On Artificial Regulatory Networks

January 2006

Wolfgang Banzhaf
Memorial University of Newfoundland
St. John's, Canada

A Regulatory Toy Model

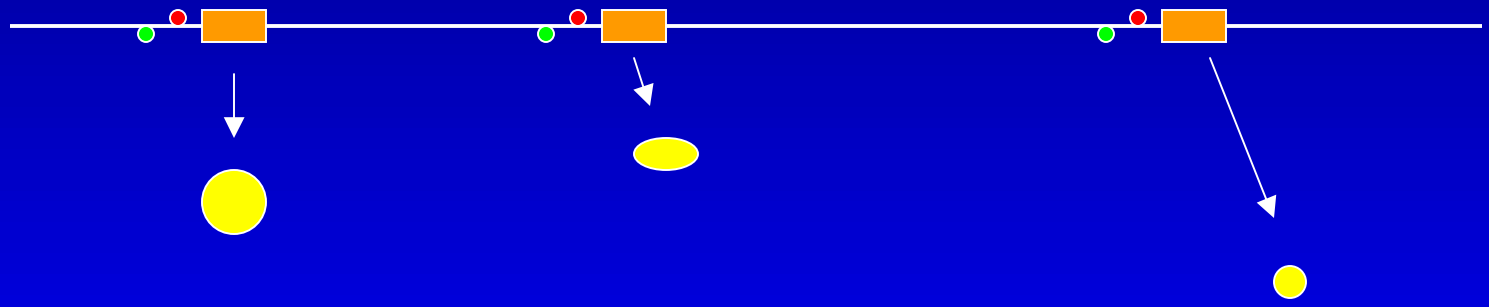
- The Model
- Network structure
- Network dynamics
- Heterochronic control
- Evolution
- Stability, communication

Joint work with Dwight Kuo and Andre Leier

The Model: Genome

A simple model of interaction between genes and proteins

Linear Genome ⟶ Reading direction

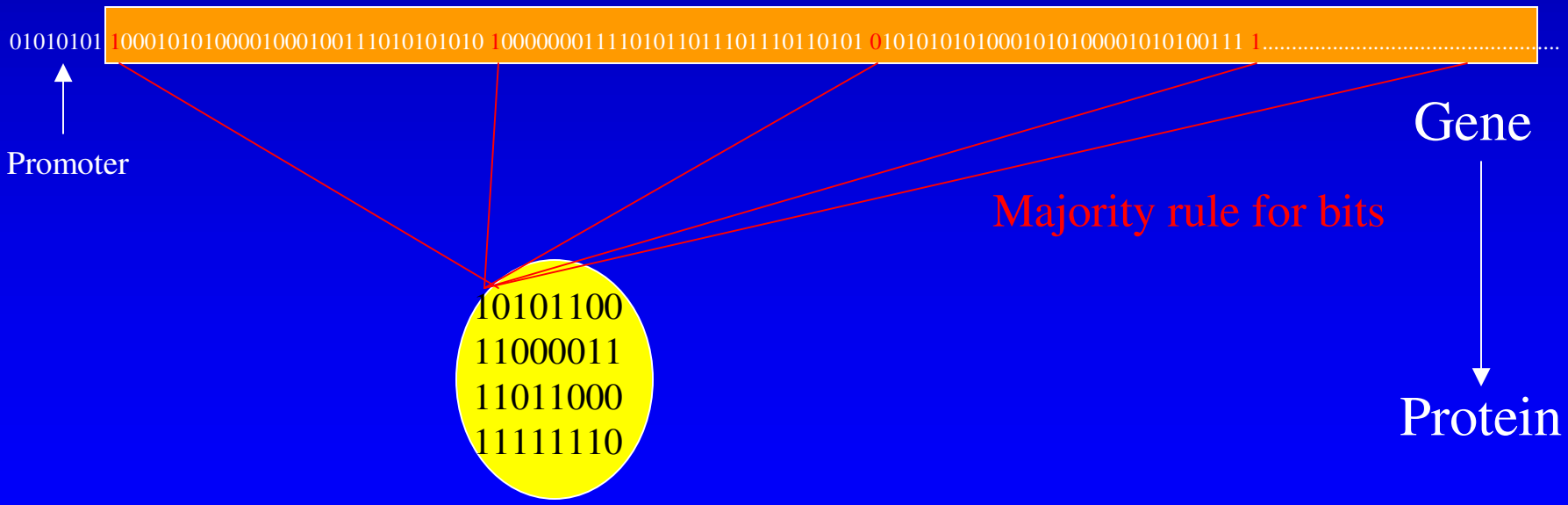


- Gene, including promoter site
- Protein
- Enhancer Site, upstream from gene
- Inhibitor Site, upstream from enhancer

The Model: Proteins from Genes

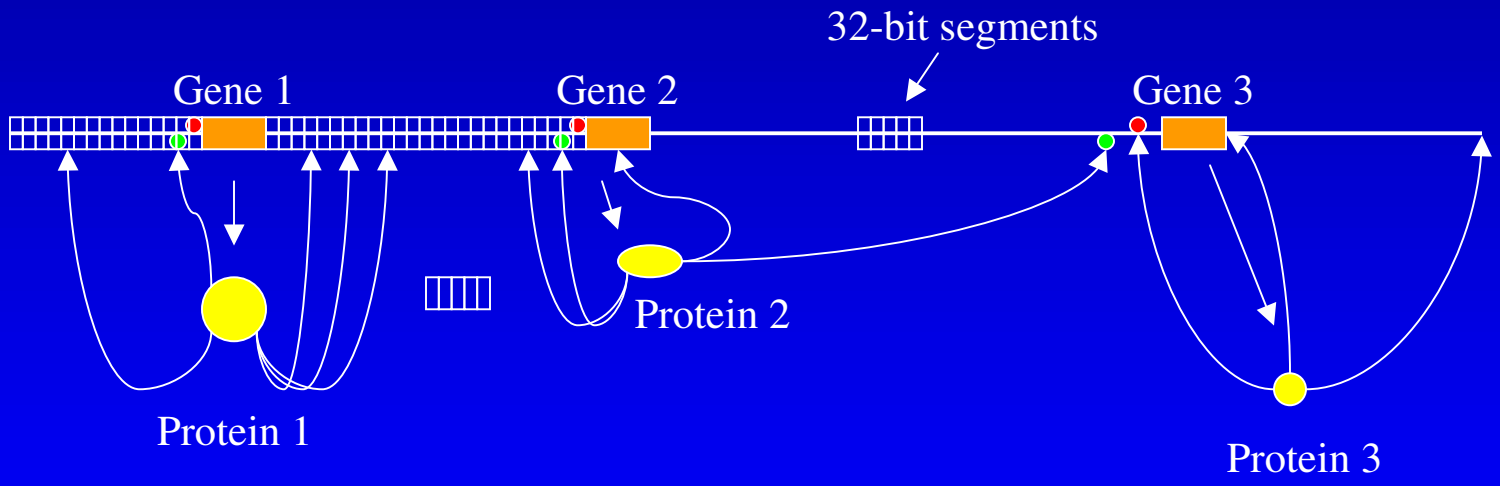
Gene: Bit string of fixed length (5 x 32 bits)
Protein: Bit string of fixed length (1 x 32 bits)
Promoter: 8 bit sequence 01010101
Probability of promoter and gene: 2^(-8) = 0,39 %

Genotype – Phenotype Mapping (Many-to-one-mapping)



The Model: Genome-Protein-Interaction

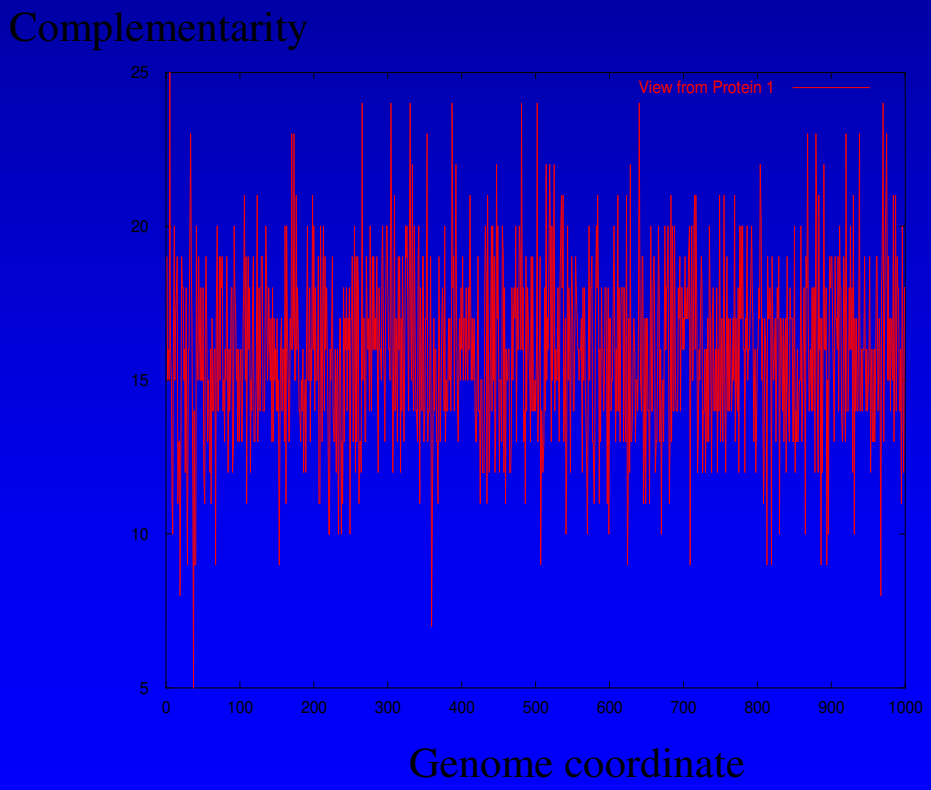
- Both, proteins and genome made from the same material: bit strings
- 32-bit segments might form complementary segments
- Complementarity controls interaction (binding) in a non-linear way



