



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 04:44 am GMT

PDB ID : 3K5B  
Title : Crystal structure of the peripheral stalk of *Thermus thermophilus* H<sup>+</sup>-ATPase/synthase  
Authors : Lee, L.K.; Stewart, A.G.; Donohoe, M.; Bernal, R.A.; Stock, D.  
Deposited on : 2009-10-07  
Resolution : 3.10 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 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	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

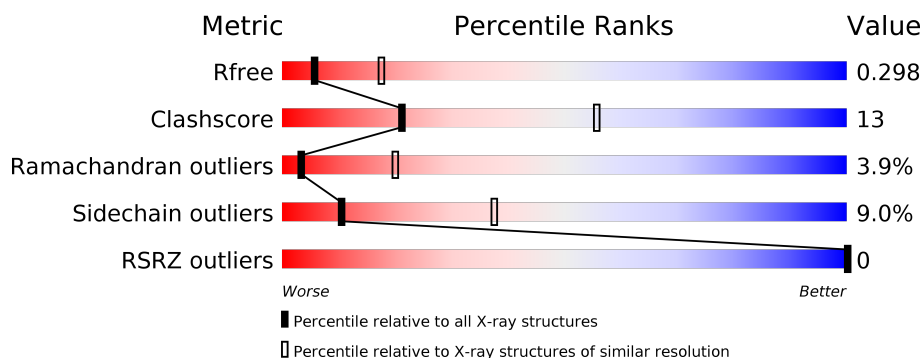
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	111664	1115 (3.12-3.08)
Clashscore	122126	1042 (3.10-3.10)
Ramachandran outliers	120053	1010 (3.10-3.10)
Sidechain outliers	120020	1010 (3.10-3.10)
RSRZ outliers	108989	1089 (3.12-3.08)

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.

Mol	Chain	Length	Quality of chain
1	B	104	
1	G	104	
2	A	188	
2	E	188	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4056 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 V-type ATP synthase, subunit (VAPC-THERM).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	100	Total	C	N	O	Se	0	0	0
			747	460	136	150	1			
1	B	100	Total	C	N	O	Se	0	0	0
			720	442	134	143	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	17	GLY	-	EXPRESSION TAG	UNP Q5SIT5
B	17	GLY	-	EXPRESSION TAG	UNP Q5SIT5

- Molecule 2 is a protein called V-type ATP synthase subunit E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	184	Total	C	N	O	Se	0	0	0
			1312	815	242	252	3			
2	A	185	Total	C	N	O	Se	0	0	0
			1277	788	240	246	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	134	MSE	LEU	ENGINEERED	UNP P74901
E	171	MSE	LEU	ENGINEERED	UNP P74901
E	178	MSE	LEU	ENGINEERED	UNP P74901
A	134	MSE	LEU	ENGINEERED	UNP P74901
A	171	MSE	LEU	ENGINEERED	UNP P74901
A	178	MSE	LEU	ENGINEERED	UNP P74901

### 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 the various outlier classes 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: V-type ATP synthase, subunit (VAPC-THERM)

Chain G: 



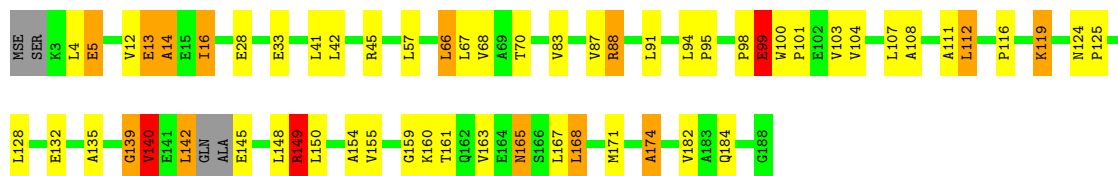
- Molecule 1: V-type ATP synthase, subunit (VAPC-THERM)

Chain B: 




