STRUCTURAL ANALYSIS OF BACTERIAL CARBONIC ANHYDRASES

Project report submitted in partial fulfilment of the requirements

for the Degree of Bachelor of Technology

In

Department of Biotechnology and Bioinformatics

By

Himani Chauhan (161512)

Sayra Sood (161516)

Under the supervision of

Dr. Narendra Kumar

Assistant Professor (Grade II)

To

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,
WAKNAGHAT,

SOLAN – 173234, HIMACHAL PRADESH

TABLE OF CONTENT

Table of Contents

CANDIDATE'S DECLARATION

I hereby declare that the work presented in the report entitled "Structural Analysis of Carbonic Anhydrase" in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Biotechnology and Bioinformatics submitted in the department of Biotechnology and Bioinformatics Jaypee University of Information Technology, Waknaghat is an authentic record of our own work carried out over a period from August 2019 to December 2019 under the supervision of Dr. Narendra Kumar, Assistant Professor, Department of Biotechnology and Bioinformatics.

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

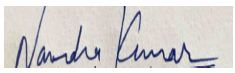
Himani Chauhan

161512

Sayra Sood

161516

This is to certify that the above statement made by the candidates is true to the best of my knowledge.



Dr. Narendra Kumar

Assistant Professor

Biotechnology and Bioinformatics

Dated : 31-08-2020

ACKNOWLEDGEMENT

We have taken efforts in this project. However, it would not have been possible without the kind support and guidance of many individuals and organizations. We would like to extend our sincere thanks to all of them.

We are thankful to **Prof. Dr. Sudhir Kumar**, Head of Department of Biotechnology and Bioinformatics for the support.

We are thankful to **Dr. Narendra Kumar** for his guidance and constant supervision as well as for providing necessary information regarding the project and also for his support in the project.

We would also like to thank **Mrs. Somlata Sharma**, senior lab assistant for her attention and time.

We would like to express our gratitude towards our university members for their kind cooperation and encouragement which help us in this project so far.

Our thanks and appreciations also go to our colleagues in the developing of the project who have willingly helped us out with their abilities.

ABSTRACT

Over the past few years, massive greenhouse gas emissions (such as CO2, CFC) from human activities into the atmosphere have contributed to an increase in global temperature often known as global warming. Consequently, avoiding this emission and designing new technology and new methods to stop such pollution have become absolutely important. Biological systems in nature are capable of fixing atmospheric CO2 but the biological system cannot absorb and use the entire CO2 in urban and industrial areas where the rate of CO2 pollution is very high. Different chemicals and synthetic materials with a CO2-absorbing property are not environmentally friendly, or rather costly.

Carbonic anhydrase (CA) is the fastest known enzyme in its active site containing zinc which converts CO2 to bicarbonate ions. It is one of the important bio catalysts for CO2 conversion. So the bio catalytic properties of microbial CA can be exploited to reduce the amount of CO2. For sequestration of CO2 more than fifty different microbial CAs were explored. CA's main benefits to sequestering CO2 are economic feasibility and low concentration carbonization of CO2. Despite the higher catalysis intensity, CA 's stability poses a major challenge for its industrial application.

The immobilization of the Carbonic Anhydrases onto the bio-inspired surface, biochar, alginate, polyurethane foam and a number of nano-textured materials have solved such problems. The use of carbon dioxide would be plausibly energized by a combination of enzyme and material that jointly absorbs and transforms carbon dioxide into a carbon-rich compound of economic value or reduced carbon derivatives.

INTRODUCTION

Carbon is a vital component of all the living beings, being the basis of all biooruanic mixes. Carbon plays a fairly important in how we see life today. A chain of organic inter conversion starts exactly when the inorganic carbon (Ci) is fixed via autotrophs — the primary producers of organic matter in the biosphere — and proceeds in heterotrophic organisms. Decomposition of the organics closes this chain and discharges carbon so as to go into another cycle of fixation and inter conversion. Disintegrated in an aqueous arrangement, Ci may exist in different forms (carbonic acid, carbon dioxide, bicarbonate and, carbonate particles) in equilibrium [1].

In living cells, acceleration of inter conversion between CO, and HCO3- is guaranteed by the metalloenzyme carbonic anhydrase (CA, carbonate dehydratase, carbonate hydrolase), which catalyses the following reversible reaction :

CO, + H,O <-> H* + HCO3*

Carbonic anhydrase , a metalloenzyme containing zinc, is an important photosynthesis enzyme due to its ability to convert CO2 to HCO3- reversibly [2]. Carbonic Anhydrases are found in all groups of living organisms. Regardless of the diversity of CAs, which are different in their amino acid sequences and in the design of the active center, they catalyze a similar reaction.

In animals, CAs are included into a wide scope of physiological reactions including global catabolic respiration, where they involve in the removal of large CO2 amounts from tissues to the surroundings. It also plays a significant part in the plant leaves stomatal closure, other than it's part in the carboxylation.

