ELSEVIER

Host–pathogen systems biology

Christian V. Forst

Bioscience Division, Los Alamos National Laboratory, Mailstop M888, P.O. Box 1663, Los Alamos, NM 87545, USA

Unlike traditional biological research that focuses on a small set of components, systems biology studies the complex interactions between a large number of genes, proteins and other elements of biological networks and systems. Host–pathogen systems biology examines the interactions between the components of two distinct organisms, either a microbial or viral pathogen and its animal host or two different microbial species in a community. With the availability of complete genomic sequences of various hosts and pathogens, together with breakthroughs in proteomics, metabolomics and other experimental areas, the investigation of host–pathogen systems on a multitude of levels of detail has come within reach.

Systems biology is a novel approach to studying, analyzing and – ultimately – controlling biological systems. Unlike traditional research that typically focuses on single genes, systems biology studies complex interactions of all levels of biological information. In light of emerging biological threats, such as the anthrax scare in Florida, USA, in 2001, the re-engineering of smallpox from old sequence data, the emergence of multi-drug resistant pathogens (e.g. *Mycobacterium tuberculosis* and strains of *Staphylococcus*) or the lurking danger of a new influenza pandemic of the avian H5N1 strain, one wants to take a step further and expand biological systems to include two organisms – a pathogen and a host.

The research of host–pathogen interactions in its broadest definition is a very mature field. It is closely linked to the understanding of the immune system and immune responses [1] (Box 1). Host–pathogen interactions can be interpreted as the battle of two systems. For example, pathogens can hijack host cells and use host cell capabilities to the pathogens' advantage [2], or they can evolve so rapidly that their sheer diversity overwhelms the immune system, as in the case of HIV infections [3].

The detailed mechanistic analysis of host–pathogen systems, encompassing all aspects of such a multilevel problem – from molecular interactions to organism responses – is still in its infancy. Components and subproblems have been addressed by theoretical and experimental approaches, often focusing on either host-response or pathogen interference by mimicking the missing 'partner'. Host–pathogen systems biology is the name given to the paradigm of integrating these different types of models into a host–pathogen system over a range of detail of descriptions. The ultimate goal for host–pathogen systems biology is not only the discovery and comprehension of underlying biology, but also the establishment of a robust framework for more efficient drug development and therapeutic intervention. Examples, approaches and perspectives of host–pathogen systems biology are given in this review.

Systems biology in drug discovery

As recent reviews indicate [4,5], systems biology approaches offer novel strategies to shorten the cumbersome path from identified target to an approved drug. Systems biology provides *in silico* models for cost-effective decision making during multimillion-dollar drug development programs [6].

The term 'systems biology' encompasses many different techniques and models for probing and understanding biological complexity, spanning multiple levels of spatial and temporal scales (Figure 1). Because biological complexity is an exponential function of the number of systems components and the interactions between them, such efforts are currently limited to simple organisms or to specific pathways in higher organisms. Limiting systems biology studies to specific functional subsystems is even more pronounced in host–pathogen system biology, which focuses on more than one organism. Where systems biology is applied to drug discovery, three

Corresponding author: Forst C.V. (chris@lanl.gov).

1359-6446/06/$ – see front matter PII: S1359-6446(05)03735-9

BOX 1

An overview of the immune system

Immune response is an essential defense mechanism against pathogens that is available to most multicellular organisms. Even unicellular organisms, such as the well-known mould *Penicillium chrysogenum*, produce chemical components to kill pathogens. Chemical agents are, among other defense mechanisms, part of an innate immune response. The following list details the defense mechanisms repertoire of the innate and adaptive immune system:

- Innate immune response (plants, animals);
- Barriers to pathogen entry;
- Mechanical responses to eliminate antigens;
- Chemical agents;
- Phagocytes;
- Fever – elevated temperature inhibits growth of microbes;
- Inflammatory responses to attract white blood cells (leukocytes) to the infection site;
- Natural killer (NK) cells to kill pathogen-infected and cancer cells;
- Adaptive immune response (higher animals);
- Synthesis of antibodies to bind antigens and promote their elimination;
- T-cell killing of virus-infected cells;
- Activation of macrophages to destroy phagocytosed pathogens.

The innate and adaptive immune responses are complementary components of multicellular host defense. The innate immune response provides the initial defense against infections with responses occurring within hours after infection. By contrast, the adaptive immune response requires several days to develop after infection. Innate immunity relies on germline-encoded receptors and is limited to some extent in its diversity, although some diversification is achieved by heterodimerization of TLRs or the semi-invariant NKT cells. NKT cells – T cells with the properties of NK cells – blur the distinction between innate and adaptive immunity by using the complex machinery of somatic recombination to produce receptors recognizing a narrow range of antigenic diversity. Conversely, the receptors of the adaptive response that are also produced by somatic recombination of gene segments experience a tremendous diversity. The adaptive immune system also produces memory cells to store receptor information for particular responses.

The innate immune response and toll-like receptor pathways

The innate immune system is essential for host defense and is responsible for early detection and containment of pathogens. The inflammatory response to pathogens is activated when the phagocyte recognizes the foreign invader using a battery of pattern recognition receptors (PRR), including toll-like receptors (TLRs) [53], members of the C-type lectin receptor family [54], scavenger receptors [55], complement receptors [56] and integrins. Conserved pathogen-specific chemical motifs recognized by these receptors include carbohydrates, glycolipids, glycoproteins, nucleic acids (DNA and double-stranded RNA), proteolipids and proteins. Stimulation of PRRs results in activation of a broad spectrum of interacting signaling pathways, revealing a system of extraordinary complexity. Additional receptors, such as cytokine, chemokine or growth factor receptors, add to the specificity of the immune response.

principal approaches can be identified [6]: (i) bioinformatic integration of 'omics' data (a bottom-up approach); (ii) integrative mathematical cell models (an intermediate approach); and (iii) computer models of disease or organ system physiology from cell and organ response information available in the literature (a top-down approach to target selection, clinical indication and clinical trial design). These complementary approaches must ultimately be integrated in the pursuit for a hierarchical molecule-to-systems-level understanding of host–pathogen interactions.

