Available online at [www.sciencedirect.com](http://www.sciencedirect.com/)



[Electronic Notes in Theoretical Computer Science 284 (2012) 35–53](http://dx.doi.org/10.1016/j.entcs.2012.05.014)

[www.elsevier.com/locate/entcs](http://www.elsevier.com/locate/entcs)

Quantitative Model Refinement as a Solution to the Combinatorial Size Explosion of Biomodels

Elena Czeizler[1](#_bookmark0)

*Research Programs Unit, Genome-Scale Biology, Faculty of Medicine, University of Helsinki FI-00014 Helsinki , Finland*

Eugen Czeizler[1](#_bookmark0)

*Department of Information and Computer Science, Aalto University School of Science*

*FI-00076 Aalto, Finland*

Bogdan Iancu Ion Petre

*Computational Biomodelling Laboratory*

*A˚bo Akademi University and Turku Centre for Computer Science FI-20520 Turku, Finland*

**Abstract**

Building a large system through a systematic, step-by-step refinement of an initial abstract specification is a well established technique in software engineering, not yet much explored in systems biology. In the case of systems biology, one starts from an abstract, high-level model of a biological system and aims to add more and more details about its reactants and/or reactions, through a number of consecutive refinement steps. The refinement should be done in a quantitatively correct way, so that (some of) the numerical properties of the model (such as the experimental fit and validation) are preserved. In this study, we focus on the data-refinement mechanism where the aim is to increase the level of details of some of the reactants of a given model. That is, we analyse the case when a model is refined by substituting a given species by several types of subspecies. We show in this paper how the refined model can be systematically obtained from the original one. As a case study for this methodology we choose a recently introduced model for the

eukaryotic heat shock response, [[19](#_bookmark18)]. We refine this model by including details about the acetylation of the heat shock factors and its influence on the heat shock response. The refined model has a significantly higher number of kinetic parameters and variables. However, we show that our methodology allows us to preserve the experimental fit/validation of the model with minimal computational effort.

*Keywords:* Model refinement, quantitative analysis, heat shock response, acetylation.

1 Work partially done while the authors were at ˚Abo Akademi University.

1571-0661 © 2012 Elsevier B.V. Open access under [CC BY-NC-ND license.](http://creativecommons.org/licenses/by-nc-nd/3.0/)

[doi:10.1016/j.entcs.2012.05.014](http://dx.doi.org/10.1016/j.entcs.2012.05.014)

# Introduction

Extensive experimental and computational effort is invested nowadays in compil- ing large, system-level models for complex biological systems, including regulatory networks, signaling pathways, metabolic pathways etc. The first step in developing such system-level models lies usually in creating an abstraction of the biological process consisting of a relatively small number of biochemical reactions describing the main mechanisms of the considered process. The chosen reactions can be ab- stract representations of some particular subprocess, encapsulating, in fact, many biochemical reactions from the considered system. A mathematical model is then associated to the molecular model comprising these chosen reactions. For this, one chooses an appropriate kinetic law, e.g., mass-action law or Michaelis-Menten kinet- ics, based on which one can then write the mathematical equations describing the dynamics of the system. The numerical setup of this mathematical model is either obtained from the literature, or, using available experimental data, it is derived through various computational model fit procedures.

Starting from this abstract model, the process of model development proceeds with a series of iterative steps involving hypothesis generation, experimental design, experimental analysis and model refinement, [[2](#_bookmark4)], [[12](#_bookmark15)]. Particularly, the simplifica- tions and abstractions included in the initial model might be refined later on, in- cluding more accurate details of the process. One approach for this would be to simply repeat the whole model development procedure in order to include all the in- tended changes. However, this can be extremely inefficient since it requires to re-fit the model, a step which is both time-consuming and computationally-intensive, [[3](#_bookmark5)]. Another approach, that seems little investigated so far, is to refine the initial model step-by-step making sure that the experimental (numerical) model fit is preserved. In other words, the numerical setup of the refined model should be obtained from that of the initial model, see [[16](#_bookmark19)] for a recent case study regarding self-assembly models.

In this study, we focus on the model refinement step within the above model de- velopment cycle. In particular, we analyse the case when the model is adjusted by refining some of its reactants, i.e., by replacing a given species by some of its subspecies. This is the case, for instance, when more details about the post- translational modifications of proteins, e.g., acetylation or phosphorylation, are re- quired. In such a case, the model is refined by replacing a given protein *P* with its variants indicating whether *P* withstood some post-translational modification or not, e.g., whether *P* was acetylated or not. This substitution of species also leads to a refinement of all the reactions in which protein *P* was involved. Then, we show how we can attain the values of the parameters of the refined model from those of the initial model. Hence, we make sure that we preserve all previously obtained systemic properties, such as numerical fit of the model.

The paper is organized as follows. First, we present the model refinement proce- dure consisting in replacing a species with several subspecies while still preserving the previously obtained model fit with respect to some experimental data. Then, we illustrate this technique by considering as a case study the regulatory role of

protein acetylation within the eukaryotic heat shock response. In particular, we re- fine a recent model for the eukaryotic heat shock response, see [[19](#_bookmark18)], by considering also the acetylation of some particular proteins.

# Quantitative model refinement

There are several types of refinement that can be applied on a given model. For instance, one could focus on the data of a given model and refine it by replacing one (or more) of its species with a number of subspecies. This way, the refined model would illustrate various differences in the behaviour of those subspecies. We call this *data reﬁnement*. Another type of refinement, that we call *process reﬁne- ment*, concerns the model reactions. That is, the model is refined by replacing a generic reaction describing a particular process with several reactions detailing on the intermediary steps of that process.

The problem of formal refinement has been considerably documented in the field of software engineering, particularly related to the concurrent computing paradigm [[24](#_bookmark26)], [[1](#_bookmark6)]. It aroused from the need to prove (in a formal/mathematical way) that the final implementation of a system corresponds to its original specifica- tions. In systems biology, the problem of quantitative model refinement has already been investigated within the framework of rule-based modelling, see [[17](#_bookmark20)], [[10](#_bookmark13)], [[5](#_bookmark8)]. This approach is not concerned with the *data reﬁnement* we have previously de- fined, since this type of refinement is already built in the system through agent resolution [[10](#_bookmark13)]. However, the main consideration in this respect is rule refinement, a method to refine set rules so that the dynamic behaviour of the model is pre- served [[17](#_bookmark20)]. Rule based modelling allows the construction of more detailed models through resolution augmentation in the space of agents and rules, comprising model variants in more finely-grained hierarchical structures [[10](#_bookmark13)]. This type of framework enables writing feasible models and operates on their perturbation spaces in order to evaluate perturbation effects over a particular model [[5](#_bookmark8)].

However, independently of the type of refinement chosen, the refinement process should preserve the systemic properties of the original model, e.g., the model fit. We present here an approach for data model refinement which preserves the previously obtained numerical properties of the model.

