

2

Some Aspects of Nutritional Biochemistry

The Central Carbon Pathway

Why is Glucose the Central Fuel Molecule?

Nature has invented several ways to synthesize ATP. One solution is a series of coupled chemical reactions catalyzed by soluble enzymes in the cytoplasm that result in substrate-level phosphorylation. The most prevalent starting substrate for such an energy-yielding pathway is glucose, which is decomposed in all higher cells by a carefully orchestrated and evolutionarily fixed way called glycolysis.

There is apparently something special about glucose since it is so widely used by organisms of all kind. In our diet, glucose comes in fruit juice, as starch and glycogen (the polymeric storage form of glucose in plants and animals, respectively), and in the disaccharides saccharose (table sugar) and lactose. For herbivores glucose comes also as cellulose, a glucose polymer from plant cell walls. The latter is arguably the quantitatively most abundant biological molecule on earth. Hemicellulose, which consists mostly of xylose and arabinose, represents 15–30% of plant material and is thus of great importance to herbivores. Even if they are part of our food, we cannot deal with cellulose and hemicellulose. Only few other hexoses are found in our diet. These are galactose (from lactose), mannose (mainly from glycoproteins), fucose (in milk oligosaccharides), and fructose (in fruit juice, saccharose). In the series of pentoses, ribose and deoxyribose dominate since they constitute the backbone of all nucleic acids that we eat with our food.

An organism that does not know to handle glucose deprives itself of the most important organic carbon source on our planet. The prominent role of glucose is also reflected by the fact that nearly all other ingested sugars are transformed in the body into glucose before they can be used for energy production (fructose is an exception). Conversely, glucose is the starting material for all monosaccharide synthesis in our body, and also the carbon skeleton of glucose provides the starting material for the synthesis of amino acids and lipids. *In fact, the first cells were probably more concerned to make glucose than to degrade it. Although*

it looks different today, the initial function of glycolysis was gluconeogenesis. We get into the realm of evolutionary speculation when we ask why glucose became the universal cellular carbon currency, but a few arguments can be given. The following argument for glucose is of course circular but nevertheless true: Glucose is used by so many organisms because it is so prevalent in the organic world. Any newly evolving biological system finds itself in the tradition of life that existed before and with it. Even historical accidents thus become a biochemical necessity. However, there are also a few chemical arguments for the prominence of glucose in biology. Take glucose in the chair conformation. Glucose has the smallest axial substituents imaginable for hexoses—hydrogen atoms. This gives it an energetically privileged position within hexoses. This stable chemical structure of glucose makes it attractive for metabolism; compared with other hexoses, it has a lower tendency to react nonenzymatically with proteins. That this concern is not a mute point is demonstrated by the problems of diabetic patients, who do not manage to keep their blood glucose levels at a carefully regulated level. Excess blood glucose results in the glucosylation of hemoglobin and of proteins from the vascular tissue, which is at the basis of the medical problems in diabetic patients. Finally, glucose can be formed from formaldehyde under simulated prebiotic conditions. *Glucose was thus simply around when Nature started to tinker with primitive metabolism.*

Glycolysis

Glycolysis is the central energy-providing process in an astonishing diversity of organisms. In many organisms, it is also the sole source of energy. This also applies to several tissues of our own body. The Greek word “glycolysis” can be translated as “sweet-splitting,” and this is a very precise description of its chemistry. Glycolysis is the controlled degradation of the six-carbon sugar glucose into the central intermediate of metabolism, pyruvate. The conversion is linked to the gain of two molecules of ATP and two of NADH, but glycolysis is more than just energy gaining. Recall from your biochemistry courses that many intermediates of glucose degradation are also the starting points for several biosynthetic pathways.

Its Origin

The design of the central carbon pathway thus evolved under dual constraints. The conservation of this pathway in so many organisms could suggest that there is only one possible chemical solution to the central metabolic pathway. Alternatively, it could reflect a “frozen accident” of evolution. In this second scenario, the pathway developed over a long time period, where it probably showed many variations, but when the chemical reactions fitted neatly together, it was fixed by selective pressure. No organism could tinker with alternative pathways when living on carbohydrates. They could only refine its regulation to their peculiar needs (what most organisms actually did), but they were not

allowed to redesign its chemical path because of selective constraints. As all living organisms are evolutionarily linked (as expressed in the concept of the universal phylogenetic tree), protoglycolysis was then inherited from the ancestor cell, which first successfully fixed this invention, by all its descendents. It is probably safer to restrict this argument to eukaryotes. Many, if not most, microbes use the Entner–Doudoroff pathway for “glycolysis” and the glycolytic pathway for gluconeogenesis. The very fact that glycolysis is found in all kingdoms of life (archaea, eubacteria, and eukaryotes) speaks in favor of its antiquity. We should therefore treat this reaction sequence with much respect as we might look through it deep into the biological past. Fitting with the above arguments, the chemical intermediates are exactly the same over all organisms; only the cofactors and enzymes show variations. For example, some bacteria, protists and perhaps all plants have a phosphofructokinase that uses pyrophosphate instead of ATP in glycolysis. The very fact that the Entner–Doudoroff pathway replaces the Embden–Meyerhof pathway (glycolysis) in so many microbes (Fuhrer et al. 2005) is a frequently used argument for its even older origin.

The Reactions

So what happens in glycolysis? Ten enzymes transform glucose into pyruvate. The six-carbon sugar glucose is symmetrically split at about halfway of the reaction pathway into two three-carbon compounds. This splitting reaction separates the preparation phase of glycolysis from the energy-yielding phase. The motto of this first part is again the old dictum “there is no free meal in biology.” Before you can extract the chemical energy stored in the glucose molecule, you have to invest energy in the form of two phosphorylation steps. There is chemical reason in these two steps, but the motivation differs for the two steps. If a cell wants to keep glucose for its own use, it has to mark it as its own. The cell labels it with a tag that prevents the flow of glucose across the plasma membrane following the concentration gradient of glucose. Cells found a very simple, although costly measure: phosphoryl transfer to glucose by hexokinase. This adds chemical charges to the otherwise electrically neutral glucose molecule. Glucose needs a transport protein to get across the membrane. The argument is different for the next ATP energy consumer step in glycolysis. The enzyme phosphofructokinase adds a second phosphate group at the opposing end of the six-carbon sugar. The electrostatic repulsion of the two negative charges puts the ring structure under strain. This strain is exploited in the following reaction where aldolase catalyzes the splitting of the C6 sugar into two C3 sugars.

With glyceraldehyde 3-phosphate (G3P) starts the second phase of glycolysis, which you can call the payoff phase. Chemically something has already happened. The entropy has increased: One molecule became two molecules, but the glyceraldehyde is still at the approximately same oxidation level as the starting compound glucose. However, if you look at the end product of the glycolytic pathway, pyruvate, you see a definitively more oxidized compound. Pyruvate shows a carboxylate and a keto-group. At these positions, G3P has an

aldehyde and a hydroxyl group. Two separate oxidation steps (only one uses NADH) have thus occurred, and the cells have learned to harness the energy released by each oxidation step in a chemically usable form. A crucial tool of the enzyme G3P dehydrogenase is a critical cysteine residue. Its thiol group establishes a covalent linkage to the aldehyde carbon of G3P. An adjacent histidine residue helps here as a base catalyst: It accepts the hydrogen from the thiol and stabilizes the resolution of the double bond between the C and O atoms in G3P, which allows the covalent binding of the substrate to the enzyme. Now comes the enzyme cofactor NAD^+ into play. It abstracts a hydride (a hydrogen atom that took away the two binding electrons) from the substrate. To fill the hole in the electronic shell of the terminal carbon of G3P, an intramolecular electron transfer has to occur, which reestablishes the $\text{C}=\text{O}$ double bond. You have now a thioester intermediate. This thioester bond has a very high standard free energy of hydrolysis, and the enzyme keeps this high-energy bond for energy fixation. It achieves this goal by excluding water from the reactive center, and it allows only a phosphate group to detach the substrate from the enzyme surface. Instead of hydrolysis a phosphorolysis occurs and creates 1,3-bisphosphoglycerate. This is an anhydride between two acids, namely the carboxylate and the phosphoric acid. Its standard energy of hydrolysis is -49.3 kJ/mol , much higher than that of ATP hydrolysis (-30.5 kJ/mol). The next enzyme in the glycolytic pathway then does the logical step: It transfers the phosphate group from 1,3-bisphosphoglycerate to ADP creating ATP (although it is not sure whether it is physically the same P_i).

Energetics

Now let's look at the stoichiometry of glycolysis and the energy sheet: $\text{glucose} + 2\text{NAD}^+ \rightarrow 2 \text{ pyruvate} + 2\text{NADH} + 2\text{H}^+$; $\Delta G'^{\circ} = -146 \text{ kJ/mol}$. This free-energy change is conserved in two ATP molecules, which need 61 kJ/mol for their synthesis. Under the actual concentrations of the reactants in the cell, the efficiency of the process is more than 60%. However, if you stop the reaction here, only a small percentage of the total chemical energy contained in glucose is recovered (recall the complete burning of glucose with oxygen to H_2O and CO_2 yields $-2,840 \text{ kJ/mol}$). However, not all cells have this option because either they do not have access to oxygen or they lack a respiratory chain. Furthermore, early in the history of life, molecular oxygen was not present in the atmosphere precluding glucose degradation using oxygen as electron acceptor.

Variations on a Theme

Johann Sebastian Bach has written variations on a single musical theme that should especially appeal to scientists by their near mathematical logic. Mother Nature is an equally creative composer and has tried a lot of variations around the theme of glycolysis. Let's first take a small variation where only a few side notes are altered.

Pyrococcus

The first theme is played by an extreme character: *Pyrococcus abyssi*. This is quite an exotic prokaryote, what biologists actually call an extremophile. It belongs to Archaea, the third kingdom of life. However, it is not just its phylogenetic affinity that makes this organism interesting. It is a hyperthermophile; its optimal growth temperature is 96 °C, the minimum and maximum temperatures of growth are 67 and 102 °C, respectively. You might raise your eyebrows since this is above the boiling point of water. This is, however, not an issue for *P. abyssi*. As the name suggests, (“the fire globule of the deepest depth”) its habitat is a hot area of the seafloor. This location solves the paradox: The hydrostatic pressure at the seafloor exceeds 600 atm, and at this pressure, 102 °C hot water cannot boil. Pyrococci are motile by flagella and reduce elemental sulfur to sulfide under strictly anaerobic conditions. All this would sound to our forefathers as hell and like a proof for the most dreadful beliefs of the Dark Ages. Yet, some like it hot, and life might actually have started under such conditions. Even under hellish conditions, you can make your living comfortable if you are adapted to it. In pyrococci all metabolites of glycolysis are identical to those of humans; the differences concern only the cofactors (Sapra et al. 2003). *Pyrococcus* uses ADP as phosphoryl donor in the first two phosphorylation steps of glycolysis, and instead of NAD⁺ it uses ferredoxin and tungsten. Pyruvate is decarboxylated to acetyl-CoA (here again pyrococci use a ferredoxin oxidoreductase). The free energy of the CO₂ release reaction is stored in the thioester bond to coenzyme A and is used for the synthesis of a further ATP with the concomitant release of acetate. The wide phylogenetic and ecological distribution and conservation of glycolysis speaks for one of the oldest sugar degradation pathways invented in biological systems on earth.

Entner–Doudoroff Pathway

Now comes another variation on the glycolysis scheme, but this time, new elements are introduced. This second variation is called the Entner–Doudoroff pathway, but remember that many microbiologists believe that this pathway preceded the glycolytic pathway in evolution. This series of reactions starts quite similar to the entrance reaction of glucose phosphorylation by ATP like in glycolysis. Since the enzyme phosphofructokinase is missing in many bacteria, an alternative path has to be taken. The C1 glucose position is oxidized from the aldehyde to the carboxylate oxidation level. Then follows a dehydration step with a familiar keto–enol tautomerization in a six-carbon sugar acid. In the next step, the keto group of the substrate forms a Schiff’s base with a lysine from the enzyme, which catalyzes the splitting of the compound into pyruvate and G3P. With the latter, we are again at the midway of the glycolysis, and the next steps follow as usual. However, there is a difference. The other half of the molecule is pyruvate and thus already at the end of the glycolytic pathway. You gain therefore only two ATP from the transformation of G3P to pyruvate. Since you had to invest one ATP in the initial phosphorylation of glucose, your net gain is

only one ATP from this pathway, which is a meager half of the exploitation of glucose in glycolysis.

Due to this energetic limitation, this pathway is only used in aerobic bacteria that can use the NADH produced in the oxidation of glucose to gluconate and in the oxidation of G3P in a respiratory chain. The only fermentative bacterium using this pathway is *Zymomonas*. This bacterium is adapted to environments with high sugar concentrations permitting such a wasteful metabolism (Seo et al. 2005). This variation also contains a lesson. The Entner–Doudoroff pathway demonstrates clearly that the glycolytic pathway is not the only way of sugar degradation and in fact two further alternatives exist in microbes (e.g., the phosphoketolase and the *Bifidobacterium bifidum* pathways).

The Pentose Phosphate Pathway

I have on purpose chosen the Entner–Doudoroff pathway because it is actually a hybrid. The early steps of this path up to 6-phosphogluconate are practically identical to the pentose phosphate pathway; the later steps are identical to glycolysis. In biochemistry the same elements are frequently used in new combinations. The pentose phosphate cycle is itself a new variation on the glucose degradation scheme. Such variations are important to cells since they allow them to maintain several pathways with distinct and even competing metabolic goals in parallel. Nature must only introduce regulated enzymes that are responding to metabolic signals that allow an appropriate channeling of the substrate flows. If this is not possible, higher cells also have the option of locating competing pathways into different cellular compartments. Actually, the function of the pentose phosphate pathway is not energy metabolism (catabolism). Its goal is to provide precursors (ribose 5-phosphate) and cofactors (NADPH) for biosynthetic pathways (anabolism). The destiny of ribose is clear: It is the precursor to nucleotides that make RNA, DNA, and a number of coenzyme nucleotides. NADPH is the cellular currency of readily available reducing power. In biochemistry textbooks, it is frequently stated that catabolic reactions are generally oxidative, while anabolic reactions are generally reductive. Catabolic enzymes like glucose 6-phosphate dehydrogenase from the Entner–Doudoroff pathway use NAD^+ , while the enzyme catalyzing the same reaction in the pentose phosphate pathway uses NADP^+ . This is, however, often not true. In *E. coli* and many other bacteria, it is the same enzyme that does both jobs with one cofactor. In animals the intracellular ratio NAD^+/NADH is high (actually 700 in a well-fed rat), which favors hydride transfer from the food substrate to NAD^+ and thus channels electrons into the respiratory chain. In contrast, the ratio $\text{NADP}^+/\text{NADPH}$ is low (0.014 in the same rat), which favors the hydride transfer from NADPH into biosynthetic pathway.

The biochemical details of the pentose phosphate pathway are actually quite complicated despite the fact that it consists basically of a combination of a few basic chemical reaction types. If you see the written partition of this pathway (the interested reader is encouraged to consult standard biochemistry books), it is all too apparent that it shares one major biochemical motif, namely the

interconnected *trans*-ketolase and *trans*-aldolase reactions, with the Calvin cycle in the dark reaction of photosynthesis for CO₂ fixation. Actually, the Calvin cycle is not an invention of photosynthesis since it also represents the major, although not the only, CO₂ fixation pathway in nonphotosynthetic prokaryotes. *It can thus also claim substantial antiquity.*

Gluconeogenesis

Nature tried something, which is also very popular in music, namely a reversal of the glycolytic theme. This movement also has a biochemical name and is called gluconeogenesis. On the biochemical partition, this pathway reads like a reversal of glycolysis at first glance. But this is of course thermodynamically not possible.

