The JAK-STAT Signaling Network in the Human B-Cell: An Extreme Signaling Pathway Analysis

Jason A. Papin and Bernhard O. Palsson
Department of Bioengineering, University of California, San Diego, La Jolla, California

ABSTRACT Large-scale models of signaling networks are beginning to be reconstructed and corresponding analysis frameworks are being developed. Herein, a reconstruction of the JAK-STAT signaling system in the human B-cell is described and a scalable framework for its network analysis is presented. This approach is called extreme signaling pathway analysis and involves the description of network properties with systemically independent basis vectors called extreme pathways. From the extreme signaling pathways, emergent systems properties of the JAK-STAT signaling network have been characterized, including 1), a mathematical definition of network crosstalk; 2), an analysis of redundancy in signaling inputs and outputs; 3), a study of reaction participation in the network; and 4), a delineation of 85 correlated reaction sets, or systemic signaling modules. This study is the first such analysis of an actual biological signaling system. Extreme signaling pathway analysis is a topologically based approach and assumes a balanced use of the signaling network. As large-scale reconstructions of signaling networks emerge, such scalable analyses will lead to a description of the fundamental systems properties of signal transduction networks.

INTRODUCTION

Reconstructed biochemical reaction networks are foundational to systems analysis in biology. Large-scale reconstruction efforts have been successful for metabolic and regulatory networks (Covert and Palsson, 2002; Forster et al., 2003; Pramanik and Keasling, 1997; Reed and Palsson, 2003; Selkov et al., 1998); however, such efforts for large-scale signaling processes are in their infancy (Bhalla and Iyengar, 1999; Gilman et al., 2002). *In silico* analysis frameworks for these reconstructed signaling networks will need to be scalable and able to describe emergent properties that arise from the interconnectivity of the network constituents. A recent approach has been developed to study the topological properties of signaling networks (Papin and Palsson, 2003), called extreme signaling pathway analysis (ExSPA). This approach uses extreme pathway analysis (Schilling et al., 2000) to characterize the properties of signaling networks. ExSPA has been applied to a prototypic signaling network to define and study properties of signaling networks (Papin and Palsson, 2003). Systems properties including input/output relationships and crosstalk were mathematically defined and described, and other emergent properties were characterized, including correlated reaction sets, pathway redundancy, and the participation of reactions in network-based pathways.

Network reconstruction involves the integration of multiple datasets to generate increasingly more accurate models of biological processes (Herrgard et al., 2004; Price et al., 2003; Reed and Palsson, 2003). The B-cell was recently selected to apply large-scale approaches to elucidate signaling networks (Gilman et al., 2002). The JAK-STAT signaling network, particularly important for many immune responses, is well-characterized in the human B-cell as well as many other cell types (Aaronson and Horvath, 2002) (Fig. 1). Typically, the binding of a corresponding ligand to its receptor induces dimerization of the receptor, which in turn results in the activation of an associated kinase called a JAK. The activated JAK protein then induces the phosphorylation of a protein from the family of signal transducers and activators of transcription (STATs). These STATs can form homo- and heterodimers. Following the STAT dimerization event, these proteins translocate into the nucleus and induce expression of their target genes.

This study presents a large-scale reconstruction of the JAK-STAT signaling network in the human B-cell. The extreme signaling pathways were computed and an analysis of the systems properties of the reconstructed JAK-STAT network was then performed based on methods previously developed (Papin and Palsson, 2003).

CONCEPTUAL FRAMEWORK AND METHODS

Stoichiometric formalism of signaling networks

The constraint-based modeling framework allows for the analysis of biological networks by successively applying known constraints such as mass balance, maximum capacity, and reaction irreversibility (Price et al., 2003). After the application of these known constraints, the remaining solution space can be characterized by calculating convex basis vectors that provide a way to represent every possible flux state of the network (Schilling et al., 2000). These convex basis vectors are fundamental pathways of the network, and studying them for genome-scale metabolic networks has yielded biologically meaningful results (Papin et al., 2003). Signaling network events are subject to mass balance and thermodynamic constraints.

Submitted June 23, 2003, and accepted for publication March 9, 2004.

Address correspondence to Bernhard O. Palsson, Dept. of Bioengineering, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0412. Tel.: 858-534-5668; E-mail: palsson@ucsd.edu.

0006-3495/04/07/37/10 $2.00

doi: 10.1529/biophysj.103.029884

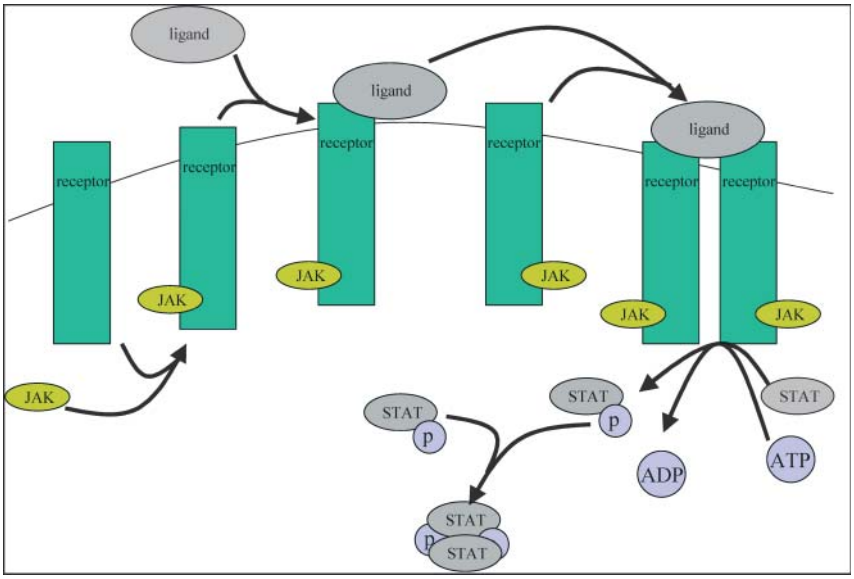

FIGURE 1 Schematic of generalized reactions for the JAK-STAT signaling network.

Consequently, *in silico* analysis methods within the constraint-based framework developed for metabolic and regulatory networks can be applied to signaling networks (Papin and Palsson, 2003).

