



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 17, 2025 – 12:11 pm GMT

PDB ID : 9GN3  
Title : The 3D structure of MsrR from Streptococcus pneumoniae  
Authors : Moche, M.; Sala, B.M.; Sandalova, T.; Achour, A.  
Deposited on : 2024-08-30  
Resolution : 2.15 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

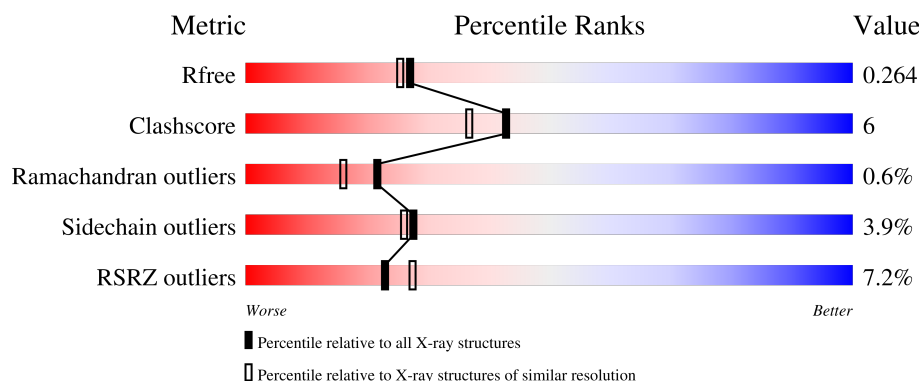
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)

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  $\geq 3$ , 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	307	<div> <div>9%</div> <div>78%</div> <div>17%</div> <div>• •</div> </div>
1	B	307	<div> <div>5%</div> <div>81%</div> <div>14%</div> <div>• •</div> </div>
1	C	307	<div> <div>7%</div> <div>83%</div> <div>12%</div> <div>• •</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7158 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 Regulatory protein MsrR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	1	0
			2295	1440	387	460	8			
1	B	294	Total	C	N	O	S	0	5	0
			2307	1445	389	465	8			
1	C	298	Total	C	N	O	S	0	3	0
			2330	1458	395	469	8			

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	129	MET	-	initiating methionine	UNP Q8DPD6
A	130	GLY	-	expression tag	UNP Q8DPD6
A	425	ALA	-	expression tag	UNP Q8DPD6
A	426	ALA	-	expression tag	UNP Q8DPD6
A	427	ALA	-	expression tag	UNP Q8DPD6
A	428	LEU	-	expression tag	UNP Q8DPD6
A	429	GLU	-	expression tag	UNP Q8DPD6
A	430	HIS	-	expression tag	UNP Q8DPD6
A	431	HIS	-	expression tag	UNP Q8DPD6
A	432	HIS	-	expression tag	UNP Q8DPD6
A	433	HIS	-	expression tag	UNP Q8DPD6
A	434	HIS	-	expression tag	UNP Q8DPD6
A	435	HIS	-	expression tag	UNP Q8DPD6
B	129	MET	-	initiating methionine	UNP Q8DPD6
B	130	GLY	-	expression tag	UNP Q8DPD6
B	425	ALA	-	expression tag	UNP Q8DPD6
B	426	ALA	-	expression tag	UNP Q8DPD6
B	427	ALA	-	expression tag	UNP Q8DPD6
B	428	LEU	-	expression tag	UNP Q8DPD6
B	429	GLU	-	expression tag	UNP Q8DPD6
B	430	HIS	-	expression tag	UNP Q8DPD6
B	431	HIS	-	expression tag	UNP Q8DPD6
B	432	HIS	-	expression tag	UNP Q8DPD6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	433	HIS	-	expression tag	UNP Q8DPD6
B	434	HIS	-	expression tag	UNP Q8DPD6
B	435	HIS	-	expression tag	UNP Q8DPD6
C	129	MET	-	initiating methionine	UNP Q8DPD6
C	130	GLY	-	expression tag	UNP Q8DPD6
C	425	ALA	-	expression tag	UNP Q8DPD6
C	426	ALA	-	expression tag	UNP Q8DPD6
C	427	ALA	-	expression tag	UNP Q8DPD6
C	428	LEU	-	expression tag	UNP Q8DPD6
C	429	GLU	-	expression tag	UNP Q8DPD6
C	430	HIS	-	expression tag	UNP Q8DPD6
C	431	HIS	-	expression tag	UNP Q8DPD6
C	432	HIS	-	expression tag	UNP Q8DPD6
C	433	HIS	-	expression tag	UNP Q8DPD6
C	434	HIS	-	expression tag	UNP Q8DPD6
C	435	HIS	-	expression tag	UNP Q8DPD6

