wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 10:43 AM EDT

PDB ID	:	2IPC
Title	:	Crystal structure of the translocation ATPase SecA from Thermus thermophilus reveals a parallel, head-to-head dimer
Authors	:	Vassylyev, D.G.; Mori, H.; Vassylyeva, M.N.; Tsukazaki, T.; Kimura, Y.; Tahirov, T.H.; Ito, K.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2006-10-12
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

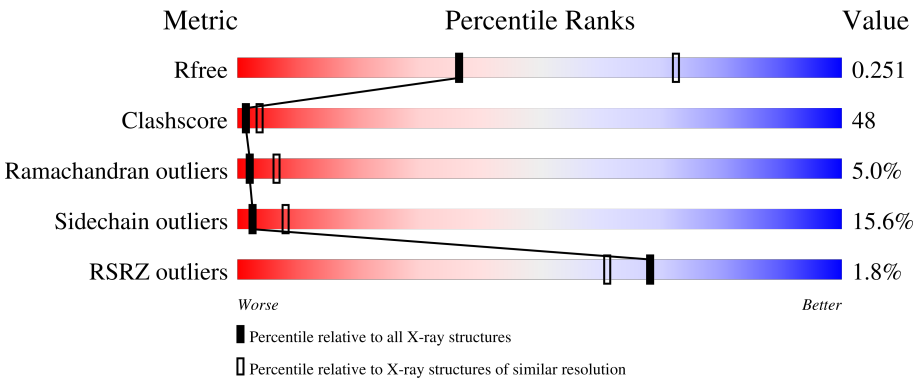
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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 >=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 <=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	997	2% 31% 51% 11% • 6%
1	B	997	2% 32% 48% 12% • 6%
1	C	997	2% 32% 50% 11% • 6%
1	D	997	% 29% 51% 12% • 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 31502 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	939	7551	4774	1354	1399	24	0	0	0
1	B	934	7515	4749	1349	1394	23	0	0	0
1	C	939	7551	4774	1354	1399	24	0	0	0
1	D	934	7515	4749	1349	1394	23	0	0	0

- Molecule 2 is water.

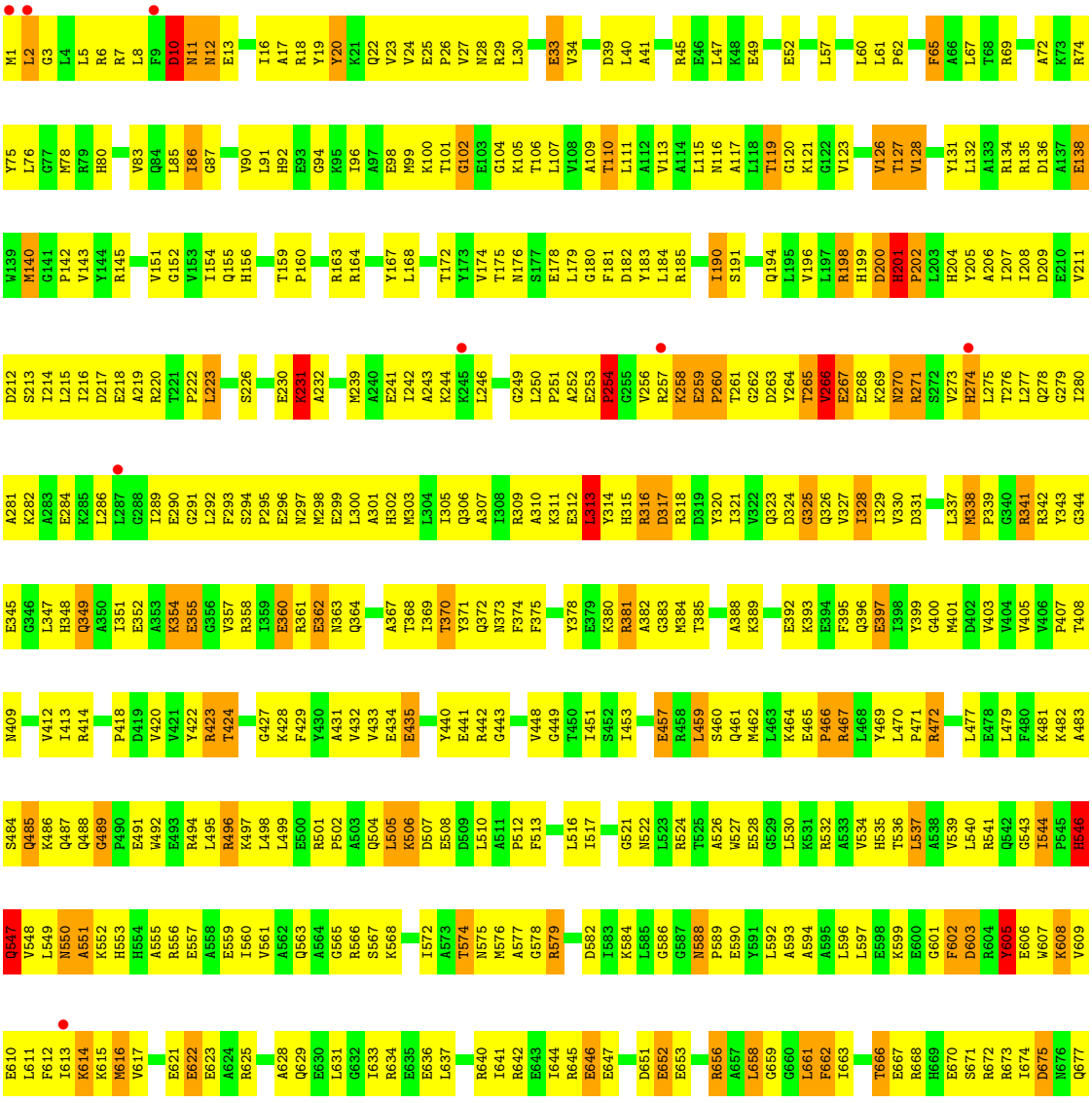
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	359	359	359	0	0
2	B	320	320	320	0	0
2	C	344	344	344	0	0
2	D	347	347	347	0	0