A recently developed analysis method within the constraint-based framework is extreme pathway analysis (Papin et al., 2003). With this approach, the first step is the creation of a stoichiometric matrix to represent the primary chemical events that take place within a network. The rows of this matrix correspond to network components (e.g., adenosine triphosphate (ATP), interleukin 2 receptor, STAT6, etc.). The columns of this matrix correspond to reactions (e.g., binding of interleukin 2 to its receptor, homodimerization of STAT6, etc.). Each element of the matrix contains the stoichiometric coefficient of the given component in the associated reaction.

With a stoichiometric formalism, the underlying reactions of a network model are explicitly defined and forced to be chemically consistent. Since this formalism requires explicit description of the reaction mechanisms, each ''state'' of a component must be accounted for in the network. For example, a protein is differentially represented if it is phosphorylated than if it is not phosphorylated. With such an explicit description of all the chemical transformations in a network, the systemic effects of each component of a network can be readily assessed.

Extreme signaling pathway analysis

Once the stoichiometric matrix has been defined, topological analyses can be used to make characterizations of network properties. Extreme pathways (Schilling et al., 2000), elementary modes (Schuster et al., 2000), and extreme currents (Clarke, 1988) are topological analysis methods based on convex analysis (Rockafellar, 1970). Convex analysis allows for the study of systems of equations with inequality constraints. Extreme pathways are the minimal set of conically independent basis vectors that completely characterize the fundamental functional states of a given reaction network and that satisfy constraints on the directions of the network reactions (Papin et al., 2003). Since extreme pathways are calculated directly from a stoichiometric matrix, subject to constraints on reaction direction, they can directly correspond to integrated functions of the signaling network as a whole.

Extreme pathways have the following important characteristics: 1), they are a unique and systemically independent set of basis vectors; 2), all routes through the network can be described by nonnegative linear combinations of the extreme pathways; and 3), time-invariant network properties are characterized (Papin et al., 2003). Recently, this approach was extended to describe signaling networks (Papin and Palsson, 2003), and herein the application of extreme pathways to signaling networks is called ExSPA.

Time-scale separation in signaling networks

Extreme pathway analysis has been extensively used for the analysis of metabolic networks (Papin et al., 2003). Since metabolic processes occur on a scale of milliseconds to seconds and regulatory and growth processes occur on a scale of minutes to hours, a quasi-steady-state assumption allows for the interpretation of the extreme pathways as steady-state flux distributions in a metabolic network (Covert et al., 2001; Varma and Palsson, 1994). The quasi-steady-state assumption has allowed for simulations of dynamic concentration profiles for *Escherichia coli* metabolism (Covert and Palsson, 2002; Varma and Palsson, 1994) using standard temporal decomposition approaches. The timescales for events in signaling networks like kinase and phosphatase activity (Goodman et al., 1998; Lillemeier et al., 2001; Theurkauf, 1994; Vuong and Chabre, 1991), receptor internalization (Ferguson, 2001; Jullien et al., 2002), and regulatory processes (McAdams and Arkin, 1998; Rivett, 1986; Zubay, 1973) are approximately known (see Table 1). Consequently, extreme pathways of signaling networks may be interpreted as steady-state flux distributions. If these transients are too rapid to be of importance they can be relaxed and only the eventual transcription state is analyzed.

This approximation is also valid for the JAK-STAT signaling network. Although inhibitory reactions have not been delineated in the reconstructed network presented herein, their activity is an important consideration for the

TABLE 1 Approximation of the order of magnitudes for signaling processes

Process	Order of magnitude	References
Signaling reactions	$<10^{\circ}$ s	Goodman et al. (1998); Vuoung et al. (1991); Lillemeier et al. (2001); Theurkauf et al. (1994)
Transcriptional regulation	10^2 s	Zubay (1973); Rivett (1986); McAdams and Arkin (1998)
Receptor internalization	10^2 s	Jullien et al. (2002); Ferguson et al. (2001)

application of extreme signaling pathways to its analysis and the steady-state flux approximation. The genes for the well-characterized suppressor of cytokine signaling (SOCS) inhibitors of JAK-STAT signaling (Shuai and Liu, 2003) are expressed as a result of STAT transcriptional activity. The regulation of other inhibitors of JAK-STAT signaling (e.g., PIAS (protein inhibitor of activated STAT) and SHP (Src homology phosphotase) proteins) is less characterized (Shuai and Liu, 2003; Wormald and Hilton, 2004). Some data suggest that JAK-STAT signaling peaks at ~30 min to 1 h and can be sustained for multiple hours (Kalvakolanu, 2003), suggesting that these mechanisms of inhibition operate over approximately an hour time frame. Combined, these data suggest that the time frame for chemical transformations involved in JAK-STAT signaling is an order of magnitude slower than transcriptional events and the inhibition of the signaling. Since the activated receptor-ligand complex can continue to phosphorylate STAT proteins, there is a constant ''source'' of flux in the network. The ''sink'' for this flux may be the transport of STAT proteins into the nucleus where they lead to transcriptional events.

RESULTS

The JAK-STAT signaling network of the human B-cell was reconstructed and ExSPA was applied to it. Emergent properties can be studied using ExSPA, including 1), crosstalk; 2), pathway redundancy; 3), reaction participation; and 4), correlated reaction sets. These properties represent novel structural characterizations of the JAK-STAT signaling network in the human B-cell.

Reconstruction of JAK-STAT signaling network in the human B-cell

A stoichiometric matrix was constructed for the chemical reactions that characterize the known JAK-STAT signaling system in the human B-cell. There are 15 receptors and 15 corresponding ligands accounted for in this reconstruction. Each of these receptors has been identified in the human B-cell. There are a total of 297 reactions (216 internal and 81 irreversible exchange) in the reconstructed JAK-STAT signaling network (the corresponding reactions and associated references can be found in the supplementary material). These reactions involve 15 ligand inputs and seven STAT homo- and heterodimer outputs as listed in Table 2. This set of reactions can be represented with a stoichiometric matrix with reactions as columns and components as rows. The metabolic and protein components of the signaling network are also shown in Table 2. The full list of components illustrates the interconnectivity of the signaling network with other cellular processes, namely metabolism and protein synthesis/degradation. For example, the cofactor conversion of ATP to ADP drives the signaling system, and ATP resynthesis is a primary function of metabolism. Additionally, receptors will need to be synthesized. Such interconnectivity between ''disparate'' cellular processes emphasizes the need for integrated modeling approaches.

