wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 12:15 am BST

PDB ID	:	3W1K
Title	:	Crystal structure of the selenocysteine synthase SelA and tRNASec complex
Authors	:	Itoh, Y.; Sekine, S.; Yokoyama, S.
Deposited on	:	2012-11-15
Resolution	:	7.50 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

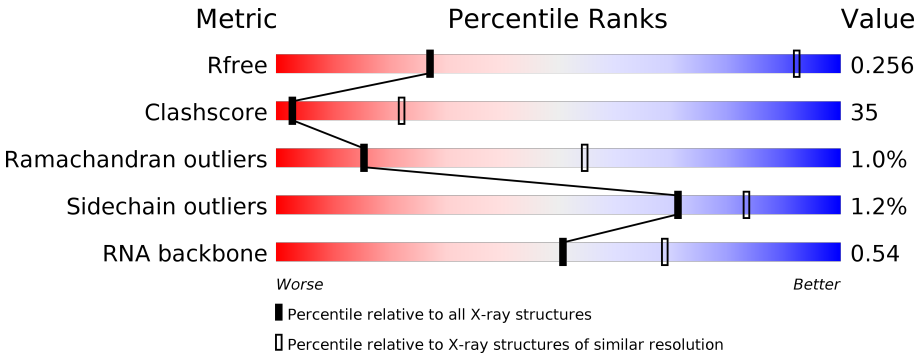
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 7.50 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.



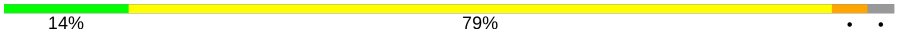
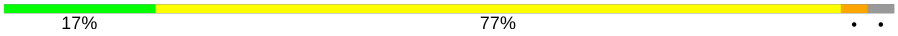
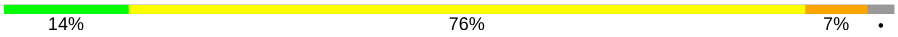
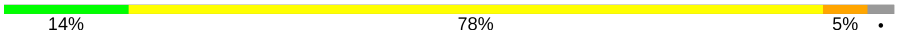
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)
RNA backbone	3102	1079 (11.50-3.00)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	452	44% 54% .
1	B	452	46% 54%
1	C	452	48% 51% .
1	D	452	47% 52% .
1	E	452	46% 53% .
2	F	95	16% 78% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	95	
2	H	95	
2	I	95	
2	J	95	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 27660 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 L-seryl-tRNA(Sec) selenium transferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	452	Total	C	N	O	P	S	0	0	0
			3575	2278	620	665	1	11			
1	B	452	Total	C	N	O	P	S	0	0	0
			3575	2278	620	665	1	11			
1	C	452	Total	C	N	O	P	S	0	0	0
			3575	2278	620	665	1	11			
1	D	452	Total	C	N	O	P	S	0	0	0
			3575	2278	620	665	1	11			
1	E	452	Total	C	N	O	P	S	0	0	0
			3575	2278	620	665	1	11			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	ALA	LYS	ENGINEERED MUTATION	UNP O67140
A	21	ALA	LYS	ENGINEERED MUTATION	UNP O67140
A	46	ALA	LYS	ENGINEERED MUTATION	UNP O67140
A	48	ALA	LYS	ENGINEERED MUTATION	UNP O67140
B	19	ALA	LYS	ENGINEERED MUTATION	UNP O67140
B	21	ALA	LYS	ENGINEERED MUTATION	UNP O67140
B	46	ALA	LYS	ENGINEERED MUTATION	UNP O67140
B	48	ALA	LYS	ENGINEERED MUTATION	UNP O67140
C	19	ALA	LYS	ENGINEERED MUTATION	UNP O67140
C	21	ALA	LYS	ENGINEERED MUTATION	UNP O67140
C	46	ALA	LYS	ENGINEERED MUTATION	UNP O67140
C	48	ALA	LYS	ENGINEERED MUTATION	UNP O67140
D	19	ALA	LYS	ENGINEERED MUTATION	UNP O67140
D	21	ALA	LYS	ENGINEERED MUTATION	UNP O67140
D	46	ALA	LYS	ENGINEERED MUTATION	UNP O67140
D	48	ALA	LYS	ENGINEERED MUTATION	UNP O67140
E	19	ALA	LYS	ENGINEERED MUTATION	UNP O67140
E	21	ALA	LYS	ENGINEERED MUTATION	UNP O67140
E	46	ALA	LYS	ENGINEERED MUTATION	UNP O67140

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	48	ALA	LYS	ENGINEERED MUTATION	UNP O67140

- Molecule 2 is a RNA chain called selenocysteine tRNA.

