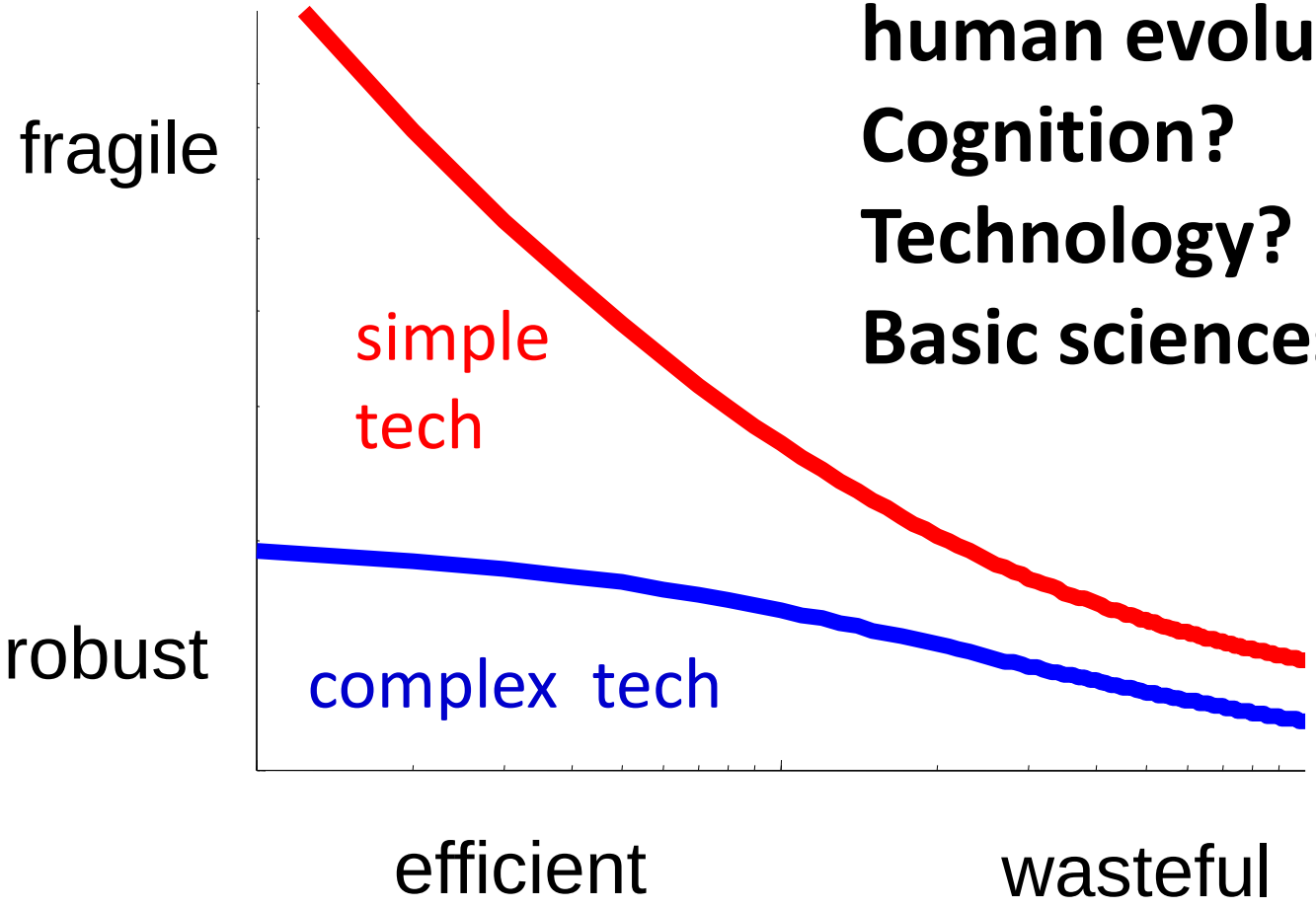
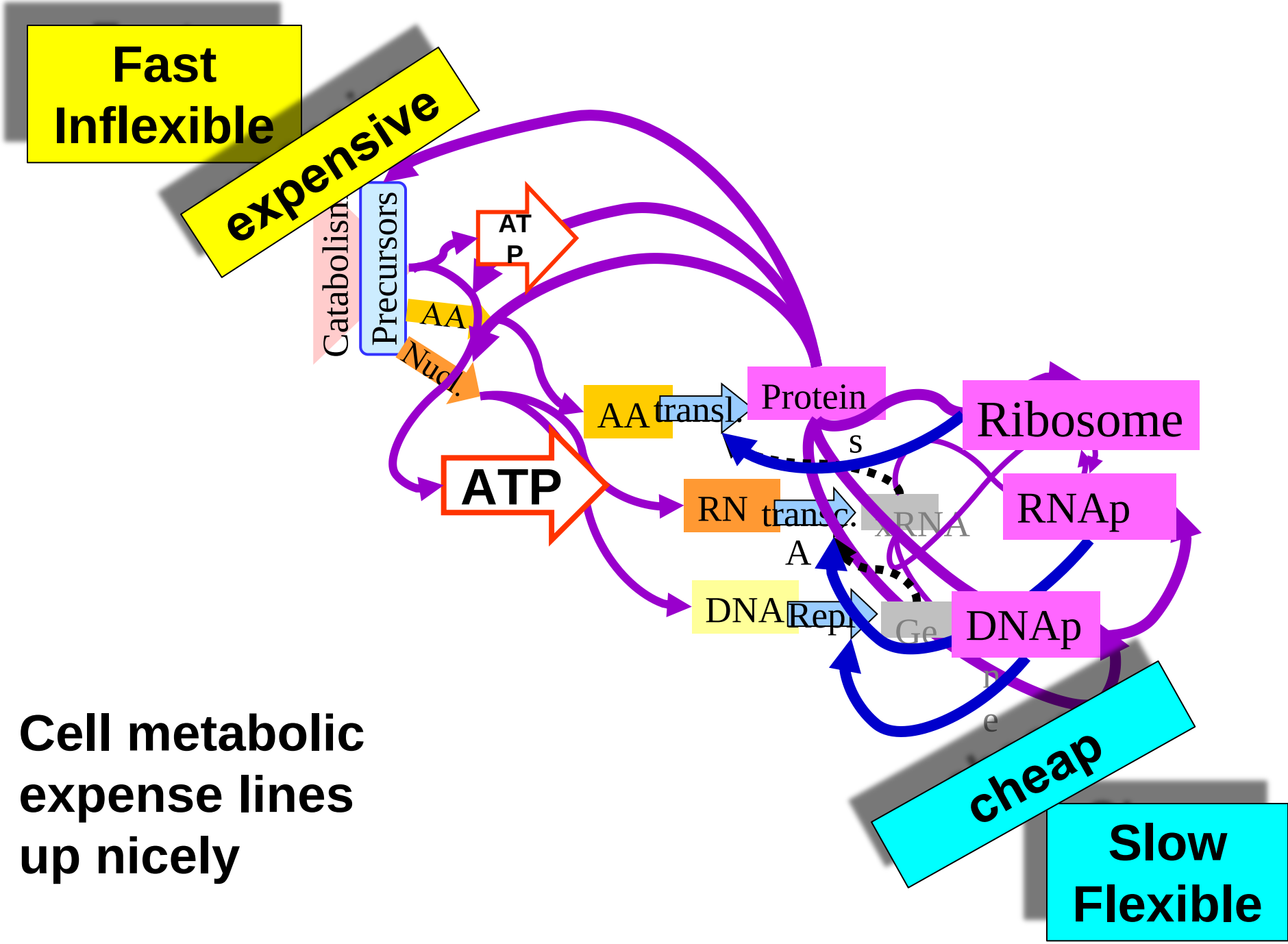
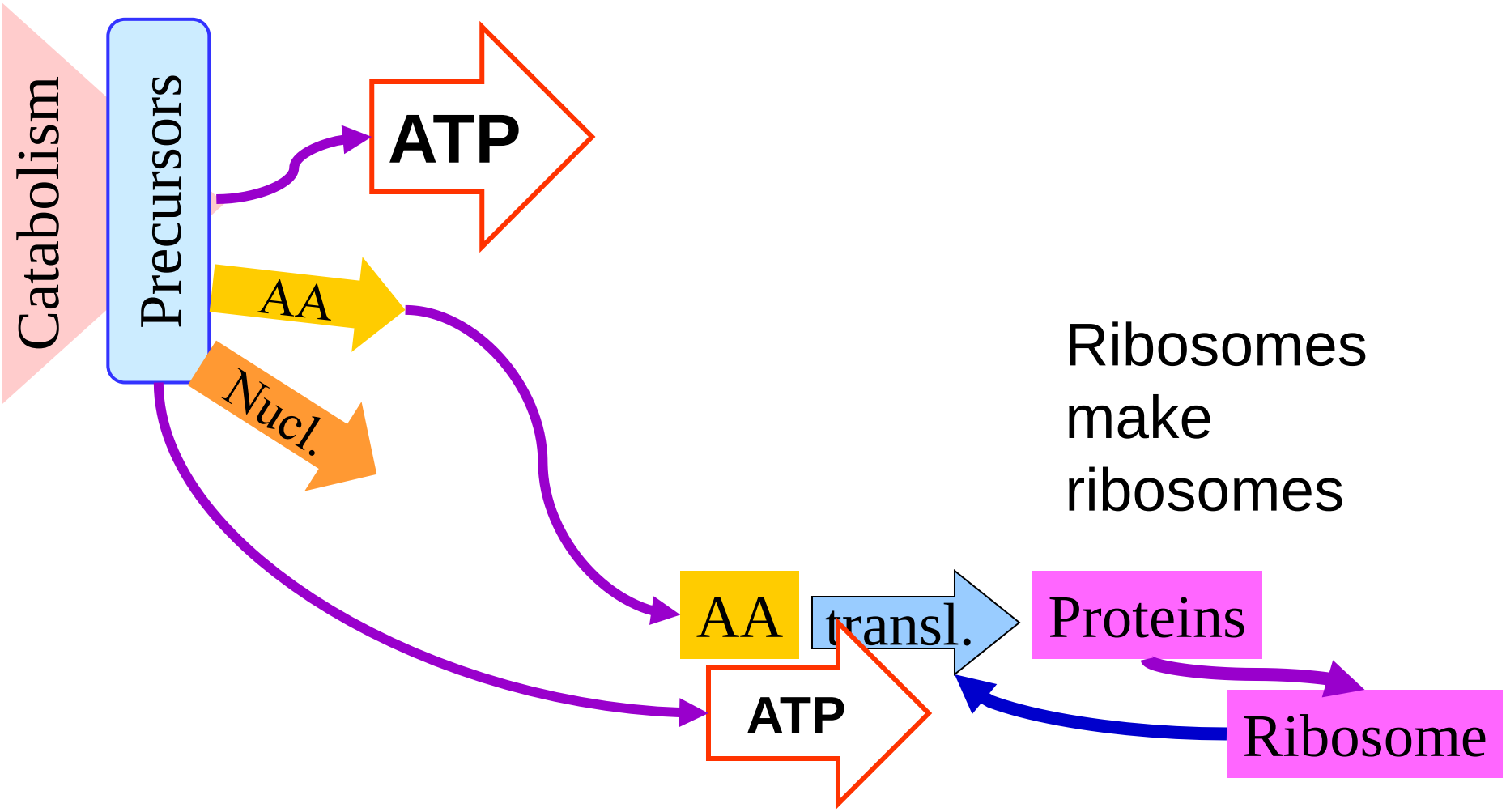
How general is this picture?



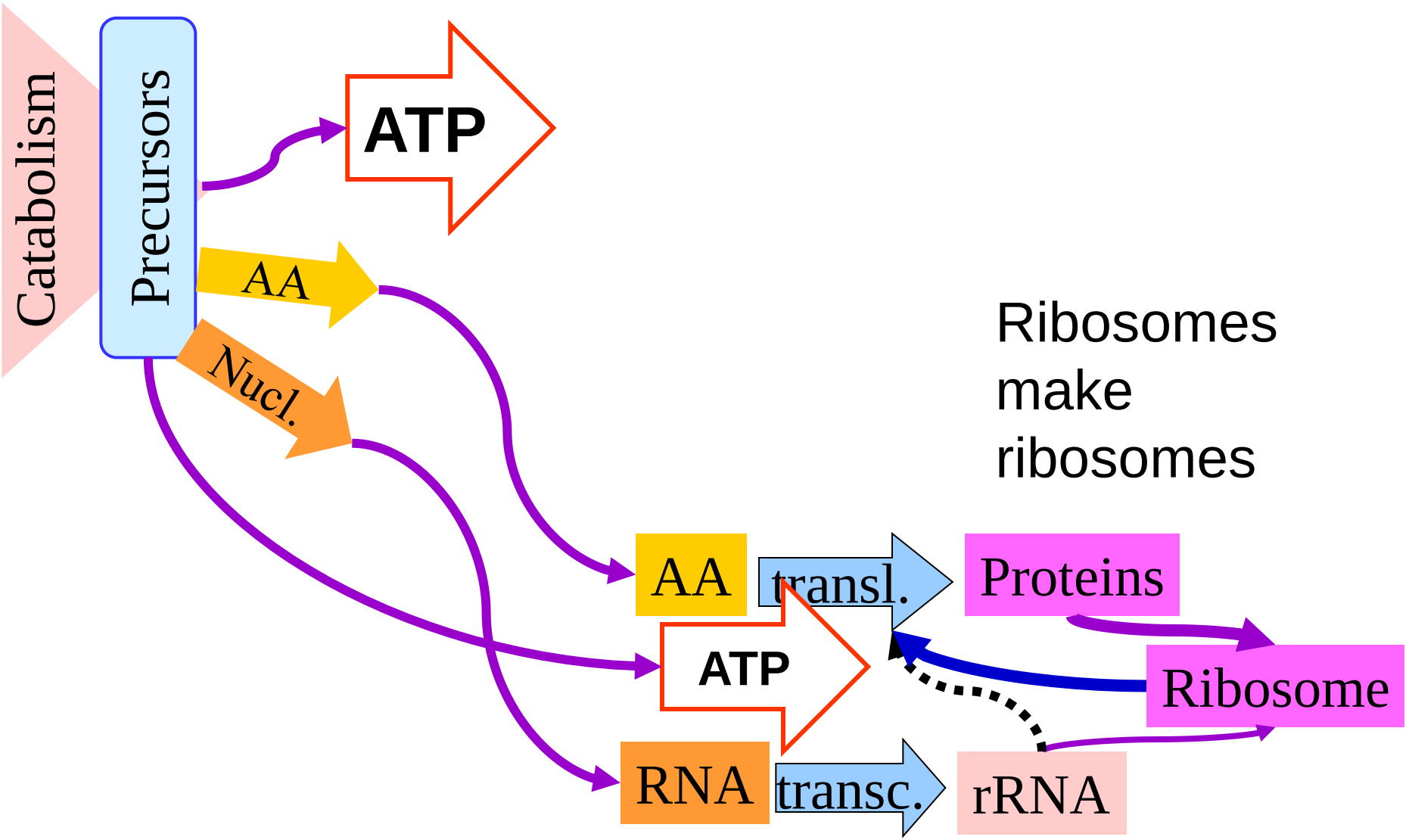
Implications for human evolution?
Cognition?
Technology?
Basic sciences?



Cell metabolic expense lines up nicely



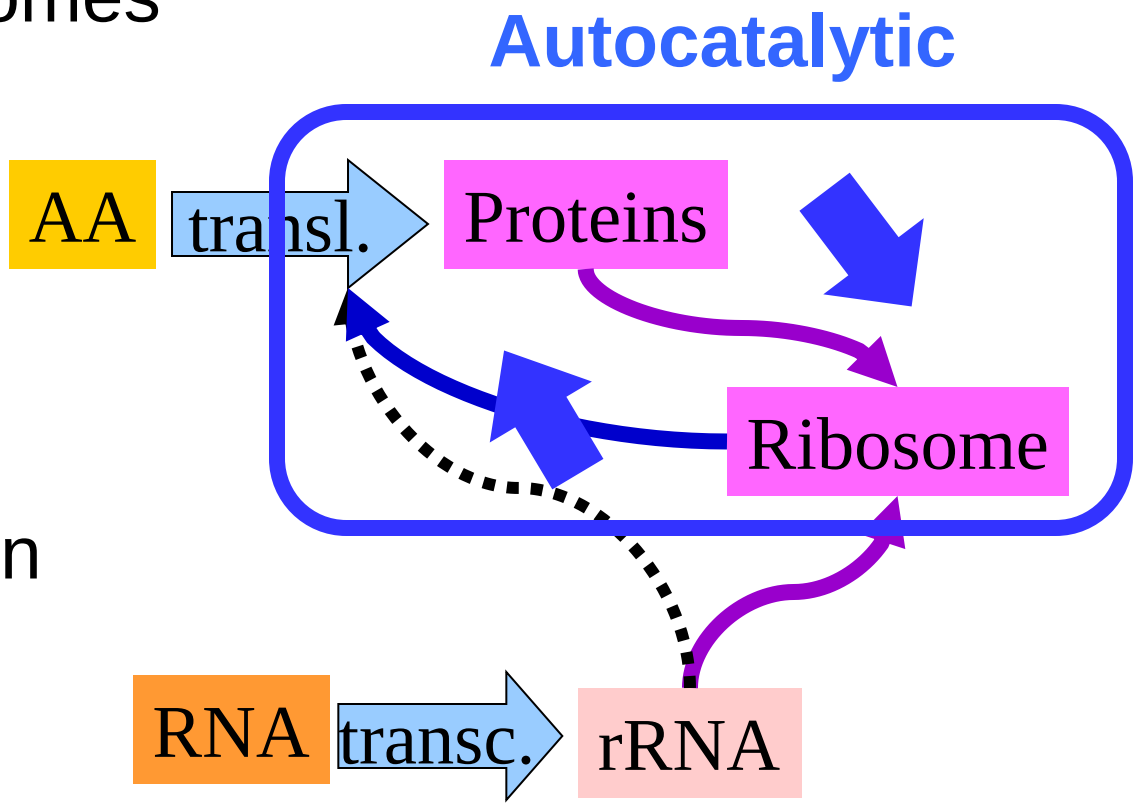
Translation: Amino acids polymerized into proteins



Ribosomes make ribosomes

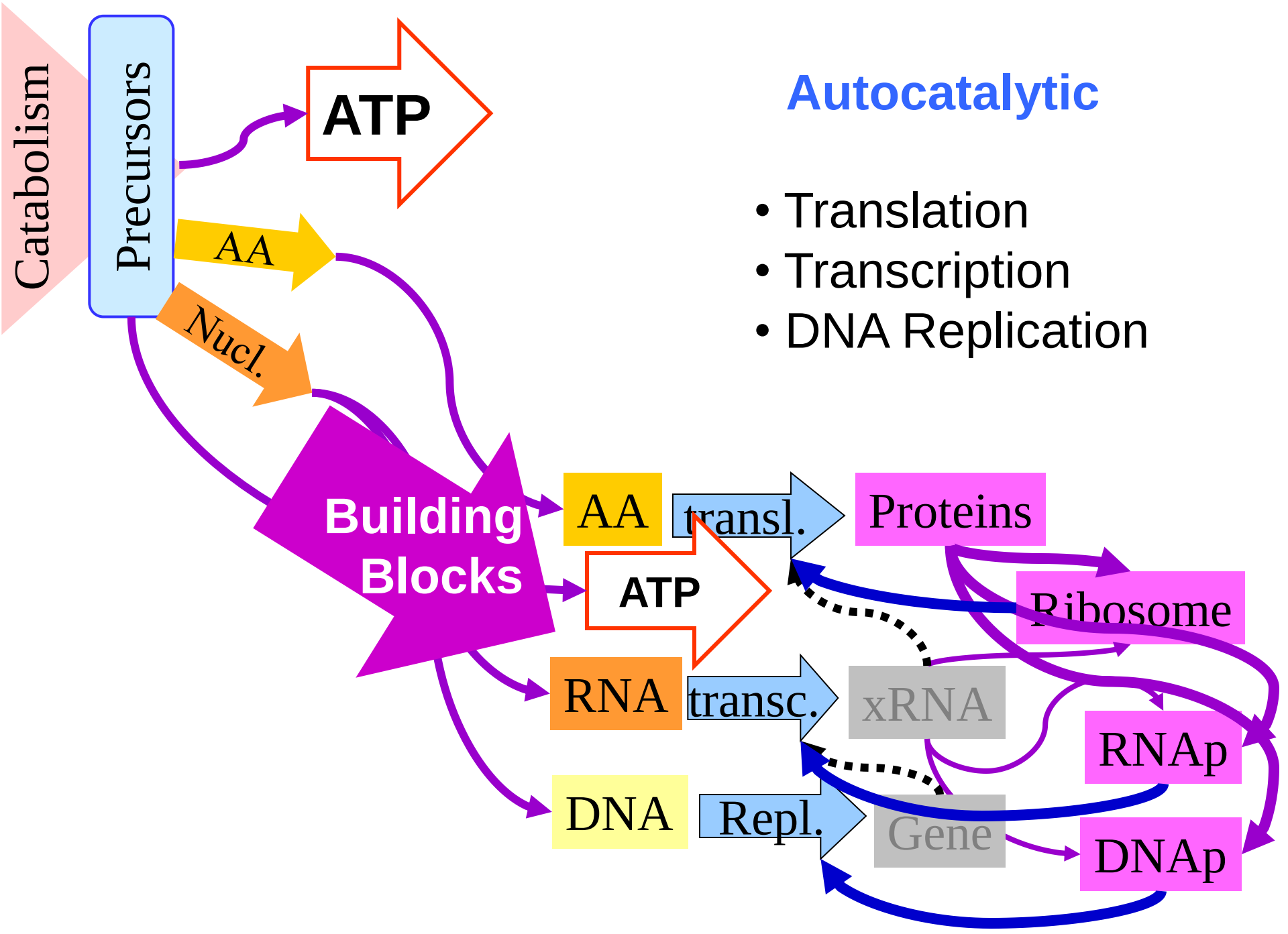
Ribosomes are made of proteins and RNA

Ribosomes
make
ribosomes



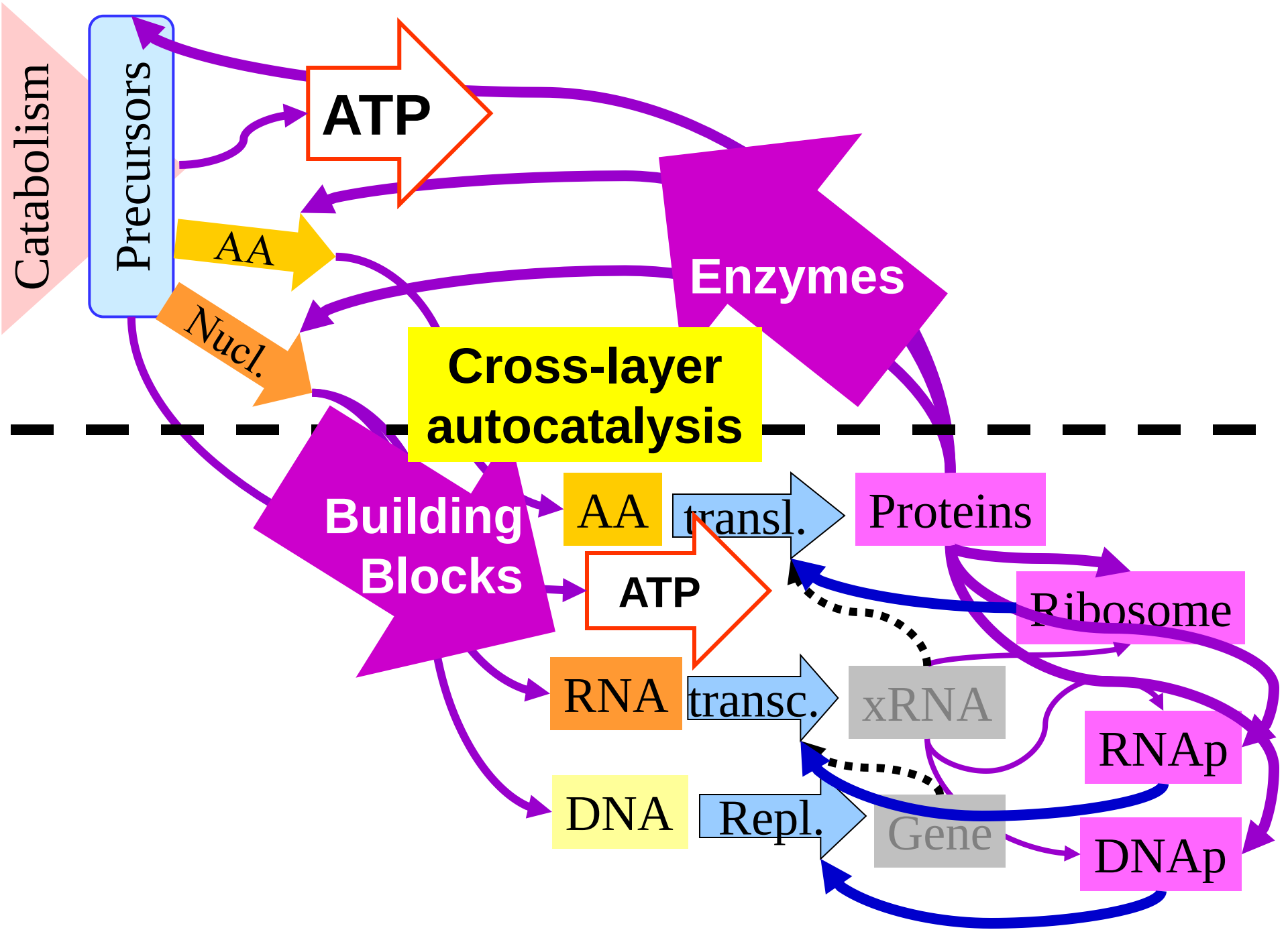
Organisms differ in the proportion of ribosomal protein vs rRNA

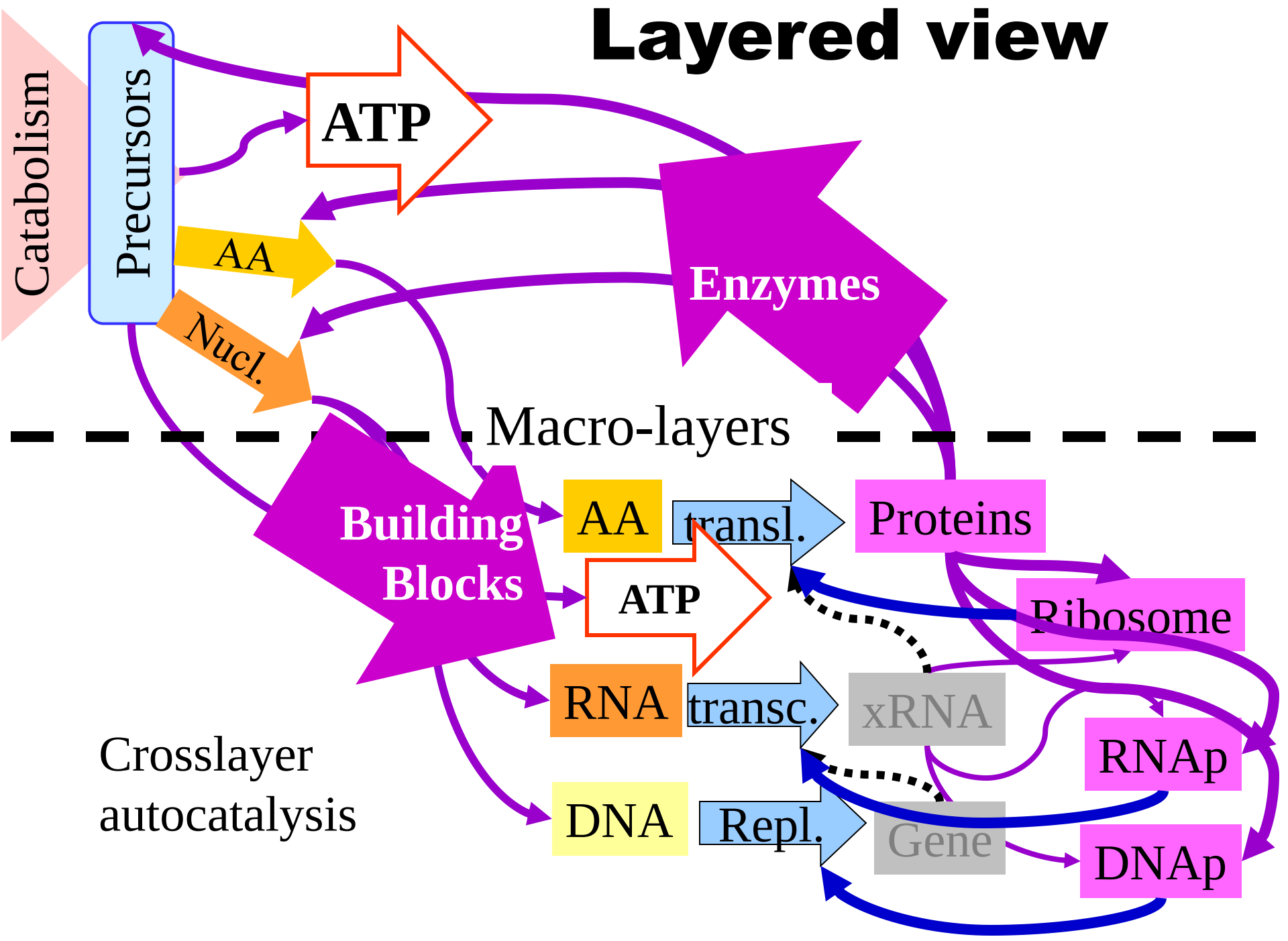
Ribosomes are made of proteins and rRNA



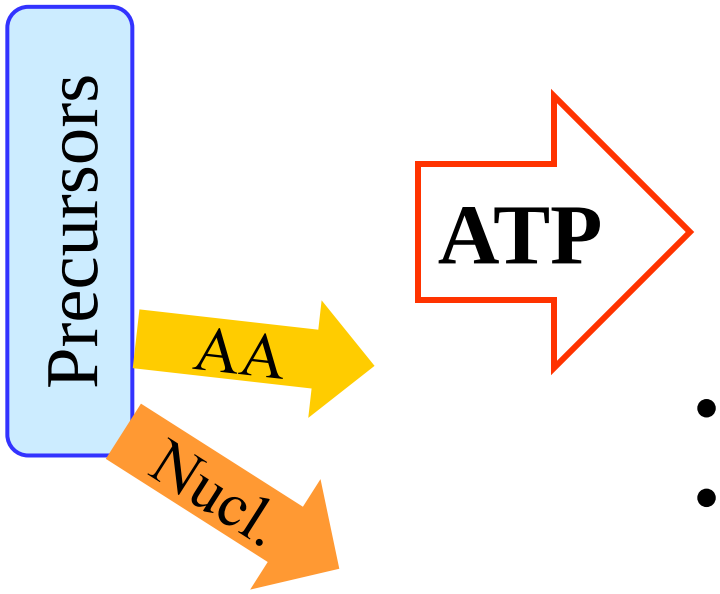
Autocatalytic

- Translation
- Transcription
- DNA Replication

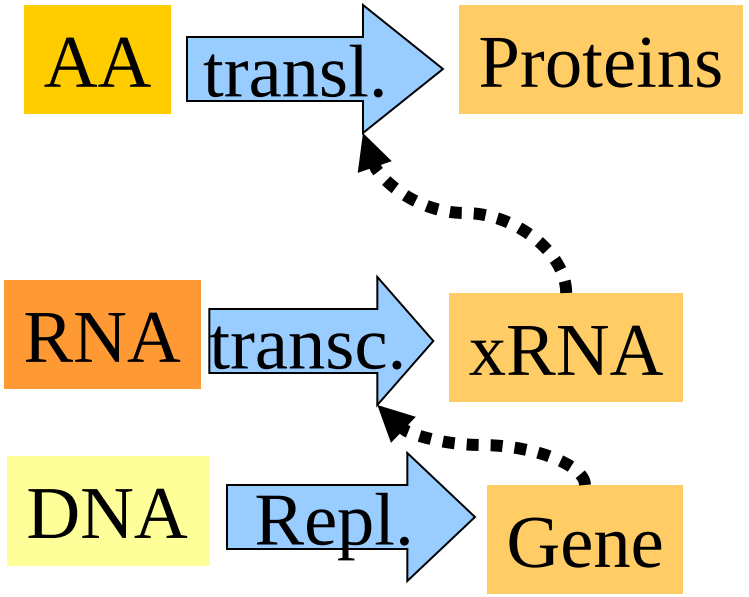

Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Cross-layer autocatalysis
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp

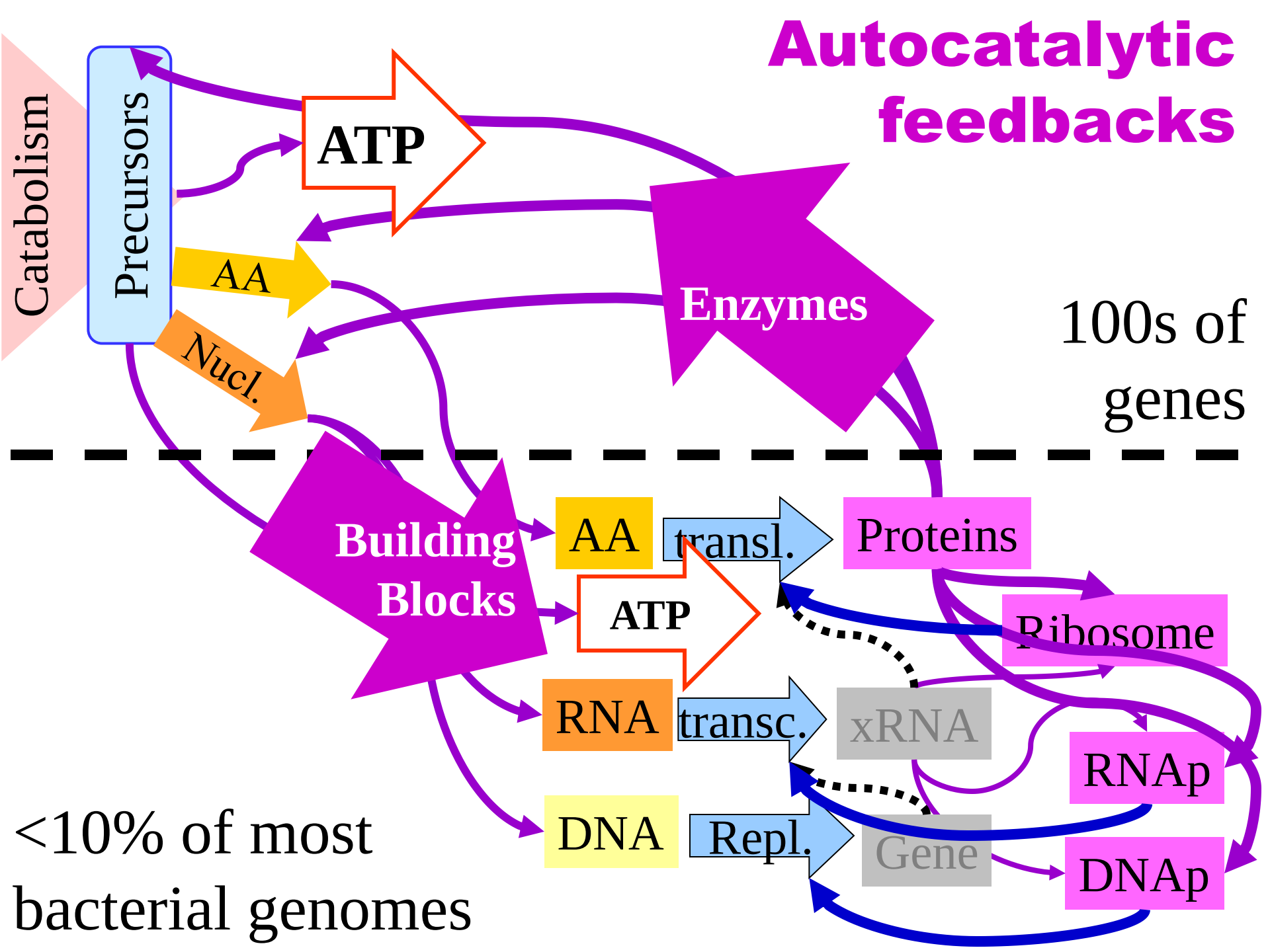

Layered view
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Macro-layers
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
Crosslayer autocatalysis

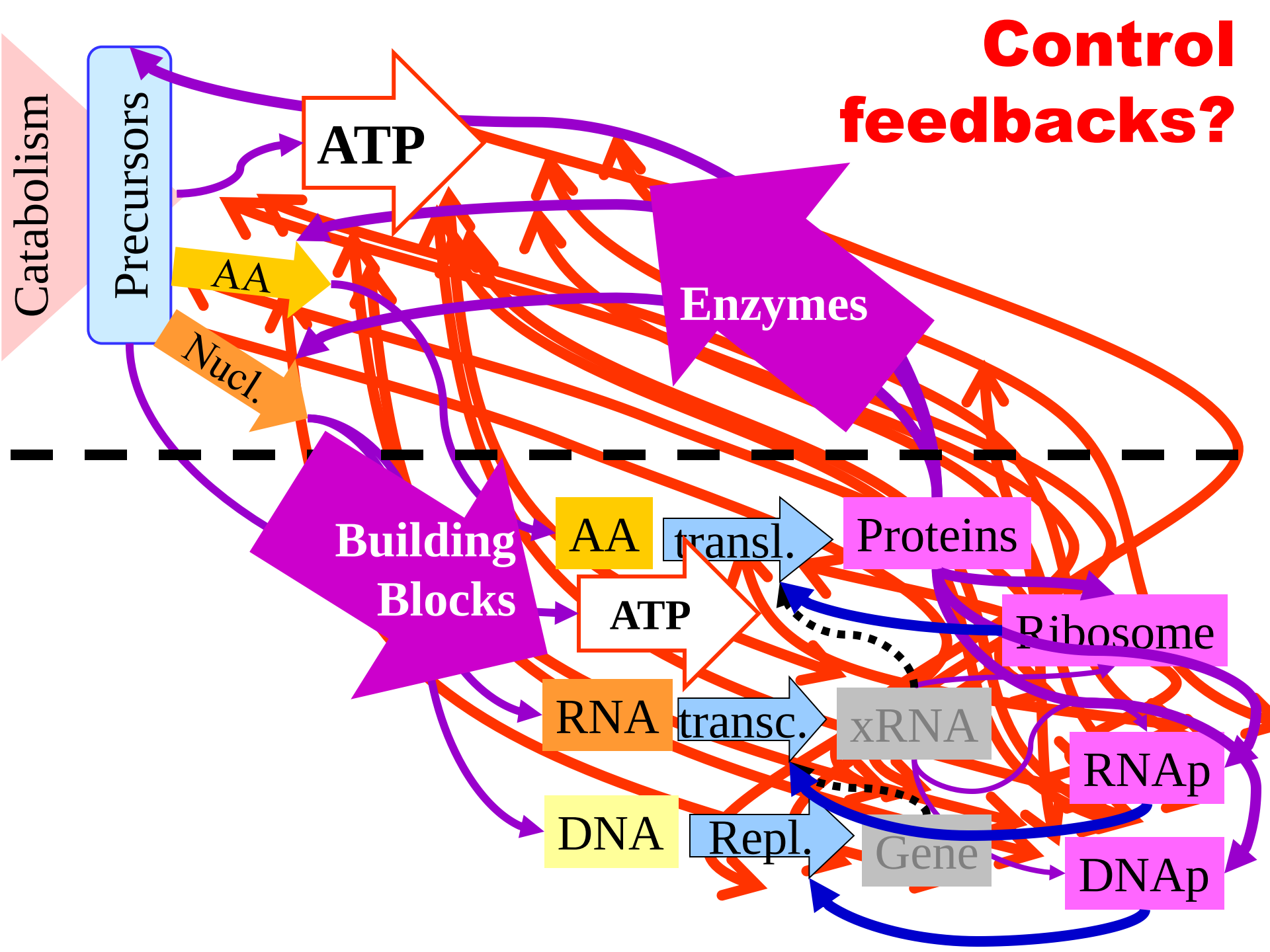
Shared protocols

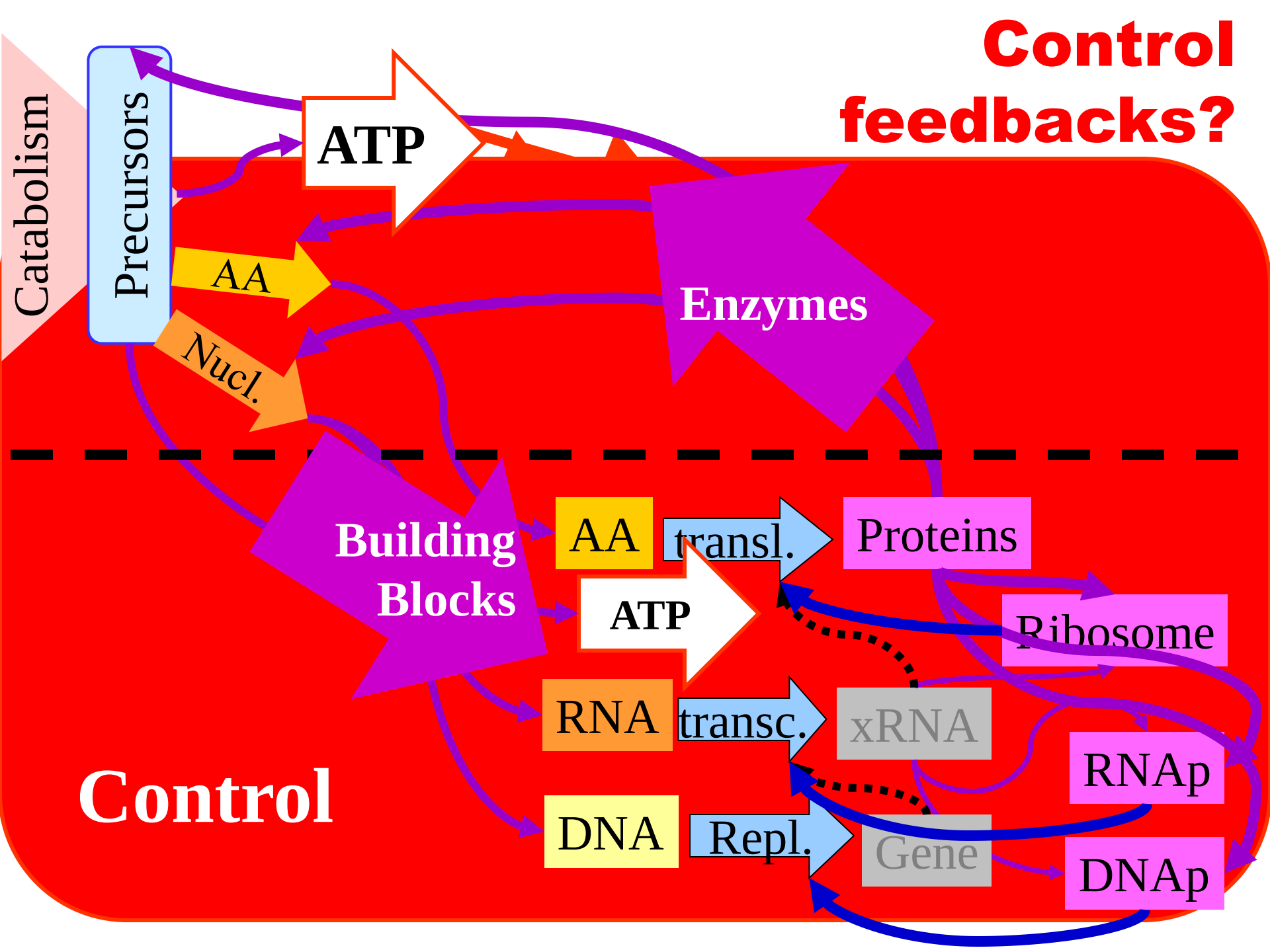


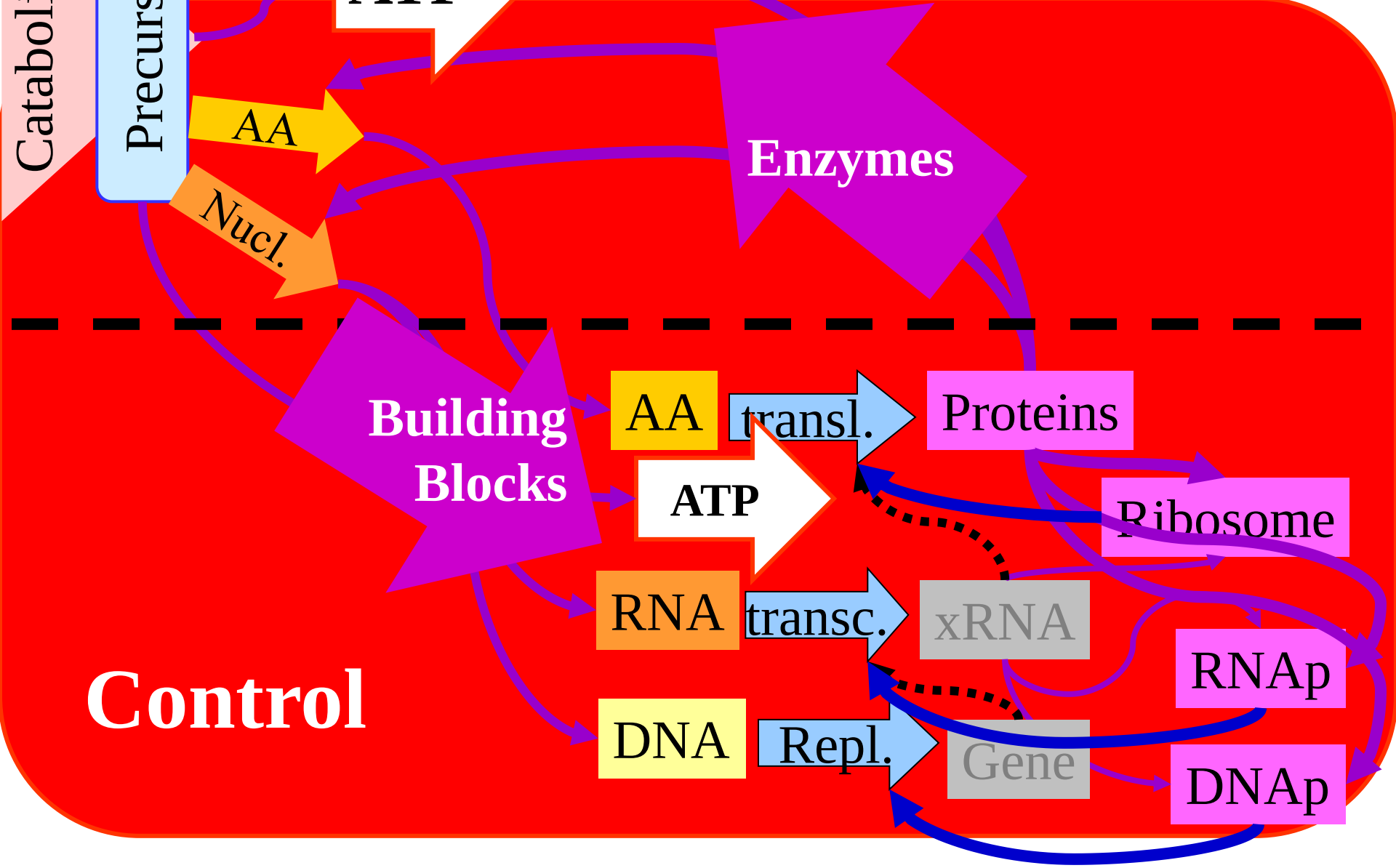
- Universal core constraints
- “virtual machines”



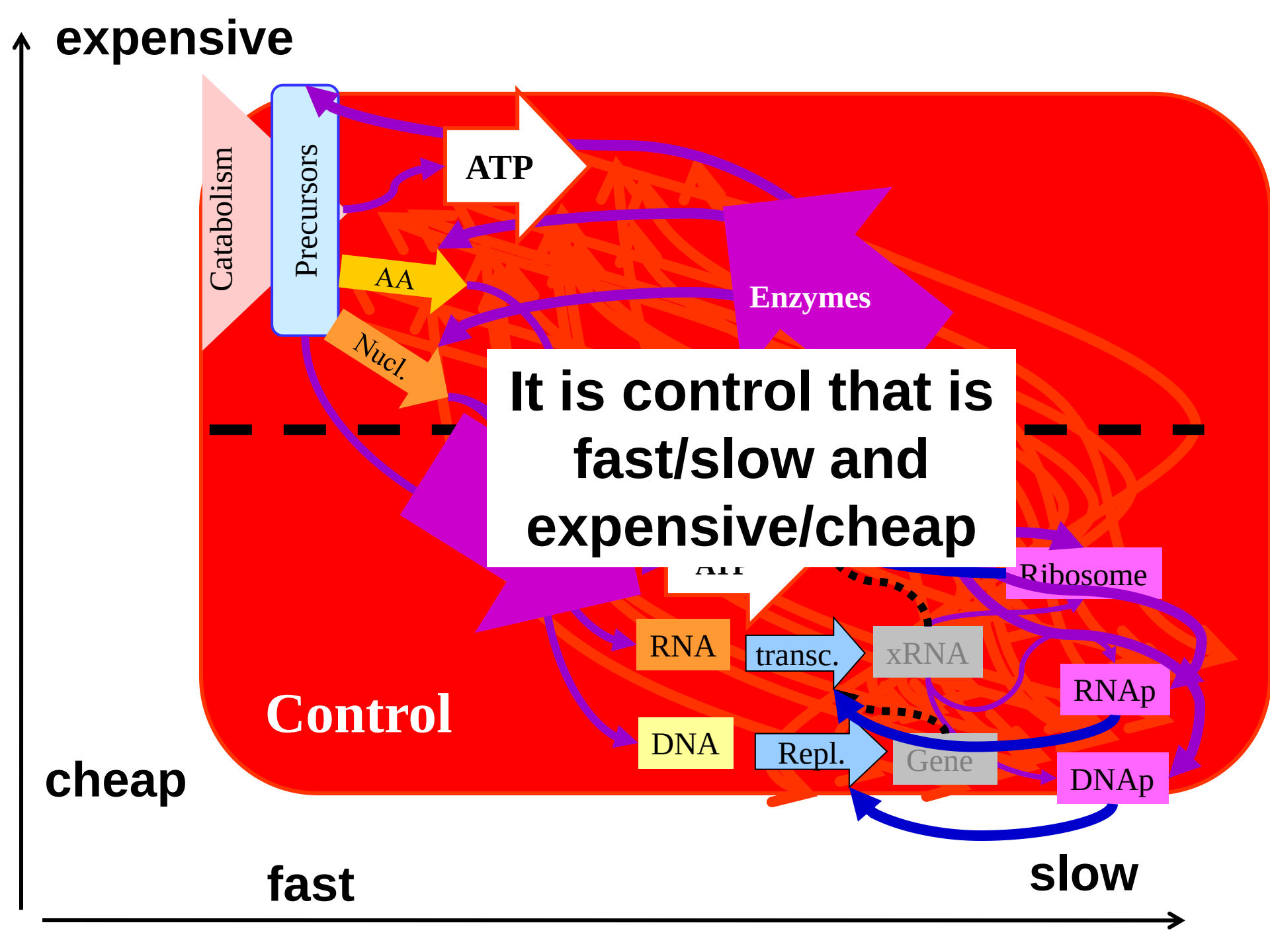

Autocatalytic feedbacks
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
100s of genes
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
<10% of most bacterial genomes

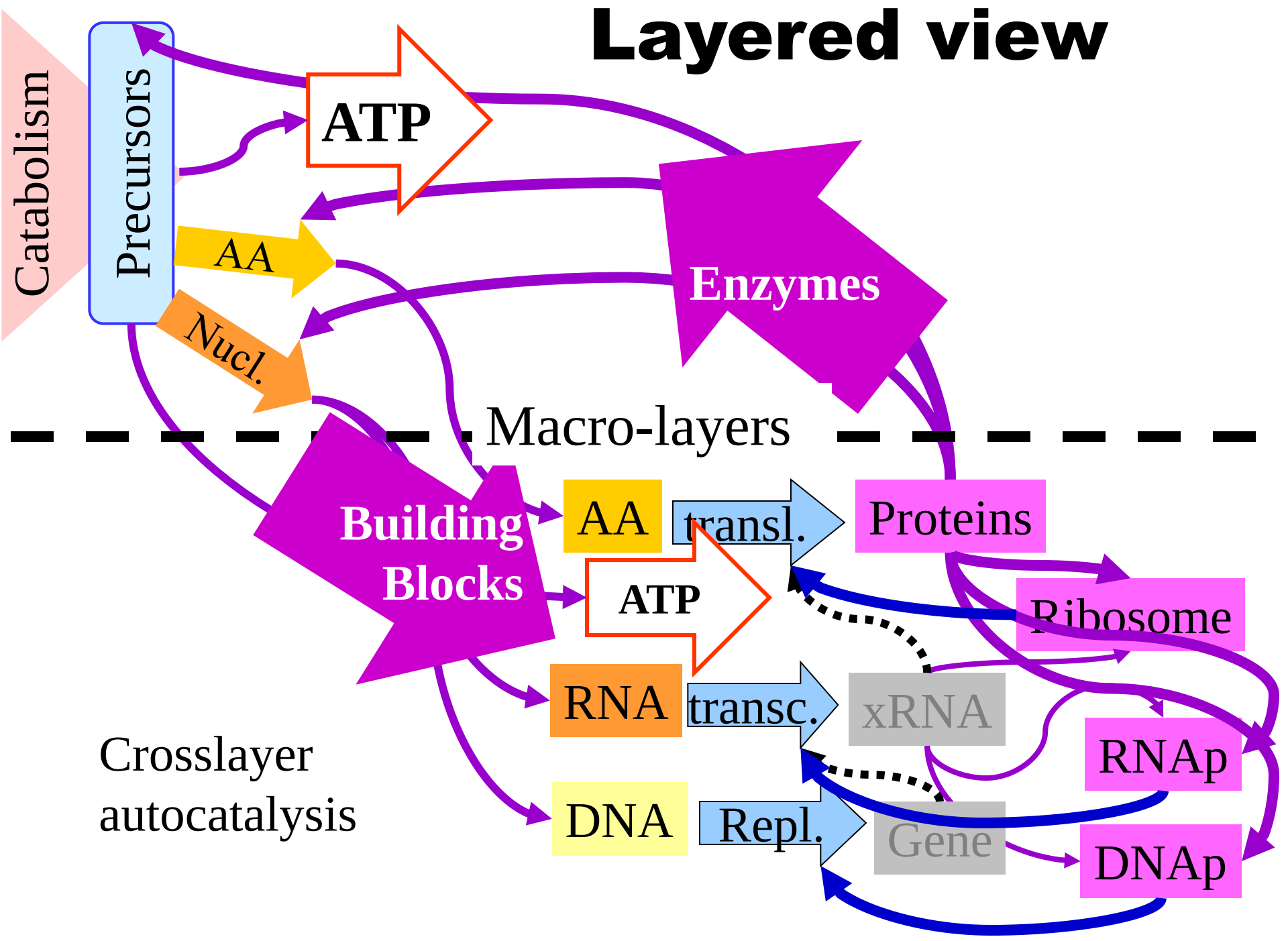

Control feedbacks?
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp

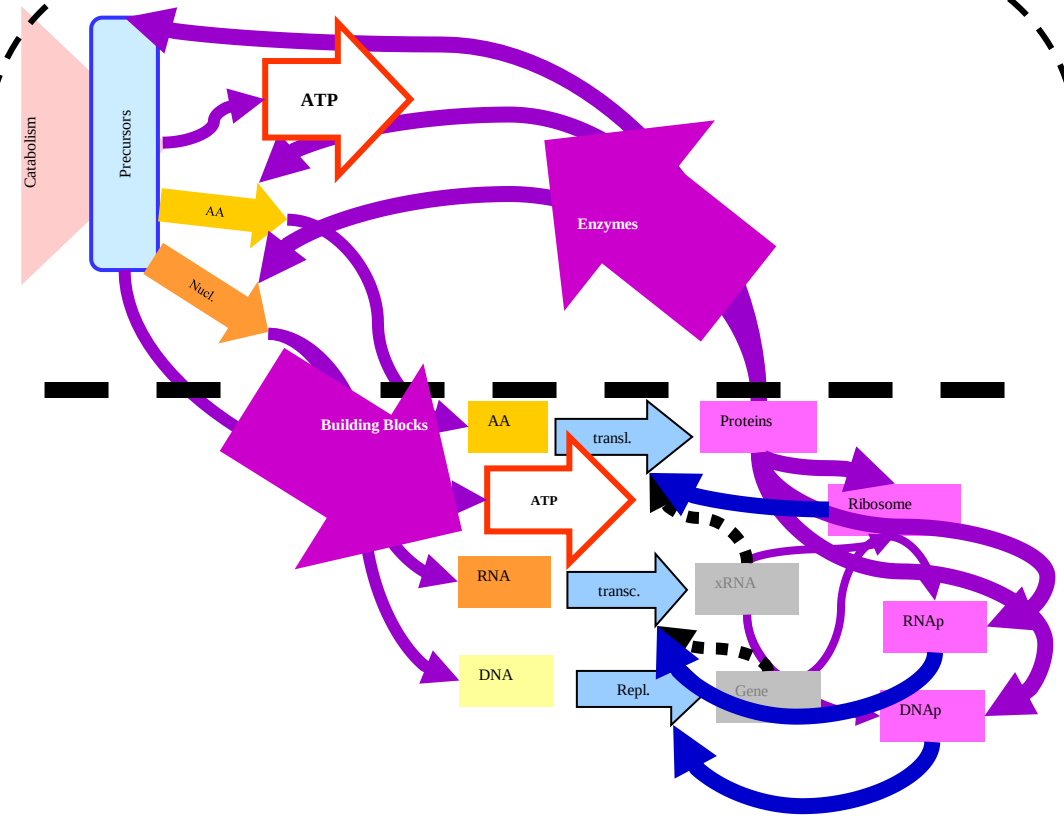

Control feedbacks?
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
Control
DNA
Repl.
Gene
DNAp



Complexity of control is huge and poorly studied.

