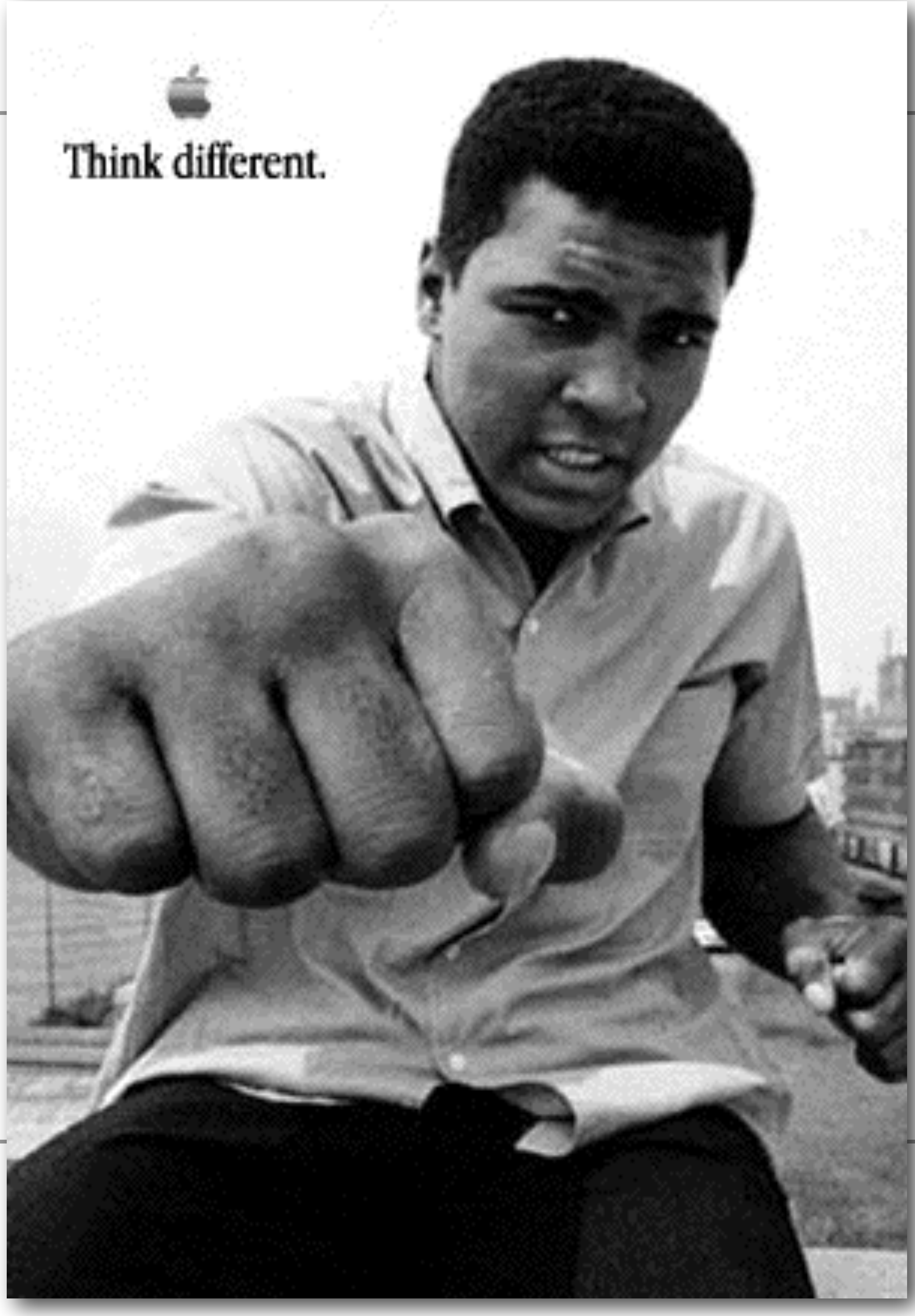
LB145-Fall *2023*

!



1. **Pick up** Name Folder
- Pick up name folder and set it up at seat.

2. **Sit** with your group.
- laptops on outer perimeter (avoid distracting)

3. **Clicker** Attendance
- Launch your Top Hat, and get ready to click.

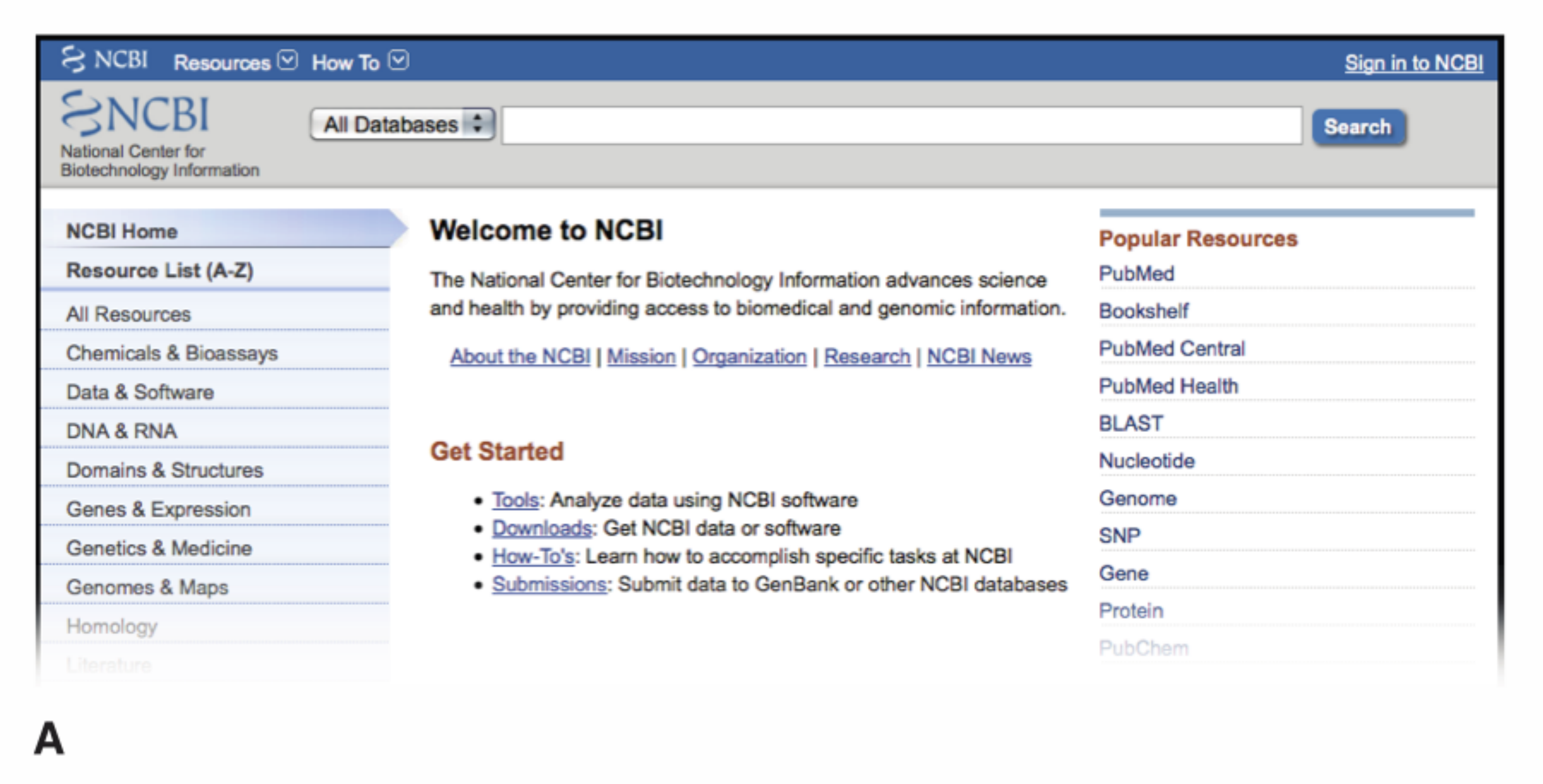
Integrating Concepts in Biology

2.4 Can cells pick and choose information?

Biology Learning Objectives

- Apply the genetic code to deduce the protein encoded by a mRNA.
- Review different examples of non-linear information, and determine how they apply to the central dogma.

NCBI Biological Information Database



Welcome to NCBI

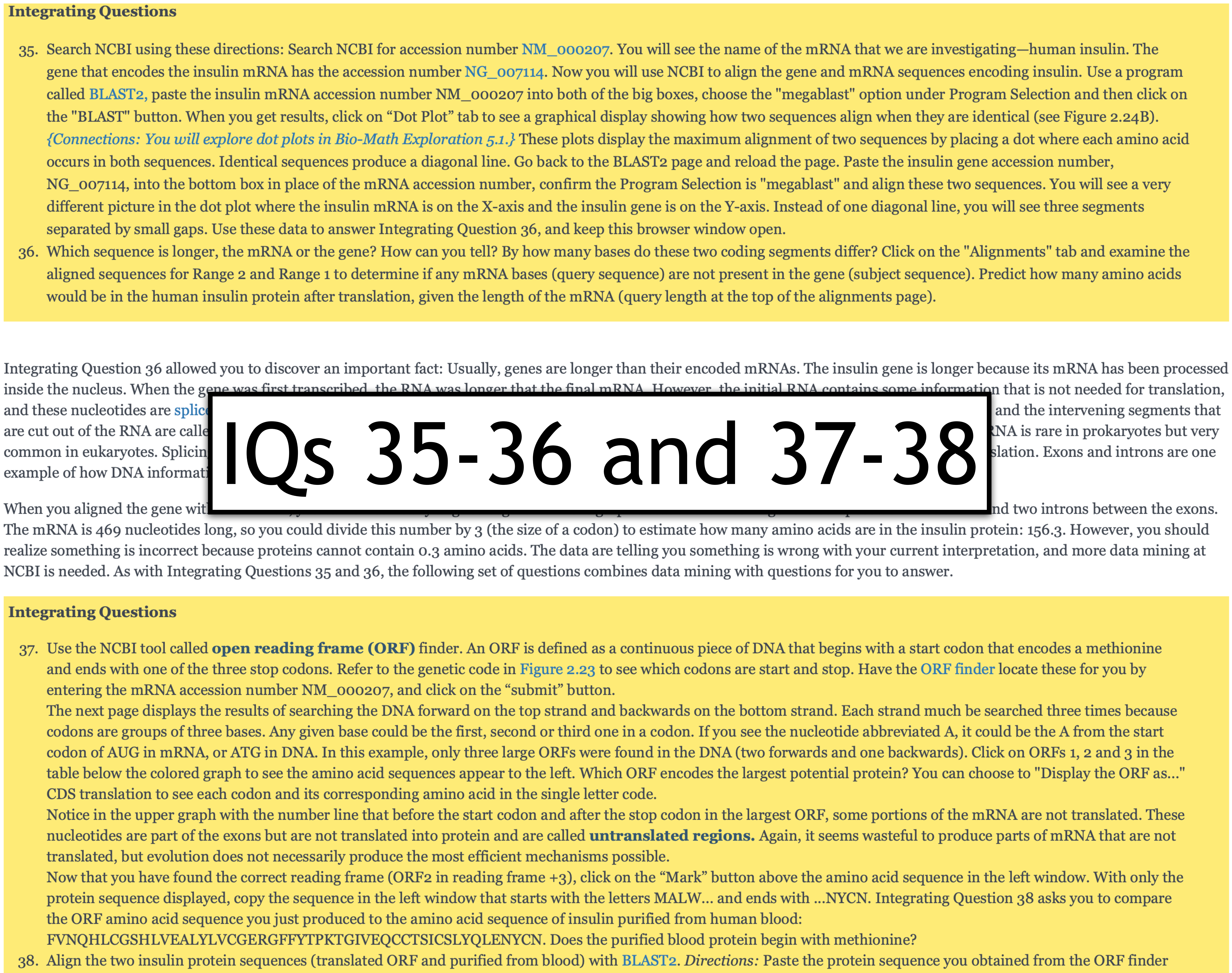
The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | Research | NCBI News

Get Started

- Tools: Analyze data using NCBI software
- Downloads: Get NCBI data or software
- How-To's: Learn how to accomplish specific tasks at NCBI
- Submissions: Submit data to GenBank or other NCBI databases

Fig. 2.24A



Integrating Questions

35. Search NCBI using these directions: Search NCBI for accession number NM_000207. You will see the name of the mRNA that we are investigating—human insulin. The gene that encodes the insulin mRNA has the accession number NG_007114. Now you will use NCBI to align the gene and mRNA sequences encoding insulin. Use a program called BLAST2, paste the insulin mRNA accession number NM_000207 into both of the big boxes, choose the "megablast" option under Program Selection and then click on the "BLAST" button. When you get results, click on “Dot Plot” tab to see a graphical display showing how two sequences align when they are identical (see Figure 2.24B). *{Connections: You will explore dot plots in Bio-Math Exploration 5.1.}* These plots display the maximum alignment of two sequences by placing a dot where each amino acid occurs in both sequences. Identical sequences produce a diagonal line. Go back to the BLAST2 page and reload the page. Paste the insulin gene accession number, NG_007114, into the bottom box in place of the mRNA accession number, confirm the Program Selection is "megablast" and align these two sequences. You will see a very different picture in the dot plot where the insulin mRNA is on the X-axis and the insulin gene is on the Y-axis. Instead of one diagonal line, you will see three segments separated by small gaps. Use these data to answer Integrating Question 36, and keep this browser window open.
36. Which sequence is longer, the mRNA or the gene? How can you tell? By how many bases do these two coding segments differ? Click on the "Alignments" tab and examine the aligned sequences for Range 2 and Range 1 to determine if any mRNA bases (query sequence) are not present in the gene (subject sequence). Predict how many amino acids would be in the human insulin protein after translation, given the length of the mRNA (query length at the top of the alignments page).

IQs 35-36 and 37-38

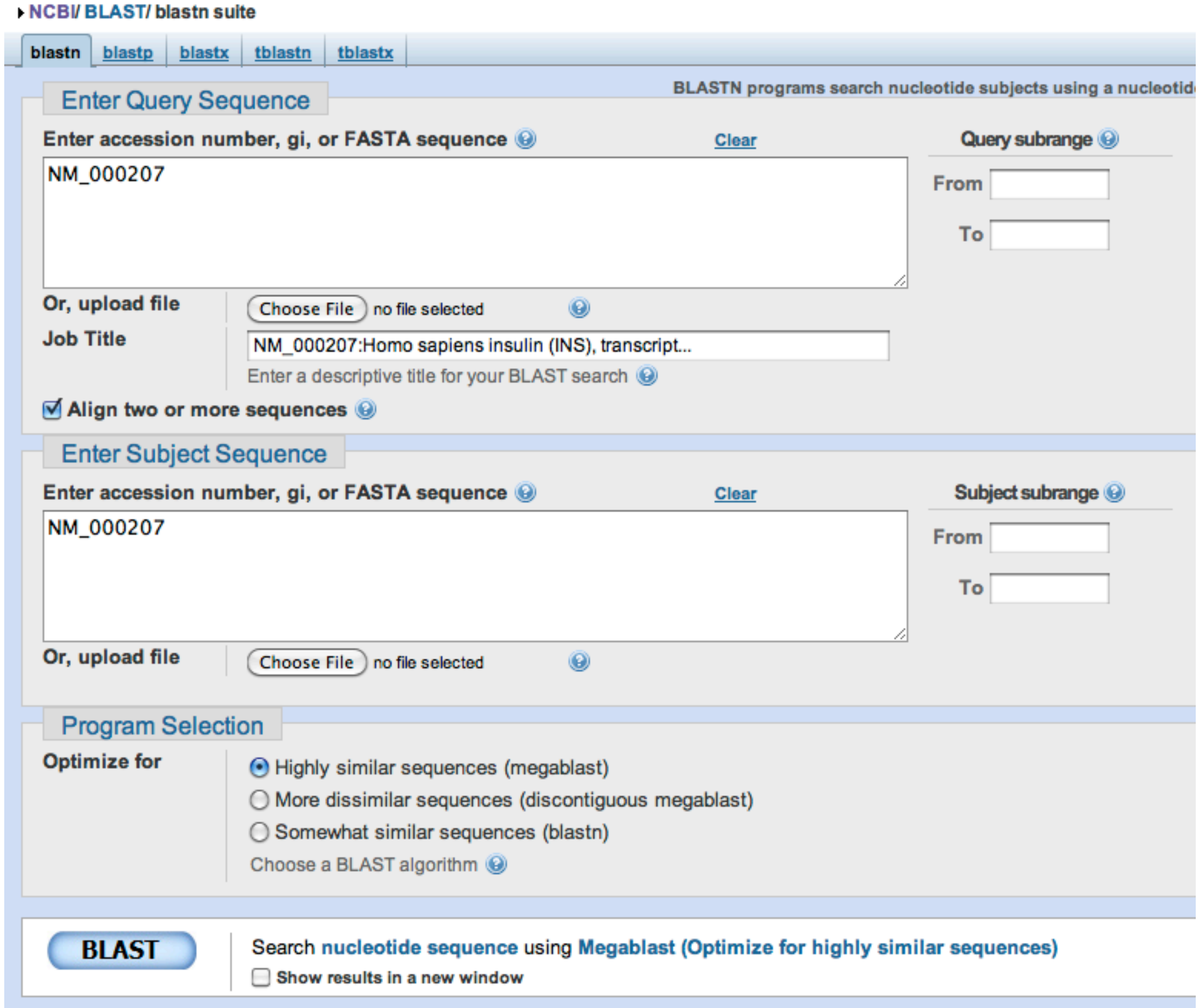
Integrating Question 36 allowed you to discover an important fact: Usually, genes are longer than their encoded mRNAs. The insulin gene is longer because its mRNA has been processed inside the nucleus. When the gene was first transcribed, the RNA was longer that the final mRNA. However, the initial RNA contains some information that is not needed for translation, and these nucleotides are splic … and the intervening segments that are cut out of the RNA are calle … NA is rare in prokaryotes but very common in eukaryotes. Splicin … slation. Exons and introns are one example of how DNA informati …

When you aligned the gene wit … nd two introns between the exons. The mRNA is 469 nucleotides long, so you could divide this number by 3 (the size of a codon) to estimate how many amino acids are in the insulin protein: 156.3. However, you should realize something is incorrect because proteins cannot contain 0.3 amino acids. The data are telling you something is wrong with your current interpretation, and more data mining at NCBI is needed. As with Integrating Questions 35 and 36, the following set of questions combines data mining with questions for you to answer.

