CG020 Genomika
Bi7201 Základy genomiky

10. Systémová biologie
10. Systems biology

Kamil Růžička

Funkční genomika a proteomika rostlin,
Mendelovo centrum genomiky a proteomiky rostlin,
Středoevropský technologický institut (CEITEC), Masarykova univerzita, Brno
kamil.ruzicka@ceitec.muni.cz, www.ceitec.muni.cz

Přehled

- What is systems biology
  - System theory
  - Omics
  - Reductionism vs. holism
  - Networks
  - Modular concept
- Regulation of gene expression – example task for systems biology
  - Gene regulation X->Y
  - Transcriptional network of E. coli
  - Negative autoregulatory networks
  - Robustness of negative autoragulatory networks
  - (Positive autoregulatory networks)

What is systems biology

- fashionable catchword?
- a real new (philosophical) concept?
- new discipline in biology?
- just biology?

What is systems biology

- fashionable catchword?
- a real new (philosophical) concept?
- new discipline in biology?
- just biology?

http://www.ceitec.eu/programs/genomics-and-proteomics-of-plant-systems/

Systems theory

- The behavior of a system depends on:
  - Properties of the components of the system
  - The interactions between the components

Systems theory

- The behavior of a system depends on:
  - Properties of the components of the system
  - The interactions between the components

Forget about **reductionism**, think **holistically**.

ὅλος [hol'-os] – greek. all, the whole, entire, complete

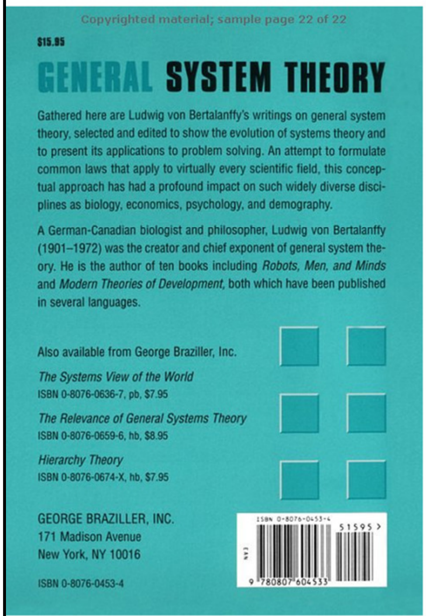
Systems biology

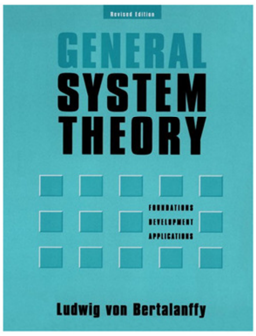
meeting of old and new

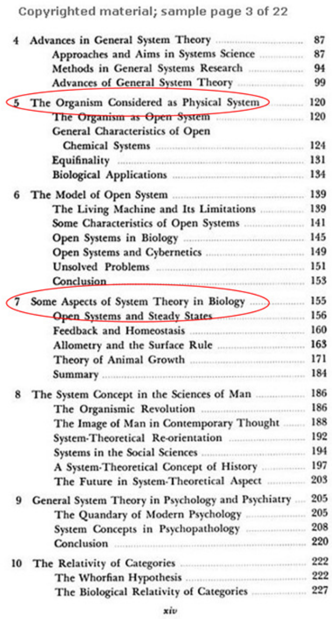
- Systems theory and theoretical biology are old
- Experimental and computational possibilities are new

Ludwig von Bertalanffy

(1901-1972)







4 Advances in General System Theory 87
Approaches and Aims in Systems Science 87
Methods in General Systems Research 94
Advances of General System Theory 99

5 The Organism Considered as Physical System 120
The Organism as Open System 120
General Characteristics of Open Chemical Systems 124
Equifinality 131
Biological Applications 134

6 The Model of Open System 139
The Living Machine and Its Limitations 139
Some Characteristics of Open Systems 141
Open Systems in Biology 145
Open Systems and Cybernetics 149
Unsolved Problems 151
Conclusion 153

7 Some Aspects of System Theory in Biology 155
Open Systems and Steady States 156
Feedback and Homeostasis 160
Allometry and the Surface Rule 163
Theory of Animal Growth 171
Summary 184

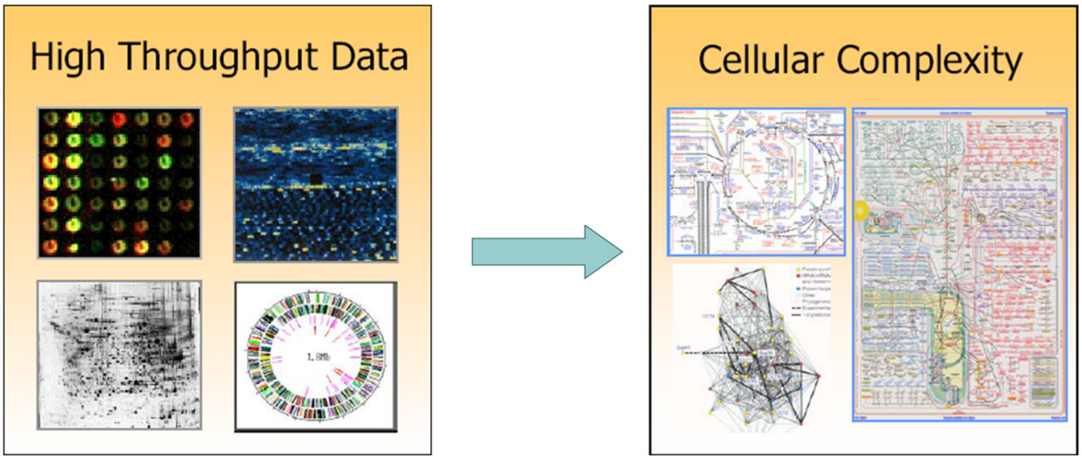
8 The System Concept in the Sciences of Man 186
The Organismic Revolution 186
The Image of Man in Contemporary Thought 188
System-Theoretical Re-orientation 192
Systems in the Social Sciences 194
A System-Theoretical Concept of History 197
The Future in System-Theoretical Aspect 203

9 General System Theory in Psychology and Psychiatry 205
The Quandary of Modern Psychology 205
System Concepts in Psychopathology 208
Conclusion 220

10 The Relativity of Categories 222
The Whorfian Hypothesis 222
The Biological Relativity of Categories 227

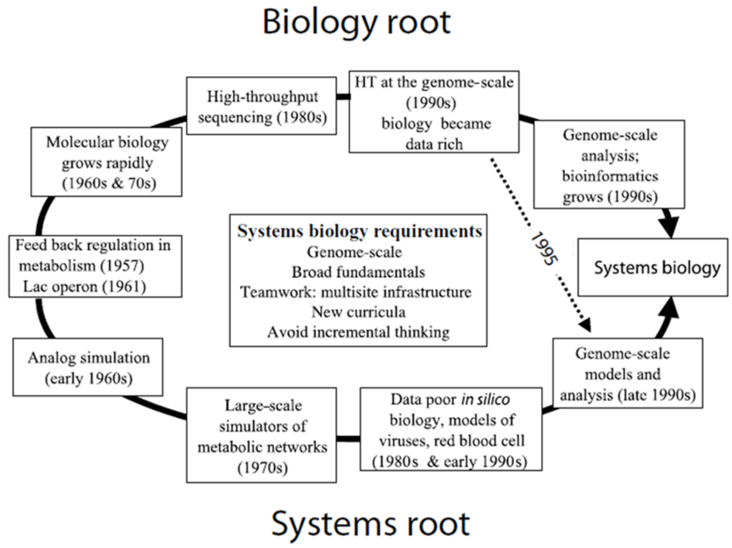
xiv

Omics-revolution shifts paradigm to large systems



- Integrative bioinformatics
- (Network) modeling

Two roots of systems biology



Palsson 2007

Associated disciplines

- Genomics
- Epigenomics
- Transcriptomics
- Translatomics / Proteomics
- Interactomics
- Metabolomics
- Fluxomics
- NeuroElectroDynamics
- Phenomics
- Biomics

Associated disciplines

- Genomics
- Epigenomics
- Transcriptomics
- Translatomics / Proteomics
- Interactomics
- Metabolomics
- Fluxomics
- NeuroElectroDynamics
- Phenomics
- Biomics

Jozef Mravec's term:
multidimensional biology

How I understand systems biology

- **Genetics** – you have one or few RNA processing genes where you show their importance in protoxylem development
- **Functional genomics** – you find in e.g protoxylem expression profiles numerous RNA processing genes and demonstrate which are important for protoxylem developments
- **Systems biology** – based on obtained large scale data you propose model how genes (and/or other components) collectively regulate protoxylem development

How I understand systems biology

- Good biology – you explain why just some genes regulate protoxylem development

(sorry for aphorisms)

