



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2014 – 11:31 PM GMT

PDB ID : 4H27  
Title : Modulating the function of human serine racemase and human serine dehydratase by protein engineering  
Authors : Wang, C.Y.; Wang, A.H.  
Deposited on : 2012-09-12  
Resolution : 1.30 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

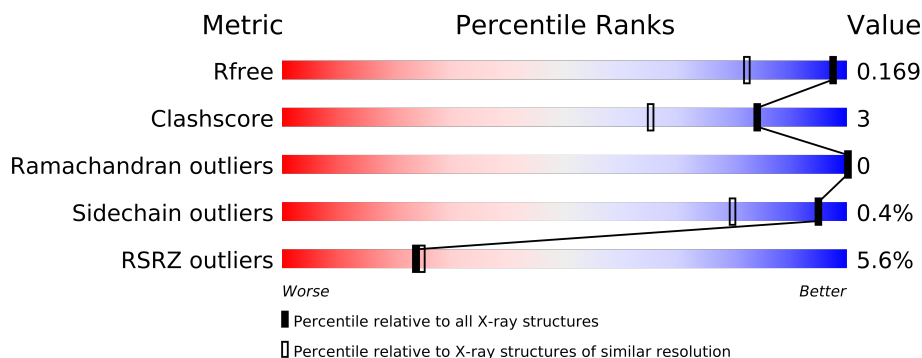
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 1.30 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))
$R_{free}$	66092	1025 (1.34-1.26)
Clashscore	79885	1140 (1.34-1.26)
Ramachandran outliers	78287	1093 (1.34-1.26)
Sidechain outliers	78261	1092 (1.34-1.26)
RSRZ outliers	66119	1025 (1.34-1.26)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	364	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	401	-	X
2	SO4	A	402	-	X
2	SO4	A	403	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2846 atoms, of which 0 are hydrogen and 0 are deuterium.

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 L-serine dehydratase/L-threoninedeaminase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	318	2369	1518	401	437	1	12	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

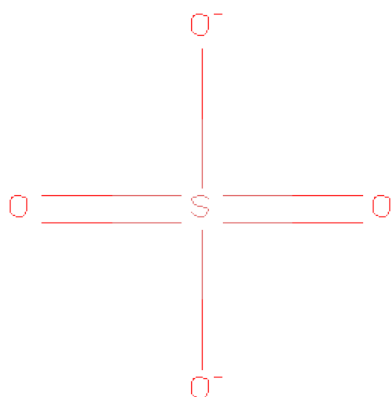
Chain	Residue	Modelled	Actual	Comment	Reference
A	-35	MET	-	EXPRESSION TAG	UNP P20132
A	-34	GLY	-	EXPRESSION TAG	UNP P20132
A	-33	SER	-	EXPRESSION TAG	UNP P20132
A	-32	SER	-	EXPRESSION TAG	UNP P20132
A	-31	HIS	-	EXPRESSION TAG	UNP P20132
A	-30	HIS	-	EXPRESSION TAG	UNP P20132
A	-29	HIS	-	EXPRESSION TAG	UNP P20132
A	-28	HIS	-	EXPRESSION TAG	UNP P20132
A	-27	HIS	-	EXPRESSION TAG	UNP P20132
A	-26	HIS	-	EXPRESSION TAG	UNP P20132
A	-25	SER	-	EXPRESSION TAG	UNP P20132
A	-24	SER	-	EXPRESSION TAG	UNP P20132
A	-23	GLY	-	EXPRESSION TAG	UNP P20132
A	-22	LEU	-	EXPRESSION TAG	UNP P20132
A	-21	VAL	-	EXPRESSION TAG	UNP P20132
A	-20	PRO	-	EXPRESSION TAG	UNP P20132
A	-19	ARG	-	EXPRESSION TAG	UNP P20132
A	-18	GLY	-	EXPRESSION TAG	UNP P20132
A	-17	SER	-	EXPRESSION TAG	UNP P20132
A	-16	HIS	-	EXPRESSION TAG	UNP P20132
A	-15	MET	-	EXPRESSION TAG	UNP P20132
A	-14	ALA	-	EXPRESSION TAG	UNP P20132
A	-13	SER	-	EXPRESSION TAG	UNP P20132
A	-12	MET	-	EXPRESSION TAG	UNP P20132
A	-11	THR	-	EXPRESSION TAG	UNP P20132
A	-10	GLY	-	EXPRESSION TAG	UNP P20132
A	-9	GLY	-	EXPRESSION TAG	UNP P20132

