Eugen Czeizler | Vladimir Rogojin | Ion Petre

The phosphorylation of the heat shock factor as a modulator for the heat shock response

Turku Centre *for* Computer Science

TUCS Technical Report
No 1041, April 2012

The phosphorylation of the heat shock factor as a modulator for the heat shock response

Eugen Czeizler
Department of Information and Computer Science
Aalto University School of Science, FI-00076 Aalto, Finland
eugen.czeizler@aalto.fi
Vladimir Rogojin
Faculty of Medicine, University of Helsinki
FI-00014 Helsinki, Finland
vladimir.rogojin@helsinki.fi
Ion Petre
Computational Biomodelling Laboratory, Åbo Akademi University
and Turku Centre for Computer Science
FI-20520 Turku, Finland
ipetre@abo.fi

TUCS Technical Report
No 1041, April 2012

Abstract

The heat shock response is a well conserved defence mechanism against the accumulation of misfolded proteins due to prolonged elevated heat. The cell responds to heat shock by raising the levels of heat shock proteins (hsp), which are responsible for chaperoning protein refolding. The synthesis of hsp is highly regulated at the transcription level by specific heat shock (transcription) factors (hsf). One of the regulation mechanisms is the phosphorylation of hsf's. Experimental evidence shows a connection between the hyper-phosphorylation of hsf's and the transactivation of the hsp-encoding genes. In this paper we incorporate several (de)phosphorylation pathways into an existing well validated computational model of the heat shock response. We analyze the quantitative control of each of these pathways over the entire process. For each of these pathways we create detailed computational models which we subject to parameter estimation in order to fit them to existing experimental data. In particular, we find conclusive evidence supporting only one of the analyzed pathways. Also, we corroborate our results with a set of computational models of a more reduced size.

Keywords: Model refinement, quantitative analysis, heat shock response, phosphorylation.

TUCS Laboratory
Computational Biomodelling Laboratory

1 Introduction

The heat shock response (HSR) is a well conserved cellular regulatory mechanism dealing with increased accumulations of misfolded proteins due to elevated temperature or other forms of stress, such as physical, environmental, or chemical [34, 21]. Under elevated temperatures proteins start misfolding and forming enlarging cytotoxic aggregates [1]. Such bulk complexes may have very harmful implications for the cell and the organism; it is known for example that the generation of protein aggregates is associated with several neurodegenerative diseases such as Alzheimer and Parkinson [30, 41]. The cell's response is given by an immediate increase of the level of chaperons, also called heat shock proteins (hsp) which bind to the misfolded proteins and assist them in reacquiring their functional fold. The synthesis of hsp is highly regulated at the transcription level, by use of transcription factors. Once the stress disappears and the level of misfolded proteins returns to normal, the entire HSR mechanism is switched-off, and the cell returns to its physiological level of chaperons, see [23, 35, 40]. One of the most remarkable aspects about HSR is how evolutionary well conserved this mechanism is, as it is present in almost all living cells. Some of the species lacking the classical HSR can be found in the Antarctic ocean such as the *Odontaster validus* sea star, the *Euplotes focardii* ciliate ([25]), the *Trematomus bernacchii* notothenioid fish ([16]), and the *Paraceradocus gibber* gammarid ([7]). However, the relevance of this regulatory mechanism goes beyond HSR, as heat shock proteins are associated and play a key role also in many other cellular signaling and regulatory processes [22, 39]. Due to their contribution to the resilience of cancer cells, hsp also became a cancer treatment target [6, 44].

It is known that hsp transcription is regulated also through the phosphorylation of its transcription factors (hsf) [43, 26, 42]. While some of the hsf phosphorylation sites are constitutively phosphorylated and have an active role in modulating the repression of the transcriptional activity [23], other sites, such as *Serine 230* of mammalian HSF1, become phosphorylated only during the heat shock, and have a demonstrated role in promoting the inducible transcriptional activity of hsf [17]. In particular, it has been experimentally observed in [23] that upon heat shock, the overall level of hsf phosphorylation exhibits a clear transient behavior, which follows closely (while slightly delayed) the DNA binding measurements. However, little to nothing has been made until now for trying to understand the way in which the phosphorylation process acts as a control mechanism over the hsp transcription process. In [40], the phosphorylation of hsf is included into an HSR molecular model, but merely as an intermediate step of the process, without giving much insight into the type of control it actually enables.

In this study we performed a detailed analysis of several phosphorylation-driven control mechanisms for the regulation of the HSR, namely for the regulation of the transcription process of new hsp. Recently, a simple and well validated molecular and computational model for the HSR has been proposed in [37, 36]. Because of its reduced size, this model was a good base for an extension which incorporates also the hsf's phosphorylation level. Using this extended model, we investigated three different phosphorylation pathways to uncover their control over the HSR process. For each of these pathways we created detailed computational models which we subjected to parameter estimation in order to fit them to existing experimental data. Our results show that only one of these pathways is able to explain the experimental data. However, for one of the remaining pathways, our results suggest that although it is not

able to explain the experimental data by itself, it may have a role in fine-tuning the control over the transactivation of the hsp-encoding genes.

2 Models

2.1 The heat shock response

Throughout this study, we used as reference the model recently introduced in [37] and [36] for the eukaryotic HSR. Upon elevated temperatures, proteins (denoted in the model as prot) start to misfold. In response, the cell quickly raises the level of a special type of chaperons, called heat shock proteins (hsp). These chaperons, sequestrate the misfolded proteins (mfp) and help them in reassuming their functional fold. The control over the hsp synthesis is regulated at the gene level through some specialized transcription factors called heat shock factors (hsf). Upon heat shock, the hsf monomers migrate into the nucleus, and form dimers (hsf_2) and then trimers (hsf_3). These trimers are able to bind to the heat shock elements (hse), which are specific DNA sequences upstream of the hsp-encoding gene. The hsf trimers then facilitate the transcription of the gene and the eventual synthesis of new hsp.

When the level of hsp is high enough, a feedback mechanism ensures the down-regulation of their further synthesis. This mechanism consists of the sequestration of hsf by hsp, breaking in this way all the complexes in which hsf takes part. In particular, hsp breaks also the hsf_3: hse complex, unbinding the hsf monomers from the DNA and effectively stopping in this way the transcription of the hsp-encoding gene. The reactions of the model are presented in Table 1, while the numerical setup of the associated computational model, i.e. the initial concentrations of the reactants and the values of the kinetic parameters (as reported in [37] and [36]), are given in Table 6, see the Appendix in the supplementary information. We assumed that the only temperature-dependent reaction is the protein misfolding, i.e., reaction 9 from Table 1. In particular, we assumed that all the other kinetic constants, except for the protein misfolding rate, are unchanged within the temperature range $37 - 42^{\circ}C$, avoiding more complicated considerations based on the Arrhenius equation [29]. This assumption is in line with the way the kinetic constants were deduced in [36] (by fitting to heat shock data at $42^{\circ}C$, while simultaneously constraining the steady-state behavior at $37^{\circ}C$). It is also in line with our expectation that the massive changes in the protein misfolding rate caused by a $5^{\circ}C$ increase in temperature and their effects on the system dynamics, far exceed those incurred by very small changes in the kinetic constants of the other reactions.

The exponentially temperature-dependent rate law of the protein misfolding reaction, ϕ_T, was deduced in [37, 36], by adapting a similar law from [28], based on experimental studies of [27]:

$$\phi_T = \left(1 - \frac{0.4}{e^{T-37}}\right) \times 1.4^{T-37} \times 1.45 \times 10^{-5} \quad s^{-1},$$

where T is the temperature of the environment in Celsius degrees. According to [27], this law is valid for temperatures between 37 and 45°C.

Table 1: The molecular model for the eukaryotic heat shock response proposed in [36].

Reaction		Reaction	
$2\,\mathsf{hsf} \leftrightarrows \mathsf{hsf}_2$	(1)	$\mathsf{hsp} + \mathsf{hsf}_3\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf} + \mathsf{hse}$	(7)
$\mathsf{hsf} + \mathsf{hsf}_2 \leftrightarrows \mathsf{hsf}_3$	(2)	$\mathsf{hsp} \rightarrow \emptyset$	(8)
$\mathsf{hsf}_3 + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3\colon \mathsf{hse}$	(3)	$\mathsf{prot} \rightarrow \mathsf{mfp}$	(9)
$\mathsf{hsp} + \mathsf{hsf} \leftrightarrows \mathsf{hsp}\colon \mathsf{hsf}$	(4)	$\mathsf{hsp} + \mathsf{mfp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mfp}$	(10)
$\mathsf{hsp} + \mathsf{hsf}_2 \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{hsf}$	(5)	$\mathsf{hsp}\colon \mathsf{mfp} \rightarrow \mathsf{hsp} + \mathsf{prot}$	(11)
$\mathsf{hsp} + \mathsf{hsf}_3 \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf}$	(6)	$\mathsf{hsf}_3\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3\colon \mathsf{hse} + \mathsf{hsp}$	(12)

2.2 Phosphorylation-induced activation of the heat shock factors

The hsf proteins have several phosphorylation sites, some of which are known to significantly influence their activity, see e.g. [15, 23, 18]. While some of these sites are constitutively phosphorylated, others become phosphorylated only as a result of the heat shock, see [17, 38]. In particular, Serine 230 is a site that is hyperphosphorylated in reaction to stress, which has a positive effect on the hsf transactivation capacity, as shown in [17]. Our models for the phosphorylation of the heat shock factor only keep track of one single site per monomer (two sites in hsf dimers, three in hsf trimers), focusing on the role played by phosphorylation at S230. This choice also allows us to generate a series of tractable computational models. Such a computational modeling approach is typical for these situations, see e.g. [5, 40], as modeling all phosphorylation sites would generate an untractable combinatorial explosion in the number of different species and reactions. (Note that modeling explicitly k phosphorylation sites for each hsf would generate $O(2^{3k})$ new species in the model due to the hsf dimer and trimer species.) As a result, each hsf in our model can be either phosphorylated (hsf^p), or un-phosphorylated (hsf). An un-phosphorylated hsf can be phosphorylated only as a result of a catalytic reaction, involving a kinase (ki) enzyme. This is a double step reaction, where the first step is the (reversible) binding of hsf to the kinase ($\mathsf{hsf}\colon \mathsf{ki}$) and the second (irreversible) step is the phosphorylation of hsf and the unbinding of ki. If hsf is in a dimer or trimer complex, then each hsf from this complex becomes phosphorylated as a result of a separate kinase-mediated phosphorylation reaction. We marked multiple phosphorylated complexes by writing in the superscript a corresponding number of ^{p} symbols.) For example, in the case of dimer complexes hsf_2, we considered the following phosphorylation reactions:

$$\mathsf{hsf}_2 + \mathsf{ki} \rightleftarrows \mathsf{hsf}_2\colon \mathsf{ki} \rightarrow \mathsf{hsf}_2^p + \mathsf{ki}$$
$$\mathsf{hsf}_2^p + \mathsf{ki} \rightleftarrows \mathsf{hsf}_2^p : \mathsf{ki} \rightarrow \mathsf{hsf}_2^{pp} + \mathsf{ki}\,.$$

The dephosphorylation process is very similar to the phosphorylation one, only that the catalytic reaction involves in this case a phosphatase (ph) enzyme; e.g., for the case of hsf_2^{pp} species we considered the following dephosphorylation reactions:

$$\mathsf{hsf}_2^{pp} + \mathsf{ph} \rightleftarrows \mathsf{hsf}_2^{pp} : \mathsf{ph} \rightarrow \mathsf{hsf}_2^p + \mathsf{ph}$$
$$\mathsf{hsf}_2^p + \mathsf{ph} \rightleftarrows \mathsf{hsf}_2^p : \mathsf{ph} \rightarrow \mathsf{hsf}_2 + \mathsf{ph}\,.$$

Experimental observations [17, 38], suggest that the hyper-phosphorylation of the $\mathsf{hsf}_3\colon \mathsf{hse}$ complexes acts as an up-regulator for the transcription of new hsp. In our analysis, we investigated two possible mechanisms for modeling the

phosphorylation-dependent transcription.

In the first case, we assumed that the transcription proceeds linearly depending on the level of phosphorylation of the $\mathsf{hsf_3{:}\,hse}$ complex. Since each hsf has only one phosphorylation site, the complex $\mathsf{hsf_3{:}\,hse}$ has four phosphorylation states: not phosphorylated, or phosphorylated on one, two, or thee of the hsfs, respectively. Thus we replaced reaction (12) from the reference model with the following four reactions:

$$\mathsf{hsf_3{:}\,hse} \to \mathsf{hsf_3{:}\,hse} + \mathsf{hsp} \qquad \mathsf{hsf}_3^{pp} : \mathsf{hse} \to \mathsf{hsf}_3^{pp} : \mathsf{hse} + \mathsf{hsp}$$
$$\mathsf{hsf}_3^{p} : \mathsf{hse} \to \mathsf{hsf}_3^{p} : \mathsf{hse} + \mathsf{hsp} \qquad \mathsf{hsf}_3^{ppp} : \mathsf{hse} \to \mathsf{hsf}_3^{ppp} : \mathsf{hse} + \mathsf{hsp}\,.$$

We assumed that the speed of these reactions, i.e. the corresponding kinetic rate constants, depends linearly on the phosphorylation state of the reactant.

In the second case, we assumed that hsp synthesis is activated only by the hyper-phosphorylation of the $\mathsf{hsf_3{:}\,hse}$ compound. Namely, the transcription of the hsp encoding gene proceeds only after at least two of the hsfs from this compound become phosphorylated. Thus, we replaced reaction (12) from the reference model by:

$$\mathsf{hsf}_3^{pp} : \mathsf{hse} \to \mathsf{hsf}_3^{pp} : \mathsf{hse} + \mathsf{hsp} \qquad \mathsf{hsf}_3^{ppp} : \mathsf{hse} \to \mathsf{hsf}_3^{ppp} : \mathsf{hse} + \mathsf{hsp}\,.$$

By tuning the speed of the (de)phosphorylation reactions and of the hsp synthesis reactions, we employed an additional control over the production of new hsp proteins.

