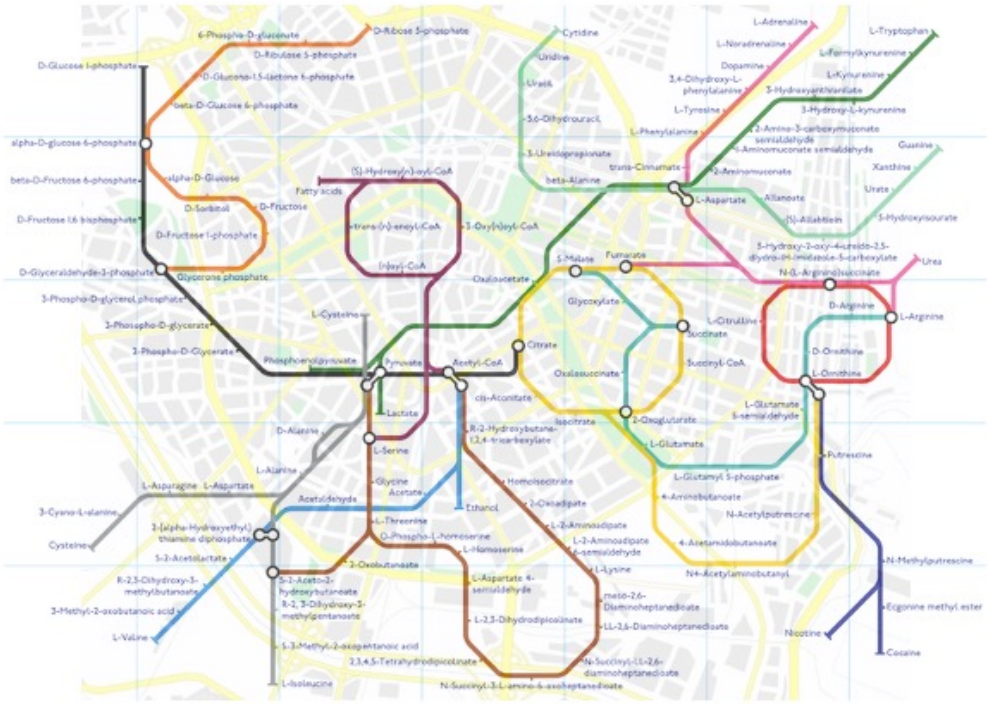
Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the required knowledge on metabolism during the course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

1

Overview of metabolism



Refresh your background

- ✓ Enzymology concepts (structure, mechanism, cofactors and coenzymes, enzymatic classes, kinetic parameters).
- ✓ Molecular mechanisms for enzyme regulation, both on activity (allosterism, covalent modification) and quantity (transcription and translation).
- ✓ Chemical functions in organic compounds.
- ✓ Microbial metabolic modes.

Before you start, you may prefer to solve the "refresh concepts quizzes" in Aula Virtual.

Keywords

- ✓ adenylate kinase
- ✓ amphibolic pathway
- ✓ anabolism
- ✓ autocatalytic cycle
- ✓ autopoiesis
- ✓ autotrophy
- ✓ bow-tie (network) structure
- ✓ catabolism
- ✓ catalytic cycle (simple)
- ✓ covalent modification (of enzymes)
- ✓ elasticity
- ✓ elementary mode
- ✓ enzyme cascade
- ✓ enzyme cooption
- ✓ enzyme promiscuity
- ✓ enzyme recruitment
- ✓ flux control coefficient
- ✓ fluxome
- ✓ genome-scale metabolic model
- ✓ heterotrophy
- ✓ isotopomer
- ✓ lithotrophy
- ✓ metabolic channeling
- ✓ Metabolic Control Analysis
- ✓ Metabolic Flux Analysis
- ✓ metabolic matrix
- ✓ metabolic mode
- ✓ metabolic pathway
- ✓ metabolome
- ✓ molecular tinkering
- ✓ natural product
- ✓ phototrophy
- ✓ primary and secondary metabolism
- ✓ recursivity or metabolic circularity (or closure)
- ✓ steady-state
- ✓ supply-demand systems
- ✓ summation property

Learning objectives

- ✓ Fundamental notions about cell metabolism, how it is organized, studied, controlled and evolved.
- ✓ Metabolism can be rationally modified, or why it is important for a biotechnologist to learn about metabolism.
- ✓ Understanding the best strategies to learn about metabolism and become familiar with the representations and metaphors associated to this field.

Contents

Overview

Metabolism refers to the dynamic state of a living cell. Every complex system, be it a cell or a city, needs a continuous supply of energy and material resources. These permanent fluxes are needed not only to manufacture all their own components, but also for growth, innovation and evolution. There is always a price to pay when processing energy, a fundamental, universal property expressed by the second law of thermodynamics. Even the most complex and optimized cell machinery does not escape degradation. Thus, to maintain and repair structures in a dynamical system requires a continuous supply of energy, and the inescapable release of disorder and by-products. Biotechnology benefits from this dynamic. Through a better understanding of its logic and the modification of its properties, we may use biomolecules in medicine, food or industry. Now that the importance of studying metabolism is clear, let's start.

1.1 What is metabolism?

Metabolism refers to the chemical machinery that allows cells to transform specific sources of chemical elements into chemical energy (ATP, NADPH) and building blocks (small organic molecules, activated coenzymes) for biomass production[1]. In fact, (virtually) all metabolic reactions are *enzyme-catalyzed* and work in a *steady state*. Very few cell reactions occur non-enzymatically.

Traditionally, subsets of enzymatic reactions leading from specific starting material(s) to a product(s) are known as *metabolic pathways*. They can be lineal, branched or cyclical (roundabout). When they participate in energy-conserving reactions and building blocks production, pathways are named *catabolic*. Pathways involved in biosynthesis and biomass production are known as *anabolic*. Anabolic pathways are not (for good thermodynamic reasons) the exact reverse of catabolic pathways, although some remarkable exceptions to this principle have been described[2]. This fact opens opportunities for the evolutionary emergence of regulatory mechanisms[3]. Central pathways are also called *amphibolic* because their intermediate metabolites can be used for both anabolic and catabolic purposes (Figure 1.1.1).

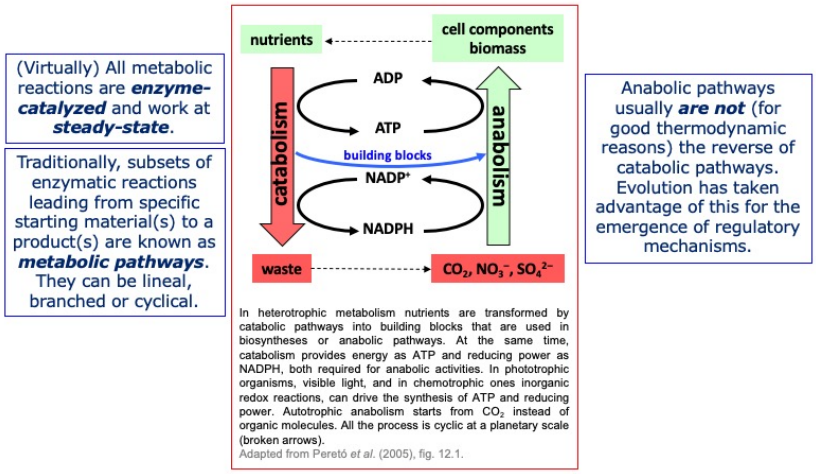


Figure 1.1.1 General overview of metabolic organization: catabolism *vs.* anabolism.

Metabolic modes. Metabolisms are classified according to carbon (auto-, hetero-), energy (photo-, chemo-) and electron (chemo-, organo-) sources (Figure 1.1.2).

[1] See PBL1.1
[2] Wait and follow the discussion on reverse citric acid (Krebs) cycles in bacteria in the following chapter.
[3] See PBL1.2

Mode	Energy source	Reducing power	Carbon source	Examples
chemo-litho-autotrophic	oxidation of inorganic compounds	inorganic compounds	CO_2	hydrogen, methanogen, denitrificant bacteria
photo-litho-autotrophic	visible light	inorganic compounds	CO_2	photosynthetic bacteria, plantas
photo-organo-heterotrophic	visible light	organic compounds	organic compounds	nonsulfur purple bacteria
chemo-organo-heterotrophic	oxidation of organic compounds	organic compounds	organic compounds	animals, fungi, many microorganisms

Peretó *et al.* 2007

Figure 1.1.2 A classification of metabolic modes.

1.2 Why should I study metabolism?

There is a *fundamental reason*: metabolism can be considered an integral part of life. Living beings grow and reproduce, and this means that they must build their own components from external sources. Definitions of a *living being* often include the idea of self-construction – *autopoiesis*, *autonomy* – which is a key aspect for any progress towards a theory of biological organization. The idea of *metabolic circularity* (also known as *metabolic closure* or *recursivity*, and theoretically and mathematically developed by Robert Rosen) starts with the usually ignored fact that enzymes are product of the very same metabolic network in which they act as catalysts (Cárdenas et al. 2013, Letelier et al. 2011, Peretó 2008). Very briefly, metabolic closure expresses the idea that cells are open systems from a thermodynamic perspective (there is a continuous flow of energy and matter through them) but they are *closed to efficient causes*. As a thinking exercise in terms of recursivity, compare a cell with an artificial machine (e.g., a computer)[4].

You could also feel a particular interest on learning about metabolism because of many everyday *applications in biomedical research and biotechnology*. Metabolic enzymes attract a growing interest as evident drug targets and metabolic disorders in diseases such as diabetes or cancer. Considering cells as chemical factories is the core of biotechnology. Most biotechnology projects are exercises of metabolic engineering in practice. Figure 1.2.1 presents a recent example of diagram showing the potential of the bacterium *Pseudomonas putida* as a "global industrial cell factory".

[4] See PBL1.3

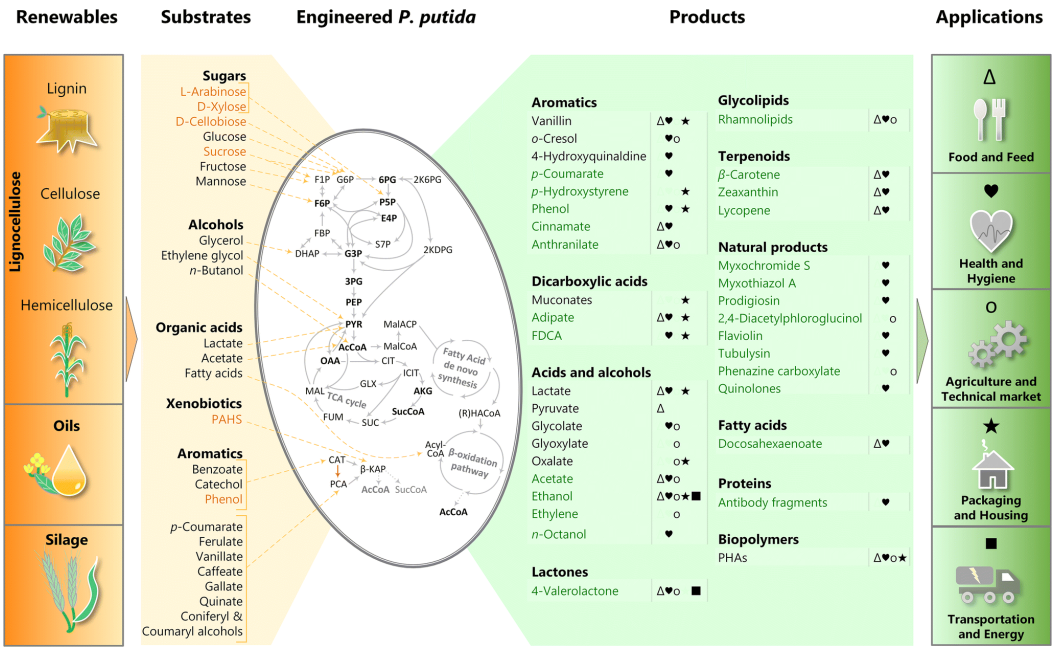


Figure 1.2.1 Figure 5 from Weimer et al. (2020): Bioproduction using engineered *P. putida* strains. Substrates generated from renewable feedstocks, such as lignocellulose, oils and silage can be used to produce value-added products for application in food and feed (Δ), health and hygiene (❤), packaging and housing (★), the transportation and energy sector (■), and for agriculture and technical application (o). The entry points of the respective substrates in the metabolism new to P. putida substrates (orange) and products (green) are marked with a dashed orange line.

You can learn more about my personal vision on metabolism and why it is important to study it in **Reading 1**.

1.3 How should I learn about metabolism?

ROTE *VS.* LOGIC. In his work on evolutionary aspects of biochemistry *The Pursuit of Perfection* (Oxford University Press, 2004) Athel Cornish-Bowden advocates for a logic and evolutionary view on metabolism, instead of using crude memory: [Metabolism] *has been a major component of teaching biochemistry to medical and life-science students, and has typically been taught as if the whole complicated organization was arbitrary or haphazard, the result of a whole series of accidents* [...] *over the long period of evolution since the origin of life some thousands of millions of years ago. I shall try to convince you that this view is wrong* [...].

One way to look at the metabolic logic is to search for *unifying principles*. Probably, one of the most fundamental is *autocatalysis* (Peretó 2012). There are three autocatalytic subsystems relevant for the understanding of life that probably played a key role during the emergence of life on Earth. Small molecules catalyze their own synthesis using external substrates and generating waste products (self-maintenance). Template chemistry allows for the generation of copies of double stranded polymers using activated monomers and generating waste products (self-replication). Finally, amphihilic molecules self-assemble in vesicles with bilayer membranes. Adding more amphiphiles induces the growing and spontaneous division of vesicles (self-reproduction).

Metabolic cycles can be classified as *simple* or *autocatalytic*. Simple cycles allow the chemical transformation of substances with the stoichiometric regeneration of one of the reactants (i.e., the feeder). In terms of standard chemical equations, the catalyst is ignored. For example, the transformation x→y shown in Fig. 1.3.1a requires the participation of a feeder A, which is also a product of the reaction. In this case, the molecule A produced replaces the one consumed. An autocatalytic cycle (Fig. 1.3.1b) exhibits an additional yield of the feeder A. An autocatalytic system must not only contain n>1 elements among the set of products equivalent to the feeder A but, what is more important, the rate of production of A must exceed the rate of its degradation or decay to secondary by-products.

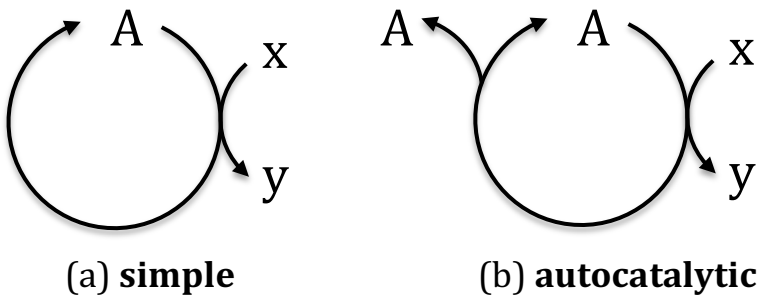


Figure 1.3.1 Catalytic cycles can be simple (2a) or autocatalytic (2b).

At this point you have to **radically change your point of view**: two kinds of catalysts are operative in metabolism, namely, conventional macromolecular enzymes (either protein enzymes or ribozymes) and catalytic cycles.

EVOLUTIONARY CONSTRAINTS also confer a logic of simplicity to metabolism. All living organisms on Earth share common ancestors. This commonality has impacted all organizational levels, including biomolecules. Metabolic pathways are a *theme with variations* in many organisms, further evidence of a common origin that unites all forms of terrestrial life. Yet organisms also contain specific metabolic pathways, such as the secondary metabolism of fungi and plants, that reflect beautiful environmental adaptations. Metabolic organization is the result of optimization processes that have been developed along evolution, taking into account the constraints imposed by the physicochemical framework and the repertoire of feasible enzymatic mechanisms, as well as historical contingencies (Figure 3).

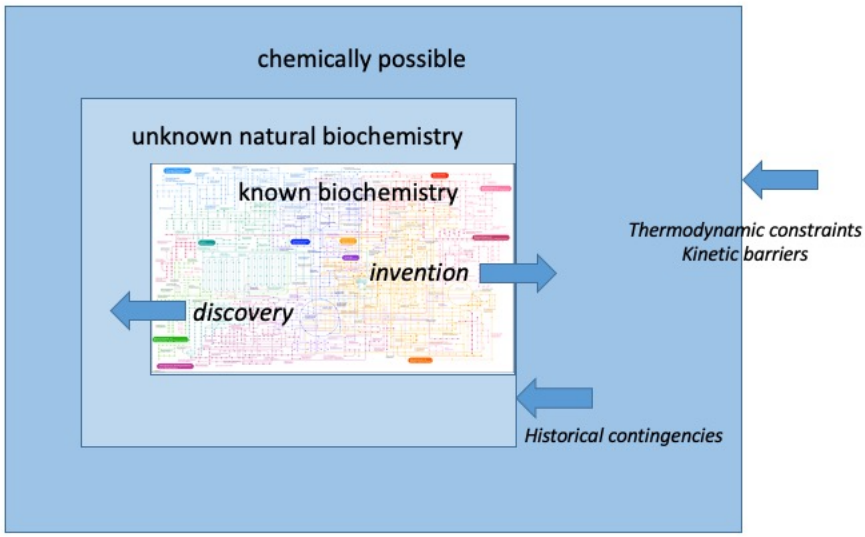


Figure 1.3.2 The possible and the actual. Metabolism represents a subset of the chemically possible. Current research, for instance in environmental metabolism, widens known biochemistry, whereas synthetic biology ventures into territories of the chemically possible that are, however, unexplored by evolution.

Meléndez-Hevia showed that certain metabolic pathways exhibit a simple structure, in the sense that they occur with the lowest possible number of enzymatic steps, whereas others appear as a unique solution to a specific chemical problem. The structure of metabolism is based on specific enzymatic actions that exploit a small repertoire of different mechanisms. Thus, few amino acids or a limited range of cofactors (coenzymes) are involved in the catalytic sites of enzymes. Another example of simplification is the use of common metabolites, especially redox coenzymes and nucleotides, that act as connectors and coupling devices between different reactions. Finally, *substrate ambiguity* and *enzyme promiscuity* play a central role in pathway evolvability (further discussions and original references can be found in Peretó 2012 and in section 1.7).

In biochemical terms, evolution has operated under strict chemical constraints. A non-exhaustive list should include limits with regard to pH, osmotic pressure, and solvent capacity in cells. Cells represent very crowded spaces, and, for instance, the storage of metabolic fuels is accomplished in such a way to minimize excess mass (e.g., polymeric chains). Furthermore, the occurrence of highly reactive chemical groups, such as aldehydes, with potential negative effects on many other cell components, is also minimized (e.g., the linear form of glucose only represents 0.02% of the total). In general, unwanted side-reactions are minimized by virtue of maintaining very low levels of metabolic intermediates (range μmol/L–nmol/L). This fact has consequences on reactivity and the necessity of activated forms of metabolites (i.e., coenzymes) and on the rate of reactions. Activated intermediates (e.g., ATP *vs.* P_i, whereas intracellular [ATP] is 5-8 mM, the [P_i] necessary to phosphorylate glucose would be 1000 M) should be viewed not only as a means of permitting reactions that would otherwise be thermodynamically unfavourable, but of doing so at low concentrations, thus considering solvent capacity and reducing unwanted side-reactions. We will study different strategies of priming or activating metabolites, including phosphorylation, carboxylation, and nucleotidylation.

As an immediate consequence of very low concentrations of metabolites, powerful enzymes are absolutely required, in terms of both specificity and catalytic efficiency. Albeit the variation of kinetic constants of enzymes is also limited by a number of physical constraints (e.g., thermodynamic equilibrium condition, diffusional limitations, steric hindrance) enzymes exhibit high proficiency (a typical increase of reaction rates over uncatalyzed reactions, $x10^{10}$–10^{16}) to catalyse the required reactions, but at surprisingly very low substrate concentrations.

SHOULD YOU LEARN *all* the reactions and pathways by heart? This question is equivalent to the following one: should you know the name of all the streets of a city to move around, navigate it, or indicate a friend where to meet downtown? Look for global traits, remarkable milestones, and landmarks to familiarize yourself with the metabolic map and discover the beauty of the city and the logic behind.

A related aspect, one that has much to do with the importance of imagery in biochemical concepts representation, is *learning metabolism by drawing*. Make your own drawings and diagrams for every metabolic process, once or however many times you need. Design your own metabolic map containing all the information that you consider relevant.

1.4 Organization of metabolism

As stated before, the view of metabolism has followed the path of discovery: individual enzymes organized in pathways with a given physiological function and, more recently, pathways inserted in big networks of reactions. According to the expenditure or conservation of energy, use of reducing power, and the role in the synthesis of building blocks, pathways are classified as catabolic or anabolic. Pathways in the core of the network, like glycolysis, pentose phosphate pathway and the citric acid cycle are also known as *amphibolic*, since many of their intermediates can be both the starting points for biosynthetic pathways or the final products of catabolic pathways.

Many authors have highlighted the existence of simple and universal architectural principles behind the fabulous metabolic diversity. As in many other natural and artificial complex adaptive systems, like ecosystems or cities, metabolic networks hide physical regularities that could be the result of a selective process of optimization – though optimality in metabolic networks is a rather subtle notion, as we will discuss with the case of design diversity in glycolytic pathways. In any case, metabolism is much more than just the physical organization of enzymes and metabolites: beyond its physicality, metabolism is a dynamic state of living, the flow of matter and energy that keeps cells alive.

When observed as a whole, metabolism shows a structure called *bow-tie* (Csete and Doyle 2004). Convergent (catabolic) pathways generate common intermediates and cofactors that are used by divergent (anabolic) pathways. Theoretical studies on the bow-tie structure show that it confers a robust yet fragile architecture and follows general principles that are used by engineers when studying the Internet or the energy (e.g., electricity) distribution network. Understanding the systemic properties of a metabolic network is an essential step for characterizing its functional

behaviour, either in a healthy or pathological state. It is obvious that these approaches also have remarkable biotechnological applications.

Traditionally, people also differentiate between *primary metabolism* and *secondary metabolism*. Albeit somewhat arbitrary, this distinction can be useful to separate those almost universal pathways, directly involved in the survival of the cell and the synthesis of biomass components, from specific pathways emerged during evolution as a result of selective pressures related to behavioural or environmental constraints, like signalling molecules, antibiotics or venoms. Primary metabolism has its roots in external sources of bioelements, whereas secondary metabolism starts from intermediate metabolites. This is the case for flavonoids, alkaloids, terpenes or polyketides and a wide array of the so-called *natural products*, most of them with an industrial or applied interest. Since animals and other eukaryotic lineages have lost their ability to synthesize (in whole or in part) many enzymatic cofactors (vitamins), the biosynthesis of these molecules sometimes is considered (with an anthropocentric perspective) as a part of secondary metabolism.

You can learn more about the importance and significance of secondary metabolism and topics related to the so called 'natural products' in **Reading 2**.

1.5 Discovering metabolism: a systems biology perspective

Since the pioneering works by Buchner in 19th century, biochemists have taken advantage of the characterization of isolated pieces of metabolism: purified enzymes, sets of enzymes working in pathways, isolated organelles and cell compartments, etc. The following advice was very popular: don't waste clean thinking on dirty enzymes. But a fundamental limitation of this reductionist approach was the lack of information about system behaviour and the global control of fluxes. It is now possible to obtain the complete genomic information of a cell, the complete set of enzymes and metabolites, to infer a complete metabolic network and use a mathematical reduction of its complexity to gain information on the system's function. Constraint-based metabolic modelling is one of the main tools in current metabolic studies, especially in biomedical and biotechnological research.

Historically, methods based on the use of inhibitors and labels (chemical or isotopic) have allowed to outline the metabolic pathways and propose intricate bidimensional maps of cell metabolism. As roadmaps of metabolism, those representations have been both immensely useful for research and extremely indigestible for students. The use of stable isotopes and mass spectrometry has opened a new era in metabolic studies, including the discovery the new pathways and systemic approaches (metabolomics, fluxomics).

In a simplistic way, we could say that classical enzymology focuses on individual, purified enzymes, whereas metabolomics is interested in (large) assemblies of reactions. Thus, while in the first case the attention focuses on local properties, metabolomics is interested in global properties. For instance, individual enzyme study searches the characterization of ligand affinities, whereas typical metabolomic studies try to determine the metabolic profile of a sample and the distribution of fluxes throughout the system. Studying hundreds of metabolites or fluxes at the same time implies

ignoring many kinetic details, although in cases where more dynamic approaches are required, traditional biochemistry and metabolomics need each other. The first has accumulated a great deal of data on enzyme kinetics, currently available in databases such as BRENDA. Typical metabolomics data sets consist in comprehensive metabolic concentration or flux profiles (at steady state or dynamically changing in response to a stimulus).

The analysis of metabolic pathways follows one of the two main routes: *stoichiometric analysis* on a more static network basis (Wiechert 2002), and *metabolic control analysis* on a more dynamic view. In this course, we will focus on stoichiometric thinking and its applications in biotechnology. Freely accessible databases such as KEGG and BioCyc are enormously helpful for the construction of a stoichiometric map. They contain collections of pathways and links to the main repositories of biochemical and genomic data. In this course, computer-based sessions will focus on enzyme and metabolic databases (session 1) and a short introduction to stoichiometric analysis through the use of the METATOOL software (session 2).

What is a metabolic pathway?

One central concept derived from the stoichiometric analysis of metabolism is *elementary mode*. This can be defined as the minimal (non-reducible) set of enzymatic transformations between two given substrates that operate in a steady state. Some authors have suggested that elementary modes mathematically define "pathways" inside the metabolic network. In fact, the pathways you can find in the different chapters of a textbook emerge as elementary modes in stoichiometric analysis. Furthermore, enumerating the elementary modes of a network allows the discovery of new "pathways" that, beyond mere mathematical curiosity, can be of physiological relevance, as has been demonstrated in some instances. Thus, a pathway can be a chapter in your textbook or an elementary mode derived from the stoichiometric analysis. But how were the "textbook pathways" initially defined? Most metabolic pathways have historical roots, yet they correspond to well defined physiological conditions accessible to empirical studies (i.e., constrained by the technology of the moment) in model or simplified systems, for example, glucose consumption under hypoxia (alcohol fermentation in yeast, lactate formation in active muscle), organic acids complete oxidation during oxygen respiration in muscle cells (which led to Krebs cycle discovery), photosynthetic C fixation in green algae (Calvin cycle from CO_2 to C3 molecules), etc. In summary, the arbitrary subdivision of the whole metabolic network in pathways can be just a matter of convenience for research or study (as proof, check how many pathways for human metabolism are defined in databases like KEGG or HumanCyc) but also a fabulous source of pragmatic problems (Stobbe et al. 2011).

IN CONTRAST TO CLASSICAL ENZYMOLOGY, modern high-throughput metabolomics involves the simultaneous generation of large metabolite or flux profiles which sometimes consist of hundreds of data points. The difference in complexity between metabolomics, on the one hand, and genomics and proteomics, on the other, is remarkable. Thus, the genome is the combination of (not-so-different) four letters, and proteomics, of twenty, whereas the metabolome consists of thousands of metabolites that are chemically very diverse and come in all sizes, chemical families and affinities to water or lipids. In general, metabolite analysis entails a series of procedures: (1) Sample preparation and quenching of metabolite activity. (2) Extraction procedure. (3) Separation of metabolites. (4)

Quantification and/or profiling with methods such as nuclear magnetic resonance (NMR) and mass spectrometry (MS). (5) Data analysis.

Stoichiometric and flux balance analysis (FBA) requires the experimental determination of fluxes. Metabolic flux analysis (MFA) is based on the study of *isotopomers*. The term is a combination of the words isotope and isomer and it means one of the different labelling states in which a particular metabolite can be encountered (for a complete discussion, see Wiechert 2001). The various possible isotopomers (for instance, 2^n for a molecule with n carbons in a ^{13}C-labeling experiment) are usually not equally likely and are instead produced with a frequency distribution that contains clues on the internal fluxes of the system. In general, the system is assumed as an open steady state, where material is fed to and metabolized by the system, but all metabolite concentrations and fluxes are constant. Evaluation of the isotopomer distribution requires a mathematical model of the pathways to describe how the labelled material is distributed throughout the system (Wiechert 2002). The parameters of the model must be estimated from the isotopomer distribution and, then, internal fluxes can be quantified. Thus, the whole approach requires skilfulness in complex techniques (NMR, MS), mathematical modelling and statistics.

Most of the above text follows Voit[2]. For a general and extended discussion on metabolic systems, see ch. 8. Voit's text devotes ch. 3 to static networks, more specifically to stoichiometric analysis (p. 70) and metabolic control analysis (pp. 74). More on control analysis can be found in Cornish-Bowden[4] (see the following discussion in section 1.6).

1.6 Regulation and control

A central aspect of the study of metabolism is to understand how metabolic fluxes are modulated under diverse conditions: normal, pathological or manipulated with technological purposes. The molecular mechanisms of enzyme regulation have been revised in the *Biochemistry* course, and you can find good introductions in general textbooks (see, e.g., Peretó et al.[5] ch. 6).

Video #1.1 | Molecular mechanisms of enzyme regulation

A series of commented slides on allosterism and covalent modification of enzymes are available at the AV.

The excellent enzymology textbook by Cornish-Bowden devotes ch. 12 to enzyme regulation and ch. 13 to multi-enzyme systems, and is a good introduction to the topic of metabolic regulation and control. The following text reproduces most of the summary of ch. 13 (p. 377-379) of Cornish-Bowden's text[4], as well as some sentences from Nelson and Cox[5] (p. 582) and Hofmeyr and Cornish-Bowden (2000).

Although enzymes have usually been studied in the absence of other relevant catalytic activity, to understand their physiological roles, it is important to first consider their behaviour as components of *metabolic pathways*.

Certain pairs or triplets of metabolites, such as NAD and the adenine nucleotides, participate as *moieties* in numerous different metabolic reactions. Their total concentrations can often be considered to be maintained constant by mechanisms outside the pathways of interest.

Metabolic control analysis (MCA) is a way of quantifying the sensitivity of metabolic fluxes and metabolite concentrations to each enzyme in a pathway. MCA shows that control of the rate of metabolite flux through a pathway is distributed among several of the enzymes in that path.

Key terminology in metabolic control analysis: *Elasticity*, ε, is an experimental measure of the sensitivity of the rate of an isolated enzyme-catalyzed reaction to the concentration of any molecule (a substrate, product or another metabolite). The *flux control coefficient*, *C*, is an experimental measure of the sensitivity of a metabolic flux to the activity of one enzyme, and it is a characteristic of the whole system, not intrinsic to the enzyme. The *concentration control coefficient* is the corresponding measure for a metabolite's concentration.

In studies of an isolated enzyme, it is common to consider the rate to be determined by the concentrations of substrates, products and effectors. When the enzyme is embedded in a metabolic system, it is sometimes more accurate (though still an approximation) to consider these concentrations to be determined by the flux through the reaction. The *sum of all the flux control coefficients* for a particular flux is 1.0: this leads to the idea that *all the enzymes share flux control* in a system. This summation property has obvious (albeit poorly understood) consequences for metabolic engineering.

The relationship between *control* and *regulation* of a biosynthetic pathway can be analysed in terms of a *supply block* (the part of the pathway that produces the "end product") and a *demand block*. The latter is often omitted from textbook representations of biosynthesis, which makes it impossible to understand regulation. Cellular metabolism is a molecular economy that is functionally organised into supply and demand blocks linked by metabolic products and cofactor cycles. The central regulatory problem of metabolism is how to satisfy a varying demand for its products, from low to high values, while maintaining these products within narrow concentration ranges far from equilibrium. Supply-demand analysis shows that these two functions are linked: the more control either block has over flux, the less it determines the degree of homeostasis and the distance from equilibrium where homeostasis is maintained, which becomes the function of the other block. A common solution to this design problem in living cells is that the flux is largely controlled by the demand block, whereas the supply block determines homeostasis of the linking metabolite. Supply-demand analysis also has major implications for biotechnology, biomedicine and drug design because it shows that what were thought to be 'rate-limiting' steps catalysed by allosteric enzymes actually have nothing to do with flux control, but are responsible for the metabolite homeostasis.

Metabolic channeling is the direct transfer of a metabolite from the enzyme that produces it to another enzyme that uses it as substrate.

Interconvertible enzyme cascades allow very much higher sensitivity to a metabolite concentration than is possible for a single enzyme, at the expense of ATP consumption.

The equilibrium catalysed by *adenylate kinase* allows small variations in ATP concentration to be amplified into large relative changes in AMP concentration. The role of AMP in enzymatic regulation based on the energetic status of the cell will be discussed in Chapter 6 of the course.

1.7 Evolution of metabolism

We still do not know when and how life originated. But useful hints can be inferred from extant metabolic pathways, as well as from their correlation with environmental changes through planetary history. Although we still lack a narrative for the origin and evolution of metabolic pathways, a true natural history of biochemistry, we are gaining insights from comparative genomics and molecular cladistic analyses of individual enzymes. This approach has been fruitful because of the vision of evolution as a tinkerer, rather than an engineer. Using Jacob's words (1977), inspired by Darwin (1862): *In contrast to the engineer, evolution does not produce innovations from scratch. It works on what already exists,* [...] *like a tinkerer who, during millions of years, has slowly modified his products,* [...] *using all opportunities to transform and create.* In summary, evolution hardly builds anything new from scratch.

Metabolism in extant organisms is an astonishingly complex and adaptable chemical machinery. For instance, the genome of a typical free-living bacterium feeding on organic matter like *Escherichia coli* has the potential to express enzymes catalysing c. 1,000 chemical reactions, and the functions of almost 20% of its genes have yet to be annotated. Indeed, a single *E. coli* cell, growing exponentially under standard aerobic conditions, contains 100 million metabolite molecules and 2.4 million enzyme molecules. The density of proteins inside a cell seems to be optimised to maximise the speed of biochemical reactions: enzyme-catalysed transformations are hardly rate-limiting for a cell, which appears to be more dependent on the kinetics of physical processes like molecular diffusion and protein folding. Gaining an understanding of the harmonious steady state flux balance still represents a challenge to biochemists, biotechnologists, and systems biologists. Even more challenging is trying to delineate the processes that led from presumably small dirty protometabolic networks of chemical transformations to the sophisticated metabolic networks found in modern cells. Nevertheless, despite the emphasis on enzymatic specificity, from what we learn from textbooks and the complete history of enzymology, it is clear that metabolic accuracy has its limits, and that stochastic factors, broad substrate specificity, and *catalytic promiscuity* contribute to a somewhat fuzzy chemistry of metabolism, making it essential for evolution to occur.

Video #1.2 | Enzyme promiscuity and evolution

Commented slides on enzyme promiscuity and its role in evolutionary innovation are available at the AV.

Fragment of my lecture "Bricolatge i oportunisme en l'evolució del metabolisme" (2020).

There are several models describing how metabolic pathways diversified after protein enzymes became established. It is generally assumed that the *patchwork* or *recruitment model*, which proposes that metabolism was initially performed by a small repertoire of enzymes with low

specificity, offers the most comprehensive explanation. The oldest metabolic pathways may have been assembled by recruiting primitive enzymes showing remarkable substrate ambiguity and catalytic promiscuity. Specificity, specialisation and efficiency in metabolism were explored and expanded by duplication and divergence of the genes coding for metabolic enzymes. However, this must have occurred without completely eroding some of the ancestral characteristics of the primitive catalysers, because many modern enzymes also exhibit broad substrate specificity and multifunctionality. In fact, the co-option of latent enzymatic activities has been tested through experimental evolution (e.g., protein resurrection or the artificial synthesis of a putative protein ancestor) and genetic studies of microorganisms (e.g., forcing the emergence of promiscuous activities by overexpressing the corresponding coding genes, see below). Enzyme recruitment has successfully been invoked to explain the evolution of many pathways, including the urea and citric acid cycles, the autotrophic pathways, and the biosynthesis of amino acids, cofactors and membrane phospholipids.

Recent easy access to massive data on metabolic processes present in organisms from the three domains of life has boosted the development of new systemic approaches to the evolution of metabolism. Such approaches reach well beyond the arbitrary borders of individual metabolic pathways that researchers have delineated during the historical development of biochemistry. Thus, phylogenetic methods and functional analyses of genome-wide metabolic reconstructions allow us to test the predictions of the patchwork model regarding the mosaic distribution of homologous protein domains throughout contemporary metabolic networks. Furthermore, it is remarkable that the distribution of kinetic constant values is dependent on the metabolic context, with the enzymes of the central intermediate metabolism (e.g., glycolysis, citric acid cycle, and pentose phosphate pathway) being those with higher catalytic (k_{cat}) and specificity constants, whereas those of the secondary metabolism exhibit lower kinetic constant values, that is, they are relatively slower and less specific.

At any rate, the uneven distribution of substrate ambiguity and catalytic promiscuity in the metabolic networks could reflect the evolutionary history of enzymes but could also be of adaptive value to organisms, working as a reservoir of metabolic innovations. In experiments performed with a collection of knockout mutants of *E. coli*, 20% of auxotrophs were rescued by overexpressing multifunctional suppressors, either isozymes or enzymes (or transporters) showing substrate ambiguity or promiscuous enzymes. In other words, extant metabolic networks are plastic and evolvable, thanks to the intrinsic structural and functional flexibility of enzymes.

Since living beings are open chemical systems that use chemicals from the environment and produce and release other compounds, there is a co-evolution of microbial metabolisms and the biogeochemical cycles implying the coupling and consistency of any evolutionary scheme with the geochemical record and the planetary chemistry. In this sense, chemical and isotopic studies in ancient rocks are fundamental to provide the best estimates for the minimal timing of any metabolic establishment.

ONE OF THE MOST DEBATED ISSUES in this context is the appearance and atmospheric accumulation of molecular oxygen and its metabolic and geological effects. About 2.4 Ga ago, atmospheric oxygen increased in what is called the great oxidation event (GOE), to a level reaching about 10% of the present atmospheric level (PAL). Most authors agree that the invention of oxygenic photosynthesis by the ancestors of cyanobacteria was the main source of oxygen, so GOE would set a minimum age for this metabolism. Some authors have postulated an earliest (more than 3.0 Ga ago) origin for cyanobacteria and thus the lag period for the oxygen increase would be explained by diverse sink reactions, like oxygen consumption by methane of biological origin or the oxidation of oceanic iron (with the consequent deposition of banded iron formations, BIF). Eventually, the marginalization of methanogenesis (e.g., due to the scarcity of an essential metal like Ni for the enzymes involved in methane synthesis) would contribute to the accumulation of oxygen. This extremely reactive molecule would also offer an opportunity for the evolution of new metabolic processes, like oxygen respiration or the synthesis of new metabolites, like steroids. Many authors also suggest that oxygen accumulation in the atmosphere was a requisite for a further increase in life complexity, i.e., the origin of biomineralization and multicellularity.

The mechanism of emergence of new metabolic processes would imply recycling and modifying old molecular devices that were in place during the anoxic phase of the metabolism, e.g., oxygenic from anoxygenic photosyntheses, aerobic from anaerobic respirations or oxygen-consuming reactions instead of the anoxic versions. The comparative study of the oldest metabolic steps and how they adapted (or whether or not they adapted at all) to the new situation in an oxygenated planet is a useful exercise in evolutionary biochemistry. For instance, why some metabolic pathways like methanogenesis never adapted to the presence of oxygen, whereas others, like carbon fixation by Rubisco, tolerated its presence; or the case of nucleotide reductase, which opportunistically coopted this molecule as a stronger oxidant.

Video #1.3 | Oxygen and metabolic evolution

Commented slides on the effect of oxygen on metabolic evolution are available at the AV.

Further information and original references in Peretó (2011, 2012, 2019).

1.8 Representations and metaphors in metabolism

Every textbook or research paper in biochemistry or molecular biology is full of graphical representations and imagery of the concepts and topics under discussion. Metabolism has also its own way of representing processes and concepts, basically using diagrams where molecules are represented either by skeleton formulas or just names (in general, common names are used instead of IUPAC nomenclature for historical reasons and, overall, for the sake of simplicity), whereas arrows (single/irreversible, double/reversible) indicate chemical transformations catalysed by enzymes. The use of curved arrows to represent the participation of more than one substrate or a cofactor in a reaction is also common. Regarding the molecular structure, Fisher projections and skeleton

formulas are very popular in metabolic representations. Figure 1.8.1 takes as an example a single enzymatic reaction (although one with a two-step mechanism) from the citric acid cycle catalysed by isocitrate dehydrogenase, and several ways to represent the intermediate isocitrate (official IUPAC name 1-hydroxypropane-1,2,3-tricarboxylic acid). The common name isocitrate derives from the fact that it is a structural isomer of citrate, a weak organic acid that exists in a variety of fruits and vegetables, most notably citrus fruits.