The actual process of network reconstruction can lead to useful biological results. For example, it is important to characterize the specificity of STAT substrates to the proper JAK and the receptor. In this reconstruction, the phosphorylation of the STATs is specific to a particular JAK-receptor complex. If the STATs were specific solely to the receptor, or solely to the JAK, the analysis would yield different results. These characterizations illustrate the hypotheses that can be generated (and subsequently tested) through the process of network reconstruction.

TABLE 2 Inputs and outputs of the JAK-STAT signaling network in the human B-cell

Inputs	Outputs
ATP	ADP
JAK1	Prolactin-JAK2 receptor-ligand complex
JAK2	Interleukin 2-JAK1 receptor-ligand complex
JAK3	Interleukin 2-JAK3 receptor-ligand complex
TYK2	Interleukin 3-JAK2 receptor-ligand complex
STAT1	Interleukin 5-JAK2 receptor-ligand complex
STAT2	Interleukin 6-JAK1 receptor-ligand complex
STAT3	Interleukin 7-JAK1 receptor-ligand complex
STAT4	Interleukin 7-JAK3 receptor-ligand complex
STAT5A	Interleukin 9-JAK1 receptor-ligand complex
STAT6	Interleukin 9-JAK3 receptor-ligand complex
Prolactin	Interleukin 10-JAK1 receptor-ligand complex
Interleukin 2	Interleukin 10-TYK2 receptor-ligand complex
Interleukin 3	Interleukin 10-TYK2-JAK1 receptor-ligand complex
Interleukin 4	Interleukin 11-JAK1 receptor-ligand complex
Interleukin 5	Interleukin 12-JAK2 receptor-ligand complex
Interleukin 6	Interleukin 12-TYK2 receptor-ligand complex
Interleukin 7	Interleukin 12-TYK2-JAK2 receptor-ligand complex
Interleukin 9	Interleukin 13 receptor-ligand complex
Interleukin 10	Interleukin 13-JAK1 receptor-ligand complex
Interleukin 11	Interleukin 13-JAK2 receptor-ligand complex
Interleukin 12	Interleukin 13-JAK2-JAK1 receptor-ligand complex
Interleukin 13	Interleukin 14-JAK1 receptor-ligand complex
Interleukin 15	Interleukin 14-JAK3 receptor-ligand complex
Interferon α/β	Interleukin 14-JAK3-JAK1 receptor-ligand complex
Interferon γ	Interleukin 15-JAK1 receptor-ligand complex
Prolactin receptor	Interleukin 15-JAK3 receptor-ligand complex
Interleukin 2 receptor	Interferon α/β-JAK1 receptor-ligand complex
Interleukin 3 receptor	Interferon α/β-TYK2 receptor-ligand complex
Interleukin 4 receptor	Interferon α/β-TYK2-JAK1 receptor-ligand complex
Interleukin 5 receptor	Interferon γ-JAK1 receptor-ligand complex
Interleukin 6 receptor	Interferon γ-JAK2 receptor-ligand complex
Interleukin 7 receptor	Interferon γ-JAK2-JAK1 receptor-ligand complex
Interleukin 9 receptor	STAT1 homodimer
Interleukin 10 receptor	STAT3 homodimer
Interleukin 11 receptor	STAT4 homodimer
Interleukin 12 receptor	STAT5A homodimer
Interleukin 13 receptor	STAT6 homodimer
Interleukin 15 receptor	STAT1-STAT2 heterodimer
Interferon α/β receptor	STAT1-STAT3 heterodimer
Interferon γ receptor	

Enumeration of extreme pathways

The extreme pathways were calculated for the JAK-STAT signaling network (see Schilling et al., 2000 for a description of the algorithm). There are 147 extreme pathways in this network and they can be categorized in two groups (Fig. 2). The first category represents the classical signaling mechanism involving one input and one output; there are 37 of the

147 extreme pathways in this group. The second category represents a concatenation of inputs to generate one output; there are 110 of the 147 extreme pathways in this group. Other categories of extreme signaling pathways, described elsewhere (Papin and Palsson, 2003), are not found in this set of extreme pathways, including signal pleiotropy (one signaling input generates multiple signaling outputs) and complex signaling events (multiple signaling inputs generate multiple signaling outputs).

The complete set of extreme pathways of the JAK-STAT signaling network are listed in the supplementary material. Two representative pathways are shown in Fig. 3. Pathway 119 (the numbers represent the order in which the extreme pathways are computed) is the formation of the activated ligand-receptor complex. Pathway 1 demonstrates some of the interconnectivity that exists in the JAK-STAT signaling network herein (Fig. 3). This pathway describes the activation of different STATs that then form heterodimers (e.g., STAT1-STAT2 heterodimer is activated as a result of the binding of ligands interferon γ and interferon α/β to their respective receptors). Although these pathways generally describe ''silos'' that connect the activated ligand-receptor complex to activated transcription factor, the reconstruction of the JAK-STAT signaling network is easily expanded to include inhibitory reactions and additional signaling moieties (e.g., mitogen-activated protein kinase (MAPK) reactions) which will lead to further interconnectivity. These pathways also depict the interconnectivity between signaling, metabolic, and transcriptional regulatory networks.

Crosstalk in the JAK-STAT signaling network

''Crosstalk'' in signaling networks has been defined as network states that are a nonnegative linear combination of extreme signaling pathways (Papin and Palsson, 2003) and as such represents the interaction of systemically independent routes in a reaction network. This system property can be characterized by making pairwise comparisons of the entire set of extreme signaling pathways. A pair of pathways may have identical, overlapping, or disjoint input sets (e.g., a pathway with interleukin-2 and interleukin-4 as input signals would overlap with a pathway that had interleukin-2 and interleukin-15 as input signals). Each pair of pathways may likewise have identical, overlapping, or disjoint output sets. Consequently, there are nine categories of crosstalk. Since the extreme signaling pathways are a unique and minimal set of basis vectors, this unambiguous definition and categorization of crosstalk describes the interaction of fundamental functional states in a signaling network.

