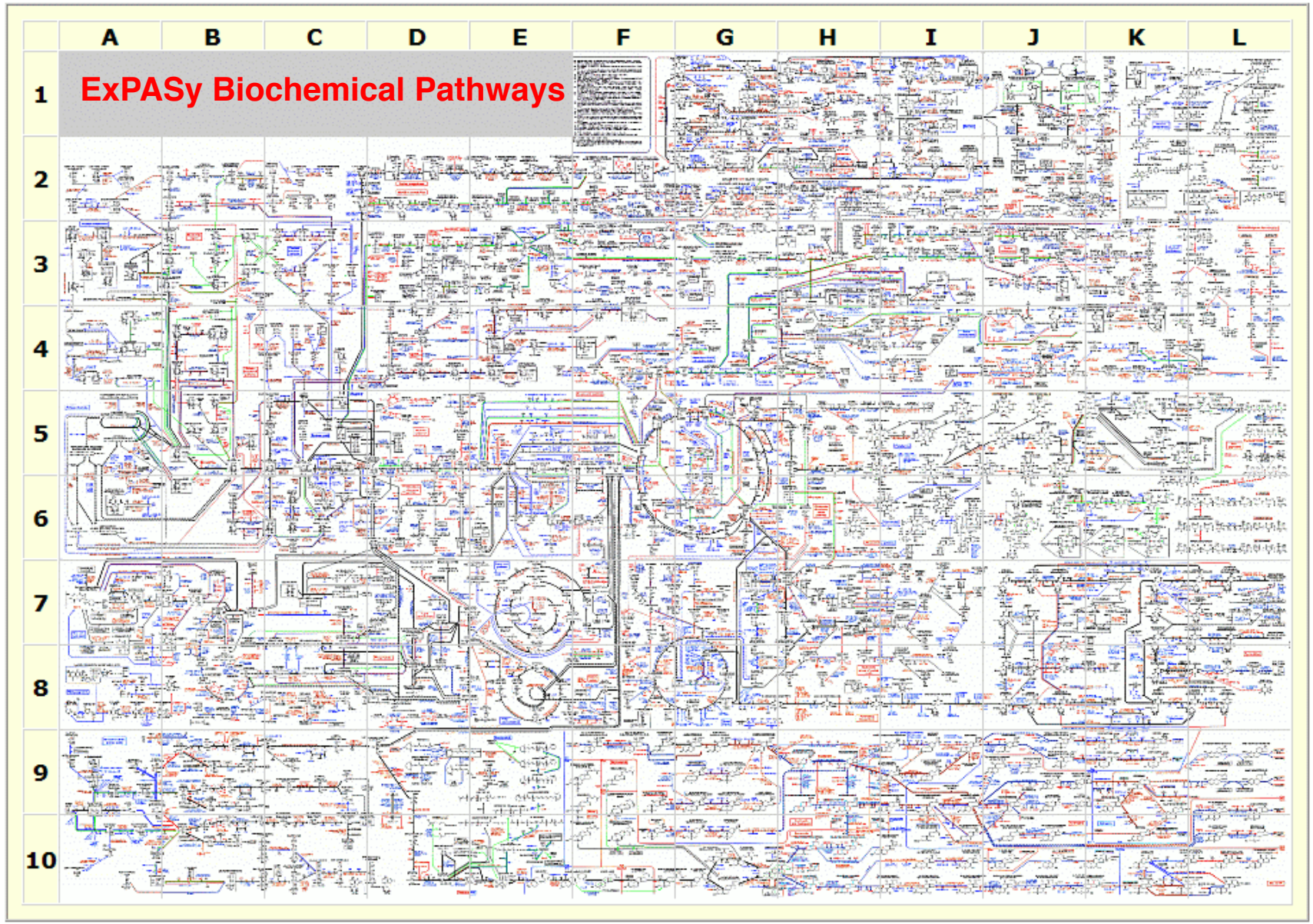
# Evolution of metabolic pathways

Shelley Copley

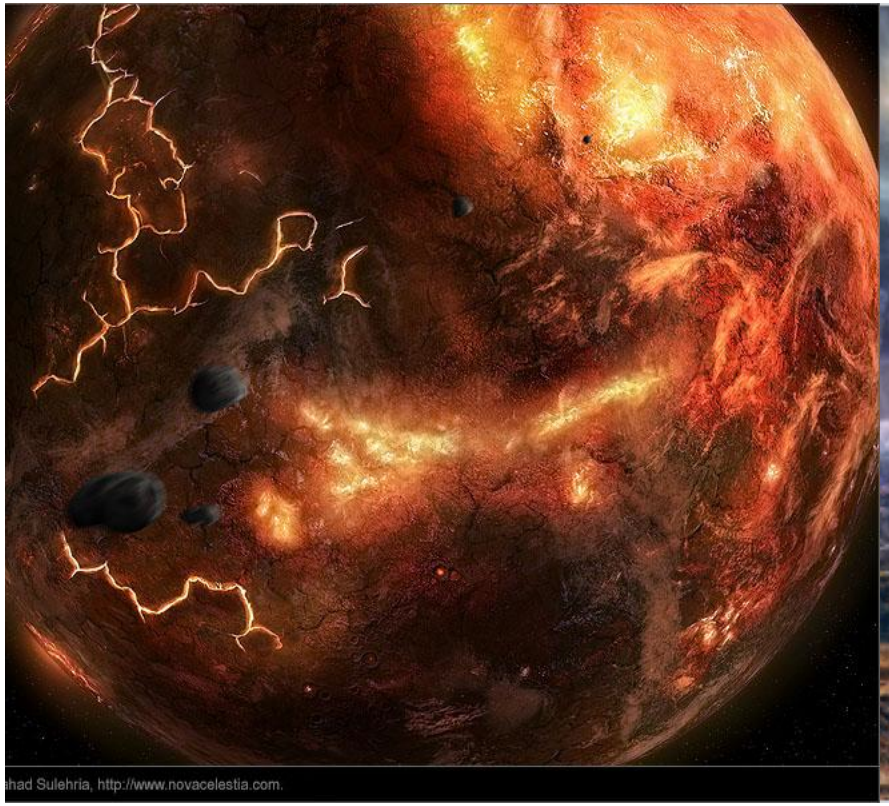
University of Colorado Boulder

# Where did all this come from?



# The Early Earth

4.5 billion years ago

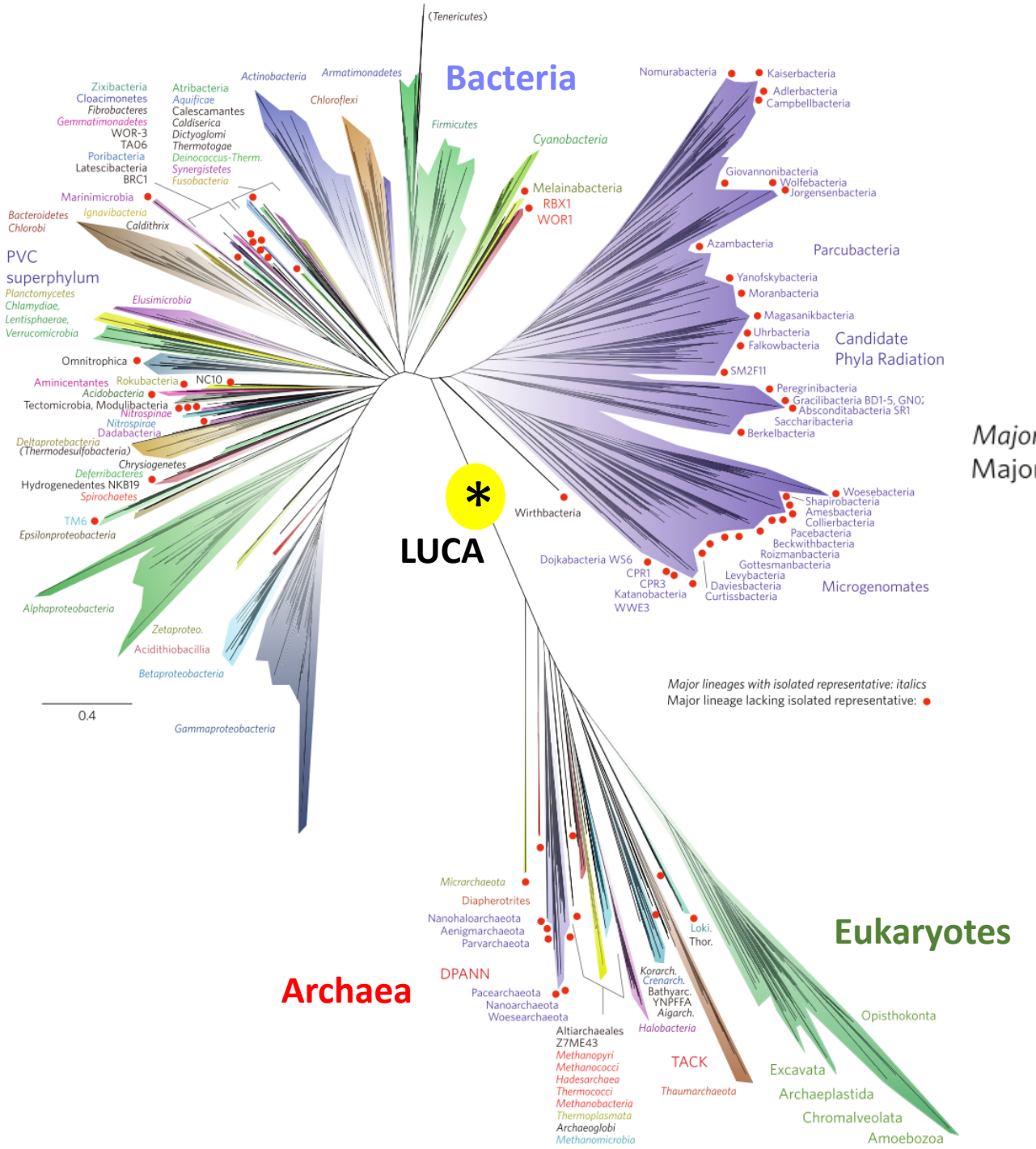



http://www.novacelestia.com/images/hadean_earth_space_art.jpg&imgrefurl

4.2 billion years ago

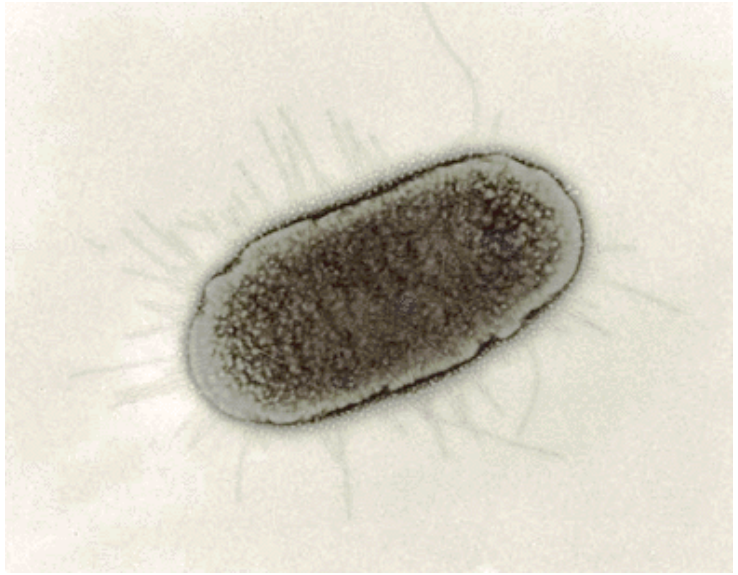
http://www.nytimes.com/2008/12/02/science/02eart.html

# The last universal common ancestor (LUCA)




*Major lineages with isolated representative: italics*
Major lineage lacking isolated representative: ●

# The LUCA



**What we know:**

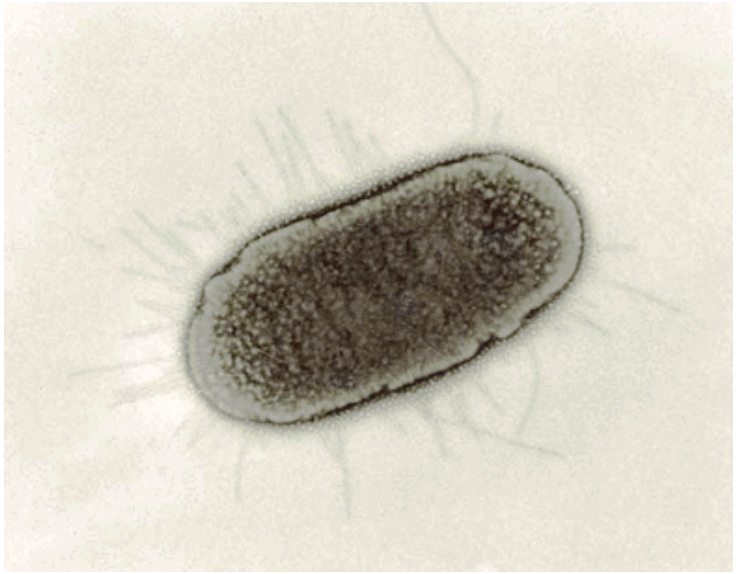
**It was microbial**
**It had ribosomes**
**proteins**
**compartmentalization**

# The proteome of the LUCA

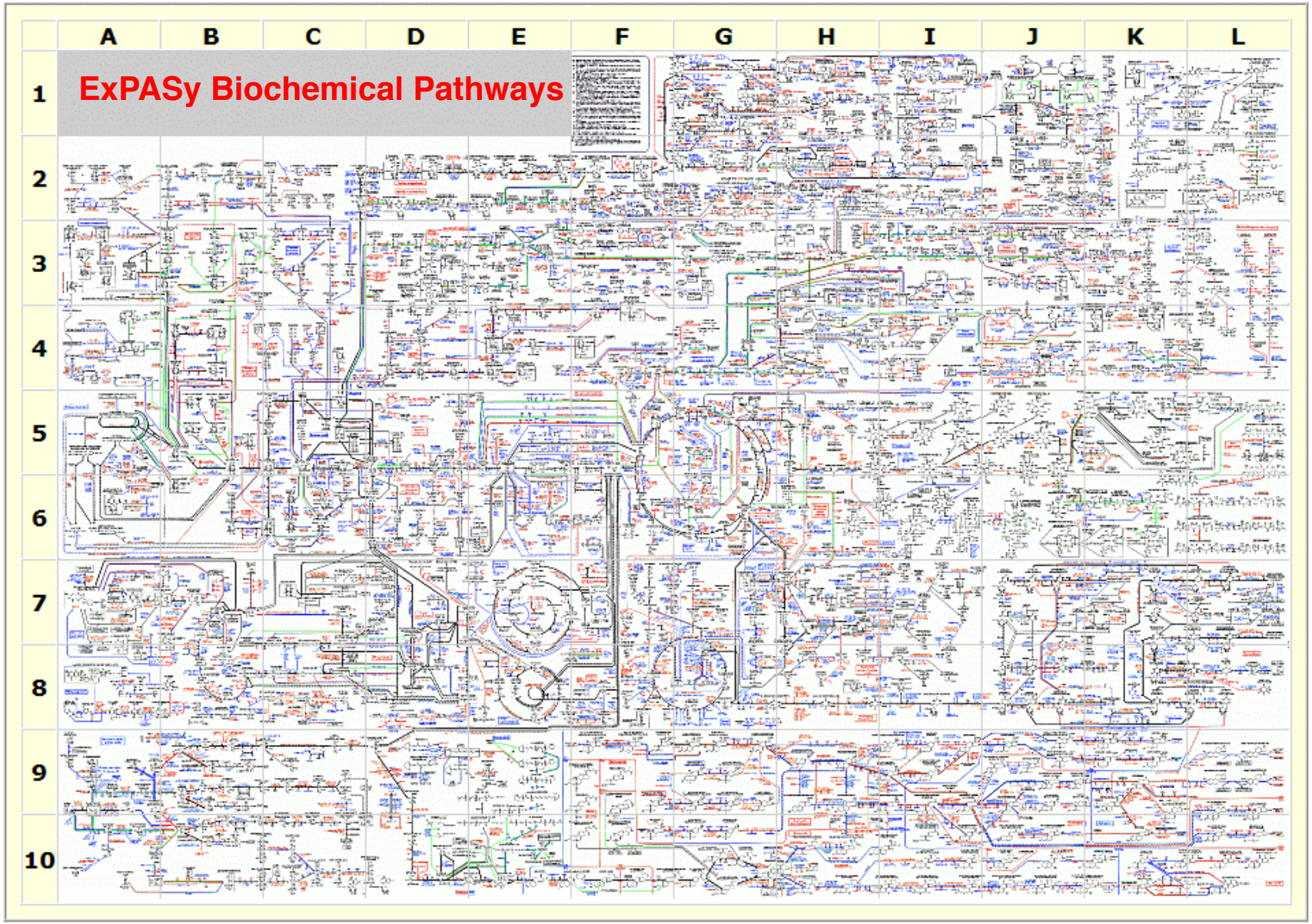
**Approx 669 genes** (*Res. Microbiol. 157*, 57-68, 2006)

**Including enzymes for synthesis of**

- amino acids
- nucleotides
- sugars
- fatty acids
- cofactors
- ATP



# Almost every reaction in this network is catalyzed by an enzyme

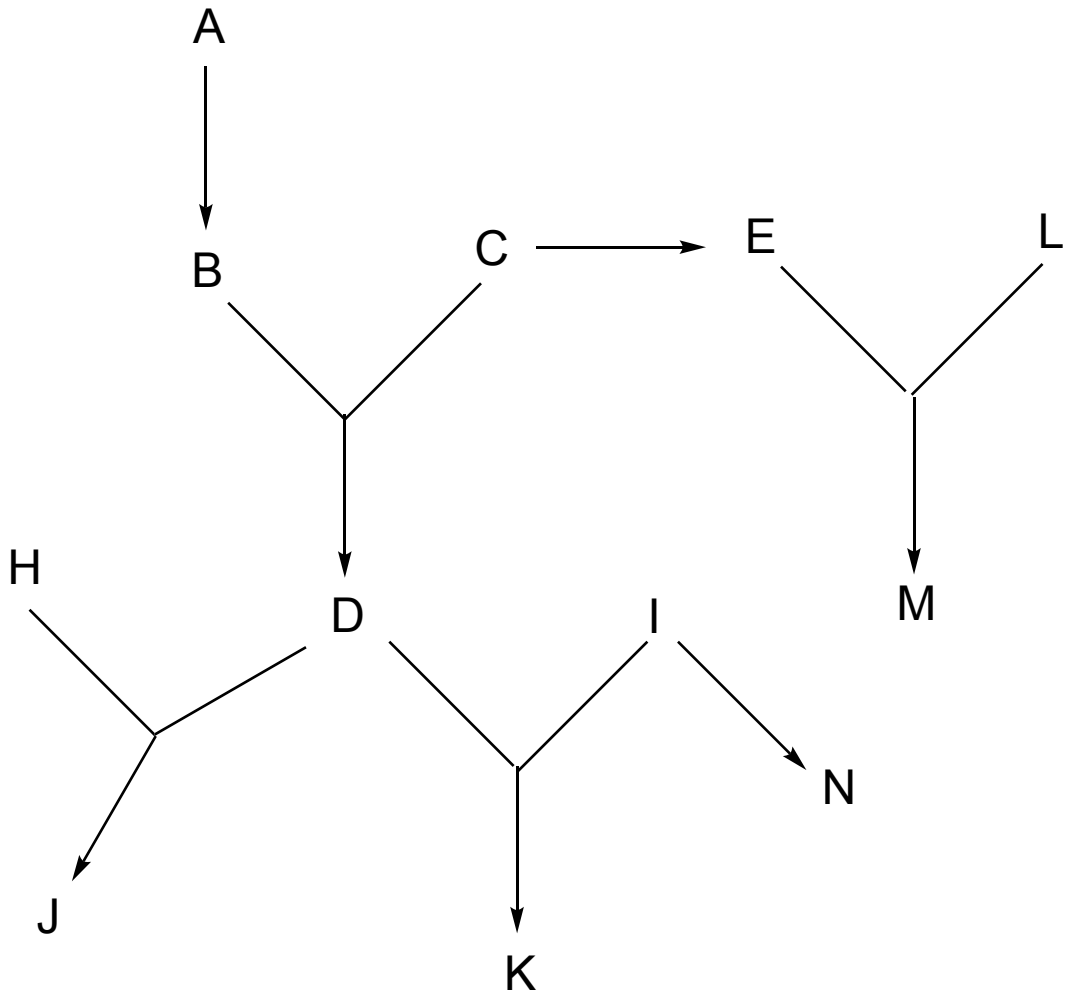


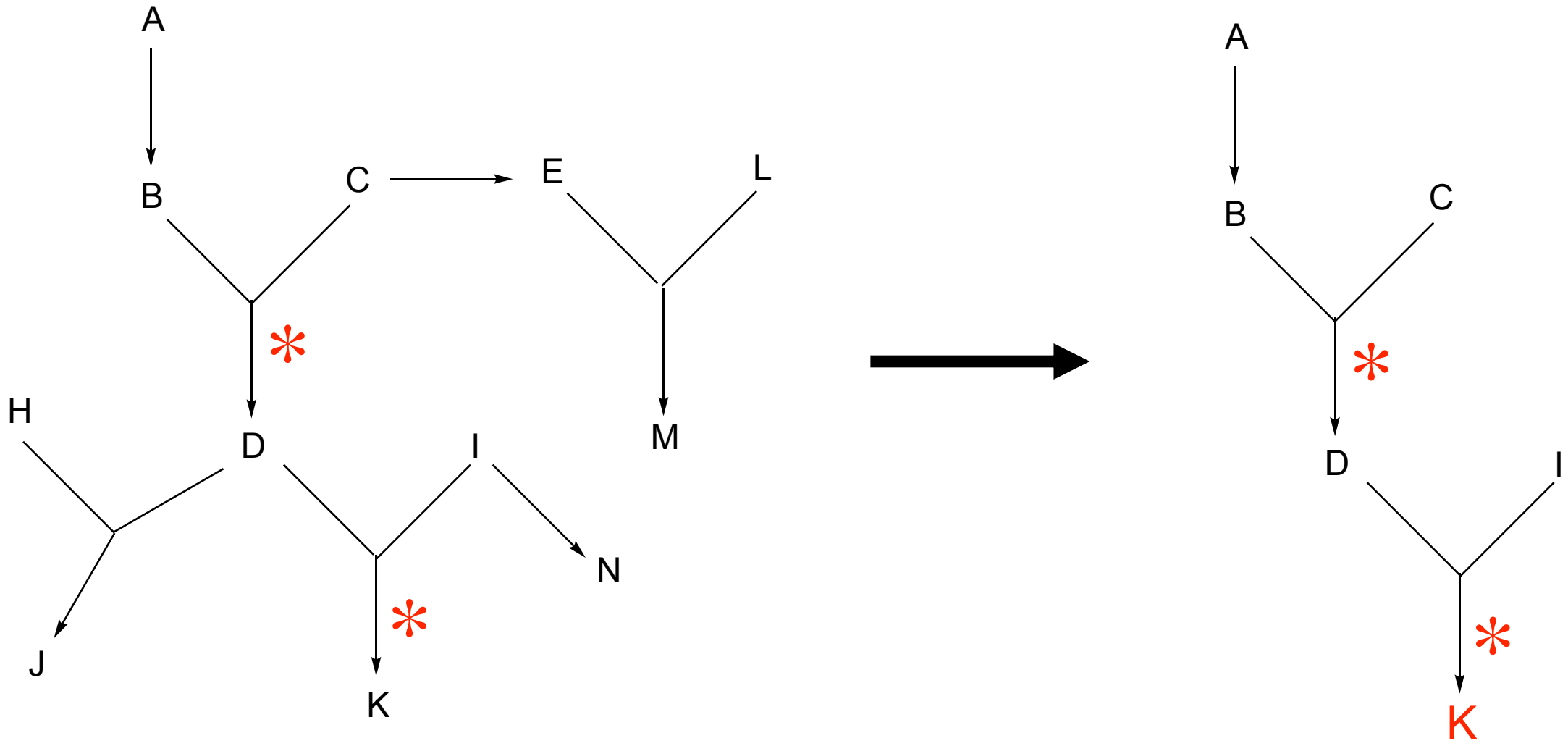
# Why catalysis is needed

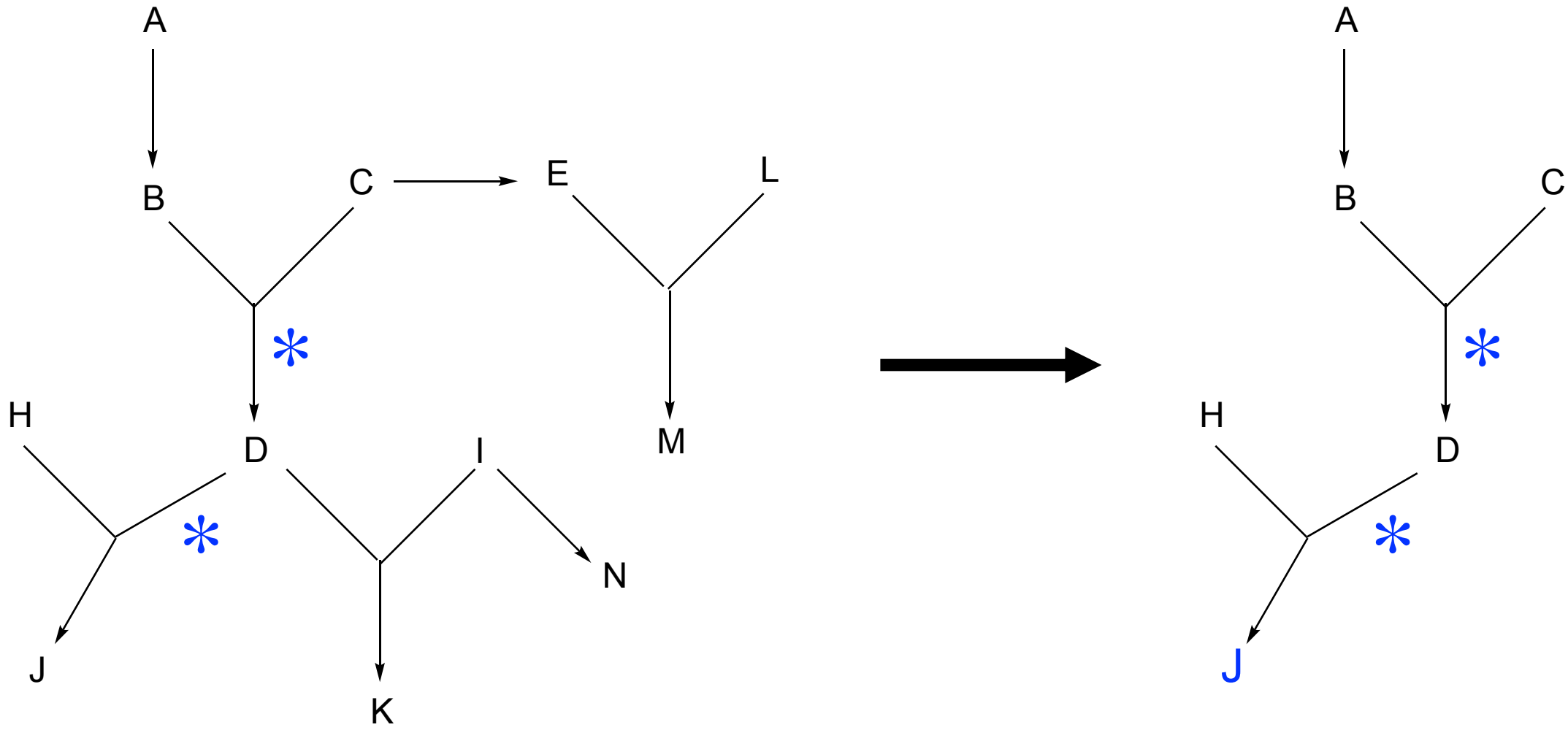
| reaction | $t_{1/2}$ |
|---|---|
| triose isomerization | 2 days |
| ester hydrolysis | 4 years |
| phosphomonoester hydrolysis | >500,000 years |
| fumarate hydration | 700,000 years |
| phosphodiester hydrolysis | >13 million years |
| OMP decarboxylation | 1.1 billion years |

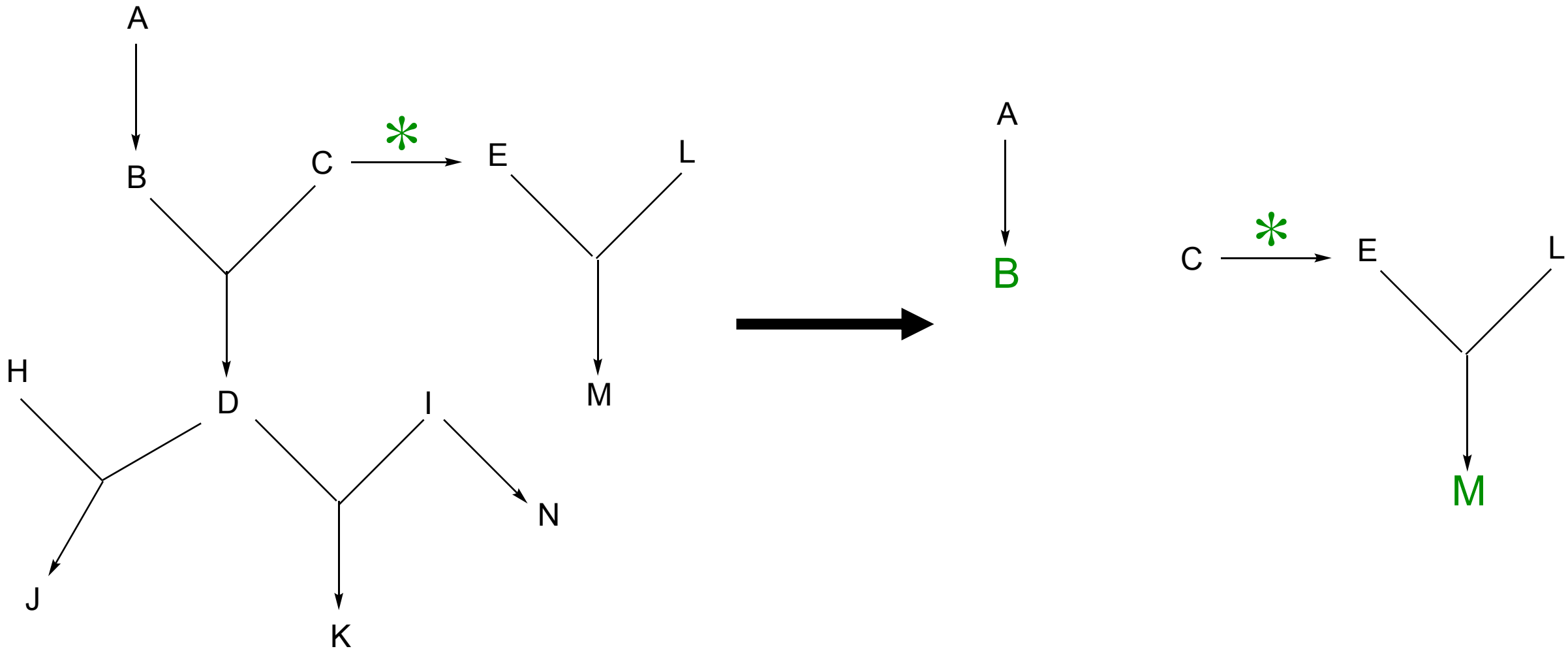
Wolfenden, Acc. Chem. Res. 34, 938, 2001

# Catalysts structure metabolism by accelerating specific reactions



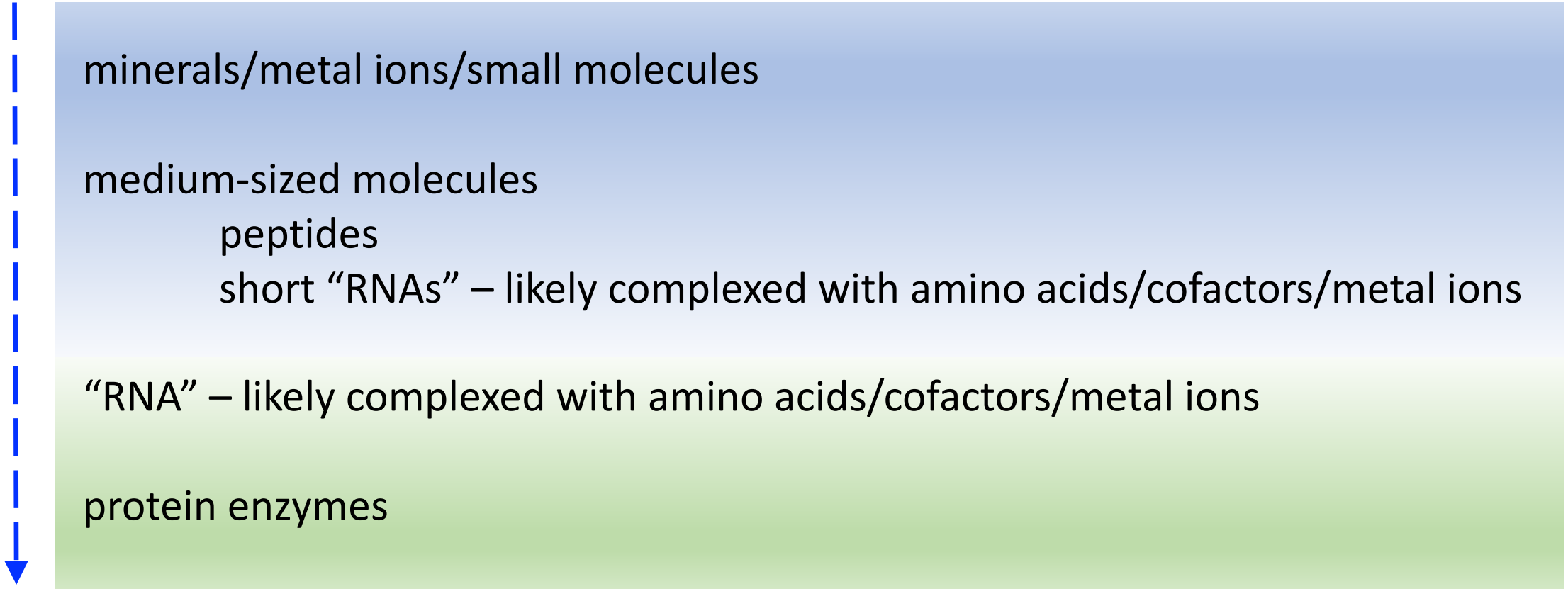

A
B
C
E
L
M
H
D
I
N
J
K
A
B
C
D
I
K


A
B
C
E
L
M
H
D
I
N
J
K
A
B
C
H
D
J

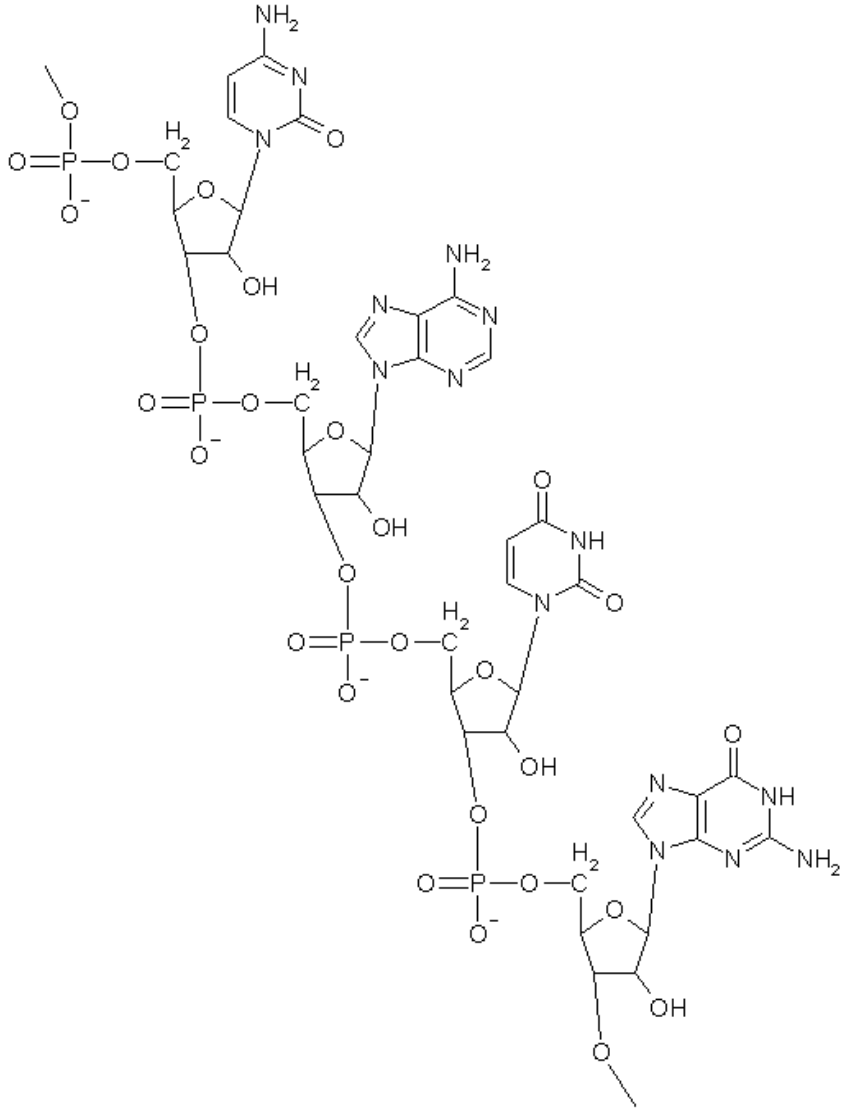

A
B
C
*
E
L
M
H
D
I
N
J
K
A
B
C
*
E
L
M

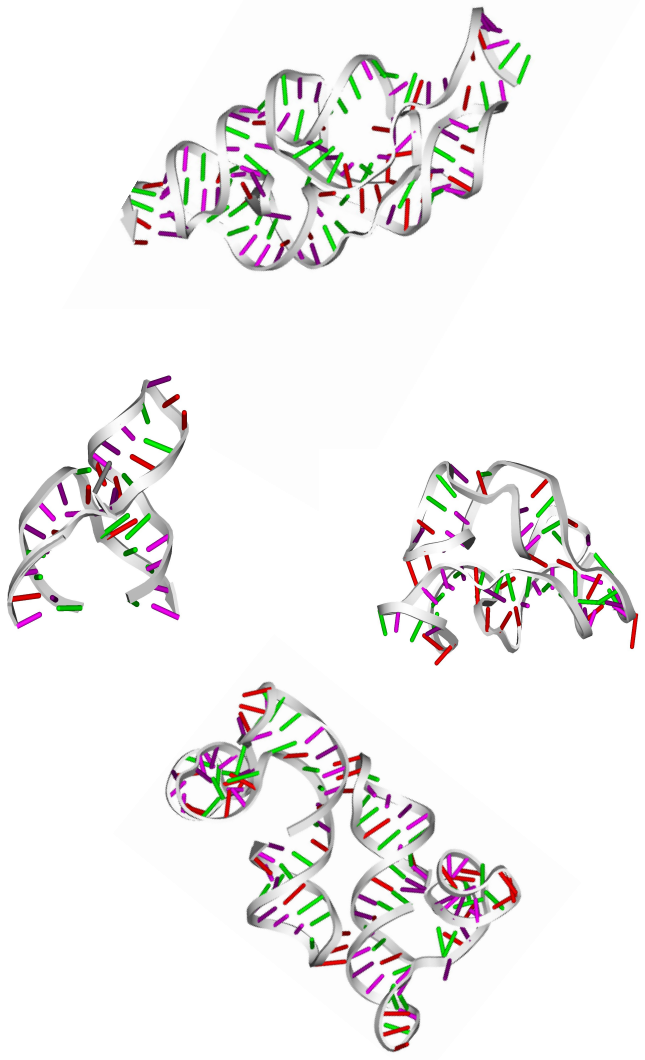
# The “evolution” of catalysts

**time**



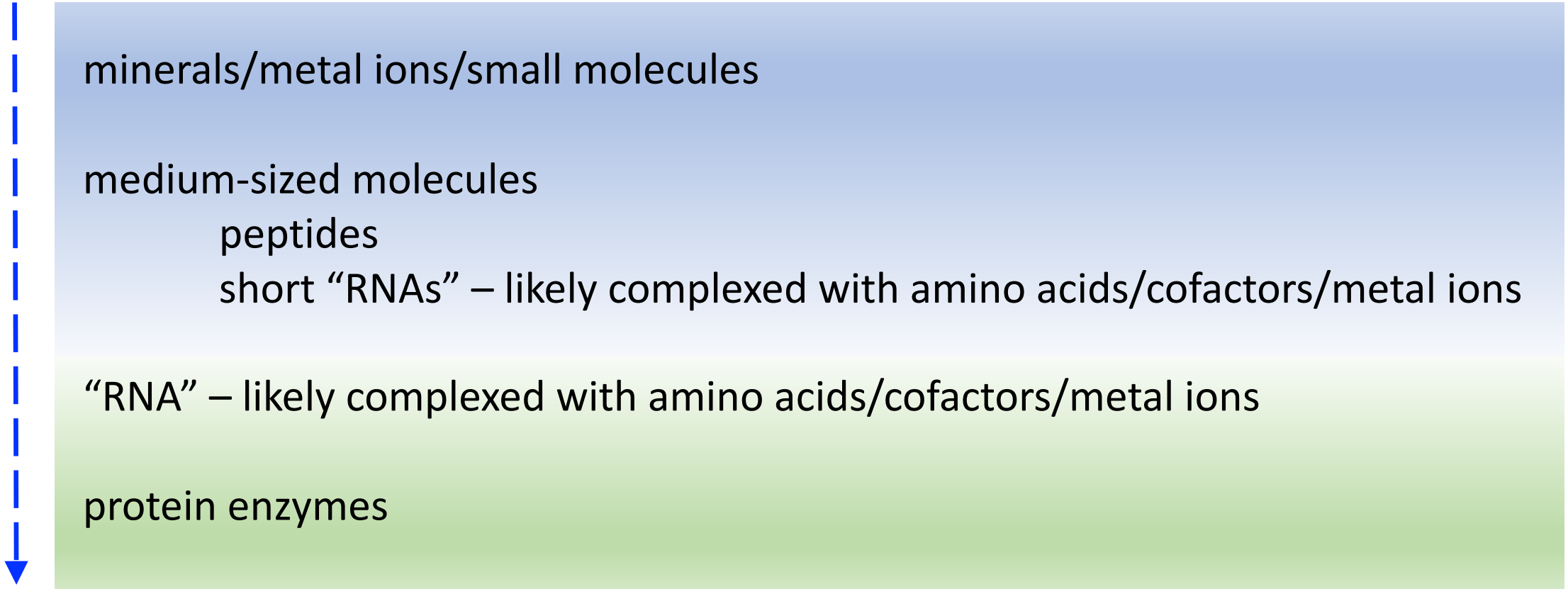
# RNA





# The “evolution” of catalysts

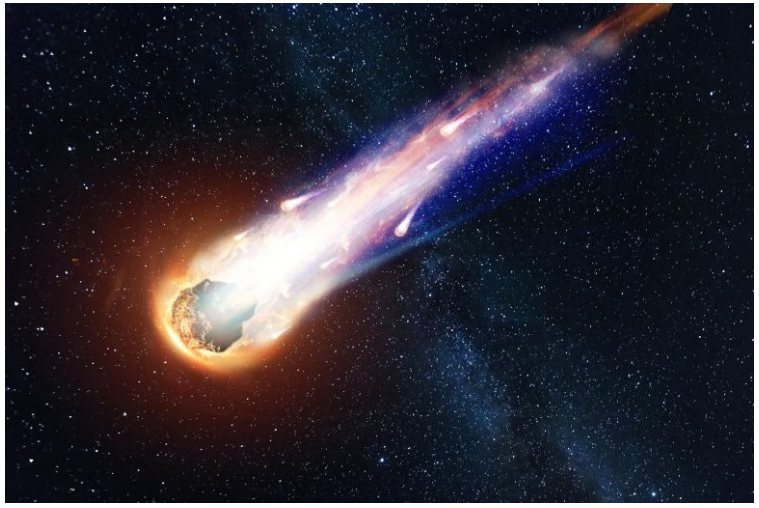
**time**



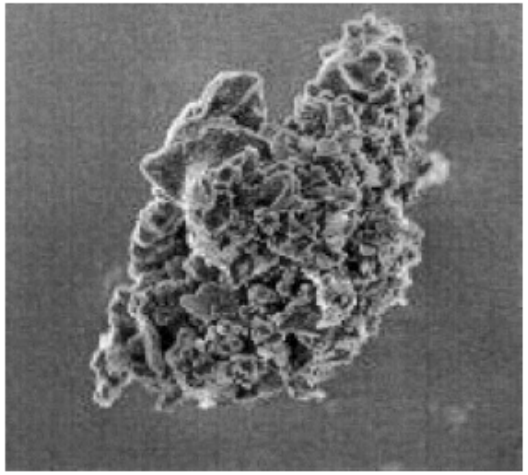
# What did the first catalysts have to work with?

http://www.skyshooter.net/images/Comets/Comet-Hale-Bopp.jpg

www.moonraker.com.au/techni/lightning-marine.htm

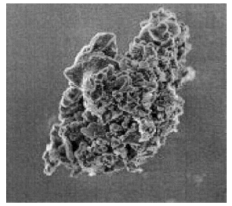


https://scitechdaily.com/first-evidence-of-bio-essential-sugars-in-meteorites/



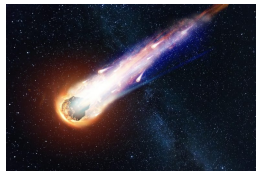
*Adv. Space Res.* 33, 57-66, 2004

# Estimated delivery of organic carbon 4.4 Gyr ago



Interplanetary dust particles $10^9$

Comets $> 10^6$



Meteorites $>10^4$

kg C $yr^{-1}$

*Nature* 355, 125, 1992

# Some meteorites contain thousands of organic compounds

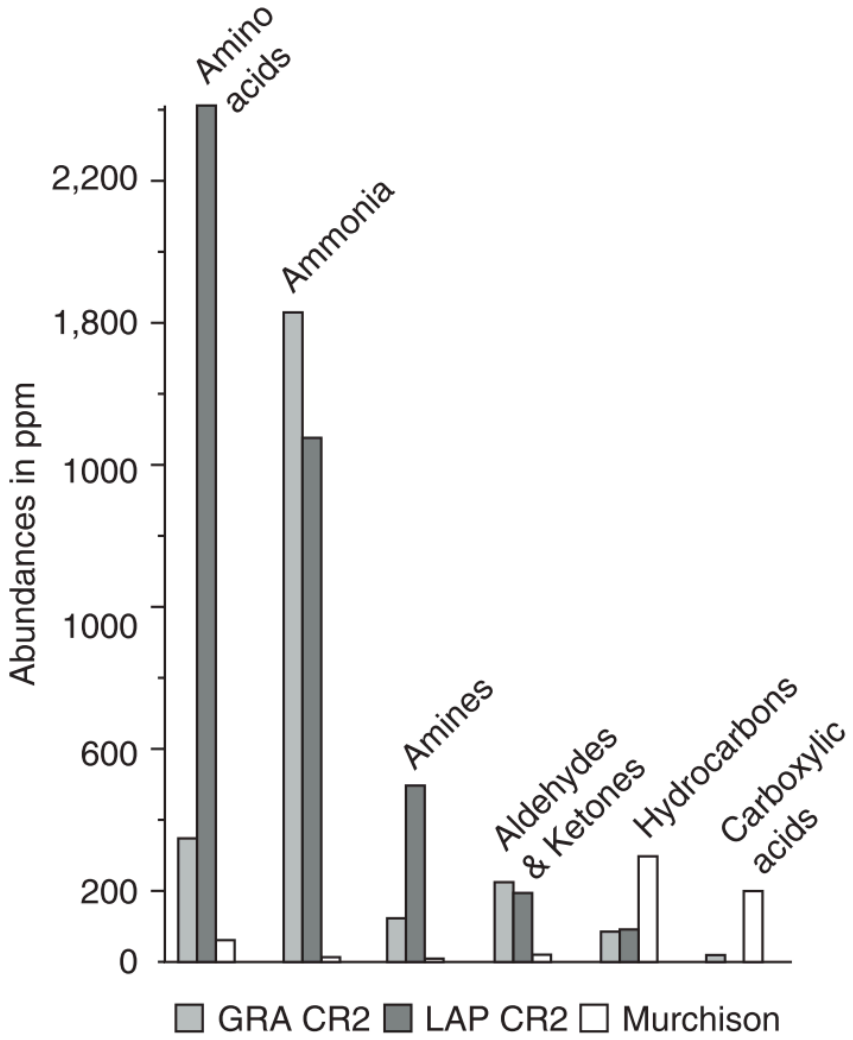
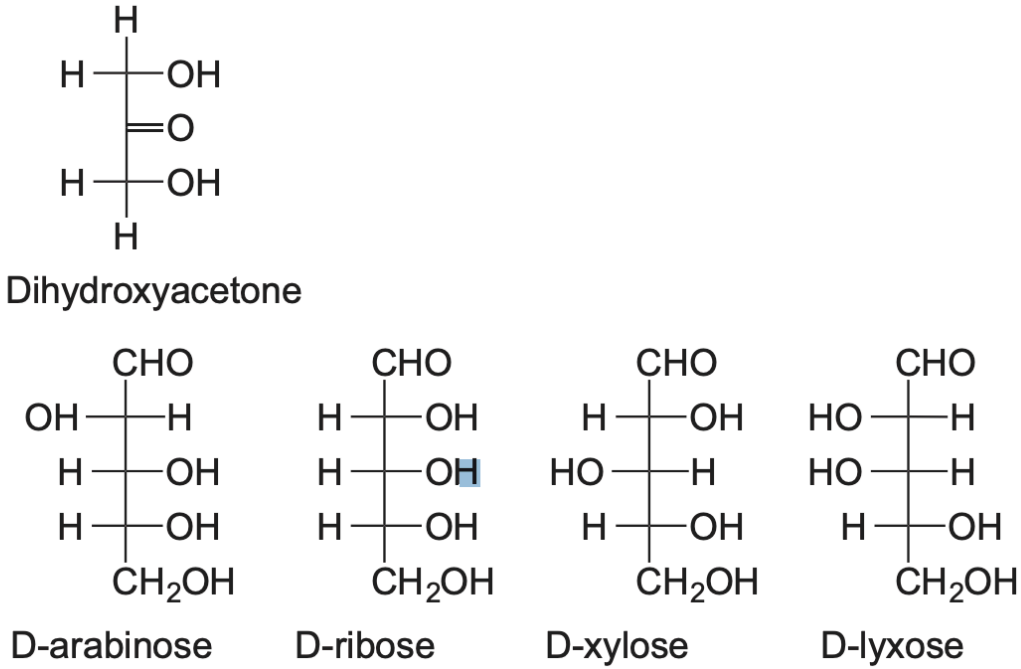



**Figure 4.** Comparative plot of major soluble organic compound abundances in the Murchison and CR2 meteorites (GRA 95229 and LAP 02342 shown).

