SBM 2

28-09-09

Bintu et al. (Hwa group)

Transcriptional regulation by the numbers: applications

next 5 classes

- Bintu 1: actual modeling and calibration of simple TF systems

- Bintu 2: the thermodynamic model (theory of the above)

- Elowitz*: combinatorial synthesis of Bintu-like promoters

- model-driven design, Jim Collins

- Marchisio & Stelling (Bioinformatics 24, 1903 - 2008): compositional building of bio-brick models

... and then we do detailed modeling for another 5 classes

(prok) transcriptional logic - Terminology

- The activity of a gene might be regulated by that of other genes with products called transcription factors (TFs)

- depending on concentrations and other parameters TFs bind more or less regions on DNA strands upstream of a gene: called promoters

- specific TF targets on DNA are called operators (subsets of promoter, terminology varies).

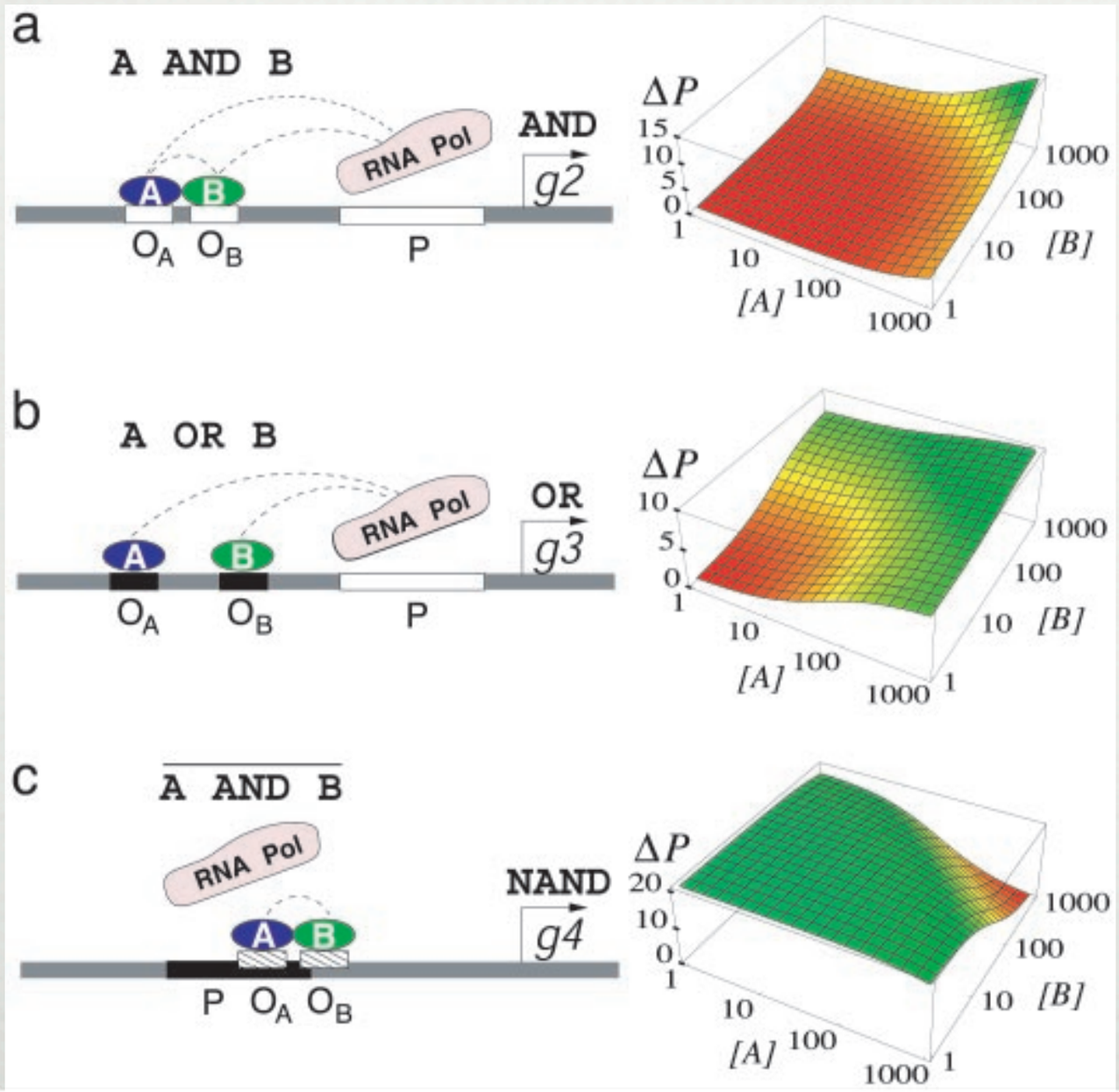
(prok) transcriptional logic - Terminology 2

- TFs can pair with small molecules: CRP with cAMP, lacR with allolactose, MelR with melibiose, etc; usually this activates activators, and represses repressors (monotony)

- once in place TFs interact with the RNA polymerase (RNAP) determining the promoter activity (roughly the transcription rate, but transcription is complex in detail/more later)

- the TF active form is said to be induced, and the chemical partner is called the inducer

let us look at a few examples (theoretic) - binary "gates"

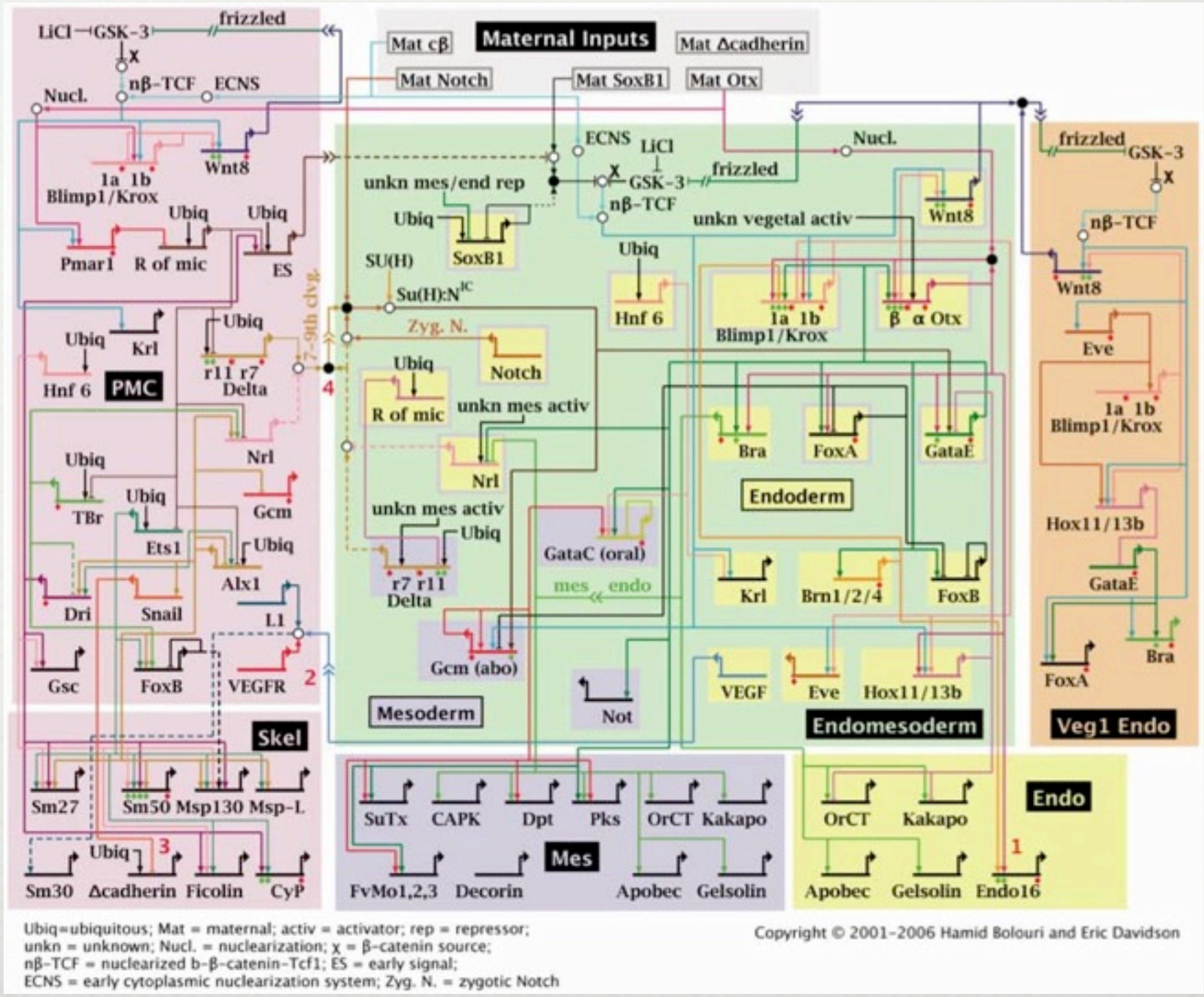


Buchler et al, PNAS 100(9)

deciding whether AND or OR needs a discretisation - more later

caveat: these simple parts can be combined!

