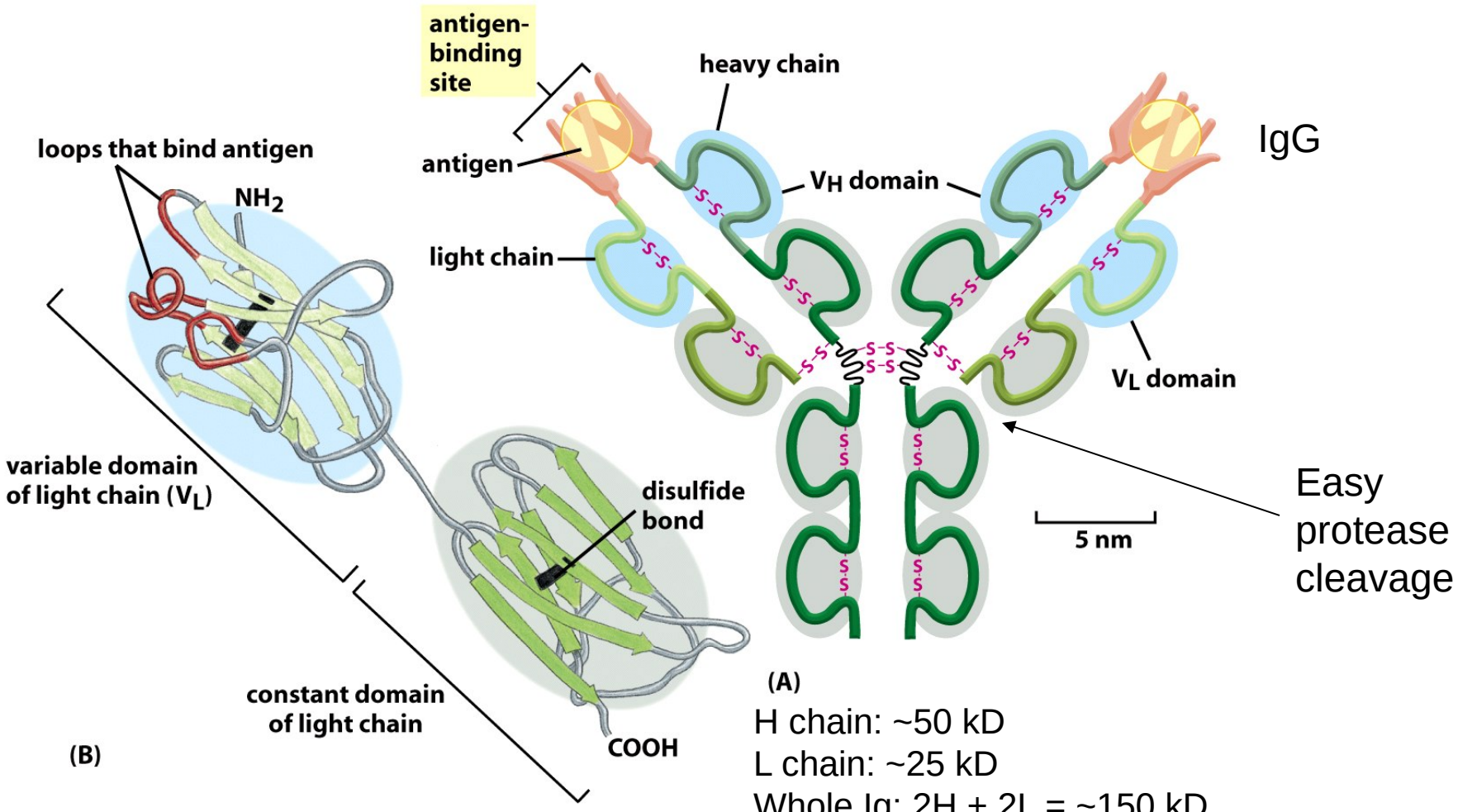
BI 8 LECTURE 8
EVOLUTION AND GENOME MANIPULATION; & BRIEF COMMENT ON QUANTITATING BIOLOGICAL REACTIONS

Ellen Rothenberg

28 January 2016

Reading: Ch. 3, pp. 109-144; panels 3-1, 3-2; Ch. 2, pp. 51-63
review Ch. 8 pp. 494-504; optional: Doudna/Charpentier

Domain structures order folding even along single polypeptide chains



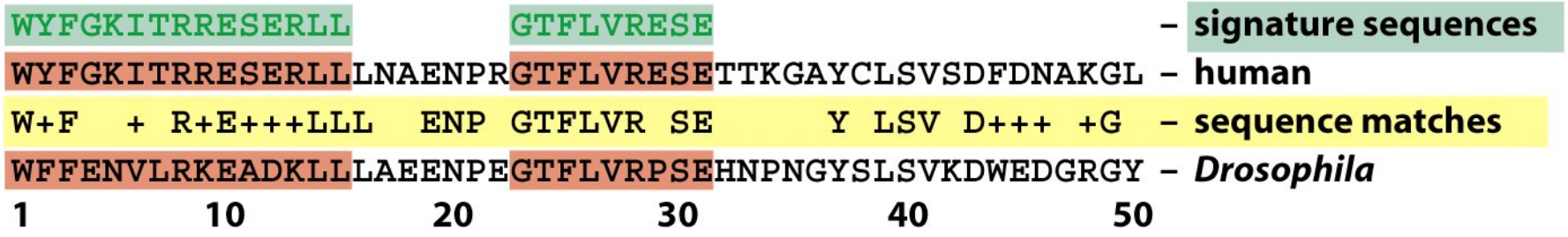
H chain: ~50 kD
L chain: ~25 kD
Whole Ig: 2H + 2L = ~150 kD
+ βME: 2 x 50 kD, 2 x 25 kD; none bind antigen
+ protease: 3 x ~50 kD, some bind antigen
+ both: all 25 kD pieces, none bind antigen

Figure 3-41 *Molecular Biology of the Cell* (© Garland Science 2008)

Related functions and structures can be predicted for proteins in distant species or among distant family members just based on sequences of genes that code for them…

… educated guessing even before starting biochemistry!

- Find possible exons predicted from continuous reading frames
- Virtual translation "*in silico*"
- Similarities can reveal matches to other known proteins in database



Human and Drosophila Src: SH2 (protein interaction) domain

Generally substantially the same folding even if only 20-25% sequence identity! (especially if conserved residues are known to be in key points of structure based on well-studied examples)

Figure 3-14 *Molecular Biology of the Cell* (© Garland Science 2008)

Deep evolutionary conservation of protein structure even when amino acid sequence has drifted (homeodomain proteins)

Conservation of ***particular*** amino acids that are needed for structural function in particular domains is more important than ***overall*** conservation

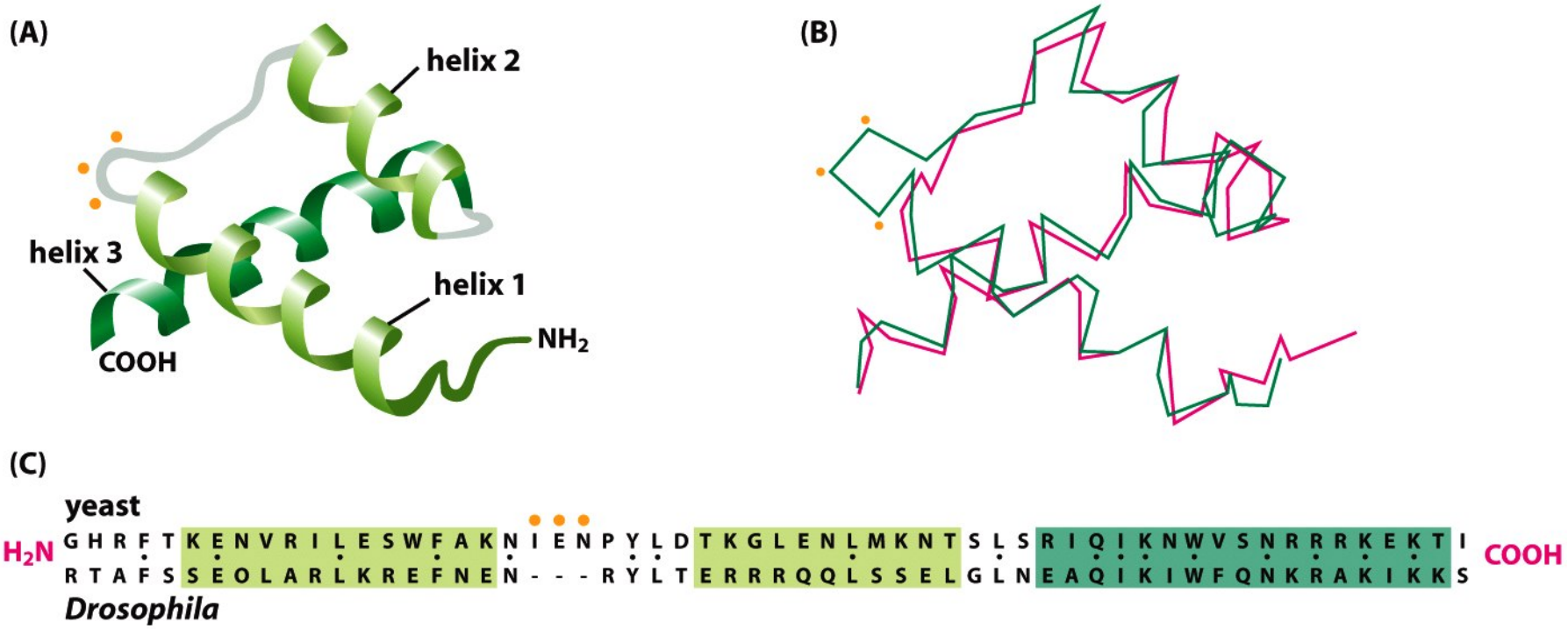
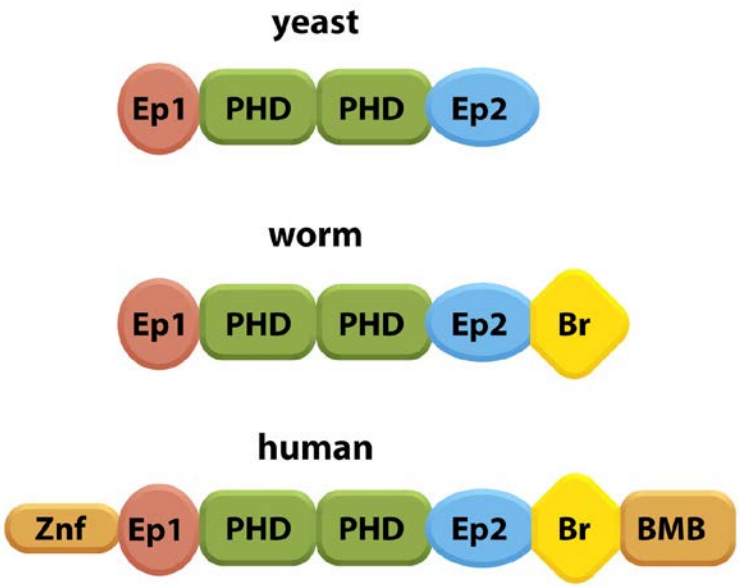
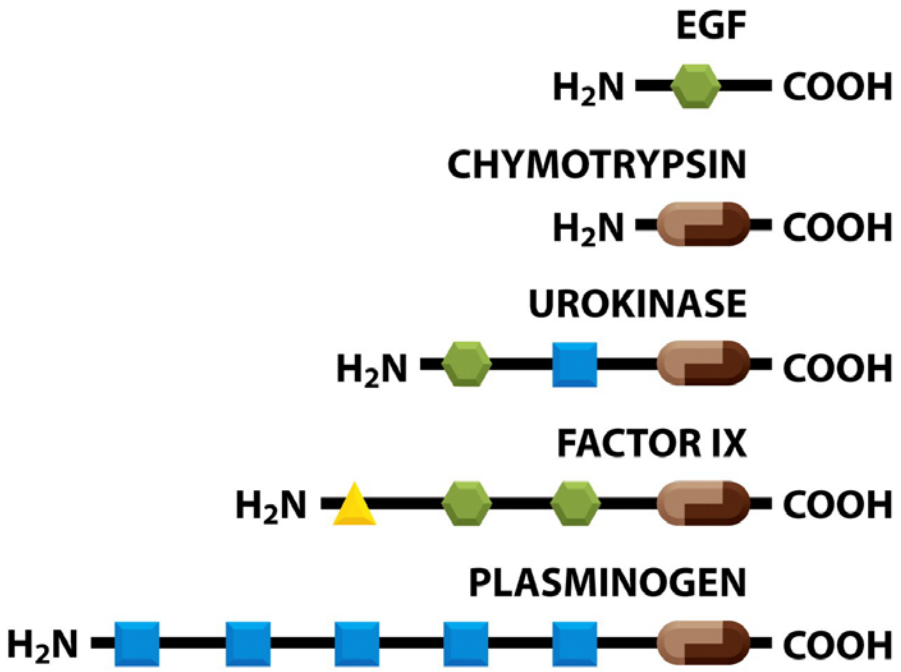