Integrating Questions

37. Use the NCBI tool called **open reading frame (ORF)** finder. An ORF is defined as a continuous piece of DNA that begins with a start codon that encodes a methionine and ends with one of the three stop codons. Refer to the genetic code in Figure 2.23 to see which codons are start and stop. Have the ORF finder locate these for you by entering the mRNA accession number NM_000207, and click on the “submit” button.
    The next page displays the results of searching the DNA forward on the top strand and backwards on the bottom strand. Each strand much be searched three times because codons are groups of three bases. Any given base could be the first, second or third one in a codon. If you see the nucleotide abbreviated A, it could be the A from the start codon of AUG in mRNA, or ATG in DNA. In this example, only three large ORFs were found in the DNA (two forwards and one backwards). Click on ORFs 1, 2 and 3 in the table below the colored graph to see the amino acid sequences appear to the left. Which ORF encodes the largest potential protein? You can choose to "Display the ORF as..." CDS translation to see each codon and its corresponding amino acid in the single letter code.
    Notice in the upper graph with the number line that before the start codon and after the stop codon in the largest ORF, some portions of the mRNA are not translated. These nucleotides are part of the exons but are not translated into protein and are called **untranslated regions.** Again, it seems wasteful to produce parts of mRNA that are not translated, but evolution does not necessarily produce the most efficient mechanisms possible.
    Now that you have found the correct reading frame (ORF2 in reading frame +3), click on the “Mark” button above the amino acid sequence in the left window. With only the protein sequence displayed, copy the sequence in the left window that starts with the letters MALW... and ends with ...NYCN. Integrating Question 38 asks you to compare the ORF amino acid sequence you just produced to the amino acid sequence of insulin purified from human blood:
    FVNQHLCGSHLVEALYLVCGERGFFYTPKTGIVEQCCTSICSLYQLENYCN. Does the purified blood protein begin with methionine?
38. Align the two insulin protein sequences (translated ORF and purified from blood) with BLAST2. *Directions:* Paste the protein sequence you obtained from the ORF finder

BLAST Align Nucleotide Sequences



BLASTN programs search nucleotide subjects using a nucleotid

Enter Query Sequence

Enter accession number, gi, or FASTA sequence

NM_000207

Query subrange

From:

To:

Or, upload file: no file selected

Job Title: NM_000207:Homo sapiens insulin (INS), transcript...

Enter a descriptive title for your BLAST search

☑ Align two or more sequences

Enter Subject Sequence

Enter accession number, gi, or FASTA sequence

NM_000207

Subject subrange

From:

To:

Or, upload file: no file selected

Program Selection

Optimize for

● Highly similar sequences (megablast)
○ More dissimilar sequences (discontiguous megablast)
○ Somewhat similar sequences (blastn)

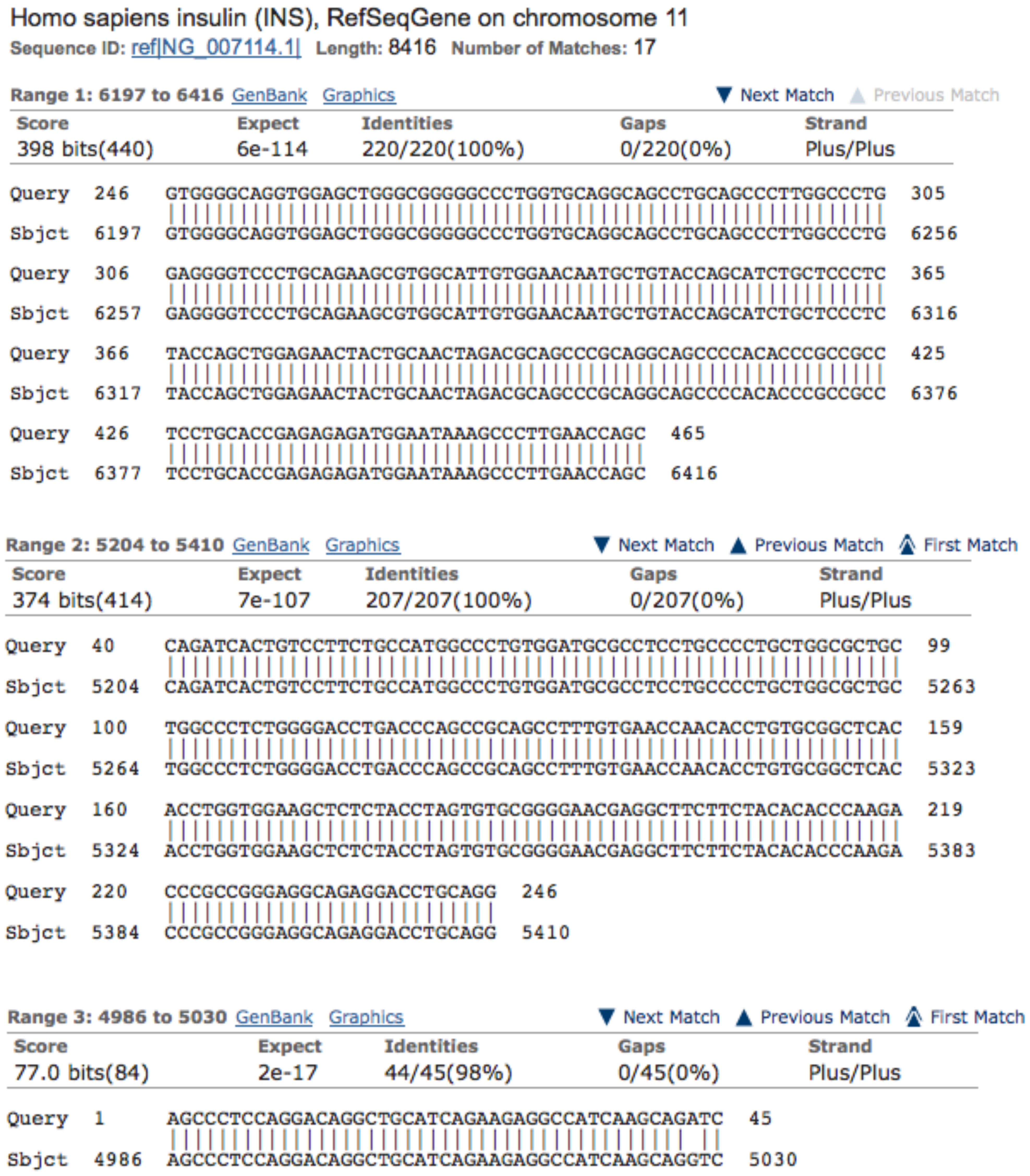
Choose a BLAST algorithm

Search nucleotide sequence using Megablast (Optimize for highly similar sequences)

☐ Show results in a new window

IQ #35

What did you get with mRNA vs gene?



Homo sapiens insulin (INS), RefSeqGene on chromosome 11
Sequence ID: ref|NG_007114.1| Length: 8416 Number of Matches: 17

1st

```
Range 1: 6197 to 6416 GenBank Graphics                    ▼ Next Match  ▲ Previous Match
Score           Expect    Identities        Gaps         Strand
398 bits(440)   6e-114    220/220(100%)     0/220(0%)    Plus/Plus

Query  246   GTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCTG  305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  6197  GTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGCCCTTGGCCCTG  6256

Query  306   GAGGGGTCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCTC  365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  6257  GAGGGGTCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCTC  6316

Query  366   TACCAGCTGGAGAACTACTGCAACTAGACGCAGCCCGCAGGCAGCCCCACACCCGCCGCC  425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  6317  TACCAGCTGGAGAACTACTGCAACTAGACGCAGCCCGCAGGCAGCCCCACACCCGCCGCC  6376

Query  426   TCCTGCACCGAGAGAGATGGAATAAAGCCCTTGAACCAGC  465
             ||||||||||||||||||||||||||||||||||||||||
Sbjct  6377  TCCTGCACCGAGAGAGATGGAATAAAGCCCTTGAACCAGC  6416
```

2nd

```
Range 2: 5204 to 5410 GenBank Graphics                    ▼ Next Match  ▲ Previous Match  ⩓ First Match
Score           Expect    Identities        Gaps         Strand
374 bits(414)   7e-107    207/207(100%)     0/207(0%)    Plus/Plus

Query  40    CAGATCACTGTCCTTCTGCCATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGC  99
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5204  CAGATCACTGTCCTTCTGCCATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGC  5263

Query  100   TGGCCCTCTGGGGACCTGACCCAGCCGCAGCCTTTGTGAACCAACACCTGTGCGGCTCAC  159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5264  TGGCCCTCTGGGGACCTGACCCAGCCGCAGCCTTTGTGAACCAACACCTGTGCGGCTCAC  5323

Query  160   ACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCAAGA  219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5324  ACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCAAGA  5383

Query  220   CCCGCCGGGAGGCAGAGGACCTGCAGG  246
             |||||||||||||||||||||||||||
Sbjct  5384  CCCGCCGGGAGGCAGAGGACCTGCAGG  5410
```

3rd

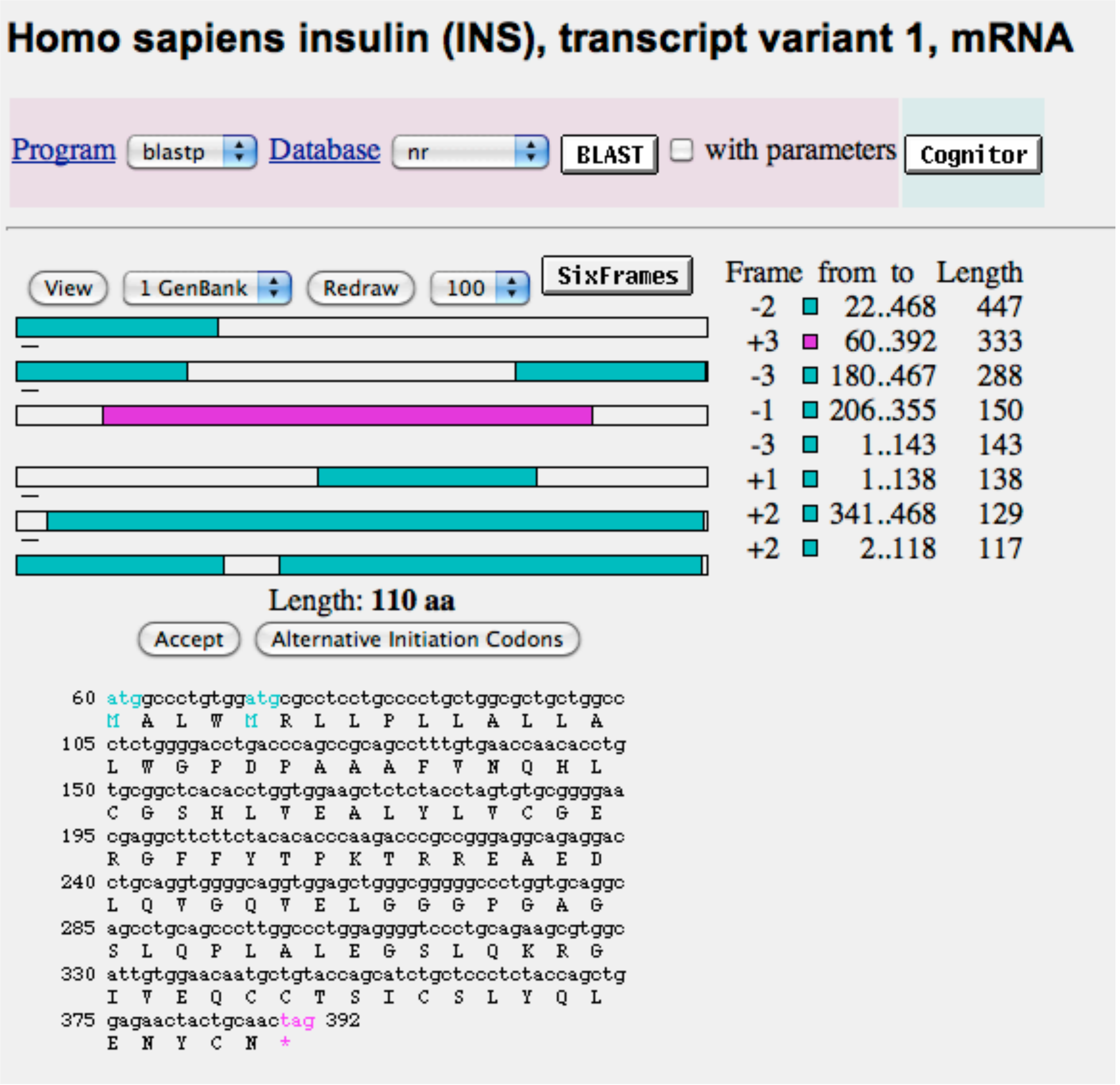
```
Range 3: 4986 to 5030 GenBank Graphics                    ▼ Next Match  ▲ Previous Match  ⩓ First Match
Score           Expect    Identities        Gaps         Strand
77.0 bits(84)   2e-17     44/45(98%)        0/45(0%)     Plus/Plus

Query  1     AGCCCTCCAGGACAGGCTGCATCAGAAGAGGCCATCAAGCAGATC  45
             |||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  4986  AGCCCTCCAGGACAGGCTGCATCAGAAGAGGCCATCAAGCAGGTC  5030
```

3 segments
with nearly
perfect
sequence
alignments

IQ #35

ORF (AUG!->STOP) Finder



Homo sapiens insulin (INS), transcript variant 1, mRNA

Program: blastp Database: nr ☐ with parameters

View: 1 GenBank 100

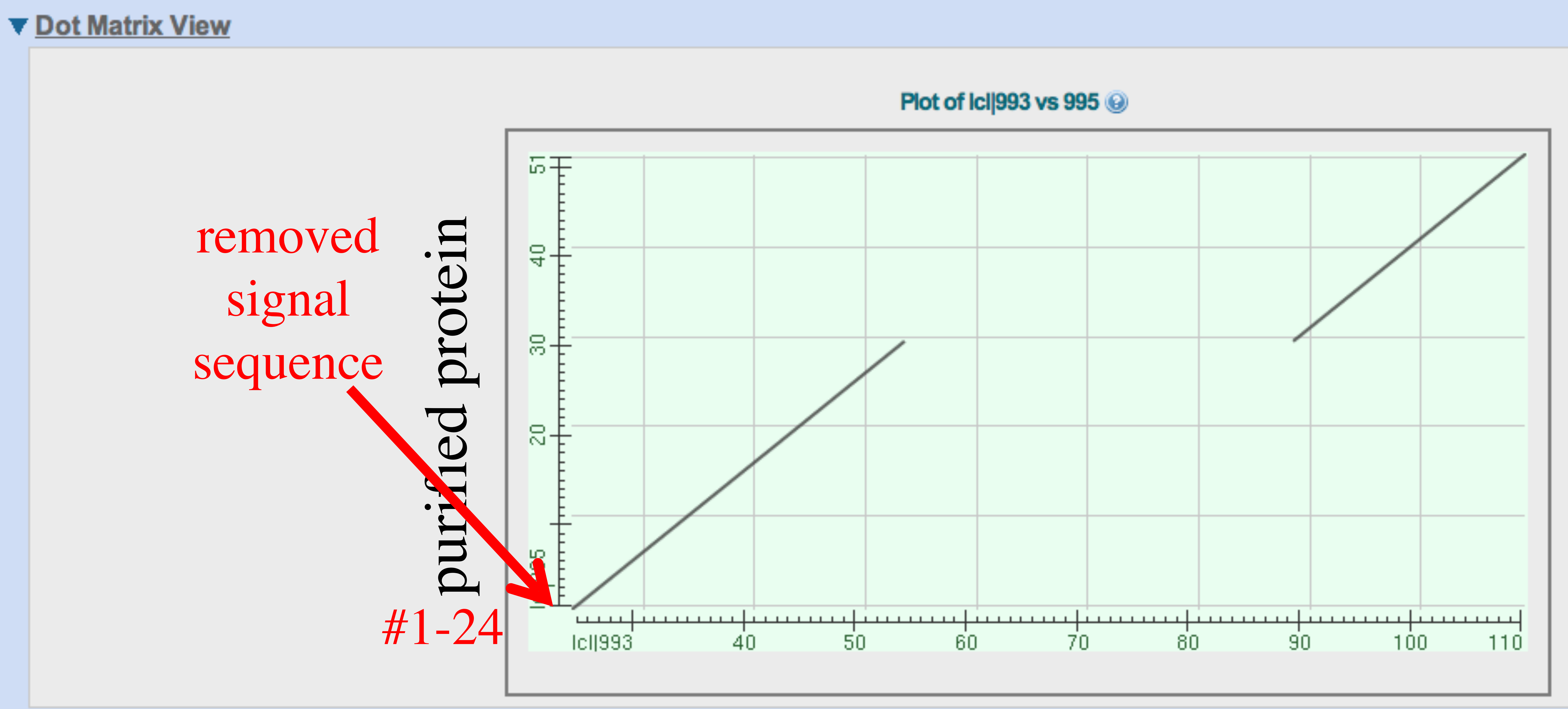
Frame		from to	Length
-2	🟦	22..468	447
+3	🟪	60..392	333
-3	🟦	180..467	288
-1	🟦	206..355	150
-3	🟦	1..143	143
+1	🟦	1..138	138
+2	🟦	341..468	129
+2	🟦	2..118	117

Length: **110 aa**

```
 60 atggccctgtggatgcgcctcctgcccctgctggcgctgctggcc
    M  A  L  W  M  R  L  L  P  L  L  A  L  L  A
105 ctctggggacctgacccagccgcagcctttgtgaaccaacacctg
    L  W  G  P  D  P  A  A  A  F  V  N  Q  H  L
150 tgcggctcacacctggtggaagctctctacctagtgtgcggggaa
    C  G  S  H  L  V  E  A  L  Y  L  V  C  G  E
195 cgaggcttcttctacacacccaagacccgccgggaggcagaggac
    R  G  F  F  Y  T  P  K  T  R  R  E  A  E  D
240 ctgcaggtggggcaggtggagctgggcgggggccctggtgcaggc
    L  Q  V  G  Q  V  E  L  G  G  G  P  G  A  G
285 agcctgcagcccttggccctggaggggtccctgcagaagcgtggc
    S  L  Q  P  L  A  L  E  G  S  L  Q  K  R  G
330 attgtggaacaatgctgtaccagcatctgctccctctaccagctg
    I  V  E  Q  C  C  T  S  I  C  S  L  Y  Q  L
375 gagaactactgcaactag 392
    E  N  Y  C  N  *
```

