



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 06:27 pm BST

PDB ID : 3SR9  
Title : Crystal structure of mouse PTPsigma  
Authors : Wang, J.; Hou, L.; Li, J.; Ding, J.  
Deposited on : 2011-07-07  
Resolution : 2.40 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  
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

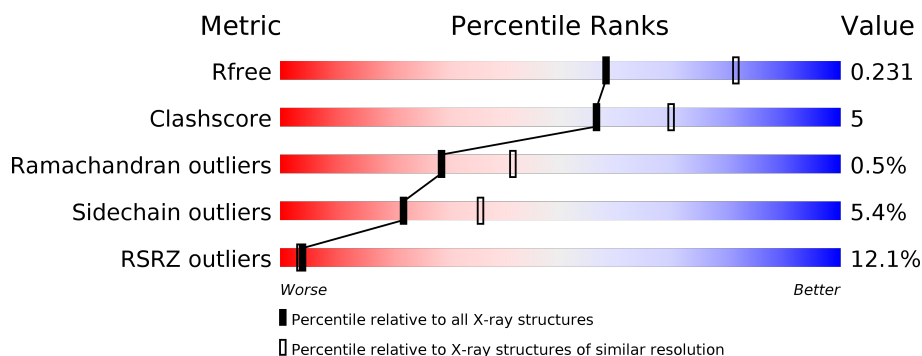
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	583	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4686 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 Receptor-type tyrosine-protein phosphatase S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	568	4597	2915	799	854	29	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1319	MET	-	EXPRESSION TAG	UNP B0V2N1
A	1320	HIS	-	EXPRESSION TAG	UNP B0V2N1
A	1321	HIS	-	EXPRESSION TAG	UNP B0V2N1
A	1322	HIS	-	EXPRESSION TAG	UNP B0V2N1
A	1323	HIS	-	EXPRESSION TAG	UNP B0V2N1
A	1324	HIS	-	EXPRESSION TAG	UNP B0V2N1
A	1325	HIS	-	EXPRESSION TAG	UNP B0V2N1

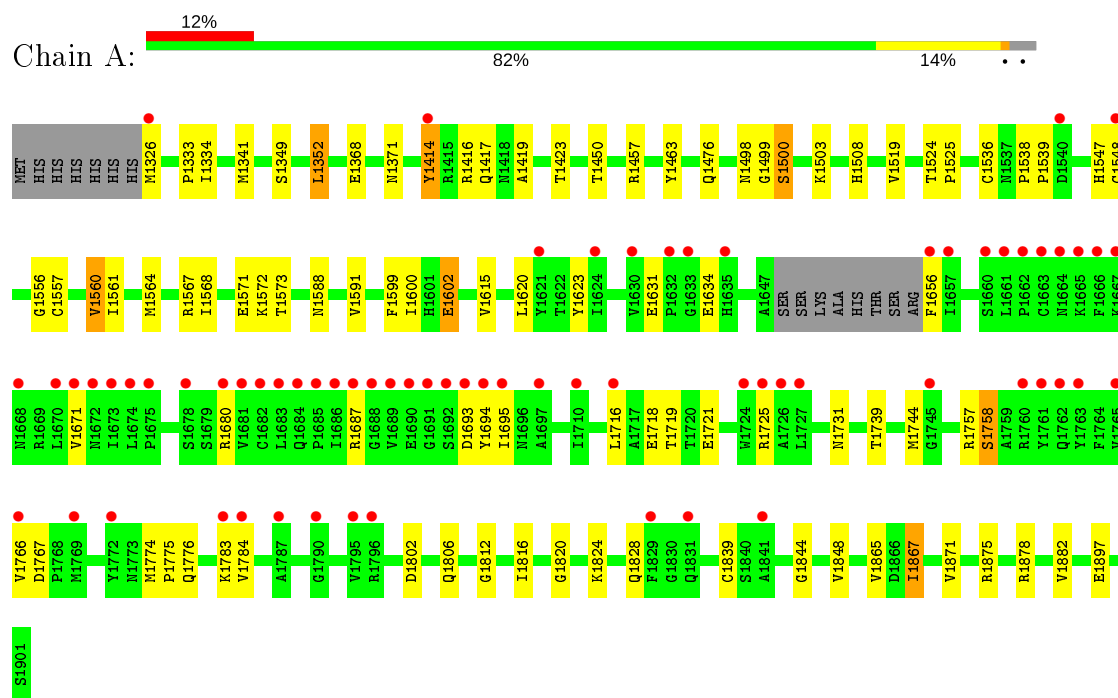
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	89	Total	O	0	0
			89	89		