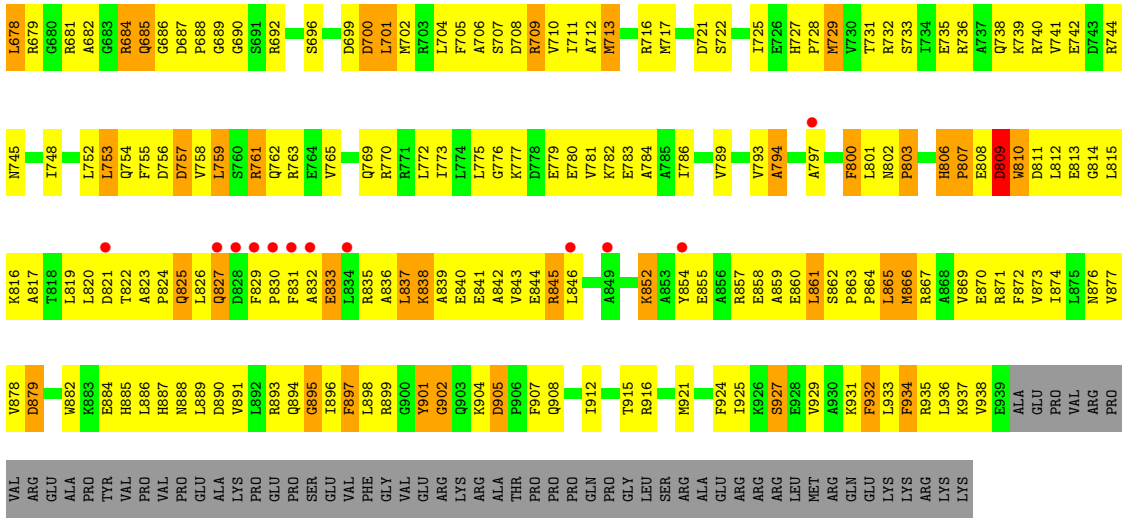
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: Preprotein translocase SecA subunit

