SIMPLIFYING MODELS WITHOUT LOSING BIOCHEMICAL COMPLEXITY

Natalie Arkus, Michael Brenner
Division of Engineering and Applied Sciences, Cambridge, Massachusetts

Models in Biology

There is an abundance of biological models containing many equations and free parameters.

Two examples:

Heat Shock Response (HSR) in *E. coli:*
(El Samad et al., PNAS, 102, 2736-2741)

$$\frac{d[mRNA(DnaK)]}{dt} = K_{tr1}\frac{[\sigma^{32}:RNAP:ph]}{ph_t} - \alpha_{mRNA}[mRNA(DnaK)]$$
$$\frac{d[DnaK_t]}{dt} = K_{TL}[mRNA(DnaK)] - \alpha_{prot}[DnaK_t]$$
$$\frac{d[mRNA(FtsH)]}{dt} = K_{tr2}\frac{[\sigma^{32}:RNAP:ph]}{ph_t} - \alpha_{mRNA}[mRNA(FtsH)]$$
$$\frac{d[FtsH_t]}{dt} = K_{TL}[mRNA(FtsH)] - \alpha_{prot}[FtsH_t]$$
$$\frac{d[mRNA(protease)]}{dt} = K_{tr3}\frac{[\sigma^{32}:RNAP:ph]}{ph_t} - \alpha_{mRNA}[mRNA(protease)]$$
$$\frac{d[protease_t]}{dt} = K_{TL}[mRNA(protease)] - \alpha_{prot}[protease_t]$$
$$\frac{d[mRNA(HslVU)]}{dt} = K_{tr4}\frac{[\sigma^{32}:RNAP:ph]}{ph_t} - alpha_{mRNA}[mRNA(HslVU)]$$
$$\frac{d[HslVU_t]}{dt} = K_{TL}[mRNA(HslVU)] - \alpha_{prot}[HslVU_t]$$
$$\frac{d[mRNA(\sigma^{32})]}{dt} = K_{tr5}\frac{[\sigma^{70}:RNAP:pg]}{pg_t} - \alpha_{mRNA}[mRNA(\sigma^{32})]$$
$$\frac{d[\sigma^{32}]}{dt} = K_{TL}\eta(T)[mRNA(\sigma^{32})] - \alpha_{prot}[\sigma^{32}_t] - \alpha_{FtsH}[\sigma^{32}:DnaK:FtsH] - \alpha_{protease}(T)[\sigma^{32}:DnaK:protease] - \alpha_{HslVU(T)}[\sigma^{32}:HslVU]$$
$$\frac{d[P_{folded}]}{dt} = K_{fold}[P_{unfolded}:DnaK] - K(T)[P_{folded}]$$

$$[\sigma^{70}:RNAP] = K_1[\sigma^{70}_f][RNAP_f]$$
$$[\sigma^{32}:RNAP] = K_2[\sigma^{32}_f][RNAP_f]$$
$$[RNAP:D] = K_3[RNAP_f][D_t]$$
$$[\sigma^{32}:DnaK:FtsH] = K_4[\sigma^{32}:DnaK][FtsH_f]$$
$$[\sigma^{32}:DnaK] = K_5[\sigma^{32}_f][DnaK_f]$$
$$[\sigma^{32}:DnaK:protease] = K_6[\sigma^{32}:DnaK][protease_f]$$
$$[P_{unfolded}:DnaK] = K_8[P_{unfolded}][DnaK_f]$$
$$[\sigma^{32}:RNAP:ph] = K_9[\sigma^{32}:RNAP]([ph_t] - [\sigma^{32}:RNAP:ph])$$
$$[\sigma^{70}:RNAP:pg] = K_{10}[\sigma^{70}:RNAP]([pg_t] - [\sigma^{70}:RNAP:pg])$$
$$[\sigma^{70}:RNAP:D] = K_{11}[\sigma^{70}:RNAP][D_t]$$
$$[\sigma^{32}:RNAP:D] = K_{12}[\sigma^{32}:RNAP][D_t]$$
$$[RNAP_t] = [RNAP_f] + [\sigma^{70}:RNAP] + [\sigma^{32}:RNAP] + [RNAP:D] + [\sigma^{70}:RNAP:D] + [\sigma^{32}:RNAP:D] + [\sigma^{70}:RNAP:pg] + [\sigma^{32}:RNAP:ph]$$
$$[\sigma^{70}_t] = [\sigma^{70}_f] + [\sigma^{70}:RNAP] + [\sigma^{70}:RNAP:D] + [\sigma^{70}:RNAP:pg]$$
$$[\sigma^{32}_t] = [\sigma^{32}_f] + [\sigma^{32}:DnaK:protease] + [\sigma^{32}:RNAP] + [\sigma^{32}:RNAP:D] + [\sigma^{32}:DnaK:FtsH] + [\sigma^{32}:DnaK] + [\sigma^{32}:RNAP:ph] + [\sigma^{32}:HslVU]$$
$$[DnaK_t] = [DnaK_f] + [\sigma^{32}:DnaK:FtsH] + [\sigma^{32}:DnaK] + [P_{unfolded}:DnaK] + [\sigma^{32}:DnaK:protease]$$
$$[FtsH_t] = [FtsH_f] + [\sigma^{32}:DnaK:FtsH]$$
$$[HslVU_t] = [HslVU_f] + [\sigma^{32}:HslVU]$$
$$[protease_t] = [protease_f] + [\sigma^{32}:DnaK:protease]$$
$$[Protein_t] = [P_{unfolded}] + [P_{unfolded}:DnaK] + [P_{folded}]$$

