Modelling Complex Systems in the Life Sciences: Mathematical Elegance v. Biological Relevance

Jaroslav Stark
Centre for Nonlinear Dynamics and its Applications
UCL

Special Thanks to
K. Hardy (Imperial College School of Medicine)
R. Callard (Institute of Child Health)



Overview

- **Why?**
  - Exciting science
  - Huge variety of untouched problems
  - Demand from biologists, students
  - Funding pressures
- **How?**
  - Driven by explicit biological goals
  - Unforeseen mathematical payoffs
  - Mathematics and biology as partners
  - How do we deal with complexity?
  - How do we compare model to data?
- **What?**
  - Focus on cell signalling
  - Broad philosophical presentation

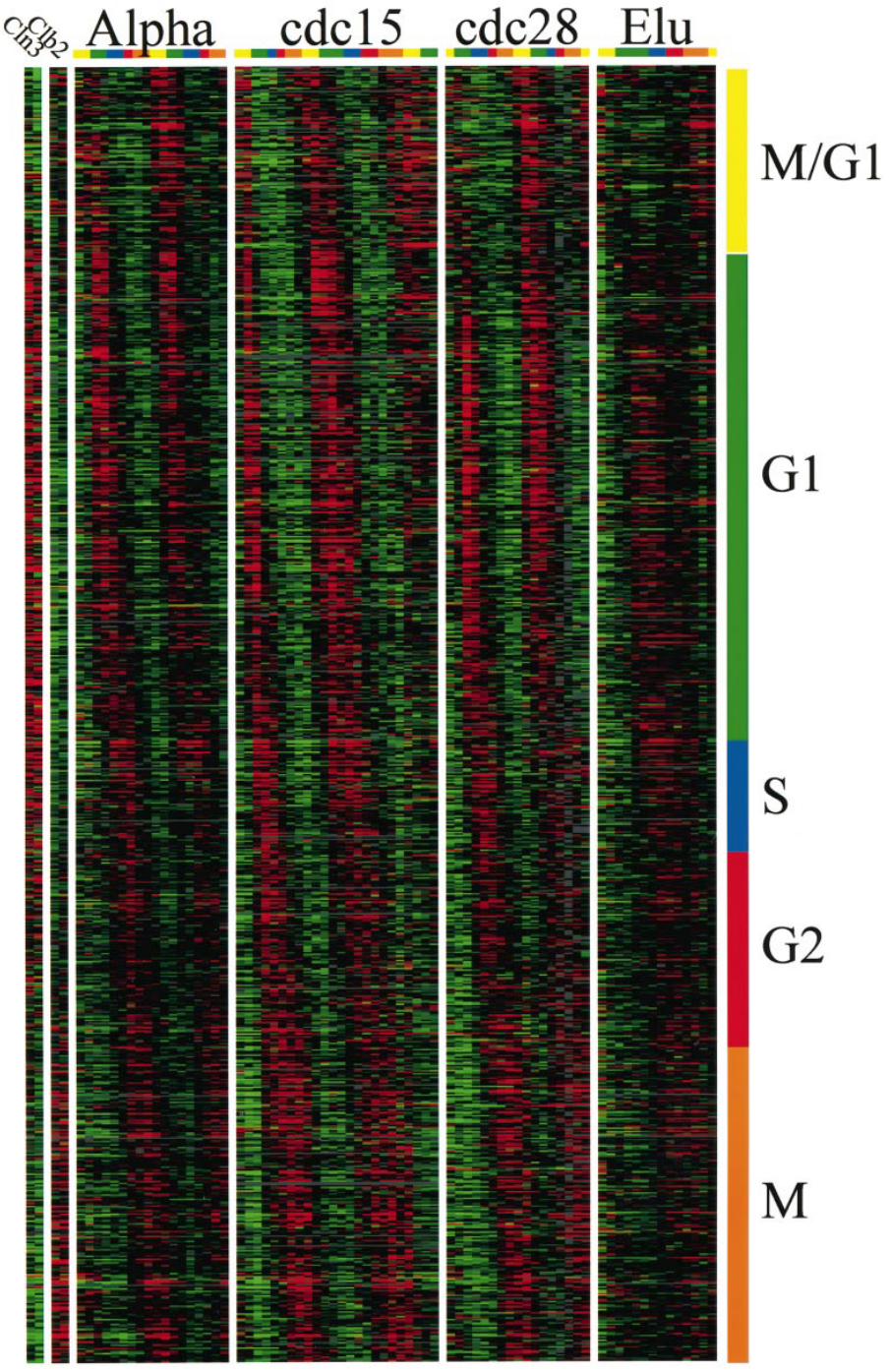


Why?

- **Life Sciences**
  - Possesses **vast** amounts of information about individual interactions
  - Needs to **integrate** this into an understanding of whole system
  - Many potential experiments or treatments are difficult or impossible
- **Mathematics**
  - Framework for constructing models
  - Techniques for their analysis
  - Methods for their simplification
- **Together**
  - New insights into behaviour of system
  - Possibility of "virtual laboratory" to supplement *in vivo* and *in vitro* work

