Network Based Models in Bioinformatics and Biocomputing

Instructor: Dr. Weller

Lecture 5

Assignments

- Finish Chapter 2 in The Regulatory Genome, first 20 pages of Chapter 3
- Homework 3 is posted.

Agenda

- Methods: measuring and interpreting the turnover of proteins and mRNA
  - Protein Note: please see the paper by Robert Beynon "Technique Review - The Dynamics of the Proteome: Strategies for measuring protein turnover on a proteome-wide scale" in Briefings in Functional Genomics and Proteomics, vol 3(4) 382-390 (2005) for full information, and credit for the figures.
  - RNA note: please see the paper by Roy Parker and Haiwei Song "Review - The enzymes and control of eukaryotic mRNA turnover" in Nature Structural and Molecular Biology, vol 11(2) 121-128 (2004).
- Chapter 2 from Davidson – part 2
  - Logic circuit representation
- Begin Chapter 3 from Davidson

Proteome Dynamics

- The correlation between the abundance of proteins and their mRNAs is not strong.
  - Steady-state measurements don't allow you to make the connection because the dynamics of both processes is missing
- The dynamic balances are between synthesis and degradation processes with many intermediate regulated steps.
  - Most published proteomics studies are comparative steady-state comparisons of relative abundance
  - A decrease in the amount of a protein may be due to less mRNA production, slower processing, slower translation on the ribosome by initiation or elongation, enhanced mRNA degradation or enhanced protein degradation. → the observation does not give information about the mechanism
    - Synthesis is linked more closely to the transcriptome and degradation to the metabolome

Protein Turnover kinetics

- The kinetics of protein metabolism combines a concentration of mRNA, a rate of ribosome initiation and a kinetic factor for ribosome activity.
  - Translation is not sensitive to protein concentration (unless it is a self-regulator)
- Degradation is a first-order process, whose rate constant (half-life) is a fractional value, e.g. 0.1 h^{-1}, or 10% per hour.
  - Flux through the pathway is related to the concentration, or pool, in the steady state balanced by the flux coming in through synthesis
    - $dP/dt = [k_s - k_d][P]$
    - If the rates are equal then the system is at steady-state and the rate of change is zero and $[P]=k_s/k_d$
    - If you can measure [P] quantitatively, and if you can find ways to measure one of the other parameters then you can resolve the system.

Measuring Turnover

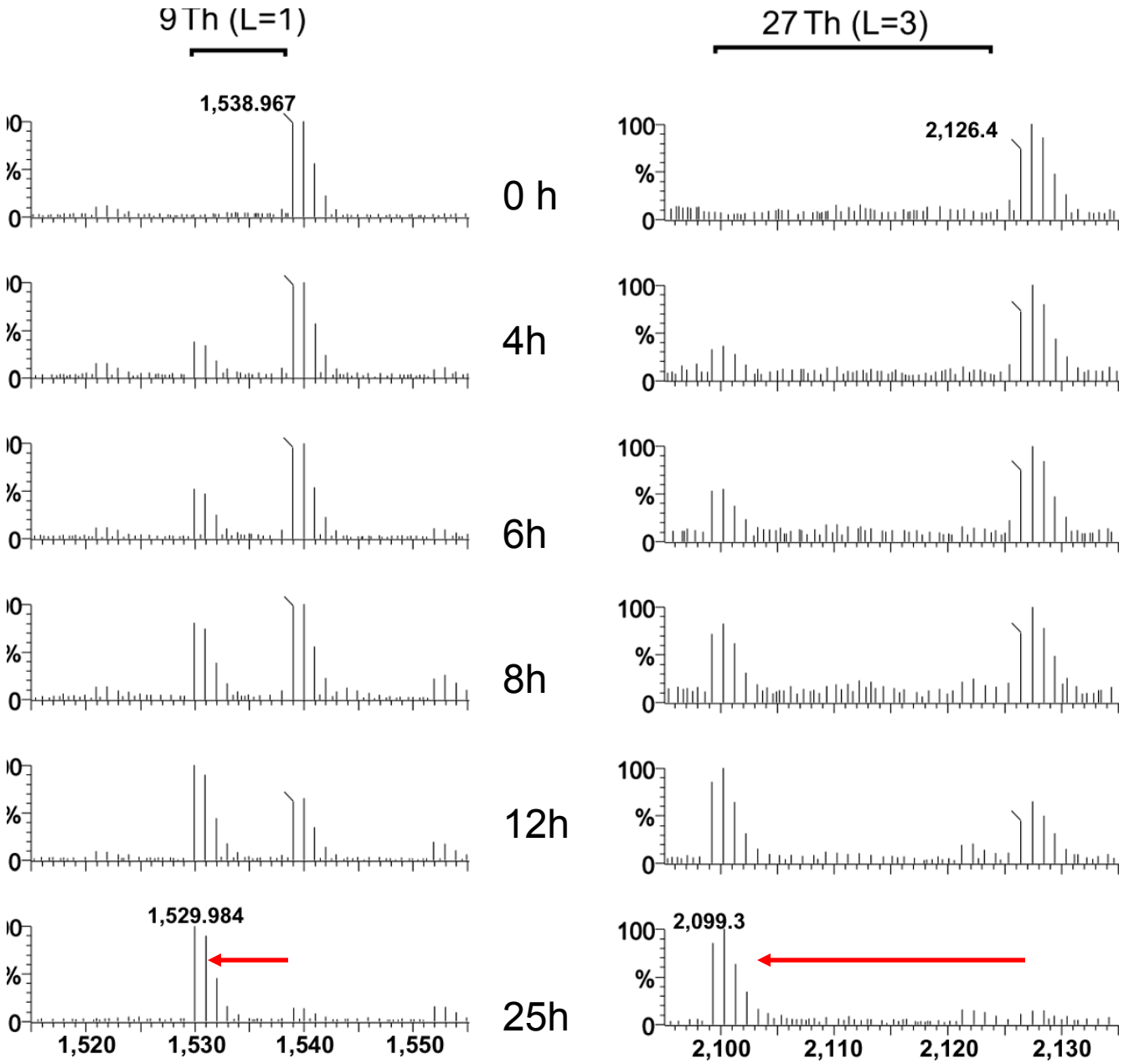
- Measuring Protein turnover requires synthesis and degradation information
  - A protein pool may not change, but the input and output might change in tandem
- Metabolic tracers are used – classically radioisotopes were used.[14C], [32P],[3H] as inputs cause all amino acids and hence all proteins to be radiolabeled
  - Separation of individual proteins then requires gel electrophoresis and/or immunoprecipitation
  - Radioisotopes are sensitive and do not affect enzyme activities significantly
  - Stable isotopes are safer to work with [18O], [15N],[13C], in amino acids, then you use the difference in mass as the signal.
    - SILAC : stable isotope labeling with amino acids in cell culture
      - Cells are fully labeled with amino acids carrying the stable isotopes
      - With each generation after replacement of the amino acid pool with stable isotope, then at least half of the proteins will be labeled.
        - » If there is high turnover within the cell generation then more than half of that protein will carry the label
      - A mass spectrometer is used to analyze the proteins and compare the mass of two peptides, these methods are not sensitive to a few percent, like radioisotopes
      - Because cells are fully labeled turnover mechanism is not discerned – to get turnover rates partial labeling must be used.

