15 Dual Causality

Nothing in biology makes sense, except in the light of evolution – Theodosius Dobzhansky

The stoichiometric matrix and the information associated with it fundamentally represent a biochemically, genetically, and genomically structured knowledge base. It can be used to analyze network properties and to relate the components of a network and its genetic bases to network or phenotypic functions. Biology is subject to *dual causality*, or *dual causation* [261]. It is governed not only by the physical laws but also by genetic programs. Thus, while biological functions obey the physical laws, their functions are not predictable by the physical laws alone. Biological systems function and evolve under the confines of the physical laws and environmental constraints. How organisms operate within these constraints is a function of their evolutionary history and their survival strategy.

15.1 Causation in Physics and Biology

Physics Classically, 'cause and effect' is established by formulating mathematical descriptions of conceptual models of fundamental physical phenomena. One example is molecular diffusion (see Figure 15.1). The fundamental process underlying diffusion is the random walk process that a collection of molecules undergoes. The statistical properties of the random walk process can be assessed quantitatively, and its macroscopic consequences are described with Fick's law. This law is described by a simple equation that is used as the basis to describe mass transfer processes from regions of high concentration to regions of low concentration. The established causality is the basis for computations that reliably predict mass transfer processes. The Boltzman and Nernst equations provide other specific cases of causality in physics, and there are many more examples.

Engineering design can be based on such predictions. Thus, in engineering, "there is nothing more practical than a good theory," as the physical laws can be used for design, often with minimal experimentation and prototyping.

Cause and effect for physical phenomena are often well established and can be described mathematically. Mathematical descriptions are in the form of equations and inequalities. An interesting discussion of the character of physical law is found in [114].

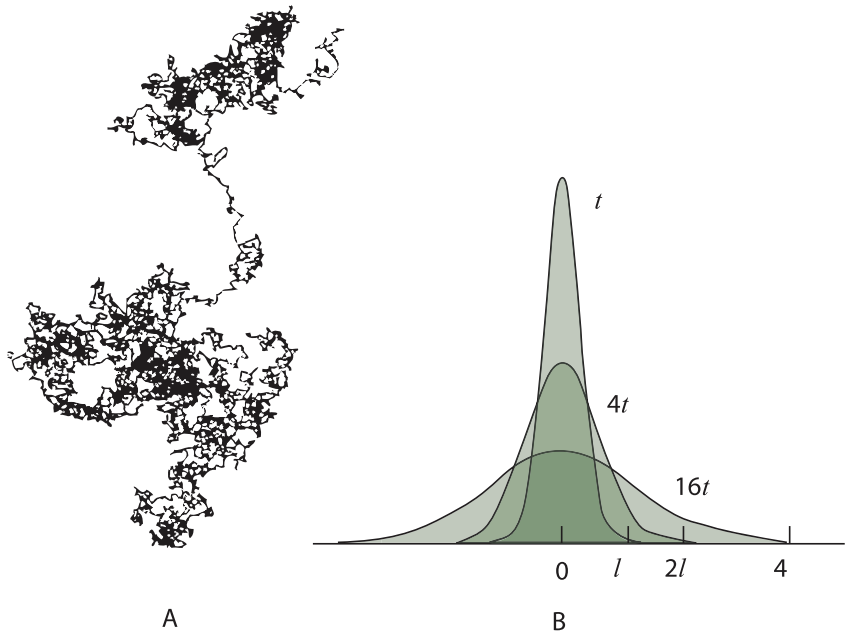

Figure 15.1 Causation in physics: an example of a random walk process and diffusion of molecules. Panel A: the simulated random walk trajectory of a single molecule. Panel B: the probability distribution for the molecule's location as a function of time when it was located at $l = 0$ at $t = 0$. The width of the distribution, l, increases with the square root of time, t. Modified from [318].

Biology Causation in biology is much different from physics. Biological causation originates fundamentally from the evolutionary process leading to genetic variation within a population. There are four key parts to the conception of an evolutionary process, shown in Figure 15.2:

1. initial phenotype resulting from a genotype;
2. natural selection of the new organism leading to the ability to produce offspring successfully;
3. successful mating leads to the possible formation of a new genotype; and
4. processes, such as mutation and recombination, that lead to the formation of a new genotype.