Consider that a model *M* consists of a list of *m* species Σ = *{A*1*, A*2*, ..., Am}*

and *n* reactions *ri*, 1 *≤ i ≤ n*, of the form:

*r* : *S*

*A* + *S*

*A* + *...* + *S A*

*ki Sj A*

+ *Sj A*

+ *...* + *Sj A* ,

*i i,*1 1

*i,*2 2

*i,m*

*m −→*

*i,*1 1

*i,*2 2

*i,m m*

where *Si,*1*, ..., Si,m, Sj*

*i,*1

*, ..., Sj*

*≥* 0 are some integers called *the stoichiometric co-*

*efficients* of *ri* and *ki ≥* 0 is *the kinetic rate constant* of *ri*. A number of differ- ent mathematical formulations can be associated to model *M* , in terms of con- tinuous or discrete variables, deterministic or non-deterministic evolution etc. We choose in this study a continuous, mass-action formulation, where to each variable *Ai,* 1 *≤ i ≤ n*, we associate a time-dependent function [*Ai*]: R+ *→* R+ describing its concentration level. In particular, for each species *Ai*, we denote by [*Ai*](*t*) its

*i,m*

concentration at time *t*. The dynamics of the system is then described through a system of differential equations [[13](#_bookmark16)] in which, for each reaction, we assumed the principle of mass action, originally introduced in [[8](#_bookmark11)], [[9](#_bookmark12)].

In particular, for model *M* we obtain the following system of ODE:

*n m n m*

*d*[*Aj* ] = *−* Σ(*k S* Y[*A* ]*Si,k* )+ Σ(*k Sj* Y[*A* ]*Si,k* )*,* 1 *≤ j ≤ m.* (1)

*dt*

*i*

*i,j*

*k*

*i*

*i,j*

*k*

*i*=1

*k*=1

*i*=1

*k*=1

The next lemma tackles the existence and uniqueness of solutions of systems of ODEs derived following the mass action law.

**Lemma 2.1** *[*[*11*](#_bookmark14)*] Given a molecular model and its associated system of ODEs derived based on the principle of mass action, for any ﬁxed initial condition, there exists an interval of the form* [0*, x*)*,x ∈* R+ *∪ {*+*∞} and a solution φ such that any other solution is a restriction of φ.*

Assume that the model M is to be detailed by distinguishing several subspecies of *A*1. Such subspecies may be several different forms of *A*1, several biochemical configurations of *A*1 (e.g., caused by some post-translational modifications) etc. Each of these subspecies may participate in all reactions, where *A*1 participated (in model *M* ), possibly with different kinetics. If *A*1 is to be replaced with subspecies *B*1*,..., Bl*, then we derive a new model *MR*, in which the set of species is denoted

through the new variables *{Aj , Aj , ..., Aj }∪ {B*1*, ..., Bl}*, for some *l ≥* 2. Variables

2 3 *m*

*Aj* , 2 *≤ i ≤ m*, correspond to species *Ai* from model *M* , whereas *B*1*, ..., Bl* are to replace species *A*1 in model *MR*. Moreover, each reaction *ri* from *M* is replaced in *MR* by reaction *rj* of the following type:

*i*

*i*

*rj* : (*Ti,*1*B*1 + *...* + *Ti,lBl*)+ *Si,*2*Aj* + *...* + *Si,mAj*

*k′*

*−→i*

*i*

(*T j*

*B*1 + *...* + *T j Bl*)+ *Sj*

2

*Aj* + *...* + *Sj*

*m*

*Aj* ,

*i,l*

*i,*1

*i,l*

*i,*2 2

*i,m m*

with *kj*

*i*

its kinetic rate constant, and *Ti,*1*, ..., Ti,l,Tj*

*, ..., T j*

nonnegative integers

such that *Ti,*1 + *...* + *Ti,l* = *Si,*1 and *T j*

*i,*1

*i,*1

*i,l*

+ *...* + *T j*

*j*

*i,*1

= *S*

. We say now that *the*

*model MR is a data reﬁnement of M on variable A*1 if and only if the following two

conditions are satisfied:

[*Ai*](*t*)= [*Aj* ](*t*)*,* for all 2 *≤ i ≤ m,* (2)

*i*

[*A*1](*t*)= [*B*1](*t*)+ *...* + [*Bl*](*t*)*,* for all *t ≥* 0*.* (3)

The refined model *MR* has *m* + *l −* 1 species, whereas *M* consists of m species; thus, *MR* has a linear-increase in the size of its data set. The number of reactions replacing in *MR* the reaction *ri* of M is given by the number of non-negative integer solutions of the following system of equations:

*Ti,*1 + *Ti,*2 + *...* + *Ti,l* = *Si,*1;

*i,l*

= *S*

*j*

*j*

*T*

*i,*1

*j*

*i,*2

+ *T*

+ *...* + *T j*

*i,*1;

over the independent unknowns *Ti,j,Tj*

*i,j*

*,* 1 *≤ j ≤ l*. The number of solutions of the

first equation is the *multinomial coefficient* “l multichooses *Si,*1”, see [[6](#_bookmark9)]:

l

= *l* + *Si,*1 *−* 1 = (*l* + *Si,*1 *−* 1)! *.*

*Si,*1

Since the two equations in the system are independent, the number of solutions

! *·* ,

*Si,*1

*Si,*1!(*l −* 1)!

of the system is , *l*

*Si,*1

*l S'*

*i,*1

!. This gives the increase in the number of reactions

in the model refinement from *M* to *MR*. In terms of kinetic parameters, *MR* will

have , *l* ! new free parameters, given by the number of possible combinations of

*Si,*1

*B*1*, ..., Bl* as reactants.

Arguing on the basis of biokinetics, see [[18](#_bookmark21)], we may assume that *Si,*1 *≤* 2: any reaction with a ternary (or higher) stoichiometric coefficient would be so slow that its effects may be ignored. In case *Si,*1 = 1, *MR* will have ,*l* != *l* new kinetic

1

parameters, i.e., a linear increase in the parameter space. In case *Si,*1 = 2, *MR*

will have ,*l* !=l(l+1)/2 new parameters, i.e., a quadratic increase in the parameter

2

space.

Some of the new kinetic parameters of *MR* may be known from the literature, or they can be experimentally measured. For the rest of them, for which no previ- ous knowledge and no *direct* kinetic measurements are available, a computational procedure is needed to calculate them so that ([2](#_bookmark2)) and ([3](#_bookmark2)) 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 wit- nesses a (potentially) quadratic increase in every step of the refinement is compu- tationally very expensive. Moreover, such a procedure makes little sense since the fit of an intermediate model is lost in the next refinement step.

We propose in this paper an approach where we systematically set the values of the unknown kinetic parameters of the refined model so that relations ([2](#_bookmark2)) and ([3](#_bookmark2)) hold. Clearly, some of the potential choices are unreasonable, such as those where we would set to 0 the kinetic parameters of all reactions involving *B*2*, B*3*, ...,* or *Bl*; such a choice would eliminate the idea of refinement and it would only rename all the variables of *M* in *MR*. Instead, we take an approach where the refined subspecies *B*1*, ..., Bl* 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.

Consider next an example of such a data refinement procedure, in which we show how we can obtain the kinetic rate constants of the refined model depending on the type of equations included in the original model. We consider the Lotka-Volterra system, [[15](#_bookmark22)], [[22](#_bookmark25)], composed of the species *A* (the prey) and *B* (the predator), and the following reactions:

*A −k→*1

2*A*, *A* + *B −k→*2

2*B*, *B −k→*3 *∅*.

