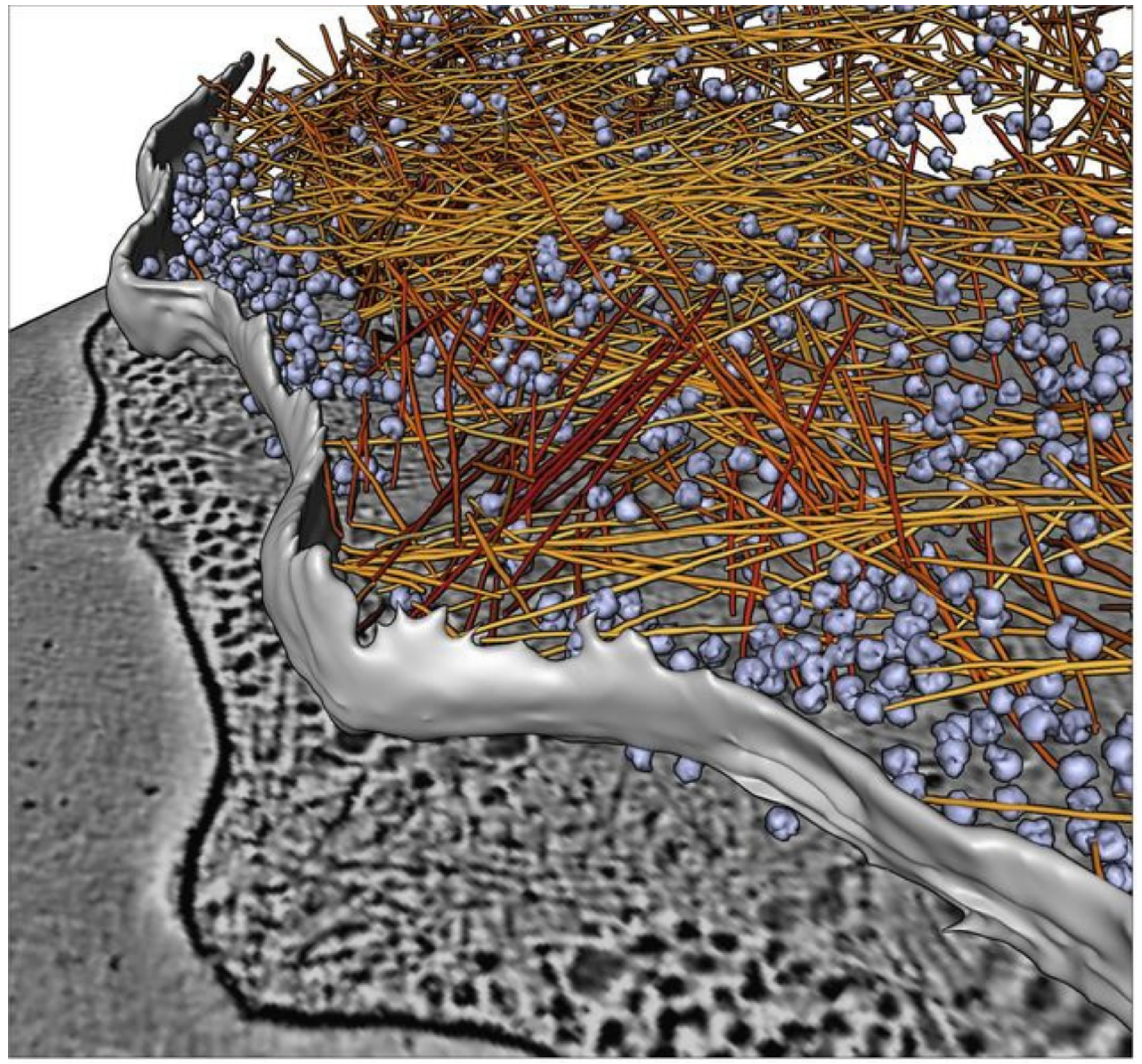
The length, time and energy coordinate system to find your way in the cell



Coordinates for the seminar

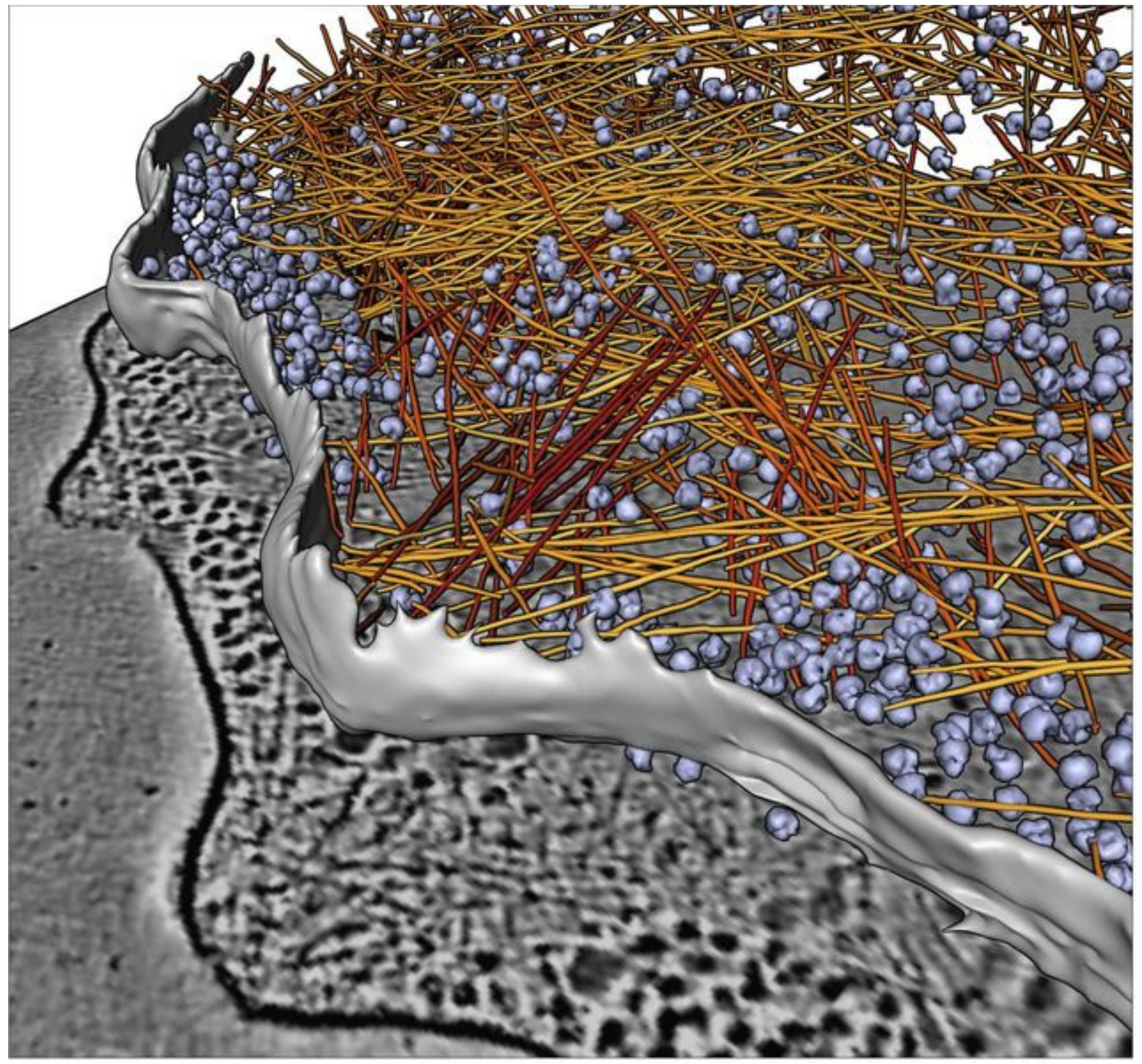
Karsten Rippe
Division of Chromatin Networks
Bioquant, Room 645, 6th floor
Telefon: 54-51376
e-mail: karsten.rippe@bioquant.uni-heidelberg.de

Overview on learning Biophysics in Heidelberg
http://malone.bioquant.uni-heidelberg.de/teaching/index_teaching.html

