Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-26 (2016) 312–317

On the personalised modelling of cancer signalling ⋆

Dirk Fey * **Axel Kuehn** ** **Boris N. Kholodenko** *

* *Systems Biology Ireland, School of Medicine, University College Dublin, Ireland (e-mails: dirk.fey@ucd.ie, boris.kholodenko@ucd.ie).*
** *Cancer Bioinformatics and systems biology, Institute of Cancer Research of Montpellier, France (e-mail: axel.kuehn@inserm.fr)*

Abstract: Dynamic modelling has long been used to understand fundamental principles of cell signalling and its dysregulation in cancer. More recently, these models have also been used to understand the individual risks of cancer patients, and predict their survival probabilities. However, the current methodologies for integrating tumour data and generating patient-specific simulations suffer from the lack of general applicability; they only work for cell signalling models in which only posttranslational protein modifications are considered, so that the total protein concentrations are conserved. Here, we present novel, generally applicable method. The method is based on a simple theoretical framework for modelling gene-regulation, and the indirect estimation of patient-specific parameters from tumour data. Because our method does not require time-invariance of the total-protein concentrations, it can be applied to models of any nature, including the many cancer signalling models involving gene-regulation.

Keywords: Systems biology, ordinary differential equations, parameter estimation.

1. INTRODUCTION

The dynamic states of cancer cell signalling are intimately linked to cancer cell fate decisions and clinical outcomes (Kolch et al., 2015; Fey et al., 2012). Recently, multiple studies have shown that patient-specific differences in the dynamic behaviour of these signalling networks underlie individual pathogenetic changes and disease manifestation (Fey et al., 2015; Flanagan et al., 2015; Lindner et al., 2013; Murphy et al., 2013). For example, a dynamic model of the JNK pathway could predict the survival probabilities of cancer patients in neuroblastoma, a common childhood cancer (Fey et al., 2015). These predictions were based on personalised simulations for over 700 patients, and revealed that a high amplitude, switch-like JNK activation was associated with neuroblastoma cell death, and better patient survival.

To generate the patient-specific simulations, the signalling model was personalised by adjusting the total protein concentration of each model component to the measured values from the patient's tumour sample. This approach works for purely posttranslational signalling models in which the total protein concentration of each gene is conserved (constant), and the only dynamic changes arise from the regulation of protein-protein interactions and enzyme activities, for example by phosphorylation.

In fact, all the personalised models mentioned are based on this simple principle of directly using the measured mRNA or protein concentrations as static parameters in the model (Fey et al., 2015; Flanagan et al., 2015; Lindner et al., 2013; Murphy et al., 2013). Thus, all these personalised models are based on the assumption that the total protein concentrations of the modelled genes are constant for all genes. Because only then do these measured concentrations constitute static parameters in the model that can readily be adjusted using the measured tumour data. Unfortunately, this assumption is often violated and prevents the application of this approach to systems in which the protein levels themselves are regulated.

Regulation of gene expression on the level of mRNA and total protein is quite common (Kolch et al., 2015; Nakakuki et al., 2010). Examples that are particularly relevant for cancer are regulation of the cell-cycle or the DNA-damage response. In these systems, both mRNA and total protein concentrations are regulated, and changing over time. Moreover, these dynamic changes are critical aspects of the systems behaviour. So far, the assumption of time-invariant total protein levels has hampered personalised modelling of these important processes. Thus, there is an urgent need for novel model-personalisation methodologies that are generally applicable.

Here we propose such a methodology.

In contrast to the current approach, the here proposed methodology does not require that the total-protein concentrations are constant. Instead, we propose a theoretical framework in which the patient-specific mRNA and total protein concentrations arise naturally from patient-specific mRNA synthesis parameters. Rather than time-invariant parameters, the mRNA and total protein concentrations are now dynamic states described by ordinary differential equations. This framework allows us to personalise models of any nature, involving signalling, gene-regulation or both, by indirectly estimating the patient-specific synthesis parameters.

⋆ Supported by EU FP7 grant "SynSignal" (No. 613879).

10.1016/j.ifacol.2016.12.145

The current manuscript is organised as follows. First, we will present a simple, yet generally applicable theoretical framework for the personalised modelling of gene expression. Second, we show how the patient-specific parameters can be estimated from expression data measured in the patients' tumour samples. Third, we illustrate the methodology by testing it on simulated data from a p53 DNA-damage response model.

2. A MODELLING FRAMEWORK FOR PERSONALISED GENE EXPRESSION

Although often considered separately, the posttranslational and gene-regulatory levels of cell signalling are intrinsically linked (Fig. 1). On the one hand, signalling on the protein level affects the gene-regulatory level, for instance by triggering mRNA synthesis trough the activation of transcription factors. On the other hand, the gene-regulatory level affects the signalling level by controlling the expression of the proteins that mediate and process the signals.