In reverting from phosphoenolpyruvate (PEP) to glucose, gluconeogenesis uses the same intermediates as glycolysis, but two enzymes differ. These are two phosphatases, which reverse the phosphorylations done by phosphofructokinase and hexokinase in glycolysis. The hydrolysis of glucose 6-phosphate to glucose appears to be a chemical child's play. In reality it is a surprisingly complex reaction. Glucose 6-phosphate is transported into still another compartment (the lumen of the endoplasmic reticulum), where a complex of five proteins catalyze the reaction.

A popular start point for gluconeogenesis in the liver is lactate coming from the exercising muscle and erythrocytes. Lactate is first converted into pyruvate by lactate dehydrogenase. This pyruvate becomes the starting material for glucose resynthesis. Liver glucose travels back to the muscle where glycolysis powers the movement and generates lactate. This physiologically important metabolic highway between muscle and liver is the Cori cycle.

Glycolysis takes place in the cytoplasm, which is not so for gluconeogenesis, which calls different cellular compartments into action. Cytoplasmic pyruvate is first transported into the mitochondrion where it is carboxylated to oxaloacetate. This is an important biochemical reaction since oxaloacetate is not only a stoichiometric intermediate in gluconeogenesis, but also the carbon skeleton for the *trans*-amination reaction creating the amino acid aspartate and a catalytic intermediate of the citric acid cycle. Oxaloacetate leaves the mitochondrial matrix after reduction by NADH to malate. In the cytoplasm, malate is reoxidized to oxaloacetate. Oxaloacetate is then simultaneously decarboxylated and phosphorylated to yield PEP. To drive this reaction, one GTP has to be spent. If not enough GTP is available, oxaloacetate cannot be used for gluconeogenesis and goes into the citric acid cycle, which creates GTP. Thus the energy status of the cell decides in what pathway oxaloacetate is actually used.

Evolution of Metabolism

All the above-mentioned pathways are either chemically related or occur even in the same cell. Apparently, a relatively small set of basic chemical reactions provides the very fabric of life as we know it on our planet. These basic

reactions come in many variations and are also variously reassorted into new biochemical themes. Nature is very modular in its construction and always reuses old clothes to make new suits. It is actually not correct to say that nature uses the same chemicals; in fact, nature uses a rather limited set of enzymes that creates these chemical intermediates. This is a subtle, but important difference. The question is thus not whether these compounds are special and cannot be replaced by other chemical compounds. In the following section we will see two variant glycolytic intermediates that do not figure in the main chemical pathway, but are important chemical regulators of intermediary metabolism. The crucial contribution was the invention of the protein enzymes that learned to handle a minimal set of chemical reactions. The cell does not play biochemistry like a student who learns pathway by pathway when advancing from chapter to chapter in the biochemistry textbook. If you take the simplest case of a unicellular prokaryote, you have actually a bag filled with a viscous solution containing a high protein concentration and a large number of small chemical compounds. Metabolism means that a food chemical is introduced into the system, turned over by the protein enzymes, and useless end products of the cellular chemical exercises are excreted from the cell. Some enzymes come in complexes and the substrate is actually reached from one enzyme to the next. However, many other enzymes from metabolic pathways that are so neatly assembled in the biochemistry textbook occur without much architectural order somewhere in the cytoplasm. They are exposed to all substrates and product chemicals at the same time and must do their job. Diffusion is thus the limiting factor for chemical communication in the living cell. Diffusion is not such a barrier for small substrate molecules that reach every point in the relatively large mammalian cell within a tenth of a second. Proteins have a much lower diffusion coefficient and are frequently retained by protein–protein interactions. This means that a pathway is in fact an event that is only reconstructed by the human mind.

Most enzymes do not know much about the other enzymes in the same pathway and do not know whom to deliver what. The only thing they have learned in evolution is to do the more or less specific catalytic reaction and to interpret the chemical environment from the small molecules it meets. If every enzyme would only execute its function every time it meets its substrate, only a rather primitive metabolism could be built. This was certainly the way primitive protein enzymes reacted early in the biochemical evolution. At this level, protein enzymes function like a perception. This term comes from a device that portrays the basic reaction of synapses, which process an input signal into an output signal. In the case of enzymes the input is the substrate, the output the product chemical. Already this system has some flexibility since the input–output relationship can take a linear, hyperbolic, or sigmoidal form according to the construction of the enzyme. A cross talk between proteins is possible only if the product of one enzyme becomes the substrate of another enzyme, and this relationship must hold for many proteins. It is evident that this restriction necessitates that early metabolism could only be built with enzymes that spoke a common chemical language. One can surmise that the chemicals of the central intermediary pathway must belong

to this Esperanto chemical vocabulary understood by all living systems. To get a primitive metabolism the enzymes had to travel together through organisms if the communication between them should not be interrupted resulting in chemical deadlocks in the cell with the accumulation of a product that could not any longer be processed. This need for the chemical Esperanto is probably also the reason why the variant ways of glucose handling share so many common chemical intermediates. A network can only be created if sufficient numbers of common nodes are shared. Only stepwise could variations be introduced which had to use the common elements. To remain in the language picture, in later steps of evolution when the organisms reached already some maturity, they could differentiate and develop in addition to the Esperanto their local language, which is only understood in their corner of the biological world.

Evolution of Complexity

Nature apparently soon discovered that it could use proteins as computational elements in the living cell (Bray 1995). The individual protein elements became stepwise more complicated. The next step in the development of metabolic networks was taken when proteins learned to read more than one chemical signal as in allosteric enzymes. They did not automatically process the substrate into the product; if, for example, the end product of the biosynthetic pathway accumulated (thus indicating no need for further synthesis), it was sensed by the allosteric enzyme and the catalytic activity dropped. Aspartate transcarbamoylase from the de novo pyrimidine nucleotide synthesis pathway fits into this category. It “reads” with its six catalytic subunits the two substrates aspartate and carbamoyl phosphate and transforms them into the product N-carbamoylaspartate. Concomitantly it “reads” with its six regulatory subunits the end product of this pathway CTP. Binding of CTP shifts the K_M for aspartate to a higher concentration. The next step in the evolution of computational devices is proteins that function as molecular switches; an example is CaM kinase. It binds Ca^{2+} complexed with calmodulin. This binding activates the kinase activity and CaM kinase phosphorylates many different other target proteins. Another variation of the scheme is represented by glycogen synthase, the enzyme making the storage form of glucose in animals. It is the target of six protein kinases and several protein phosphatases that add or take away a phosphate group from the enzyme affecting its enzymatic activity. The kinases and phosphatases acting on glycogen synthase themselves come under the control of other signal chemicals, including hormones, such that different organs can now speak with the regulated enzyme. The chemical cross talk possibilities could still be increased when nucleotide-binding proteins were included into the network or protein phosphorylation cascades were invented. Now much more sophisticated switches than simple ON/OFF decisions could be constructed; AND and OR or NOT gates could be built into the metabolic network. It appears that the picture of the blind watchmaker applies also to the design of metabolism. Over long evolutionary periods, the coordination of metabolism was optimized by changing rate and binding constants of enzymes and then their chemical cross talk, all step by step in an

interactive random way until the system as a whole performed in a selectively advantageous way. Redesigning the basic rules became from a given degree of complexity impossible. There was only a single way forward: you could only overlay new layers of complexity on the old layers. Interestingly, the increasing complexity in biological systems was actually not achieved by substantially increasing the number of enzymes. In fact, we differ from our gut bacterium *E. coli* only by a factor of 10 with respect to gene number. Apparently, there are constraints on how many different molecular nodes you can introduce into a metabolic network. The control circuits increased substantially in complexity, but less so the number of chemical reactions.

Glycolysis from the side of the chemical intermediates is very similar in our gut bacterium *E. coli* as in our own body. However, *E. coli* has only to integrate this pathway into the needs of a single cell, which is already a great job. In contrast, glycolysis in humans needs the coordination of something like 10^1 organs, 10^2 tissues and about 10^{13} cells (to speak only of orders of magnitudes). It is apparent that the same set of chemical reactions came here under the control of enzymes that have to integrate a far more complex set of signals. The mathematical description of glucose flow in *E. coli* is a showcase of systems biology (more on it further down), while glucose handling in human diabetes is still beyond a detailed understanding despite the large number of biomedical researchers working in this high priority field of contemporary medicine. Before presenting more data on metabolic networks as revealed by systems biology approaches, I want to review some more “traditional” issues around glycolysis.

Variant Glycolytic Intermediates

2,3-Bisphosphoglycerate

First, I want to illustrate the role of variant glycolytic intermediates with two examples. The first one is 2,3-bisphosphoglycerate (BPG) from erythrocytes. The fact that erythrocytes, the dedicated oxygen transporters of our body, derive their energy from glycolysis might sound paradoxical. Apparently, nature found that respiration and dedicated oxygen transport could create conflicts of interest and preferred to muzzle erythrocytes energetically. However, this is not a major handicap because erythrocytes swim in a carefully controlled glucose solution—blood. Glucose is taken up and phosphorylated to glucose 6-phosphate. Then comes an interesting metabolic split into competing pathways within the same cell. Five to ten percent of the glucose is used in the pentose phosphate pathway for NADPH production. The reason is clear: NADPH is needed to cope with the oxidation stress imposed by its transport functions. The remaining 90% of the glucose goes into glycolysis where erythrocytes show an interesting sideway. Part of 1,3-BPG is not transformed in an ATP-generating step to 3-phosphoglycerate. Instead a mutase transforms part of it into 2,3-BPG. *Quite substantial amounts of this compound can accumulate in erythrocytes. A specific phosphatase transforms it back into the glycolytic intermediate 3-phosphoglycerate, but this means a missed opportunity with respect to*

energy gain. The strong selective forces acting on all living systems will assure economical solutions to organisms. If this principle is violated (it is in fact frequently violated because organisms do not search necessarily the cheapest solution, they also have to search adaptable solutions), you can generally make the bet that nature had a hindsight. Against common wisdom, wasteful solutions can even be imposed by selection. If you look at this question at the whole animal level, you can think of the tail of the peacock or the antlers of the extinct Irish elk, which is a metabolically wasteful, even harmful development with respect to predation or nutrition but nevertheless imposed by selection forces (this time by the sexual preferences of females for a big male sexual display organ). Mean and lean is thus not necessarily the optimal solution in biology. In fact BPG is an important regulator of oxygen transport. It binds to the hemoglobin and lowers hemoglobin's affinity for oxygen by stabilizing the T state. The physiological level of 5 mM BPG in the blood (this is a substantial amount and equals the steady-state concentration of blood glucose) assures that 38% of the oxygen cargo is delivered in the peripheral tissue. If you now climb without adaptation to high altitudes, you have a problem. At 4,500 m above sea level, the partial oxygen pressure is only 7 kPa, less oxygen is bound in the lungs, and the hemoglobin would only release 30% of its cargo in the tissue. As a quick fix solution, the blood BPG level will have risen after a few hours to 8 mM. This rise shifts the hemoglobin oxygen binding, assuring again 37% oxygen release. After returning to the sea level, the BPG levels decrease again. A sideway of glycolysis thus becomes an important regulator of respiration.

Fructose 2,6-Bisphosphate

Another example is fructose 2,6-bisphosphate. The enzyme, which carries the impossible name of 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase, leads to the synthesis of this compound that resembles another glycolytic intermediate. As even biochemists cannot pronounce easily this tongue twister, this enzyme is shortly called PFK-2. This enzyme is really a maverick: if you read its name it does one thing (kinase: a phosphate transfer) and its opposite (phosphatase: hydrolysis of a phosphate group). Actually, if you look at its domain structure, these are two fused enzymes. Not enough with that oddity, it synthesizes fructose 2,6-bisphosphate, which closely resembles the glycolytic intermediate fructose-1,6 bisphosphate.

Phosphofructokinase, also called PFK-1, is the committing step into glycolysis and the most important regulator of the flux of glycolysis. The enzyme is inhibited by ATP; quite logically, glycolysis decreases when the energy charge of the cell is high. PFK-1 is also inhibited by H^+ . This control loop prevents lactate dehydrogenase excessively reducing pyruvate to lactate. Phosphofructokinase is further inhibited by citrate. High citrate concentrations signal ample supply of biosynthetic precursors to the citrate cycle. Additional glucose is thus not needed to fill up these pools. However, the most powerful activator of PFK-1, which overrides the other signals, is fructose 2,6- bisphosphate synthesized by PFK-2. This compound increases the affinity of PFK-1 for its substrate from a

sigmoidal to a hyperbolic velocity–concentration curve and offsets the inhibitory effect of ATP on PFK-1. This is now another case where a close relative of a glycolytic intermediate fulfills a regulatory role. *Are these relics from the try-and-error phase of early glycolysis when many compounds were explored for their suitability to construct glycolysis, the major metabolic highway of cellular life? Did one compound actually made it to a pathway intermediate and the other got the consolation prize to be an important regulator? Or did Nature try on purpose the related compounds for regulation that are sufficiently similar to existing intermediates that they can be made and used by modifications of the available enzymes, while at the same time being sufficiently different to allow separate control?*

Lactate and Ethanol Fermentation: A Bit of Biotechnology

Pyruvate is a very versatile intermediate. In the textbooks of biochemistry, it is the end product of glycolysis, but due to the need for reoxidation of NADH, the carbon metabolism cannot stop here. A one-step reaction fulfilling this requirement is the reduction of pyruvate to lactic acid by lactate dehydrogenase. In our body, only L-lactate is synthesized because our lactate dehydrogenase strongly prefers hydrogen transfer from the A site of NADH. This is not the case in all organisms. In lactic acid bacteria, some lactate dehydrogenases prefer the A site, others the B site of NADH; the latter lead to the production of D-lactate. Some bacteria produce mixtures of D,L-lactic acid.

Lactic Acid Bacteria

Streptococcus thermophilus, which ferments yogurt in symbiosis with *Lactobacillus bulgaricus*, is a homolactic starter bacterium (Figure 2.1). This means it produces primarily lactate from lactose. It excretes this lactate as a waste product together with protons despite the fact that it still contains a lot of chemical energy. This excretion of protons is driven by the membrane-bound bacterial ATPase. Lactate excretion also leads to an acidification of the fermented milk, leading to a precipitation of milk proteins and the buildup of a semisolid food matrix. Due to their peculiar metabolism, lactic acid bacteria develop an acid resistance (down to a pH of 3.5) that is greater than that of most other bacteria. In fact, during the spontaneous fermentation of cabbage to sauerkraut, which relies on the bacteria naturally associated with the vegetables, one sees a succession of different populations of lactic acid bacteria (*Leuconostoc* sp. followed by *Lactobacillus plantarum*, which is likely driven by a phage that kills *Leuconostoc*). The succession is also dictated by their different degrees of acid resistance. The final products, sauerkraut or yogurt, are relatively stable and can be stored much longer than fresh lettuce or milk, respectively. However, the exclusive production of lactate by streptococci is observed only when the cells grow in the excess of substrate like lactose in milk. When the substrate is offered in growth-limiting amounts, some streptococci change to further exploitation of pyruvate as far as

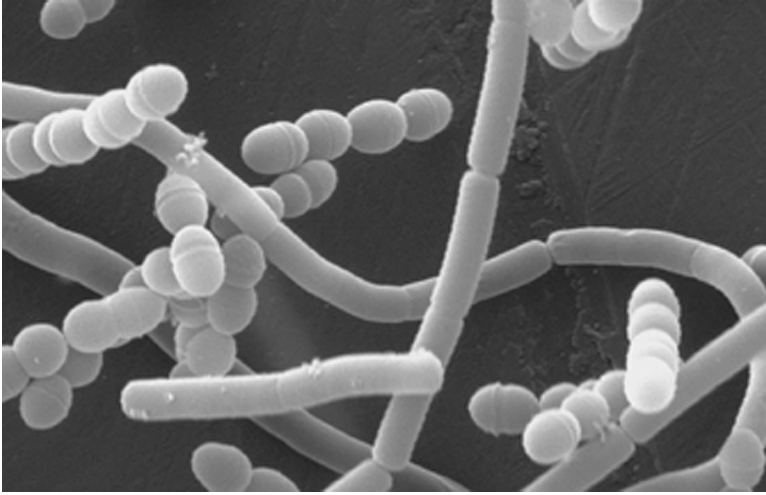


FIGURE 2.1. The picture shows the bacterial consortium (starter bacteria) which achieves yogurt fermentation. The culture consists of *Streptococcus thermophilus*, the short chains of globular cocci in the picture, and *Lactobacillus bulgaricus*, the chains of elongated cells. Despite their different morphology, both bacteria are phylogenetically closely related low GC content Gram-positive bacteria. The scanning electron microscope reveals cell division, but nothing from the interior of the bacterial cell.

their catalytic activities reach. Pyruvate is, for example, decarboxylated, and the energy of the reaction is stored in the thioester bond of acetyl-CoA. However, streptococci do not possess a Krebs cycle. They had to invent another way to exploit the energy contained in acetyl-CoA. They achieve this by the transfer of the acetyl group on a phosphate. The acetylphosphate created in this reaction can then transfer the phosphate on ADP, creating another molecule of ATP and acetate. Many other chemical end products of fermentation exist that give the specific fermented food its flavor and attraction to the human consumer. Some are also more decorative like in Swiss cheese production where the end products are propionic acid, giving this cheese its particular flavor, and CO_2 , which gives decorative holes in the Swiss cheese. Lactic acid bacteria can even produce ethanol by two reduction steps of acetyl-CoA to acetaldehyde and then to ethanol.