The set of mass-action based ODEs describing the dynamics of this system are:

*d*[*A*]*/dt* = *k*1[*A*] *− k*2[*A*][*B*] *d*[*B*]*/dt* = *k*2[*A*][*B*] *− k*3[*B*].

After refining the Lotka-Volterra model on variable *A* into subspecies *A*1 and

*A*2, we obtain the following set of reactions:

*A*1 *−r→*1 *A*2 *−r→*4

*A*1 + *A*1, *A*1 *−r→*2

*A*1 + *A*1, *A*2 *−r→*5

*A*1 + *A*2, *A*1 *−r→*3

*A*1 + *A*2, *A*2 *−r→*6

*A*2 + *A*2, *A*2 + *A*2,

*A*1 + *Bj −r→*7

2*Bj*, *A*2 + *Bj −r→*8

2*Bj*, *Bj −r→*9 *φ*.

This leads to the following system of ODEs describing the dynamics of the variables *A*1*, A*2*,* and *Bj*:

*d*[*A*1]*/dt* = (*r*1 *− r*3)[*A*1]+ (2*r*4 + *r*5)[*A*2] *− r*7[*A*1][*Bj*],

*d*[*A*2]*/dt* = (*r*2 + 2*r*3)[*A*1]+ (*−r*4 + *r*6)[*A*2] *− r*8[*A*2][*Bj*],

*d*[*Bj*]*/dt* = *r*7[*A*1][*Bj*]+ *r*8[*A*2][*Bj*] *− r*9[*Bj*].

Thus, *d*([*A*1]+ [*A*2])*/dt* = (*r*1 *− r*3 + *r*2 + 2*r*3)[*A*1]+ (2*r*4 + *r*5 *− r*4 + *r*6)[*A*2] *−* (*r*7 +

*r*8)[*A*1][*Bj*]. Consequently, if we choose *r*1 = *r*2 = *r*3 = *r*4 = *r*5 = *r*6 = *k*1*/*3, *r*7 = *r*8 = *k*2 and *r*9 = *k*3 then we obtain

*d*([*A*1]+ [*A*2])*/dt* = *k*1([*A*1]+ [*A*2]) *− k*2([*A*1]+ [*A*2])[*Bj*],

*d*[*Bj*]*/dt* = *k*2([*A*1]+ [*A*2])[*Bj*] *− k*3[*Bj*],

which is identical to the initial system up to a renaming of variables where [*A*] is replaced by [*A*1]+ [*A*2] and [*B*] is replaced by [*Bj*]. For any *x*0 *≥* 0, if we set the initial values of the variables *A*1, *A*2, and *Bj* such that [*A*](*x*0)= [*A*1](*x*0)+[*A*2](*x*0) and [*B*](*x*0)= [*Bj*](*x*0), it follows from Lemma [2.1](#_bookmark2) that there exists an open interval *I* including *x*0 such that [*A*](*t*)= [*A*1](*t*)+ [*A*2](*t*) and [*B*](*t*)= [*Bj*](*t*), for all *t ∈ I*. That is, the second model is indeed a quantitative refinement of the initial one (on the interval *I*).

# Models

* 1. *The eukaryotic heat shock response: a molecular model*

The heat shock response in eukaryotes is an evolutionarily conserved mechanism that controls the cellular response to proteotoxicity originating from environmental stressors such as elevated temperatures. When subjected to increased tempera- tures, proteins in the cell tend to misfold and accumulate in large aggregates that eventually induce cell death. Survival of the cell is promoted by a mechanism that restores protein homeostatis, i.e. the equilibrium between synthesis, folding and degradation of proteins.

We describe the molecular model for the heat shock response proposed in [[20](#_bookmark23)] as follows. The key factors within this process are the heat shock proteins (hsp), that

act as chaperones, assisting the misfolded proteins (mfp) in their refolding process. The response is regulated by the transactivation of the hsp-encoding genes. The transcription of the gene is mediated by specific heat shock transcription factors (hsf). The hsf’s trimerize (hsf3) and subsequently bind to a promoter site of the hsp-encoding gene, called heat shock element (hse). The trimerization phase of the heat shock factors is preceded by a dimerization stage resulting in the constitu- tion of dimers (hsf2). Hsf trimers bind to heat shock elements forming hsf3: hse complexes. Therefore, protein synthesis is activated and new hsp molecules are ul- timately formed. When the level of hsp’s is sufficiently elevated, hsp synthesis is turned off through an ingenious mechanism [[19](#_bookmark18), [20](#_bookmark23)]. Heat shock proteins bind to the hsf molecules forming hsp: hsf complexes, thus impeding hsf’s to trimerize and to bind to the heat shock elements. The sequestration of hsf’s can be done in three different ways: by breaking dimers, trimers, and by unbinding hsf3 from the heat shock elements combined with the simultaneous breaking of the trimer. However, an increase in the temperature causes some of the proteins (prot) to misfold, which drives hsp away from hsf. This in turn quickly turns on the heat shock response since the heat shock factors are again free and thus able to promote the synthe- sis of more heat shock proteins. The reactions of the molecular model in [[19](#_bookmark18)] are presented in Table [1](#_bookmark3).

Table 1

The molecular model for the eukaryotic heat shock response proposed in [[19](#_bookmark18)].

2 hsf hsf2 hsp + hsf3 *→* hsp: hsf +2 hsf

hsf + hsf2 hsf3 hsp + hsf3: hse *→* hsp: hsf +2 hsf + hse

hsf3 + hse hsf3: hse hsp *→∅*

hsf3: hse *→* hsf3: hse + hsp prot *→* mfp

hsp + hsf hsp: hsf hsp + mfp hsp: mfp hsp + hsf2 *→* hsp: hsf + hsf hsp: mfp *→* hsp + prot

* 1. *Mathematical model*

Associated to the molecular model in Table [1](#_bookmark3), we consider a mathematical model formulated in terms of mass-action based ODEs, see [[13](#_bookmark16)]. The model was originally considered in [[19](#_bookmark18)]. We include the system of ODEs in Table [A.1](#_bookmark28) and we refer the reader to [[19](#_bookmark18)] for more details.

Both the kinetic rate constants and the initial values of all reactants were esti- mated in [[19](#_bookmark18)], by imposing the following three conditions:

1. the system is in a steady state if the temperature is 37*◦C*. This is a natural consequence of the fact that the model should not exhibit any response in the absence of the heat shock, i.e., at 37*◦C*;
2. the numerical predictions of the model for [hsf3: hse](*t*) should agree with the experimental data from [[14](#_bookmark17)], for a temperature of 42*◦C*;
3. the numerical prediction of the model for [hsp](*t*) should confirm the data obtained in [[19](#_bookmark18)] through a de-novo fluorescent reporter-based experiment, for a temperature of 42*◦C*.
   1. *The role of protein acetylation within the eukaryotic heat shock response*

It has been recently shown that the acetylation of the heat shock factors (hsf), i.e., the transcription factors for the hsp-encoding genes, plays an important role in regulating the heat shock response [[23](#_bookmark27)]. The acetylation process consists in substituting an acetyl group for a hydrogen atom within a chemical compound. The reverse process, i.e., the deacetylation, represents the suppression of an acetyl group from a compound. Protein acetylation can occur at the alpha-amino group of the amino-terminus (N-terminal acetylation) or on the lysine residues at the epsilon- amino group (lysine acetylation) [[7](#_bookmark10)]. The lysine acetylation in particular is known to play a significant role in gene regulation by changing the charge of histone tails. Due to the neutralization of the positive charge of the histones, lysine acetylation diminishes their DNA binding affinity [[4](#_bookmark7), [21](#_bookmark24)].

