

## Origin and evolution of metabolisms

*Juli Peretó***Introduction: some conceptual remarks on metabolism**

Metabolism is the set of enzymatic reactions that allow living beings to use external energy sources to drive the building of their biochemical components from external chemical sources and also to carry out energy-consuming functions, such as osmotic and mechanical work. The role played by gene-encoded enzymatic catalysts is one of the essential properties of life. Since their stability is finite and there is a need for constant replacement, enzymes are themselves products of metabolism (Cornish-Bowden *et al.*, 2004). Thus, the proteome (i.e. the totality of proteins and their concentrations that exists in a particular cell state) is a product of the metabolome (i.e. the totality of metabolites and their concentrations that exists in a particular metabolic state). This situation gives rise to the ‘metabolic circularity’ or ‘recursivity’; a concept needed for a complete understanding of metabolism (Cornish-Bowden *et al.*, 2007).

Extant metabolic networks are certainly complex, with hundreds or thousands of concatenated enzymatic reactions. Since there are a limited number of coenzymes (i.e. the special reactants that help enzymes in performing their catalytic mechanisms), recurrently used by different enzymes, and some central metabolites are true crossroads between different lines of chemical transformation, complex networks emerge. From a topological perspective, metabolic networks show a power-law distribution of connectivity (Fell and Wagner, 2000). In other words, most metabolites are poorly connected whereas a few of them (coenzymes and metabolic crossroads) support many connections. One of the challenges of the evolutionary study of metabolism is to unravel how these complex networks emerged, from an expectedly simpler chemical network, based on non-enzymatic reactions.

Traditionally, biochemists have identified certain transformations, from sources to sinks, with physiological significance, named ‘metabolic pathways’. These pathways can be either ‘catabolic’ (i.e. using energy- and electron-rich metabolites to produce useful forms of metabolic energy) or ‘anabolic’ (i.e. energy and electron-consuming pathways aimed at building the cellular components). Obviously, behind the necessary simplification of traditional biochemical studies, many more chemical transformations exist in the cell as a result of the complexity of connections between pathways. Current systems-biology approaches,

Table 18.1. The metabolic modes. Metabolisms are classified taking into account the external sources of energy (visible light and either inorganic or organic oxidative reactions), reducing power (electrons from inorganic or organic compounds) and carbon.

Mode	Energy source	Reducing power	Carbon source	Examples
chemo-litho-autotrophic	oxidation of inorganic compounds	inorganic compounds	CO <sub>2</sub>	hydrogen-oxidizing bacteria, methanogens, denitrifying bacteria
photo-litho-autotrophic	visible light	inorganic compounds	CO <sub>2</sub>	photosynthetic bacteria, plants
photo-organo-heterotrophic	visible light	organic compounds	organic compounds	nonsulphur purple bacteria, heliobacteria
chemo-organo-heterotrophic	oxidation of organic compounds	organic compounds	organic compounds	animals, fungi, many prokaryotic and eukaryotic microorganisms

which try to understand the cell functioning from a holistic perspective, using the data from 'omics' technologies (genomics, transcriptomics, proteomics, metabolomics etc., i.e. the totality of genes, RNAs, proteins, metabolites etc. associated with one biological system) are unveiling, in part, this complexity (Wagner, 2007).

On a global scale, life is sustained by a flux of energy – either visible light or from chemical sources – throughout the metabolic networks. The different combinations of energy, electron and carbon sources offer a diversity of metabolic modes (Table 18.1).

Darwin (1859) cited, as one of the difficulties of his natural selection theory, the existence of organs of extreme perfection and complication. As a case study he used the eye. But the metabolic machinery is by far more complicated. The eye evolved no fewer than forty times – and probably more than sixty times – independently in various animal lineages (Lane, 2009), whereas oxygenic photosynthesis, for example, was invented only once by the ancestors of extant cyanobacteria (Barber, 2008). As natural selection fully explains the origin of the eyes, we are confident that it will be possible to learn the natural causes for the origin and evolution of metabolism.

### The chemical roots of metabolism

Although the problem of the origin of metabolism is intimately intertwined with the problem of the origin of life itself, most authors have instead focused their attention on the origin of the genetic information. Following the Oparin tradition, it is generally assumed that life originated with a heterotrophic metabolic mode, i.e. from a 'primitive soup' rich in organic compounds from which the first organisms gained energy and matter (Oparin,

1924; Haldane, 1929; Lazcano *et al.*, 1992; Peretó *et al.*, 1997). During the last few decades scientists have accumulated many data favouring the contributions of volcanic, atmospheric and cosmic chemistries to the inventories of abiotic compounds and processes in the early Earth. Out of this abiotic chemistry, the emergence of suprachemical (or infra-biological) subsystems exhibiting basic lifelike performances, such as self-reproductive vesicles, self-replicative polymers and self-maintained chemical networks, was a necessary step during the chemical evolution or prebiotic phase. Eventually, the harmonic articulation of those three prebiotic subsystems in the same functional framework could be considered as the beginning of biological evolution. Thus, any evolutionary scenario, to be complete, must give an explicit account of the mechanisms of energy and matter fluxes through these primitive systems.

One of the first questions to be addressed is whether there is a correlation or continuity between prebiotic chemistry and biochemistry. Several authors consider the existence of chemomimetic processes, i.e. the occurrence of organo-chemical mechanisms in abiotic processes that anticipate the kind of chemistry to be used by the first biochemical transformations (Morowitz, 1992; Zubay, 1993; Meléndez-Hevia *et al.*, 1996; Costanzo *et al.*, 2007; Eschenmoser, 2007) leading to a true chemical continuity. Christian de Duve's protometabolic model is a useful theoretical framework to explore this continuity between the chemical world of a prebiotic phase, basically determined by the participation of non-genetically instructed catalysts (e.g. minerals, short oligomers), and the chemistry of a protometabolism using replicative polymers (e.g. RNA) as catalysts (de Duve, 2005). Since the very moment that replicative polymers appear on the scene, natural selection would automatically emerge and historical contingency would add to the purely chemical processes.

