



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:08 AM GMT

PDB ID : 2FSF
Title : Escherichia coli SecA, the preprotein translocase dimeric ATPase
Authors : Papanikolau, Y.; Petratos, K.; Economou, A.
Deposited on : 2006-01-23
Resolution : 2.00 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
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

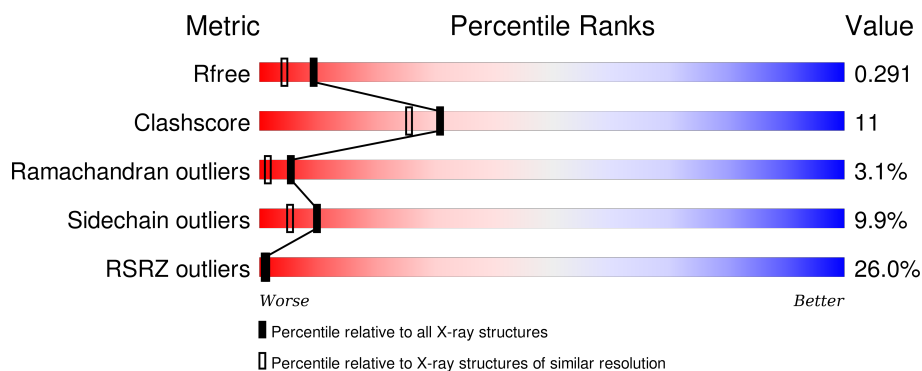
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 ≥ 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	853	<div> <div>15%</div> <div>61%</div> <div>15%</div> <div>• •</div> <div>20%</div> </div>
1	B	853	<div> <div>28%</div> <div>55%</div> <div>21%</div> <div>7%</div> <div>•</div> <div>15%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11637 atoms, of which 0 are hydrogens and 0 are deuteriums.

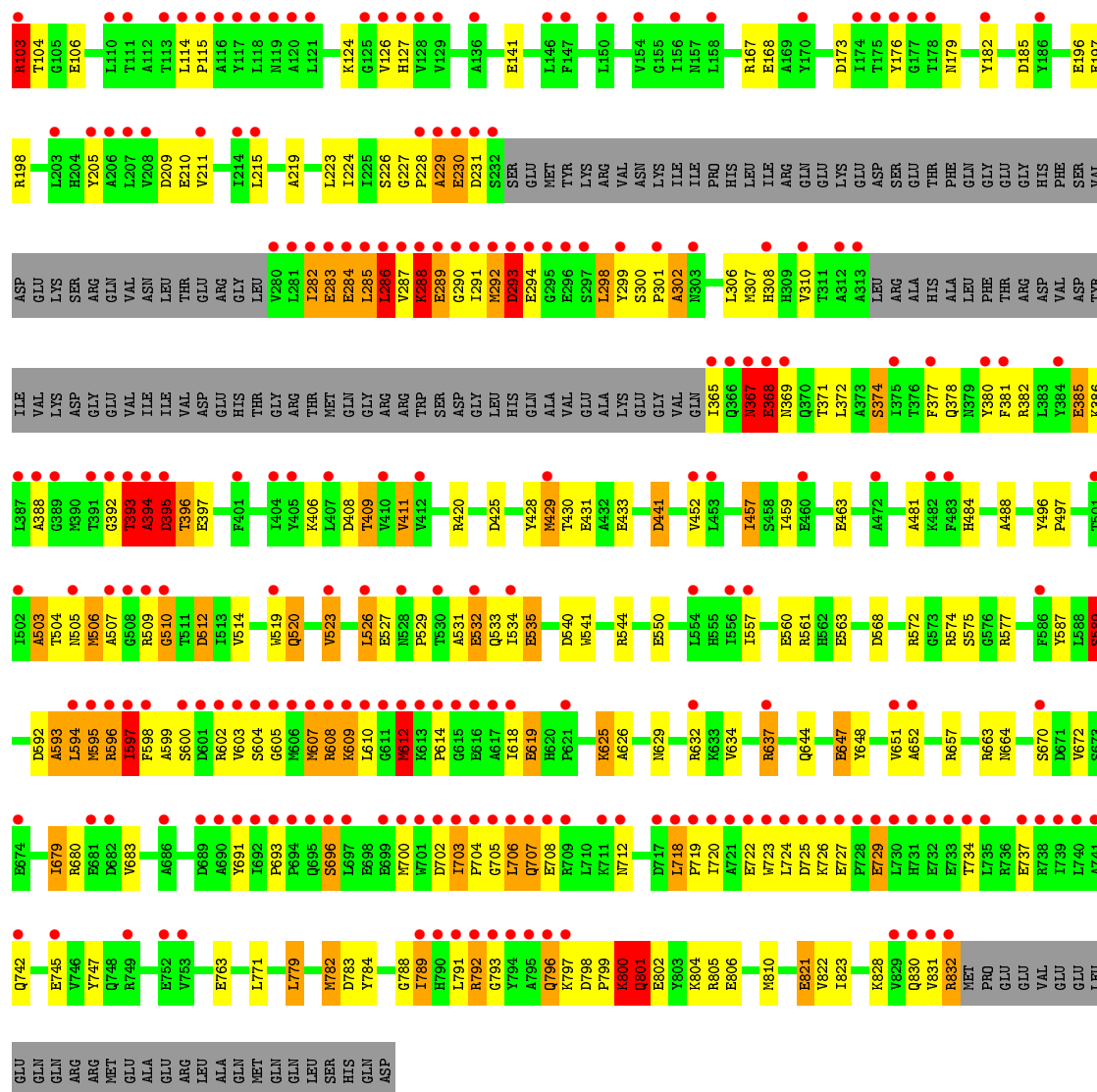
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Preprotein translocase secA subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	686	Total	C	N	O	S	0	0	0
			5462	3426	966	1044	26			
1	B	723	Total	C	N	O	S	0	0	0
			5741	3600	1009	1104	28			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	177	Total	O	0	0
			177	177		
2	B	257	Total	O	0	0
			257	257		