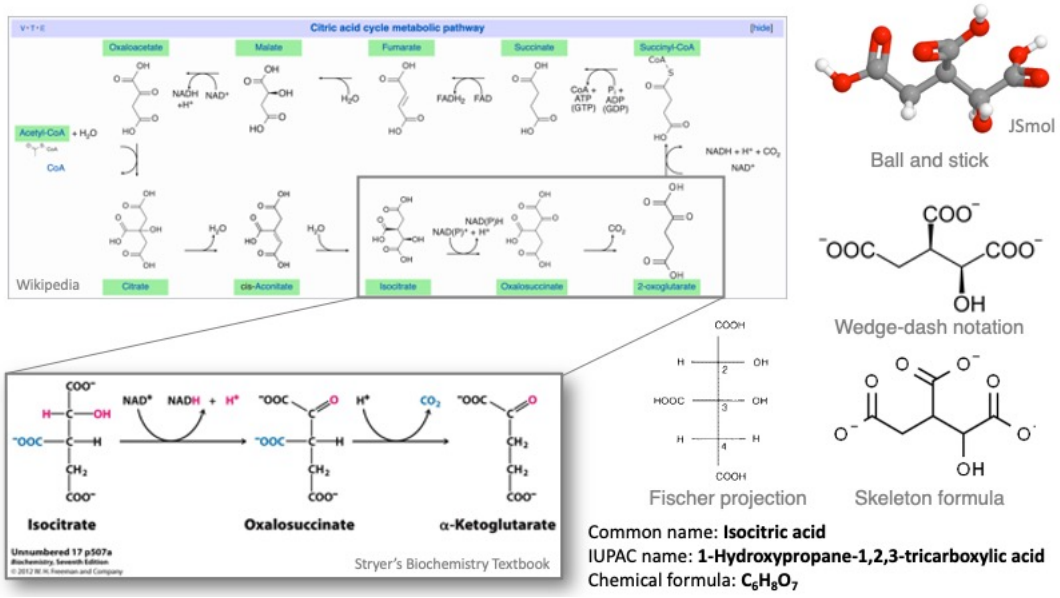


Figure 1.8.1 The diversity of forms to represent molecules in the context of a metabolic diagram. Inset: the CAC in Wikipedia and the detail of the isocitrate dehydrogenase reaction as represented in Stryer's Biochemistry.

In the field of metabolism, maps are the most powerful form of representation. Metabolic maps are sets of connected reactions trying to present all the possible enzymatic transformations in a system. Maps can recollect the reactions of a given kind of cell (e.g., a bacterium or a hepatocyte), but can also have a meta-metabolic meaning, i.e., the set of all known reactions, regardless of the presence or not in a specific cell type. Figure 1.8.2 is an example of meta-metabolic map.

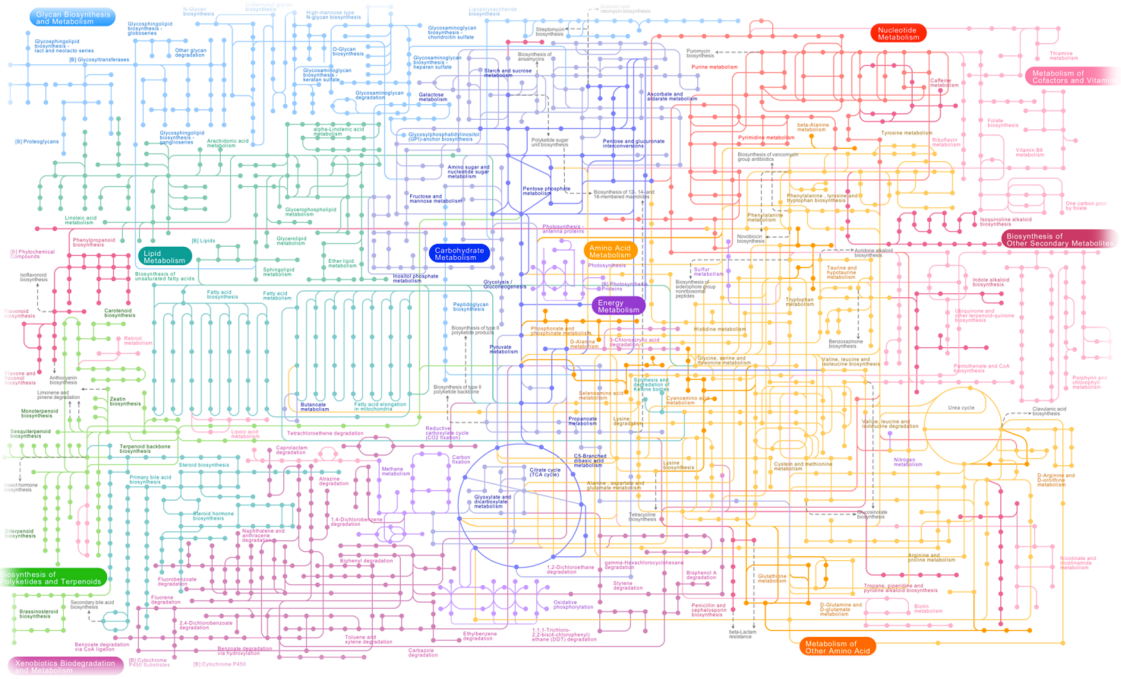


Figure 1.8.2 The meta-metabolic map of the KEGG database. Dots represent metabolites, edges, reactions, and colors, arbitrary metabolic modules or pathways. This map is built from the collection of reactions deposited in the database and is interactive (connects with all the information regarding biochemical and genomic data behind each metabolic step).

Metabolism regulation and control are concepts that need also their map representations. Hence, arrows will represent not only chemical reactions, but also actions (like hormones and second messengers in enzyme cascades) and transfer of information (including allosteric interactions and the transcriptional/translational regulation of enzymes). With the emergence of systems biology, new forms of representation and visualization have been popularized, and images have also become tools for mathematical and computational modeling. Particularly in metabolism, pathways have been gradually replaced by networks and their mathematical representations (e.g., stoichiometric matrix) as a step towards metabolic models that simulate the distribution of flow among the reactions.

Unfortunately, we are far from a standard system to representat metabolic maps. There is a huge diversity of ways to show the relationships between metabolites and pathways and integrate the deluge of data that we withstand today. As an example, in Fig. 1.8.3, I compare the diverse imagery representing the CAC in four different databases: KEGG, MetaCyc, BRENDA and WikiPathways.

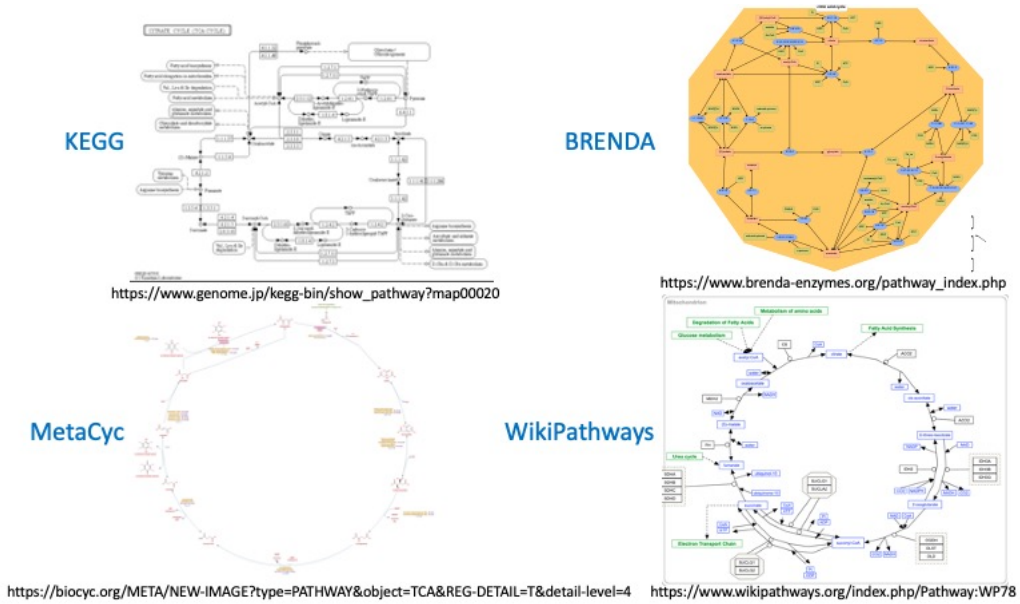


Figure 1.8.3 A range of representations of the CAC in popular data banks.

Metaphors are so rooted in biochemists' way of thinking that, although we do use them, they tend to go unnoticed. Come back to any text on metabolism (this chapter, for instance) and try to be aware of metaphors: pathway, step, building block, map, network, cascade, tinkering, machinery, bow-tie, factory, etc. Even the word cell is, in its origin, a metaphor. Historians and philosophers of science have explored how metaphors go beyond pure analogy or description. They have true epistemic value and may have provided insight on (or limited) the way we think and discover in science (Morange 2020).

1.9 How is this course organized?

Before we start, we should justify how this course on metabolism, addressed to biotechnologists, is organized. After recognizing the suitability to subdivide the network into several arbitrary pathways in order to facilitate the study, I have chosen a classical organization of topics starting with the quasi-universal central pathways dealing with C compounds, which coincide with the minimal network necessary to supply the precursors to all biomass components in a model system like *E. coli*. The following Table presents the list of 12 necessary precursors for the synthesis of all biomass components in a bacterial cell: polysaccharides (cell wall), lipids (membranes), nucleotides (nucleic acids), amino acids (proteins). The 12 precursors belong to three central pathways: citric acid or tricarboxylic acid cycle (CAC/TCA cycle), glycolysis, and pentose phosphate pathway (PPP). The connectivity between these pathways and the diversity of links with environmental sources of bioelements (i.e., nutrients) can explain the different metabolic modes compatible with a given physiological situation, type of C source, presence or absence of oxygen, etc.

Biomass precursors in *Escherichia coli*

TABLE 2.2 Twelve Major Precursor Metabolites and the Primary Pathway Used for their Production

Precursor Metabolite		Metabolic Pathway	Amount Required for Biosynthesis of *E. coli* (μmol/g cell)
Glucose-6-phosphate	G6P	Glycolysis	205
Fructose-6-phosphate	F6P	Glycolysis	71
Triose phosphate	GAP (glyceraldehyde 3P)	Glycolysis	129
3-Phosphoglycerate	3PG	Glycolysis	1,496
Phosphoenolpyruvate	PEP	Glycolysis	519
Pyruvate	PYR	Glycolysis	2,833
Ribose-5-phosphate	R5P	PPP	898
Erythrose-4-phosphate	E4P	PPP	361
Acetyl-CoA	ACA	TCA cycle	3,748
α-Ketoglutarate	2OG (2-oxoglutarate)	TCA cycle	1,079
Oxaloacetate	OXA	TCA cycle	1,787
Succinyl-CoA[a]	SCA	TCA cycle	–

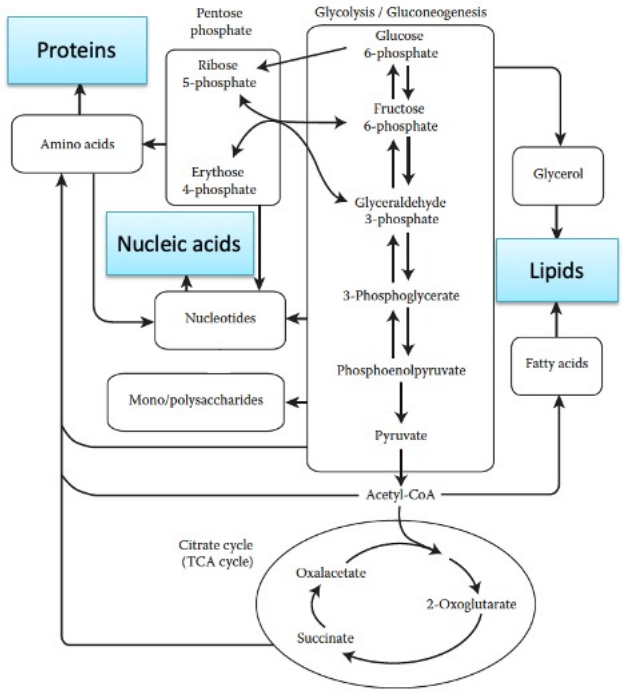
Source: Ingraham, J.L., Maaloe, O., and F.C. Neidhardt. *Growth of the Bacterial Cell*, Sinaeur Associates, Inc, Sunderland, MA, 1983. With permission.

Note: Also included are the specific proportions of these metabolites required in the synthesis of *E. coli*.

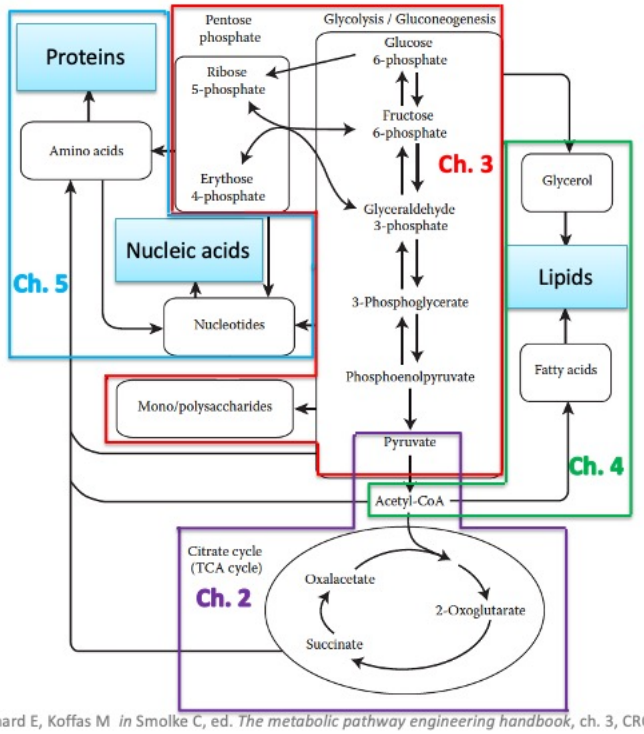
[a] Succinyl-CoA is a precursor metabolite in the synthesis of tetrapyrroles.

Akinterinwa O, Cirina PC *in* Smolke C, ed. *The metabolic pathway engineering handbook*, ch. 2, CRC Press, 2010

The following figures present the central pathways, the 12 biomass precursors and their connections with biomass components, as well as the proposed subdivision of pathways in chapters during this course: CAC (chapter 2), glycolysis/gluconeogenesis (chapter 3), lipid pathways (chapter 4), amino acid and nucleotide metabolism (chapter 5). The course is complemented with a discussion on metabolism regulation and control (chapter 6) and a final presentation on metabolic engineering and biotechnological applications (chapter 7).



Fowler ZL, Leonard E, Koffas M *in* Smolke C, ed. *The metabolic pathway engineering handbook*, ch. 3, CRC Press, 2010



Fowler ZL, Leonard E, Koffas M *in* Smolke C, ed. *The metabolic pathway engineering handbook*, ch. 3, CRC Press, 2010

Figure 1.8.3 Major metabolic pathways and how this course on metabolism is organized.

Video #1.4 | Metabolic overview

In these slides I remind you of some basic concepts and explain how the course is organized.

Bibliography and resources

Textbooks

Berg JM, Tymoczko J, Gatto G, Stryer L (2019) *Biochemistry*, 9th ed., New York: WH Freeman [Stryer[9]], ch. 10, ch. 15.

Cornish-Bowden A (2012) *Fundamentals of Enzymology*, 4th ed., Weinheim: Wiley-VCH [Cornish-Bowden[4]], ch. 13.

Metzler DE (2003, 2004) *Biochemistry: the chemical reactions of living cells*, 2nd ed., New York: Academic Press [Metzler[2]], vol. 1 ch. 10, vol. 2 ch. 17.

Nelson DL, Cox MM (2017) *Lehninger Principles of Biochemistry*, 7th ed. New York: WH Freeman [Nelson and Cox[7]], ch. 1, ch. 15.

Peretó J, Sendra R, Pamblanco M, Bañó MC (2005) *Fonaments de Bioquímica*, 5ena ed., València: PUV [Peretó et al[5]], ch. 12 [Spanish edition, 2007].

Voit EO (2017) *A First Course in Systems Biology*, 2nd ed., New York: Garland Science [Voit[2]], ch. 8.

Papers and book chapters

Cárdenas ML, Piedrafita G, Montero F, Cornish-Bowden A (2013) ¿Qué es la vida? *SEBBM* **175**:14–17 (accessed 5 July 2020).

Csete M, Doyle J (2004) Bow ties, metabolism and disease. *Trends Biotechnol* **22**: 446–450.

Hofmeyr JHS, Cornish-Bowden A (2000) Regulating the cellular economy of supply and demand. *FEBS Lett* **476**:47–51 (accessed 5 July 2020).

Letelier JC, Cárdenas ML, Cornish-Bowden A (2011) From *L'Homme Machine* to metabolic closure: steps towards understanding life. *J Theor Biol* **286**:100–113 (accessed 5 July 2020).

Morange M (2020) *The black box of biology. A history of the molecular revolution*. Ch. 29: Images, representations, and metaphors in molecular biology. Harvard University Press.

Peretó J (2008) Sobre la naturaleza y la fabricación de seres vivos. *Apuntes de Ciencia y Tecnología* **27**:30–38 (accessed 5 July 2020).

Peretó J (2011) Origins and evolution of metabolisms. In: *Origins and Evolution of Life. An Astrobiological Perspective* (Gargaud M et al., eds.), Cambridge University Press.

Peretó J (2012) Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem Soc Rev* **41**:5394–5403 (accessed 5 July 2020).

Peretó J (2019) Prebiotic chemistry that led to life. In: *Handbook of Astrobiology* (Kolb V, ed.) Ch. 5, CRC Press.

Stobbe MD, Houten SM, Jansen GA, van Kampen AH, Moerland PD (2011) Critical assessment of human metabolic pathway databases: a stepping stone for future integration. *BMC Syst Biol* **5**:165.

Weimer A, Kohlstedt M, Volke DC, Nikel PI, Wittmann PC (2020) Industrial biotechnology of *Pseudomonas putida*: advances and prospects. *Appl Microbiol Biotechnol*.

Wiechert W (2001) ^{13}C Metabolic Flux Analysis. *Metab Eng* **3**:195–206.

Wiechert W (2002) Modeling and simulation: tools for metabolic engineering. *J Biotechnol* **94**:37–63 (accessed 5 July 2020).

On the net

The classical and famous Roche poster on biochemical pathways by Gerhard Michal is now available online. On the other hand, Merck has adapted the iconic IUBMB-Nicholson Metabolic Pathways Chart to an interactive version.

WikiPathways is an open, collaborative platform dedicated to the curation of biological pathways.

MACiE is a useful webpage for exploring the biochemical diversity in terms of enzymatic mechanisms and coenzymes.

At this moment (August 2020), the WikiProject Molecular Biology/Metabolic Pathways is inactive. Maybe you can consider reactivating it.

In his famous series Molecule of the Month for the PDB database, biochemist and artist David S. Goodsell offers fascinating images and descriptions of many molecules and enzymes.

Study Questions

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ1.1 Define *amide*, *lactone*, *ester*, *thioester*, *ether bond*, *anhydride bond*, *hemiacetal*, *Schiff base*, *keto-enol equilibrium*, *pyranose* and *furanose rings*, *enantiomer*, *anomer*.

SQ1.2 The reaction of NADH with oxygen to produce NAD^+ and H_2O is very exergonic, yet the reaction of NADH and oxygen takes place very slowly. Why does a thermodynamically favorable reaction not take place rapidly?

SQ1.3 Match the following two columns:

1. Electron acceptor in oxidations	a. Autocatalytic cycle
2. Electron donor in reductions	b. NAD^+
3. It can promote exponential growth	c. Biosynthesis
4. Anabolism	d. NADPH

SQ1.4 What is the structural feature common to ATP, FAD, NAD^+, and CoA? Suggest an evolutionary explanation to this fact.

SQ1.5 Trehalose is a highly effective cryoprotectant that has been selected as "blood" sugar in insects. Give a chemical reason why a disaccharide such as trehalose is abundant in insect haemolymph, instead of a different disaccharide such as maltose.

SQ1.6 Imagine that you are cooking paella. While frying the chicken and rabbit meat, you realize that the surface is browning. This *sofregit* is an essential source of flavor and the result of a non-enzymatic, heat-induced food-browning process named Maillard reaction. Identify the two types of biomolecules that participate in this reaction and which sort of bond is initially established between them.

SQ1.7 One letter makes a big difference. Define *autotroph* and *auxotroph*.

SQ1.8 Under laboratory conditions, the green alga *Chlamydomonas reindhartii* can be grown in an illuminated liquid mineral medium with bubbled carbon dioxide. However, it can grow in darkness by consuming acetate aerobically. Determine the sources of carbon, energy and electrons in each case and provide the complete name of the corresponding metabolic mode.

SQ1.9 Differentiate between *steady-state* and *flux* in a metabolic pathway. To illustrate these concepts, hydraulic metaphors can be very useful.

SQ1.10 ***Be a cultivated biotechnologist.*** Radioactive ^{14}C was discovered in 1940 and was available in significant quantities by 1946 as a product of nuclear reactors. For instance, in 1946, Melvin Calvin and Andrew A. Benson began their studies that elucidated the mechanism to incorporate CO_2 into organic materials during photosynthesis, by combining metabolite separation through two-dimensional paper chromatography with radioautography. Historians generally agree that the rapid and widespread adoption of the use of radiolabeled molecules in biochemical research was an indirect consequence of the Cold War. Why?

Problem-based learning 1 (PBL1)

You will be asked to present written answers for these questions. Eventually, some of the questions will be discussed in classroom.

Reading

1 En el siguiente texto expongo mi visión personal sobre el metabolismo y los aspectos más relevantes que justifican su estudio e investigación.

En la Biología contemporánea casi todo ha girado alrededor de los genes. El DNA ha sido preeminente en la investigación gracias a los grandes éxitos de la Biología Molecular y, sin embargo, ni un bit de la información contenida en los genomas se haría realidad sin el concurso de la maquinaria química celular que los descodifica: el metabolismo. La Biología de Sistemas es un réquiem por la Biología Molecular: nos prometieron que secuenciar los genomas nos traería definitivamente el secreto de la vida y veinte años después es hora de reconocer que una cosa es la información y otra muy diferente el conocimiento, y que el pensamiento lineal y reduccionista no basta. La irreversibilidad del flujo de la información genética, junto con la cadena lineal de causas y efectos implícita en el Dogma Central de Crick, favoreció el espejismo de que todo lo que necesitamos saber está en los genes. Pero a medida que investigamos más nos percatamos de que en este esquema hay una parte importante de la complejidad biológica que se nos escapa. Apreciar la centralidad del metabolismo equivale a darle la vuelta al Dogma Central y considerar los procesos de flujo de la información como subsidiarios de los flujos de materia y energía que sostienen las células.

La comprensión de la vida a todas las escalas, desde las moléculas hasta los ecosistemas, contiene un *leitmotiv* inevitable: la lucha de los organismos por conseguir las materias primas (los elementos químicos) y las fuentes de energía, necesarias para persistir lo suficiente como para crecer y hacer copias de sí mismos. Es, pues, una lucha contra el tiempo –ya lo señaló Erwin Schrödinger cuando observó que los seres vivos "hacen algo" durante más tiempo de lo que cabría esperar de un "pedazo de materia inanimada en circunstancias similares". Es fascinante preguntarse cómo surgió esa factoría química incansable de la vida y sus peculiares ciclos autocatalíticos a partir de la geoquímica del planeta primitivo (lo que Christian de Duve llamó protometabolismo) y hasta qué punto el determinismo físico-químico en esa transición crucial prefiguró el metabolismo actual. De manera trascendental, la exploración operada por la selección natural se superpuso al determinismo. Determinismo y contingencia histórica dibujaron el paisaje de esa química encarnada en las células que llamamos metabolismo.

El metabolismo está configurado como una red de reacciones químicas entrelazadas que permiten la transformación de materiales ambientales en biomasa, es decir, en todos y cada uno de los componentes celulares. Las redes metabólicas pueden nutrirse de materiales más o menos elaborados por otras redes metabólicas o partir de compuestos químicos simples, en el caso extremo de la autotrofia, de CO_2, N_2 y sales minerales disueltas en agua. Un aspecto notable del metabolismo es por qué los centenares o, a veces, miles de reacciones que operan simultáneamente lo hacen de manera que el sistema no colapsa ni genera un caos químico inescrutable. El hecho de que las reacciones metabólicas estén catalizadas por enzimas más o menos específicas y regulables permite transformaciones principales y flujos coordinados hacia los productos deseados, sin excluir la existencia de un ruido químico producido por la promiscuidad enzimática. Este ruido puede llegar a ser adaptativo o puede representar un daño metabólico que sea necesario reparar. En cualquier caso, la historia del estudio del metabolismo se ha centrado, en su primera etapa, en revelar la lógica estequiométrica y de diseño global de las rutas metabólicas. Primero fue de manera aislada, estudiando aquellos procesos que dominan determinadas situaciones fisiológicas (glicólisis durante las fermentaciones, ciclo del ácido cítrico durante el metabolismo aeróbico). Actualmente se puede abordar un estudio sistémico del metabolismo gracias a la convergencia de técnicas experimentales y computacionales.

Durante los últimos 3.500 millones de años, el metabolismo no ha dejado de evolucionar. Hubo momentos de exploración de nuevos territorios, como cuando el oxígeno molecular, producido por la fotosíntesis, abrió nuevas oportunidades de transformación química y, por tanto, nuevos metabolismos. A través de los eones, el metabolismo ha dado con una fracción de las soluciones químicas posibles. Los límites de la exploración evolutiva los han impuesto las condiciones físico-químicas (reacciones posibles en las condiciones ambientales y celulares concretas) y las restricciones históricas: una célula busca soluciones reutilizando estructuras y catalizadores preexistentes. Esta es una visión darwinista del reciclaje de mecanismos y dispositivos que François Jacob adaptó a la escala celular con su idea de bricolaje molecular. Un bricolaje que incluye el

aprovechamiento del ruido biológico, como ilustra la promiscuidad enzimática como materia prima de la evolución o la complementación metabólica entre consorcios de simbiontes edificada sobre el escape accidental de moléculas a través de las membranas celulares.

La vida se ha abierto paso hasta en las condiciones ambientales más sorprendentes, explotando la capacidad de un aprovechamiento energético y material suficiente para dar lugar a una biodiversidad extraordinaria. Suficiente, pero no necesariamente óptimo. Las soluciones químicas encontradas por la vida no suelen ser siempre las mejores posibles, sino las que bastan para la existencia y la reproducción. La exploración de la diversidad metabólica actual –la bioprospección– nos permite descubrir nuevos procesos y moléculas que pueden tener utilidad en nuestras vidas. El caso más obvio es el de los medicamentos, como los antibióticos. Pero también es de enorme interés conocer la dinámica de los procesos metabólicos, especialmente en humanos, para entender las alteraciones metabólicas asociadas con las patologías. Así pues, el enfoque predominante se basa en el estudio y descubrimiento del metabolismo terrestre, una fracción de lo químicamente posible.

¿Podemos alcanzar soluciones fuera de las exploradas por la evolución? De algún modo, se trata de extrapolar lo que la biotecnología ha estado haciendo desde hace siglos, primero a base de una estrategia de prueba y error, más tarde con un mayor fundamento genético y bioquímico. La ingeniería metabólica ha permitido modificar aquel metabolismo descrito previamente aplicando parámetros de optimización acordes con las necesidades prácticas, limitándose hasta épocas recientes a trasplantar módulos metabólicos de un organismo a otro. Sin embargo, un enfoque inconformista de la biología sintética consiste en ir más allá de lo que nos ofrece la naturaleza y reinventar el metabolismo, con soluciones químicas que, que sepamos, jamás fueron el resultado de ninguna trayectoria evolutiva. El transmetabolismo –neologismo propuesto por el autor junto con los microbiólogos Víctor de Lorenzo y Pablo Níkel– son los procesos químicos llevados a cabo por sistemas metabólicos sintéticos que exploran nuevas regiones de lo químicamente posible. En el futuro se podrán diseñar metabolismos a la carta que, por ejemplo, resuelvan problemas de contaminación química ambiental, mejoren la productividad agrícola o mitiguen la acumulación de gases de efecto invernadero. En este sentido, me pregunto si el transmetabolismo podría ayudar a aplazar nuestra extinción.

Peretó, J. (2018) Justificación general: la dimensión evolutiva y estequiométrica del metabolismo. En: *Proyecto Docente*. Universitat de València.

2 The discussion on the meaning and significance of “secondary metabolism” is relevant in the context of a course on metabolism and the biotechnological applications of this knowledge.

It has long been recognized that plants and animals profoundly affect the characteristics of one another during the course of their evolution. We have only to consider the case of pollination of flowering plants to see the truth of this statement. Both flowers and their pollinators, whether they are insects, birds, or bats, have developed specialized adaptations to promote success of the operation. Flowers attracting animals with bright colors ensure that pollinators enter their tissue by rewarding the animals with pollen and nectar. This form of signaling between plants and animals is really the essence of biochemical coevolution.

Chemical signals between organisms are largely due to compounds that are not part of primary metabolism. Primary metabolites are absolute requirements for existence and reproduction, such as carbohydrates, nucleic acids and their components (nucleotides, nucleosides, etc.), proteins, amino acids, and lipids. Chemical signals emitted by organisms have been described in the literature as "secondary compounds" or "secondary metabolites," presumably to indicate that they are not as important as primary metabolites. Organic chemists have traditionally referred to these compounds as "natural products" to distinguish them from synthetic compounds made in laboratories. From the ecological and evolutionary point of view, the term semiochemicals (from the Greek word *semio*, meaning "signal" or "sign") is more appropriate than these other terms, since it emphasizes their role as chemical signals between organisms. Semiochemicals can serve many ecological roles, such as prey or mate attractants, feeding deterrents, or warnings and deterrents to potential predators or

competitors. Semiochemicals are produced by an organism; then either they are emitted into the environment or they remain within the tissues of the producer and eventually evoke a response in another organism.

[...] In this chapter we examine only those semiochemicals produced by plants. We ask how some may have changed during the course of evolution in response to the insects that coevolved with the plants. Among extant organisms, we find that land plants, by and large, produce the largest number of semiochemicals. Plant semiochemicals can be assigned to severa! different classes on the basis of their chemi-cal structures. We consider here one or two examples from the three majar classes: phenols, terpenes, and alkaloids (table 1). The nearly 11,000 known structures are only the tip of the iceberg; the number of compounds present in nature is probably 10 or even 20 times greater.

Table 1
Major classes of semiochemicals.

	Approximate number of structures known
Phenols	
flavonoids[a]	} 1,000
tannins	
lignins	not well characterized (structurally complex)
Terpenes[b]	
monoterpenes	1,000
sesquiterpenes	600
diterpenes	1,000
triterpenes	800
tetraterpenes	350
polyterpenes	?
Alkaloids	6,000

Excerpt from: Swain, T. (2000) Chemical signals from plants and phanerozoic evolution. In: *Environmental Evolution: Effects of the origin and evolution of life on planet Earth* (L. Margulis, C. Matthews, A. Haselton, eds.), ch. 11, pp. 201–218. Cambridge: MIT Press.

To further complicate matters, in 1891, the German Physiological Chemist Albrecht Kössel (who won the 1910 Nobel Prize for Medicine) unknowingly aided this fission of the subject of metabolism when he proposed that plants had two distinct types of metabolism 'primary' and 'secondary'. He proposed that *primary metabolites* were involved in basic processes of the cell and were common to many organisms. *Secondary metabolites* were made by distinct pathways limited to some organisms; hence they served a less vital role Kössel's primary metabolism was the subject taken over by biochemists and his secondary metabolites were the Natural Products being studied by chemists. A consequence of Kössel's categorization is that the terms Natural Products and Secondary metabolites are still used synonymously (Figure 1.3). The former term is used mainly by chemists and the lalter mainly by scientists with a biological background. Could there be a better example of the fragmentation of a scientific subject than one where two different groups have the different names for the same subject?

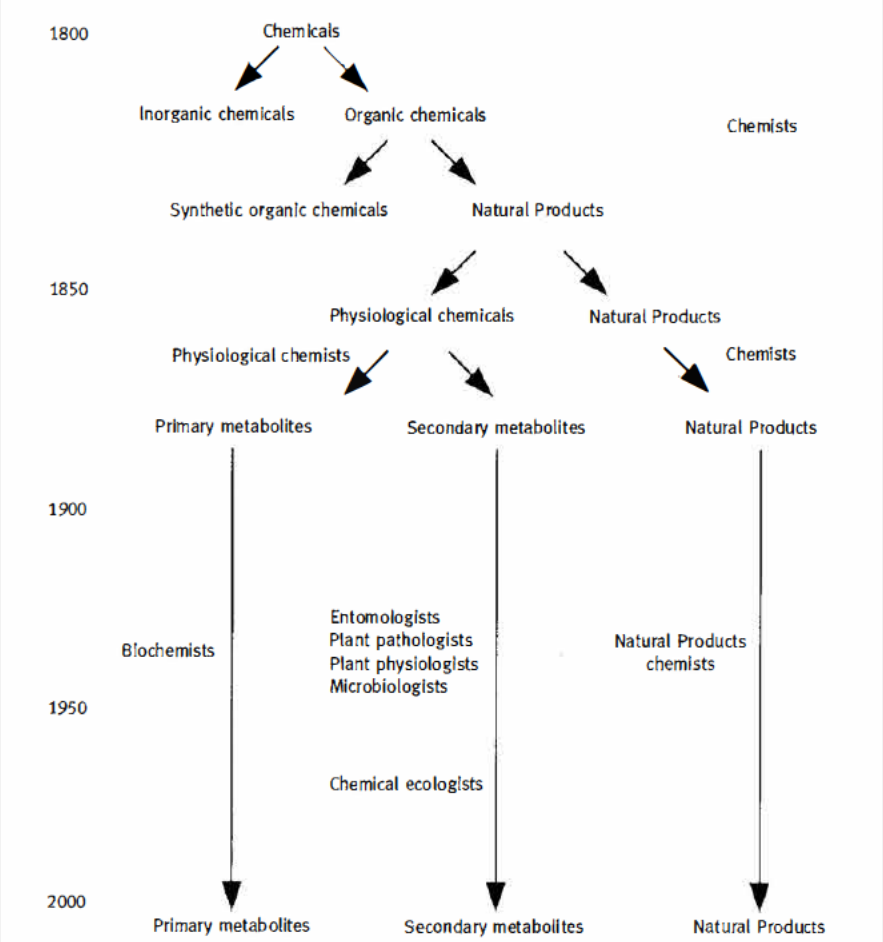


Figure 1.3. The time scale of fragmentation of the study of NPs—a fragmentation that has resulted in the current unsatisfactory situation where different groups of scientists use their own names for the same class of substances.

Excerpt from: Firn, R. (2010) *Nature's Chemicals. The Natural Products that Shaped Our World*. Oxford: Oxford University Press.

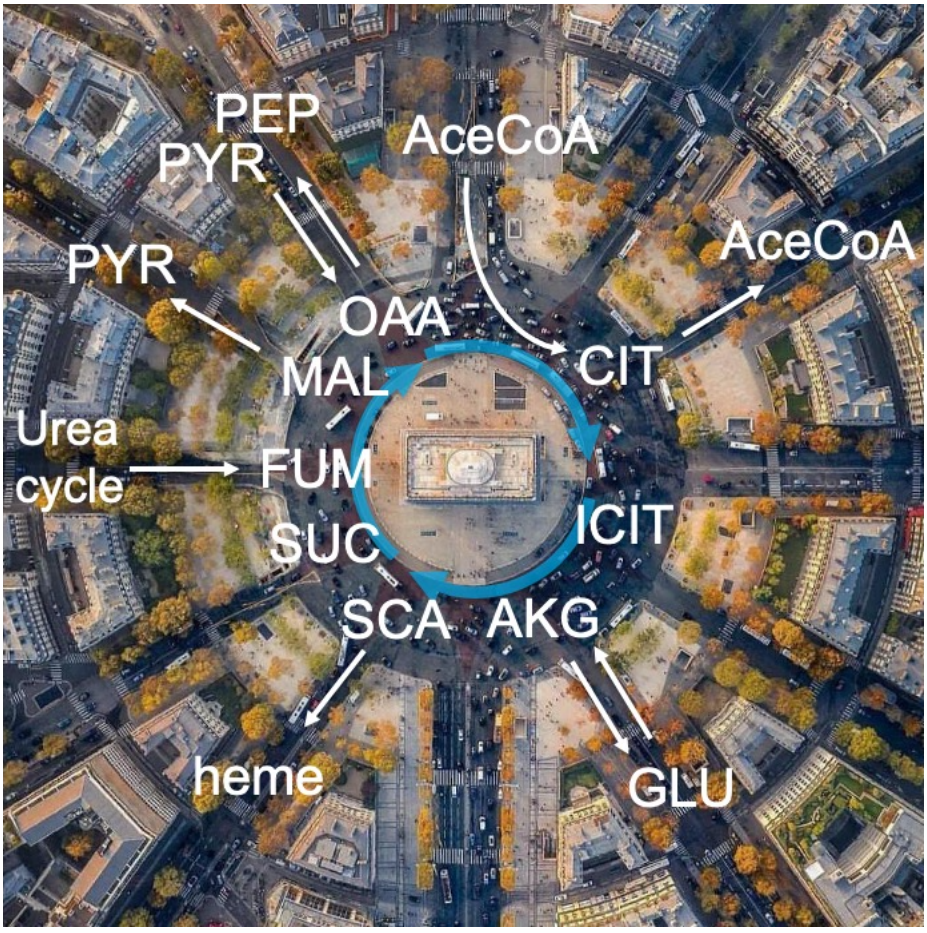
Químicamente, las cualidades que distinguen a una especia son sus raros aceites esenciales y sus oleorresinas, compuestos muy volátiles que dan su sabor, aroma y propiedades conservantes. Los botánicos consideran estas substancias componentes secundarios porque desempeñan un papel secundario en el metabolismo de la planta, lo que equivale a decir que no participan en la fotosíntesis ni intervienen en la absorción de nutrientes. Pero secundario no es sinónimo de irrelevante. Por lo general, se acepta que su *raison d'être* es una forma de respuesta evolutiva, la manera que tiene la planta de contrarrestar las amenazas de los parásitos, las bacterias, los hongos o los patógenos presentes en el medio tropical donde habita. En pocas palabras, la química de las especias –lo que en último extremo las convierte en especias– es, en términos evolutivos, lo mismo que las púas para el puercoespín o la concha para la tortuga. En su estado natural, la canela es una elegante armadura; el apetitoso aroma de la nuez moscada es, para determinados insectos, un montón de toxinas. La ironía elemental de su historia es que el atractivo de las especias es (desde el punto de vista de la planta) un tiro por la culata darwiniano. Lo que hace tan atractivas a las especias para el ser humano es repulsivo para otros miembros del reino animal.

Extracto de: Turner, J. (2004) *Spice: The History of a Temptation*. Trad. de M. Temprano García. Acantilado, Barcelona 2018.

Cautionary note on these handouts *By no means do these Study Guides aim to present all the required knowledge on metabolism during the course. Please, according to your previous background and your specific necessities, refer to the cited bibliography.*

2

Acetyl CoA and the CAC



Refresh your background

- ✓ Enzymology (mechanisms, cofactors and coenzymes, stereochemical properties).
- ✓ Organic functions and the effect of a carbonyl group on the decarboxylation of an organic acid.
- ✓ Aerobic *vs.* anaerobic growth conditions in microorganisms, autotrophic *vs.* heterotrophic growth.