This process repeats itself. The result is diversity: a *biopopulation* of non-identical individuals. Therefore, living systems are time-variant: they evolve and change over time. In contrast, physical phenomena are time-invariant; e.g., oxygen, a homogeneous population of identical molecules, always diffuses the same way in water under a given set of circumstances, the unit charge on the electron does not change, and so on.

The outcome of the selection process is, in part, stochastic, and is influenced by environmental variables. The selection process in biology gives the appearance of

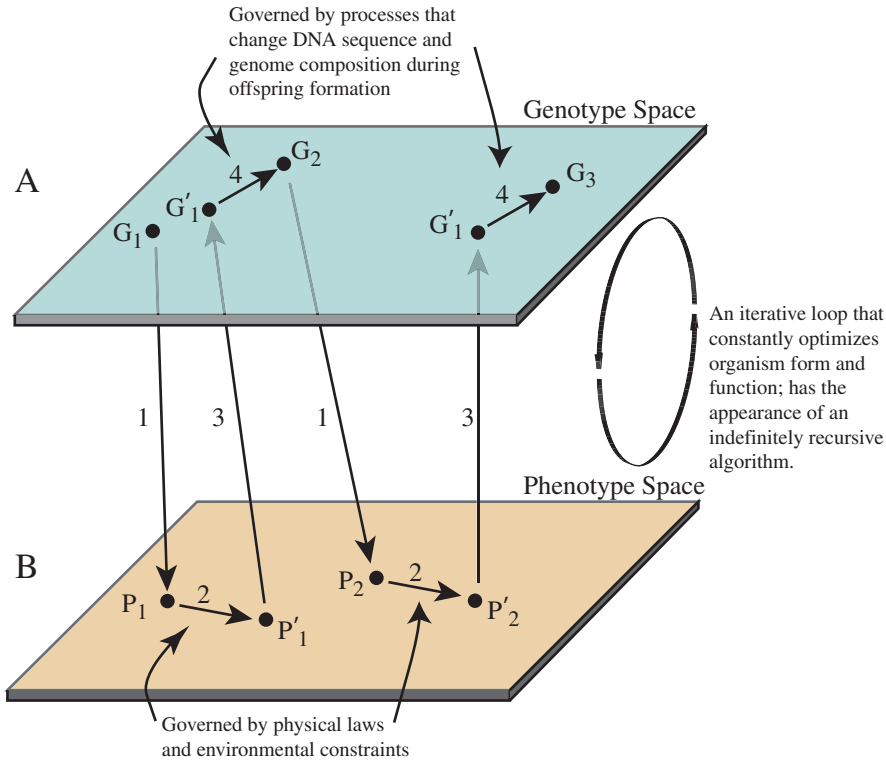

Figure 15.2 Causation in biology. The genotype–phenotype relationship conceptualized as an iterative mapping process. Panel A depicts the position of genotype (G) and Panel B a phenotype (P) in their respective spaces. Iteration over two generations is illustrated. Redrawn based on [173].

'sense of purpose' that, fundamentally, is survival. The sense of purpose for quantitative model-building is represented by an objective function (see Chapter 21) that is meant to describe the basis of selection. In general, it is hard to know what the detailed objective is underlying a selection process. Thus, the objective function becomes the focus of study as one seeks to understand the selection process and the distal causation it represents. The objective function itself is now subject to an experimental investigation through adaptive laboratory evolution in a controlled setting.

Causation in biology, therefore, is in some respect an endless iterative process that seeks to find an optimal 'solution' for survival. Alterations in the genetic program with each iteration have the potential to induce changes in phenotypic functions. This recursive process takes place within the constraints imposed by physics and chemistry under given environmental conditions. Necessary ingredients to understand distal causation mechanistically are thus constraints and optimality.