- can become hugely complex in euk. development (Davidson's et al on sea urchin)



this class ...

various combinations of activators

- with explicit formulas for transcriptional boost, F

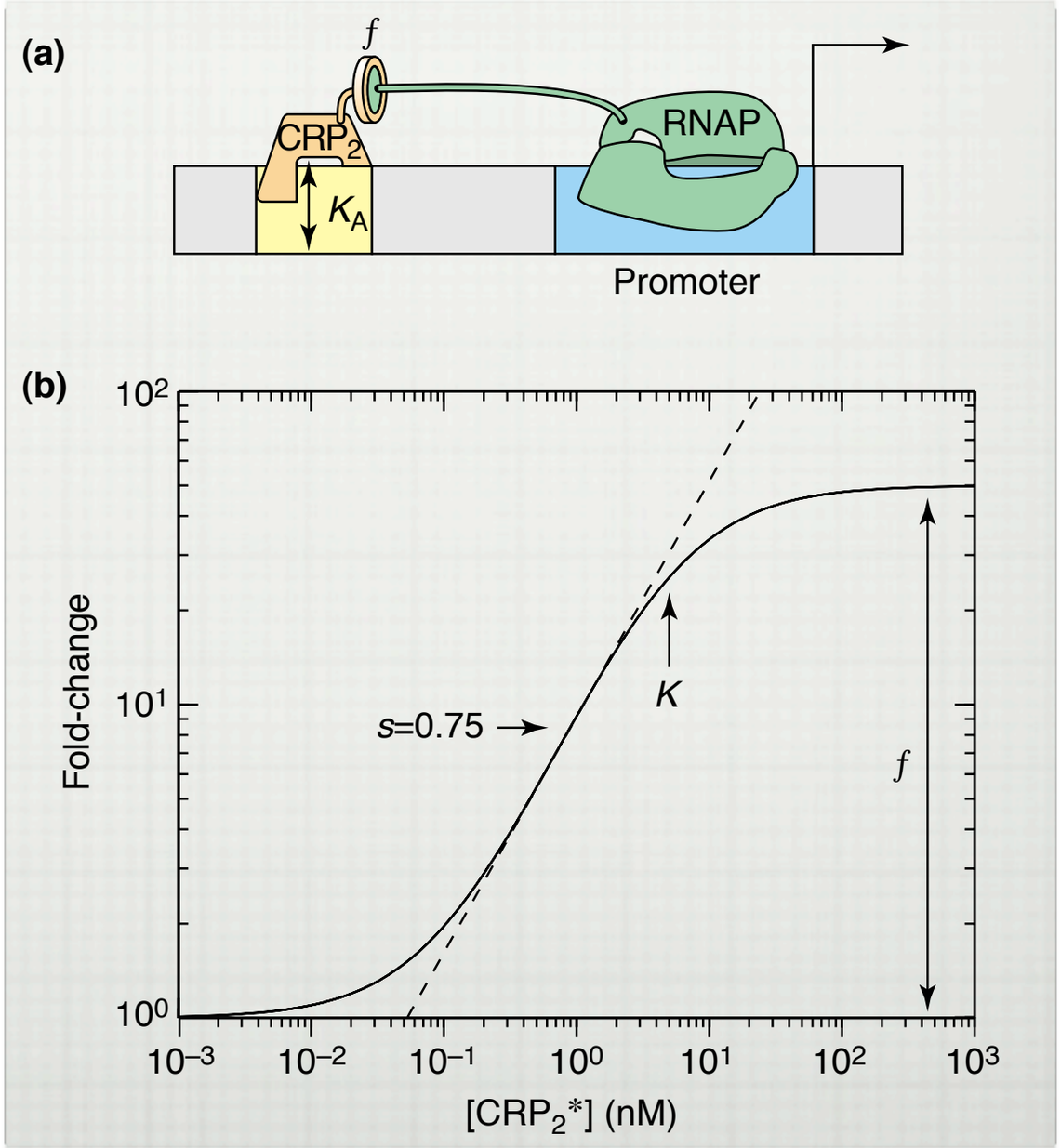
1. simple activation: CRP/P_{lac}

$$F = (1 + f\,[A]/K_A)/(1 + [A]/K_A)$$

a ratio, ie a multiplicative factor

log-log plot:
$[CRP_2^*]$ = active (dimer) TF
fold-change F

params:
K_A = eq dissociation constant
f = cooperation > 1
s = sensitivity/steepness