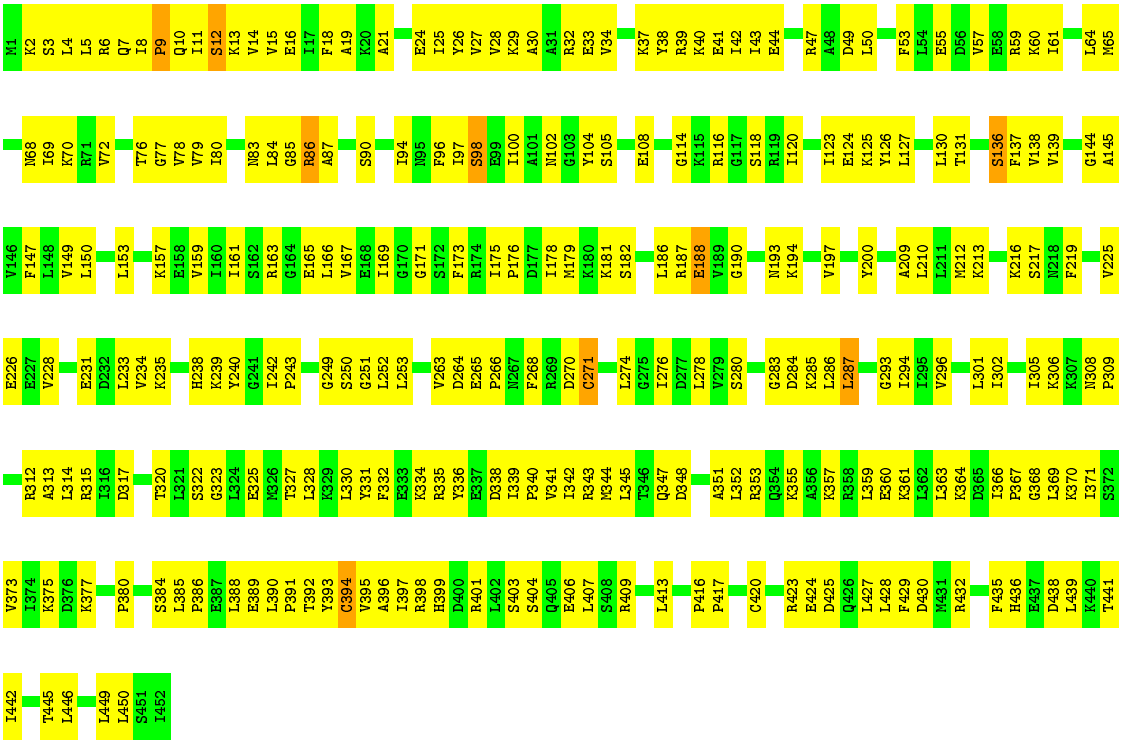
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	92	Total	C	N	O	P	0	0	0
			1957	873	342	650	92			
2	G	92	Total	C	N	O	P	0	0	0
			1957	873	342	650	92			
2	H	92	Total	C	N	O	P	0	0	0
			1957	873	342	650	92			
2	I	92	Total	C	N	O	P	0	0	0
			1957	873	342	650	92			
2	J	92	Total	C	N	O	P	0	0	0
			1957	873	342	650	92			

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. 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. 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.