Computational systems biology models, methods and tools

Scales and models

The goal of host–pathogen systems biology is to understand physiology and infectious disease at the level of molecules, cellular networks (e.g. metabolic, regulatory and signaling networks), cells (host cells as well as various viruses and bacterial pathogens), tissues, organs and ultimately whole organisms. A comprehensive systems model can span approximately ten orders of magnitudes in scale, and even more in time (Figure 1). Two distinct strategies for modeling along many levels of description can be recognized, a bottom-up and a top-down approach which are integrated in a third, hybrid strategy.

1. Bottom-up approach

It could be argued that a full understanding of a host–pathogen system requires knowledge of all of its components. A bottom-up approach focuses on the measurement and description of complex systems using the building blocks – their interactions and dynamic properties, such as kinetic parameters. In molecular biology, bottom-up modeling started after the genomic revolution with a plethora of 'omic' information available. The bottom-up approach can be used to investigate which genes, proteins or phosphorylation states of proteins are expressed or upregulated in an infection process, leading to testable hypotheses that the regulated species are important to disease induction or progression. By integrating of genomic, proteomic and metabolomic data, models have been developed that mechanistically describe intra- and inter-cellular processes (e.g. during drug response or disease progression).

2. Top-down approach

The top-down modeling approach attempts to develop integrative and predictive multiscale models of biological processes. A long-term goal would be an *in silico* model of human–pathogen physiology and infection. However, with the current technology, such modeling focuses on relatively specific problems at particular scales, for example, at the pathway, immune cell system or organ level.

3. Hybrid models

Bottom-up models serve as scaffolds for top-down models by providing information of possible and potential interactions and subprocesses, how these subprocesses respond to drugs and infection and how matter and information is passed between subprocesses and through different scales. Such hybrid approaches benefit from bottom-up molecular measurements and knowledge as well as top-down predictive modeling. A 'postgenomic physiology' could span many different levels of biology, from molecules to whole organisms, moving away from 'naïve reductionism' towards a discipline that fosters integration and synthesis, as Strange [7] envisioned in a recent review.

Methods

Complementary to the biological hierarchy of host–pathogen systems, methodological descriptions and simulations of such systems have been performed on different level of detail. Interaction networks and network models of biological systems have been studied at the level of: (i) topological connections; (ii) qualitative connections; (iii) quantitative connections; and (iv) higher order interactions, as reviewed by Bower and Bolouri [8].

1. Topological connections

Networks are assembled from interaction data, physical measurements or computational predictions of protein–protein, protein–DNA,

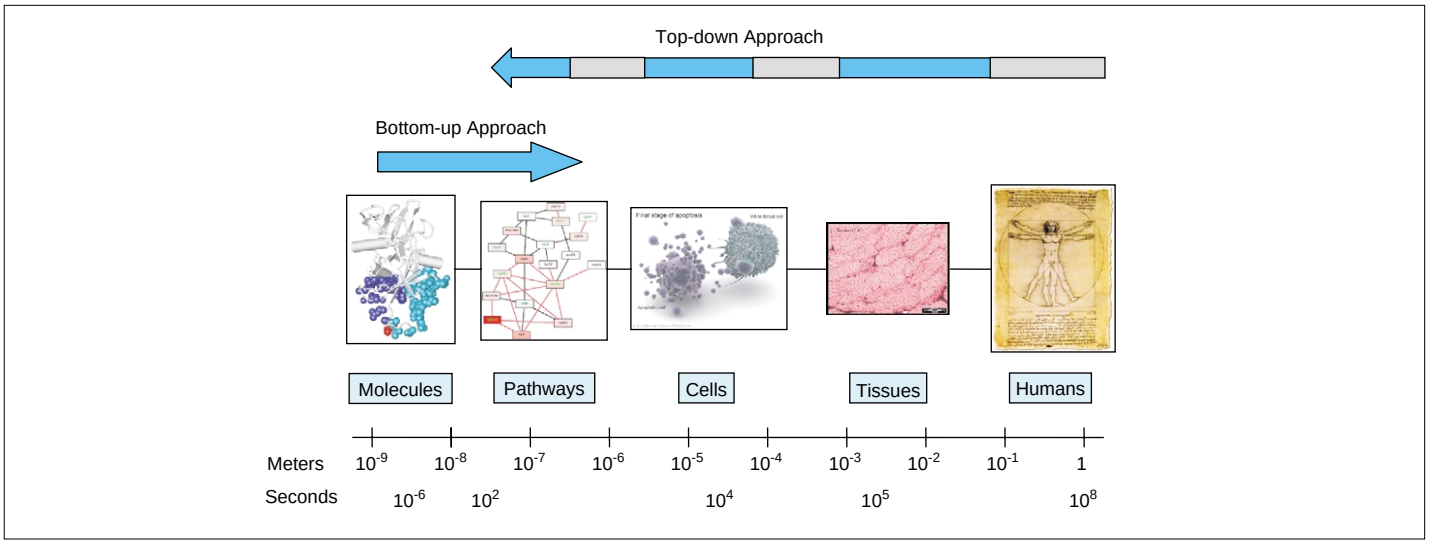

FIGURE 1

Multiscale approaches in biological systems modeling, from molecules, pathways, cells and cell-cell interactions and tissues to the whole organism. Life takes place on many different temporal and spatial scales. The spatial scale ranges in orders of magnitudes from nanometers, for chemicals and proteins, to meters for the whole body. The temporal scale covers fast biochemical reactions happening in microseconds to the lifespan of an organism in years. The bottom-up approach ('omics') focuses on large scale identification of molecular components. The top-down approach (modeling) attempts to form integrative (multilevel) models of human physiology and pathogen infection, which typically focus on relatively specific questions at particular scales, due to the limitations of current technologies. Adapted, with permission, from [6].

protein–small-chemical or other identified interactions; or by inference of correlations between cellular components. Undirected edges indicate interactions between components. Networks of this kind are often referred to as interaction maps or networks.