IQ #37

Insulin Translated *vs* Circulating



removed
signal
sequence

purified protein

#1-24

translated ORF

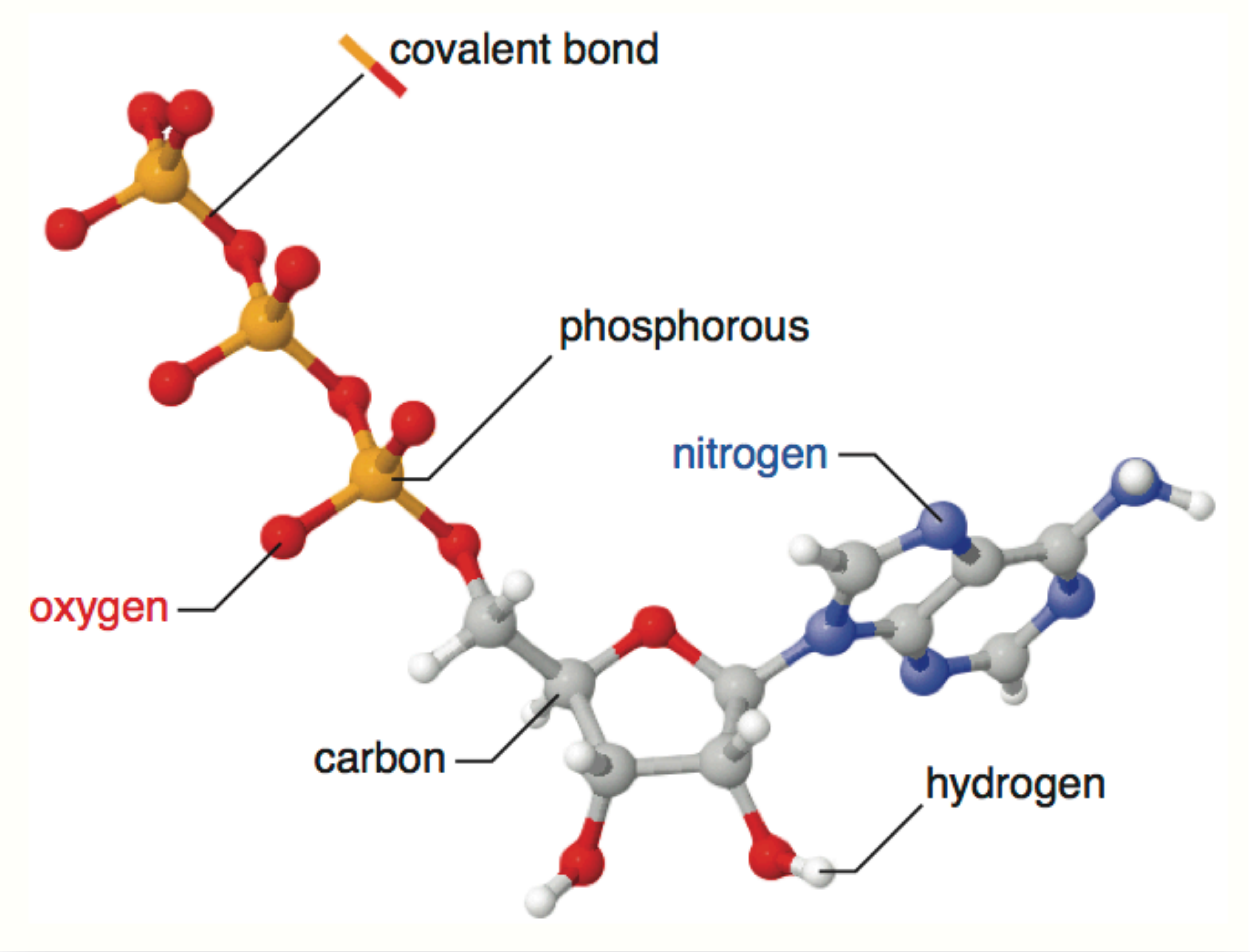
IQ #38

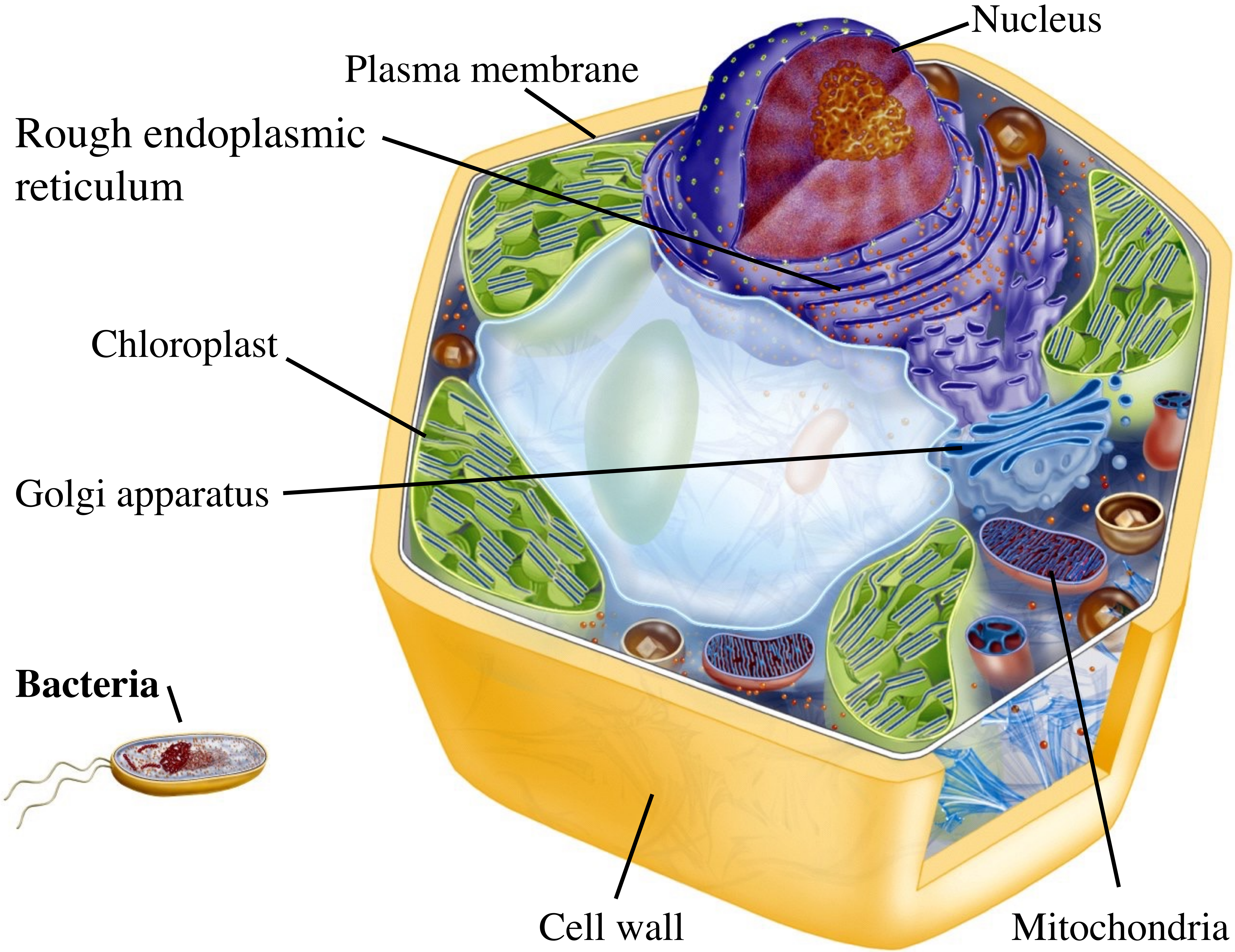
Integrating Concepts in Biology

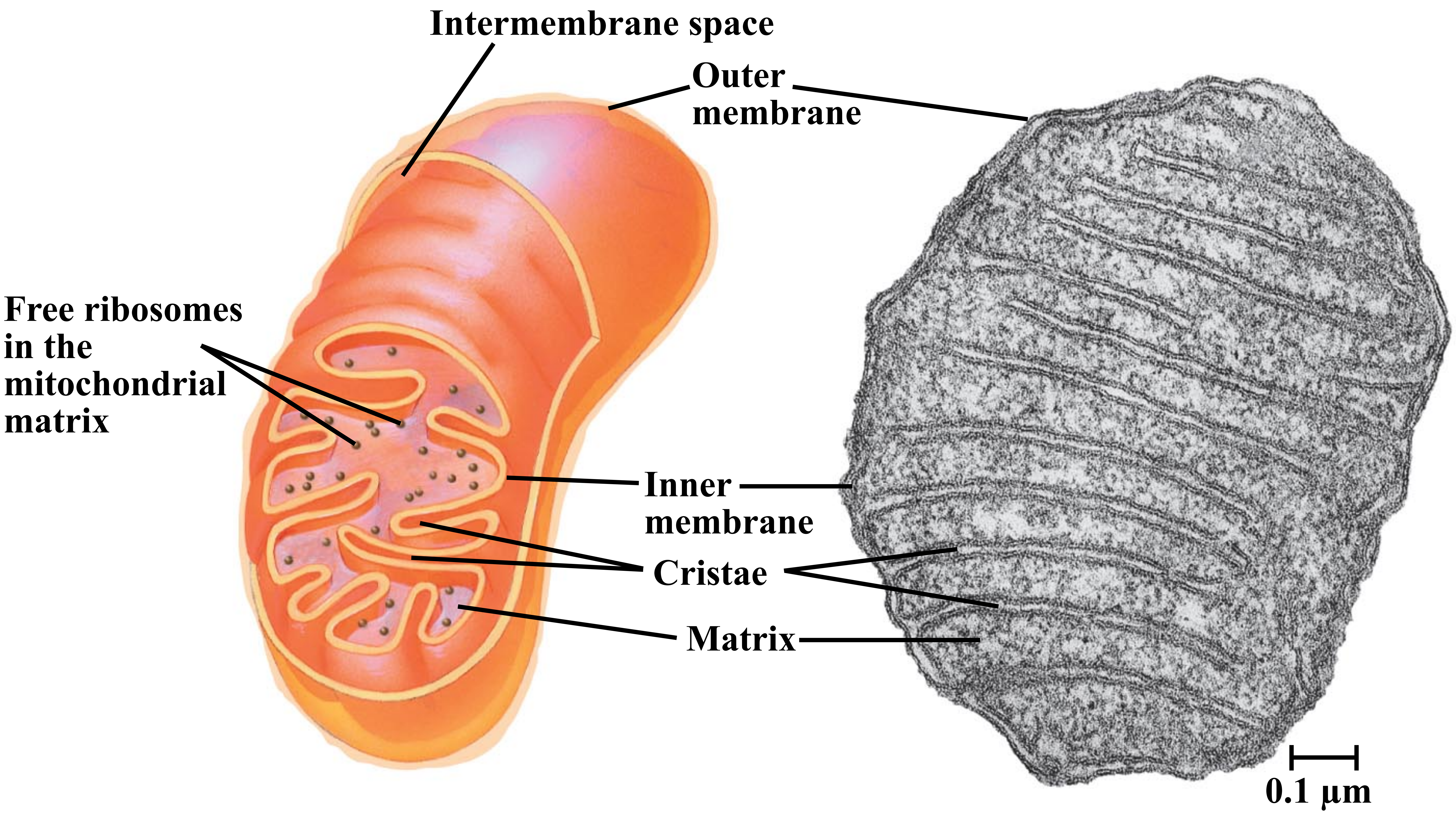
10.1 How do y?

Biology Learning Ob

- Discuss energy tran onnection between covalent b




Nucleus
Plasma membrane
Rough endoplasmic reticulum
Chloroplast
Golgi apparatus
Bacteria
Cell wall
Mitochondria



Intermembrane space

Outer membrane

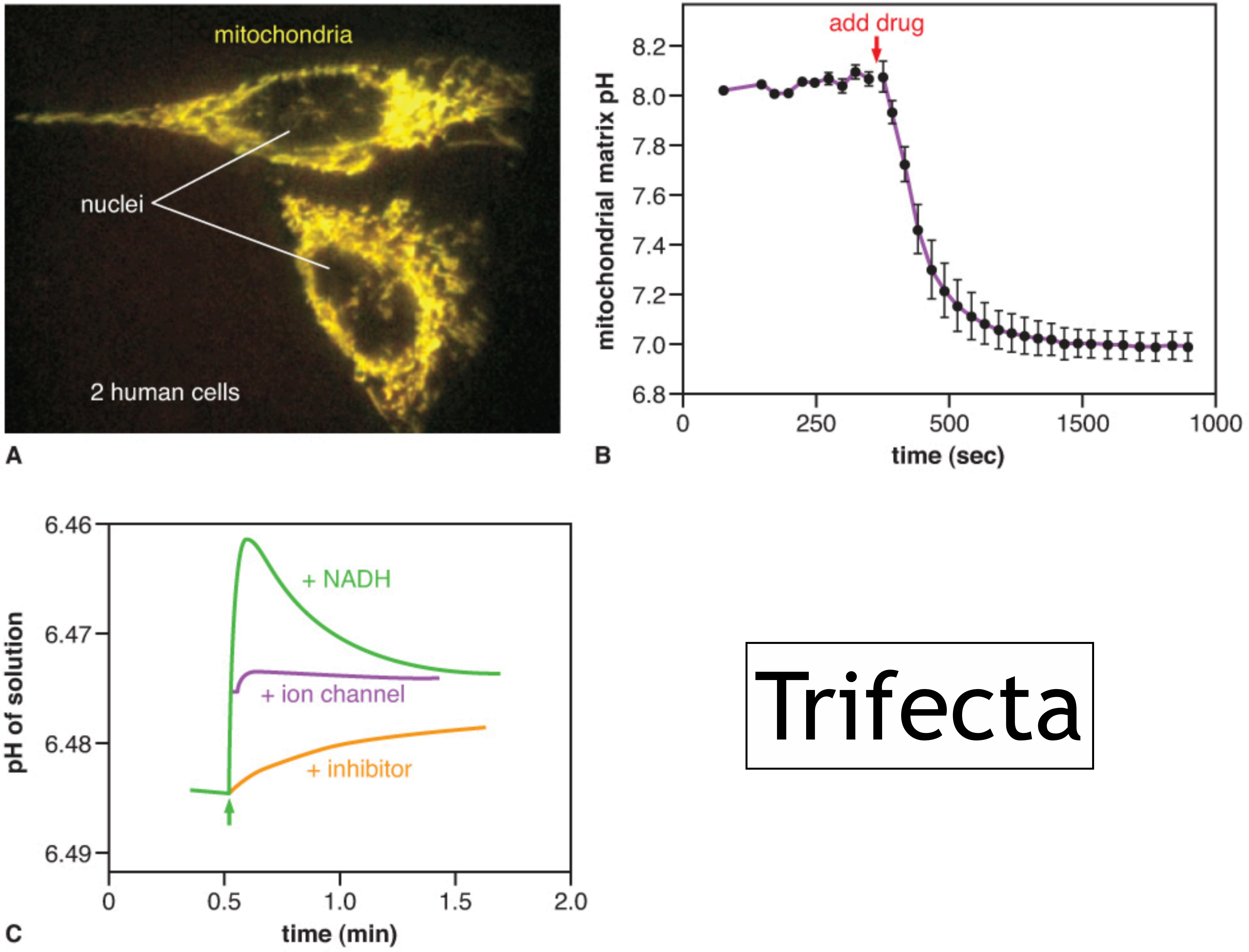
Free ribosomes in the mitochondrial matrix

