



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:21 pm BST

PDB ID : 5KHN  
Title : Crystal structures of the Burkholderia multivorans hopanoid transporter HpnN  
Authors : Su, C.-C.; Yu, E.W.  
Deposited on : 2016-06-15  
Resolution : 3.44 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

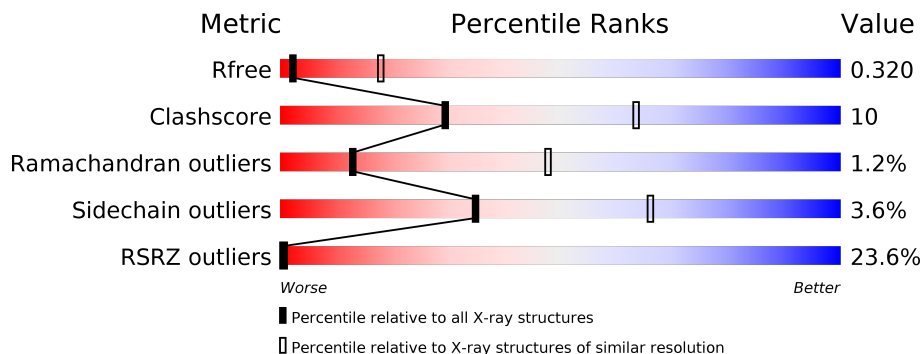
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.44 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	1278 (3.50-3.38)
Clashscore	141614	1361 (3.50-3.38)
Ramachandran outliers	138981	1327 (3.50-3.38)
Sidechain outliers	138945	1328 (3.50-3.38)
RSRZ outliers	127900	1192 (3.50-3.38)

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	883	
1	B	883	

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 12612 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 RND transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	848	6303	4069	1080	1136	18	0	0	0
1	A	848	6309	4073	1081	1137	18	0	0	0

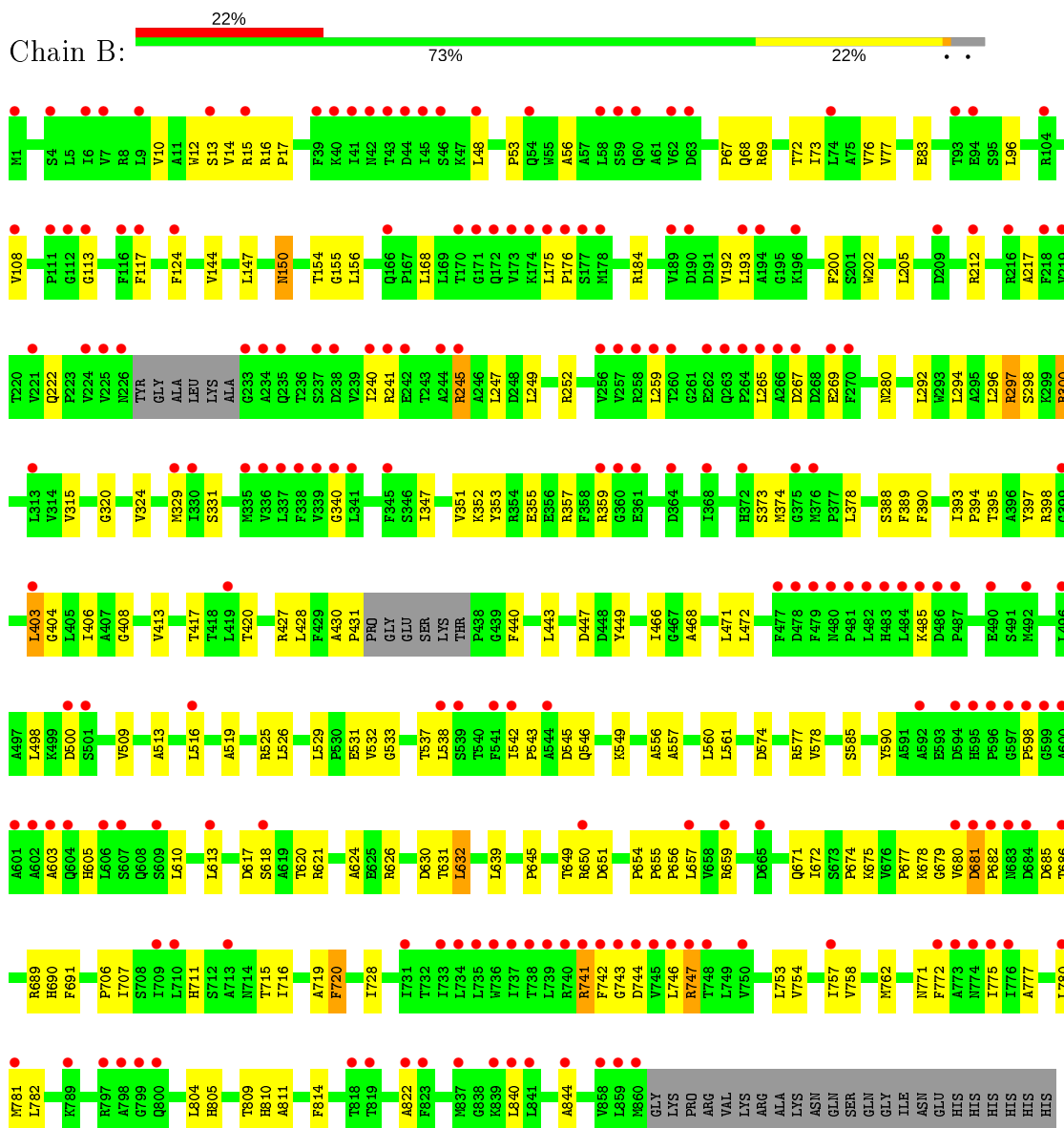
There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	360	GLY	ASP	engineered mutation	UNP A0A1B4TSD3
B	362	ALA	ARG	engineered mutation	UNP A0A1B4TSD3
B	365	ALA	HIS	engineered mutation	UNP A0A1B4TSD3
B	878	HIS	-	expression tag	UNP A0A1B4TSD3
B	879	HIS	-	expression tag	UNP A0A1B4TSD3
B	880	HIS	-	expression tag	UNP A0A1B4TSD3
B	881	HIS	-	expression tag	UNP A0A1B4TSD3
B	882	HIS	-	expression tag	UNP A0A1B4TSD3
B	883	HIS	-	expression tag	UNP A0A1B4TSD3
A	360	GLY	ASP	engineered mutation	UNP A0A1B4TSD3
A	362	ALA	ARG	engineered mutation	UNP A0A1B4TSD3
A	365	ALA	HIS	engineered mutation	UNP A0A1B4TSD3
A	878	HIS	-	expression tag	UNP A0A1B4TSD3
A	879	HIS	-	expression tag	UNP A0A1B4TSD3
A	880	HIS	-	expression tag	UNP A0A1B4TSD3
A	881	HIS	-	expression tag	UNP A0A1B4TSD3
A	882	HIS	-	expression tag	UNP A0A1B4TSD3
A	883	HIS	-	expression tag	UNP A0A1B4TSD3

