wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 05:46 am BST

PDB ID	:	2WL3
Title	:	crystal structure of catechol 2,3-dioxygenase
Authors	:	Cho, H.J.; Kim, K.J.; Kang, B.S.
Deposited on	:	2009-06-21
Resolution	:	2.20 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

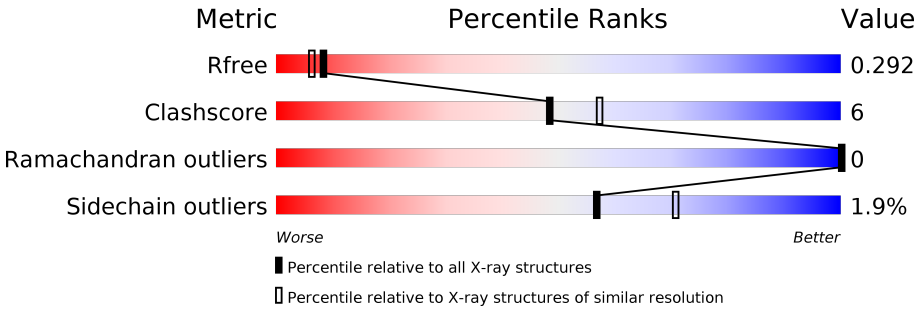
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	305	80% 13% • 6%
1	B	305	79% 15% 6%
1	C	305	77% 16% 7%
1	D	305	78% 13% 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	1291	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9238 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 CATECHOL 2,3-DIOXYGENASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	Se	0	1	0
			2270	1437	401	423	2	7			
1	B	287	Total	C	N	O	S	Se	0	0	0
			2262	1432	398	422	2	8			
1	C	284	Total	C	N	O	S	Se	0	0	0
			2227	1412	390	416	2	7			
1	D	280	Total	C	N	O	S	Se	0	0	0
			2193	1389	387	408	2	7			

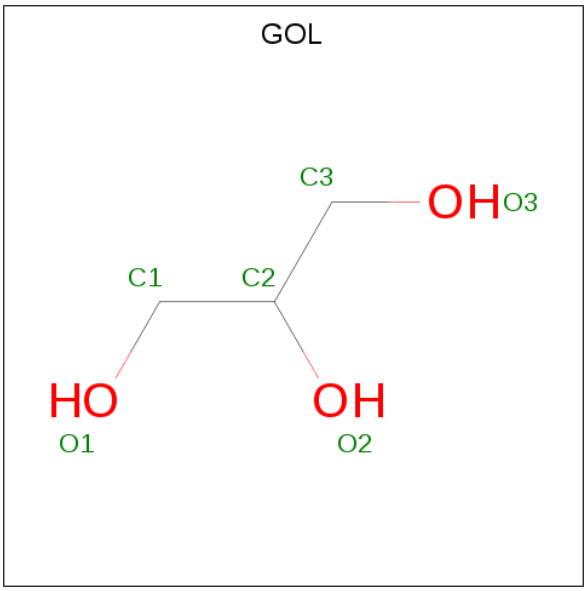
- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	D	1	Total C O 6 3 3	0	0

- Molecule 5 is water.

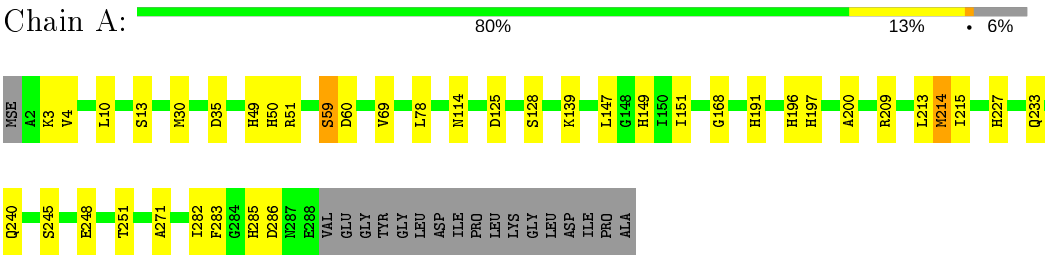
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	76	Total O 76 76	0	0
5	B	63	Total O 63 63	0	0
5	C	53	Total O 53 53	0	0
5	D	59	Total O 59 59	0	0

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.

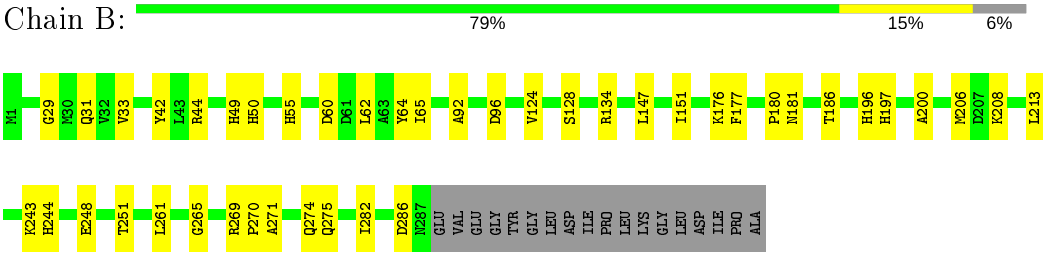
- Molecule 1: CATECHOL 2,3-DIOXYGENASE

Chain A:



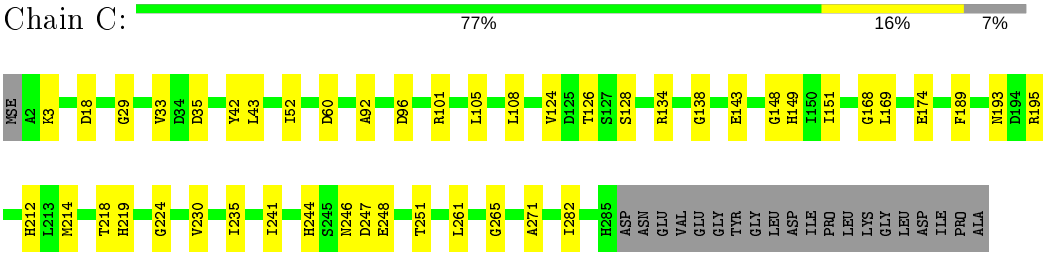
- Molecule 1: CATECHOL 2,3-DIOXYGENASE

Chain B:



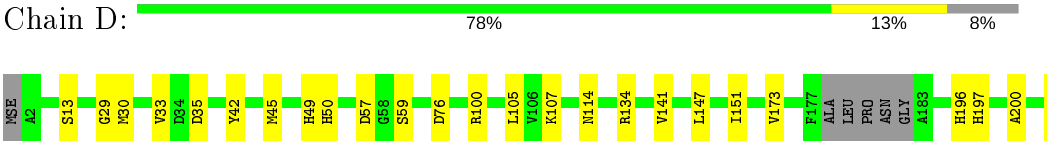
- Molecule 1: CATECHOL 2,3-DIOXYGENASE

Chain C:



- Molecule 1: CATECHOL 2,3-DIOXYGENASE

Chain D:




Q233
G242
K243
H244
D247
E248
A249
F252
Y253
S258
E263
R269
P270
A271
F283
D286
ASN
GLU
VAL
GLU
GLY
TYR
GLY
LEU
ASP
ILE
PRO
LEU
LYS
GLY
LEU
ASP
ILE
PRO
ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, α, β, γ	102.37Å 102.37Å 142.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.24 – 2.20 29.23 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.9 (29.24-2.20) 96.3 (29.23-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	16.53 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.155 , 0.178 0.257 , 0.292	Depositor DCC
R_{free} test set	3601 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	27.7	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 40.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.38$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-l	Xtriage
Reported twinning fraction	0.436 for H, K, L 0.564 for -H, K, -L	Depositor
Outliers	0 of 71651 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9238	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.71% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GOL, CA, FE

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	0.53	0/2329	0.62	1/3151 (0.0%)
1	B	0.59	1/2318 (0.0%)	0.65	0/3138
1	C	0.56	0/2283	0.63	0/3094
1	D	0.58	2/2246 (0.1%)	0.63	0/3040
All	All	0.57	3/9176 (0.0%)	0.63	1/12423 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	124	VAL	CB-CG1	-6.17	1.39	1.52
1	D	141	VAL	CB-CG2	-5.78	1.40	1.52
1	D	173	VAL	CB-CG1	-5.54	1.41	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	51	ARG	NE-CZ-NH1	5.03	122.82	120.30

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	2270	0	2151	28	0
1	B	2262	0	2140	26	0
1	C	2227	0	2094	30	0
1	D	2193	0	2060	23	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	1	0	0	0	0
4	A	12	0	16	4	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	76	0	0	2	0
5	B	63	0	0	3	0
5	C	53	0	0	1	0
5	D	59	0	0	0	0
All	All	9238	0	8485	105	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 105 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:149:HIS:CE1	1:A:214:MSE:HG2	1.98	0.99
1:B:49:HIS:HD2	1:B:50:HIS:ND1	1.82	0.78
1:C:174:GLU:HG3	1:C:189:PHE:HB2	1.69	0.75
1:B:206:MSE:HE1	5:B:2045:HOH:O	1.86	0.74
1:C:151:ILE:HG13	1:C:212:HIS:CE1	2.23	0.73

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	286/305 (94%)	278 (97%)	8 (3%)	0	100	100
1	B	285/305 (93%)	279 (98%)	6 (2%)	0	100	100
1	C	282/305 (92%)	272 (96%)	10 (4%)	0	100	100
1	D	276/305 (90%)	268 (97%)	8 (3%)	0	100	100
All	All	1129/1220 (92%)	1097 (97%)	32 (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	233/240 (97%)	228 (98%)	5 (2%)	53	67
1	B	232/240 (97%)	229 (99%)	3 (1%)	69	81
1	C	227/240 (95%)	221 (97%)	6 (3%)	46	58
1	D	222/240 (92%)	219 (99%)	3 (1%)	67	80
All	All	914/960 (95%)	897 (98%)	17 (2%)	57	71

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	286	ASP
1	C	18	ASP
1	C	169	LEU
1	B	274	GLN
1	D	59	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	233	GLN
1	C	114	ASN
1	D	114	ASN
1	B	212	HIS
1	D	197	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

Of 10 ligands modelled in this entry, 5 are monoatomic - leaving 5 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	GOL	C	1287	-	5,5,5	0.44	0	5,5,5	0.40	0
4	GOL	B	1289	-	5,5,5	0.47	0	5,5,5	0.31	0
4	GOL	A	1292	-	5,5,5	0.35	0	5,5,5	0.32	0
4	GOL	A	1291	-	5,5,5	0.42	0	5,5,5	0.20	0
4	GOL	D	1288	-	5,5,5	0.34	0	5,5,5	0.38	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
4	GOL	C	1287	-	-	0/4/4/4	-
4	GOL	B	1289	-	-	4/4/4/4	-
4	GOL	A	1292	-	-	0/4/4/4	-
4	GOL	A	1291	-	-	4/4/4/4	-
4	GOL	D	1288	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	1288	GOL	O1-C1-C2-C3
4	D	1288	GOL	C1-C2-C3-O3
4	B	1289	GOL	C1-C2-C3-O3
4	A	1291	GOL	O1-C1-C2-C3
4	A	1291	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1291	GOL	4	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.