



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 07:16 pm BST

PDB ID : 4WWV  
Title : Aminopeptidase APDkam598 from the archaeon *Desulfurococcus kamchatkensis*  
Authors : Petrova, T.; Boyko, K.M.; Rakitina, T.V.; Korzhenevskiy, D.A.; Gorbacheva, M.A.; Popov, V.O.  
Deposited on : 2014-11-12  
Resolution : 3.01 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 : 2.11  
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) : 2.11

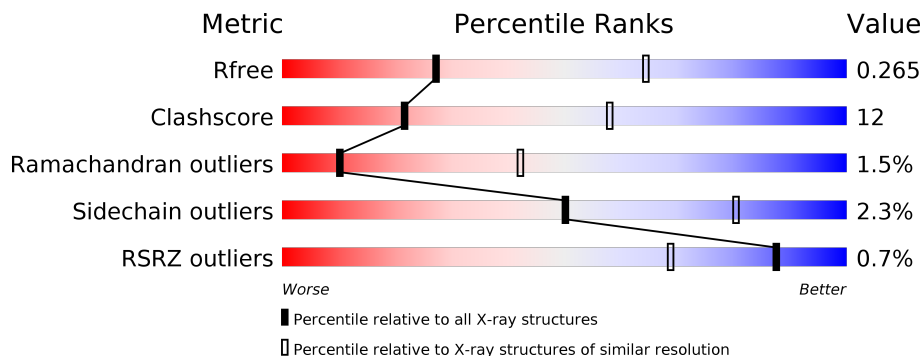
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.01 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 on 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	394	
1	B	394	

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 5505 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 Aminopeptidase from family M42.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	362	2743	1756	473	508	6	0	0	1
1	B	361	2736	1753	474	503	6	0	0	1

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-25	MET	-	initiating methionine	UNP B8D484
A	-24	GLY	-	expression tag	UNP B8D484
A	-23	SER	-	expression tag	UNP B8D484
A	-22	SER	-	expression tag	UNP B8D484
A	-21	HIS	-	expression tag	UNP B8D484
A	-20	HIS	-	expression tag	UNP B8D484
A	-19	HIS	-	expression tag	UNP B8D484
A	-18	HIS	-	expression tag	UNP B8D484
A	-17	HIS	-	expression tag	UNP B8D484
A	-16	HIS	-	expression tag	UNP B8D484
A	-15	SER	-	expression tag	UNP B8D484
A	-14	SER	-	expression tag	UNP B8D484
A	-13	GLY	-	expression tag	UNP B8D484
A	-12	LEU	-	expression tag	UNP B8D484
A	-11	VAL	-	expression tag	UNP B8D484
A	-10	PRO	-	expression tag	UNP B8D484
A	-9	ARG	-	expression tag	UNP B8D484
A	-8	GLY	-	expression tag	UNP B8D484
A	-7	SER	-	expression tag	UNP B8D484
A	-6	HIS	-	expression tag	UNP B8D484
A	-5	MET	-	expression tag	UNP B8D484
A	-4	LEU	-	expression tag	UNP B8D484
A	-3	GLU	-	expression tag	UNP B8D484
A	-2	ASP	-	expression tag	UNP B8D484
A	-1	PRO	-	expression tag	UNP B8D484