Example: View from Protein

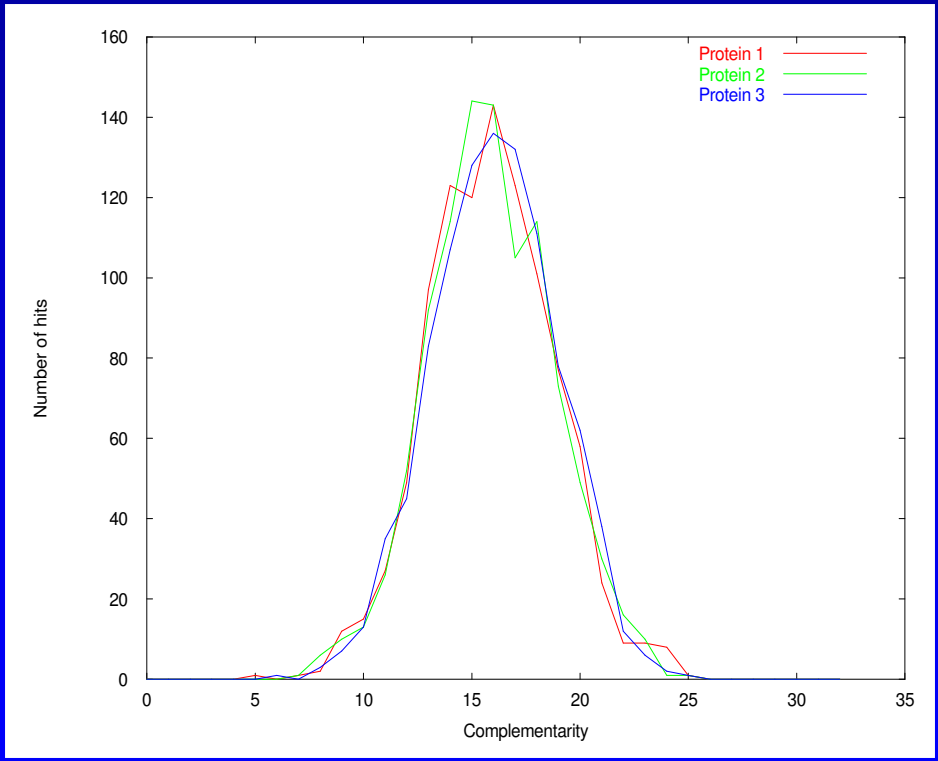
- Complementarity of random genome (1000 x 32-bit) and a protein
- Viewed from the protein perspective

- Average complementarity: 16
- C_min: 5; C_max=25



Example: View from Protein II

- Distribution of complementarities in random genome (1000 x 32-bit)
- 3 proteins

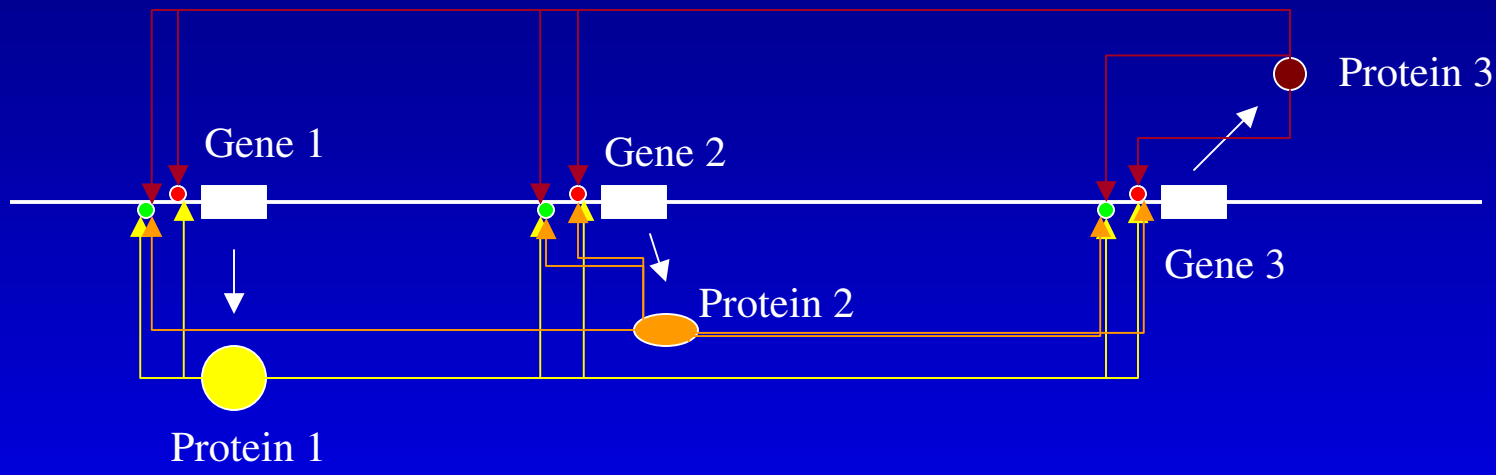


Example: Specificity

- Specificity of complementarities in random genomes (1000 x 32-bit)
- Appr. constant number of most specific interaction

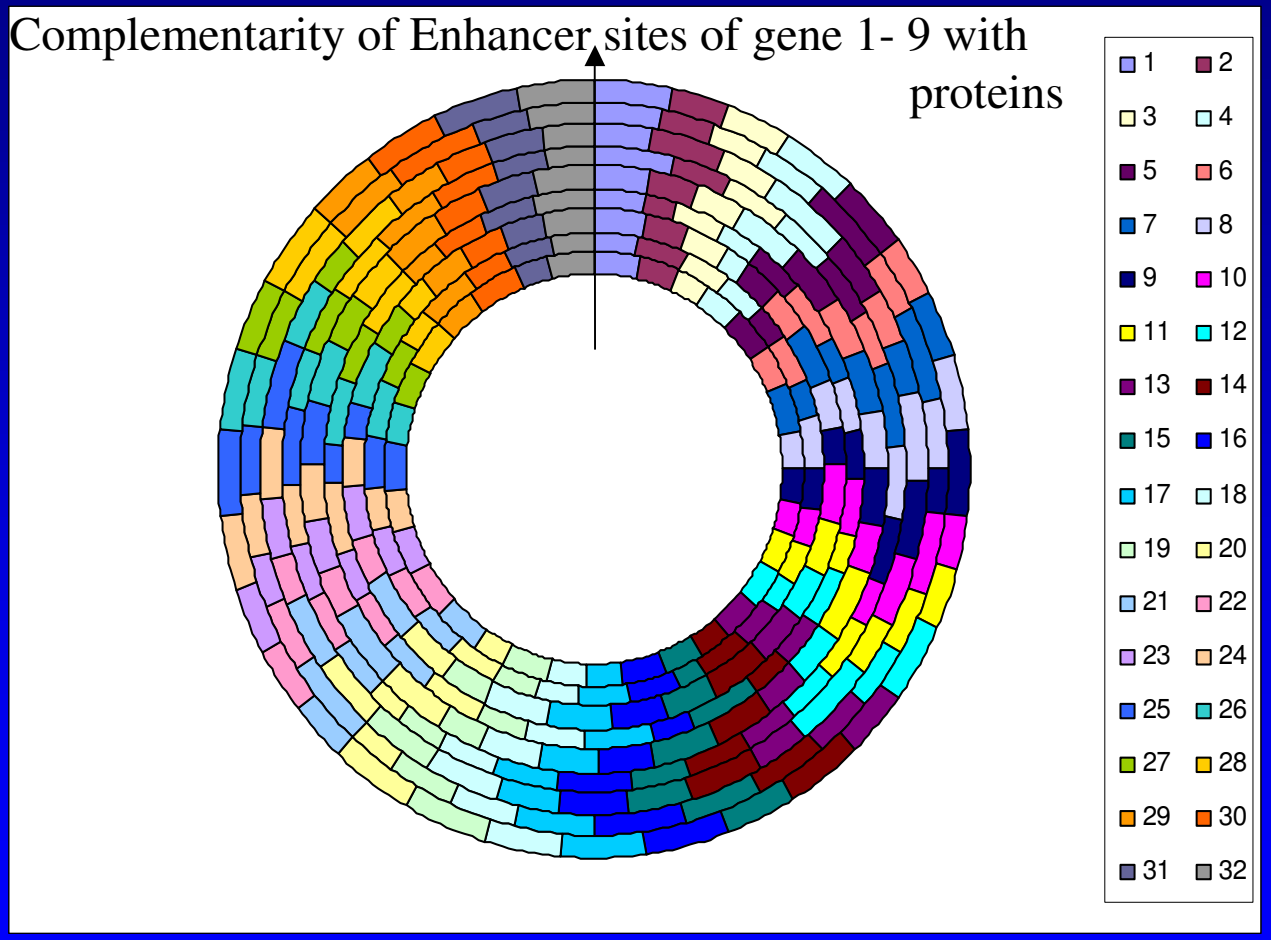
Genome Length	Number of Genes/Proteins	Max. Compl.	Occurence
1,000	3	25	3
10,000	37	28	4
100,000	409	30	3