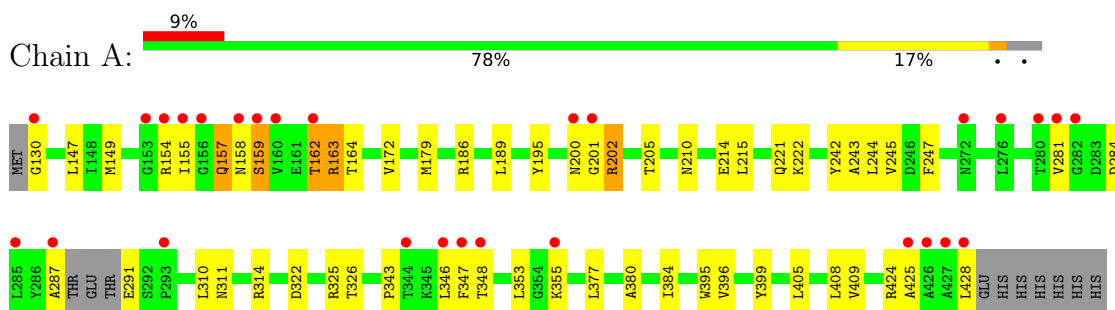
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	72	Total O 72 72	0	0
2	B	81	Total O 81 81	0	0
2	C	73	Total O 73 73	0	0

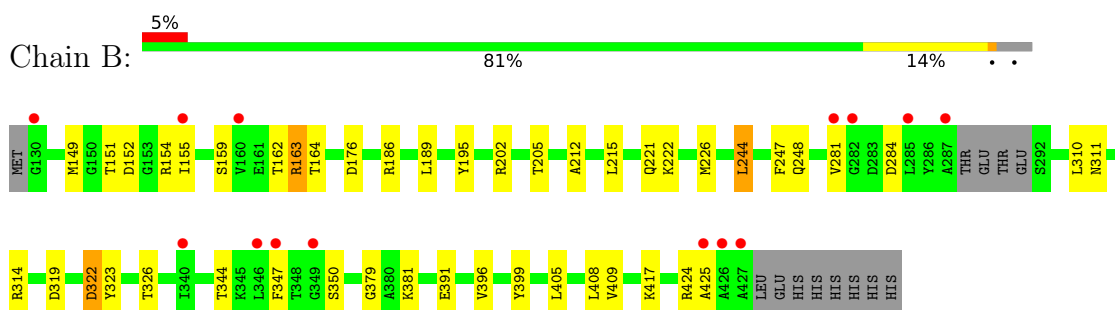
### 3 Residue-property plots [i](#)

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.

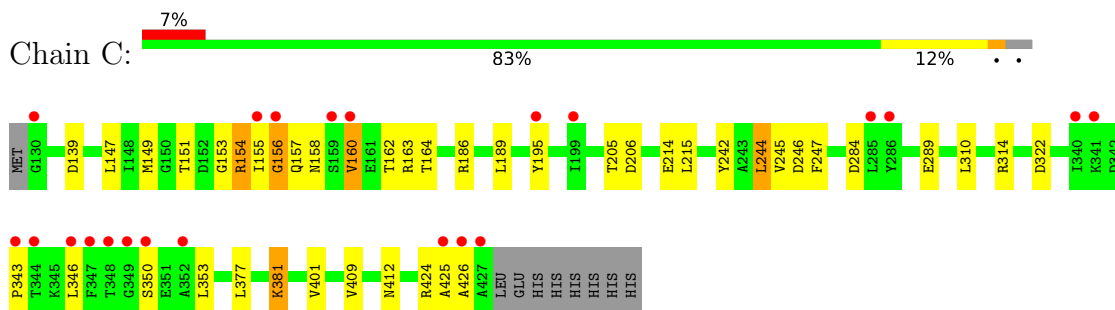
#### • Molecule 1: Regulatory protein MsrR