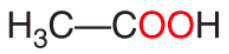
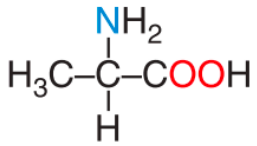
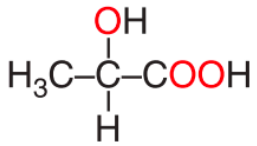
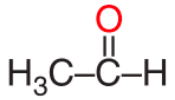
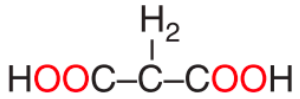
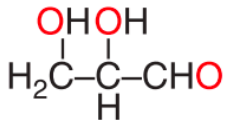
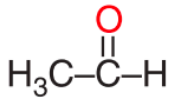
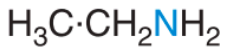
Cold Spring Harb Perspect Biol 2010;2:a002105

sugars in the Murchison meteorite




PNAS 116, 24440–24445, 2019

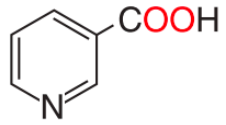
**Table 1.** Classes of organic compounds in the Murchison meteorite.

| Compound Class | Structure & Example Molecule | |
|---|---|---|
| Carboxylic acids |  | Acetic acid |
| Amino acids |  | Alanine |
| Hydroxy acids |  | Lactic acid |
| Ketoacids |  | Pyruvic acid |
| Dicarboxylic acids |  | Succinic acid |
| Sugar alcohols & acids |  | Glyceric acid |
| Aldehydes & Ketones |  | Acetaldehyde |
| Amines & Amides |  | Ethyl amine |

found in life

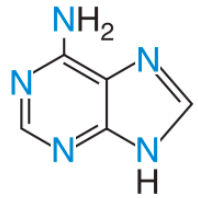
Cold Spring Harb Perspect Biol 2010;2:a002105

Pyridine carb. acids




Nicotinic acid

Purines & Pyrimidines




Adenine

found in life

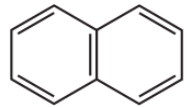
Hydrocarbons:

Alyphatic

$H_3C–CH_2–CH_3$

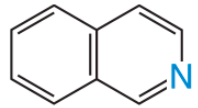
Propane

Aromatic

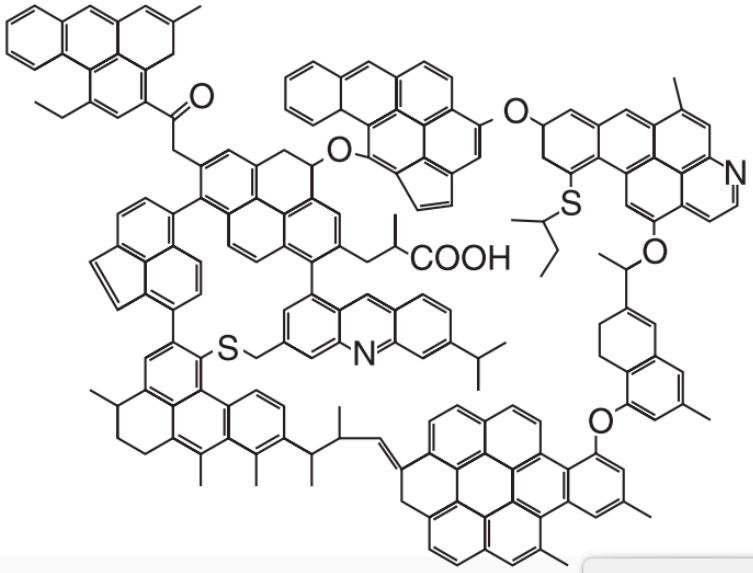


Naphthalene

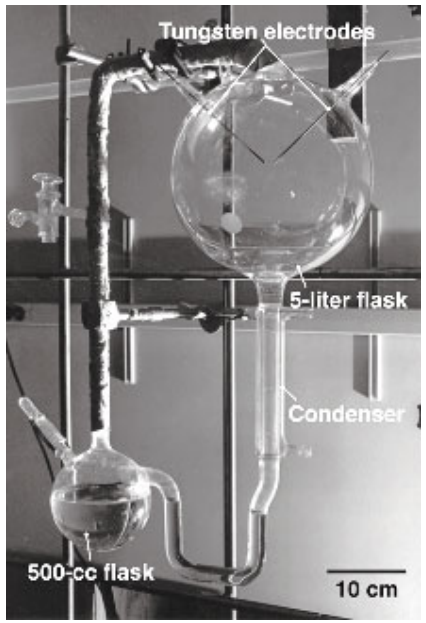
Polar




Isoquinoline




Insoluble Material (estimated)




http://complex.upf.es/
~josep/research.html

## Proteinogenic amino acids

glycine
aspartate
glutamate
valine
proline
leucine
isoleucine
serine
threonine

## Non-proteinogenic Amino acids

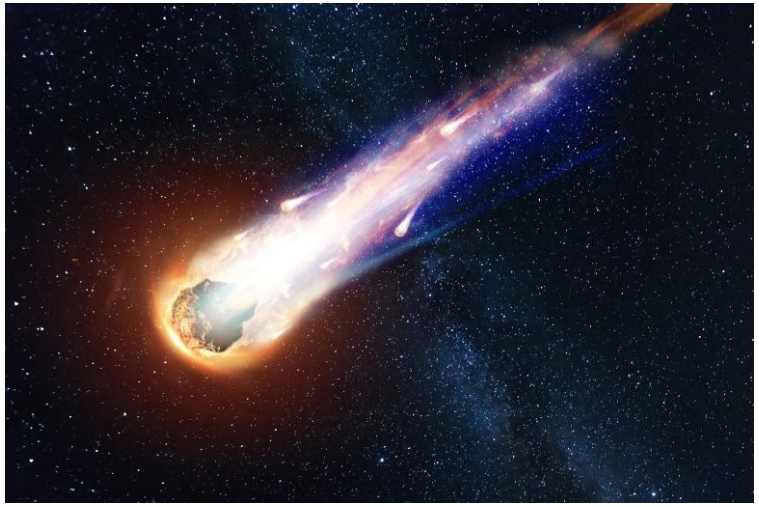
sarcosine
β-alanine
α-amino-n-butyric acid
α -aminoisobutyric acid
isovaline
norvaline
alloisoleucine
norleucine
allothreonine
α -hydroxy-γ-aminobutyric acid
α,γ-diaminobutyric acid
α, β -diaminopropionic acid
isoserine

# Life may have originated in a warm little pond

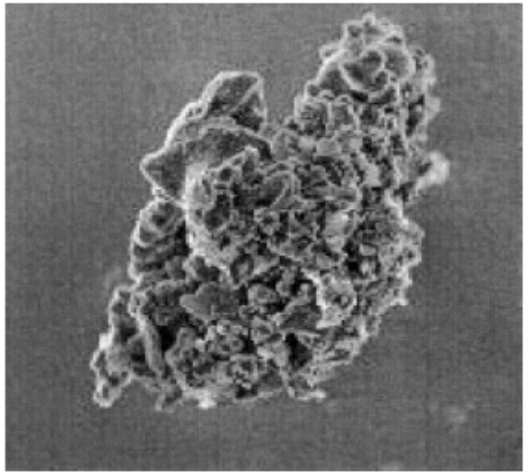
# complex mixtures; only partial overlap with biomolecules

http://www.skyshooter.net/images/Comets/Comet-Hale-Bopp.jpg

www.moonraker.com.au/techni/lightning-marine.htm

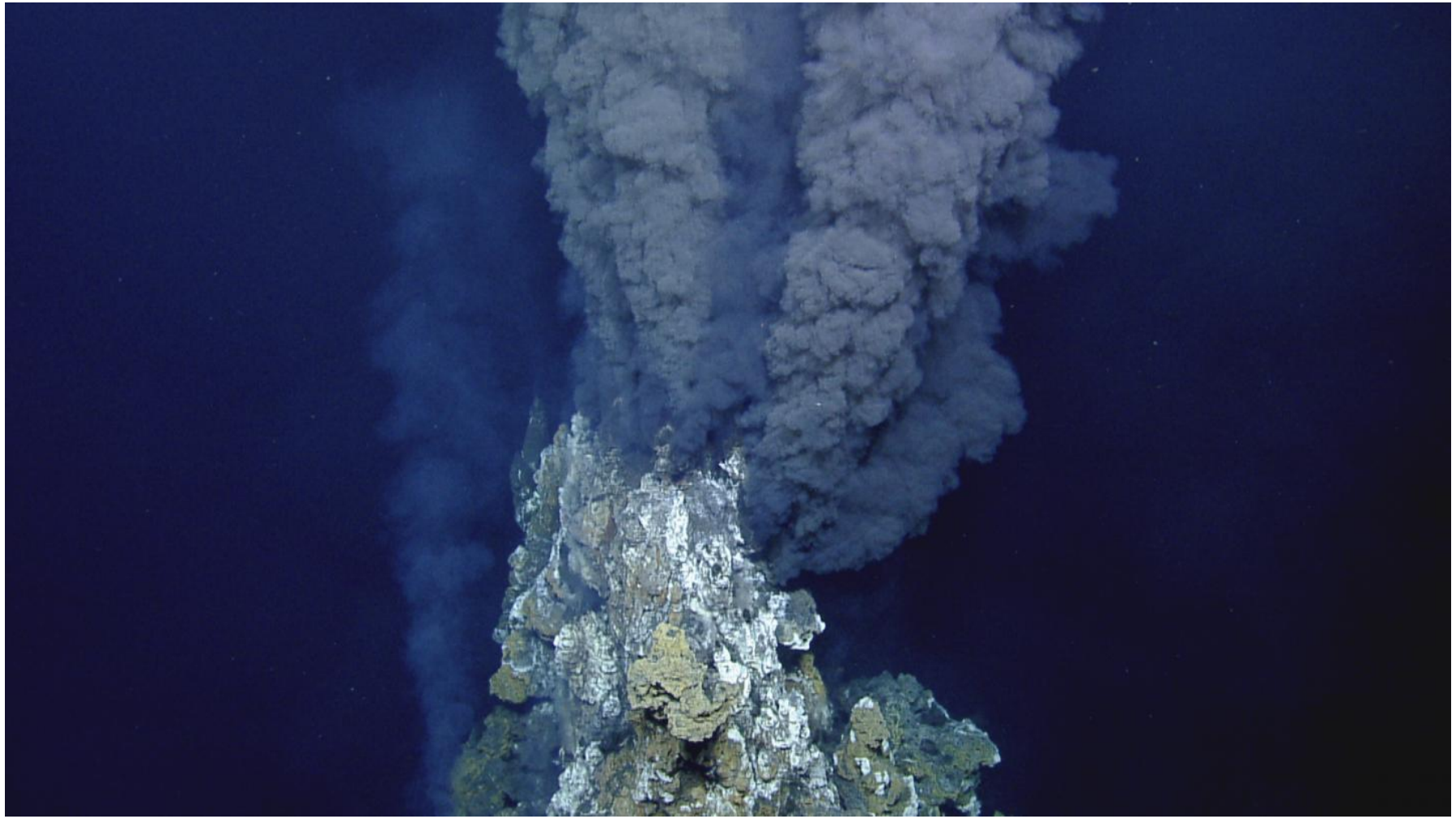


https://scitechdaily.com/first-evidence-of-bio-essential-sugars-in-meteorites/

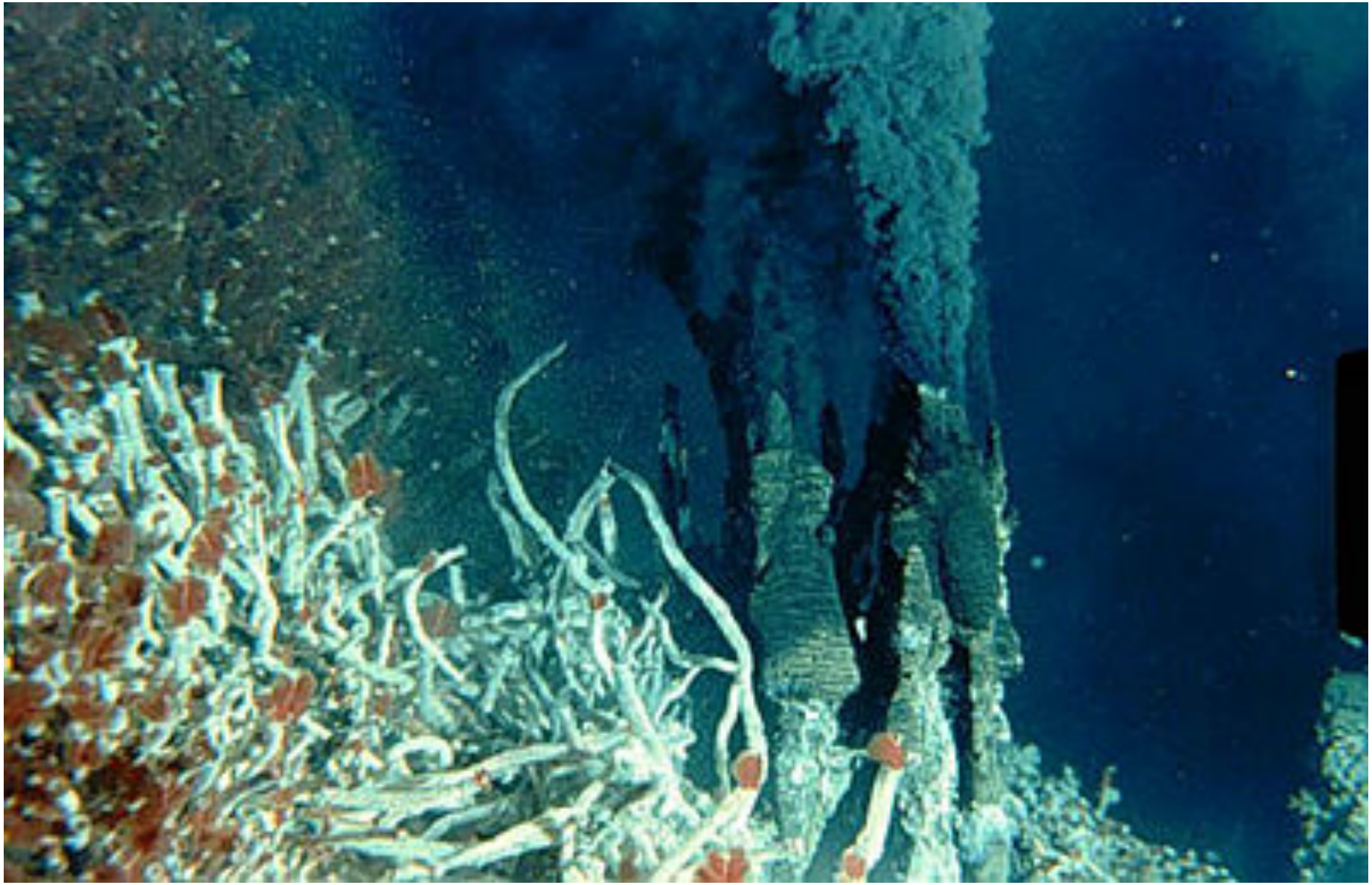


*Adv. Space Res.* 33, 57-66, 2004

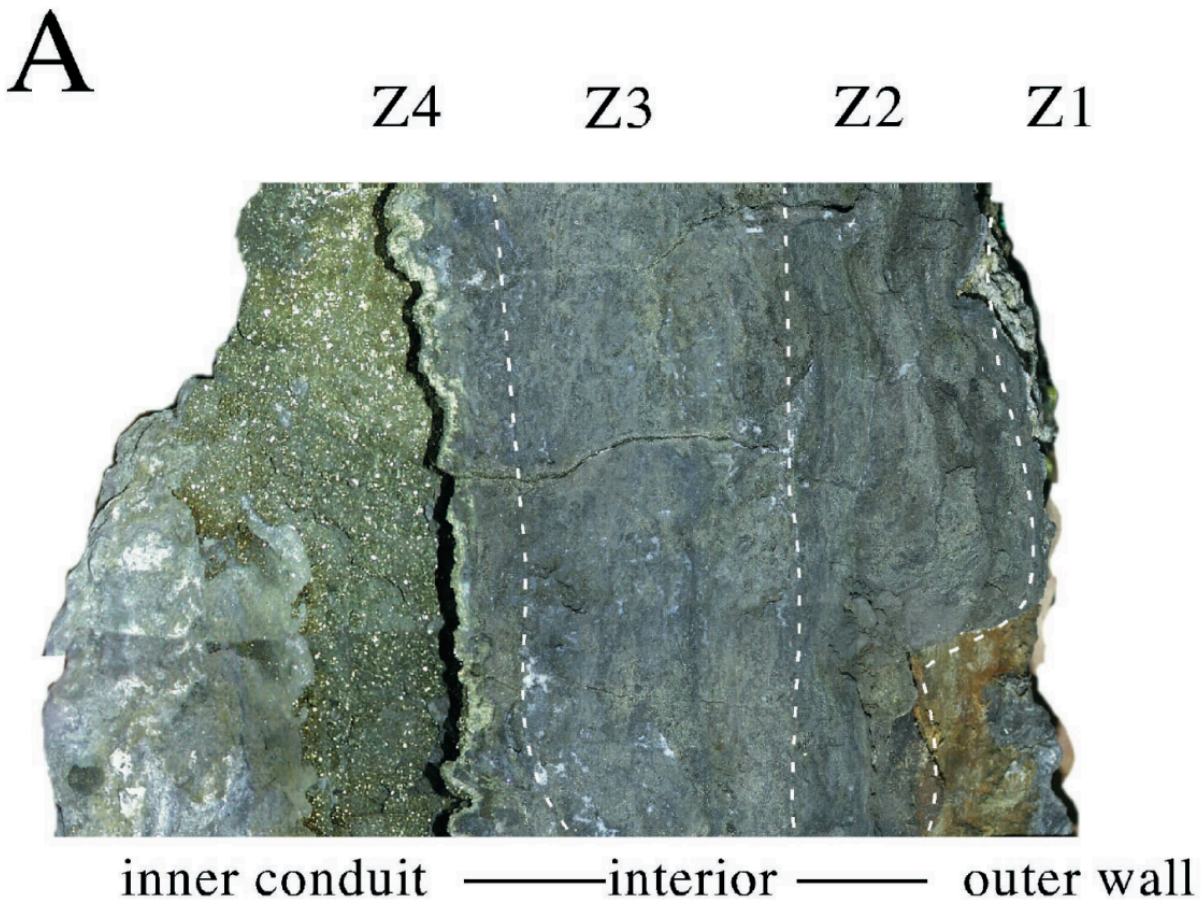
# An alternative – an origin in hydrothermal vents
# Discovered 1979

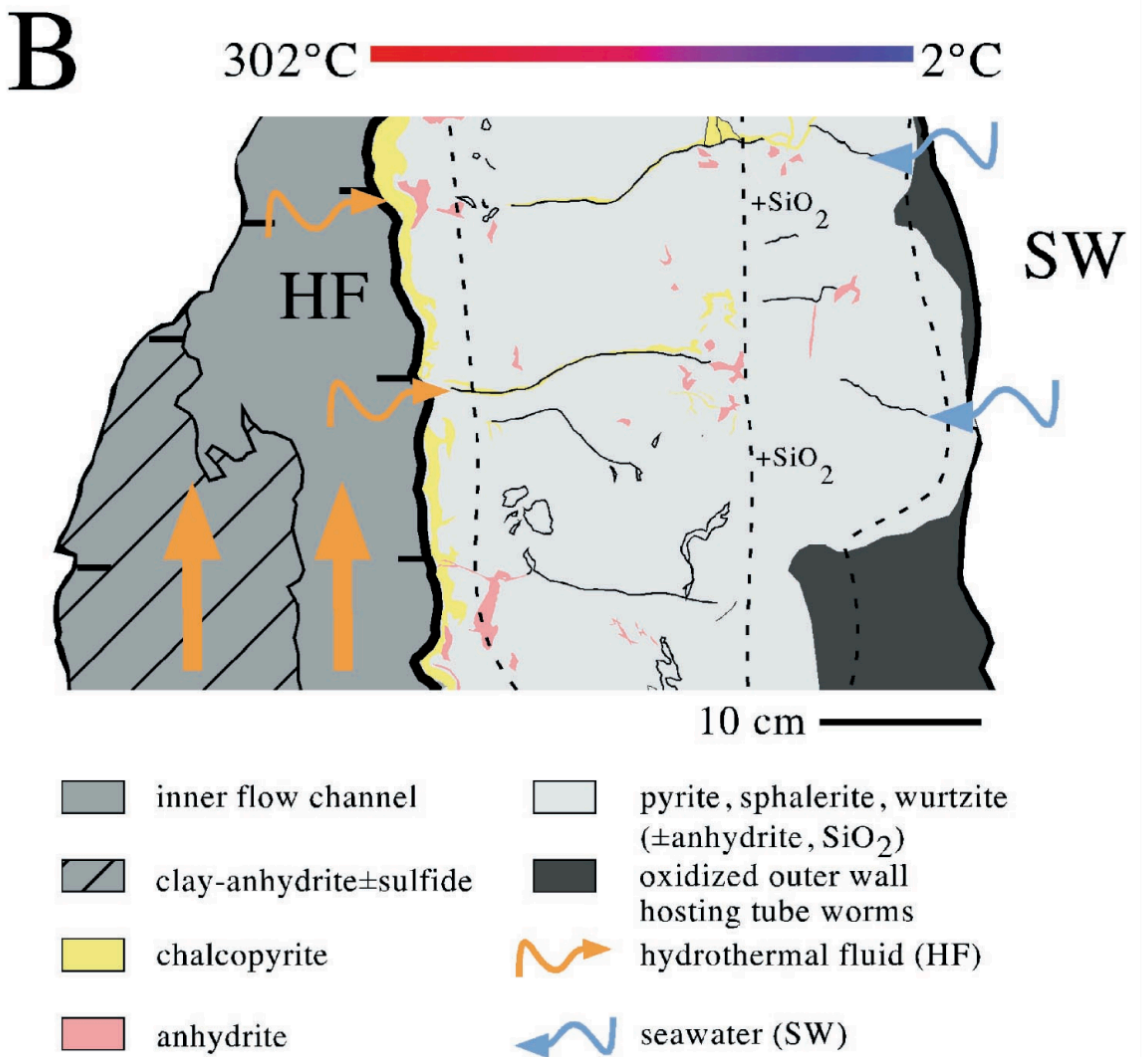


# Hydrothermal vents host lush ecosystems



# Cross-section through a cut-face of the sulfide chimney “Finn”



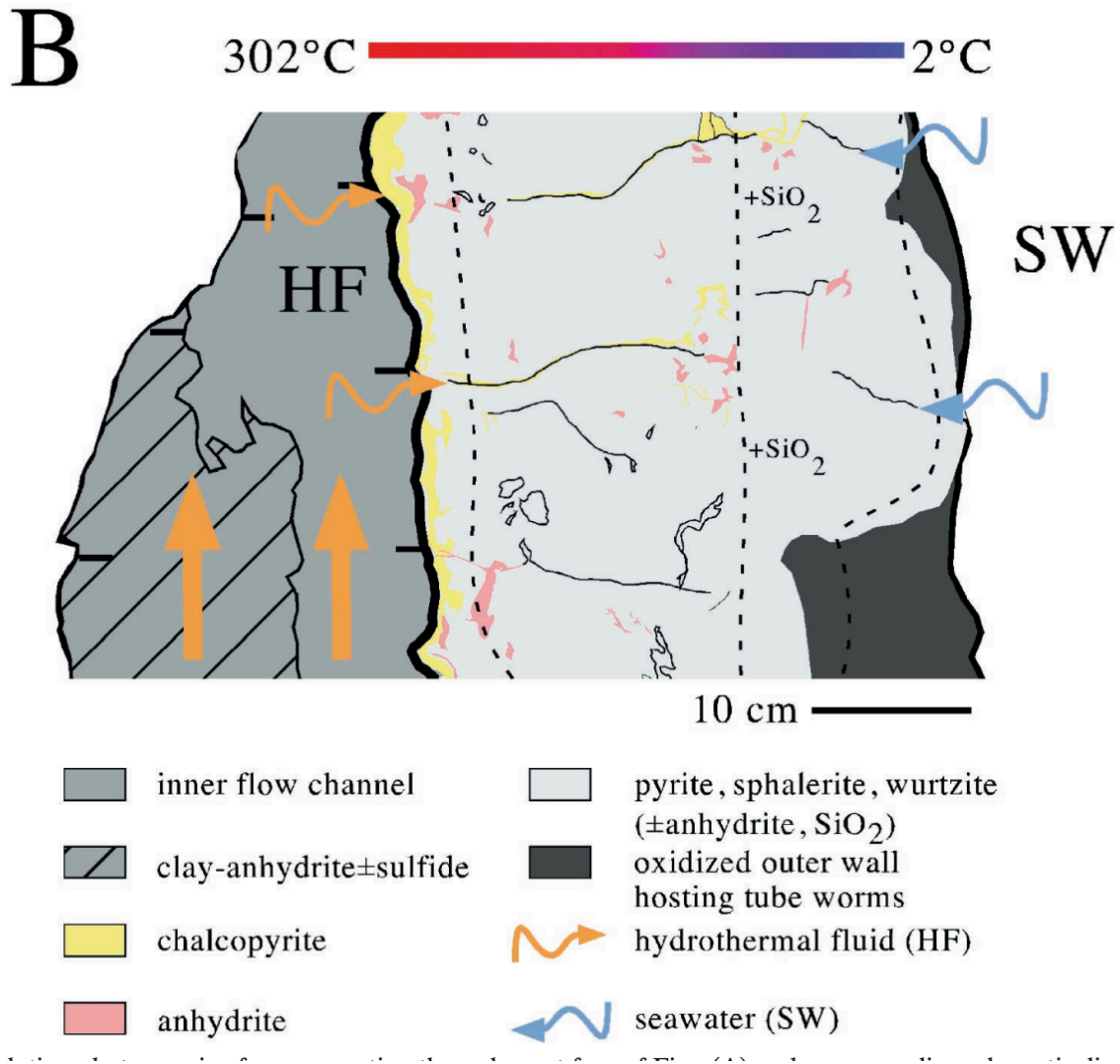



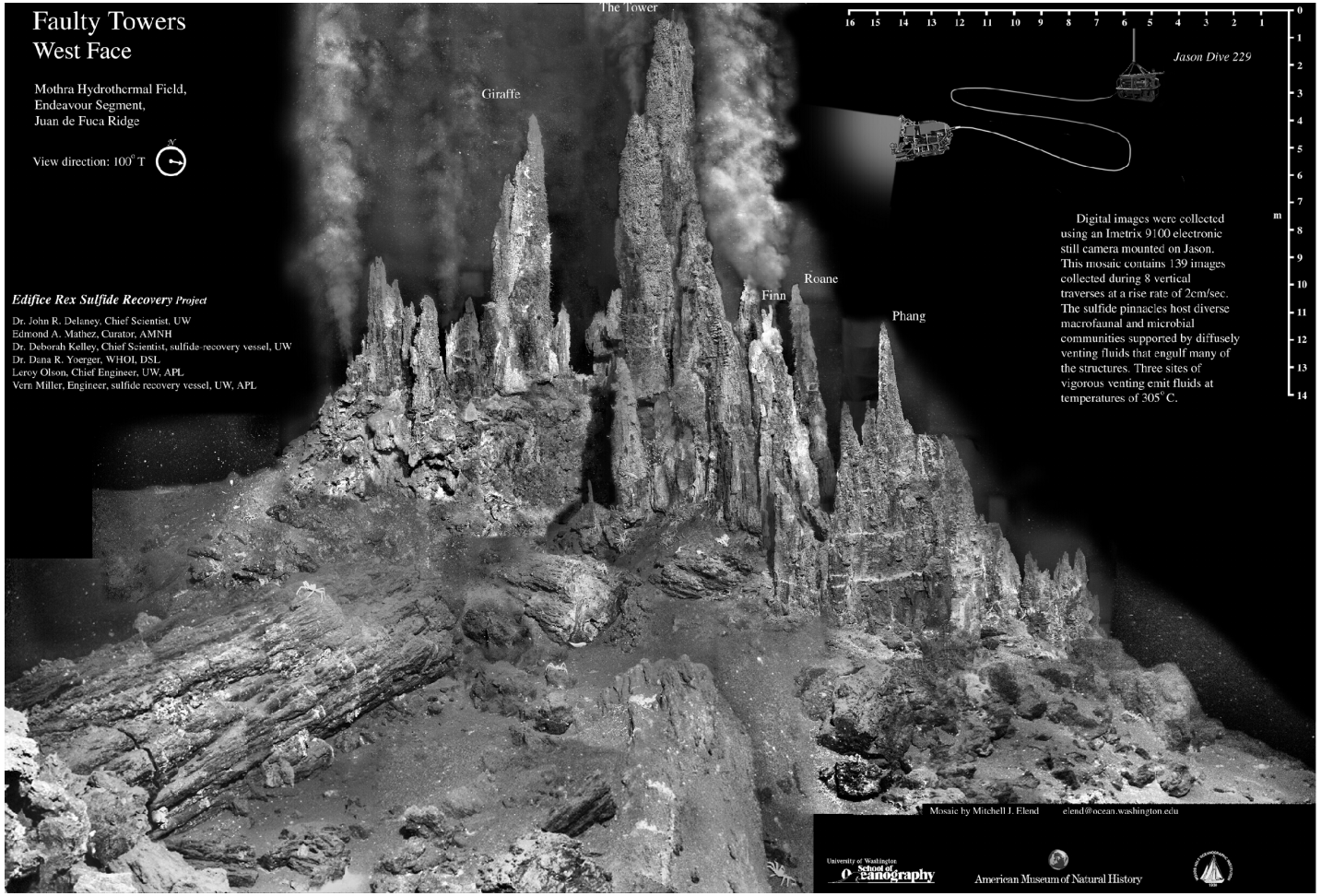

Zones 1-3 – 2 x$10^6$ – 2 x $10^8$ microbes/g sulfide
Zone 4 – 2 x $10^5$ microbes/g sulfide

•Applied and Environmental Microbiology 69(6):3580-92, 2003

Porous walls provide

1) compartmentalization
2) catalytic surfaces
3) continuous supply of reactive small molecules






# Black smokers

Acidic vent fluids contain

$H_2$

$CH_4$

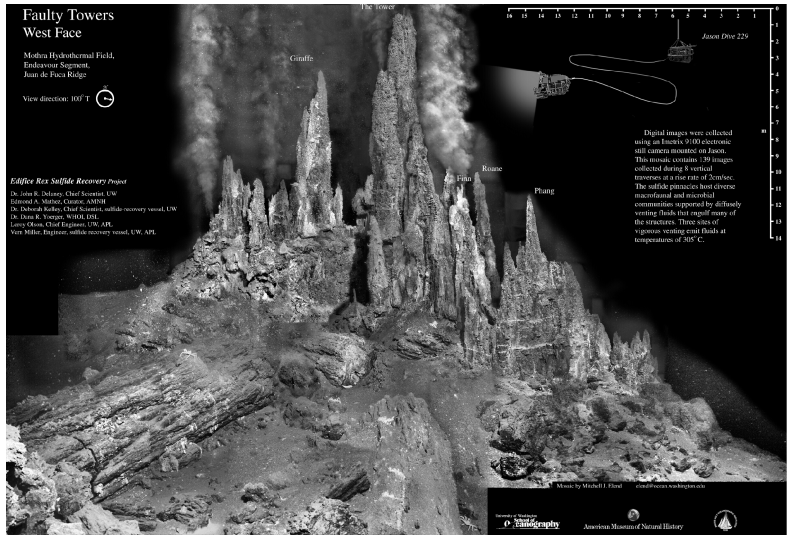
$CO_2$

$H_2S$

$NH_4^+$

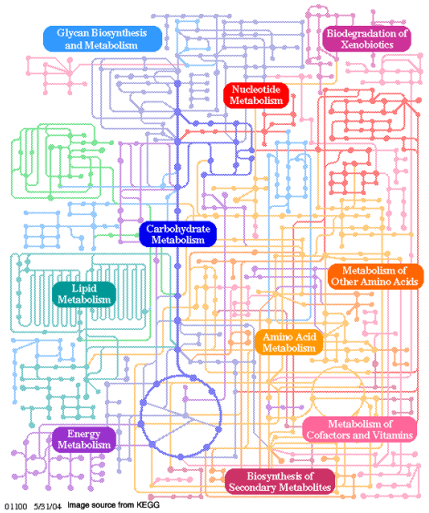
# The Lost City Hydrothermal Field Discovered 2000

- 60 m tall carbonate towers
- Venting <40-90°C diffuse fluids, low metals, low silica, pH's 9-11
- Fluids are enriched in methane, hydrogen, and other hydrocarbons




requires build-up of a sparse set of metabolites

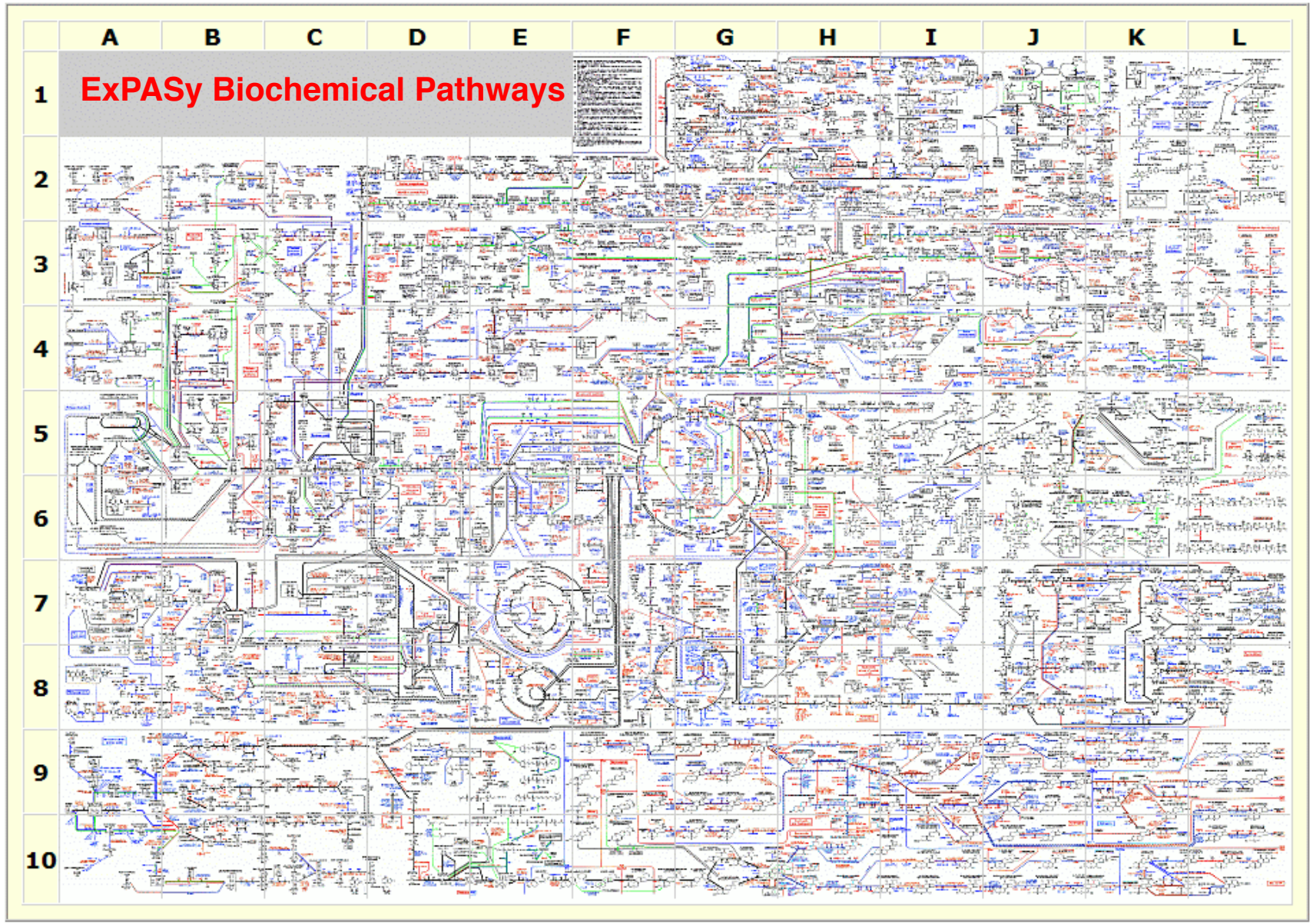







requires pruning of a complex set of metabolites

# Where did all this come from?



1945

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 31 June 15, 1945 Number 6

*Copyright 1945 by the National Academy of Sciences*

---

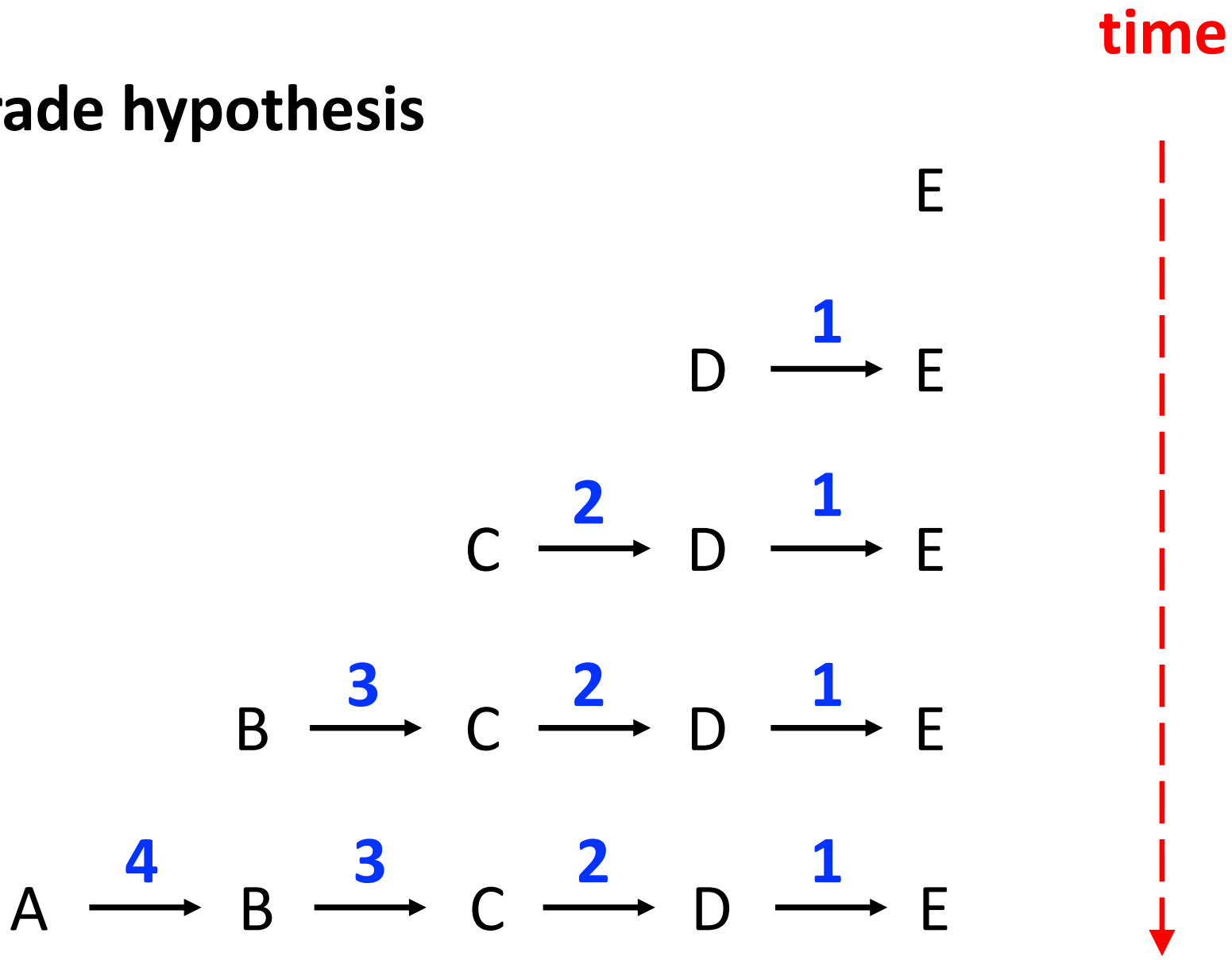
# *ON THE EVOLUTION OF BIOCHEMICAL SYNTHESES*

By N. H. Horowitz

School of Biological Sciences, Stanford University, Calif.

Communicated April 23, 1945

# The retrograde hypothesis



1965

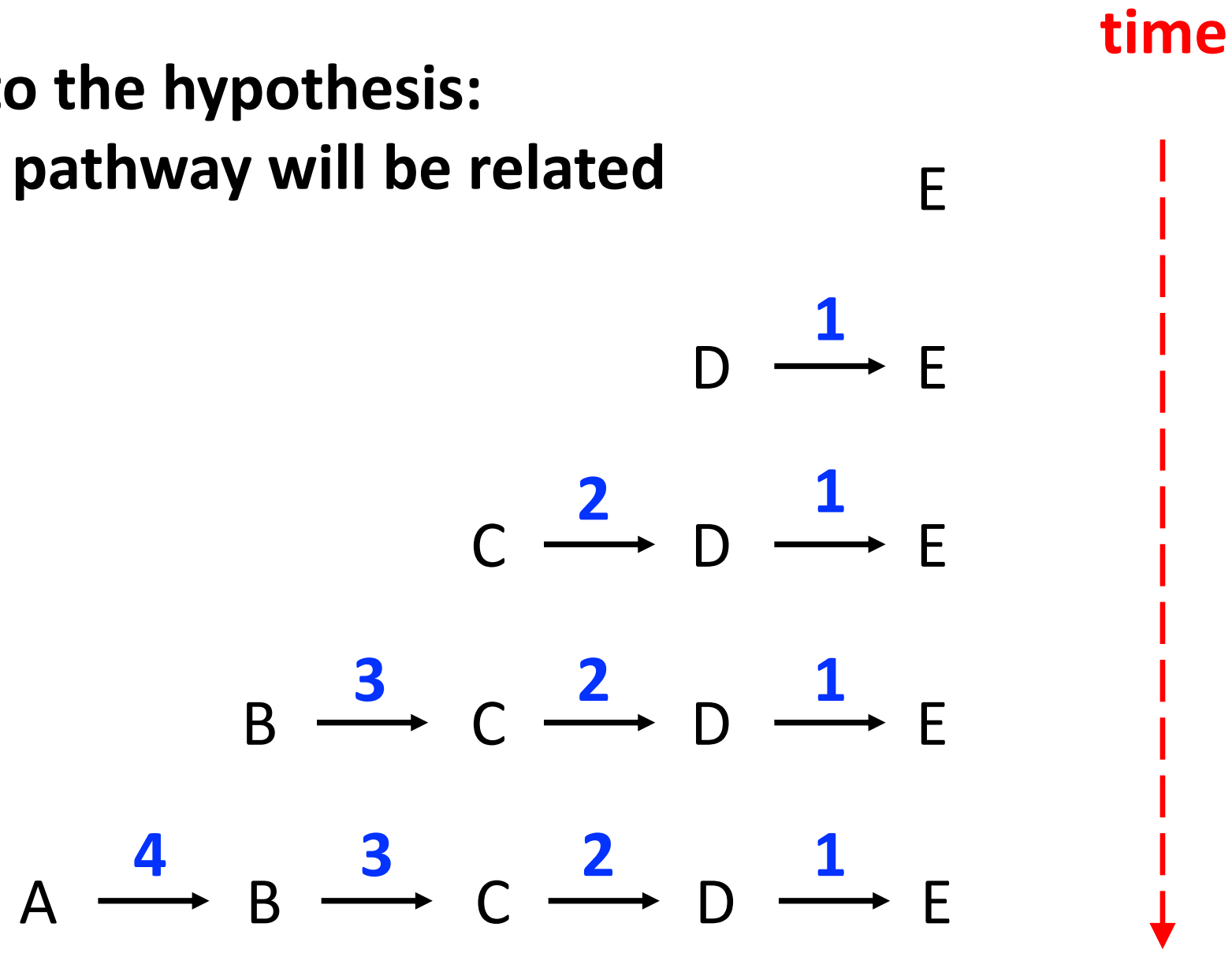
# The Evolution of Biochemical Syntheses — Retrospect and Prospect

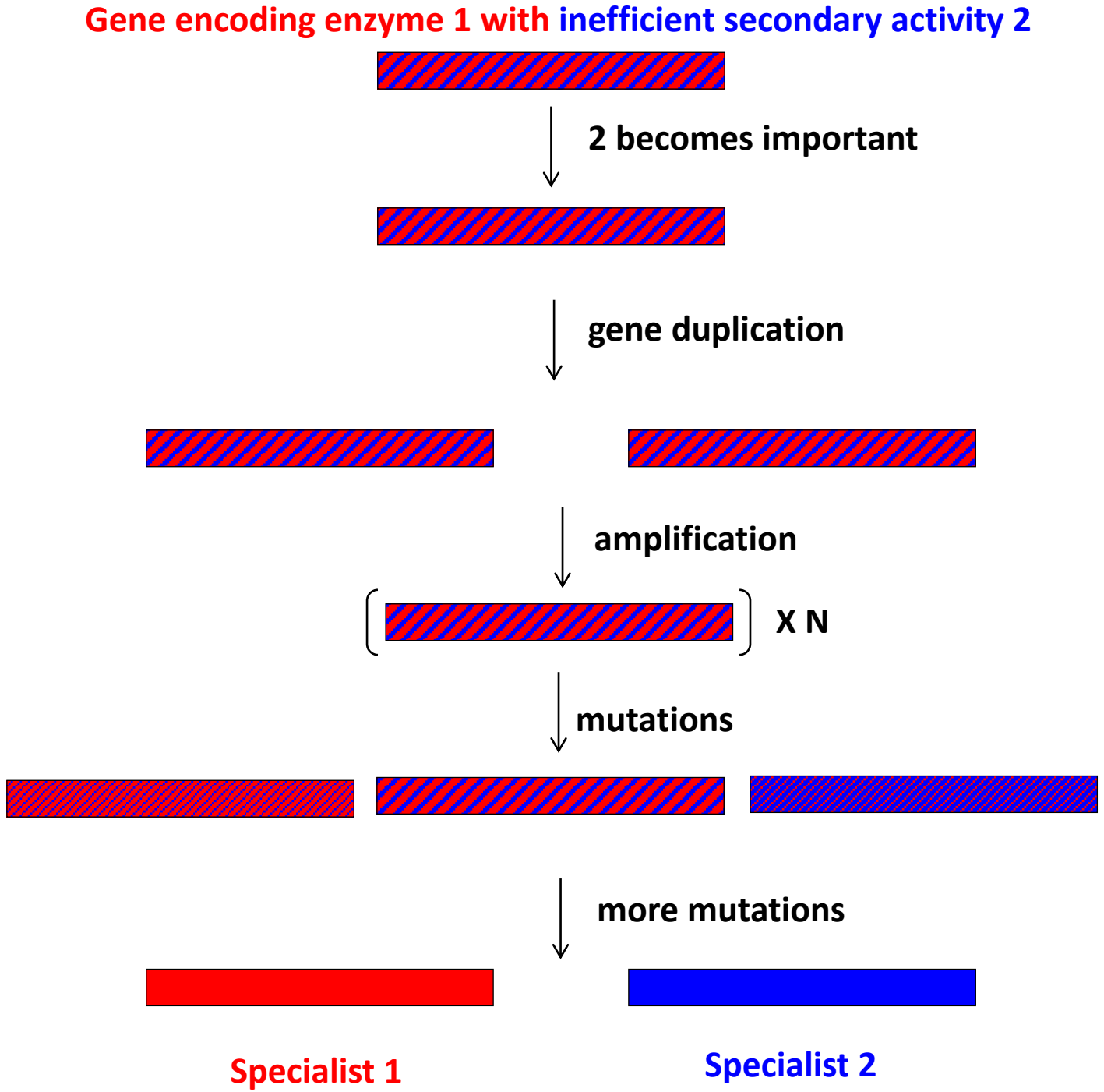
N. H. Horowitz

*Biology Division, California Institute of Technology, Pasadena, California*

**An addition to the hypothesis:**

**Enzymes in a pathway will be related**




Gene encoding enzyme 1 with inefficient secondary activity 2
2 becomes important
gene duplication
amplification
X N
mutations
more mutations
Specialist 1
Specialist 2

1974

*J. theor. Biol.* (1974) **44,** 145–160

## On Earlier States of the Biochemical System

MARTYNAS YČAS

*Department of Microbiology, Upstate Medical Center, State University of New York, Syracuse, New York* 13210, *U.S.A.*

(*Received* 18 *May* 1973, *and in revised form* 14 *August* 1973)

1976

*Ann. Rev. Microbiol. 1976. 30:409-25*
*Copyright © 1976 by Annual Reviews Inc. All rights reserved*

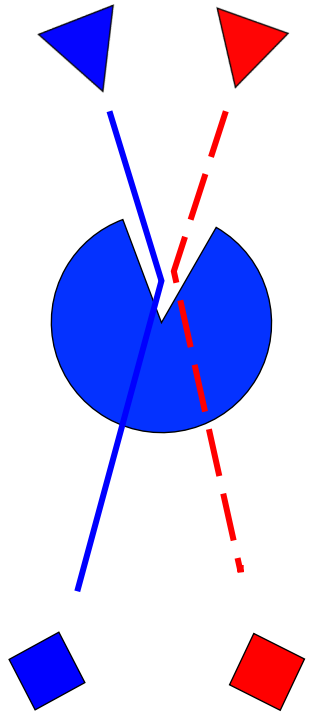
## ENZYME RECRUITMENT IN EVOLUTION OF NEW FUNCTION

*Roy A. Jensen*

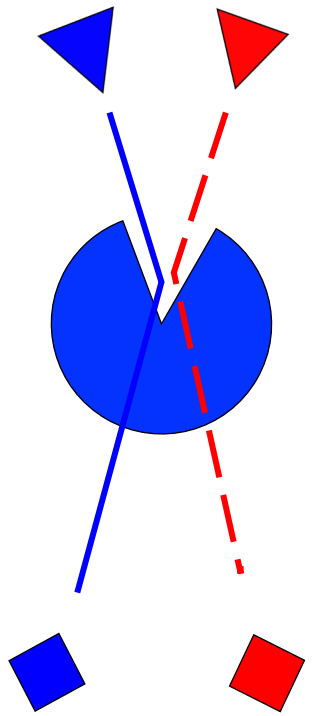
Department Biological Sciences, State University of New York at Binghamton, Binghamton, New York 13901

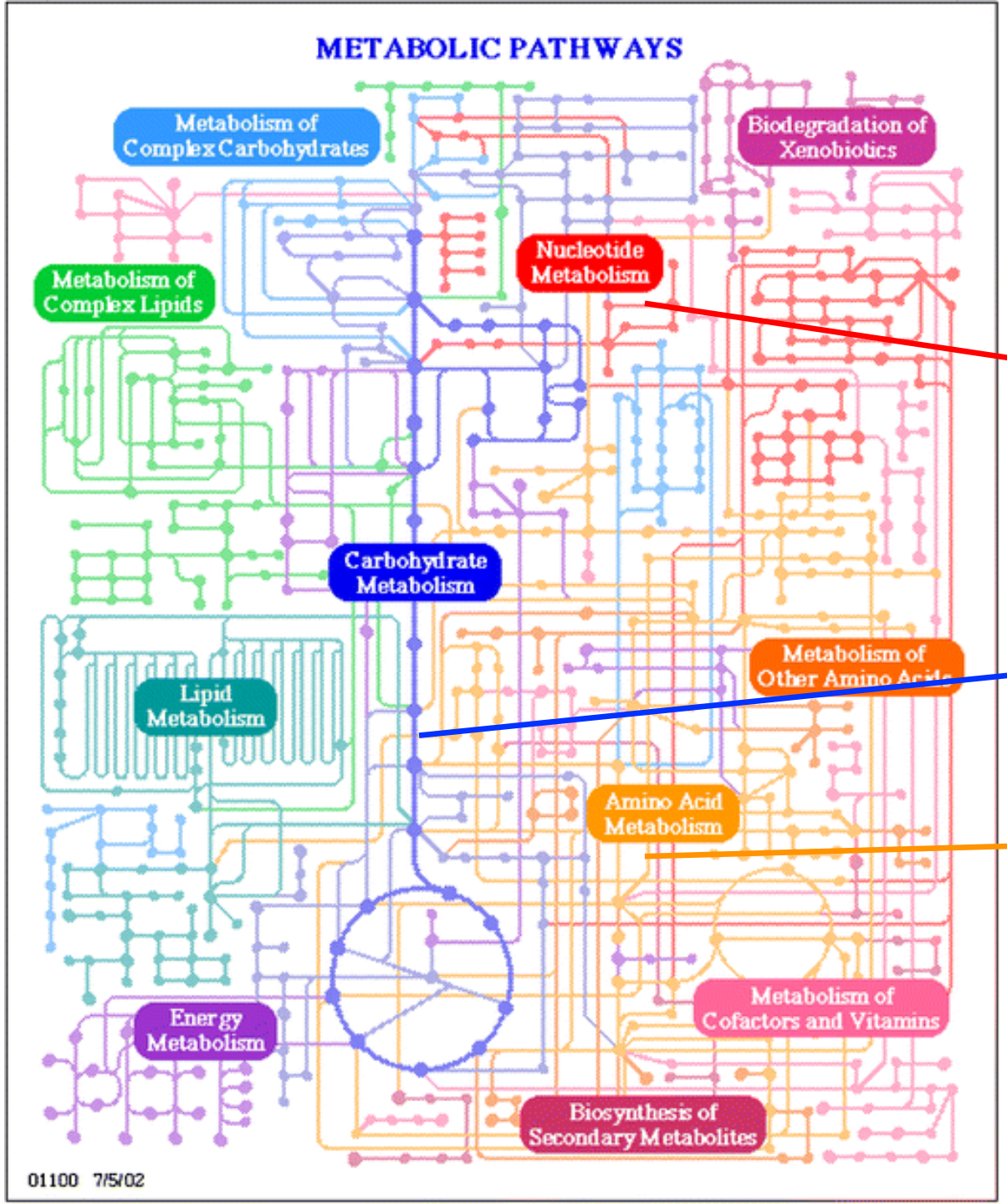
# Key idea: early enzymes were promiscuous