Example: View from Genome



Interaction strength of enhancer / inhibitor site with proteins

Example: View from Genome II



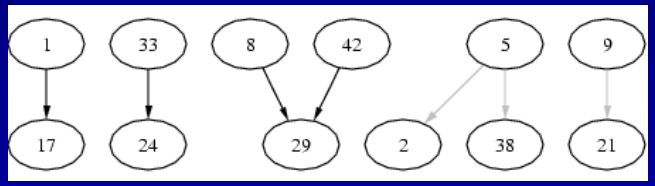
The Model: Network View

- Protein binding controls expression of genes.
- The higher the binding affinity, the stronger the control
- Network picture of the resulting interaction
- Difference between various kinds of generation of the (random) genome:
- Fully random genome vs. genome generated by duplication and divergence

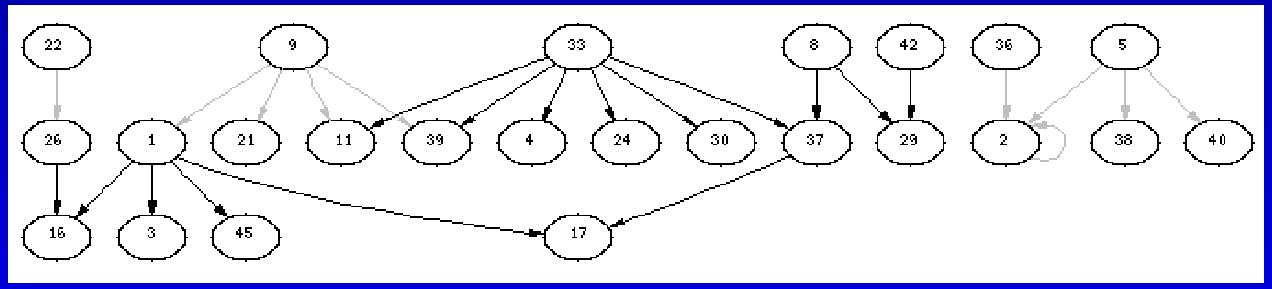
Network Example: Genome with 32 genes

C = Complementarity

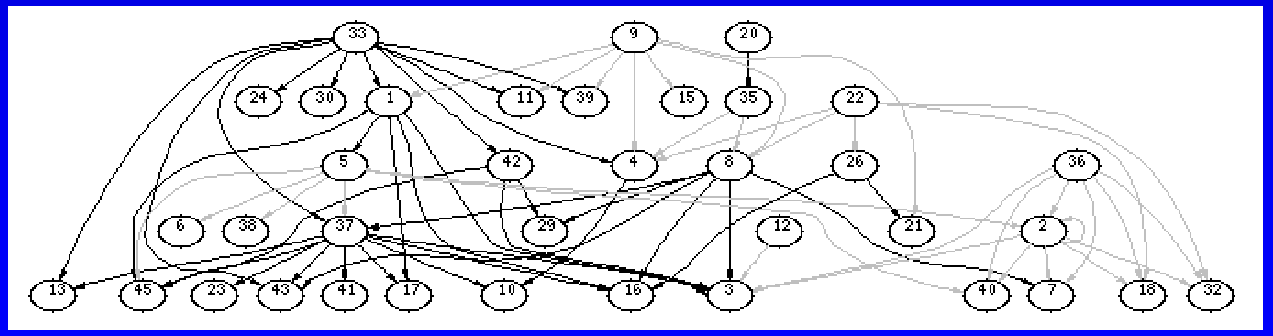
Shown are connections with C>24



Shown are connections with C>23

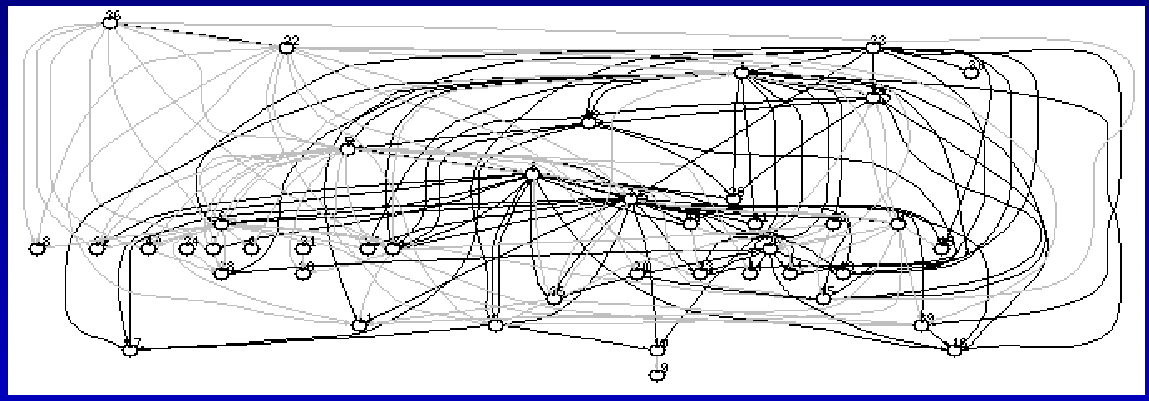


Shown are connections with C>22



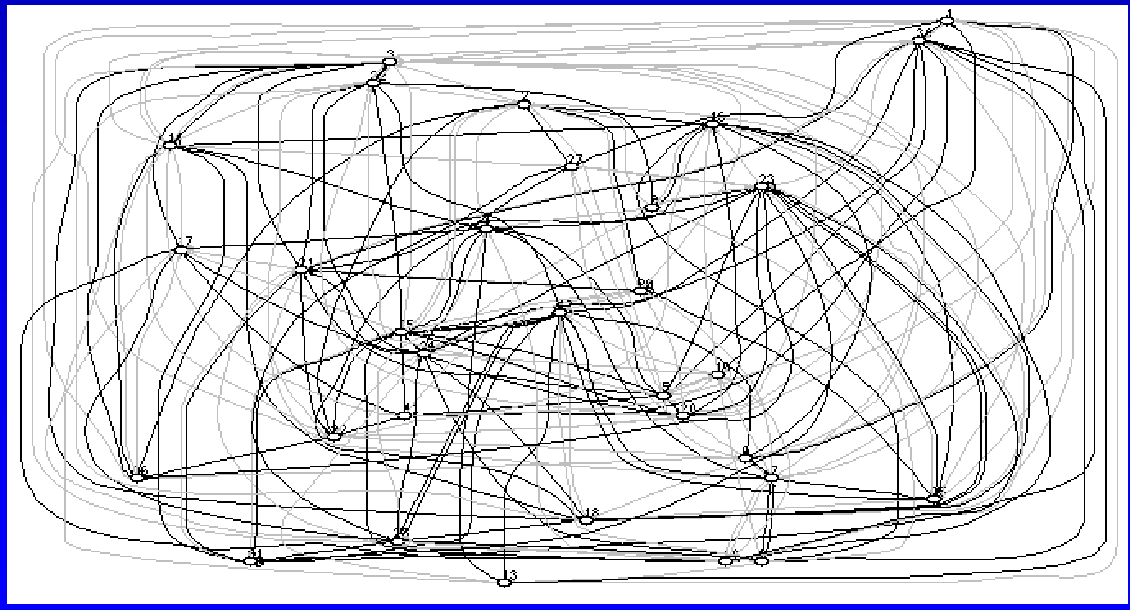
Example for Network View: Genome with 32 genes

