



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 06:56 PM GMT

PDB ID : 1D8I  
Title : X-RAY CRYSTAL STRUCTURE OF YEAST RNA TRIPHOSPHATASE IN COMPLEX WITH A SULFATE ION.  
Authors : Lima, C.D.; Wang, L.K.; Shuman, S.  
Deposited on : 1999-10-24  
Resolution : 2.05 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

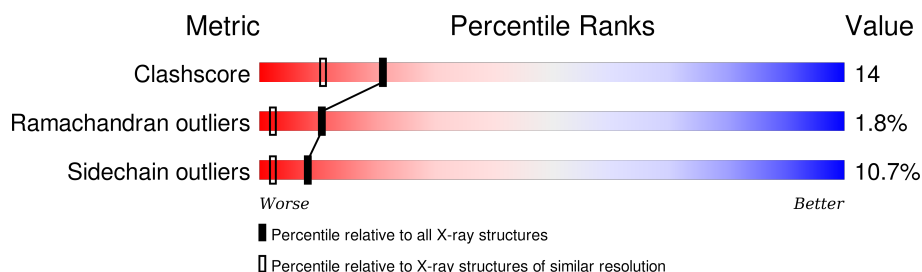
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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))
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	311	
1	B	311	
1	C	311	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7329 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.

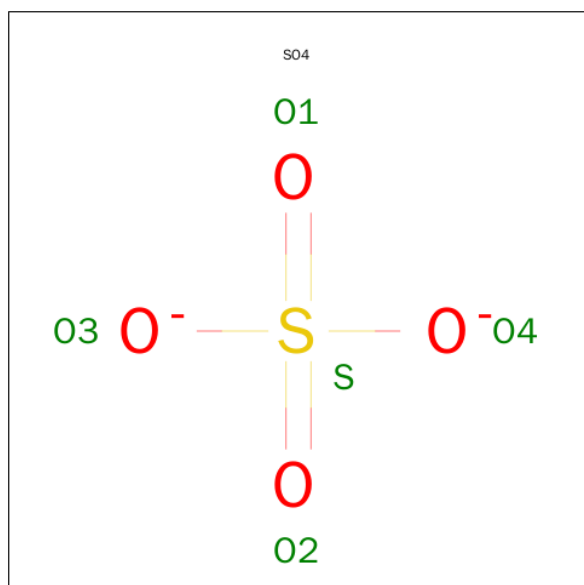
- Molecule 1 is a protein called MRNA TRIPHOSPHATASE CET1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	0	0
			2295	1457	393	440	5			
1	B	288	Total	C	N	O	S	0	0	0
			2295	1457	393	440	5			
1	C	288	Total	C	N	O	S	0	0	0
			2295	1457	393	440	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	242	ARG	LYS	CONFLICT	UNP O13297
B	242	ARG	LYS	CONFLICT	UNP O13297
C	242	ARG	LYS	CONFLICT	UNP O13297

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0

- Molecule 3 is water.

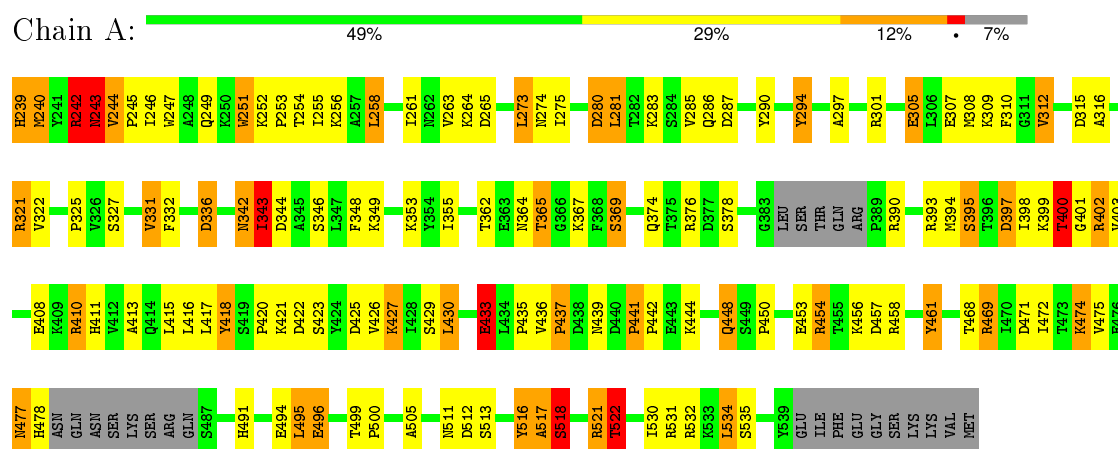
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	121	Total O 121 121	0	0
3	B	134	Total O 134 134	0	0
3	C	159	Total O 159 159	0	0