## i.e. able to catalyze secondary reactions



# “blue” enzyme could be recruited to perform ”red” reaction in another pathway






# The patchwork hypothesis

A

Enz1

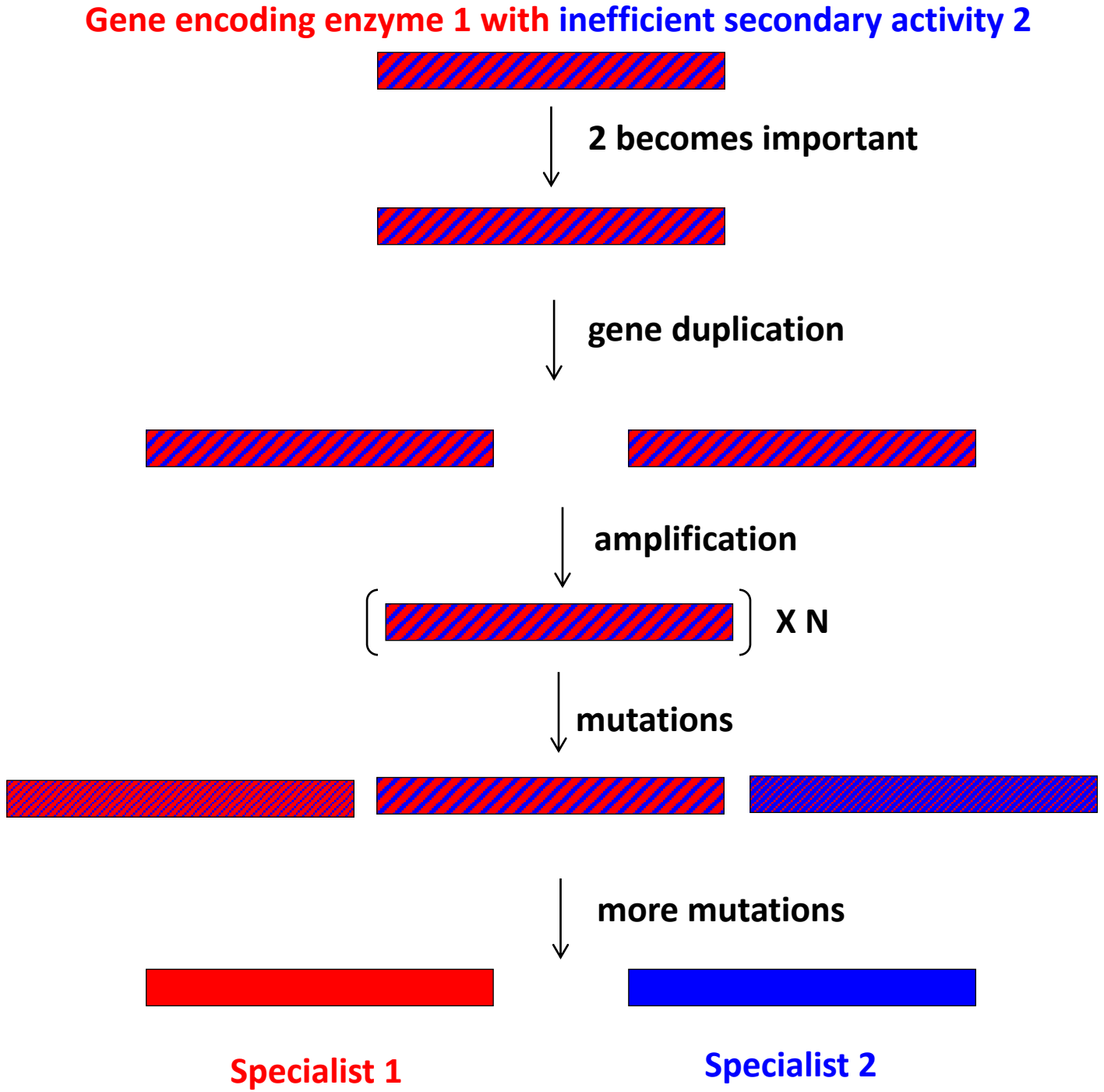
B

Enz2

C

Enz3

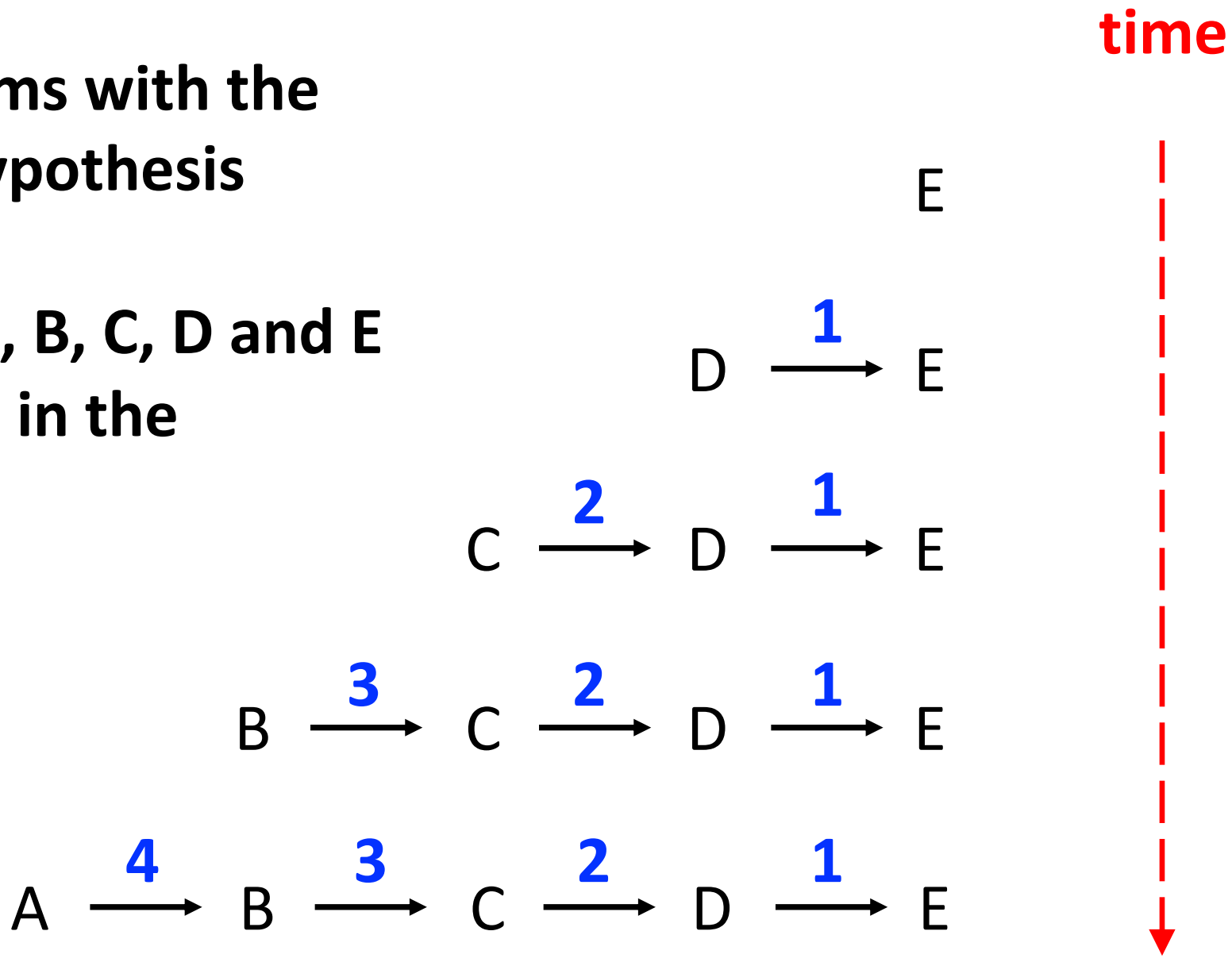
D


Gene encoding enzyme 1 with inefficient secondary activity 2
2 becomes important
gene duplication
amplification
X N
mutations
more mutations
Specialist 1
Specialist 2

Questions?

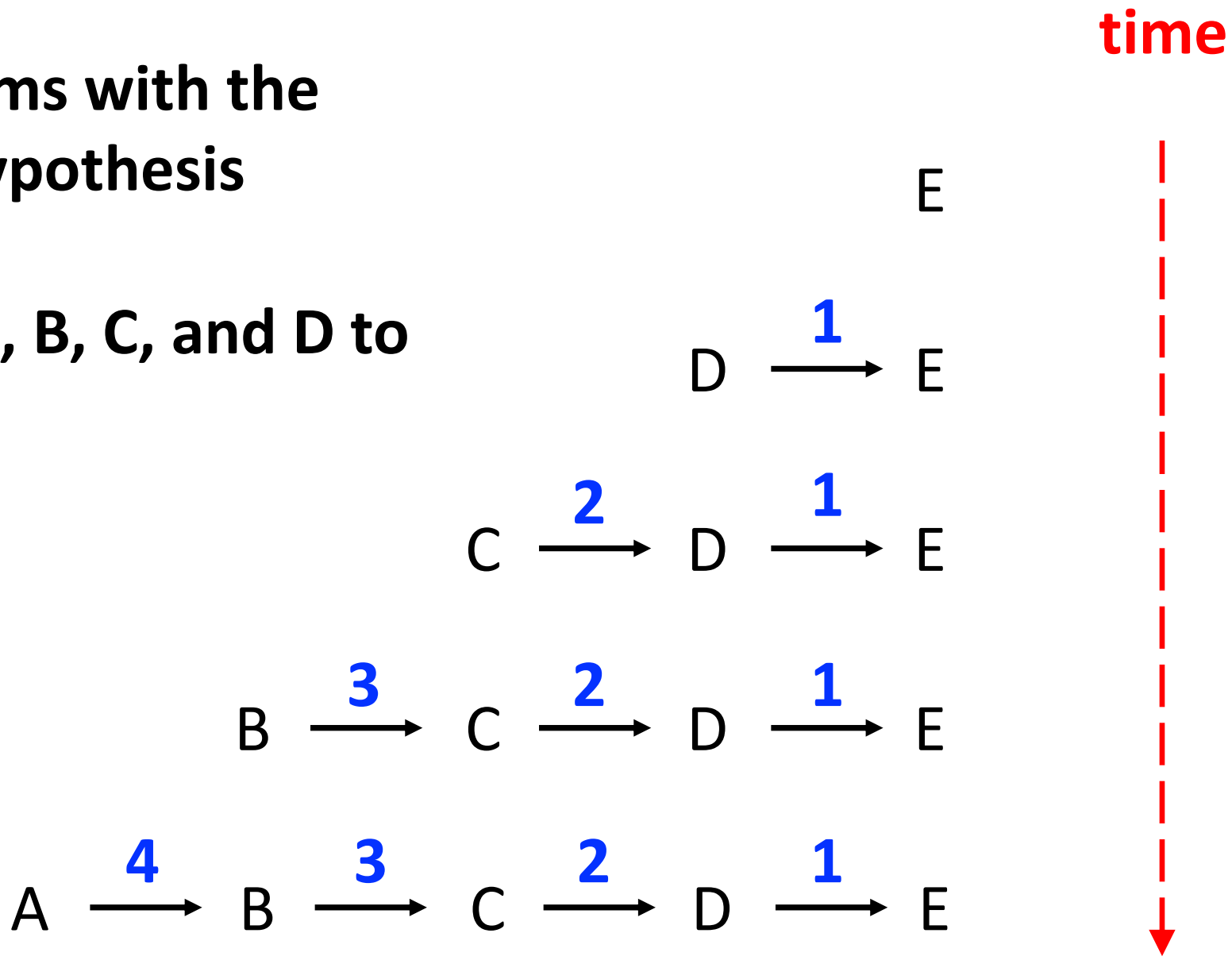
# Three problems with the retrograde hypothesis

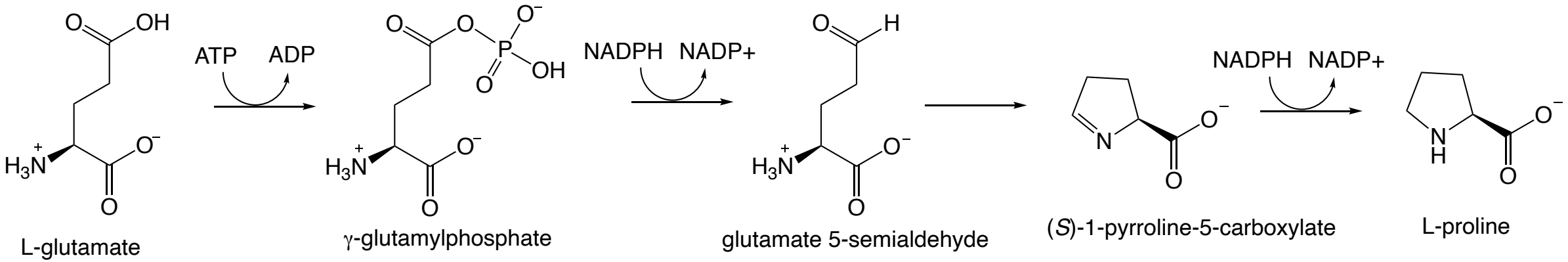
**1) Requires A, B, C, D and E to be present in the environment**



# Three problems with the retrograde hypothesis

## 2) Requires A, B, C, and D to be stable





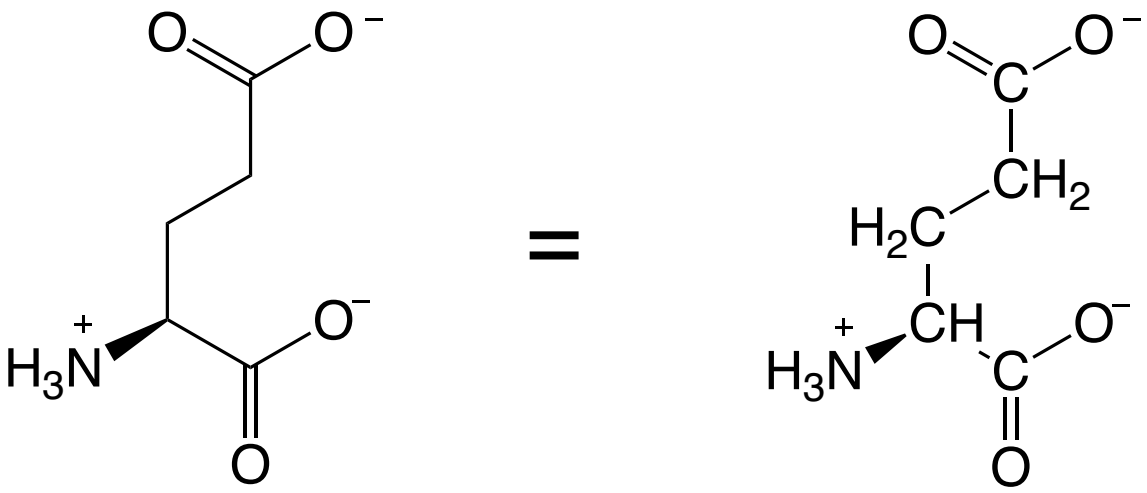

L-glutamate

γ-glutamylphosphate

glutamate 5-semialdehyde

(*S*)-1-pyrroline-5-carboxylate

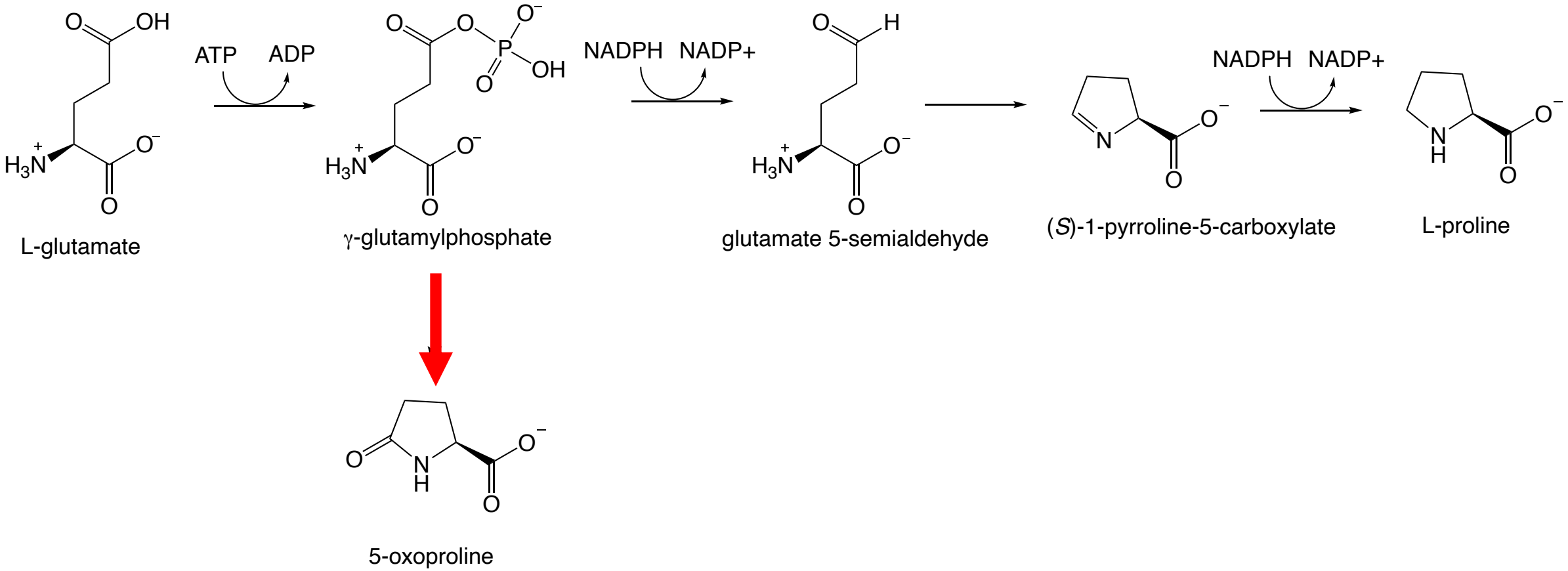
L-proline



## Conventions:

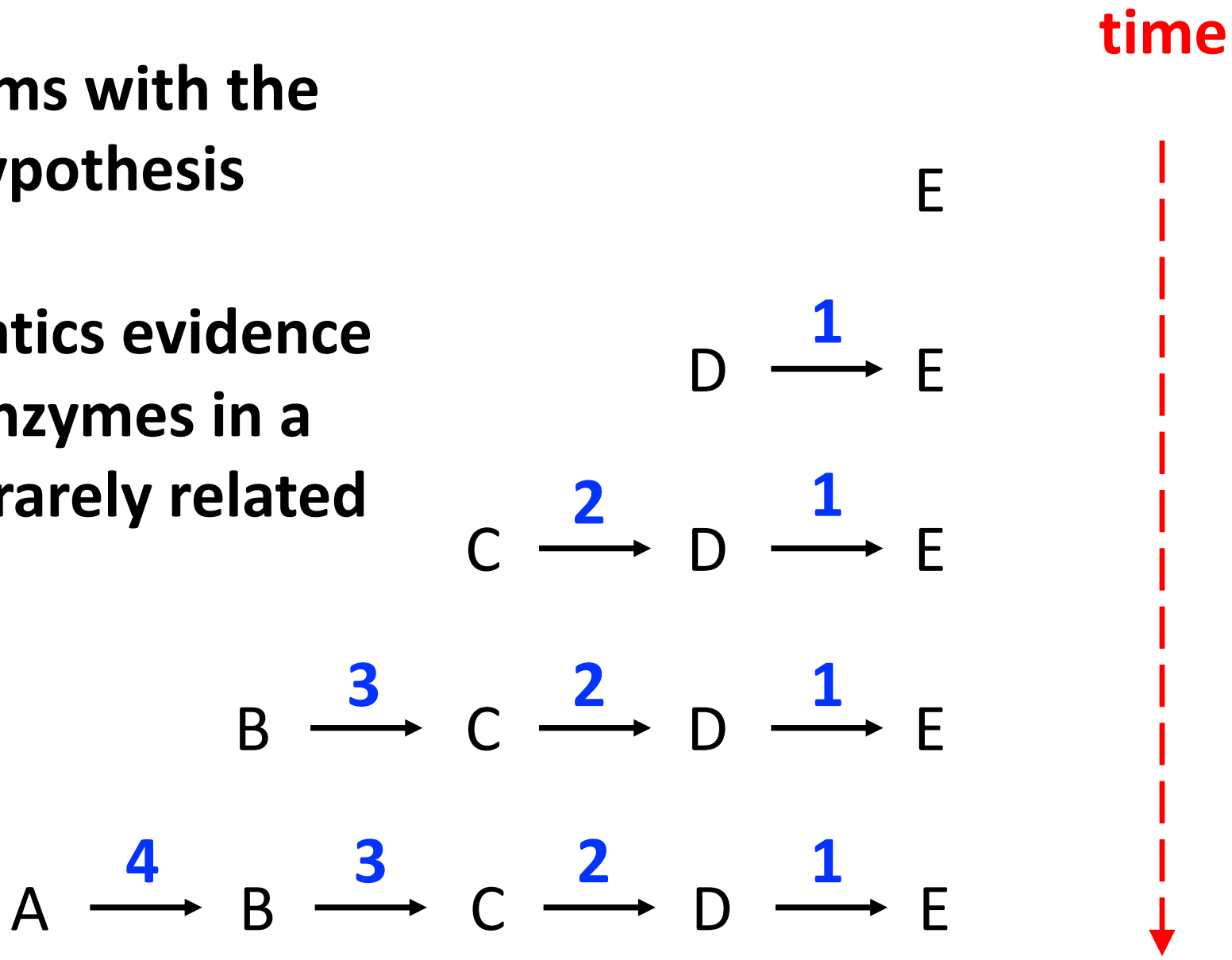
1) Unlabeled vertex = C

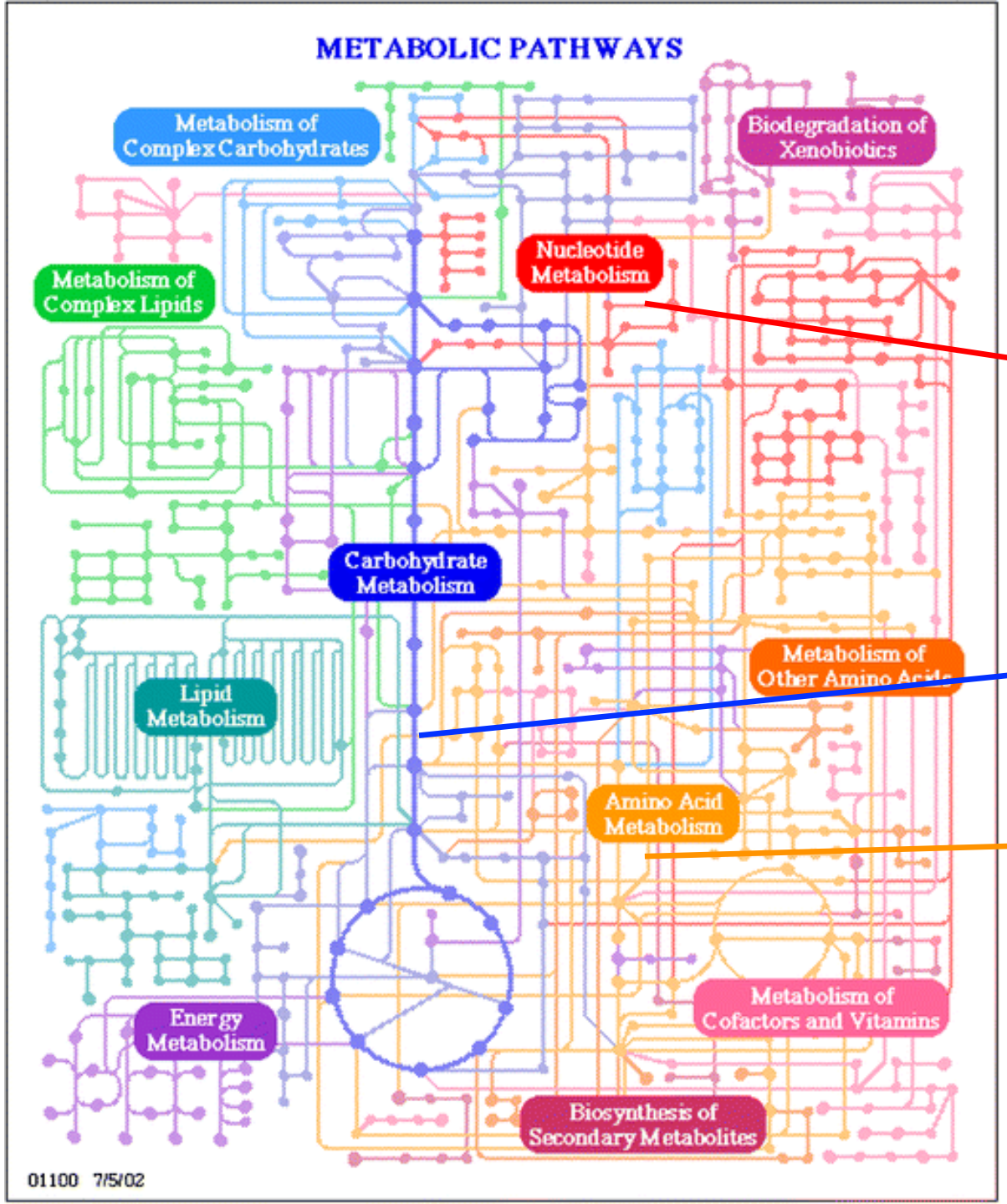
2) H atoms attached to C are not shown


ATP
ADP
NADPH
NADP+
NADPH
NADP+
L-glutamate
γ-glutamylphosphate
glutamate 5-semialdehyde
(S)-1-pyrroline-5-carboxylate
L-proline
5-oxoproline

**Three problems with the retrograde hypothesis**

**3) Bioinformatics evidence shows that enzymes in a pathway are rarely related**






# The patchwork hypothesis

A

Enz1

B

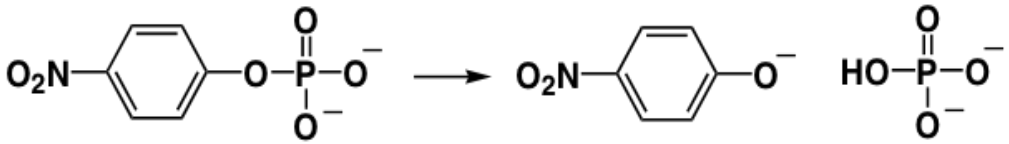
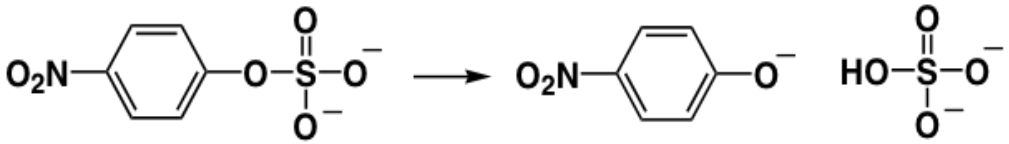
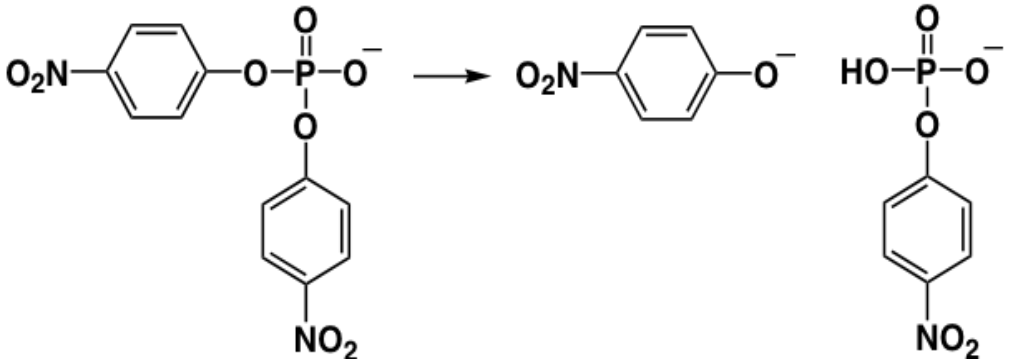
Enz2

C

Enz3

D

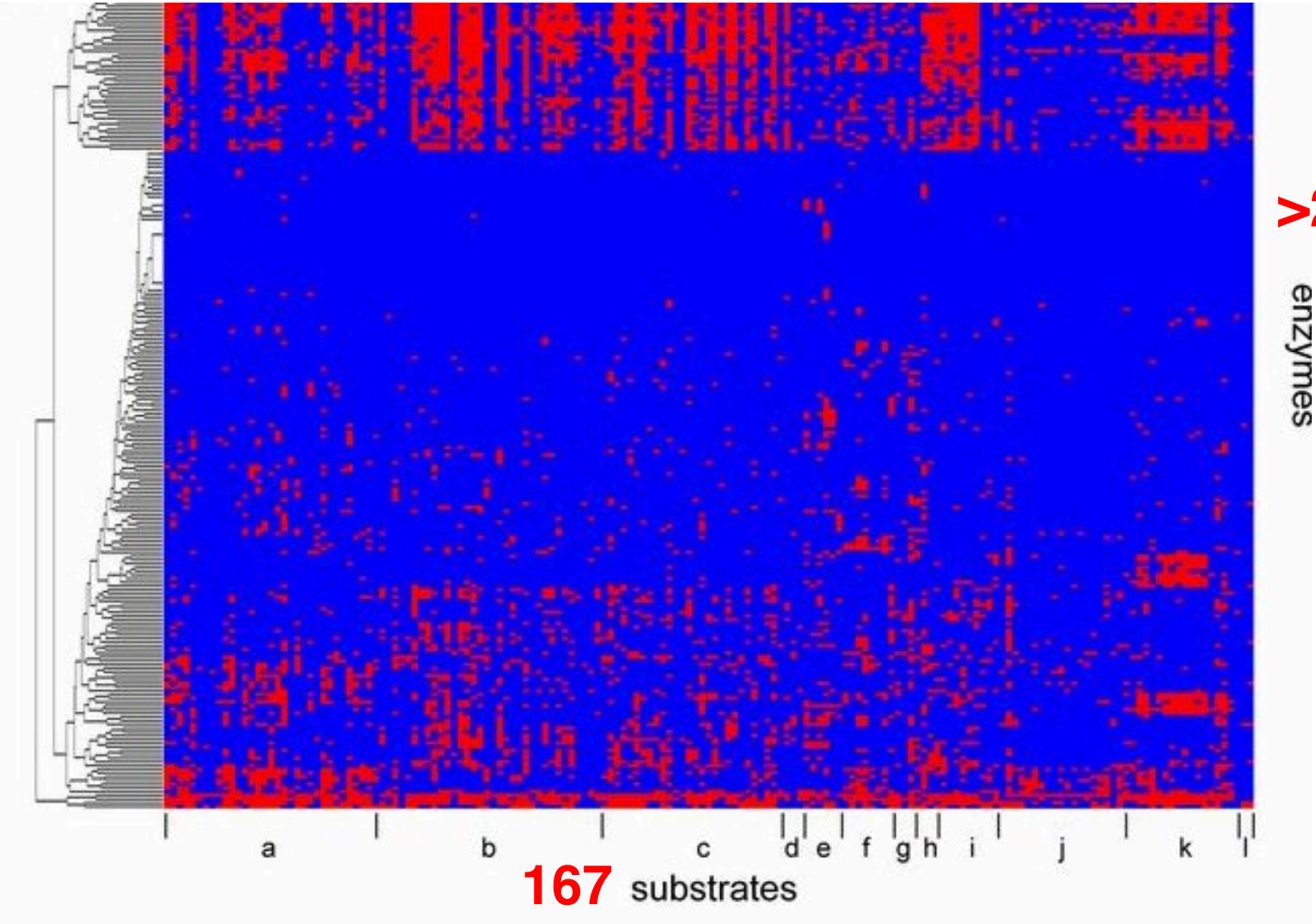
# Even inefficient promiscuous activities can accelerate reactions by orders of magnitude

| e.g. *E. coli* alkaline phosphatase | $k_{cat} / K_M$ ($M^{-1}s^{-1}$) | "proficiency" |
|---|---|---|
|  | $3 \times 10^7$ | $10^{17}$ |
|  | $1 \times 10^{-2}$ | $10^9$ |
|  | $1 \times 10^2$ | $10^{11}$ |

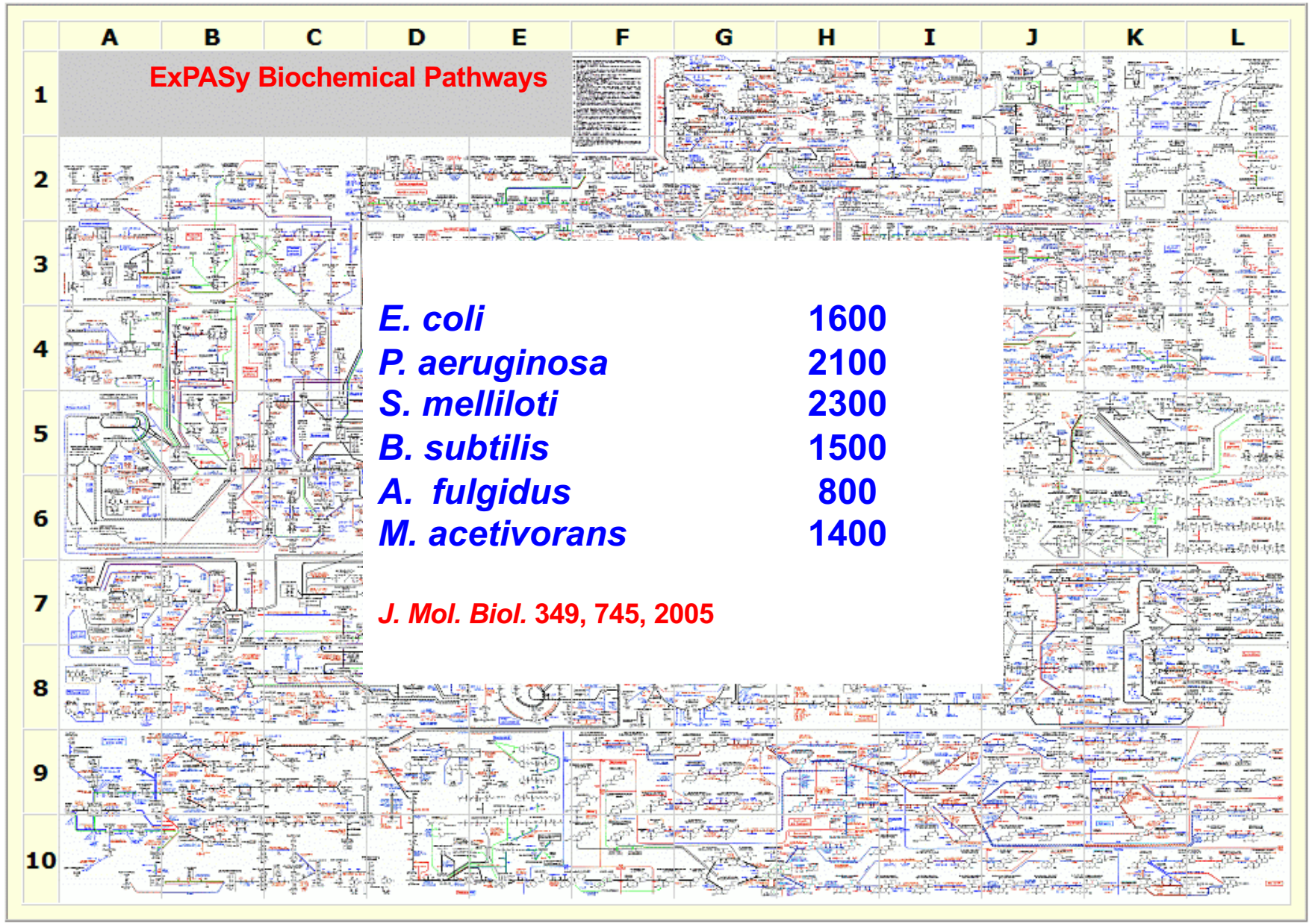
O'Brien and Herschlag, *J. Am. Chem. Soc.* 1998, 120-12369 and *Biochemistry* 2001, 40, 5691

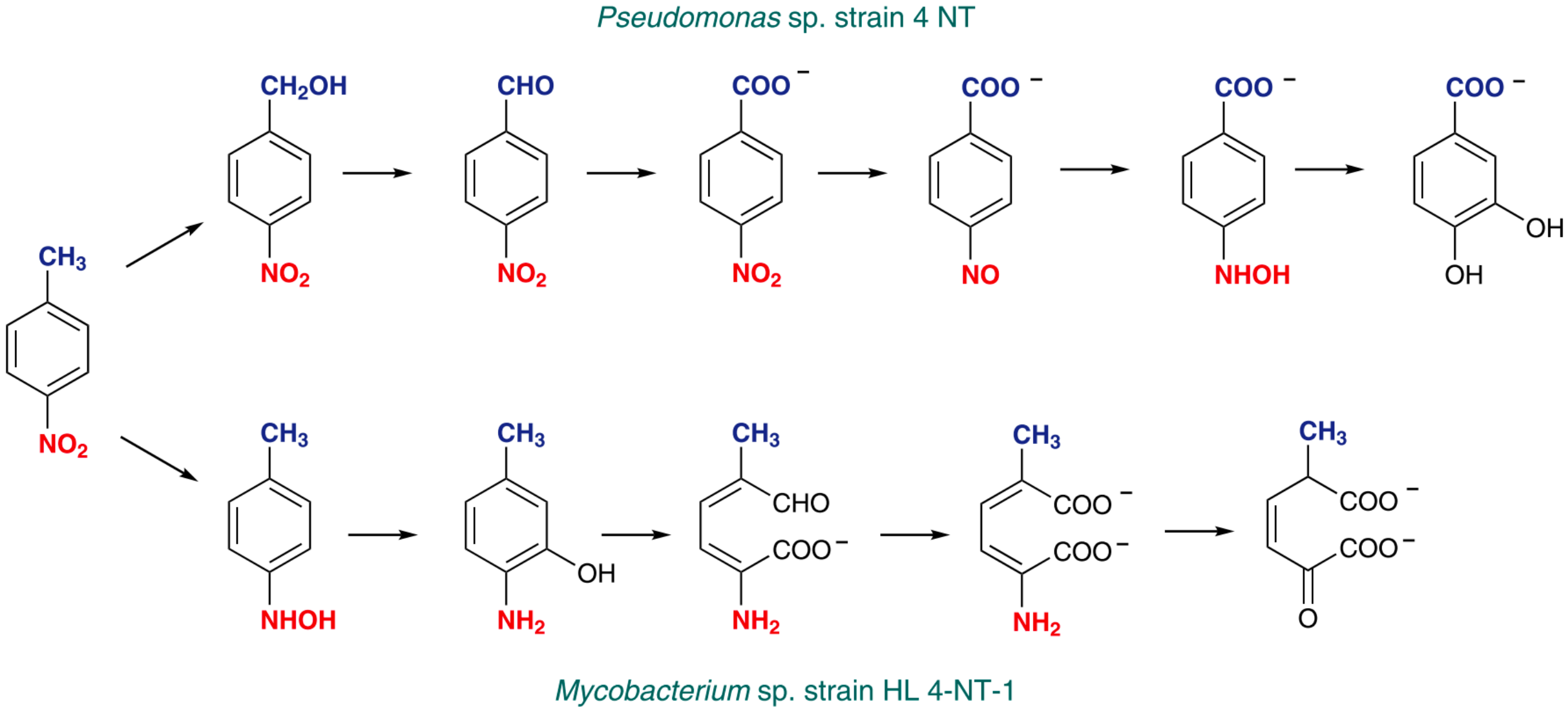
# Promiscuity is common

Panoramic view of a superfamily of phosphatases through substrate profiling,
*PNAS* **112**, E1974 – E1983, 2015



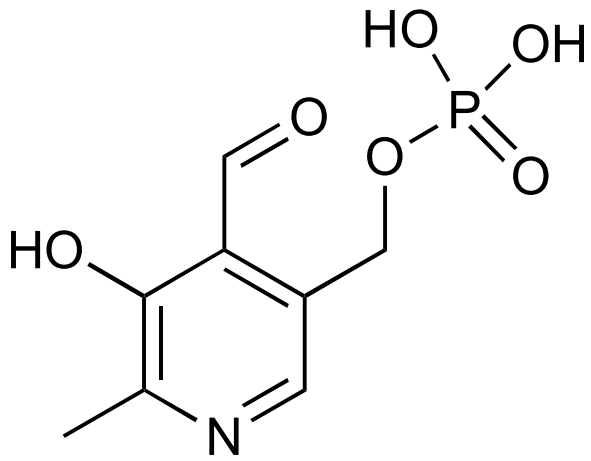
# The number of promiscuous activities is unknown but undoubtedly huge






*Nature Chem Biol* **5**, 559-566, 2009

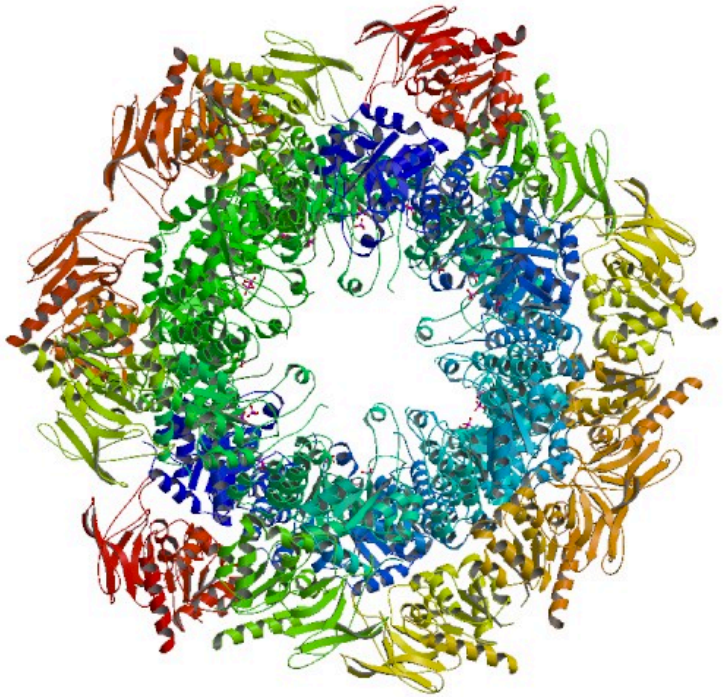
# Pyridoxal 5’-phosphate (vitamin B6)
# (PLP)



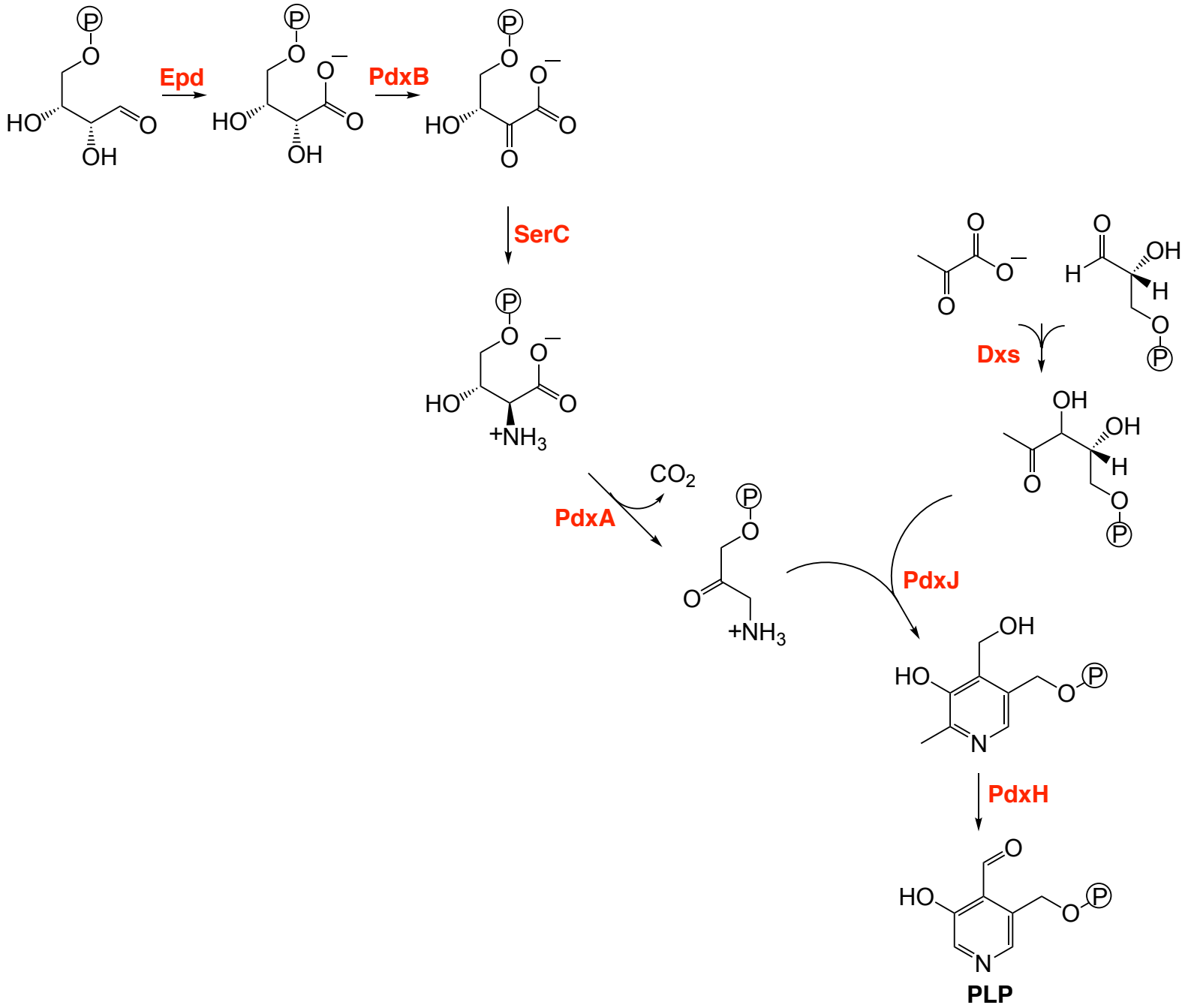

**Transamination**
**Racemization**
**β-elimination**
**Retro aldol cleavage**
**Radical reactions**

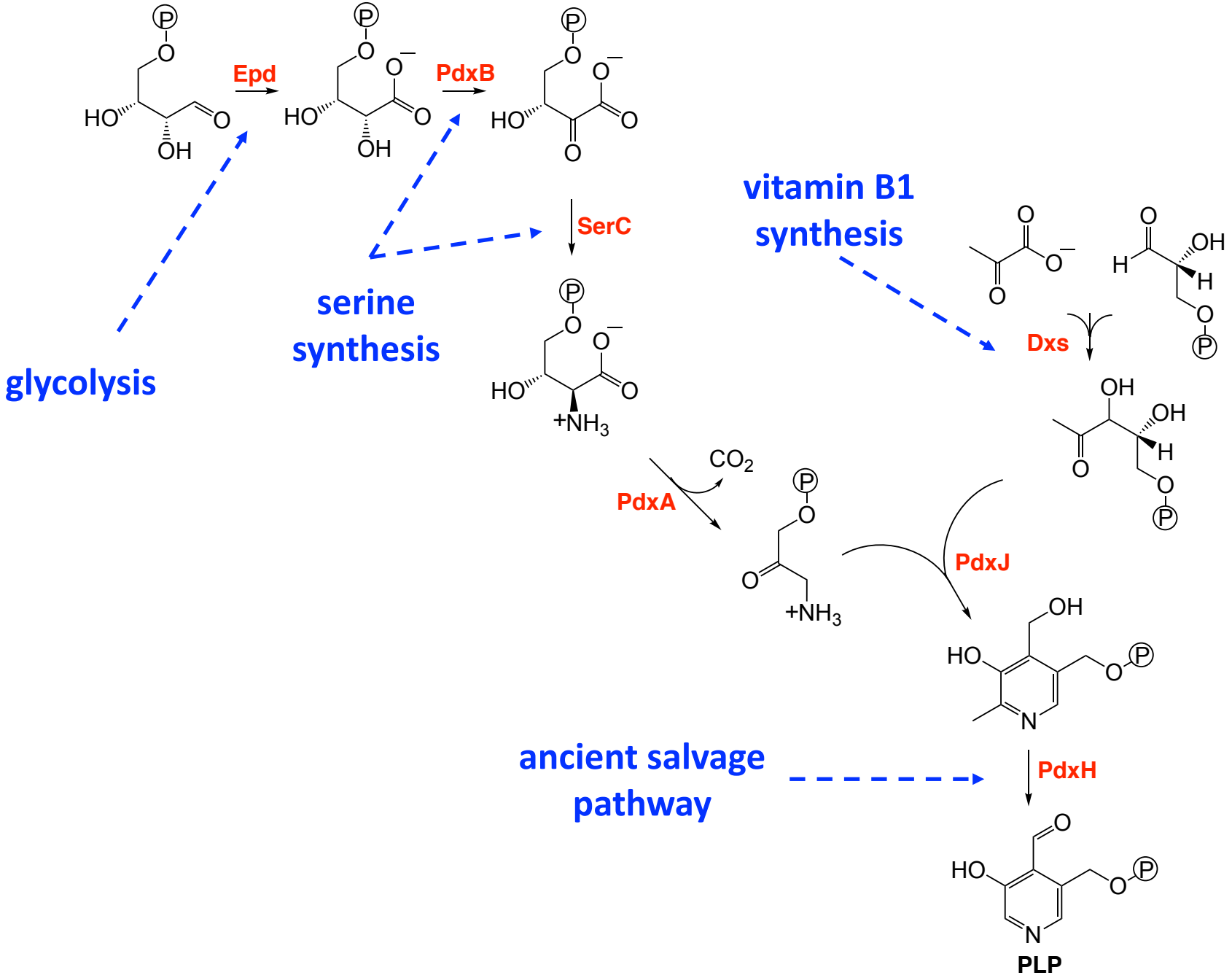
# PLP synthesis in most organisms

glutamine
glyceraldehyde 3-phosphate
ribose 5-phosphate

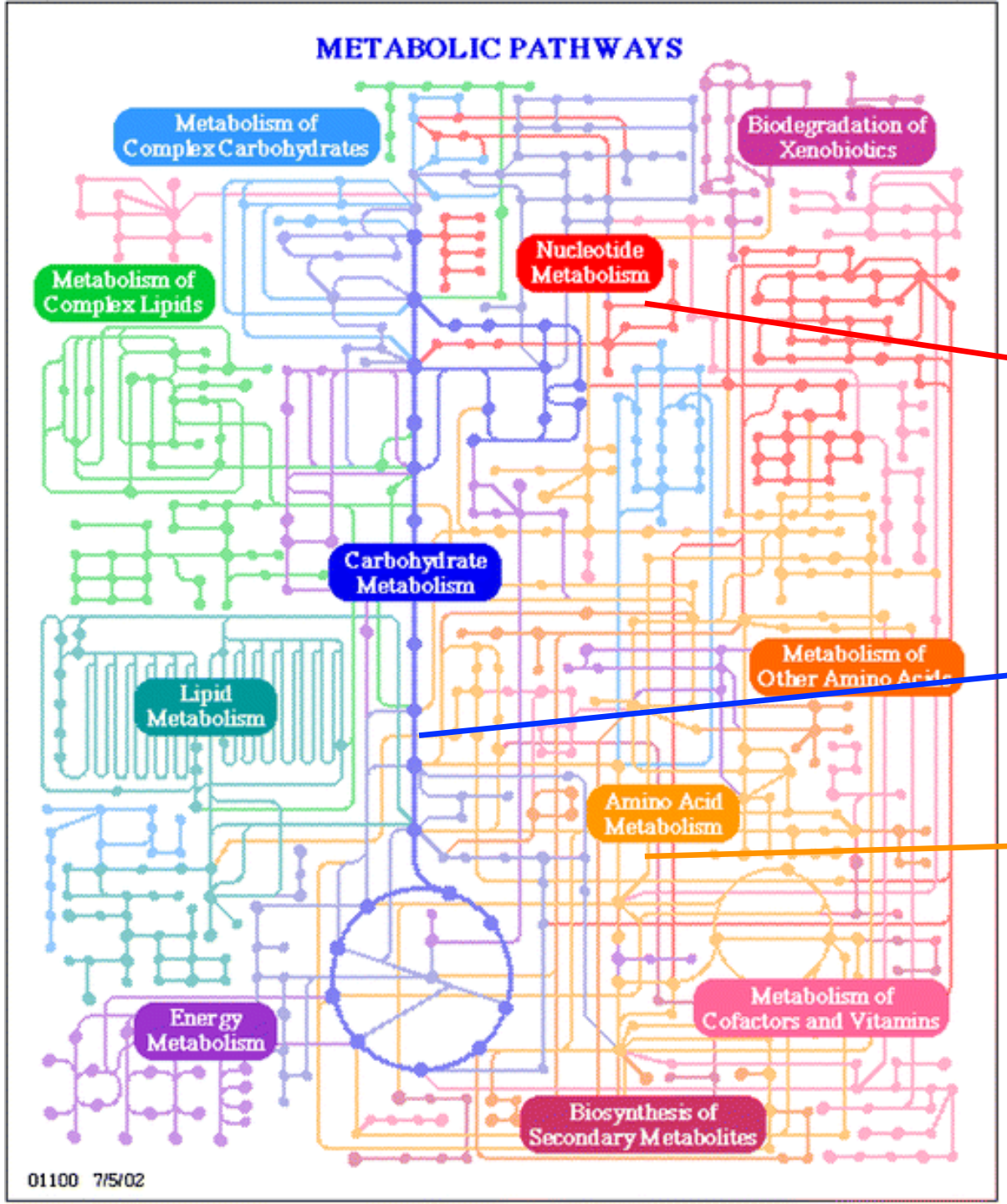


**PLP**

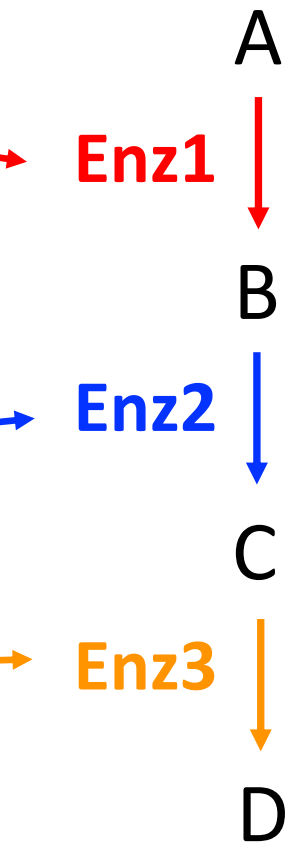

Epd
PdxB
SerC
Dxs
PdxA
$CO_2$
PdxJ
PdxH
PLP


Epd
PdxB
SerC
glycolysis
serine synthesis
vitamin B1 synthesis
Dxs
PdxA
PdxJ
PdxH
ancient salvage pathway
PLP

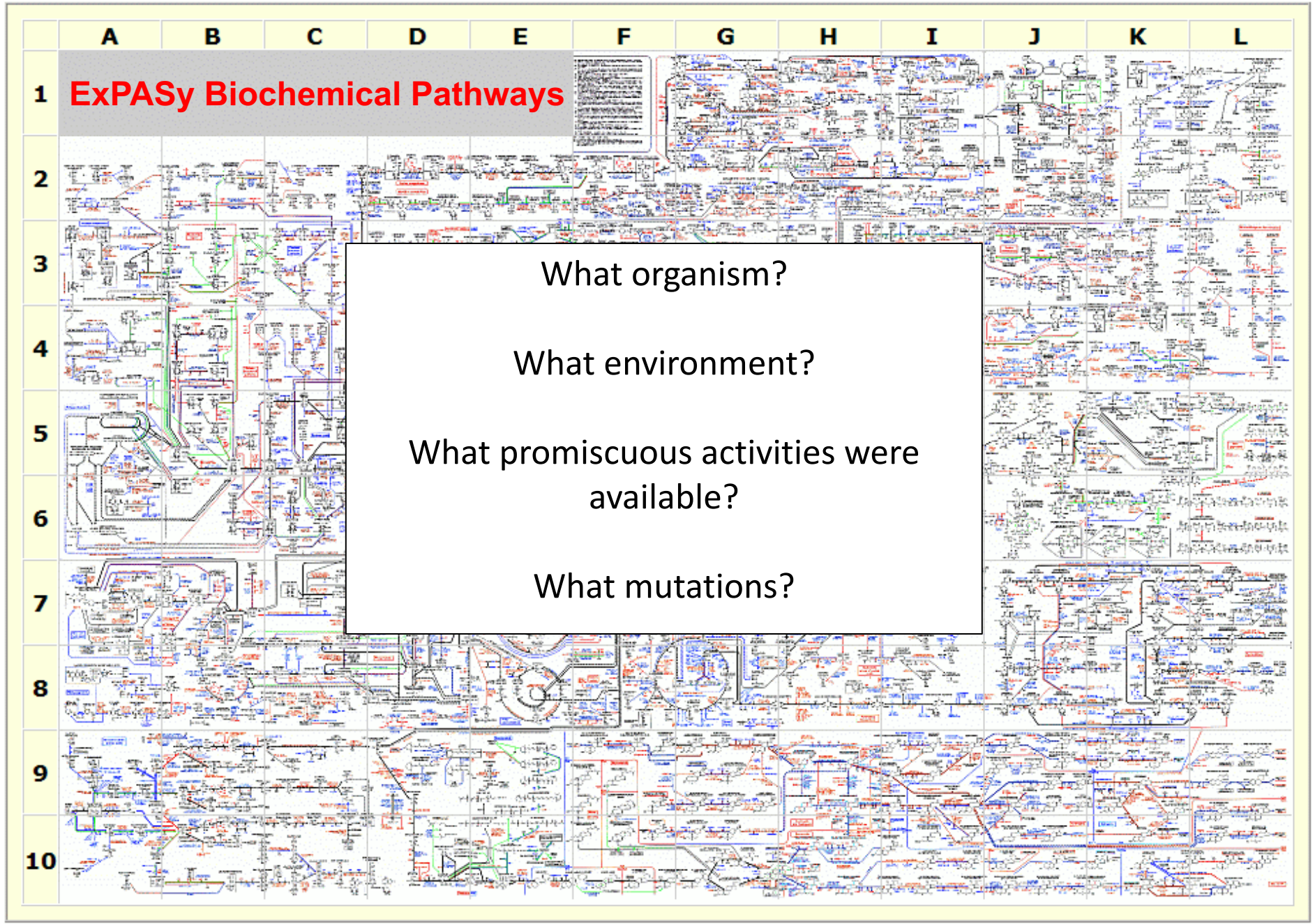
Questions?




