



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 19, 2023 – 11:59 PM EDT

PDB ID : 2GBW  
Title : Crystal Structure of Biphenyl 2,3-Dioxygenase from *Sphingomonas yanoikuyae* B1  
Authors : Ferraro, D.J.; Brown, E.N.; Yu, C.; Parales, R.E.; Gibson, D.T.; Ramaswamy, S.  
Deposited on : 2006-03-12  
Resolution : 1.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

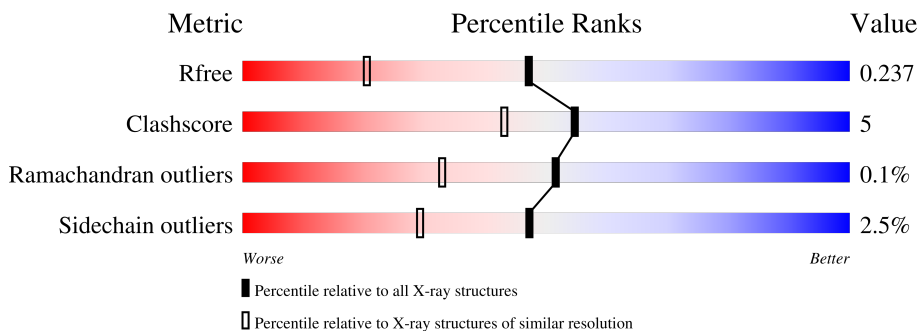
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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  $\geq 3$ , 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	454	
1	C	454	
1	E	454	
2	B	174	
2	D	174	
2	F	174	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 17030 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 Biphenyl 2,3-Dioxygenase Alpha Subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	449	Total 3569	C 2265	N 622	O 665	S 17	0	1	0
1	C	446	Total 3553	C 2257	N 620	O 659	S 17	0	1	0
1	E	449	Total 3578	C 2271	N 622	O 668	S 17	0	3	0

- Molecule 2 is a protein called Biphenyl 2,3-Dioxygenase Beta Subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	170	Total 1446	C 914	N 260	O 265	S 7	0	2	0
2	D	170	Total 1449	C 916	N 260	O 266	S 7	0	3	0
2	F	170	Total 1436	C 906	N 260	O 265	S 5	0	0	0

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

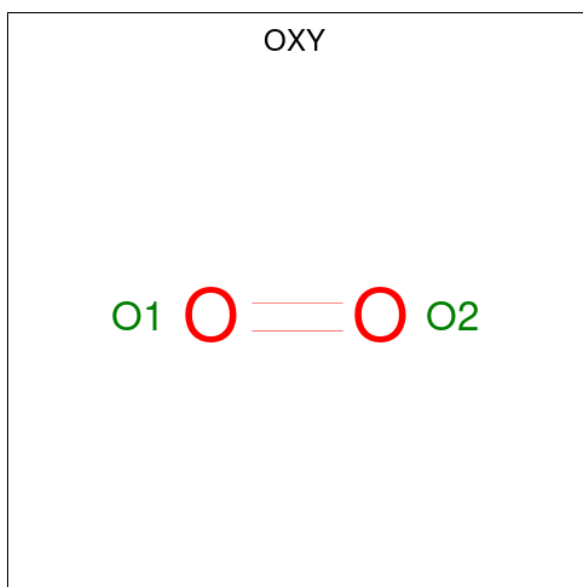
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Fe 1	0	0
3	C	1	Total 1	Fe 1	0	0
3	E	1	Total 1	Fe 1	0	0

- Molecule 4 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe<sub>2</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	Fe	S	0	0
			4	2	2		
4	C	1	Total	Fe	S	0	0
			4	2	2		
4	E	1	Total	Fe	S	0	0
			4	2	2		

- Molecule 5 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O<sub>2</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	O	0	0
			2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total O 2 2	0	0
5	E	1	Total O 2 2	0	0