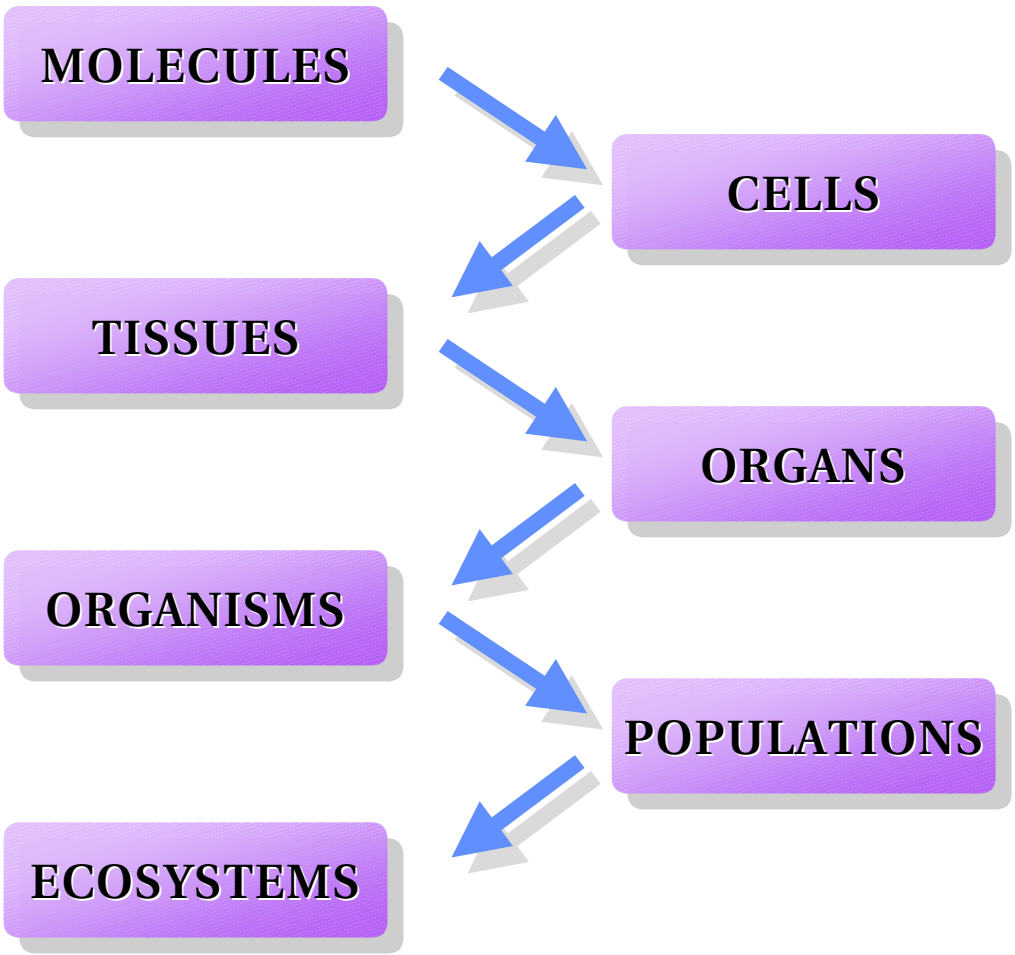


Gene Expression Data



Components

- Every biological system consists of many components



Interactions

- As important as components

COMPONENTS

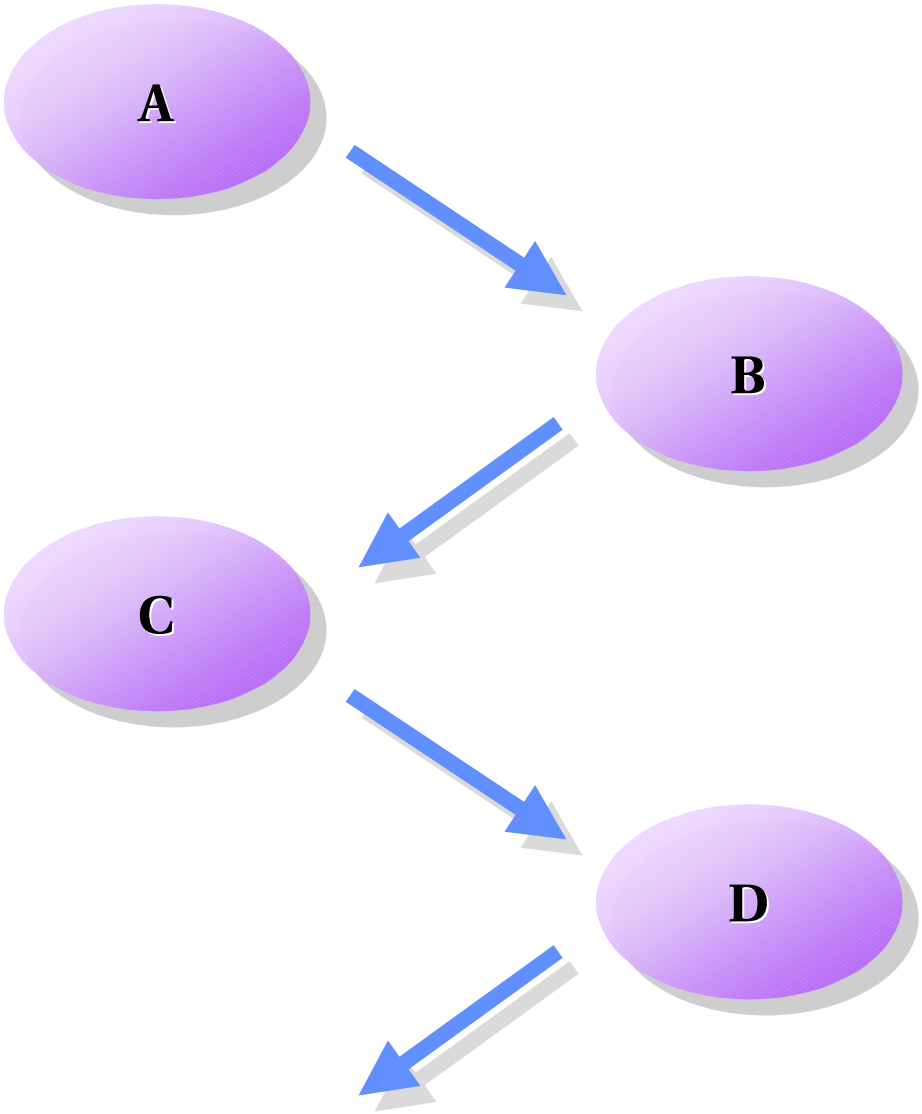
STRUCTURE

INTERACTIONS

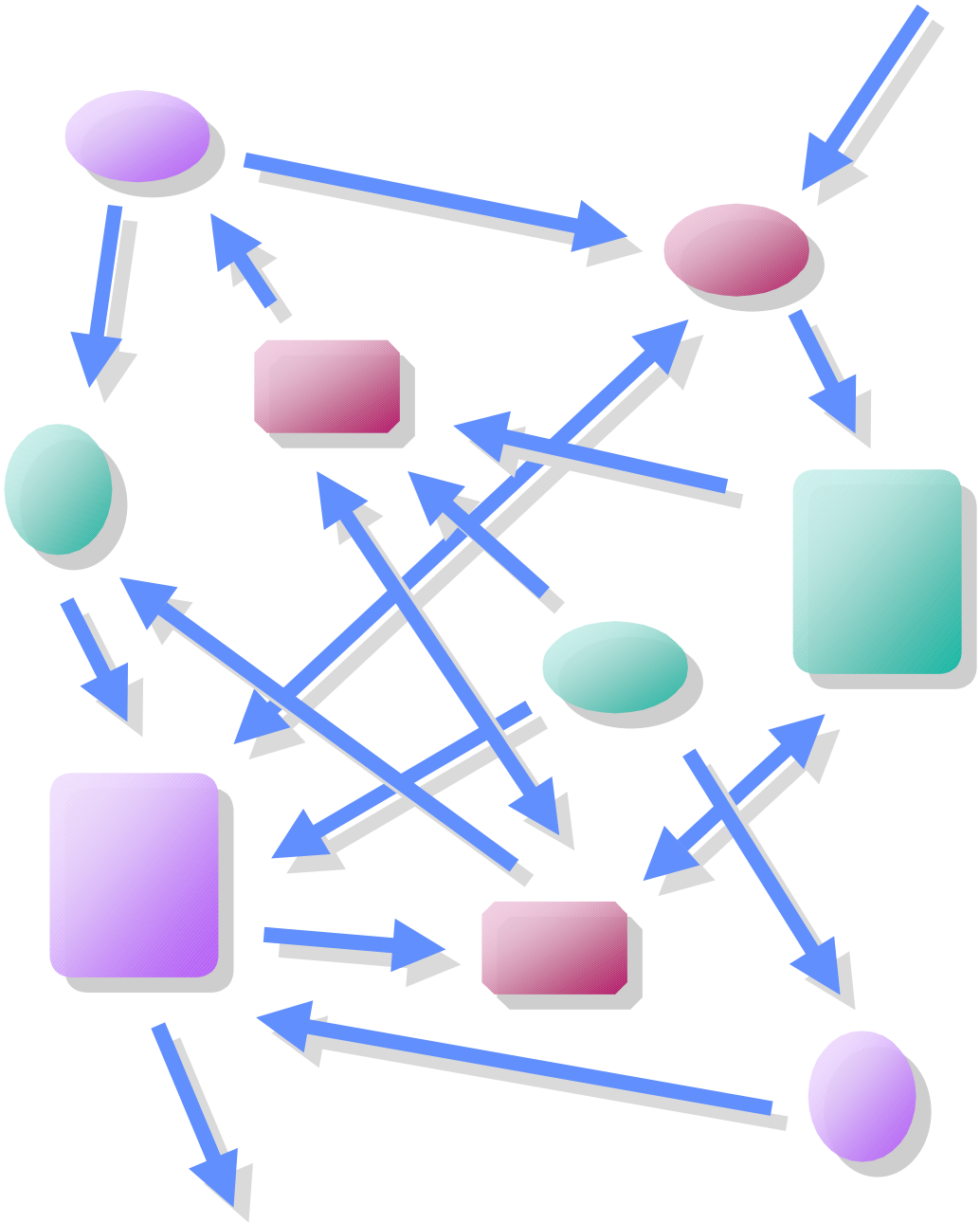


Simple Interactions

- Classical view



More Realistic Picture



