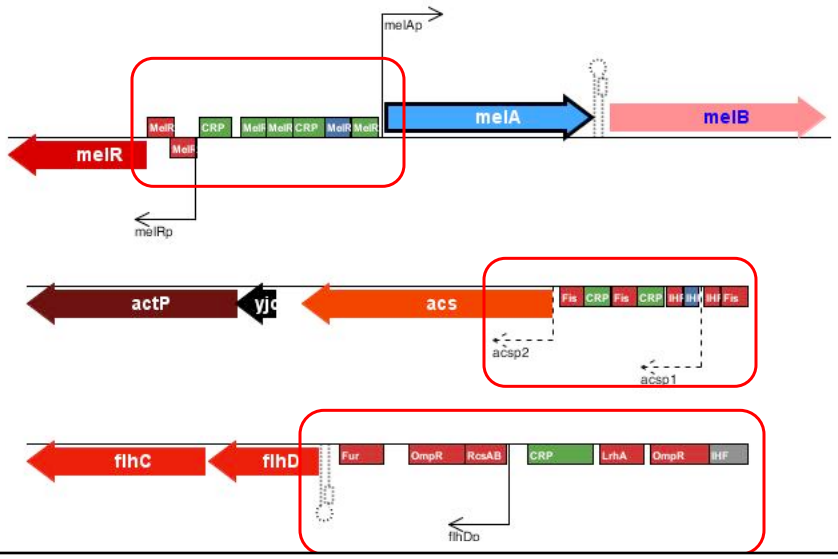
D. Combinatorial transcriptional control

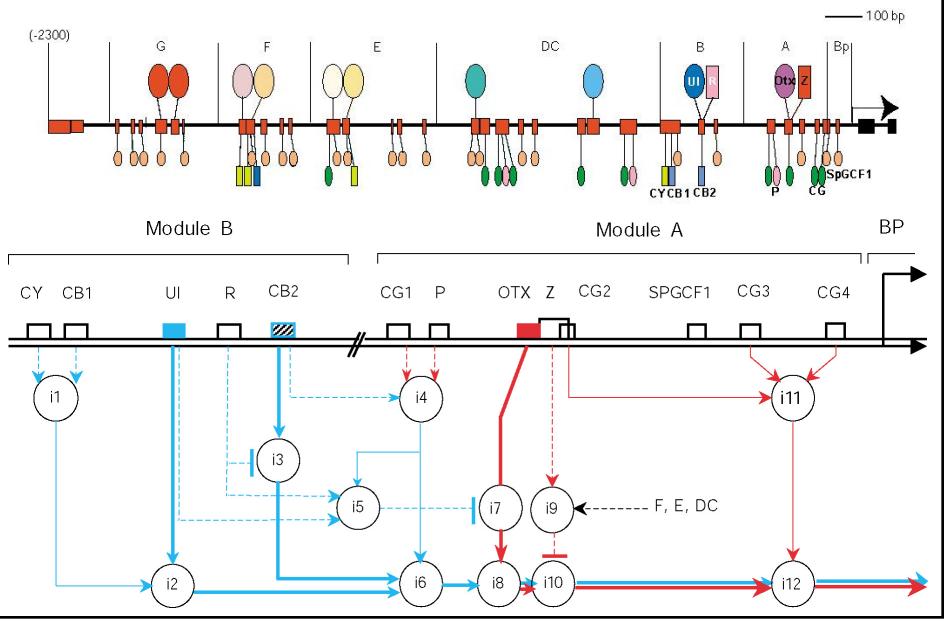
Complex bacterial promoters with multiple inputs:



1

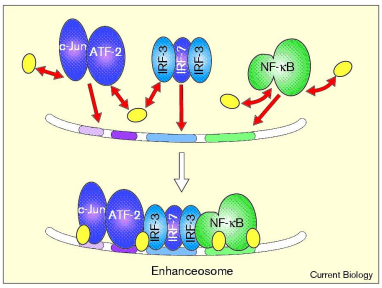
Complex eukaryotic promoter

[Davidson et al]



2

Mechanisms of complex transcriptional control?



