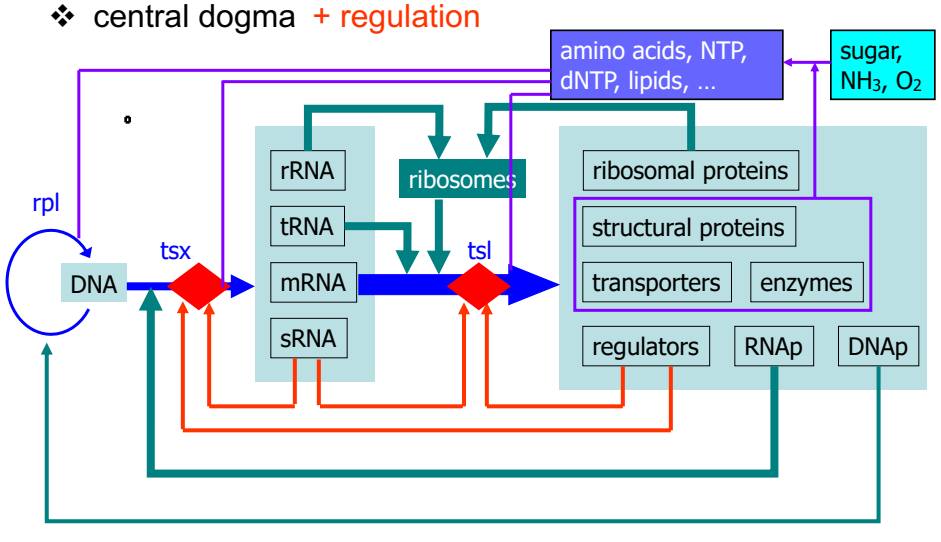
❖ central dogma + regulation

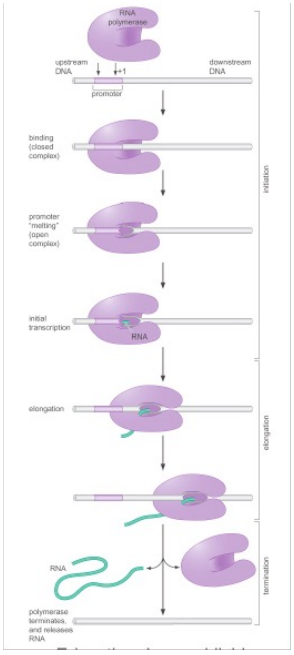


- tsx initiation control by transcription factors (TF)
- tsl initiation control by sRNA and RNA-binding proteins
- tsx termination control by sRNA and anti-terminators
- control of mRNA and protein degradation

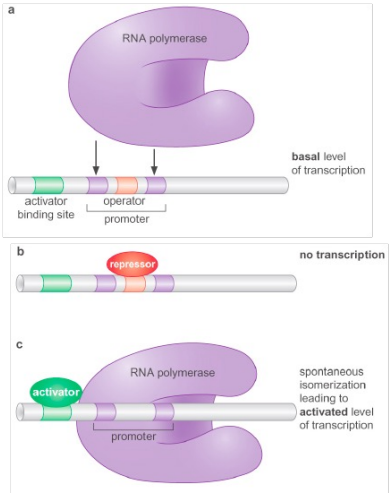
} coupled to environmental signals

1

transcriptional initiation and termination



tsx init control by activators, repressors



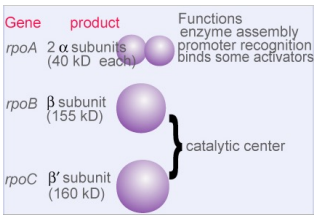
2

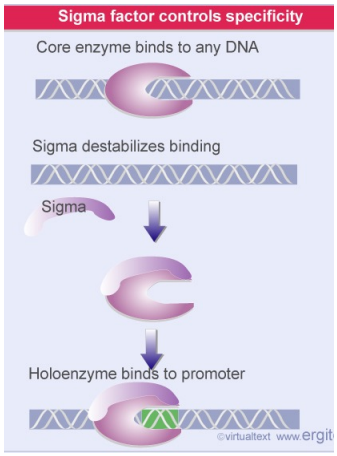
Topic 2: Transcription Initiation Control