### 3 Residue-property plots

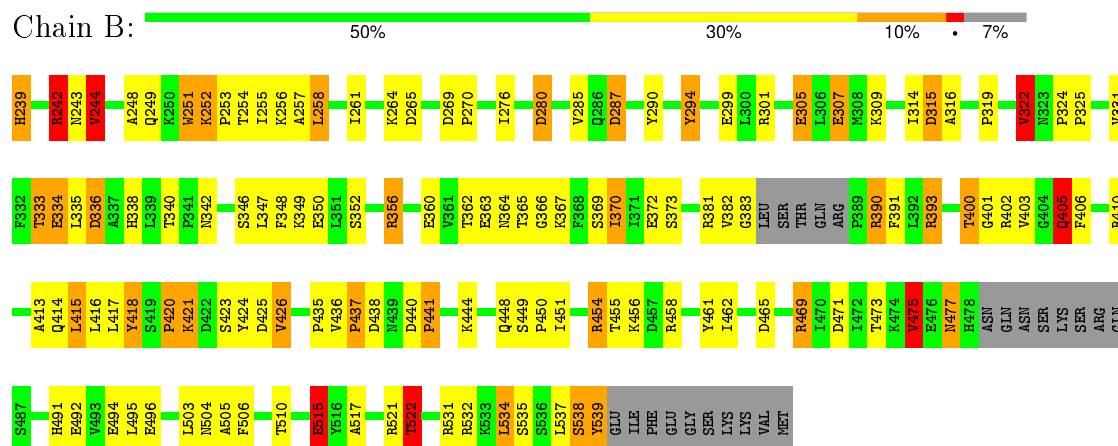
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

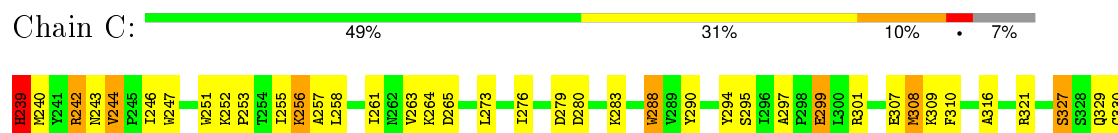
#### • Molecule 1: MRNA TRIPHOSPHATASE CET1



#### • Molecule 1: MRNA TRIPHOSPHATASE CET1



#### • Molecule 1: MRNA TRIPHOSPHATASE CET1



V331	F332	T333	E334	I335	D336	K337	H338		N342	T343		K349	E350	L351	S352	K353	Y354	I355	K356	G357	I358		T362	F363	N364	T365	G366	K367		I370	I371		Q374		D377		Y380	R381	V382	G383	LEU	SER	THR	GLN	ARG		P389	R390	F391	L392	R393		T400	G401	R402		Q405	F406		R410																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
	L415	L416	L417	Y418	S419	P420	K421	D422	S423	Y424	D425	V426	K427		E433	L434	P435	V436	P437	D438	N439	D440	P441	P442		K446		I451		R454	T455	R456	D457	R458		Y461		N464		T468	R469		I472		V475	E476	N477	H478	ASN	GLN	ASN	SER	LYS	SER	ARG	GLN	S487		T490																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
H491		E494		I497		A501	L502	L503	N504	A505	F506	D507	H508	L509	T510	N511	D512	S513	R514	E515	Y516	A517	S518		R521	T522	F523	L524	N525	N526	G527	T528		R531	R532	K533	L534		S538	Y539	GLU	ILE	PHE	GLU	GLY	SER	LYS	LYS	VAL	MET																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.36Å 116.09Å 82.31Å 90.00° 110.25° 90.00°	Depositor
Resolution (Å)	25.00 – 2.05	Depositor
% Data completeness (in resolution range)	82.2 (25.00-2.05)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.203 , 0.270	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7329	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.99	2/2342 (0.1%)	2.17	84/3178 (2.6%)
1	B	0.93	1/2342 (0.0%)	2.00	72/3178 (2.3%)
1	C	1.00	6/2342 (0.3%)	2.08	85/3178 (2.7%)
All	All	0.98	9/7026 (0.1%)	2.08	241/9534 (2.5%)