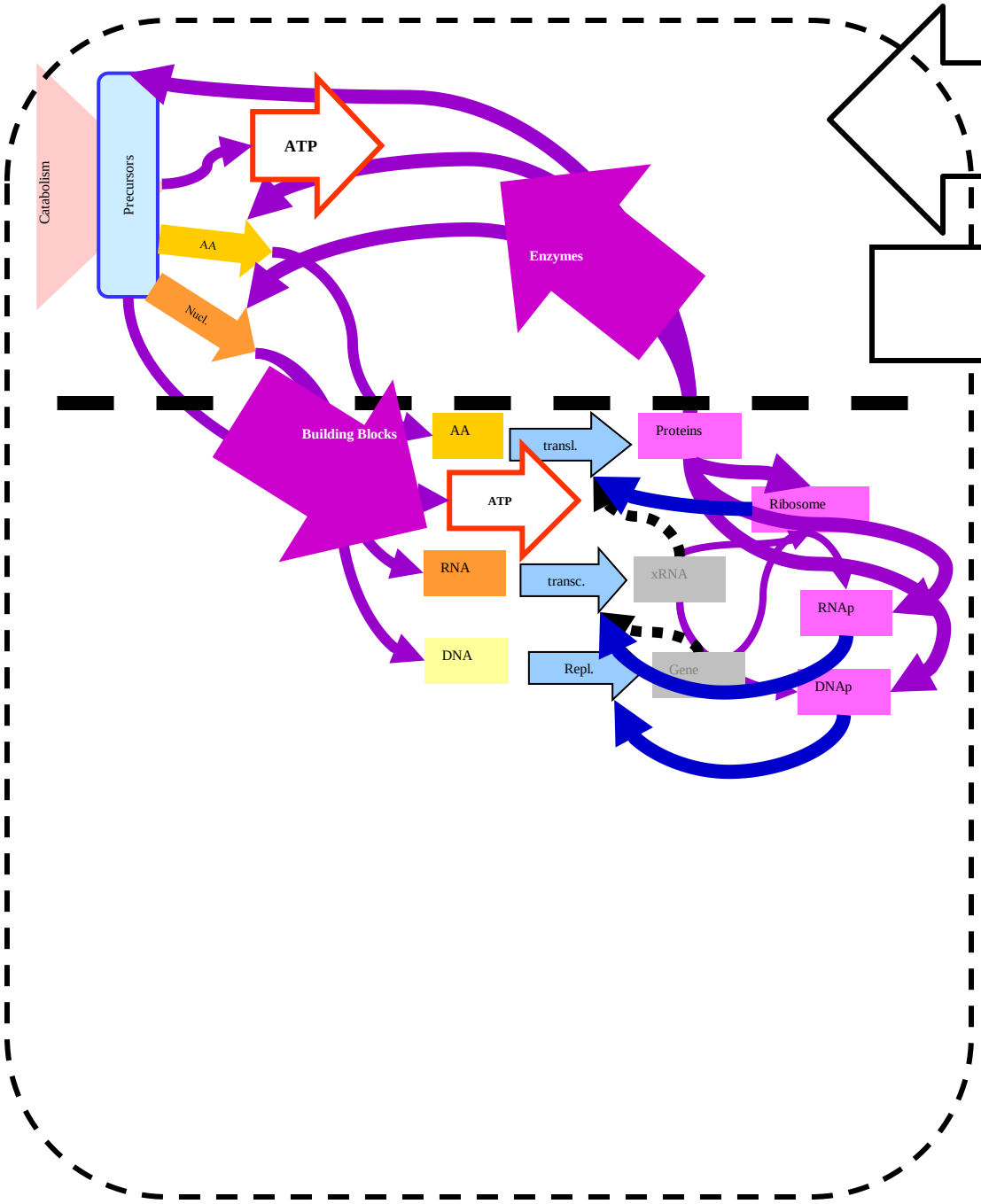

expensive
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
It is control that is fast/slow and expensive/cheap
Ribosome
RNA
transc.
xRNA
RNAp
Control
DNA
Repl.
Gene
DNAp
cheap
fast
slow


Layered view
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Macro-layers
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
Crosslayer autocatalysis



<10% of most bacterial genomes

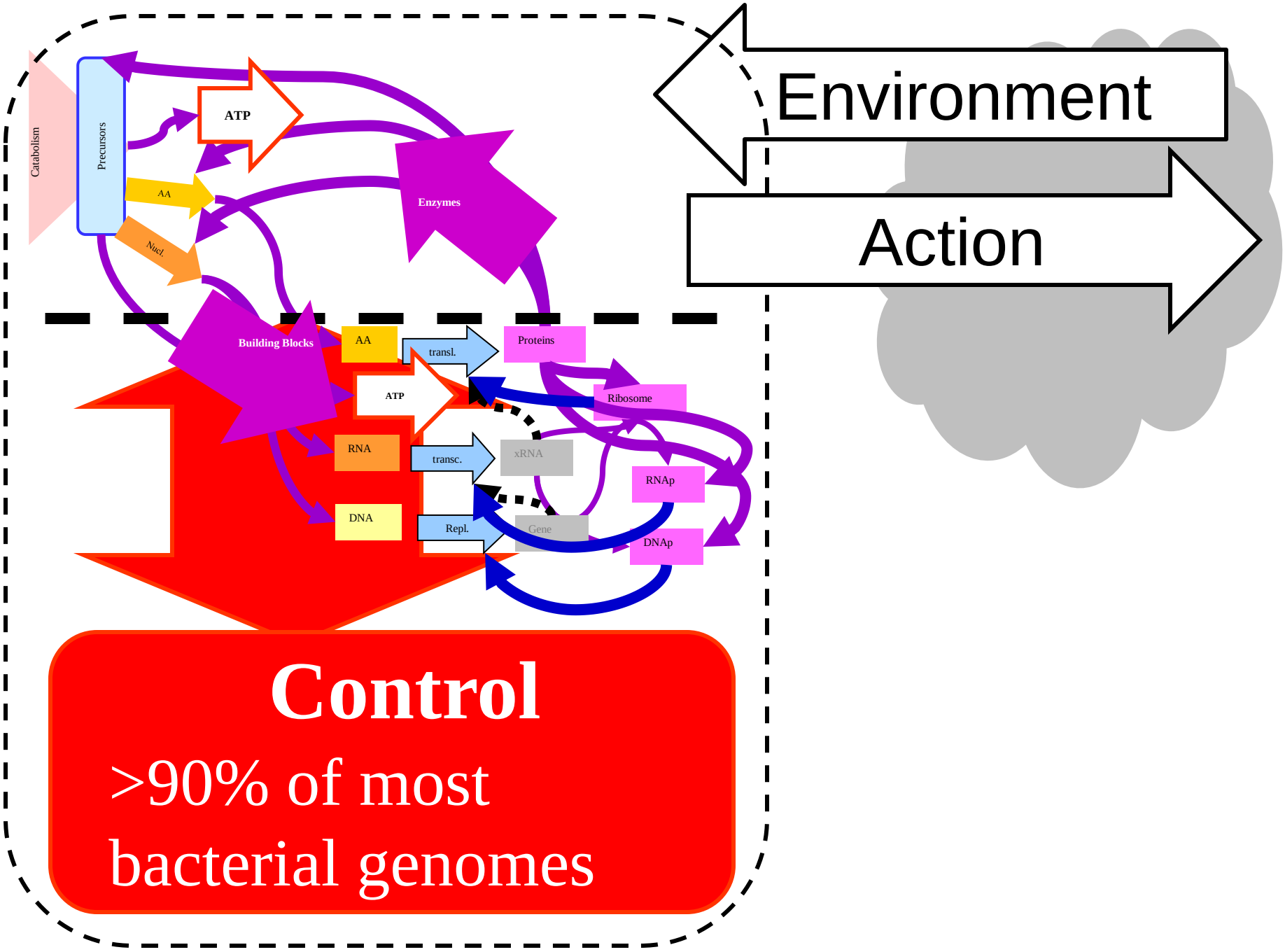
~300 genes, ~minimal genome, requires idealized environment

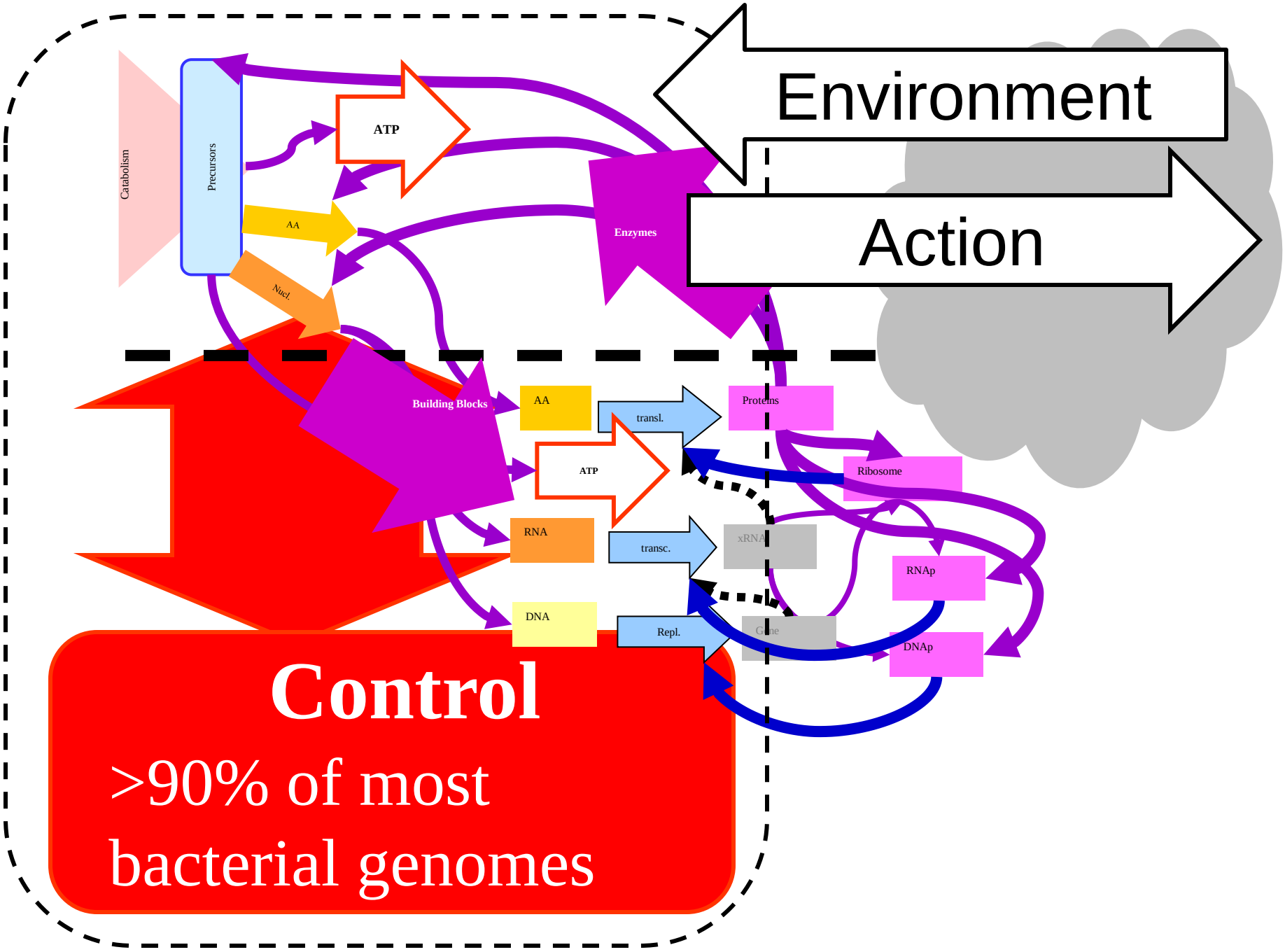


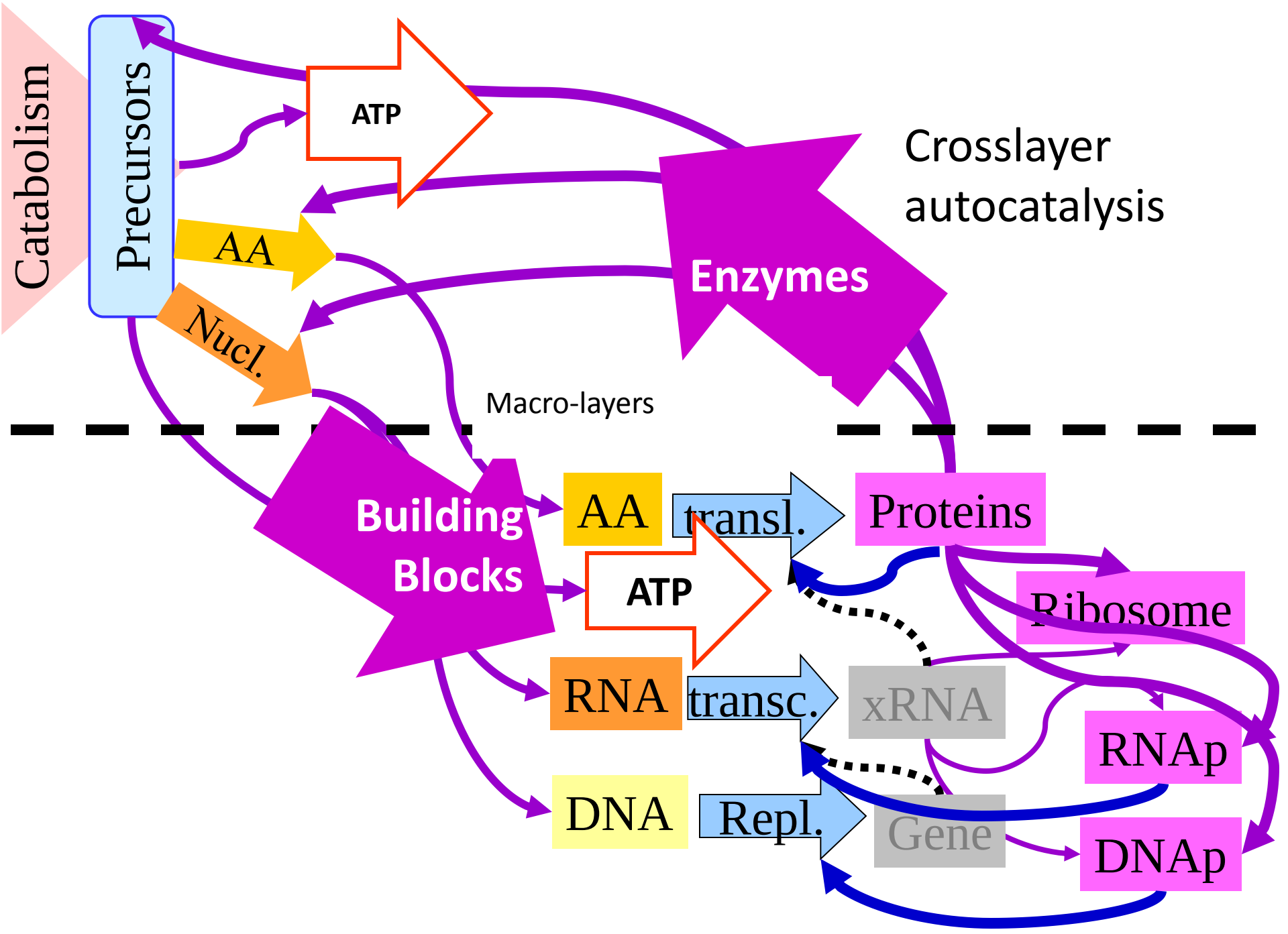
Environment

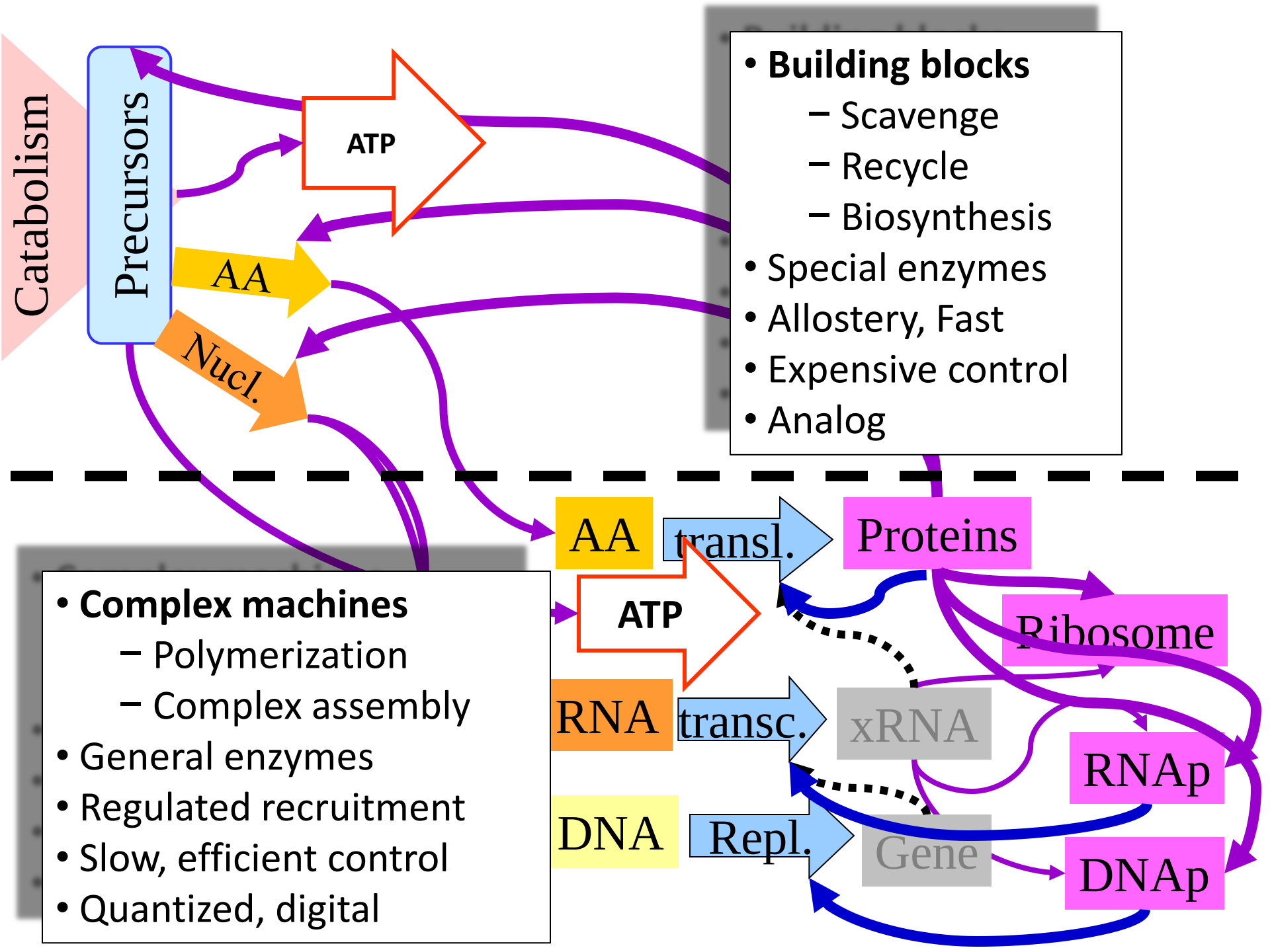
Action

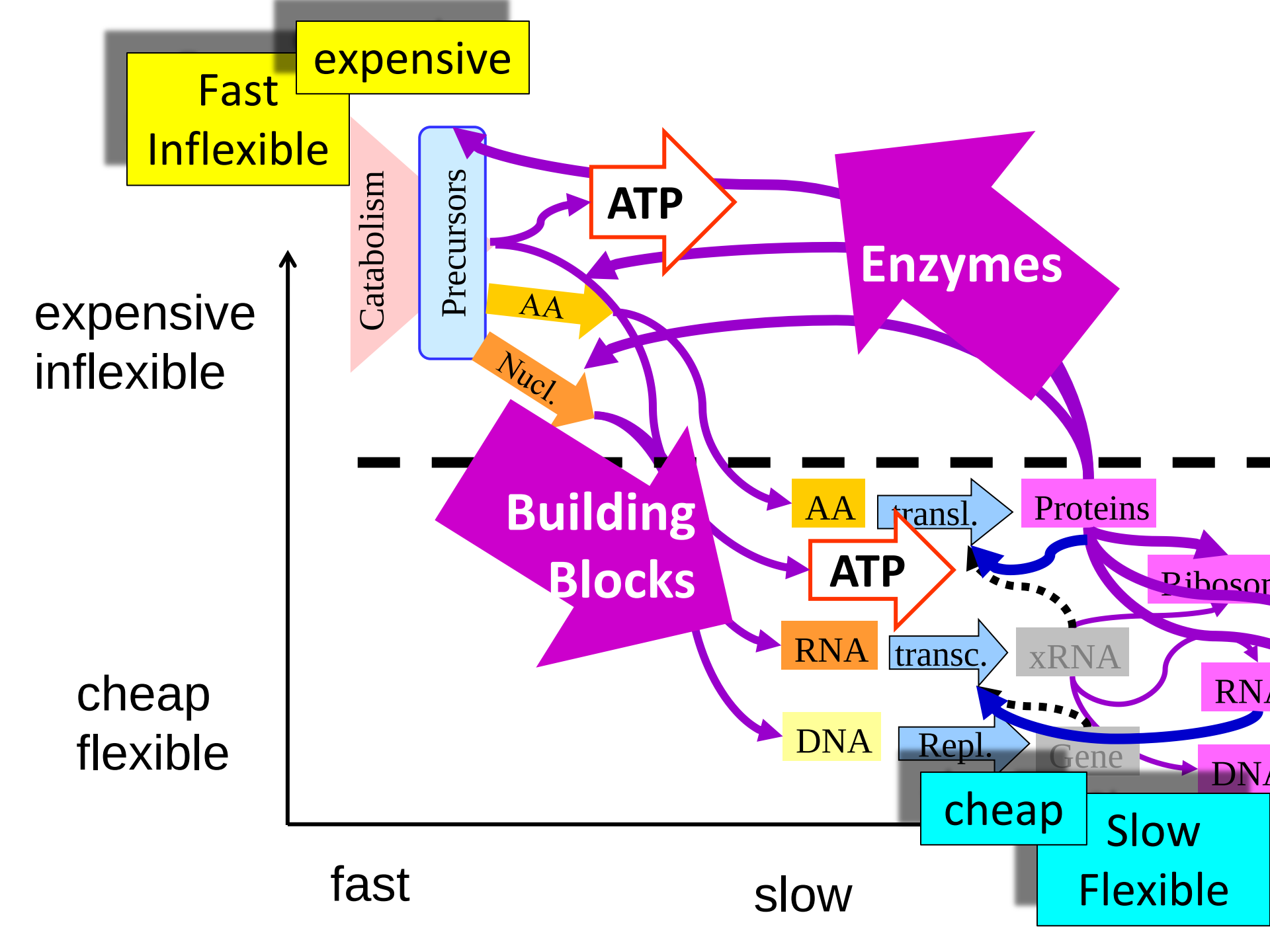
What about "natural" environments?

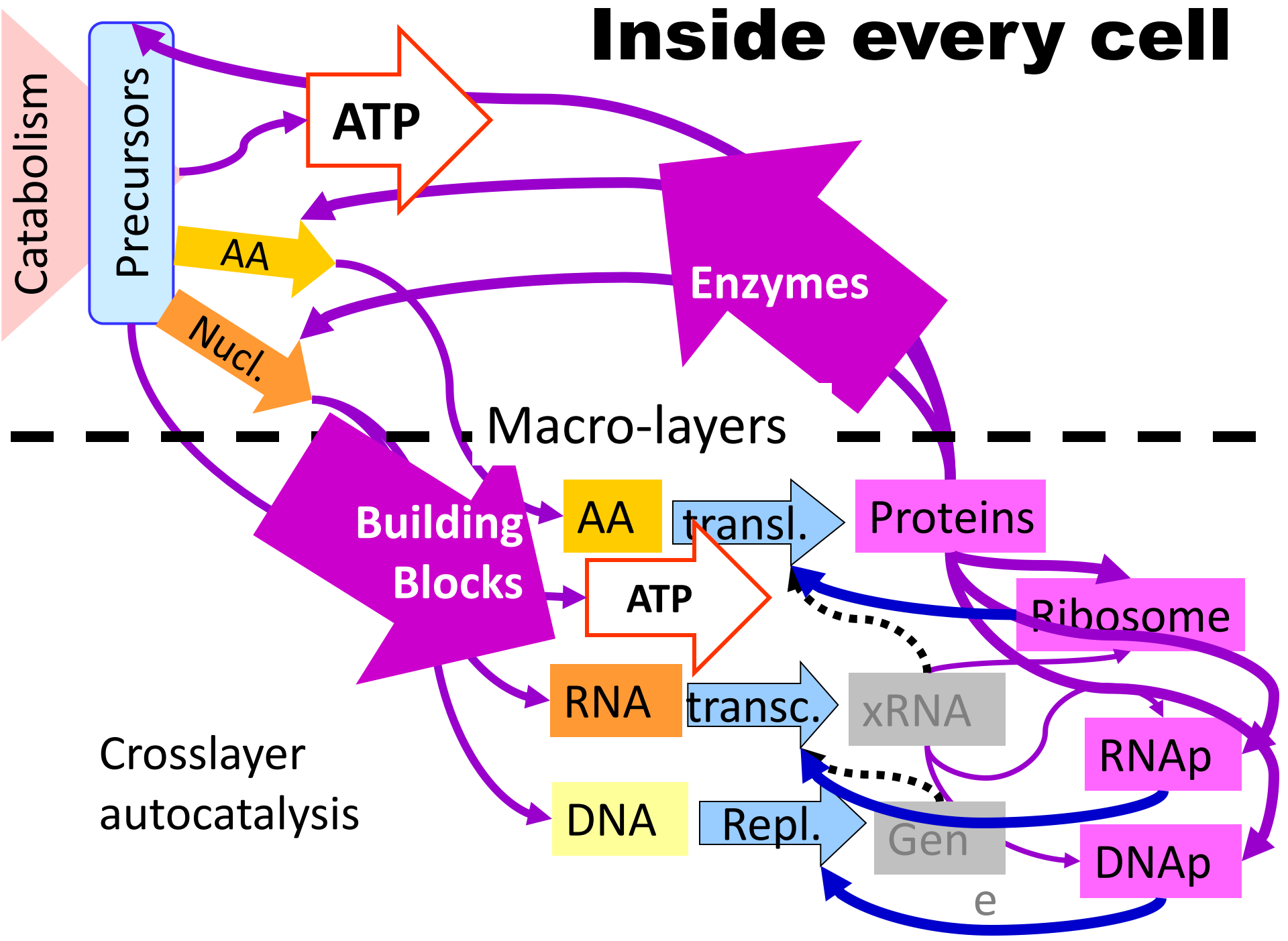

Environment
Action
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
Control
>90% of most
bacterial genomes


Environment
Action
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
Control
>90% of most bacterial genomes


Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Crosslayer
autocatalysis
Macro-layers
Building
Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp


Catabolism
Precursors
ATP
AA
Nucl.
• Building blocks
– Scavenge
– Recycle
– Biosynthesis
• Special enzymes
• Allostery, Fast
• Expensive control
• Analog
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
• Complex machines
– Polymerization
– Complex assembly
• General enzymes
• Regulated recruitment
• Slow, efficient control
• Quantized, digital


expensive
Fast
Inflexible
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
expensive
inflexible
Building
Blocks
AA
transl.
Proteins
ATP
RNA
transc.
xRNA
DNA
Repl.
Gene
cheap
flexible
cheap
Slow
Flexible
fast
slow

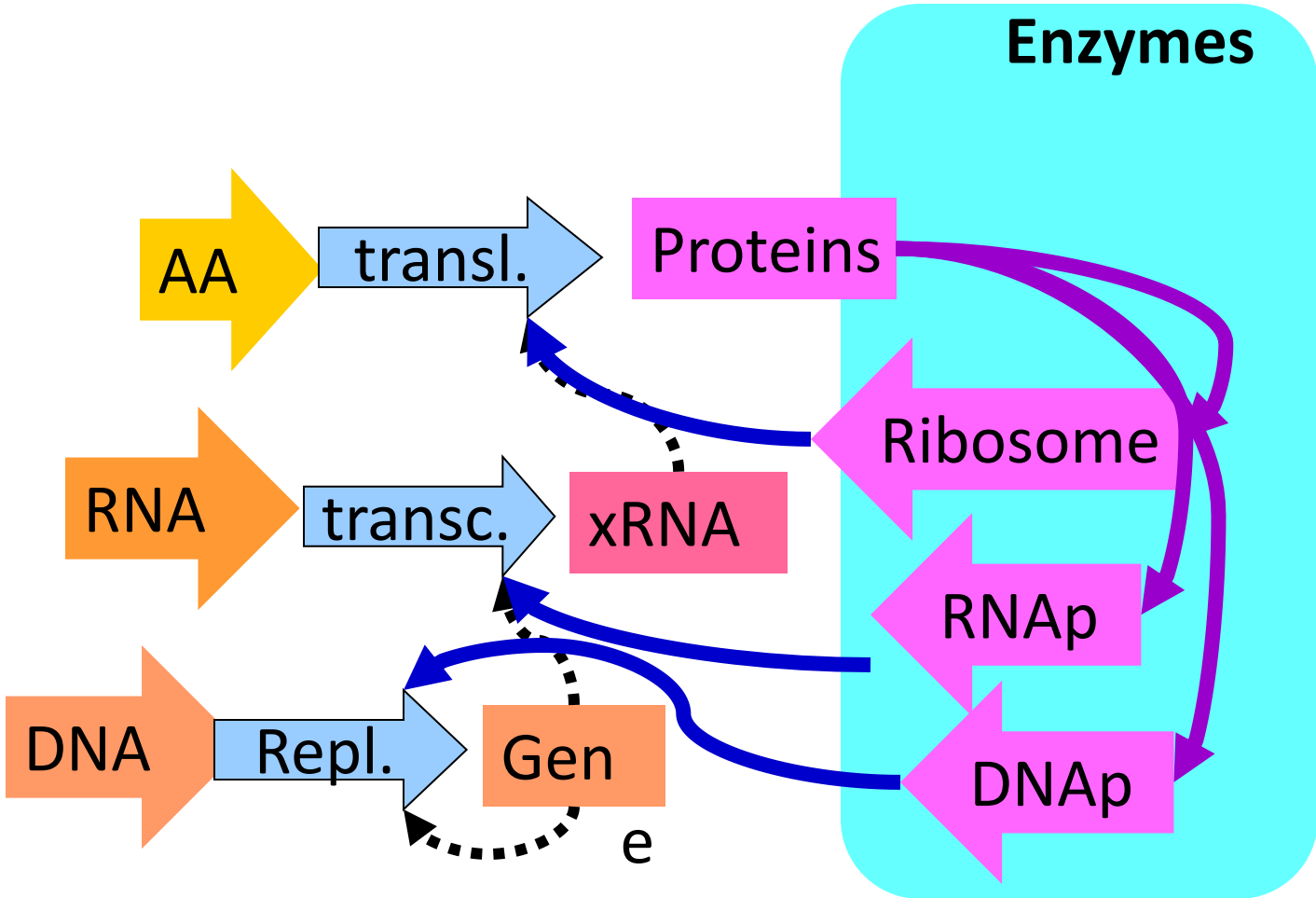

Inside every cell
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Macro-layers
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gen
e
DNAp
Crosslayer autocatalysis

Lower layer autocatalysis

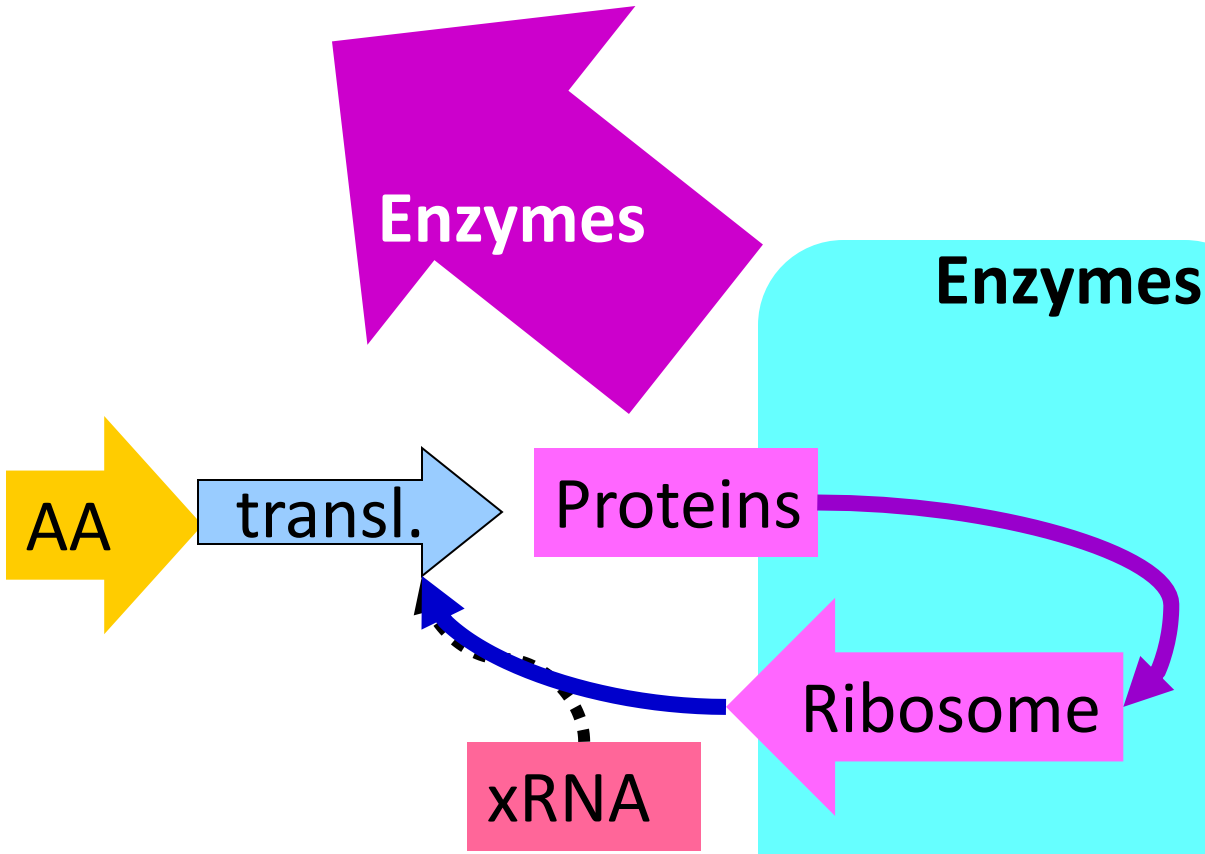
Macromolecules making …

Three lower layers? Yes:

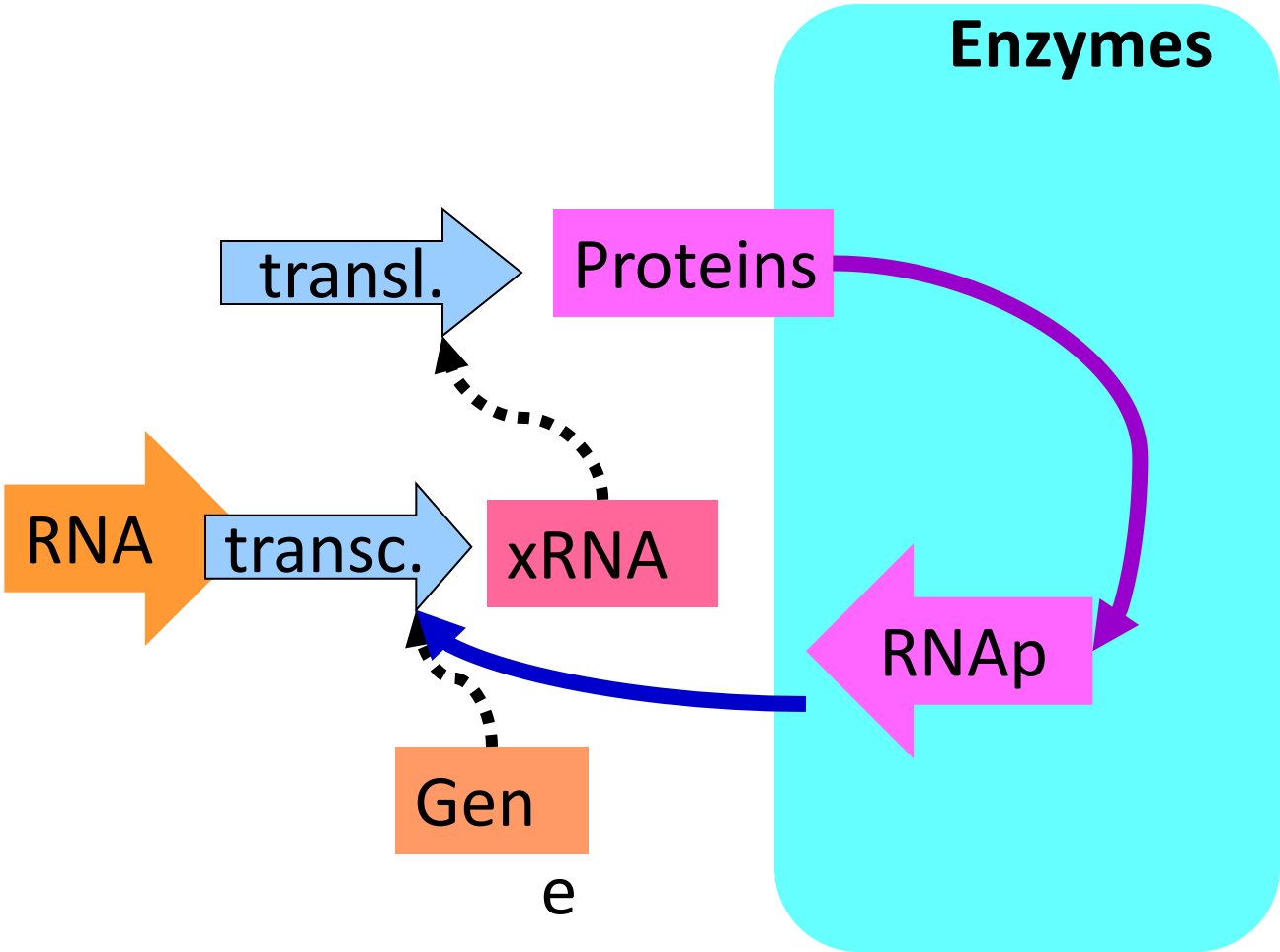
- Translation
- Transcription
- Replication



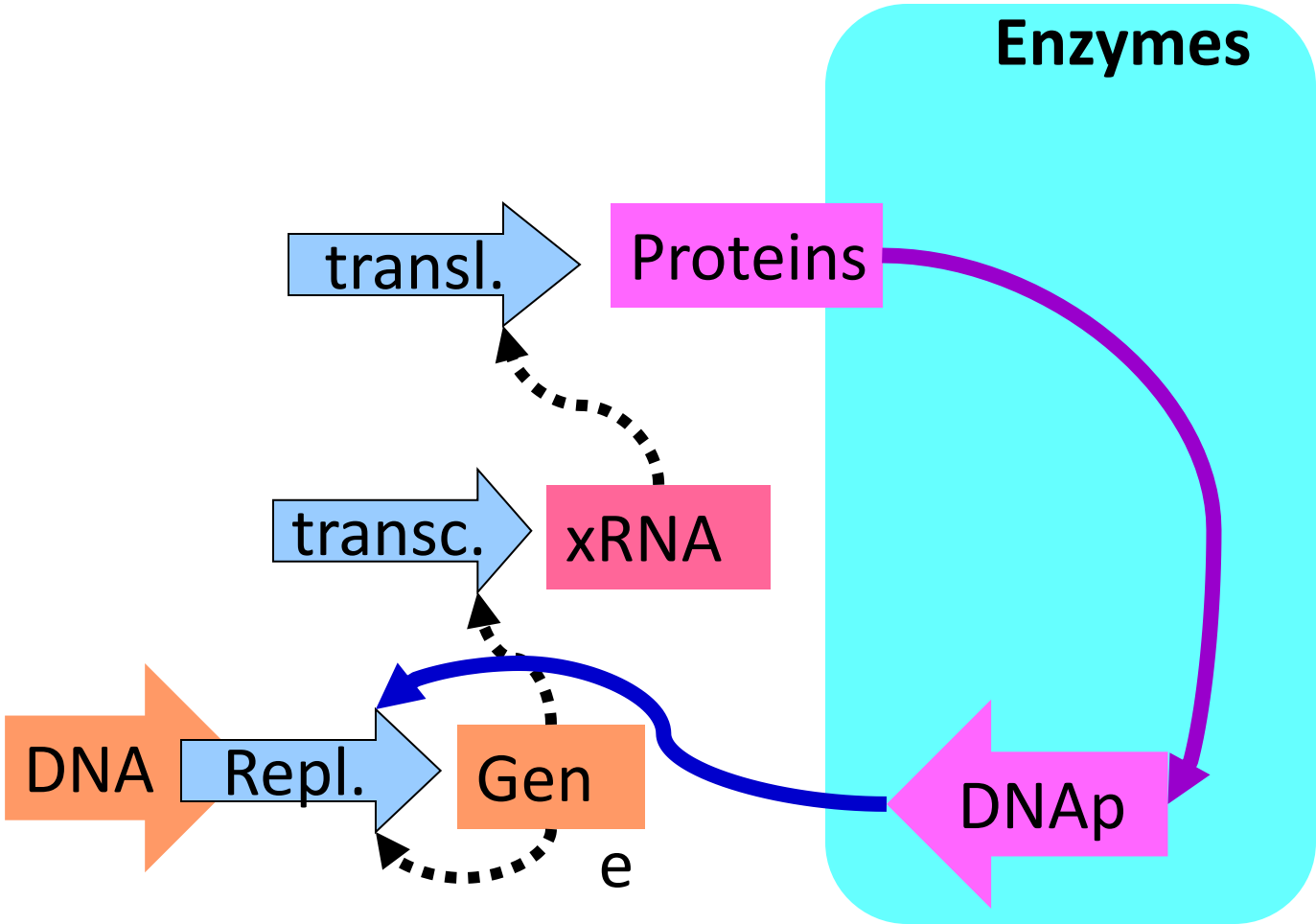
- **Translation**
- Transcription
- Replication

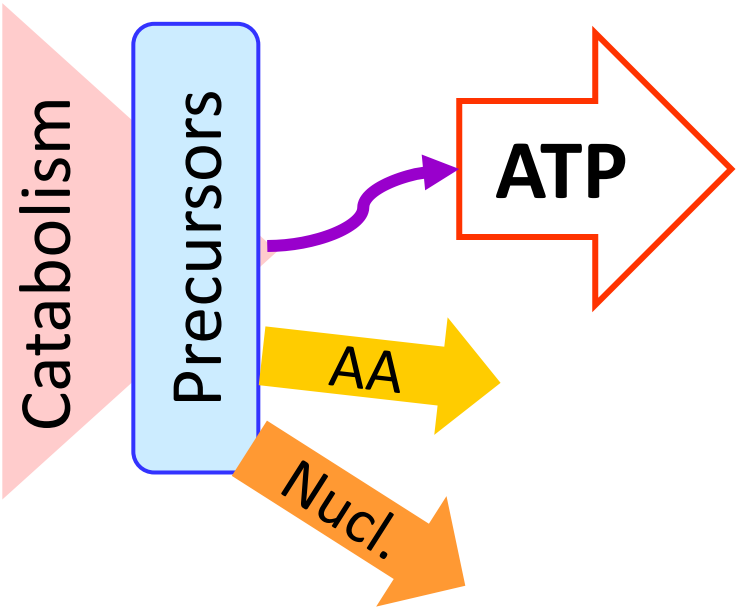


- Translation
- **Transcription**
- Replication

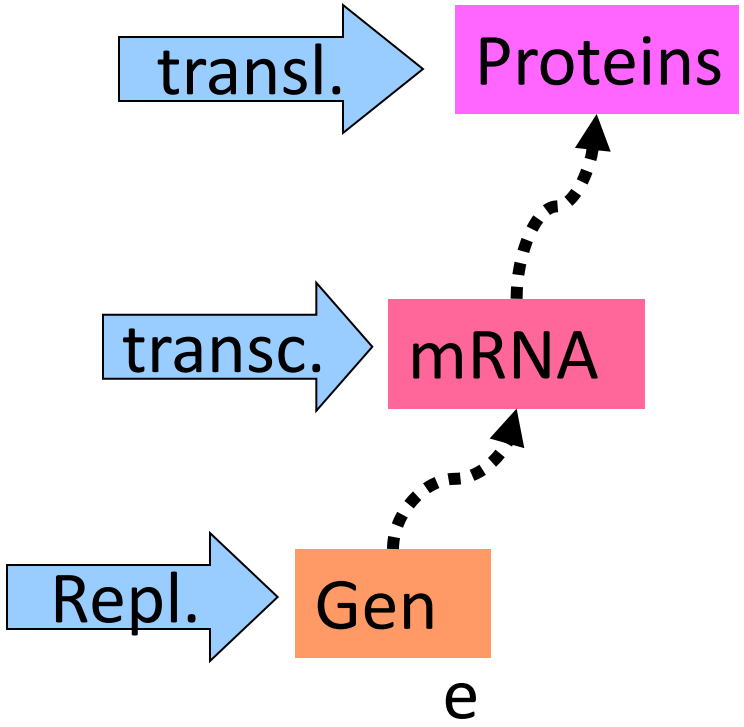


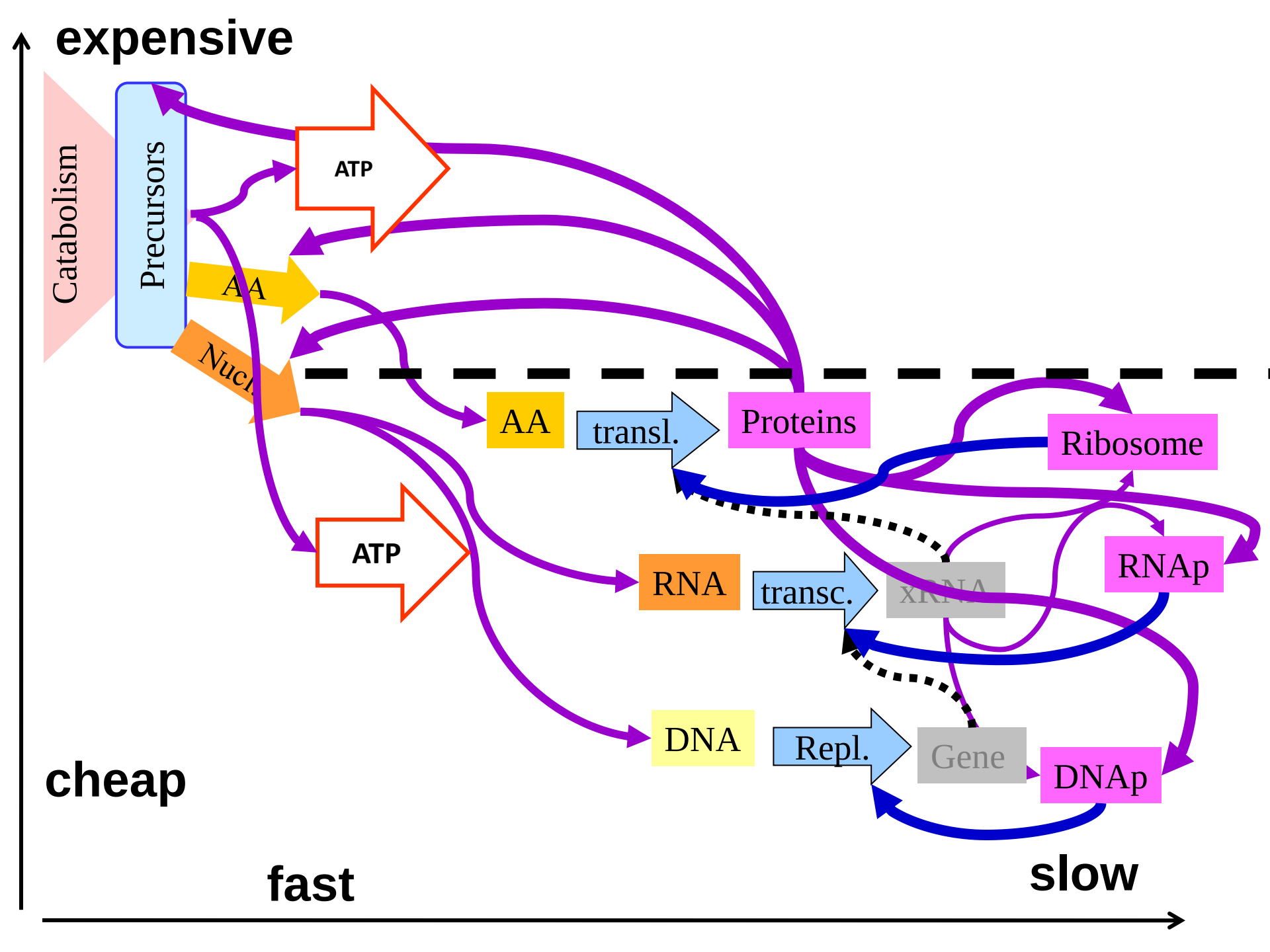
- Translation
- Transcription
- **Replication**

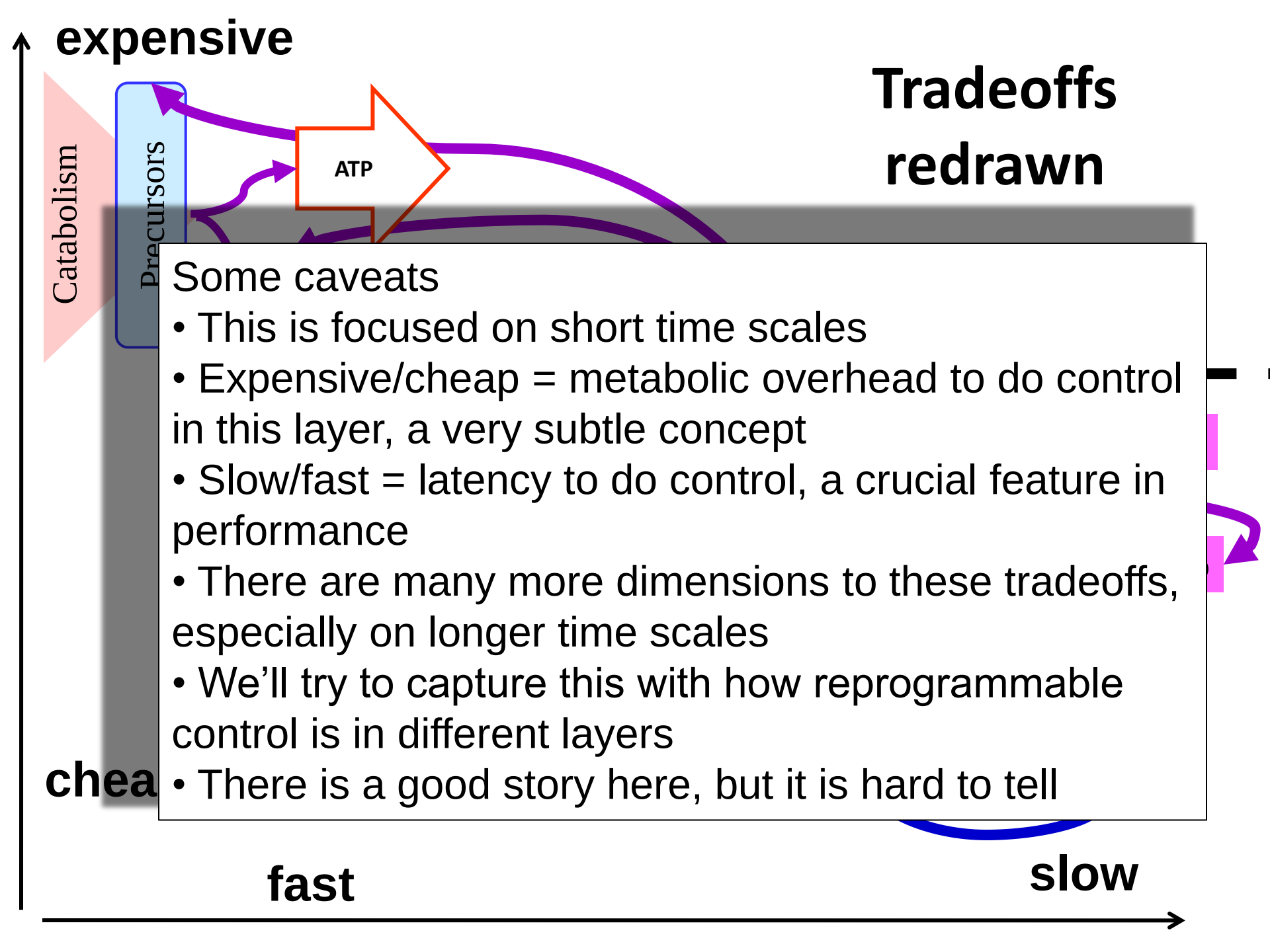


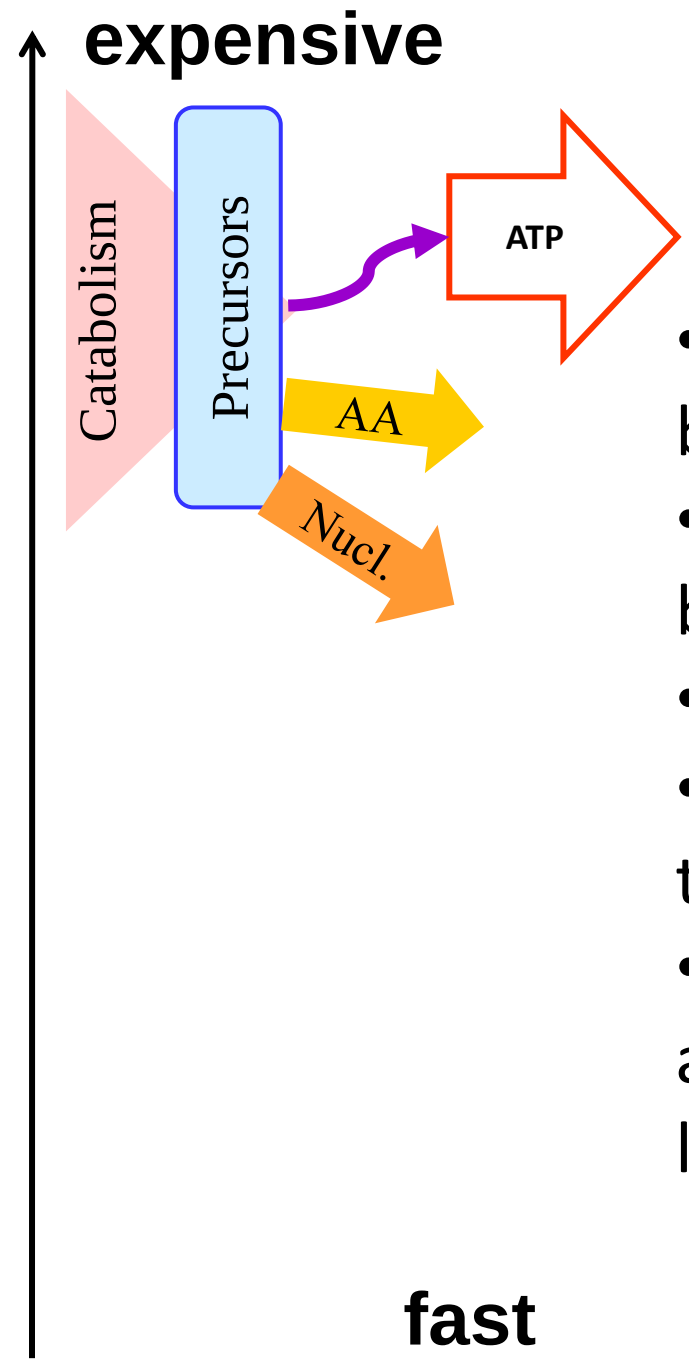


Pathway views




expensive
Catabolism
Precursors
ATP
AA
Nucl.
AA
transl.
Proteins
Ribosome
ATP
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
cheap
fast
slow


expensive
Tradeoffs redrawn
Catabolism
Precursors
ATP
Some caveats
• This is focused on short time scales
• Expensive/cheap = metabolic overhead to do control in this layer, a very subtle concept
• Slow/fast = latency to do control, a crucial feature in performance
• There are many more dimensions to these tradeoffs, especially on longer time scales
• We'll try to capture this with how reprogrammable control is in different layers
• There is a good story here, but it is hard to tell
chea
fast
slow



~~upper protein layer~~

metabolism?

