



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:28 PM GMT

PDB ID : 1XDV  
Title : Experimentally Phased Structure of Human the Son of Sevenless protein at 4.1 Å.  
Authors : Sondermann, H.; Soisson, S.M.; Boykevisch, S.; Yang, S.S.; Bar-Sagi, D.; Kuriyan, J.  
Deposited on : 2004-09-08  
Resolution : 4.10 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

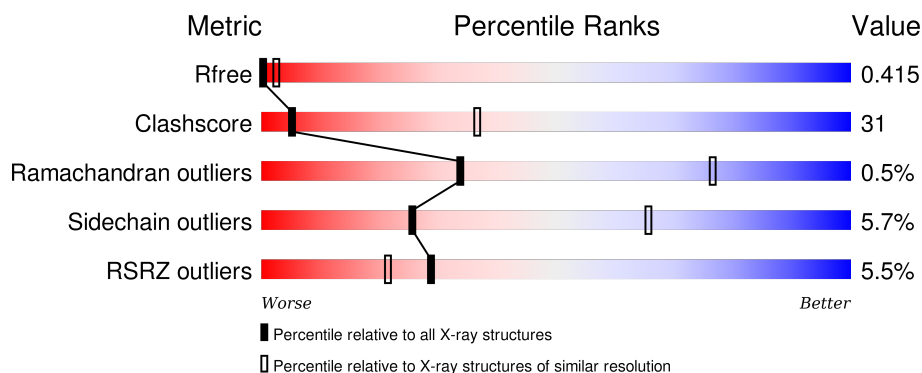
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.10 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))
$R_{free}$	91344	1018 (4.60-3.60)
Clashscore	102246	1117 (4.60-3.60)
Ramachandran outliers	100387	1063 (4.60-3.60)
Sidechain outliers	100360	1049 (4.60-3.60)
RSRZ outliers	91569	1022 (4.60-3.60)

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. 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	847	<div> <div>5%</div> <div>59%</div> <div>28%</div> <div>•</div> <div>11%</div> </div>
1	B	847	<div> <div>5%</div> <div>56%</div> <div>29%</div> <div>•</div> <div>10%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12516 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.

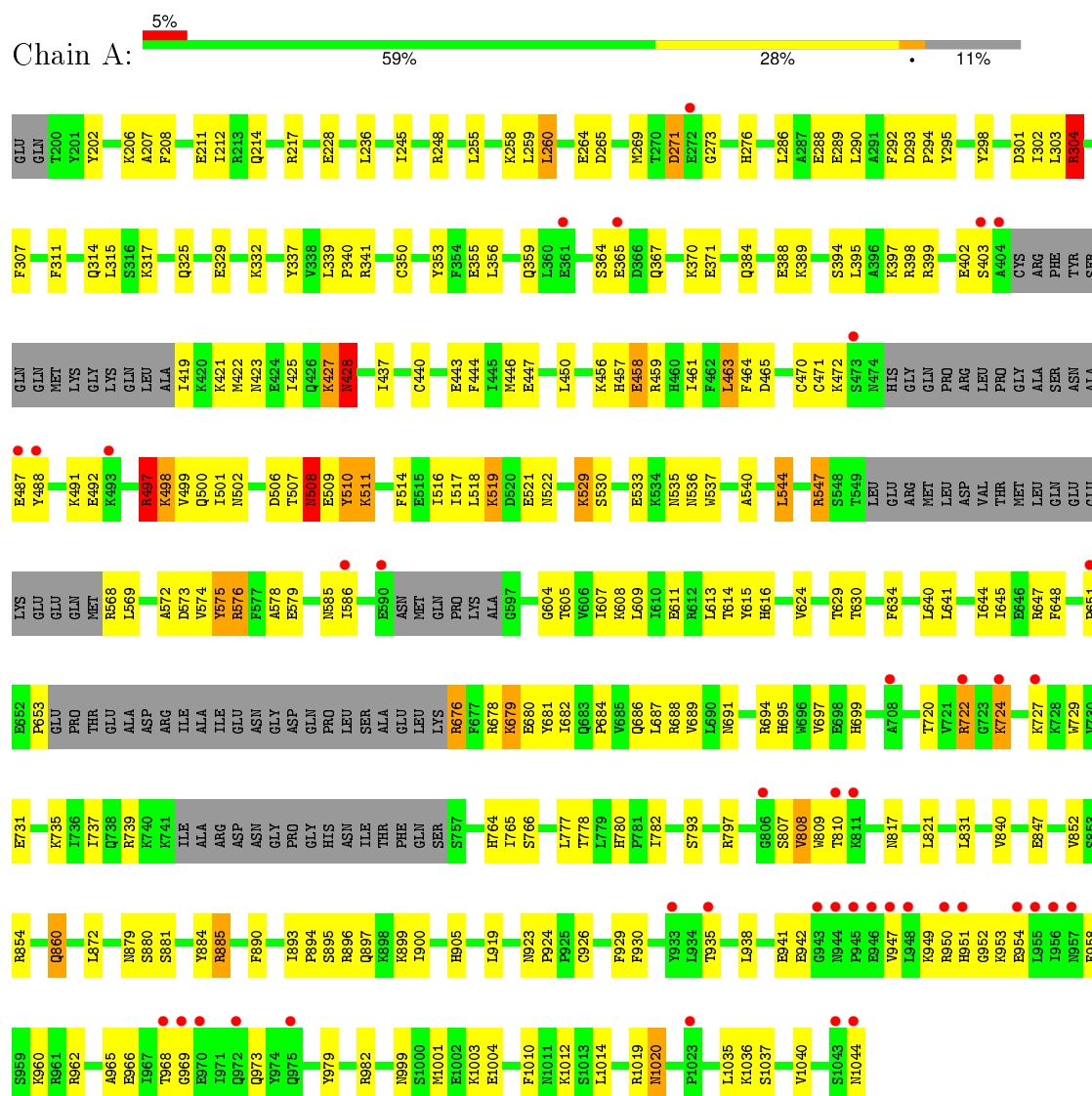
- Molecule 1 is a protein called Son of sevenless protein homolog 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	758	Total	C	N	O	S	0	0	0
			6254	4009	1066	1151	28			
1	B	759	Total	C	N	O	S	0	0	0
			6262	4015	1067	1152	28			

