Available online at www.sciencedirect.com
SCIENCE DIRECT
BioSystems 83 (2006) 136–151
BioSystems
www.elsevier.com/locate/biosystems

A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity

Michael L. Blinov, James R. Faeder,
Byron Goldstein, William S. Hlavacek *

Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received 8 January 2005; received in revised form 6 May 2005; accepted 21 June 2005

Abstract

We consider a model of early events in signaling by the epidermal growth factor (EGF) receptor (EGFR). The model includes EGF, EGFR, the adapter proteins Grb2 and Shc, and the guanine nucleotide exchange factor Sos, which is activated through EGF-induced formation of EGFR–Grb2–Sos and EGFR–Shc–Grb2–Sos assemblies at the plasma membrane. The protein interactions involved in signaling can potentially generate a diversity of protein complexes and phosphoforms; however, this diversity has been largely ignored in models of EGFR signaling. Here, we develop a model that accounts more fully for potential molecular diversity by specifying rules for protein interactions and then using these rules to generate a reaction network that includes all chemical species and reactions implied by the protein interactions. We obtain a model that predicts the dynamics of 356 molecular species, which are connected through 3749 unidirectional reactions. This network model is compared with a previously developed model that includes only 18 chemical species but incorporates the same scope of protein interactions. The predictions of this model are reproduced by the network model, which also yields new predictions. For example, the network model predicts distinct temporal patterns of autophosphorylation for different tyrosine residues of EGFR. A comparison of the two models suggests experiments that could lead to mechanistic insights about competition among adapter proteins for EGFR binding sites and the role of EGFR monomers in signal transduction.

Keywords: Computational systems biology; Receptor tyrosine kinase; Protein complexes; Rule-based modeling; Automatic network generation; BioNetGen

1. Introduction

Processes by which a cell senses and responds to its environment are often marked by combinatorial complexity (Hlavacek et al., 2003). Cellular signaling (Hunter, 2000) generally involves protein–protein interactions and enzymatic activities that imply a diversity of potential protein complexes and phosphoforms, which are difficult to simply enumerate let alone assay or understand. For example, the number of possible phosphoforms of a protein is 2^n, where n is the number of amino acid residues that are subject to phosphorylation and dephosphorylation by kinases and phosphatases, at least nine tyrosines in the case of epidermal growth factor (EGF) receptor (EGFR) (Jorissen et al., 2003). Additional molecular diversity can arise from the multivalent character of protein–protein interactions. A protein

* Corresponding author. Present address: Mail Stop K710, LANL, Los Alamos, NM 87545, USA. Tel.: +1 505 665 1355; fax: +1 505 665 3493.
E-mail address: wish@lanl.gov (W.S. Hlavacek).

0303-2647/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2005.06.014

involved in signaling typically consists of multiple protein interaction domains (Pawson and Nash, 2003), such as the Src homology 2 (SH2) and 3 (SH3) domains of the Grb2 adapter protein. Each of these domains may interact with a partner that also contains multiple domains. As a result, proteins can combine in a variety of ways to form a spectrum of heterogeneous complexes. Proteomic studies confirm that diverse molecular species arise during signal transduction (Husi et al., 2000; Bunnell et al., 2002; Blagoev et al., 2003, 2004).

Given the protein–protein interactions and enzymatic activities involved in the cellular response to a signal, how do we catalog the potential molecular species implied by these interactions and activities? How do we predict which of the possible molecular species might actually arise during signaling? How do we determine the functional implications of these molecular species or the relative importance of processes that influence them? How can we best use large-scale proteomic measurements to obtain mechanistic insights? These questions are being asked in the emerging field of systems biology, and mathematical models have an important role to play in addressing such questions (Bhalla and Iyengar, 1999; Endy and Brent, 2001; Wiley et al., 2003; Hlavacek et al., 2003; Goldstein et al., 2004). A mathematical model requires an explicit statement of our understanding (or assumptions) about how a signal transduction system operates in a form that allows, through computational analysis, the behavior of the system to be predicted and compared with experimental observations. Here, we provide a demonstration of how a mathematical model, incorporating detail at the level of protein sites and domains, can be used to study signal transduction with a comprehensive treatment of protein complexes and phosphoforms implied by protein interactions.

We develop and analyze a mathematical model for early events in signaling by EGFR, which is a well-studied cell-surface receptor involved in cell proliferation (Schlessinger, 2000; Jorissen et al., 2003). It has been the subject of numerous model-based studies (Wiley et al., 2003). Our model, which we will call the network model, provides a description of EGFR signaling that accounts for the spectrum of molecular species (356) and the reactions among these species (3749) implied by specified interactions and activities of EGF, EGFR, the adapter proteins Grb2 and Shc, and the guanine nucleotide exchange factor Sos. These interactions and activities are the same as those considered in the seminal model of Kholodenko et al. (1999), which is based on assumptions (simultaneous phosphorylation and dephosphorylation of receptor tyrosines, inability of phosphorylated receptors in a dimer to dissociate, and competition among cytosolic proteins for receptor binding) that significantly limit, a priori, the number of molecular species that can arise during signaling. We will call the model of Kholodenko et al. (1999) the pathway-like model because it represents the signaling system as a set of reaction sequences rather than as a highly branched reaction network.

The rest of this report is organized as follows. First, we describe how the network model is constructed based on the proteins, interactions, and model parameters considered in the work of Kholodenko et al. (1999). Notably, the network model involves no more parameters than the pathway-like model. We then compare the predictions of the two models with the experimental observations of Kholodenko et al. (1999). We find both models are equally consistent. We also present new predictions of the network model and testable predictions that distinguish the two models. A comparison of the models allows us to evaluate the simplifying assumptions of Kholodenko et al. (1999). These assumptions have not been tested so far, even though this model has served as the starting point for a number of modeling studies of EGFR signaling (Schoeberl et al., 2002; Gong and Zhao, 2003; Hatakeyama et al., 2003; Resat et al., 2003; Conzelmann et al., 2004; Liu et al., 2005). We suggest experiments that could lead to insights into the mechanisms of signaling and determine which of the two models better represents signaling. Finally, we use the network model to predict the dynamics of the protein complexes and protein phosphorylation states that are generated during signaling. These predictions provide a picture of molecular diversity that is more detailed than could be currently obtained using the most sophisticated proteomic assays. For example, the model predicts which molecular species containing membrane-proximal Sos are prevalent at different time points. The model could also be used to predict how the population of these species depends on reaction dynamics and concentrations of components. As proteomic technologies mature, testing such predictions will become feasible.

2. The network model

2.1. Basis of the model

The network model (Fig. 1) is based on the same proteins, enzymatic activities and protein–protein interactions considered in the model of Kholodenko et al. (1999). The focus of this model is the cascade of signaling events that lead to recruitment of cytosolic Sos to the inner cell membrane (Fig. 1A and B), which can be described as follows. EGF binds to EGFR, which leads