Modeling Genetic & Biochemical Processes

Tomorrow's forecast: Mostly sunny with a chance of cell division

(0) Intro

(1) Bacteriophage T7 model

(2) Thoughts on modeling

(3) Yeast pheromone signal transduction model

(4) Other

Drew Endy

ITP

April 18, 2001

Santa Barbara

Complementing(?) Approaches

(1) **Top-down, existing data**

e.g., analysis of mRNA expression array data leads to underlying genetic regulatory network architecture.

(2) **Bottom-up, existing data**

e.g., bacteriophage λ genetic-switch model.

(3) **Bottom-up, driving experiments**

e.g., bacteriophage T7 development, yeast pheromone mating response.

(4) **Bottom-up, designed/re-factored biological systems**

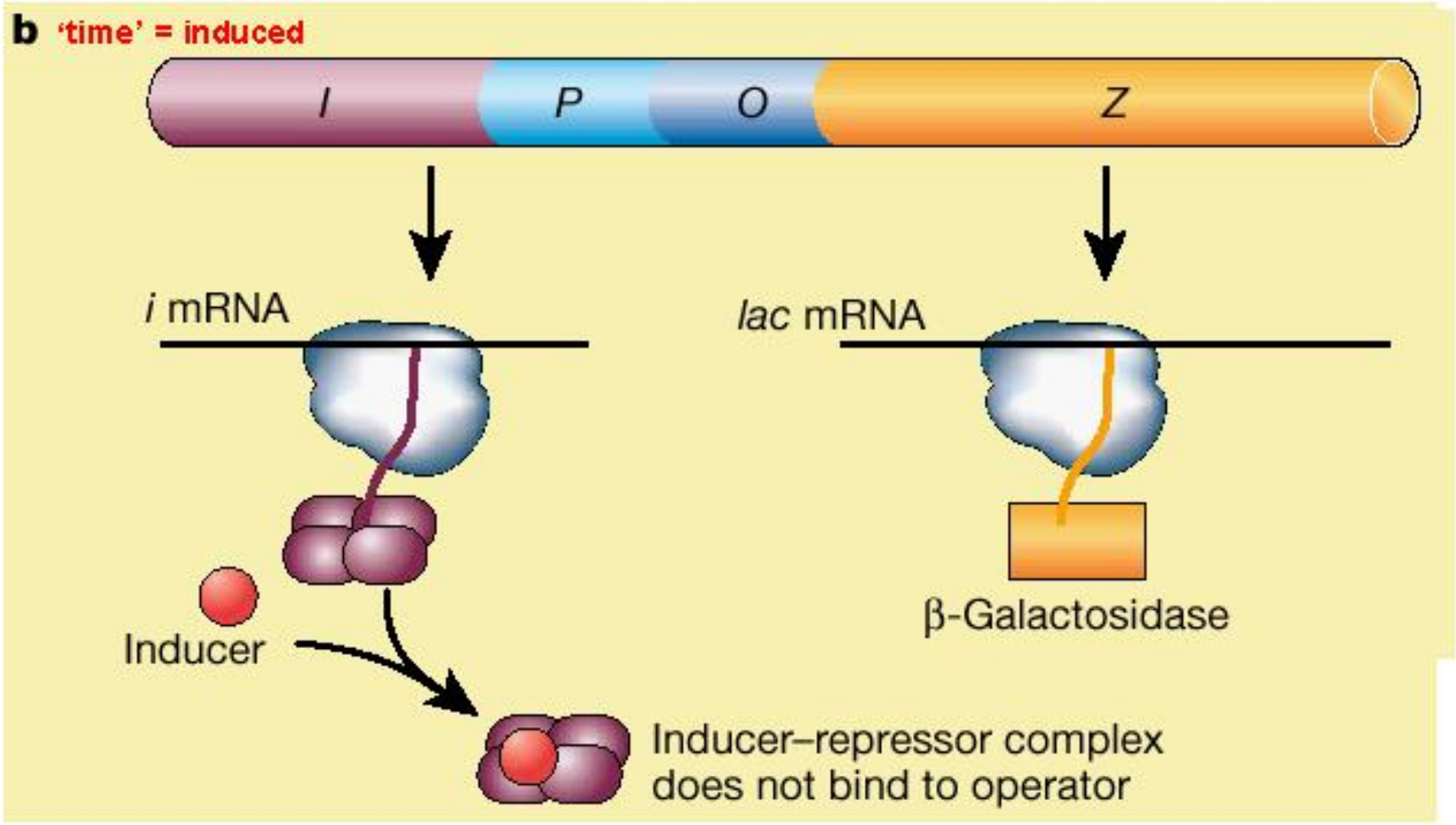
e.g., bistable switch, repressillator, frankenphage.

(5) **Module/sub-module (whatever) analysis**

e.g., "use-it-or-lose-it", adaptation in chemotaxis

Biology has a good history of model building

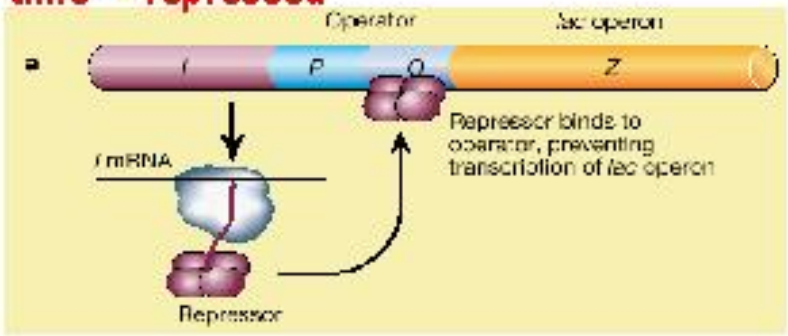
For example: ***lac* operon** (see Benno Müller-Hill, 1996; Dienert (1900) Ann Inst Pasteur **14**:139-189)



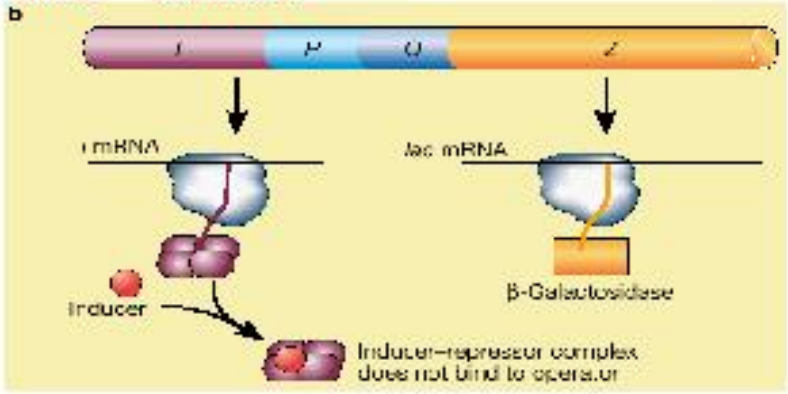
Quantitative Models?

Model

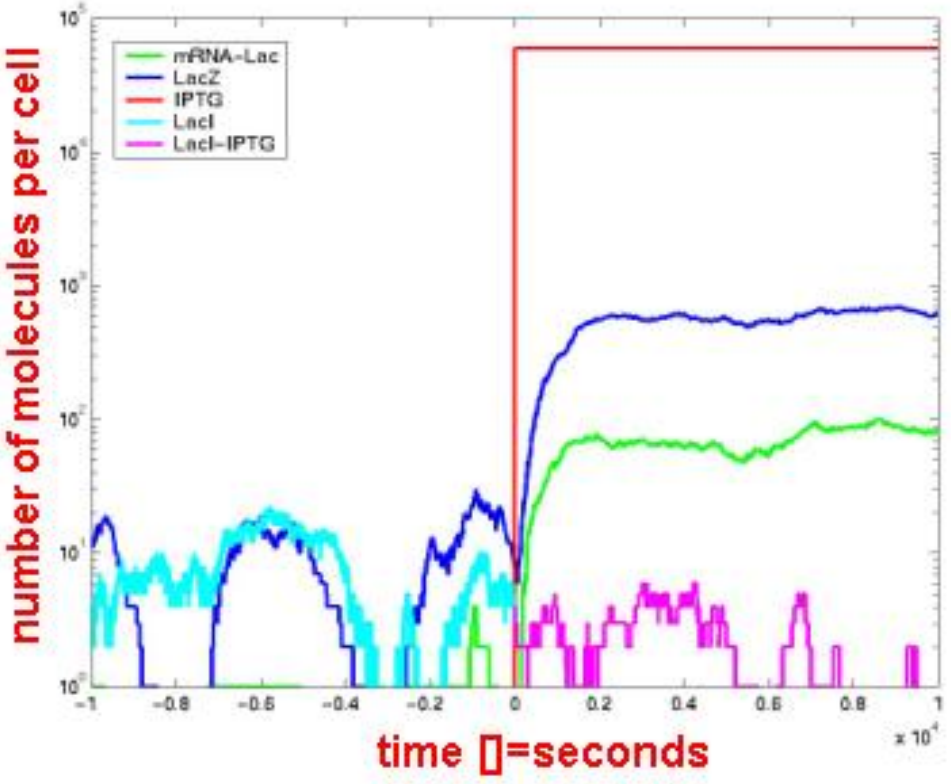
'time' = repressed



'time' = induced



Model output



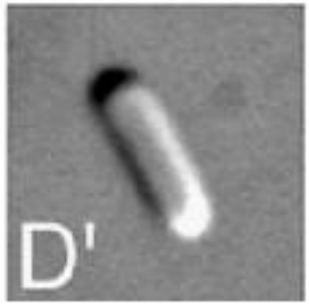
Needed: Better description of process physics

Even when the system is "specified", it is still ***very hard*** to represent a biological system realistically.

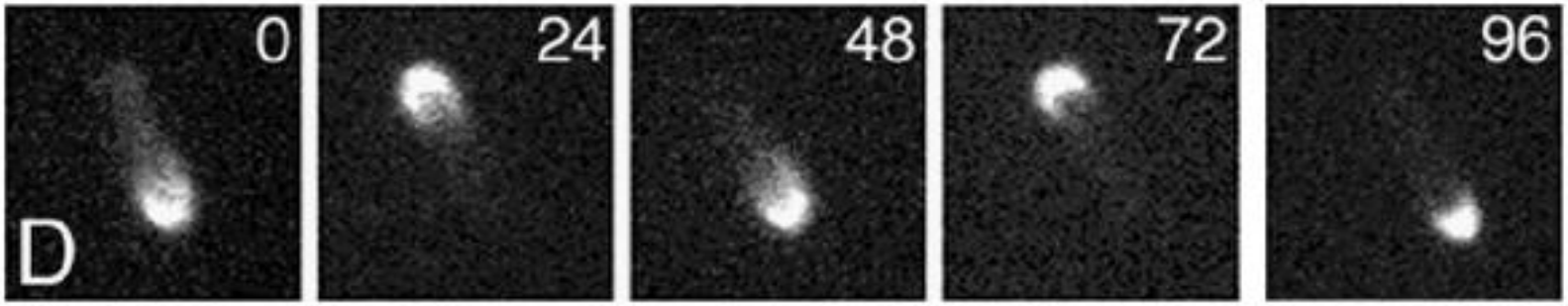
E. coli cytoplasm is ~200-320 mg/ml protein

GFP has an apparent diffusion coefficient **11-fold** lower in *E. coli* cytoplasm than in water.

Elowitz, *J. Bact* **181**:197 (1999).



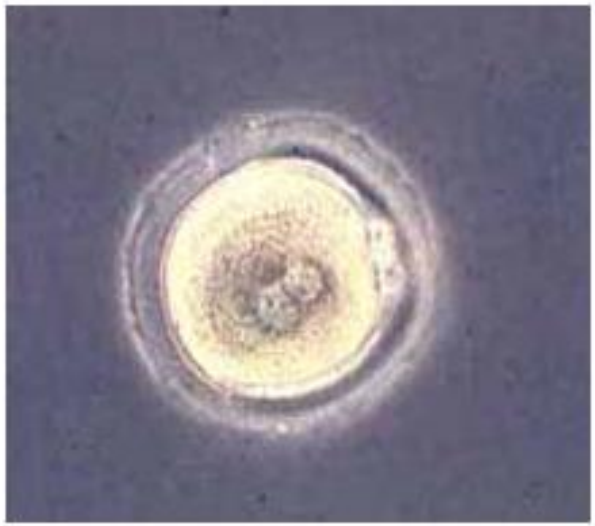
minC-GFP translocation



Raskin & DeBoer, *J.Bact. & PNAS-USA* (1999)

E-human 1.0?

Human



time < 1 week

time > 70 years

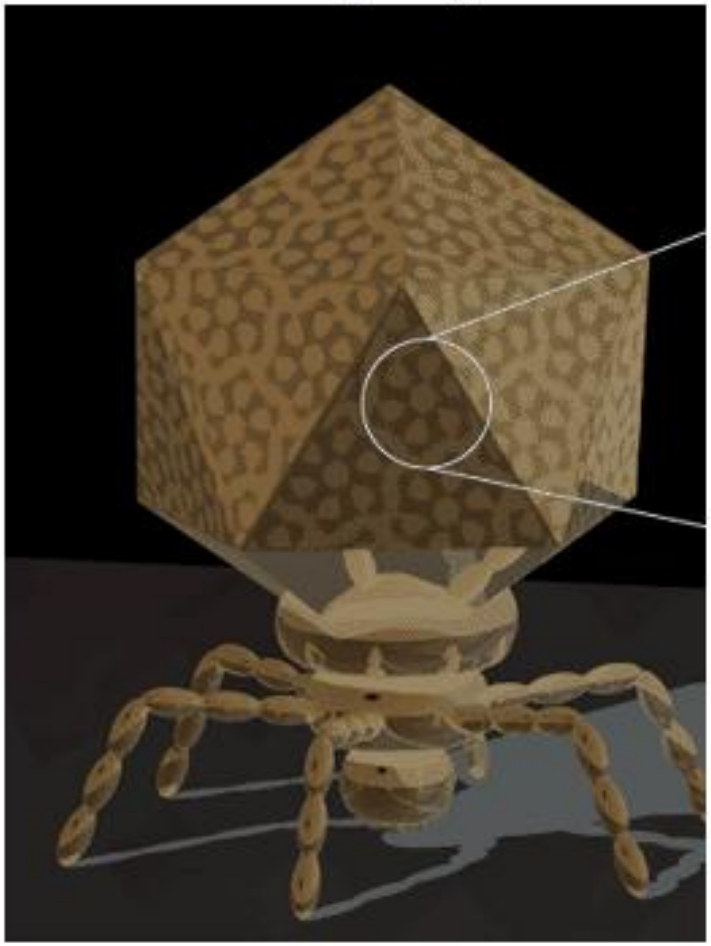
nope

why?