### 3 Residue-property plots [i](#)

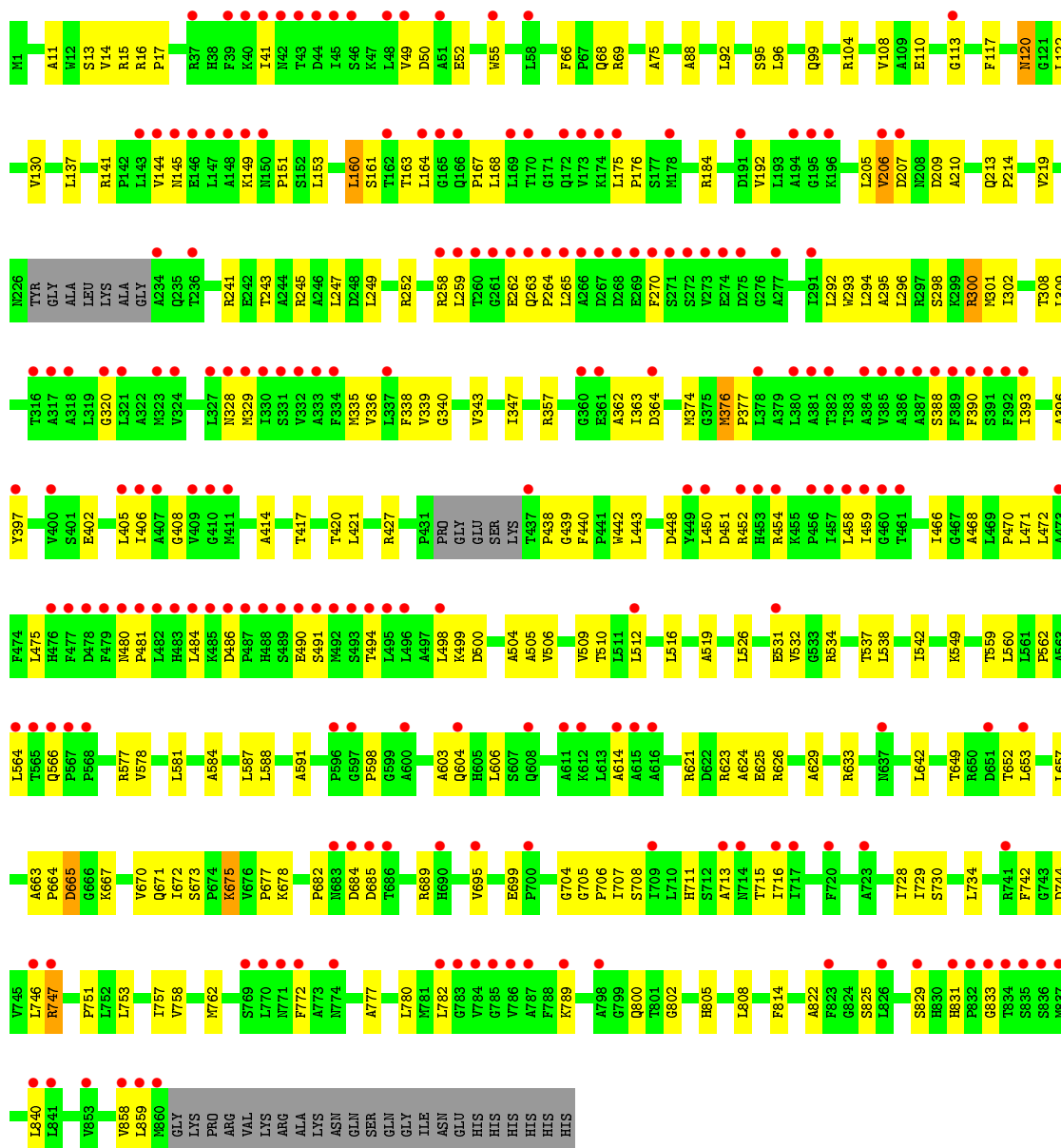
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: RND transporter



