WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 27, 2024 – 07:59 PM EST

PDB ID	:	1D8H
Title	:	X-RAY CRYSTAL STRUCTURE OF YEAST RNA TRIPHOSPHATASE IN COMPLEX WITH SULFATE AND MANGANESE IONS.
Authors	:	Lima, C.D.; Wang, L.K.; Shuman, S.
Deposited on	:	1999-10-24
Resolution	:	2.00 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.00 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	Percentile Ranks	Value
Clashscore		26
Ramachandran outliers		3.2%
Sidechain outliers		13.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
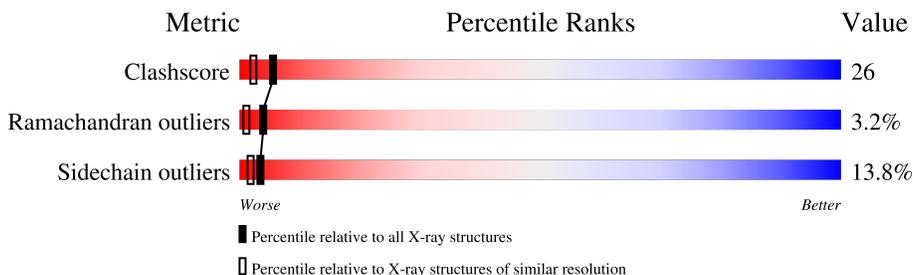
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	311	27% 39% 21% 6% 7%
1	B	311	27% 36% 23% 6% 7%
1	C	311	23% 39% 22% 8% 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	C	603	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 7626 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
			2313	1470	395	443	5			
1	B	288	Total	C	N	O	S	0	0	0
			2313	1470	395	443	5			
1	C	288	Total	C	N	O	S	0	0	0
			2313	1470	395	443	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 MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mn	0	0
			1	1		
2	B	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2
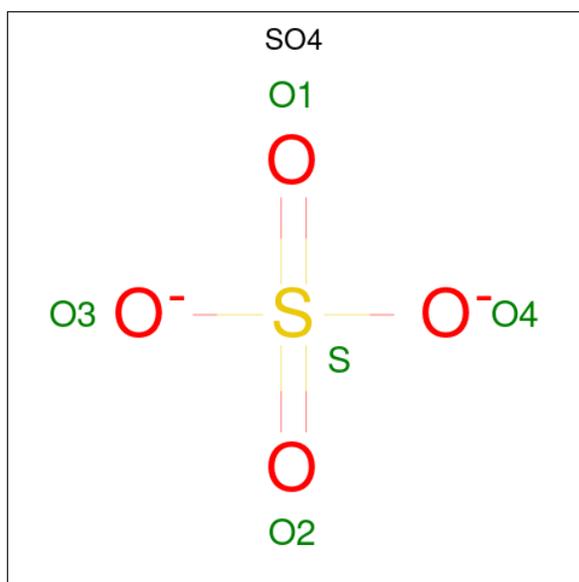
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0
3	C	1	Total 5	O 4	S 1	0	0

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	222	Total 222	O 222	0	0
4	B	226	Total 226	O 226	0	0
4	C	221	Total 221	O 221	0	0

3 Residue-property plots ⓘ

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. 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.

Note EDS was not executed.

