



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 3, 2019 – 07:12 PM EDT

PDB ID : 2C7X  
Title : Crystal structure of narbomycin-bound cytochrome P450 PikC (CYP107L1)  
Authors : Sherman, D.H.; Li, S.; Yermalitskaya, L.V.; Kim, Y.; Smith, J.A.; Waterman, M.R.; Podust, L.M.  
Deposited on : 2005-11-29  
Resolution : 1.75 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  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : rb-20031633  
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) : rb-20031633

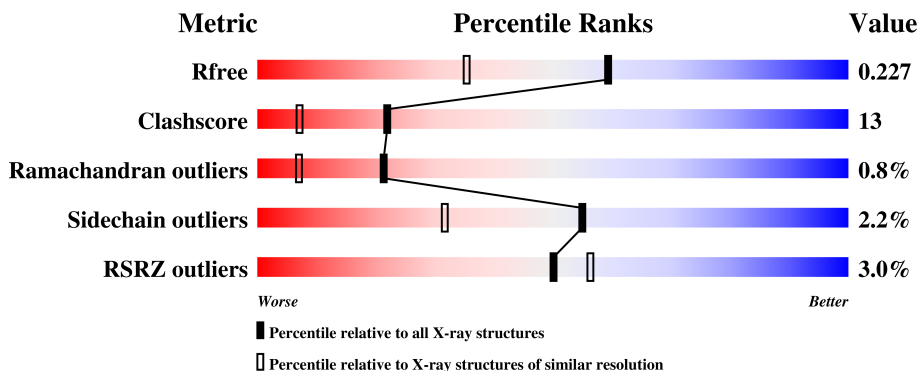
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	1952 (1.76-1.76)
Clashscore	122126	2072 (1.76-1.76)
Ramachandran outliers	120053	2050 (1.76-1.76)
Sidechain outliers	120020	2050 (1.76-1.76)
RSRZ outliers	108989	1913 (1.76-1.76)

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	A	436	

## 2 Entry composition [i](#)

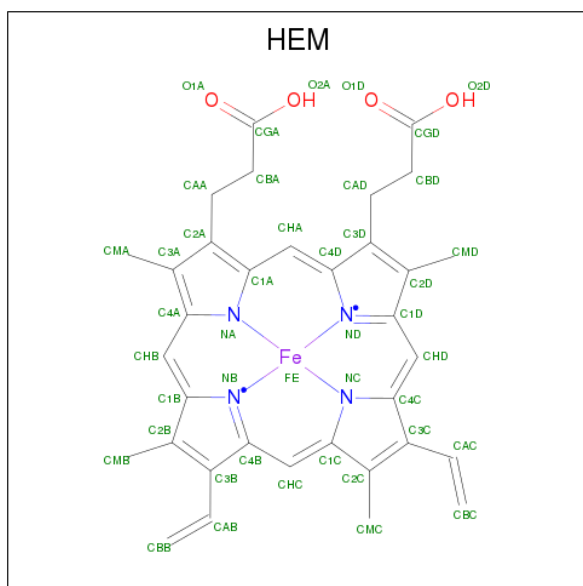
There are 4 unique types of molecules in this entry. The entry contains 3296 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 CYTOCHROME P450 MONOOXYGENASE.

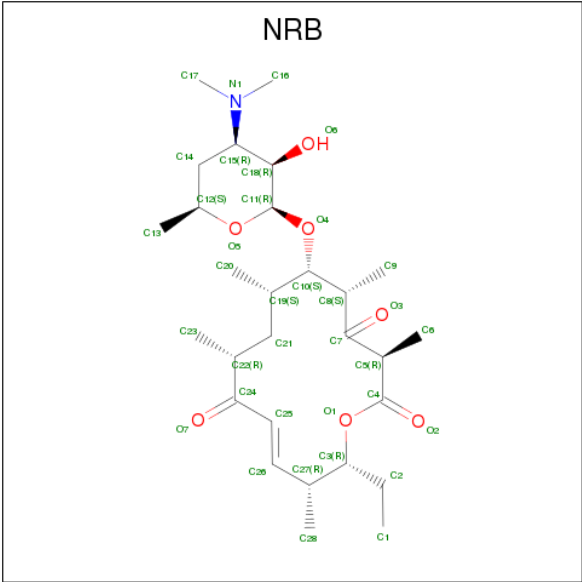
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	397	Total	C	N	O	S	0	0	1
			3051	1930	546	562	13			

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is NARBOMYCIN (three-letter code: NRB) (formula:  $C_{28}H_{47}NO_7$ ).