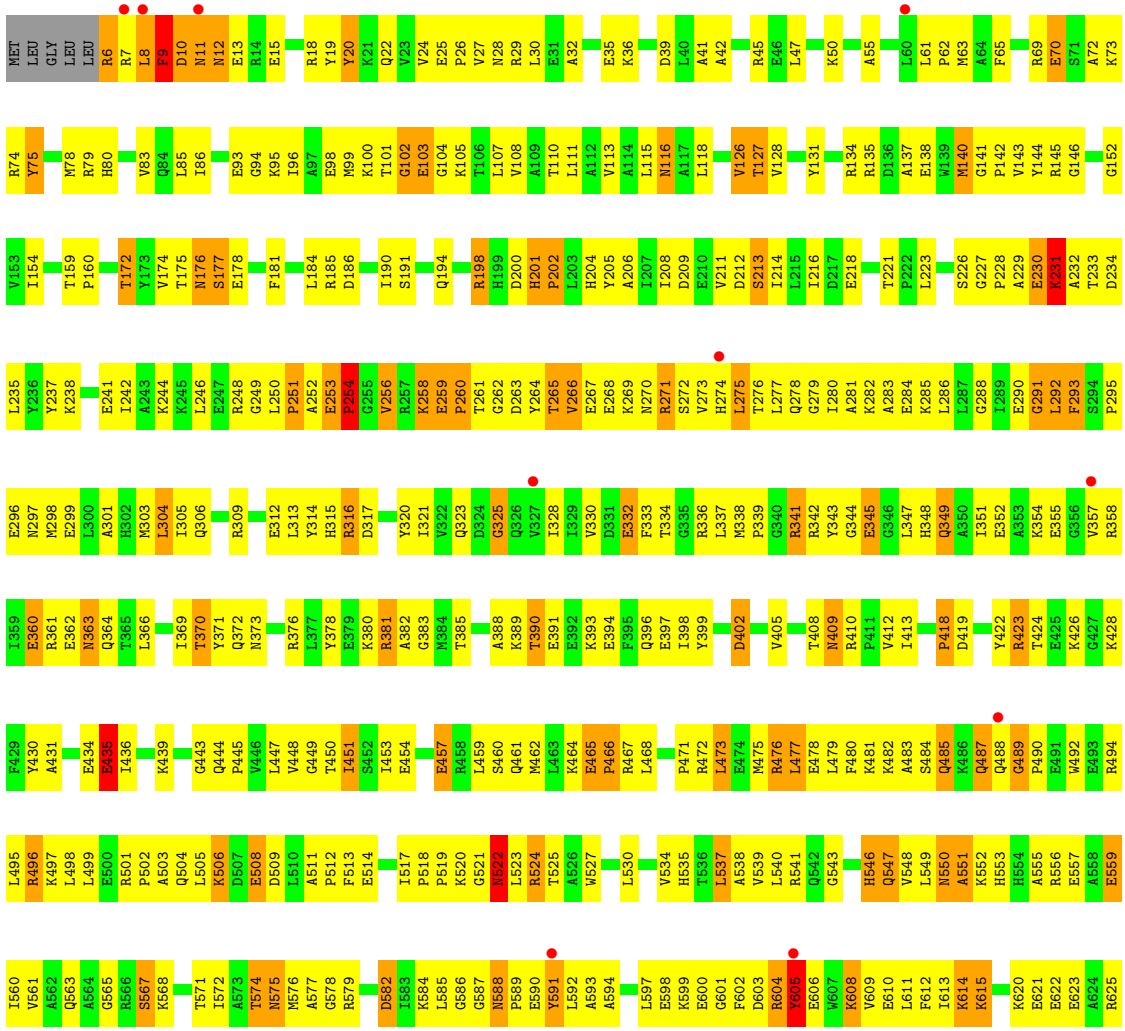
Chain A:

