Noise reduction and information transfer in the cell

Glenn Vinnicombe (Dept. of Engineering, Cambridge)
in collaboration with
Johan Paulsson (Harvard Medical School)
Ioannis Lestas (Dept. of Engineering, Cambridge)

Overview

- The cell is a noisy cell, in spite over being packed full of feedback loops.
- Why?

ARTICLES

Fundamental limits on the suppression of molecular fluctuations

Ioannis Lestas[1], Glenn Vinnicombe[1] & Johan Paulsson[2]

REVIEWS

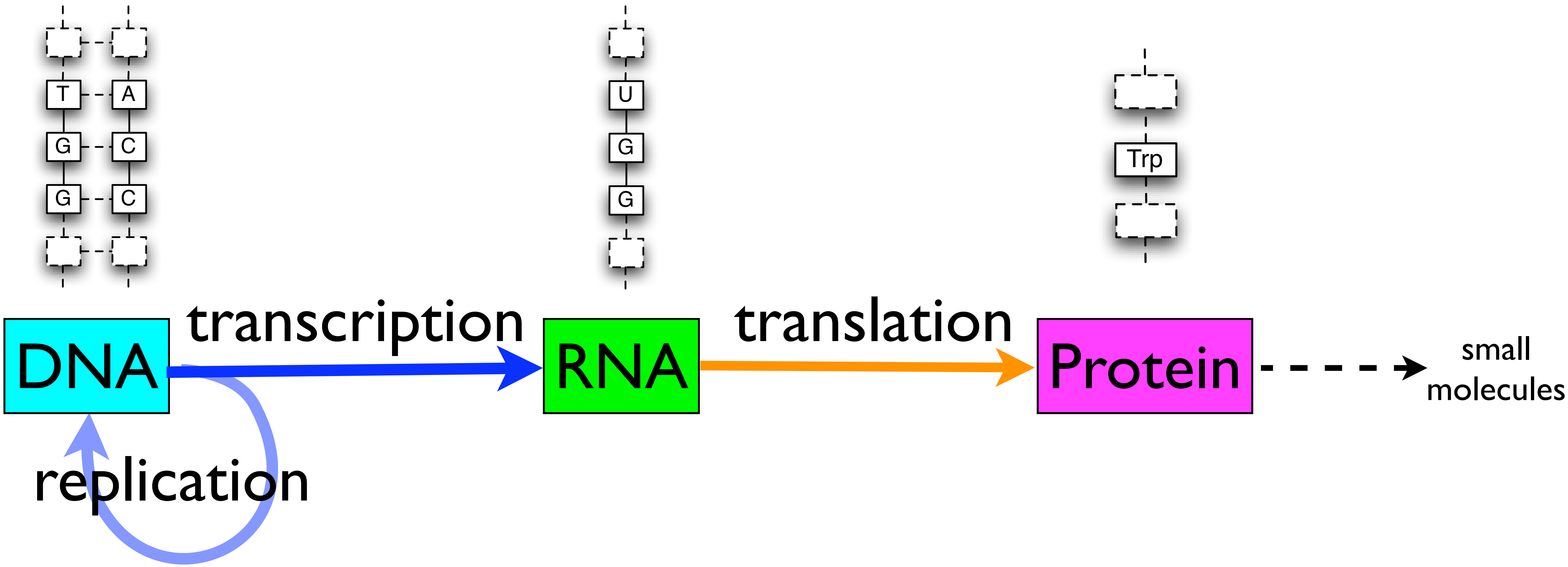
Functional roles for noise in genetic circuits

Avigdor Eldar[1]† & Michael B. Elowitz[1]

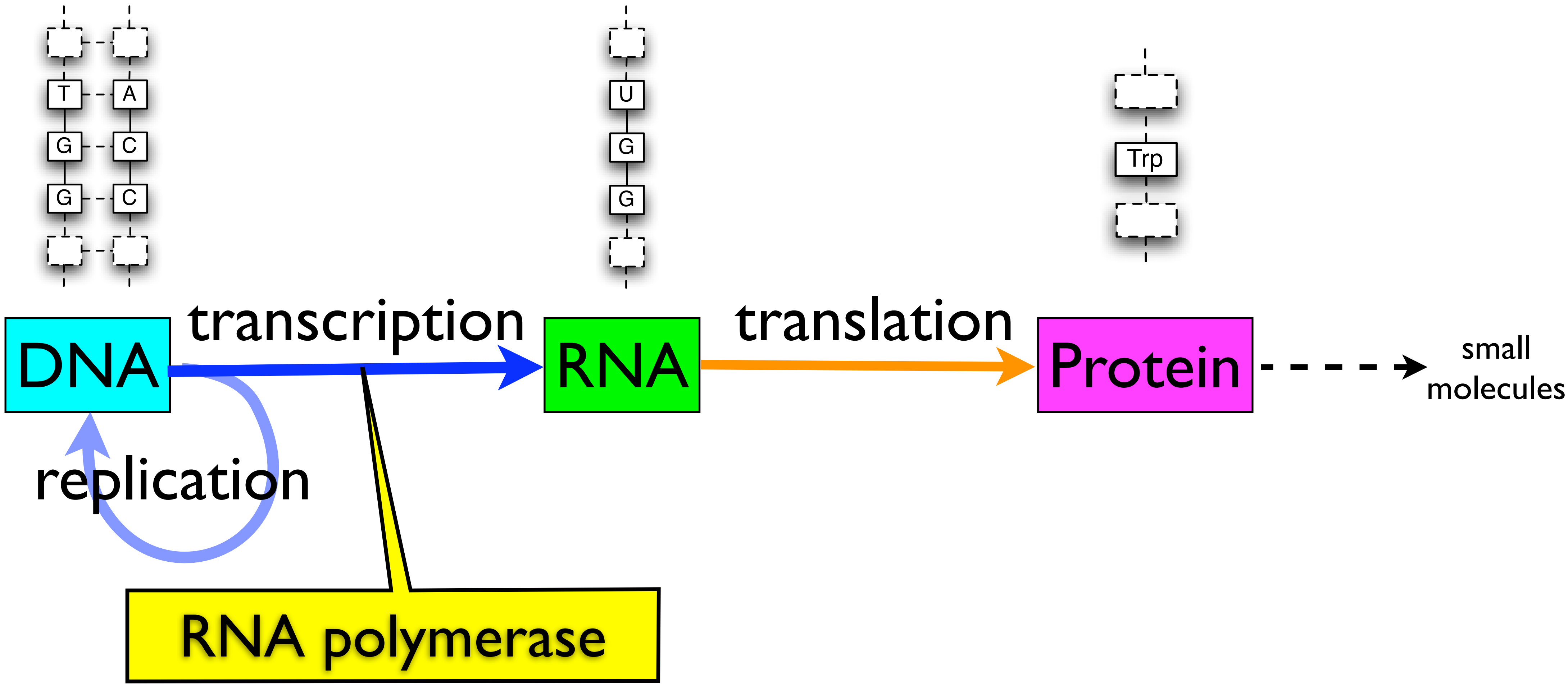
Overview

- The cell is a noisy cell, in spite over being packed full of feedback loops.
- Why?
- I shall present some fundamental limitations for biological systems, in terms of minimum achievable variances.
- These limits apply to the regulation of a single species within an arbitrarily complex network, and the suggest that the cost of reducing noise can be extremely high.