A. Mechanisms of tsx initiation in bacteria

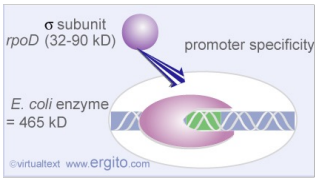
1. Components:

• core enzymes of RNA polymerase:





• sigma factor:



3

• *E. coli* has 6 different σ-factors

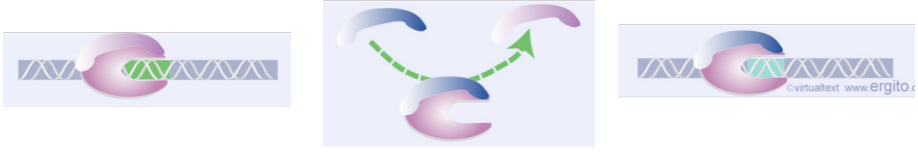
70-80% of genes; further regulated by TFs

Gene	Factor	Use
rpoD	σ^{70}	general
rpoS	σ^{S}	stress
rpoH	σ^{32}	heat shock
rpoE	σ^{E}	heat shock
rpoN	σ^{54}	nitrogen
fliA	σ^{28} (σ^{F})	flagellar

• core promoter recognition sequences

Factor	-35 Sequence	Separation	-10 Sequence
σ^{70}	TTGACA	16-18 bp	TATAAT
σ^{32}	CCCTTGAA	13-15 bp	CCCGATNT
σ^{54}	CTGGNA	6 bp	TTGCA
σ^{28} (σ^{F})	CTAAA	15 bp	GCCGATAA

• substitution of σ-factors ➔ recognize different set of promoters



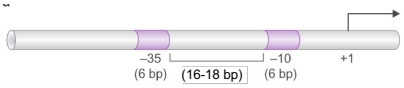
• *B. subtilis* has ~20 σ-factors (include sporulation, competence, …)

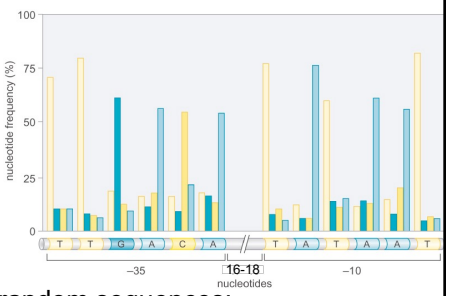
• generally, more complex the life style of organism, more sigma factors

4

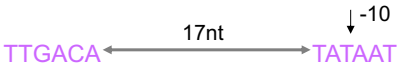
• core promoter recognized by σ^{70}-factor:

canonical promoter has fuzzy motif





• consensus sequence:



occurrence of fuzzy promoter motifs in random sequences:

• 3 out 6 matches in -35 region $\binom{6}{3} \times 0.25^3 \times 0.75^3 \approx 13\%$

• degeneracy in spacing (16-18bp) $\times 3$

• 4 out 6 matches in -10 region $\binom{6}{4} \times 0.25^2 \times 0.75^4 \approx 3\%$

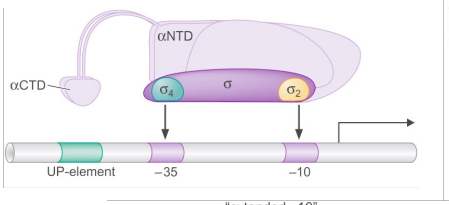
➔ at given position in the genome, motif occurrence probability

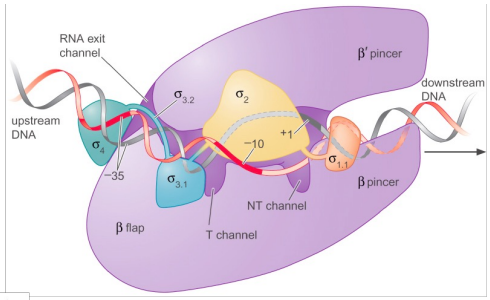
$\approx 13\% \times 3 \times 3.3\% \approx 1.3\%$

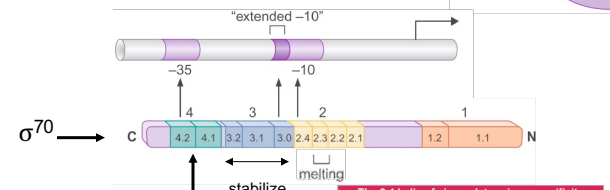
(one occurrence every ~ 80bp, i.e., everywhere!)