Carbonic anhydrase is one of the most pleniful enzyme in the chloroplast. Which accounts for almost twenty percent of total soluble protein. Carbonic Anhydrasess of photosynthetic cells have a great importance at both the ecological as well as the biospheric levels.

CA is controlled by its important function in the composition of the modern atmosphere.

Carbonic anhydrase act as a quite significant enzyme for bone physiology as well as kidney and brain physiology. Carbonic anhydrase catalyses the production of carbonic acid from CO2 and H2O. The Carbonic acid produced in the procedure separates spontaneously resulting in discharge of protons. This is quite fundamental in making the acidic condition that is required for the disintegration of minerals of the bones in the resorption lacunae. CA II insufficiency is an autosomal recessive inborn error of the metabolism, resulting in various diseases like Albers-Schönberg disease, cerebral calcification, renal tubular acidosis.

Since 1932, after the discovery of mammalian CAs, seven genetically unique families have been recognized, namely alpha, beta, gamma, delta, zeta, eta and theta. All of the Carbonic Anhydrase families catalyse the reversible hydration of carbon dioxide & they are arranged by contrasts in structural fold in this manner.

Three evolutionarily distinct classes of carbonic anhydrase

Primary sequences of CA were found from the low activity of two hundred sixty residues CA I and the high activity of two hundred fifty-nine residues CA II that were secluded from the human erythrocytes. Full amino acid sequence of isoenzymes Carbonic Anhydrase I & Carbonic Anhydrase II from various species like equine & bovine have followed. Moreover, in humans, 10 isoenzymes have now been described as clearly resulting from a mutual ancestral enzyme.

For human Carbonic Anhydrase I, human Carbonic Anhydrase II, bovine Carbonic Anhydrase III, human Carbonic Anhydrase IV and a condensed form of murine Carbonic Anhydrase V, crystal structures have been found. The General folds of monomeric human isozymes are very similar to the chiefly antiparallel L-sheet as the leading secondary characteristic. IN 1984, the amino acid sequence of a single spinach chloroplast carbonic anhydrase small peptide fragment was taken & it could not be aligned to any mammalian enzymes amino acid sequence. The cDNA sequences which encoded various plant chloroplast carbonic anhydrases were subsequently determined & shown to be belonging to a genetically different class.

It proved to be quite an astonishing discovery, as the amino acid sequences of 2 periplasmic CA from Chlamydomonas had formerly been proven to be more analogous as that of animal CAs. A 3-dimensional structure for a plant enzyme has yet to be reported, though the Circular Dichroism spectra recommends dominance of K-helix structure in blunt divergence with the human isozymes made of L-sheet structures. Also, in its quaternary form, the plant class appears to be more complex in that the enzymes from dicotyledonous plants are assumed to be octamers in which two tetramers are bound by disulphide bridges, whereas monocotyledonous enzymes are advised to be dimers. Hence, the CAs found in the animal kingdom belong to a single gene class now known as K class carbonic anhydrases, while the class of plant enzymes is classified as L class carbohydrases. Originally purified in 1972, Sicca was not published for another twenty years as the first sequence of a prokaryote CA was.

The development of the cynT gene in Escherichia coli emerged in 1992 as having CA Function. When associating the amino acid sequence with any other CAs, CynT was found to be most similar to members of the plant-type or L-class. A gene displaying differentiation with L-class genes was found to be essential for the photosynthetic fixation of CO2 by Synechococcus PCC7942, briefly afterwards.

Both prokaryotic carbonic anhydrases therefore seemed to fit in with class L. When the amino acid derived in the enzyme sequence is from M. Thermophila was associated with the known CAs amino acid sequences, not finding any major similarities. In addition, there was a trimmer

in the crystal structure with a novel L-helix fold in the middle. These implications suggested that this CA is a discrete 3rd class called Q class. It would have been tempting to conclude that all of the Bacteria domain's carbonic anhydrases belonged to the L class, and that all Archaea domain CAs matched the Q class. But the assumed M. The thermophilic enzyme showed sequence discriminating of the presumed amino acid sequence of the Synechococcus cyanobacterium ccmM gene by 35 per cent.

Subsequently, Neisseria gonorrhoeae's inferred amino acid sequence of a purified CA showed that this enzyme goes to the K class. Quite newly, the methanoarchaeon Methanobacterium thermoautotrophicum has acknowledged an alleged L class carbonic anhydrase gene & the enzyme has been categorized and has been exposed to have CA activity. L class has thus now been recognized in all the 3 kingdoms of life. Hence, however the L and the K classes visibly preponderate in the eukaryotes, the prokaryotes characterize all 3 genetically separate classes of carbonic anhydrase.