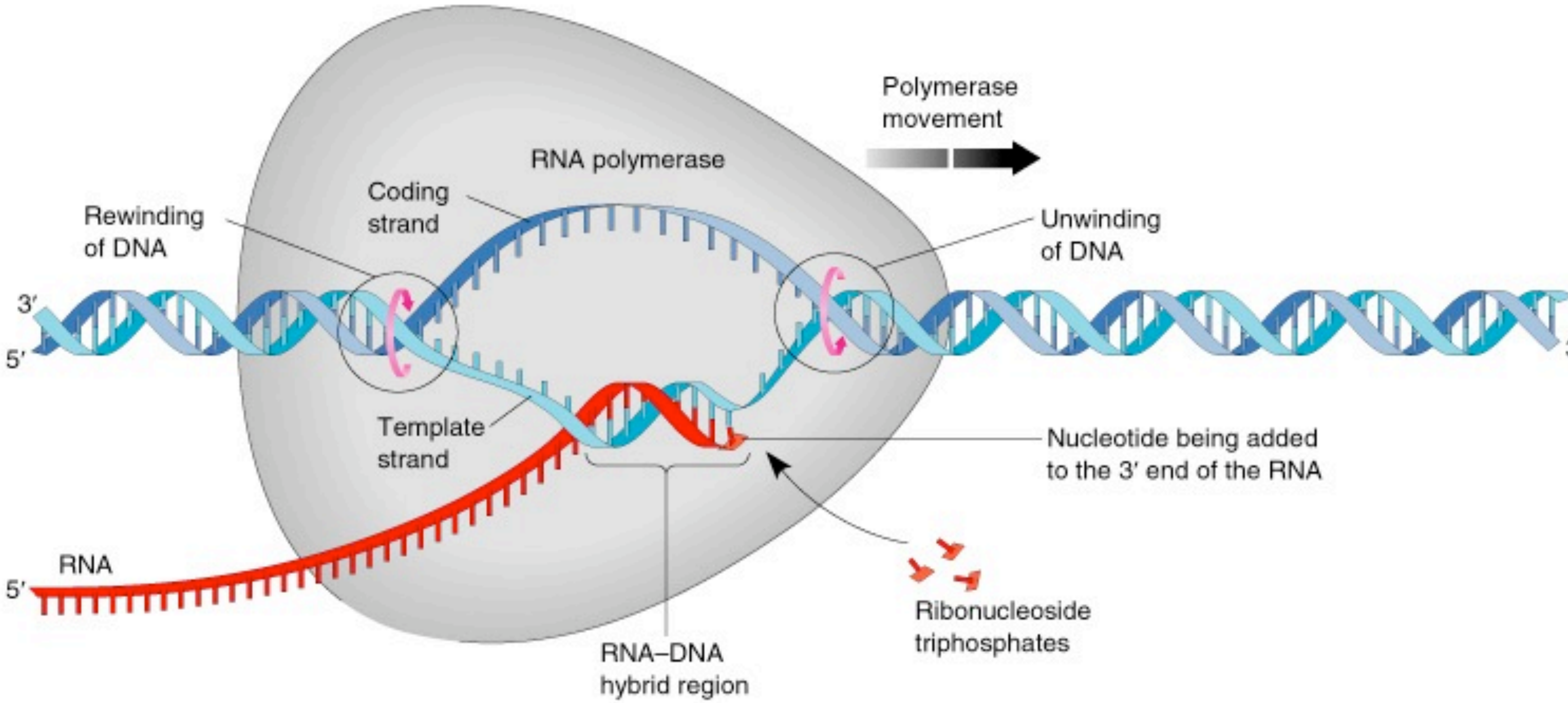
The Central Dogma



The Central Dogma



Transcription



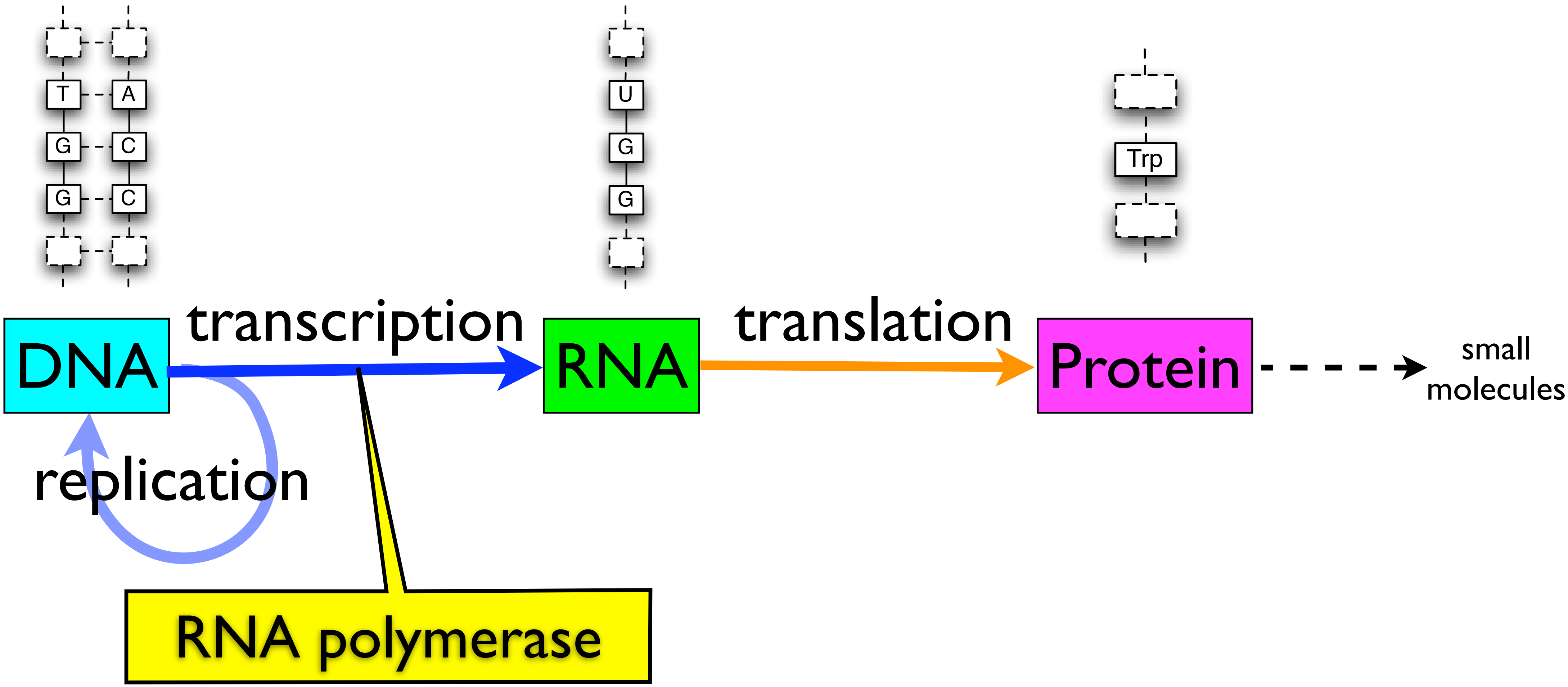
1. RNA polymerase slides along the DNA, creating an open complex as it moves.

2. The DNA strand known as the template strand is used to make a complementary copy of RNA as an RNA–DNA hybrid.

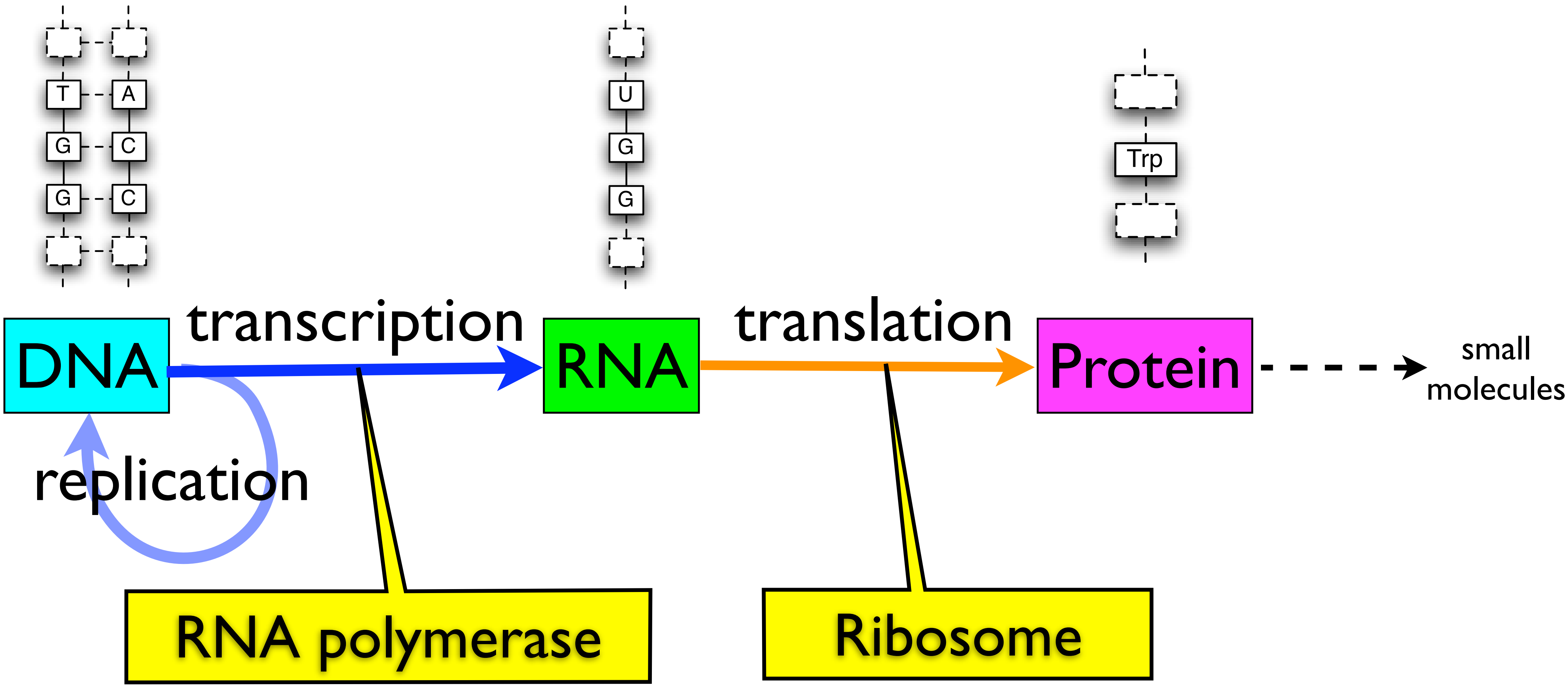
3. The RNA is synthesized in a 5′ to 3′ direction using ribonucleoside triphosphates as precursors. Pyrophosphate is released (not shown).

4. The complementarity rule is the same as the A–T and G–C rule except that U is substituted for T in the RNA.

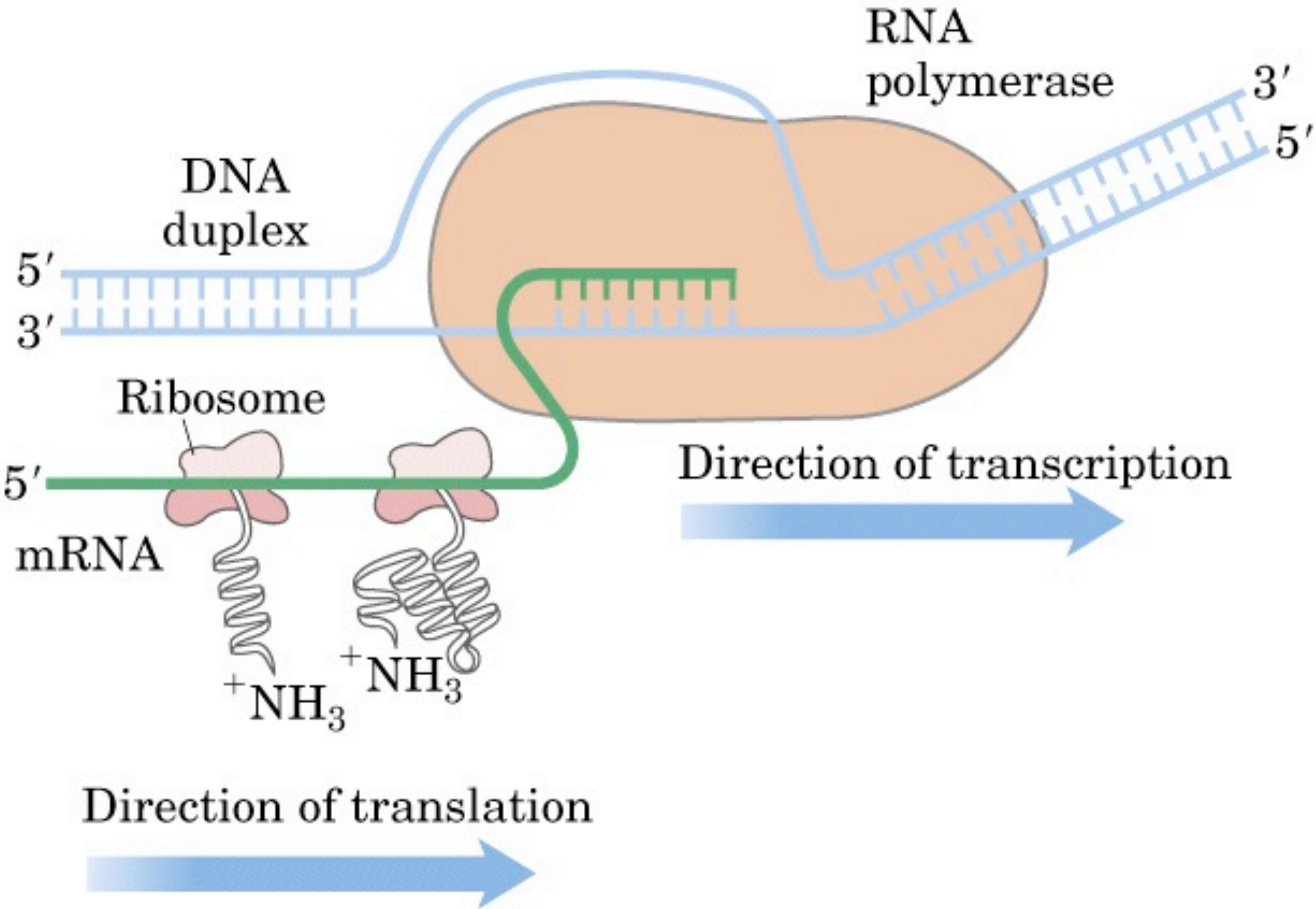
The Central Dogma



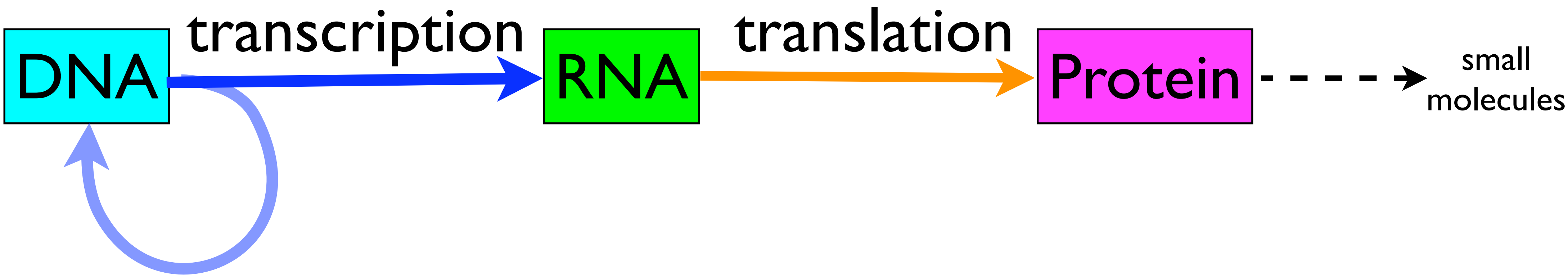
The Central Dogma



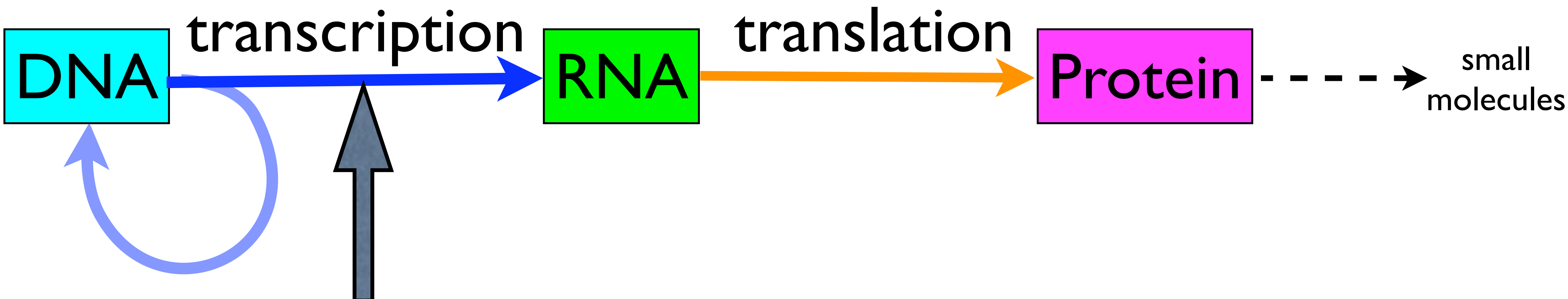
Transcription & Translation in Prokaryotes