Figure 3-13 *Molecular Biology of the Cell* (© Garland Science 2008)

Domain based modules are the foundation of protein evolution: match, multiply, and mix

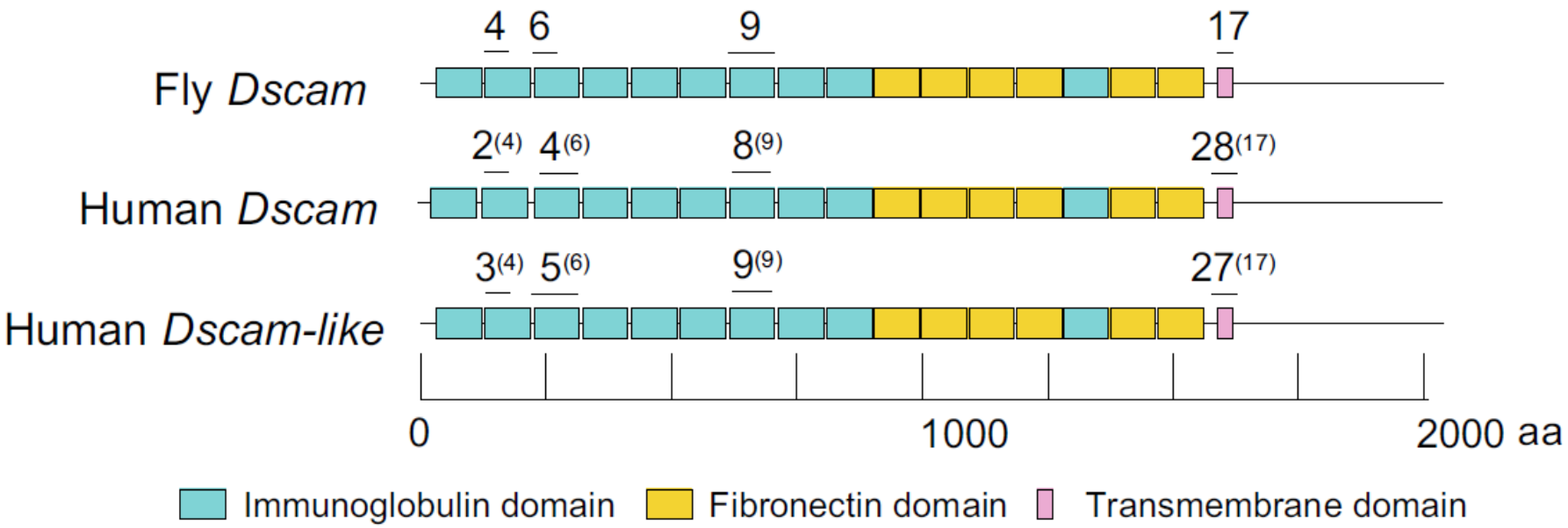




Modularity is easy to encode in genome when domains are encoded by discrete exons… introns give lots of room for copying,cutting & pasting

Figure 3-19 *Molecular Biology of the Cell* (© Garland Science 2008)

A fascinating story of evolution *and* alternative splicing: Down's syndrome cell adhesion molecule (Dscam)

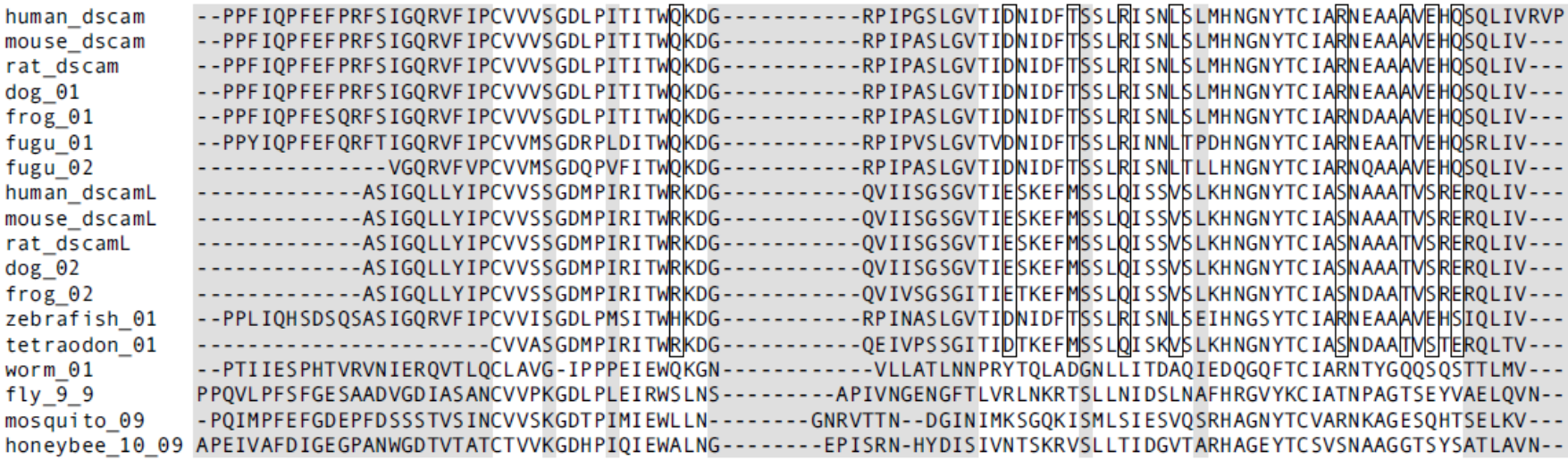


As in many cases, exon/ domain boundaries coincide… a useful mechanism for evolution