* 1. *Data reﬁnement of the model*

We discuss, in this section, how to extend the heat shock model of [[19](#_bookmark18)] to account for the acetylation of hsf and its influence on the response. For this, we refine all species and complexes involving hsf to account for two subtypes of hsf: one where its K80 residue is not acetylated, and one where it is acetylated. Consequently, the complex hsp: hsf will also be refined to two subtypes, depending on the acetylation of its hsf component. In the case of hsf2, hsf3 and hsf3: hse, the refinement will be accomplished by counting how many of the hsf components in that complex (two in the case of hsf2 and three in the case of hsf3 and hsf3: hse) have their K80 residue acetylated. Thus, we perform the following data refinements:

hsf *→ {*rhsf*,* rhsf(1)*}*;

hsf2 *→ {*rhsf2*,* rhsf2(1)*,* rhsf2(2)*}*;

hsf3 *→ {*rhsf3*,* rhsf3(1)*,* rhsf3(2)*,* rhsf3(3)*}*;

hsf3: hse *→ {*rhsf3: rhse*,* rhsf3(1): rhse*,* rhsf3(2): rhse*,* rhsf3(3): rhse*}*;

hsp: hsf *→ {*rhsp: rhsf*,* rhsp: rhsf(1)*}.*

These data refinements imply several changes in the list of reactions of our model. For example, the reaction hsp + hsf hsp: hsf is replaced by two reactions : rhsp + rhsf rhsp: rhsf, and rhsp +rhsf(1) rhsp: rhsf(1). As another example, reaction 2 hsf hsf2 is replaced by three reactions, based on the method described in Section [2](#_bookmark1):

2 rhsf rhsf2; rhsf + rhsf(1) rhsf2(1); 2rhsf(1) rhsf2(2)*.*

Due to space limitations, the complete list of reactions is given in Appendix [B](#_bookmark29). The refined model consists of 39 reactions, some of which are reversible.

* 1. *Quantitative equivalence of the basic and the reﬁned model*

We discuss now the numerical setup of the refined model for the heat shock response in such a way that the refinement procedure is quantitatively correct as defined in ([2](#_bookmark2))-([3](#_bookmark2)). The initial values of the refined variables are set so that the following 10 conditions (derived from the data refinement relationships) are satisfied:

[hsf](0) = [rhsf](0) + [rhsf(1)](0);

[hsf2](0) = [rhsf2](0) + [rhsf2(1)]+ [rhsf2(2)](0);

[hsf3](0) = [rhsf3](0) + [rhsf3(1)](0) + [rhsf3(2)](0) + [rhsf3(3)](0); [hsp: hsf](0) = [hsp: rhsf](0) + [rhsp: rhsf(1)](0);

[hsf3: hse](0) = [rhsf3: rhse](0) + [rhsf3(1): rhse](0) + [rhsf3(2): rhse](0)+

+ [rhsf3(3): rhse](0);

[hsp](0) = [rhsp](0);

[hsp: mfp](0) = [rhsp: rmfp](0);

[mfp](0) = [rmfp](0);

[prot](0) = [rprot](0);

[hse](0) = [rhse](0)*.*

The system of mass-action based ODEs for the refined model is in Table [C.1](#_bookmark30). The refined model consists of 20 species, 39 reactions (some of them reversible), 54 kinetic parameters. We recall that the basic model had 10 species, 12 reactions and 16 kinetic parameters. We focus now on the numerical setup of the refined model so that it is a quantitative refinement of the model in [[19](#_bookmark18)], as defined in Section [2](#_bookmark1). We first introduce the following notations:

Rhsf = rhsf + rhsf(1);

Rhsf2 = rhsf2 + rhsf2(1) + rhsf2(2);

Rhsf3 = rhsf3 + rhsf3(1) + rhsf3(2) + rhsf3(3);

Rhsf3: Rhse = rhsf3: rhse + rhsf3(1): rhse + rhsf3(2): rhse + rhsf3(3): rhse;

Rhsp: Rhsf = rhsp: rhsf + rhsp: rhsf(1).

We aim to identify some values for the kinetic parameters of the refined model in such a way that its system of differential equations is identical to the system associated to the initial model, modulo a variable renaming where hsf, hsf2, hsf3, hsf3: hse, and hsp: hsf are replaced by Rhsf, Rhsf2, Rhsf3, Rhsf3: Rhse, and Rhsp: Rhsf, respectively. To drive this process, we write the ODEs for [Rhsf], [Rhsf2],[Rhsf3], [Rhsf3: Rhse] and for [Rhsp: Rhsf], see Table [C.2](#_bookmark31), based on the system of ODEs in Table [C.1](#_bookmark30). We then choose the values of the kinetic parameters in such a way that the right hand side of each ODE in Table [C.2](#_bookmark31) becomes identical to the right hand side of the corresponding ODE in Table [A.1](#_bookmark28), modulo the variable renaming above. For example, we aim to choose the kinetic parameters of the refined model in such

a way that the ODE corresponding to Rhsf in Table [C.2](#_bookmark31) is identical to the ODE corresponding to hsf in Table [A.1](#_bookmark28), modulo the variable renaming above. To identify the first term of the ODE for Rhsf (in the form written in Table [C.2](#_bookmark31)) with the first term of the ODE for hsf (in the form written in Table [A.1](#_bookmark28)), it is enough to set *r*+ = *k*+, *r*+ = 2*k*+, and *r*+ = *k*+. A similar reasoning for all terms of all ODEs

1 1 2 1 3 1

leads to a solution.

Due to space limitations, we skip all details and only include the table listing the identified values for the parameters in Appendix [D](#_bookmark32). Clearly, the solution is not unique. However, to find one, we cannot count on solving the systems of ODEs in Tables [A.1](#_bookmark28) and [C.2](#_bookmark31); in fact these systems cannot be solved analytically. Instead, we simply chose the values of the kinetic parameters of the refined model as expressions of the kinetic parameters of the original model, in such a way that the ODEs in Table [C.2](#_bookmark31) can be rewritten to the equations in Table [A.1](#_bookmark28), modulo the variable renaming above. Based also on the way we set up the initial values of the refined variables, it follows that the systems of differential equations associated to the initial model and to the refined one have identical initial conditions. Thus, it follows from Lemma [2.1](#_bookmark2) that conditions ([2](#_bookmark2)) and ([3](#_bookmark2)) are satisfied, i.e., the second model is indeed a quantitative refinement of the initial model in Table [1](#_bookmark3).

# Discussion

We focus in this paper on quantitative model refinement as an essential stage within the complex process of model development in systems biology. In particular, we analyse the case when the model is refined by replacing one species with several subspecies. Specifically, we show that, by attaining the numerical setup of the refined model from that of the initial model, we obtain a quantitative refinement which preserves previously obtained numerical properties of the model, e.g., model fit and validation, see [[3](#_bookmark5)].