- The layer names are an attempt to bridge to traditional terms
- Which arose in the "pathway" view, before layering
- 3 layers?: protein, RNA, and DNA
- 4 layers?: metabolism, translation, transcription, replication
- Named for the macromolecules that are catalysts or "instructions" for their layers, or the process

expensive

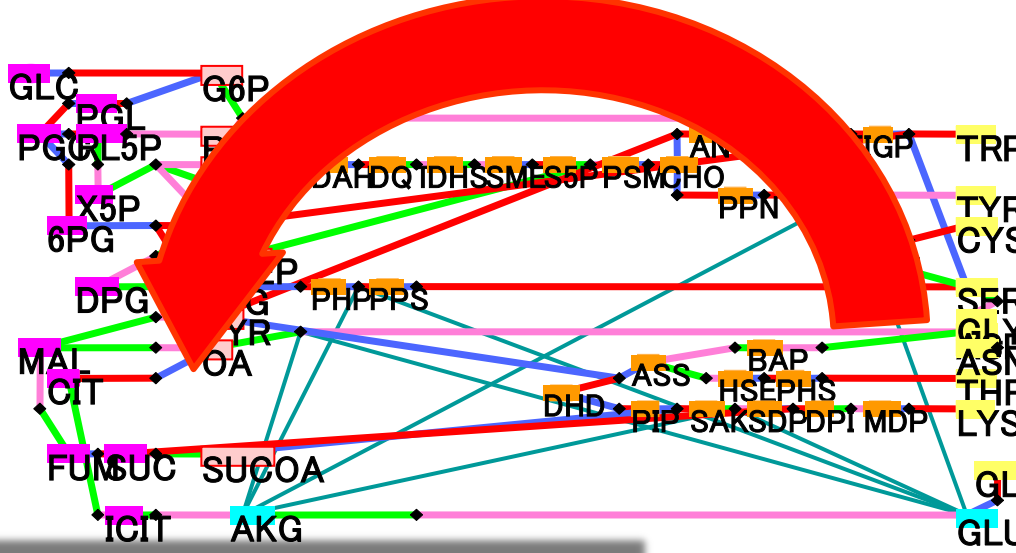
Catabolism

Precursors

ATP

AA

Nucl.



- Fastest allosteric control
- Complex special proteins
- High metabolic overhead
- **Hard to reprogram**

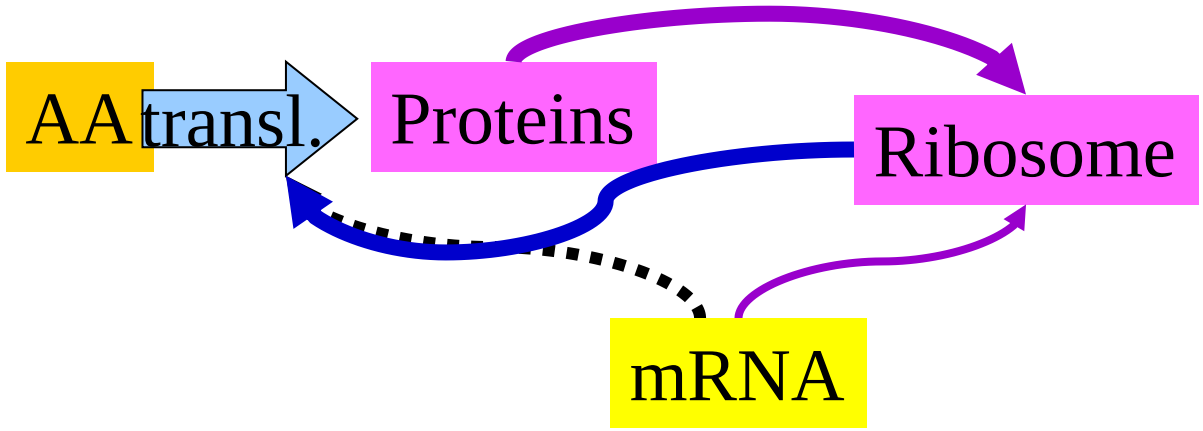
- Layer of “action”
- Sensing and actuation in this layer

fast

expensive

middle RNA layer

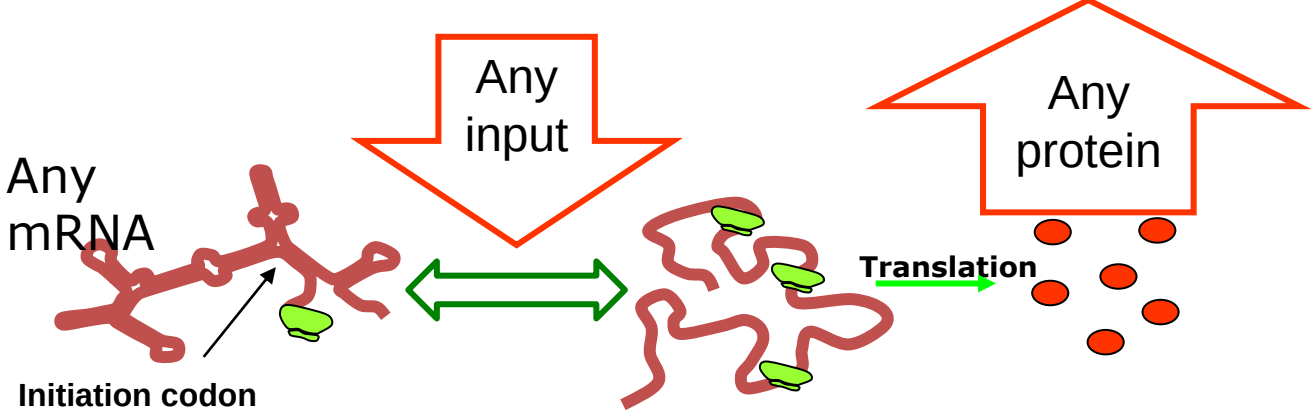
translation?

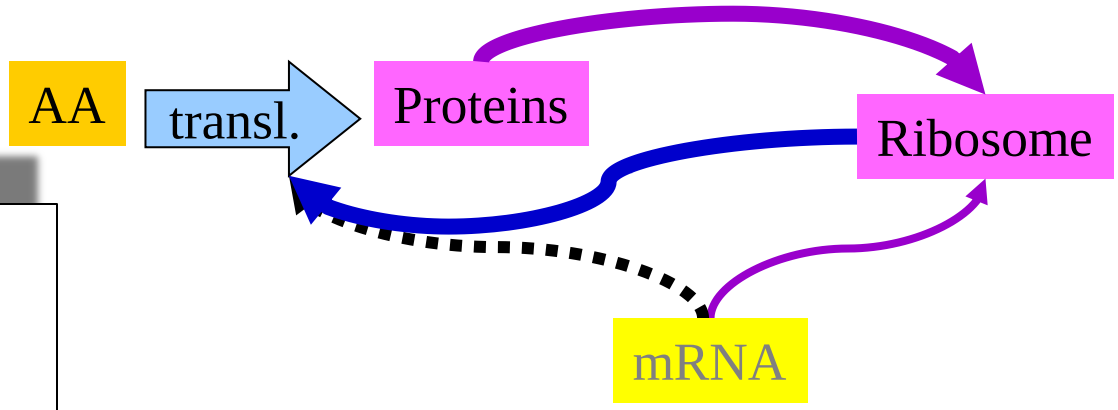


cheap

fast

slow



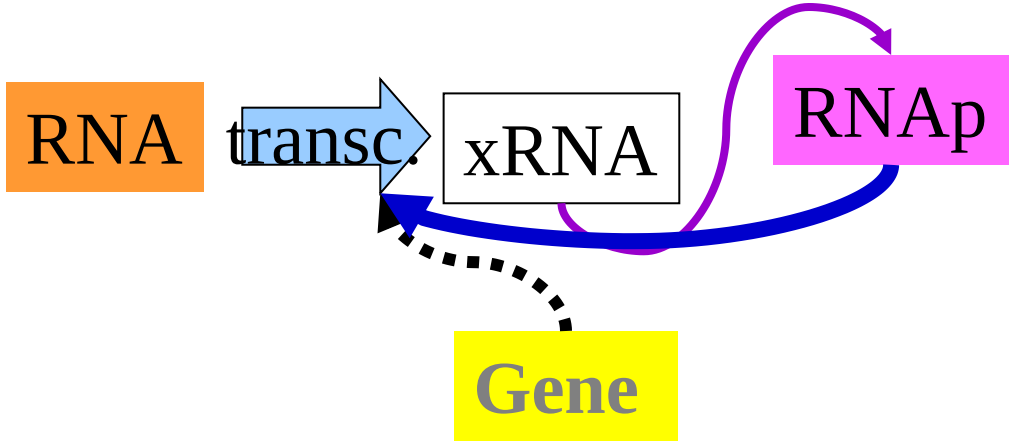


- Fast translation control
- Complex RNAs
- General polymerases
- Medium metabolic overhead?
- Highly reprogrammable?

- Lots of control happens here
- This is the “heart” and “brain” of the cell
- **Complexity and importance is underrated**

expensive

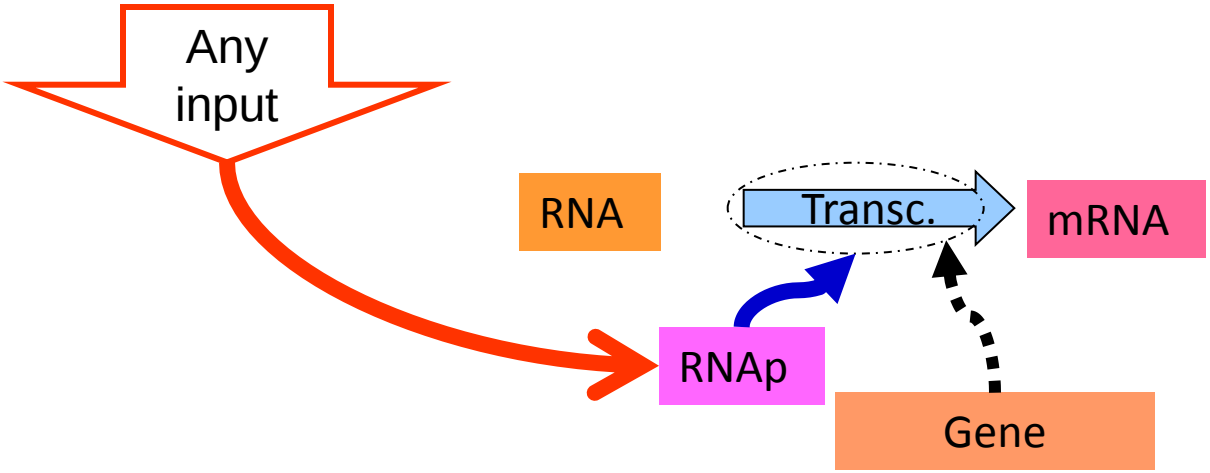
Transcription layer

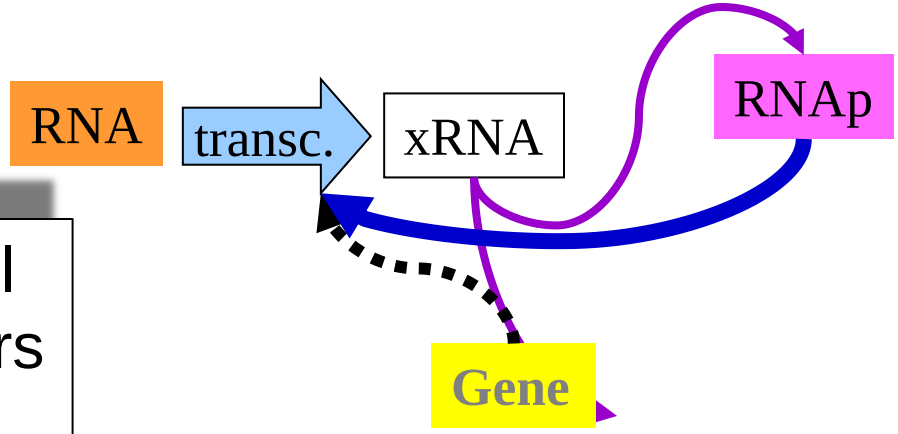


cheap

fast

slow



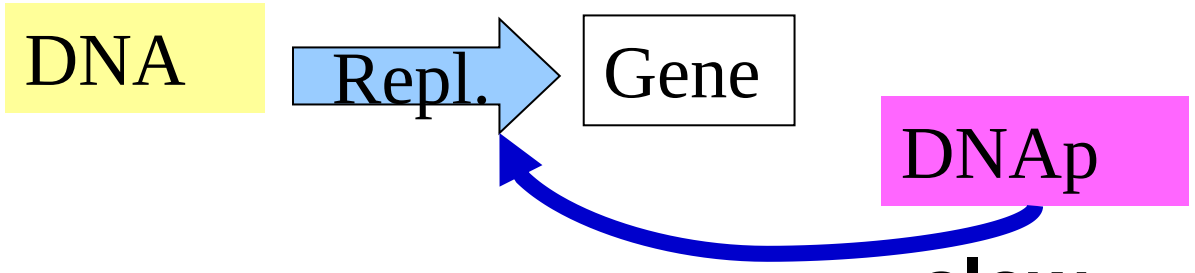


- Slowest transcription control
- Complex transcription factors
- General polymerases
- Lowest metabolic overhead
- **Easily reprogrammed**

expensive

Replication layer

- Amount of control here ***extremely underrated***
- Getting better
- Bacterial genome is highly dynamic
- Source of astonishing evolvability
- Note: horizontal gene transfer works because of whole “protocol stack” not just shared codons



cheap

fast

slow

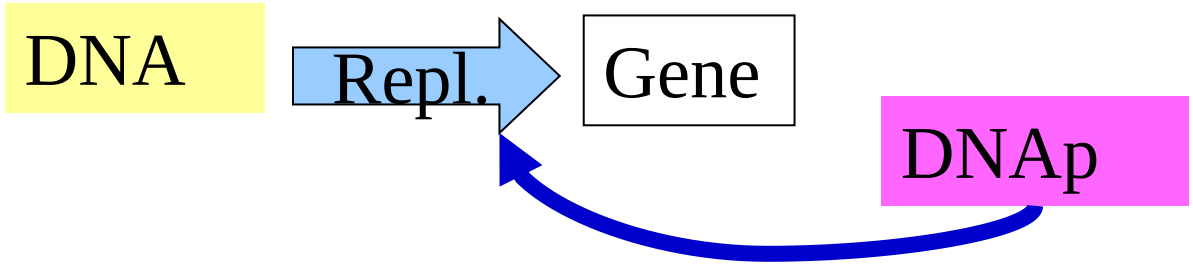
Architecture = protocols = "constraints that deconstrain"

Bacterial biosphere

- carriers: ATP, NADH, etc
- Precursors, …
- Enzymes
- Translation
- Transcription
- Replication

Protocols

- **Horizontal gene transfer works because of whole "protocol stack" not just shared codons**



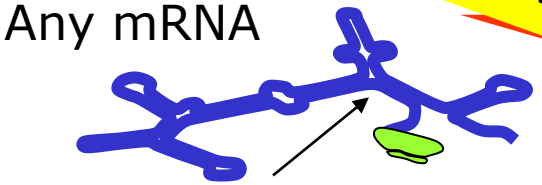
Summary so far

- Fastest allosteric feedback control
- Complex specialized proteins
- High metabolic overhead
- Hard to reprogram

- Fast translation control
- Complex RNAs
- Med. metabolic overhead
- Highly reprogrammable?

- General polymerases

- Slowest transcription control
- Complex transcription factors
- Lowest metabolic overhead
- Easily reprogrammed



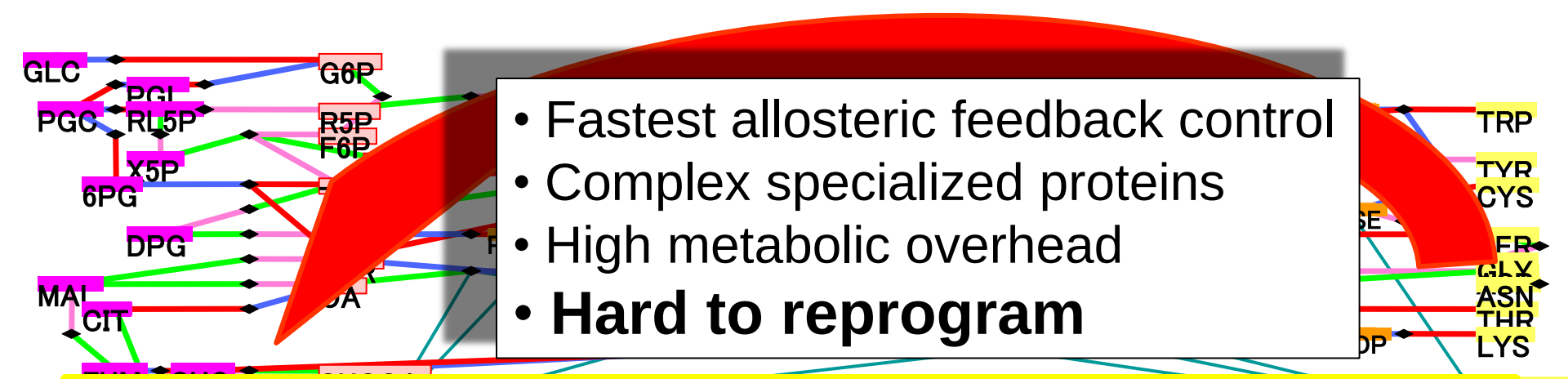
Any input

Any

Enzymes

mRNA

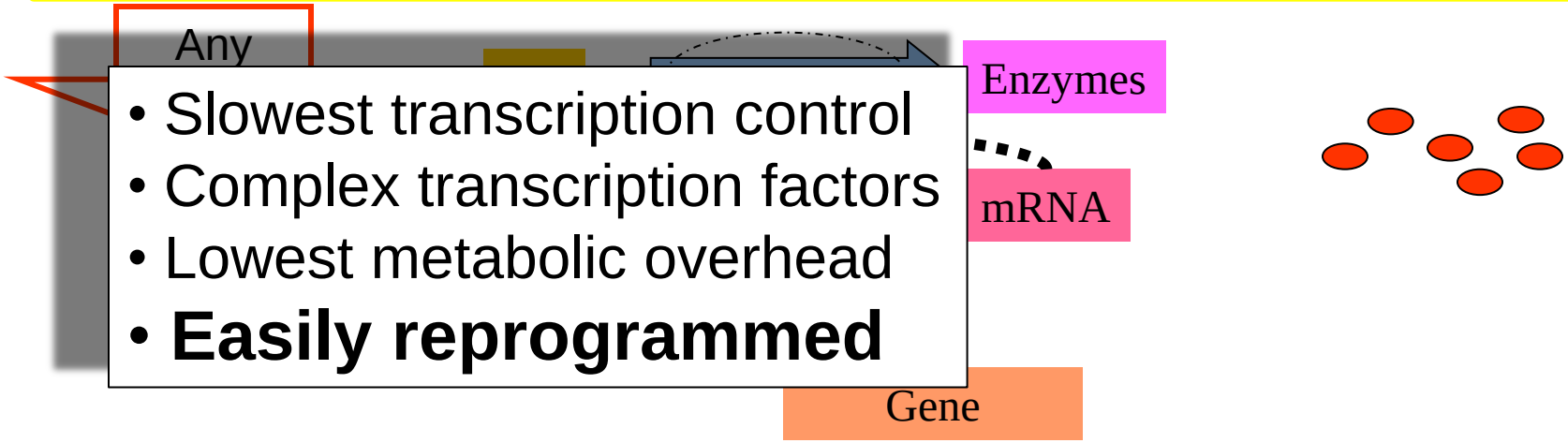
Gene



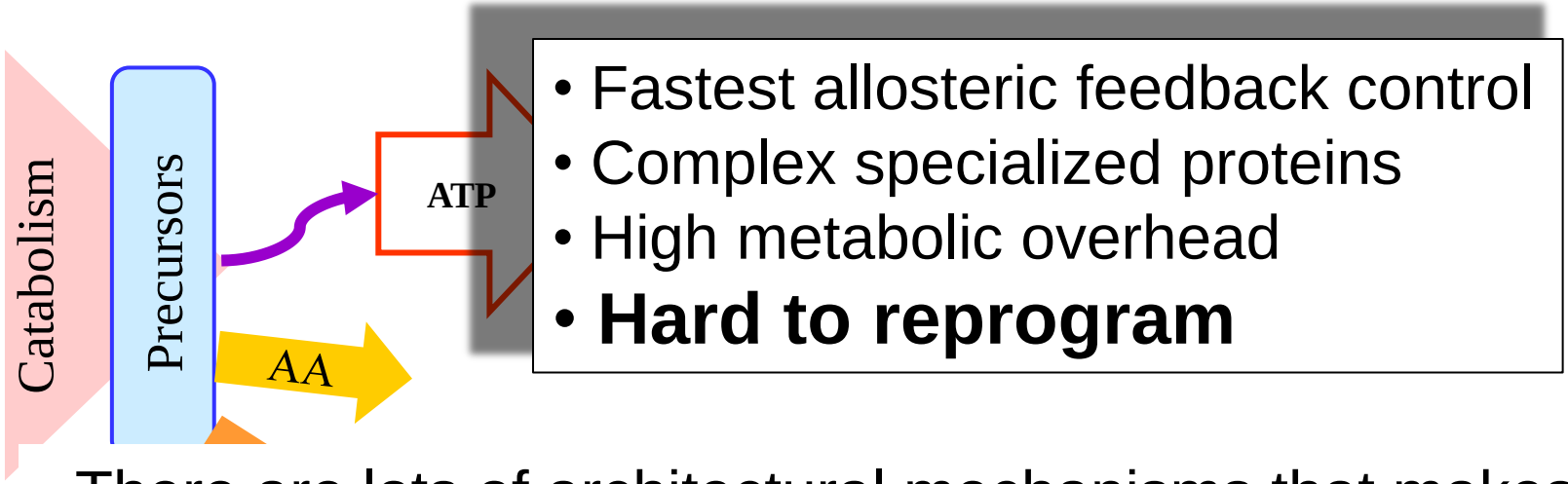
- Fastest allosteric feedback control
- Complex specialized proteins
- High metabolic overhead
- **Hard to reprogram**

This is hard to explain. Reprogramming the protein layer involves changing the genome, so they are in some sense "the same," but…

What I mean specifically, is that it is easier to change ***control*** of transcription than to change ***control*** in protein interaction circuits. This needs lots of details to make clear.



- Slowest transcription control
- Complex transcription factors
- Lowest metabolic overhead
- **Easily reprogrammed**



• There are lots of architectural mechanisms that makes this surprisingly reprogrammable, e.g. see the discussion on two-component signal transduction…. Nevertheless…

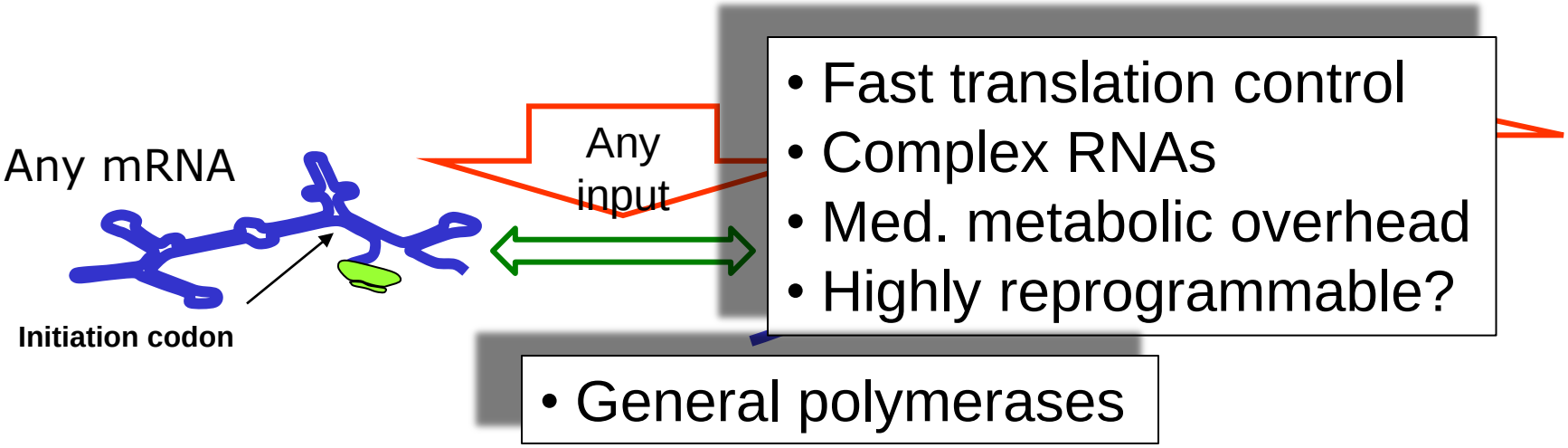
•… changes here require changes in protein function (in addition to sequence), which is complicated difficult.

• Changing the allosteric properties of proteins is really hard

• E.g. synthetic biology barely touches this because relation between sequence and function is complex

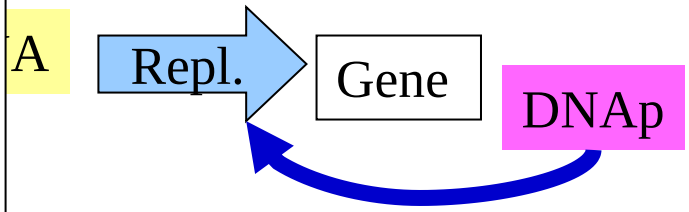
• Here the distinction (a la Ptashne) of allostery versus regulated recruitment is also essential (again illustrated by 2comp signal transduction, but also transcription control)

- **Control in RNA is underrated**, but getting more attention
- RNA polymers are versatile
- Can interact with all layers
- Control is fast and cheap
- Even greater use in higher eukaryotes



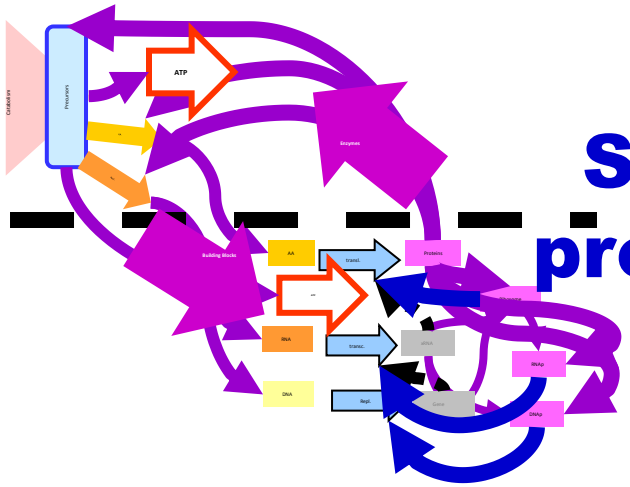
• As reprogrammable as everything else is, this part is the most reprogrammable.
• All transcription control is regulated recruitment, and promoter regions are easily mutated to new function since the relation between sequence and function is direct
• Horizontal gene transfer means this can also be changed by large amounts that are nevertheless functional
• **The extent to which microbial genomes are actively controlled is underrated but evidence is growing.**

- Slowest transcription control
- Complex transcription factors
- Lowest metabolic overhead
- **Easily reprogrammed**



Bacterial biosphere

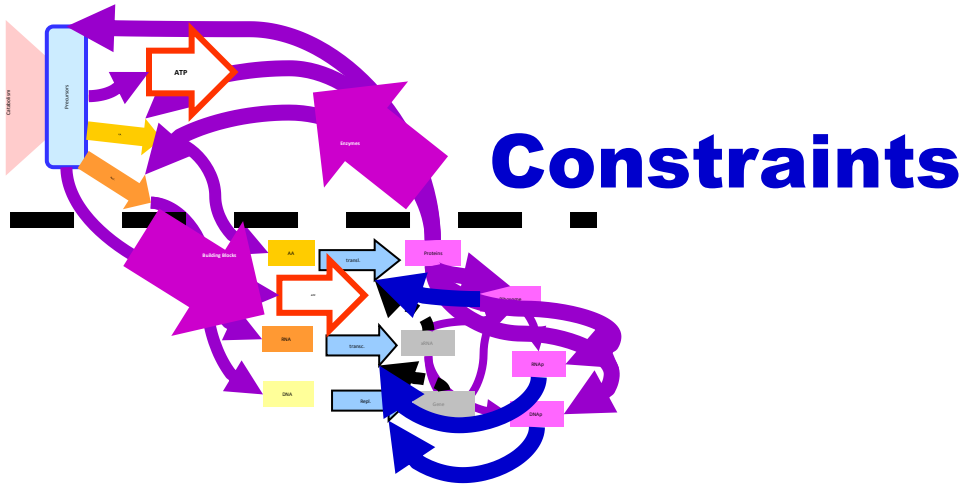
Diverse Environments



Shared protocols

Diverse Genomes

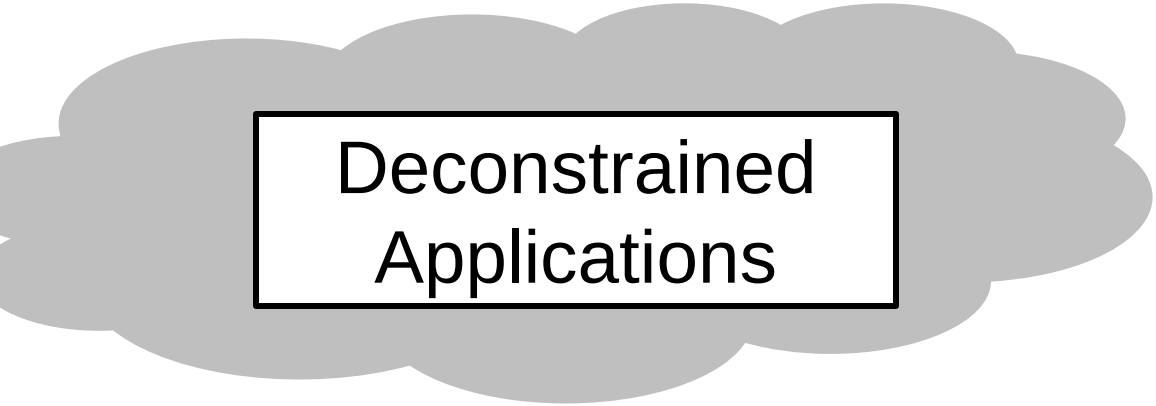
Deconstrained



Architecture
=
Constraints
that
Deconstrain

Deconstrained

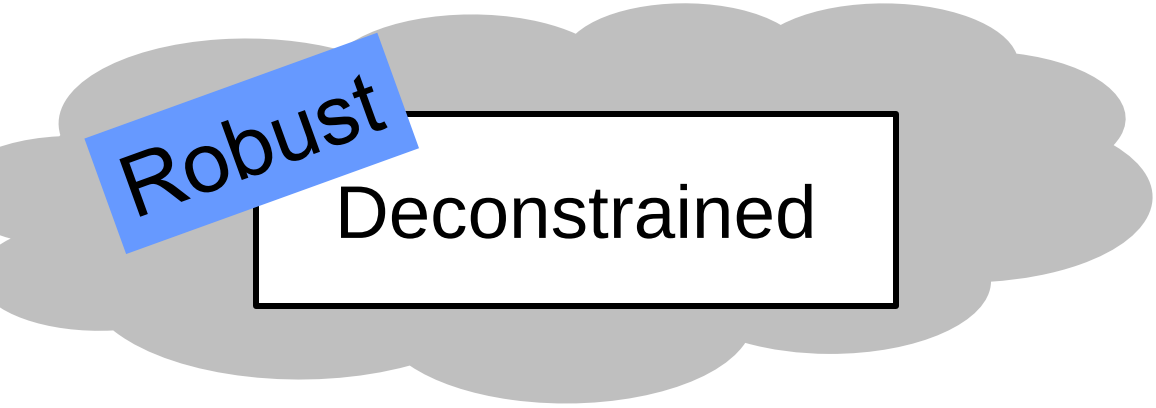
The Technium



Constraints
= Protocols

Architecture
=
Constraints
that
Deconstrain

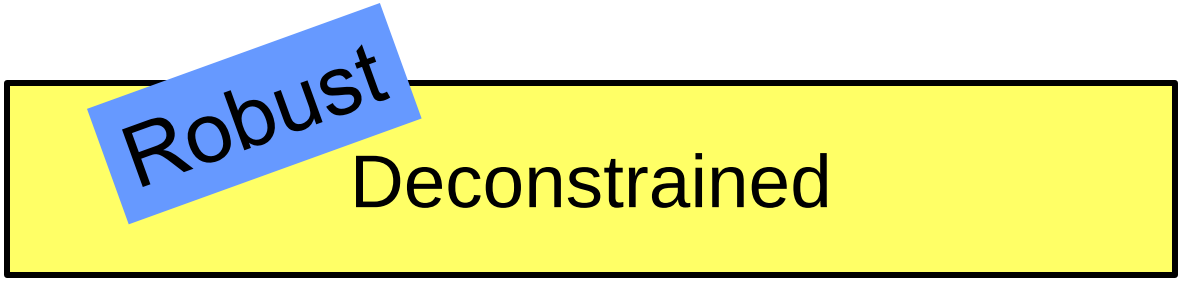
Deconstrained Hardware



Fragile?

Constraints
= Protocols

Hijacking
Parasites
Predators

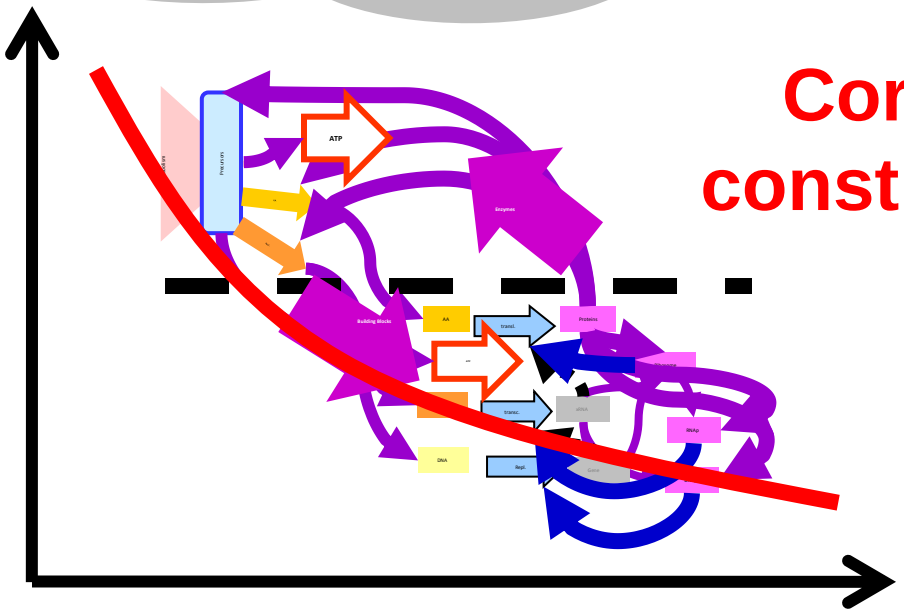


What makes the bacterial biosphere so adaptable?

Deconstrained

Environment

Action

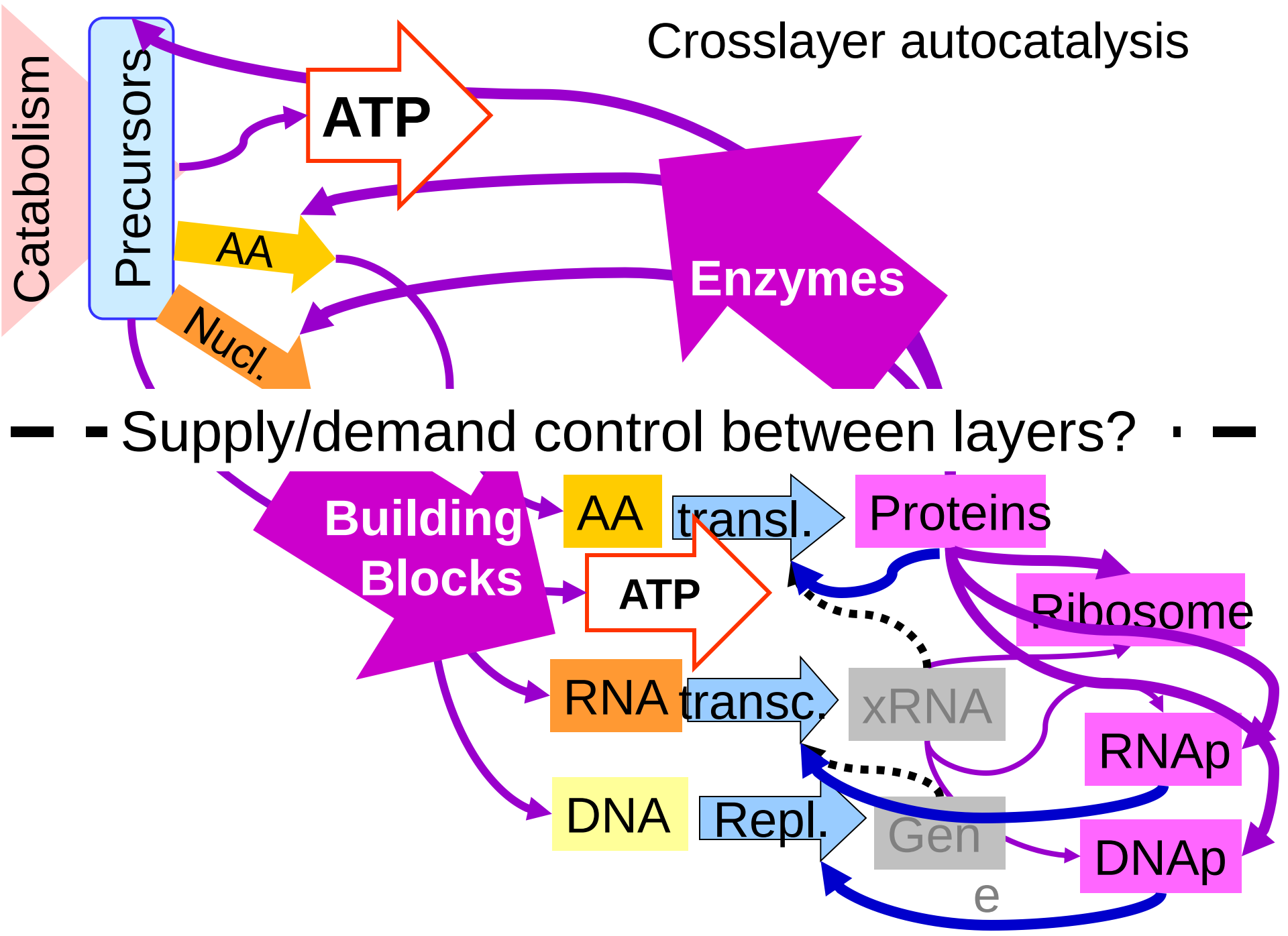


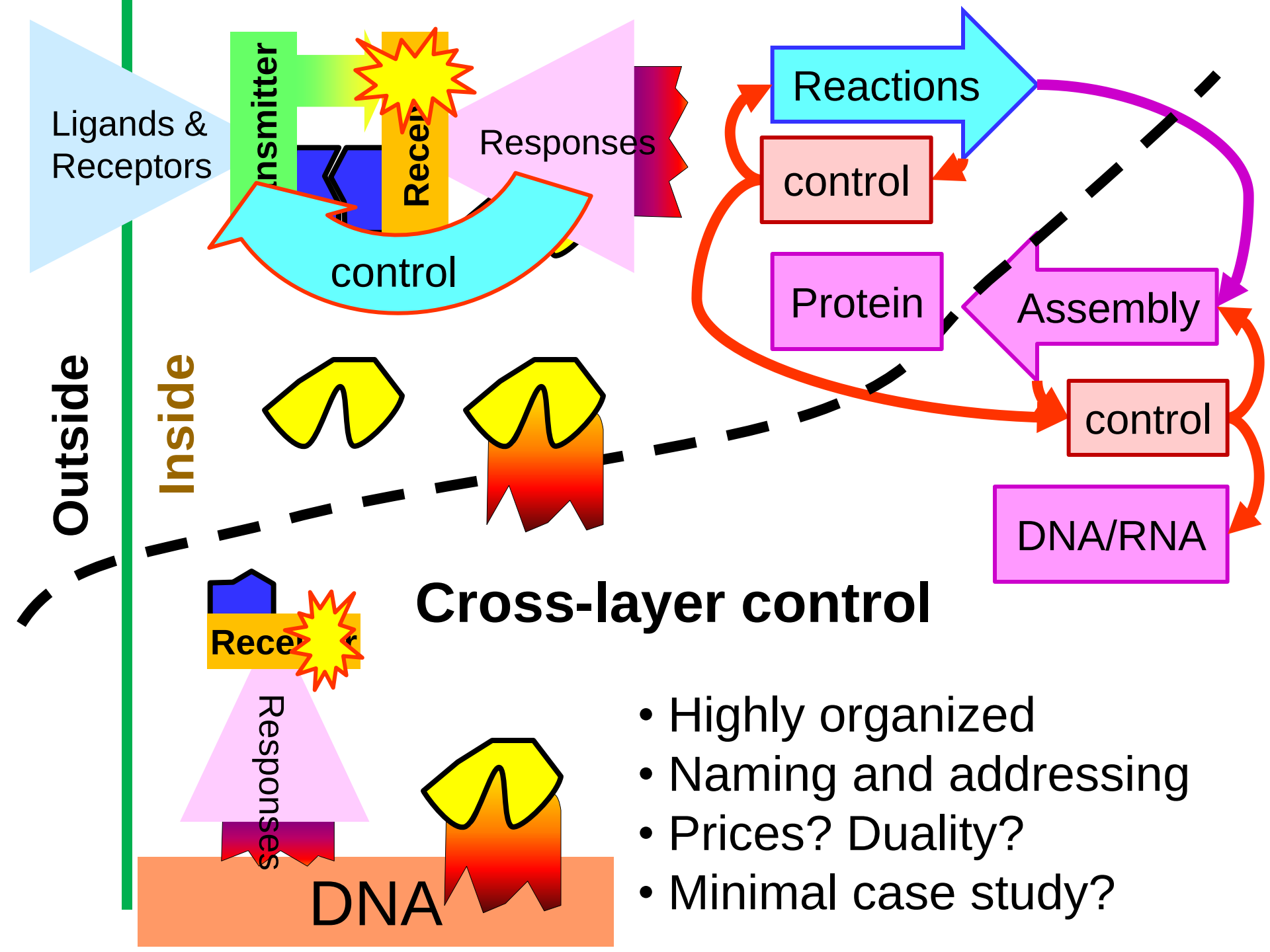
Core conserved constraints facilitate tradeoffs

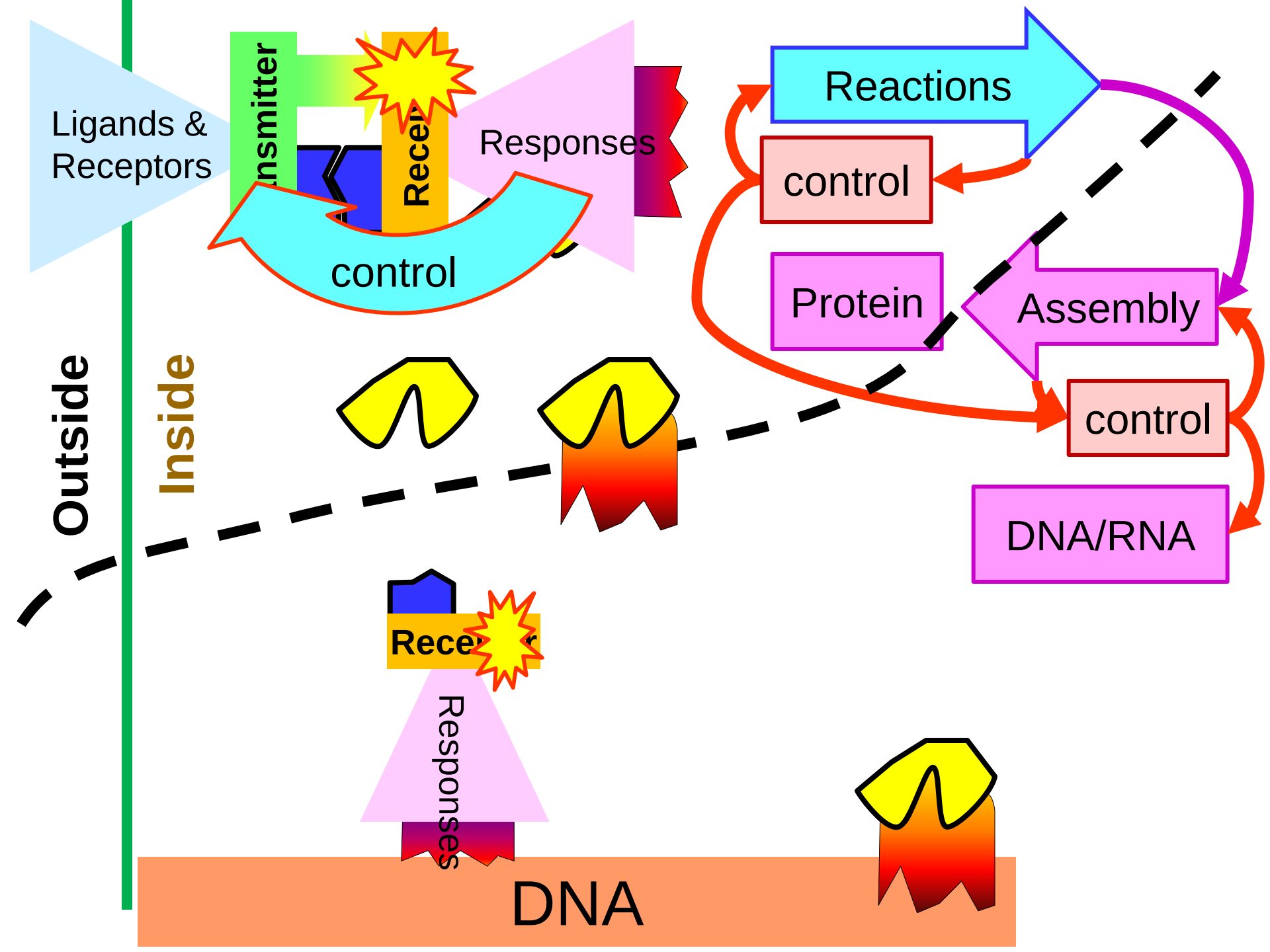
+

Active control of the genome (facilitated variation)

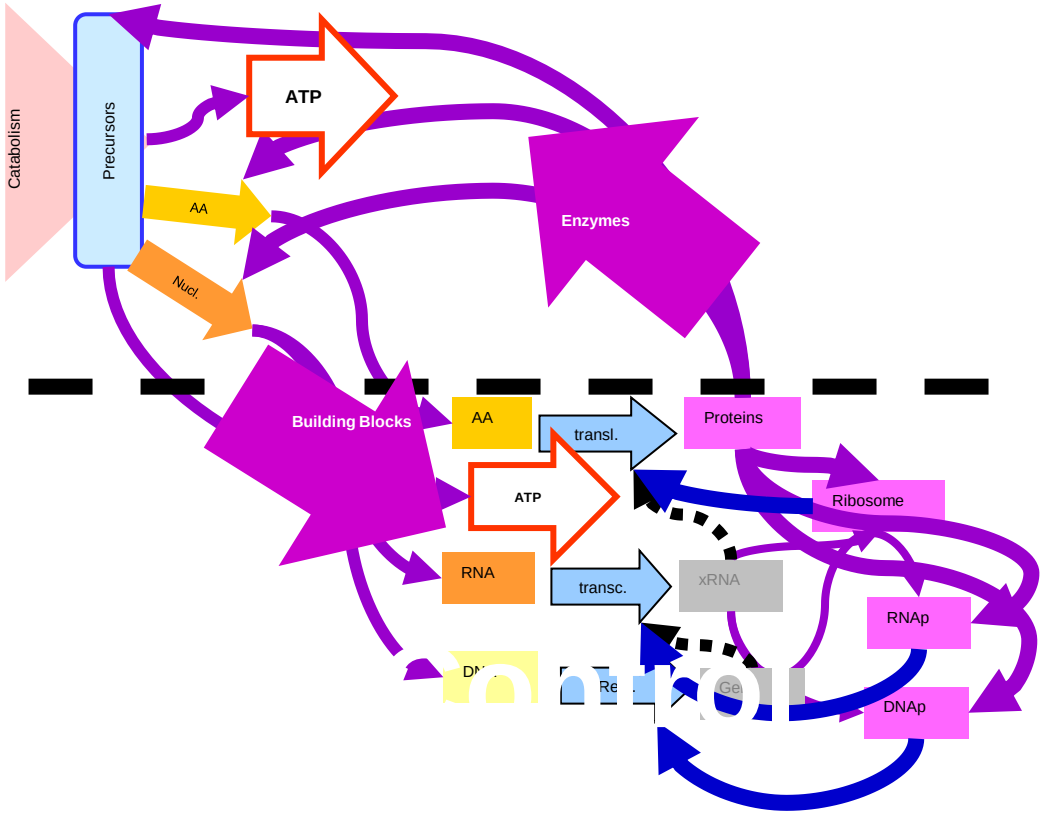
Deconstrained


Crosslayer autocatalysis
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Supply/demand control between layers?
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gen
e
DNAp


Ligands &
Receptors
Responses
control
Reactions
control
Protein
Assembly
control
DNA/RNA
Outside
Inside
Responses
DNA
Cross-layer control
• Highly organized
• Naming and addressing
• Prices? Duality?
• Minimal case study?