Keywords

- ✓ Amphibolic pathway
- ✓ Anaplerosis
- ✓ Cataplerosis
- ✓ Citric acid cycle (CAC)/Tricarboxylic acids (TCA) cycle or Krebs cycle
- ✓ Enzyme enantioselectivity
- ✓ Glyoxylate cycle
- ✓ Glyoxysome
- ✓ Horseshoe mode
- ✓ Lipoic acid
- ✓ Ogston model
- ✓ Prochiral
- ✓ Reverse or reductive CAC/TCA/Krebs cycle or Arnon cycle
- ✓ Thiamine
- ✓ Thioester bond

Learning objectives

- ✓ Acetyl CoA as metabolic crossroads
- ✓ CAC as an example of a simple catalytic cycle
- ✓ Role of anaplerotic and cataplerotic reactions in a catalytic cycle
- ✓ Amphibolic character of the CAC cycle
- ✓ Stereochemical properties of the CAC cycle
- ✓ CAC as a reversible roundabout: oxidative and reductive modes
- ✓ Horseshoe configuration: the biosynthesis of CAC intermediates under anaerobic conditions
- ✓ The glyoxylate cycle avoids decarboxylation and allows the net synthesis of C4 from C2
- ✓ Autocatalytic character of the reverse CAC and glyoxylate cycle

Contents

Overview

The oxidation of biomolecular building blocks can be considered as a three-stage process. The first is the formation of activated fragments of two carbons: the acetyl group of acetyl CoA. The second is the complete oxidation of these two carbon atoms by the citric acid cycle (CAC) and the third is the reoxidation in the electronic transport chain of the cofactors reduced in the previous stages. In fact, the first stage involves a range of metabolic pathways and acetyl CoA behaves as a true metabolic crossroads. Besides the abovementioned function, the Krebs cycle also generates numerous intermediaries for biosynthesis, which makes it amphibolic, that is, it works both catabolically and anabolically. Since the CAC is an example of simple catalytic cycle, modulating its rate is achieved by changing the level of its intermediates: the restitution of intermediaries in the cycle is carried out by means of anaplerotic reactions, whereas their consumption is attained by cataplerotic processes. The operation of a canonical CAC occurs under oxidative conditions, although in a diversity of microorganisms, the cycle can run in the reductive direction, allowing the net synthesis of acetyl CoA from CO_2. Under anaerobic (fermentative) conditions, the cycle reactions can adopt a horseshoe configuration that makes the cycle intermediates available without the participation of an external oxidant. Finally, avoiding the decarboxylating steps of the CAC and with the help of two additional enzymes, the glyoxylate cycle allows the net synthesis of C4 intermediates from acetyl CoA. Both the reverse CSC and the glyoxylate cycle are examples of autocatalytic cycles.

2.1 Origins and fates of acetyl CoA

Many of the catabolic pathways eventually give rise to a two-carbon unit activated with coenzyme A, acetyl CoA.

Slide #2 | Acetyl CoA – a closer look

CoA is a complex coenzyme, but the reactive moiety is the free thiol group – SH. In acetyl CoA, the thioester bond between the CoA thiol group and the carboxylate group of acetate is a bond with high group transfer potential (namely, its hydrolysis is associated with a high release of free energy; some authors refer to this with the unfortunate term of "high energy bond"). CoA is involved in numerous acyl group transfer reactions.

Acetyl CoA is a crossroads of several metabolic pathways. It can be generated by the β-oxidation of fatty acids, by the degradation of some amino acids, and by the oxidative decarboxylation of pyruvate, the final product of glycolysis. On the other hand, under oxidative conditions, the catabolic fate of acetyl CoA is the complete oxidation of the acetyl group to CO_2 in the citric acid cycle (CAC). However, depending on the physiological situation of the organism, acetyl CoA can be used for the biosynthesis of fatty acids, sterols or ketone bodies. The direct synthesis of acetyl CoA from acetate, which consumes ATP pyrophosphorolytically and is catalysed by acetate thiokinase or acetyl CoA-synthase, can occur in some animal tissues, seeds and micro-organisms.

Slide #3 | Acetyl CoA – origins and fates

Mitochondrial and nucleo-cytosolic bioenergetic metabolism of acetyl CoA in mammalian cells.

Slide #4 | CAC – respiratory context

The CAC as an oxidative pathway allows the total oxidation of acetate to CO_2, and the corresponding electrons are channeled to the respiratory chain (via redox coenzymes). In mammalian cells, the CAC occurs in the mitochondria (the enzymes are soluble, located in the matrix, except for succinate dehydrogenase, also known as Complex II of the electron transport chain. In prokaryotes, the CAC occurs in the cytoplasm.

2.2 Oxidative decarboxylation of pyruvate

The oxidative decarboxylation of pyruvate to acetyl CoA is catalyzed by a multi-enzyme complex called pyruvate dehydrogenase complex (PDC). PDC consists of three catalytic activities: pyruvate dehydrogenase (E1), dihydrolipoamide-S-acetyltransferase (E2) and dihydrolipoamide-dehydrogenase (E3). In *E. coli*, it is a 4,600-kDa complex formed by 24 polypeptide chains of E1, 24 of E2 and 12 of E3. The mitochondrial bovine heart muscle chain is 8,400 kDa and consists of 30 tetramers of E1 and 6 dimers of E3, around a nucleus of 60 monomers of E2.

Slide #5 | Acetyl CoA – synthesis by the PDH complex

This enzyme complex is a good example of the metabolite-channelling effect. The mechanism of this enzyme illustrates the role of thiamine in the catalysis of decarboxylation of an α-ketoacid, such as pyruvate.

Slide #6 | PDC – multi-step mechanism and metabolite channelling

Slide #7 | PDC – multi-step mechanism and metabolite channelling (video)

2.3 The CAC is a simple catalytic cycle with a chemical logic behind

The CAC is a simple catalytic cycle, but one highly connected with the rest of the metabolic network (see below anaplerotic and cataplerotic reactions, section 2.6). Remember from chapter 1 our discussion on catalysis in metabolism: macromolecular catalysts (enzymes) and multi-molecular catalysts (catalytic cycles) coexist in living cells. The CAC serves as the catalyst for the complete oxidation of acetate (C2) to CO_2.

Slide #8 | CAC – a roundabout, not a carrousel

As discussed in chapter 1, analogies and metaphors not only are useful pedagogical tools but may also influence research programs. Regarding the CAC, the most popular iconography found in textbooks corresponds to the rounded distribution of reaction steps, usually without making explicit in the drawing the multiplicity of connections in and out the cycle (the so-called anaplerotic and catapletoric reactions, section 2.6). The interconnectedness of the CAC with the rest of the metabolic network justifies the analogy of the cycle with a traffic roundabout, much more realistic and meaningful than the immediate metaphor of a carrousel.

Slide #9 | CAC – the chemical logic

As stated by David Metzler, the CAC is "a clever way to cleave a reluctant bond" (Metzler[2], p. 950). Oxidation of the chemically resistant two-carbon acetyl group to CO_2 presents a chemical problem. Cleavage of a C-C bond occurs most frequently between atoms that are α or β to a carbonyl group. Such β cleavage is clearly impossible within the acetyl group. The only other common type of cleavage is that of a C-C bond adjacent to a carbonyl group (α cleavage), a thiamine-dependent process. However, α cleavage would require oxidation (hydroxylation) of the methyl group of acetate. Although many hydroxylation reactions occur, they are rarely used in the central pathways. According to Metzler, perhaps this is because the overall yield of energy obtainable via hydroxylation is less than that gained from dehydrogenation and use of an electron transport chain. In addition, think about how difficult hydroxylation reactions are in the absence of oxygen, at the beginning of metabolic evolution (consider that the CAC is, allegedly, one of the oldest pathways). The solution, then, to the chemical problem of oxidising acetyl groups is one very commonly found in nature: a catalytic cycle. Although direct cleavage is impossible, the two-carbon acetyl group of acetyl CoA can undergo a Claissen condensation with a second compound that contains a carbonyl group. The condensation product has more than two carbon atoms, and a β cleavage to yield CO_2 is now possible.

2.4 Enzymatic steps of the CAC

Szent-Györgyi described the catalytic effect that the addition of succinate, fumarate or oxaloacetate has on oxygen consumption by muscle tissue when it oxidizes glucose. Krebs and Johnson added citrate and α-ketoglutarate (2-oxoglutarate) to the list. The synthesis of citrate from oxaloacetate and a two-carbon molecule, as well as the cyclic nature of the pathway, were the keys to the explanation given by Krebs and Johnson (1937) to the previous observations – together with Henseleit, Krebs had already described the urea cycle in 1932, and with Kornberg, he would go on to describe the glyoxylate cycle in 1957.

Thus, the first step of the CAC is the binding of the two activated carbons of acetyl CoA to a four-carbon substrate, oxaloacetate, to form a six-carbon molecule, citrate. The subsequent reactions oxidize two carbon atoms to CO_2 and regenerate the cycle's priming substrate. Stoichiometrically, three molecules of NADH (or 2 NADH + 1NADPH), one of $FADH_2$ and one of GTP (or ATP) are generated.

The first of the two oxidative decarboxylations of the cycle is catalysed by isocitrate dehydrogenase, which transforms the isocitrate into α-ketoglutarate. The reaction catalysed by α-ketoglutarate dehydrogenase is similar to the pyruvate dehydrogenase reaction (and the enzymes are homologues, that it, they share a common evolutionary ancestor and are functionally interchangeable). First, an oxidative decarboxylation of an α-keto (2-oxo) acid and then the formation of an acyl CoA. At this point in the cycle, two carbons have entered as acetyl CoA and two have left as CO_2. The rest of the cycle from this stage on allows the regeneration of the oxaloacetate priming substrate from succinyl CoA. The succinyl CoA-synthase allows coupling the high transfer potential of the thioester in the succinyl CoA with a substrate level phosphorylation. Succinate dehydrogenase – also known as respiratory complex II – is a flavoprotein that introduces the electrons of the succinate oxidation directly into the electronic transport chain, at the level of quinones. The stereospecific hydration of the double bond of fumarate to form L-malate is catalysed by fumarate hydrate or fumarase. Malate dehydrogenase catalyses the final reaction of the CAC, the formation of oxaloacetate by oxidation of L-malate.

Slides #10 and #11 | CAC – enzymatic steps
Slide #12 | CAC – Overall stoichiometry

Read the description of the CAC enzymes by David S. Goodsell (*Molecule of the Month*, October 2012) Citric acid cycle: Eight enzymes form a cyclic pathway for energy production and biosynthesis.

Citrate synthase catalyses the condensation of one acetyl CoA and one oxaloacetate molecule to form citrate. The functional group of the citrate molecule is a very difficult tertiary alcohol to oxidise. The intervention of aconitase, an enzyme that catalyses the reversible

isomerisation of citrate to form isocitrate, generates a substrate with a more easily oxidizable secondary alcohol. The reaction of aconitase involves successive dehydration and hydration through an intermediate, cis-aconitate. Of the four possible diastereoisomers of the isocitrate, only one is produced. Citrate, the substrate of aconitase, is a symmetrical molecule. However, using radioactive markers, it has been shown that aconitase acts on citrate asymmetrically, always dehydrating the part of the molecule that comes from oxaloacetate. Ogston suggested in 1948 that the chiral active centre of an enzyme could make an achiral substrate behave asymmetrically: citrate binds to the active centre of aconitase at three different points; in this way, the enzyme can differentiate the two groups $-CH_2-COO^-$ from citrate, as they are on different planes of space. The molecules that behave like citrate are called prochiral.

Slide #13 | CAC – labelling experiments and stereochemistry

Slide #14 | CAC – more stereochemical details and the Ogston postulates

2.5. Amphibolic character of the CAC

The CAC does not function exclusively as a catabolic pathway for acetyl CoA oxidation, but is also an important source of biosynthetic pathway precursors. Since the cycle can be seen as both a catabolic and anabolic pathway, it is said to be an amphibolic pathway. Thus, citrate participates in the biosynthesis of fatty acids and steroids. Thanks to the presence of specific transporters, in eukaryotic cells, citrate can pass through the internal mitochondrial membrane and generate in the cytosol one molecule of oxaloacetate and one of acetyl CoA, which will be used in the synthesis of fats. The α-ketoglutarate can be reversibly converted to glutamate, which can become part of proteins or be a precursor for other amino acids or nucleotides. Succinyl CoA can be condensed with glycine to start the biosynthesis of porphyrins. Oxaloacetate is convertible to aspartate – which can be used in the synthesis of urea, proteins and nucleotides – and is also a precursor of sugars through the gluconeogenic pathway.

Slide #15 | CAC – amphibolic character

IF THE CAC IS AN AEROBIC PATHWAY, in the sense that it requires the reoxidation of coenzymes in order to maintain function, how does an anaerobic organism obtain these biosynthetic precursors? How do the oxidative decarboxylations that generate acetyl CoA and α-ketoglutarate take place? We have already seen that one molecule of NADH is formed for each molecule of acetyl CoA generated from pyruvate, and two for each of α-ketoglutarate. In the absence of oxygen or an alternative electron acceptor, the electrons can be used for the reduction of some metabolite. The solution in *E. coli*, and probably in other species of bacteria under anaerobic conditions, is to use parallel reduction reactions, with the cycle adopting an open, horseshoe configuration (from

oxaloacetate to succinyl CoA, running this part of the CAC in reverse) thus consuming these electrons.

These reactions consume NADH produced in the reactions described above, while allowing the synthesis up to succinyl CoA. In *E. coli*, there is a NAD^+-dependent fumarate reductase different from FAD-dependent succinate dehydrogenase and upregulated only in the absence of O_2. Therefore, the starting materials needed for the biosynthetic pathways can be obtained anaerobically from sugars. These reactions, now used by bacterial species, were most likely to occur when there was no O_2 in the atmosphere, over two billion years ago. The appearance of fairly strong external oxidants (O_2) and the addition of a single enzyme (for instance, succinyl CoA synthase) would be sufficient to explain the evolutionary origin of the CAC. On the other hand, with sufficiently strong reducers – for example, ferredoxin – the cycle could be made to work in reverse, and instead of oxidising carbon to CO_2, it would reduce this gas to organic matter (see below).

Slide #16 | CAC – horseshoe mode or how to obtain intermediates in anoxia

2.6 Anaplerotic and cataplerotic reactions

The reactions that replace the cycle intermediaries consumed in the biosynthetic pathways are called anaplerotic.

In mammals, one such reaction is catalysed by pyruvate carboxylase:

$$\text{pyruvate} + HCO_3^- + \text{ATP} + H_2O \rightarrow \text{oxaloacetate} + \text{ADP} + \text{Pi}$$

The distribution of pyruvate between decarboxylation to give acetyl CoA and carboxylation to form oxaloacetate is the decision point for using this metabolite in oxidation and obtaining ATP or for biosynthetic purposes.

In plants and bacteria, oxaloacetate is introduced into the cycle by the carboxylation of phosphoenolpyruvate catalysed by phosphoenolpyruvate carboxylase:

$$\text{phosphoenolpyruvate} + CO_2 \rightarrow \text{oxaloacetate} + \text{Pi}$$

Transamination reactions that transfer the α-amino group from an amino acid to a α-keto acid can also be considered as anaplerotic reactions: the deamination of glutamate and aspartate yields, respectively, α-ketoglutarate and oxaloacetate.

Slide #17 | CAC – anaplerotic reactions

Note that the net effect of anaplerotic reactions, when increasing the concentrations of the cycle intermediates, is the increase of the concentration of the catalytic cycle. Thus, the net rate of acetyl CoA oxidation increases.

If intermediates can be added to the CAC, it is equally important to remove them to avoid the accumulation of anions in the cell. Reactions that deplete intermediaries of the cycle are called catapletoric, and the net result will be the reduction of the rate of acetyl CoA oxidation. There

areseveral cataplerotic enzymes; these include PEPCK (phosphoenolpyruvate carboxykinase is an important gluconeogenic enzyme that transforms OAA in PEP), Glu dehydrogenase, and Glu and Asp aminotransferase (transaminase). Cataplerosis is required when catabolic processes need some of the CAC reactions during the overall transformation, for example, the complete oxidation of glutamine by cancer cells involves cataplerotic reactions.

Slide #18 | CAC – cataplerotic reactions

2.7 Building biomass from acetate: the glyoxylate cycle

Animal tissues do not have a metabolic pathway for the net transformation of acetyl CoA into glucose. Instead, in other systems – yeasts, bacteria, seeds of germinating plants – the glyoxylate cycle allows the net synthesis of malate – and thus of oxaloacetate and sugars – from acetyl CoA derived from the degradation of fatty acids or of two-carbon compounds, for instance, acetate or ethanol. The net synthesis of C4 compounds from C2 is possible by combining the enzymes isocitrate lyase and malate synthase with some of the reactions of the CAC, bypassing the decarboxylative steps of the cycle.

Slide #19 | Glyoxylate cycle – building biomass from acetate
Slide #20 | Glyoxylate cycle – as an autocatalytic cycle
Slide #21 | Glyoxylate cycle –physiological context example in plants

2.8 Autotrophy by reversing the CAC

A reductive or reverse CAC (also known as Arnon cycle) is the mechanism used by some autotrophic bacteria to fix carbon. These examples also illustrate how metabolic pathways may have evolved from existing reaction sequences, which are very conveniently used in different ways and offer new metabolic possibilities to the cell: what was first, the oxidative or the reductive CAC? This is still a matter of controversy.

Slide #22 | CAC in reverse – autotrophic pathway
Slide #23 | CAC *literally* in reverse – autotrophy against the thermodynamic dogma

Bibliography and resources

Textbooks

Metzler[2], vol. 1 ch. 10, vol. 2 ch. 17.
Nelson and Cox[7], ch. 16.
Peretó et al.[5], ch. 14.

Stryer[9], ch. 17.

Papers and book chapters

Peretó J (2011) Origins and evolution of metabolisms. In: *Origins and Evolution of Life. An Astrobiological Perspective* (Gargaud M et al., eds.), Cambridge University Press

Peretó J (2012) Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem Soc Rev* **41**:5394–5403

On the net

The Nobel Foundation is a good source of information on the awarded scientists. Learn about some biochemists related to this chapter: Hans Krebs, Fritz Lipmann, and Albert Szent-Györgyi.

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ2.1 Find the EC numbers for acetate thiokinase and acetyl CoA synthase, write the adjusted reactions catalysed by these enzymes, and give some hints about their phylogenetic distribution.

SQ2.2 *Learn by mapping* Build a concept map of chapter 2.

SQ2.3 *Learn by drawing* Draw a complete pathway for the net synthesis of oxaloacetate from: (a) acetyl CoA; (b) CO_2 via a reverse CAC.

SQ2.4 Formulate a hypothesis on why malonate inhibits succinate dehydrogenase. How will the concentrations of CAC intermediates change immediately after the addition of malonate?

SQ2.5 Define *suicide inhibitor* and apply the definition to the inhibition of aconitase by fluoroacetate.

SQ2.6 Match the following two columns:

1. Anaplerotic reaction	a. Arnon cycle
2. Autocatalytic cycle	b. Citrate
3. Cataplerotic reaction	c. Glu deamination
4. Prochiral molecule	d. PEP carboxykinase

SQ2.7 Why does the CAC qualify as an aerobic pathway although no O_2 participates in any reaction of the cycle?

SQ2.8 Before any oxidation can take place in the CAC, citrate must be isomerized into isocitrate. Why?

SQ2.9 Distinguish between catalytic coenzymes and stoichiometric coenzymes in the pyruvate dehydrogenase complex.

SQ2.10 How does the decarboxylation of α-ketoglutarate resemble that of pyruvate decarboxylation? Draw the product of the reaction between pyruvate (or α-ketoglutarate) and thiamine pyrophosphate (TPP, carbanionic form) catalysed by the E1 component of the multienzyme complex, and compare with the structure of a β-ketoacid undergoing spontaneous decarboxylation.

SQ2.11 *Be a cultivated biotechnologist* Hans Krebs (1900-1981) studied and completed his PhD in Germany in 1925. Then, he started working in the University of Berlin, at Otto Warburg's lab. However, in 1933 he migrated to England, where he lived and worked until his death. Why did he leave Germany?

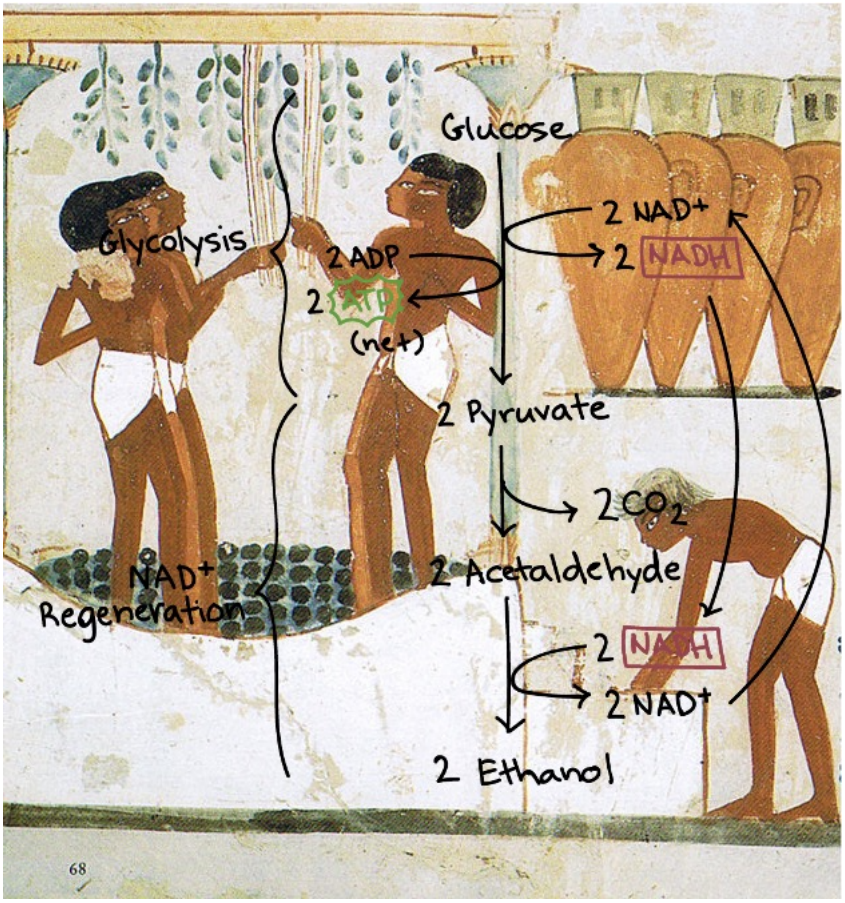
Problem-Based Learning (PBL)

PBL2 offers the opportunity to learn and consolidate the basic concepts related to chapter 2.

Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the required knowledge on metabolism during the course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

3

Central carbon pathways



Refresh your background

- ✓ Enzymology concepts (cofactors and coenzymes, enzymatic classes, kinetic parameters, thermodynamic constraints).
- ✓ Chemical functions in organic compounds and their redox status.
- ✓ Microbial metabolic modes (autotrophy, heterotrophy), fermentations.

Keywords

- ✓ Arnon cycle (reverse/reductive CAC)
- ✓ Biotin
- ✓ Calvin-Benson (reductive pentose phosphate) cycle
- ✓ Cori cycle
- ✓ Dismutation (or disproportionation)
- ✓ Embden-Meyerhof-Parnas (EMP) pathway
- ✓ Entner-Doudoroff (ED) pathway
- ✓ Fermentation (biochemical sense)
- ✓ Glycogen
- ✓ Hetero- and homopolysaccharide
- ✓ Pareto front (or surface)
- ✓ Pentose phosphate/Hexose monophosphate pathway
- ✓ Photorespiration
- ✓ Shuttles (for cytosolic NADH)
- ✓ Starch
- ✓ Sulphoglycolysis
- ✓ Wood-Ljundahl pathway

Learning objectives

- ✓ Glycolysis (EMP and ED) and gluconeogenesis: enzymatic steps and connections with the rest of the network
- ✓ Pareto optimality in biology: application to glycolytic diversity (why EMP and ED?)
- ✓ How sugars other than glucose can be catabolized
- ✓ Fermentation in the biochemical sense, and diversity of fermentation pathways
- ✓ Some applications of fermentations (you are contemplating the earliest birth of biotechnology)
- ✓ Pentose phosphate pathway: oxidative and non-oxidative branches
- ✓ Stoichiometric combinatorics in central pathways
- ✓ Structure and metabolism of glycogen/starch
- ✓ How to make biomass from CO_2 through a diversity of autotrophic pathways: recognizing the autocatalytic character of some pathways

Contents

Summary

Organisms share, almost universally, some central metabolic transformations that give access to biomass precursors. In addition to the CAC (see chapter 2) these pathways include different versions of glycolysis (the canonical Embden-Meyerhof-Parnas EMP, the Entner-Doudoroff ED pathways, and other variations) and the pentose phosphate pathway (PPP). Glycolytic pathways are diverse ways of utilization of hexoses and other sugars that allow obtaining energy and biosynthetic precursors under both anaerobic and aerobic conditions. The diversity of stoichiometric yields of glycolytic pathways (EMP *vs.* ED, aerobic *vs.* anaerobic) offers an excellent study case on Pareto optimality in metabolic systems. The synthesis of glucose through gluconeogenesis, from non-sugar precursors, requires energy and supplies glucose to fasting animals, and glucose-6-phosphate, to microorganisms that grow in the absence of glucose. The oxidative branch of PPP allows the conversion of hexose into pentoses (the starting point for nucleotide biosynthesis, for instance) and the synthesis of NADPH (reducing power), required for biosynthetic reductions and coping with oxidative stress. On the other hand, the non-oxidative branch of PPP allows the interconversion of hexoses and pentoses. There is a diversity of pathways allowing the synthesis of central metabolites from CO_2 (autotrophic pathways), a key step in the global circulation of C in the biosphere and perhaps an opportunity to discover new ways to alleviate the climatic crisis.

3.1 Central C metabolism in context

Before starting our journey around the pathways managing carbohydrates metabolism (the traditional name for molecules composed of C, H and O), it is worth providing some context: in the city of metabolism, where are these pathways located? What are the main connections with bioenergetics (ATP synthesis, redox status)? And how do these pathways furnish anabolism with the essential building blocks? Some diagrams in Metzler's textbook are good guides for this visit.

Slide #2 | Central C metabolism in context – the bioenergetics connections

Slide #3 | Central C metabolism in context – the biosynthetic connections

3.2 Glycolysis: A theme with variations

Glycolysis, in the so-called Embden-Meyerhof-Parnas (EMP) version of the pathway, allows the metabolic transformation of hexose into pyruvate, and serves as a basis for the anaerobic synthesis of ATP in many cells. This widely distributed pathway was elucidated in the 1930s, when attempting to answer two apparently different questions related to anaerobic metabolism: how does yeast produce ethanol from glucose? And how does skeletal muscle use glucose as an energy source for contraction?

Slide #4 | Glycolysis stages – the EMP pathway

Slide #5 | Upper glycolysis – stages 1 and 2

Slide #6 | Lower glycolysis – stage 3

Slide #7 | EMP glycolysis – the perfect ten?

Read the description of Glycolytic Enzymes by David S. Goodsell in his series "Molecule of the Month" (February 2004). Think critically about the section "A Perfect Ten": do you agree with his evolutionary view on glycolysis?

Importance of phosphorylated intermediates

Have you realised that metabolic intermediates are, under cellular conditions, either anionic (e.g., CAC intermediates) or organo-phosphate intermediates carrying negative charges? The anionic state has several consequences: (1) The plasma membrane supports an electric potential negative in the inside face; this fact, together with the absence of transporters for phosphorylated metabolites, favors the retention of these molecules inside the cell, despite the large difference in their intracellular and extracellular concentrations. (2) The energy released by the consumption of phosphoanhydride bonds is partially conserved in phosphoester bonds; additionally, phosphoanhydride compounds exhibit a high group transfer potential allowing substrate level phosphorylations. (3) The binding of phosphate residues by enzymes increases the potential specificity and affinity; many glycolytic enzymes are specific for the Mg (II) complexes of the phosphorylated intermediates.

In most textbooks, some elementary energetic calculations are proposed to compare the ATP yield of substrate-level phosphorylation and the complete oxidation of C (coupling glycolysis with the oxidative decarboxylation of pyruvate to acetyl CoA and a complete CAC. Several questions emerge. For instance, is it conceivable that other glycolytic designs could yield higher amounts of ATP per hexose? At least in theory (nobody has described natural representatives of such pathways), other glycolytic pathways are possible yielding up to 5 ATP per hexose, but they present several unsurmountable issues regarding their physiological and thermodynamic feasibility (see the discussion by Ng et al. 2019).

Under aerobic conditions, the glycolytic pathway is supplied with NAD^+ obtained from the oxidation of NADH in the respiratory chain, and the pyruvate is transformed into acetyl CoA by the multienzyme PDC. Finally, the catabolic fate of acetyl CoA is the complete oxidation to CO_2 via CAC. Back-of-the-envelope calculations indicate that we could expect 16 times more ATP per hexose in aerobic conditions, in comparison to anoxic operation of the glycolytic pathway – provided that there is an efficient way to oxidize NADH (see the section 3.4 on fermentations). In animal cells, the mitochondrial oxidation of NADH produced in the cytosol by glycolysis requires the use of so-called shuttles, since NADH cannot diffuse through the mitochondrial inner membrane. Two classes of shuttles are known: glycerol phosphate and malate-aspartate. The energetic yield of the former has to be corrected because the NADH electrons enter the transport chain at the level of quinones.

Slide #8 | Glycolysis – back-of-the-envelope energetic calculations

Slide #9 | NADH shuttles – an example of vectorial metabolism

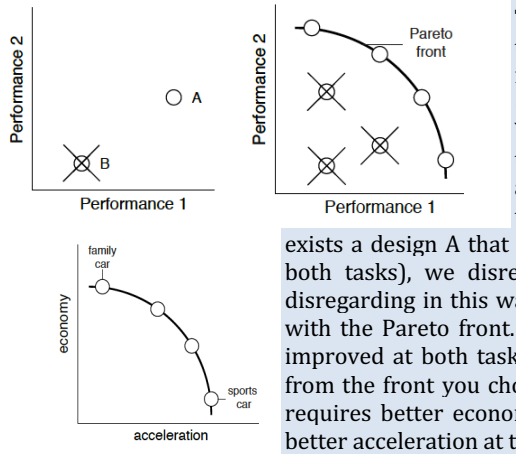
In the 1950s, a different version of glycolysis was described in bacteria, the so-called Entner-Doudoroff (ED) pathway, also known as semi-phosphorylative glycolysis: the ATP yield of the EMP pathway is 2ATP per glucose, whereas the ED pathway yields 1ATP/glucose. More recently, the ED pathway has been demonstrated also in plants. The existence of non-phosphorylative glycolysis – in bacteria and also in plants – raises the question about the optimality of the energetic yield of a catabolic pathway (see below), up to the point that some authors and textbooks consider ED a "minor version" of EMP, the "good one". In a moment you will realize that this kind of qualifications are pure nonsense.

Slide #10 | Glycolysis – the ED pathway

Slide #11 | Glycolysis – the good, the bad and the ugly

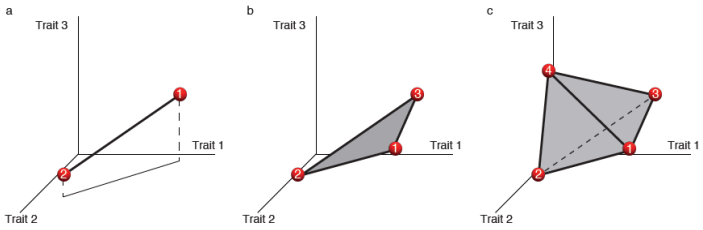
Pareto optimality in biological systems

In nature, biological systems usually need to carry out multiple tasks. Bacteria, for example, need to grow quickly and survive stresses. Multiple tasks lead to a fundamental trade-off: no design can be optimal at all tasks at once. No animal can fly like an eagle, swim like a dolphin and run like a cheetah. How can evolution optimize when multiple tasks coincide?

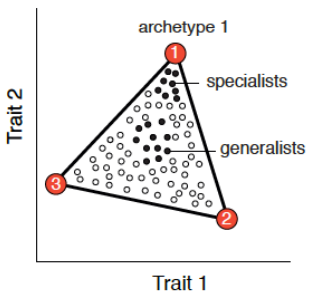


Engineers routinely need to solve this type of problem. They use an approach called *Pareto optimality*. Suppose that you want to design a car. The design specifications require performance at two tasks, let's say acceleration (time from 0 to 100km/h) and fuel economy (km/liter). You take all possible designs and plot them according to their performance at the two tasks. In this plot, whose axes are the two performances, each design is a point in the performance space. Now consider a design B. If there exists a design A that is better than B at both tasks (has higher performance at both tasks), we disregard design B. We say that B is dominated by A. By disregarding in this way all points which are dominated by another, we are left with the Pareto front. This is the set of designs that cannot be simultaneously improved at both tasks. This front is what engineers care about. Which design from the front you choose is based on the market niche of the car: a family car requires better economy at expense of acceleration, and a sports car requires better acceleration at the expense of economy.

PARETO OPTIMIZATION has been applied to several biological problems in the last decade (Schuetz et al. 2012, Shoval et al. 2012), including human breast cancer, animal behavior and shape, longevity-mass relationship, gene expression, ammonite shapes, complex networks, protein folds, proteome performance in different cell compartments, and yield-cost trade-off and resource allocation in microbial metabolism (Basan 2018). In biology, unlike engineering, we usually don't know what the tasks are in advance. We can make an educated guess, but we cannot be sure. Thus, we cannot directly use performance space to do Pareto optimality, because we don't know what tasks to compare, and, even if we did, we usually cannot evaluate the performance of each phenotype at each task. We simply plot the data in trait space, using all the traits that we can measure directly. The axes are the traits, and each phenotype is a point in this space. We will now see that the evolution under several tasks makes the data show particular geometric shapes.



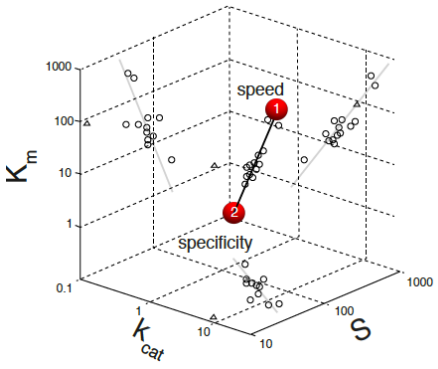
These shapes can help us discern the number of tasks, and even what these tasks might be. For example, when there are two tasks at play, the data will fall on a line segment with the ends occupied by the *archetypes*, i.e., the phenotype (trait combination) that is best at a particular task. The two archetypes of the segment give us clues about what the tasks are. If there are three tasks, we expect the optimal phenotypes to fall inside a triangle, whose three vertices are the three archetypes. If there are four tasks, the phenotypes will fall inside a tetrahedron. In general, a trade-off between k tasks will result in a Pareto front shaped as a polytope with k vertices (a polytope is the generalization of a polygon or polyhedron to any dimension). Each vertex is an archetype for one of the tasks.



The neat contribution of this approach is to help us discover what the tasks are directly from biological data. The polytopes can help us to infer the tasks: The phenotypes closest to a vertex should be specialists at something, and that something gives clues to what the task might be. Phenotypes near the centre of the polytope should be generalists. This approach of inferring the tasks from the geometric shape of the data in the trait space is called Pareto task inference.

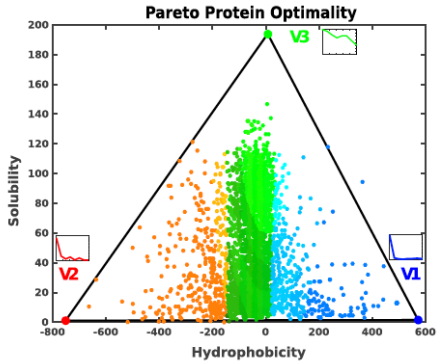
EXAMPLE 1 Rubisco seems to evolve under a speed-specificity trade-off

A protein can also have multiple tasks. For example, Rubisco, one of the most abundant proteins, is tasked with capturing CO_2 from the air and adding it to a sugar molecule that can be used to build biomass. Rubisco can be characterized by a trait space with four kinetic parameters. Two of these parameters are catalytic speed k_{cat} and affinity K_m to CO_2. The other two are catalytic speed and affinity, k'_{cat} and K'_m, for the main competitor of CO_2, oxygen O_2. Capturing O_2 instead of CO_2 is a mistake that requires energy to correct (photorespiration). To study trade-offs in Rubisco, Savir et al. (2010) compiled these four kinetic traits from 30 photosynthetic organisms. They found that the 30 Rubiscos fall approximately on a line in the 4D trait space. The figure shows the data in the space of three traits, k_{cat}, K_m, and the specificity $S = k_{cat}\ K'_m / k'_{cat}\ K_m$, together with the projections of the data on the three planes. At one end of the line segment are the fastest Rubiscos, which occur in organisms like corn, a C4 plant, that can concentrate CO_2. Since these plants reach a high CO_2 concentration inside their leaves, they do not need to worry about oxygen. At the other end are the slowest Rubiscos, which, conversely, bind CO_2 most strongly. These occur in organisms which do not concentrate CO_2 and face competition from O_2. Thus, this protein seems to evolve under a speed-specificity trade-off. Source: Alon (2019).



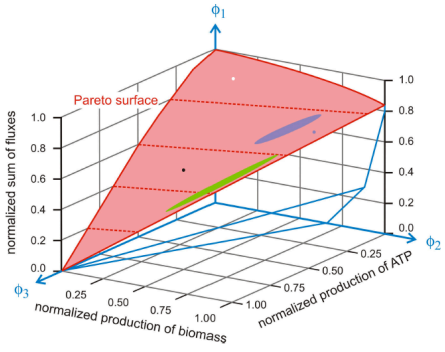
EXAMPLE 2 The *E. coli* proteome shows a solubility-hydrophobicity trade-off

Proteins have coevolved with cellular environments to improve or preserve their functions, maintaining at the same time the degree of hydrophobicity necessary to fold correctly and enough solubility to perform their biological roles. The *E. coli* proteome has been studied using a Pareto front analysis in the solubility-hydrophobicity space. The results indicate the existence of a Pareto optimal front, a triangle whose vertices correspond to archetypal proteins specialized in distinct tasks, such as regulatory processes, membrane transport, outer-membrane pore formation, catalysis, and binding. The vertices are further enriched with proteins that occupy different subcellular compartments, namely cytoplasmic, inner membrane, outer membrane, and outer membrane bounded periplasmic space. The combination of various enriching features offers an interpretation of how bacteria use the physico-chemical properties of proteins, both to drive them into their final destination in the cell and to have their tasks accomplished. The figure shows a scatter plot of the 3,172 proteins of the *E. coli* proteome. Each protein is represented as a point whose coordinates are the values of its hydrophobicity and solubility. Proteins whose points lie inside the triangle are the best compromise in the multi-objective optimization of the three tasks, which are better performed by the corresponding archetypes located at the three vertices. Points outside the triangle would have a better counterpart inside the triangle in at least one of the tasks. The RGB colors identify the distribution of the integral inner membrane (blue), outer membrane, and outer membrane bounded periplasmic (red) and cytoplasmic (green) proteins, which also characterize the vertices. Source: Koçillari et al. (2018).



EXAMPLE 3 Dominant fluxes in *E. coli* show a Pareto surface

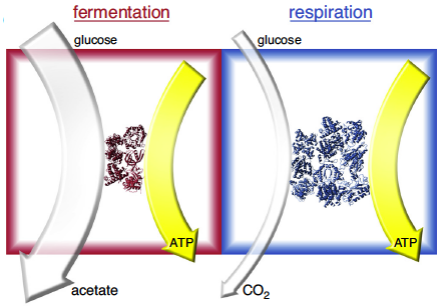
The surface of the Pareto optimal fluxes that best describe the metabolism of *E. coli*, the sum of all individual fluxes (ϕ_1), the production of biomass (ϕ_2), and the production of energy in form of the ATP yield (ϕ_3), was determined by means of a ^{13}C-based flux analysis (red surface corresponding to minimal ϕ_1, maximal ϕ_2, and maximal ϕ_3. The points and areas near the red surface were calculated from measured ^{13}C distributions in metabolites of *E. coli* under different growth conditions (aerobe cultures: excess glucose, blue; glucose deficiency, green; nitrogen deficiency, black, and, for comparison, anaerobe culture in white). All points lie systematically somewhat below the Pareto surface. The Cartesian coordinate system represents normalized fluxes: "1" means best possible fulfilment of the optimization criterion – minimum of the sum of fluxes,



maximum biomass, and maximum ATP yield – and "0" characterizes the solution at maximal distance from the optimum. The origin, accordingly, symbolizes the poorest possible solution. Source: Schuetz et al. 2012; the figure is redrawn by Schuster (2013).

