



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 03:21 PM EDT

PDB ID : 6AZU  
Title : Holo IDO1 crystal structure  
Authors : Lewis, H.A.; Yan, C.  
Deposited on : 2017-09-13  
Resolution : 2.82 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

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

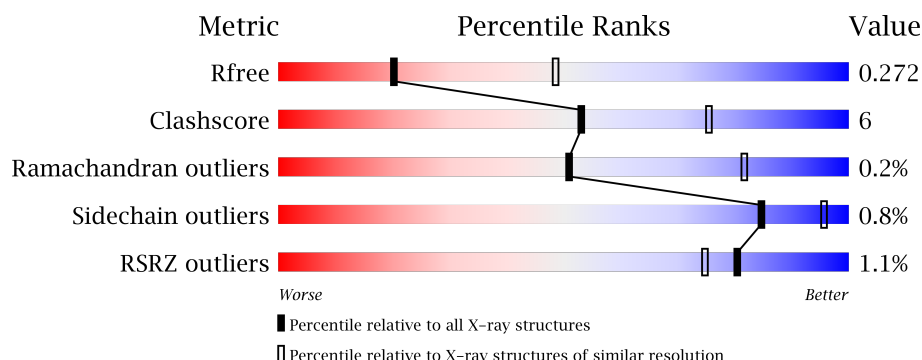
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.82 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))
$R_{free}$	100719	2917 (2.84-2.80)
Clashscore	112137	3382 (2.84-2.80)
Ramachandran outliers	110173	3324 (2.84-2.80)
Sidechain outliers	110143	3326 (2.84-2.80)
RSRZ outliers	101464	2948 (2.84-2.80)

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. 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	402	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 14%, green 78%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>1%</span> <span>78%</span> <span>14%</span> <span>8%</span> </div> </div>
1	B	402	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 79%, yellow 12%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>79%</span> <span>12%</span> <span>9%</span> </div> </div>
1	C	402	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 11%, green 80%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>1%</span> <span>80%</span> <span>11%</span> <span>9%</span> </div> </div>
1	D	402	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 13%, green 78%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>1%</span> <span>78%</span> <span>13%</span> <span>8%</span> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10993 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.

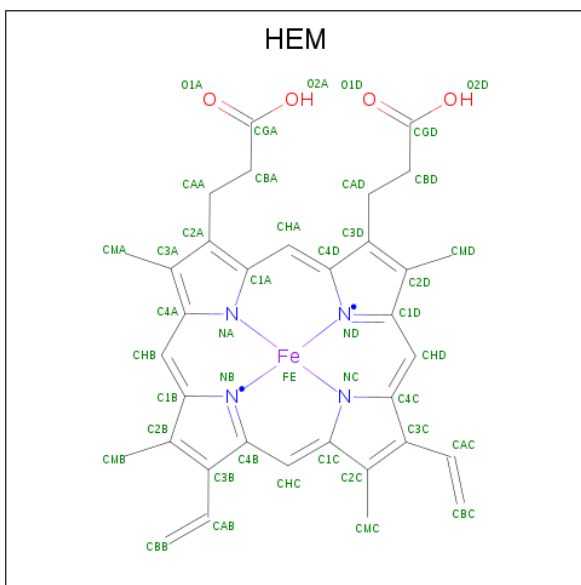
- Molecule 1 is a protein called Indoleamine 2,3-dioxygenase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	370	Total	C	N	O	S	0	0	0
			2730	1777	458	479	16			
1	B	366	Total	C	N	O	S	0	0	0
			2709	1763	455	475	16			
1	C	366	Total	C	N	O	S	0	0	0
			2686	1746	449	475	16			
1	D	370	Total	C	N	O	S	0	0	0
			2676	1736	450	474	16			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	-	expression tag	UNP P14902
A	3	SER	-	expression tag	UNP P14902
A	4	HIS	-	expression tag	UNP P14902
B	2	GLY	-	expression tag	UNP P14902
B	3	SER	-	expression tag	UNP P14902
B	4	HIS	-	expression tag	UNP P14902
C	2	GLY	-	expression tag	UNP P14902
C	3	SER	-	expression tag	UNP P14902
C	4	HIS	-	expression tag	UNP P14902
D	2	GLY	-	expression tag	UNP P14902
D	3	SER	-	expression tag	UNP P14902
D	4	HIS	-	expression tag	UNP P14902

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	D	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

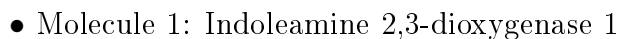
- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		



- Molecule 1: Indoleamine 2,3-dioxygenase 1



SER  
GLU  
ASP  
PRO  
SER  
LYS  
LEU  
GLU  
ALA  
LYS  
GLY  
THR  
GLY  
G381  
M385  
T394  
K401  
GLU  
GLY

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.11Å 177.51Å 101.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.02 – 2.82 34.02 – 2.82	Depositor EDS
% Data completeness (in resolution range)	99.8 (34.02-2.82) 99.8 (34.02-2.82)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.81Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.211 , 0.272 0.212 , 0.272	Depositor DCC
$R_{free}$ test set	2664 reflections (5.10%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.2	Xtriage
Anisotropy	0.872	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 73.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10993	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.86% 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: HEM, SO4

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	0.46	0/2796	0.61	1/3813 (0.0%)
1	B	0.49	0/2774	0.63	0/3779
1	C	0.43	0/2751	0.61	0/3751
1	D	0.44	0/2742	0.61	0/3742
All	All	0.45	0/11063	0.62	1/15085 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	62	LEU	CA-CB-CG	5.65	128.29	115.30