Inner membrane

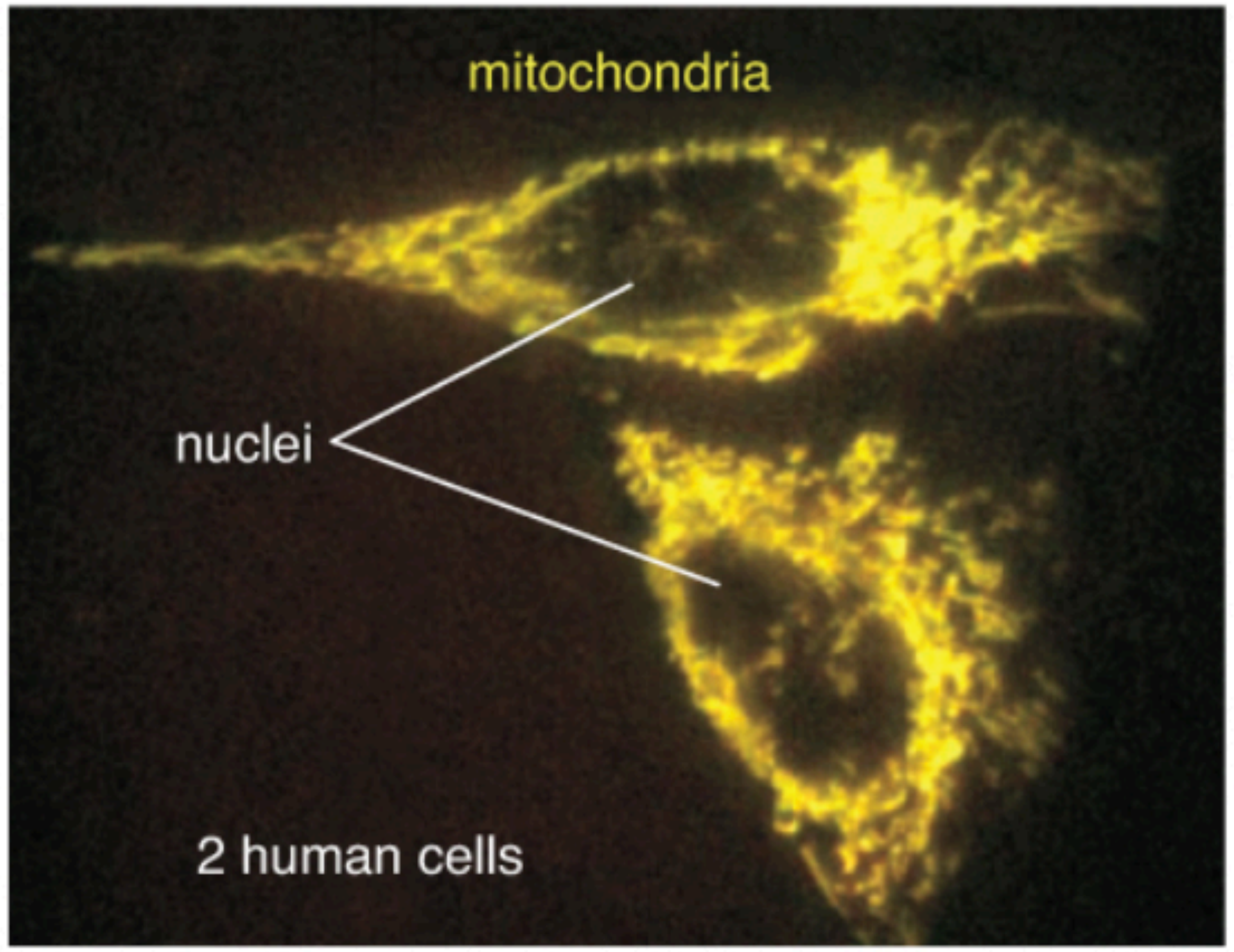
Cristae

Matrix

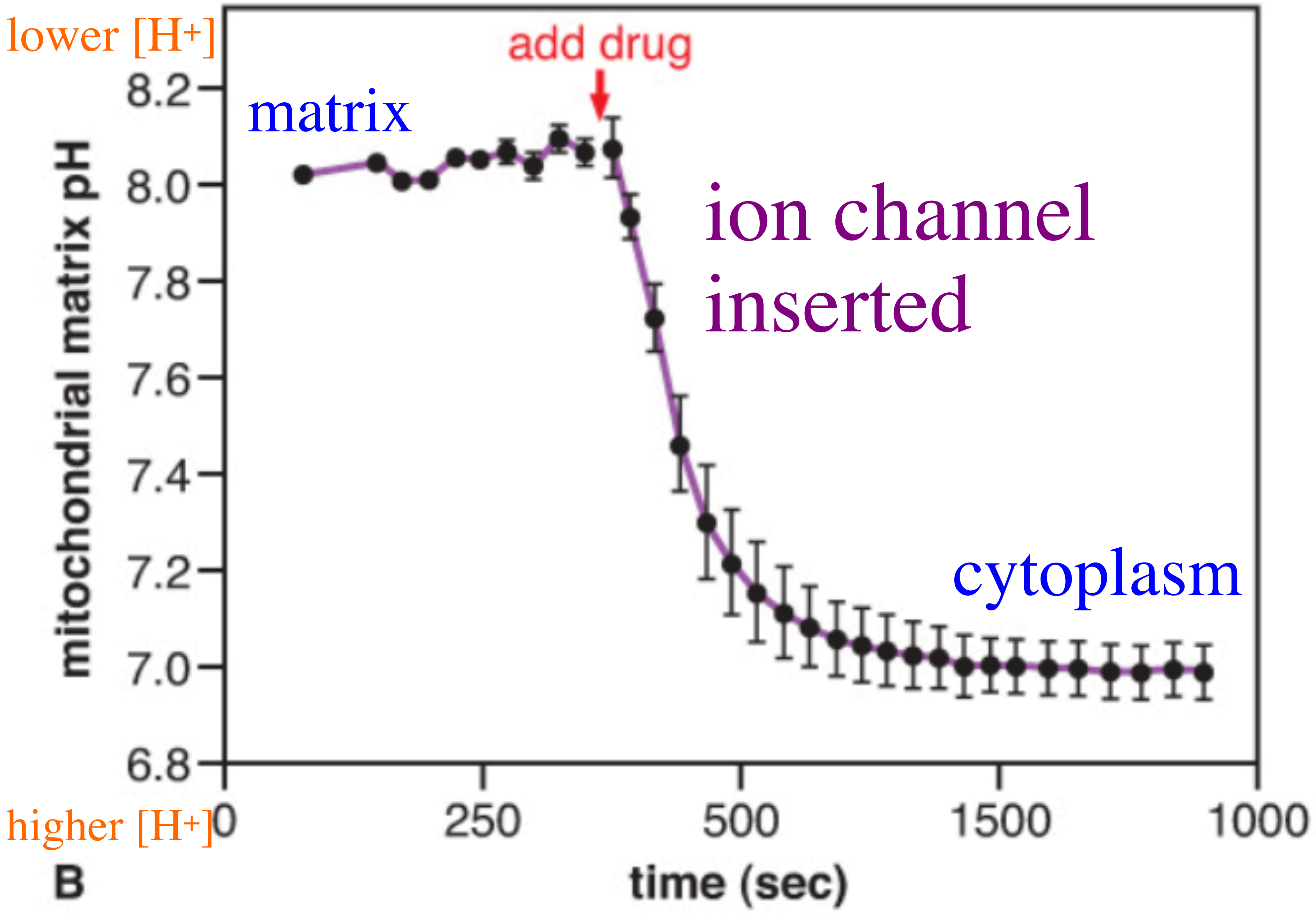
0.1 µm


mitochondria
nuclei
2 human cells
A
add drug
mitochondrial matrix pH
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
0
250
500
1500
1000
time (sec)
B
pH of solution
6.46
6.47
6.48
6.49
+ NADH
+ ion channel
+ inhibitor
0
0.5
1.0
1.5
2.0
time (min)
C

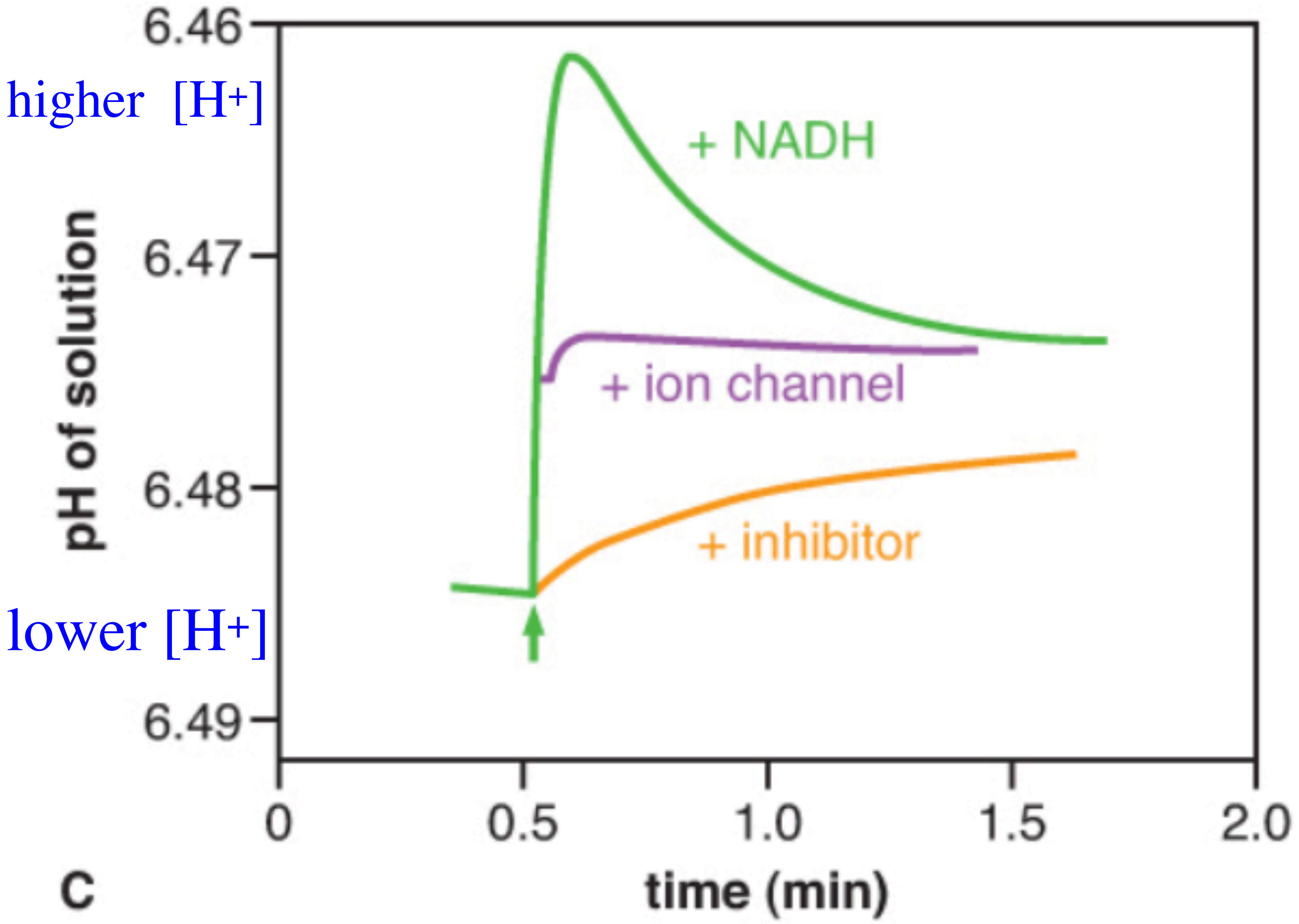
Trifecta


mitochondria
nuclei
2 human cells

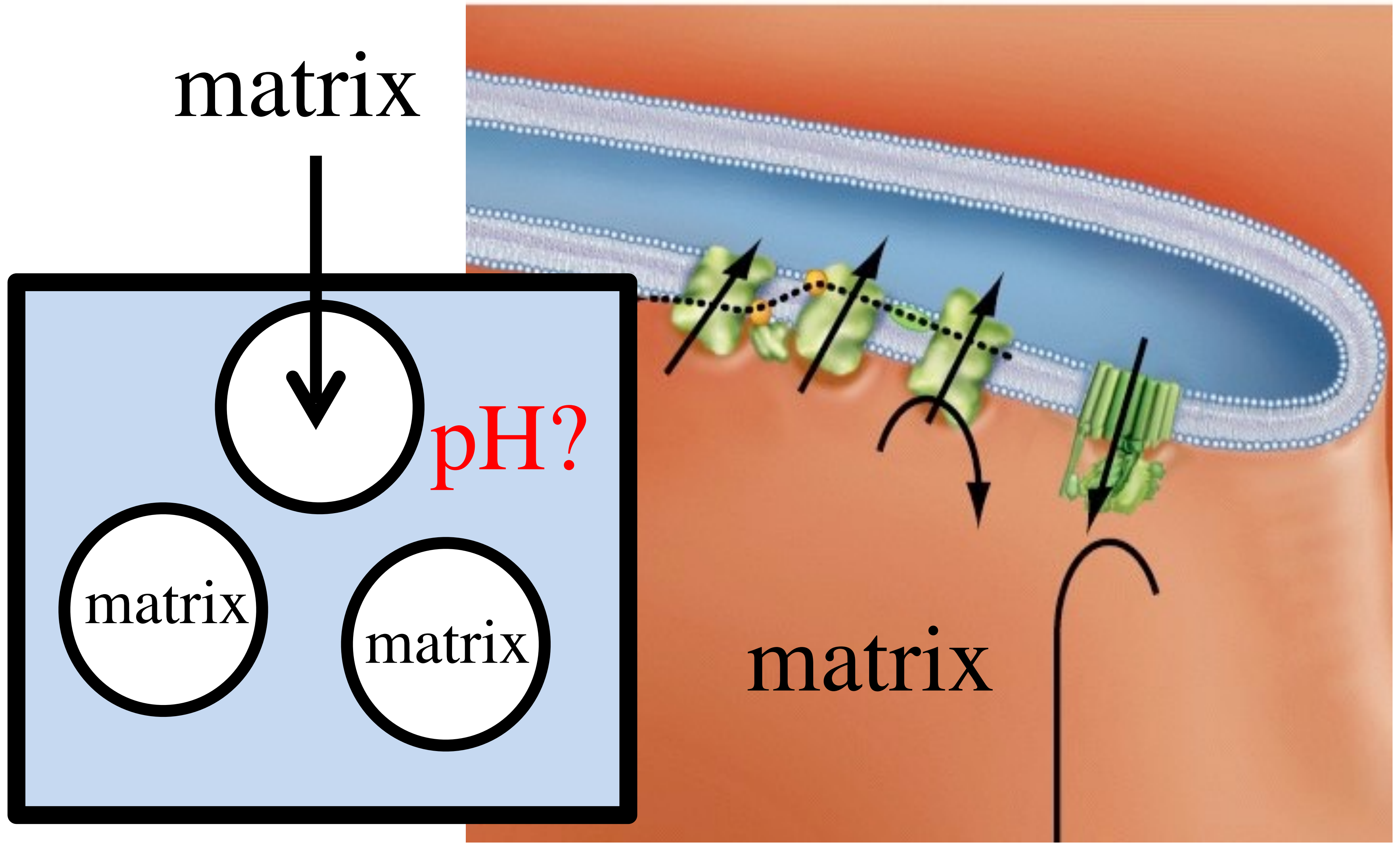
A


lower [H+]
higher [H+]
add drug
matrix
ion channel inserted
cytoplasm
mitochondrial matrix pH
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
0
250
500
1500
1000
time (sec)

B


higher [H+]
lower [H+]
pH of solution
6.46
6.47
6.48
6.49
+ NADH
+ ion channel
+ inhibitor
0
0.5
1.0
1.5
2.0
time (min)

C


matrix
pH?
matrix
matrix
matrix

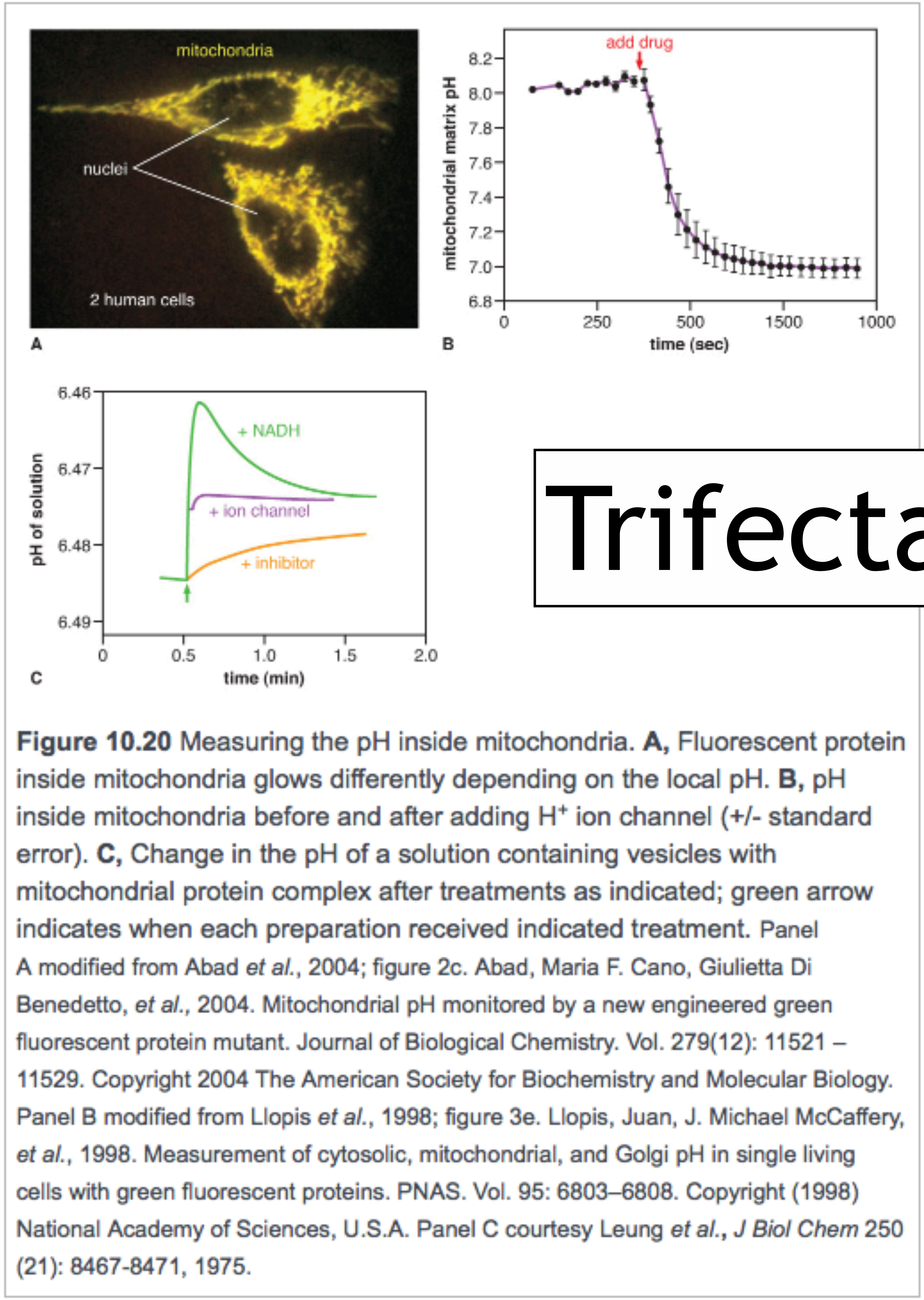
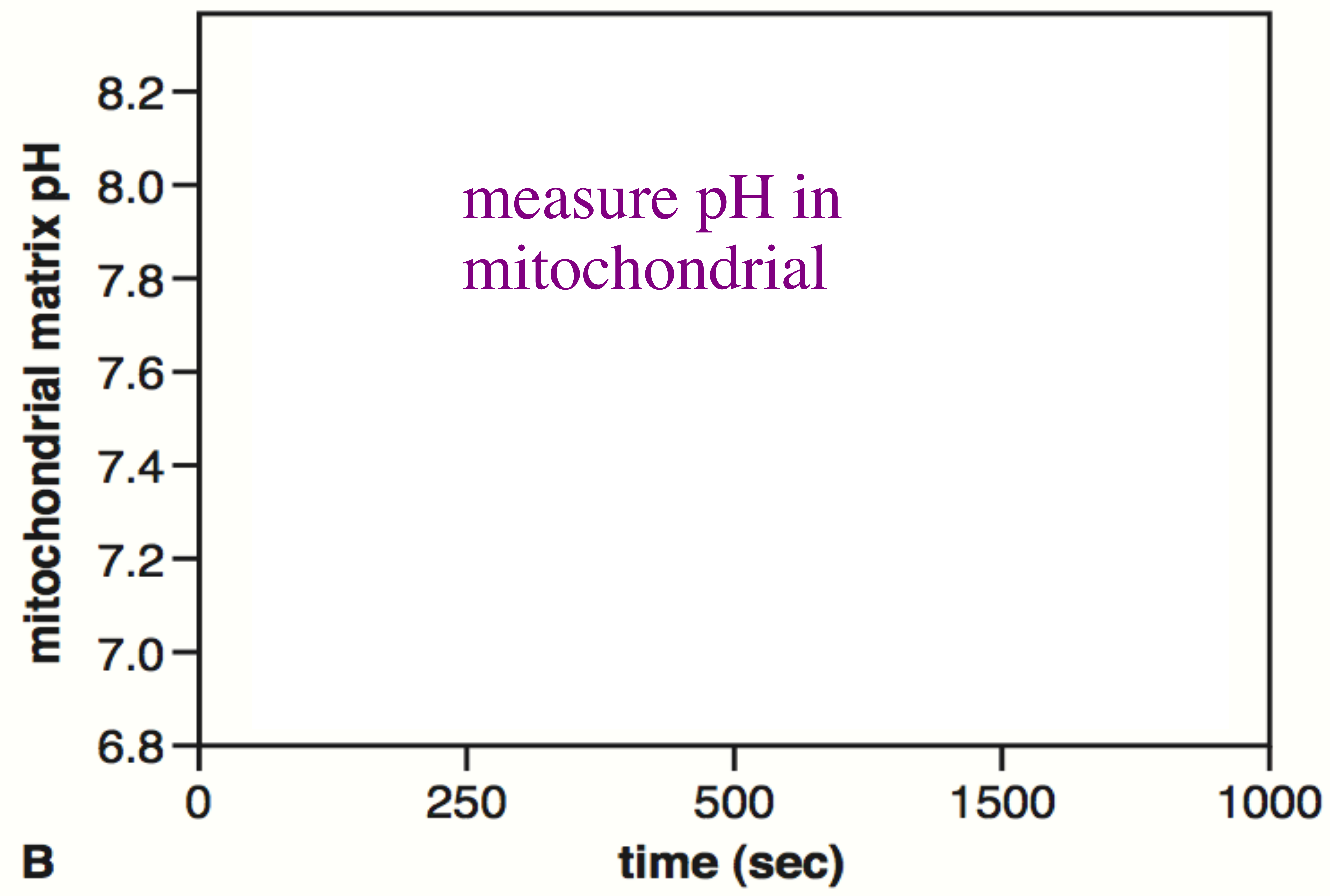


Figure 10.20 Measuring the pH inside mitochondria. **A,** Fluorescent protein inside mitochondria glows differently depending on the local pH. **B,** pH inside mitochondria before and after adding H^+ ion channel (+/- standard error). **C,** Change in the pH of a solution containing vesicles with mitochondrial protein complex after treatments as indicated; green arrow indicates when each preparation received indicated treatment. Panel A modified from Abad *et al.*, 2004; figure 2c. Abad, Maria F. Cano, Giulietta Di Benedetto, *et al.*, 2004. Mitochondrial pH monitored by a new engineered green fluorescent protein mutant. Journal of Biological Chemistry. Vol. 279(12): 11521 – 11529. Copyright 2004 The American Society for Biochemistry and Molecular Biology. Panel B modified from Llopis *et al.*, 1998; figure 3e. Llopis, Juan, J. Michael McCaffery, *et al.*, 1998. Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. PNAS. Vol. 95: 6803–6808. Copyright (1998) National Academy of Sciences, U.S.A. Panel C courtesy Leung *et al.*, *J Biol Chem* 250 (21): 8467-8471, 1975.

