Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme

Tobias J Erb and Jan Zarzycki

Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) is arguably one of the most abundant proteins in the biosphere and a key enzyme in the global carbon cycle. Although RubisCO has been intensively studied, its evolutionary origins and rise as Nature's most dominant carbon dioxide (CO_2)-fixing enzyme still remain in the dark. In this review we will bring together biochemical, structural, physiological, microbiological, as well as phylogenetic data to speculate on the evolutionary roots of the CO_2-fixation reaction of RubisCO, the emergence of RubisCO-based autotrophic CO_2-fixation in the context of the Calvin-Benson-Bassham cycle, and the further evolution of RubisCO into the 'RubisCOsome', a complex of various proteins assembling and interacting with the enzyme to improve its operational capacity (functionality) under different biological and environmental conditions.

Address
Max Planck Institute for Terrestrial Microbiology, Department of Biochemistry and Synthetic Metabolism, Karl-von-Frisch-Str. 10, 35043 Marburg, Germany

Corresponding author: Erb, Tobias J (toerb@mpi-marburg.mpg.de)

Current Opinion in Biotechnology 2018, **49**:100–107

This review comes from a themed issue on **Plant biotechnology**

Edited by **Alisdair Fernie** and **Joachim Kopka**

http://dx.doi.org/10.1016/j.copbio.2017.07.017

More than 90% of the inorganic carbon that is converted into biomass is fixed by the enzyme RubisCO that catalyzes the carboxylation and cleavage of ribulose-1,5-bisphosphate (RuBP) into two molecules of 3-phosphoglycerate (3PG). RubisCO is found in all three domains of life: bacteria, archaea and eukaryotes. The enzyme makes up 30–50% of the soluble protein in plant leaf and it has been estimated that for every person on earth there is 5 kg of RubisCO [1]. Altogether, this makes RubisCO one of the most abundant enzymes in the global carbon cycle that literally feeds life on earth.

Despite its dominant role in carbon fixation, the enzyme has some peculiarities. First, RubisCO requires a post-translational activation to perform the carboxylation reaction: a conserved lysine residue in the active site needs to be carbamylated in order to complex an Mg^{2+} ion that is in turn required for activity. Second, the enzyme is a rather slow catalyst. The turnover frequency of an average RubisCO is only between 1 and 10 s^{-1} (http://brenda-enzymes.org), which can make it a limiting factor in photosynthetic CO_2-fixation under optimal conditions. Finally, RubisCO makes mistakes. Besides the carboxylation reaction RubisCO catalyzes a non-productive oxygenation side-reaction that leads to the formation of 2-phosphoglycolate (2PG). 2PG is a toxic compound that inhibits several enzymes in central carbon metabolism [2–4]. An average C_3-plant RubisCO has an error rate of more than 20%. This number can even add up to more than 40% at high temperatures and/or low intracellular CO_2, which results in substantial amounts of 2PG that are formed during photosynthesis [5–7]. 2PG is recycled in an energy-demanding process called photorespiration. It has been estimated that approximately 30% of the photosynthetic energy in plants is wasted through photorespiration [7,8].

Why is RubisCO so inefficient? Evidence accumulated that enzyme activity and specificity are reciprocally linked with each other in RubisCO [9•,10•,11,12]. A faster RubisCO has a higher error rate and a more specific RubisCO has a lower catalytic rate. This link has mechanistic and evolutionary reasons. RubisCO evolved before the first great oxygenation event in an atmosphere without oxygen (O_2), so that its mechanism was not constrained by O_2. However, with the rise of atmospheric O_2 concentrations to modern-day levels, as a result of the second great oxygenation event, RubisCO had to learn to discriminate between CO_2 and O_2. Because discrimination usually comes at the cost of reduced catalytic rate, a more specific enzymes almost inevitably becomes a slower catalyst [13]. As a consequence RubisCO had to evolve along a Pareto front of enzyme activity and specificity, a trade-off in which the modern enzyme apparently became trapped.

Although this part of RubisCO's evolutionary history is generally accepted, the answer to the question how RubisCO arose to become Nature's predominant CO_2-fixing enzyme is less well known. In this article we leave firm grounds to illuminate the emergence of 'proto-RubisCO', the first RubisCO homolog that catalyzed a carboxylation reaction, its integration into the Calvin-Benson-Bassham cycle and its further evolution into the modern day CO_2-fixing enzyme complex, called the 'RubisCOsome'. The picture we draw is based on works, ideas, discussions and comments of many

colleagues and is summarized in three main hypothesis ('opinions') that are discussed in the following.