In de Duve's proposal there is an explicit role for thioesters as the first kind of molecule involved in energy transduction that would also serve as a bridge between sulphur and phosphate chemistries. In fact, the low solubility and reactivity of the main mineral source of phosphate (apatite) has obscured the prebiotic origin of the energy-transduction mechanisms based on phosphate-dependent chemical couplings for a long time, whose universality in extant cells is a strong indication of its ancestral character. Nevertheless, interesting candidates for an early bioenergetics based on phosphorous are pyrophosphate and polyphosphate generated in volcanic settings (Baltscheffsky and Baltscheffsky, 1994; Gedulin and Arrhenius, 1994) and meteoritic phosphonates (Pasek, 2008).

However, others prefer a radically different point of departure for early metabolisms. Instead of a heterotrophic scenario (where a chemically-rich environment is the cradle for simple protometabolic networks) an autotrophic origin (i.e. based on the energy and electrons released from redox reactions involving mineral compounds and the synthesis of organic matter from CO<sub>2</sub>) is postulated. Thus, the complete organochemical diversity originates from geochemical simplicity (e.g. CO<sub>2</sub> and electron-rich compounds such as H<sub>2</sub>S). Morowitz (1992) suggested that a primitive form of the reductive tricarboxylic-acid cycle would be the initial seed of a prebiotic chemistry that prefigured most of modern metabolism. Another model, with partial experimental support, postulates that the anaerobic synthesis of pyrite from H<sub>2</sub>S and FeS would serve as a source of electrons and energy

for chemical syntheses from CO<sub>2</sub> throughout the same autotrophic route (Wächtershäuser, 2006). Russell and Martin (2004) postulate a geochemical version of the metabolic synthesis of acetate from CO or CO<sub>2</sub> present in certain anaerobic microorganisms. In these cases, the submarine hydrothermal settings would offer the thermal and chemical gradients required by the models.

### The coevolution of metabolism and the environment

Living beings are chemical open systems that use chemicals from the environment and produce and release other compounds. Coevolution of microbial metabolisms and the elemental cycles implies the coupling and consistency of any evolutionary scheme with the geochemical record and the planetary chemistry. In this sense, chemical and isotopic studies in ancient rocks are fundamental in providing the best estimates for the minimal timing of any metabolic establishment. Nevertheless, the current studies show an apparent temporal uncoupling between the origin of metabolic innovations and their signals in the rock record (Nealson and Rye, 2005).

#### Box 18.1. The bioenergetic processes

Every life form conserves energy using basic and universal mechanisms. (A) Whether the external or primary energy source is visible light or chemical (inorganic or organic) reactions, electrons are channeled to soluble carriers or are associated to membranes. The soluble reduced carriers (e.g. NAD[P]H) are used for biosynthetic purposes, whereas carriers in membranes usually generate electrochemical potential gradients (either of protons or sodium ions), a type of useful energy invested in ATP synthesis, membrane transport, motility etc. Electrochemical potential gradients and ATP are enzymatically interchangeable energy pools. (B) Electron-transport chains are arranged around several common enzymatic complexes that recurrently use the same redox cofactors such as quinones (Q), cytochromes (e.g. *b*, *c*) or iron–sulphur proteins (FeS). In photoelectronic chains, the photochemical reaction centers based on bacteriochlorophyll (BChl) or chlorophyll (Chl) use the electromagnetic energy to drive the uphill electronic transport (gray lines). Peripheral electron-transport proteins are cytochrome *c* (cyt *c*) and plastocyanin (PC). Photosystem II (PSII) has an associated water-oxidative oxygen-releasing enzymatic complex. The archetypical oxygenic photosynthesis present in cyanobacteria (and plastids) could be the result of evolutionary tinkering on previously emerged anoxygenic photosyntheses (as in the red- and green-bacteria versions shown here). On the other hand, respiratory chains accept electrons through oxidative complexes such as NADH dehydrogenase (NDH) and donate electrons to external acceptors such as molecular oxygen (cytochrome oxidase, cyt *aa*<sub>3</sub>), nitrate (nitrate reductase, NAR), nitrite (NO<sub>2</sub> Rase), nitrous oxide (N<sub>2</sub>O Rase), or nitric oxide (NO Rase). The free-quinone pool is indicated as Q inside a dashed-line circle.



