



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:23 PM GMT

PDB ID : 4RSJ  
Title : Pyrococcus furiosus Smc hinge domain with an extended coiled coil  
Authors : Soh, Y.M.; Shin, H.C.; Oh, B.H.  
Deposited on : 2014-11-08  
Resolution : 3.50 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

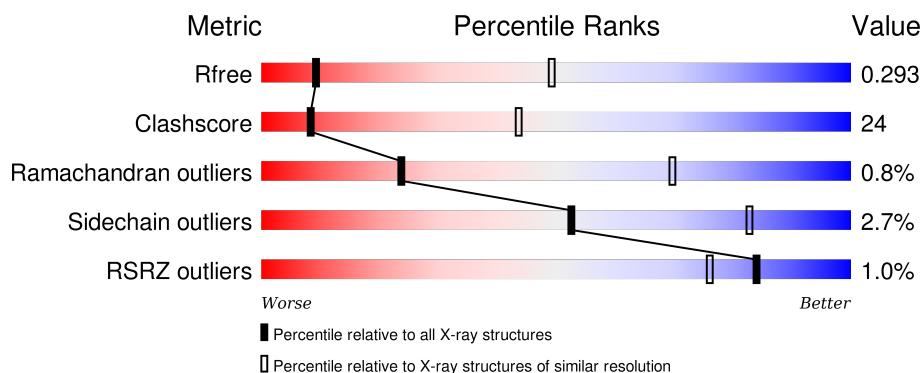
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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	276	 60% 37% ..
1	B	276	 53% 41% . .
1	C	276	 62% 33% . .
1	D	276	 58% 36% • 5%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7715 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.

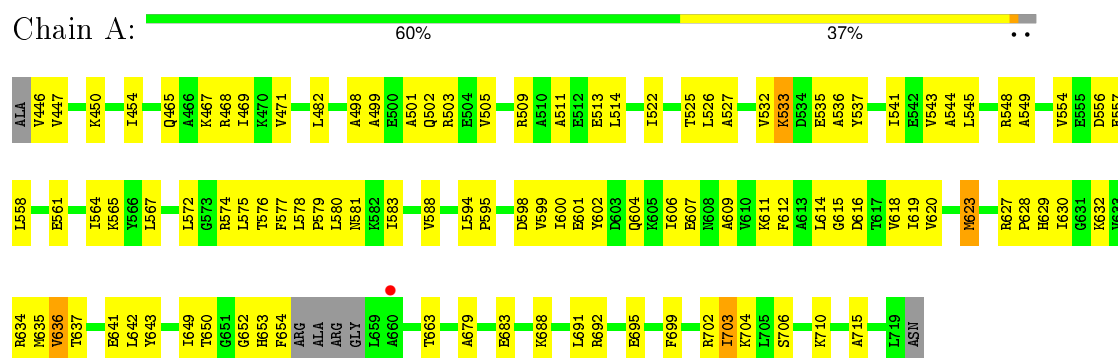
- Molecule 1 is a protein called Chromosome partition protein Smc.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	30	0	0
			1954	1217	346	389	2			
1	B	266	Total	C	N	O	S	36	0	0
			1994	1246	360	386	2			
1	C	265	Total	C	N	O	S	32	0	0
			1908	1187	336	383	2			
1	D	263	Total	C	N	O	S	36	0	0
			1859	1155	340	362	2			