Regulation of gene expression



Regulation of gene expression



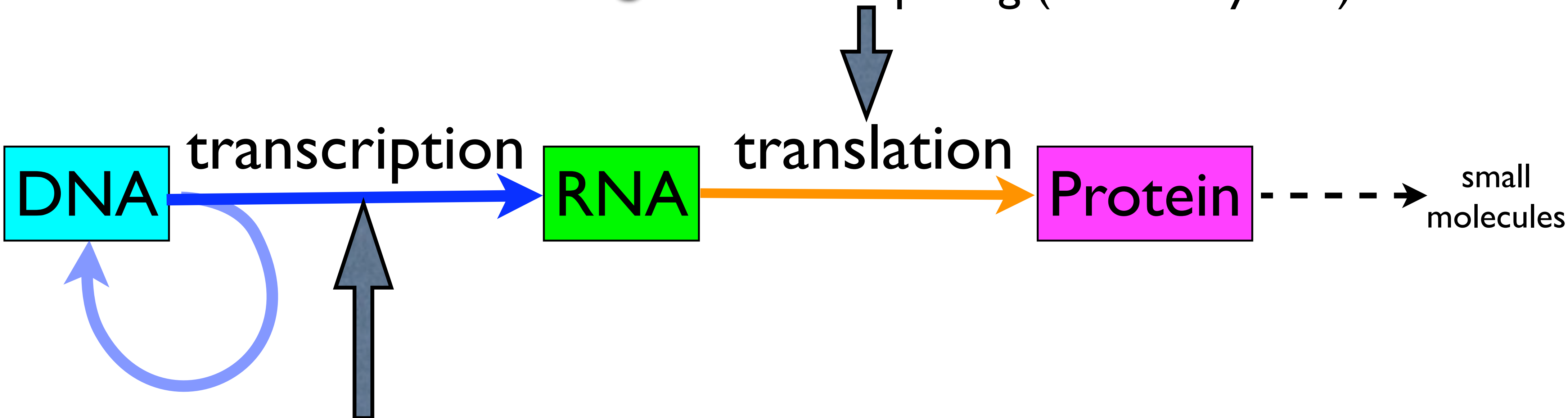
- Gene silencing
- Gene regulation by

DNA binding proteins:

- Activators/repressors
- Sigma factors

Regulation of gene expression

- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)



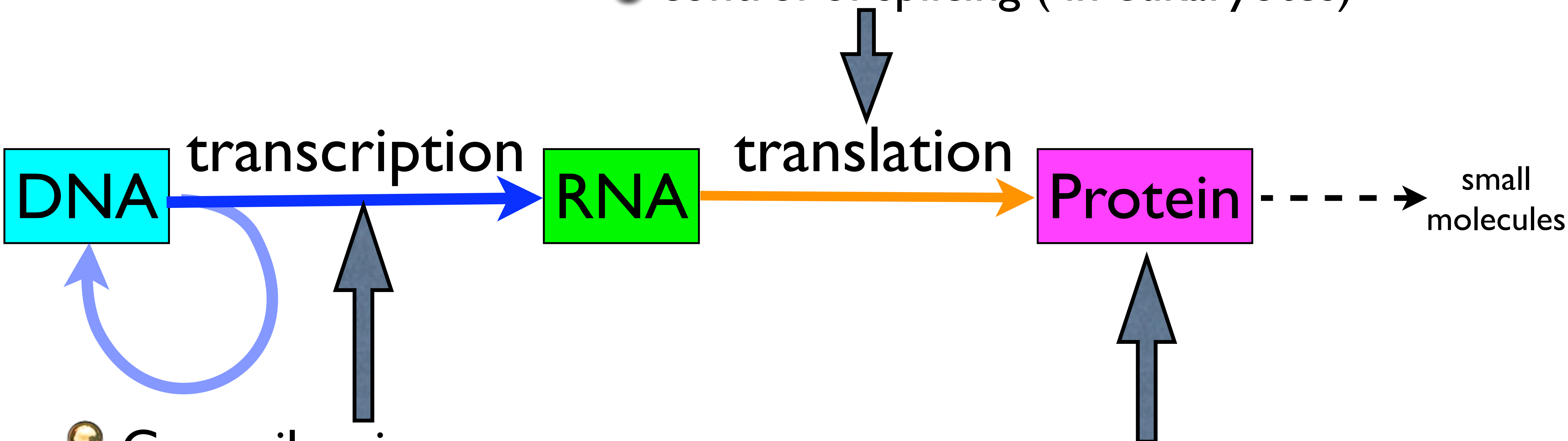
- Gene silencing
- Gene regulation by

DNA binding proteins:

  - Activators/repressors
  - Sigma factors

Regulation of gene expression

- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)

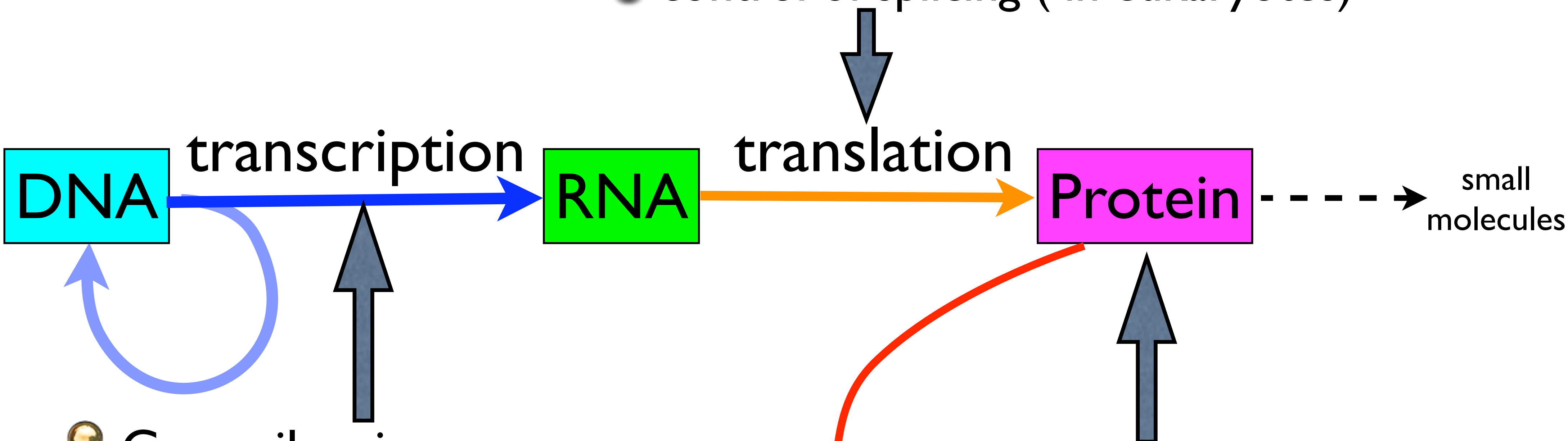


- Gene silencing
- Gene regulation by DNA binding proteins:
  - Activators/repressors
  - Sigma factors

- post-translational modification e.g. phosphorylation

Regulation of gene expression

- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)

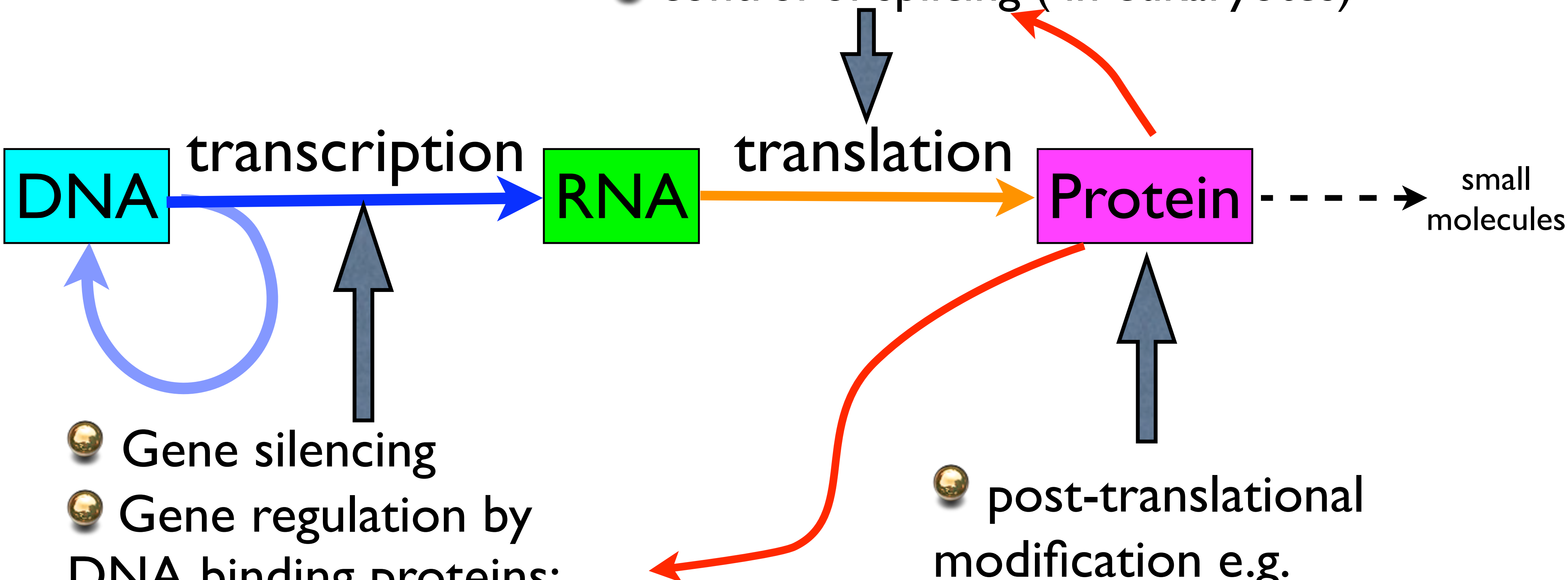


- Gene silencing
- Gene regulation by DNA binding proteins:
  - Activators/repressors
  - Sigma factors