**The patchwork hypothesis is supported by the evidence – but how does it happen?**



# Lost in time…….


ExPASy Biochemical Pathways
What organism?
What environment?
What promiscuous activities were available?
What mutations?

# A story about the evolution of a novel metabolic pathway

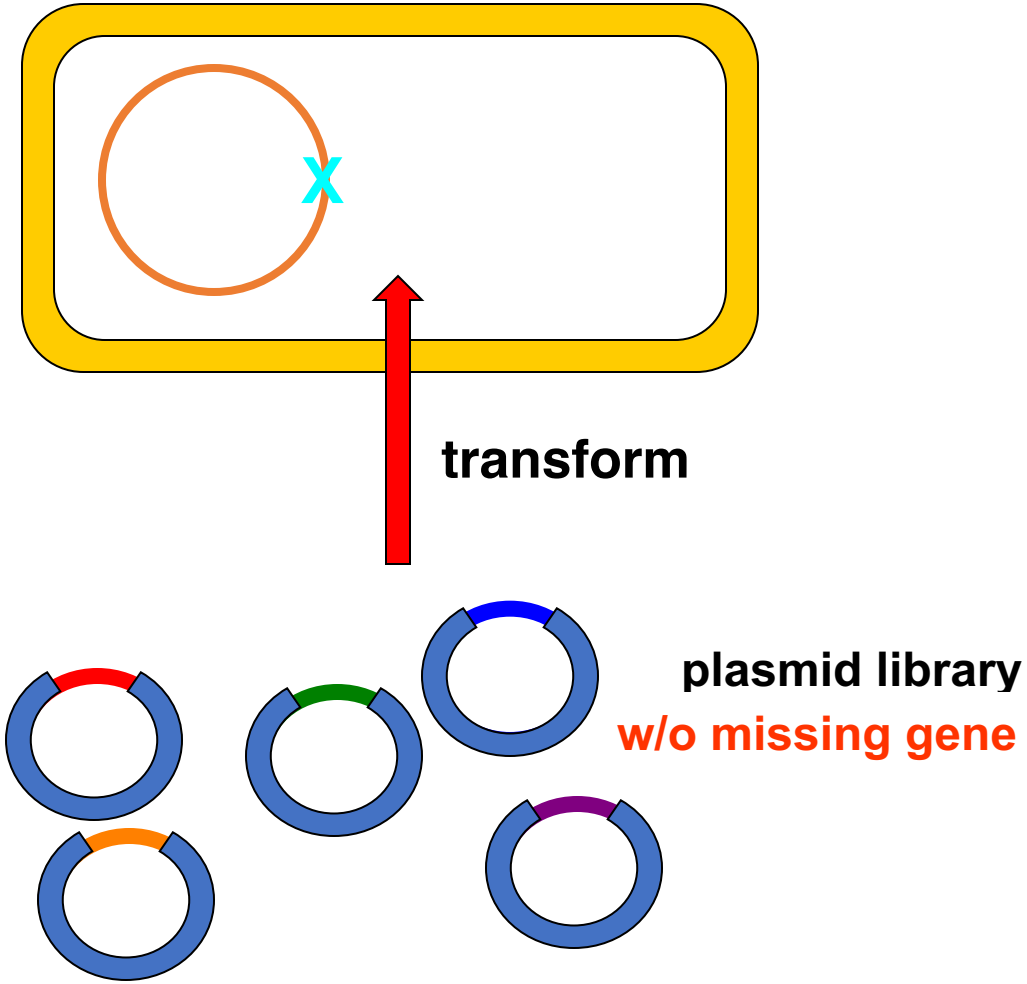
# Hidden resources in the *Escherichia coli* genome restore PLP synthesis and robust growth after deletion of the essential gene *pdxB*

**Juhan Kim[a,b,1], Jake J. Flood[a,b,1], Michael R. Kristofich[a,b], Cyrus Gidfar[a,b], Andrew B. Morgenthaler[a,b], Tobias Fuhrer[c], Uwe Sauer[c], Daniel Snyder[d], Vaughn S. Cooper[d], Christopher C. Ebmeier[a], William M. Old[a], and Shelley D. Copley[a,b,2]**

[a]Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, Boulder, CO 80309; [b]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO 80309; [c]Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland; and [d]Center for Evolutionary Biology and Medicine, University of Pittsburgh, Pittsburgh, PA 15260

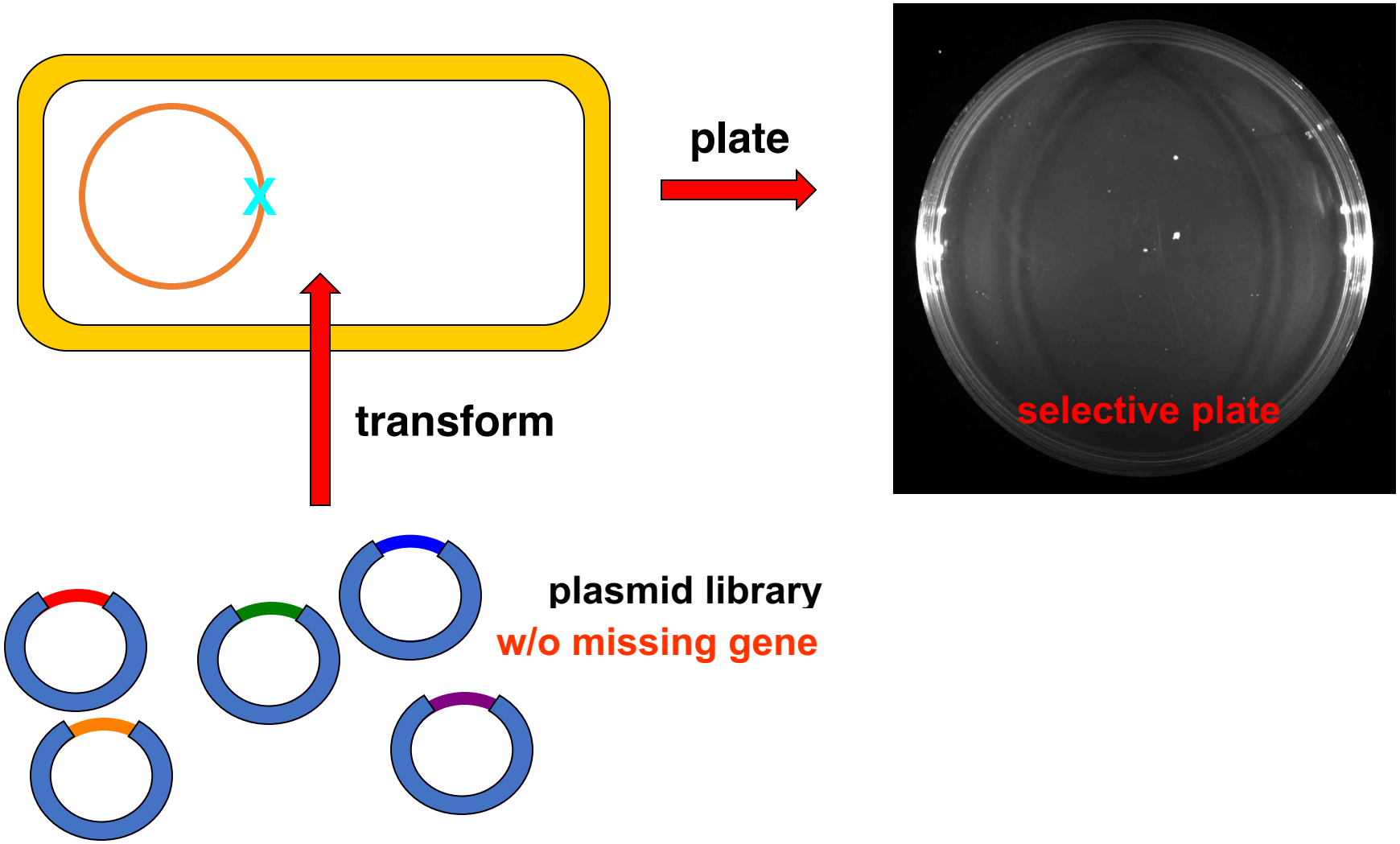
Edited by Michael Lynch, Arizona State University, Tempe, AZ, and approved October 11, 2019 (received for review September 7, 2019)

# Multicopy suppression



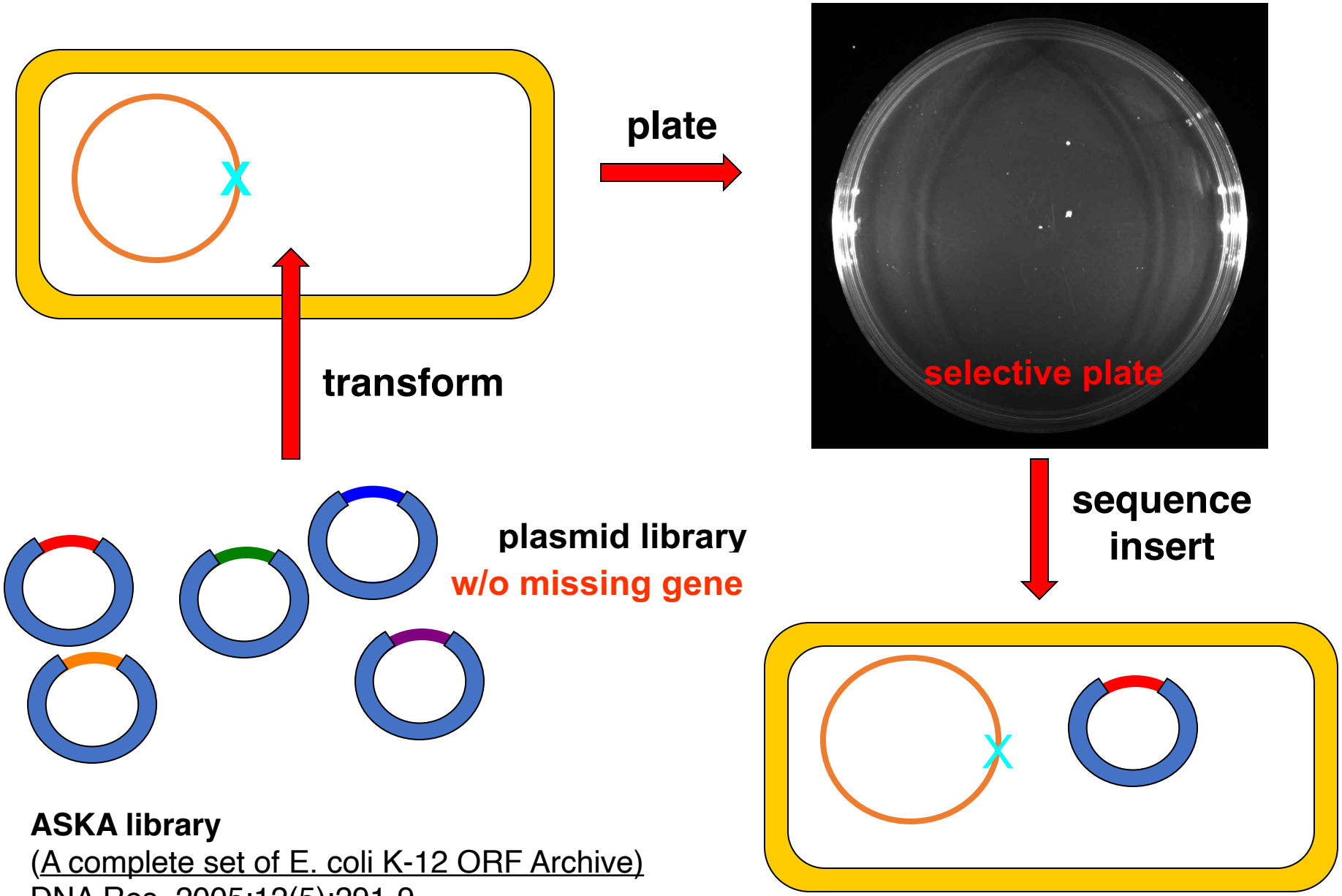

**ASKA library**
(A complete set of E. coli K-12 ORF Archive)
DNA Res. 2005;12(5):291-9

# Multicopy suppression



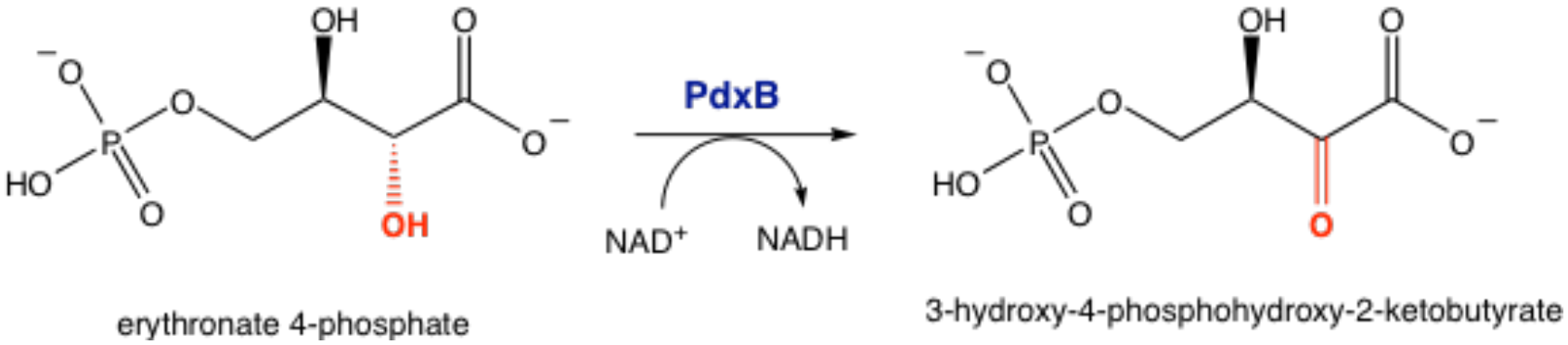

**ASKA library**
(A complete set of E. coli K-12 ORF Archive)
DNA Res. 2005;12(5):291-9

# Multicopy suppression




**ASKA library**
(A complete set of E. coli K-12 ORF Archive)
DNA Res. 2005;12(5):291-9

# PdxB: Erythronate-4-phosphate dehydrogenase

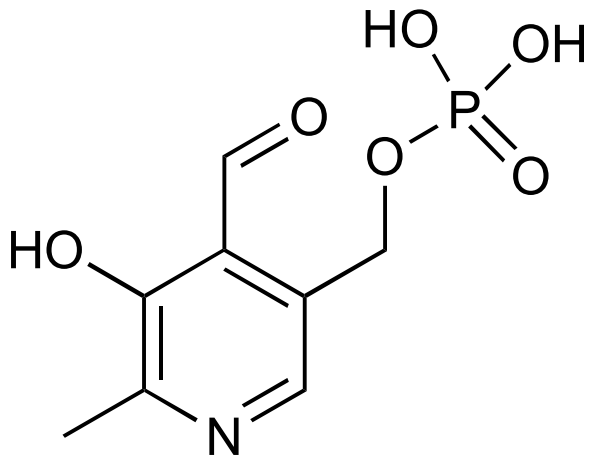



Note: bacterial genes are designated by italics and lower case

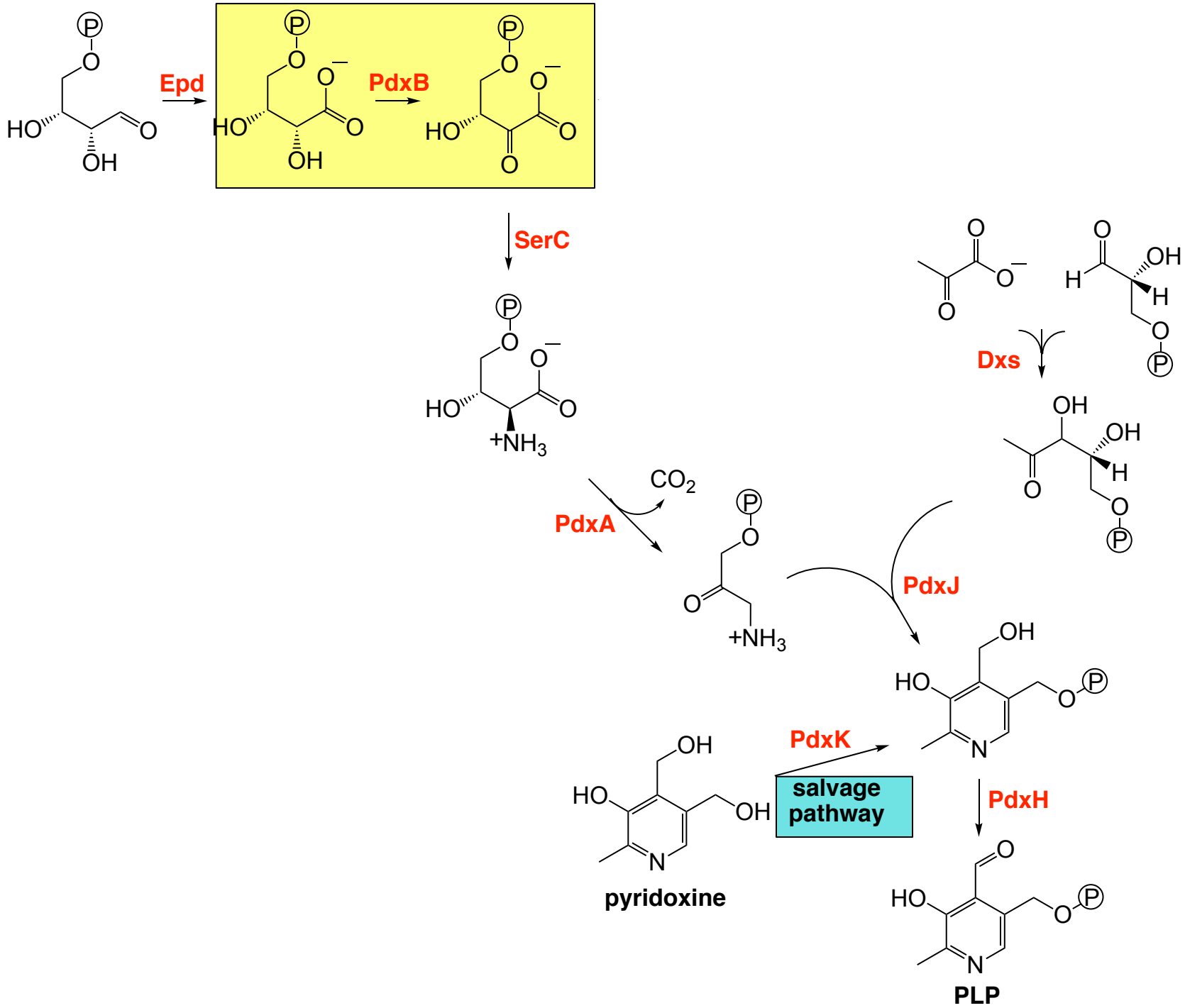
Proteins encoded by genes are capitalized and not italicized

E.g. the gene *pdxB* encodes the protein PdxB (which is erythronate 4-phosphate dehydrogenase)

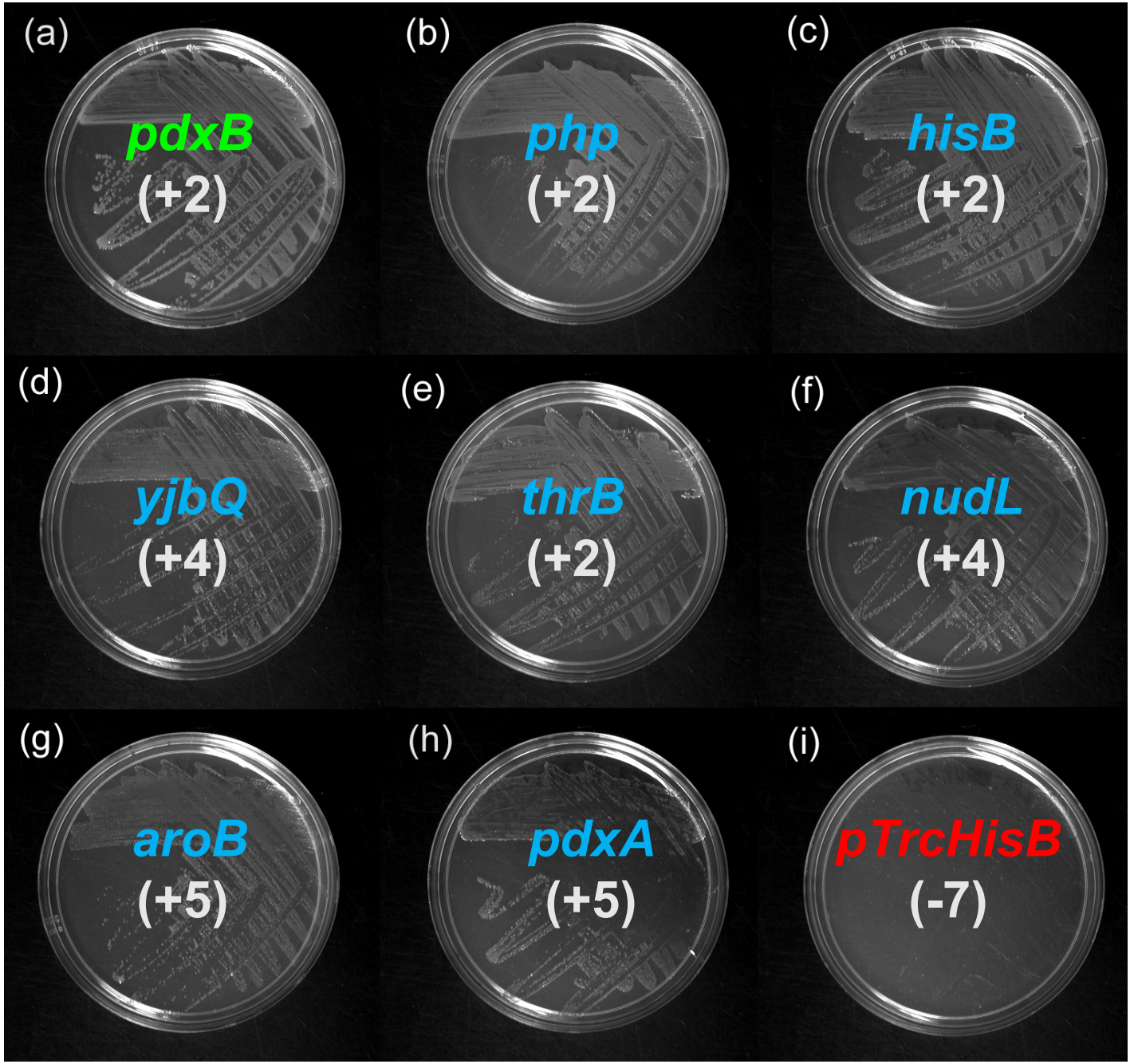
# Pyridoxal 5’-phosphate (vitamin B6)
# (PLP)




**Transamination**
**Racemization**
**β-elimination**
**Retro aldol cleavage**
**Radical reactions**


Epd
PdxB
SerC
Dxs
PdxA
PdxJ
PdxK
salvage pathway
PdxH
pyridoxine
PLP

Δ*pdxB E. coli*
M9/glucose plates




Note: bacterial genes are designated by italics and lower case

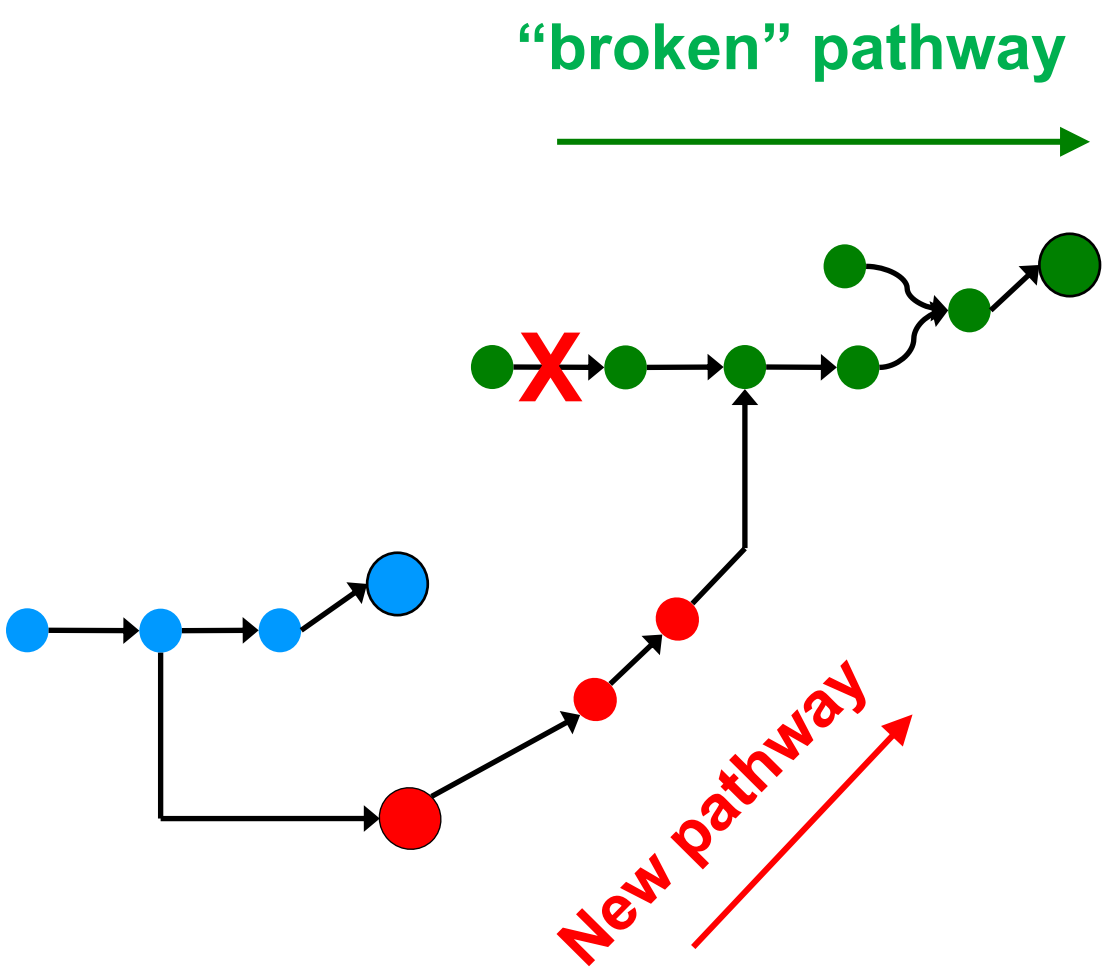
Proteins encoded by genes are capitalized and not italicized

E.g. the gene *pdxB* encodes the protein PdxB (which is erythronate 4-phosphate dehydrogenase)

| enzyme | activity |
| --- | --- |
| PdxA | dehydrogenase |
| AroB | synthase |
| ThrB | kinase |
| HisB | dehydratase |
| Php | predicted hydrolase |
| NudL | hydrolase |
| YjbQ | conserved protein of unknown function |

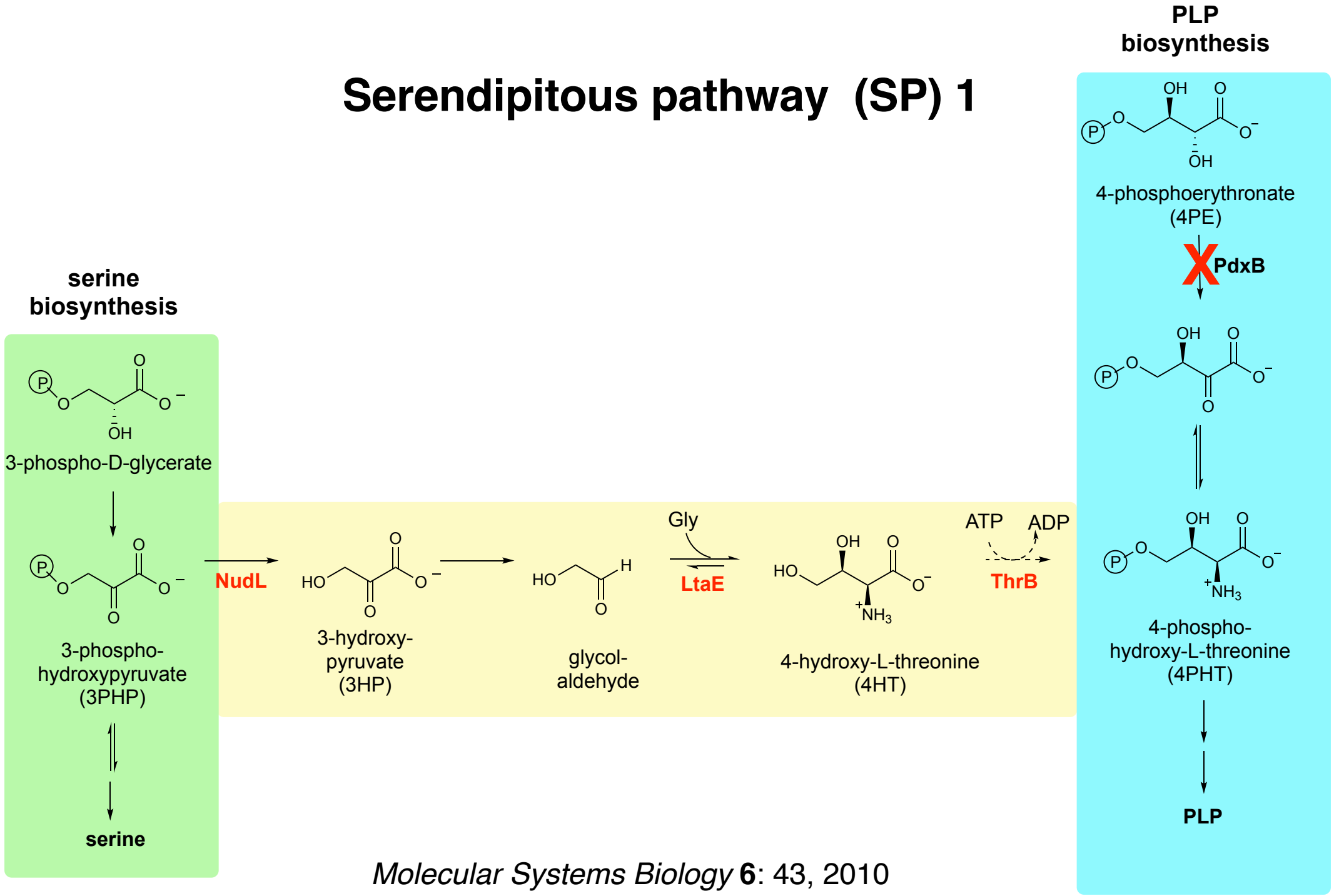
| enzyme | activity |
| --- | --- |
| PdxA | dehydrogenase |
| AroB | synthase |
| ThrB | kinase |
| HisB | dehydratase |
| Php | predicted hydrolase |
| NudL | hydrolase |
| YjbQ | conserved protein of unknown function |

# Serendipitous pathway

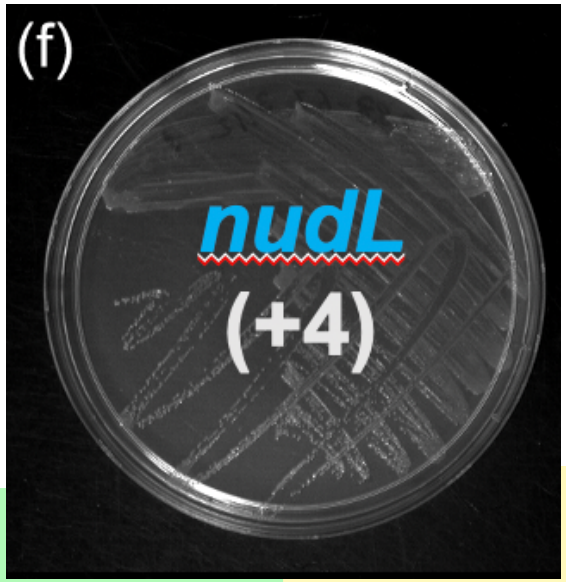



# Metabolic innovation

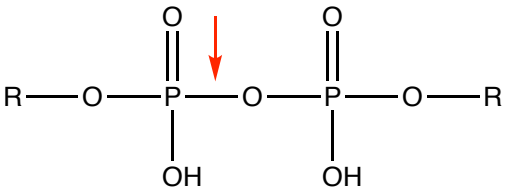
# Serendipitous pathway (SP) 1

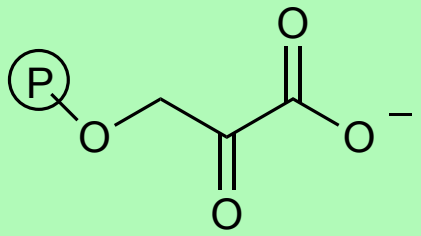



*Molecular Systems Biology* **6**: 43, 2010

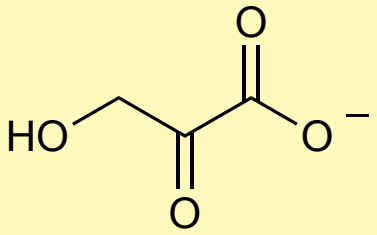



**predicted CoA pyrophosphorylase**



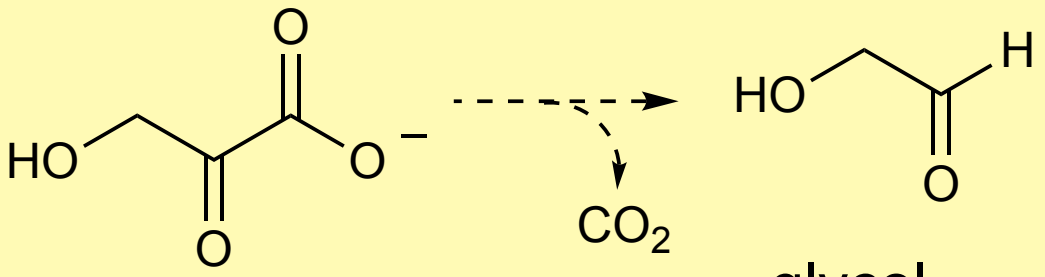




3-phospho-hydroxypyruvate




3-hydroxy-pyruvate

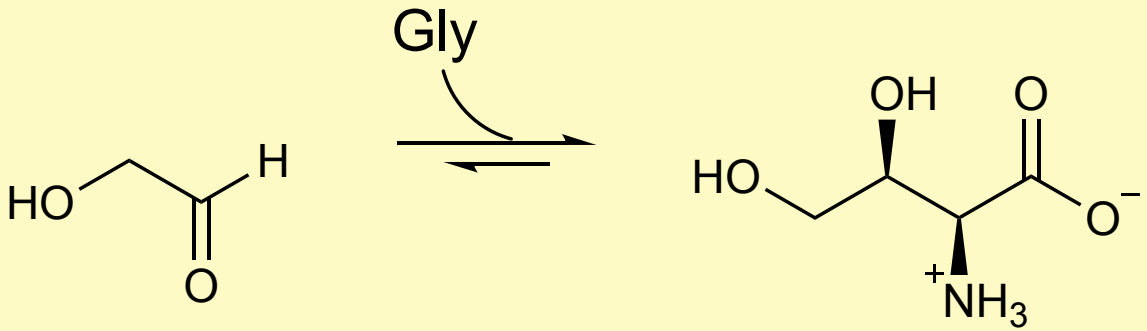
**non-enzymatic**




3-hydroxy-pyruvate

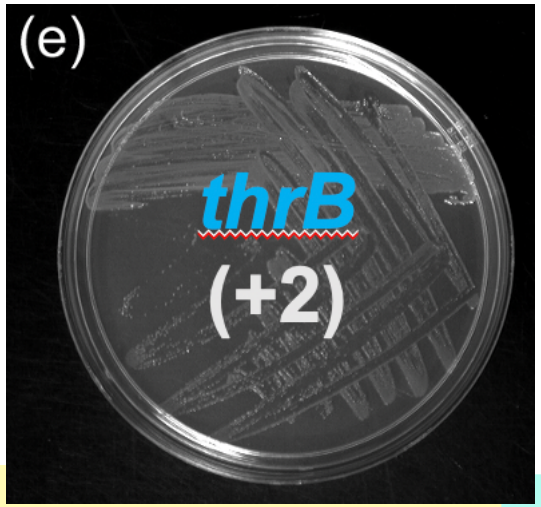
glycol-aldehyde

**low-specificity**
**threonine aldolase**
**LtaE**

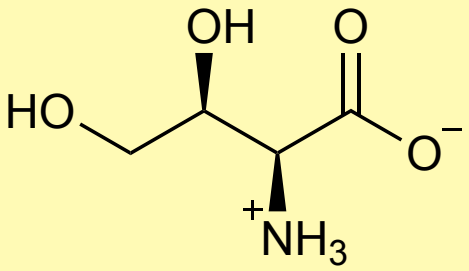



glycol-
aldehyde

4-hydroxy-L-threonine
(4HT)

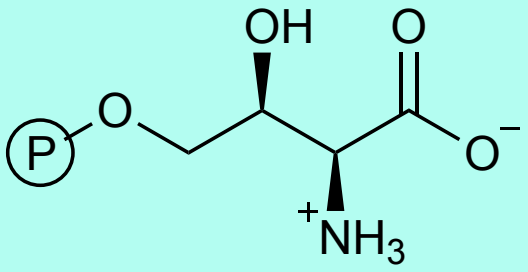



**homoserine kinase**



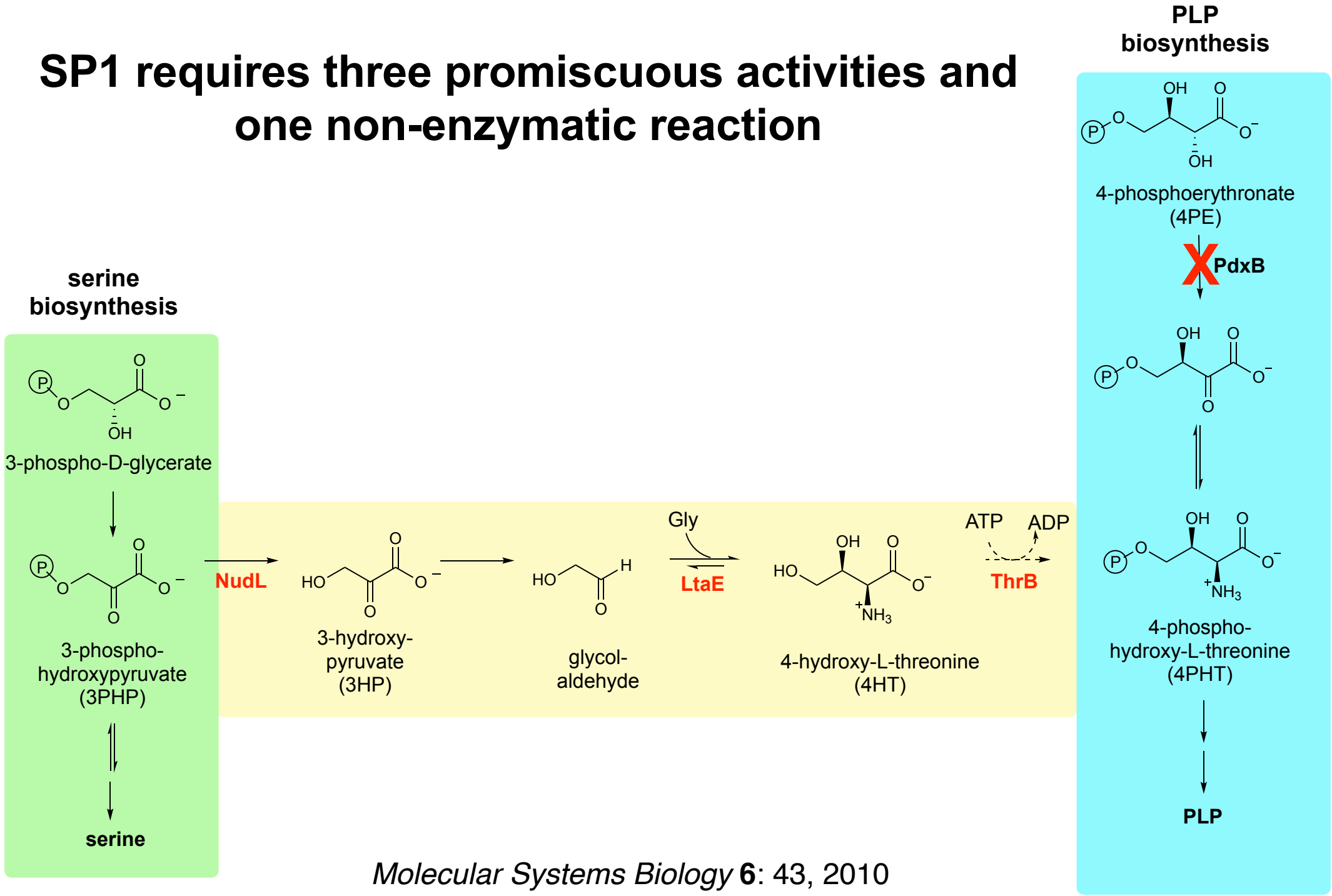

4-hydroxy-L-threonine
(4HT)

ATP ADP



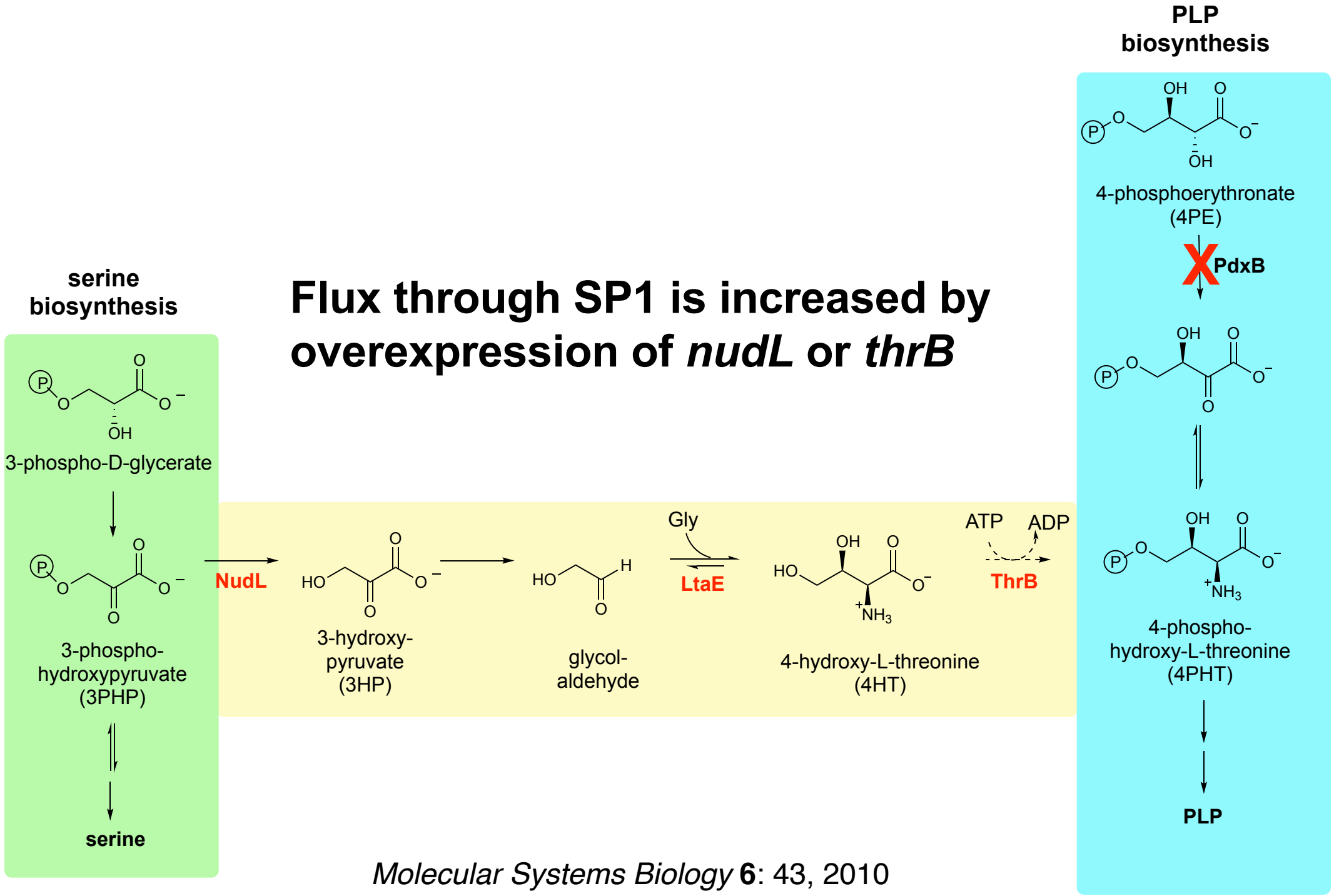

4-phospho-
hydroxy-L-threonine
(4PHT)

# SP1 requires three promiscuous activities and one non-enzymatic reaction




*Molecular Systems Biology* **6**: 43, 2010

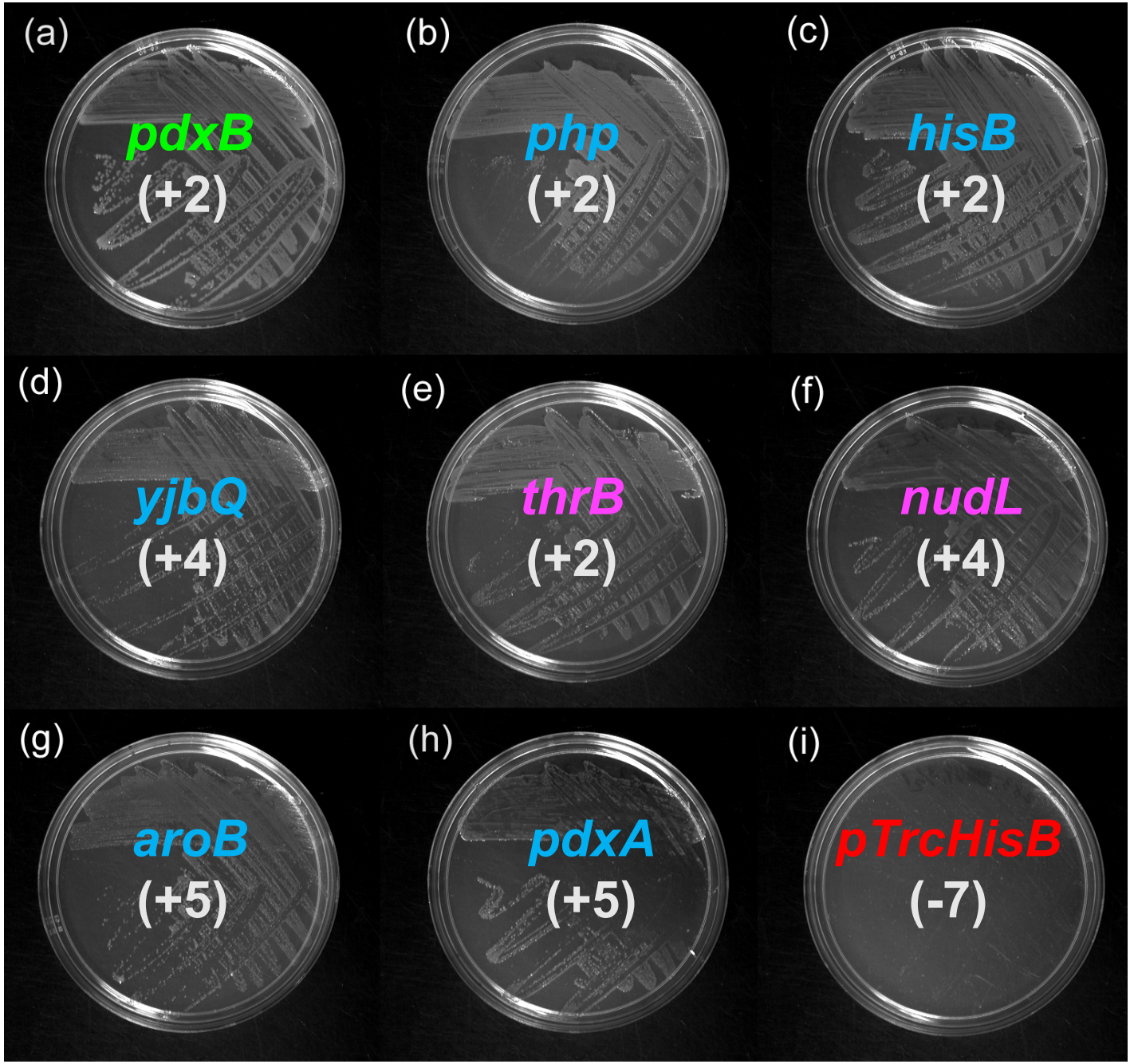
# Flux through SP1 is increased by overexpression of *nudL* or *thrB*



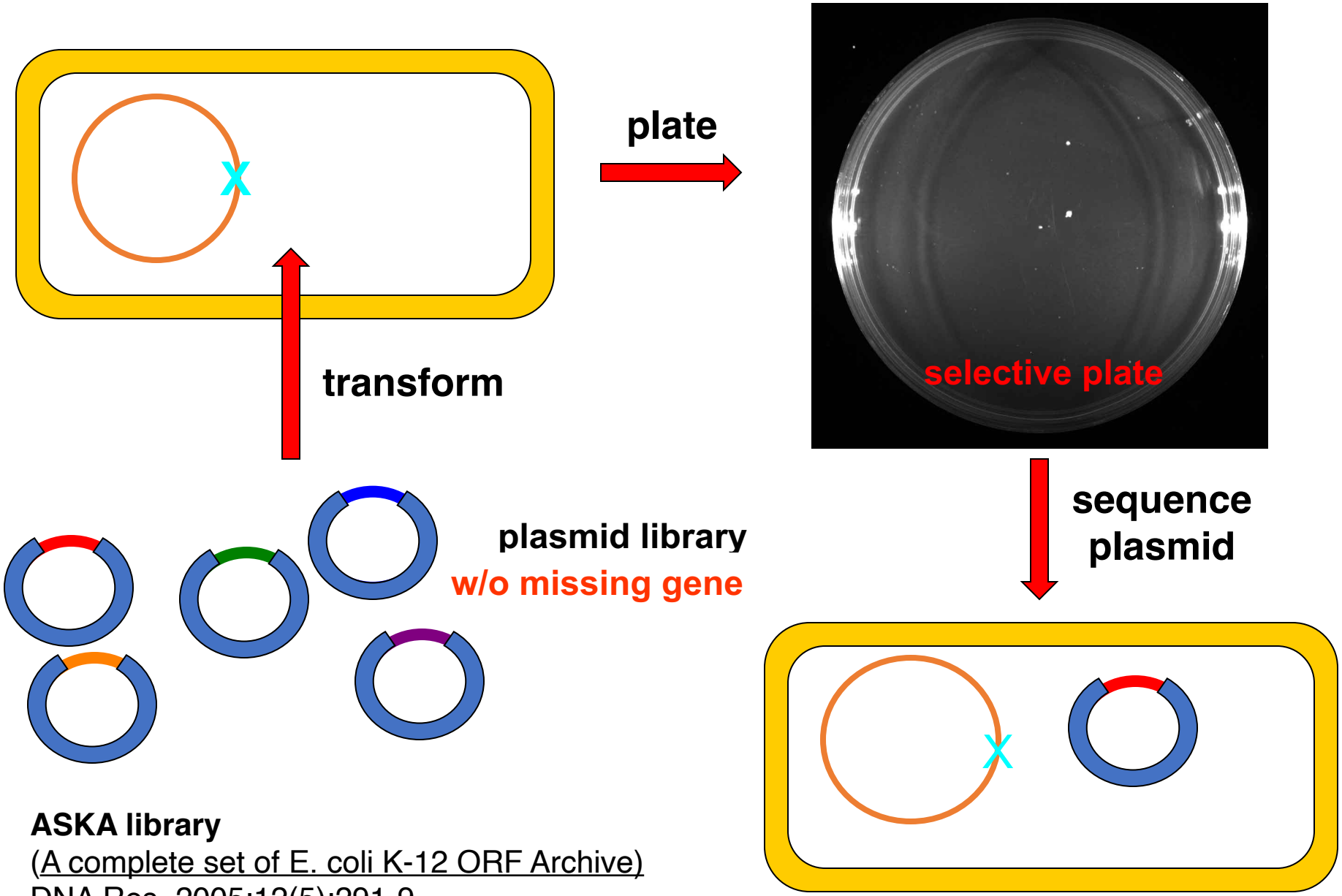

*Molecular Systems Biology* **6**: 43, 2010

Questions?