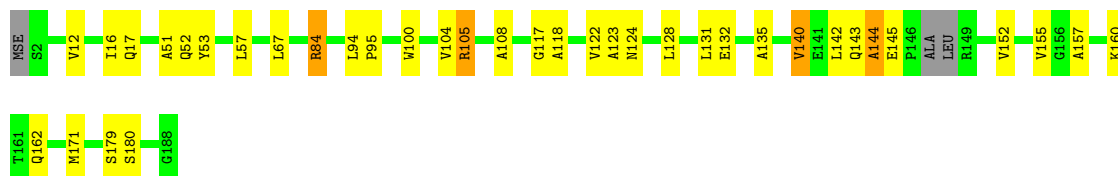
- Molecule 2: V-type ATP synthase subunit E

Chain E: 



- Molecule 2: V-type ATP synthase subunit E

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.78Å 79.75Å 75.80Å 90.00° 97.64° 90.00°	Depositor
Resolution (Å)	29.66 – 3.10 29.65 – 3.10	Depositor EDS
% Data completeness (in resolution range)	94.4 (29.66-3.10) 98.4 (29.65-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 3.11Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, $R_{free}$	0.232 , 0.288 0.254 , 0.298	Depositor DCC
$R_{free}$ test set	859 reflections (5.31%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	106.8	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 57.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.139 for l,-k,h	Xtriage
Reported twinning fraction	0.536 for H, K, L 0.464 for L, -K, H	Depositor
Outliers	1 of 16171 reflections (0.006%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4056	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.80% 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	B	0.32	0/720	0.50	0/968
1	G	0.34	0/748	0.48	0/1001
2	A	0.32	0/1287	0.49	0/1748
2	E	0.34	0/1322	0.53	0/1789
All	All	0.33	0/4077	0.50	0/5506

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 [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	B	720	0	678	17	0
1	G	747	0	726	21	0
2	A	1277	0	1161	29	0
2	E	1312	0	1244	51	0
All	All	4056	0	3809	102	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 13.

The worst 5 of 102 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)
2:E:87:VAL:HG12	2:E:171:MSE:HE2	1.20	1.09
1:G:110:ALA:HB1	2:E:182:VAL:HG23	1.57	0.85
2:E:149:ARG:O	2:E:150:LEU:HD13	1.84	0.76
2:E:41:LEU:O	2:E:41:LEU:HD13	1.86	0.76
2:E:88:ARG:HA	2:E:171:MSE:HE1	1.69	0.73

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	B	98/104 (94%)	91 (93%)	5 (5%)	2 (2%)	8	35
1	G	98/104 (94%)	90 (92%)	6 (6%)	2 (2%)	8	35
2	A	181/188 (96%)	151 (83%)	22 (12%)	8 (4%)	3	17
2	E	180/188 (96%)	156 (87%)	14 (8%)	10 (6%)	2	12
All	All	557/584 (95%)	488 (88%)	47 (8%)	22 (4%)	3	20

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	22	LEU
2	E	160	LYS
1	B	23	ILE
2	A	140	VAL
2	E	5	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	B	55/75 (73%)	48 (87%)	7 (13%)	5	20
1	G	62/75 (83%)	59 (95%)	3 (5%)	28	63
2	A	97/137 (71%)	92 (95%)	5 (5%)	25	59
2	E	108/137 (79%)	94 (87%)	14 (13%)	4	19
All	All	322/424 (76%)	293 (91%)	29 (9%)	10	37

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	149	ARG
2	E	184	GLN
2	A	84	ARG
2	E	165	ASN
1	B	30	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 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 [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	B	99/104 (95%)	-0.57	0 100 100	52, 63, 81, 92	0
1	G	99/104 (95%)	-0.51	0 100 100	45, 56, 73, 86	2 (2%)
2	A	182/188 (96%)	-0.42	0 100 100	49, 64, 76, 77	1 (0%)
2	E	181/188 (96%)	-0.42	0 100 100	36, 58, 74, 79	4 (2%)
All	All	561/584 (96%)	-0.47	0 100 100	36, 60, 76, 92	7 (1%)

There are no RSRZ outliers to report.

### 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 carbohydrates 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.