Pulse-Chase experiments

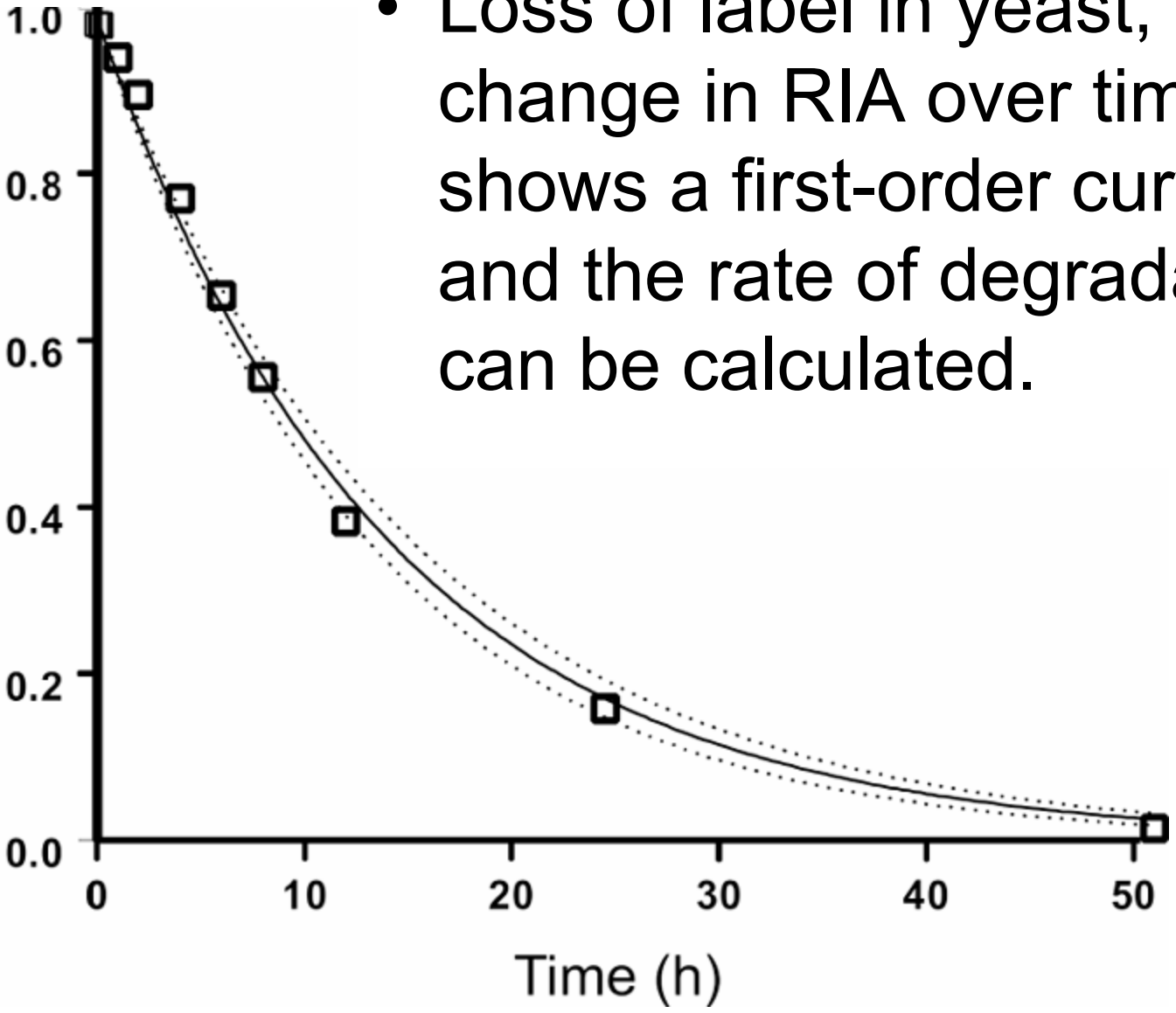
- If you use stable-isotope labeled precursors, the amino acids will not get homogeneous amounts of label because the different amino acids have different amounts of N, O and C.
- Labeled amino acids can be produced to carry one modification each
  - The site is picked for minimal turnover in the metabolome but reasonable representation in the proteins (Arg on the guanidino group is rapidly exchanged, for example)
    - The precursor pool needs to be converted as much as possible to the variant isotope and then chase with standard material so that recycling is unlikely.

Example Yeast experiment

- Single-cell culture systems
  - Start in medium with no label, then change to medium in which all of one type of amino acid carries the alternate isotope.
  - Start in medium with the alternate isotope and transfer to unlabeled medium
  - The precursor RIA (relative isotope abundance) needs to change between 0 and 1 very rapidly, and you use an excess of the targeted amino acid in the chase phase (so usually you chase with unlabeled amino acid).
  - For example, grow a yeast strain that is a Leu auxotroph in a chemostat (always in log phase)
    - It could not synthesize its own Leu and so dilute the pool
    - Provide it with $^{2}H_{10}$-Leu in the medium, waited for 7 doublings, replaced the medium with regular Leu.
      - Remove samples at specified times, purify the proteins and resolve on 2-D gels
      - Spots were recovered, proteolyzed and subjected to MALDI-ToF MS
      - Transition rates for individual proteins could then be approximated by the loss of heavy isotope from peptides.



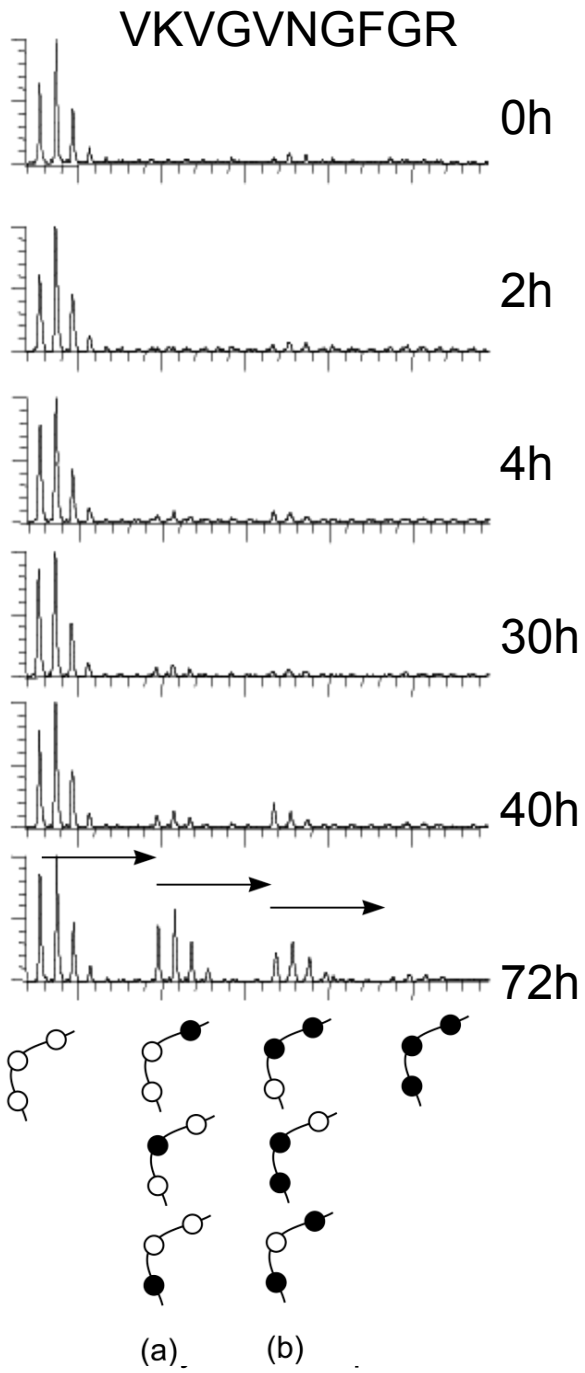
L=1 means the peptide has 1 Leu. Th=thompson, a unit of mass/charge



- Loss of label in yeast, using change in RIA over time shows a first-order curve, and the rate of degradation can be calculated.

Proteome dynamics in multi-cell experiments

- You cannot blitz the amino acid input pool of an intact animal.
  - How do you provide an RIA = 1 to tissue in an organism, for the necessary period of time (e.g. turnover of muscle in a mouse is 10% per day).
  - Usually a dietary supplement is provided, and even then the isotope is only part of the input.
  - An experiment with chickens used supplemented meal
    - Constant food ingestion through the day and continuous absorption at a fairly constant rate – the RIA was estimated to be 0.5, using labeled Val as crystals in the meal, which included some unlabeled Val as part of the protein.
      - You can see that when the RIA is only 0.5, for each Val there is a 50-50 chance of incorporating the H version, so the overall pattern becomes much more complicated.
        - » For VVV you can have LLL, HHH, LLH, LHH, and for the two mixed isomers there are three positions where the modified Val can be.