Systems biology As described above, causation differs in physics and biology. However, both are relevant to systems biology and they represent opposite ends of a hierarchical process (Figure 15.3). Systems biology tries to bridge the two ends of

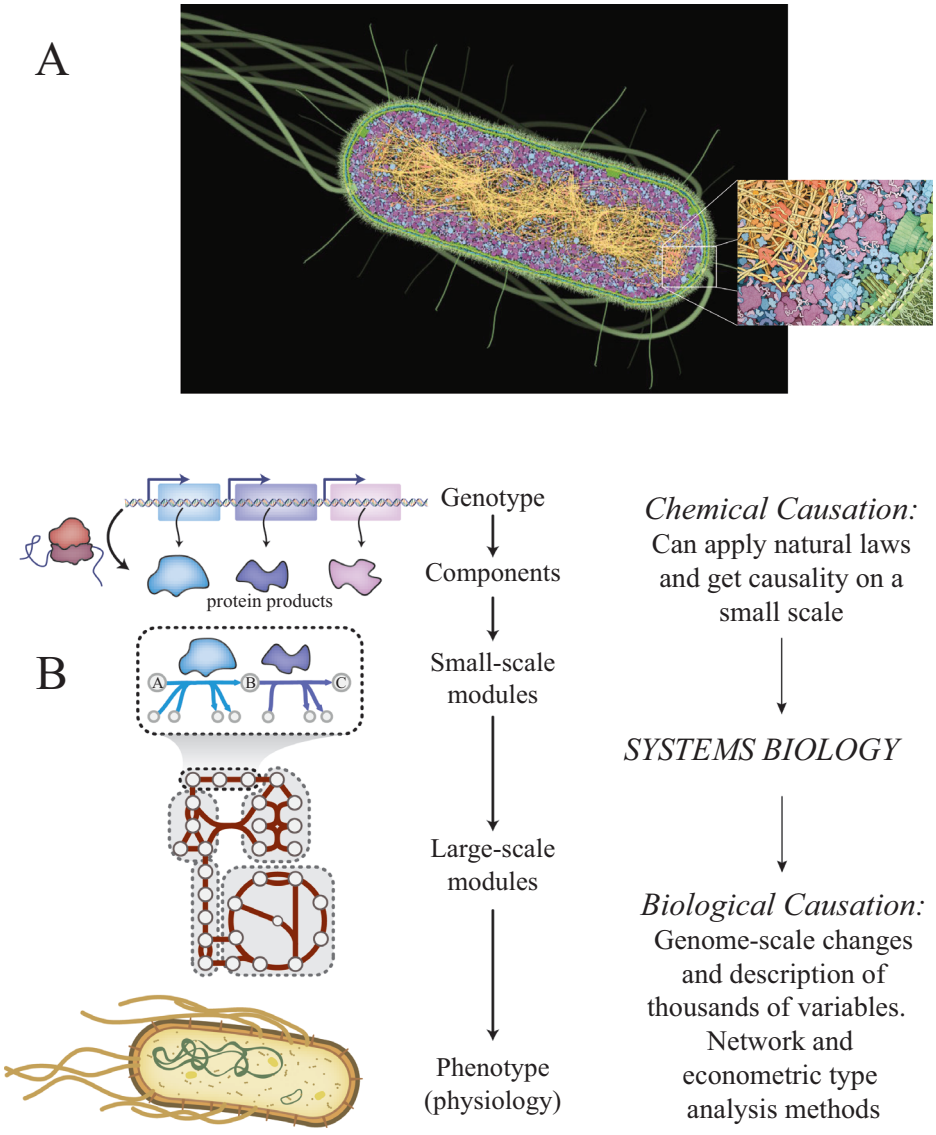

Figure 15.3 The hierarchical nature of living systems and multi-level causation. (A) Living systems are complex and hard to define precisely in biophysical terms. (B) Systems biology tries to provide structure to the hierarchical relationship between molecular and physiological events. Inserted image from [148] used with kind permission from Springer Science+Business Media B.V. Prepared by Nathan Lewis.

this spectrum and develop what amounts to a *quantitative and mechanistic genotype–phenotype relationship*.

