



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 07:04 AM GMT

PDB ID : 1G5C
Title : CRYSTAL STRUCTURE OF THE 'CAB' TYPE BETA CLASS CARBONIC ANHYDRASE FROM METHANOBACTERIUM THERMOAUTOTROPHICUM
Authors : Strop, P.; Smith, K.S.; Iverson, T.M.; Ferry, J.G.; Rees, D.C.
Deposited on : 2000-10-31
Resolution : 2.10 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

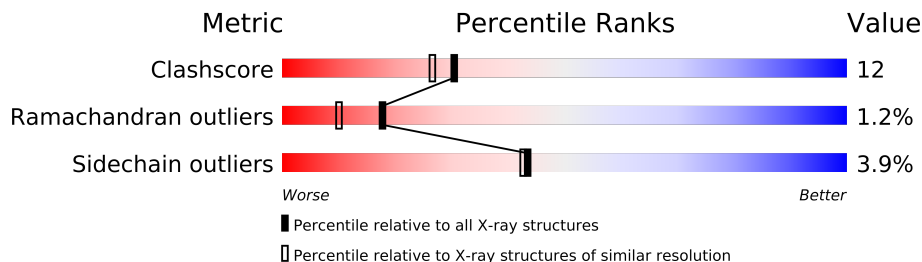
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 21963
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

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	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	170	
1	B	170	
1	C	170	
1	D	170	
1	E	170	
1	F	170	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8010 atoms, of which 0 are hydrogen and 0 are deuterium.

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 BETA-CARBONIC ANHYDRASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	61	0	0
			1311	820	228	257	6			
1	B	158	Total	C	N	O	S	31	0	0
			1210	755	208	241	6			
1	C	169	Total	C	N	O	S	78	0	0
			1311	820	228	257	6			
1	D	158	Total	C	N	O	S	59	0	0
			1210	755	208	241	6			
1	E	169	Total	C	N	O	S	151	0	0
			1311	820	228	257	6			
1	F	155	Total	C	N	O	S	42	0	0
			1187	743	204	234	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	ASP	VAL	CONFLICT	UNP Q50565
B	79	ASP	VAL	CONFLICT	UNP Q50565
C	79	ASP	VAL	CONFLICT	UNP Q50565
D	79	ASP	VAL	CONFLICT	UNP Q50565
E	79	ASP	VAL	CONFLICT	UNP Q50565
F	79	ASP	VAL	CONFLICT	UNP Q50565

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Zn	0	0
			1	1		
2	E	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		

Continued on next page...

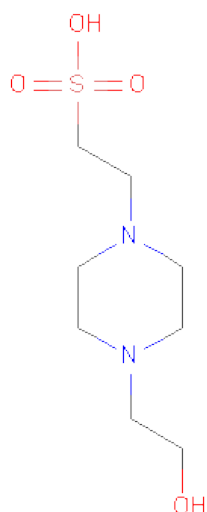
Continued from previous page...

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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		
3	A	2	Total	Ca	0	0
			2	2		
3	C	1	Total	Ca	0	0
			1	1		
3	F	1	Total	Ca	0	0
			1	1		
3	E	1	Total	Ca	0	0
			1	1		

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINEETHANESULFONIC ACID (three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	C	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
4	E	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	83	Total	O	0	0
			83	83		
5	B	73	Total	O	0	0
			73	73		
5	C	58	Total	O	0	0
			58	58		
5	D	49	Total	O	0	0
			49	49		
5	E	84	Total	O	0	0
			84	84		
5	F	66	Total	O	0	0
			66	66		

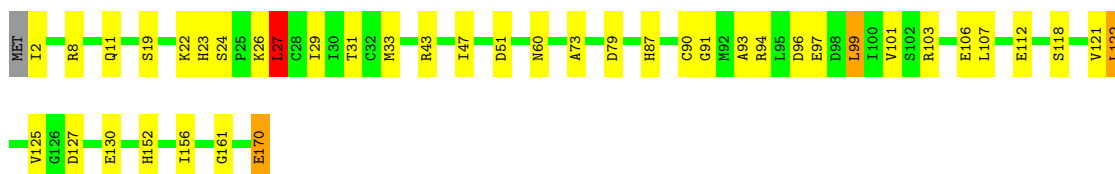
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: BETA-CARBONIC ANHYDRASE

