Full wwPDB X-ray Structure Validation Report (i)

Jan 30, 2024 – 07:02 PM EST

PDB ID	:	1HVA
Title	:	ENGINEERING THE ZINC BINDING SITE OF HUMAN CARBONIC ANHYDRASE II: STRUCTURE OF THE HIS-94-> CYS APOENZYME IN A NEW CRYSTALLINE FORM
Authors	:	Alexander, R.S.; Christianson, D.W.
Deposited on	:	1992-10-27
Resolution	:	2.30 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:

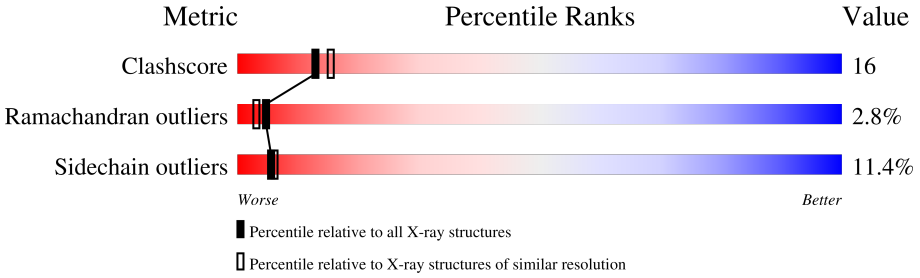
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.30 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	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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	260	53% 37% 5% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2062 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	2024	1300	345	376	3	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	94	CYS	HIS	conflict	UNP P00918

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	38	38	38	0	0

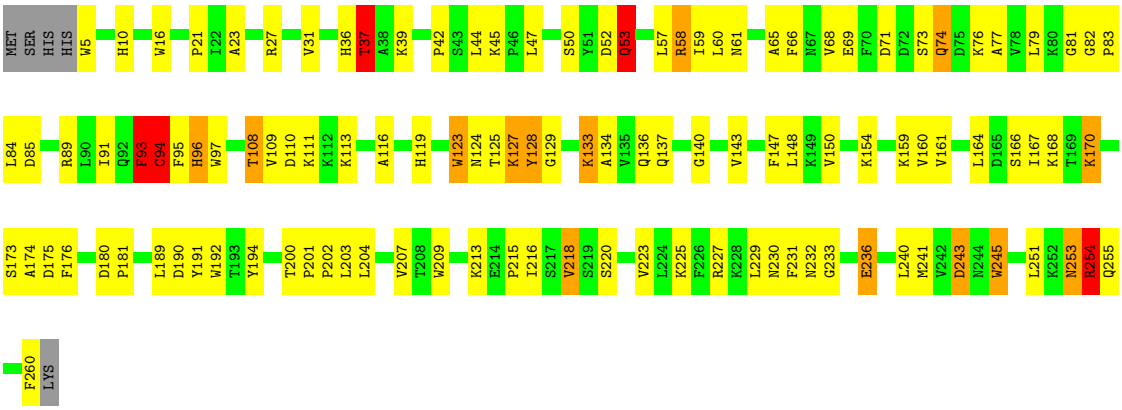
3 Residue-property plots (i)

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: 53% 37% 5%



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	42.40Å 72.20Å 75.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.50 – 2.30	Depositor
% Data completeness (in resolution range)	(Not available) (6.50-2.30)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.155 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2062	wwPDB-VP
Average B, all atoms ($Å^2$)	13.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

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	1.09	2/2083 (0.1%)	1.91	53/2828 (1.9%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	220	SER	CA-CB	-5.99	1.44	1.52
1	A	94	CYS	CA-CB	5.43	1.65	1.53