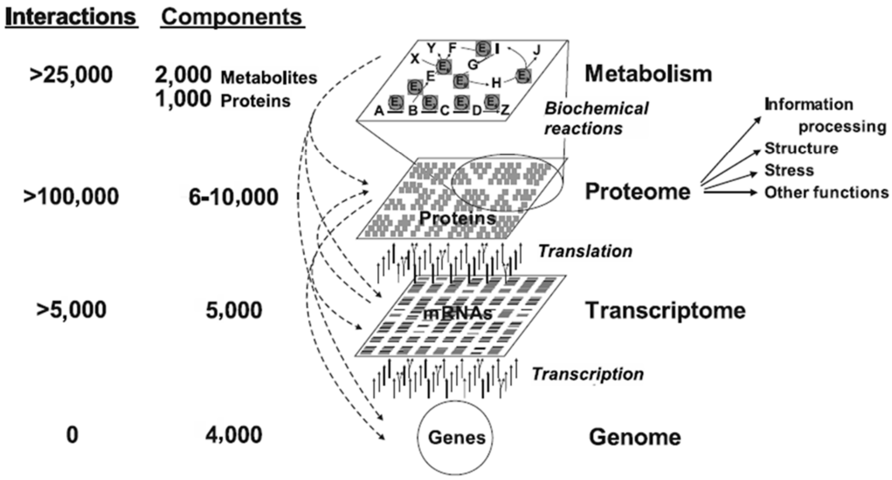
Reconstructed genome-scale networks

Species	#Reactions	#Genes	Reference
Escherichia coli	2077	1260	Feist AM. *et al.* (2007), *Mol. Syst. Biol.*
Saccharomyces cerevisiae	1175	708	Förster J. *et al.* (2003), *Genome Res.*
Bacillus subtilis	1020	844	Oh YK. *et al.* (2007), *J. Biol. Chem.*
Lactobacillus plantarum	643	721	Teusink B. *et al.*, (2006), *J. Bio. Chem.*
Human	3673	1865	Duarte NC. *et al.*, (2007), *PNAS*
Arabidopsis			Arabidopsis Interactome Mapping Consortium (2011), *Science*

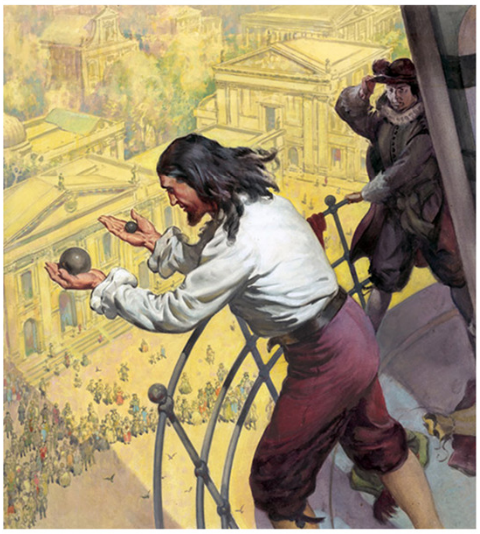
Reconstructed genome-scale networks

Species	#Reactions	#Genes	Reference
Escherichia coli	2077	1260	Feist AM. *et al.* (2007), *Mol. Syst. Biol.*
Saccharomyces cerevisiae	1175	708	Förster J. *et al.* (2003), *Genome Res.*
Bacillus subtilis	1020	844	Oh YK. *et al.* (2007), *J. Biol. Chem.*
Lactobacillus plantarum	643	721	Teusink B. *et al.*, (2006), *J. Bio. Chem.*
Human	3673	1865	Duarte NC. *et al.*, (2007), *PNAS*
Arabidopsis			Arabidopsis Interactome Mapping Consortium (2011), *Science*

Complexity of cellular networks in *E. coli*



Sometimes the <u>things</u> are different than we just <u>think</u>



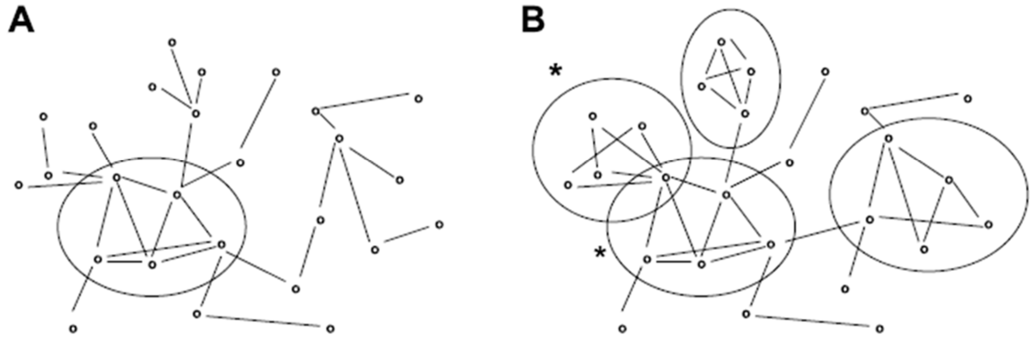
Reconstruction of networks from -omics for systems analysis

- Gene expression networks: based on transcriptional profiling and clustering of genes
- Protein-protein interaction networks (Y2H, TAP etc).
- Metabolic networks: network of interacting metabolites through biochemical reactions.

Reconstruction of networks from -omics for systems analysis

- Gene expression networks: based on transcriptional profiling and clustering of genes
- Protein-protein interaction networks (Y2H, TAP etc).
- Metabolic networks: network of interacting metabolites through biochemical reactions.

How to simplify.
Modularity concept.



Lets e.g. assume that transcription and translation is one module.

E. coli

Binding of a small molecule (a signal) to a transcription factor, causing a change in transcription factor activity	~1 msec
Binding of active transcription factor to its DNA site	~1 sec
Transcription + translation of the gene	~5 min
Timescale for 50% change in concentration of the translated protein (stable proteins)	~1 h (one cell generation)
Generation time	20 min

Description of gene regulation

Transcription factor X regulates gene Y:

$$X \rightarrow Y$$

$$(X \rightarrow transcription \rightarrow translation \rightarrow Y)$$

Description of gene regulation

$$X \rightarrow Y$$

Rate of production: *ß* [units .time^{-1}]

Rate of degradation: α [time^{-1}]

Description of gene regulation

$$X \rightarrow Y$$

Rate of production: *ß* [units .time^{-1}]

Rate of degradation: α [time^{-1}]

$$\alpha = \alpha_{dil} + \alpha_{deg}$$

Description of gene regulation

$$X \rightarrow Y$$

Rate of production: β [units .time^{-1}]

Rate of degradation: α [time^{-1}]

$$\alpha = \alpha_{dil} + \alpha_{deg}$$

cells grow → α_{dil}

protein is degraded → α_{deg}

Description of gene regulation

$$X \rightarrow Y$$

Rate of production: *ß* [units.time^{-1}]

Rate of degradation: α [time^{-1}]

Change of concentration:

$$\frac{dY}{dt} = \beta - \alpha Y$$

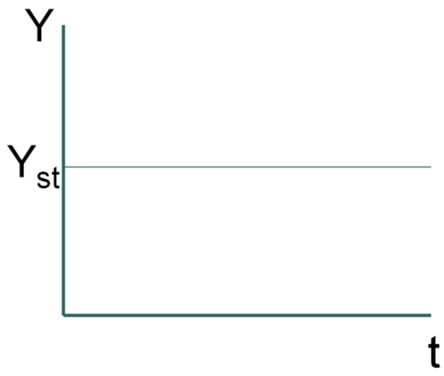
1. Steady state – ustálený stav

$$\frac{dY}{dt} = \beta - \alpha Y$$

$$\frac{dY}{dt} = 0$$

↓

$$Y_{st} = \frac{\beta}{\alpha}$$



2. Production of Y stops

$$\frac{dY}{dt} = \beta - \alpha Y$$

$$\beta = 0$$

$$\downarrow$$

$$Y_t = Y_{st} e^{-\alpha t}$$

The decay is exponential.

2. Production of Y stops:

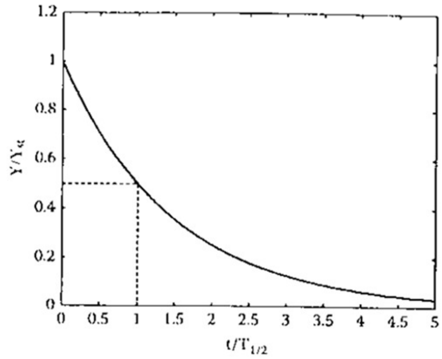
Measure of Y decay – response time ($T_{1/2}$).

$$Y_t = Y_{st} e^{-\alpha t}$$

$$Y_t = \frac{1}{2} Y_{st}$$

↓

$$T_{1/2} = \frac{\log 2}{\alpha}$$



(log => ln [.CZ])

2. Production of Y stops:

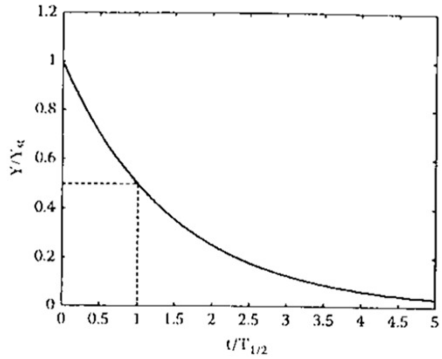
Measure of Y decay – response time ($T_{1/2}$).

$$Y_t = Y_{st}\mathrm{e}^{-\alpha t}$$

$$Y_t = \frac{1}{2}Y_{st}$$

↓

$$T_{1/2} = \frac{\log 2}{\alpha}$$



Large $\alpha \rightarrow$ rapid degradation

(log => ln [.CZ])

Stable proteins

(most of E. coli proteins)

$$T_{1/2} = \frac{\log 2}{\alpha}$$

$$\alpha = \alpha_{\text{dil}} + \alpha_{\text{deg}}$$

$$\alpha \approx \alpha_{\text{dil}}$$

τ – cell generation

$$T_{1/2} = \frac{\log 2}{\alpha_{dil}} = \tau$$

Stable proteins

$$T_{1/2} = \frac{\log 2}{\alpha}$$

$\alpha = \alpha_{\text{dil}} + \alpha_{\text{deg}}$

$\alpha \approx \alpha_{\text{dil}}$

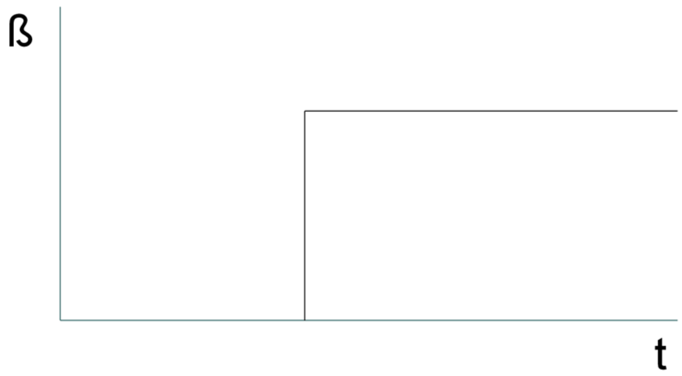
τ – cell generation

$$T_{1/2} = \frac{\log 2}{\alpha_{dil}} = \tau$$

Response time is one generation.

