



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 03:16 pm GMT

PDB ID : 1WSU  
Title : C-terminal domain of elongation factor selB complexed with SECIS RNA  
Authors : Yoshizawa, S.; Rasubala, L.; Ose, T.; Kohda, D.; Fourmy, D.; Maenaka, K.  
Deposited on : 2004-11-11  
Resolution : 2.30 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

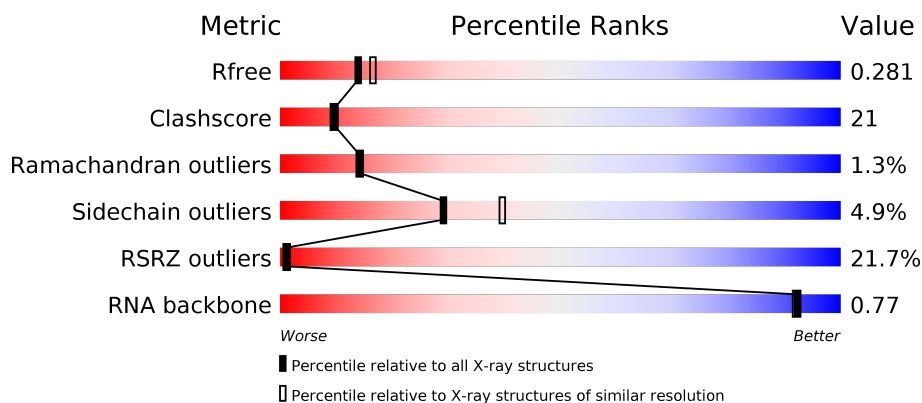
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
$R_{free}$	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)
RNA backbone	2435	1106 (2.80-1.80)

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 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	E	23	
1	F	23	
1	G	23	
2	A	124	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	124	<div><div><div>2%</div><div>74%</div><div>22%</div><div>...</div></div></div>
2	C	124	<div><div><div>33%</div><div>42%</div><div>37%</div><div>•</div><div>18%</div></div></div>
2	D	124	<div><div><div>42%</div><div>45%</div><div>49%</div><div>•</div><div>•</div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5484 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 RNA chain called 5'-R(\*GP\*GP\*CP\*GP\*UP\*UP\*GP\*CP\*CP\*GP\*GP\*UP\*CP\*U\*GP\*GP\*CP\*AP\*AP\*CP\*GP\*CP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	22	Total	C	N	O	P	0	0	0
			465	209	85	151	20			
1	F	22	Total	C	N	O	P	0	0	0
			465	209	85	151	20			
1	G	21	Total	C	N	O	P	0	0	0
			445	200	82	144	19			

- Molecule 2 is a protein called Selenocysteine-specific elongation factor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	124	Total	C	N	O	0	0	0
			1007	642	178	187			
2	B	122	Total	C	N	O	0	0	0
			994	636	175	183			
2	C	102	Total	C	N	O	0	0	0
			845	547	149	149			
2	D	121	Total	C	N	O	0	0	0
			990	633	175	182			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	511	GLY	-	CLONING ARTIFACT	UNP Q46455
B	511	GLY	-	CLONING ARTIFACT	UNP Q46455
C	511	GLY	-	CLONING ARTIFACT	UNP Q46455
D	511	GLY	-	CLONING ARTIFACT	UNP Q46455

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	114	Total 114	O 114	0	0
3	B	75	Total 75	O 75	0	0
3	C	21	Total 21	O 21	0	0
3	D	11	Total 11	O 11	0	0
3	E	22	Total 22	O 22	0	0
3	F	16	Total 16	O 16	0	0
3	G	14	Total 14	O 14	0	0

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: 5'-R(\*GP\*GP\*CP\*GP\*UP\*UP\*GP\*CP\*CP\*GP\*GP\*UP\*CP\*U\*GP\*GP\*CP\*A P\*AP\*CP\*GP\*CP\*C)-3'

