



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 08:14 am BST

PDB ID : 4D8B  
Title : High resolution structure of monomeric S. progenies SpeB reveals role of glycine-rich active site loop  
Authors : Gonzalez, G.E.; Wolan, D.W.  
Deposited on : 2012-01-10  
Resolution : 1.06 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
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

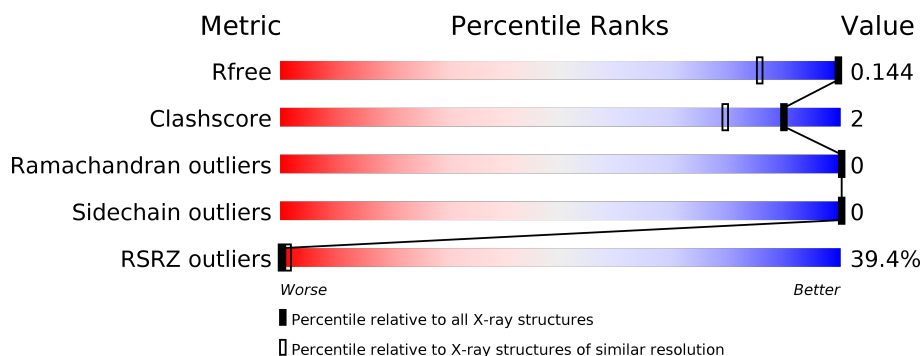
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.06 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	1202 (1.10-1.02)
Clashscore	141614	1252 (1.10-1.02)
Ramachandran outliers	138981	1204 (1.10-1.02)
Sidechain outliers	138945	1202 (1.10-1.02)
RSRZ outliers	127900	1178 (1.10-1.02)

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	261	<div> <div>38%</div> <div>95%</div> <div>..</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4274 atoms, of which 1865 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.

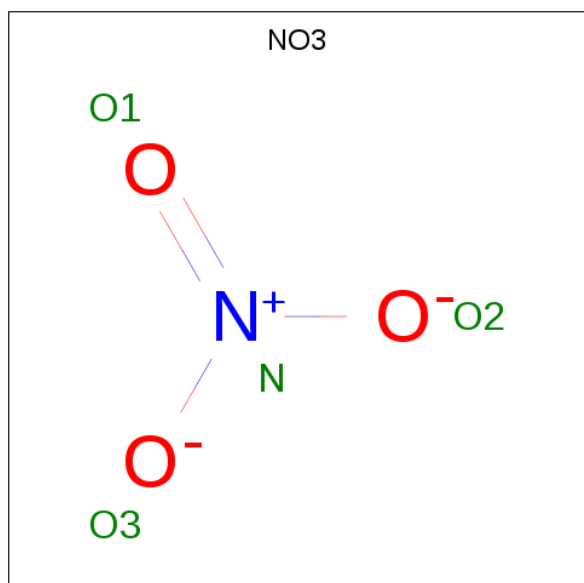
- Molecule 1 is a protein called Streptopain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	254	3821	1235	1865	341	375	5	0	2	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	399	LEU	-	EXPRESSION TAG	UNP E0PTS8
A	400	GLU	-	EXPRESSION TAG	UNP E0PTS8
A	401	HIS	-	EXPRESSION TAG	UNP E0PTS8
A	402	HIS	-	EXPRESSION TAG	UNP E0PTS8
A	403	HIS	-	EXPRESSION TAG	UNP E0PTS8
A	404	HIS	-	EXPRESSION TAG	UNP E0PTS8
A	405	HIS	-	EXPRESSION TAG	UNP E0PTS8
A	406	HIS	-	EXPRESSION TAG	UNP E0PTS8

- Molecule 2 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	N	O	0	0
			4	1	3		

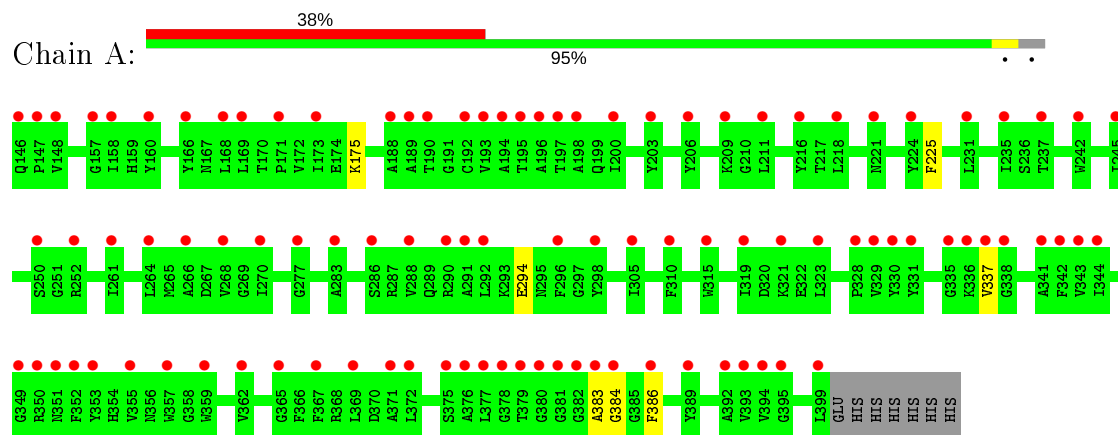
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	449	Total	O	0	0
			449	449		

### 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 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: Streptopain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.99Å 50.11Å 41.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.68 – 1.06 19.68 – 1.06	Depositor EDS
% Data completeness (in resolution range)	96.4 (19.68-1.06) 96.4 (19.68-1.06)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 1.06Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.144 , 0.160 0.127 , 0.144	Depositor DCC
$R_{free}$ test set	5359 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	9.8	Xtriage
Anisotropy	0.177	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.46 , 51.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	4274	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.06% 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: NO3

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	0.40	0/2013	0.60	0/2733

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	1956	1865	1858	7	0
2	A	4	0	0	0	0
3	A	449	0	0	3	1
All	All	2409	1865	1858	7	1

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 2.