3. Production of Y starts from zero

$$\frac{dY}{dt} = \beta - \alpha Y$$

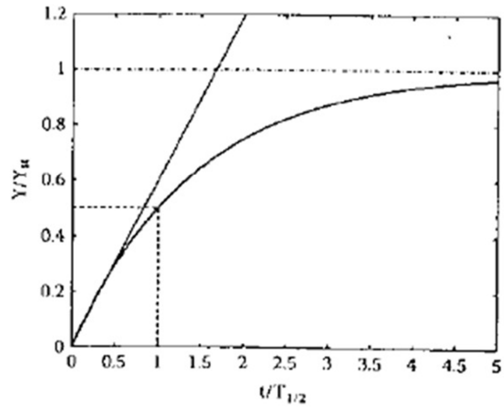


3. Production of Y starts from zero

$$\frac{dY}{dt} = \beta - \alpha Y$$

↓

$$Y_t = \frac{ß}{\alpha}(1 - e^{-\alpha t})$$

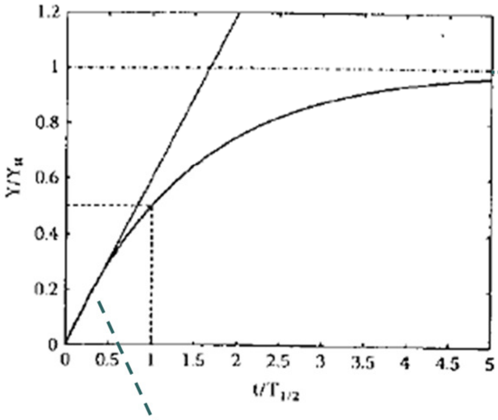


3. Production of Y starts from zero

$$\frac{dY}{dt} = \beta - \alpha Y$$

(magic)

$$Y_t = \frac{ß}{\alpha}(1 - e^{-\alpha t})$$



$$Y_{st} = \frac{ß}{\alpha}$$

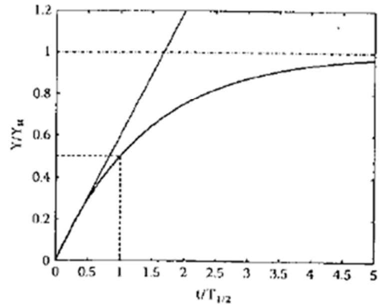
Y grows almost linearly initially

3. Production of Y starts from zero

Response time:

$$Y_t = Y_{st}(1 - e^{-\alpha t})$$

$$Y_t = \frac{1}{2} Y_{st}$$



$$T_{1/2} = \frac{\log 2}{\alpha}$$

The same response time as in case 2.
Response time does not depend on production rate!

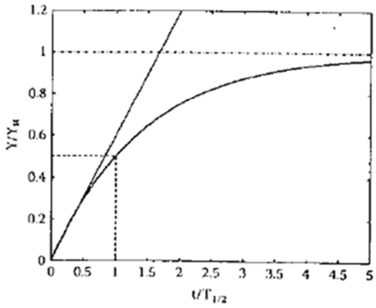
3. Production of Y starts from zero

Response time:

$$Y_t = Y_{st}(1 - e^{-\alpha t})$$

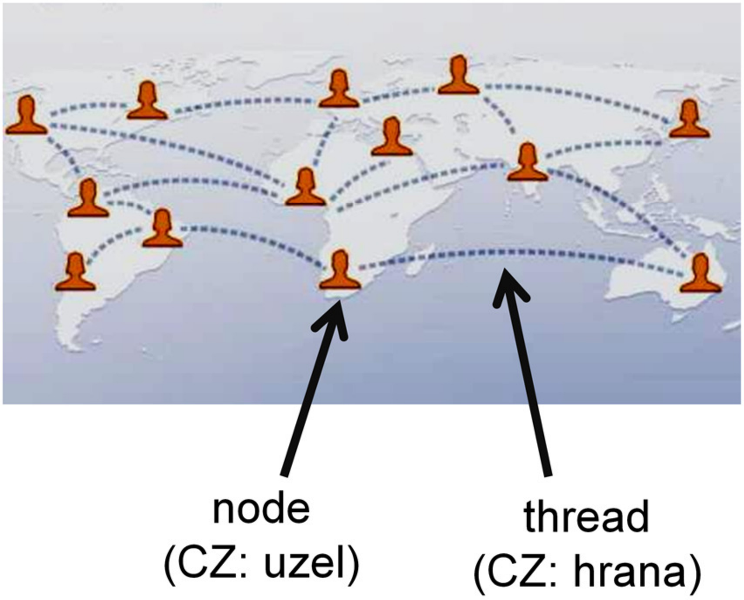
$$Y_t = \frac{1}{2} Y_{st}$$

$$T_{1/2} = \frac{\log 2}{\alpha}$$

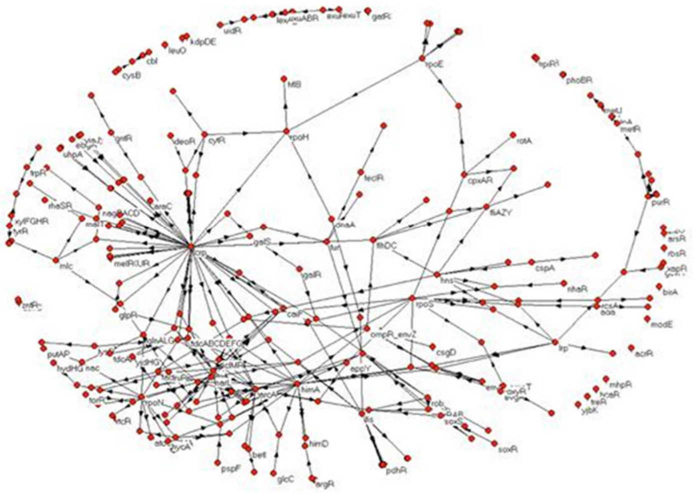


Not many degradation mechanisms in E. coli (energy consuming).
Perhaps in plants?

Networks



Transcriptional network of *E. coli*

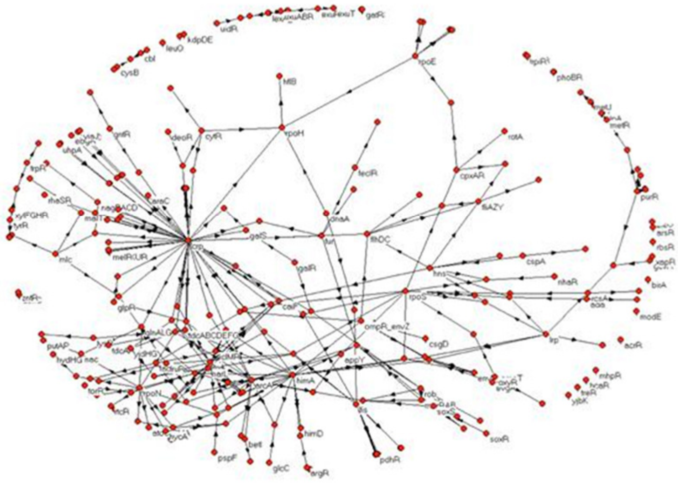


420 nodes, 520 edges
How may self-edges? (CZ: samohrana?)

Likelihood of the self-edge

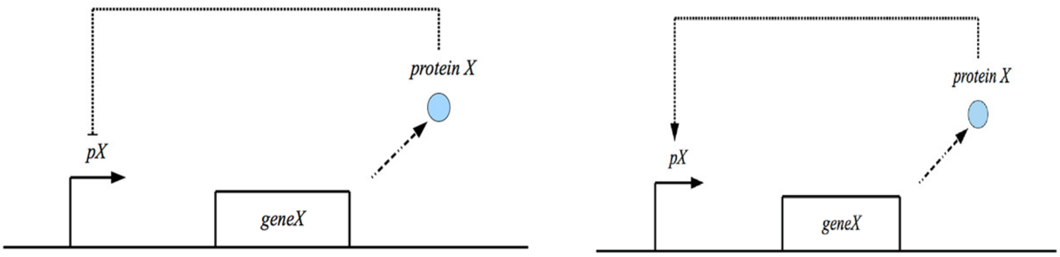
- Assumptions from random network (400 nodes (N), 500 edges (E)). How many self-edges?
- $P_S = E \cdot \frac{1}{N} = 500 \cdot \frac{1}{400} = 1.2 \quad (\pm 1.1)$

Autoregulation is a <u>network motif</u>



420 nodes, 520 edges. <u>40</u> self-edges!