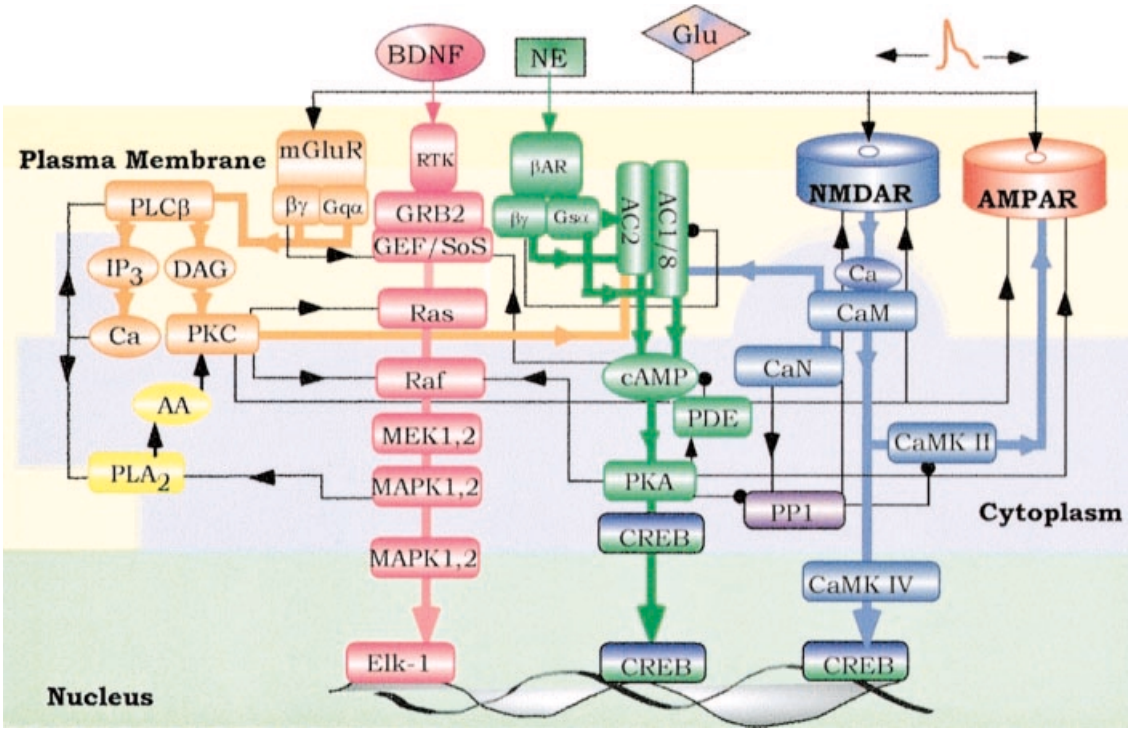
Real Example

- Traditional view

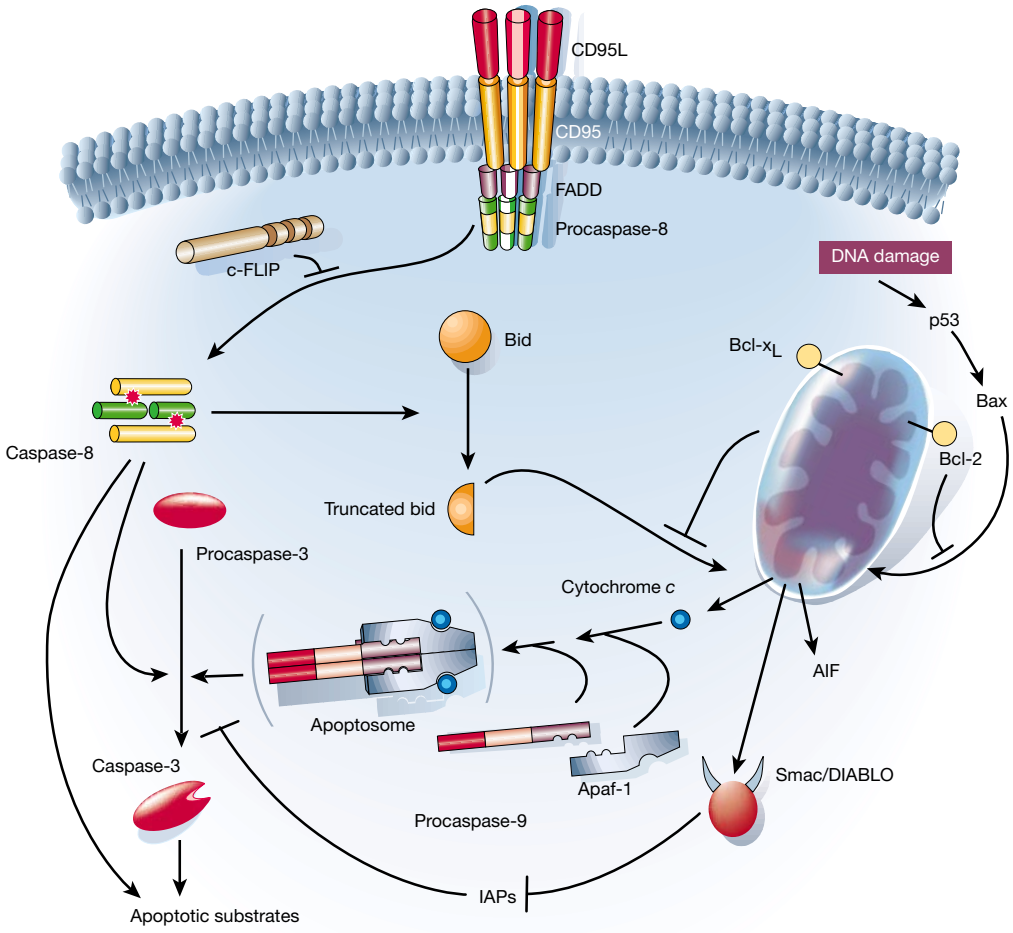


Real Example

- Modern view



Real Example II

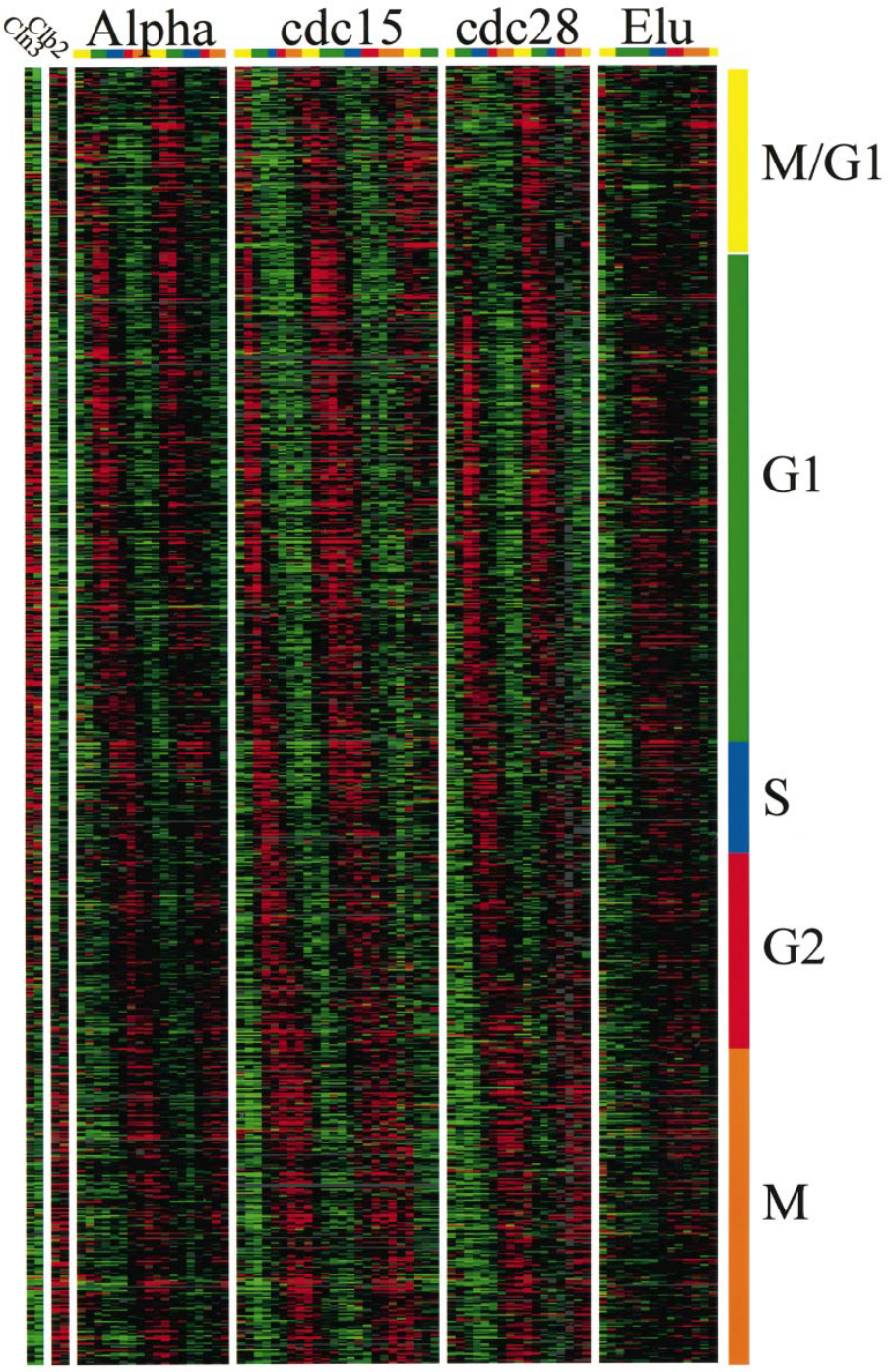


How?

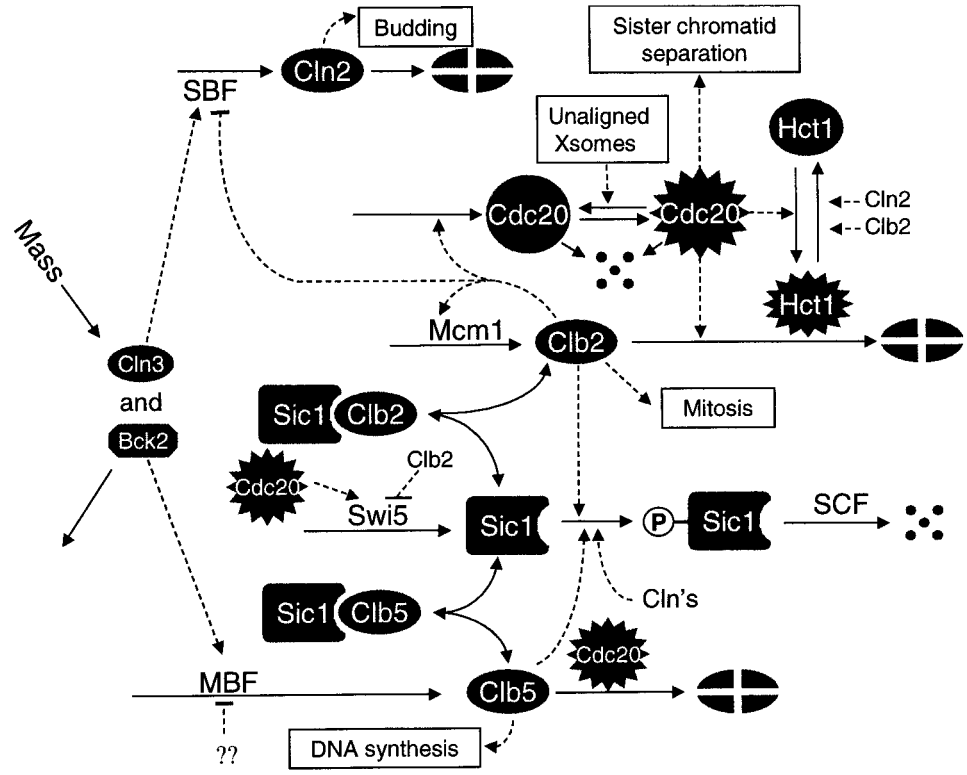
- **Pitfalls**
  - Mathematically sophisticated but biologically useless models
    - Lacking biological input
    - Lacking biological output
  - Biologically realistic but mathematically intractable models
    - Biologists want to include every effect
- **Explicit biological goals**
  - Biologically sound model
  - Experimentally verifiable model
  - Generate new biological hypotheses
    - It is not enough to simply reproduce the data