Duplication and divergence
Shown are connections with
C>21



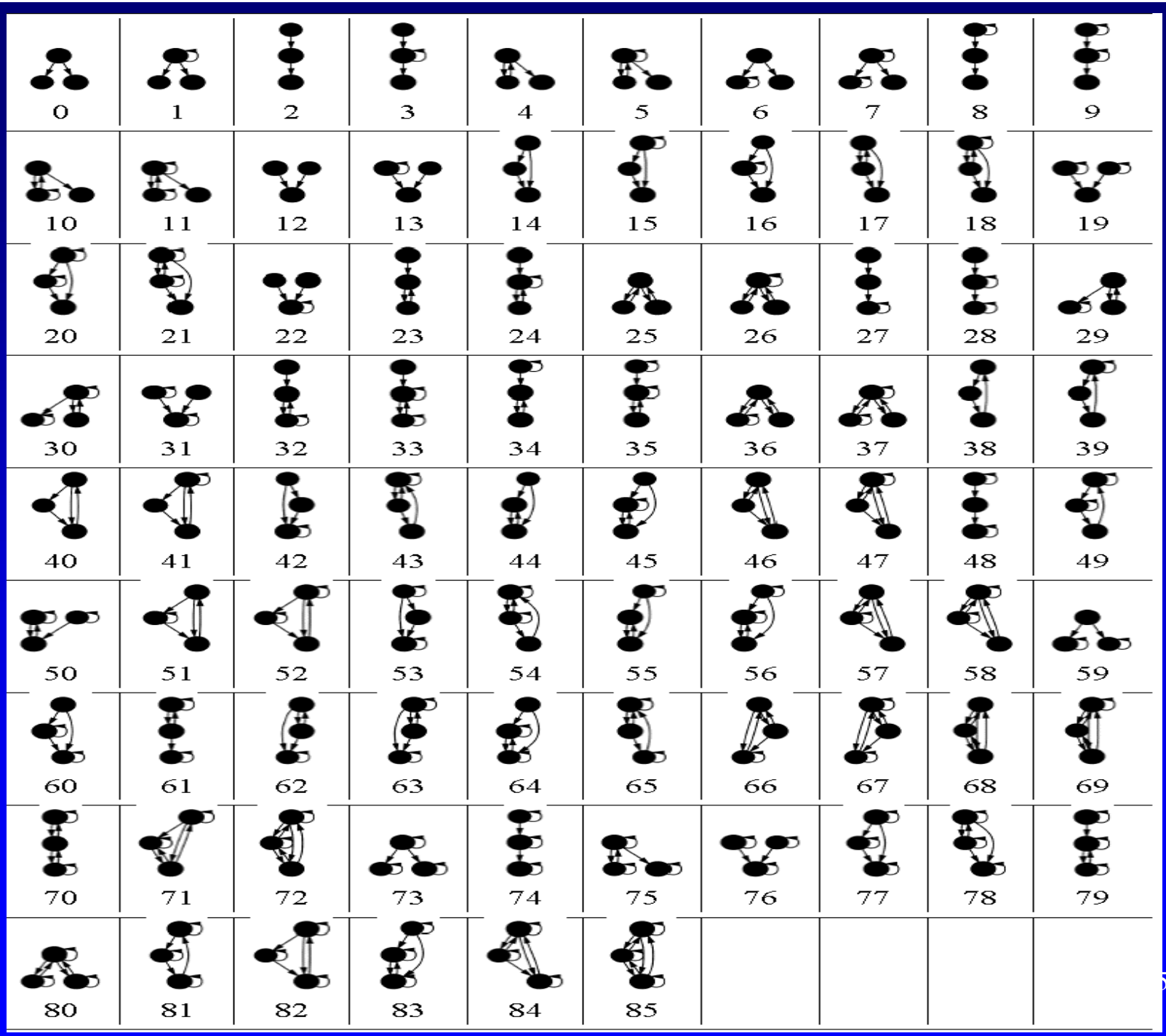
C = Complementarity

Random generation
Shown are connections
with C>19

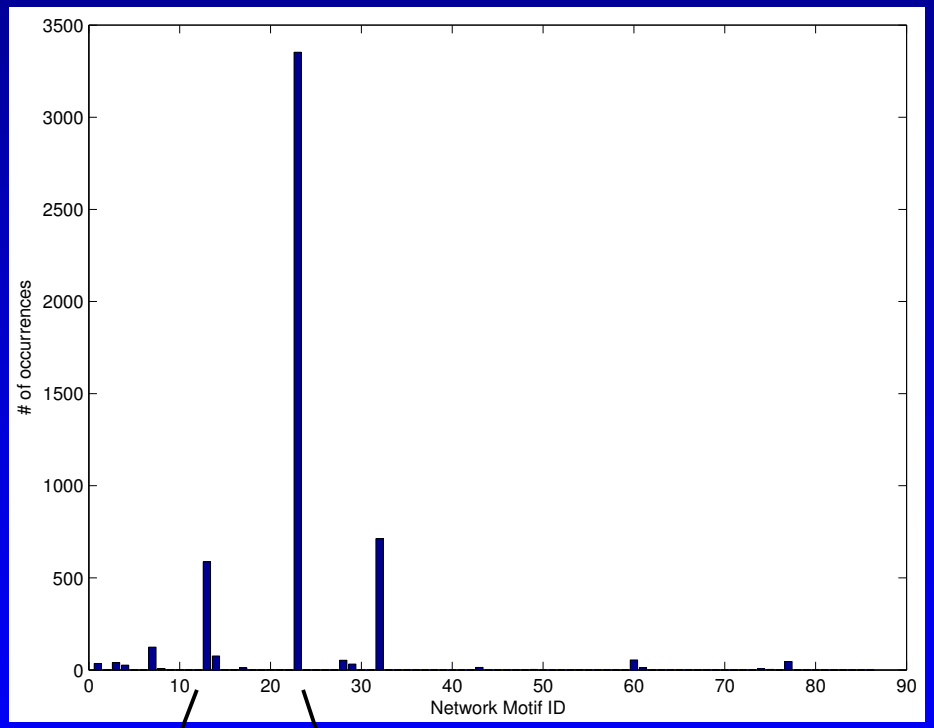


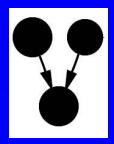
Network Motifs

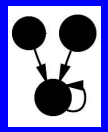
- Structural elements (subgraphs) which are basic elements of complex networks
- Certain motifs occur with significantly higher probability than in random networks
- Network Motifs have been shown to be conserved over evolutionary time
- We look at motifs that include feedback-connections


0 1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18 19
20 21 22 23 24 25 26 27 28 29
30 31 32 33 34 35 36 37 38 39
40 41 42 43 44 45 46 47 48 49
50 51 52 53 54 55 56 57 58 59
60 61 62 63 64 65 66 67 68 69
70 71 72 73 74 75 76 77 78 79
80 81 82 83 84 85