EXAMPLE 4 Resource allocation may explain overflow metabolism

Metabolic reactions are catalysed by enzymatic machinery. Therefore, to achieve optimal growth rates, the cell must balance the fluxes from metabolic reactions and polymerization, while minimizing the investment in enzymatic machinery and investing the 'right' proteome fractions in these processes. For example, if the cell were to invest too many resources into catabolic processes, this would result in the production of more precursors and energy than could be processed by the ribosomes (and some of the investment in the catabolic processes would be effectively wasted). On the other hand, too much investment in ribosomes and biosynthetic machinery would result in a situation where biosynthesis could not be adequately supplied with precursors and energy (part of the investment in ribosomes and biosynthetic machinery would be effectively wasted). Moreover, because the cell has finite proteomic resources (fractions of total proteome), any increase in investment in one process must coincide with a corresponding decrease in investment in another. Maximum growth rate is therefore achieved when the cell invests optimal proteome fractions in different cellular processes, such that it balances fluxes from different processes while minimizing the resource investment in each of these processes. The figure illustrates the lean production hypothesis for the example of energy metabolism and acetate excretion taken from Basan et al. (2015). For the same ATP production flux (yellow arrow), fermentation consumes a much larger carbon flux (gray arrow) as compared to respiration. However, fermentation requires a smaller absolute investment in enzymatic machinery (red and blue proteins) to catalyze this flux, which enables faster growth. Other examples of overflow metabolism include the Warburg effect observed in cancer cells. Source: Basan (2018).



EXAMPLE 5 Are EMP and ED the best possible glycolytic pathways?

Thermodynamics and kinetics arguments, energy production (ATP yield), precursor synthesis, minimization of toxic intermediates, and minimal pathway hypotheses have been proposed as explanations for the prevalence of the ED and EMP designs over other would-be glycolytic pathways (Heinrich et al. 1997, Meléndez-Hevia et al. 1997, Court et al. 2015, Noor et al. 2010, Bar-Even et al. 2012). More recent studies have reaffirmed the minimization of enzymes production as a key driver of resource allocation optimization. For example, fast-growing cells have to invest more resources for the synthesis of growth-related proteins (i.e., proteins associated with translational and transcriptional machinery). Instead of simply making proportionally more protein to accommodate higher growth requirements, they often shift metabolism towards pathways with more modest catalytic resources per unit of growth at the expense of energy efficiency. Basan et al. further verified that *E. coli* switches from respiration to the more proteome-efficient fermentation under high growth rate conditions. Cellular metabolism has been shaped by evolution to ensure that carbon catabolic pathways are carefully selected to be in tune with both growth rate requirements and resource availability. Optimal glycolytic pathways must thus be able to balance high ATP production capacity while generating important intermediates and redox molecules at minimal proteome cost. These requirements are in direct conflict with one another, requiring the establishment of Pareto optimal curves to decipher the relative "weights" between objectives that nature responds to when selecting different pathway designs. In that context of minimizing proteome costs, Flamholz et al. (2013) have proposed that ED and EMP designs represent the best solutions to the trade-off between energy yields and proteome costs. On the other hand, Ng et al. (2019) have identified over 11,916 possible routes between glucose and pyruvate at different pre-determined stoichiometric ATP yields (up to 5 ATP per glucose). Pareto optimality analysis between energy efficiency and protein cost reveals that the naturally evolved ED and EMP pathways are indeed among the most protein cost-efficient pathways in their respective ATP yield categories and remain thermodynamically feasible across a wide range of ATP/ADP ratios and pathway intermediate metabolite concentration ranges. In contrast, pathways with higher ATP yields (>2), while feasible, are bound within stringent and often extreme operability ranges of cofactor and intermediate metabolite concentrations. The preponderance of EMP and ED is thus consistent not only with optimally balancing energy yield vs. enzyme cost, but also with ensuring operability for wide metabolite concentration ranges and ATP/ADP ratios. In other words, while optimizing protein resource is an important factor in pathway selection, robustness to extracellular changes (e.g, high/low glucose supply, alternate carbon substrates, stress, etc.) also plays a major role in shaping metabolic pathways. Furthermore, pathways need to retain operability beyond the realms of exponential (i.e., steady-state) growth phase, where cofactors and intracellular metabolites in organisms undergo high fluctuations (e.g., during lag, stationary, cell-division phases, etc.). Thus, robustness of a pathway to different intracellular concentrations may ultimately determine whether a pathway is universally adopted. At least for glycolytic pathways, the Pareto optimal surface for

protein cost vs. energetic yield must be extended to include the dimension of robustness. It is plausible that robustness imperatives that apply to naturally selected pathways may also be important for synthetic ones.

Slide #12 | EMP and ED are widely distributed

Slide #13 | Glycolysis – trade-off between energy yield and proteome cost

Slide #14 | Robustness of natural glycolyses

In summary, cells seem to operate close to Pareto optimal but the performance with respect to every single criterion is almost always improvable (Schuster 2013). In fact, several questions remain unanswered. Is it true that only a handful of tasks affect the fitness of organisms, or is this an artifact of our limited experimental data and measurement abilities? Further advancing our understanding of evolution as a multi-objective problem will require deciphering how contrasting goals are combined to define fitness, determining why specific points on the Pareto front are chosen instead of others (Noor and Milo 2012). Rational design and evolutionary methods are routinely used for the production of biomolecules with optimized properties (see chapter 7). The most colossal synthetic biology challenge is whether we can improve what evolution offers us.

References

Alon U (2018) Chapter 11: Evolution and multi-objective optimality (video). Systems Biology Course. Uri Alon Lab, Weizmann Institute of Science.

Alon U (2019) Chapter 14: Multi-objective optimality in biology. In: *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall/CRC Computational Biology Series) 2nd ed.

Bar-Even A, Flamholz A, Noor E, Milo R (2012) Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nat Chem Biol 8, 509–517.

Basan M (2018) Resource allocation and metabolism: the search for governing principles. Curr Op Microbiol 45, 77-83.

Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. Nature 528, 99-104.

Court SJ, Waclaw B, Allen RJ (2015) Lower glycolysis carries a higher flux than any biochemically possible alternative. Nature communications, 6(1), 1-8.

Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo, R (2013) Glycolytic strategy as a tradeoff between energy yield and protein cost. Proc Natl Acad Sci USA 110, 10039–10044.

Heinrich R, Montero F, Klipp E, Waddell TG, Meléndez-Hevia E (1997) Theoretical approaches to the evolutionary optimization of glycolysis: thermodynamic and kinetic constraints. Eur J Biochem 243, 191–201 (1997).

Koçillari L, Fariselli P, Trovato A, Seno F, Maritan A (2018) Signature of Pareto optimization in the *Escherichia coli* proteome. Sci Rep 8, 9141.

Meléndez-Hevia E, Waddell TG, Heinrich R, Montero F (1997) Theoretical approaches to the evolutionary optimization of glycolysis–chemical analysis. Eur J Biochem 244, 527–543.

Ng CY, Wang L, Chowdhury A, Maranas CD (2019) Pareto optimality explanation of the glycolytic alternatives in nature. Sci Rep 9, 2633.

Noor E, Milo R (2012) Efficiency in evolutionary trade-offs. Science 336, 1114-1115.

Noor E, Eden E, Milo R, Alon, U (2010) Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. Mol Cell 39, 809–820.

Savir Y, Noor E, Milo R, Tlusty, T (2010) Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. Proc. Natl Acad Sci 107, 3475-3480.

Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U (2012). Multidimensional optimality of microbial metabolism. Science 366, 601–604.

Schuster P (2013) Designing living matter: can we do better than evolution? Complexity 18, 21-33.

Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, Dekel E, Kavanagh K, Alon U (2012) Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. Science, 336, 1157-1160.

SEVERAL GLYCOLYTIC INTERMEDIATES are building blocks of biomass or biosynthetic precursors, as discussed in chapter 1. Thus, glycolysis qualifies also as an amphibolic pathway, although hardly any textbook will call attention to this aspect.

Slide #15 | Glycolysis – amphibolic character

3.3 Catabolism of other sugars

Beyond glucose, there are a variety of sugars that can be the substrate for glycolysis. For instance, the disaccharide sucrose, made up of glucose and fructose, must be first hydrolyzed in glucose and fructose. Fructose can be alternately incorporated by different enzymes: via fructokinase and F1P aldolase, or taking advantage of the ambiguity for substrate recognition of hexokinase (yielding F6P). Mannose can also be a substrate of hexokinase and phosphomannose isomerase yields F6P. Milk sugar, lactose, represents a case with a stereochemical issue. Lactose is composed by glucose and galactose, which differs from glucose in position 4 (i.e., glucose and galactose are epimers in position 4). This isomerism blocks the direct access of galactose to glycolysis and, then, two additional steps are required to metabolize galactose: first, galactose is a substrate of galactokinase yielding galactose 1-phosphate (Gal1P). Then, Gal1P reacts with UDP-glucose to give G1P and UDP-Gal (galactose 1-phosphate uridyl transferase) and finally UDP-Gal is a substrate of UDP-gal 4-epimerase that changes the isomerism of carbonyl 4. In this way, G1P is the net product of the process and can already be incorporated to glycolysis. This set of essential reactions allows to take advantage of 50% of lactose. A defect in this machinery causes the accumulation of galactose in the first moments of life with very dramatic consequences (the galactosemia condition). A seemingly banal metabolic reaction can actually be very important. On the other hand, lactose tolerance (or lactase permanence) is a relatively recent adaptation of humans. In general, mammals in adulthood do not suckle. Why do humans drink milk? Diverse human groups have fixed mutations in the promoter region of lactase allowing the expression of an enzyme that would normally be silenced in adulthood. Human population genetics show that the ability to digest lactose in adulthood has been fixed several times during human coevolution with domesticated mammals.

Slide #16 | Several sugars can be glycolytic substrates

DO YOU THINK THAT WE KNOW every metabolic detail of a model bacterium like *E. coli*? As recently as 2014, researchers have described a new metabolic pathway in this bacterium that accounts for the catabolism of an abundant biomolecule: sulphoquinovose (SQ, 6-deoxy-6-sulphoglucose). SQ is the polar headgroup of the plant sulpholipid in the photosynthetic membranes of all higher plants, mosses, ferns, algae and most photosynthetic bacteria. It is also found in some non-photosynthetic bacteria, and SQ is part of the surface layer of some Archaea. The estimated annual production of SQ is 10,000,000,000 tonnes (10 petagrams), which represents a major portion of the organo-sulphur in nature, where SQ is degraded by bacteria. Sulphoglycolysis is a modified version of the EMP pathway with a series of enzymes that are homologous to the glycolytic ones. Furthermore, the corresponding ED version of sulphoglycolysis has been described in *P. putida*. Note that, in both cases, only half of the SQ is incorporated in the bacterium biomass. The other half (DHPS, 2,3-dihydroxypropane sulphonate, or SL, 3-sulpholactate) is excreted and must be mineralized by other bacteria, thus closing the sulphur cycle within a bacterial community.

Slide #17 | Sulphoglycolysis –EMP and ED versions

3.4 Fermentations

The glycolytic pathways as such do not require oxygen (or any other external oxidant) to function, provided that the NADH produced in the lower part of the pathway is reoxidized. This is possible thanks to the process of fermentation (note the diversity of meanings of this term). In biochemistry, fermentation is any process in which the energy released by the catabolism of organic substrates is invested in the synthesis of ATP by substrate level phosphorylation, while maintaining the redox balance.

Maintaining the redox balance is essential and can be achieved (a) by internal redox compensation or dismutation (or disproportionation), which results in end products with different oxidation states, or (b) with the involvement of an exogenous oxidant (independent of an electron transport chain).

Slide #18 | Fermentation – in the biochemical sense

Slide #19 | Fermentation – achieving an internal redox balance by dismutation

Homolactic fermentation is an example of the first type. The glycolytic transformation of glucose into pyruvate is coupled to the reduction of pyruvate to L-lactate by the action of lactate dehydrogenase. The NADH produced in glycolysis is reoxidized during lactate synthesis.

Slide #20 | Dismutation in homolactic fermentation

In alcoholic fermentation, pyruvate is transformed into CO_2 and acetaldehyde by pyruvate decarboxylase, and finally NADH is reoxidized during the alcohol dehydrogenase catalyzed step (ethanol formation from acetaldehyde).

Slide #21 | Dismutation in alcoholic fermentation

Certain microorganisms – such as the genera *Lactobacillus* and *Streptococcus* – ferment glucose to lactate – and also galactose derived from lactose present in milk – during the formation of cheese or yogurt. The yeast *Saccharomyces cerevisiae* – among other microorganisms – and some plants ferment sugars to ethanol. *Clostridium acetobutyricum* ferments starch to butanol and acetone. Other microbial fermentation products of industrial interest are formic, acetic, propionic, butyric and succinic acids and the alcohols methanol, glycerol, isopropanol, butanol and butanediol.

An example of type (b) fermentation is the formation of lactate by *C. lactoacetophilum*, which occurs only in the presence of acetate, which acts as an external electron acceptor.

Slide #22 | Fermentation – diversity of end products
Slide #23 | Fermentation – diversity of end products
Slide #24 | Heterolactic fermentations

In addition to sugars, microorganisms ferment other substrates, such as organic acids (e.g., citrate), amino acids and nitrogen bases (see chapter 5), etc.

Slide #25 | Fermentations – beyond sugars
Slide #26 | Citrate fermentation

Fermentations are metabolic processes with a remarkable interest in biomedicine and biotechnology. Many pathogenic bacteria are anaerobic and rely in fermentation as a main bioenergetic support. On the other hand, many fermentations are also behind food processing and the creation of a fabulous diversity of flavors and tastes in foodstuff preserved from decay, from cheese to wine, bread or soya sauce. As historians of the Ancient world pointed out, every society that has learned how to grow carbohydrates (grains or grapes, for instance) discovers, sooner or later, how to ferment it.

Slide #27 | Fermentations – biomedical and biotechnological applications

LACTATE IS ALSO PRODUCED under conditions of hypoxia in water-immersed plants, intensely active skeletal muscle, erythrocytes, and the eye cornea. In these cases, it is preferable to refer to this process as *anaerobic glycolysis*. As an example, the Cori cycle is the result of the coupling between anaerobic glycolysis in active muscle and gluconeogenesis (see next section) in the liver. Read about lactate dehydrogenase in David S. Goodsell (Molecule of the Month, June 2008) Lactate Dehydrogenase: Our cells temporarily build lactate when supplies of oxygen are low.

Slide #28 | Cori cycle

3.5 Gluconeogenesis

Glucose is an important nutrient for many organisms. Thus, along with fatty acids, it is the main source of energy for animals – note, however, remarkable exceptions, such as migrating birds, in chapter 4. Even some cell types, such as erythrocytes and neurons, cannot oxidize fatty acids and therefore have an absolute requirement for glucose. Most of this sugar is supplied in the diet in the form of sucrose or polysaccharides. Plants obtain it, phosphorylated or not, mainly from the reserve

polysaccharide – starch, see below – or from sucrose. Gluconeogenesis is the synthesis of glucose from non-glycidic precursors. These precursors are mainly lactate, which is a product of amino acid and protein metabolism in muscle and other tissues, and glycerol, which is derived from the degradation of triacylglycerides in fat tissue. In ruminants, the propionate produced in the rumen when bacteria ferment glucose is a quantitatively important gluconeogenic precursor.

Unlike glycolysis, gluconeogenesis can function in only some animal tissues – kidney and especially liver – in certain plants and in microorganisms, when they grow in the absence of glucose. Gluconeogenesis from lactate is especially important in the liver. In this tissue, the G6P formed by gluconeogenesis can be used for glycogen synthesis or be transformed into glucose, which through the bloodstream will reach the tissues that consume it as energy fuel. In mammals, the lactate used during gluconeogenesis comes from extrahepatic tissues, especially active skeletal muscle and erythrocytes. The conversion of glucose to lactate in muscle followed by the conversion of lactate to glucose in the liver is called the Cori cycle (see slide #28). The use of lactate as a gluconeogenic precursor involves lactate dehydrogenase to oxidize it to pyruvate.

Slide #29 | Glycolysis *vs.* gluconeogenesis

The carboxylating stage of pyruvate to oxaloacetate involves biotin as a cofactor, covalently bound to the enzyme pyruvate-carboxylase through a Lys residue.

Slide #30 | Biotin-dependent pyruvate carboxylase

The transfer to the oxaloacetate cytosol produced in the mitochondrial matrix uses two shuttles: one in the form of aspartate, and one in the form of malate. This is not the only issue related to the distribution of the gluconeogenic enzymes in different eukaryotic cell compartments. For instance, the essential step of G6P hydrolysis is performed by G6Pase, an enzyme located in the endoplasmic reticulum.

Slide #31 | The role of cell compartments in gluconeogenesis

In plants, G6P derived from photosynthetic carbon fixation is used for the synthesis of starch – in the chloroplast stroma – or of the transportable disaccharide sucrose – in the cytosol.

Glycerol is an important gluconeogenic substrate in fasting conditions. It is produced by hydrolysis of triacylglycerols in adipose tissue and reaches the liver through the bloodstream. Glycerol rapidly enters the cell by diffusion and is transformed into glycerol-3-phosphate by glycerol kinase. Glycerol-3-phosphate dehydrogenase – the enzyme of the glycerol phosphate shuttle – catalyses the formation of dihydroxyacetone phosphate, a glycolytic and gluconeogenic intermediate, in the reaction:

$$\text{glycerol-3-phosphate} + NAD^+ \leftrightarrow \text{dihydroxyacetone phosphate} + NADH + H^+$$

Slide #32 | Summary – multiple sources for gluconeogenesis

ANIMAL CELLS DO NOT HAVE a metabolic pathway for the net transformation of acetyl CoA into glucose. Instead, in other systems – yeasts, bacteria, seeds of germinating plants – the glyoxylate cycle works, resulting in the net synthesis of malate – and thus of oxaloacetate and sugars – from acetyl CoA resulting from the degradation of fatty acids or two-carbon compounds – e.g., acetate, ethanol. As shown in chapter 2, this autocatalytic cycle is possible by the combination of isocitrate lyase and malate synthase with some of the reactions of the CAC.

3.6 Pentose phosphate pathway (PPP)

When the chemistry of nucleic acids began to be known in detail in the 1960s, the question of the origin of the pentoses that form their skeleton was raised, since the supply of these sugars in the diet is insufficient. A group of enzymes were discovered that oxidize G6P to CO_2 and ribose-5-phosphate (R5P), in the so-called *oxidative branch of the pentose phosphate pathway* (PPP). The main functions of this glucose oxidation pathway are the formation of ribose, and the supply of reducing power (NADPH), which is necessary for biosynthetic reactions (e.g., fatty acid and steroid synthesis) and protection against oxidative stress.

Slide #33 | PPP – oxidative branch

The reactions supplying NADPH and R5P are considered to be the oxidative part of the pathway, which is irreversible, while the non-oxidizing, reversible phase transforms pentose phosphates into hexose phosphates. This pathway works in all organisms and is especially active in some mammalian tissues, such as the liver, mammary glands and adipose tissue.

Slide #34 | PPP – non-oxidative branch

The PPP is a paradigm for stoichiometric combinations: connections of both the oxidative and non-oxidative branches, as well as glycolytic reactions, generate a potential diversity of transformations: all electrons delivered by the complete oxidation of G6P conserved in NADPH, balanced synthesis of NADPH and pentoses, interconversion of hexoses and pentoses, synthesis of NADPH, pentoses and ATP, etc.

Slide #35 | PPP – total oxidation of G6P and the synthesis of NADPH

Slide #36 | PPP – diversity of stoichiometric designs, catabolic and anabolic

Slide #37 | A new metabolic cycle in *P. putida* – combining EMP, ED and PPP

Session 2 of Computer Lab is devoted to calculating combinatorial stoichiometries in a small metabolic network (enumeration of elementary modes).

3.7 Polysaccharides: glycogen and starch

Excess of C (e.g., glucose) may be transformed into reserve polymers: glycogen – in vertebrates and many other organisms such as bacteria and fungi – or starch – in plants. Polysaccharides also play a central role as structural polymers.

Slide #38 | Polysaccharides

In plant cells, starch is a mixture of amylose and amylopectin. Amylose is a linear polymer of about 100 to 1,000 glucose residues connected by glycosidic bonds (α1-4). Amylopectin is a branched polymer of between 300 and 6,000 residues, with branching points located every 15 to 25 residues. Exo and endo-amylases are necessary for the mobilisation of glucose stored in starch.

Slide #39 | Starch – structure and degradative activities

Glycogen contains about 600,000 glucose residues linked by glycosidic bonds (α1-4) with about 2 branching points (α1-6) every 10–14 residues. The presence of branches gives solubility to the glycogen molecule.

Slide #40 | Glycogen – structure

The glucose units of the external branches of the polysaccharide are released thanks to the combined action of two enzymes: glycogen phosphorylase – or starch phosphorylase in plants – and (α1-6)-glycosidase. Glycogen phosphorylase requires as a cofactor pyridoxal phosphate which forms a Schiff's base by binding the aldehyde group to the side chain of a Lys of the enzyme. The reactive group is the phosphate that catalyzes the phosphorolysis of the glycosidic bond (α1-4) that binds two glucose residues, with the release of the phosphorylated terminal glucose residue in the C1, G1P. The fact that the glycosidic bond is broken by the incorporation of inorganic phosphate – i.e., by phosphorolysis – instead of water – hydrolysis – saves one molecule of ATP per glucose, since it makes it unnecessary to phosphorylate it to be incorporated into the catabolism. However, the use of G1P molecules in glycolysis requires the action of a mutase that transforms them into G6P. The action of phosphorylases is repetitive from the non-reducing ends of the branches to four residues before a branching point (α1-6). At this point, the activities of the bifunctional debranching enzyme intervene, transfering a chain of three residues from one of the shortened chains to the end of the other, and hydrolyses the residue linked by the link α1-6. Therefore, the

products of glycogenolysis are G1P – 9 to 13 out of every 10 to 14 residues – and free glucose – one molecule out of every 10 to 14 residues.

Slide #41 | Glycogen – mobilization of glucose

Read on glycogen phosphorylase in David S. Goodsell (Molecule of the Month, December 2001): Glycogen phosphorylase releases sugar from its cellular storehouse.

The substrate for glycogen biosynthesis – glycogenogenesis – is not a phosphorylated glucose, but UDP-glucose. In this way, it is possible to differentiate within the same cell compartment – cytosol – two sets of hexoses with opposite metabolic destinies. The molecule that gives glucose residues for the formation of glycogen is synthesized in the reaction catalyzed by UDP-glucose-pyrophosphorylase:

$$\text{UTP + glucose 1-phosphate} \rightarrow \text{UDP-glucose+ PPi}$$

Immediate hydrolysis of the pyrophosphate significantly increases the performance of the reaction (compare with other biosynthetic precursors that are also obtained by pyrophosphorolytic use of an NTP).

The incorporation of glucose residues starts from glycogenin, a glycoprotein initiator, through an autocatalytic reaction whereby the initiator incorporates a maximum of 8 glucose residues, linked to each other by links (α1-4) and to the polypeptide chain by a Tyr residue. The successive incorporations of glucose residues (α1-4) are catalysed by glycogen synthase, while a transferase is responsible for introducing the branching points by catalysing the transfer of a chain of between five and nine glucose residues to a branching point four to six residues away from any other branch, and forming a link (α1-6). In bacteria, the donor molecule producing glucosyl residues in glycogen is ADP-glucose. The stages of starch synthesis in plants are very similar to those followed for glycogen. In contrast, the initiator are fragments of degraded starch and the main donor molecule for glucose residues is ADP-glucose.

Slide #42 | Glycogen – biosynthesis

Slide #43 | Glycogen – summary of synthesis and degradation

Slide #44 | Glycogen – metabolism in physiological context

Slide #45 | Glycogen – a Nobel summary

3.8 Autotrophic pathways

The autotrophic mode of metabolism is required for planetary life as we know it. Metabolic pathways leading to the net synthesis of cell carbon-based chemicals from CO_2 in autotrophic organisms provide biochemical food to the other consumers of organic molecules (heterotrophic organisms). Thus, autotrophic organisms are also called primary producers. A large number of microorganisms, as well as green plants, algae, and some protists, are phototrophic. Most phototrophs use energy conserved in ATP and electrons in NADPH for the assimilation of carbon

dioxide as the carbon source for biosynthesis. These phototrophs are called photoautotrophs. In the 1880s, Sergei Winogradsky (1856–1953) proposed the concept of chemolithotrophy, the oxidation of inorganic compounds as a source of energy and electrons for the autotrophic growth. Studying sulfur bacteria, he concluded that these organisms obtained their carbon from CO_2 and they were called autotrophs. The discovery of autotrophy in chemolithotrophic bacteria was of major significance in the advance of our understanding of cells physiology, because it showed that CO_2 could be converted to organic carbon without photosynthesis. Chemolithotrophy is shown by members of both the Bacteria and Archaea domains and is important in the cycling of inorganic compounds on Earth. Considering the capacity to fix C, autotrophic microorganisms are seen as targets for the search of potential mechanisms of C sequestration to alleviate the climate crisis.

Slide #46 | Autotrophic microorganisms and the climate crisis

At least six biochemical mechanisms are known for the autotrophic fixation of CO_2 into cell material. The pathways differ in the participating enzymes, ATP, and reducing power requirements, as well as in carbon isotope fractionation. For this repertoire to be complete, we must mention the description of a CAC operating literally (with the same enzymes of the oxidative version) in reverse (see chapter 2).

Slide #47 | Autotrophic pathways – a phylogenetic distribution

1 The Calvin-Benson cycle, discovered in the 1950s in Melvin Calvin's lab, starts with the condensation of a 5C sugar (ribulose 1,5-bisphosphate) with CO_2 to yield two 3C molecules (3-phosphoglycerate). From these molecules, the initial 5C sugar is regenerated and organic materials are biosynthesized. The cycle is operative in plastids of plants, algae, and protists, as well as in cyanobacteria, some aerobic or facultative anaerobic proteobacteria, CO-oxidizing mycobacteria, some iron- and sulfur-oxidizing firmicutes, and green sulfur bacteria. This pathway's ability to fix carbon is conferred by the activity of two enzymes (together with fragments of central metabolism like gluconeogenesis and the pentose phosphate pathway): ribulose 1,5-bisphosphate carboxylase-oxygenase (Rubisco) and phosphoribulokinase. Although Rubisco activity has been detected in some Archaea, it has not been possible to demonstrate their Calvin-Benson-dependent autotrophic growth (Berg et al. 2010).

Read David S. Goodsell on Rubisco (Molecule of the Month, November 2000): Rubisco fixes atmospheric carbon dioxide into bioavailable sugar molecules. Although you have studied the Calvin-Benson cycle in other courses, here are some minimal reminders.

Slide #48 | Calvin-Benson cycle

Slide #49 | Rubisco – two enzymatic faces

Slide #50 | Photorespiration – metabolic tinkering to the rescue

Slide #51 | Sugar biosynthesis in plants

2 In the green sulfur bacterium *Chlorobium*, CO_2 fixation occurs by a reversal (or reductive) citric acid cycle, also known as Arnon cycle. This is an analogue of a Krebs cycle operating in reverse (see chapter 2). Since the Arnon cycle involves some enzymes (e.g., the carboxylating and reducing steps) that are inhibited by oxygen, this pathway is only found in microorganisms growing under anaerobic conditions. These include some proteobacteria, green sulfur bacteria, and microaerophilic bacteria like *Aquifex*. In sulfur-reducing delta-proteobacteria, the operation of the CAC has been described to be literally in reverse. Remind that we have discussed the Arnon cycle and reverse CAC in chapter 2.

Slide #52 | Arnon cycle

Slide #53 | Arnon cycle – two enzymatic solutions for the simplest stoichiometry

Slide #54 | Fixing C by literally reversing the CAC

3 In some Gram-positive bacteria and methanogenic archaea, the ability to synthesize acetyl CoA from CO and/or CO_2, the Wood-Ljungdahl (WL) pathway was identified. One CO_2 molecule is reduced to CO and another one to a methyl group (attached to a cofactor). Then, acetyl CoA is synthesized from CO and the methyl group. The key enzymes of this pathway are inhibited by oxygen, thus, it is restricted to force anaerobic microorganisms. These include some proteobacteria, planctomycetes, spirochaetes, and archaea. Albeit the WL pathways in bacteria and archaea are analogous, some enzymes are neither evolutionarily related nor they use the same coenzyme repertoire. Whether the WL pathway is the oldest autotrophic pathway or it was operative in LUCA is still a debated issue (Peretó 2012).

Slide #55 | Synthesis of acetyl CoA from CO/CO_2

Slide #56 | WL pathway in more detail – the fates of acetyl CoA

Slide #57 | WL pathway and methanogenesis are related

Slide #58 | Coenzymes and cofactors

4 The 3-hydroxypropionate bi-cycle is present in some green non-sulfur phototrophic bacteria like *Chloroflexus*. A succinyl CoA molecule is synthesized from acetyl CoA and two bicarbonate molecules. Although the same intermediates as the hydroxypropionate-hydroxybutyrate cycle (see below) are used, most of the participating enzymes are different. The final product of the cycle is glyoxylate and its assimilation requires a second metabolic cycle. This pathway is restricted to the family Chloroflexaceae and might represent an early attempt of autotrophy in anoxygenic phototrophs.

5 The hydroxypropionate-hydroxybutyrate cycle occurs in some aerobic archaea, like *Sulfolobus*. Although this pathway is formally the same as the 3-hydroxypropionate bicycle, the

nonhomologous participating enzymes indicate that both pathways evolved independently in a remarkable case of evolutionary convergence in metabolism.

6 The dicarboxylate-hydroxybutyrate cycle occurs in some anaerobic archaea such as Thermoproteales and Desulfurococcales. The pathway can be divided into two parts: (1) one acetyl CoA, one CO_2, and one bicarbonate are converted into succinyl CoA; (2) this C4 molecule is transformed into two molecules of acetyl CoA (one serves as biosynthetic precursor, the other as acceptor of the cycle).

Slide #59 | Archaeal autotrophic pathways
Slide #60 | Autotrophic pathways – an outline
Slide #61 | Autotrophic pathways – a stoichiometric summary

Finally, I represent the six autotrophic pathways in terms of the kind of catalytic cycles (simple or autocatalytic) they show. All the pathways, except the reductive acetyl-CoA pathway, are autocatalytic in the sense that the net product they generate represents a stoichiometric increase in the feeder (or a molecule equivalent to it). Thus, the Arnon cycle generates a C4 molecule (OAA) from four CO_2 molecules and regenerates OAA for the following turn of the cycle. The net product of carbon fixation in the Calvin-Benson cycle is a triose C3. Although the substrate necessary to initiate the cycle is a pentose C5, in metabolic terms, C3 and C5 are fully equivalent, since there is a complete set of enzymes performing the non-oxidative pentose-phosphate pathway that allows the interconversion of C3 and C5 (see section 3.6).

Both hydroxybutyrate cycles (the dicarboxylate/4-hydroxybutyrate cycle, and the 3-hydroxypropionate/4-hydroxybutyrate cycle) are variations on a theme: the net synthesis of acetyl-CoA from CO_2 and the regeneration of acetyl-CoA as the feeder of the cycle. Finally, the 3-hydroxybutyrate bi-cycle can synthesise a C3 compound (pyruvate) from HCO_3^- through the concatenation of two autocatalytic cycles. Although the feeder of both cycles is acetyl-CoA, the two net products (glyoxylate and pyruvate) are its equivalents: glyoxylate is converted into pyruvate and pyruvate can easily be converted into acetyl-CoA via oxidative decarboxylation. Thus, these five autotrophic pathways qualify as autocatalytic cycles, because all of them synthesise the feeder, or an immediate precursor, from CO_2.

The reductive acetyl-CoA pathway represents quite a different case. The net synthesis of acetyl-CoA from CO_2 is a linear pathway that proceeds through the consecutive action of three simple catalytic cycles. The cycles use cofactors as catalytic reagents: THF, Co-containing corrinoid, and a Ni-FeS cluster. The net product of C fixation in this case (acetyl-CoA) does not meet the criterion of being equivalent to any of the regenerated feeders, so the reductive acetyl-CoA pathway is not an autocatalytic cycle.

Slide #62 | Autotrophic pathways – catalytic cycles in context

Bibliography and resources

Textbooks

Metzler[2], vol. 1 ch. 10, vol. 2 ch. 17
Nelson and Cox[7], ch. 14.
Peretó et al.[5], ch. 14.
Stryer[9], ch. 16, ch. 20, ch. 21.

Papers and book chapters

Basan M (2018). Resource allocation and metabolism: the search for governing principles. *Curr Op Microbiol* **45**:77-83.
This is a review on optimality principles in metabolism, a fundamental knowledge for any biotechnologist.
Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber BE, Fuchs G (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* **8**:447–460
González-Toril E, Peretó J (2015) Autotrophy. In: *Encyclopedia of Astrobiology* (2nd ed., Springer) Vol. 1, p. 218–223.
Peretó J (2012) Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem Soc Rev* **41**:5394–5403

On the net

CoFactor - The organic enzyme cofactor database is a good resource on enzyme cofactors. Part of organic cofactors are vitamins.
Learn about the pioneers in glycolysis studies, some of them Nobel prize awardees: Eduard Buchner, Otto F. Meyerhof, Jakub K. Parnas, Gustav Embden, Michael Doudoroff, and Nathan Entner (who has no Wikipedia article: does anyone dare to write it down?).

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ3.1 *Learn by mapping* Build a concept map of chapter 3 and establish connections with the map you did in chapter 2.

SQ3.2 Discuss the following sentence with the help of a graphic representation: glycolysis is a pathway that allows the autocatalytic synthesis of ATP.

SQ3.3 Lactic acid fermentation and alcoholic fermentation are oxidation-reduction reactions. Identify the ultimate electron donor and electron acceptor.

SQ3.4 *Learn by drawing* Draw the complete pathways and write global balanced equations for alcohol fermentation of sucrose in *Saccharomyces cerevisiae* (EMP pathway) and in *Zymomonas mobilis* (ED pathway).

SQ3.5 Match the following two columns:

1. Pyruvate decarboxylation	a. ED pathway
2. Semiphosphorylative glycolysis	b. Glucose 6-phosphatase
3. Substrate-level phosphorylation	c. Thiamine
4. Present in liver	d. Fermentation

SQ3.6 What energetic barrier prevents glycolysis from simply running in reverse to synthesize glucose? How is this barrier overcome in gluconeogenesis?

SQ3.7 Write balanced equations for complete glucose oxidation in CO_2: (a) through the combination of EMP pathway, PDC and CAC; (b) through the PPP (oxidative and non-oxidative branches combined). List the main differences regarding the energetic output of both processes.

SQ3.8 How many ATP equivalents are needed to convert glycerol in a glycosyl residue in glycogen?

SQ3.9 In the conversion of galactose in glucose 1-phosphate, UDP-glucose acts catalytically. Explain.

SQ3.10 In the conversion of glucose into two molecules of lactate, the NADH generated earlier in the pathway is oxidized to NAD^+. Why is it not to the cell's advantage to simply make more NAD^+ so that the regeneration would not be necessary? However, the recommended daily allowance of niacin (B3 vitamin) is 15 mg per day. Explain.

SQ3.11 *A useful exercise during the learning process.* Draw a detailed map including glycolysis/gluconeogenesis/fermentation, PPP and the CAC. Be as thorough as possible detailing the origins and fates of pyruvate.

SQ3.12 *Be a cultivated biotechnologist.* What is the oldest evidence of fermentation processes used by humans? Given the geographical and temporal distribution of the remnants, what does this suggest about the "invention" of fermentation? List some potential beneficial properties of fermented foods in primitive societies.

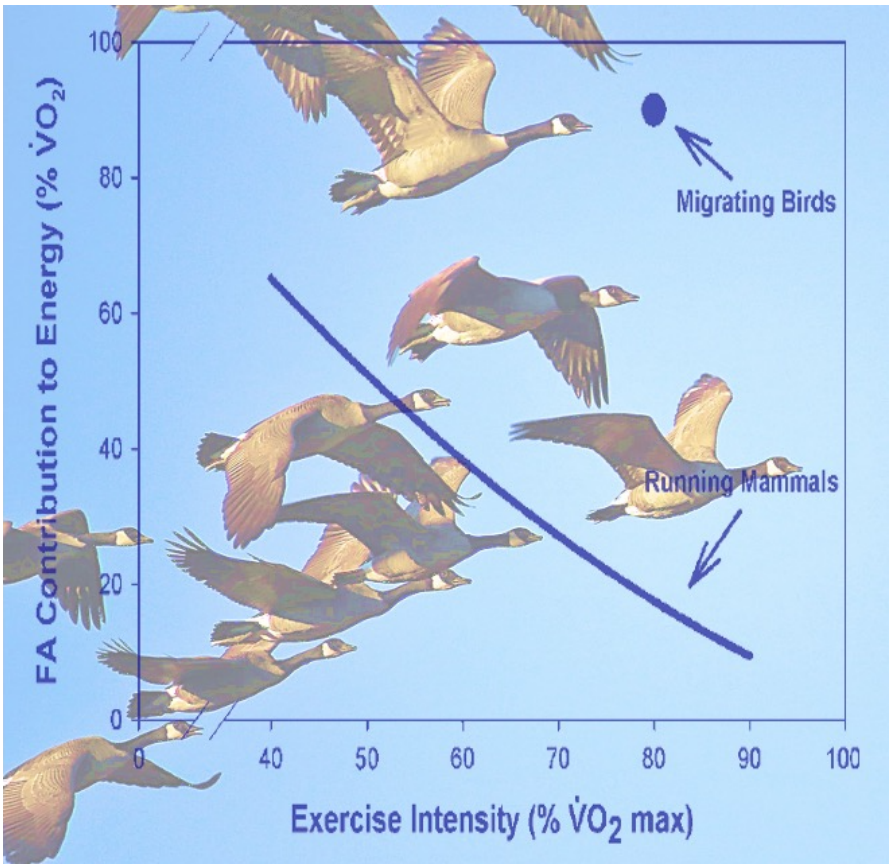
Problem-Based Learning (PBL)

PBL3 offers the opportunity to learn and consolidate the basic concepts related to chapter 3.

Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the required knowledge on metabolism during the course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

4

Lipid metabolism



Refresh your background

- ✓ Enzymology concepts (structure, mechanism, cofactors and coenzymes, enzymatic classes, kinetic parameters).
- ✓ Chemical functions in organic compounds and lipid nomenclature.

Keywords

- ✓ Alpha-oxidation
- ✓ Beta-oxidation
- ✓ Coenzyme B12 (cobalamin)
- ✓ Essential fatty acid
- ✓ Isoprenic unit
- ✓ Isoprenoid
- ✓ Ketone bodies
- ✓ Lipogenesis
- ✓ Lipolysis
- ✓ Lipoprotein
- ✓ Megasynthase
- ✓ Multifunctional FA synthase
- ✓ Nonribosomal peptide
- ✓ Phosphatidate
- ✓ Phospholipid
- ✓ Polyketide
- ✓ Statin
- ✓ Steroid
- ✓ Terpenes
- ✓ Triacylglycerol (triglyceride)

Learning objectives

- ✓ Compare fats vs. carbohydrates as metabolic fuels
- ✓ Compare biochemical aspects between catabolism and anabolism of FA
- ✓ Know particular aspects of FA oxidation: unsaturated and odd-chain FA
- ✓ Know that some FA must be acquired through the diet (essential FA)
- ✓ Structure vs. storage: biosynthesis of membrane lipids and TAG
- ✓ Diversity of natural products derived from acetyl CoA: polyketides, terpenes (steroids)
- ✓ Metabolism of cholesterol as a recent pathway innovation
- ✓ Solubility and metabolic fates of fats: lipoproteins, ketone bodies and their physiological roles

Contents

Overview

Lipids, like glycides, are important structural elements and sources of energy. In most organisms, lipids are mainly found in the form of triacylglycerols (TAG), in which most of the carbon is in a state of maximum reduction. Therefore, the metabolic oxidation of fats consumes more oxygen than that of carbohydrates, and more energy is released. β-oxidation, the mechanism of fatty acid degradation, produces acetyl CoA molecules which, among other metabolic destinations, can circulate between tissues after the formation of ketone bodies (KB). Acetyl CoA is also the precursor of a fabulous diversity of natural products, such as polyketides and terpenes (and terpenoids) with a growing biotechnological value. From the metabolism of terpenes, we highlight the synthesis of cholesterol, a component of eukaryotic membranes and a biosynthetic precursor of remarkable metabolites, such as bile acids and some hormones. The metabolism of lipoprotein particles illustrates how the challenge of water insolubility of fats is addressed in a physiological context.