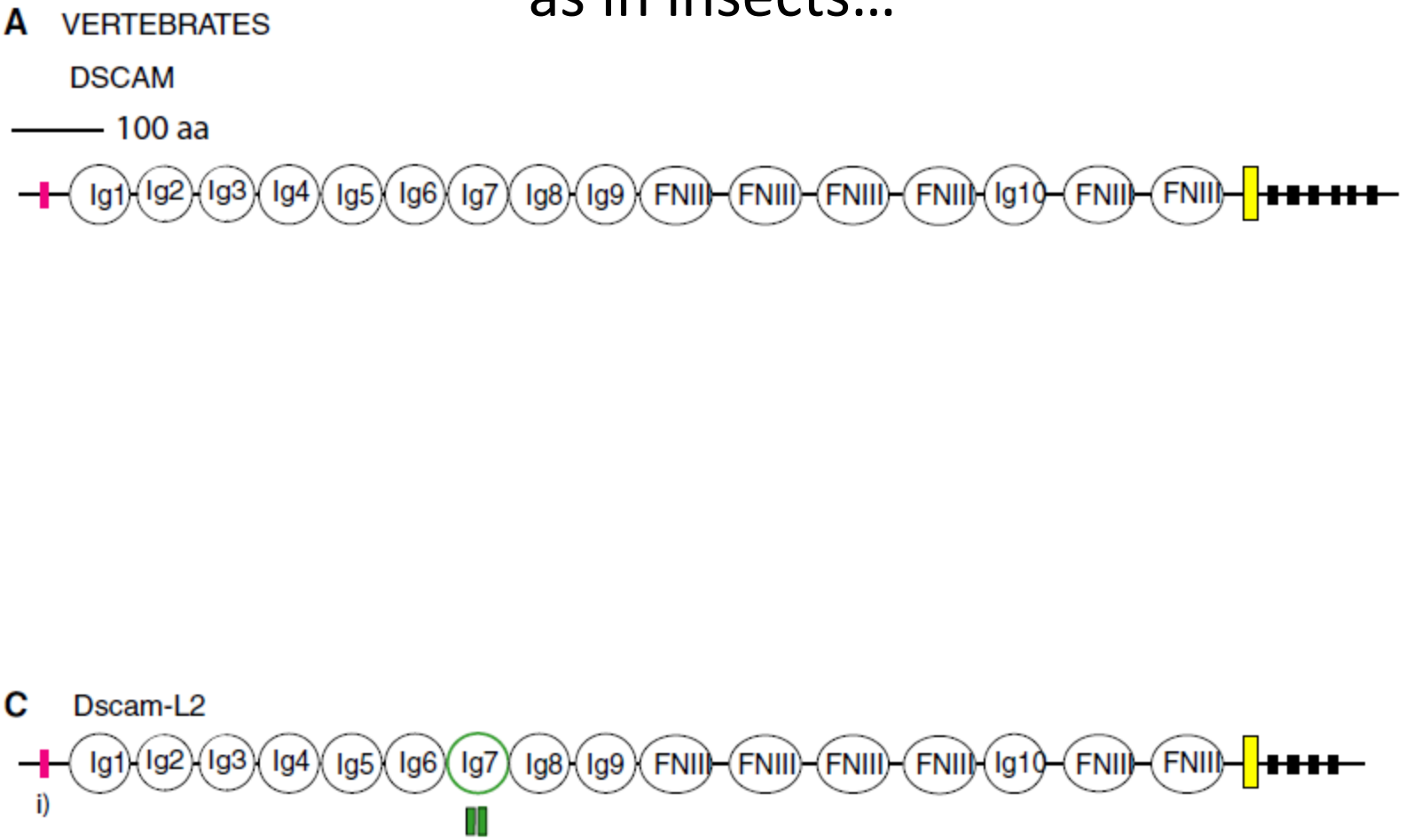
(Crayton et al. 2006 BMC Evolutionary Biology)

Conservation can be spotty… yet key residues are a clue to similar function in corresponding domains (Dscam)

(Crayton et al. 2006
BMC Evolutionary Biology)

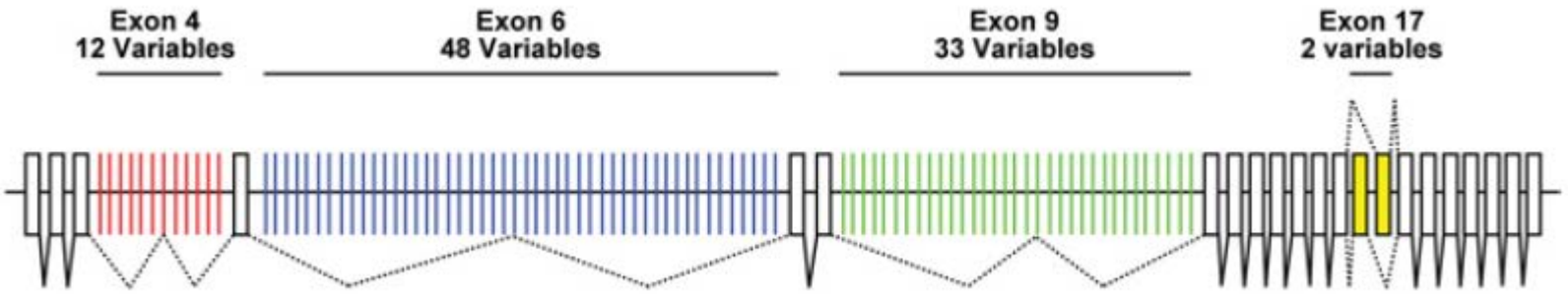


Yet the use of the exons is NOT the same in vertebrates as in insects…



(Brites et al. 2013 Evolution)

In insects – not vertebrates – RAMPANT mutually exclusive alternative splicing to choose each of four exons → ***38,016 potential isoforms!***



Thought to be used for immune recognition in insects
Diversity helps to create a wide repertoire of possible defense structures… but neuronal pathfinding needs diversity too

(Hemani, Soller 2012 Biochem Soc Transact)

Nucleic acid manipulation can be key to establishing function of proteins and protein domains

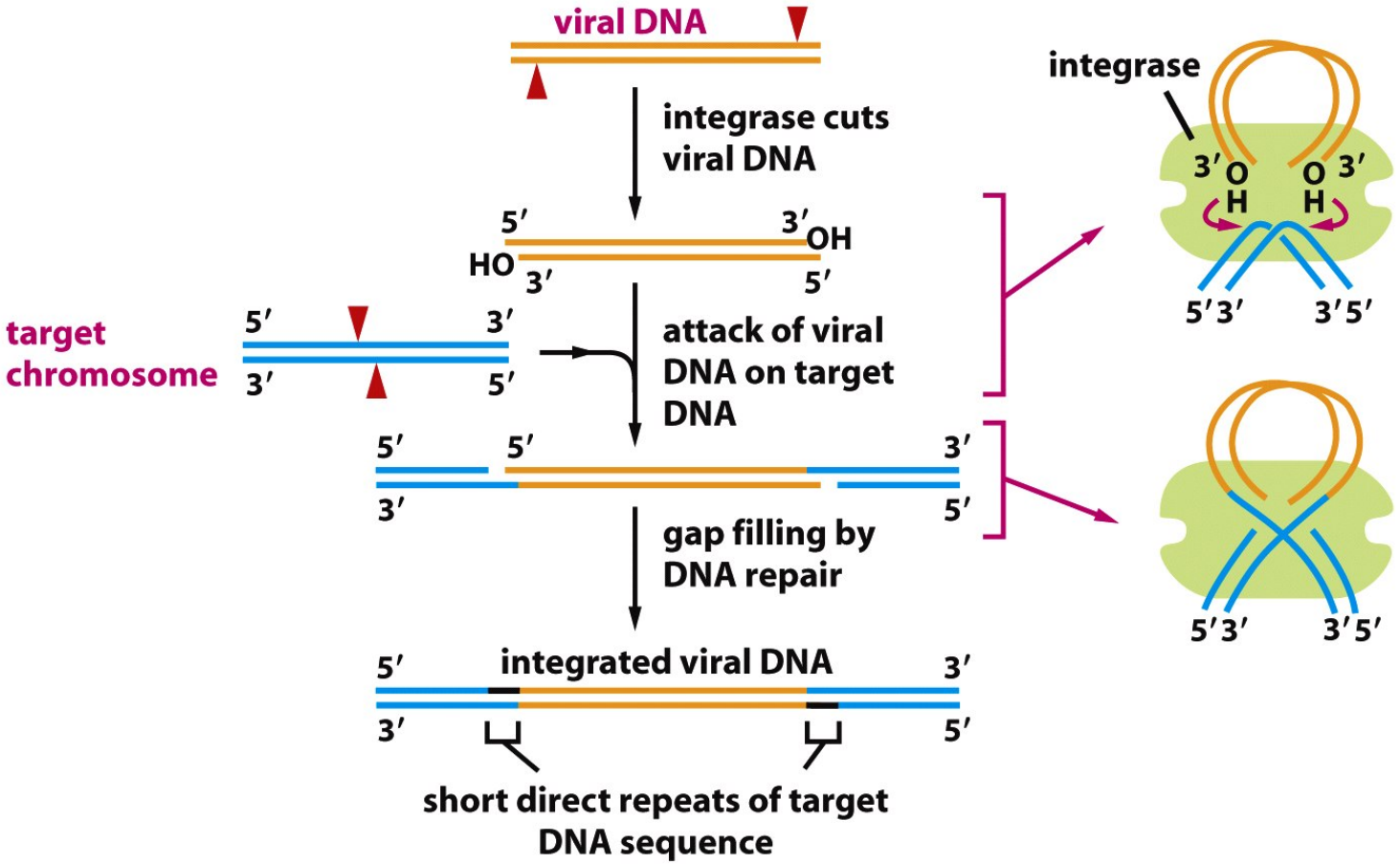
- DNA can be expressed into proteins of interest in vivo or in vitro
  - Transfection into cells in culture (cDNA or genomic)
  - Change in function or protein expression relative to background (hopefully low or zero!)
  - Also available: in vitro, coupled transcription/ translation systems: put in DNA (cDNA cloned in vector with promoter), get out protein
- If you have a functional assay, you can see effect of adding a protein on a system
- With a functional assay, compare effect of adding wildtype vs. protein coded by deleted or mutated cDNA
  => infer key domains, key amino acids!!

