



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2017 – 10:39 AM EST

PDB ID : 2CA2  
Title : CRYSTALLOGRAPHIC STUDIES OF INHIBITOR BINDING SITES IN HUMAN CARBONIC ANHYDRASE II. A PENTACOORDINATED BINDING OF THE SCN-ION TO THE ZINC AT HIGH P\*H  
Authors : Eriksson, A.E.; Kylsten, P.M.; Jones, T.A.; Liljas, A.  
Deposited on : unknown  
Resolution : 1.90 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

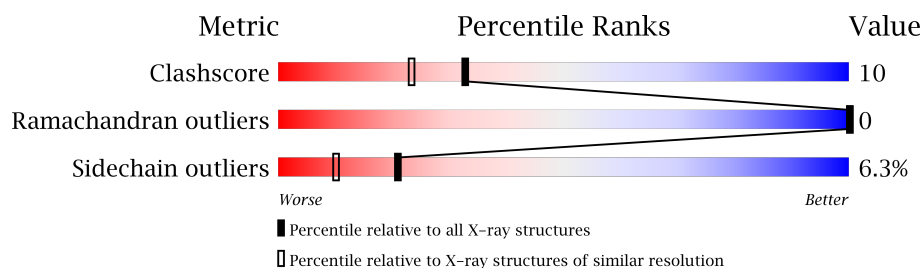
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)

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 [i](#)

There are 5 unique types of molecules in this entry. The entry contains 2216 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.

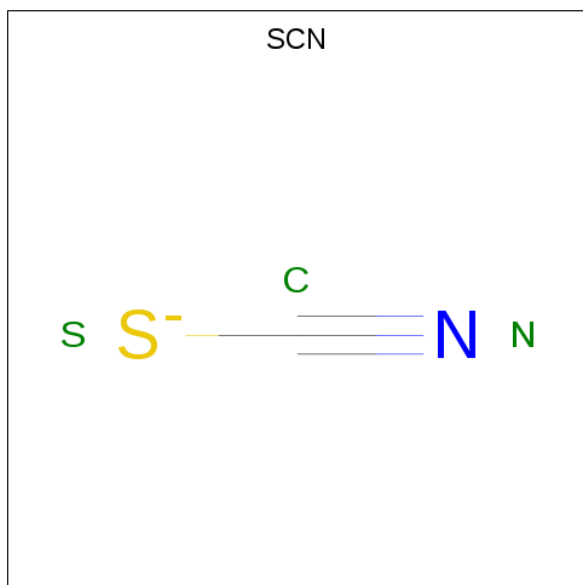
- Molecule 1 is a protein called CARBONIC ANHYDRASE II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	256	Total	C	N	O	S	0	0	0
			2039	1309	350	378	2			

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

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

- Molecule 3 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 5 is water.

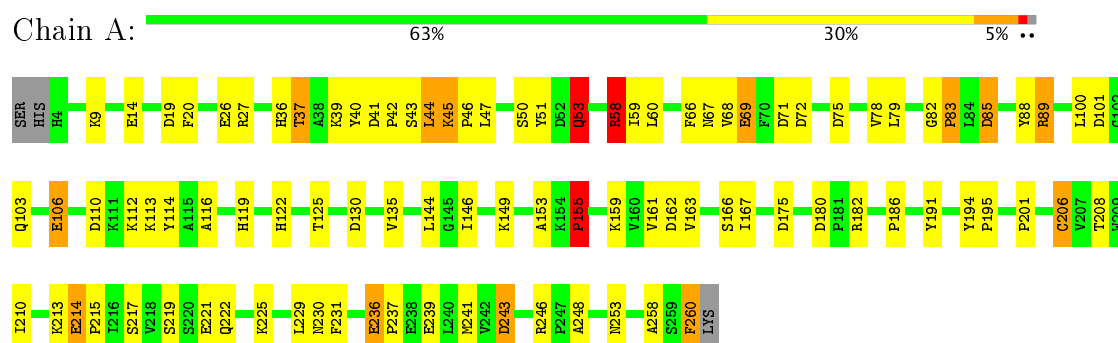
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	172	Total 172	O 172	0	0

### 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 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



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore 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)	(Not available) – 1.90	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-1.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.176 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2216	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	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, SCN, 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	1.32	3/2100 (0.1%)	2.02	64/2851 (2.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	125	THR	C-N	14.02	1.66	1.34
1	A	69	GLU	CD-OE1	-5.11	1.20	1.25
1	A	14	GLU	CD-OE2	5.06	1.31	1.25

