Full wwPDB X-ray Structure Validation Report (i)

Mar 3, 2024 – 07:06 AM EST

PDB ID	:	6CA2
Title	:	ENGINEERING THE HYDROPHOBIC POCKET OF CARBONIC ANHYDRASE II
Authors	:	Alexander, R.S.; Christianson, D.W.
Deposited on	:	1991-07-09
Resolution	:	2.50 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:

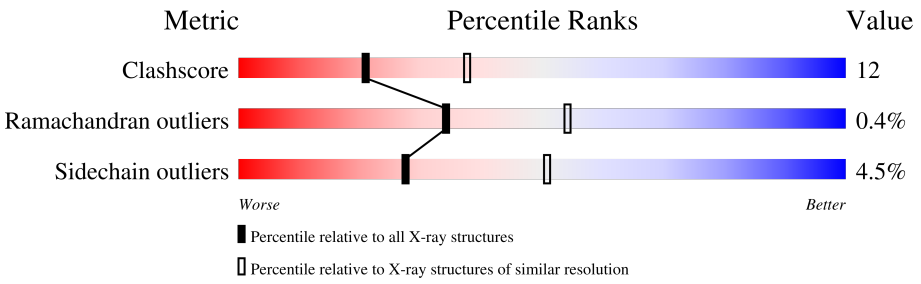
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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	67% 25% 5% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2118 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	2033	1307	347	377	2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	PHE	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	83	83	83	0	0

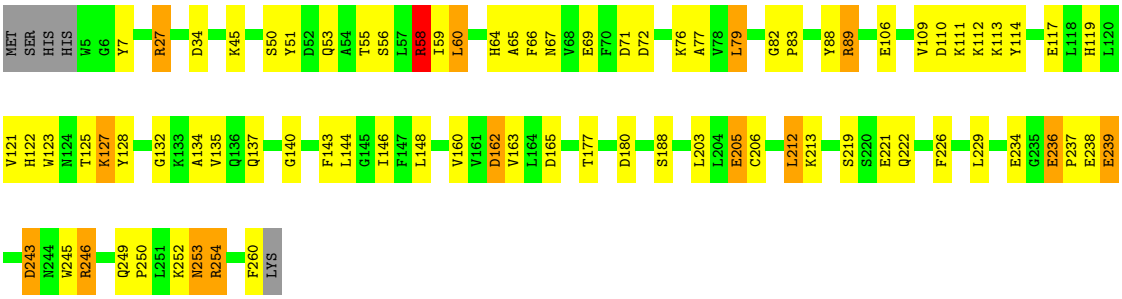
3 Residue-property plots (i)

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: 67% 25% 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 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) – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.180 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2118	wwPDB-VP
Average B, all atoms ($Å^2$)	13.0	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, 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	1.45	12/2093 (0.6%)	1.59	16/2839 (0.6%)

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	238	GLU	CD-OE1	10.60	1.37	1.25
1	A	106	GLU	CD-OE2	9.27	1.35	1.25
1	A	238	GLU	CD-OE2	-7.67	1.17	1.25
1	A	234	GLU	CD-OE2	-6.83	1.18	1.25
1	A	221	GLU	CD-OE2	-6.59	1.18	1.25
1	A	234	GLU	CD-OE1	6.42	1.32	1.25
1	A	205	GLU	CD-OE1	-6.05	1.19	1.25
1	A	236	GLU	CD-OE2	6.03	1.32	1.25
1	A	239	GLU	CD-OE2	6.00	1.32	1.25
1	A	221	GLU	CD-OE1	5.75	1.31	1.25
1	A	188	SER	CB-OG	-5.64	1.34	1.42
1	A	226	PHE	CG-CD2	5.28	1.46	1.38

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	162	ASP	CB-CG-OD1	8.68	126.11	118.30
1	A	71	ASP	CB-CG-OD1	8.11	125.60	118.30
1	A	34	ASP	CB-CG-OD2	6.91	124.52	118.30
1	A	34	ASP	CB-CG-OD1	-6.88	112.11	118.30
1	A	79	LEU	CA-CB-CG	6.59	130.46	115.30
1	A	246	ARG	NE-CZ-NH1	6.56	123.58	120.30
1	A	27	ARG	NE-CZ-NH1	6.54	123.57	120.30
1	A	254	ARG	NE-CZ-NH1	6.32	123.46	120.30
1	A	243	ASP	CB-CG-OD2	6.26	123.94	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	162	ASP	CB-CG-OD2	-6.04	112.86	118.30
1	A	7	TYR	CB-CG-CD2	-5.77	117.54	121.00
1	A	89	ARG	CD-NE-CZ	5.48	131.27	123.60
1	A	58	ARG	NE-CZ-NH1	5.43	123.02	120.30
1	A	180	ASP	CB-CG-OD1	5.37	123.13	118.30
1	A	165	ASP	CB-CG-OD2	-5.17	113.65	118.30
1	A	71	ASP	CB-CG-OD2	-5.08	113.73	118.30

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	2033	0	1981	48	2
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	83	0	0	3	1
All	All	2118	0	1981	48	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 12.

