1	A	189	GLN
1	A	203	GLN
1	A	219	GLN
1	A	221	ASP
1	A	222	ARG
1	A	228	LYS
1	A	232	TYR
1	A	252	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	254	ASN
1	A	270	ARG
1	A	278	LYS
1	A	282	ASN
1	A	319	LYS
1	A	327	LYS
1	A	337	LYS
1	A	364	ARG
1	A	375	ARG
1	A	380	VAL
1	A	383	LEU
1	A	404	TYR
1	A	427	LYS
1	A	434	LYS
1	A	444	LYS
1	A	468	ASN
1	A	490	GLU
1	A	492	GLU
1	A	495	GLN
1	A	529	LEU
1	A	557	ASN
1	A	626	LYS
1	A	720	GLU
1	A	748	LYS
1	A	780	LEU
1	A	808	GLN
1	A	813	LEU
1	A	822	GLU
1	A	849	ASN
1	A	855	LYS
1	A	869	LYS
1	A	874	ASN
1	A	889	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	202	ASN
1	A	254	ASN
1	A	279	HIS
1	A	282	ASN
1	A	379	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	386	GLN
1	A	436	GLN
1	A	468	ASN
1	A	483	HIS
1	A	557	ASN
1	A	667	HIS
1	A	698	GLN
1	A	759	GLN
1	A	789	ASN
1	A	820	ASN
1	A	825	GLN
1	A	832	GLN
1	A	874	ASN

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 monosaccharides 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, 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	721/721 (100%)	0.34	58 (8%) 12 14	11, 20, 40, 61	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	232	TYR	9.0
1	A	224	TYR	8.9
1	A	223	ILE	8.8
1	A	221	ASP	5.8
1	A	427	LYS	5.8
1	A	225	LEU	5.7
1	A	219	GLN	5.7
1	A	340	ASP	5.2
1	A	233	LYS	4.9
1	A	433	ASP	4.7
1	A	226	TRP	4.3
1	A	171	LYS	4.1
1	A	170	VAL	4.0
1	A	375	ARG	4.0
1	A	341	ASN	3.8
1	A	216	ILE	3.7
1	A	866	ASP	3.6
1	A	278	LYS	3.5
1	A	229	PHE	3.5
1	A	228	LYS	3.4
1	A	490	GLU	3.3
1	A	721	GLN	3.3
1	A	174	TYR	3.3
1	A	865	GLY	3.2
1	A	704	SER	3.2
1	A	849	ASN	3.2
1	A	218	SER	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	689	THR	3.1
1	A	279	HIS	3.1
1	A	277	HIS	3.0
1	A	766	ILE	2.9
1	A	867	GLU	2.8
1	A	230	SER	2.7
1	A	890	LEU	2.7
1	A	284	GLU	2.6
1	A	220	ALA	2.5
1	A	196	ASP	2.5
1	A	837	ILE	2.4
1	A	864	GLU	2.4
1	A	431	ASP	2.4
1	A	368	GLN	2.4
1	A	492	GLU	2.3
1	A	288	VAL	2.3
1	A	491	GLY	2.3
1	A	705	ASN	2.2
1	A	276	MET	2.2
1	A	239	THR	2.1
1	A	227	GLU	2.1
1	A	194	LYS	2.1
1	A	289	GLY	2.1
1	A	275	TRP	2.1
1	A	703	SER	2.0
1	A	287	ILE	2.0
1	A	772	ASP	2.0
1	A	189	GLN	2.0
1	A	280	VAL	2.0
1	A	217	SER	2.0
1	A	290	ASN	2.0

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 monosaccharides 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.