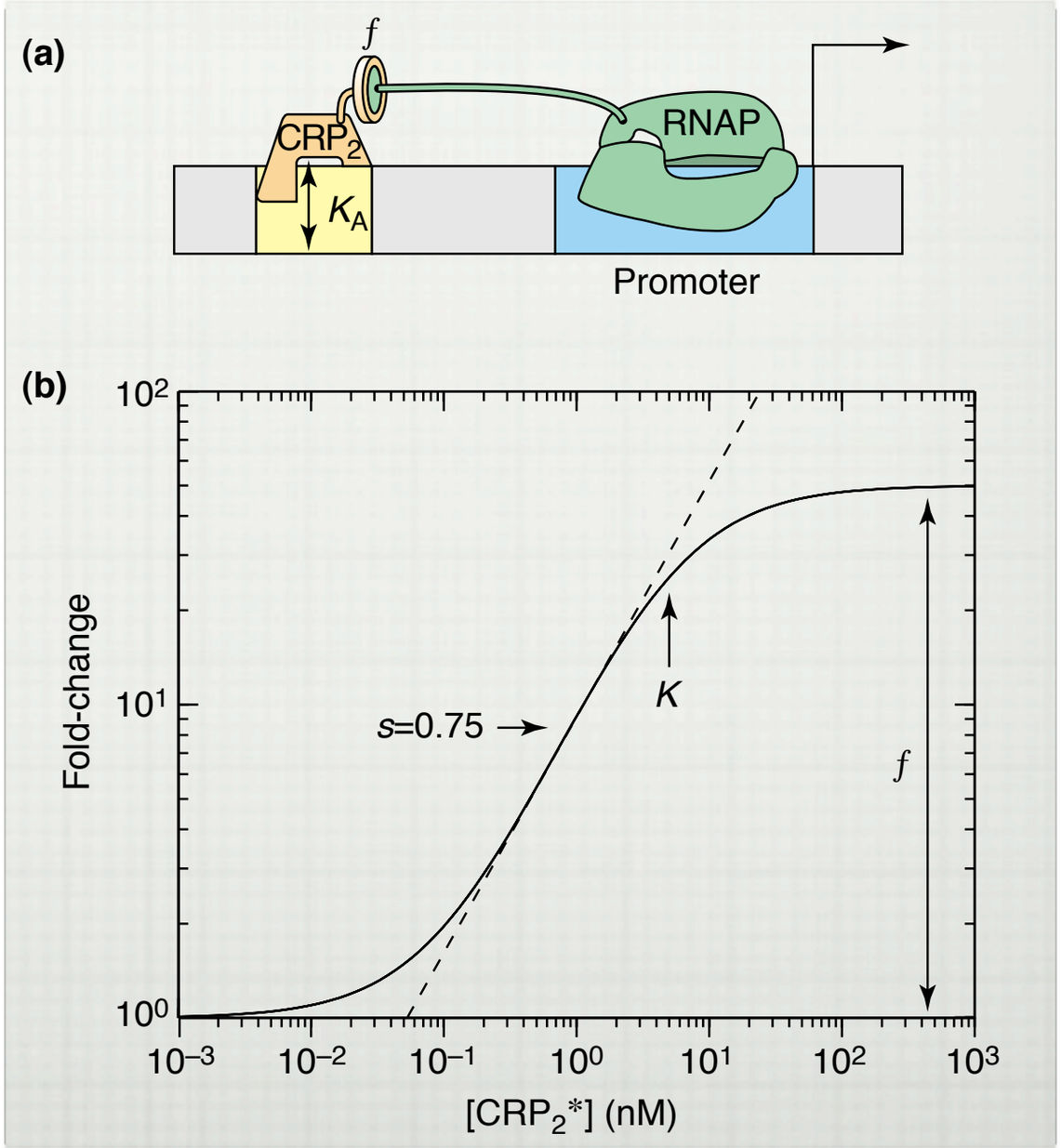
1. simple activation: CRP/P_{lac}

$$F = (1 + f\,[A]/K_A)/(1 + [A]/K_A)$$

a ratio, ie a multiplicative factor

log-log plot:
$[CRP_2^*]$ = active (dimer) TF
fold-change F

params:
K_A = eq dissociation constant
f = cooperation > 1
s = sensitivity/steepness

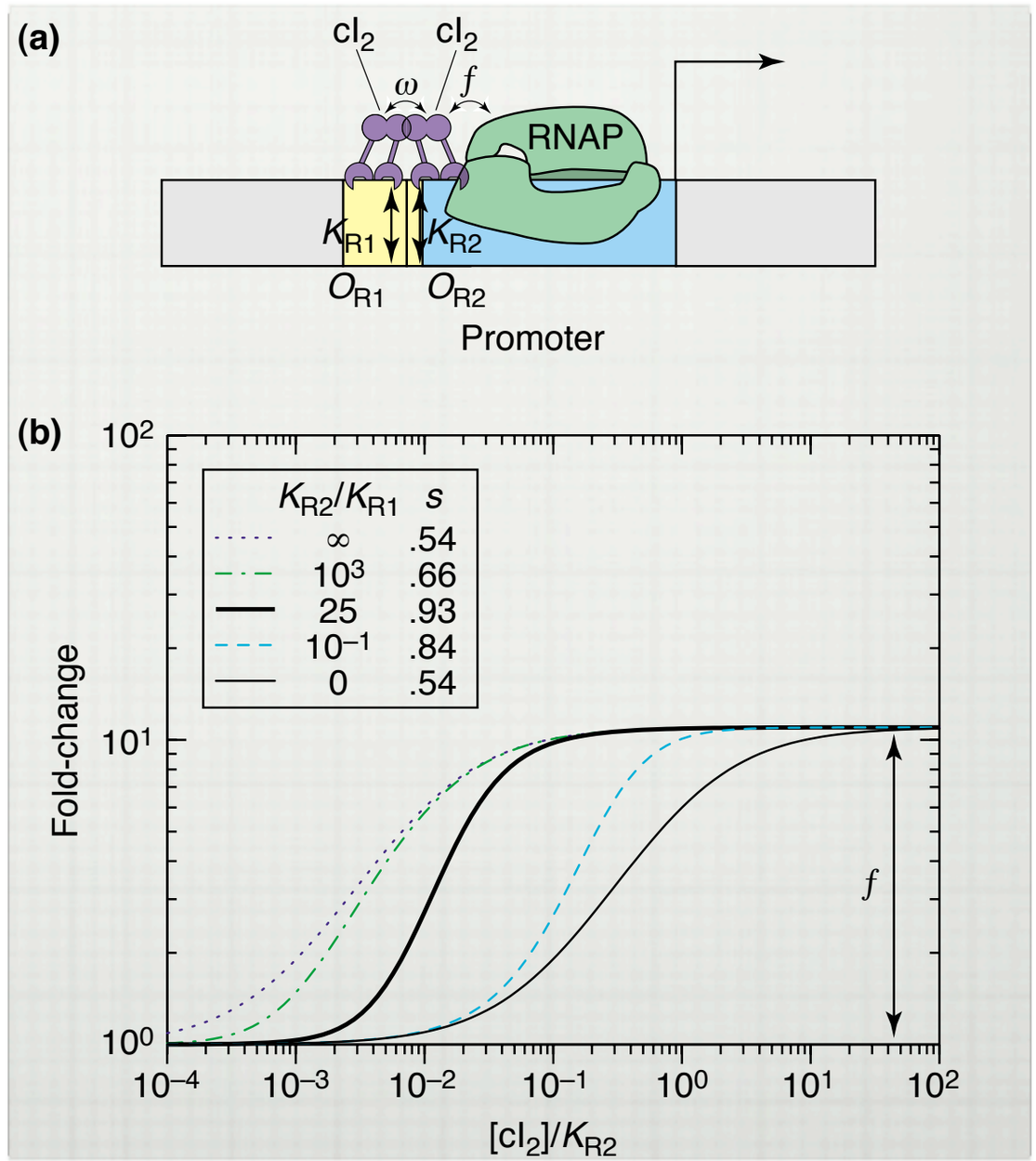


2. activation and helper: cI/P_{RM}

$$F = \frac{(1 + [H]/K_H + f\,[A]/K_A + f\,\Omega\,[H][A]/K_H K_A)}{(1 + [H]/K_H + [A]/K_A + \Omega\,[H]/K_{H^*}[A]/K_A)}$$

log-log plot of F vs
$[cI_2]$ = (dimer) TF
for various values of K_{R2}/K_{R1}

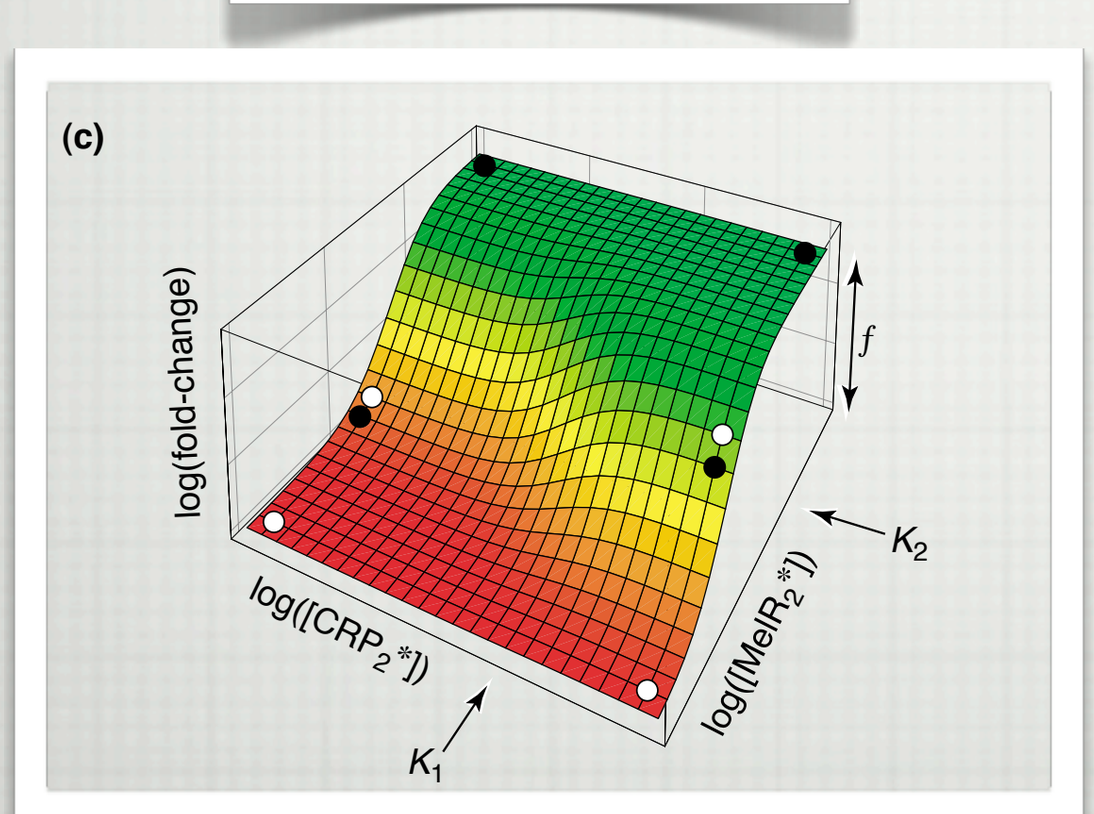
parameters: K_{R2}, K_{R1}, f, omega
sensitivity s