The genotype–phenotype relationship has been the subject of argument and speculation. However, since the first genome sequence appeared in 1995, we have learned how to build such relationships with a mechanistic basis and have applied it

to metabolism and growth of microorganisms and metabolism in mammals. Genome sequences provide comprehensive, albeit not yet complete, information about the genetic elements that create the form and function of an organism. A constraint-based analysis provides a framework within which these basic considerations can be accommodated. Before we describe this framework, we will discuss the challenges with applying theory-based approaches to large-scale model-building in biology.

15.2 Building Quantitative Models

15.2.1 The physical sciences

Proximal causation *In silico* model-building in the physico-chemical sciences starts with basic principles such as thermodynamics, chemical potential, the diffusion equation, mass conservation, or the Nernst and Boltzman equations. These equations are based on well-developed fundamental physical theories, and they typically contain a large number of parameters, most of which can be measured individually under defined conditions. These parameters, such as the diffusivity of oxygen or the unit charge on the electron, are time-invariant. These equations then form the basis for computer models and simulation.

Limitations of theory-based modeling approaches in biology Traditional theory-based models of large-scale biological processes are faced with fundamental challenges.

First, the intracellular chemical environment is complex (e.g., see Figure 15.3A) and hard to define in terms needed for the formulation of equations that describe the physics of the intracellular milieu.

Second, assuming that we had all the governing equations defined, we would have to find numerical values for all the parameters that appear in these equations. These values would have to be accurate for intracellular conditions.

Third, even if we could overcome the first two challenges, we have to face the fact that evolution changes the numerical values of kinetic constants over time. In addition, in a biopopulation, even if we had a perfect *in silico* model for one individual organism it would not apply perfectly to other individuals in the biopopulation due to genetic and epigenetic differences between individuals. Such time-dependency and diversity of parameter values are key distinguishing features between biological and physico-chemical systems.

15.2.2 The life sciences

Distal causation: the selection process The process of evolution is fundamental to the biological sciences. Organisms exist in particular environments and, as they replicate, they produce offspring that are not genetically identical to the parent, thus generating a biopopulation of individuals that are each slightly different from one another (Figure 15.2). Over time, natural selection favors those individuals in the biopopulation that have more *fit* functions than other members of the biopopulation.

To survive in a given environment, organisms must satisfy myriad constraints, which limits the range of available phenotypes. The better an organism can achieve a relatively fit function in a given environment, the more likely it is to survive.

Constraining behaviors Because of dual causality, mathematical model-building in biology at the network and genome-scale will need to differ from that practiced in the physico-chemical sciences. The third limitation listed above is due to the dual causality that needs to be accounted for in realistic models of biological processes. An approach to the *in silico* analysis of cellular functions can be formulated based on the fact that cells are subject to governing constraints that limit their possible behaviors. Imposing these constraints can determine what functional states can and cannot be achieved by a reconstructed network. Imposing a series of successive constraints can limit allowable cellular behavior, but will never predict it precisely.

The imposition of constraints leads to the formulation of *solution spaces* rather than the computation of a *single solution* (or a discrete set of a few solutions), the hallmark of theory-based models. Cellular behaviors (i.e., functional states of networks) within the defined solution space can be attained, those outside cannot. Each allowable behavior basically represents a different candidate phenotype based on the component list, the biochemical properties of the components, their interconnectivity, and the imposed constraints. The constraint-based approach leads to *in silico* analysis procedures that are helpful in analyzing, interpreting, and even predicting the genotype–phenotype relationship.

Thinking about constraints Cells are subject to a variety of constraints. There are both *non-adjustable* (i.e., invariant or hard) and *adjustable* constraints (Table 15.1). The former can be used to bracket the range of possible phenotypic functions. The latter can be used to further limit allowable behavior, but these constraints can adjust through an evolutionary process or through changing environmental conditions. In addition, the adjustable constraints may vary slightly from one individual to another in a biopopulation. Together, these constraints define a range of possible functions, described mathematically as a solution space, and direct the realization of phenotypic expression.

