



# wwPDB X-ray Structure Validation Summary Report

Jun 17, 2024 – 01:14 AM EDT


PDB ID : 3D5N  
Title : Crystal structure of the Q97W15\_SULSO protein from Sulfolobus solfataricus. NESG target SsR125.  
Authors : Vorobiev, S.M.; Chen, Y.; Seetharaman, J.; Lee, D.; Foote, R.E.; Maglaqui, M.; Janjua, H.; Xiao, R.; Acton, T.B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)  
Deposited on : 2008-05-16  
Resolution : 2.80 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

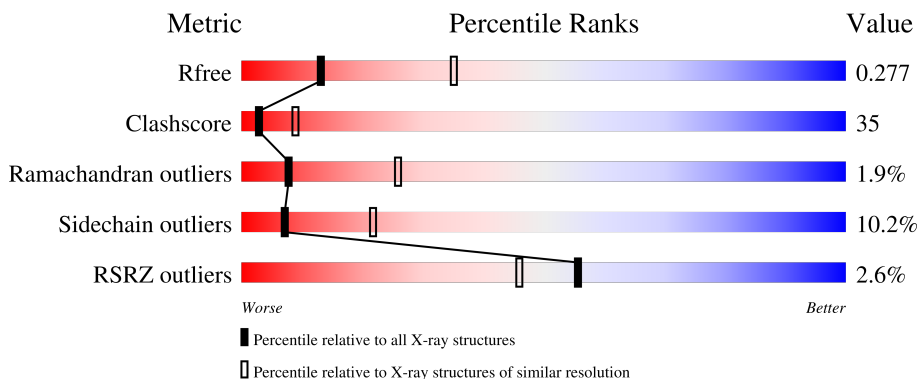
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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.



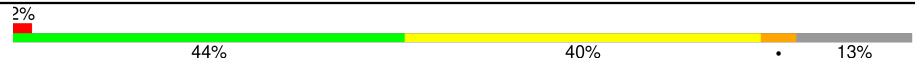

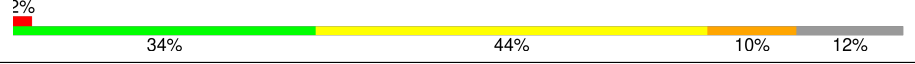
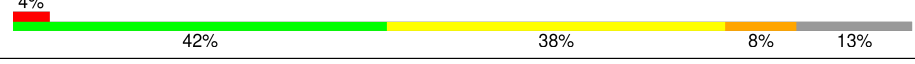
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of 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	197	
1	B	197	
1	D	197	
1	E	197	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	197	
1	G	197	
1	H	197	
1	I	197	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 10496 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 Q97W15\_SULSO.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	178	1356	882	223	243	3	5	0	0	0
1	B	167	1268	831	201	229	3	4	0	0	0
1	D	178	1373	894	226	245	3	5	0	0	0
1	E	172	1300	848	208	236	3	5	0	0	0
1	F	172	1292	847	214	223	3	5	0	0	0
1	G	166	1269	829	207	225	3	5	0	0	0
1	H	173	1315	855	212	241	3	4	0	0	0
1	I	171	1290	842	206	235	3	4	0	0	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	190	LEU	-	expression tag	UNP Q97W15
A	191	GLU	-	expression tag	UNP Q97W15
A	192	HIS	-	expression tag	UNP Q97W15
A	193	HIS	-	expression tag	UNP Q97W15
A	194	HIS	-	expression tag	UNP Q97W15
A	195	HIS	-	expression tag	UNP Q97W15
A	196	HIS	-	expression tag	UNP Q97W15
A	197	HIS	-	expression tag	UNP Q97W15
B	190	LEU	-	expression tag	UNP Q97W15
B	191	GLU	-	expression tag	UNP Q97W15
B	192	HIS	-	expression tag	UNP Q97W15
B	193	HIS	-	expression tag	UNP Q97W15
B	194	HIS	-	expression tag	UNP Q97W15