Trifecta

pH Inside Mitochondria



measure pH in
mitochondrial

Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

pH Inside Mitochondria

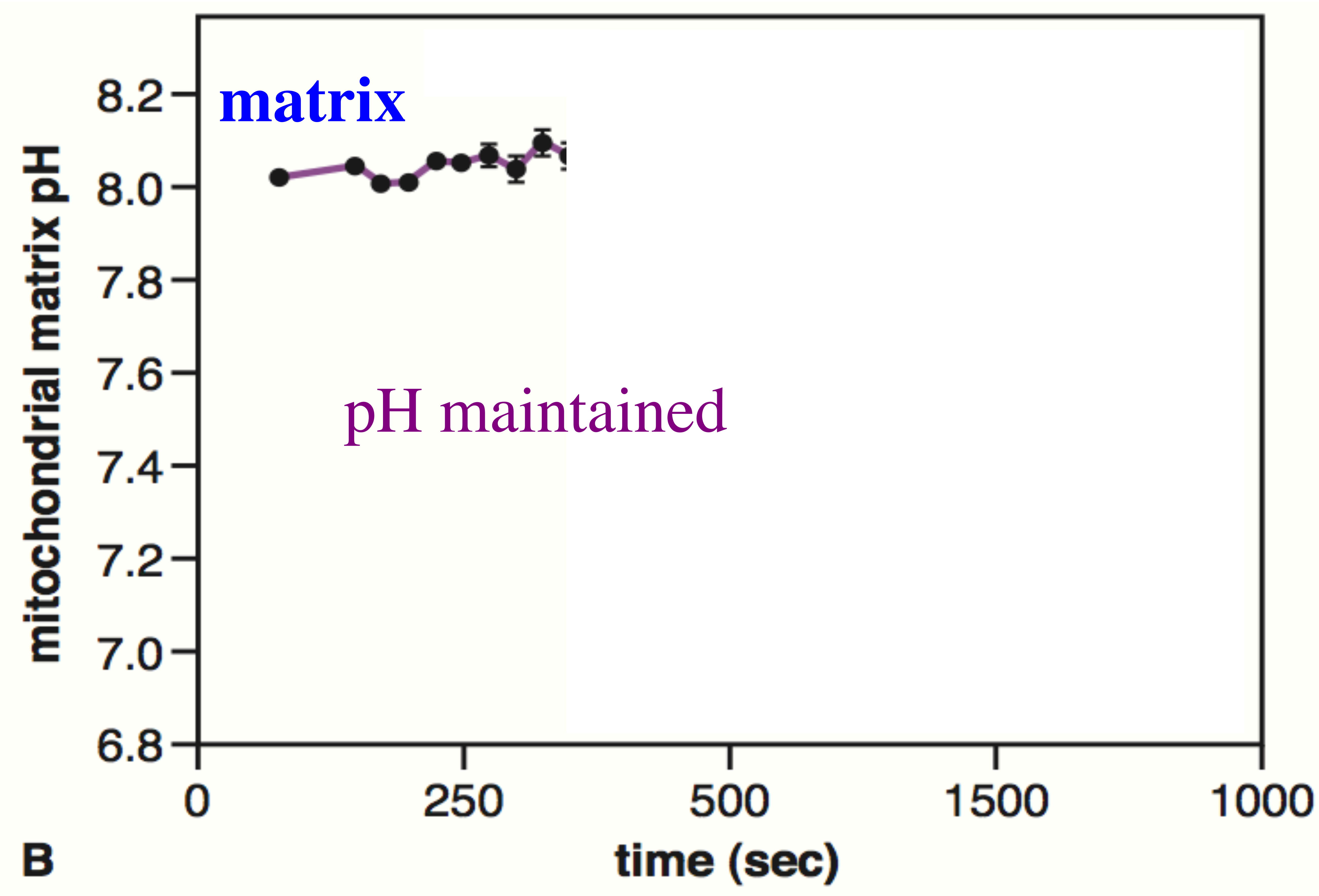


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

pH Inside Mitochondria

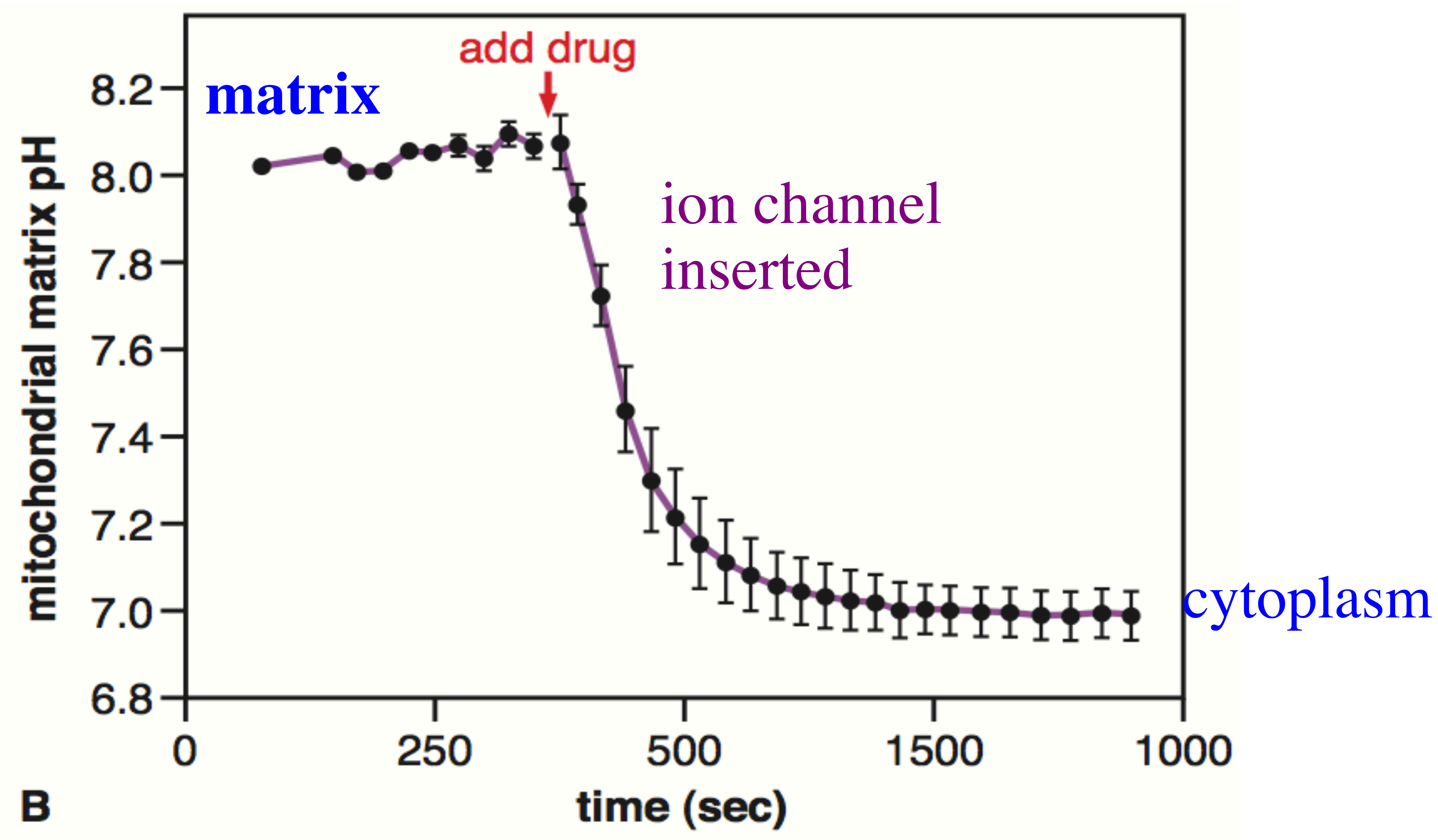


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Which way does the H^+ flow?

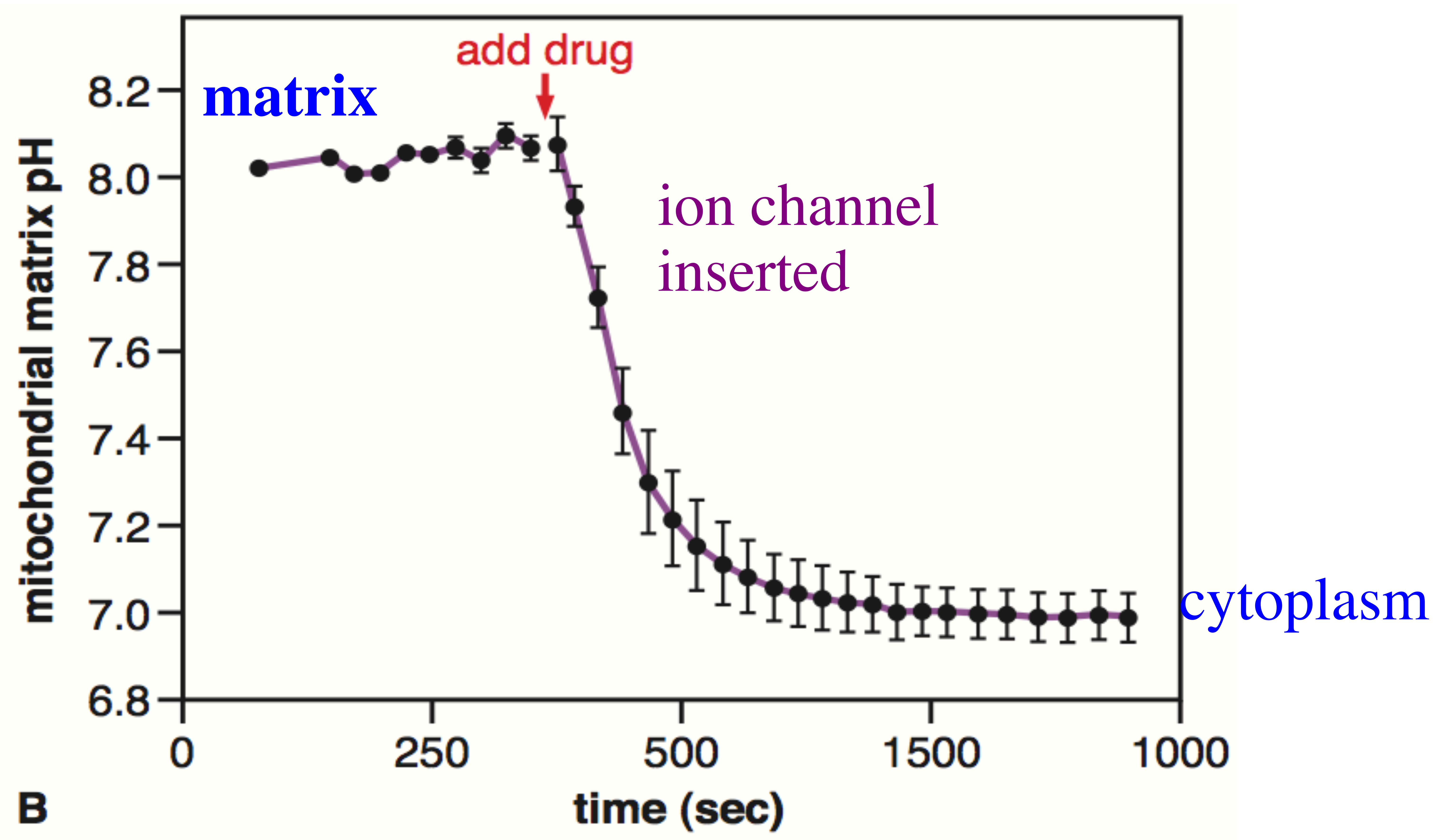
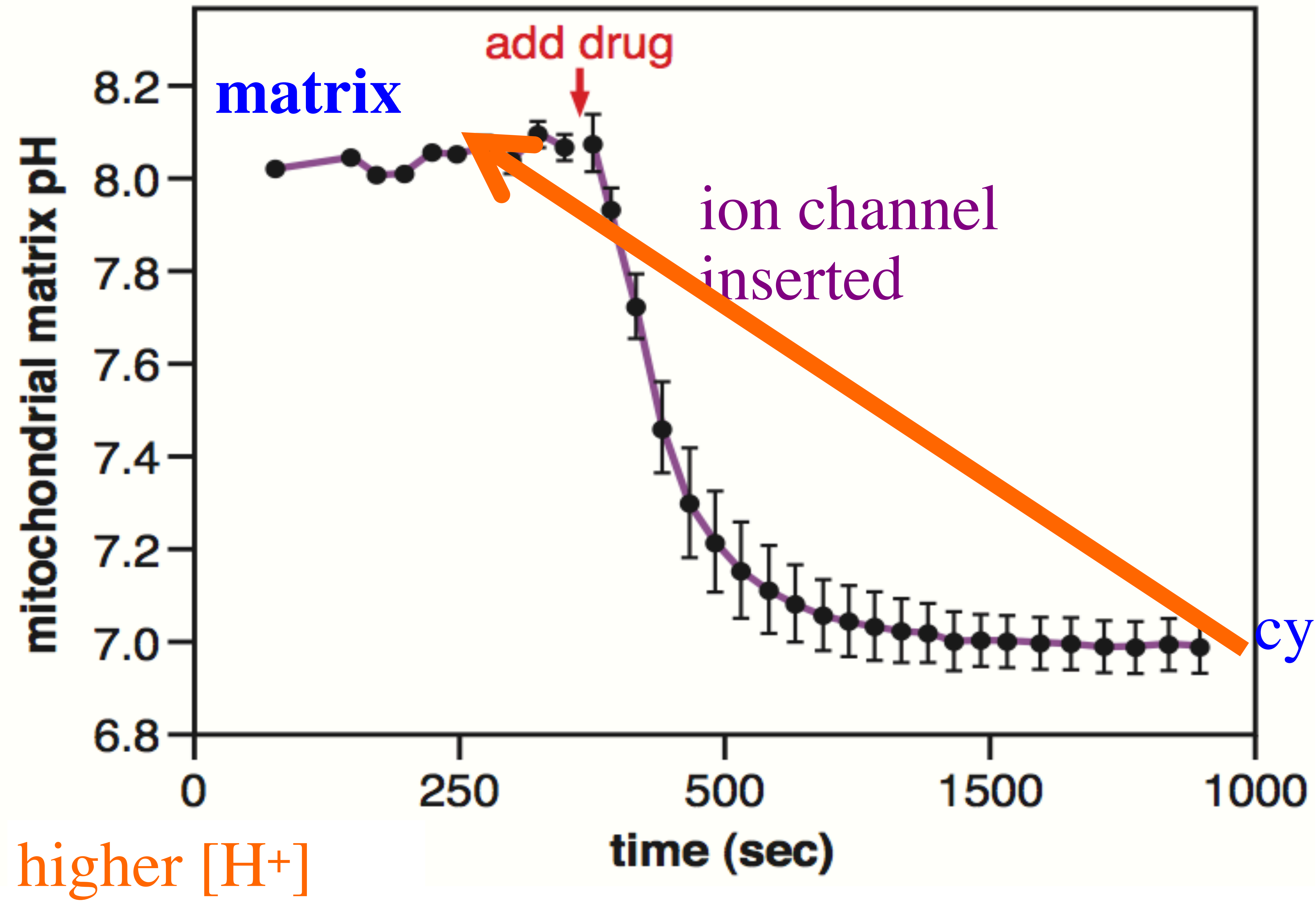


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Which way does the H^+ flow?