1. History

In vertebrate erythrocytes, CA was first discovered. A catalyst that might dehydrate HCO3 & let it escape as carbon dioxide was presumed to be present in blood. CA became the first recognized metalloenzyme (K) after its discovery that Zn act as an intrinsic cofactor. It was also found that sulfanilamide inhibits enzymes which contributed to pharmacological investigations & the subsequent development of Carbonic Anhydrase inhibitors to treat glaucoma. It took CA more than twenty years to receive some attention of the researchers & it did because of the simple substrates on which it works. It got recognised among the most active catalysts for an acidic base. It also gained interest in the enzyme after Lindskog substituted Zn(2+) with cobalt and found that enzyme activity was unchanged. This has led to numerous studies of metal oxidation and hypotheses that the enzyme's behavior is related to nearby group ionization . In l970s experiments, human erythrocytes were used to isolate & fetch the 1st sequences of Carbonic Anhydrase - I amino acids and Carbonic Anhydrase - II. John Edsall carried out extensive studies on human erythrocyte Carbonic Anhydrase kinetics & also discovered the first effective antagonist, imidazole.

Current CA research continues to explore the receptor binding mechanism and the role of CA in inhibiting ectopic cardiac calcification.

2. Carbonic Anhydrase (CA) Action Mechanism

Active sites contain a zinc ion in carbonic anhydrase & they have a quite similar type of catalytic activity.

A zinc prosthetic group is coordinated by histidine side chains in 3 positions in carbonic anhydrase enzyme. The enzyme's active site contains a specific CO2 pocket that brings it closer to the $Zn(OH)_2$ group. This results in an attack of electron- rich hydroxide ion on carbon dioxide and thus create a molecule of bicarbonate [2].

3. Structure

The structures of Carbonic Anhydrase are evolutionary and are also unique, distinguished by dissimilar groups such as alpha anhydrases (α), beta anhydrases (β), gamma anhydrases (γ) and delta anhydrases (δ). Most of these are found, whereas there are also two additional groups of anhydrases called epsilon anhydrases (ε) and zeta anhydrases (ζ), both restricted to the bacterial kingdom and protists. The α -class can be find in the periplasm throughout the animal kingdom. As reported by Chirica et al, the alpha anhydrases were commonly found in eubacteria, vertebrates & algae. Most commonly found in higher plants are β - carbonic anhydrases; micro-algae; eubacteria; archaebacteria; cyanobacteria and fungi. γ-carbonic anhydrases in the algae are found. Only some diatoms haven identified δ-carbonic anhydrases. The families of δ-carbonic anhydrase seem to be a case of the convergent evolution, with almost zero sequence similarity to the types of alpha, beta or gamma carbonic anhydrase. ε-carbonic anhydrase is a part of carboxysome shell & has some more domains that perform bacterial function. Z-carbonic anhydrase is restricted to malignant protists and resembles a family of β-carbonic anhydrase with some other metals like zinc or cadmium substitution.

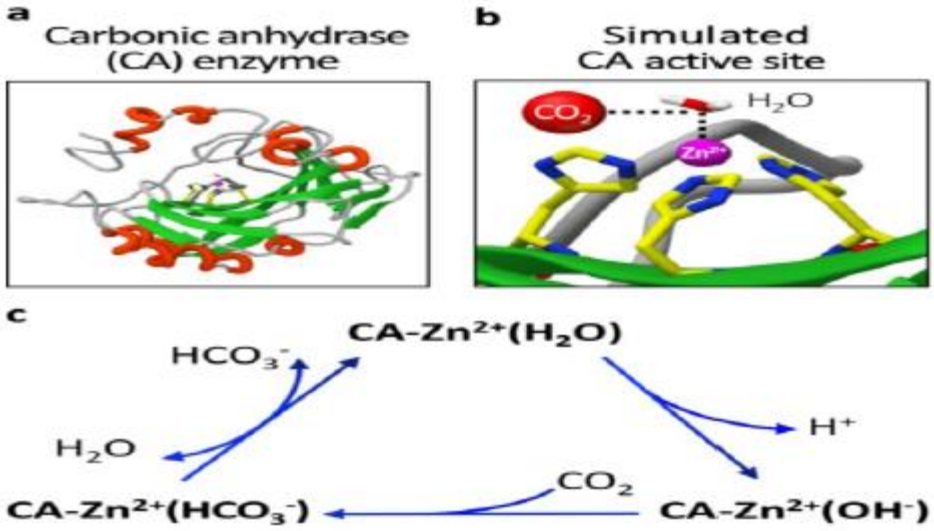


Fig 1: Carbonic Anhydrase enzyme and its mechanism for storing and regenerating CO2.

a) *Carbonic Anhydrase (CA) enzyme ribbon representation.*
b) *Molecular simulated active CA site (vide infra). This active site contains Zn^{2+} ions surrounded by 3 histidine coordinators and an H_2O molecule.*

c) *Depiction of Carbonic Anhydrase active site's overall catalytic process for CO2 hydration to HCO3- with zinc as the metal. This reaction is powered by a gradient of concentration: clockwise when there is more CO2 than HCO3 - and counterclockwise when there is more HCO3-.*

Alpha-class Carbonic Anhydrase (α-CA)

In Anabaena sp, a gene ecaA was acknowledged. Strain PCC7120 which codes an almost 29 kDa protein with a deduced amino acid sequence screening identity with numerous isozymes of the human K class. Even though the EcaA expression is heteroloeous in the E. Coli lead to progress of present entities with almost no noticeable CA activity, antisera consequent from heterologically established EcaA cross-reacted chicken CA II. Polyclonalantisera ensuing from the heterologically molded EcaA also cross-reacted in whole Anabaena cell lysates with a 29-kDa protein.