This is a trivaline peptide – the intermediate isotopic variants can be seen. The intermediate isotopomers a and b can be used to calculate the functional RIA of the precursor pool. This then lets you obtain the rate of synthesis.

mRNA Turnover

- mRNA synthesis is dictated primarily by the CREs that we have been studying
- The degradation of mRNA plays an important role in modulating gene expression
  - There is an additional role in making sure new-made mRNA is correctly synthesized
- The process is controlled by enzymes, with a variety of specificities and control mechanisms
  - Classes of mRNAs are controlled by particular enzymes, and translation factors can affect the sensitivity of a transcript to degradation
  - Ribonucleprotein (RNP) complexes, which perform functions like splicing, also recruit particular nucleases via specific mRNA-binding proteins.

Roles of mRNA turnover

- mRNA turnover plays three physiological roles
  - Sets the basal level of gene expression and provides a site for regulatory responses
  - mRNA decay pathways are a primary route for antiviral defenses
    - Regulatory sequence mechanisms or RNAi mechanisms
  - Specialized systems exist to rapidly recognize and degrade aberrant mRNAs

Turnover pathways

- Two principle pathways of turnover
  - 3' processive degradation from a polyA tail
    - Specific deadenylases perform this activity
    - The exosome complex carries out 3'→5' exonuclease digestion
      - The pppG cap remains and is hydrolyzed by a scavenger enzyme DcpP
  - 5' decapping so that 5'→3' exonucleases can function (such as Xrn1p)
    - The enzyme has two subuits: Dcp1p and Dcp2p
- There are also mRNA-specific degradation pathways involving endonucleases, triggered by miRNA or siRNA sequences.

Turnover Mechanisms

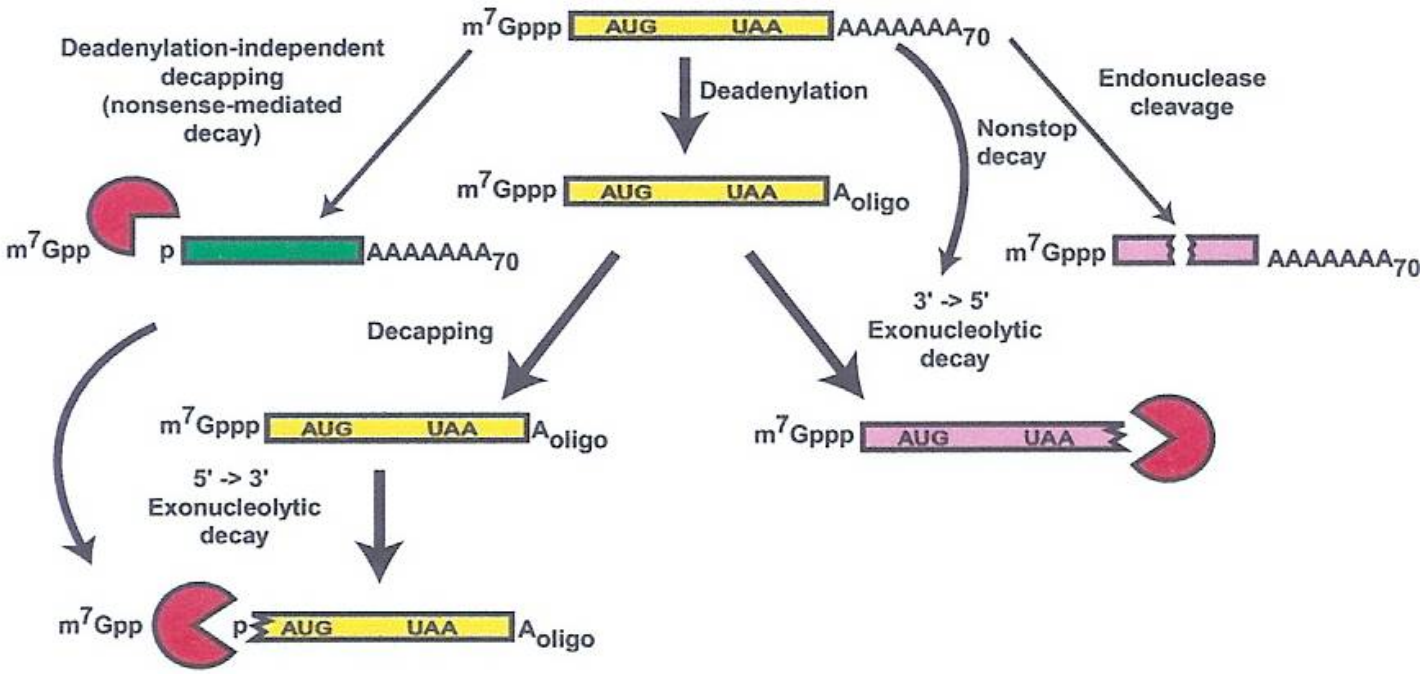


Figure 1 Pathways of eukaryotic mRNA turnover.

Turnover as error checking

- Quality-control steps
  - Pre-mRNAs with incompletely excised introns, UTRs and other signals that indicate incomplete nuclear processing are degraded
  - An internal stop codon triggers the decapping enzymes, or undergoes extremely fast deadenylation/exonuclease degradation, as are mRNAs with no stop codon at all.

Control of deadenylation

- Three complexes are known
  - In yeast the dominant complex contains two nucleases and ~5 accessory proteins
    - The nuclease Ccr4p is highly conserved across large phylogenetic distances
      - An ExoIII family nuclease
    - The Pop2p nuclease is an RNAse D enzyme (also conserved, uses a metal ion)
      - This has its own deadenylation activity and also enhances the Ccr4p activity – stabilizing the complex
  - A complex that deadenlates cytoplasmic mRNA contains two maim proteins, Pan2p (RNAase D also) and Pan3p
    - The main job is to shorten polyA tails (from hundreds down to 55-75nt)
    - Under certain conditions this seems to be the main deadenylase (RAD5 replicative stress)
    - Pan3p regulates the nuclease, binding Dun1p, which binds to specific mRNAs, and also Mex67p, which is involved in shuttling mRNA during transport from the nucleus to the RNP
  - PolyA RiboNuclease (PARN) is an exonuclease, also part of the RNAase D superfamily
    - It has a domain related to single-stranded na binding proteins
      - This is the major deadenylase in mammalian cells, homologs are present in most eukaryotic genomes, but is not present in yeast or flies.
      - This is the major mechanism for nonsense-mediated degradation and AU-rich instability recognition

Deadenylases

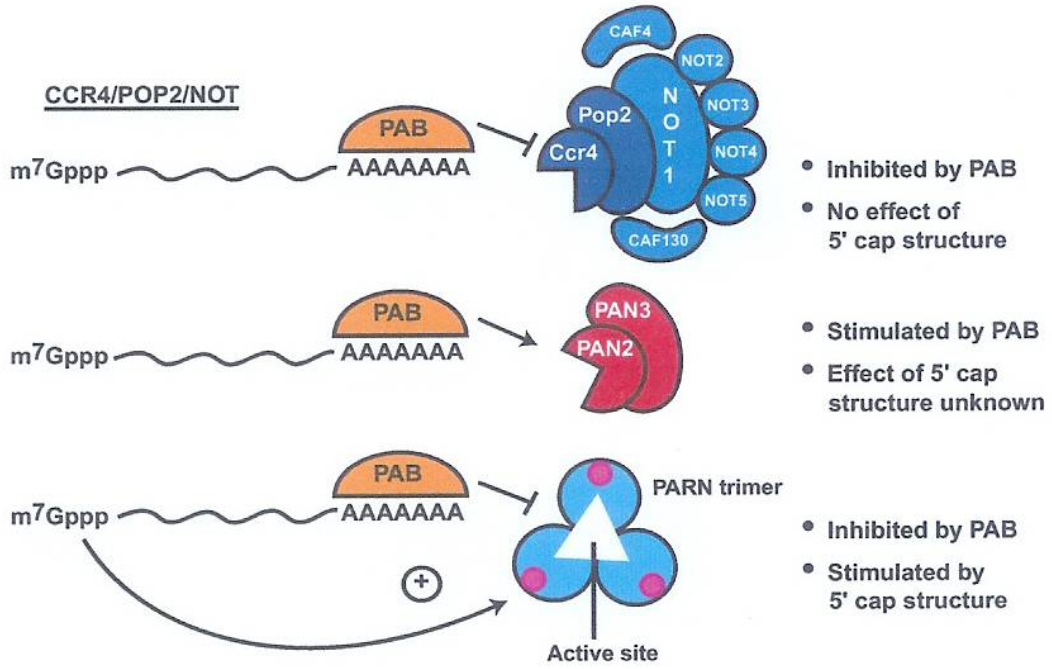


Figure 2 Multiple eukaryotic mRNA deadenylases.