lower $[H^+]$



matrix

ion channel inserted

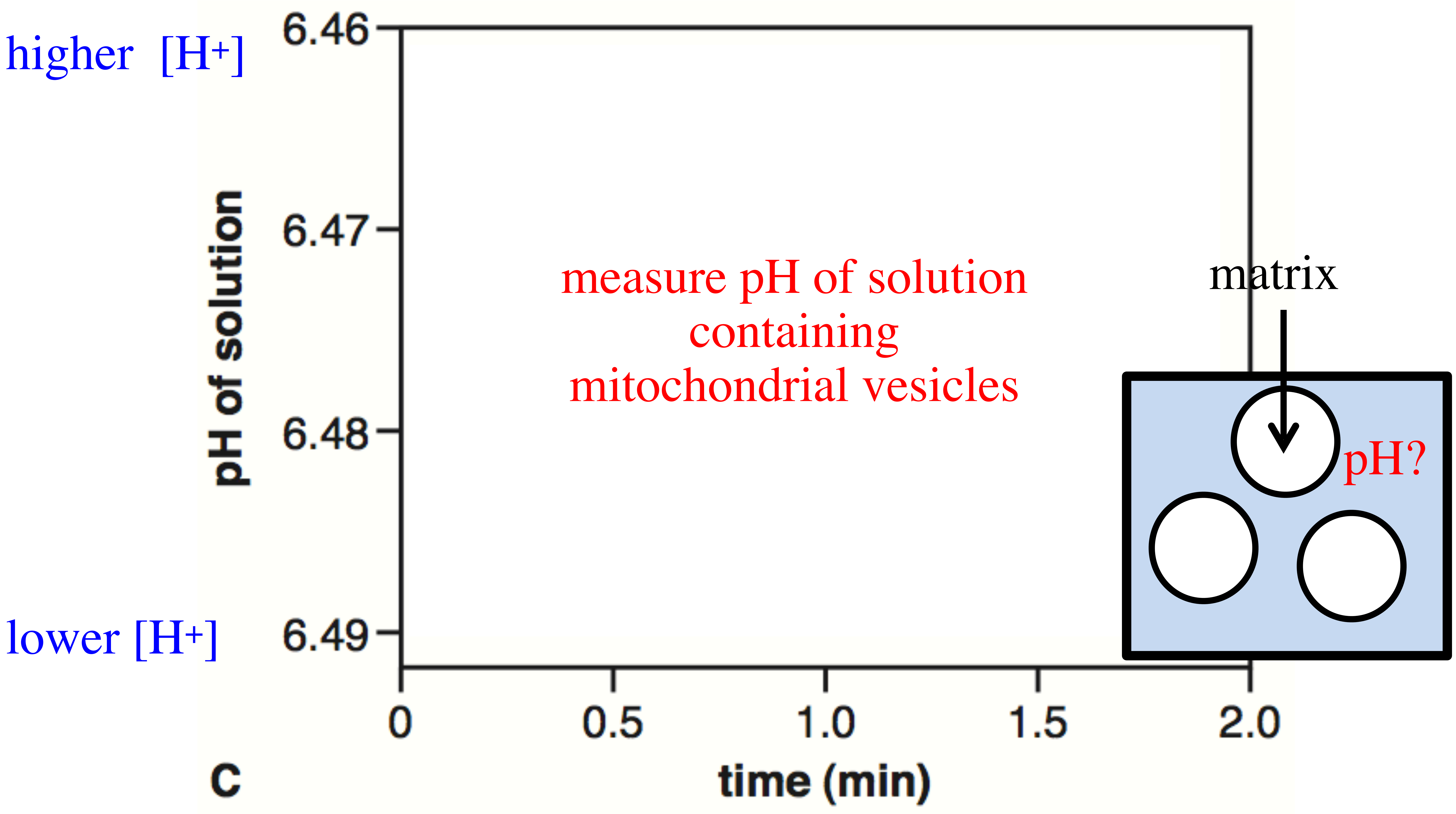
cytoplasm

higher $[H^+]$

Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH



measure pH of solution containing mitochondrial vesicles

Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH

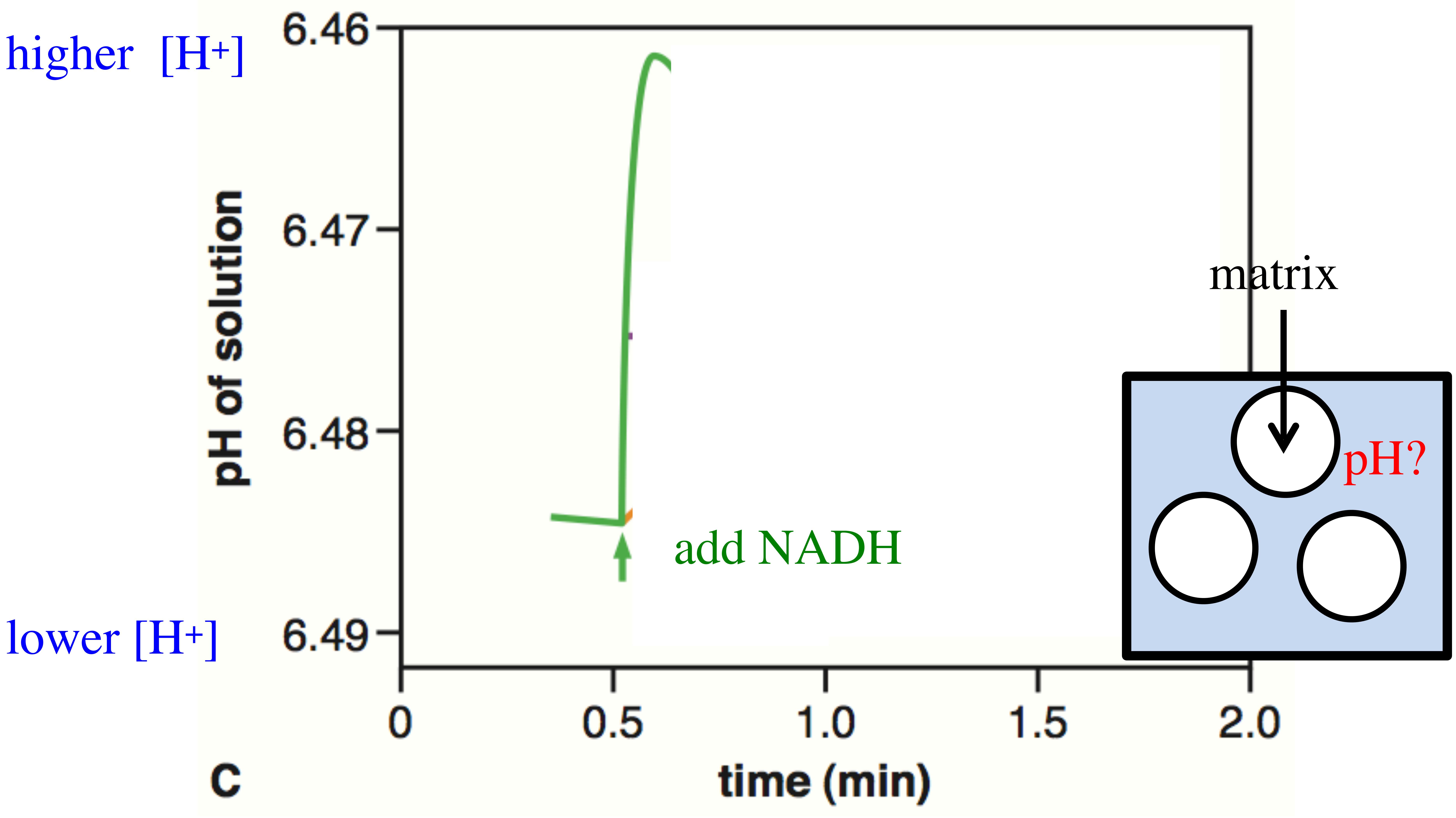


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH

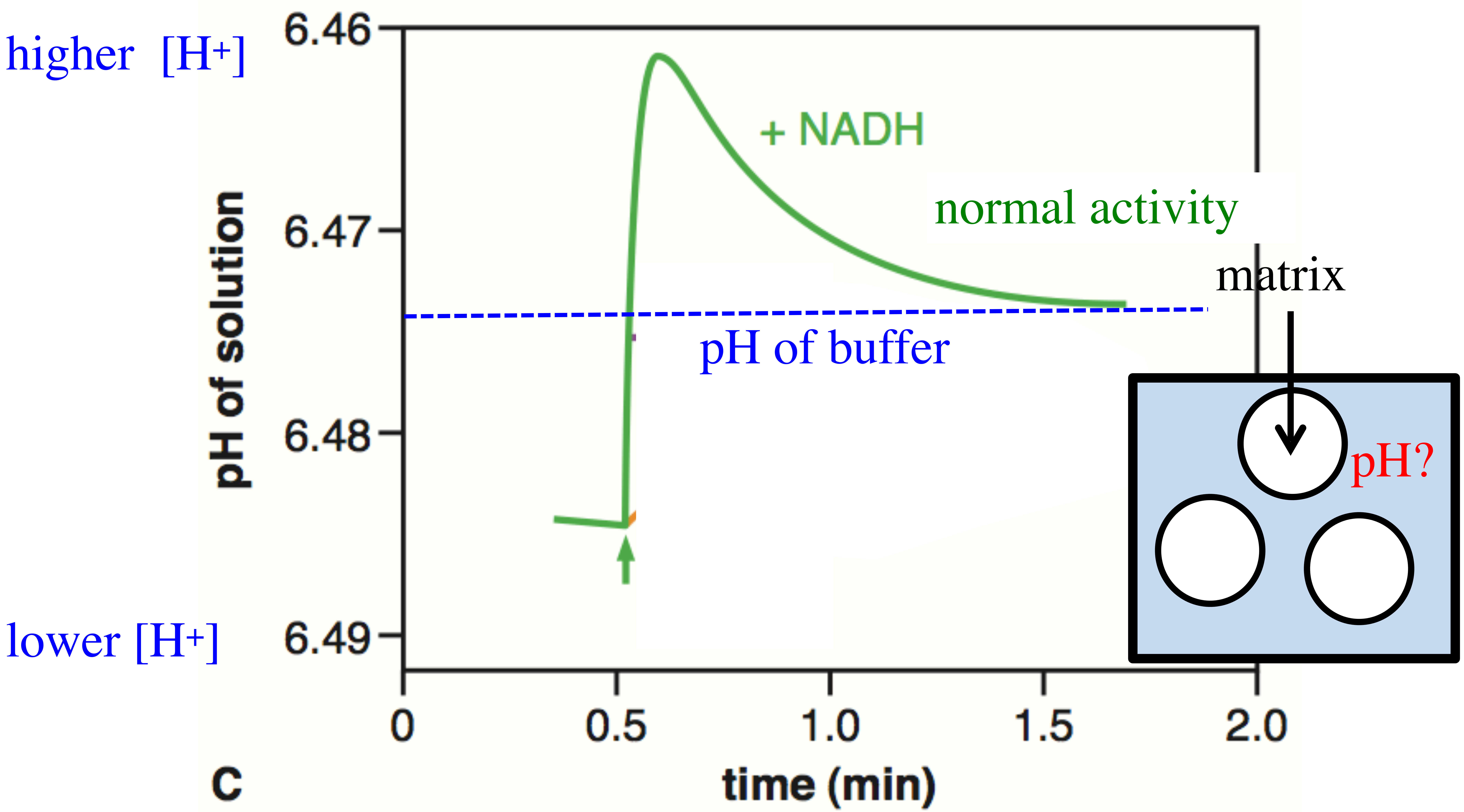


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH

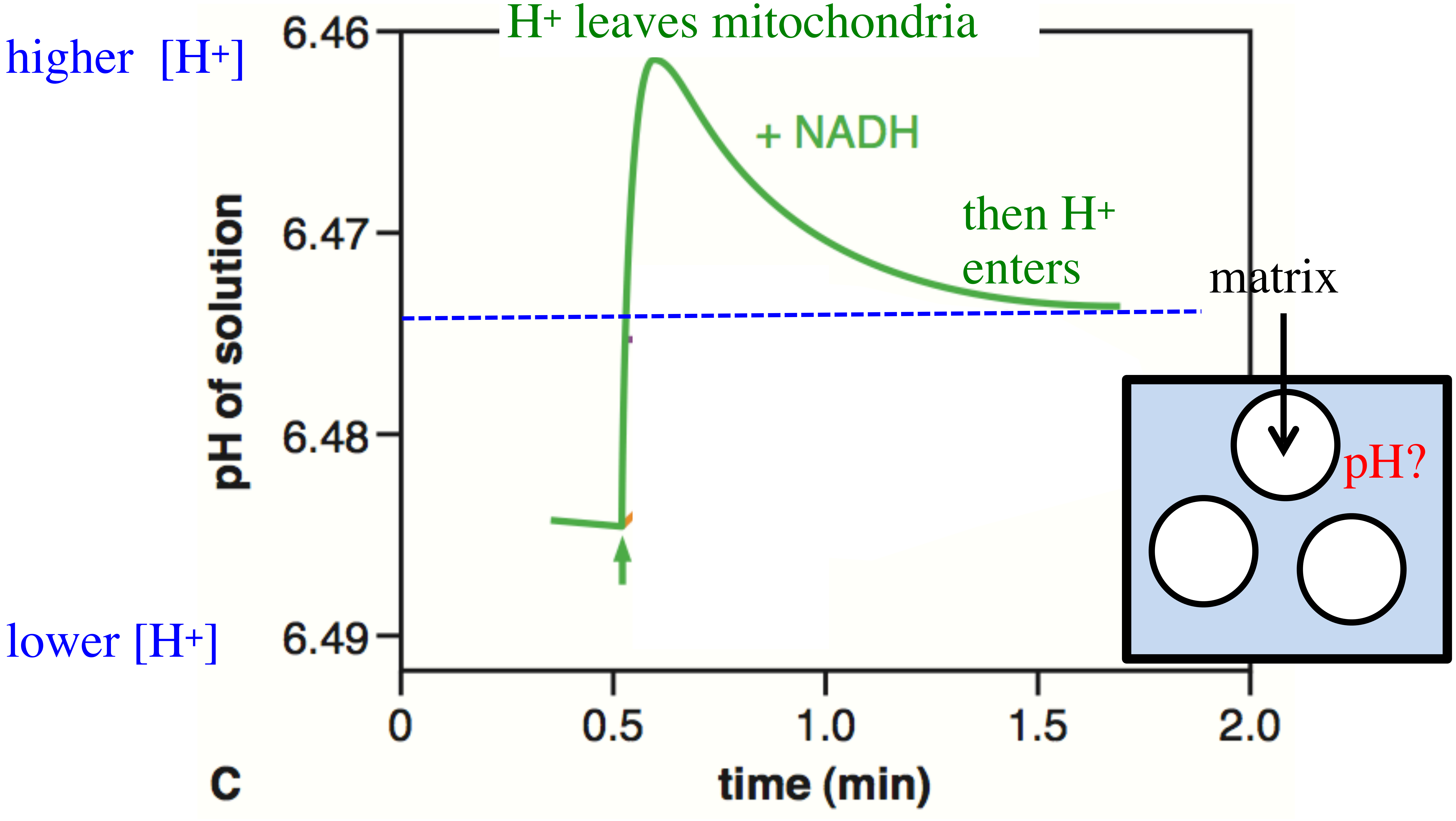


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH

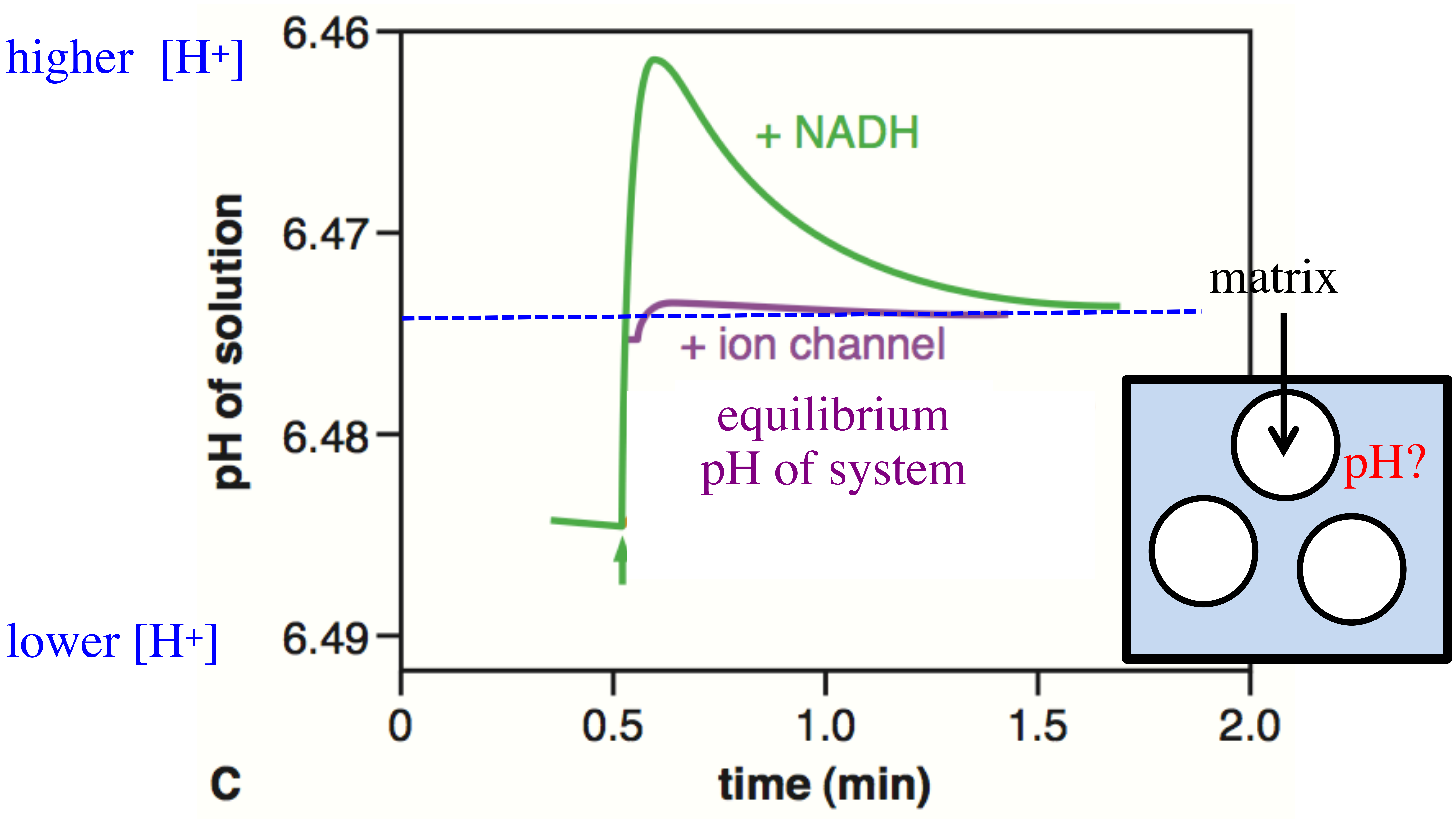
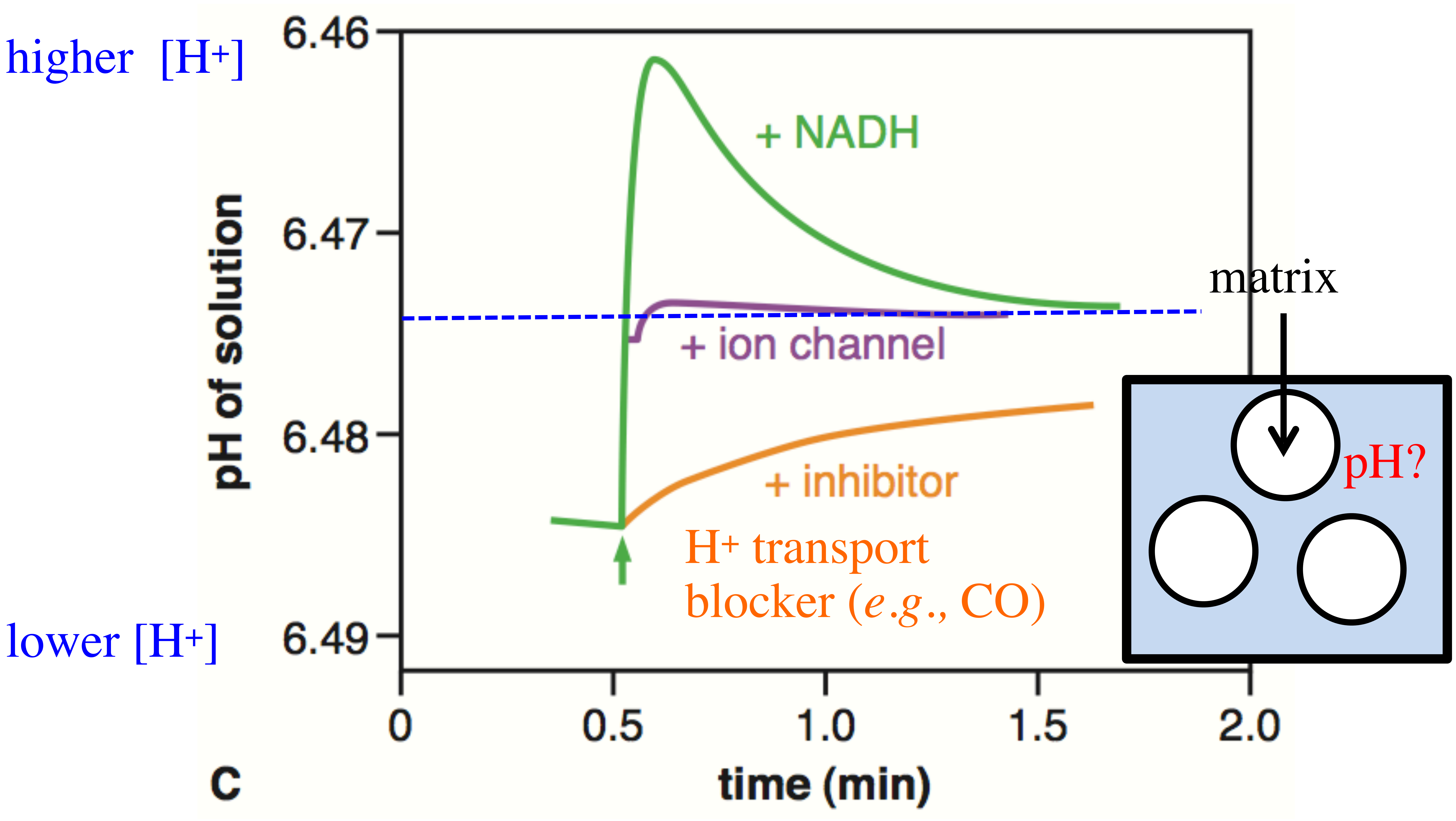


Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Δ of Mitochondrial pH



higher [H^+]

lower [H^+]

matrix

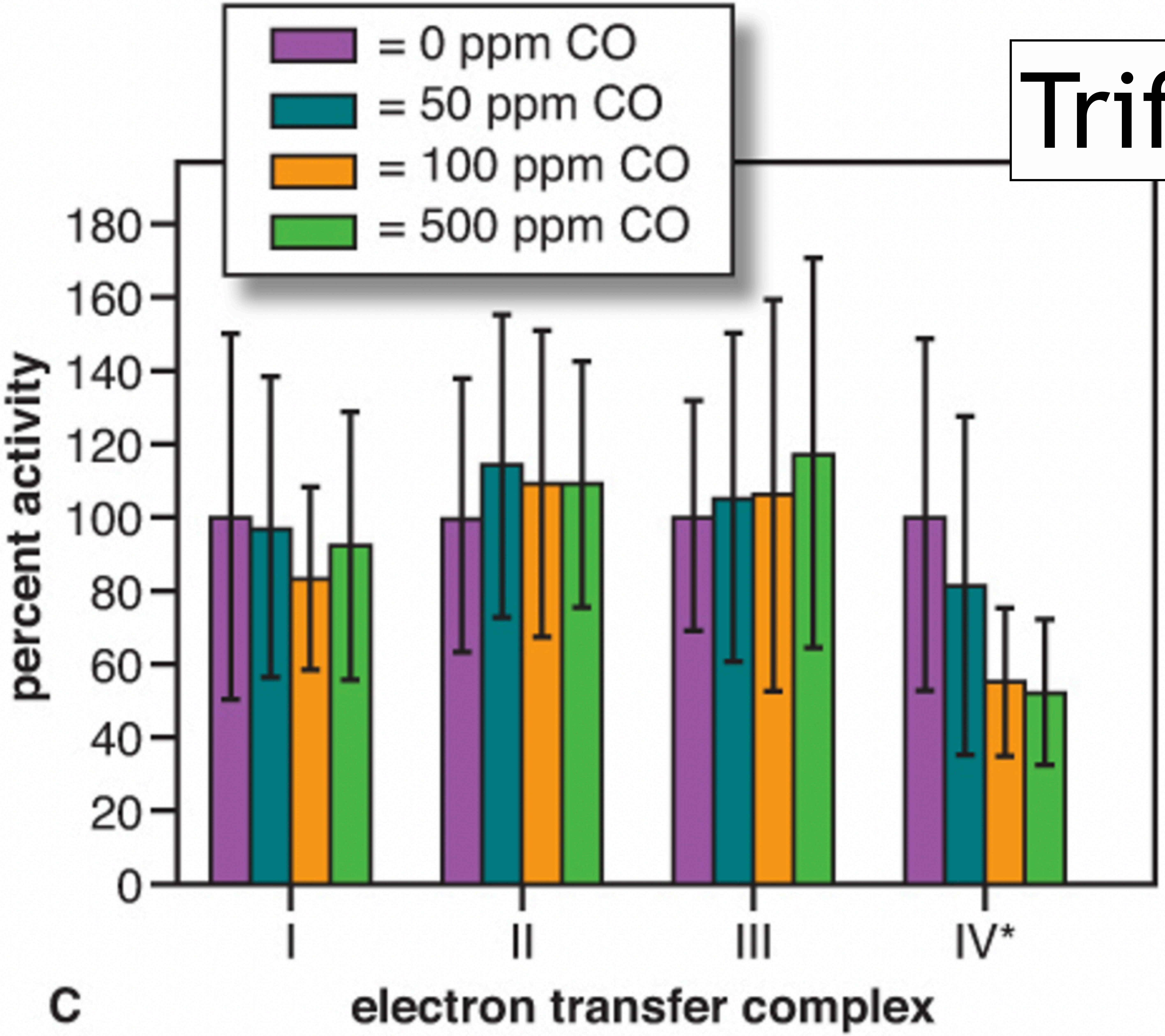
pH?

H^+ transport blocker (*e.g.*, CO)

Fig. 10.20

A: from Abad *et al.*, 2004; B: modified from Llopis *et al.*, 1998; C: Leung *et al.*, 1975

Trifecta


= 0 ppm CO
= 50 ppm CO
= 100 ppm CO
= 500 ppm CO
percent activity
180
160
140
120
100
80
60
40
20
0
I
II
III
IV*
C
electron transfer complex

Trifecta

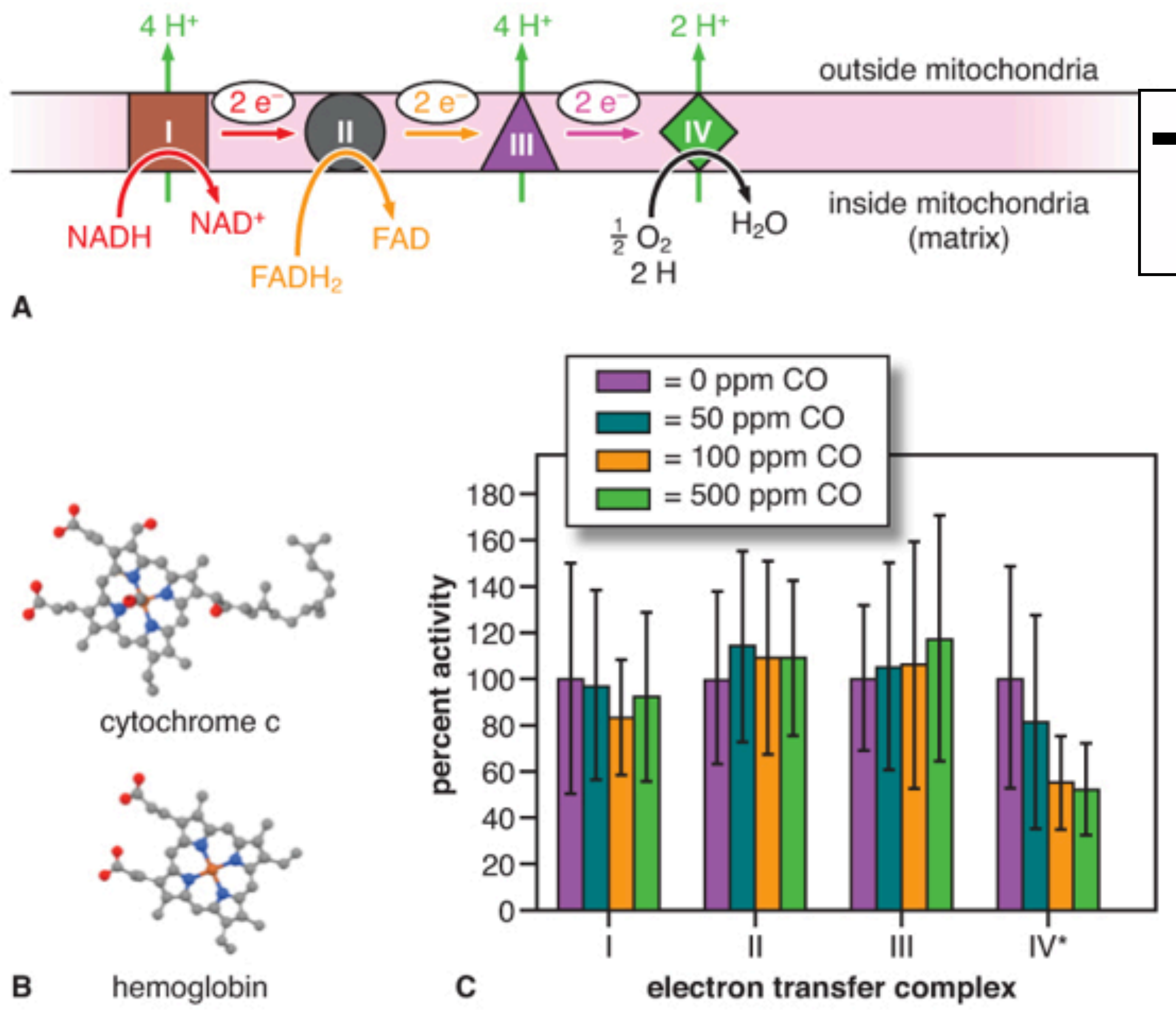
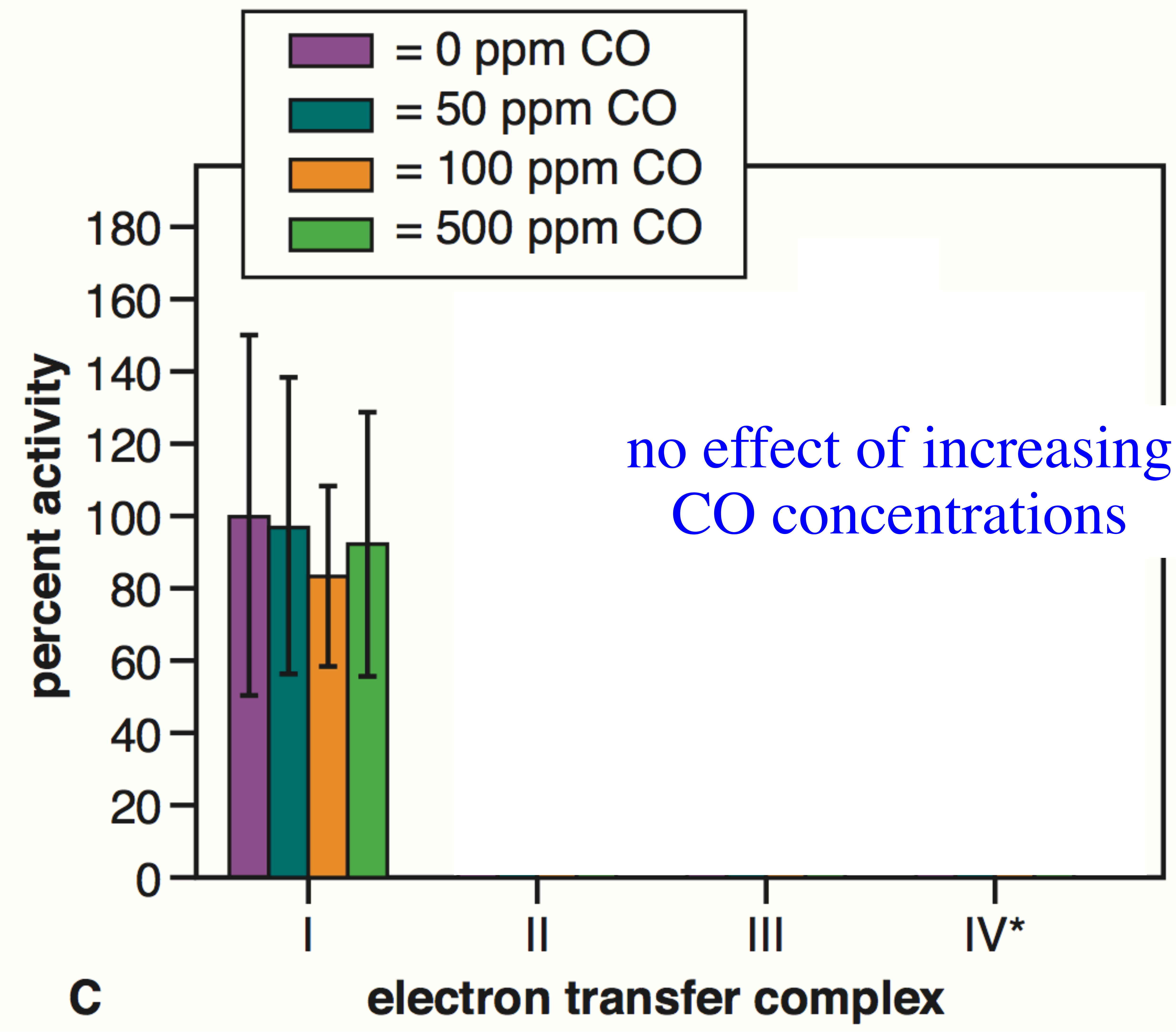


Figure 10.21 Electron transfer pathway in mitochondrial membrane. **A,** Four protein complexes contain hemes and transfer electrons from NADH and $FADH_2$ donors to the final acceptor molecule. **B,** Atomic structures of two similar hemes from cytochrome c of complex IV and human hemoglobin. Nitrogen atoms are blue, centrally located iron ions are orange, carbons are gray, oxygens are red. **C,** Effect of carbon monoxide (CO) on electron transfer by the four complexes shown in **A,** +/- standard deviation. * indicates $p < 0.001$. Panels A & B from common knowledge – original art. Panel C modified from Alonso *et al.*, 2003; part of their figure 2. Alonso, Jose-Ramon, Francesc Cardellach, *et al.* 2003. Pharmacology and Toxicology. Vol. 93: 142 – 146. Carbon Monoxide Specifically Inhibits Cytochrome C Oxidase of Human Mitochondrial Respiratory Chain.