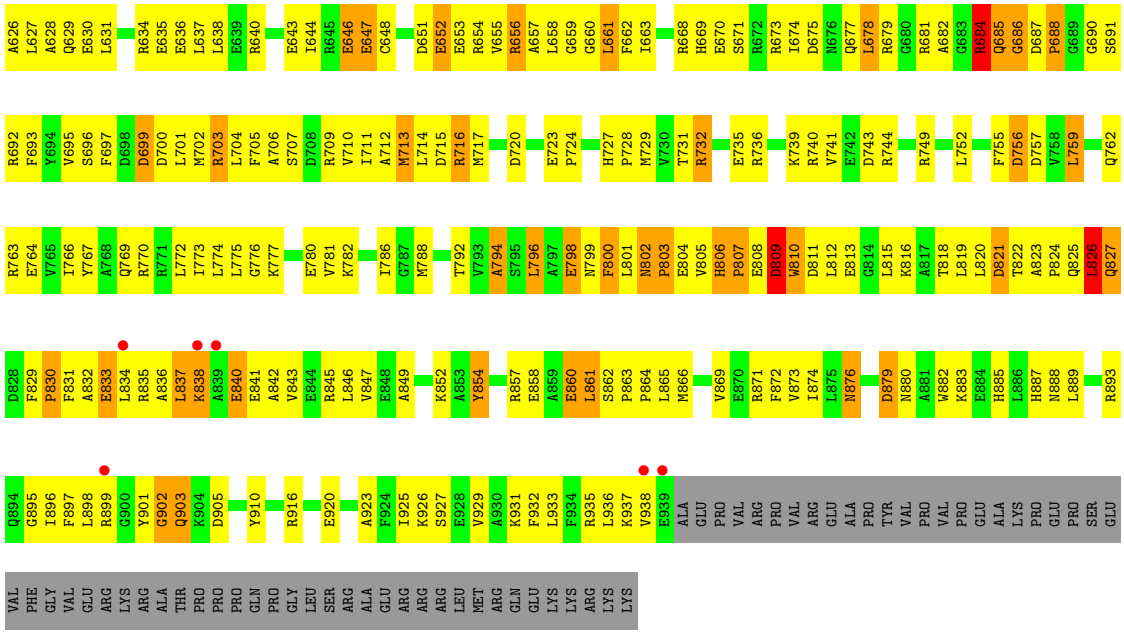




- Molecule 1: Preprotein translocase SecA subunit

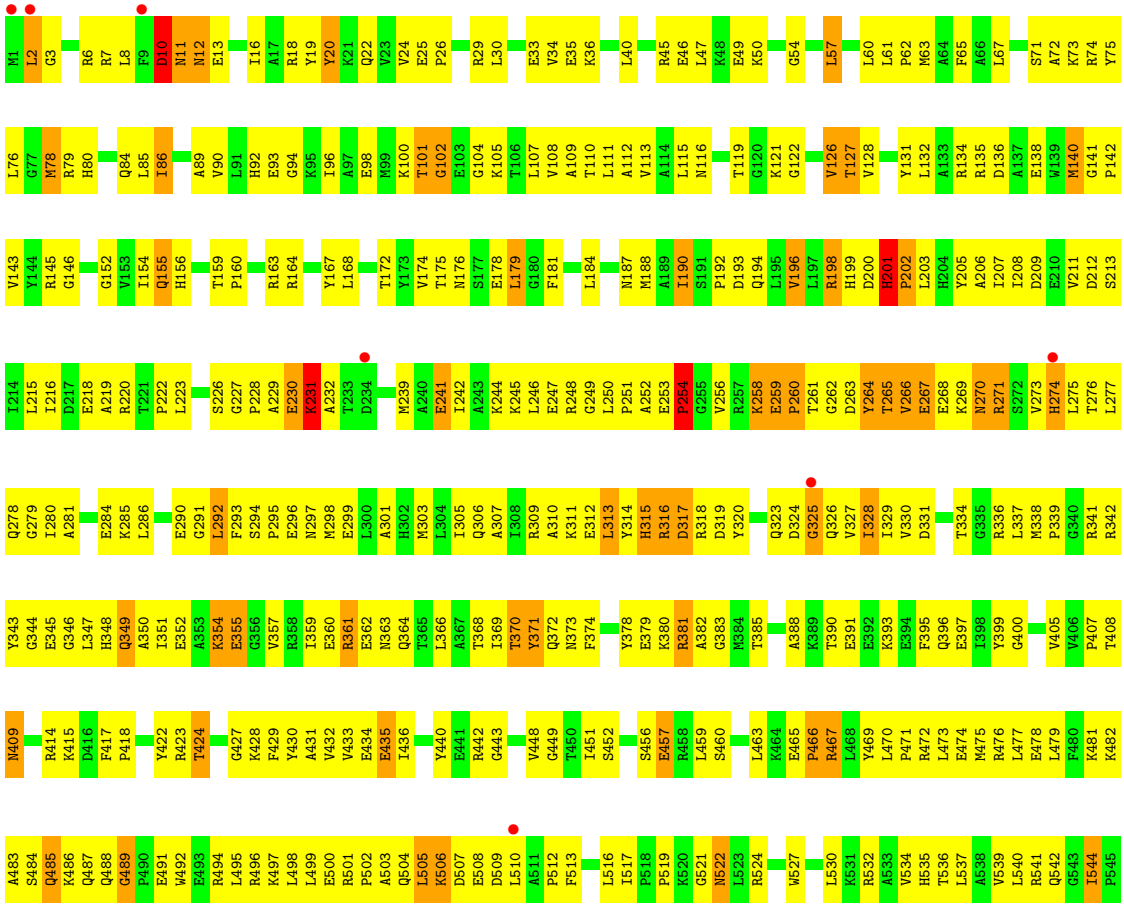
Chain B:

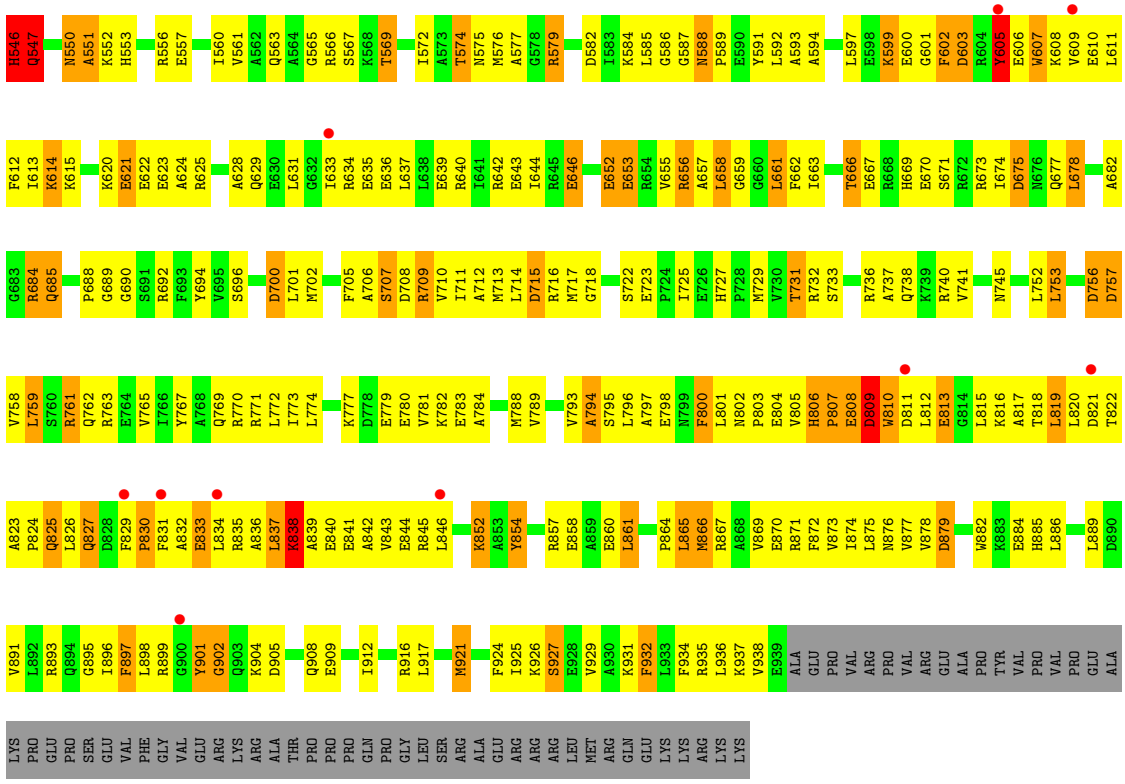




• Molecule 1: Preprotein translocase SecA subunit

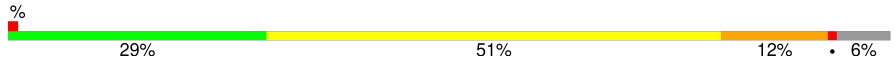
Chain C:

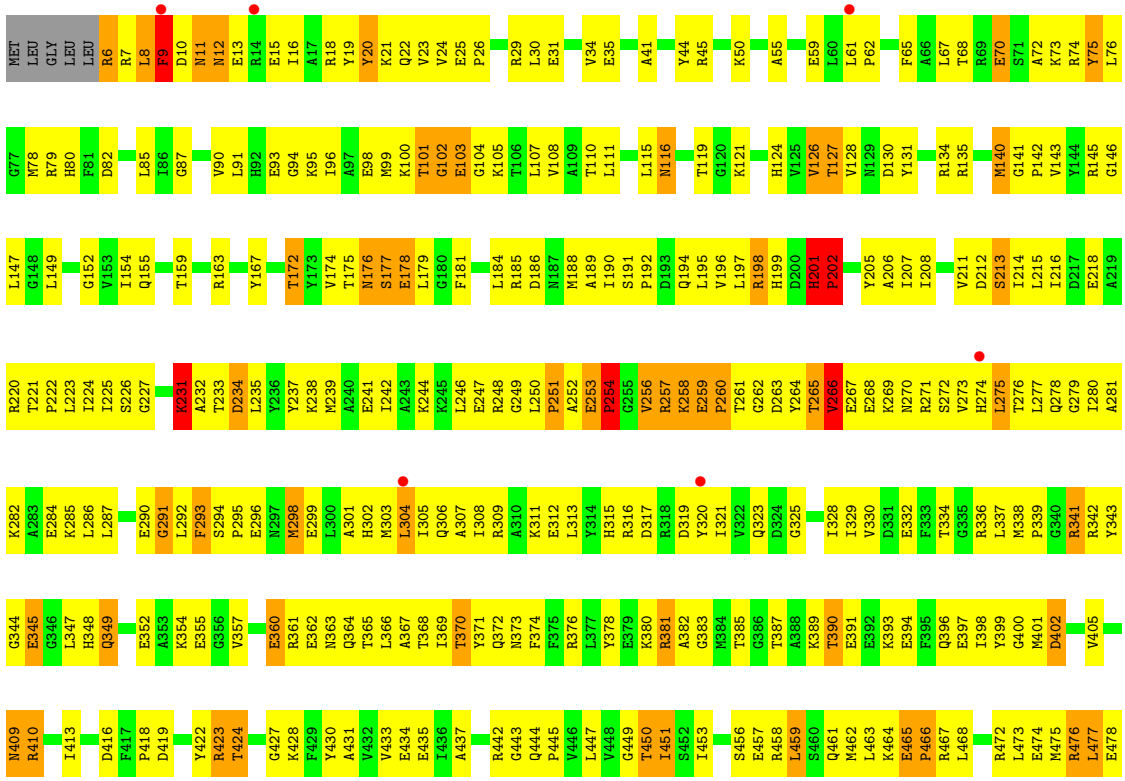


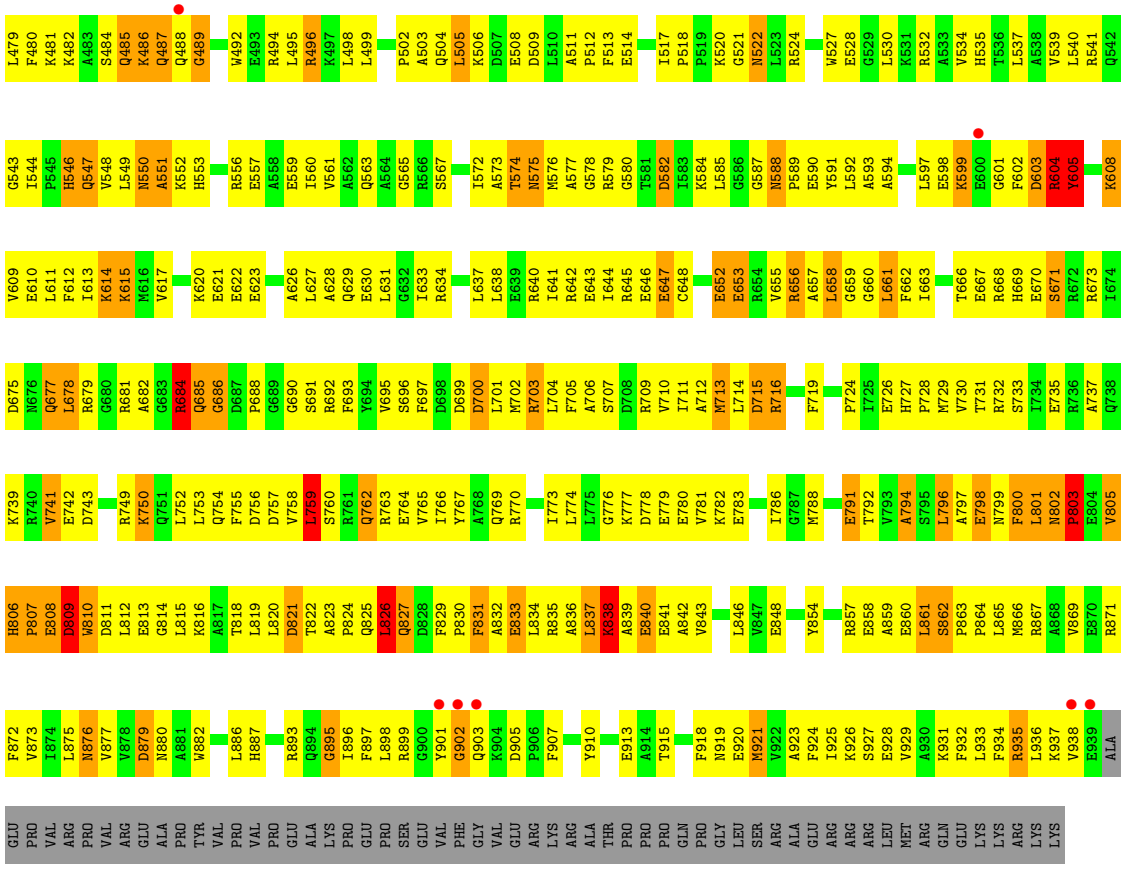


• Molecule 1: Preprotein translocase SecA subunit

Chain D:






L479 F480 K481 K482 A483 S484 Q485 K486 Q487 Q488 G489 W492 E493 R494 L495 R496 K497 L498 L499 P502 A503 Q504 L505 K506 D507 E508 D509 L510 A511 P512 F513 E514 I517 P518 P519 K520 G521 N522 L523 R524 W527 E528 G529 L530 K531 R532 A533 V534 H535 T536 L537 A538 V539 L540 R541 Q542
G543 I544 F545 H546 Q547 V548 L549 N550 A551 K552 H553 R556 E557 A558 E559 I560 V561 A562 Q563 A564 G565 R566 S567 I572 A573 T574 N575 M576 A577 G578 R579 G580 T581 D582 I583 K584 L585 G586 G587 N588 P589 E590 Y591 L592 A593 A594 L597 E598 K599 E600 G601 F602 D603 R604 Y605 K608
V609 E610 L611 F612 I613 K614 K615 M616 V617 K620 E621 E622 E623 A626 L627 A628 Q629 Q630 L631 G632 I633 R634 L637 L638 E639 R640 I641 R642 E643 I644 R645 E646 E647 C648 E652 E653 R654 V655 R656 A657 L658 G659 G660 L661 F662 I663 T666 E667 R668 H669 E670 S671 R672 R673 I674
D675 M676 Q677 L678 R679 G680 R681 A682 G683 R684 Q685 G686 D687 P688 G689 G690 S691 R692 F693 Y694 V695 S696 F697 D698 D699 D700 L701 M702 R703 L704 F705 A706 S707 D708 R709 V710 I711 A712 M713 L714 D715 R716 F719 P724 I725 E726 H727 P728 M729 V730 T731 R732 S733 I734 E735 R736 A737 Q738
K739 R740 V741 E742 D743 R749 K750 Q751 L752 L753 Q754 F755 D756 D757 V758 L759 S760 R761 Q762 R763 E764 V765 I766 Y767 A768 Q769 R770 I773 L774 L775 G776 K777 D778 E779 E780 V781 K782 E783 I786 G787 M788 E791 T792 V793 A794 S795 L796 A797 E798 N799 F800 L801 N802 P803 E804 V805
H806 P807 E808 D809 W810 D811 L812 E813 G814 L815 K816 A817 T818 L819 L820 D821 T822 A823 P824 Q825 L826 Q827 D828 F829 P830 F831 A832 E833 L834 R835 A836 L837 K838 A839 E840 E841 A842 V843 L846 V847 E848 Y854 R857 E858 A859 E860 L861 S862 P863 P864 L865 M866 R867 A868 V869 E870 R871