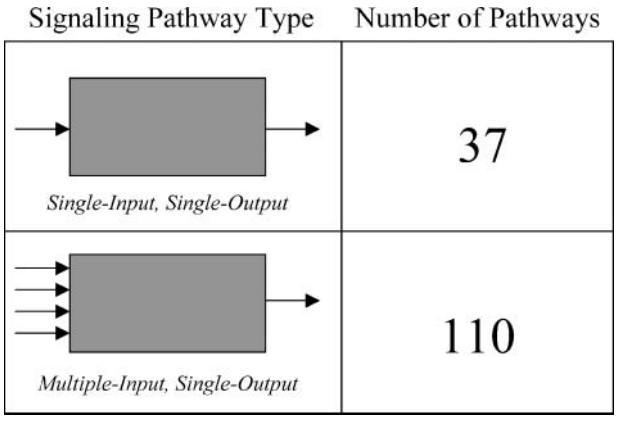

FIGURE 2 Signaling pathway types in the JAK-STAT network. Other signaling types, not present in this network, include signal pleiotropy (one-input, multiple-outputs) and cross talk (multiple inputs, multiple outputs).

Various forms of crosstalk in the JAK-STAT signaling network can be characterized with the ExSPA approach (Papin and Palsson, 2003). For the JAK-STAT system, the pairwise combinations of all extreme pathways (crosstalk) have been grouped into one of the nine categories described above. These categories and their respective biological interpretations have been described (Papin and Palsson, 2003). For example, pairwise combinations of pathways with disjoint inputs and disjoint outputs may correspond to completely independent functions of a network even though intracellular reaction may be shared. Pairwise combinations of pathways with overlapping inputs and disjoint outputs may correspond to economized uses of the network allowing a small difference in a signaling input to generate a distinct biological response.

The percentages of each type of crosstalk found in the reconstructed JAK-STAT network are listed in Fig. 4. With 147 extreme pathways, there are 10,731 pairwise combinations ($(147^2 - 147)/2$). Approximately 99.8% of the pairs of extreme signaling pathways have disjoint outputs and nearly 0.2% have identical outputs. There are no pairwise combinations of extreme pathways with overlapping outputs since all the extreme pathways of the JAK-STAT network have only one signaling output. Approximately 63.9%, 21.3%, and 14.8% of the pairwise combinations have disjoint, overlapping, and identical input sets, respectively. The high percentage of pairwise combinations with disjoint sets of inputs and disjoint sets of outputs indicates a fairly deterministic signaling network; there is very little ''classical'' crosstalk (i.e., identical signaling molecules used in different signaling pathways (Schwartz and Baron, 1999)) as one input signal typically corresponds to one output signal. As more signaling networks are comprehensively reconstructed, the percentages of types of crosstalk in different systems will be characterized and comparative systems properties will be described.

Signaling redundancy

''Pathway redundancy'' is the multiplicity of routes through a network by which identical inputs can generate identical outputs. This emergent property has been characterized in genome-scale metabolic networks (Papin et al., 2002a; Price

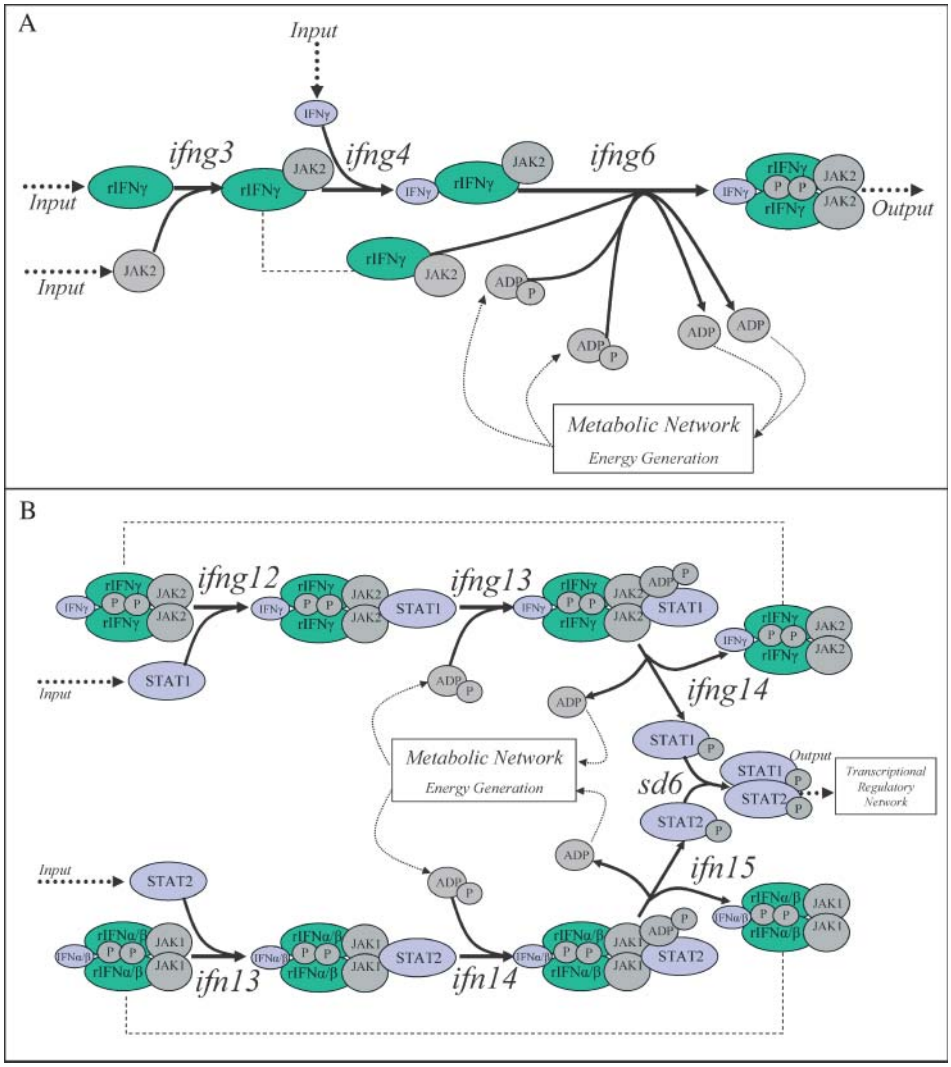

FIGURE 3 Examples of extreme signaling pathways of the JAK-STAT signaling network. Pathway 119 (*A*) corresponds to the binding of the interferon γ ligand to its receptor and the subsequent formation of an activated receptor-ligand complex. Pathway 1 (*B*) corresponds to the activation of the STAT1-STAT2 heterodimer by the interferon γ and interferon α/β bound receptors. The names of the reactions are in italics above the reaction arrows (see supplementary material).

et al., 2002) and a prototypic signaling network (Papin and Palsson, 2003). A significant amount of redundancy in signaling networks would suggest a lack of specificity for generating a particular output with multiple different signaling inputs. A small degree of redundancy in a signaling network would suggest a highly specific relationship between inputs and outputs.