4.1 Lipid metabolism in context

Before starting our journey through the pathways of lipid metabolism, let's have some context in the city of metabolism. Schemes in Metzler's textbook are good guides for the visit to the acetyl CoA neighbourhood, with direct connections with the central C pathways, glycolysis, PPP (as a source of NADPH for biosynthetic pathways), and CAC.

Slide #2 | The acetyl CoA neighbourhood – connections with central pathways

As a reminder, it is important to refresh your knowledge on fatty acid (FA) nomenclature.

Slide #3 | FA nomenclature

Lipids have remarkable structural roles and biochemical functions (e.g., phospholipids or cholesterol in membranes) but we will focus on the energetic aspect of fats (i.e., triacylglycerols TAG and FA) as energy reservoirs.

Slide #4 | Fats provide efficient means for storing energy for later use

Since carbohydrates (CHO) are in a more advanced state of oxidation than FA, the energy yield associated with the oxidation of glycides, in terms of ATP obtained from carbon oxidized to CO_2, is 25% lower than that corresponding to FA[1]. But when it comes to highlighting the superior energy storage capacity of fats over CHO, we must note that in the cell, fats form practically anhydrous aggregates, whereas reserve CHO, such as glycogen, are hydrated – about 2 g of water/g of glycogen. However, there is a contradictory aspect that must be taken into account: fats are not soluble in water so the transport of the fuel from storage sites (e.g., adipose tissue, adipocytes) to the utilization destination (e.g., muscle cells, myocites) may represent a logistic issue. In this context, metabolic adaptations have been necessary to convert fats in the main energy source during flying exercise in migrating birds (see "The extraordinary case of the 'obese athletes'" at the end of this chapter).

4.2 Metabolism of triacylglycerols

Triacylglycerols (TAG) are important energy stores in many organisms. In animals, metabolized fatty acids derive from two primary sources: the diet – about 90 % of total ingested lipids are TAG – and the mobilization of the fat stored in the adipocytes. The digestion, absorption and transport of dietary lipids pose various problems because of their insolubility in an aqueous medium. During

[1] The energy released in fat oxidation does not necessarily lead to ATP synthesis: in brown adipose tissue, the decoupling of mitochondrial electronic transport makes FA catabolism a thermogenic process.

the process of lipogenesis, TAG are formed from fatty acids and 2-monocylglycerols absorbed into intestinal cells, at different stages catalyzed by acyl CoA-transferases:

2-monoacylglycerol + acyl CoA → 1,2-diacylglycerol + HSCoA

1,2-diacylglycerol + acyl CoA → triacylglycerol + HSCoA

When TAGs are formed from three molecules of acyl CoA and one of glycerol-phosphate, a phosphatase is also involved that removes the phosphate group of 1,2-dacylglycerol 3-phosphate (phosphatidic acid), an intermediate that is also part of the phospholipid biosynthesis route. Glycerol 3-phosphate can be produced by reducing the glycolytic intermediate dihydroxyacetone phosphate or (in the liver) also by phosphorylation of glycerol catalyzed by glycerol kinase.

Slide #5 | Dietary fats and lipolysis

4.3 Catabolism of fatty acids

Knoop (1904) conducted the first metabolic marking experiments that revealed how fatty acids are degraded. He fed dogs with fatty acid derivatives containing a phenyl group attached to the terminal carbon. When these derivatives had an odd number of carbon atoms, the compound detected in the urine was hippuric acid (a combination of benzoate and Gly); however, when the derivatives ingested were an even number of carbon atoms, phenylacetic acid was recovered. This result shows that fatty acids do not degrade carbon to carbon, since in this case only hippuric acid would have been obtained. Knoop proposed that fatty acids are degraded by sequential oxidations of carbon β, and at each step there is one acyl chain with two less carbons. The catabolism of fatty acids in animal cells can be divided into three stages: (1) activation of fatty acids in the cytosol, (2) transport to the mitochondria and (3) β-oxidation. The oxidation will be complete if the acetyl CoA is a substrate of the citric acid cycle (see chapter 2).

(1) Fatty acid activation

Before they are oxidized, the fatty acids have to be activated to form the corresponding acyl CoA in a reaction catalyzed by a family of at least three acyl CoA-synthases (systematic name: AMP-forming acyl CoA-ligase– classified according to their specificity towards the length of the fatty acid hydrocarbon chain. These enzymes are associated with the external mitochondrial membrane and the endoplasmic reticulum and catalyze the formation of a thioester bond between the fatty acid and coenzyme A. In the cell, this reaction is driven by the hydrolysis of PPi, by pyrophosphatase, a general mechanism that increases the driving force in biosynthetic processes. Compare it with other mechanisms of chemical activation, e.g., amino acid activation during aa-tRNA synthesis.

(2) Transport of fatty acids to the mitochondrial matrix

Most fatty acids are found in the cytosol, where their activation takes place because that is the place where they are synthesized or because they are transported there from outside the cell;

but the oxidation of these acids takes place in the mitochondrial matrix. The carnitine-dependent transport system allows access of fatty acids to the site of oxidation.

Slide #6 | FA activation and intracellular transport

(3) β-oxidation

Acyl CoA derivatives are successively oxidized into the carbon β with the sequential release of two carbon fragments in the form of acetyl CoA. There is a chemical logic behind: weakening the C-C bond by introducing a keto group in β position. In this sense, the combined action of attracting electrons from the thioester bond and the new keto group makes the excision of the covaoent bond easier.

Slide #7 | FA β-oxidation – chemical logic and enzymatic steps

Each cycle of β-oxidation comprises four reactions: (i) formation of a double trans-α,β link through dehydrogenation catalysed by a flavin-dependent enzyme, acyl CoA-dehydrogenase; (ii) hydration of the double bond by enoyl CoA-hydratase to form 3-L-hydroxyacyl CoA; (iii) dehydrogenation of β-hydroxyacyl CoA by 3-L-hydroxyacyl CoA dehydrogenase, with formation of the corresponding β-oxoacyl CoA; (iv) excision of the link Cα-Cβ in a thiolysis reaction with HSCoA, catalysed by the β-oxacyl CoA-thiolase, to form acetyl CoA and a new acyl CoA with two carbons less than the original. This series of reactions occurs in each cycle until the last two molecules of acetyl CoA are obtained.

The β-oxidation of unsaturated fatty acids (UFA) requires additional enzymes in mammals, to solve the stereochemical problem posed by double bonds in cis. Take into account the unsaturation in fatty acids (for instance, to modulate the membrane fluidity or the transport of FA) always has cis configuration, whereas the unsaturated intermediates during FA catabolism show a trans configuration. Two additional enzymes (cis-Δ^3-enoyl isomerase and 2,4-dienoyl CoA reductase) are sufficient to catabolise the most abundant UFA.

Slide #8 | UFA β-oxidation – additional enzymatic steps

FA with an odd number of carbon atoms are also substrates of the β-oxidation, but the product of the last cycle is propionyl CoA instead of acetyl CoA. This compound is converted to succinyl CoA and may have a gluconeogenic fate. The series of involved reactions show a peculiar mechanism with the participation of cobalamin (also known as vitamin B_{12}).

Slide #9 | Odd-chain FA β-oxidation – what happens with the C3 remnant?

Propionyl CoA may have diverse metabolic fates: anabolic, catabolic (cataplerotic) and anaplerotic (of the CAC).

Slide #10 | Metabolic fates of propionyl CoA

Propionyl CoA can be derived directly form propionic acid, the main product of gut microbiome fermentation in ruminants. In fact, in cows, propionic acid is the main gluconeogenic precursor.

Slide #11 | Propionate-acetate fermentation

Although β-oxidation is quantitatively the most significant route of fatty acid catabolism, fatty acids can also be degraded by carbon oxidation α instead of β. The process starts with a hydroxylation of the carbon in α position, followed by a dehydrogenation that generates the corresponding α-oxo acid and the consequent decarboxylation of carbon 1. The α-oxidation is an obligate step for processing branched FA (i.e., isoprenoid-derived organic acids). It is the case of phytanic acid derived from phytol, the main side-chain of chlorophyll. The importance of α-oxidation in human metabolism is reflected by Refsum disease, a pathology derived from its malfunction and the accumulation of phytanic acid.

Slide #12 | Branched FA α-oxidation

The oxidation of fatty acids has traditionally been considered as an exclusive activity of the mitochondria. But the discovery of another system of β-oxidation in peroxisomes from animals, plants and fungi has forced a change in this idea. The peroxisome pathway has the same intermediaries as the mitochondrial one, although the enzymes are different. For example, the first reaction of the peroxisome pathway is catalyzed by a FAD-dependent acyl CoA-oxidase – instead of a dehydrogenase. Because there is no electronic transport chain in peroxisomes, the electrons of the reduced FAD are transferred directly to oxygen, with the formation of H_2O_2, which is destroyed by catalase. The other fundamental difference is that carnitine is not used for the transport of acyl CoA to the peroxisome, but the fatty acids enter by simple diffusion, are activated by a peroxisome acyl CoA-synthase, and finally oxidized. As for the function of β-peroxisomal oxidation, in animals it is related to the oxidation of long chain fatty acids, which are shortened and finally degraded in the mitochondria. In plants and yeasts, the fatty acids are totally degraded in the peroxisomes up to acetyl CoA, which is incorporated to the glyoxylate cycle (see chapter 2) or exported to be oxidized by the CAC. Within the eukaryotic domain, the β-oxidation of peroxisome has a wider phylogenetic distribution than the mitochondrial pathway.

4.4 Anabolism of fatty acids

At the beginning of the 20th century, it was already known that most fatty acids have an even number of carbon atoms and therefore their biosynthetic process would involve successive additions of a compound of two carbons, the opposite of what happened with oxidation. This hypothesis was demonstrated in the 1940s by Ritenberg and Bloch, using radioactive markers for the first time. But the discovery of fatty acid biosynthesis as a totally different route to degradation was made by Wakil in the late 1950s. The synthesis of fatty acids is very similar in most organisms. In mammals it takes place mainly in the cytosol of liver cells, in fat tissue and in the mammary glands during lactation.

Acetyl CoA generated in the mitochondrial matrix by β-oxidation or by oxidative decarboxylation of pyruvate must be transported to the cytosol if it is to be used in FA synthesis. Thus, acetyl CoA and OAA condense in citrate, which is exchanged by a carrier using orthophosphate or malate. In the cytosol, ATP citrate-lyase releases acetyl CoA and OAA, which is reduced to malate by an isozyme of malate dehydrogenase. This is transformed into pyruvate by the action of the malic enzyme and with the formation of NADPH, which can be used in the synthesis of fatty acids.

Slide #13 | Moving acetyl CoA from mitochondria to cytosol

A previous step to fatty acid biosynthesis is the formation of malonyl CoA, an activated form of acetyl CoA that is the main substrate of the pathway, in a reaction catalysed by acetyl CoA-carboxylase, which uses biotin as a prosthetic group.

The biosynthesis of fatty acids can be divided into two stages: (1) the de novo synthesis of palmitate (16:0), and (2) subsequent modifications of the chain, including lengthening or shortening and the introduction of one or more double bonds C=C (unsaturated fatty acids). Some organisms, including humans, are unable to synthesise all the classes of fatty acids they need due to a lack of metabolic ability in the stages of introducing unsaturation. These fatty acids must be ingested in the diet and are called essential fatty acids.

There are seven enzymatic reactions involved in the conversion of acetyl CoA to fatty acid by the action of fatty acid synthase. This enzyme is a multifunctional protein that contains the acyl carrier protein (ACP), which is the binding site for biosynthetic intermediates. The thioesterase activity of fatty acid-synthase has a preference for 16:0 and 18:0 substrates, and most products of de novo synthesis are 16:0. For the formation of longer fatty acids, elongation occurs, which in eukaryotes takes place in the mitochondria and especially in the endoplasmic reticulum. Microsomal elongation is generated by the condensation of malonyl CoA in the already formed chain, and the formed β-oxoacyl CoA is reduced with NADPH, dehydrated and reduced again to yield an acyl CoA saturated with two more carbons than the original, and so on until the appropriate length is achieved. The mitochondrial elongation route is essentially the reverse of β-

oxidation – i.e., from acyl CoA and acetyl CoA – except for double bond reduction, where NADPH is used instead of $FADH_2$.

Slide #14 | Oxidation and synthesis of FA compared
Slide #15 | FA biosynthesis – the role of ACP
Slide #16 | An iterative process – bacterial FA synthesis
Slide #17 | An iterative process – eukaryotic FAS is a multifunctional enzyme
Slide #18 | Catalytic cycle of mammalian FAS

Read David S. Goodsell's description of FA synthase (Molecule of the Month, June 2007): Fatty acids are constructed in many sequential steps by a large protein complex.

Lipoic acid (required as a cofactor by the pyruvate dehydrogenase complex, as well as by α-ketoglutarate dehydrogenase) is synthesized in mitochondria from an 8-carbon long FA. Thus, although the FA biosynthetic process is mainly located in cytosol, there is a FA synthase specific of octanoic acid located in mitochondria.

Slide #19 | Mitochondrial FA (8:0) biosynthesis

THERE ARE TWO CHEMICALLY DIFFERENTIATED routes for introducing a double cis bond in saturated fatty acids: the anaerobic route, well studied in *E. coli*, and the aerobic route, which is found in many eukaryotes and has been studied mainly in mammalian livers. In the aerobic pathway, unsaturated fatty acids are synthesised from the saturated ones by a microsomal system called acyl fatty CoA desaturase. For example, stearoyl CoA desaturase (SCD):

$$\text{stearoyl CoA (18:0)} + NADH + H^+ + O_2 \rightarrow \text{oleoyl CoA (18:1, } \Delta^9\text{)} + NAD^+ + 2H_2O$$

Another enzyme, flavin-dependent cytochrome b_5 reductase, is involved in electron transfer. Elongation and further unsaturation give a rise to a huge diversity of UFAs. Some of those have essential functions and humans although we do not have the metabolic pathways to obtain them. Thus, α-linolenic (18:3, ω-3) and linoleic (18:2, ω-6) acids are required dietary components (essential FAs). Both essential FAs are the starting points for the synthesis of a variety of other UFAs (e.g., arachidonic acid is the metabolic precursor of eicosanoids, a family of signalling molecules like, among others, prostaglandins, thromboxanes, and leukotrienes).

The enzyme SCD is an example of oxidase, enzymes that catalyse oxidations in which molecular oxygen is the electron acceptor, but oxygen atoms do not appear in the oxidized product. On the other hand, oxygenases catalyse oxidative reactions in which oxygen atoms are directly incorporated into the product molecule, forming a new hydroxyl or carboxyl group, for example. Depending on whether one or the two atoms of oxygen are incorporated, there are mono- or dioxygenases. Most monooxygenases catalyse reactions in which the main substrate becomes hydroxylated, so they are also called hydroxylases.

Slide #20 | FA unsaturation

Slide #21 | Essential FAs

4.5 Biosynthesis of membrane lipids

Phosphatidic acid is a common precursor for storage lipids (TAG) and many membrane lipids (phospholipids). It is formed by successive acylations of glycerol 3-phosphate by acyl CoA (the monoacylated form is named lysophosphatidic acid). Hydrolysis of the phosphoric ester followed by a third acylation yields a TAG. CDP-diacylglycerol – activated precursor in the de novo synthesis of many phospholipids – is formed from phosphatidic acid and CTP. The activated phosphatidyl unit is then transferred to the hydroxyl group of a polar alcohol (e.g., inositol), to form a phospholipid (e.g., phosphatidylinositol). Alternatively, the biosynthesis of membrane lipids can proceed through the NTP-activation of the precursor alcohol.

Slide #22 | TAG and phospholipid biosynthesis

Slide #23 | Phosphatidic acid – precursor of lipid biosynthesis

Slide #24 | Two major strategies in phospholipid biosynthesis

Slide #25 | Lipid biosynthesis in plants

4.6 Polyketides

Polyketides (PK) represent a wide range of natural products – secondary metabolites –, containing alternating carbonyl and methylene groups ('β-polyketones'), biosynthetically derived from repeated condensation of acetyl coenzyme A (via malonyl coenzyme A), and usually the compounds derived from them by further condensations. There are many PK of commercial interest such as antibiotics (macrolides), antifungals, insecticides, cytostatics, immunosuppressive, or anticholesteremic compounds. Among PK of pharmaceutical relevance, the are several top-selling small molecule drugs, for example, antibiotics (e.g., erythromycin and tetracycline), antineoplastics (e.g., daunorubicin), and immunosuppressants (e.g., rapamycin).

PKs are synthesized by a specific class of enzymes (PK synthases PKSs) that are classified based on their sequence information, structural organization, and product specificity. Type I PKSs are polyproteins with multi-domain architecture that ultimately form bulky multi-functional complexes involved in PK biosynthesis in a modular or iterative fashion. Classical examples of type I PKS include erythromycin biosynthesis gene clusters from bacterium *Saccharopolyspora erythraea* and 6-methylsalicylic acid (6-MSA) biosynthesis in fungus *Penicillium patulum*. Individual modules in the biosynthetic proteins of type I PKS include acyltransferase domain (AT domain) and acyl carrier protein domain (ACP domain), as well as a thioesterase domain (TE domain). Type II PKSs are multienzyme complexes consisting of distinct, separable proteins similar to type II fatty

acid synthase system in plants and bacteria. Despite being dissociable, type II enzymes form in vivo complexes similar to those of the type I PKS systems. Classical examples of type II PKSs include antibiotic biosynthesis in *Streptomyces*. Computational methods and genome mining have led to the identification of >150 putative type II PKS gene clusters in actinobacterial genomes. Type III PKSs have a simple architecture that is a homodimer of identical ketoacyl synthase (KS) monomeric domains. The active site in each monomer catalyzes the entire reactions of priming, extension, and cyclization of acetate units into a polyketide chain. Primarily present in plants, type III PKSs are also found in bacteria and fungi as well. Unlike the type I and II PKSs, type III PKSs generally utilize CoA thioesters as substrates and do not involve ACP domains. Furthermore, type III PKSs can accomplish an entire series of decarboxylative condensations and cyclization reactions in a single active site. Polyketides produced by type III PKSs have applications in medicine, agriculture, and other fields, due to their notable structural diversity and biological activities. The first type III PKS system was discovered in plants in 1970 and chalcone synthase (CHS) is a typical and well-studied plant type III PKS involved in flavonoid biosynthesis. Flavonoids are major pigmentation compounds in plants. CHS is among the several type III PKSs that use 4-coumaroyl CoA (derived from the amino acid Phe) as the starter unit. CHS, assisted by other synthetic enzymes such as chalcone isomerase and flavanone isomerase, forms a flavonoid scaffold to adapt reactions such as hydroxylation, O-methylation, glycosylation, acylation, prenylation, and conjugation, resulting in the biosynthesis of >6,000 naturally occurring flavonoids. In addition to pigmentation, flavonoids play a vital role in plant–microbe interaction, defence mechanism due to their antimicrobial property, UV protection, etc. Since the discovery of CHS in plants, at least 15 families of type III PKSs have been characterized in bacteria, fungi and plants, highlighting the utility of PKSs in the development of natural product libraries for therapeutic development.

Slide #26 | Polyketides PK

Slide #27 | PKS – structural diversity

Slide #28 | PKS III – chalcone synthase and flavonoids

Structural, phylogenetic and functional studies on the diversity of enzyme complexes involved in the synthesis of specific products through iterative mechanistic steps (like in an assembly line) have led to the proposal of the term *megasynthases* to refer to the complexes involved in that biosynthesis. In addition to PKSs and FASs (in fact, some authors consider FAS as a particular case of PKS), it is worth considering among megasynthases also the synthesis of nonribosomal peptides (Nonribosomal Peptide Synthase NRPS). NRP are a very diverse family of natural products with a broad range of biological activities (toxins, siderophores, pigments, etc.). NRP antibiotics, cytostatics, and immunosuppressants are in commercial use.

Slide #29 | Megasynthases

4.7 Terpenes and terpenoids

Allegedly the biggest family of natural products, the acetyl CoA metabolically-derived terpenes and terpenoids (or polyprenoids) are the products of condensation of a basic C5 scaffold called isoprene unit.

Slide #30 | Terpenes and terpenoids

The sequence of terpene biosynthesis by the mevalonate (MV) pathway begins with the condensation in the cytosol of three molecules of acetyl CoA to form hydroxymethylglutaryl CoA (HMG CoA), an intermediate which is also an intermediate in the synthesis of ketone bodies (KB, see below). Two enzymes are involved in the synthesis of HMG CoA: acetyl CoA thiolase (which condenses two acetyl CoA to form acetoacetyl CoA) and HMGCoA synthase (which condenses acetoacetyl CoA with acetyl acetyl CoA). The enzyme that transforms HMG CoA into MV, HMG CoA reductase, is found in the endoplasmic reticulum membrane, while the enzyme that forms KB is mitochondrial. MV is the precursor of the two isoprenoid precursors: isopentenyl pyrophosphate (IPP) and its isomer dimethylalyl pyrophosphate (DMAPP).

IPP and DMAPP are the two C5 starting units that, combined head-to-tail, generate geranyl pyrophosphate (GPP, C10), the basic unit for monoterpenes. An additional C5 unit (IPP) added to GPP generates farnesyl pyrophosphate (FPP C15), the basic unit of sesquiterpenes. Triterpenes C30 (e.g., steroids) form from two FPP. Two GPP form geranylgeranyl pyrophosphate (C20 GGPP) that give rise to diterpenes (C20) or tetraterpenes (C40). Additional condensations and combinatorics of the basic units generate a fabulous variety of compounds and polymers.

In addition to the MV pathway, there is a second major pathway for the synthesis of IPP: the DXP (1-deoxyxylulose 5-phosphate) or MEP (2-*C*-methyl-D-erythritol 4-phosphate) pathway (in an uninspired way, also known as non-mevalonate pathway). The main difference here is that in the MV pathway, IPP is a product of acetyl CoA condensation, whereas in the DXP pathway, IPP derives from glycolytic intermediates (namely, pyruvate and glyceraldehyde 3-phosphate). This different metabolic origin has consequences on the efficiency of the IPP biosynthesis regarding mass conservation when we start from sugar feedstocks, because they involve a different number of decarboxylative steps.

Slide #31 | Mevalonate pathway
Slide #32 | DXP (or MEP) pathway

Terpenes are hydrocarbons and attract current industrial interest as potential components of biodiesel (see chapter 7).

4.8 Cholesterol

Steroids are tetracyclic hydrocarbons that can be considered as derived from the cyclopentanoperhydrophenanthrene core. They are molecules that are found, with a very few exceptions, in eukaryotes. Cholesterol is a major member of the steroid family, a subset of the terpenoid group of natural products. In addition to being part of the cell membranes modulating their fluidity, it is the precursor of steroid hormones and bile acids. The function of cholesterol in the membranes is carried out in the bacteria by hopanoids, molecules with a similar structure to cholesterol but in which, unlike cholesterol, O_2 does not participate during its biosynthesis. This indicates that these metabolic pathways predate the accumulation of O_2 in the atmosphere and that cholesterol biosynthesis was incorporated into the metabolism of eukaryotic cells over the last 2 billion years (see chapter 1).

All the carbon atoms in cholesterol are derived from acetate. Bloch already proposed, in the 1940s, that first acetate is converted to five-carbon isoprenoid units and then these units are condensed to form a linear cholesterol precursor of 30 carbons, squalene. The squalene is folded to form four rings. This process, which requires NADPH and O_2, is part of a branched route that produces other indispensable isoprenoids such as ubiquinone, dolichol – essential for initiating N-glycosylation of proteins in the rough endoplasmic reticulum – and isopentenyladenosine – a modified tRNA base. Squalene is then cycled in two steps to form a tetracyclic steroid skeleton, lanosterol. The last sequence of reactions involves 19 enzymatic steps, starting with lanosterol and ending with cholesterol. These steps involve oxidations and losses of methyl groups with the participation of 11 O_2 molecules per one lanosterol molecule converted in cholesterol.

Slide #33 | Biosynthesis of cholesterol
Slide #34 | Impact of oxygen in metabolic evolution

Cholesterol is the precursor of steroid hormones and bile salts. Hydroxylations by cytochrome P450 monooxygenases that use NADPH and O_2 play a major role in the biosynthesis of steroid hormones and bile salts from cholesterol. The cytochrome P450 system also plays a role in the metabolism of drugs and other foreign substances, such as caffeine.

Vitamin D, which is important in the control of calcium and phosphorous metabolism, is formed from a derivative of cholesterol (7-dehydrocholesterol) by the action of UV light.

Slide #35 | Cholesterol as a biosynthetic precursor

4.9 Ketone bodies and polyhydroxybutyrate

A significant fraction of the acetyl CoA produced by β-oxidation in the mitochondria is converted to ketone bodies, KB (acetoacetate and β-hydroxybutyrate), by a process known as ketogenesis. This process of KB formation is important in situations where the concentration of oxaloacetate is low because gluconeogenesis is activated and therefore the flow through the citrate synthase is diminished. Acetoacetate, a β-oxo acid, can also undergo non-enzymatic decarboxylation and be transformed into acetone and CO_2. Acetoacetate, β-hydroxybutyrate and acetone are called KB. The first two are important metabolic fuels for many tissues, especially the heart and skeletal muscle. Under extreme conditions, when glucose is insufficient, KBs become the main source of energy for the brain. The liver releases acetoacetate and β-hydroxybutyrate, which it synthesises, and these are transported by the bloodstream. Peripheral tissues degrade the KBs to acetyl CoA by, among other activities, 3-oxoacyl CoA-transferase. The liver does not express this enzyme, which makes it an exporter of ketone bodies, without the possibility of consuming them. Ketosis is a pathological condition in which acetoacetate is produced faster than it is metabolised. It is also a symptom of diabetes, and diabetic individuals' breath carries the characteristic smell of acetone, which also manifests itself in situations of prolonged fasting.

Slide #36 | Ketone bodies

HYDROXYBUTYRATE IS ALSO A PRECURSOR of polyhydroxybutyrate (PHB), a polymer produced by bacteria to store carbon and energy and also a biobased polymer considered by the industry as a suitable biodegradable plastic.

Read Luigi Di Costanzo and David S. Goodsell's Molecule of the Month (December 2017): Bacteria build biodegradable plastic that could be better for the environment.

Slide #37 | PHB – a reserve polymer used as biodegradable plastic

4.10 Lipoproteins

Combined with proteins, lipids form lipoproteins, which serve to transport them in an aqueous medium. In the intestine, TAG from the diet forms chylomicrons, which are released into the bloodstream through the lymphatic system to be carried to various tissues. TAG synthesized in the liver forms very low density lipoproteins (VLDL) which are released directly into the blood. The different classes of lipoproteins are basically classified according to their density, which is inversely proportional to their lipid content. The different lipoproteins contain characteristic apoproteins. Some apoproteins show specific biochemical, catalytic or receptor recognition activities, and also act as passive lipid transporters.

Thus, apoC2 in chylomicrons is an activator of lipoprotein-lipase, an enzyme on the surface of endothelial cells in adipose tissue capillaries and muscle tissue that hydrolyses chylomicron TAG and VLDL. The fatty acids that penetrate into the adipocytes are substrates for the synthesis of reserve TAG. The glycerol produced passes into the bloodstream and is transformed in the liver into glycerol-3-phosphate. The fatty acids are complexed with serum albumin in order to circulate in the blood and be metabolized by various tissues.

Slide #38 | Lipoproteins

Cholesterol synthesised in the liver can be converted into bile salts to be used in digestion, or esterified with fatty acids by the action of acyl CoA:cholesterol O-acyltransferase (ACAT), an enzyme in the endoplasmic reticulum that forms cholesterol esters. These are secreted into the bloodstream as part of the VLDL. As they circulate, their TAG and apolipoprotein components are captured in muscle and fat capillaries, and the VLDLs are converted to IDL and LDL. Peripheral tissues obtain most of their cholesterol from these LDLs through receptor-mediated endocytosis. Inside the cell, cholesterol esters are hydrolyzed by a lysosomal lipase to free cholesterol, which is incorporated into the cell membrane or re-esterified by ACAT and stored.

Dietary cholesterol is transported by the blood to the chylomicrons. Thus, the liver and peripheral tissues have two ways of obtaining cholesterol: either by synthesizing it from acetyl CoA (endogenous cholesterol), as we have seen, or by capturing it from the blood by receptor-mediated endocytosis (exogenous cholesterol). Cholesterol circulates from the liver to the peripheral tissues in the LDL and returns to the liver transported by the HDL.

Slide #39 | Lipoprotein metabolism

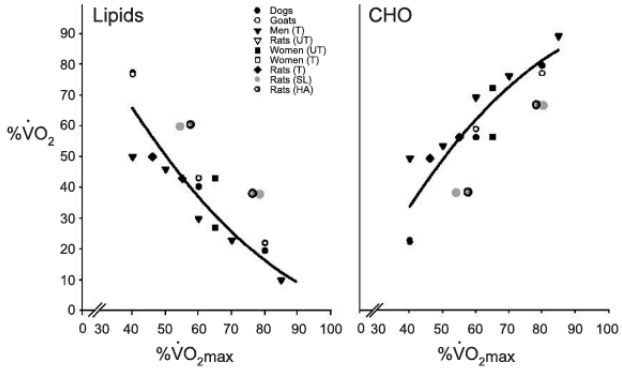
Excess cholesterol is used by the liver, which transforms it into bile acids, thus protecting the body from accumulation of water-insoluble substances that can cause the atheromas or blockages in the arteries that accompany atherosclerosis. Absence of the LDL receptor in the homozygous form of familial hypercholesterolemia leads to an elevated plasma level of LDL cholesterol and the deposition of cholesterol on blood-vessel walls, which, in turn, may result in childhood heart attacks. Cholesterol biosynthesis can effectively be blocked by statins, potent competitive inhibitors of HMG-CoA reductase. The best-selling statin is atorvastatin (Lipitor) which in 2003 became the best-selling pharmaceutical in history.

Slide #40 | Statins block cholesterol biosynthesis

The extraordinary case of the "obese athletes"

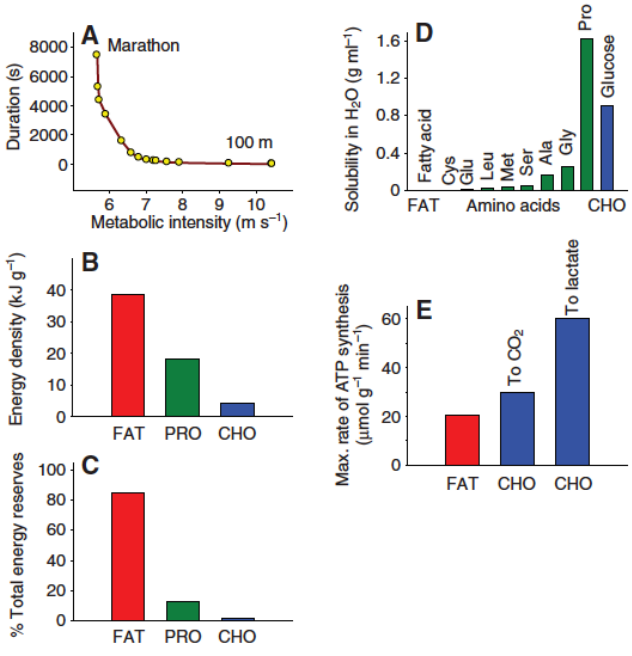
Conserved aspects of lipid metabolism in mammals

Because animals display a wide variety of locomotor skills and occupy diverse environments, is it reasonable to suggest that there are conserved aspects in their fuel selection patterns? Taylor, Weber, and Weibel systematically examined the matching of oxygen and fuel delivery using dogs and goats as models of adaptive variation. As the result of this work and other studies, it is apparent that one of the most conserved aspects of fuel selection in mammals is the proportion of lipids and carbohydrates (CHO) used during locomotion, relative to their aerobic maximum (VO_2max). Exercise data taken across taxa, body size, varying aerobic capacities and after acclimation to chronic hypoxia all show surprisingly similar patterns of fuel selection relative to exercise intensity (figure). As early as the 1930s, Edwards et al. (1934) at the Harvard Fatigue Laboratory made the observation that there was an increased reliance on glycolysis with increasing exercise intensity in humans. This phenomenon has recently been described and refined in the human exercise physiology literature as the Crossover Concept of Brooks and Mercier (1994).



Source McClelland GB (2004) Fat to the fire: the regulation of lipid oxidation with exercise and environmental stress. Comp Biochem Physiol Part B 139, 443–460.

Metabolic fuel diversity



(A) Relationship between maximum metabolic intensity and duration for different physiological tasks. To illustrate this general relationship, world track records are plotted as running time (s) vs. average speed (metabolic intensity; m s^{-1}). Running records for men [year 2009 for distances ranging from 100 m to 26.2 miles (marathon)] are indicated. (B–E) Key differences between the fuels available to support physiological tasks of varying intensity and duration [values are adapted from diverse published studies; for references, see the original paper]. The energy densities of lipids (FAT), proteins (PRO) and carbohydrates (CHO) are indicated in panel B. The relative contributions of the different fuels to the total energy reserves of the organism as a whole are presented in panel C. The solubility of metabolic fuels in water (25 °C) is shown in panel D because they must be transported through aqueous fluids (plasma, interstitial fluid and cytosol). Fatty acids are hydrophobic and have extremely low solubility in water, whereas carbohydrates are very soluble. Amino acids cover a wide range of solubilities, from cysteine (as low as fatty acids) to proline (higher than glucose). Maximal rates of ATP synthesis from different pathways (mol ATP g^{-1} wet

muscle mass min^{-1}) are indicated in panel E: lipid oxidation, carbohydrate oxidation and anaerobic glycolysis.
Source Figure 1 in Weber JM (2011) Metabolic fuels: regulating fluxes to select mix. J Exp Biol 214, 286–294.

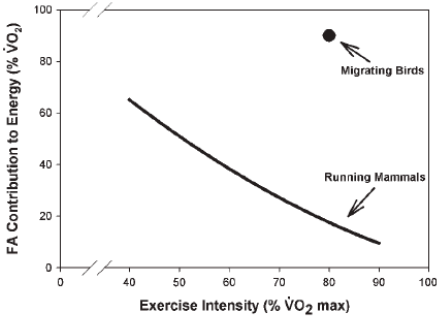


Fig. 1 The contribution of fatty acid oxidation to total energy demand as a function of relative exercise intensity for a variety of running mammals (dogs, goats, rats, and humans). A hypothetical data point for migratory birds is plotted to illustrate that unlike mammals, birds can exercise at a high intensity and use fatty acids as fuel. Modified from Fig. 2 in McClelland (2004).

Migrating birds: fly me to the moon

The metaphor of marathon running is inadequate to fully capture the magnitude of birds' long-distance migratory flight. In some respects, a journey to the Moon seems more appropriate. Birds have no access to supplementary water or nutrition during a multi-day flight, and they must carefully budget their body fat and protein stores to provide both fuel and life support. Fatty acid transport is crucial to successful non-stop migratory flight in birds. Although fat is the most energy-dense metabolic fuel, the insolubility of its component fatty acids makes them difficult to transport to working muscles fast enough to support the highly aerobic exercise required to fly. Recent evidence indicates that migratory birds compensate for this by expressing large amounts of fatty acid transport proteins on their muscle membranes (FAT/CD36 and FABPpm) and in the cytosol (H-FABP). Through endogenous mechanisms and/or diet, migratory birds may alter the fatty acid composition of the fat stores and muscle membranes to improve endurance during flight. Fatty acid chain length, degree of unsaturation, and placement of double bonds can affect the rate of mobilization of fatty acids from adipose tissue, utilization of fatty acids by muscles, and whole-animal performance. However, there is great uncertainty regarding how important fatty acid composition is to migration success or whether particular types of fatty acids (e.g., omega-3 or omega-6) are most beneficial. Migratory bats provide an interesting example of evolutionary convergence with birds, which may provide evidence for the generality of the bird model to the evolution of migration by flight in vertebrates. Yet only recently have attempts been made to study bat migration physiology. Many aspects of their fuel metabolism are predicted to be more similar to those of migrant birds than to those of non-flying mammals. Bats may be distinct from most birds in their potential to conserve energy by using torpor between flights, and in the behavioural and physiological trade-offs they may make between migration and reproduction, which often overlap.
Source Abstract and figure 1 from Guglielmo CG (2010) Move That Fatty Acid: Fuel Selection and Transport in Migratory Birds and Bats. Integr Comp Biol 50, 336–345.

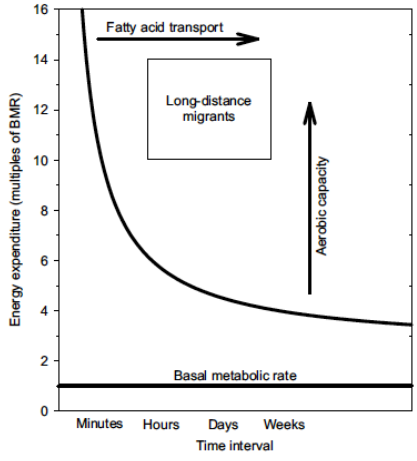


Fig. 2. The generalized relationship between rate of energy expenditure expressed as multiples of basal metabolic rate (BMR) and duration of activity is shown. Very high power output can be maintained for very short periods of seconds to minutes, whereas maximum sustained metabolic rates of ~4–5×BMR can be maintained for weeks. Long-distance migratory birds, and perhaps bats, can exercise at 10–14×BMR for periods ranging from hours to more than a week. This requires increases in both muscle aerobic capacity and fatty acid transport. Redrawn from fig. 35 in Piersma and van Gils, 2011.

Migratory birds are physiologically specialized to accumulate massive fat stores (up to 50–60% of body mass), and to transport and oxidize fatty acids at very high rates to sustain flight for many hours or days. Target gene, protein and enzyme analyses and recent -omic studies of bird flight muscles confirm that high capacities for fatty acid uptake, cytosolic transport, and oxidation are consistent features that make fat-fueled migration possible. Augmented circulatory transport by lipoproteins is suggested by field data but has not been experimentally verified. Migratory bats have high aerobic capacity and fatty acid oxidation potential; however, endurance flight fueled by adipose-stored fat has not been demonstrated. Patterns of fattening and expression of muscle fatty acid transporters are inconsistent, and bats may partially fuel migratory flight with ingested nutrients. Changes in energy intake, digestive capacity, liver lipid metabolism and body temperature regulation may contribute to migratory fattening. Although control of appetite is

similar in birds and mammals, neuroendocrine mechanisms regulating seasonal changes in fuel store set-points in migrants remain poorly understood. Triacylglycerol in birds and bats contains mostly 16 and 18 carbon fatty acids with variable amounts of 18:2n-6 and 18:3n-3, depending on diet. Unsaturation of fat converges near 70% during migration, and unsaturated fatty acids are preferentially mobilized and oxidized, making them good fuel. Twenty and 22 carbon n-3 and n-6 polyunsaturated fatty acids (PUFA) may affect membrane function and peroxisome proliferator-activated receptor signalling. However, evidence for dietary PUFA as doping agents in migratory birds is equivocal and requires further study.

Source Abstract and figure 2 from Guglielmo CG (2018) Obese super athletes: fat-fueled migration in birds and bats. J Exp Biol 221, jeb165753.

Bibliography and resources

Textbooks

Metzler[2] ch. 10, ch. 17, ch. 21.
Nelson and Cox[7], ch. 17, ch. 21.
Peretó et al.[5] ch. 15.
Stryer[9] ch. 22, ch. 26.

Papers and book chapters

Bloch K (1994) Evolutionary perfection of a small molecule. In: *Blondes in Venetian Paintings, the Nine-Banded Armadillo, and other Essays in Biochemistry*. New Haven: Yale University Press, pp. 14–36.
A very compelling exposition of how there is a fine adaptation between the structure and function of the cholesterol molecule that has been revealed by numerous in vitro and in vivo experiments.
Chan DI, Vogel HJ (2010) Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem J* **430**:1–19

On the net

LipidWeb is a good source of information on lipid classes, nomenclature, functions, etc. It belongs to a database of lipidomics (LIPID MAPS® Lipidomics Gateway).
"Cholesterol is the most highly decorated small molecule in biology. Thirteen Nobel prizes have been awarded who devoted major parts of their careers to cholesterol" (M. Brown and J. Goldstein, Nobel Lecture, 1985). Learn about some of the pioneers in the study of fats and cholesterol: Heinrich Otto Wieland, Adolf Otto Reinhold Windaus, Konrad Bloch and Feodor Lynen, Dorothy C. Hodgkin, Robert B. Woodward, Sune Bergström, Bengt Samuelsson and John Vane, Michael S. Brown and Joseph L. Goldstein, Satoshi Ōmura, Thomas C. Südhof.

