WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 01:49 PM EST

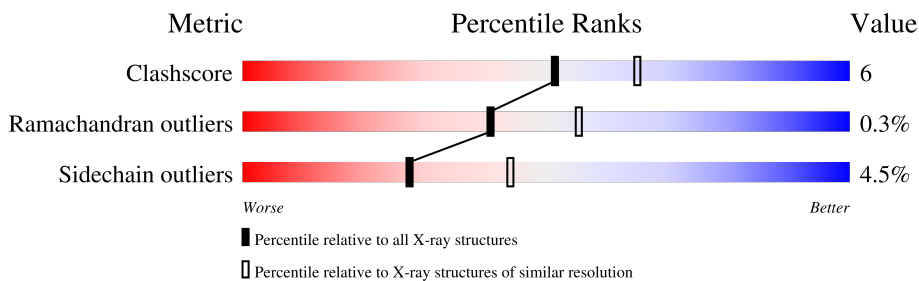|		
PDB ID	:	1CJX
Title	:	CRYSTAL STRUCTURE OF PSEUDOMONAS FLUORESCENS HPPD
Authors	:	Serre, L.; Sailland, A.; Sy, D.; Boudec, P.; Rolland, A.; Pebay-Peroulla, E.; Cohen-Addad, C.
Deposited on	:	1999-04-20
Resolution	:	2.40 Å(reported)

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 2.40 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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0.3%
Sidechain outliers		4.5%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	357	78% 17% • •
1	B	357	80% 17% • •
1	C	357	81% 16% • • •
1	D	357	80% 15% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11821 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
1	A	353	Total	C	N	O	S	0	4	1
			2801	1791	474	522	14			
1	B	353	Total	C	N	O	S	34	4	1
			2805	1793	477	521	14			
1	C	353	Total	C	N	O	S	41	4	1
			2806	1794	477	521	14			
1	D	353	Total	C	N	O	S	40	5	1
			2810	1796	477	523	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	355	ALA	SER	SEE REMARK 999	UNP P80064
B	355	ALA	SER	SEE REMARK 999	UNP P80064
C	355	ALA	SER	SEE REMARK 999	UNP P80064
D	355	ALA	SER	SEE REMARK 999	UNP P80064

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

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

- Molecule 3 is ETHYL MERCURY ION (three-letter code: EMC) (formula: C_2H_5Hg).

EMC

HG

Hg$^+$

C2

C1
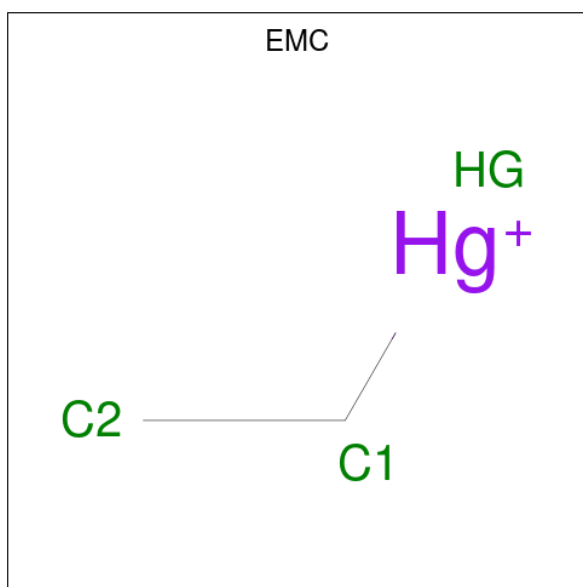
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 3	C 2	Hg 1	0	0
3	B	1	Total 3	C 2	Hg 1	0	0
3	C	1	Total 3	C 2	Hg 1	0	0
3	D	1	Total 3	C 2	Hg 1	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

OXT

O$^-$

C

CH3

O

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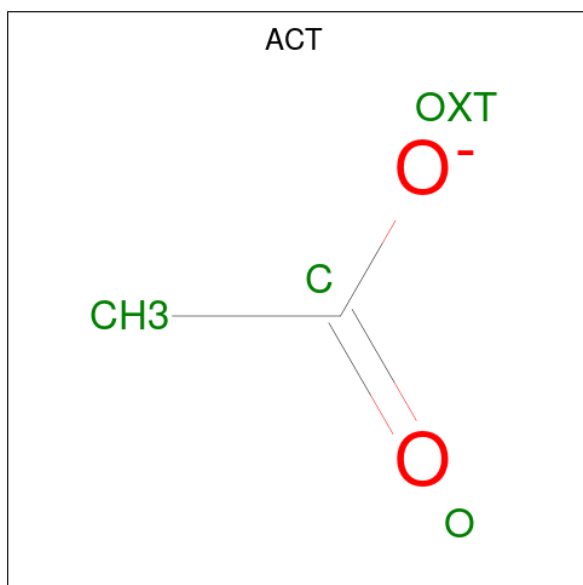
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	D	1	Total 4	C 2	O 2	0	0