F872 V873 I874 L875 N876 V877 V878 D879 N880 A881 W882 L886 H887 R893 Q894 G895 I896 F897 L898 R899 G900 Y901 G902 Q903 K904 D905 P906 F907 Y910 E913 A914 T915 F918 N919 E920 M921 V922 A923 F924 I925 K926 S927 E928 V929 A930 K931 F932 L933 F934 R935 L936 K937 V938 E939 ALA
GLU PRO VAL ARG PRO VAL ARG GLU ALA PRO TYR VAL PRO VAL PRO GLU ALA LYS PRO GLU PRO SER GLU VAL PHE GLY VAL GLU ARG LYS ARG ALA THR PRO PRO PRO GLN PRO GLY LEU SER ARG ALA GLU ARG ARG ARG LEU MET ARG GLN GLU LYS LYS ARG LYS LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	168.62Å 168.62Å 149.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 2.80 37.02 – 2.80	Depositor EDS
% Data completeness (in resolution range)	97.4 (40.00-2.80) 97.5 (37.02-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 2.81Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.221 , 0.255 0.215 , 0.251	Depositor DCC
R_{free} test set	6730 reflections (5.89%)	wwPDB-VP
Wilson B-factor (Å^2)	52.4	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 33.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.038 for -h,-k,l 0.499 for h,-h-k,-l 0.039 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	31502	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.08% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.65	1/7682 (0.0%)	0.86	23/10350 (0.2%)
1	B	0.68	2/7646 (0.0%)	0.86	18/10302 (0.2%)
1	C	0.65	0/7682	0.85	19/10350 (0.2%)
1	D	0.66	1/7646 (0.0%)	0.86	21/10302 (0.2%)
All	All	0.66	4/30656 (0.0%)	0.86	81/41304 (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	C	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	716	ARG	CG-CD	8.29	1.72	1.51
1	D	716	ARG	CG-CD	8.19	1.72	1.51
1	B	804	GLU	CB-CG	-5.44	1.41	1.52
1	A	338	MET	CG-SD	5.10	1.94	1.81