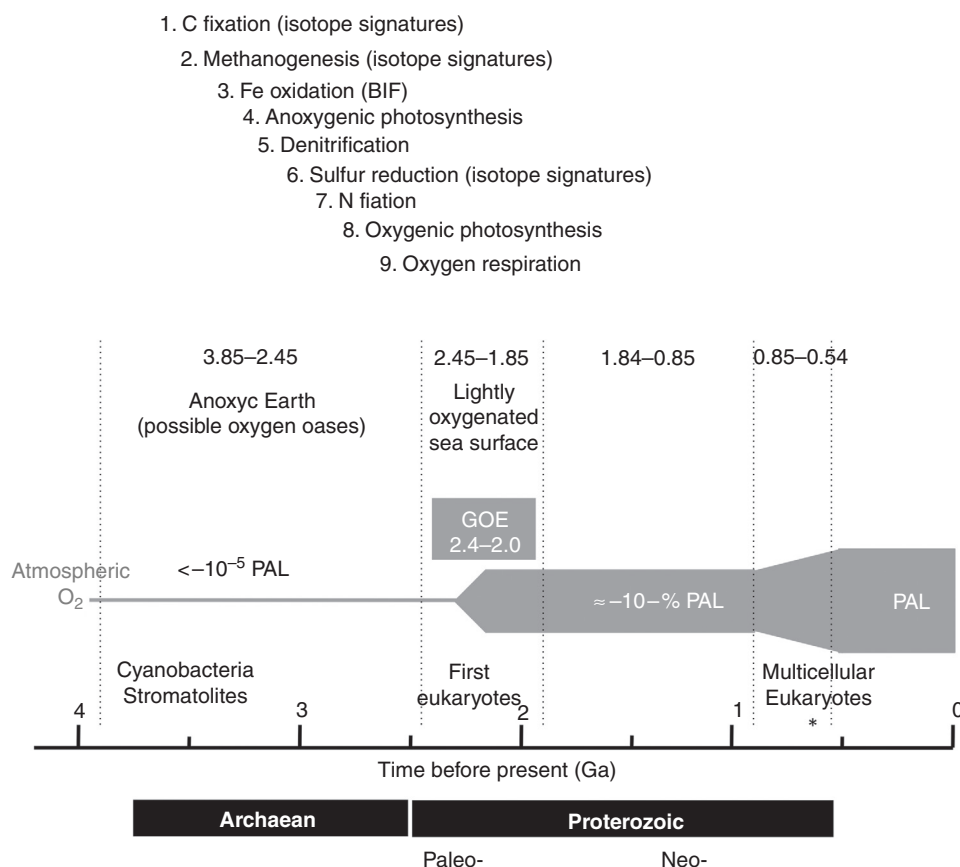


Figure 18.1. A hypothesis on the evolution of metabolisms and planetary history of oxygen. The accumulation of atmospheric oxygen along the geological time (in Ga =  $10^9$  years) in terms of the fraction of the present atmospheric level (PAL) is a consensus from several sources (Holland, 2006; Falkowski and Isozaki, 2008; Kump, 2008) based on geochemical data. Some paleontological landmarks (\*) and the great oxidation event (GOE) are indicated. The order of appearance of metabolisms is based on different, sometimes controversial, data from geochemistry, biochemistry and molecular phylogeny. BIF: banded-iron formations.

One of the most debated issues is the appearance and atmospheric accumulation of molecular oxygen and its metabolic and geological effects. About 2.4 Ga ago the atmospheric oxygen rose in what is called the great oxidation event (GOE), to a level reaching about 10 % of the present atmospheric level (PAL, Figure 18.1). Most authors agree that the invention of oxygenic photosynthesis (see Box 18.1) by the ancestors of cyanobacteria was the main source of oxygen, so GOE would set a minimum age for this metabolism. Some authors have postulated an earliest (more than 3.0 Ga ago) origin of cyanobacteria, and thus the lag period for the oxygen increase would be explained by diverse sink reactions, such as from oxygen consumption by methane of biological origin

or the oxidation of oceanic iron (with consequent deposition of the banded-iron formations, BIFs). Eventually, the marginalization of methanogenesis (e.g. by the scarcity of an essential metal such as Ni for the enzymes involved in the synthesis of methane; Saito, 2009) would contribute to the accumulation of oxygen. This extremely reactive molecule would also offer the opportunity for the evolution of new metabolic processes, such as oxygen respiration (see Box 18.1) or the synthesis of new metabolites, such as steroids (see below). Many authors also suggest that oxygen accumulation in the atmosphere was a requisite for a further increase in life complexity, i.e. the origin of biomineralization and multicellularity.

The mechanism of emergence of new metabolic processes would imply recycling and modifying of old molecular devices, e.g. oxygenic from anoxygenic photosyntheses or aerobic from anaerobic respirations. In both cases, the evolutionary schemes proposed remain controversial. At any rate, the crucial step in the origin of oxygenic photosynthesis was the addition of a catalyst able to oxidize water (associated to photosystem II, see Box 18.1) to an older light-dependent anoxygenic electron-transport chain (Allen and Martin, 2007). On the other hand, the enzymatic step necessary to reduce oxygen to water (cytochrome oxidase) would emerge from the tinkering of older reductases, most likely NO reductase (Ducluzeau *et al.*, 2009), responsible for an early denitrification process (Box 18.1).

### The evolution of metabolic pathways

The origin and evolution of metabolic pathways allowed primitive cells to become more chemically independent from the prebiotic sources of essential molecules. It is reasonable to assume that during the early stages of cell evolution, the primitive metabolism was based on a limited number of rudimentary (i.e. unspecific) enzymes. Several mechanisms observed in current cells (Box 18.2) may account for a rapid expansion of metabolic abilities. This ‘mode’ of evolution would justify a rapid ‘tempo’ in metabolic innovation, as observed in the spread of antibiotic resistances or the appearance of new metabolic activities within very short periods of time (i.e. from days to months) in experimental evolution with bacterial populations (Blount *et al.*, 2008).

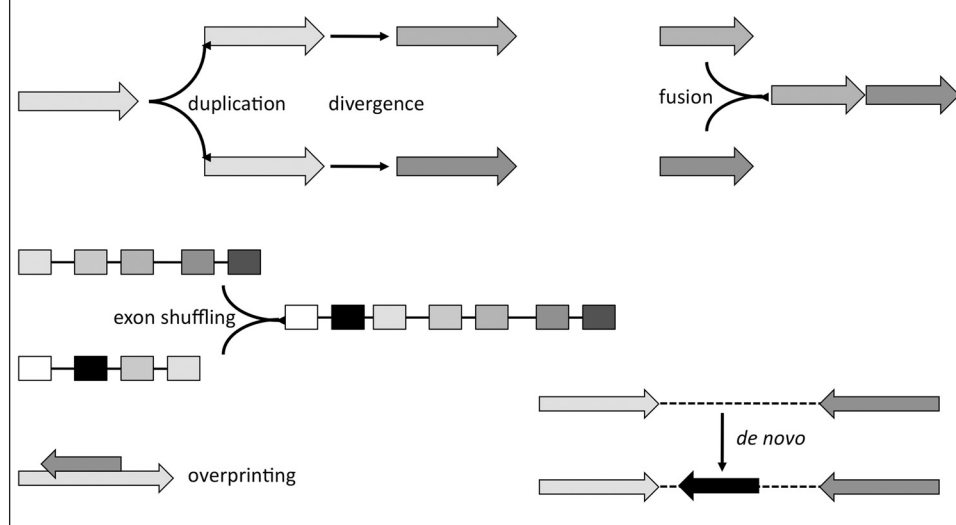
#### Box 18.2. How do new genes appear?

Let us suppose that we begin with a certain number of (not yet known) ‘starter-type’ genes (Lazcano and Miller, 1994). Several mechanisms have been discovered in extant cells that could explain a rapid expansion of the enzymatic repertoire.

1. **Duplication and divergence.** DNA duplication may involve from gene fragments (internal duplications) to whole genomes. Two duplicated genes may structurally diverge (paralogs) or may also fuse forming an elongated gene (Fani *et al.*, 2007). Regarding their functional fates,

gene divergence after duplication may originate new metabolic functions (James and Tawfik, 2003).

2. **Exon shuffling.** Most eukaryotic genes are fragmented, i.e. are organized as mosaics of coding sequences (exons) interrupted by non-coding segments (introns). Since many exons encode functional domains in proteins, new proteins may originate by the shuffling of exons (Long *et al.*, 2003).
3. **Overprinting and *de novo* from non-coding sequences.** Some mutations could give rise to new sequences for transcription and translation initiation and termination, either inside a pre-existing coding sequence (Delaye *et al.*, 2008) or within non-coding regions (Cai *et al.*, 2008). Wide arrows indicate coding regions; boxes, exons; lines, introns, and; dashed lines, intergenic regions.