Material for the lecture: Biophysical concepts and theoretical descriptions

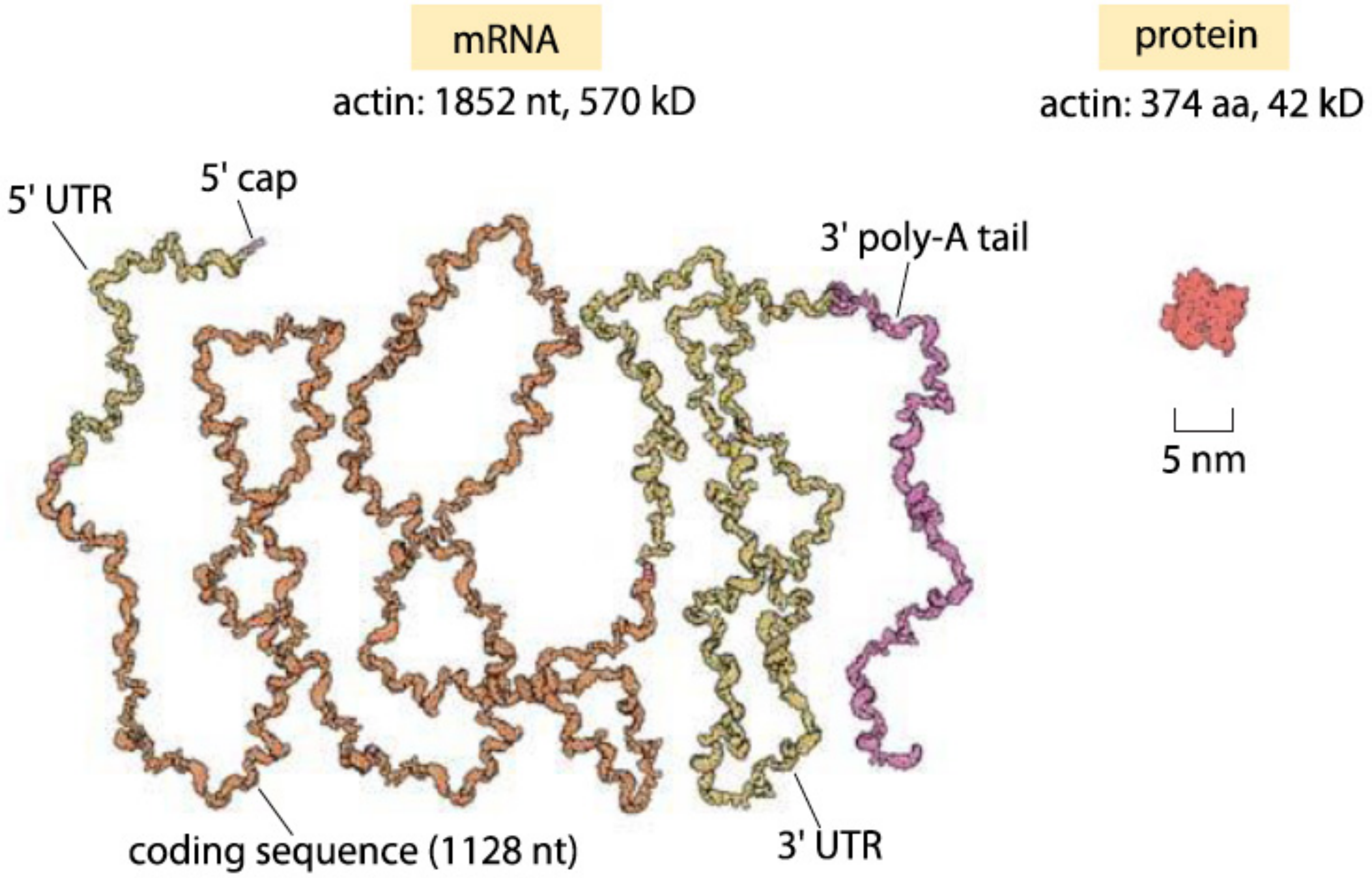
http://malone.bioquant.uni-heidelberg.de/teaching/BPC_lectures/BPC_1+2.html
Username: teaching
Password: nonukes

The length, time and energy coordinate system to find your way in the cell

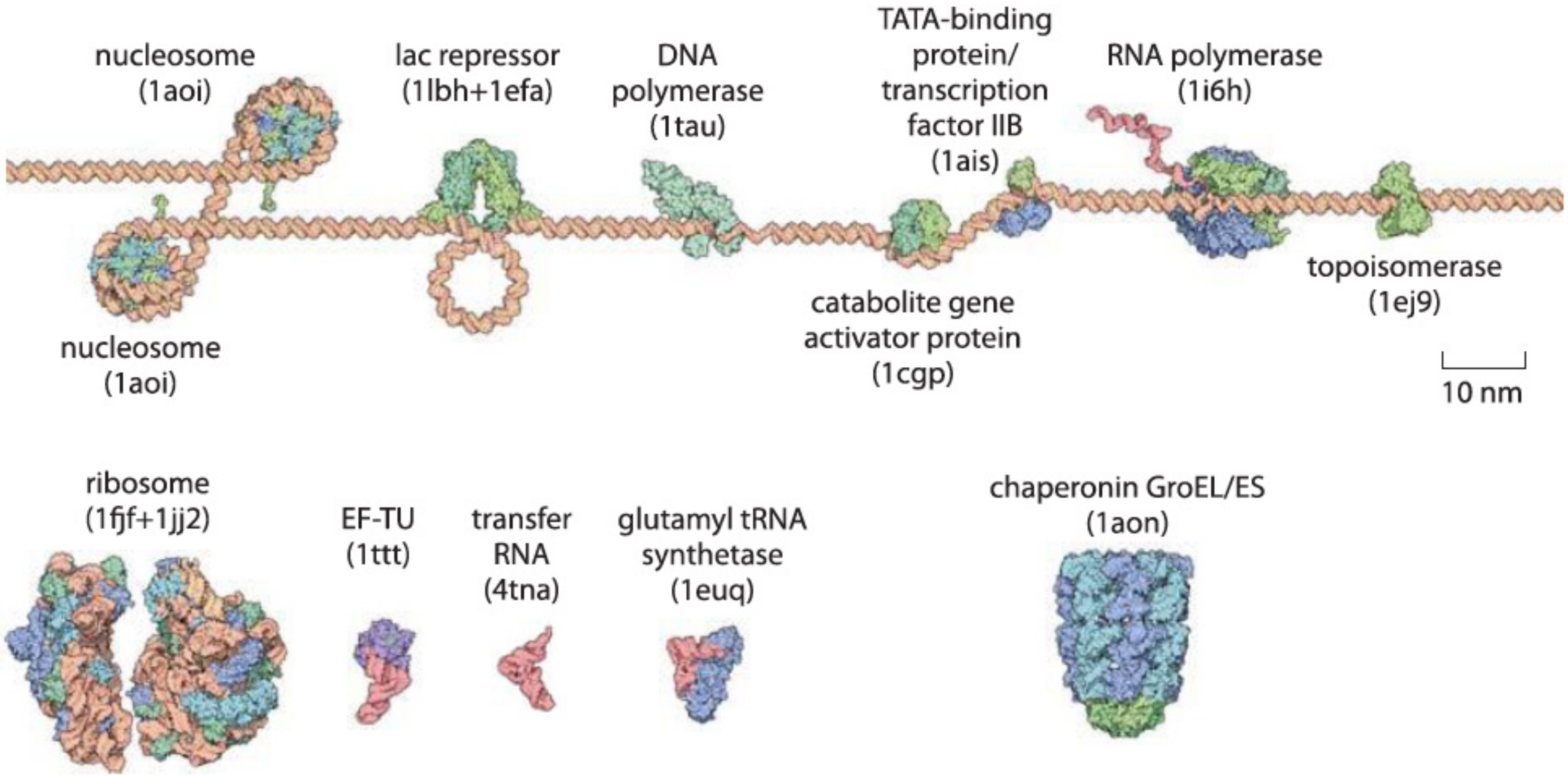


Length scales: What is the size of biological macromolecules?

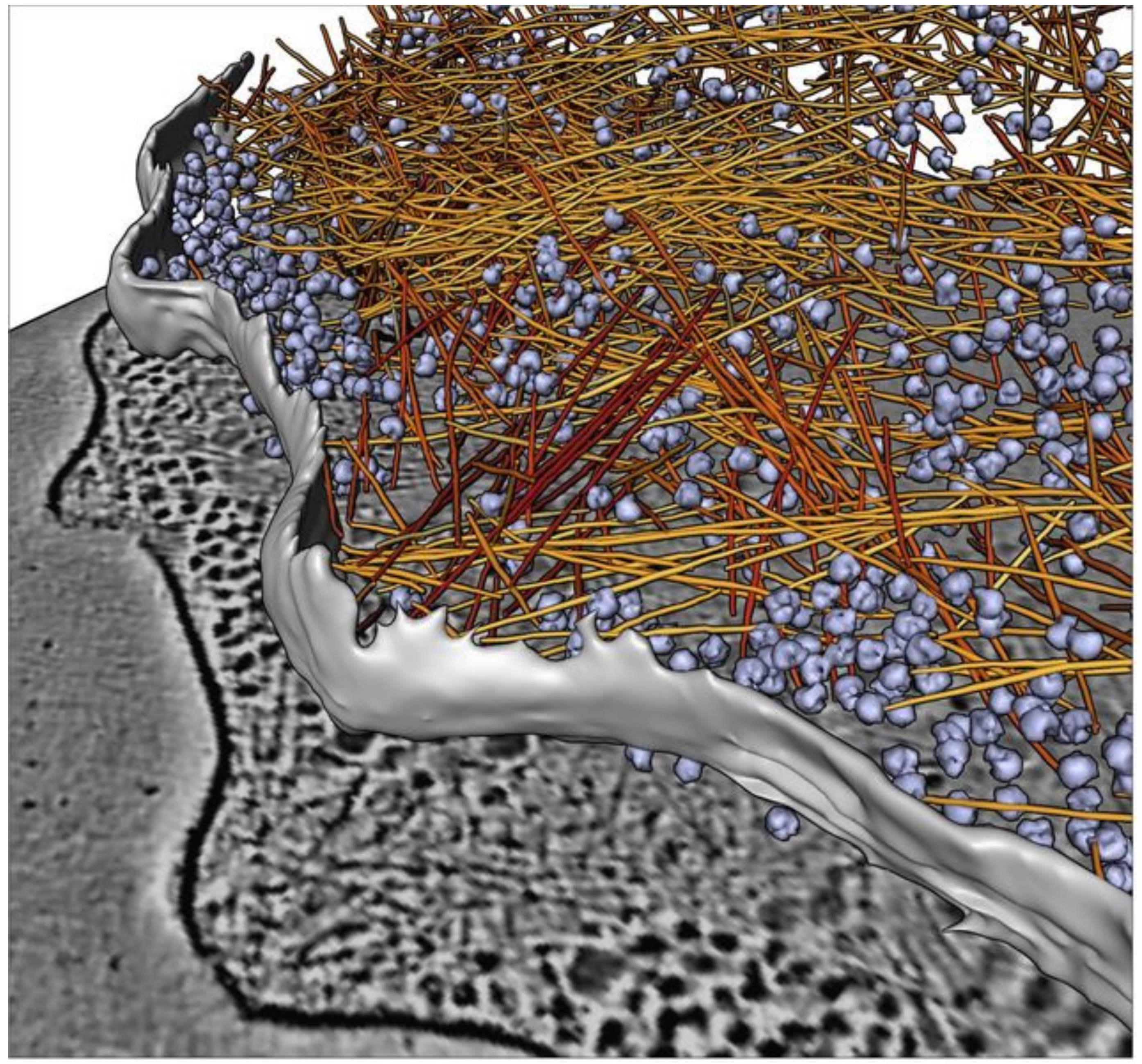
Actin: How does the mRNA size compare to protein size?



