Lecture 7: Processes and Timescales: Rates for the fundamental processes

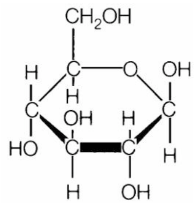
Reading Assignment for Lectures 7--9: Phillips, Kondev, Theriot (PKT), Chapter 5
(Problem Set 1 due Friday, Sept. 24)

Energy inventory for building one E. coli cell:

Cell needs food for at least three purposes:

1. to provide materials for building a new cell (replicating itself).
2. to provide the energy input for the building process.
3. to provide the energy for its other metabolic processes while the building is going on.

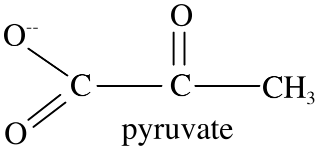
With PKT, I will assume that both materials and energy come from the metabolism of glucose ($C_6H_{12}O_6$), which is a typical carbohydrate.



The energy release from "burning" one molecule via the glycolytic pathway follows directly from the figure 4 kcal/g=4x10^6 cal/kg (Lecture 2), since MW is $6(12)+12(1)+6(16)=180\ Da$. Thus,

$$\text{energy/molecule}=\frac{4\times10^{6}}{N_{\text{molecules/kg}}}=\frac{4\times10^{6}}{\frac{1000}{180}\cdot6.02\times10^{23}}=1.2\times10^{-18}\ cal=5\times10^{-18}\ J$$

Dividing this by E_{ATP}, $\frac{5\times10^{-18}}{8.4\times10^{-20}}\approx60\ E_{ATP}$, i.e., in principle, there is sufficient energy to "charge up" about 60 ATP's. In practice, the glycolytic pathway can produce something like 30--35 ATP or ATP-equivalent units of useful energy, i.e., it has an efficiency of 30/60 or close to 50%. It achieves this by "degrading" the high-quality energy of the glucose in many small steps by use of a complex enzyme-regulated pathway.



pyruvate

How is the metabolic energy of the glucose extracted inside the cell?

First step is glycolysis: 1 glucose + 2ATP$\longrightarrow$2 pyruvate + 4ATP + 2NADH
The glycolytic pathway consists of a complicated network of 10 reactions. (see next page)
The pyruvate thus formed contributes to cell building in at least two ways:

- It is a major building block for making amino acids necessary to manufacture proteins.
- It is enters the mitochondria, where it is further metabolized ("burned") to CO_2 and water, producing further ATP in the process.

Note: All these processes tend to extract the energy in small increments, thus leading to greater efficiency in the conversion of the stored chemical energy in the glucose molecule into useable "fuel" for cellular processes (ATP, NADH).
With this background, we return to the three items in the food/energy budget of building an E. coli cell:

1. At Lecture 4.7 I calculated the number of C atoms/cell at about 10^{10}. Each glucose supplies 6 C's, so we need about 2x10^9 glucose molecules just for materials.