Zymomonas and Ethanol

In fact, the dairy industry is second in the food industries and follows the largest branch, producing alcoholic beverages. Yeasts play the dominant role in industrial alcohol fermentation, but the bacterium *Zymomonas* might play an increasing role in the future. Both organisms reach the same maximum alcohol concentration of 12% by fermentation. However, under batch fermentation and continuous fermentation conditions, *Zymomonas* produces about 10-fold higher

amounts of ethanol per biomass. In other words, less biomass has to be disposed of for the same amount of ethanol produced, which is of technical interest. *Zymomonas* uses the Entner–Doudoroff pathway for sugar degradation, and it yields only half of the ATP than glycolysis provides to yeast, hence the produced biomass is much smaller.

The brewing of beer is a complicated process, some would say even an art. The yeast *Saccharomyces cerevisiae* can ferment glucose to ethanol. Under anaerobic conditions, this is the only mode of energy production. In the presence of oxygen, respiration occurs. Since glucose respiration yields much more ATP than glucose fermentation, the yeast cell must compensate this yield difference by a higher glucose consumption under anaerobic condition when compared to aerobic condition (“Pasteur effect”). However, alcoholic fermentation may set in even under aerobic conditions if the glucose concentration surpasses a critical threshold value (“Crabtree effect”). This metabolic flexibility makes yeast a difficult organism for the food industry. Mastering its ethanol production capacities by a timely change from oxygenic to anoxygenic metabolism in yeast makes beer making an art. Industrially yeast is thus not an easy beast to tame.

Genetic Engineering

Zymomonas has an exclusively anaerobic metabolism and lends itself to a more straightforward industrial processing. *Zymomonas* has traditionally been used for the production of alcoholic beverages. For example, the popular Mexican drink pulque is made with *Zymomonas* from the sap of the agave plant. It has thus some industrial potential, but it can only use a relatively limited number of sugars as carbon substrates for alcohol production: glucose, fructose, and sucrose. These are valuable sugars also needed in food and feed, but what about less valuable carbohydrate sources like xylose. The latter is a waste material produced as a by-product of industrial pulp and papermaking. What is the prospect of metabolic engineering? Will it be possible to introduce genes into *Zymomonas* that allow its growth on waste containing xylose? Genetic engineers constructed a shuttle vector that can travel between *E. coli* and *Zymomonas*, which carries two xylose assimilation genes and a *trans*-ketolase and an aldolase that can funnel the pentose sugar via the pentose pathway into the Entner–Doudoroff pathway (Zhang et al. 1995). The xylose genes were placed under the control of a strong constitutive *Zymomonas* promoter, namely that of the glyceraldehyde dehydrogenase gene. The recombinant bacterium did what the researchers hoped—it could grow on glucose or on xylose, and it converted xylose to ethanol at high yield. This was not a trivial result. Not only does it allow gaining a valued compound from a waste product (perhaps not in your beer, but as bio-fuel in heating or in a car), it has even important theoretical implications. The experiment tells that you can add a few genes, which open new nutritional possibilities to a microbe. By acquiring a few crucial genes, it can conquer a new environment. The metabolism of this organism is not in such a poised equilibrium that any new metabolic trafficking would upset the cell. This result is also of substantial theoretical interest for the metabolic network

discussion, which we will touch in a different section. Encouraged by these results, the genetic engineer tried to open other food sources for *Zymomonas*. Xylose is mainly found in hardwood, but there are also valuable energy crops like switchgrass that contain large amounts of arabinose. The researchers introduced into the same shuttle plasmid three arabinose-degrading genes. The strategy was the same, and it paid off: *Zymomonas* could also grow on arabinose. Vice versa, Lonnie Ingram has tried to introduce *Zymomonas* genes into *E. coli* to produce ethanol in this workhorse of the biotechnology industry (Ohta et al. 1991; Tao et al. 2001). Since *E. coli* has many pathways starting from pyruvate, the construction of such strains also necessitated the inactivation of *E. coli* genes to channel the metabolism into the desired direction. *Klebsiella oxytoca*, a cousin of *E. coli*, thus became an ethanol-producing bacterium by genetic engineering.

A Short Running Exercise

Energy Stores for Muscles

Animal life in contrast to plant life is defined by a fundamental property: locomotion. To move around you need muscles. The movement of muscles is powered by molecular motors, which consist of thin and thick filaments that slide past each other during contraction. The molecular interplay of myosin and actin is powered by ATP. We run differently when we anticipate a short distance sprint intended for catching a prey or escaping a predator (or nowadays for a sports event) or when we envision a medium or long distance run. We dose the running speed differently for the simple reason that we cannot sustain the sprinter's speed for very long. The reason for this behavioral adaptation is our empirical knowledge of our physiology. Yet this empirical knowledge can be rationalized in the light of our biochemical knowledge. The stores of ATP in the muscles are low. They were calculated to be 200 mmol for a 70-kg person with a total muscle mass of 28 kg. This keeps you running at full speed, perhaps for 2 s. Therefore for a 100-m catch or escape sprint you need to tap another energy resource. This is creatine phosphate; our reference person contains of it the equivalent of 400 mmol ATP. It is biochemically more inert than ATP; hence, the cell exploits it as a transient energy buffer. As creatine phosphate can directly transfer its high-energy phosphate to ADP, ATP is quickly regenerated. However, this store does not bring you to the 100-m mark. Therefore you need to exploit the next energy store: glycogen, a polymeric form of glucose residues. Muscles have by far the greatest glycogen store of the human body followed by the liver. According to the energetics of the glycolytic and respiratory metabolism, the equivalent of the same amount of muscle glycogen of our test person is worth either 7,000 or 80,000 mmol ATP. However, there is a difference between both modes of glycogen use. Your running speed will decide about life or death: you get the prey animal you want to eat, which assures your nutritional survival over the next days or weeks, or you don't get it with potential dire nutritional consequences. Or stated more dramatically, you fall victim to a predator or you happily escape from this attack, assuring your

physical survival. Running speed is a direct function of the maximal *rate* of ATP production, and it differs between glycolysis and respiration. If you use glycolysis, you get ATP relatively quickly from substrate-level phosphorylation. If you bet on the respiratory chain, you get more for your money, but it takes longer to get to the superior level of ATP. The reason is simple: Metabolizing glucose to CO_2 needs more time because there are more chemical reactions and more cellular compartments involved. The glycolytic rate is about twofold higher than the oxidative rate of ATP production. A human sprinter will thus opt for glycolytic use of glycogen, and this brings him over the 100-m distance. The muscle recalls this running spree: After the run, its ATP level is down from 5.2 to 3.7 mM, its creatine phosphate has diminished from 9.1 to 2.6 mM, while the blood lactate level is up to 8.3 from initially 1.6 mM. During intensive exercise, the blood is flooded with lactate from the muscle. It is the job of the liver (and partly the kidney) to take care of lactate. Both organs synthesize glucose from lactate via gluconeogenesis. They deliver this newly synthesized glucose into the blood stream. The muscle can move on in using glucose to sustain its strenuous exercise.

Liaison Dangereuse: Lactate, Cancer, and the Warburg Effect

Imaging Techniques

Modern imaging techniques in medicine have revolutionized the biochemical investigation that can now also be made with human beings. Here I will discuss only one technique, positron-emission tomography (PET). This technology uses short-lived isotopes of carbon (^{11}C), nitrogen (^{13}N), oxygen (^{15}O), and fluorine (^{18}F), which are produced in a cyclotron. Their half-life ranges from 2 (^{15}O) to 110 min (^{18}F) and represents therefore only a low-radiation burden to the subject. These instable isotopes loose a positron and fall back to a stable nuclear level. The positron carries the mass of an electron, but in contrast to the negatively charged electron, the positron carries a positive charge. After emission from the instable isotope, the positron travels a few millimeters in the tissue until it meets an electron. This leads to a matter–antimatter collision, mass annihilation, and emission of gamma rays. A positron camera, which is arranged cylindrically around the subject, measures this emission. A computer calculates the position of the collision and constructs virtual cuts of the patient, which are called tomographs. The instable isotopes allow the labeling of many biochemical substrates. Take glucose: Glucose is transformed into 2-deoxyglucose (a hydrogen atom replaces the hydroxyl group) and then into 2- ^{18}F -2-deoxyglucose, or FDG (the instable fluorine replaces this hydrogen). Fluorine takes about the size of the hydroxyl group and can thus fool the cells. It is taken up like glucose and is phosphorylated by hexokinase like glucose. However, the next enzyme of glycolysis

is cleverer, it denies further metabolism. In tissues with low glucose 6-phosphatase, the glucose analogue is now confined to the cell (metabolic trapping). Tissues with higher glucose turnover can now be localized. The turnover can even be quantified and expressed as $\mu\text{mol glucose}/100\text{ g tissue} \times \text{min}$. Since the brain gets more than 99% of its energy from glucose and since the technique has a substantial anatomical resolution when combined with nuclear spin tomography, PET can now localize brain areas in the active subject performing different tasks (calculations, speaking, reading, etc.). FDG-PET has another important clinical application: It detects malignant tumors with a specificity and sensitivity near to 90% if the cancer is greater than 0.8 cm^3 .

Warburg Effect

The biochemical basis for this diagnostic method goes back to one of the founding fathers of biochemistry. Nearly 70 years ago, Otto Warburg observed that in cancer cells the energy metabolism deviates substantially from that of the normal surrounding tissues. The activity of a number of glycolytic enzymes, like hexokinase, phosphofructokinase, and pyruvate kinase, and the glucose transporter are consistently and significantly increased. Glycolysis proceeds about 10-times faster in most solid tumors than in neighboring healthy cells. The Warburg effect is in fact the negation of the Pasteur effect: While normal cells change their energy metabolism from glycolysis to aerobic respiration as oxygen becomes available, the accelerated glycolysis of the cancer cell is maintained in the presence of oxygen. It was hypothesized that this phenotype is a consequence of early tumor growth, when clusters of cells grow without any vascularization. The absence of blood vessels means no oxygen, and the precancerous cells get an energy problem that they try to fix by an increased glycolysis. In fact, the center of these cancerous lesions frequently shows a necrotic zone. It was therefore logical to implicate an adaptation to hypoxia (low oxygen pressure) as the key to this phenomenon. Indeed, hypoxia-inducible factor 1 (HIF-1) turned out to be a key transcription factor. It upregulates a series of genes involved in glycolysis, angiogenesis, and cell survival. HIF-1 is a heterodimer of two constitutively transcribed subunits HIF-1 α and HIF-1 β . The complex is regulated via instability of HIF-1 α . Under normal oxygenic conditions, HIF-1 α undergoes ubiquitination, i.e., it gets a chemical death certificate fixed at it such that it is proteolytically degraded in a proteasome—a destruction complex that dutifully destroys all proteins with this ubiquitin signal. As one would expect for such a destructive signal, its allocation is under careful control. Before the death tag can be fixed, a critical proline residue must first be hydroxylated by one enzyme, and this enzyme is active only in the presence of oxygen and iron. After this first chemical modification, a tumor suppressor protein binds, and ubiquitination can occur (Lu et al. 2002). So far, so good. HIF-1 plays an important role in normal cellular metabolism when the oxygen level falls as in the exercising muscle where it is only one tenth of that in the nearby capillary blood. Here it makes sense when HIF-1 increases the flux through glycolysis. Mice, which got their HIF-1 gene knocked out, do

not show this increased glycolysis; they do not accumulate lactate while muscle ATP remains at the same level, thanks to an increased activity of enzymes in the mitochondria. Mutant mice show first a better endurance under exercise conditions. However, under repeated exercise, the muscles of these mutants showed extensive histological damage. HIF-1 thus achieves a self-protection of the muscle (Mason et al. 2004). However, in tumors HIF-1 is not inactivated, even when oxygen is present. Tumor cells do not change for oxygenic respiration. Glycolysis remains high, and lactate levels elevated. It was even argued that lactate became a chemical club for the cancer cell, not unlike how lactic acid bacteria eliminate competing bacteria by lactate production. The malignant cell developed during carcinogenesis a marked acid resistance, which is not found in the healthy surrounding cells. The cancer cell can thus poison the tissue and erode its way through the tissue into the next blood vessel to build metastases (Gatenby and Gillies 2004). We see here how a metabolic end product gets, perhaps, a selective function in the fight for resources, here for a malignant cell clone against its parent organism. We will later see another comparable case in cyanobacteria, which evolved oxygen in their new metabolism, which became a poison for competing, but oxygen-sensitive bacteria. Otto Warburg predicted that we would understand malignant transformation when we understand the cause of increased glycolysis in cancer cells. We understand now some parts of the control cycle from hypoxia over HIF-1 to increased glycolytic flux. One of the latest news at the time of the writing was the binding of HIF-1 to a DNA consensus sequence called hypoxia response element (HRE) in the promoter region of PFK-2. Binding of HIF-1 increases the transcription of this important regulator of glycolysis.

Glucokinase at the Crossroad of Cellular Life and Death

Apoptosis

Cancer cells not only have an increased metabolic rate but are also less likely to commit suicide when damaged. Cell suicide is an everyday event in multicellular organisms and occurs by a carefully orchestrated process called apoptosis. That energy metabolism and cell death are intimately linked processes was already indicated by the fact that mitochondria play a key role in apoptosis. Many apoptotic signals converge on mitochondria where they cause the release of cytochrome *c* from mitochondria. Cytochrome *c* in the cytoplasm in turn triggers the activation of a group of proteins called caspases, which lead to an ordered cell destruction from within. Recently glucose and a key enzyme of glycolysis became implicated into another pathway leading to apoptosis, further strengthening the links between energy metabolism and programmed cell death. The proapoptotic player is designated with the acronym BAD. The activity of BAD is regulated by phosphorylation in response to growth and survival factors (Downward 2003). In liver mitochondria, BAD exists in a large protein complex consisting of two proteins that decide on the phosphorylation status of BAD. One is the protein kinase A, which still needs an anchoring protein to get into this complex, and the

other is protein phosphatase 1. The fifth partner was, however, a surprise (Danial et al. 2003). It was also a kinase; specifically it was glucokinase, a member of the hexokinase family. This enzyme catalyzes the first step of glycolysis, it phosphorylates glucose to glucose 6-phosphate, and is also involved in glycogen synthesis in the liver cell. Addition of glucose to purified mitochondria induced the phosphorylation of BAD. Liver cells deprived of glucose go into apoptosis unless BAD is also missing. Glucokinase activity was markedly blunted when the three phosphorylation sites on BAD were mutated. Insulin stimulates in the liver cell not only glucose transport, but recruits hexokinase to the mitochondria and stimulates glycolysis and ATP production. Conversely, withdrawal of growth factors decreases the glycolysis rate, O₂ consumption, and closes the VDAC channel, which mediates the ADP for ATP exchange over the mitochondrial membrane. This channel closure is also linked to cytochrome *c* release from the mitochondria. In this way, the glucose metabolism is crucially linked to cell survival at least in liver cells.