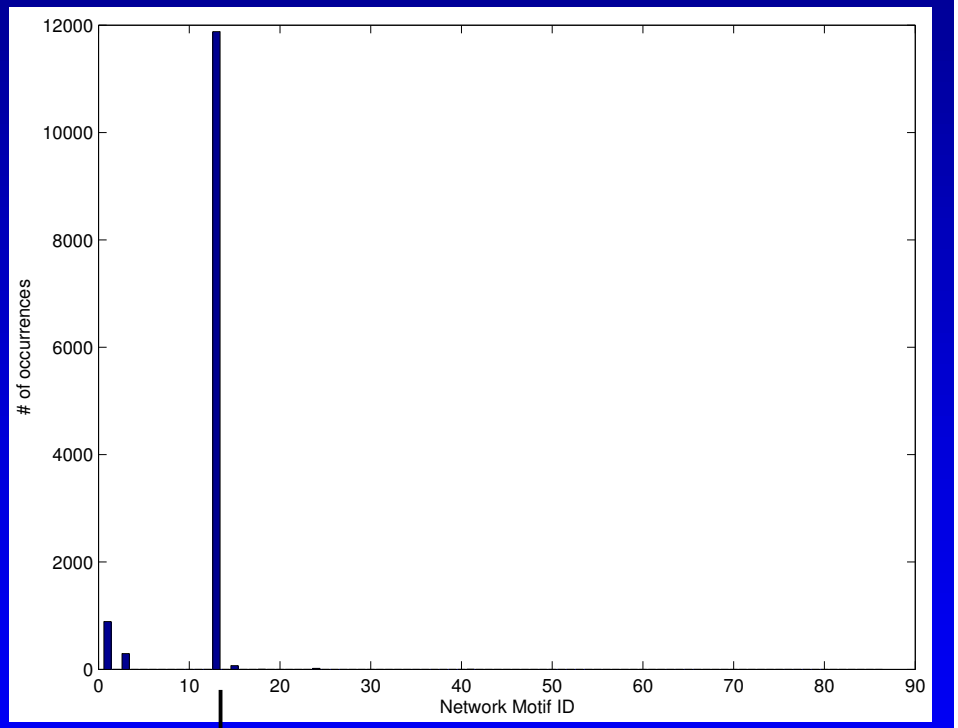
E. Coli

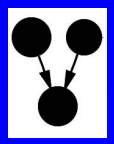




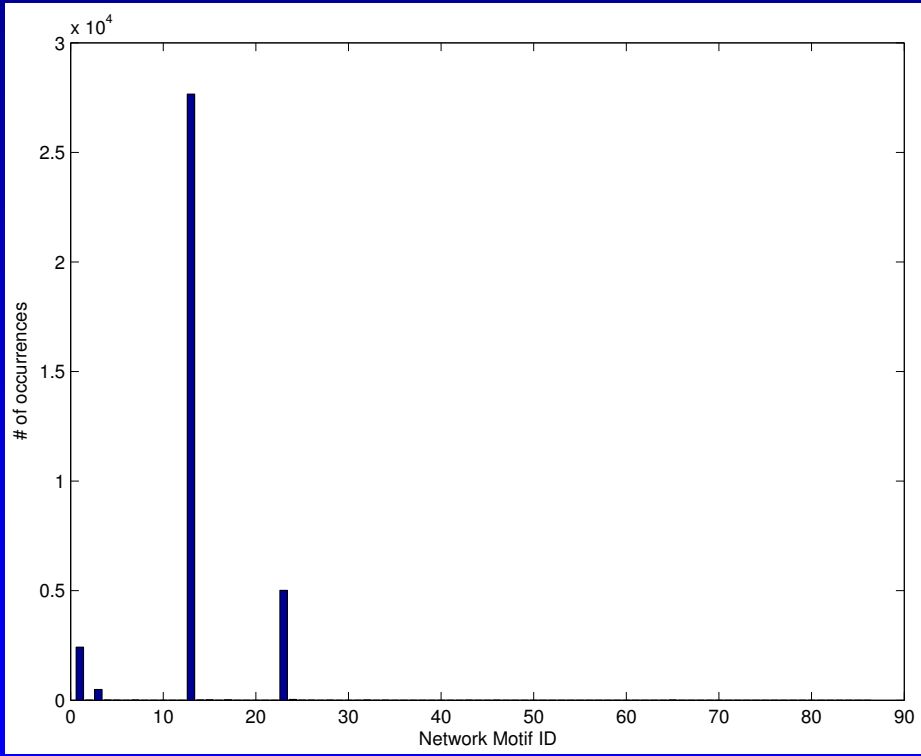


S. Cervisiae

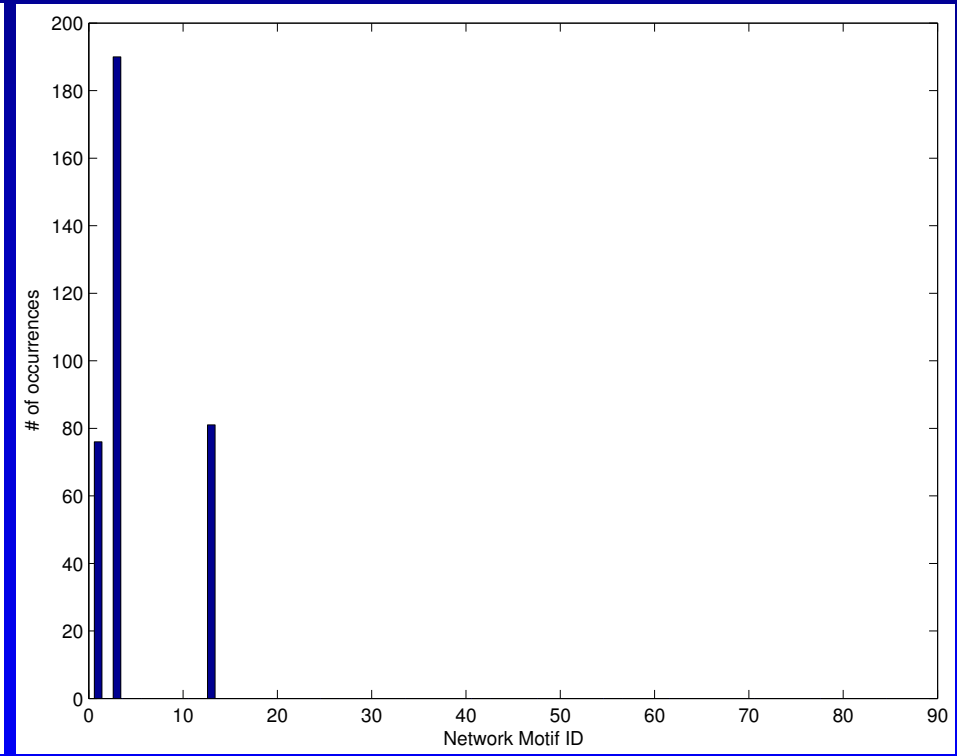




Duplication / Divergence



Random



The Model: Protein-Genome-Interaction

- Protein binding to enhancer / inhibitor sites leads to expression of the site gene.
- Influence on expression of gene is non-linear with binding strength

$$en_i = \sum_j c_j e^{\beta(u_j^+ - u_{\max})} / N \qquad in_i = \sum_j c_j e^{\beta(u_j^- - u_{\max})} / N$$

with c_j the concentration of protein j and $u_{\max}$ maximum complementarity.

- en and in determine the (competitive) occupation probability of site i
- Both signals contribute to expression of gene i

$$\frac{dc_i}{dt} = \delta(en_i - in_i)c_i - \Phi$$

with Φ a flow term normalizing production of proteins.

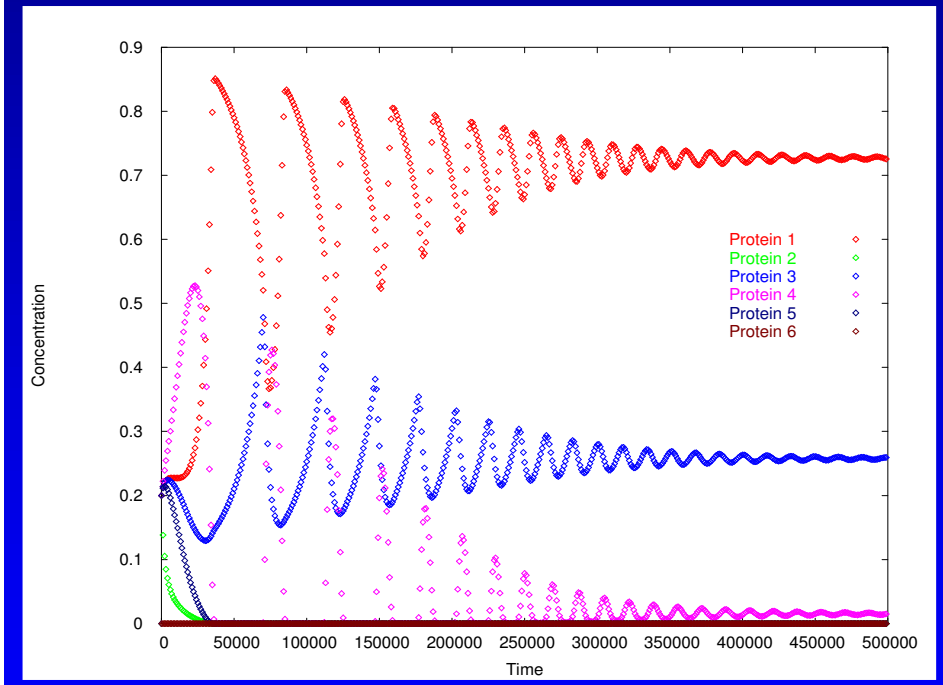
The Model: Protein-Genome-Interaction II

3 sorts of competition:

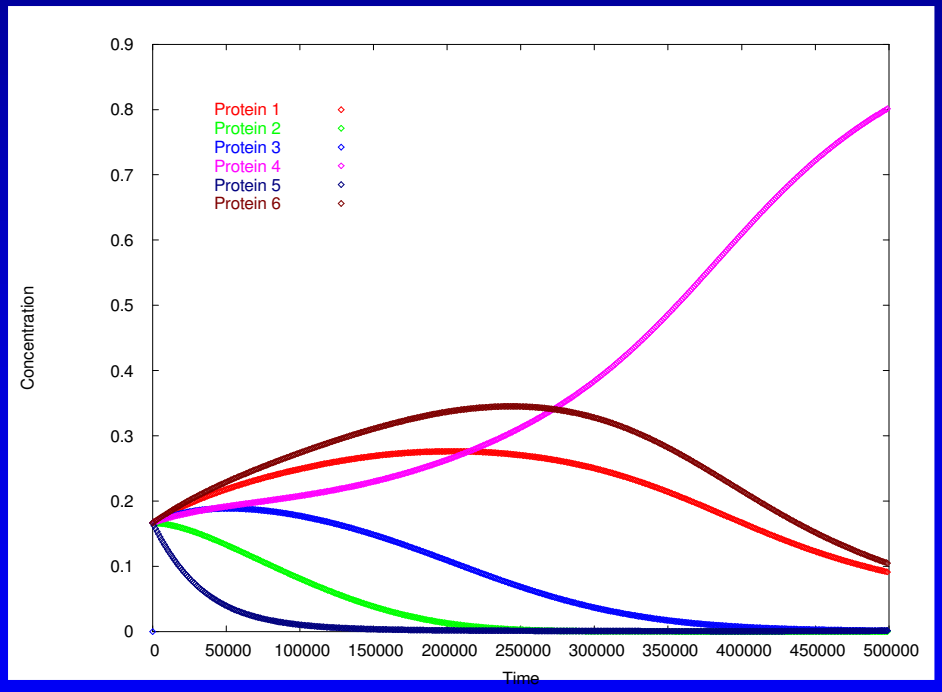
1. Competition of proteins for binding sites
2. Competition of binding sites for proteins
3. Competition of genes for raw material for expression