Kinetic measurements

- How do you measure turnover?
  - Pulse-label with α-P32-ATP, remove samples over specific time intercals, purify the RNA and then
    - How long does it take label to appear in mRNA
    - Under what conditions does the label decrease
    - Recognition of the cap structure is from antibody staining
    - Recognition of specific mRNAs is by hybridization to a complementary probe
      - You can capture mRNA using probes complementary to 5', middle and 3' ends
      - Then transform to cDNA, clone and sequence
      - Perform nested PCR experiments
  - Do knock-out experiments on different genes to see if rate, specificity of response changes
    - For example, there is a polyA binding protein (Pab1p) that inhibits degradation by Ccr4p but stimulates PAN.
  - On purified enzyme you can measure rates of activity using classical enzyme kinetic biochemistry experiments
    - Rate of hydrolysis (release of isotope) given certain highly-defined substrates
      - Different length of polyA tail
      - With and without cap
      - With and without various cofactors and binding proteins

mRNP substrate specificities

- Each mRNA deadenylase has a preferred mRNP substrate
- The flux between different mRNP states alters the susceptibility of mRNAs to different deadenylases – there is global control of mRNA half-life exerted by controlling mRNP composition dynamics
  - Ex. PAN rapidly degrades mRNA that has Pab1p bound to the tail,, while PARN rapidly degrades mRNA with an exposed cap and a polyA tail that lacks Pab1p.
  - Thus mRNA-specific binding proteins can displace Pab1p and increase the availability of the mRNA to the Ccr4p or PARN complexes (leading to decreased translation), or may specifically bind to the complex and so increase the deadenylation.

Exosomes

- Exosome: After deadenylation, a large complex of nucleases degrades the mRNA
  - There are 9 core subunits, forming a ring-like structure with 6 RNAasePH domains and extensions with S1 and KH RNAase domains.
    - Other associated proteins have activities such as RNAase II hydrolase that degrade all the way to mononucleosides
- Cytoplasmic exosome interacts with a heterotrimeric complex (Ski2,3,8p), related to RNA helicases, the interaction to the complex is mediated by Ski7p.
  - There may be direct recruitment of an mRNA to the exosome by a protein, such as TTP and KuR, which bind to the ARE element in the 3'UTR and recruit the mRNA to the exosome
  - Without a stop codon, the mRNA stalls the ribosome, and the Ski7p protein acts like a translation factor, releasing it from the ribosome but also directing it to the exosome.

Exosome structure models

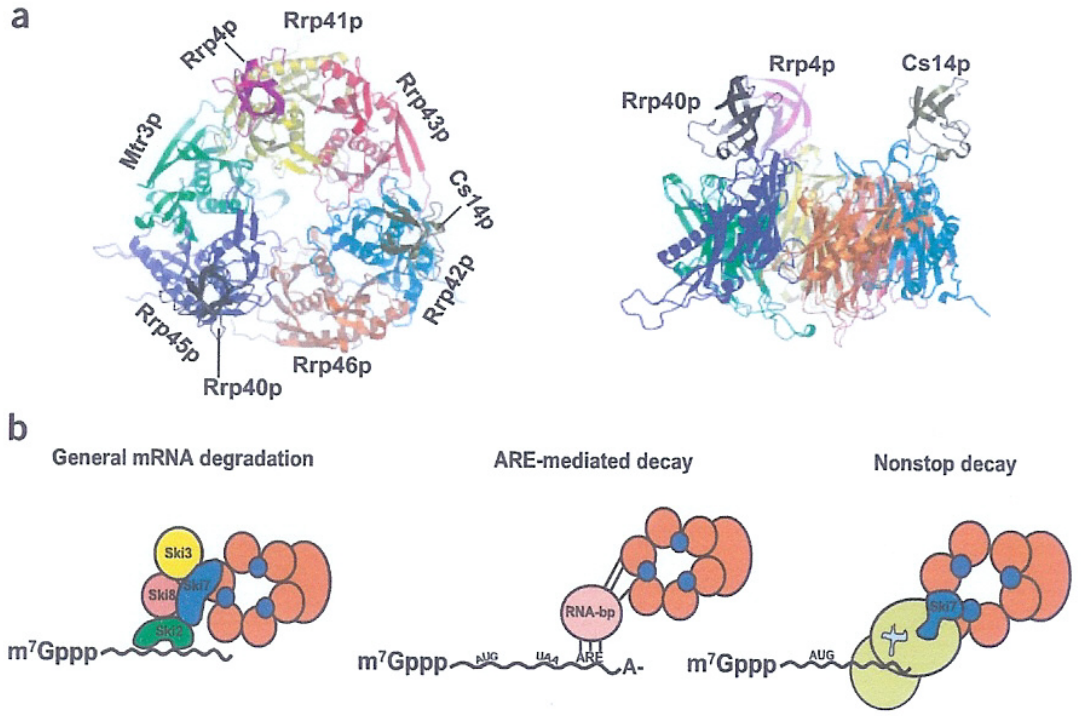


Figure 3 Models for exosome structure and function. (**a**) A hypothesized structure for the exosome. (**b**) Possible mechanisms of exosome recruitment to different mRNA substrates.

Decapping

- Decapping – 2 principle enzymes
  - Scavenger decapping(DcpS) decaps short oligonucleotides only, to release m^7GMP.
    - This is on the product of the 3'→5' exonuclease exosome activity, and it coprecipitates with the exosome
    - It also hydrolyzes m^7GDP → m^7GMP
  - For long polymers, after deadenylation and decapping by Dcp1p(regulatory, may be bi-functional)+Dcp2p(catalytic), a 5'→ 3' exonuclease can act
    - The resulting monoP substrate is acted on by Xrn1p

Decapping control

- Control of decapping has three inputs
  - polyA tail
    - Pab1p binding inhibits decapping
      - It has a physical interaction with eIF4G (part of the cap-binding complex) which may prevent activity or may enhance translation
    - Other types of regulation must be in place since lack of a stop codong but presence of polyA tail does not prevent decapping
  - Translational status (decapping and translation are in competition)
    - Recruitment of translation initiation factors is important and eIF4E,G are cap-complex binding factors that in the translational RNP inhibit decapping in vitro
      - There is a rearrangement so that the mRNA binds Lsm-1-7p (part of the RNA-binding complex called Pat1p) and not these IFs or Pab1p.
    - After translation the mRNA enters a P body, complexes with the decapping factors and exonucleases.
      - You can hang up mRNAs in the P-bodies by inserting strong secondary structure or deleting Xrn1p activity, and these end up in P bodies.
  - Rate of recruitment/assembly of the decapping complex to the mRNA (there is competition for the complex)
    - Certain proteins (Upfp1) have activity that involves recruiting the mRNA to the decapping complex.
    - The Rps28a mRNA has a strong loop in the 3'UTR binds its own product and then a second protein which then recruits the subunits of the decapping complex.