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	GLN	-	EXPRESSION TAG	UNP P20132
A	-7	GLN	-	EXPRESSION TAG	UNP P20132
A	-6	MET	-	EXPRESSION TAG	UNP P20132
A	-5	GLY	-	EXPRESSION TAG	UNP P20132
A	-4	ARG	-	EXPRESSION TAG	UNP P20132
A	-3	GLY	-	EXPRESSION TAG	UNP P20132
A	-2	SER	-	EXPRESSION TAG	UNP P20132
A	-1	GLU	-	EXPRESSION TAG	UNP P20132
A	0	PHE	-	EXPRESSION TAG	UNP P20132
A	65	SER	ALA	ENGINEERED MUTATION	UNP P20132
A	89	GLY	SER	SEE REMARK 999	UNP P20132

- Molecule 2 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	462	Total O 462 462	0	0

i

- Molecule 1: L-serine dehydratase/L-threoninedeaminase

Sequence logo showing the conservation of amino acids across 100 positions. The y-axis represents information content in bits (0 to 1.5). The x-axis shows residues 1 to 100. Residues 1-10 are highly conserved, with Gln10, Glu11, and Leu12 being the most prominent. Residues 11-15 show a mix of conserved and variable positions. Residues 16-20 are mostly conserved, with Val16 and Ile17 being notable. Residues 21-25 are highly conserved, with Gln21, Thr22, and Asn23 being the most prominent. Residues 26-30 are mostly conserved, with Thr26, Asn27, and Arg28 being notable. Residues 31-35 are mostly conserved, with Thr31, Asn32, and Arg33 being notable. Residues 36-40 are mostly conserved, with Thr36, Asn37, and Arg38 being notable. Residues 41-45 are mostly conserved, with Thr41, Asn42, and Arg43 being notable. Residues 46-50 are mostly conserved, with Thr46, Asn47, and Arg48 being notable. Residues 51-55 are mostly conserved, with Thr51, Asn52, and Arg53 being notable. Residues 56-60 are mostly conserved, with Thr56, Asn57, and Arg58 being notable. Residues 61-65 are mostly conserved, with Thr61, Asn62, and Arg63 being notable. Residues 66-70 are mostly conserved, with Thr66, Asn67, and Arg68 being notable. Residues 71-75 are mostly conserved, with Thr71, Asn72, and Arg73 being notable. Residues 76-80 are mostly conserved, with Thr76, Asn77, and Arg78 being notable. Residues 81-85 are mostly conserved, with Thr81, Asn82, and Arg83 being notable. Residues 86-90 are mostly conserved, with Thr86, Asn87, and Arg88 being notable. Residues 91-95 are mostly conserved, with Thr91, Asn92, and Arg93 being notable. Residues 96-100 are mostly conserved, with Thr96, Asn97, and Arg98 being notable.

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	158.50Å 158.50Å 59.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.30 29.73 – 1.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.00-1.30) 99.8 (29.73-1.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.30 (at 1.30Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.144 , 0.168 0.144 , 0.169	Depositor DCC
$R_{free}$ test set	4608 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.1	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 92128 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2846	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: LLP, 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	0.60	0/2392	0.74	3/3247 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	51	ARG	NE-CZ-NH2	-8.29	116.16	120.30
1	A	51	ARG	NE-CZ-NH1	7.37	123.99	120.30
1	A	51	ARG	CG-CD-NE	-5.88	99.46	111.80

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2369	0	0	6	0
2	A	15	0	0	1	0
3	A	462	0	0	5	2
All	All	2846	0	0	7	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including

hydrogens) of the entry. The overall clashscore for this entry is 3.