### Models of metabolic-pathway evolution

Although we do not know how and when the central metabolic pathways originated, there are several models trying to explain their evolution (for reviews, see Peretó *et al.* 1997; Fani and Fondi, 2009). The classical hypotheses include (1) the retrograde model (Horowitz, 1945, 1965), (2) the forward model (Granick, 1957, 1965) and (3) the patchwork model (Ycas, 1974; Jensen, 1976). In earlier stages, some combination of enzymatic and non-enzymatic reaction steps could coexist (Lazcano and Miller, 1999).

- (1) The first attempt to explain the evolution of metabolic pathways was developed by Howoritz following Oparin's heterotrophic model and the one-gene-one-protein correspondence suggested by Beadle and Tatum (1941). Horowitz proposed that biosynthetic enzymes were acquired via gene duplication and in the reverse order as found in extant pathways, under the selective pressure of the exhaustion of prebiotic materials. Several criticisms of this model have been addressed elsewhere (Lazcano *et al.*, 1992; Peretó *et al.*, 1997) and, as a matter of fact, it can be applied only to very few cases.



- (2) The Granick model assumes that simpler metabolites are older than more complex compounds and, consequently, the enzymes involved in earlier steps in a pathway are older than the latter ones. Albeit originally proposed for explaining the evolution of heme and chlorophyll biosyntheses, this model fits better to the biosynthesis of several membrane components, including steroids (Peretó *et al.*, 1997). In general, the metabolic expansion due to the accumulation of oxygen in the atmosphere could be the result of the extension of an ancestral anaerobic metabolic core (Raymond and Segré, 2006).
- (3) According to the patchwork model, metabolic pathways may have been assembled by the recruitment of primitive enzymes that could react with a wide range of chemically related substrates. New enzymes with narrow specificities would result from gene duplication and divergence events. Several pieces of evidence give support to this model, including the analysis of whole genome sequences showing a high proportion of paralogous duplications corresponding to different enzymes evolved from common ancestral sequences and acting today in different pathways, i.e. with diverse substrate specificities, reaction mechanisms or both (Schmidt *et al.*, 2003). Accordingly, the distribution of protein domains across genomes supports the patchwork scenario (Caetano-Anollés *et al.*, 2009). Furthermore, the emergence of new activities during experimental evolution of bacterial populations also implies the recruitment of old enzymes to serve new functions (Mortlock, 1992).

Enzyme recruitment is a pervasive mechanism invoked to explain the evolution of metabolic pathways. For instance, the urea cycle of terrestrial animals likely evolved by the addition of arginase to the biosynthesis of arginine (Takiguchi *et al.*, 1989). Autotrophy could emerge by adding just one or two new activities to the previous heterotrophic pathways (Peretó *et al.*, 1999). The Krebs cycle might be the result of the assembly of enzymes involved in amino-acid metabolism (Meléndez-Hevia *et al.*, 1996). Patchwork assembly can also explain the evolution of biosynthetic pathways of amino acids, such as histidine (Fani *et al.*, 1995), lysine (Velasco *et al.*, 2002), tryptophan (Xie *et al.*, 2003), and lysine, arginine and leucine (Fondi *et al.*, 2007). The recruitment of stereospecific dehydrogenases from families of less efficient primitive enzymes has been invoked to explain the origin of homochiral lipid membranes (Peretó *et al.*, 2004). Furthermore, the patchwork model offers more explanatory power when studying pathways with chemically complex intermediates, such as in the biosynthesis of cofactors (Holliday *et al.*, 2007). Finally, directed-evolution experiments support this model, since the ability to use new carbon sources (Mortlock, 1992) or to catabolise toxic compounds (Copley, 2000) usually emerges via combinatorial pre-existent pathways, deregulation of transcriptional circuits and selection of enzyme variants with altered catalytic parameters.

Besides the mode of construction of pathways, another question of interest is the order of appearance of the different metabolic modules during evolution. To address this problem one must admit, in agreement with Morowitz (1992), that extant metabolic networks have traces of their history imprinted in their own structure. According to this author, a metabolic ancient core organized around a prebiotic reductive tricarboxylic-acid cycle followed by the addition of 'enzymatic shells', successively increases the complexity from central

pathways (such as in the Krebs cycle, glycolysis and fatty-acid metabolism) to amino-acid biosynthesis, sulphur metabolism and purine, pyrimidine and cofactor biosyntheses. A striking similar structure emerges from the studies of protein domain distribution, although in this case the nucleotide metabolism appears as an older module, suggesting a possible chemical link to an early RNA World (Caetano-Anollés *et al.*, 2009). Using a chemical and biochemical retrospective reasoning applied to the history of the appearance of metabolic pathways, Meléndez-Hevia and colleagues (2008) have proposed that early metabolism was directly connected to prebiotic chemistry and was built by radial growth from a central nucleus (glycolysis and an open or horseshoe Krebs cycle) and also by adding other peripheral subnetworks to it.

### *Enzymatic properties and evolvability*

The molecular tinkering associated with protein-function evolution has long been recognized (Jacob, 1977), one classic example being the use of some metabolic enzymes as lens proteins in animals. The conventional view of an extremely specific and proficient enzyme performing a well-defined and unique function must be substituted by the appreciation of several properties – such as ‘ambiguity’ and ‘plasticity’ – of crucial importance for the comprehension of enzyme evolvability, defined as ‘the ability of proteins to rapidly adopt (i.e. within a few sequence changes) new functions within existing folds or even adopt entirely new folds’ (Tokuriki and Tawfik, 2009).

Catalytic ambiguity refers to the ability of one enzyme to promote a secondary activity in the same active center responsible for the primary one. This is compatible with a view of proteins whereby one primary sequence can adopt a flexible structure with certain conformational diversity (James and Tawfik, 2003). Catalytic ambiguity also implies a certain plasticity or malleability of active sites, i.e. the alteration of activity throughout a small number of amino-acid substitutions. Jensen (1976) already emphasized the role of substrate ambiguity in enzyme evolution, since it opens up the possibility of recruiting new activities after duplication and divergence of the gene encoding for the enzyme (Figure 18.2). The evolution of new functions is possible by point mutations with large effects on the secondary activity without compromising the native structure and the primary activity, before the gene duplication, divergence and selection of the new protein (Aharoni *et al.*, 2005; Tokuriki and Tawfik, 2009).