- specific, complex protein-protein interaction

  … but different TFs can work together to implement different functions

  ➡ combinatorial control favors simpler, less specific interaction

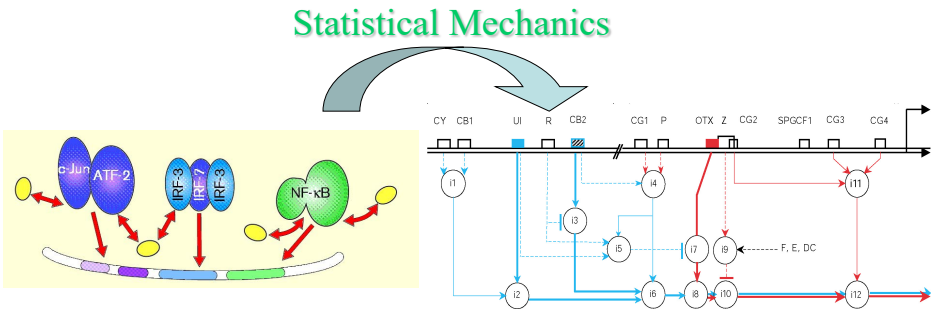
- alternative: **regulated recruitment** [Ptashne & Gann '97]
  - simple, glue-like interaction between TFs/RNAP
  - arrange DNA binding sites/strengths to accomplish desired control functions

  … but how to implement? and what are the limitations?

3

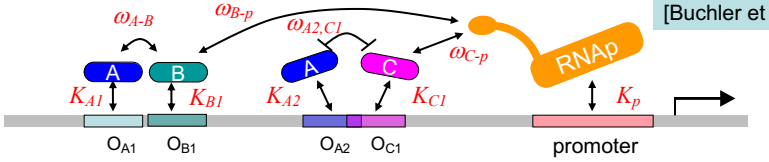
Theory of Combinatorial Transcription Control

[Buchler et al, PNAS '03]



4

Theory of Combinatorial Transcription Control



[Buchler et al, PNAS '03]

quantitative description:

via changes in regulatory sequences **alone**

- occupation of site j: $\sigma_j = \{0,1\}$
- TF-operator interaction: $K_j = \exp(-\beta\Delta G_j) \quad = 1 \sim 1000$ nM
- TF-TF and TF-RNAP interaction: $\omega_{i,j} = \exp(-\beta E_{ij}) \quad = \{0, 1, 10 \sim 100\}$
  - **regulated recruitment** [M. Ptashne]
  - long-distance interaction possible via DNA looping
- promoter activity ~ **equilibrium** promoter occupation prob $\mathcal{P}\left([TF_j]\right) \approx \langle \sigma_p \rangle$ [Shea & Ackers, '85]

➔ thermodynamic model

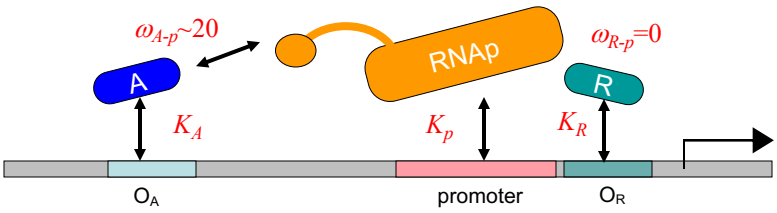
$$Z = \sum_{\{\sigma_i\}} \prod_{j=1}^{N} \left([TF_j]/K_j\right)^{\sigma_j} \cdot \prod_{i<j} \omega_{i,j}^{\sigma_i \sigma_j}$$

➡ programmable molecular Boltzmann machine!

➡ What kind of control functions $\mathcal{P}([TF_j])$ are implementable via the appropriate choices of $\{K_j, \omega_j\}$?

