Computational Biology and Chemistry 61 (2016) 121–129

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/compbiolchem

Research Article

On origin and evolution of carbonic anhydrase isozymes: A phylogenetic analysis from whole-enzyme to active site

Srijoni Banerjee[a], Parag A. Deshpande[a,b,*]

[a] Quantum and Molecular Engineering Laboratory, Department of Chemical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India

[b] Centre for Theoretical Studies, Indian Institute of Technology Kharagpur, Kharagpur 721302, India

ARTICLE INFO

Article history:
Received 5 August 2015
Received in revised form 9 November 2015
Accepted 19 January 2016
Available online 23 January 2016

Keywords:
Kingdom-specificity
Active site
Sequence alignment
Carbonic anhydrases
Biological function
Structural similarity

ABSTRACT

Genetic evolution of carbonic anhydrase enzyme provides an interesting instance of functional similarity in spite of structural diversity of the members of a given family of enzymes. Phylogenetic analysis of α-, β- and γ-carbonic anhydrase was carried out to determine the evolutionary relationships among various members of the family with the enzyme marking its presence in a wide range of cellular and chromosomal locations. The presence of more than one class of enzymes in a particular organism was revealed by phylogenetic time tree. The evolutionary relationships among the members of animal, plant and microbial kingdom were developed. The study revises a long-established notion of kingdom-specificity of the different classes of carbonic anhydrases and provides a new version of the presence of multiple classes of carbonic anhydrases in a single organism and the presence of a given class of carbonic anhydrase across different kingdoms.

1. Introduction

The name carbonic anhydrase (CA) represents a family of enzymes which is found in all three forms of life, *i.e.*, animals, plants and micro-organisms. The classical notion about CA described it to be derived from a single gene, now known as the α-CA gene, which later proved to be wrong with the acceptance of β- and γ-CA as genetically distinct. A comparison of amino acid sequences and crystal structures of three main classes of CA, *viz.*, α-, β-, and γ-CA, shows that the three classes are evolutionarily different and have evolved independently (Tripp et al., 2001). Yet, interestingly, all of them catalyze the reversible hydration of CO_2 to a bicarbonate ion and a proton ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$). The interchange $CO_2 \leftrightarrow HCO_3^-$ in a human body under different biological conditions has been reported as early as in 1930s by Meldrum and Roughton (1933), an era when α-CA was synonymous to CA. However, with the discovery of β-CA in plants (Hewett-Emmett et al., 1984) and γ-CA in micro-organisms (Alber and Ferry, 1994), the genetic characteristics and diversity of CA became a subject of interest. It is needless to say, however, that due to interesting structural and functional spectra provided by carbonic anhydrase isozymes, newer classes of this enzyme are continuously being discovered with δ-, ζ-, and η- carbonic anhydrases (Zimmerman and Ferry, 2008; Kumar and Ferry, 2014; Del Prete et al., 2014) reported in the recent past. However, the we confine the scope of the present study to α-, β- and γ- carbonic anhydrases mainly due to the fact that the current study is based upon bioinformatics tools and only these three classes of carbonic anhydrases have been reported well-enough in the literature to have the critical mass for the database to be worthwhile for any bioinformatics study.

The CA family is distinct from other families of enzymes because of the diversity exhibited by them both in cellular distribution and in the putative or biological functions of its member enzymes which range from photosynthesis, to respiration, pH homeostasis, and ion transport (Tashian, 1989). Although α-, β- and γ-CA have no amino acid sequence similarity, they are functionally identical biocatalysts. The sixteen isozymes of α-CA possess an active site consisting of a Zn^{2+} ion coordinated tetrahedrally to three His residues and a water molecule (Eriksson et al., 1988). The X-ray structure of β-CA, derived from red alga *Porphyridium purpureum*, reveals that the dimeric enzyme consists of two identical structural motifs with each motif carrying three projecting α helices (Mitsuhashi et al., 2000). It also consists of a Rossmann fold and an anti-parallel β-strand. This motif is remarkably different from

* Corresponding author at: Quantum and Molecular Engineering Laboratory, Department of Chemical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India.
E-mail address: parag@che.iitkgp.ernet.in (P.A. Deshpande).

http://dx.doi.org/10.1016/j.compbiolchem.2016.01.003
1476-9271/

those in α and γ-CA. Two zinc binding are sites present on both sides of a monomer of β-CA. The zinc atom is coordinated to Cys, Asp, His and Cys residues. No water molecule is present within the coordination radius of the zinc ions. X-ray crystallographic structure of γ-CA from archaeon *Methanosarcina thermophila* (Kisker et al., 1996) reveals a unique structure. The trimeric molecule has a completely different fold when compared to α and β-CA. Each subunit of the trimeric molecule is dominated by seven turns of a left handed β-helix, with three short strands per turn. The zinc ions are located between the subunits and the ions are coordinated by two His residues from one subunit and another His residue from the neighboring subunit. A water molecule is also present in the active site.