- post-translational modification e.g. phosphorylation

Regulation of gene expression

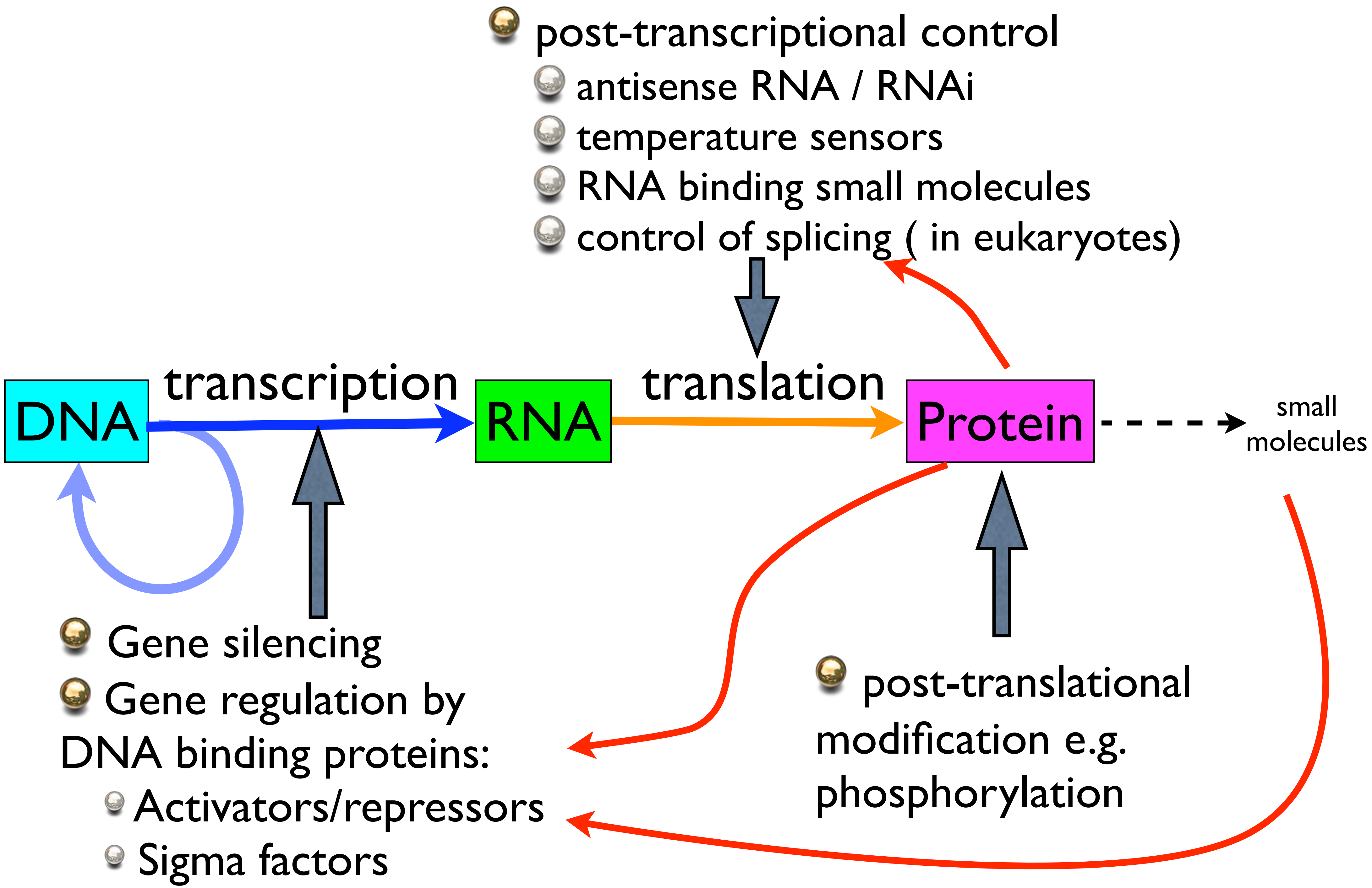
- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)



- Gene silencing
- Gene regulation by DNA binding proteins:
  - Activators/repressors
  - Sigma factors

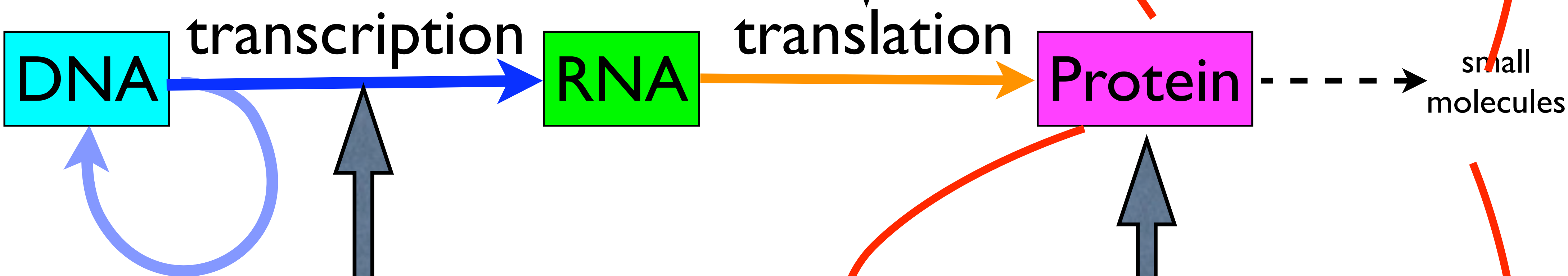
- post-translational modification e.g. phosphorylation

Regulation of gene expression



Regulation of gene expression

- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)



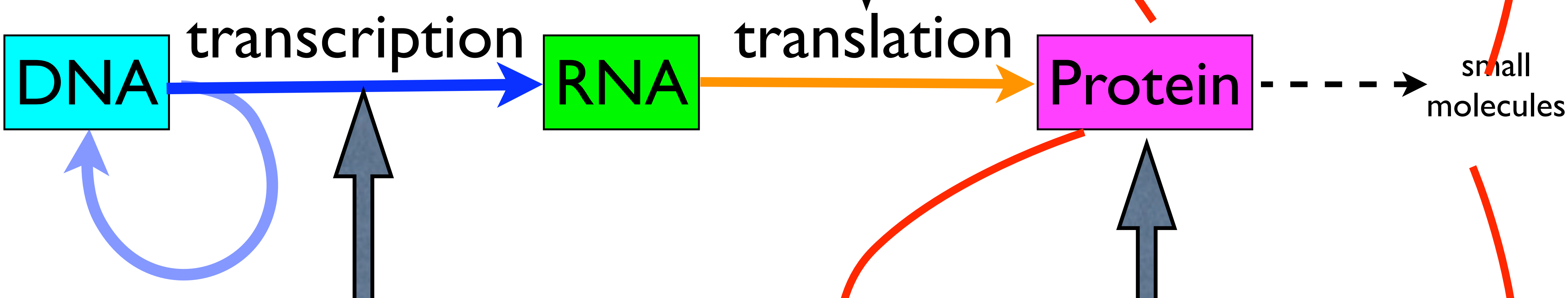
- Gene silencing
- Gene regulation by DNA binding proteins:
  - Activators/repressors
  - Sigma factors

- post-translational modification e.g. phosphorylation

Regulation of gene expression

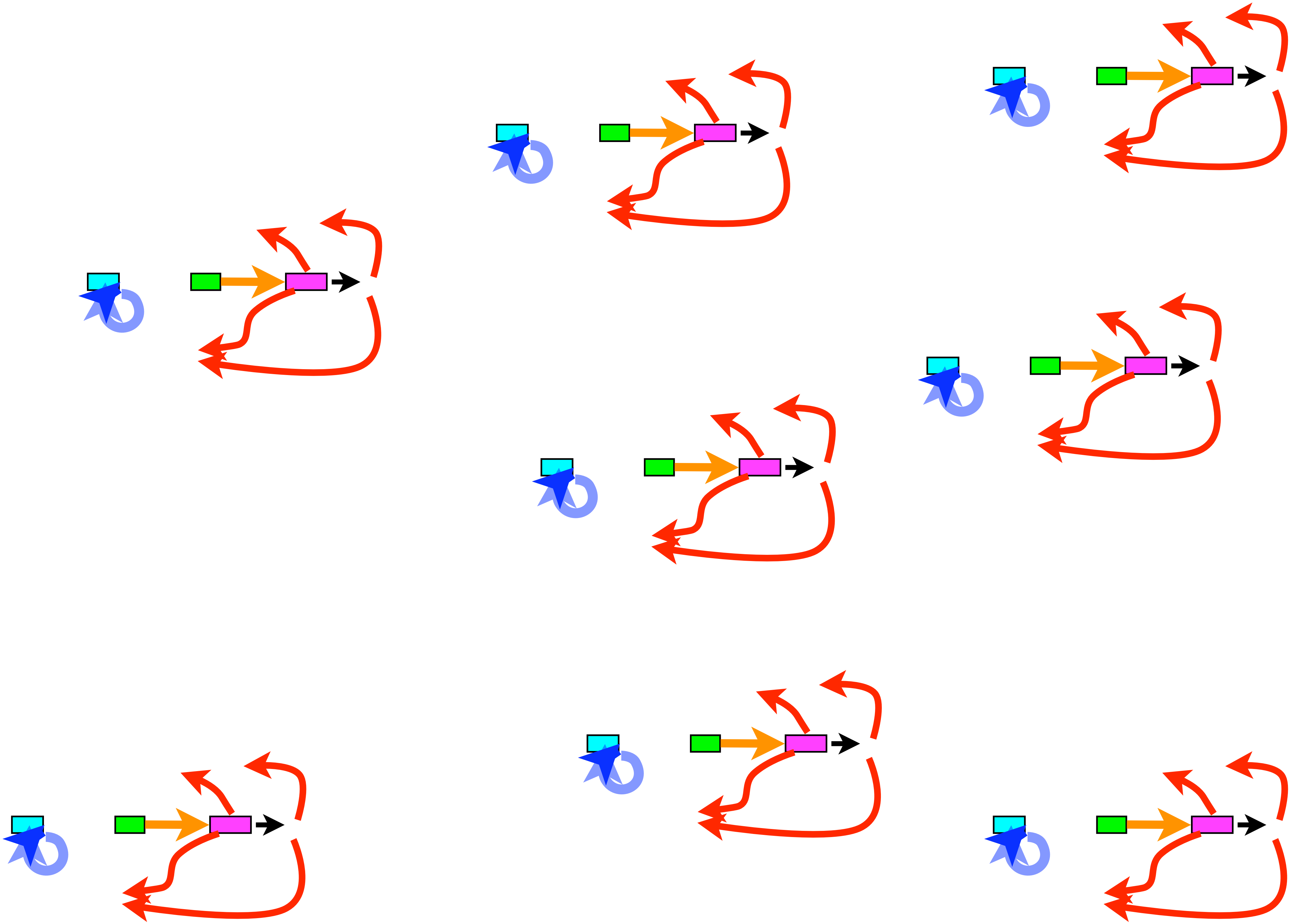
- post-transcriptional control
  - antisense RNA / RNAi
  - temperature sensors
  - RNA binding small molecules
  - control of splicing (in eukaryotes)

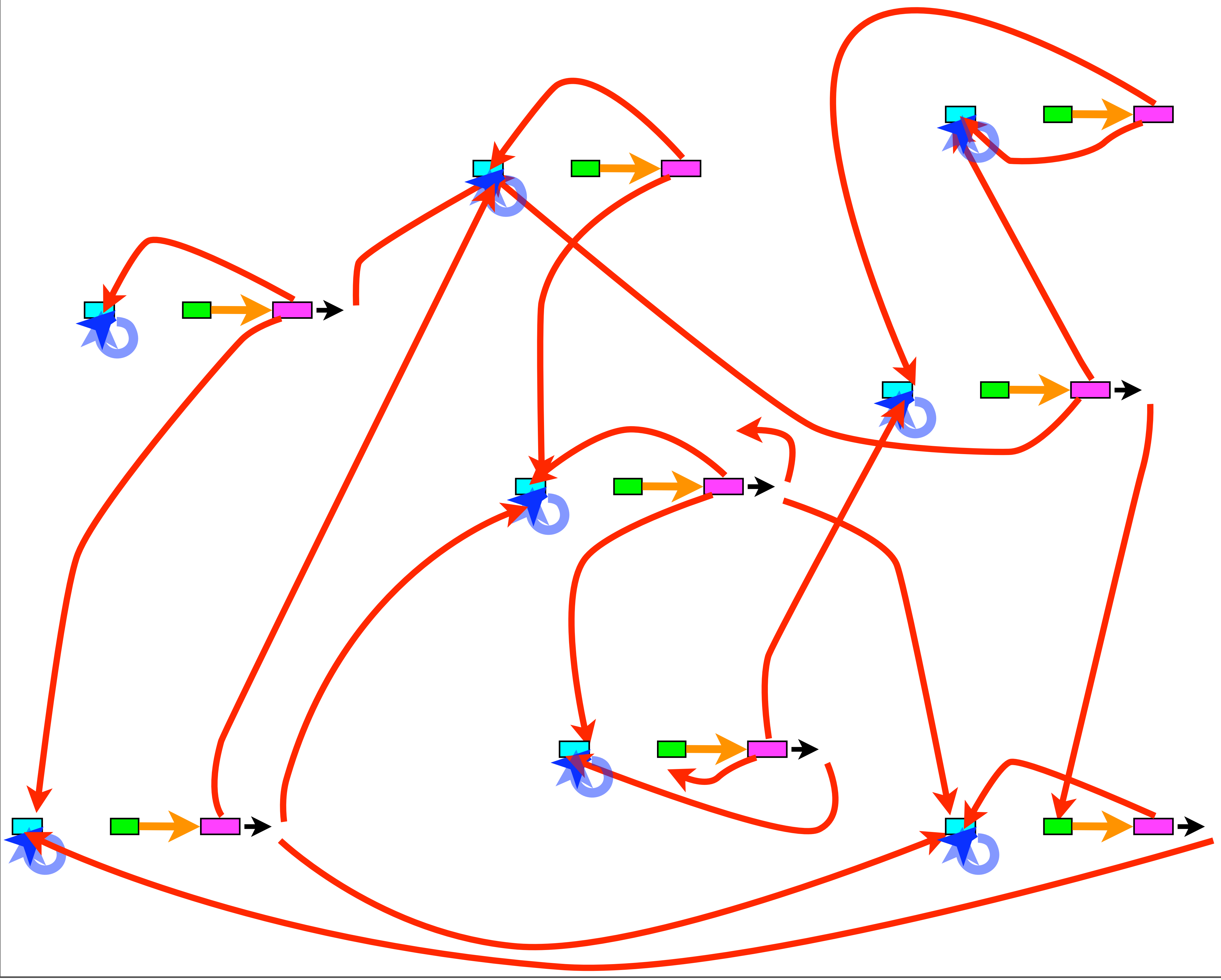
(and this is just part of the story!)