Signaling pathway redundancy has been characterized for the JAK-STAT signaling network (Table 3). The systemic signaling pathway redundancy calculation indicated that on average there were four independent routes (extreme pathways) with identical signaling inputs and identical signaling outputs (data not shown). The signaling pathway redundancy can be further discriminated by the level of output and input redundancy. Output redundancy is the number of extreme pathways with identical outputs; a high output redundancy corresponds to output signals that can be generated by way of multiple systemically independent routes. Input redundancy is the number of extreme pathways with identical inputs; it is an indication of the inputs that have a high degree of influence in the signaling network as each corresponding extreme pathway corresponds to a unique function in the network. Output redundancy and input redundancy are related to the previous concepts of crosstalk between pairs of extreme pathways with identical output and input signals, respectively (Fig. 4, 3rd column and 3rd row). The output redundancy and input redundancy values have been calculated. The synthesis of the STAT1-STAT3 heterodimer is the most redundant and the synthesis of the STAT3, STAT4, and STAT6 homodimers is the least redundant. High redundancy values are indicative of signal inputs and outputs to which the signaling network is particularly robust under network modifications.

Reaction participation in the JAK-STAT signaling network

The ''reaction participation'' is the percentage of extreme pathways in which a given reaction is used (Papin et al., 2002b). The reaction participation values have been

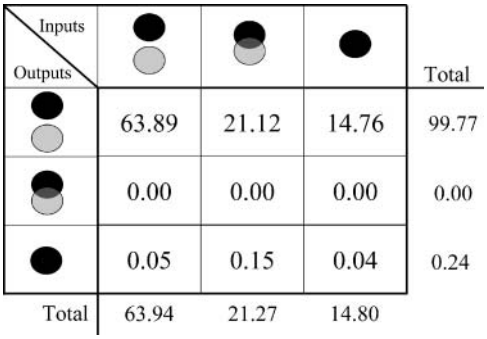

FIGURE 4 Cross talk in the JAK-STAT signaling network of the human B-cell. There are 10,731 pairwise combinations of the 147 extreme pathways. Each value is the percentage of all pairwise combinations that belong to the corresponding category.

calculated for the JAK-STAT signaling network (Fig. 5). The reactions with participation values greater than 10% are shown in the inset table. Plotting the data on a log-log scale does not result in a strict linear relationship. However, there is a decreasing relationship between the number of reactions and their corresponding participation numbers, which indicates the importance of a small number of reactions in determining the phenotypic potential of the signaling network.

For example, the exchange reactions of ATP and ADP have 100% participation. These components are essential for all the states of the signaling network since phosphate transfer is the mechanism by which the signal propagates. The exchange reactions (i.e., reactions that describe the transfer of components across the system boundary) for STAT1, STAT3, the STAT1-STAT3 heterodimer, and the reaction *SD7*, which describes the formation of the STAT1-STAT3 heterodimer, have the next highest participation values. This result is in part a reflection of the calculations described above regarding the high degree of output redundancy for the STAT1-STAT3 heterodimer; the network is structured such that there are multiple routes to synthesize the STAT1-STAT3 heterodimer. There are 168 reactions that participate in only one extreme pathway. Reactions with low participation values correspond to reactions that have very specific network functions. Manipulating reactions with low participation values could allow for control of very specific functions, potentially important for drug targeting.

The exchange reactions for STAT1, STAT3, STAT4, STAT5A, and STAT6 homodimers participated in 5%, 4%, 4%, 7%, and 4% of the extreme signaling pathways, respectively. These reaction participation values do not correlate with the number of cytokines that activate the respective STATs (i.e., the STATs activated by a greater number of cytokines do not necessarily participate in a greater number of reactions (see supplementary material)). The reaction participation values for the exchange reactions of heterodimers are significantly higher than the exchange reactions for homodimers. The exchange reactions for STAT1-STAT3 and STAT1-STAT2 heterodimers participated in 33% and 16% of the extreme signaling pathways, respectively. This result suggests that the combinatorial expansion around a small number of already existing components can allow for a significant increase in the number of distinct phenotypes. Similarly, the crosstalk analysis indicated that 21% of the pairwise comparisons consist of overlapping inputs and disjoint outputs. Together these results suggest that signaling networks may be designed to expand around an existing repertoire to generate additional responses to environmental stimuli.

Correlated reaction sets

Reactions that always appear together in the set of extreme pathways for a given network have been called ''correlated reaction sets'' (Papin et al., 2002b); in other words, all the possible states of the network that use these reactions use them together. The correlated reaction sets for the JAK-STAT signaling network in the human B-cell have been calculated (Table 4). There are 85 correlated reaction sets. An obvious correlated reaction set is the exchange reactions for ATP and ADP; whenever ATP goes into the system, ADP consequently has to come out (set 1). Additional correlated reaction sets include the exchange reactions for the ligand and receptor pairs (e.g., set 4). However, there are exchange reactions for ligands and receptors that belong to reaction sets with additional reactions (e.g., set 2). The receptors associated with these ligands only bind to one member of the JAK family in this network and consequently

TABLE 3 Signaling redundancy

STAT1 Homodimer	STAT1-STAT2 Heterodimer	STAT1-STAT3 Heterodimer	STAT3 Homodimer	STAT4 Homodimer	STAT5A Homodimer	STAT6 Homodimer	Number
0	0	1	0	0	0	0	48
0	1	0	0	0	0	0	24
0	0	0	0	0	1	0	11
1	0	0	0	0	0	0	8
0	0	0	0	0	0	1	6
0	0	0	0	1	0	0	6
0	0	0	1	0	0	0	6

The number of extreme signaling pathways with equivalent signal outputs and their respective signaling output are shown.