All (53) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	94	CYS	CA-CB-SG	14.04	139.27	114.00
1	A	209	TRP	CD1-CG-CD2	11.06	115.15	106.30
1	A	27	ARG	NE-CZ-NH2	9.88	125.24	120.30
1	A	5	TRP	CD1-CG-CD2	9.32	113.76	106.30
1	A	68	VAL	CG1-CB-CG2	-9.16	96.24	110.90
1	A	123	TRP	CD1-CG-CD2	8.68	113.25	106.30
1	A	227	ARG	NE-CZ-NH2	-8.66	115.97	120.30
1	A	245	TRP	CD1-CG-CD2	8.65	113.22	106.30
1	A	97	TRP	CG-CD2-CE3	8.44	141.50	133.90
1	A	254	ARG	CA-CB-CG	8.36	131.79	113.40
1	A	209	TRP	CE2-CD2-CG	-8.14	100.79	107.30
1	A	123	TRP	CE2-CD2-CG	-8.09	100.83	107.30
1	A	245	TRP	CE2-CD2-CG	-8.06	100.85	107.30
1	A	192	TRP	CD1-CG-CD2	7.88	112.60	106.30
1	A	16	TRP	CE2-CD2-CG	-7.82	101.05	107.30
1	A	5	TRP	CE2-CD2-CG	-7.73	101.11	107.30
1	A	207	VAL	CG1-CB-CG2	-7.33	99.17	110.90
1	A	16	TRP	CD1-CG-CD2	7.17	112.04	106.30
1	A	243	ASP	CB-CG-OD1	7.16	124.75	118.30
1	A	97	TRP	CE2-CD2-CG	-6.98	101.71	107.30
1	A	192	TRP	CE2-CD2-CG	-6.86	101.81	107.30
1	A	128	TYR	CB-CG-CD2	-6.83	116.90	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	27	ARG	NH1-CZ-NH2	-6.82	111.89	119.40
1	A	93	PHE	CB-CG-CD2	-6.80	116.04	120.80
1	A	91	ILE	O-C-N	-6.61	112.12	122.70
1	A	69	GLU	CA-CB-CG	6.54	127.80	113.40
1	A	97	TRP	CD1-CG-CD2	6.40	111.42	106.30
1	A	209	TRP	CG-CD2-CE3	6.34	139.61	133.90
1	A	209	TRP	CG-CD1-NE1	-6.31	103.79	110.10
1	A	229	LEU	CA-CB-CG	6.05	129.22	115.30
1	A	245	TRP	CG-CD2-CE3	6.04	139.34	133.90
1	A	192	TRP	CG-CD1-NE1	-6.00	104.10	110.10
1	A	227	ARG	NE-CZ-NH1	5.99	123.30	120.30
1	A	143	VAL	CG1-CB-CG2	-5.97	101.35	110.90
1	A	123	TRP	CG-CD2-CE3	5.94	139.25	133.90
1	A	254	ARG	NE-CZ-NH1	5.92	123.26	120.30
1	A	16	TRP	CG-CD2-CE3	5.91	139.22	133.90
1	A	123	TRP	CB-CG-CD1	-5.90	119.33	127.00
1	A	253	ASN	O-C-N	5.87	132.09	122.70
1	A	97	TRP	CB-CG-CD1	-5.80	119.46	127.00
1	A	5	TRP	CG-CD1-NE1	-5.72	104.38	110.10
1	A	109	VAL	CG1-CB-CG2	-5.67	101.83	110.90
1	A	73	SER	CA-C-N	5.62	129.57	117.20
1	A	53	GLN	CA-CB-CG	5.57	125.65	113.40
1	A	37	THR	O-C-N	-5.38	114.10	122.70
1	A	218	VAL	CA-CB-CG1	-5.31	102.93	110.90
1	A	89	ARG	NE-CZ-NH2	-5.27	117.66	120.30
1	A	96	HIS	CA-CB-CG	-5.24	104.70	113.60
1	A	36	HIS	CB-CA-C	5.18	120.77	110.40
1	A	77	ALA	N-CA-C	-5.18	97.01	111.00
1	A	245	TRP	CB-CG-CD1	-5.18	120.27	127.00
1	A	36	HIS	N-CA-C	-5.12	97.18	111.00
1	A	123	TRP	CG-CD1-NE1	-5.11	105.00	110.10

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	2024	0	1980	66	0
2	A	38	0	0	5	0
All	All	2062	0	1980	66	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 16.