Polyclonal antisera resultant from the heterologically formed EcaA also cross-reacted in whole Anabaena cell lysates with a 29-kDa protein. The expression of EcaA seems to be controlled on the growth medium by the level of CO2. High levels of EcaA occur in cells grown at increased carbon dioxide level (1% CO2) and low but still perceptible levels are detected after twenty-four hours of low carbon dioxide exposure (0.01%).

EcaA transcription imprints an identical governing design, with high levels of messenger Ribonucleic acid being renowned only in high Carbon dioxide-grown cells. Studies of immunoelectron microscopy using antisera derived from EcaA have uncovered that this protein is present outside of the cell & is linked to cytoplasmic membrane, periplasmic space or the cell wall. But, whole cell lysates and membrane preparations both do not exhibited CA activity below idyllic conditions for EcaA expression; consequently, there is no vigorous signal for the location where this CA is present.

A homologous gene was recognized via primers using Anabaena ecaA sequence and secluded from the Synechococcus sp cyanobacterium. Forme PCC7942. Captivatingly, finding a whole synechocystis genome sequence PCC6803 quantified that this microbe lacks a gene training a carbonic anhydrase of class K. Synechococcus ecaA encodes a 26-kDa protein using some huge quantity of positively charged residues in the amino-terminal domain that suggests existence of the sequence of indications to target membranes. Synechococcus exposed the exact pattern of EcaA expression as Anabaena showed with higher levels of EcaA present inside the cells twisted at increased CO2 level. Deletion mutants of the Synechococcus ecaA were built in an e-ort to recognize the role of EcaA in the CCM of cyanobacteria. In CA activity restrained at surface of the cell as calculated by 180 exchange process, no significant difference was detected among wild-type cells & mutant. The ecaA mutant did not distress active carbon dioxide transport and Na-dependent bicarbonate & Na-independent bicarbonate 3 transport. At low and high Ci, the

ecaA mutant primarily grew slower as compared to wild-type, however the rates were almost blurry during the steady state.

Maximum levels of Cidependent photosynthesis were also identical; but, the mutant grown in low pCO2 displayed a faster increase of oxygen evolution (29 per cent faster) when the small CO2 concentrations were applied to the wild form. Wild-type cells as well as mutant cells both exhibited similar designs of photosynthetic oxygen evolution when they were added to limit concentrations of HCO3 3. An earlier onset of photosynthesis in the EcaA mutant after adding limiting CO2 concentrations advises that the instant external CO2 concentration available for transportation is greater than that available for wildtype.

As carbon dioxide transport is far more rapid than bicarbonate 3 transport it would result in rapid accretion of Ci. Inside wild-type, EcaA's activity quickly hydrates the added carbon dioxide; hence, there is a smaller quantity of carbon dioxide available and bicarbonate 3 would be major transported species. These results advise that EcaA has no role in the CCM; though, under situations of more acidic and alkaline development, the possibility that EcaA can play a significant part. Some other potential roles could be envisaged for EcaA in cyanobacteria. To maximize the availability of substrates for either the CO2 or HCO3 3 transporter, EcaA may be needed for fully equilibrating Ci speciation at transport site. On the other hand, EcaA can also act as a sensor, either detecting or signaling deviations in environmental carbon dioxide levels. Recent signals advises that bicarbonate 3 is a signal for rapid induction of CCM's high aanity state & consequently it is essential to detect the CO2 level.

Beta-class Carbonic Anhydrases (β-CA)

Synechococcus mutants cannot be used for Carbon dioxide[51,130,131] have been secluded but could assemble Ci inside the cell. 30-fold reduction in the activity of carboxysomal CA was seen, this activity complemented the isfA gene and a high partial pressure of carbon dioxide was required. A protein comprising of 272 amino acides was encoded by this gene only, which has 31% of sequence identical to E-CynT, L class CA. Mutants were internally accumulated Ci to wild-type levels and for carbon dioxide fixation, that internal Ci pool was not used and it was all resulted from insertional mutagenesis inactivation of isfA gene.

