



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 11:42 PM GMT

PDB ID : 1YDA  
Title : STRUCTURAL BASIS OF INHIBITOR AFFINITY TO VARIANTS OF HUMAN CARBONIC ANHYDRASE II  
Authors : Nair, S.K.; Christianson, D.W.  
Deposited on : 1994-12-22  
Resolution : 2.10 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

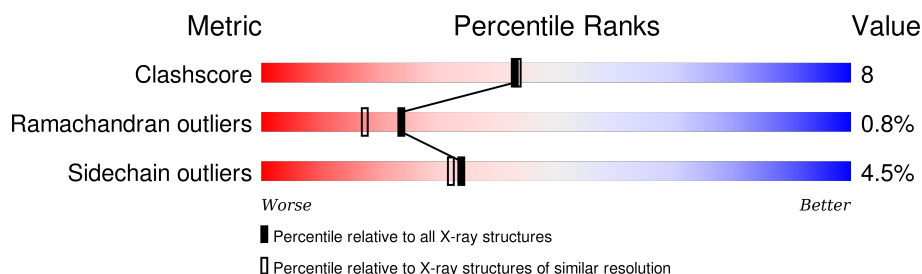
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	259	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2123 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 CARBONIC ANHYDRASE II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	0	0	0
			2029	1302	347	378	2			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	198	GLU	LEU	CONFLICT	UNP P00918

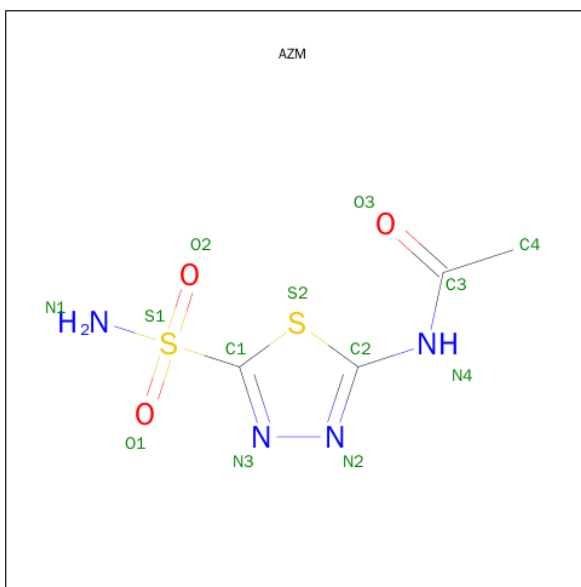
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

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

- Molecule 4 is 5-ACETAMIDO-1,3,4-THIADIAZOLE-2-SULFONAMIDE (three-letter code: AZM) (formula: C<sub>4</sub>H<sub>6</sub>N<sub>4</sub>O<sub>3</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	A	1	13	4	4	3	2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	79	Total	O	0	0
			79	79		

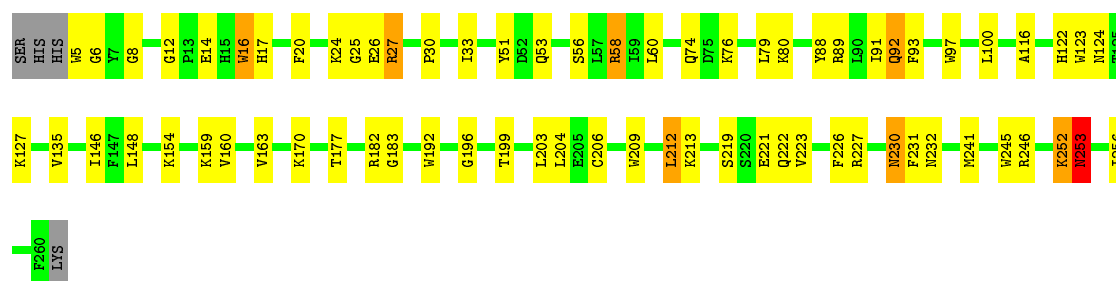
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

#### • Molecule 1: CARBONIC ANHYDRASE II

Chain A: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.70 Å   41.70 Å   73.00 Å 90.00°   104.60°   90.00°	Depositor
Resolution (Å)	6.50 – 2.10	Depositor
% Data completeness (in resolution range)	(Not available) (6.50-2.10)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.181 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2123	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	11.0	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: ZN, AZM, HG

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	0.88	0/2089	1.64	40/2836 (1.4%)

There are no bond length outliers.

