



wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 03:18 pm BST

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

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

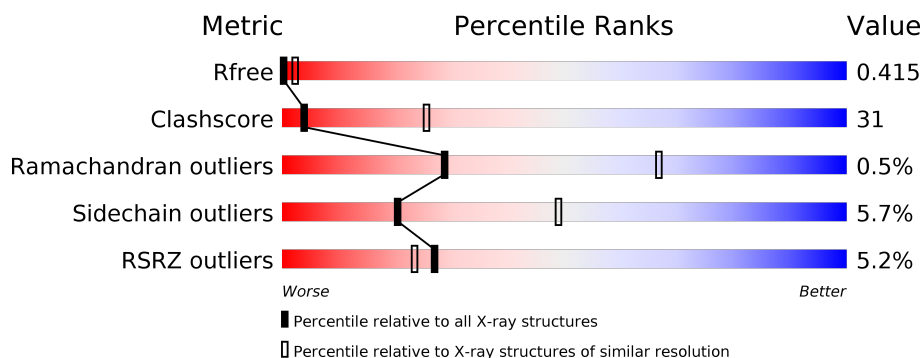
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	130704	1193 (4.50-3.70)
Clashscore	141614	1003 (4.44-3.76)
Ramachandran outliers	138981	1005 (4.48-3.72)
Sidechain outliers	138945	1199 (4.50-3.70)
RSRZ outliers	127900	1034 (4.50-3.70)

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 ≥ 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	847	<div> <div>4%</div> <div> <div></div> <div>59%</div> <div>28%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	847	<div> <div>5%</div> <div> <div></div> <div>56%</div> <div>29%</div> <div>•</div> <div>10%</div> </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.

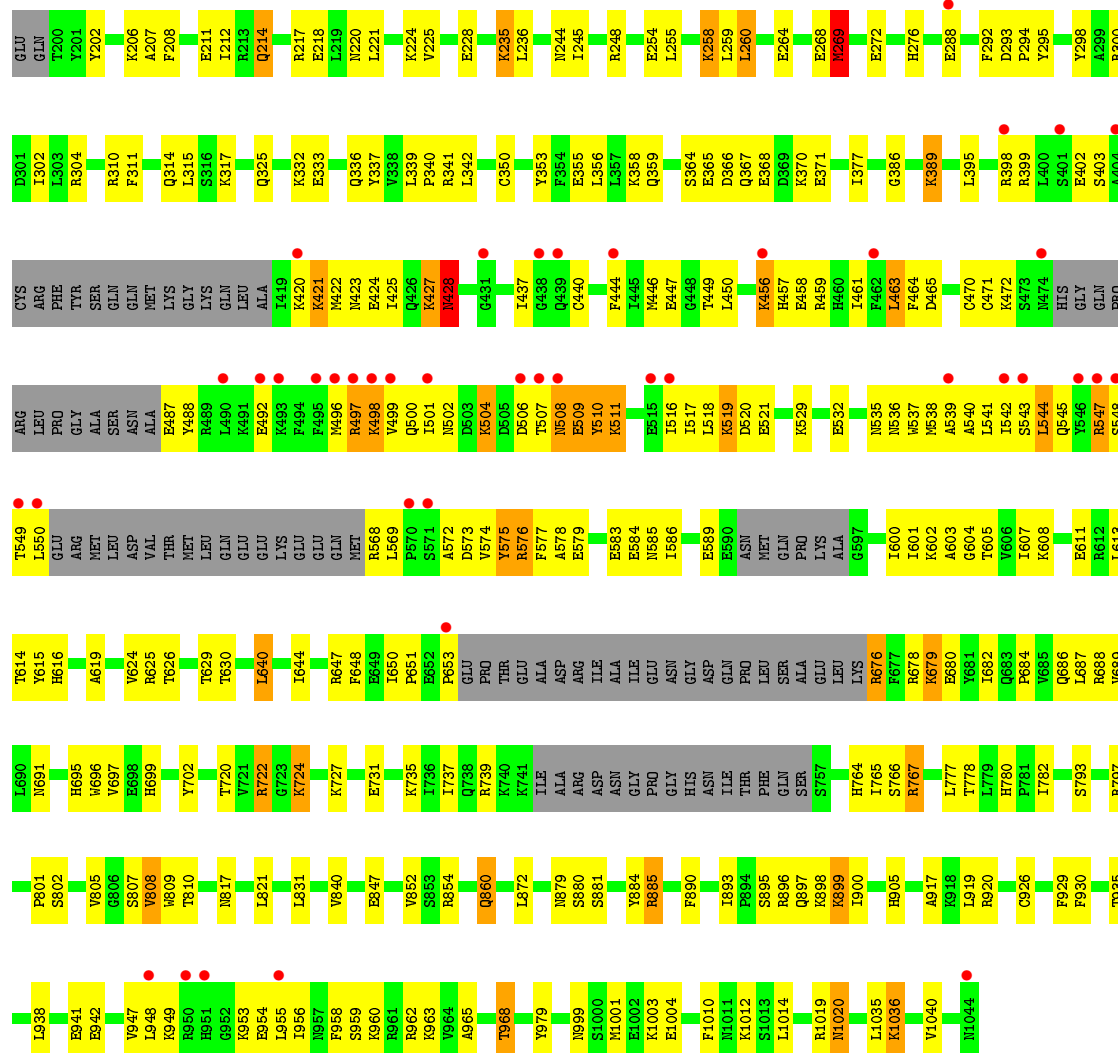
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- 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			

- Molecule 1: Son of sevenless protein homolog 1





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	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
$\langle I/\sigma(I) \rangle$ ¹	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	1858 reflections (9.84%)	wwPDB-VP
Wilson B-factor (Å ²)	110.0	Xtriage
Anisotropy	0.632	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , -10.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.75	EDS
Total number of atoms	12516	wwPDB-VP
Average B, all atoms (Å ²)	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.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.39	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

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%)	29 67
1	B	745/847 (88%)	710 (95%)	31 (4%)	4 (0%)	29 67

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%)	29 67

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%)	22 51
1	B	699/773 (90%)	656 (94%)	43 (6%)	18 46
All	All	1397/1546 (90%)	1317 (94%)	80 (6%)	20 49

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, 95th 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 ²)	Q<0.9
1	A	758/847 (89%)	0.26	37 (4%) 29 25	27, 27, 27, 27	0
1	B	759/847 (89%)	0.31	42 (5%) 25 21	27, 27, 27, 27	0
All	All	1517/1694 (89%)	0.28	79 (5%) 27 23	27, 27, 27, 27	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	969	GLY	5.9
1	B	498	LYS	5.8
1	B	549	THR	5.6
1	B	497	ARG	5.2
1	B	496	MET	5.0

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.