Ligands &
Receptors
nsmitter
Recei
Responses
control
Reactions
control
Protein
Assembly
control
DNA/RNA
Outside
Inside
Recei r
Responses
DNA

Upper megalayer/metalayer performs all cell functions, behaviors, scope is functional, distributed

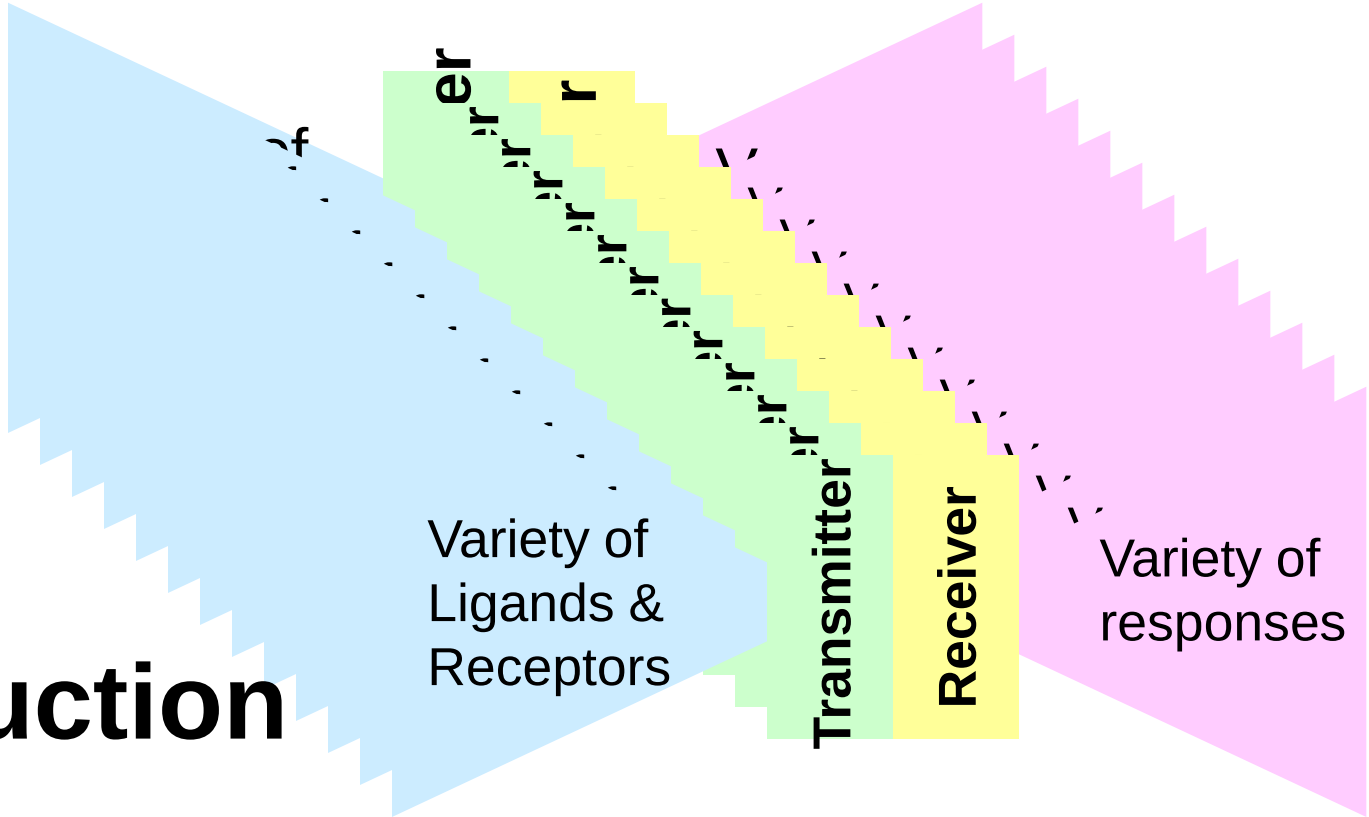


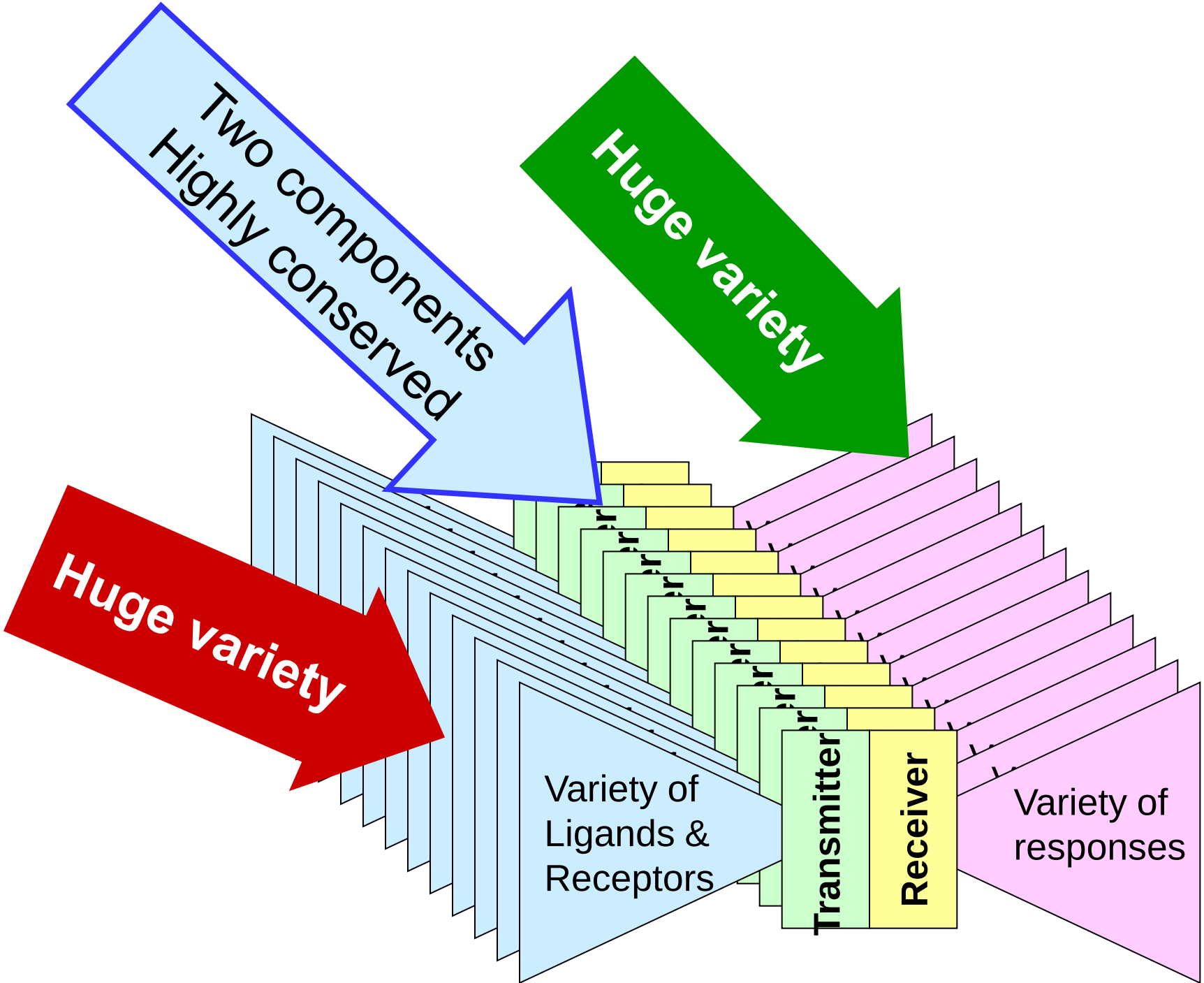
Signal transduction and transcription factors do name/address translation

Genome is physical, scope is location

- ≈50 such “two component” systems in *E. Coli*
- All use the same protocol
  - Histidine autokinase transmitter
  - Aspartyl phospho-acceptor receiver
- Huge variety of receptors and responses
- Also multistage (phosphorelay) versions

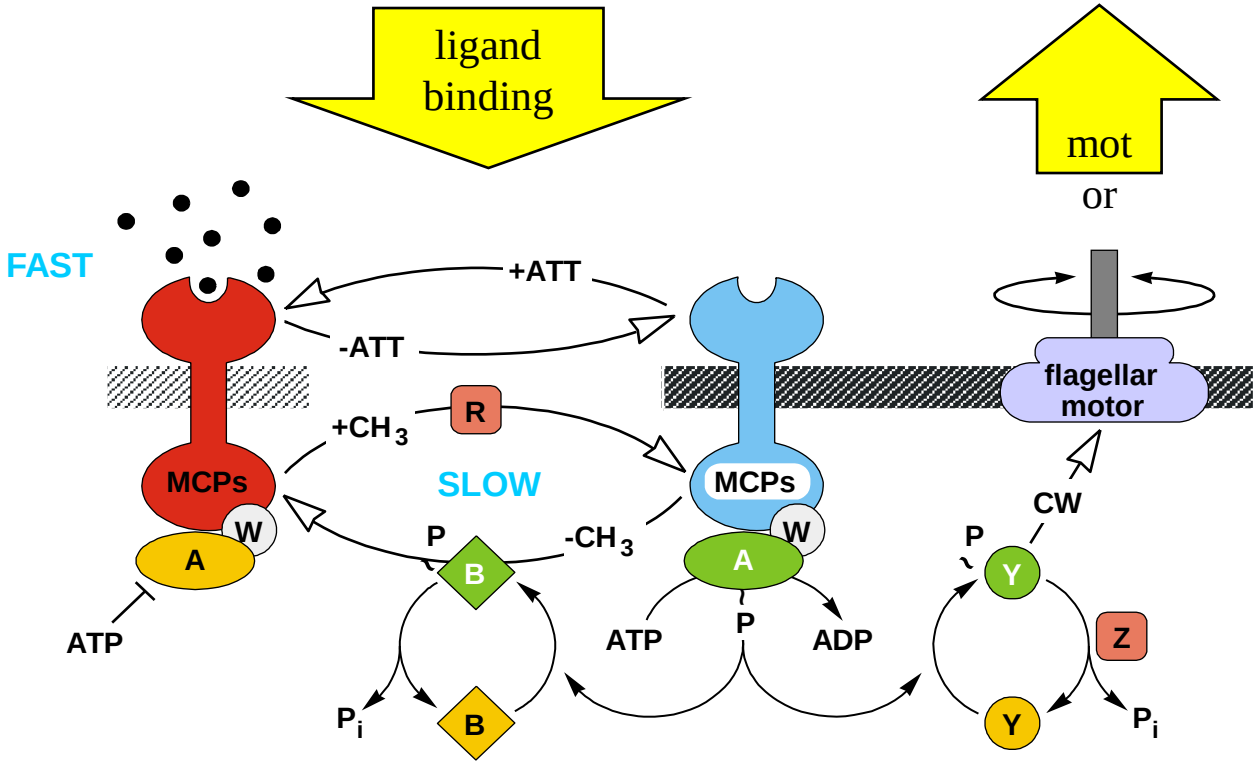
Signal transduction



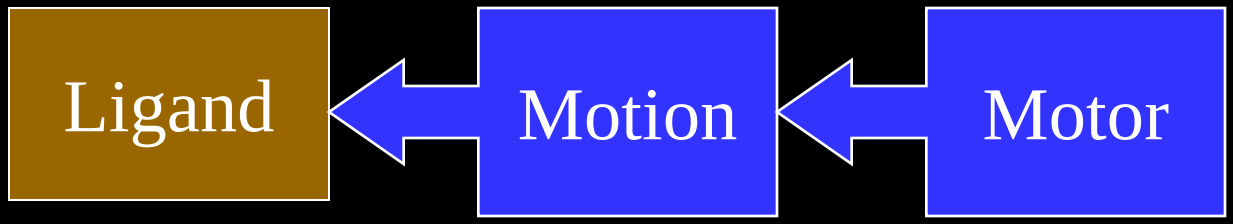

Two components
Highly conserved
Huge variety
Huge variety
Variety of
Ligands &
Receptors
Transmitter
Receiver
Variety of
responses

More necessity and robustness

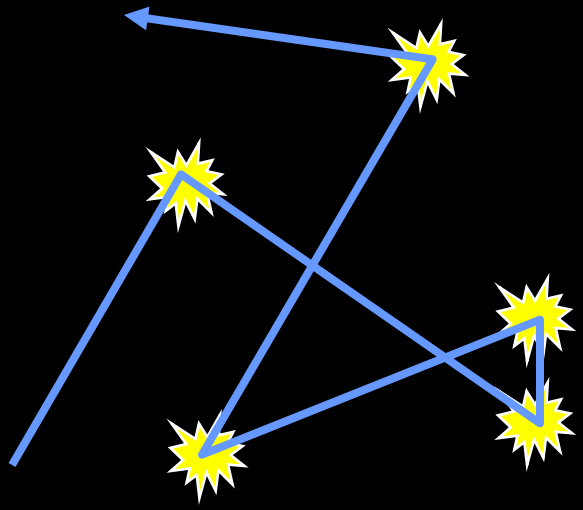
- Integral feedback and signal transduction (bacterial chemotaxis, G protein) (Yi, Huang, Simon)
- Example of "exploratory process"

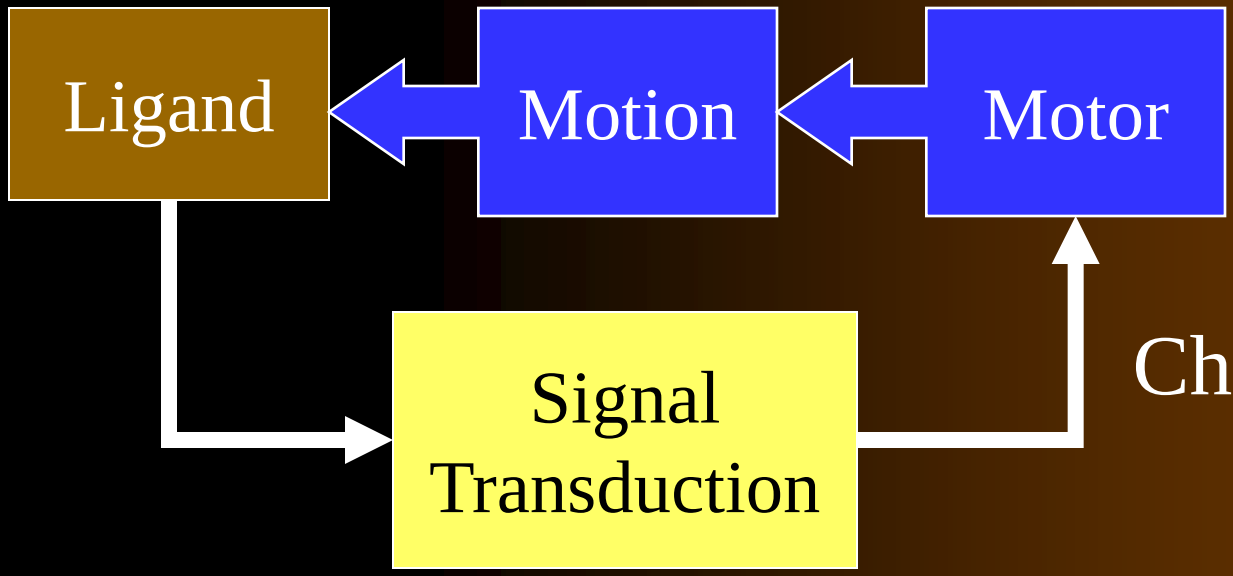


Bacterial chemotaxis

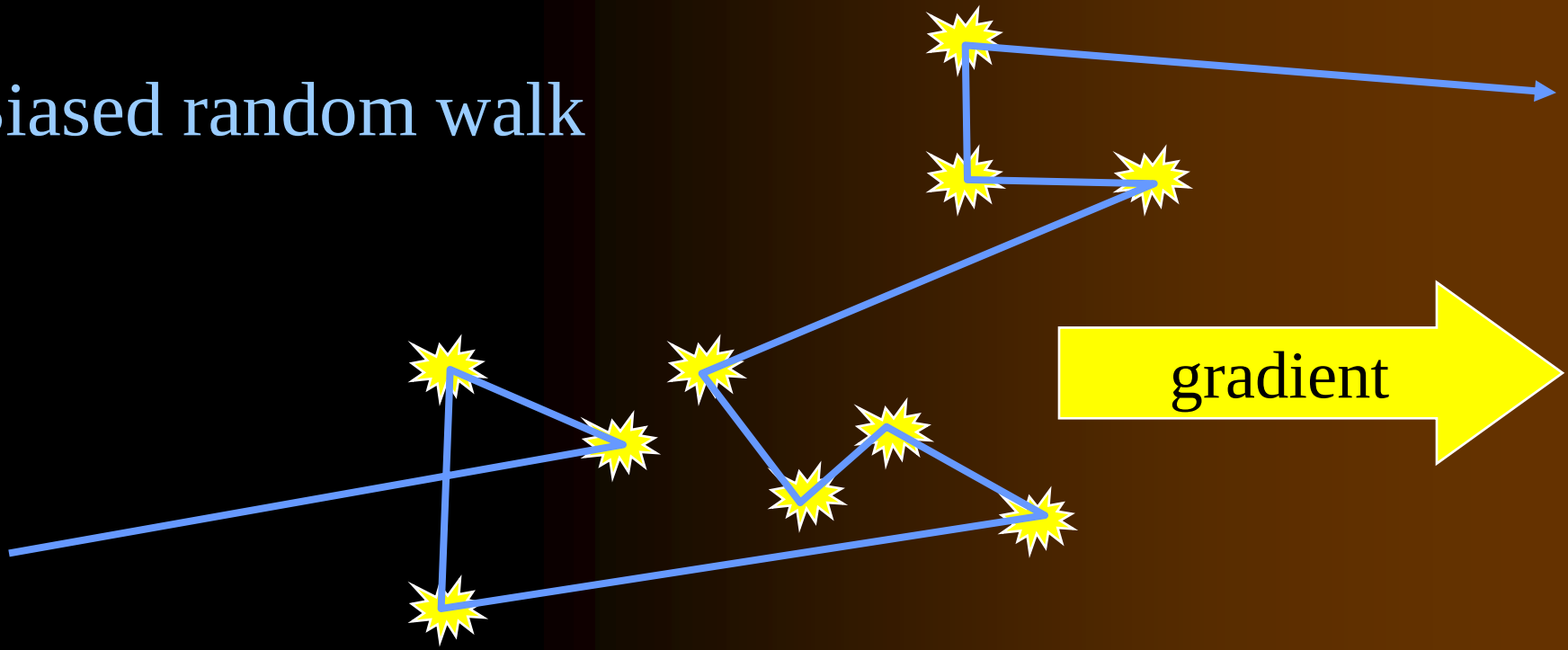


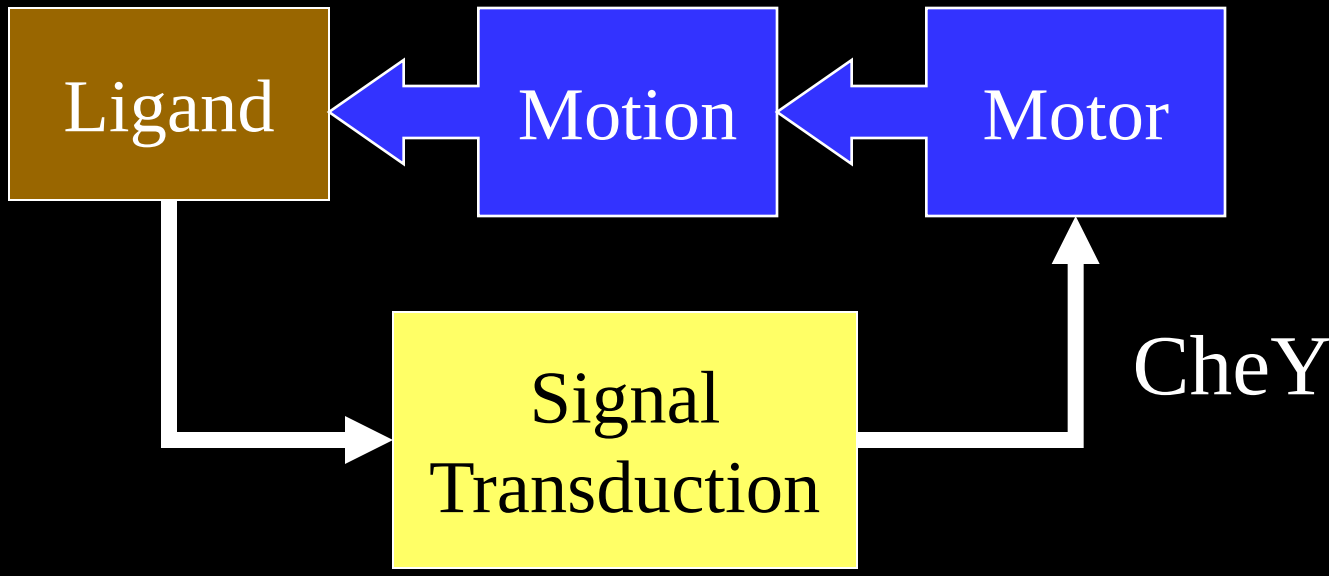
Random walk

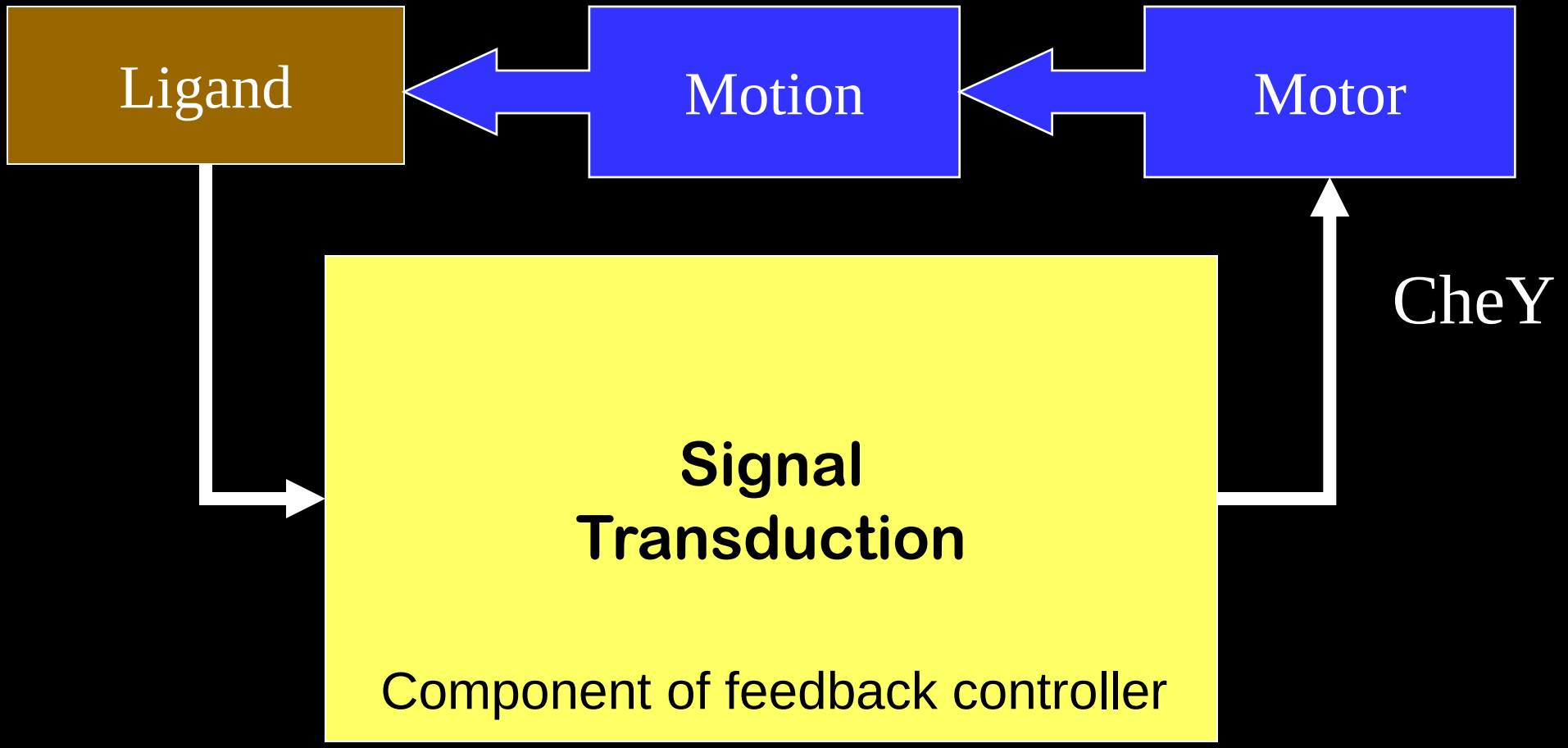


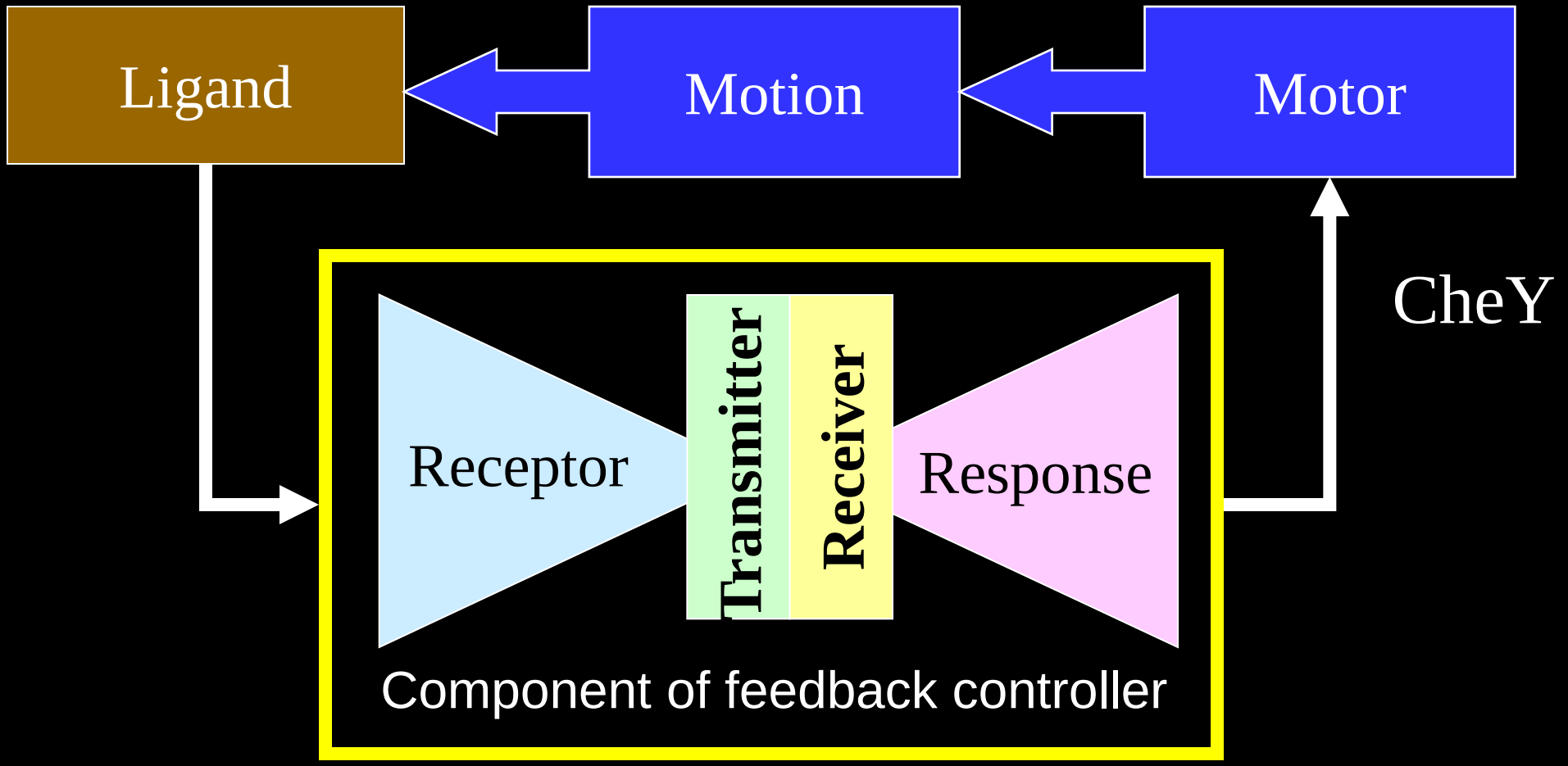


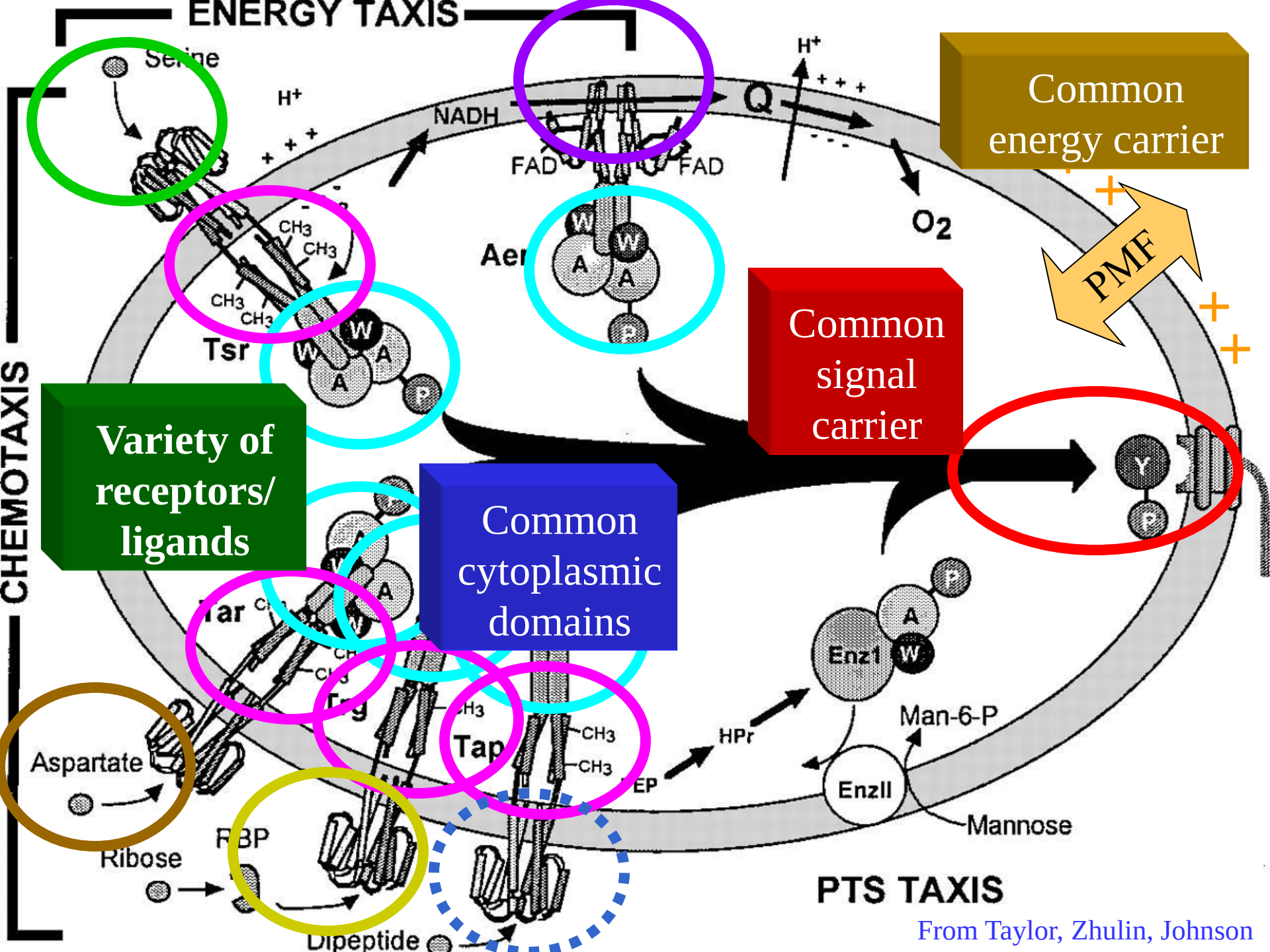
Biased random walk



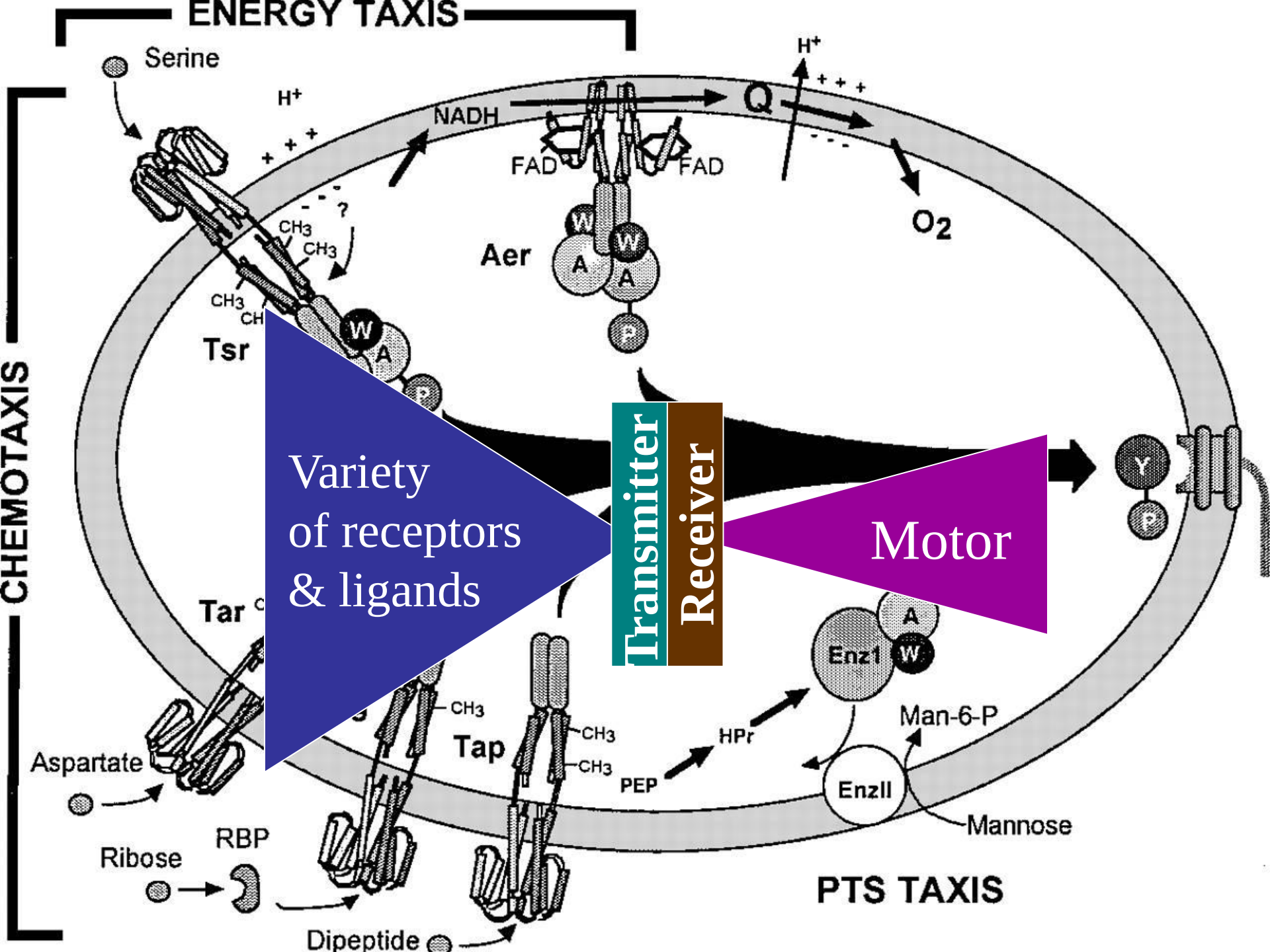

Ligand
Motion
Motor
Signal
Transduction
CheY

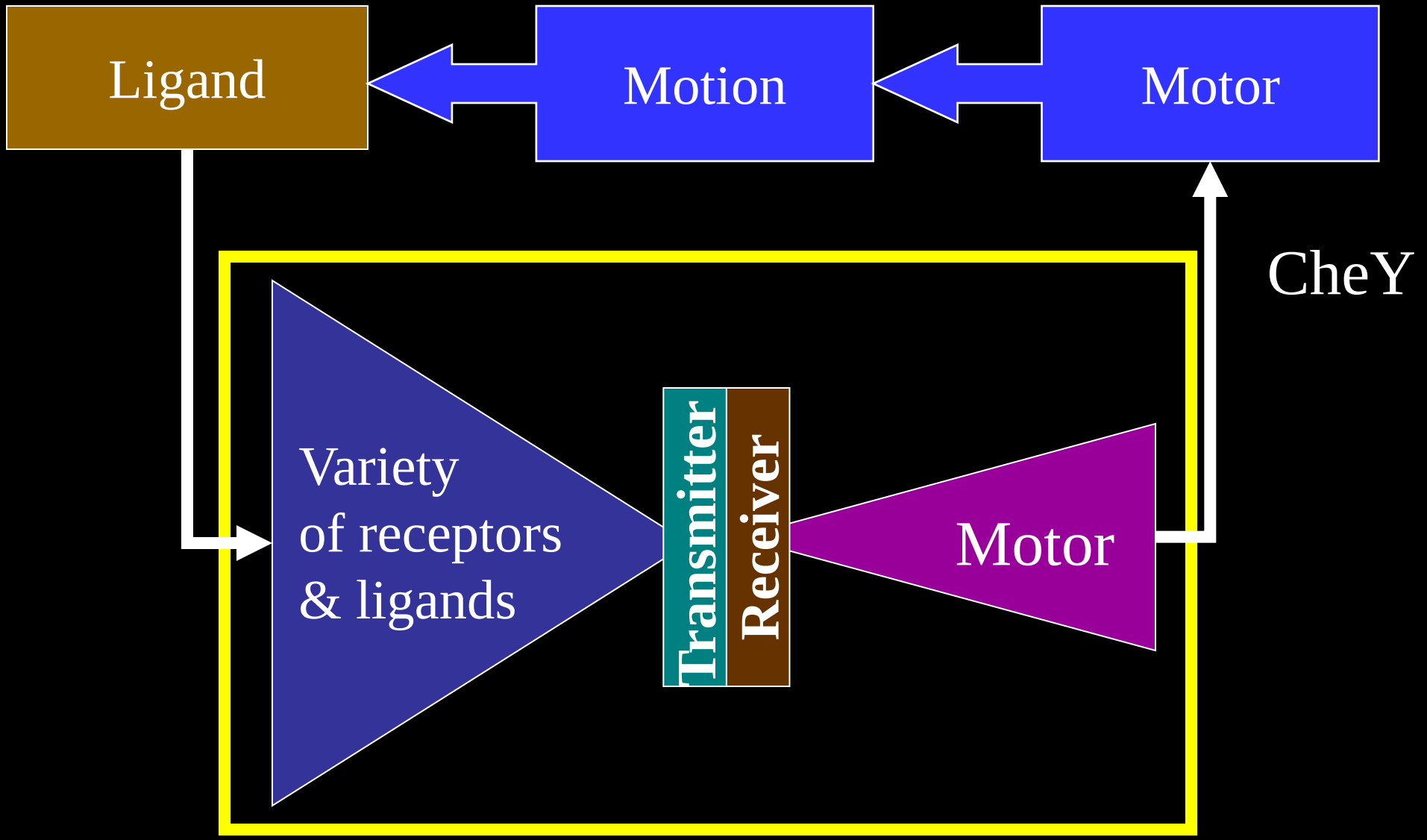

Ligand
Motion
Motor
CheY
Signal
Transduction
Component of feedback controller

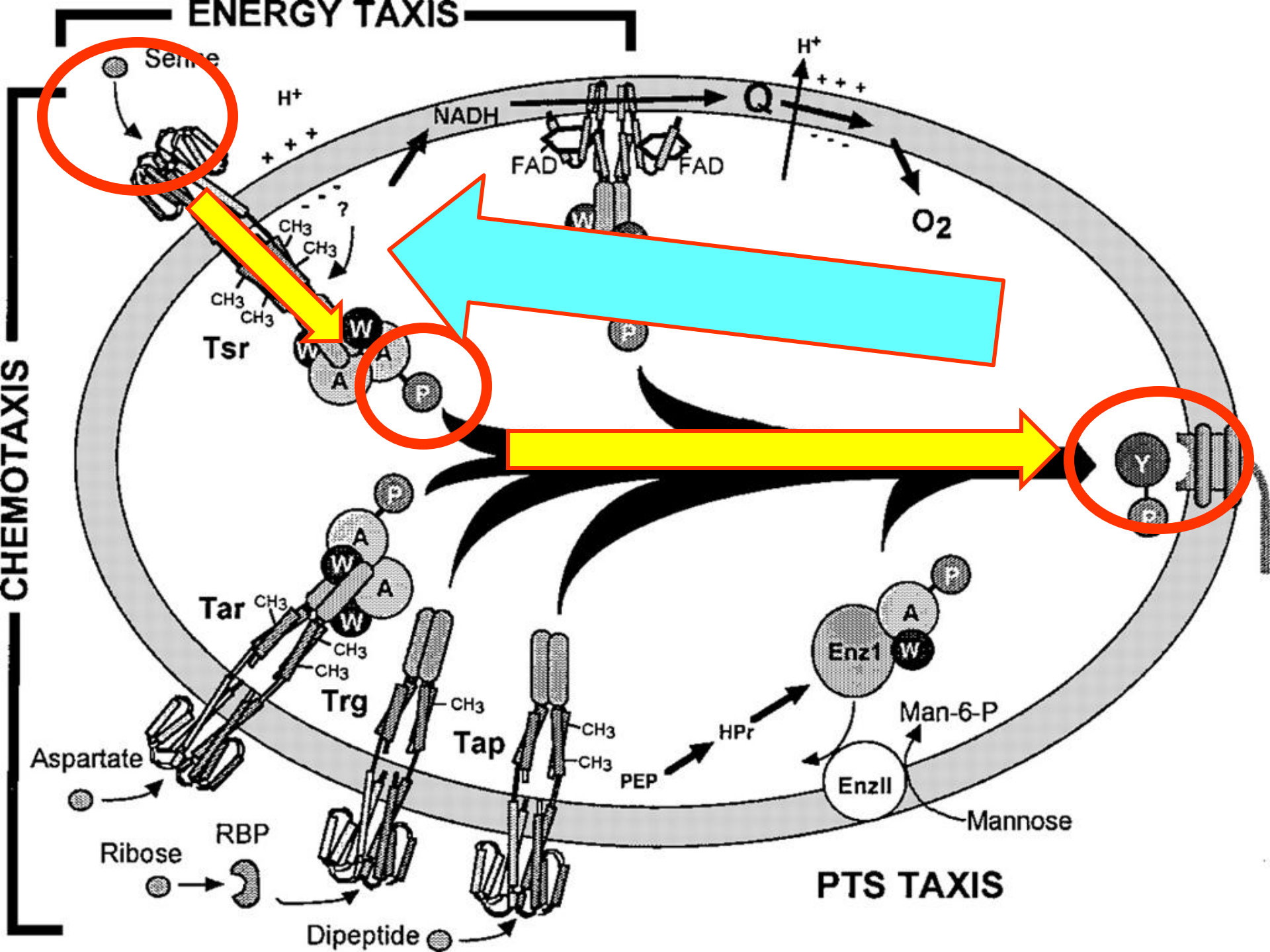

Ligand
Motion
Motor
CheY
Receptor
Transmitter
Receiver
Response
Component of feedback controller

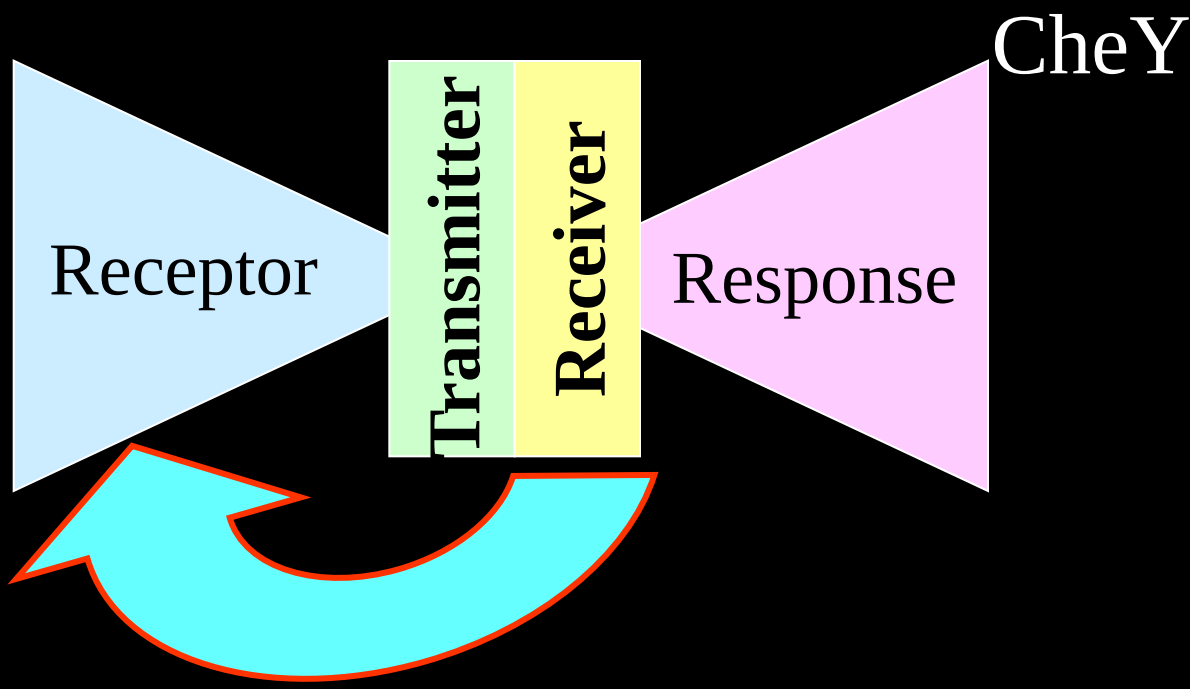


From Taylor, Zhulin, Johnson

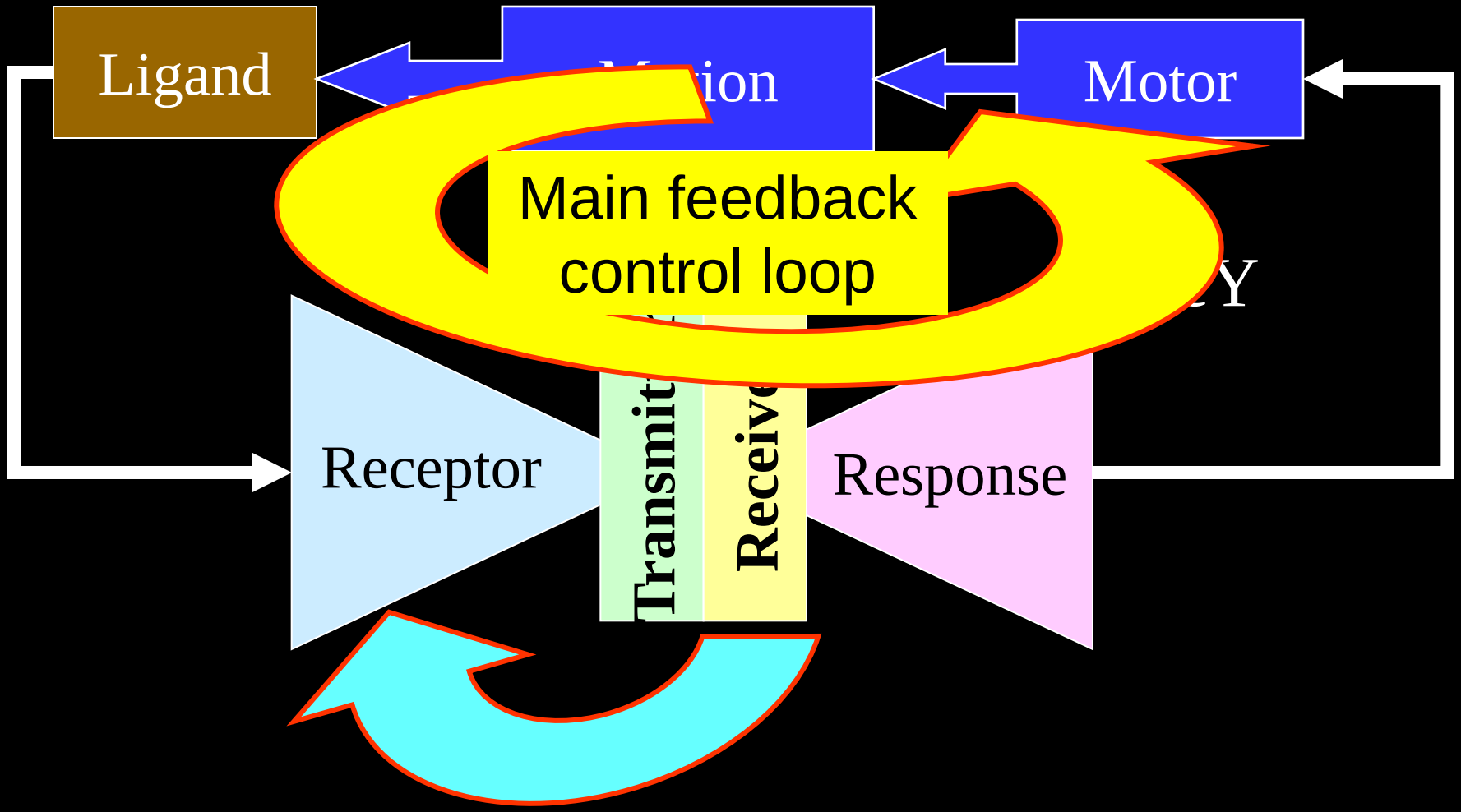

ENERGY TAXIS
CHEMOTAXIS
Serine
H+
NADH
Q
FAD
FAD
O2
Aer
Tsr
Tar
Tap
Variety
of receptors
& ligands
Transmitter
Receiver
Motor
Aspartate
Ribose
RBP
Dipeptide
PEP
HPr
Enz1
EnzII
Man-6-P
Mannose
PTS TAXIS


Ligand
Motion
Motor
CheY
Variety
of receptors
& ligands
Transmitter
Receiver
Motor

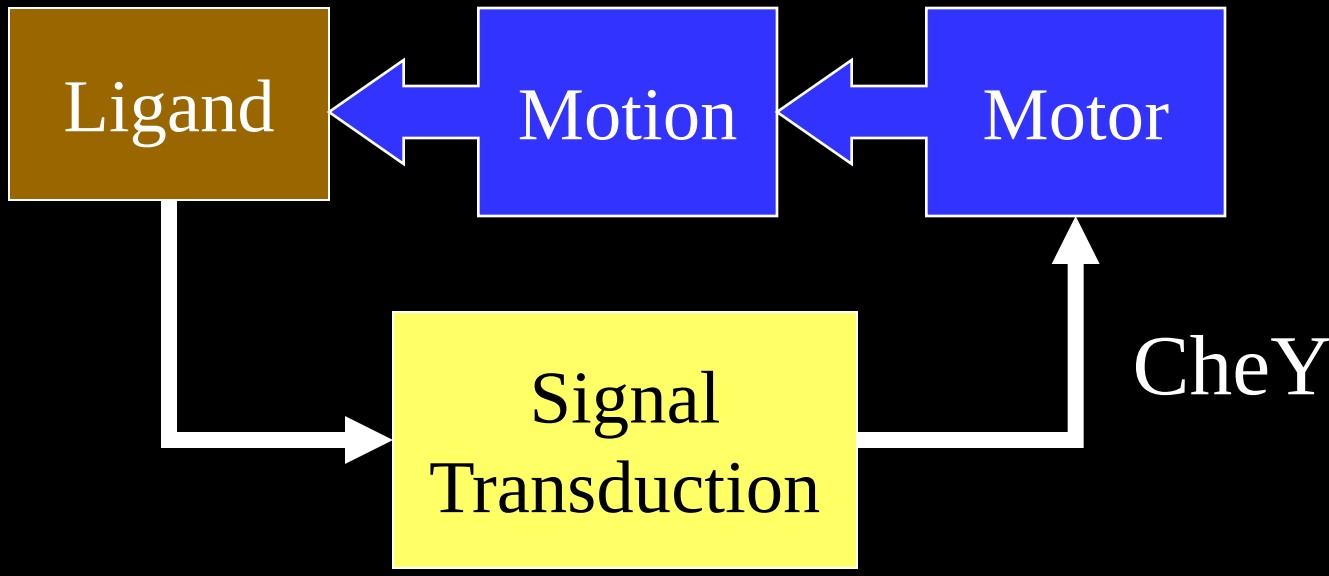

ENERGY TAXIS
CHEMOTAXIS
Sense
H+
NADH
FAD
FAD
Q
H+
O2
CH3
CH3
CH3
CH3
?
Tsr
W
A
P
Tar
Trg
Tap
CH3
Aspartate
RBP
Ribose
Dipeptide
PEP
HPr
Enz1
EnzII
Man-6-P
Mannose
Y
PTS TAXIS

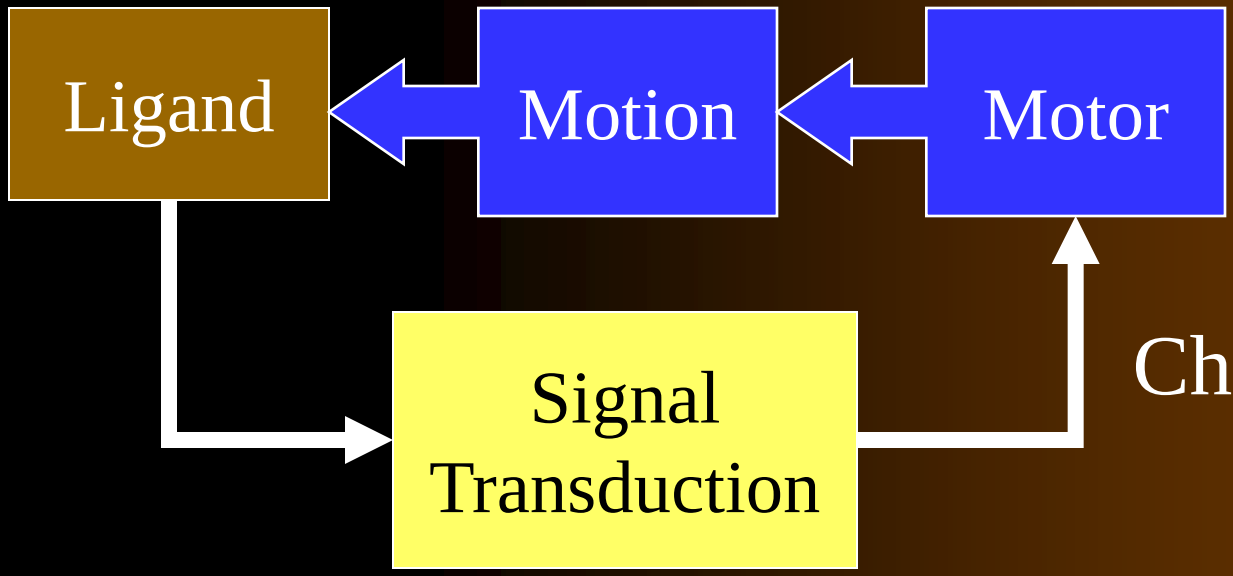


Integral feedback internal to signaling network

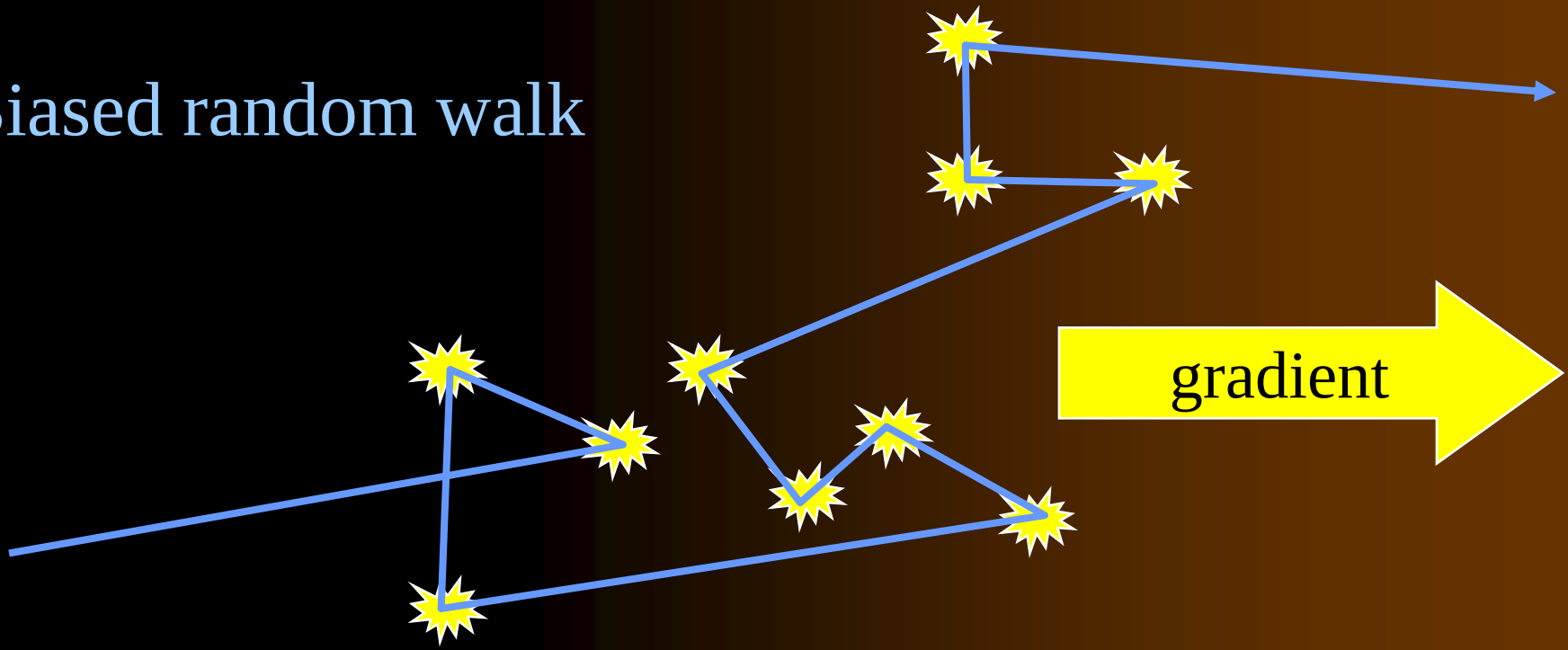


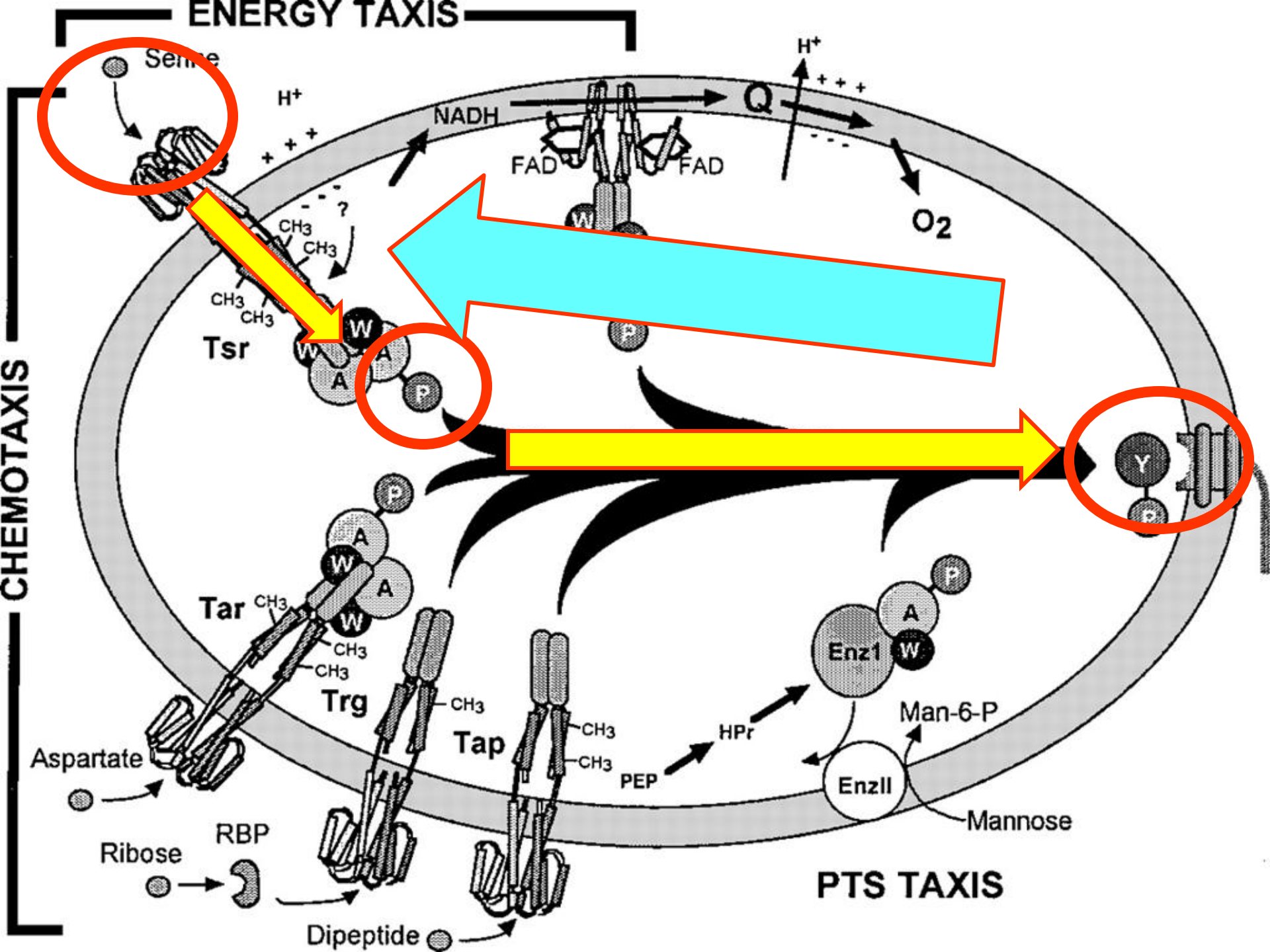
Integral feedback internal to signaling network


Ligand
Motion
Motor
Signal
Transduction
CheY

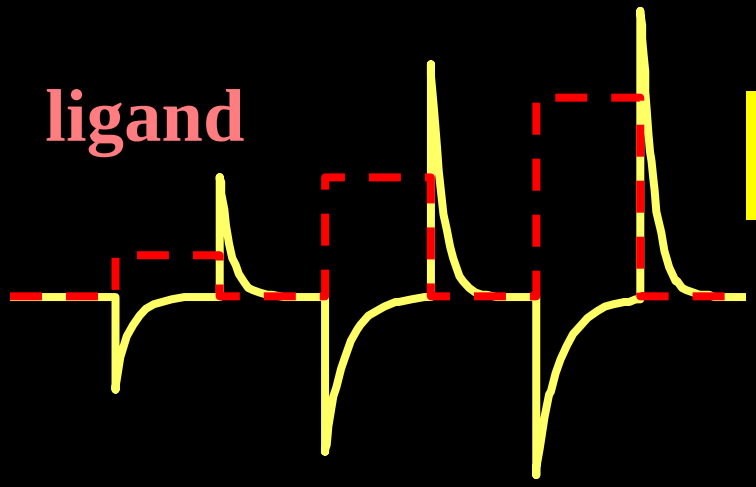


Biased random walk

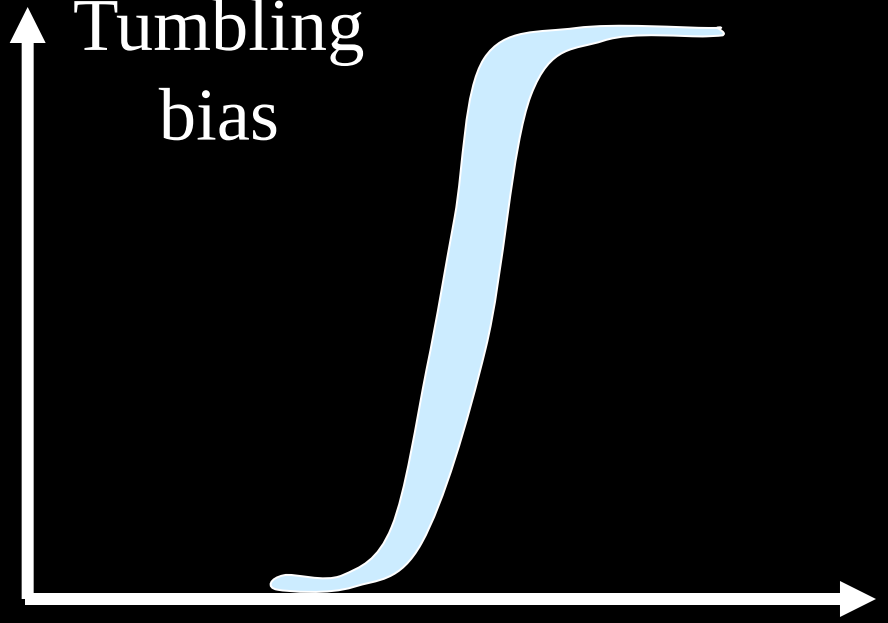



ENERGY TAXIS
CHEMOTAXIS
Sense
H+
NADH
FAD
FAD
Q
H+
O2
CH3
CH3
CH3
CH3
?
Tsr
W
A
P
Tar
Trg
Tap
CH3
Aspartate
RBP
Ribose
Dipeptide
PEP
HPr
Enz1
EnzII
Man-6-P
Mannose
Y
PTS TAXIS

