wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:55 PM GMT

PDB ID	:	1Z2Z
Title	:	Crystal Structure of the Putative tRNA pseudouridine synthase D (TruD) from Methanosarcina mazei, Northeast Structural Genomics Target MaR1
Authors	:	Forouhar, F.; Yong, W.; Ciano, M.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2005-03-10
Resolution	:	2.60 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 (i) symbol.

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

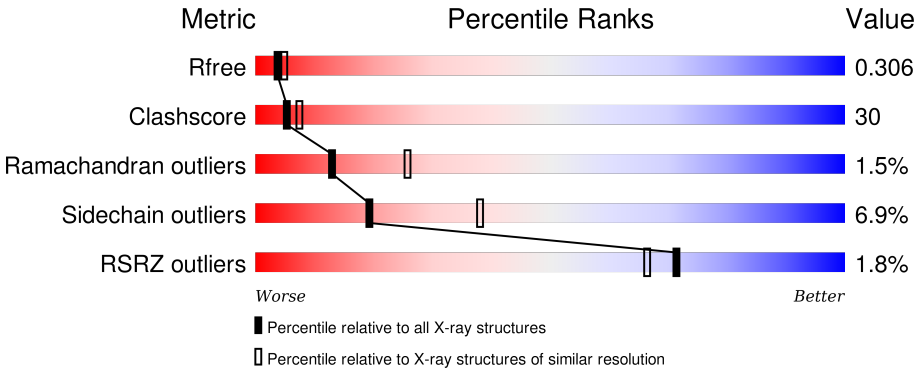
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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 <=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	446	2% 45% 48% 5%
1	B	446	% 55% 37% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7176 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 Probable tRNA pseudouridine synthase D.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	437	Total	C	N	O	S	Se	0	0	0
			3478	2204	599	664	2	9			
1	B	437	Total	C	N	O	S	Se	0	0	0
			3478	2204	599	664	2	9			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	243	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	244	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	268	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	326	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	379	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	415	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	430	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	437	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
A	439	LEU	-	CLONING ARTIFACT	UNP Q8Q0M2
A	440	GLU	-	CLONING ARTIFACT	UNP Q8Q0M2
A	441	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
A	442	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
A	443	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
A	444	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
A	445	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
A	446	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	243	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	244	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	268	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	326	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	379	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	415	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	430	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	437	MSE	MET	MODIFIED RESIDUE	UNP Q8Q0M2
B	439	LEU	-	CLONING ARTIFACT	UNP Q8Q0M2
B	440	GLU	-	CLONING ARTIFACT	UNP Q8Q0M2
B	441	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	442	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	443	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	444	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	445	HIS	-	EXPRESSION TAG	UNP Q8Q0M2
B	446	HIS	-	EXPRESSION TAG	UNP Q8Q0M2

- Molecule 2 is water.