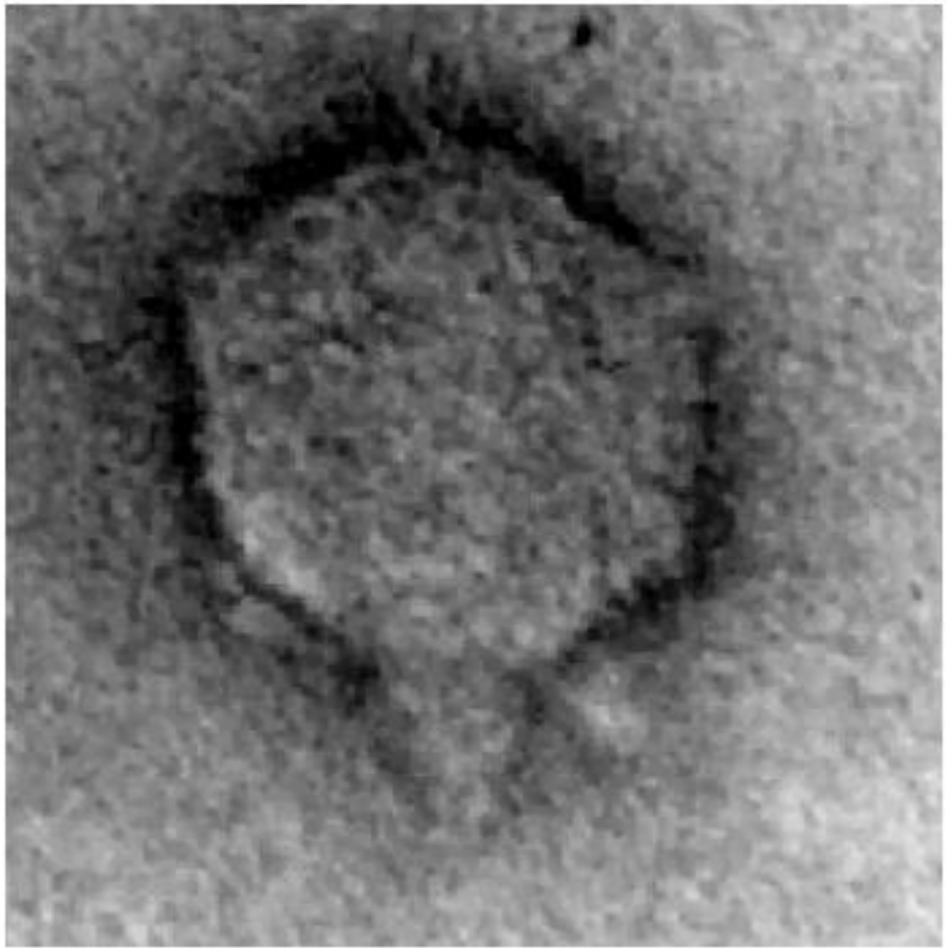
(1) incomplete specification
(2) unknown transition rules
(3) uncertain physics
(4) unknown levels of resolution
(5) computational methods/power
(6) unknown function

Possible now? Maybe bacteriophage T7 model…

What is phage T7?



Novagen, Inc



Spires & Brown, unpublished

Possible now? Bacteriophage T7 model…

What do we know about T7?

"…T7 was isolated from standard anti-coli-phage mixture prepared by Dr. W.J. MacNeal…" Demerec & Fano (1944).

Genome sequenced (100%) Dunn & Studier (1982).
- 39,937 bp linear dsDNA genome
- 56 genes thought to encode 60 proteins
- Primary function assigned to ~33 genes
- > 50 regulatory elements

'Protein' linkage map via yeast two-hybrid, Bartel et al. (1996).

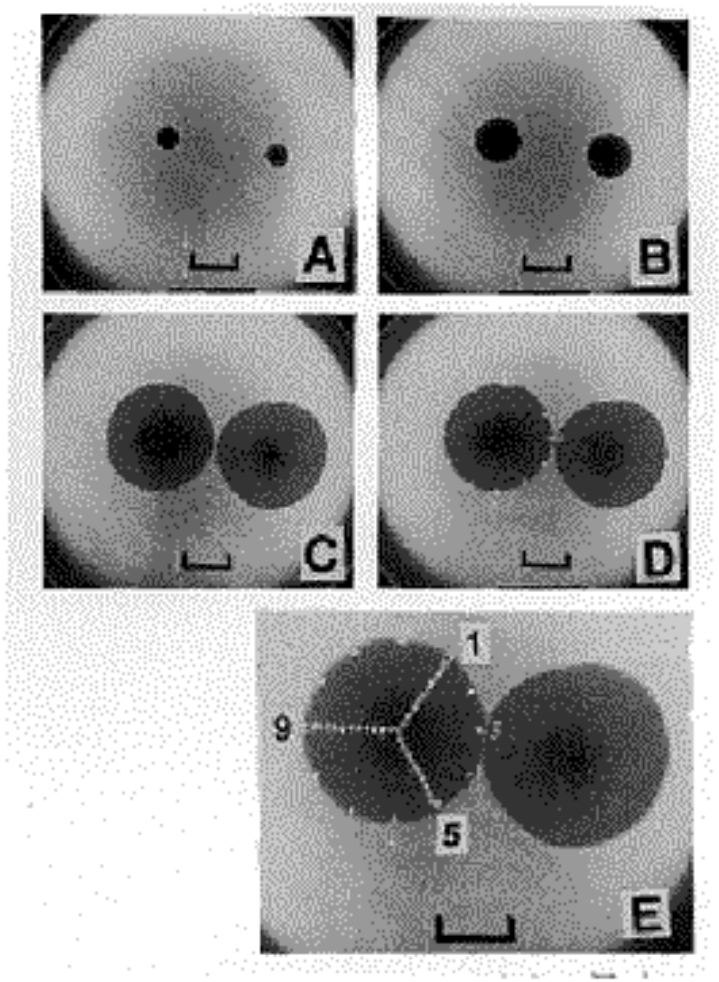
~62nm diameter particle, short tail.

T7 RNA and DNA polymerases extremely well characterized.

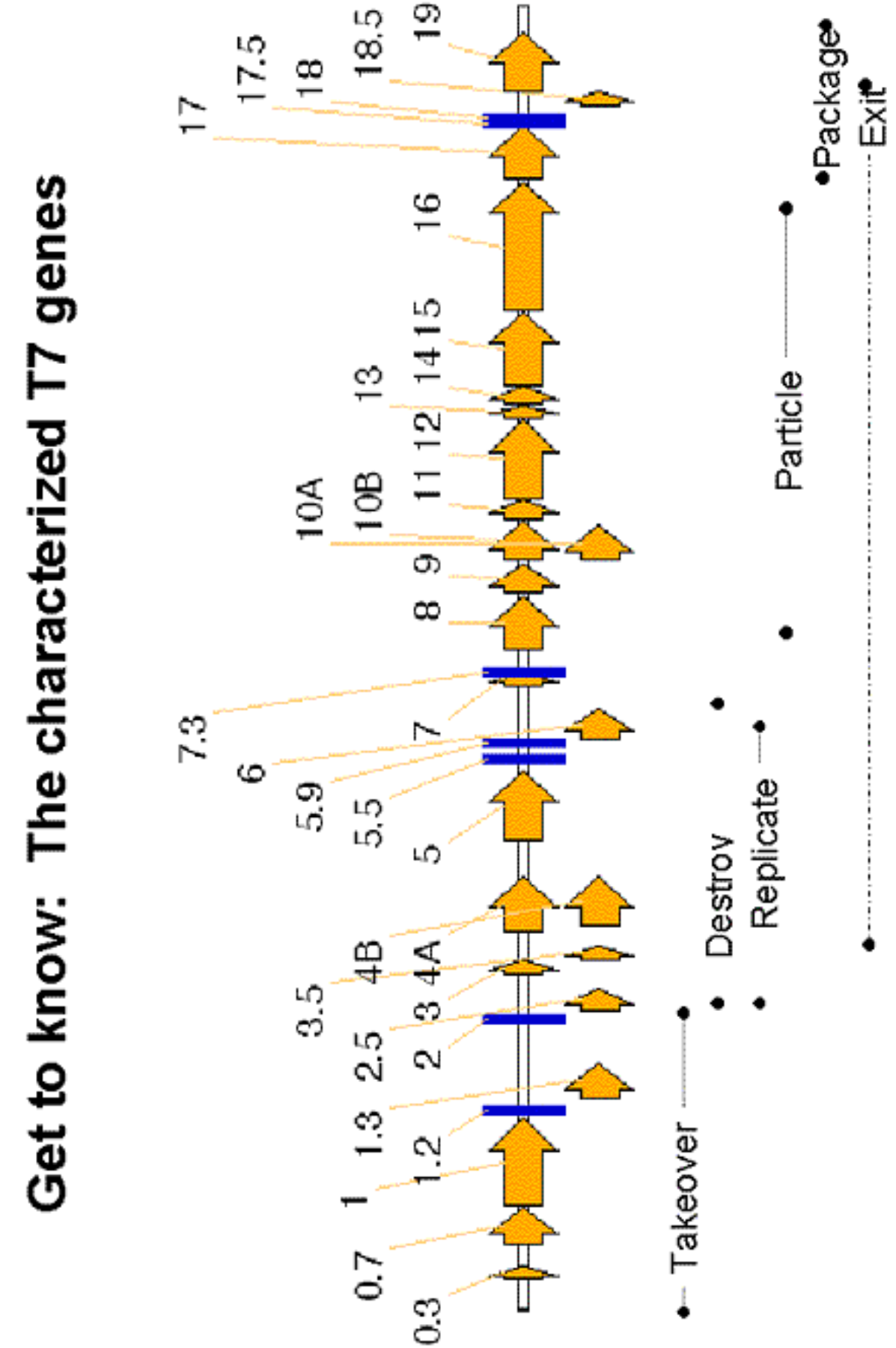
And much more…

T7 is a relatively well studied system.

Innocent beginnings



John Yin, *J.Bact.* **175**:1272 (1993)

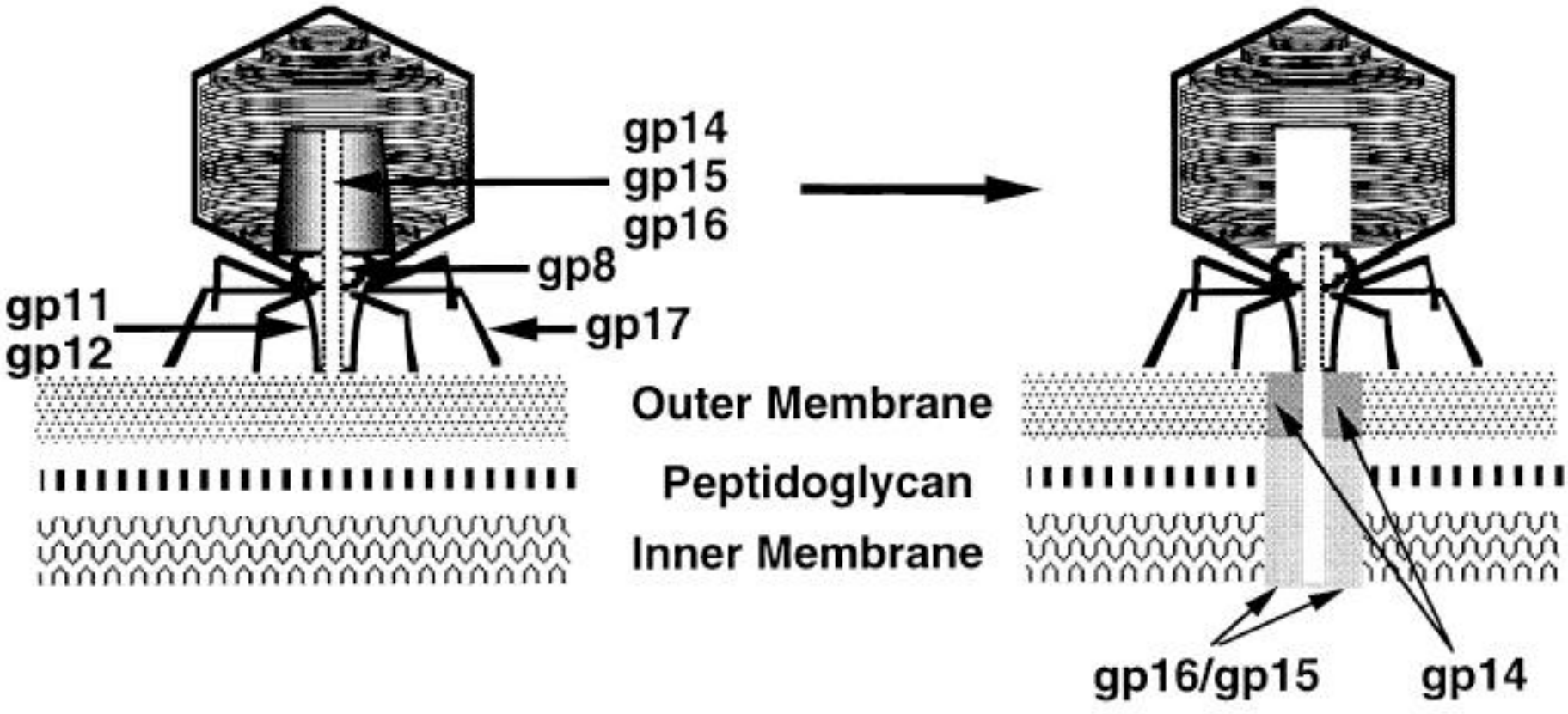

Get to know: The characterized T7 genes
0.3
0.7
1
1.2
1.3
2
2.5
3
3.5
4A
4B
5
5.5
5.9
6
7
7.3
8
9
10A
10B
11
12
13
14
15
16
17
17.5
18
18.5
19
Takeover
Destroy
Replicate
Particle
Package
Exit

T7 model is a coupled series of smaller models

(1) DNA entry
(2) Transcription
(3) Translation
(4) Replication
(5) Prohead assembly
(6) DNA packaging
(7) Particle assembly
(8) Lysis

Given a T7 particle bound to host…

T7 DNA entry:



Molineux, *Molec Microbio.* **40**:1 (2001)

