



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 01:01 pm BST

PDB ID : 3HWC
Title : Crystal Structure of Chlorophenol 4-Monooxygenase (TftD) of Burkholderia cepacia AC1100
Authors : Ballinger, J.W.; Kang, C.H.
Deposited on : 2009-06-17
Resolution : 2.50 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

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) : 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

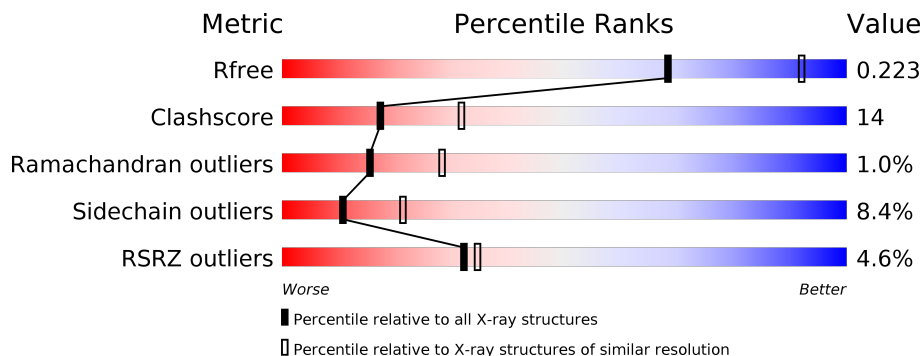
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 ≥ 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	515	 6% 64% 24% 5% 6%
1	B	515	 6% 65% 24% • 6%
1	C	515	 4% 64% 26% • 6%
1	D	515	 70% 20% • 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16516 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 Chlorophenol-4-monooxygenase component 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	482	3834	2429	684	708	13	0	0	0
1	B	482	3834	2429	684	708	13	0	0	0
1	C	482	3834	2429	684	708	13	0	0	0
1	D	482	3834	2429	684	708	13	0	0	0

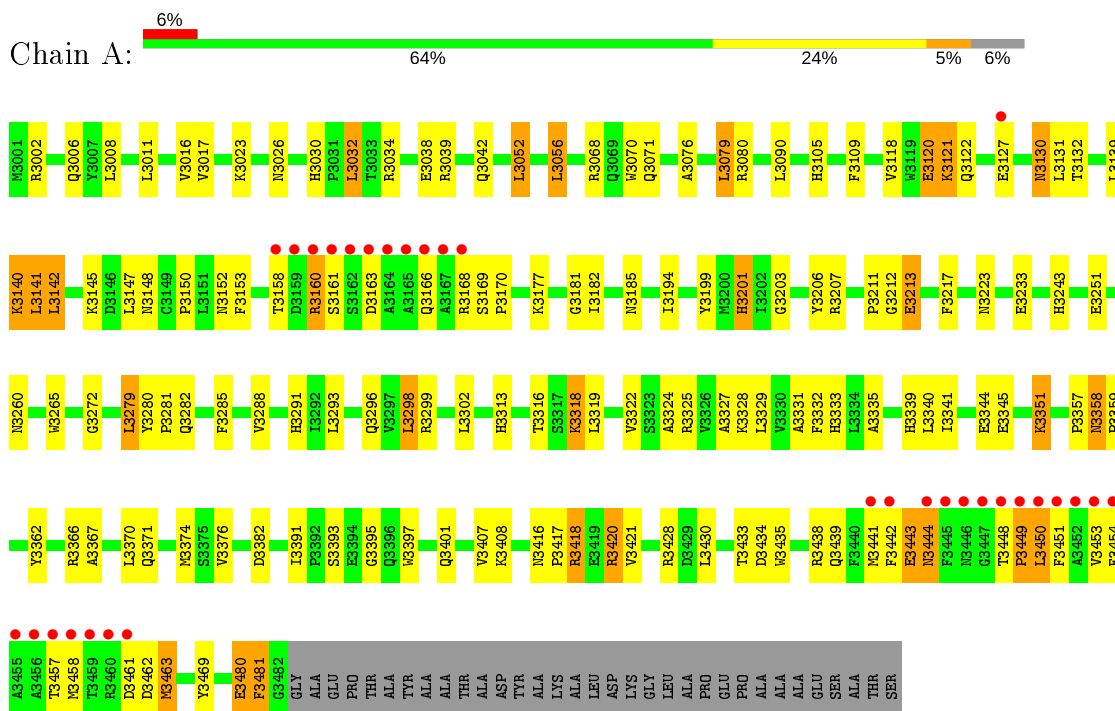
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	289	Total	O	0	0
			289	289		
2	B	271	Total	O	0	0
			271	271		
2	C	245	Total	O	0	0
			245	245		
2	D	375	Total	O	0	0
			375	375		

