Full wwPDB X-ray Structure Validation Report (i)

Jan 27, 2024 – 12:58 PM EST

PDB ID	:	1CAY
Title	:	WILD-TYPE AND E106Q MUTANT CARBONIC ANHYDRASE COMPLEXED WITH ACETATE
Authors	:	Hakansson, K.; Briand, C.; Zaitsev, V.; Xue, Y.; Liljas, A.
Deposited on	:	1993-02-26
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

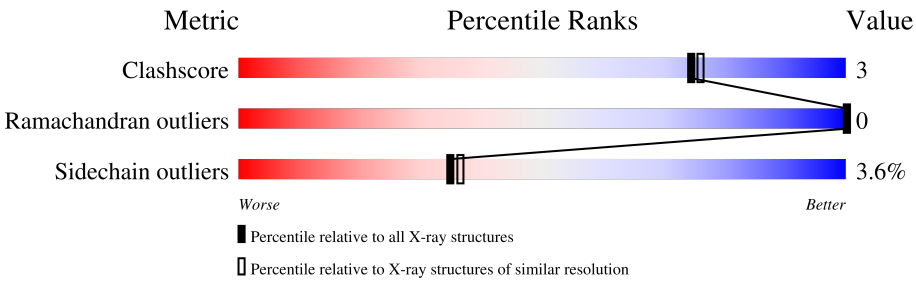
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	259	78% 20% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACY	A	500	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2301 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.

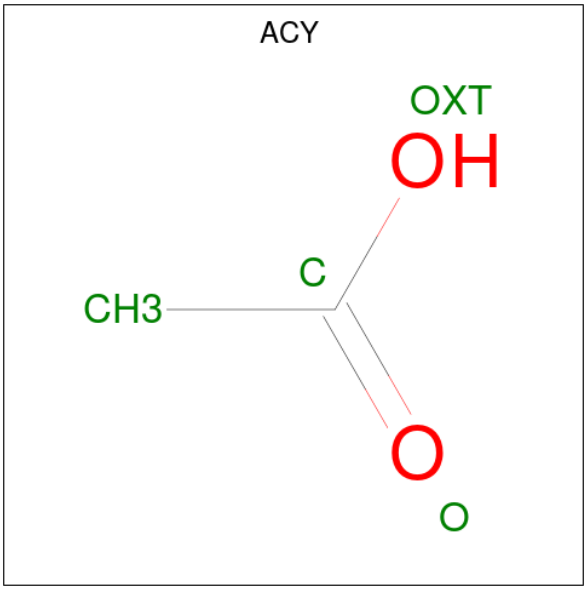
- Molecule 1 is a protein called CARBONIC ANHYDRASE II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	2079	1333	360	384	2	0	4	0

- 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 ACETIC ACID (three-letter code: ACY) (formula: $C_2H_4O_2$).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	217	217	217	0	0

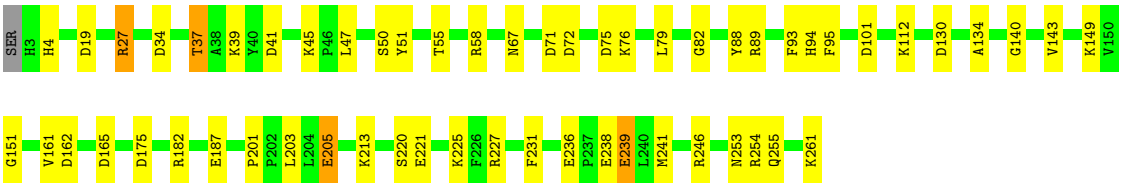
3 Residue-property plots (i)

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

Chain A:



4 Data and refinement statistics (i)

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, α, β, γ	42.70Å 41.70Å 73.00Å 90.00° 104.60° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.10	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROFFT	Depositor
R, R_{free}	0.141 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2301	wwPDB-VP
Average B, all atoms ($Å^2$)	15.0	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: ZN, ACY

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	1.13	1/2161 (0.0%)	1.81	49/2931 (1.7%)

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	2	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	236	GLU	CD-OE2	5.95	1.32	1.25