2.3 Mathematical models

Our computational models are described in terms of ordinary differential equations (ODEs), and they were derived based on the principle of mass-action kinetics. To each species we associated a time-dependent continuous variable representing its concentration level. For each variable, depending on the reactions in which the associated reactant is involved, the differential equation gives the cumulated production and consumption rates. According to the principle of mass-action, the rate of a reaction is proportional to the concentration of the involved reactants, see [11]. In some of the cases we simplify our analysis by considering Michaelis-Menten kinetics to reduce the number of molecular species considered in the model; when using this kinetics, the ki and ph enzymes are not modeled explicitly anymore, and thus the computational model is significantly reduced.

In our analysis, we considered several possible mechanisms which could explain the experimentally observed correlation between phosphorylation and the DNA binding levels. Each of these mechanisms was embedded into the extended HSR model. Moreover, we derived appropriate ODE models, which we subjected to parameter estimation procedures. All of our considered models are freely available at [9], using the COPASI [19] computational environment.

2.4 Experimental data

We fitted our computational models using two sets of experimental data of [23]. The first is a set of experimental measurements regarding the DNA binding activity of HeLa cells, while subjected to a continuous $42°C$ heat shock for 4 hours. The second set reports the total number of phosphorylated hsf proteins during the heat shock, i.e. also at 42°C. As a final numerical prerequisite, we required that the initial values of all species form a steady state of the model for a temperature of 37°C. This condition was justified by the fact that the system is supposed to be in a steady state in the absence of the heat shock, see [37, 36].

3 Methods

3.1 Quantitative model refinement of biomodels

The concept of quantitative model refinement was discussed in [10, 14] for rule-based modeling and demonstrated on two larger case studies based on ODEs in [8, 32]. The idea is to build a large, detailed computational model by starting from a generic (and smaller) model, and subsequently adding more details about its reactants and/or reactions. The aim is to perform the refinement in a step-wise quantitatively correct way, to preserve all previously obtained systemic properties of the generic model, such as the numerical fit and validation. In particular, a data-refinement mechanism was considered in [8, 32], where a model was refined by substituting a given species by several types of subspecies, and the refined model was algorithmically constructed from the original one. Note that, once a species was substituted with several subspecies, all the reactions in which the initial species was involved became also subject to refinement.

In the case of the HSR model analyzed here, we started from the basic model of [36] where hsf has no post-translational modifications, and we extended this model to the case where hsf, whether in a monomer or inside a complex, can be phosphorylated. Moreover, since the basic model was thoroughly fitted and validated to experimental data in [36], we aimed to preserve as much as possible from that model's kinetic rate parameters.

For each species that is refined by subspecies substitution, all the reactions involving the initial species either as a reactant or as a product also had to be refined. For instance, reaction (1) from the reference model (i.e., $\mathsf{hsf} + \mathsf{hsf} \rightleftarrows \mathsf{hsf}_2$) was replaced in the extended model by the following reactions:

$$\mathsf{hsf} + \mathsf{hsf} \rightleftarrows \mathsf{hsf}_2 \qquad \mathsf{hsf}^p + \mathsf{hsf} \rightleftarrows \mathsf{hsf}_2^p \qquad \mathsf{hsf}^p + \mathsf{hsf}^p \rightleftarrows \mathsf{hsf}_2^{pp} .$$

We proceeded similarly with all reactions involving hsf or complexes containing hsf. The final refined model is given in Table 2. At the end of this step we had a model detailing the phosphorylation status of hsf, hsf_2, hsf_3, $\mathsf{hsf_3{:}\,hse}$ and $\mathsf{hsp{:}\,hsf}$. It is important to note however that the model does not include details about how any of these species can be (de)phosphorylated through interactions with kinases and phosphatases, nor does it detail the effect of the heat shock on these enzymes. These non-refinement extensions were added in the second step of the model extension, see Section 4 and their kinetic constants were estimated through model fit algorithms, see Section 3.2. The reactions added in the non-refinement extension are given in Table 3.

Since the final model contains a large number of kinetic constants, our aim was to derive as many of them as possible from the basic model rather than to run parameter estimation for all of them, a computationally challenging problem. To do this, we required that the overall behavior of the refined model, i.e. the model obtained after the first step of the extension, did not change in the following sense: for each variable X of the basic model, we required that the sum of all variables replacing X in the refined model coincided with X as a function of time. In this way, we made sure that the numerical fit of the basic model (expressed without the phosphorylation details) was preserved in the refined model. We also ensured in this way that all mass conservation relations of the basic model were preserved in the refined model. In this formulation, we essentially needed to solve a system of functional relations (expressed through ODEs), one for each variable of the basic model, where the unknowns were the kinetic rate constants of the refined model. We discuss the data refinement technique in more details in the supplementary information. We also refer for

Table 2: The molecular model for the eukaryotic heat shock response, refined with the details of the phosphorylation status of hsf.

Reaction	
$2\,\mathsf{hsf} \leftrightarrows \mathsf{hsf}_2$	(1')
$\mathsf{hsf}^p + \mathsf{hsf} \leftrightarrows \mathsf{hsf}_2^p$	(2')
$2\,\mathsf{hsf}^p \leftrightarrows \mathsf{hsf}_2^{pp}$	(3')
$\mathsf{hsf} + \mathsf{hsf}_2 \leftrightarrows \mathsf{hsf}_3$	(4')
$\mathsf{hsf}^p + \mathsf{hsf}_2 \leftrightarrows \mathsf{hsf}_3^p$	(5')
$\mathsf{hsf} + \mathsf{hsf}_2^p \leftrightarrows \mathsf{hsf}_3^p$	(6')
$\mathsf{hsf}^p + \mathsf{hsf}_2^p \leftrightarrows \mathsf{hsf}_3^{pp}$	(7')
$\mathsf{hsf} + \mathsf{hsf}_2^{pp} \leftrightarrows \mathsf{hsf}_3^{pp}$	(8')
$\mathsf{hsf}^p + \mathsf{hsf}_2^{pp} \leftrightarrows \mathsf{hsf}_3^{ppp}$	(9')
$\mathsf{hsf}_3 + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3\colon \mathsf{hse}$	(10')
$\mathsf{hsf}_3^p + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse}$	(11')
$\mathsf{hsf}_3^{pp} + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse}$	(12')
$\mathsf{hsf}_3^{ppp} + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse}$	(13')
$\mathsf{hsp} + \mathsf{hsf} \leftrightarrows \mathsf{hsp}\colon \mathsf{hsf}$	(14')
$\mathsf{hsp} + \mathsf{hsf}^p \leftrightarrows \mathsf{hsp}\colon \mathsf{hsf}^p$	(15')
$\mathsf{hsp} + \mathsf{hsf}_2 \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{hsf}$	(16')
$\mathsf{hsp} + \mathsf{hsf}_2^p \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{hsf}^p$	(17')
$\mathsf{hsp} + \mathsf{hsf}_2^p \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{hsf}$	(18')
$\mathsf{hsp} + \mathsf{hsf}_2^{pp} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{hsf}^p$	(19')
$\mathsf{hsp} + \mathsf{hsf}_3 \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf}$	(20')

Reaction	
$\mathsf{hsp} + \mathsf{hsf}_3^p \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + 2\,\mathsf{hsf}$	(21')
$\mathsf{hsp} + \mathsf{hsf}_3^p \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{hsf}^p + \mathsf{hsf}$	(22')
$\mathsf{hsp} + \mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf}^p$	(23')
$\mathsf{hsp} + \mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{hsf}^p + \mathsf{hsf}$	(24')
$\mathsf{hsp} + \mathsf{hsf}_3^{ppp} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + 2\,\mathsf{hsf}^p$	(25')
$\mathsf{hsp} + \mathsf{hsf}_3\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf} + \mathsf{hse}$	(26')
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + 2\,\mathsf{hsf} + \mathsf{hse}$	(27')
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{hsf}^p + \mathsf{hsf} + \mathsf{hse}$	(28')
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{hsf}^p + \mathsf{hsf} + \mathsf{hse}$	(29')
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + 2\,\mathsf{hsf}^p + \mathsf{hse}$	(30')
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + 2\,\mathsf{hsf}^p + \mathsf{hse}$	(31')
$\mathsf{hsp} \rightarrow \emptyset$	(32')
$\mathsf{prot} \rightarrow \mathsf{mfp}$	(33')
$\mathsf{hsp} + \mathsf{mfp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mfp}$	(34')
$\mathsf{hsp}\colon \mathsf{mfp} \rightarrow \mathsf{hsp} + \mathsf{prot}$	(35')
$\mathsf{hsf}_3\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3\colon \mathsf{hse} + \mathsf{hsp}$	(36')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} + \mathsf{hsp}$	(37')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} + \mathsf{hsp}$	(38')
$\mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse} + \mathsf{hsp}$	(39')

Table 3: The model for the (de)phosphorylation of hsf and of complexes containing hsf

Reaction	**Kinetic rate**	
$\mathsf{hsf} + \mathsf{ki} \leftrightarrows \mathsf{hsf}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(40')
$\mathsf{hsf}\colon \mathsf{ki} \rightarrow \mathsf{hsf}^p + \mathsf{ki}$	k'_{17}	(41')
$\mathsf{hsf}^p + \mathsf{ph} \leftrightarrows \mathsf{hsf}^\mathsf{p}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(42')
$\mathsf{hsf}^\mathsf{p}\colon \mathsf{ph} \rightarrow \mathsf{hsf} + \mathsf{ph}$	k'_{19}	(43')
$\mathsf{hsf}_2 + \mathsf{ki} \leftrightarrows \mathsf{hsf}_2\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(44')
$\mathsf{hsf}_2\colon \mathsf{ki} \rightarrow \mathsf{hsf}_2^p + \mathsf{ki}$	k'_{17}	(45')
$\mathsf{hsf}_2^p + \mathsf{ki} \leftrightarrows \mathsf{hsf}_2^\mathsf{p}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(46')
$\mathsf{hsf}_2^\mathsf{p}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_2^{pp} + \mathsf{ki}$	k'_{17}	(47')
$\mathsf{hsf}_2^{pp} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_2^\mathsf{pp}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(48')
$\mathsf{hsf}_2^\mathsf{pp}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_2^p + \mathsf{ph}$	k'_{19}	(49')
$\mathsf{hsf}_2^p + \mathsf{ph} \leftrightarrows \mathsf{hsf}_2^\mathsf{p}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(50')
$\mathsf{hsf}_2^\mathsf{p}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_2 + \mathsf{ph}$	k'_{19}	(51')
$\mathsf{hsf}_3 + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(52')
$\mathsf{hsf}_3\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^p + \mathsf{ki}$	k'_{17}	(53')
$\mathsf{hsf}_3^p + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(54')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^{pp} + \mathsf{ki}$	k'_{17}	(55')
$\mathsf{hsf}_3^{pp} + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(56')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^{ppp} + \mathsf{ki}$	k'_{17}	(57')
$\mathsf{hsf}_3^{ppp} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(58')
$\mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3^{pp} + \mathsf{ph}$	k'_{19}	(59')
$\mathsf{hsf}_3^{pp} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(60')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3^p + \mathsf{ph}$	k'_{19}	(61')
$\mathsf{hsf}_3^p + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(62')

Reaction	**Kinetic rate**	
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3 + \mathsf{ph}$	k'_{19}	(63')
$\mathsf{hsp}\colon \mathsf{hsf} + \mathsf{ki} \leftrightarrows \mathsf{hsp}\colon \mathsf{hsf}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(64')
$\mathsf{hsp}\colon \mathsf{hsf}\colon \mathsf{ki} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{ki}$	k'_{17}	(65')
$\mathsf{hsp}\colon \mathsf{hsf}^p + \mathsf{ph} \leftrightarrows \mathsf{hsp}\colon \mathsf{hsf}^\mathsf{p}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(66')
$\mathsf{hsp}\colon \mathsf{hsf}^\mathsf{p}\colon \mathsf{ph} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{ph}$	k'_{19}	(67')
$\mathsf{hsf}_3\colon \mathsf{hse} + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3\colon \mathsf{hse}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(68')
$\mathsf{hsf}_3\colon \mathsf{hse}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} + \mathsf{ki}$	k'_{17}	(69')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(70')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} + \mathsf{ki}$	k'_{17}	(71')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} + \mathsf{ki} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse}\colon \mathsf{ki}$	k'^{+}_{16}, k'^{-}_{16}	(72')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse}\colon \mathsf{ki} \rightarrow \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse} + \mathsf{ki}$	k'_{17}	(73')
$\mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(74')
$\mathsf{hsf}_3^\mathsf{ppp}\colon \mathsf{hse}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} + \mathsf{ph}$	k'_{19}	(75')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(76')
$\mathsf{hsf}_3^\mathsf{pp}\colon \mathsf{hse}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} + \mathsf{ph}$	k'_{19}	(77')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse} + \mathsf{ph} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse}\colon \mathsf{ph}$	k'^{+}_{18}, k'^{-}_{18}	(78')
$\mathsf{hsf}_3^\mathsf{p}\colon \mathsf{hse}\colon \mathsf{ph} \rightarrow \mathsf{hsf}_3\colon \mathsf{hse} + \mathsf{ph}$	k'_{19}	(79')
$\mathsf{ki} \rightarrow \mathsf{mki}$	$k'_{20} \cdot \phi_T$	(80')
$\mathsf{mki} + \mathsf{hsp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mki}$	k'^{+}_{21}, k'^{-}_{21}	(81')
$\mathsf{hsp}\colon \mathsf{mki} \rightarrow \mathsf{ki} + \mathsf{hsp}$	k'_{22}	(82')
$\mathsf{ph} \rightarrow \mathsf{mph}$	$k'_{23} \cdot \phi_T$	(83')
$\mathsf{mph} + \mathsf{hsp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mph}$	k'^{+}_{24}, k'^{-}_{24}	(84')
$\mathsf{hsp}\colon \mathsf{mph} \rightarrow \mathsf{ph} + \mathsf{hsp}$	k'_{25}	(85')

more details to [8], where the HSR model was extended to model the effects of hsf acetylation on the control of the heat shock response. Similarly as in the case of the present study, that extension consisted of a refinement step where each hsf is replaced with two versions depending on whether it is acetylated (on a specific) site or not, plus a non-refinement step detailing the control of the acetylation on the heat shock response.