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ4.1 *Learn by mapping* Build a concept map for chapter 4 and be careful about the connections with earlier chapters.

SQ4.2 Construct a table and compare fatty acid β-oxidation and biosynthesis with respect to: cell compartments involved in eukaryotic cells, acyl carrier, organization of the enzyme system, redox cofactors, stereochemistry of the hydroxylated intermediate, and direction of synthesis or degradation.

SQ4.3 Cells often use the same reaction pattern for analogous metabolic conversions. Sometimes this is the result of divergence from common ancestry (e.g., the oxidative decarboxylation of pyruvate or a-ketoglutarate). Compare the chemical steps of fatty acid oxidation with the second half of reactions (from succinate on) in the CAC. How would you address the question of whether this is a case of evolutionary convergence, or the pathways share an evolutionary past?

SQ4.4 Write a balanced equation for the synthesis of tripalmitin starting from glucose.

SQ4.5 Match the following two series of terms:

1. Cofactor of carboxylases	a. Squalene
2. Electron donor in FA biosynthesis	b. Vitamin B12
3. Propionate metabolism	c. Biotin
4. Cholesterol	d. NADPH

SQ4.6 The liver is the primary site for ketone-body (KB) synthesis. However, KBs are not used by the liver, but released for other tissues to use, albeit the liver actually does gain energy in the process of synthesizing and releasing KBs. Calculate the number of molecules of ATP generated by the liver in the conversion of palmitate (16:0) into acetoacetate.

SQ4.7 Draw the two major ways to make a glycerol-based phospholipid in eukaryotic cells using lecithin as a model (phosphatidylcholine). Which one is known as the Kennedy pathway – the predominant one in mammalian cells?

SQ4.8 What is a potential disadvantage of having many catalytic sites together on one very long polypeptide chain, like in megasynthases?

SQ4.9 Cholesterol biosynthesis depends on the activity of ATP-citrate lyase. Explain.

SQ4.10 In FA biosynthesis from glucose in mammalian cells, almost half of the reducing equivalents come from glycolysis and the rest from the oxidative pentose phosphate pathway. Explain this situation with every possible detail. Calculate the contribution of each pathway.

SQ4.11 Asthaxanthin is a red-orange carotenoid pigment produced by algae. Marine animals, such as salmons, shrimps or lobsters, that feed on the algae get their colour from the ingested carotenoid. Draw the molecular structure of astaxanthin, a C40 tetraterpene, and circle the 8 isoprene units in its structure.

SQ4.12 An old biochemistry adage is that "fats burn in the flame of carbohydrates". What is the molecular basis of this adage?

SQ4.13 *Learn by drawing* Draw a complete diagram for palmitate biosynthesis (from glucose) and catabolism (to CO_2). Include in your map the reactions of KB metabolism, as well as the connection with cholesterol biosynthesis.

SQ4.14 *Be a cultivated biotechnologist* In 1888, the poet Rubén Darío described the color blue as "the color of dream, the color of art, a Hellenic and Homeric color, an oceanic and firmament color". But he forgot to mention blue as the colour of mountain landscapes. Light scattering by tree-released isoprene is the reason of that bluish haze hanging over the mountains. In the Magnificent Molecules series (Education in Chemistry, RSC), Simon Cotton's (2015) text Isoprene relates this romantic view with the effect of isoprene derivatives on atmosphere chemistry and with applications of isoprene-derived polymers, like rubber. David S. Goodsell and Luigi di Costanzo have described isoprene synthase and the terpenoid synthase fold: Plants release a billion metric tons of isoprene and other organic gases every year (Molecule of the month, September 2016). Thus, isoprene is considered the most abundant biogenic volatile organic compound (BVOC), yet it is also a neglected greenhouse gas. Likely, fungal and microbial soil communities play a crucial role as sinks for this BVOC. Look for information on isoprene degradation (which can be aerobic or anaerobic), a neighbourhood of metabolism that remains partially unexplored.

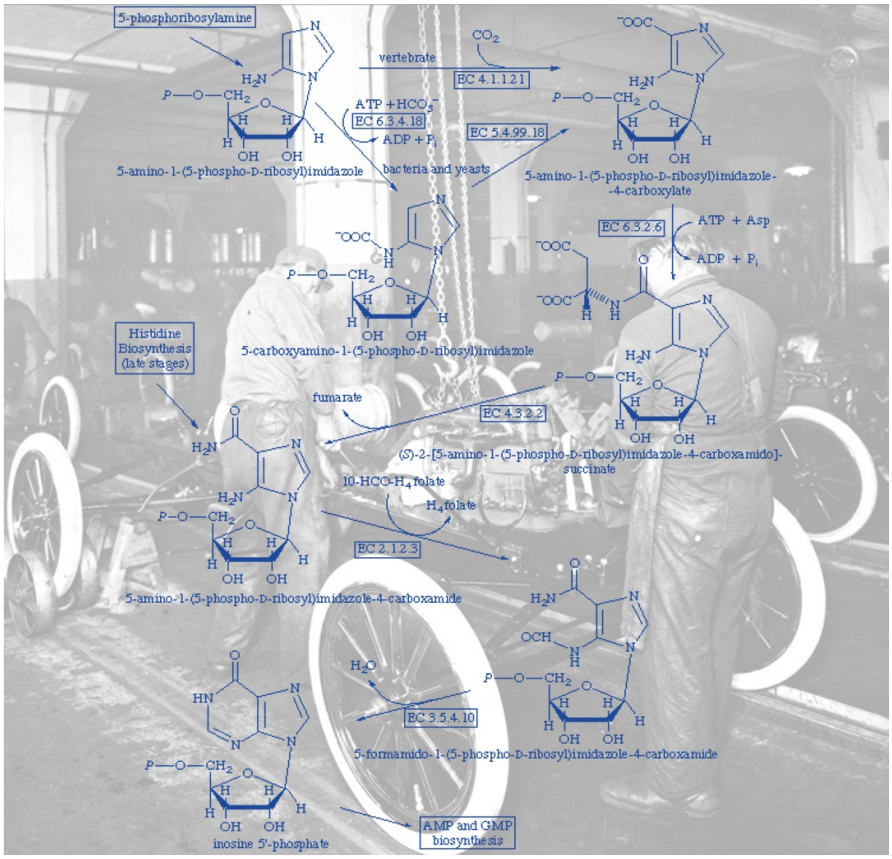
Problem-Based Learning (PBL)

PBL4 offers the opportunity to learn and consolidate the basic concepts related to chapter 4.

Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the knowledge that it is required to learn on metabolism during this course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

5

Metabolism of N compounds



Refresh your background

- ✓ Enzymology concepts (structure, mechanism, cofactors and coenzymes, enzymatic classes, kinetic parameters).
- ✓ Chemical functions in organic compounds.

Keywords

- ✓ Aminotransferase (transaminase)
- ✓ Anticancer drug targets (in thymidylate synthesis)
- ✓ *De novo* biosynthesis (of nucleotides)
- ✓ Dihydrofolate reductase
- ✓ Essential amino acid
- ✓ Glucogenic amino acid
- ✓ Glucose-alanine cycle
- ✓ Glyphosate
- ✓ Gout
- ✓ Isopeptide bond
- ✓ Ketogenic amino acid
- ✓ Lesch-Nyhan syndrome
- ✓ Metabolic channelling
- ✓ Nicotinic acid
- ✓ Nitrogen fixation
- ✓ Polyamine
- ✓ Proteasome
- ✓ PRPP (5-phosphopribosyl-1-pyrophosphate)
- ✓ Pyridoxal phosphate
- ✓ Ribonucleotide reductase
- ✓ Salvage pathways (salvage pathways)
- ✓ SAM (S-adenosylmethionine)
- ✓ Schiff base
- ✓ Shikimate pathway
- ✓ Stereoelectronic effect
- ✓ Stickland reaction
- ✓ Suicide inhibitor
- ✓ Transamination
- ✓ Ubiquitin
- ✓ Urea cycle

Learning objectives

- ✓ Identify the role of ubiquitin and explain the function of the proteasome.
- ✓ Describe the fate of nitrogen that is removed when amino acids are used as fuels.
- ✓ Explain how the carbon skeletons of the amino acids are metabolized after nitrogen removal.
- ✓ Acknowledge that amino acids can be fermentable substrates.
- ✓ Explain the centrality of nitrogen fixation to life and describe how atmospheric nitrogen is converted into biologically useful forms of nitrogen.
- ✓ Identify the sources of carbon atoms for amino acid synthesis.
- ✓ Explain the importance of one-carbon units' metabolism.
- ✓ Identify important biomolecules that are derived from amino acids.
- ✓ Explain how pyridoxal is able to catalyze so many different reactions.
- ✓ Describe the two main strategies for nucleotide biosynthesis.
- ✓ Describe the major characteristics of pathways for purine and pyrimidine nucleotide synthesis.
- ✓ Explain how deoxyribonucleotides are formed.
- ✓ Explain the ecological and biotechnological importance of aromatic compounds catabolism by bacteria.

Contents

Overview

The degradation of most amino acids begins with the transfer of the α-amino group to α-ketoglutarate, which is converted to glutamate. Carbon skeletons (α-keto acids) are incorporated to the central pathways. The ammonium ion, product of the oxidative deamination of glutamate, is quickly eliminated or transformed into a non-toxic product, such as urea, before being excreted. The availability of nitrogen, in an assimilable form, is one of the most important limiting factors for the growth of living beings. Although N_2 is the most abundant gas in the atmosphere, only some bacteria can reduce it and incorporate it into organic compounds, in a process known as biological nitrogen fixation. All the nitrogen that forms part of the amino acids (and in general all the nitrogenous compounds of living beings) is incorporated through glutamate or glutamine. The carbon skeletons of the amino acids are synthesized from a few precursors that are intermediaries of the central metabolic pathways. Because nucleotides are very important constituents (for instance, monomers for nucleic acids), all living organisms can synthesize them using similar metabolic pathways, starting form precursors supplied by the central metabolism (de novo biosynthesis of ribonucleotides). Also, nucleosides and nitrogenous bases from the digestion of nucleic acids present in the diet or from intracellular turnover can be recovered and used for the synthesis of new nucleotides (this is the so-called salvage pathway). Those that are not reused are degraded and their catabolic products are excreted.

5.1 Amino acids and nucleotide neighborhoods

Slide #2 | Amino acid and nucleotide neighborhoods

Amino acids are metabolic derivatives from intermediates belonging to central pathways (pentose phosphate pathway, glycolysis and CAC). In turn, nucleobases are direct descendants from amino acids, basically Gly, Asp and the lateral chain of Ser (mono carbon units).

5.2. Amino acids as products of protein turnover

In animals, when the amino acid content of the diet exceeds the requirements of their own proteins and other compounds, the excess amino acids ingested are degraded and the carbonaceous skeletons are metabolized for energy. When the calorie content of the diet is insufficient to meet the body's basic energy needs, endogenous proteins are hydrolyzed and the released amino acids are used as an energy source. In plants and micro-organisms, which can synthesize all the amino acids they need – unlike animals, which can only make some and for whom some amino acids become essential – the anabolic pathways are regulated so that they produce the amino acids in the right amounts (see chapter 6).

Slide #3 | Amino acids as dietary components

A PROTEIN PRODUCED AT A PARTICULAR STAGE of an organism's development, or in response to a particular environmental situation, may later be no longer needed and will have to be degraded. The continued existence of certain proteins in the cell could be harmful, which is why they are eliminated by proteolysis. Proteins have different half-lives, ranging from many days to minutes or even less than a minute. The enzymes that degrade most rapidly are those that occupy key points in metabolism regulation. This means that the levels of these enzymes are rapidly controlled. When the cell no longer needs them, they are immediately destroyed, but they can also be rapidly resynthesized when required. Other enzymes that do not perform regulatory functions in the metabolism may have longer half-lives.

The intracellular stability of proteins depends on structural features of the proteins themselves that are signs of degradation. Some of these structural determinants and the proteolytic

activities involved in protein turnover are known. In eukaryotes, where the degradation of intracellular proteins has been studied further, a protein complex (multi-catalytic and with diverse endoproteolytic activities) called the proteasome is the most important system in the elimination of proteins. The proteasome rapidly and selectively hydrolyses those proteins that have been marked by ubiquitin, a protein of only 76 amino acids with a highly conserved sequence. Ubiquitin binds to proteins, through an ATP-consuming reaction, in an isopeptide bond. This bond is established between its terminal carboxyl group and the subsequent amino of one or more specific Lys of the proteins to be degraded. The same bound ubiquitin is a target for ubiquitination (on its Lys48), which leads to the formation of ubiquitin polymers. And it is really these polymers that label the proteins for their elimination by the proteasome. The proteasome is barrel shaped, and proteolysis occurs within it. In a process that consumes ATP, the complex unfolds the proteins to facilitate entry into the barrel structure and produce more efficient proteolysis. The complex also contains an isopeptidase activity that, prior to proteolytic degradation, releases ubiquitin for further use. The ubiquitination of proteins also has other functions, different from labelling for destruction. One example is its involvement in transcriptional regulation.

The structural signals of the target proteins for ubiquitination or other proteolytic clearance mechanisms are diverse. Those mentioned below participate, more or less directly, as degradation signals: a) the so-called N-terminal rule relates the half-life of a protein to the identity of its N-terminal amino acid residue. This rule operates in all organisms, although with different versions, from bacteria to mammals. As we have already mentioned, in eukaryotes it is part of the ubiquitous/proteasome system. b) In short-half-life proteins, there are one or more segments rich in the amino acids Pro, Glu, Ser and Thr – PEST with the symbolism of a letter – flanked by basic residues. c) Oxidation of residues of Lys, Arg, Pro and Cys is also a mark for intracellular proteolysis. d) Certain post-translational chemical modifications, such as phosphorylation, are also recognition signals for degradation systems.

There are homologues of proteasome proteins in prokaryotic organisms, but no ubiquitin has been found. The labelling of prokaryotic proteins for intracellular degradation is different from eukaryotic and possibly non-covalent in nature. Nevertheless, in Asgard archaea (allegedly the living representatives of the archaeal partner of the eukaryotic last common ancestor) harbor a ubiquitin modification system.

Read David Goodsell's texts on ubiquitin and proteasome: Ubiquitin is used to tag obsolete proteins for destruction (Molecule of the month, December 2004). Proteasomes destroy damaged or obsolete proteins inside cells (Molecule of month, October 2013).

Slide #4 | Amino acids as products of protein turnover

Slide #5 | Ubiquitin modification system

The carboxy-terminal Gly residue of Ub becomes covalently attached to the e-amino groups of several Lys residues on a protein destined to be degraded (or to the K48 of another Ub, left image shows a Ub dimer). The energy for the formation of the isopeptide bonds comes from ATP.
Ubiquitin conjugation. The Ub-activating enzyme E1 adenylates Ub (1) and transfers the Ub to one of its own Cys residues (2). Ub is then transferred to a Cys residue in the Ub-conjugating enzyme E2 (3). Finally, the Ub-protein ligase E3 transfers the Ub to a Lys residue on the target protein (4a and 4b).

Slide #6 | Ubiquitin – the kiss of death

The proteasome digests the Ub-tagged proteins. The 20S proteasome comprises 28 homologous subunits arranged in four rings of 7 subunits each. A 19S cap is attached to each end of the 20S catalytic unit.
The proteasome and other proteases generate free amino acids. Ubiquitinated proteins are processed to peptide fragments from which the Ub is subsequently removed and recycled. The peptide fragments are further digested to yield free amino acids, which can be used for biosynthetic reactions, most notably protein synthesis. Alternatively, the amino group can be removed, processed (e.g., to urea) and excreted, and the carbon skeleton can be used to synthesize carbohydrate or fats or used directly as a fuel for cellular respiration.

5.3 Amino acid catabolism

Despite the range of metabolic pathways of amino acids, there are many common stages. The degradation of all amino acids – either from intestinal digestion of exogenous proteins provided by the diet (see slide #2) or from intracellular protein turnover (slide #6) – starts with the elimination of the α-amino group. This is achieved by transferring the amino groups to an α-keto acid, usually α-ketoglutarate, which delivers the corresponding α-keto acid of the amino acid, and the amino group remains in the glutamate molecule. This reaction is catalyzed by a family of enzymes called aminotransferases or transaminases. They all use pyridoxal phosphate (PLP, see below for details) as a cofactor and differ in their preference for the amino acid on which they act. Two transaminases that are particularly important due to their abundance are Asp transaminase and Ala transaminase, which catalyze, respectively, the reactions:

aspartate + α-ketoglutarate ↔ oxaloacetate + glutamate

alanine + α-ketoglutarate ↔ pyruvate + glutamate

Transamination reactions are reversible and are also considered to be anaplerotic reactions of the CAC (chapter 2, section 2.6). Through transamination, the α-amino groups of most of the amino acids that undergo degradation are collected into a single type of molecule, glutamate. Subsequently, in mitochondria, glutamate will undergo oxidative deamination by a glutamate dehydrogenase catalyzed reaction, in which α-ketoglutarate is regenerated to be involved in new transamination reactions, and the amino group is released as an ammonium ion. The reaction requires NAD^+ or $NADP^+$ as an oxidizing agent and is fully reversible, reflecting the fact that the opposite reaction can be involved in the synthesis of amino acids.

Slide #7 | Amino acid catabolism – nitrogen removal

Slide #8 | Direct deamination of Ser and Thr

As a result of these two reactions – transamination and oxidative deamination of glutamate – all the amino acids are converted into the corresponding α-ketoacids and ammonium ion. The latter is eliminated by excretion, as it is toxic for the cells, while carbon skeletons can undergo different types of degradation reactions, depending on the nature of the amino acid. There are amino acids whose carbon skeletons generate glycolytic or CAC intermediates and are precursors of carbohydrates by gluconeogenesis: they are therefore called glucogenic. On the other hand, some

produce acetyl CoA or acetoacetate and are therefore are named ketogenic. It must be said that this classification is only valid for organisms that do not perform the glyoxylate cycle (chapter 2, section 2.7).

Slide #9 | Carbon skeletons of degraded aa emerge as major metabolic intermediates

Slide #10 | Glucogenic and ketogenic aa

5.4 Nitrogen excretion and the urea cycle

As stated before, the catabolic processes of the nitrogenous compounds, mainly the deamination of the amino acids, generate ammonium ion, which in physiological conditions immediately establishes the balance: $NH_4^+ \leftrightarrow NH_3 + H^+$. The accumulation of ammonia is toxic for the cells. Thus, those cells with an active amino acid catabolism must be able to excrete or detoxify ammonia as quickly as it is produced. Many aquatic organisms simply excrete ammonia into the surrounding water, where it is immediately diluted. Because they eliminate nitrogen in this way, they are said to be ammonotelic organisms. By contrast, terrestrial organisms, which live in an environment where water is not as abundant, have evolved to develop metabolic processes that convert ammonia into a non-toxic end product that requires less water for excretion. Thus, ureotelic organisms, which are the majority of terrestrial vertebrates, are those that transform ammonia into urea before excretion. Others, including birds and reptiles, excrete nitrogen in the form of uric acid and are therefore called uricotelic organisms. Uric acid, which is synthesized by the same route followed for the synthesis of purine nucleotides (see below), is quite insoluble in water and is excreted as a semi-solid paste containing very little water. In contrast, urea, which is very soluble, removes a significant amount of water by osmosis during excretion.

Slide #11 | Strategies for N excretion

UREA IS ALMOST EXCLUSIVELY SYNTHESIZED IN THE LIVER. Its biosynthetic pathway is cyclical and was discovered by Krebs and Henseleit in 1932. The cycle is a shared pathway between mitochondria and the cytosol of the hepatocytes. The arginase enzyme is responsible for the cyclic nature of the urea formation pathway. All organisms can synthesize arginine from ornithine through the same reactions, but only ureotelic organisms form urea, because they possess arginase.

The overall reaction of the cycle can be written as

$CO_2 + NH_4^+ + 3ATP + aspartate + 2H_2O \rightarrow urea + fumarate + 2ADP + 2P_i + AMP + PP_i$

The synthesis of urea is energetically expensive, because the equivalent of four ATP is consumed in each cycle. Urea is transported by the bloodstream to the kidneys for excretion.

Slide #12 | Urea synthesis begins with the formation of carbamoyl phosphate

The consumption of 2 ATP molecules makes the synthesis of CP essentially irreversible. The mammalian enzyme requires N-acetylglutamate (NAG) for its activity. CP synthesis occurs along a tunnel inside CP synthase.

Slide #13 | Urea cycle

The carbamoyl group in CP has a high transfer potential because of its anhydride bond. This group is transferred to Orn to form Cit (by Orn transcarbamoylase). Note the mitochondrial localization of Glu deamination, and CP and Cit syntheses. Asp is the second N donor in the cycle (synthesis of AS, driven by the pyrophosphorolytic cleveage of ATP; remember, once more, the role of pyrophosphatase). Argininosuccinase cleaves AS into Arg and fumarate (corresponding to the carbon skeleton of Asp). Arginase hydrolyzes Arg to form urea and regenerate Orn. The synthesis of fumarate in the urea cycle is important because it is a precursor for glucose synthesis. Fumarate is hydrated to malate, which is in turn oxidized to OAA. This metabolite can be converted to glucose by gluconeogenesis or transaminated to Asp.

Slide #14 | Other forms of N excretion

Although urea is manufactured only in the liver, virtually all animal tissues and organs can degrade amino acids and, therefore, generate ammonia. There are mechanisms involved in the transport of ammonium – in a non-toxic form – from the different tissues to the liver, to transform it into an excretion product. For most cells, the enzyme glutamine synthase incorporates ammonium into glutamate and produces glutamine. This is transported in the bloodstream to the liver, where it is hydrolyzed by the enzyme glutaminase. The ammonium thus released is incorporated into the urea cycle. The muscle also uses another process to eliminate ammonium, the so-called glucose-alanine cycle. Muscle cells in intense activity generate large quantities of pyruvate by glycolysis, a part of which is transformed into alanine by transamination. The α-ketoglutarate, produced in the transamination reaction, is transformed back into glutamate by the action of glutamate dehydrogenase. The alanine released into the blood is transported to the liver, where it is transformed into pyruvate by reverse transamination. The amino groups are used for the synthesis of urea, while the pyruvate, through gluconeogenesis, is converted into glucose, which returns to the blood. This cyclical process is an example of metabolic economy, because in the form of a single molecule, alanine, pyruvate and ammonium ion are transported from the muscle to the liver, which will then use them.

Slide #15 | Peripheral tissues transport N to the liver

5.5 Amino acid fermentations

Many N compounds (amino acids, purines, pyrimidines) are fermented by a variety of microoganisms. A number of single amino acids can serve as energy and carbon source for anaerobes: Ala, Gly, Thr, Lys, Asp and Glu. Glutamic acid seems to be the preferred substrate in several *Clostridium* species. Glu may be converted to CO_2, ammonia, acetate and butyrate according to this diagram:

$$2\ \text{Glu} + 2\ H_2O + H^+ \rightarrow 2\ CO_2 + 2\ NH_4^+ + 2\ \text{acetate} + \text{butyrate}$$

Although a number of clostridial species grow with some single amino acids, many proteolytic clostridia (for instance, clostridiums such as: *C. perfringens, C. difficile, C. sporogenes*, and *C. botulinum*) prefer to ferment mixtures of amino acids, through the so-called Stickland reaction. They carry out coupled oxidation-reductions between pairs of amino acids. One amino acid, e.g., Ala, is oxidized, and a second one, e.g., Gly, is reduced.

Slide #16 | Amino acid fermentations

5.6 Nitrogen fixation

One of the limiting factors for the growth and development of living beings is the availability of nitrogen in a usable form. All organisms can convert ammonia into organic nitrogen, i.e., they can form a diversity of C-N bonds. But only a few have the ability to synthesize NH_3 and organic nitrogen compounds starting from N_2, the most abundant atmospheric gas. The reduction of N_2 to NH_3, called biological nitrogen fixation, is carried out only by some bacterial species, sometimes in a symbiotic relationship with plants. Among N_2 fixing bacteria, there are some that live in the soils – such as those of the genus *Klebsiella* or *Azotobacter* – others are cyanobacteria, and there are others that live in symbiosis – for example the genus *Rhizobium* – with leguminous plants and induce the formation of the root nodules where N_2 fixation occurs.

Slide #17 | N fixation and amino acid biosynthesis

Slide #18 | N biological cycle

All living beings can use NH_3 to generate organic nitrogen compounds. Bacteria of the genus *Nitrosomonas* use ammonia as a respiration electron donor and convert it into nitrite (NO_2^-). Another bacterial genus, *Nitrobacter*, obtains energy from the oxidation of nitrite to nitrate (NO_3^-). Nitrate and nitrite can be reduced to NH_3 by most plants and microorganisms that possess the enzymes nitrate and nitrite reductase. Some soil bacteria use nitrate as a final electron acceptor in respiratory chains and produce N_2 gas, which escapes into the atmosphere, while at the same time causing denitrification of the soil. Finally, and closing the cycle, only some prokaryotic organisms have the capacity to reduce N_2 to NH_3, a process known as biological nitrogen fixation.

The triple bond of the nitrogen molecule ($N \equiv N$) has a bond energy of 945 kJ/mol, which makes N_2 extraordinarily stable and difficult to reduce. In fact, industrial reduction of dinitrogen is carried out under very high pressures and temperatures above 500 ^oC – in what is known as the Haber-Bosch process.

The stoichiometry of the biological fixation reaction of N_2 can be written as follows:

$$N_2 + 8\ e^- + 8\ H^+ + 16\ ATP + 16\ H_2O \rightarrow 2\ NH_3 + H_2 + 16\ ADP + 16\ P_i$$

where you can see that it is a really energy expensive process. Reduction equivalents are electron transport proteins, such as ferredoxin. H_2 is a by-product of the reaction, although there are species that can recycle it to provide electrons for subsequent N_2 reduction cycles through the hydrogenase system. The enzymatic machinery responsible for fixing N_2 is called the nitrogenase system and has two components: component I or nitrogenase, a protein with several subunits containing iron and

molybdenum atoms, which catalyzes the reduction of N_2, and component II or nitrogenase reductase, which transfers electrons from ferredoxin to component I. Nitrogenase is rapidly inactivated by O_2; therefore, the different N_2-fixing species have developed distinct systems to get protection against this molecule, which is so abundant in the atmosphere. For example, in the root nodules of legumes infected by the bacterium of the genus *Rhizobium*, protection is achieved through the plant's production of leghemoglobin. This protein strongly binds O_2 and, in this way, an anaerobic environment is maintained inside the nodule. In other cases, there are mechanisms of spatial compartmentalization (cyanobacterial heterocysts, for instance) or temporal separation between the process of N_2 fixation and oxygenic photosynthesis.

Slide #19 | Nitrogenase

Electrons flow from ferredoxin to the reductase (iron protein, or Fe protein) to nitrogenase (molybdenum-iron protein, or MoFe protein), in order to reduce nitrogen to ammonia. ATP hydrolysis within the reductase drives the conformational changes necessary for the efficient transfer of electrons. Reductase and nitrogenase are iron–sulfur proteins. Reductase, also called Fe protein, transfers electrons from ferredoxin to the nitrogenase. A site on the nitrogenase (the P cluster) accepts the electrons. From the P cluster, the electrons flow to the FeMo cofactor, the site where N2 is reduced to ammonia. The FeMo cofactor contains both molybdenum and iron, which are required for the reduction.

David S. Goodsell on nitrogenase: Nitrogenase uses an exotic cluster of metals to fix atmospheric nitrogen into bioavailable ammonia (Molecule of the months, February 2002).

5.7 Amino acid biosynthesis

All the α-amino groups of amino acids and, in general, all the organic nitrogen atoms that are part of living beings, are ultimately derived from the ammonium ion by means of glutamate or glutamine formation reactions. The most important, and simplest, route of ammonium incorporation is the generation of glutamate by reductive amination of α-ketoglutarate through a glutamate dehydrogenase catalyzed reaction:

$$\alpha\text{-ketoglutarate} + NH_4^+ + NAD(P)H + H^+ \leftrightarrow \text{glutamate} + H_2O + NAD(P)^+$$

This reaction works in the opposite direction in the deamination of amino acids (section 5.3). In bacteria and plants, there is a glutamate dehydrogenase specific for NADPH, which probably acts only in the direction of glutamate synthesis. In animals, on the other hand, this enzyme must intervene exclusively in the catabolism of amino acids. Animals can incorporate nitrogen from the ammonium ion, but most, if not all, of what they need is obtained from preformed organic nitrogen compounds, mainly protein amino acids, present in the diet.

Another reaction involved in the assimilation of the ammonium ion is the one catalyzed by glutamine synthase:

$$\text{Glutamate} + NH_4^+ + ATP \leftrightarrow \text{glutamine} + H_2O + ADP + P_i$$

Glutamine synthase is a key enzyme in the flow of ammonium nitrogen into organic compounds, whose activity is subject to various regulatory systems (see chapter 6).

The Km of glutamate dehydrogenase for ammonium is higher than that for glutamine synthase, and perhaps that is why glutamate biosynthesis – through the coupled reactions catalyzed by glutamine synthase and glutamate synthase – constitutes a major factor, in prokaryotes and plants, and an important alternative to the glutamate dehydrogenase catalyzed ammonia incorporation reaction, especially at low NH_4^+ concentrations, a condition in which most cells are often found. Glutamate generated by any of the above routes can then transfer the amino group for the synthesis of the other amino acids, through the reactions catalyzed by aminotransferases (section 5.3).

Slide #20 | Glu and Gln – the biological doors for N

Living beings vary widely in their ability to synthesize amino acids. Many bacteria and most plants can synthesize protein amino acids. In contrast, mammals can only make about ten in sufficient quantities to maintain their development. According to that, sometimes amino acids are classified into essential and non-essential. The former amino acids are those that must be provided by the diet, while the others can be synthesized by the animal.

Slide #21 | Amino acid biosynthetic families in plants and bacteria

The carbon skeletons for amino acid synthesis are provided by intermediates of the glycolytic pathway, the citric acid cycle, and the pentose phosphate pathway.

Slide #22 | Essential and non-essential amino acids

Human beings can synthesize 11 amino acids from simple precursors, whereas 9 amino acids cannot be synthesized and must be obtained from the diet. These two groups of amino acids are referred to as nonessential and essential amino acids, respectively. Some amino acids are nonessential to human beings because they can be biosynthesized in a small number of steps. Amino acids that require a large number of steps for their synthesis are essential in the diet because some of the enzymes for these steps have been lost in the course of evolution.
The 1993 film Jurassic Park (based on the 1990 Michael Crichton novel of the same name) features dinosaurs that were genetically altered so that they could not produce the amino acid lysine. This was known as the "lysine contingency" and was supposed to prevent the cloned dinosaurs from surviving outside the park, forcing them to be dependent on lysine supplements provided by the park's veterinary staff. In reality, no animals are capable of producing lysine (it is an essential amino acid).

Slide #23 | Some easy-to-make amino acids

Slide #24 | Shikimate pathway – aromatic aa among many other things

Slide #25 | Shikimate and chorismate – precursors of aromatic amino acids

Shikimate and chorismate are synthesized from central metabolites (PEP and E4P)*.

Slide #26 | Biosynthesis of Phe and Tyr

Note that Tyr is non-essential in humans**.

Slide #27 | Trp biosynthesis

Slide #28 | Trp synthase – a striking case of substrate channeling

Substrate channeling in enzymes could emerge under different selective pressures. In this case, indole is a molecule with a high membrane permeability and channeling is a smart strategy to avoid indole leaking from the cell.

Slide #29 | His biosynthesis

Note the connection of this pathway with the de novo biosynthesis of purines.

Slide #30 | Sulfur metabolism

*David S. Goodsell on the biosynthesis of aromatic amino acids: The weedkiller Roundup attacks a key enzyme involved in the construction of aromatic compounds (Molecule of the months, February 2018).

**Goodsell on Phe hydroxylase: An unusual cofactor is used in the synthesis of aromatic amino acids (Molecule of the month, January 2005); and on tetrahydrobiopterin: Tetrahydrobiopterin plays an essential role in the production of aromatic amino acids, neurotransmitters and nitric oxide (Molecule of the month, August 2015).

5.8 One-carbon groups' metabolism

Tetrahydrofolate, THF, a derivative of folic acid (vitamin B_9), consists of three components: a pteridine ring, p-aminobenzoate and glutamate. In nature, folates can contain a variable number of glutamate residues, between 3 and 8 in most cases, linked by amide bonds between the α-amino group and the γ-carboxyl group. THF binds single carbon units, with different oxidation yet interconvertible states, through N^5, N^{10} or both. The mono-carbon groups are captured by THF from different molecules, for example, from formate to give N^{10}-formyl-THF, from His to give N^5-formyl-THF, or from Ser to give N^5, N^{10}-methylen-THF. N^5, N^{10}-methylene-THF can also donate a one-carbon unit in an alternative synthesis of glycine. The reaction, catalyzed by glycine synthase, also requires CO_2, NH_4^+, and NADH.

THF-bound one-carbon fragments can, enzymatically, be interconverted, and NADP(H) acts as acceptor/giver of reduction equivalents in the transformations between different oxidation states. The different forms derived from THF intervene as fragment givers of a carbon in various biosynthetic reactions, related to amino acid and nucleotide metabolism.

Slide #31 | One-carbon groups metabolism

Slide #32 | SAM is the major donor of methyl groups

S-Adenosylmethionine (SAM) is synthesized from methionine and ATP in an unusual reaction in which the triphosphate of ATP is cleaved to pyrophosphate and phosphate. After donation of a methyl group by SAM, the resulting *S*-adenosylhomocysteine is cleaved to yield adenosine and homocysteine. Methionine is regenerated with the transfer of a methyl group from N^5-methyl-THF, a reaction catalyzed by methionine synthase. The use of SAM and its regeneration constitute the activated methyl cycle. SAM is a substrate for DNA methylases, enzymes that methylate DNA, and it is also a precursor for the plant hormone ethylene that induces fruit ripening.

5.9 Amino acids as biosynthetic precursors

Video #5.1 | Amino acid are precursors of many biomolecules

Commented slides on amino acids as precursors of a wide variety of biochemicals, in addition to proteins. Amino acids are precursors to immune system signals, hormones, membrane lipid constituents, and electron carriers, as well as nucleotide bases.

5.10 Why is pyridoxal so multifaceted?

Video #5.2 | Pyridoxal – my favorite coenzyme

Commented slides on PLP catalytic virtuosity. PLP-dependent enzymes catalyze, among others, decarboxylations, deaminations, racemizations, and aldol cleavages at the α-carbon of amino acids.

5.11 Nucleotide biosynthesis

The nucleotide biosynthesis pathways are essentially similar in all living organisms. The synthesis of all nucleotides involves the activated sugar 5-phospho-α-D-ribosyl-1-pyrophosphate (PRPP), which is formed by the action of the enzyme PRPP synthase, which in turn catalyzes the reaction:

ribose-5-phosphate + ATP ↔ 5-phospho-α-D-ribosyl-1-pyrophosphate (PRPP) + AMP

in which the carbon 1 of ribose-5-phosphate is activated by binding the pyrophosphate group donated by ATP.

There are two major strategies for ribonucleotide biosynthesis: de novo synthesis (directly connected to central pathways) and the salvage pathways. Deoxyribonucleotides are metabolic derivatives of ribonucleotides.

Slide #34 | Nucleotide metabolism – major strategies

In the de novo synthesis of ribonucleotides, it should be noted that the precursor of the purine bases is progressively synthesized on the phosphoribose moiety, while the precursor of the pyrimidine bases (orotate) is synthesized in free form before joining the phosphorylated pentose. In the pyrimidine ring, the origin of the amino group for the synthesis of carbamoyl phosphate is usually ammonia from glutamine. In prokaryotes, the same carbamoyl phosphate synthase (CPS) is involved in the synthesis of both pyrimidines and arginine, while in eukaryotes there are two isozymes: one (CPS I) is located in the mitochondria and is involved in the synthesis of urea and arginine (see section 5.4), whereas the other one (CPS II) is cytosolic and involved in the synthesis of pyrimidines.

Slide #35 | The de novo biosynthesis of ribonucleotides

Slide #36 | Metabolic origins of the nucleobases

Slide #37 | Pyrimidine ring is assembled from bicarbonate, glutamine and aspartate

Slide #38 | Stages of the de novo biosynthesis of pyrimidine ribonucleotides

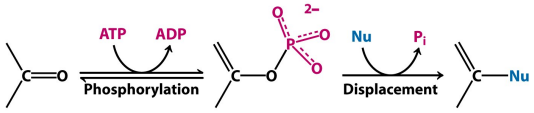
Cytidine is formed from the uracil base of UMP, but the synthesis can take place only after UMP has been converted to UTP. NMPs are converted into NTPs in stages: (1) NMPs are converted to NDPs by specific nucleoside monophosphate kinases, for instance, UMP kinase,

$$UMP + ATP \leftrightarrow UDP + ADP$$

(2) NDPs and NTPs are interconverted by nucleoside diphosphate kinase, an enzyme that has broad specificity, in contrast to NMP kinases. X and Y represent any of several ribonucleosides or even deoxyribonucleosides:

$$XDP + YTP \leftrightarrow XTP + YDP$$

DURING THE DE NOVO BIOSYNTHESIS OF PURINE NUCLEOTIDES, the purine ring system is assembled on the ribose phosphate moiety. The initial step is the displacement of pyrophosphate by ammonia, rather than by a preassembled base, catalyzed by Gln phosphoribosyl amidotransferase. The enzyme has two domains: one is homologous to the phosphoribosyltransferases participating in salvage pathways, see below; the other produces ammonia from Gln (albeit it is distinct from the domain that performs the same function in CP synthase). Ammonia is channeled through a tunnel in the enzyme. The purine ring is assembled by successive steps of activation by phosphorylation, followed by displacement by nucleophiles (nu):



Nine additional steps are required to assemble the purine ring. Remarkably the first six steps are analogous reactions. Most of these steps are catalyzed by enzymes with ATP-grasp domains that are homologous to those in CP synthase. Each step consists of the activation of a carbon-bound oxygen atom (typically a carbonyl oxygen atom) by phosphorylation, followed by the displacement of the phosphoryl group by ammonia or an amine group acting as a nucleophile (Nu).

Slide #39 | de novo biosynthesis of purine nucleotides – the strategy

Slide #40 | de novo biosynthesis of purine nucleotides – in more detail

(1) Gly is coupled to the amino group of phosphoribosylamine. (2) N10-formyl-THF transfers a formyl group to the amino group of the Gly residue. (3) The inner amide group is phosphorylated and converted into an amidine by the addition of ammonia derived from Gln. (4) An intramolecular coupling reaction forms the five-membered imidazole ring. (5) Bicarbonate adds first to the exocyclic amino group and then to a carbon atom of the imidazole ring. (6) The imidazole carboxylate is phosphorylated, and the phosphate is displaced by the amino group of Asp.

Slide #41 | de novo biosynthesis of purine nucleotides – in more detail (2)

Inosinate formation. The removal of fumarate, the addition of a second formyl group, and cyclization complete the synthesis of inosinate, a purine nucleotide.
Generating AMP and GMP. Inosinate is the precursor of AMP and GMP. AMP is formed by the addition of Asp followed by the release of fumarate. GMP is generated by the addition of water, dehydrogenation by NAD^+, and the replacement of the carbonyl oxygen by $-NH_2$ derived in the hydrolysis of Gln.

MOST ORGANISMS ALSO HAVE THE ABILITY TO SYNTHESIZE NUCLEOTIDES FROM FREE NUCLEOSIDES OR BASES that come from the diet or from the replacement of their nucleic acids, through *salvage pathways*.

Thus, phosphoribosyl-transferases catalyze reactions in which 5'-nucleosides are generated from free nitrogenous bases and PRPP. For instance, free purine bases (from the turnover of nucleotides or from the diet) can be attached to PRPP to form the corresponding NMPs (the reaction is analogous to the formation of orotidylate). Two salvage enzymes, with different specificities, recover purine bases.

Adenine phosphoribosyltransferase:

$$\text{adenine} + \text{PRPP} \leftrightarrow \text{adenylate (AMP)} + PP_i$$

Hypoxanthine-guanine phosphoribosyltransferase (HGPRT):

$$\text{guanine} + \text{PRPP} \leftrightarrow \text{guanylate (GMP)} + PP_i$$

$$\text{hypoxanthine} + \text{PRPP} \leftrightarrow \text{inosinate (IMP)} + PP_i$$

Apparently, those reactions are trivial ones, but some disorders associated to their malfunction evidence their vital importance. For example, the Lesch-Nyhan syndrome, a genetic disease characterized by self-mutilation, mental deficiency, and gout, is caused by the absence of HGPRT. Similar salvage pathways exist for pyrimidines. Pyrimidine phosphoribosyltransferase will reconnect uracil, but no cytosine, to PRPP.

