



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2016 – 05:00 AM EDT

PDB ID : 5D8E  
Title : crystal structure of SSB from homo sapiens  
Authors : Li, Y.H.; Gao, Z.Q.; Dong, Y.H.  
Deposited on : 2015-08-17  
Resolution : 3.00 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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
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)	:	rb-20027939

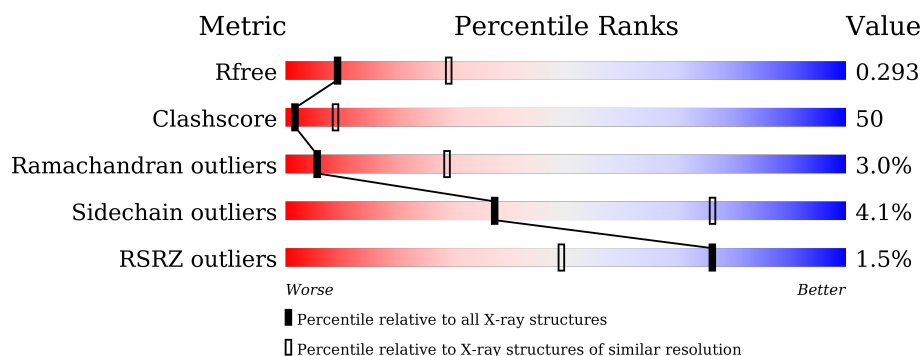
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	115	<div> <div>0%</div> <div> <div>43%</div> <div>44%</div> <div>8%</div> </div> </div>
1	B	115	<div> <div>3%</div> <div> <div>32%</div> <div>50%</div> <div>6%</div> <div>10%</div> </div> </div>
1	C	115	<div> <div>44%</div> <div>43%</div> <div>9%</div> </div>
1	D	115	<div> <div>2%</div> <div> <div>30%</div> <div>55%</div> <div>10%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3235 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 SOSS complex subunit B1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	Se	0	0	0
			821	524	138	154	3	2			
1	B	103	Total	C	N	O	S	Se	0	0	0
			800	510	134	151	3	2			
1	C	105	Total	C	N	O	S	Se	0	0	0
			814	520	137	152	3	2			
1	D	103	Total	C	N	O	S	Se	0	0	0
			800	510	134	151	3	2			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	MSE	LEU	engineered mutation	UNP Q9BQ15
A	110	HIS	-	expression tag	UNP Q9BQ15
A	111	HIS	-	expression tag	UNP Q9BQ15
A	112	HIS	-	expression tag	UNP Q9BQ15
A	113	HIS	-	expression tag	UNP Q9BQ15
A	114	HIS	-	expression tag	UNP Q9BQ15
A	115	HIS	-	expression tag	UNP Q9BQ15
B	70	MSE	LEU	engineered mutation	UNP Q9BQ15
B	110	HIS	-	expression tag	UNP Q9BQ15
B	111	HIS	-	expression tag	UNP Q9BQ15
B	112	HIS	-	expression tag	UNP Q9BQ15
B	113	HIS	-	expression tag	UNP Q9BQ15
B	114	HIS	-	expression tag	UNP Q9BQ15
B	115	HIS	-	expression tag	UNP Q9BQ15
C	70	MSE	LEU	engineered mutation	UNP Q9BQ15
C	110	HIS	-	expression tag	UNP Q9BQ15
C	111	HIS	-	expression tag	UNP Q9BQ15
C	112	HIS	-	expression tag	UNP Q9BQ15
C	113	HIS	-	expression tag	UNP Q9BQ15
C	114	HIS	-	expression tag	UNP Q9BQ15
C	115	HIS	-	expression tag	UNP Q9BQ15