It is clear that the problem of deducing the kinetic constant of the refined model based on the refinement conditions formulated above is not unique. For example, the problem admits the trivial solution where the constants of all reactions involving at least one phosphorylated species are set to 0, thus effectively reducing the refined model to the basic model, see also the discussion in the supplementary information. Since the systems of ODEs corresponding to the original and to the refined model are in general impossible to solve analytically, we adopt a symbolic approach similar to that in [8] where we set the kinetic rate constants of the refined model as in Table 4. This particular choice of constants is mathematically elegant since it is symmetrical with respect to the phosphorylation status of the reactants: for example, a reaction involving fully phosphorylated reactants have the same constant as the similar reaction involving the same reactants being non-phosphorylated. We refer to [8] for a mathematical proof that the choice of constants as in Table 4 yields a refined model that satisfies the refinement conditions formulated above. We also refer to [32] for a different study, on protein self-assembly, based on the same technique of quantitative refinement. In the absence of biological knowledge regarding the kinetic parameters of the refined model, our solution is only driven by the symbolic approach described above. If there is biological knowledge about some of the values of the parameters of the refined model, then such knowledge can be taken into consideration in the form of constraints in our symbolic approach. The existence of a solution in such a case, where some of the refined reactions' kinetic constant are know, as well as effectively constructing one, appears as an interesting open mathematical problem in this context.

Table 4: The kinetic rate constants of the extended HSR model with respect to corresponding rate constants of the reference model from Table 1. In this setup the extended HSR model preserves the numerical fit of the reference model.

React.	**Kin. const.**	**React.**	**Kin. const.**	**React.**	**Kin. const.**
$(1')$	k_1^+, k_1^-	$(14')$	k_4^+, k_4^-	$(27')$	$1/2 \cdot k_7$
$(2')$	$2 \cdot k_1^+, k_1^-$	$(15')$	k_4^+, k_4^-	$(28')$	$1/2 \cdot k_7$
$(3')$	k_1^+, k_1^-	$(16')$	k_5	$(29')$	$1/2 \cdot k_7$
$(4')$	k_2^+, k_2^-	$(17')$	$1/2 \cdot k_5$	$(30')$	$1/2 \cdot k_7$
$(5')$	$k_2^+, 1/2 \cdot k_2^-$	$(18')$	$1/2 \cdot k_5$	$(31')$	k_7
$(6')$	$k_2^+, 1/2 \cdot k_2^-$	$(19')$	k_5	$(32')$	k_8
$(7')$	$k_2^+, 1/2 \cdot k_2^-$	$(20')$	k_6	$(33')$	Φ_T
$(8')$	$k_2^+, 1/2 \cdot k_2^-$	$(21')$	$1/2 \cdot k_6$	$(34')$	k_{10}^+, k_{10}^-
$(9')$	k_2^+, k_2^-	$(22')$	$1/2 \cdot k_6$	$(35')$	k_{11}
$(10')$	k_3^+, k_3^-	$(23')$	$1/2 \cdot k_6$	$(36')$	k_{12}
$(11')$	k_3^+, k_3^-	$(24')$	$1/2 \cdot k_6$	$(37')$	k_{12}
$(12')$	k_3^+, k_3^-	$(25')$	k_6	$(38')$	k_{12}
$(13')$	k_3^+, k_3^-	$(26')$	k_7	$(39')$	k_{12}

3.2 Parameter estimation

The problem of estimating the parameters of kinetic models in systems biology is difficult, see e.g., [3, 31, 33]. It can be formulated as a mathematical optimization problem where the goal is to minimize a cost function that quantifies the differences between the model predictions and the experimental measurements. The high number of variables in a typical biomodel makes an exact solution to the problem unfeasible in practice. There are however many approximation methods that can be used to, each with its own advantages and disadvantages. For instance, while local approximation algorithms are faster, they tend to converge to local optima. On the other hand, global optimization algorithms are typically slower, but they tend to converge to a global optimum. The global optimization methods can be further divided into deterministic [12, 20] and stochastic approaches [2, 13]. Although the deterministic methods guaranty the convergence to a global optimum, they cannot ensure the termination of this process within a finite time interval [33] and usually come with no guarantees on the speed of the convergence. On the other hand, the inherent randomness of the stochastic approaches makes it difficult to guaranty their convergence to global optima [33]. However, many stochastic methods are capable of locating the vicinity of global solutions with relative efficiency in practice [33]. We choused COPASI, [19], as a computational environment for parameter fitting since it includes a number of optimization algorithms, both local and global. This software is a widely used tool in the computational systems biology modeling community, having a documented good performance, see e.g. [3, 31, 33]. In particular, for determining the best numerical fits of our models, a suite of various global, stochastic parameter estimation procedures was used, comprising of methods such as simulated annealing, genetic algorithm, evolution strategy using stochastic ranking, and particle swarm. Each of these methods uses its own specific strategy for sampling the parameter space looking for combinations of parameter numerical values that give better and better fits of the model predictions to the experimental data.

We concentrated our numerical estimation procedures for deriving the kinetic rate constants corresponding on the second step of the model extension (see the discussion in Section 3.1): (de)phosphorylation reactions, the misfolding and refolding reactions of the associated ki and ph enzymes, and the phosphorylation-dependent transcription reactions. The remaining kinetic rate constants were derived from the corresponding values from the reference model, by using the quantitative refinement procedure, see Section 3.1.

We also preserved several other aspects from the reference model. We assumed that the total number of hsf (phosphorylated or not) in monomers, dimers, trimers, bound to DNA, or bound to hsp, is constant; the actual concentration of the total initial population was taken from [37]. Also, the total number of hse, either free or bound to hsf_3, is constant. As above, the total concentration of the initial hse population was taken from [37].

The kinetic constants of all (de)phosphorylation reactions were considered to be equal. We made this computational assumption since all of these reactions refer to the (de)phosphorylation of the same protein, namely hsf.

The model fit consisted in searching for a set of parameter values that minimizes three different sums of squared deviations: (i) that of the model predictions for DNA binding activity at $42°C$ from the experimental data of [23]; (ii) that of the model predictions for total phosphorylation at $42°C$ from the experimental data of [23]; (iii) that of the model predictions for hsf_3: hse at $37°C$

from the initial value $\mathsf{hsf_3}\colon \mathsf{hse}(0)$. The first two criteria seek to fit the model to the available experimental data of [23] on DNA binding and heat shock factor phosphorylation at $42°C$, while the last condition seeks to set the initial state of the model to an approximation of a steady state of the model at $37°C$, an internal restriction of our model proposed in [36], see also Section 2.4. Since our fit condition involved two different temperature values, we duplicated the model and ran the two models independently for temperature values of $37°C$ and $42°C$, resp. We made sure however that the two models use the same initial values and the same kinetic constants. After each fit iteration, we approximated the steady state of the model at $37°C$ using Copasi and set it as the initial state of both models; we then compared the model predictions with the two experimental data sets.

3.3 Quality of model fit

To estimate how good a model fit is, or in other words how well the global minimum of the scoring function was approximated, a notion of model fit quality was introduced in [24]. The fit quality considers only one set of experimental data at a time and attempts to give a measure of the average deviation of the model prediction from the experimental data, while relating it to the (average of the) absolute values of the model predictions. In this way, the fit quality does not discriminate against model deviations that are large in their absolute value, but relatively small in comparison with the (also large) values that the model aims to explain. This method also allows for comparisons of different models and different data sets. The formula for the quality of the fit with respect to the experimental data exp, denoted $qual(exp)$, is:

$$qual(exp) = \frac{\sqrt{SS_f/N_f}}{\text{mean of predicted values}} \cdot 100\%, \tag{1}$$

where SS_f is the sum of squared deviation of the model predictions from the experimental values exp and N_f is the number of experimental data points. It was argued in [24] that a low (say, lower than 20%) value of $qual(exp)$ could be considered as an indicator of a successful fit. We discuss the quality of our best fits of all our models in Section 5.

4 Hypothesis for the correlation between the heat shock response and the hsf hyperphosphorylation

We proposed and analyzed in this study three alternative mechanisms for the phosphorylation-induced regulation of HSR. In the first setting, we analyzed whether the kinase and phosphatase dynamics for the heat-induced misfolding and refolding could lead to the experimentally observed evolution of the total level of phosphorylated hsf. In the second setting, on top of the the heat-induced misfolding of both kinase and phosphatase, we also assumed that hsf can be both phosphorylated and dephosphorylated while they are bound in hsp: hsf complexes. In the third setting we assumed that hsp: hsf can only be dephosphorylated, not phosphorylated.

4.1 First scenario: heat-induced misfolding of the kinase and phosphatase enzymes

The first phosphorylation-controlled mechanism for HSR involves the dynamics of ki and ph enzymes. Similarly to other proteins, these enzymes are also subject to misfolding due to heat shock. We hypothesized here that the increase in the phosphorylation level at 42°C could be explained if the phosphatase enzymes were more prone to misfolding than the kinase. In order to investigate this control mechanism, we incorporated it into the extended molecular model, and derived the corresponding mathematical (ODE) model. The newly obtained computational model was subjected to parameter estimation procedures, in order to fit it to the available experimental data.

The new molecular reactions included into the extended model, are:

- heat-induced misfolding of kinase (mki) and phosphatase (mph): $\mathsf{ki} \rightarrow \mathsf{mki}$, $\mathsf{ph} \rightarrow \mathsf{mph}$;
- sequestration of the misfolded enzymes by the chaperons: $\mathsf{mki} + \mathsf{hsp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mki}$, $\mathsf{mph} + \mathsf{hsp} \leftrightarrows \mathsf{hsp}\colon \mathsf{mph}$;
- refolding of the enzymes: $\mathsf{hsp}\colon \mathsf{mki} \rightarrow \mathsf{ki} + \mathsf{hsp}$, $\mathsf{hsp}\colon \mathsf{mph} \rightarrow \mathsf{ph} + \mathsf{hsp}$.

When deriving the associated computational extended model, the rate on which the ki and ph enzymes are misfolded was inferred from the generic misfolding function of proteins, i.e. reaction (9) from Table 1, by multiplying it with a constant, which remains to be estimated. We introduced this multiplication with a constant to have a simple model for the case where the the ki and ph populations are much more sensitive to heat than the average protein.

4.2 Second and third scenarios: (De)Phosphorylation of hsf in hsp: hsf complexes

In the first scenario, hsf could be (de)phosphorylated while present in all complexes, except when bound to hsp. However, as observed from the numerical simulation of our reference HSR computational model, many hsfs are bound in these complexes most of the time. By allowing the hsf proteins from these complexes to be (de)phosphorylated, the influence of the phosphorylation process over the control of the HSR increases. Due to the uncertainties regarding the situations in which hsf is prone to (de)phosphorylation, we analyzed two possible scenarios regarding the (de)phosphorylation of hsp: hsf. In scenario II, in order to assume an un-biased approach, we allowed for both phosphorylation and dephosphorylation. In scenario III, following the biological mechanism proposed in [40], we considered that while within the hsp: hsf complex, hsf can only be dephosphorylated.

In the first setting, six new reactions (containing also two new species) had to be added to the molecular model for the (de)phosphorylation of the hsp: hsf complex:

$$\mathsf{hsp}\colon \mathsf{hsf} + \mathsf{ki} \rightleftarrows \mathsf{hsp} : \mathsf{hsf} : \mathsf{ki} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^{\mathsf{p}} + \mathsf{ki},$$

$$\mathsf{hsp}\colon \mathsf{hsf}^{\mathsf{p}} + \mathsf{ph} \rightleftarrows \mathsf{hsp} : \mathsf{hsf}^{\mathsf{p}} : \mathsf{ph} \rightarrow \mathsf{hsp}\colon \mathsf{hsf} + \mathsf{ph}\,.$$

In the second setting, since we allowed only for the dephosphorylation of the $\mathsf{hsp}\colon \mathsf{hsf}^{\mathsf{p}}$ complex, only the last three of the above reactions were included into the model.

No new kinetic rate constants were added to the models since the newly introduced hsp: hsf (de)phosphorylation reactions had the same rate constants as the already existing (de)phosphorylation reactions.

4.3 Gene transcription models

As we discussed in Section 2.2, we tested two different hypothesis for the phosphorylation-induced transcription process. In the first case, we assumed that the transcription proceeds linearly depending on the level of phosphorylation of the $\mathsf{hsf_3}$: hse complex. In the second case, we assumed that hsp synthesis is activated only by the hyper-phosphorylation of the $\mathsf{hsf_3}$: hse compound, i.e., only when at least two of the hsf proteins from this compound are phosphorylated.

5 Results

We subjected all the computational models from the previous section to parameter estimation procedures. As previously mentioned, we estimated only part of the model's parameters, while the remaining ones were preserved from our reference HSR model. In particular, we estimated the rate constants of reactions (36')-(39'), denoted $k'_{12} - k'_{15}$, resp., as described below. The parameters which were (re)estimated during these processes are presented below:

- the three kinetic rate constants corresponding to the phosphorylation reactions, e.g., for reactions (40') and (41') (Table 2);
- the three kinetic rate constants corresponding to the dephosphorylation reactions, e.g., for reactions (42') and (43') (Table 2);
- the kinetic rate constants $k'_{12} - k'_{15}$ corresponding to the hsp synthesis reactions:
  - in the case when the flux of the hsp synthesis is linearly dependent on the $\mathsf{hsf_3}$: hse phosphorylation level, we estimated two values: the base rate a giving the transcription rate in the case of non-phosphorylated hsf trimers and the increment b giving the linear dependency of the transcription activity on the phosphorylation level of the hsf trimer. We then set $k'_{12+i} = a + i * b$, for $i = 0, 1, 2, 3$. (Table 2);
  - in the case when hsp synthesis is activated only when the $\mathsf{hsf_3}$: hse complex is hyper-phosphorylated (i.e., phosphorylated on either two or all three sites) we estimated only one rate constant r, shared by reactions (38')-(39') (Table 2). In this case, we set $k'_{12} = k'_{13} = 0$ and $k'_{14} = k'_{15} = r$.
- the multiplication factor for the ki misfolding function in reaction (80') (Table 3);
- the two kinetic rate constants corresponding to the hsp (reversible) sequestration of mki, i.e. reaction (81') (Table 3);
- the kinetic rate constant corresponding to the hsp refolding of mki, i.e. reaction (82') (Table 3);
- the multiplication factor for the ph misfolding function in reaction (83') (Table 3);

Table 5: The twelve computational models tested in this study.

