



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 08:58 am BST

PDB ID : 5NOE  
Title : Anthranilate phosphoribosyltransferase from *Thermococcus kodakaraensis*  
Authors : Perveen, S.; Rashid, N.; Papageorgiou, A.C.  
Deposited on : 2017-04-12  
Resolution : 1.91 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

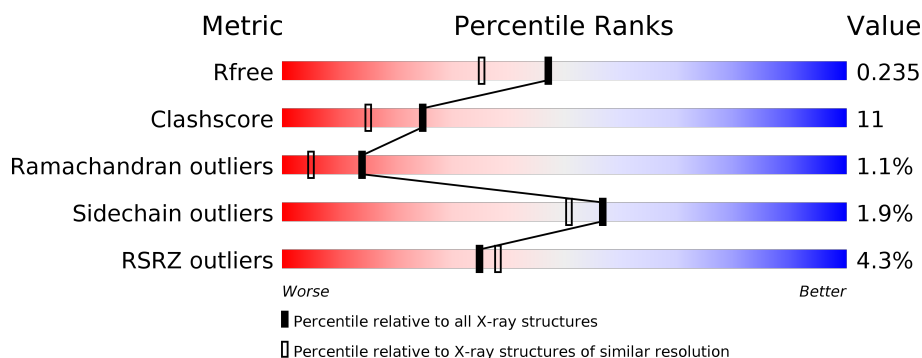
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.91 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	325	<div> <div>4%</div> <div> <div></div> <div>87%</div> <div>12%</div> <div>.</div> </div> </div>
1	B	325	<div> <div>8%</div> <div> <div></div> <div>73%</div> <div>24%</div> <div>..</div> </div> </div>
1	C	325	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>8%</div> <div>..</div> </div> </div>
1	D	325	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>8%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10610 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 Anthranilate phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2405	1513	412	471	9			
1	B	325	Total	C	N	O	S	0	1	0
			2393	1507	408	469	9			
1	C	325	Total	C	N	O	S	0	2	0
			2419	1523	413	474	9			
1	D	324	Total	C	N	O	S	0	2	0
			2422	1524	418	471	9			

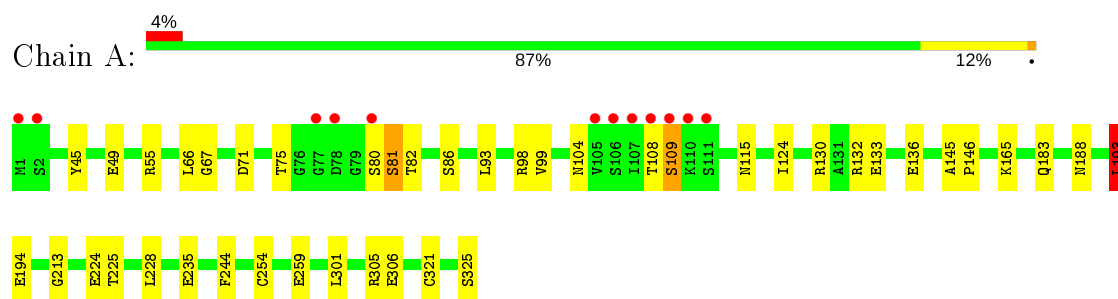
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	203	Total	O	0	0
			203	203		
2	B	133	Total	O	0	0
			133	133		
2	C	318	Total	O	0	0
			318	318		
2	D	317	Total	O	0	0
			317	317		

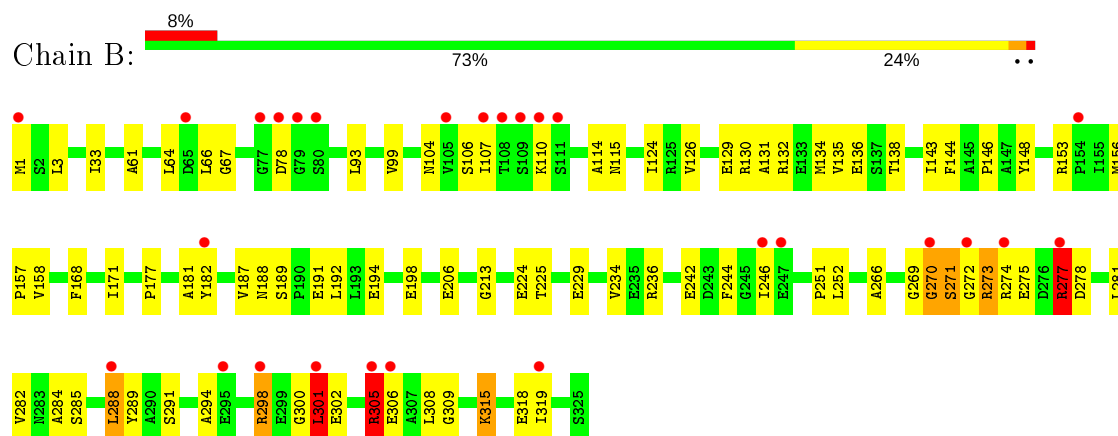
### 3 Residue-property plots

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.