β-catenin Degradation in the Wnt Signaling Pathway:
(Lee et al., PloS Biology, 1, 116-132)

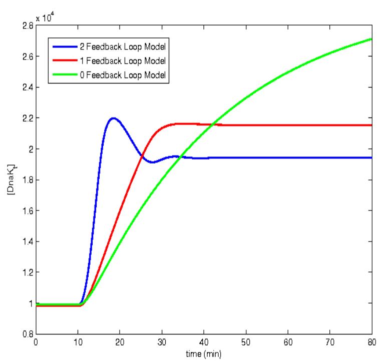
$$\frac{dX_1}{dt} = -v_1 + v_2$$
$$\frac{dX_2}{dt} = v_1 - v_2$$
$$\frac{dX_3}{dt} = v_4 - v_5 - v_8 + v_{10}$$
$$\frac{dX_4}{dt} = -v_3 - v_4 + v_5 + v_6$$
$$\frac{dX_5}{dt} = v_3 - v_6$$
$$\frac{dX_6}{dt} = v_3 - v_6 + v_7$$
$$\frac{dX_7}{dt} = -v_7 - v_{17}$$
$$\frac{dX_8}{dt} = v_8 - v_9$$
$$\frac{dX_9}{dt} = v_9 - v_{10}$$
$$\frac{dX_{10}}{dt} = v_{10} - v_{11}$$
$$\frac{dX_{11}}{dt} = -v_8 + v_{12} - v_{13} - v_{16} - v_{17}$$
$$\frac{dX_{12}}{dt} = -v_7 + v_{14} - v_{15}$$
$$\frac{dX_{13}}{dt} = -v_{16}$$
$$\frac{dX_{14}}{dt} = v_{16}$$
$$\frac{dX_{15}}{dt} = v_{17}$$
$$X_1 + X_2 = const$$
$$X_3 + X_4 + X_5 + X_8 + X_9 = const$$
$$X_3 + X_4 + X_6 + X_7 + X_8 + X_9 + X_{15} = const$$
$$X_{13} + X_{14} = const$$

where, for example

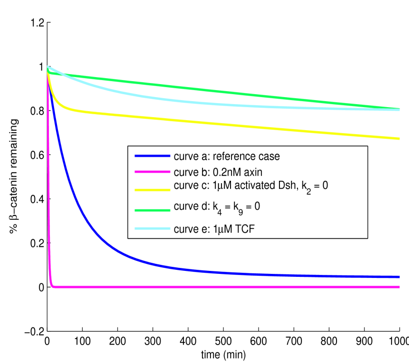
$$v_i = k_{+i}X_jY_l - k_{-i}(X_j.Y_l)$$
$$\frac{dX_{11}}{dt} = \frac{v_{12} - \left(\frac{k_9X_3}{K_8} + K_{13}\right)X_{11} - \frac{X_{11}}{K_8+X_{11}}\left(k_4X_4 - k_5X_3 - \frac{k_9X_3X_{11}}{K_8} + k_{10}X_9\right)}{1 + \frac{X_3}{K_8} + \frac{TCF^0K_{16}}{(K_{16}+X_{11})^2} + \frac{APC^0K_{17}}{(K_{17}+X_{11})^2} - \frac{X_{11}X_3}{K_8(K_8+X_{11})}}$$

Yet, the biological data to which the models are fit are relatively simple:

HSR:



Wnt Signaling:



Can these systems be simplified while maintaining all relevant biological information?