It was due to such observation Fukuzawa et al. advised for Carbon dioxide based enzyme. The task of the Carboxysomal carbonic anhydrases which is Rubsico, IsfA is to dehydrate bicarbonate to carbon dioxide. No physiological studies have been issued with respect to the localization of IsfA within Synechococcus cells or biochemical characterization of the purified

enzyme. It isolated portion of Deoxyribonucleic Acid from a Synechocystis sp. Subgenomic plasmid library strain PCC6803, which sturdily hybridized to a fragment of the Synechococcus sp isfA gene. Forme PCC7942. The isolated gene, referred to as ccaA (localized cyanobacterial carboxysome), encodes a polypeptide of 274 amino acids with a predictable Mm of 30.7 kDa.

55% amino acid sequence inferred is similar to that of the synechococcus IsfA gene product sequence. Therefore, an improved subcellular fractionation technique used to obtain intact carboxysomes which is free from both the Soluble enzymes and thylakoid membranes was Espie. The 97% of CcaA was associated with integral secluded carboxysomes and this information was revealed by Immunoassays which used CcaA-derived polyclonal antisera. Hence, CcaA is the CA in synechocystis and delivers Rubisco with carbon dioxide. The activity of carboxysomes is inacitivated by dithioreitol and it requires a maximum activity of 20 mM Earlier studies have shown that the activity of carboxysomes CA is disabled by dithiothreitol and needs a maximum activity of 20 mM. However, these properties have not been tested with sanitized CcaA. No CA activity was observed and No cross-reactive anti-CcaA-derived proteins were noticed with the soluble segment or associated with the thylakoid membranes. Irrespective of the other CA arise in the cytoplasm or are connected with the thylakoid membrane, it remains to be fixed. An additional orf (slr0051) with a high degree of sequence similarity to the E in Synechocystis PCC6803 Coli CynT L has been discovered as a CA class. A 263-amino acid polypeptide with a putative lipoprotein lipid membrane attachment site was encoded by the gene, known as EcaB (external CA beta class) which indicates the polyprotein may be either associated with the membrane. But the location of the EcaB has not been identified till now. Thus, the E. inclusion bodies were formed without any noticeable CA activity with coli. The Synechococcus PCC7042 EcaA mutant appeared to be similar to the phenotype of Synechococcus PCC6803 EcaB mutant and no noticeable difference in enzyme activity was found. The K class enzyme EcaA, EcaB do not play any important catalytic role in the CCM 's working. Therefore, CA EcaA class Synechocystis PCC7942 K may have the identical function as the EcaB class Synechocystis PCC6803 L. Thus, for the similar purpose of the above two diverse genera of cyanobacteria may use different classes of CA.

Gamma class carbonic anhydrase (γ-CA)

Synechococcus PCC7942 encodes genes for Rubisco subunits. In this Synechococcus PCC7942, a large proportion of carbon concentration mechanism (CCM) genes were recognized inside an upstream region of the rbcLS operon. Each carbon concentration mechanism gene is vital for low pCO2 growth and for proper carboxysome assembly. CcmM codes a bifunctional gene containing of five hundred and thirty-nine amino acids, of which roughly two hundred and nineteen are thirty-five per cent identical to Cam, the prototype Q CA of M. THERMOFIL.

3 internal repeats in the C-terminal three hundred and twenty amino acids display 45^51 per cent identity to the Anabaena Rubisco activase C-terminal region.

Rubisco activase catalyzes the Rubisco^ribulose 1,5-biphosphate complex's ATP-dependent stimulation that is dynamically coupled with the photosynthesis light reactions. These internal repeats also demonstrate similarity to RbcS. The homologous area contains the small subunit of the Anabaena Rubisco amino acids Arg-86 and Asp-91, which were presumed to be intricate in binding together large and small Rubisco subunits. Genes with identical inferred amino acid sequences were also present in Synechocystis sp, but with four RbcS domains. Coli, CcmM is adequately plentiful protein located in carboxysomes even if it is not known whether it is located in or on the carboxysome peripheral shell.

In carboxysome-enriched grounding, the CcmM antisera also cross-reacts with a 35-kDa protein, but it isn't recognized whether this polypeptide is a processed form of CcmM, a 2nd site translation product, or an artifact. CA activity for the heterologously expressed CcmM in E has not been reported yet. It remains to be strongminded coli & its role in carbon dioxide emission. Aside from a structural role, CcmM may also avoid carbon dioxide from dripping out of the carboxysome. Any carbon dioxide from the carboxysome gets converted to HCO3 3 that futher gets dehydrated to stock Rubisco with CO2 through the carboxysomal L class CA.

1. Composition

Sixteen CA isozymes have been described for mammals. Carbonic anhydrases I, II are the most well-known erythrocyte carbonic anhydrases. In saliva, Carbonic anhydrases VI is secreted. The mitochondrial Carbonic anhydrases VA and VB. CARP-VIII, X, XI are the three forms of carbonic anhydrases related proteins.