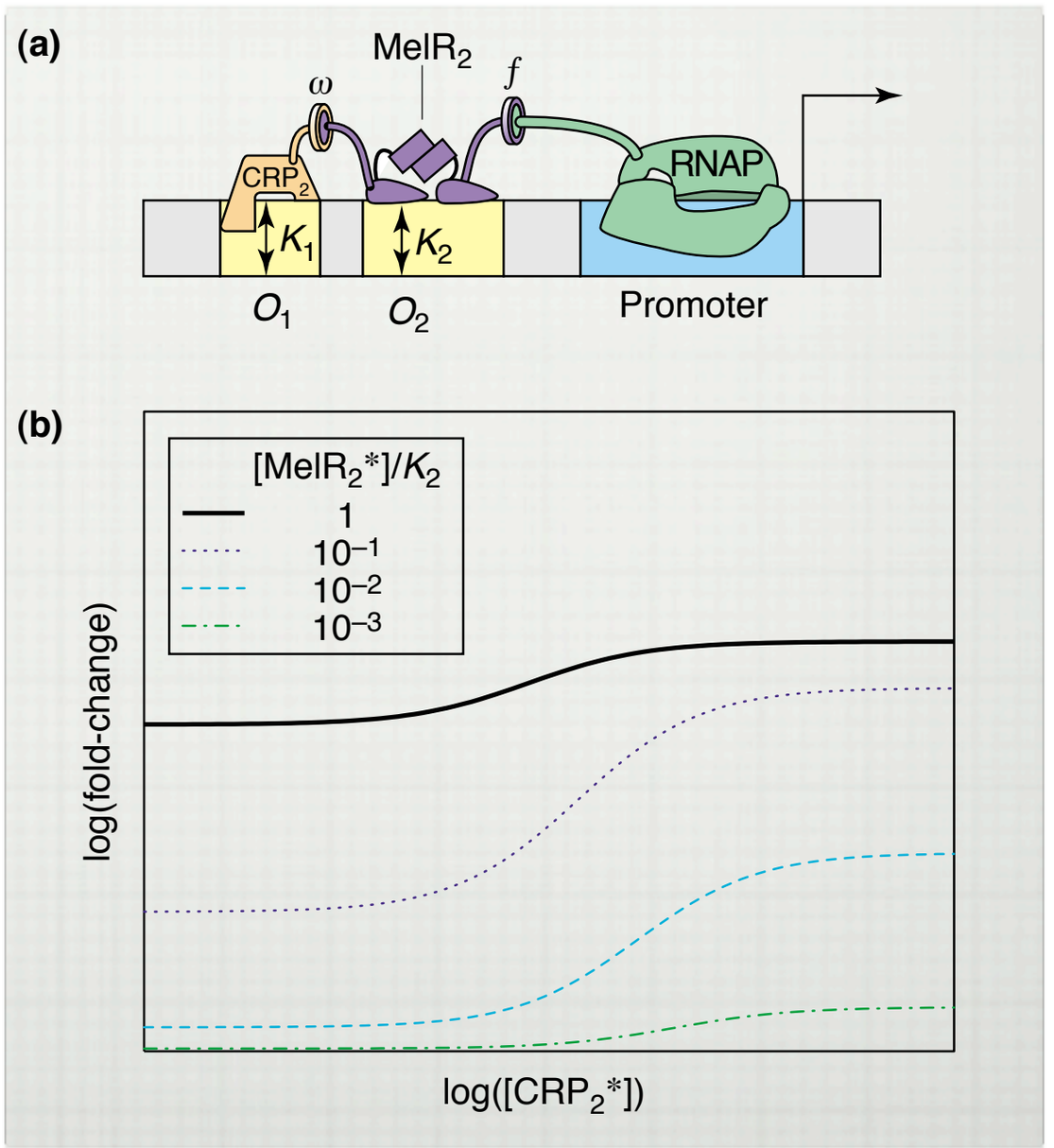


3. co-activation: CRP,MelR/P_{RM}

F and params same as in 2.