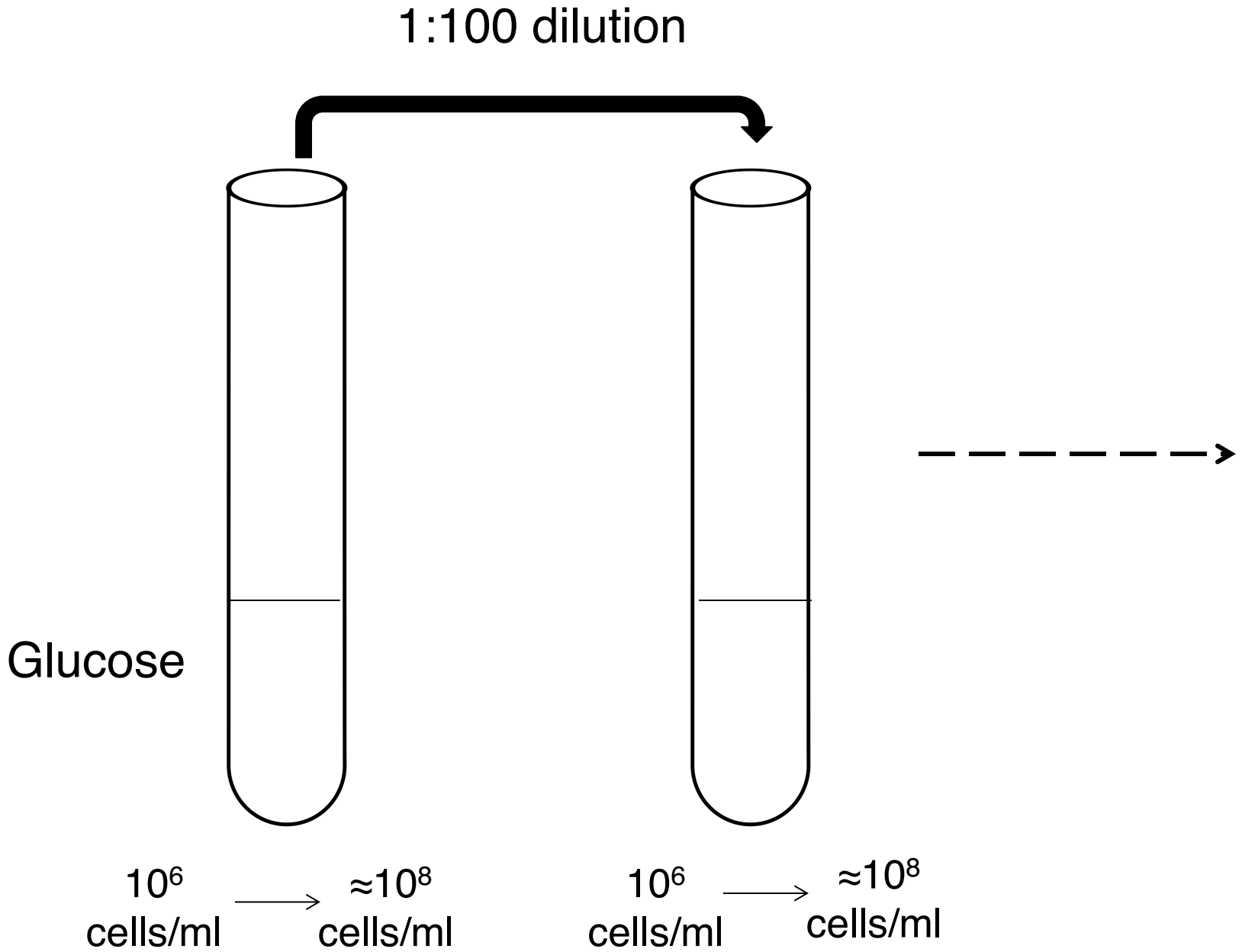
**At least two other SPs are facilitated by overexpression of single enzymes**

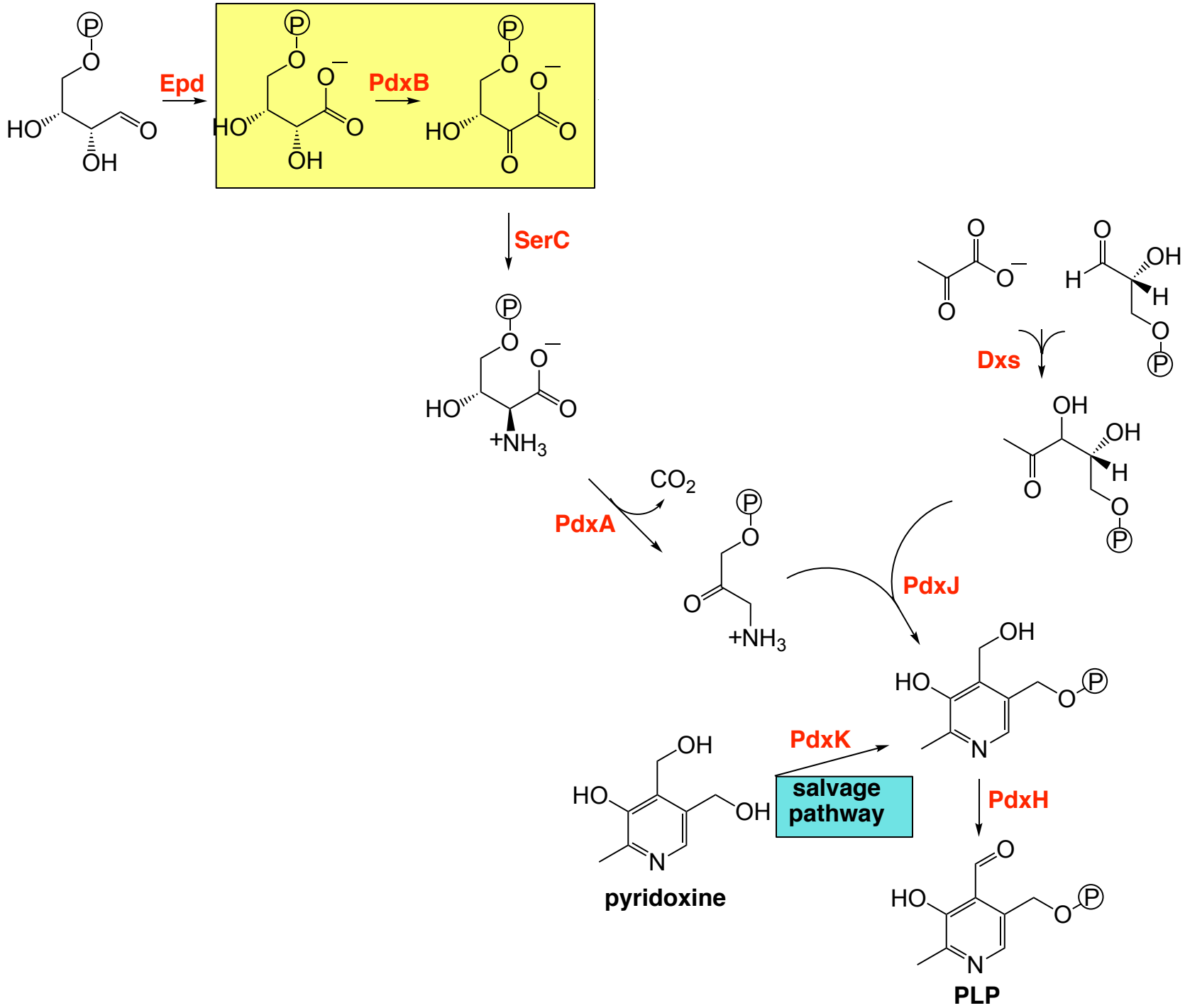


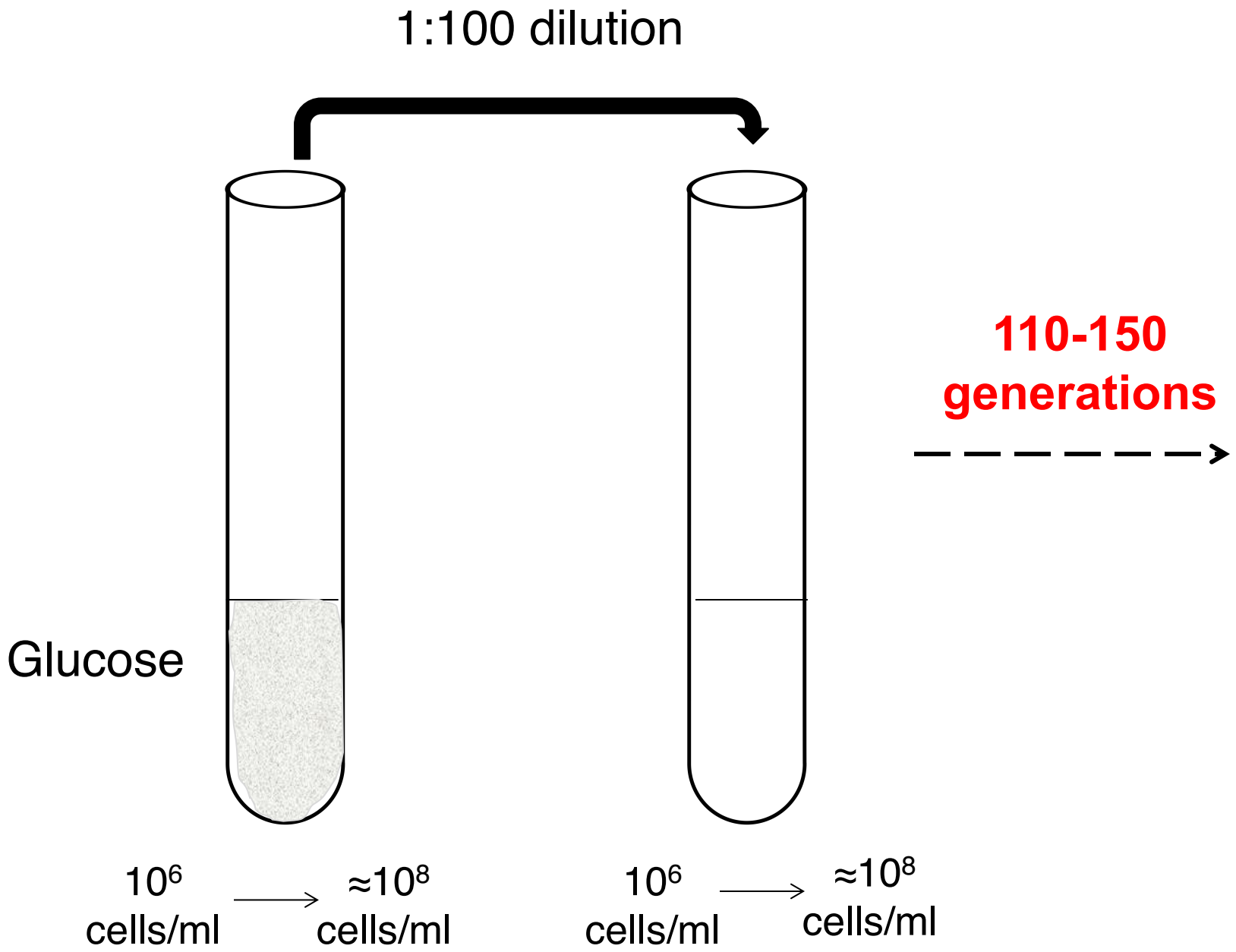
# Multicopy suppression

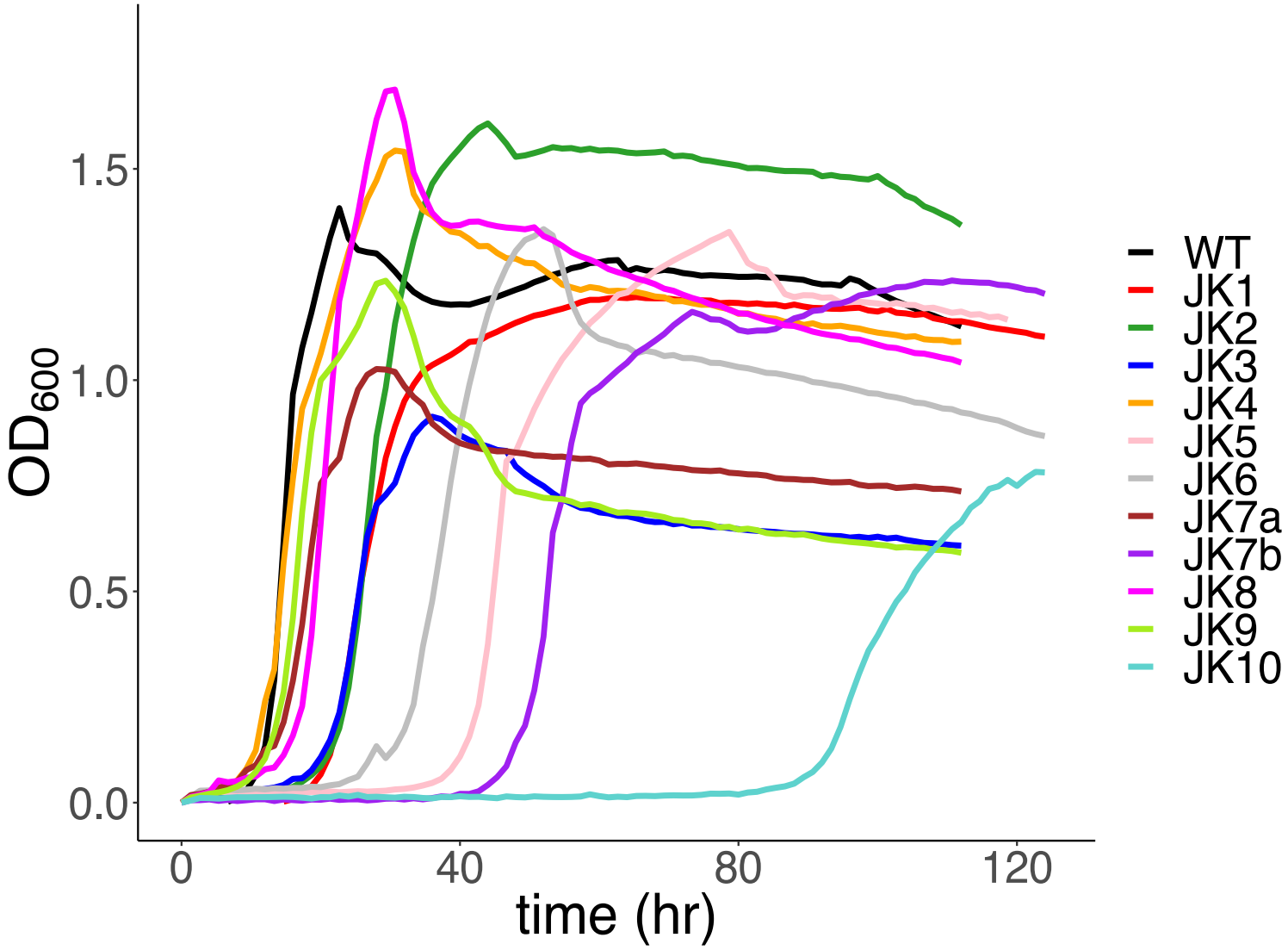



**ASKA library**
(A complete set of E. coli K-12 ORF Archive)
DNA Res. 2005;12(5):291-9


1:100 dilution
Glucose
$10^6$ cells/ml
$\approx 10^8$ cells/ml
$10^6$ cells/ml
$\approx 10^8$ cells/ml


Epd
PdxB
SerC
Dxs
PdxA
PdxJ
PdxK
salvage pathway
PdxH
pyridoxine
PLP

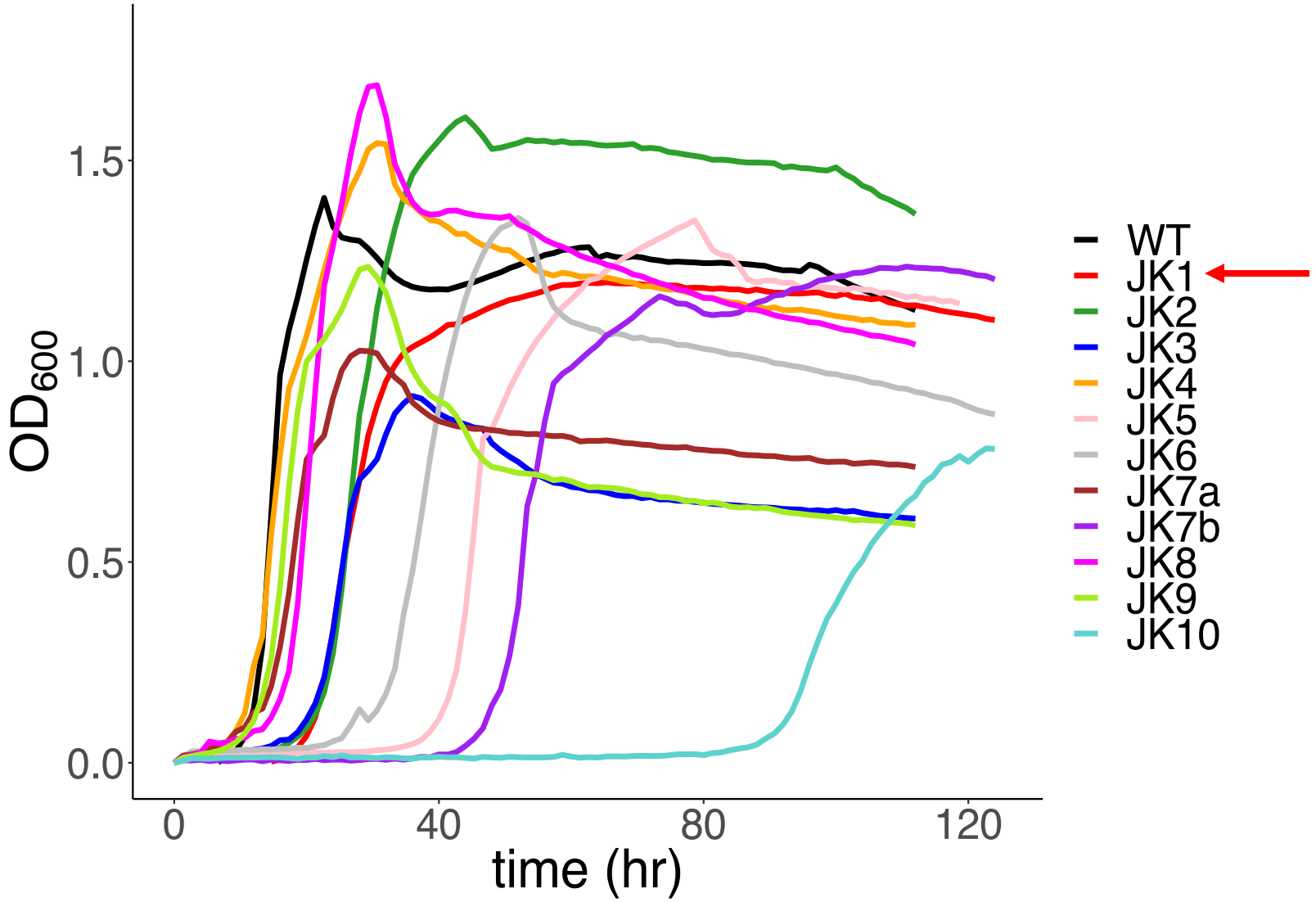

1:100 dilution
110-150
generations
Glucose
10^6
cells/ml
≈10^8
cells/ml
10^6
cells/ml
≈10^8
cells/ml


1.5
1.0
0.5
0.0
$OD_{600}$
0
40
80
120
time (hr)
WT
JK1
JK2
JK3
JK4
JK5
JK6
JK7a
JK7b
JK8
JK9
JK10

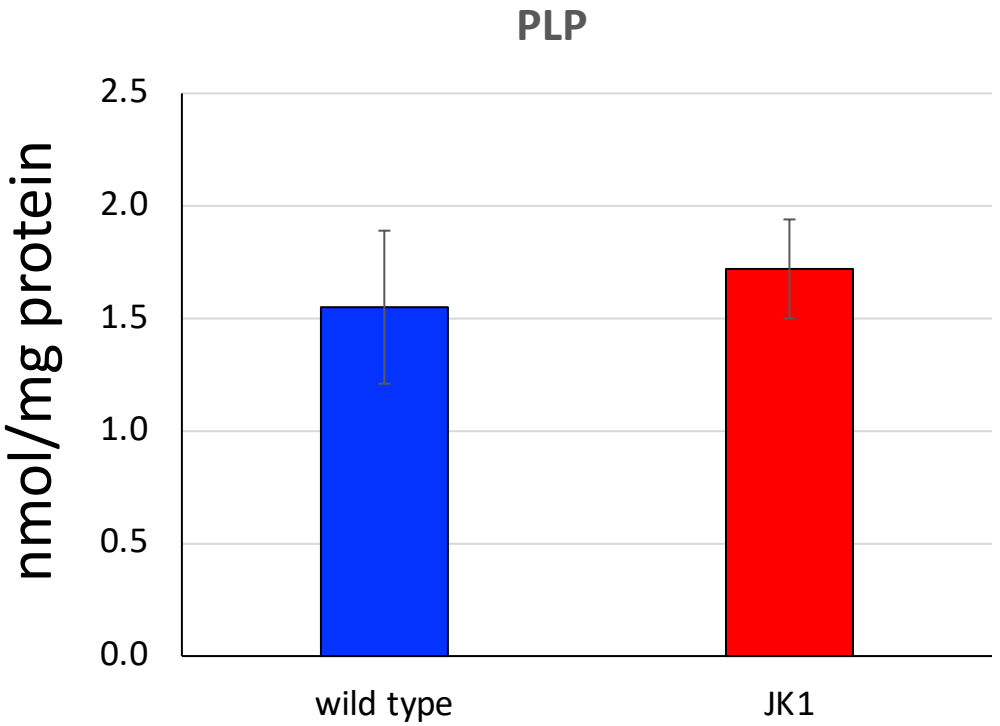
Juhan Kim

1) How are the evolved strains making PLP?

2) How do mutations improve PLP synthesis?

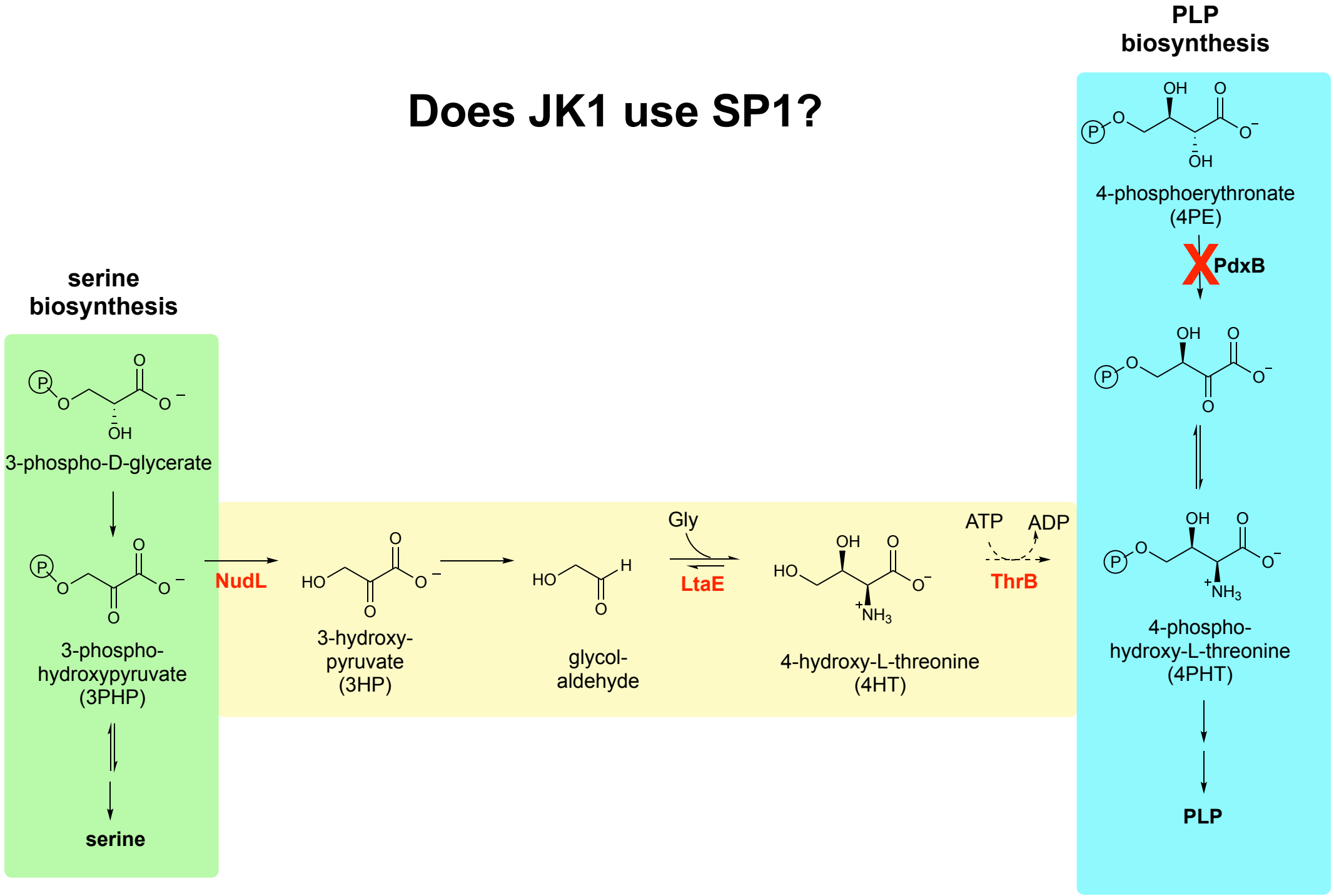



*PNAS 116,* 24164-24173, 2019



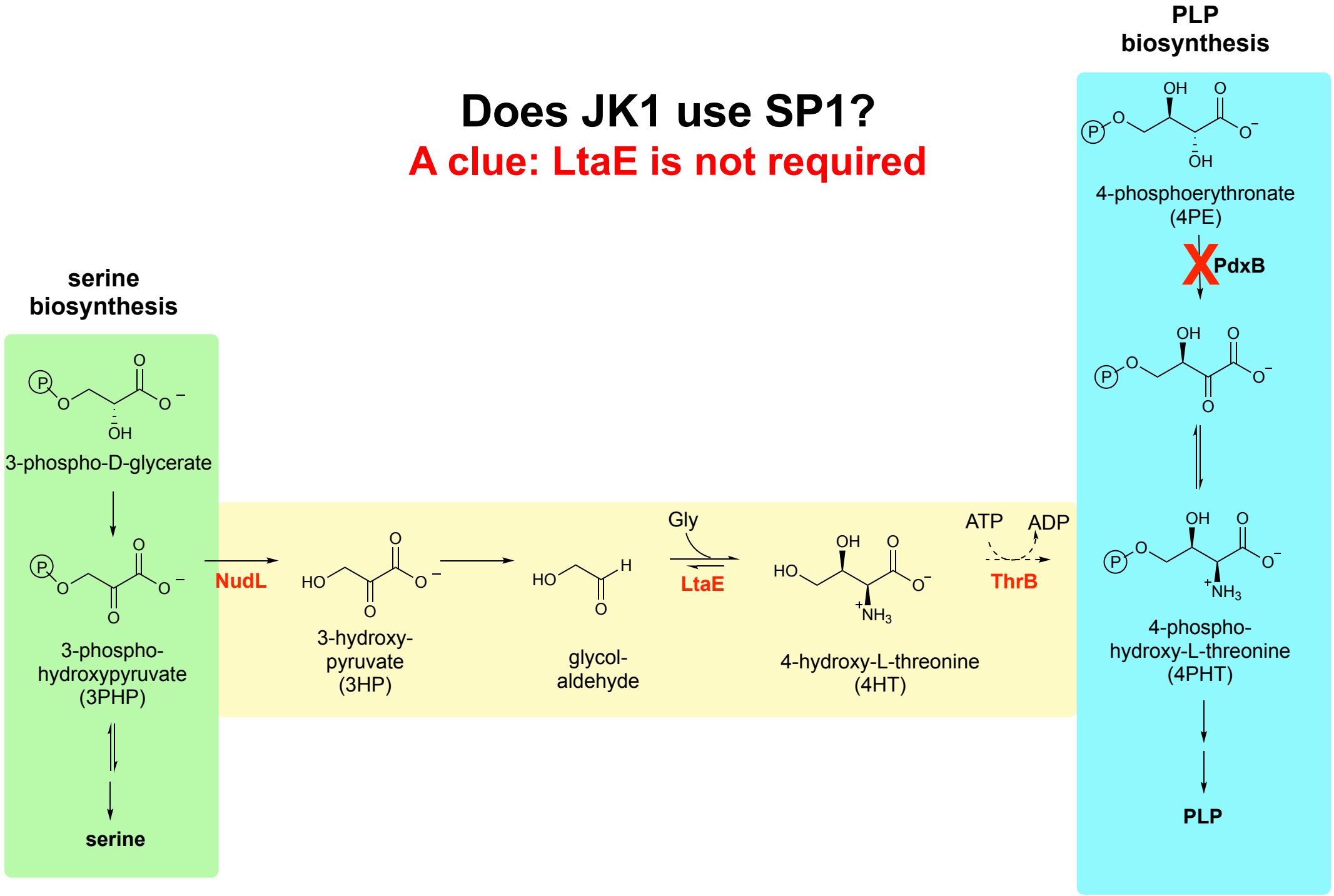

*PNAS 116,* 24164-24173, 2019

# Does JK1 use SP1?



# Does JK1 use SP1?

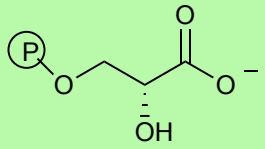
## A clue: LtaE is not required



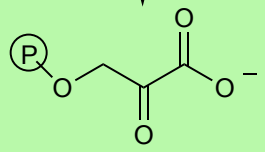
# Does JK1 use SP1?

## A clue: LtaE is not required but ThrB *is* required (even when threonine is supplied)

**serine biosynthesis**



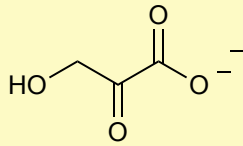

3-phospho-D-glycerate



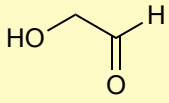

3-phospho-hydroxypyruvate (3PHP)

**serine**

**NudL**



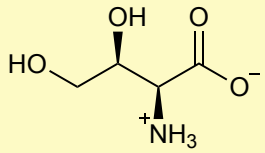

3-hydroxy-pyruvate (3HP)




glycol-aldehyde

Gly

**LtaE**

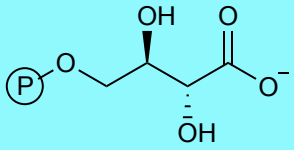



4-hydroxy-L-threonine (4HT)

ATP ADP

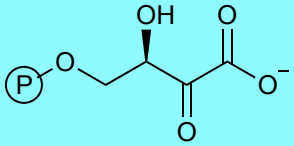
**ThrB**

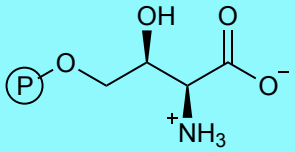
**PLP biosynthesis**




4-phosphoerythronate (4PE)

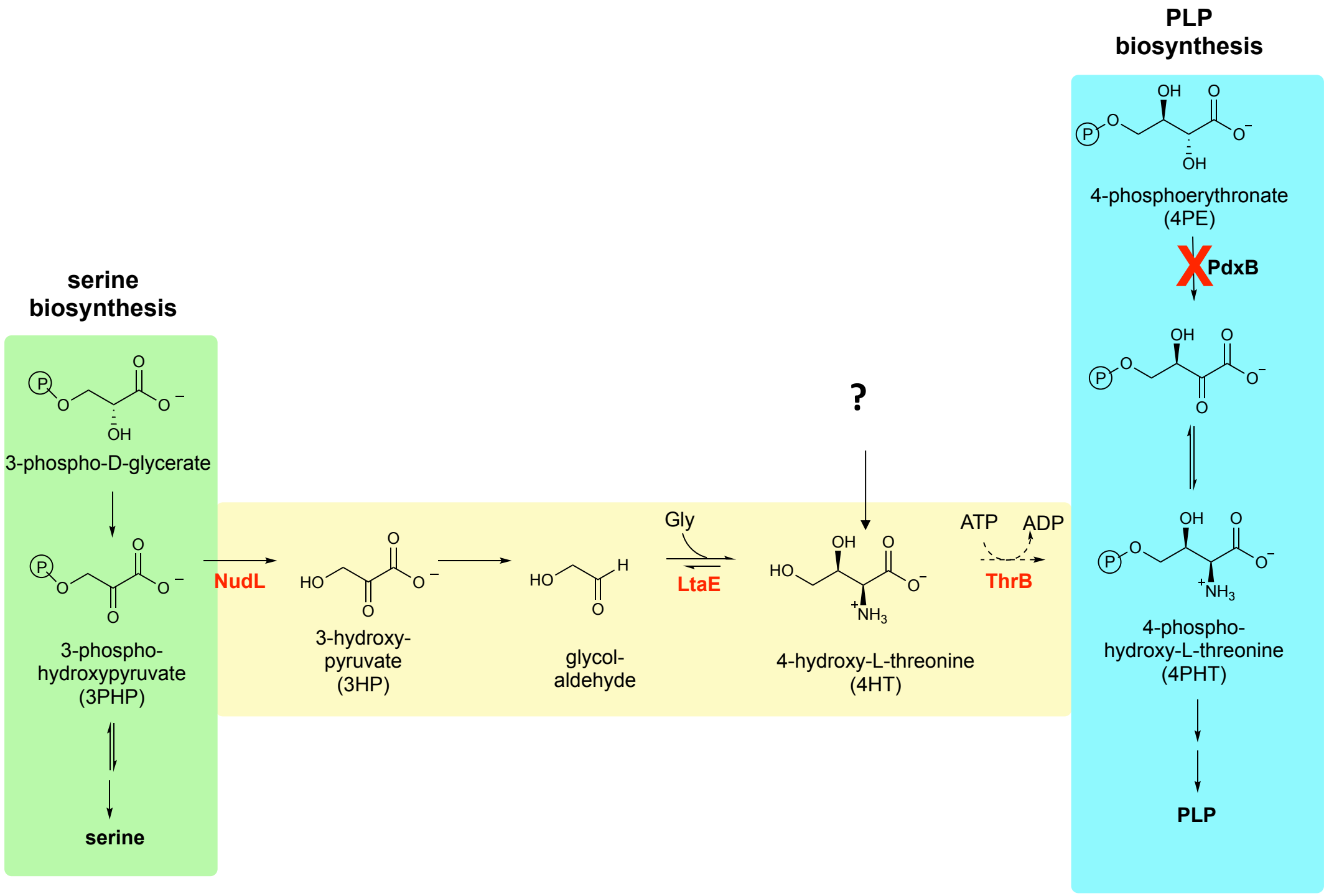
**PdxB**

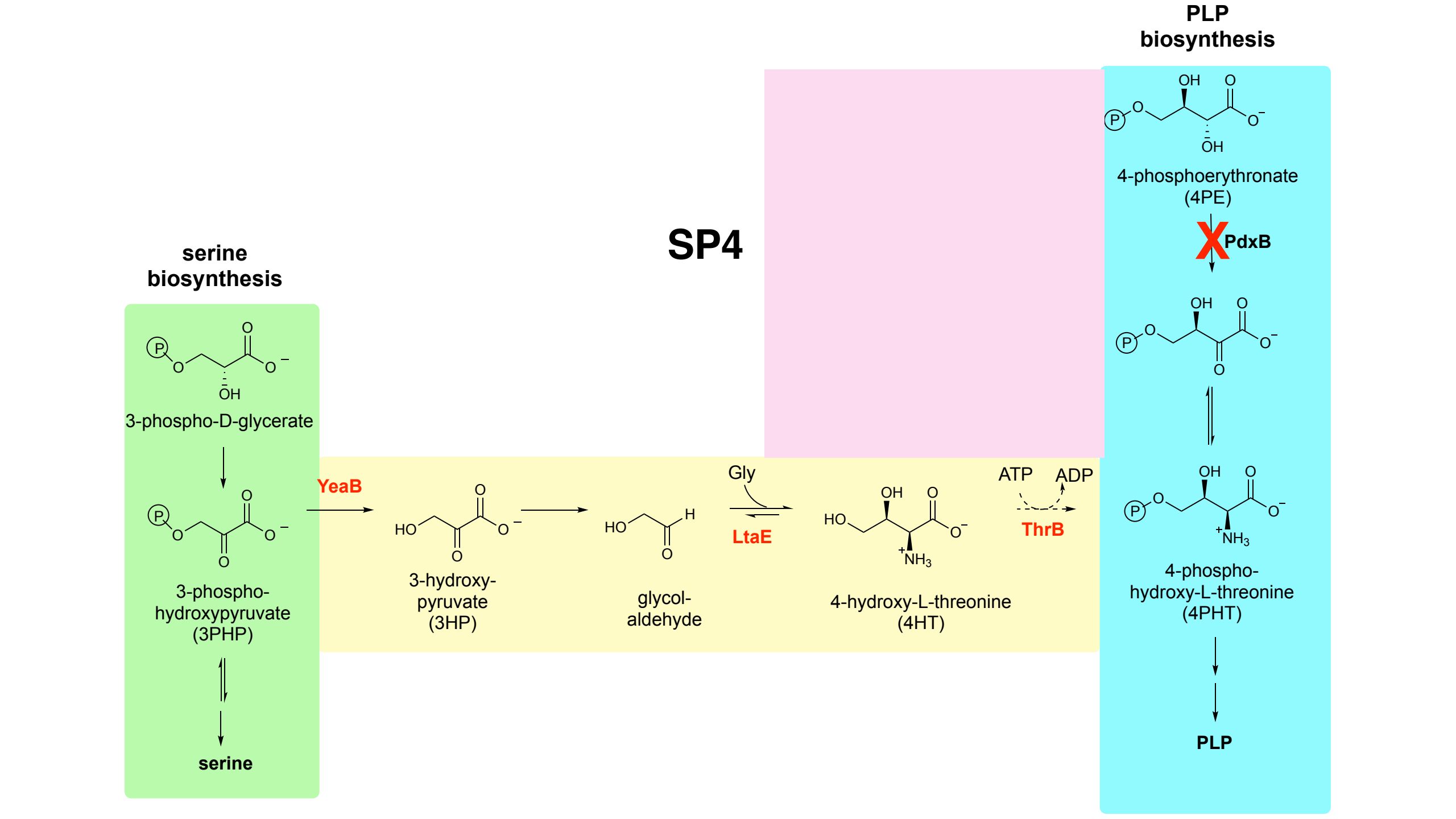


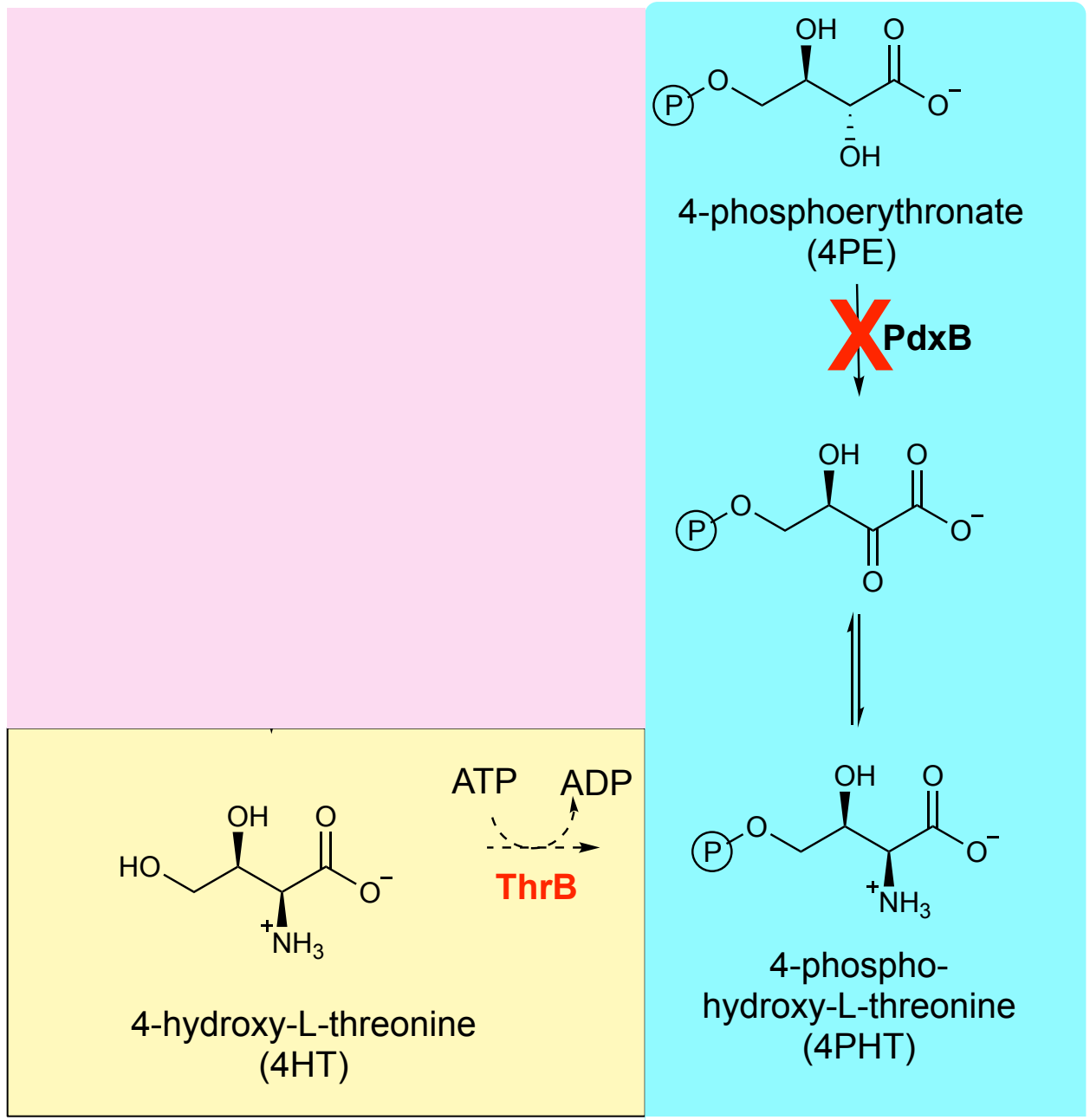



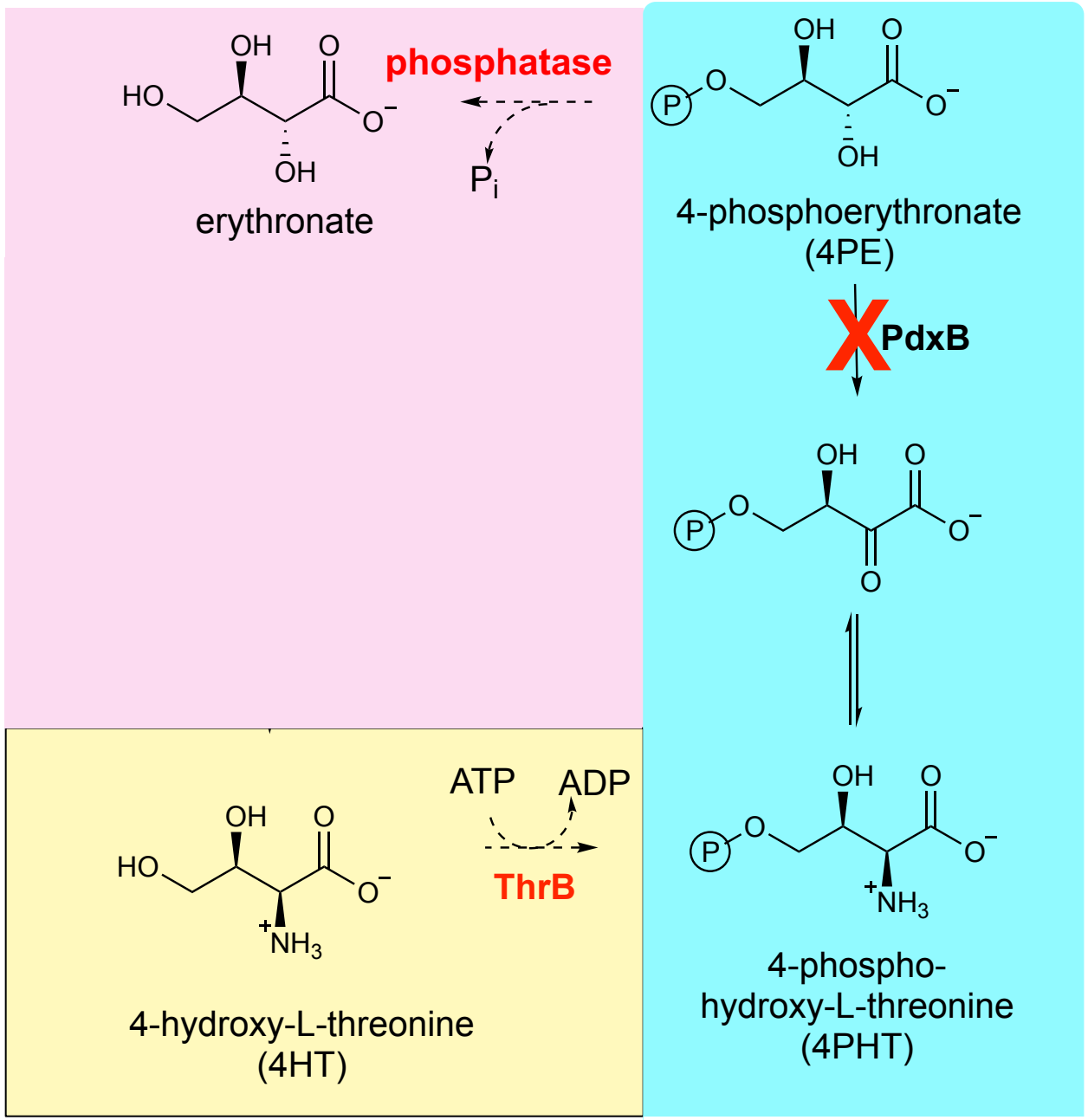

4-phospho-hydroxy-L-threonine (4PHT)