Example: Protein Dynamics

- Large variety of behavior
  - Oscillations, transient states, point attractor

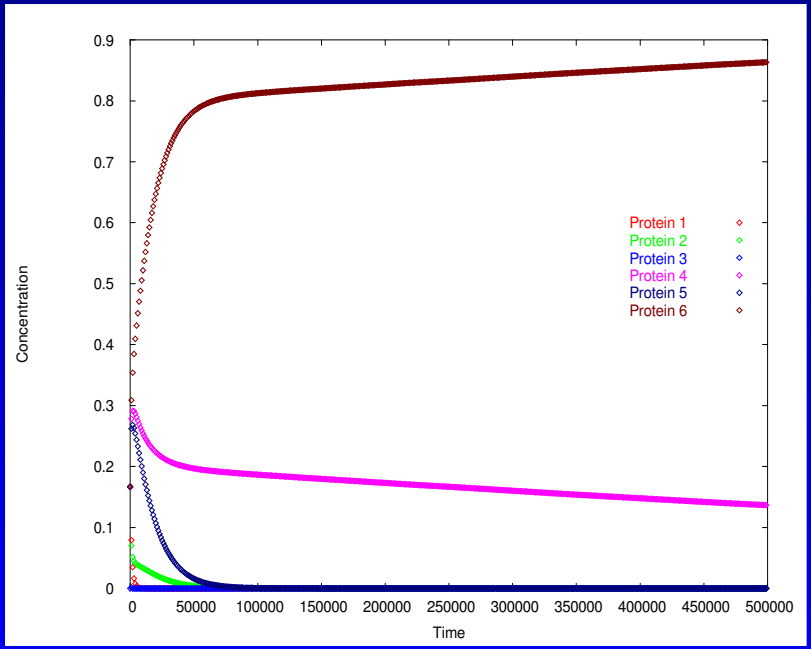


Run1

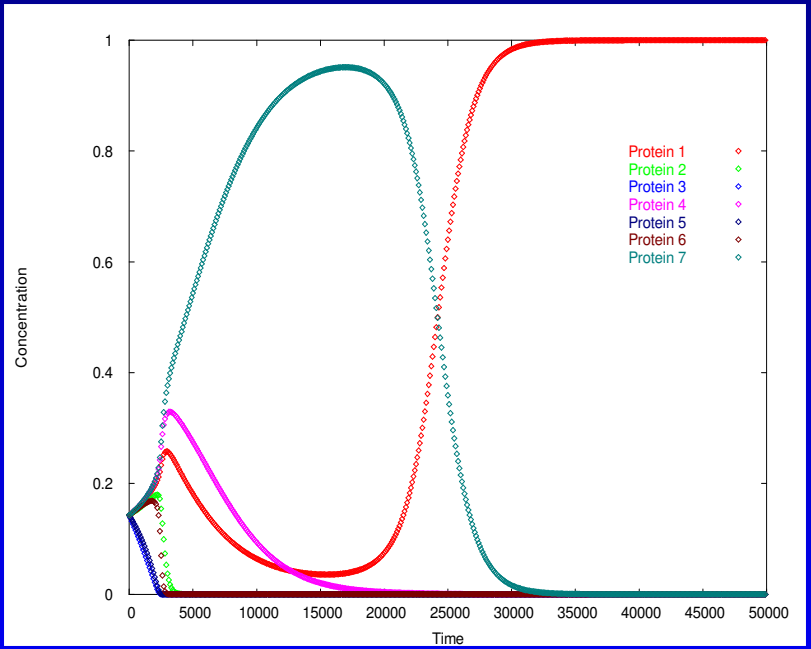


Run2

Example: Protein Dynamics II



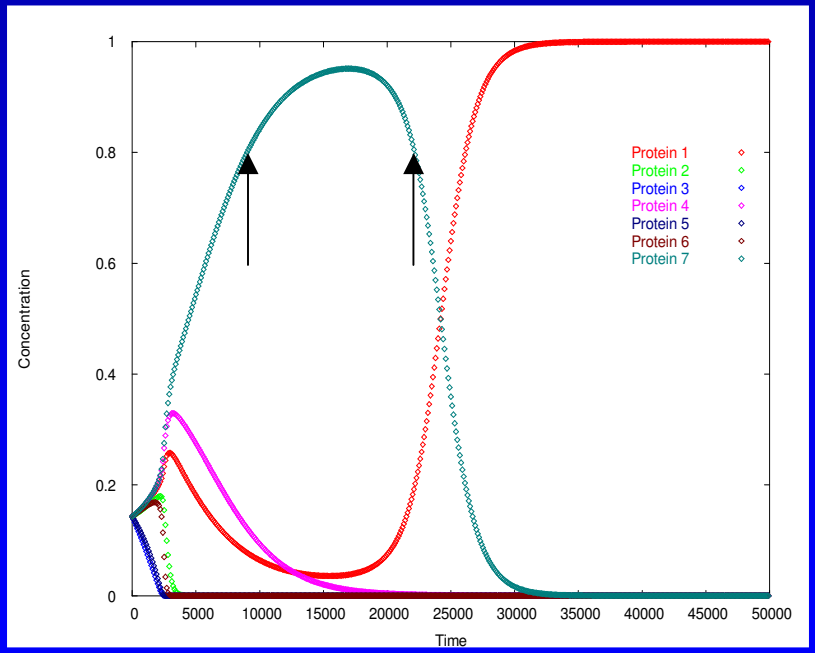
Run3



Run4

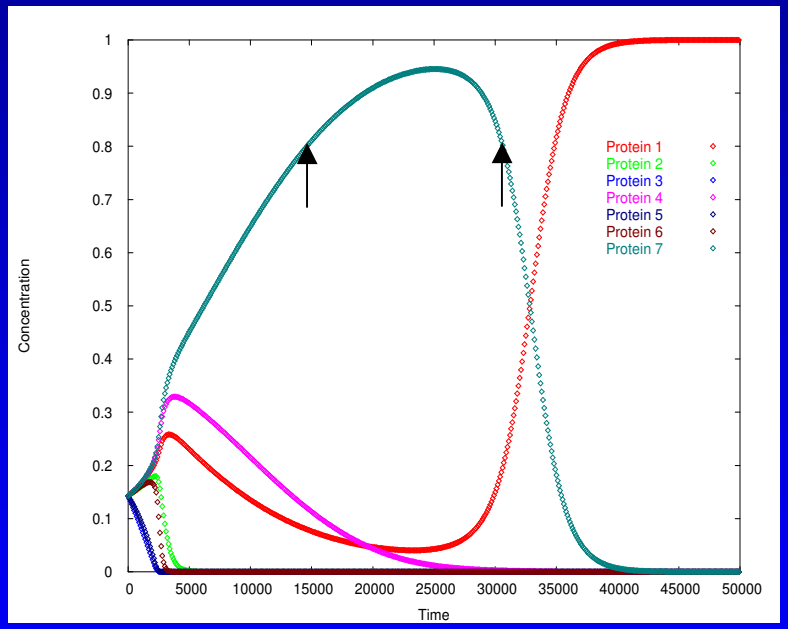
Example: Heterochronic Control

- Certain proteins must be present in particular concentrations at determined moments (t= 9,500, t‘=21,000)
- These proteins are used to control production of other proteins in turn
- Some of these proteins are used to build structures of the body

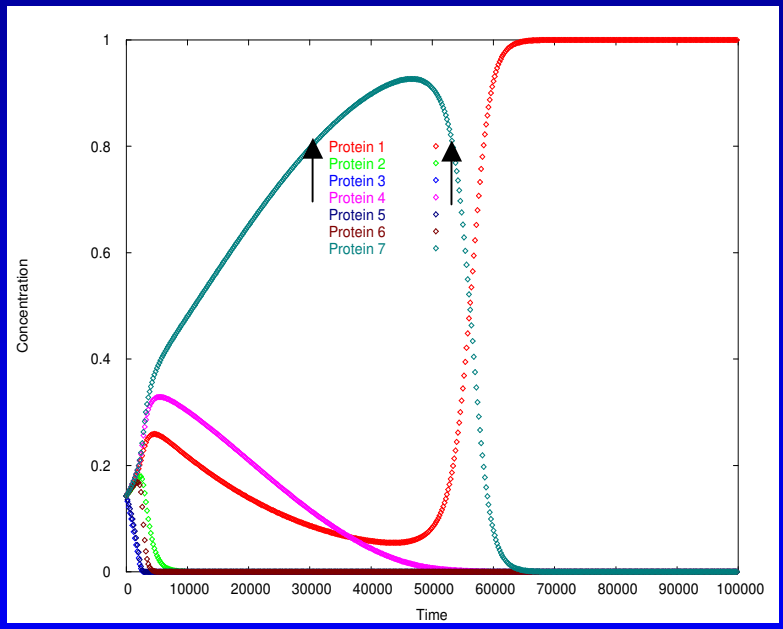


Example: Heterochronic Control II

- Single bit flips change timing seriously
- 1 bit mutation: t=15,000 , t‘=30,000
- 1 bit mutation: t=30,000 , t‘=53,000