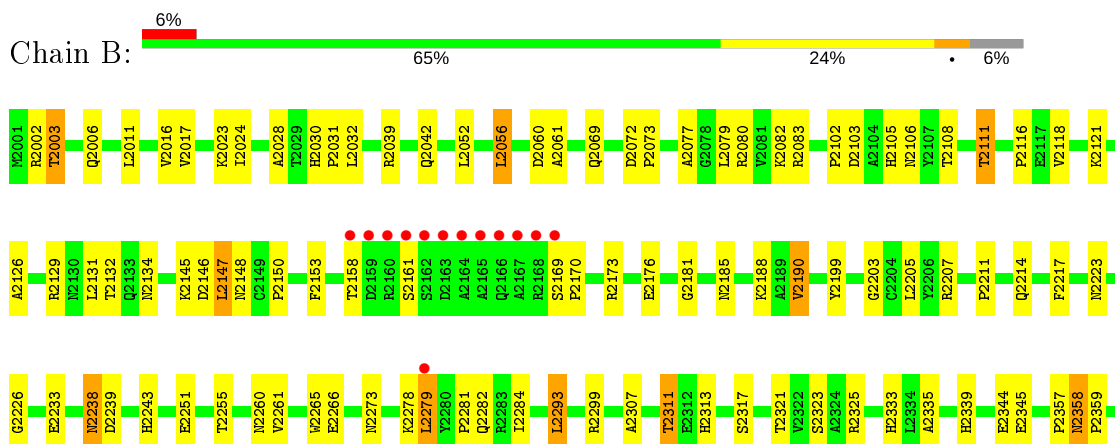
3 Residue-property plots [i](#)

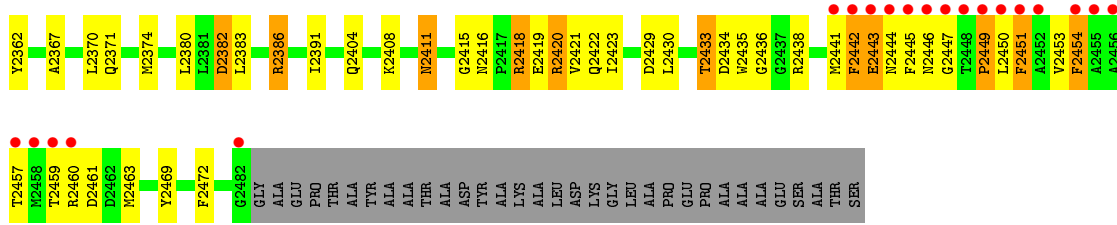
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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: Chlorophenol-4-monooxygenase component 2