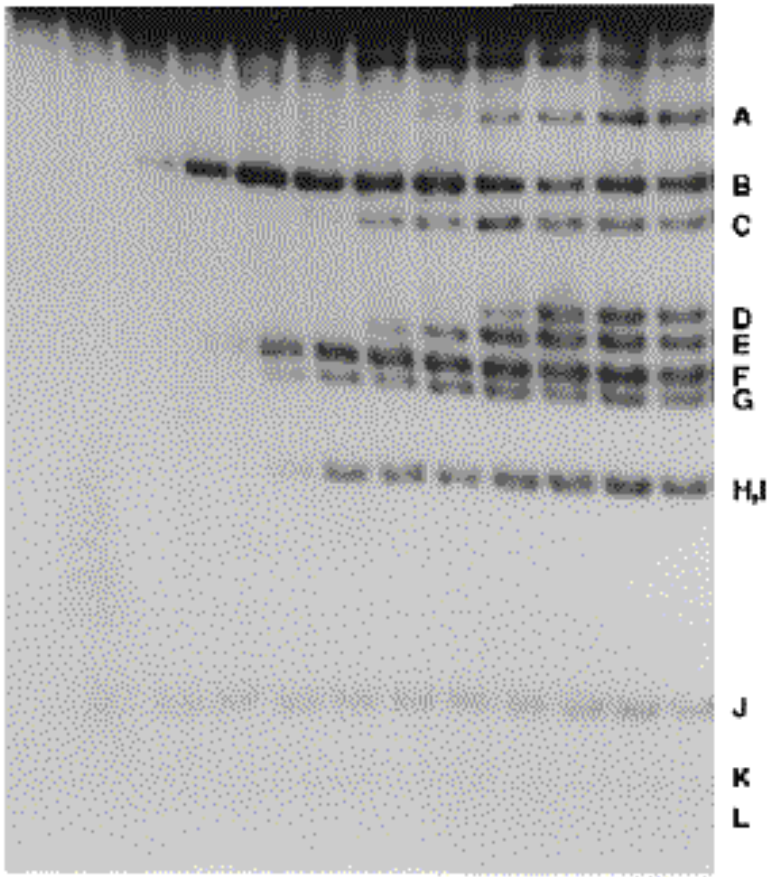
Example DNA entry experiment

(1) J K B M F G H I C E A L D — *Dpn*I sites (cuts at GATC given M+)

restriction fragment time course

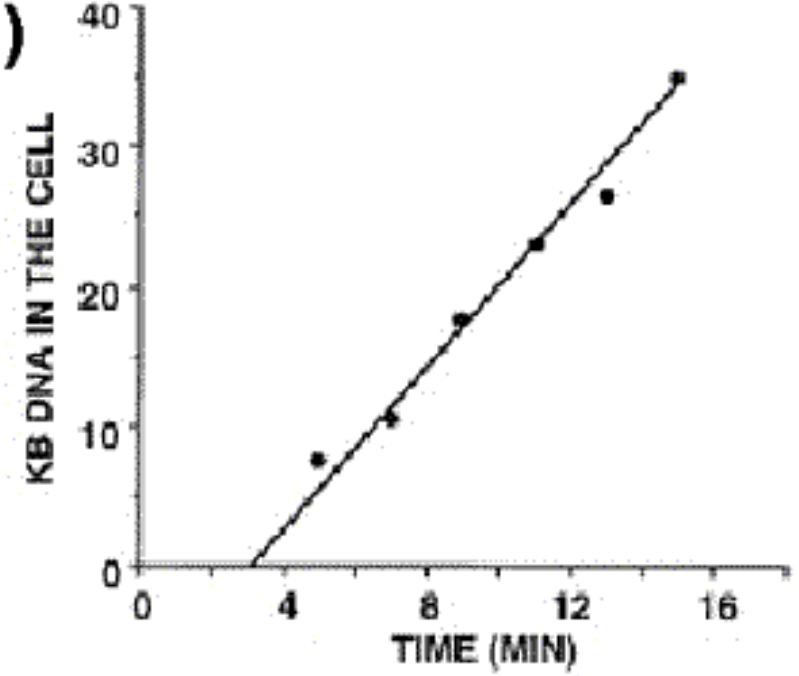
1 3 5 7 9 11 13 15 17 19 21 23

(2)



RNAP mediated entry

(3)



Garcia & Molineux, *J.Bact.* **177**:4066 (1995)

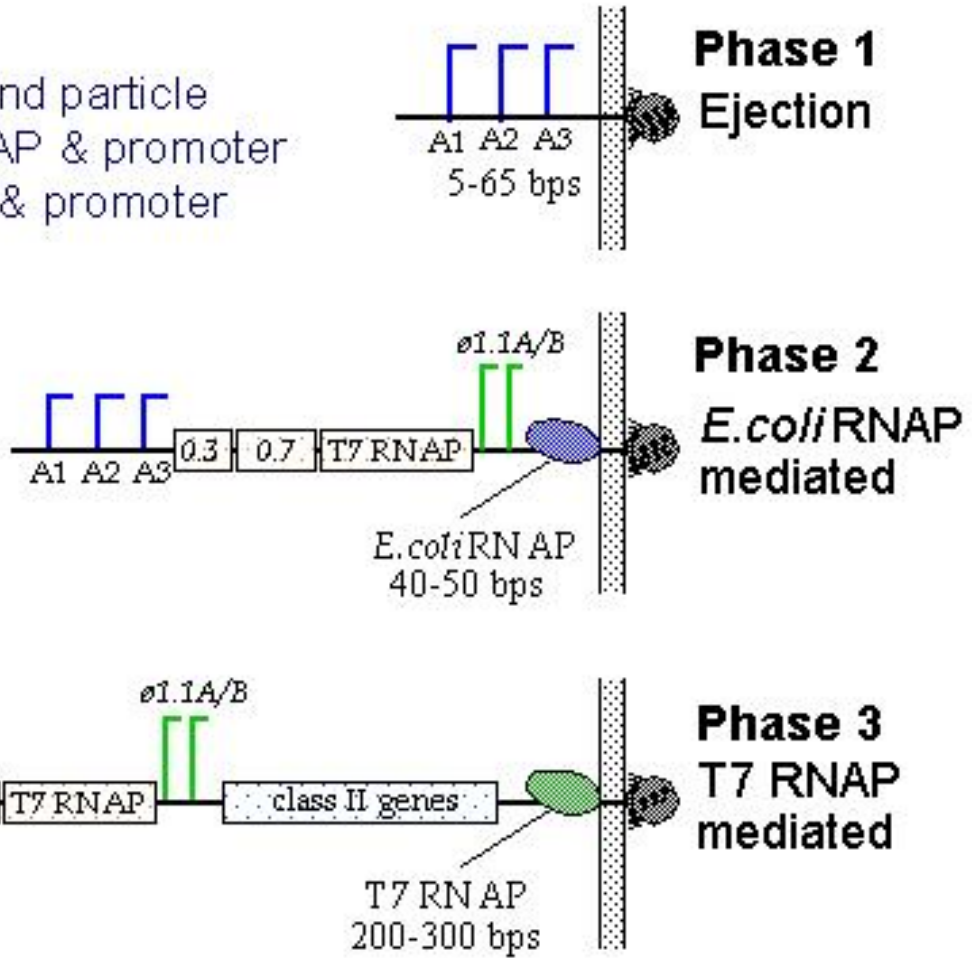
Series of models

Process

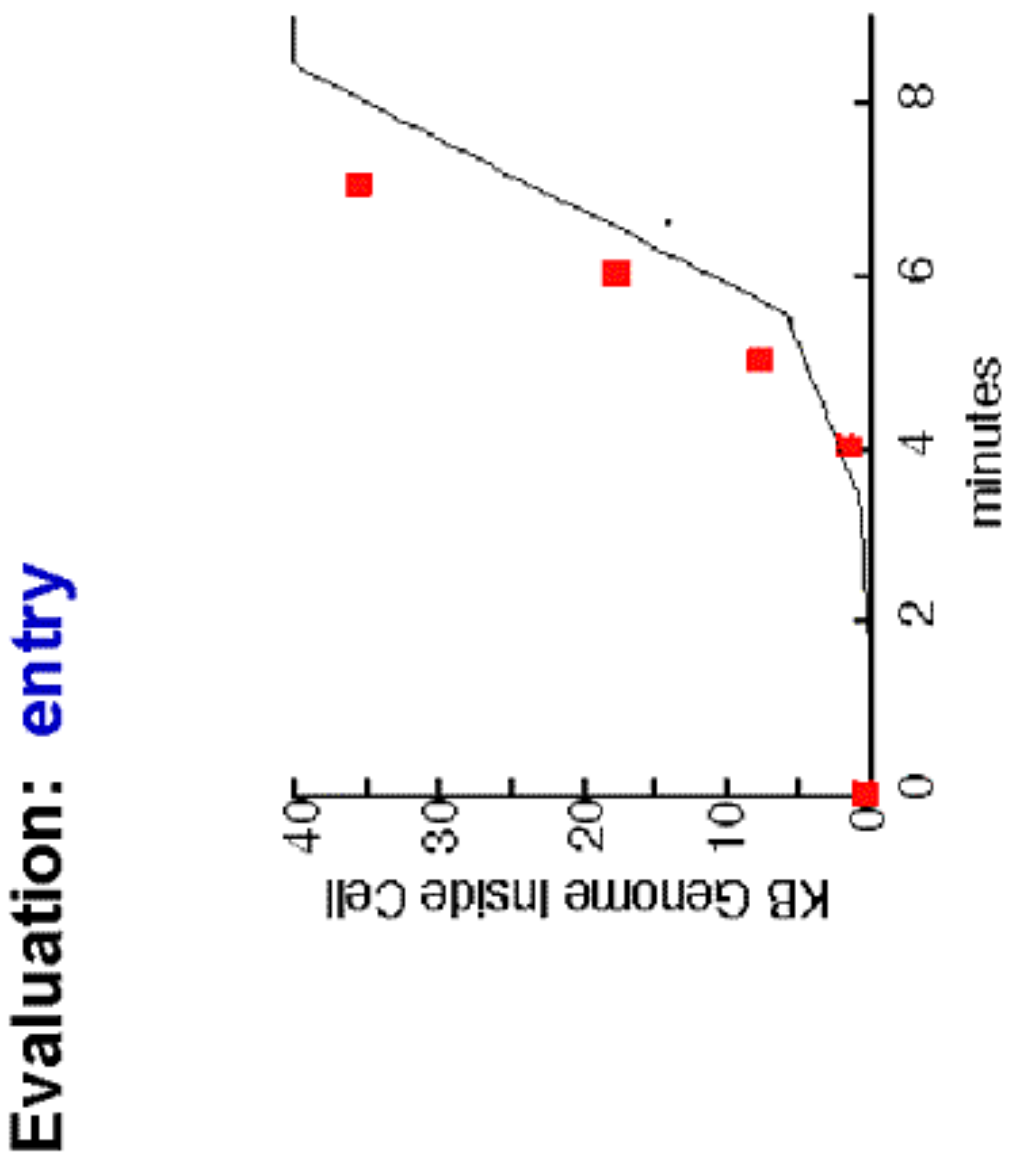
DNA entry
Transcription
Translation
DNA replication
Capsid assembly
DNA packaging
Particle assembly
[Lysis]

Model

α, bound particle
L' = β, RNAP & promoter
γ, gp1 & promoter



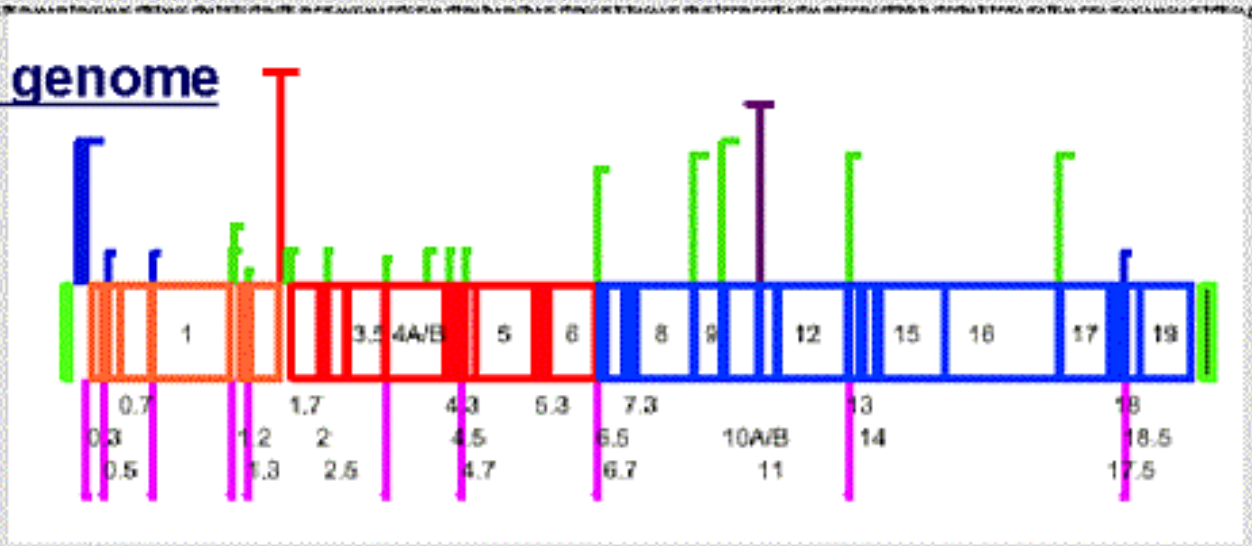
[15] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development



Model resolution = genetic elements

T7 sequence

T7 *model* genome

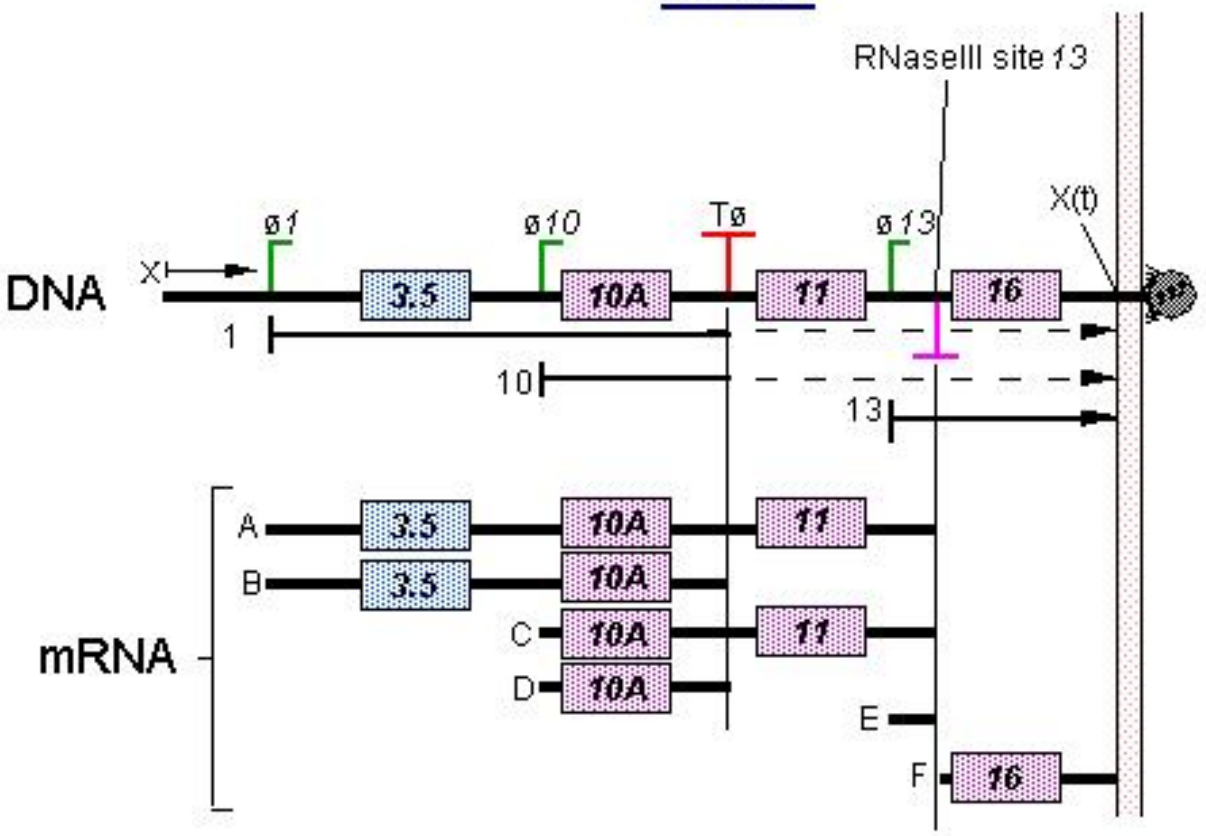


Series of models

Process

DNA entry
Transcription
Translation
DNA replication
Capsid assembly
DNA packaging
Particle assembly
[Lysis]