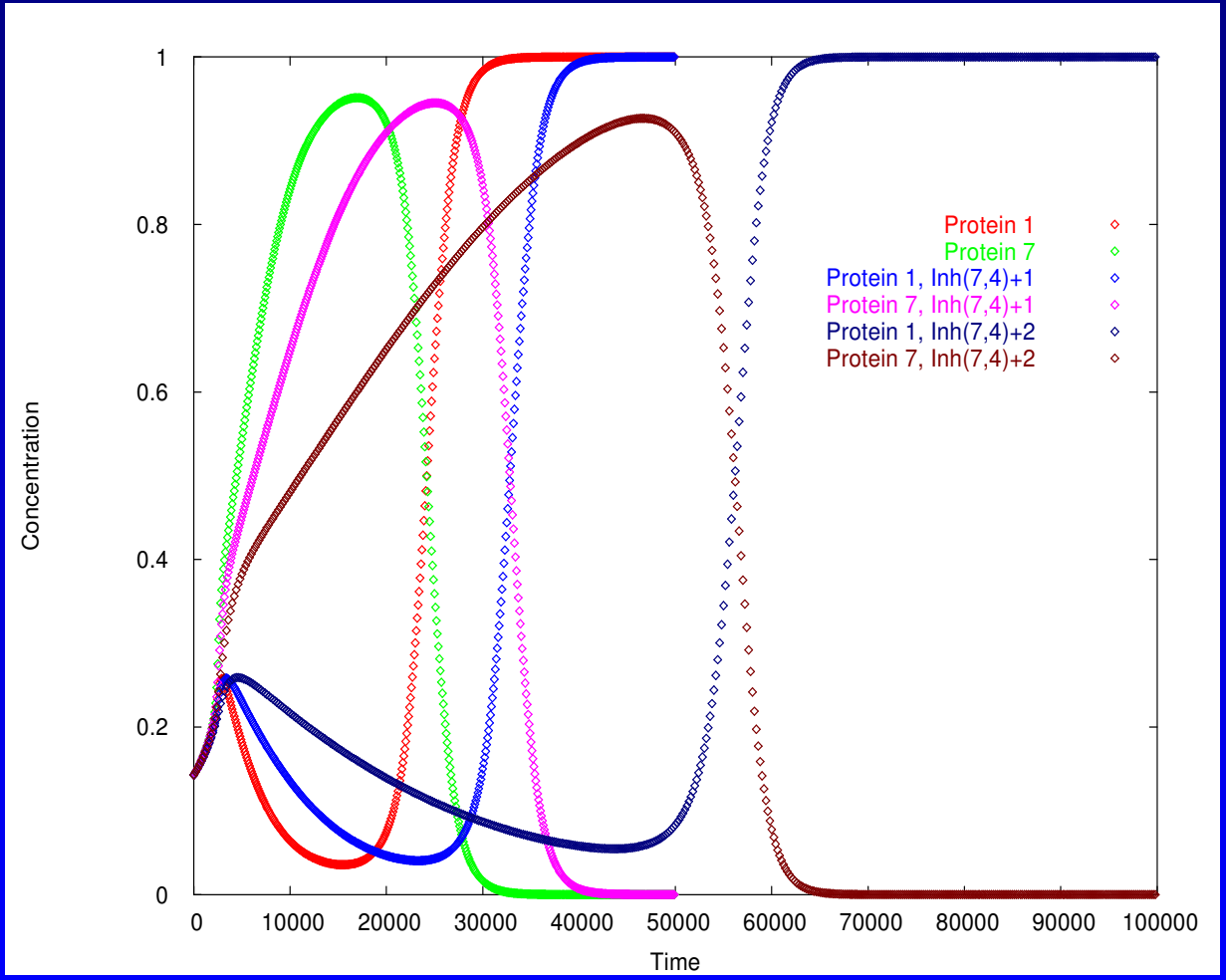
Run5: Inh(4,7)+1



Run6 : Inh(4,7)+2

Example: Heterochronic Control III

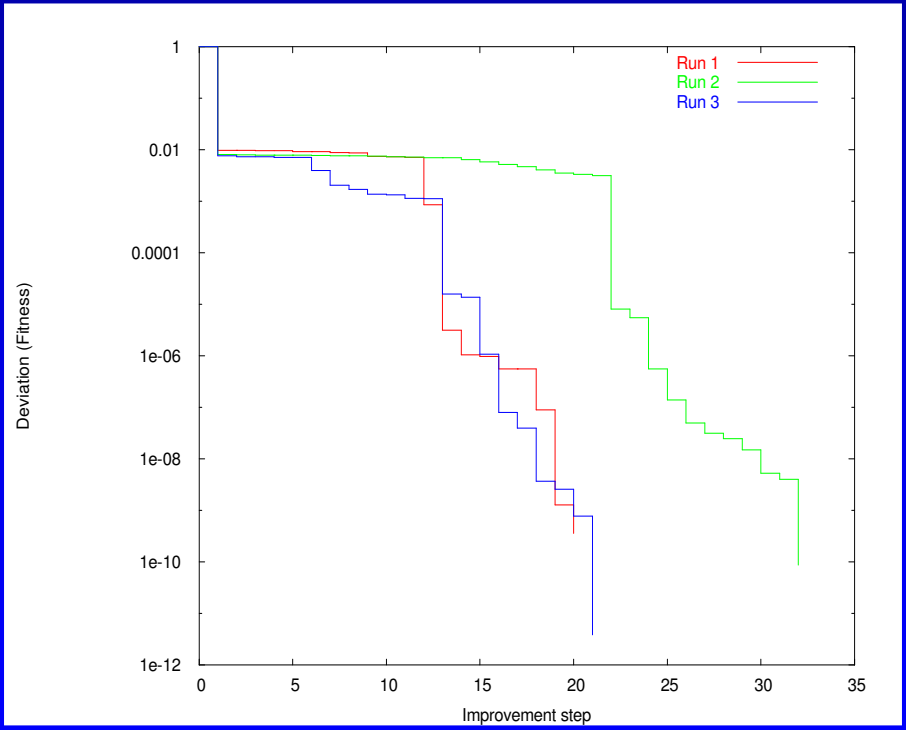
- Relative shift due to single bit flips

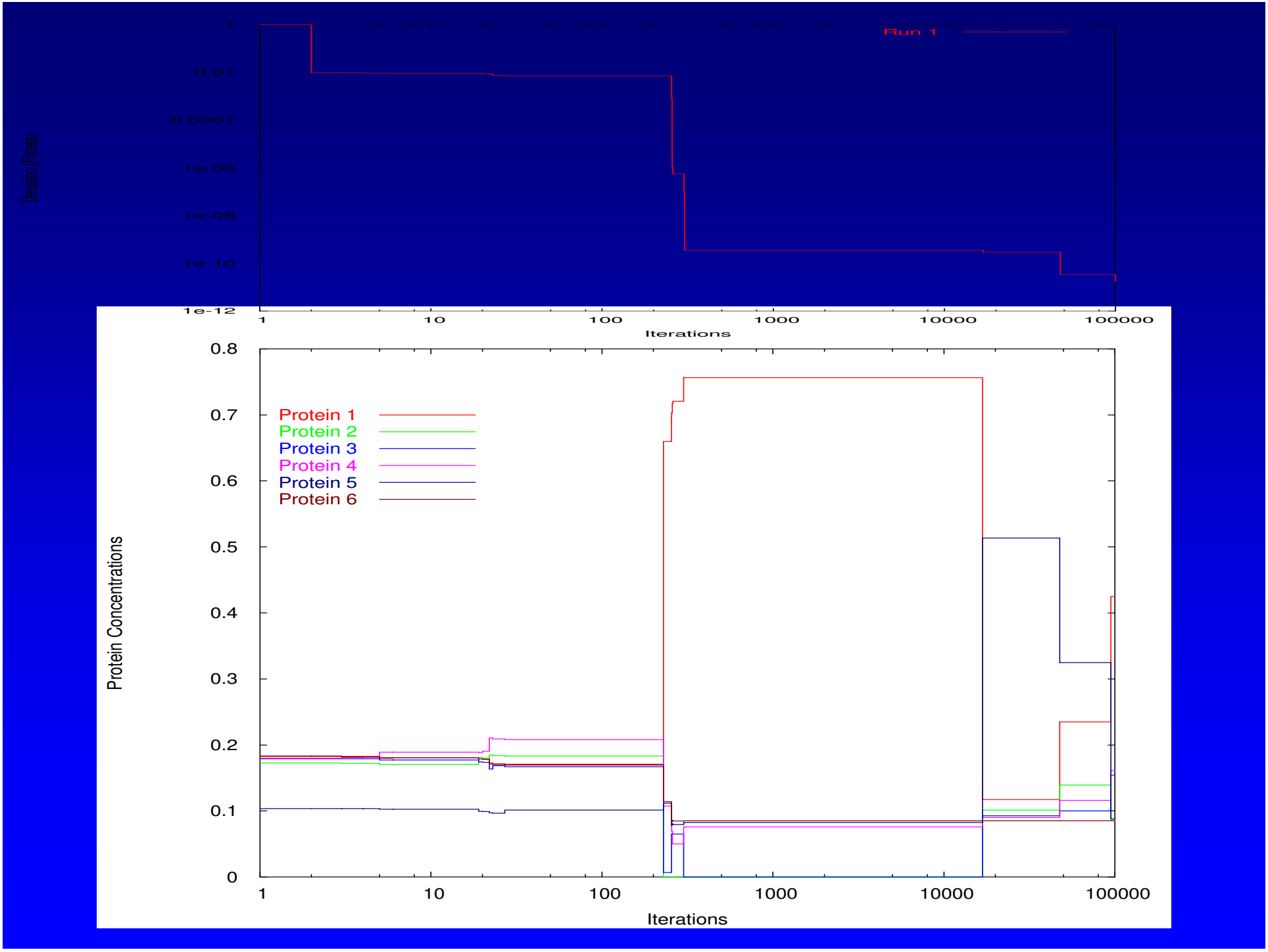


Example: Evolution

Evolutionary problem:

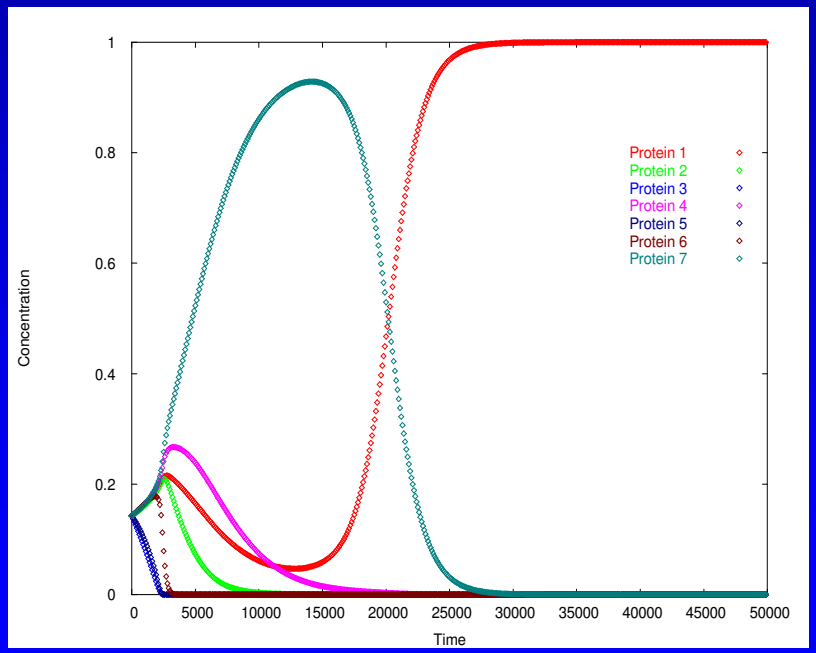
Find a configuration of protein expression levels allowing Protein 6 to be expresses at C=0.085