		Scenarios for the (de)phosphorylation of hsf				
		I	*II*	*III*		
Transcription level	*Linear*	M1.1	M2.1	M3.1	*Mass-action*	**Kinetic model**
	Threshold	M1.2	M2.2	M3.2	*Mass-action*	
	Linear	M1.3	M2.3	M3.3	*Michaelis-Menten*	
	Threshold	M1.4	M2.4	M3.4	*Michaelis-Menten*	

- the two kinetic rate constants corresponding to the hsp (reversible) sequestration of mph, i.e. reaction (84') (Table 3);

- the kinetic rate constant corresponding to the hsp refolding of mph, i.e. reaction (85') (Table 3).

In addition to these kinetic rate constants, we also estimated the initial concentrations (given in numbers of molecules) for: [hsp], [mfp], [hsp : mfp], as well as for the total populations of ki and ph enzymes, either free, or within a complex.

We investigated 12 models spanning the three scenarios for the (de)phosphorylation of hsf, see Section 4.1 – 4.2, the two mechanisms for the influence of the phosphorylation on the transcription level of the hsp-gene, see Section 4.3 and two different kinetic models for the (de)phosphorylation reactions (mass-action and Michaelis-Menten), see Section 6. For each model, we performed a number of parameter estimation iterations using several algorithms implemented in Copasi. The estimation procedure used two experimental data sets, one on DNA binding and the other on phosphorylation, see Section 2.4. For the best fit of each model we calculated the fit quality with respect to each of the two data sets, as discussed in Section 3.3. We collected in Table 14 (supplementary information) the average fit quality for each of the 12 models we tested. We also illustrate the fits in Figures 2-4 included in the supplementary information.

5.1 Evaluation of the control mechanism based on heat induced misfolding of kinase and phosphatase (scenario I)

Despite numerous rounds of parameter estimations, the models M1.1 and M1.2 in which the control over the hsf phosphorylated populations (and thus over the transcription level of the hsp encoding genes) is performed through a differentiated ki and ph misfolding (and subsequent refolding) rate, could not be fit to the experimental data. The result was independent of the two mechanisms for the transactivation of the hsp-gene. The fit quality of the best estimated models M1.1 and M1.2 were 61% and 41%, resp., which were considerably higher than the 20% proposed threshold for a numerical indicator of a successful fit. We considered this as a clear sign that the heat induced misfolding of ki and ph is not the main driver for generating the transient phosphorylation of hsf, and thus, not for the phosphorylation dependent regulation of the HSR.

5.2 Evaluation of the control mechanisms based on the (de)phosphorylation of hsp: hsf complexes (scenarios II and III)

As a first subcase of this setting, we considered the un-biased assumption that the complex hsp: hsf can be both phosphorylated and dephosphorylated (scenario II). In this case, the computational models M2.1 and M2.2 could not be fitted such that their predictions would be in agreement with the existing experimental data. This was again independent of whether the transcription of new hsp is activated by the hyper-phosphorylation of the hsf_3: hse complex, or whether it is linearly dependent on the phosphorylation level of hsf_3: hse. For these cases, the fit quality of the best estimated models M2.1 and M2.2 were 54% and 42%, respectively. However, in comparison with the estimations from the previous setting, i.e., where the level of phosphorylated molecules is modified only due to the differentiated misfolding and refolding of kinase and phosphatase, the fitted model came closer to the experimental data. This suggested that our initial reasoning of considering the hsf from the complex hsp: hsf as a candidate for (de)phosphorylation did increase the influence of the phosphorylation process over the control of the HSR.

The reason this setup could not be fitted to the experimental values may be the following. At 37°C, we required that the phosphorylation level is almost 0%, where 100% represents the maximum phosphorylation level reached at 42°C. Moreover, at steady state at 37°C, most of the hsfs are inside the hsp: hsf complex, which allows for both phosphorylation and dephosphorylation of hsf. Thus, in order for the phosphorylation level to be 0 at $37°C$, the activity of phosphatases must dominate over the activity of the kinases on the complex hsp: hsf. This leads to a tendency of heat shock factors to have low phosphorylation, which continues unhindered also under stress, thus making it difficult for the model to predict high phosphorylation levels at $42°C$.

As a second subcase, we considered the situation when the complex hsp: hsf can only be dephosphorylated (scenario III). Compared with the previous case, the phosphorylated hsfs from within this setting had a different dynamics. Even if in this setting the kinetic constants of the phosphorylation reactions are greater than the ones of the dephosphorylation reaction, at steady state, most of the hsfs are dephosphorylated because they are in the hsp: hsf complex. The values of the quality of the best model fits in this setting were 16% and 18% for models M3.1 and M3.2, resp., under the 20% quality threshold value. The numerical setups for the best fits are given in Tables 7 and 8 for model M3.1 and Tables 9 and 10 for model M3.2 (see the Appendix in the supplementary information).

6 Discussion

We performed in this paper a detailed analysis of the phosphorylation-mediated transcription of heat shock proteins. The three distinct pathways considered in here are each enforcing a different dynamics of the overall process. In the first setting, we analyze whether the experimentally observed evolution of the level of phosphorylated hsf can be explained by differentiated kinase and phosphatase dynamics for heat-induced misfolding and subsequent refolding. In this pathway, we assume that once a hsf enters a hsp: hsf complex, it can be neither phosphorylated nor de-phosphorylated. This particular aspect, i.e., the hsf being (de)phos-

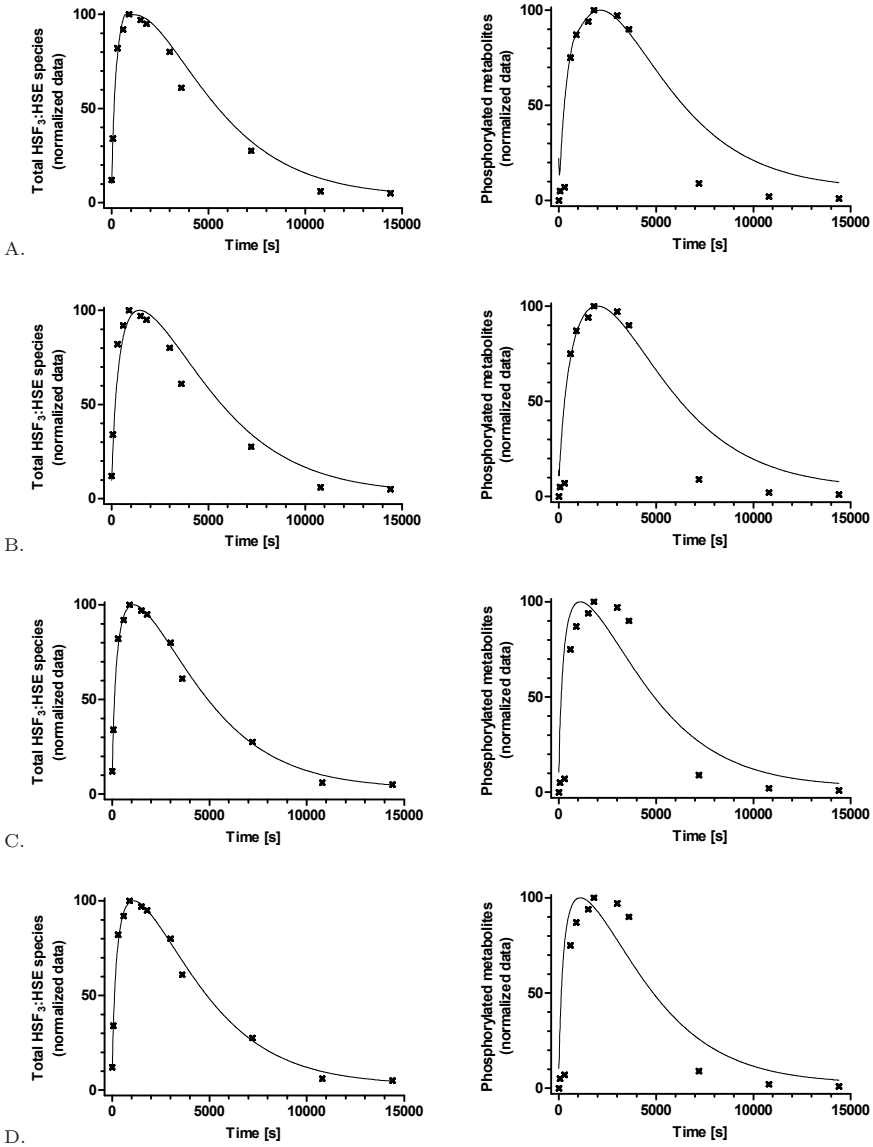

Figure 1: The normalized time-dependent levels of DNA binding and phosphorylated hsf metabolites for the case when the hsp:hsf complex can only be dephosphorylated. The continuous lines are the model predictions, whereas the experimental data (from [23]) are showed with crossed points. Figures A, B, C, D correspond to models M3.1, M3.2, M3.3 and M3.4, resp.

phorylated when within the hsp: hsf complex, is addressed in the following two scenarios. Thus, in the second pathway we consider both hsp: hsf phosphorylation and de-phosphorylation reactions, while in the third pathway we consider only hsp: hsf de-phosphorylation. For each of the three settings we perform a detailed computational analysis including parameter estimations for fitting the models to available experimental data. Despite numerous estimation rounds, only the computational model corresponding to the third setting could be fitted to the experimental data. A similar conclusion is obtained after a separate analysis is performed on a reduced computational model.

In order to further verify our results, we repeated our analysis using a reduced mathematical model employing Michaelis-Menten kinetics for the (de)phosphorylation of the heat shock factors, see Table 11 (the Appendix in the supplementary information). This type of kinetics has been previously employed in the analysis of other (de)phosphorylation pathways, see e.g. [4]. Using Michaelis-Menten kinetics, the computational model becomes significantly reduced, as the ki and ph enzymes are not modeled explicitly anymore. Thus, even if this computational model cannot capture the biological process in as many details as the previous model, it has the advantage of having a considerably smaller size. Because of this, it is easier to apply numerical procedures for parameter estimation, and it can thus give relevant predictions on the actual design of the regulation mechanism.

We considered for modelling each of the three control mechanisms: the case when the hsf from within the hsp: hsf complex can be neither phosphorylated nor dephosphorylated, the case when it can be both phosphorylated and dephosphorylated, and the case when it can be only dephosphorylated, respectively. Moreover, in each of the above three cases, we considered both of the proposed mechanisms for modeling the phosphorylation-dependent transcription of new hsp: the case when the transcription is activated by the hyper-phosphorylation of the hsf_3: hse complex (namely at least two of the hsf proteins from this compound are phosphorylated), and the case when the transcription speed is linearly dependent on the phosphorylation level of hsf_3: hse. Thus, six independent computational models have been created and fitted to the experimental data: M1.3, M1.4, M2.3, M2.4, M3.3 and M3.4.

The results of our second approach are in complete agreement with our previous findings. Both scenarios I and II could not be fitted to the experimental data. This was independent of the way we modeled the phosphorylation-dependent transcription. In particular, the fit quality values of the best estimated models were: 30% and 27% for models M1.3 and M1.4, resp.; 27% and 37% for models M2.3 and M2.4, resp. In the case of scenario III we obtained a good fitting of the experimental data (albeit not as good as the fits in Figures 1 A. and B.), see Figure 1 C. This result also proved to be independent of the way we modeled the phosphorylation-dependent transcription, see Figure 1 D. for the case when the transcription of new hsp is activated by the hyper-phosphorylation of the hsf_3: hse complex. The two associated fit quality values are 17% and 16% for models M3.3 and M3.4, resp.

A possible reason for the fact that scenario I could not be fitted is that the hsf from within the complex hsp: hsf do not participate at all in the (de)phosphorylation process. Carrying out several numerical simulations (modifying only part of the kinetic parameters) we observe a relatively high phosphorylation level in the steady state at 37°C, much more than experimentally observed. This might be explained by the following mechanism. As observed experimentally, at steady state at 37°C, only a few hsfs are in hsf_3: hse complexes. Thus, in

order to produce the smallest amount of hsp (needed even at 37°C to achieve an equilibrium), some (very few) $\mathsf{hsf_3}$: hse complexes must be highly phosphorylated. Because the hsf from within the complex hsp: hsf do not participate at all in the (de)phosphorylation processes, the total level of phosphorylated hsf at equilibrium must be high in order to have a small portion of the $\mathsf{hsf_3}$: hse complexes phosphorylated.

In scenario II, by allowing the hsf from hsp: hsf complexes to be (de)phosphorylated, the influence of the phosphorylation process over the control of the HSR increases. Thus, for example, at 42°C, the phosphorylation level is higher than in the previous setting in several of our numerical simulations. Although the heat induced increase in the phosphorylation level (i.e., from steady state at 37°C to the peak of the heat shock from 42°C) becomes more pronounced in this setting, it is still not as considerable as experimentally observed, where we have an increase from almost 0% before the heat shock to 100% during the peek of the heat shock. This is because the phosphorylation level from 37°C is still too high. Moreover, most hsf are in hsp: hsf complexes, which allows for both phosphorylation and dephosphorylation. Thus, in order to have a small portion of the $\mathsf{hsf_3}$: hse complexes highly phosphorylated, the total level of phosphorylated hsf at equilibrium must be high.

Contrary to the previous two cases, in the third scenario we can distinguish a mechanism which allows for few $\mathsf{hsf_3}$: hse complexes to be phosphorylated at 37°C without increasing too much the overall hsf phosphorylation level at this temperature. At 37°C, most hsf are bound to hsp; some are phosphorylated before binding to hsp, while others are not. By allowing only the dephosphorylation of hsp: hsf^p we ensure a lower overall phosphorylation level, even if the phosphorylation reaction is slightly more rapid than the dephosphorylation one. Then, at 42°C at the beginning of the heat shock (the first approximately 30 minutes) the phosphorylation level increases considerably as hsp: hsf complexes are broken and hsfs become available for phosphorylation. This is intuitively why the dephosphorylation-only of hsp: hsf^p is the most probable phosphorylation-based control mechanism.