Network biology [9,10] is a method that studies inter- and intracellular networks and their genomic, proteomic and metabolomic foundations. Network biology forms the basis of systems biology by providing information on biological components, their interactions and their functional interplay in biological networks. One particular aspect of network biology focuses on the graph structure of the underlying interaction map by providing quantifiable measures, such as node-degree distribution, mean path length and clustering coefficients, as well as by identifying architectural features, such as the existence of motifs and modules and their hierarchical structure [10]. Such measures are particularly interesting when linked to phenotypic properties of the biological system, such as system survival. Jeong *et al.* [11] have shown that proteins essential for survival are highly connected in a yeast protein-interaction network.

2. Qualitative connections

At the level of qualitative connections, directionality and causality indicate how input nodes affect the output nodes Directional edges indicate causal relationship as well as qualitative interactions (e.g. activating or inhibitory interactions). Qualitative models include metabolic flux models that assume steady state conditions [12,13], models that consider gene regulation [14] or Boolean network models [15,16].

3. Quantitative connections

At the level of quantitative connections, quantitative functions are assigned sets of interactions that describe the dynamic co-dependence of the dynamic behavior of an output, depending on the dynamic behavior of inputs. Methods of choice include power law models, such as S-systems [17,18], reaction kinetics modeled by ordinary differential equations [19,20] or stochastic simulations [21,22].

4. Higher order interactions

At the level of higher order interactions and reaction rules, higher level nodes and connections represent abstract concepts that can be expanded into hierarchical sublevel subgraphs based on reaction rules. Examples include signaling networks and metabolic reactions, with context dependent or rule-based interactions and different types of nodes [23].

Other methods

Response networks

Response networks were first mentioned in connection with biomolecular networks by Magasanik [24]. Groundwork for a systematic theoretical analysis of response networks has been laid by Zien *et al.* [25] and has been further developed independently by Ideker *et al.* [26]. The idea behind response network analysis is the analysis of experimental data, such as expression profiles, in the context of biological networks. By a superposition of experimental data with network information, networks are identified that best represent the system response according to the experimental conditions tested.

Comparative network analysis

In principle, graph comparison is an NP (non-deterministic polynomial-time) hard problem, which typically can only be addressed by exhaustive enumeration techniques. However, methods for comparative network analysis of biological systems have been developed in the past. Such methods have proven to be powerful in several applications including metabolic [27–30] and protein interaction networks [31], as well as correlation of protein interaction networks with gene expression [32]. Recently, a method has been developed to correlate and compare response networks for identification of common and specific responses [33].

Tools

A plethora of tools and software for biological systems modeling have been developed and are available for download, often free for academic users. Since its development by Hucka and co-workers [34], many modeling tools use the systems biology markup language (SBML; http://sbml.org) for portable model description. Owing to the number of tools available, the potential user is referred to the systems biology website by 'Kitano's Symbiotic Systems Project' (www.systems-biology.org) that lists model editors for graphic-assisted model construction, simulation tools for deterministic and stochastic simulations, analysis tools and utilities. Physiology modeling software is not yet well-integrated with molecular and cellular modeling tools. A list of physiological modeling groups and tools can be found at the Federation of American Scientists website (www.fas.org).

Network Models

Host–pathogen system models that fall in the 'omics' category comprise interaction maps or interaction networks that show components of a network and their interactions for further analysis .

Genomic foundation of host–pathogen interactions: a Chlamydia psittaci*–Human metabolic interaction network*

One example of a true host-pathogen network model is the tryptophan (trp) biosynthesis network of a class of obligate intracellular pathogens, *Chlamydiae* (Figure 1a) [9].

Chlamydia primarily infects mucosal epithelial cells with consecutive infection of subepithelial tissue [35]. *Chlamydia* infections progress although a lifecycle of three distinct stages. The host is invaded by elementary bodies (EBs), which represent the extracellular infectious stage. After infection, EBs develop into intracellular reticular bodies (RBs), which replicate and further mature into EBs, which then lyse the host cell and initiate another round of pathogen infection. The cycle between EBs, RBs and lysis of host cells characterizes the acute disease state. A third state of development, the persistence state, describes the chronic disease progress. In tissue culture, the persistence state of *Chlamydia* can be induced by various factors, specifically by introducing interferon-γ (IFN-γ), nutrient limitations or other environmental stress. For example, it is well documented that tryptophan levels in host cells decrease as an effect of IFN-γ [36]. It has also been recognized that tryptophan depletion might have a role in the development of chronic disease conditions [37].