log-log plot of F vs $[cI_2]$ and $[MelR_2]$

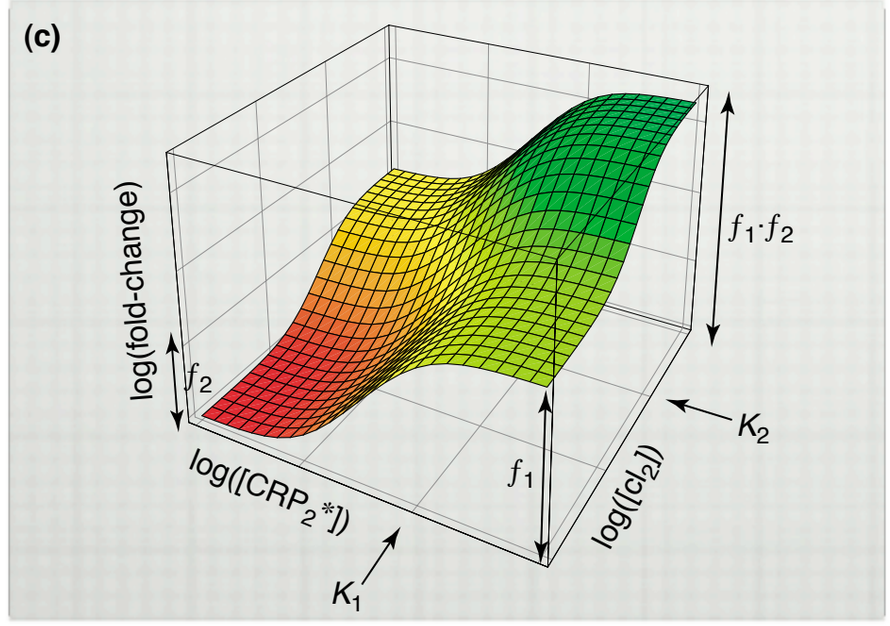


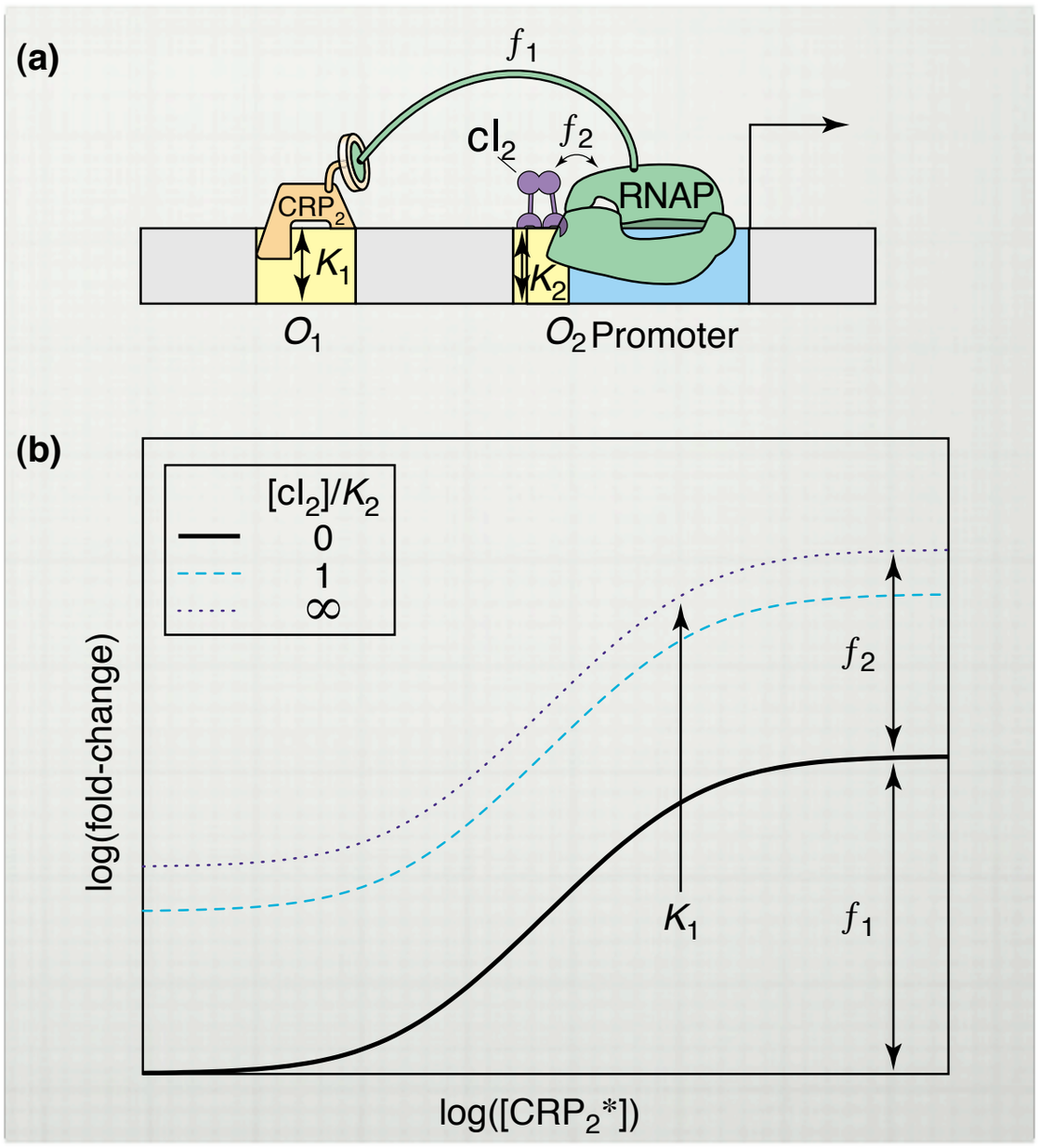


4. dual activation: CRP,cI/P

$$F = \frac{(1 + f_1[A_1]/K_1 + f_2[A_2]/K_2 + \Omega f_1 f_2 [A_1][A_2]/K_1K_2)}{(1 + [A_1]/K_1 + [A_2]/K_2 + \Omega[A_1][A_2]/K_1K_2)}$$

here $\Omega=1$, no direct interaction between A_1 and A_2





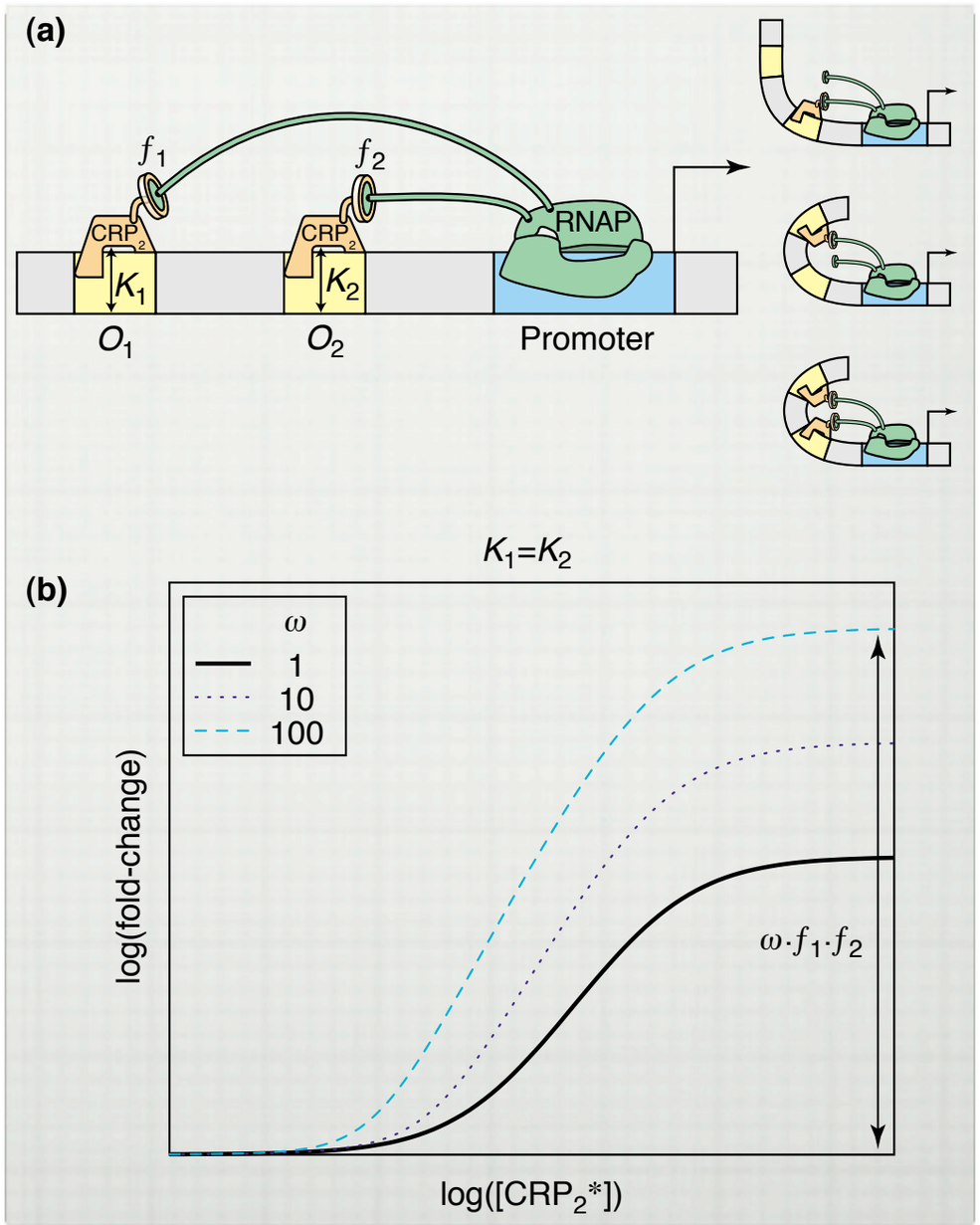
5. dual activation with direct cooperation: CRP,cI/P

CRP is known to activate at different distances

same F as in 4.

$$F = (1 + f_1[A_1]/K_1 + f_2[A_2]/K_2 + \Omega f_1 f_2 [A_1][A_2]/K_1K_2) / (1 + [A_1]/K_1 + [A_2]/K_2 + \Omega[A_1][A_2]/K_1K_2)$$

here $\Omega > 1$, there is direct interaction between A_1 and A_2: looping DNA



next class ...

- various combinations of repressors

- thermo model for the derivation of the expressions for F

deliberately simple promoters

Elowitz et al. Mol Syst Biol. 2007 3:145

- ☐ The paper reports the synthesis of about 200 promoters glued to a reporter gene;
- ☐ the obtained DNA constructs can be seen as binary functions (most have 2 operators so 2 TF they can interact with).
- ☐ The constructs are simple promoter architectures, a priori no TF-TF contacts and no operator overlap.
- ☐ Constructs are classified in an original way as real-valued binary functions and then sequenced (why?)