6

2. RNAp-promoter interaction

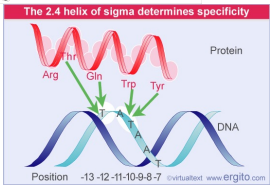


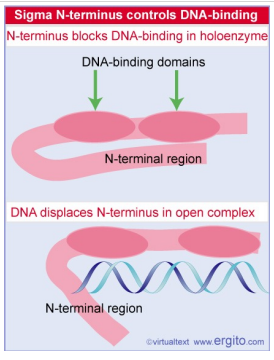




stabilize open complex

bends DNA at -35; facilitates interaction with upstream activators





8

- avail. RNAp conc ~ 30 nM [McClure, 1983]
  ≈ 0.5 ~ 1 μM [Klumpp & Hwa, PNAS 2008]
  ≈ 1 μM (glc medium) [Balakrishnan et al, Science 2022]

- typical binding constants and rates

(rapid equilibria) (slow)

$$\mathrm{R + P} \underset{}{\overset{K_1}{\rightleftharpoons}} \mathrm{I_1} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{I_2} \overset{K_3}{\rightleftharpoons} \mathrm{RP_O}$$

(slow) (rapid equilibria)

$$K_1 / K_{ns} = 10^{-3} \sim 1 \quad \Rightarrow \quad \widetilde{K}_1 = \sum_{n \neq j}^{N} e^{(G_1 - G_{ns})/kT} = \frac{N \cdot K_1}{K_{ns}} = 10^4 \sim 10^7 \text{ nM}$$

  – promoter binding typically very weak, i.e., $[\mathrm{RNAp}] / \widetilde{K}_1 \ll 1$
  – opportunity for regulation, e.g., boost promoter binding probability

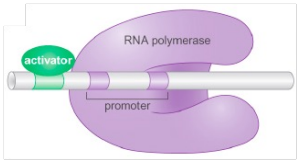
$$k_2 = 10^{-3} \sim 10^{-1} \text{ sec}^{-1}$$

  – fast end need not be faster
    [cf: search kinetics]
  – another opportunity for regulation

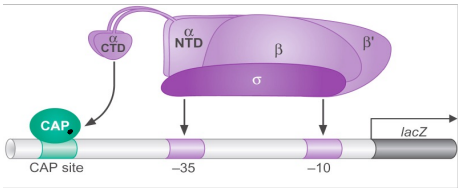
12

3. TF-RNAp interaction

- Recruitment:

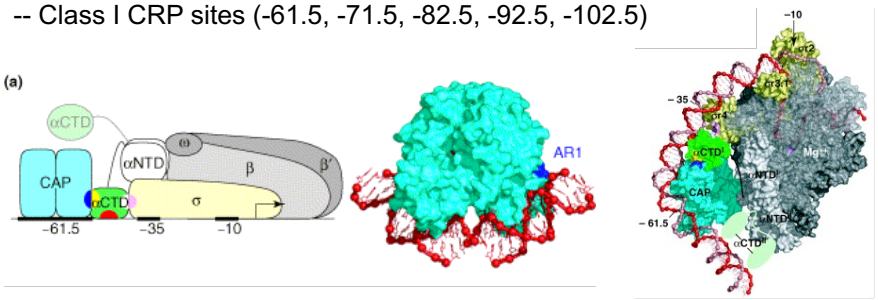


e.g., CRP (activated by cAMP; aka CAP)

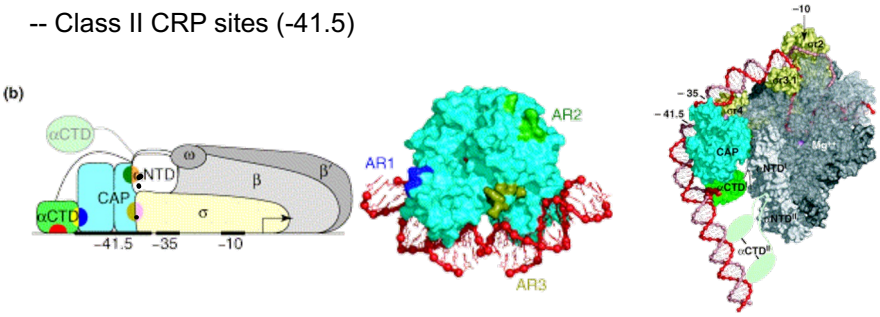


13

-- Class I CRP sites (-61.5, -71.5, -82.5, -92.5, -102.5)