Protein and DNA size scales

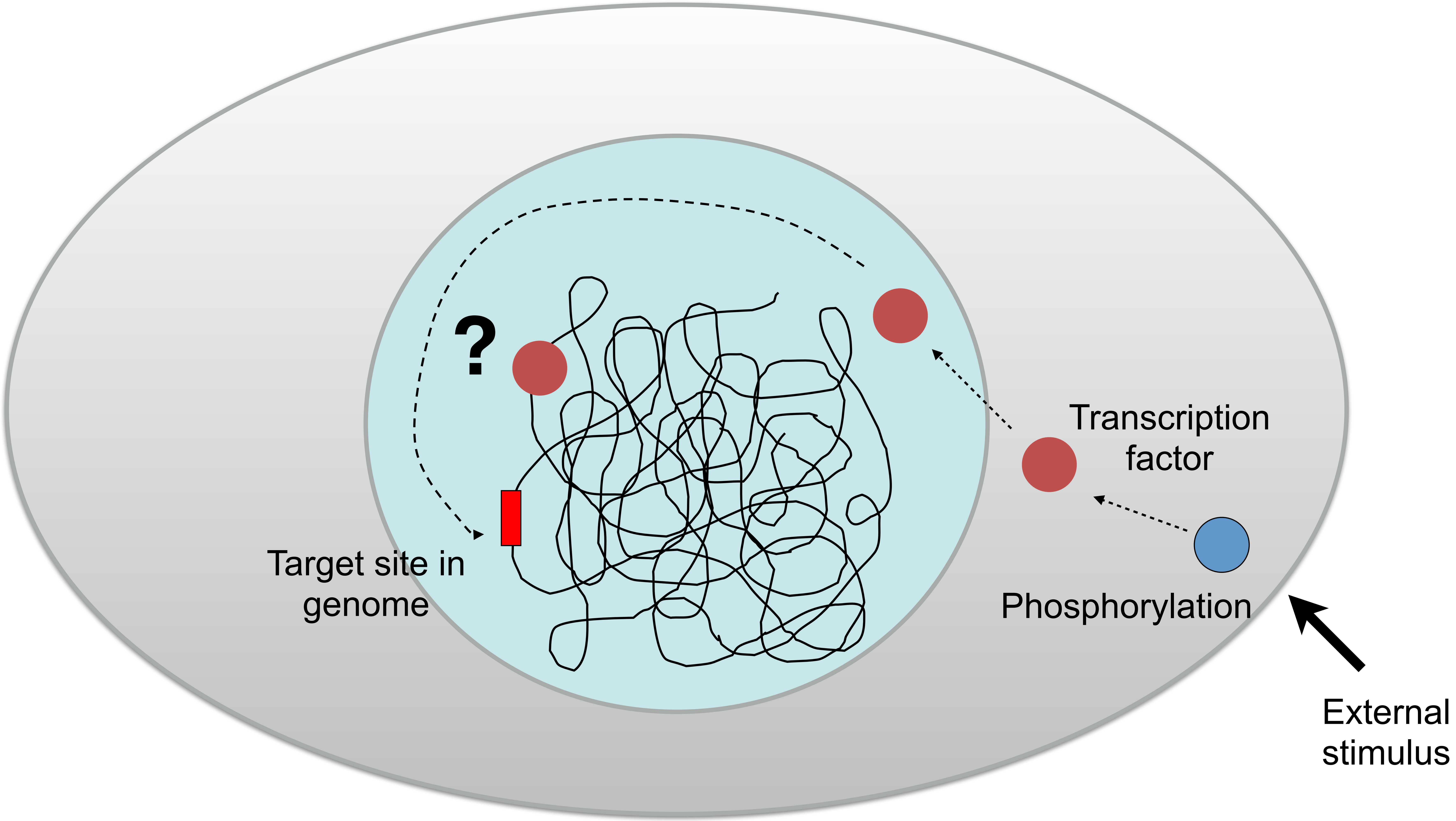


The length, time and energy coordinate system to find your way in the cell

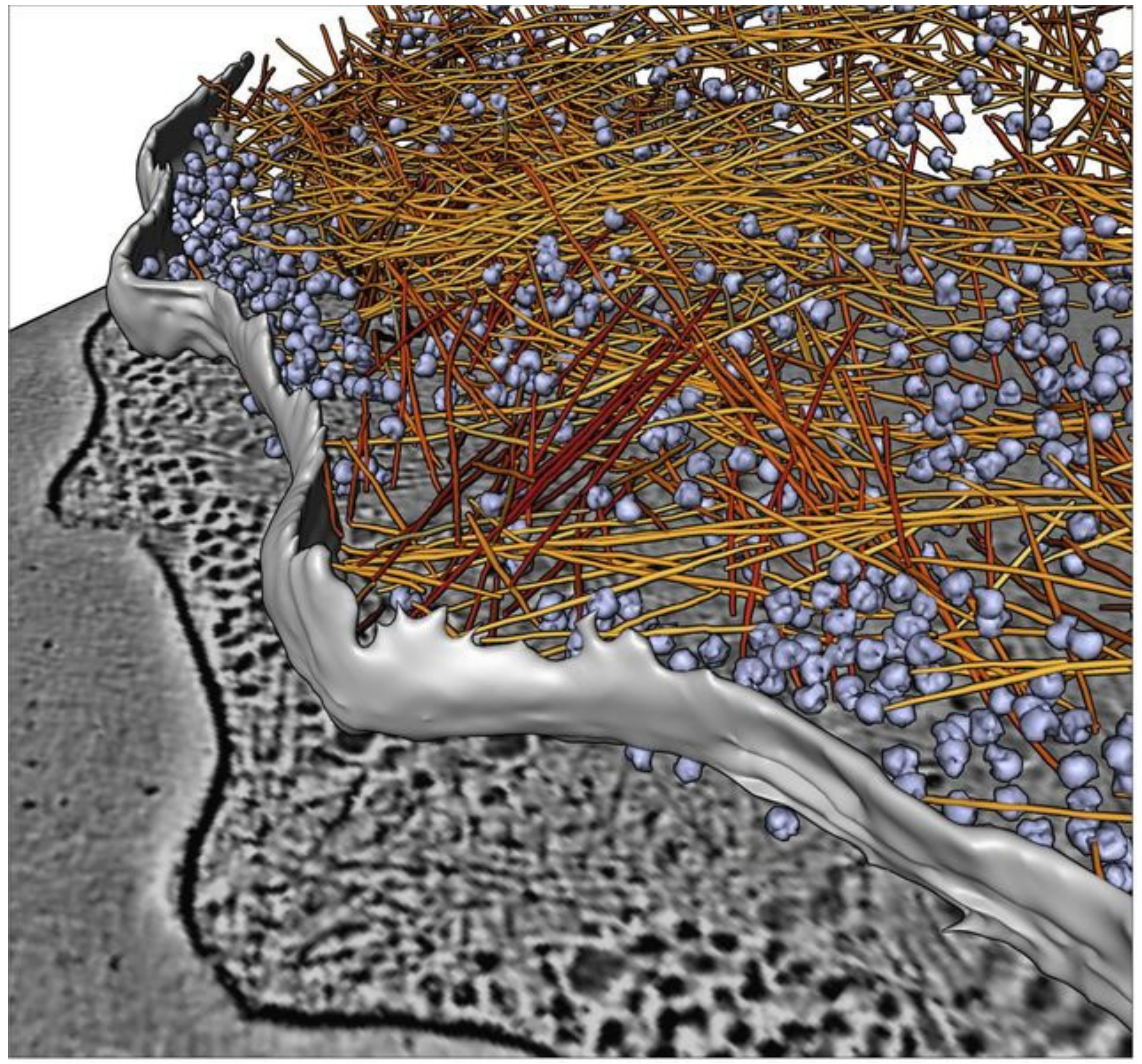


Time scales: How long does it take for a transcription factor (TF) to find its target sequence in a human cell once it has entered the nucleus?

What is the search/accessible volume?
What is the target size?
How fast does the transcription factor move?



The length, time and energy coordinate system to find your way in the cell



Energy scales: How many ATP molecules are hydrolyzed in a cell/day? How much is this in kJ/mol?