Metabolic Networks

Quantitative Biology

*To understand the complexity of biological systems, we must analyze it. This means that we have to dissect it either literally with a lancet or intellectually by teasing apart the constituting elements. In this dissected way, we learn biology in the university textbooks. Individual pathways are presented chapter by chapter; only in later chapters of the textbooks, some integration of metabolism is offered. So far as learning is concerned, this is a fine method. However, humans tend to mix the representation of a thing with the real thing. In a word of Zen philosophy, the finger, which points to the moon, is not the moon. The analytical method needs the synthetic method as a necessary complementation. This is still relatively easy in chemistry where analytic and synthetic chemistry are classical and complementary branches. In view of the awful complexity of biological systems in chemical terms, biologists tend to be highly satisfied when they have pulled the molecules apart. However, in simple, but highly investigated systems like the gut bacterium *E. coli*, engineers and physicists started to synthesize the whole picture from its constituents by using computational methods. The mathematical treatment of the metabolism has a tradition that dates back into the 1970s. To predict cellular behavior, each individual step in a biochemical network must be described with a rate equation. As these data cannot be read from genome sequences, and since empirical data are lacking, Edwards and Palsson (2000) tried to model metabolic flux distributions under a steady-state assumption. Based on stoichiometric and capacity constraints, the *in silico* analysis was able to qualitatively predict the growth potential of mutant *E. coli* strains in 86% of the cases examined. In follow-up studies, *in silico* predictions of *E. coli* metabolic capabilities for optimal growth on alternative carbon substrates like acetate were confirmed by experimental data (Edwards et al. 2001).*

Optimal Growth Rate Planes

Annotated genome sequences can be used to construct whole cell metabolic pathways. In particularly well-investigated systems like *E. coli* and *S. cerevisiae*, limiting constraints can be imposed on the *in silico* calculated model based on mass conservation, thermodynamics, biochemical capacity, and nutritional environment such that optimal growth rates can be calculated for common carbon substrates. For *E. coli* strain K-12, optimal growth rate planes (phenotype phase plane analysis) were calculated for substrate and oxygen uptake at different temperatures, which avoided both futile cycling and excessive acidic waste excretion. When they did this calculation with different substrates, they found growth rates on the line of optimality for a number of substrates (Ibarra et al. 2002). The absolute growth rate decreased successively when going from glucose to malate, succinate, and acetate. When the cells were serially subcultured on a given substrate, a small but significant increase in growth rate of about 20% was observed. This means that the cells were overall more or less well poised to make their living from these carbon sources, but small increases could still be obtained by selection procedures. However, the predicted optimal growth of *E. coli* on the carbon source glycerol was not achieved by the cells. Was the model wrong? Interestingly, after 700 generations of *E. coli* growth on glycerol the growth rate nearly doubled and came close to the optimal growth rate predicted by the *in silico* model. Apparently, the starting organism was never adapted to growth on glycerol, but could achieve the adaptation by a trial-and-error process demonstrating substantial flexibility in the metabolic network of *E. coli* toward changes in its food source.

Fiat Flux

Other scientists used the central metabolism of *E. coli* for elementary flux mode. This major pathway contains 89 substances and 110 reactions connected in 43,000 elementary flux modes (glucose is involved in 27,000 of these modes). Corresponding to biological intuition, glucose is a much more versatile substrate than acetate: Glucose, in this model, can be used in more than 45 different ways than acetate (Stelling et al. 2002). Glucose was also the only substrate that could be used anaerobically without additional terminal electron acceptors. Their calculations showed a remarkable robustness of the central metabolism. Mutants with significantly reduced metabolic flexibility (important nodes were removed affecting the flow along certain lines) still showed a growth yield similar to wild type. This calculation fits well with gene knockout experiments in *E. coli*: fewer than 300 out of the 4,000 genes are essential in the sense that the deletion of one of them prevented growth in a rich medium (Csete and Doyle 2002). Also, the network diameter did not change substantially. The calculations showed that the cell has to search a trade-off between two contradictory challenges. On one side is flexibility, i.e., the capacity to realize a maximal flux distribution via redundant node connections, and on the other side is efficiency, i.e., an optimal outcome like cell growth with a minimum of constitutive elements like genes and

proteins. This trade-off control should correlate with messenger RNA (mRNA) levels. The metabolic network structure should help to understand the large amount of mRNA expression data, nowadays provided by microarray analysis. *In silico* prediction fitted relatively well with mRNA expression data from *E. coli* growing on glucose, glycerol, and acetate. An editorial on this report noted that biochemical progress has been driven in the past mainly by new observations rather than theories. Stoichiometric analysis will not only push biology more into the direction of the “exact” sciences like chemistry, but also boost hypothesis-driven experimental research in biology (Cornish and Cárdenas 2002). Recent data underline that this might not be a vain hope. Bioengineers developed the first integrated genome-scale computational model of a transcriptional regulatory and metabolic network for *E. coli*. This model accounts for 1,010 genes of *E. coli* including 104 regulatory genes that control the expression of 479 genes (Covert et al. 2004). The model was tested against a data set of 13,000 growth conditions. The growth matrix tested a large number of knockout mutants for growth on media that varied for carbon and nitrogen food sources. Remarkably, the experimental and computational data agreed in 10,800 of the cases. The discrepancies can be used to update the *in silico* model by successive iterations, and it can also generate hypotheses that lead to uncharacterized enzymes or noncanonical pathways. The updated *E. coli* model can now predict a much higher percentage of the observed mRNA expression changes and highlights a new paradox like the reduction of mRNA levels for a regulatory gene when the protein was in fact activated. System biologists can now start to guide the experimentalist at least in well-investigated prokaryotic systems such as *E. coli*.

Highways ...

In silico predictions of *E. coli* metabolic activities must account for 537 metabolites and 739 chemical reactions (*E. coli* has a 4 Mb genome). Their interconnection results in a complex web of molecular interactions that defies the classical metabolic maps commonly pinned to the walls of research laboratories. The group around Barabasi used the Edwards and Palsson model on *E. coli* but used a less detailed analysis (topology versus stoichiometry). They ran the explicit calculations with two out of the 89 potential input substrates of *E. coli*, namely on a glutamate-rich or a succinate-rich substrate. The biochemical activity of the metabolism is dominated by several “hot” reactions, which resemble metabolic superhighways (Almaas et al. 2004). These highways are embedded in a network of mostly small-flux reactions. For example, the succinyl-coenzyme A synthetase reaction exceeds the flux of the aspartate oxidase reaction by four orders of magnitude. The results are somewhat disappointing because they are obvious and confirm what we knew from the older metabolic wall charts.

... And Small Worlds

In a follow-up study, the same group worked with the genome sequences from 43 organisms representing all three domains of life and compared their metabolic

networks (Jeong et al. 2000). On the basis of the gene annotations, pathways were predicted and data from the biochemical literature were integrated into the model. They based the analysis on a mathematical theory developed in 1960, the classical random network theory. The biochemical network was built up of nodes, the substrates, which are connected to one another through links, which are the actual metabolic reactions. Their first question concerned the structure of the network. Was it exponential or scale free? In biological parlance, this reads, Do most nodes have the same number of links, which follow a Poisson distribution, or does this distribution follow a power-law? That is, most nodes have only a few links while a few, called hubs, have a very large number of links. The answer—according to this study—was clear-cut. In all three domains of life, the power-law described the network structure. This means that a handful of substrates link the metabolic ensemble. Then the physicists investigated the biochemical pathway lengths. Many complex networks show a small-world character, i.e., any two nodes can be connected by relative short paths along existing links. The physicists calculated in their graph-theoretic analysis an average path length of about 3. However, their computer program considered enzymatic reactions as simple links between metabolites. When another group used a more detailed approach considering the pattern of structural changes of metabolites (from the traditional biochemical perspective a much better alternative), the average path length became 8. They summarized the difference in the title of their paper: “The Metabolic World of *Escherichia coli* is Not Small” (Arita 2004). The discussions become somewhat difficult to follow for biologists having just a biochemical education level, but no training in mathematics. However, problems even with sophisticated computer programs become evident for simple-minded biologists. For example, you can grow *E. coli* in the laboratory using just D-glucose as sole carbon source. Thus in reality, all carbon atoms in any metabolite must be reachable from D-glucose. In the best current calculations, this applies only for 450 out of the 900 *E. coli* metabolites. The small-world character of the programs might thus reflect more of their incapacity to reach the other half of the metabolites from glucose than their real short pathway length. Arita (2004) suspected that many carbon atoms in metabolites become reachable from D-glucose via cyclic metabolic pathways. Especially the TCA cycle might be the maelstrom that equalizes the carbon atoms for all cellular metabolites pointing just to another central function of this central turning wheel of metabolism.

Yeast: Many Nonessential Genes

To avoid a too theoretical section, I will mention two biological problems on which metabolic network analysis was applied in a way that is directly understandable for mainstream biologists. To set out the problem, let's start with a basic statement. Gene disruption is a fundamental tool of molecular genetics, where the phenotypic consequence of the loss of the gene function can be determined. We are so much impressed by the effect of some mutations especially in the field of medical genetics that one might a priori expect that most genes are needed for survival of a given organism in its habitat. For organisms with

facile genetic methods and known genome sequence, a systematic experimental approach to this problem can be done. For example, 96% of the 6,000 genes of the baker's yeast *Saccharomyces cerevisiae* (Figure 2.2) were neatly replaced by a deletion cassette, and the resulting mixture of mutants was tested in a rich medium for growth (Giaever et al. 2002). This procedure allowed to assess growth defects as small as a 12% decreased growth rate compared with the parental wild type. The surprise was great: Under laboratory conditions, 80% of yeast genes seem not to be essential for viability. As the rich medium might hide the need for genes, the researchers conducting this genetic tour de force used growth conditions that asked adaptive responses from the yeast cells. The yeasts were grown with restrictions in amino acid availability and changes in sugar carbon source, osmolarity/salinity, and alkaline pH. These more selective conditions added only few further genes to the list of essential functions; this was most marked in the alkali selection protocol, which revealed 128 alkali-sensitive mutants. Another surprise was the observation that genes that are upregulated in expression studies during these selective conditions were mostly dispensable for the survival of the cell.

Yeast: Genome Duplication

There are different possible explanations for these surprising results. One was actually proposed by the scientists of this study: They ascribed the low number of essential genes to the highly duplicated nature of the yeast genome. Indirect

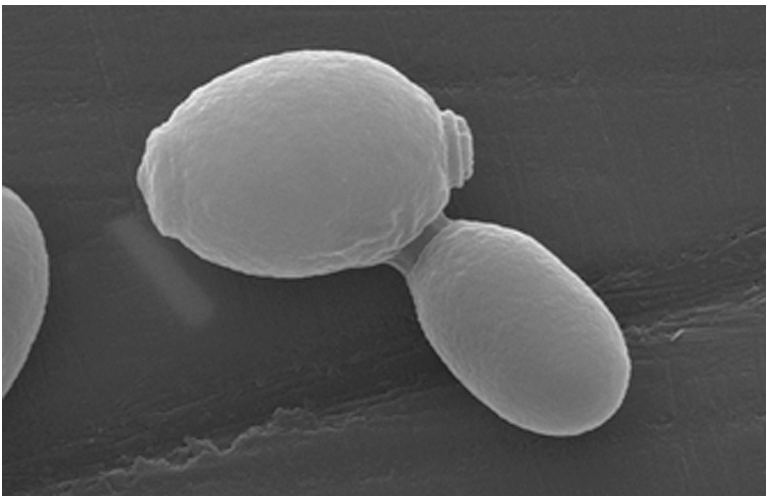


FIGURE 2.2. *Saccharomyces cerevisiae*, a yeast used in wine, beer, and bread production, rivals the role of lactic acid bacteria for importance in food processing. This yeast has been dubbed the *E. coli* of the eukaryotes because it is the best-investigated “higher” organism. The picture shows how daughter cells bud from a yeast mother cell. Note that *S. cerevisiae* is not much longer than many bacteria.

molecular evidence that the entire yeast genome has been duplicated was already around since 1997 (Wolfe and Shields 1997). An elegant confirmation of this hypothesis was provided by scientists from Boston, who sequenced another yeast, *Kluyveromyces waltii*, and compared it to the genome of *S. cerevisiae* (Kellis et al. 2004). The trick was that *Kluyveromyces* had diverged from the ancestor line of *Saccharomyces* before this duplication event. By a comparative genome analysis, it became apparent that the two genomes were related by a 1:2 mapping, with each region of *Kluyveromyces* corresponding to two regions of *Saccharomyces*, as expected for whole genome duplication. The analysis demonstrated that *S. cerevisiae* arose from complete duplication of eight ancestral chromosomes and subsequently returned to functionally normal ploidy by massive loss of nearly 90% of the duplicated genes. The *S. cerevisiae* genome is today only 13% greater than the *K. waltii* genome. The loss occurred in small steps involving at the average two genes. The losses were balanced and complementary in paired regions, preserving at least one copy and virtually each gene in the ancestral gene set. When they looked more carefully into the 450 duplicate gene pairs, they observed frequently accelerated evolution in one member of the gene pair. Apparently, one copy tended to preserve the original function, while the other copy would be free to diverge. The derived gene tends to be specialized in function, in their cellular localization or in their temporal expression and sometimes developed a new function. This interpretation fits nicely with genetic analysis conducted 1 year earlier by other yeast geneticists (Gu et al. 2003b). They distinguished singleton and duplicate genes in the yeast genome. When they created null mutants for these genes, they observed that duplicate genes had a significantly lower proportion of genes with a lethal effect of deletion than singleton genes (12 vs. 29%). They found, with lesser statistical support, that duplicate genes showed a lesser degree of mutual functional compensation when their degree of genetic divergence had increased. Duplicate genes were also more flexible with respect to singletons when tested through different growth conditions.

Yeast Metabolic Network

However, the authors of this study admitted that the high frequency of genes that have weak or no fitness effects of deletion call for alternative interpretations like compensation through alternative pathways or network branching. And with this hypothesis, we are back to metabolic network analysis. L. Hurst and colleagues analyzed this alternative explanation by using an *in silico* flux model of the yeast metabolic network (Papp et al. 2004). The model included 809 metabolites as nodes, which are connected by 851 different biochemical reactions when transport processes of external metabolites were included. Their analysis indicated that seemingly dispensable genes might be important, but only under conditions not yet examined in the laboratory. This was their dominant explanation for apparent dispensability accounting for a maximum of 68% of the dispensable genes. Compensation by duplicate genes accounted for at most 28% of these events, while buffering by metabolic network flux reorganization

was the least important process. In fact, they observed that the yeast metabolic network had difficulties in tolerating large flux reorganizations. Isoenzymes were selected in order to enhance gene dosage, which then provides higher enzymatic flux instead of maintaining alternative pathways. These data have important implications for our perception of genome organization. Apparently, despite its ancient duplication event, yeast is not a peculiar case. Other model organisms like *E. coli* (Gerdes et al. 2003), *Bacillus subtilis* (Kobayashi et al. 2003), and *Caenorhabditis elegans* (Kamath et al. 2003) yielded only 7–19% of essential genes for laboratory growth. This contrasts well with the high percentage of the essential genes in the *Mycoplasma genitalium* genome of up to 73% (Hutchison et al. 1999). In this series, *Mycoplasma* is clearly the odd man out. Its genome is apparently the result of secondary genome reduction occurring in an intracellular parasite with strict host and tissue specificity. This parasite has given up most of its “dispensable” genes simply because it has given up the idea to conquer alternative habitats beyond its super-specialized niche.