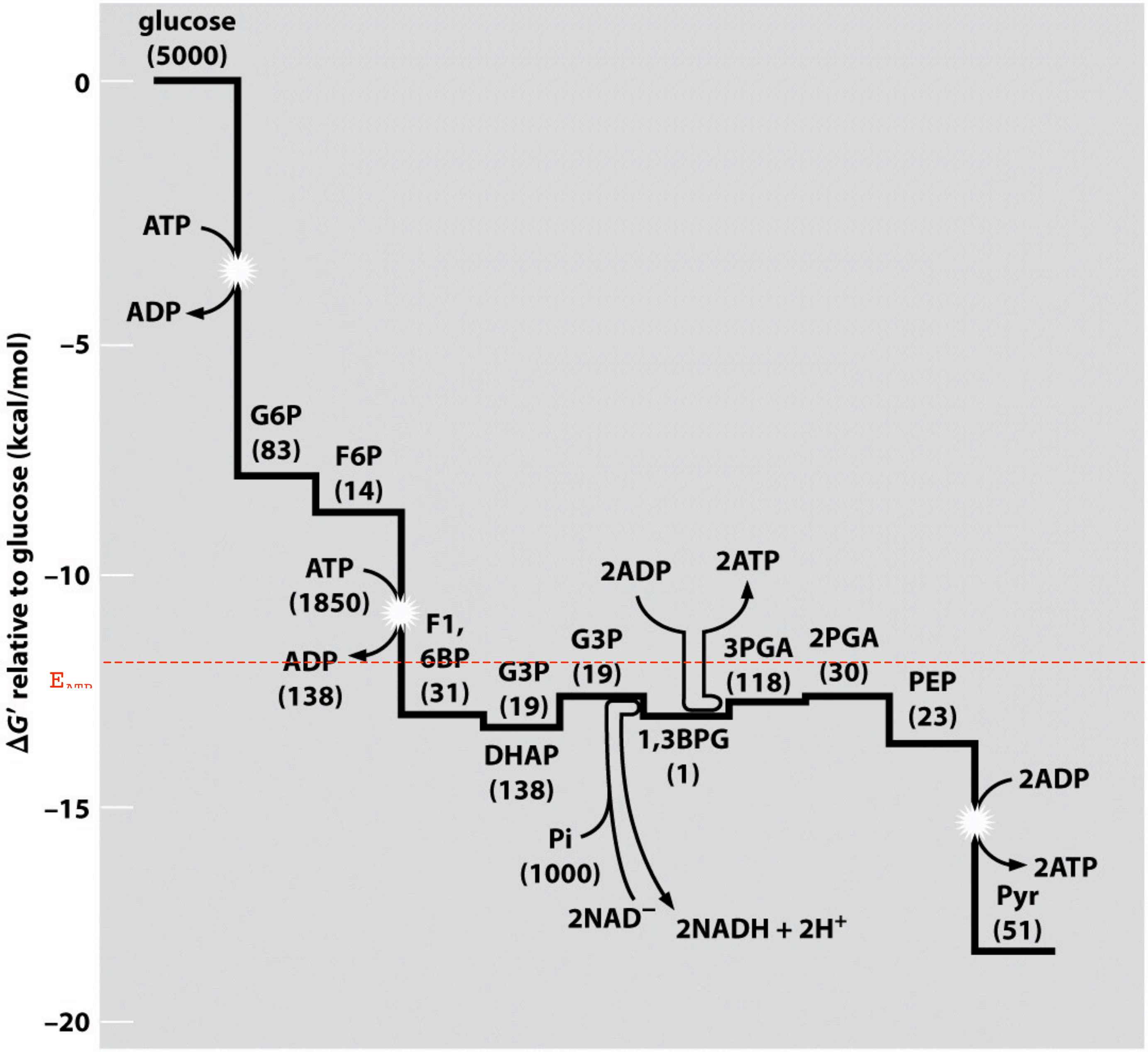


Figure 5.5b Physical Biology of the Cell (© Garland Science 2009)

(Note how little of the overall energy is used up to this point.)

2. In addition, we need the energy to synthesize all the different kinds of molecules in our materials inventory (Chapter 2, Table 2.1), including:

Proteins:	$3\text{x}10^6$	(we estimated this at Lecture 4.6)
Phospholipids:	$2.2\text{x}10^7$	
Lipopolysaccharides:	$1.2\text{x}10^6$	

plus RNA, DNA, etc

We have to follow the production pathway for each of these molecules to count up the number of ATP equivalents required in the building process:
Comment: I assume aerobic growth conditions.