Decapping enzymes

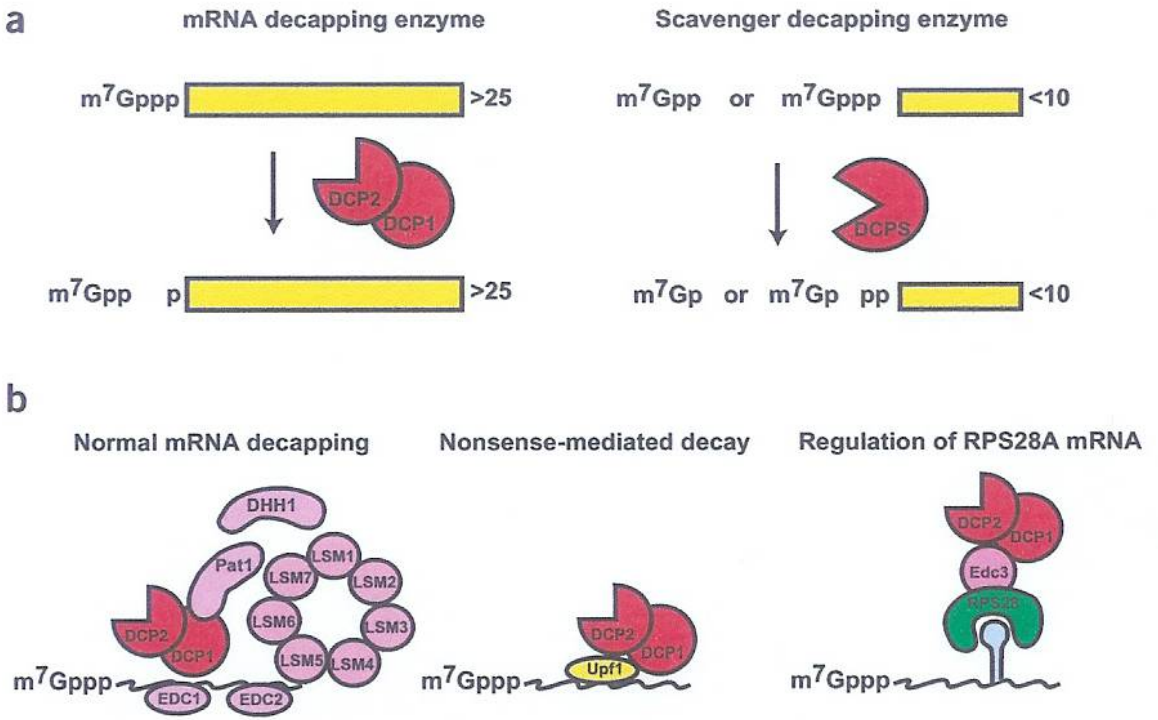


Figure 4 Eukaryotic decapping enzymes. (**a**) mRNA decapping enzymes and their substrates. (**b**) Possible decapping complexes recruiting Dcp1–Dcp2p to different mRNA substrates.

Regulation of exonuclease

- Exonuclease regulation
  - Xrn1p in the cytoplasm, and Rat1p in the nuclease have high rate of activity
    - Mutations in eIF5A block degradation after decapping, so regulation is possible
    - The Xrn1p of Drosophila, Pacman, is developmentally regulated.
- The larger scenario is that mRNA degradation is one part of a system that uses export, localization and translation as mechanisms for fine-tuning RNA physiology.
  - Mutations in proteins whose phenotype affects secretion, cytoskeletal dynamics and translation all affect mRNA decay.

Davidson – chapter 2 cont

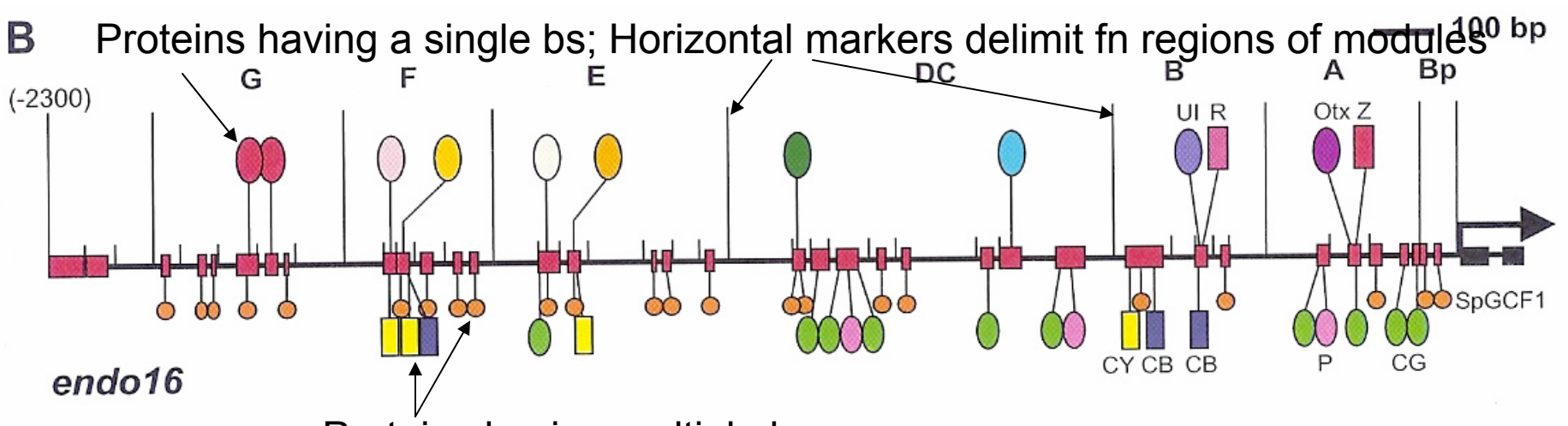
- Revisit the endo-16 control
- Discuss the circuit diagram representation and NetBuilder diagram conventions

- Processing functions are encoded in the DNA and underlie spatial control of expression, using 2 classes of input
  - Those that vary by time/location, the active TFs
    - The spatial cues are generally signalling ligands produced by other cells.
  - Those that modulate the first set but can be at constant concentration – these are 'black box' effectors.
    - To understand how these work you must perform kinetic studies – single time points will not be informative

Endo-16

- In the embryo there are about 50 genes expressed in the endoderm network, out of ~8500 expressed in total.
- Endo-16 is a large polyfunctional protein of sea urchins. It is a secreted protein found in the endoderm of early embryo and the inner wall of the midgut of the late embryo.
  - In the early embryo it is expressed in two types of cells, but there is a ring of non-expression between them (the polar patch).
  - There are two different subregions governing the expression, services by nine tfs and 54 regulatory proteins.
  - Sites are distributed in 6 CREs (G-A).
- There are repressive interactions mediated by the domains DC, E, F
  - E and F are not found in all individuals
- There are enhancer regions, G, B, A
  - A is mediated by the tf Otx, which acts as its driver
    - A causes earlier activation and is the core processing unit for the other CREs
  - The spatio-temporal activation of B for late expression is driven by the tf Brn1/2/4

Endo-16: Necessary and sufficient segment of DNA to govern correct expression in an egg



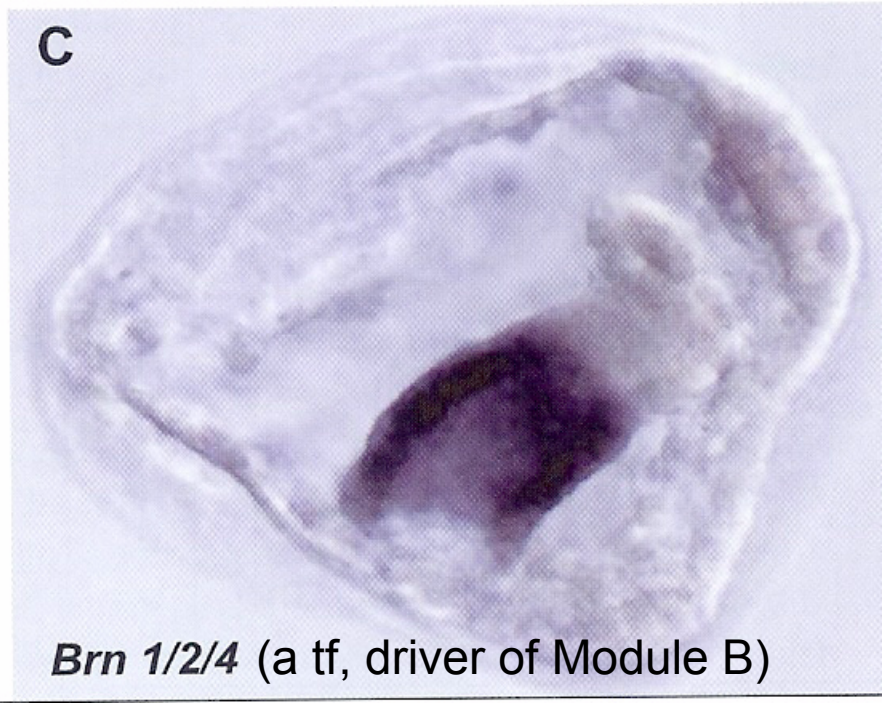
G Positive booster

F
E } Repression in adjacent ectoderm

DC Repression in skeletogenic mesenchyme

B Expression in midgut of late embryo
- Controls late rise in expression
- Activates switch resulting in exclusive use of its own input

