



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 03:06 am BST

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

<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

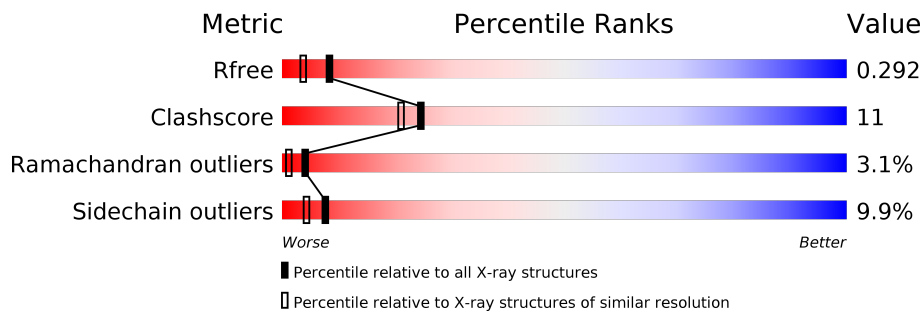
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (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 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	853	 61% 15% 7% 1% 20%
1	B	853	 55% 21% 7% 1% 15%

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
			Total	C	N	O	S			
1	A	686	5462	3426	966	1044	26	0	0	0
1	B	723	5741	3600	1009	1104	28	0	0	0

- Molecule 2 is water.

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

Q830	L718	L610	Q820	A388	TYR	GLY
V831	F719	G611	V523	G392	ILE	HIS
R832	I720	R612	V523	G392	VAL	PHE
MET	A721	R613	L526	T393	LYS	SER
PRO	E722	P614	L526	A394	ASP	VAL
GLU	W723	E723	E527	D395	GLY	ASP
GLU	L724	I618	R628	T396	GLU	VAL
VAL	D725	E619	F529	E397	VAL	LYS
GLU	K726	F530	F530	ILE	SER	SER
GLU	E727	R625	A531	K406	ILE	ARG
LEU	F728	A626	E532	L407	VAL	GLN
GLU	E729	I634	Q533	D408	VAL	GLN
GLN	T734	N629	I634	T409	ASP	VAL
ARG	E737	R632	E535	V410	GLU	ASN
MET	Q742	V634	D540	V411	THR	LEU
GLU	Q742	R637	W541	R420	THR	THR
ALA	E745	Q644	R544	D425	THR	GLY
GLU	V746	E647	E550	Y428	GLY	LEU
ARG	Y747	E648	I557	M429	GLY	LEU
LEU	E763	Y648	E560	T430	ARG	ARG
GLN	L771	V651	R561	E431	TRP	E284
GLN	L779	A652	E562	A432	SER	L285
SER	M782	R657	E563	E433	ASP	L286
HIS	D783	R663	D568	D441	GLY	V287
GLN	Y784	M664	R572	V452	GLY	K288
ASP	G788	S670	G573	I457	HIS	E289
	I789	D671	S488	S488	HIS	G290
	H790	V672	I459	I459	ALA	L291
	L791	I679	R577	E463	VAL	M292
	R792	R680	R577	A481	GLU	E294
	G793	V683	Y587	H481	LYS	L298
	Q796	Y691	L588	H484	GLY	Y299
	K797	I692	S589	A488	VAL	S300
	D798	P693	D592	A488	GLN	F301
	P799	P693	A593	Y486	ASP	A302
	K800	S696	L594	P497	LEU	L306
	Q801	M700	M595	A503	ARG	M307
	Y803	D702	R596	T504	ALA	H308
	K804	I703	F598	N505	ALA	B309
	R805	P704	A599	M506	LEU	V310
	E806	G705	A599	A507	LEU	A313
	M810	L706	S600	F377	ALA	ARG
	E821	Q707	D601	Q378	HIS	ALA
	V822	E708	R602	N379	ALA	ALA
	I823	M712	V603	Y380	LEU	LEU
	K828		S604	T511	PHE	PHE
	V829		G605	D512	THR	THR
			R606	E515	ARG	ARG
			M607	V514	ASP	ASP
			R608	W519	VAL	VAL
			K609	L387	ASP	ASP

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)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.213 , 0.261 0.260 , 0.292	Depositor DCC
R_{free} test set	7010 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	35.8	Xtrriage
Anisotropy	0.109	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 64.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtrriage
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

Xtrriage'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.*

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 [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	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 [\(i\)](#)

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	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 [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	682/853 (80%)	635 (93%)	31 (4%)	16 (2%)	6	2
1	B	717/853 (84%)	650 (91%)	39 (5%)	28 (4%)	3	1
All	All	1399/1706 (82%)	1285 (92%)	70 (5%)	44 (3%)	4	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 [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	582/728 (80%)	532 (91%)	50 (9%)	10	6
1	B	613/728 (84%)	545 (89%)	68 (11%)	6	3
All	All	1195/1456 (82%)	1077 (90%)	118 (10%)	8	4

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.