All (48) 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:53:GLN:O	1:A:76:LYS:HD3	1.61	1.00
1:A:125:THR:C	1:A:127:LYS:N	2.17	0.98
1:A:58:ARG:HD2	1:A:69:GLU:OE1	1.68	0.93
1:A:253:ASN:C	1:A:253:ASN:HD22	1.78	0.85
1:A:113:LYS:NZ	4:A:271:HOH:O	2.21	0.73
1:A:53:GLN:HG3	1:A:76:LYS:CE	2.23	0.69
1:A:53:GLN:HG3	1:A:76:LYS:NZ	2.08	0.69
1:A:89:ARG:O	1:A:122:HIS:HA	1.96	0.65
1:A:128:TYR:CE1	1:A:137:GLN:HG3	2.32	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:253:ASN:C	1:A:253:ASN:ND2	2.51	0.61
1:A:89:ARG:HG3	1:A:125:THR:CG2	2.30	0.61
1:A:253:ASN:HD22	1:A:254:ARG:N	1.98	0.60
1:A:45:LYS:O	1:A:82:GLY:HA2	2.02	0.60
1:A:55:THR:HG23	4:A:336:HOH:O	2.01	0.60
1:A:128:TYR:CZ	1:A:137:GLN:HG3	2.37	0.60
1:A:160:VAL:O	1:A:163:VAL:HG12	2.03	0.59
1:A:89:ARG:HG3	1:A:125:THR:HG22	1.85	0.56
1:A:72:ASP:OD2	1:A:123:TRP:NE1	2.37	0.55
1:A:132:GLY:O	1:A:135:VAL:HG22	2.08	0.54
1:A:135:VAL:O	1:A:206:CYS:SG	2.67	0.53
1:A:236:GLU:HB3	1:A:237:PRO:HD2	1.90	0.53
1:A:51:TYR:CD1	1:A:77:ALA:HB1	2.44	0.53
1:A:219:SER:OG	1:A:222:GLN:HG3	2.09	0.53
1:A:203:LEU:HD22	1:A:246:ARG:NH1	2.25	0.52
1:A:58:ARG:NH1	1:A:60:LEU:HD13	2.25	0.52
1:A:56:SER:O	1:A:177:THR:HA	2.11	0.50
1:A:121:VAL:HG22	1:A:143:PHE:CE2	2.47	0.50
1:A:146:ILE:HG12	1:A:212:LEU:HD23	1.94	0.49
1:A:117:GLU:OE2	1:A:119:HIS:NE2	2.34	0.48
1:A:88:TYR:HB3	1:A:122:HIS:HB3	1.95	0.48
1:A:60:LEU:O	1:A:66:PHE:HA	2.14	0.48
1:A:64:HIS:ND1	4:A:334:HOH:O	2.33	0.48
1:A:236:GLU:HB3	1:A:237:PRO:CD	2.44	0.48
1:A:134:ALA:O	1:A:140:GLY:HA3	2.16	0.45
1:A:114:TYR:HB3	1:A:148:LEU:O	2.17	0.45
1:A:213:LYS:HD3	1:A:260:PHE:CZ	2.51	0.45
1:A:59:ILE:HA	1:A:67:ASN:O	2.17	0.44
1:A:146:ILE:HG12	1:A:212:LEU:CD2	2.48	0.44
1:A:112:LYS:HE3	1:A:114:TYR:CZ	2.53	0.43
1:A:249:GLN:HB3	1:A:250:PRO:CD	2.48	0.43
1:A:110:ASP:O	1:A:111:LYS:HB2	2.19	0.42
1:A:53:GLN:CG	1:A:76:LYS:CE	2.95	0.42
1:A:53:GLN:HG3	1:A:76:LYS:HZ2	1.81	0.42
1:A:252:LYS:O	1:A:253:ASN:CG	2.58	0.42
1:A:243:ASP:HA	1:A:245:TRP:CD1	2.54	0.41
1:A:109:VAL:O	1:A:112:LYS:HB3	2.21	0.41
1:A:53:GLN:CG	1:A:76:LYS:HE2	2.51	0.41
1:A:27:ARG:HG3	1:A:205:GLU:HB3	2.04	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:58:ARG:NH2	1:A:239:GLU:OE1[2_555]	2.13	0.07
1:A:162:ASP:OD2	4:A:316:HOH:O[2_445]	2.14	0.06

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%)	239 (95%)	11 (4%)	1 (0%)	34 54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	ALA

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%)	210 (96%)	10 (4%)	27 51

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

Mol	Chain	Res	Type
1	A	50	SER
1	A	58	ARG
1	A	60	LEU
1	A	79	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	83	PRO
1	A	127	LYS
1	A	144	LEU
1	A	212	LEU
1	A	229	LEU
1	A	253	ASN

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	67	ASN
1	A	136	GLN
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	2.17

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.