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	HIS	-	expression tag	UNP Q97W15
B	196	HIS	-	expression tag	UNP Q97W15
B	197	HIS	-	expression tag	UNP Q97W15
D	190	LEU	-	expression tag	UNP Q97W15
D	191	GLU	-	expression tag	UNP Q97W15
D	192	HIS	-	expression tag	UNP Q97W15
D	193	HIS	-	expression tag	UNP Q97W15
D	194	HIS	-	expression tag	UNP Q97W15
D	195	HIS	-	expression tag	UNP Q97W15
D	196	HIS	-	expression tag	UNP Q97W15
D	197	HIS	-	expression tag	UNP Q97W15
E	190	LEU	-	expression tag	UNP Q97W15
E	191	GLU	-	expression tag	UNP Q97W15
E	192	HIS	-	expression tag	UNP Q97W15
E	193	HIS	-	expression tag	UNP Q97W15
E	194	HIS	-	expression tag	UNP Q97W15
E	195	HIS	-	expression tag	UNP Q97W15
E	196	HIS	-	expression tag	UNP Q97W15
E	197	HIS	-	expression tag	UNP Q97W15
F	190	LEU	-	expression tag	UNP Q97W15
F	191	GLU	-	expression tag	UNP Q97W15
F	192	HIS	-	expression tag	UNP Q97W15
F	193	HIS	-	expression tag	UNP Q97W15
F	194	HIS	-	expression tag	UNP Q97W15
F	195	HIS	-	expression tag	UNP Q97W15
F	196	HIS	-	expression tag	UNP Q97W15
F	197	HIS	-	expression tag	UNP Q97W15
G	190	LEU	-	expression tag	UNP Q97W15
G	191	GLU	-	expression tag	UNP Q97W15
G	192	HIS	-	expression tag	UNP Q97W15
G	193	HIS	-	expression tag	UNP Q97W15
G	194	HIS	-	expression tag	UNP Q97W15
G	195	HIS	-	expression tag	UNP Q97W15
G	196	HIS	-	expression tag	UNP Q97W15
G	197	HIS	-	expression tag	UNP Q97W15
H	190	LEU	-	expression tag	UNP Q97W15
H	191	GLU	-	expression tag	UNP Q97W15
H	192	HIS	-	expression tag	UNP Q97W15
H	193	HIS	-	expression tag	UNP Q97W15
H	194	HIS	-	expression tag	UNP Q97W15
H	195	HIS	-	expression tag	UNP Q97W15
H	196	HIS	-	expression tag	UNP Q97W15

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	197	HIS	-	expression tag	UNP Q97W15
I	190	LEU	-	expression tag	UNP Q97W15
I	191	GLU	-	expression tag	UNP Q97W15
I	192	HIS	-	expression tag	UNP Q97W15
I	193	HIS	-	expression tag	UNP Q97W15
I	194	HIS	-	expression tag	UNP Q97W15
I	195	HIS	-	expression tag	UNP Q97W15
I	196	HIS	-	expression tag	UNP Q97W15
I	197	HIS	-	expression tag	UNP Q97W15

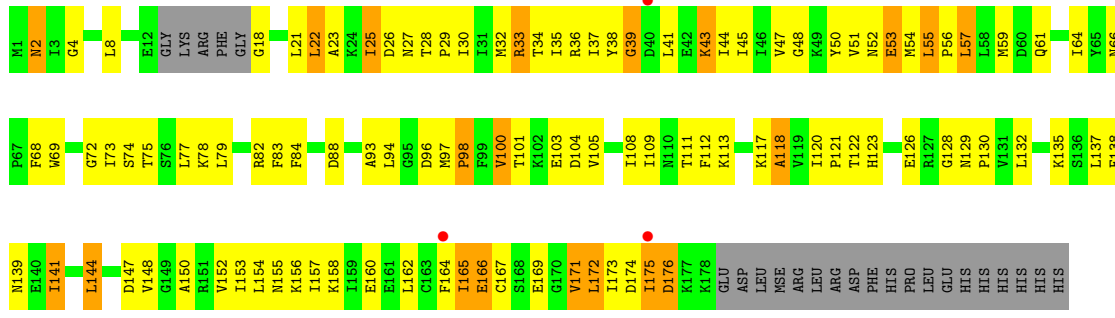
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total O 3 3	0	0
2	B	4	Total O 4 4	0	0
2	D	6	Total O 6 6	0	0
2	E	4	Total O 4 4	0	0
2	F	4	Total O 4 4	0	0
2	G	2	Total O 2 2	0	0
2	H	5	Total O 5 5	0	0
2	I	5	Total O 5 5	0	0