Investigating the tryptophan biosynthesis pathway in a particular *Chlamydia* species, *Chlamydia psittaci*, is interesting because it shows the interdependence and connectivity between pathogen and host and thus helps to explain the development of the chronic disease. This pathway assembles an almost complete biosynthetic unit lacking the first step (the conversion of chorismate into anthranilate). Interestingly, genes encoding the enzymatic subunits *trpAα* and *trpAβ* that are typically present in tryptophan operons and that are responsible for catalyzing the conversion of chorismate to anthranilate are absent in the *C. psittaci* tryptophan operon and are not encoded elsewhere in the genome. Instead, the *C. psittaci* tryptophan operon includes two genes *kynU* and *kprS*, both of which are atypical components of the classic tryptophan operon (Figure 2b). This can only be understood through systematic metabolic network analysis. *KynU* encodes kynureninase, an enzyme that converts kynurenine into anthranilate. K*prS* codes for 5-phospho-D-ribosyl-1-pyrophosphate (PRPP) synthetase, a component needed in the first steps of tryptophan biosynthesis (Figure 2a). The complete tryptophan network, including the tryptophan-salvage pathway of the host, is shown in Figure 2a. For tryptophan biosynthesis, *C. psittaci* obtains an alternative source of anthranilate by hijacking the host's tryptophan depletion pathway by intercepting the byproduct kynurenine. At first, the tryptophan depletion pathway of the host is activated by inducing indoleamine-2,3-dioxygenase using IFN-γ (reaction with EC number* 1.13.11.11 in Figure 2a). Then, *C. psittaci* uses host kynurenine through kynU to produce its own tryptophan, enabling intracellular growth and causing chronic infections. Knowledge of such a metabolic host–pathogen system enables accelerated drug development of successful antibiotics against chronic *Chlamydia* infection.

Dynamic models and simulations

Immune–receptor signaling

On the host side, mathematical and computational network models have been used to study the process of signaling through receptors of the immune system [38]. Mathematical dynamic models are essentially used on two distinct levels of description: 'simple models' and 'detailed models' (see earlier section titled 'Computational systems biology models, methods and tools').

Simple models attempt to capture some features of the signal transduction system but make no attempt at mechanistically describing the signaling cascades that are activated. In these models, the actual signaling cascade is replaced by one or more arbitrary transitions. These models are simple in the sense that they have few components and a simple mathematical description. This implies that such models are also intrinsically more abstract than detailed models – their components and parameters often do not correspond directly to well-defined physical quantities, such as measured binding constants or chemical reaction rates. Nonetheless, simple models can provide insights into the behavior of a system and drive experimental and detailed modeling efforts by suggesting further more-detailed experiments or models.

Simple models capture two basic aspects of immune-receptor signaling: serial engagement and serial triggering [39,40], and kinetic proofreading [41]. Serial engagement [42] and kinetic proofreading models [41] associate the measured properties of ligand–receptor interactions with the amplitude of signaling responses, but these models do not describe molecular interactions beyond ligand–receptor binding.

Building detailed models of cell signaling cascades involves the selection of a limited set of protein components that participate in signaling. Goldstein *et al.* [38] propose a three-part protocol for defining such a mathematical model:

1. Selection of a set of components and identification of their interactions based on what is known about the system.
2. Selection of parameters that quantify the cellular concentrations of the components and the strength of the interactions between components (known as rate constants).
3. Selection of a mathematical formulation of a method of simulation.

* The classification system for enzymes and biochemical reactions by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB, www.chem.qmul.ac.uk/iubmb/enzyme).