Thymine, a product of DNA degradation, is salvaged by first being incorporated into a nucleoside by thymidine phosphorylase.

$$\text{thymine} + \text{deoxyribose-1-phosphate} \leftrightarrow \text{thymidine} + P_i$$

Thymidine kinase then generates the nucleotide (TMP).

Slide #42 | Nucleobases can be recycled by salvage pathways

David S. Goodsell on salvage pathways: Cells salvage and recycle their obsolete DNA and RNA (Molecule of the month, July 2012).

DEOXYRIBONUCLEOTIDES ARE THE BIOSYNTHETIC PRECURSORS of DNA in all cells. They are synthesized in virtually all organisms through the reduction of the corresponding ribonucleotide diphosphate, thanks to the action of the ribonucleotide diphosphate reductase (rNDP reductase). The same enzyme reduces the four ribonucleoside diphosphates using NADPH as the ultimate electron donor, although they are transferred to rNDP reductase by thiorredoxin or glutaredoxin. The deoxyribonucleotide diphosphates generated by this reaction are phosphorylated by the enzyme dinucleotide-diphosphokinase, with the production of the corresponding dNTPs. Because dNTPs are used almost exclusively for DNA synthesis, there are regulatory mechanisms that coordinate the synthesis of DNA and dNTP precursors.

dTTP, the precursor nucleotide required for DNA synthesis, is derived from dUTP. dUTP is rapidly hydrolyzed by the activity of dUTP diphosphohydrolase (dUTPase), after its formation. In addition, this prevents uracil from being incorporated into the DNA. dUMP now serves as a

substrate for the synthesis of dTMP, in a reaction catalyzed by thymidylate synthase. By means of methylation and reduction, with N^5,N^{10}-methylene-THF acting as a donor of the methyl groups and reduction equivalents, dUMP is transformed into dTMP. The dihydrofolate (DHF) generated in this reaction is reduced to THF by NADPH in the reaction catalyzed by the dihydrofolate-reductase. The THF is re-charged with an activated monocarbon group (see section 5.8).

Subsequently, dTMP is phosphorylated twice by the action of nucleoside monophosphate kinase and nucleoside diphosphokinase, to give rise to dTTP. This is the substrate for DNA polymerase. There is an alternative route for the synthesis of dTMP in which dCDP is defosphorylated to dCMP and dCDP is deaminated by the action of dCMP deaminase, which transforms it into dUMP. As before, dUMP is methylated and thus converted into dTMP. Whether thymine nucleotides are generated more by one route than the other depends on the type of cell and the organism.

Slide #43 | Ribonucleotides are metabolic precursors of deoxyribonucleotides*
Slide #44 | RNRase diversity
Slide #45 | Thymidylate is formed by methylation of deoxyuridylate
Slide #46 | Several anticancer drugs block the synthesis of thymidylate**
Slide #47 | Folate-dependent NADPH production
Slide #48 | Primacy of oxPP and folate pathways in NADPH production

*David S. Goodsell on ribonucleotide reductase: Ribonucleotide reductase creates the building blocks of DNA (Molecule of the month, October 2019).
**Goodsell on dihydrofolate reductase: DHFR is a target for cancer chemotherapy and bacterial infection (Molecule of the month, October 2002).

5.12 Aspects of nucleobases and aromatics catabolism

Nucleotides are produced by the intracellular degradation of nucleic acids (e.g., mRNA turnover) or, in animals, by the digestion of nucleic acids ingested with food. This digestion takes place in the intestine by the action of endonucleases (deoxyribonuclease and pancreatic ribonuclease), phosphodiesterase exonucleases and nucleotidases. The nucleosides produced can be absorbed directly by the cells of the intestinal mucosa or can be hydrolyzed to free and pentose bases by the action of nucleosidases, or, more frequently, undergo phosphorolysis by the action of nucleoside-phosphorylases, which catalyze the reaction in which the nucleoside is converted to a free base and ribose-1-phosphate. Inside the cells, the replacement of the nucleic acids takes place through similar reactions. The nitrogenous bases and nucleosides, products of these two processes, can be used for the synthesis of new nucleic acid molecules through the recovery routes, as we will see later, but an important part is completely degraded and their catabolic products excreted.

Nucleotide catabolism generates uric acid and several metabolic intermediates. In a first step, the ribonucleotides are dephosphorylated by the 5'-nucleotidases. Later, and mainly by the activity of the nucleoside-phosphorylases, the separation between the nitrogenous bases and ribose takes place. The latter, in the form of ribose-1-phosphate, after conversion to ribose-5-phosphate by the action of phosphoribomutase, can be incorporated into the pentose phosphate pathway. The nitrogenous bases are catabolized, and the degradation products are excreted. The degradation of deoxyribonucleotides is similar, but generates deoxyribose-1-phosphate, which is oxidized to ribose-1-phosphate prior to its incorporation into the pentose phosphate pathway.

Xanthine oxidase catalyzes the conversion of xanthine into uric acid, the final common pathway for the degradation of purine nucleotides. Uric acid ionizes to form urate. Primates directly excrete the uric acid produced in the degradation of the purines, while other organisms process it further before excreting it. Thus, other mammals oxidize it to allantoin; teleostean fish transform it to allantoic acid; amphibians and cartilaginous fish to urea; and marine invertebrates to ammonia and carbon dioxide.

Slide #49 | Purine catabolism

High blood levels of urate induce gout, a painful disease that results from the accumulation of urate crystals in the joints. Administration of allopurinol, a suicide inhibitor of the oxidase, relieves the symptoms of gout. Purines are then excreted as xanthine and hypoxanthine. Urate is a potent antioxidant, so the selective advantage to having relatively high blood urate levels may be the prevention of oxidative damage. Adenosine deaminase, an important enzyme in the degradation of AMP, deaminates adenosine to form inosine. A lack of adenosine deaminase results in severe combined immunodeficiency disorder (SCID). Patients suffering from SCID have compromised immune systems and are susceptible to recurring infections that can result in death.

Slide #50 | Disruptions in purine catabolism can cause pathological conditions

David S. Goodsell on xanthine oxidoreductase: Xanthine oxidoreductase helps break down obsolete purine nucleotides (Molecule of the month, September 2009).

MICROBIAL BIODEGRADATION OF ENVIRONMENTAL POLLUTANTS is a field of growing importance because of its potential use in bioremediation and biocatalysis. Some microorganisms and microbial communities have developed the ability to process recalcitrant, often xenobiotic compounds that are not part of their central metabolism (CM) by transforming them into compounds that can enter their CM. Such biodegradation processes have enormous potential for environmental cleanup (bioremediation) and in biocatalysis (green chemistry). The main properties found by Pazos et al. (2003) include the position of central hubs and basic ancient functions close to the CM, with large and difficult-to-degrade compounds more concentrated in the periphery. All the analyses point to a

model of growth from the CM towards the more diversified reactions, a model that may be connected with the history of biodegradation on Earth.

Many bacteria, including species of *Azotobacter*, *Ralstonia* and numerous species of *Pseudomonas*, possess genes for the degradation of aromatic compounds through the *meta pathway*. These genes are often found on plasmids, such as the TOL plasmids, that contain all of the genes required for the complete degradation of toluene, xylene, naphthalene and other aromatic compounds to TCA cycle intermediates. This degradative pathway consists of two parts: an "upper pathway" that converts the compounds to their carboxylic acid derivative, and a "lower pathway" that converts the (substituted) benzoic acids to CAC intermediates via catechol. The lower pathway, also known as the meta cleavage pathway, is induced by benzoic acids, or by upper pathway substrates. The most comprehensively studied meta-cleavage pathway is that of the TOL plasmid pWWO, which encodes a toluene degradation pathway in *P. putida*.

Slide #51 | Catabolism of aromatic substances

Bibliography and resources

Textbooks

Metzler[2], ch. 10, 17, 21.
Nelson and Cox[7], ch. 18, ch. 22.
Peretó et al.[5], ch. 16.
Stryer[9], ch. 23, ch. 24, ch. 25.

Papers and book chapters

Pazos F, Valencia A, de Lorenzo V (2003) The organization of the microbial biodegradation network from a systems biology perspective. *EMBO reports* **4**: 994-999.

On the net

According to Wikipedia, inborn errors of metabolism form a large class of genetic diseases involving congenital disorders related to metabolism. The majority are due to defects in single genes that code for metabolic enzymes. In most of the disorders, problems arise due to accumulation of substances which are toxic or interfere with normal functions, or to the effects of a reduced ability to synthesize essential compounds. Inborn errors of metabolism are now often referred to as congenital metabolic diseases or inherited metabolic disorders. The term 'inborn errors of metabolism' was coined by a British physician, Archibald Garrod (1857–1936), in 1908. He is known for work that prefigured the "one gene-one enzyme" hypothesis, based on his studies on the nature and inheritance of alkaptonuria. His seminal text, *Inborn Errors of Metabolism*, was published in 1923. There are several genetic diseases associated to enzymes belonging to the N compounds metabolism: phenylketonuria, alkaptonuria, CPS I deficiency, porphyria, Lesch-Nyhan syndrome, etc. DDIEM is a manually curated, ontologically formalized, freely available knowledge base of drugs, therapeutic procedures, and mitigated phenotypes related to inborn errors of metabolism.

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ5.1 *Learn by mapping* Build a concept map of chapter 5 and try to establish connections with other chapters.

SQ5.2 Name the ketoacids that form by transamination of the following amino acids: alanine, aspartate, glutamate, and isoleucine. Justify whether or not these amino acids can be classified as glucogenic.

SQ5.3 Match the two columns:

1. Ribonucleotide reductase	a. Labels protein to be degraded
2. Pyridoxal	b. Carrier of one-carbon units
3. Ubiquitin	c. Deoxyribonucleotide biosynthesis
4. Tetrahydrofolate	d. Transfers amino groups between ketoacids

SQ5.4 How many molecules of ATP are required to synthesize one molecule of CTP from scratch?

SQ5.5 *Learn by drawing* Draw the complete pathway: (a) from aspartate to ribose; (b) from palmitate to glutamate.

SQ5.6 Write a balanced equation for the synthesis of TMP from dUMP that is coupled to the conversion of serine to glycine.

SQ5.7 Four high-transfer-potential phosphoryl groups are consumed in the synthesis of urea (see slide #13). In this reaction, Asp is converted into fumarate. Suppose that fumarate is converted to oxaloacetate (the so-called urea bi-cycle). What is the resulting stoichiometry of urea synthesis? If

oxidative phosphorylation is taking place, how many high-transfer-potential phosphoryl groups are really spent?

SQ5.8 Folic acid deficiency anemia is a condition characterized by pale skin, irritability, loss of appetite, and fatigue. A common cause is a deficient diet lacking in foods like green leafy vegetables, beans and whole grains. The condition results in a decrease of hemoglobin synthesis and subsequent loss of red cells. Explain the relationship between folic acid and decreased hemoglobin synthesis.

SQ5.9 What is substrate channeling and how does it affect enzyme efficiency?

SQ5.10 *Be a cultivated biotechnologist* Our sense of taste is a specialized chemosensory system devoted to the evaluation of food and drink. We now have evidence that humans detect six basic tastes: sweet, salty, bitter, sour, umami (Japanese word meaning 'savory'), and fat ('oleogustus'). Accepting that taste evolved as a mechanism for survival, suggest the evolutionary pressures that could justify this diversity of genetically-coded mechanisms to detect tastes. Identify particular chemicals that elicit those tastes in an omnivorous human diet. Note the special relationship of amino acids (or their derivatives) with umami and bitter taste.

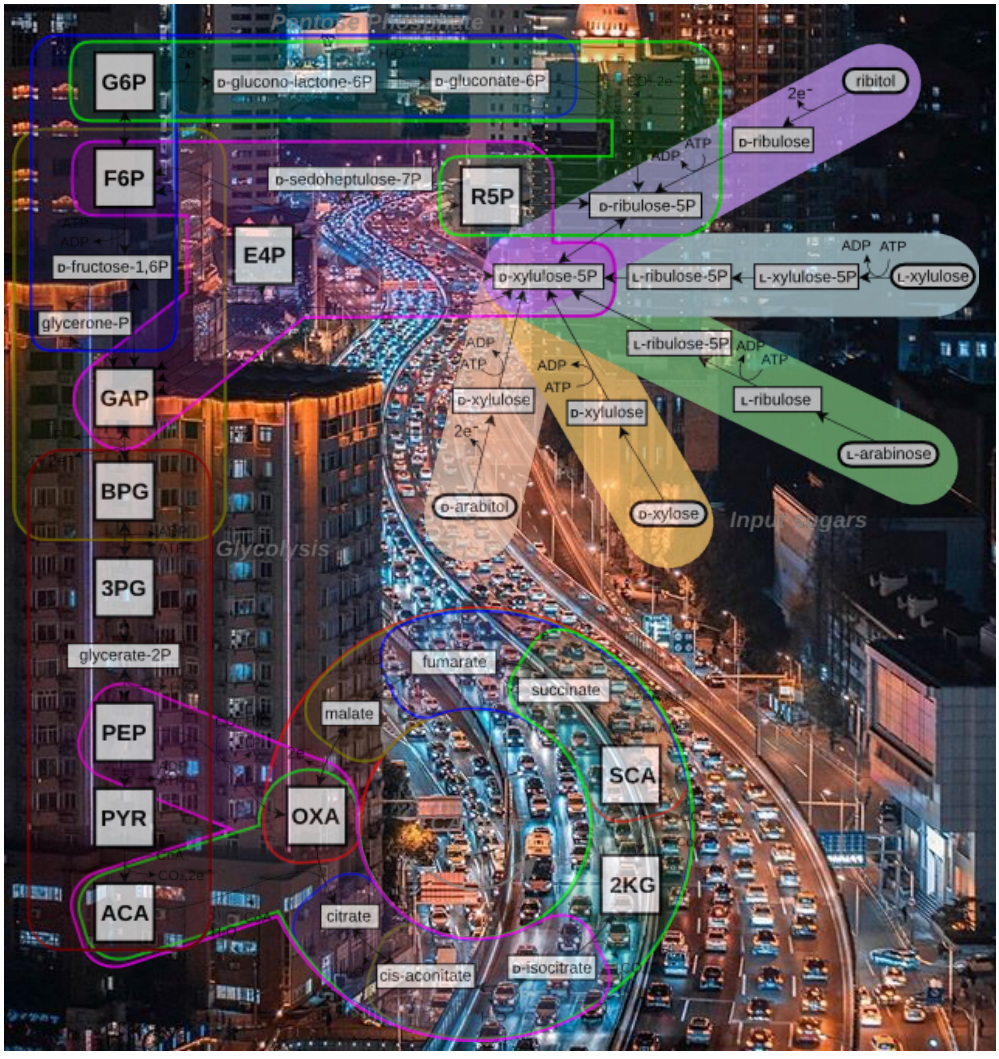
Problem-Based Learning (PBL)

PBL5 offers the opportunity to learn and consolidate the basic concepts related to chapters 5–7.

Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the knowledge that it is required to learn on metabolism during this course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

6

Integration of metabolism



Refresh your background

- ✓ Regulation of enzymes (allosterism, R and T forms, covalent modification, enzyme cascades, etc.).
- ✓ Basics on signal transduction pathways.
- ✓ Microbial metabolic modes.

Keywords

- ✓ Allosteric effect
- ✓ AMP-dependent kinase
- ✓ Bifunctional enzyme
- ✓ Catabolite repression
- ✓ Crabtree effect
- ✓ Cumulative feedback effects
- ✓ Diauxic growth
- ✓ Enzyme cascade
- ✓ Epinephrine
- ✓ Feedback and feedforward regulation
- ✓ First messenger
- ✓ Flux push / Flux pull
- ✓ Fructose-2,6-bisphosphate
- ✓ Glucagon
- ✓ Insulin
- ✓ Metabolite damage
- ✓ Modulon
- ✓ Pasteur effect
- ✓ Protein kinase A (cAMP-dependent)
- ✓ Regulon
- ✓ Second messenger
- ✓ Two-component system
- ✓ Warburg effect

Learning objectives

- ✓ Describe the major enzyme regulatory mechanisms.
- ✓ Describe feedback inhibition as a strategy in pathway regulation.
- ✓ Describe the main regulatory points in the central pathways.
- ✓ Explain the regulatory mechanism of glycogen metabolism under energy or hormonal control.
- ✓ Define the global regulation dependent on environmental clues in bacterial metabolism.
- ✓ Describe biochemical adaptations related to exercise or food intake.

Contents

Overview

Up to now, we have been studying metabolism one pathway at a time: the extraction of energy from fuels and the use of building blocks to generate all biomass constituents. Now is the moment to integrate and study aspects of metabolic regulation and control that are essential for living beings: surviving in changing environments requires rapid changes in metabolic fluxes. Sequences of reactions of the central metabolic pathways are almost the same in all organisms. Yet, in order to meet the specific needs of different cell types, the enzymes that catalyze these reactions exhibit different regulatory patterns in different organisms or at different stages in cell development. We will examine examples of regulation of central pathways, the regulation of glycogen metabolism in different physiological contexts (energy-dependent in muscle, hormone-dependent in liver), and global regulation in microorganisms as a response to environmental challenges.

6.1 Overview of regulatory mechanisms

In order to adapt to variations in environmental conditions, cells have different mechanisms for modulating the speed of metabolic reactions. On the one hand, the control of enzyme concentration depends on the balance between their synthesis and their degradation. Regulated synthesis processes include transcription of the genes encoding the enzymes and translation of the corresponding mRNA. There is a range of proteolytic enzymes and chemical labelling processes of proteins destined to destruction, which regulate degradation. However, the fastest responses are achieved by directly modulating enzymatic activity. Allosteric enzymes, those that are regulated by non-covalent interaction of ligands with allosteric binding sites, play an important role in controlling metabolic pathways. On the other hand, the greatest changes in activity are achieved through the reversible or irreversible covalent modification of enzymes. These modifications are catalysed by other enzymes. Further amplification of metabolic signals is generated when the modified enzymes are sequentially connected, forming so-called activation enzyme cascades.

Another level of metabolic organization involves the spatial distribution of pathways and enzymes in the different cell compartments. Classically, textbooks emphasize the interplay and segregation of metabolic processes between cytoplasm, mitochondria and other compartments like peroxisomes or the endoplasmic reticulum. In recent times, evidence has accumulated on the key interaction between metabolic enzymes and epigenetics in the context of essential mechanisms related to the behavior and interplay of organisms with environmental cues.

Slide #2 | Metabolic pathways and compartments
Slide #3 | Overview of metabolic regulation
Slide #4 | Mechanisms of metabolic regulation
Slide #5 | Examples of allosteric regulation

Fructose-2,6-bisphosphate (F-2,6-BP) is the most potent effector of some glycolytic and gluconeogenic enzymes. The bifunctional enzyme PFK-2/F-2,6-BPase is responsible for both the synthesis and the degradation of this effector.

Slide #6 | PFK regulation by F-2,6-BP

The regulation of the pyruvate dehydrogenase complex (PDC) influences the production of acetyl CoA from pyruvate and, therefore, the final catabolic steps of sugars and lipids, among other biomolecules. In general, the substrates of the enzyme complex are activators and the products are inhibitors. Thus, the activity is inhibited by NADH and acetyl-CoA. Furthermore, pyruvate dehydrogenase in plants and mammals is also subject to regulation by covalent modification. The complex contains a protein kinase and a phosphatase: the former phosphorylates E1 and

inactivates it, while phosphatase catalyzes the dephosphorylation and, therefore, the activation of E1. At the same time, kinase is allosterically activated by NADH and acetyl CoA, and inhibited by pyruvate. The other effector that regulates the phosphorylation state of E1 is Ca^{2+}, a phosphatase activator.

Slide #7 | Regulation of PDC

The CAC is sensitive to the availability of substrates and to the relative levels of the main products: the [NAD^+]/[NADH] ratio and the cellular energy charge. If the [NAD^+]/[NADH] ratio decreases – for example, due to an inhibition of the electron transport chain or in the case of excessive consumption of ethanol – the low concentration of NAD^+ could limit the activity of dehydrogenases. Other parameters that can affect the cycle are the amount of acetyl CoA relative to HSCoA, the amount of acetyl CoA relative to succinyl CoA, and the amount of citrate relative to oxaloacetate. Allosteric effectors and reversible covalent modification processes of cycle enzymes are also known. Thus, isocitrate dehydrogenase in mammals is allosterically activated by Ca^{2+} and ADP and inhibited by NADH, whereas in *E. coli*, phosphorylation of a Ser residue of the enzyme almost completely inactivates it. The same kinase also exhibits phosphatase activity in a different domain, which is responsible for reactivating isocitrate dehydrogenase. It is a bifunctional enzyme in which the kinase and phosphatase activities are reciprocally regulated, that is, isocitrate, oxaloacetate, pyruvate, and the glycolytic intermediates 3-phosphoglycerate and phosphoenolpyruvate allosterically activate phosphatase and inhibit at the same time. kinase. Thus, when the concentrations of glycolytic intermediates and the CAC are high in *E. coli*, isocitrate dehydrogenase is active, whereas when phosphorylation cancels the activity, the accumulated isocitrate is diverted towards the glyoxylate cycle. Furthermore, in facultative anaerobic organisms, the transcription of some genes that encode CAC enzymes is repressed in the absence of oxygen (see section 6.8).

Slide #8 | CAC regulation

6.2 Metabolic damage and repair

It is increasingly clear that (a) many metabolites undergo spontaneous or enzyme-catalyzed in vivo side reactions of 'promiscuous' enzymes or spontaneous chemical reactions, (b) the damaged metabolites formed by these reactions can be harmful, and (c) organisms have biochemical systems that limit the buildup of damaged metabolites. Second, genetic and genomic evidence from prokaryotes and eukaryotes is implicating a network of new, conserved enzymes that return a damaged molecule to its pristine state (metabolite repair) or somehow convert harmful molecules to harmless ones (damage preemption). Metabolite (that is, small-molecule) repair is analogous to macromolecule (DNA and protein) repair and seems, from comparative genomic evidence, to be equally widespread. Comparative genomics also implies that metabolite repair could be the

function of many conserved protein families lacking known activities. How – and how well – cells deal with metabolite damage affects fields ranging from medical genetics to metabolic engineering (from Linster et al. 2013 and Hanson et al. 2016).

Slide #9 | Metabolite damage and repair

Slide #10 | Damage-prone metabolites

Slide #11 | Damage products may affect regulation

6.3 Regulation in a physiological context

The metabolic patterns of each organ (in mammals, for instance, brain, muscle, adipose tissue, kidney or liver) are strikingly different. Take, for example, the liver. The function of the liver is to maintain adequate levels of combustible metabolites in the blood for use by other tissues. The liver receives all the nutrients absorbed by the intestines, except fatty acids, which are delivered to it through the portal vein from the lymph. The liver, therefore, has many biochemical functions: (1) synthesis and transformation of glucose; (2) synthesis and degradation of TAG; (3) synthesis of ketone bodies (KBs); (4) other specialised functions, including: synthesis of serum proteins, degradation of porphyrins and the nitrogenous bases of nucleic acids, storage of iron and detoxification of biologically active substances such as drugs, poisons and hormones. On the other hand, the brain does not contribute to the body's energy needs. The brain tissue has a remarkably high respiration rate. Despite constituting only 2% of the human body mass, it is responsible for 20% of the body's basal oxygen consumption. This consumption is independent of mental activity. Under normal physiological conditions, glucose is the only fuel for the brain, which consumes about 50% of all glucose used by the entire body. Nerve cells need a continuous supply of glucose. A blood concentration of glucose below half the normal level (5 mM) causes a brain dysfunction and, if it is even lower (1 mM), it causes death. Under fasting conditions, the brain obtains most of the ATP by oxidising (via CAC) the acetyl CoA, which is caused by the catabolism of KB. Fatty acids (FA) are not metabolised by the brain because they cannot cross the blood-brain barrier.

Skeletal muscle can function aerobically and anaerobically. The energy consumption of the muscle varies greatly depending on the activity it performs. Its main fuels are glucose from blood or glycogen, FA and KB. Glycogen is a readily available fuel reservoir, because it can quickly be transformed into G6P and enter the glycolytic pathway. During exercise, muscle tissue uses oxygen more rapidly than it receives the blood supply and therefore its metabolism is anaerobic and glucose is only degraded to lactate. This lactate is exported to the liver, where, through gluconeogenesis, it is transformed into glucose (Cori cycle), or reaches the heart muscle to be used as a source of energy. Muscle cells also use FA and KB as fuels, which are degraded in the mitochondria. In a fasting situation, the degradation of muscle proteins supplies most of the carbon required for glucose synthesis. On the other hand, heart muscle uses FA preferably as metabolic fuel.

Finally, adipose tissue contains large fuel reserves in the form of TAG. Adipocytes are distributed throughout the body, although they are most abundant under the skin, in the abdominal cavity, in the skeletal muscle, around the blood vessels and in the chest. A normal person weighing 70 kg has about 15 kg of fat. Fat tissue is, after the liver, the most important for the maintenance of energy homeostasis. It obtains most of the FA components of TAG from the liver or the diet. The synthesis of TAG also depends on blood glucose levels. Insulin controls the anabolism and catabolism of TAG by promoting glucose to enter adipocytes and inhibiting TAG degradation. TAG contribute to the fasting glucose supply by supplying the glycerol released from its mobilisation as a gluconeogenic substrate.

Slide #12 | Metabolic profile of organs

THE STARVED–FED CYCLE IS THE PHYSIOLOGICAL RESPONSE TO FASTING. Blood-glucose level must be maintained during fasting and subsequent eating. The typical response features the following steps: (a) the well-fed state is characterized by insulin secretion, (b) the presence of insulin stimulates glucose uptake and glycogen synthesis in muscle, adipose tissue, and liver, while suppressing gluconeogenesis in the liver, (c) insulin instead stimulates glycolysis in the liver. The early fasting state is characterized by a drop in blood glucose levels. The physiological response is a decrease in insulin secretion and an increase in glucagon secretion. Glucagon restores blood-glucose levels by stimulating glycogen breakdown and gluconeogenesis in the liver. Glucagon also stimulates fatty-acid mobilization from adipose tissue, causing a shift in fuel utilization in muscle from glucose to fatty acid. The refed state begins with the ingestion of a meal. FA are processed normally. However, the liver does not initially absorb glucose, leaving this fuel in the blood for use by other tissues. The liver remains in the gluconeogenic mode in order to replenish its own glycogen store. As glucose levels continue to rise, the liver begins to remove it from the blood and use it for FA synthesis.

Metabolic adaptations in prolonged starvation minimize protein degradation. A key metabolic priority during prolonged fasting is to maintain glucose homeostasis. During the initial stages of prolonged fasting, proteins are degraded and the carbon skeletons are used as gluconeogenic precursors. Another metabolic priority is to preserve protein. This is accomplished by shifting fuel use from glucose to FA that are mobilized by the adipose tissues for use by peripheral tissues, to allow continued use of glucose by the brain. The liver converts FA into KBs, which after several weeks of starvation become the major fuel for the brain.

Slide #13 | Metabolic profile of organs – starvation

FUEL CHOICE DURING EXERCISE is determined by activity intensity and duration. ATP is the immediate source of energy for muscle contraction. ATP can be rapidly replenished by creatine phosphate. Together, ATP and creatine phosphate can power muscle contraction for only a few seconds. Longer bouts of intense exercise require the conversion of glycogen into lactate in aerobic glycolysis, but lactate accumulation limits the duration of the effort. Exercise for longer duration

requires aerobic metabolism of glycogen and fatty acids. Because aerobic metabolism is slower, the velocity of the exercise slows with duration. Consumption of fats requires more oxygen than the combustion of glucose. The respiratory quotient (RQ) is the ration of CO_2 produced to O_2 utilized. The RQ for glucose is 1, whereas the RQ for fats is approximately 0.7. As exercise intensity increases, the RQ increases from 0.7 to 1.0.

$C_6H_{12}O_6 + 6\,O_2 \longrightarrow 6\,CO_2 + 6\,H_2O$
Glucose

$C_{16}H_{32}O_2 + 23\,O_2 \longrightarrow 16\,CO_2 + 16\,H_2O$
Palmitate

Slide #14 | Fuel choice during exercise

At rest, glycolysis is not very active. The high concentration of ATP inhibits PFK, pyruvate kinase and hexokinase. G6P is converted into glycogen. During exercise, the decrease in the ATP/AMP ratio resulting from muscle contraction activates PFK and hence glycolysis. The flux down the pathway is increased.

Slide #15 | Glycolysis – the energy charge paradigm

6.4 AMPK vs. hormone-regulated PK

AMP-activated protein kinase (AMPK) is a fuel gauge that acts as a key regulator of metabolism. In general, AMPK activates ATP-generating pathways and inhibits ATP-requiring pathways. Why is AMP a good indicator of a low energy charge? The equilibrium catalysed by adenylate kinase (ATP + AMP = 2 ADP) allows small variations in ATP concentration to be amplified into large relative changes in AMP concentration. Natural selection has favored the emergence of high affinity sites that bind AMP as a regulatory signal. Many metabolic enzymes sense the energetic status of the cell through [AMP] oscillations.

Slide #16 | Why is AMP a good regulator?
Slide #17 | Small changes in [ATP] are mirrored by large changes in [AMP]
Slide #18 | AMPK is a key regulator of metabolism

SIGNAL TRANSDUCTION PATHWAYS depend on molecular circuits connecting cell-surface receptors sensitive to chemical signals with metabolic responses. The primary messenger is a chemical signal, like a hormone, that is sensed by a receptor, usually proteins spanning the cell membrane able to connect extracellular signals with intracellular transducers. Usually, second messengers relay information from protein-ligand complexes. Second messengers are intracellular molecules that change in concentration in response to environmental signals. This is the case, for instance, of cyclic AMP and its role transducing hormonal signals that regulate central metabolic pathways. The case of glucagon effects on glycolysis and gluconeogenesis is presented.

Slide #19 | Hormone-dependent protein kinase

Slide #20 | PKA regulates PKK2/FBPase2

In summary, protein kinases that depend on intracellular signals (AMPK) control the metabolic fluxes as a response of internal energy status. On the other hand, hormone-dependent protein kinases can affect metabolic pathways in response to external information and independently of the energetic status of the cell. This is crucial for a coherent metabolic dialogue between different tissues in pluricellular organisms.

6.5 Regulation of glycogen metabolism

Glycogen is a highly branched homopolymer of glucose present in the cytoplasm of many cells, including bacteria (see chapter 3). In humans, the largest stores of glycogen are in the liver and in skeletal muscle. The liver breaks down glycogen and releases glucose into the blood to provide energy for the brain and red blood cells. Muscle glycogen stores are mobilized to provide energy for muscle contraction.

The key regulatory enzyme for glycogen degradation is glycogen phosphorylase, which is regulated by allosteric interactions and reversible phosphorylation. Phosphorylase exists in two forms: a less active *b* form and a more active *a* form. The *a* form differs from the *b* form in that a serine residue is phosphorylated. Both forms display an equilibrium between the R and T states: in the *b* form, the T state is favored, whereas in the *a* form, the R state is favored.

Liver phosphorylase and muscle phosphorylase are isozymes with different regulatory patterns. A key role of the liver is to maintain adequate blood glucose levels. As a result, the default state of liver phosphorylase is the *a* form in the R state. In essence, liver phosphorylase is prepared to generate blood glucose unless signaled otherwise. Glucose is a negative regulator of liver phosphorylase, facilitating the transition from the R state to the T state. In muscle, the default form of phosphorylase is the *b* form in the T state. When energy is needed, as signaled by an increase in the concentration of AMP, the phosphorylase binds AMP, which stabilizes the R state. The T state of phosphorylase is stabilized by ATP and glucose 6-phosphate.

Slide #21 | Muscle vs. liver phosphorylase – allosteric regulation

Phosphorylase kinase is the regulatory enzyme that converts glycogen phosphorylase *b* to phosphorylase *a* with the addition of a phosphate. This covalent modification removes a peptide loop from the active site of the *b* form, rendering the enzyme more active. Phosphorylase kinase has a subunit composition of $(\alpha\beta\gamma\delta)_4$, with the active site on the γ subunit. Phosphorylase kinase itself is activated first by Ca^{2+} binding and then by phosphorylation. The δ (delta) subunit of phosphorylase is the calcium sensor calmodulin. Phosphorylase kinase is phosphorylated by

protein kinase A. Phosphorylase kinase is maximally active when bound to calcium and phosphorylated.

Slide #22 | Phosphorylase kinase is multiregulated

Slide #23 | Muscle phosphorylase – covalent modification

Epinephrine and glucagon signal the need for glycogen breakdown. G proteins transmit the signal for the initiation of glycogen breakdown. The hormones epinephrine and glucagon bind to specific seven-transmembrane (7TM) receptors in the plasma membranes of target cells. Epinephrine binds to the β-adrenergic receptor in muscle; glucagon binds to the glucagon receptor in the liver. This causes Gs protein activation. The GTP-bound subunit of Gs adenylate cyclase, catalyzes the formation of the second messenger cAMP from ATP. The elevated cAMP concentration activates protein kinase A. Protein kinase A phosphorylates phosphorylase kinase, which activates glycogen phosphorylase, leading to glycogen degradation.

Glycogen synthase is usually inactive when in the phosphorylated *b* form and is usually active when in the unphosphorylated *a* form. Protein kinase A and glycogen synthase kinase modify the synthase. Another regulatory process for glycogen synthase is the conversion of the *b* form in the T state to the active R state by binding glucose 6-phosphate. Note that phosphorylation has opposite effects on glycogen synthase compared to glycogen phosphorylase. Glycogen synthesis is inhibited by the same glucagon- and epinephrine-signaling pathways that stimulate glycogen breakdown. Phosphorylation of glycogen synthase *a* by protein kinase A to form glycogen synthase *b* inhibits glycogen synthesis. Glycogen synthase kinase also phosphorylates and inhibits glycogen synthase.

Slide #24 | Pathway integration – The liver controls glycemia

Slide #25 | Hormone-regulated glycogen metabolism

Protein phosphatase 1 (PP1) shifts glycogen metabolism from the degradation mode to the synthesis mode. PP1 removes phosphoryl groups from glycogen synthase b, converting it into the more active a form. PP1 also removes phosphoryl groups from phosphorylase kinase and glycogen phosphorylase, inhibiting glycogen degradation.

Slide #26 | The role of protein phosphatase

6.6 Lipid metabolism regulation

Triacylglycerols are stored in adipocytes in a lipid droplet. Epinephrine and glucagon, acting through 7TM receptors, stimulate lipid breakdown (lipolysis). Protein kinase A phosphorylates perilipin, which is associated with the lipid droplet, and hormone-sensitive lipase. Phosphorylation

of perilipin results in the activation of adipocyte triacylglyceride lipase (ATGL). ATGL initiates lipids breakdown. The glycerol released during lipolysis is absorbed by the liver for use in glycolysis or gluconeogenesis.

Slide #27 | Hormone-regulated TAG metabolism

Slide #28 | TAG synthesis and mobilization

The complex regulation of cholesterol biosynthesis takes place at several levels. The liver is the major site of cholesterol biosynthesis. Regulation of cholesterol synthesis occurs by controlling the amount and activity of HMG-CoA reductase, the enzyme that catalyzes the synthesis of mevalonate. One way to do this is to control the rate of synthesis of HMG-CoA reductase mRNA with the sterol regulatory element binding protein (SREBP). SREBP resides in the ER in association with the SREBP cleavage activating protein (SCAP). When cholesterol levels fall, SCAP escorts SREBP to the Golgi complex, where it is proteolytically processed and activated. The activated SREBP moves to the nucleus to stimulate reductase mRNA synthesis. When cholesterol levels are adequate, SCAP binds cholesterol, which causes a structural change in SCAP that allows SCAP to bind to Insig, another endoplasmic reticulum membrane protein. The interaction between SCAP and Insig traps SCAP and SREBP in the endoplasmic reticulum membrane.

There are additional regulatory strategies for cholesterol biosynthesis. The rate of translation of the reductase mRNA is inhibited by nonsterol metabolites of mevalonate. Increases in cholesterol concentration result in the proteolytic degradation of reductase. In the presence of sterols, a subset of Insig proteins that are associated with ubiquitinating enzymes binds to reductase. The membrane-bound portion of the reductase is ubiquitinated, extracted from the membrane, and destroyed by the proteasome. Phosphorylation of the reductase by an AMP-activated protein kinase (AMPK) inactivates the enzyme. It allows cholesterol synthesis to cease when the ATP level is low.

Slide #29 | Multiregulation of cholesterol metabolism

6.7 N compounds metabolism regulation

Feedback inhibition regulates amino acid biosynthesis. Feedback inhibition is a common means of regulating metabolic flux. In feedback inhibition, the final product in a pathway inhibits the enzyme catalyzing the committed step. For instance, the committed step in serine synthesis is catalyzed by 3-phosphoglycerate dehydrogenase, which is inhibited by serine. Branched pathways are regulated by one of several different methods. (1) Feedback inhibition and activation: If two pathways have an initial common step, one pathway is inhibited by its own product and stimulated by the product of the other pathway (e.g., threonine deaminase). (2) Enzyme multiplicity: The committed step is catalyzed by two or more enzymes with differing regulatory properties (e.g., three distinct aspartate kinases for control of synthesis of threonine, methionine, and lysine). (3) Cumulative

feedback inhibition: A common step for several pathways is partly inhibited independently by each of the various end products (e.g., glutamine synthase).

Slide #30 | Feedback regulation in aa metabolism
Slide #31 | Branched pathways
Slide #32 | Cumulative allosteric regulation – Gln synthase

Key steps in nucleotide biosynthesis are regulated by feedback inhibition. Aspartate transcarbamoylase (ATCase) regulates the synthesis of pyrimidine nucleotides. ATCase is inhibited by CTP, a classical example of feedback inhibition. Purine nucleotide synthesis is controlled in the following ways: (1) Glutamine phosphoribosyl amidotransferase catalyzes the committed step in purine synthesis and is inhibited by a number of ribonucleotides, notably AMP and GMP. (2) AMP inhibits the formation of adenylosuccinate, a precursor to AMP, whereas GMP inhibits the formation of xanthylate, a precursor to GMP. (3) GTP stimulates the synthesis of AMP, whereas ATP stimulates the synthesis of GMP. (4) PRPP synthetase is also highly regulated even though it is not the committed step in purine synthesis.

The synthesis of deoxyribonucleotides is controlled by the regulation of ribonucleotide reductase (RNRase). Each R1 subunit of the ribonucleotide reductase from *E. coli* has two allosteric sites: one that regulates enzyme activity and one that regulates substrate specificity. The binding of dATP to the activity site inhibits enzyme activity, an effect reversed by ATP. The binding of ATP or dATP to the specificity site stimulates the reduction of UDP and CDP. The binding of thymidine triphosphate (TTP) to the specificity site stimulates the reduction of GDP while inhibiting all other reductions. The increase in dGDP stimulates the reduction of ATP to dADP.

Slide #33 | Feedback regulation in nucleotide metabolism
Slide #34 | Key steps under control in purine metabolism
Slide #35 | RNRase regulation – activity and specificity

6.8 Global responses to environmental challenges

“Beyond fuelling cellular activities with building blocks and energy, metabolism also integrates environmental conditions into intracellular signals. The underlying regulatory network is complex and multifaceted: it ranges from slow interactions, such as changing gene expression, to rapid ones, such as the modulation of protein activity via post-translational modification or the allosteric binding of small molecules. In this Review, we outline the coordination of common metabolic tasks, including nutrient uptake, central metabolism, the generation of energy, the supply of amino acids and protein synthesis. Increasingly, a set of key metabolites is recognized to control individual regulatory circuits, which carry out specific functions of information input and regulatory output.

Such a modular view of microbial metabolism facilitates an intuitive understanding of the molecular mechanisms that underlie cellular decision making" (Abstract, Chubukov et al. 2014).