A Expression in vegetal plate in early embryo
- Sole communication to BTA for whole system
- Synergistic amplification of B input
- Transduction of FE, DC repression



Brn 1/2/4 (a tf, driver of Module B)

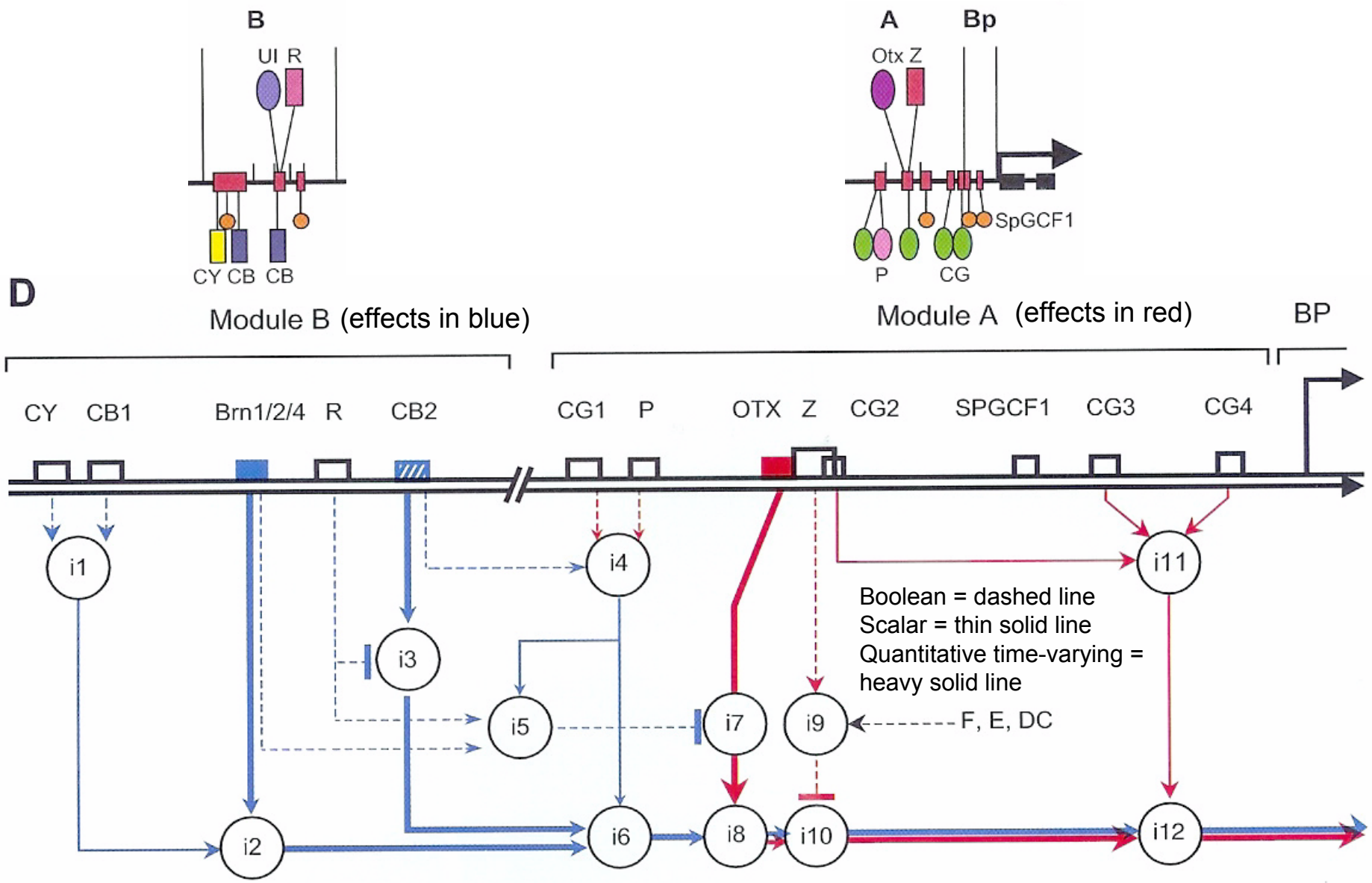
Describing the endo-16 Logic Circuit

- The basal promoter (Bp) is where the RNA polymerase and associated factors assemble in order to make the transcript.
  - All inputs to the Bp are channeled through Module A, which determines whether or not transcription will occur and the rate, when it does.
    - Mutations at specific sites in Module A have been shown to block all upstream effects from reaching the Bp.
- Repressive Modules: these are DC, E and F
  - They are known to be essential since certain regions must have a lack of the product – these constitute negative inputs to Module A
    - The tf Otx is present and active in most cells in early development, and is a positive driver of Module A.
      - It is needed for many functions → it cannot be repressed (ie the concentration modulated) – another type of preventative action must be used.
    - The inner boundary requires repression with DC and the outer boundary interactions of A with E, F
      - The F repression factor is a Creb, or signaling protein

Endo-16 Effects

- Positive inputs to Module A
  - Later in development it is not necessary to modulate Otx, the repressor modules are not needed and all effects are due to modules A+B.
    - A new tf, Brn1/2/4 is the main driver at the later stage, which binds to a site in Module B.
  - Two new interactions must occur between modules A and B:
    - A switch sensitive to the concentration of Brn1/2/4 must be present, which is the R regulator in B
      - Brn1/2/4 is activated after the gut forms (pos gastrulation). The switch turns off the input to the Otx regulator of A → only inputs to B will be effective
    - Amplification in A of the B input, a set-up function is needed.
      - Several proteins (having specific binding sites) are needed: CB2 in B and CG1 and P in A.

The Logic Circuit Diagram



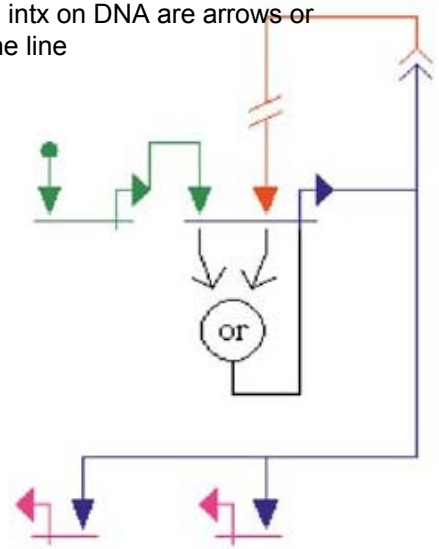
Time-varying driver intx (colored box), time-and-space invariant intx (open boxes)

Kinetic vs logical values

- The functions are conditional on the inputs, including linear amplifications, non-linear effects like switches and power functions (cooperativity or Hill coefficients), detection of input thresholds as well as logic operations.

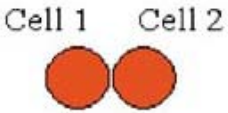
NetBuilder

Interactions outside of the nucleus are shown above genes and CRE intx on DNA are arrows or bars that impinge on the line



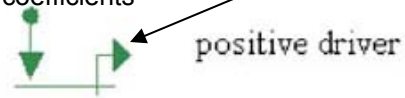
Interaction symbols are AND, OR, NOT, or addition and multiplication. NOT might mean repression, change in sign or reciprocal of input

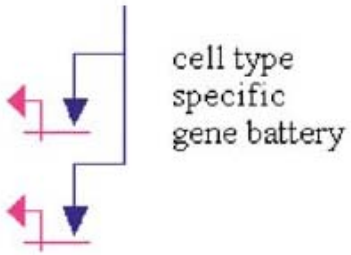
Switch elements include a bistable switch; threshold symbols and connectors that indicate outcomes.