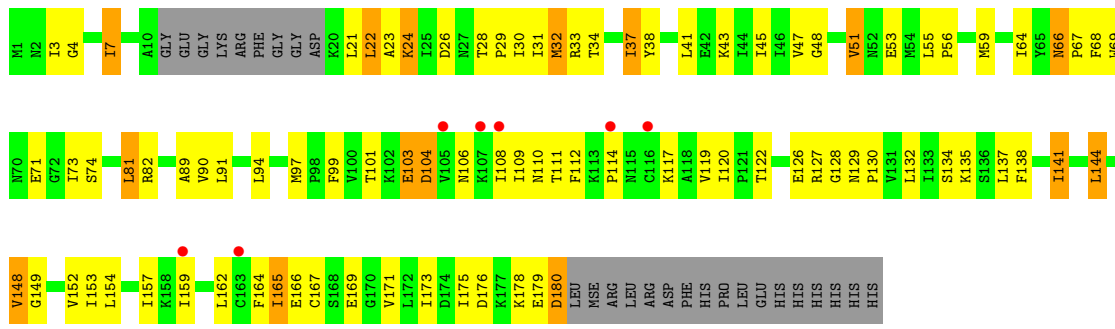








• Molecule 1: Q97W15\_SULSO



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.96Å 169.23Å 173.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.47 – 2.80 48.47 – 2.77	Depositor EDS
% Data completeness (in resolution range)	86.2 (48.47-2.80) 96.7 (48.47-2.77)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.77Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.231 , 0.266 0.244 , 0.277	Depositor DCC
$R_{free}$ test set	6108 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	74.0	Xtrriage
Anisotropy	0.088	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 71.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.034 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	10496	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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	0.41	0/1373	0.70	0/1852
1	B	0.41	0/1283	0.71	1/1734 (0.1%)
1	D	0.40	0/1390	0.69	0/1870
1	E	0.45	0/1315	0.70	1/1778 (0.1%)
1	F	0.41	0/1307	0.68	0/1765
1	G	0.41	0/1284	0.71	0/1736
1	H	0.48	0/1330	0.70	0/1797
1	I	0.38	0/1305	0.68	0/1767
All	All	0.42	0/10587	0.70	2/14299 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	132	LEU	CA-CB-CG	5.60	128.19	115.30
1	B	144	LEU	CA-CB-CG	5.57	128.11	115.30

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	1356	0	1391	78	0
1	B	1268	0	1300	76	0
1	D	1373	0	1430	91	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1300	0	1324	105	0
1	F	1292	0	1335	96	0
1	G	1269	0	1309	98	0
1	H	1315	0	1339	122	0
1	I	1290	0	1305	110	0
2	A	3	0	0	0	0
2	B	4	0	0	1	0
2	D	6	0	0	0	0
2	E	4	0	0	0	0
2	F	4	0	0	0	0
2	G	2	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0
All	All	10496	0	10733	740	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 35.

The worst 5 of 740 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:G:111:THR:HG21	1:G:165:ILE:HG21	1.20	1.10
1:A:18:GLY:HA3	1:A:50:TYR:HB2	1.31	1.09
1:H:66:ASN:HD21	1:H:75:THR:HG23	1.17	1.09
1:E:68:PHE:HB2	1:E:75:THR:HG21	1.34	1.08
1:G:81:LEU:HD22	1:G:141:ILE:HD11	1.37	1.06

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	176/197 (89%)	155 (88%)	18 (10%)	3 (2%)	9	29
1	B	163/197 (83%)	141 (86%)	19 (12%)	3 (2%)	8	28
1	D	176/197 (89%)	157 (89%)	18 (10%)	1 (1%)	25	56
1	E	168/197 (85%)	150 (89%)	14 (8%)	4 (2%)	6	20
1	F	168/197 (85%)	152 (90%)	16 (10%)	0	100	100
1	G	162/197 (82%)	142 (88%)	15 (9%)	5 (3%)	4	14
1	H	169/197 (86%)	149 (88%)	14 (8%)	6 (4%)	3	11
1	I	167/197 (85%)	147 (88%)	17 (10%)	3 (2%)	8	28
All	All	1349/1576 (86%)	1193 (88%)	131 (10%)	25 (2%)	8	26

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	85	LYS
1	B	171	VAL
1	D	159	ILE
1	G	49	LYS
1	H	155	ASN

### 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	144/169 (85%)	128 (89%)	16 (11%)	6	19
1	B	136/169 (80%)	120 (88%)	16 (12%)	5	16
1	D	148/169 (88%)	132 (89%)	16 (11%)	6	19
1	E	139/169 (82%)	123 (88%)	16 (12%)	5	17
1	F	135/169 (80%)	125 (93%)	10 (7%)	13	37
1	G	137/169 (81%)	128 (93%)	9 (7%)	16	44
1	H	141/169 (83%)	125 (89%)	16 (11%)	6	18
1	I	137/169 (81%)	122 (89%)	15 (11%)	6	19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1117/1352 (83%)	1003 (90%)	114 (10%)	<b>7</b> <b>22</b>

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	132	LEU
1	I	144	LEU
1	F	141	ILE
1	I	141	ILE
1	I	7	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	66	ASN
1	F	52	ASN
1	E	155	ASN
1	G	52	ASN
1	B	66	ASN

### 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 monosaccharides 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	173/197 (87%)	0.19	5 (2%) 51 41	42, 66, 92, 98	0
1	B	163/197 (82%)	0.35	6 (3%) 41 31	39, 70, 99, 100	0
1	D	173/197 (87%)	0.26	3 (1%) 70 63	45, 69, 94, 100	0
1	E	167/197 (84%)	0.28	4 (2%) 59 49	28, 69, 89, 93	0
1	F	167/197 (84%)	0.26	4 (2%) 59 49	34, 74, 99, 100	0
1	G	161/197 (81%)	0.22	3 (1%) 66 59	45, 70, 90, 96	0
1	H	168/197 (85%)	0.21	3 (1%) 68 61	33, 68, 93, 95	0
1	I	166/197 (84%)	0.39	7 (4%) 36 26	44, 77, 100, 100	0
All	All	1338/1576 (84%)	0.27	35 (2%) 56 46	28, 70, 96, 100	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	175	ILE	5.2
1	A	18	GLY	4.7
1	A	17	GLY	4.6
1	I	108	ILE	4.1
1	H	175	ILE	3.6

### 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 monosaccharides in this entry.



## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.