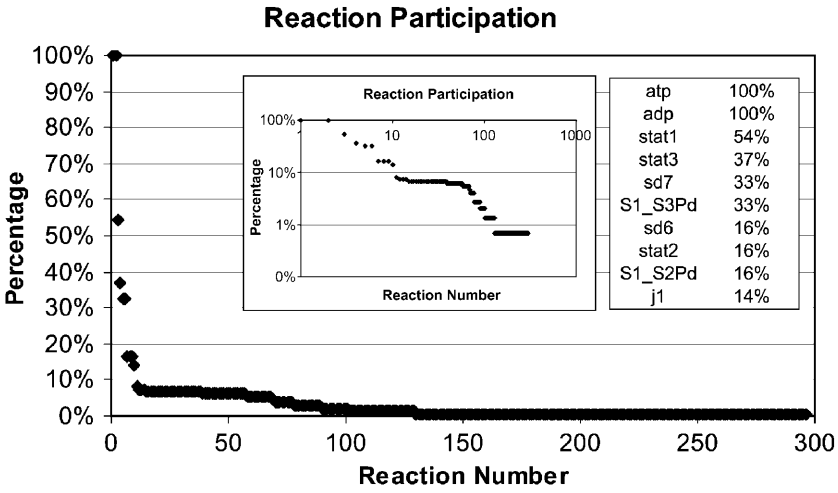

FIGURE 5 Reaction participation in JAK-STAT signaling network. The percentage of the extreme signaling pathways in which the corresponding reaction participates is indicated on the *y* axis. The inset table only lists the reaction participation values greater than 20% (abbreviations in supplementary material). The inset plot is the reaction participation plotted on a log-log scale.

the additional reactions are only present with the input of the ligand and its corresponding receptor.

There are additional distinctions between the correlated reaction sets that correspond to biological properties of the JAK-STAT signaling network. One separation of the reaction sets occurs with the ligands that result in the activation of STAT proteins by way of different JAKs. For example, reaction sets 5 and 6 each include reactions that describe the phosphorylation of STAT6 by way of JAK1-associated and JAK3-associated IL-4 dimerized receptors, respectively. The supplementary material contains the complete list of correlated reaction sets and associated reaction and compound abbreviations, as well as corresponding references.

TABLE 4 Representative correlated reaction sets

Set number	Reaction name	Stoichiometry/Exchange
1	atp	INPUT
	adp	OUTPUT
2	prl1	rPRL + j2 → rPRLj2
	prl2	rPRLj2 + PRL → rlPRLj2
	prl3	rlPRLj2 + rPRLj2 + 2 atp → 2 adp + 2PRLj2
	PRL	INPUT
	rPRL	INPUT
	2PRLj2	OUTPUT
3	prl4	2PRLj2 + stat5A → 2PRLj2s5A
	prl5	2PRLj2s5A + atp → PRLj2s5Ap
	prl6	PRLj2s5Ap → 2PRLj2 + stat5AP + adp
4	I4	INPUT
	rI4	INPUT
5	il4i	2I4j1 + stat6 → 2I4j1_s6
	il4j	2I4j1_s6 + atp → I4j1s6p
	il4k	I4j1s6p → 2I4j1 + adp + stat6P
6	il4l	2I4j3 + stat6 → 2I4j3_s6
	il4m	2I4j3_s6 + atp → I4j3s6p
	il4n	I4j3s6p → 2I4j3 + adp + stat6P

Each set of reactions always appears together in a set of extreme pathways for the JAK-STAT signaling network. The complete list of correlated reaction sets, as well as the reaction and component abbreviations, are provided in the supplementary material.

These correlated reaction sets provide hypotheses regarding regulatory control of the JAK-STAT signaling network. It has been hypothesized that correlated reaction sets in metabolic networks correspond to enzymes located on the same operon or controlled by the same regulon (Papin et al., 2002b). In signaling networks, correlated reaction sets may correspond to proteins that are coordinately regulated. If a defined regulatory rule does not correspond to a correlated reaction set, there may be additional functions of the signaling components within the set that have not yet been characterized. Thus, the correlated reaction set could be used to generate hypotheses regarding regulatory rules and network function of signaling components.

DISCUSSION

The JAK-STAT signaling network in the human B-cell has been reconstructed. The extreme signaling pathways of this network have been calculated and described. From the extreme signaling pathways, emergent systems properties have been characterized. These properties of the JAK-STAT signaling network include 1), a definition of network crosstalk; 2), an analysis of redundancy in signaling inputs and outputs; 3), a study of reaction participation in the network; and 4), a delineation of correlated reaction sets. This integrated, mathematical analysis of network properties has been previously performed for a prototypic signaling network; this study is the first such analysis of an actual biological signaling system, and novel properties have been described.

Extreme signaling pathway analysis provides a concise and unique description of network properties. Consequently, extreme signaling pathways enable unambiguous descriptions of systems properties. The extreme signaling pathways are a conically independent basis set (Schilling et al., 2000)

and thus all physiological network steady states can be decomposed into this set of pathways (Wiback et al., 2003). Since they are derived from the application of mass balance constraints, extreme signaling pathways represent actual signaling states in a network, and thus described emergent properties correspond to actual biological characteristics.

One such biological property is crosstalk, which has been defined as the nonnegative linear combination of extreme signaling pathways (Papin and Palsson, 2003). Since the JAK-STAT network studied herein is fairly deterministic, the types of cross talk that occur are limited; nearly 64% of pairwise combinations of extreme signaling pathways fall in the category of disjoint inputs and disjoint outputs. Nonetheless, with ~21% of the pairwise comparisons in the category with overlapping inputs and disjoint outputs, there is more complex control as the particular combination of inputs can generate distinct signaling outputs. This analysis of crosstalk is an explicit description of the important signaling network property of interconnectivity.

In addition to the novel description of crosstalk for this signaling network, the analyses of pathway redundancy, reaction participation, and correlated reaction sets have generated insights into network properties. For the JAK-STAT network described herein, the formation of the STAT1-STAT3 heterodimer has the highest degree of pathway redundancy. This result suggests that the formation of this transcription factor complex is more robust to modifications in the network structure. Measures of pathway redundancy indicate network functions that are both highly specialized (low redundancy) and very generalized (high redundancy), perhaps suggestive of system objectives. Reaction participation calculations make a parallel analysis of reaction function. The removal of reactions with low participation in the extreme pathways would result in specific control of a signaling network. The identification of correlated reaction sets provides hypotheses regarding the coordination of regulatory control. These emergent properties have been discussed in detail for metabolic networks and have been described for signaling networks (Papin and Palsson, 2003).