Zn metal is connected to his 93, 95, and 118 (numbering the mature chain). The zinc bonded water molecule is connected with a bonded hydrogen network, comprising 28-Ser,91-Glu, 105-Glu, 106-His, 116-His, 193-Tyr, 198-Thr, 208-Trp, and 223-Asn and these compounds were easily found. The unique C-terminal knot structure found in Bovine and human carbonic anhydrases I and II has been denoted to be significant in enzymatic and mechanical properties.

2. CAs in Human Genome

By sequence comparison and gene mapping, Carbonic anhydrases I, II, III are confirmed as unique carbonic anhydrases.

The three genes were mapped in the order of Carbonic anhydrases II, III, I to a gene cluster on chromosome.

After the completion of the human genome, a total of 15 Carbonic anhydrases isoforms have been identified. These carbonic anhydrases isozymes differentiate among the cell position and catalytic efficiency. The cytosolic Carbonic anhydrases are I, II, III, IV, VIII, X, XI and

XIII. The membranous Carbonic anhydrases are IV, IX, XII and the mitochondrial are Carbonic anhydrases VA and VB.

3. CA Crystal Structures

For the families of alpha, beta, gamma and carbonic anhydrases only crystal structures have been found till date. The Protein Database consists of more than 900 Carbonic anhydrases structures, resolved through X-ray or neutron crystallography. Majority of these structures constitute the alpha class, including the structures that are catalytically active except for carbonic anhydrases V for all human isoforms. A carbonic anhydrases along with specific inhibitor group includes the great count of structures of carbonic anhydrases complexes, depicting the importance of the drug targets on the carbonic anhydrases.

4. Specificity

In erythrocytes, the transport and excretion of carbon dioxide in the blood totally depends on carbonic anhydrase's fast catalysis of carbon dioxide reactions. Bovine carbonic anhydrases exhibit hydratase activity against a large range of substrates and pyruvates reversibly.

5. Catalytic Activity

When the catalyst is absent the spontaneous reversible carbon dioxide hydration has an effective rate of n the absence of the catalyst the spontaneous reversible 0.15 s-1, whereas the constant rate of 50s-1 is shown by the reverse reaction. The processes like carbon dioxide hydration and bicarbonate dehydration are linked to quick processes which are related to secretory or transport method. The most important role of carbonic anhydrase's is to catalyze the hydration of carbon dioxide at a high rate with Kcat which ranges between 104 to 106 s-1. Generally, the effective catalysis for the carbon dioxide hydration reaction are the bacterial carbonic anhydrases, which belongs to the three known classes of carbonic anhydrases i.e. alpha, beta and gamma. During the studies of 3D bacterial carbonic anhydrases, it was found that the catalytic pocket is small for alpha class and larger for beta class and largest for gamma class.[5]

6. Engineered Bacteria with thermostable CA for trapping CO2

A method for carbon dioxide reduction was required essentially. In order to develop a working catalyst for capturing the carbon dioxide, by the high-density fermentation of E coli. cultures with recombinant expression of Ssp carbonic anhydrases was represented. The covalent immobilization method was done by using the carbodiimide activation reaction on the surface of MNP. This provided two main advantages: 1) The MNP – immobilized Ssp carbonic anhydrases via carbodiimide enhanced the biocatalyst's stability and storage. 2) By adding a magnet or an electromagnetic field having the ferromagnetic properties could be retrieved and reused. But this process requires a large amount of price as the method consists two process one is purification of biocatalysts and the second is the immobilization of the enzymes. This is on of the main aspect through which this method could be discouraged for industrial applications. [5]

PROCEDURE

1. Make a PDB for our enzyme of interest i.e. Bacterial Carbonic Anhydrases.
2. Filter the search results for the organism of interest, in our case, E coli.
3. The completion of this step will get us some PDB ids. We got seven.
4. From these seven, one PDB id was selected and we ran blast for PDB database and setting the E-value to $10e^{-6}$.
5. The result of this step gave us 29 similar PDB ids from different organisms.
6. Removed the het atoms and water molecules of all the PDB ids using a python code.
7. A graph between PDB id and percentage identity was plotted using Python.
8. A table containing the details of the PDB ids that were obtained was formed. The parameters included in the table were: PDB id, Organism name, percentage identity, sequence length, resolution, e value, maximum scores, sequence lengths, metal ion and chains.
9. The structures of all the PDB ids were analysed and studied in detail.