Model

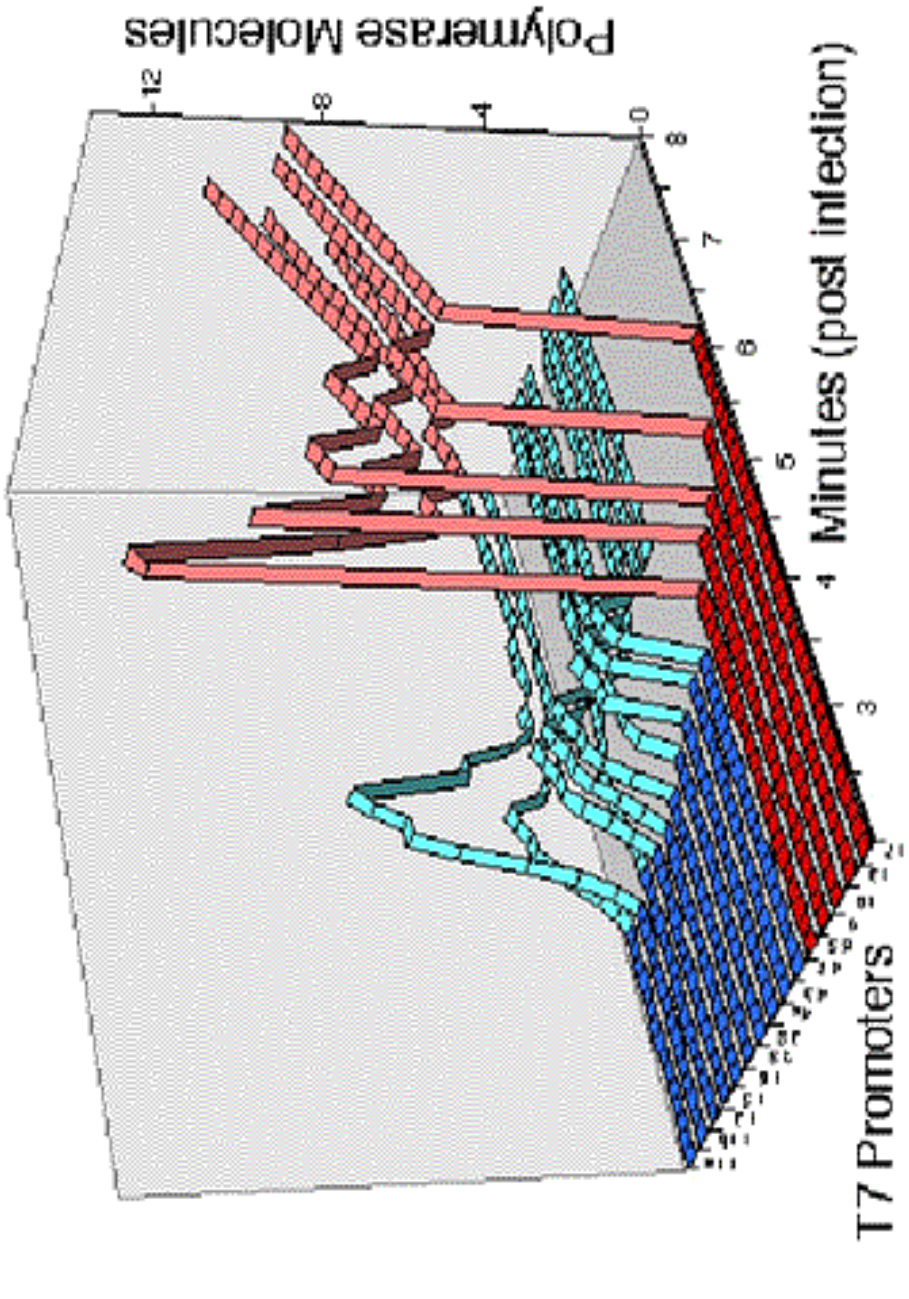


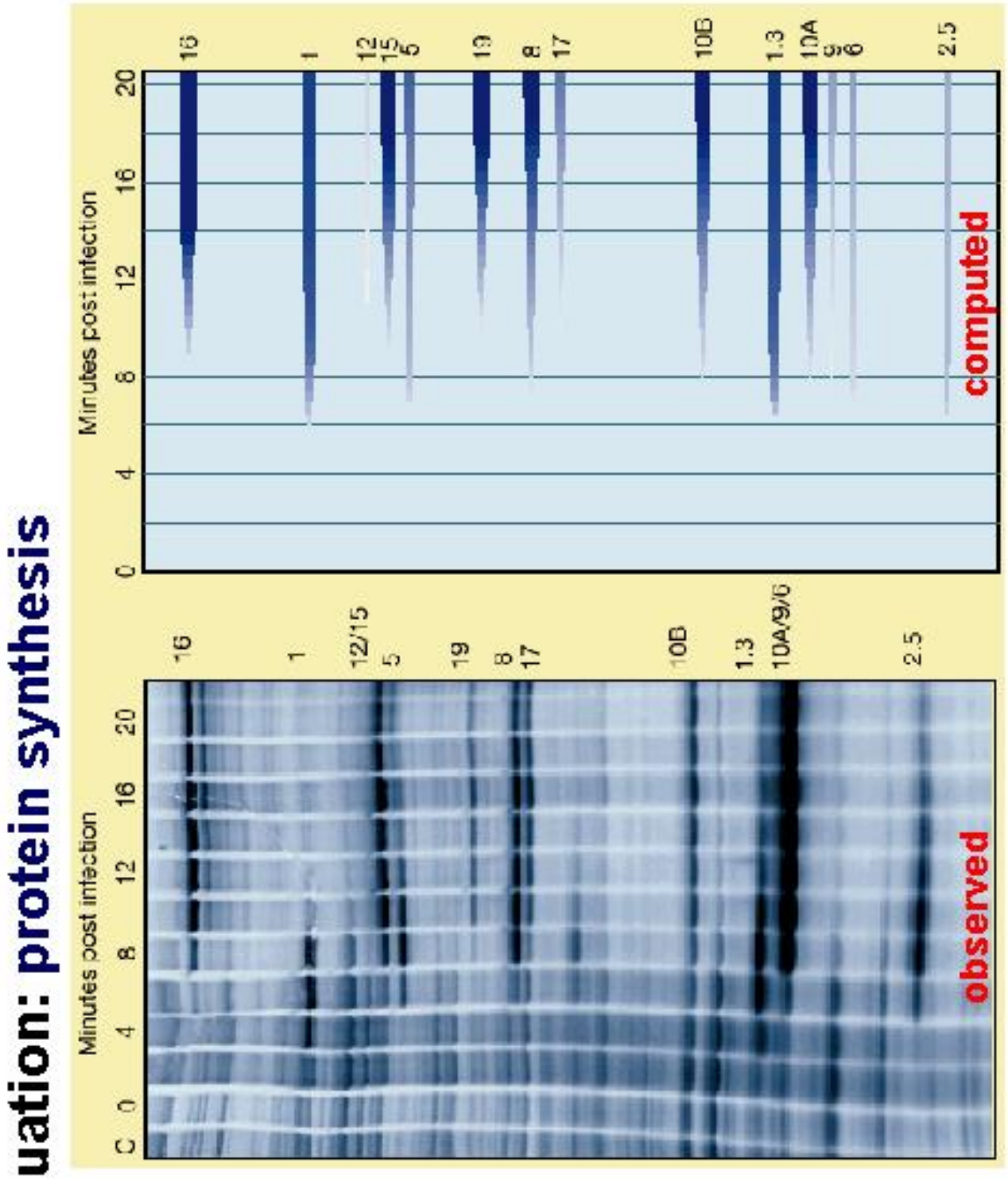
Step 1: Assign RNA polymerase to each promoter
Step 2: Set RNA polymerase density for each mRNA
Step 3: Set translation level based on mRNA species

[18] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development

$$\frac{d(mRNAi)}{dt} = [(k_{Ph})(S_e i) + (k_{PT7})(S_l i)] - (k_{dm})(mRNAi) \text{ for } i = 1 \text{ to } N_m$$

$$\frac{d(gpi)}{dt} = [(k_T)(R_d)(mGPi) - (k_{dgp})(gpi) + T_l \quad \text{for } i = 1 \text{ to } N_{gp}$$




Evaluation: protein synthesis
Minutes post infection
0
4
8
12
16
20
16
1
12/15
5
19
8
17
10B
1.3
10A/9/6
2.5
observed
computed

Series of models

Process	Model
DNA entry	
Transcription	
Translation	
DNA replication	"In the simulation we assume that a replisome consists of a gp5:thioredoxin complex and six gp4A molecules. We assume that replisomes form instantly given stoichiometric amounts of these components. Furthermore, we assume thiroredoxin is in excess at all times and all T7 DNA polymerase molecules (gp5) become processive after forming a 1:1 complex with thioredoxin. We also assume that up to two replication forks can form per complete T7 genome, given sufficient numbers of replisomes. This last assumption allows for an exponential increase in the total rate of DNA synthesis as replication proceeds. The roles of gp1, gp2.5, and gp3.5 are accounted for by incorporating "hard" switches such that DNA synthesis occurs only if the concentrations of gp2.5 and the gp1:gp3.5 complex are non-zero. Having met these constraints, the simulation treats replisome elongation as the rate-limiting step in DNA synthesis. This rate is set as a function of dNTP concentrations using Michaelis-Menten kinetics. The simulation ignores any intermediate steps between the formation of acid-soluble DNA fragments created from the phage-mediated digestion of the host genome and their eventual conversion to dNTPs. We assume that the digestion of the host genome proceeds at a constant rate if the concentrations of the T7 endo- (gp3) and exonucleases (gp6) are non-zero" [L. You, Thesis Proposal, University of Wisconsin-Madison, Madison, WI (1999)].
Capsid assembly	
DNA packaging	
Particle assembly	
[Lysis]	

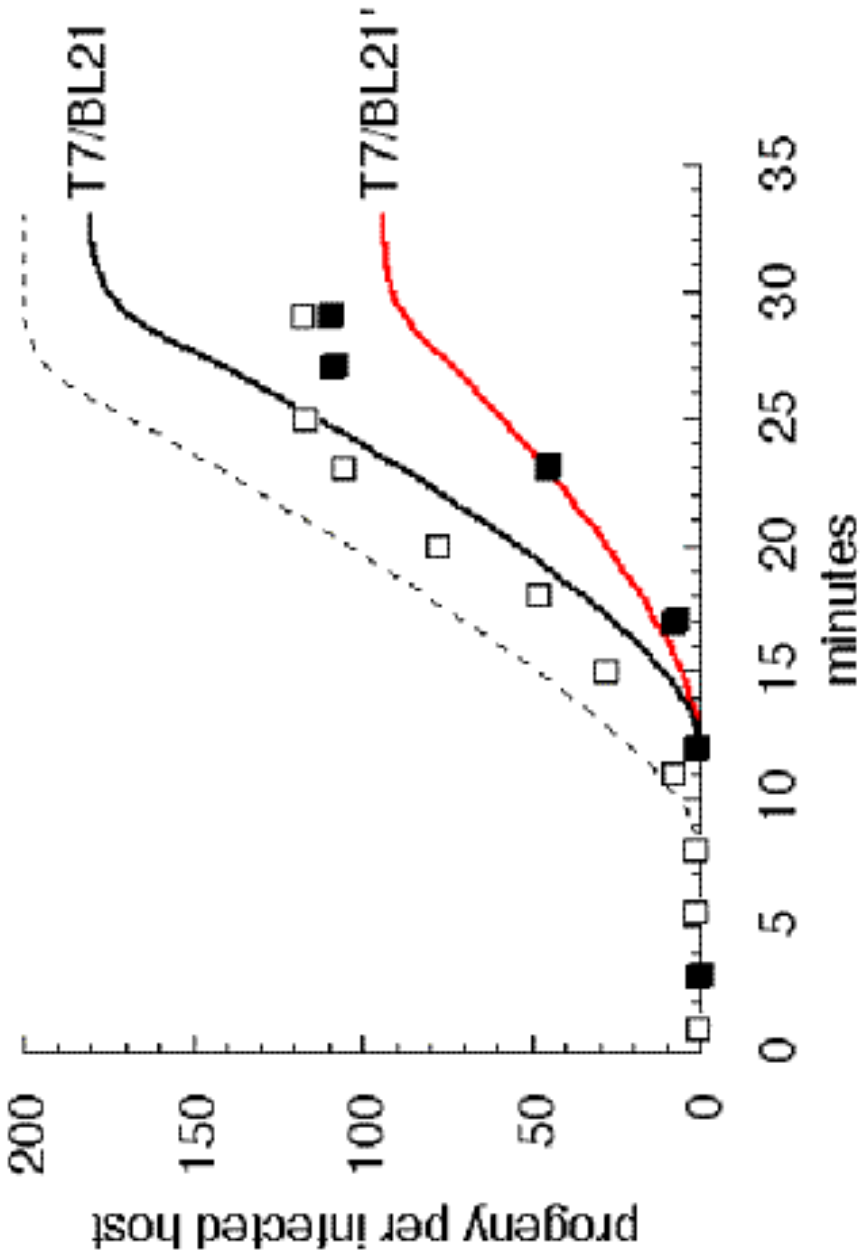
Series of models

Process	Model
DNA entry	
Transcription	
Translation	
DNA replication	
Capsid assembly	"Procapsid assembly is simulated with a 4.78-order nucleation-limited reaction derived from data for phage P22 (Prevelige et al., 1993). T7 and P22, both Podoviridae, have dissimilar genomes and growth cycles (Hausmann, 1988 in The Bacteriophages, R. Calendar, Ed., Plenum, NY), but their icosahedral capsids are each approximately 60 nm in diameter and are attached to short noncontractile tails (Ackermann & Berthiaume, 1995, Atlas of Virus Diagrams, CRC, NY). The P22 kinetic data for procapsid assembly is the most comprehensive for any phage and allows the development of a procapsid rate expression. This representation for the formation of procapsids includes the requirement that the major capsid protein concentration is above a nucleating concentration before assembly initiates. The consumption of procapsids as progeny are formed requires complete procapsids, T7 DNA, and enough of each structural protein to complete the phage. As procapsids and progeny phage particles are assembled the simulation accounts for the depletion of various T7 proteins. The simulation also accounts for the utilization of scaffold protein, gp9, during procapsid assembly. Unlike P22, which is known to recycle its scaffold protein, T7 does not (Roeder & Sadowski, 1977), despite the fact that gp9 does not remain in the final phage particle (Steven & Trus, 1983 in Electron Microscopy of Proteins: V, J.R.Harris & R.W. Horne, Eds., Academic Press, London). The simulation assumes that packaging of DNA into the procapsid is the rate limiting step for T7 progeny formation (given DNA, procapsids, and non-zero concentrations for all particle proteins and the DNA maturation proteins, gp18 and gp19)." *adapted from D. Endy, Dissertation, Dartmouth College, Hanover NH (1998)*
DNA packaging	
Particle assembly	
[Lysis]	