Multifaceted proteins (or moonlighting) are a manifestation of the evolution of more than one function on the same protein. Thus, a protein moonlights when it has more than one functional role, e.g. an enzyme with additional structural or regulatory functions. Since amino-acid residues at the catalytic site usually represent a small fraction of the total enzymatic structure there are ample opportunities to evolve those additional functions.

Adaptive changes in the catalytic performance of enzymes have been reported regarding environmental conditions such as oxygen availability, osmotic relationships and

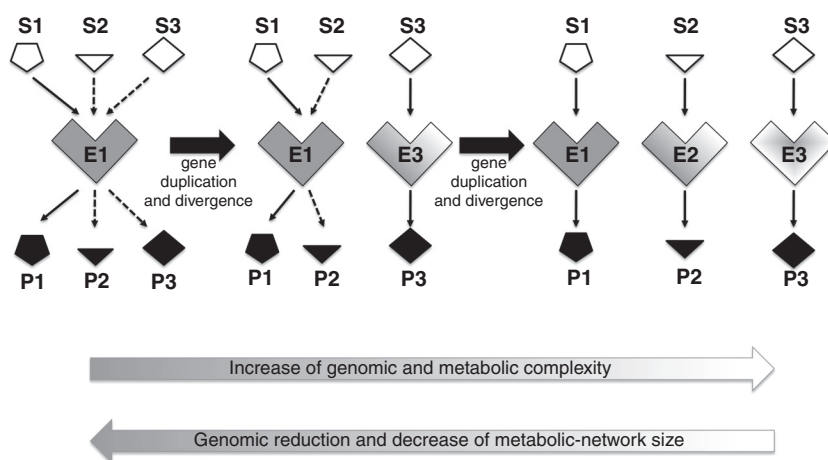


Figure 18.2. Increase of metabolic complexity throughout gene duplication and divergence. The enzyme E1 shows substrate ambiguity, i.e. it acts on three different substrates S1 to S3 – open pentagon, triangle and rhomboid, being the transformation of S1 in P1 the main or primary activity and the activities on S2 and S3 (dashed arrows) the secondary ones. A first duplication event and divergence specializes E3 in the use of S3 whereas E1 maintains the activity on S1 and, with less extension, on S2. A second duplication event generates three specialized enzymes fully active on each of the three substrates. On the other hand, the reverse phenomenon might occur during genome and metabolic minimization in endosymbionts (Zientz *et al.*, 2004) and pathogens (Yus *et al.*, 2009).

temperature regimes (Hochachka and Somero, 2002). Several authors have also addressed the question of optimality of the enzymatic function, either at the level of the catalytic constants of individual enzymes (Cornish-Bowden, 1976; Heinrich *et al.*, 2002), the architecture of the metabolic pathways (Meléndez-Hevia and Torres, 1988; Heinrich *et al.*, 1999; Cornish-Bowden, 2004) or the fitness of a final metabolic product such as cholesterol (Bloch, 1994) and glycogen (Meléndez *et al.*, 1997) as well-adapted structures able to perform their physiological function.

### Examples of mosaic evolution in eukaryotic metabolism

Protein sequence analysis, molecular cladistics and comparative genomics have enhanced our understanding of metabolic evolution. Their applicability cannot be extended far beyond the universal ancestor, although important insights on early metabolism can be achieved through comparative biochemistry over the broadest representation of genomes. Nevertheless, such backtrack studies are hindered by polyphyletic secondary losses, lateral gene transfers, replacements and redundancies. These difficulties are also patent in the study of less ancient stages of metabolic evolution, e.g. during eukaryotic evolution since about 2 Ga before present. Molecular cladistics and comparative genomics show a mosaic evolution of metabolic pathways during the merging of genomes from different origins giving rise to the diversity of eukaryotic compartments. Furthermore, ongoing endosymbioses

(see Chapter 21) are also the scenario for evolutionary opportunism and molecular tinkering in metabolic innovation. The biochemical and genetic causes of the metabolic mosaicism observed in eukaryotic cells are still poorly understood (Ringemann *et al.*, 2006).

### ***Mosaic pathways in eukaryotic compartments***

The bacterial origin of mitochondria and plastids is well established, although some controversies remain (de Duve, 2007). The symbiogenetic acquisition of genomes implies that, initially, many pathways could be redundant between the endosymbiont and the host. The comparative-genomics' and proteomics' studies by Gabaldón and Huynen (2003, 2004) have reconstructed the ancestral metabolism of the protomitochondrion: a quite complex bacterium whose original metabolism has been almost completely lost or substituted by new functions.

Molecular cladistic analyses on individual enzymes also show that the enzymes involved in a pathway have different phylogenetic origins. For instance, the Calvin cycle, an idiosyncratic pathway from the cyanobacterial ancestor of modern plastids, is catalyzed in plants by a mixture of enzymes with different origins, including mitochondrial ones (Martin and Schnarrenberger, 1997). Also, the redundancy of isoenzymes in the cell compartments has been solved differently in the different eukaryotic lineages, as shown by the case of the universally distributed anabolic and catabolic thiolases (Peretó *et al.*, 2005).

### ***Opportunistic metabolic evolution in endosymbionts***

Ongoing endosymbioses are excellent case studies for the evolution of metabolic pathways. The reduction of genomes and metabolic networks observed during the endosymbiotic event (Moya *et al.*, 2008) preferentially affects the redundant pathways with the host, e.g. the biosynthesis of fatty acids and membrane lipids. Meanwhile, the biosynthetic pathways for essential metabolites (e.g. amino acids, cofactors) are kept by the endosymbiont and result in a mutual metabolic interdependence with the host. However, sometimes the reduction in the repertoire of enzymatic activities favours a return to a situation similar to an ancestral state, i.e. fewer enzymes have to do the work and recurrence to substrate ambiguity could be a solution (Figure 18.2). In fact, some generic enzymes, like transaminases and kinases, have been reduced in number and most likely the substrate specificity has been relaxed (Zientz *et al.*, 2004).