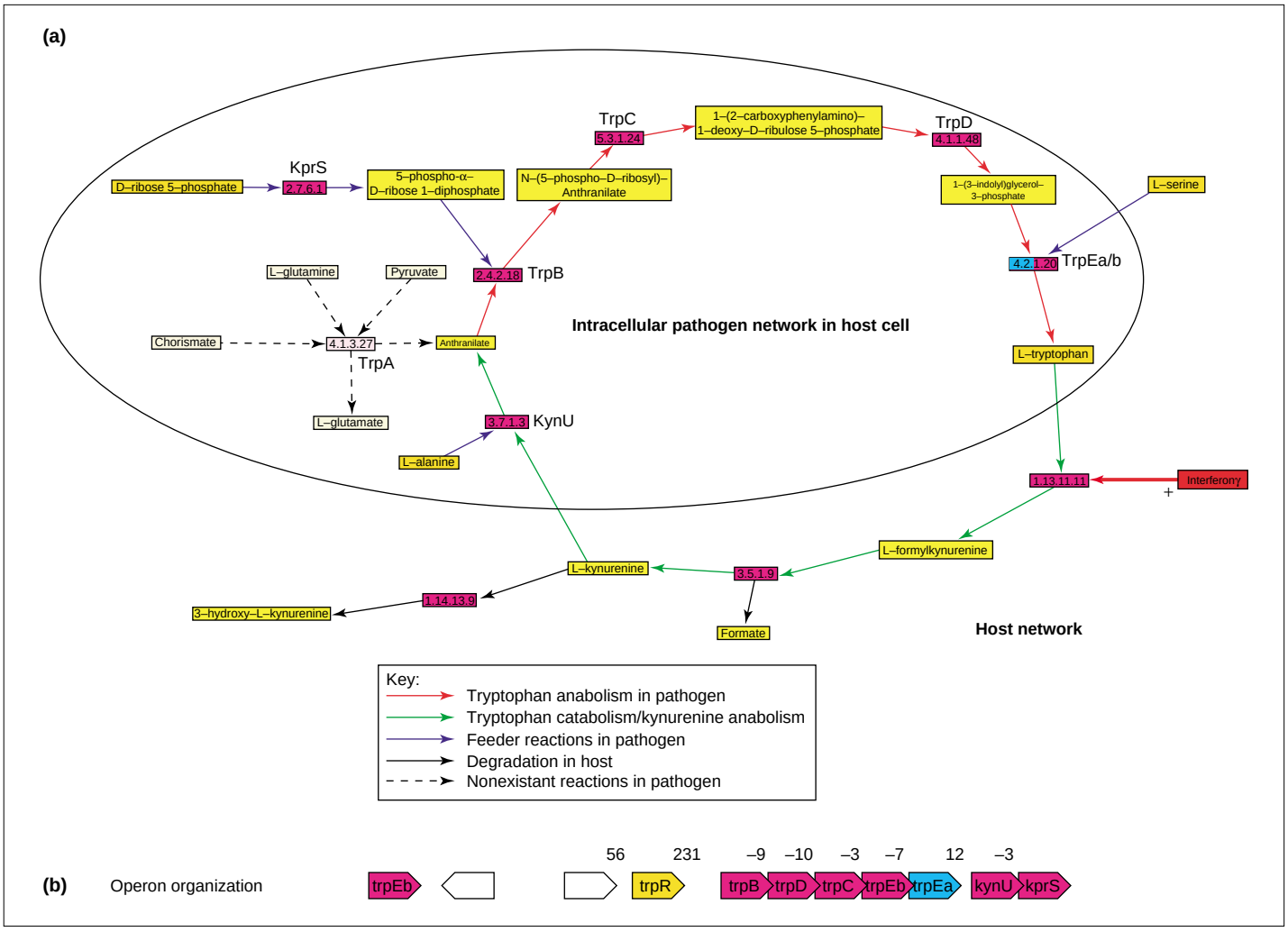

FIGURE 2

***Chlamydia psittaci*–host trp-metabolic network. (a)** A shared trp metabolic network between the pathogen *Chlamydia psittaci* and a human host. The subnetwork in the oval denotes the pathogen part of the network. Thin red arrows indicate anabolic trp reactions in the pathogen, green arrows are catabolic reactions, blue arrows refer to feeder reaction that provide required compounds [L-alanine (ala), ser and 5-phospho-D-ribosyl-1-pyrophosphate (PRPP)] for trp biosynthesis, and solid black arrows denote degradation reactions in the host. The dashed black arrows are reactions not found in the pathogen. The bold red arrow describes the activation of indoleamine-2,3-dioxygenase (EC 1.13.11.11) by IFN-γ. Reproduced, with permission, from [9]. **(b)** The operon organization of the trp operon in *C. psittaci* is shown. The presence of *kprS* and *kynU* is unusual in typical trp operons. Numbers above the genes refer to intergenic distances.

Reaction-network models are based on the assumption that each species is uniformly distributed throughout the cell. Reaction-diffusion models allow for the variation of species concentrations in different cellular compartments.

Using the FcεRI receptor (the high-affinity IgE receptor) as an example, Faeder *et al.* [23] have developed a detailed signaling model that takes into account downstream components affecting the signaling cascade (Figure 3). Figure 3a shows the four components of the receptor, the ligand (IgE dimer), the receptor (FcεRI) and the two kinases Lyn and Syk. The nine basic interactions are shown in Figure 3b, which include association and dissociation, transphosphorylation (i.e. catalysis of phosphorylation) and dephosphorylation. A surprising aspect of this model is that, because of combinatorial complexity, four components and nine interactions expand to a signaling network with 354 species and 3680 reactions (one particular reaction species is depicted in Figure 3c). The simulation results of the FcεRI signaling model show complex behavior of phosphorylation profiles as a function of the ligand–receptor off-rate, because the balance between kinetic proofreading and serial engagement changes, moving down the signaling cascade. Because serial engagement increases with the off-rate, an increase in phosphorylation with off-rate indicates that serial engagement is the dominant effect, whereas a decrease indicates that kinetic proofreading is dominant. Depending on the timely occurrence of a particular phosphorylation event, either kinetic proofreading or serial engagement is dominant. For example, the phosphorylation profile of γ immunoreceptor tyrosine-based activation motif (γ-ITAM) passes through a maximum, which indicates a transition between control by kinetic proofreading and control by serial engagement. Thus, the detailed signaling model shows that kinetic proofreading and serial engagement are emergent properties and the interplay of these mechanisms gives rise to an optimal off rate at which the highest response is achieved.

The ultimate goal of immune-receptor signaling models is to understand how the components of a signaling cascade work together to direct cellular responses to changes in the extracellular