5

1. non-interacting



simple activation:

$$W_{\text{off}} = 1 + \frac{[A]}{K_A}$$

$$W_{\text{on}} = \frac{[RNAp]}{K_p} + \omega_{A-p} \cdot \frac{[A]}{K_A} \cdot \frac{[RNAp]}{K_p}$$

$$\mathcal{P} \approx \frac{W_{\text{on}}}{W_{\text{off}}} = \frac{[RNAp]}{K_p} \cdot \frac{1 + \omega_{A-p}[A]/K_A}{1 + [A]/K_A}$$

(for typical weak promoters)

simple repression:

$$\mathcal{P} \approx \frac{[RNAp]}{K_p} \cdot \frac{1}{1 + [R]/K_R}$$

co-regulation multiplicative

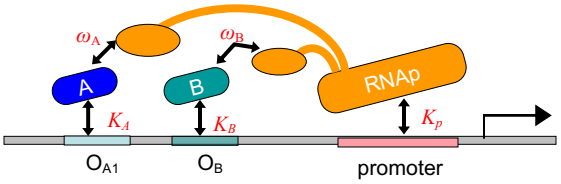
$$\mathcal{P} \propto \frac{1 + \omega_{A-p}[A]/K_A}{1 + [A]/K_A} \cdot \frac{1}{1 + [R]/K_R}$$

[example: lac promoter (details later)]

6

2. Synergistic activation

RNAp can simultaneously contact two TFs (e.g., Crp and Fnr)



statistical weight W for each configuration $\{\sigma_A, \sigma_B, \sigma_p\}$, with $q_X = [X]/K_X$

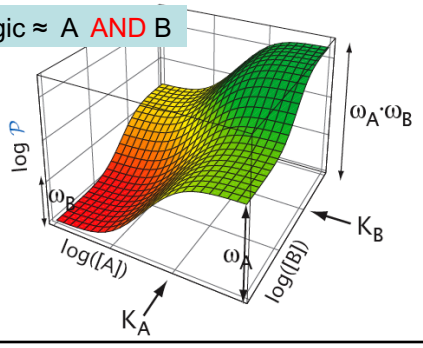
$$W_{off}\begin{cases} W(0,0,0)=1 \\ W(1,0,0)=q_A \\ W(0,1,0)=q_B \\ W(1,1,0)=q_A \cdot q_B \end{cases}$$

$$W_{on}\begin{cases} W(0,0,1)=q_p \\ W(1,0,1)=\omega_A \cdot q_A \cdot q_p \\ W(0,1,1)=\omega_B \cdot q_B \cdot q_p \\ W(1,1,1)=\omega_3 \cdot q_A \cdot q_B \cdot q_p \end{cases}$$

consider $\omega_3 \approx \omega_A \cdot \omega_B$

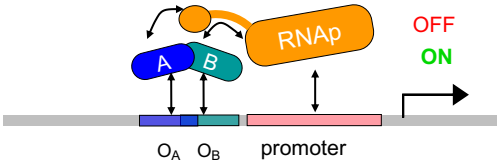
$$\mathcal{P}([A],[B]) \approx \frac{W_{on}}{W_{off}} \approx q_p \frac{1+\omega_A q_A}{1+q_A} \cdot \frac{1+\omega_B q_B}{1+q_B}$$

logic ≈ A AND B

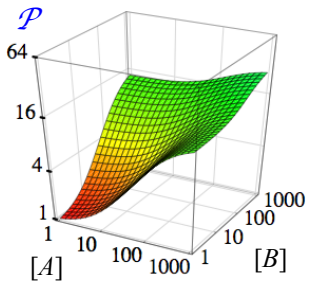


7

3. Competitive (or "independent") activation

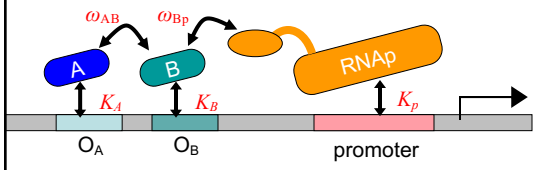


$$\mathcal{P}([A],[B]) \approx q_p \frac{1+\omega_A q_A + \omega_B q_B}{1+q_A+q_B}$$

