Supplementary information

Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks:

Model Parameters

Daniel E. Zak[α,β]

Gregory E. Gonye[β]

James S. Schwaber[β,⋆]

Francis J. Doyle III[γ]

[α] Department of Chemical Engineering, University of Delaware, Newark, DE 19716

[β] Department of Pathology, Cell Biology and Anatomy, Thomas Jefferson University Philadelphia, PA 19107

[γ] Department of Chemical Engineering, University of California, Santa Barbara, CA 93106

⋆ *Please address questions or comments about this material to: James Schwaber,* james.schwaber@mail.tju.edu*, (215) 503–7823, FAX: (215) 923–3808*

1 Introduction

For all genes in the network, rate parameters must be specified for the processes of transcription, transcript degradation, translation, and protein degradation. For the genes that encode transcription factors, parameters for dimerization, undimerization, promoter binding, and promoter unbinding must also be specified. For the ligand binding mechanism, parameters must be specified for receptor binding, ligand concentration, and ligand degradation. In the paragraphs below, the derivation of each of these parameters that were used in the model is described, followed by a tabulation of the parameter values.

2 Concentration

Stochastic simulations were used in the present study to describe reactions at an intracellular level. For this reason, the appropriate concentration unit for the rate parameters is *molecules/cell*. To

convert parameters that were reported in molar units, it was necessary to estimate the volume of a typical mammalian cell. This was taken as 125 $micrometers^3$ [1]. Some values were reported in terms of *gram–tissue* [4]. To convert this to units of *molecules/cell*, it was assumed that there are approximately 10^8 cells per gram of tissue [2].

3 Transcription, transcript degradation, translation, and protein degradation rate parameters for metabolic enzymes

Kenney and Lee [4] (1982) provide a comprehensive study of absolute kinetic parameters in mammalian gene expression. They investigated the metabolic enzymes tyrosine aminotransferase (TAT) and alanine aminotransferase (AAT) in rat liver tissue. Despite their similar functions, the rate parameters of these two genes vary greatly. For transcript levels, they report 0.025 and 0.003 *picomole/gram–tissue* for TAT and AAT respectively. These correspond to 151 and 18 *transcripts/cell*, respectively. For protein levels, they report 80 and 350 *picomole/gram–tissue* for TAT and AAT respectively, which correspond to 4.8×10^5 and 2.1×10^6 *proteins/cell*. For first–order transcript degradation rate parameters, they report 9.7×10^{-3} and 4.8×10^{-4} $1/minutes$ for TAT and AAT, respectively. For the proteins, the first–order degradation rate parameters were 7.7×10^{-3} and 1.6×10^{-4} $1/minutes$ for TAT and AAT, respectively. For transcript synthesis, they report absolute rates of 0.348 and 0.002 *picomole/gram–tissue/day* for TAT and AAT, respectively. Assuming that there are two copies of each promoter for each gene per cell ($P_T = 2$), these correspond to transcription rates of 0.725 and 4.2×10^{-3} *transcripts/promoter/minute* for TAT and AAT respectively. For protein synthesis, they report absolute rates of 888 and 81 *picomole/gram–tissue/day* for TAT and AAT respectively. Using the mRNA levels per cell calculated above, these can be expressed as 24.5 and 18.9 *proteins/transcript/minute* for TAT and AAT, respectively. In the genetic regulatory network model, rate parameters for AAT were assigned to gene H and rate parameters for TAT were assigned to gene J.

4 Other transcription, transcript degradation, translation, and protein degradation rate parameters

Kenney and Lee (1982) [4] provided rate parameters for metabolic enzymes, but this is not sufficient for all of the genes in the model network, which include transcription factors and the receptor.

Ouali *et al.* (1997) provide half–lives for angiotensin II type–1 receptors (AT_1) and their transcripts in bovine adrenocortical (BAC) cells [7]. From these, first–order degradation rate parameters of 1.4×10^{-3} and 7.8×10^{-4} $1/minute$ for the AT_1 transcript and protein, respectively, were obtained. They also provide a typical receptor count of $10^5/cell$. These rate parameters were assigned to gene E.