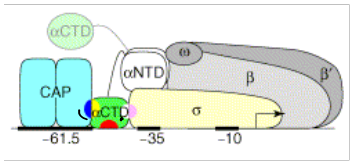
-- Class II CRP sites (-41.5)

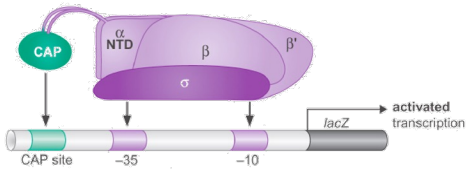


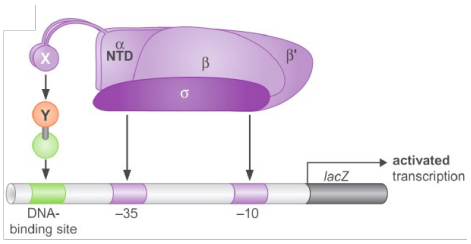
14

-- mechanism of activation? activator bypass experiments:

➔ glue-like attraction between CRP and α-CTD





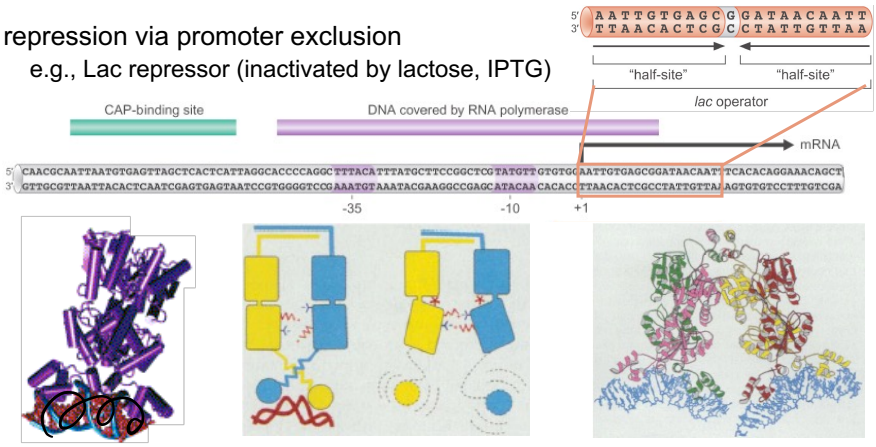


(bacterial two-hybrid system)

15

-- repression via promoter exclusion

e.g., Lac repressor (inactivated by lactose, IPTG)

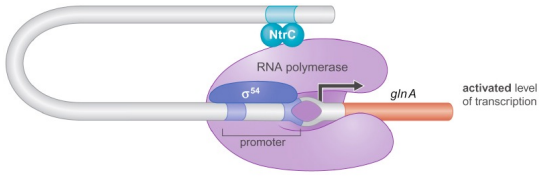


dimer of Lac repressor

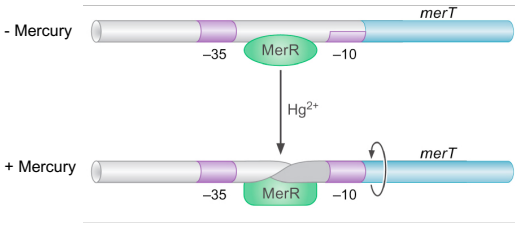
- minimal unit for DNA binding
- large portion of molecule for ligand binding and allosteric control of the DNA-binding domain
- wt Lac tetramerizes

16

- Allosteric mechanisms of activation
  - NtrC (activated by phosphorylation under low nitrogen level): can activate σ^{54} from 1-2 kbp away; has ATPase activity