All (64) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	ASP	CB-CG-OD1	18.15	134.64	118.30
1	A	58	ARG	NE-CZ-NH1	17.44	129.02	120.30
1	A	27	ARG	NE-CZ-NH2	11.95	126.27	120.30
1	A	182	ARG	NE-CZ-NH1	11.77	126.18	120.30
1	A	71	ASP	CB-CG-OD1	11.44	128.60	118.30
1	A	246	ARG	NE-CZ-NH2	-11.30	114.65	120.30
1	A	19	ASP	CB-CG-OD2	-10.60	108.76	118.30
1	A	101	ASP	CB-CG-OD2	10.33	127.60	118.30
1	A	162	ASP	CB-CG-OD2	-10.25	109.07	118.30
1	A	58	ARG	NE-CZ-NH2	-9.13	115.74	120.30
1	A	51	TYR	CB-CG-CD2	8.93	126.36	121.00
1	A	44	LEU	CA-CB-CG	8.56	134.98	115.30
1	A	89	ARG	NE-CZ-NH1	8.53	124.57	120.30
1	A	85	ASP	CB-CG-OD1	8.45	125.90	118.30
1	A	125	THR	C-N-CA	-8.27	101.02	121.70
1	A	26	GLU	CG-CD-OE1	8.26	134.83	118.30
1	A	110	ASP	CB-CG-OD1	7.95	125.46	118.30
1	A	36	HIS	CA-CB-CG	-7.84	100.27	113.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	ASP	CB-CG-OD2	-7.19	111.83	118.30
1	A	89	ARG	NE-CZ-NH2	-7.07	116.77	120.30
1	A	51	TYR	CB-CG-CD1	-7.06	116.76	121.00
1	A	26	GLU	OE1-CD-OE2	-6.95	114.96	123.30
1	A	100	LEU	CA-CB-CG	6.95	131.29	115.30
1	A	72	ASP	CB-CG-OD1	6.86	124.47	118.30
1	A	71	ASP	CB-CG-OD2	-6.77	112.20	118.30
1	A	180	ASP	CB-CG-OD2	6.67	124.30	118.30
1	A	88	TYR	CB-CG-CD2	-6.55	117.07	121.00
1	A	182	ARG	NE-CZ-NH2	-6.54	117.03	120.30
1	A	162	ASP	CB-CG-OD1	6.50	124.15	118.30
1	A	155	PRO	C-N-CA	6.48	135.91	122.30
1	A	258	ALA	N-CA-CB	6.47	119.16	110.10
1	A	116	ALA	CB-CA-C	6.45	119.77	110.10
1	A	246	ARG	NE-CZ-NH1	6.39	123.50	120.30
1	A	214	GLU	OE1-CD-OE2	-6.17	115.89	123.30
1	A	58	ARG	CD-NE-CZ	6.16	132.22	123.60
1	A	27	ARG	CA-CB-CG	6.14	126.91	113.40
1	A	85	ASP	CB-CG-OD2	-6.06	112.85	118.30
1	A	106	GLU	CG-CD-OE2	-6.00	106.29	118.30
1	A	130	ASP	CB-CG-OD1	5.99	123.69	118.30
1	A	191	TYR	CB-CG-CD1	5.91	124.54	121.00
1	A	110	ASP	OD1-CG-OD2	-5.78	112.31	123.30
1	A	208	THR	N-CA-CB	5.71	121.15	110.30
1	A	243	ASP	CB-CG-OD1	5.70	123.43	118.30
1	A	103	GLN	CG-CD-NE2	5.67	130.31	116.70
1	A	75	ASP	CB-CG-OD1	5.63	123.37	118.30
1	A	260	PHE	CA-C-O	-5.63	108.28	120.10
1	A	144	LEU	CA-CB-CG	5.62	128.23	115.30
1	A	112	LYS	CD-CE-NZ	-5.55	98.93	111.70
1	A	88	TYR	CB-CG-CD1	5.49	124.30	121.00
1	A	85	ASP	CB-CA-C	5.49	121.38	110.40
1	A	236	GLU	OE1-CD-OE2	5.45	129.84	123.30
1	A	19	ASP	O-C-N	5.43	131.38	122.70
1	A	53	GLN	OE1-CD-NE2	5.39	134.30	121.90
1	A	230	ASN	CA-CB-CG	5.37	125.21	113.40
1	A	166	SER	CA-CB-OG	-5.30	96.89	111.20
1	A	37	THR	CA-CB-CG2	5.24	119.74	112.40
1	A	239	GLU	CG-CD-OE1	5.24	128.77	118.30
1	A	68	VAL	N-CA-C	-5.21	96.92	111.00
1	A	14	GLU	CA-CB-CG	5.21	124.87	113.40
1	A	114	TYR	CB-CG-CD2	-5.21	117.88	121.00

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	206	CYS	CA-CB-SG	5.19	123.35	114.00
1	A	175	ASP	OD1-CG-OD2	-5.15	113.52	123.30
1	A	78	VAL	CA-CB-CG1	5.08	118.51	110.90
1	A	69	GLU	OE1-CD-OE2	-5.00	117.30	123.30

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	2039	0	1988	40	0
2	A	1	0	0	0	0
3	A	3	0	0	0	0
4	A	1	0	0	0	0
5	A	172	0	0	9	0
All	All	2216	0	1988	40	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 10.