A remarkable aspect of symbiogenetic innovation in metabolism is syntrophy or metabolic complementation. The biochemical integration of hosts and symbionts can be extremely tight since some metabolic pathways are shared and the participating enzymes are encoded by different genomes. In other words, fragments of a metabolic pathway are spatially separated into different compartments and organisms. One striking case is the heme biosynthesis by *Bradyrhizobium japonicum* from a metabolic precursor ( $\delta$ -aminolevulinic acid) supplied by the host plastid (soybean) (Figure 18.3A).

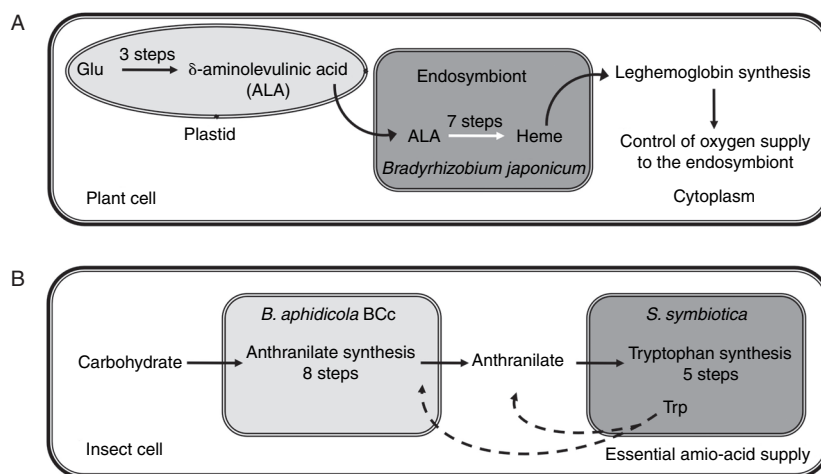


Figure 18.3. Metabolic complementation in endosymbionts. Some metabolic pathways are fragmented between hosts and endosymbionts. (A) The synthesis of heme in soybean is shared between plastid and the endosymbiont *Bradyrhizobium japonicum*. The plant leghemoglobin synthesis uses the bacterial heme. Leghemoglobin controls oxygen availability to the bacterium. (B) The synthesis of Trp in the cedar aphid is distributed between the two bacterial endosymbionts: *Buchnera aphidicola* supplies the anthranilate necessary for Trp biosynthesis by *Serratia symbiotica*. Trp is an essential amino acid for both the host and the endosymbiont *B. aphidicola*, whereas *S. symbiotica* could be considered an auxotrophic mutant for anthranilate.

Eventually the plant cell uses the heme group for leghemoglobin synthesis (Sangwan and O'Brian, 1991).

Another excellent example is provided by the biosynthesis of the essential amino acid Trp in cedar aphids. The common ancestor of animals lacked the complete pathway for Trp biosynthesis. Some 200 million years ago the association of the aphid ancestor with  $\gamma$ -proteobacteria overcame the need for an external supply of Trp, allowing the adaptation of the insect host to new ecological niches and diets devoid of Trp. The association with a second symbiont solved the redundancies in such a way that today there is a need for an interspecific anthranilate supply between the two endosymbionts (Figure 18.3B). *Buchnera aphidicola* contains the biosynthetic pathway up to anthranilate, whereas *Serratia symbiotica* uses this anthranilate to synthesize the Trp needed by all three partners (Gosalbes *et al.*, 2008). Thus, aphids synthesize Trp by using bacterial pathways fully integrated in a vertical inheritance.

### Concluding remarks

We still do not know when and how life originated. However, useful hints can be inferred from extant metabolic pathways, as well as from their correlation with environmental changes through planetary history. Although we still lack a narrative for the origin and evolution of metabolic pathways – a true natural history of biochemistry – we

are gaining insights from comparative genomics and molecular cladistic analyses of individual enzymes. The fruitfulness of this approach is indebted to the vision of evolution as a tinkerer rather than an engineer. Using the words of Jacob (1977), inspired by Darwin (1862): ‘In contrast to the engineer, evolution does not produce innovations from scratch. It works on what already exists ... like a tinkerer who, during millions of years, has slowly modified his products ... using all opportunities to transform and create.’

### References

- Aharoni, A., Gaidukow, L., Khersonsky, O, Gould, S. McQ., Roodveldt, C. and Tawfik, D. S. (2005). The ‘evolvability’ of promiscuous protein functions. *Nature Genetics*, **37**, 73–6.
- Allen, J. F. and Martin, W. (2007). Out of thin air. *Nature*, **445**, 610–12.
- Baltscheffsky, H. and Baltscheffsky, M. (1994). Molecular origin and evolution of early biological energy conversion. In *Early life on Earth*, ed. S. Bengston. New York: Columbia University Press, Nobel symposium no. 84, pp. 81–90.
- Barber, J. (2008). The photosynthetic generation of oxygen. *Philosophical Transactions of the Royal Society B*, **363**, 2665–74.
- Beadle, G.W. and Tatum, E.L. (1941). Genetic control of biochemical reactions in *Neurospora*. *Proceedings of the National Academy of Sciences USA*, **27**, 499–506.
- Bloch, K. (1994). Blondes in Venetian Paintings: The Nine-banded Armadillo, and Other Essays in Biochemistry. New Haven: Yale University Press.
- Blount, Z. D., Borland, C. Z. and Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, **105**, 7899–906.
- Caetano-Anollés, G., Yafremava, L. S., Gee, H., Caetano-Anollés, D., Kim, H. S. and Mittenthal, J. E. (2009). The origin and evolution of modern metabolism. *The International Journal of Biochemistry and Cell Biology*, **41**, 285–97.
- Cai, J., Zhao, R., Jiang, H. and Wang, W. (2008). *De novo* origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics*, **179**, 487–96.
- Copley, S. D. (2000). Evolution of a metabolic pathway for degradation of a toxic xenobiotic: the patchwork approach. *Trends in Biochemical Sciences*, **25**, 261–5.
- Cornish-Bowden, A. (1976). The effect of natural selection on enzymic catalysis. *Journal of Molecular Biology*, **101**, 1–9.
- Cornish-Bowden, A. (2004). *The Pursuit of Perfection: Aspects of Biochemical Evolution*. Oxford: Oxford University Press.
- Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C. and Soto-Andrade, J. (2007). Beyond reductionism: metabolic circularity as a guiding vision for a real biology of systems. *Proteomics*, **7**, 839–45.
- Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C., Soto-Andrade, J. and Abarzúa, F. G. (2004). Understanding the parts in terms of the whole. *Biology of the Cell*, **96**, 713–17.
- Costanzo, G., Saladino, R., Crestini, C., Ciciriello, F. and Di Mauro, E. (2007). Formamide as the main building block in the origin of nucleic acids. *BMC Evolutionary Biology*, **7**, S1.
- Darwin, C. R. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray, ch. 6.