Perfect adaptation is ***necessary*** …



$CheY_p$

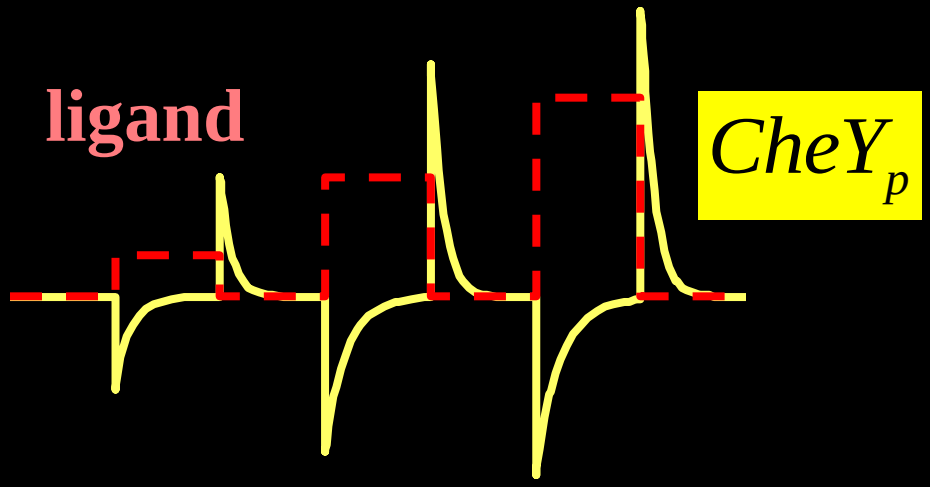


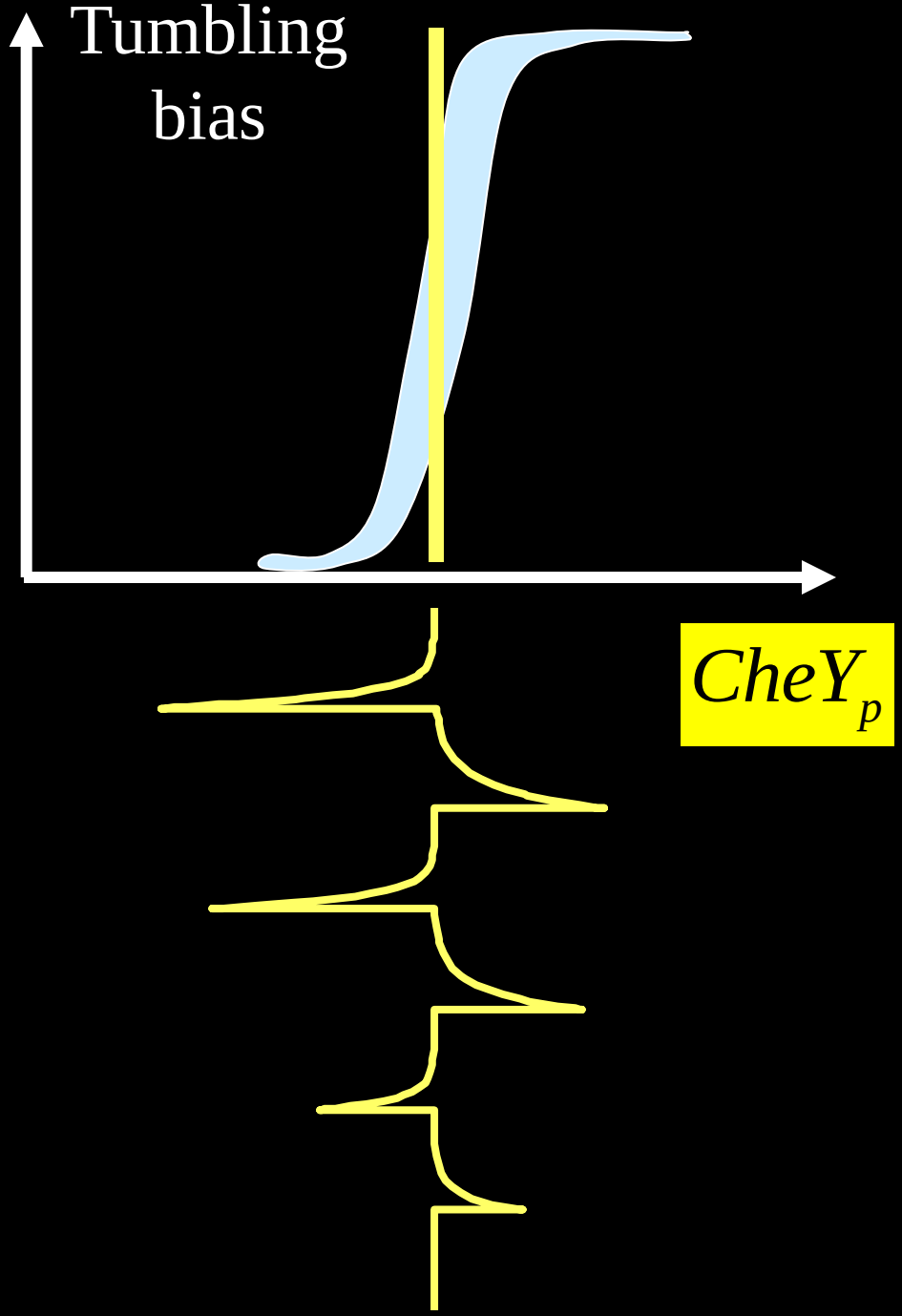
Motor

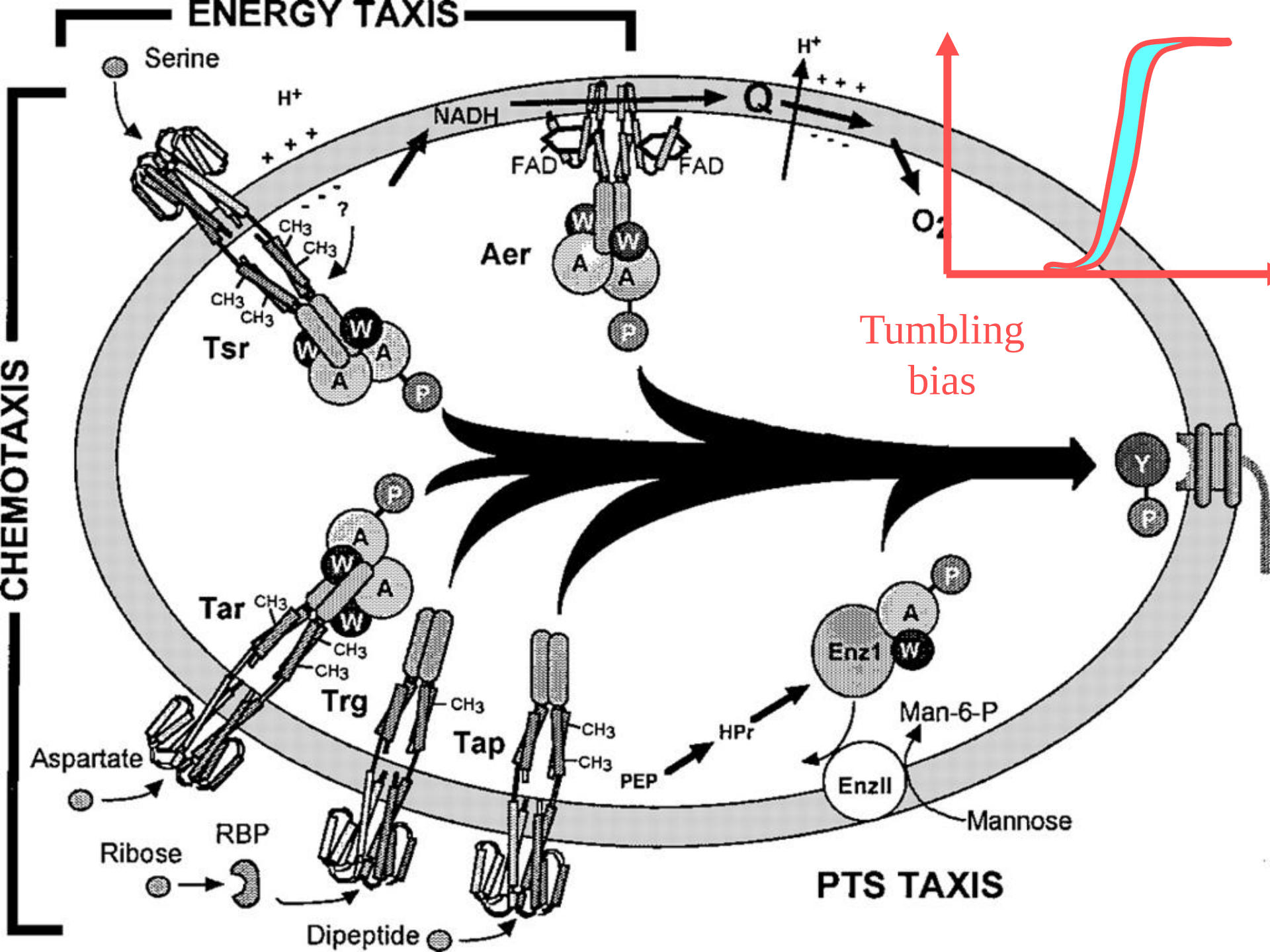
Signal Transduction

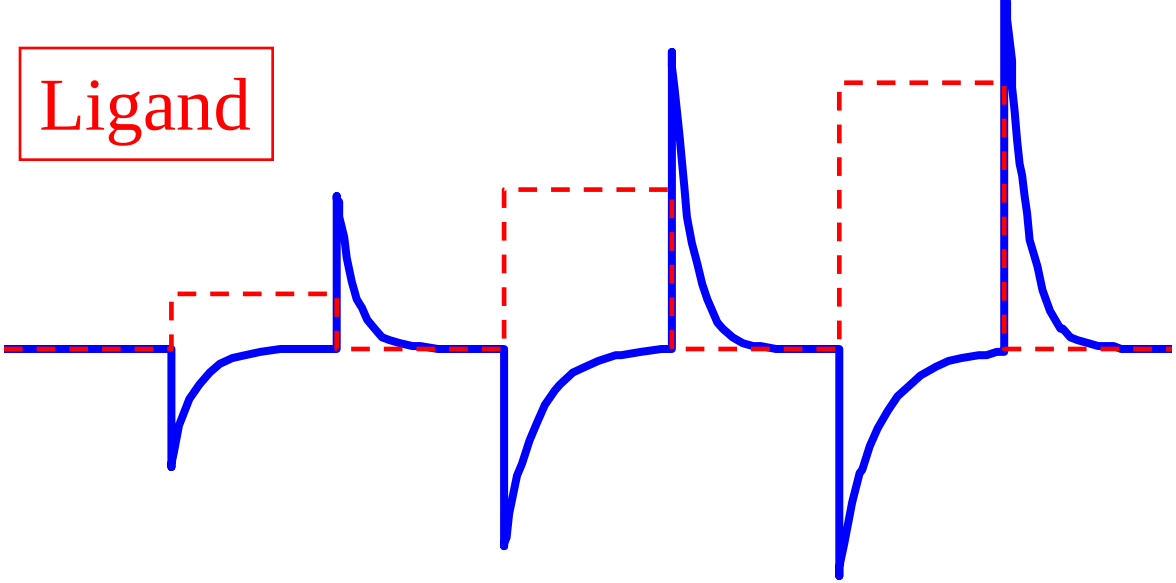
Perfect adaptation is ***necessary*** …

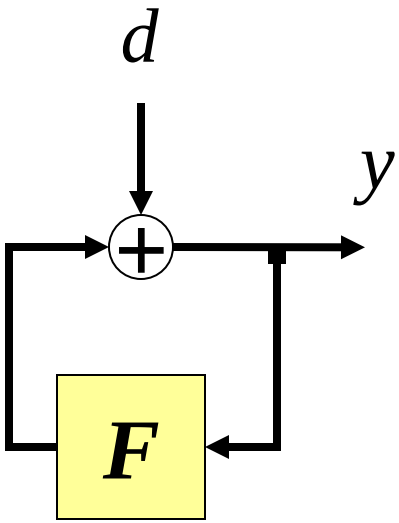
…to keep CheYp in the responsive range of the motor.






ENERGY TAXIS
CHEMOTAXIS
Serine
H+
NADH
Q
FAD
FAD
O2
Aer
Tsr
Tar
Trg
Tap
CH3
W
A
P
Tumbling
bias
Y
Enz1
HPr
PEP
Man-6-P
EnzII
Mannose
Aspartate
RBP
Ribose
Dipeptide
PTS TAXIS



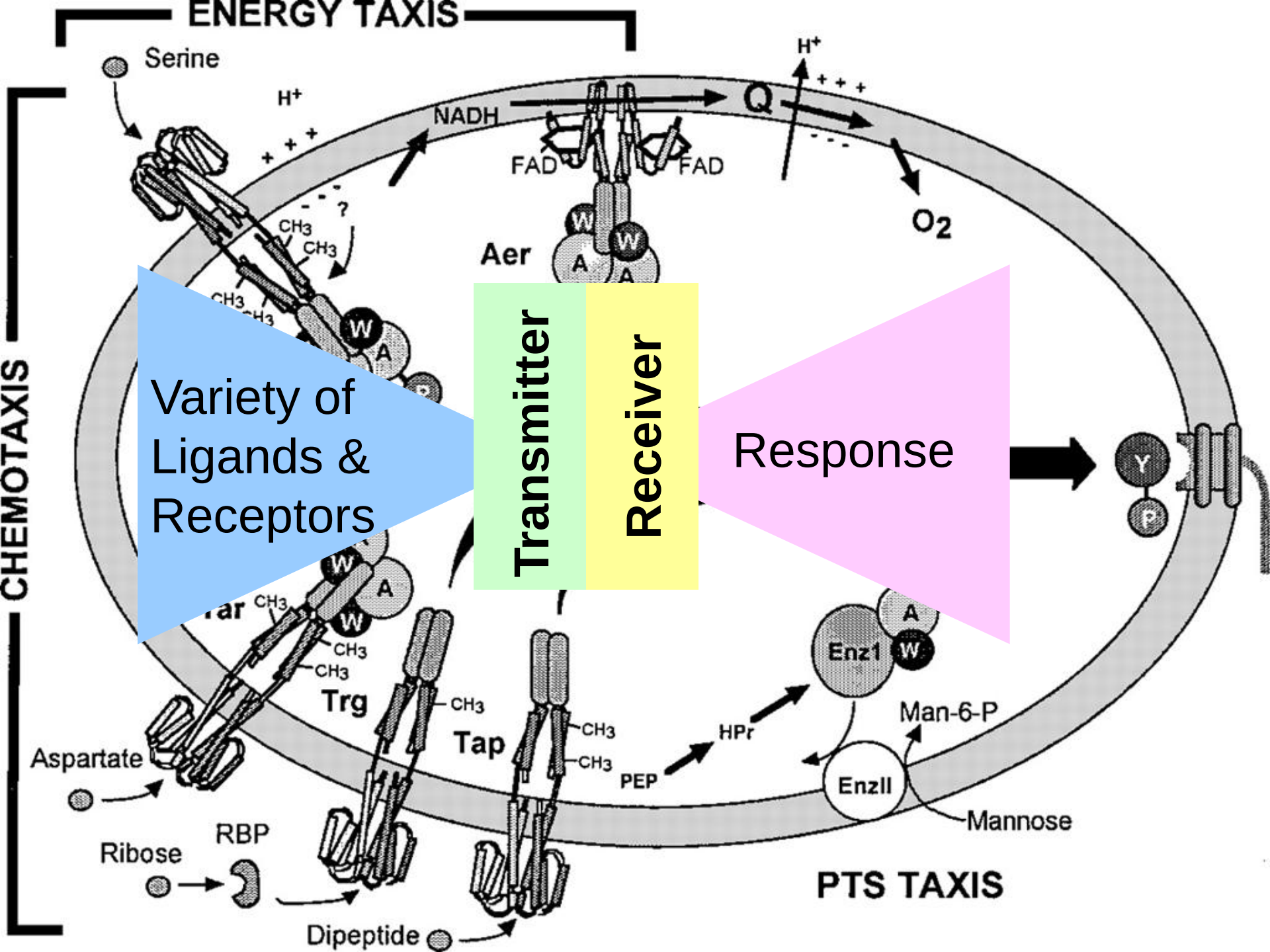


Integral feedback

$$F \to -\infty$$
$$\ln(S) \to -\infty$$

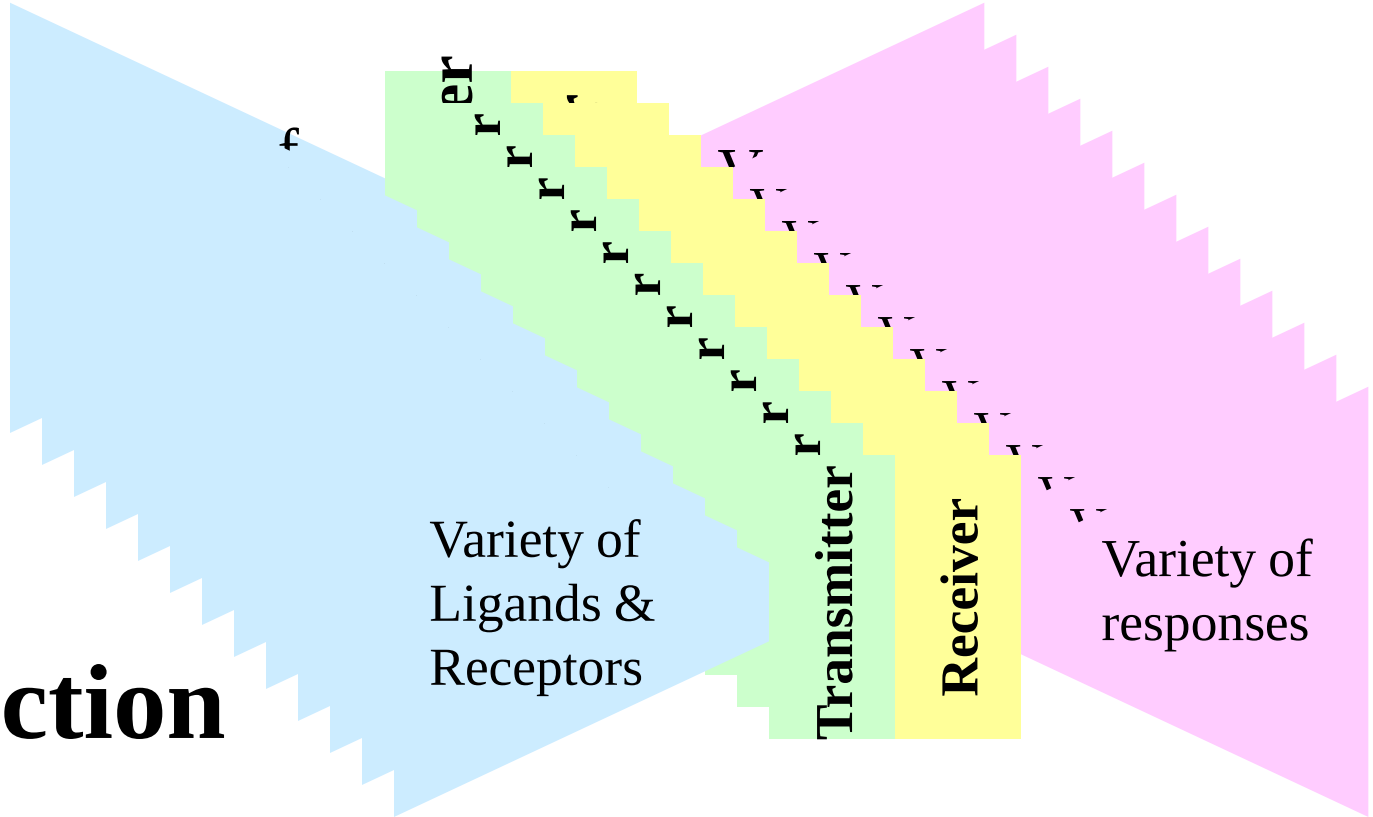
$$F(s) = \frac{\hat{F}(s)}{s}, \quad \hat{F}(0) < 0$$
$$\Updownarrow$$
$$F(0) = -\infty$$
$$\Updownarrow$$
$$S(0) = 0$$

$$S \equiv \frac{y}{d} = \frac{1}{1-F}$$


ENERGY TAXIS
CHEMOTAXIS
Serine
H+
NADH
Q
FAD
FAD
O2
Aer
CH3
Variety of Ligands & Receptors
Transmitter
Receiver
Response
Y
P
Tar
Trg
Tap
Aspartate
RBP
Ribose
Dipeptide
Enz1
HPr
PEP
Man-6-P
EnzII
Mannose
PTS TAXIS

- ≈50 such “two component” systems in *E. Coli*
- All use the same protocol
  - Histidine autokinase transmitter
  - Aspartyl phospho-acceptor receiver
- Huge variety of receptors and responses
- Also multistage (phosphorelay) versions

Signal transduction



Flow of “signal”

Recognition, specificity

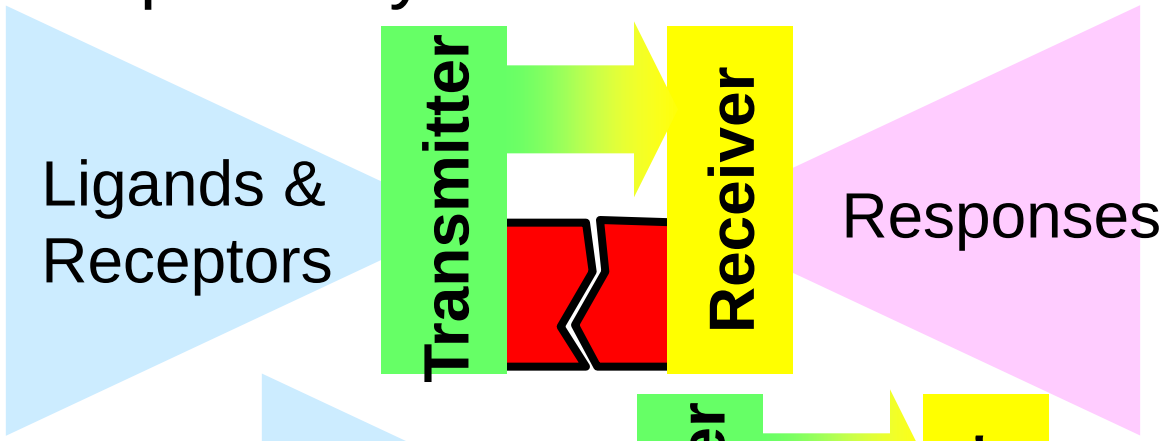
Shared protocols

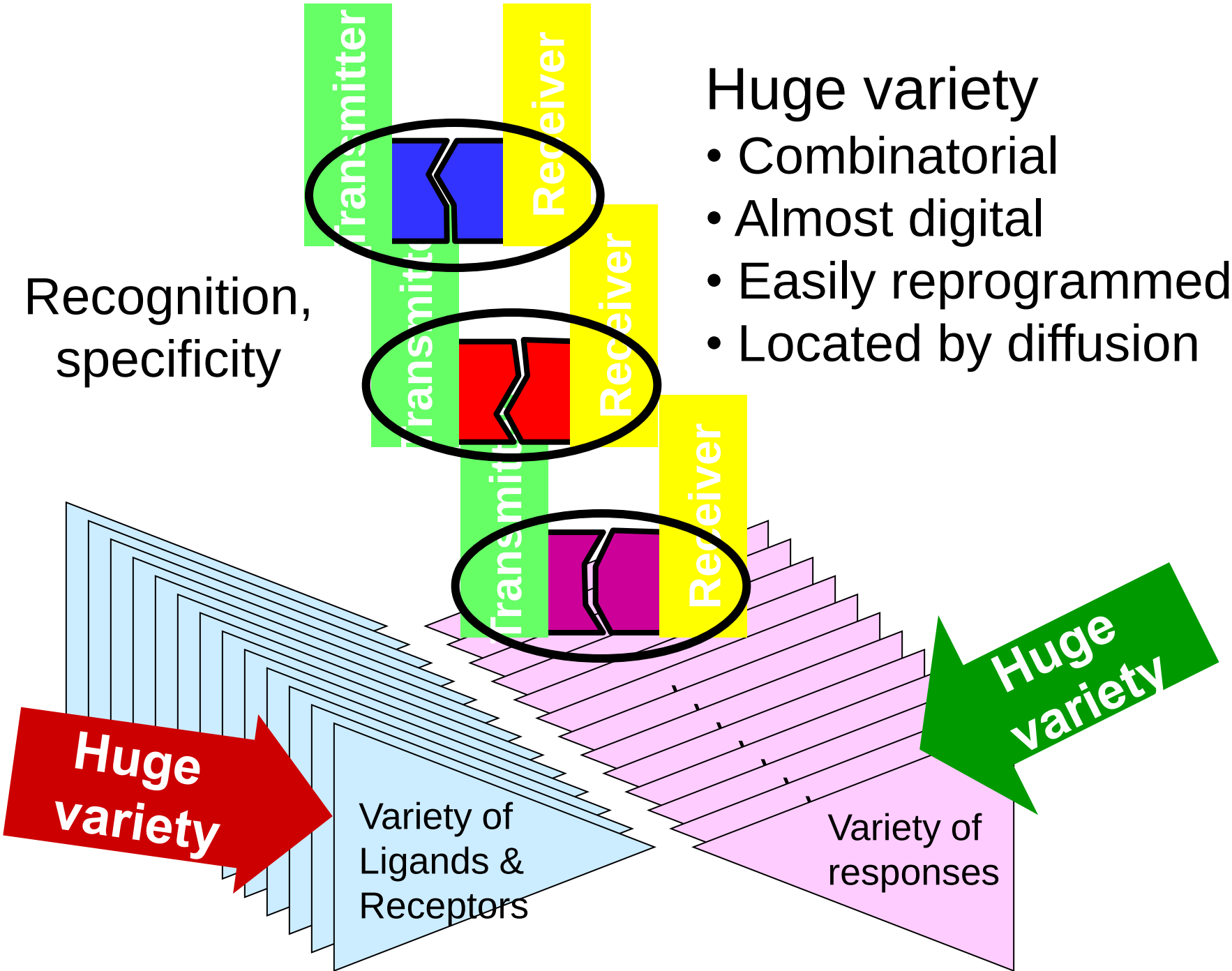
- “Name resolution” within signal transduction
- Transmitter must locate “cognate” receiver and avoid non-cognate receivers
- Global search by rapid, local diffusion
- Limited to very small volumes

Flow of “signal”

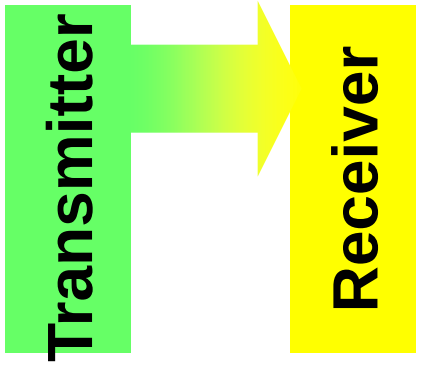
Shared protocols

Recognition, specificity



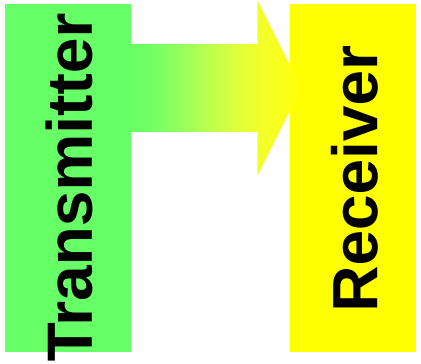

Transmitter
Receiver
Recognition,
specificity
Huge variety
• Combinatorial
• Almost digital
• Easily reprogrammed
• Located by diffusion
Huge variety
Variety of
Ligands &
Receptors
Huge variety
Variety of
responses

Flow of “signal”

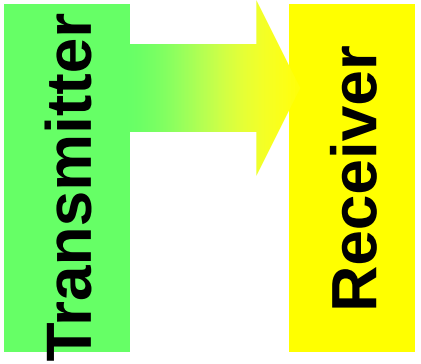


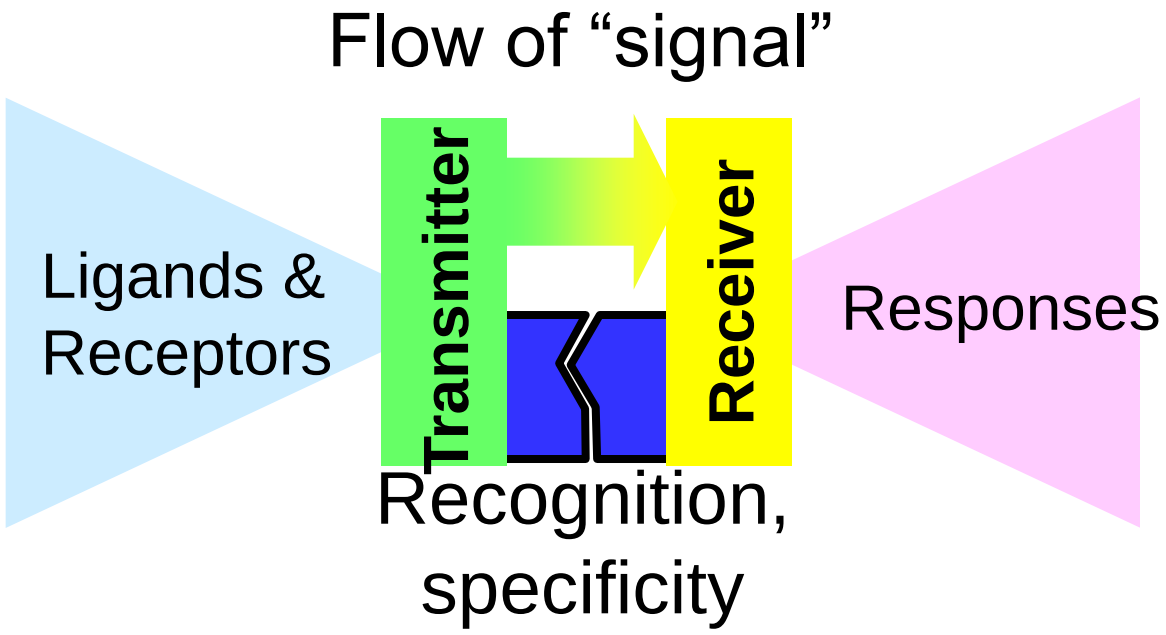
Limited variety

- Fast, analog (via #)
- Hard to change



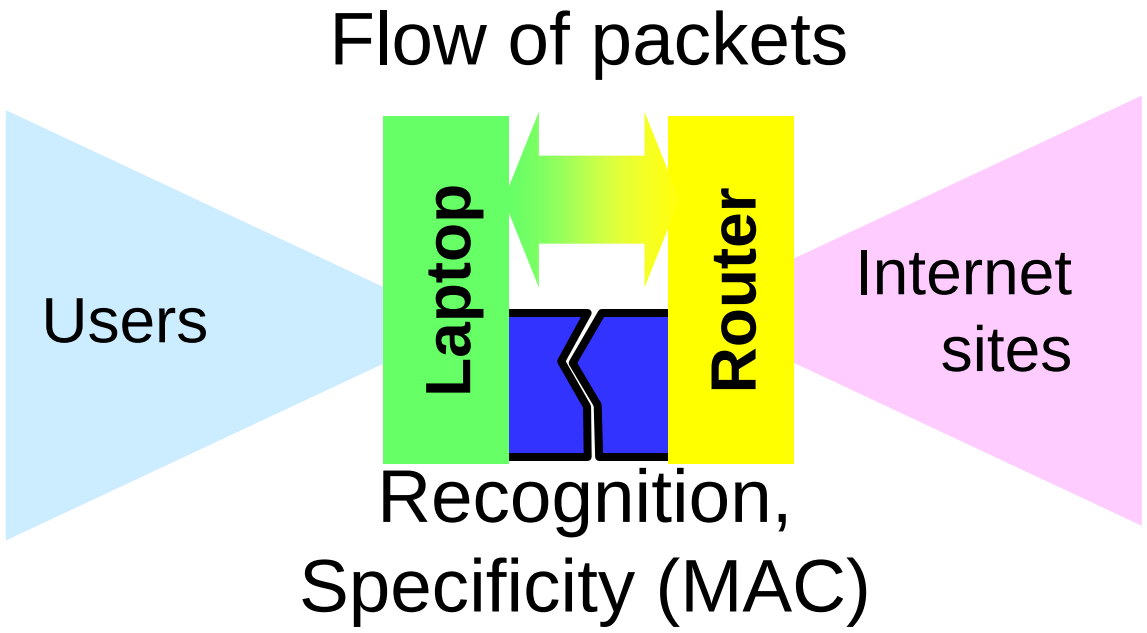
Reusable in different pathways





Shared protocols

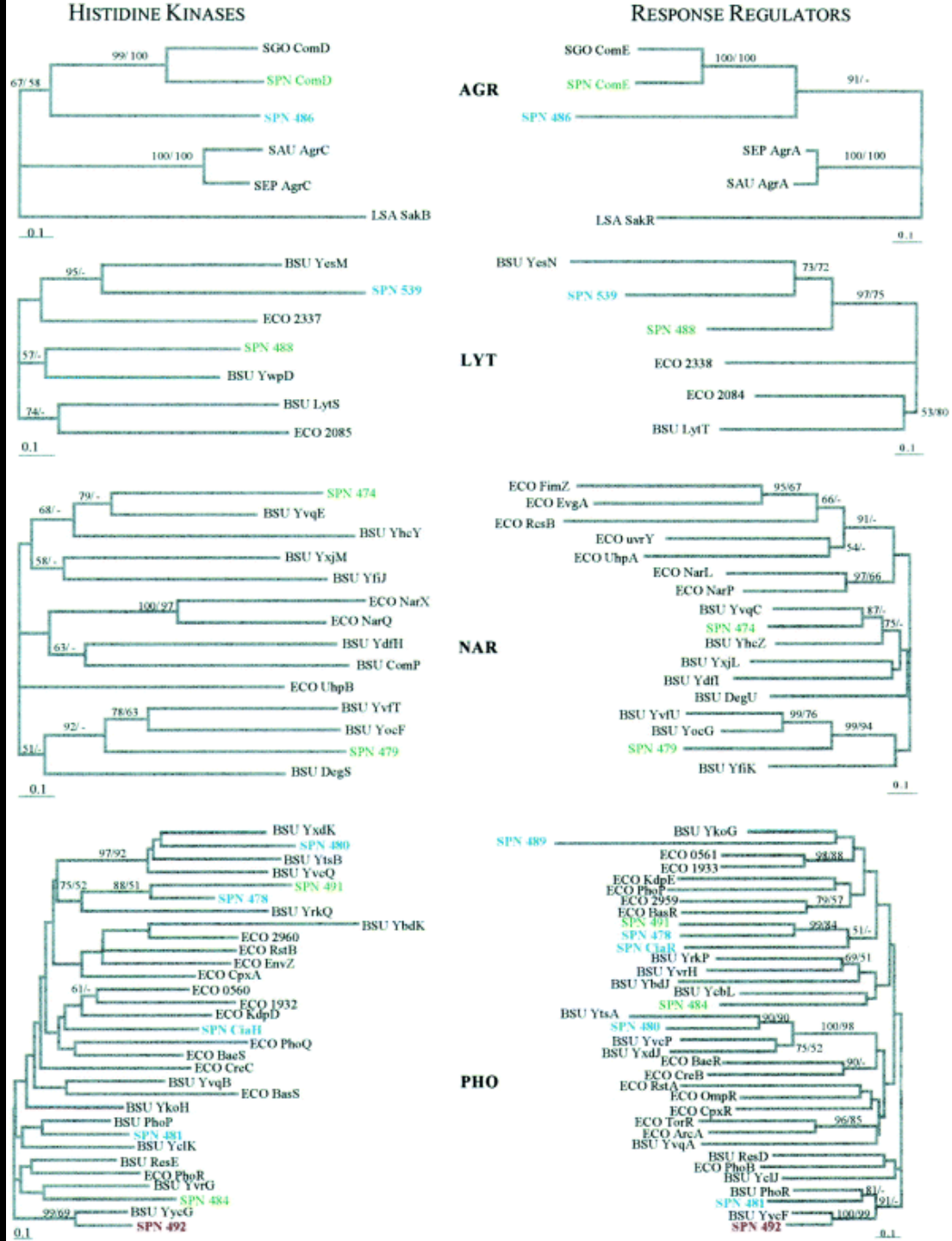
Note: Any wireless system and the Internet to which it is connected work the same way.

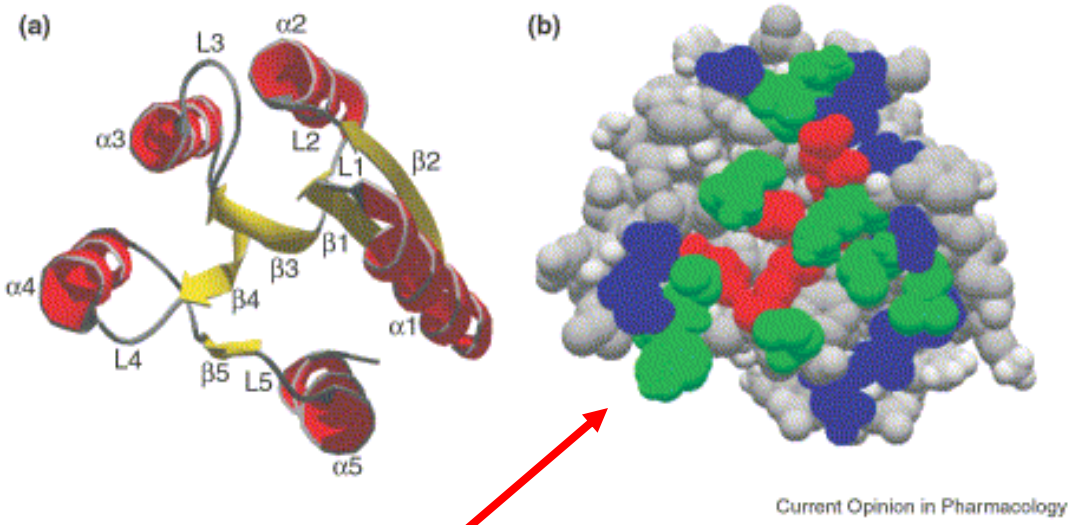


Molecular phylogenies show evolvability of this bowtie architecture.

“Name” recognition is almost digital.

Response regulators can translate these names to DNA addresses with another DNA-binding domain (also digital).

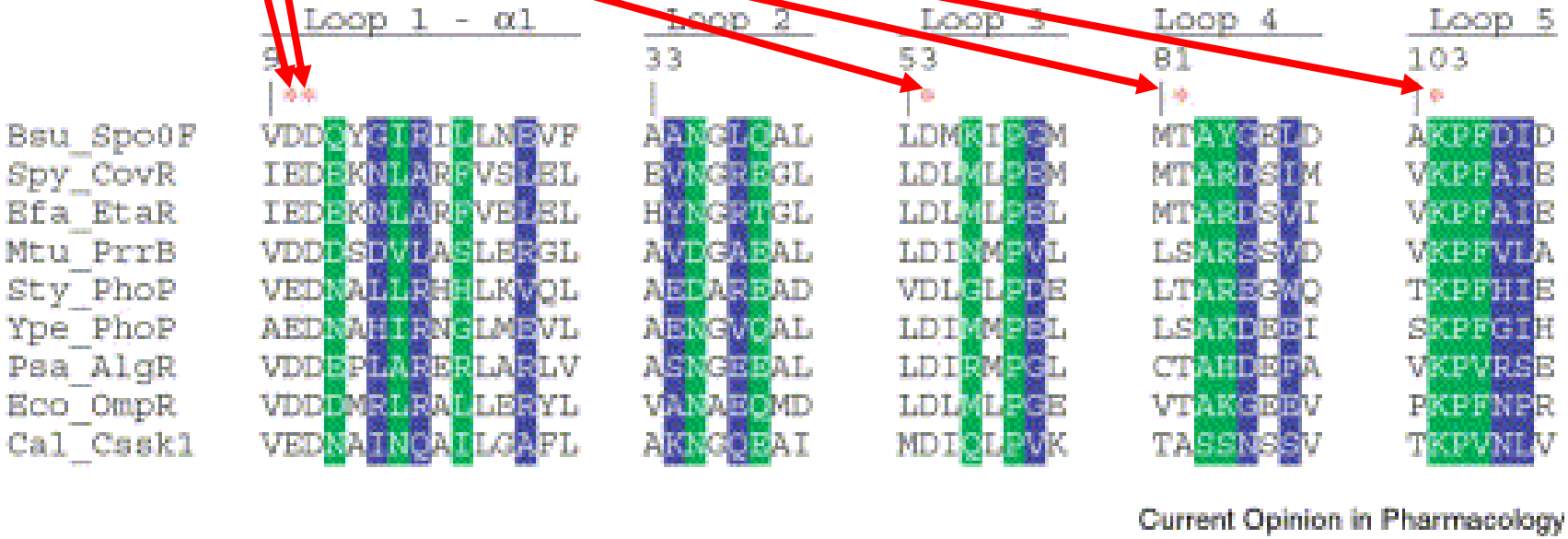


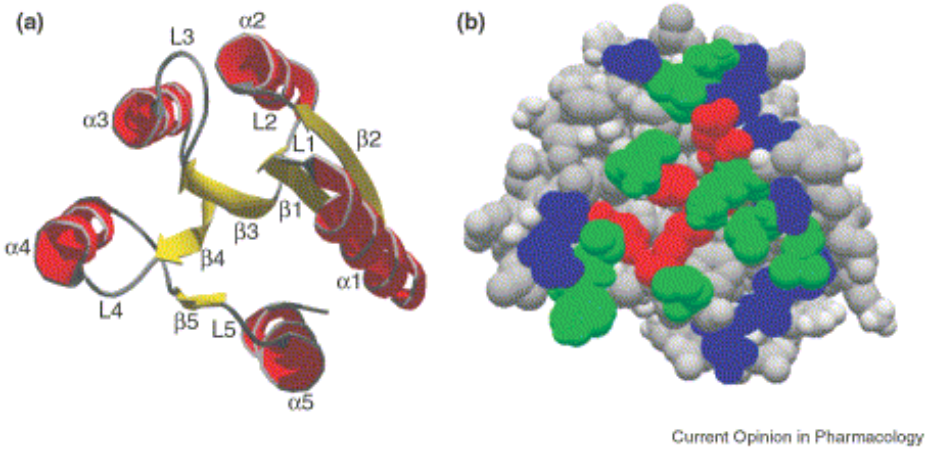


conserved residues of interaction surface with phosphotransferase domains

highly variable amino acids of the interaction surface that are responsible for specificity of the interaction

invariant active-site residues





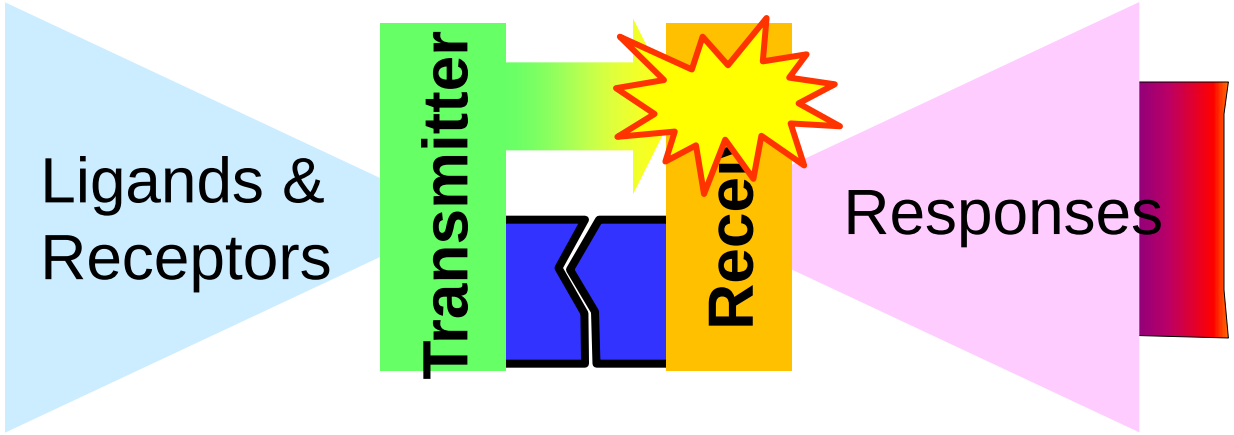
conserved functional domains

invariant active-site residues

highly variability for specificity of the interaction

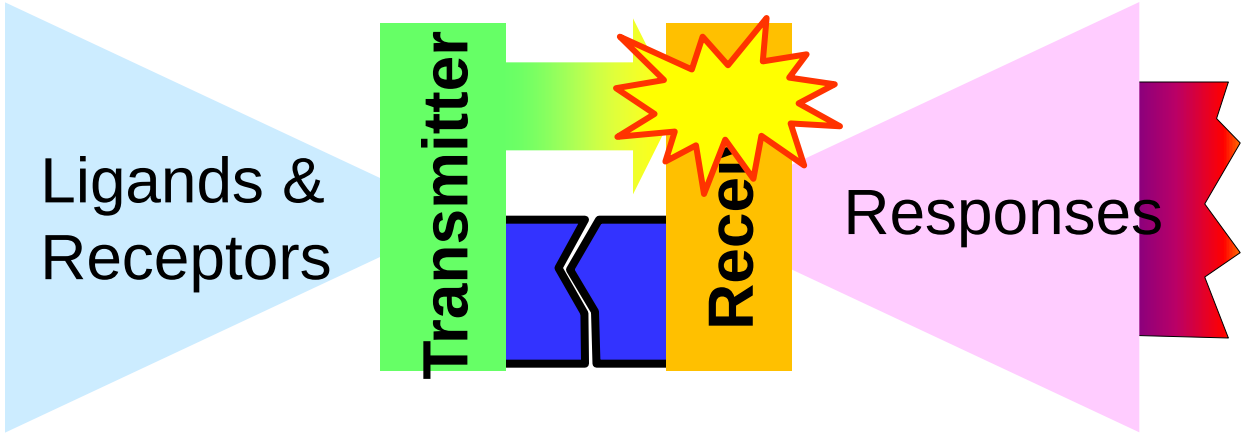
- **Automobiles**: Keys provide specificity but no other function. Other function conserved, driver/vehicle interface protocol is "universal."
- **Ethernet cables**: Specificity via MAC addresses, function via standardized protocols.

MAC



“Name” recognition
= molecular recognition
= localized functionally
= global spatially

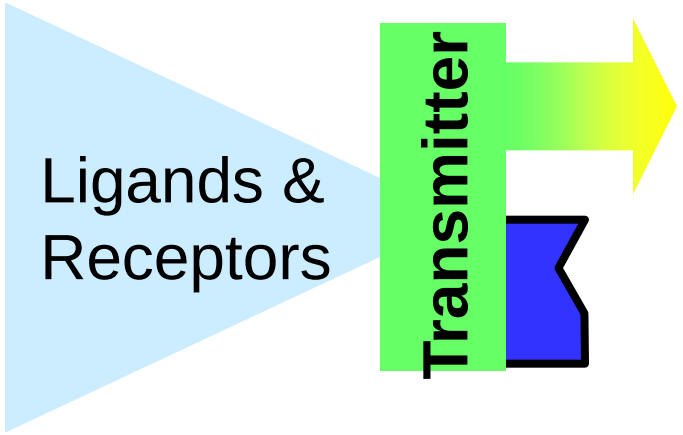
Transcription factors
do “name” to “address”
translation



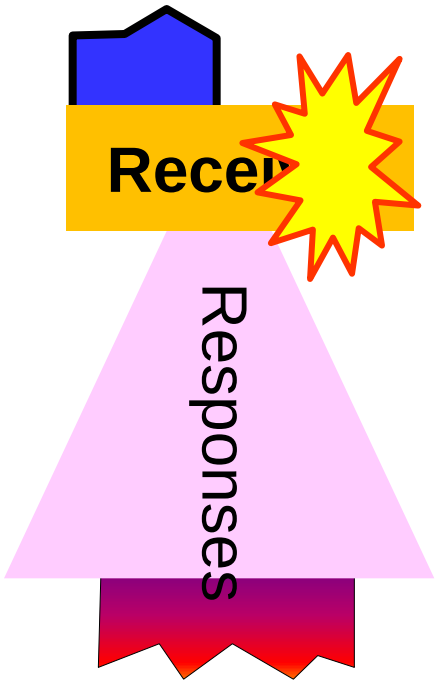
“Name” recognition
= molecular recognition
= localized functionally

Transcription factors
do “name” to “address”
translation

DNA



“Name” recognition
= molecular recognition
= localized functionally

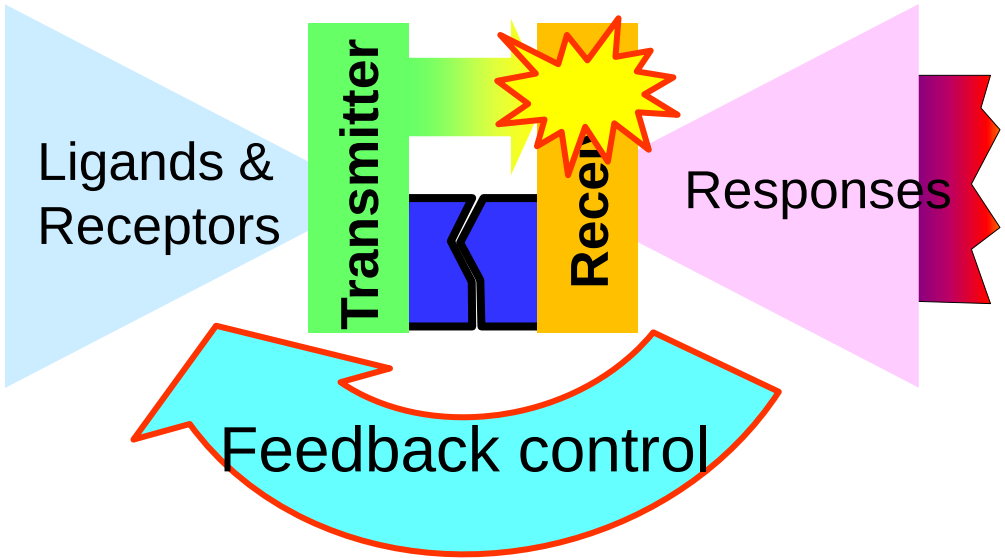


Transcription factors
do “name” to “address”
translation

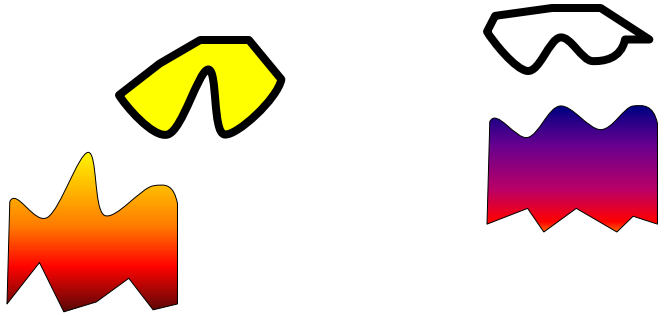
“Addressing”
= molecular recognition
= localized spatially

Both are
• Almost digital
• Highly programmable

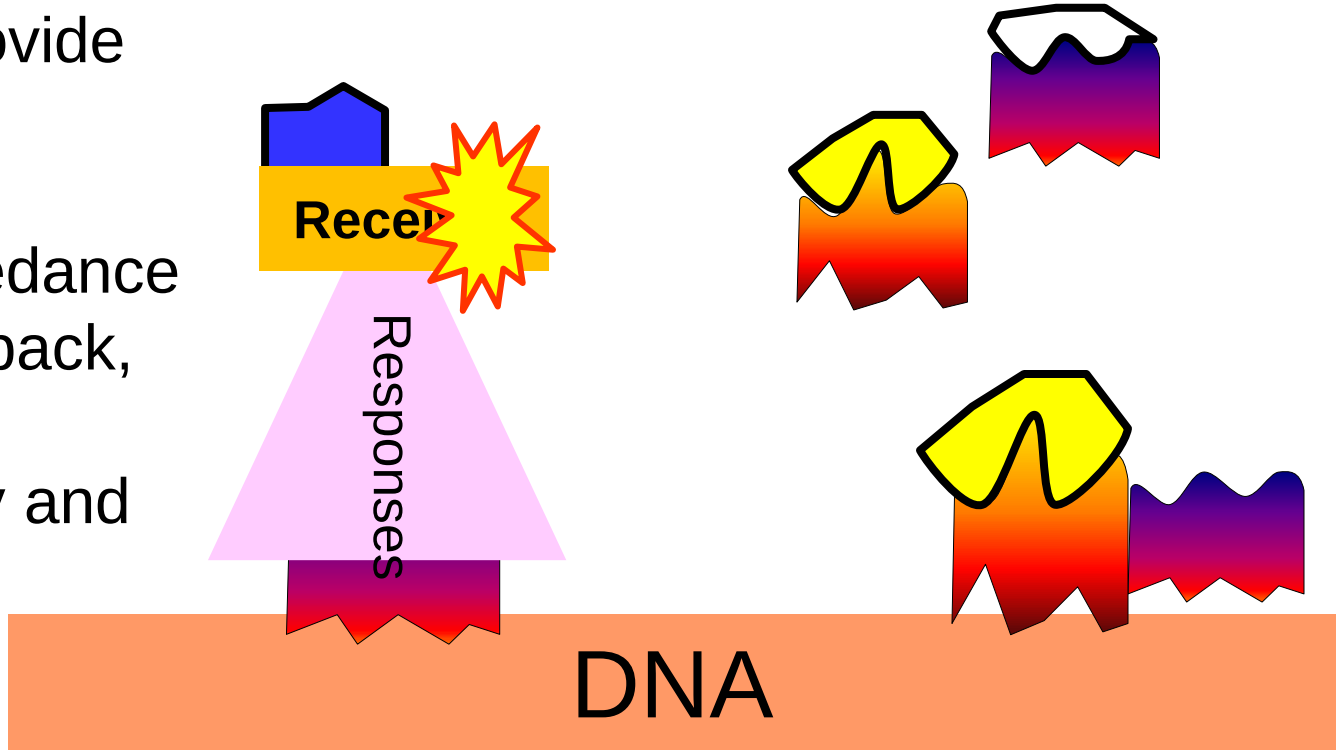
DNA



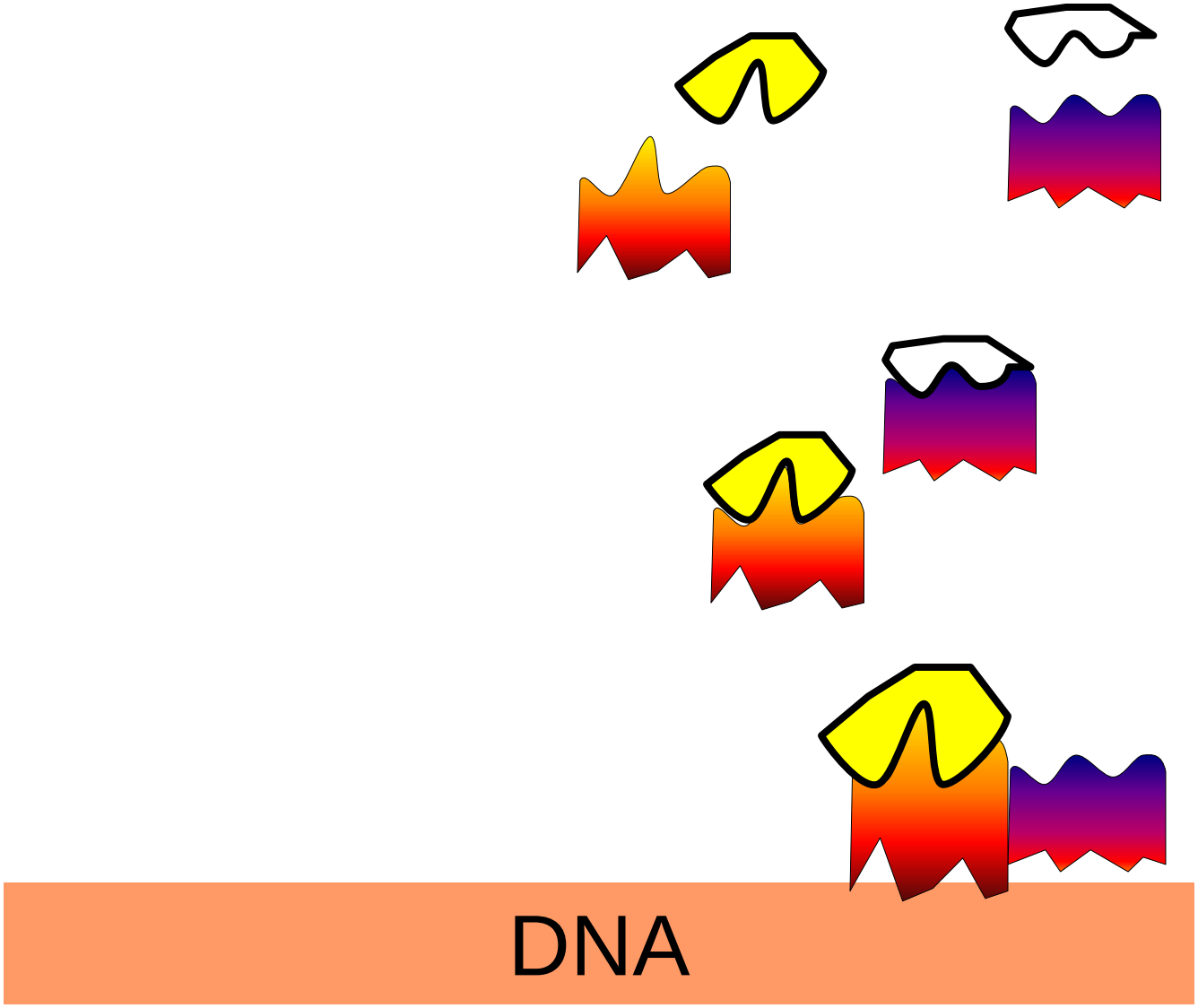
There are simpler transcription factors for sensing internal states



2CST systems provide speed, flexibility, external sensing, computation, impedance match, more feedback, but
greater complexity and overhead

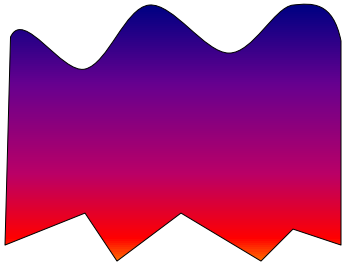


There are simpler transcription factors for sensing internal states



There are simpler transcription factors for sensing internal states

Sensor domains

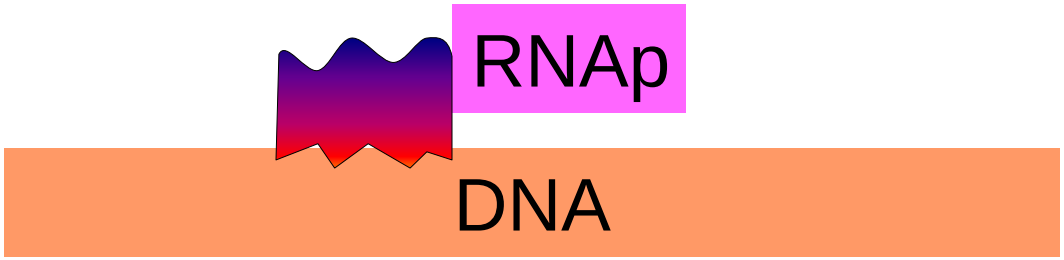


DNA and RNAp binding domains

Domains can be evolved independently or coordinated.