Input/output relationships for the JAK-STAT signaling network are relatively simple. Consequently, the emergent properties calculated for the previous prototypic signaling network (Papin and Palsson, 2003) have not all been calculated for the JAK-STAT signaling network. Such properties should arise in larger signaling networks once they are reconstructed. For example, the JAK-STAT signaling network can be expanded to include other signaling components like STAT inhibitory proteins (e.g., SOCS protein (Yamada et al., 2003)), as well as MAPK proteins and G-proteins which have both been implicated in STAT activation (Jain et al., 1998; Pelletier et al., 2003). Importantly, this modeling approach is scalable. Further refinements and modifications will not change the types of results that have been presented here. Rather, discrepancies between results in the different models will generate hypotheses that can be tested and resolved.

There is much interest in describing and analyzing ''modules'' and ''motifs'' of biochemical networks (Hartwell et al., 1999; Milo et al., 2002; Rives and Galitski, 2003). Signaling modules often consist of arbitrary groups of adjacent signaling reactions in the network that may operate together (e.g., (Bhalla and Iyengar, 1999; Hoffmann et al., 2002)). These signaling modules may serve as important conceptual tools for studying related signaling events. Their definition is based on local network topology and a priori definition of the investigator. In contrast, the extreme signaling pathways are network-based characteristics and the emergent properties calculated based on them carry no investigator biases. Since extreme pathways describe balanced uses of a biochemical network, the correlated reaction sets described in this study are effectively ''systemic modules'' because they are generated with consideration of all network demands and the reactions in a set always function together. Such systemic modules may contain reactions that are not topologically adjacent in a visual representation of the network (Papin et al., 2002b).

Extreme signaling pathway analysis details structural properties of the mass-balanced representation of the JAK-STAT signaling network. To date, the approach discussed herein does not allow for dynamic analyses of concentration profiles like recent analyses of signaling modules (e.g., Schoeberl et al., 2002). The JAK-STAT signaling model does not yet include genetic feedback mechanisms that will lead to a description of additional complex behavior important for the physiology of the human B cell. Accounting for genetic feedback and other dynamic processes in a signaling network (e.g., an activated transcription factor induces the synthesis of a protein that in turn inhibits the corresponding signaling pathway) will require further work on timescale separation of signaling, regulatory, and metabolic processes. The JAK-STAT signaling network herein is an example of the interconnectivity that exists between signaling, metabolic, and transcriptional regulatory networks (Fig. 6). This interconnectivity can be accounted for with stoichiometric representations of biochemical transformations and their subsequent analysis.

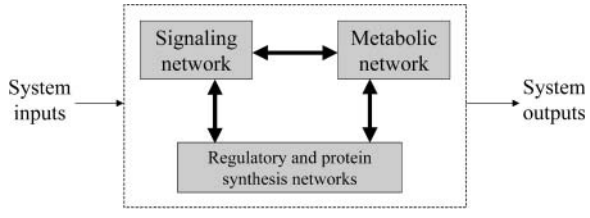

FIGURE 6 Signaling, metabolic, and transcriptional regulatory networks are interconnected. For example, ATP and GTP, primary currencies of signaling networks, are principal products of metabolic networks.

Taken together, this study presents the first constraint-based analysis of an actual signaling network. The JAK-STAT signaling system in the human B-cell is of central importance in immune response and other vital cellular processes. As the signaling network is further characterized and reconstructed, the calculated extreme signaling pathways may allow for the description of novel biological properties that will be important in medical and biotechnological applications as they previously have for metabolic networks (Carlson et al., 2002; Liao et al., 1996).

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

We thank the Whitaker Foundation (Graduate Research Fellowship to JP) for financial support. We also thank Scott Becker, Nathan Price, and Shankar Subramaniam for useful discussions.

REFERENCES

Aaronson, D. S., and C. M. Horvath. 2002. A road map for those who know JAK-STAT. *Science.* 296:1653–1655.

Bhalla, U. S., and R. Iyengar. 1999. Emergent properties of networks of biological signaling pathways. *Science.* 283:381–387.

Carlson, R., D. Fell, and F. Srienc. 2002. Metabolic pathway analysis of a recombinant yeast for rational strain development. *Biotechnol. Bioeng.* 79:121–134.

Clarke, B. L. 1988. Stoichiometric network analysis. *Cell Biophys.* 12: 237–253.

Covert, M. W., and B. O. Palsson. 2002. Transcriptional Regulation in Constraints-based Metabolic Models of Escherichia coli. *J. Biol. Chem.* 277:28058–28064.

Covert, M. W., C. H. Schilling, and B. Palsson. 2001. Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* 213: 73–88.

Ferguson, S. S. 2001. Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. *Pharmacol. Rev.* 53:1–24.

Forster, J., I. Famili, P. Fu, B. B. Palsson, and J. Nielsen. 2003. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res.* 13:244–253.

Gilman, A. G., M. I. Simon, H. R. Bourne, B. A. Harris, R. Long, E. M. Ross, J. T. Stull, R. Taussig, A. P. Arkin, M. H. Cobb, J. G. Cyster, P. N. Devreotes, J. E. Ferrell, D. Fruman, M. Gold, A. Weiss, M. J. Berridge, L. C. Cantley, W. A. Catterall, S. R. Coughlin, E. N. Olson, T. F. Smith, J. S. Brugge, D. Botstein, J. E. Dixon, T. Hunter, R. J. Lefkowitz, A. J. Pawson, P. W. Sternberg, H. Varmus, S. Subramaniam, R. S. Sinkovits, J. Li, D. Mock, Y. Ning, B. Saunders, P. C. Sternweis, D. Hilgemann, R. H. Scheuermann, D. DeCamp, R. Hsueh, K. M. Lin, Y. Ni, W. E. Seaman, P. C. Simpson, T. D. O'Connell, T. Roach, S. Choi, P. Eversole-Cire, I. Fraser, M. C. Mumby, Y. Zhao, D. Brekken, H. Shu, T. Meyer, G. Chandy, W. D. Heo, J. Liou, N. O'Rourke, M. Verghese, S. M. Mumby, H. Han, H. A. Brown, J. S. Forrester, P. Ivanova, S. B. Milne, P. J. Casey, T. K. Harden, J. Doyle, M. L. Gray, S. Michnick, M. A. Schmidt, M. Toner, R. Y. Tsien, M. Natarajan, R. Ranganathan, and G. R. Sambrano. 2002. Overview of the Alliance for Cellular Signaling. *Nature.* 420:703–706.