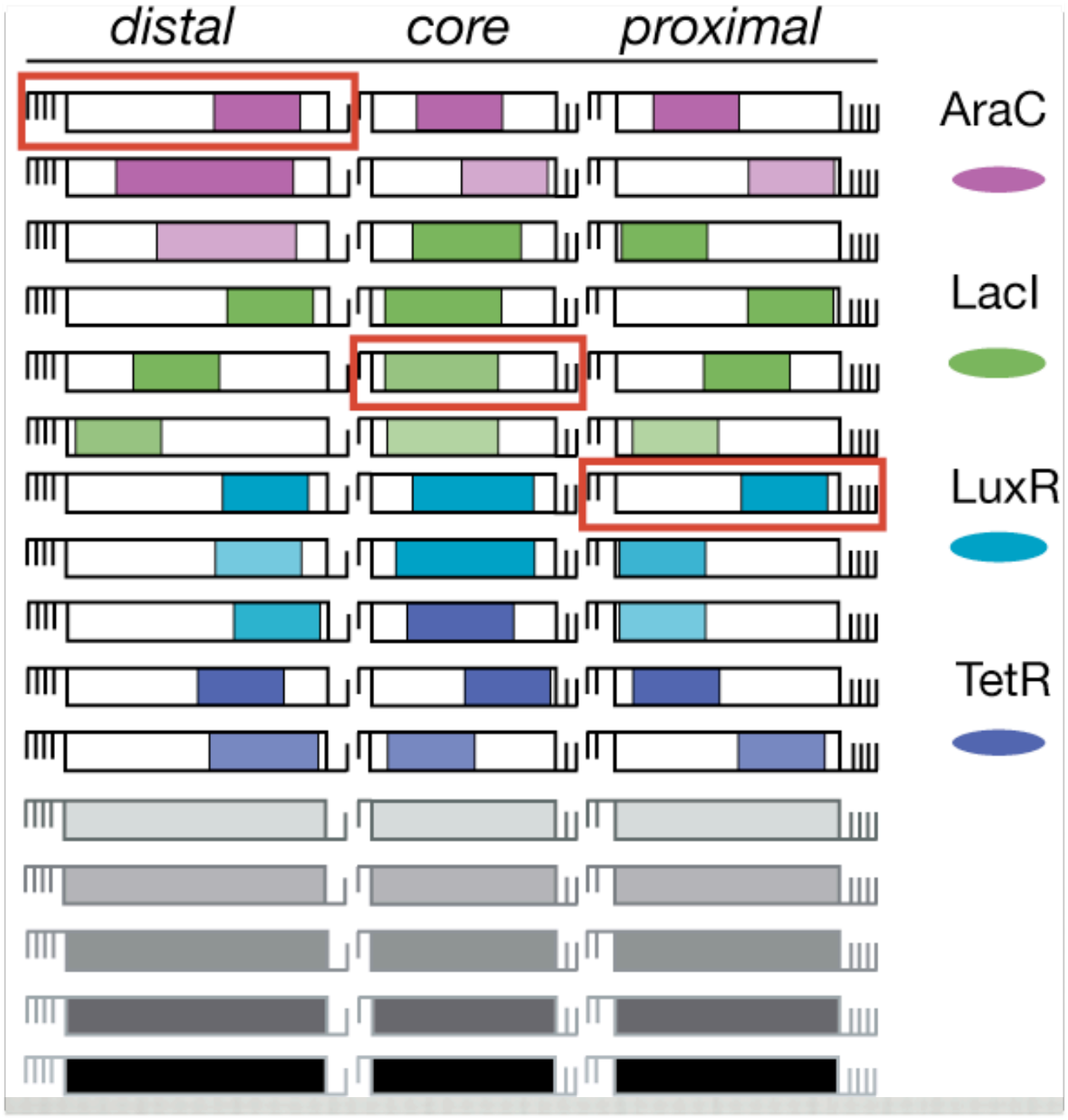
remarks

- A self-documenting automated bio-brick factory!
- yet output of a promoter::gene not a Boolean valued function of the concentrations of its TF/inputs (lac promoter has 4 output levels):
- could take "low" value of a few molecules per bacterium (1 nM)
- "high" value 1,000 molecules per bacterium (1M)

a combinatorial library of random promoter architectures

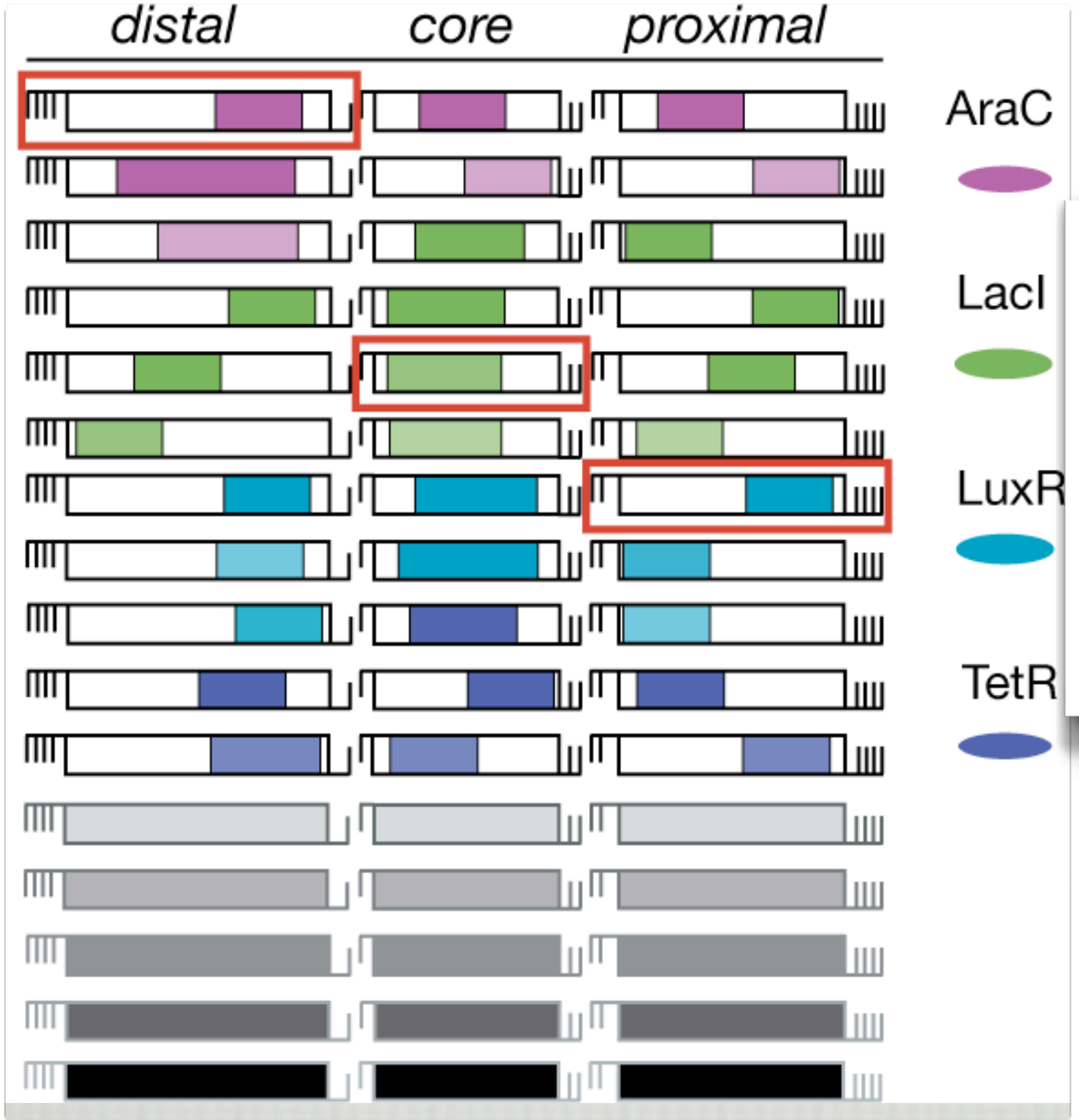
- [ ] 4096 of which 288 sequenced cassettes:
- [ ] 217 unique
- [ ] of which 27 binary (twofold response under 2 TFs)
- [ ] promoter = distal::core::proximal, device = promoter::G-luciferase
- [ ] TFs= AraC, LuxR (activators) -activated by Lara, VAI
- [ ] TetR, lacI (repressors) -inactivated by aTc, IPTG

