

## Regulatory effects on RNA polymerase

Binding constants  
Rate constants  
Measuring  $K_B$

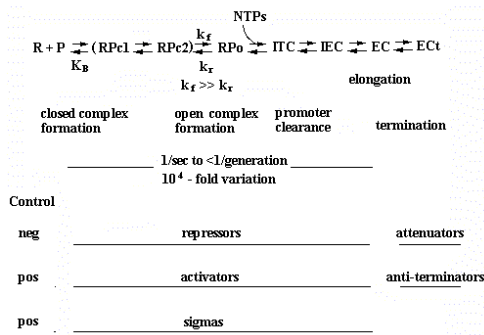
## Multiple steps in initiation and elongation by RNA polymerase are targets for regulation

RNA Polymerase has to

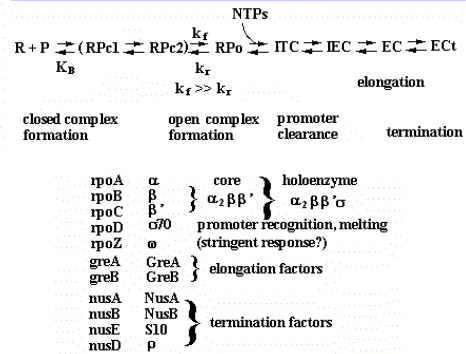
- \* bind to promoters,
- \* form an open complex,
- \* initiate transcription,
- \* escape from the promoter,
- \* elongate
- \* terminate transcription.

All are potential targets for regulation.

## Steps at which RNA polymerase is regulated



## Factors associated with each step



## Effects on $K_B$ , $k_f$ , $k_r$

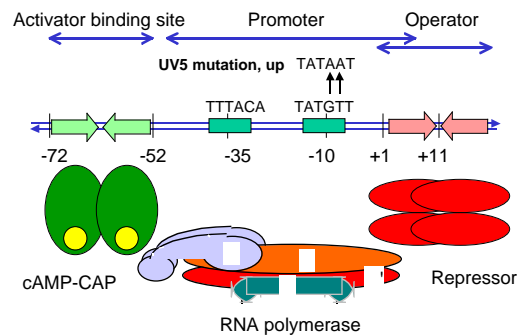
Summarizing a lot of work, we know that:

strong promoters have high  $K_B$ , high  $k_f$ , low  $k_r$ , and high rates of promoter clearance.

weak promoters have low  $K_B$ , low  $k_f$ , high  $k_r$ , and low rates of promoter clearance.

moderate promoters have one or more "weak" spots.

## *lac* regulatory region



## Measurement of $K_B$

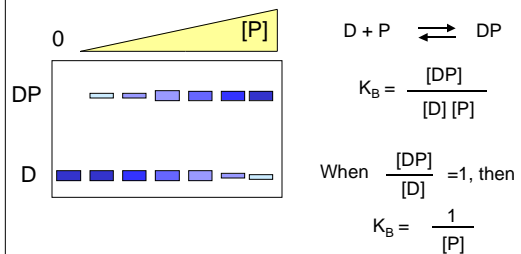
## Synonymous and related terms

$K_B = K_b = K_{eq}$  = equilibrium constant for binding  
 $K_S = K_B$  for binding of protein to a *specific* DNA sequence  
 $K_{NS} = K_B$  for binding of protein to *nonspecific* DNA  
 $[P] = [P_2]$  = molar concentration of protein  
 $[R_4]$  = molar concentration of repressor  
 $[D]$  = molar concentration of free DNA  
 $[D_S]$  = concentration of free specific DNA  
 $[D_{NS}]$  = concentration of free **nonspecific** DNA  
 $[DP]$  = molar concentration of DNA-protein complex  
 $[R_4D_S]$  = concentration of repressor-operator

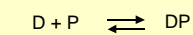
## Techniques to measure amount of protein bound to DNA

- Need:
  - Radioactively labeled DNA (usually a specific sequence)
  - Purified DNA-binding protein
- Combine them and measure the amount of protein-DNA complex and free DNA by:
  - Electrophoretic mobility shift assays
  - DNase I footprinting
  - Retention of protein-DNA complexes on filters

## Measure $K_B$ by EMSA



## Measure $K_B$ from $[DP]/[D]$



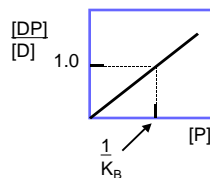
$$K_B = \frac{[DP]}{[D][P]}$$

$$\frac{[DP]}{[D]} = K_B [P]$$

When  $\frac{[DP]}{[D]} = 1$ , then

$$K_B = \frac{1}{[P]}$$

If you could measure  $[DP]$  and  $[D]$  at each  $[P]$ , you could measure  $K_B$ :