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	VAL	MET	engineered mutation	UNP B8D484
A	180	SER	GLY	conflict	UNP B8D484
A	208	LEU	PHE	conflict	UNP B8D484
A	280	GLY	VAL	conflict	UNP B8D484
B	-25	MET	-	initiating methionine	UNP B8D484
B	-24	GLY	-	expression tag	UNP B8D484
B	-23	SER	-	expression tag	UNP B8D484
B	-22	SER	-	expression tag	UNP B8D484
B	-21	HIS	-	expression tag	UNP B8D484
B	-20	HIS	-	expression tag	UNP B8D484
B	-19	HIS	-	expression tag	UNP B8D484
B	-18	HIS	-	expression tag	UNP B8D484
B	-17	HIS	-	expression tag	UNP B8D484
B	-16	HIS	-	expression tag	UNP B8D484
B	-15	SER	-	expression tag	UNP B8D484
B	-14	SER	-	expression tag	UNP B8D484
B	-13	GLY	-	expression tag	UNP B8D484
B	-12	LEU	-	expression tag	UNP B8D484
B	-11	VAL	-	expression tag	UNP B8D484
B	-10	PRO	-	expression tag	UNP B8D484
B	-9	ARG	-	expression tag	UNP B8D484
B	-8	GLY	-	expression tag	UNP B8D484
B	-7	SER	-	expression tag	UNP B8D484
B	-6	HIS	-	expression tag	UNP B8D484
B	-5	MET	-	expression tag	UNP B8D484
B	-4	LEU	-	expression tag	UNP B8D484
B	-3	GLU	-	expression tag	UNP B8D484
B	-2	ASP	-	expression tag	UNP B8D484
B	-1	PRO	-	expression tag	UNP B8D484
B	0	VAL	MET	engineered mutation	UNP B8D484
B	180	SER	GLY	conflict	UNP B8D484
B	208	LEU	PHE	conflict	UNP B8D484
B	280	GLY	VAL	conflict	UNP B8D484

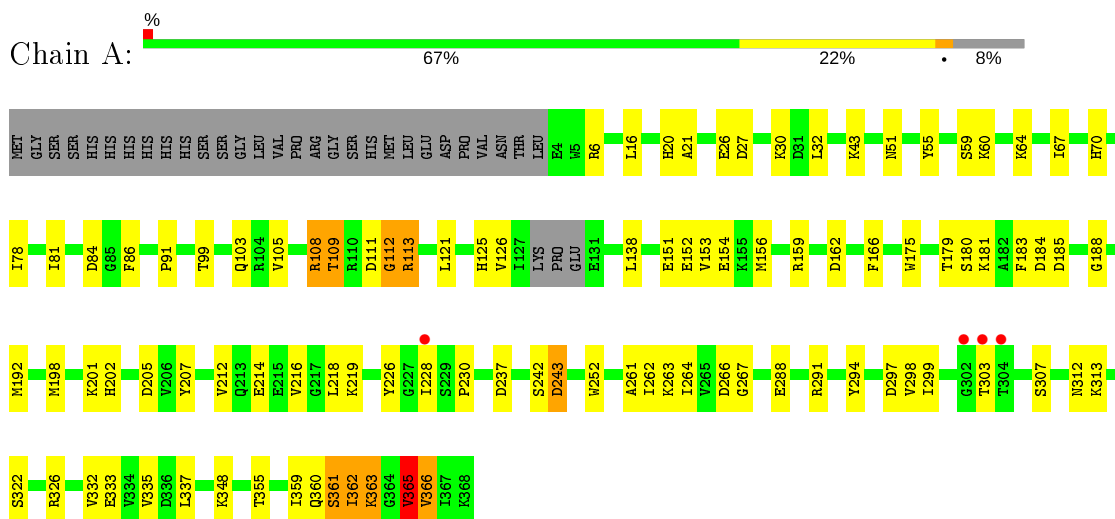
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	15	Total O 15 15	0	0
2	B	11	Total O 11 11	0	0