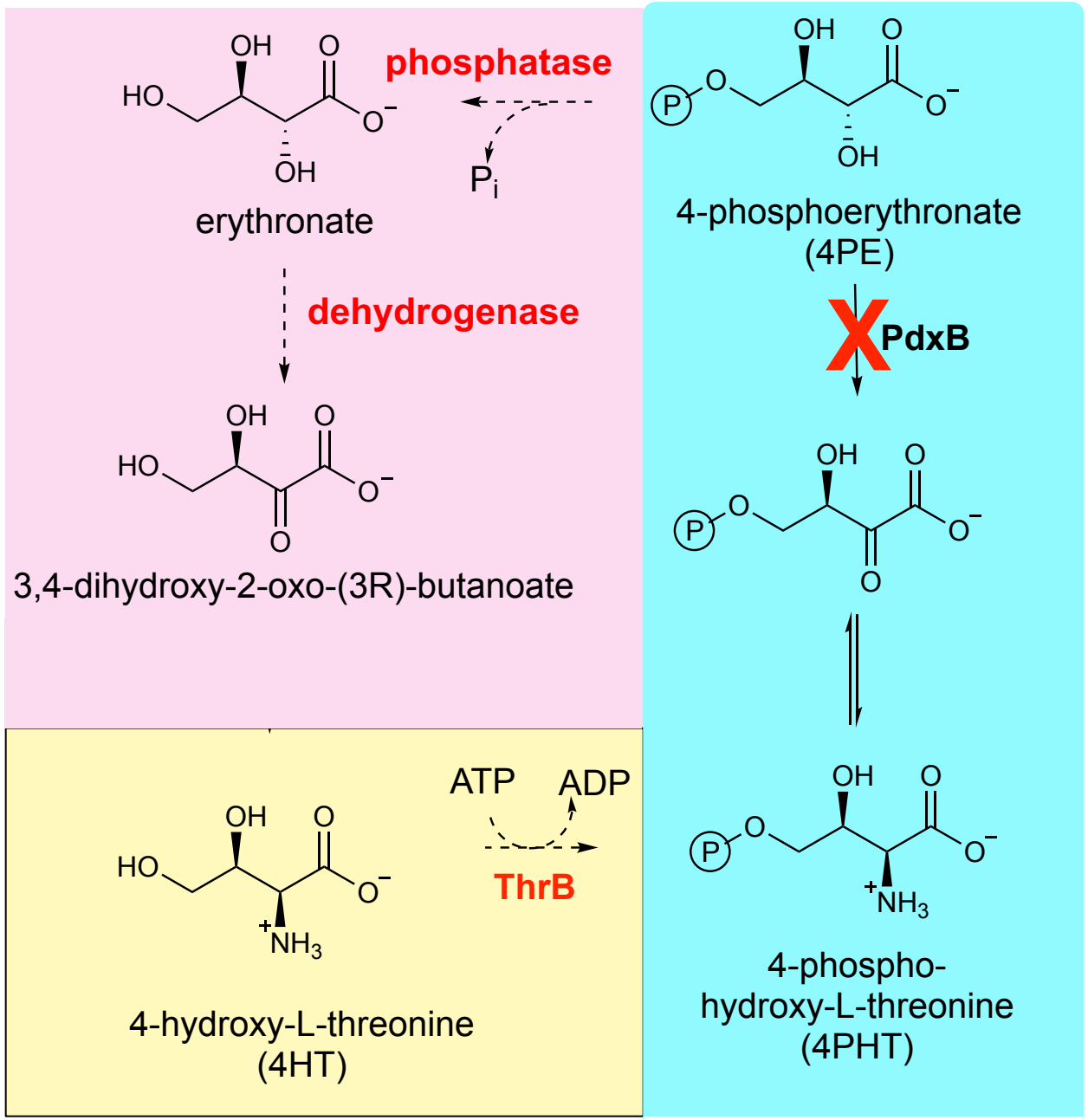
**PLP**

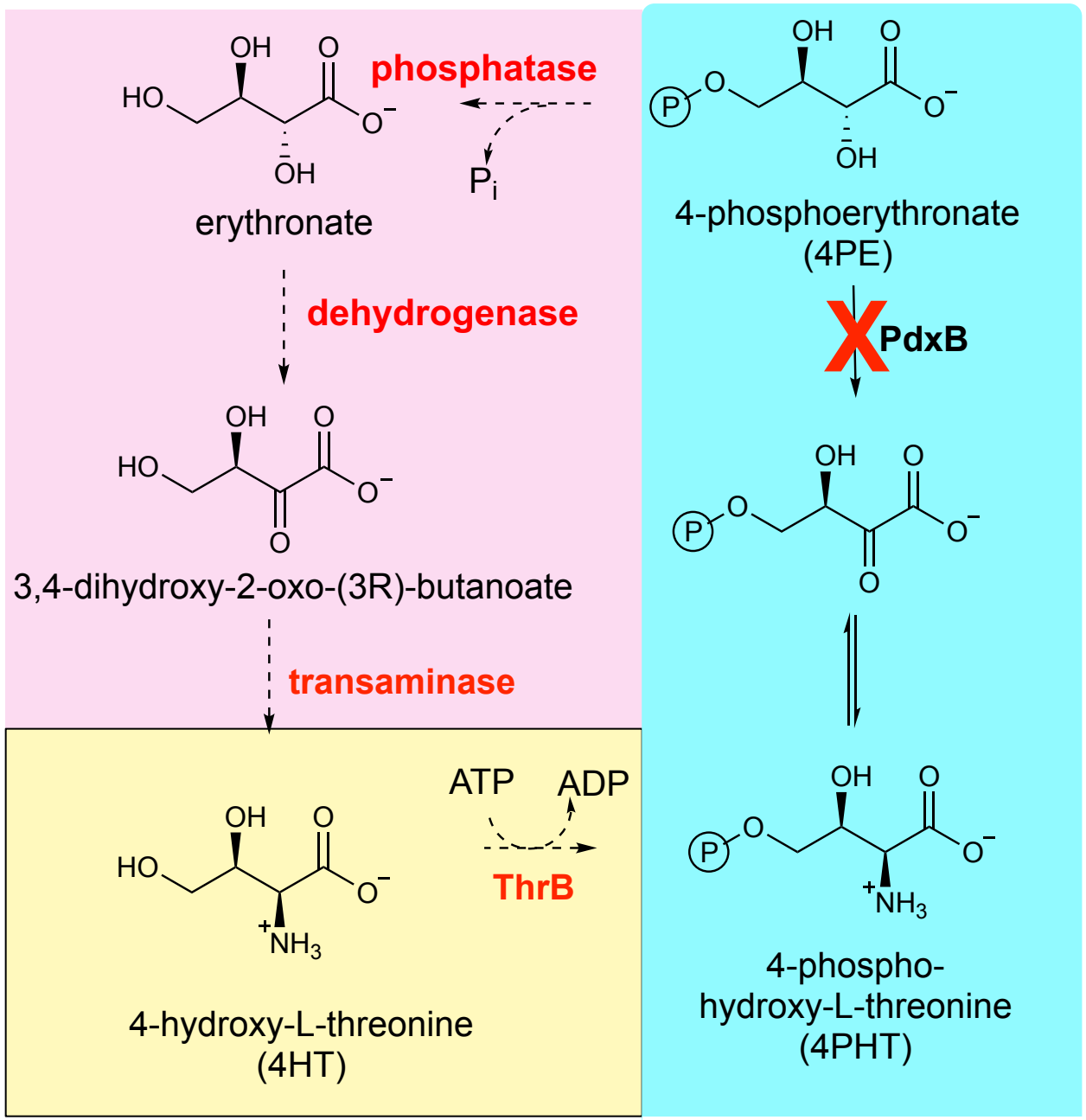

PLP
biosynthesis
serine
biosynthesis
3-phospho-D-glycerate
3-phospho-
hydroxypyruvate
(3PHP)
serine
NudL
3-hydroxy-
pyruvate
(3HP)
glycol-
aldehyde
Gly
LtaE
?
4-hydroxy-L-threonine
(4HT)
ATP
ADP
ThrB
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)
PLP

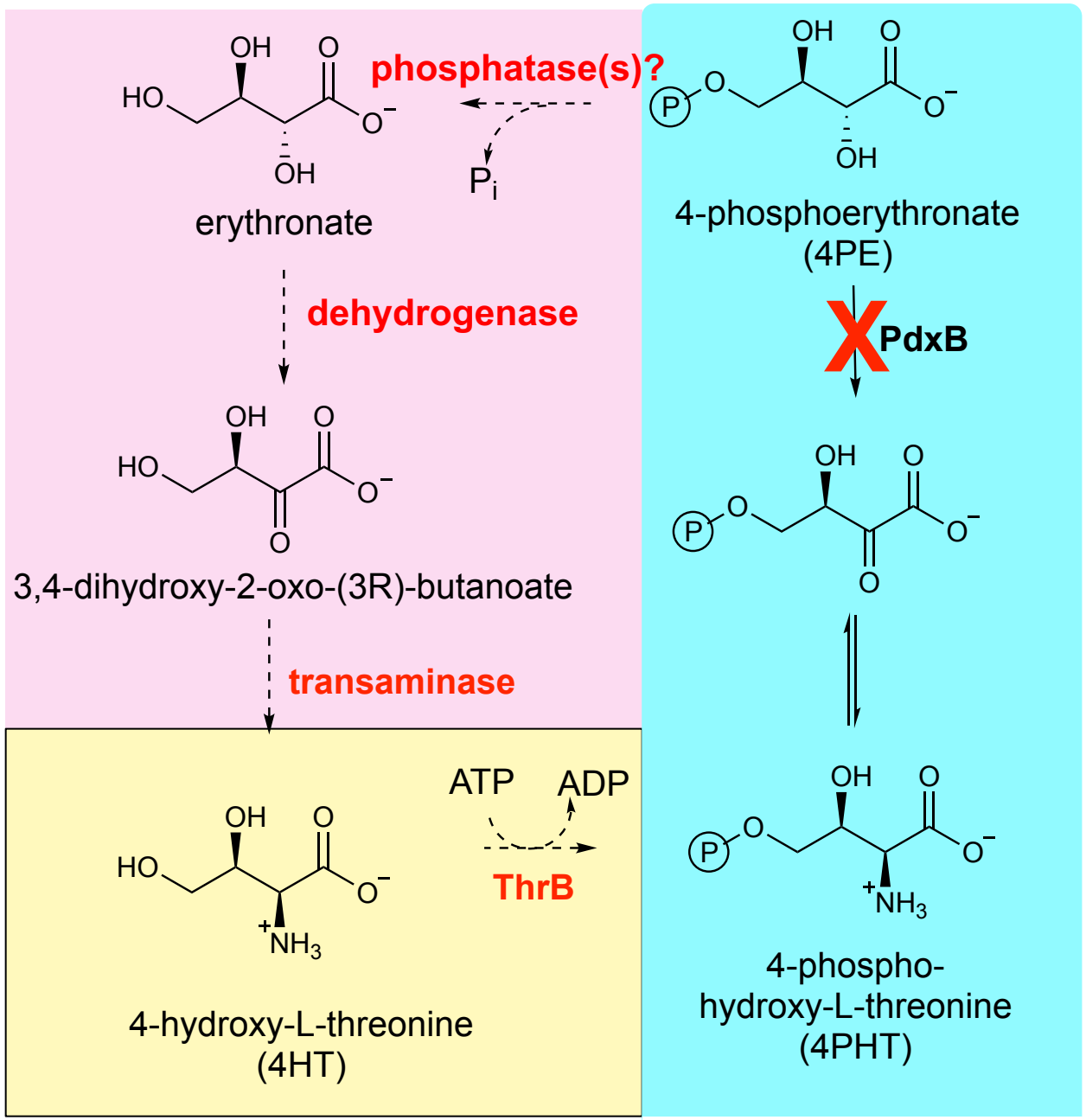

PLP
biosynthesis
SP4
serine
biosynthesis
4-phosphoerythronate
(4PE)
PdxB
3-phospho-D-glycerate
YeaB
3-phospho-
hydroxypyruvate
(3PHP)
3-hydroxy-
pyruvate
(3HP)
glycol-
aldehyde
Gly
LtaE
4-hydroxy-L-threonine
(4HT)
ATP
ADP
ThrB
4-phospho-
hydroxy-L-threonine
(4PHT)
serine
PLP

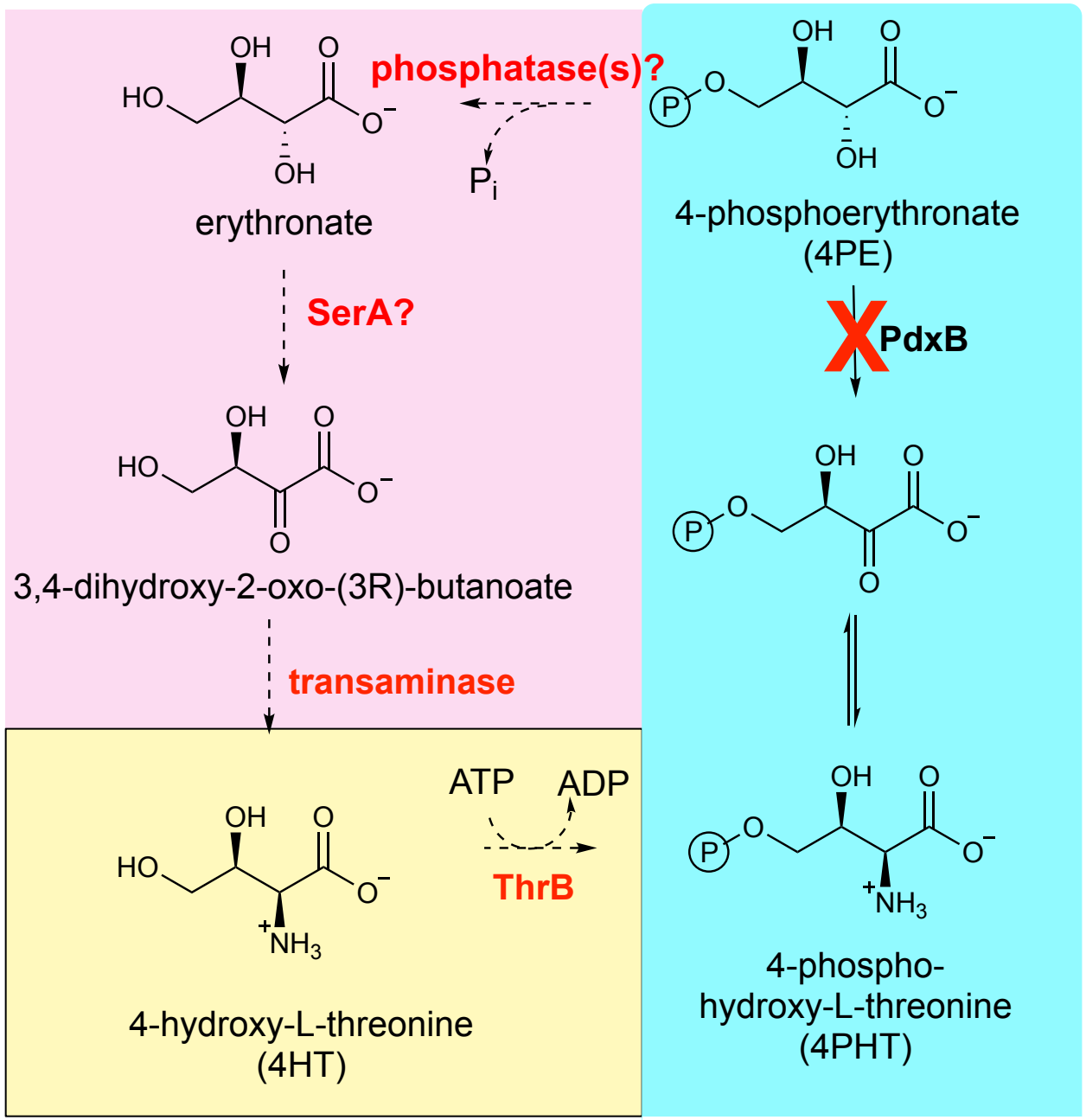

OH
O
O
P
O⁻
OH
4-phosphoerythronate
(4PE)
PdxB
OH
O
O
P
O⁻
O
4-phospho-
hydroxy-L-threonine
(4PHT)
ATP
ADP
ThrB
OH
O
HO
O⁻
+NH3
4-hydroxy-L-threonine
(4HT)


phosphatase
erythronate
$P_i$
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)


phosphatase
erythronate
$P_i$
dehydrogenase
3,4-dihydroxy-2-oxo-(3R)-butanoate
4-phosphoerythronate
(4PE)
PdxB
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phospho-
hydroxy-L-threonine
(4PHT)


phosphatase
erythronate
dehydrogenase
3,4-dihydroxy-2-oxo-(3R)-butanoate
transaminase
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)


phosphatase(s)?
Pi
erythronate
dehydrogenase
3,4-dihydroxy-2-oxo-(3R)-butanoate
transaminase
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)


phosphatase(s)?
Pi
erythronate
SerA?
3,4-dihydroxy-2-oxo-(3R)-butanoate
transaminase
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

**SerA = 3-phospho-glycerate dehydrogenase**

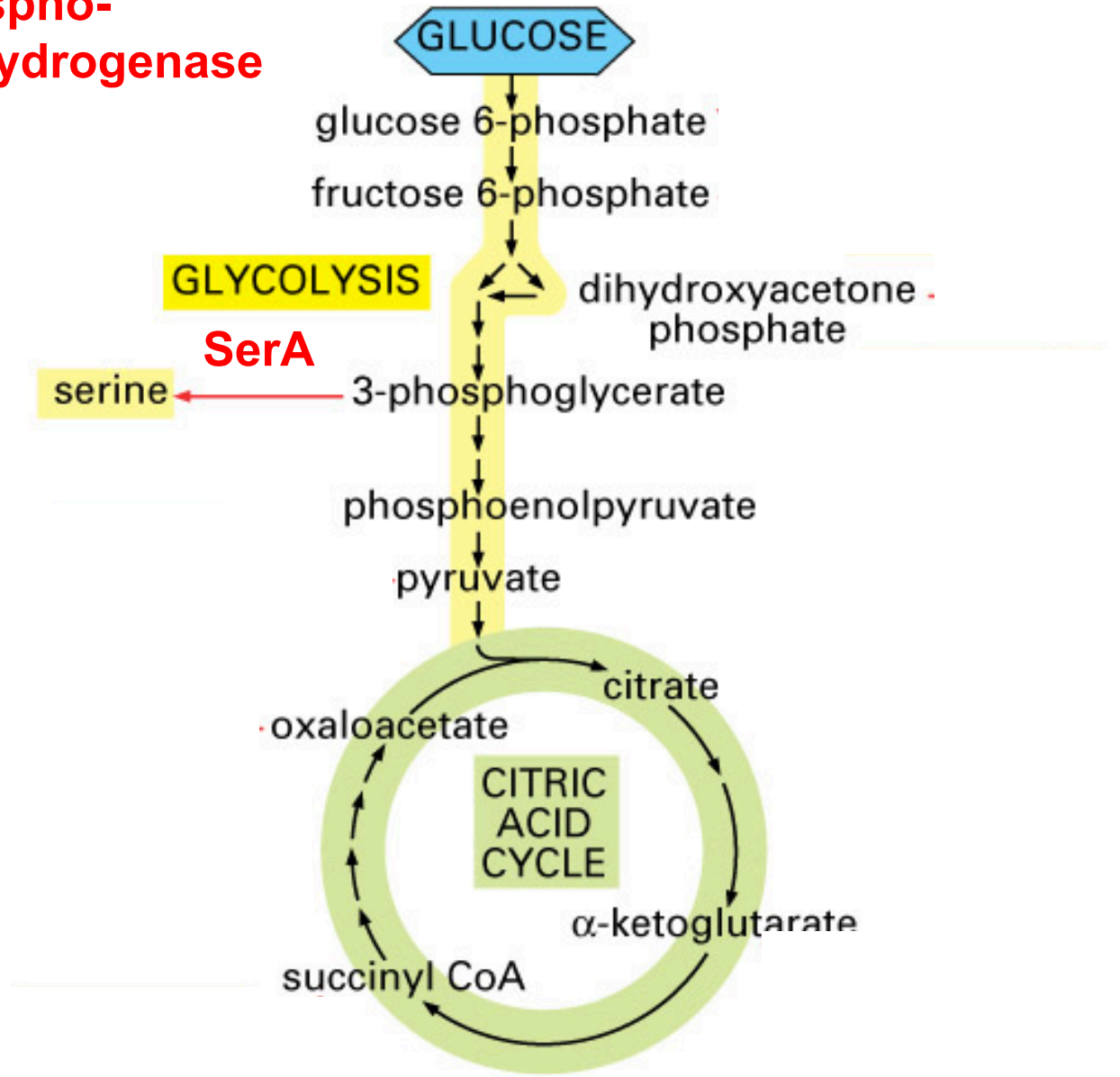



Figure 13-23 Essential Cell Biology, 2/e. (© 2004 Garland Science)

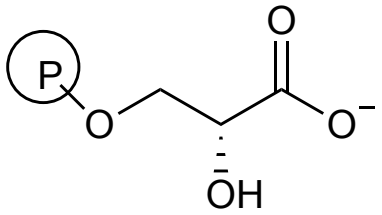
# Mutations in *serA* were found in multiple strains

| JK1 | JK2 | JK3 | JK4 | JK5 | JK6 |
|---|---|---|---|---|---|
| *ybhA/pgl* | *ybhA/pgl* | *ybhA/pgl* | *ybhA* | *ybhA/pgl* | *gapA* |
| *gapA* | *gapA* | *gapA* | *rpe* | *gapA* | ***serA*** |
| *rpoS* | *purF* | *ilvH* | *sdhA* | *yjjK* | *yjjK* |
| *rpoC* | *gltB* | *rng* | *rho* | *purF* | |
| | *ypjA* | | *lon* | *ilvH* | |
| | | | | *nadR* | |

| JK7a | JK7b | JK8 | JK9 | JK10 |
|---|---|---|---|---|
| *ybhA* | *ybhA/pgl* | *gapA* | *ybhA/pgl* | *ybhA/pgl* |
| *gapA* | ***serA*** | ***serA*** | *gapA* | *gapA* |
| *purF* | *gapA* | *yjjK* | ***serA*** | *rpe* |
| *nadR* | *pykF* | *gltB* | *pykF* | *ilvH* |
| *rpoS* | *pyrE* | *livH* | | *rng* |

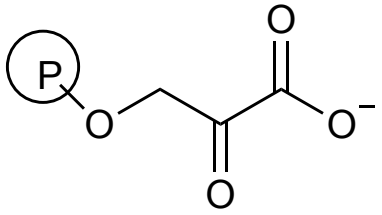
# Both reactions require oxidation of an alcohol alpha to a carboxylate

**serine biosynthesis**



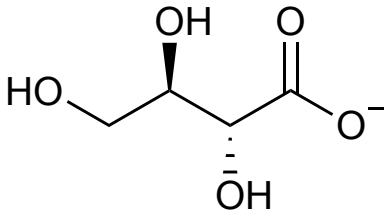

3-phospho-D-glycerate

**SerA**



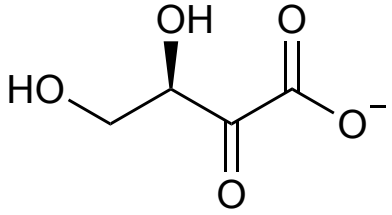

3-phosphooxypyruvate

**SP4**




D-erythronate

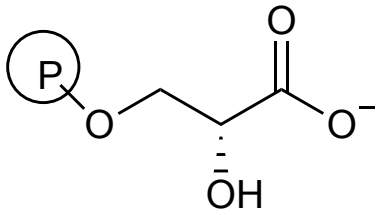
**SerA**




(3*R*)-3,4-dihydroxy-2-oxobutanoate
(DHOB)

# SerA has weak activity with erythronate

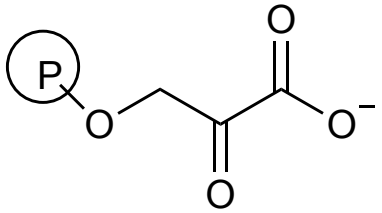
**serine biosynthesis**




3-phospho-D-glycerate

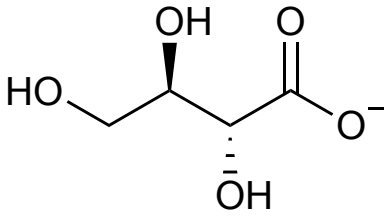
$k_{cat}/K_M$ = $1.6 \times 10^3\ M^{-1}s^{-1}$

**SerA**




3-phosphooxypyruvate

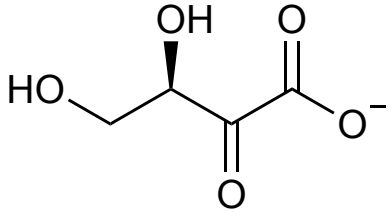
**SP4**




D-erythronate

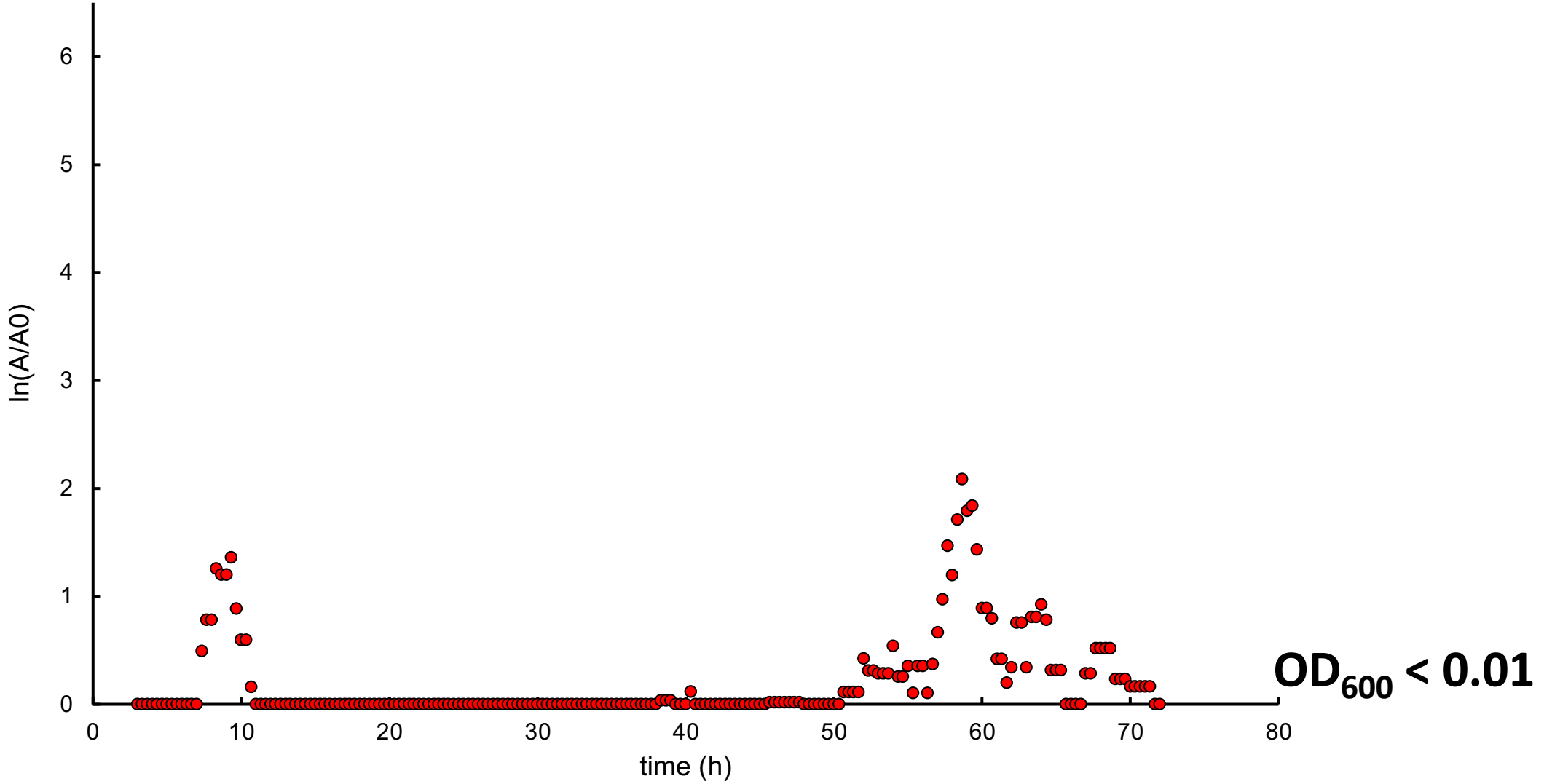
$k_{cat}/K_M$ = $0.07\ M^{-1}s^{-1}$

**SerA**

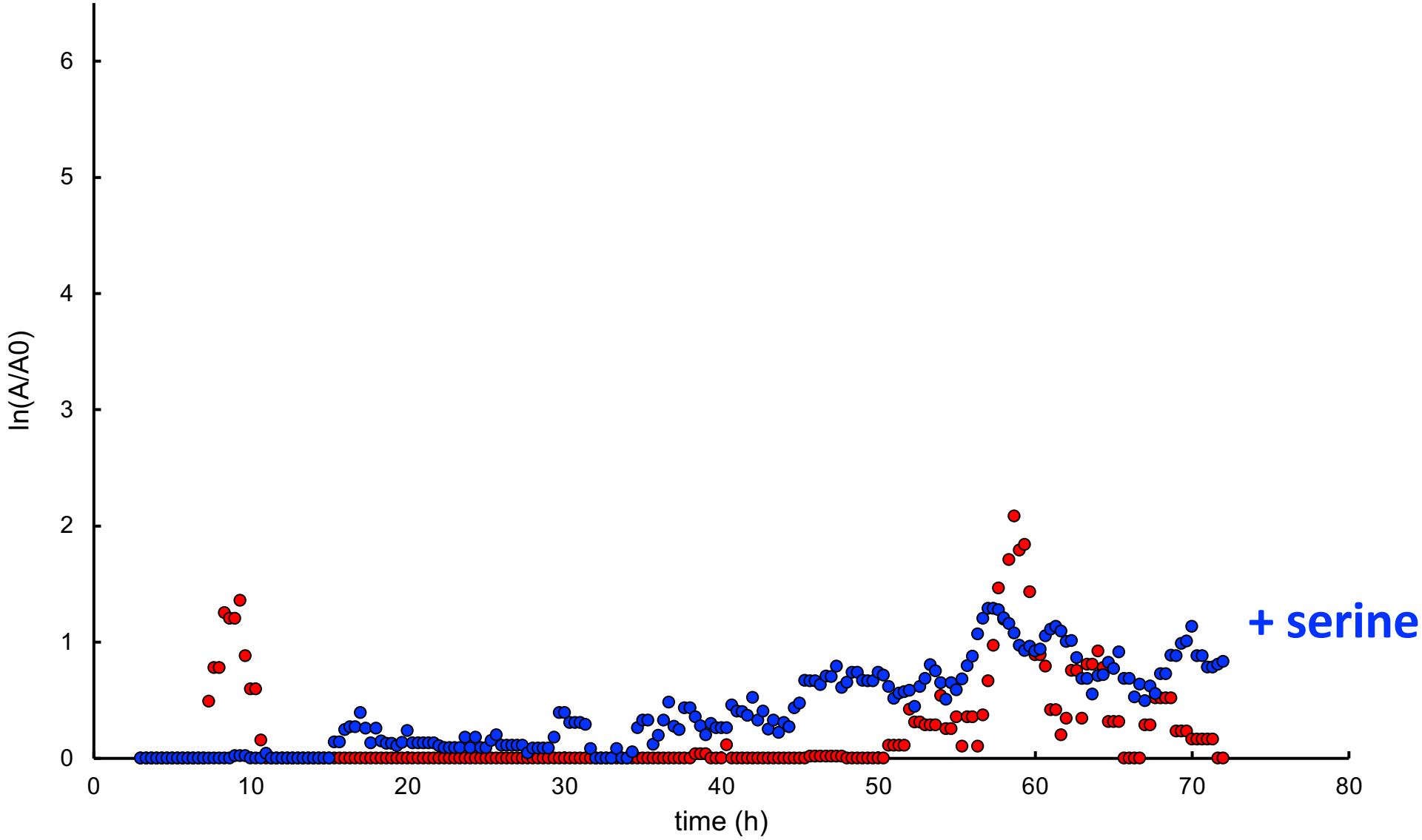



(3*R*)-3,4-dihydroxy-2-oxobutanoate

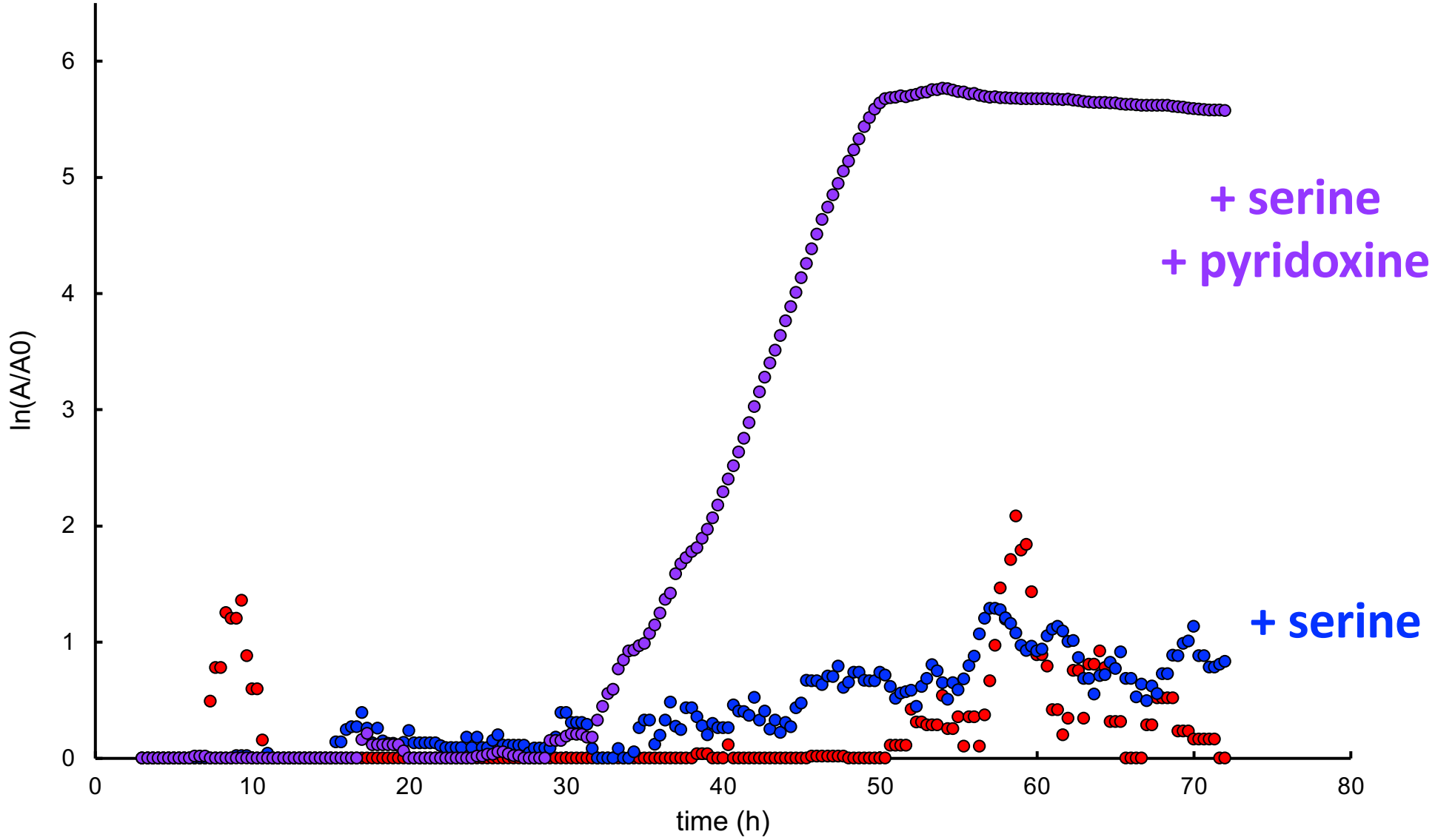
# *ΔserA* JK1

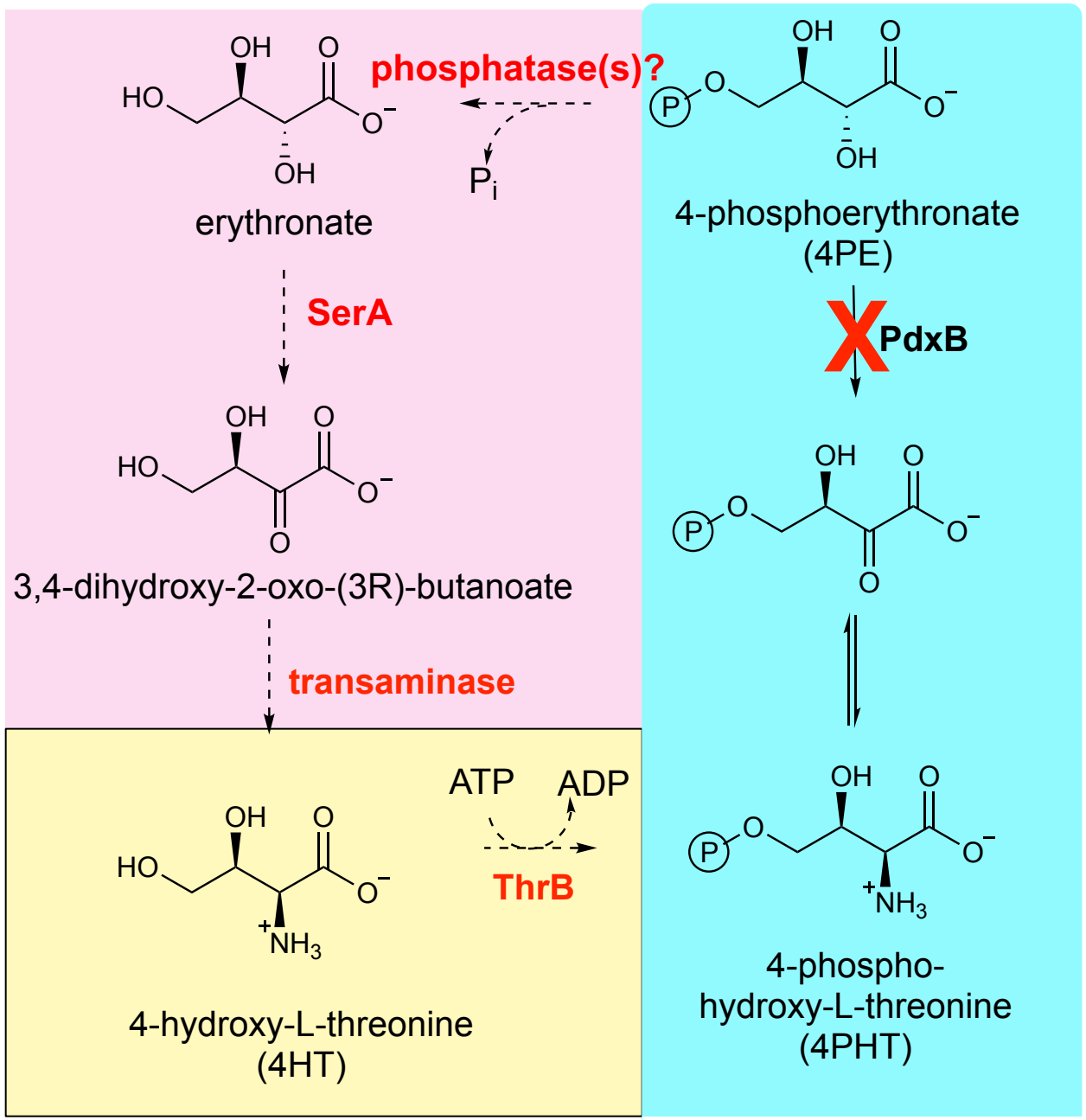


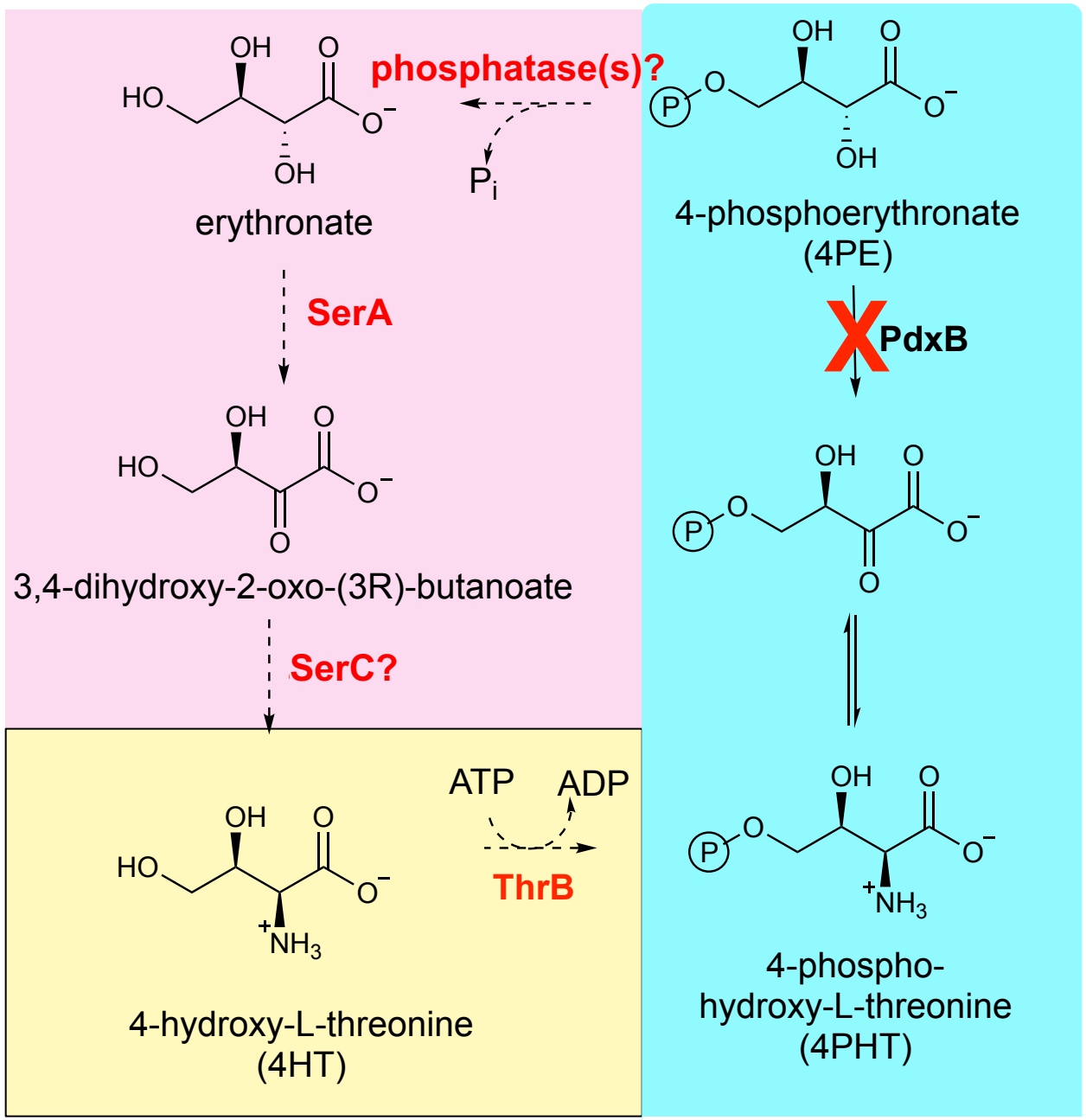
# *ΔserA* JK1



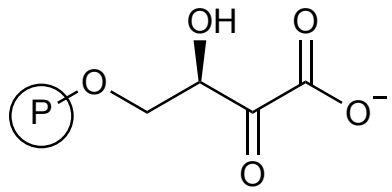
# *ΔserA* JK1


ln(A/A0)
0
1
2
3
4
5
6
0
10
20
30
40
50
60
70
80
time (h)
+ serine
+ pyridoxine
+ serine


phosphatase(s)?
Pi
erythronate
SerA
3,4-dihydroxy-2-oxo-(3R)-butanoate
transaminase
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

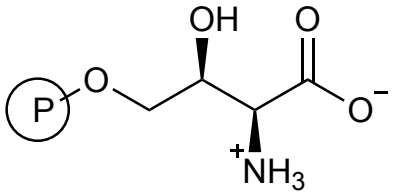

phosphatase(s)?
Pi
erythronate
SerA
3,4-dihydroxy-2-oxo-(3R)-butanoate
SerC?
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

## PLP biosynthesis



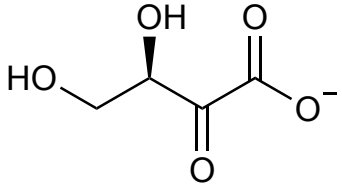

(3*R*)-3-hydroxy-2-oxo-4-phosphooxybutanoate

**SerC**



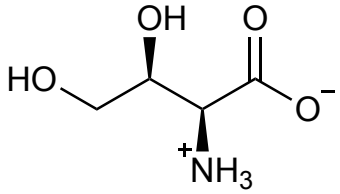

4-phosphooxy-L-threonine

## SP4



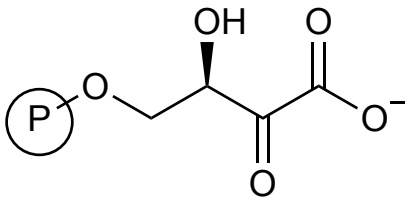

(3*R*)-3,4-dihydroxy-2-oxobutanoate

**SerC**

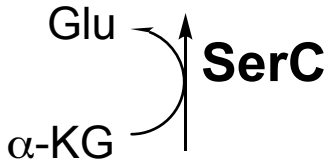



4-hydroxy-L-threonine
(4HT)

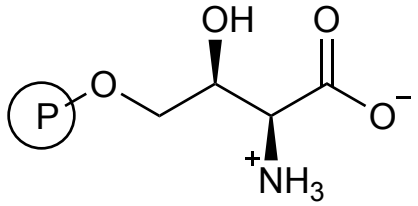
## PLP biosynthesis




(3*R*)-3-hydroxy-2-oxo-4-phosphooxybutanoate

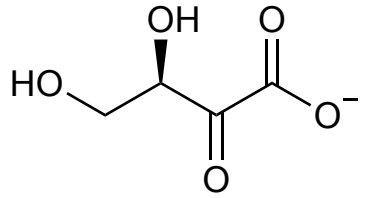



$k_{cat}/K_M$ = 727 $M^{-1}s^{-1}$

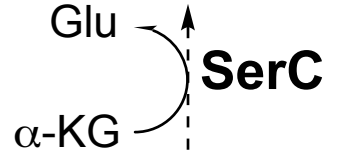



4-phosphooxy-L-threonine

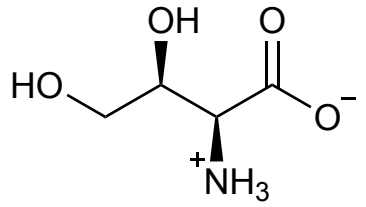
## SP4



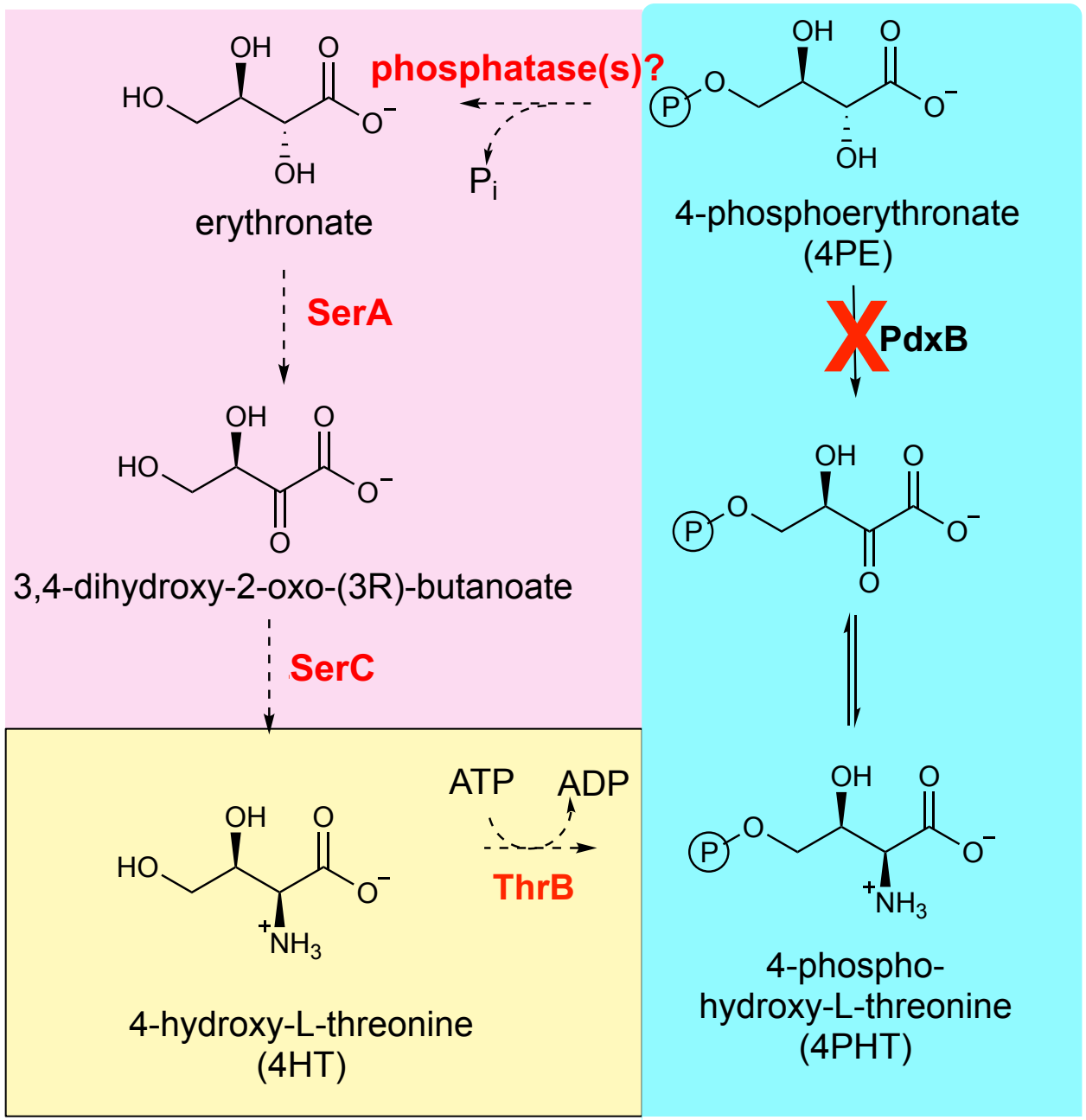

(3*R*)-3,4-dihydroxy-2-oxobutanoate




$k_{cat}/K_M$ = 0.007 $M^{-1}s^{-1}$




4-hydroxy-L-threonine
(4HT)


phosphatase(s)?
Pi
erythronate
SerA
3,4-dihydroxy-2-oxo-(3R)-butanoate
SerC
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

Questions?

1) How are the evolved strains making PLP?