This distinction between adjustable and non-adjustable constraints and their role in understanding dual causation is illustrated schematically in Figure 15.4. The feasible space of steady-state reaction fluxes is determined by non-adjustable constraints (the outer octagon). A narrower subset of functional states within this space is defined by the regulation of kinetic properties, or adjustable constraints. These adjustable constraints can be modified further through evolution to alter the limits of the subspace of expressed functional states. The direction of such an evolutionary expansion of the regulated subspace is driven by a need for improved performance, which can be described by an objective function (see Chapter 21).

15.2.3 Genome-scale models

Phenotypic functions are the results of the interactions of multiple gene products. In principle, all expressed gene products under a given condition affect the phenotypic

Table 15.1 Constraints on the functions of biochemical reaction networks. Adapted from [311].

Factor	Type of constraint
Physico-chemical constraints	
Osmotic pressure, electroneutrality, solvent capacity, membrane space, molecular diffusion, thermodynamics	Hard, non-adjustable constraints
Connectivity	
Systemic stoichiometry Causal relationships	Hard, non-adjustable constraints, but can be adjusted by horizontal gene transfer
Capacity	
Maximum/minimum flux	Non-adjustable maximum based on maximum association rates Adjustable by transcriptional regulation
Rates	
Mass action, enzyme kinetics, regulation	Highly adjustable by an evolutionary process

state of an organism. Thus, to develop mechanistic genotype–phenotype relationships a genome-scale model is needed. The availability of whole-genome sequences made the construction of such models possible.

Three generations of genome-scale models The successive application of constraints lends itself to a step-wise development of increasingly refined data-driven *in silico* models [313]. These models broaden in scope with the establishment and imposition of additional constraints. Constraint-based models can address questions relating to determining the possible functions of a network, which of these functions the cell actually chooses, and how such choices are made (Table 15.2).

The *first generation* of constraint-based models for microbial metabolism appeared at the turn of the century [105, 106]. The 'omics' data type on which they are based is *genomic*. Literature (*bibliomic*) data are used as well as the formulation of hard physico-chemical constraints, such as mass, energy and redox balance, thermodynamic, and maximal reaction rates. These constraints collectively define all the possible functional states of a reconstructed network. They are confined mathematically to a solution space. The properties of this space can be studied by the methods described in the following chapters.

The *second generation* of constraint-based models includes the imposition of transcriptional regulatory networks, leading to the shrinking of the allowable states of metabolic networks (see Figure 15.4). In response to environmental queues

Table 15.2 Generations of genome-scale models. From [313].

Generation	Usage	Type of data used
First	What states are possible	Genomic and bibliomic
Second	What states are chosen	Condition-dependent transcriptomic and genome location data
Third	How states are chosen	Time-dependent proteomic and metabolomic

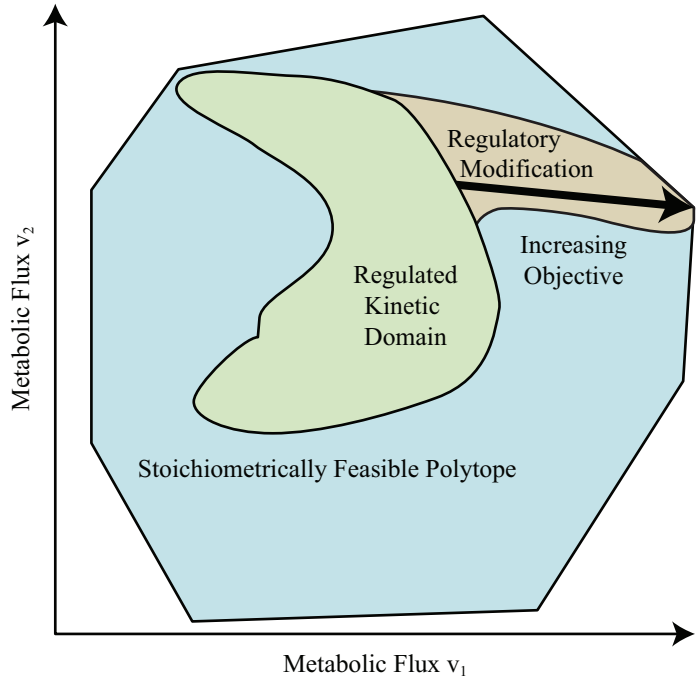

Figure 15.4 Early conceptualization of hard and adjustable constraints and how distal causation drives the choice of functional states within the hard constraints. Redrawn from [439].

and built-in regulation, the solution space is shrunk [85, 86] to contain network functions that the cell has chosen through an evolutionary process. The choices that a cell makes can then be identified and analyzed. Through the reconstruction of transcriptional regulatory networks, we can now begin to impose condition-dependent constraints, or restraints on reconstructed metabolic networks and formulate the second-generation models.