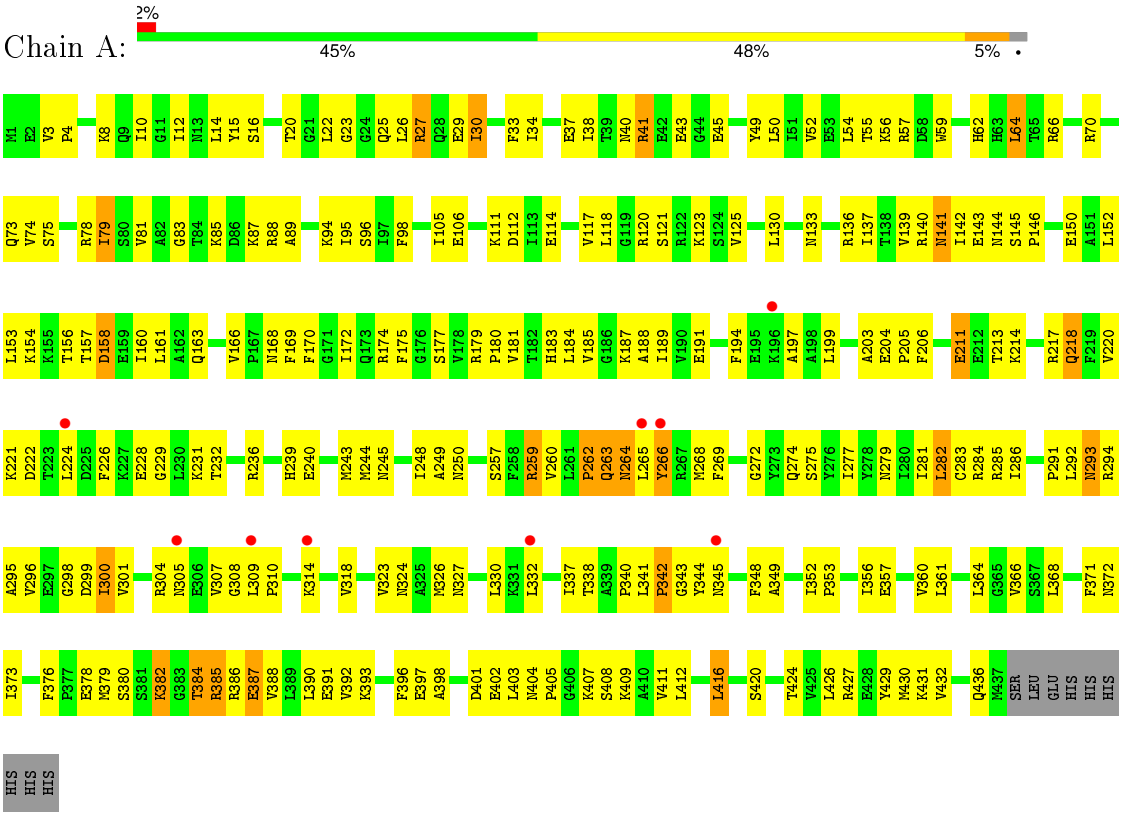
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	106	Total O 106 106	0	0
2	B	114	Total O 114 114	0	0

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 errors 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.

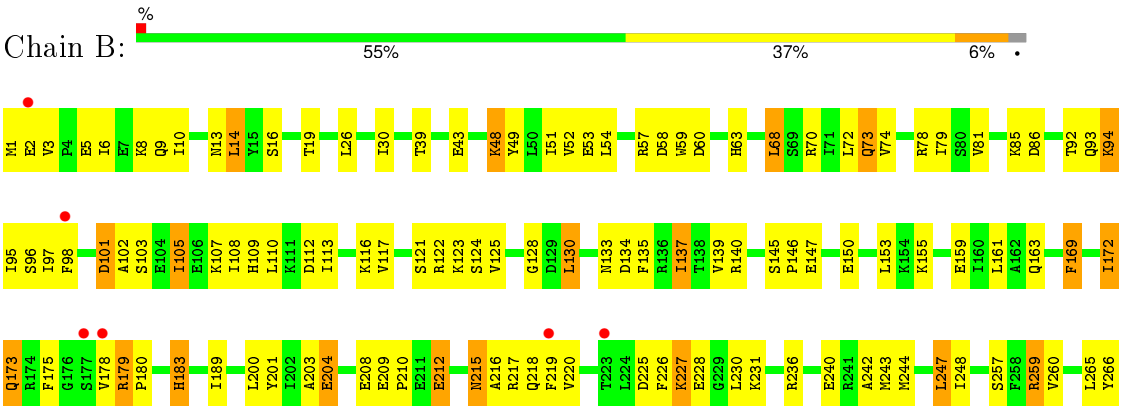
- Molecule 1: Probable tRNA pseudouridine synthase D