**2) How do mutations improve PLP synthesis?**

# Mutations in evolved strains

**JK1**
*ybhA*
*pgl*
*gapA*
*rpoS*
*rpoC*

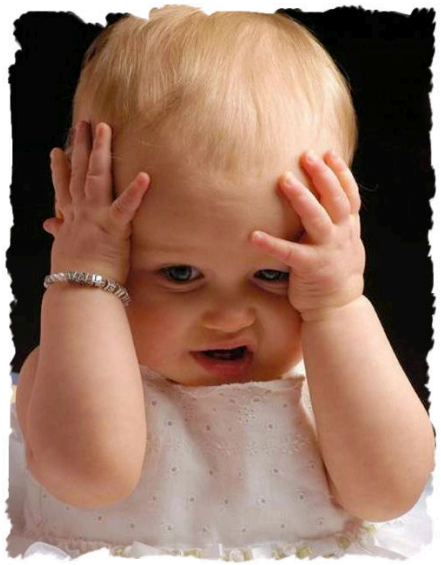
**JK2**
*ybhA*
*pgl*
*gapA*
*purF*
*gltB*
*ypjA*

**JK3**
*ybhA*
*pgl*
*gapA*
*ilvH*
*rng*

**JK4**
*ybhA*
*rpe*
*sdhA*
*rho*
*lon*

**JK5**
*ybhA*
*pgl*
*gapA*
*yjjK*
*purF*
*ilvH*
*nadR*

**JK6**
*gapA*
*serA*
*yjjK*



**JK7a**
*ybhA*
*gapA*
*purF*
*nadR*
*rpoS*

**JK7b**
*ybhA*
*pgl*
*serA*
*gapA*
*pykF*
*pyrE*

**JK8**
*gapA*
*serA*
*yjjK*
*gltB*
*livH*

**JK9**
*ybhA*
*pgl*
*gapA*
*serA*
*pykF*

**JK10**
*ybhA*
*pgl*
*gapA*
*rpe*
*ilvH*
*rng*

# Mutations in evolved strains

**JK1**
*ybhA*
*pgl*
*gapA*
*rpoS*
*rpoC*

**JK2**
*ybhA*
*pgl*
*gapA*
*purF*
*gltB*
*ypjA*

**JK3**
*ybhA*
*pgl*
*gapA*
*ilvH*
*rng*

**JK4**
*ybhA*
*rpe*
*sdhA*
*rho*
*lon*

**JK5**
*ybhA*
*pgl*
*gapA*
*yjjK*
*purF*
*ilvH*
*nadR*

**JK6**
*gapA*
*serA*
*yjjK*

**JK7a**
*ybhA*
*gapA*
*purF*
*nadR*
*rpoS*

**JK7b**
*ybhA*
*pgl*
*serA*
*gapA*
*pykF*
*pyrE*

**JK8**
*gapA*
*serA*
*yjjK*
*gltB*
*livH*

**JK9**
*ybhA*
*pgl*
*gapA*
*serA*
*pykF*

**JK10**
*ybhA*
*pgl*
*gapA*
*rpe*
*ilvH*
*rng*

# Mutations in *ybhA* cause loss of function

| JK1 | JK2 | JK3 | JK4 | JK5 | JK6 |
|---|---|---|---|---|---|
| *ybhA* | *ybhA* | *ybhA* | *ybhA* | *ybhA* | *gapA* |
| *pgl* | *pgl* | *pgl* | *rpe* | *pgl* | *serA* |
| *gapA* | *gapA* | *gapA* | *sdhA* | *gapA* | *yjjK* |
| *rpoS* | *purF* | *ilvH* | *rho* | *yjjK* | |
| *rpoC* | *gltB* | *rng* | *lon* | *purF* | |
| | *ypjA* | | | *ilvH* | |
| | | | | *nadR* | |

| JK7a | JK7b | JK8 | JK9 | JK10 |
|---|---|---|---|---|
| *ybhA* | *ybhA* | *gapA* | *ybhA* | *ybhA* |
| *gapA* | *pgl* | *serA* | *pgl* | *pgl* |
| *purF* | *serA* | *yjjK* | *gapA* | *gapA* |
| *nadR* | *gapA* | *gltB* | *serA* | *rpe* |
| *rpoS* | *pykF* | *livH* | *pykF* | *ilvH* |
| | *pyrE* | | | *rng* |

**Deletions**

**Premature stop codons**

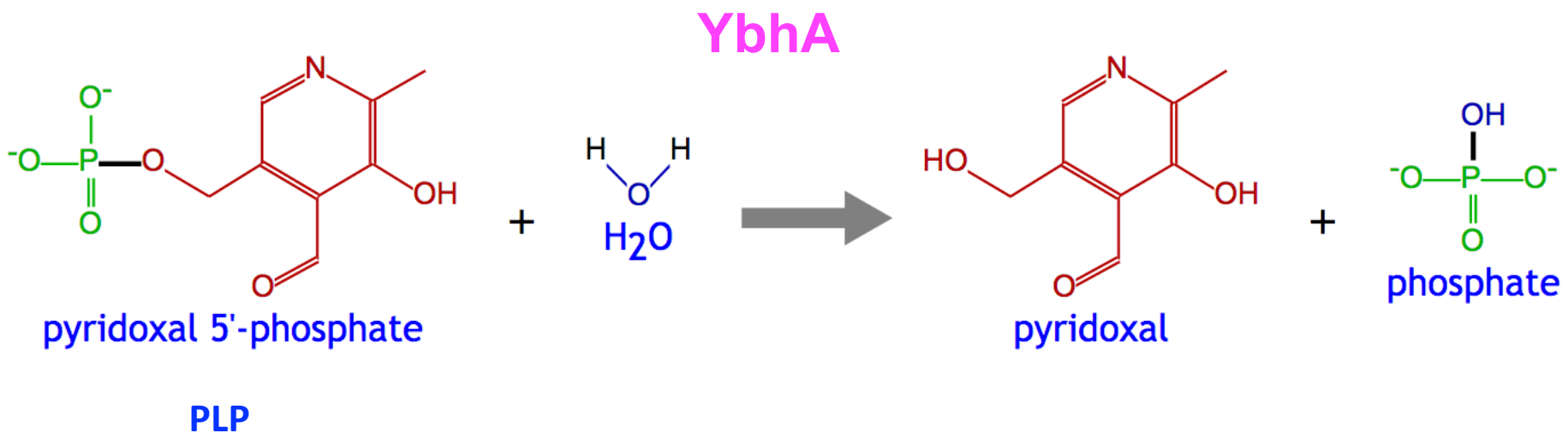
**Frameshift mutations**

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 47, pp. 36149–36161, November 24, 2006
Printed in the U.S.A.

# Genome-wide Analysis of Substrate Specificities of the *Escherichia coli* Haloacid Dehalogenase-like Phosphatase Family*[S]

Received for publication, June 7, 2006, and in revised form, September 21, 2006 Published, JBC Papers in Press, September 21, 2006, DOI 10.1074/jbc.M605449200

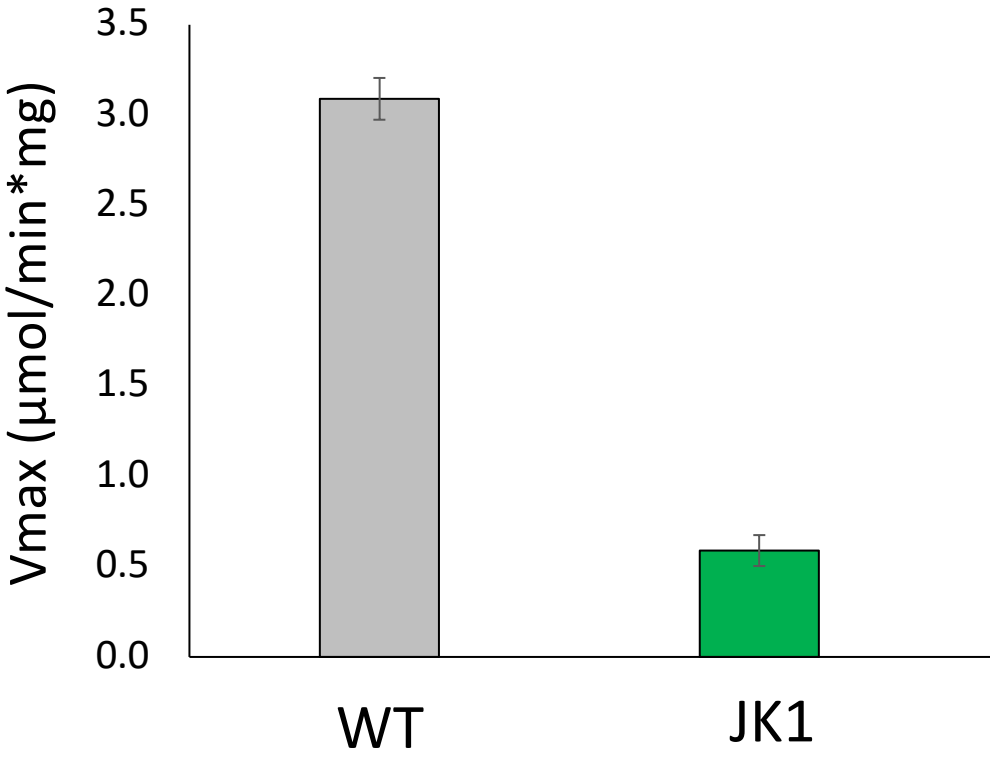
**Ekaterina Kuznetsova[‡§¶], Michael Proudfoot[‡], Claudio F. Gonzalez[‡], Greg Brown[‡], Marina V. Omelchenko[‖], Ivan Borozan[‡], Liran Carmel[‖], Yuri I. Wolf[‖], Hirotada Mori[**], Alexei V. Savchenko[‡§¶], Cheryl H. Arrowsmith[‡ ‡‡1], Eugene V. Koonin[‖], Aled M. Edwards[‡§¶2], and Alexander F. Yakunin[‡¶3]**



# Mutations in *gapA* decrease catalytic activity

| JK1 | JK2 | JK3 | JK4 | JK5 | JK6 |
|---|---|---|---|---|---|
| *ybhA* | *ybhA* | *ybhA* | *ybhA* | *ybhA* | *gapA* |
| *pgl* | *pgl* | *pgl* | *rpe* | *pgl* | *serA* |
| *gapA* | *gapA* | *gapA* | *sdhA* | *gapA* | *yjjK* |
| *rpoS* | *purF* | *ilvH* | *rho* | *yjjK* | |
| *rpoC* | *gltB* | *rng* | *lon* | *purF* | |
| | *ypjA* | | | *ilvH* | |
| | | | | *nadR* | |

| JK7a | JK7b | JK8 | JK9 | JK10 |
|---|---|---|---|---|
| *ybhA* | *ybhA* | *gapA* | *ybhA* | *ybhA* |
| *gapA* | *pgl* | *serA* | *pgl* | *pgl* |
| *purF* | *serA* | *yjjK* | *gapA* | *gapA* |
| *nadR* | *gapA* | *gltB* | *serA* | *rpe* |
| *rpoS* | *pykF* | *livH* | *pykF* | *ilvH* |
| | *pyrE* | | | *rng* |




(WT = strain with no mutations)

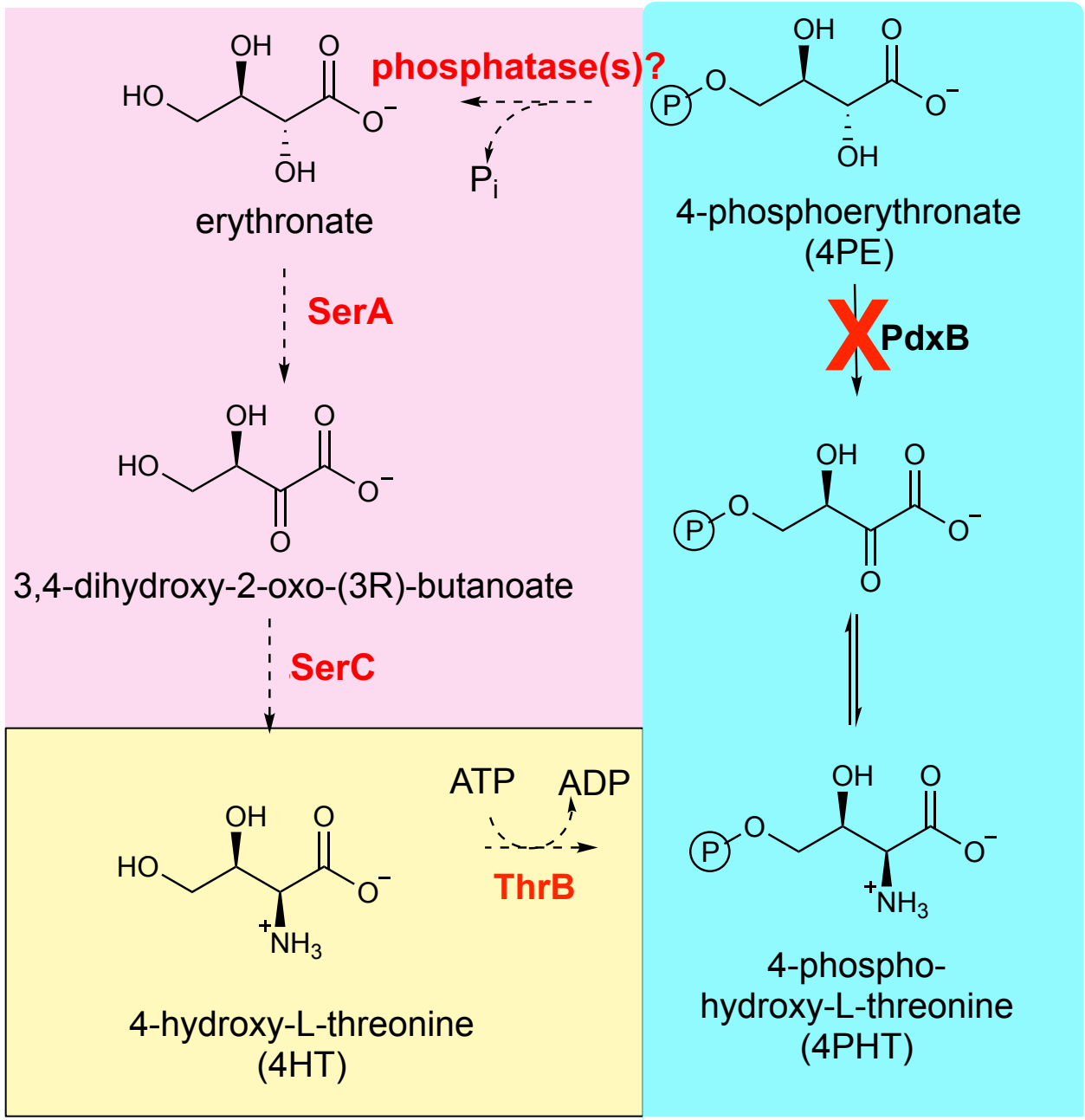
# Mutations in *ybhA* cause loss of function

| JK1 | JK2 | JK3 | JK4 | JK5 | JK6 |
|---|---|---|---|---|---|
| *ybhA* | *ybhA* | *ybhA* | *ybhA* | *ybhA* | *gapA* |
| *pgl* | *pgl* | *pgl* | *rpe* | *pgl* | *serA* |
| *gapA* | *gapA* | *gapA* | *sdhA* | *gapA* | *yjjK* |
| *rpoS* | *purF* | *ilvH* | *rho* | *yjjK* | |
| *rpoC* | *gltB* | *rng* | *lon* | *purF* | |
| | *ypjA* | | | *ilvH* | |
| | | | | *nadR* | |

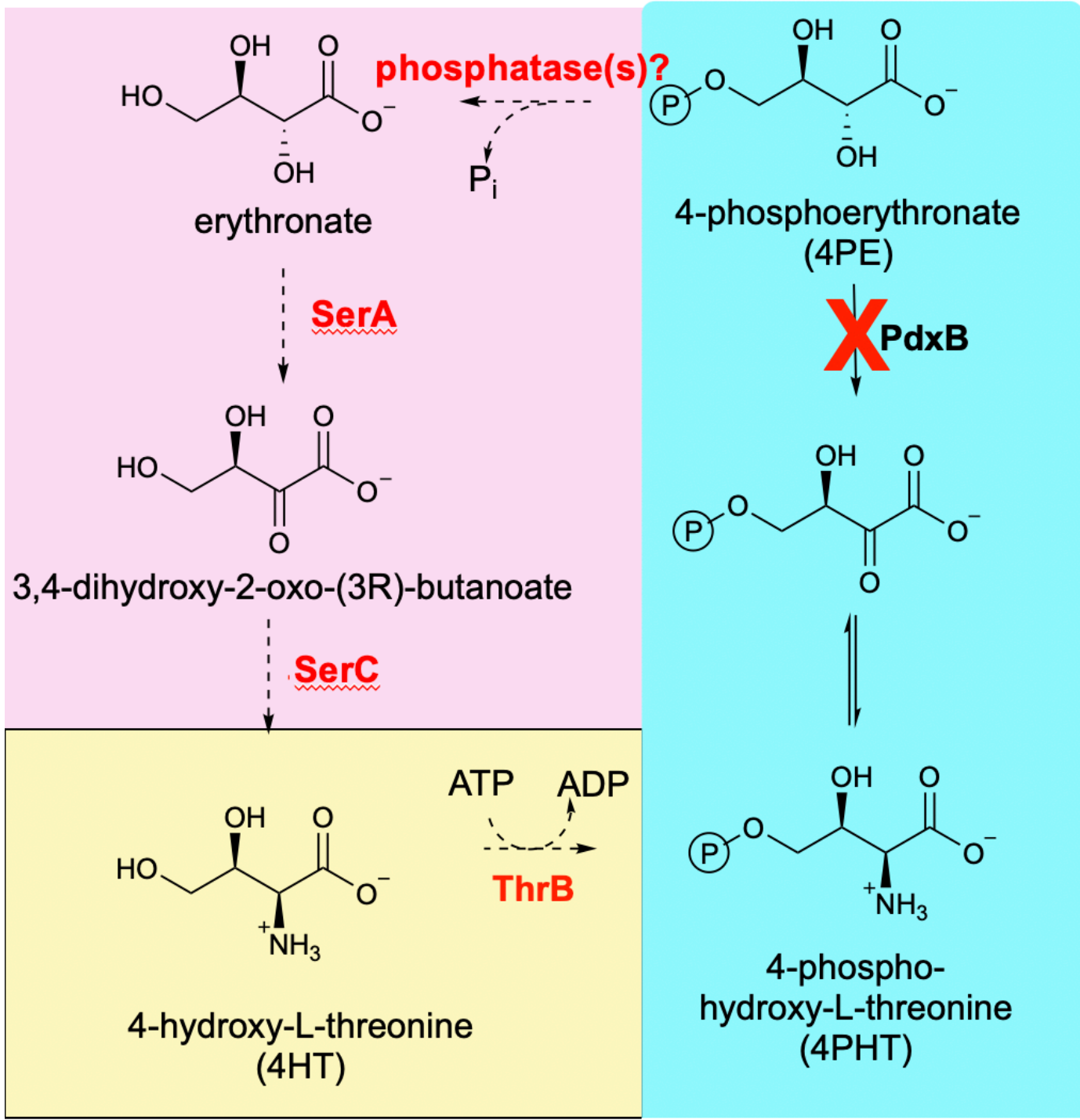
| JK7a | JK7b | JK8 | JK9 | JK10 |
|---|---|---|---|---|
| *ybhA* | *ybhA* | *gapA* | *ybhA* | *ybhA* |
| *gapA* | *pgl* | *serA* | *pgl* | *pgl* |
| *purF* | *serA* | *yjjK* | *gapA* | *gapA* |
| *nadR* | *gapA* | *gltB* | *serA* | *rpe* |
| *rpoS* | *pykF* | *livH* | *pykF* | *ilvH* |
| | *pyrE* | | | *rng* |

**Deletions**

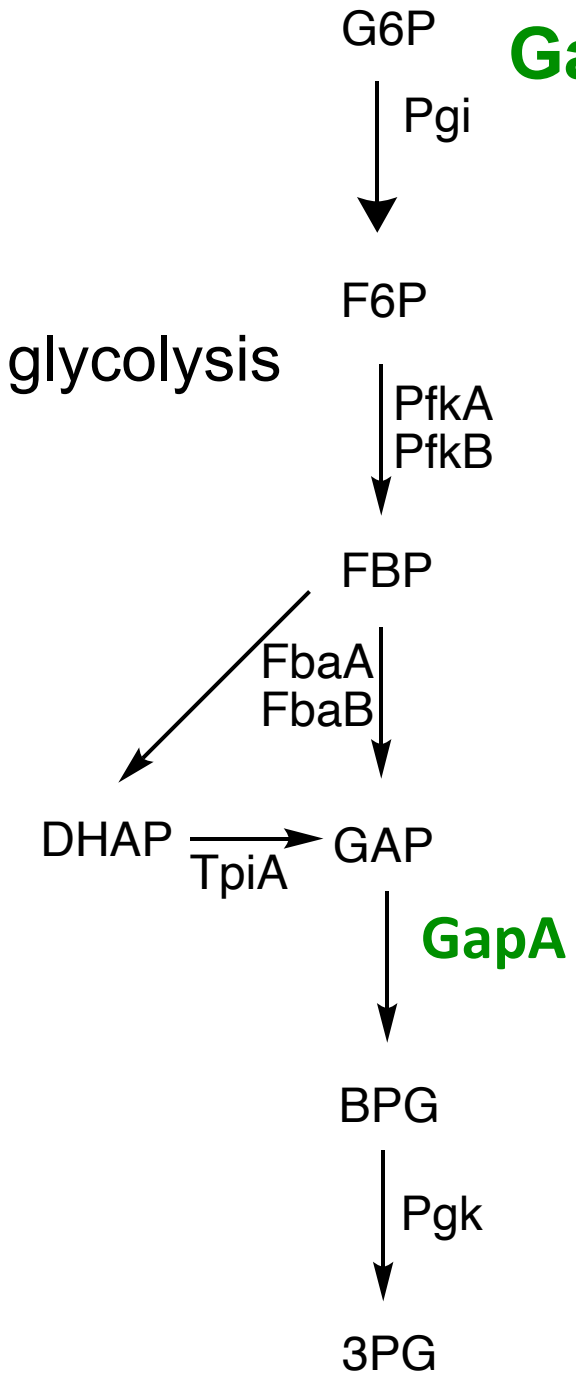
**Neither Pgl nor GapA catalyzes a reaction in SP4**

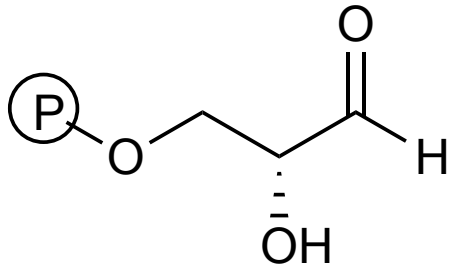


**Punchline: *pgl* and *gapA* mutations increase the ability of SerA to perform its new function in SP4**

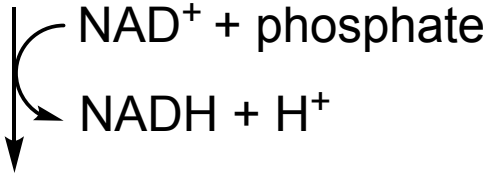


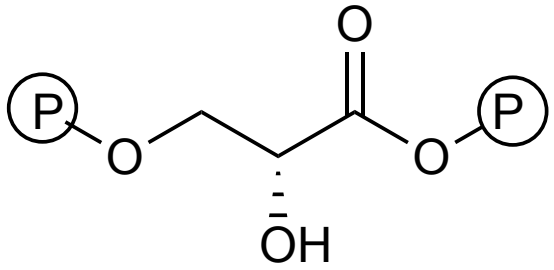
# GapA = D-glyceraldehyde 3-phosphate dehydrogenase




D-glyceraldehyde 3-phosphate (GAP)



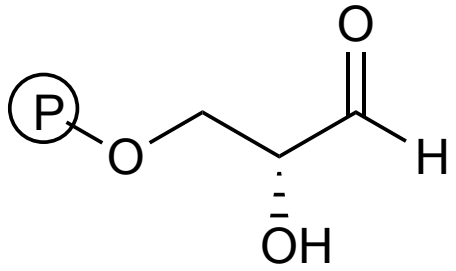




3-phospho-D-glyceroyl phosphate (BPG)

# GapA = D-glyceraldehyde 3-phosphate dehydrogenase

glycolysis

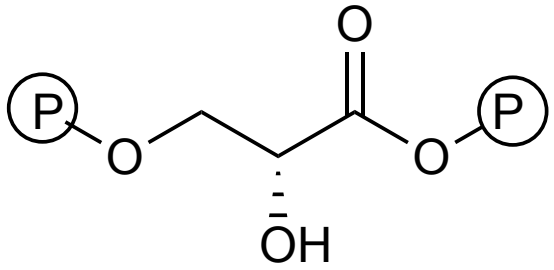
G6P
Pgi
F6P
PfkA
PfkB
FBP
FbaA
FbaB
DHAP
TpiA
GAP
GapA
BPG
Pgk
3PG




D-glyceraldehyde 3-phosphate (GAP)

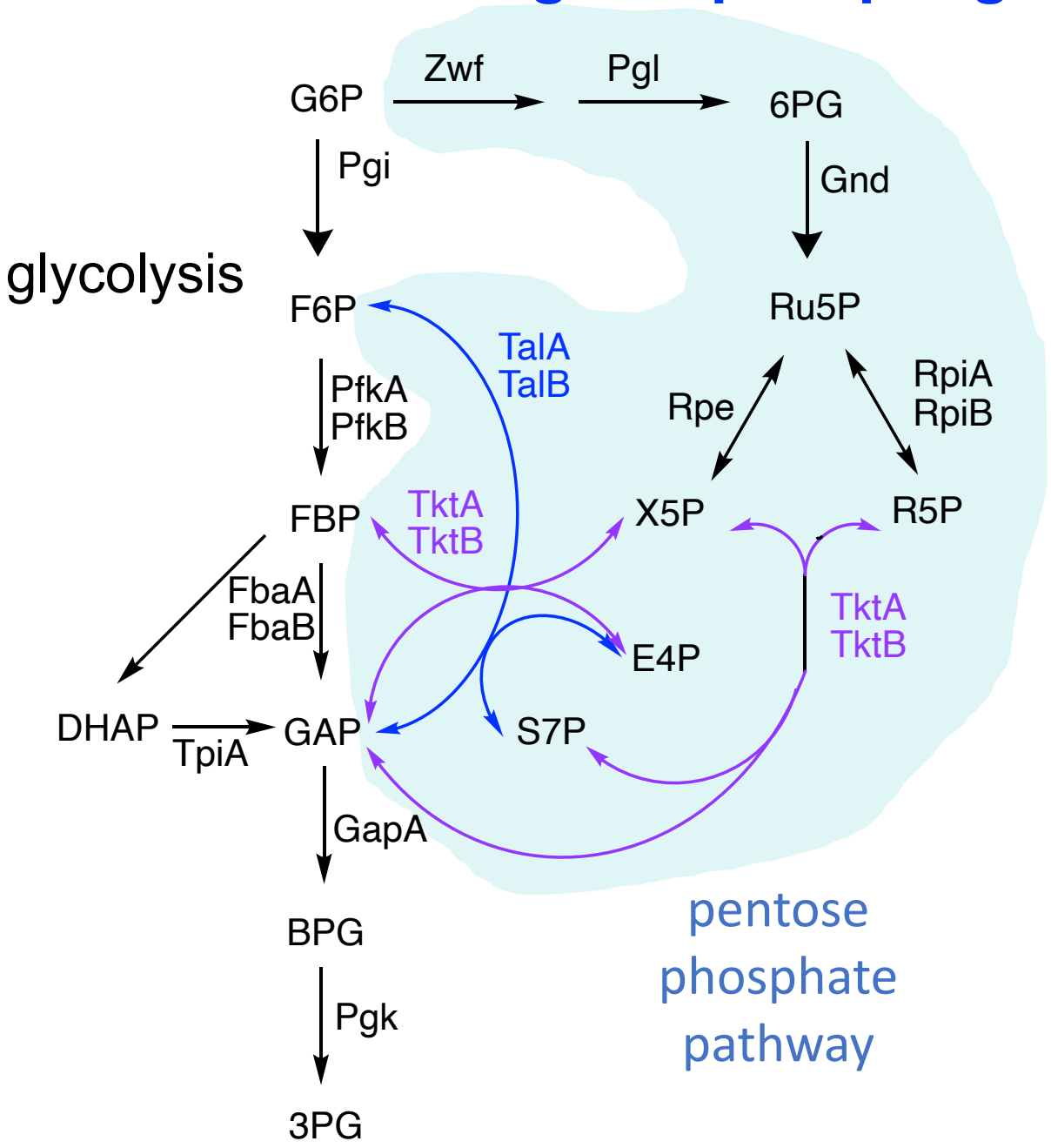
$NAD^+$ + phosphate

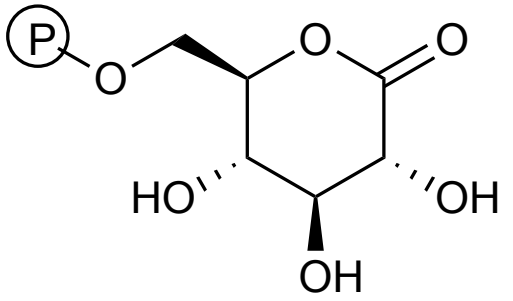
NADH + $H^+$



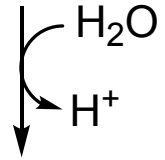

3-phospho-D-glyceroyl phosphate (BPG)

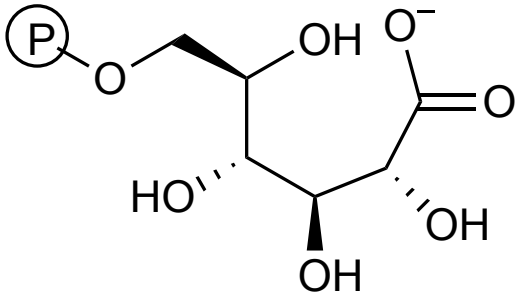
# Pgl = 6-phosphogluconolactonase



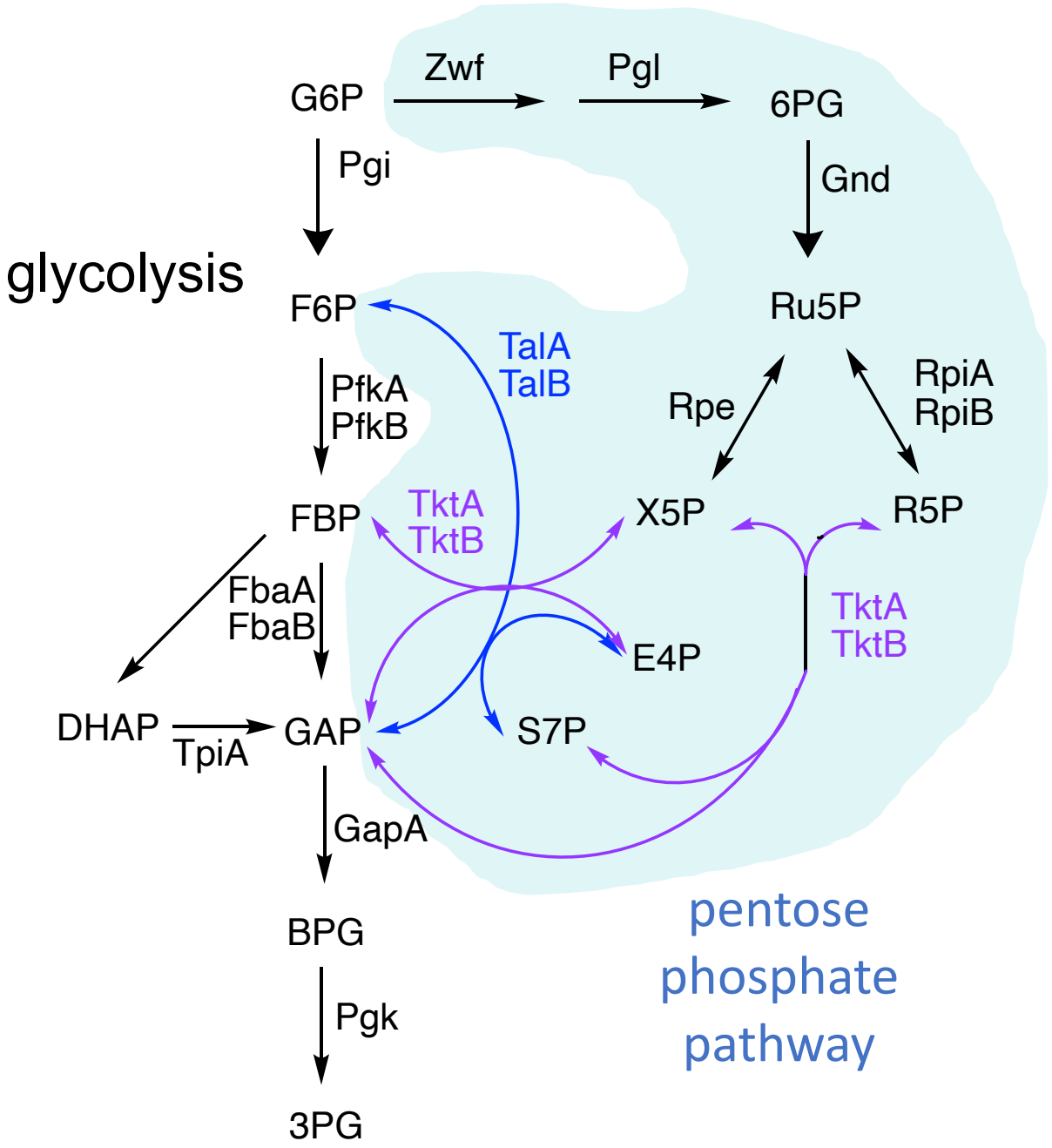




D-glucopyranose 6-phosphate



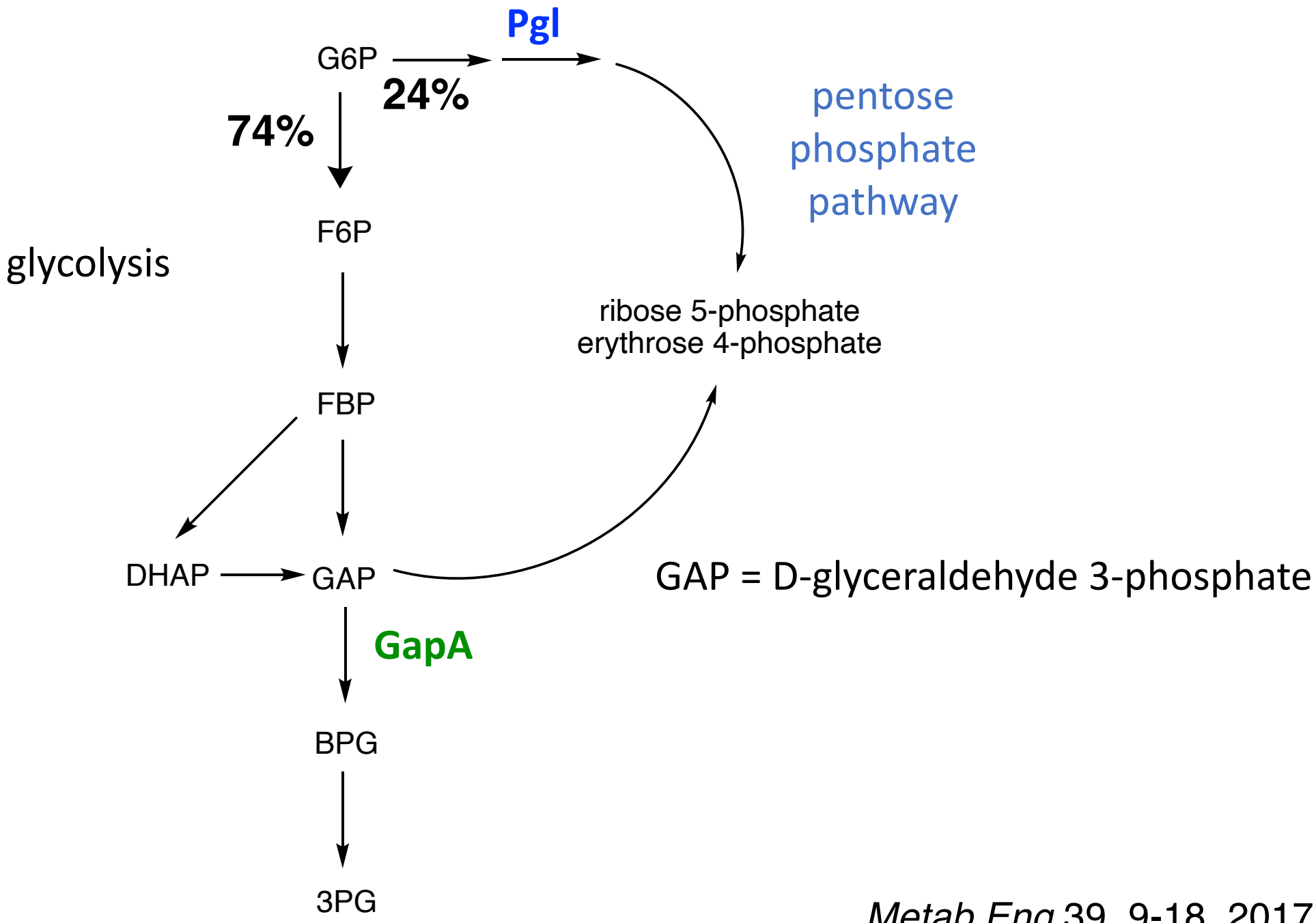




D-gluconate 6-phosphate

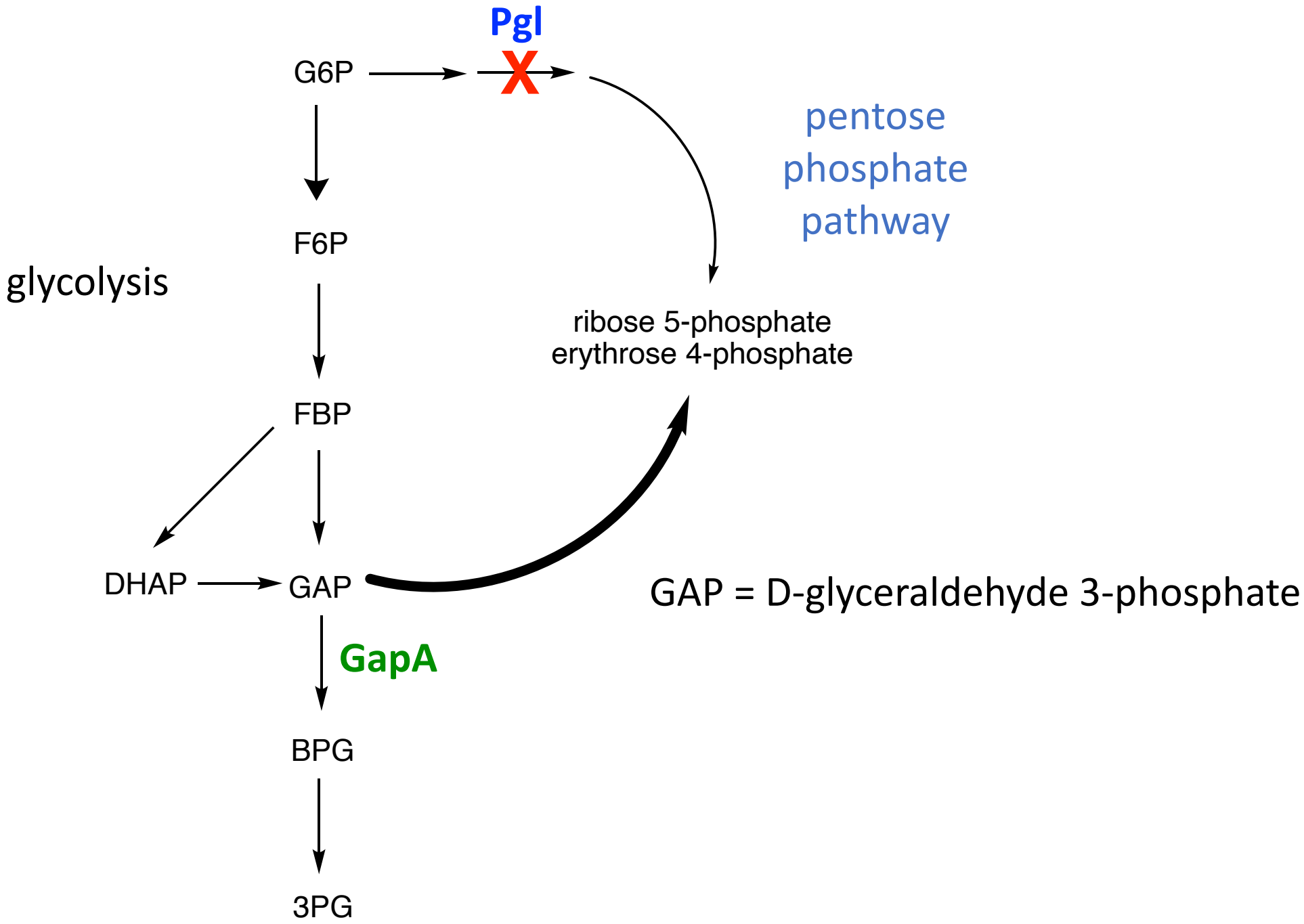


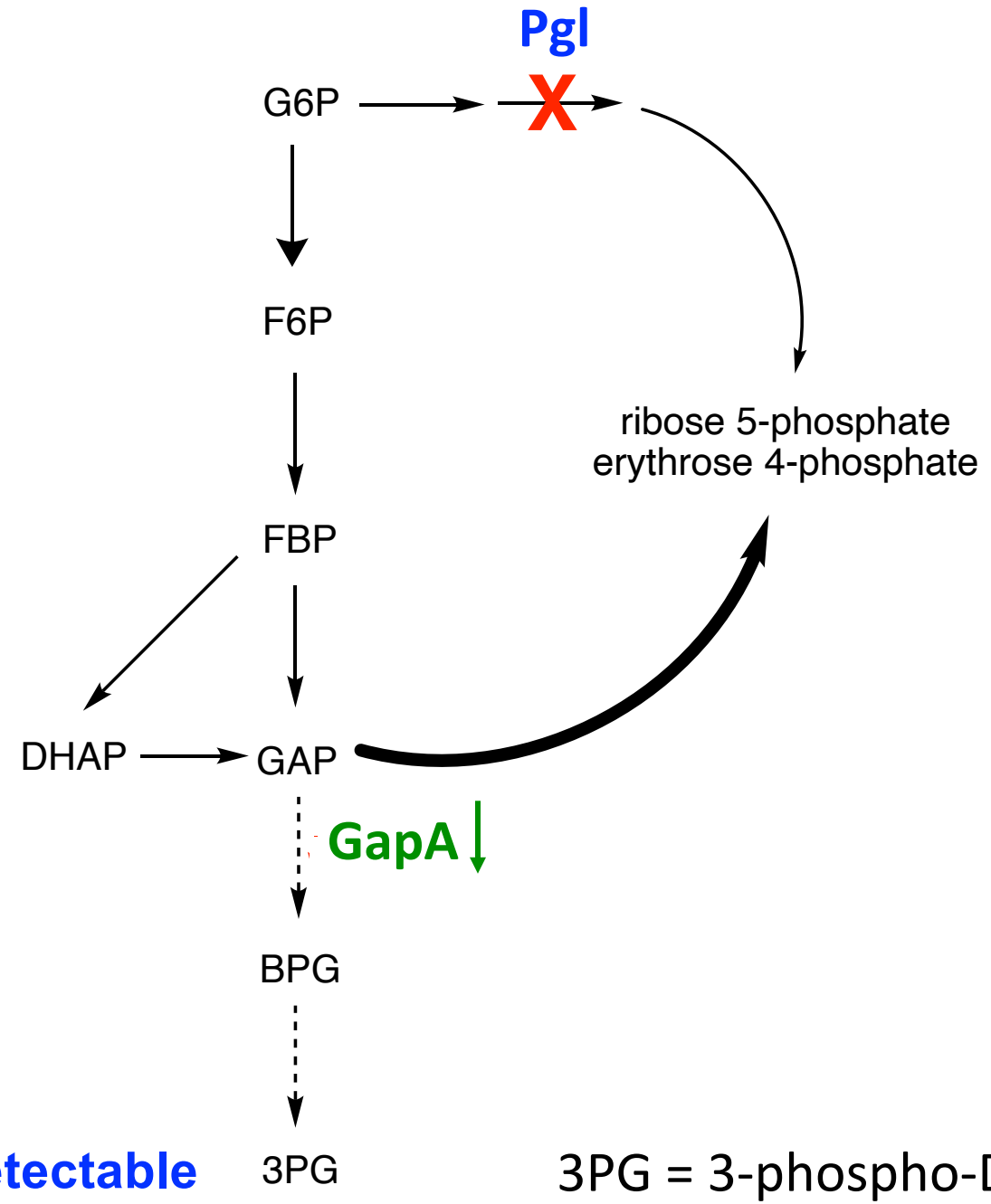

R5P = ribose 5-phosphate

E4P = erythrose 4-phosphate

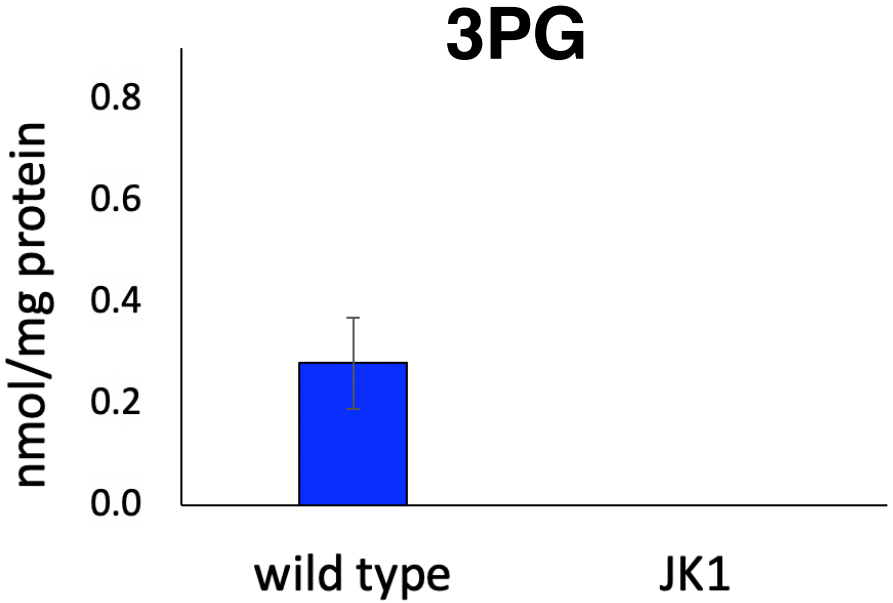



*Metab Eng* 39, 9-18, 2017


Pgl
G6P
pentose
phosphate
pathway
F6P
glycolysis
ribose 5-phosphate
erythrose 4-phosphate
FBP
DHAP
GAP
GAP = D-glyceraldehyde 3-phosphate
GapA
BPG
3PG



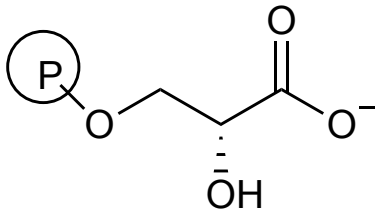

**undetectable**




3PG = 3-phospho-D-glycerate

# 3PG is a competitive inhibitor of the newly needed reaction

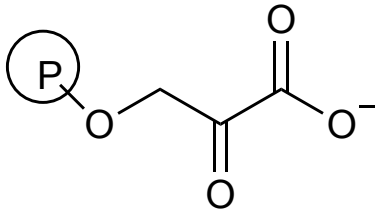
**serine biosynthesis**




3-phospho-D-glycerate

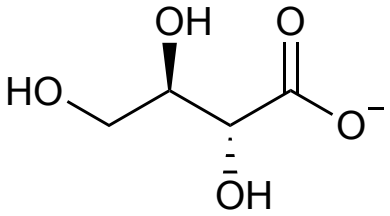
$k_{cat}/K_M$ = $1.6 \times 10^3\ M^{-1}s^{-1}$