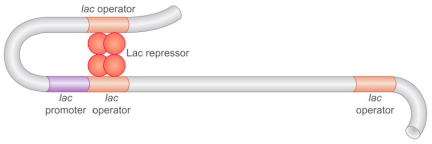


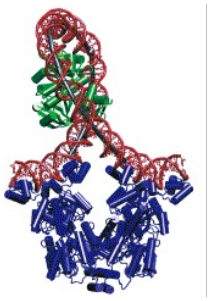
  - MerR: activate σ^{70}-mediated tsx by twisting promoter DNA



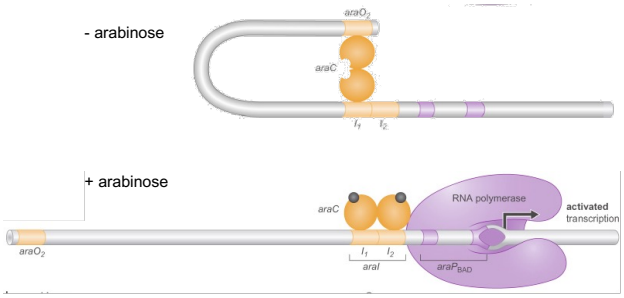
17

- DNA looping
  - tight repression by lac tetramers





  - strong regulation by AraC



18

B. Basic Models of Transcriptional Control

1. tsx init by RNAp alone

$$\underbrace{\text{RNAp} + \text{promoter}}_{\text{“}P\text{”}} \underset{}{\overset{K_p}{\rightleftharpoons}} \text{RNAp} \cdot \text{promoter} \xrightarrow{\alpha} \text{RNAp} + \text{promoter} + \underbrace{mRNA}_{\text{“}m\text{”}}$$

- mRNA level: $\frac{d}{dt}[m] = \alpha \cdot \mathcal{P} - \beta \cdot [m]$

  α: probability of promoter occupation by RNAp

  β: mRNA degradation

- steady-state mRNA level (measurable): $[m^*] = \alpha \cdot \mathcal{P} / \beta$
- from protein-DNA interaction, expect $\mathcal{P} = 1 \Big/ \left(1 + \widetilde{K}_p / [P]_{av}\right)$

  where $[P]_{av}$ = avail RNAp conc ≈ 0.5 ~ 1 μM

  $\widetilde{K}_p = N \cdot K_p / K_{ns} = 10^4 \sim 10^7$ nM

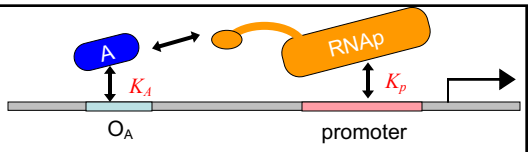
➔ for RNAp by itself, $\mathcal{P} \approx [P]_{av} / \widetilde{K}_p \ll 1$

➔ TF can modulate $\mathcal{P}$ or α

19

2. Activation by recruitment

How does gene expression depend on [A]?



Strategy: [Shea & Ackers, 1985]

-- assume $[m^*] = \alpha \cdot \mathcal{P}([A],[P])/\beta$

-- $\mathcal{P}$ computed according to thermodynamics (assumes thermal equilibrium)

Recall for operator site alone: $p_A = [A]_{tot} \Big/ \left([A]_{tot} + \widetilde{K}_A\right)$

[will drop tilde and subscript "tot" from here on]

Total probability of RNAp binding to promoter in the presence of A:

$$\mathcal{P}([A],[P]) = \frac{W(0,1)+W(1,1)}{W(0,0)+W(0,1)+W(1,0)+W(1,1)}$$

where $W(\sigma_A, \sigma_P)$ = weight of { operator A is occupied (σ_A =1) or unoccupied (σ_A =0); promoter is occupied (σ_P =1) or unoccupied (σ_P =0) }

20

Dependence of the total probability of RNAp-promoter binding on A:

$$\mathcal{P}([A],[P]) = \frac{W(0,1)+W(1,1)}{W(0,0)+W(0,1)+W(1,0)+W(1,1)}$$

Form of $W(\sigma_A, \sigma_P)$: let $W(0,0)=1$ (since only ratio of weights matter)

$$W(0,1) = [P]/K_p, \quad W(1,0) = [A]/K_A$$
$$W(1,1) = \omega \cdot ([A]/K_A) \cdot ([P]/K_P)$$

derived from stat mech

$\omega = e^{-E_{int}/k_BT}$ ("cooperativity factor")

check: P by itself, i.e., [A]=0, $p_P = \dfrac{W(0,1)}{W(0,0)+W(0,1)} = \dfrac{[P]/K_P}{1+[P]/K_P}$