Autoregulation is a network motif



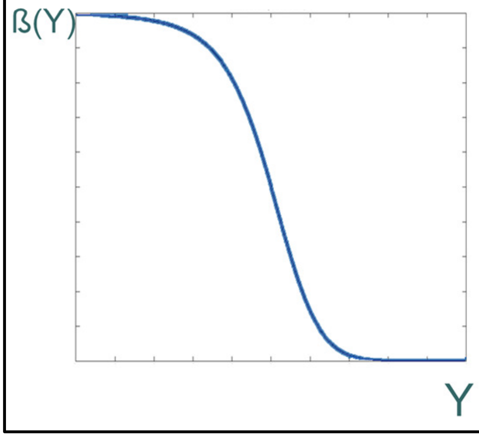
negative regulation positive regulation

E. coli:40 autoregulatory loops: 36 negative, 4 positive

Negative autoregulatory loop is best described by Hill's function

$$\frac{dY}{dt} = \beta - \alpha Y$$

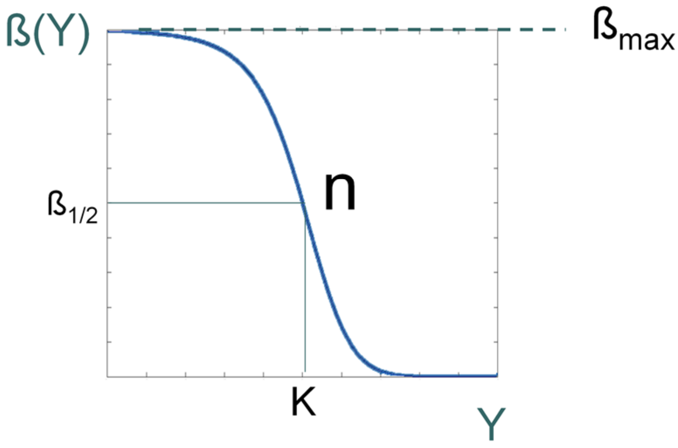
$$\frac{dY}{dt} = \beta(Y) - \alpha Y$$



$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

Negative autoregulatory loops

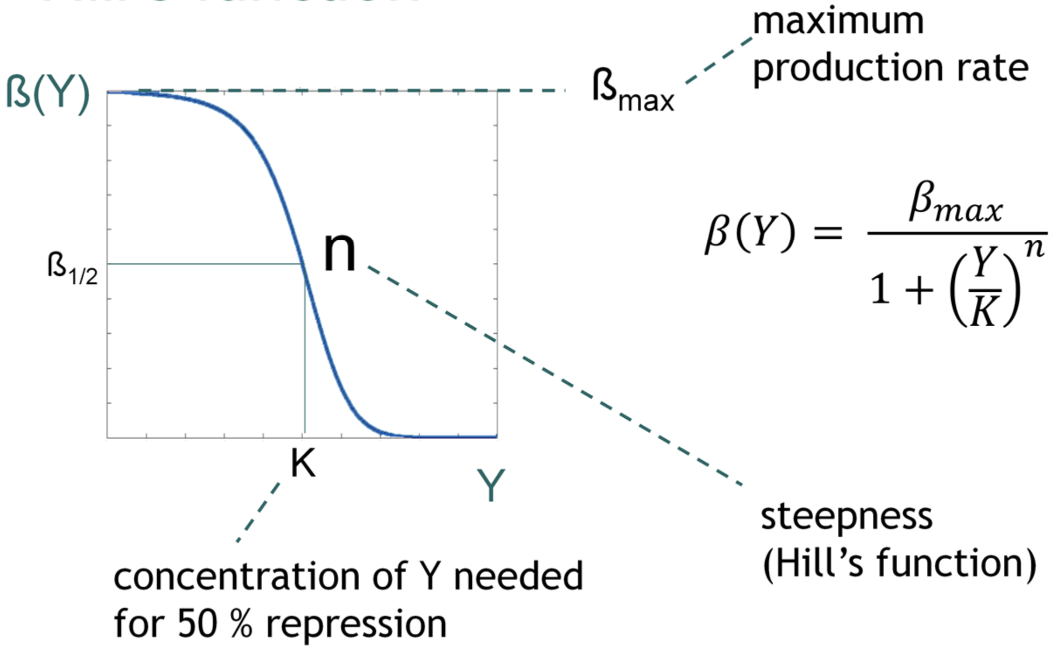
Hill's function



$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

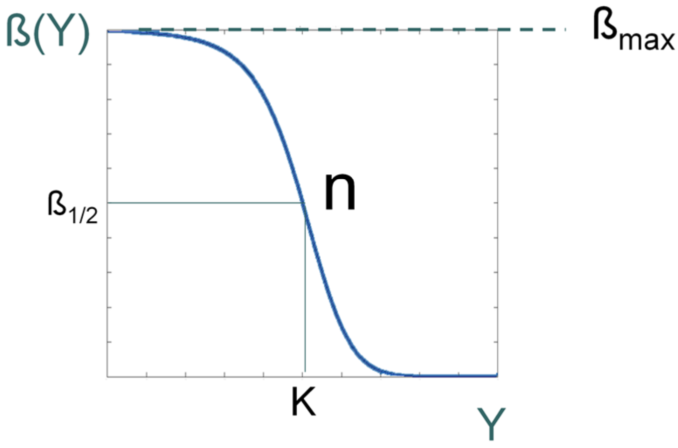
Negative autoregulatory loops

Hill's function



$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

Negative autoregulatory loops
Hill's function



maximum (initial) production rate

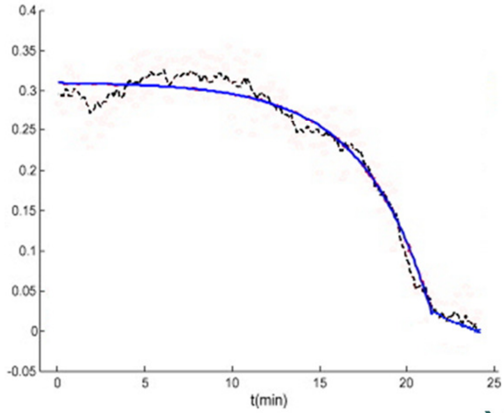
$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

Negative autoregulatory loops

ß synthesis rate – stochastic noise

(stochastický ruch)

ß(Y)

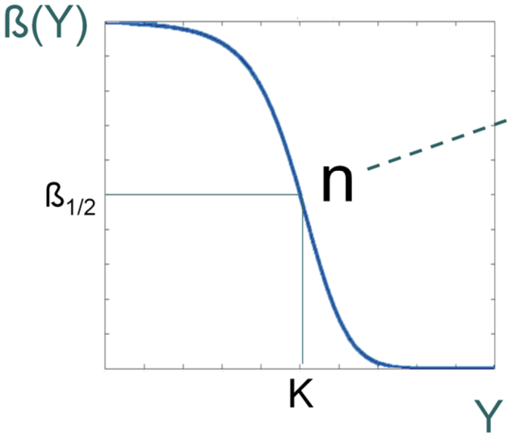


Y

$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

ß may vary by 10 - 30 % (other parameters stable)

Negative autoregulatory loops
Hill's coeficcient

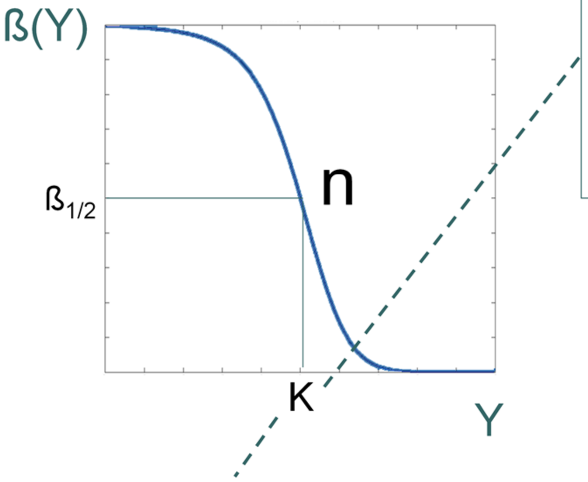


- varies between 1 – 4, the higher the steeper
- important factor: multimerization

$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

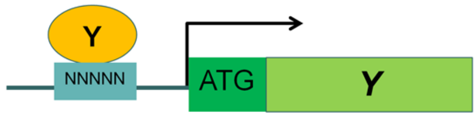
Negative autoregulatory loops

K – repression coefficient



concentration of Y needed for 50 % repression

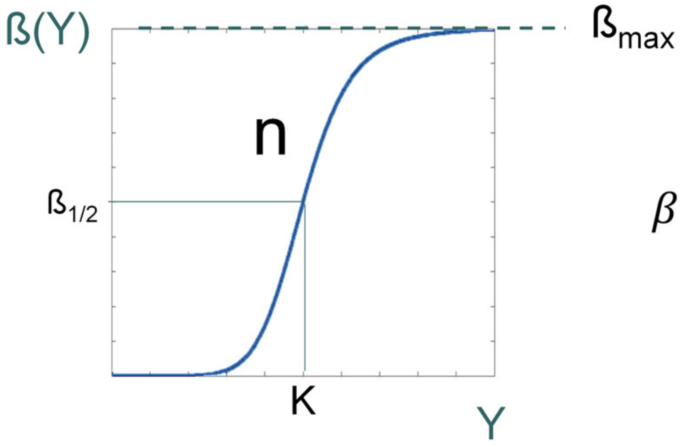
- depends on chemical bonds between Y and its binding sites
- a point mutation can increase K ~10 times