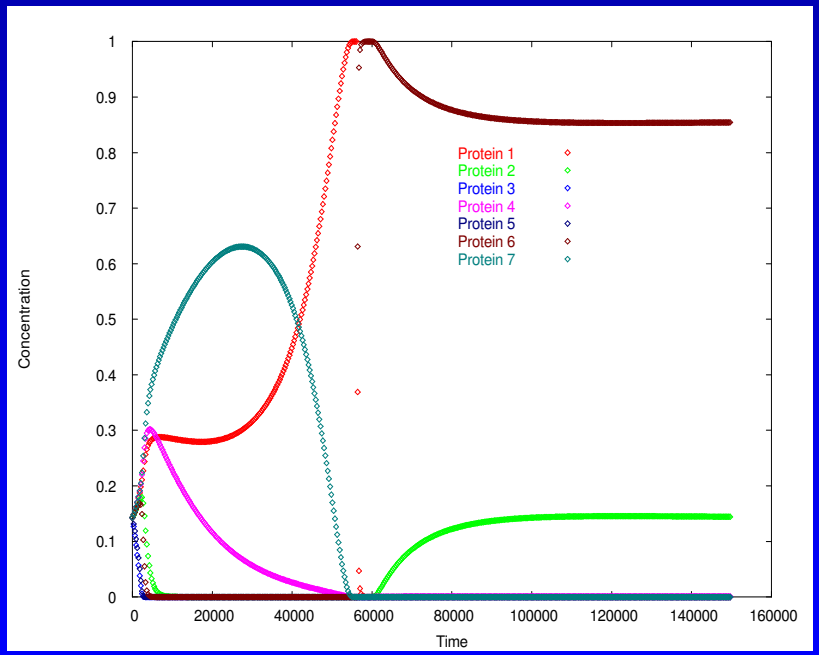

Run 1
Deviation (Fitness)
1
0.01
0.0001
1e-06
1e-08
1e-10
1e-12
1
10
100
1000
10000
100000
Iterations
Protein 1
Protein 2
Protein 3
Protein 4
Protein 5
Protein 6
Protein Concentrations
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
10
100
1000
10000
100000
Iterations

Example: Evolution II

- Neutral steps
- Jumps in variation
- Innovation



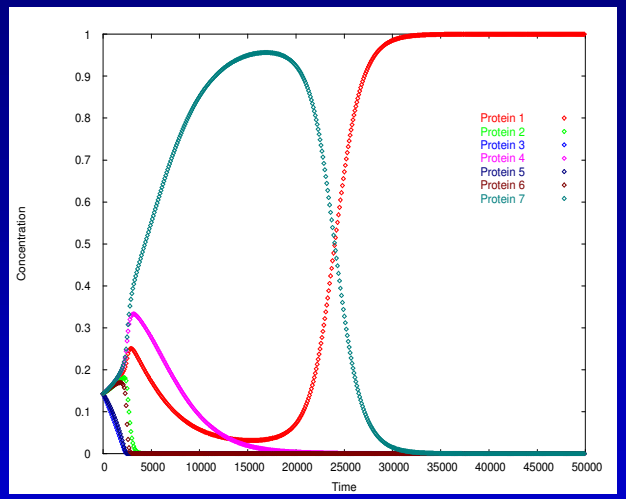
Run7: 1 mutation in exon of protein 4



Run8 : 2 mutations in exon of protein 1

Example: Stability and Communication

Figure 1
Protein 1: - 10 %



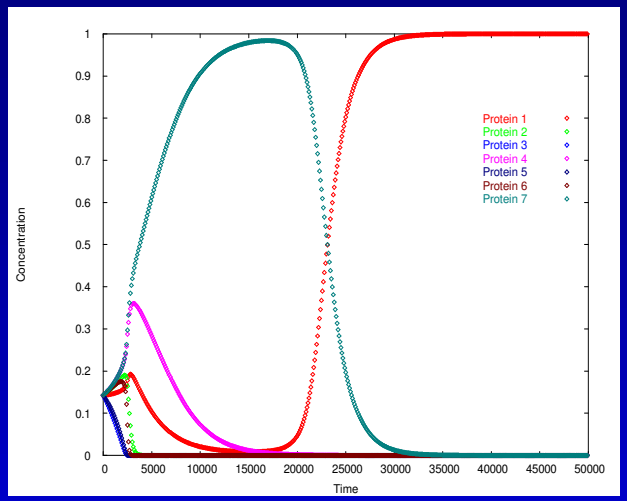
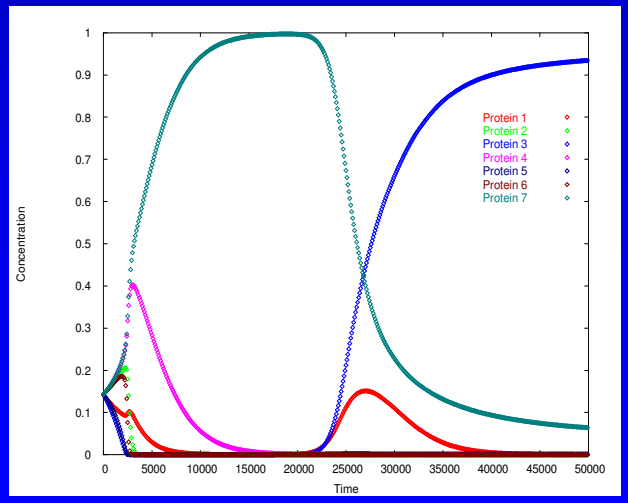


Figure 2
Protein 1: - 100 %

Figure 3
Protein 1: - 300 %



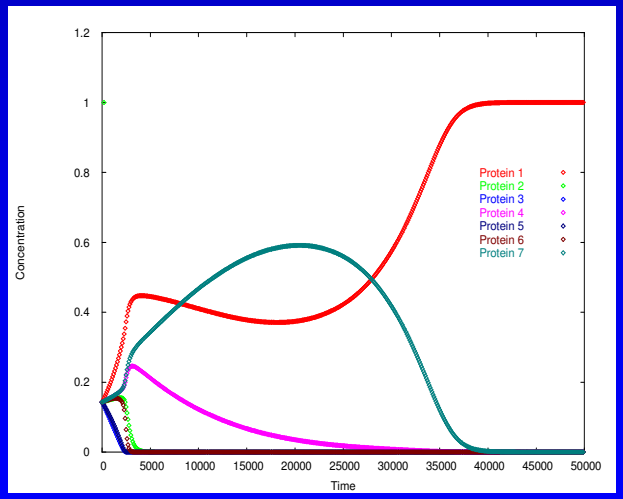
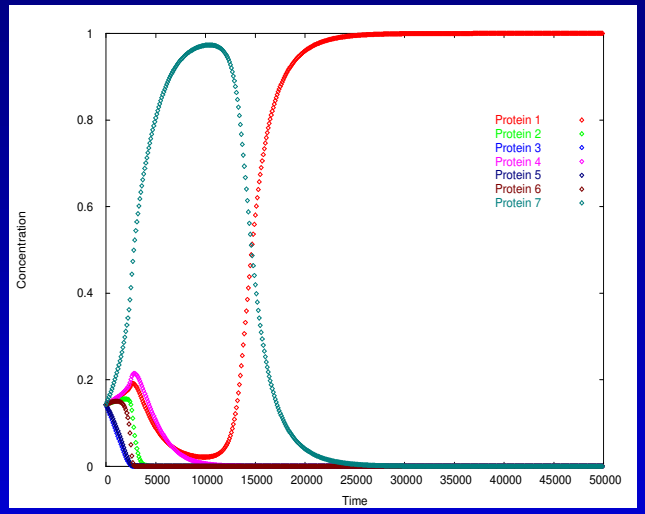


Figure 4
Protein 1: + 200 %

Example: Stability and Communication II

Figure 5
Protein 7: + 200 %



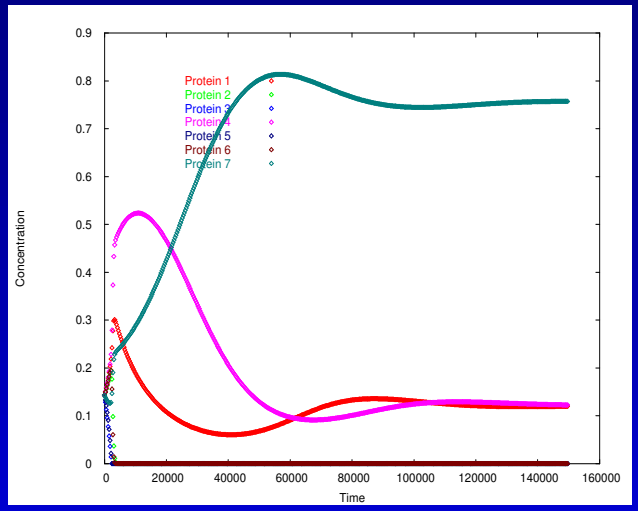
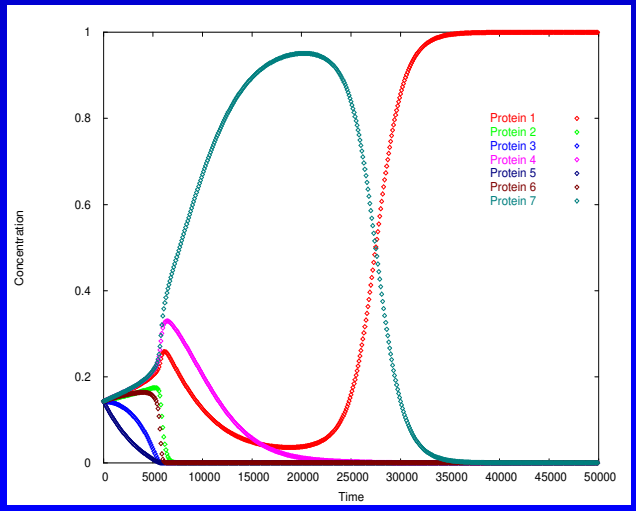


Figure 6
Protein 7: - 200 %

Figure 7
Protein 3: + 300 %



Summary

- Simplicity of the model
- Stability of production
- Extraordinary evolutionary plasticity
- Similarity of behavior
- Time as an essential variable

References

- http://www.cs.mun.ca/~banzhaf/