### 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: Receptor-type tyrosine-protein phosphatase S



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.65Å 94.65Å 124.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.40 49.50 – 2.40	Depositor EDS
% Data completeness (in resolution range)	90.4 (50.00-2.40) 90.4 (49.50-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.233 , 0.272 0.227 , 0.231	Depositor DCC
$R_{free}$ test set	1156 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.4	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 48.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.047 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4686	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.28% 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](#)

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	0.37	0/4711	0.50	0/6382

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	4597	0	4470	45	0
2	A	89	0	0	2	0
All	All	4686	0	4470	45	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 (45) 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:1757:ARG:HG3	1:A:1758:SER:H	1.36	0.90
1:A:1757:ARG:HG3	1:A:1758:SER:N	1.90	0.86
1:A:1556:GLY:O	1:A:1560:VAL:HG23	1.80	0.81
1:A:1620:LEU:HD13	1:A:1865:VAL:HG11	1.67	0.77

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1744:MET:SD	1:A:1806:GLN:NE2	2.60	0.74
1:A:1349:SER:HB3	1:A:1352:LEU:HD22	1.73	0.70
1:A:1414:TYR:HE1	2:A:22:HOH:O	1.78	0.66
1:A:1725:ARG:HH11	1:A:1725:ARG:HG2	1.61	0.65
1:A:1716:LEU:H	1:A:1719:THR:HB	1.61	0.65
1:A:1463:TYR:OH	1:A:1508:HIS:HE1	1.80	0.64
1:A:1557:CYS:O	1:A:1561:ILE:HG12	1.97	0.63
1:A:1561:ILE:HD11	1:A:1600:ILE:HA	1.80	0.63
1:A:1820:GLY:O	1:A:1824:LYS:HG2	1.98	0.63
1:A:1417:GLN:CD	1:A:1417:GLN:H	2.02	0.63
1:A:1680:ARG:NH2	1:A:1694:TYR:HA	2.16	0.61
1:A:1739:THR:CG2	1:A:1839:CYS:O	2.51	0.58
1:A:1334:ILE:HG12	1:A:1572:LYS:O	2.05	0.56
1:A:1767:ASP:HB2	1:A:1783:LYS:HB3	1.86	0.56
1:A:1416:ARG:HB3	1:A:1419:ALA:HB2	1.87	0.56
1:A:1757:ARG:CG	1:A:1758:SER:N	2.67	0.55
1:A:1766:VAL:HG22	1:A:1784:VAL:HG12	1.90	0.54
1:A:1615:VAL:HG21	1:A:1623:TYR:CG	2.45	0.52
1:A:1341:MET:HE1	1:A:1602:GLU:HG3	1.92	0.51
1:A:1414:TYR:CE2	1:A:1567:ARG:HA	2.47	0.49
1:A:1333:PRO:HA	1:A:1573:THR:HG22	1.94	0.49
1:A:1423:THR:O	1:A:1547:HIS:HB2	2.13	0.49
1:A:1561:ILE:HD13	1:A:1600:ILE:HG23	1.95	0.48
1:A:1739:THR:HG21	1:A:1839:CYS:O	2.15	0.47
1:A:1508:HIS:HD2	2:A:2:HOH:O	1.98	0.45
1:A:1812:GLY:H	1:A:1897:GLU:CD	2.20	0.45
1:A:1774:MET:HB3	1:A:1775:PRO:HD2	1.99	0.44
1:A:1844:GLY:O	1:A:1848:VAL:HG23	2.18	0.44
1:A:1631:GLU:HB3	1:A:1634:GLU:HG3	2.00	0.43
1:A:1620:LEU:HD13	1:A:1865:VAL:CG1	2.42	0.42
1:A:1524:THR:HB	1:A:1525:PRO:HD3	2.01	0.42
1:A:1725:ARG:HH11	1:A:1725:ARG:CG	2.31	0.42
1:A:1739:THR:HG22	1:A:1839:CYS:O	2.18	0.42
1:A:1875:ARG:HA	1:A:1878:ARG:O	2.20	0.42
1:A:1519:VAL:HG21	1:A:1599:PHE:HB2	2.01	0.41
1:A:1564:MET:O	1:A:1568:ILE:HG23	2.20	0.41
1:A:1538:PRO:HA	1:A:1539:PRO:HD3	1.95	0.41
1:A:1414:TYR:OH	1:A:1571:GLU:HG3	2.20	0.41
1:A:1867:ILE:O	1:A:1871:VAL:HG23	2.21	0.40
1:A:1499:GLY:O	1:A:1500:SER:C	2.60	0.40
1:A:1450:THR:HG21	1:A:1548:CYS:O	2.22	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	564/583 (97%)	536 (95%)	25 (4%)	3 (0%)	29	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1500	SER
1	A	1591	VAL
1	A	1882	VAL

