



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 08:28 pm BST

PDB ID : 1SP8  
Title : 4-Hydroxyphenylpyruvate Dioxygenase  
Authors : Fritze, I.M.; Linden, L.; Freigang, J.; Auerbach, G.; Huber, R.; Steinbacher, S.  
Deposited on : 2004-03-16  
Resolution : 2.00 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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

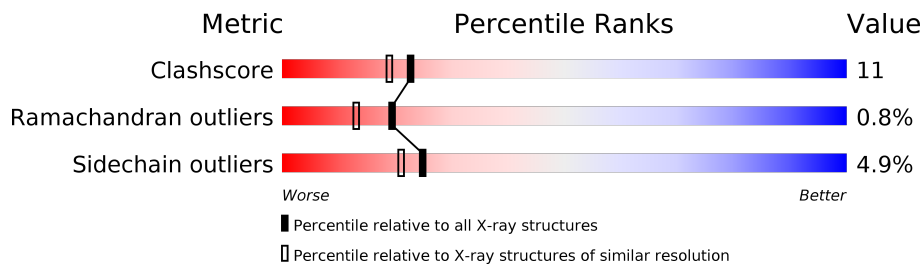
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	418	
1	B	418	
1	C	418	
1	D	418	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12660 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 4-Hydroxyphenylpyruvate Dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	388	Total 2938	C 1854	N 524	O 550	S 10	137	0	0
1	B	392	Total 2971	C 1876	N 528	O 557	S 10	125	0	0
1	C	386	Total 2920	C 1844	N 519	O 547	S 10	129	0	0
1	D	393	Total 2980	C 1881	N 529	O 560	S 10	130	0	0

- Molecule 2 is FE (II) ION (three-letter code: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Fe 1 1	0	0
2	A	1	Total Fe 1 1	0	0
2	D	1	Total Fe 1 1	0	0
2	C	1	Total Fe 1 1	0	0

- Molecule 3 is water.

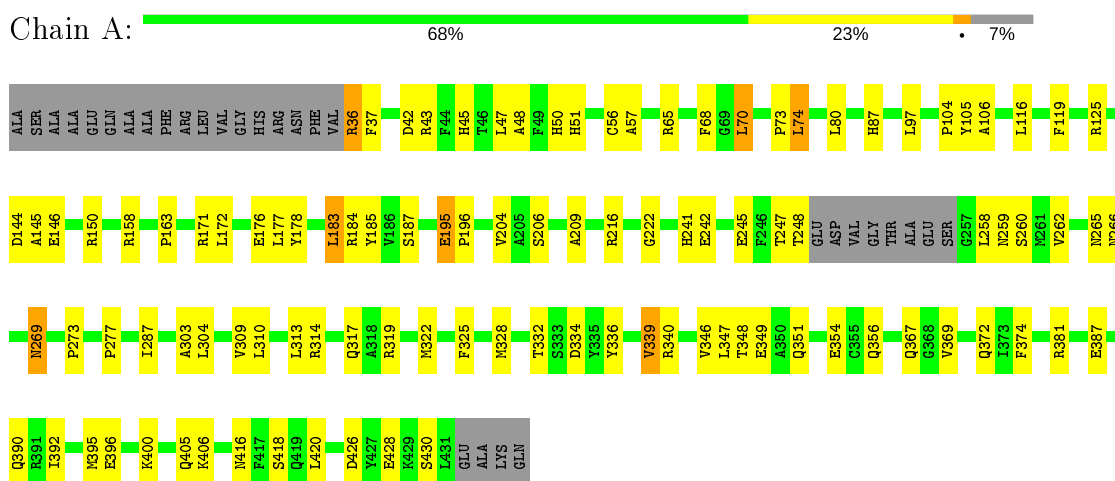
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	226	Total O 226 226	0	0
3	B	215	Total O 215 215	0	0
3	C	201	Total O 201 201	0	0
3	D	205	Total O 205 205	1	0

### 3 Residue-property plots

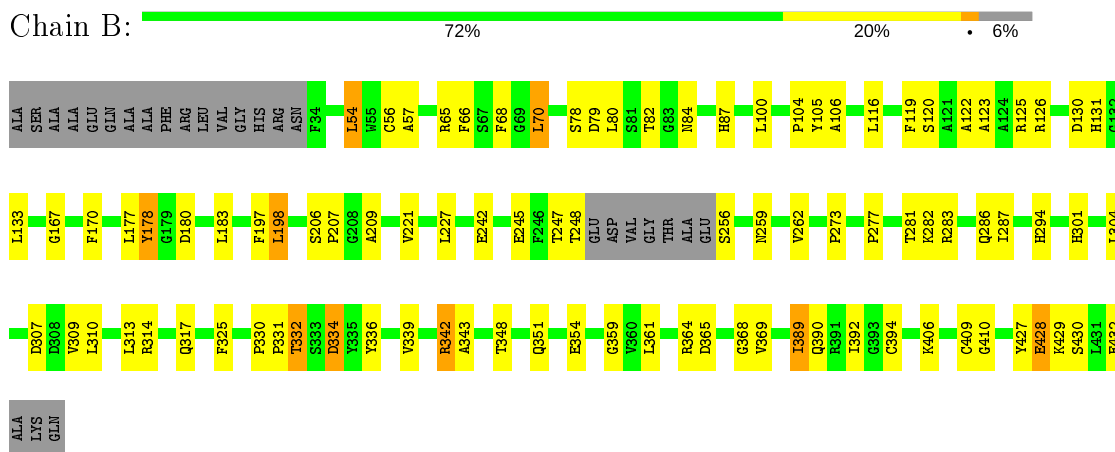
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

Note EDS was not executed.