**SerA**




3-phosphooxypyruvate

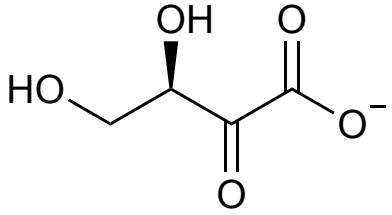
**SP4**



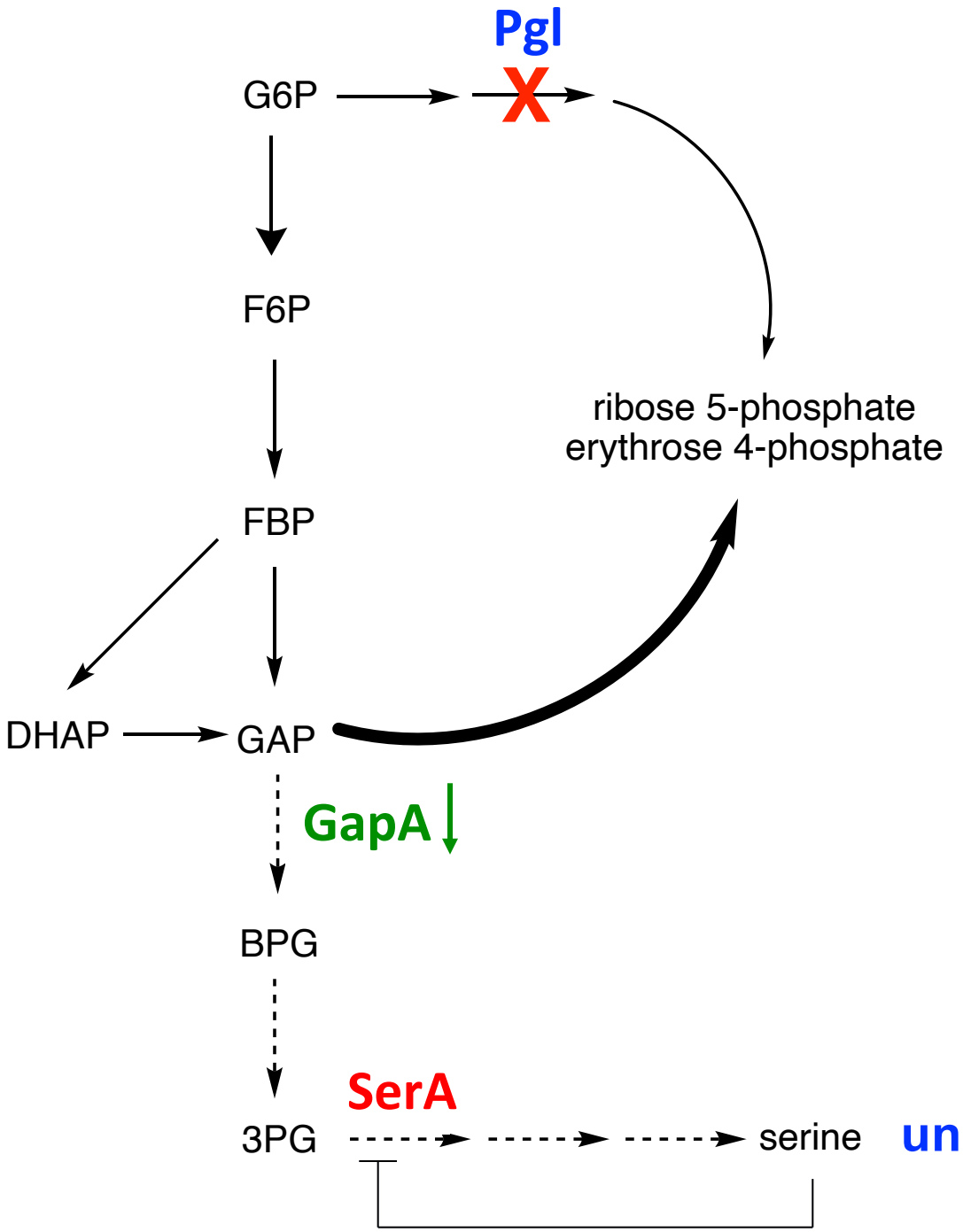

D-erythronate

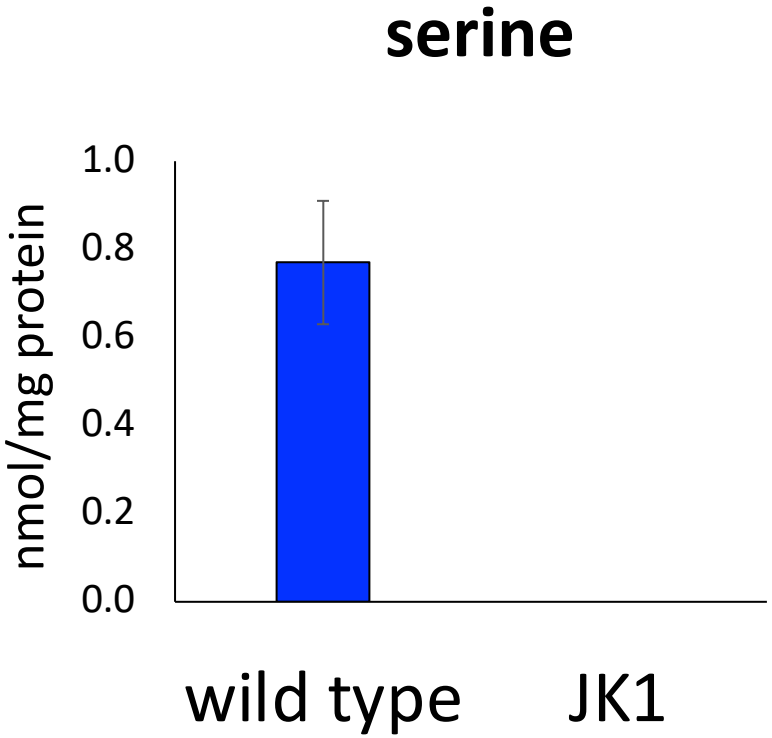
$k_{cat}/K_M$ = $0.07\ M^{-1}s^{-1}$

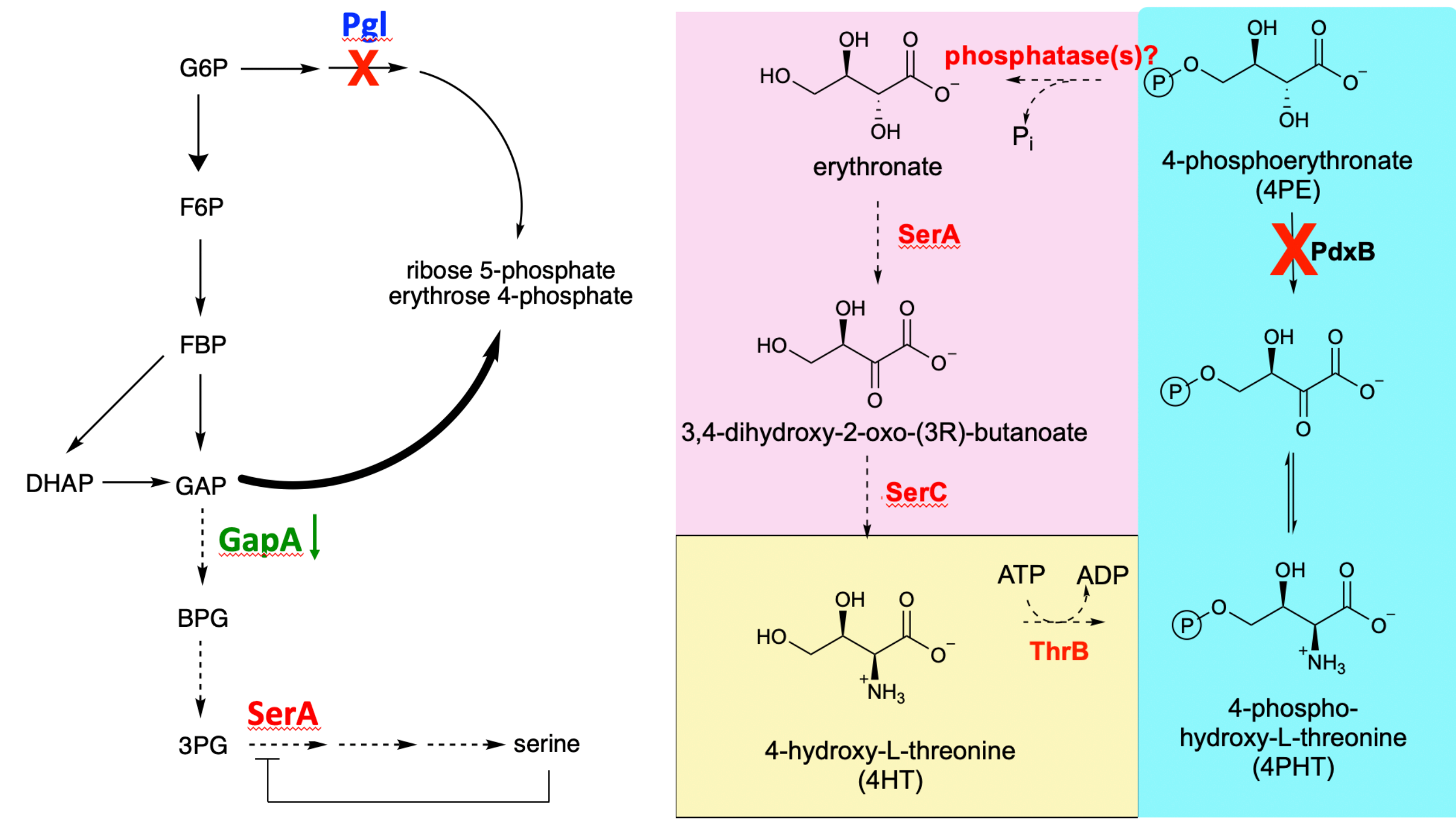
**SerA**




(3*R*)-3,4-dihydroxy-2-oxobutanoate


Pgl
G6P
F6P
FBP
ribose 5-phosphate
erythrose 4-phosphate
DHAP
GAP
GapA
BPG
SerA
3PG
serine
undetectable


serine
nmol/mg protein
1.0
0.8
0.6
0.4
0.2
0.0
wild type
JK1

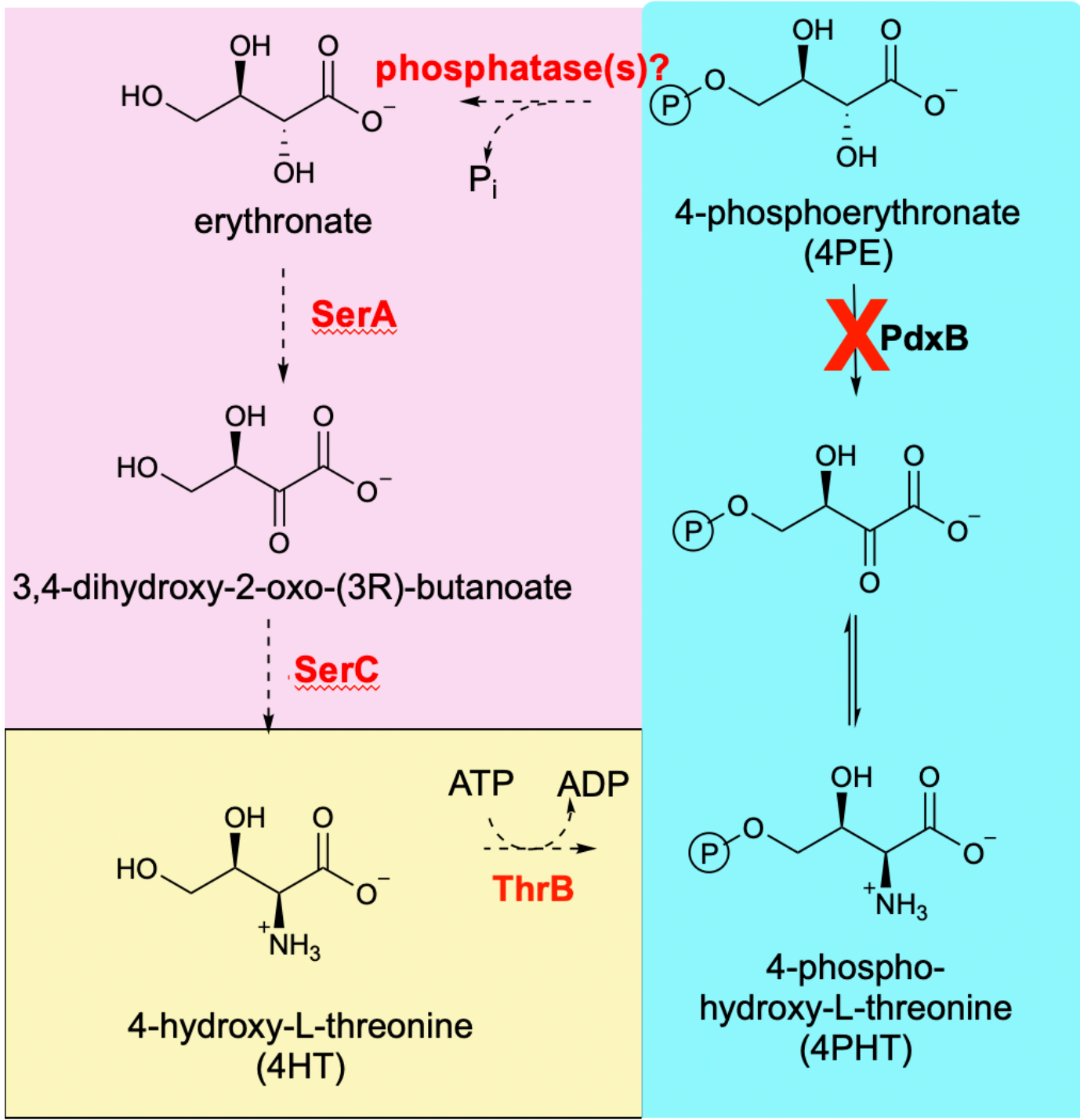

Pgl
G6P
F6P
FBP
DHAP
GAP
BPG
3PG
serine
ribose 5-phosphate
erythrose 4-phosphate
GapA
SerA
phosphatase(s)?
Pi
erythronate
SerA
3,4-dihydroxy-2-oxo-(3R)-butanoate
SerC
4-hydroxy-L-threonine
(4HT)
ATP
ADP
ThrB
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

**Punchline: *pgl* and *gapA* mutations increase the ability of SerA to perform its new function in SP4 by**

**1) decreasing inhibition by the native substrate (3PG)**

***and***

**2) decreasing feedback inhibition by serine**



# *serA* mutations

**JK1**
*ybhA*
*pgl*
*gapA*
*rpoS*
*rpoC*

**JK2**
*ybhA*
*pgl*
*gapA*
*purF*
*gltB*
*ypjA*

**JK3**
*ybhA*
*pgl*
*gapA*
*ilvH*
*rng*

**JK4**
*ybhA*
*rpe*
*sdhA*
*rho*
*lon*

**JK5**
*ybhA*
*pgl*
*gapA*
*yjjK*
*purF*
*ilvH*
*nadR*

**JK6**
*gapA*
*serA*
*yjjK*

**JK7a**
*ybhA*
*gapA*
*purF*
*nadR*
*rpoS*

**JK7b**
*ybhA*
*pgl*
*serA*
*gapA*
*pykF*
*pyrE*

**JK8**
*gapA*
*serA*
*yjjK*
*gltB*
*livH*

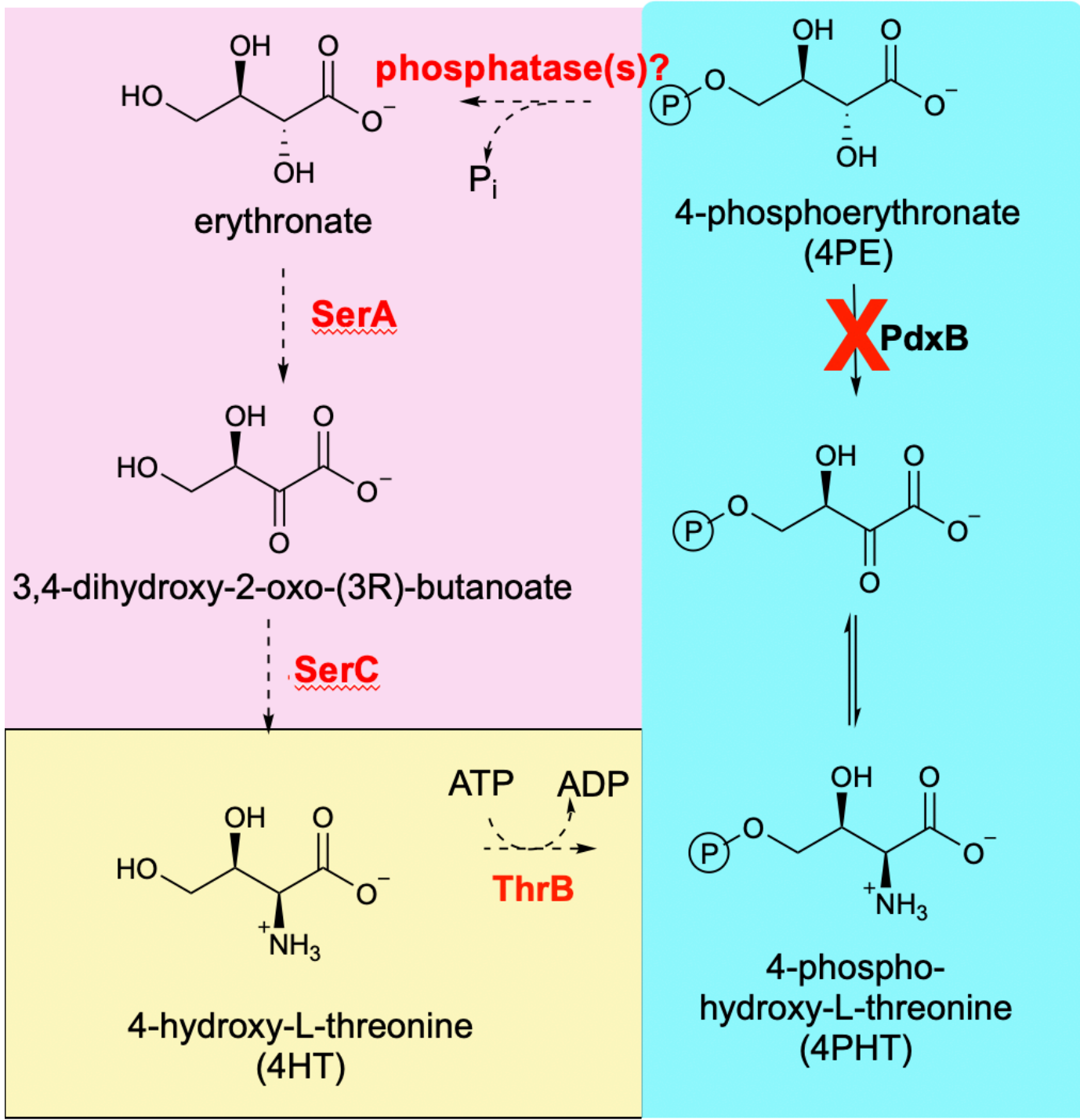
**JK9**
*ybhA*
*pgl*
*gapA*
*serA*
*pykF*

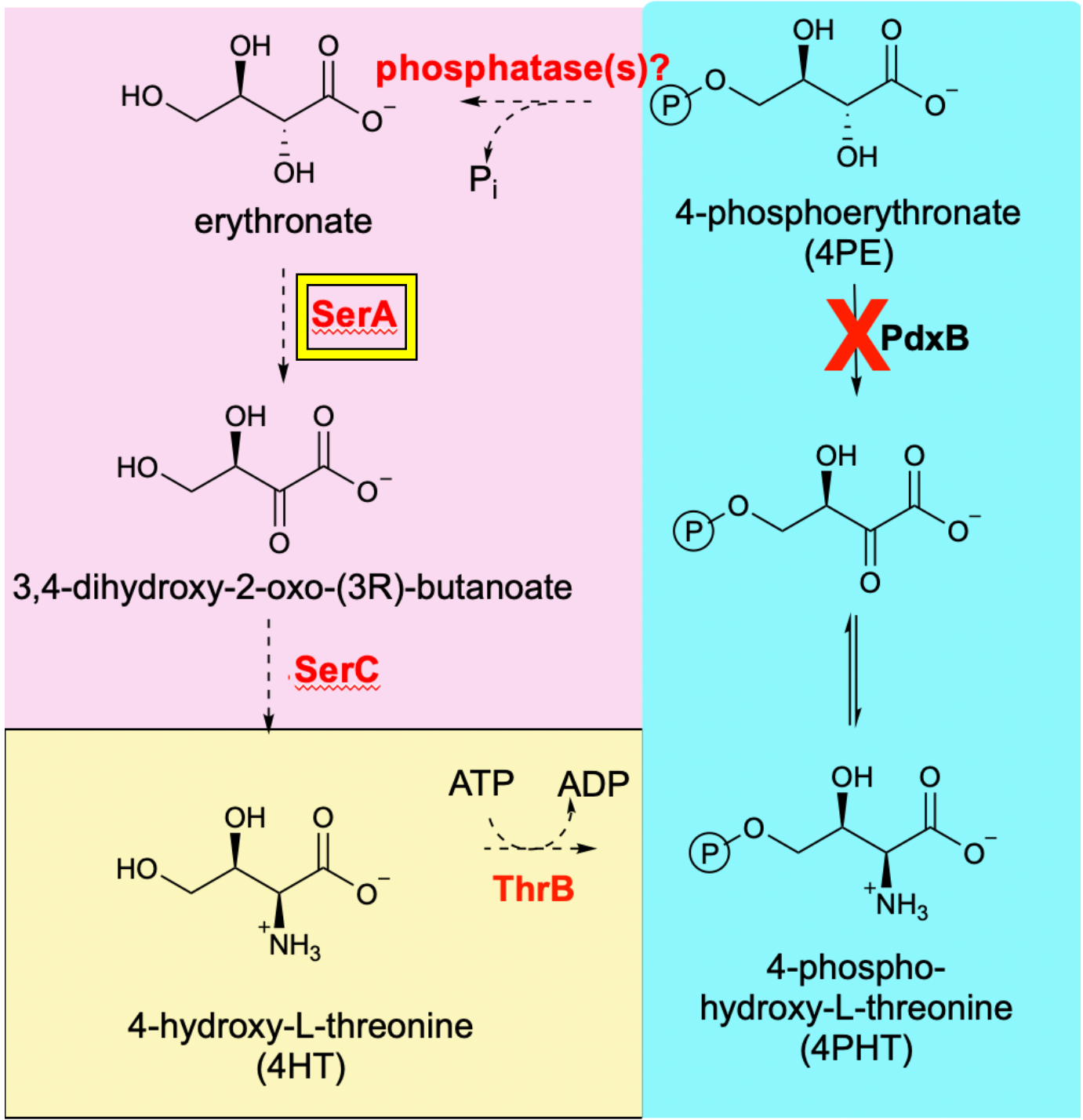
**JK10**
*ybhA*
*pgl*
*gapA*
*rpe*
*ilvH*
*rng*

**Punchline: *serA* mutations increase the ability of SerA to perform its new function in SP4 by**

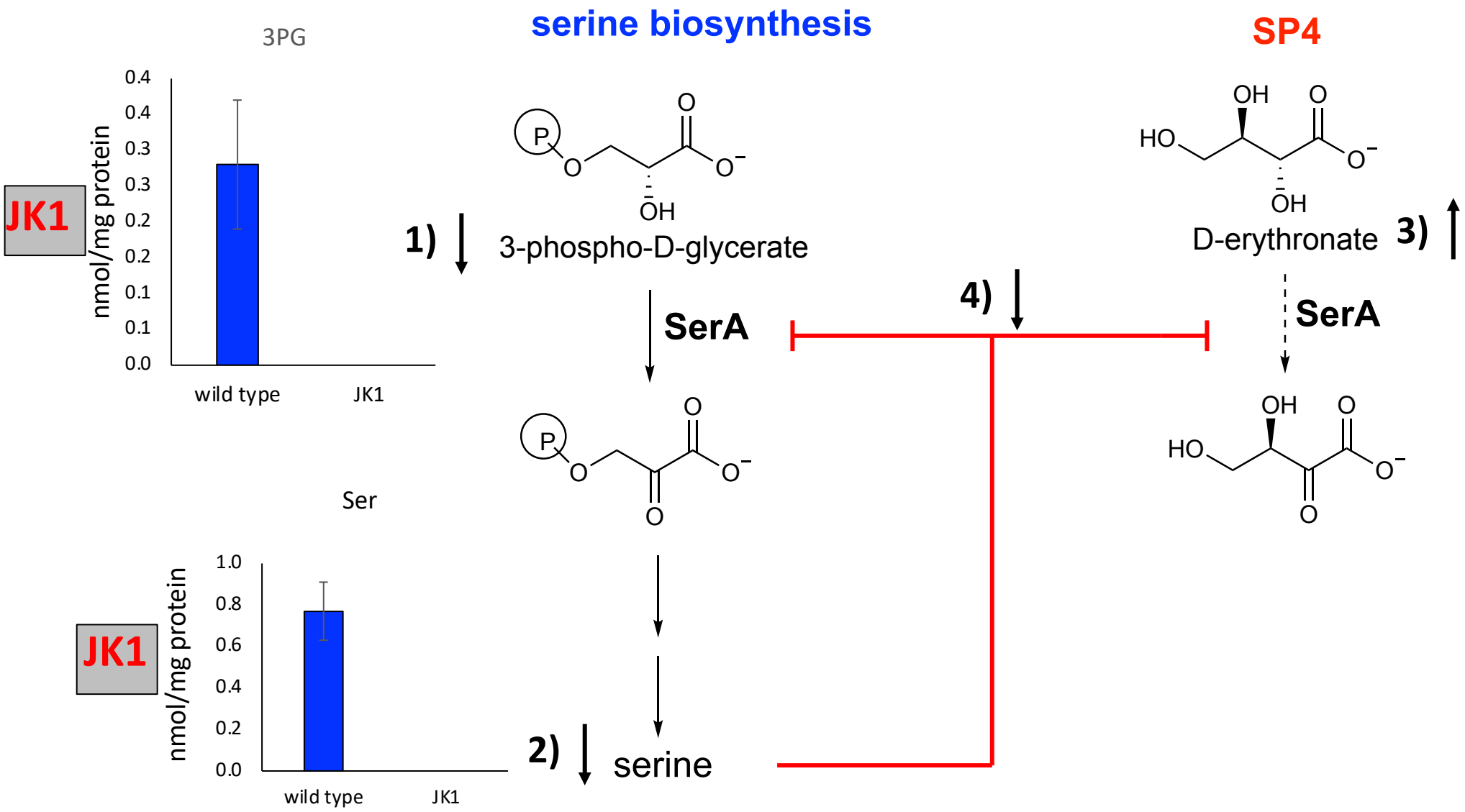
**1) decreasing feedback inhibition by serine**

**2) improving new activity**

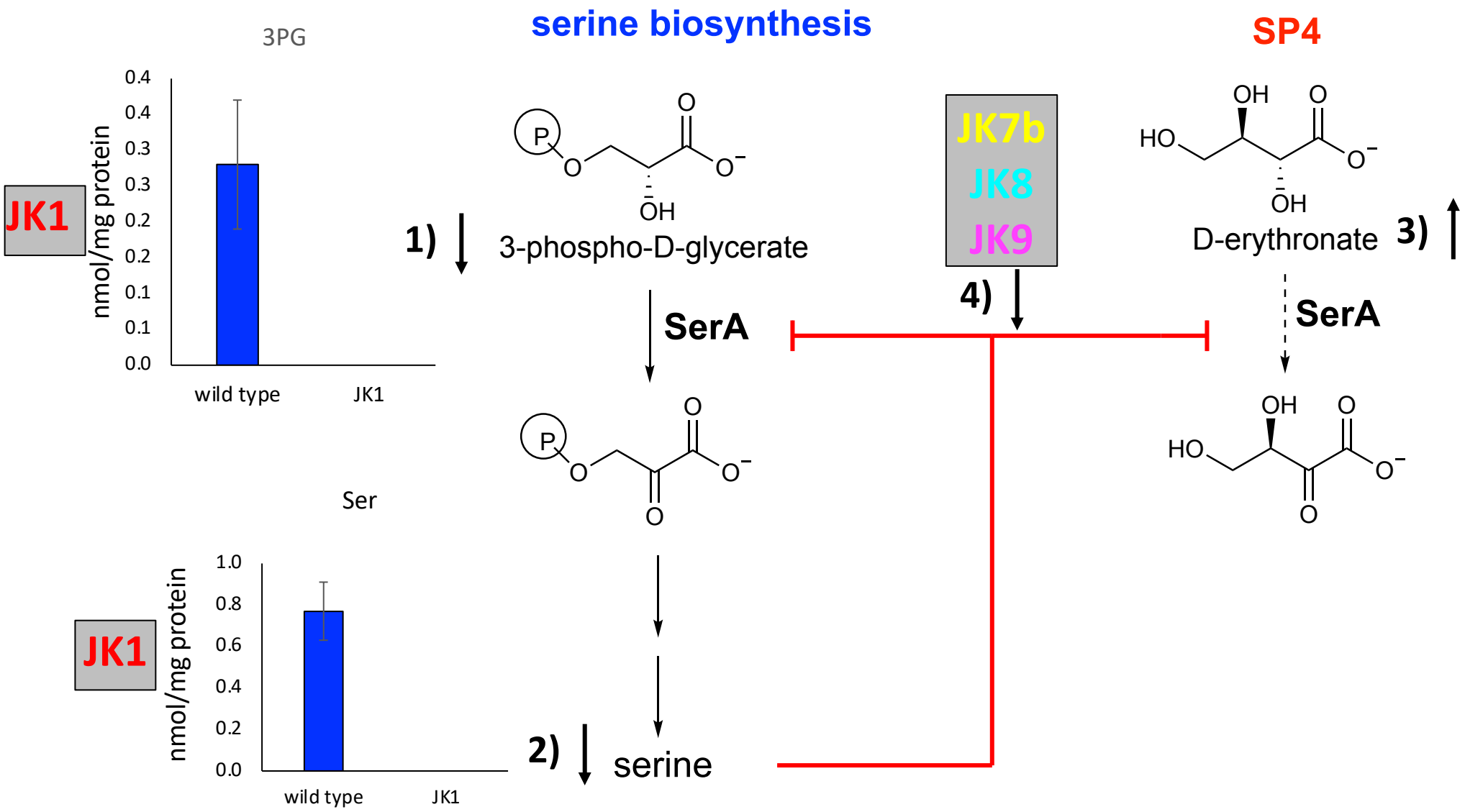



phosphatase(s)?
$P_i$
erythronate
SerA
3,4-dihydroxy-2-oxo-(3R)-butanoate
SerC
ATP
ADP
ThrB
4-hydroxy-L-threonine
(4HT)
4-phosphoerythronate
(4PE)
PdxB
4-phospho-
hydroxy-L-threonine
(4PHT)

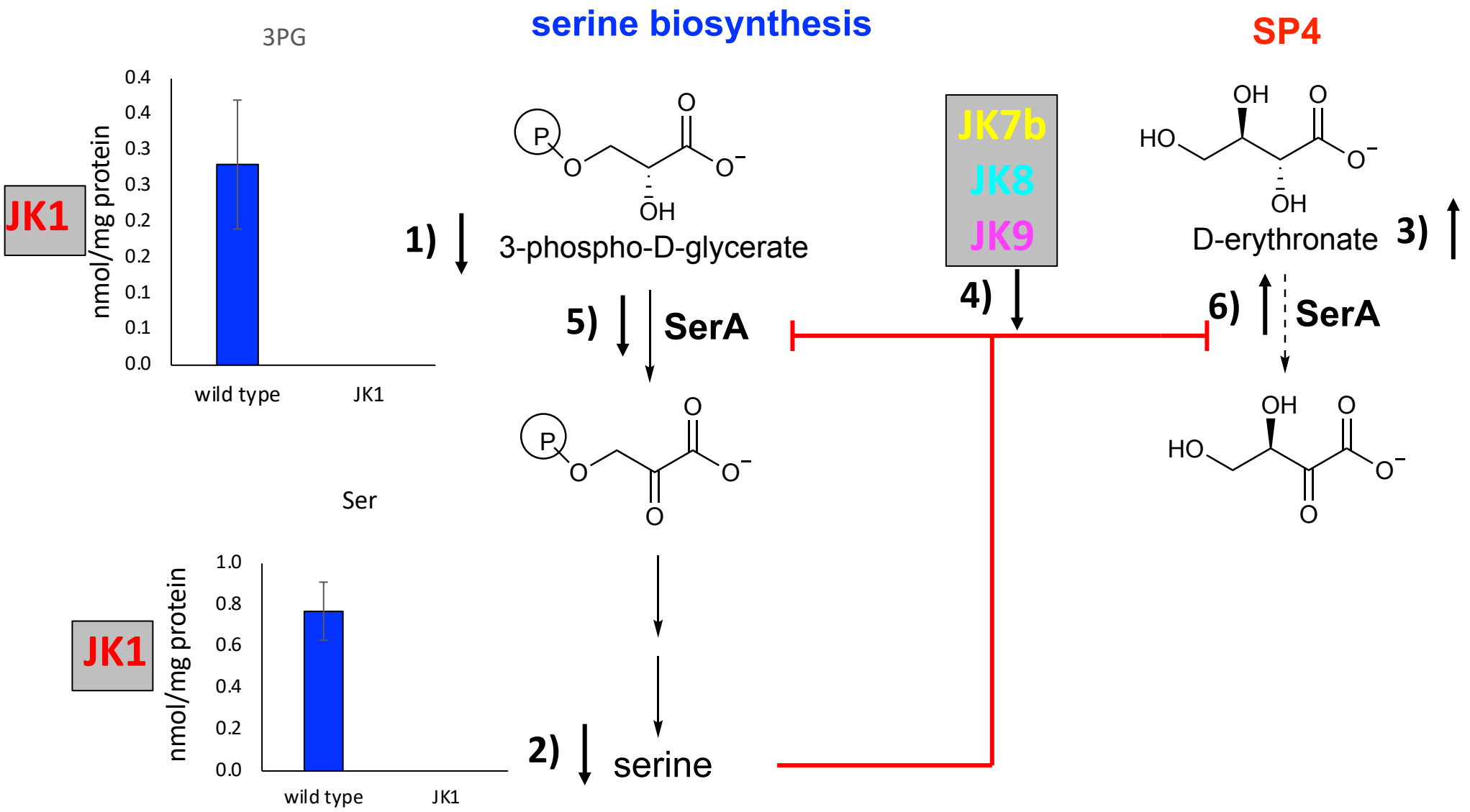
# Seven (!) ways to improve oxidation of erythronate



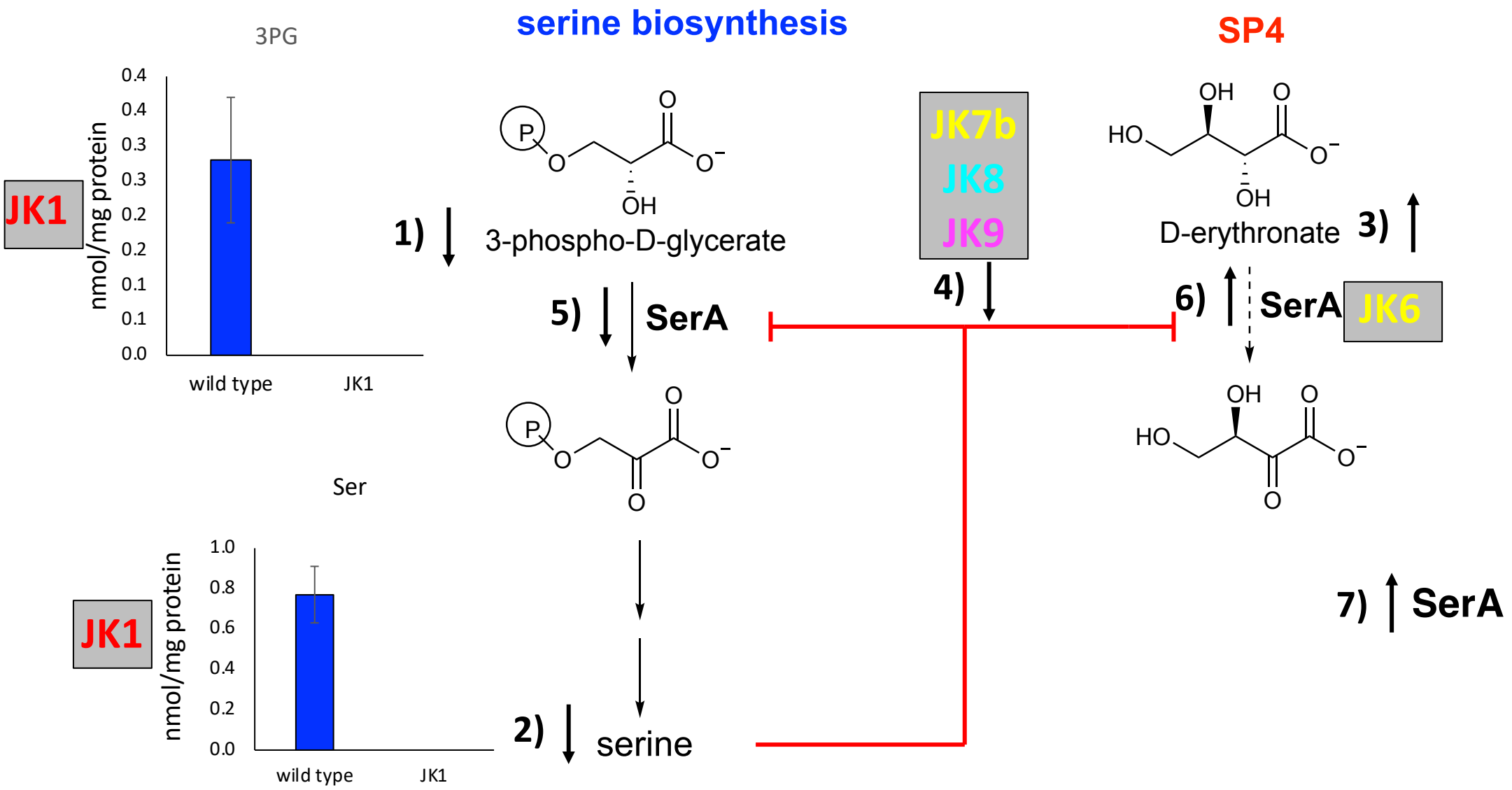
# Seven (!) ways to improve oxidation of erythronate



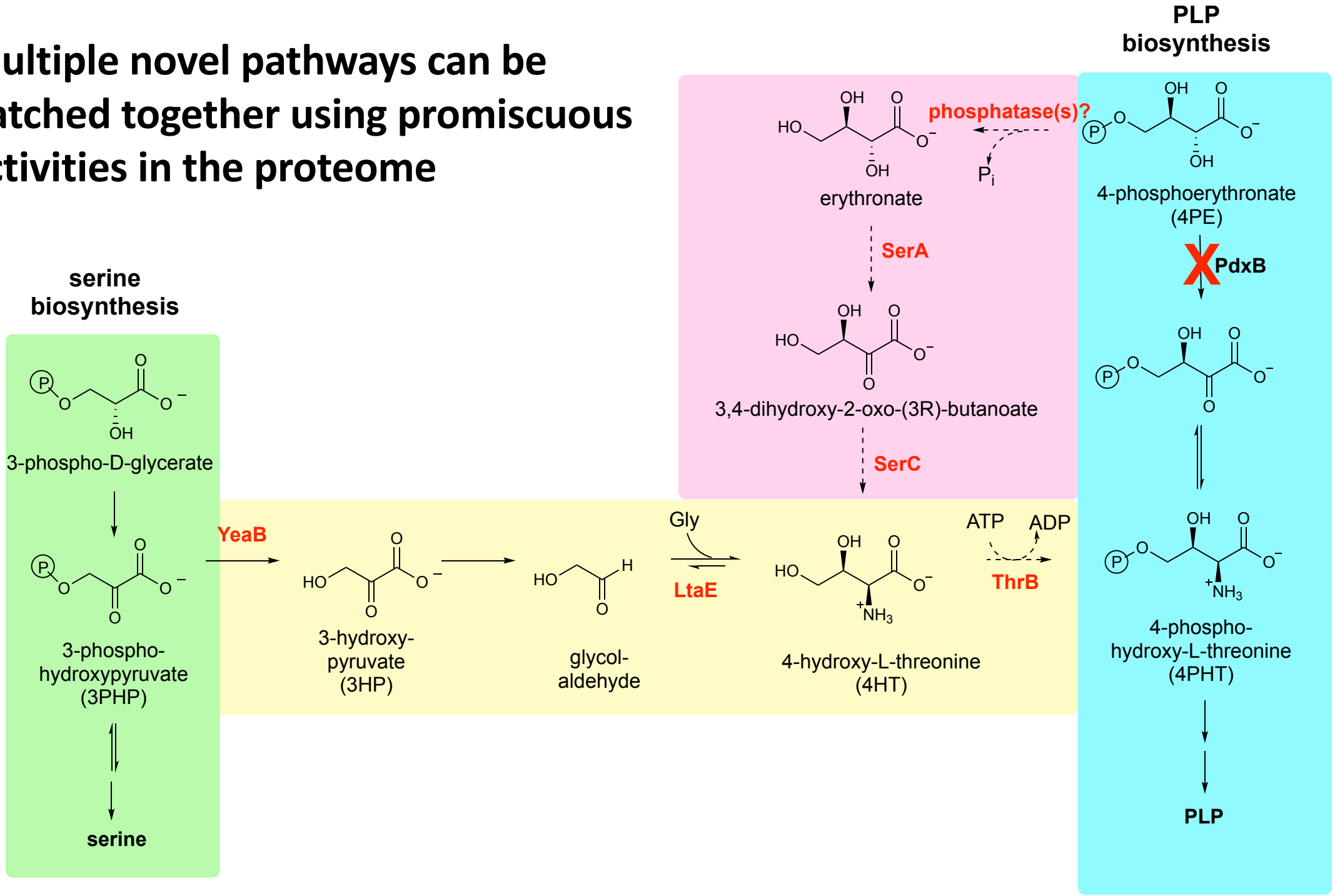
# Seven (!) ways to improve oxidation of erythronate



# Seven (!) ways to improve oxidation of erythronate



# Multiple novel pathways can be patched together using promiscuous activities in the proteome

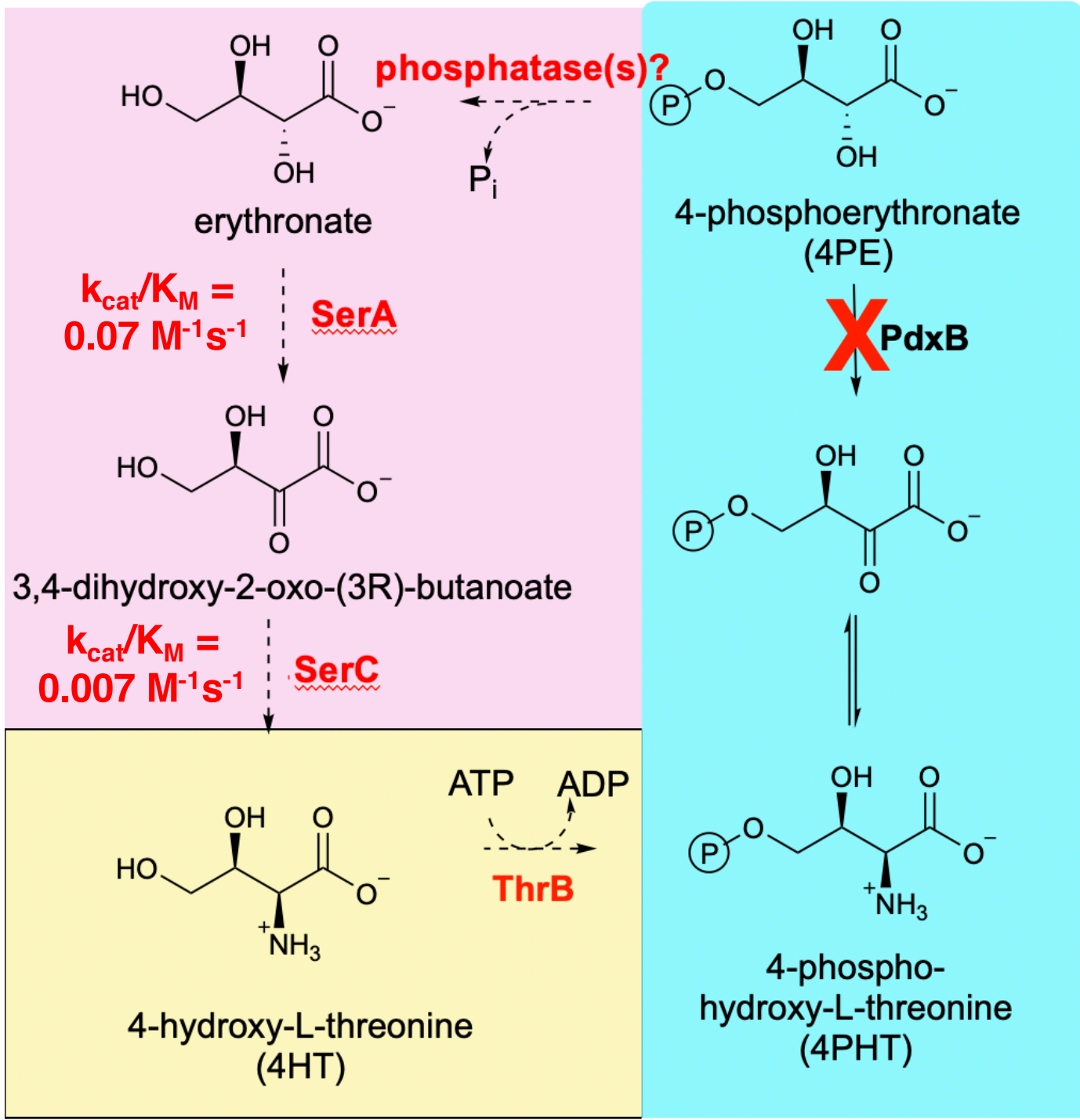


**Mutations that elevate flux through a novel pathway need not occur in genes encoding enzymes in the pathway**

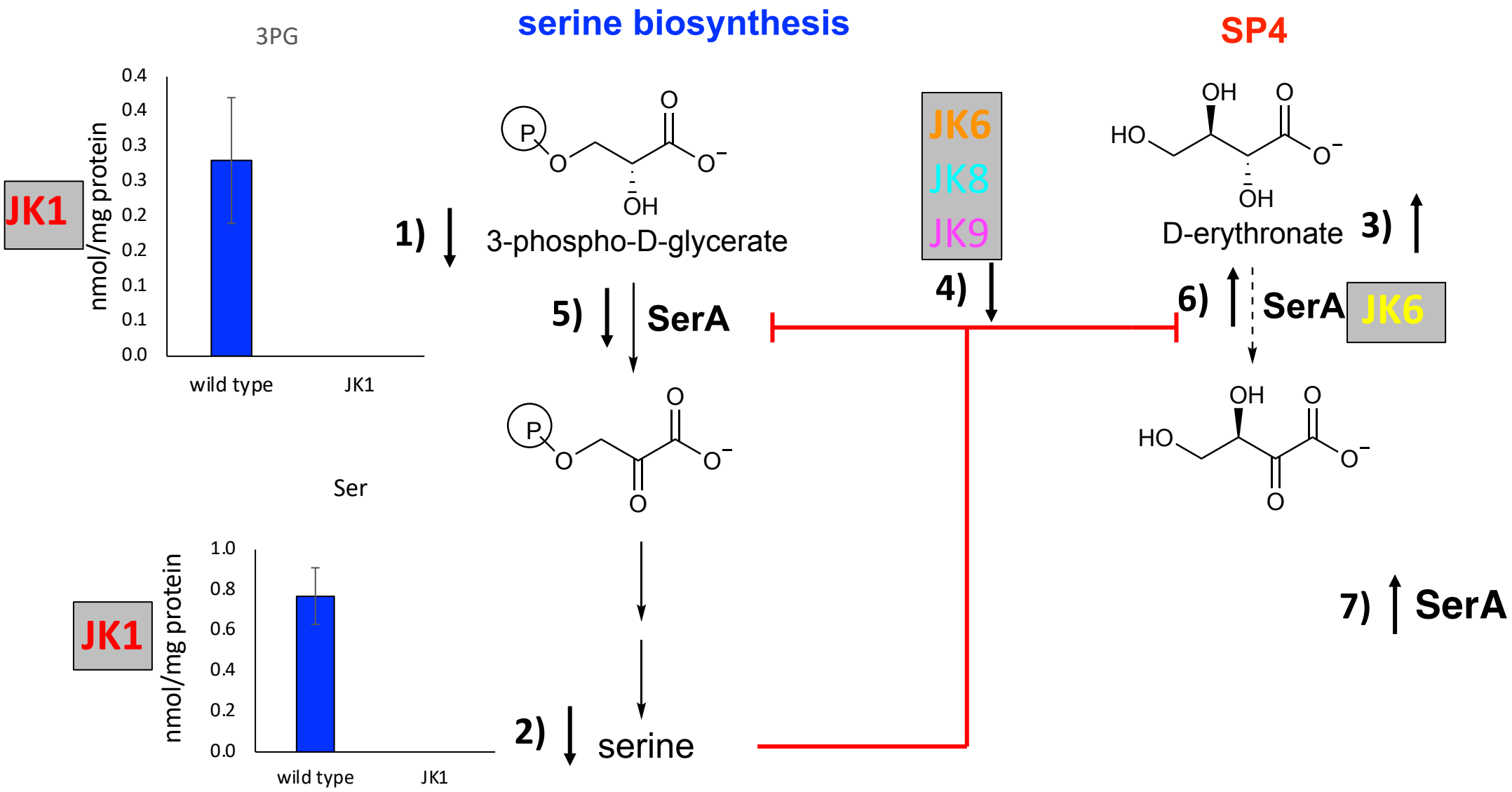
| JK1 | JK2 | JK3 | JK4 | JK5 | JK6 |
|---|---|---|---|---|---|
| *ybhA/pgl* | *ybhA/pgl* | *ybhA/pgl* | *ybhA* | *ybhA/pgl* | *gapA* |
| *gapA* | *gapA* | *gapA* | *rpe* | *gapA* | ***serA*** |
| *rpoS* | *purF* | *ilvH* | *sdhA* | *yjjK* | *yjjK* |
| *rpoC* | *gltB* | *rng* | *rho* | *purF* | |
| | *ypjA* | | *lon* | *ilvH* | |
| | | | | *nadR* | |

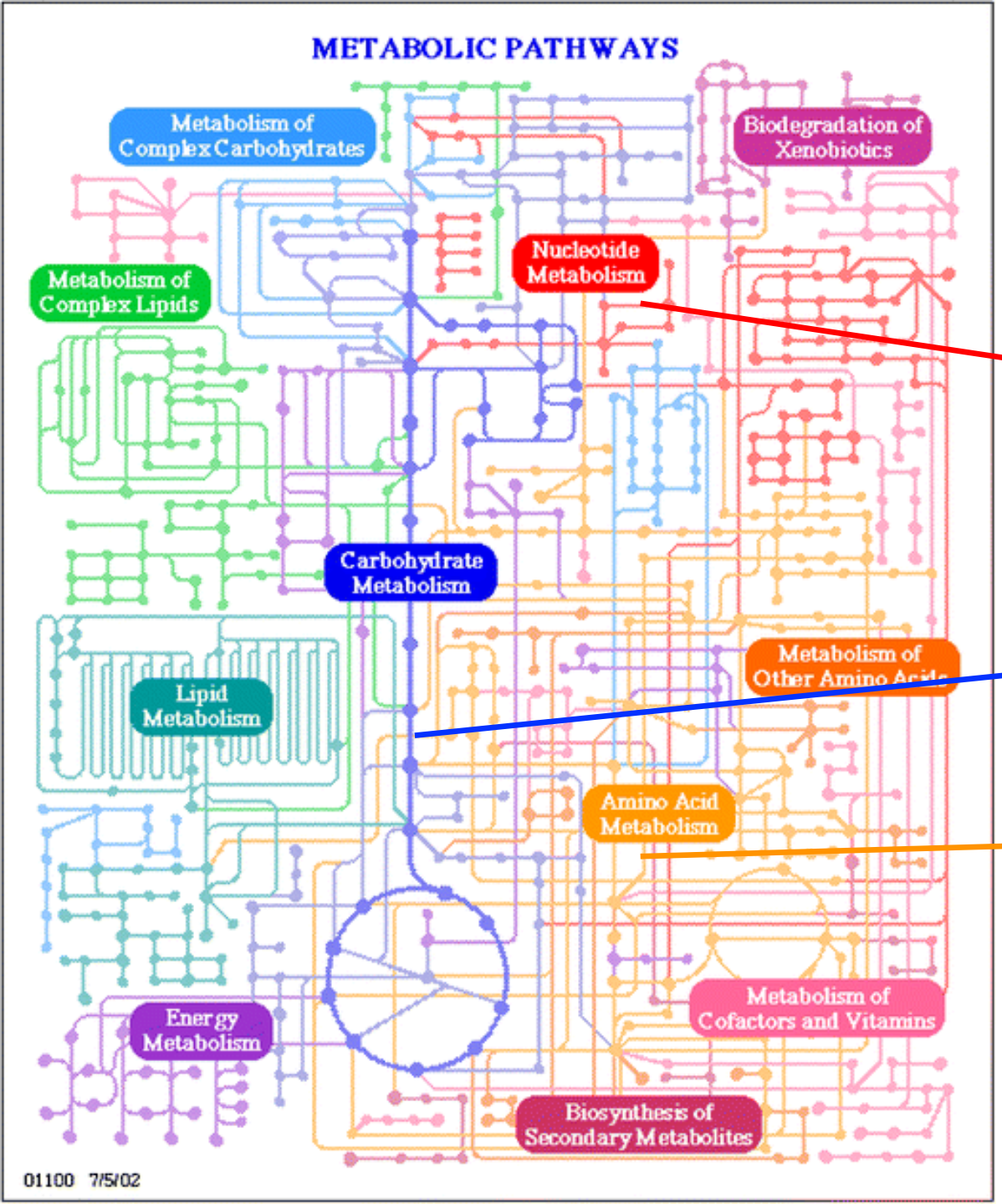
| JK7a | JK7b | JK8 | JK9 | JK10 |
|---|---|---|---|---|
| *ybhA* | *ybhA/pgl* | *gapA* | *ybhA/pgl* | *ybhA/pgl* |
| *gapA* | ***serA*** | ***serA*** | *gapA* | *gapA* |
| *purF* | *gapA* | *yjjK* | ***serA*** | *rpe* |
| *nadR* | *pykF* | *gltB* | *pykF* | *ilvH* |
| *rpoS* | *pyrE* | *livH* | | *rng* |

**Inefficient promiscuous activities may be sufficient to launch a new pathway**



# The same phenotypic result can be achieved in multiple ways






# The patchwork hypothesis

A

Enz1

B

Enz2

C

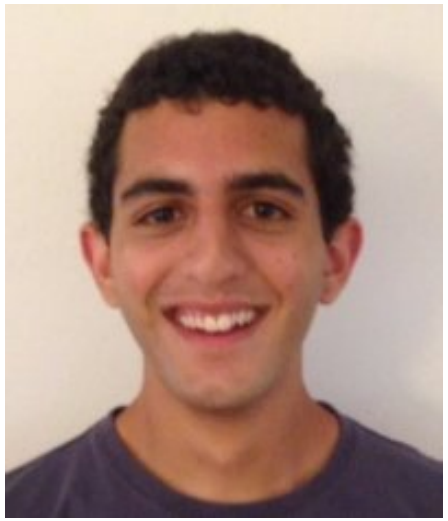
Enz3

D

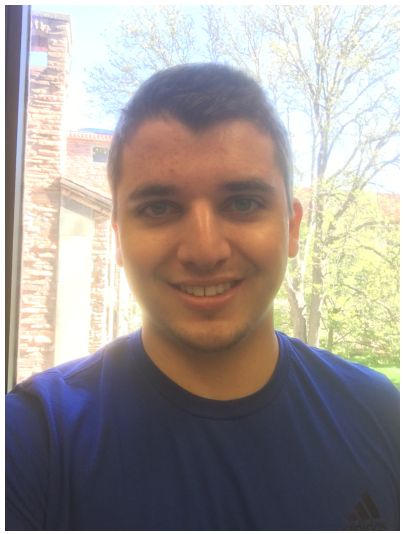
Jake Flood

Dr. Juhan Kim

Andrew Morgenthaler



Cyrus Gidfar



Michael Kristofich

NIGMS

NASA Astrobiology Institute

**Proteomics:** Will Old, University of Colorado Boulder
Chris Ebmeier
**Metabolomics:** Uwe Sauer, ETH Zurich
Tobias Fuhrer
**Sequencing:** Vaughn Cooper, University of Pittsburgh
Dan Snyder