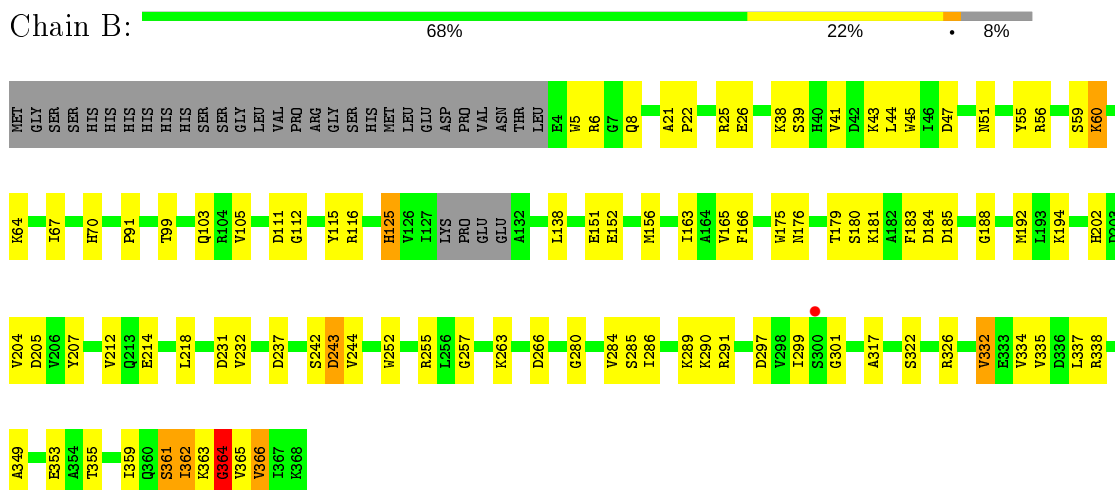
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Aminopeptidase from family M42



- Molecule 1: Aminopeptidase from family M42



## 4 Data and refinement statistics

Property	Value	Source
Space group	F 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	234.32Å 234.32Å 234.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.83 – 3.01 47.83 – 3.01	Depositor EDS
% Data completeness (in resolution range)	95.0 (47.83-3.01) 95.0 (47.83-3.01)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 3.01Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.204 , 0.263 0.207 , 0.265	Depositor DCC
$R_{free}$ test set	1044 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 11.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.477 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5505	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.56% 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	0.60	0/2793	0.70	1/3790 (0.0%)
1	B	0.60	0/2786	0.71	1/3779 (0.0%)
All	All	0.60	0/5579	0.70	2/7569 (0.0%)

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	8
1	B	0	4
All	All	0	12

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	364	GLY	N-CA-C	6.56	129.50	113.10
1	A	121	LEU	CA-CB-CG	5.08	126.99	115.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	108	ARG	Peptide
1	A	112	GLY	Peptide
1	A	228	ILE	Peptide
1	A	267	GLY	Peptide
1	A	360	GLN	Peptide

## 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	2743	0	2759	65	0
1	B	2736	0	2746	71	0
2	A	15	0	0	0	0
2	B	11	0	0	0	0
All	All	5505	0	5505	131	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 12.

The worst 5 of 131 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:21:ALA:HB1	1:A:26:GLU:HA	1.50	0.94
1:A:109:THR:HG21	1:A:113:ARG:HG3	1.47	0.93
1:B:21:ALA:HB1	1:B:26:GLU:HA	1.54	0.88
1:B:263:LYS:NZ	1:B:297:ASP:OD2	2.08	0.87
1:A:263:LYS:NZ	1:A:297:ASP:OD2	2.13	0.79

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	358/394 (91%)	333 (93%)	19 (5%)	6 (2%)	9 39
1	B	357/394 (91%)	333 (93%)	19 (5%)	5 (1%)	11 43

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	715/788 (91%)	666 (93%)	38 (5%)	11 (2%)	10	42

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	109	THR
1	A	362	ILE
1	B	365	VAL
1	A	365	VAL
1	B	243	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	281/328 (86%)	275 (98%)	6 (2%)	53	82
1	B	277/328 (84%)	270 (98%)	7 (2%)	47	79
All	All	558/656 (85%)	545 (98%)	13 (2%)	50	80

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	335	VAL
1	B	105	VAL
1	B	204	VAL
1	A	332	VAL
1	B	179	THR

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	A	362/394 (91%)	0.08	4 (1%) 80 56	22, 33, 63, 78	1 (0%)
1	B	361/394 (91%)	0.05	1 (0%) 94 84	22, 32, 61, 79	1 (0%)
All	All	723/788 (91%)	0.06	5 (0%) 87 69	22, 33, 62, 79	2 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	303	THR	5.2
1	A	304	THR	2.8
1	B	300	SER	2.5
1	A	302	GLY	2.2
1	A	228	ILE	2.1

### 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.