P given A, i.e., [A]=∞, $p_{P|A} = \dfrac{W(1,1)}{W(1,0)+W(1,1)} = \dfrac{\omega \cdot [P]/K_P}{1+\omega \cdot [P]/K_P}$

➔ promoter strength effectively increased (K_p ➔ K_p/ω)

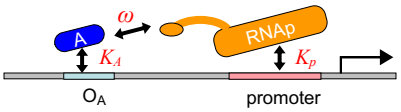
Compact notation: $W(\sigma_A,\sigma_P) = ([A]/K_A)^{\sigma_A} \cdot ([P]/K_P)^{\sigma_P} \cdot \omega^{\sigma_A \sigma_P}$

then $\mathcal{P}([A],[P]) = \sum_{\sigma_A} W(\sigma_A,\sigma_P=1) \Big/ \sum_{\sigma_A,\sigma_P} W(\sigma_A,\sigma_P)$

21

$$\mathcal{P}([A],[P]) = \frac{[P]/K_P + \omega\cdot([A]/K_A)\cdot([P]/K_P)}{1+[A]/K_A + [P]/K_P + \omega\cdot([A]/K_A)\cdot([P]/K_P)}$$

- function of $[A]$ and $[P]$, parameterized by K_A, K_B, ω
- typical parameter range:
  - promoters weak: $[P]/K_P \ll 1$
  - TF concentration: $[A] = 1 \sim 1000$ nM
  - operators tunable: $K_A = 1 \sim 1000$ nM
  - cooperativity weak: $\omega = 10 \sim 100$ (typically ~ 20)



➔want promoter activity as function of $[A]$

"basal level" = $\mathcal{P}_{lo}$

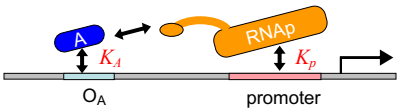
- expected behavior
  - low state: for $[A]=0$, $\mathcal{P} = \frac{[P]/K_P}{1+[P]/K_P} \approx [P]/K_P \ll 1$

$$\Rightarrow \quad \mathcal{P} \approx \mathcal{P}_{lo} \text{ as long as } \omega\cdot[A]/K_A \ll 1$$

22

$$\mathcal{P}([A],[P]) = \frac{[P]/K_P + \omega\cdot([A]/K_A)\cdot([P]/K_P)}{1+[A]/K_A + [P]/K_P + \omega\cdot([A]/K_A)\cdot([P]/K_P)}$$

- function of $[A]$ and $[P]$, parameterized by K_A, K_B, ω
- typical parameter range:
  - promoters weak: $[P]/K_P \ll 1$
  - TF concentration: $[A] = 1 \sim 1000$ nM
  - operators tunable: $K_A = 1 \sim 1000$ nM
  - cooperativity weak: $\omega = 10 \sim 100$ (typically ~ 20)



➔want promoter activity as function of $[A]$

"basal level" = $\mathcal{P}_{lo}$

- expected behavior
  - low state: for $[A]=0$, $\mathcal{P} = \frac{[P]/K_P}{1+[P]/K_P} \approx [P]/K_P \ll 1$

$$\Rightarrow \quad \mathcal{P} \approx \mathcal{P}_{lo} \text{ as long as } \omega\cdot[A]/K_A \ll 1$$

  - high state: for $[A] \gg K_A$, can consider A always bound to O_A

$$\Rightarrow \quad \mathcal{P}_{hi} \approx \frac{\omega\cdot[P]/K_P}{1+\omega\cdot[P]/K_P} \le 1$$