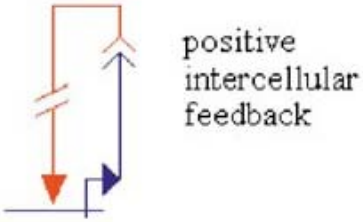


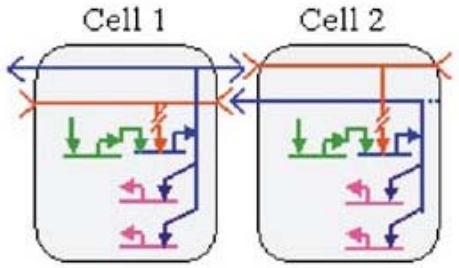
Scalars (include linear amplification and hill coefficients

Gene demarcation









if CY & CB1	i1 = 1
else	i1 = 0.5
	i2 = i1 ï Brn124(t)
if R	i3 = CB2(t)
else	i3 = k ï CB2(t) (1<k<2)
if P & CG1 & CB2	i4 = 2
else	i4 = 0
if Brn124(t)>threshold & R & i4 á 0	i5 = 1
else	i5 = 0
	i6 = i4 ï (i2+i3)
if i5 = 0	i7 = OTX(t)
else	i7 = 0
	i8 = i6 + i7
if (F or E or DC) & Z	i9 = 1
else	i9 = 0
if i9 = 1	i10 = 0
else	i10 = i8
if (CG2 & CG3 & CG4)	i11 = 2
else	i11 = 1
	i12 = i11 ï i10

- If CY (target site CY), and i≠0 (here it is 1) means the site is present and occupied.
- i=0 means that either the site was mutated, or the factor was inactivated so productive occupation of the site is not possible – this is 'else'
- CB1 and CY1 interactions together are the i1 circuit, and cause an increase in the output of the spatio-temporal regulator of Module B, which binds at the Brn1/2/4 site
- The output of the Brn1/2/4 module is at i2.
- There is another time-varying positive input generated by the interaction at CB2 (peaking at 40h).
- An interaction at R is required for the B-A intermodule switch, which shuts off the otx input (i5,i7); this operates only if there is input from Brn1/2/4 AND CB2 is present and occupied.

Wikipedia – the Diacritic and i

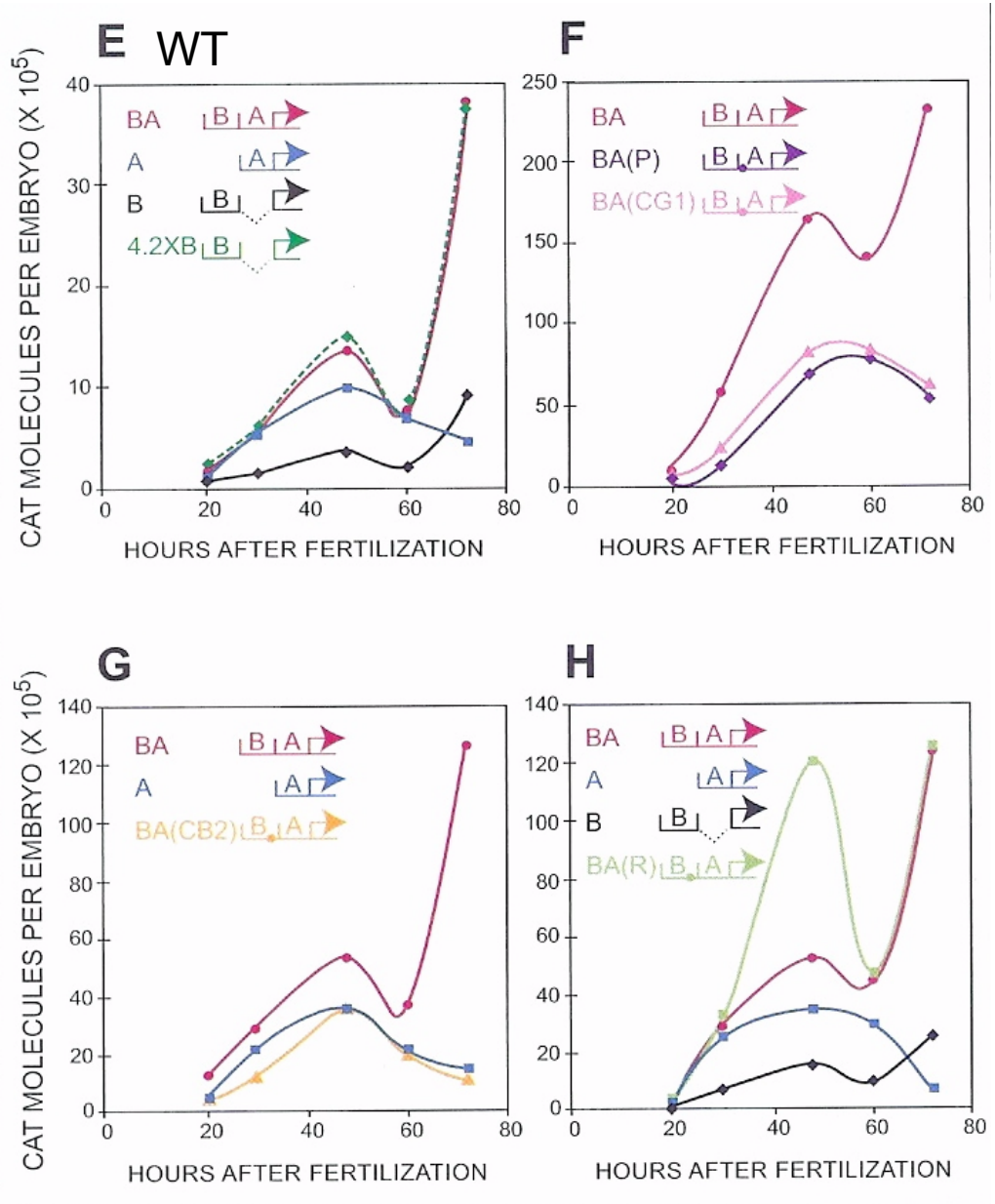
- The derivative with respect to time is often represented as a dot above a variable. Two dots represents the second derivative.

$$\dot{a} = \frac{\mathrm{d}}{\mathrm{d}t}a \qquad\qquad \ddot{a} = \frac{\mathrm{d}^2}{\mathrm{d}t^2}a$$

- This may be contrasted with the more common notation for a derivative using a prime:

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}f(x) \qquad\qquad f''(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x)$$

- In physics, a dot typically represents a (partial) time derivative $\frac{\mathrm{d}}{\mathrm{d}t}$ while a prime represents a spatial derivative $\frac{\mathrm{d}}{\mathrm{d}x}$.