Series of models

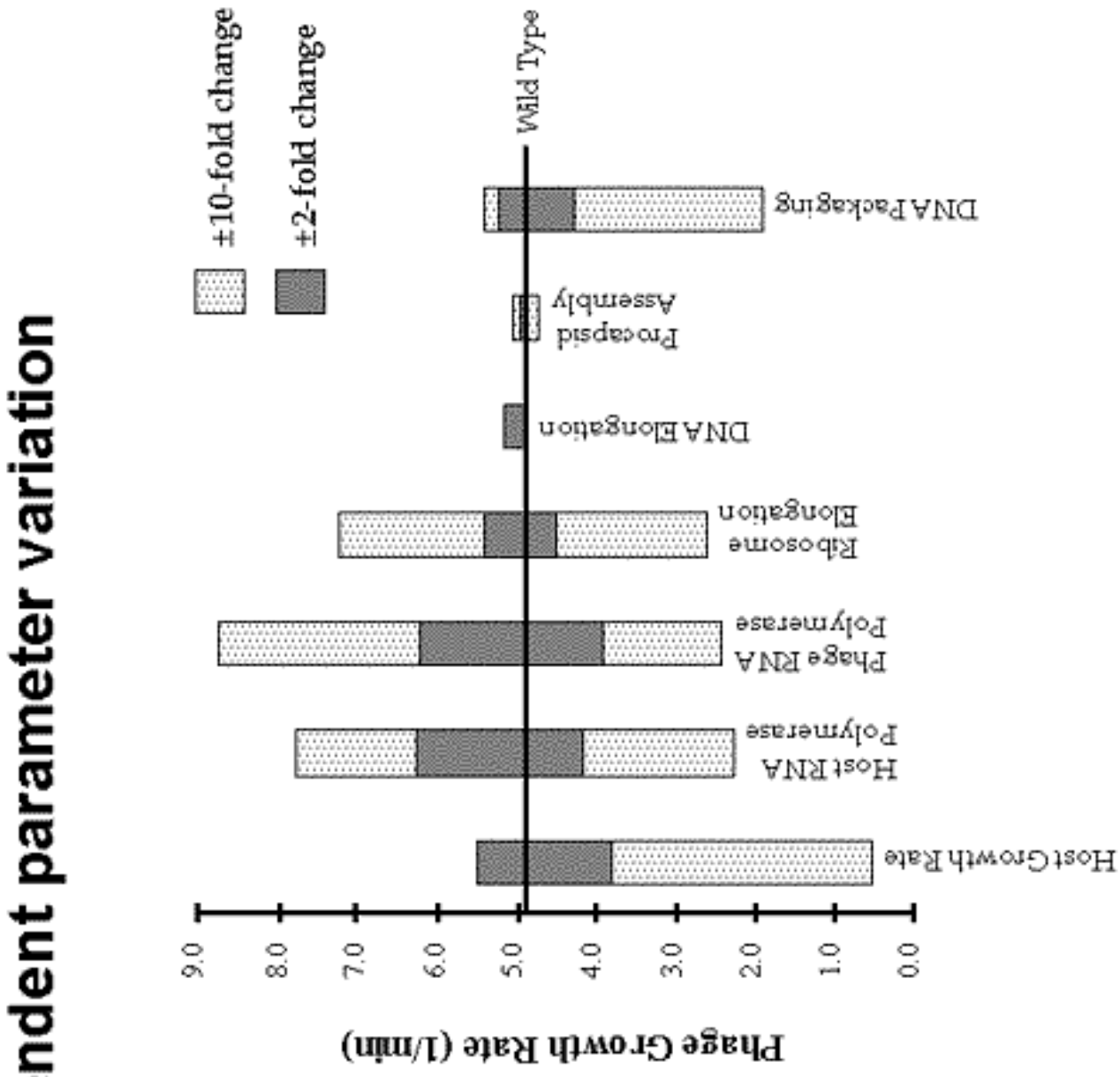
Process	Model
DNA entry	
Transcription	*not included*
Translation	
DNA replication	Concentration dependent relationship between degradation of host membrane(s) and phage proteins is unknown at present.
Capsid assembly	
DNA packaging	
Particle assembly	
[Lysis]	

Evaluation: intracellular progeny

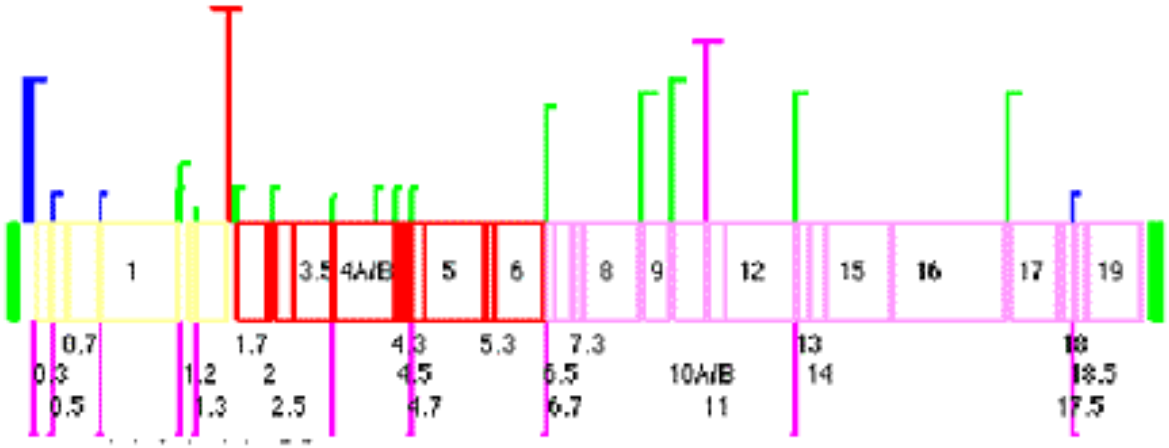


Experimental data and simulation results* from Endy et al., (1997)

*Simulated red line from T7v2 using $\eta_{pack} = 0.5$


Independent parameter variation
±10-fold change
±2-fold change
Wild Type
Host Growth Rate
Host RNA Polymerase
Phage RNA Polymerase
Ribosome Elongation
DNA Elongation
Procapsid Assembly
DNA Packaging
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Phage Growth Rate (1/min)

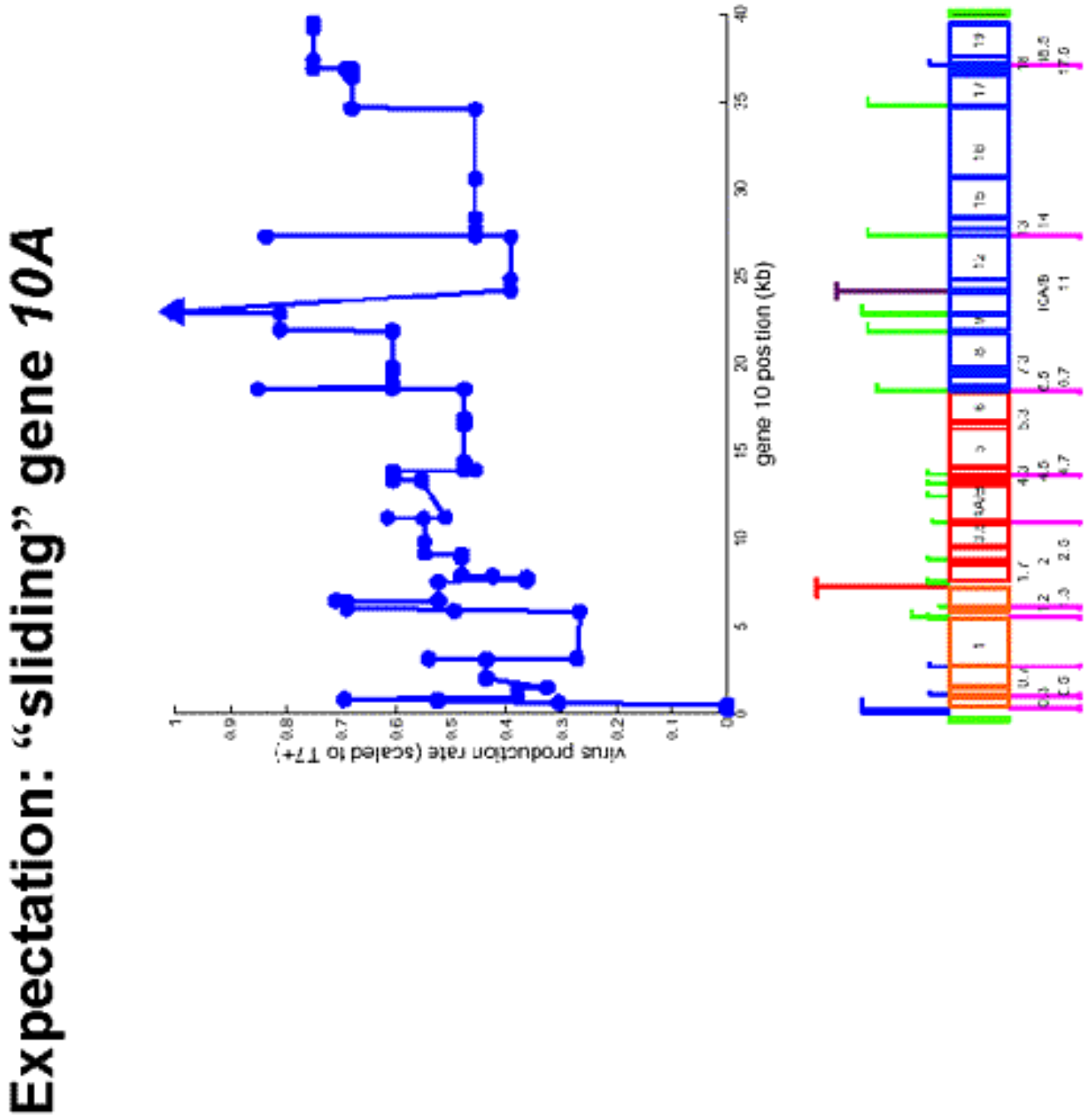
All this trouble for what?



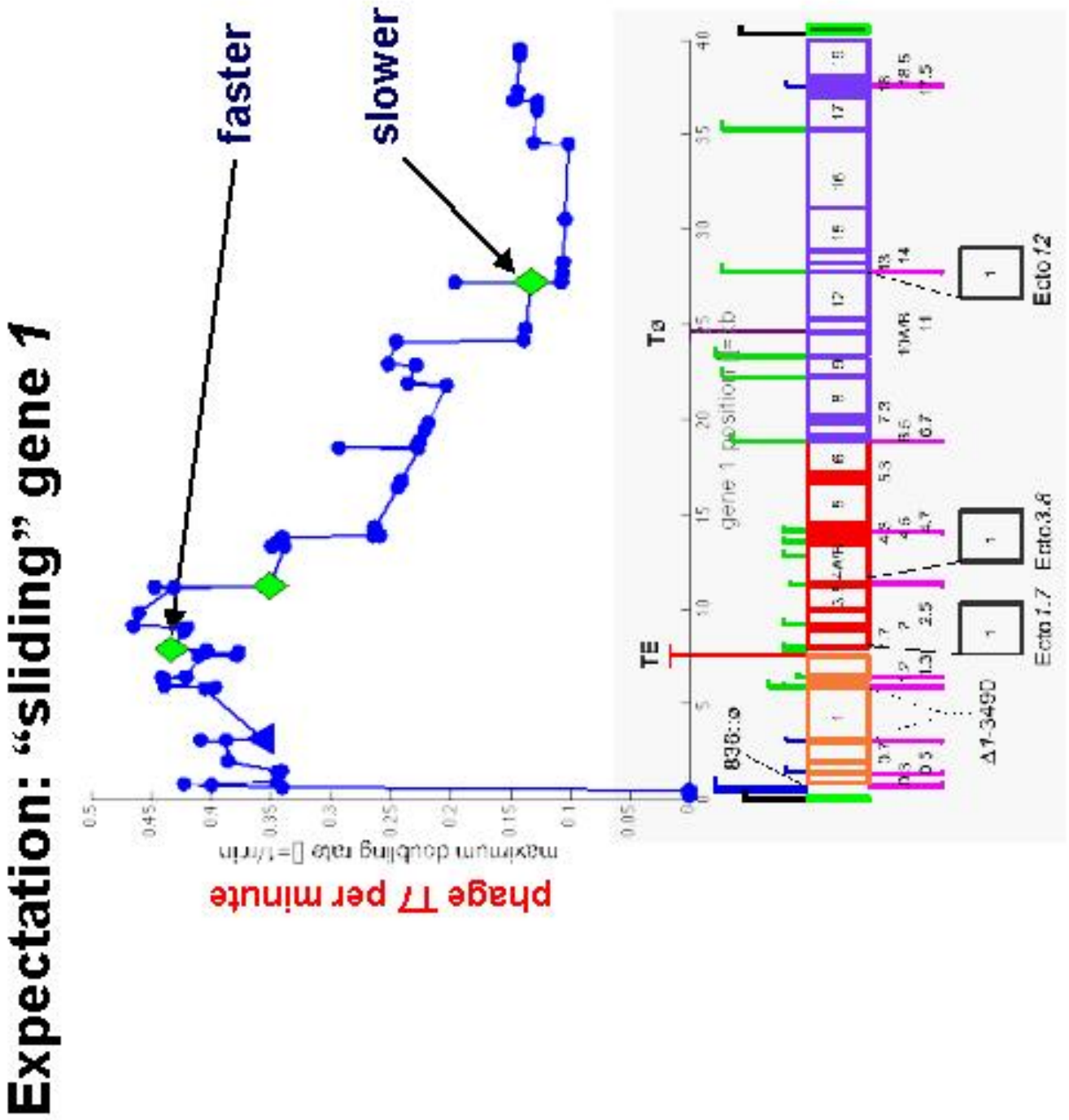
In T7, the position of a gene on the genome directly regulates the timing and level of gene expression.

If I have a model based on what is known, can I compute what happens when I move genes around?