Lateral Transfers into Networks

The situation is different for *E. coli*. Against popular belief, its niche is not the genetics and molecular biology laboratory, but the intestine of mammals and birds. This is a complicated and highly competitive microbial community in a dynamic physiological context where the host takes precautions to expel the gut commensals. Once out of this niche, *E. coli* must survive in the environment before it reaches the next gut; recent data suggest that it does so in soil. *E. coli* became a successful pathogen in mammals and birds, small wonder that different *E. coli* strains differ by 0.5–1 Mb in their genome content. Most of these gene acquisitions were by horizontal gene transfer. The question is now how metabolic networks incorporate new genes acquired during adaptive evolution in bacteria. Gene duplication, the main source of evolutionary novelty in eukaryotes, plays only a minor role in *E. coli*. *E. coli* K-12 contains few duplicated enzymes in its metabolic network, and all but one seem to be ancient (Pal et al. 2005). Most changes to the metabolic network of *E. coli*, which occurred in the past 100 million years since its divergence from *Salmonella*, are due to horizontal gene transfer (Figure 2.3). Only 7% of the genes that are horizontally transferred into the metabolic network of *E. coli* are essential under nutrient-rich laboratory growth conditions. Genes that contributed most to the evolution of the metabolic network and thus to the evolution of proteobacteria were generally environment specific. The flux balance analysis of these authors also explains why most of these horizontally transferred genes are not expressed under laboratory conditions. They concluded that the evolution of networks is largely driven by adaptation to new environments and not by optimization in fixed environments. Proteins contributing to peripheral reactions (nutrient uptake, first metabolic step) were more likely to be transferred than enzymes catalyzing intermediate steps and biomass production.



FIGURE 2.3. *Salmonella typhimurium*, a close relative of the gut bacterium *Escherichia coli*, is the prototype of Gram-negative bacteria belonging to the Proteobacteria group. They resemble each other morphologically, but differ in their pathogenic potential.

Salmonella in Mice

It is a laboratory abstraction to analyze the metabolic network of a microorganism growing as an isolated culture in a broth medium. In nature most microorganisms occur in association where metabolites are exchanged between different bacterial species. In the case of commensals or pathogens, bacteria also acquire metabolites from the eukaryotic host. Metabolic network reconstruction for an isolated microbe thus gives only an incomplete picture. This was recently realized by German scientists who tried to identify targets for new antimicrobials derived from a model of in vivo *Salmonella* metabolism (Becker et al. 2006). They used two mouse *Salmonella* infection types, the typhoid fever and the enteritis model. Proteome analysis identified 228 and 539 metabolic enzymes in the spleen and the cecum from the typhoid and enteritis model, respectively. The *Salmonella* genome probably encodes 2,200 proteins with putative metabolic functions. The proteome analysis underestimated the true number of expressed proteins due to the technical challenge to differentiate minor bacterial proteins against a background of high host proteins. However, a large discrepancy remains between both figure sets, which is explained by the authors with two interpretations. First, the metabolic network of *Salmonella* shows extensive metabolic redundancies reducing the number of truly essential bacterial genes. Second, *Salmonella* has apparently an access to a surprisingly diverse pool of host nutrients. Auxotrophic mutants of *Salmonella*, which could not grow without supplements in broth culture, retained substantial virulence in the typhoid fever model. Apparently *Salmonella*, which resides in the supposed

nutrient-poor murine macrophage phagosome, still finds access to various amino acids, purine and pyrimidine nucleosides, glycerol, sialic acid, hexoses, pentoses, vitamins, and electron acceptors for both aerobic and anaerobic catabolism. With this section, we get to the next level of complexity where metabolic networks of interacting organisms must be described to achieve a realistic model.

De Revolutionibus Orbium Metabolorum

Revolutionary Histories

The introduction of the heliocentric system into astronomy is hailed as the Copernican revolution and the section heading is paraphrasing the title of the famous book of this Polish canonicus. Actually our understanding of the word “revolution” has changed over time. For people like Nicolaus Copernicus, Latin was their scientific language as scientists are nowadays speaking English. In Latin, “revolution” has a direct physical meaning, “revolvere,” from which it is derived, means simply to rotate. However, the idea that the earth is revolving around the sun was such a rotating and revolutionary idea at his time that Nicolaus Copernicus preferred—against all instincts of modern scientists—to have his work only published on his deathbed. The concepts underlying oxidative phosphorylation—the subject of our next section—are also revolutionary for modern biology in the double sense. They taught us how energy is mechanistically extracted from foodstuff, and this discovery revolutionized biochemistry. In addition, two rotating biochemical devices stand at the beginning and at the end of the process. On the start, this is the Krebs or citric acid cycle revolving intermediates that are organized into a cyclic pathway. At the end is a small molecular motor that mechanically rotates when fuelled by a proton gradient and which thereby synthesizes ATP. This process of oxygenic respiration is true molecular biology, i.e., the description of fundamental biological processes at a molecular level. If one considers the electron transport chain in the inner mitochondrial membrane, part of the description is even at the subatomic level. Basic life processes can now be understood in chemical detail: We understand now where the oxygen we breathe is used in our metabolism, where CO₂ we exhale is created, and where and how the 40kg of ATP is created that our metabolism has to produce day by day to power our life processes. The site of all the above-mentioned activities is the mitochondrion. We are dead within minutes when the blood circulation or lung respiration stops, reflecting the fact that more than 90% of our ATP derived from foodstuff is gained from oxidative phosphorylation. Much of the anatomy and physiology of animals is explained by the biochemical needs of mitochondria. We are thus leaving the cytoplasm, the site of the glycolytic enzymes. This change of place of action from the cytoplasm to a cell organelle means also the crossing of a watershed.

Mitochondria as Bacterial Endosymbionts

Endosymbiont Hypothesis

Mitochondria resemble prokaryotes in numerous important properties. Like the bacterium *E. coli*, mitochondria are enveloped by two membranes. The inner membrane, which is much larger in area than the outer membrane, folds into the matrix building the so-called cristae. The matrix contains soluble enzymes, including those of the citric acid cycle, the β -oxidation of fatty acids, prokaryotic type ribosomes, tRNAs, and a small circle of DNA. Nowadays only a smaller part of the mitochondrial proteins are encoded on this mitochondrial DNA chromosome. The inner membrane contains complexes I–IV of the electron transport chain and the ATP synthase also called complex V. The mitochondrion is wrapped by a second membrane, creating an intermembrane space corresponding to the periplasmic space of bacteria. Like bacteria, mitochondria reproduce by fission. All these similarities make the conclusion inescapable that mitochondria were once eubacteria that invaded the ancestor of the modern cells.

Proof for this endosymbiotic origin of mitochondria came from the genome of mitochondria. At first glance, this seems not to be an obvious source of information: The mitochondrial genomes are linear or circular and range in size from <6kb (kilobase, 1,000 bp) in the malaria parasite to 360 kb in the model plant *Arabidopsis*, the water cress. The latter encodes despite its large size only 32 proteins; >80% of its DNA is actually noncoding. Protozoa at the basis of the eukaryotic tree were considered as the most interesting study objects for the origin of mitochondria since some of them lack mitochondria (more on that subject in a later section). Therefore many mitochondrial genomes from protozoa were sequenced. The breakthrough was achieved in 1997 when a *Nature* report entitled “An Ancestral Mitochondrial DNA Resembling a Eubacterial Genome in Miniature” was published. This genome belonged to a free-living freshwater protist called *Reclinomonas americana* (Lang et al. 1997). Its 69 kb genome encodes 97 genes. Protein-coding genes contributed to complexes I–IV of the respiratory chain, the ATP synthase, the mitochondrial-specific translation apparatus (ribosomal proteins), and most spectacularly a eubacterial RNA polymerase. The mitochondria from higher eukaryotes use strangely a phage-like RNA polymerase of totally unknown origin. The informational content of this mitochondrial genome is greater than that of the sum of all other sequenced mitochondrial genomes (Gray et al. 1999). Thus it is not only the most bacteria-like genome, but also the closest relative to the ancestral protomitochondrial genome.

Rickettsia

One year later, another genome sequence was published in *Nature* and this is now the most mitochondria-like genome within eubacteria (Andersson et al. 1998; Gray 1998). If one considers that the evolution of animals could not

have occurred without the capture of this powerhouse of the cellular energy metabolism, the closest bacterial relative of mitochondria is an unlikely character. It is *Rickettsia prowazekii*, the causative agent of epidemic, louse-borne typhus. *This is a dreadful human disease, which wrote many chapters in human military history. Pericles died from it in Athens at 430 BC during the Peloponnesian War; it caused the retreat of Napoleon's troops from Russia; and infected 30 millions and killed three million people mainly in the Soviet Union in the wake of the First World War. If you are dissatisfied with the tongue-twisting nature of many species names in biology, you should do justice to that of this microbe since it honors the names of a US pathologist and a Czech microbiologist, who both died when investigating this pathogen.*

Rickettsia's life cycle belongs to the annals of the history of eating, so close is its relationship with the subject of this book. *Rickettsia* are transmitted by lice. When lice feed on the blood of a human infected with rickettsias, the bacteria infect the gut of the insect, where they multiply and are excreted in the feces a week later. If the insect then takes a blood meal on an uninfected person, it defecates on the skin. This irritation causes the person to scratch and thereby the bite wound is contaminated by rickettsias. The bacterium spreads with the bloodstream and then infects the endothelial cells of the blood vessel, causing all the clinical sequels, the rash and the high mortality of 50% if untreated. This does not sound like a close relative of the energy-producing machine of the cell. Yet its genome sequence reveals the parallels. *Rickettsia* enters the host cell, escapes the phagosome and multiplies in the cytoplasm of the eukaryotic cell. It is an intracellular bacterium, and this lifestyle caused a dramatic reshuffling of its genome and of its metabolism. The amino acid metabolism has practically disappeared, so did the nucleotide biosynthesis pathway. Small wonder since these metabolites and ATP are amply provided by the host cell. With its 1,100 kb size, the genome of rickettsia is small, yet other intracellular bacteria went even further with their genome reduction as demonstrated by *Mycoplasma* with only 470 kb large genomes. Selection could not maintain nonessential genes, and the process of gene loss is apparently still ongoing as demonstrated by the unusual 24% of noncoding DNA and frequent point mutations in the rickettsia genome. However, rickettsia maintained their capacity to produce energy. Later in the infection process, when the cells get depleted of cytosolic ATP, the intracellular pathogen switches to its own energy system. Like in mitochondria, pyruvate is imported into the bacterium, converted by pyruvate dehydrogenase (PDH) into acetyl-CoA. The TCA cycle enzymes are present in rickettsia as well as the respiratory chain and the ATP synthesizing complex. However, the glycolytic pathway is conspicuously absent. Suddenly, although in pathological disguise, we have here a metabolic parallel to mitochondria. Phylogenetic analysis of the cytochrome *c* oxidase gene indicated that the respiratory systems of rickettsia and mitochondria diverged about 1.5–2 Gy (giga years, one billion years) ago, shortly after the amount of oxygen in the atmosphere began to rise. This is

the most appropriate moment for the evolution of oxygen-based respiratory systems. However, mitochondria are not derived from rickettsia. Instead both genomes derive from an α -Proteobacterium ancestor and both lines followed an independent path of genome reduction. Mitochondria went much farther down this way. In fact, many mitochondrial proteins are now encoded by genes located in the chromosomes of the nucleus. The current endosymbiont theory envisions a successive transfer of genes from the protomitochondrial genome to the host genome. This transfer stabilizes the endosymbiotic relationship since mitochondria that lost essential parts of their gene content can no longer escape from this close metabolic relationship between the ex-bacterium and its eukaryotic host. *Actually, mitochondria became slaves of the modern cells and the successive loss of genetic autonomy can be read from the decreasing size and information content of mitochondrial DNA in animals. However, when we explore the energy metabolism of mitochondria in the following chapters, we should not forget that we look into a basic blueprint invented by α -Proteobacteria that was only perfected in the extant mitochondria to suit the needs of the eukaryotic master.*

Pyruvate Dehydrogenase: The Linker Between Pathways

Haeckel's Principle

Glycolysis is confined to the cytosol, at least as far as animals are concerned. The Krebs or citric acid cycle in contrast takes place in the matrix of the mitochondrion. One might suspect that this location in different cellular compartments might be a practical solution to separate pathways, intermediates, and metabolic fluxes allowing their separate regulation. This is certainly true; however, we should not forget that biochemistry is very conservative, and we might see here still the metabolic endowments of the different cells that made up the eukaryotic cell. Glycolysis was the energy-providing process of the ancestor cell providing the cytoplasm of the eukaryotic cell, whereas the citric acid cycle was the heritage of the α -Proteobacteria leading to mitochondria. This conservative nature of biochemistry becomes even more evident when two pathways leading to a comparable end product are maintained in two cellular locations. Take the synthesis of membrane lipids. In plants, chloroplasts, the home of the photosynthetic apparatus, synthesize membrane lipids in a typical prokaryotic pathway. Apparently, chloroplasts in Ernst Haeckel's words recall phylogeny and remember that they were derived from bacteria, in their case cyanobacteria. In parallel, plant cells synthesize membrane lipids along a typical eukaryotic pathway in the endoplasmic reticulum. Membrane lipid synthesis in plants still reflects the distinct origin of the cellular compartments. This separation is maintained even if it obliges modern plants to a collaboration between different cellular compartments. We should thus be cautious when arguing from teleological principles in biology, history might be an equally strong argument.

Pyruvate

Back to pyruvate: If the cell wants the full pay for its glucose food, it has to change the compartment. Pyruvate has to cross the two membranes of the mitochondrion to reach the place of its complete oxidation to CO_2 and H_2O . The outer mitochondrial membrane is not a major barrier to its transport, but the inner membrane is selectively permeable—and for good reason as we will hear soon. Gases and water move rapidly across this membrane; small uncharged molecules like protonated acetic acid achieve this too. However, pyruvate is charged under physiological conditions and needs therefore a carrier to cross the membrane. Pyruvate gets into the mitochondrial matrix by an antiporter: It takes pyruvate in and OH^- out. It is thus an electroneutral transport because no net charges are moved across the membrane. The transport is driven by the pH gradient across the membrane created during electron transport in the respiratory chain. In the matrix, pyruvate is received by PDH. This is a remarkable enzyme for several reasons: It is so large that its structure can be visualized by cryoelectron microscopy (Gu et al. 2003a). In fact, it is five times larger than ribosomes, the cellular protein synthesis machines. The inner core consists of 60 molecules of enzyme E2. This enzyme contains a long flexible linker molecule, made from lipoic acid fixed to a lysine side chain. The active components are two thiol groups that can undergo reversible oxidation to a disulfide bond, leading to a five-membered heterocyclic ring. The reduced linker swings to the E3 enzyme, which is also located in the core of the multienzyme complex with 12 copies. The task of this enzyme is the reoxidation of the reduced linker dihydrolipoyl with FAD, hence its name dihydrolipoyl dehydrogenase. However, NAD^+ is the ultimate electron acceptor in this oxidation reaction, the reduced NADH is fed into the electron transport chain. Then the reoxidized disulfide linker swings back from E3 and touches the outer rim of the complex consisting of a shell of many E1 enzymes. The lipoate linker assures substrate channeling: the five-reaction sequence on the enzyme complex never releases the intermediates. This measure prevents other pathways keen on this central metabolite from stealing it from PDH.