Example: Proteins
To synthesize a typical amino acid requires 1.2 E_{ATP}.
To join each amino acid residue to the growing peptide chain requires 4 E_{ATP}.
Thus, synthesizing 3×10^6 protein molecules requires energy of about:

$$300 \cdot (1.2 + 4) \cdot 3 \times 10^6 \approx 4.5 \times 10^9 \, E_{ATP}$$

Similar calculation for other required molecules leads to Table 5.2:

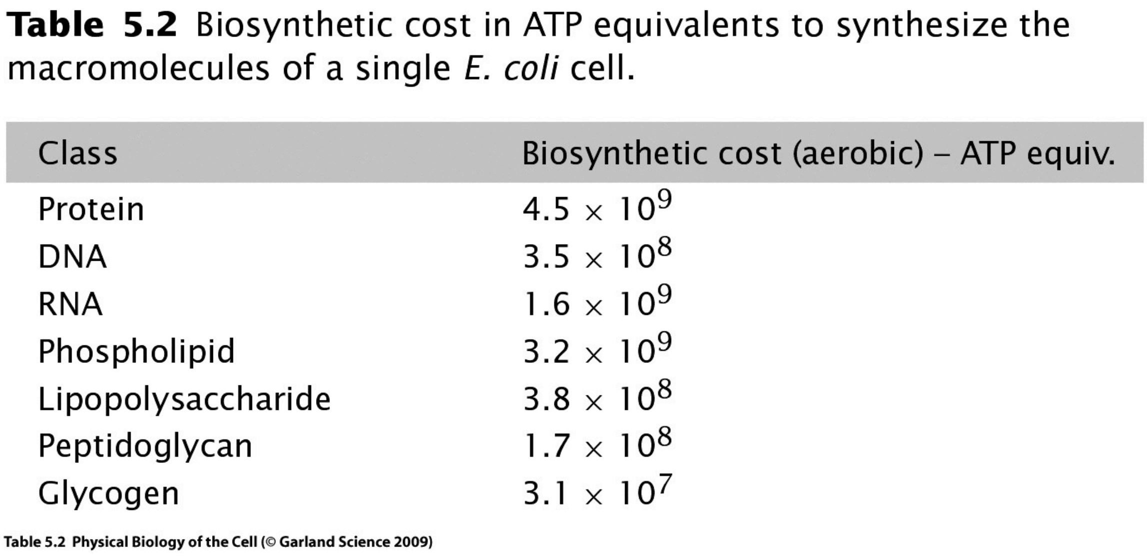


Table 5.2 Biosynthetic cost in ATP equivalents to synthesize the macromolecules of a single *E. coli* cell.

Class	Biosynthetic cost (aerobic) – ATP equiv.
Protein	4.5×10^9
DNA	3.5×10^8
RNA	1.6×10^9
Phospholipid	3.2×10^9
Lipopolysaccharide	3.8×10^8
Peptidoglycan	1.7×10^8
Glycogen	3.1×10^7

Table 5.2 Physical Biology of the Cell (© Garland Science 2009)

Upshot: about 1×10^{10} E_{ATP}. Assume 30 E_{ATP} per glucose (p. 6.6) shows that total number of glucose molecules required is around 0.3×10^9 (compared to 2×10^9 for materials).
Comment: PKT use figure of 15 E_{ATP} per glucose (p. 179) even though they quote the number 30 E_{ATP} on p. 171. I do not understand this apparent inconsistency.

3. Assume a number comparable to 2. (i.e., normal metabolism account for half the energy needed).
Upshot: Glucose required for energy is significantly less than for raw-material, even alculated at, say, 50% efficiency.

At this point I am going to depart briefly from the ordering of material in the text.
(Some of this material is in PKT Chapter 8.1—8.2.)
Partly, I just want to establish a language and notation for discussing probability ideas.
Partly, I want to have a very simple example of a very general result (the "Central Limit Theorem") in front of us when I discuss the ideas of thermodynamics (next).

Introduction to Probability ideas
Experiment or observation has outcomes labeled by a single real ("random") variable x.
Comment: Easy to generalize to multiple variables!
Two common cases:

(a) Discrete statistics: $x \in (x_1, x_2, \ldots, x_M) \equiv \{x_k\}_{k=1}^{M}$, each with probability $p(x_k) \equiv p_k$ subject to

$$p_k \geq 0 \text{ and } \sum_{k=1}^{M} p_k = 1.$$

Comment: In many common cases of discrete statistics, the variables are integers.