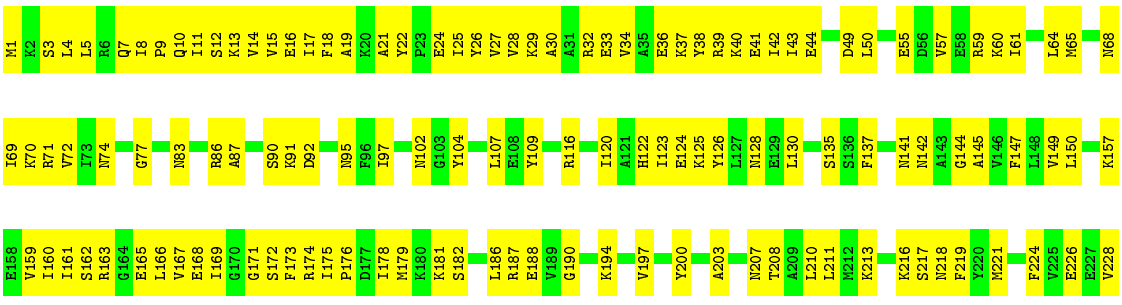
- Molecule 1: L-seryl-tRNA(Sec) selenium transferase

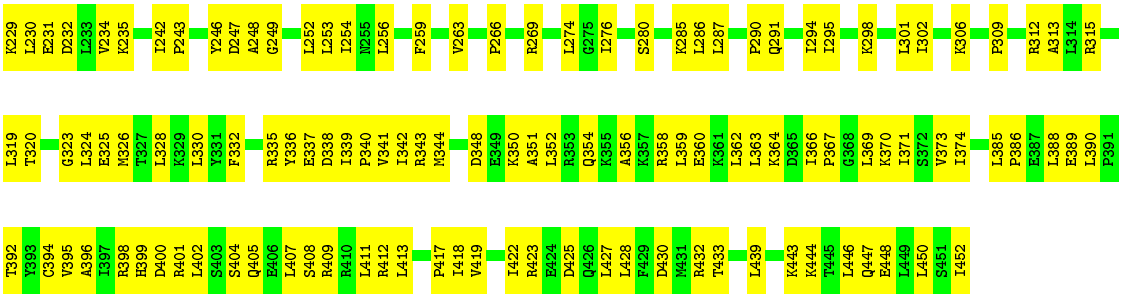
Chain A:



- Molecule 1: L-seryl-tRNA(Sec) selenium transferase

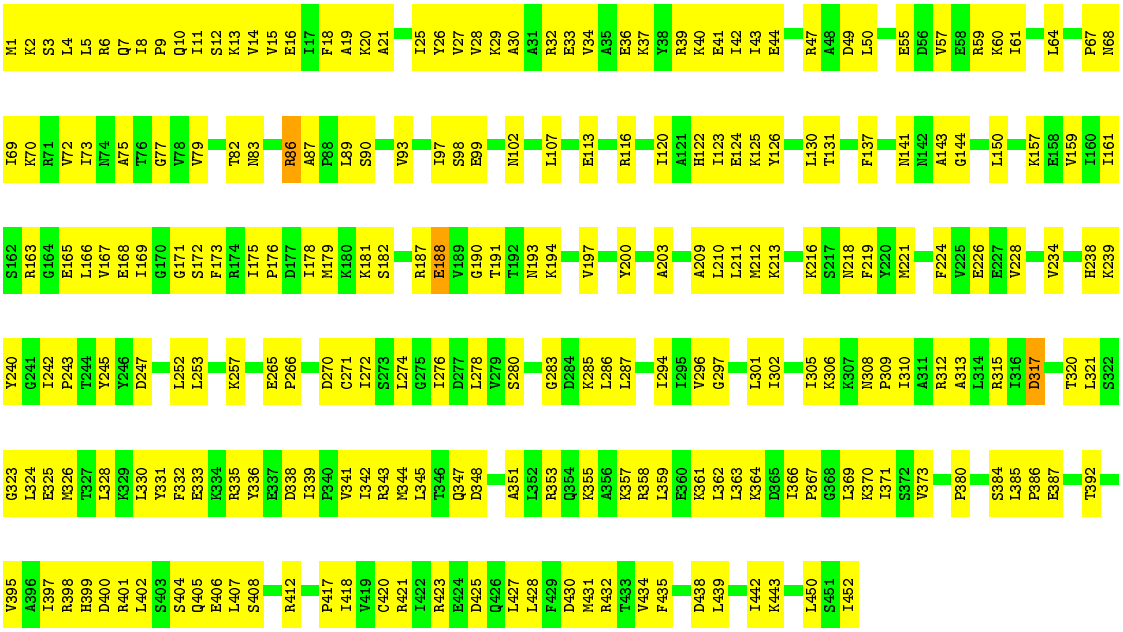
Chain B:





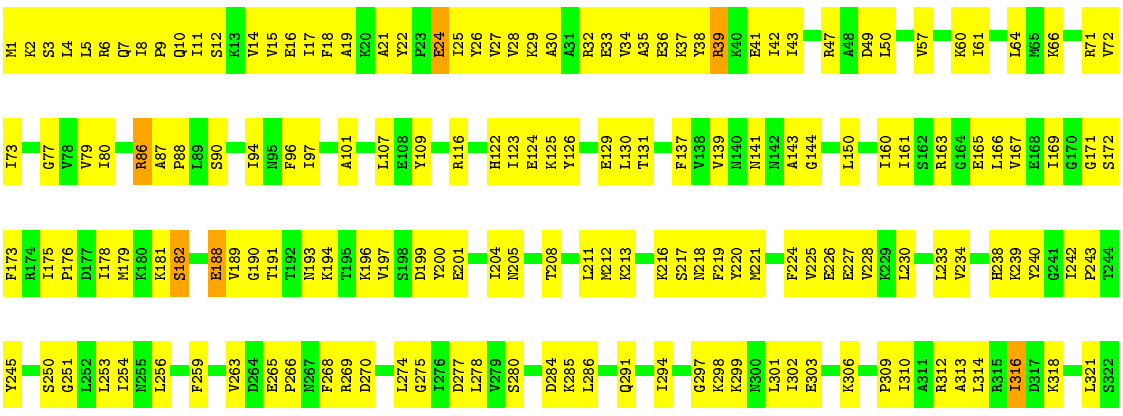
• Molecule 1: L-seryl-tRNA(Sec) selenium transferase

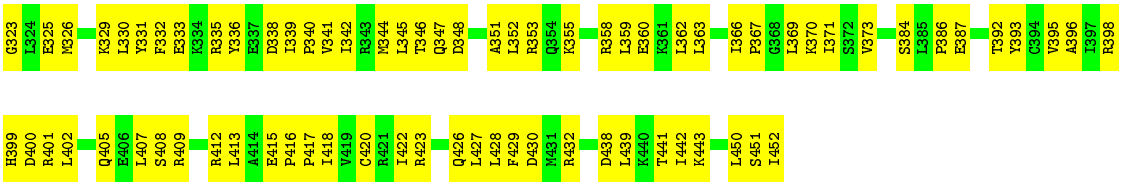
Chain C:



• Molecule 1: L-seryl-tRNA(Sec) selenium transferase

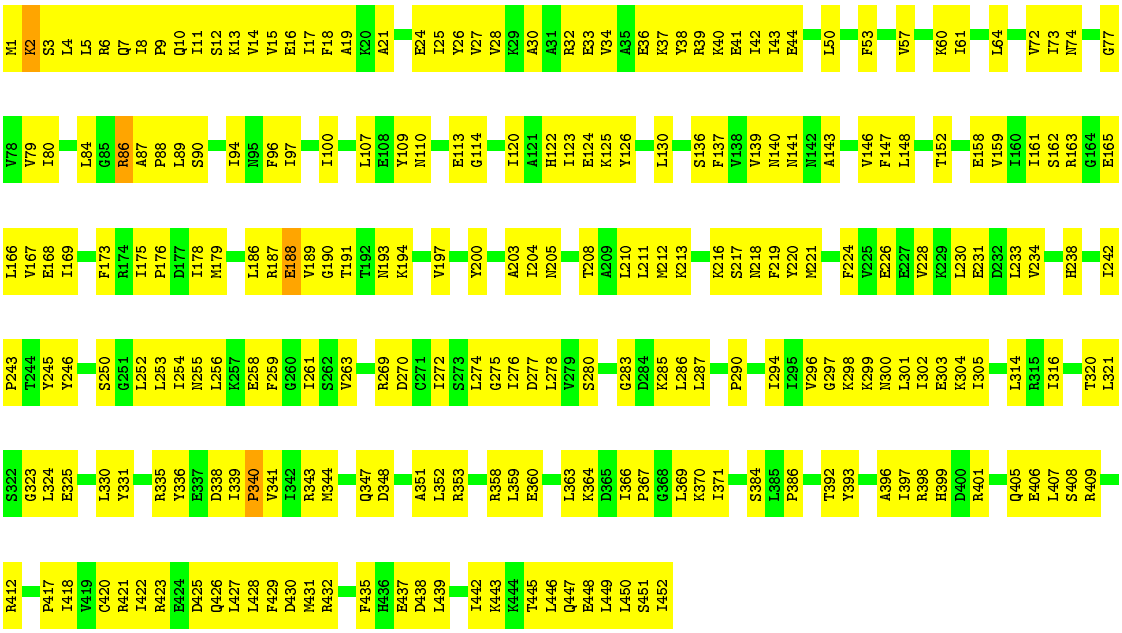
Chain D:





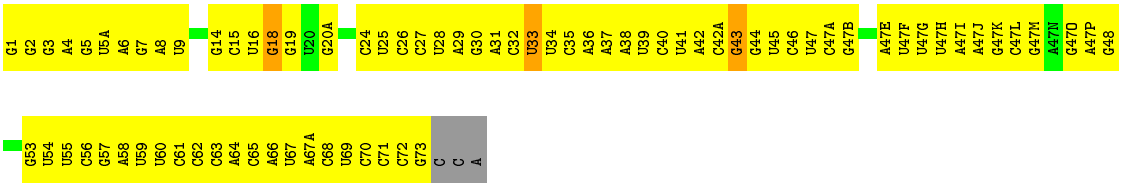
• Molecule 1: L-seryl-tRNA(Sec) selenium transferase

Chain E:



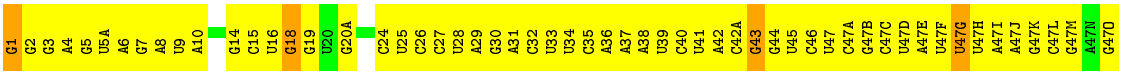
• Molecule 2: selenocysteine tRNA

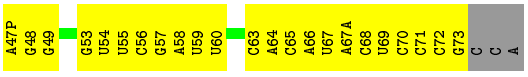
Chain F:



• Molecule 2: selenocysteine tRNA

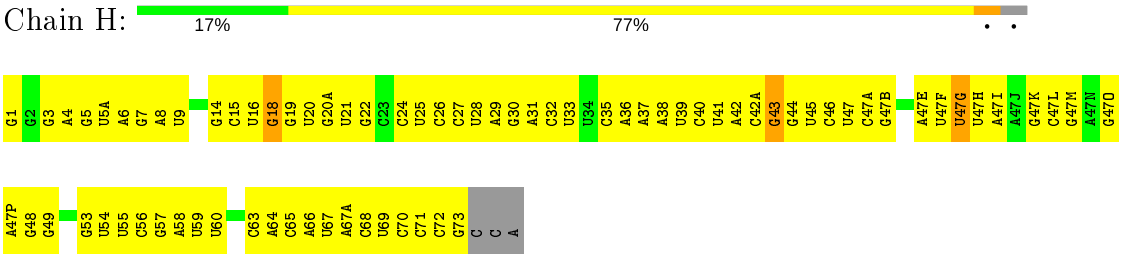
Chain G:





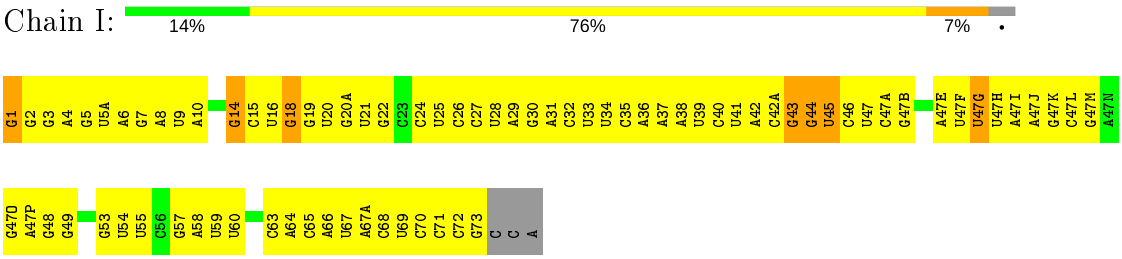
• Molecule 2: selenocysteine tRNA

Chain H:



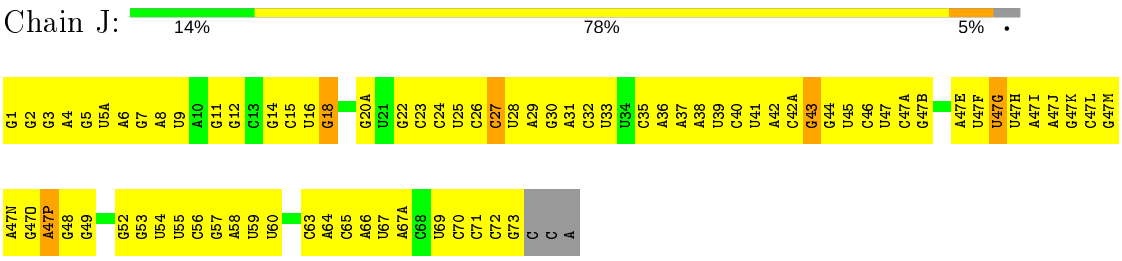
- Molecule 2: selenocysteine tRNA

Chain I:



- Molecule 2: selenocysteine tRNA

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	148.55Å 355.97Å 165.50Å 90.00° 115.41° 90.00°	Depositor
Resolution (Å)	49.94 – 7.50 49.94 – 7.49	Depositor EDS
% Data completeness (in resolution range)	98.2 (49.94-7.50) 98.0 (49.94-7.49)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 7.37Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.194 , 0.240 0.205 , 0.256	Depositor DCC
R_{free} test set	475 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	391.5	Xtriage
Anisotropy	0.388	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 562.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.085 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.71	EDS
Total number of atoms	27660	wwPDB-VP
Average B, all atoms (Å^2)	525.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.56% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: LLP

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	0.41	0/3597	0.65	0/4831
1	B	0.39	0/3597	0.65	0/4831
1	C	0.39	0/3597	0.64	0/4831
1	D	0.41	0/3597	0.65	0/4831
1	E	0.40	0/3597	0.64	0/4831
2	F	0.49	1/2185 (0.0%)	0.75	0/3401
2	G	0.44	1/2185 (0.0%)	0.73	0/3401
2	H	0.42	1/2185 (0.0%)	0.73	0/3401
2	I	0.43	1/2185 (0.0%)	0.72	0/3401
2	J	0.43	1/2185 (0.0%)	0.72	2/3401 (0.1%)
All	All	0.42	5/28910 (0.0%)	0.68	2/41160 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	G	OP3-P	-7.24	1.52	1.61
2	I	1	G	OP3-P	-7.05	1.52	1.61
2	J	1	G	OP3-P	-7.05	1.52	1.61
2	H	1	G	OP3-P	-6.86	1.52	1.61
2	G	1	G	OP3-P	-6.77	1.53	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	47(P)	A	OP2-P-O3'	6.24	118.93	105.20
2	J	27	C	OP2-P-O3'	5.73	117.81	105.20

There are no chirality outliers.

There are no planarity outliers.

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	3575	0	3758	296	0
1	B	3575	0	3758	290	0
1	C	3575	0	3758	284	0
1	D	3575	0	3758	300	0
1	E	3575	0	3758	256	0
2	F	1957	0	989	102	0
2	G	1957	0	989	124	0
2	H	1957	0	989	105	0
2	I	1957	0	989	112	0
2	J	1957	0	989	91	0
All	All	27660	0	23735	1788	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 35.