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	201	HIS	C-N-CD	-13.17	91.63	120.60
1	B	201	HIS	C-N-CD	-12.87	92.30	120.60
1	A	201	HIS	C-N-CD	-12.44	93.22	120.60
1	C	201	HIS	C-N-CD	-11.34	95.65	120.60
1	D	809	ASP	N-CA-C	7.75	131.91	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	591	TYR	Sidechain
1	C	131	TYR	Sidechain

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	7551	0	7648	744	0
1	B	7515	0	7600	695	0
1	C	7551	0	7648	715	0
1	D	7515	0	7600	809	0
2	A	359	0	0	138	0
2	B	320	0	0	151	0
2	C	344	0	0	147	0
2	D	347	0	0	158	0
All	All	31502	0	30496	2909	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 48.

The worst 5 of 2909 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:D:252:ALA:HB3	1:D:812:LEU:HD11	1.21	1.16
1:D:273:VAL:HG13	1:D:820:LEU:HB2	1.23	1.12
1:C:435:GLU:OE1	1:C:692:ARG:NH1	1.85	1.09
1:A:269:LYS:HD3	1:A:270:ASN:H	1.08	1.08
1:A:273:VAL:HG13	1:A:820:LEU:HB2	1.33	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	937/997 (94%)	789 (84%)	101 (11%)	47 (5%)	2 6
1	B	932/997 (94%)	778 (84%)	108 (12%)	46 (5%)	2 7
1	C	937/997 (94%)	784 (84%)	108 (12%)	45 (5%)	2 7
1	D	932/997 (94%)	777 (83%)	107 (12%)	48 (5%)	2 6
All	All	3738/3988 (94%)	3128 (84%)	424 (11%)	186 (5%)	2 6

5 of 186 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	LEU
1	A	11	ASN
1	A	12	ASN
1	A	231	LYS
1	A	254	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	788/840 (94%)	667 (85%)	121 (15%)	2 8
1	B	784/840 (93%)	660 (84%)	124 (16%)	2 8
1	C	788/840 (94%)	672 (85%)	116 (15%)	3 9
1	D	784/840 (93%)	655 (84%)	129 (16%)	2 7
All	All	3144/3360 (94%)	2654 (84%)	490 (16%)	2 8

5 of 490 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	827	GLN
1	D	661	LEU
1	C	362	GLU
1	D	646	GLU
1	D	811	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	685	GLN
1	D	187	ASN
1	C	745	ASN
1	D	11	ASN
1	D	323	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 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, 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	939/997 (94%)	-0.30	20 (2%)	63	54	18, 61, 86, 103	0
1	B	934/997 (93%)	-0.35	16 (1%)	70	63	16, 59, 85, 112	0
1	C	939/997 (94%)	-0.30	17 (1%)	68	61	20, 61, 88, 102	0
1	D	934/997 (93%)	-0.36	13 (1%)	75	70	21, 59, 87, 112	0
All	All	3746/3988 (93%)	-0.33	66 (1%)	68	61	16, 60, 87, 112	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	274	HIS	5.7
1	A	832	ALA	4.9
1	B	11	ASN	4.5
1	D	938	VAL	4.5
1	D	600	GLU	4.5

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 monosaccharides 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.