We refined the model from [[19](#_bookmark18)] by considering only one acetylation site for each hsf molecule. This led to a significant augmentation in the number of reactions of the refined model and, consequently, in the number of parameters. While the basic model comprises a set of 12 reactions involving 10 different species and 16 kinetic rate constants, the refined model contains a number of 39 reactions involving 20 species and 54 kinetic rate constants. Fitting a model of this proportion implies a lot of time and computational resources. However, with our approach, we were able to build a refined model, with a satisfactory numerical behaviour (as defined in our notion of quantitative refinement), avoiding any supplementary model fit.

Our solution to the problem of setting the kinetic parameters of the refined model is clearly not unique. 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 where we make sure that the two systems of ODEs are isomorphic through the variable renaming given by the data refinement relationship. In the absence of any 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 constraint in our symbolic approach. The existence of a solution in such a case, as well as effectively constructing one appear as interesting problems in this context.

We only focused in this paper on refining the basic heat shock response model of [[19](#_bookmark18)] in a quantitatively correct way to include some of the details of the acetylation of hsf. Due to lack of space, we did not include in the refined model the details regarding the role that the acetylation of hsf plays in fine-tuning the heat shock response, as described in [[23](#_bookmark27)]. We plan to return to these aspects in a separate study.

# Acknowledgement

We are grateful to the anonymous referees for their detailed comments that helped improve our presentation. This work has been partially supported by Academy of Finland (grants 129863, 108421, and 122426) and the Foundation of ˚Abo Akademi.

# References

1. L. Aceto, *Action refinement in process algebras*, Distinguished Dissertations in Computer Science, Cambridge University Press, 1992.
2. F. J. Bruggeman and H. V. Westerhoff. The nature of systems biology. Trends in Microbiology, 15(1):45- 50, 2007.
3. W. W. Chen, B. Schoeberl, P. J. Jasper, M. Niepel, U. B. Nielsen, D. A. Lauffenburger, P. K. Sorger. Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. Molecular Systems Biology, 5(239), 2009.
4. C. Choudhary, C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. C. Walther, J. V. Olsen, M. Mann, Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. Science, 325(5942), 834-840, 2009.
5. 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.
6. I. M. Gessel and R. P. Stanley (1995) ”Algebraic enumeration” in Graham, R. L., M. Gro¨tschel, L. Lovasz (Eds.), Handbook of combinatorics, Vol. 2. Elsevier: 10211061.
7. M. A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins, In: Gene, 363: 15-23, 2005.
8. C. M. Guldberg and P. Waage, Studies Concerning Affinity. *C. M. Forhandlinger: Videnskabs-Selskabet* *i Christiana* **35**, 1864.
9. C. M. Guldberg and P. Waage, Concerning Chemical Affinity. *Erdmann’s Journal fu¨r Practische Chemie*

**127**: 69-114, 1879.

1. 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.
2. M. W. Hirsch, S. Smale, R. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*, Second Edition, Academic Press, 2004.
3. H. Kitano. Systems biology: A brief overview. Science, 295(5560):1662-1664, 2002.
4. E. Klipp, R. Herwig, A. Kowald, C. Wierling, H. Lehrach, Systems Biology in Practice, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005.
5. M. P. Kline, R. I. Morimoto, Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. Molecular and cellular biology,17(4),21072115, 1997.
6. A. J. Lotka, Elements of physical biology. Williams and Wilkins, Baltimore, 1925.
7. A. Mizera, E. Czeizler, I. Petre, Self-assembly models of variable resolution, submitted, 2011.
8. E. Murphy, V. Danos, J. Feret, J. Krivine, R. Harmer, Rule-based modelling and model refinement. In Huma M. Lodhi and Stephen H. Muggleton, editors, *Elements of Computational Systems Biology.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010.
9. D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry* Third Edition, Worth Publishers, New York, 2000.
10. I. Petre, A. Mizera, C. L. Hyder, A. Meinander, A. Mikhailov, R.I. Morimoto, L. Sistonen, J. E. Eriksson, 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.
11. T. R. Rieger, R. I. Morimoto, V. Hatzimanikatis, Mathematical Modeling of the Eukariotic Heat-Shock Response: Dynamics of the hsp70 Promoter, Biophysical Journal, vol. 88, 1646-1658, 2005.
12. K. Struhl, Histone acetylation and transcriptional regulatory mechanisms,Genes & Development, 12:599-606, 1998.
13. V. Volterra, In R. N. Chapman 1931. Animal ecology. McGraw-Hill, New York, 409-448, 1926.
14. S. D. Westerheide, J. Anckar, S. M. Stevens Jr., L. Sistonen, R. I. Morimoto, Stress-Inducible Regulation of Heat Shock Factor 1 by the Deactylase SIRT1, Science, vol. 323, 2009.
15. N. Wirth, Program Development by Stepwise Refinement, Communications of the ACM, vol. 14, iss. 4, New York, USA, 1971.

# The original ODE-based model for the heat shock response

Table A.1

The system of ODEs associated with the biochemical model proposed in [[19](#_bookmark18)].

*d*[hsf]*/dt* = *−*2*k*+[hsf]2 + 2*k−*[hsf2] *− k*+[hsf][hsf2]+ *k−*[hsf3]

1 1 2 2

*−k*+[hsf][hsp]+ *k−*[hsp: hsf]+ *k*6[hsf2][hsp]

5 5

+2*k*7[hsf3][hsp]+ 2*k*8[hsf3: hse][hsp]

*d*[hsf2]*/dt* = *k*+[hsf]2 *− k−*[hsf2] *− k*+[hsf][hsf2]+ *k−*[hsf3]

1 1 2 2

*−k*6[hsf2][hsp]

*d*[hsf3]*/dt* = *k*+[hsf][hsf2] *− k−*[hsf3] *− k*+[hsf3][hse]+ *k−*[hsf3: hse]

2 2 3 3

*−k*7[hsf3][hsp]

*d*[hse]*/dt* = *−k*+[hsf3][hse]+ *k−*[hsf3: hse]+ *k*8[hsf3: hse][hsp]

3 3

*d*[hsf3: hse]*/dt* = *k*+[hsf3][hse] *− k−*[hsf3: hse] *− k*8[hsf3: hse][hsp]

3 3

*d*[hsp]*/dt* = *k*4[hsf3: hse] *− k*+[hsf][hsp]+ *k−*[hsp: hsf] *− k*6[hsf2][hsp]

5 5

*−k*7[hsf3][hsp] *− k*8[hsf3: hse][hsp] *− k*+ [hsp][mfp]

11

+(*k−* + *k*12)[hsp: mfp] *− k*9[hsp]

11

*d*[hsp: hsf]*/dt* = *k*+[hsf][hsp] *− k−*[hsp: hsf]+ *k*6[hsf2][hsp]

5 5

+*k*7[hsf3][hsp]+ *k*8[hsf3: hse][hsp]

*d*[mfp]*/dt* = *φT* [prot] *− k*+ [hsp][mfp]+ *k−* [hsp: mfp]

11 11

*d*[hsp: mfp]*/dt* = *k*+ [hsp][mfp] *−* (*k−* + *k*12)[hsp: mfp]

11 11

*d*[prot]*/dt* = *−φT* [prot]+ *k*12[hsp: mfp]

# The refined model for the heat shock response

Table B.1: The list of reactions for the refined model. For an irreversible reaction *qi*, *ri* denotes its kinetic rate constant. For a reversible reaction *qi*, we denote by *r*+ and *r−* the kinetic rate constants of its

*i* *i*

‘left-to-right’ and ‘right-to-left’ directions, respectively.