Since the reversible hydration of CO_2 catalyzed by CA has been of widespread adaptive value to the early organisms, it is possible that it was one of the first enzymes to appear (Tashian, 1989). The original gene family of CA probably arose before the divergence of prokaryotes and eukaryotes more than half a billion years ago and for this prolonged period in the evolutionary history of CA, the CA genes underwent several rounds of erroneous duplication resulting

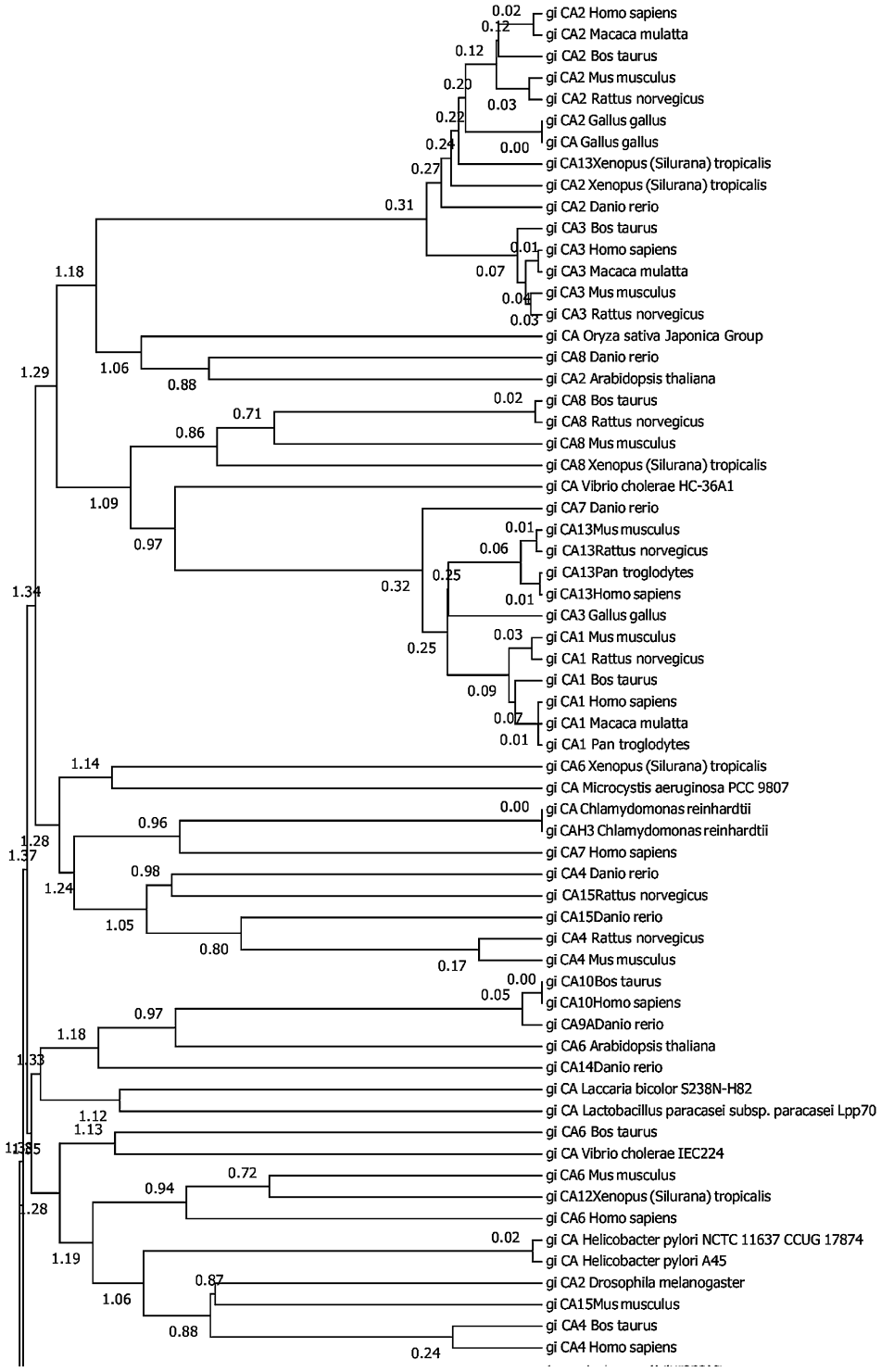

Fig. 1. Phylogenetic time tree of α-CA.