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	209	TRP	CD1-CG-CD2	10.09	114.37	106.30
1	A	51	TYR	CB-CG-CD2	-10.03	114.98	121.00
1	A	182	ARG	NE-CZ-NH1	10.01	125.31	120.30
1	A	27	ARG	NE-CZ-NH1	9.82	125.21	120.30
1	A	227	ARG	NE-CZ-NH2	-8.99	115.81	120.30
1	A	209	TRP	CE2-CD2-CG	-8.31	100.65	107.30
1	A	182	ARG	NE-CZ-NH2	-8.13	116.24	120.30
1	A	97	TRP	CD1-CG-CD2	8.12	112.80	106.30
1	A	5	TRP	CD1-CG-CD2	7.96	112.67	106.30
1	A	123	TRP	CD1-CG-CD2	7.76	112.51	106.30
1	A	16	TRP	CD1-CG-CD2	7.66	112.43	106.30
1	A	97	TRP	CE2-CD2-CG	-7.49	101.31	107.30
1	A	245	TRP	CE2-CD2-CG	-7.27	101.49	107.30
1	A	100	LEU	CA-CB-CG	6.98	131.36	115.30
1	A	123	TRP	CE2-CD2-CG	-6.97	101.72	107.30
1	A	192	TRP	CD1-CG-CD2	6.92	111.84	106.30
1	A	227	ARG	NE-CZ-NH1	6.89	123.74	120.30
1	A	16	TRP	CE2-CD2-CG	-6.88	101.80	107.30
1	A	5	TRP	CE2-CD2-CG	-6.83	101.84	107.30
1	A	27	ARG	NE-CZ-NH2	-6.80	116.90	120.30
1	A	192	TRP	CE2-CD2-CG	-6.70	101.94	107.30
1	A	245	TRP	CD1-CG-CD2	6.61	111.58	106.30
1	A	245	TRP	CB-CG-CD1	-6.54	118.50	127.00
1	A	24	LYS	CB-CG-CD	-6.45	94.84	111.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	88	TYR	CB-CG-CD2	-6.41	117.15	121.00
1	A	209	TRP	CG-CD1-NE1	-6.41	103.69	110.10
1	A	51	TYR	CB-CG-CD1	6.29	124.77	121.00
1	A	25	GLY	CA-C-N	-6.16	103.65	117.20
1	A	123	TRP	CG-CD1-NE1	-6.08	104.02	110.10
1	A	253	ASN	OD1-CG-ND2	-6.07	107.94	121.90
1	A	14	GLU	CA-CB-CG	5.77	126.10	113.40
1	A	97	TRP	CG-CD1-NE1	-5.76	104.34	110.10
1	A	159	LYS	CA-CB-CG	5.73	126.01	113.40
1	A	246	ARG	NE-CZ-NH2	-5.70	117.45	120.30
1	A	230	ASN	CB-CG-ND2	5.51	129.94	116.70
1	A	5	TRP	CG-CD1-NE1	-5.49	104.61	110.10
1	A	16	TRP	CG-CD1-NE1	-5.49	104.61	110.10
1	A	245	TRP	CG-CD2-CE3	5.09	138.49	133.90
1	A	16	TRP	CG-CD2-CE3	5.04	138.43	133.90
1	A	97	TRP	CG-CD2-CE3	5.01	138.41	133.90

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	2029	0	1977	31	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	13	0	6	2	0
5	A	79	0	0	1	0
All	All	2123	0	1983	32	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 8.