• Molecule 1: mRNA TRIPHOSPHATASE CET1
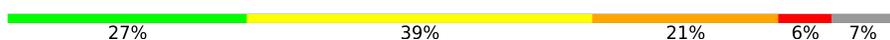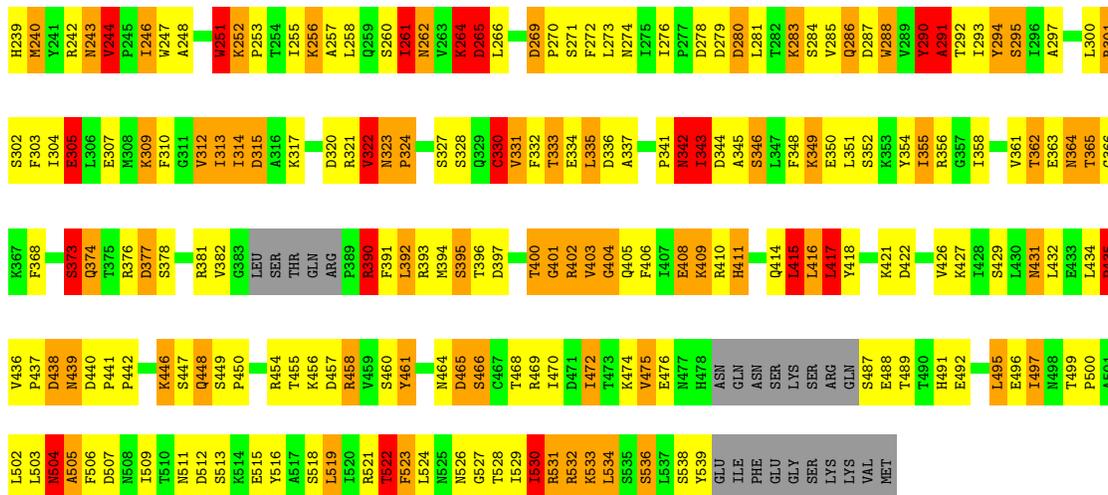
• Molecule 1: mRNA TRIPHOSPHATASE CET1
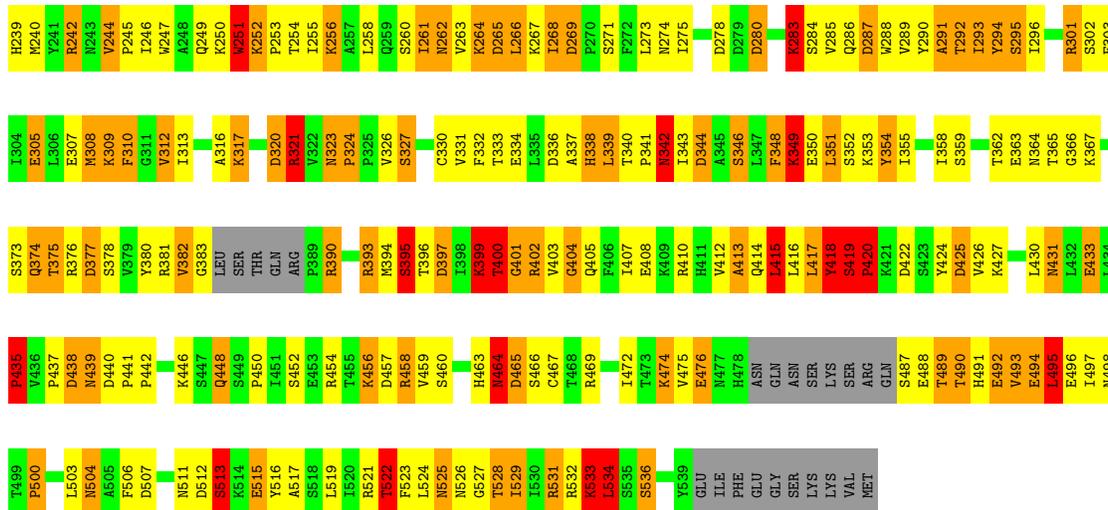
• Molecule 1: mRNA TRIPHOSPHATASE CET1

Chain C:
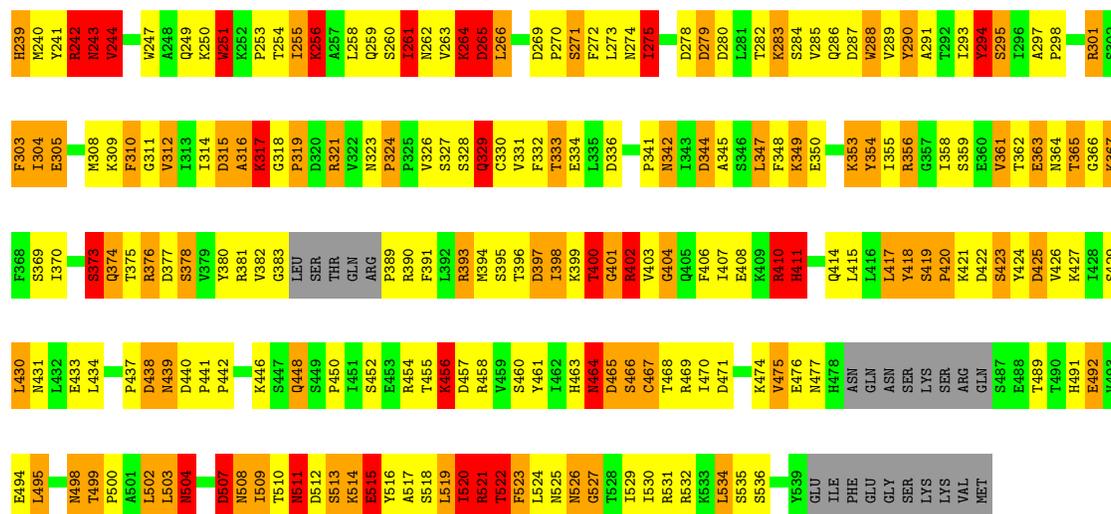
4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	116.44Å 118.72Å 85.14Å 90.00° 108.06° 90.00°	Depositor
Resolution (Å)	25.00 – 2.00	Depositor
% Data completeness (in resolution range)	88.1 (25.00-2.00)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.229 , 0.310	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7626	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