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Son of sevenless protein homolog 1



- Molecule 1: Son of sevenless protein homolog 1



F929	I782	V685	K608	L544	PRO	CYS	B301	GLU
F930	S793	Q686	B611	Q545	ARG	ARG	I302	GLN
T935	R797	R687	R612	Y546	LEU	PHE	L303	Y200
L938	R801	V689	L613	S548	GLY	SER	R304	Y202
E941	S802	N691	T614	T549	ALA	GLN	R310	K206
E942	S806	H695	Y615	L550	SER	GLN	F311	A207
V947	V805	W696	H616	GLU	ASN	MET	Q314	F208
L948	S807	E698	V624	ASP	ALA	LYS	L315	E211
K949	W808	R625	R626	VAL	LEU	LYS	S316	I212
H950	T810	H699	T629	THR	K491	GLN	K317	E213
H951	N817	Y702	T630	MET	K492	LEU	Q325	Q214
K952	N817	T720	T630	GLN	K493	ALA	K332	R217
K953	L821	T721	T630	GLU	F494	K421	E333	E218
E954	L821	R722	L640	GLU	F495	M422	L219	L219
L955	L831	G723	I644	LYS	M496	N423	N220	N220
I956	L831	G724	I644	GLU	K497	E424	L221	L221
F957	V840	K727	R647	GLU	K498	I425	Y337	Y337
F958	E847	E731	R648	MET	V499	Q426	Y338	K224
S959	E847	E731	E649	GLN	Q500	K427	P340	V225
K960	E847	E731	E649	GLN	I501	N428	R341	K224
K961	V852	K735	L650	R568	N502	G431	L342	E228
R962	S853	I736	P651	L568	D503	I437	C350	K235
K963	R854	I737	E652	P670	K504	G438	L236	L236
A965	Q860	I738	P653	A572	D506	F354	N244	N244
T968	L872	K739	GLU	D573	T507	Q439	L245	L245
Y979	N879	K740	THR	V574	N508	C440	R248	R248
S1000	S880	K741	GLU	R576	E509	E443	L356	L356
M1001	S881	I741	ALA	A577	Y510	F444	L357	L357
E1002	Y884	ALA	ASP	F577	K511	I445	K358	E254
K1003	R885	ARG	ARG	E579	F514	M446	Q359	L255
E1004	R890	ASN	ILE	E583	E515	E447	S364	K258
F1010	F890	HIS	ALA	E584	I516	G448	E385	K258
M1011	I893	ASN	ALA	E585	I517	T449	D366	L259
K1012	P894	THR	ILE	N586	L518	L450	Q367	L260
S1013	S895	GLY	GLU	I586	K519	K456	E368	E264
L1014	R896	PRO	ASN	E589	D520	H457	D369	E264
R1019	Q897	SER	GLY	E590	E521	E458	K370	E268
M1020	K898	GLU	LEU	ASN	F526	R459	E371	E268
L1035	I900	H764	LYS	E597	S527	H460	I377	K269
K1036	H905	I765	R676	GLN	A528	I461	G386	E272
V1040	A917	R767	R677	PRO	K529	F463	K389	H276
K918	K918	L777	K679	LEU	E532	L464	L395	E288
L919	L919	T778	E680	ALA	N535	D465	L395	E288
R920	R920	T779	R681	ALA	N536	C470	L395	F292
K926	C926	H780	I682	LYS	W537	C471	R398	D293
		P781	P684	LYS	N538	K472	R399	P294
					A539	S473	I400	Y295
					A540	N474	S401	
					L541	HIS	E402	Y298
					I542	GLY	S403	A299
					Q683	GLN	A404	R300