2.1 Combined modelling of gene-regulation and signalling

Mathematically, we can think of signalling as a system consisting of two layers, linked by feedback (Fig. 1). Each layer consists of a set of ordinary differential equations. The signalling layer has the concentrations of all post-translational protein forms as states, and takes the vector of mRNA concentrations as input. Vice versa, the gene-regulatory layer has the mRNA concentrations as states and takes the vector of protein concentrations as input.

Formally, we have the following system

$$\frac{d}{dt}y = f(y,p) + B\,\mathrm{diag}(k_{transl})x \tag{1a}$$

$$\frac{d}{dt}x = \rho + g(y,k) - \mathrm{diag}(k_{rdeg})x, \tag{1b}$$

where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ are vectors of mRNA and protein concentrations, respectively. $f(y,p)$ and $g(y,k)$ are functions describing posttranslational modifications and gene-regulatory effects, respectively. Note that $f(y,p)$ should contain a degradation term, i.e. $f(y,p) = \tilde{f}(y,p) - \mathrm{diag}(k_{pdeg})y$. $B \in \mathbb{R}^{n\times m}$ is a matrix containing m different unit vectors $e_i \in \mathbb{R}^n$, indicating the subset of unmodified proteins y_i that are translated (synthesised) from their corresponding mRNA templates. $\mathrm{diag}\,k_{transl} \in \mathbb{R}^{m\times m}$ and $\mathrm{diag}\,k_{rdeg} \in \mathbb{R}^{m\times m}$ are diagonal matrices containing the protein-translation and mRNA-degradation rate constants. $\rho \in \mathbb{R}^m$ is a vector of patient-specific parameters.

Remark on the notation. In the following, and for simplicity of presentation, we often do not explicitly denote the dependency of f and g on the kinetic parameters p and k. That is, when referring to f and g we mean the functions $f : y \mapsto f(y,p)$ and $g : y \mapsto g(y,p)$, respectively.

2.2 Rationale and assumptions behind the approach

The method is based on linking the measured patient-specific gene-expression differences on the systems level to parameter changes in the model on individual gene level. Because we want the patient-specific parameters to be identifiable, we assume that the basal mRNA-synthesis rate for each gene varies between patients, but all other parameters do not.

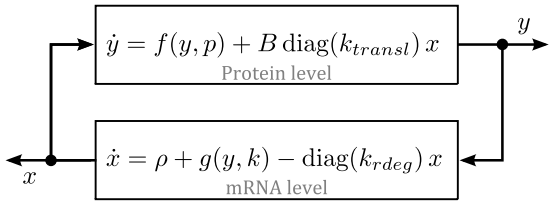

Fig. 1. The dynamic system used for combined modelling of posttranslational and gene-regulatory networks: y describes the protein level, and x the mRNA level.

Assumption 1. The observed gene-expression differences in tumour samples are caused by patient-specific differences of the basal mRNA-synthesis rates.

Although some other parameters, most notably mRNA-degradation and protein-translation rate constants, might also vary between patients, we only estimate the patient-specific mRNA-synthesis rates. All other kinetic parameters not directly relating to gene expression – such as catalytic rate constants and binding affinities – are presumed not to vary between patients.

Remark on violating Assumption 1. Although it is a theoretical requirement for formulating the methodology, violating Assumption 1 has limited practical consequences. As we will see later in Sec. 4.3, accurate, albeit less precise simulation results can be obtained even when the rates of mRNA-degradation and protein-translation for each gene differ between patients.

3. ESTIMATING THE PATIENT-SPECIFIC PARAMETERS

Our aim is to estimate the patient specific parameter ρ from gene expression measurements of the patients' tumours. Such tumour data are usually obtained from tissue samples collected in biopsies or surgery, and would usually reflect the homeostatic, unperturbed state of the patient's tumour. Thus, these tumour data would correspond to the unstimulated steady-state of our model (1).

Assumption 2. Measured gene expression values from tumour samples reflect basal (unstimulated) steady-state values.

This is a fair assumption considering that most tumour data are collected prior to pharmacological treatment (de Gramont et al., 2015; Juty et al., 2015).

Put more formally, the goal is to estimate ρ from steady state measurements of either x or y, based on the model in (1) with given functions f, g and given kinetic parameters p, k, k_{transl} and k_{rdeg}.

For the steady-state we have

$$0 = f(y,p) + B\,\mathrm{diag}(k_{transl})x, \tag{2a}$$

$$0 = \rho + g(y,k) - \mathrm{diag}(k_{rdeg})x. \tag{2b}$$

The next step is to solve (2) for the patient-specific parameters ρ. Two scenarios can be distinguished.