All (7) 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:294:GLU:OE1	3:A:892:HOH:O	2.12	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:175:LYS:HE2	1:A:225:PHE:O	2.09	0.53
1:A:337:VAL:HG23	1:A:337:VAL:O	2.11	0.51
1:A:175:LYS:CE	3:A:764:HOH:O	2.58	0.50
1:A:337:VAL:HG22	1:A:383:ALA:HB3	1.99	0.45
1:A:384:GLY:HA2	1:A:386:PHE:CZ	2.55	0.41
1:A:175:LYS:NZ	3:A:764:HOH:O	2.12	0.40

All (1) 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)
3:A:955:HOH:O	3:A:955:HOH:O[2_555]	2.13	0.07

## 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	254/261 (97%)	250 (98%)	4 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	205/213 (96%)	205 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

1 ligand is 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	NO3	A	501	-	1,3,3	3.34	1 (100%)	0,3,3	0.00	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	NO3	O1-N	3.34	1.39	1.24

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	254/261 (97%)	2.02	100 (39%) 0 1	6, 9, 16, 24	0

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	337	VAL	13.5
1	A	377	LEU	8.1
1	A	381	GLY	8.1
1	A	383	ALA	7.1
1	A	382	GLY	6.9
1	A	399	LEU	6.3
1	A	389	TYR	5.9
1	A	379	THR	5.8
1	A	336	LYS	5.6
1	A	380	GLY	5.3
1	A	147	PRO	5.2
1	A	148	VAL	5.2
1	A	376	ALA	4.9
1	A	384	GLY	4.7
1	A	349	GLY	3.9
1	A	335	GLY	3.9
1	A	338	GLY	3.5
1	A	321	LYS	3.4
1	A	343	VAL	3.4
1	A	355	VAL	3.3
1	A	211	LEU	3.2
1	A	378	GLY	3.1
1	A	315	TRP	3.1
1	A	357	TRP	3.1
1	A	352	PHE	3.0
1	A	158	ILE	3.0
1	A	266	ALA	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	206	TYR	2.9
1	A	393	VAL	2.9
1	A	231[A]	LEU	2.8
1	A	190	THR	2.8
1	A	192	CYS	2.8
1	A	330	TYR	2.8
1	A	173	ILE	2.7
1	A	331	TYR	2.7
1	A	292	LEU	2.6
1	A	310	PHE	2.6
1	A	168	LEU	2.6
1	A	209	LYS	2.6
1	A	291	ALA	2.6
1	A	329	VAL	2.6
1	A	394	VAL	2.6
1	A	264	LEU	2.6
1	A	296	PHE	2.6
1	A	252	ARG	2.5
1	A	235	ILE	2.5
1	A	351	ASN	2.5
1	A	350	ARG	2.5
1	A	237	THR	2.5
1	A	359	TRP	2.5
1	A	203	TYR	2.5
1	A	298	TYR	2.5
1	A	342	PHE	2.5
1	A	367	PHE	2.5
1	A	386	PHE	2.5
1	A	196	ALA	2.4
1	A	290	ARG	2.4
1	A	160	TYR	2.4
1	A	198	ALA	2.4
1	A	375	SER	2.4
1	A	344	ILE	2.4
1	A	169	LEU	2.4
1	A	372	LEU	2.4
1	A	146	GLN	2.4
1	A	189	ALA	2.4
1	A	283	ALA	2.4
1	A	195	THR	2.4
1	A	288	VAL	2.4
1	A	157	GLY	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	369	LEU	2.4
1	A	261	ILE	2.3
1	A	216	TYR	2.3
1	A	224	TYR	2.3
1	A	362	VAL	2.3
1	A	221	ASN	2.3
1	A	200	ILE	2.3
1	A	270	ILE	2.3
1	A	353	TYR	2.3
1	A	197	THR	2.2
1	A	268	VAL	2.2
1	A	166	TYR	2.2
1	A	250	SER	2.2
1	A	323	LEU	2.2
1	A	305	ILE	2.2
1	A	194	ALA	2.2
1	A	341	ALA	2.2
1	A	328	PRO	2.2
1	A	319	ILE	2.2
1	A	371	ALA	2.2
1	A	193	VAL	2.2
1	A	277	GLY	2.2
1	A	395	GLY	2.2
1	A	245	ILE	2.2
1	A	286[A]	SER	2.1
1	A	188	ALA	2.1
1	A	392	ALA	2.1
1	A	218	LEU	2.1
1	A	242	TRP	2.0
1	A	171	PRO	2.0
1	A	365	GLY	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

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	NO3	A	501	4/4	0.82	0.20	14,16,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.