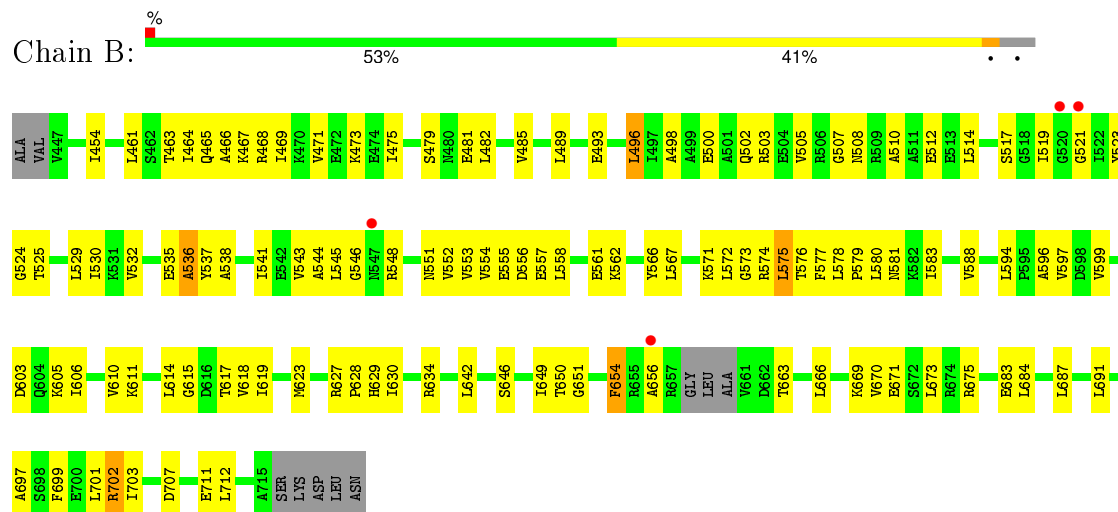
### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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.

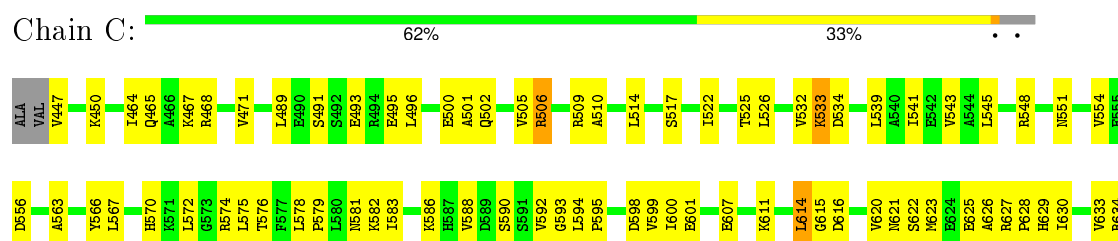
#### • Molecule 1: Chromosome partition protein Smc