5   Model quality (i)

5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MN, 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	1.66	28/2360 (1.2%)	3.02	221/3199 (6.9%)
1	B	1.60	28/2360 (1.2%)	2.95	237/3199 (7.4%)
1	C	1.57	24/2360 (1.0%)	3.07	233/3199 (7.3%)
All	All	1.61	80/7080 (1.1%)	3.02	691/9597 (7.2%)

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	A	1	29
1	B	0	26
1	C	0	29
All	All	1	84

The worst 5 of 80 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	460	SER	CA-CB	10.79	1.69	1.52
1	A	321	ARG	CZ-NH1	10.26	1.46	1.33
1	C	327	SER	CA-CB	9.87	1.67	1.52
1	B	288	TRP	CD2-CE2	-9.27	1.30	1.41
1	A	302	SER	CA-CB	8.84	1.66	1.52

The worst 5 of 691 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	301	ARG	NE-CZ-NH1	43.95	142.27	120.30
1	A	532	ARG	NE-CZ-NH1	29.98	135.29	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	531	ARG	NE-CZ-NH2	27.91	134.26	120.30
1	A	532	ARG	NE-CZ-NH2	-25.79	107.41	120.30
1	A	242	ARG	NE-CZ-NH2	-25.72	107.44	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	522	THR	CB

5 of 84 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	255	ILE	Mainchain
1	A	269	ASP	Mainchain
1	A	276	ILE	Mainchain
1	A	280	ASP	Mainchain
1	A	286	GLN	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	2313	0	2318	112	0
1	B	2313	0	2315	128	0
1	C	2313	0	2316	129	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	C	5	0	0	2	0
4	A	222	0	0	24	2
4	B	226	0	0	47	3
4	C	221	0	0	36	1
All	All	7626	0	6949	359	4

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 26.

The worst 5 of 359 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:239:HIS:N	4:C:675:HOH:O	1.91	1.02
1:C:403:VAL:HG21	1:C:446:LYS:HE2	1.48	0.95
1:C:264:LYS:O	1:C:266:LEU:HG	1.73	0.89
1:C:502:LEU:HA	1:C:519:LEU:HD21	1.57	0.87
1:C:256:LYS:HB2	4:C:793:HOH:O	1.75	0.87

All (4) 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)
4:A:730:HOH:O	4:B:656:HOH:O[2_556]	2.06	0.14
4:A:682:HOH:O	4:C:773:HOH:O[4_557]	2.12	0.08
4:B:684:HOH:O	4:B:747:HOH:O[2_556]	2.17	0.03
4:B:678:HOH:O	4:B:747:HOH:O[2_556]	2.18	0.02

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%)	253 (90%)	19 (7%)	10 (4%)	3	1
1	B	282/311 (91%)	247 (88%)	27 (10%)	8 (3%)	5	1
1	C	282/311 (91%)	252 (89%)	21 (7%)	9 (3%)	4	1
All	All	846/933 (91%)	752 (89%)	67 (8%)	27 (3%)	4	1

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	264	LYS

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	A	401	GLY
1	A	404	GLY
1	A	438	ASP
1	B	400	THR

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	265/291 (91%)	229 (86%)	36 (14%)	3	2
1	B	265/291 (91%)	236 (89%)	29 (11%)	6	3
1	C	265/291 (91%)	220 (83%)	45 (17%)	2	1
All	All	795/873 (91%)	685 (86%)	110 (14%)	3	2

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	456	LYS
1	C	261	ILE
1	C	534	LEU
1	C	474	LYS
1	B	476	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	448	GLN
1	C	364	ASN
1	B	477	ASN
1	C	448	GLN
1	C	338	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 for Mogul analysis.

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	#$|Z| > 2$
3	SO4	A	601	-	4,4,4	0.77	0	6,6,6	2.67	2 (33%)
3	SO4	B	602	-	4,4,4	0.80	0	6,6,6	2.05	3 (50%)
3	SO4	C	603	-	4,4,4	0.79	0	6,6,6	1.52	1 (16%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	SO4	O4-S-O1	5.41	137.56	109.31
3	C	603	SO4	O3-S-O1	3.21	126.08	109.31
3	B	602	SO4	O4-S-O1	3.14	125.70	109.31
3	A	601	SO4	O4-S-O3	-2.92	96.61	109.06
3	B	602	SO4	O4-S-O3	-2.59	97.99	109.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	603	SO4	2	0

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.