Chain E: 



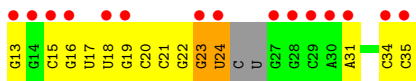
- Molecule 1: 5'-R(\*GP\*GP\*CP\*GP\*UP\*UP\*GP\*CP\*CP\*GP\*GP\*UP\*CP\*U\*GP\*GP\*CP\*A P\*AP\*CP\*GP\*CP\*C)-3'

Chain F: 




- Molecule 1: 5'-R(\*GP\*GP\*CP\*GP\*UP\*UP\*GP\*CP\*CP\*GP\*GP\*UP\*CP\*U\*GP\*GP\*CP\*A P\*AP\*CP\*GP\*CP\*C)-3'

Chain G: 



- Molecule 2: Selenocysteine-specific elongation factor

Chain A: 

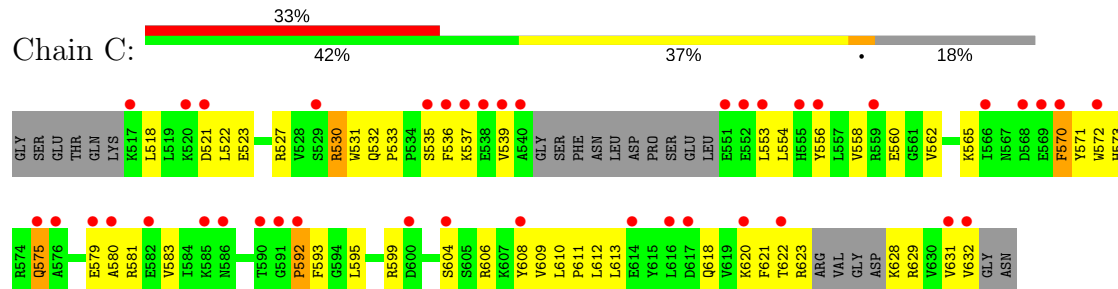


- Molecule 2: Selenocysteine-specific elongation factor

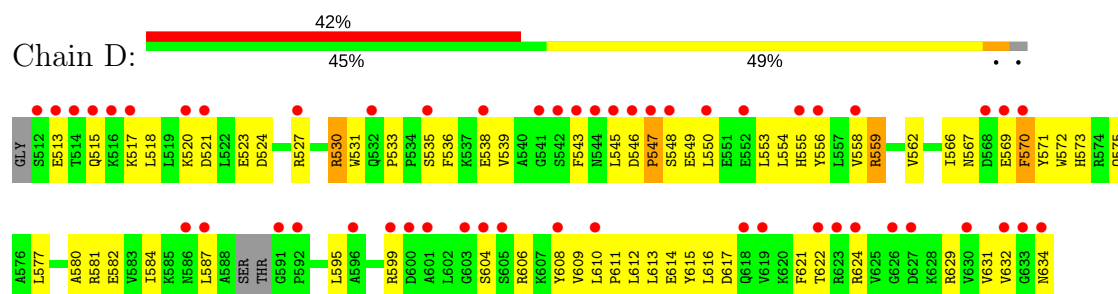
Chain B: 



- Molecule 2: Selenocysteine-specific elongation factor



- Molecule 2: Selenocysteine-specific elongation factor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.86Å 169.81Å 71.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.30 9.99 – 2.30	Depositor EDS
% Data completeness (in resolution range)	93.2 (10.00-2.30) 93.2 (9.99-2.30)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 2.31Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.224 , 0.279 0.227 , 0.281	Depositor DCC
$R_{free}$ test set	4099 reflections (10.91%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.0	Xtriage
Anisotropy	0.539	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 86.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5484	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	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 57.57 % 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 2.3220e-05. 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, 0.333 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	E	0.43	0/518	0.74	0/805
1	F	0.35	0/518	0.67	0/805
1	G	0.23	0/496	0.65	0/771
2	A	0.46	0/1027	0.66	0/1383
2	B	0.44	0/1014	0.65	0/1367
2	C	0.29	0/861	0.50	0/1158
2	D	0.30	0/1009	0.54	0/1357
All	All	0.37	0/5443	0.63	0/7646

There are no bond length outliers.