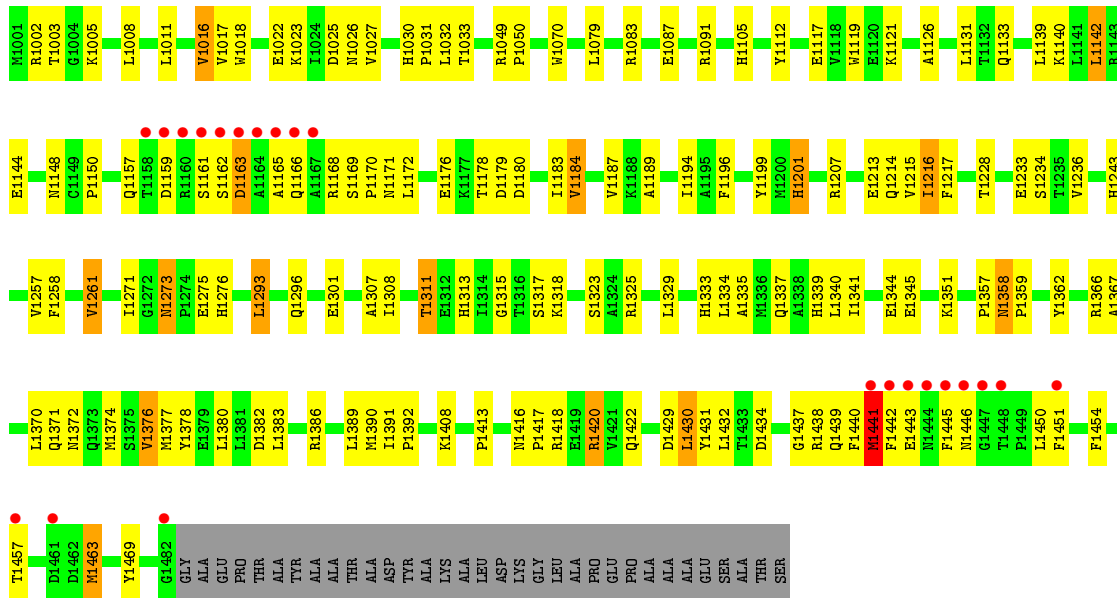


- Molecule 1: Chlorophenol-4-monooxygenase component 2

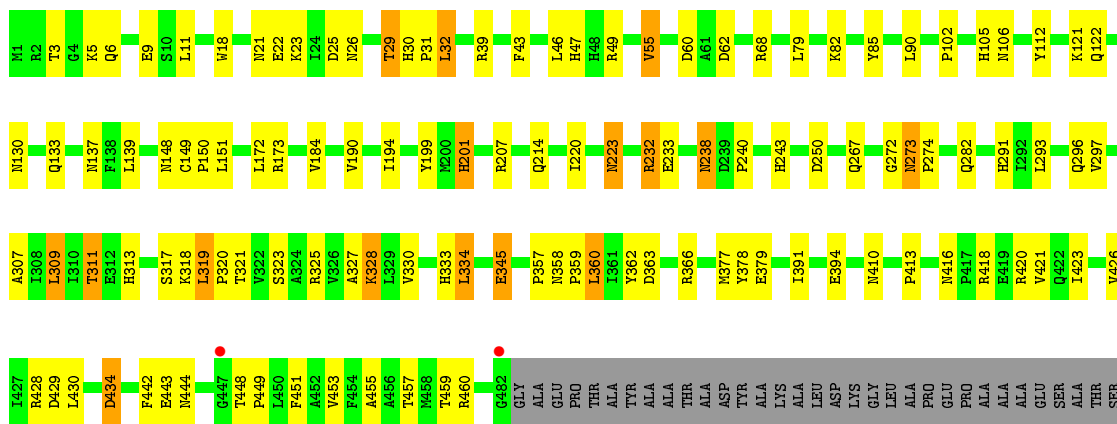




• Molecule 1: Chlorophenol-4-monoxygenase component 2



• Molecule 1: Chlorophenol-4-monoxygenase component 2



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	148.20Å 149.87Å 212.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 2.50 47.34 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.0 (19.94-2.50) 94.8 (47.34-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.51 (at 2.00Å)	Xtrriage
Refinement program	PHENIX ?	Depositor
R, R_{free}	0.168 , 0.224 0.167 , 0.223	Depositor DCC
R_{free} test set	1996 reflections (1.27%)	wwPDB-VP
Wilson B-factor (Å ²)	23.7	Xtrriage
Anisotropy	0.412	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.009 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16516	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.25	0/3930	0.44	0/5335
1	B	0.25	0/3930	0.43	0/5335
1	C	0.24	0/3930	0.42	0/5335
1	D	0.26	0/3930	0.44	0/5335
All	All	0.25	0/15720	0.43	0/21340

There are no bond length outliers.

There are no bond angle outliers.

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	3834	0	3734	123	0
1	B	3834	0	3734	124	0
1	C	3834	0	3734	119	0
1	D	3834	0	3737	94	0
2	A	289	0	0	6	0
2	B	271	0	0	8	0
2	C	245	0	0	3	0
2	D	375	0	0	5	0
All	All	16516	0	14939	419	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 14.

The worst 5 of 419 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:C:1178:THR:HG22	1:C:1180:ASP:H	1.14	1.05
1:A:3079:LEU:HD11	1:A:3357:PRO:HG3	1.44	0.97
1:A:3451:PHE:HD1	1:D:321:THR:HG1	0.98	0.97
1:B:2433:THR:HG22	1:B:2436:GLY:H	1.31	0.96
1:A:3185:ASN:HD22	1:A:3260:ASN:H	1.13	0.92

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	480/515 (93%)	453 (94%)	22 (5%)	5 (1%)	15 28
1	B	480/515 (93%)	446 (93%)	24 (5%)	10 (2%)	7 11
1	C	480/515 (93%)	458 (95%)	19 (4%)	3 (1%)	25 43
1	D	480/515 (93%)	472 (98%)	7 (2%)	1 (0%)	47 68
All	All	1920/2060 (93%)	1829 (95%)	72 (4%)	19 (1%)	15 28

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3449	PRO
1	A	3463	MET
1	A	3480	GLU
1	A	3481	PHE
1	B	2190	VAL

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	406/425 (96%)	365 (90%)	41 (10%)	7 14
1	B	406/425 (96%)	379 (93%)	27 (7%)	16 31
1	C	406/425 (96%)	369 (91%)	37 (9%)	9 18
1	D	406/425 (96%)	374 (92%)	32 (8%)	12 24
All	All	1624/1700 (96%)	1487 (92%)	137 (8%)	11 21

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2420	ARG
1	C	1142	LEU
1	D	319	LEU
1	B	2433	THR
1	C	1008	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 104 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2333	HIS
1	C	1201	HIS
1	D	313	HIS
1	B	2339	HIS
1	B	2411	ASN

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](#)

There are no ligands in this entry.

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, 95th 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 ²)	Q<0.9
1	A	482/515 (93%)	-0.19	32 (6%) 18 19	14, 29, 87, 134	0
1	B	482/515 (93%)	-0.15	33 (6%) 17 17	15, 30, 92, 129	0
1	C	482/515 (93%)	-0.24	22 (4%) 32 34	16, 33, 69, 119	0
1	D	482/515 (93%)	-0.61	2 (0%) 92 93	12, 22, 45, 68	0
All	All	1928/2060 (93%)	-0.30	89 (4%) 32 34	12, 29, 69, 134	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3459	THR	14.2
1	B	2450	LEU	12.3
1	C	1165	ALA	10.6
1	A	3162	SER	10.4
1	A	3165	ALA	10.1

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](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.