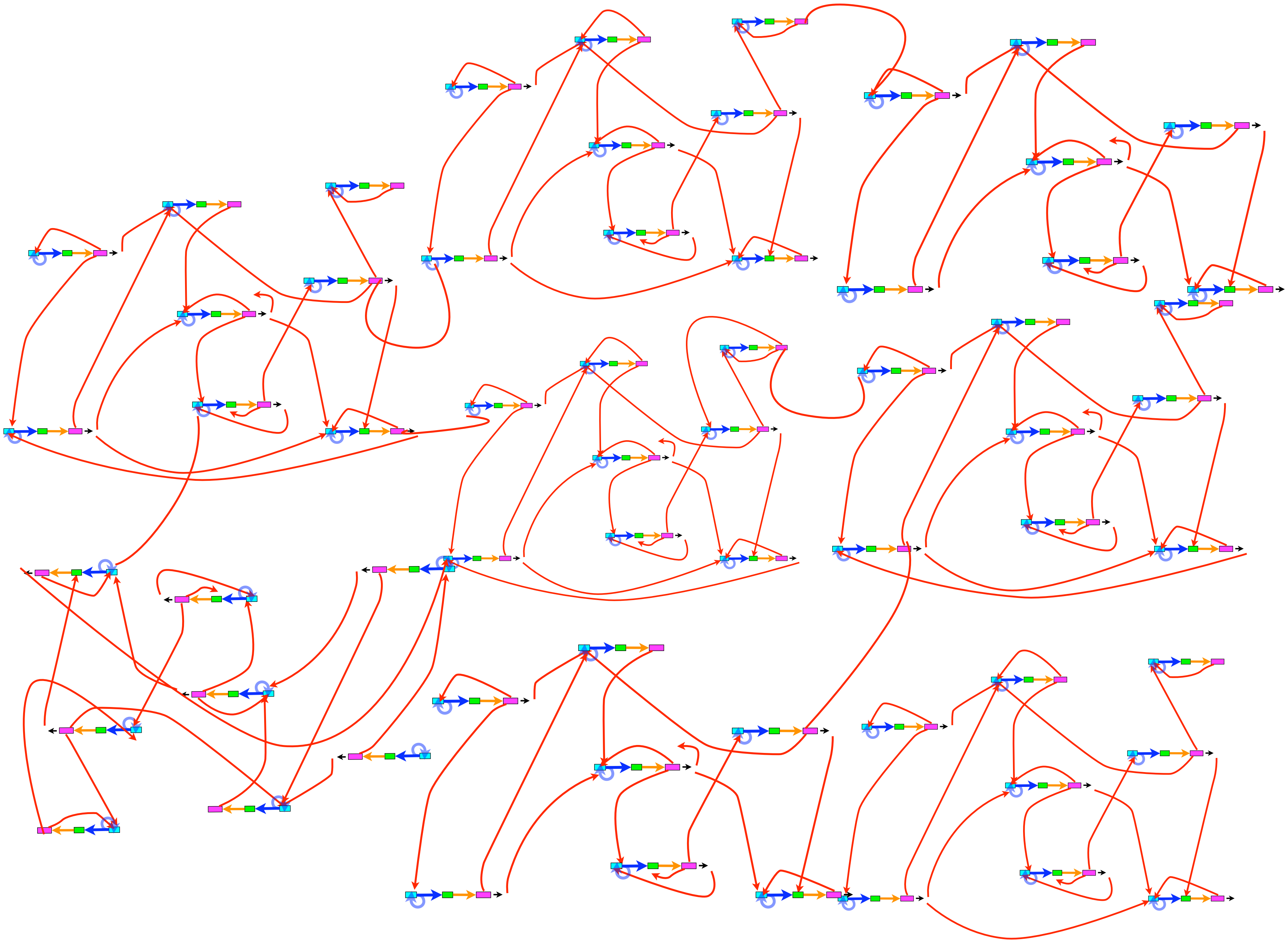


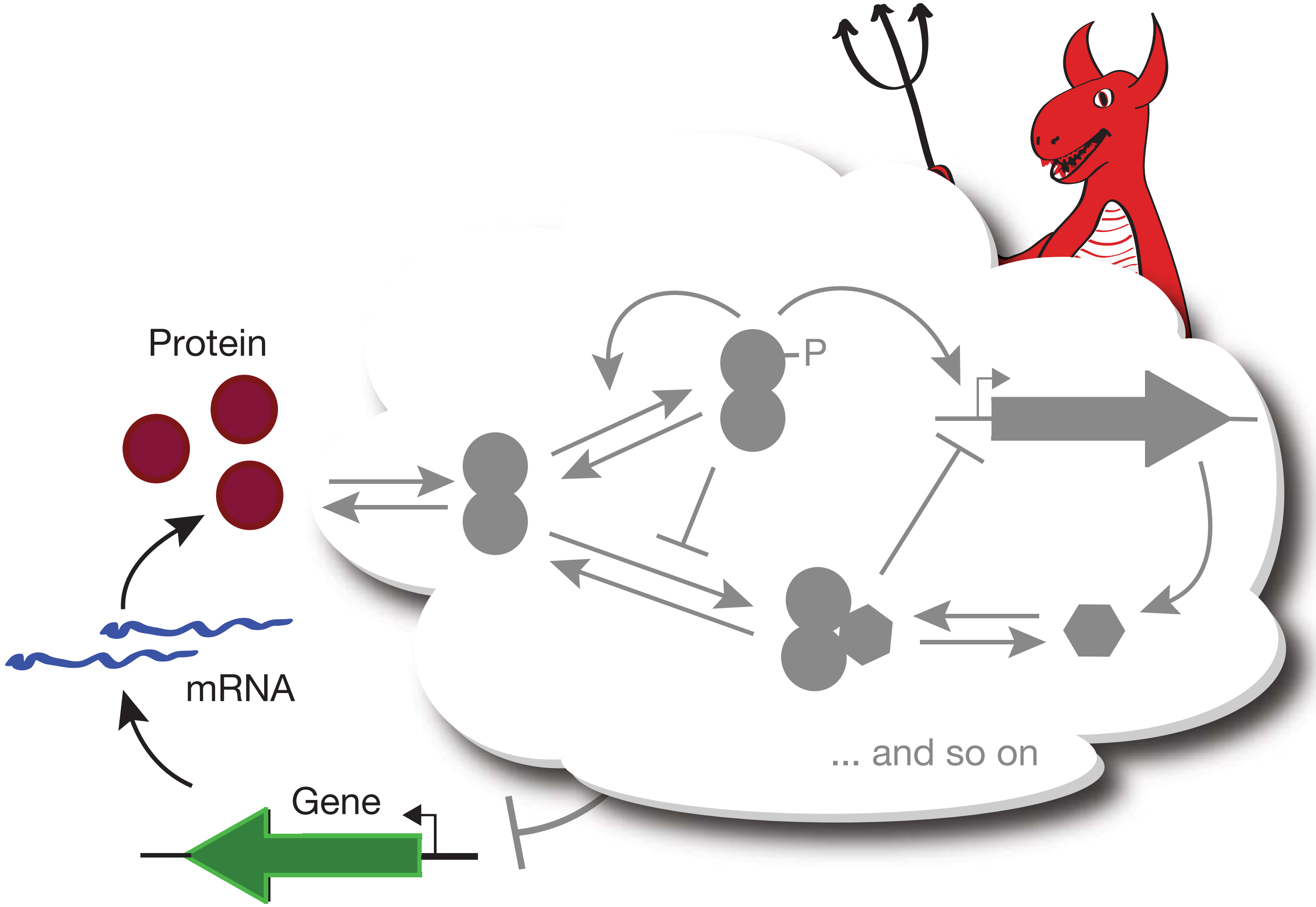
- Gene silencing
- Gene regulation by DNA binding proteins:
  - Activators/repressors
  - Sigma factors

- post-translational modification e.g. phosphorylation



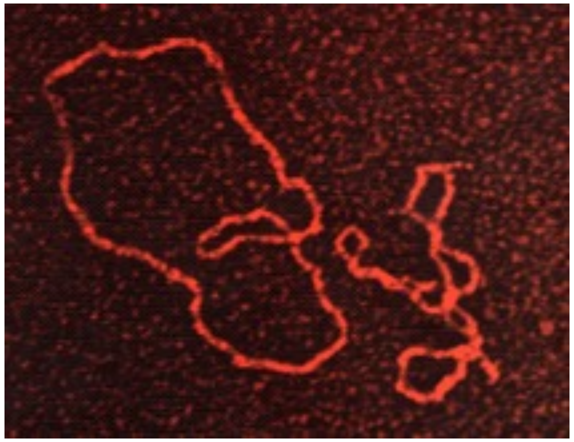


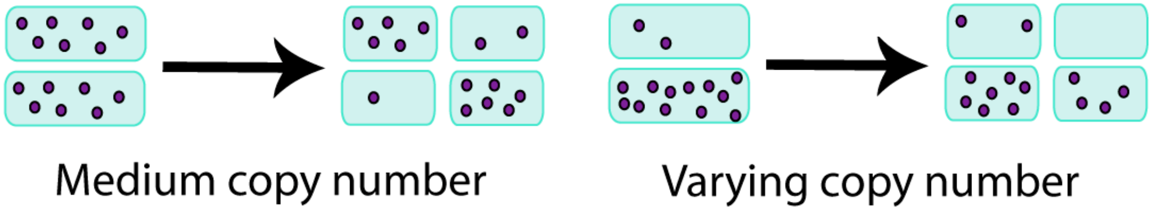


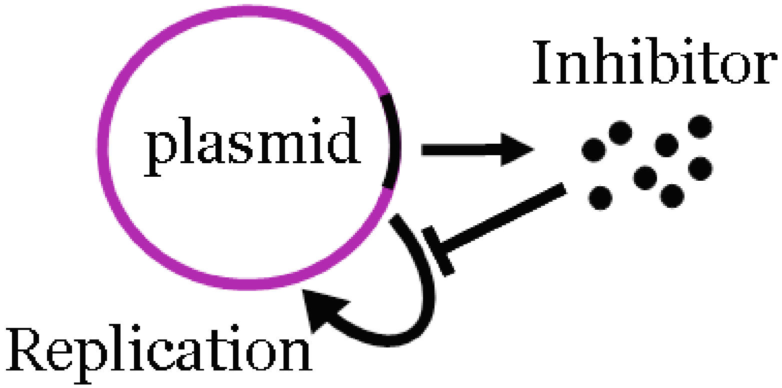

Protein
P
mRNA
Gene
... and so on

Minimizing variance: ColE1 replication control

- Approx 16 plasmid copies per cell
- Partitioned randomly at cell division
- Under strong selection for small variance