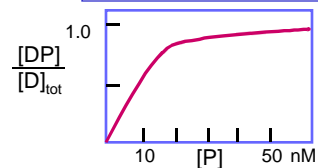
slope =  $K_B$

## Measure $K_B$ from $[DP]/[D]_{tot}$

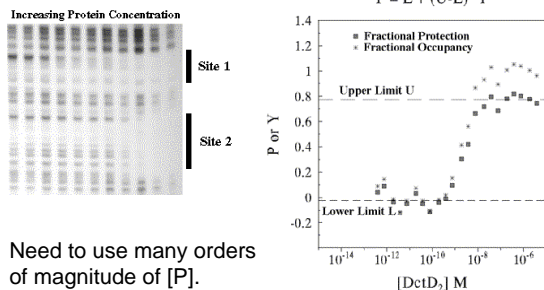
It is more reliable to measure the fraction of labeled DNA in complex with protein, i.e.  $[DP]/[D]_{tot}$

Substitution of  $[D] = [D]_{tot} - [DP]$  into equation for  $K_B$  gives:

$$\frac{[DP]}{[D]_{tot}} = \frac{K_B [P]}{1 + K_B [P]}$$



### Protein binding assayed by DNase I footprinting



Need to use many orders of magnitude of [P].

Fr. Dr. Tracy Nixon

### What value for $K_B$ provides the best fit?

- Classical methods:
  - Transform the data into a line
  - Or, e.g., at  $[DP]/[D]_{tot} = 0.5$ , then  $K_B = 1/[P]$
  - Problems:
    - Where to draw the line?
    - No accurate estimate of error
- Nonlinear, least squares regression analysis = NLLS
  - A computer program calculates the goodness of fit for many values of  $K_B$ , then one can choose the best fit (least error).

### Modeling binding reactions for NLLS

We can model binding reactions by

1. tabulating the different states that exist in a system,
2. associating each state with a fractional probability based on the Boltzmann partition function and the Gibb's free energy for that state ( $G_s$ ),  
and
3. determine the probability of any observed measurement by the ratio of
  - a) the sum of fractional probabilities that give the observation, and
  - b) the sum of the fractional probabilities of all possible states.

### Fractional occupancy from fractional probabilities

Where  $j$  is the number of ligands bound, the fractional probability of a particular state is:

$$f_s = \frac{e^{-G_s/RT} \times [P_2]^j}{e^{-G_s/RT} \times [P_2]^j}$$

For a single protein binding to DNA (one site), the fractional occupancy,  $\bar{Y}$ , is:

$$\bar{Y} = \frac{f_1}{f_s} \quad \text{and} \quad \bar{Y} = \frac{e^{-G/RT} \times [P_2]}{1 + e^{-G/RT} \times [P_2]}$$

### Fractional occupancy in terms of $K_B$

Since  $\Delta G = -RT \ln (Keq)$

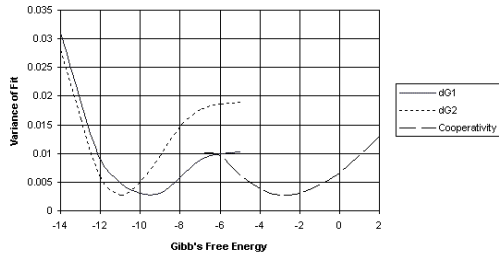
$$\text{then } \bar{Y} = \frac{K_B \times [P_2]}{1 + (K_B \times [P_2])}$$

$$\text{Same as: } \frac{[DP]}{[D]_{tot}} = \frac{K_B [P]}{1 + K_B [P]}$$

### Data analysis by NLLS

- After collecting binding data, one uses a nonlinear, least squares regression analysis to find the values of  $G$  (or  $K_B$ ) that generate a function that best predicts the data.
- Uses maximum likelihood theory to find the value most likely to be correct.
- Produces plot of the variance of fit (or error) over a wide range of possible values for the parameter being measured, e.g.  $G$ .
- Reproducible by different investigators
- Provides a rigorous estimate of error.

### Variance of fit plotted vs. free energy



The G value with the smallest error is the most accurate

### Example of calculating $K_B$ from plot of variance of fit vs. G

$G_1 = -9.5$  kcal/mol gives the minimum variance (or error).

$$\Delta G = -RT \ln(K_{eq})$$

$$\ln K_B = -G/RT = \frac{-(-9.5 \text{ kcal/mol})}{0.59 \text{ kcal/mol}} = 16.1017$$

$$K_B = 9.8 \times 10^6 \text{ M}^{-1}$$

$$R = 1.98 \times 10^{-3} \text{ kcal deg}^{-1} \text{ mol}^{-1}$$

$$T = 298^\circ \text{ K}$$

$$RT = 0.59 \text{ kcal/mol}$$