Goodman, O. B., Jr., J. G. Krupnick, F. Santini, V. V. Gurevich, R. B. Penn, A. W. Gagnon, J. H. Keen, and J. L. Benovic. 1998. Role of arrestins in G-protein-coupled receptor endocytosis. *Adv. Pharmacol.* 42:429–433.

Hartwell, L. H., J. J. Hopfield, S. Leibler, and A. W. Murray. 1999. From molecular to modular cell biology. *Nature.* 402:C47–C52.

Herrgard, M. J., M. W. Covert, and B. O. Palsson. 2004. Reconstruction of microbial transcriptional regulatory networks. *Curr Opin Biotechnol* 15.

Hoffmann, A., A. Levchenko, M. L. Scott, and D. Baltimore. 2002. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science.* 298:1241–1245.

Jain, N., T. Zhang, S. L. Fong, C. P. Lim, and X. Cao. 1998. Repression of Stat3 activity by activation of mitogen-activated protein kinase (MAPK). *Oncogene.* 17:3157–3167.

Jullien, J., V. Guili, L. F. Reichardt, and B. B. Rudkin. 2002. Molecular kinetics of nerve growth factor receptor trafficking and activation. *J. Biol. Chem.* 277:38700–38708.

Kalvakolanu, D. V. 2003. Alternate interferon signaling pathways. *Pharmacol. Ther.* 100:1–29.

Liao, J. C., S. Y. Hou, and Y. P. Chao. 1996. Pathway analysis, engineering and physiological considerations for redirecting central metabolism. *Biotech. Bioeng.* 52:129–140.

Lillemeier, B. F., M. Koster, and I. M. Kerr. 2001. STAT1 from the cell membrane to the DNA. *EMBO J.* 20:2508–2517.

McAdams, H. H., and A. Arkin. 1998. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* 27:199–224.

Milo, R., S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. Network motifs: simple building blocks of complex networks. *Science.* 298:824–827.

Papin, J. A., and B. O. Palsson. 2003. Topological analysis of mass-balanced signaling networks: a framework to obtain emergent properties including crosstalk. *J. Theor. Biol.* 227:283–297.

Papin, J. A., N. D. Price, J. S. Edwards, and B. O. Palsson. 2002a. The genome-scale metabolic extreme pathway structure in Haemophilus influenzae shows significant network redundancy. *J. Theor. Biol.* 215: 67–82.

Papin, J. A., N. D. Price, and B. O. Palsson. 2002b. Extreme pathway lengths and reaction participation in genome-scale metabolic networks. *Genome Res.* 12:1889–1900.

Papin, J. A., N. D. Price, S. J. Wiback, D. Fell, and B. O. Palsson. 2003. Metabolic pathways in the post-genome era. *Trends Biochem. Sci.* 28: 250–258.

Pelletier, S., F. Duhamel, P. Coulombe, M. R. Popoff, and S. Meloche. 2003. Rho family GTPases are required for activation of Jak/STAT signaling by G protein-coupled receptors. *Mol. Cell. Biol.* 23:1316–1333.

Pramanik, J., and J. D. Keasling. 1997. Stoichiometric model of Escherichia coli metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotech. Bioeng.* 56: 398–421.

Price, N. D., J. A. Papin, and B. O. Palsson. 2002. Determination of redundancy and systems properties of the metabolic network of Helicobacter pylori using genome-scale extreme pathway analysis. *Genome Res.* 12:760–769.

Price, N. D., J. A. Papin, C. H. Schilling, and B. O. Palsson. 2003. Genome-scale microbial in silico models: the constraints-based approach. *Trends Biotechnol.* 21:162–169.

Reed, J. L., and B. O. Palsson. 2003. Thirteen years of building constraint-based in silico models of Escherichia coli. *J. Bacteriol.* 185:2692–2699.

Rives, A. W., and T. Galitski. 2003. Modular organization of cellular networks. *Proc. Natl. Acad. Sci. USA.* 100:1128–1133.

Rivett, A. J. 1986. Regulation of intracellular protein turnover: covalent modification as a mechanism of marking proteins for degradation. *Curr. Top. Cell. Regul.* 28:291–337.

Rockafellar, R. T. 1970. Convex Analysis. Princeton University Press, Princeton. N.J.

Schilling, C. H., D. Letscher, and B. O. Palsson. 2000. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theor. Biol.* 203:229–248.

Schoeberl, B., C. Eichler-Jonsson, E. D. Gilles, and G. Muller. 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* 20:370–375.

Schuster, S., D. A. Fell, and T. Dandekar. 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* 18:326–332.

Schwartz, M. A., and V. Baron. 1999. Interactions between mitogenic stimuli, or, a thousand and one connections. *Curr. Opin. Cell Biol.* 11:197–202.

Selkov, E., Jr., Y. Grechkin, N. Mikhailova, and E. Selkov. 1998. MPW: the Metabolic Pathways Database. *Nucleic Acids Res.* 26:43–45.

Shuai, K., and B. Liu. 2003. Regulation of JAK-STAT signalling in the immune system. *Nat. Rev. Immunol.* 3:900–911.

Theurkauf, W. E. 1994. Premature microtubule-dependent cytoplasmic streaming in cappuccino and spire mutant oocytes. *Science.* 265:2093–2096.

Varma, A., and B. O. Palsson. 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Appl. Environ. Microbiol.* 60:3724–3731.

Vuong, T. M., and M. Chabre. 1991. Deactivation kinetics of the transduction cascade of vision. *Proc. Natl. Acad. Sci. USA.* 88:9813–9817.

Wiback, S. J., R. Mahadevan, and B. O. Palsson. 2003. Reconstructing metabolic flux vectors from extreme pathways: defining the alpha-spectrum. *J. Theor. Biol.* 224:313–324.

Wormald, S., and D. J. Hilton. 2004. Inhibitors of cytokine signal transduction. *J. Biol. Chem.* 279:821–824.

Yamada, S., S. Shiono, A. Joo, and A. Yoshimura. 2003. Control mechanism of JAK/STAT signal transduction pathway. *FEBS Lett.* 534:190–196.

Zubay, G. 1973. In vitro synthesis of protein in microbial systems. *Annu. Rev. Genet.* 7:267–287.