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	182	ARG	NE-CZ-NH1	17.61	129.10	120.30
1	A	89	ARG	NE-CZ-NH2	-14.55	113.03	120.30
1	A	175	ASP	CB-CG-OD1	12.77	129.79	118.30
1	A	27	ARG	NE-CZ-NH1	12.16	126.38	120.30
1	A	89	ARG	NE-CZ-NH1	11.25	125.92	120.30
1	A	182	ARG	NE-CZ-NH2	-9.29	115.66	120.30
1	A	75	ASP	CB-CG-OD1	9.07	126.47	118.30
1	A	162	ASP	CB-CG-OD2	-8.94	110.26	118.30
1	A	221	GLU	CG-CD-OE1	8.80	135.89	118.30
1	A	72	ASP	CB-CG-OD2	8.71	126.14	118.30
1	A	101	ASP	CB-CG-OD2	-7.38	111.66	118.30
1	A	71	ASP	CB-CG-OD1	7.36	124.93	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	19	ASP	CB-CG-OD2	-7.01	111.99	118.30
1	A	58	ARG	N-CA-CB	-6.91	98.17	110.60
1	A	162	ASP	CB-CG-OD1	6.43	124.08	118.30
1	A	175	ASP	CB-CG-OD2	-6.39	112.55	118.30
1	A	75	ASP	CB-CG-OD2	-6.25	112.67	118.30
1	A	101	ASP	CB-CG-OD1	6.17	123.85	118.30
1	A	130	ASP	CB-CG-OD2	6.14	123.83	118.30
1	A	205	GLU	OE1-CD-OE2	-6.06	116.03	123.30
1	A	51	TYR	CB-CG-CD1	6.04	124.63	121.00
1	A	187	GLU	OE1-CD-OE2	5.89	130.37	123.30
1	A	34	ASP	CB-CG-OD1	5.88	123.59	118.30
1	A	227	ARG	NE-CZ-NH2	5.87	123.24	120.30
1	A	238	GLU	CG-CD-OE2	5.85	130.00	118.30
1	A	37	THR	CA-CB-CG2	5.81	120.54	112.40
1	A	236	GLU	CG-CD-OE1	5.79	129.89	118.30
1	A	165	ASP	CB-CG-OD2	-5.78	113.09	118.30
1	A	72	ASP	OD1-CG-OD2	-5.73	112.41	123.30
1	A	221	GLU	CG-CD-OE2	-5.68	106.93	118.30
1	A	254	ARG	NE-CZ-NH2	-5.66	117.47	120.30
1	A	27	ARG	CA-CB-CG	5.63	125.79	113.40
1	A	246	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	A	95	PHE	CB-CG-CD1	-5.54	116.92	120.80
1	A	227	ARG	CD-NE-CZ	5.49	131.28	123.60
1	A	130	ASP	OD1-CG-OD2	-5.42	112.99	123.30
1	A	220	SER	CA-CB-OG	-5.42	96.58	111.20
1	A	93	PHE	N-CA-CB	5.38	120.29	110.60
1	A	130	ASP	CB-CG-OD1	5.26	123.03	118.30
1	A	41	ASP	CB-CG-OD2	5.24	123.01	118.30
1	A	19	ASP	CB-CG-OD1	5.17	122.95	118.30
1	A	239	GLU	CG-CD-OE2	5.16	128.61	118.30
1	A	236	GLU	CA-CB-CG	5.12	124.66	113.40
1	A	221	GLU	OE1-CD-OE2	-5.11	117.16	123.30
1	A	88	TYR	CB-CG-CD1	5.08	124.05	121.00
1	A	4[A]	HIS	CA-CB-CG	5.04	122.17	113.60
1	A	4[B]	HIS	CA-CB-CG	5.04	122.17	113.60
1	A	261	LYS	N-CA-CB	5.02	119.64	110.60
1	A	151	GLY	N-CA-C	-5.01	100.58	113.10

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	4[A]	HIS	CA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
1	A	4[B]	HIS	CA

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	2079	0	2021	13	0
2	A	1	0	0	0	0
3	A	4	0	3	2	0
4	A	217	0	0	2	0
All	All	2301	0	2024	13	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 3.

All (13) 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:201:PRO:HA	1:A:203:LEU:HG	1.88	0.55
1:A:143:VAL:HG21	3:A:500:ACY:CH3	2.38	0.54
1:A:47:LEU:HD22	1:A:79:LEU:HD11	1.90	0.52
1:A:112:LYS:NZ	4:A:423:HOH:O	2.45	0.50
1:A:161:VAL:CG1	1:A:225:LYS:HD2	2.42	0.50
1:A:27:ARG:HG3	1:A:205:GLU:HB3	1.93	0.48
1:A:55:THR:OG1	1:A:76:LYS:HE2	2.16	0.46
1:A:45:LYS:O	1:A:82:GLY:HA2	2.17	0.44
1:A:27:ARG:HD3	4:A:294:HOH:O	2.17	0.43
1:A:67:ASN:HD22	1:A:94:HIS:HB3	1.83	0.43
1:A:143:VAL:HG21	3:A:500:ACY:H1	2.01	0.41
1:A:134:ALA:O	1:A:140:GLY:HA3	2.20	0.40
1:A:231:PHE:CE1	1:A:241:MET:HG3	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	260/259 (100%)	250 (96%)	10 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	227/224 (101%)	219 (96%)	8 (4%)	36 38

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	THR
1	A	39	LYS
1	A	50	SER
1	A	149	LYS
1	A	213	LYS
1	A	239	GLU
1	A	253	ASN
1	A	255	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
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 monosaccharides in this entry.

5.6 Ligand geometry (i)

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	ACY	A	500	2	3,3,3	1.40	0	3,3,3	2.58	2 (66%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	500	ACY	O-C-CH3	3.65	136.55	122.33
3	A	500	ACY	OXT-C-CH3	-2.28	105.78	115.18

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
3	A	500	ACY	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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.