Slide #36 | Coarse-grained regulation of metabolism

"Nutrient limitation drives competition for resources across organisms. However, much is unknown about how selective pressures resulting from nutrient limitation shape microbial coding sequences. Here, we study this 'resource-driven selection' by using metagenomic and single-cell data of marine microbes, alongside environmental measurements. We show that a significant portion of the selection exerted on microbes is explained by the environment and is associated with nitrogen availability. Notably, this resource conservation optimization is encoded in the structure of the standard genetic code, providing robustness against mutations that increase carbon and nitrogen incorporation into protein sequences. This robustness generalizes to codon choices from multiple taxa across all domains of life, including the human genome." (Abstract, Shenhav and Zeevi 2020, Science **370**:683–687)

Slide #37 | Nutrient availability has shaped the genetic code

GLOBAL REGULATION. Most mechanisms of metabolic regulation through changes in gene expression modulate more than one *operon* or *regulon*. For example, the cAMP-CRP complex regulates enzymes in more than 200 metabolic processes. These regulatory processes are referred to as *global regulation*, a *multigene system* or *pleiotropic control*. The term *modulon* is used to define a collection of operons and regulons regulated by a common effector. Stresses such as heat shock and oxidative stress result in regulation of many operons and regulons. It is not known whether this regulation is modulated by a common effector. In this case, the term *stimulon* is used to define a collection of modulons, regulons, and operons regulated by a common environmental stimulation.

Slide #38 | Global regulation – responses to environmental stimuli

Slide #39 | Global regulation – facultative anaerobes

Slide #40 | Global regulation – facultative anaerobes

Slide #41 | Control of C and N metabolism in *E. coli*

Regulatory circuits that control carbon and energy metabolism in *Escherichia coli*. Diagram of central carbon metabolism (a). Inset panels display regulatory modules controlling parts of the net glucose as a carbon source (b), coordination of local (lactose) and global (carbon supply) signals for carbon uptake by LacI and Crp, respectively (c), the fructose-1,6-bisphosphate (FBP)–Cra circuit, which regulates the switch between glycolysis and gluconeogenesis (d), control of carbon catabolism via ATP demand (e), regulation of respiration by the availability of electron acceptors (f) and the oxygen-sensing switch between aerobic and anaerobic respiration (g). αKG, α-ketoglutarate; acetyl-P, acetyl-phosphate; cAMP, cyclic AMP; EMP, Embden–Meyerhof pathway; G6P, glucose-6- phosphate; GLX, glyoxylate; OAA, oxaloacetate; PEP, phosphoenolpyruvate; PTM, post-translational modification.

"In contrast to the independently operating and mostly well-understood regulatory circuits of uptake pathways, the densely connected network of regulatory interactions within central

carbon metabolism has hampered the investigation of its regulatory circuits. Knowing the transcriptional regulatory network is mostly insufficient for understanding central metabolic operations, as metabolic control relies heavily on allosteric regulation by metabolite binding and post-translational protein modifications. The limited relevance of transcription is unsurprising given that metabolism might need to change rapidly – *E. coli* can adapt to environmental changes that reverse central fluxes in a matter of seconds. However, methods to deduce metabolite–enzyme interactions or the effect of a covalent modification on protein activity have lagged far behind expression-focused research. Nevertheless, an emerging theme is that cells rely on internal signals that are mostly independent of the exact source of carbon, and the crucial cues seem to come from a limited number of central metabolites" (Text and Table 1 from: Chabukov et al. 2014).

Slide #42 | Regulatory metabolites in *E. coli*

"Many of the tasks in the regulation of N uptake and metabolism, such as the detection of various nitrogen sources and their preferential usage, mirror those in carbon metabolism. The preferred nitrogen source for *E. coli* in most conditions is ammonium, but a wide range of organic nitrogen-containing molecules can be also used. In parallel to the internal detection of carbon substrates, *E. coli* detects many of these nitrogen sources after they are taken up (via basal expression of their transporters or via non-specific transport) and upregulates enzymes for their transport and catabolism." (Chabukov et al. 2014)

Slide #43 | Control of N metabolism in *E. coli*

"[The coordination of anabolism and catabolism of amino acids] is mostly achieved by end-product inhibition, which is a ubiquitous regulatory mechanism that balances the production of a specific amino acid with its demand but minimally affects the rest of metabolism. All 20 amino acids in *E. coli* have either been shown to inhibit the first committed step in their synthesis via allosteric regulation or, in the case of single-reaction pathways, can be assumed to do this via product inhibition. This mechanism ensures a rapid increase in synthesis in response to a higher demand, or repression of synthesis in response to an excess supply. […] To a large extent, similar principles also govern the transcriptional regulation of amino acid biosynthesis: at least ten amino acids negatively regulate the transcription of their own biosynthesis pathways. […] One of the endpoints of metabolism is the assembly of amino acids into proteins. In a fast-growing cell, protein synthesis and ribosome production account for most nutrient and energy consumption. However, as fast growth can require the devotion of up to 75% of cellular transcription to the production of ribosomes, such a program would be highly deleterious when nutrients are limited, and thus *E. coli* devotes resources to ribosome biogenesis only when resources for protein synthesis are abundant. Making this decision requires the integration of several metabolic signals, and *E. coli* uses the availability of energy sources (such as ATP and GTP) and amino acids, which are the major substrates of protein synthesis, to determine the rate of ribosome biogenesis. […] The reason that

superfluous ribosome biosynthesis would be deleterious is that cells can only modulate their total protein allocation, taking resources from the expression of one protein and devoting them to another. Given that ribosomal proteins and metabolic enzymes are the major protein fractions at high growth rates, decreased ribosome synthesis would enable increased enzyme synthesis. Conversely, this constraint of total protein allocation is a plausible explanation for why rapidly growing cells primarily rely on glycolytic energy generation and not on the more protein-intensive respiration. Complementary to the common perception of protein synthesis regulating metabolism, *E. coli* thus coordinates proteome partitioning between ribosome and metabolic enzyme synthesis via metabolic signal-dependent transcription factors – for example, by Crp and the global carbon/nitrogen availability reporter α-ketoglutarate." (Chabukov et al. 2014)

Slide #44 | Control of amino acid metabolism and translation in *E. coli*

Slide #45 | The logic of metabolic regulation

The high-level logic of *Escherichia coli* metabolic regulation. The specific regulatory interactions are summarized and overlaid on the coarse-grained metabolic network (slide #36). a | Coordination within sectors is obtained by a high-level combination of activation and inhibition by means of different molecular implementations. Individual modules maintain homeostasis and inter-pathway balance by activating their consumption and repressing production, either directly or indirectly. b | The logic that is implemented by regulatory circuits across sectors can be summarized by two major motifs implementing negative feedback loops: the 'flux push' motif, by which the high abundance of a metabolite activates its own consumption, and the 'flux pull' motif, by which the low abundance of a metabolite activates its production.

6.9 Overflow metabolism and the Warburg effect

"In our ideal view of metabolism, we think of catabolism and anabolism. In catabolism nutrients break down to carbon dioxide and water to generate biochemical energy. In anabolism nutrients break down to generate building blocks for cell biosynthesis. Yet, when cells are pushed to high metabolic rates, they exhibit incomplete catabolism of nutrients, with a lower energy yield and excretion of metabolic byproducts. This phenomenon, characterized by the excretion of metabolic byproducts that could otherwise be used for catabolism or anabolism, is generally known as *overflow metabolism*. Overflow metabolism is a ubiquitous phenotype that has been conserved during evolution. Examples are the acetate switch in the bacterium *E. coli*, Crabtree effect in unicellular eukaryote yeasts, the lactate switch in sports medicine, and the Warburg effect in cancer. Several theories have been proposed to explain this seemingly wasteful phenotype. Yet, there is no consensus about what determines overflow metabolism and whether it offers any selective advantage" (Abstract, Vazquez 2017).

Slide #46 | Fermentative metabolism in yeast

THE WARBURG EFFECT "at a glance" (Koppenol et al. 2011).

- Otto Warburg was a pioneering biochemistry researcher who made substantial contributions to our early understanding of cancer metabolism. Warburg was awarded the Nobel Prize in Physiology or Medicine in 1931 for his discovery of cytochrome *c* oxidase, not for his work on cancer and the formulation of the Warburg hypothesis.
- The Warburg effect is the reverse of the Pasteur effect (the inhibition of fermentation by O_2) exhibited by cancer cells; alteration of the Pasteur effect in cancer is linked to prolyl hydroxylases and hypoxia-inducible factor (HIF).
- Tumour suppressors and oncogenes converge on HIF to reverse the Pasteur effect and thereby induce the Warburg effect.
- Cancer cells carry out aerobic glycolysis and respiration concurrently.
- Tumour suppressors and oncogenes exert direct effects on metabolism: p53 promotes the pentose phosphate pathway and oxidative phosphorylation; MYC induces glycolysis and glutamine metabolism.
- Mutations in metabolic enzymes, specifically isocitrate dehydrogenase 1 (IDH1) and IDH2 and other citric acid cycle enzymes, are causally linked to familial and spontaneous cancers.

Slide #47 | The Warburg effect
Slide #48 | Overflow metabolism in cancer cells
Slide #49 | Metabolism of glucose and glutamine in cancer cells

6.10 Metabolic regulation and industrial fermentations

Fermentation, in the biochemical sense, is the anaerobic microbial growth based on substrate-level phosphorylation. In a more general sense, this term is also used by biotechnologists and microbiologists to describe processes producing useful materials to a large industrial scale using microorganisms. Microorganisms have acquired through evolution complex regulatory mechanisms for efficient growth, but not for the production of specific materials. To improve fermentation efficiency, the industry has developed and uses various mutants defective in regulatory mechanisms. When Fleming discovered penicillin in 1929, the fungus *Penicillium notatum* produced 1.2 mg/L penicillin. At present, the fermentation industry uses strains producing over 50 g/L penicillin. These industrial strains are mutants derived from wild-type strains, but with altered regulatory mechanisms.

Coryneform bacteria of the genera *Brevibacterium* and *Corynebacterium* are the most commonly used industrial strains in amino acid production. Industrial strains are mutants with defects in regulation. They are selected based on their resistance properties to analogues. Auxotrophic mutants are used for the fermentative production of intermediates. Metabolic

engineering approaches could target an increasing number of enzyme activities, relieve regulatory mechanisms and/or improve membrane permeability.

"With a world market of more than four million tons per year, L-amino acids are among the most important products in industrial biotechnology. The recent years have seen a tremendous progress in the development of tailor-made strains for such products, intensively driven from systems metabolic engineering, which upgrades strain engineering into a concept of optimization on a global scale. This concept seems especially valuable for efficient amino acid production, demanding for a global modification of pathway fluxes – a challenge with regard to the high complexity of the underlying metabolism, superimposed by various layers of metabolic and transcriptional control." (Abstract from Becker J and Wittmann C 2012. Systems and synthetic metabolic engineering for amino acid production – the heartbeat of industrial strain development. *Curr Op Biotech* **23**:718–726)

Slide #50 | Penicillin production

Slide #51 | Metabolic regulation and industrial fermentation

Bibliography and resources

Textbooks

Berg et al.[9] ch. 16, 17, 21, 22, 24, 25, 26 and 27
Metzler[2] ch. 10, 17 and 21
Peretó et al.[5] ch. 6, 14, 15, 16 and 17

Papers and book chapters

Chubukov V et al. (2014) Coordination of microbial metabolism. *Nature Rev Microbiol* **12**:327–340.
Hanson AD et al. (2016) Metabolite damage and metabolite damage control in plants. *Annu Rev Plant Biol* **67**:131–152.
Koppenol W, Bounds P, Dang C (2011) Otto Warburg's contributions to current concepts of cancer metabolism. *Nat Rev Cancer* **11**:325–337. https://doi.org/10.1038/nrc3038
Linster CL et al. (2013) Metabolite damage and its repair or pre-emption. *Nature Chem Biol* **9**: 72–80.
Vazquez A (2017). *Overflow Metabolism: From Yeast to Marathon Runners*. Academic Press.

On the net Databases collected under MetaCyc/BioCyc (https://metacyc.org/) websites contain detailed information on metabolic regulation, specially from model organisms like *E. coli* (EcoCyc: https://ecocyc.org/).

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions are can be found in a separated PDF file.

SQ6.1 *Learn by mapping* Build a concept map of chapter 6 and try to establish connections with the rest of chapters.

SQ6.2 Phosphorylation has opposite effects on glycogen synthesis and breakdown. What is the advantage of its having opposing effects?

SQ6.3 Match the two columns:

1. Major strategy of amino acid biosynthesis regulation	a. Glucagon
2. Allosteric activator of PFK-1	b. Feedback inhibition
3. Stimulates glycogen breakdown in the liver	c. Epinephrin
4. Blocks glycogen synthesis in muscle	d. F-2,6-BP

SQ6.4 The respiratory quotient (RQ) is a classic metabolic index defined as the volume of CO_2 released divided by the volume of O_2 consumed. (a) Calculate the RQ values for the complete oxidation of glucose and of tripalmitoylglycerol. (b) What do RQ measurements reveal about the contributions of different energy sources during intense exercise? (Assume that protein degradation is negligible).

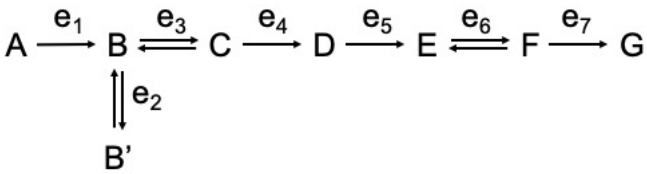
SQ6.5 How does the activation of AMPK during aerobic exercise foster the switch to fatty acid oxidation for distance running?

SQ6.6 Suppose that a promoter mutation leads to the overproduction of protein kinase A in adipose cells. How might fatty acid metabolism be altered by this mutation?

SQ6.7 What are the regulatory means that prevent high levels of activity in glycolysis and gluconeogenesis simultaneously?

SQ6.8 Define overflow metabolism. Why do the Warburg effect observed in cancer cells and the Crabtree effect in *Saccharomyces cerevisiae* qualify as examples of overflow metabolism?

SQ6.9 Examine the metabolic pathway shown below. Which of the enzymes is likely to be the allosteric enzyme that controls the synthesis of G? *Hint*: find the "committed step" in that pathway.



SQ6.10 *Be a cultivated biotechnologist* According to the Oxford Dictionary, *serendipity* is "the occurrence and development of events by chance in a happy or beneficial way". There are many instances of serendipity in the history of technology invention and scientific discovery. Search the etymology of the term *serendipity* and look for information about the serendipitous discovery of some drugs, for instance, Viagra (an inhibitor of PDE5, the phosphodiesterase that hydrolyzes cGMP) and compactin (an analogous to lovastatin, inhibitor of HMG-CoA reductase that lowers serum cholesterol levels).

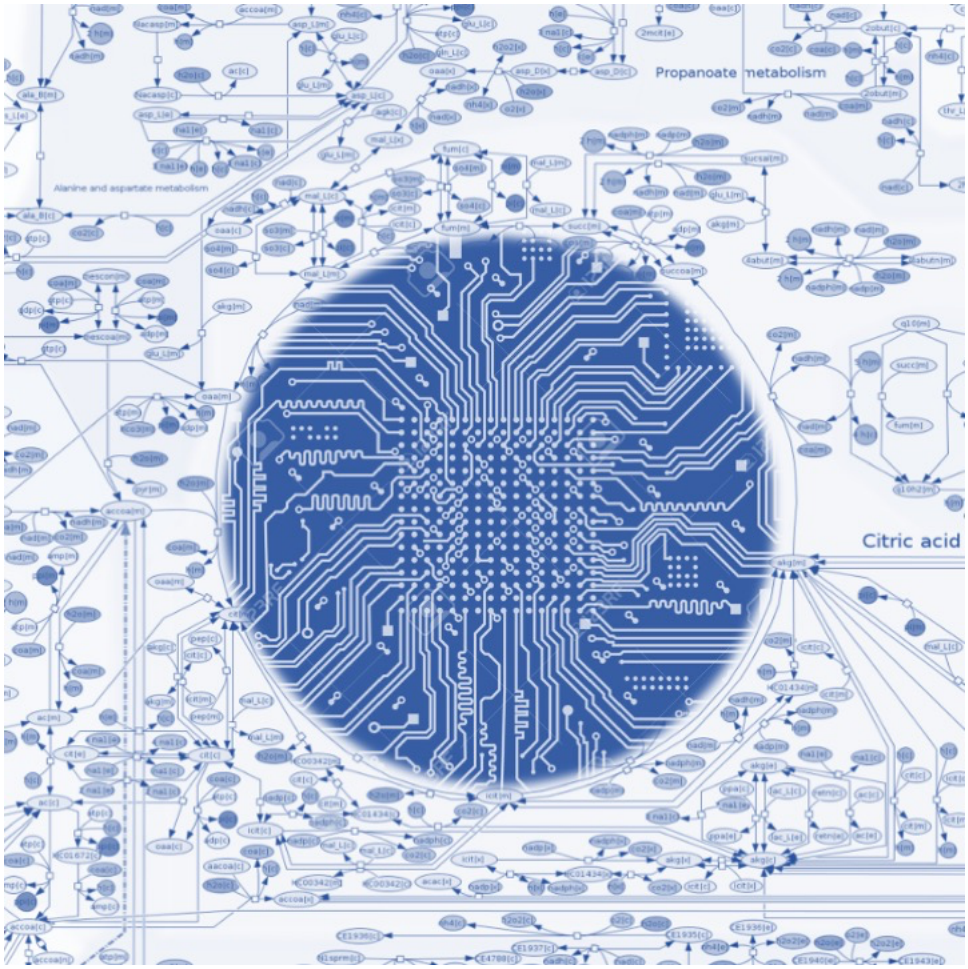
Problem-Based Learning (PBL)

PBL5 offers the opportunity to learn and consolidate the basic concepts related to chapters 5–7.

Cautionary note on these handouts *By no means do these Study Guides on Metabolism and Regulation aim to present all the knowledge that it is required to learn on metabolism during this course. Please, according to your previous knowledge and your specific necessities, refer to the cited bibliography.*

7

Metabolic engineering



Refresh your background

- ✓ General overview of the stoichiometric relationships in central pathways..
- ✓ General principles of gene and genome modification.

Keywords

- ✓ Artemisinin
- ✓ Biofuel
- ✓ Chassis
- ✓ Cofactor balance
- ✓ Energy 'surplus' pathways
- ✓ Experimental evolution
- ✓ Genome-scale metabolic model
- ✓ Golden rice
- ✓ Industrial systems biology
- ✓ Native (inherent) pathway
- ✓ Noninherent (synthetic) pathway
- ✓ Nonoxidative glycolysis
- ✓ Recursive (iterative) biosynthesis
- ✓ Reverse glyoxylate pathway
- ✓ Scaffolding
- ✓ Synthetic biology
- ✓ Synthetic microcompartments
- ✓ Theoretical (stoichiometric) pathway yield
- ✓ Transmetabolism

Learning objectives

- ✓ Describe major strategies to improve fluxes in known metabolic networks
- ✓ Identify the main problems related to the modification of fluxes
- ✓ Characterize alternative pathways according to their stoichiometric efficiency
- ✓ Describe the main strategies of transmetabolism
- ✓ Learn that metabolic systems engineering accelerates the development of efficient microbial cell factories by integrating traditional metabolic engineering with systems biology, synthetic biology, and evolutionary engineering
- ✓ Compare the current strategies in metabolic systems engineering with the ideal of synthetic biology

Contents

Overview

Metabolic engineering considers cells as true chemical factories. The first approach to the control and modification of metabolic fluxes was purely intuitive: which fluxes to close and which to promote in order to obtain a greater yield of the desired product. Recombinant DNA technology was of great help at this stage, but we have no idea how many projects failed for every success story we know. Metabolic modelling and experimental evolution have enabled systemic strategies that bring us closer to the ideal of synthetic biology. Metabolic designs absent in nature and solutions that are the result of artificial selection open new paths to transmetabolism, the exploration of what is possible beyond the natural.

7.1 Overview of metabolic engineering

In our more remote past, hundreds of thousands of years ago, hunter-gatherers were content with what nature offered them. Since the Neolithic Revolution, with the domestication of plants and animals, artificial selection (i.e., natural selection manipulated by humans) has offered us an extraordinary variety of food sources, shelter, transport assistance, etc. Today, modern biotechnology has greater knowledge, tames microorganisms and molecules, and open to us a wide range of opportunities to solve issues related to agriculture, industry or the environmental crisis, to give but a few examples.

Slide #2 | Domestication

Slide #3 | A recurrent environmental problem with a potential biotechnological solution

WHAT IS METABOLIC ENGINEERING?

Abstract from Bailey (1991)

Application of recombinant DNA methods to restructure metabolic networks can improve production of metabolite and protein products by altering pathway distributions and rates. Recruitment of heterologous proteins enables extension of existing pathways to obtain new chemical products, alter posttranslational protein processing, and degrade recalcitrant wastes. Although some of the experimental and mathematical tools required for rational metabolic engineering are available, complex cellular responses to genetic perturbations can complicate predictive design.

Abstract from Stephanopoulos and Vallino (1991)

In order to enhance the yield and productivity of metabolite production, researchers have focused almost exclusively on enzyme amplification or other modifications of the product pathway. However, overproduction of many metabolites requires significant redirection of flux distributions in the primary metabolism, which may not readily occur following product deregulation because metabolic pathways have evolved to exhibit control architectures that resist flux alterations at branch points. This problem can be addressed through the use of some general concepts of metabolic rigidity, which include a means for identifying and removing rigid branch points within an experimental framework.

Excerpt from Stephanopoulos et al. (1998)

Metabolic engineering is about the analysis and modification of metabolic pathways. [...] The concept of metabolic pathway manipulation for the purpose of endowing microorganisms with desirable properties is a very old one indeed. We have many outstanding examples of this strategy in the areas of amino acids, antibiotics, solvents, and vitamin production. [...] Shortly after the feasibility of recombinant DNA technology was established, various terms were coined to represent the potential applications of this technology to directed pathway modification. Some of the terms suggested were molecular breeding, in vitro evolution, pathway engineering, cellular engineering and metabolic engineering (Bailey 1991, Stephanopoulos and Vallino 1991). [...] Here we define metabolic engineering as *the directed improvement of product formation or cellular properties through the modification of specific biochemical reaction(s) or the introduction of new one(s) with the use of recombinant DNA technology*.

Slide #4 | Metabolic engineering – two seminal papers and a foundational book

As a general summary, we can highlight some keywords from the above texts, such as *improvement*, *directionality*, *specificity* or *recombinant DNA*. With the perspective of the last three decades, we can add that metabolic engineering also involves computational and modelling approaches, as well as considering the evolvability of biological systems as an ally for their engineering. What remains true, in any case, is that cells are seen as true chemical factories.

Slide #5 | Metabolic engineering – key words

Metabolism is the result of the evolutionary exploration of space of the chemically possible. We only know a fraction of the metabolic map generated by evolution. Metabolic engineering has moved within the limits of natural metabolism with the benefit of bioprospecting actual mechanism. More recently, biotechnologists' nonconformity has led them to invent new metabolic designs in an endeavor that we can call transmetabolism (see below and Peretó 2020).

Slide #6 | The possible and the actual

Slide #7 | Bioprospecting – discovery within the actual

Slide #8 | Metabolic engineering and transmetabolism

7.2 The cell factory

Growing concerns over limited fossil resources and associated environmental problems are motivating the development of sustainable processes for the production of chemicals, fuels and materials from renewable resources. Metabolic engineering is a key enabling technology for transforming microorganisms into efficient cell factories for these compounds. Systems metabolic engineering, which incorporates the concepts and techniques of systems biology, synthetic biology and evolutionary engineering at the systems level, offers a conceptual and technological framework to speed the creation of new metabolic enzymes and pathways or the modification of existing pathways for the optimal production of desired products. (Excerpt from the abstract, Lee *et al.* 2012).

Slide #9 | Microbial chemical factories

Slide #10 | Examples of commercialized bioproducts

Slide #11 | From chemically produced (D.L) Met to fermentative L-Met

Slide #12 | Vanillin – natural non-inherent

Slide #13 | Taxol – natural non-inherent by consortia

Slide #14 | Bioproduced building blocks

Slide #15 | Biosynthetic potential for biofuels

Video #7.1 | Biofuels

Commented slides on biofuels as a study case in metabolic engineering. The discussion includes general principles on the biosynthetic potential for biofuels, the question of theoretical vs. real yields, and the issue of pathway choice.

7.3 Improving the known

Slide #17 | Metabolic engineering – improving the known

There is wide variety of approaches to metabolic engineering:

–Modification of an extant, native (inherent) pathway: Increased production of a metabolite by altering structural/regulatory genes resulting in change and/or redirection of metabolic flux.

–Expanding metabolism by building a new pathway: by expressing additional new enzymes (from other species), transfer of a completely new pathway, modifying natural enzymes by protein engineering, designing a completely new, non-natural pathway, optimizing cofactor use/recycling or redesign the spatiotemporal relationships between enzymes.

–Experimental evolution and systems biology. Further optimization steps to improve yields and performance are possible through directed evolution and computer simulations (metabolic systems engineering).

Slide #18 | A variety of approaches to metabolic engineering

Slide #19 | Modification of an extant pathway – the intuitive approach

Slide #20 | Amino acid production

Slide #21 | Increased clavulanic acid production

Slide #22 | Microbial production of SCAs

Slide #23 | Introducing a new pathway – natural, noninherent

Slide #24 | Golden Rice (1)

Slide #25 | Golden Rice (2)

Slide #26 | Golden Rice (3)

Slide #27 | A non-native pathway to artemisinic acid

Slide #28 | Semisynthesis of artemisinin

Slide #29 | Yeast-producing opiates

Slide #30 | Sugar synthesis from CO_2 in a hemiautotrophic *E. coli*

Slide #31 | Full autotrophic *E. coli*

Slide #32 | Acquired mutations from heterotrophy to autotrophy

Slide #33 | Discover your inner autotroph

Carbon fixation is one of the most important biochemical processes. Most natural carbon fixation pathways are thought to have emerged from enzymes that originally performed other metabolic tasks. Can we recreate the emergence of a carbon fixation pathway in a heterotrophic host by recruiting only endogenous enzymes? In

this study, we address this question by systematically analyzing possible carbon fixation pathways composed only of *Escherichia coli* native enzymes. We identify the GED (Gnd–Entner–Doudoroff) cycle as the simplest pathway that can operate with high thermodynamic driving force. This autocatalytic route is based on reductive carboxylation of ribulose 5-phosphate (Ru5P) by 6-phosphogluconate dehydrogenase (Gnd), followed by reactions of the Entner–Doudoroff pathway, gluconeogenesis, and the pentose phosphate pathway. We demonstrate the in vivo feasibility of this new-to-nature pathway by constructing *E. coli* gene deletion strains whose growth on pentose sugars depends on the GED shunt, a linear variant of the GED cycle which does not require the regeneration of Ru5P. Several metabolic adaptations, most importantly the increased production of NADPH, assist in establishing sufficiently high flux to sustain this growth. Our study exemplifies a trajectory for the emergence of carbon fixation in a heterotrophic organism and demonstrates a synthetic pathway of biotechnological interest.

7.4 Transmetabolism: into the nonnatural possible

"[…] Evolution does not deliver fully optimal metabolic solutions, yet just those sufficient for survival and reproduction. The fuzzy discreteness of metabolic circuits has allowed the classical strategy of metabolic engineering: the pursuit of the enhancement of cell performances upgrading metabolite fluxes of native pathways, even by means of transplantation of enzymes or pathways wired within the metabolic core of a heterologous host. Thus, metabolic engineering navigates inside the possible, playing within the limits of biological and physicochemical trade-offs: usually it is not possible to improve one trait without worsening other cellular performances. In addition to these constraints, some sub-optimal solutions in extant metabolisms are also the result of historical contingencies brought to the present biosphere by irreversible evolutionary trajectories. […]

The community of microbial biotechnologists is ready to foster a non-conformist approach to metabolic engineering, I mean, a biotechnology that goes beyond nature's subset of the chemically possible. Today, there are many successful attempts of rational design of unnatural pathways with new stoichiometric balances, as well as repurposed enzymes performing new-to-nature reactions, including the incorporation of non-biogenic chemical elements to metabolism. Powerful computational algorithms that search for new theoretical stoichiometries, and experimental evolution, allowing an efficient empirical evaluation of the possible with the help of natural selection under artificial conditions, are key approaches for the non-rational exploration of the chemically available to the biochemical factory. We can invent enzymatic arrangements that, as far as it is known by the scientific community, have not been reached by any natural system due to historical constraints in the evolutionary trajectories. Although sometimes imagination precedes discovery. […]

According to the Oxford English Dictionary, transhumanism is "the belief or theory that the human race can evolve beyond its current physical and mental limitations, especially by means of science and technology." In the next decades, transmetabolism will allow us to reach innovative and useful chemical solutions never explored by evolution, and beyond our "current physical and mental limitations". In fact, part of these limits have started to melt by searching reaction networks with algorithms, empowered by artificial intelligence, for the discovery of new unnatural pathways leading to (natural or unnatural) compounds of industrial interest. Furthermore, computer-generated networks of potential chemistries may open a window to the past, enabling empirical

tests of plausible prebiotic scenarios. The smart fusion of dry and wet biochemistries is our best baggage for this fascinating journey through transmetabolism. I am convinced that – paraphrasing an apocryphal Carl Sagan quotation – somewhere in the space of the possible, an incredible biochemical transformation is waiting to be known." Read the complete article in Peretó (2020) Microbial Biotech (DOI 10.1111/1751-7915.13691) (Open Access paper).

Slide #34 | Transmetabolism – into the nonnatural possible

Slide #35 | Metabolic engineering vs. transmetabolism

Slide #36 | Metabolic engineering vs. transmetabolism

Slide #37 | Synthetic pathways – the nonconformist approach

Slide #38 | Reverse beta-oxidation

Advanced (long-chain) fuels and chemicals are generated from short-chain metabolic intermediates through pathways that require carbon-chain elongation. The condensation reactions mediating this carbon–carbon bond formation can be catalysed by enzymes from the thiolase superfamily, including β-ketoacyl-acyl-carrier protein (ACP) synthases, polyketide synthases, 3-hydroxy-3-methylglutaryl-CoA synthases, and biosynthetic thiolases. Pathways involving these enzymes have been exploited for fuel and chemical production, with fatty-acid biosynthesis (β-ketoacyl-ACP synthases) attracting the most attention in recent years. Degradative thiolases, which are part of the thiolase superfamily and naturally function in the β-oxidation of fatty acids 5,6, can also operate in the synthetic direction and thus enable carbon-chain elongation. Here we demonstrate that a functional reversal of the β-oxidation cycle can be used as a metabolic platform for the synthesis of alcohols and carboxylic acids with various chain lengths and functionalities. This pathway operates with coenzyme A (CoA) thioester intermediates and directly uses acetyl-CoA for acyl-chain elongation (rather than first requiring ATP-dependent activation to malonyl-CoA), characteristics that enable product synthesis at maximum carbon and energy efficiency. The reversal of the β-oxidation cycle was engineered in *Escherichia coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize n-alcohols, fatty acids and 3-hydroxy-, 3-keto- and trans-Δ2-carboxylic acids. The superior nature of the engineered pathway was demonstrated by producing higher-chain linear n-alcohols ($C \geq 4$) and extracellular long-chain fatty acids ($C > 10$) at higher efficiency than previously reported. The ubiquitous nature of β-oxidation, aldehyde/alcohol dehydrogenase and thioesterase enzymes has the potential to enable the efficient synthesis of these products in other industrial organisms (Abstract from Dellomonaco C et al. (2011) Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**:355–359).

Slide #39 | Dealing with C loss

Slide #40 | Non-oxidative glycolysis (NOG)

Glycolysis, or its variations, is a fundamental metabolic pathway in life that functions in almost all organisms to decompose external or intracellular sugars. The pathway involves the partial oxidation and splitting of sugars to pyruvate, which in turn is decarboxylated to produce acetyl-coenzyme A (CoA) for various biosynthetic purposes. The decarboxylation of pyruvate loses a carbon equivalent, and limits the theoretical carbon yield to only two moles of two-carbon (C2) metabolites per mole of hexose. This native route is a major source of carbon loss in biorefining and microbial carbon metabolism. Here we design and construct a non-oxidative, cyclic pathway that allows the production of stoichiometric amounts of C2 metabolites from hexose, pentose and triose phosphates without carbon loss. We tested this pathway, termed non-oxidative glycolysis (NOG), in vitro and in vivo, in *Escherichia coli*. NOG enables complete carbon conservation in sugar catabolism to acetyl-CoA, and can be used in conjunction with CO_2 fixation and other one-carbon (C1) assimilation pathways to achieve a 100% carbon yield to desirable fuels and chemicals (Abstract from Bogorad IW et al. (2013) Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* **502**:693–697).

Slide #41 | C rearrangement in natural pathways vs. NOG

Slide #42 | Reverse glyoxylate cycle – C4 to C2

Most central metabolic pathways such as glycolysis, fatty acid synthesis, and the TCA cycle have complementary pathways that run in the reverse direction to allow flexible storage and utilization of resources. However, the glyoxylate shunt, which allows for the synthesis of four-carbon TCA cycle intermediates from acetyl-CoA, has not been found to be reversible to date. As a result, glucose can only be converted to acetyl-CoA via the decarboxylation of the three-carbon molecule pyruvate in heterotrophs. A reverse glyoxylate shunt (rGS) could be extended into a pathway that converts C4 carboxylates into two

molecules of acetyl-CoA without loss of CO_2. Here, as a proof of concept, we engineered in *Escherichia coli* such a pathway to convert malate and succinate to oxaloacetate and two molecules of acetyl-CoA. We introduced ATP-coupled heterologous enzymes at the thermodynamically unfavorable steps to drive the pathway in the desired direction. This synthetic pathway in essence reverses the glyoxylate shunt at the expense of ATP. When integrated with central metabolism, this pathway has the potential to increase the carbon yield of acetate and biofuels from many carbon sources in heterotrophic microorganisms, and could be the basis of novel carbon fixation cycles (Abstract from Mainguet SE et al. (2013) A reverse glyoxylate shunt to build a non-native route from C4 to C2 in *Escherichia coli*. *Metabolic engineering* **19**:116–127).

Slide #43 | A synthetic autotrophic pathway

Optimizing designer metabolisms in vitro Biological carbon fixation requires several enzymes to turn CO_2 into biomass. Although this pathway evolved in plants, algae, and microorganisms over billions of years, many reactions and enzymes could aid in the production of desired chemical products instead of biomass. Schwander et al. constructed an optimized synthetic carbon fixation pathway in vitro by using 17 enzymes – including three engineered enzymes – from nine different organisms across all three domains of life. The pathway is up to five times more efficient than the in vivo rates of the most common natural carbon fixation pathway. Further optimization of this and other metabolic pathways by using similar approaches may lead to a host of biotechnological applications.
Carbon dioxide (CO_2) is an important carbon feedstock for a future green economy. This requires the development of efficient strategies for its conversion into multicarbon compounds. We describe a synthetic cycle for the continuous fixation of CO_2 in vitro. The crotonyl–coenzyme A (CoA)/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle is a reaction network of 17 enzymes that converts CO_2 into organic molecules at a rate of 5 nanomoles of CO_2 per minute per milligram of protein. The CETCH cycle was drafted by metabolic retrosynthesis, established with enzymes originating from nine different organisms of all three domains of life, and optimized in several rounds by enzyme engineering and metabolic proofreading. The CETCH cycle adds a seventh, synthetic alternative to the six naturally evolved CO_2 fixation pathways, thereby opening the way for in vitro and in vivo applications (Abstract from Schwander T et al. (2016) A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**:900–904).

Slide #44 | Cell-free synthetic biochemistry

Synthetic biochemistry, the cell-free production of biologically based chemicals, is a potentially high-yield, flexible alternative to in vivo metabolic engineering. To limit costs, cell-free systems must be designed to operate continuously with minimal addition of feedstock chemicals. We describe a robust, efficient synthetic glucose breakdown pathway and implement it for the production of bioplastic. The system's performance suggests that synthetic biochemistry has the potential to become a viable industrial alternative (Abstract Opgenorth PH et al. (2016) A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nature Chem Biol* **12**:393–395).

Slide #45 | Synthetic pathways – scaffolding

Slide #46 | DNA origami and swinging arms

Slide #47 | Synthetic pathways – microcompartments

7.5 Metabolic systems engineering and synthetic biology

Sustainable production of chemicals from renewable non-food biomass has become a promising alternative to overcome environmental issues caused by our heavy dependence on fossil resources. Systems metabolic engineering, which integrates traditional metabolic engineering with systems biology, synthetic biology, and evolutionary engineering, is enabling the development of microbial cell factories capable of efficiently producing a myriad of chemicals and materials including biofuels, bulk and fine chemicals, polymers, amino acids, natural products and drugs. In this paper, many tools and strategies of systems metabolic engineering, including in silico genome-scale metabolic simulation, sophisticated enzyme engineering, optimal gene expression modulation, in vivo biosensors, de novo pathway design, and genomic engineering, employed for developing microbial cell factories, are reviewed. Also, detailed procedures of systems metabolic engineering

used to develop microbial strains producing chemicals and materials are showcased. Finally, future challenges and perspectives in further advancing systems metabolic engineering and establishing biorefineries are discussed (Abstract from Ko et al. 2020).

Slide #48 | Evolutionary approaches

Slide #49 | Industrial systems biology

Slide #50 | Systems metabolic engineering – examples

Slide #51 | The (future) synthetic biology approach

Slide #52 | Emerging technologies

Slide #53 | European missions

Slide #54 | Future directions in metabolic systems engineering

Slide #55 | Are we doing synthetic biology?

Slide #56 | Why should I study metabolism?

Bibliography and resources

Textbooks

Stephanopoulos GN, Aristidou AA, Nielsen J (1998) *Metabolic engineering. Principles and methodologies*. Academic Press, San Diego

Papers and book chapters

Bailey, JE (1991) Toward a science of metabolic engineering. *Science* **252**:1668–1675.

Choi KR, Jang WD, Yang D, Cho JS, Park D, Lee SY (2019). Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotech* **37:**817–837.

Ko YS, Kim JW, Lee JA, Han T, Kim GB, Park JE, Lee SY (2020). Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chem Soc Rev* **49**:4615–4636.

Lee JW, Na D, Park JM, Lee J, Choi S, Lee SY (2012) Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol* **8**:536–546.

Peretó J (2021) Transmetabolism: the non-conformist approach to biotechnology. *Microb Biotechnol* https://doi.org/10.1111/1751-7915.13691

Stephanopoulos, G, Vallino JJ (1991). Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**:1675–1681.

Woolston BM, Edgar S, Stephanopoulos G (2013) Metabolic engineering: past and future. *Annu Rev Chem Biomol Eng* **4**:259–288.

On the net<https://www.nature.com/subjects/metabolic-engineering> is a good source of news on metabolic engineering housed by the *Nature* publishing group.

Study Questions (SQ)

These questions can be solved by looking for hints in textbooks and other resources. They won't be discussed in the classroom. Answers to the questions can be found in a separated PDF file.

SQ7.1 *Learn by mapping* Build a concept map of chapter 7 and try to establish connections with the rest of chapters.

SQ7.2 Explain how genome sequencing and an accurate functional annotation of genes may be of great help in metabolic engineering.

SQ7.3 Match the two columns:

1. Imitates substrate channeling	a. Reverse glyoxylate shunt
2. Optimizes the conversion of C4 in C2 molecules	b. Transmetabolism
3. Optimizes the conversion of C6 in C2 molecules	c. Scaffolding
4. Engineering metabolism beyond the natural	d. Non-oxidative glycolysis (NOG)

SQ7.4 Emilia Matallana has reviewed Ed Regis's book "Golden Rice: the imperiled birth of a GMO superfood" (https://www.sebbm.es/revista/articulo.php?id=688&url=un-nacimiento-todavia-deseado). Look for information on the difficulties for the introduction of Golden Rice in the fields of concerned countries.

SQ7.5 *Be a cultivated biotechnologist* Read the article: "A Way to Brew Morphine Raises Concerns Over Regulation" by Donald G. McNeil in the *New York Times* (May 18, 2015) <https://www.nytimes.com/2015/05/19/health/a-way-to-brew-morphine-raises-concerns-over-regulation.html>. (1) Discuss the potential risks and benefits of this kind of research; (2) pick and defend one of the following positions: **(a)** We should control this research by restricting access to the yeast strains and DNA because… **(b)** Maybe we should restrict access, but it's just not feasible because… **(c)** This technology shouldn't be regulated yet, but access should be restricted in the future if… (be specific about your conditions!) **(d)** Leave the scientists alone! There is no need to regulate this research because… Do you think that scientists have an ethical obligation to proactively engage in debate about the potential misuse of their research?