For the transcription factors, values for the inducible transcription factor $c-jun$ were used. Herdegen and Leah (1998) provide half lives of 20–25 and 90–120 minutes for $c-jun$ mRNA and protein, respectively [3]. These can be converted to first–order degradation rate parameters of 3.1×10^{-2} and 6.6×10^{-3} $1/minute$, respectively. For lack of specific information about the degradation rates of transcription factors dimers, they were assumed to be equal to the degradation rates of the monomers. These rate parameters were assigned to genes A, C, D, F, G, and K.

Herdegen and Leah (1998) also provide turnover information for the transcription factor $Fra-2$ that binds to $c-jun$ protein to inhibit its transcription activating function. They state that upon induction, $Fra-2$ transcripts reach peak levels after 60 minutes and decline to supra–basal levels in 8 hours. This allows an approximate first–order degradation rate constant of 1.1×10^{-2} $1/minute$ to be calculated for the transcript. For $Fra-2$ protein, the half–life of 4 hours gives a degradation rate constant of 2.9×10^{-3} $1/minute$. These rate parameters were assigned to gene B.

Data for the remaining parameters for the transcription factors (translation and transcription rates) were unavailable. For this reason, the parameter values were approximated from the values for TAT and AAT in Kenney and Lee (1982) [4] to give reasonable final concentrations.

5 Transcription factor dimerization and promoter binding

The parameters that describe the dimerization and promoter binding kinetics of transcription factors are not widely available. Kohler and Schepartz (2001) performed a comprehensive study of

transcription factor ($c-fos$ and $c-jun$) dimerization and DNA–binding kinetics *in vitro* using stopped–flow fluorescence resonance energy transfer (FRET) [6]. They observed rate parameters of 7.36×10^{6} $1/mole/liter/second$ and 6.4×10^{-2} $1/second$ for the forward and reverse dimerization reactions in solution, respectively. These correspond to rate parameters of 5.9×10^{-3} $cell/molecule/minute$ and 3.8 $1/minute$, respectively. These parameters were used for all transcription factors in the model network.

Kohler and Schepartz (2001) found that their data was best explained by a model in which transcription factor dimer ($c-fos/c-jun$) complexes with DNA were assembled by transcription factor monomers first forming an unstable complex with DNA, which then recruited the other transcription factor to form the stable trimer. This differs slightly from the transcriptional regulatory module of the present work, where transcription factors dimerize first and then bind DNA. This difference is not likely to have a large impact on the overall behavior of the system because there only need to be about 150 molecules of one transcription factor to be present in solution before the rate of formation of transcription factor/transcription factor/DNA trimer exceeds the rate of dissociation of the other transcription factor from the DNA. This is a condition likely to be met *in vivo*, where upon induction there may be 10^3 – 10^5 molecules of transcription factor protein per cell (ex: $c--fos$, [3]). Following this argument, the forward rate parameter for transcription factor binding to promoter was taken as that for binding of $c-fos$ monomer to DNA (7.36×10^{6} $1/mole/liter/second$ = 7.5×10^{-4} $cell/molecule/minute$). The rate parameter for the unbinding of the transcription factor from the promoter was taken as that for dissociation of the $c-fos/c-jun$/DNA complex (0.39 $1/minute$). These rate parameters were used for all transcription factors in the model network. It must also be noted that there are known examples where transcription factors dimerize before binding DNA [9].

6 Ligand and ligand binding

For the binding of the ligand to the receptor, the kinetics for the epidermal growth factor receptor (EGFR) were used. Kholodenko *et al.* (1999) give a rate constant of 3.0×10^{-3} $1/nanomole/liter/second$ for the second–order ligand binding reaction, which corresponds to 2.4×10^{-3} $cell/molecule/minute$. They give a first–order unbinding reaction rate constant of 3.6 $1/minute$.

Studies that investigate the downstream consequences of ligand binding often use nanomolar ligand concentrations (Kholodenko *et al.* (1999) used 0.2, 2, and 20 *nanomolar* [5]; Reagan *et al.* (1993) used 10 *nanomolar* [8], and Ouali *et al.* (1997) used 10–100 *nanomolar* [7]). For this reason, 10 nanomolar ligand concentrations were considered in the present study. It is difficult to estimate the amount of ligand that is actually presented to the surface of the cell given its concentration in solution. To overcome this, it was assumed that this amount corresponds to the amount displaced by the cell volume (125 $micrometers^3$). For 10 nanomolar ligand concentration, this gives 752.5 $molecules/cell$.