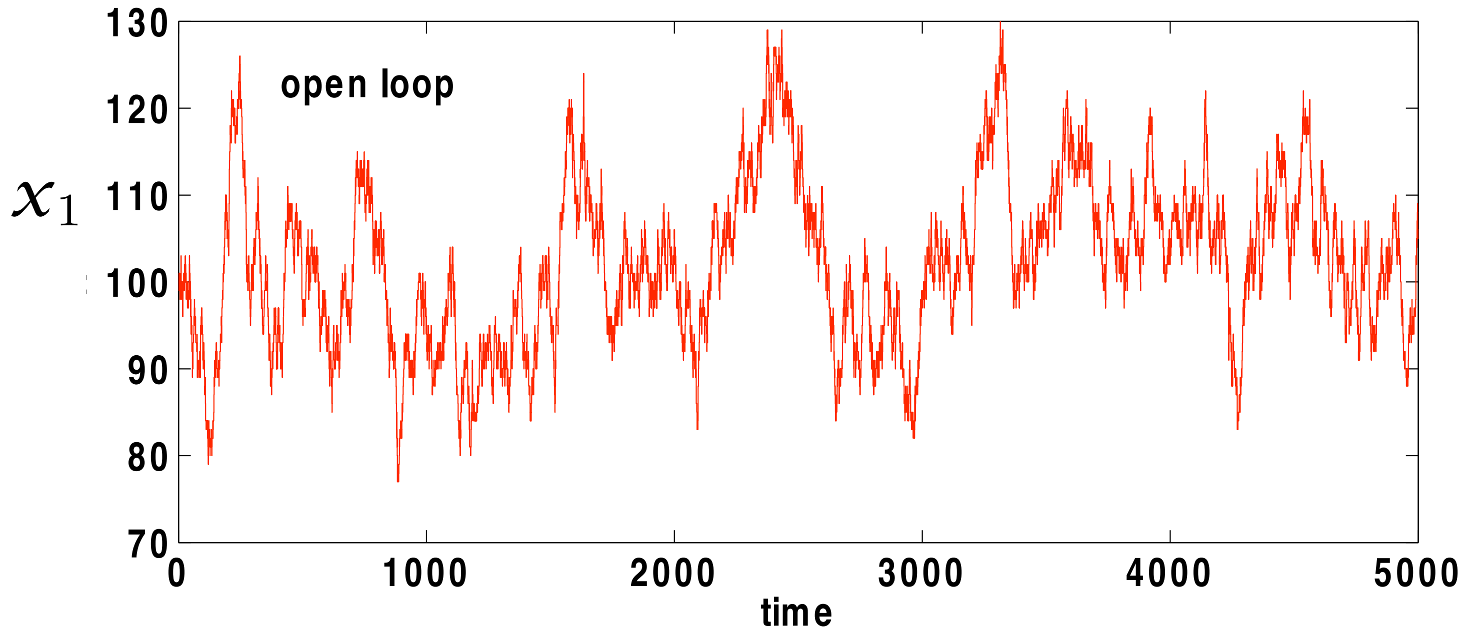


Regulation of molecule numbers

Consider a single species: e.g. mRNA of a constitutively expressed gene

$$x_1 \xrightarrow{c} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

$$\emptyset \underset{k_2}{\overset{k_1}{\leftrightarrows}} X_1$$



$$\sigma_1^2 = \langle x_1 \rangle$$

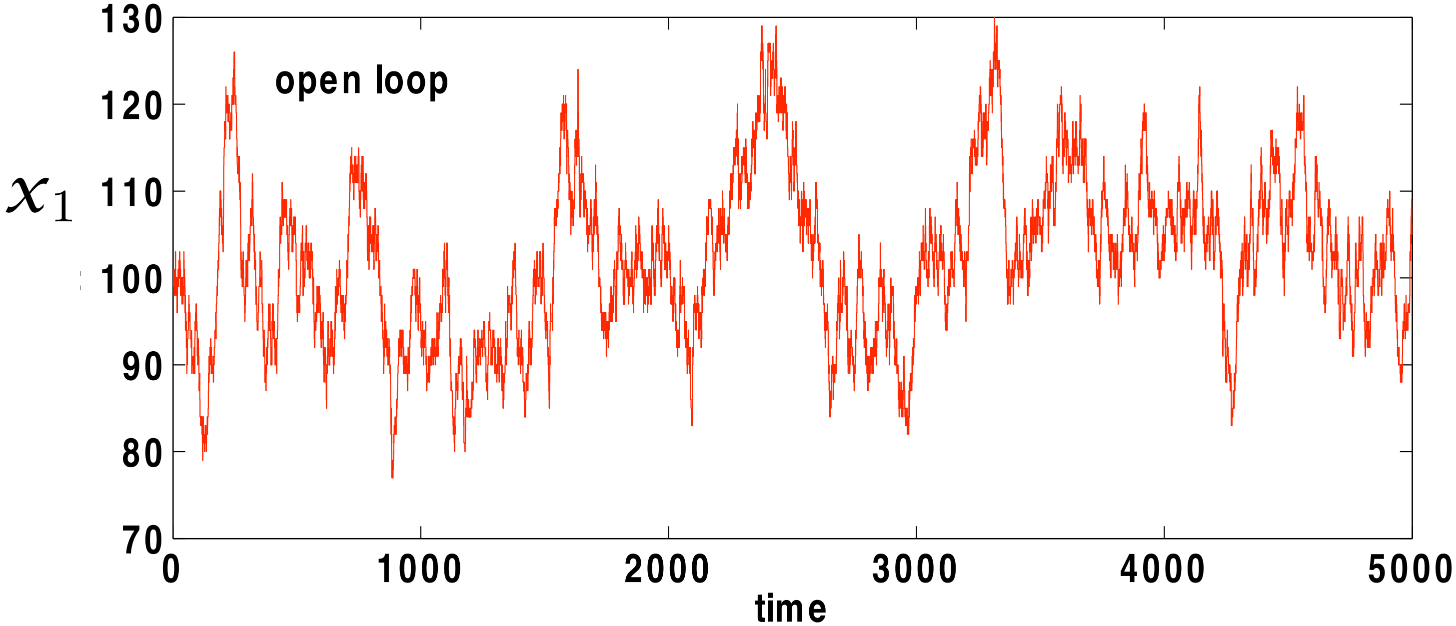
Regulation of molecule numbers

Consider a single species: e.g. mRNA of a constitutively expressed gene

$$x_1 \xrightarrow{c} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

$$\emptyset \underset{k_2}{\overset{k_1}{\leftrightarrows}} X_1$$

$$\Pr\big(x_1(t+dt) = N + 1 | x_1(t) = N\big) = c\,dt$$



$$\sigma_1^2 = \langle x_1 \rangle$$

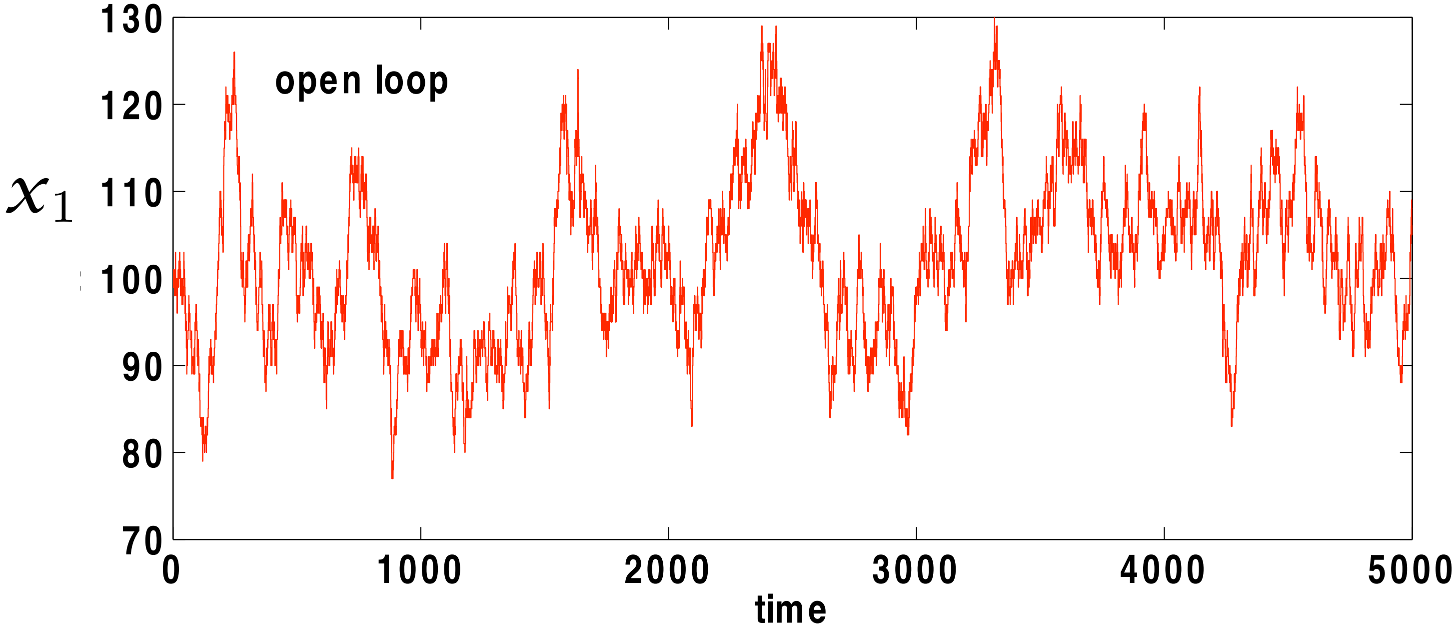
Regulation of molecule numbers

Consider a single species: e.g. mRNA of a constitutively expressed gene

$$x_1 \xrightarrow{c} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

$$\emptyset \underset{k_2}{\overset{k_1}{\leftrightarrows}} X_1$$

$$\Pr(x_1(t+dt) = N+1 | x_1(t) = N) = c\,dt$$
$$\Pr(x_1(t+dt) = N-1 | x_1(t) = N) = N/\tau_1\,dt$$



$$\sigma_1^2 = \langle x_1 \rangle$$

Regulation of molecule numbers

(with feedback)

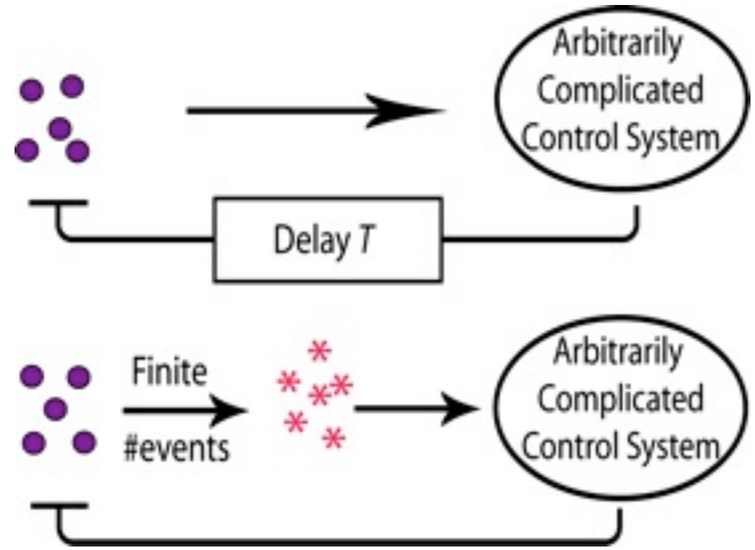
$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

where $u_t = f(\{x_1(t') : t' < t\})$

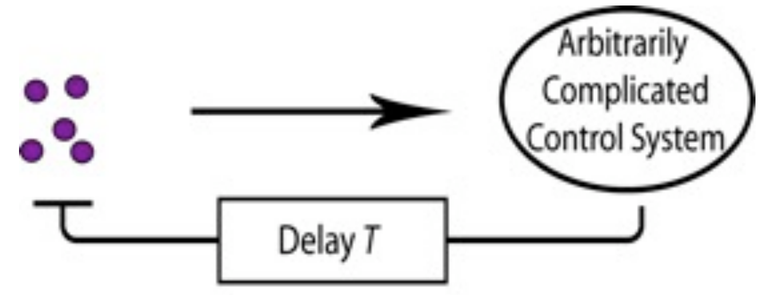
- Can make $\frac{\sigma_1^2}{\langle x_1 \rangle}$ arbitrarily small by appropriate choice of f.