### 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	496/510 (97%)	469 (95%)	27 (5%)	22	36

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1326	MET
1	A	1352	LEU
1	A	1368	GLU
1	A	1371	ASN
1	A	1414	TYR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1457	ARG
1	A	1476	GLN
1	A	1498	ASN
1	A	1503	LYS
1	A	1536	CYS
1	A	1560	VAL
1	A	1588	ASN
1	A	1602	GLU
1	A	1656	PHE
1	A	1671	VAL
1	A	1687	ARG
1	A	1693	ASP
1	A	1695	ILE
1	A	1718	GLU
1	A	1721	GLU
1	A	1731	ASN
1	A	1758	SER
1	A	1776	GLN
1	A	1802	ASP
1	A	1816	ILE
1	A	1828	GLN
1	A	1867	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1476	GLN
1	A	1508	HIS
1	A	1579	HIS
1	A	1625	GLN
1	A	1750	HIS
1	A	1776	GLN
1	A	1823	HIS
1	A	1828	GLN
1	A	1831	GLN

### 5.3.3 RNA ⓘ

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 carbohydrates 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

### 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	568/583 (97%)	0.65	69 (12%) <b>4</b> <b>3</b>	21, 59, 114, 148	5 (0%)

All (69) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1686	ILE	9.4
1	A	1689	VAL	8.3
1	A	1692	SER	8.0
1	A	1663	CYS	8.0
1	A	1662	PRO	7.6
1	A	1691	GLY	6.8
1	A	1685	PRO	6.8
1	A	1665	LYS	6.6
1	A	1687	ARG	6.4
1	A	1673	ILE	6.2
1	A	1657	ILE	6.2
1	A	1761	TYR	6.2
1	A	1690	GLU	6.2
1	A	1666	PHE	5.8
1	A	1656	PHE	5.4
1	A	1688	GLY	4.7
1	A	1763	TYR	4.7
1	A	1635	HIS	4.5
1	A	1694	TYR	4.5
1	A	1660	SER	4.5
1	A	1661	LEU	4.0
1	A	1680	ARG	4.0
1	A	1716	LEU	4.0
1	A	1672	ASN	4.0
1	A	1630	VAL	3.8
1	A	1695	ILE	3.7
1	A	1727	LEU	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1841	ALA	3.7
1	A	1668	ASN	3.7
1	A	1684	GLN	3.4
1	A	1762	GLN	3.2
1	A	1633	GLY	3.2
1	A	1683	LEU	3.2
1	A	1540	ASP	3.2
1	A	1667	LYS	3.1
1	A	1674	LEU	3.0
1	A	1414	TYR	3.0
1	A	1766	VAL	3.0
1	A	1682	CYS	3.0
1	A	1783	LYS	3.0
1	A	1326	MET	3.0
1	A	1548	CYS	2.9
1	A	1829	PHE	2.9
1	A	1664	ASN	2.8
1	A	1765	VAL	2.8
1	A	1784	VAL	2.8
1	A	1697	ALA	2.7
1	A	1760	ARG	2.7
1	A	1726	ALA	2.7
1	A	1671	VAL	2.6
1	A	1681	VAL	2.6
1	A	1787	ALA	2.5
1	A	1670	LEU	2.5
1	A	1769	MET	2.3
1	A	1621	TYR	2.3
1	A	1675	PRO	2.3
1	A	1693	ASP	2.3
1	A	1790	GLY	2.2
1	A	1678	SER	2.2
1	A	1796	ARG	2.2
1	A	1724	TRP	2.2
1	A	1795	VAL	2.2
1	A	1710	ILE	2.2
1	A	1632	PRO	2.2
1	A	1725	ARG	2.1
1	A	1745	GLY	2.1
1	A	1772	TYR	2.1
1	A	1624	ILE	2.0
1	A	1831	GLN	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 carbohydrates 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.