#### • Molecule 1: Regulatory protein MsrR



#### • Molecule 1: Regulatory protein MsrR



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.63Å 130.63Å 100.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.00 – 2.15 46.00 – 2.15	Depositor EDS
% Data completeness (in resolution range)	64.2 (46.00-2.15) 64.2 (46.00-2.15)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.204 , 0.259 0.211 , 0.264	Depositor DCC
$R_{free}$ test set	2628 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.0	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 57.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.013 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7158	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.17 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.7669e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.42	0/2324	0.81	1/3141 (0.0%)
1	B	0.41	0/2336	0.79	0/3157
1	C	0.40	0/2360	0.80	2/3191 (0.1%)
All	All	0.41	0/7020	0.80	3/9489 (0.0%)

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	4

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	244[A]	LEU	CB-CG-CD1	-5.10	102.33	111.00
1	C	244[B]	LEU	CB-CG-CD1	-5.10	102.33	111.00
1	A	325	ARG	NE-CZ-NH2	-5.01	117.80	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	154	ARG	Sidechain
1	A	163	ARG	Sidechain
1	A	186	ARG	Sidechain
1	A	202	ARG	Sidechain

## 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	2295	0	2272	30	0
1	B	2307	0	2271	25	0
1	C	2330	0	2301	34	0
2	A	72	0	0	1	0
2	B	81	0	0	3	0
2	C	73	0	0	2	2
All	All	7158	0	6844	84	2

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.

The worst 5 of 84 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:322:ASP:OD2	2:B:501:HOH:O	1.84	0.94
1:A:322:ASP:OD2	2:A:501:HOH:O	1.87	0.93
1:B:379:GLY:O	2:B:502:HOH:O	1.91	0.89
1:C:158:ASN:HB3	1:C:289:GLU:OE2	1.74	0.88
1:A:130:GLY:HA3	1:A:158:ASN:HB3	1.62	0.82

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:523:HOH:O	2:C:553:HOH:O[5_556]	2.07	0.13
2:C:560:HOH:O	2:C:560:HOH:O[5_556]	2.07	0.13

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	293/307 (95%)	276 (94%)	15 (5%)	2 (1%)	19	13
1	B	295/307 (96%)	284 (96%)	10 (3%)	1 (0%)	37	34
1	C	299/307 (97%)	283 (95%)	14 (5%)	2 (1%)	19	13
All	All	887/921 (96%)	843 (95%)	39 (4%)	5 (1%)	22	16

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	155	ILE
1	A	155	ILE
1	C	156	GLY
1	C	154	ARG
1	A	157	GLN

### 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	247/258 (96%)	237 (96%)	10 (4%)	27	25
1	B	248/258 (96%)	237 (96%)	11 (4%)	24	22
1	C	251/258 (97%)	242 (96%)	9 (4%)	30	29
All	All	746/774 (96%)	716 (96%)	30 (4%)	27	25

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	244[A]	LEU
1	C	381	LYS
1	B	322	ASP
1	C	412	ASN
1	C	186	ARG

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

Mol	Chain	Res	Type
1	A	157	GLN
1	B	272	ASN
1	C	157	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 oligosaccharides 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	296/307 (96%)	0.43	28 (9%)	15 18	16, 46, 111, 144	1 (0%)
1	B	294/307 (95%)	0.32	14 (4%)	36 42	15, 46, 104, 132	5 (1%)
1	C	298/307 (97%)	0.43	22 (7%)	22 26	17, 49, 96, 134	3 (1%)
All	All	888/921 (96%)	0.40	64 (7%)	23 27	15, 48, 105, 144	9 (1%)

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	347	PHE	5.3
1	B	427	ALA	4.9
1	A	427	ALA	4.6
1	A	428	LEU	4.5
1	B	155	ILE	4.4

### 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.