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	148	Total 148	O 148	0	0
5	B	138	Total 138	O 138	0	0
5	C	137	Total 137	O 137	0	0
5	D	144	Total 144	O 144	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.
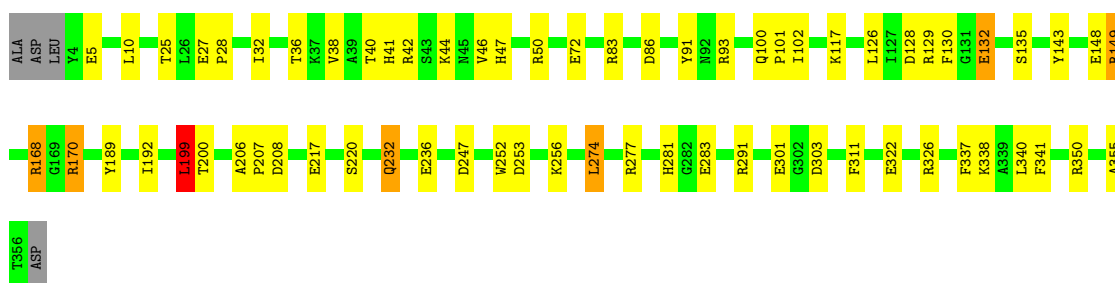
• Molecule 1: 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE

Chain A:
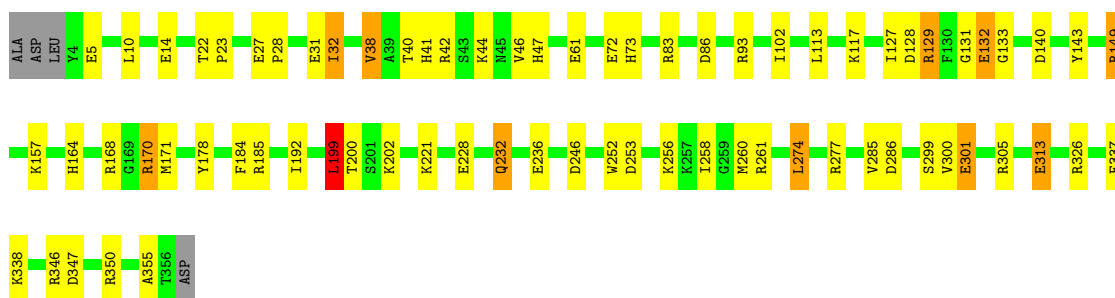
• Molecule 1: 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE

Chain B:

• Molecule 1: 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE

Chain C:
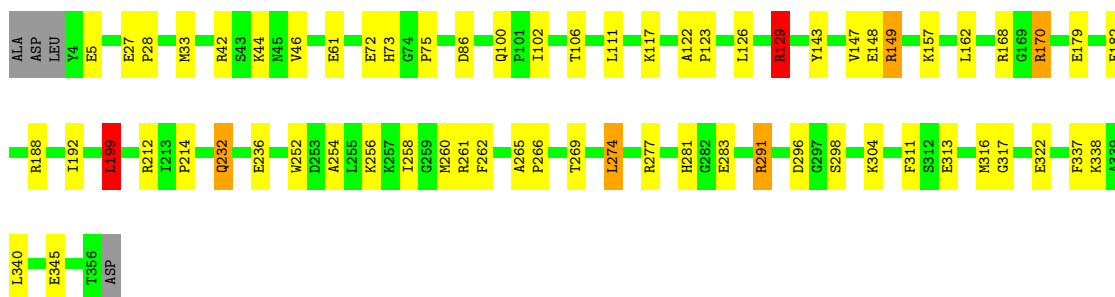
• Molecule 1: 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE

Chain D:
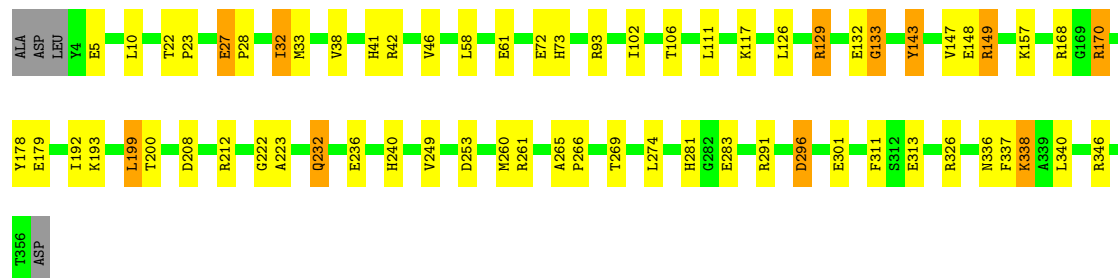
4 Data and refinement statistics (i)

