



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 24, 2024 – 02:01 PM EST

PDB ID : 7CA2  
Title : ENGINEERING THE HYDROPHOBIC POCKET OF CARBONIC ANHYDRASE II  
Authors : Nair, S.K.; Christianson, D.W.  
Deposited on : 1991-07-09  
Resolution : 2.40 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 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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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) : 2.36

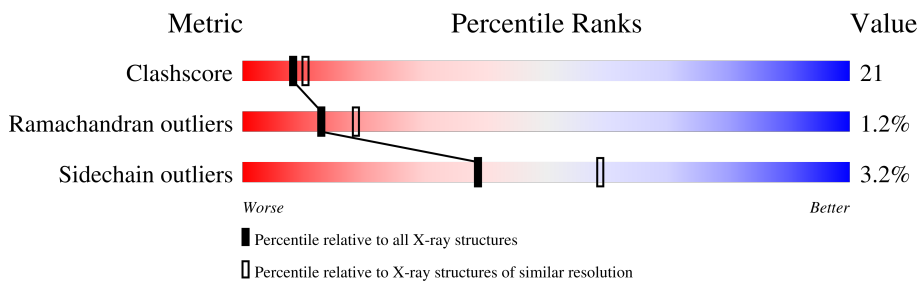
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	260	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2117 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
			Total	C	N	O	S			
1	A	255	2026	1300	347	377	2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	GLY	VAL	conflict	UNP P00918

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

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

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

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

- Molecule 4 is water.

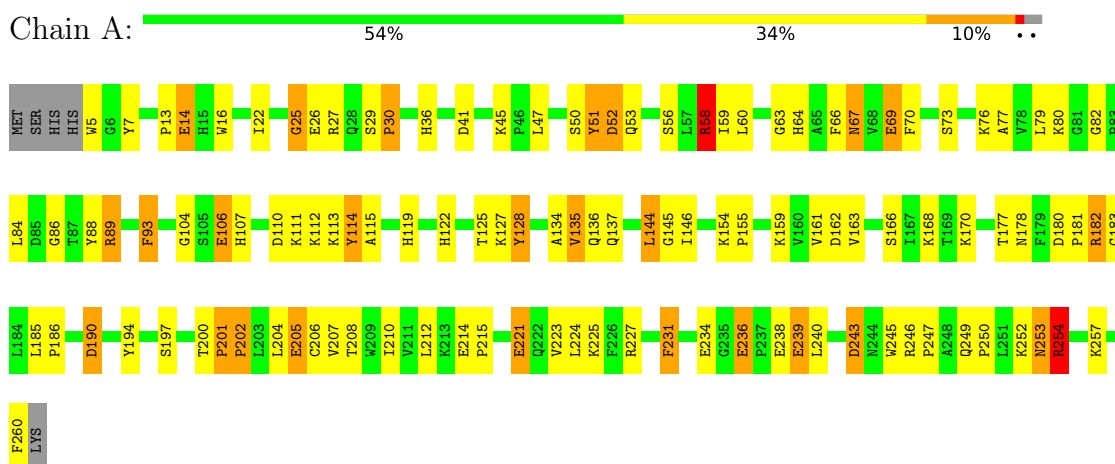
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	89	89	89	0	0

### 3 Residue-property plots

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.

Note EDS was not executed.

- Molecule 1: CARBONIC ANHYDRASE II



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	42.70Å 41.70Å 73.00Å 90.00° 104.60° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.40	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.40)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.170 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2117	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

## 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: HG, ZN

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	1.56	15/2085 (0.7%)	1.77	37/2828 (1.3%)