From Data to Model



Signaling Network



Mathematical Model

Equations governing cyclin-dependent kinases

$$\frac{d}{dt}[Cln2] = (k'_{s,n2} + k''_{s,n2}[SBF]) \cdot mass - k_{d,n2}[Cln2]$$

$$\frac{d}{dt}[Clb2]_T = (k'_{s,b2} + k''_{s,b2}[Mcm1]) \cdot mass - V_{d,b2}[Clb2]_T,\ V_{d,b2} = k'_{d,b2}([Hct1]_T - [Hct1]) + k''_{d,b2}[Hct1] + k'''_{d,b2}[Cdc20]$$

$$\frac{d}{dt}[Clb5]_T = (k'_{s,b5} + k''_{s,b5}[MBF]) \cdot mass - V_{d,b5} \cdot [Clb5]_T,\ V_{d,b5} = k'_{d,b5} + k''_{d,b5}[Cdc20]$$

$$[Bck2] = [Bck2]^0 \cdot mass, \qquad [Cln3]^* = [Cln3]_{max} \frac{D_{n3} \cdot mass}{J_{n3} + D_{n3} \cdot mass}$$

$$[Clb2]_T = [Clb2] + [Clb2/Sic1], \quad [Clb5]_T = [Clb5] + [Clb5/Sic1]$$

$$[Sic1]_T = [Sic1] + [Clb2/Sic1] + [Clb5/Sic1]$$

Equations governing the inhibitor of Clb-dependent kinases

$$\frac{d}{dt}[Sic1]_T = k'_{s,c1} + k''_{s,c1}[Swi5] - \left(k_{d1,c1} + \frac{V_{d2,c1}}{J_{d2,c1} + [Sic1]_T}\right) \cdot [Sic1]_T$$

$$\frac{d}{dt}[Clb2/Sic1] = k_{as,b2}[Clb2] \cdot [Sic1] - \left(k_{di,b2} + V_{d,b2} + k_{d1,c1} + \frac{V_{d2,c1}}{J_{d2,c1} + [Sic1]_T}\right) \cdot [Clb2/Sic1]$$

$$\frac{d}{dt}[Clb5/Sic1] = k_{as,b5}[Clb5] \cdot [Sic1] - \left(k_{di,b5} + V_{d,b5} + k_{d1,c1} + \frac{V_{d2,c1}}{J_{d2,c1} + [Sic1]_T}\right) \cdot [Clb5/Sic1]$$

$$V_{d2,c1} = k_{d2,c1}(\varepsilon_{c1,n3}[Cln3]^* + \varepsilon_{c1,k2}[Bck2] + [Cln2] + \varepsilon_{c1,b5}[Clb5] + \varepsilon_{c1,b2}[Clb2])$$

Equations governing the Clb degradation machinery

$$\frac{d}{dt}[Cdc20]_T = (k'_{s,20} + k''_{s,20}[Clb2]) - k_{d,20}[Cdc20]_T$$

$$\frac{d}{dt}[Cdc20] = k_{a,20}([Cdc20]_T - [Cdc20]) - (V_{i,20} + k_{d,20}) \cdot [Cdc20]$$

$$V_{i,20} = \begin{cases} k'_{i,20}, & \text{for END_M + 12 min} < t < \text{START_S} \\ k''_{i,20}, & \text{for START_S} < t < \text{END_M} \end{cases}$$

$$\frac{d}{dt}[Hct1] = \frac{(k'_{a,t1} + k''_{a,t1}[Cdc20]) \cdot ([Hct1]_T - [Hct1])}{J_{a,t1} + [Hct1]_T - [Hct1]} - \frac{V_{i,t1}[Hct1]}{J_{i,t1} + [Hct1]}$$

$$V_{i,t1} = k'_{i,t1} + k''_{i,t1}([Cln3]^* + \varepsilon_{i,t1,n2}[Cln2] + \varepsilon_{i,t1,b5}[Clb5] + \varepsilon_{i,t1,b2}[Clb2])$$

Equations for growth, DNA synthesis, budding and spindle formation

$$\frac{d}{dt}mass = \mu \cdot mass, \quad \frac{d}{dt}[ORI] = k_{s,ori}([Clb5] + \varepsilon_{ori,b2}[Clb2]) - k_{d,ori}[ORI]$$

$$\frac{d}{dt}[BUD] = k_{s,bud}([Cln2] + [Cln3]^* + \varepsilon_{bud,b5}[Clb5]) - k_{d,bud}[BUD], \quad \frac{d}{dt}[SPN] = k_{s,spn}\frac{[Clb2]}{J_{spn} + [Clb2]} - k_{d,spn}[SPN]$$

Equations governing transcription factors

$$[SBF] = [MBF] = G(V_{a,sbf}, k'_{i,sbf} + k''_{i,sbf}[Clb2], J_{a,sbf}, J_{i,sbf}),\ V_{a,sbf} = k_{a,sbf}([Cln2] + \varepsilon_{sbf,n3}([Cln3]^* + [Bck2]) + \varepsilon_{sbf,b5}[Clb5]),$$

$$[Mcm1] = G(k_{a,mcm}[Clb2], k_{i,mcm}, J_{a,mcm}, J_{i,mcm}),\ [Swi5] = G(k_{a,swi}[Cdc20], k'_{i,swi} + k''_{i,swi}[Clb2], J_{a,swi}, J_{i,swi})$$

Bifurcation Analysis

Fig. 4

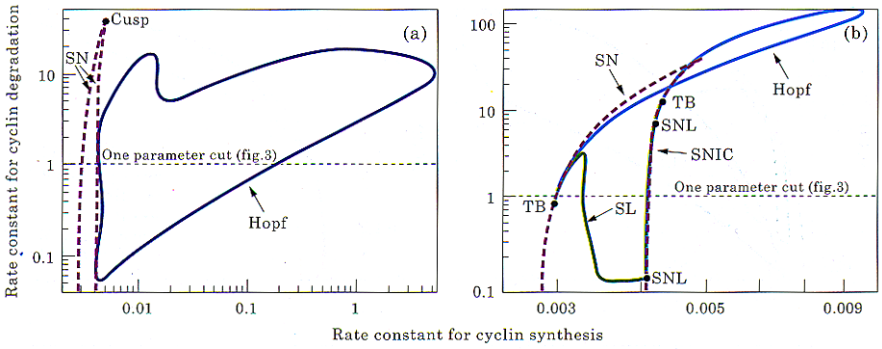


Fig. 5

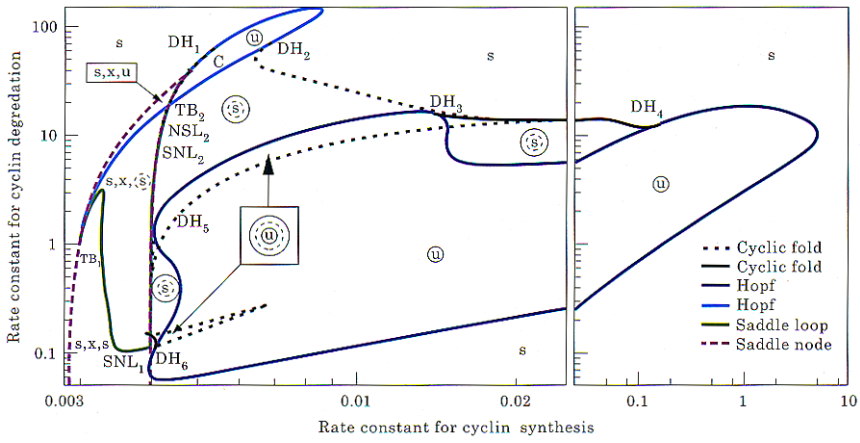
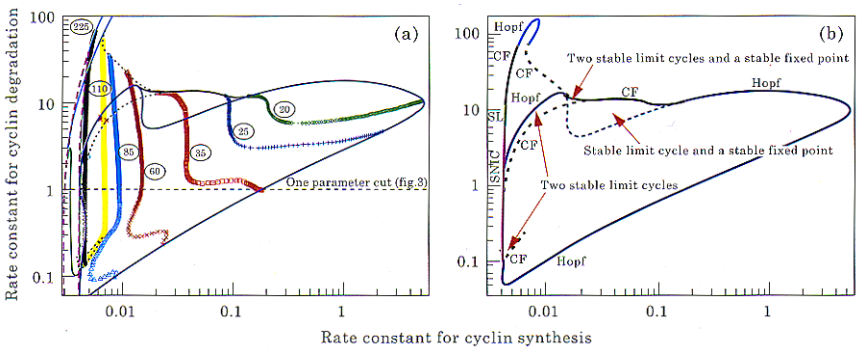