- Molecule 1: 4-Hydroxyphenylpyruvate Dioxygenase

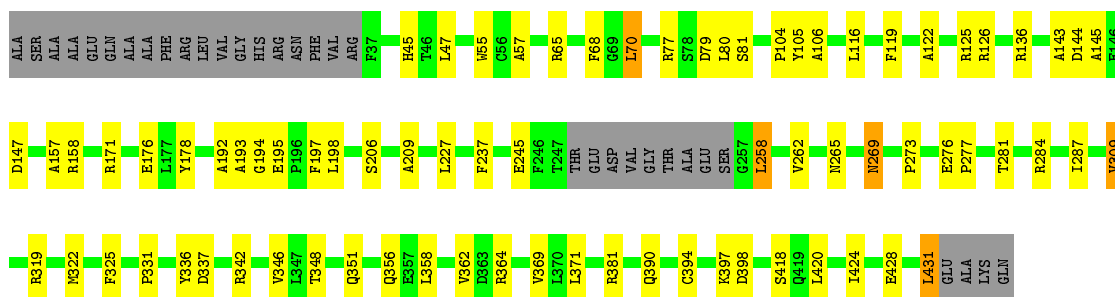


- Molecule 1: 4-Hydroxyphenylpyruvate Dioxygenase



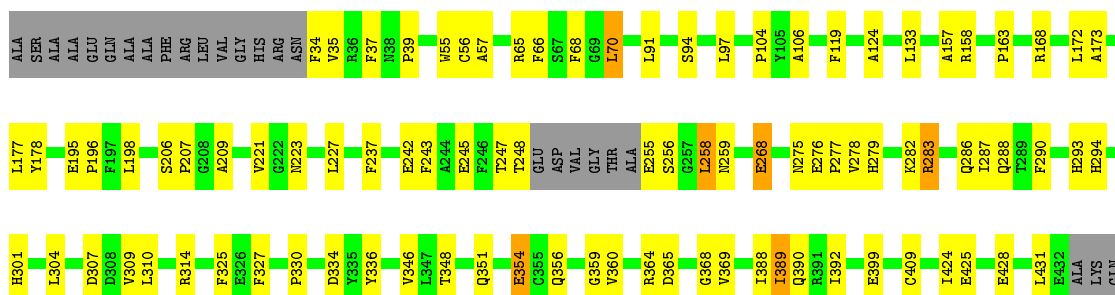
- Molecule 1: 4-Hydroxyphenylpyruvate Dioxygenase





- Molecule 1: 4-Hydroxyphenylpyruvate Dioxygenase

Chain D:  72%  21%  6%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.00Å 110.90Å 174.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.84 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (19.84-2.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.275 , 0.324	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	12660	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

## 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: FE2

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	0.52	0/3007	0.72	0/4069
1	B	0.50	0/3041	0.71	1/4115 (0.0%)
1	C	0.51	0/2989	0.70	0/4045
1	D	0.51	1/3050 (0.0%)	0.72	0/4127
All	All	0.51	1/12087 (0.0%)	0.71	1/16356 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	56	CYS	CB-SG	-5.01	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	54	LEU	CA-CB-CG	5.53	128.02	115.30

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	2938	0	2830	75	0
1	B	2971	0	2859	70	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2920	0	2810	42	0
1	D	2980	0	2865	71	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	226	0	0	4	0
3	B	215	0	0	3	0
3	C	201	0	0	4	0
3	D	205	0	0	8	0
All	All	12660	0	11364	251	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 11.

The worst 5 of 251 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:206:SER:HB2	1:A:209:ALA:HB3	1.48	0.95
1:D:206:SER:HB2	1:D:209:ALA:HB3	1.48	0.94
1:B:282:LYS:HG3	1:B:283:ARG:H	1.38	0.87
1:C:265:ASN:HD21	1:C:269:ASN:ND2	1.74	0.85
1:C:265:ASN:HD21	1:C:269:ASN:HD21	1.25	0.84

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	384/418 (92%)	362 (94%)	20 (5%)	2 (0%)	29 23
1	B	388/418 (93%)	368 (95%)	18 (5%)	2 (0%)	29 23

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	382/418 (91%)	364 (95%)	13 (3%)	5 (1%)	12	6
1	D	389/418 (93%)	365 (94%)	20 (5%)	4 (1%)	15	9
All	All	1543/1672 (92%)	1459 (95%)	71 (5%)	13 (1%)	19	13

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	430	SER
1	C	281	THR
1	C	397	LYS
1	C	193	ALA
1	D	282	LYS

### 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	292/313 (93%)	277 (95%)	15 (5%)	24	19
1	B	296/313 (95%)	283 (96%)	13 (4%)	28	25
1	C	290/313 (93%)	272 (94%)	18 (6%)	18	13
1	D	297/313 (95%)	286 (96%)	11 (4%)	34	32
All	All	1175/1252 (94%)	1118 (95%)	57 (5%)	25	21

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	389	ILE
1	C	178	TYR
1	D	334	ASP
1	B	409	CYS
1	C	65	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	288	GLN
1	B	356	GLN
1	D	351	GLN
1	B	294	HIS
1	B	419	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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.