Example: M=2 ("Two-state system")

$x_1 = +1$: heads, spin up, step to the right $p_1 = p$ or σ=1 "spin"

$x_2 = -1$: tails, spin down, step to the left $p_2 = q = 1 - p$ σ=-1

(b) Continuous statistics: $-\infty < x < \infty$ (or some other interval) with probability density $p(x)$ (i.e., $dxp(x)$=probability that outcome is between x and x+dx) subject to $p(x) \geq 0$ and $\int_{-\infty}^{\infty} dxp(x) = 1$.

Example: "Gaussian distribution" with "mean" x_0 and "variance" σ^2:

$$p_G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}$$

, which is symmetrical about x_0 with a characteristic "width" σ.

Terminology: Averages and Moments

(a) Discrete	**(b) Continuous**	
$\langle f(x)\rangle \equiv \sum_{k=1}^{M} p_k f(x_k)$	$\langle f(x)\rangle \equiv \int_{-\infty}^{\infty} dxp(x)f(x)$	Definition of "average"

Comment: This commonly used notation can be a little confusing. The x variable is summed/integrated over, so $\langle f(x)\rangle$ does not depend on x. But, it does depend on the probability distribution. Thus, it would be more transparent to write $\langle f\rangle_p$. But, this is not usually done.

$\langle 1\rangle = \sum_{k=1}^{M} p_k \cdot 1 = 1$	$\langle 1\rangle = \int_{-\infty}^{\infty} dxp(x) \cdot 1 = 1$	Normalisation of prob. distrib.
$\langle x\rangle = \sum_{k=1}^{M} p_k x_k$	$\langle x\rangle = \int_{-\infty}^{\infty} dxp(x)x$	"Mean" of x, first moment of distrib.
$\langle x^2\rangle = \sum_{k=1}^{M} p_k x_k^2$	$\langle x^2\rangle = \int_{-\infty}^{\infty} dxp(x)x^2$	"Second moment"

Definition: $\sigma^2 \equiv \left\langle \left(x - \langle x\rangle\right)^2 \right\rangle = \left\langle x^2 - 2x\langle x\rangle + \langle x\rangle^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x\rangle^2$ is called the "variance."

Thus, σ measures the rms deviation of x from its mean <x>.

Comment: Will show below that the parameter σ in the Gaussian distribution agrees with this definition.

Generally, the n-th moment is defined as $\left\langle x^n \right\rangle = \int_{-\infty}^{\infty} dxp(x)x^n$, with n=0,1,2,…

The n=0 moment must exist in order for the probability distribution P(x) to be normalisable.
Higher moments are increasingly dominated by larger values of |x|.
In some cases, all moments exist, e.g., when P(x) vanishes outside some finite interval.
In other cases, some or all higher moments may diverge.

Distributions which behave as $p(x) \sim e^{-\alpha|x|}$ at large |x| are always OK, since the exponential dominates any power law.

However, suppose $p(x) \sim |x|^{-\tau}$ at large |x|: $\left\langle x^n \right\rangle \sim \int_{-\infty}^{\infty} dx |x|^{-\tau} x^n$.

This integral only converges when τ-n>1. (why?) 7.5
Thus, we need τ>1 in order that the probability distribution be normalizable.
however, if, for example, τ=1.5, then p(x) is a perfectly good distribution but none of the moments, n=1,2,3…, exist. (not even the mean!)
In order for first and second moments to exist, you must have τ>3. One common distribution which does NOT meet this criterion is the Lorenzian, $p(x) \sim \frac{1}{a^2 + x^2}$.

Apply this to examples above:

Two-state system:	Gaussian:
$\langle x \rangle = p - q = 2p - 1$	$\langle x \rangle = x_0$
$\left\langle x^2 \right\rangle = p + q = 1$	$\left\langle x^2 \right\rangle = x_0^2 + \sigma^2$
$\left\langle \left(x - \langle x \rangle\right)^2 \right\rangle = 4p(1-p)$	$\left\langle \left(x - \langle x \rangle\right)^2 \right\rangle = \sigma^2$

For the two-state system, this is easy; for the Gaussian distribution, we need to do integrals.