



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 04:30 am BST

PDB ID : 1VL7  
Title : Crystal structure of a putative heme oxygenase (alr5027) from nostoc sp. pcc 7120 at 1.50 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2004-07-13  
Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.8.5 (274361), CSD as541be (2020)  
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

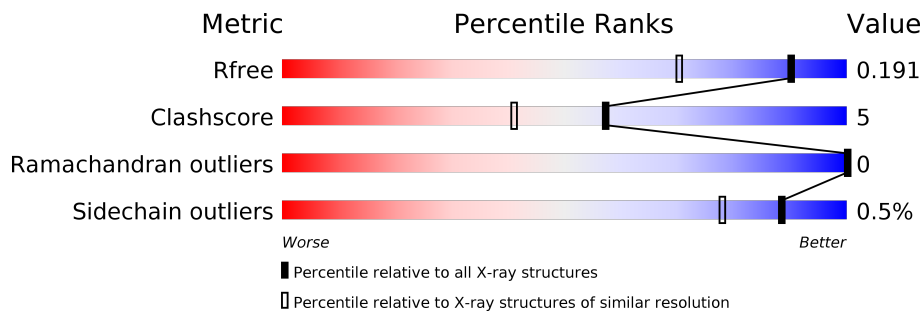
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	157	 78%      8% •      14%
1	B	157	 78%      7% •      14%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 2519 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.

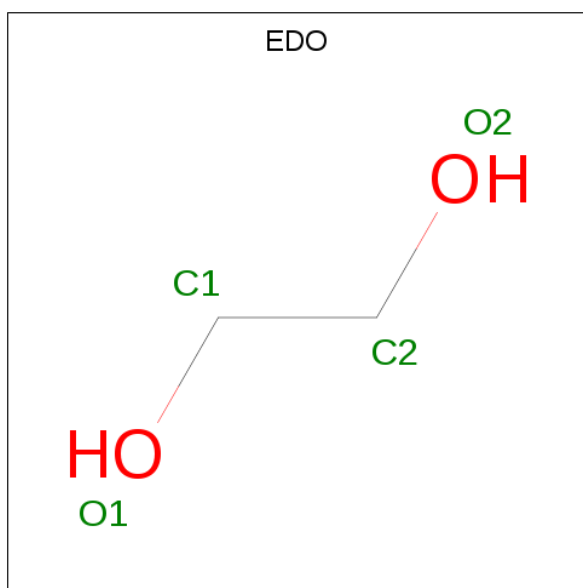
- Molecule 1 is a protein called hypothetical protein alr5027.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	135	1087	703	182	201	1	0	4	0
1	B	135	1131	727	191	212	1	0	11	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q8YMA7
A	-10	GLY	-	LEADER SEQUENCE	UNP Q8YMA7
A	-9	SER	-	LEADER SEQUENCE	UNP Q8YMA7
A	-8	ASP	-	LEADER SEQUENCE	UNP Q8YMA7
A	-7	LYS	-	LEADER SEQUENCE	UNP Q8YMA7
A	-6	ILE	-	LEADER SEQUENCE	UNP Q8YMA7
A	-5	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
A	-4	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
A	-3	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
A	-2	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
A	-1	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
A	0	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-11	MET	-	LEADER SEQUENCE	UNP Q8YMA7
B	-10	GLY	-	LEADER SEQUENCE	UNP Q8YMA7
B	-9	SER	-	LEADER SEQUENCE	UNP Q8YMA7
B	-8	ASP	-	LEADER SEQUENCE	UNP Q8YMA7
B	-7	LYS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-6	ILE	-	LEADER SEQUENCE	UNP Q8YMA7
B	-5	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-4	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-3	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-2	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	-1	HIS	-	LEADER SEQUENCE	UNP Q8YMA7
B	0	HIS	-	LEADER SEQUENCE	UNP Q8YMA7

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	134	Total O 134 134	0	0
3	B	135	Total O 135 135	0	0

### 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. 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. 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: hypothetical protein alr5027

Chain A: 



- Molecule 1: hypothetical protein alr5027

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.00Å 100.33Å 47.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.05 – 1.50 43.05 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (43.05-1.50) 99.5 (43.05-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.57 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.2.0000	Depositor
R, $R_{free}$	0.154 , 0.183 0.164 , 0.191	Depositor DCC
$R_{free}$ test set	2570 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.4	Xtrriage
Anisotropy	0.443	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 50.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2519	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.51% 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](#)

