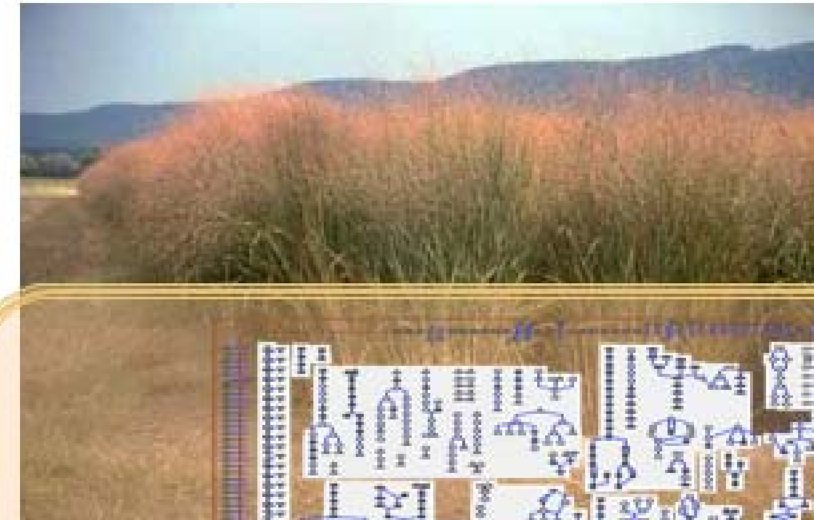
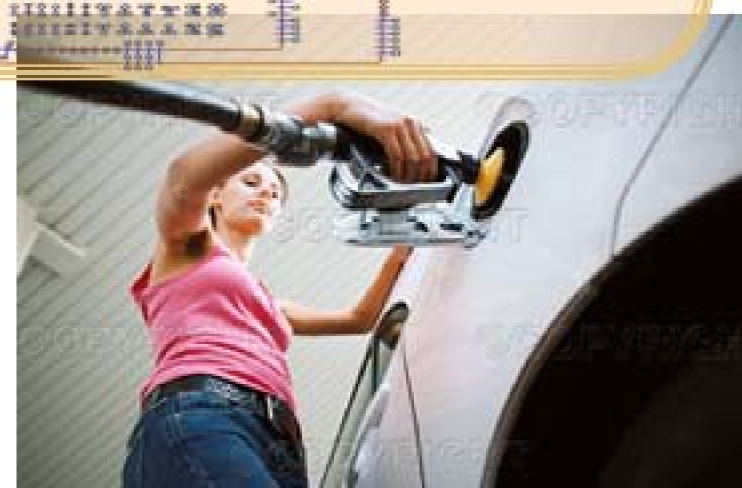
Synthetic biology and the race to new biofuels

