



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2023 – 04:01 AM EDT

PDB ID : 1V4N  
Title : Structure of 5'-deoxy-5'-methylthioadenosine phosphorylase homologue from *Sulfolobus tokodaii*  
Authors : Kitago, Y.; Yasutake, Y.; Sakai, N.; Tsujimura, M.; Yao, M.; Watanabe, N.; Kawarabayasi, Y.; Tanaka, I.  
Deposited on : 2003-11-14  
Resolution : 2.45 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

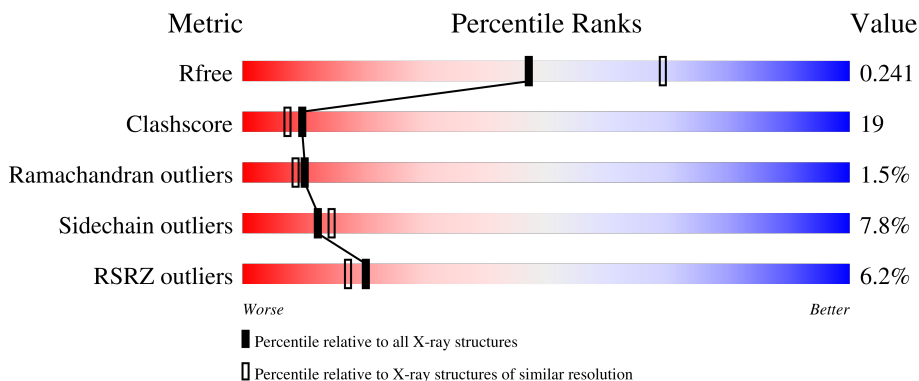
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	A	281	 5% 62% 26% 6% 5%
1	B	281	 8% 66% 25% 5% 5%
1	C	281	 5% 65% 23% 6% 5%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6539 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 271aa long hypothetical 5'-methylthioadenosine phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	266	2077	1333	350	381	13	0	0	0
1	B	269	2102	1349	354	386	13	0	0	0
1	C	267	2085	1338	352	382	13	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	272	LEU	-	expression tag	GB 15920700
A	273	GLU	-	expression tag	GB 15920700
A	274	HIS	-	expression tag	GB 15920700
A	275	HIS	-	expression tag	GB 15920700
A	276	HIS	-	expression tag	GB 15920700
A	277	HIS	-	expression tag	GB 15920700
A	278	HIS	-	expression tag	GB 15920700
A	279	HIS	-	expression tag	GB 15920700
A	280	HIS	-	expression tag	GB 15920700
A	281	HIS	-	expression tag	GB 15920700
B	272	LEU	-	expression tag	GB 15920700
B	273	GLU	-	expression tag	GB 15920700
B	274	HIS	-	expression tag	GB 15920700
B	275	HIS	-	expression tag	GB 15920700
B	276	HIS	-	expression tag	GB 15920700
B	277	HIS	-	expression tag	GB 15920700
B	278	HIS	-	expression tag	GB 15920700
B	279	HIS	-	expression tag	GB 15920700
B	280	HIS	-	expression tag	GB 15920700
B	281	HIS	-	expression tag	GB 15920700
C	272	LEU	-	expression tag	GB 15920700
C	273	GLU	-	expression tag	GB 15920700
C	274	HIS	-	expression tag	GB 15920700

*Continued on next page...*

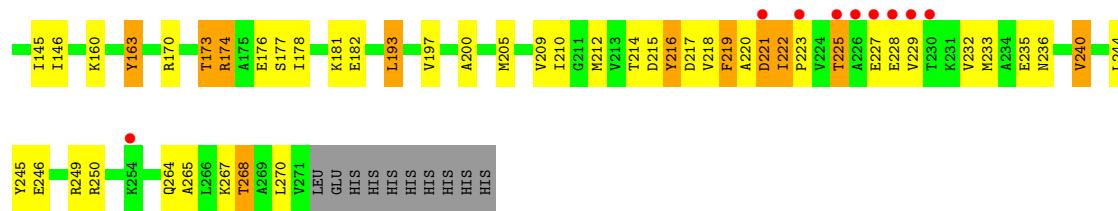
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	275	HIS	-	expression tag	GB 15920700
C	276	HIS	-	expression tag	GB 15920700
C	277	HIS	-	expression tag	GB 15920700
C	278	HIS	-	expression tag	GB 15920700
C	279	HIS	-	expression tag	GB 15920700
C	280	HIS	-	expression tag	GB 15920700
C	281	HIS	-	expression tag	GB 15920700