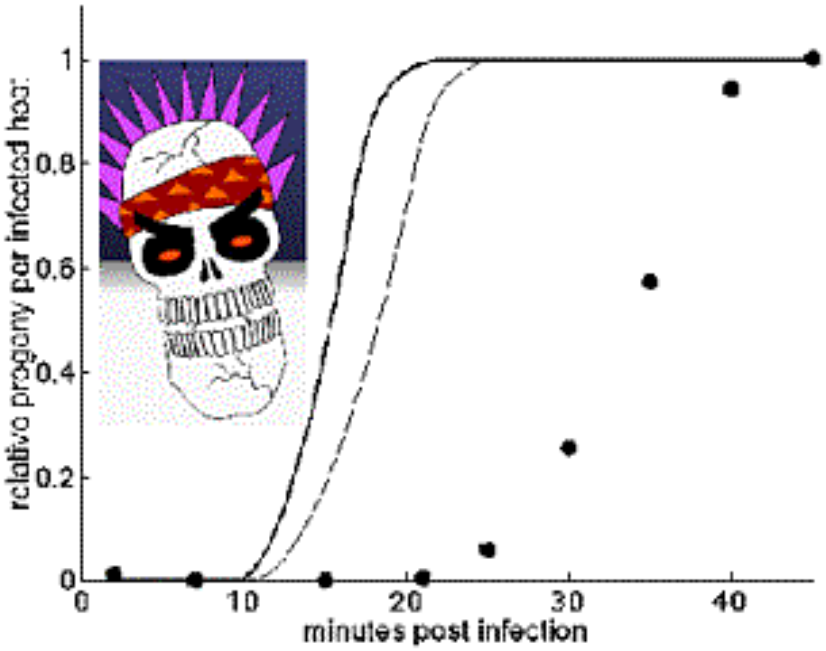
[26] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development



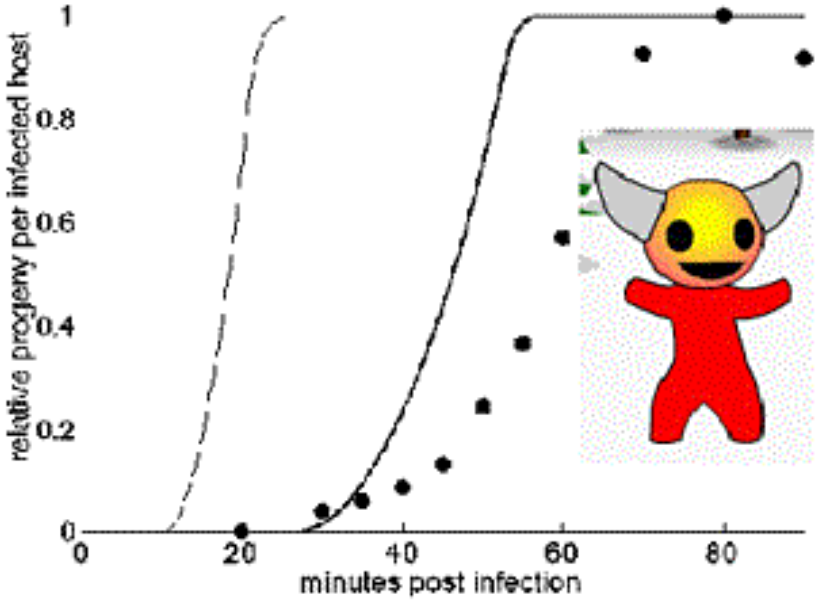
[27] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development



Reality check

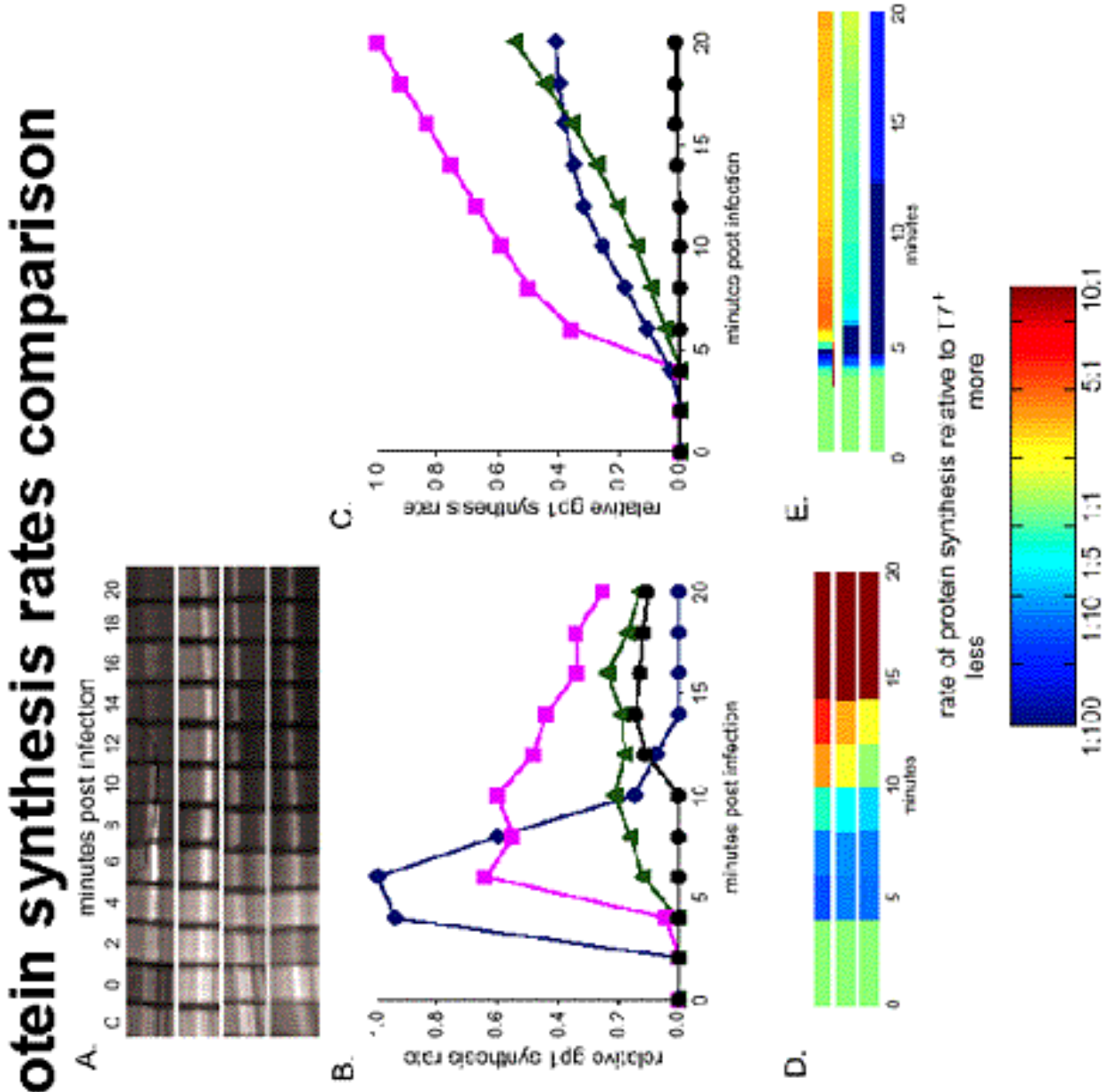


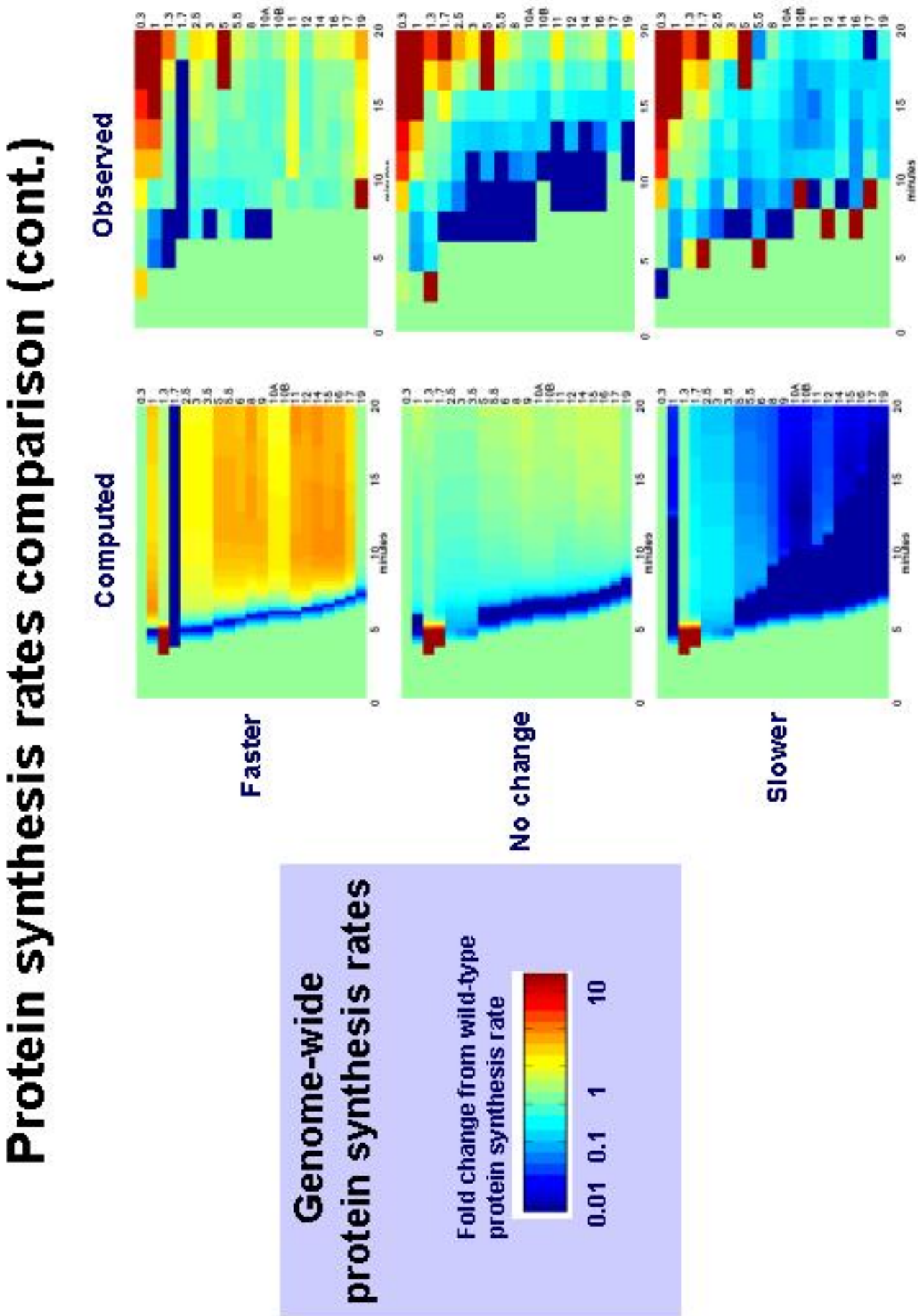
should be faster... blah!



should be slower... cool!

[29] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development



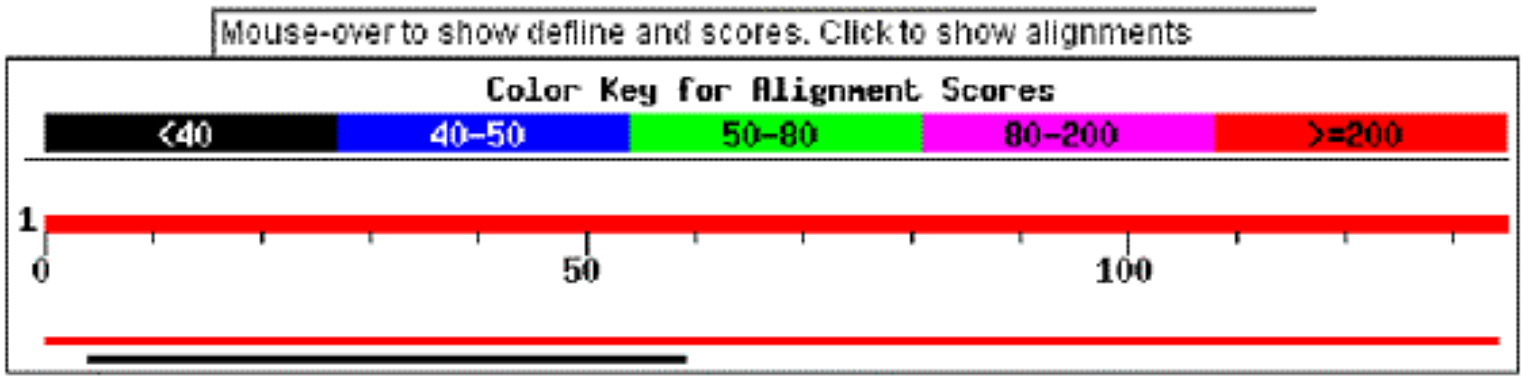

Protein synthesis rates comparison (cont.)
Observed
Computed
Faster
No change
Slower
Genome-wide protein synthesis rates
Fold change from wild-type protein synthesis rate
0.01
0.1
1
10

Acknowledge: The uncharacterized T7 genes

4.7

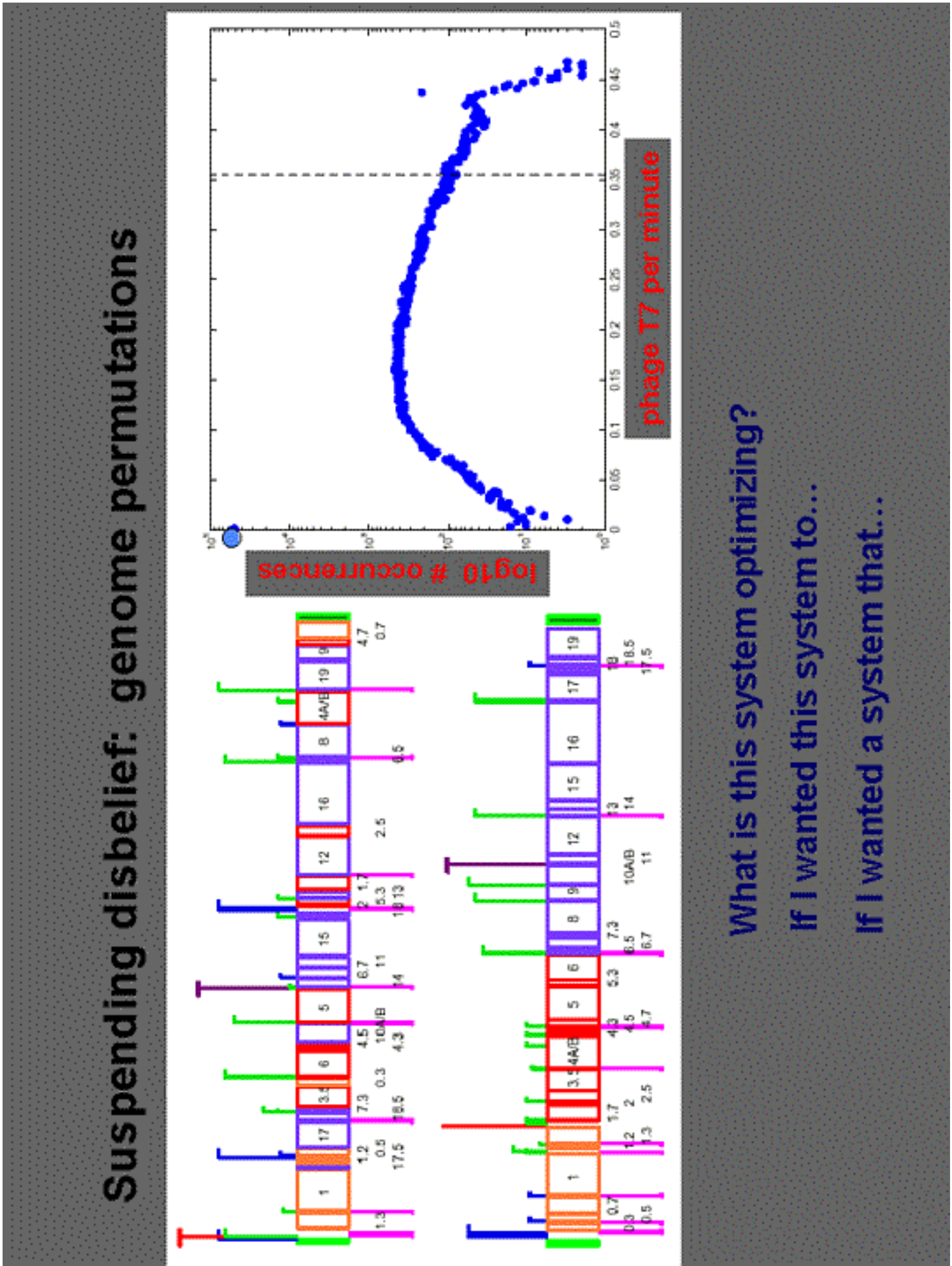
Query=
(135 letters)