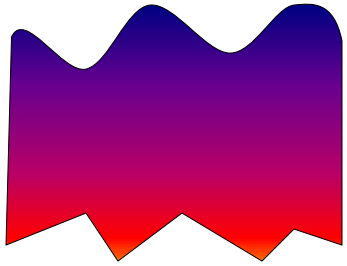
Application layer cannot access DNA directly.

Highly evolvable architecture.



This is like a “name to address” translation.

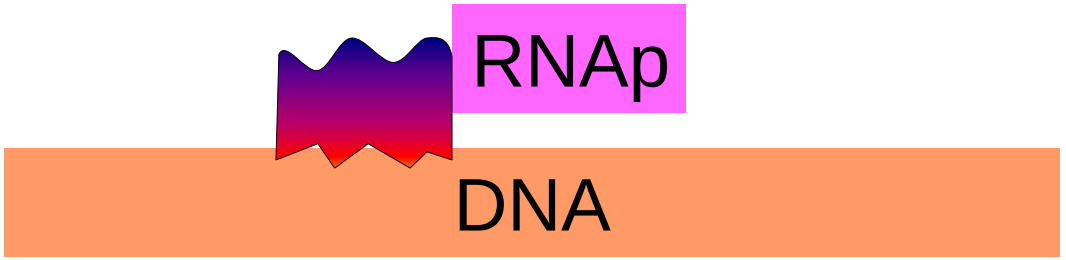
Sensor domains

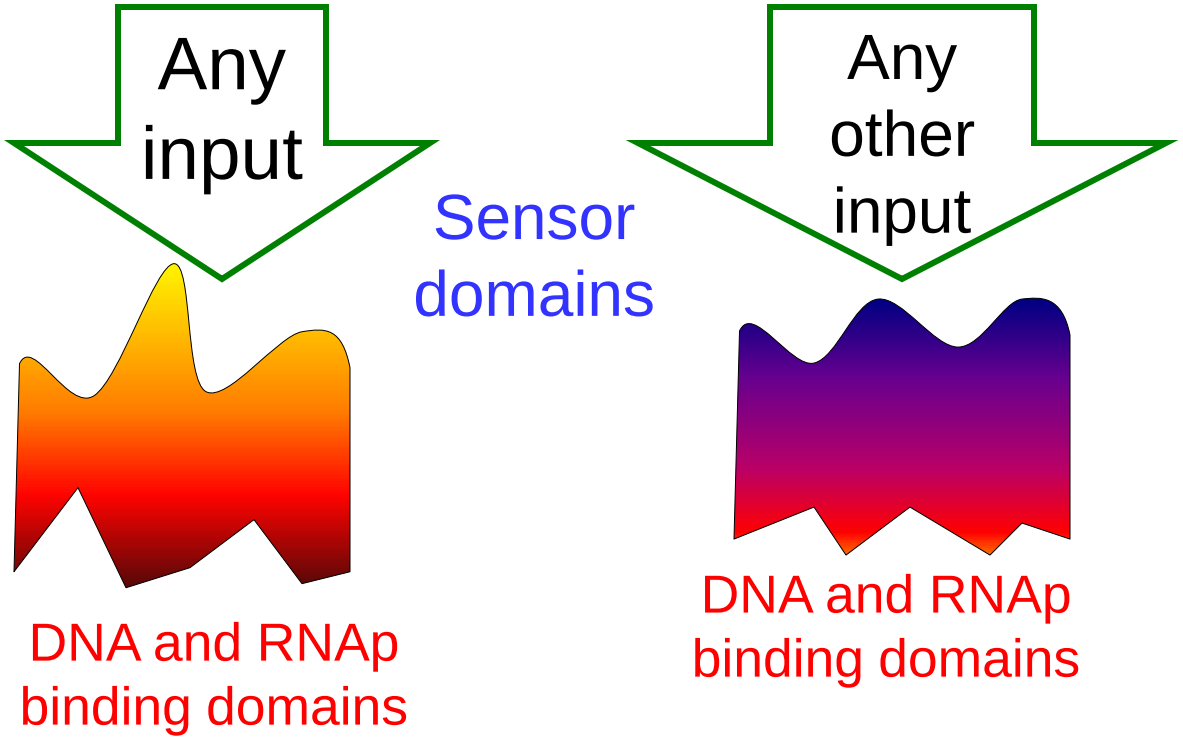


DNA and RNAp binding domains

Sensing the demand of the application layer

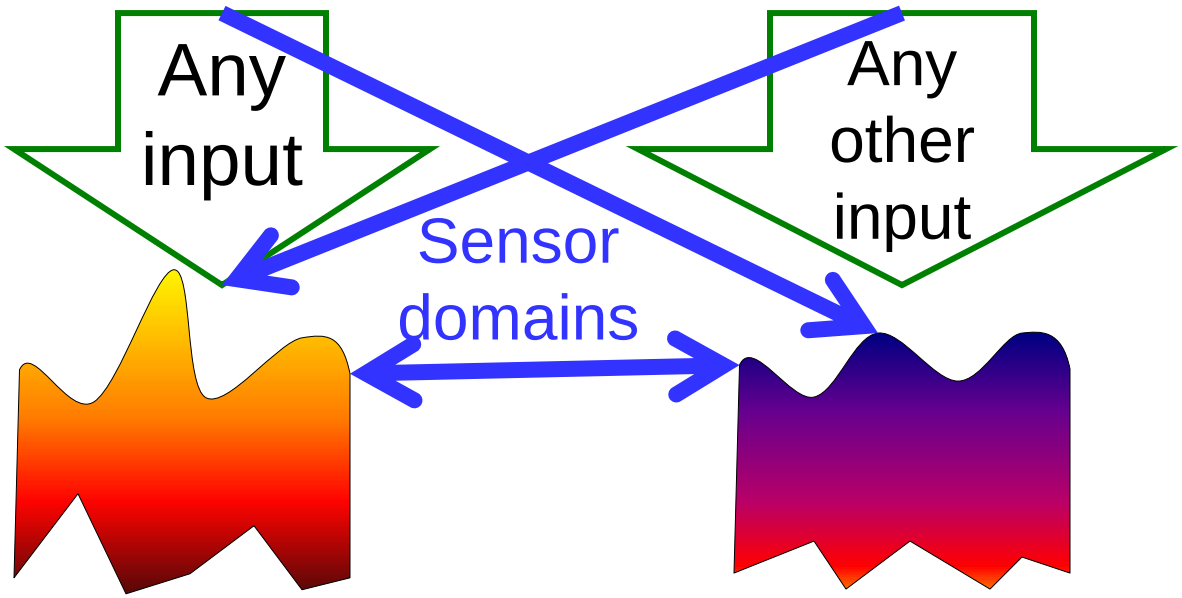
Initiating the change in supply





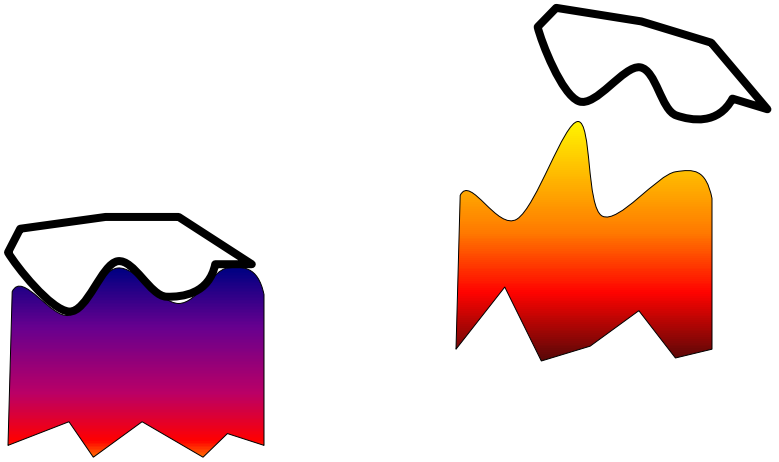
- Sensor sides attach to metabolites or other proteins
- This causes an allosteric (shape) change
- (Sensing is largely analog (# of bound proteins))
- Effecting the DNA/RNAp binding domains
- Protein and DNA/RNAp recognition is more digital
- Extensively discussed in both Ptashne and Alon

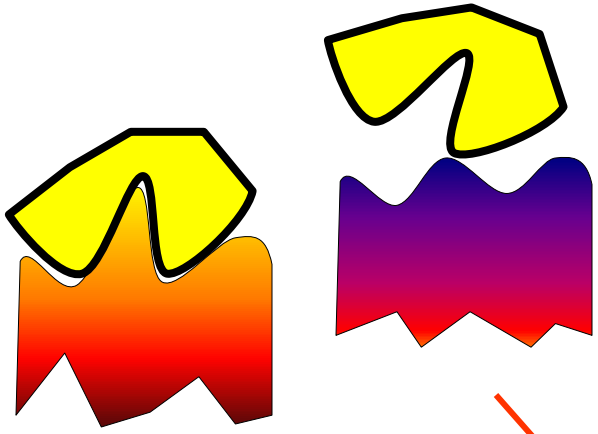
“Cross talk” can be finely controlled



- Application layer signals can be integrated or not
- Huge combinatorial space of (mis)matching shapes
- A functionally meaningful “name space”
- Highly adaptable architecture
- Interactions are fast (but expensive)
- Return to this issue in “signal transduction”

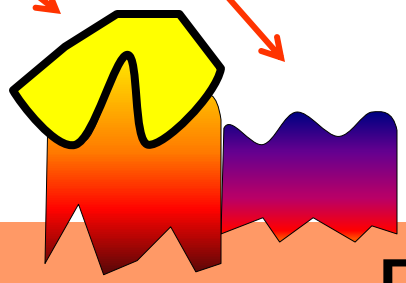
“Name” recognition
= molecular recognition
= localized functionally
= global spatially





Transcription factors
do “name” to “address”
translation

Both are
• Almost digital
• Highly programmable

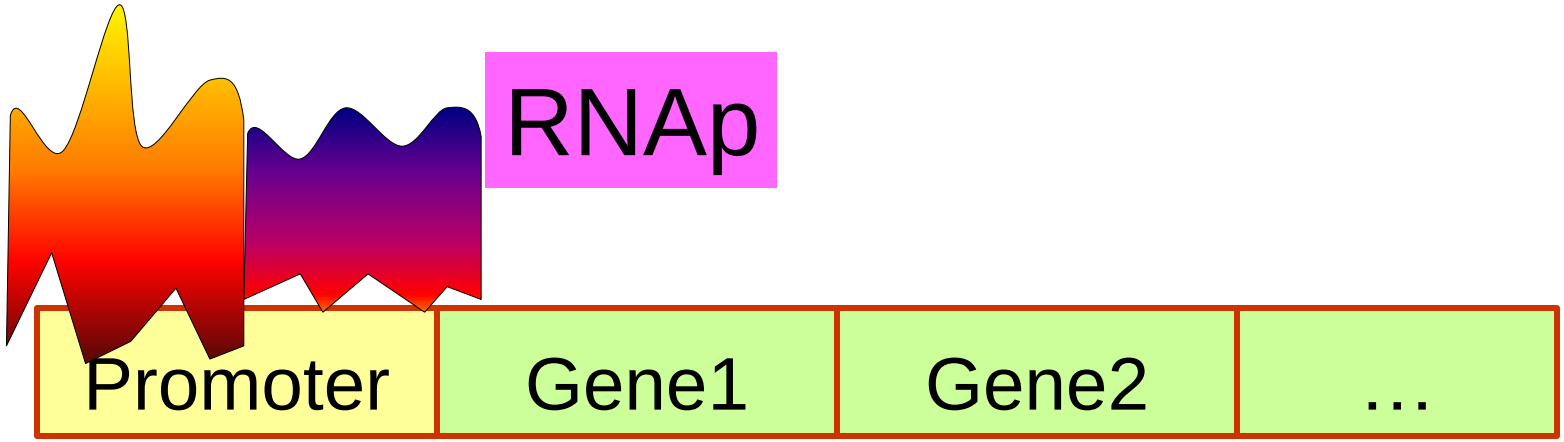


“Addressing”
= molecular recognition
= localized spatially

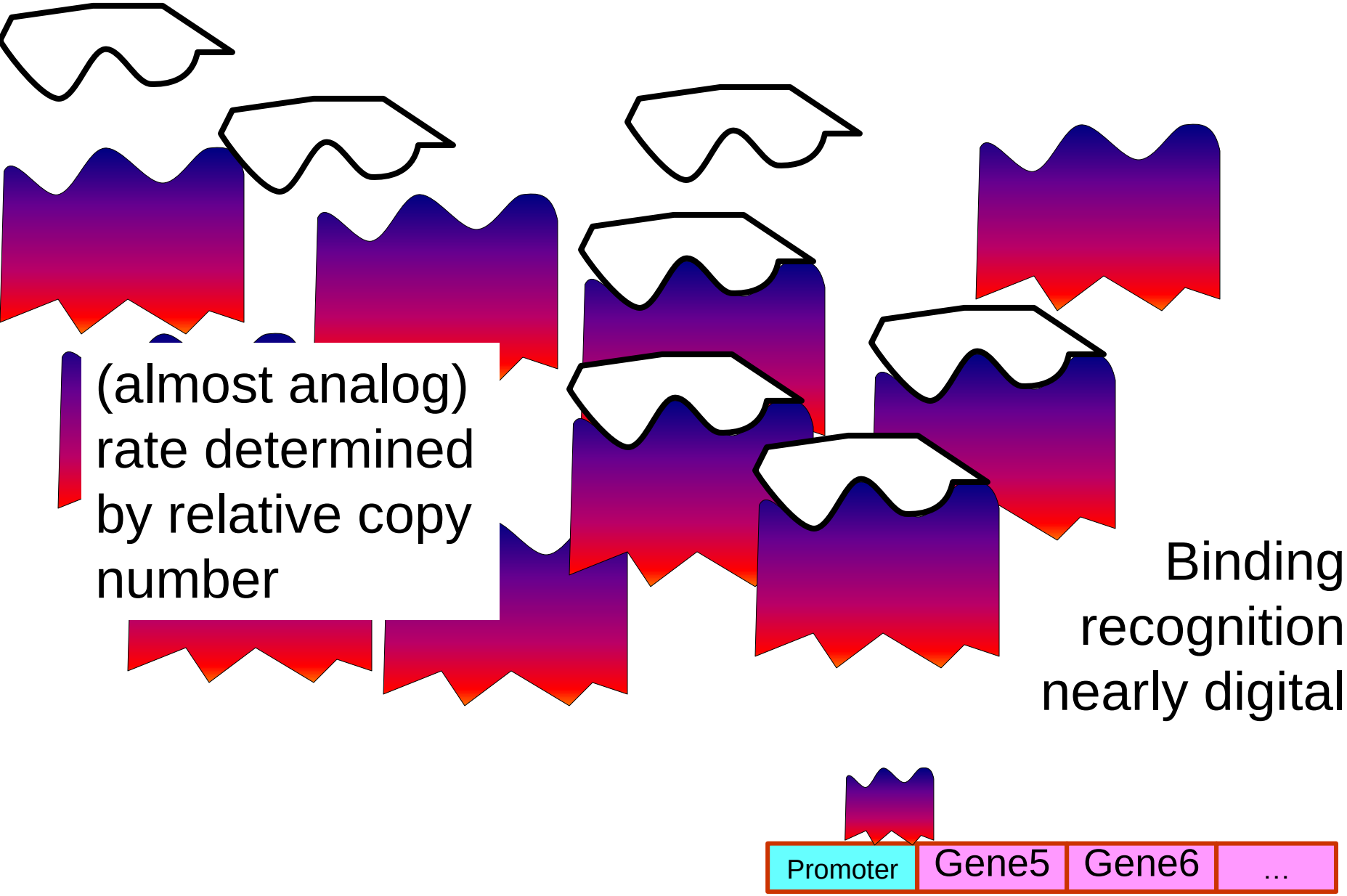
DNA

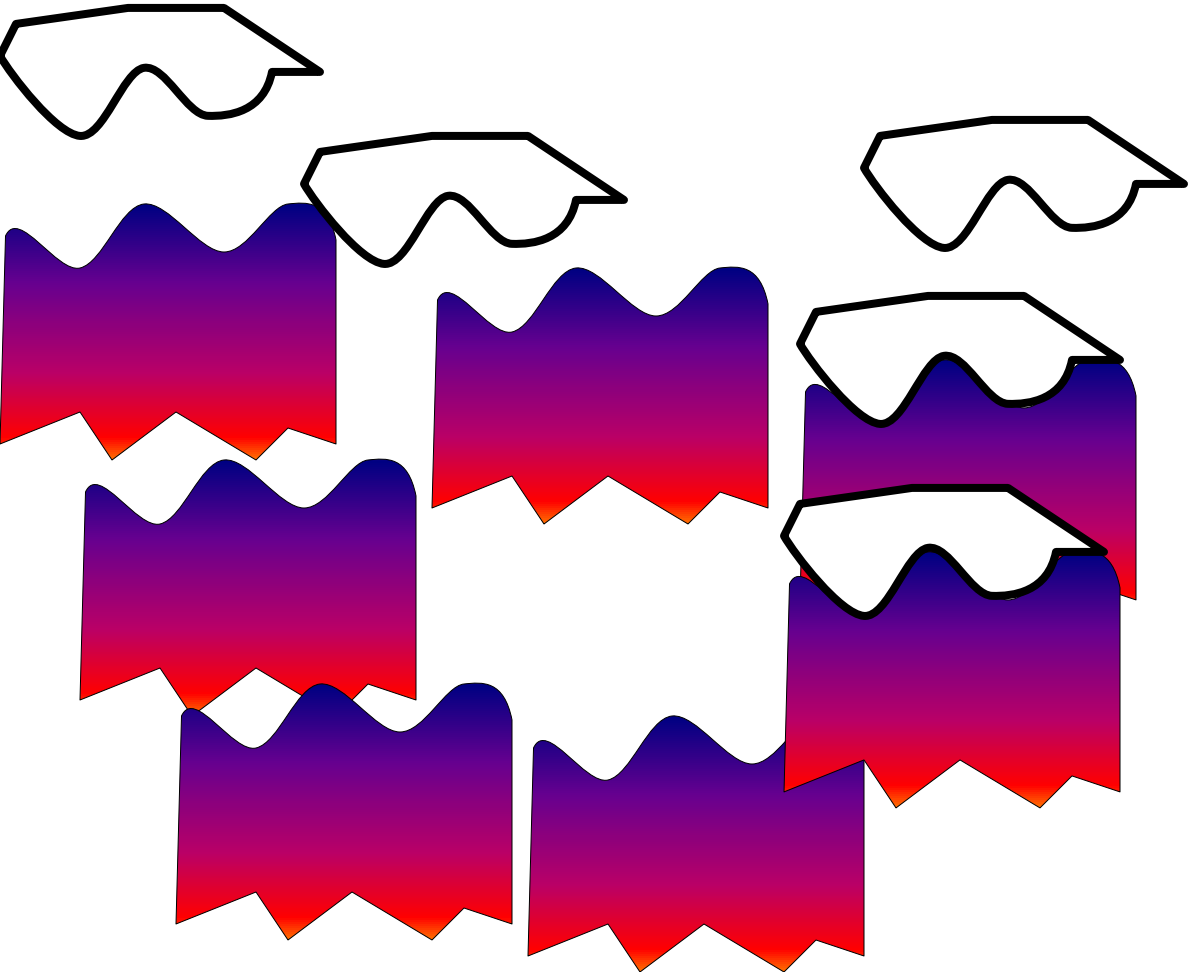
Can activate or repress

And work in complex logical combinations



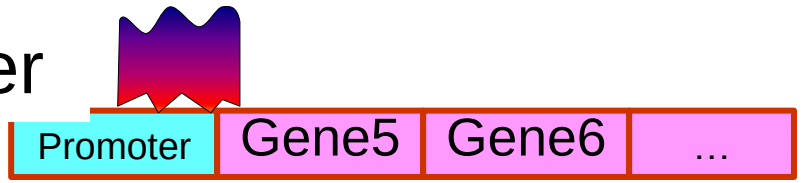
- Both protein and DNA sides have sequence/shape
- Huge combinatorial space of “addresses”
- Modest amount of “logic” can be done at promoter
- Transcription is very noise (but efficient)
- Extremely adaptable architecture

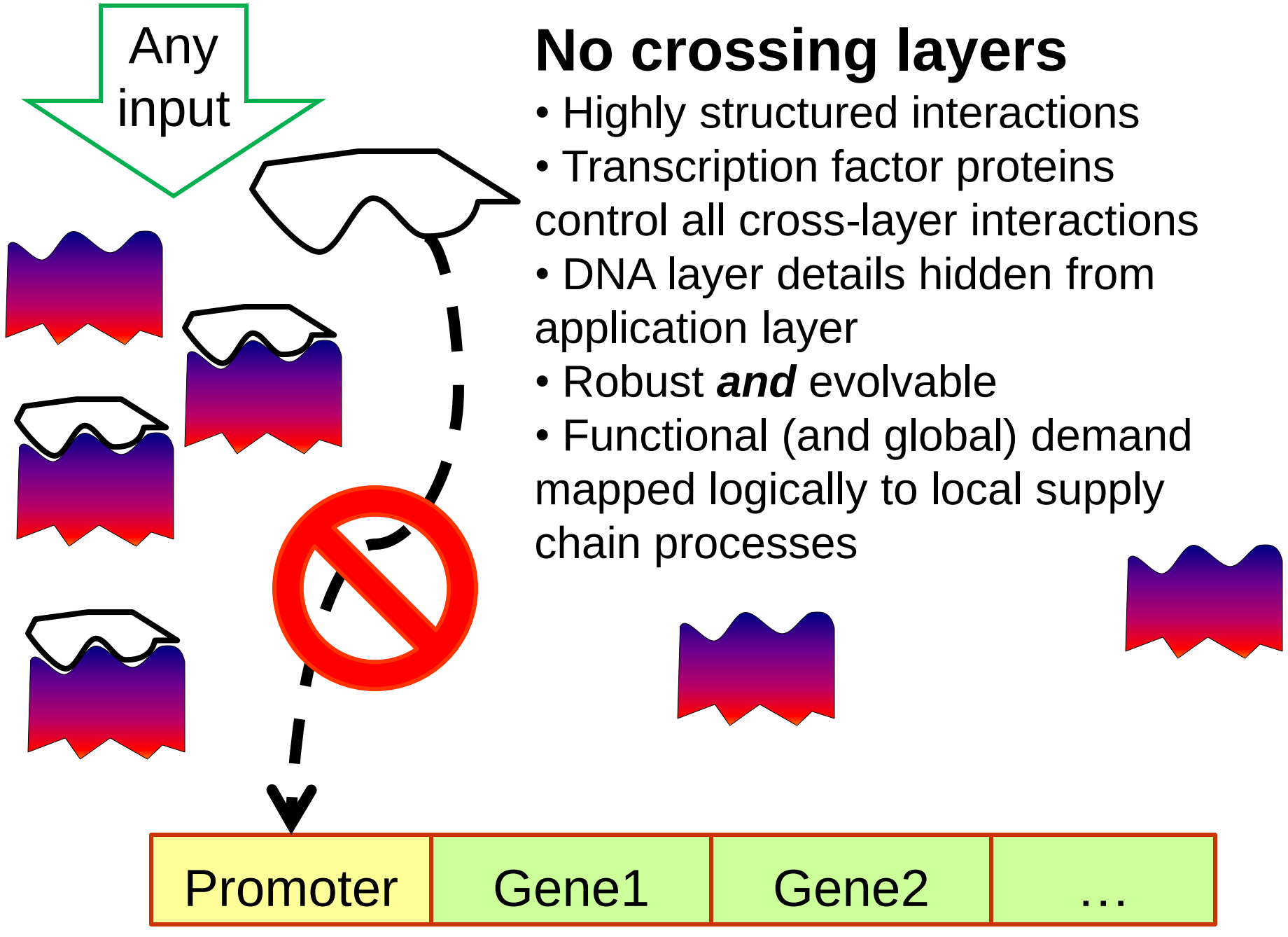

(almost analog)
rate determined
by relative copy
number
Binding
recognition
nearly digital
Promoter
Gene5
Gene6
…



rate (almost analog)
determined by copy number

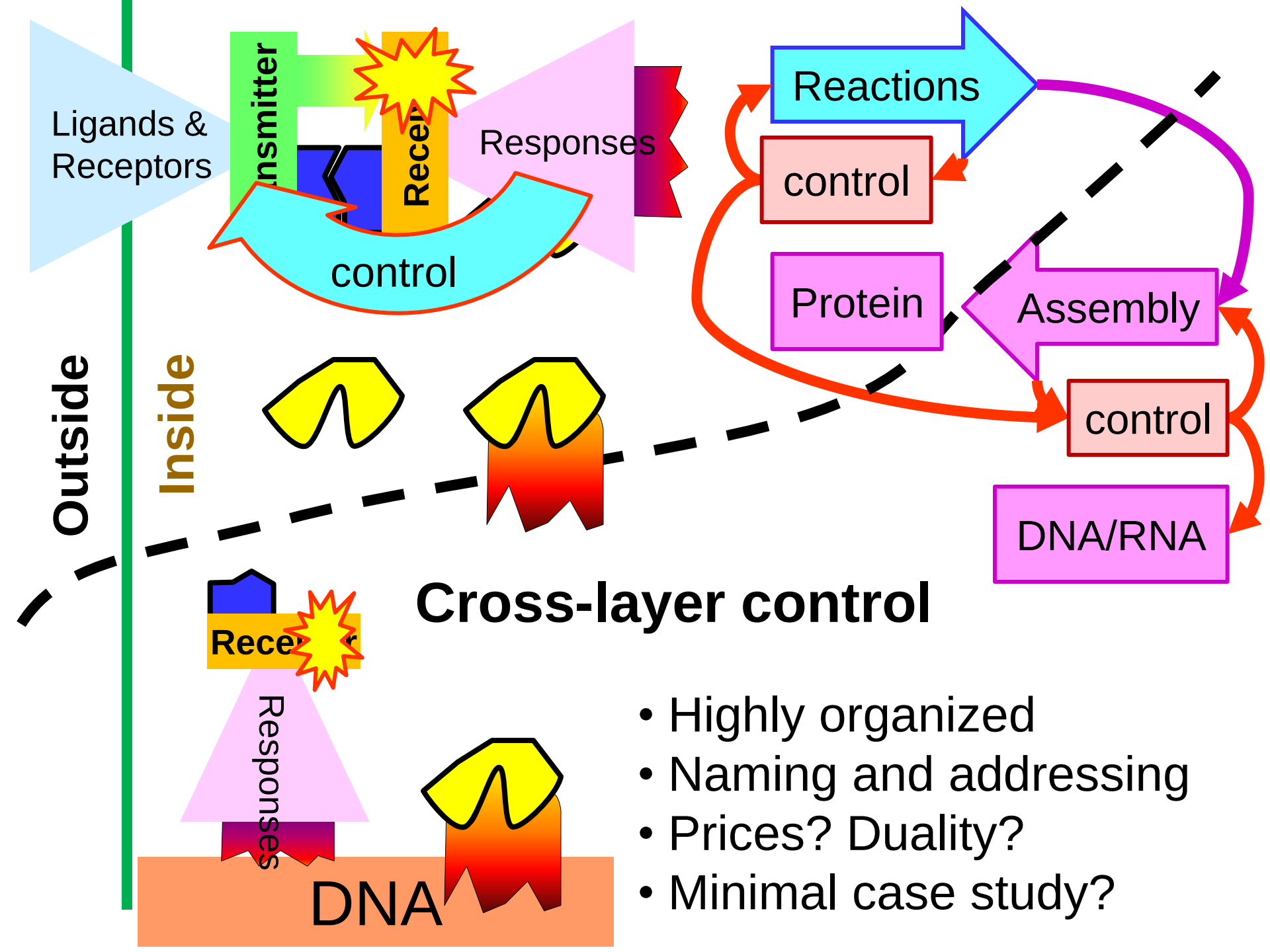
Recall: can work by pulse code modulation so for small copy number does digital to analog conversion

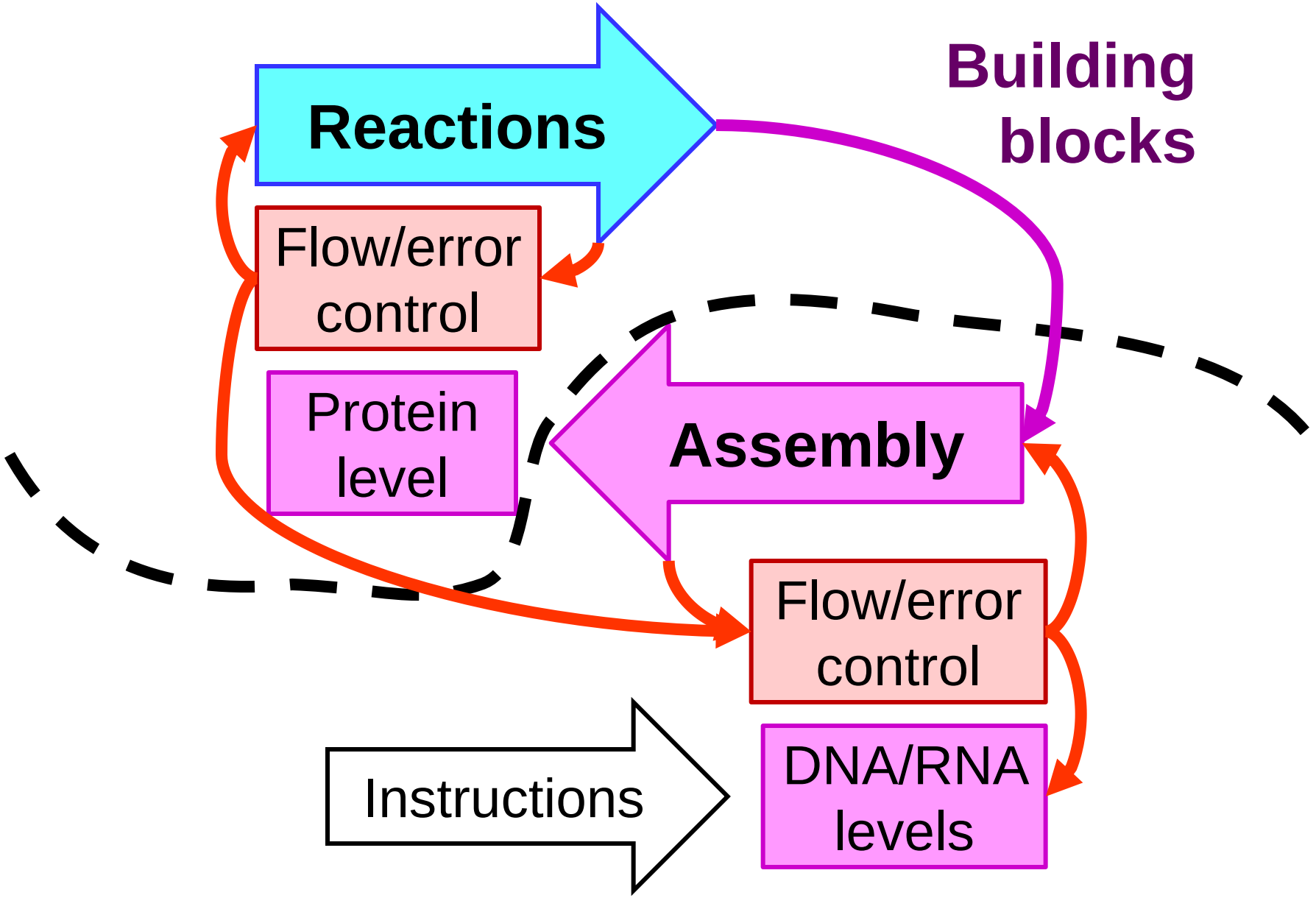


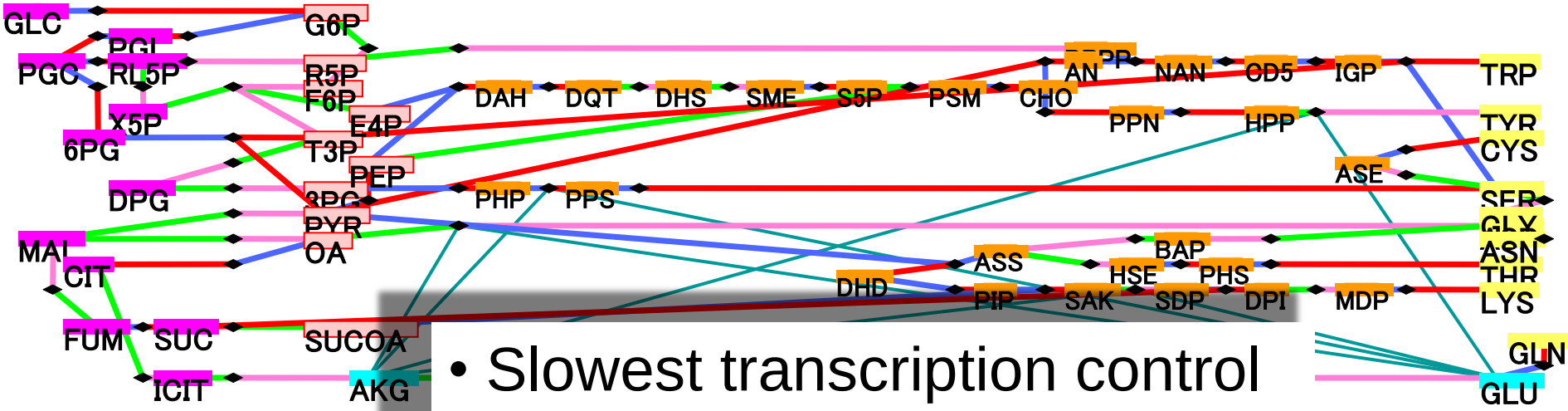


No crossing layers

- Highly structured interactions
- Transcription factor proteins control all cross-layer interactions
- DNA layer details hidden from application layer
- Robust ***and*** evolvable
- Functional (and global) demand mapped logically to local supply chain processes


Ligands &
Receptors
Responses
control
Reactions
control
Protein
Assembly
control
DNA/RNA
Outside
Inside
Responses
DNA
Cross-layer control
• Highly organized
• Naming and addressing
• Prices? Duality?
• Minimal case study?


Reactions
Building blocks
Flow/error control
Protein level
Assembly
Flow/error control
Instructions
DNA/RNA levels



- Slowest transcription control
- Complex transcription factors
- Lowest metabolic overhead
- Easily reprogrammed

Any input

RNA

Transc.

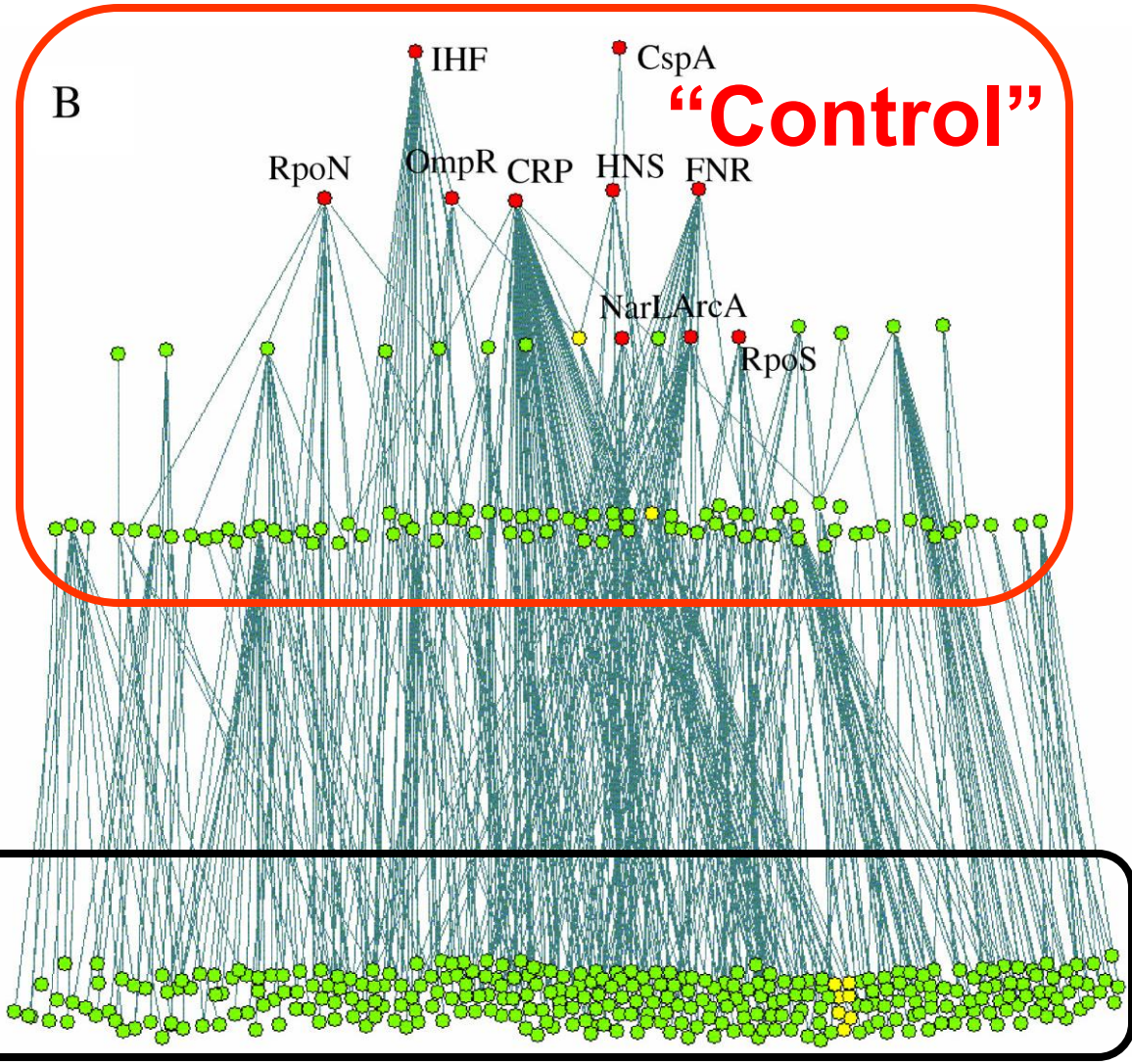
mRNA

RNAp

Gene

This architecture has limited scalability:

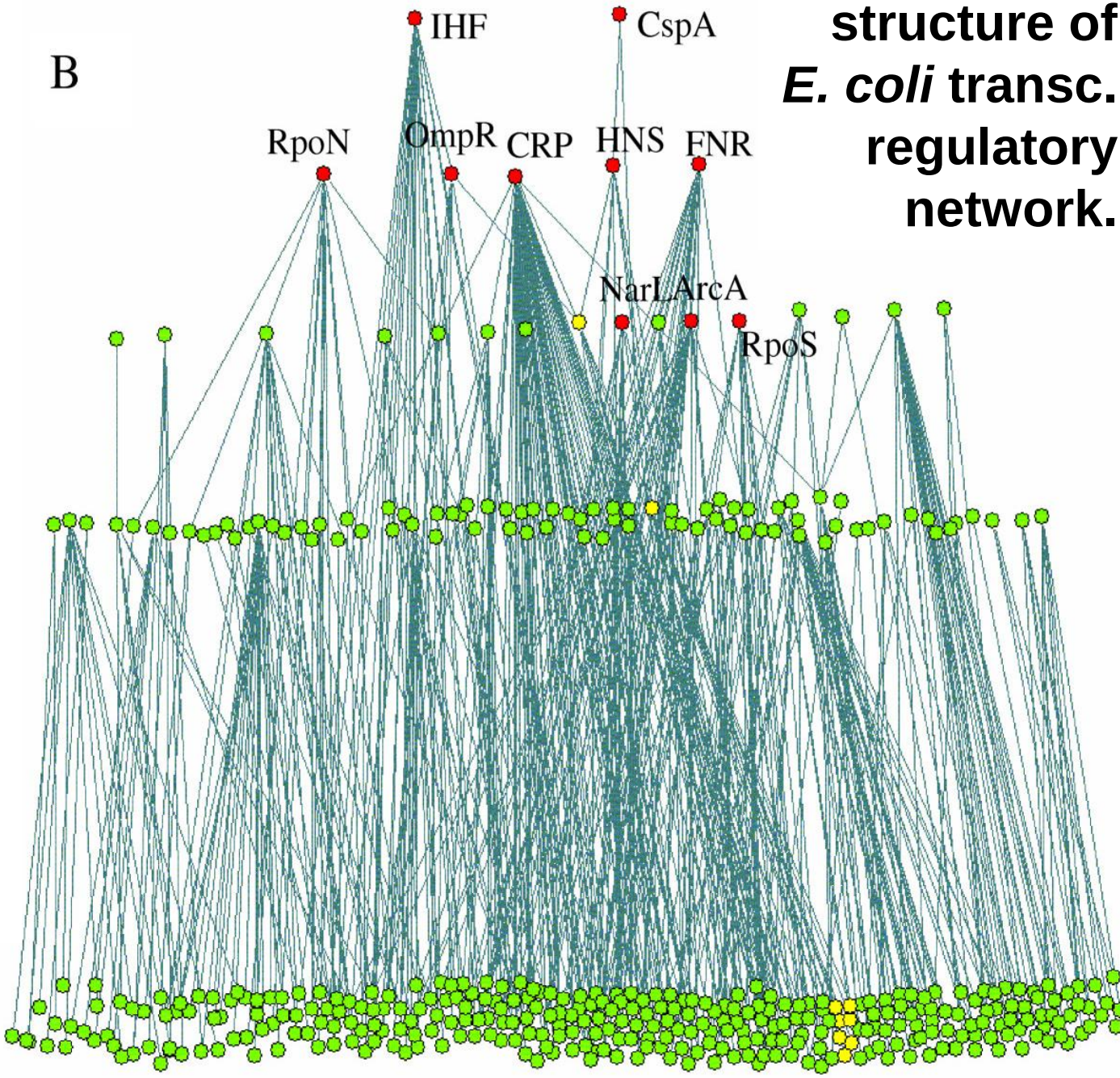
1) Fast diffusion can only work in small volumes
2) The number of proteins required for control grows superlinearly with the number of enzymes (Mattick)



Hierarchical structure of *E. coli* transc. regulatory network.

All transcriptional regulatory links are downward. Nodes are operons. Global regulators are red. Yellow marked nodes are operons in the longest regulatory pathway related with flagella motility.

Ma *et al. BMC Bioinformatics* 2004 **5**:199 doi:10.1186/1471-2105-5-199



heat shock

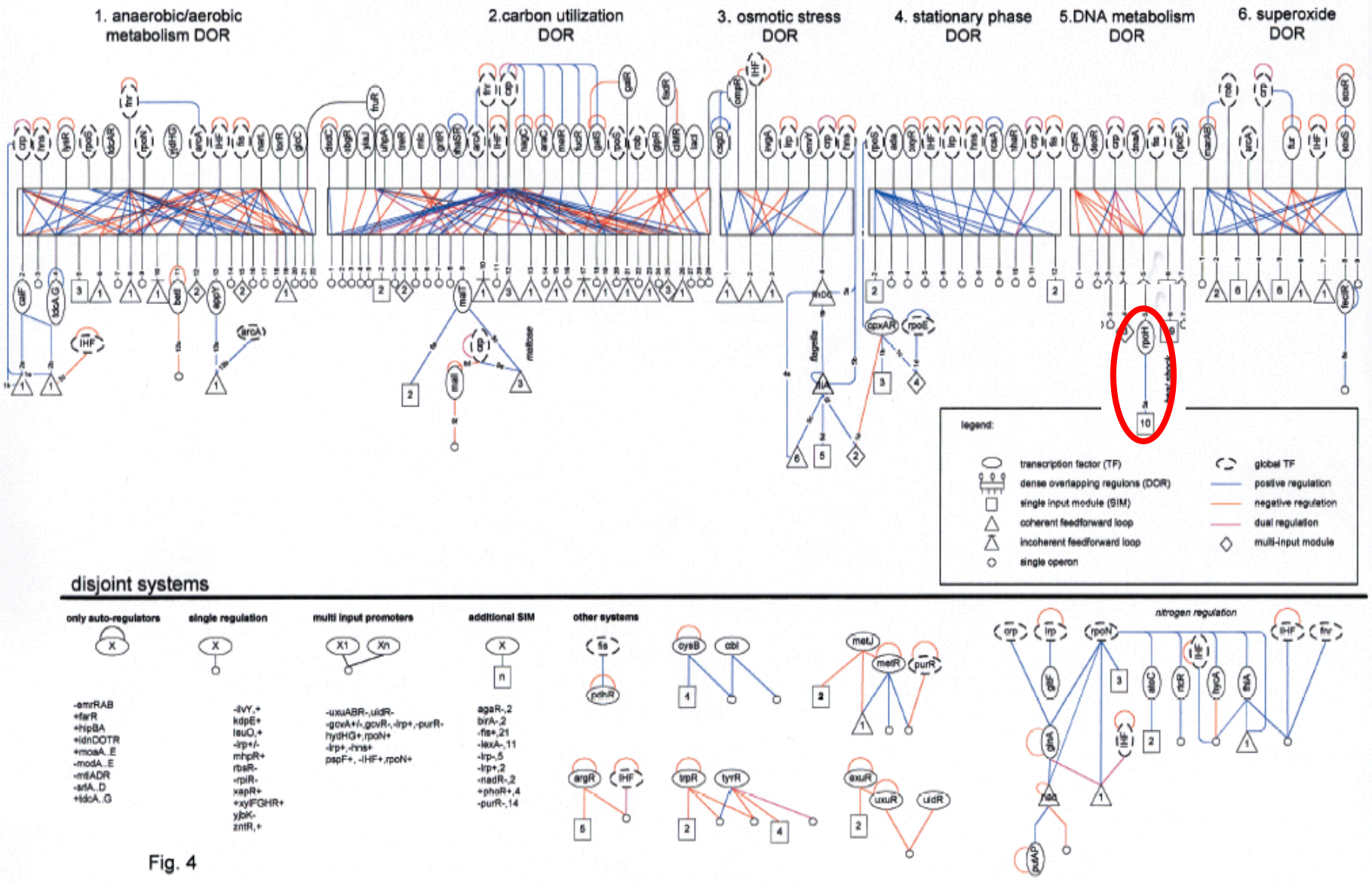
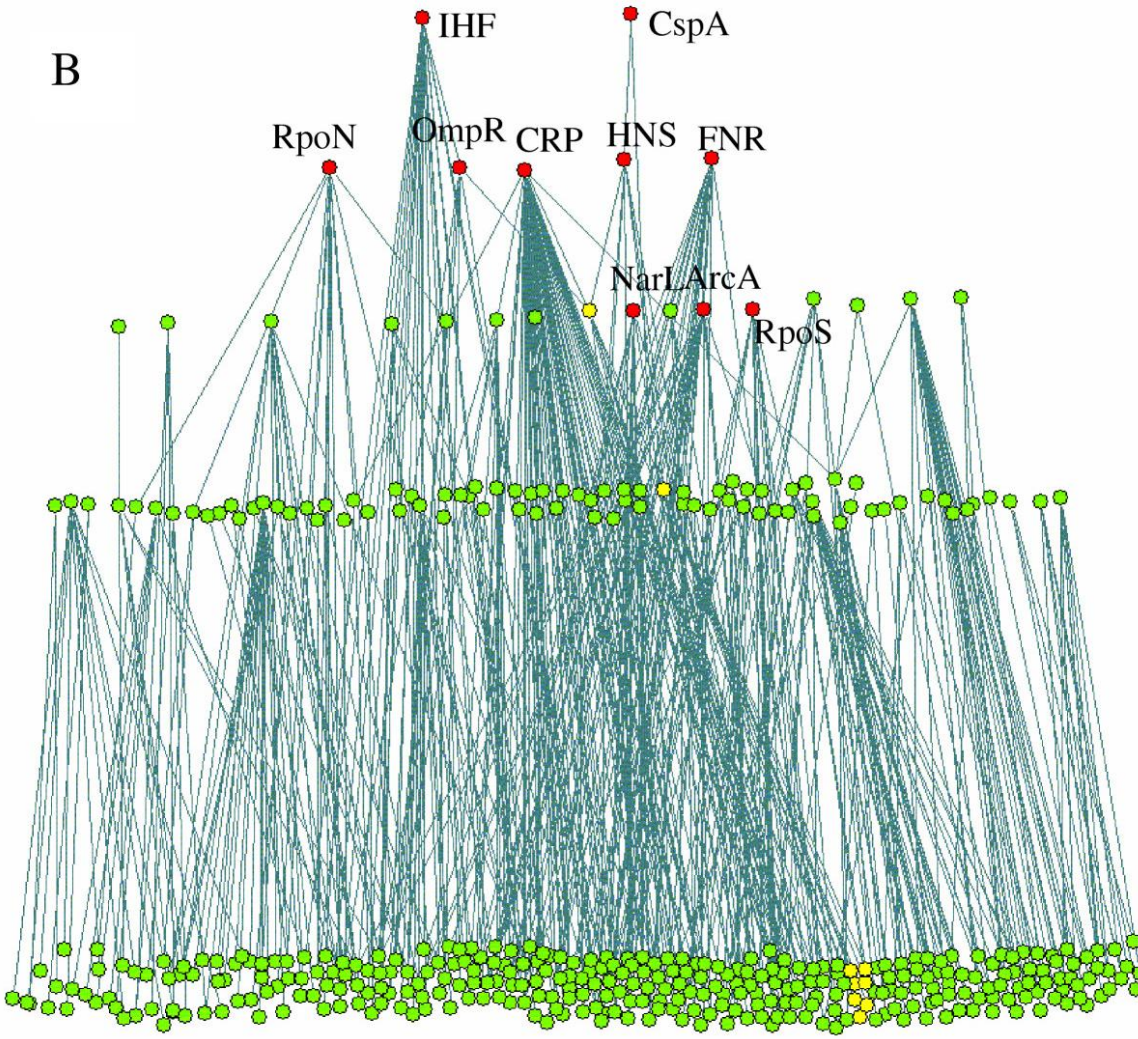


Fig. 4

Note: all feedback in this picture has been removed in two ways:

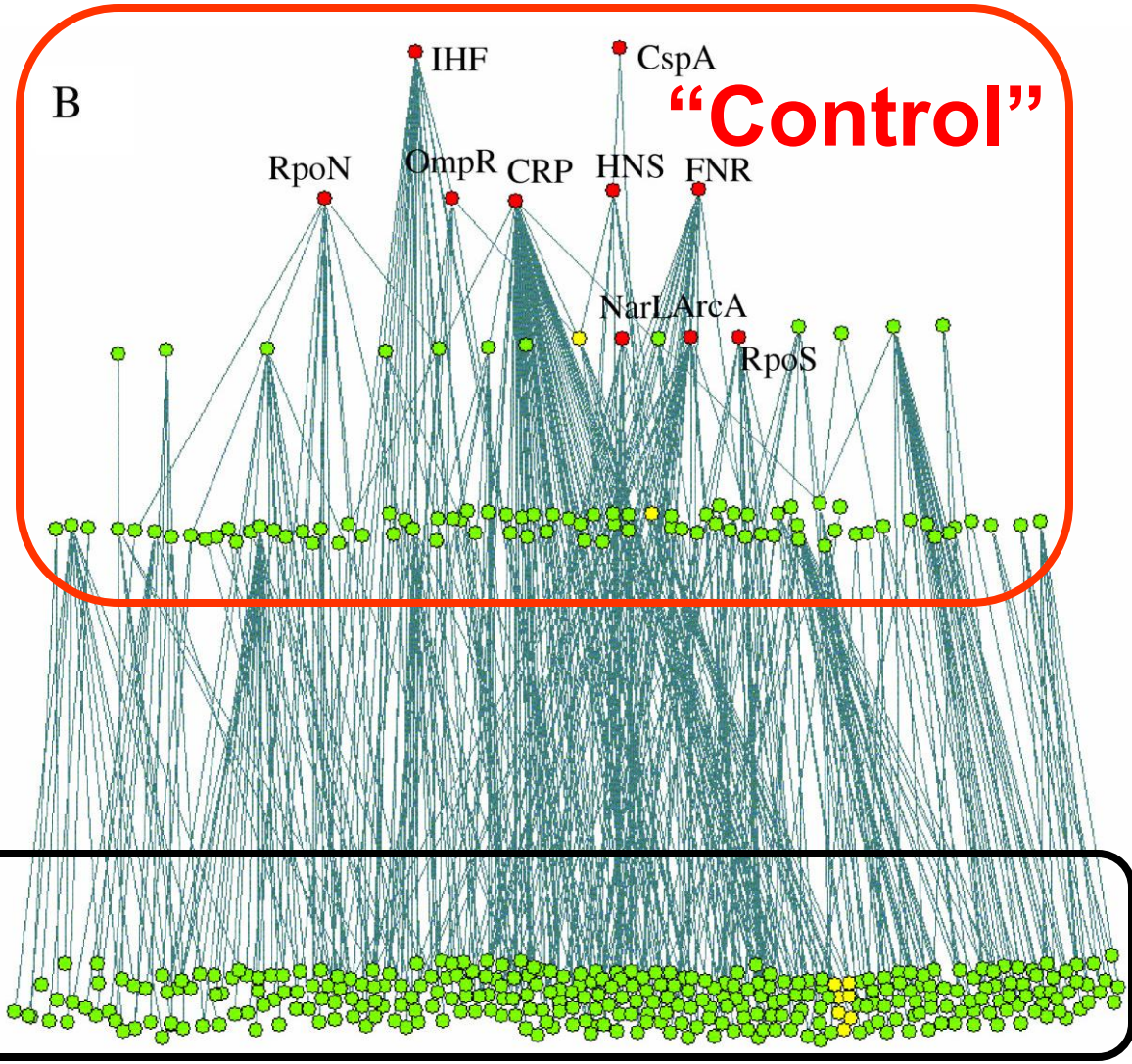
1) There are self-loops where an operon is controlled by one it's own genes
2) All the real complex control is in the protein interactions not shown (e.g. see heat shock details)

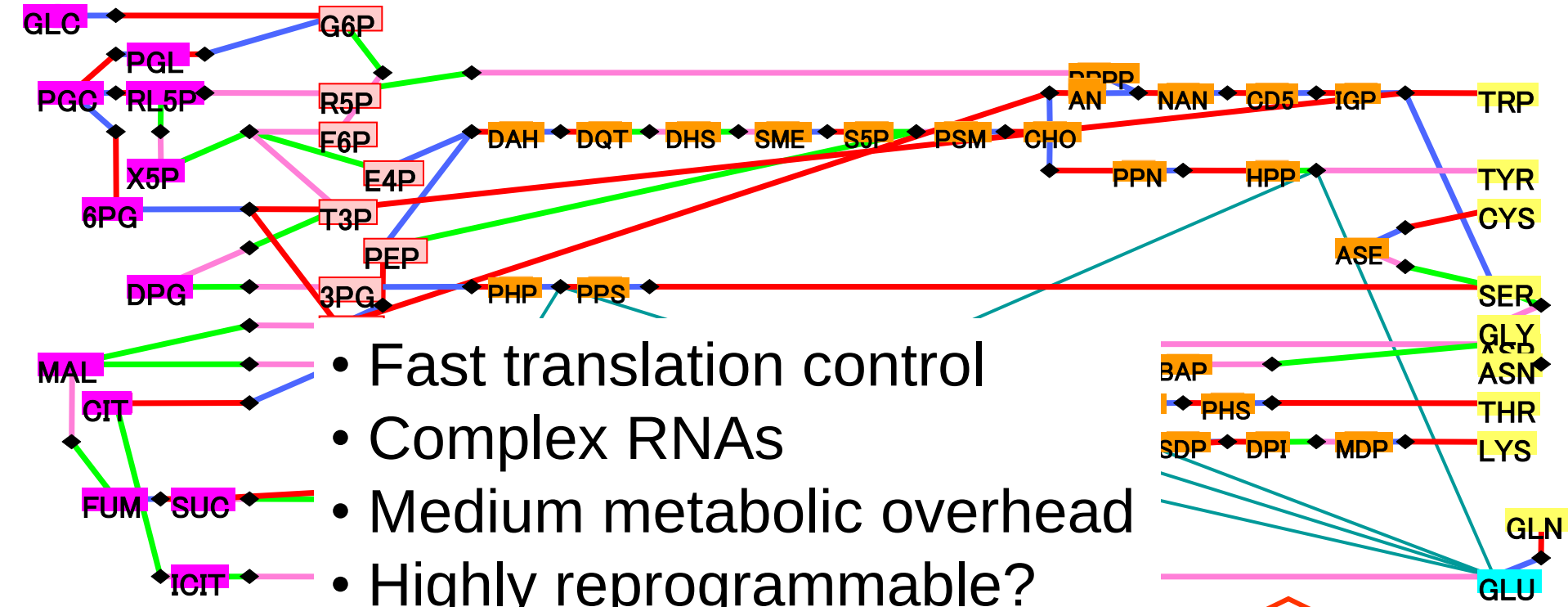
These are not really ***control*** systems, they just initiate manufacturing



This architecture has limited scalability:

1) Fast diffusion can only work in small volumes
2) The number of proteins required for control grows superlinearly with the number of enzymes (Mattick)





- Fast translation control
- Complex RNAs
- Medium metabolic overhead
- Highly reprogrammable?