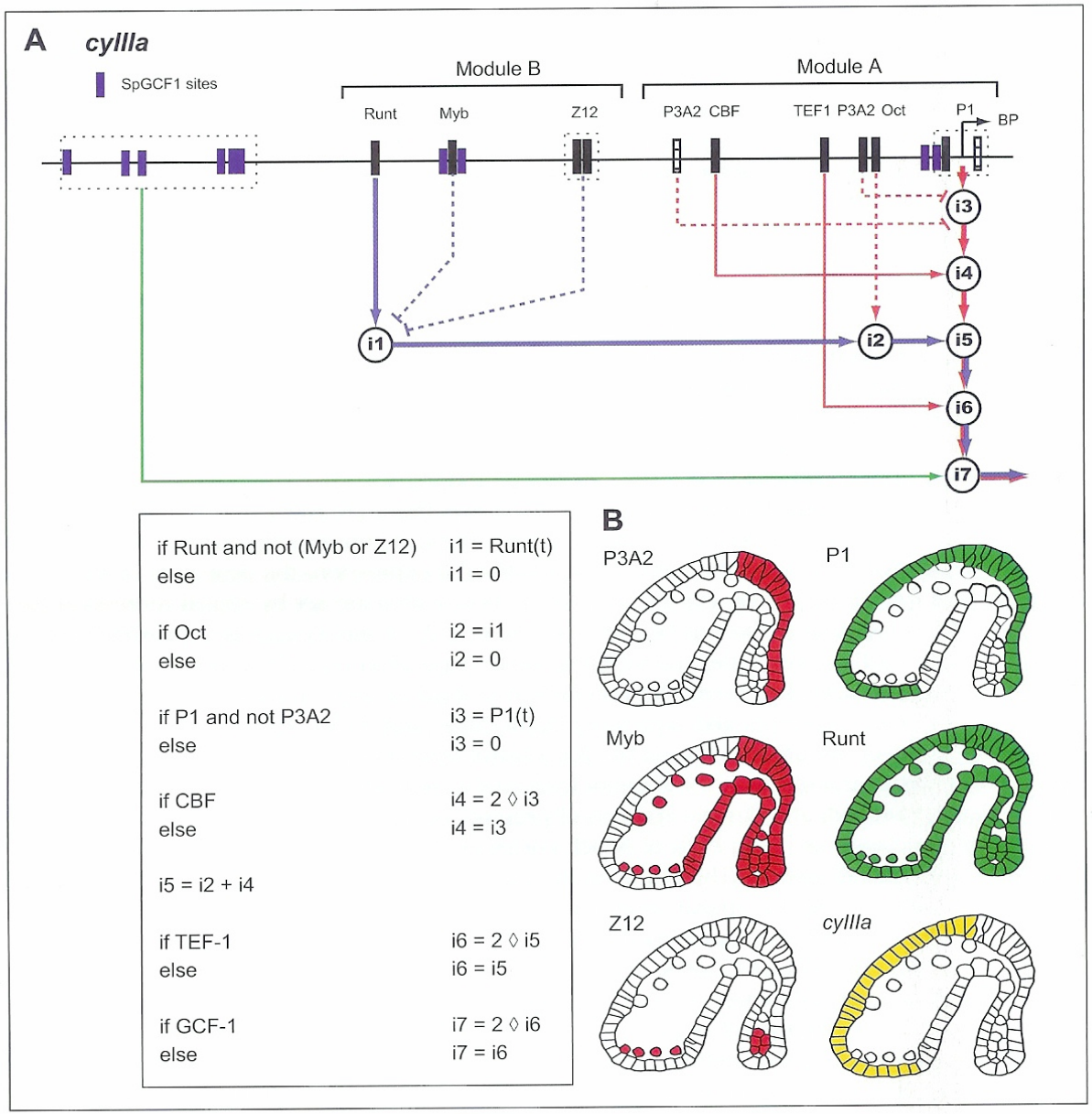
- How the logic is tested: kinetic experiments for induction and repression of the basal reporter gene chloramphenicol acyl transferase (CAT)
- The construct in its normal architecture, BA, is red
- B alone is black
- A alone is blue
- Targeted deletions/mutations are given in violet, with parentheses and dots

Factors vs Drivers

- In the endo-16 regulatory modules, only 2 of 9 proteins having specific binding sites function as 'drivers' of gene transcription, having spatio-temporal regulatory functions: Otx and Brn1/2/4.
  - Absolute concentration or activation/inactivation is used to fill the need for these drivers
- The other proteins process and modulate the effects of the two drivers, having repressor, switching and amplification effects.
    - Their action is still directly encoded by DNA binding sites.
    - These proteins can be ubiquitous and at generally high concentrations all the time: they can only act when the drivers are present anyway.

Computing and Genomic Regulatory Logic

- What would be needed to compute rather than experimentally test the logic circuits described?
  - First, truly be able to identify all target binding sites
    - These are not yet fully characterized for any species
  - Have available for interpretation the set of functions that occur given a particular pattern of site occupation and factors
    - There are closely related families of tfs that have considerable binding site overlap and functions that are not identical but the distinctions are not well understood
  - Have available the rules for combining the functions in order to infer the output – e.g. when amplification will occur and by what factor
    - Almost none of the combinatorial rules are yet understood
    - Some factors can have opposite effects given a new partner or co-factor
- So computational predictions are a goal, not a reality. In the meantime the approach of using characterized modules with saturated tf characterization allows dissection of complete functional effects and modulators
  - The differences between closely related CREs may start to reveal the rules we seek.


A
cyIIIa
SpGCF1 sites
Module B
Module A
Runt
Myb
Z12
P3A2 CBF
TEF1 P3A2 Oct
P1
BP
i1
i2
i3
i4
i5
i6
i7
if Runt and not (Myb or Z12) i1 = Runt(t)
else i1 = 0
if Oct i2 = i1
else i2 = 0
if P1 and not P3A2 i3 = P1(t)
else i3 = 0
if CBF i4 = 2 ◊ i3
else i4 = i3
i5 = i2 + i4
if TEF-1 i6 = 2 ◊ i5
else i6 = i5
if GCF-1 i7 = 2 ◊ i6
else i7 = i6
B
P3A2
P1
Myb
Runt
Z12
cyIIIa

Saturated Characterization of a CRE

- What is needed to completely characterize a CRE?
  - Mutational analysis of the binding sites for sequence and distance
  - Mutational analysis of transcription factor DNA- and protein- and co-factor-binding regions
  - Analysis of functional results
    - Transcriptional activation or repression (microinjection of antisense oligos, or mRNA not expressed at that time/space)
    - Intermodule linking
    - Amplification of input
    - Concentration dependence
    - Types of quantitation transformations
      - Scalar
      - Multiplier
      - Power
      - Boolean

Common Circuit Responses

- For linking between modules it is clear that Boolean logic most commonly applies:
  - AND meaning that a positive output only occurs in domains where two different regulatory factors are coincidentally bound – where the spatial distribution overlaps.
- Switch-like behavior is the most frequent response to signal transduction events
  - Intercellular ligands are used to modify a tf, for example if there is no signal ligand and the tf is present then it will act as a dominant repressor
  - If a ligand and a tf are present (and usually a co-activator as well) then the tf cannot repress the Bp and transcription will occur
  - The repression of transcription is usually Boolean and usually dominant.

Target site occupancy

- Transcription factors have a binding constant in the formation of the complex with the DNA binding site.
- This is a bimolecular chemical reaction and thus depends on the concentration of the two reactants
  - The more of a reactant is present the more you push the reaction toward the product
  - Thus occupancy is determined by the binding constant (an intrinsic property) and the nuclear concentration of the DNA sites (invariant) and the transcription factor (which can be modulated).
  - The binding equilibrium of the DNA-tf duplex may be modulated by additional interactions with adjacent proteins, expressed as cooperativity constants.

Measuring Binding Constants

- It is possible to label both the protein and DNA and measure the on- and off-rates of duplex formation
  - You can do competition experiments with labeled and unlabeled protein.
- It turns out that on-rates (the protein finding the recognition site and binding, which is diffusion driven) do not vary much over a wide range of tf classes (for a given concentration) but the off-rates vary by orders of magnitude.
- In terms of relative concentrations, in animal cells, there are hundreds of sites and thousands to tens of thousands of a given tf at a given time.
  - The tfs bind transiently to DNA for which no strong interaction occurs (very transient) mostly.
  - For positively acting factors the site must be occupied if you see initiation of transcription (which has an invariant rate once it has begun).
    - It is possible to measure transcript production rates, and processing and cytoplasmic transport are not rate-limiting, it is possible to measure protein production rates; if you know transcript and protein production and turnover rates you will know relative concentrations of proteins at any given time.

Cis-Regulatory Design

- In development a primary function of cis-regulatory information processing is to integrate diverse driver inputs so that the output is always unique – AND logic again.
  - Two factors must be present together – why?
    - The interaction may give cooperative binding (greater stability)
    - The interaction may give the complex the ability to bind a third element, a co-factor
    - Each factor may be correctly oriented to give needed input to the basal transcriptional apparatus (BTA)