Chain A: 



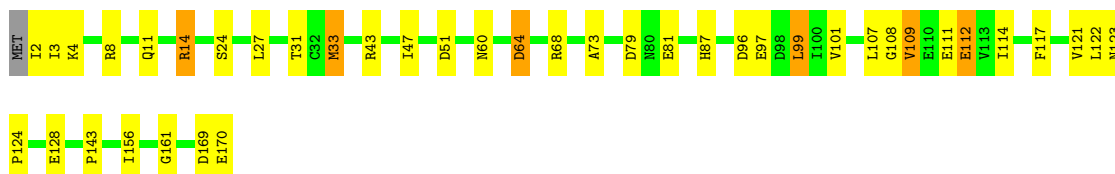
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain B: 



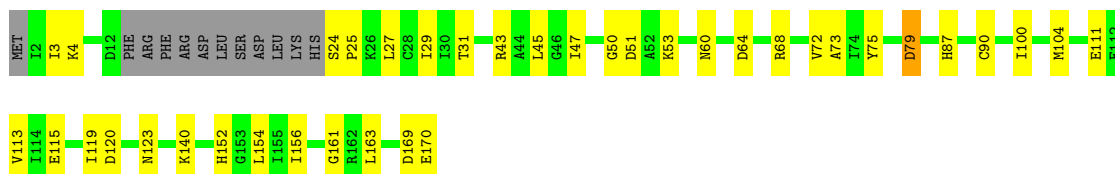
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain C: 



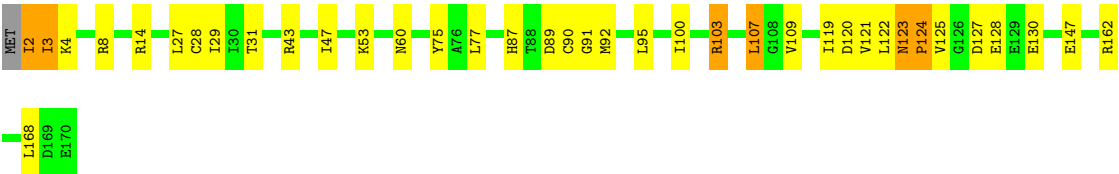
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain D: 



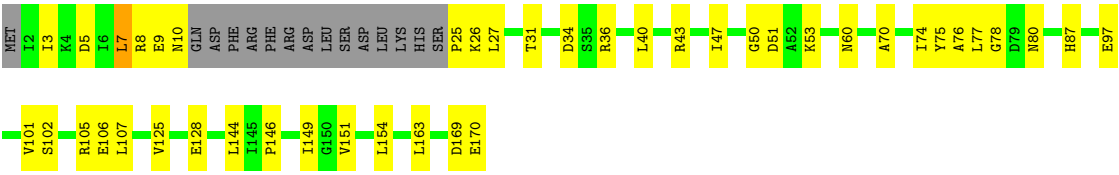
• Molecule 1: BETA-CARBONIC ANHYDRASE

Chain E: 



● Molecule 1: BETA-CARBONIC ANHYDRASE

Chain F: 



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	54.67Å 113.21Å 156.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	18.00 – 2.10	Depositor
% Data completeness (in resolution range)	98.8 (18.00-2.10)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.211 , 0.248	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	8010	wwPDB-VP
Average B, all atoms (Å ²)	30.0	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, EPE, CA