Modifying genomes within cells

- Use sequence-specific recombination to introduce desired mutations into target DNA plasmids …or cellular genomes

- Recombination is a key natural aspect of DNA maintenance in cells as well as an artificial result of DNA cleavage and ligation in vitro
- Starts like restriction digestion with a nick or staggered break in the DNA
- Local homology promotes rejoining
- But recombination can introduce new sequences …***or delete original sequences***

Viruses in prokaryotes and eukaryotes modify host genomes "for a living": they encode their own equivalents of restriction enzymes



Integrases: high specificity for non-disruptive sites in viral DNA, varying specificity for sites in host genome

Figure 5-73 *Molecular Biology of the Cell* (© Garland Science 2008)

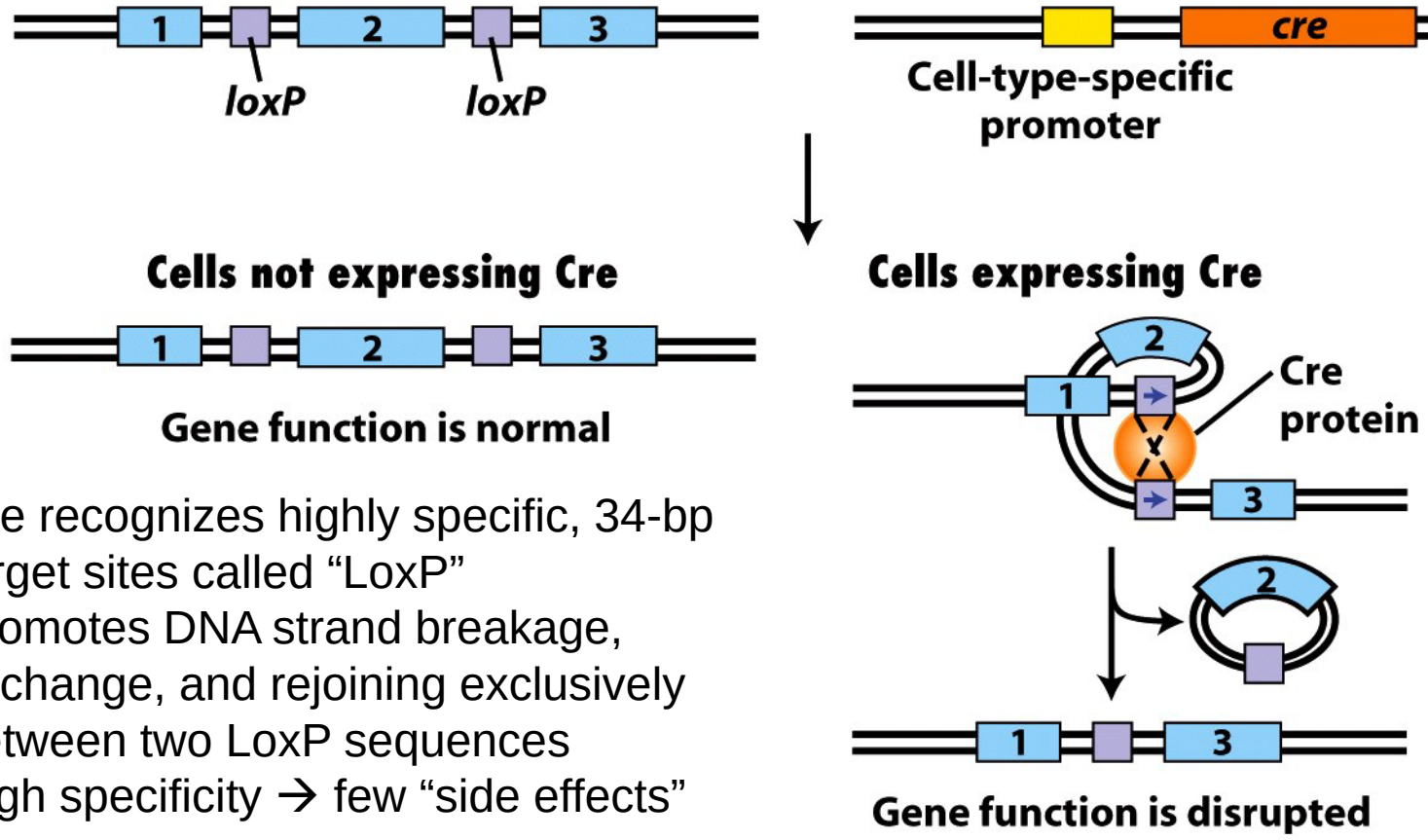
Manipulating recombinase (integrase) specificity for our own uses

- Recombinases made by viruses and parasites have exquisite site specificity … though sometimes only for the “willing” recombination partner (viral genome)

- To use, learn the target site for specific cutting

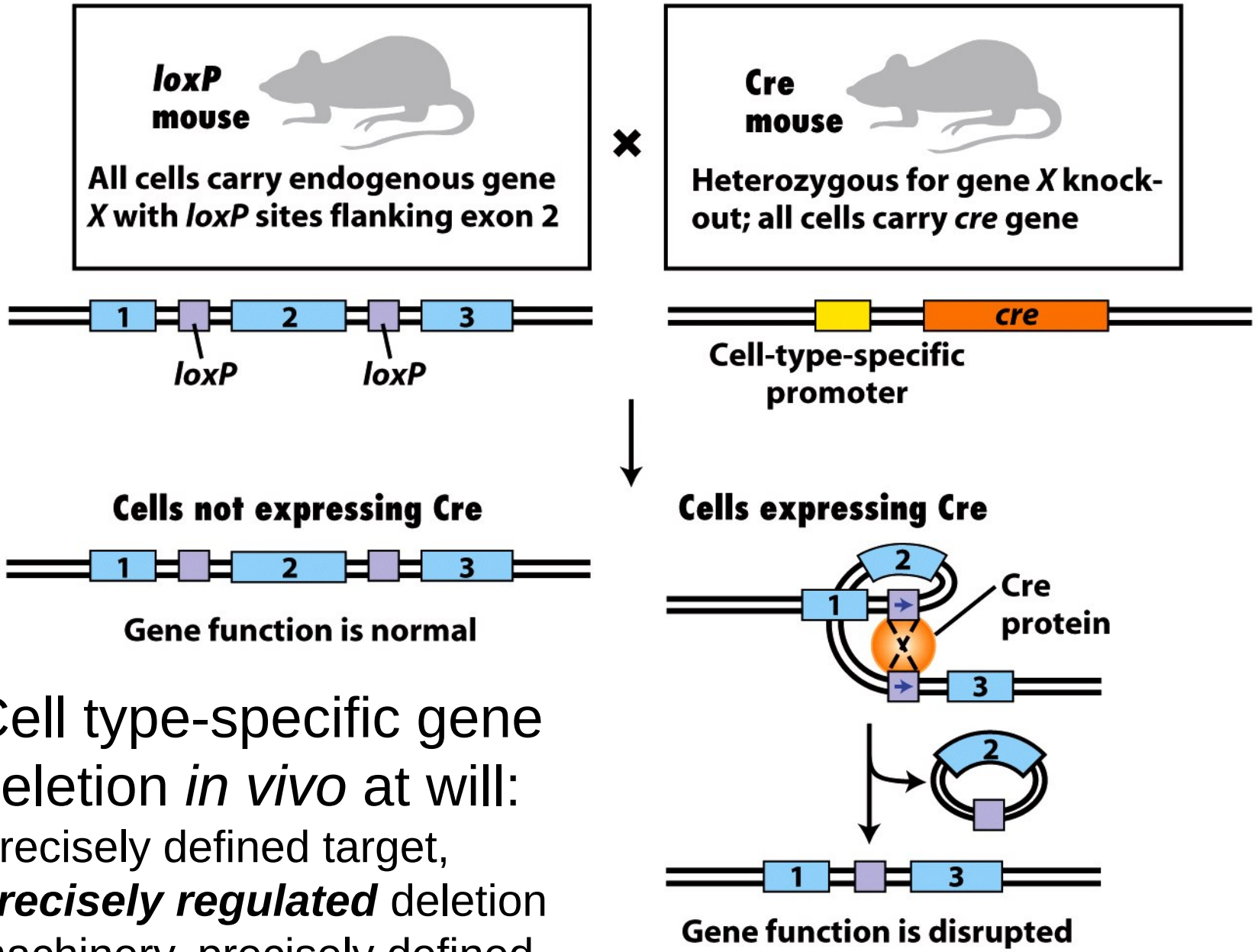
- To control, clone the recombinase gene for expression on YOUR demand

Cre: A particularly useful recombination enzyme – site-specific cutting and rejoining from a single enzyme



Cre recognizes highly specific, 34-bp target sites called "LoxP"
Promotes DNA strand breakage, exchange, and rejoining exclusively between two LoxP sequences
High specificity → few "side effects"

Figure 5-42
Molecular Cell Biology, Sixth Edition



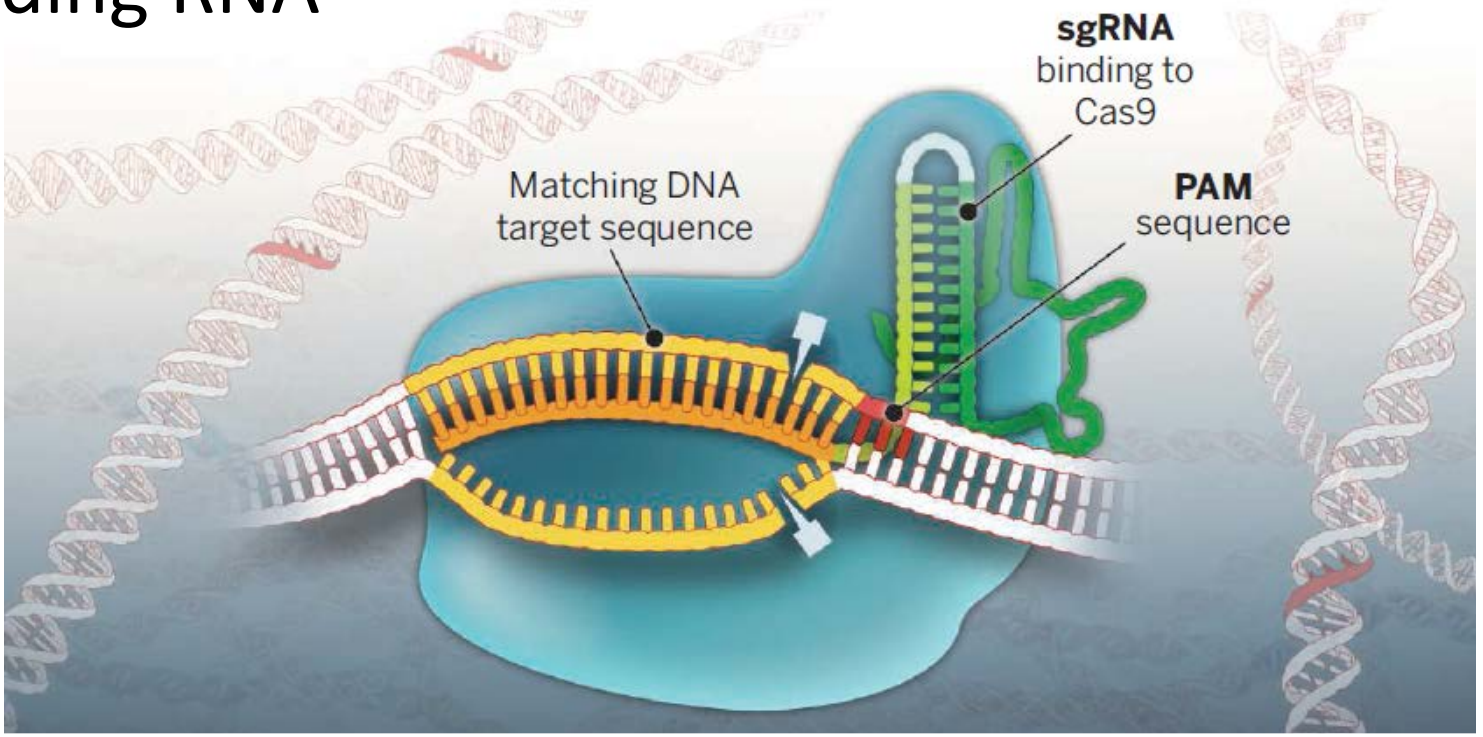
Cell type-specific gene deletion *in vivo* at will:

Precisely defined target, ***precisely regulated*** deletion machinery, precisely defined structure for end product

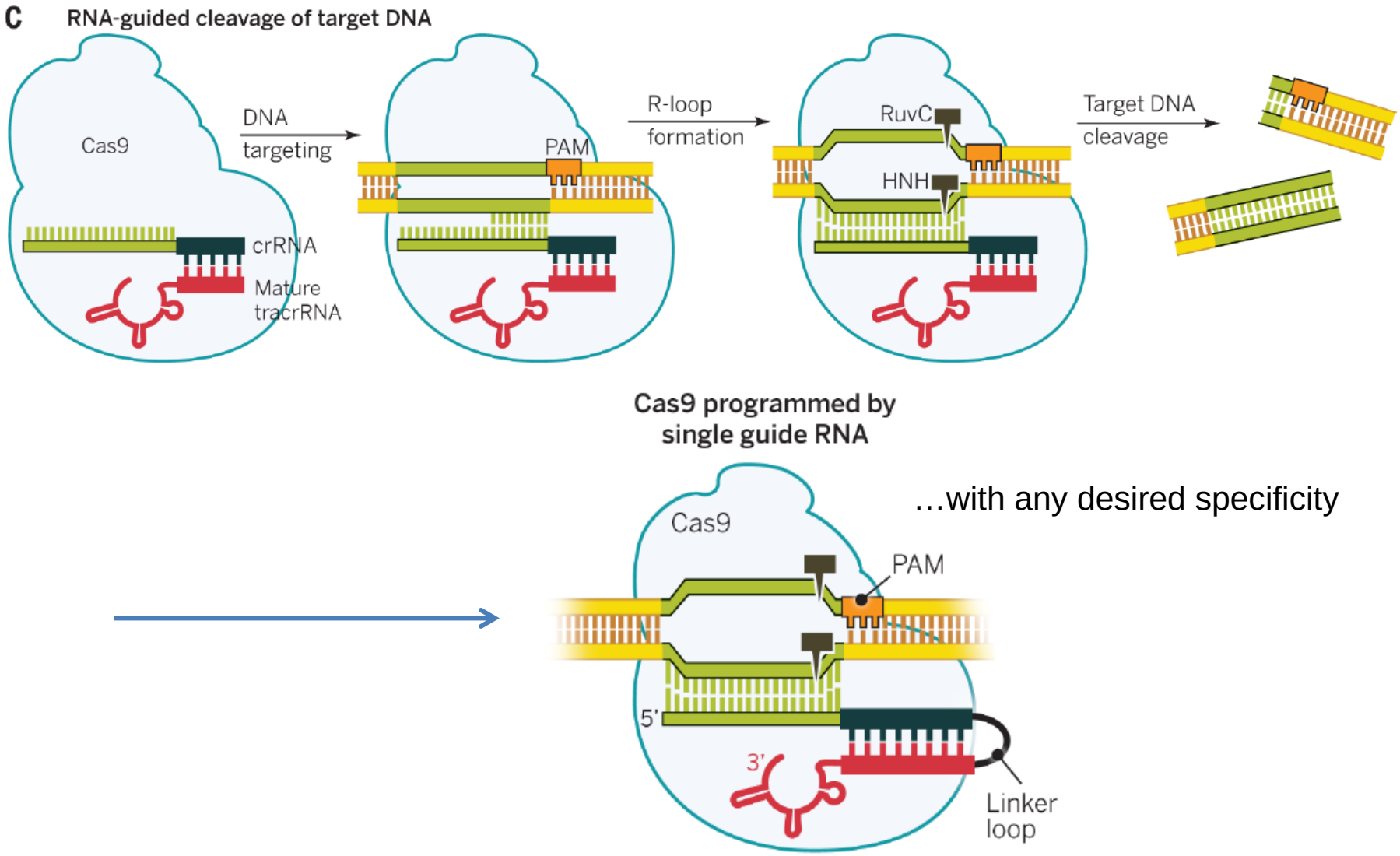
New hottest approach around: CRISPr

- Logistical issues with Cre/LoxP system (and related systems, Flp/Frt, ϕ31, etc.)
  - Good: long-sequence target sites (~30 bp) are very uncommon in mammalian genomes naturally, adding specificity
  - Bad: You have to make a targeted recombinant animal first, with LoxP or Frt sites in right places, before you can use Cre to cut & rejoin…. Years of work, high cost
- CRISPr: a natural bacterial defense system that can target ANY site (almost) if a small cheap complementary RNA is introduced into system

The heart of CRISPr… an obedient protein and a guiding RNA

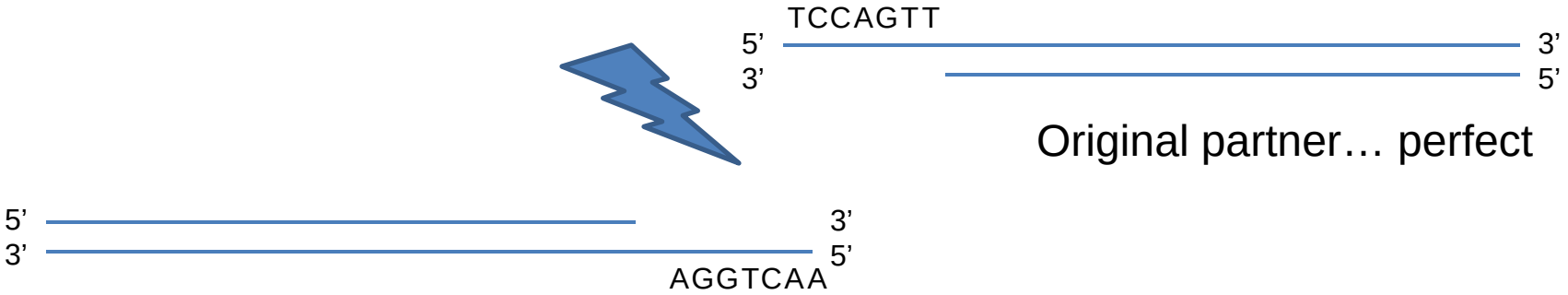


A mechanism in bacteria to use past viral infections to create “immunity” can be made into a tool



…with any desired specificity

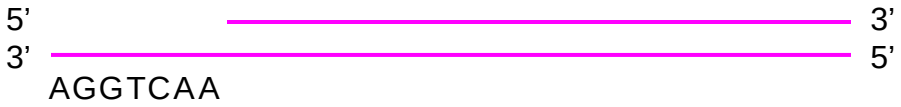
Targeting a new DNA sequence to a DNA break at will: Sequence homology, even over short distance, can enhance DNA break repair



Poor partner for resealing

Imperfect … but adequate!

Poor partner for resealing



Key points for quantitation of biological reactions

- Synthesis / turnover
- Equilibrium vs. Kinetics
- Specific binding
- Enzyme kinetics

Levels of all macromolecules are determined by balance between synthesis and decay rates