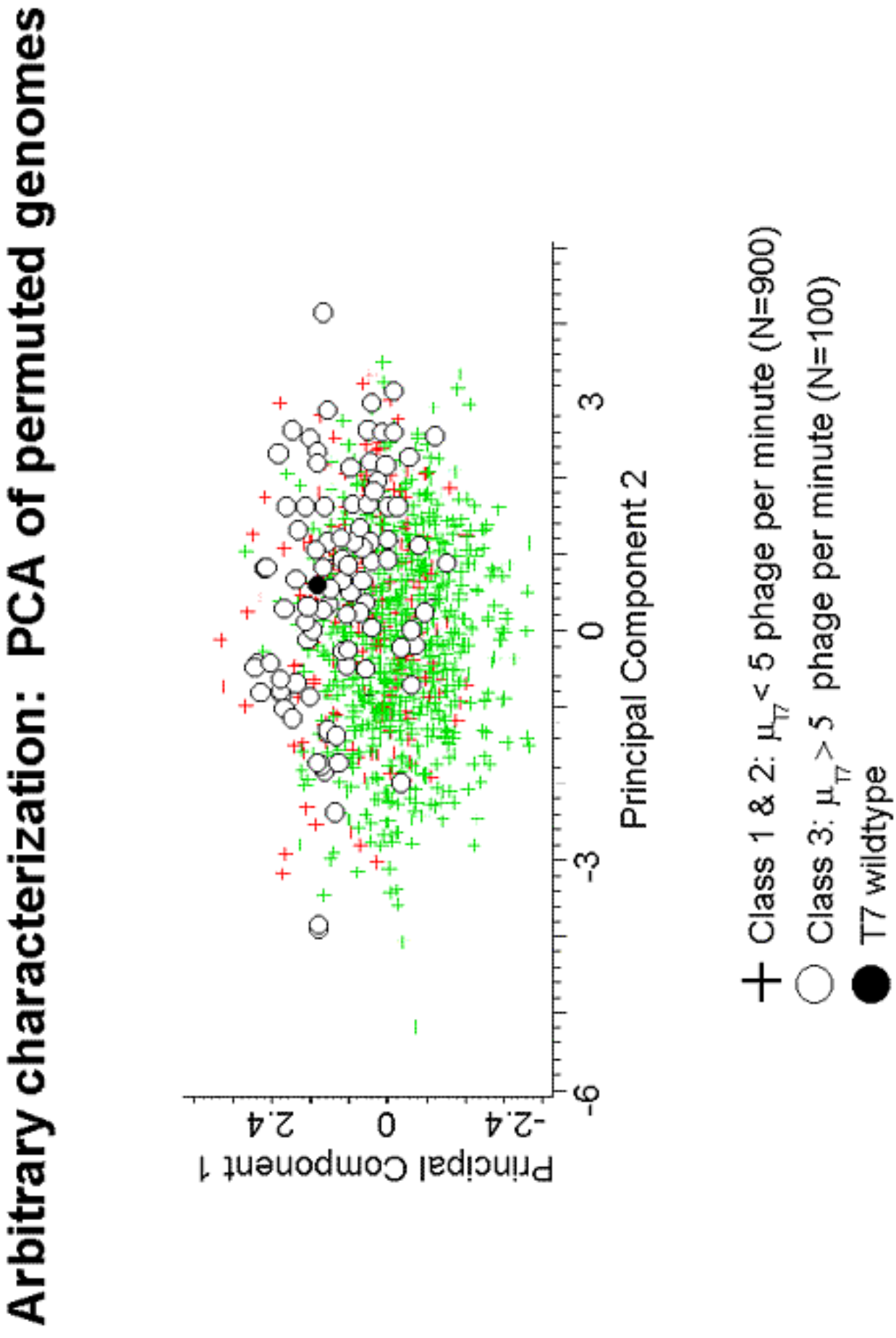
Distribution of 2 Blast Hits on the Query Sequence



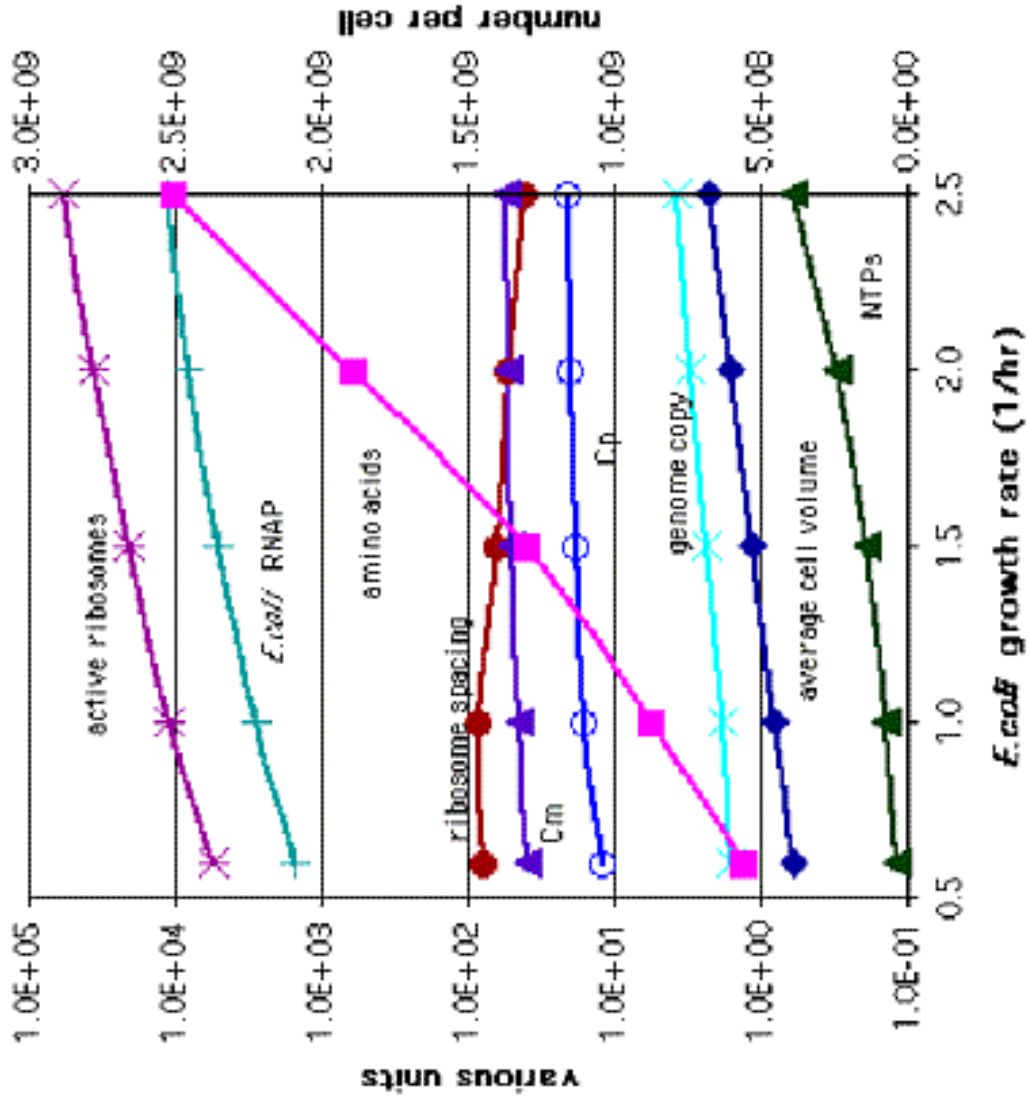
Sequences producing significant alignments:		Score (bits)	E Value
gi|9627453|ref|NP_041981.1|	gene 4.7 [Bacteriophage T7] >gi...	221	4e-57
gi|9759587|dbj|BAB11444.1|	(AB010070) SMC-like protein [Ara...	31	8.3

[32] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development



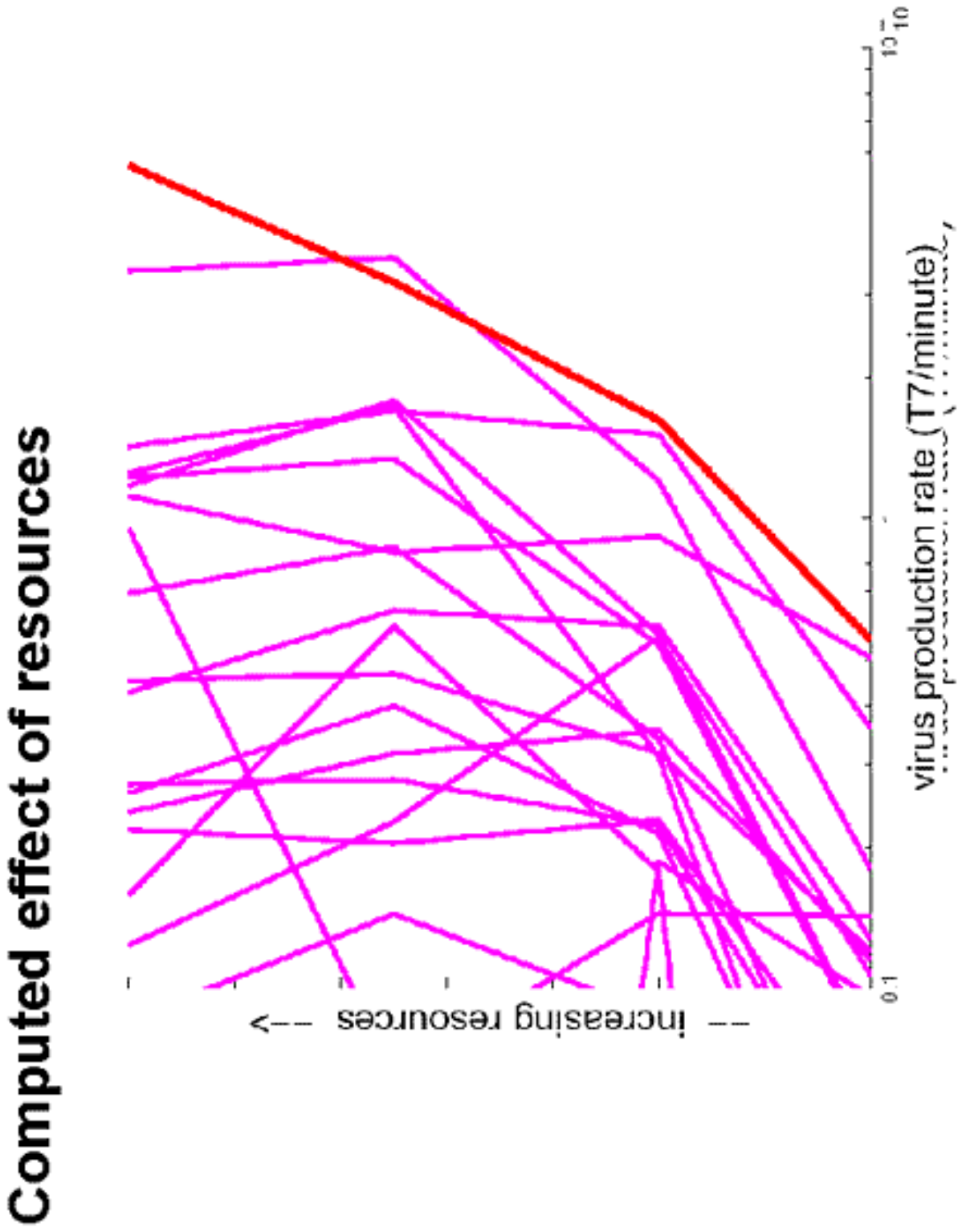

Arbitrary characterization: PCA of permuted genomes
Principal Component 1
-2.4
0
2.4
Principal Component 2
-6
-3
0
3
Class 1 & 2: μ_{T7} < 5 phage per minute (N=900)
Class 3: μ_{T7} > 5 phage per minute (N=100)
T7 wildtype