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	3.39	1/1323 (0.1%)	0.67	1/1783 (0.1%)
1	B	0.39	0/1218	0.64	0/1642
1	C	0.36	0/1323	0.66	1/1783 (0.1%)
1	D	0.34	0/1218	0.62	0/1642
1	E	0.45	0/1323	0.73	0/1783
1	F	0.37	0/1195	0.66	0/1610
All	All	1.46	1/7600 (0.0%)	0.66	2/10243 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	170	GLU	C-OXT	122.24	3.55	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	33	MET	CG-SD-CE	6.71	110.94	100.20
1	A	27	LEU	CA-CB-CG	5.31	127.52	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1311	0	1341	32	0
1	B	1210	0	1241	39	0
1	C	1311	0	1341	31	0
1	D	1210	0	1241	30	0
1	E	1311	0	1341	27	0
1	F	1187	0	1225	29	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	15	0	17	3	0
4	C	15	0	17	0	0
4	E	15	0	17	4	0
5	A	83	0	0	2	0
5	B	73	0	0	1	0
5	C	58	0	0	4	0
5	D	49	0	0	0	0
5	E	84	0	0	0	0
5	F	66	0	0	1	0
All	All	8010	0	7781	172	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 12.

The worst 5 of 172 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:24:SER:HB2	1:D:25:PRO:HD3	1.44	0.97
1:A:118:SER:HA	1:A:122:LEU:HD23	1.50	0.94
1:C:8:ARG:O	1:C:11:GLN:HG2	1.70	0.90
1:E:2:ILE:HD12	1:E:3:ILE:HD12	1.58	0.85
1:C:73:ALA:HB1	1:C:79:ASP:OD1	1.76	0.85

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	167/170 (98%)	157 (94%)	7 (4%)	3 (2%)	13	5
1	B	154/170 (91%)	152 (99%)	2 (1%)	0	100	100
1	C	167/170 (98%)	154 (92%)	10 (6%)	3 (2%)	13	5
1	D	154/170 (91%)	146 (95%)	8 (5%)	0	100	100
1	E	167/170 (98%)	158 (95%)	4 (2%)	5 (3%)	7	2
1	F	151/170 (89%)	142 (94%)	8 (5%)	1 (1%)	30	23
All	All	960/1020 (94%)	909 (95%)	39 (4%)	12 (1%)	18	10

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	125	VAL
1	C	109	VAL
1	C	123	ASN
1	E	121	VAL
1	A	94	ARG

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	147/148 (99%)	143 (97%)	4 (3%)	57	60
1	B	136/148 (92%)	132 (97%)	4 (3%)	55	57
1	C	147/148 (99%)	141 (96%)	6 (4%)	41	40
1	D	136/148 (92%)	132 (97%)	4 (3%)	55	57
1	E	147/148 (99%)	138 (94%)	9 (6%)	26	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	133/148 (90%)	127 (96%)	6 (4%)	38	35
All	All	846/888 (95%)	813 (96%)	33 (4%)	43	43

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	27	LEU
1	E	2	ILE
1	F	40	LEU
1	D	79	ASP
1	D	120	ASP

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

Mol	Chain	Res	Type
1	B	123	ASN
1	D	123	ASN
1	E	123	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 12 are monoatomic - leaving 3 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	$\# Z > 2$
4	EPE	A	2201	-	15,15,15	1.42	4 (26%)	20,20,20	1.14	2 (10%)
4	EPE	C	2202	-	15,15,15	1.16	1 (6%)	20,20,20	1.32	1 (5%)
4	EPE	E	2203	-	15,15,15	1.65	4 (26%)	20,20,20	2.30	7 (35%)

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
4	EPE	A	2201	-	-	0/9/19/19	0/1/1/1
4	EPE	C	2202	-	-	0/9/19/19	0/1/1/1
4	EPE	E	2203	-	-	0/9/19/19	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	2203	EPE	C6-N1	4.43	1.59	1.47
4	A	2201	EPE	C9-N1	2.83	1.54	1.47
4	E	2203	EPE	C2-N1	2.56	1.54	1.47
4	A	2201	EPE	C2-N1	2.45	1.53	1.47
4	E	2203	EPE	C7-C8	-2.36	1.40	1.51

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	2203	EPE	C5-N4-C3	5.13	121.59	108.86
4	E	2203	EPE	C3-C2-N1	-5.06	100.64	110.61
4	C	2202	EPE	O1S-S-C10	-4.23	103.18	106.81
4	A	2201	EPE	O1S-S-C10	-2.86	104.36	106.81
4	E	2203	EPE	O8-C8-C7	-2.73	98.83	111.24

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.