Available online at www.sciencedirect.com
SCIENCE DIRECT®
BioSystems 83 (2006) 152–166
BioSystems
www.elsevier.com/locate/biosystems

Trading the micro-world of combinatorial complexity for the macro-world of protein interaction domains

Nikolay M. Borisov [a], Nick I. Markevich [a,b], Jan B. Hoek [a], Boris N. Kholodenko [a,*]

[a] *Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, 1020 Locust St., Philadelphia, PA 19107, USA*
[b] *Institute of Theoretical and Experimental Biophysics, Pushchino, Moscow Region 142290, Russia*

Received 4 January 2005; received in revised form 16 March 2005; accepted 23 March 2005

Abstract

Membrane receptors and proteins involved in signal transduction display numerous binding domains and operate as molecular scaffolds generating a variety of parallel reactions and protein complexes. The resulting combinatorial explosion of the number of feasible chemical species and, hence, different states of a network greatly impedes mechanistic modeling of signaling systems. Here we present novel general principles and identify kinetic requirements that allow us to replace a mechanistic picture of all possible micro-states and transitions by a macro-description of states of separate binding sites of network proteins. This domain-oriented approach dramatically reduces computational models of cellular signaling networks by dissecting mechanistic trajectories into the dynamics of macro- and meso-variables. We specify the conditions when the temporal dynamics of micro-states can be exactly or approximately expressed in terms of the product of the relative concentrations of separate domains. We prove that our macro-modeling approach equally applies to signaling systems with low population levels, analyzed by stochastic rather than deterministic equations. Thus, our results greatly facilitate quantitative analysis and computational modeling of multi-protein signaling networks.

Keywords: Scaffold; Tyrosine kinase receptor; Adapter protein; Complex signaling networks; Time series; Model reduction

1. Introduction

Processing of extracellular signals involves covalent modification of amino acid residues on cell-surface receptors and cytoplasmic signaling proteins (Bray, 1998; Hunter, 2000). For instance, receptors that belong to the large family of receptor tyrosine kinases (RTK) modify tyrosine residues on proteins by attaching a phosphate group (Schlessinger, 2000). Tyrosine phosphorylation delivers a message for a plethora of binding partners, including adaptor proteins and effector enzymes, such as protein and lipid kinases and phosphatases (Pawson and Nash, 2003).

Receptors and various signaling proteins contain a number of different domains that display a multitude of phosphorylation states and generate a large variety of heterogeneous multi-protein complexes (Pawson and Scott, 1997). Even for a few initial steps in transduction of extracellular cues, a combinatorial variety of signaling processes involving receptors, scaffolds and adapters may generate hundreds of thousands of molecular species (Faeder et al., 2003; Hlavacek et al., 2003). For instance, following activation, insulin receptor (IR) and insulin-like growth factor receptor-1 (IGF-1R) bind various combinations of downstream targets, including

* Corresponding author. Tel.: +1 215 503 1614; fax: +1 215 923 2218.
E-mail address: Boris.Kholodenko@jefferson.edu (B.N. Kholodenko).

0303-2647/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2005.03.006

growth factor receptor binding protein-2 (Grb2), the Src homology and collagen domain protein (Shc), the p85 subunit of phosphatidylinositol 3-kinase (PI3K) and scaffolding adaptor proteins known as insulin receptor substrates, IRS1–IRS6 (Cai et al., 2003; Paz et al., 1996; Saltiel and Pessin, 2002; White, 2002). Likewise, receptor-mediated phosphorylation of multiple tyrosine residues on the IRS molecules creates docking sites for a variety of SH2 domain-containing proteins, such as Grb2, PI3K, soluble tyrosine kinases Src and Fyn, Ras-GAP and the protein tyrosine phosphatase SHP-2 (Cai et al., 2003; Myers et al., 1996; White, 1998). In addition, both IR and IGF-1R can bind and phosphorylate the scaffold adaptors known as Grb2-associated binders (GAB-1 and GAB-2) (Ingham et al., 1998; Lehr et al., 2000; Rodrigues et al., 2000). In turn, phosphorylated GABs bind numerous targets at their docking sites, including Shc, Grb2, p85, phospholipase Cγ, SHP-2 and the Crk adapter protein (Nishida and Hirano, 2003; Yamasaki et al., 2003).