$$\beta(Y) = \frac{\beta_{max}}{1 + \left(\frac{Y}{K}\right)^n}$$

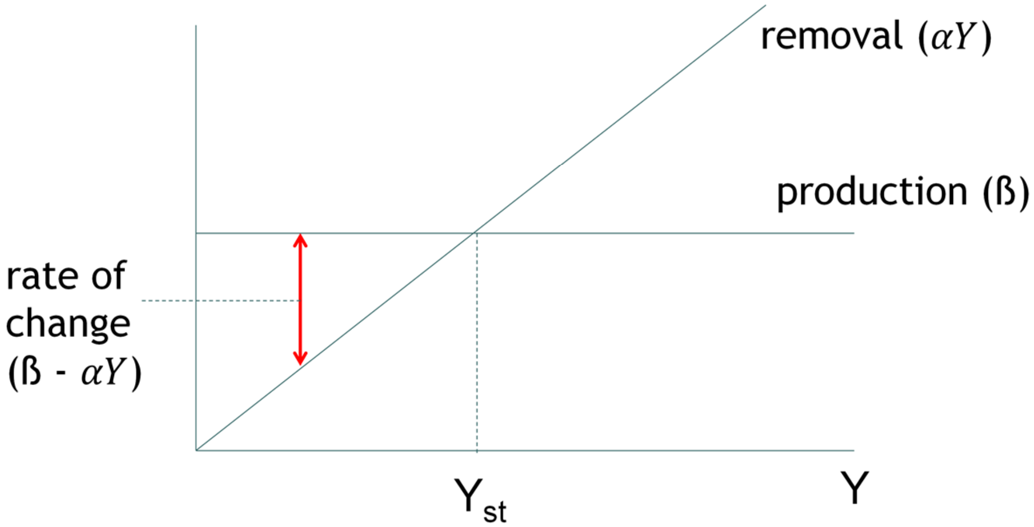
Positive autoregulatory loops

Hill's function

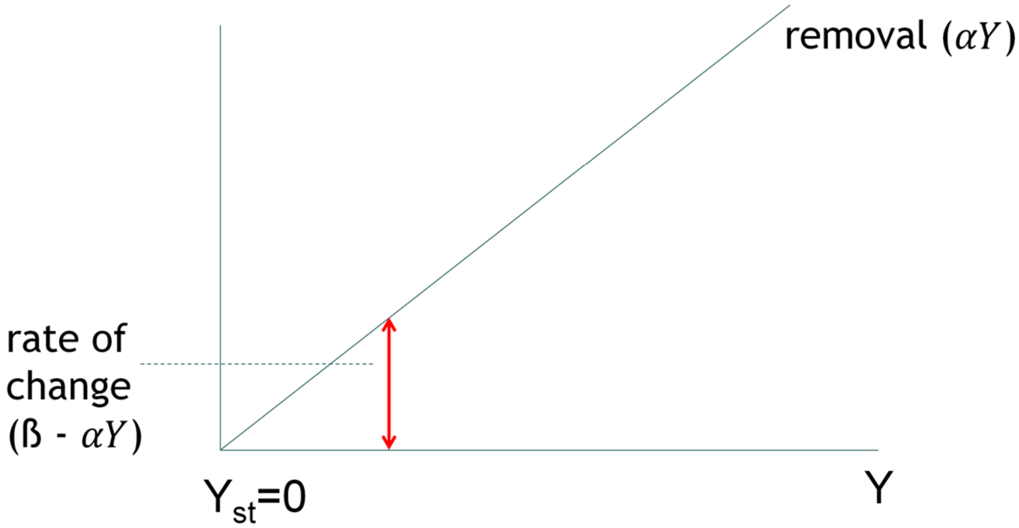


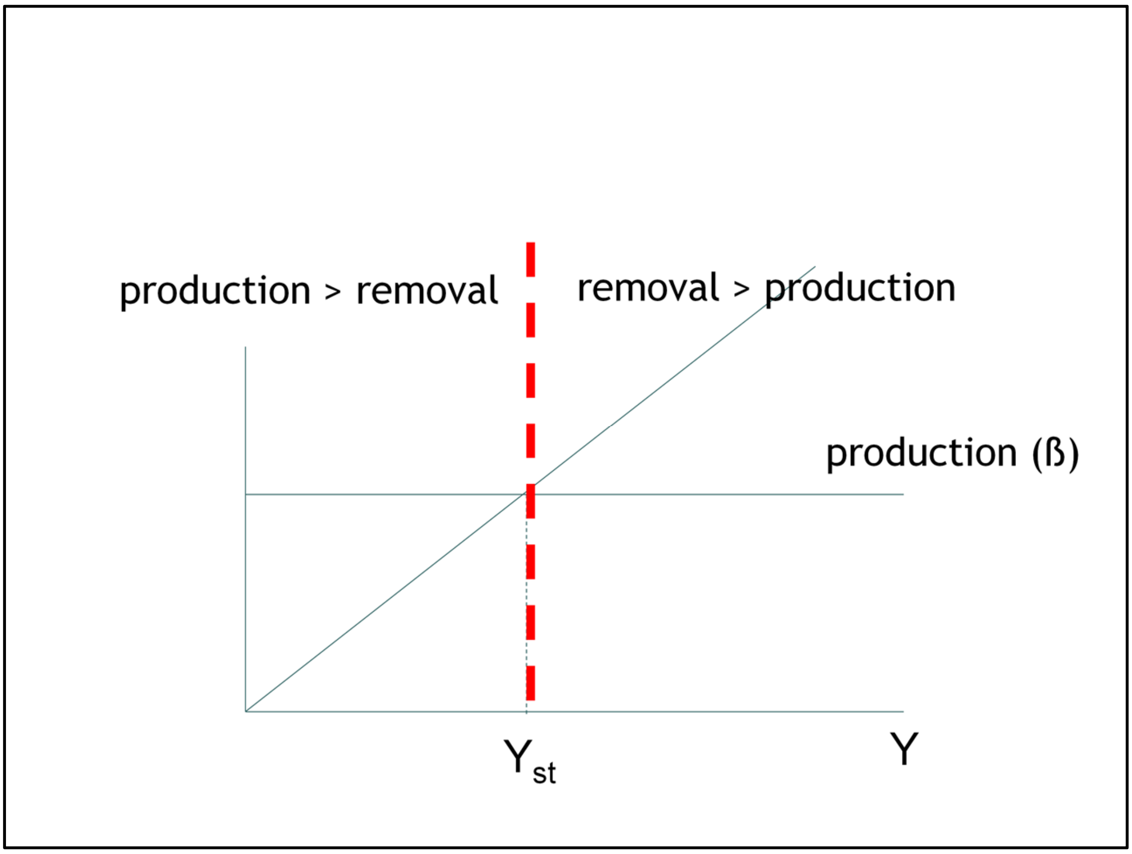
$$\beta(Y) = \frac{\beta_{max}\left(\frac{Y}{K}\right)^n}{1+\left(\frac{Y}{K}\right)^n}$$

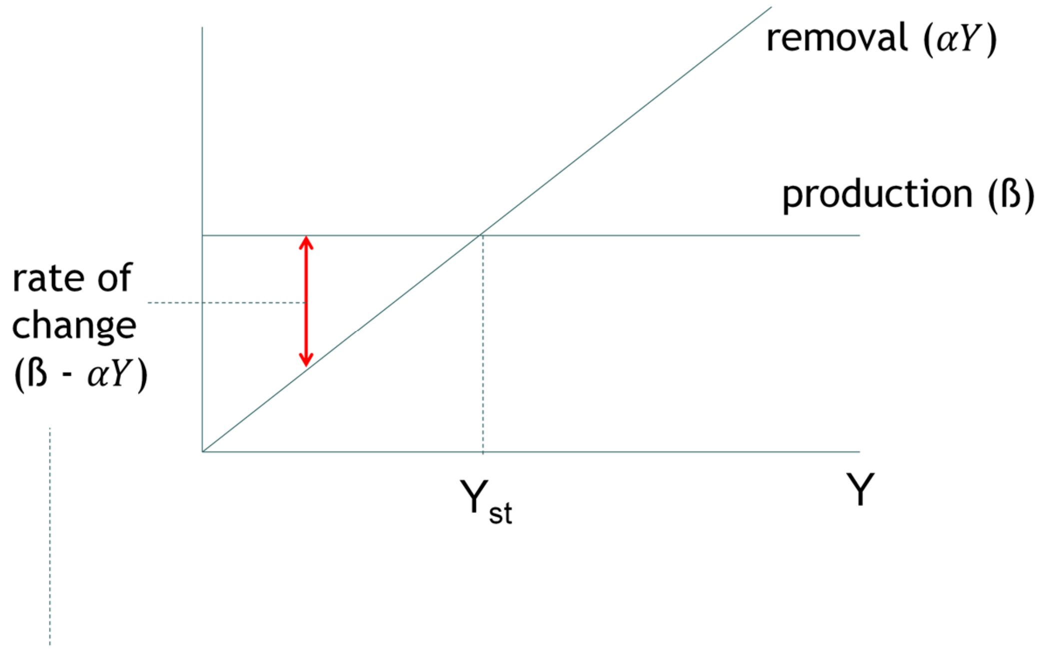
Back to simple regulation



Example: if ß=0 (no production)