First-order approximate description of resources

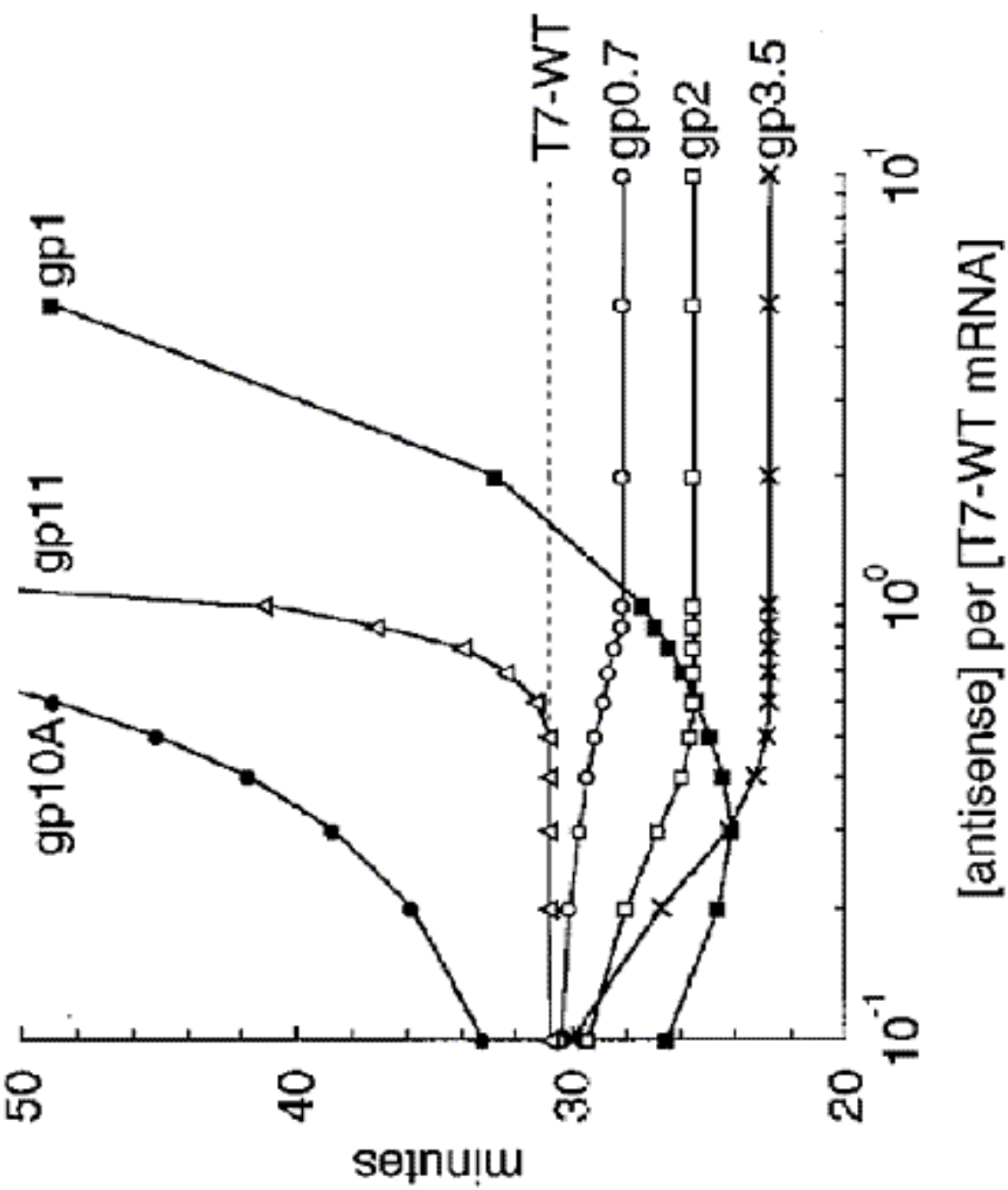


Dennis & Bremer (*E.coli* & *S.typh* 2)

[35] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development

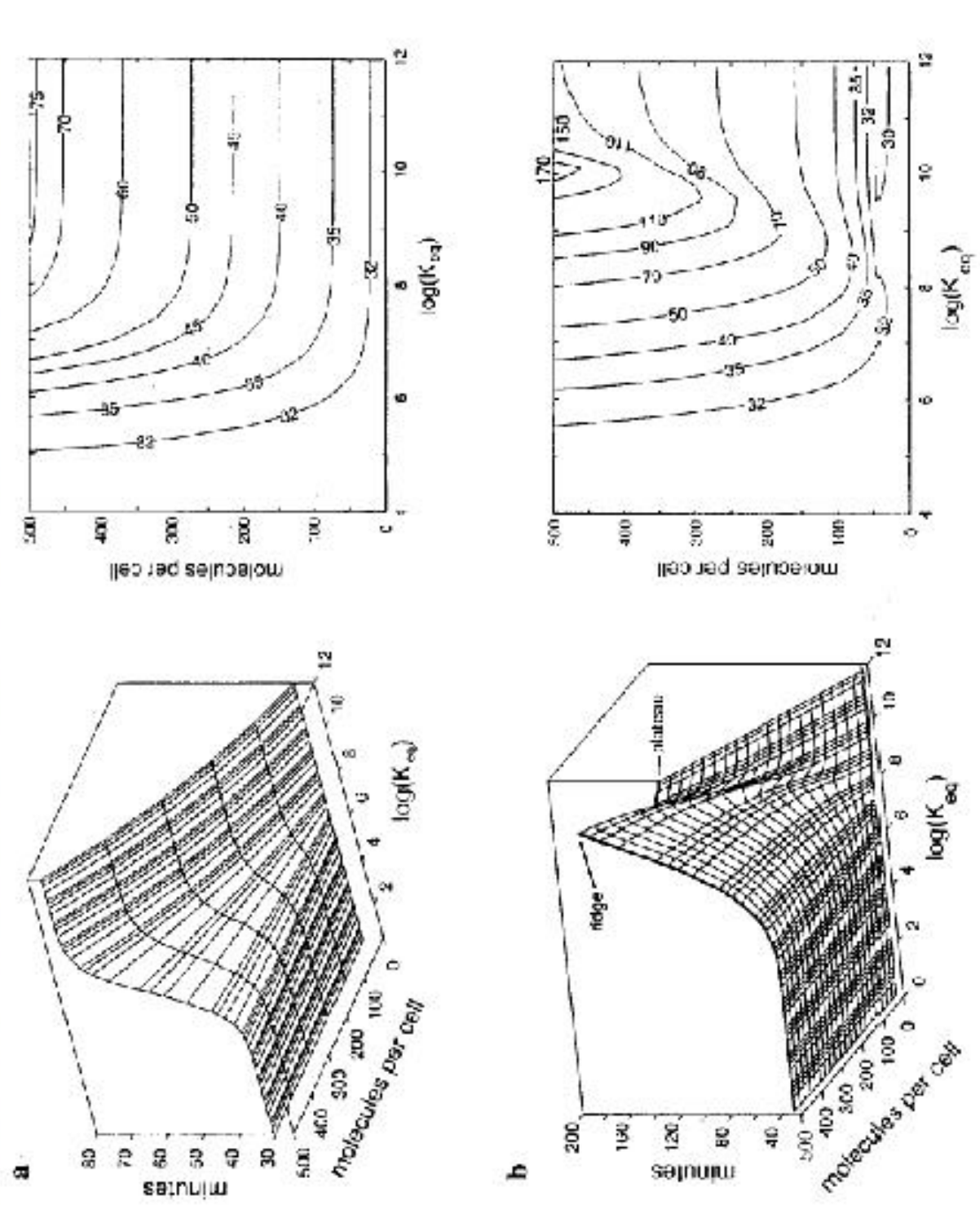


Computing consequences of intervention...

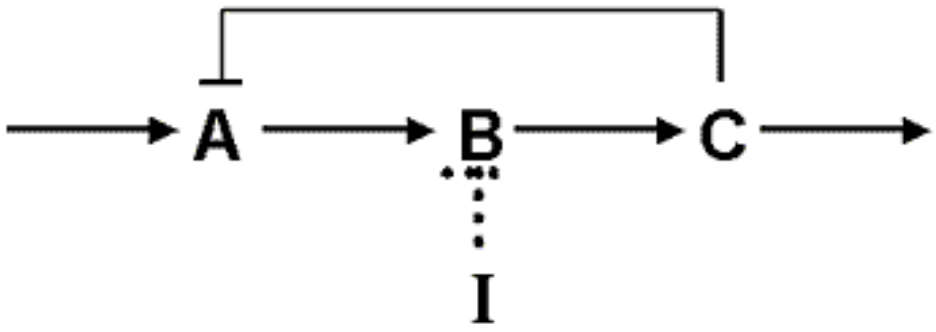


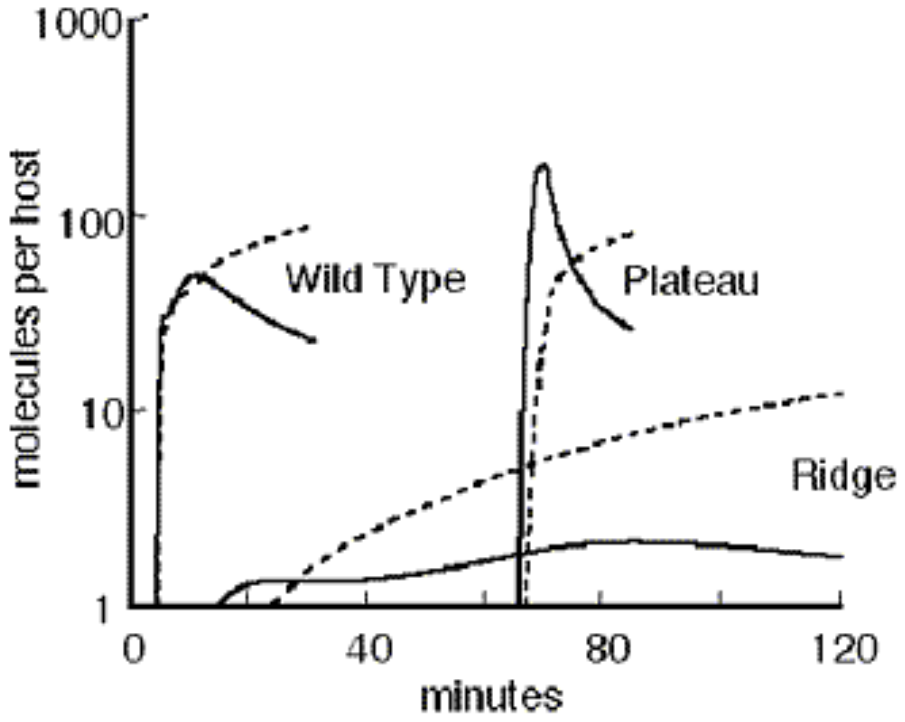
[37] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development

Radically different ways of choosing targets...



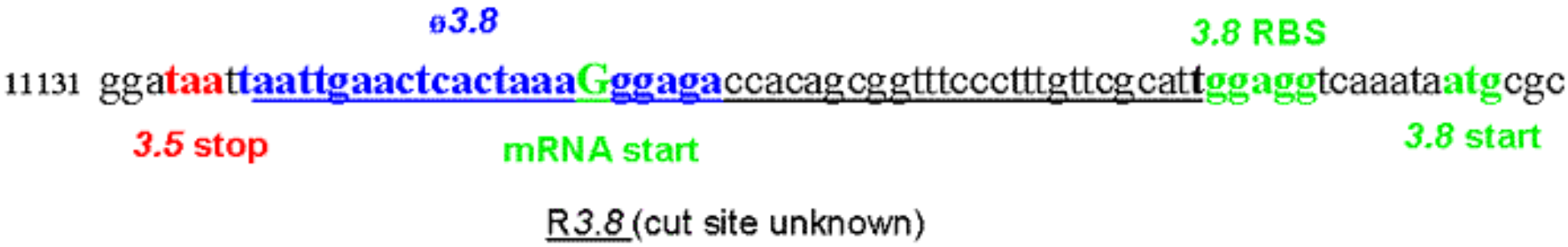
Partially inhibit component of negative feedback loop





Thinking about doing experiments "right"

T7 sequence near ø*3.8*

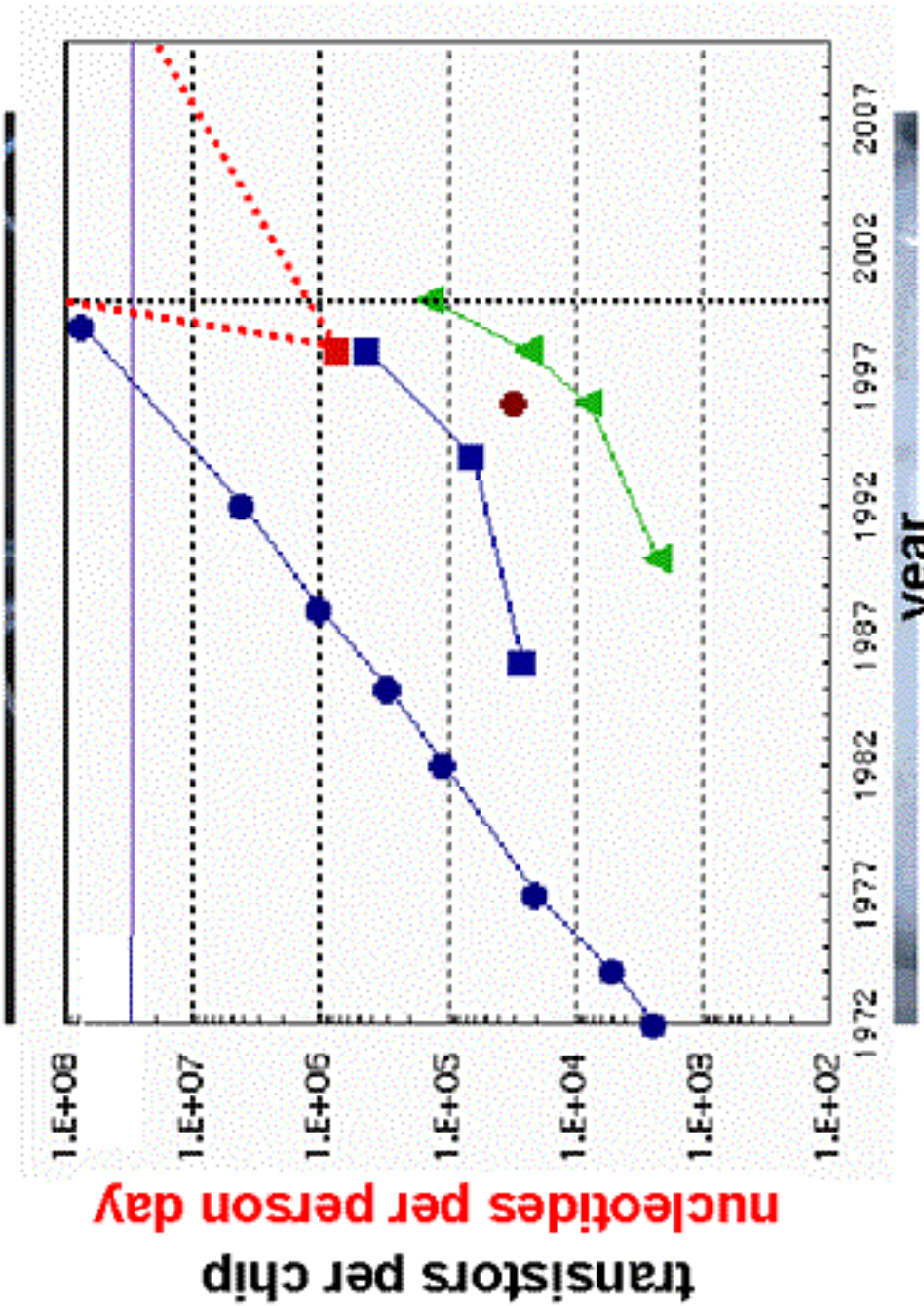


Two approaches

(1) "Refactor" T7

(2) Build a new phage (ϕf) that can be modeled

Aside: technology infrastructure



Rob Carlson (http://www.molsci.org/~rcarlson/exp_biol_tech.html)

Modules defined by "function", what is function?

At least as important as having complete description of components

Better description of environment

How do biological systems encode environmental information (time scales)?

Better description of constraints (from physics)

Biology is a medium for creation

Evolution is the default design algorithm

e.g., constraint of continued existence

Acknowledgements

Eric Lyons
Lynn Thomason

John Yin
Lingchong You
Ian Molineux

Roger Brent
[Sydney Brenner]
Rob Carlson
Alejandro Colman-Lerner
Larry Lok

http://virus.molsci.org/t7/
http://yeast.molsci.org/alpha/

[43] Dr. D. Endy, Molecular Science Inst (ITP 4/18/01) Modeling Bacteriophage T7 Development