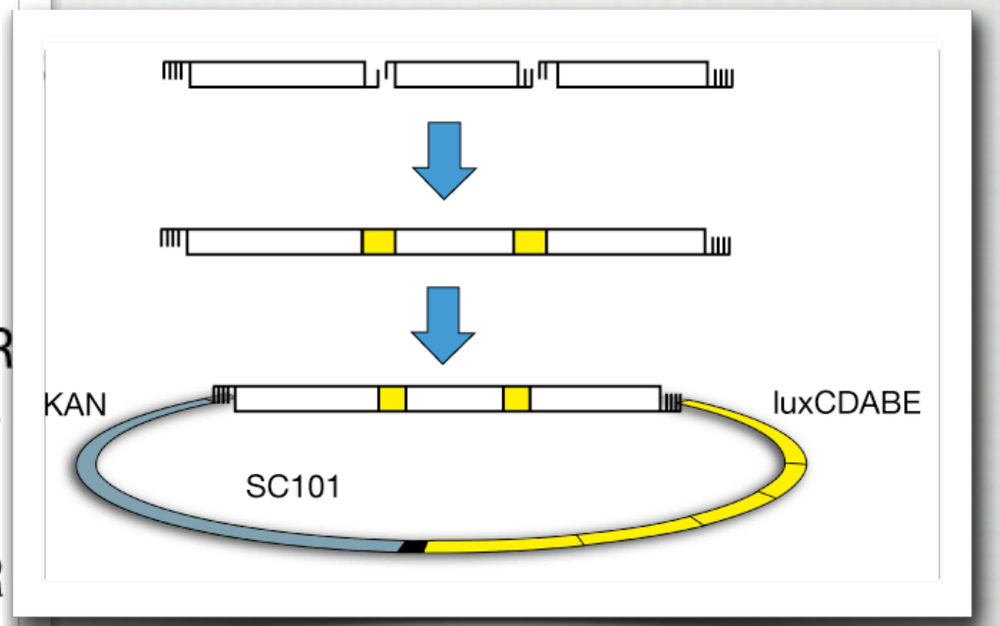
The library - sequence work



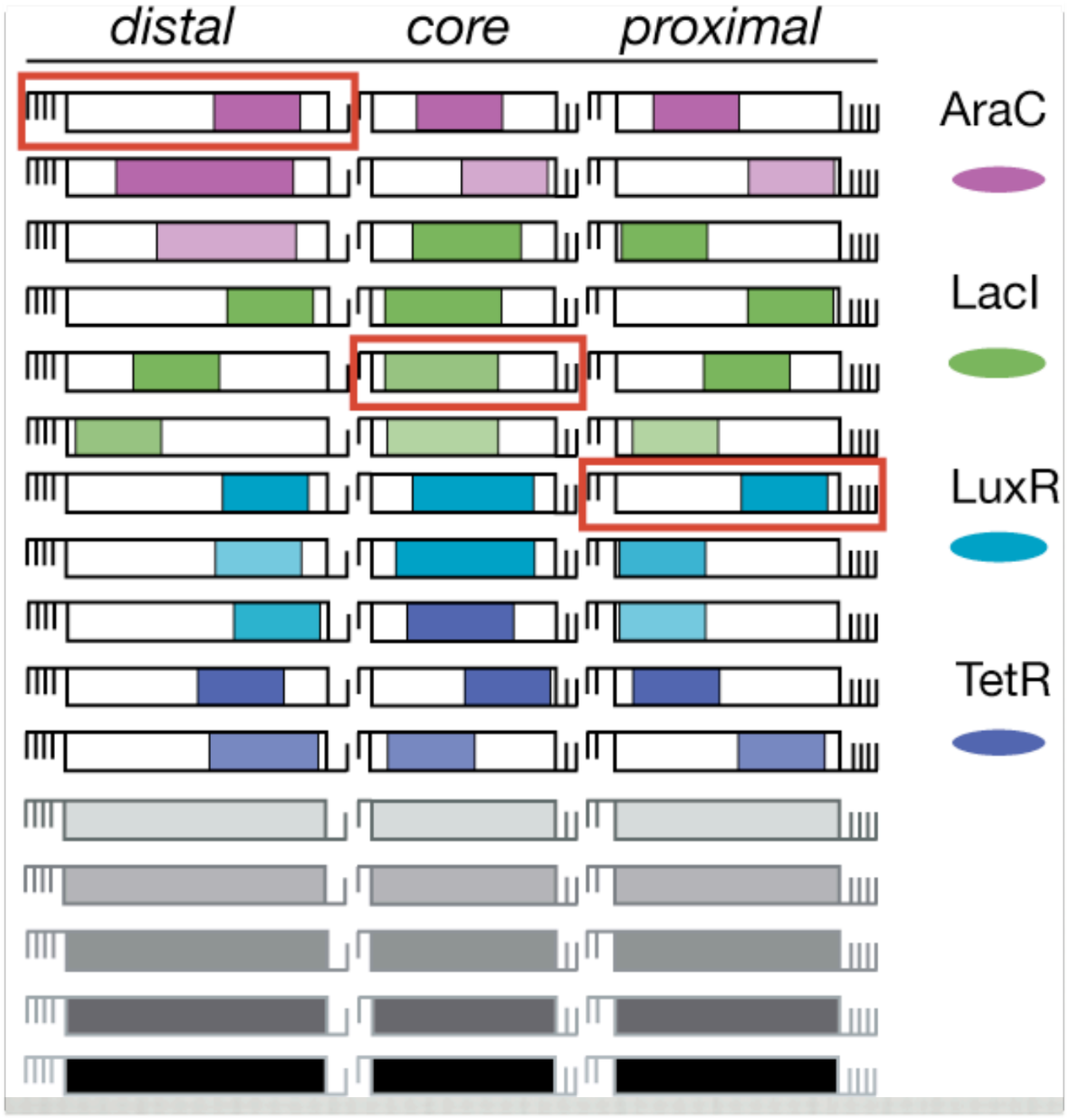
The library - sequence work

The library - sequence work

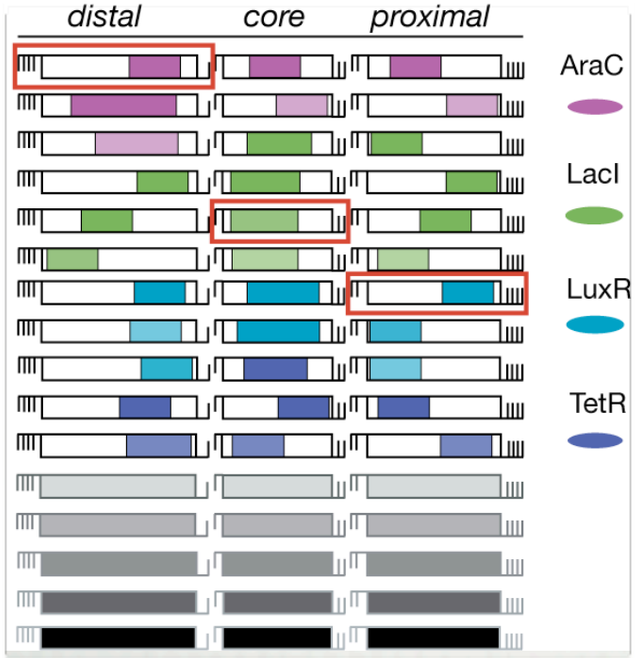




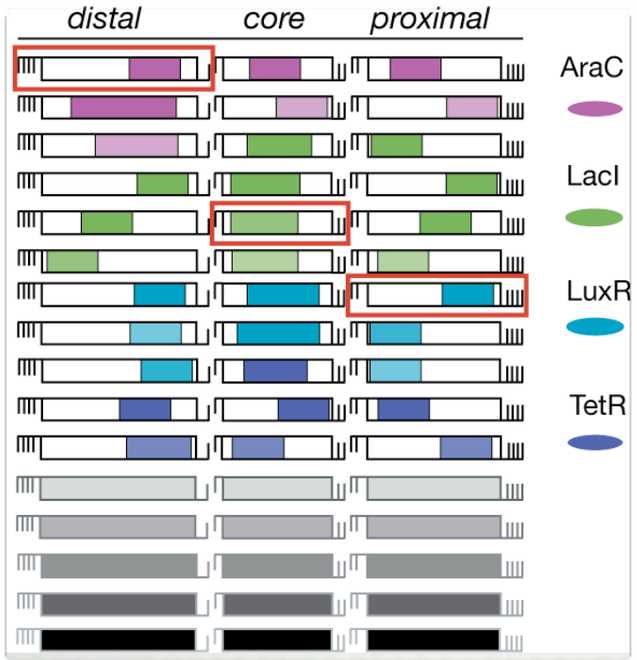
The library - sequence work



The library - sequence work



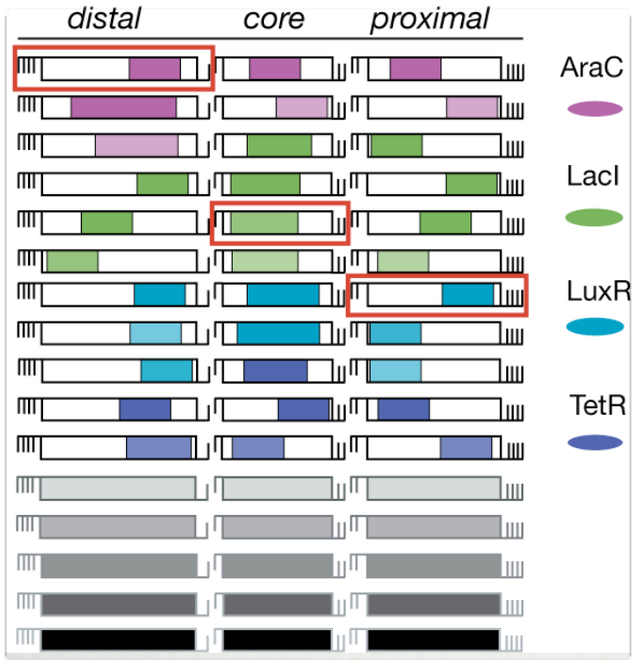
The library - sequence work




```
tcgagtacaacgtcgtgttagctgcctttagcaattttatccatagacttgtgagcgctcacaatttataattcgtgcaatTtttaacctgtaggatcgtacaggtg
    catgttgcagcacaatcgacggaaaatcgttaaaatataggtatctgaacactcgcgagtgttaaatattaagcacgttaAaaatttggacatcctagcatgtccacctag
```

AraC I1 site -35 LacI Os site -10 +1 LuxR box

The library - sequence work




```
tcgagtacaacgtcgtgttagctgcctttagcaattttatccatagacttgtgagcgctcacaatttataattcgtgcaatTtttaacctgtaggatcgtacaggtg
    catgttgcagcacaatcgacggaaaatcgttaaaatagGtatctgaacactcgcgagtgttaaatattaagcacgttaAaaatttggacatcctagcatgtccacctag
```

AraC I1 site -35 LacI Os site -10 +1 LuxR box

288/4096 SEQUENCED

remarks

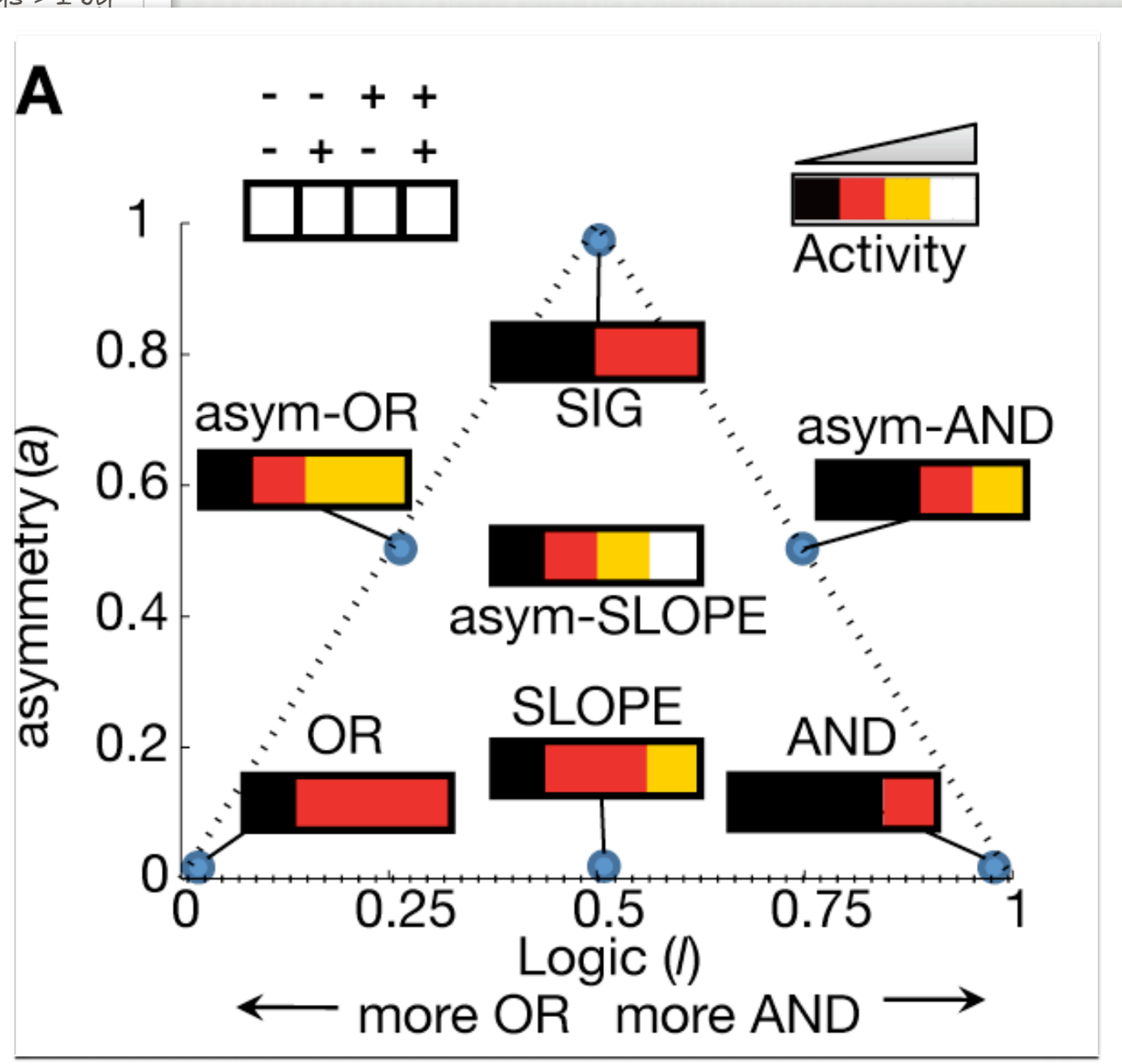
- [ ] The search is limited to neighbourhood of existing operators; it is really variation
- [ ] The observation is discretized (how robust is that?); who is listening to the outputs intervals;
- [ ] lacks a composition/impedance study; endogenousness?
- [ ] specificity, name space: possible to engineer chemical/TF specificity? wrt what is this complete?

typology (construction of the phenotype)

- ☐ regulatory range: exp-on/EXP-off [caveat: this is always >1 by def]
- ☐ logic type: from or l=0, to and l=1