There are no bond angle outliers.

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	E	465	0	243	13	0
1	F	465	0	243	14	0
1	G	445	0	232	23	0
2	A	1007	0	1014	24	0
2	B	994	0	1005	24	0
2	C	845	0	865	51	0
2	D	990	0	998	71	0
3	A	114	0	0	10	0
3	B	75	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	21	0	0	3	0
3	D	11	0	0	2	0
3	E	22	0	0	0	0
3	F	16	0	0	0	0
3	G	14	0	0	3	0
All	All	5484	0	4600	207	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 21.

The worst 5 of 207 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)
2:C:613:LEU:HB3	2:C:629:ARG:HH12	0.91	1.07
2:D:535:SER:OG	2:D:538:GLU:HG3	1.58	1.02
2:D:613:LEU:HB3	2:D:629:ARG:HH12	1.23	1.02
2:C:613:LEU:HB3	2:C:629:ARG:NH1	1.74	1.01
1:G:17:U:H3	1:G:31:A:H62	1.16	0.90

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	
2	A	122/124 (98%)	117 (96%)	5 (4%)	0	100	100
2	B	120/124 (97%)	116 (97%)	3 (2%)	1 (1%)	22	26
2	C	96/124 (77%)	86 (90%)	9 (9%)	1 (1%)	18	20
2	D	117/124 (94%)	104 (89%)	9 (8%)	4 (3%)	4	2
All	All	455/496 (92%)	423 (93%)	26 (6%)	6 (1%)	14	14

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	547	PRO
2	D	548	SER
2	B	530	ARG
2	D	587	LEU
2	D	631	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	
2	A	108/108 (100%)	103 (95%)	5 (5%)	31	42
2	B	107/108 (99%)	101 (94%)	6 (6%)	25	33
2	C	90/108 (83%)	85 (94%)	5 (6%)	25	33
2	D	106/108 (98%)	102 (96%)	4 (4%)	38	52
All	All	411/432 (95%)	391 (95%)	20 (5%)	29	39

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

Mol	Chain	Res	Type
2	B	623	ARG
2	B	624	ARG
2	D	530	ARG
2	B	566	ILE
2	B	575	GLN

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

Mol	Chain	Res	Type
2	B	544	ASN
2	B	573	HIS
2	D	544	ASN
2	B	515	GLN
2	C	573	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	E	20/23 (86%)	3 (15%)	0
1	F	20/23 (86%)	3 (15%)	0
1	G	19/23 (82%)	2 (10%)	0
All	All	59/69 (85%)	8 (13%)	0

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	E	23	G
1	E	24	U
1	E	25	C
1	F	23	G
1	F	24	U

There are no RNA pucker outliers to report.

## 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(Å²)	Q<0.9
1	E	22/23 (95%)	0.16	0	100100	38, 57, 70, 73	0
1	F	22/23 (95%)	0.54	0	100100	50, 72, 88, 101	0
1	G	21/23 (91%)	2.62	15 (71%)	00	89, 110, 143, 161	0
2	A	124/124 (100%)	0.06	5 (4%)	3946	17, 36, 69, 102	0
2	B	122/124 (98%)	0.09	3 (2%)	5865	21, 40, 63, 100	0
2	C	102/124 (82%)	1.78	41 (40%)	00	60, 94, 113, 136	0
2	D	121/124 (97%)	2.22	52 (42%)	00	62, 90, 127, 137	0
All	All	534/565 (94%)	1.01	116 (21%)	11	17, 68, 118, 161	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	633	GLY	10.3
2	C	590	THR	9.1
2	D	514	THR	8.4
2	D	627	ASP	7.2
1	G	27	G	7.2

### 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.