All (66) 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:57:LEU:HD21	1:A:71:ASP:HB2	1.63	0.81
1:A:161:VAL:HG13	1:A:225:LYS:HG3	1.69	0.75
1:A:94:CYS:HG	1:A:96:HIS:CE1	2.12	0.63
1:A:94:CYS:HG	1:A:96:HIS:HE2	0.67	0.63
1:A:124:ASN:HD21	1:A:127:LYS:HE2	1.64	0.63
1:A:113:LYS:HD3	2:A:264:HOH:O	2.01	0.61
1:A:190:ASP:HB2	1:A:213:LYS:NZ	2.17	0.59
1:A:230:ASN:HB3	1:A:232:ASN:OD1	2.03	0.58
1:A:124:ASN:OD1	1:A:127:LYS:HG2	2.05	0.57
1:A:53:GLN:HG3	1:A:76:LYS:HE2	1.86	0.57
1:A:94:CYS:SG	1:A:96:HIS:NE2	2.43	0.56
1:A:170:LYS:HB2	1:A:231:PHE:O	2.05	0.55
1:A:150:VAL:HA	1:A:218:VAL:O	2.06	0.55
1:A:110:ASP:HA	2:A:280:HOH:O	2.05	0.55
1:A:61:ASN:HB2	1:A:167:ILE:O	2.07	0.54
1:A:147:PHE:HB2	1:A:215:PRO:HB3	1.90	0.54
1:A:154:LYS:NZ	1:A:216:ILE:HG22	2.22	0.54
1:A:93:PHE:HA	1:A:119:HIS:O	2.08	0.53
1:A:251:LEU:O	1:A:254:ARG:HB3	2.10	0.52
1:A:21:PRO:HD2	2:A:293:HOH:O	2.08	0.51
1:A:94:CYS:SG	1:A:119:HIS:ND1	2.66	0.50
1:A:231:PHE:CD1	1:A:241:MET:HG3	2.47	0.50
1:A:47:LEU:HB2	1:A:189:LEU:HD23	1.93	0.50
1:A:10:HIS:HB2	2:A:292:HOH:O	2.12	0.50
1:A:123:TRP:CZ3	1:A:125:THR:HA	2.46	0.49
1:A:124:ASN:ND2	1:A:127:LYS:HE2	2.28	0.49
1:A:128:TYR:CZ	1:A:137:GLN:HG2	2.47	0.49
1:A:59:ILE:CG2	1:A:174:ALA:HB3	2.43	0.48
1:A:45:LYS:HD2	1:A:81:GLY:HA2	1.96	0.47
1:A:108:THR:HG22	1:A:113:LYS:HG2	1.96	0.47
1:A:45:LYS:NZ	1:A:84:LEU:O	2.48	0.47
1:A:60:LEU:O	1:A:66:PHE:HA	2.14	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:23:ALA:HB2	1:A:203:LEU:HD13	1.97	0.47
1:A:74:GLN:O	1:A:76:LYS:HG2	2.15	0.46
1:A:136:GLN:NE2	1:A:137:GLN:OE1	2.49	0.46
1:A:213:LYS:HD3	1:A:260:PHE:CE2	2.50	0.46
1:A:251:LEU:HG	1:A:254:ARG:HB3	1.97	0.46
1:A:134:ALA:O	1:A:140:GLY:HA3	2.15	0.46
1:A:65:ALA:HB2	2:A:291:HOH:O	2.16	0.45
1:A:154:LYS:HZ3	1:A:216:ILE:HG22	1.80	0.45
1:A:154:LYS:HD2	1:A:154:LYS:HA	1.82	0.44
1:A:159:LYS:HG2	1:A:176:PHE:CE1	2.52	0.44
1:A:170:LYS:HB3	1:A:233:GLY:HA2	2.00	0.44
1:A:45:LYS:HE3	1:A:82:GLY:O	2.17	0.44
1:A:116:ALA:HB3	1:A:148:LEU:HD22	2.00	0.43
1:A:93:PHE:CE1	1:A:95:PHE:HE2	2.37	0.43
1:A:133:LYS:O	1:A:136:GLN:OE1	2.36	0.43
1:A:58:ARG:NH1	1:A:60:LEU:HD13	2.34	0.43
1:A:243:ASP:HA	1:A:245:TRP:CD1	2.54	0.43
1:A:251:LEU:HD21	1:A:254:ARG:O	2.19	0.43
1:A:60:LEU:HD23	1:A:60:LEU:N	2.34	0.43
1:A:213:LYS:HD3	1:A:260:PHE:CZ	2.53	0.43
1:A:59:ILE:HG23	1:A:174:ALA:HB3	2.01	0.43
1:A:160:VAL:O	1:A:164:LEU:HG	2.19	0.42
1:A:31:VAL:HG22	1:A:194:TYR:CZ	2.53	0.42
1:A:168:LYS:NZ	1:A:230:ASN:HD21	2.18	0.42
1:A:59:ILE:HG12	1:A:167:ILE:HD13	2.02	0.41
1:A:82:GLY:HA2	1:A:191:TYR:OH	2.20	0.41
1:A:45:LYS:HE3	1:A:45:LYS:HB2	1.97	0.41
1:A:124:ASN:CG	1:A:127:LYS:HG2	2.40	0.41
1:A:180:ASP:HA	1:A:181:PRO:HD2	1.92	0.41
1:A:202:PRO:HB2	1:A:204:LEU:HG	2.03	0.41
1:A:161:VAL:HG13	1:A:225:LYS:CG	2.47	0.40
1:A:233:GLY:O	1:A:236:GLU:HG2	2.21	0.40
1:A:251:LEU:O	1:A:251:LEU:HG	2.21	0.40
1:A:232:ASN:HB2	1:A:236:GLU:HG3	2.03	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	253/260 (97%)	229 (90%)	17 (7%)	7 (3%)	5 3

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	THR
1	A	254	ARG
1	A	170	LYS
1	A	111	LYS
1	A	129	GLY
1	A	53	GLN
1	A	42	PRO

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	220/225 (98%)	195 (89%)	25 (11%)	5 6

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	THR
1	A	39	LYS
1	A	44	LEU
1	A	50	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	52	ASP
1	A	58	ARG
1	A	74	GLN
1	A	79	LEU
1	A	83	PRO
1	A	85	ASP
1	A	93	PHE
1	A	94	CYS
1	A	108	THR
1	A	127	LYS
1	A	133	LYS
1	A	166	SER
1	A	173	SER
1	A	175	ASP
1	A	200	THR
1	A	201	PRO
1	A	223	VAL
1	A	236	GLU
1	A	240	LEU
1	A	253	ASN
1	A	255	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	HIS
1	A	178	ASN
1	A	230	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)

There are no ligands in this entry.

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.