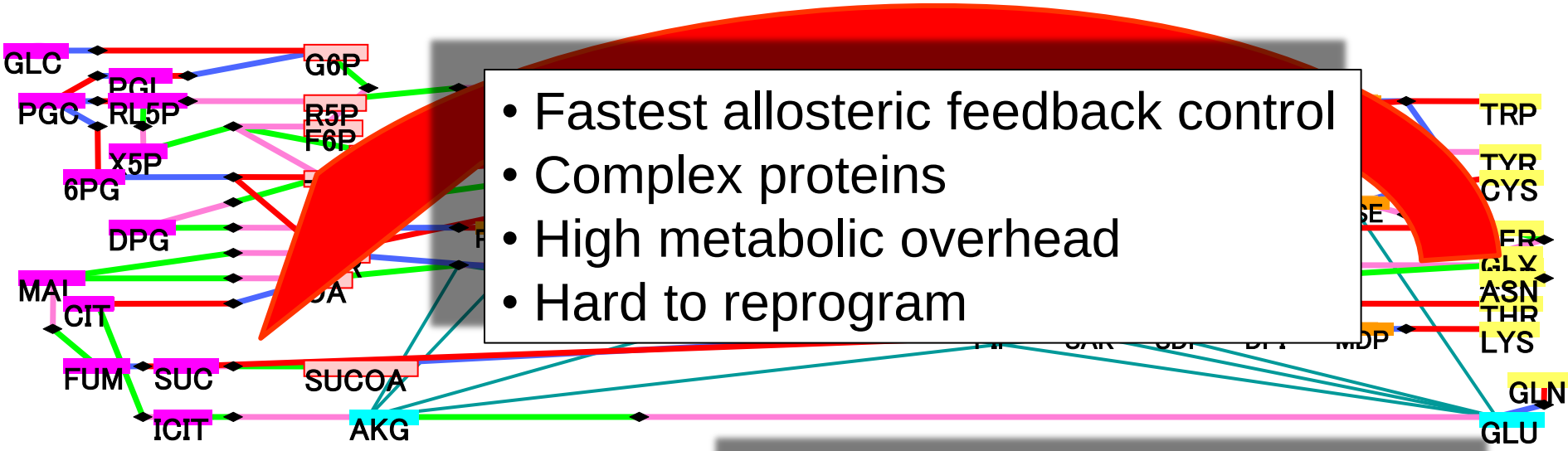
Any mRNA

Any input

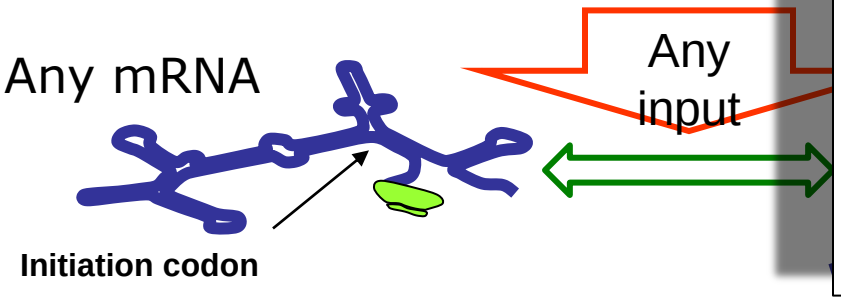
Any protein

Translation

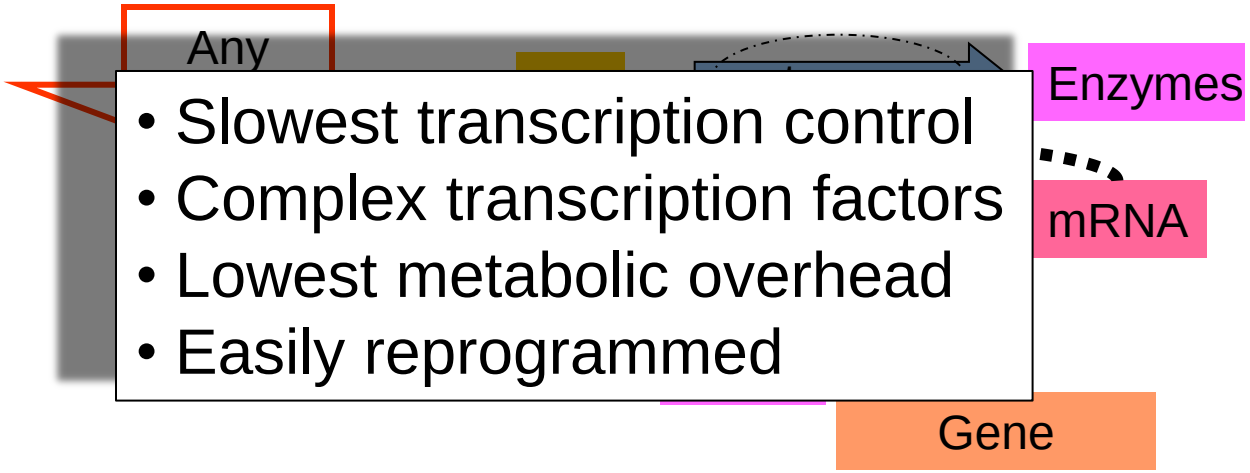
Initiation codon



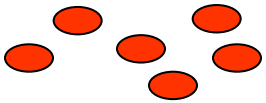
- Fastest allosteric feedback control
- Complex proteins
- High metabolic overhead
- Hard to reprogram

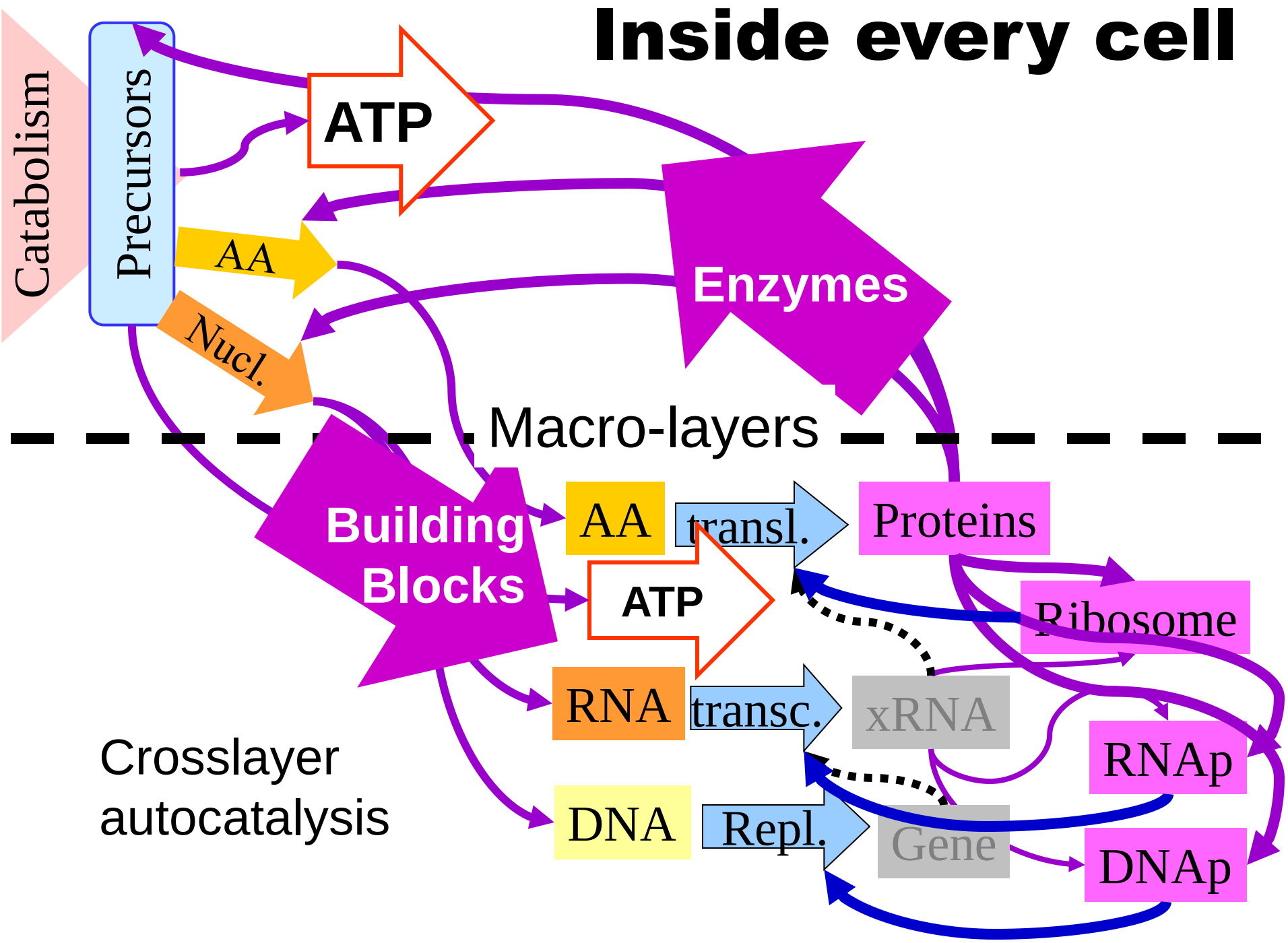


- Fast translation control
- Complex RNAs
- Medium metabolic overhead
- Highly reprogrammable?



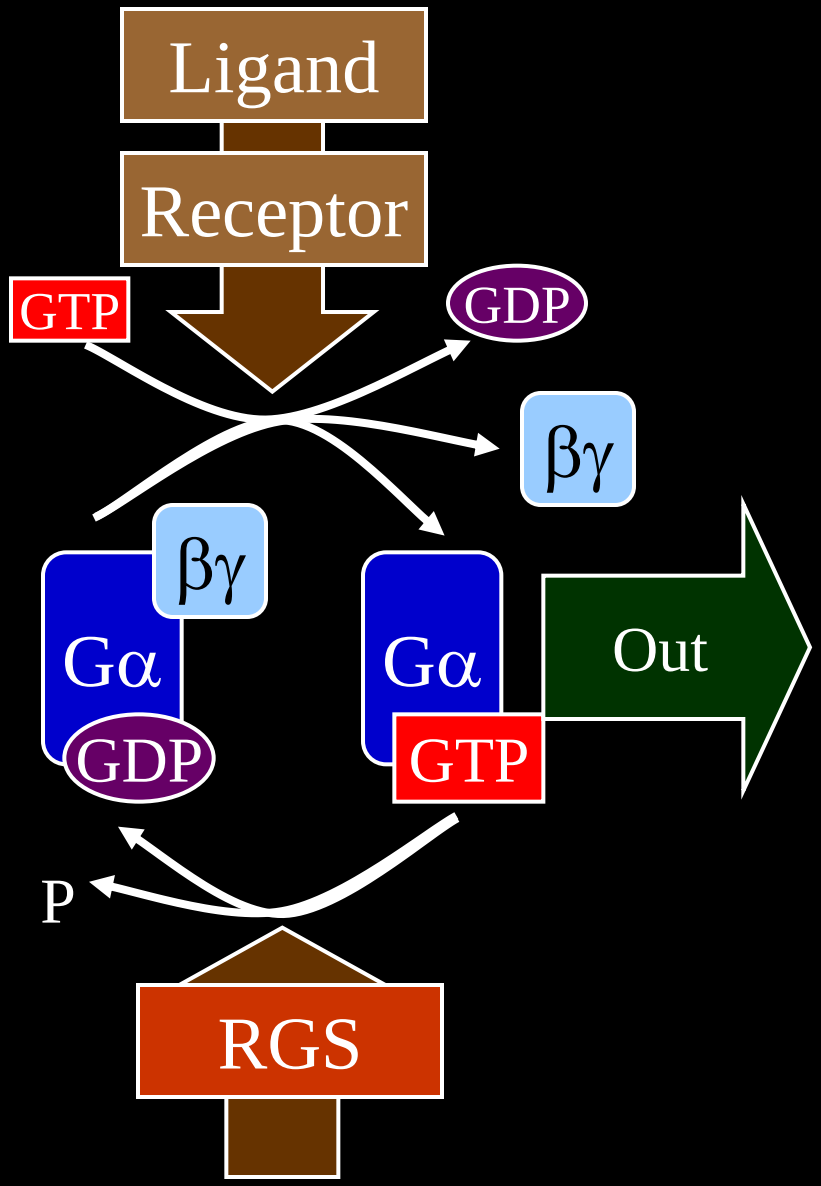
- Slowest transcription control
- Complex transcription factors
- Lowest metabolic overhead
- Easily reprogrammed

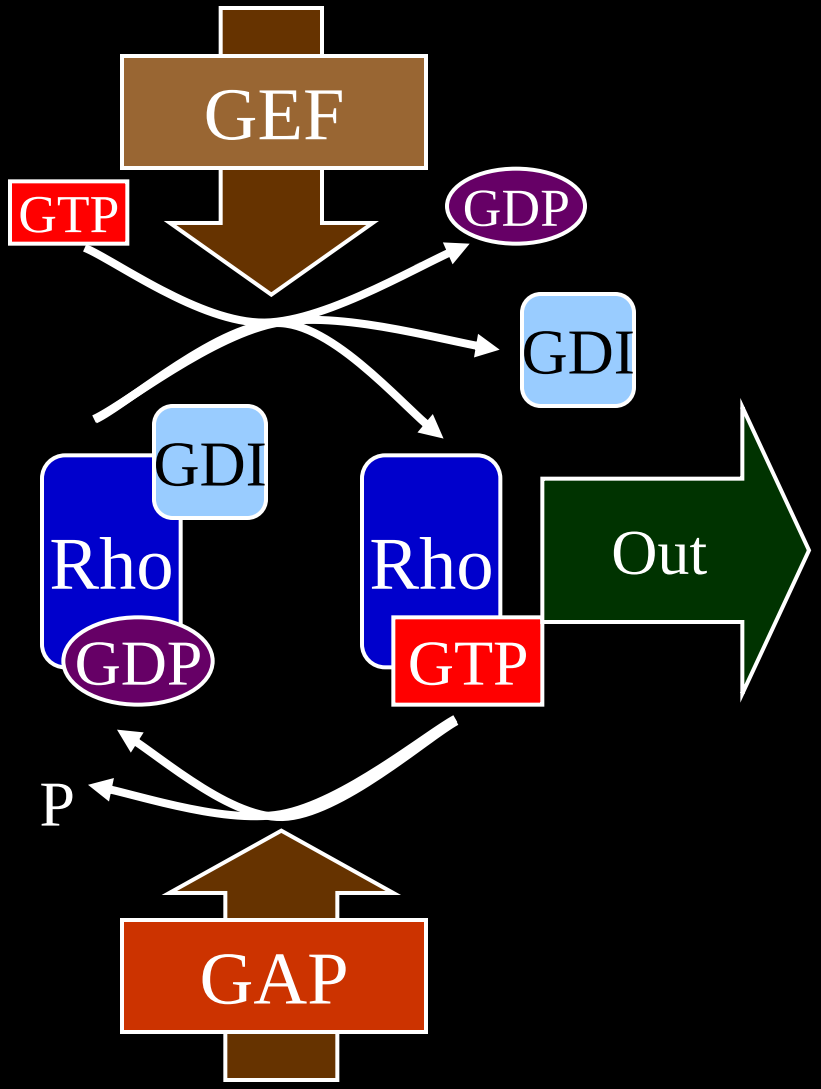


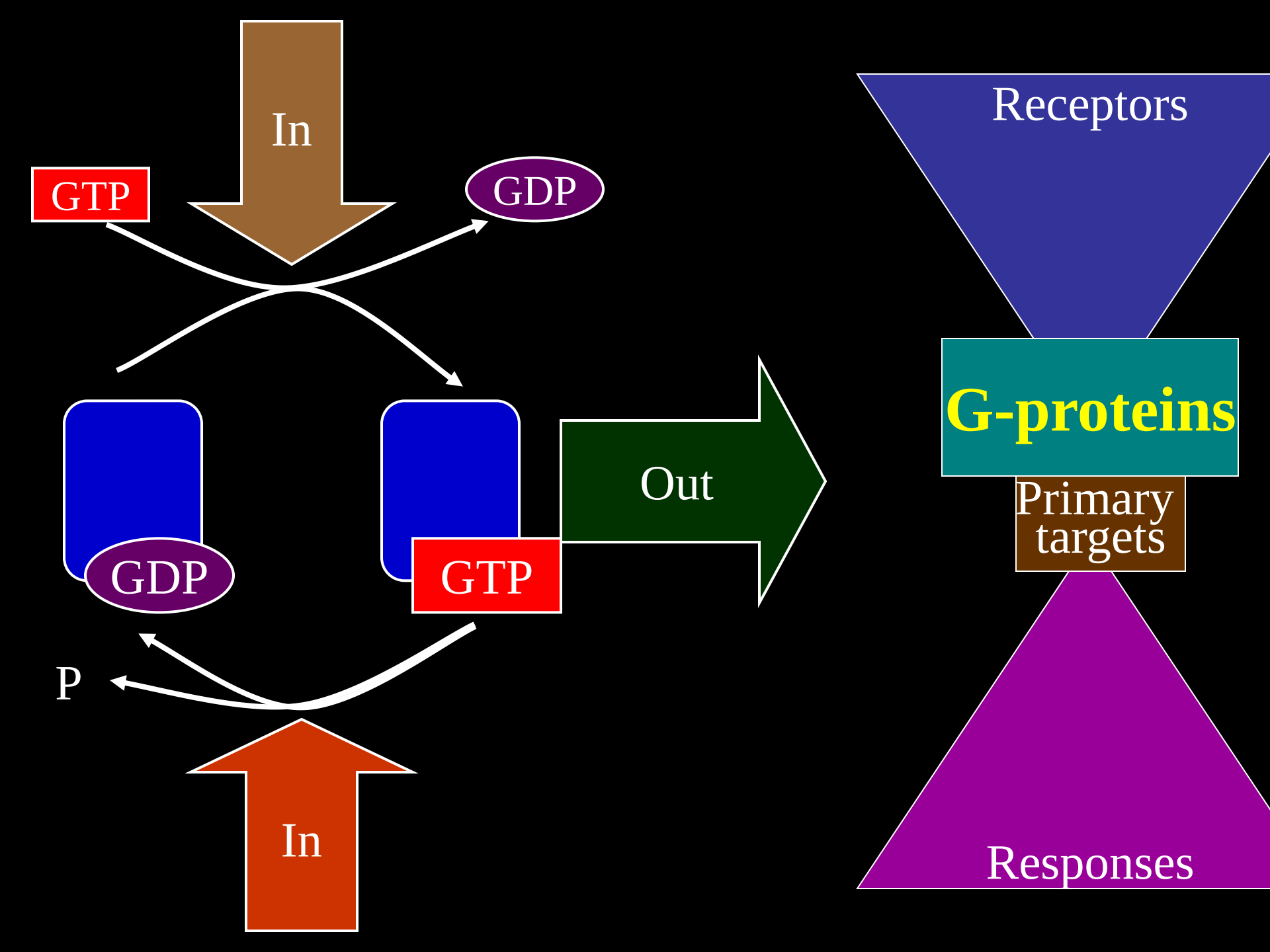

Inside every cell
Catabolism
Precursors
ATP
AA
Nucl.
Enzymes
Macro-layers
Building Blocks
AA
transl.
Proteins
ATP
Ribosome
RNA
transc.
xRNA
RNAp
DNA
Repl.
Gene
DNAp
Crosslayer autocatalysis

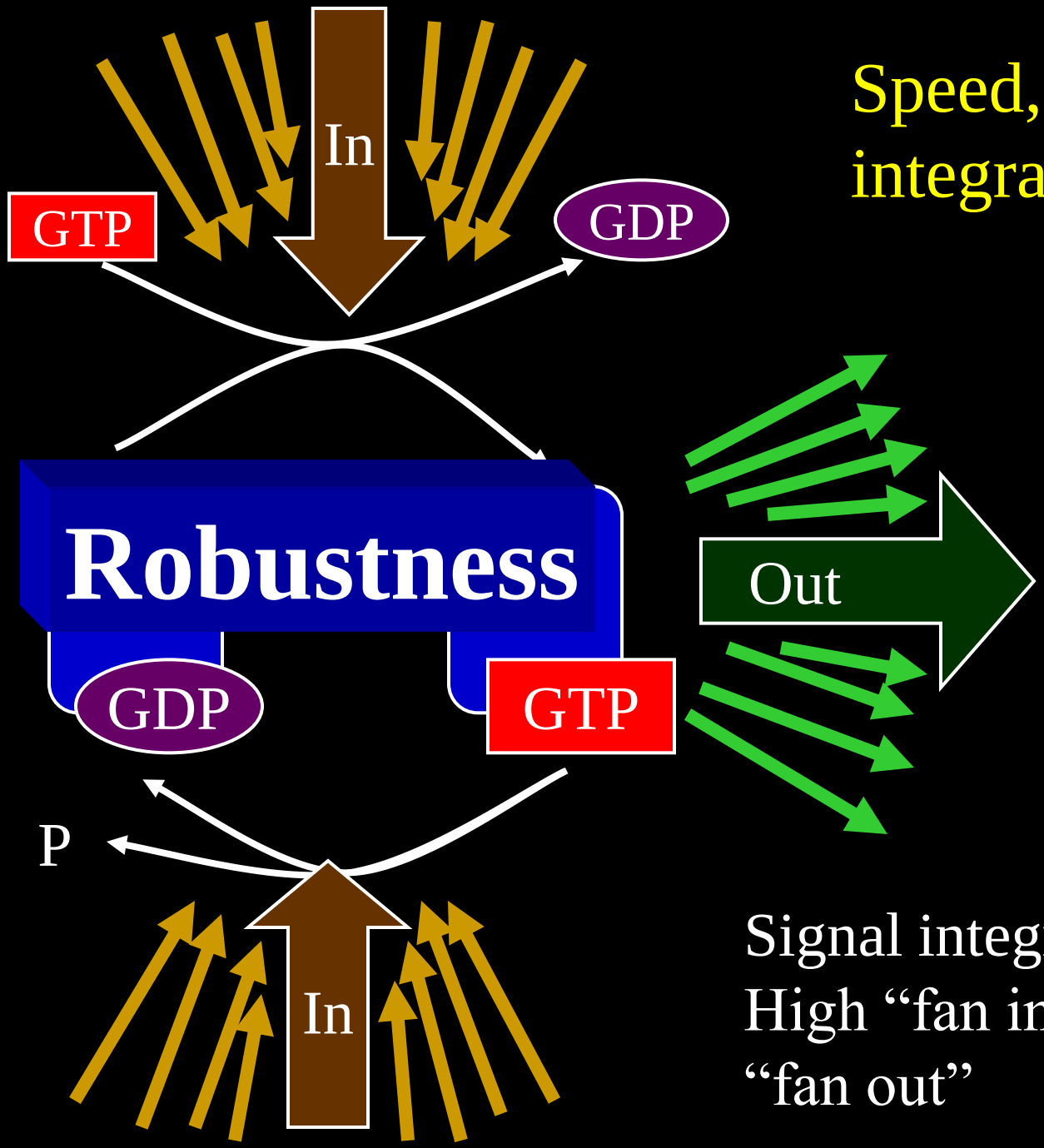
Eukaryotes have lots more bowties

- More elaborate organization at every level
- Surprise: stoichiometry is not that much more complicated
- But complexity of regulation appears almost arbitrarily greater
- Analogous to analog versus digital control systems? (e.g. from hundreds to billions of transistors?)
- GPCRs and NFκB are extreme and extremely important examples


Ligand
Receptor
GTP
GDP
βγ
βγ
Gα
GDP
Gα
GTP
Out
P
RGS


GEF
GTP
GDP
GDI
GDI
Rho
GDP
Rho
GTP
Out
P
GAP

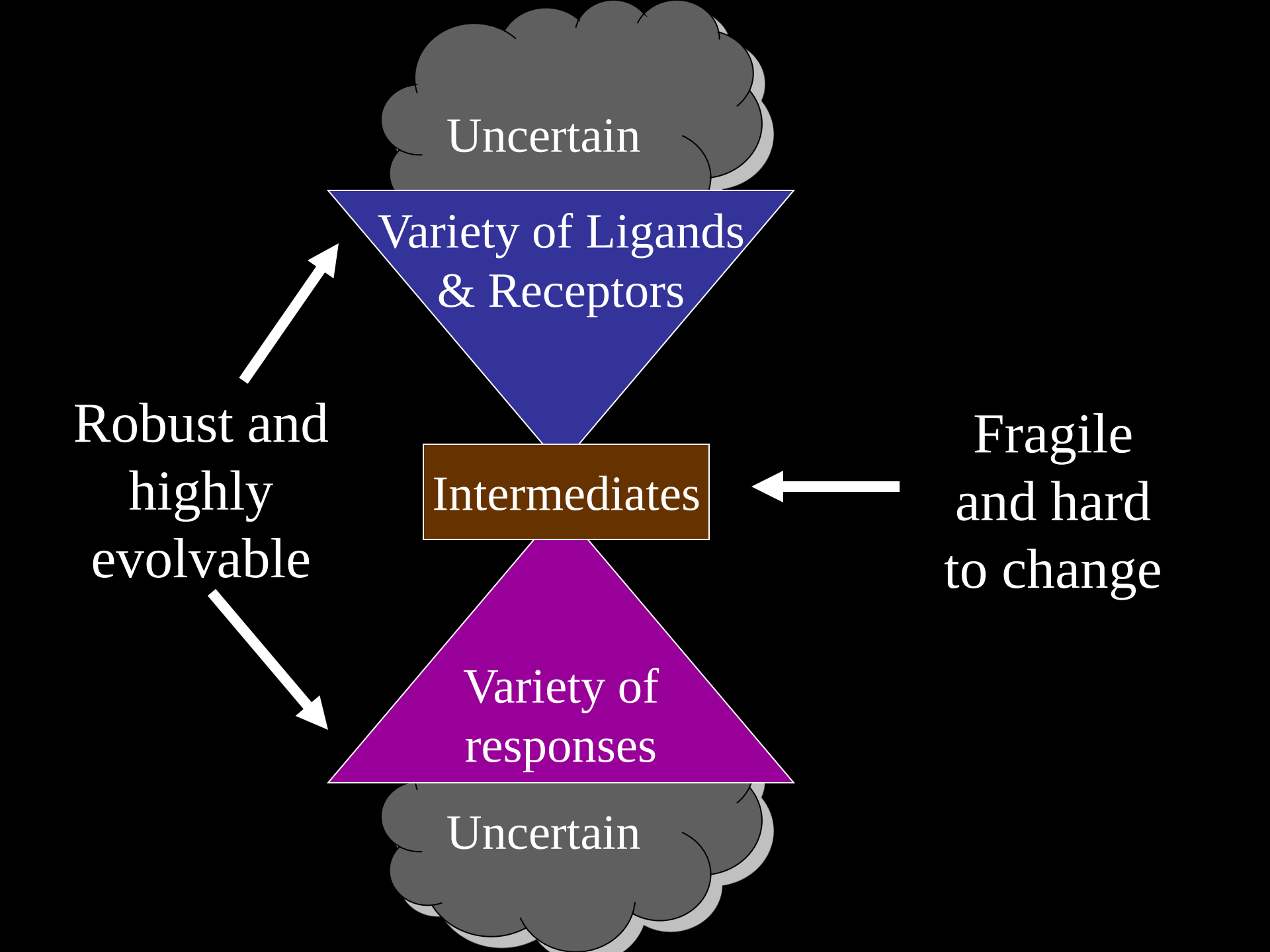

In
GTP
GDP
GDP
GTP
Out
P
In
Receptors
G-proteins
Primary targets
Responses



Speed, adaptation, integration, evolvability

Impedance matching: Independent of ΔG of inputs and outputs

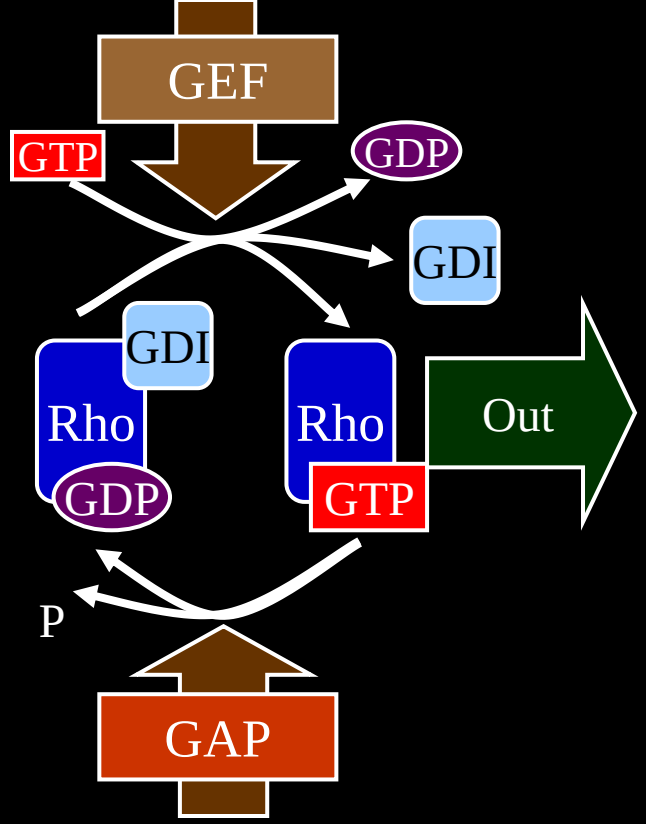
Signal integration: High “fan in” and “fan out”


Uncertain
Variety of Ligands & Receptors
Robust and highly evolvable
Intermediates
Fragile and hard to change
Variety of responses
Uncertain

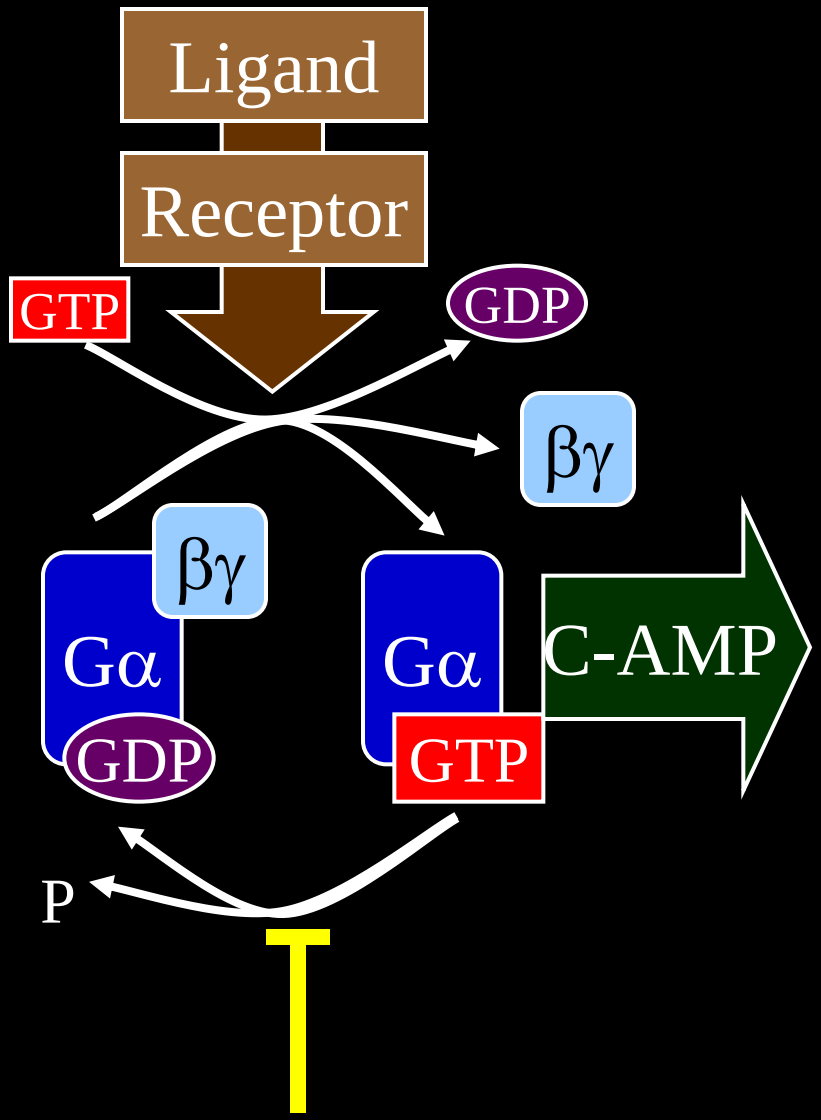
Fragility?

- A huge variety of pathogens attack and ***hijack*** GTPases.
- A huge variety of cancers are associated with altered (***hijacked***) GTPase pathways.

- The GTPases may be the least evolvable elements in signaling pathways, in part because they facilitate evolvability elsewhere



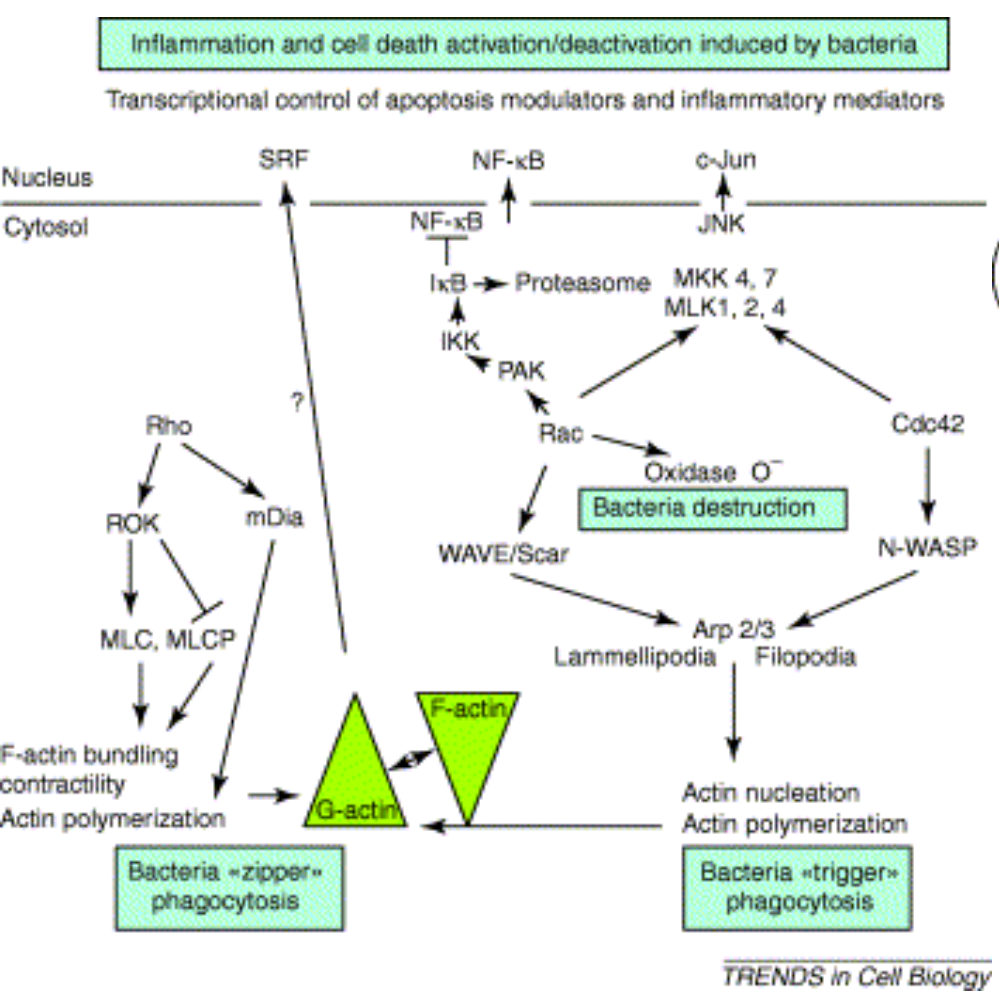
Hijacking

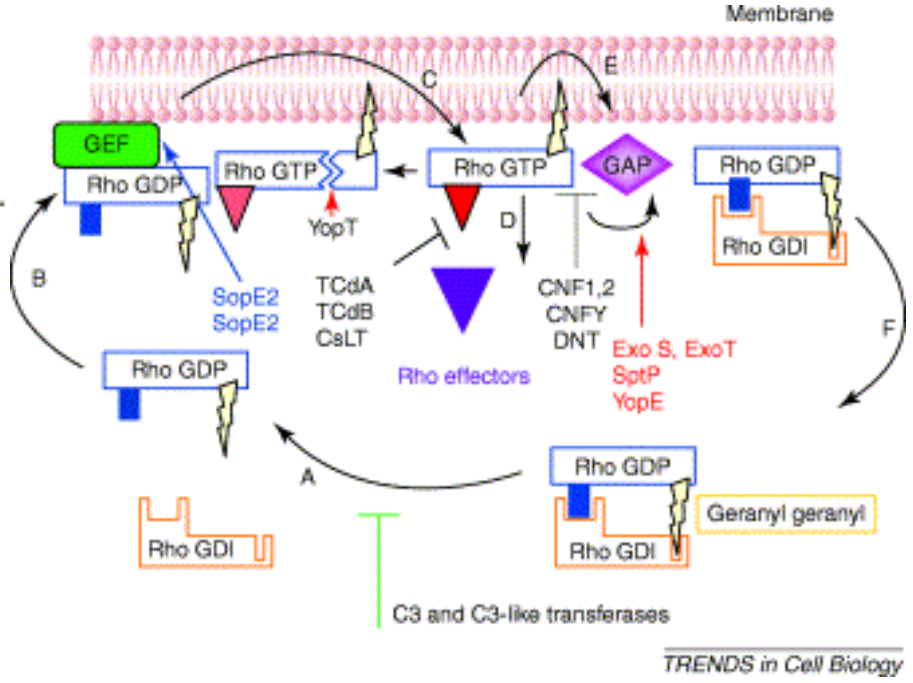


Cholera toxins hijack the signal transduction by blocking a GTPase activity.

Bacterial virulence factors targeting Rho GTPases: parasitism or symbiosis?

Patrice Boquet and Emmanuel Lemichez



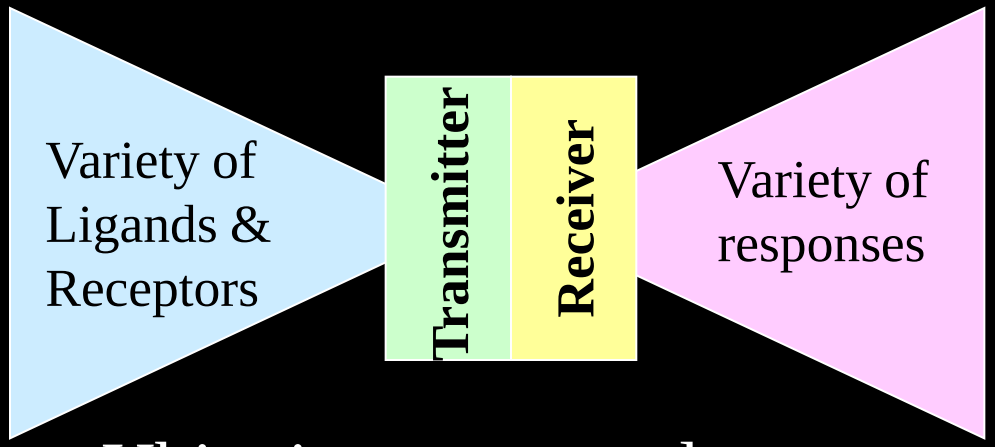


TRENDS in Cell Biology May 2003

Toxin and virulence factors	Biochemical activity	Cellular targets/effects	Pathogens[a]
Toxins			
Toxin A	UDP-glucosyl transferase	Rho, Rac, Cdc42, RhoG, TC10 inactivation	C *difficile*
Toxin B	UDP-glucosyl transferase	Rho, Rac, Cdc42, RhoG, TC10 inactivation	C *difficile*
Toxin B-1470	UDP-glucosyl transferase	Rac1, Ral, Rap1, Ras, Cdc42, RhoG, TC10 inactivation	C *difficile*
Lethal toxin	UDP-glucosyl transferase	Rac1, Ral, Rap1, Ras, RhoG, TC10 inactivation	C *sordellii*
Hemorrhagic toxin	UDP- glucosyl transferase	Rho, Rac, Cdc42 inactivation	C *sordellii*
α toxin	UDP-N-acetyl-glucosamine transferase	Rho, Rac, Cdc42 inactivation	C *novyi*
CNF1 and CNF 2, CNFY	Glutamine deamidase	Rho, Rac, Cdc42 activation/ degradation	E, Y
DNT	Glutamine deamidase/ transglutaminase	Rho, Rac, Cdc42 activation/ (degradation?)	Bo
Virulence factors with unknown type of translocation			
C3 transferase	ADP-ribosyl transferase	RhoA, B, C inactivation	C *botulinum*
C3-related transferase	ADP-ribosyl transferase	RhoA, B, C inactivation	C *limosum*
C3-related transferase	ADP-ribosyl transferase	RhoA, B, C inactivation	*B. cereus*
EDIN	ADP-ribosyl transferase	RhoA, B, C inactivation	St
Stau	ADP-ribosyl transferase	RhoA, B, C, Rnd3 inactivation	St
CDT	ADP-ribosyl transferase	RhoA, B, C inactivation	C *difficile*

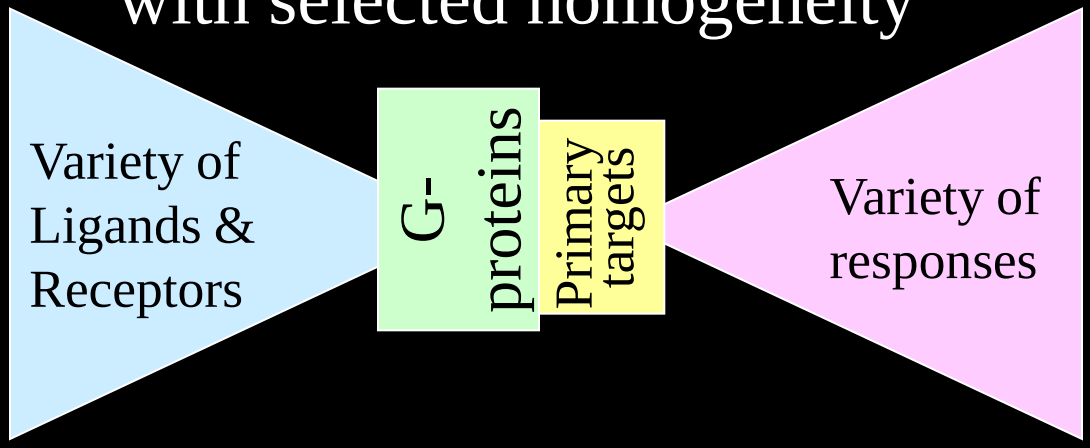
Type 3 translocated virulence factors			
SopE and SopE2	GDP-GTP exchange factor	Cdc42, Rac activation	Sa
SptP	GTPase activating protein (N-ter)	Cdc42, Rac inactivation. No activity on small GT-Pases	Sa
	Phosphatase (C-ter)		
YopT	Cysteine protease	Rho, Rac, Cdc42 inactivation	Y
YopE	GTPase activating protein	Rho, Rac, Cdc42 inactivation	Y
YpkA/YopO	Ser/Thr kinase RhoA and Cdc42 binding	RhoA and Cdc42 (activity unknown)	Y
IpaC	Unknown	Rac, Cdc42 activation	Sh
ExoS	GTPase activating protein (N-ter)	RhoA, Cdc42, Rap1 inactivation	P
	ADP-ribosyltransferase (C-ter)	Ras, Rap1, Rap2, Ral, Rac1, RhoA, Cdc42 inactivation	
ExoT	GTPase activating protein (N-ter)	Rho, Rac, Cdc42	P
	ADP-ribosyltransferase (C-ter)	No activity on small GT-Pases tested	
SopB/SigD	PtdIns(4,5)P_2 phosphatase	Cdc42 Indirect activation?	Sa
Type 4 secretory mechanism and bacterial adhesions			
CagA pathogenicity island (PAI)	Unknown	Rac1, Cdc42 activation	Hp
Opacity proteins (Opa 52)	Activation of Rac1 via Hck/Fgr kinase stimulation	Rac1 activation	Ng
Type IV pilus	Receptor clustering?	Rho, Cdc42 activation	Nm
Type 1 (FimH adhesin)	Receptor clustering?	Rho, Rac, Cdc42 activation	E

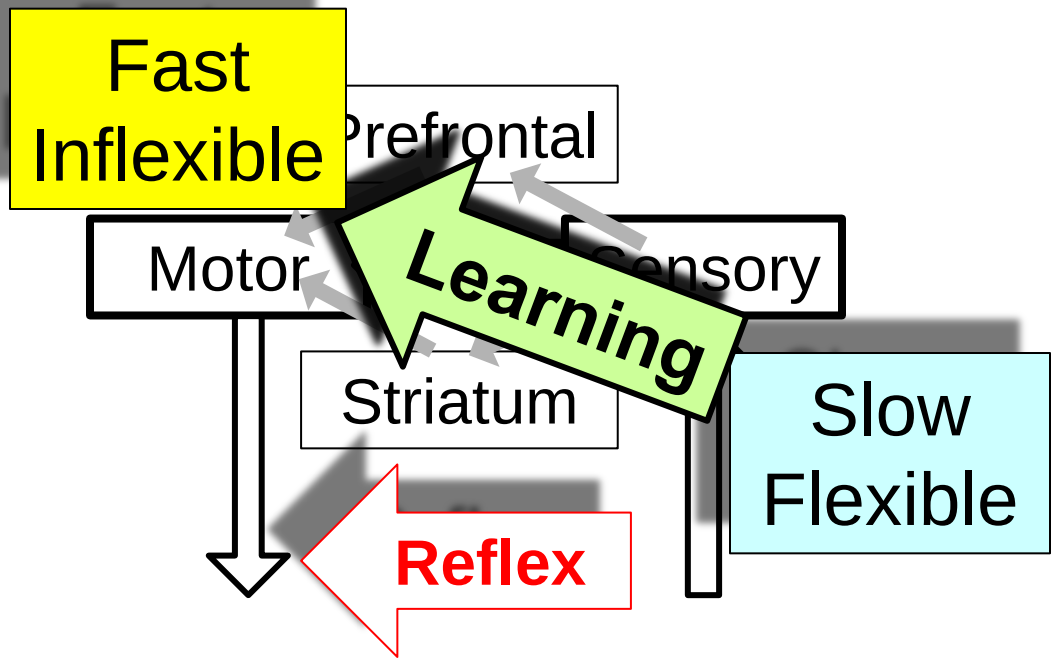
Signal transduction



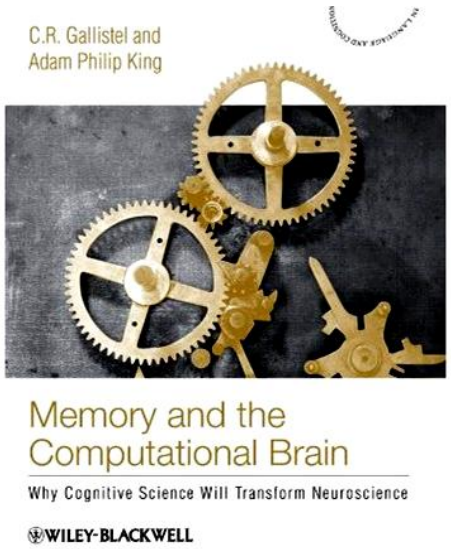
- Ubiquitous protocol
- “Robust yet fragile”
- Robust & evolvable
- Fragile to “hijacking”
- Manages extreme heterogeneity with selected homogeneity

~~Accident~~ or necessity?

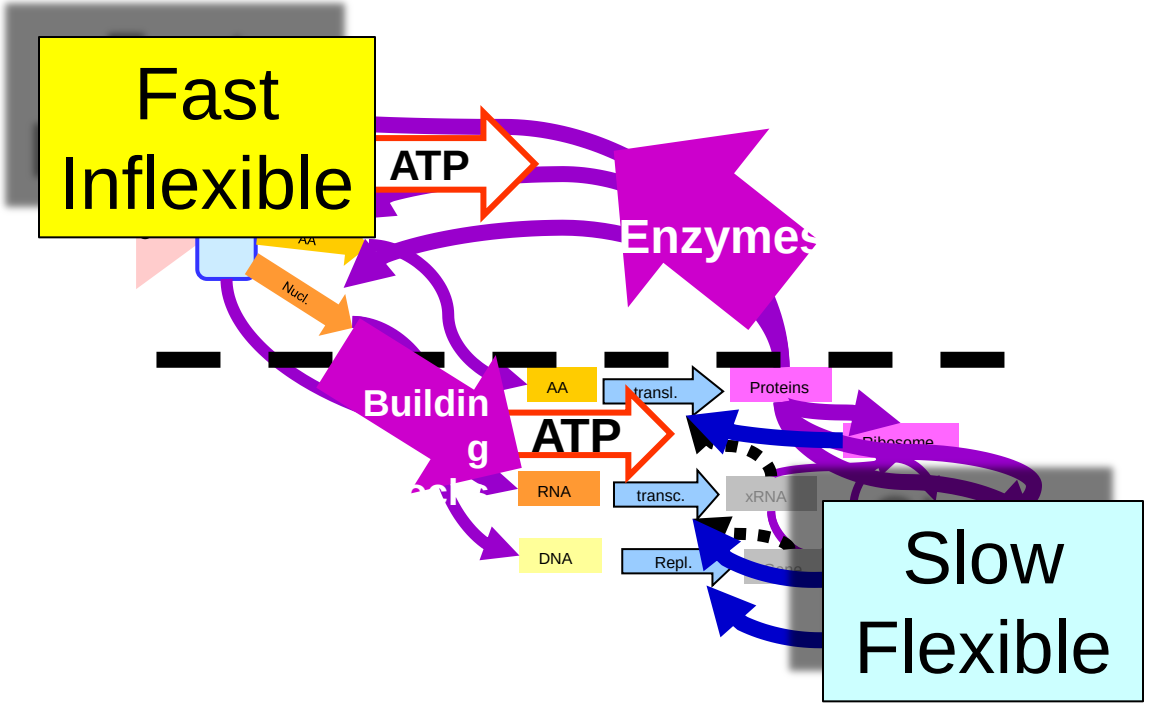




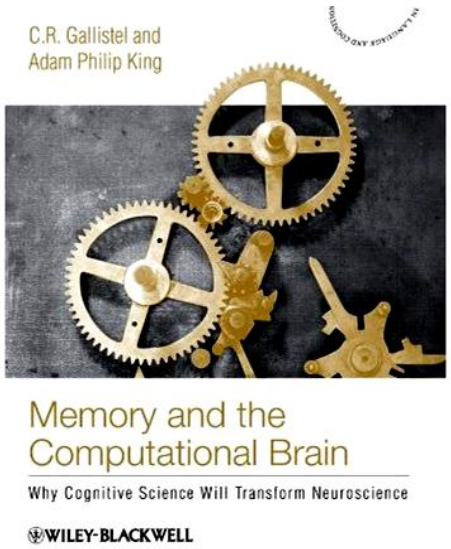
Gallistel and King



- Sensori-motor memory potential $\approx \infty$
- Limits are on ***speed*** of
  - nerve propagation delays
  - learning
- But control is ***never*** centralized
- Where is R/W random access memory (RAM)?