All (32) 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:160:VAL:O	1:A:163:VAL:HG12	2.02	0.60
1:A:58:ARG:HH12	1:A:60:LEU:HD13	1.69	0.57
1:A:74:GLN:HB2	1:A:76:LYS:HE3	1.87	0.56
4:A:264:AZM:O3	4:A:264:AZM:S2	2.64	0.55
1:A:58:ARG:HH11	1:A:60:LEU:HD22	1.71	0.55
1:A:146:ILE:HG12	1:A:212:LEU:HD23	1.89	0.54
1:A:199:THR:OG1	4:A:264:AZM:N1	2.42	0.52
1:A:56:SER:O	1:A:177:THR:HA	2.10	0.52
1:A:219:SER:OG	1:A:222:GLN:HG3	2.09	0.52
1:A:89:ARG:O	1:A:122:HIS:HA	2.11	0.51
1:A:33:ILE:HG12	1:A:256:ILE:HD13	1.94	0.49
1:A:91:ILE:HG23	1:A:92:GLN:HG3	1.94	0.48
1:A:58:ARG:NH1	1:A:60:LEU:HD13	2.29	0.48
1:A:253:ASN:C	1:A:253:ASN:HD22	2.19	0.47
1:A:116:ALA:HB3	1:A:148:LEU:HD22	1.97	0.45
1:A:124:ASN:OD1	1:A:127:LYS:HB3	2.16	0.45
1:A:17:HIS:HA	1:A:20:PHE:O	2.16	0.44
1:A:154:LYS:HE3	1:A:183:GLY:O	2.16	0.44
1:A:27:ARG:HH21	1:A:196:GLY:H	1.65	0.43
1:A:6:GLY:O	1:A:12:GLY:HA2	2.18	0.43
1:A:213:LYS:HE2	1:A:213:LYS:HB3	1.73	0.42
1:A:135:VAL:HB	1:A:204:LEU:HD13	2.01	0.42
1:A:252:LYS:O	1:A:253:ASN:HB3	2.19	0.42
1:A:253:ASN:C	1:A:253:ASN:ND2	2.73	0.42
1:A:230:ASN:HB3	1:A:232:ASN:OD1	2.20	0.42
1:A:231:PHE:CE2	1:A:241:MET:HA	2.54	0.41
1:A:170:LYS:HB2	1:A:231:PHE:O	2.20	0.41
1:A:80:LYS:HB2	5:A:285:HOH:O	2.20	0.41
1:A:16:TRP:HE3	1:A:203:LEU:HD11	1.85	0.41
1:A:223:VAL:O	1:A:226:PHE:HB2	2.20	0.41
1:A:204:LEU:HB3	1:A:206:CYS:SG	2.61	0.40
1:A:8:GLY:O	1:A:12:GLY:N	2.54	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	253/259 (98%)	235 (93%)	16 (6%)	2 (1%)	24	17

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	253	ASN
1	A	252	LYS

### 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	220/224 (98%)	210 (96%)	10 (4%)	34	32

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	GLU
1	A	30	PRO
1	A	53	GLN
1	A	58	ARG
1	A	79	LEU
1	A	92	GLN
1	A	93	PHE
1	A	212	LEU
1	A	221	GLU
1	A	253	ASN

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	67	ASN
1	A	137	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	253	ASN

### 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 3 ligands modelled in this entry, 2 are 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
4	AZM	A	264	2	9,13,13	1.71	2 (22%)	7,19,19	7.37	5 (71%)

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	AZM	A	264	2	-	0/2/10/10	0/0/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	264	AZM	C2-N4	-3.09	1.31	1.36
4	A	264	AZM	O2-S1	3.36	1.49	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	264	AZM	O2-S1-O1	-16.53	95.57	118.80
4	A	264	AZM	O1-S1-N1	-2.44	104.13	107.28
4	A	264	AZM	O3-C3-C4	-2.32	117.81	122.06
4	A	264	AZM	C4-C3-N4	2.20	118.10	114.97
4	A	264	AZM	O2-S1-N1	9.49	119.57	107.28

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	264	AZM	2	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

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers

EDS was not executed - this section will therefore be empty.