



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 02:34 pm BST

PDB ID : 2G19  
Title : Cellular Oxygen Sensing: Crystal Structure of Hypoxia-Inducible Factor Prolyl Hydroxylase (PHD2)  
Authors : Syed, R.S.; Li, V.  
Deposited on : 2006-02-14  
Resolution : 1.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
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

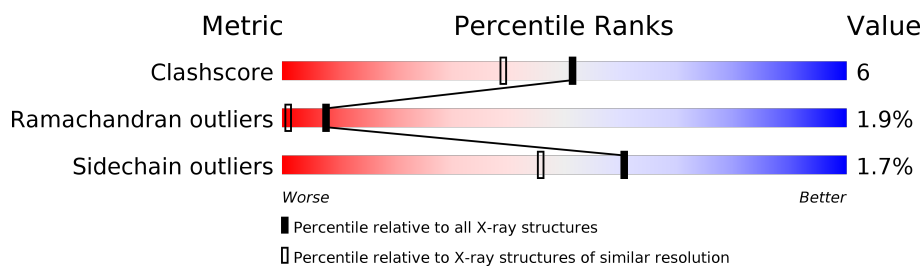
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
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 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	244	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1863 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 Egl nine homolog 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	216	1708	1080	300	317	11	0	0	0

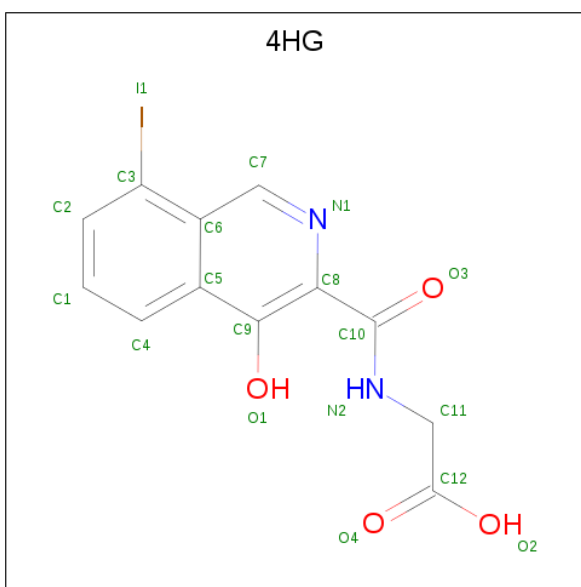
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	418	HIS	-	EXPRESSION TAG	UNP Q9GZT9
A	419	HIS	-	EXPRESSION TAG	UNP Q9GZT9
A	420	HIS	-	EXPRESSION TAG	UNP Q9GZT9
A	421	HIS	-	EXPRESSION TAG	UNP Q9GZT9
A	422	HIS	-	EXPRESSION TAG	UNP Q9GZT9
A	423	HIS	-	EXPRESSION TAG	UNP Q9GZT9

- 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		

- Molecule 3 is N-[(4-HYDROXY-8-ODOISOQUINOLIN-3-YL)CARBONYL]GLYCINE (three-letter code: 4HG) (formula: C<sub>12</sub>H<sub>9</sub>IN<sub>2</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	I	N	O	0	0
			19	12	1	2	4		

- Molecule 4 is water.

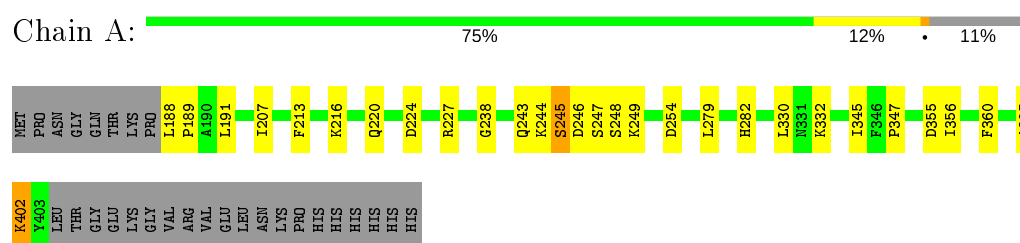
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	135	Total	O	0	0
			135	135		

### 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.

Note EDS was not executed.

- Molecule 1: Egl nine homolog 1



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.85Å 110.85Å 40.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.00 – 1.70	Depositor
% Data completeness (in resolution range)	95.6 (48.00-1.70)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	CNX	Depositor
R, $R_{free}$	0.216 , 0.253	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1863	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 4HG, 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	0.35	0/1747	0.66	0/2358

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

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	1708	0	1676	18	0
2	A	1	0	0	0	0
3	A	19	0	7	1	0
4	A	135	0	0	2	0
All	All	1863	0	1683	19	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.

All (19) 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)
3:A:801:4HG:I1	4:A:602:HOH:O	2.56	0.92
1:A:245:SER:HB3	1:A:249:LYS:HB2	1.58	0.84
1:A:399:ALA:O	1:A:402:LYS:HB3	2.02	0.59
1:A:245:SER:HB3	1:A:249:LYS:CB	2.30	0.59
1:A:188:LEU:HB3	1:A:189:PRO:HD3	1.92	0.51
1:A:207:ILE:HD11	1:A:347:PRO:HG3	1.94	0.50
1:A:243:GLN:O	1:A:245:SER:N	2.47	0.48
1:A:355:ASP:O	1:A:356:ILE:HD13	2.15	0.46
1:A:238:GLY:HA2	1:A:254:ASP:O	2.18	0.44
1:A:216:LYS:HD3	1:A:360:PHE:HE1	1.83	0.44
1:A:345:ILE:O	1:A:347:PRO:HD3	2.18	0.44
1:A:191:LEU:HD12	1:A:282:HIS:HB3	2.00	0.43
1:A:398:ARG:O	1:A:401:VAL:HG23	2.18	0.43
1:A:246:ASP:O	1:A:248:SER:N	2.52	0.43
1:A:332:LYS:HA	1:A:360:PHE:HB2	2.00	0.42
1:A:243:GLN:C	1:A:245:SER:N	2.72	0.41
1:A:385:ALA:HB3	4:A:617:HOH:O	2.20	0.41
1:A:224:ASP:OD1	1:A:227:ARG:NH2	2.53	0.41
1:A:213:PHE:CE1	1:A:279:LEU:HD11	2.57	0.40