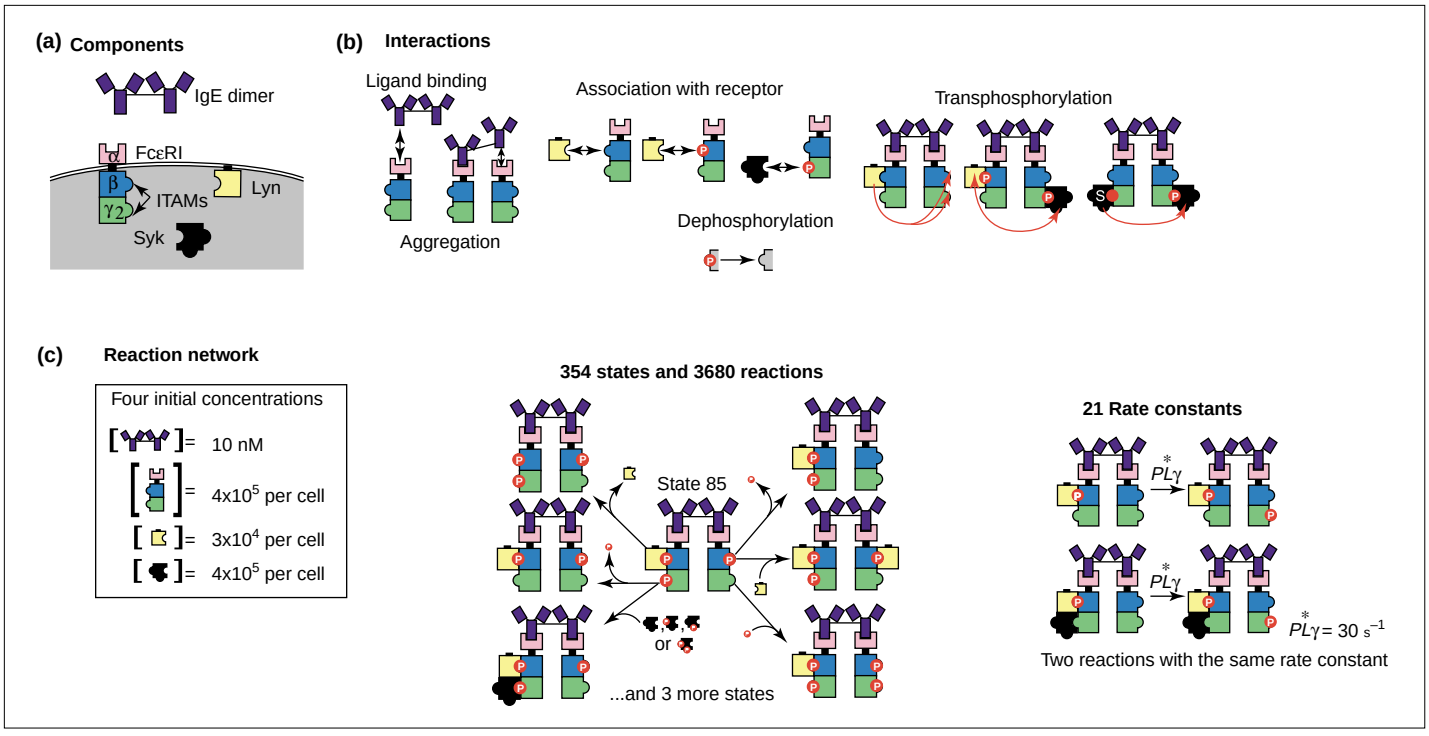

FIGURE 3

A detailed model of early events in FcεRI signaling. (a) The four components in the model are the IgE dimer, the receptor (FcεRI) and the kinases Lyn and Syk. The two cytosolic domains of the receptor each contain an immunoreceptor tyrosine-based activation motif (ITAM). **(b)** There are nine basic interactions, five for association–dissociation between signaling components, three transphosphorylation reactions, and one for spontaneous dephosphorylation of phosphorylated sites. **(c)** All possible combinations between components and basic interactions yields 354 complexes and phosphorylation states, each of which is tracked as separate species. The species are connected by 3680 reactions assembling a large biological network that is defined by a small number of parameters (the initial conditions of 4 proteins and 21 rate constants). One typical species is illustrated along with nine different reactions, of which six are explicitly shown. Reactions seven to nine are generated by using different phosphorylation states of Syk (gray square) to form additional states from the complex in the center to the bottom-left complex (indicated by '… and 3 more states'). The states are connected by a large biochemical reaction network (composed of 3680 reactions). A small number of parameters (the initial concentrations of the 4 proteins and 21 rate constants) define this network because the same rate constant can be used for many similar reactions. The figure shows two of the 24 reaction in which Lyn transphosphorylates γ-ITAM. The $p^*_{L\gamma}$ indicate the reaction rate of these two reactions. Reproduced, with permission, from [38]

environment. Simple and detailed mathematical models have contributed to the understanding of essential host–pathogen signaling events through immune receptors by identifying the fundamental mechanisms that are involved in determining, regulating and therapeutically modifying immune responses.

Immune system modeling

Complementing the molecular biology modeling approach described earlier, the following models capture the dynamics of the immune response to infectious pathogens at the cellular level in host–pathogen systems biology. A comprehensive review on mathematic modeling techniques of such systems has been presented by Perelson and Weisbuch [43].

The basic idea of viral infection models is simple and lead to the development of viral dynamics as a research field [44]. These infection models consist of three types of cells, target cells (T), infected cells (I), and virus particles (virions, V). Infected cells produce new virus particles at a constant rate p and die at rate δ. Virions are cleared by the immune system at rate c. The rate at which a target cell is infected is k.

$$dT/dt = \lambda - dT - kVT \qquad \text{[Eqn 1]}$$

$$dT/dt = kVT - \delta I \qquad \text{[Eqn 2]}$$

$$dT/dt = pI - cV \qquad \text{[Eqn 3]}$$

These equations (Equations 1, 2 and 3), developed by Perelson *et al.* [45], describe the basic model of viral dynamics and have been used to study primary HIV infection. More-complex models include specific components of the immune systems, such as the cytotoxic T lymphocytes [46], other nontoxic lymphocytes and cytokines–chemokines [47]. These models essentially include specific expressions for resting, active, memory and cytotoxic T cells.

Mathematical models of HIV infections have also proved to be useful in exploring the response of HIV to antiviral therapy. Specifically, the evolution and evasion by HIV under selection pressure from the immune system and drug treatments have been extensively studied. A review by Frost [3] discusses the benefit of evolutionary dynamic HIV models for the understanding of HIV response to highly active antiretroviral therapy. Frost specifically discusses the role of such models in the design and analysis of structured treatment-interruption studies in reducing drug toxicity, boosting HIV-specific immune responses and to allowing reversion of drug-resistance mutation in highly drug-experienced patients.

Building on such pharmacokinetic models of HIV evolution in human hosts, Dixit and Perelson [48] developed a hybrid model of HIV dynamics under antiretroviral therapy that combines

pharmacokinetics and intracellular delay (the time required for an infected cell to replicate a virus). This model helps to accurately determine the pharmacological delay and the time-dependent efficacy of drug action.