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
1	A	0	1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	236	GLU	CD-OE2	11.45	1.38	1.25
1	A	205	GLU	CD-OE2	9.88	1.36	1.25
1	A	239	GLU	CD-OE2	9.84	1.36	1.25
1	A	238	GLU	CD-OE1	9.16	1.35	1.25
1	A	234	GLU	CD-OE1	8.45	1.34	1.25
1	A	26	GLU	CD-OE2	7.58	1.33	1.25
1	A	236	GLU	CD-OE1	-6.57	1.18	1.25
1	A	69	GLU	CD-OE2	6.30	1.32	1.25
1	A	221	GLU	CD-OE2	-5.93	1.19	1.25
1	A	221	GLU	CD-OE1	5.62	1.31	1.25
1	A	106	GLU	CD-OE2	5.61	1.31	1.25
1	A	107	HIS	CG-CD2	5.22	1.44	1.35
1	A	73	SER	CA-CB	-5.18	1.45	1.52
1	A	14	GLU	CD-OE1	5.16	1.31	1.25
1	A	30	PRO	N-CA	-5.06	1.38	1.47

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	182	ARG	NE-CZ-NH2	-11.16	114.72	120.30
1	A	202	PRO	N-CD-CG	10.40	118.80	103.20
1	A	30	PRO	N-CD-CG	9.91	118.07	103.20
1	A	30	PRO	N-CA-CB	9.60	114.82	103.30
1	A	201	PRO	CA-C-O	-9.33	97.81	120.20
1	A	202	PRO	CA-N-CD	-9.11	98.75	111.50
1	A	30	PRO	CA-N-CD	-9.10	98.76	111.50
1	A	182	ARG	NE-CZ-NH1	8.82	124.71	120.30
1	A	246	ARG	NE-CZ-NH2	-8.53	116.03	120.30
1	A	29	SER	CA-C-O	-8.51	102.22	120.10
1	A	201	PRO	O-C-N	8.29	136.86	121.10
1	A	29	SER	O-C-N	8.10	136.49	121.10
1	A	202	PRO	N-CA-CB	7.79	112.65	103.30
1	A	227	ARG	NE-CZ-NH1	-7.73	116.43	120.30
1	A	243	ASP	CB-CG-OD1	-7.73	111.34	118.30
1	A	254	ARG	NE-CZ-NH1	7.56	124.08	120.30
1	A	128	TYR	CB-CG-CD1	-6.64	117.02	121.00
1	A	41	ASP	CB-CG-OD2	6.57	124.21	118.30
1	A	236	GLU	CG-CD-OE1	6.03	130.35	118.30
1	A	14	GLU	OE1-CD-OE2	-6.00	116.10	123.30
1	A	14	GLU	CA-CB-CG	5.95	126.49	113.40
1	A	58	ARG	NE-CZ-NH1	5.73	123.17	120.30
1	A	114	TYR	CB-CG-CD1	5.72	124.43	121.00
1	A	223	VAL	CA-CB-CG1	5.61	119.31	110.90
1	A	207	VAL	N-CA-C	5.45	125.72	111.00
1	A	239	GLU	CA-CB-CG	5.45	125.39	113.40
1	A	190	ASP	CB-CG-OD2	5.23	123.01	118.30
1	A	221	GLU	CG-CD-OE2	5.20	128.69	118.30
1	A	197	SER	O-C-N	5.16	130.96	122.70
1	A	52	ASP	CB-CG-OD2	-5.12	113.69	118.30
1	A	231	PHE	CA-CB-CG	5.09	126.12	113.90
1	A	190	ASP	CB-CG-OD1	-5.06	113.75	118.30
1	A	89	ARG	CD-NE-CZ	-5.04	116.54	123.60
1	A	239	GLU	CG-CD-OE1	5.03	128.36	118.30
1	A	89	ARG	NE-CZ-NH1	5.02	122.81	120.30
1	A	51	TYR	CA-CB-CG	5.02	122.93	113.40
1	A	80	LYS	CA-CB-CG	5.00	124.40	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	58	ARG	Sidechain

## 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	2026	0	1975	83	2
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	89	0	0	18	2
All	All	2117	0	1975	83	2

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 21.