*Continued on next page...*

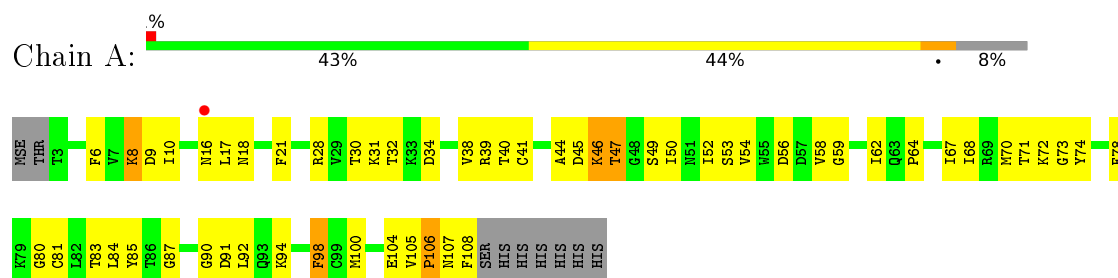
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	70	MSE	LEU	engineered mutation	UNP Q9BQ15
D	110	HIS	-	expression tag	UNP Q9BQ15
D	111	HIS	-	expression tag	UNP Q9BQ15
D	112	HIS	-	expression tag	UNP Q9BQ15
D	113	HIS	-	expression tag	UNP Q9BQ15
D	114	HIS	-	expression tag	UNP Q9BQ15
D	115	HIS	-	expression tag	UNP Q9BQ15

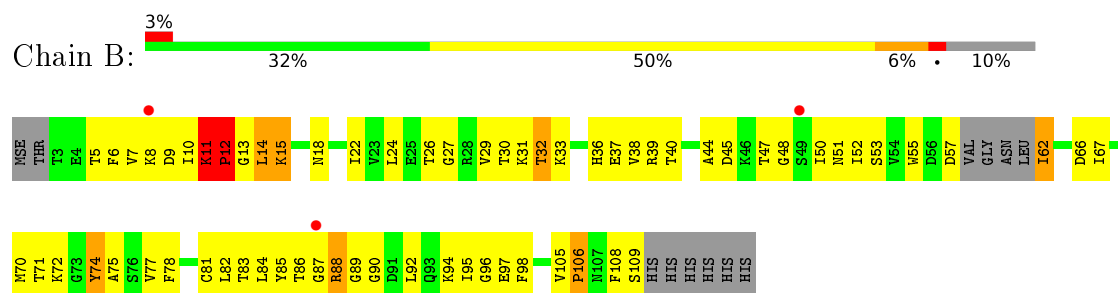
### 3 Residue-property plots [i](#)

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

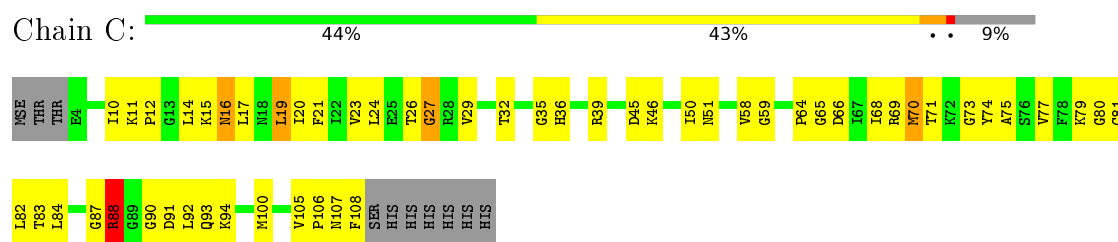
- Molecule 1: SOSS complex subunit B1



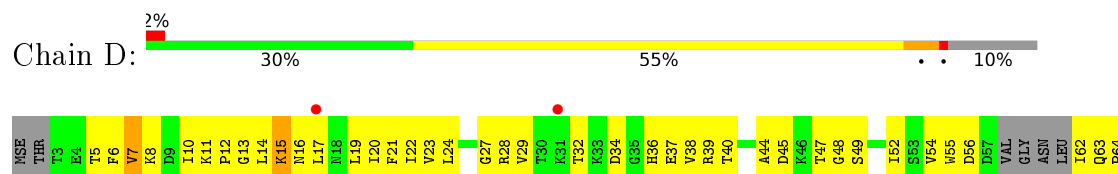
- Molecule 1: SOSS complex subunit B1



- Molecule 1: SOSS complex subunit B1



- Molecule 1: SOSS complex subunit B1



665	666	167	168	669	670	Y74	A75	S76	V77	F78	A79	G80	C81	L82	T83	L84	Y85	T86	G87	R88	G89	G90	G91	L92	G93	G94	I95	F98	C99	M100	E104	V105	P106	S109	HIS	HIS	HIS	HIS	HIS	HIS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	137.69 Å 137.69 Å 83.82 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.91 – 3.00 35.91 – 2.99	Depositor EDS
% Data completeness (in resolution range)	97.0 (35.91-3.00) 96.9 (35.91-2.99)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.04 (at 3.00 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.264 , 0.289 0.276 , 0.293	Depositor DCC
$R_{free}$ test set	925 reflections (5.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	75.0	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 70.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.38$	Xtriage
Estimated twinning fraction	0.450 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3235	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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