Fig. 7

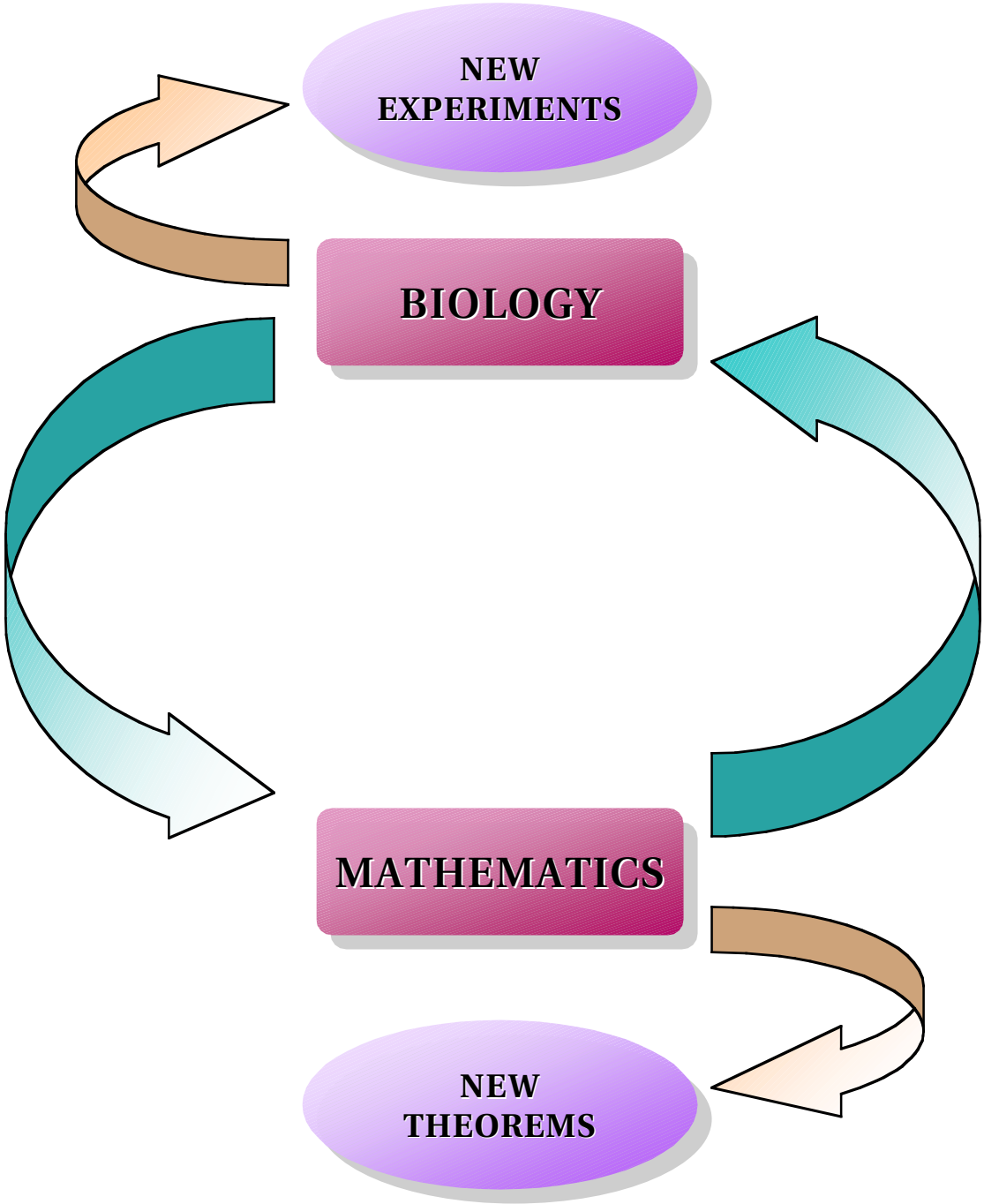


How? II

- **Prerequisites**
  - Must be done in collaboration
  - Clear communication
  - Takes time to develop
  - Iteration between biology and mathematics
  - Both sides must benefit equally
- **Mathematical Payoffs**
  - Always there
    - Hodgin-Huxley equations
    - Turing instability
  - Usually unpredictable



The Modelling Process



Modelling Issues

- **Nature of Biological Data**
  - Noisy, inaccurate
  - Often contradictory
  - Non-stationary
  - Huge crude data sets
    - Genome, DNA arrays
    - Neural activity
  - Crucial parameters may be difficult to obtain
- **Types of Models**
  - Maps, ODE's, PDE's, Delay DE's …
  - Cellular automata, logical networks
  - Stochastic versions of the above two



Mathematical Problems

- Complexity
- Equivalence of Systems
- Non-Autonomy



Complexity

- **Appropriate Level of Model**
  - Comprehensive realistic models
  - Simple idealized models
  - Why are simple models often useful?
- **Complex Models**
  - How to analyse complex models?
  - How to simplify complex models?
  - How to compare simply and complex models?
- **Which Effects to Include**
  - Biological data is very imprecise
  - Often the biological information is confusing or even contradictory
- **Automatically Generated Models**