Pyruvate Dehydrogenase

The outer shell made of E1 actually takes care of the pyruvate. E1 is called a PDH, but the chemical reaction it mediates is actually an oxidative decarboxylation. The carboxyl group of pyruvate leaves as CO_2 ; this is the first carbon from glucose leaving our body as a gas in our breath. We should think on this brave enzyme when we exchange oxygen from the air surrounding us and trade it against CO_2 in our expiration. The task is actually not simple: E1 needs the cofactor thiamine pyrophosphate for this task. As many cofactors of enzymes it has a somewhat complicated structure and our body does not take the pain to synthesize it. However, this cellular economy comes at a price, we have to take it up from our food where it is better known as vitamin B1. The active group in this reaction is the C-2 carbon from the central thiazolium ring with

a relative acidic proton. If this proton dissociates, a carbanion (a negatively charged carbon) is created, which is a powerful nucleophil, a chemical substance searching for a positively charged compound. It finds its partner in the partially positive carbonyl group of pyruvate. The positively charged nitrogen atom of the thiazolium ring acts now as an electron sink. The traveling of the electrons leads to the leaving of the carboxyl group as CO_2 . In pyruvate decarboxylation, we now get a hydroxyethylgroup at thiamine, which leaves as acetaldehyde. However, in PDH, a different path is followed: The hydroxyethyl group is oxidized to the level of a carboxylic acid creating acetate. The two electrons, which are extracted from the substrate, are taken up by the disulfide group of the flexible E2 linker. The acetyl group created by this oxidation–reduction reaction is first esterified to the thiol group of the linker and then transferred to the thiol group of coenzyme A. Now we have the acetyl-CoA, the fuel of the citric acid cycle. The acetyl group is linked in a high-energy thioester group, which actually conserves the energy of the oxidation reaction. We understand now also why E2 is called a transacetylase, more specifically a dihydrolipoyl transacetylase.

PDH is also an important site of regulation: The enzyme complex is inhibited by ATP, acetyl-CoA, NADH, and fatty acids. This is a logical design, all these compounds indicate that fuel supply is high and the energy charge of the cell is also high. There is thus no need to feed further pyruvate into the catabolic pathway. Equally logical is the activation of the PDH by AMP, CoA, NAD^+ , and Ca^{2+} . These compounds signal low energy and food charge. Calcium is released from contracting muscles: Energy is thus urgently needed and PDH must channel pyruvate into the citric acid cycle.

Deficiency of Cofactor and Enzyme

PDH is a remarkable enzyme not only for its size but also because it needs five different coenzymes. Four cannot be synthesized by our body and are thus vitamins: thiamine, riboflavin (in FAD), niacin (in NAD), and pantothenate (in CoA). Pantothenate deficiency is very rare in humans and has only been observed in prisoners of war suffering from severe malnutrition. The symptoms were a strange numbness in the feet. More severe and more prevalent is actually thiamine deficiency. Polished rice lacks thiamine, which is mainly found in the removed hulls of the rice. In populations that live mainly from rice, thiamine deficiency is known as beriberi. There are different forms of it, a cardiac form in infants, and dry beriberi in adults with peripheral neuropathy. Also alcoholics show a form of thiamine deficiency known as Wernicke–Korsakoff syndrome, which is linked to neurological symptoms. The problem with alcohol is that it represents what is called in nutrition “empty calories.” Food comes here in a relatively pure chemical form of ethanol and is only minimally “contaminated” by vitamins as in conventional food items.

I now want to mention a fortunately very rare human condition, PDH deficiency. Human diseases that affect tissues with high-energy requirements are often caused by defects in the mitochondrial functions. Mitochondrial dysfunctions are also discussed as causes of type-2 diabetes (Lowell and Shulman 2005).

PDH deficiency affects mainly the central nervous system: Key symptoms are developmental delay, feeding difficulties, lethargy, ataxia, and blindness. Early death is the inevitable outcome. The chemical signs can directly be understood from the biochemistry. The decreased production of acetyl-CoA results in reduced energy production. Pyruvate and lactate accumulate in the body and lead to metabolic acidosis. With the block in the link from glycolysis to the citric acid cycle, why is the energy production down? There are in fact other feeder pathways into the TCA cycle. A major alternative is the β -oxidation of fatty acids. Fatty acids linked via their carboxyl group to CoA in the cytosol can be *trans*-esterified to carnitine in the outer mitochondrial membrane (or the intermembrane space, the biochemical details have not yet been settled), and they get there via a carnitine transporter in the inner mitochondrial membrane into the matrix. Here a second carnitine acyltransferase transfers the fatty acid back to CoA. The next steps are straightforward: Four stereotype successive enzyme reactions release electrons and acetyl-CoA in multiple rounds from the imported fatty acids. In the first step, a dehydrogenase abstracts electrons from the fatty acid and introduces a double bond between the α - and β -carbon atom, hence the name β -oxidation. In the second step, a hydratase adds water to the double bond. The resulting hydroxyl group at the β -carbon position is oxidized by another dehydrogenase to a keto-group. In the last step, a free CoA group attacks the bond between the α and β carbon and splits it by a thiolysis. This means that at every turn of the cycle, again, an activated, but shortened acyl-CoA group is created. The reduced FADH_2 and NADH feed their electrons into the respiratory chain and the split acetyl-CoA enters the TCA cycle. Why should then energy production be a problem? High energy demanding tissues like the heart are actually covering their energy needs nearly exclusively by β -oxidation of fatty acids followed by oxidative respiration. In fact, infants with PDH deficiency do not show cardiac insufficiency; they suffer mainly from neurological symptoms. That neurons have a high-energy demand is not a sufficient explanation. However, in contrast to the heart, the brain lives nearly exclusively from glucose as carbon source, fatty acids can thus not replace the glucose in the brain mitochondria.

On the Value of Mutants

Pediatricians have then searched solutions to this problem in nutritional biochemistry literature. Actually, during starvation, glucose becomes limiting because the body can only store a limited amount of glucose as glycogen. As the brain is an absolutely vital function for survival, we can of course not give up this organ simply because it can only use glucose as carbon fuel. In fact, nature knows about this problem and after a few days of starving, the brain learns to use ketone bodies as an alternative fuel. Ketone bodies are produced in the liver essentially by the condensation of two acetyl-CoA molecules leading to compounds like aceton, acetoacetate, or hydroxybutyrate. However, ketone diets have not yielded the remedy expected from textbook biochemistry knowledge (Wexler

et al. 1997). *What you do when your biochemistry model fails is to develop an animal model of the disease and study what went wrong. This was actually done in the case of PDH deficiency. The investigated mutant is called noa. The names of mutants are left to the discretion of the researchers. Some of them are fanciful like “bobbed” (a fly with short bristles linked to a female short hair dress in the 1920s), others descriptive as “krüppel” or “fushi tarazu” but only understandable for parts of the scientific community due to language barriers.* “Noa” follows a tradition in microbiology where the mutant names tend to be acronyms: “noa” in full reads *no optokinetic response a*. This mutant is in the zebrafish, a popular pet animal of geneticists. The major advantage of this animal is the fact that biologists can study the living mutant animal under the microscope. However, the observation of *noa* does not necessitate a microscope. The zebrafish shows expanded melanophores (pigmented cells), no feeding behavior, lethargy and premature death. Noa has a defective E2 subunit of PDH (Taylor et al. 2004). The phenotype is relatively easy to understand: Like children, the fish larvae show elevated levels of pyruvate and lactate. Interestingly, the ATP/ADP ratio is normal despite the lower energy production (the TCA cycle is blocked). The ratio can only be maintained because the ATP consumption is decreased. A similar regulation is done by fish in anoxic water: It reduces swimming activity and visual function. In fact, photoreceptors belong to the most energy-demanding cells of the vertebrate body. As the receptors do not see light under these conditions, the brain gets the information of darkness and the fish adapts to this misperceived night by a darkening of its skin via the expansion of the melanine deposits. Now comes the ketogenic diet and with it an arousal of the lethargic animals: They start to swim and eat (their aquarium delicacy are paramecia, a protist known to all children getting their first light microscope). The retina resumes its activity; the fish larvae get brighter again, survive, and resume growth. However, they show increasing growth retardation and mimic thus the inefficiency of the ketogenic diet already observed in infants. This is now an important message: Our understanding of the vertebrate metabolism is not yet so developed that we could easily do nutritional engineering in humans. However, the stakes are set, and there are many pathological situations where nutritional interventions would be a highly wanted addition to the medical toolbox.

Why is the Citric Acid Cycle so Complicated?

Principles

In colloquial physiological speech, we speak of burning the food when we extract energy out of it. If you burn sucrose chemically, you get a lot of energy because this is a strongly exergonic reaction, but you get it as heat. Heat is a relatively worthless form of energy because you cannot convert heat into other forms of energy. And organisms have multiple tasks to perform, and all need to be powered by an energy input. The pervading principle is to do small, but controlled steps downhill the overall exergonic reaction pathway of food burning

and to conserve the redox energy in energy forms usable for the cell. The principle of the small chemical steps of energy extraction from food is realized in the citric acid cycle. Actually the most striking aspect of this pathway is that it is not linear like glycolysis, starting with one compound and ending up with another. It is a truly cyclic process.

The Chemical Steps

A four-carbon compound, oxaloacetate, condenses with the two-carbon compound acetyl-CoA, produced by PDH to give a six-carbon compound, citrate. After a dehydration step, a hydration step follows, and citrate is nearly reconstituted in isocitrate. *Here you should protest: If you were taught that nature uses the most economical of all possible solutions, then you would not expect that the oxidative degradation of acetyl-CoA takes the detour to first lengthen the molecule to a six-carbon compound and then doing the illogical steps of first a dehydration and then its inverse, namely a hydration.* The reaction is catalyzed by the enzyme aconitase. Notably, an aconitase is also found in the cytoplasm, but here its function is not in the citric acid cycle (there is none in the cytoplasm), but here it is an iron-responsive element. It is a bifunctional enzyme: It can catalyze the citrate to isocitrate reaction, but it can also bind specific mRNAs and interfere with their translation. As iron metabolism was most likely an earlier activity than oxygenic respiration, we might suspect that aconitase is a late recruit to the citric cycle, but this is a pure speculation. Back to the cycle: Next follow two successive decarboxylation steps. Decarboxylations are very popular steps in biochemistry; they often drive reactions that would otherwise be highly endergonic. Two molecules of CO_2 leave the cycle and counterbalance the addition of the two carbon atoms introduced in the acetyl group of acetyl-CoA. Notably, the two carbon atoms that leave the cycle are not those that have just entered the cycle with acetyl-CoA. It is of central importance to the following steps of energy conservation that a hydroxyl group in isocitrate is oxidized to the ketone oxidation level (α -ketoglutarate). The abstracted electrons are recovered by higher organisms in one molecule of NADH; in microbes, the acceptor is always NADPH. Then follows a nearly exact copy of the reaction catalyzed by the PDH complex, the E3 subunit is in fact identical between both enzymes. In contrast, the E1 and E2 subunits differ because they must display a distinct binding specificity, but the cofactors and the reaction mechanisms are absolutely identical. Here we see again the modular organization of the metabolism. PDH and α -ketoglutarate dehydrogenase certainly derive from a common ancestor (and related enzymes are also found in amino acid degradation pathways). The energy gained by oxidation is conserved in the energy-rich thioester bond (succinyl-CoA). Then comes a difference to acetyl-CoA. The thioester bond in acetyl-CoA is used to drive the condensation of oxaloacetate with acetyl-CoA to citrate, while the energy in the succinyl-CoA bond is recovered in the anhydride bond of GTP. The remaining reactions of the citric acid cycle follow two goals: First, the extraction of further electrons from the food molecules, and second, the reconstruction of oxaloacetate, the starting

compound of the cycle. The central C–C bond in succinate is oxidized to a C=C bond in fumarate, the abstracted electrons create FADH_2 . Water is added to the double bond yielding malate. The hydroxyl group in malate is oxidized to a keto-group, which creates another NADH and most importantly leads again to oxaloacetate. The cycle is closed and oxaloacetate is ready to fuse again with acetyl-CoA to restart the cycle. In principle only catalytic amounts of intermediates of the citric acid cycle are needed, and the steady-state concentrations of oxaloacetate are definitively extremely low ($<10^{-6}$ M). Here we are back again.

Alternatives?

As a critical reader you might ask whether this cycle is not unduely complicated. Is it not possible to oxidize pyruvate directly? The answer is yes, of course. Chemically, these are two successive decarboxylations: first to formate (HCOOH) and then to molecular hydrogen H_2 . On paper this reaction looks quite simple, but the enzyme catalyzing this reaction, pyruvate-formate lyase is strictly anaerobic, the reaction doesn't work in the presence of oxygen. This is an important restriction because it prevents cells like *E. coli* from using this pathway in the presence of oxygen where it can metabolize glucose via glycolysis and the citric acid cycle. This gives obviously the most energy from glucose. If oxygen is lacking, but an alternative electron acceptor like nitrate is present, *E. coli* uses first nitrate respiration. In the absence of an electron acceptor, *E. coli* changes to fermentation pathways. In fact, *E. coli* has different options, which testify its metabolic versatility. In one it transforms pyruvate into acetyl-CoA, which is then degraded to ethanol by two reduction steps and, not oxidation steps. Actually you sacrifice reducing equivalents (NADH), and the cell has to excrete an energy-rich two-carbon compound, ethanol. This is, by the way, the reason that absolute abstinent people still have a detectable blood alcohol level produced in the gut by *E. coli*.

Beside other options (e.g., excreting acetate), *E. coli* has the above-mentioned oxygen-sensitive pyruvate-formate lyase that catalyzes the first decarboxylation. However, this enzymatic reaction is not simple at all: this enzyme is under a complicated network of transcriptional control (Fnr, NarL, ArcA repressors/activators) and posttranscriptionally regulated by an activase (Act) and a deactivase (AdhE). The explication of this control web would lead too far, but its essence is the control for the most efficient glucose use under different metabolic conditions. Then it needs a partner: the decarboxylation of formate, i.e., its splitting into CO_2 and H_2 , is mediated by formate hydrogen lyase. Again this is a complicated enzyme. Formate-hydrogen lyase is a multicomponent membrane-associated complex, and at least 12 genes contributed by two operons (*hpc* and *hpy*) are involved. The enzyme complex requires a molybdenum cofactor, Ni and Fe. It requires in addition that an internal stop codon be read by an unusual tRNA that recognizes the stop codon as a signal for the insertion of a selenocysteine. The expression of the enzyme is also under transcriptional control. It depends on the alternative sigma factor σ^{54} that changes the promoter recognition of the RNA polymerase.

So a possible answer is this: The citric acid cycle is in fact not so complicated at all; the seemingly simpler chemical pathway of two successive decarboxylations is with respect to the enzymes probably even more complicated. At the end, the citric acid cycle uses only nine enzymes. In addition, strictly anaerobic enzymes lost a lot of importance when the atmosphere of the earth got increasingly rich in oxygen. Furthermore, the TCA cycle took over a number of biosynthetic service function and became thus the hub of the central metabolism of higher organisms.