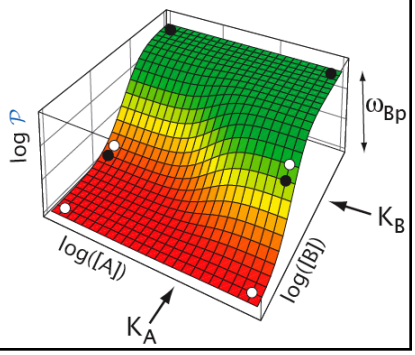


logic ≈ A OR B

4. Cooperative activation

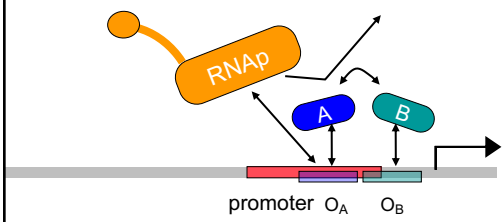


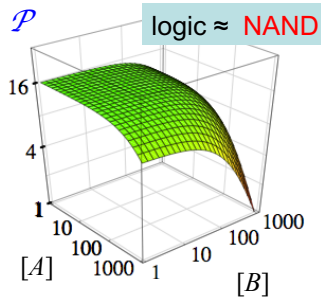
could work as AND or OR by choosing K_A, K_B



8

5. Cooperative repression

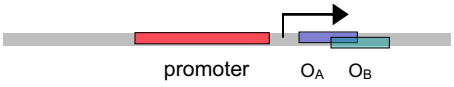




logic ≈ NAND

note: direct interaction between A and B
not necessary (cf "collaborative competition")

6. Competitive repression



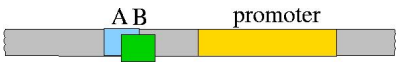
logic ≈ NOR

9

More complex control function, e.g., **XOR** ?

A/B	AND	OR	NAND	**XOR**	EQ
lo/lo	OFF	OFF	**ON**	**OFF**	**ON**
lo/hi	OFF	**ON**	**ON**	**ON**	OFF
hi/lo	OFF	**ON**	**ON**	**ON**	OFF
hi/hi	**ON**	**ON**	OFF	**OFF**	**ON**

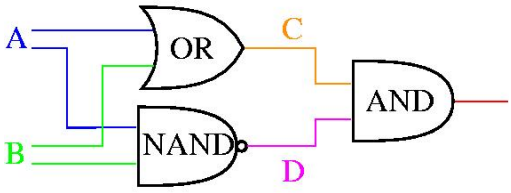
cannot be implemented by overlapping A and B sites



[cf: linear perceptron (Minsky '69)]

XOR(A,B) = (A **OR** B) **AND NOT**(A **AND** B)

Gene cascade



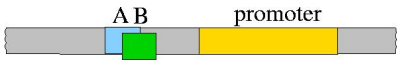
problems:
- need a gene for each intermediate result
- multiple rounds of gene expression: noise + delay
- synchronization difficult
- amplification nontrivial

10

More complex control function, e.g., **XOR** ?

A/B	AND	OR	NAND	**XOR**	EQ
lo/lo	OFF	OFF	ON	**OFF**	ON
lo/hi	OFF	ON	ON	ON	OFF
hi/lo	OFF	ON	ON	ON	OFF
hi/hi	ON	ON	OFF	**OFF**	ON

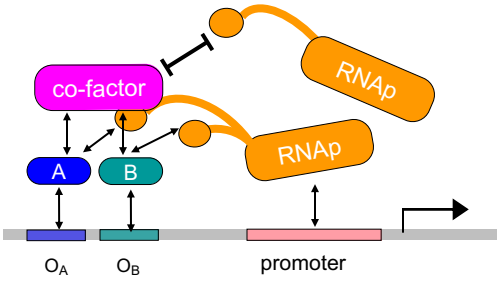
cannot be implemented by overlapping A and B sites



[cf: linear perceptron (Minsky '69)]

XOR(A,B) = (A **OR** B) **AND NOT**(A **AND** B)

Allosteric or co-factor mediated



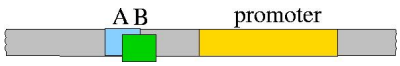
problem:
- need dedicated component
- lose combinatorial control
  e.g., can't implement **AND** elsewhere

11

More complex control function, e.g., **XOR** ?