All (7) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:403:SO4:O2	3:A:827:HOH:O	1.93	0.85
1:A:100:LYS:NZ	3:A:873:HOH:O	2.17	0.78
1:A:51:ARG:NH2	3:A:788:HOH:O	2.27	0.68
1:A:89:GLY:CA	3:A:909:HOH:O	2.43	0.67
1:A:237:GLN:CG	3:A:943:HOH:O	2.49	0.61
1:A:59:HIS:CE1	1:A:84:THR:OG1	2.70	0.44
1:A:188:VAL:O	1:A:239:HIS:CE1	2.71	0.43

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	Distance(Å)	Clash(Å)
3:A:820:HOH:O	3:A:820:HOH:O[6_556]	1.50	0.70
3:A:669:HOH:O	3:A:816:HOH:O[15_456]	2.17	0.03

## 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	315/364 (86%)	308 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	246/283 (87%)	245 (100%)	1 (0%)	95	80



All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	112	LEU

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 chains in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\# Z  > 2$
1	LLP	A	41	1	24,24,25	3.92	5 (20%)	30,32,34	1.08	3 (10%)

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
1	LLP	A	41	1	-	0/15/17/19	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	41	LLP	O-C	17.71	1.23	1.11
1	A	41	LLP	O3-C3	-4.64	1.25	1.37
1	A	41	LLP	C2-N1	2.82	1.39	1.33
1	A	41	LLP	C6-N1	2.14	1.39	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	41	LLP	C4'-NZ	-2.01	1.35	1.45

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	41	LLP	C6-C5-C4	2.43	119.94	118.10
1	A	41	LLP	C2'-C2-C3	2.10	123.56	121.02
1	A	41	LLP	OP4-C5'-C5	2.00	113.34	109.26

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates i

There are no carbohydrates in this entry.

## 5.6 Ligand geometry i

3 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	401	-	4,4,4	0.51	0	6,6,6	0.43	0
2	SO4	A	402	-	4,4,4	0.47	0	6,6,6	1.15	1 (16%)
2	SO4	A	403	-	4,4,4	1.01	0	6,6,6	1.31	1 (16%)

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	401	-	-	0/0/0/0	0/0/0/0
2	SO4	A	402	-	-	0/0/0/0	0/0/0/0
2	SO4	A	403	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	402	SO4	O2-S-O1	-2.64	100.79	109.53
2	A	403	SO4	O4-S-O3	2.54	119.79	109.08

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	318/364 (87%)	0.01	18 (5%) 23 24	7, 14, 31, 38	2 (0%)

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	113	LEU	12.2
1	A	112	LEU	10.8
1	A	157	TRP	9.3
1	A	94	LEU	8.4
1	A	301	ILE	6.1
1	A	111	GLU	5.3
1	A	110	GLY	4.7
1	A	90	THR	4.5
1	A	89	GLY	4.3
1	A	5	GLU	3.9
1	A	108	VAL	3.8
1	A	318	GLU	2.8
1	A	93	ALA	2.7
1	A	91	THR	2.6
1	A	88	PRO	2.5
1	A	115	GLU	2.4
1	A	114	ASP	2.3
1	A	321	GLY	2.3

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

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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LLP	A	41	24/25	0.07	0.59	8,15,20,22	0

### 6.3 Carbohydrates

There are no carbohydrates in this entry.

### 6.4 Ligands

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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	A	403	5/5	0.29	17.87	22,27,32,33	0
2	SO4	A	401	5/5	0.13	5.79	18,20,25,27	0
2	SO4	A	402	5/5	0.18	2.60	27,31,32,32	0

### 6.5 Other polymers

There are no such residues in this entry.