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.03Å 124.71Å 245.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 4.10 49.19 – 4.10	Depositor EDS
% Data completeness (in resolution range)	91.0 (8.00-4.10) 94.5 (49.19-4.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.84 (at 4.14Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.433 , 0.449 0.403 , 0.415	Depositor DCC
$R_{free}$ test set	1584 reflections (9.85%)	DCC
Wilson B-factor (Å <sup>2</sup> )	110.0	Xtriage
Anisotropy	0.632	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , -10.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 18877 reflections	Xtriage
$F_o, F_c$ correlation	0.75	EDS
Total number of atoms	12516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.39	2/6390 (0.0%)	0.61	2/8615 (0.0%)
1	B	0.40	3/6398 (0.0%)	0.61	2/8626 (0.0%)
All	All	0.39	5/12788 (0.0%)	0.61	4/17241 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	269	MET	CG-SD	6.62	1.98	1.81
1	B	427	LYS	CA-C	-5.53	1.38	1.52
1	A	427	LYS	CA-C	-5.50	1.38	1.52
1	A	428	ASN	N-CA	-5.18	1.35	1.46
1	B	428	ASN	N-CA	-5.18	1.35	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	269	MET	CG-SD-CE	7.27	111.83	100.20
1	B	1020	ASN	N-CA-C	5.13	124.86	111.00
1	A	1020	ASN	N-CA-C	5.13	124.84	111.00
1	A	304	ARG	NE-CZ-NH1	5.00	122.80	120.30

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	6254	0	6262	384	122
1	B	6262	0	6267	477	116
All	All	12516	0	12529	766	122

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 31.

The worst 5 of 766 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:269:MET:CE	1:A:729:TRP:CZ2	1.75	1.68
1:A:1019:ARG:HH22	1:B:905:HIS:CE1	1.17	1.60
1:B:342:LEU:HD21	1:B:550:LEU:CD1	1.20	1.58
1:B:269:MET:CG	1:B:691:ASN:HD21	0.94	1.58
1:A:269:MET:HE2	1:A:729:TRP:CZ2	1.25	1.56

The worst 5 of 122 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)
1:A:293:ASP:OD2	1:B:300:ARG:NE[2_765]	0.38	1.82
1:A:894:PRO:C	1:B:575:TYR:OH[3_755]	0.54	1.66
1:A:949:LYS:CA	1:B:573:ASP:OD2[1_655]	0.56	1.64
1:A:953:LYS:CA	1:B:574:VAL:CG2[1_655]	0.66	1.54
1:A:953:LYS:N	1:B:574:VAL:CA[1_655]	0.67	1.53

## 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	744/847 (88%)	709 (95%)	31 (4%)	4 (0%)	34 76
1	B	745/847 (88%)	710 (95%)	31 (4%)	4 (0%)	34 76

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1489/1694 (88%)	1419 (95%)	62 (4%)	8 (0%)	34 76

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1020	ASN
1	B	1020	ASN
1	A	497	ARG
1	A	508	ASN
1	B	497	ARG

### 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	698/773 (90%)	661 (95%)	37 (5%)	28 67
1	B	699/773 (90%)	656 (94%)	43 (6%)	23 63
All	All	1397/1546 (90%)	1317 (94%)	80 (6%)	25 65

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1040	VAL
1	B	356	LEU
1	B	885	ARG
1	B	214	GLN
1	B	258	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	888	HIS
1	B	244	ASN
1	B	866	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	936	ASN
1	B	214	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 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 [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	758/847 (89%)	0.28	41 (5%) 29 22	27, 27, 27, 27	0
1	B	759/847 (89%)	0.32	43 (5%) 27 20	27, 27, 27, 27	0
All	All	1517/1694 (89%)	0.30	84 (5%) 29 21	27, 27, 27, 27	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	496	MET	6.5
1	B	549	THR	5.6
1	B	498	LYS	5.4
1	A	969	GLY	5.0
1	B	497	ARG	4.7

### 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.