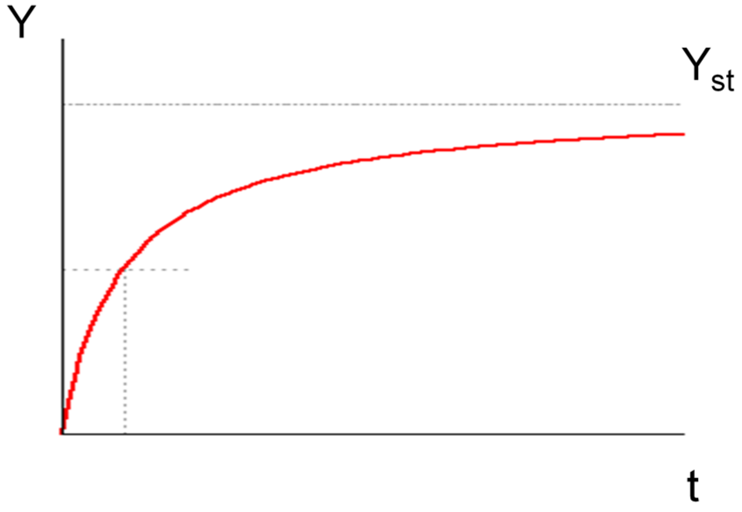



production > removal
removal > production
production (ß)
Y_{st}
Y

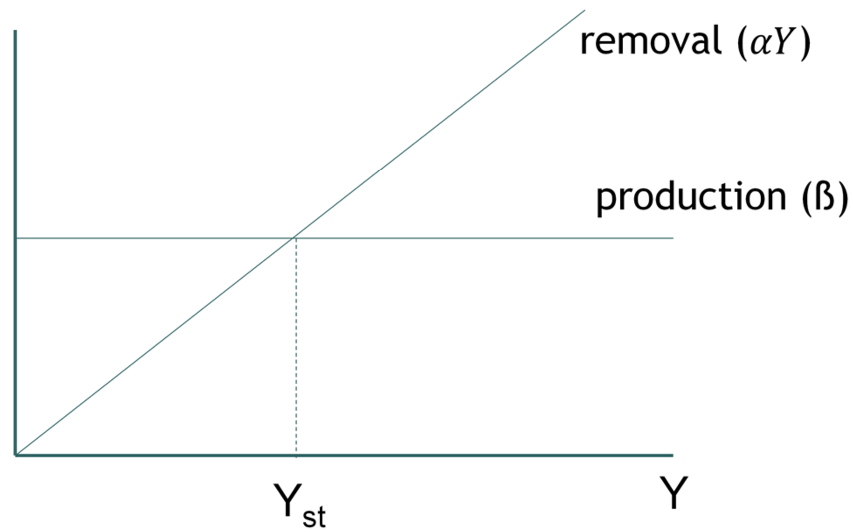


the longer the distance, the faster the change

Therefore more difficult to come to Y_{st} with time



Autoregulation vs. simple production



Comparison

Lets assume that these values are the same:

1. Y_{st}
2. α

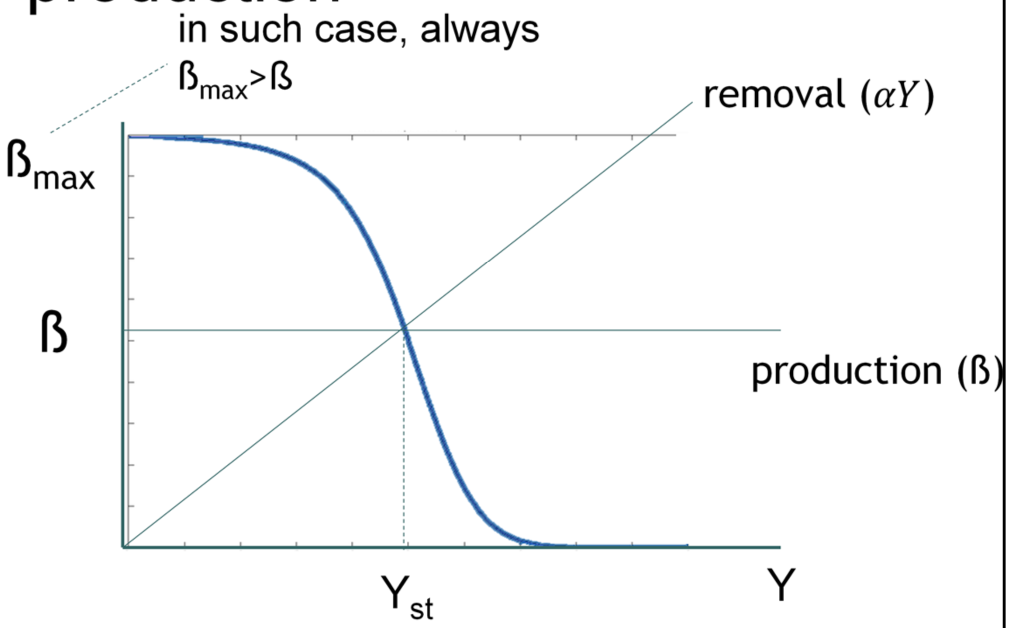
Comparison

Lets assume that these values are the same:

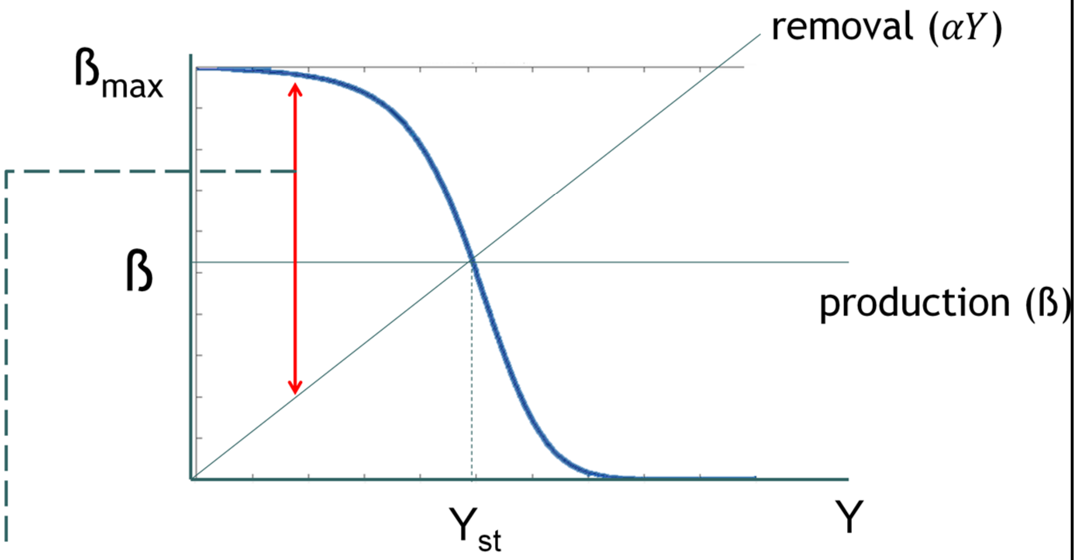
1. Y_{st}
2. α

Lets put it in one graph.

Autoregulation vs. simple production

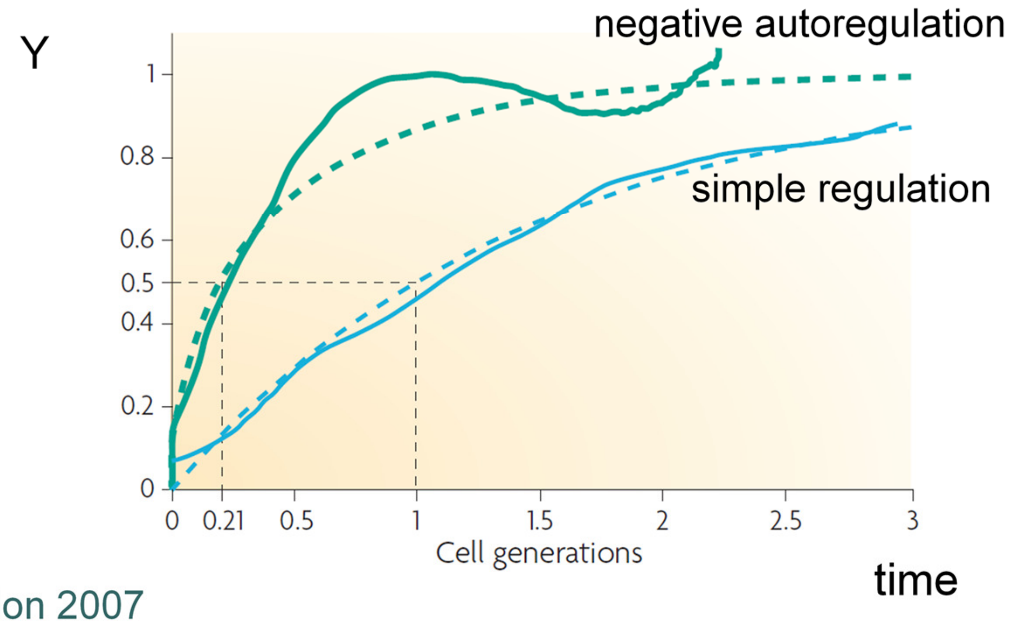


Autoregulation vs. simple production



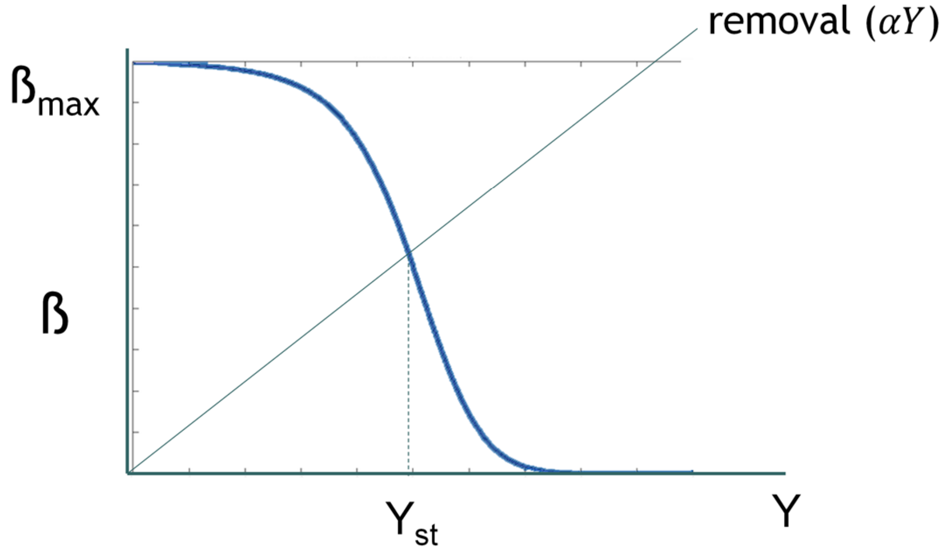
the distance always more far => the reactions are faster