Xtriage (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, α, β, γ	79.59Å 142.75Å 159.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.40	Depositor
% Data completeness (in resolution range)	79.8 (20.00-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	8.00	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.219 , 0.276	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11821	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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: EMC, ACT, 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	6.72	6/2892 (0.2%)	2.08	41/3901 (1.1%)
1	B	0.47	1/2895 (0.0%)	1.46	33/3904 (0.8%)
1	C	0.50	1/2895 (0.0%)	1.36	23/3904 (0.6%)
1	D	0.49	0/2904	1.37	27/3917 (0.7%)
All	All	3.39	8/11586 (0.1%)	1.60	124/15626 (0.8%)

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	2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	313[A]	GLU	CD-OE1	233.44	3.82	1.25
1	A	313[B]	GLU	CD-OE1	233.44	3.82	1.25
1	A	313[A]	GLU	CD-OE2	84.97	2.19	1.25
1	A	313[B]	GLU	CD-OE2	84.97	2.19	1.25
1	A	313[A]	GLU	CG-CD	57.69	2.38	1.51

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	313[A]	GLU	CG-CD-OE1	-43.26	31.79	118.30
1	A	313[B]	GLU	CG-CD-OE1	-43.26	31.79	118.30
1	A	313[A]	GLU	OE1-CD-OE2	-41.84	73.09	123.30
1	A	313[B]	GLU	OE1-CD-OE2	-41.84	73.09	123.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	313[A]	GLU	CG-CD-OE2	-27.68	62.94	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	132	GLU	Mainchain
1	A	31	GLU	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	2801	0	2725	31	0
1	B	2805	0	2736	26	0
1	C	2806	0	2740	35	0
1	D	2810	0	2742	33	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	3	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	3	0	0	0	0
4	A	4	0	3	0	0
4	B	4	0	3	0	0
4	C	4	0	3	0	0
4	D	4	0	3	0	0
5	A	148	0	0	4	0
5	B	138	0	0	4	0
5	C	137	0	0	4	0
5	D	144	0	0	3	0
All	All	11821	0	10955	124	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 6.

The worst 5 of 124 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:D:261[B]:ARG:NH1	1:D:313[B]:GLU:OE2	1.90	1.04
1:A:313[A]:GLU:CD	1:A:313[A]:GLU:CG	2.38	0.92
1:A:102:ILE:HD11	1:A:117:LYS:HG2	1.53	0.90
1:D:281:HIS:HD2	1:D:283:GLU:H	1.22	0.85
1:A:313[A]:GLU:CD	1:A:313[A]:GLU:OE2	2.19	0.81

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	355/357 (99%)	339 (96%)	15 (4%)	1 (0%)	41	55
1	B	355/357 (99%)	340 (96%)	13 (4%)	2 (1%)	25	36
1	C	355/357 (99%)	343 (97%)	12 (3%)	0	100	100
1	D	356/357 (100%)	339 (95%)	16 (4%)	1 (0%)	41	55
All	All	1421/1428 (100%)	1361 (96%)	56 (4%)	4 (0%)	41	55

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	355	ALA
1	B	132	GLU
1	D	301	GLU
1	B	301	GLU

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	296/296 (100%)	282 (95%)	14 (5%)	26	42
1	B	296/296 (100%)	282 (95%)	14 (5%)	26	42
1	C	296/296 (100%)	285 (96%)	11 (4%)	34	53
1	D	297/296 (100%)	283 (95%)	14 (5%)	26	42
All	All	1185/1184 (100%)	1132 (96%)	53 (4%)	27	44

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	5	GLU
1	C	274	LEU
1	D	269	THR
1	C	126	LEU
1	C	232	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	281	HIS
1	C	92	ASN
1	D	281	HIS
1	D	92	ASN
1	D	232	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 for Mogul analysis.

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	#$|Z| > 2$
4	ACT	A	631	2	3,3,3	0.70	0	3,3,3	1.36	0
4	ACT	B	631	2	3,3,3	0.64	0	3,3,3	1.17	0
3	EMC	A	630	1	1,2,2	0.40	0	-		
3	EMC	B	630	1	1,2,2	0.64	0	-		
4	ACT	D	631	2	3,3,3	0.76	0	3,3,3	1.04	0
3	EMC	C	630	1	1,2,2	0.42	0	-		
4	ACT	C	631	2	3,3,3	0.82	0	3,3,3	1.30	0
3	EMC	D	630	1	1,2,2	0.71	0	-		

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 ⓘ

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 ⓘ

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.