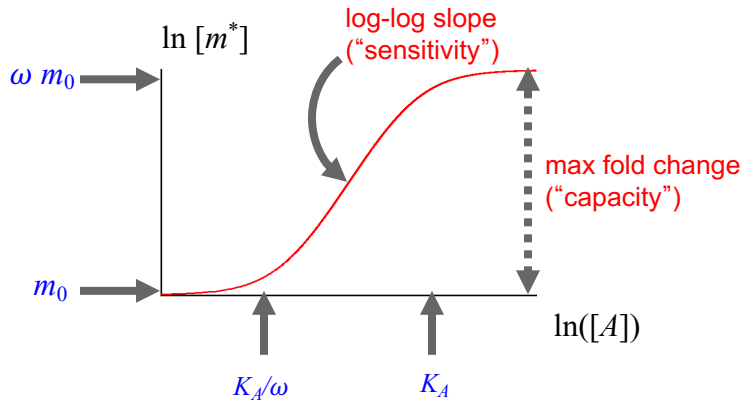
  - maximal fold-change ("capacity"): $\mathcal{P}_{hi}/\mathcal{P}_{lo} \approx \omega\cdot\frac{1+[P]/K_P}{1+\omega\cdot[P]/K_P} \le \omega$

➔ for maximal control, want weak promoter such that $\omega\cdot[P]/K_P \ll 1$

23

take $\omega[P]/K_P \ll 1$ from here on, then $\mathcal{P} \approx \frac{[P]}{K_P} \cdot \frac{1+\omega\cdot[A]/K_A}{1+[A]/K_A}$

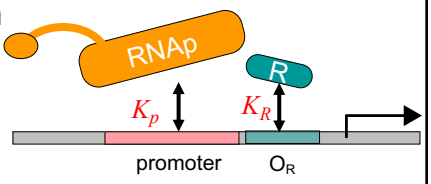
$\Rightarrow \quad [m^*] = \alpha\cdot\mathcal{P}/\beta \approx m_0 \cdot \frac{1+\omega\cdot[A]/K_A}{1+[A]/K_A}, \qquad m_0 \equiv \frac{\alpha\,[P]}{\beta\,K_P}$



24

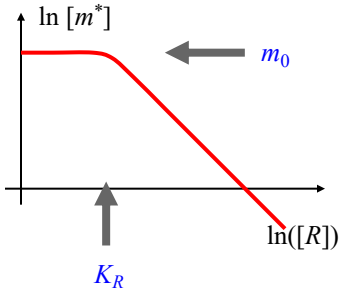
3. Repression by promoter occlusion

$W(\sigma_R = 1, \sigma_P = 0) = [R]/K_R,$
$W(\sigma_R = 0, \sigma_P = 1) = [P]/K_P,$
$W(\sigma_R = 1, \sigma_P = 1) = 0$



[promoter and O_R cannot be simultaneously occupied]

$$\mathcal{P} = \frac{W(0,1)+W(1,1)}{W(0,0)+W(0,1)+W(1,0)+W(1,1)} = \frac{[P]/K_P}{1+[P]/K_P+[R]/K_R} \propto \frac{1}{1+[R]/K_R}$$



-- large $[R]$ can provide arbitrarily strong repression according to model
-- "promoter leakage" provides the lower limit on $[m^*]$
-- high TF conc often generate toxic side effects

25

4. Activation by catalysis (rather than recruitment)

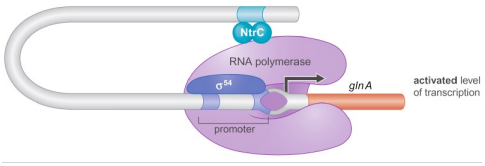
$$\underbrace{\text{RNAp} + \text{promoter}}_{\text{“}\mathcal{P}\text{”}} \underset{}{\overset{K_P}{\rightleftharpoons}} \text{RNAp} \cdot \text{promoter} \xrightarrow{\alpha} \text{RNAp} + \text{promoter} + \underbrace{mRNA}_{\text{“}m\text{”}}$$

- mRNA level: $\frac{d}{dt}[m] = \alpha \cdot \mathcal{P} - \beta \cdot [m]$

  tsx init rate ↑ ↑ probability of promoter occupation by RNAp; ↓ mRNA degradation

- steady-state mRNA level (measurable): $[m^*] = \alpha \cdot \mathcal{P} / \beta$

for σ54 promoters, the rate of promoter opening catalyzed by activator



26

4. Activation by catalysis (rather than recruitment)

$$\underbrace{\text{RNAp} + \text{promoter}}_{\text{“}\mathcal{P}\text{”}} \underset{}{\overset{K_P}{\rightleftharpoons}} \text{RNAp} \cdot \text{promoter} \xrightarrow{\alpha} \text{RNAp} + \text{promoter} + \underbrace{mRNA}_{\text{“}m\text{”}}$$