- Darwin, C. R. (1862). On the Various Contrivances by which British and Foreign Orchids Are Fertilized by Insects, and on the Good Effects of Intercrossing. London: John Murray, ch. 9.
- de Duve, C. (2005). *Singularities: Landmarks on the Pathways of Life*. Cambridge: Cambridge University Press.
- de Duve, C. (2007). The origin of eukaryotes: a reappraisal. *Nature Reviews Genetics*, **8**, 395–403.
- Delaye, L., Deluna, A., Lazcano, A. and Becerra, A. (2008). The origin of a novel gene through overprinting in *Escherichia coli*. *BMC Evolutionary Biology*, **8**, 31.
- Ducluzeau, A.-L., van Lis, R., Duval, S., Schoepp-Cothenet, B., Russell, M. J. and Nitschke, W. (2009). Was nitric oxide the first deep electron sink? *Trends in Biochemical Sciences*, **34**, 9–15.
- Eschenmoser, A. (2007). The search for the chemistry of life's origin. *Tetrahedron*, **63**, 12821–44.
- Falkowski, P. G. and Isozaki, Y. (2008). The story of O<sub>2</sub>. *Science*, **322**, 540–2.
- Fani, R., Liò, P. and Lazcano, A. (1995). Molecular evolution of the histidine biosynthetic pathway. *Journal of Molecular Evolution*, **41**, 760–74.
- Fani, R., Brilli, M., Fondi, M. and Liò, P. (2007). The role of gene fusions in the evolution of metabolic pathways: the histidine biosynthesis case. *BMC Evolutionary Biology*, **7**, S4.
- Fani, R. and Fondi, M. (2009). Origin and evolution of metabolic pathways. *Physics of Life Reviews*, **6**, 23–52.
- Fell, D. and Wagner, A. (2000). The small world of metabolism. *Nature Biotechnology*, **18**, 1121–2.
- Fondi, M., Brilli, M., Emiliani, G., Paffetti, D. and Fani, R. (2007). The primordial metabolism: an ancestral interconnection between leucine, arginine and lysine biosynthesis. *BMC Evolutionary Biology*, **16**, S3.
- Gabaldón, T. and Huynen, M. A. (2003). Reconstruction of the proto-mitochondrial metabolism. *Science*, **301**, 609.
- Gabaldón, T. and Huynen, M. A. (2004). Shaping the mitochondrial proteome. *Biochimica et Biophysica Acta*, **1659**, 212–20.
- Gedulin, B. and Arrhenius, G. (1994). Sources and geochemical evolution of RNA precursor molecules: the role of phosphate. In *Early life on Earth*, ed. S. Bengtson. New York: Columbia University Press, Nobel symposium no. 84, pp. 91–106.
- Gosalbes, M. J., Lamelas, A., Moya, A. and Latorre, A. (2008). The striking case of tryptophan provision in the cedar aphid *Cinara cedri*. *Journal of Bacteriology*, **190**, 6026–9.
- Granick, S. (1957). Speculations on the origins and evolution of photosynthesis. *Annals of the New York Academy of Sciences*, **69**, 292–308.
- Granick, S. (1965). The evolution of heme and chlorophyll. In *Evolving Genes and Proteins*, eds. V. Bryson and H. J. Vogel. New York: Academic Press, pp. 67–88.
- Haldane, J. B. S. (1929). The origin of life. *The Rationalist Annual*, 3–10.
- Heinrich, R., Meléndez-Hevia, E., Montero, F., Nuño, J. C., Stephani, A., Waddell, T. G. (1999). The structural design of glycolysis: an evolutionary approach. *Biochemical Society Transactions*, **27**, 294–8.
- Heinrich, R., Meléndez-Hevia, E. and Cabezas, H. (2002). Optimization of kinetic parameters of enzymes. *Biochemistry and Molecular Biology Education*, **30**, 184–8.

- Hochachka, P. W. and Somero, G. N. (2002). *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. Oxford: Oxford University Press.
- Holland, H. D. (2006). The oxygenation of the atmosphere and oceans. *Philosophical Transactions of the Royal Society B*, **361**, 903–15.
- Holliday, G. L., Thornton, J. M., Marquet, A., Smith, A. G., Rébeillé, F., Mendel, R., Schubert, H. L., Lawrence, A. D. and Warren, M. J. (2007). Evolution of enzymes and pathways for the biosynthesis of cofactors. *Natural Product Reports*, **24**, 972–87.
- Horowitz, N. H. (1945). On the evolution of metabolic pathways. *Proceedings of the National Academy of Sciences USA*, **31**, 153–7.
- Horowitz, N. H. (1965). The evolution of biochemical syntheses: retrospect and prospect. In *Evolving Genes and Proteins*. New York: Academic Press, pp. 15–23.
- Jacob, F. (1977). Evolution and tinkering. *Science*, **196**, 1161–6.
- James, L. C. and Tawfik, D. S. (2003). Conformational diversity and protein evolution: a 60-year-old hypothesis revisited. *Trends in Biochemical Sciences*, **28**, 361–8.
- Jensen, R. A. (1976). Enzyme recruitment in evolution of new function. *Annual Review of Microbiology*, **30**, 409–25.
- Kump, L. R. (2008). The rise of atmospheric oxygen. *Nature*, **451**, 277–8.
- Lane, N. (2009). *Life Ascending: The Ten Great Inventions of Evolution*. London: Profile Books.
- Lazcano, A., Fox, G. and Oró, J. (1992). Life before DNA: the origin and evolution of early Archaean cells. In *The Evolution of Metabolic Function*, ed. R. P. Mortlock. Boca Raton: CRC Press, pp. 237–95.
- Lazcano, A. and Miller, S. L. (1994). How long did it take for life to begin and evolve to cyanobacteria? *Journal of Molecular Evolution*, **39**, 546–54.
- Lazcano, A. and Miller, S. L. (1999). On the origin of metabolic pathways. *Journal of Molecular Evolution*, **49**, 424–31.
- Long, M., Deutsch, M., Wang, W., Betrán, E., Brunet, F. G. and Zhang, J. (2003). Origin of new genes: evidence from experimental and computational analyses. *Genetica*, **118**, 171–82.
- Martin, W. and Schnarrenberger, C. (1997). The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Current Genetics*, **132**, 1–18.
- Meléndez-Hevia, E. and Torres, N. V. (1988). Economy of design in metabolic pathways: further remarks on the game of the pentose phosphate cycle. *Journal of Theoretical Biology*, **132**, 97–111.
- Meléndez, R., Meléndez-Hevia, E. and Cascante, M. (1997). How did glycogen structure evolve to satisfy the requirement for rapid mobilization of glucose? A problem of physical constraints in structure building. *Journal of Molecular Evolution*, **45**, 446–55.
- Meléndez-Hevia, E., Montero-Gómez, N. and Montero, F. (2008). From prebiotic chemistry to cellular metabolism. The chemical evolution of metabolism before Darwinian natural selection. *Journal of Theoretical Biology*, **252**, 505–19.
- Meléndez-Hevia, E., Waddell, T. G. and Cascante, M. (1996). The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *Journal of Molecular Evolution*, **43**, 293–303.
- Morowitz, H. J. (1992). *Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis*. New Haven: Yale University Press.