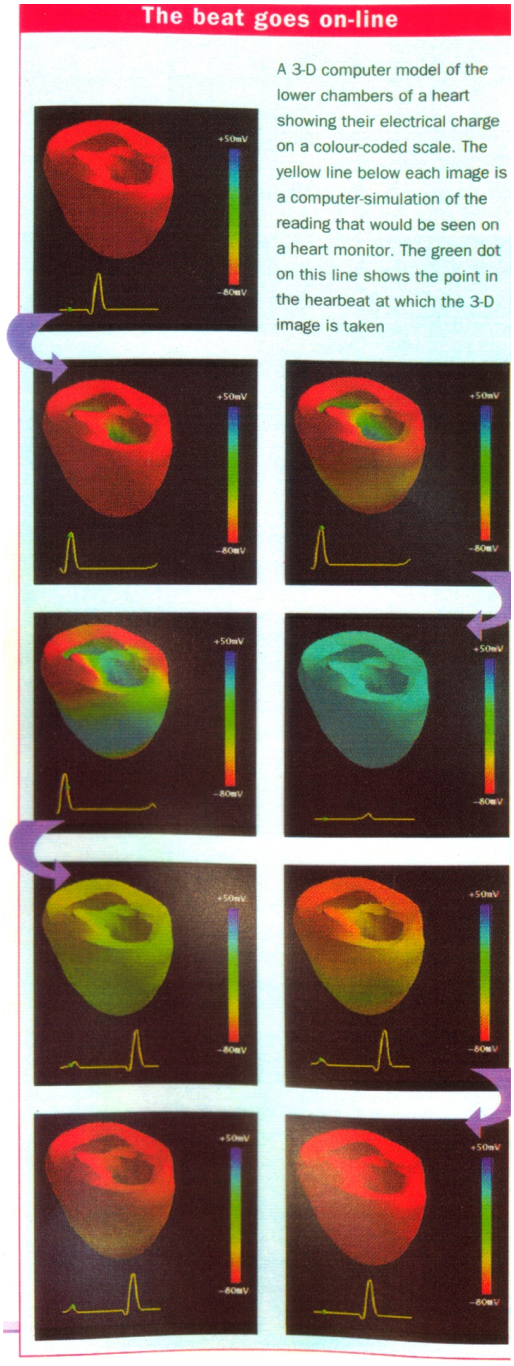
Virtual Heart

D. Noble *et al*

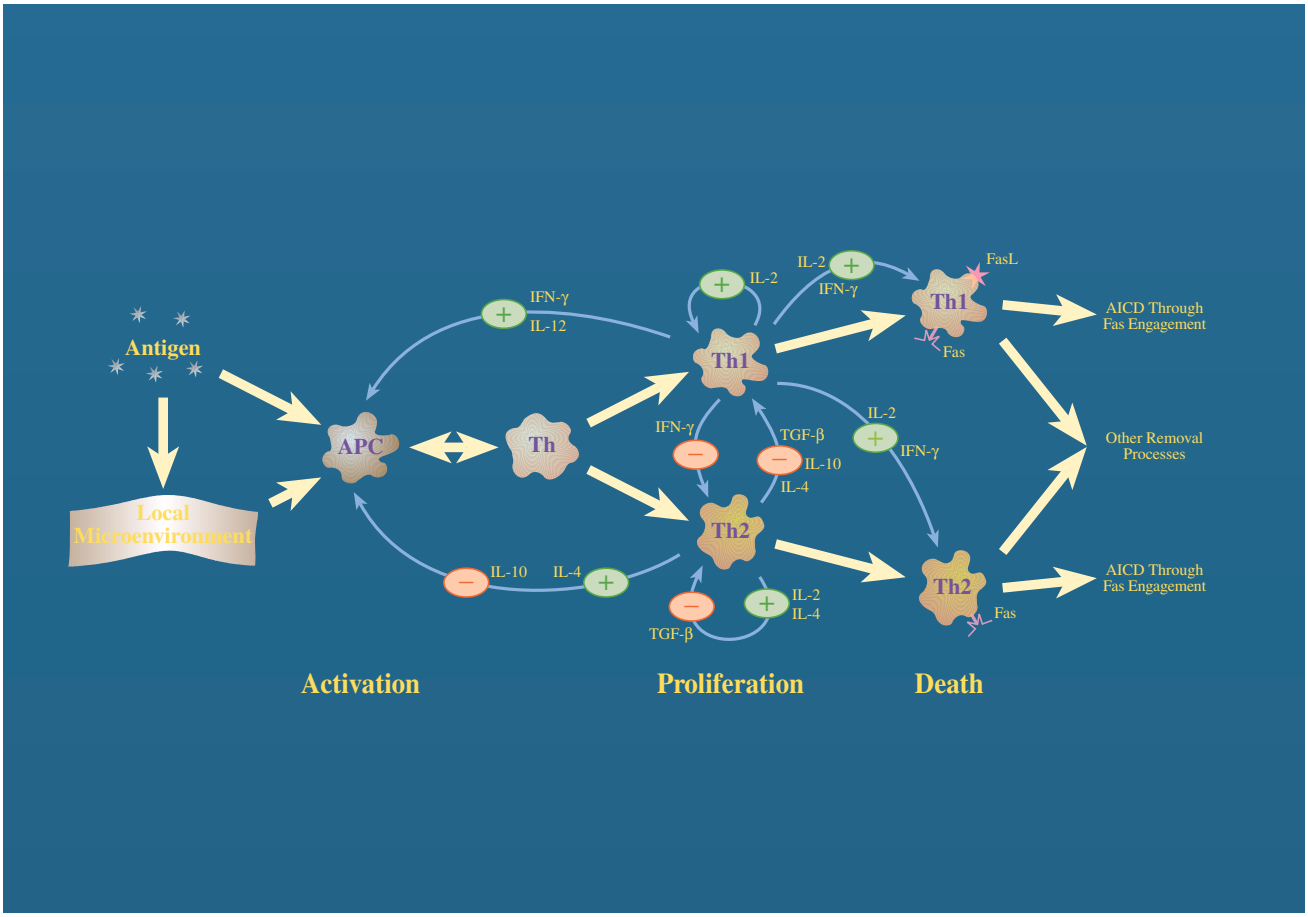
- Anatomically correct model of heart tissue
- 50+ membrane transport channels
- Resolution down to 10^3 cells $\approx 1mm^3$
- Can Correctly Reproduce Normal ECG
- Can Correctly Reproduce Arrhythmias Due to Drug Action or Ischaemia