In addition to removal of ligand via receptor binding, it was assumed that ligand was removed by a mechanism with first–order kinetics, as would be the case in a flow chamber or an unstable ligand. In the case of the flow chamber, the rate parameter for removal is the reciprocal of the residence time of the chamber. A residence time of 1 minute was assumed for the chamber in the present study, giving a rate parameter for ligand removal of $k_{DQ} = 1\ 1/minute$. Since it was desired that the steady–state concentration of ligand be 10 nanomolar (neglecting consumption of ligand by receptor binding), this required an input of 752.5 $molecules/cell/minute$. This quantity is represented by the constant C_Q in the model ($C_Q = 752.5\ molecules/cell/minute$). Thus in the single step study, the ligand input was held constant at 752.5 $molecules/cell/minute$, and for the one hour pulses the ligand input was held at 752.5 $molecules/cell/minute$ for one hour.

References

[1] A. P. Davenport and D. J. Nunez. Quantification of radioactive mRNA *in situ* hybridization signals. In *In Situ Hybridization, Principles and Practice (JM Polak and J O'D McGee, eds)*, pages 849–854, 1990.

[2] J. L. Hargrove. *Kinetic Modeling of Gene Expression.* R.G. Landes Co., Austin, Texas, 1994.

[3] T. Herdegen and J. D. Leah. Inducible and constitutive transcription factors in the mammalian nervous system: Control of gene expression by Jun, Fos, and Krox, and CREB/ATF proteins. *Brain. Res. Rev.*, 28:370–490, 1998.

[4] F. T. Kenney and K. L. Lee. Turnover of gene products in the control of gene expression. *Bioscience*, 32(3):181–184, March 1982.

[5] B. N. Kholodenko, O. V. Demin, G. Moehren, and J. B. Hoek. Quantification of short term signaling by the epidermal growth factor receptor. *J. Biol. Chem.*, 274(42):30169–30181, 1999.

[6] J. J. Kohler and A. Schepartz. Kinetic studies of Fos.Jun.DNA complex formation: DNA binding prior to dimerization. *Biochemistry–US*, 40(1):130–142, 2001.

[7] R. Ouali, M. C. Berthelon, M. Begeot, and J. M. Saez. Angiotensin II receptor subtypes AT1 and AT2 are down–regulated by Angiotensin II through AT1 receptor by different mechanisms. *Endocrinology*, 138(2):725–733, 1997.

Table 1: **Degradation, translation, dimerization, and promoter binding parameters in the genetic regulatory network simulator**: (**p** = proteins, **t** = transcripts, **m** = molecules, **c** = cell)

Parameter	**Value**	**Units**	**Reference**
Protein degradation		$1/minute$	
k_{DA},k_{DC},k_{DegD}, k_{DF},k_{DG},k_{DK}	6.6×10^{-3}		[3]
k_{DB}	2.9×10^{-3}		[3]
k_{DE}	7.8×10^{-4}		[7]
k_{DH}	1.6×10^{-4}		[4]
k_{DJ}	7.7×10^{-3}		[4]
Transcript degradation		$1/minute$	
k_{DMA},k_{DMC},k_{DMD}, k_{DMF},k_{DMG},k_{DMK}	3.1×10^{-2}		[3]
k_{DMB}	1.1×10^{-2}		[3]
k_{DME}	1.4×10^{-3}		[7]
k_{DMH}	4.8×10^{-4}		[4]
k_{DMJ}	9.7×10^{-3}		[4]
Translation		$p/t/minute$	
k_{TA},k_{TC},k_{TD}, k_{TF},k_{TG},k_{TK}	2.5		*calculated*
k_{TB}	1.2		*calculated*
k_{TE}	18.9		[4]
k_{TH}	18.9		[4]
k_{TJ}	24.5		[4]
Dimerization		$c/m/minute$	
k_{C_2},k_{D_2},k_{F_2}, k_{G_2},k_{K_2},k_{AB}	5.9×10^{-3}		[6]
k_{EQ}	2.4×10^{-3}		[5]
Undimerization		$1/minute$	
k_{UC_2},k_{UD_2},k_{UF_2}, k_{UG_2},k_{UK_2},k_{UAB}	3.8		[6]
k_{UEQ}	3.6		[5]
Dimer degradation		$1/minute$	
k_{DC_2},k_{DD_2},k_{DF_2}, k_{DG_2},k_{DK_2},k_{DAB}, k_{DEQ}	6.6×10^{-3}		*estimated*
Promoter binding		$c/m/minute$	
$k_{P_{AA}}$,$k_{P_{BAF}}$,$k_{P_{CAD}}$, $k_{P_{DCF}}$,$k_{P_{GC}}$,$k_{P_{KC}}$, $k_{P_{ED}}$,$k_{P_{HG}}$,$k_{P_{JK}}$, $k_{P_{FEQ}}$	7.5×10^{-4}		[6]
Promoter unbinding		$1/minute$	
$k_{UP_{AA}}$,$k_{UP_{BAF}}$,$k_{UP_{CAD}}$, $k_{UP_{DCF}}$,$k_{UP_{GC}}$,$k_{UP_{KC}}$, $k_{UP_{ED}}$,$k_{UP_{HG}}$,$k_{UP_{JK}}$, $k_{UP_{FEQ}}$	0.39		[6]