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	93	Total O 93 93	0	0
2	B	87	Total O 87 87	0	0
2	C	95	Total O 95 95	0	0





## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.87Å 142.89Å 169.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.45 44.33 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-2.45) 100.0 (44.33-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.58 (at 2.39Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.190 , 0.241 0.191 , 0.241	Depositor DCC
$R_{free}$ test set	1918 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtrriage
Anisotropy	0.649	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 47.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.015 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.023 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6539	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% of the height of the origin peak. No significant pseudotranslation is detected.*

<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	A	0.99	1/2122 (0.0%)	0.99	2/2877 (0.1%)
1	B	0.97	1/2148 (0.0%)	1.01	2/2912 (0.1%)
1	C	1.00	1/2131 (0.0%)	1.04	6/2888 (0.2%)
All	All	0.99	3/6401 (0.0%)	1.01	10/8677 (0.1%)

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	B	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	39	GLU	CD-OE2	10.07	1.36	1.25
1	C	39	GLU	CD-OE2	8.22	1.34	1.25
1	A	39	GLU	CD-OE2	7.96	1.34	1.25

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	174	ARG	NE-CZ-NH1	-7.81	116.39	120.30
1	C	174	ARG	NE-CZ-NH1	-7.19	116.70	120.30
1	C	26	THR	C-N-CA	-6.82	104.64	121.70
1	B	117	LYS	C-N-CA	-6.25	109.17	122.30
1	C	26	THR	N-CA-C	6.24	127.85	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	B	121	TYR	Sidechain
1	B	219	PHE	Sidechain

## 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	A	2077	0	2111	92	0
1	B	2102	0	2137	86	0
1	C	2085	0	2121	80	0
2	A	93	0	0	8	0
2	B	87	0	0	7	0
2	C	95	0	0	5	0
All	All	6539	0	6369	242	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 19.

The worst 5 of 242 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:225:THR:HG22	1:C:228:GLU:H	1.20	1.02
1:B:173:THR:HG22	1:B:176:GLU:H	1.24	1.00
1:C:214:THR:CG2	1:C:229:VAL:HA	1.94	0.98
1:A:214:THR:CG2	1:A:229:VAL:HA	1.93	0.97
1:A:214:THR:HG21	1:A:229:VAL:HA	1.46	0.96

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	264/281 (94%)	252 (96%)	7 (3%)	5 (2%)	8	6
1	B	267/281 (95%)	261 (98%)	3 (1%)	3 (1%)	14	14
1	C	265/281 (94%)	256 (97%)	5 (2%)	4 (2%)	10	9
All	All	796/843 (94%)	769 (97%)	15 (2%)	12 (2%)	10	9

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	217	ASP
1	A	219	PHE
1	B	218	VAL
1	C	22	PRO
1	A	223	PRO

### 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	228/243 (94%)	209 (92%)	19 (8%)	11	12
1	B	231/243 (95%)	211 (91%)	20 (9%)	10	11
1	C	229/243 (94%)	214 (93%)	15 (7%)	16	20
All	All	688/729 (94%)	634 (92%)	54 (8%)	12	15

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	141	HIS
1	B	218	VAL
1	C	219	PHE
1	B	160	LYS
1	B	178	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	264	GLN
1	C	23	GLN
1	B	141	HIS
1	B	27	ASN
1	B	264	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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 monosaccharides 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(Å <sup>2</sup> )	Q<0.9
1	A	266/281 (94%)	-0.02	15 (5%) 24 21	16, 27, 67, 85	0
1	B	269/281 (95%)	0.07	22 (8%) 11 8	16, 28, 68, 93	0
1	C	267/281 (95%)	-0.03	13 (4%) 29 27	16, 28, 64, 85	0
All	All	802/843 (95%)	0.01	50 (6%) 20 17	16, 28, 67, 93	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	24	ILE	7.3
1	B	224	VAL	5.1
1	A	225	THR	5.0
1	B	229	VAL	4.9
1	B	230	THR	4.7

### 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 monosaccharides 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.