- Molecule 1: RND transporter





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.91Å 129.55Å 111.73Å 90.00° 114.04° 90.00°	Depositor
Resolution (Å)	93.80 – 3.44 93.80 – 3.44	Depositor EDS
% Data completeness (in resolution range)	99.6 (93.80-3.44) 99.7 (93.80-3.44)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 3.41Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.283 , 0.319 0.285 , 0.320	Depositor DCC
$R_{free}$ test set	1928 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	124.5	Xtrriage
Anisotropy	0.135	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 66.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.378 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	12612	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	127.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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	0.24	0/6443	0.42	0/8814
1	B	0.27	1/6437 (0.0%)	0.43	0/8804
All	All	0.25	1/12880 (0.0%)	0.43	0/17618

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	150	ASN	C-N	5.30	1.44	1.34

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	A	6309	0	6537	138	1
1	B	6303	0	6528	115	1
All	All	12612	0	13065	244	1

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 10.

The worst 5 of 244 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:685:ASP:HB2	1:B:689:ARG:HD3	1.64	0.79
1:B:83:GLU:HG3	1:B:631:THR:HG23	1.69	0.74
1:B:300:ARG:H	1:B:300:ARG:HD2	1.52	0.73
1:B:351:VAL:HG11	1:B:811:ALA:HB2	1.70	0.73
1:A:685:ASP:HB2	1:A:689:ARG:HD3	1.72	0.72

All (1) 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)
1:B:590:TYR:OH	1:A:451:ASP:OD2[1_554]	2.14	0.06

## 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	842/883 (95%)	774 (92%)	56 (7%)	12 (1%)	11	44
1	B	842/883 (95%)	778 (92%)	56 (7%)	8 (1%)	15	52
All	All	1684/1766 (95%)	1552 (92%)	112 (7%)	20 (1%)	13	48

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	297	ARG
1	A	665	ASP
1	B	679	GLY
1	A	206	VAL
1	A	802	GLY



### 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	661/690 (96%)	638 (96%)	23 (4%)	36 67
1	B	659/690 (96%)	635 (96%)	24 (4%)	35 66
All	All	1320/1380 (96%)	1273 (96%)	47 (4%)	35 66

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	720	PHE
1	A	120	ASN
1	A	742	PHE
1	B	747	ARG
1	A	160	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	222	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 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	848/883 (96%)	1.44	204 (24%) <b>0</b> <b>1</b>	71, 125, 165, 201	0
1	B	848/883 (96%)	1.25	197 (23%) <b>0</b> <b>1</b>	77, 127, 165, 204	0
All	All	1696/1766 (96%)	1.35	401 (23%) <b>0</b> <b>1</b>	71, 126, 165, 204	0

The worst 5 of 401 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	597	GLY	28.0
1	B	44	ASP	24.9
1	A	272	SER	23.0
1	A	267	ASP	19.8
1	B	43	THR	19.1

### 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.