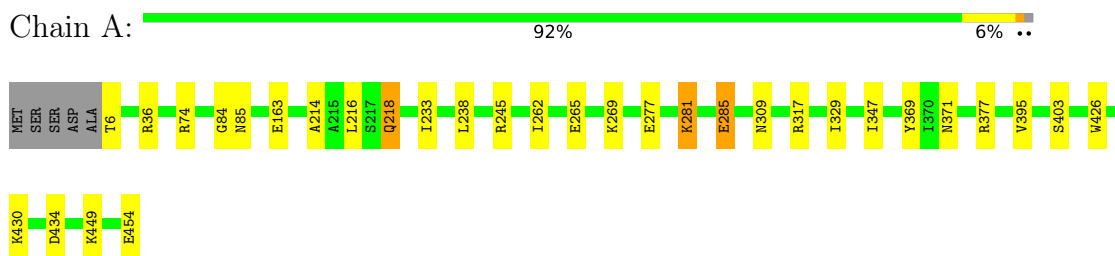
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	502	Total O 502 502	0	0
6	B	250	Total O 250 250	0	0
6	C	441	Total O 441 441	0	0
6	D	232	Total O 232 232	0	0
6	E	354	Total O 354 354	0	0
6	F	199	Total O 199 199	0	0

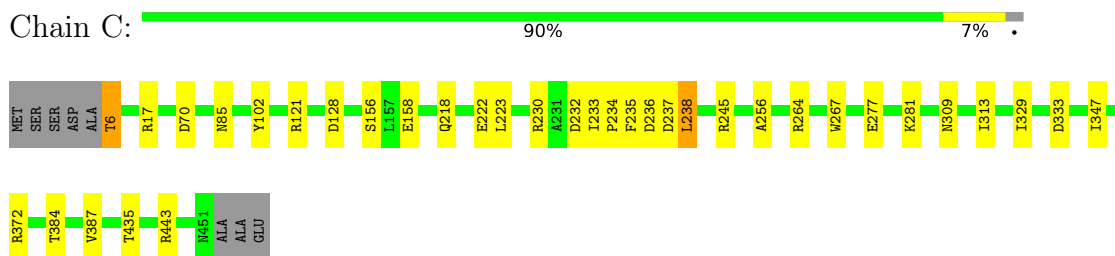
### 3 Residue-property plots [i](#)

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.

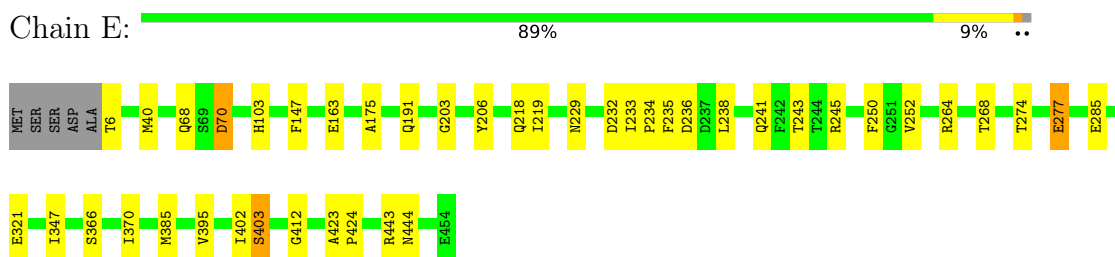
- Molecule 1: Biphenyl 2,3-Dioxygenase Alpha Subunit



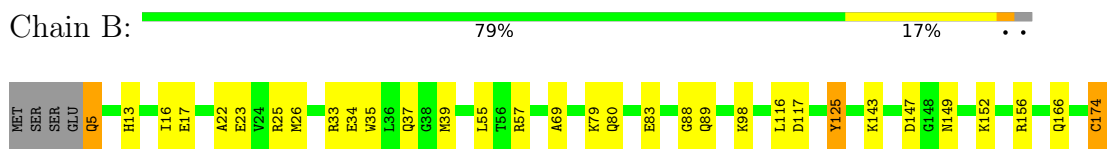
- Molecule 1: Biphenyl 2,3-Dioxygenase Alpha Subunit




- Molecule 1: Biphenyl 2,3-Dioxygenase Alpha Subunit



- Molecule 2: Biphenyl 2,3-Dioxygenase Beta Subunit




- Molecule 2: Biphenyl 2,3-Dioxygenase Beta Subunit

Chain D:  83% 13% ...



- Molecule 2: Biphenyl 2,3-Dioxygenase Beta Subunit

Chain F:  83% 11% ..