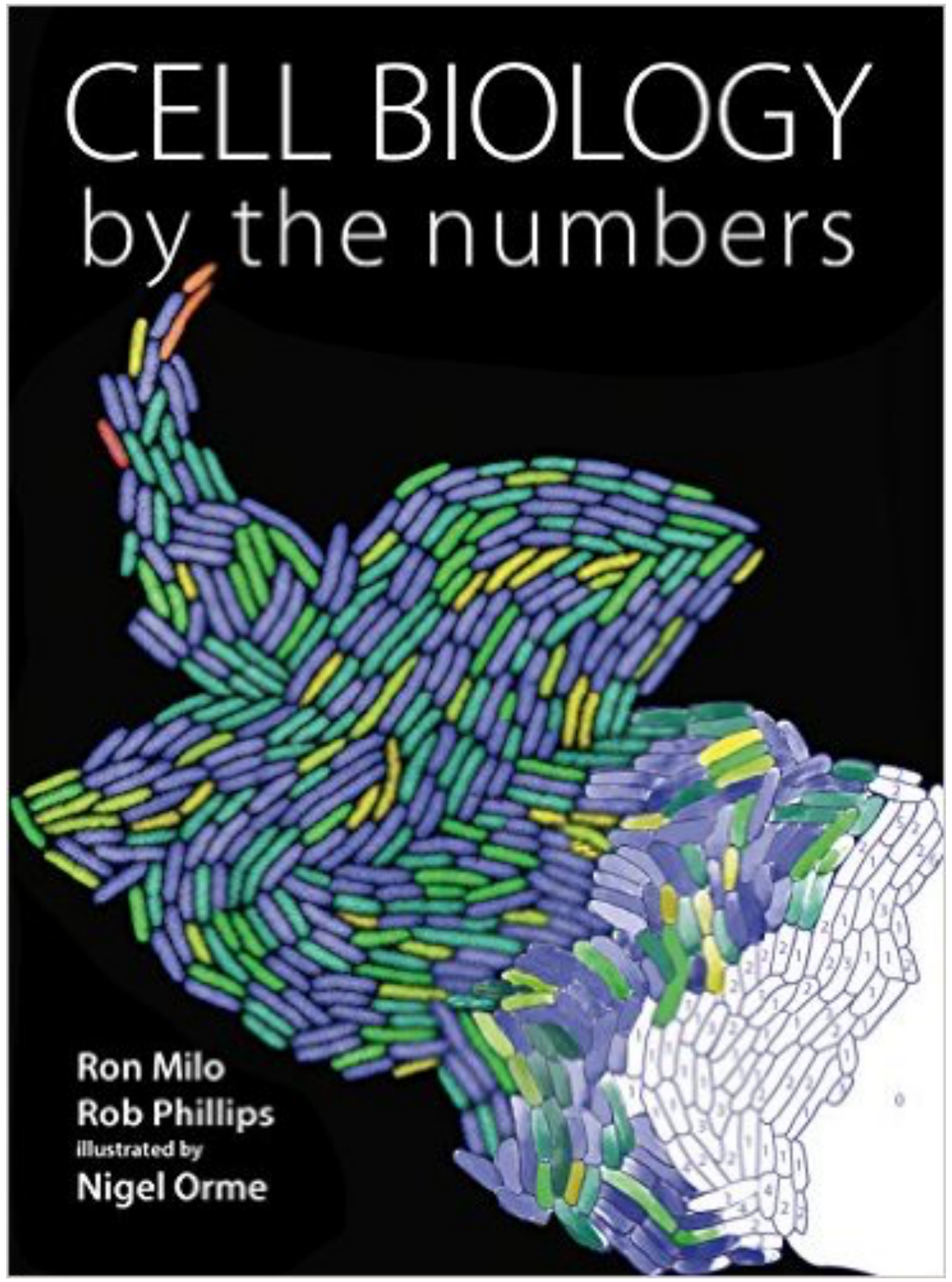
ATP numbers: hydrolysis per day / cell and ATP turnover

Energy metabolism of a student 19-25 years, 74 kg:

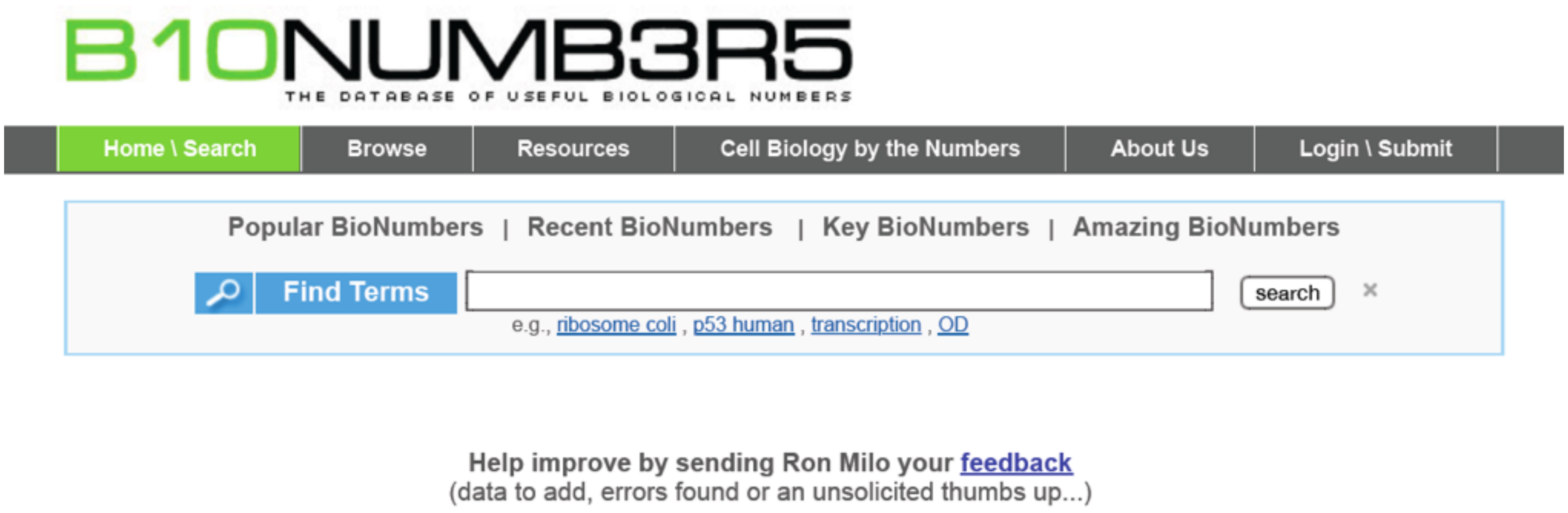
- 1 820 kcal/day basal x 1.65 (typical activity coefficient due to movements on campus)
- 3 000 kcal/day or 12 000 kJ/day
- ATP: 507 g/mol, hydrolysis ~60 kJ/mol
- 12 000 kJ/60 kJ/mol ATP = 200 mol ATP or 100 kg ATP or $1.2 \cdot 10^{26}$ ATP molecules
- $3.2 \cdot 10^{13}$ ATP/cell with $3.7 \cdot 10^{13}$ cells per human body (Bianconi, Ann Hum Biol 40, 463-471, 2013)

Quantitative biology needs numbers…

Cell biology by the numbers (book by Rob Philipps & Ron Milo, Lecture by Ron Milo).
Book draft:
http://book.bionumbers.org



Bionumbers database:
http://bionumbers.hms.harvard.edu



What we would like to do…

- Develop a coordinate system and intuition about the cellular world

- Apply it to “sanity” checks, e.g., in discussions of the research results or for the design of experiments.

- Understand principles that rationalize how the cell operates

Some examples for basic numbers in cell biology and how to use them

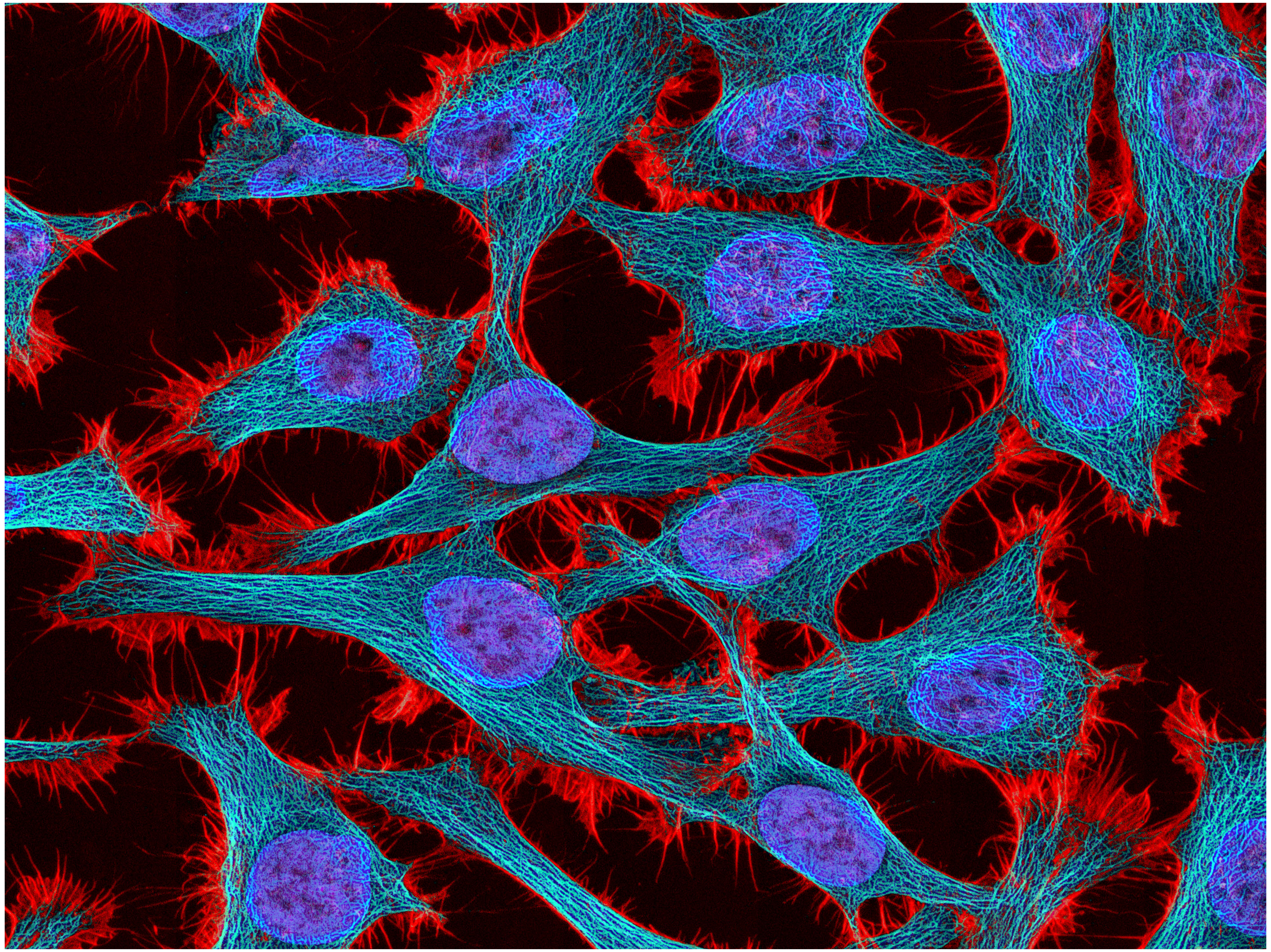
What is the volume of a cell?