The Horseshoe TCA Pathway

E. coli's Problem

Like glycolysis, the TCA cycle comes in many variants. What at first glance appears as a complicated cycle is in fact a simple and malleable device. I will present an interesting variant of the TCA cycle. It describes how it turns in the absence of oxygen as electron acceptor. If higher organisms use the TCA cycle in the catabolic mode, it is linked to energy production in the respiratory chain using molecular oxygen as electron acceptor. In fact, even if *E. coli* prefers oxygen for growth, under physiological conditions, i.e., in its gut ecological niche, *E. coli* does not see much oxygen. Except for a few niches in tropical ecosystems offering high nutrient concentrations and high temperature, *E. coli* is not known to grow in freshwater of temperate ecosystems (Winfield and Groisman 2003). This inability of *E. coli* to grow in freshwater is actually the underlying logic to use *E. coli* as an indicator of fecal contamination in recreational water. This poses a dilemma for *E. coli* when it gets expelled from the gut by peristalsis and defecation and has to find a new host before it starves to death. *Some freshwater microbiologists seem to suggest that a residual growth of E. coli is also observed in freshwater from temperate ecosystems, but I have not found published reports supporting this claim. In fact, many crucial questions in the ecology of E. coli are still seriously under-investigated.*

Problems with the Reductionist Principle

There are two reasons for this ignorance. One is a research principle called reductionism. This working principle states that a biological process should be studied under the simplest and most standardized conditions and many laboratories should work with the same standardized organisms to create maximal synergism. This principle was introduced into biology by physicists that went in the 1940s into biology to tackle the question of Erwin Schrödinger, "What is life?" by using the approaches of physics. Schrödinger raised only the question, but the next generation of physicists like Max Delbrück actually worked with E. coli and its phages to address this question with concrete experiments. Their work was a scientific bombshell and led to a new discipline, molecular biology, which made biology for the first time an exact science. The success of this approach was dramatic. In the nonspecialized research journals, the biological

sciences overshadowed over the last decades all other natural science branches combined. But also triumphs have their price. The extremely successful reductionist approach with E. coli discouraged researchers to ask questions like, What is the metabolism of E. coli in the gut? Where does it grow actually in the gut? E. coli is found in the gut lumen, but here it is apparently starving and has a very long generation time.

E. coli's Solution

If *E. coli* does not grow outside of the intestine (pathological conditions set aside), why is then the TCA cycle maintained which needs a respiratory chain? We know that *E. coli's* growing fraction in the gut is not found in the gut lumen, but as small microcolonies within the mucus overlaying the gut epithelia. *It is possible that the mucus-associated E. coli microcolonies still capture enough oxygen from the blood vessels of the gut mucosa. However, the metabolic fluxes in these E. coli microcolonies have not yet been investigated.* Researchers have studied the metabolism of *E. coli* grown in vitro under anoxic conditions with intestinal mucins as sole carbon source. The microarrays showed that under these conditions the enzymes of the citric acid cycle are not active, while the glycolytic pathway is fully activated (Chang et al. 2004). However, the TCA cycle is not totally down under anoxic conditions when tested under in vitro conditions: Most enzymes work at 5–20% of their aerobic (aerobic and oxic is largely synonymous as is anoxic and anaerobic) activity level. There is only one enzyme of the cycle, which is really nonfunctional: This is the ketoglutarate dehydrogenase complex, which catalyzes the reaction from ketoglutarate to succinyl-CoA. This enzyme is the cousin of the PDH complex discussed above. Notably, the two decarboxylation reactions in the TCA cycle are essentially irreversible reactions, while the rest of the cycle except for the citrate synthase reaction are perfectly reversible reactions. This allows now an important reorganization of the cycle. Our chemical wizard *E. coli* does not give up the cycle, but it splits it into two arms. One branch is oxidative and leads via the normal pathway of the TCA cycle until ketoglutarate. As the reactions stop here, ketoglutarate gets a new mission. The keto-group becomes the acceptor of amino groups leading to the amino acid glutamate and thus into anabolic pathways. As the glutamate cannot take up the entire metabolic flux from glucose, the PDH activity is downregulated, too (small wonder because it is so similar to the ketoglutarate dehydrogenase (KDH) complex in its reaction mechanism) and the pyruvate-formate lyase mentioned above takes over and channels part of the glycolytic flow into various excreted products. However, since the TCA pathway is not cycling, an anaplerotic (fill-up) reaction is needed to keep the two separate arms under substrate flow. This is achieved in *E. coli* by carboxylation of PEP to oxaloacetate. The latter can now lead into the oxidative part via condensation to citrate or can feed into the reductive branch of the TCA cycle, which runs now in reverse direction to fumarate or succinyl-CoA. Recall that all these TCA reactions are fully reversible. The two metabolites play different roles: Fumarate

with its double C=C bond is a suitable electron acceptor for *E. coli* yielding succinate with the reduced C–C single bond under conditions when neither nitrate or oxygen are available as electron acceptors. Succinyl-CoA is also an important precursor to anabolic pathways, like heme synthesis.

The Horseshoe Cycle

Other bacteria have in fact permanently lost the KDH and run the TCA cycle like *E. coli* under anoxic conditions. This is the so-called horseshoe TCA pathway operated by obligate chemolithoautotrophs. The name “horseshoe” is meant in the literal sense since the open cycle printed on a paper resembles now a horseshoe. These bacteria use the TCA enzymes only for biosynthetical purposes and not for energy generation. These bacteria cannot grow on common carbon sources because they have lost the transporters for sugar import. They must fix inorganic CO₂ via the Calvin cycle and gain energy from membrane-bound cytochromes. Small surprise that they are only competitive in environments that are especially poor in organic nutrients.

Split TCA cycles are in fact common among bacteria. Out of 17 microbial genomes surveyed *in silico*, only four appeared to encode all the genes necessary for a complete, canonical TCA cycle. Lack of KDH often accompanies anaerobic or microaerophilic metabolism. In some bacteria, the TCA half-cycles are joined by alternative enzymes. In *Mycobacterium tuberculosis*, which adapted to persistence in human macrophages, the lack of KDH is apparently imposed by the need to synthesize high amounts of glutamate from ketoglutarate as a “compatible solute” and as osmoprotectant to withstand the high osmotic pressure an intracellular bacterium experiences in its host cell. The two branches of the TCA cycle are joined by succinic semialdehyde, which is synthesized by α -ketoglutarate decarboxylase (Tian et al. 2005b). As this bypass enzyme does not exist in humans, it represents an excellent target for chemotherapy of tuberculosis.

This modified TCA pathway in *E. coli* and these chemolithotrophs provide another answer to the complicated chemical design of the TCA cycle. Even in organisms that live most of their time under anoxic conditions, the TCA cycle is important since it has crucial anabolic functions. This is also the case for animals like us, which run the cycle under double mission. The TCA cycle is called amphibolic, from Greek ἀμφι, meaning “on both sides.” It serves catabolic and anabolic processes. If you draw precursors from the TCA cycle, you drain the cycle by diminishing the flow of matter through it. It becomes therefore of crucial importance to replenish the intermediates. The lowest steady-state concentrations are measured for oxaloacetate. Therefore nature has logically decided for three different replenishing reactions for oxaloacetate leaving from the glycolytic intermediates pyruvate or PEP. These reactions are called anaplerotic, from Greek ἀνα-πληρω, meaning “to fill up.” A fourth pathway practiced by plants, some invertebrates, and microorganisms (including *E. coli* and yeast) is still another way of creating oxaloacetate by condensing two acetyl-CoA into oxaloacetate via the glyoxylate cycle.

History Might Matter: An Argument on Chance and Necessity

Determinism?

In our metabolism, the TCA cycle runs exclusively clockwise in the conventional biochemical representations even if it is used both for energy-delivering (catabolic) and for substrate-providing (anabolic) reactions. This puts of course substantial constraints on the chemical design of the intermediates from such a pathway and it seems not unreasonable to anticipate that there might not be many chemical compounds that could fill in the ticket. Are there only one or a few chemical solutions to construct such a cycle that has to fulfill so many chemical constraints or in other words: Is the chemistry of the cycle deterministic?

Caveats

Before embarking on this question, I must perhaps play down our concept of the TCA cycle as the only pathway mediating the complete oxidation of carbohydrates to CO_2 . New pathways are even described for *E. coli* when you investigate it under nutritional conditions, which come close to the natural situation. Microbes typically subsist under conditions of starvation (absence of nutrients) or hunger (suboptimal supply of nutrients) in their natural environment. In laboratory media (conditions of feast with excess of glucose), *E. coli* experiences catabolite repression, and its metabolism cannot be compared to the hunger situation. The hungry *E. coli* shows a hitherto unknown pathway, called the phosphoenolpyruvate-glyoxylate cycle (Fischer and Sauer 2003). This pathway combines the glyoxylate shunt with PEP carboxykinase to oxidize PEP completely to CO_2 . Thus, under different physiological conditions, different types of metabolisms are observed with the same organism. I recall that the metabolism of *E. coli* in the gut has to my knowledge not yet been explored. There is another caveat. Metabolic flux analysis has emerged as key technology to quantify the in vivo distribution of molecular fluxes through the metabolism of model microbes with industrial relevance. When a wider range of bacteria was investigated, it turned out that the generally held view that the Embden–Meyerhof–Parnas pathway (“glycolysis”) is the major route of glucose catabolism may be a misconception (Fuhrer et al. 2005). In this study, the Entner–Doudoroff pathway was the almost exclusive route of glucose catabolism, whereas the EMP pathway was mostly absent and the pentose phosphate pathway served exclusive biosynthetic functions. If we consider, in the following, the TCA cycle as written in stone, we should keep these caveats about “textbook biochemistry” in mind.

We have seen several times that history eminently matters in biology. As I am writing a natural history of eating, the question when the TCA cycle originated is of some importance and it might also provide arguments for the historical chance versus necessity debate. Some biochemists argue that respiration is much more complicated than glycolysis and was possible only after the rise of oxygen in the atmosphere. It should therefore be a much later invention. Personally,

I do not really buy these arguments. Oxygen is not the only electron acceptor in respiration as demonstrated by many prokaryotes. There is thus no reason to wait on the arrival of oxygen, which might mean several 100 millions years after the evolution of cyanobacteria. It is furthermore not said that the TCA was invented for its current use. We have seen at several occasions that nature never discards old inventions, but uses and reuses them in a new context. That might well have been the case also for the TCA cycle.

Chlorobium

The case is made by the green sulfur bacterium *Chlorobium*. This is not an exotic bacterium; it is widely distributed and a prominent member in the cycling of sulfur in the biosphere. This organism is a photoautotrophic organism with a very peculiar organization of the photosynthetic apparatus. In many respects, this apparatus looks quite primitive and is considered by some microbiologists as a possible precursor of photosystem I (PSI). The characteristic feature of *Chlorobium* is vesicles attached to the inside of the cytoplasmic membrane. The attachment surface of the vesicle shows a crystalline baseplate structure (Figure 2.4). Apposed to the baseplate membrane of the vesicle is a type I photosynthetic reaction center associated with a bacteriochlorophyll *a* containing light-harvesting protein. The light energy is stored in a proton gradient, and this gradient is used to drive ATP synthesis in an F_1F_0 -type ATP synthase as in many other photosynthetic prokaryotes. The vesicle is also called chlorosome because it is filled with rod-like structures that consist of stacked aggregates of bacteriochlorophyll *c*, obviously not in a protein-bound form. Bacteriochlorophyll *c* molecules collect the light and channel it via bacteriochlorophyll *a*, embedded

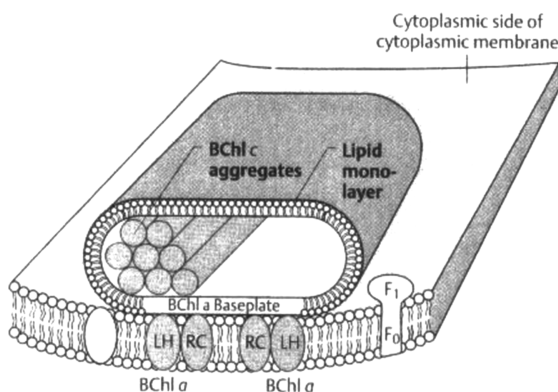


FIGURE 2.4. *Chlorobium*: chlorosome organization. Chlorosomes are attached by a proteinaceous baseplate to the cytoplasmic side of the cytoplasmic membrane. They absorb light via the linearly arrayed bacteriochlorophyll *c*, *d*, or *e*, which are arranged as rod-shaped elements in the chlorosome. The baseplate harbors bacteriochlorophyll *a* containing light-harvesting complexes LH. Below the baseplate are the reaction centers RC. F_1F_0 is the ATP synthase. (courtesy of Thieme Publisher).

in the baseplate, and then to the reaction centers located in the membrane. About 1,000 bacteriochlorophylls in the chlorosome serve a single reaction center. Later we discuss about purple bacteria that have a photosynthesis apparatus resembling photosystem II (PSII). Cyanobacteria have in contrast an already evolved photosynthesis apparatus consisting of a PSII/cytochrome b_6f /PSI like in modern plants, which probably represents a combination of the more ancient photosystems evolved in purple and green sulfur bacteria. It is thus relatively safe to conclude that the photosynthesis system in cyanobacteria evolved later than the photosynthesis system we see in *Chlorobium*.

The Reductive TCA Cycle

What does this indirect conclusion mean for our argument about the age of the TCA cycle? In fact, a lot: *Chlorobium* has a TCA cycle that runs with exactly the same intermediates like the TCA cycle in many other forms of life. As this bacterium obtains energy from its photosystem, it can use its TCA cycle for other purposes. You might suspect that it is used in this organism as a hub of the intermediary metabolism to provide precursors for anabolic pathways. In fact, despite the fact of using the same intermediates, it shows a dramatic difference with modern TCA cycles: It turns counterclockwise. We have seen that the TCA cycle releases the CO_2 from the molecules of our food that we recover as CO_2 gas in our breath. If the cycle turns in the opposite sense, it must also do the opposite with CO_2 , namely CO_2 fixation. This so-called reductive citric acid cycle is nothing exotic in microbiology. Indeed, it is one of the four basic mechanisms of CO_2 fixation in prokaryotes. Identical reductive TCA pathways were identified in sulfate-reducing (*Desulfobacter*) and Knallgas bacteria (*Hydrogenobacter*) as well as archaea (*Thermoproteus*). There are several indirect reasons that speak in favor of the antiquity of this pathway and that the TCA cycle as we know it now is only an adaptation of this old invention to the rise of oxygen in the atmosphere. This would again be an illustration that Mother Nature never rejects an old invention, but reuses them in a way like the Greek god Proteus, who was constantly changing his form with each emergence from the sea.

Let's look somewhat into the reductive TCA cycle. Despite using the same intermediates, some enzymes must be different. In order to reverse the cycle, three irreversible reactions of the oxidative TCA cycle must be catalyzed by alternative enzymes. The use of these alternative enzymes, which work with NAD(P)H and ferredoxin as reductant for the reductive carboxylations, has a remarkable consequence: The entire cycle is now fully reversible. This means that this cyclic pathway can serve both purposes in the same organism. It can mediate CO_2 fixation when powered by an independent ATP supply (the reductive TCA cycle is actually less energy devouring than the more widely distributed Calvin cycle, using only five instead of nine ATP per triose phosphate synthesized). The outlet of the reductive cycle is of course also different. ATP citrate lyase catalyzes the reversible reaction: $\text{citrate} + \text{CoASH} + \text{ATP} \leftrightarrow \text{oxaloacetate} + \text{acetyl-CoA} + \text{ADP} + \text{P}_i$. The first triose is created by a third carboxylation reaction: $\text{acetyl-CoA} + \text{CO}_2 + \text{ferredoxin red} + 2 \text{H}^+ \leftrightarrow \text{pyruvate} +$

CoASH + ferredoxin ox. Alternatively, it can serve as a pathway for the end oxidation of acetyl-CoA. The decision of the direction is made depending on the substrate supply and the energy charge of the cell.