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	18
1	B	0	15
1	C	0	19
All	All	0	52

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	332	PHE	N-CA	-10.36	1.25	1.46
1	C	330	CYS	CB-SG	-9.36	1.66	1.82
1	C	247	TRP	NE1-CE2	-6.76	1.28	1.37
1	C	295	SER	CB-OG	6.32	1.50	1.42
1	C	419	SER	CB-OG	5.82	1.49	1.42

The worst 5 of 241 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	321	ARG	NE-CZ-NH1	-20.80	109.90	120.30
1	B	531	ARG	NE-CZ-NH2	18.95	129.78	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	521	ARG	NE-CZ-NH1	-17.89	111.35	120.30
1	A	321	ARG	NE-CZ-NH2	17.59	129.10	120.30
1	A	458	ARG	NE-CZ-NH1	-16.57	112.02	120.30

There are no chirality outliers.

5 of 52 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	240	MET	Mainchain
1	A	242	ARG	Mainchain
1	A	243	ASN	Mainchain
1	A	273	LEU	Mainchain
1	A	294	TYR	Mainchain

## 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	2295	0	2274	74	0
1	B	2295	0	2273	66	0
1	C	2295	0	2274	60	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	0	0
3	A	121	0	0	18	0
3	B	134	0	0	16	1
3	C	159	0	0	8	1
All	All	7329	0	6821	197	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 14.

The worst 5 of 197 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:A:441:PRO:HB2	1:A:442:PRO:HD3	1.47	0.94

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:441:PRO:HB2	1:C:442:PRO:HD3	1.51	0.91
1:C:441:PRO:HB2	1:C:442:PRO:CD	2.04	0.86
1:C:538:SER:O	1:C:539:TYR:HB3	1.74	0.85
1:A:263:VAL:HA	3:A:701:HOH:O	1.77	0.84

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)
3:B:723:HOH:O	3:C:662:HOH:O[2_657]	2.13	0.07

## 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	282/311 (91%)	260 (92%)	18 (6%)	4 (1%)	14	4
1	B	282/311 (91%)	264 (94%)	13 (5%)	5 (2%)	11	2
1	C	282/311 (91%)	260 (92%)	16 (6%)	6 (2%)	9	2
All	All	846/933 (91%)	784 (93%)	47 (6%)	15 (2%)	11	2

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	401	GLY
1	A	437	PRO
1	B	401	GLY
1	B	437	PRO
1	C	401	GLY

### 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	
1	A	259/291 (89%)	232 (90%)	27 (10%)	9	3
1	B	259/291 (89%)	230 (89%)	29 (11%)	7	2
1	C	259/291 (89%)	232 (90%)	27 (10%)	9	3
All	All	777/873 (89%)	694 (89%)	83 (11%)	8	2

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	307	GLU
1	B	423	SER
1	C	487	SER
1	B	334	GLU
1	B	363	GLU

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	342	ASN
1	B	364	ASN
1	C	342	ASN
1	A	511	ASN
1	B	243	ASN

### 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

6 ligands 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	$\# Z  > 2$
2	SO4	A	601	-	4,4,4	0.98	0	6,6,6	0.28	0
2	SO4	A	604	-	4,4,4	1.04	0	6,6,6	1.28	1 (16%)
2	SO4	B	602	-	4,4,4	0.93	0	6,6,6	0.22	0
2	SO4	B	605	-	4,4,4	1.29	1 (25%)	6,6,6	1.57	1 (16%)
2	SO4	C	603	-	4,4,4	0.91	0	6,6,6	0.45	0
2	SO4	C	606	-	4,4,4	0.62	0	6,6,6	0.49	0

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
2	SO4	A	601	-	-	0/0/0/0	0/0/0/0
2	SO4	A	604	-	-	0/0/0/0	0/0/0/0
2	SO4	B	602	-	-	0/0/0/0	0/0/0/0
2	SO4	B	605	-	-	0/0/0/0	0/0/0/0
2	SO4	C	603	-	-	0/0/0/0	0/0/0/0
2	SO4	C	606	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	605	SO4	O1-S	2.45	1.55	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	605	SO4	O2-S-O1	-3.55	98.24	109.50
2	A	604	SO4	O2-S-O1	2.73	118.14	109.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.