The *third generation* of constraint-based models will account for the abundance or concentration of the cellular components. Various 'omic' data types can now be obtained in a time-resolved fashion. Such data will help clarify just how

the cell implements the choices it has made and how it evolves to find new choices. This approach is likely to lead to the definition of the rate constants of the network as a whole rather than constants for the individual underlying biochemical events. Initial efforts in this direction are represented by the MASS modeling procedure [317].

Some properties of genome-scale models Biological networks have several fundamental properties that need to be considered when interpreting large-scale data sets and building models to describe their functions (see Figure 15.5).

Redundancy. Biochemical reaction systems have redundancy built into them at many levels. Often, individual steps in a network can be carried out in more than one way. Isozymes represent different enzymes that carry out the same reaction. Similarly, some codons can be translated by more than one tRNA. There are also network-level redundancies. The overall function of a network to support a phenotype can be achieved in more than one way. Thus, there are multiple equivalent outcomes from the same biological selection process. The mathematical aspect of this feature, *equivalent optimal solutions*, is detailed in Chapter 20. Biologically, these may be called *silent phenotypes*.

Multi-functionality. There are components in biochemical networks that can carry out more than one function. Examples include generalist of *promiscuous* enzymes that can catalyze many related chemical reactions. Similarly, some tRNA molecules can translate more than one codon. At the network level, there could be global network states that would give similar phenotypes even if the environments were different. This feature would be called a *generalist phenotype*. The notion of a high-flux backbone in metabolism [14] is composed of a set of reactions that lead to optimal growth on different substrates. A high-flux backbone is an example of a large correlated set of reactions that function together in optimal solutions [325].

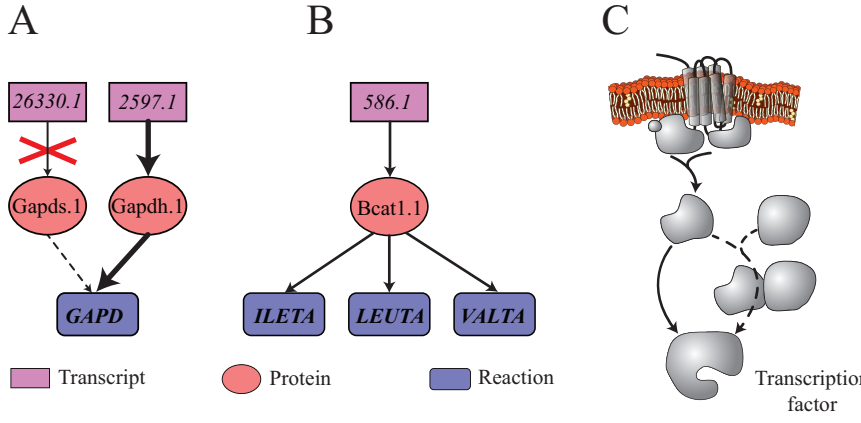

Figure 15.5 Some properties of genome-scale models. Panel A: redundancy; panel B: multi-functionality; panel C: non-causality. Prepared by Nathan Lewis.

Non-causality. Due to the hierarchical organization of organisms, changes on one level may not percolate up to functions at a higher-level of organization, and would thus be non-causal. A well-known example of non-causality are *hitch-hiker mutations* that co-select with a causal mutation located nearby on the genome. In the field of signal transduction, many are interested in knowing 'who-talks-to-whom,' meaning that one wants to know all possible chemical interactions between two components. Protein–protein interaction maps provide one example. In this case, however, the biologically meaningful question is 'who-*listens*-to-whom,' as we are only interested in knowing if chemical interactions are a part of a higher-order biological function. Thus, there can be many non-causal (biologically), but detectable, chemical interactions among macromolecules.