Reaction

Kinetic rate constants

2 rhsf rhsf2 *r*+*, r−*

1 1

rhsf + rhsf(1) rhsf2(1) *r*+*, r−*

2 2

2rhsf(1) rhsf2(2) *r*+*, r−*

3 3

rhsf + rhsf2 rhsf3 *r*+*, r−*

4 4

rhsf(1) + rhsf2 rhsf3(1) *r*+*, r−*

5 5

rhsf +rhsf2(1) rhsf3(1) *r*+*, r−*

6 6

rhsf(1) + rhsf2(1) rhsf3(2) *r*+*, r−*

7 7

rhsf +rhsf2(2) rhsf3(2) *r*+*, r−*

8 8

rhsf(1) + rhsf2(2) rhsf3(3) *r*+*, r−*

9 9

rhsf3 + rhse rhsf3: rhse *r*+ *, r−*

10 10

rhsf3(1) + rhse rhsf3(1): rhse *r*+ *, r−*

11 11

rhsf3(2) + rhse rhsf3(2): rhse *r*+ *, r−*

12 12

rhsf3(3) + rhse rhsf3(3): rhse *r*+ *, r−*

13 13

rhsf3: rhse *→* rhsf3: rhse + rhsp *r*14

rhsf3(1): rhse *→* rhsf3(1): rhse + rhsp *r*15

rhsf3(2): rhse *→* rhsf3(2): rhse + rhsp *r*16

rhsf3(3): rhse *→* rhsf3(3): rhse + rhsp *r*17

rhsp + rhsf rhsp: rhsf *r*+ *, r−*

18 18

rhsp +rhsf(1) rhsp: rhsf(1) *r*+ *, r−*

19 19

rhsp + rhsf2 *→* rhsp: rhsf + rhsf *r*20

Table B.1: The list of reactions for the refined model - Continued

rhsp +rhsf2(1) *→* rhsp: rhsf +rhsf(1) *r*21

rhsp +rhsf2(1) *→* rhsp: rhsf(1) + rhsf *r*22

rhsp +rhsf2(2) *→* rhsp: rhsf(1) +rhsf(1) *r*23

rhsp + rhsf3 *→* rhsp: rhsf +2 *∗* rhsf *r*24

rhsp +rhsf3(1) *→* rhsp: rhsf +rhsf(1) + rhsf *r*25

rhsp +rhsf3(1) *→* rhsp: rhsf(1) +2 *∗* rhsf *r*26

rhsp +rhsf3(2) *→* rhsp: rhsf +2rhsf(1) *r*27

rhsp +rhsf3(2) *→* rhsp: rhsf(1) +rhsf(1) + rhsf *r*28

rhsp +rhsf3(3) *→* rhsp: rhsf(1) +2rhsf(1) *r*29

rhsp + rhsf3: rhse *→* rhsp: rhsf +2 rhsf + rhse *r*30 rhsp + rhsf3(1): rhse *→* rhsp: rhsf(1) +2 rhsf + rhse *r*31 rhsp + rhsf3(1): rhse *→* rhsp: rhsf +rhsf(1) + rhsf + rhse *r*32 rhsp + rhsf3(2): rhse *→* rhsp: rhsf(1) +rhsf(1) + rhsf + rhse *r*33 rhsp + rhsf3(2): rhse *→* rhsp: rhsf +2rhsf(1) + rhse *r*34 rhsp + rhsf3(3): rhse *→* rhsp: rhsf(1) +2rhsf(1) + rhse *r*35 rhsp *→∅ r*36

rprot *→* rmfp *r*37

rhsp + rmfp rhsp: rmfp *r*+ *, r−*

38 38

rhsp: rmfp *→* rhsp + rprot *r*39

# The ODE-based model of the refined heat shock re- sponse model

Table C.1: The system of differential equations of the mathematical model associated with the refined heat shock response model

*d*[rhsf]*/dt* = *−*2*r*+[rhsf]2 + 2*r−*[rhsf2] *− r*+[rhsf][rhsf(1)]+ *r−*[rhsf2(1)]

1 1 2 2

*−r*4+[rhsf][rhsf2]+ *r*4*−*[rhsf3] *− r*6+[rhsf][rhsf2(1)]+ *r*6*−*[rhsf3(1)]

*−r*8+[rhsf][rhsf2(2)]+ *r*8*−*[rhsf3(2)] *− r*18+[rhsp][rhsf]

+*r*18*−*[rhsp: rhsf]+ *r*20[rhsp][rhsf2]+ *r*22[rhsp][rhsf2(1)]

+2*r*24[rhsp][rhsf3]+ *r*25[rhsp][rhsf3(1)]+ 2*r*26[rhsp][rhsf3(1)]

+*r*28[rhsp][rhsf3(2)]+ 2*r*30[rhsp][rhsf3: rhse]

+2*r*31[rhsp][rhsf3(1): rhse]+ *r*32[rhsp][rhsf3(1): rhse]

+*r*33[rhsp][rhsf3(2): rhse]

*d*[rhsf(1)]*/dt* = *−r*+[rhsf][rhsf(1)]+ *r−*[rhsf (1)] *−* 2*r* +[rhsf(1)]2

2 2 2 3

+2*r*3*−*[rhsf2(2)] *− r*5+[rhsf(1)][rhsf2]+ *r*5*−*[rhsf3(1)]

*−r*7+[rhsf(1)][rhsf2(1)]+ *r*7*−*[rhsf3(2)] *− r*9+[rhsf(1)][rhsf2(2)]

+*r*9*−*[rhsf3(3)] *− r*19+[rhsp][rhsf(1)]+ *r*19*−*[rhsp: rhsf(1)]

+*r*21[rhsp][rhsf2(1)]+ *r*23[rhsp][rhsf2(2)]+ *r*25[rhsp][rhsf3(1)]

+2*r*27[rhsp][rhsf3(2)]+ *r*28[rhsp][rhsf3(2)]+ 2*r*29[rhsp][rhsf3(3)]

+*r*32[rhsp][rhsf3(1): rhse]+ *r*33[rhsp][rhsf3(2): rhse]

+2*r*34[rhsp][rhsf3(2): rhse]+ 2*r*35[rhsp][rhsf3(3): rhse]

*d*[rhsf2]*/dt* = *r*+[rhsf]2 *− r−*[rhsf2] *− r*+[rhsf][rhsf2]+ *r−*[rhsf3]

1 1 4 4

*−r*+[rhsf(1)][rhsf2]+ *r−*[rhsf3(1)] *− r*20[rhsp][rhsf2]

5 5

*d*[rhsf2(1)]*/dt* = *r*+[rhsf][rhsf(1)] *− r−*[rhsf2(1)] *− r*+[rhsf][rhsf2(1)]

2 2 6

+*r−*[rhsf3(1)] *− r*+[rhsf(1)][rhsf2(1)]+ *r−*[rhsf3(2)]

6 7 7

*−r*21[rhsp][rhsf2(1)] *− r*22[rhsp][rhsf2(1)]

*d*[rhsf2(2)]*/dt* = *r*+[rhsf(1)]2 *− r−*[rhsf2(2)] *− r*+[rhsf][rhsf2(2)]