Different docking sites initiate separate signaling pathways that propagate distinct cellular responses. For instance, a pathway initiated by the binding of Grb2 to tyrosine phosphorylated GAB and the recruitment of the GDP/GTP exchange factor SOS enables activation of the small GTPase Ras and leads to activation of the mitogen-activated protein kinase (MAPK) cascade, which promotes mitogenesis and differentiation. Independent binding of p85 to GAB initiates the PI3K/AKT pathway implicated in glucose and lipid metabolism and cell survival (Luo et al., 2003; Shepherd et al., 1998). Owing to the multiplication of all different possibilities, the number of feasible molecular species in the IR and IGF1-R pathways readily reaches hundreds of thousands and even a million, hampering a rigorous quantitative analysis and mathematical modeling of these signaling pathways.

An entire collection of potential molecular species that emerge as different forms of receptors, scaffold proteins and signaling enzymes will be referred to as a set of “micro-states”. In a framework of a standard mechanistic description, all these possible micro-states and their chemical transformations are taken into account (Faeder et al., 2003; Hlavacek et al., 2003; Levchenko et al., 2000). Because of exceedingly high numbers of micro-states, previous models of large signaling networks merely ignored a combinatorial variety of feasible state combinations, focusing on experimentally detected protein complexes (Asthagiri and Lauffenburger, 2001; Hatakeyama et al., 2003; Haugh et al., 1999, 2000; Kholodenko et al., 1999; Moehren et al., 2002; Schoeberl et al., 2002). The potential problem for network modeling that arises from the combinatorial complexity of signal transduction has been recognized (Kholodenko et al., 1999; Morton-Firth and Bray, 1998; Shimizu et al., 2000), and several approaches toward handling it have been suggested. An elegant algorithm to account for all potential species and reactions was developed and implemented in the rule-based software tool BioNetGen (Blinov et al., 2004). However, due to the enormous number of distinct chemical species and a lack of knowledge of the kinetics for each possible transformation, such a detailed “micro-description” rapidly becomes impractical. The structure of reduced models that concentrate on the predominant species is still difficult to predict, and this structure can be highly dependent on the values of kinetic parameters.

An appealing stochastic approach to modeling multistate signaling systems involved in bacterial chemotaxis was suggested by Bray and co-workers (Morton-Firth and Bray, 1998; Shimizu et al., 2000). In the computer program StochSim, individual multistate complexes present distinct software objects. Consequently, a combinatorial explosion of the number of micro-states is circumvented by following stochastic changes in the states of individual, distinguishable molecules, the number of which does not increase in the course of simulations (Le Novere and Shimizu, 2001). Still, for large networks of hundreds of different proteins, the StochSim calculation time would be exceedingly slow, increasing proportionally to the number of molecules squared.

Recently, a novel approach has been introduced that replaces a mechanistic picture of all possible states by a macro-model that analyzes the states of individual protein domains/sites, such as the phosphorylation levels and the fractions occupied by binding partners (Borisov et al., 2005; Conzelmann et al., 2005). In the present paper this approach is extended to include macro-models of receptors operating as scaffolds and stochastic simulations of signaling systems with low population numbers. A general modeling framework proposed here drastically reduces the number of states and differential equations to be solved and, therefore, the computational cost of both deterministic and stochastic simulations. We demonstrate that a necessary prerequisite for the reduction of a mechanistic model is the presence of protein domains/sites that do not influence other sites, allosterically or through interactions with bound partners. Importantly, the existence of additional sites involved in allosteric interactions does not impede the reduction of combinatorial complexity of multicomponent receptor-mediated signal transduction. We show when and how a mechanistic description of a signaling network can be expressed explicitly or