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.96Å 134.96Å 219.89Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.80 – 1.70 19.79 – 1.70	Depositor EDS
% Data completeness (in resolution range)	95.2 (19.80-1.70) 95.2 (19.79-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 1.70Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.188 , 0.228 0.199 , 0.237	Depositor DCC
$R_{free}$ test set	12180 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.6	Xtrriage
Anisotropy	0.185	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 50.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17030	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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: OXY, FES, 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	0.80	1/3677 (0.0%)	0.76	2/4990 (0.0%)
1	C	0.81	0/3661	0.80	3/4967 (0.1%)
1	E	0.70	0/3692	0.72	0/5010
2	B	0.86	1/1487 (0.1%)	0.85	0/2009
2	D	0.82	0/1493	0.84	0/2017
2	F	0.80	1/1471 (0.1%)	0.86	0/1989
All	All	0.79	3/15481 (0.0%)	0.79	5/20982 (0.0%)

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
2	D	0	1
2	F	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	163	GLU	CD-OE2	-5.51	1.19	1.25
2	F	125	TYR	CE1-CZ	-5.46	1.31	1.38
2	B	174	CYS	C-OXT	-5.44	1.13	1.23

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	70	ASP	CB-CG-OD1	7.42	124.98	118.30
1	A	74	ARG	NE-CZ-NH1	6.23	123.42	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	17	ARG	NE-CZ-NH1	6.07	123.33	120.30
1	C	333	ASP	CB-CG-OD1	5.57	123.31	118.30
1	A	74	ARG	NE-CZ-NH2	-5.15	117.73	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	87	SER	Peptide
2	F	147	ASP	Peptide

## 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	3569	0	3368	26	0
1	C	3553	0	3360	29	0
1	E	3578	0	3379	37	0
2	B	1446	0	1400	35	0
2	D	1449	0	1405	27	0
2	F	1436	0	1382	19	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
4	A	4	0	0	0	0
4	C	4	0	0	0	0
4	E	4	0	0	1	0
5	A	2	0	0	0	0
5	C	2	0	0	0	0
5	E	2	0	0	0	0
6	A	502	0	0	7	0
6	B	250	0	0	7	0
6	C	441	0	0	4	0
6	D	232	0	0	5	0
6	E	354	0	0	8	0
6	F	199	0	0	3	0
All	All	17030	0	14294	152	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.

The worst 5 of 152 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)
2:D:26[B]:MET:SD	2:D:39[B]:MET:CE	2.31	1.19
2:B:26[A]:MET:HG3	2:B:39[A]:MET:HE1	1.10	1.04
2:F:146:ARG:HH11	2:F:146:ARG:HG2	1.19	1.03
2:B:26[A]:MET:HG3	2:B:39[A]:MET:CE	1.91	1.00
1:C:233:ILE:N	1:C:233:ILE:HD12	1.80	0.96

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	448/454 (99%)	434 (97%)	14 (3%)	0	100	100
1	C	445/454 (98%)	427 (96%)	18 (4%)	0	100	100
1	E	450/454 (99%)	433 (96%)	16 (4%)	1 (0%)	47	30
2	B	170/174 (98%)	166 (98%)	4 (2%)	0	100	100
2	D	171/174 (98%)	166 (97%)	4 (2%)	1 (1%)	25	11
2	F	168/174 (97%)	164 (98%)	4 (2%)	0	100	100
All	All	1852/1884 (98%)	1790 (97%)	60 (3%)	2 (0%)	51	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	88	GLY
1	E	203	GLY

### 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	372/375 (99%)	363 (98%)	9 (2%)	49 31
1	C	371/375 (99%)	362 (98%)	9 (2%)	49 31
1	E	374/375 (100%)	369 (99%)	5 (1%)	69 56
2	B	155/157 (99%)	150 (97%)	5 (3%)	39 20
2	D	156/157 (99%)	151 (97%)	5 (3%)	39 20
2	F	153/157 (98%)	145 (95%)	8 (5%)	23 8
All	All	1581/1596 (99%)	1540 (97%)	41 (3%)	47 28

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	70	ASP
2	F	125	TYR
1	E	245	ARG
1	E	403	SER
2	F	143	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	125	ASN
1	E	246	HIS
1	C	218	GLN
1	E	378	ASN
1	E	200	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 3 are monoatomic - leaving 6 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	FES	C	455	1	0,4,4	-	-	-		
4	FES	A	455	1	0,4,4	-	-	-		
4	FES	E	455	1	0,4,4	-	-	-		
5	OXY	E	457	3	1,1,1	0.21	0	-		
5	OXY	C	457	3	1,1,1	0.15	0	-		
5	OXY	A	457	3	1,1,1	0.13	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	FES	A	455	1	-	-	0/1/1/1
4	FES	E	455	1	-	-	0/1/1/1
4	FES	C	455	1	-	-	0/1/1/1

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	455	FES	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.