



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2021 – 04:01 AM EDT

PDB ID : 1JPI  
Title : Phe232Leu mutant of human UROD, human uroporphyrinogen III decarboxylase  
Authors : Phillips, J.D.; Parker, T.L.; Schubert, H.L.; Whitby, F.G.; Hill, C.P.; Kushner, J.P.  
Deposited on : 2001-08-02  
Resolution : 2.30 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 : 2.23.2  
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) : 2.23.2

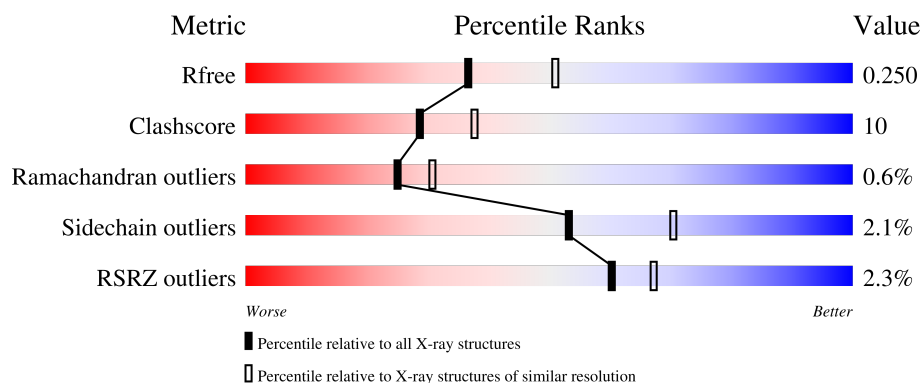
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	388	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3086 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 UROPORPHYRINOGEN DECARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	357	Total	C	N	O	S	101	0	0
			2799	1791	489	504	15			

There are 22 discrepancies between the modelled and reference sequences:

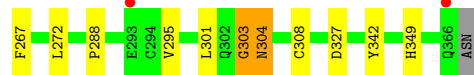
Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP P06132
A	-19	GLY	-	expression tag	UNP P06132
A	-18	HIS	-	expression tag	UNP P06132
A	-17	HIS	-	expression tag	UNP P06132
A	-16	HIS	-	expression tag	UNP P06132
A	-15	HIS	-	expression tag	UNP P06132
A	-14	HIS	-	expression tag	UNP P06132
A	-13	HIS	-	expression tag	UNP P06132
A	-12	HIS	-	expression tag	UNP P06132
A	-11	HIS	-	expression tag	UNP P06132
A	-10	HIS	-	expression tag	UNP P06132
A	-9	HIS	-	expression tag	UNP P06132
A	-8	SER	-	expression tag	UNP P06132
A	-7	SER	-	expression tag	UNP P06132
A	-6	GLY	-	expression tag	UNP P06132
A	-5	HIS	-	expression tag	UNP P06132
A	-4	ILE	-	expression tag	UNP P06132
A	-3	GLU	-	expression tag	UNP P06132
A	-2	GLY	-	expression tag	UNP P06132
A	-1	ARG	-	expression tag	UNP P06132
A	0	HIS	-	expression tag	UNP P06132
A	232	LEU	PHE	engineered mutation	UNP P06132

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	287	Total 287	O 287	0	0



• Molecule 1: UROPORPHYRINOGEN DECARBOXYLASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.32Å 103.32Å 72.03Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.30 16.46 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.0 (20.00-2.30) 99.0 (16.46-2.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 2.30Å)	Xtriage
Refinement program	X-PLOR 3.843	Depositor
R, $R_{free}$	0.200 , 0.243 0.215 , 0.250	Depositor DCC
$R_{free}$ test set	958 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.3	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 65.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3086	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% 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	0.36	0/2869	0.66	4/3898 (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	A	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	303	GLY	N-CA-C	12.34	143.96	113.10
1	A	304	ASN	N-CA-C	10.19	138.52	111.00
1	A	303	GLY	CA-C-N	-7.82	99.99	117.20
1	A	303	GLY	C-N-CA	5.66	135.85	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	GLY	Mainchain

### 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	2799	0	2788	52	0
2	A	287	0	0	4	0
All	All	3086	0	2788	52	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 10.

The worst 5 of 52 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:114:GLU:HG3	1:A:116:GLN:H	1.48	0.79
1:A:73:LEU:HD11	1:A:81:ALA:HB2	1.68	0.73
1:A:308:CYS:SG	1:A:342:TYR:HE1	2.13	0.72
1:A:83:ILE:HD13	1:A:84:PHE:O	1.90	0.71
1:A:194:ILE:HD13	2:A:1205:HOH:O	1.95	0.67

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	355/388 (92%)	343 (97%)	10 (3%)	2 (1%)	25 31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	304	ASN
1	A	57	SER



### 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	291/316 (92%)	285 (98%)	6 (2%)	53 70

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	255	PRO
1	A	260	ILE
1	A	349	HIS
1	A	83	ILE
1	A	50	ARG

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

Mol	Chain	Res	Type
1	A	143	GLN
1	A	319	GLN
1	A	349	HIS

### 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

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	343/388 (88%)	-0.19	8 (2%) 60 67	14, 30, 56, 94	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	59	CYS	7.1
1	A	106	PRO	5.0
1	A	48	GLU	2.7
1	A	366	GLN	2.7
1	A	87	ILE	2.4

### 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.