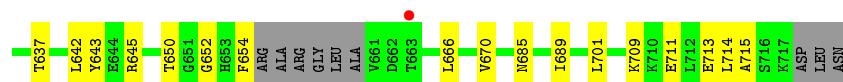


#### • Molecule 1: Chromosome partition protein Smc

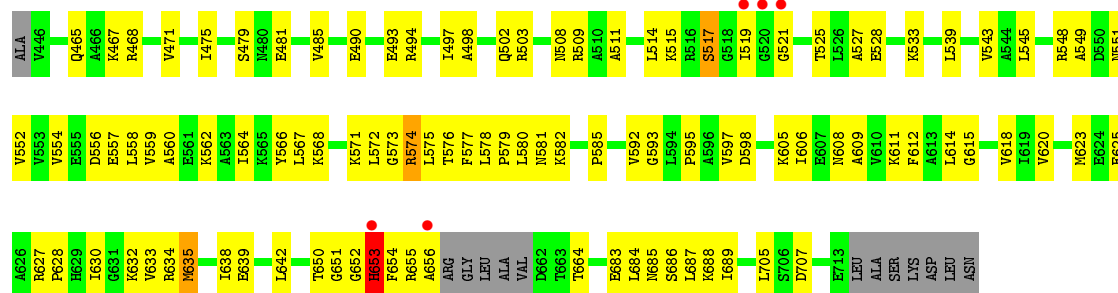


#### • Molecule 1: Chromosome partition protein Smc





# • Molecule 1: Chromosome partition protein Smc



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.92Å 116.88Å 145.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.50 37.63 – 2.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.50) 73.5 (37.63-2.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.75 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.236 , 0.284 0.251 , 0.293	Depositor DCC
$R_{free}$ test set	928 reflections (4.75%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.2	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 65.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	2 of 26379 reflections (0.008%)	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	7715	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.69 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.7565e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 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	0.24	0/1969	0.46	0/2661
1	B	0.24	0/2009	0.48	0/2705
1	C	0.24	0/1922	0.48	0/2598
1	D	0.43	2/1873 (0.1%)	0.66	4/2535 (0.2%)
All	All	0.30	2/7773 (0.0%)	0.52	4/10499 (0.0%)

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	D	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	653	HIS	C-N	11.06	1.59	1.34
1	D	517	SER	C-N	-10.31	1.14	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	653	HIS	O-C-N	-17.11	95.33	122.70
1	D	653	HIS	CA-C-N	12.02	143.65	117.20
1	D	653	HIS	C-N-CA	7.43	140.28	121.70
1	D	517	SER	O-C-N	-7.20	110.96	123.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	517	SER	Mainchain
1	D	653	HIS	Mainchain,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	1954	0	1861	103	0
1	B	1994	0	1990	116	0
1	C	1908	0	1818	81	0
1	D	1859	0	1746	85	0
All	All	7715	0	7415	362	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 24.

The worst 5 of 362 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:533:LYS:HG2	1:A:601:GLU:HG3	1.23	1.19
1:C:533:LYS:HE3	1:C:601:GLU:CG	1.78	1.12
1:A:699:PHE:HZ	1:B:702:ARG:HD2	1.19	1.07
1:A:465:GLN:HG2	1:A:468:ARG:HH11	1.19	1.05
1:C:500:GLU:HB3	1:C:509:ARG:HH21	1.22	1.04

There are no symmetry-related clashes.

## 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	266/276 (96%)	247 (93%)	17 (6%)	2 (1%)	24	70
1	B	262/276 (95%)	237 (90%)	23 (9%)	2 (1%)	24	70
1	C	261/276 (95%)	240 (92%)	21 (8%)	0	100	100
1	D	259/276 (94%)	230 (89%)	25 (10%)	4 (2%)	13	56
All	All	1048/1104 (95%)	954 (91%)	86 (8%)	8 (1%)	24	70

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	536	ALA
1	A	604	GLN
1	D	533	LYS
1	D	605	LYS
1	A	636	VAL

### 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	184/232 (79%)	178 (97%)	6 (3%)	45	79
1	B	198/232 (85%)	193 (98%)	5 (2%)	55	84
1	C	181/232 (78%)	178 (98%)	3 (2%)	68	89
1	D	167/232 (72%)	161 (96%)	6 (4%)	42	77
All	All	730/928 (79%)	710 (97%)	20 (3%)	52	83

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	654	PHE
1	B	702	ARG
1	D	623	MET
1	B	575	LEU
1	B	623	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	465	GLN
1	B	570	HIS
1	D	502	GLN
1	A	581	ASN
1	D	480	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 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	270/276 (97%)	-0.47	1 (0%) 93 90	44, 88, 132, 195	26 (9%)
1	B	266/276 (96%)	-0.40	4 (1%) 76 67	48, 82, 134, 157	26 (9%)
1	C	265/276 (96%)	-0.50	1 (0%) 93 90	37, 88, 120, 148	26 (9%)
1	D	263/276 (95%)	-0.51	5 (1%) 70 60	46, 83, 121, 184	26 (9%)
All	All	1064/1104 (96%)	-0.47	11 (1%) 84 76	37, 85, 127, 195	104 (9%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	520	GLY	4.9
1	B	520	GLY	4.7
1	D	521	GLY	3.8
1	B	656	ALA	3.7
1	C	663	THR	3.1

### 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.