E. coli	1 μm^3
budding yeast	30 μm^3
HeLa cell line	3 000 μm^3

Cell volumes

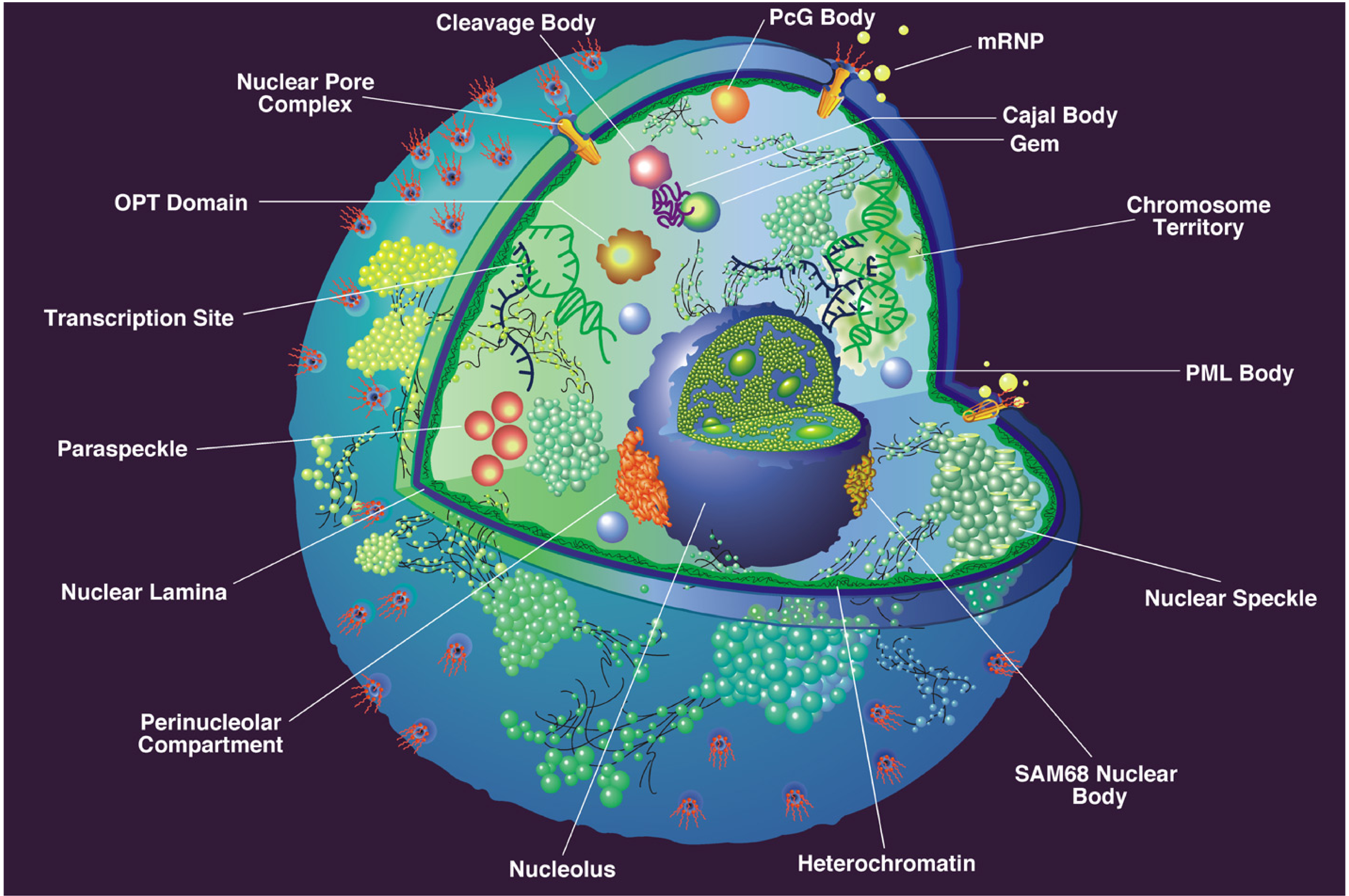
cell type	average volume (μm^3)	BNID
sperm cell	30	109891, 109892
red blood cell	100	107600
lymphocyte	130	111439
neutrophil	300	108241
beta cell	1,000	109227
enterocyte	1,400	111216
fibroblast	2,000	108244
HeLa, cervix	3,000	103725, 105879
hair cell (ear)	4,000	108242
osteoblast	4,000	108088
alveolar macrophage	5,000	103566
cardiomyocyte	15,000	108243
megakaryocyte	30,000	110129
fat cell	600,000	107668
oocyte	4,000,000	101664

HeLa cell (wikipedia): DNA, blue; microtubuli, cyan; actin, red

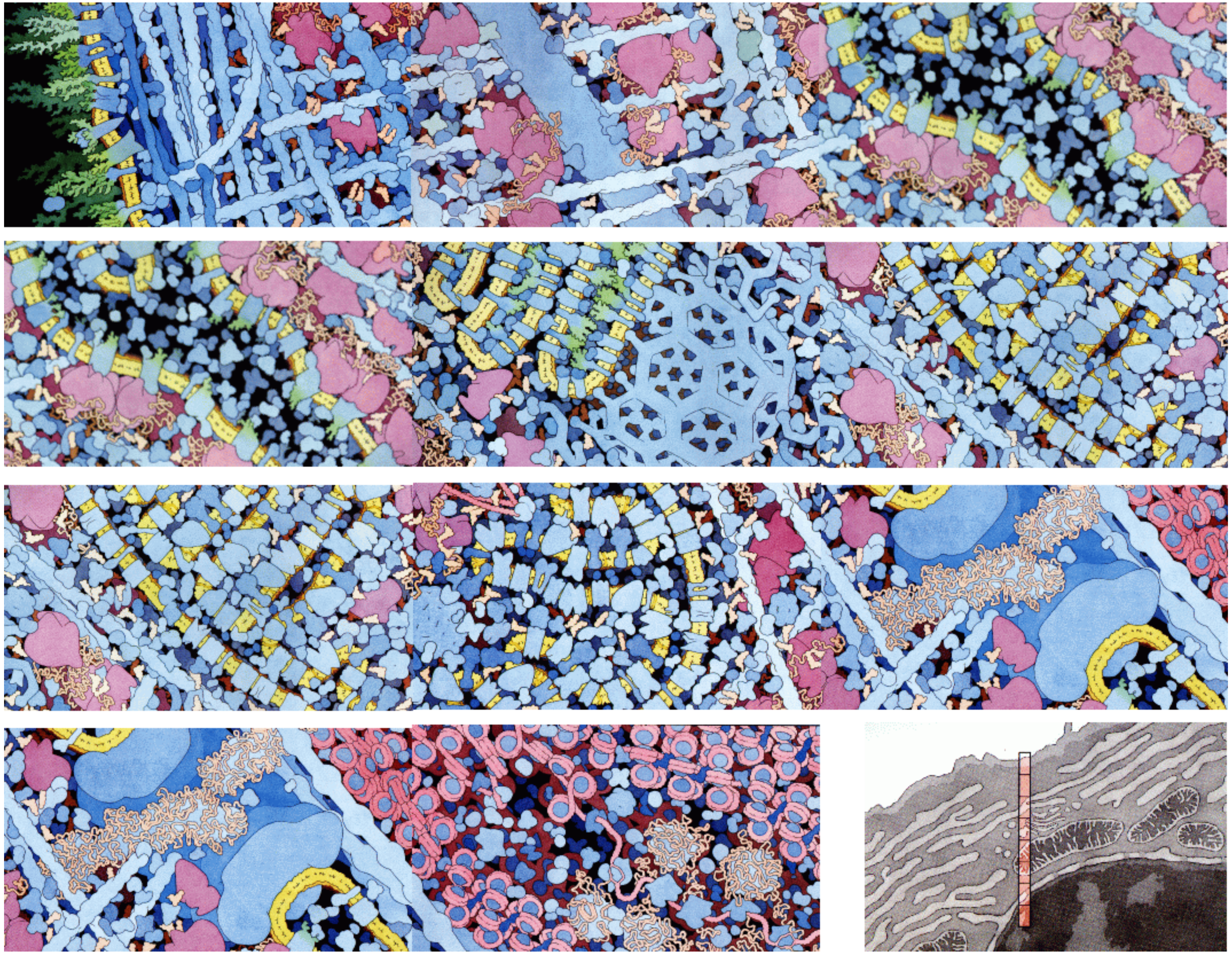


Nucleus diameter: ~10 µm, 0.4 pl volume

The mammalian cell nucleus



The cell is a very crowded place (David Goodsell)



from left to right:

cell surface

cytoplasm

synthesis of proteins from the endoplasmic reticulum

Golgi apparatus,

coated vesicle

mitochondrion

nucleus

proteins: blue

DNA and RNA: red and orange

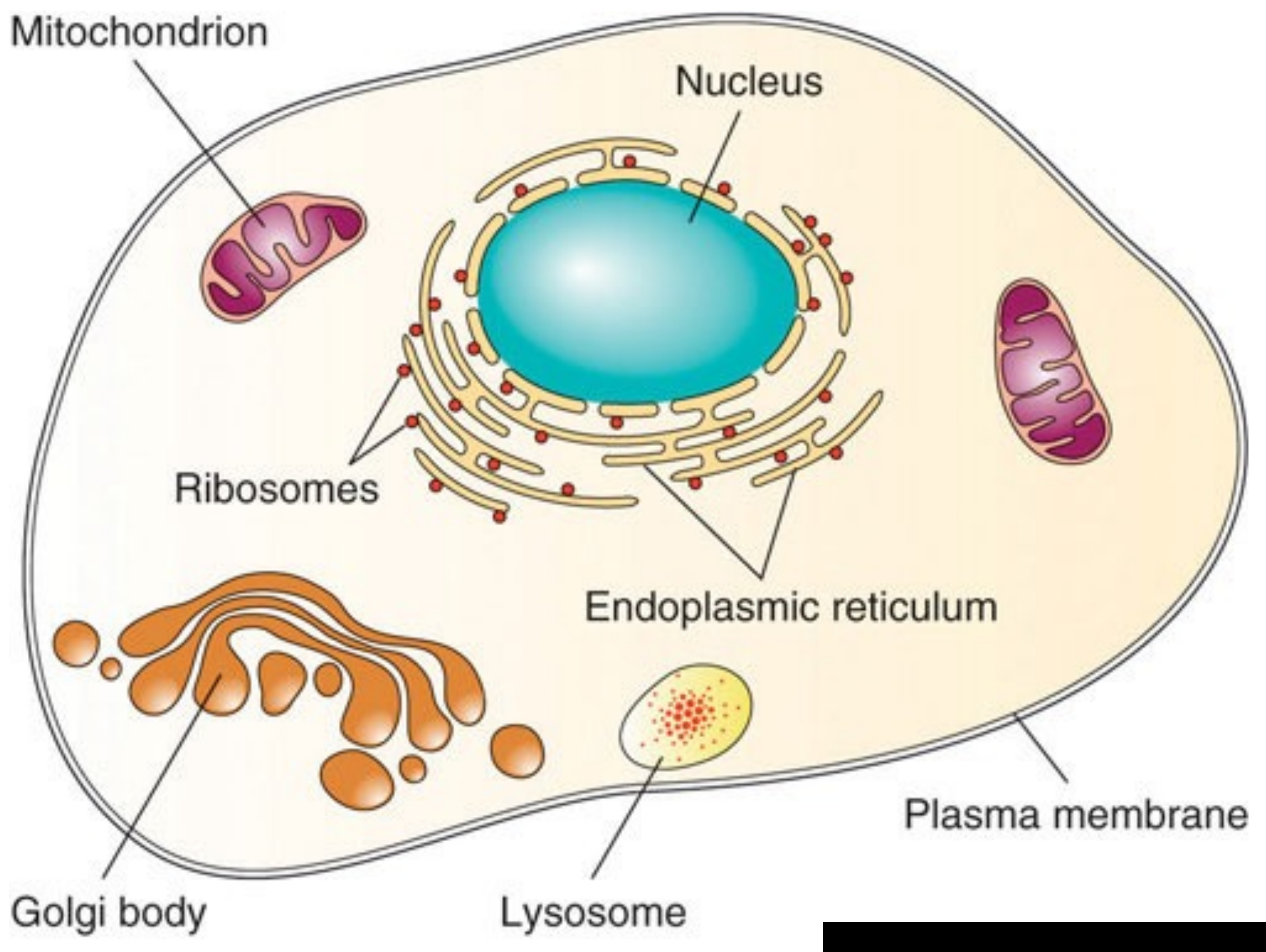
lipids: yellow

carbohydrates:

green

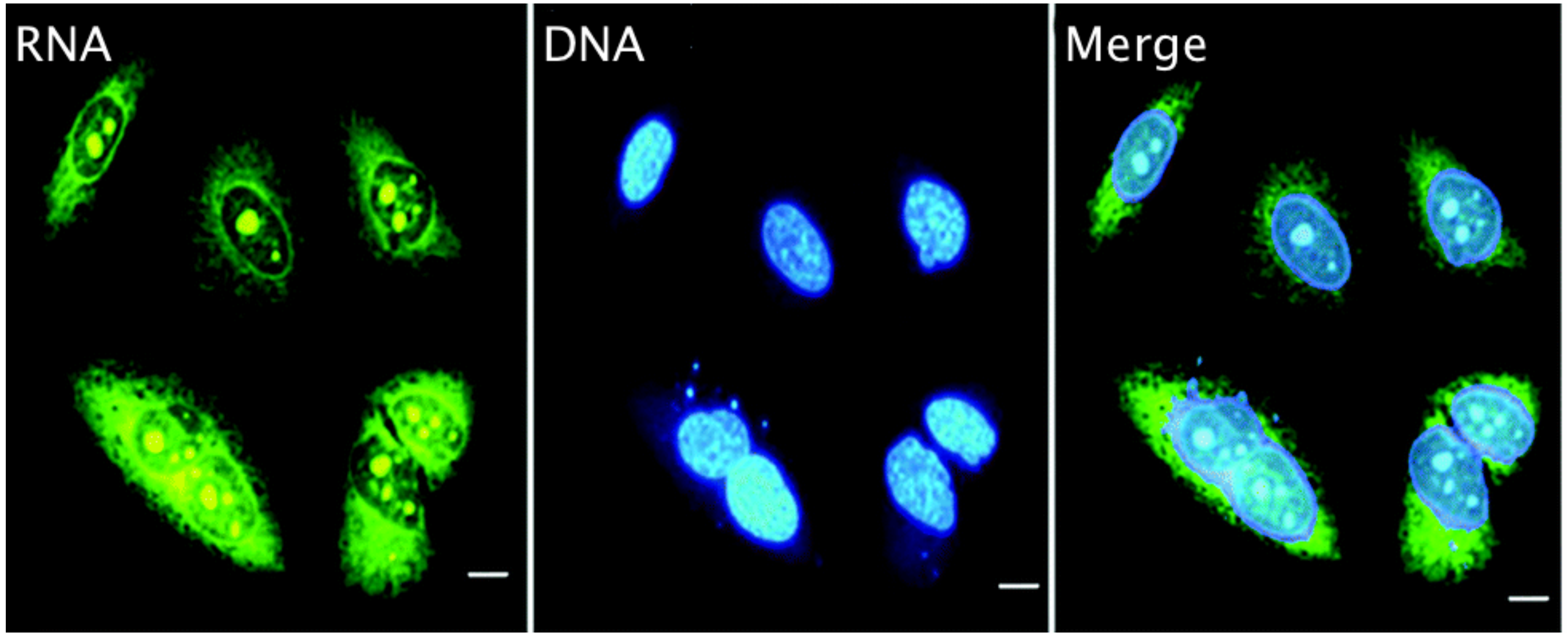
Ribosomes: magenta

Cells and their nucleic acids



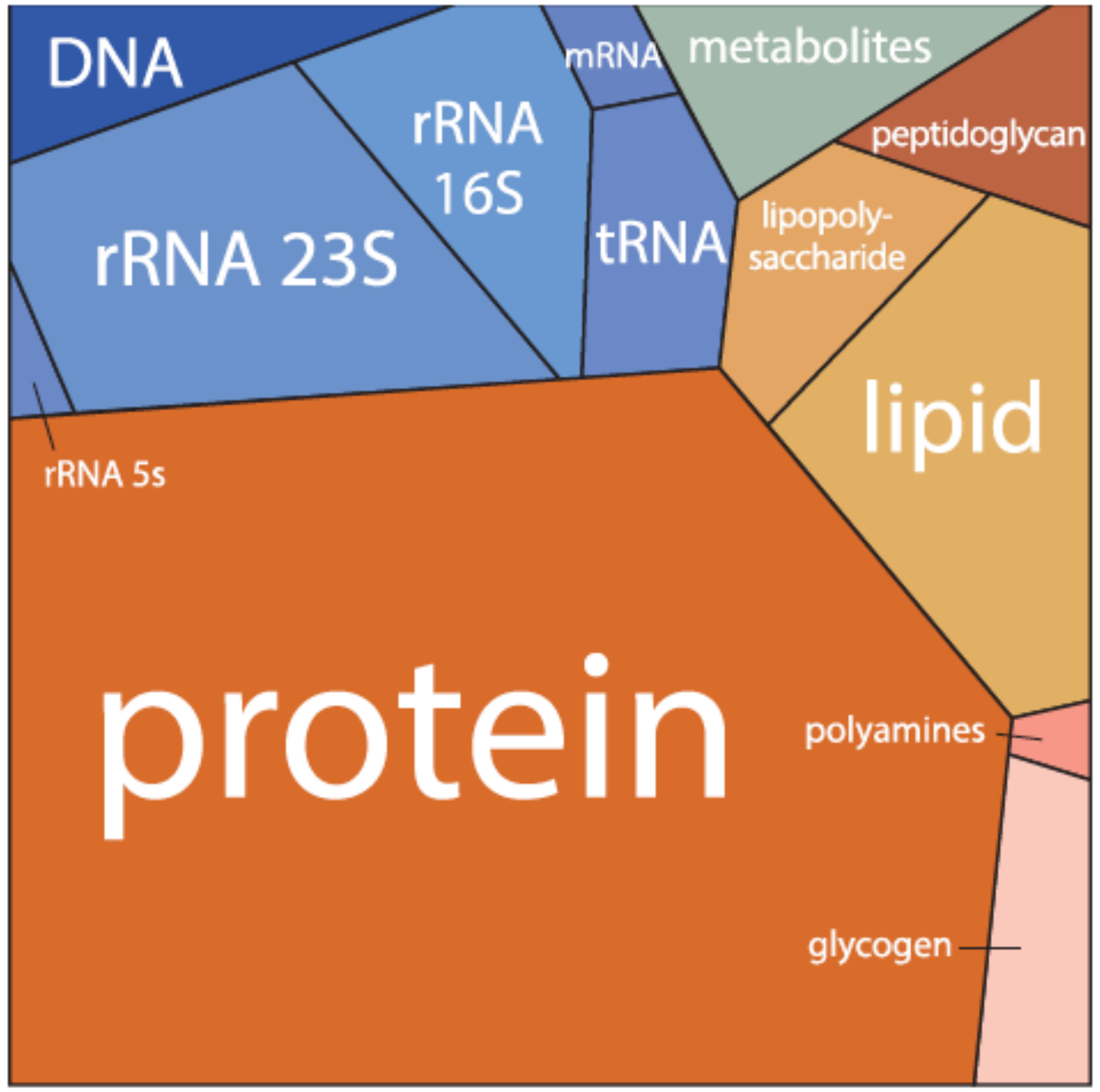
DNA:
99% genomic DNA
1% mtDNA

RNA:
80% ribosomal RNA
15% transfer RNA
5% messenger RNA and
other non-coding RNAs