A/B	AND	OR	NAND	**XOR**	EQ
lo/lo	OFF	OFF	ON	**OFF**	ON
lo/hi	OFF	ON	ON	ON	OFF
hi/lo	OFF	ON	ON	ON	OFF
hi/hi	ON	ON	OFF	**OFF**	ON

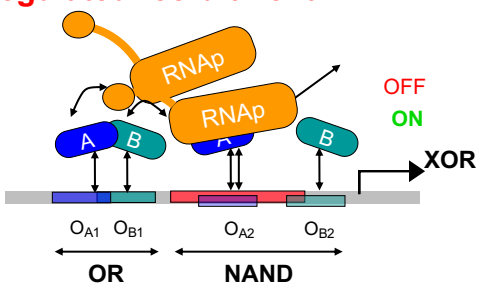
cannot be implemented by overlapping A and B sites



[cf: linear perceptron (Minsky '69)]

XOR(A,B) = (A **OR** B) **AND NOT**(A **AND** B)

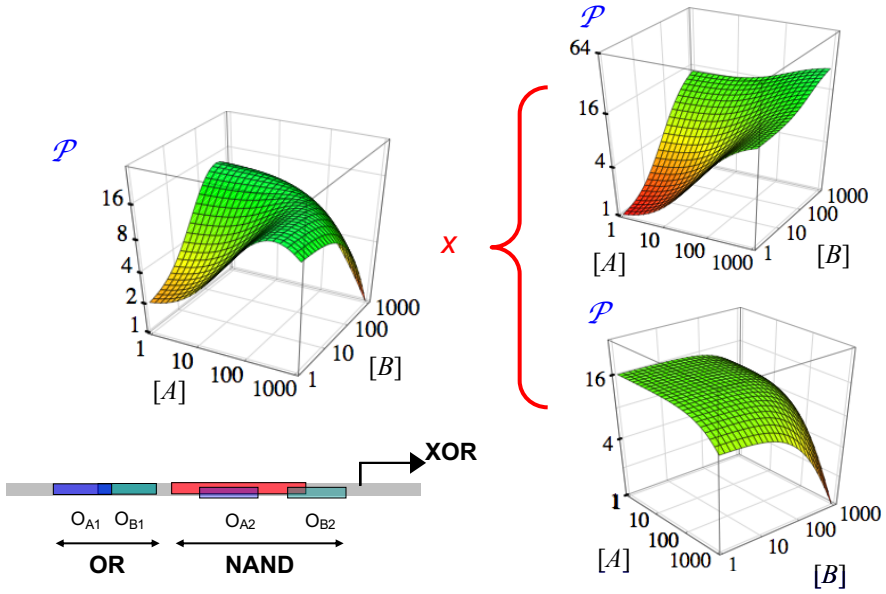
Regulated recruitment



➡ integrates **OR** and **NAND** into a single regulatory region
➡ no need for special proteins
➡ modular and evolvable

12

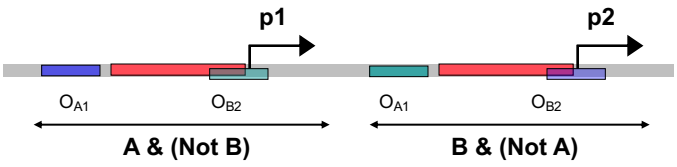
More complex control function, e.g., **XOR** ?



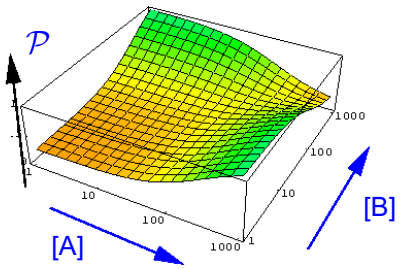
13

More complex control function, e.g., **XOR** ?