3 3 8

+*r−*[rhsf3(2)] *− r*+[rhsf(1)][rhsf2(2)]+ *r−*[rhsf3(3)]

8 9 9

*−r*23[rhsp][rhsf2(2)]

*d*[rhsf3]*/dt* = *r*+[rhsf][rhsf2] *− r−*[rhsf3] *− r*+ [rhsf3][rhse]

4 4 10

+*r−* [rhsf3: rhse] *− r*24[rhsp][rhsf3]

10

Table C.1: The system of differential equations of the mathematical model associated with the biochemical model - Continued

*d*[rhsf3(1)]*/dt* = *r*+[rhsf(1)][rhsf2] *− r−*[rhsf3(1)]+ *r*+[rhsf][rhsf2(1)]

5 5 6

*−r−*[rhsf3(1)] *− r*+ [rhsf3(1)][rhse]+ *r−* [rhsf3(1): rhse]

6 11 11

*−r*25[rhsp][rhsf3(1)] *− r*26[rhsp][rhsf3(1)]

*d*[rhsf3(2)]*/dt* = *r*+[rhsf(1)][rhsf2(1)] *− r−*[rhsf3(2)]+ *r*+[rhsf][rhsf2(2)]

7 7 8

*−r−*[rhsf3(2)] *− r*+ [rhsf3(2)][rhse]+ *r−* [rhsf3(2): rhse]

8 12 12

*−r*27[rhsp][rhsf3(2)] *− r*28[rhsp][rhsf3(2)]

*d*[rhsf3(3)]*/dt* = *r*+[rhsf(1)][rhsf2(2)] *− r−*[rhsf3(3)] *− r*+ [rhsf3(3)][rhse]

9 9 13

+*r−* [rhsf3(3): rhse] *− r*29[rhsp][rhsf3(3)]

13

*d*[rhse]*/dt* = *−r*+ [rhsf3][rhse]+ *r−* [rhsf3: rhse] *− r*+ [rhsf3(1)][rhse]

10 10 11

+*r−* [rhsf3(1): rhse] *− r*+ [rhsf3(2)][rhse]+ *r−* [rhsf3(2): rhse]

11 12 12

*−r*+ [rhsf3(3)][rhse]+ *r−* [rhsf3(3): rhse]+ *r*30[rhsp][rhsf3: rhse]

13 13

+*r*31[rhsp][rhsf3(1): rhse]+ *r*32[rhsp][rhsf3(1): rhse]

+*r*33[rhsp][rhsf3(2): rhse]+ *r*34[rhsp][rhsf3(2): rhse]

+*r*35[rhsp][rhsf3(3): rhse]

*d*[rhsf3: rhse]*/dt* = *r*+ [rhsf3][rhse] *− r−* [rhsf3: rhse]

10 10

*−r*30[rhsp][rhsf3: rhse]

*d*[rhsf3(1): rhse]*/dt* = *r*+ [rhsf3(1)][rhse] *− r−* [rhsf3(1): rhse]

11 11

*−r*31[rhsp][rhsf3(1): rhse] *− r*32[rhsp][rhsf3(1): rhse]

*d*[rhsf3(2): rhse]*/dt* = *r*+ [rhsf3(2)][rhse] *− r−* [rhsf3(2): rhse]

12 12

*−r*33[rhsp][rhsf3(2): rhse] *− r*34[rhsp][rhsf3(2): rhse]

*d*[rhsf3(3): rhse]*/dt* = *r*+ [rhsf3(3)][rhse] *− r−* [rhsf3(3): rhse]

13 13

*−r*35[rhsp][rhsf3(3): rhse]

*d*[rhsp]*/dt* = *r*14[rhsf3: rhse]+ *r*15[rhsf3(1): rhse]+ *r*16[rhsf3(2): rhse]

+*r*17[rhsf3(3): rhse] *− r*+ [rhsp][rhsf]+ *r−* [rhsp: rhsf]

18 18

*−r*+ [rhsp][rhsf(1)]+ *r−* [rhsp: rhsf(1)] *− r*20[rhsp][rhsf2]

19 19

*−r*21[rhsp][rhsf2(1)] *− r*22[rhsp][rhsf2(1)] *− r*23[rhsp][rhsf2(2)]

*−r*24[rhsp][rhsf3] *− r*25[rhsp][rhsf3(1)] *− r*26[rhsp][rhsf3(1)]

*−r*27[rhsp][rhsf3(2)] *− r*28[rhsp][rhsf3(2)] *− r*29[rhsp][rhsf3(3)]

Table C.1: The system of differential equations of the mathematical model associated with the biochemical model - Continued

*−r*30[rhsp][rhsf3: rhse] *− r*31[rhsp][rhsf3(1): rhse]

*−r*32[rhsp][rhsf3(1): rhse] *− r*33[rhsp][rhsf3(2): rhse]

*−r*34[rhsp][rhsf3(2): rhse] *− r*35[rhsp][rhsf3(3): rhse]

*−r*36[rhsp] *− r*+ [rhsp][rmfp]+ *r−* [rhsp: rmfp]

38 38

+*r*39[rhsp][rmfp]

*d*[rhsp: rhsf]*/dt* = *r*+ [rhsp][rhsf] *− r−* [rhsp: rhsf]+ *r*20[rhsp][rhsf2]

18 18

+*r*21[rhsp][rhsf2(1)]+ *r*24[rhsp][rhsf3]+ *r*25[rhsp][rhsf3(1)]

+*r*27[rhsp][rhsf3(2)]+ *r*30[rhsp][rhsf3: rhse]

+*r*32[rhsp][rhsf3(1): rhse]+ *r*34[rhsp][rhsf3(2): rhse]

*d*[rhsp: rhsf(1)]*/dt* = *r*19+[rhsp][rhsf(1)] *− r*19*−*[rhsp: rhsf(1)]

+*r*22[rhsp][rhsf2(1)]+ *r*23[rhsp][rhsf2(2)]

+*r*26[rhsp][rhsf3(1)]+ *r*28[rhsp][rhsf3(2)]

+*r*29[rhsp][rhsf3(3)]+ *r*31[rhsp][rhsf3(1): rhse]

+*r*33[rhsp][rhsf3(2): rhse]+ *r*35[rhsp][rhsf3(3): rhse]

*d*[rhsp: rmfp]*/dt* = *r*+ [rhsp][rmfp] *−* (*r−* + *r*39)[rhsp: rmfp]

38 38

*d*[rmfp]*/dt* = *r*37[rprot] *− r*+ [rhsp][rmfp]+ *r−* [rhsp: rmfp]

38 38

*d*[rprot]*/dt* = *−r*37[rprot]+ *r*39[rhsp: rmfp]

Table C.2: The ODEs corresponding to Rhsf, Rhsf2, Rhsf3, Rhsf3: Rhse, and Rhsp: Rhsf in the refined model

*d*[Rhsf]*/dt* = *−*2(*r*+[rhsf]2 + *r*+[rhsf][rhsf(1)]+ *r*3+[rhsf(1)]2)+ 2(*r−*[rhsf2]

1 2 1

+*r−*[rhsf2(1)]+ *r*3*−*[rhsf2(2)]) *−* (*r*4+[rhsf][rhsf2]+ *r*6+[rhsf][rhsf2(1)]