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

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	0.81	0/1118	0.95	2/1514 (0.1%)
1	B	0.77	0/1181	0.97	3/1600 (0.2%)
All	All	0.79	0/2299	0.96	5/3114 (0.2%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	98	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	A	73	ASP	CB-CG-OD1	5.31	123.08	118.30
1	B	73	ASP	CB-CG-OD1	5.10	122.89	118.30
1	A	43	ASP	CB-CG-OD2	5.09	122.89	118.30
1	B	113	ARG	NE-CZ-NH1	5.06	122.83	120.30

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	1087	0	1066	13	0
1	B	1131	0	1095	9	0
2	A	12	0	18	0	0
2	B	20	0	30	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	134	0	0	1	0
3	B	135	0	0	3	0
All	All	2519	0	2209	22	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 5.

All (22) 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:63:ALA:HB3	3:A:209:HOH:O	1.73	0.87
1:A:139:PHE:CZ	1:A:141:ILE:HD11	2.17	0.80
1:B:109:GLN:NE2	3:B:260:HOH:O	2.33	0.61
1:B:41[A]:ILE:HA	1:B:46[A]:ASN:O	2.01	0.59
1:B:135:LYS:HE3	3:B:250:HOH:O	2.02	0.58
1:B:139:PHE:CZ	1:B:141:ILE:HD11	2.39	0.57
1:A:41:ILE:HA	1:A:46[A]:ASN:O	2.04	0.57
1:A:11:TYR:C	1:A:11:TYR:CD1	2.80	0.54
1:A:85:ARG:HG3	1:A:141:ILE:HD13	1.90	0.54
1:B:139:PHE:CE2	1:B:141:ILE:HD11	2.44	0.53
1:A:11:TYR:C	1:A:11:TYR:HD1	2.14	0.51
1:B:41[B]:ILE:HG22	1:B:42[B]:ASP:O	2.16	0.46
1:B:44[B]:ALA:O	1:B:45[B]:LYS:HB2	2.15	0.45
1:A:46[B]:ASN:HD22	1:A:47:ILE:N	2.01	0.45
1:A:76:ALA:HB2	3:B:197:HOH:O	2.18	0.44
1:A:15:ILE:HG21	1:A:41:ILE:CD1	2.48	0.44
1:A:84:ARG:HH12	2:B:149:EDO:C1	2.32	0.43
1:B:42[A]:ASP:HB3	1:B:109:GLN:HE22	1.83	0.42
1:A:85:ARG:HG3	1:A:141:ILE:CD1	2.49	0.41
1:B:113:ARG:HD2	2:B:148:EDO:H21	2.02	0.40

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	137/157 (87%)	135 (98%)	2 (2%)	0	100	100
1	B	144/157 (92%)	138 (96%)	6 (4%)	0	100	100
All	All	281/314 (90%)	273 (97%)	8 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	113/134 (84%)	112 (99%)	1 (1%)	78	61
1	B	119/134 (89%)	119 (100%)	0	100	100
All	All	232/268 (87%)	231 (100%)	1 (0%)	88	82

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	TYR

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	109	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](#)

8 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	EDO	B	148	-	3,3,3	0.53	0	2,2,2	0.34	0
2	EDO	A	148	-	3,3,3	0.40	0	2,2,2	0.29	0
2	EDO	B	149	-	3,3,3	0.47	0	2,2,2	1.12	0
2	EDO	A	147	-	3,3,3	0.39	0	2,2,2	0.89	0
2	EDO	B	147	-	3,3,3	0.64	0	2,2,2	0.11	0
2	EDO	A	146	-	3,3,3	0.69	0	2,2,2	0.53	0
2	EDO	B	146	-	3,3,3	0.72	0	2,2,2	0.24	0
2	EDO	B	150	-	3,3,3	0.54	0	2,2,2	0.70	0

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	EDO	B	148	-	-	0/1/1/1	-
2	EDO	A	148	-	-	1/1/1/1	-
2	EDO	B	149	-	-	1/1/1/1	-
2	EDO	A	147	-	-	1/1/1/1	-
2	EDO	B	147	-	-	0/1/1/1	-
2	EDO	A	146	-	-	0/1/1/1	-
2	EDO	B	146	-	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	B	150	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	148	EDO	O1-C1-C2-O2
2	B	149	EDO	O1-C1-C2-O2
2	A	147	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	148	EDO	1	0
2	B	149	EDO	1	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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.