Response time was confirmed indeed faster (~5 times)

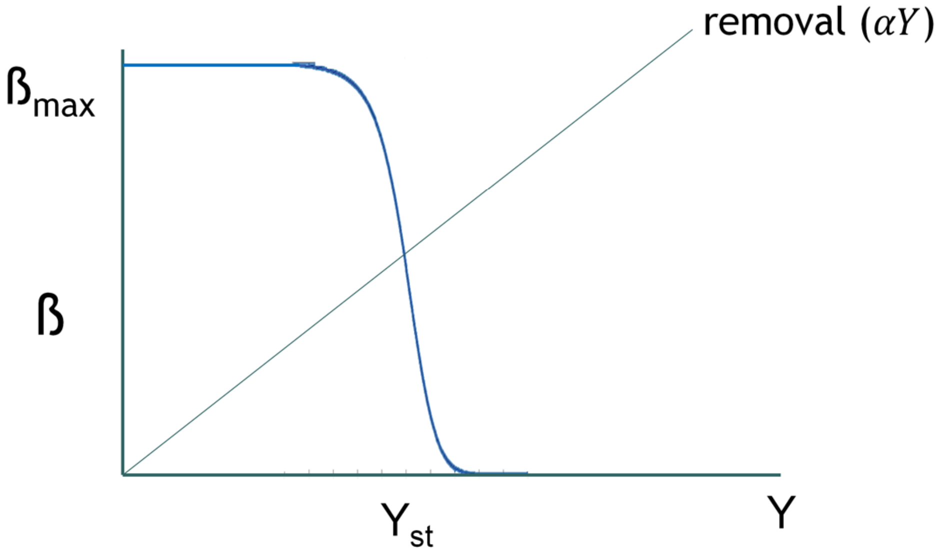


Alon 2007

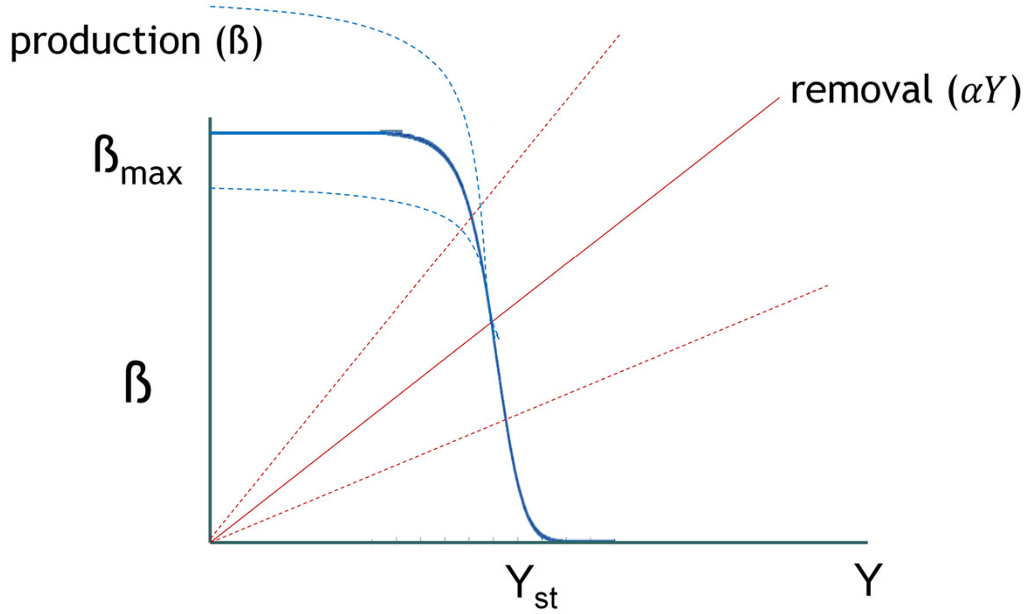
Cases of sharp curve



Cases of sharp curve

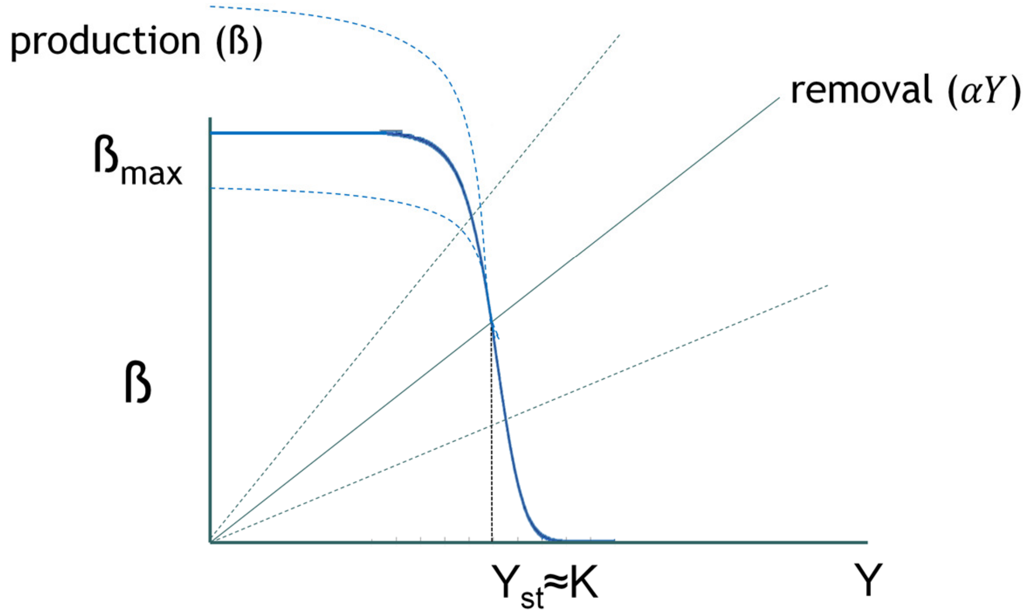


Cases of sharp curve



Fluctuations in synthesis or removal don't change much.

Cases of sharp curve

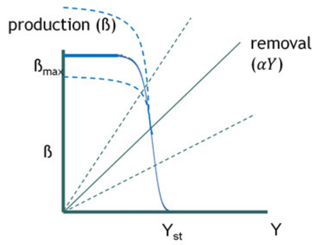


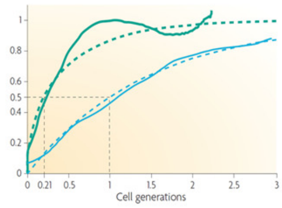
Y_{st} depends only on K – on protein-DNA binding properties.

Conclusions

Negative autoregulation

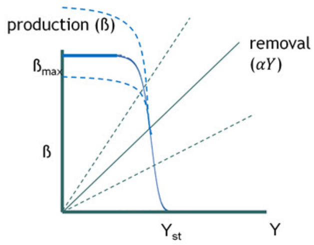
- speeds up response time
- is robust (for α, ß) => basically on/off
- bypasses stochastic noise

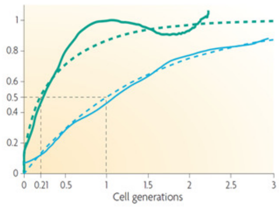




Conclusions

The model explains why negative autoregulation is a common network motif in E. coli.



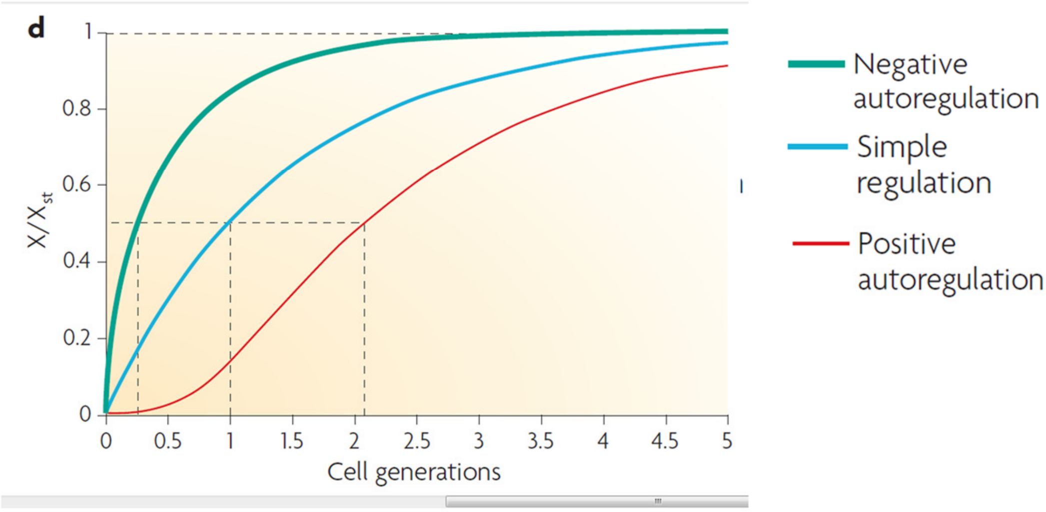


Conclusions

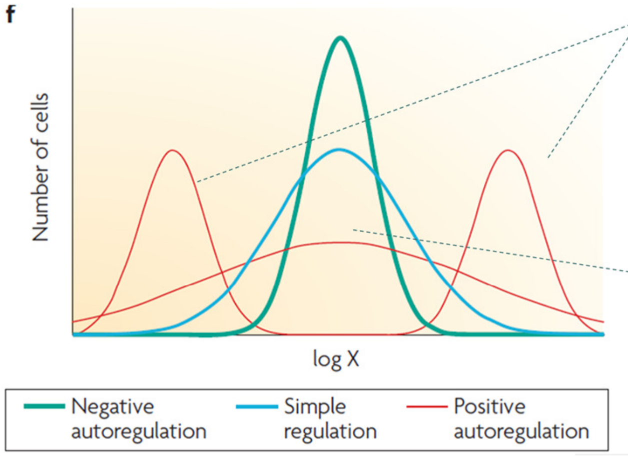
The model explains why negative autoregulation is network motif.