- ☐ symmetry: from a=0 (complete symmetry) to A=1 (dependency in only 1 input) [works only for binary functions]

typology (construction of the phenotype)

- regulatory range: exp-on/EXP-off [caveat: this is always >1 by def]
- logic type: from or l=0, to and l=1
- symmetry: from a=0 (complete symmetry) to A=1 (d in only 1 input) [works only for binary functions]



typology (2)

- The level of TF is controlled undirectly by chemicals, and repressors are repressed, while activators are activated.
- So whatever the construct is, the attached function is monotonic increasing.
- The classification scheme -writing b1<b2<b3<b4 for the increasing sequence of responses (by monotony b1, and b4 are obtained for 00 and 11 inputs)- is:
- - the dynamic range r= log(b4/b1) in log scale
- - the asymmetry a=log(b3/b2)/r the b3 to b2 gap normalised to r so in 0 (fully symmetric) to 1 (unary function)
- - the and-ity l= (log(b4) -1/2(log(b3) + log(b2)))/r which is 0 if b4=b3=b2, 1 (an OR) if b3=b2=b1 (an AND)

model of RR promoter activity under dual repression (Bintu)

- [ ] The r, a, and l trinity above can be defined in terms of the micro-trinity c1, c2, omega measuring the joint activity of a pair of repressors

- [ ] $P(R1,R2)=A/(c1\ R1+c2\ R2+omega\ c1\ c2\ R1\ R2)$

- [ ] A max promoter activity

- [ ] c1, c2 TF efficiencies (at excluding RNApol)

- [ ] omega=coooperation (>1)

Looking for Mr Nice component

- [ ] computational models of transcription (eg "Transcriptional regulation by the numbers" Curr Opin Genet Dev -2005)
- [ ] evolution driven design (eg "Directed evolution of a genetic circuit" PNAS 2002)
- [ ] combinatorial approach (this paper)