Questions

Some biochemistry books come with a chart of the metabolic pathways. The very fact that the central metabolic pathway is formed by the connected glycolysis and TCA cycle for so many extant organisms speaks in favor of the existence of this pathway in the universal ancestor of life. Some biochemists argued that we can treat it as a virtual fossil. It could not be changed after its invention due to the high selective advantage conferred by this invention and the high interconnectivity of the central metabolism. However, its initial design—so goes the argument—was a chance event, and its conservation reflects only a historical accident. The alternative is the hypothesis that it represents an optimally successful chemical solution to designing biochemical networks and if life would be recreated under comparable environmental conditions as on the young earth, it would end up with a rather similar solution. Harold Morowitz and colleagues (2000) argued in that sense. They imagined a shell structure for the metabolism of autotrophs. In their view, the core and hence the oldest biochemical fossil is the reductive TCA cycle. The first outer shell is the synthesis of amino acids derived from amination of the ketoacids generated in the core metabolism. The second shell of the metabolic chart contains the reactions that incorporate sulfur into amino acids. The third shell reactions deal with the synthesis of dinitrogen heterocycles leading to the invention of bases and from there to nucleic acids. This onion-type view of the metabolism is in itself nothing very heretical and could even reflect the temporal order of the evolution of biochemistry on earth. The interesting argument is that in their view the inner core of the metabolism might be necessary and deterministic. Their suggestion is that any aqueous carbon-based life anywhere in our universe will resemble the intermediates of the TCA cycle. They did a bold chemical approach by applying a simple set of a priori rules. For example, they based the core chemistry on $C_xH_yO_z$ compounds with certain permitted indices, favoring for feasibility reasons small molecules. They preferred water-soluble compounds and those having low heats of combustion. While making some other selection rules based on chemical plausibility, they run the rules with the online Beilstein, an enormous encyclopedia on organic chemistry where the print Encyclopedia Britannica version is a pocket edition in comparison. From the 3.5 million entries in the Beilstein emerged 153 molecules, and hold and believe, all 11 members of the reductive citric acid cycle were in. Leslie Orgel (2000), an eminence in biochemical evolution, was not impressed by the result. He doubted that some of the rules were a priori not necessary and unknowingly selected for the intermediates they wanted to identify. In fact, is it chemically compulsory to base the selection procedure on the reactivity on the carbonyl function in $C_xH_yO_z$ compounds? Can alternative worlds not be built on boron or silicon chemistry? Evolutionary analysis with computer-generated

virtual organisms is now a subject in the biological research literature. Why should physical organisms not be based on silicon bonds?

Metabolic Crossroads in Ancient Landscapes: NAD or NADP—That's the Question

Logic of Life?

*Is there a logic of life in the biological observations or just an enormous variation of unique solutions somewhat tamed by frozen accidents? Many philosophical answers to the organization of the living might tell you more about the state of mind and temperament of the thinker than about the state of nature. Most biologists will probably lack a clear global answer to the Lucretian “De natura rerum,” but in their zeal for details they might search an answer in a particular biological phenomenon to which they can easily and without regret dedicate decades of their professional life. Psychologically, this concentration on details of a single organism chosen from millions of species makes sense only if you start from the underlying hypothesis that your observations will provide you data that are also of relevance to other researchers or even the philosophical orientation of a thoughtful layman. Practically, you have in the biological research literature a large number of specialty journals and only a handful of nonspecialized journals. However, one might suspect that this differentiation does not split the biological observations into those that apply to a single organism and those that are of wide applicability. This differentiation will reflect more the depths of the experimental approach and the brilliance of the research group, which determines the degree of generalization that you derive from your observations. In fact, the evolution theory puts at the same time large limits (at the phenotypical level where an enormous space of ecological possibilities remains to be exploited) and narrow limits (by the very meaning of descent, all organisms remain linked and thus related). The choice of model organisms like *E. coli*, the yeast, or the cress *Arabidopsis* is a logical choice only if the solution found with these individual organisms are of heuristic value for millions of other species. As biologists are generally not too keen on theoretical discussions, let's take a few examples from model organisms and you decide yourself from the experimental details what perception of nature you derive from it.*

Isocitrate Dehydrogenase

To stick to the theme of the previous sections, I have chosen isocitrate dehydrogenase, the enzyme in the TCA cycle that catalyzes the oxidative decarboxylation of isocitrate to form α -ketoglutarate. This enzyme is one of the contributors of the CO_2 in your breath. However, before the enzyme removes CO_2 from the oxalosuccinate reaction intermediate, it abstracts a hydride, which it transfers to NAD^+ or NADP^+ . Normally, enzymes do not like this ambiguity, either you work as a decent catabolic enzyme and you use then the nonphosphorylated form of the coenzyme or you are dedicated to anabolism, then you use the

phosphorylated coenzyme. We have discussed that the TCA cycle can turn in both directions, and even in the standard clockwise representation, it can fulfill catabolic and anabolic functions. Isocitrate dehydrogenase belongs to a large, ubiquitous and very ancient family of enzymes. Most family members use NAD^+ to oxidize their substrate, but not all. *E. coli*, for example, uses NADP^+ . None is able to use both. The phylogenetic family tree demonstrates that NAD^+ use is the ancient trait. NADP^+ use evolved independently several times, but this “later” use is still very old since it probably dates to the oxygenization of the atmosphere (Zhu et al. 2005). Interestingly, there is some convergent evolution in these independent events since the same suite of amino acids changes occurred in both eubacterial and archaeal lineages to adapt to the binding of NADP^+ . The geometry of the binding pocket for both NAD^+ and NADP^+ is known from high-resolution crystallographic structures, which allowed US scientists to engineer the binding specificity of the *E. coli* enzyme from NADP^+ to NAD^+ . What does this switch mean to the *E. coli* cell? To answer this question, the scientists grew the cell with the NAD^+ and the NADP enzyme, and the cell with the altered enzyme turned out to be fitter when the cells were grown on glucose. This is a rather surprising result because it would mean that the change of coenzyme specificity was maladaptive. As such a result is a violation of the most fundamental laws of Darwinism, the researchers searched further. To the reassurance of the traditional picture, the adaptive value of the NADP -enzyme form became clear when the cells were grown on acetate on which the NADP^+ -enzyme outcompeted the NAD^+ -enzyme in less than 10 generations.

Glucose Versus Acetate Food

Does this result make sense with our knowledge of biochemistry? Stated otherwise, is there logic in this design? Let’s compare glucose and acetate. Glucose is a highly reduced and thus an energy-rich compound. NADPH , the reducing power for anabolic reactions, is obtained by diverting energy-rich carbon from glycolysis into the oxidative branch of the pentose phosphate pathway. Acetate, a highly oxidized and thus an energy-poor compound, is, for example, derived from pyruvate, the end product of glycolysis. NADPH can thus not be produced from acetate via the pentose phosphate pathway. Alternative enzymes must supply this compound for anabolism. Metabolic flux analysis in *E. coli* showed that isocitrate dehydrogenase provides 90% of the NADPH for anabolism.

E. coli’s Problem

Taking this observation at face value it would mean that the ancestor of *E. coli* has not seen much glucose and had to live on less energetic food like acetate. This is probably also true for *E. coli* living today, and on theoretical reasons this must be so because any organisms must be adapted to the conditions it encounters today or to be precise, it had encountered yesterday. Since *E. coli*’s only known ecological niche is the intestine of mammals and birds, where it

populates the large intestine with modest titers, it will not see much glucose. The vertebrate body has already absorbed glucose from food into the bloodstream. The digestion of complex carbohydrate, which remains in the gut, is not the strong side of *E. coli*—it must leave this job to better carbohydrate digesters in the colon like *Bacteroides* and *Bifidobacterium*. The half-oxidized waste of other bacteria now becomes the food for *E. coli*. I must stress here that this sketch is only a probable scenario, not the result of experimental measurements. We know surprisingly little about the metabolism of *E. coli* in its natural niche, the gut. To give you a taste for the beauty of biochemistry—or the logic of life, but here I do not want to anticipate your judgment—let's go back to the Zhu paper (Zhu et al. 2005).

Glyoxylate Cycle

These authors searched the genomes of prokaryotes and found a strong correlation between strains possessing an NADP⁺-dependent isocitrate dehydrogenase and those showing an isocitrate lyase. The latter is the enzyme that diverts isocitrate from the TCA cycle into the glyoxylate cycle. A lyase is by definition an enzyme that catalyzes cleavage (or in reverse direction, additions) in which electronic arrangements occur. Isocitrate lyase cleaves the six-carbon compound isocitrate into the four-carbon compound succinate and the two-carbon compound glyoxylate. To keep the cycle running, glyoxylate condenses in two reactions with acetyl-CoA to reconstitute isocitrate in a shortcut of the TCA cycle which avoids two CO₂-releasing steps of the TCA cycle. It now becomes understandable why cells that want to grow on acetate need isocitrate lyase to provide carbon for biosynthesis and NADP⁺-dependent isocitrate dehydrogenase for the supply of reducing equivalents for anabolism. However, we now have a problem: Isocitrate can take two ways—one releasing CO₂ from organic acids in catabolism, the other sparing carbons for anabolic use. *E. coli* must regulate both ways to respond to its actual needs, but it cannot clamp down either pathway entirely when growing on acetate. The solution is as simple as appealing. The cell encodes in the *aceK* gene a kinase/phosphatase that phosphorylates the isocitrate dehydrogenase and inactivates it. During growth on acetate, about 75%, but not 100%, of the enzyme is inhibited. The kinase/phosphatase activity of the regulating enzyme is regulated by intermediates of glycolytic and TCA pathways and the energy level of the cell such that *E. coli* can now channel its metabolic flow according to its needs.

When Zhu et al. looked through the genomes of prokaryotes, the tight association between these two isocitrate-handling enzymes was independent of taxonomical group (archaea, firmicutes, proteobacteria), metabolic lifestyle, and habitat (you find there auto-, hetero-, chemo-, and lithotrophs). Proudly, the authors declared at the end of their research article that it is apparently possible to reconstruct not only what occurred, but also how it occurred and why it occurred. An ancient adaptive event that occurred billions of years ago can thus be reconstructed by a combination of genomics, protein engineering, and chemostat experiments.

From NAD to NADP

In a follow-up study, the lab of Antony Dean worked with another enzyme involved in the biosynthesis of leucine, which uses NAD^+ . They engineered six amino acids in the protein, which progressively transformed the enzyme into an NADP^+ -binding instead of an NAD^+ -binding enzyme, and they investigated the adaptive value of the mutant enzymes (Lunzer et al. 2005). The genotype—phenotype fitness map showed that NAD use is the global optimum for this enzyme. The reason can also be understood with basic knowledge in biochemistry. To perform optimally both for catabolic and anabolic pathways, cells keep NADPH concentrations higher than those of NADP^+ , while NAD^+ concentrations are maintained higher than those of NADH. The enzymatic reaction leading to leucine involves the familiar scheme of an oxidative decarboxylation, creating NADH and CO_2 as side-products. If the modified enzyme uses NADP^+ as coenzyme, it will experience product inhibition by the high cellular concentration of NADPH, which will not occur from NADH because it is kept at a low level in the cell.

The Logic and Adaptive Value of Metabolic Cycles

Plant Cycles

Life has evolved on earth. This statement sounds pretty trivial. Nevertheless, one should not overlook that the development of life on a planet that revolves around itself with a periodicity, which we call a day, puts strong constraints on life. Biological clocks have therefore evolved so that clock outputs are in phase with the Earth's rotation. Circadian clocks—by definition—synchronize biological events with the day–night cycles and they have evolved at least four times in organisms. Nowhere is the adaptive value of this clock more apparent than in photosynthetic organisms that “feed” on light. Already cyanobacteria and even more plants coordinate their metabolism with the light–dark cycle. In higher plants, there are circadian rhythms in transcript abundance of genes associated with chlorophyll synthesis and the light-harvesting apparatus. One should expect that this synchronization between light offer and metabolic demand improves the competitiveness of plants. The proof that this synchronization improves photosynthetic effectiveness was provided recently (Dodd et al. 2005). The researchers varied the light–dark periods in their experimental setting from 10–10 to 12–12 to 14–14 h, yielding artificial “days” of 20-, 24-, and 28-h duration, respectively. In this light “feeding” regime they tested wild-type *Arabidopsis*, long- and short-period clock mutants, and an arrhythmic plant overexpressing the oscillator component. The answers were clear. Leaves contained more chlorophyll when the oscillator period matched that of the environment, but did this improve photosynthesis? The answer was a clear yes: net carbon fixation increased with the matching of the rhythm and the rhythmic stomatal opening and closure played a key role. The fixed carbon was actually used for increasing the aerial biomass of the plants; the leaf area was visibly greater and the plants looked clearly

different. The enhanced growth and survival became more pronounced when the plant mutants competed with each other. Importantly, no mutant was favored under all light regimes, but only when a match between the endogenous and exogenous cycle was achieved. Circadian clock function was apparently selected during plant evolution.

Yeast Oscillations

Not all biological cycles are imposed by the 24-h rhythm of the Earth's rotation. The budding yeast *Saccharomyces* shows cycles of alternating glycolytic and respiratory activity that was described already 40 years ago. After growth to high concentration, followed by a starvation period, and then continuous supply of low-level glucose, the cells started to cycle with a rhythm of about 5 h. Biologists followed the events with expression studies. Microarray analysis was done every 25 min and revealed a periodic expression of about half of the yeast genome (3,500 genes), with a periodicity of 5 h (Tu et al. 2005). The oxidative phase, characterized by an intense burst of respiration, shows increased expression of genes involved in amino acid synthesis, RNA metabolism, and protein translation. All these processes have high energy demands that are best covered by ATP produced in abundance in the respiratory chain. The oxidative phase is followed by a reductive phase, which comes in two forms. The first is when the cell begins to cease its oxygen consumption. Genes involved in mitochondria biogenesis, DNA replication, and cell division show now peak expression levels. The second phase is characterized by nonrespiratory modes of metabolism. Protein degradation, autophagy, peroxisome, and vacuolar functions are now transcribed. Notably, cell division is confined to the nonrespiratory phase, which would allow minimizing oxidative damage to DNA. The observed fluctuations really reflect a metabolic cycle where respiration is followed by a glycolytic fermentative metabolism resulting in an accumulation of ethanol and acetate in the medium. The acetate is then charged on acetyl-CoA and prepares the cell for the next respiratory burst, consuming the metabolites accumulated in the previous period—the researchers distinguished therefore a building and a charging reductive phase. In contrast to prokaryotic cells, the eukaryotic yeast cell contains organelles. Some of them like the vacuoles experience even a visible change during the yeast metabolic cycle, the mitochondria cycle in biochemical terms. However, the yeast cell has evolved with the metabolic cycling a temporal compartmentalization of cellular processes, which are mutually exclusive. Tu and colleagues speculated that this peculiar temporal separation of cellular processes was a means of coordinating incompatible biochemical activities contributed by the two-cell types that probably fused during the birth of the eukaryotic cell, i.e., the reductive nonrespiratory cell and the oxidative respiratory cell. With respect to the logic-of-life argument, it is interesting to note that temporal compartmentalization of metabolic function might also take place during the circadian cycle of flies and mice. In fact, useful inventions tend to be made independently in nature. The cyanobacteria have learned relatively early in the evolution two lessons in biochemistry that are crucial for life on earth: oxigenic

photosynthesis and nitrogen fixation. The first develops oxygen, and the second must occur under strictly reductive conditions. *Synechococcus elongatus* has found a solution that is later reused in yeast and higher plants. These two biochemically incompatible pathways are executed at temporally distinct phases of the circadian cycle: photosynthesis in the light, nitrogen fixation in the dark (Nagoshi et al. 2004).

The Quest for Food

A Natural History of Eating

Brüssow, H.

2007, XVIII, 870 p. 207 illus., Hardcover

ISBN: 978-0-387-30334-5