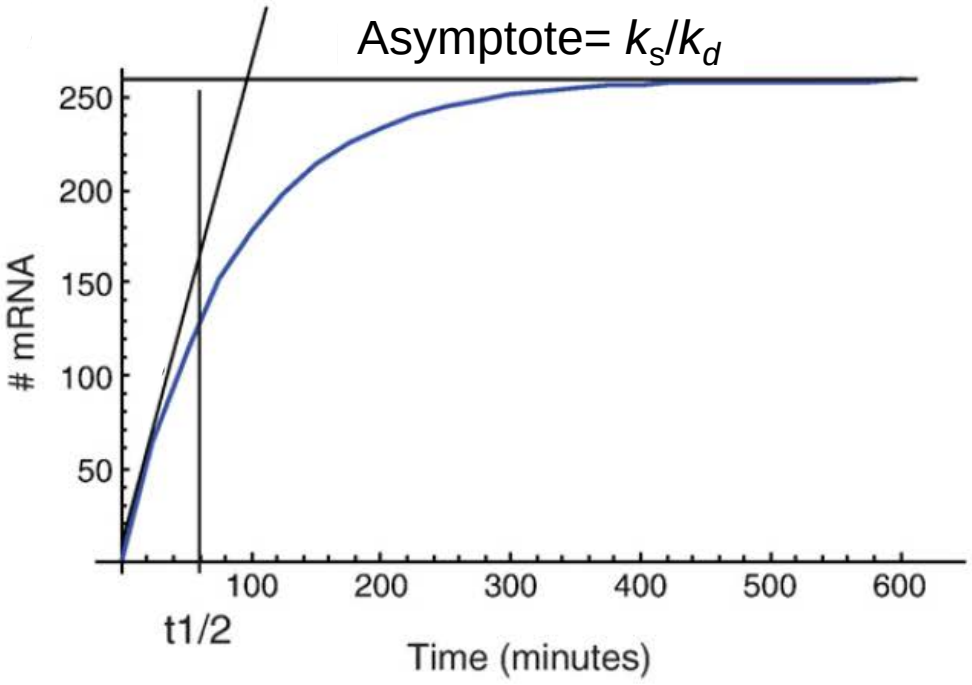
- The rate of increase in any molecule X is

$$dX/dt = k_s - k_d X(t)$$

- $X(t) = (k_s/k_d)(1 - e^{-k_d t})$

- The steady state level for any molecule, when dX/dt = 0, is at

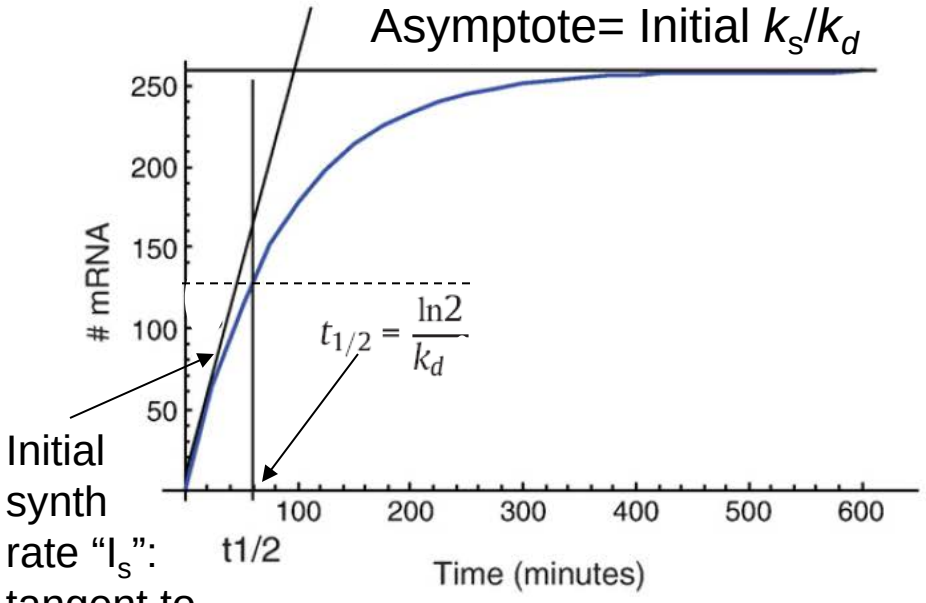
$$X = k_s/k_d$$



Here, Initial Synthesis rate = k_s:
tangent to slope at t= 0

Rates of synthesis and turnover both matter:

Synthesis kinetics and decay kinetics affect both maximum level and time constants of responses



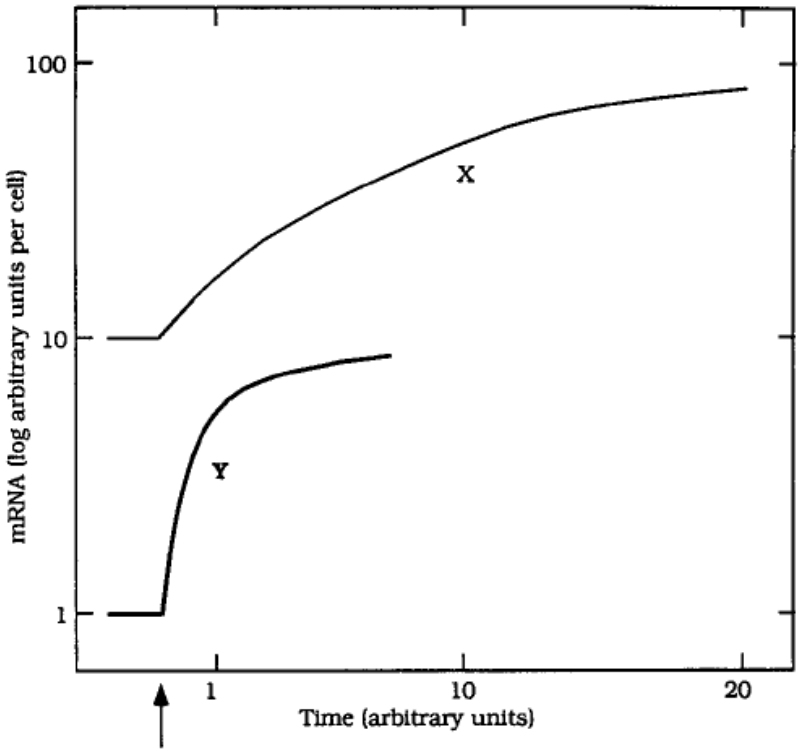


FIG. 3. Rate of change of two mRNAs (stable and unstable) following transcriptional up-regulation (137–139). mRNAs X and Y are transcribed at the same rate. mRNA X is more abundant, because it is 10-fold more stable. The arrow indicates the time at which transcription of both mRNAs increases 10-fold.

Time to half-maximal accumulation depends on decay rate k_d: $t_{1/2} = (\ln 2)/k_d$

From Ben-Tabou de Leon & Davidson, Devel. Biol. 325: 317-328 (2009)

Decay rates have enormous impact on biological response ***kinetics*** as well as molecular levels

For molecules differing only in **decay rates**:

Molecules with same steady-state levels but different **turnover rates** can respond to differences in synthesis rate very differently … **Key point is what fraction of molecules at each timepoint are new!**

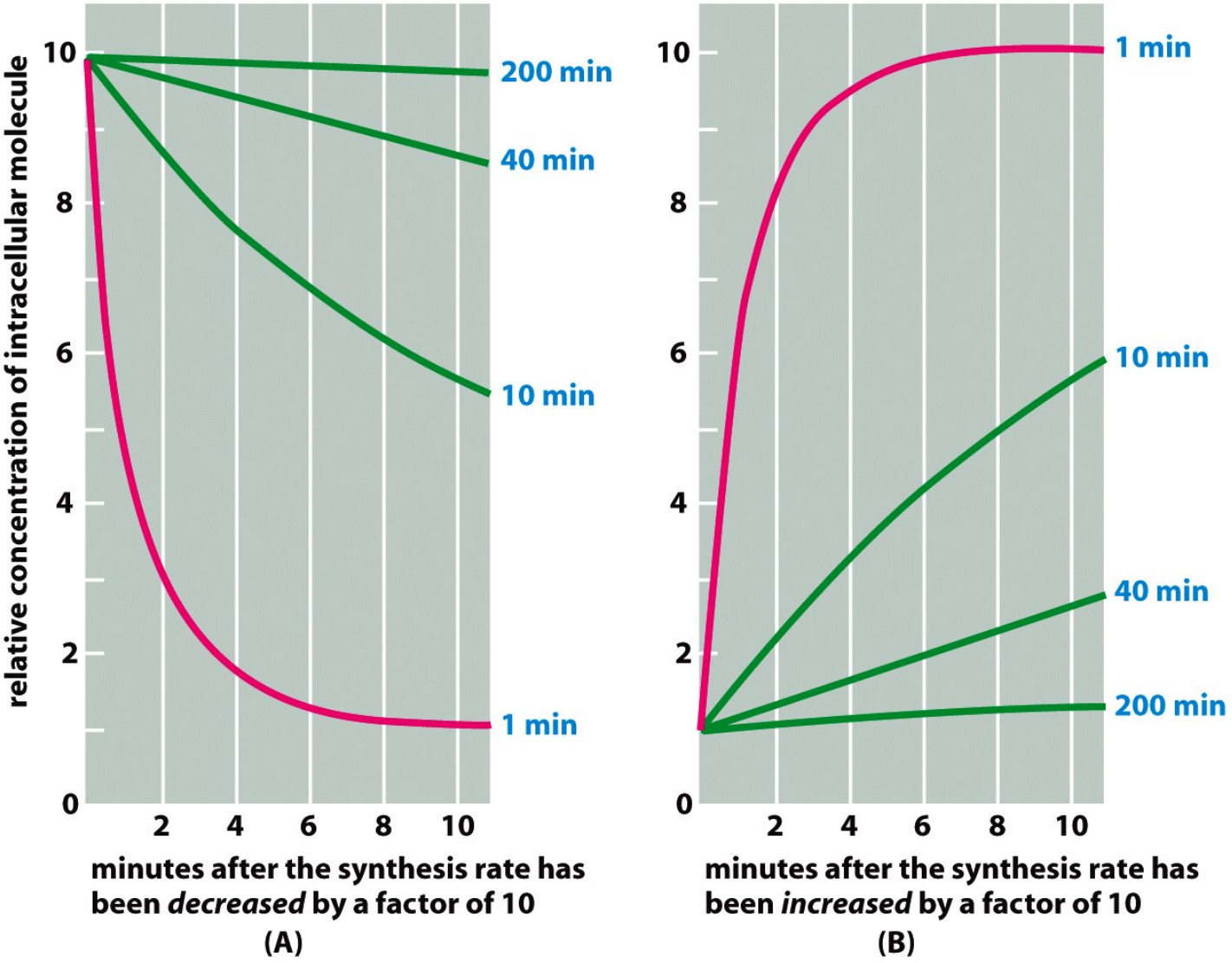


Figure 15-11 *Molecular Biology of the Cell* (© Garland Science 2008)

Equilibria between alternative states are the "constraints" but also the tools of biology

- For aggregates of individual molecules, equilibria between synthesis and degradation
- For chemical recognition, between bound and free
- For reactions, between forward and reverse reaction

Reactions run to equilibrium: level of product at which forward and reverse reactions are balanced