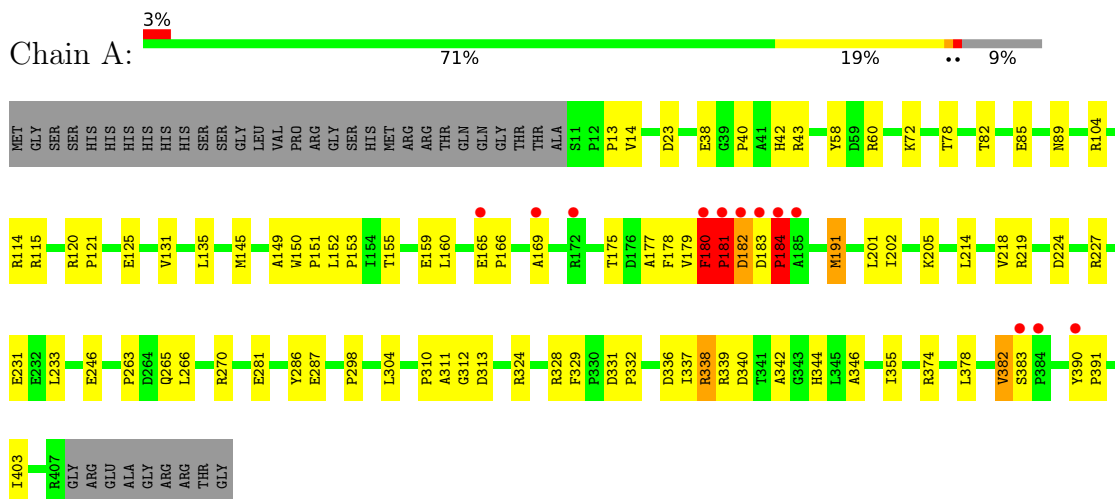
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			36	28	1	7		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	166	Total	O	0	0
			166	166		



- Molecule 1: CYTOCHROME P450 MONOOXYGENASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.96Å 64.90Å 92.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.77 – 1.75 49.77 – 1.75	Depositor EDS
% Data completeness (in resolution range)	94.1 (49.77-1.75) 94.2 (49.77-1.75)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.52 (at 1.75Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.207 , 0.232 0.202 , 0.227	Depositor DCC
$R_{free}$ test set	3512 reflections (9.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.9	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3296	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 6.06% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, NRB

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	3/3124 (0.1%)	0.65	5/4268 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	13	PRO	N-CD	-25.41	1.12	1.47
1	A	263	PRO	N-CD	-5.76	1.39	1.47
1	A	180	PHE	C-N	-5.75	1.23	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	13	PRO	N-CD-CG	8.35	115.73	103.20
1	A	180	PHE	CB-CA-C	-5.94	98.52	110.40
1	A	184	PRO	CA-N-CD	-5.80	103.38	111.50
1	A	181	PRO	CA-C-N	-5.78	104.49	117.20
1	A	180	PHE	CB-CG-CD1	-5.38	117.03	120.80

There are no chirality outliers.

There are no planarity outliers.

### 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	3051	0	3007	77	0
2	A	43	0	30	1	0
3	A	36	0	47	3	0
4	A	166	0	0	9	0
All	All	3296	0	3084	79	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 79 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:342:ALA:HA	4:A:2142:HOH:O	1.48	1.09
1:A:175:THR:HG21	1:A:246:GLU:OE1	1.60	1.01
1:A:180:PHE:CZ	1:A:390:TYR:HD2	1.81	0.97
1:A:78:THR:HG21	4:A:2129:HOH:O	1.63	0.97
1:A:131:VAL:HB	1:A:374:ARG:HH21	1.29	0.97

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	395/436 (91%)	384 (97%)	8 (2%)	3 (1%)	21 7

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	181	PRO
1	A	182	ASP
1	A	382	VAL



### 5.3.2 Protein sidechains ⓘ

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	315/355 (89%)	308 (98%)	7 (2%)	55 32

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	181	PRO
1	A	338	ARG
1	A	184	PRO
1	A	160	LEU
1	A	191	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	89	ASN
1	A	208	GLN
1	A	238	HIS
1	A	265	GLN
1	A	344	HIS

### 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

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEM	A	1407	1	27,50,50	1.75	10 (37%)	17,82,82	1.06	1 (5%)
3	NRB	A	1408	-	37,37,37	1.76	11 (29%)	44,53,53	2.15	11 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	A	1407	1	-	0/6/54/54	0/0/8/8
3	NRB	A	1408	-	-	0/51/67/67	0/1/2/2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1407	HEM	C3B-CAB	-3.73	1.40	1.47
2	A	1407	HEM	C3C-C2C	-3.39	1.35	1.40
2	A	1407	HEM	C3C-CAC	-3.26	1.41	1.47
2	A	1407	HEM	C3B-C2B	-2.20	1.37	1.40
2	A	1407	HEM	C1A-NA	2.04	1.40	1.36

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1408	NRB	O2-C4-C5	-3.85	119.61	124.78
3	A	1408	NRB	C11-O4-C10	-3.57	109.06	117.97
3	A	1408	NRB	C11-O5-C12	-2.08	109.62	112.92
3	A	1408	NRB	C14-C15-N1	-2.07	109.83	115.67
3	A	1408	NRB	O5-C12-C13	2.08	111.21	106.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1407	HEM	1	0
3	A	1408	NRB	3	0

## 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	397/436 (91%)	0.12	12 (3%) 50 56	21, 33, 53, 74	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	182	ASP	5.1
1	A	180	PHE	4.7
1	A	183	ASP	3.8
1	A	185	ALA	3.5
1	A	384	PRO	3.0

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NRB	A	1408	36/36	0.90	0.16	36,44,54,54	0
2	HEM	A	1407	43/43	0.98	0.08	19,22,25,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.