Results

PDB_ID	ORG_NAME	% IDENTI	METAL_ATM	E VALUE	MAX_SCO	CHAINS	RESOLUTIO	SEQ_LEN
1I6P_A	Escherichia coli]	100	ZN	3.23E-166	457	A	2	220
1I6Q_A	Escherichia coli]	98.174	ZN, MSE	3.55E-161	444	A, B	2.2	220
3OY1_A	Salmonella enterica					A, B	1.54	223
2ARC_A	Haemophilus influenz	84.084	ZN, SO4	2.91E-102	296	A, B, C, D, E, F	2.3	229
3MF3_A	Haemophilus influenz	84.084	CO, ACT	3.07E-102	295	A, B, C, D, E, F	2.5	221
3E2X_A	Haemophilus influenz	83.617	ZN, SO4	9.93E-102	294	A, B	2.55	229
3E28_A	Haemophilus influenz	83.617	ZN, SO4	1.03E-101	294	A, B, C, D, E, F	2.5	229
3E3G_A	Haemophilus influenz	83.617	ZN, SO4	1.39E-101	294	A, B, C, D, E, F	2.3	229
3EIV_A	Haemophilus influenz	83.617	ZN	1.66E-101	294	A, B	2.8	229
3E24_A	Haemophilus influenz	83.617	ZN, PO4	3.64E-101	293	A, B	2.301	229
4WAJ_A	Haemophilus influenz	82.15	ZN, SO4	2.62E-100	291	A, B	2.7	229
4WAK_A	Haemophilus influenz	82.15	ZN, BCT, K	6.62E-100	290	A, B	2.49	229
5OJK_A	Vibrio cholerae]	61.818	ZN, BCT	1.21E-99	289	A, B, C, D, E, F, G, H	1.9	222
4WAM_A	Haemophilus influenz	82.215	ZN, PO4	7.16E-98	285	A, B	2.2	229
4RXY_A	Pseudomonas aerugin		ZN, GOL	9.89E-86	253	A	1.901	215
5BOJ_A	Pseudomonas aerugin	64.612	ZN, CO2	3.33E-84	249	A	1.6	209
4O1K_A	Sordaria macrospora]	51.053	ZN	2.32E-69	213	A	1.6	209
1DDZ_A	Porphyridium purpure	46.675	ZN	1.89E-68	218	A, B	1.83	233
3UCJ_A	Coccomyxa sp. PA]	40.201	ZN, ACT, GOL		166	A, B	1.85	227
4O1J_A	Sordaria macrospora]	37.879	ZN, CL	1.79E-42	144	A, B	2.695	255
2W3N_A	Cryptococcus neoformans	33.796	ZN, ACT, GOL	1.36E-35	126	A, B, C	2.05	239
2W3O_A	neoformans]	33.796	ZN, CL	1.36E-35	126	A	1.3	243
6GWU_A	Candida albicans]	38.596	ZN, BME, PEG		115	A, B, C, D	2.2	208
3EYX_A	Saccharomyces cerevisiae]	36.5	ZN, ACT, EDO	3.03E-23	95.1	A, B	2.04	216
2A5V_A	Mycobacterium tuberc	34.969	ZN, SO4	6.87E-16	74.3	A, B, C, D	2.2	213
1YM3_A	Mycobacterium tuberc	34.969	ZN, MG	6.75E-16	74.3	A	1.75	215

PDB_ID	ORG_NAME	% IDENTI	METAL_ATM	E_VALUE	MAX_SCOR	CHAINS	RESOLUTIO	SEQ_LEN
1I6P_A	Escherichia coli]	100	ZN	3.23E-166	457	A	2	220
1I6O_A	Escherichia coli]	98.174	ZN, MSE	3.55E-161	444	A, B	2.2	220
3QY1_A	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2]					A, B	1.54	223
2ABC_A	Haemophilus influenzae]	81.084	ZN, SO4	2.91E-102	296	A, B, C, D, E, F	2.3	229
3MF3_A	Haemophilus influenzae]	81.084	CO, ACT	3.07E-102	295	A, B, C, D, E, F	2.5	221
3F2X_A	Haemophilus influenzae]	81.617	ZN, SO4	9.90E-102	294	A, B	2.55	229
3F28_A	Haemophilus influenzae]	81.617	ZN, SO4	1.03E-101	294	A, B, C, D, E, F	2.5	229
3E3G_A	Haemophilus influenzae]	81.617	ZN, SO4	1.39E-101	294	A, B, C, D, E, F	2.3	229
3E1V_A	Haemophilus influenzae]	81.617	ZN	1.60E-101	294	A, B	2.8	229
3E24_A	Haemophilus influenzae]	81.617	ZN, PO4	3.64E-101	293	A, B	2.301	229
4WAJ_A	Haemophilus influenzae]	82.15	ZN, SO4	2.82E-100	291	A, B	2.7	229
4WAK_A	Haemophilus influenzae]	82.15	ZN, BCT, K	6.62E-100	290	A, B	2.49	229
5O3K_A	Vibrio cholerae]	61.818	ZN, BCT	1.21E-99	289	A, B, C, D, E, F, G, H	1.9	222
4WAM_A	Haemophilus influenzae]	81.215	ZN, PO4	7.16E-98	285	A, B	2.2	229
4RXY_A	Pseudomonas aeruginosa]	54.930	ZN, GOL	9.85E-86	253	A	1.901	215
5BOJ_A	Pseudomonas aeruginosa]	54.412	ZN, CO2	3.33E-84	249	A	1.6	209
4O1K_A	Sordaria macrospora]	51.053	ZN	2.32E-69	213	A	1.6	209
1DDZ_A	Porphyridium purpureum]	46.875	ZN	1.89E-68	218	A, B	1.83	233
3UCJ_A	Coccomyxa sp. PA]	40.201	ZN, ACT, GO	8.68E-51	166	A, B	1.85	227
4O1J_A	Sordaria macrospora]	37.879	ZN, CL	1.79E-42	144	A, B	2.695	255
2W3N_A	Cryptococcus	33.796	ZN, ACT, GO	1.36E-35	126	A, B, C	2.05	239
2W3O_A	neoformans]	33.796	ZN, CL	1.36E-35	126	A	1.3	243
6GWU_A	Candida albicans]	38.596	ZN, BME, PE	6.09E-32	115	A, B, C, D	2.2	208
3EYX_A	Saccharomyces cerevisiae]	35.5	ZN, ACT, EDO	3.03E-23	95.1	A, B	2.04	216
2A5V_A	Mycobacterium tuberculosis]	31.884	ZN, SO4	6.87E-16	74.3	A, B, C, D	2.2	213
1YM3_A	Mycobacterium tuberculosis]	31.884	ZN, MG	6.75E-16	74.3	A	1.75	215