2

+*r*8+[rhsf][rhsf2(2)]+ *r*5+[rhsf(1)][rhsf2]+ *r*7+[rhsf(1)][rhsf2(1)]

+*r*9+[rhsf(1)][rhsf2(2)]) + (*r*4*−*[rhsf3]+ (*r*5*−* + *r*6*−*)[rhsf3(1)]

+(*r*7*−* + *r*8*−*)[rhsf3(2)]+ *r*9*−*[rhsf3(3)]) *−* [rhsp](*r*18+[rhsf]

+*r*19+[rhsf(1)]) + (*r*18*−*[rhsp: rhsf]+ *r*19*−*[rhsp: rhsf(1)])

+[rhsp](*r*20[rhsf2]+ (*r*21 + *r*22)[rhsf2(1)]+ *r*23[rhsf2(2)])

+2[rhsp](*r*24[rhsf3]+ (*r*25 + *r*26)[rhsf3(1)]+ (*r*27 + *r*28)[rhsf3(2)]

Table C.2: The ODEs corresponding to Rhsf, Rhsf2, Rhsf3, Rhsf3: Rhse, and Rhsp: Rhsf in the refined model

- Continued

+*r*29[rhsf3(3)]) + 2[rhsp](*r*30[rhsf3: rhse]+ (*r*31 + *r*32)[rhsf3(1): rhse]

+(*r*33 + *r*34)[rhsf3(2): rhse]+ *r*35[rhsf3(3): rhse])

*d*[Rhsf2]*/dt* = (*r*+[rhsf]2 + *r*+[rhsf][rhsf(1)]+ *r*+[rhsf(1)]2) *−* (*r−*[rhsf2]

1 2 3 1

+*r−*[rhsf2(1)]+ *r−*[rhsf2(2)]) *−* (*r*+[rhsf][rhsf2]+ *r*+[rhsf][rhsf2(1)]

2 3 4 6

+*r*+[rhsf][rhsf2(2)]+ *r*+[rhsf(1)][rhsf2]+ *r*+[rhsf(1)][rhsf2(1)]

8 5 7

+*r*+[rhsf(1)][rhsf2(2)]) + (*r−*[rhsf3]+ (*r−* + *r−*)[rhsf3(1)])

9 4 5 6

+(*r−* + *r−*)[rhsf3(2)]+ *r−*[rhsf3(3)]) *−* [rhsp](*r*20[rhsf2]

7 8 9

+(*r*21 + *r*22)[rhsf2(1)]+ *r*23[rhsf2(2)])

*d*[Rhsf3]*/dt* = (*r*+[rhsf][rhsf2]+ *r*+[rhsf][rhsf2(1)]+ *r*+[rhsf][rhsf2(2)]

4 6 8

+*r*+[rhsf(1)][rhsf2]+ *r*+[rhsf(1)][rhsf2(1)]+ *r*+[rhsf(1)][rhsf2(2)])

5 7 9

*−*(*r−*[rhsf3]+ (*r−* + *r−*)[rhsf3(1)]+ (*r−* + *r−*)[rhsf3(2)]+ *r−*[rhsf3(3)])

4 5 6 7 8 9

*−*[rhse](*r*+ [rhsf3]+ *r*+ [rhsf3(1)]+ *r*+ [rhsf3(2)]+ *r*+ [rhsf3(3)])

10 11 12 13

+(*r−* [rhsf3: rhse]+ *r−* [rhsf3(1): rhse]+ *r−* [rhsf3(2): rhse]

10 11 12

+*r−* [rhsf3(3): rhse]) *−* [rhsp](*r*24[rhsf3]+ (*r*25 + *r*26)[rhsf3(1)]

13

+(*r*27 + *r*28)[rhsf3(2)]+ *r*29[rhsf3(3)])

*d*[Rhsf3: Rhse]*/dt* = [rhse](*r*+ [rhsf3]+ *r*+ [rhsf3(1)]+ *r*+ [rhsf3(2)]

10 11 12

+*r*+ [rhsf3(3)]) *−* (*r−* [rhsf3: rhse]+ *r−* [rhsf3(1): rhse]+ *r−* [rhsf3(2): rhse]

13 10 11 12

+*r−* [rhsf3(3): rhse]) *−* [rhsp](*r*30[rhsf3: rhse]+ (*r*31 + *r*32)[rhsf3(1): rhse]

13

+(*r*33 + *r*34)[rhsf3(2): rhse]+ *r*35[rhsf3(3): rhse])

*d*[Rhsp: Rhsf]*/dt* = [rhsp](*r*+ [rhsf]+ *r*19+[rhsf(1)]) *−* (*r−* [rhsp: rhsf]

18 18

+*r*19*−*[rhsp: rhsf(1)]) + [rhsp](*r*20[rhsf2]+ (*r*21 + *r*22)[rhsf2(1)]

+*r*23[rhsf2(2)]) + [rhsp](*r*24[rhsf3]+ (*r*25 + *r*26)[rhsf3(1)]

+(*r*27 + *r*28)[rhsf3(2)]+ *r*29[rhsf3(3)])

+[rhsp](*r*30[rhsf3: rhse]+ (*r*31 + *r*32)[rhsf3(1): rhse] (*r*33 + *r*34)[rhsf3(2): rhse]+ *r*35[rhsf3(3): rhse])

# The numerical setup of the refined model

Table D.1

The numerical values of the parameters of the refined model

*r*+ = *k*+; *r*+ = *k*+; *r*16 = *k*4; *r*28 = *k*7*/*2;

1 1 8 2

*r−* = *k−*; *r−* = *k−/*2; *r*17 = *k*4; *r*29 = *k*7;

1 1 8 2

*r*+ =2 *· k*+; *r*+ = *k*+; *r*18+ = *k*+; *r*30 = *k*8;

2 1 9 2 5

*r−* = *k−*; *r−* = *k−*; *r*18*−* = *k−*; *r*31 = *k*8*/*2;

2 1 9 2 5

*r*+ = *k*+; *r*10+ = *k*+; *r*19+ = *k*+; *r*32 = *k*8*/*2;

3 1 3 5

*r−* = *k−*; *r*10*−* = *k−*; *r*19*−* = *k−*; *r*33 = *k*8*/*2;

3 1 3 5

*r*+ = *k*+; *r*11+ = *k*+; *r*20 = *k*6; *r*34 = *k*8*/*2;

4 2 3

*r−* = *k−*; *r*11*−* = *k−*; *r*21 = *k*6*/*2; *r*35 = *k*8;

4 2 3

*r*+ = *k*+; *r*12+ = *k*+; *r*22 = *k*6*/*2; *r*36 = *k*9;

5 2 3

*r−* = *k−/*2; *r*12*−* = *k−*; *r*23 = *k*6; *r*37 = Φ*T* ;

5 2 3

*r*+ = *k*+; *r*13+ = *k*+; *r*24 = *k*7; *r*+ = *k*11+;

6 2 3 38

*r−* = *k−/*2; *r*13*−* = *k−*; *r*25 = *k*7*/*2; *r−* = *k*11*−*;

6 2 3 38

*r*+ = *k*+; *r*14 = *k*4; *r*26 = *k*7*/*2; *r*39 = *k*12

7 2

*r−* = *k−/*2; *r*15 = *k*4; *r*27 = *k*7*/*2;

7 2