Table 2: **Transcriptional parameters in the genetic regulatory network simulator:** [*transcripts/promoter/minute*].

Parameter	**Value**	**Reference**
Gene A		
$k_{RP_{AA}}$	7.3×10^{-4}	*estimated*
$k_{RAP_{AA}}$	7.3×10^{-1}	[4]
Gene B		
$k_{RP_{BAF}}$	3.6×10^{-4}	*estimated*
$k_{RAP_{BAF}}$	3.6×10^{-2}	*estimated*
$k_{RF_2P_{BAF}}$	3.6×10^{-1}	*estimated*
$k_{RAF_2P_{BAF}}$	7.2×10^{-1}	[4]
Gene C		
$k_{RP_{CAD}}$	7.3×10^{-4}	*estimated*
$k_{RAP_{CAD}}$	7.3×10^{-1}	[4]
$k_{RD_2P_{CAD}}$	0	*estimated*
$k_{RAD_2P_{CAD}}$	7.3×10^{-2}	*estimated*
Gene D		
$k_{RP_{DCF}}$	7.3×10^{-4}	*estimated*
$k_{RC_2P_{DCF}}$	0	*estimated*
$k_{RF_2P_{DCF}}$	7.3×10^{-1}	[4]
$k_{RC_2F_2P_{DCF}}$	7.3×10^{-1}	[4]
Gene E		
$k_{RP_{ED}}$	4.2×10^{-3}	[4]
$k_{RD_2P_{ED}}$	4.2×10^{-7}	*estimated*
Gene F		
$k_{RP_{FEQ}}$	7.3×10^{-4}	*estimated*
$k_{REQP_{FEQ}}$	7.3×10^{-1}	[4]
Gene G		
$k_{RP_{GC}}$	7.3×10^{-1}	[4]
$k_{RC_2P_{GC}}$	7.3×10^{-4}	*estimated*
Gene H		
$k_{RP_{HG}}$	4.2×10^{-3}	[4]
$k_{RG_2P_{HG}}$	4.2×10^{-6}	*estimated*
Gene J		
$k_{RP_{JK}}$	7.3×10^{-4}	*estimated*
$k_{RK_2P_{JK}}$	7.3×10^{-1}	[4]
Gene K		
$k_{RP_{KC}}$	7.3×10^{-1}	[4]
$k_{RC_2P_{KC}}$	7.3×10^{-4}	*estimated*

[8] L. P. Reagan, X. Ye, C. H. Maretzki, and S. J. Fluharty. Down–regulation of Angiotensin II receptor subtypes and desensitization of cyclic GMP production in neuroblastoma N1E-115 cells. *J. Neurochem.*, 60(1):24–31, 1993.

[9] D. S. Spinner, S. Liu, S. W. Wang, and J. Schmidt. Interaction of the myogenic determination factor myogenin with E12 and a DNA target: Mechanism and kinetics. *J. Mol. Biol.*, 317(3):431–445, 2002.