Application of Dominant Balance

Example of Dominant Balance:

$$\begin{aligned} 10000 &= 6000 + 3900 + 10 + 20 + 60 + .01 + .0004 + 5 + 3 + 1.9896 \\ &\approx 6000 + 3900 \end{aligned}$$

Applied to Biological Models:

Step 1: Look for a separation of time or concentration flux scales → this will allow differential equations to be approximated as algebraic.

Example:

$$\frac{dx_1}{dt} = x_1x_2 - 0.003x_1$$
$$\frac{dx_2}{dt} = x_1^2 - 0.3x_2$$

x_2 degrades and thus equilibrates relatively quickly

$$\Longrightarrow 0 \approx x_1^2 - 0.3x_2$$

The system can therefore be simplified from 2 ODEs to 1:

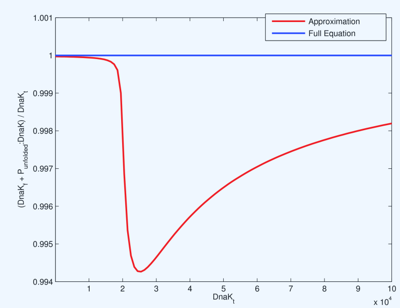
$$\frac{dx_1}{dt} = \frac{x_1^3}{0.3} - 0.003x_1$$

Step 2: Approximate algebraic equations by keeping only the most dominant terms.

Example:

$$\begin{aligned} [DnaK_f] &= [DnaK_t] - [P_{unfolded}:DnaK] - [\sigma^{32}:DnaK] \\ &\approx [DnaK_t] - [P_{unfolded}:DnaK] \end{aligned}$$

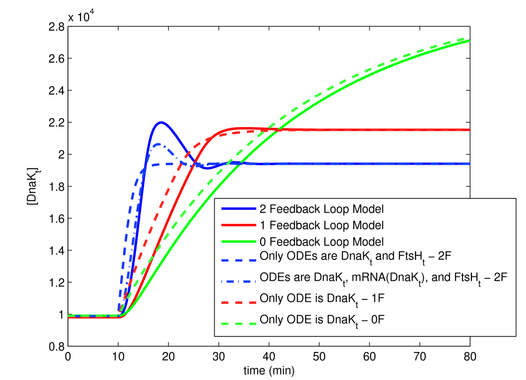
This approximation yields 99.4% accuracy!



Simplified Biological Models that Retain all Biological Information

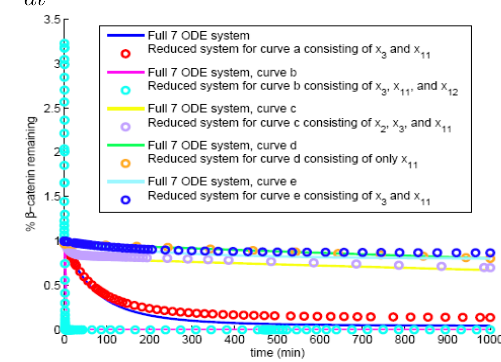
The reduced HSR system:

$$\frac{d[DnaK_t]}{dt} = K_{tr1}\frac{A}{B + C[DnaK_t] + D[FtsH_t]} - \alpha_{prot}[DnaK_t]$$
$$\frac{d[FtsH_t]}{dt} = K_{tr2}\frac{A}{B + C[DnaK_t] + D[FtsH_t]} - \alpha_{prot}[FtsH_t]$$



The reduced model of the Wnt signaling pathway:

$$\frac{dx_3}{dt} = Ax_{12} - k_5x_3$$
$$\frac{dx_{11}}{dt} = \frac{-x_{11}(k_9x_3 + k_{13})(K_{16} + x_{11})^2}{K_8(1 + TCF^0K_{16})}$$
$$\frac{dx_{12}}{dt} = -Bx_{12}$$



For both systems, A, B, C, and D are simply constants that are functions of the original parameters and variables of the system.

For example, for the Wnt signaling pathway

$$A = \frac{k_4k_6GSK^0x_{12}APC^0}{K_7(k_3x_2 + k_4 + k_{-6})}$$