Models of host–pathogen physiology

A feasible approach to address the computational issues of integrating molecular, cellular and organ levels in a top-down approach is to put in place an organ-level framework and add increasing complexity in a modular format. Regarding the immune response, for example, one could begin with models of host–pathogen interactions that examine cell–cell communications through bacterial secretion systems as well as cytokine networks, and then start replacing the cells in the model, which are initially regarded as 'black boxes', with simulation of cell behavior modeled from intracellular network modules (e.g. models of cytoskeleton motility, proliferation or cytokine response and release), ultimately replacing black-box modules with bottom up-approaches. The Physiome Project (www.pysiome.org) [49], conceived by the Commission on Bioengineering in Physiology and presented at the International Union of Physiological Sciences (IUPS) council at their 32nd world congress in Glasgow, UK, in 1993, is a worldwide effort to define the physiome through the development of databases and models that will facilitate the understanding of the integrative function of cells, organs and organisms. The project is focused on compiling and providing a central repository of databases, linking experimental information and computational models from many laboratories into a single self-consistent framework. In the vision of the physiome project, such a framework enables and promotes multiscale modeling of physiological processes.

Entelos have developed complex simulations of disease physiology using a framework, PhysioLab (www.entelos.com/science/physiolabtech.html), for determining differential equations based on empirical data in humans [50]. In these models, cells, or even tissues, are represented as black boxes without explicit internal network models that respond to inputs by providing specified dynamic outputs. Using such an organ level framework of disease physiology, Stokes *et al.* [51] have developed a computational model of chronic asthma that includes interactions between cells and their response to each other and their environment. Different steady states of this disease, such as chronic asthma including chronic eosinophilic inflammation, chronic airway obstruction, airway hyper-responsiveness and elevated IgE levels, can be induced in the model. These *in silico* asthmatic models respond as expected to various drugs, such as β_2-agonists, glucocorticoids and leukotriene antagonists [51]. Furthermore, this model accurately predicts a decrease in airway eosinophils without much therapeutic improvement in airway conduction after reduction in interleukin (IL)-5 protein, as observed in clinical trials of an anti-IL-5 antibody in asthmatics.

Conclusion

Host–pathogen systems biology is still in its infancy. Although many aspects of host–pathogen systems have been addressed by experimental discoveries and mathematical and computational models, a comprehensive analysis of all aspects of host–pathogen interaction (including pathogen interference, host response, and pathogen response) is a far-fetched goal. Current model systems either focus on host response to pathogen infections or pathogen biology in a simulated host environment, and the size of these model systems depend essentially on the detail of description. They can be large interaction maps (bottom-up models) with thousands of components and interactions, conceptual top-down cellular or organ models consisting of interacting black boxes without detailed knowledge of internal processes within each individual black box, or small mechanistic dynamic models describing a few steps of a much larger system.

Efforts are currently underway to combine pathogen action and host response in a comprehensive, multiscale (hybrid) model, merging top-down approaches with 'omic' bottom-up approaches [52]. Combined with sophisticated experimental techniques, such as quantitative protein expression, tags by quantum dots for localization and nanobiotechnological measurements on single cells, new insights into the complexity of host–pathogen systems are within reach. Potential applications of host–pathogen systems biology range from biological target identification and drug discovery to bio-threat assessment and personalized health care. As with any modeling approach, theoretical models raise the challenge of experimental validation and the iterative cycle of improvement inherent to the modeling effort. Concerning drug discovery, success stories are still anecdotal – until a given model shows a track record of successful predictions it will be risky to rely on for drug development decisions. For the foreseeable future, modeling predictions will most likely be only one of many inputs into the decision-making process in the pharmaceutical industry. A long-term goal for host–pathogen systems biology, although still in the realm of science fiction, would include a full scale *in silico* model of an individualized human fighting against pathogenic infections.

Acknowledgements

Fruitful discussions with, and critical reading of the manuscript by, James Faeder are gratefully acknowledged.

References

1 Goldsby, R.A. *et al.* (1999) *Kuby Immunology*, 4th edition, W. H. Freeman & Co.
2 Kahn, R. *et al.* (2002) Cellular hijacking: a common strategy for microbial infection. *Trends Biochem. Sci.* 27, 308–314
3 Frost, S.D.W. (2002) Dynamics and evoluion of HIV-1 during structired treatment interruptions. *AIDS Rev.* 4, 119–127
4 Bugrim, A. *et al.* (2004) Early prediction of drug metabolism and toxicity: systems biology approach and modeling. *Drug Discov. Today* 9, 127–135
5 Butcher, E.C. (2005) Can cell systems biology rescure drug discovery? *Nat. Rev. Drug Discov.* 4, 461–467
6 Butcher, E.C. *et al.* (2004) Systems biology in drug discovery. *Nat. Biotechnol.* 22, 1253–1259
7 Strange, K. (2005) The end of "naïve reductionism": rise of systems biology or renaissance of physiology? *Am. J. Physiol. Cell Physiol.* 288, C968–C974
8 Bower, J.M. and Bolouri, H., eds (2001) *Computational modeling of genetic and biochemical networks*, MIT Press.
9 Forst, C.V. (2002) Network genomics – a novel approach for the analysis of biological systems in the post-genomic era. *Mol. Biol. Rep.* 29, 265–280
10 Barabási, A-L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113
11 Jeong, H. *et al.* (2001) Lethality and centrality in protein networks. *Nature* 411, 41–42
12 Schilling, C.H. (1999) Toward metabolic phenomics: analysis of genomic data using flux balances. *Biotechnol. Prog.* 15, 288–295
13 Schuster, S. (2000) A general definition of metabolic pathways useful for

systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* 18, 326–332
14 Covert, M.W. and Palsson, B.O. (2002) Transcriptional regulation in constraint-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* 277, 28058–28064
15 Kauffman, S.A. (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* 22, 437–467
16 Kauffman, S.A. (1993) *The origins of order: self-organization and selection in evolution*, Oxford University Press.
17 Savageau, M.A. (1969) Biological Systems Analysis, II. The steady-state solutions for an *n*-pool system using a power-law approximation. *J. Theor. Biol.* 25, 370–379
18 Savageau, M.A. (1970) Biological Systems Analysis, III. Dynamic solutions using a power-law approximation. *J. Theor. Biol.* 26, 215–226
19 Ackers, G.K. *et al.* (1982) Quantitative model for gene regulation by lambda repressor. *Proc. Natl. Acad. Sci. U. S. A.* 79, 1129–1133
20 Novak, B. and Tyson, J.J. (1993) Modeling the cell division cycle: M-phase trigger, oscillations, and size control. *J. Theor. Biol.* 165, 101–134
21 Gillespie, D. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361
22 McAdams, H.H. and Arkin, A. (1998) Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* 27, 199–224
23 Faeder, J.R. *et al.* (2003) Investigation of early events in FcεRI-mediated signaling using a detailed mathematical model. *J. Immunol.* 170, 3769–3781
24 Magasanik, B. (1995) Nitrogen response networks of yeast and bacteria. *J. Cell Biol.* (*suppl. 19A*), 326–326
25 Zien, A. *et al.* (2000) Analysis of gene expression data with pathway scores. In *Proceedings of ISMB'00* (Bourne, P. *et al.*, eds.), pp. 407–417, AAAI Press.
26 Ideker, T. *et al.* (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18(*Suppl. 1*), S233–S240
27 Dandekar, T. *et al.* (1999) Pathway alignment: application to the comparative analysis of glycolytic enzymes. *Biochem. J.* 343, 115–124
28 Forst, C.V. and Schulten, K. (1999) Evolution of metabolism: a new method for the comparison of metabolic pathways using genomic information. *J. Comput. Biol.* 6, 343–360
29 Forst, C.V. and Schulten, K. (2001) Phylogenetic Analysis of Metabolic Pathways. *J. Mol. Biol.* 52, 471–489
30 Ogata, H. *et al.* (2000) A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters. *Nucleic Acids Res.* 28, 4021–4028
31 Kelley, B.P. *et al.* (2003) Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc. Natl. Acad. Sci. U. S. A.* 100, 11394–11399
32 Nakaya, A. *et al.* (2001) Extraction of Correlated Gene Clusters by Multiple Graph Comparison. *Genome Inform. Ser. Workshop Genome Inform.* 12, 44–53
33 Cabusora, L. *et al.* (2005) Differential network expression during drug and stress response. *Bioinformatics* 21, 2898–2905
34 Hucka, M. *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531
35 Campbell, L.A. *et al.* (1993) Detection of *Chlamydia trachomatis* deoxyribonucleic acid in women with tubal infertility. *Fertil. Steril.* 59, 45–50
36 Byrne, G.I. *et al.* (1986) Induction of tryptophan catabolism is the mechanism for gamma-interferon-mediated inhibition of intracellular *Chlamydia psittaci* replication in T24 cells. *Infect. Immun.* 53, 347–351
37 Beatty, W.L. *et al.* (1994) Persistent Chlamydiae: from cell culture to a paradigm for chlamydial pathogenesis. *Microbiol. Rev.* 58, 686–699
38 Goldstein, B. *et al.* (2004) Mathematical and computational models of immune-receptor signaling. *Nat. Rev. Immunol.* 4, 445–456
39 Valitutti, S. *et al.* (1995) Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* 375, 148–151
40 San José, E. *et al.* (2000) Triggering the TCR complex causes the downregulation of nonengaged receptors by a signal transduction-dependent mechanism. *Immunity* 12, 161–170
41 McKeithan, T.W. (1995) Kinetic proofreading in T-cell receptor signal-transduction. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5042–5046
42 Wofsy, C. *et al.* (2001) Calculations show substantial serial engagement of T cell receptors. *Biophys. J.* 80, 606–612
43 Perelson, A. and Weisbuch, G. (1997) Immunology for physicists. *Rev. Mod. Phys.* 69, 1219–1267
44 Nowak, M.A. and May, R.M. (2000) *Virus dynamics: mathematical principles of immunology and virology*, Oxford Univ. Press.
45 Perelson, A.S. *et al.* (1996) HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271, 1582–1586
46 DeBoer, R.J. *et al.* (2001) Recruitment times, proliferation, and apoptosis rates during the $CD8^+$ T cell response to LCMV. *J. Virol.* 75, 10663–10669
47 Wodarz, D. *et al.* (2000) A new theory of cytotoxic T-lymphocyte memory: implications for HIV treatment. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 329–343
48 Dixit, N.M. and Perelson, A.S. (2004) Complex patterns of viral load decay under antiretroviral therapy: influence of pharmacokinetics and intracellular delay. *J. Theor. Biol.* 226, 95–109
49 Hunter, P.J. and Borg, T.K. (2003) Integration from proteins to organs: the Physiome Project. *Nat. Rev. Mol. Cell Biol.* 4, 237–243
50 Musante, C.J. *et al.* (2002) Small- and large-scale biosimulation applied to drug discovery and development. *Drug Discov. Today* 7, S192–S196
51 Stokes, C.L. *et al.* (2001) A computer model of chronic asthma with applications to clinical studies: example of treatment of exercise-induced asthma. *J. Allergy Clin. Immunol.* 107, 933
52 Kirschner, D. and Marino, S. (2005) *Mycobacterium tuberculosis* as viewed through a computer. *Trends Microbiol.* 13, 206–211
53 Akira, S. and Takeda, K. (2004) Toll-like receptor signaling. *Nat. Rev. Immunol.* 4, 499–511
54 Gordon, S. (2002) Pattern recognition receptors: doubling up for the innate immune response. *Cell* 111, 927–930
55 Platt, N. *et al.* (2002) The many roles of the class A macrophage scavenger receptor. *Int. Rev. Cytol.* 212, 1–40
56 Ernst, J.D. (1998) Macrophage receptors for *Mycobacterium tuberculosis*. *Infect. Immun.* 66, 1277–1281