All (40) 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:248:ALA:HB2	5:A:308:HOH:O	1.21	1.25
1:A:58:ARG:HD2	1:A:69:GLU:OE1	1.49	1.12
1:A:45:LYS:HB3	1:A:46:PRO:CD	2.16	0.76
1:A:243:ASP:OD1	5:A:398:HOH:O	2.12	0.66
1:A:236:GLU:HB3	1:A:237:PRO:HD2	1.80	0.64
1:A:53:GLN:HG3	5:A:383:HOH:O	2.00	0.61
1:A:221:GLU:OE1	5:A:305:HOH:O	2.16	0.61
1:A:85:ASP:O	5:A:339:HOH:O	2.18	0.56
1:A:146:ILE:HD13	1:A:186:PRO:HD3	1.88	0.56
1:A:113:LYS:HE3	5:A:299:HOH:O	2.05	0.55
1:A:40:TYR:HE1	1:A:42:PRO:HB3	1.71	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:45:LYS:HB3	1:A:46:PRO:HD2	1.88	0.55
1:A:225:LYS:HE3	5:A:435:HOH:O	2.06	0.55
1:A:59:ILE:HG12	1:A:167:ILE:HD13	1.90	0.54
1:A:58:ARG:HD2	1:A:69:GLU:CD	2.26	0.54
1:A:161:VAL:CG1	1:A:225:LYS:HD2	2.39	0.53
1:A:159:LYS:HG3	5:A:301:HOH:O	2.10	0.52
1:A:41:ASP:HB3	1:A:44:LEU:HD22	1.93	0.49
1:A:40:TYR:CE1	1:A:42:PRO:HB3	2.47	0.49
1:A:219:SER:OG	1:A:221:GLU:HG2	2.13	0.49
1:A:45:LYS:O	1:A:82:GLY:HA2	2.14	0.48
1:A:58:ARG:CD	1:A:69:GLU:OE1	2.41	0.48
1:A:106:GLU:OE1	1:A:119:HIS:HE1	1.96	0.48
1:A:44:LEU:HD11	1:A:83:PRO:HB3	1.95	0.48
1:A:161:VAL:HG21	1:A:222:GLN:HG2	1.97	0.46
1:A:213:LYS:HD3	1:A:260:PHE:CZ	2.50	0.46
1:A:153:ALA:O	1:A:155:PRO:HD3	2.16	0.45
1:A:231:PHE:CE2	1:A:241:MET:HG3	2.52	0.44
1:A:213:LYS:HD3	1:A:260:PHE:CE2	2.52	0.44
1:A:149:LYS:O	1:A:217:SER:HA	2.16	0.44
1:A:225:LYS:CE	5:A:435:HOH:O	2.65	0.42
1:A:236:GLU:HB3	1:A:237:PRO:CD	2.47	0.42
1:A:60:LEU:O	1:A:66:PHE:HA	2.20	0.41
1:A:135:VAL:O	1:A:206:CYS:SG	2.78	0.41
1:A:20:PHE:CE2	1:A:201:PRO:HB3	2.55	0.41
1:A:214:GLU:HA	1:A:215:PRO:HD3	1.91	0.41
1:A:59:ILE:HA	1:A:67:ASN:O	2.20	0.41
1:A:194:TYR:HA	1:A:195:PRO:HD3	1.91	0.40
1:A:89:ARG:O	1:A:122:HIS:HA	2.21	0.40
1:A:47:LEU:HD11	1:A:210:ILE:HG21	2.03	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	254/259 (98%)	243 (96%)	11 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	221/224 (99%)	207 (94%)	14 (6%)	21	10

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	LYS
1	A	37	THR
1	A	39	LYS
1	A	43	SER
1	A	45	LYS
1	A	50	SER
1	A	53	GLN
1	A	58	ARG
1	A	79	LEU
1	A	83	PRO
1	A	155	PRO
1	A	163	VAL
1	A	229	LEU
1	A	253	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	GLN
1	A	67	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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	$\# Z  > 2$
3	SCN	A	263	4	1,2,2	3.47	1 (100%)	0,1,1	0.00	-

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	SCN	A	263	4	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	263	SCN	C-N	3.47	1.27	1.15

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	125:THR	C	127:LYS	N	1.66

## 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.