➔ alternative implementations exist, e.g.,

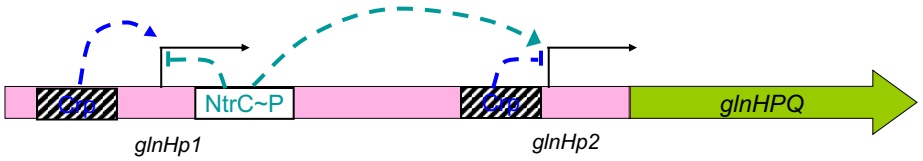


gene expression $\propto \mathcal{P}_{p1}([A],[B]) + \mathcal{P}_{p2}([A],[B]) \simeq$ XOR



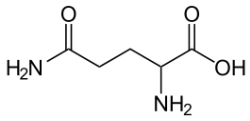
14

XOR promoter in bacteria?



Crp: senses carbon shortage
NtrC~P: senses nitrogen shortage
function of *glnHPQ* gene product: Glutamine transporter

Glutamine ($C_5H_{10}N_2O_3$)



use as carbon source under carbon shortage

use as nitrogen source under nitrogen shortage

➔ need quantitative characterization of such promoters
➔ design of synthetic promoters using exogenous regulators

15

EQ gate?

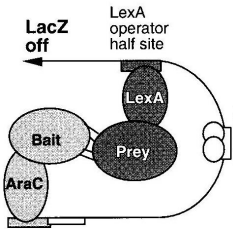
A/B	**EQ**
lo/lo	ON
lo/hi	OFF
hi/lo	OFF
hi/hi	ON

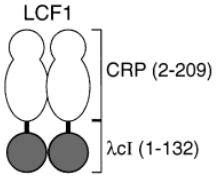
- strong promoter
- need **multiple** ways of repression
  A (low) B (high) or A (high) B (low)

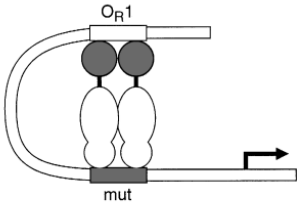
→ competitive promoter binding awkward due to limited promoter size (dead end)

→ possible solution: interaction at a distance
e.g., DNA-looping via dimers (AraC, GalR, MelR,…)

heterodimers: [A. Hochschild et al]

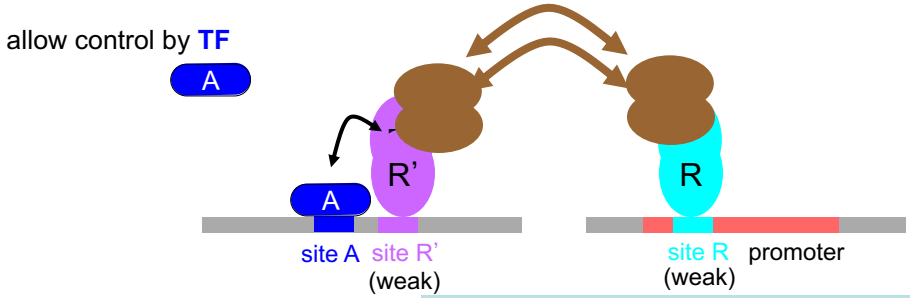




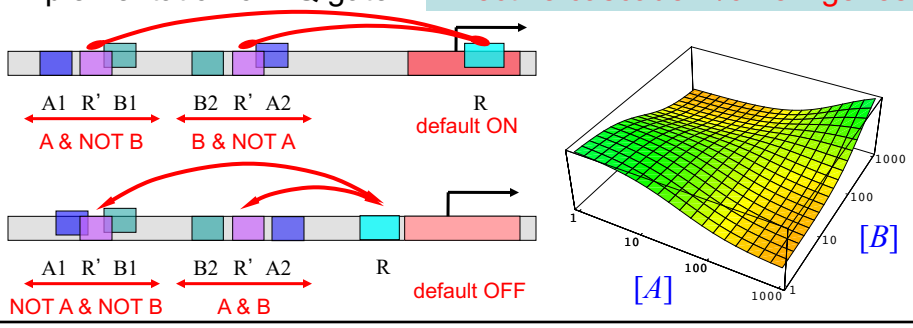


16

E.g., distal repression by heterodimer pair R and R'

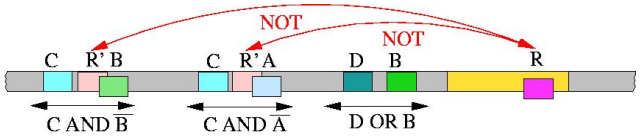


Implementation of **EQ** gate: Effective cascade w/o new genes!