CONCLUSION

Carbonic anhydrase performs a quite critical part in the accumulating all of the processes of photosynthetic carbon dioxide.

As the main culprit behind global warming, carbon dioxide is a problem both for the present generation and for future generations. The need for an hour is, therefore, to think about the present and the future's sustainability.

This study aims to show us, how the physiological role of carbonic anhydrase present in the plants to solve the issue of the atmospheric greenhouse gas such as carbon dioxide & its vital role for carbon sequestration as a biochemical marker.

In addition, thermally stable and catalytically highly efficient, bacterial Carbonic anhydrases,

due to the accumulation of this gas in the atmosphere may have some interesting applications for biomimetic carbon dioxide capture in the context of global warming because of anthropogenic activities. In addition, CA-activators of these enzymes could be a better alternative to reduce global warming.

REFERENCES

[1] Kupriyanova, E., Pronina, N. and Los, D. (2017). Carbonic anhydrase — a universal enzyme of the carbon-based life. *Photosynthetica*, 55(1), pp.1-2.

[2] Bhat, F., Ganai, B. and Uqab, B. (2017). Carbonic Anhydrase: Mechanism, Structure and Importance in Higher Plants. *Asian Journal of Plant Science and Research*, (2249-7412), pp.17-23.

[3] Alsharidi, A., Al-Hamed, M. and Alsuwaida, A. (2015). Carbonic anhydrase II deficiency: report of a novel mutation. *CEN Case Reports*, 5(1), pp.108-112.

[4] Fu, Y., Jiang, Y., Dunphy, D., Xiong, H. and Croissant, J. (2018). Ultra-thin enzymatic liquid membrane for CO2 separation and capture. *Nature Communications*, 9(1).

[5] Supuran CT, Capasso C. An Overview of the Bacterial Carbonic Anhydrases. Metabolites. 2017 Nov 11;7(4):56. doi: 10.3390/metabo7040056. PMID: 29137134; PMCID: PMC5746736.

[6] Ferry, J. G. (2015). Carbonic Anhydrases of environmentally and medically relevant anaerobic prokaryotes. Prokaryotic Carbonic Anhydrases, 325-336. https://doi.org/10.1016/b978-0-444-63258-6.00018-4.

Appendix

Python Script used:

- **Relationship between PDB IDs on x- axis and their % similarity on y- axis, using Python code.**

```
import csv
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
fh = csv.reader(open("C
    :\\Users\\161516\\Desktop\\CarbonicAnhydrase\\identity_table
    .csv"), delimiter = ",")
next(fh)
pdb_id = []
identity = []
for line in fh:
  print(line)
pdb_id.append(line[0])
identity.append(float(line[1]))
x = np.random.randn(100)
plt.hist(x, bins = range(10, 100))
plt.plot(pdb_id, identity)
plt.show()
```

- **To remove water molecules from PDB IDs**

```
import re
import os
path = "C:\\Users\\161516\\Desktop\\Carbonic Anhydrase\\ids"
for filename in os.listdir(path):
  path_r = path + "\\" + str(filename)
fh1 = open(path_r)
filename2 = filename.strip(".pdb")
filename2 = filename2 + "_new.pdb"
path_w = path + "\\" + str(filename2)
fh2 = open(path_w, "w+")
for line in fh1:
  result = re.match(r "HETATM.*", line)
result2 = re.search(r "CONECT.*", line)
if result:
  fh2.write("")
elif result2:
  fh2.write("")
else :
  fh2.write(line)

fh1.close()
fh2.close()

```