However, limitations are imposed by:

- Delays
- Feedback mechanisms/capacity



Limitations due to delay



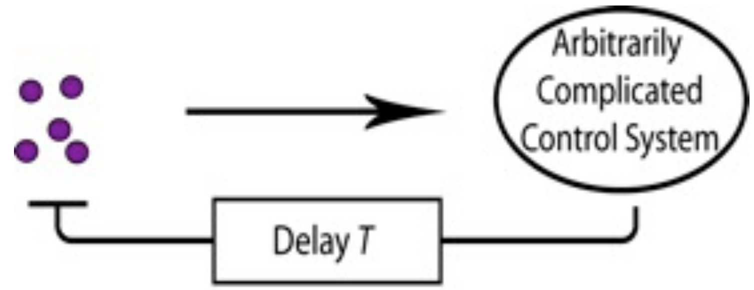
$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

where $u_t = f(\{x_1(t') : t' < t - T\})$

Theorem: If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq (1 - e^{-2T/\tau_1})$$

Limitations due to delay



$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

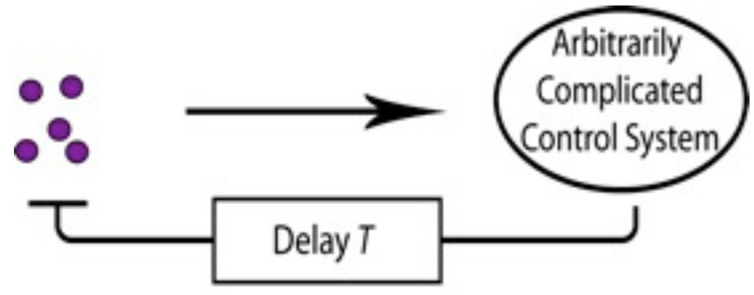
delay

where $u_t = f(\{x_1(t') : t' < t - T\})$

Theorem: If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq (1 - e^{-2T/\tau_1})$$

Limitations due to delay



$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

delay

where $u_t = f(\{x_1(t') : t' < t - T\})$

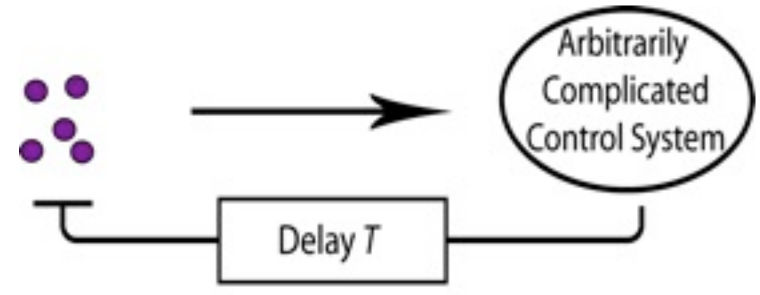
Theorem: If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq (1 - e^{-2T/\tau_1})$$

Proof: Let $\mathcal{I}_t = \{x_1(t'), t' < t - T\}$

$$\sigma_1^2 = E[(x_1 - E[x_1])^2] = E[E[x_1 - E[x_1])^2]|I_t]]$$
$$\geq E[E[(x_1 - E[x_1|I_t])^2|I_t]]$$
$$= E[E[x_1|\mathcal{I}_t] - x_1(t-T)e^{-2T/\tau_1}]$$

Limitations due to delay



$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

delay

where $u_t = f(\{x_1(t') : t' < t - T\})$

Theorem: If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq (1 - e^{-2T/\tau_1})$$

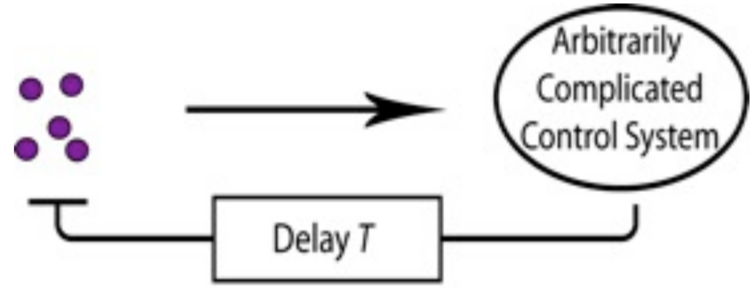
On the other hand, if

$$x_1 \xrightarrow{u_t x_1} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq 2T/\tau_1$$

Limitations due to delay



$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

delay

where $u_t = f(\{x_1(t') : t' < t - T\})$

Theorem: If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq (1 - e^{-2T/\tau_1})$$

replication

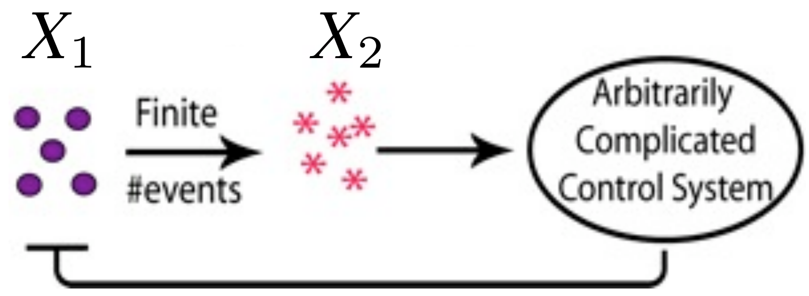
On the other hand, if

$$x_1 \xrightarrow{u_t x_1} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq 2T/\tau_1$$

Limitations due to mechanisms: molecular channels



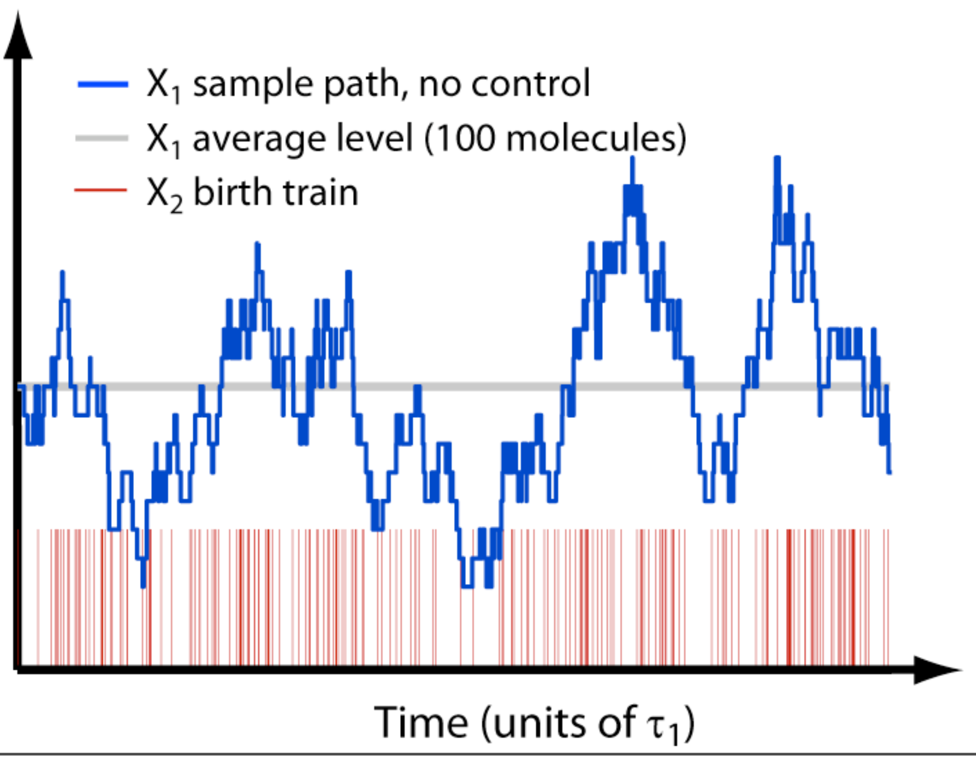
System:

$$x_1 \xrightarrow{u_t} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

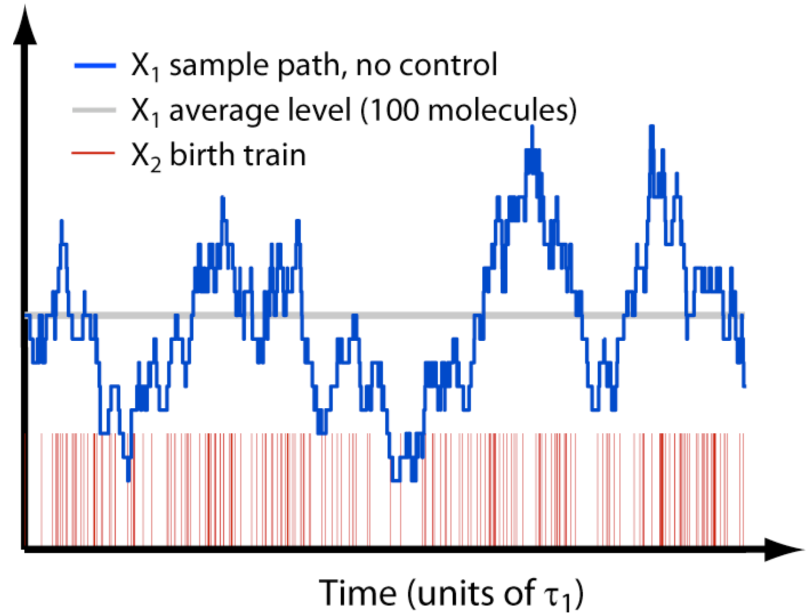
Sensor:

$$x_2 \xrightarrow{\alpha x_1} x_2 + 1$$
$$x_2 \xrightarrow{x_2/\tau_2} x_2 - 1$$

where $u_t = f(\{x_2(t') : t' < t\})$



What is the capacity of a molecular channel?



Consider the channel:

$$x_2 \xrightarrow{z} x_2 + 1$$

$$x_2 \xrightarrow{x_2/\tau_2} x_2 - 1$$

- Capacity is related to that for a photon counting channel (but depends on what constraints are put on z)
- One answer is $C = \langle z \rangle \log\left(1 + \frac{\sigma_z^2}{\langle z \rangle^2}\right) \leq \frac{\sigma_z^2}{\langle z \rangle}$ (nat/s)

$= 1.443\sigma_z^2/\langle z \rangle$ bit/s

(where $C = \max_z I(x_2; z)$)

Feedback capacity and variance

Taking the diffusion approximation of the replication case:

$$x_1 \xrightarrow{u_t x_1} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

$$dx_1 = u_t \langle x_1 \rangle dt + \sqrt{2\langle x_1 \rangle / \tau_1}\, dw$$

gives (Gorbunov and Pinsker '74)

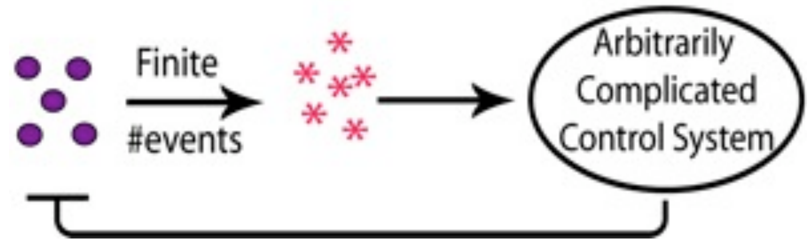
$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq \frac{1}{C\tau_1} \qquad (\text{for } I(u; x_1) \leq C)$$

Putting this together with the bound on C gives:

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq \sqrt{\frac{N_1}{N_2}}$$

where $N_2 = \langle x_2 \rangle \tau_1 / \tau_2 =$ no of molecules of X_2 made per lifetime of X_1.
$N_1 = \langle x_1 \rangle =$ no of molecules of X_1 made per lifetime of X_1.