All (83) 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:58:ARG:HD2	1:A:69:GLU:OE1	1.49	1.10
1:A:125:THR:C	1:A:127:LYS:N	2.33	0.81
1:A:252:LYS:HE2	4:A:270:HOH:O	1.81	0.80
1:A:128:TYR:CE1	1:A:137:GLN:HG3	2.18	0.79
1:A:146:ILE:HG12	1:A:212:LEU:HD23	1.66	0.76
1:A:253:ASN:HD22	1:A:254:ARG:N	1.83	0.76
1:A:89:ARG:HG3	1:A:125:THR:HG21	1.67	0.76
1:A:159:LYS:HD2	4:A:324:HOH:O	1.87	0.73
1:A:221:GLU:HA	1:A:224:LEU:HD12	1.68	0.73
1:A:30:PRO:HG3	1:A:106:GLU:HB3	1.70	0.72
1:A:253:ASN:ND2	1:A:254:ARG:N	2.38	0.71
1:A:84:LEU:O	4:A:289:HOH:O	2.10	0.68
1:A:27:ARG:NH2	4:A:310:HOH:O	2.27	0.68
1:A:128:TYR:CZ	1:A:137:GLN:HG3	2.31	0.66
1:A:52:ASP:OD1	4:A:341:HOH:O	2.13	0.65
1:A:63:GLY:HA3	1:A:170:LYS:NZ	2.13	0.63
1:A:253:ASN:HD22	1:A:253:ASN:C	2.02	0.63
1:A:110:ASP:O	1:A:111:LYS:HB2	2.00	0.61
1:A:89:ARG:HG3	1:A:125:THR:CG2	2.31	0.61
1:A:178:ASN:OD1	4:A:348:HOH:O	2.16	0.60
1:A:113:LYS:HD3	4:A:277:HOH:O	2.00	0.60
1:A:45:LYS:O	1:A:82:GLY:HA2	2.02	0.59
1:A:53:GLN:HG3	4:A:279:HOH:O	2.01	0.59