Mass fractions in dry mass of E. coli growing at doubling time of 40 min

Density relative to water



object	density	BNID
DNA (unhydrated)	2.0	107858, 111208
RNA	2.0	111208
DNA (in solution with 7M CsCl)	1.7-1.8	107857
chromatin	1.4	106492
proteins	1.2-1.4	104272, 111208
chloroplasts	1.1-1.2	106492, 109442
mammalian viruses	1.1-1.2	106492, 106494, 109442
mitochondria	1.05-1.2	106492, 106494, 109442
hepatocyte	1.05-1.15	106494, 109441
erythrocyte	1.1	101502, 109441
E. coli	1.08-1.10	103875, 102239, 110096
budding yeast	1.08-1.10	106439
skeletal muscle	1.06	111214
synaptic vesicle	1.05	101502
HeLa	1.04-1.08	109441
fibroblast	1.03-1.05	101502, 106494, 109441
membrane (including proteins)	1.02-1.18	106492, 106494, 109442
phospholipid (+ cholesterol)	1.01	108142
adipocyte tissue (fat cells)	0.92	111213

Concentration of proteins and DNA/RNA in the nucleus that take up 20-30% of its volume

DNA	RNA	Protein
~ 15mg/ml (6pg DNA per cell, [19] nucleus ~1/10 of cell volume $4x10^{-9}$ cm^3 typical) [20] **~18.5mg/ml** (56mM nucleosome concentration, [21] 200 bp/nucleosome, 2bases/bp,1Mbase/3 30g. [22] **~19 mg/ml** [23] ~**20-31 mg/ml** (8.1-12.5pg/cell, [24] nucleus ~1/10 of cell volume $4x10^{-9}$ cm^3 typical) [20]	**~11 mg/ml** (5-25pg RNA per cell, [25] 18% in nucleus,[26] nucleus ~1/10 of cell volume $4x10^{-9}$ cm^3 typical).[20] **~12-15mg/ml** (27.1-33.1pg/cell,[24] 18% in nucleus,[26] nucleus ~1/10 of cell volume $4x10^{-9}$ cm^3 typical).[20]	**~106-215 mg/ml** in various regions of the nucleus.[27] **~108mg/ml** (6pg DNA per cell,[20] protein mass 72X DNA mass and cell volume $4x10^{-9}$ cm3 typical).[20] **~200-300mg/ml** in E.coli.[28]

Zeskind Nature Methods 4, 567 - 569. https://dx.doi.org/10.1038/nmeth1053

The density of DNA and proteins expressed as their “partial specific volume”

The artial specific volume *v* bar ($\bar{v}$)

$$\bar{v} = \left(\frac{\partial V}{\partial m} \right)_{T,p}$$

V: volume
m: mass
ρ: density

$$\bar{v} = \frac{1}{\rho}$$

$\bar{v}$ for protein: 0.73 ml g^{-1}
$\bar{v}$ for DNA: 0.55 ml g^{-1}

Calculating volume and friction coefficient of a spherical protein

$$V_{\text{sphere}} = \frac{4}{3}\pi r^3$$

The volume of a sphere with radius *r*

$$V_{\text{sphere}} = \frac{M\bar{v}}{N_{\text{A}}}$$

Volume of a spherical macromolecule

$$f = 6\pi\eta r$$

Stokes law: The frictional coefficient *f* of a sphere with radius r in a solution with viscosity η

$\bar{v}$: Partial specific volume, 0.73 ml/g (protein) or 0.55 ml/g (DNA)
M: Molecular mass of the macromolecule
$\boldsymbol{N_{A}}$: Avogadros number, $6.022 \cdot 10^{23}$
ρ: density of the solution, 0.9982 g ml^{-1} for water at 20 ˚C
η: viscosity of the solution, 1.002 mPa second for water at 20 ˚C

How many mRNA molecules are in an *E. coli* cell?

a) $10^3 - 10^4$
b) $10^5 - 10^6$
c) $10^7 - 10^8$
d) $10^9 - 10^{10}$

and in a human cell?

a) same
b) 10x
c) 100x
d) 1000x

Strategies to answer the question

- Look it up in “Cell biology by the numbers book” at http://book.bionumbers.org
- Look it up in the Bionumbers database at http://bionumbers.hms.harvard.edu
- Ask the AI-powered search engine perplexity.ai
- Derive the number from other data that you have

One strategy to derive the number from other data

- Calculate the number of proteins per cell
- Estimate the protein synthesis rate R from the division time
- Estimate the protein production rate r per mRNA
- Calculate the number of mRNA = R/r

How many proteins are in a cell?

protein mass per volume (≈ 0.2 g/ml)

$$\underbrace{\frac{N}{V}}_{\text{number of proteins per cell volume}} = \frac{C_p}{l_{aa} \times m_{aa}}$$

mass aa (≈ 100 Da)

aa per protein ($\approx 400 \frac{\text{aa}}{\text{protein}}$)

Avogadro's number

$$\frac{N}{V} = \frac{0.2\,[\text{g/ml}] \times 6 \times 10^{23} \left[\frac{\text{Da}}{\text{g}}\right] \times 10^{-12} \left[\frac{\text{ml}}{\mu\text{m}^3}\right]}{400 \left[\frac{\text{aa}}{\text{protein}}\right] \times 100 \left[\frac{\text{Da}}{\text{aa}}\right]} \approx \boxed{3 \times 10^6 \frac{\text{proteins}}{\mu\text{m}^3}}$$

organism	characteristic volume	number of proteins
E. coli	1 μm^3	$\approx 3 \times 10^6$
budding yeast	≈ 30 μm^3	$\approx 100 \times 10^6$
HeLa cell line	$\approx 3{,}000$ μm^3	$\approx 10 \times 10^9$

Estimating the protein synthesis rate *R*

$$\frac{N_{protein}}{V} \approx 3 \times 10^{6} \text{ proteins}/\mu m^{3}$$

rate of protein production per cell

$$R = \frac{N_{protein}}{\tau}$$

doubling time

bacteria

$V \approx 1 \mu m^{3}$,
$\tau \approx 1$ h

$$\approx \frac{3 \times 10^{6} \text{ proteins}}{3{,}000 \text{ s}} \approx 10^{3} \text{ proteins/s}$$

mammalian cell

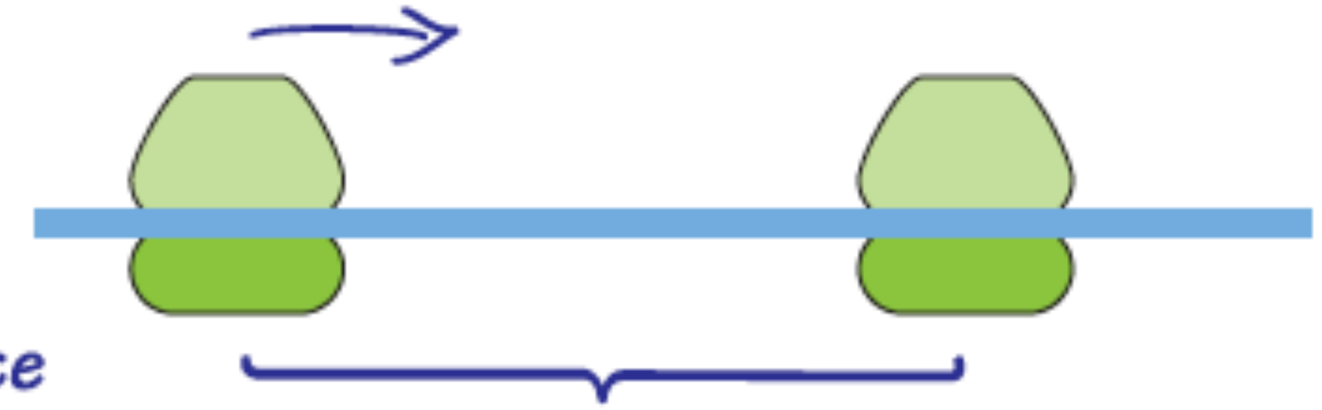
$V \approx 3{,}000 \mu m^{3}$,
$\tau \approx 24$ h

$$\approx \frac{10^{10} \text{ proteins}}{10^{5} \text{ s}} \approx 10^{5} \text{ proteins/s}$$