Although it is currently known that hsf can be phosphorylated on several sites, see e.g. [15, 23, 18], in our present computational approach we have modeled only one phosphorylation site for each hsf. Our choice is justified by both modeling and methodological reasons. The current research investigates the positive role of phosphorylation in the up-regulation of hsf transcriptional activity. Although hsf can be phosphorylated on several sites, recent findings have shown that one site particularly, namely S230, has a positive effect on hsf transactivating capacity [17, 38]. Moreover, only upon heat shock, this site is inducibly hyper-phosphorylated. At the same time, numerical modeling of several phosphorylation sites for each hsf would require some enormous computing power. Adding only one phosphorylation site for each hsf transforms the complexity of the model substantially. Namely, if the initial ODE model from [37] contains 10 species and 12 reactions, by incorporating in this model only one phosphorylation site we obtained an ODE model with 45 species and 85 reactions. Numerical simulations and partial parameter estimations on this model alone proved to be a computational challenge.

In our analysis of the phosphorylation mediated transcription of hsp, we considered three distinct (de)phosphorylation pathways for the hsf. Although by the current knowledge of the biological network, each of these pathways seems plausible, we do not exclude the possible existence of other scenarios, involving other (de)phosphorylation mechanisms. However, by our observation, the

mechanisms described here have the greatest influence over the overall process, while other mechanisms also considered during this research were able to just slightly modify the predictions. It could thus be that some new mechanism, not investigated by us, would make our predictions fit the experimental data even better, in particular regarding the steep decrease of the phosphorylation level occurring after 5000 seconds. For these data points, the predictions of our model are higher than the experimental data, which could be explained through an additional, as yet unidentified, control mechanism. The lack of data in-between the data point at 3600 seconds and the one at 7200 seconds makes it difficult to identify the possible nature of such a mechanism; additional experimental data covering this time interval is likely to help the analysis.

Although it was not enforced by our numerical construction, an interesting correlation appeared between the extended computational model for HSR where we use only mass-action kinetics, and the reduced computational model where we used Michaelis-Menten kinetics for modeling the (de)phosphorylation of hsf. As a result of our parameter estimation procedures, the time-dependent prediction of the extended model for the total number of phosphorylated molecules followed closely the available experimental data. In the case of the reduced model, although we still consider the fitting to be good, we can observe a slight delay between the predicted time-dependent value of the total number of phosphorylated hsf and the actual experimental data used to fit the model. Although each of the numerical models were fitted independently according to the available experimental data, the following correlation was observed. If from the extended model we removed the 6 reactions modeling the misfolding, sequestration by hsp, and refolding, respectively, for both ki and ph, then the new prediction of our model regarding the time-dependent value of the total number of phosphorylated hsf would be exactly as in the reduced numerical model. That is, it would be slightly delayed if compared with the actual experimental data. We conclude from here that, while the phosphorylation control mechanism given by the misfolding (and subsequent refolding) of the ki and ph enzymes is not sufficient in itself for generating the experimentally observed behavior of the system, it still has a clear role in fine-tuning the phosphorylation response.

Acknowledgment

This work was supported in part by Academy of Finland, grants 129863, 108421, and 122426.

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, D. Bray, K. Hopkin, K. Roberts, and P. Walter. *Essential Cell Biology, Second Edition.* Garland Science/Taylor & Francis Group, Sept. 2003.

[2] M. M. Ali and C. Storey and A. Törn. Application of stochastic global optimization algorithms to practical problems. *J. Optim. Theory Appl.*, 95(3): 545–563, 1997.

[3] S. M. Baker, K. Schallau, B. H. Junker. Comparison of different algorithms for simultaneous estimation of multiple parameters in kinetic metabolic models. *J Integr Bioinform*, 7(3):133, 2010.

[4] H. Câteau and S. Tanaka. Kinetic analysis of multisite phosphorylation using analytic solutions to Michaelis-Menten equations. *J. Theor. Biol.*, 217:1–14, 2002

[5] W. W. Chen, B. Schoeberl, P. J. Jasper, M. Niepel, U. B. Nielsen, D. A. Lauffenburger, and P. K. Sorger. Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.*, 5:239, 2009.

[6] D. R. Ciocca and S. K. Calderwood. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress Chaperones*, 10:86–103, 2005.

[7] M. S. Clark, K. P. Fraser, and L. S. Peck. Antarctic marine molluscs do have an HSP70 heat shock response. *Cell Stress Chaperones*, 13:39–49, 2008.

[8] El. Czeizler, Eu. Czeizler, B. Iancu, and I. Petre. Quantitative model refinement as a solution to the combinatorial state explosion of biomodels. In Proceedings of the 2nd International Workshop on Static Analysis and Systems Biology, Vennice 2011, (to appear) also as TUCS Technical Report 1015, http://tucs.fi/publications/view/?pub_id=tCzCzIaPe11a. Jun 2011.

[9] E. Czeizler, V. Rogojin, and I. Petre. The phosphorylation of the heat shock factor as a modulator for the heat shock response: computational models. *http://users.ics.tkk.fi/eugenc/HSR.*

[10] V. Danos, J. Feret, W. Fontana, R. Harmer, J. Krivine, Rule-based Modelling and Model Perturbation, Transactions on Computational Systems Biology XI, 116-137, 2009.

[11] H. C. Edwards and D. E. Penney. *Differential Equations: Computing and Modeling, Third Edition.* Prentice Hall, June 2003.

[12] Ignacio E. Grossmann. *Global optimization in engineering design.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[13] C. Guus and E. Boender and H. E. Romeijn. Stochastic Methods. In: R. Horst and P. M. Pardalos (Eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[14] R. Harmer, Rule-based modelling and tunable resolution, S. Barry Cooper & Vincent Danos, editors, Fifth Workshop on Developments in Computational Models - Computational Models from Nature EPTCS 9, 65-71, 2009.

[15] V. Hietakangas, J. K. Ahlskog, A. M. Jakobsson, M. Hellesuo, N. M. Sahlberg, C. I. Holmberg, A. Mikhailov, J. J. Palvimo, L. Pirkkala, and L. Sistonen. Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Mol. Cell. Biol.*, 23:2953–2968, Apr 2003.

[16] G. E. Hofmann, B. A. Buckley, S. Airaksinen, J. E. Keen, and G. N. Somero. Heat-shock protein expression is absent in the antarctic fish Trematomus bernacchii (family Nototheniidae). *J. Exp. Biol.*, 203:2331–2339, Aug 2000.

[17] C. I. Holmberg, V. Hietakangas, A. Mikhailov, J. O. Rantanen, M. Kallio, A. Meinander, J. Hellman, N. Morrice, C. MacKintosh, R. I. Morimoto, J. E. Eriksson, and L. Sistonen. Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. *EMBO J.*, 20:3800–3810, Jul 2001.

[18] C. I. Holmberg, S. E. Tran, J. E. Eriksson, and L. Sistonen. Multisite phosphorylation provides sophisticated regulation of transcription factors. *Trends Biochem. Sci.*, 27:619–627, Dec 2002.

[19] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. COPASI–a COmplex PAthway SImulator. *Bioinformatics*, 22:3067–3074, Dec 2006.

[20] R. Horst and H. Tuy. *Global optimization: Deterministic approaches*. Springer-Verlag, Berlin, 1990.

[21] M. Jaattela. Escaping cell death: survival proteins in cancer. *Exp. Cell Res.*, 248:30–43, Apr 1999.

[22] H. H. Kampinga. Thermotolerance in mammalian cells. Protein denaturation and aggregation, and stress proteins. *J. Cell. Sci.*, 104 (Pt 1):11–17, Jan 1993.

[23] M. P. Kline and R. I. Morimoto. Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. *Mol. Cell. Biol.*, 17:2107–2115, Apr 1997.

[24] M. Kühnel, L. S. Mayorga, T. Dandekar, J. Thakar, R. Schwarz, E. Anes, G. Griffiths, and J. Reich. Modelling phagosomal lipid networks that regulate actin assembly. *BMC Systems Biology* 2:107, 2008.

[25] A. La Terza, G. Papa, C. Miceli, and P. Luporini. Divergence between two Antarctic species of the ciliate Euplotes, E. focardii and E. nobilii, in the expression of heat-shock protein 70 genes. *Mol. Ecol.*, 10:1061–1067, Apr 2001.

[26] J. S. Larson, T. J. Schuetz, and R. E. Kingston. Activation in vitro of sequence-specific DNA binding by a human regulatory factor. *Nature*, 335:372–375, Sep 1988.

[27] J. R. Lepock, H. E. Frey, K. P. Ritchie. Protein denaturation in intact hepatocytes and isolated cellular organelles during heat shock. *The Journal of Cell Biology*, 122(6):1267–1276, 1993.

[28] A. Peper, C. A. Grimbergen, J. A. E. Spaan,J. E. M. Souren,R. van Wijk. A mathematical model of the hsp70 regulation in the cell. *Int. J. Hyperthermia*, 14(1):97–124, 1997.

[29] K. J. Laidler. Chemical Kinetics. Third Edition, Benjamin-Cummings, 1997.

[30] C. L. Masters, G. Simms, N. A. Weinman, G. Multhaup, B. L. McDonald, and K. Beyreuther. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, 82:4245–4249, 1985.

[31] P. Mendes and D. Kell. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14(10): 869–883, 1998.

[32] A. Mizera, Eugen Czeizler, and Ion Petre. Self-assembly models of variable resolution. em LNBI Transactions on Computational Systems Biology (to appear), 2012.

[33] C. G. Moles, P. Mendes and J. R. Banga. *Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. Genome Res* 13:2467–2474, 2003.

[34] R. I. Morimoto. Cells in stress: transcriptional activation of heat shock genes. *Science*, 259:1409–1410, Mar 1993.

[35] A. Peper, C. A. Grimbergen, J. A. Spaan, J. E. Souren, and R. van Wijk. A mathematical model of the hsp70 regulation in the cell. *Int J Hyperthermia*, 14:97–124, 1998.

[36] I. Petre, A. Mizera, C. L. Hyder, A. Meinander, A. Mikhailov, R. I. Morimoto, L. Sistonen, J. E. Eriksson, and R.-J. Back. A simple mass-action model for the eukaryotic heat shock response and its mathematical validation. *Natural Computing*, 10(1):595–612, 2011.

[37] I. Petre, A. Mizera, C. L. Hyder, A. Mikhailov, J. E. Eriksson, L. Sistonen, and R.-J. Back. A new mathematical model for the heat shock response. In A. Condon, D. Harel, J. N. Kok, A. Salomaa, and E. Winfree, editors, *Algorithmic Bioprocesses*, Natural Computing Series, pages 411–425. Springer Berlin Heidelberg, 2009. 10.1007/978-3-540-88869-7-21.

[38] L. Pirkkala, P. Nykanen, and L. Sistonen. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB J.*, 15:1118–1131, May 2001.

[39] A. G. Pockley. Heat shock proteins as regulators of the immune response. *Lancet*, 362:469–476, Aug 2003.

[40] T. R. Rieger, R. I. Morimoto, and V. Hatzimanikatis. Mathematical modeling of the eukaryotic heat-shock response: Dynamics of the hsp70 promoter. *Biophysical Journal*, 88(3):1646 – 1658, 2005.

[41] E. Scherzinger, R. Lurz, M. Turmaine, L. Mangiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, and E. E. Wanker. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell*, 90:549–558, Aug 1997.

[42] P. K. Sorger. Yeast heat shock factor contains separable transient and sustained response transcriptional activators. *Cell*, 62:793–805, Aug 1990.

[43] P. K. Sorger, M. J. Lewis, and H. R. Pelham. Heat shock factor is regulated differently in yeast and HeLa cells. *Nature*, 329:81–84, 1987.

[44] P. Workman and E. de Billy. Putting the heat on cancer. *Nat. Med.*, 13:1415–1417, Dec 2007.

Supplementary information

A Data refinement of a model

In the data-refinement approach, consider that a model M consists of a list of m species $\Sigma = \{A_1, A_2, ..., A_m\}$ and n reactions r_i, $1 \leq i \leq n$, of the form:

$$r_i \ : \ s_{i,1}A_1 + s_{i,2}A_2 + ... + s_{i,m}A_m \xrightarrow{k_i} s'_{i,1}A_1 + s'_{i,2}A_2 + ... + s'_{i,m}A_m,$$

where $s_{i,1}, ..., s_{i,m}, s'_{i,1}, ..., s'_{i,m} \geq 0$ are the stoichiometric coefficients of r_i and $k_i \geq 0$ is its the kinetic rate constant. Assume now that species A_1 is to be detailed in some way, differentiating it into several different subspecies $B_1, \ldots, B_l$; such subspecies may be several different forms of A_1, several biochemical configurations of A_1 (e.g., caused by some protein post-translational modifications), etc. Each of these subspecies may participate in all reactions, where A_1 participated (in model M), possibly with different kinetics. Model M is thus *refined* into a new model M_R, in which the set of species is denoted through the new variables $\{A'_2, A'_3, ..., A'_m\} \cup \{B_1, ..., B_l\}$, for some $l \geq 2$. Variables A'_i, $2 \leq i \leq m$, correspond to species A_i from model M, whereas $B_1, ..., B_l$ replace species A_1 in model M_R. Moreover, each reaction r_i from M is replaced in M_R by reaction r'_i of the following type:

$$\begin{array}{l} r'_i \ : \ (t_{i,1}B_1 + ... + t_{i,l}B_l) + s_{i,2}A'_2 + ... + s_{i,m}A'_m \xrightarrow{k'_i} \\ \qquad (t'_{i,1}B_1 + ... + t'_{i,l}B_l) + s'_{i,2}A'_2 + ... + s'_{i,m}A'_m, \end{array}$$

with k'_i its kinetic rate constant, and $t_{i,1}, ..., t_{i,l}, t'_{i,1}, ..., t'_{i,l}$ nonnegative integers such that $t_{i,1} + ... + t_{i,l} = s_{i,1}$ and $t'_{i,1} + ... + t'_{i,l} = s'_{i,1}$. We say now that *the model M_R is a data refinement of M on variable A_1* if and only if the following two conditions are satisfied:

$$[A_i](t) = [A'_i](t), \text{ for all } 2 \leq i \leq m, \tag{2}$$

$$[A_1](t) = [B_1](t) + ... + [B_l](t), \text{ for all } t \geq 0. \tag{3}$$

Some of the new kinetic parameters of M_R may be known from the literature, or they can be experimentally measured. For the rest of them, for which no previous knowledge and no direct kinetic measurements are available, a computational procedure is needed to calculate them so that (2) and (3) hold. Such a procedure should focus only on those parameters whose kinetic values are not known.