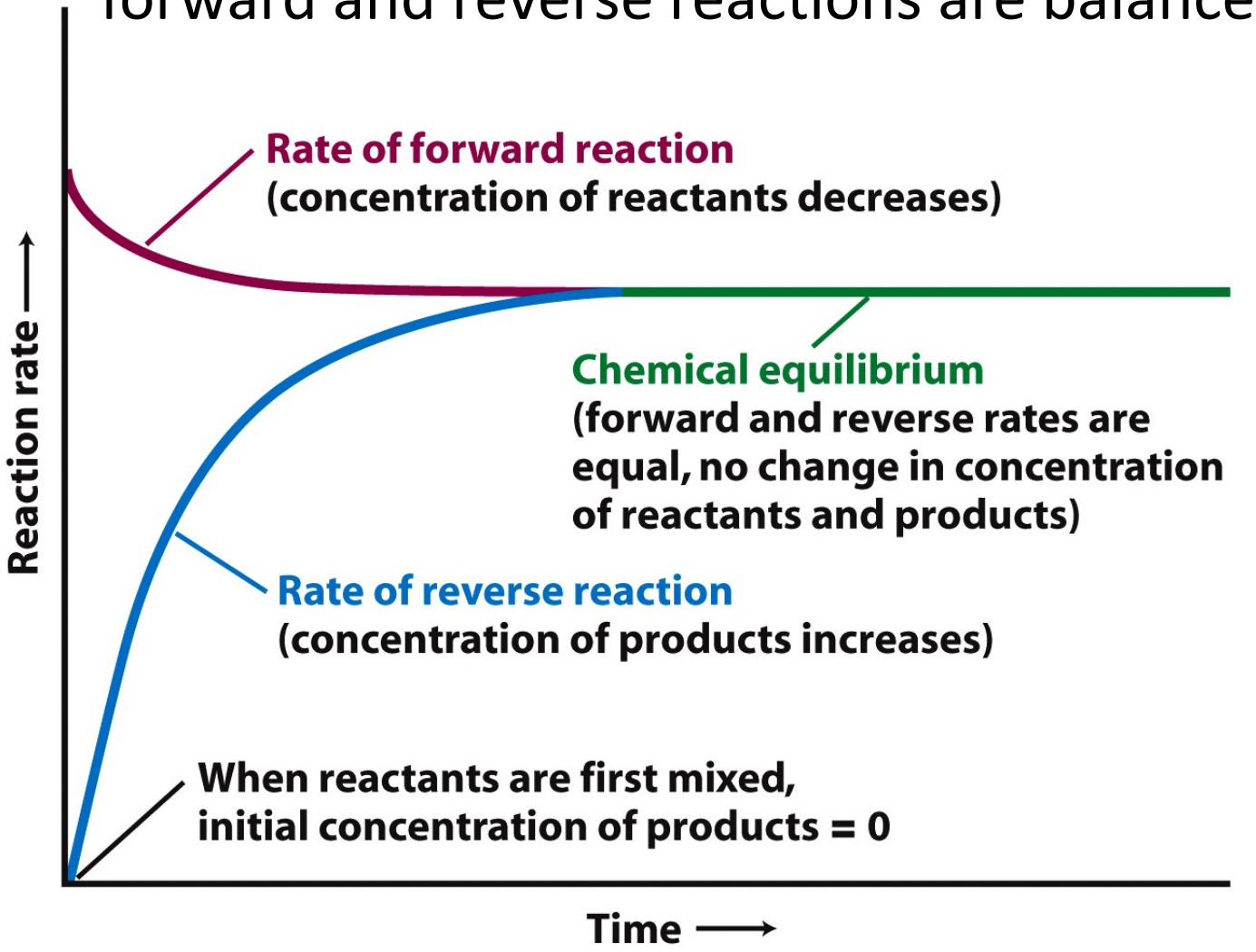


Figure 2-22
Molecular Cell Biology, Sixth Edition

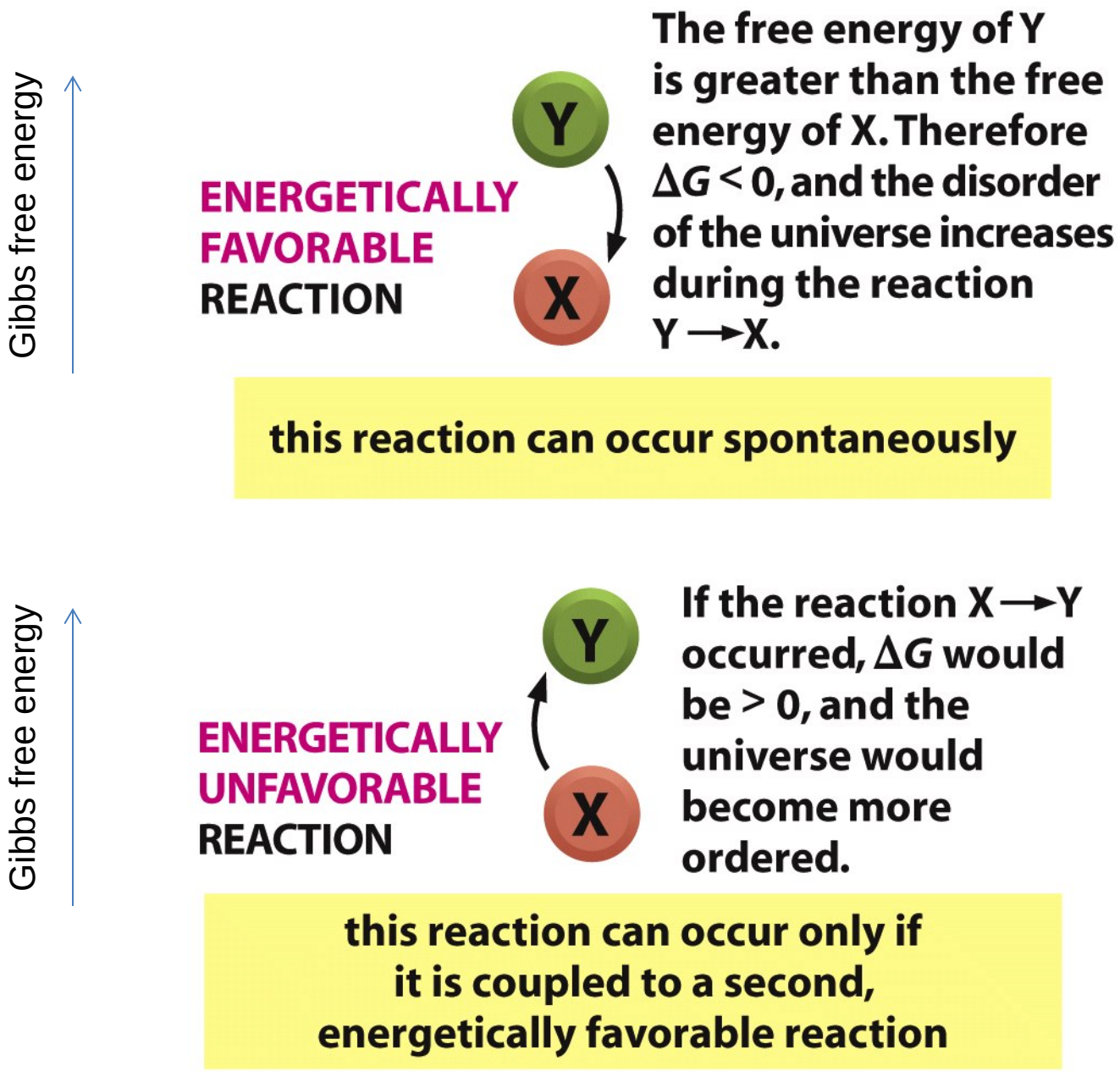
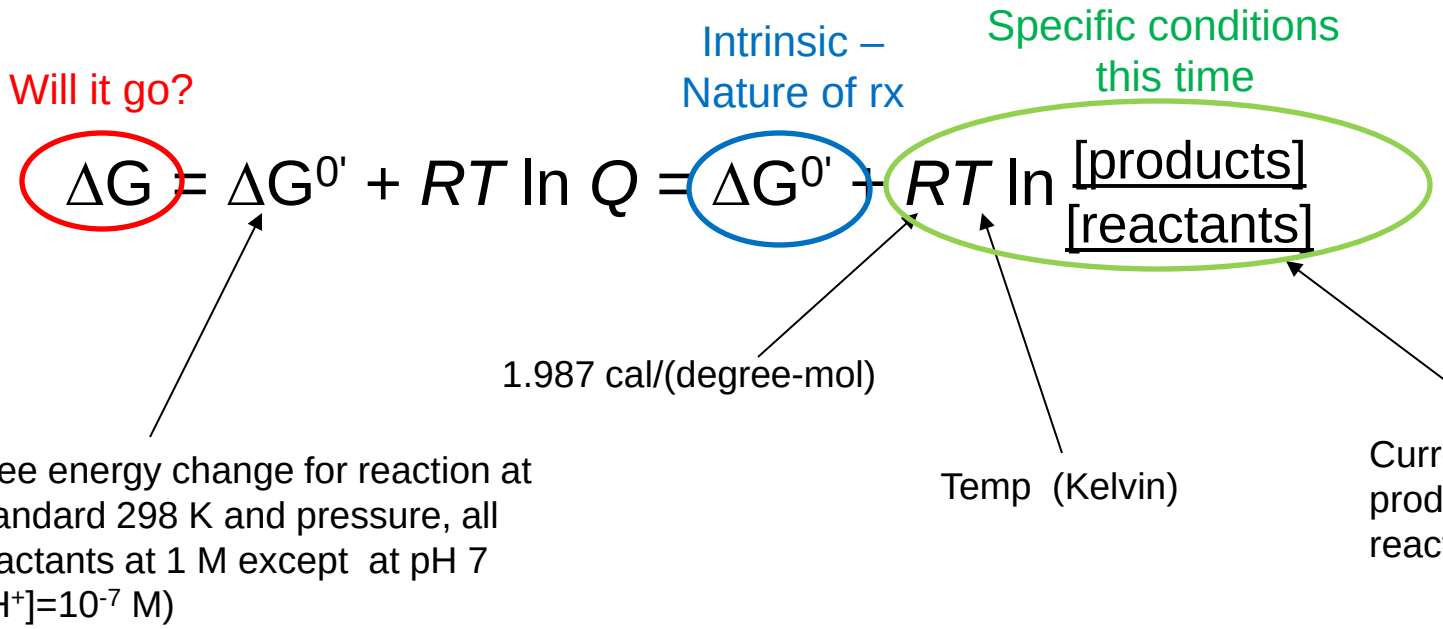