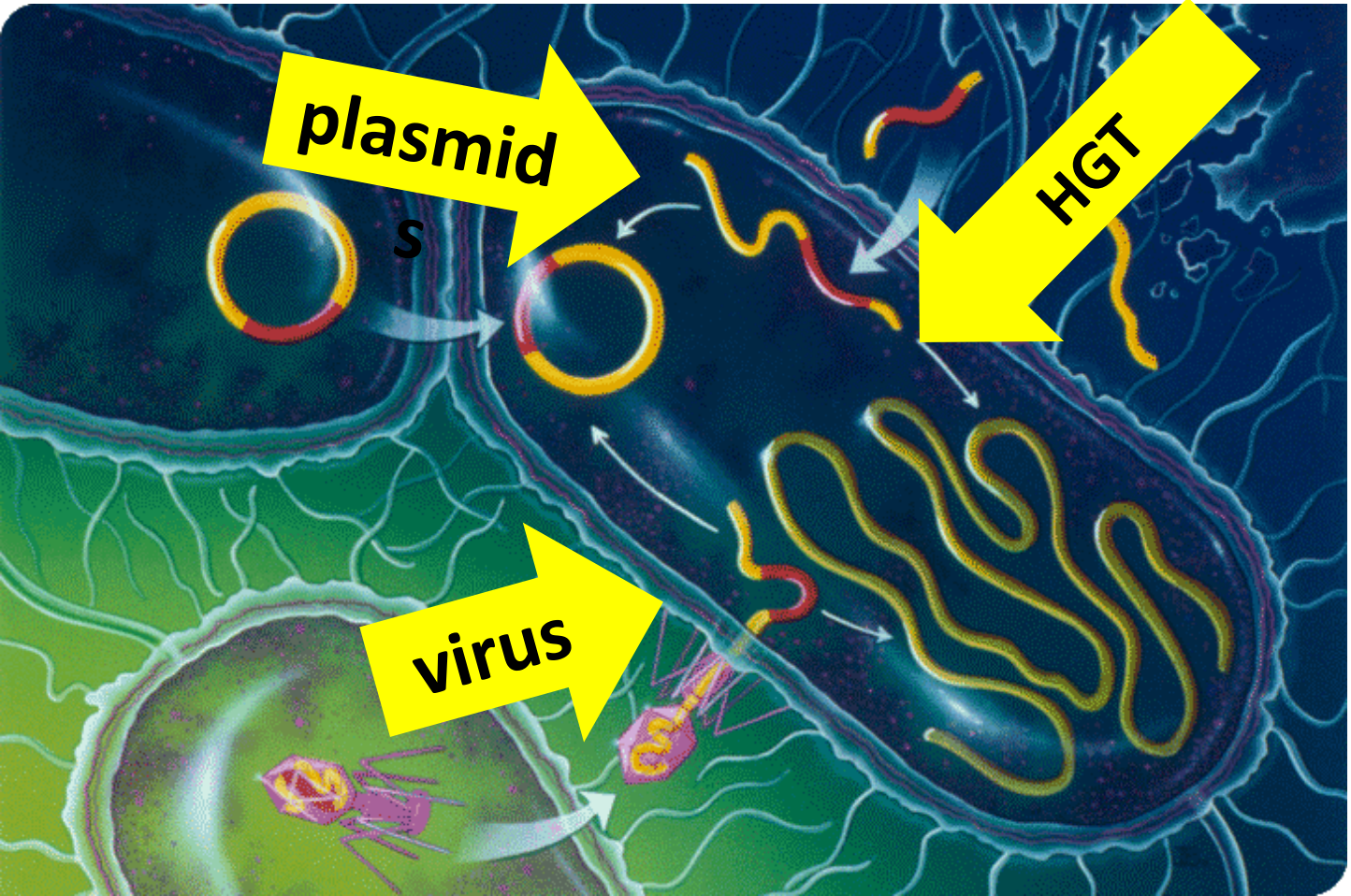
Gallistel and King



- Genome memory potential ≈ ∞
- Limits are on ***speed*** of control and learning
- Control is highly ***de***centralized
- There is a huge slow read/write RAM
- Sophisticated naming and addressing

selection + drift + mutation + gene flow

+ facilitated *variation*



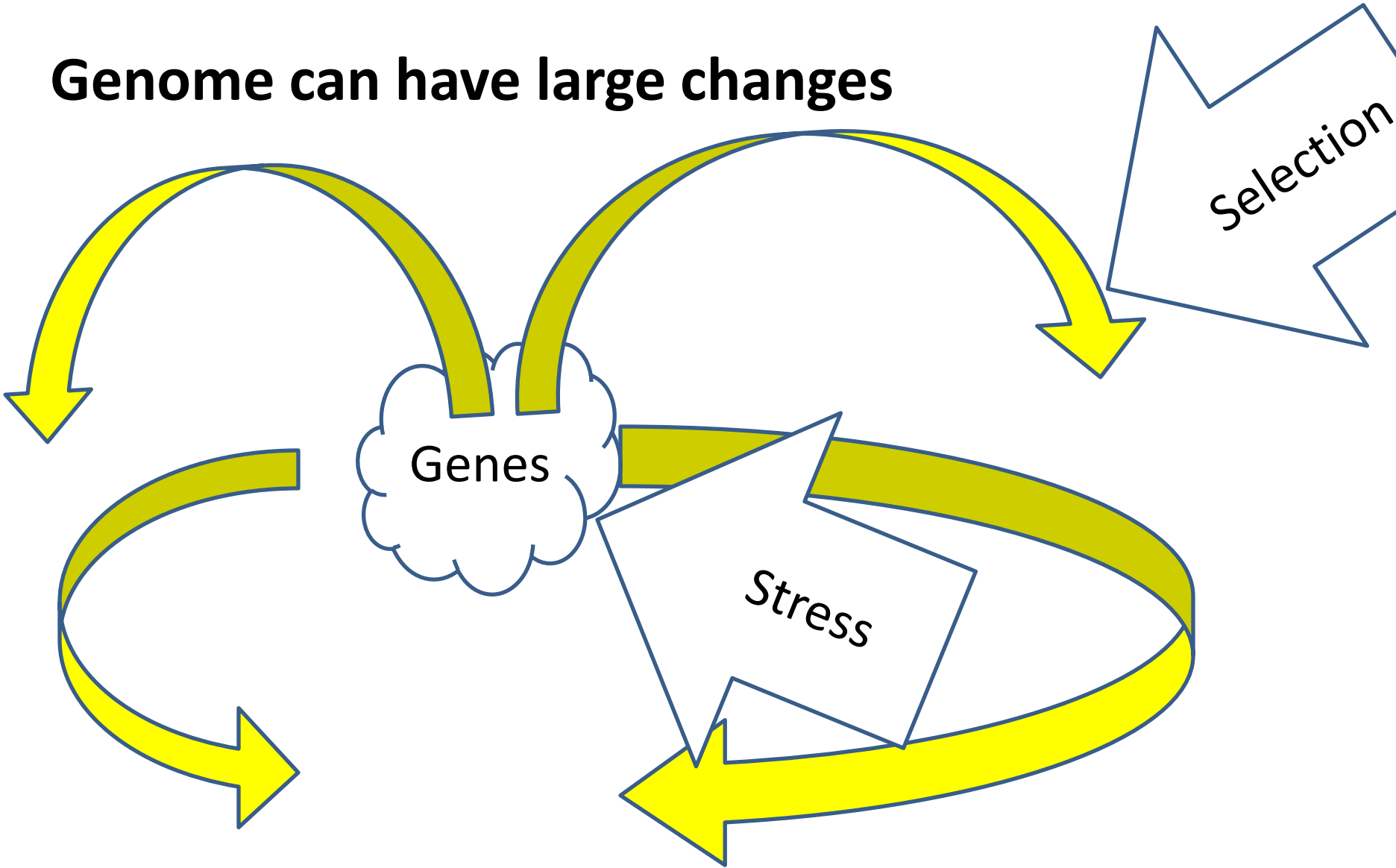
large functional changes in genomes

HGT
= horizontal gene transfer

natural selection + genetic drift + mutation + gene flow

+ facilitated *variation*

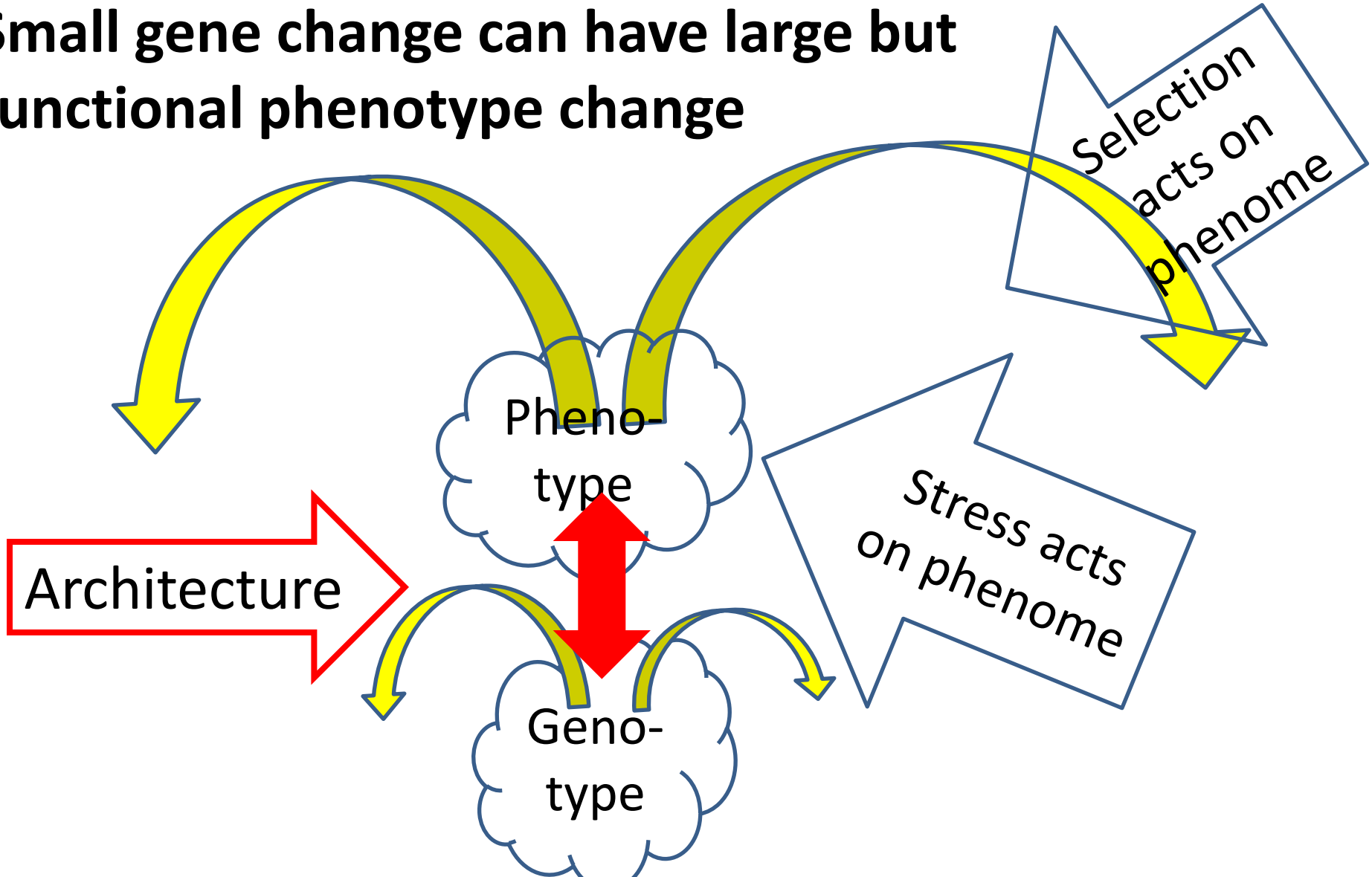
Genome can have large changes



natural selection + genetic drift + mutation + gene flow

+ facilitated *variation*

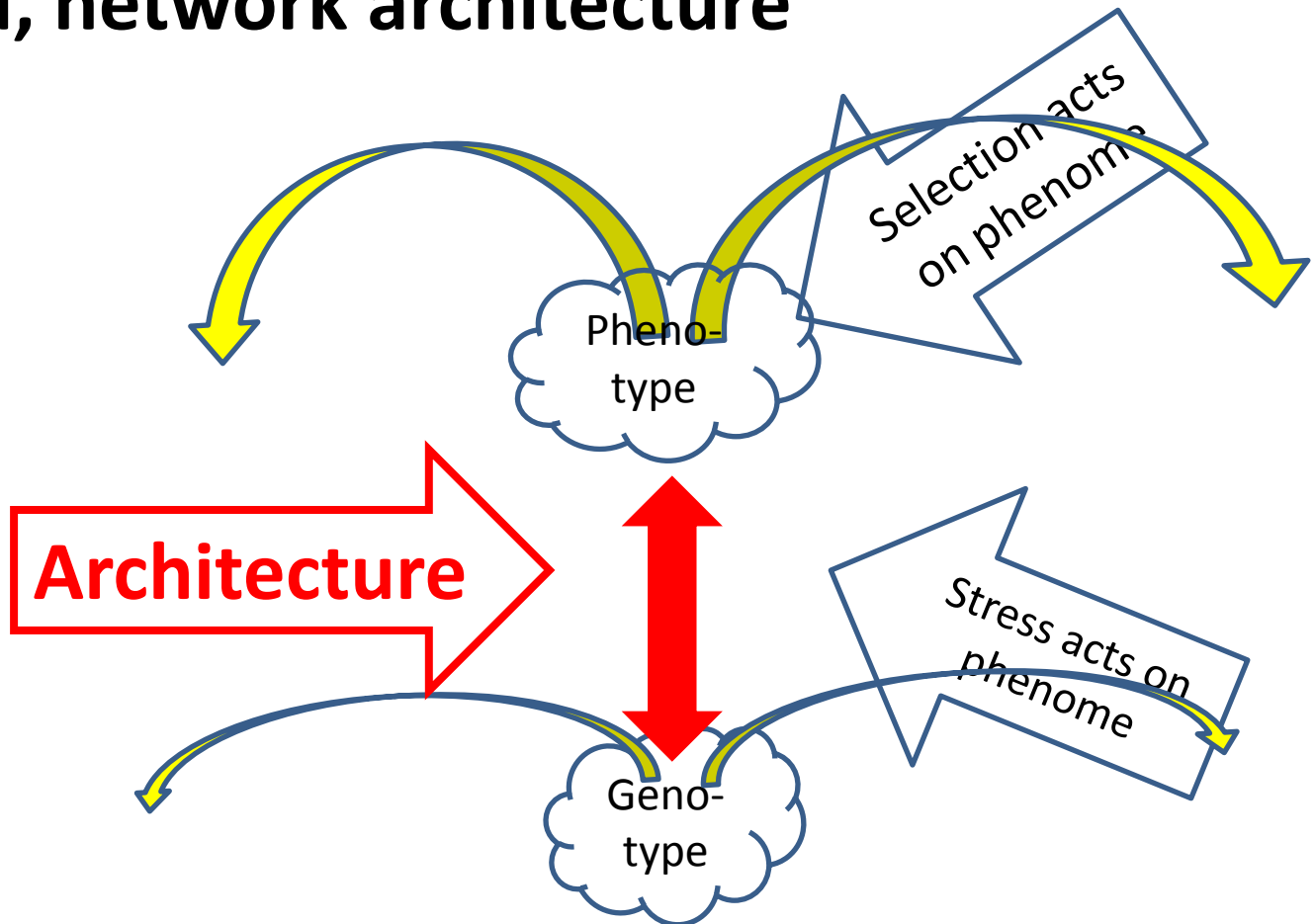
Small gene change can have large but functional phenotype change



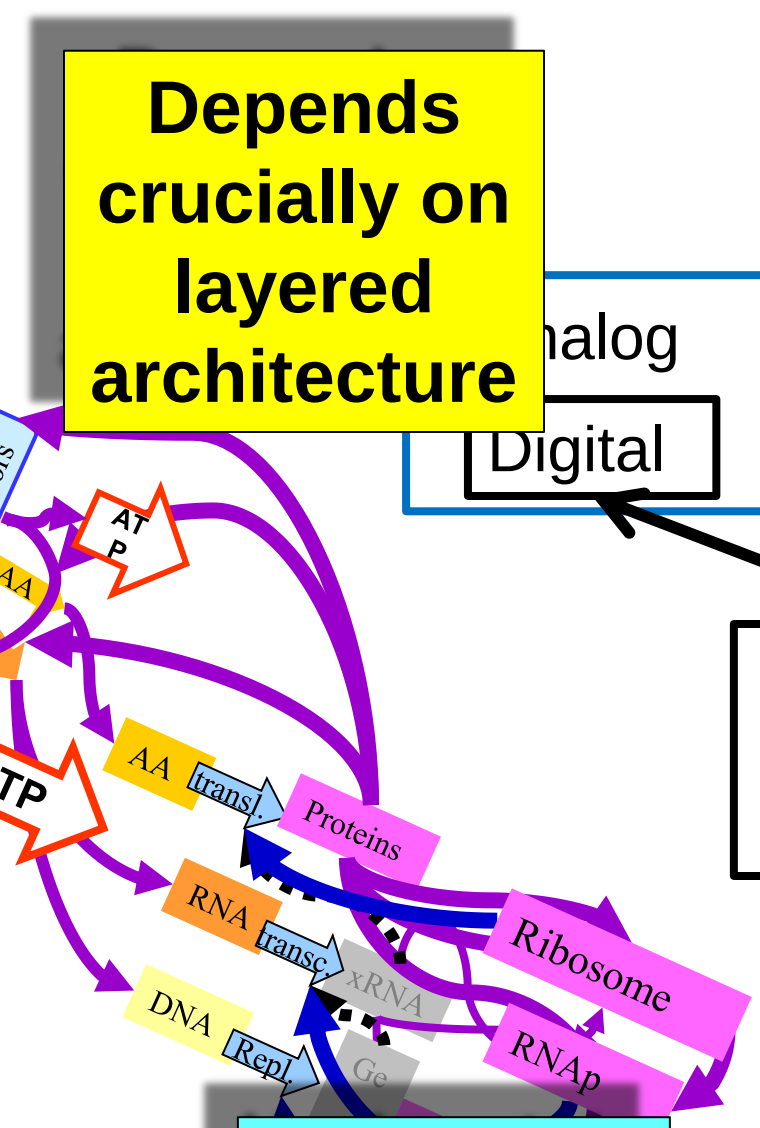
natural selection + genetic drift + mutation + gene flow

+ facilitated *variation*

Only possible because of shared, layered, network architecture



Depends crucially on layered architecture

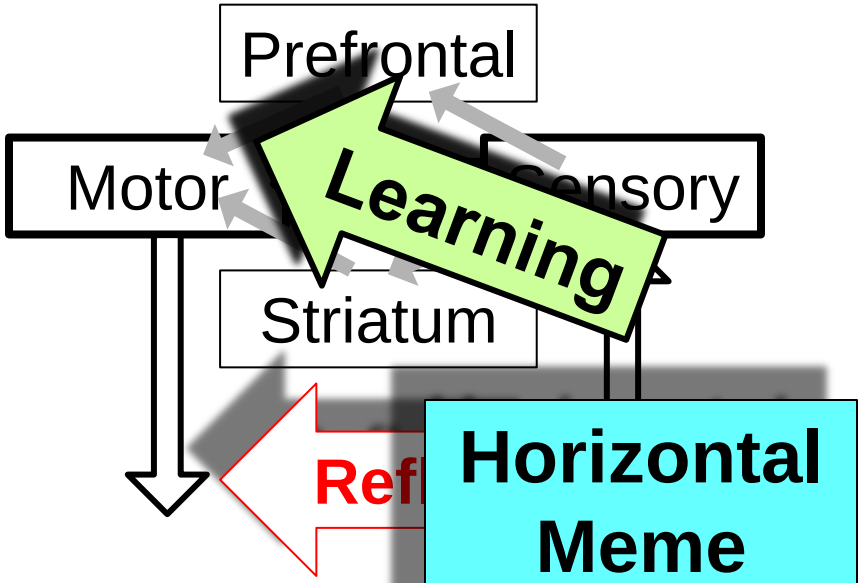


Horizontal Gene Transfer

nalog
Digital

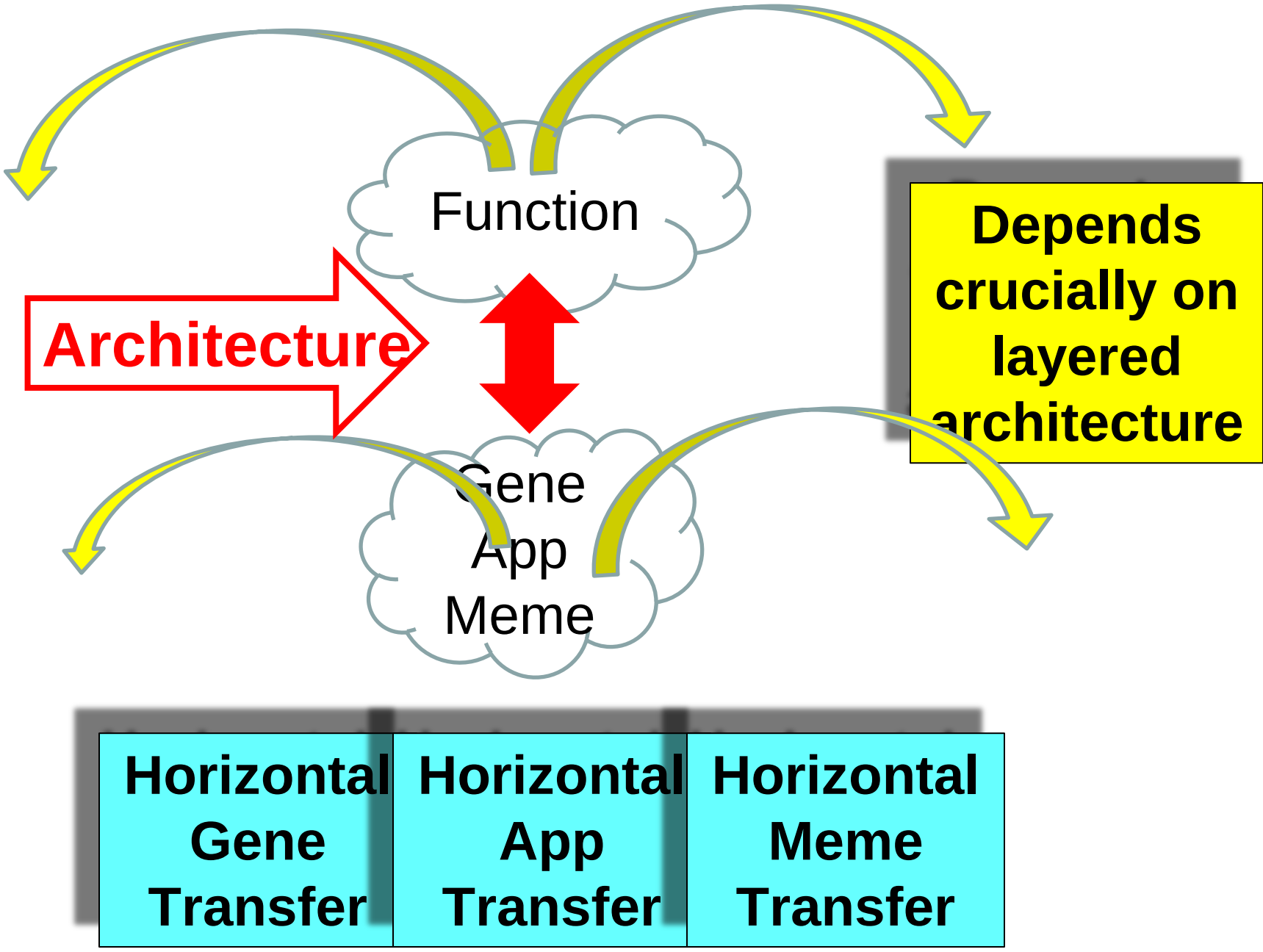
Hardware
Soft

Horizontal App Transfer



Horizontal Meme Transfer

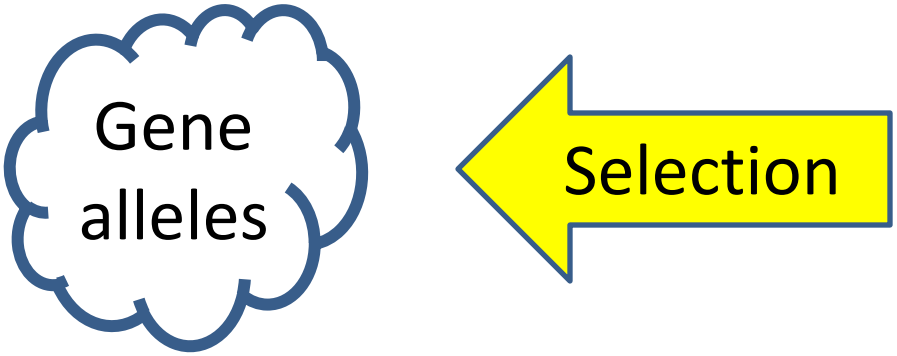
Amazingly Flexible/ Adaptable


Function
Architecture
Gene
App
Meme
Depends crucially on layered architecture
Horizontal Gene Transfer
Horizontal App Transfer
Horizontal Meme Transfer

Standard theory:

natural selection + genetic drift
+ mutation + gene flow

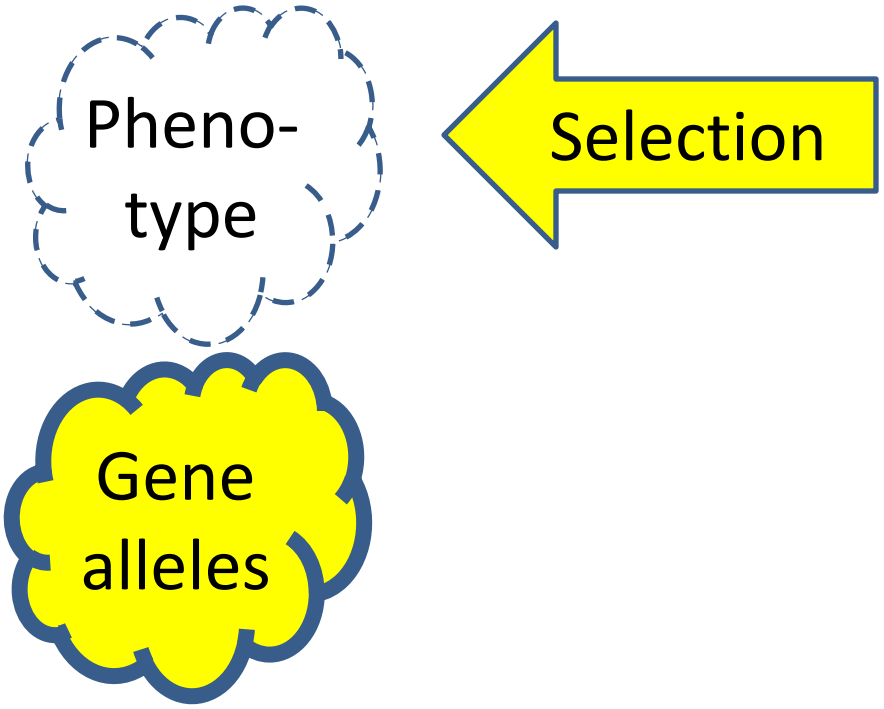
Greatly abridged cartoon here



Shapiro explains well what this is and why it's incomplete (but Koonin is more mainstream)

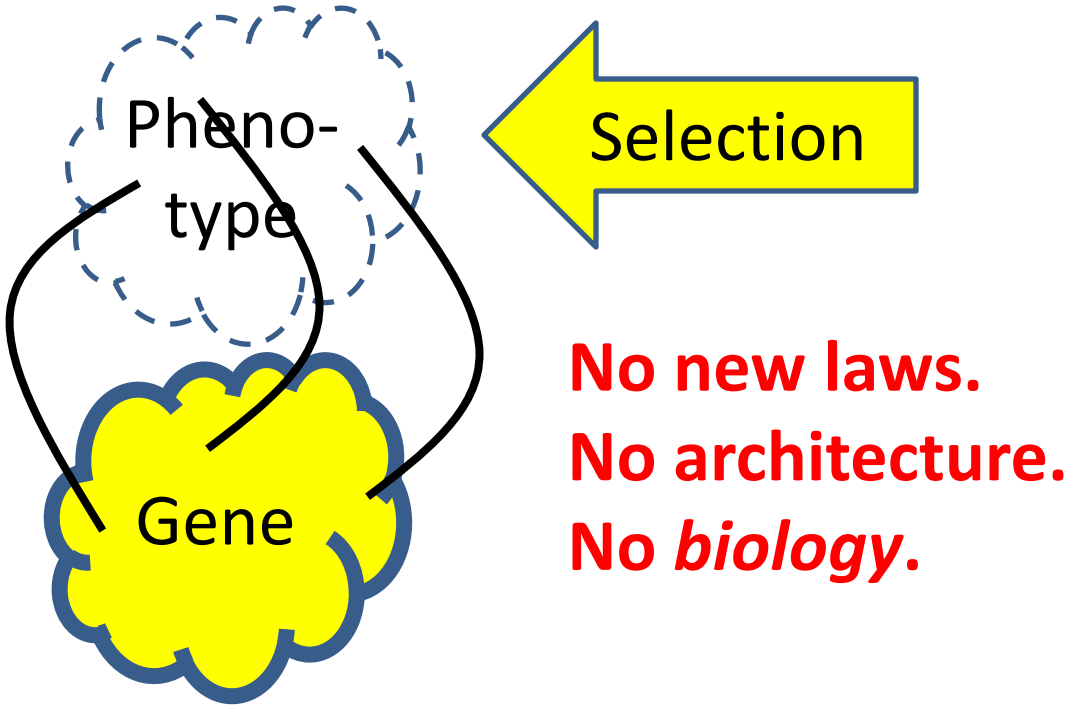
Standard theory:

selection + drift + mutation + gene flow



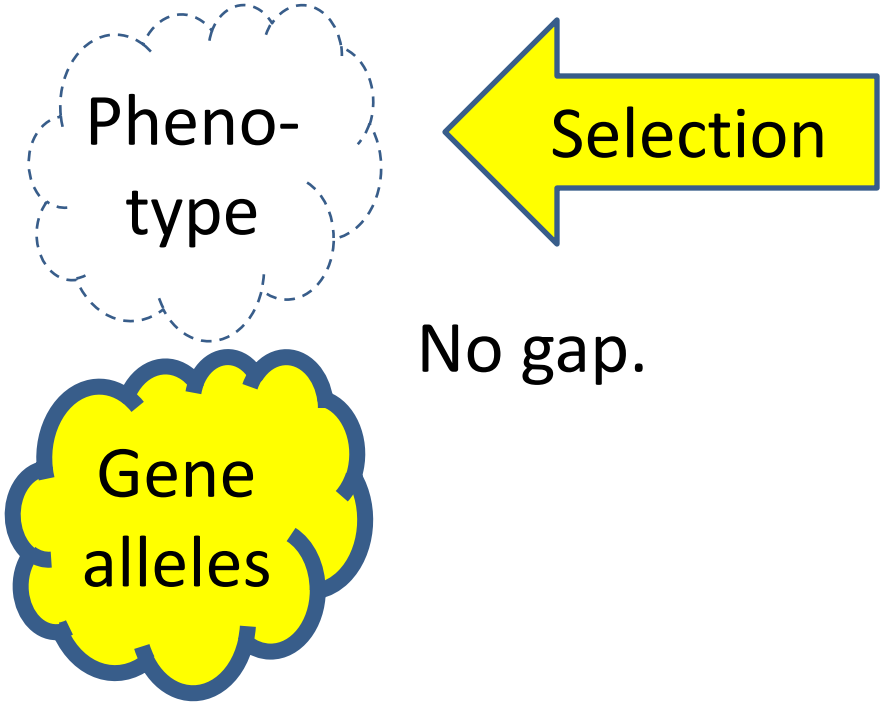
Standard theory:

selection + drift + mutation + gene flow



No new laws.
No architecture.
No *biology*.

selection +
drift +
mutation +
gene flow



***All* complexity is *emergent* from random ensembles with *minimal* tuning .**

No new laws.

No architecture.

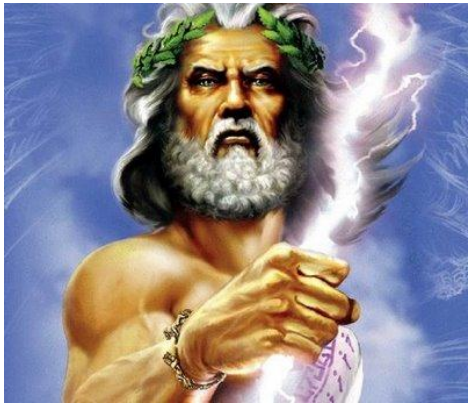
The battleground

Pheno-type

Pheno-type

Gene alleles

No gap.
Just physics.



Huge gap.
Need supernatural

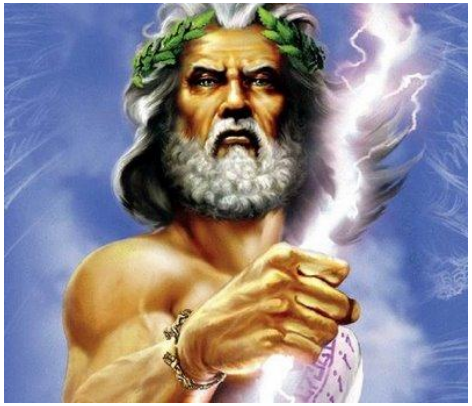
Genes?

What they agree on

Pheno-type

No new laws.
No architecture.
No biology.

Huge gap.



Pheno-type

No gap.

Gene alleles

Genes

Depends crucially on layered architecture

Prefrontal

Motor

Sensory

Striatum

Learning

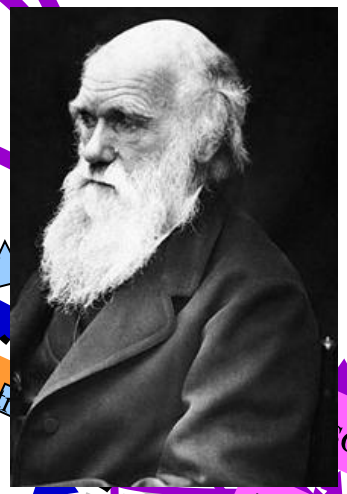
Digital

Hardware

Soft

Horizontal Meme Transfer

Horizontal App Transfer



AA

transl.

RNA

DNA

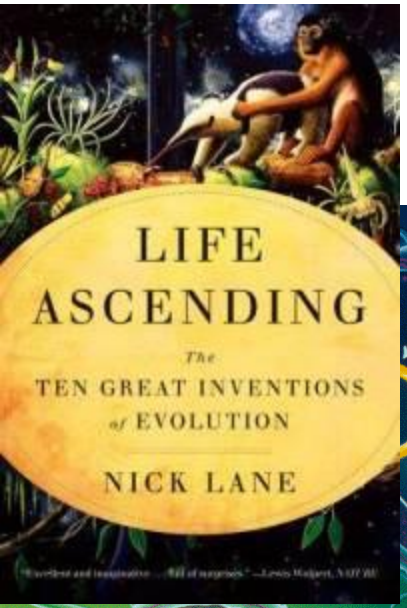
Repl.

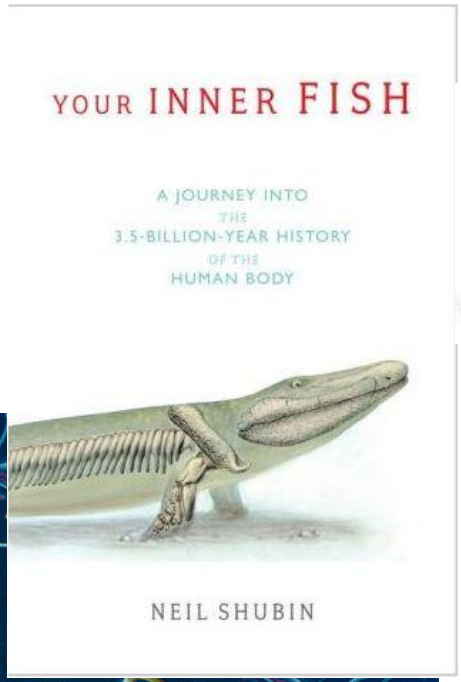
RNAp

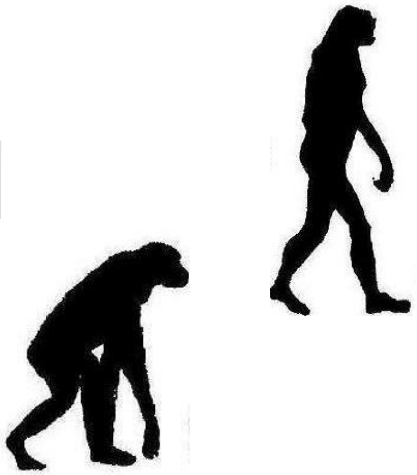
Horizontal Gene Transfer

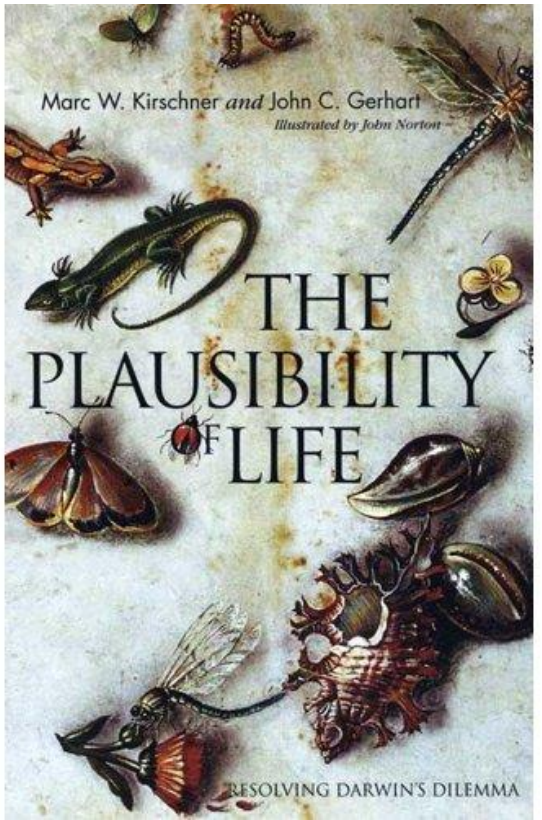
Amazingly Flexible/ Adaptable

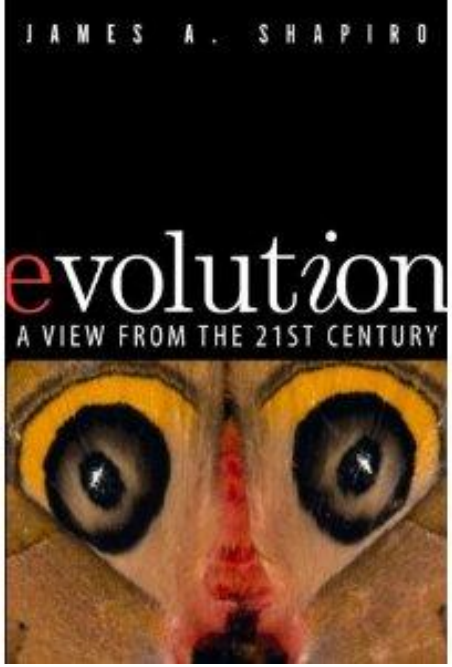
Putting biology back into evolution

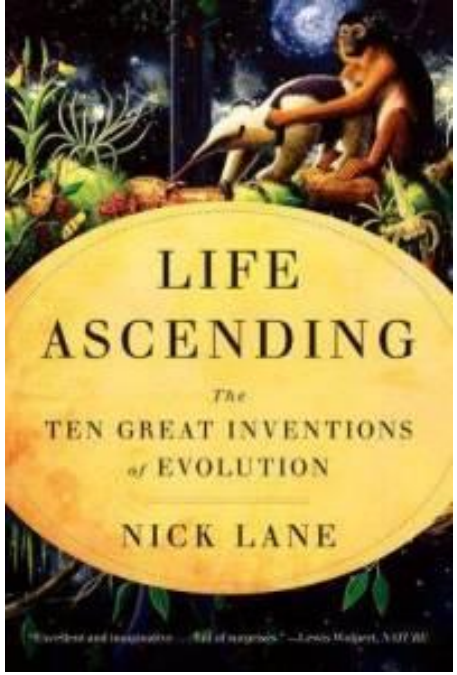


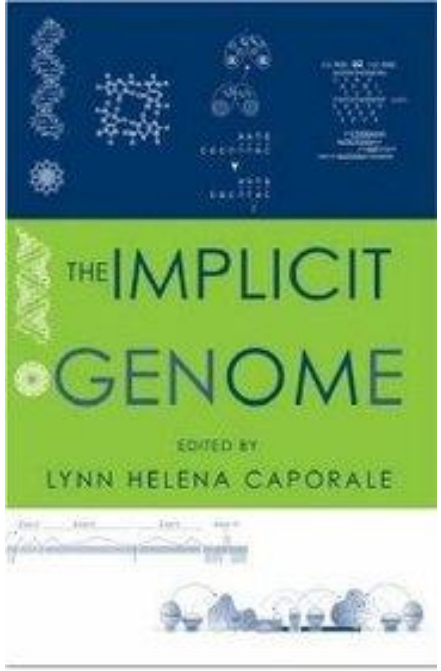


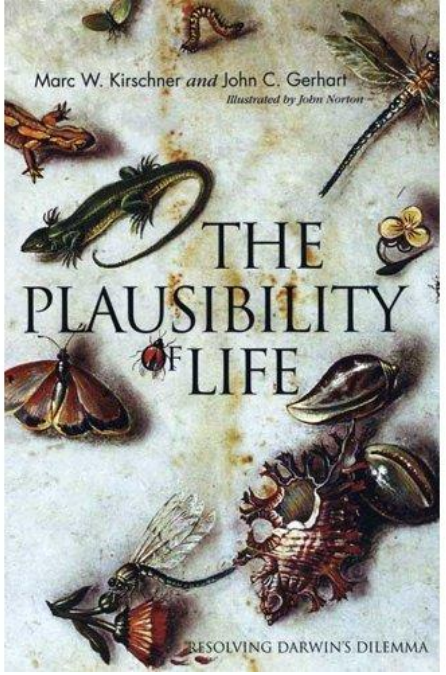






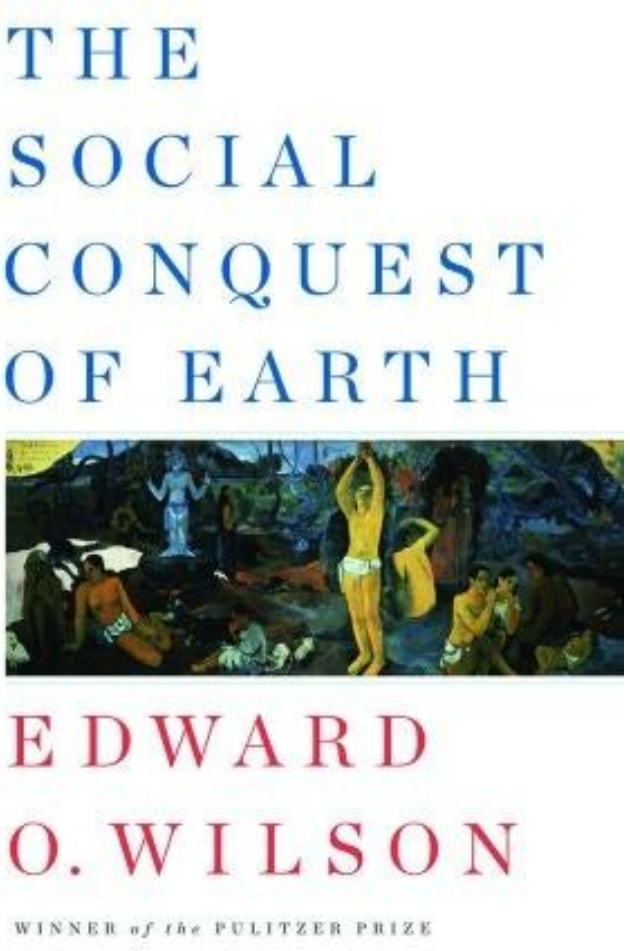




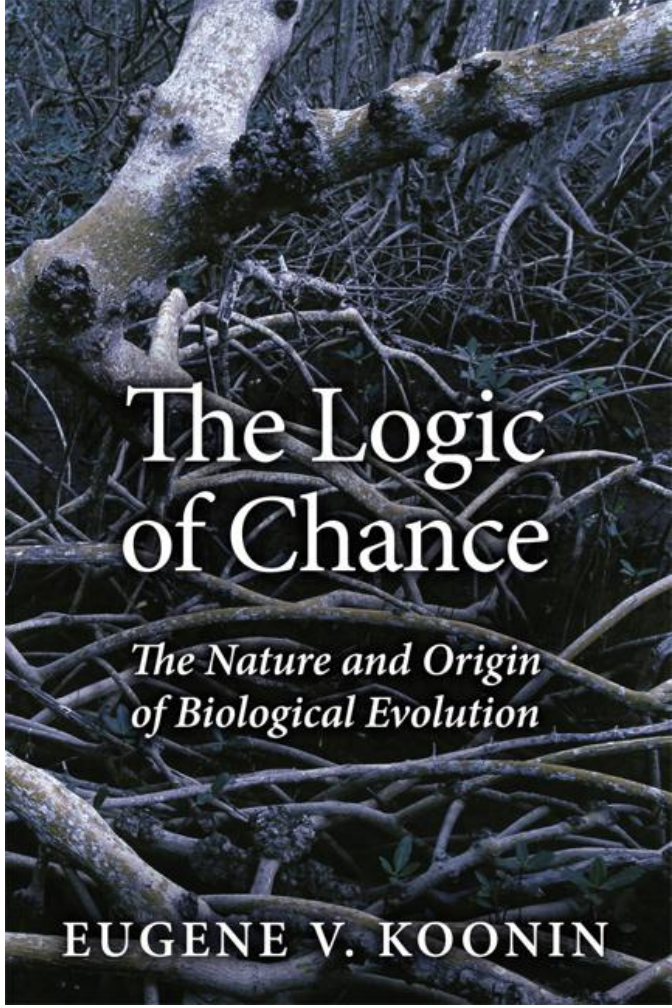


The heresies

- Many mechanisms for “horizontal” gene transfer
- Many mechanisms to create large, functional mutations
- At highly variable rate, can be huge, global
- Selection alone is a very limited filtering mechanism
- Mutations ***can*** be “targeted” within the genomes
- ***Can*** coordinate DNA change w/ useful adaptive needs
- Viruses ***can*** induce DNA change giving heritable resistance
- Still myopic about future, still produces the grotesque



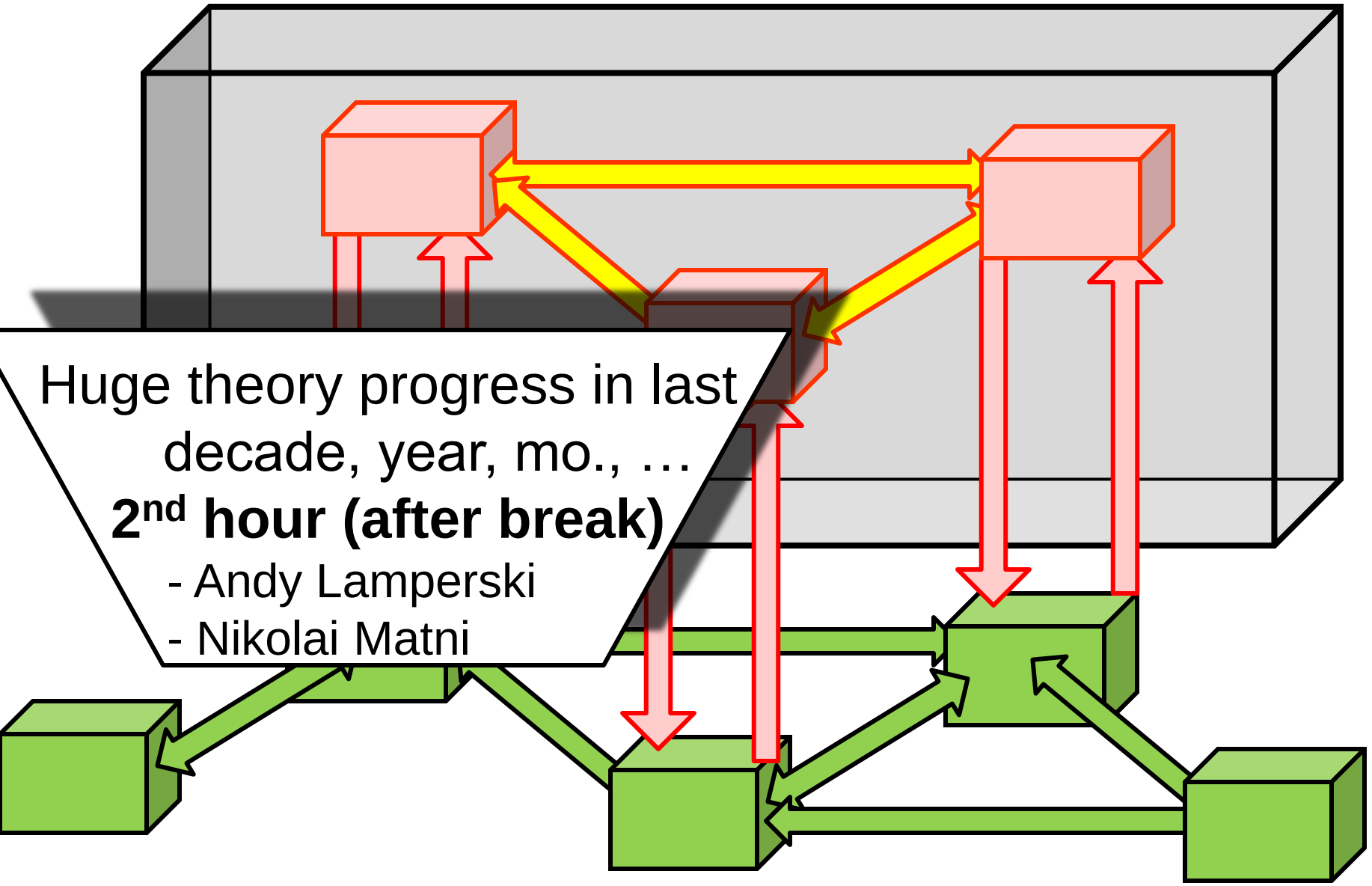
Surprising heresies from "conservatives"



~~kin selection~~

~~modern synthesis~~

Going beyond black box: control is decentralized with internal delays.

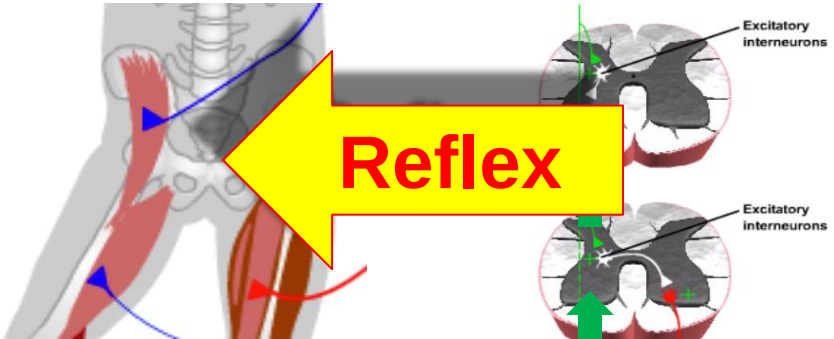


Wolpert, Grafton, etc

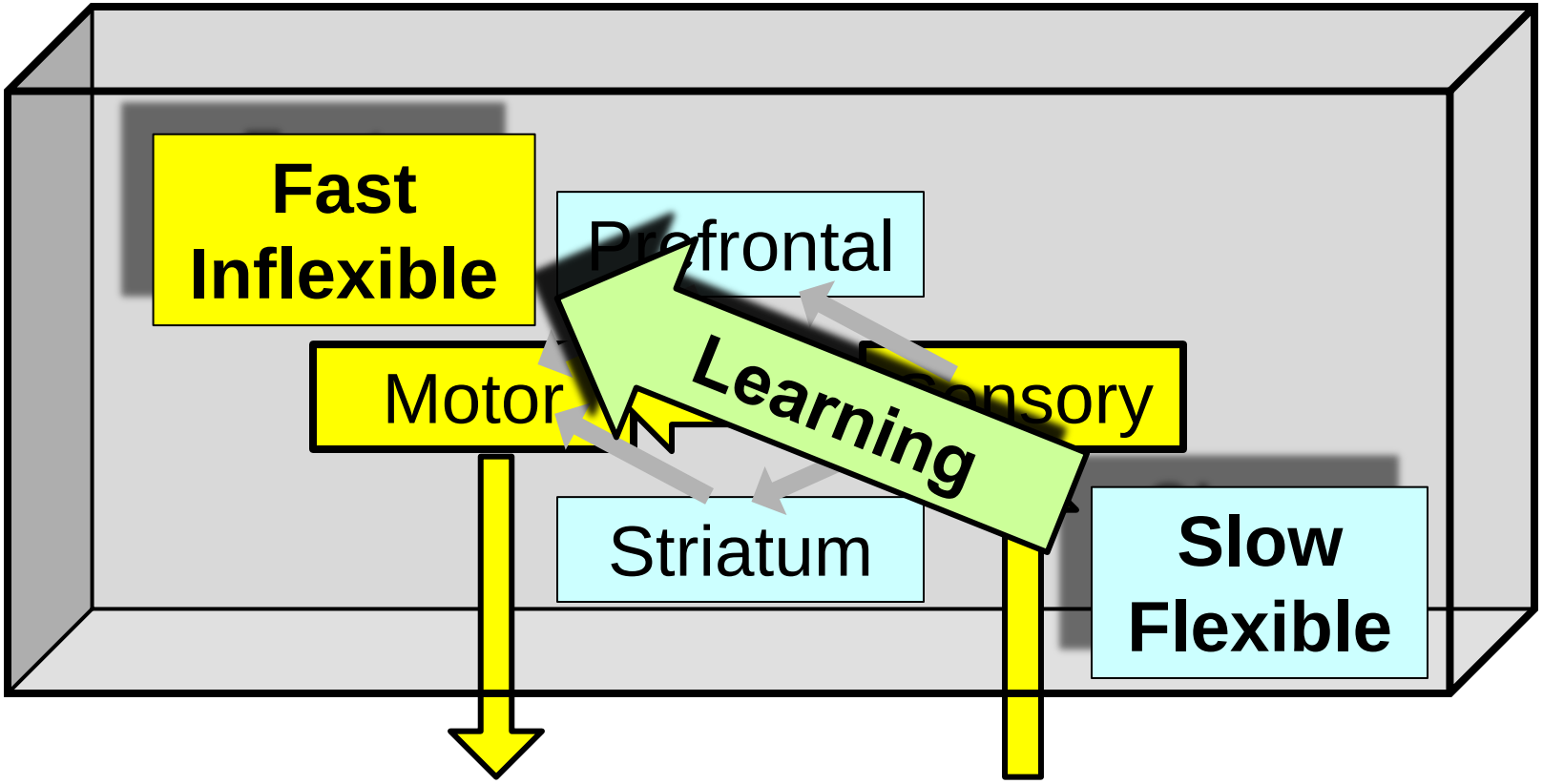
robust

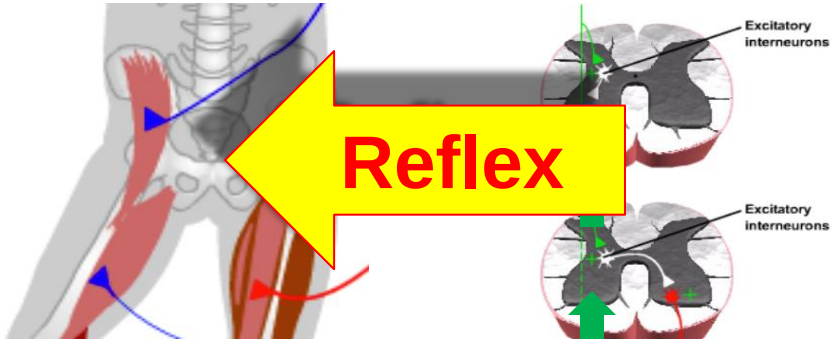
Brain as ~~optimal~~ controller

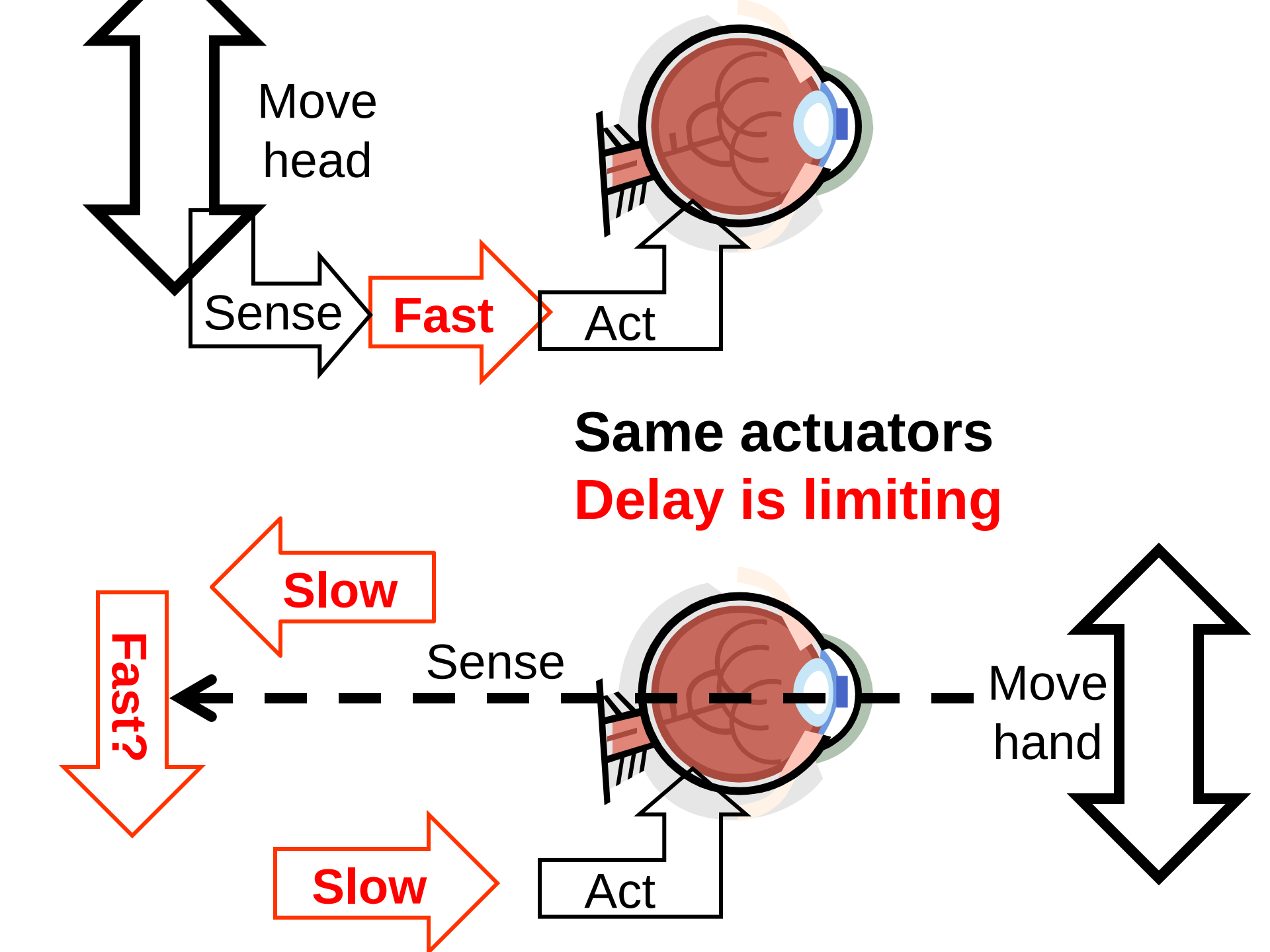
- Acquire
- Translate/ integrate
- **Automate**



Going beyond black box: control is decentralized with internal delays.






Move
head
Sense
Fast
Act
Same actuators
Delay is limiting
Slow
Fast?
Sense
Move
hand
Slow
Act