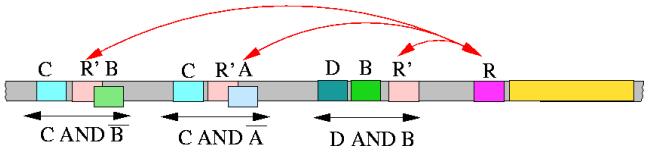
17

Generalized control architecture w/ multiple TF's



Output = […**OR**…**OR**…] **&** **NOT**[…**&**…**&**…] **&** **NOT**[…**&**…**&**…]

phenotype: dominant repression — Conjunctive Normal Form



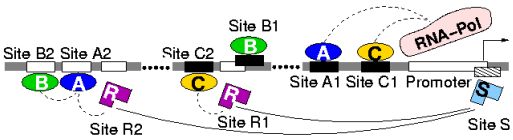
Output = […**&**…**&**…] **OR** […**&**…**&**…] **OR** […**&**…**&**…]

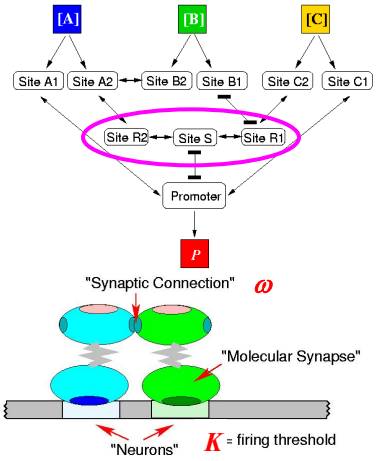
phenotype: enhancer autonomy — Disjunctive Normal Form

→ all logic functions reducible to minimal CNF or DNF
→ schemata for constructing arbitrary regulatory logic functions (programmable molecular computer!)

18

Molecular computer ⟷ "Neural network"





TF concentration (n)	input
TF binding sites (j)	neurons
promoter activity (P)	output
binding strength (K_j)	firing threshold
TF-TF interaction ($\omega_{i,j}$)	synapse

➡ single node of GRN is already a network!

- symmetric interaction
- "hidden units" (cofactors)

} **Molecular Boltzmann machine!**

"learning" ←➔evolution of regulatory sequences

19

Summarize:

A large variety of control functions $\mathcal{P}([TF_j])$ may be implementable via appropriate choices of $\{K_j, \omega_{i,j}\}$, i.e., via regulatory sequences **alone**
= programmable molecular computer

➔ **synthesize** "desired" transcriptional "logic gates"
➔ "**breed**" regulatory sequences to implement desired control functions

Potential application: cell-specific gene expression profiling

	cell A	cell B	cell C	cell X
gene 1	+	-	+	-
gene 2	-	+	+	+
gene 3	-	-	+	+
...				

→ cell type discrimination: use multiple traits
→ cell X revealed by a reporter gene
driven by designed regulatory sequences
→ cell X eliminated by promoter activating apoptosis
→ targeted delivery not required (~ smart bomb!)

20

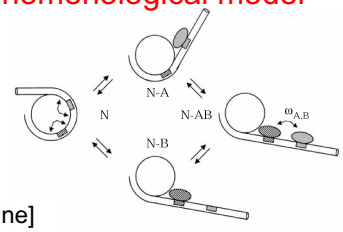
Ingredients for complex transcription control

- programmable protein-DNA interaction
- weak, glue-like interaction between nearby proteins
- long-distance activation/repression
- insulation of gene regulatory control

Eukaryotes: formalism as phenomenological model

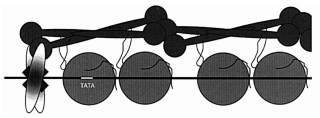
- generic cooperative interaction mediated by **nucleosomes** [Polach & Widom, '96]

➡ physical attraction not necessary



- **short-range "quenching"** [Arnosti, Levine]

- distal repression via recruitment of **chromatin modification** agents



- **insulating elements**: crucial for minimizing cross talk

similar ingredients but superior implementation platform

21