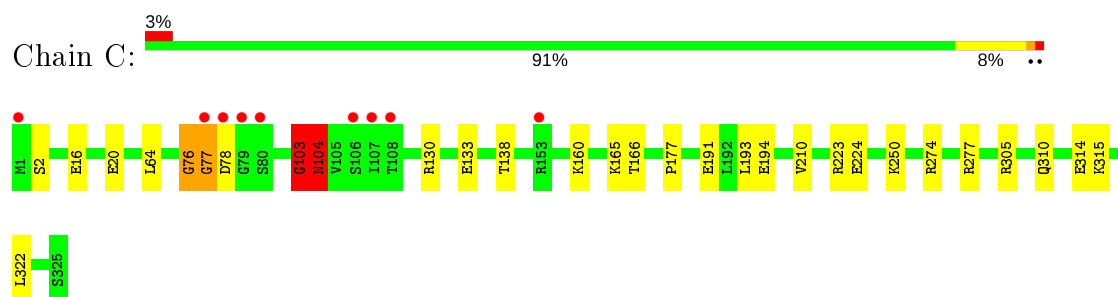
#### • Molecule 1: Anthranilate phosphoribosyltransferase



#### • Molecule 1: Anthranilate phosphoribosyltransferase



#### • Molecule 1: Anthranilate phosphoribosyltransferase



#### • Molecule 1: Anthranilate phosphoribosyltransferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.87Å 85.61Å 180.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.27 – 1.91 49.27 – 1.91	Depositor EDS
% Data completeness (in resolution range)	98.6 (49.27-1.91) 98.6 (49.27-1.91)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 1.91Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, $R_{free}$	0.187 , 0.236 0.187 , 0.235	Depositor DCC
$R_{free}$ test set	4931 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.4	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 50.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.019 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10610	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 19.95 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.6821e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

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	0.44	0/2442	0.62	2/3308 (0.1%)
1	B	0.85	10/2433 (0.4%)	1.09	14/3298 (0.4%)
1	C	0.62	0/2462	0.71	1/3334 (0.0%)
1	D	0.53	0/2465	0.68	1/3336 (0.0%)
All	All	0.63	10/9802 (0.1%)	0.80	18/13276 (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	1
1	C	0	3
All	All	0	4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	305	ARG	NE-CZ	-17.61	1.10	1.33
1	B	305	ARG	CZ-NH2	-14.68	1.14	1.33
1	B	305	ARG	CD-NE	-11.78	1.26	1.46
1	B	305	ARG	CB-CG	-10.22	1.25	1.52
1	B	301	LEU	CA-CB	8.84	1.74	1.53

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	305	ARG	NE-CZ-NH1	-23.73	108.43	120.30
1	B	277	ARG	NE-CZ-NH1	-19.30	110.65	120.30
1	B	301	LEU	CB-CG-CD1	18.01	141.62	111.00
1	B	277	ARG	NE-CZ-NH2	14.09	127.34	120.30
1	B	305	ARG	CG-CD-NE	-13.62	83.20	111.80

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	67	GLY	Peptide
1	C	103	GLY	Peptide
1	C	76	GLY	Peptide
1	C	77	GLY	Peptide

## 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	2405	0	2424	31	0
1	B	2393	0	2406	127	0
1	C	2419	0	2451	32	0
1	D	2422	0	2458	20	0
2	A	203	0	0	10	1
2	B	133	0	0	7	0
2	C	318	0	0	19	2
2	D	317	0	0	11	1
All	All	10610	0	9739	209	3

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 11.

The worst 5 of 209 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:B:301:LEU:CB	1:B:301:LEU:CG	1.75	1.57
1:B:301:LEU:CD1	1:B:301:LEU:CG	1.84	1.55
1:B:301:LEU:CB	1:B:301:LEU:HG	1.46	1.18
1:B:288:LEU:HD11	1:B:300:GLY:HA2	1.23	1.09
1:B:244:PHE:HB3	1:B:301:LEU:HD23	1.36	1.06

All (3) 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)
2:C:707:HOH:O	2:C:711:HOH:O 4_556]	2.09	0.11
2:A:593:HOH:O	2:D:672:HOH:O 3_545]	2.14	0.06
2:C:640:HOH:O	2:C:707:HOH:O 4_456]	2.18	0.02

## 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	323/325 (99%)	310 (96%)	11 (3%)	2 (1%)	25	14
1	B	324/325 (100%)	301 (93%)	16 (5%)	7 (2%)	6	1
1	C	325/325 (100%)	315 (97%)	9 (3%)	1 (0%)	41	31
1	D	324/325 (100%)	313 (97%)	7 (2%)	4 (1%)	13	4
All	All	1296/1300 (100%)	1239 (96%)	43 (3%)	14 (1%)	14	5

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	104	ASN
1	D	75	THR
1	B	270	GLY
1	B	277	ARG
1	B	308	LEU

### 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	253/255 (99%)	247 (98%)	6 (2%)	49	41
1	B	251/255 (98%)	242 (96%)	9 (4%)	35	24
1	C	257/255 (101%)	255 (99%)	2 (1%)	81	81
1	D	256/255 (100%)	254 (99%)	2 (1%)	81	81
All	All	1017/1020 (100%)	998 (98%)	19 (2%)	57	51

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	126	VAL
1	B	224	GLU
1	C	104	ASN
1	B	3	LEU
1	C	322	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	ASN
1	B	104	ASN
1	C	310	GLN

### 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 ⓘ

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	325/325 (100%)	-0.03	12 (3%) 41 44	23, 36, 54, 86	0
1	B	325/325 (100%)	0.49	27 (8%) 11 13	23, 44, 72, 96	0
1	C	325/325 (100%)	-0.12	9 (2%) 53 56	14, 25, 40, 80	0
1	D	324/325 (99%)	-0.14	8 (2%) 57 60	17, 26, 48, 95	0
All	All	1299/1300 (99%)	0.05	56 (4%) 35 38	14, 33, 63, 96	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	108	THR	8.4
1	B	108	THR	8.2
1	C	79	GLY	7.4
1	B	77	GLY	7.1
1	C	77	GLY	6.9

### 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.