Re-running parameter estimation procedures when the parameter space witnesses a (potentially) quadratic increase in every step of the refinement is computationally very expensive. Moreover, such a procedure makes little sense since the fit of an intermediate model is lost in the next refinement step.

In our data-refinement approach we systematically set the values of the unknown kinetic parameters of the refined model so that relations (2) and (3) hold. Clearly, some of the potential choices of numerical values are unreasonable, such as those where we would set to 0 the kinetic parameters of all reactions involving any of B_2, B_3, ..., B_l; such a choice would only rename all variables of M into M_R and completely elude the idea of refinement. Instead, we take an approach where the refined subspecies $B_1, ..., B_l$ of species A_1 are not distinguishable through the kinetics of the reactions they participate in. Thus, in the absence of any biological data regarding differences between some of these reactants, our choice of kinetic parameters aims to make no numerical distinctions

between their reaction kinetics. As a side effect, this leads to simpler and more elegant mathematical considerations. Note again that if some of the new kinetic parameters of M_R are known (e.g. from the literature), then their known value is to be used in M_R rather than deduced based on refinement.

Setting the kinetic constants of the refined model in the way explained above depends heavily on the details of the model. Identifying a formal set of rules for how to do it for an arbitrary model seems to be an open and highly interesting problem. For a simple model consisting of reaction $A \xrightarrow{k} B$ only, refining A into subspecies A_1 and A_2, it is enough to set the kinetic constants of the refined reactions to k: $A_1 \xrightarrow{k} B$, $A_2 \xrightarrow{k} B$. Indeed, the mass action-based ODE formulation of the refined model leads to $d(A_1 + A_2)/dt = -k(A_1 + A_2)$, i.e., the same ODE as that of A: $dA/dt = -kA$. On the other hand, for the model $2A \xrightarrow{k} B$, refining A into A_1 and A_2 leads to reactions $2A_1 \xrightarrow{k_1} B$, $A_1 + A_2 \xrightarrow{k_2} B$, $2A_2 \xrightarrow{k_3} B$. Setting $k_1 = k_2 = k_3 = k$ as in the previous example leads to an incorrect refinement where $A_1 + A_2 \neq A$. Instead, setting $k_1 = k_3 = k$, $k_2 = 2k$ leads to a correct refinement. Note also that refining species A in the model $B \xrightarrow{k} 2A$ to subspecies A_1 and A_2 leads to the refined model $B \xrightarrow{k} 2A_1$, $B \xrightarrow{k/2} A_1 + A_2$, $B \xrightarrow{k} 2A_2$. Intuitively, the non-uniform setting of the constants in the last two examples compensates for the loss of symmetry going from reaction $2A \to B$ ($B \to 2A$) to reaction $A_1 + A_2 \to B$ ($B \to A_1 + A_2$, resp.)

Table 6: The numerical values of the parameters (A) and the values of the initial populations (B) for the basic heat shock response model (Table 1). k_i denotes the kinetic rate constant of the irreversible reaction (i) in Table 1; k_i^+ denotes the 'left-to-right' direction of reaction (i); k_i^- denotes its 'right-to-left' direction. The values of the initial populations are in $\sharp \cdot ml^{-1}$.

A

Par.	**Value**	**Unit**	**Par.**	**Value**	**Unit**
k_1^+	3.49	$\frac{ml}{\#\cdot s}$	k_7	$2.73 \cdot 10^{-7}$	$\frac{ml}{\#\cdot s}$
k_1^-	0.19	s^{-1}	k_8	$3.2 \cdot 10^{-5}$	s^{-1}
k_2^+	1.07	$\frac{ml}{\#\cdot s}$	k_{10}^+	$3.32 \cdot 10^{-3}$	$\frac{ml}{\#\cdot s}$
k_2^-	10^{-9}	s^{-1}	k_{10}^-	4.44	s^{-1}
k_3^+	0.17	$\frac{ml}{\#\cdot s}$	k_{11}	13.94	s^{-1}
k_3^-	$1.21 \cdot 10^{-6}$	s^{-1}	k_{12}	$8.3 \cdot 10^{-3}$	s^{-1}
k_4^+	9.74	$\frac{ml}{\#\cdot s}$			
k_4^-	3.56	s^{-1}			
k_5	2.33	$\frac{ml}{\#\cdot s}$			
k_6	$4.31 \cdot 10^{-5}$	$\frac{ml}{\#\cdot s}$			

B

Species	**Init. pop.**
[hsf]	0.67
[hsf_2]	$8.7 \cdot 10^{-4}$
[hsf_3]	$1.2 \cdot 10^{-4}$
[hse]	29.73
[hsf_3: hse]	2.96
[hsp]	766.88
[hsp: hsf]	1403.13
[mfp]	517.352
[hsp: mfp]	71.65
[prot]	$1.15 \cdot 10^8$

Table 7: Parameter estimation results: the numerical values of the parameters of the extended heat shock response model (Tables 2 and 3), for the case when hsp: hsf can only be dephosphorylated, and the transactivation of the hsp-encoding genes is proportional with the phosphorylation level of the hsf_3: hse complex (model M3.1). We skip the values of the rate constants for reactions (1')-(35') (Table 2), which were fixed according to Tables 4 and 6. The rate constants of reactions (36')-(39'), $k'_{12} - k'_{15}$, resp., were part of the parameter estimation as described in Section 5.

Par.	**Value**	**Units**
k'_{12}	$1.85 \cdot 10^{-3}$	s^{-1}
k'_{13}	$3.73 \cdot 10^{-3}$	s^{-1}
k'_{14}	$5.61 \cdot 10^{-3}$	s^{-1}
k'_{15}	$7.49 \cdot 10^{-3}$	s^{-1}
k'^{+}_{16}	$4.12 \cdot 10^{-2}$	$\frac{ml}{\#\cdot s}$
k'^{-}_{16}	2.99	s^{-1}
k'_{17}	$5.78 \cdot 10^{3}$	s^{-1}
k'^{+}_{18}	$1.64 \cdot 10^{-3}$	$\frac{ml}{\#\cdot s}$
k'^{-}_{18}	$1.72 \cdot 10^{-2}$	s^{-1}

Par.	**Value**	**Units**
k'_{19}	$2.3 \cdot 10^{3}$	s^{-1}
k'_{20}	$5.4 \cdot 10^{3}$	
k'^{+}_{21}	$5.43 \cdot 10^{-4}$	$\frac{ml}{\#\cdot s}$
k'^{-}_{21}	9.45	s^{-1}
k'_{22}	2.87	s^{-1}
k'_{23}	$1.09 \cdot 10^{3}$	
k'^{+}_{24}	$8.77 \cdot 10^{-2}$	$\frac{ml}{\#\cdot s}$
k'^{-}_{24}	1.73	s^{-1}
k'_{25}	74.86	s^{-1}

Table 8: The values of the initial populations (given in $\sharp \cdot ml^{-1}$) from the extended heat shock response model (Tables 2 and 3), for the case when hsp: hsf can only be dephosphorylated, and the transactivation of the hsp-encoding genes is proportional with the phosphorylation level of the $\mathsf{hsf_3}$: hse complex (model M3.1).

Species	**Init. pop.**
[hsf]	0.68
[hsf^p]	$2.83 \cdot 10^{-3}$
[hsf: ki]	$8.68 \cdot 10^{-5}$
[$\mathsf{hsf^p}$: ph]	$1.15 \cdot 10^{-7}$
[$\mathsf{hsf_2}$]	$9.27 \cdot 10^{-4}$
[hsf_2^p]	$8.11 \cdot 10^{-6}$
[hsf_2^{pp}]	$1.95 \cdot 10^{-8}$
[$\mathsf{hsf_2}$: ki]	$1.18 \cdot 10^{-7}$
[$\mathsf{hsf}_2^{\mathsf{p}}$: ki]	$1.03 \cdot 10^{-9}$
[$\mathsf{hsf}_2^{\mathsf{PP}}$: ph]	$7.89 \cdot 10^{-13}$
[$\mathsf{hsf}_2^{\mathsf{p}}$: ph]	$3.28 \cdot 10^{-10}$
[$\mathsf{hsf_3}$]	$1.18 \cdot 10^{-4}$
[hsf_3^p]	$1.64 \cdot 10^{-5}$
[hsf_3^{pp}]	$2.17 \cdot 10^{-6}$
[hsf_3^{ppp}]	$1.01 \cdot 10^{-6}$
[$\mathsf{hsf_3}$: ki]	$2 \cdot 10^{-8}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: ki]	$2.80 \cdot 10^{-9}$
[$\mathsf{hsf}_3^{\mathsf{PP}}$: ki]	$3.92 \cdot 10^{-10}$
[$\mathsf{hsf}_3^{\mathsf{PPP}}$: ph]	$7.71 \cdot 10^{-11}$
[$\mathsf{hsf}_3^{\mathsf{PP}}$: ph]	$1.83 \cdot 10^{-10}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: ph]	$1.31 \cdot 10^{-9}$
[hse]	29.37
[$\mathsf{hsf_3}$: hse]	$6.05 \cdot 10^{-3}$

Species	**Init. pop.**
[$\mathsf{hsf_3}$: hse^p]	0.0435
[$\mathsf{hsf_3}$: hse^{pp}]	0.353
[$\mathsf{hsf_3}$: hse^{ppp}]	2.91
[$\mathsf{hsf_3}$: hse: ki]	$1.30 \cdot 10^{-6}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: hse: ki]	$8.53 \cdot 10^{-6}$
[$\mathsf{hsf}_3^{\mathsf{PP}}$: hse: ki]	$6.07 \cdot 10^{-5}$
[$\mathsf{hsf}_3^{\mathsf{PPP}}$: hse: ph]	$1.18 \cdot 10^{-4}$
[$\mathsf{hsf}_3^{\mathsf{PP}}$: hse: ph]	$2.83 \cdot 10^{-5}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: hse: ph]	$3.98 \cdot 10^{-6}$
[hsp]	749
[hsp: hsf]	1396
[hsp: hsf^p]	5.67
[hsp: $\mathsf{hsf^p}$: ph]	$2.3 \cdot 10^{-4}$
[prot]	$1.15 \cdot 10^{8}$
[mfp]	529.6
[hsp: mfp]	71.65
[ki]	17.87
[mki]	8.84
[hsp: mki]	0.292
[ph]	54.13
[mph]	0.008
[hsp: mph]	$6.87 \cdot 10^{-3}$

Table 9: Parameter estimation results: the numerical values of the parameters of the extended heat shock response model (Tables 2 and 3), for the case when hsp: hsf can only be dephosphorylated, and the transactivation of the hsp-encoding genes is activated only after the hsf_3: hse complex is hyperphosphorylated (model M3.2). We skip the values of the rate constants for reactions (1')-(35') (Table 2), which were fixed according to Tables 4 and 6. The rate constants of reactions (36')-(39'), $k'_{12} - k'_{15}$, resp., were part of the parameter estimation as described in Section 5.

Par.	**Value**	**Units**	**Par.**	**Value**	**Units**
k'_{12}	0	s^{-1}	k'_{19}	351	s^{-1}
k'_{13}	0	s^{-1}	k'_{20}	$2.5 \cdot 10^3$	
k'_{14}	$1.6 \cdot 10^{-2}$	s^{-1}	k'^{+}_{21}	$7.46 \cdot 10^{-3}$	$\frac{ml}{\#\cdot s}$
k'_{15}	$1.6 \cdot 10^{-2}$	s^{-1}	k'^{-}_{21}	0.117	s^{-1}
k'^{+}_{16}	1.56	$\frac{ml}{\#\cdot s}$	k'_{22}	49.86	s^{-1}
k'^{-}_{16}	0.141	s^{-1}	k'_{23}	$2.5 \cdot 10^3$	
k'_{17}	$29.92 \cdot 10^2$	s^{-1}	k'^{+}_{24}	0.367	$\frac{ml}{\#\cdot s}$
k'^{+}_{18}	1.51	$\frac{ml}{\#\cdot s}$	k'^{-}_{24}	$5.02 \cdot 10^{-2}$	s^{-1}
k'^{-}_{18}	9.98	s^{-1}	k'_{25}	8.07	s^{-1}

Table 10: The values of the initial populations (given in $\sharp \cdot ml^{-1}$) from the extended heat shock response model (Tables 2 and 3), for the case when hsp: hsf can only be dephosphorylated, and the transactivation of the hsp-encoding genes is activated only after the $\mathsf{hsf_3}$: hse complex is hyper-phosphorylated (model M3.2).