Electron Transport and Carriers



no effect of increasing CO concentrations

Fig. 10.21

C: modified from Alonso *et al.*, 2003

Electron Transport and Carriers

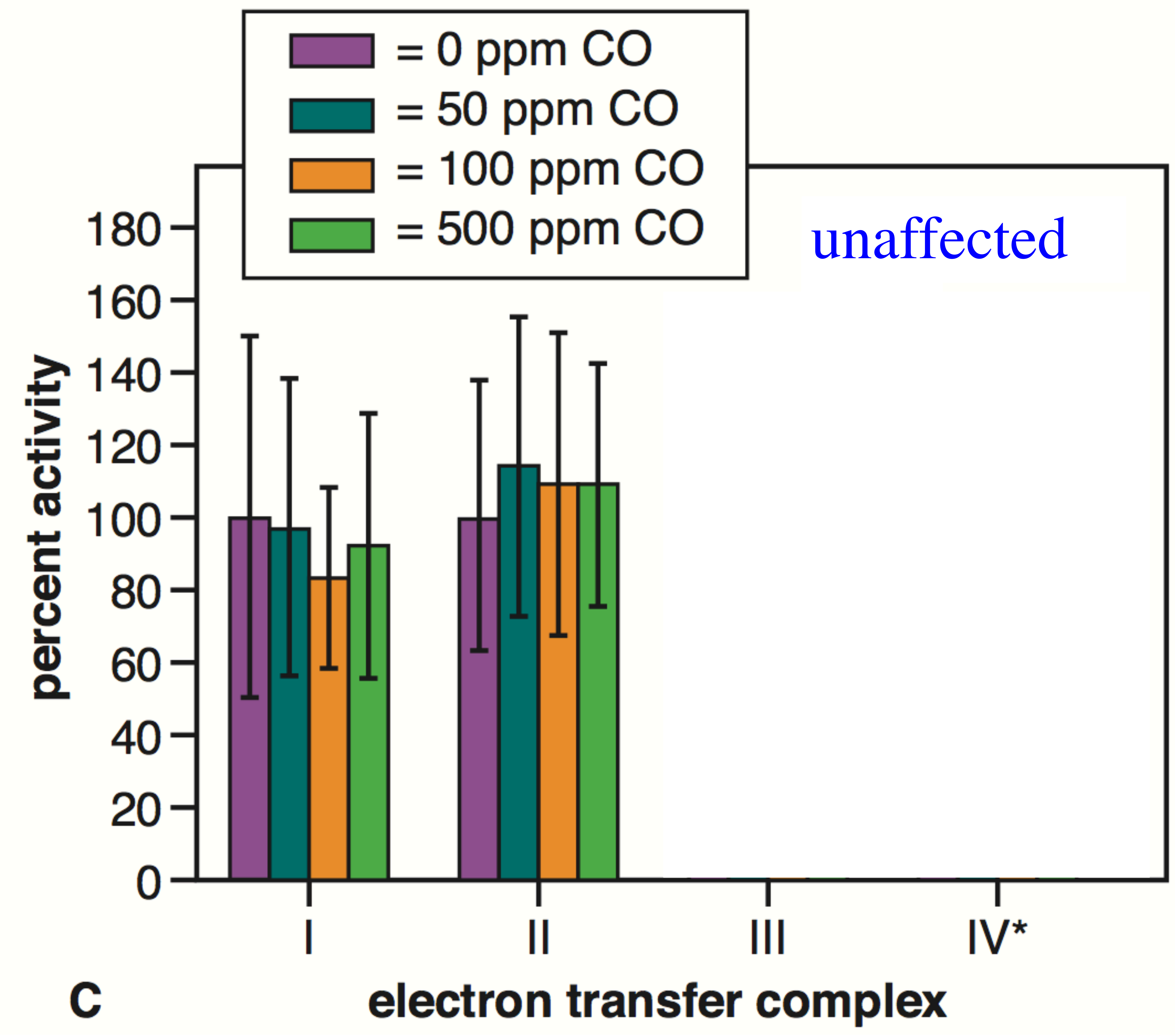


Fig. 10.21

C: modified from Alonso *et al.*, 2003

Electron Transport and Carriers

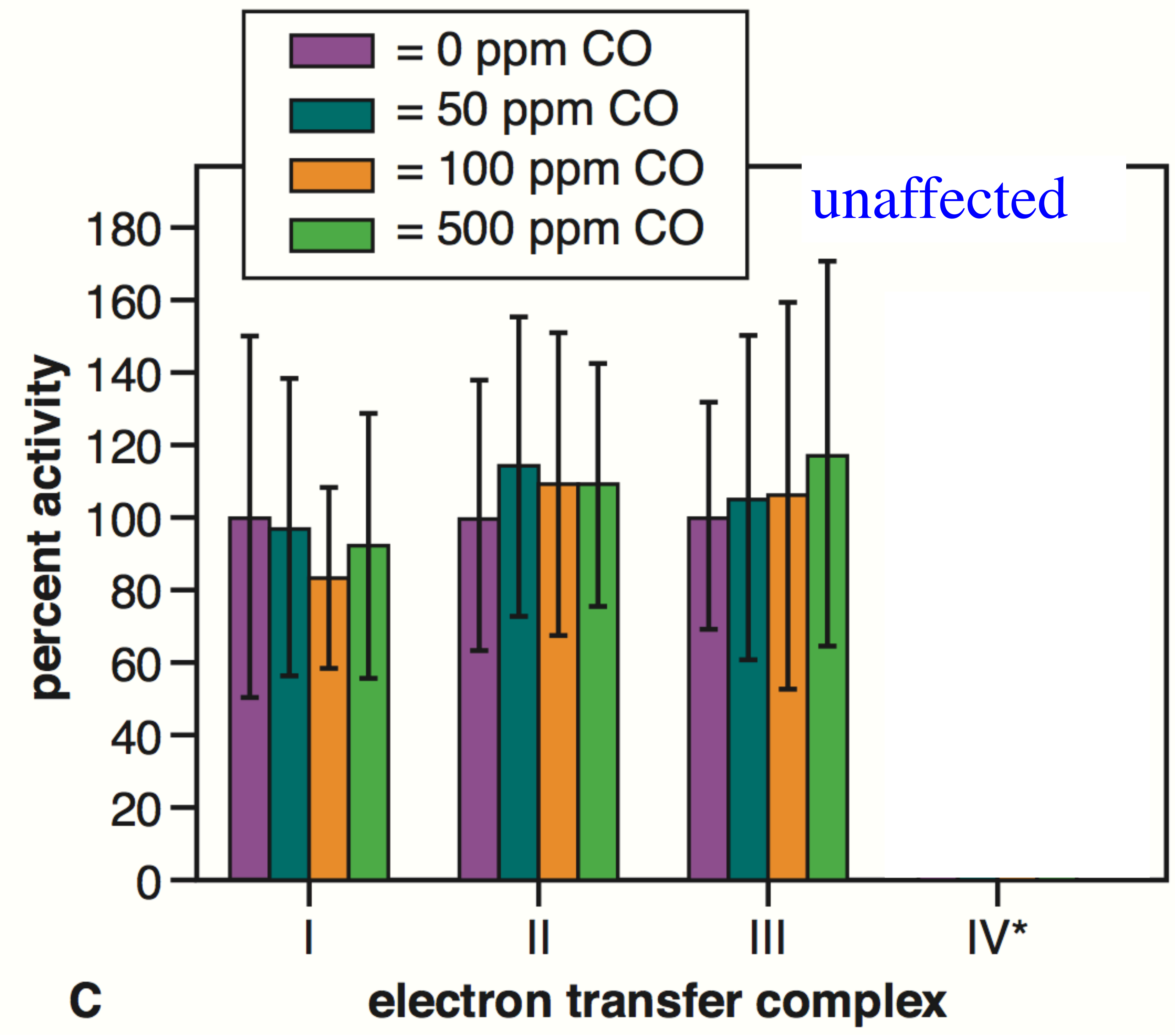
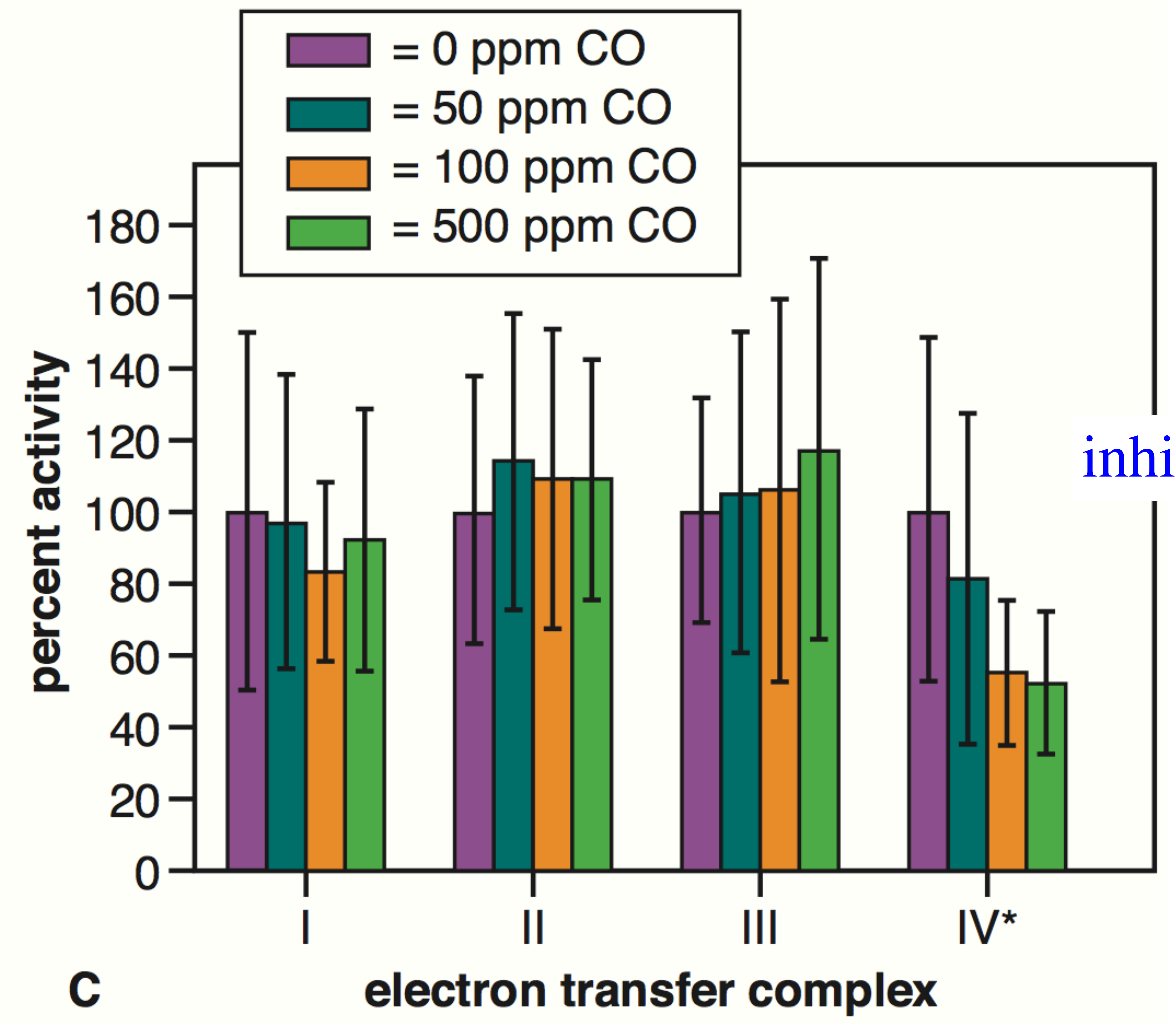


Fig. 10.21

C: modified from Alonso *et al.*, 2003

Electron Transport and Carriers



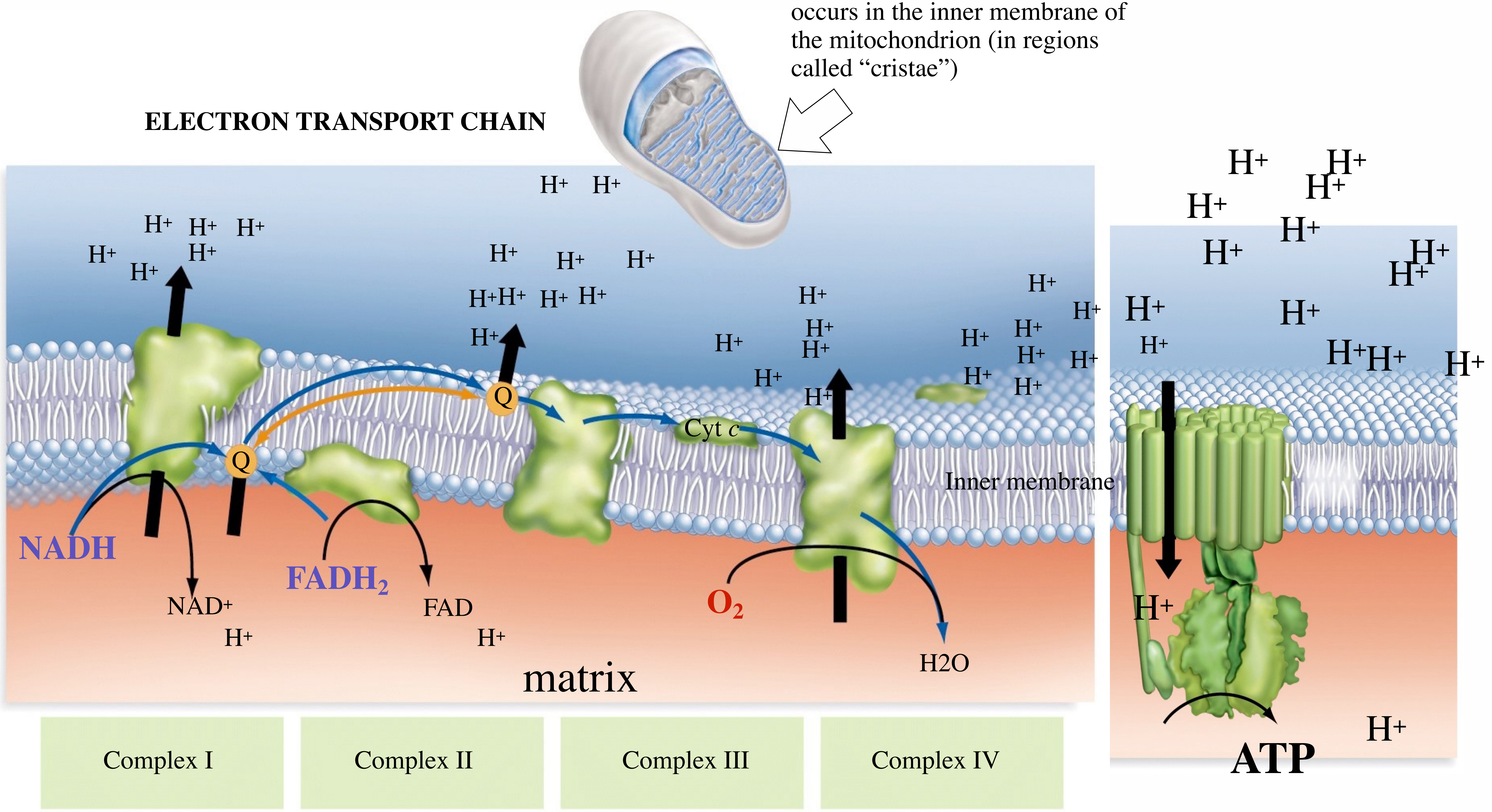
inhibited!

Fig. 10.21

C: modified from Alonso *et al.*, 2003

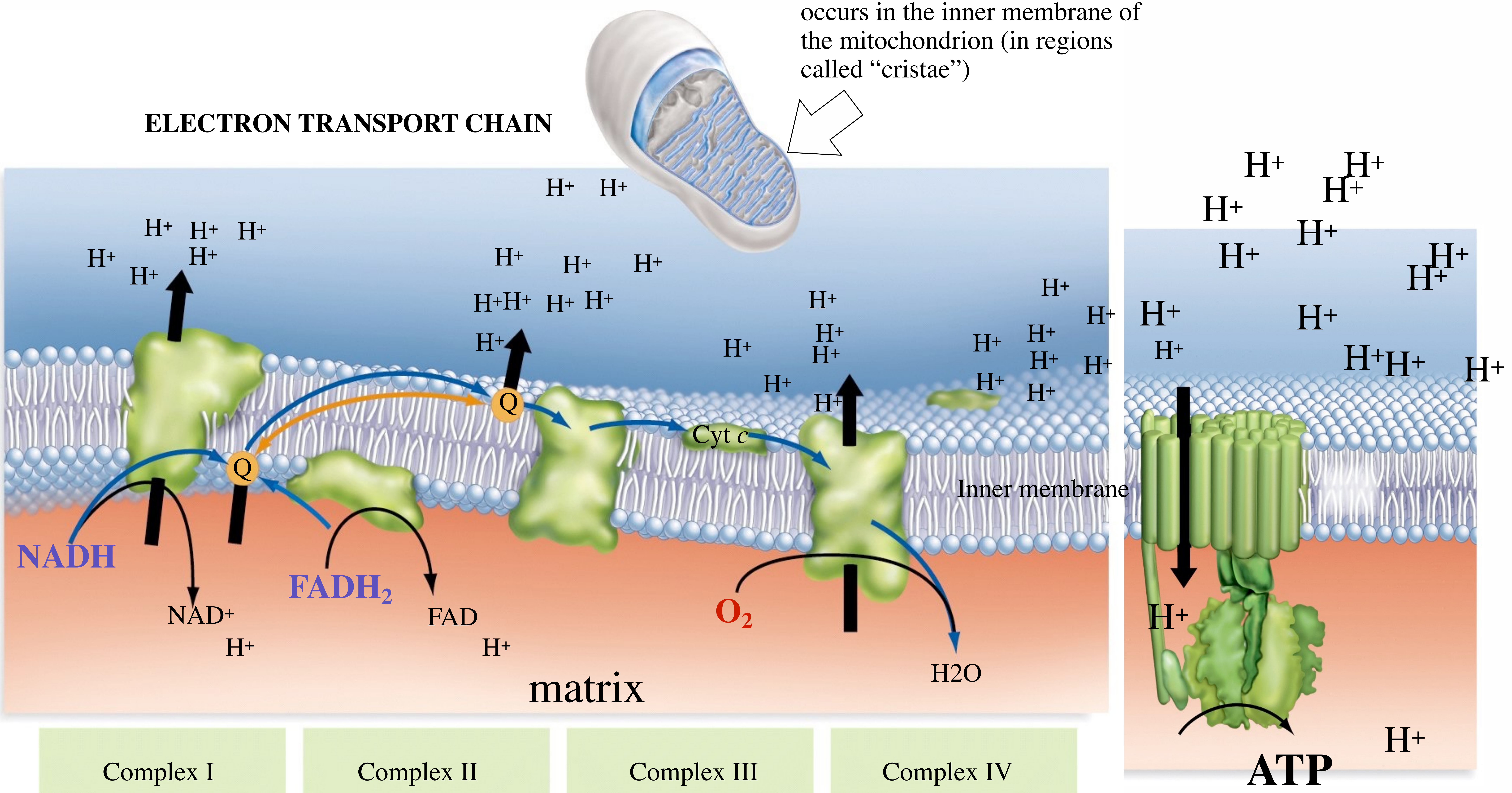
The electron transport chain occurs in the inner membrane of the mitochondrion (in regions called “cristae”)

ELECTRON TRANSPORT CHAIN



1. What is backwards? 2. How much ATP? 3. Poke hole?

ELECTRON TRANSPORT CHAIN



The electron transport chain occurs in the inner membrane of the mitochondrion (in regions called "cristae")

Complex I

Complex II

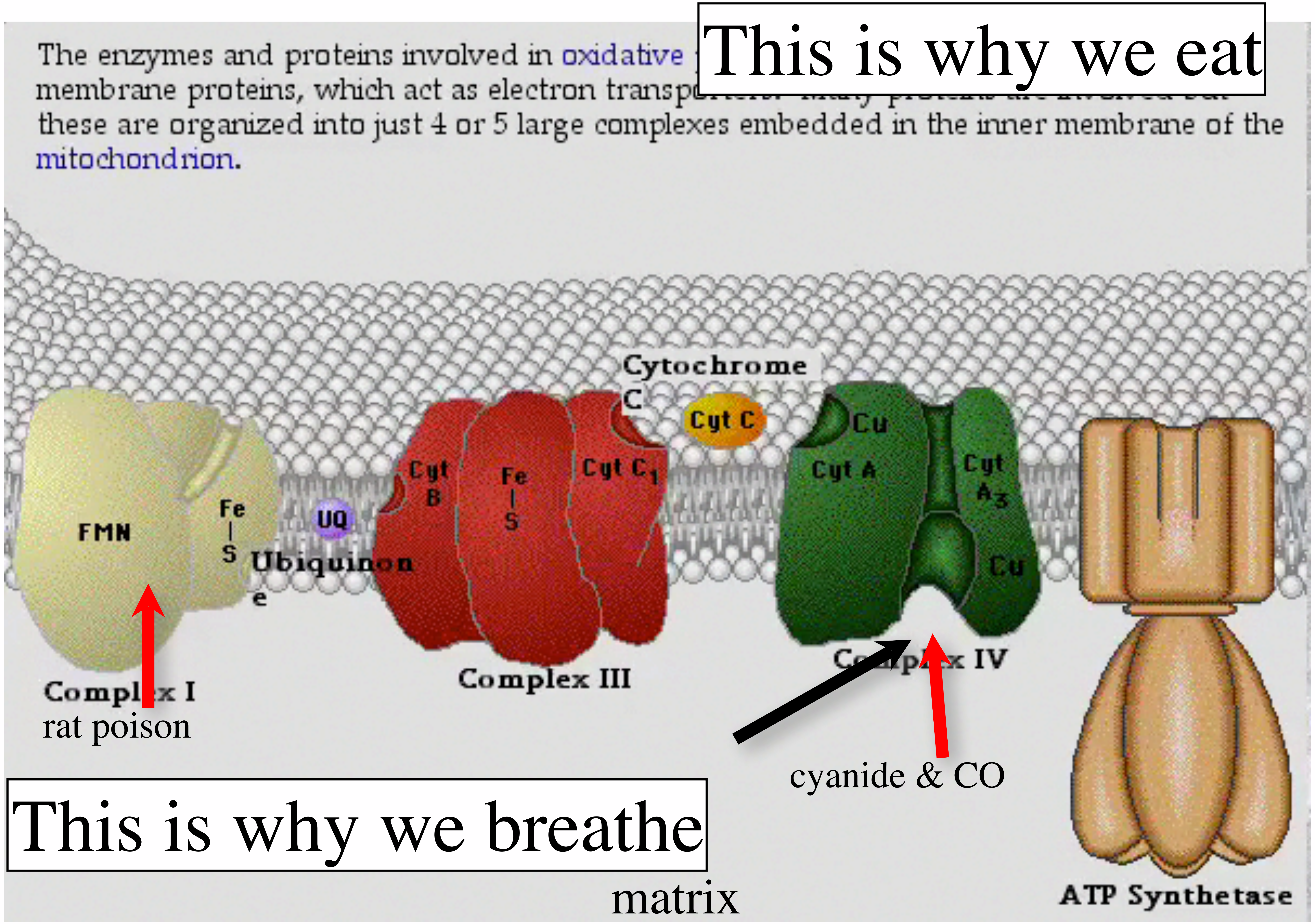
Complex III

Complex IV

NADH-Q Reductase, FADH-Q Reductase, Cytochrome Reductase, Cytochrome Oxidase

(NADH dehydrogenase, FADH dehydrogenase, Cytochrome complex, Cytochrome Oxidase)

This is why we eat



The enzymes and proteins involved in oxidative
membrane proteins, which act as electron transp
these are organized into just 4 or 5 large complexes embedded in the inner membrane of the mitochondrion.

rat poison

cyanide & CO

This is why we breathe

matrix