We will not avoid mathematics in biology.

Positive autoregulation leads to slower response



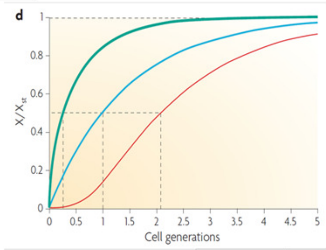
Positive autoregulation leads to higher variation



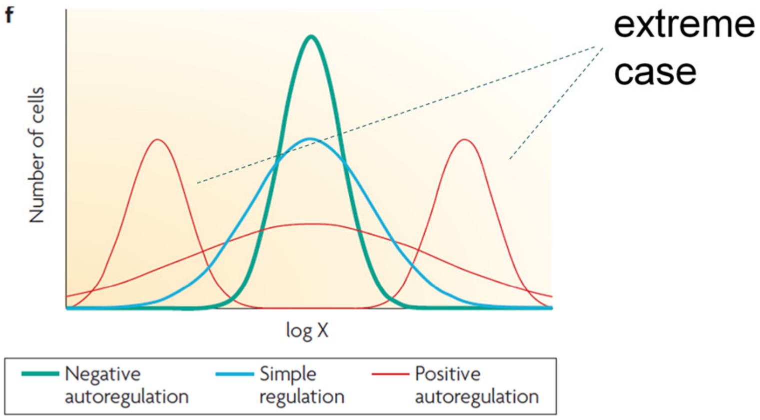
extreme case

less extreme case

=> increasing cell-cell variability



Positive autoregulation leads to higher variation



Strong variation:

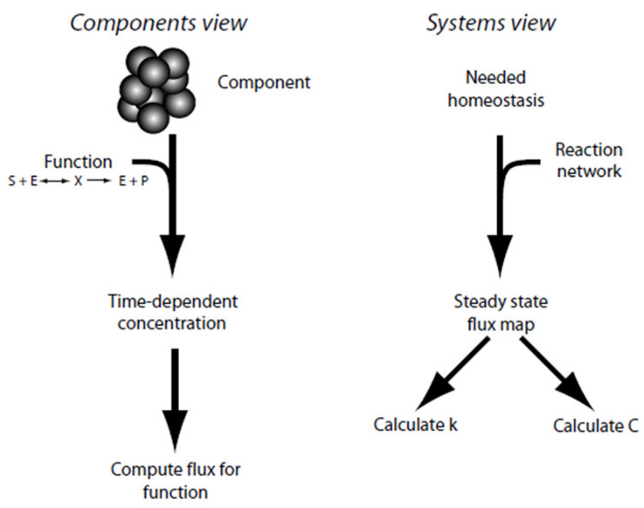
- => differentiation of cells into 2 populations (development)
- => memory for maintaining gene expression (development)
- helps with maintaining mixed phenotype for better response to changing environment

Literature

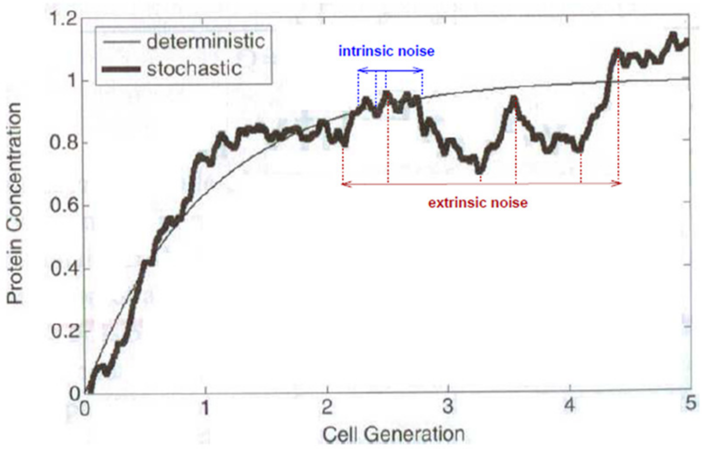
- Source literature
  - http://www.youtube.com/watch?v=Z_BHVFP0Lk and further – excellent talks about systems biology from Uri Alon (Weizman Institute)
  - Rosenfeld N, Negative autoregulation speeds the response times of transcription networks. J Mol Biol. 2002 Nov 8;323(5):785-93. – experimental testing of the data
  - Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007 Jun;8(6):450-61. Review about the same.
  - Alon, U. (2006). An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman and Hall/CRC).
  - Palsson, B.Ø. (2011). Systems Biology: Simulation of Dynamic Network States (Cambridge University Press). Most common textbook about systems biology,

- For enthusiasts
  - Zimmer (2009). Microcosm- E Coli & the New Science of Life (Vintage) (popular scientific book about E. coli as model organism and what you probably didn't know)
  - Albert-László Barabási (2005) V pavučině sítí. (Paseka) (znamenitá kniha o matematice sítí, dynamicky se rozvíjejícím oboru od předního světového vědce)
  - PA052 Úvod do systémové biologie, Přednášky. Fakulta Informatiky MU
  - http://sybila.fi.muni.cz/cz/index - obor na fakultě informatiky.

Reductionism vs. holism

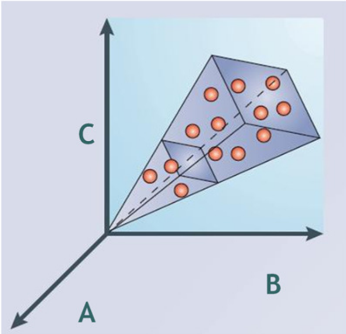


Stochastic noise (stochastický ruch)

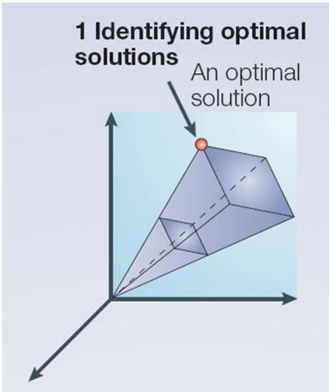


- interní ruch – transkripce, translace, post-transkripční jevy, pozice DNA v chromozómu

Flux balance analysis (FBA)



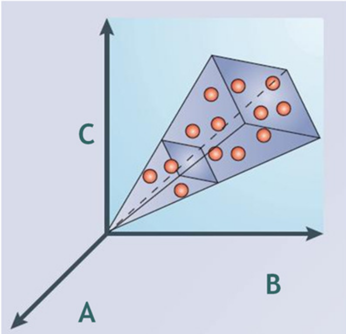
Constraints set bounds on solution space, but where in this space does the "real" solution lie?



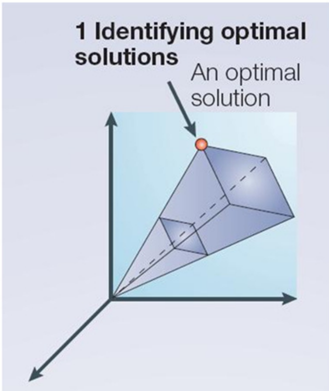
FBA: optimize for that flux distribution that maximizes an objective function (e.g. biomass flux) – subject to S.v=0 and $\alpha_j \le v_j \le \beta_j$

Thus, it is assumed that organisms are evolved for maximal growth -> efficiency!

Flux balance analysis (FBA)



Constraints set bounds on solution space, but where in this space does the "real" solution lie?

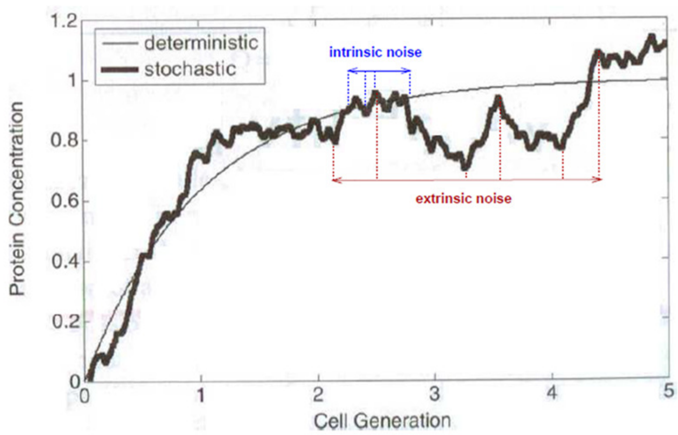


FBA: optimize for that flux distribution that maximizes an objective function (e.g. biomass flux) – subject to S.v=0 and $\alpha_j \le v_j \le \beta_j$

Thus, it is assumed that organisms are evolved for maximal growth -> efficiency!

PA052 Úvod do systémové biologie

Metagenomics



- interní ruch – transkripce, translace, post-transkripční jevy, pozice DNA v chromozómu
- externí ruch – fluktuace koncentrací regulačních faktorů