These three attributes are important considerations in studying the hierarchical nature of biological systems. Multi-scale, multi-parameter analysis methods will be needed to study this hierarchical organization. They will need to be able to deal with non-regular patterns, which will be a deviation from classical methods such as Fourier analysis that looks for repeated regular patterns. All of these features have appeared through the evolutionary process which abides by a series of constraints.

Higher-order properties There are some other notable higher-order properties of biological networks, which will not be detailed here. Such properties include *self-assembly* of components to form a functioning network spontaneously, the *selection* that seems to be at work during both distal and proximal causation, the notion of *a self* in biology (namely, is a component a part of a network, or not?), and the notion of *awareness* (that ultimately is related to the mathematical concept of a functional state of a network). It is an interesting and fundamental challenge to the field to determine if such important biological properties can be defined mathematically in the context of genome-scale models. If possible, molecular systems biology will advance notably.

15.3 Constraints in Biology

All expressed phenotypes resulting from the selection process must satisfy the governing constraints. Therefore, clear identification and statement of constraints to define ranges of allowable phenotypic states provides a fundamental approach to understanding biological systems that is consistent with our understanding of the way in which organisms operate and evolve.

Different types of constraints limit cellular functions and several authors have discussed general constraints in biology [84,93,139,179,261]. Here we start this discussion by dividing constraints into four categories [344]: (1) fundamental physico-chemical, (2) spatial or topological, (3) condition-dependent environmental, and (4) regulatory, or self-imposed constraints.

Physico-chemical constraints Many physico-chemical constraints govern cellular processes. These constraints are inviolable and thus represent *hard* constraints.

Conservation of mass, elements, energy, and momentum represent hard constraints. The interior of a cell is densely packed, forming an environment where the viscosity may be on the order of 100–1000 times that of water. Diffusion rates inside a cell may be slow, especially for macro-molecules. The confinement of a large number of molecules within a semi-permeable membrane causes high osmolarity. Thus, cells require mechanisms for dealing with the osmotic pressure generated, such as sodium–potassium pumps to balance osmolarity or a rigid cell wall to physically withstand it. Intracellular reaction rates are determined by local concentrations inside cells. Reactions have maximal reaction rates (denoted with $v_{\max}$) estimated to be about a million molecules per μm^3 per second (see Equation (17.6)). Furthermore, biochemical reactions need to have a negative free energy charge in order to proceed in the forward direction. These are some of the many basic physico-chemical constraints under which cells must operate.

Spatial constraints The crowding of molecules inside cells leads to *spatial*, or three-dimensional, constraints. The linear dimension of the bacterial genome is on the order of 1000 times that of the length of the cell. DNA must therefore be tightly packed in the nuclear region in an accessible and functional configuration because DNA is only functional if it is accessible. Thus, at least two competing needs (to be tightly packed, yet accessible) constrain the physical arrangement of the bacterial genome. DNA in eukaryotes is organized in a highly hierarchical fashion.

As a further example, we note that the ratio between the total number of tRNA molecules and the number of ribosomes in a typical *E. coli* cell is approximately 10 to 1 [282]. With 43 different types of tRNA, there is less than one full set of tRNAs per ribosome. The genome, therefore, may have to be configured such that the location of rare codons is spatially close and translated by the same ribosome. Protein localization and crowding of space in membranes represent additional topological constraints.

Identification of these constraints and analysis of their consequences will be important for the understanding of the three-dimensional organization of cells. This challenge is hard, but progress is being made [477].

Environmental constraints Environmental constraints on cells, such as nutrient availability, pH, temperature, osmolarity, etc., are typically time- and condition-dependent. For example, *H. pylori*, a human gastric pathogen, lives in a relatively constant environment, but is constrained by its low pH. It produces ammonia to sufficiently neutralize the pH in its immediate surroundings in order to stay alive. The growth of a plant is limited by the flux of incident photons as well as nitrogen and phosphorous availability in the soil.