### 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	0.70	0/832	0.80	1/1118 (0.1%)
1	B	0.99	2/810 (0.2%)	0.94	2/1086 (0.2%)
1	C	0.81	0/825	0.89	2/1108 (0.2%)
1	D	0.68	0/810	0.88	2/1086 (0.2%)
All	All	0.80	2/3277 (0.1%)	0.88	7/4398 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	D	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	106	PRO	N-CD	-18.94	1.21	1.47
1	B	12	PRO	N-CD	5.61	1.55	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	87	GLY	N-CA-C	7.26	131.25	113.10
1	C	27	GLY	N-CA-C	7.21	131.13	113.10
1	C	87	GLY	N-CA-C	6.79	130.08	113.10
1	B	106	PRO	N-CD-CG	6.20	112.50	103.20
1	B	11	LYS	N-CA-C	5.59	126.10	111.00

There are no chirality outliers.



All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	11	LYS	Peptide
1	D	78	PHE	Peptide

## 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	821	0	841	49	2
1	B	800	0	814	132	1
1	C	814	0	834	61	0
1	D	800	0	816	92	1
All	All	3235	0	3305	326	2

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 50.

The worst 5 of 326 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:B:78:PHE:CZ	1:B:85:TYR:CE1	1.95	1.55
1:B:78:PHE:CZ	1:B:85:TYR:CZ	1.95	1.54
1:B:78:PHE:CE1	1:B:85:TYR:OH	1.66	1.46
1:D:16:ASN:HB2	1:D:88:ARG:NH2	1.47	1.26
1:B:72:LYS:NZ	1:B:89:GLY:O	1.78	1.16

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	Interatomic distance (Å)	Clash overlap (Å)
1:A:80:GLY:O	1:D:74:TYR:OH[4_545]	2.05	0.15
1:A:34:ASP:OD1	1:B:33:LYS:NZ[3_654]	2.16	0.04

## 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	104/115 (90%)	93 (89%)	8 (8%)	3 (3%)	6	29
1	B	99/115 (86%)	79 (80%)	17 (17%)	3 (3%)	5	29
1	C	103/115 (90%)	90 (87%)	10 (10%)	3 (3%)	6	29
1	D	99/115 (86%)	84 (85%)	12 (12%)	3 (3%)	5	29
All	All	405/460 (88%)	346 (85%)	47 (12%)	12 (3%)	5	29

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	15	LYS
1	D	7	VAL
1	A	47	THR
1	B	11	LYS
1	B	12	PRO

### 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	92/98 (94%)	89 (97%)	3 (3%)	45	82
1	B	90/98 (92%)	84 (93%)	6 (7%)	20	57
1	C	91/98 (93%)	87 (96%)	4 (4%)	35	74
1	D	90/98 (92%)	88 (98%)	2 (2%)	60	88
All	All	363/392 (93%)	348 (96%)	15 (4%)	37	76

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	62	ILE
1	B	74	TYR
1	C	88	ARG
1	B	32	THR
1	C	70	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	63	GLN
1	B	107	ASN
1	D	16	ASN
1	D	36	HIS

### 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	104/115 (90%)	0.24	1 (0%) 84 60	49, 64, 90, 111	0
1	B	101/115 (87%)	0.40	3 (2%) 54 25	25, 91, 118, 126	0
1	C	103/115 (89%)	0.24	0 100 100	49, 65, 82, 112	0
1	D	101/115 (87%)	0.43	2 (1%) 68 39	55, 91, 118, 137	0
All	All	409/460 (88%)	0.33	6 (1%) 76 49	25, 74, 112, 137	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	31	LYS	3.5
1	B	87	GLY	3.3
1	D	17	LEU	3.0
1	B	49	SER	3.0
1	A	16	ASN	2.5

### 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.