- mRNA level: $\frac{d}{dt}[m] = \alpha \cdot \mathcal{P} - \beta \cdot [m]$

  tsx init rate ↑ ↑ probability of promoter occupation by RNAp; ↓ mRNA degradation

- steady-state mRNA level (measurable): $[m^*] = \alpha \cdot \mathcal{P} / \beta$

for σ54 promoters, the rate of promoter opening catalyzed by activator

model: $\alpha \Rightarrow \alpha_{\sigma_A}$

$$\alpha \cdot \mathcal{P} \Rightarrow \sum_{\sigma_A} \alpha_{\sigma_A} \cdot W(\sigma_A, \sigma_P = 1) \Big/ \sum_{\sigma_A, \sigma_P} W(\sigma_A, \sigma_P)$$

$$\Rightarrow \quad [m^*] \approx m_0 \cdot \frac{1 + \frac{\alpha_1}{\alpha_0} \cdot \omega \cdot [A]/K_A}{1 + [A]/K_A}, \qquad m_0 \equiv \frac{\alpha_0}{\beta} \frac{[P]}{K_P}$$

➔ same form as recruitment, but capacity increased by α_1/α_0
➔ large fold change, but dedicated components

27

- "Advantages of the σ^{54} system:
  - very low basal rate for small α_0
    (activators need to consume ATP to catalyze open complex)
  - large capacity w/o need for large ω
    (recall also that very large ω can reduce capacity)
  - can activate from a long distance away (via DNA looping -- later)
- but in most bacteria species, there is at most one σ^{54} factor
  (compared to many families of σ^{70} factors)
- possible disadvantages?
  long distance activation can create unintentional cross talk unless
  different promoters are kept far apart (require long chromosomes)
  or separated by "insulating elements" (not available for prokaryotes)

28

5. Induction of TF

$$X + I \underset{k_-}{\overset{k_+}{\rightleftharpoons}} XI$$

dissociation constant $K_I \equiv \frac{[X]\cdot[I]}{[XI]} = \frac{k_-}{k_+}$

$[X]_{tot} = [X] + [XI]$

$$[XI] = [X]_{tot}\frac{[I]}{[I]+K_I} \approx [X]_{tot}\frac{[I]_{tot}}{[I]_{tot}+K_I}$$

usually $[I]_{tot} \gg [X]_{tot}$, so $[I] \approx [I]_{tot}$

will drop the subscript "tot" from here on

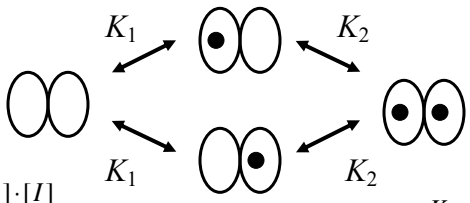
"activated TF" X^* = form of TF able to bind specifically to DNA
or able to activate RNAp

if $X^* = XI$, then $[X^*] = [X]_{tot}\frac{[I]}{[I]+K_I}$

if $X^* = X$, then $[X^*] = [X]_{tot}\frac{K_I}{[I]+K_I}$

29

often TF are dimers (X_2)



$$K_1 \equiv \frac{[X_2]\cdot[I]}{[X_2 I]} \qquad K_2 \equiv \frac{[X_2 I]\cdot[I]}{[X_2 I_2]}$$

$$[X_2]_{tot} = [X_2]\cdot\left(1 + 2\frac{[I]}{K_1} + \frac{[I]^2}{K_1 K_2}\right)$$

- non-cooperative ($K_1 = K_2$): $[X_2] = [X_2]_{tot} \Big/ \left(1 + \frac{[I]}{K_1}\right)^2$
- strongly cooperative ($K_2 \ll K_1$): $[X_2] \approx [X_2]_{tot} \Big/ \left(1 + \frac{[I]^2}{K_1 K_2}\right)$
  (e.g., binding of 2nd molecule much easier after 1st is bound)

Hill function

➔ active TF could be X_2, $X_2 I$, or $X_2 I_2$

30