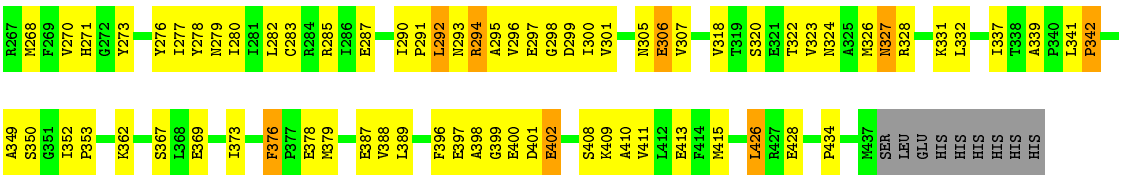
Chain A:



- Molecule 1: Probable tRNA pseudouridine synthase D

Chain B:




R267 M268 F269 V270 H271 G272 Y273 Y276 I277 Y278 N279 I280 I281 L282 C283 R284 R285 I286 E287 I290 P291 L292 N293 R294 A295 V296 E297 G298 D299 I300 V301 N305 E306 V307 V318 I319 S320 E321 T322 V323 N324 A325 N326 N327 R328 K331 L332 I337 T338 A339 P340 L341 P342 A349 S350 G351 I352 P353 K362 S367 L368 E369 I373 F376 P377 E378 M379 E387 V388 L389 F396 E397 A398 G399 E400 D401 E402 S408 K409 A410 V411 L412 E413 F414 M415 L426 R427 E428 P434 M437 SER LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.56Å 132.53Å 66.78Å 90.00° 116.98° 90.00°	Depositor
Resolution (Å)	29.03 – 2.60 29.03 – 2.60	Depositor EDS
% Data completeness (in resolution range)	90.9 (29.03-2.60) 98.5 (29.03-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	4.89 (at 2.61Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.233 , 0.296 0.246 , 0.306	Depositor DCC
R_{free} test set	2895 reflections (9.86%)	DCC
Wilson B-factor (Å^2)	40.0	Xtriage
Anisotropy	0.684	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 53.6	EDS
Estimated twinning fraction	0.021 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 58135 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7176	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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, 0.375 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	0.41	0/3531	0.59	0/4748
1	B	0.43	0/3531	0.63	0/4748
All	All	0.42	0/7062	0.61	0/9496

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

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	3478	0	3522	228	0
1	B	3478	0	3522	203	0
2	A	106	0	0	4	0
2	B	114	0	0	11	0
All	All	7176	0	7044	423	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 30.

The worst 5 of 423 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:401:ASP:HB3	1:B:409:LYS:HB3	1.44	0.97

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:6:ILE:HD12	1:B:6:ILE:H	1.31	0.95
1:A:170:PHE:H	1:A:279:ASN:HD21	1.14	0.92
1:A:309:LEU:HD12	1:A:310:PRO:HD2	1.51	0.90
1:B:291:PRO:HG2	1:B:294:ARG:HB2	1.54	0.88

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	435/446 (98%)	380 (87%)	47 (11%)	8 (2%)	11	21
1	B	435/446 (98%)	397 (91%)	33 (8%)	5 (1%)	17	36
All	All	870/892 (98%)	777 (89%)	80 (9%)	13 (2%)	13	26

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	264	ASN
1	B	73	GLN
1	B	306	GLU
1	A	143	GLU
1	B	86	ASP

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	382/382 (100%)	360 (94%)	22 (6%)	25 49
1	B	382/382 (100%)	351 (92%)	31 (8%)	15 28
All	All	764/764 (100%)	711 (93%)	53 (7%)	19 38

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	48	LYS
1	B	130	LEU
1	B	327	ASN
1	B	68	LEU
1	B	101	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	327	ASN
1	B	25	GLN
1	B	324	ASN
1	B	9	GLN
1	B	109	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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

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^{th} 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	428/446 (95%)	0.11	9 (2%)	67	61	12, 49, 76, 83	0
1	B	428/446 (95%)	-0.08	6 (1%)	78	74	14, 36, 66, 77	0
All	All	856/892 (95%)	0.02	15 (1%)	71	66	12, 41, 74, 83	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	265	LEU	3.0
1	A	345	ASN	2.9
1	B	178	VAL	2.9
1	B	219	PHE	2.8
1	A	266	TYR	2.8

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.