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	214/244 (88%)	203 (95%)	7 (3%)	4 (2%)	<b>8</b> <b>1</b>

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	244	LYS
1	A	247	SER
1	A	245	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	402	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	181/206 (88%)	178 (98%)	3 (2%)	60 46

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	220	GLN
1	A	330	LEU
1	A	401	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	205	HIS
1	A	220	GLN
1	A	221	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

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
3	4HG	A	801	2	17,20,20	2.87	11 (64%)	19,28,28	2.13	5 (26%)

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
3	4HG	A	801	2	-	4/7/9/9	0/2/2/2

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	801	4HG	C8-N1	4.44	1.42	1.34
3	A	801	4HG	C7-N1	4.25	1.41	1.32
3	A	801	4HG	C9-C8	4.22	1.46	1.39
3	A	801	4HG	C1-C4	4.19	1.46	1.36
3	A	801	4HG	C7-C6	3.99	1.46	1.42
3	A	801	4HG	C2-C3	3.58	1.46	1.37
3	A	801	4HG	C1-C2	3.07	1.45	1.38
3	A	801	4HG	C3-I1	-2.89	2.05	2.10
3	A	801	4HG	C11-N2	-2.34	1.41	1.46
3	A	801	4HG	C4-C5	2.28	1.46	1.42
3	A	801	4HG	C3-C6	2.16	1.45	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	801	4HG	C6-C3-I1	-4.73	114.90	120.86
3	A	801	4HG	C8-C10-N2	4.65	120.21	115.67

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	801	4HG	C2-C3-C6	3.17	126.68	121.08
3	A	801	4HG	C1-C2-C3	-2.72	117.11	120.78
3	A	801	4HG	O1-C9-C5	2.59	122.13	116.64

There are no chirality outliers.

All (4) torsion outliers are listed below:

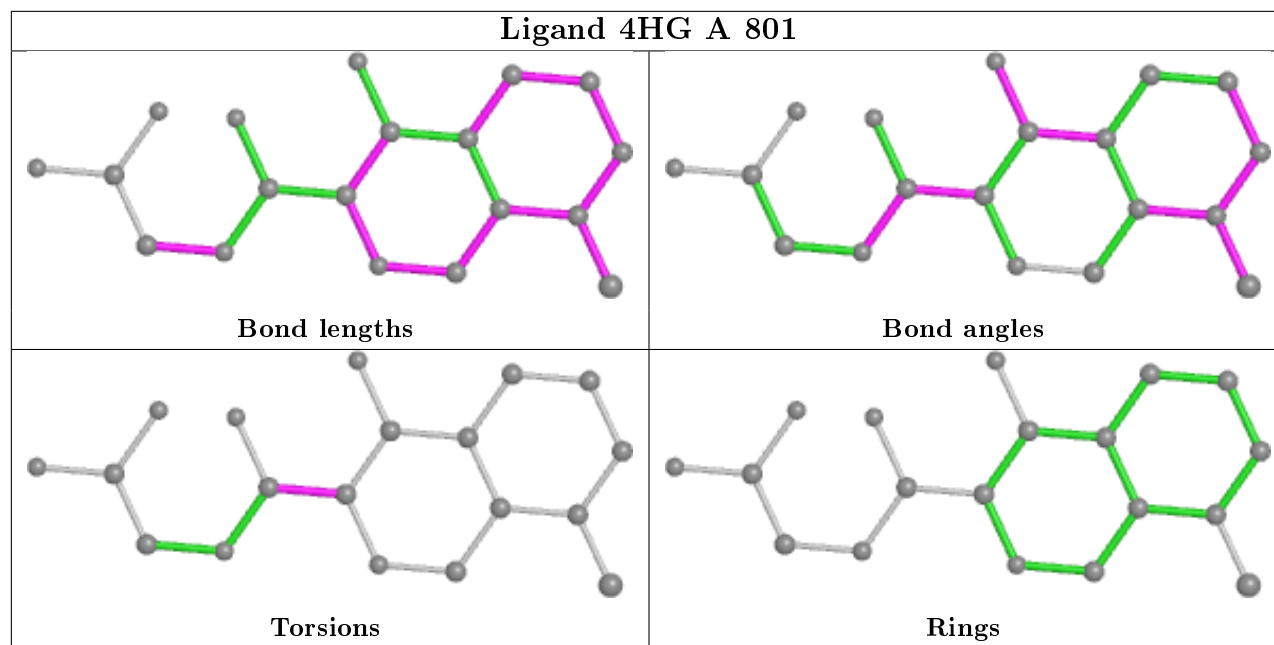
Mol	Chain	Res	Type	Atoms
3	A	801	4HG	O3-C10-C8-N1
3	A	801	4HG	N2-C10-C8-C9
3	A	801	4HG	N2-C10-C8-N1
3	A	801	4HG	O3-C10-C8-C9

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	801	4HG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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.