Estimate the protein production rate *r* per mRNA

rate of translation

$v \approx 10\text{–}50$ bases/s



distance between ribosomes

$d \approx 30\text{–}300$ bases

ribosomes touching

few ribosomes per mRNA (supported by observations)

rate of protein production per mRNA

$$r = \frac{v}{d} \approx 0.1\text{–}1 \text{ protein/mRNA/s}$$

Calculate the number of mRNAs

$$N_{mRNA} = \frac{R}{r}$$

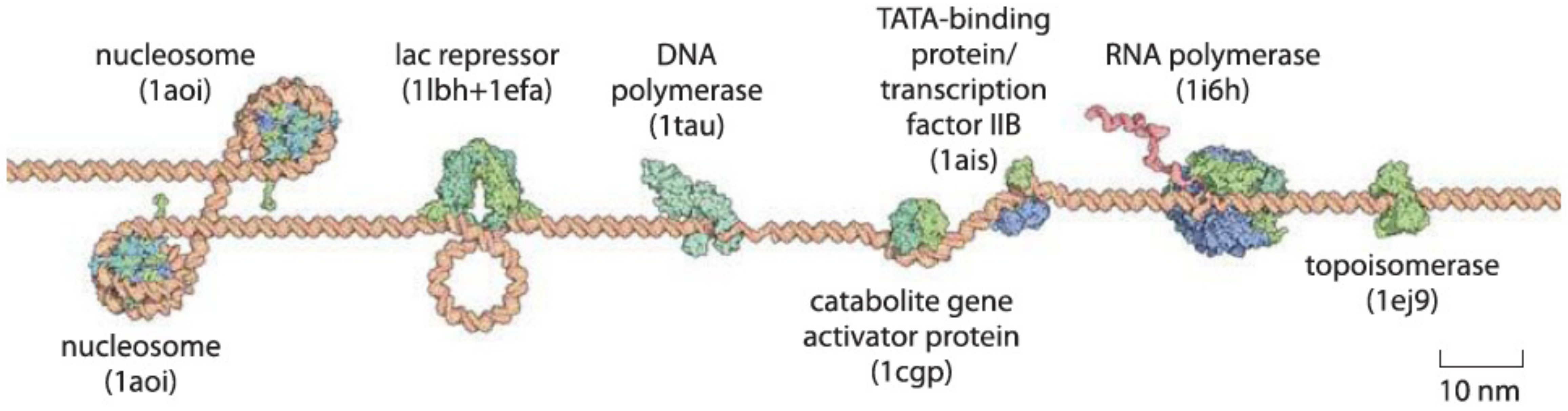
bacteria

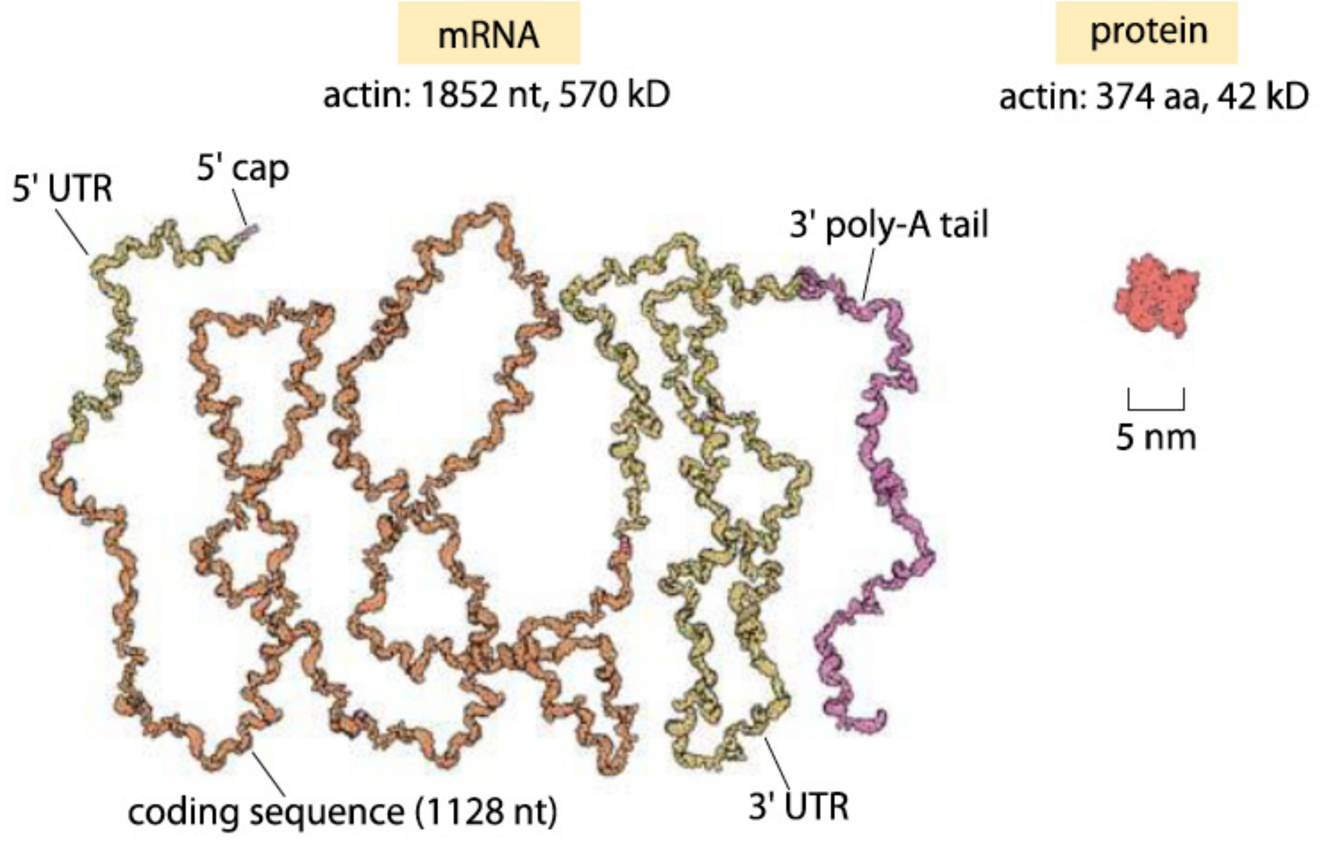
$$\approx \frac{10^3 \text{ proteins/s}}{0.1-1 \text{ protein/mRNA/s}} \approx \boxed{10^3\text{-}10^4 \text{ mRNA/bacterial cell}}$$

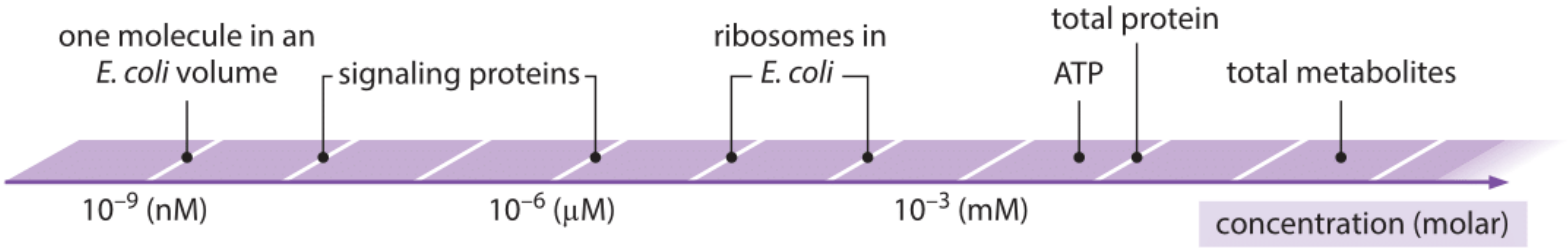
mammalian cell

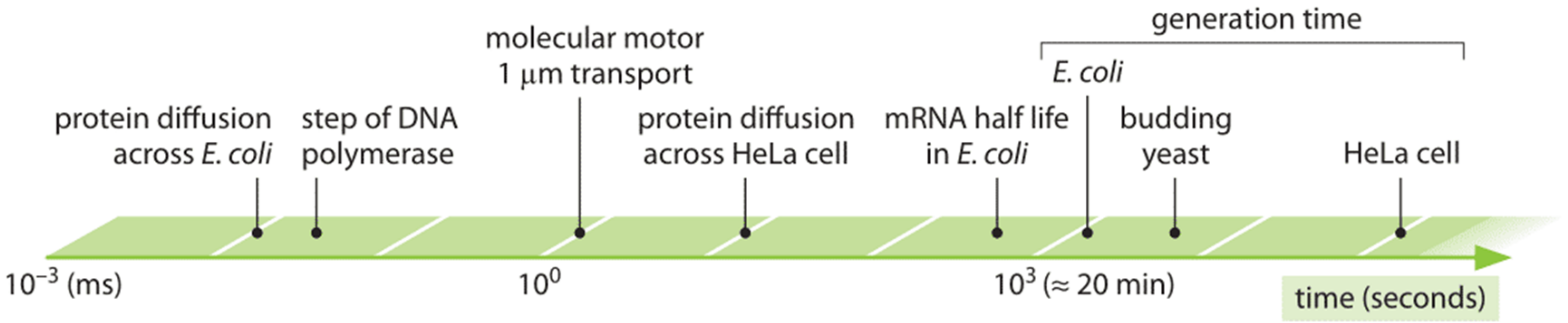
$$\approx \frac{10^5 \text{ proteins/s}}{0.1-1 \text{ protein/mRNA/s}} \approx \boxed{10^5\text{-}10^6 \text{ mRNA/mammalian cell}}$$

Summary


nucleosome
(1aoi)
lac repressor
(1lbh+1efa)
DNA
polymerase
(1tau)
TATA-binding
protein/
transcription
factor IIB
(1ais)
RNA polymerase
(1i6h)
nucleosome
(1aoi)
catabolite gene
activator protein
(1cgp)
topoisomerase
(1ej9)
10 nm


mRNA
actin: 1852 nt, 570 kD
protein
actin: 374 aa, 42 kD
5' UTR
5' cap
3' poly-A tail
5 nm
coding sequence (1128 nt)
3' UTR


one molecule in an
E. coli volume
signaling proteins
ribosomes in
E. coli
total protein
ATP
total metabolites
10^{-9} (nM)
10^{-6} (μM)
10^{-3} (mM)
concentration (molar)


generation time
E. coli
molecular motor
1 μm transport
protein diffusion
across *E. coli*
step of DNA
polymerase
protein diffusion
across HeLa cell
mRNA half life
in *E. coli*
budding
yeast
HeLa cell
10^{-3} (ms)
10^{0}
10^{3} (≈ 20 min)
time (seconds)