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	2730	0	2512	34	0
1	B	2709	0	2471	31	0
1	C	2686	0	2429	30	0
1	D	2676	0	2377	34	0
2	A	43	0	30	3	0
2	B	43	0	30	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	43	0	30	3	0
2	D	43	0	30	2	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	D	5	0	0	0	0
All	All	10993	0	9909	130	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 130 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:51:GLU:HG2	1:C:155:ARG:HH21	1.34	0.93
1:C:194:ASP:OD2	1:D:193:ARG:NH2	2.18	0.76
1:B:80:ARG:NH2	1:B:121:PRO:O	2.19	0.76
1:C:176:SER:HA	1:C:179:LYS:HE2	1.66	0.76
1:A:193:ARG:HH12	1:A:197:LEU:HD12	1.55	0.72

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	366/402 (91%)	350 (96%)	14 (4%)	2 (0%)	32	65
1	B	360/402 (90%)	349 (97%)	11 (3%)	0	100	100
1	C	360/402 (90%)	347 (96%)	13 (4%)	0	100	100
1	D	366/402 (91%)	352 (96%)	13 (4%)	1 (0%)	44	76
All	All	1452/1608 (90%)	1398 (96%)	51 (4%)	3 (0%)	51	82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	263	SER
1	A	13	LYS
1	A	285	GLY

### 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	245/348 (70%)	242 (99%)	3 (1%)	75	93
1	B	240/348 (69%)	237 (99%)	3 (1%)	73	92
1	C	237/348 (68%)	237 (100%)	0	100	100
1	D	227/348 (65%)	225 (99%)	2 (1%)	82	95
All	All	949/1392 (68%)	941 (99%)	8 (1%)	85	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	64	MET
1	D	214	PHE
1	B	323	LYS
1	A	356	ILE
1	B	214	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	280	GLN
1	A	293	GLN
1	B	293	GLN
1	C	212	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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

8 ligands are modelled in this entry.

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$
2	HEM	A	501	1	28,50,50	2.54	12 (42%)	17,82,82	1.86	7 (41%)
3	SO4	A	502	-	4,4,4	0.12	0	6,6,6	0.18	0
2	HEM	B	501	1	28,50,50	2.43	10 (35%)	17,82,82	1.45	5 (29%)
3	SO4	B	502	-	4,4,4	0.15	0	6,6,6	0.09	0
2	HEM	C	501	1	28,50,50	2.33	11 (39%)	17,82,82	1.61	4 (23%)
3	SO4	C	502	-	4,4,4	0.17	0	6,6,6	0.20	0
2	HEM	D	501	1	28,50,50	2.26	9 (32%)	17,82,82	1.67	4 (23%)
3	SO4	D	502	-	4,4,4	0.12	0	6,6,6	0.23	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
2	HEM	A	501	1	-	0/6/54/54	0/0/8/8

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SO4	A	502	-	-	0/0/0/0	0/0/0/0
2	HEM	B	501	1	-	0/6/54/54	0/0/8/8
3	SO4	B	502	-	-	0/0/0/0	0/0/0/0
2	HEM	C	501	1	-	0/6/54/54	0/0/8/8
3	SO4	C	502	-	-	0/0/0/0	0/0/0/0
2	HEM	D	501	1	-	0/6/54/54	0/0/8/8
3	SO4	D	502	-	-	0/0/0/0	0/0/0/0

The worst 5 of 42 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	HEM	C3B-C2B	-4.80	1.34	1.40
2	A	501	HEM	C3C-C2C	-4.55	1.34	1.40
2	A	501	HEM	C3B-C2B	-4.38	1.34	1.40
2	C	501	HEM	C3B-C2B	-4.38	1.34	1.40
2	D	501	HEM	C3C-C2C	-4.32	1.34	1.40

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	HEM	CMA-C3A-C4A	-3.56	123.00	128.46
2	D	501	HEM	CMA-C3A-C4A	-3.02	123.82	128.46
2	B	501	HEM	CBD-CAD-C3D	-2.95	106.83	112.47
2	C	501	HEM	CBD-CAD-C3D	-2.61	107.49	112.47
2	A	501	HEM	C1D-C2D-C3D	-2.39	105.33	107.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	HEM	3	0
2	B	501	HEM	1	0
2	C	501	HEM	3	0
2	D	501	HEM	2	0

## 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	370/402 (92%)	-0.38	5 (1%) 75 69	28, 45, 74, 161	0
1	B	366/402 (91%)	-0.38	2 (0%) 90 88	28, 43, 68, 107	0
1	C	366/402 (91%)	-0.27	5 (1%) 75 69	29, 47, 75, 122	0
1	D	370/402 (92%)	-0.31	4 (1%) 80 75	31, 51, 77, 129	0
All	All	1472/1608 (91%)	-0.34	16 (1%) 80 75	28, 47, 74, 161	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	284	GLY	4.2
1	C	324	GLY	3.5
1	D	326	ALA	3.5
1	A	285	GLY	3.1
1	D	283	ALA	2.9

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	HEM	A	501	43/43	0.94	0.19	0.50	3,28,75,83	0
2	HEM	C	501	43/43	0.93	0.18	0.30	22,45,69,108	0
2	HEM	D	501	43/43	0.94	0.16	-0.20	13,49,79,86	0
2	HEM	B	501	43/43	0.97	0.16	-0.38	9,28,49,53	0
3	SO4	B	502	5/5	0.81	0.41	-	132,132,136,137	0
3	SO4	D	502	5/5	0.87	0.18	-	102,106,110,114	0
3	SO4	A	502	5/5	0.72	0.27	-	117,120,123,127	0
3	SO4	C	502	5/5	0.95	0.18	-	100,100,103,103	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.