Virtual Heart II



T Cell Differentiation

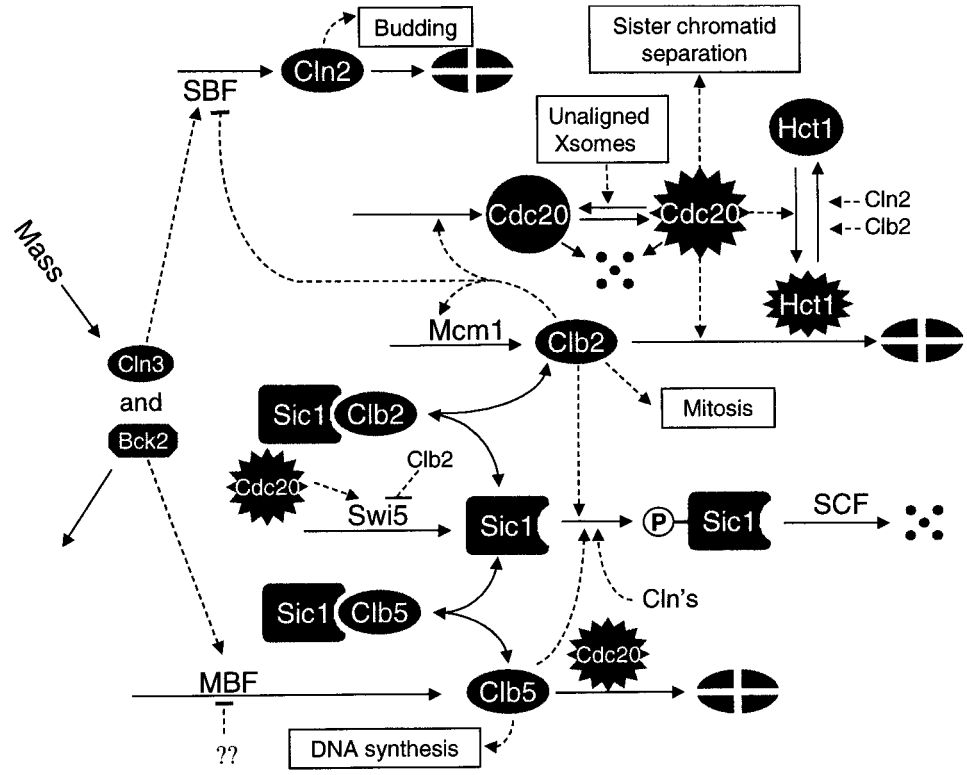


$$\frac{\mathrm{d}x_1}{\mathrm{d}\tau} = \sigma_1 \frac{\theta_1(1+x_2)+x_1}{(1+x_2)^2} + \pi_1 x_1 \frac{x_1+\chi_0}{(1+x_2)^2} - \delta_1 \left(\frac{x_1+\chi_0}{1+x_2}\right) x_1^2 - x_1,$$

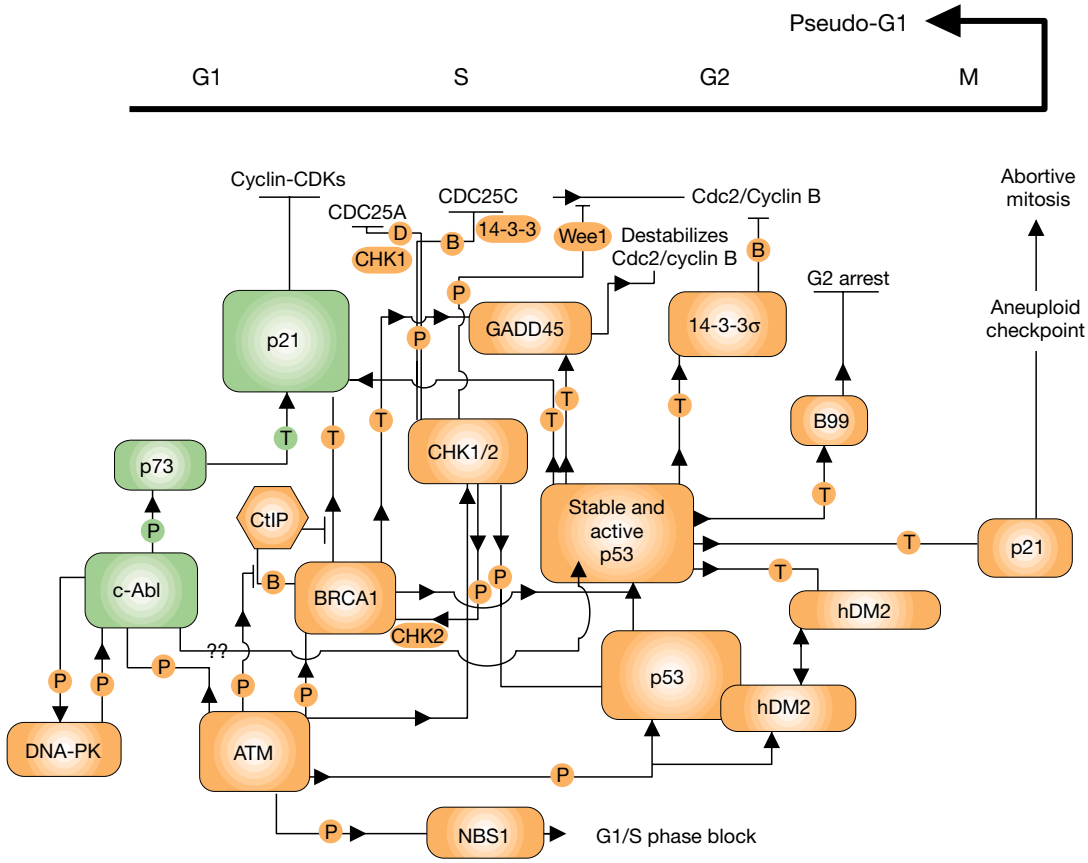
$$\frac{\mathrm{d}x_2}{\mathrm{d}\tau} = \sigma_2 \left(\frac{\theta_2+x_2}{1+x_2}\right) + \pi_2 \left[\rho + \left(\frac{x_1+\chi_0}{1+x_2}\right)\right] \frac{x_2^2}{1+x_1+x_2} - \delta_2 \left(\frac{x_1+\chi_0}{1+x_2}\right) x_1 x_2 - x_2.$$



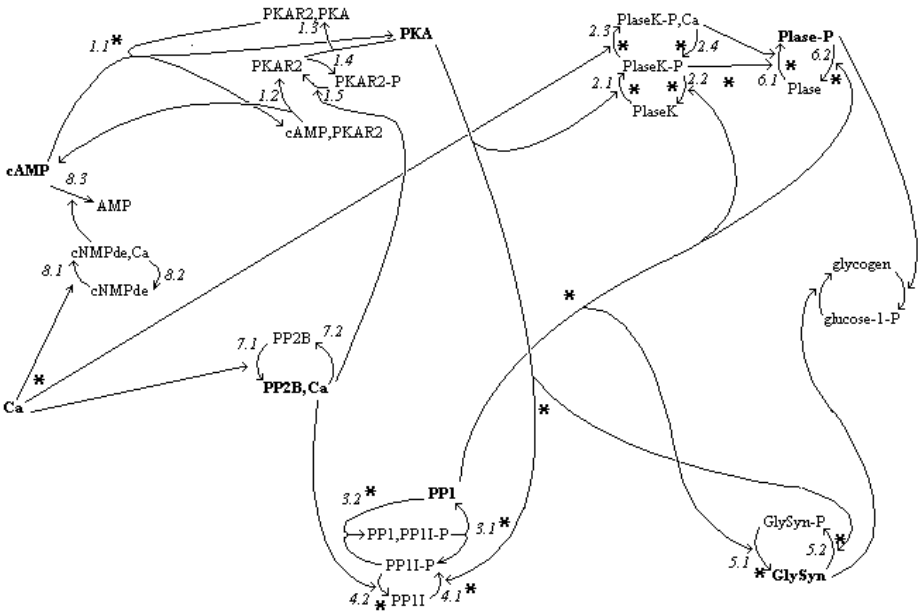
Same or Different?



Same or Different? II



Automatic Models



Equivalence of Systems

- **Models should be experimentally testable**
- **Insufficient to reproduce data**
  - Discrepancies from reality can be very informative
- **How do we compare model to data?**
  - Qualitative v. quantitative fit
  - Need for real parameters
  - Statistical approaches are often too difficult
- **How do we compare two models?**
  - Standard notions of conjugacy are not appropriate



Conclusions

- **Need for biological relevance**
- **Build strong collaborations**
  - Build up a library of good examples and role models
- **Develop tools for the analysis and simplification of complex biological models**
- **Create mathematical frameworks for comparison of models and data**
- **Have fun!**