Conversely, the life cycle of *E. coli* is characterized by a series of sudden environmental changes. Outside of an animal it lives at ambient temperature and in the presence of ample oxygen. Then it experiences a heat shock when it enters the mouth of an animal, followed by an acid shock when it reaches the stomach. Following entry into the small intestine, another pH shock is experienced, followed by a nutritionally rich anaerobic environment where it can grow rapidly in the presence of other bacterial species. Then, finally, it experiences a cold shock and ample

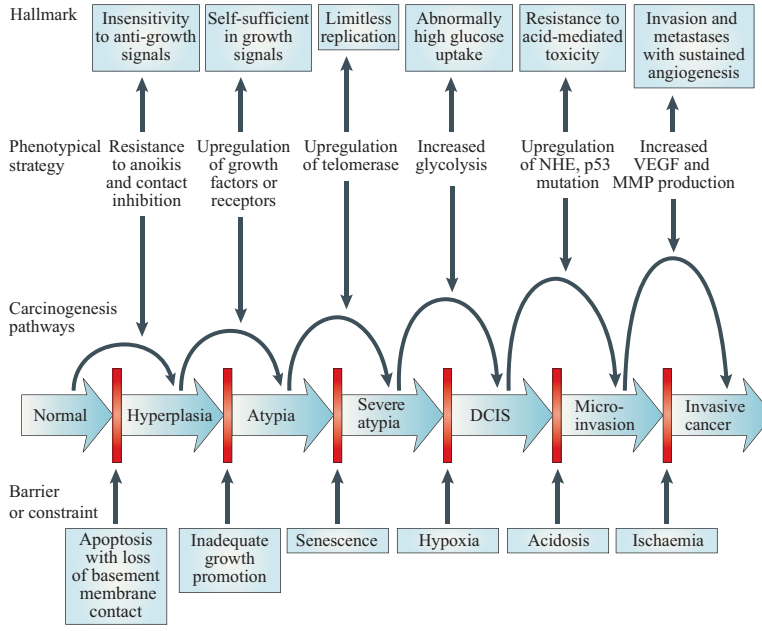

Figure 15.6 The carcinogenic process conceptualized as a series of losses of constraints. Taken from [136].

oxygen with diminishing nutrients surrounding it upon excretion. *E. coli* needs to be able to adjust its internal functional state to survive this series of environmental changes.

Knowing the environmental constraints is of fundamental importance for the quantitative analysis of microorganism functions; however, natural environments may be hard to define precisely. Conversely, in the laboratory, defined growth media can be used so that the environmental variables are known precisely.

Regulatory constraints These constraints are fundamentally different than the three types discussed above. They are *self-imposed*, are subject to evolutionary change, and can thus be time-variant. As illustrated in Figure 15.4, they work within the outer constraints defined by physico-chemical processes. For this reason, these constraints may be thought of as regulatory *restraints*, in contrast to the physico-chemical constraints, the spatial constraints, and environmental constraints. Based on environmental conditions, regulatory constraints provide a mechanism to eliminate suboptimal phenotypic states and confine cellular functions to behaviors of high fitness. Regulatory constraints are produced in a variety of ways.

The loss of ability to impose constraints through regulation to maintain a certain phenotype would lead to a loss of the desired biological function. The carcinogenic process provides a serious example (Figure 15.6). This process can be understood as a series of losses of constraints that leads to a malignant phenotype and ultimately a metastatic state.

15.4 Summary

- Systems biology bridges multiple scales in biology.
- Dual causation in biology requires us to accommodate the physico-chemical constraints under which cells operate, as well as the fundamental biological processes of natural selection and generation of alternatives when building models in systems biology.
- Organisms have to abide by a series of constraints, including those arising from basic physical laws, spatial constraints, and the environment in which they operate.
- Many possible biological functions are achievable under these constraints, and organisms willfully impose constraints through various regulatory mechanisms to select useful functional states from all allowable states.
- A constraint-based approach that enables the simultaneous analysis of physico-chemical factors and biological properties emerges from these considerations.