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:93:PHE:HA	1:A:119:HIS:O	2.02	0.59
1:A:63:GLY:HA3	1:A:170:LYS:HZ3	1.67	0.59
1:A:112:LYS:HE3	1:A:114:TYR:CZ	2.39	0.57
1:A:56:SER:O	1:A:177:THR:HA	2.05	0.56
1:A:202:PRO:HB2	1:A:204:LEU:HD21	1.88	0.56
1:A:53:GLN:O	1:A:76:LYS:HB3	2.06	0.55
1:A:252:LYS:O	1:A:253:ASN:CG	2.45	0.55
1:A:243:ASP:HA	1:A:245:TRP:CD1	2.42	0.55
1:A:50:SER:O	1:A:77:ALA:HA	2.07	0.54
1:A:86:GLY:HA3	4:A:346:HOH:O	2.08	0.54
1:A:154:LYS:HE3	1:A:183:GLY:O	2.08	0.53
1:A:113:LYS:NZ	4:A:273:HOH:O	2.41	0.53
1:A:257:LYS:HE2	4:A:315:HOH:O	2.09	0.52
1:A:60:LEU:O	1:A:66:PHE:HA	2.09	0.51
1:A:231:PHE:HD2	1:A:239:GLU:HG2	1.76	0.51
1:A:247:PRO:O	1:A:249:GLN:NE2	2.36	0.51
1:A:253:ASN:ND2	1:A:253:ASN:C	2.65	0.50
1:A:249:GLN:HB3	1:A:250:PRO:CD	2.42	0.50
1:A:59:ILE:HA	1:A:67:ASN:O	2.11	0.50
1:A:168:LYS:HE3	4:A:282:HOH:O	2.11	0.50
1:A:125:THR:O	1:A:127:LYS:CA	2.60	0.49
1:A:180:ASP:OD2	1:A:182:ARG:NH2	2.40	0.48
1:A:13:PRO:HA	1:A:16:TRP:CD2	2.49	0.48
1:A:64:HIS:CD2	4:A:352:HOH:O	2.66	0.48
1:A:89:ARG:O	1:A:122:HIS:HA	2.14	0.47
1:A:36:HIS:HD2	4:A:340:HOH:O	1.97	0.47
1:A:202:PRO:HB2	1:A:204:LEU:CD2	2.44	0.47
1:A:125:THR:C	1:A:127:LYS:CA	2.83	0.47
1:A:25:GLY:HA3	1:A:205:GLU:OE2	2.15	0.46
1:A:194:TYR:HA	4:A:308:HOH:O	2.15	0.45
1:A:88:TYR:HB3	1:A:122:HIS:HB3	1.97	0.45
1:A:190:ASP:HB3	1:A:260:PHE:CD1	2.52	0.45
1:A:250:PRO:HB2	1:A:252:LYS:HG3	1.99	0.45
1:A:145:GLY:O	1:A:212:LEU:HD22	2.17	0.45
1:A:22:ILE:HD11	1:A:205:GLU:CD	2.37	0.45
1:A:5:TRP:CD2	1:A:201:PRO:HG2	2.52	0.44
1:A:135:VAL:HB	1:A:204:LEU:CD1	2.48	0.44
1:A:70:PHE:CZ	1:A:181:PRO:HG3	2.51	0.44
1:A:144:LEU:HD23	1:A:144:LEU:HA	1.82	0.44
1:A:51:TYR:CD1	1:A:77:ALA:HB1	2.53	0.44
1:A:135:VAL:HA	1:A:206:CYS:SG	2.57	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:214:GLU:HA	1:A:215:PRO:HD3	1.72	0.44
1:A:134:ALA:C	1:A:136:GLN:H	2.20	0.44
1:A:154:LYS:HA	1:A:155:PRO:HD2	1.81	0.43
1:A:159:LYS:HG3	4:A:314:HOH:O	2.18	0.43
1:A:161:VAL:CG1	1:A:225:LYS:HD2	2.48	0.43
1:A:47:LEU:HD11	1:A:210:ILE:HG21	2.01	0.43
1:A:185:LEU:HA	1:A:186:PRO:HD3	1.91	0.43
1:A:5:TRP:CG	1:A:201:PRO:HG2	2.54	0.42
1:A:47:LEU:HD23	1:A:82:GLY:HA3	2.00	0.42
1:A:104:GLY:HA3	1:A:115:ALA:O	2.20	0.42
1:A:236:GLU:HG3	4:A:302:HOH:O	2.20	0.42
1:A:170:LYS:HD2	1:A:231:PHE:O	2.20	0.41
1:A:7:TYR:OH	1:A:200:THR:HG22	2.20	0.41
1:A:161:VAL:HG13	1:A:225:LYS:CD	2.51	0.41
1:A:163:VAL:O	1:A:166:SER:HB2	2.21	0.41
1:A:84:LEU:HD23	1:A:208:THR:HG21	2.02	0.41
1:A:64:HIS:HB3	4:A:333:HOH:O	2.19	0.41
1:A:180:ASP:HA	1:A:181:PRO:HD3	1.86	0.40
1:A:250:PRO:HB2	1:A:252:LYS:HE3	2.03	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:240:LEU:CG	4:A:347:HOH:O[2_545]	1.98	0.22
1:A:162:ASP:OD2	4:A:344:HOH:O[2_445]	2.15	0.05

## 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	251/260 (96%)	229 (91%)	19 (8%)	3 (1%)	<b>13</b> <b>19</b>

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	25	GLY
1	A	135	VAL
1	A	254	ARG

### 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	219/224 (98%)	212 (97%)	7 (3%)	39 59

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	GLU
1	A	58	ARG
1	A	67	ASN
1	A	79	LEU
1	A	93	PHE
1	A	144	LEU
1	A	253	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	36	HIS
1	A	53	GLN
1	A	67	ASN
1	A	136	GLN
1	A	178	ASN
1	A	253	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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

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	2.33

## 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.