RubisCO evolved from a non-CO_2-fixing ancestor

How did RubisCO emerge as Nature's most dominant CO_2-fixing biocatalyst? A key for understanding the evolution of RubisCO are RubisCO-like proteins (RLPs) [14•]. RLPs are phylogenetically and evolutionarily related to RubisCOs, with which together they form the RLP/RubisCO enzyme superfamily (Figure 1). While RLPs share substantial sequence identity with RubisCOs, they lack active site amino acid residues that are known to be essential for the carboxylation reaction of RubisCO.

For several RLP subfamilies a biochemical and physiological function has been determined. The tautomerase subfamily of RLPs operates in a variant of the ubiquitous 'methionine salvage' pathway that recycles the dead end metabolite methylthioadenosine (MTA) into L-methionine [15•]. The 1,3-isomerase subfamily of RLPs function in the so-called MTA-isoprenoid shunt that channels MTA into isoprenoid biosynthesis [16]. Recently, a subfamily of decarboxylase RLPs was identified that serve in a degradation pathway of the four carbon sugar acids erythronate and threonate [17].

Based on sequence diversity, functional evidences, and genomic context, at least four more RLP subfamilies can be distinguished [16–19] that are most likely also isomerases and/or epimerases [20]. Thus, the RLP/RubisCO superfamily features at least seven functionally distinct RLP subfamilies besides the subfamily of 'true RubisCOs' that can be divided into three distinct phylotypes (RubisCOs Form I, II and III).

How did the different functions in the RLP/RubisCO superfamily emerge during evolution, and among the different functions, how in particular, did the carboxylation trait evolve that is apparently restricted to the subfamily of true RubisCOs only? Two alternative hypotheses can be formulated that are mutually exclusive.

Hypothesis 1.
The carboxylation function in the RLP/RubisCO superfamily is an ancient trait [21]. According to this hypothesis, the carboxylation reaction is an original function and RLPs evolved from a CO_2-fixing, ancestral enzyme (a 'proto-RubisCO') by loss of the CO_2-fixation activity.

Hypothesis 2.
The carboxylation function in the RLP/RubisCO superfamily is a novel trait that was acquired during evolution [22,23]. According to this hypothesis, RubisCO emerged from a non-CO_2-fixing ancestor. The emergence of the carboxylation function is a secondary event and all RubisCOs are of monophyletic origin.

Although the history of evolutionary events cannot be recapitulated in retrospect, several arguments can be considered to delineate the most plausible evolutionary scenario. Valuable information lies in the various reactions that are catalyzed by the RLP/RubisCO superfamily. The enzymes of the different subfamilies catalyze distinct biochemical reactions. Yet, they share some basic aspects. First, all reactions in the RLP/RubisCO superfamily center on structurally similar substrates, a C_5 or C_4 sugar derivative that features a phosphate in the C1 and a keto group in the C2 or C3 position (Figure 1). Second, a general catalytic mechanism can be formulated that is conserved in all members of the RLP/RubisCO superfamily characterized to date. This core mechanism includes the formation of a central enolate intermediate. In most cases this happens through the acid/base catalyzed abstraction of a proton adjacent to the keto group [24,25]. In a subsequent step, the central enolate intermediate then attacks an electrophile that can be a proton, carbon dioxide or oxygen molecule [26].

While the initial steps of catalysis that lead to formation of the central enolate intermediate are mostly conserved between the different subfamilies, they diverge in the later steps of the catalytic cycle. This results in a diverse outcome of reaction products and explains the different catalytic functions that evolved in the RLP/RubisCO superfamily (Figure 1). In case of the tautomerase and the decarboxylase RLP subfamilies, the enolate intermediate simply attacks a proton to yield the reaction product. In case of the 1,3-isomerase RLP subfamily two subsequent proton abstraction and attacks take place. Finally, in RubisCOs the central enolate intermediate attacks CO_2 (or O_2), which is followed by a subsequent water-mediated hydrolysis reaction to generate the two 3PG molecules.

Apparently, there is an increasing mechanistic complexity from the simple tautomerase reaction to the multi-step reaction of RubisCO. It is fair to conclude that the increase in mechanistic complexity observed in the RubisCO subfamily reflects an evolutionary development and that the more complex mechanism of RubisCO is rather the end of an evolutionary trajectory than its starting point (Figures 1 and 2).

This line of mechanistic evidence is further supported by structural arguments. All members of the RLP/RubisCOs superfamily are eight-stranded α/β-barrels, which is the most frequent and most versatile protein fold used by evolution [27]. However, in contrast to all other members of the RLP/RubisCO superfamily, RubisCOs feature an additional β-hairpin following loop number 6 that carries residues essential for the carboxylation reaction. Applying the principle of Occam's razor, it is rather likely that the subfamily of RubisCOs gained this additional stretch of amino acids during evolution compared to the possibility that all other subfamily members lost this additional