Summary: Limitations due to channel capacity



System: $x_1 \xrightarrow{u_t x_1} x_1 + 1$, $x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$ $\Rightarrow$ $dx_1 = u_t\langle x_1\rangle dt + \sqrt{2\langle x_1\rangle/\tau_1}\, dw$

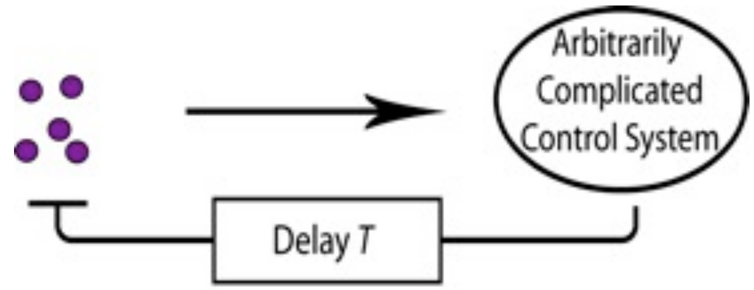
Sensor: $x_2 \xrightarrow{\alpha x_1} x_2 + 1$, $x_2 \xrightarrow{x_2/\tau_2} x_2 - 1$

where $u_t = f(\{x_2(t') : t' < t\})$

$$\frac{\sigma_1^2}{\langle x_1\rangle} \geq \sqrt{\frac{N_1}{N_2}}$$

where $N_2 = \langle x_2\rangle \tau_1/\tau_2 =$ no of molecules of X_2 made per lifetime of X_1.
$N_1 = \langle x_1\rangle =$ no of molecules of X_1 made per lifetime of X_1.

Summary: Limitations due to delay



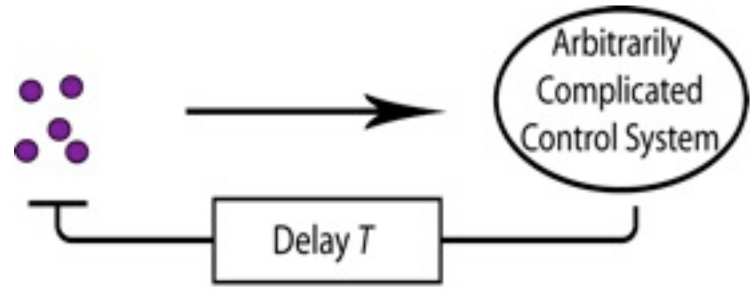
$$x_1 \xrightarrow{u_t x_1} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

where $u_t = f\left(\{x_1(t') : t' < t - T\}\right)$

If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq 2\frac{T}{\tau_1}$$

Summary: Limitations due to delay



$$x_1 \xrightarrow{u_t x_1} x_1 + 1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1 - 1$$

where $u_t = f(\{x_1(t') : t' < t - T\})$

If x_1 is a stationary process then

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq 2\frac{T}{\tau_1}$$

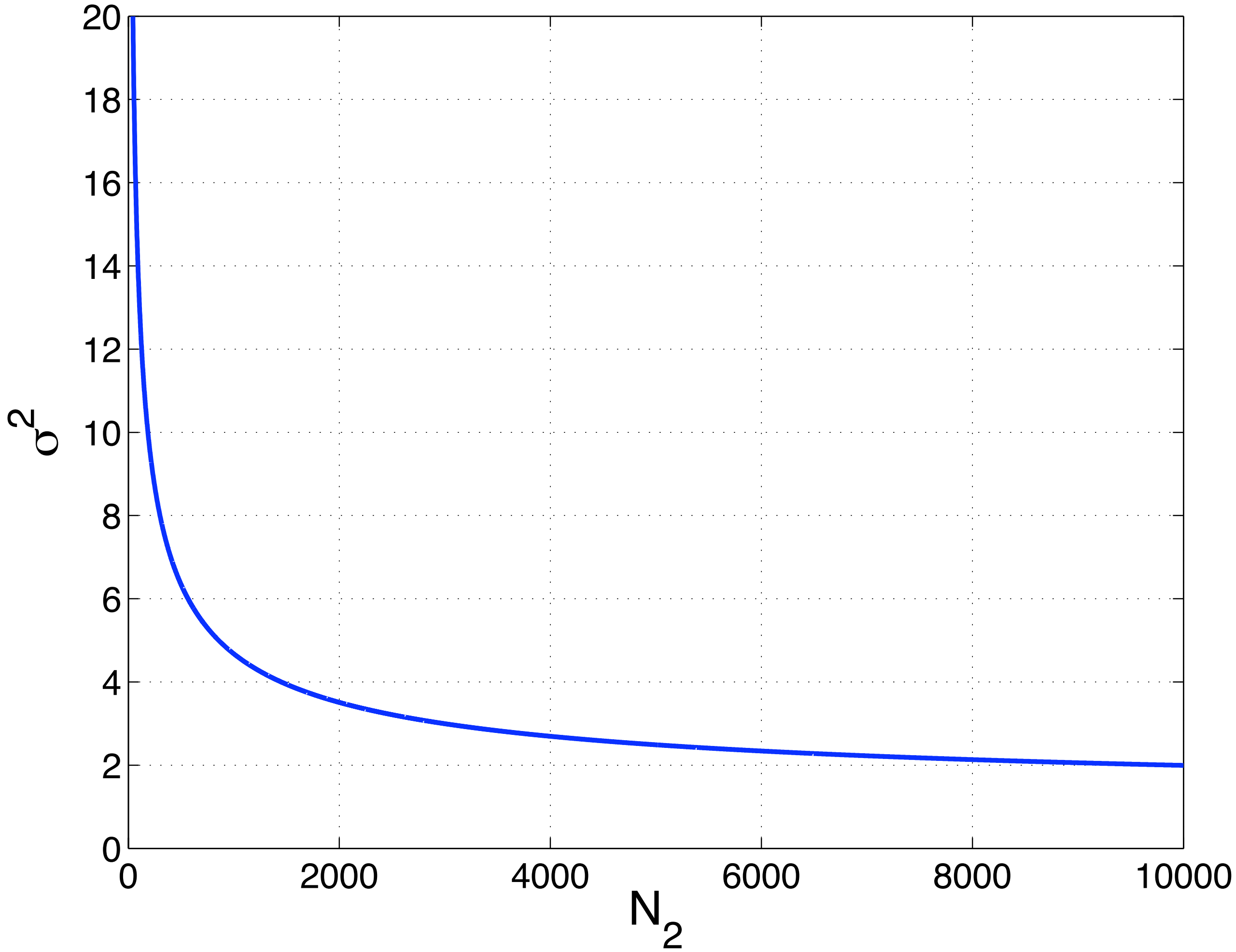
Can combine bounds:

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq \frac{T}{\tau_1} + \sqrt{\frac{N_1}{N_2} + \left(\frac{T}{\tau_1}\right)^2}$$

Some numbers

For ColE1, assuming a mean of 16 copies immediately after cell division

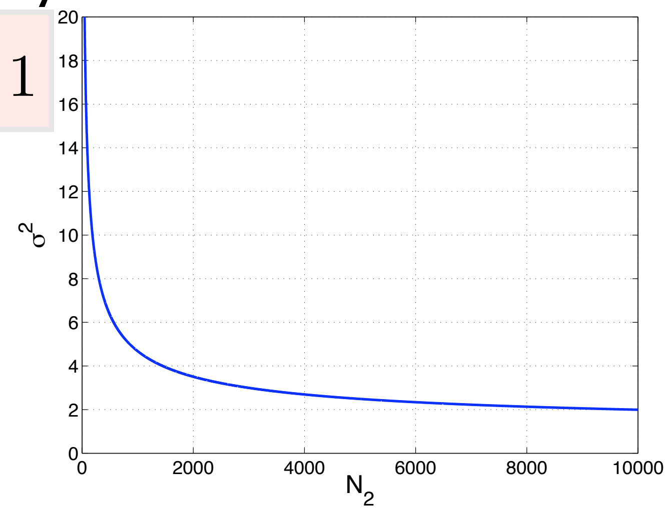
- Lower bound due to delays only ($T = 44$): $\sigma^2 \gtrsim 1$
- 10000 inhibitors/cell cycle: $\sigma^2 \gtrsim 2$



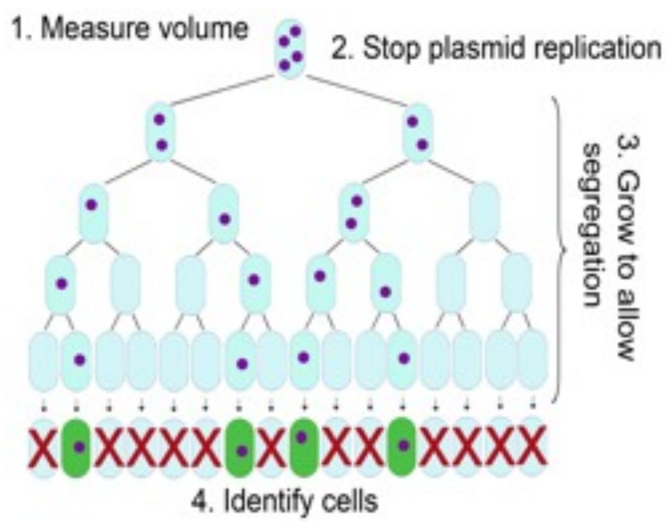
Some numbers

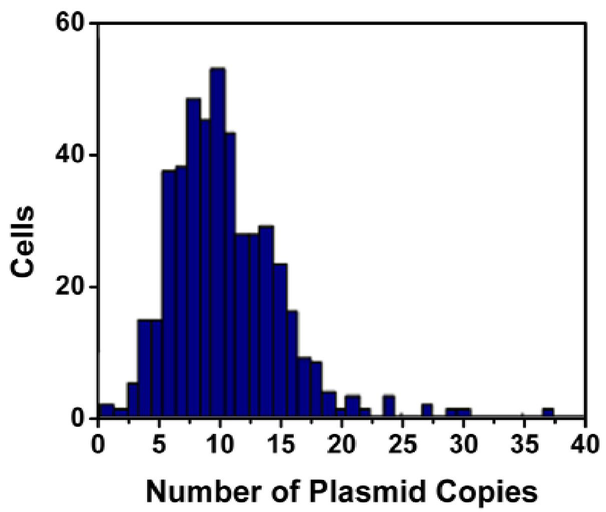
For ColE1, assuming a mean of 16 copies immediately after cell division

- Lower bound due to delays only ($T = 44$): $\sigma^2 \gtrsim 1$
- 10000 inhibitors/cell cycle: $\sigma^2 \gtrsim 2$



- Early experimental results (Paulsson) suggest $\sigma^2 < 4$ (*although for a different plasmid with the same copy number*).





Extensions

Similar results hold for non-replication case (e.g. transcription/translation:

$$x_1 \xrightarrow{u_t} x_1 + 1$$

e.g.

$$\frac{\sigma_1^2}{\langle x_1 \rangle} \geq \frac{2}{1+\sqrt{1+4N_2/N_1}}, \quad 1 - \exp(-2T/\tau_1)$$

and for nonlinear use of channel:

$$x_2 \xrightarrow{f(x_1)} x_2 + 1$$

(e.g. Hill functions $f(x_1) = v\frac{x_1^H}{K + x_1^H}$)

same bounds, but with $N_2 \rightarrow \gamma N_{2\mathrm{max}}$

Conclusions

- The ultimate bounds on feedback performance due to delays and finite numbers of synthesis events are sharp and appear biologically relevant.
- Biological questions are inspiring new theory here!

Conclusions

- The ultimate bounds on feedback performance due to delays and finite numbers of synthesis events are sharp and appear biologically relevant.
- Biological questions are inspiring new theory here!

  - and new theory is now also inspiring biological questions!