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	75.03Å 90.17Å 163.05Å 90.00° 100.48° 90.00°	Depositor
Resolution (Å)	19.98 – 2.00 19.98 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.1 (19.98-2.00) 97.1 (19.98-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.82 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.213 , 0.261 0.257 , 0.291	Depositor DCC
R_{free} test set	7014 reflections (5.01%)	DCC
Wilson B-factor (Å ²)	35.8	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 64.3	EDS
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 140049 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11637	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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	1.34	35/5552 (0.6%)	1.10	19/7491 (0.3%)
1	B	1.49	61/5833 (1.0%)	1.20	36/7870 (0.5%)
All	All	1.42	96/11385 (0.8%)	1.15	55/15361 (0.4%)

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	A	0	4
1	B	0	9
All	All	0	13

The worst 5 of 96 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	98	CYS	CB-SG	-13.49	1.59	1.82
1	A	698	GLU	CD-OE2	13.10	1.40	1.25
1	A	698	GLU	CD-OE1	10.85	1.37	1.25
1	A	560	GLU	CG-CD	9.58	1.66	1.51
1	A	697	LEU	C-N	9.56	1.56	1.34

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	167	ARG	NE-CZ-NH1	9.57	125.09	120.30
1	B	425	ASP	CB-CG-OD2	-9.12	110.09	118.30
1	B	425	ASP	CB-CG-OD1	9.05	126.45	118.30
1	A	72	ARG	NE-CZ-NH2	8.68	124.64	120.30
1	B	561	ARG	NE-CZ-NH1	8.38	124.49	120.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	ARG	Peptide
1	A	393	THR	Peptide
1	A	395	ASP	Peptide
1	A	697	LEU	Mainchain
1	B	229	ALA	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	5462	0	5460	90	0
1	B	5741	0	5731	155	1
2	A	177	0	0	12	0
2	B	257	0	0	30	0
All	All	11637	0	11191	245	1

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 11.

The worst 5 of 245 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:96:GLU:HG3	2:B:1033:HOH:O	1.46	1.13
1:A:799:PRO:HA	2:A:1026:HOH:O	1.50	1.08
1:B:800:LYS:O	1:B:801:GLN:HB2	1.61	1.01
1:B:103:ARG:HH11	1:B:103:ARG:HG3	1.26	0.97
1:A:799:PRO:C	1:A:800:LYS:O	1.95	0.93

All (1) 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:B:600:SER:OG	1:B:830:GLN:O[1_655]	1.97	0.23

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	682/853 (80%)	635 (93%)	31 (4%)	16 (2%)	8	3
1	B	717/853 (84%)	650 (91%)	39 (5%)	28 (4%)	4	1
All	All	1399/1706 (82%)	1285 (92%)	70 (5%)	44 (3%)	5	1

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	394	ALA
1	A	396	THR
1	A	507	ALA
1	A	599	ALA
1	A	614	PRO

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	582/728 (80%)	532 (91%)	50 (9%)	13	7
1	B	613/728 (84%)	545 (89%)	68 (11%)	8	4
All	All	1195/1456 (82%)	1077 (90%)	118 (10%)	10	5

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	37	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	293	ASP
1	B	779	LEU
1	B	38	LEU
1	B	282	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	528	ASN
1	B	533	GLN
1	B	664	ASN
1	B	14	ASN
1	B	520	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	686/853 (80%)	1.07	130 (18%) 2 2	40, 53, 84, 102	0
1	B	723/853 (84%)	1.68	237 (32%) 0 1	38, 54, 87, 102	0
All	All	1409/1706 (82%)	1.39	367 (26%) 1 1	38, 53, 86, 102	0

The worst 5 of 367 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	791	LEU	12.9
1	B	728	PRO	11.4
1	B	789	ILE	10.2
1	B	365	ILE	9.9
1	B	615	GLY	9.5

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.