The worst 5 of 1788 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:C:157:LYS:HE2	1:C:209:ALA:HB2	1.27	1.14
1:B:1:MET:HG3	1:B:4:LEU:HD12	1.31	1.12
1:C:102:ASN:HB3	1:D:71:ARG:HH22	1.15	1.08
1:A:10:GLN:H	1:A:13:LYS:HD2	1.13	1.08
2:J:32:C:H2'	2:J:33:U:C6	1.93	1.02

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	449/452 (99%)	423 (94%)	22 (5%)	4 (1%)	17 57
1	B	449/452 (99%)	419 (93%)	25 (6%)	5 (1%)	14 52
1	C	449/452 (99%)	421 (94%)	24 (5%)	4 (1%)	17 57
1	D	449/452 (99%)	421 (94%)	24 (5%)	4 (1%)	17 57
1	E	449/452 (99%)	420 (94%)	23 (5%)	6 (1%)	12 48
All	All	2245/2260 (99%)	2104 (94%)	118 (5%)	23 (1%)	15 54

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	9	PRO
1	E	9	PRO
1	B	124	GLU
1	C	124	GLU
1	D	124	GLU

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	390/390 (100%)	380 (97%)	10 (3%)	46 66
1	B	390/390 (100%)	389 (100%)	1 (0%)	92 95
1	C	390/390 (100%)	385 (99%)	5 (1%)	69 81
1	D	390/390 (100%)	384 (98%)	6 (2%)	65 80
1	E	390/390 (100%)	389 (100%)	1 (0%)	92 95
All	All	1950/1950 (100%)	1927 (99%)	23 (1%)	71 83

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	188	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	188	GLU
1	D	316	ILE
1	C	3	SER
1	C	317	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	68	ASN
1	C	291	GLN
1	E	347	GLN
1	C	95	ASN
1	C	122	HIS

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	F	91/95 (95%)	6 (6%)	0
2	G	91/95 (95%)	6 (6%)	0
2	H	91/95 (95%)	6 (6%)	0
2	I	91/95 (95%)	9 (9%)	1 (1%)
2	J	91/95 (95%)	6 (6%)	0
All	All	455/475 (95%)	33 (7%)	1 (0%)

5 of 33 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	F	18	G
2	F	20(A)	G
2	F	33	U
2	F	42(A)	C
2	F	43	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	I	44	G

5.4 Non-standard residues in protein, DNA, RNA chains (i)

5 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	LLP	B	285	1	23,24,25	1.24	2 (8%)	25,32,34	1.03	1 (4%)
1	LLP	E	285	1	23,24,25	1.15	1 (4%)	25,32,34	1.02	1 (4%)
1	LLP	D	285	1	23,24,25	1.14	1 (4%)	25,32,34	0.90	1 (4%)
1	LLP	A	285	1	23,24,25	1.17	2 (8%)	25,32,34	0.95	1 (4%)
1	LLP	C	285	1	23,24,25	1.14	1 (4%)	25,32,34	0.90	1 (4%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	B	285	1	-	6/16/17/19	0/1/1/1
1	LLP	E	285	1	-	2/16/17/19	0/1/1/1
1	LLP	D	285	1	-	3/16/17/19	0/1/1/1
1	LLP	A	285	1	-	2/16/17/19	0/1/1/1
1	LLP	C	285	1	-	4/16/17/19	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	285	LLP	C4-C4'	-3.15	1.40	1.46
1	E	285	LLP	C4-C4'	-2.90	1.41	1.46
1	A	285	LLP	C4-C4'	-2.78	1.41	1.46
1	C	285	LLP	C4-C4'	-2.52	1.41	1.46
1	D	285	LLP	C4-C4'	-2.44	1.42	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	285	LLP	OP4-P-OP1	2.68	114.00	106.47
1	B	285	LLP	OP4-P-OP1	2.54	113.59	106.47
1	A	285	LLP	OP4-P-OP1	2.33	113.00	106.47
1	D	285	LLP	OP4-P-OP1	2.23	112.73	106.47
1	C	285	LLP	OP4-P-OP1	2.22	112.69	106.47

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	285	LLP	C4-C5-C5'-OP4
1	B	285	LLP	C6-C5-C5'-OP4
1	B	285	LLP	C-CA-CB-CG
1	B	285	LLP	O-C-CA-CB
1	E	285	LLP	C-CA-CB-CG

There are no ring outliers.

5 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	285	LLP	6	0
1	E	285	LLP	3	0
1	D	285	LLP	3	0
1	A	285	LLP	4	0
1	C	285	LLP	7	0

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

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.