Figure 2-50 *Molecular Biology of the Cell* (© Garland Science 2008)

Gibbs free energy changes predict the direction of chemical and biochemical reactions

$$\Delta G = \Delta H - T\Delta S$$

Free energy is negative if reaction is strongly exothermic ($\Delta H < 0$), or if entropy (ΔS) increases enough to offset endothermic $\Delta H > 0$



Equilibrium constant K_{eq} for a reaction depends directly on $\Delta G^{0'}$

$$\Delta G = \Delta G^{0'} + RT \ln K_{eq}$$

So at equilibrium, when $\Delta G = 0$,

$$\Delta G^{0'} = -RT \ln K_{eq}$$

or $$K_{eq} = e^{(-\Delta G^{0'}/RT)}$$

Or for a binding reaction,

$$[AB]/([A][B]) = e^{(-\Delta G^{0'}/RT)}$$

If $\Delta G^{0'}$ itself is negative, this makes exponent positive and [AB] > [A][B]

Figure 3-43c *Molecular Biology of the Cell* (© Garland Science 2008)

A major type of reaction is binding vs. unbinding of two molecules

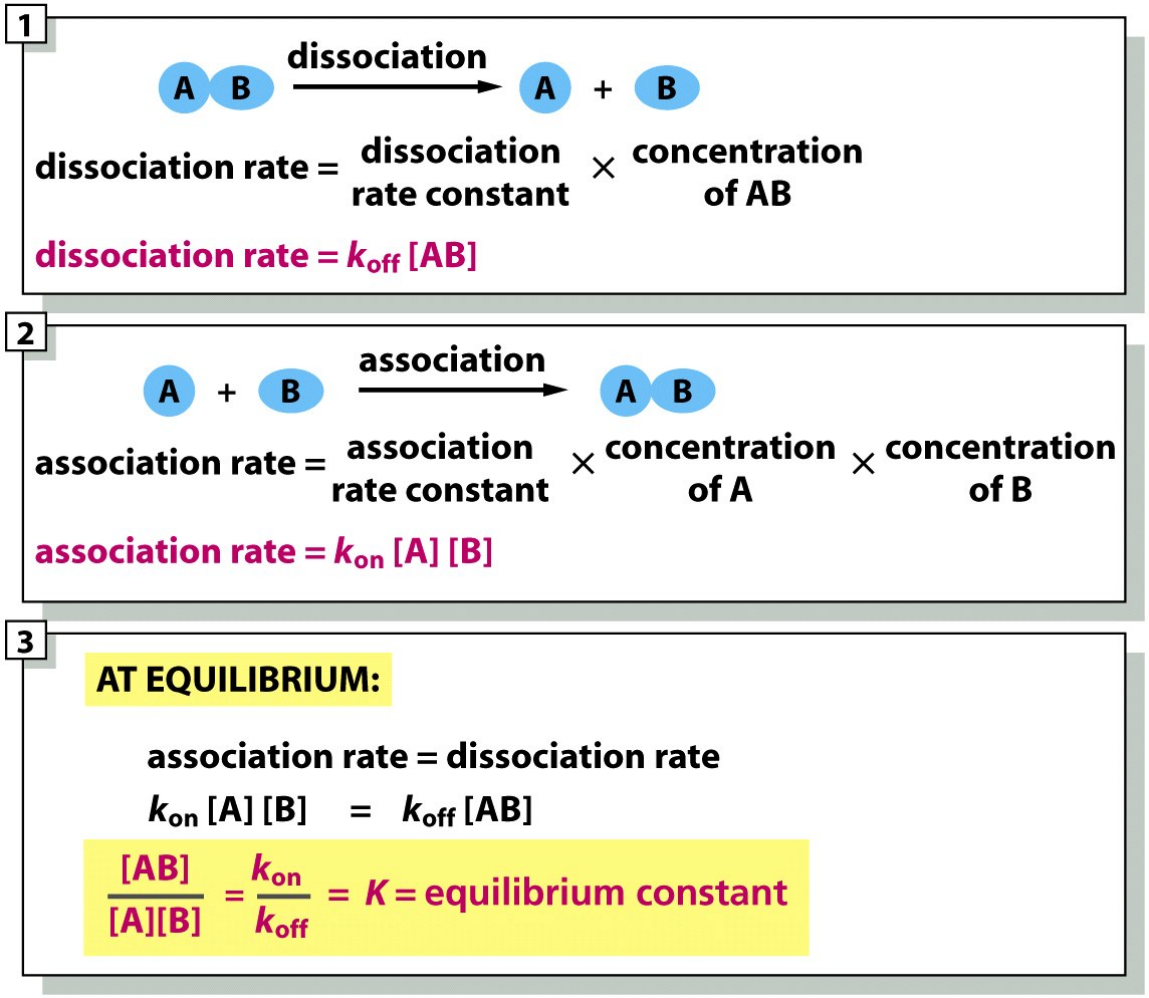


Figure 3-43a *Molecular Biology of the Cell* (© Garland Science 2008)

The power of a few noncovalent bonds in setting chemical equilibria

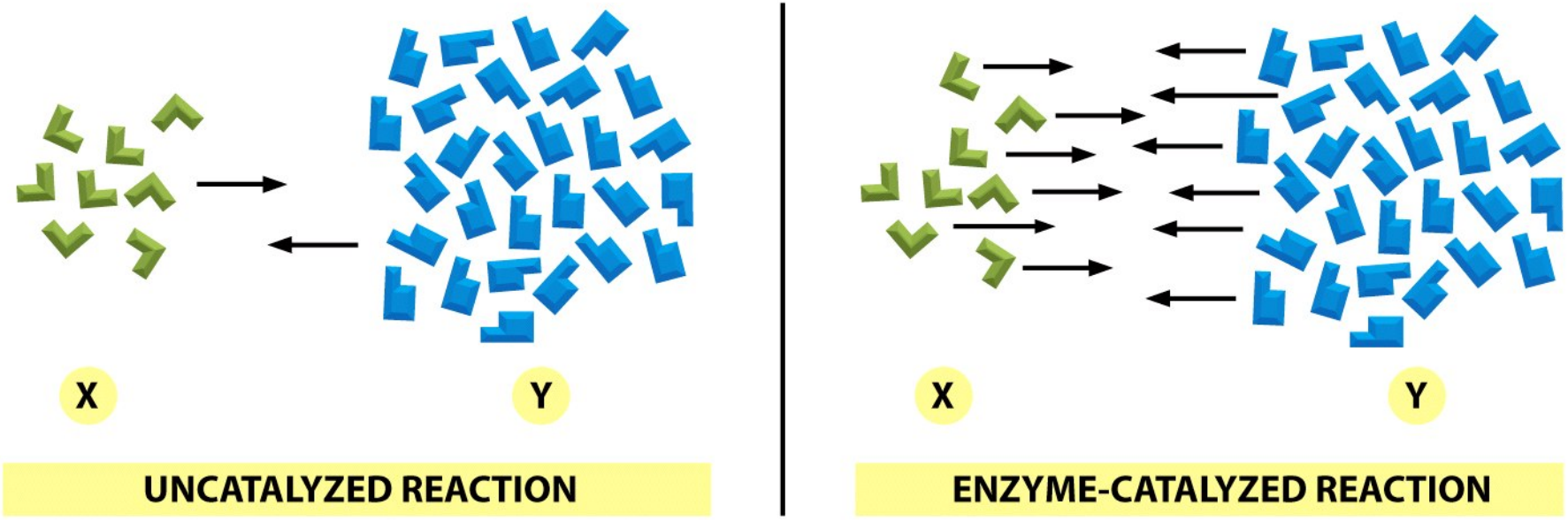
Recall: the strength of *ONE* H-bond in water is ~ 1 kcal/mole … a difference of 2-3 H-bonds in a complex can shift equilibrium constant by a factor of 30-100

The relationship between free-energy differences and equilibrium constants (37°C)

equilibrium constant $\frac{[AB]}{[A][B]} = K$ (liters/mole)	free-energy difference of AB minus free energy of A + B (kcal/mole)	free-energy difference of AB minus free energy of A + B (kJ/mole)
1	0	0
10	–1.4	–5.9
10^2	–2.8	–11.9
10^3	–4.3	–17.8
10^4	–5.7	–23.7
10^5	–7.1	–29.7
10^6	–8.5	–35.6
10^7	–9.9	–41.5
10^8	–11.3	–47.4
10^9	–12.8	–53.4
10^{10}	–14.2	–59.4
10^{11}	–15.6	–65.3

Figure 3-43b *Molecular Biology of the Cell* (© Garland Science 2008)

Enzymes speed up reactions (sometimes immensely) by lowering the free energy cost of the transition states: Direction that reaction will run depends on overall free energy change in current conditions



Catalysis can work both ways in principle… not necessarily a net increase in one direction

Figure 2-53 *Molecular Biology of the Cell* (© Garland Science 2008)

But enzymes can also add to reaction specificity and directionality by coupling them with other reactions

- Siphon off accumulated products to slow reverse reaction

- Couple a reaction physically with an energetically favorable one

  (e.g. ATP + H_2O → AMP + PP_i might be reversible… but after $PP_i + H_2O$ → 2 P_i, overall

  ATP + $2H_2O$ → AMP + 2 P_i gives ΔG ~ -8 kcal/mol)