- Mortlock, R. P. (ed.). (1992). *The Evolution of Metabolic Function*. Boca Raton: CRC Press.
- Moya, A., Peretó, J., Gil, R. and Latorre, A. (2008). Learning how to live together: genomic insights into prokaryote–animal symbioses. *Nature Reviews Genetics*, **9**, 218–29.
- Nealson, K. H. and Rye, R. (2005). Evolution of metabolism. In *Biogeochemistry*, ed. W. H. Schlesinger. Amsterdam: Elsevier, vol. 8, pp. 41–61.
- Oparin, A. I. (1924). *The Origin of Life*. Appendix to Bernal, J. D. (1967) *The Origin of Life*. London: Weidenfeld and Nicholson.
- Pasek, M. A. (2008). Rethinking early Earth phosphorus geochemistry. *Proceedings of the National Academy of Sciences USA*, **105**, 853–8.
- Peretó, J., López-García, P. and Moreira, D. (2004). Ancestral lipid biosynthesis and early membrane evolution. *Trends in Biochemical Sciences*, **29**, 469–77.
- Peretó, J., López-García, P. and Moreira, D. (2005). Phylogenetic analysis of eukaryotic thiolases suggests multiple proteobacterial origins. *Journal of Molecular Evolution*, **61**, 65–74.
- Peretó, J., Fani, R., Leguina, J. I. and Lazcano, A. (1997). Enzyme evolution and the development of metabolic pathways. In *New Beer in an Old Bottle: Eduard Buchner and the Growth of Biochemical Knowledge*, ed. A. Cornish-Bowden. València: Publicacions de la Universitat de València, pp. 173–98.
- Peretó, J., Velasco, A. M., Becerra, A. and Lazcano, A. (1999). Comparative biochemistry of CO<sub>2</sub> fixation and the evolution of autotrophy. *International Microbiology*, **2**, 3–10.
- Raymond, J. and Segré, D. (2006). The effect of oxygen on biochemical networks and the evolution of complex life. *Science*, **311**, 1764–7.
- Ringemann, C., Ebenhöf, O., Heinrich, R. and Ginsburg, H. (2006) Can biochemical properties serve as selective pressure for gene selection during inter-species and endosymbiotic lateral gene transfer? *IEE Proceedings-Systems Biology*, **153**, 212–22.
- Russell, M. J. and Martin, W. (2004). The rocky roots of the acetyl-CoA pathway. *Trends in Biochemical Sciences*, **29**, 358–63.
- Saito, M. A. (2009). Less nickel for more oxygen. *Nature*, **458**, 714–5.
- Sangwan, I. and O'Brian, M. R. (1991). Evidence for an inter-organismic heme biosynthetic pathway in symbiotic soybean root nodules. *Science*, **251**, 1220–2.
- Schmidt, S., Sunyaev, S., Bork, P. and Dandekar, T. (2003). Metabolites: a helping hand for pathway evolution? *Trends in Biochemical Sciences*, **28**, 336–41.
- Takiguchi, M., Matsubasa, T., Amaya, Y. and Mori, M. (1989). Evolutionary aspects of urea cycle enzyme genes. *Bioessays*, **10**, 163–6.
- Tokuriki, N. and Tawfik, D. S. (2009). Protein dynamism and evolvability. *Science*, **324**, 203–7.
- Velasco, A. M., Leguina, J. I. and Lazcano, A. (2002). Molecular evolution of the lysine biosynthetic pathways. *Journal of Molecular Evolution*, **55**, 445–9.
- Wächtershäuser, G. (2006). From volcanic origins of chemoautotrophic life to bacteria, Archaea and eukarya. *Philosophical Transactions of the Royal Society B*, **361**, 1787–1808.
- Wagner, A. (2007). *Robustness and Evolvability in Living Systems*. Princeton: Princeton University Press.
- Xie, G., Keyhani, N. O., Bonner, C. A. and Jensen, R. A. (2003). Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiology and Molecular Biology Reviews*, **67**, 303–42.

- Ycas, M. (1974). On earlier states of the biochemical system. *Journal of Theoretical Biology*, **44**, 145–60.
- Yus, E., Maier, T., Michalodimitrakis, K., van Noort, V., Yamada, T., Chen, W. H., Wodke, J. A., Güell, M., Martínez, S., Bourgeois, R., Kühner, S., Raineri, E., Letunic, I., Kalinina, O. V., Rode, M., Herrmann, R., Gutiérrez-Gallego, R., Russell, R. B., Gavin, A. C., Bork, P. and Serrano, L. (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science*, **326**, 1263–8.
- Zientz, E., Dandekar, T. and Gross, R. (2004). Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiology and Molecular Biology Reviews*, **68**, 745–70.
- Zubay, G. (1993). To what extent do biochemical pathways mimic prebiotic pathways? *Chemtracts Biochemistry and Molecular Biology*, **4**, 317–23.