Species	**Init. pop.**
[hsf]	0.655
[hsf^{p}]	$4.6 \cdot 10^{-3}$
[hsf: ki]	$1.08 \cdot 10^{-2}$
[$\mathsf{hsf}^{\mathsf{p}}$: ph]	$5.7 \cdot 10^{-4}$
[$\mathsf{hsf_2}$]	$8.04 \cdot 10^{-4}$
[hsf_2^{p}]	$3.19 \cdot 10^{-5}$
[hsf_2^{pp}]	$8.86 \cdot 10^{-7}$
[$\mathsf{hsf_2}$: ki]	$1.75 \cdot 10^{-5}$
[$\mathsf{hsf}_2^{\mathsf{p}}$: ki]	$5.26 \cdot 10^{-7}$
[$\mathsf{hsf}_2^{\mathsf{pp}}$: ph]	$1.1 \cdot 10^{-7}$
[$\mathsf{hsf}_2^{\mathsf{p}}$: ph]	$3.95 \cdot 10^{-6}$
[$\mathsf{hsf_3}$]	$3.37 \cdot 10^{-5}$
[hsf_3^{p}]	$2.93 \cdot 10^{-5}$
[hsf_3^{pp}]	$2.7 \cdot 10^{-5}$
[hsf_3^{ppp}]	$2.74 \cdot 10^{-5}$
[$\mathsf{hsf_3}$: ki]	$5.56 \cdot 10^{-7}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: ki]	$4.85 \cdot 10^{-7}$
[$\mathsf{hsf}_3^{\mathsf{pp}}$: ki]	$4.44 \cdot 10^{-7}$
[$\mathsf{hsf}_3^{\mathsf{ppp}}$: ph]	$3.4 \cdot 10^{-6}$
[$\mathsf{hsf}_3^{\mathsf{pp}}$: ph]	$3.34 \cdot 10^{-6}$
[$\mathsf{hsf}_3^{\mathsf{p}}$: ph]	$3.62 \cdot 10^{-6}$
[hse]	29.63
[$\mathsf{hsf_3}$: hse]	0.56

Species	**Init. pop.**
[$\mathsf{hsf_3}$: hse^{p}]	0.641
[$\mathsf{hsf_3}$: hse^{pp}]	0.727
[$\mathsf{hsf_3}$: hse^{ppp}]	0.825
[$\mathsf{hsf_3}$: hse: ki]	0.009
[$\mathsf{hsf}_3^{\mathsf{p}}$: hse: ki]	0.01
[$\mathsf{hsf}_3^{\mathsf{pp}}$: hse: ki]	0.012
[$\mathsf{hsf}_3^{\mathsf{ppp}}$: hse: ph]	0.102
[$\mathsf{hsf}_3^{\mathsf{pp}}$: hse: ph]	0.09
[$\mathsf{hsf}_3^{\mathsf{p}}$: hse: ph]	0.079
[hsp]	777
[hsp: hsf]	1402
[hsp: hsf^{p}]	0.739
[hsp: $\mathsf{hsf}^{\mathsf{p}}$: ph]	0.092
[prot]	$1.15 \cdot 10^{8}$
[mfp]	510
[hsp: mfp]	71.65
[ki]	31.64
[mki]	1.19
[hsp: mki]	0.014
[ph]	29.56
[mph]	0.002
[hsp: mph]	0.08

Table 11: The reduced molecular model for the eukaryotic heat shock response

Reaction	**Kinetic par.**	
$2\,\mathsf{hsf} \leftrightarrows \mathsf{hsf}_2$	$2k_1^+, k_1^-$	(1")
$\mathsf{hsf}^p + \mathsf{hsf} \leftrightarrows \mathsf{hsf}_2^p$	k_1^+, k_1^-	(2")
$2\,\mathsf{hsf}^p \leftrightarrows \mathsf{hsf}_2^{pp}$	k_1^+, k_1^-	(3")
$\mathsf{hsf} + \mathsf{hsf}_2 \leftrightarrows \mathsf{hsf}_3$	k_2^+, k_2^-	(4")
$\mathsf{hsf}^p + \mathsf{hsf}_2 \leftrightarrows \mathsf{hsf}_3^p$	$k_2^+, \frac{1}{2}k_2^-$	(5")
$\mathsf{hsf} + \mathsf{hsf}_2^p \leftrightarrows \mathsf{hsf}_3^p$	$k_2^+, \frac{1}{2}k_2^-$	(6")
$\mathsf{hsf}^p + \mathsf{hsf}_2^p \leftrightarrows \mathsf{hsf}_3^{pp}$	$k_2^+, \frac{1}{2}k_2^-$	(7")
$\mathsf{hsf} + \mathsf{hsf}_2^{pp} \leftrightarrows \mathsf{hsf}_3^{pp}$	$k_2^+, \frac{1}{2}k_2^-$	(8")
$\mathsf{hsf}^p + \mathsf{hsf}_2^{pp} \leftrightarrows \mathsf{hsf}_3^{ppp}$	k_2^+, k_2^-	(9")
$\mathsf{hsf}_3 + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3\mathsf{:\,hse}$	k_3^+, k_3^-	(10")
$\mathsf{hsf}_3^p + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{p}\mathsf{:\,hse}$	k_3^+, k_3^-	(11")
$\mathsf{hsf}_3^{pp} + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{pp}\mathsf{:\,hse}$	k_3^+, k_3^-	(12")
$\mathsf{hsf}_3^{ppp} + \mathsf{hse} \leftrightarrows \mathsf{hsf}_3^\mathsf{ppp}\mathsf{:\,hse}$	k_3^+, k_3^-	(13")
$\mathsf{hsp} + \mathsf{hsf} \leftrightarrows \mathsf{hsp:\,hsf}$	k_4^+, k_4^-	(14")
$\mathsf{hsp} + \mathsf{hsf}^p \leftrightarrows \mathsf{hsp:\,hsf}^p$	k_4^+, k_4^-	(15")
$\mathsf{hsp} + \mathsf{hsf}_2 \rightarrow \mathsf{hsp:\,hsf} + \mathsf{hsf}$	k_5	(16")
$\mathsf{hsp} + \mathsf{hsf}_2^p \rightarrow \mathsf{hsp:\,hsf} + \mathsf{hsf}^p$	$\frac{1}{2}k_5$	(17")
$\mathsf{hsp} + \mathsf{hsf}_2^p \rightarrow \mathsf{hsp:\,hsf}^p + \mathsf{hsf}$	$\frac{1}{2}k_5$	(18")
$\mathsf{hsp} + \mathsf{hsf}_2^{pp} \rightarrow \mathsf{hsp:\,hsf}^p + \mathsf{hsf}^p$	k_5	(19")
$\mathsf{hsp} + \mathsf{hsf}_3 \rightarrow \mathsf{hsp:\,hsf} + 2\,\mathsf{hsf}$	k_6	(20")
$\mathsf{hsp} + \mathsf{hsf}_3^p \rightarrow \mathsf{hsp:\,hsf}^p + 2\,\mathsf{hsf}$	$\frac{1}{2}k_6$	(21")
$\mathsf{hsp} + \mathsf{hsf}_3^p \rightarrow \mathsf{hsp:\,hsf} + \mathsf{hsf}^p + \mathsf{hsf}$	$\frac{1}{2}k_6$	(22")
$\mathsf{hsp} + \mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsp:\,hsf} + 2\,\mathsf{hsf}^p$	$\frac{1}{2}k_6$	(23")
$\mathsf{hsp} + \mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsp:\,hsf}^p + \mathsf{hsf}^p + \mathsf{hsf}$	$\frac{1}{2}k_6$	(24")
$\mathsf{hsp} + \mathsf{hsf}_3^{ppp} \rightarrow \mathsf{hsp:\,hsf}^p + 2\,\mathsf{hsf}^p$	k_6	(25")
$\mathsf{hsp} + \mathsf{hsf}_3\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf} + 2\,\mathsf{hsf} + \mathsf{hse}$	k_7	(26")
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{p}\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf}^p + 2\,\mathsf{hsf} + \mathsf{hse}$	$\frac{1}{2}k_7$	(27")
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{p}\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf} + \mathsf{hsf}^p + \mathsf{hsf} + \mathsf{hse}$	$\frac{1}{2}k_7$	(28")
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{pp}\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf}^p + \mathsf{hsf}^p + \mathsf{hsf} + \mathsf{hse}$	$\frac{1}{2}k_7$	(29")
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{pp}\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf} + 2\,\mathsf{hsf}^p + \mathsf{hse}$	$\frac{1}{2}k_7$	(30")
$\mathsf{hsp} + \mathsf{hsf}_3^\mathsf{ppp}\mathsf{:\,hse} \rightarrow \mathsf{hsp:\,hsf}^p + + 2\,\mathsf{hsf}^p + \mathsf{hse}$	k_7	(31")
$\mathsf{hsp} \rightarrow \emptyset$	k_8	(32")
$\mathsf{prot} \rightarrow \mathsf{mfp}$	ϕ_T	(33")
$\mathsf{hsp} + \mathsf{mfp} \leftrightarrows \mathsf{hsp:\,mfp}$	k_{10}^+, k_{10}^-	(34")
$\mathsf{hsp:\,mfp} \rightarrow \mathsf{hsp} + \mathsf{prot}$	k_{11}	(35")
$\mathsf{hsf}_3\mathsf{:\,hse} \rightarrow \mathsf{hsf}_3\mathsf{:\,hse} + \mathsf{hsp}$	k_{12}''	(36")
$\mathsf{hsf}_3^\mathsf{p}\mathsf{:\,hse} \rightarrow \mathsf{hsf}_3^\mathsf{p}\mathsf{:\,hse} + \mathsf{hsp}$	k_{13}''	(37")
$\mathsf{hsf}_3^\mathsf{pp}\mathsf{:\,hse} \rightarrow \mathsf{hsf}_3^\mathsf{pp}\mathsf{:\,hse} + \mathsf{hsp}$	k_{14}''	(38")
$\mathsf{hsf}_3^\mathsf{ppp}\mathsf{:\,hse} \rightarrow \mathsf{hsf}_3^\mathsf{ppp}\mathsf{:\,hse} + \mathsf{hsp}$	k_{15}''	(39")
$\mathsf{hsf} \rightarrow \mathsf{hsf}^p$	$K_M''^1, V_{max}''^1$	(40")
$\mathsf{hsf}^p \rightarrow \mathsf{hsf}$	$K_M''^2, V_{max}''^2$	(41")
$\mathsf{hsf}_2 \rightarrow \mathsf{hsf}_2^p$	$K_M''^1, V_{max}''^1$	(42")
$\mathsf{hsf}_2^p \rightarrow \mathsf{hsf}_2^{pp}$	$K_M''^1, V_{max}''^1$	(43")
$\mathsf{hsf}_2^{pp} \rightarrow \mathsf{hsf}_2^p$	$K_M''^2, V_{max}''^2$	(44")
$\mathsf{hsf}_2^p \rightarrow \mathsf{hsf}_2$	$K_M''^2, V_{max}''^2$	(45")
$\mathsf{hsf}_3 \rightarrow \mathsf{hsf}_3^p$	$K_M''^1, V_{max}''^1$	(46")

Continued on next page

Table 11 – continued from previous page

Reaction	Kinetic par.	
$\mathsf{hsf}_3^p \rightarrow \mathsf{hsf}_3^{pp}$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(47")
$\mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsf}_3^{ppp}$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(48")
$\mathsf{hsf}_3^{ppp} \rightarrow \mathsf{hsf}_3^{pp}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(49")
$\mathsf{hsf}_3^{pp} \rightarrow \mathsf{hsf}_3^{p}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(50")
$\mathsf{hsf}_3^{p} \rightarrow \mathsf{hsf}_3$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(51")
$\mathsf{hsp}\colon \mathsf{hsf} \rightarrow \mathsf{hsp}\colon \mathsf{hsf}^p$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(52")
$\mathsf{hsp}\colon \mathsf{hsf}^p \rightarrow \mathsf{hsp}\colon \mathsf{hsf}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(53")
$\mathsf{hsf}_3\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^{\mathsf{p}}\colon \mathsf{hse}$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(54")
$\mathsf{hsf}_3^{\mathsf{p}}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^{\mathsf{pp}}\colon \mathsf{hse}$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(55")
$\mathsf{hsf}_3^{\mathsf{pp}}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^{\mathsf{ppp}}\colon \mathsf{hse}$	$K_M'^{\prime 1}, V_{max}'^{\prime 1}$	(56")
$\mathsf{hsf}_3^{\mathsf{ppp}}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^{\mathsf{pp}}\colon \mathsf{hse}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(57")
$\mathsf{hsf}_3^{\mathsf{pp}}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3^{\mathsf{p}}\colon \mathsf{hse}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(58")
$\mathsf{hsf}_3^{\mathsf{p}}\colon \mathsf{hse} \rightarrow \mathsf{hsf}_3\colon \mathsf{hse}$	$K_M'^{\prime 2}, V_{max}'^{\prime 2}$	(59")

Table 12: Parameter estimation results: the numerical values of the parameters and the values of the initial populations for the reduced heat shock response model from Table 11, for the case when $\mathsf{hsp}\colon\mathsf{hsf}$ can only be dephosphorylated, and the transactivation of the hsp-encoding genes is proportional with the phosphorylation level of the $\mathsf{hsf}_3\colon\mathsf{hse}$ complex. A. The numerical values of the parameters. B. The values of the initial populations in $\sharp \cdot ml^{-1}$.

A

Par.	**Value**	**Units**
k''_{12}	$1.44 \cdot 10^{-3}$	s^{-1}
k''_{13}	$3.53 \cdot 10^{-3}$	s^{-1}
k''_{14}	$5.62 \cdot 10^{-3}$	s^{-1}
k''_{15}	$7.72 \cdot 10^{-3}$	s^{-1}
K''^1_M	$3.85 \cdot 10^{-2}$	$\frac{\#}{ml}$
V''^1_{max}	3.24	$\frac{\#}{ml}$
K''^2_M	$1 \cdot 10^{-5}$	$\frac{\#}{ml}$
V''^2_{max}	1.64	$\frac{\#}{ml}$

B

Species	**Init. pop.**
$[\mathsf{hsf}]$	0.679
$[\mathsf{hsf}^p]$	$2.04 \cdot 10^{-4}$
$[\mathsf{hsf}_2]$	$9.14 \cdot 10^{-4}$
$[\mathsf{hsf}_2^p]$	$4.8 \cdot 10^{-7}$
$[\mathsf{hsf}_2^{pp}]$	$2.46 \cdot 10^{-10}$
$[\mathsf{hsf}_3]$	$1.35 \cdot 10^{-4}$
$[\mathsf{hsf}_3^p]$	$6.96 \cdot 10^{-8}$
$[\mathsf{hsf}_3^{pp}]$	$5.89 \cdot 10^{-11}$
$[\mathsf{hsf}_3^{ppp}]$	$2.27 \cdot 10^{-11}$
$[\mathsf{hse}]$	29.49
$[\mathsf{hsf}_3\colon\mathsf{hse}]$	0.04
$[\mathsf{hsf}_3\colon\mathsf{hse}^p]$	$3.97 \cdot 10^{-2}$
$[\mathsf{hsf}_3\colon\mathsf{hse}^{pp}]$	$3.97 \cdot 10^{-2}$
$[\mathsf{hsf}_3\colon\mathsf{hse}^{ppp}]$	3.08
$[\mathsf{hsp}]$	754
$[\mathsf{hsp}\colon\mathsf{hsf}]$	1402
$[\mathsf{hsp}\colon\mathsf{hsf}^p]$	$1.07 \cdot 10^{-4}$
$[\mathsf{prot}]$	$1.15 \cdot 10^8$
$[\mathsf{mfp}]$	525
$[\mathsf{hsp}\colon\mathsf{mfp}]$	71.65

Table 13: Parameter estimation results: the numerical values of the parameters and the values of the initial populations for the reduced heat shock response model from Table 11, for the case when hsp: hsf can only be dephosphorylated, and the transactivation of the hsp-encoding genes is activated only after the hsf_3: hse complex is highly-phosphorylated (at leat two, out of three sites are phosphorylated). A. The numerical values of the parameters. B. The values of the initial populations in $\sharp \cdot ml^{-1}$.

A

Par.	**Value**	**Units**
k''_{12}	0	s^{-1}
k''_{13}	0	s^{-1}
k''_{14}	$8.16 \cdot 10^{-3}$	s^{-1}
k''_{15}	$8.16 \cdot 10^{-3}$	s^{-1}
K''^1_M	$2.93 \cdot 10^{-2}$	$\frac{\#}{ml}$
V''^1_{max}	123	$\frac{\#}{ml \cdot s}$
K''^2_M	$6.73 \cdot 10^{-3}$	$\frac{\#}{ml}$
V''^2_{max}	94.45	$\frac{\#}{ml \cdot s}$

B

Species	**Init. pop.**
[hsf]	0.667
[hsf^p]	$8.76 \cdot 10^{-2}$
[hsf_2]	$6.77 \cdot 10^{-4}$
[hsf_2^p]	$1.77 \cdot 10^{-4}$
[hsf_2^{pp}]	$4.71 \cdot 10^{-5}$
[hsf_3]	$9.17 \cdot 10^{-5}$
[hsf_3^p]	$2.75 \cdot 10^{-5}$
[hsf_3^{pp}]	$8.22 \cdot 10^{-6}$
[hsf_3^{ppp}]	$2.46 \cdot 10^{-6}$
[hse]	29.57
[hsf_3: hse]	0.07
[hsf_3: hse^p]	0.074
[hsf_3: hse^{pp}]	0.096
[hsf_3: hse^{ppp}]	2.88
[hsp]	758
[hsp: hsf]	1403
[hsp: hsf^p]	$1.48 \cdot 10^{-2}$
[prot]	$1.15 \cdot 10^8$
[mfp]	523
[hsp: mfp]	71.65

Table 14: The average fit quality for each of the tested models.

Model	Avg. fit quality	Model	Avg. fit quality	Model	Avg. fit quality
M1.1	61%	M2.1	54%	M3.1	16%
M1.2	41%	M2.2	42%	M3.2	18%
M1.3	30%	M2.3	27%	M3.3	17%
M1.4	27%	M2.4	37%	M3.4	16%

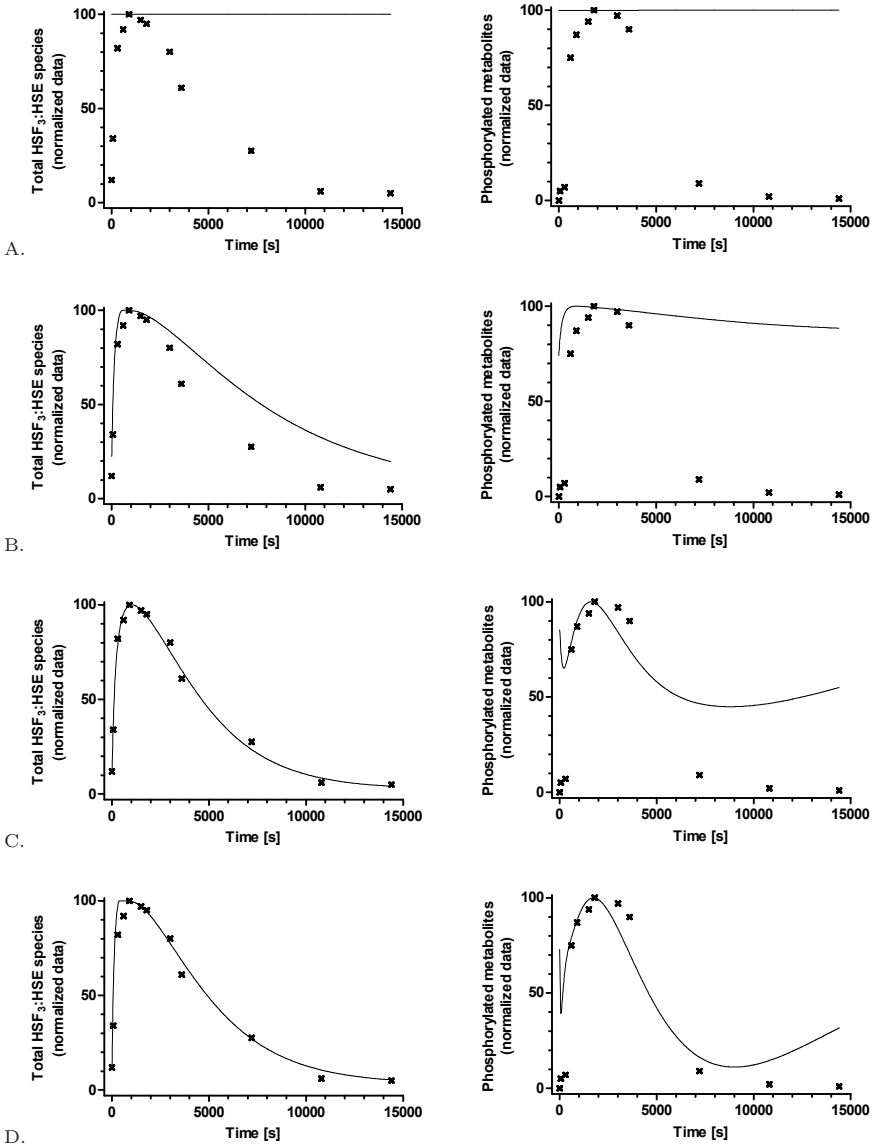

Figure 2: The normalized time-dependent levels of DNA binding and phosphorylated hsf metabolites for the best fits of models $M1.1 - M1.4$, with the fit quality indicated in Table 14. The continuous lines are the model predictions, whereas the experimental data (from [23]) are showed with crossed points. Figures A, B, C, D correspond to models M3.1, M3.2, M3.3 and M3.4, resp.

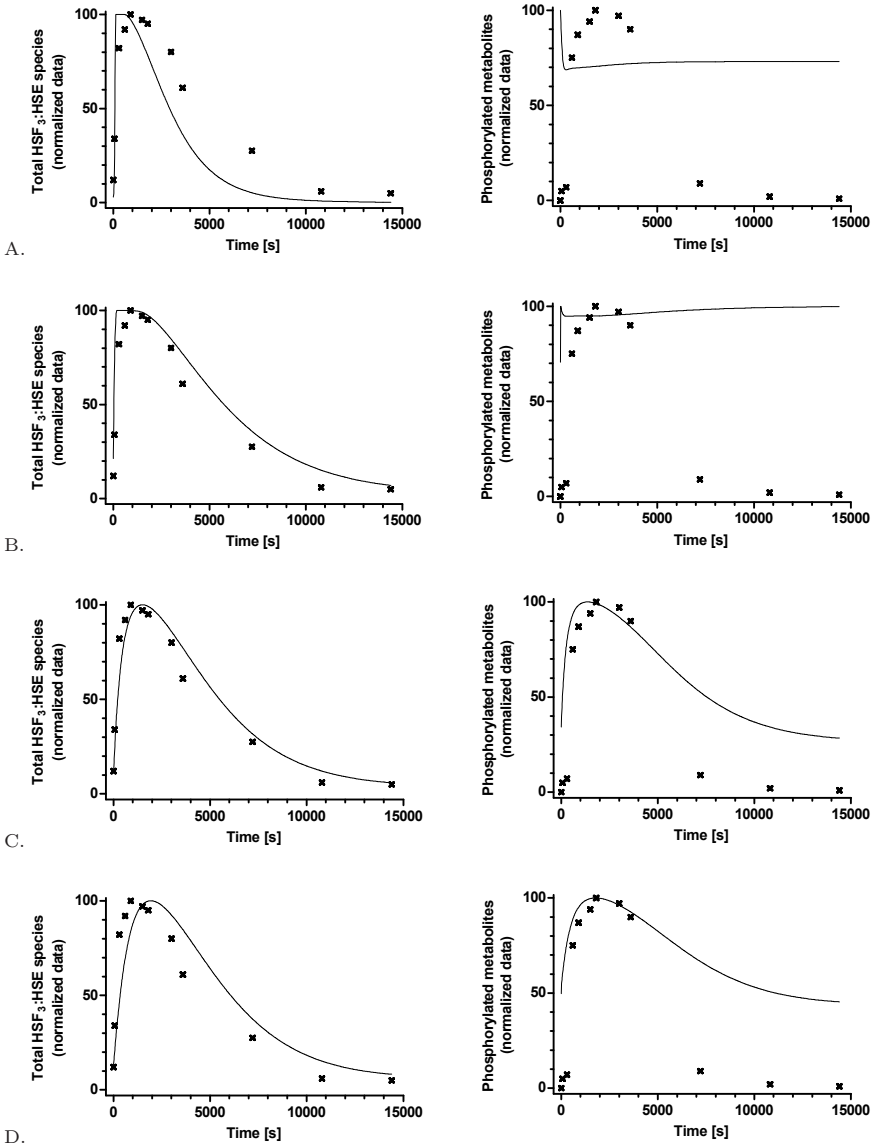

Figure 3: The normalized time-dependent levels of DNA binding and phosphorylated hsf metabolites for the best fits of models $M2.1 - M2.4$, with the fit quality indicated in Table 14. The continuous lines are the model predictions, whereas the experimental data (from [23]) are showed with crossed points. Figures A, B, C, D correspond to models M3.1, M3.2, M3.3 and M3.4, resp.

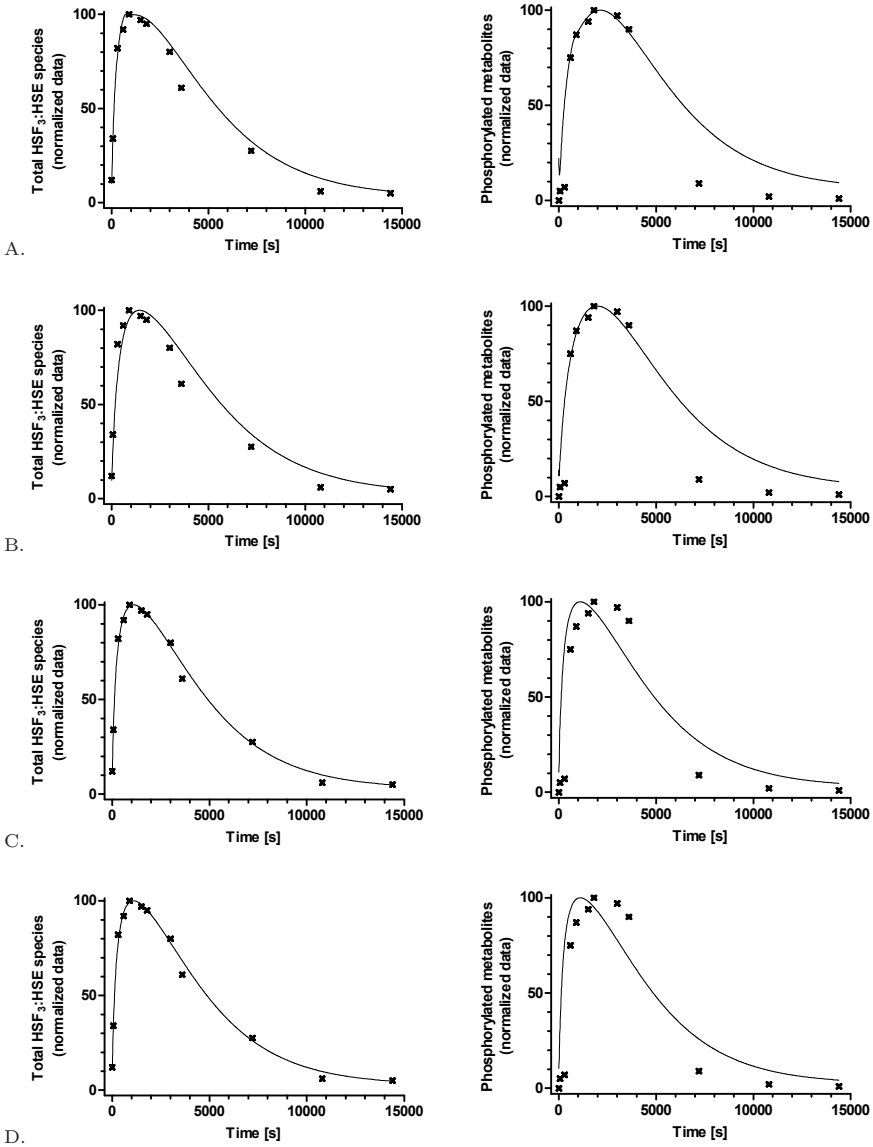

Figure 4: The normalized time-dependent levels of DNA binding and phosphorylated hsf metabolites for the best fits of models $M1.1 - M1.4$, with the fit quality indicated in Table 14. The continuous lines are the model predictions, whereas the experimental data (from [23]) are showed with crossed points. Figures A, B, C, D correspond to models M3.1, M3.2, M3.3 and M3.4, resp.

Turku
Centre *for*
Computer
Science

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi

University of Turku
- Department of Information Technology
- Department of Mathematics

Åbo Akademi University
- Department of Information Technologies

Turku School of Economics
- Institute of Information Systems Sciences

ISBN 978-952-12-2730-1
ISSN 1239-1891