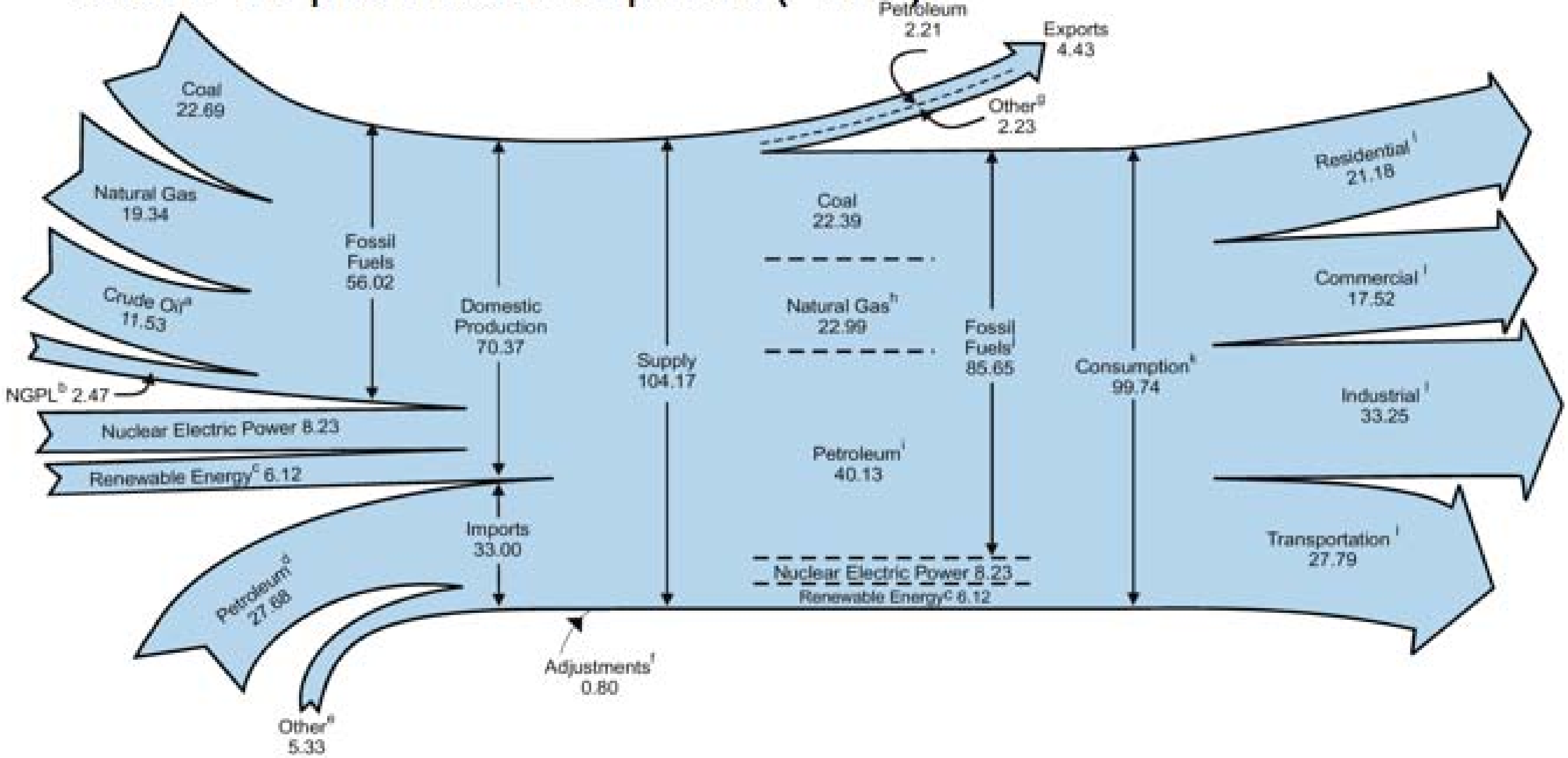


Frances Arnold
California Institute of Technology
NRG 0.1
Nov. 2, 2007



U.S. Energy Flow, 2004 (Quads)

- Transportation accounts for ~ 28% of energy use (22% globally).
- Transportation accounts for 27% of global carbon emissions.
- Most transportation energy is from petroleum fuels.
- A lot of US petroleum is imported (~60%).



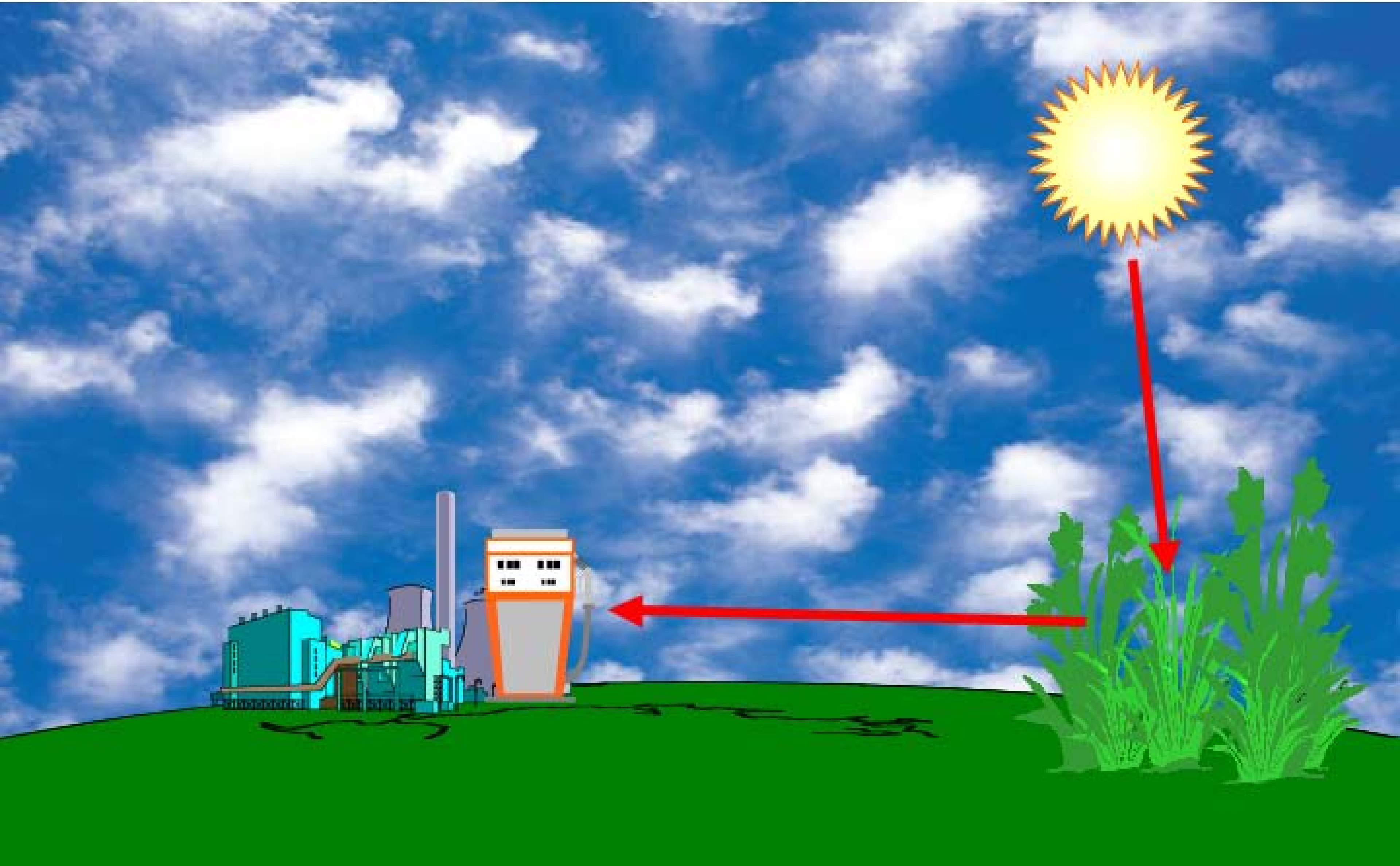
from: Patricia Dehmer; source: Energy Information Administration

Drivers for sustainable fuels:

- Security of supply & energy diversification

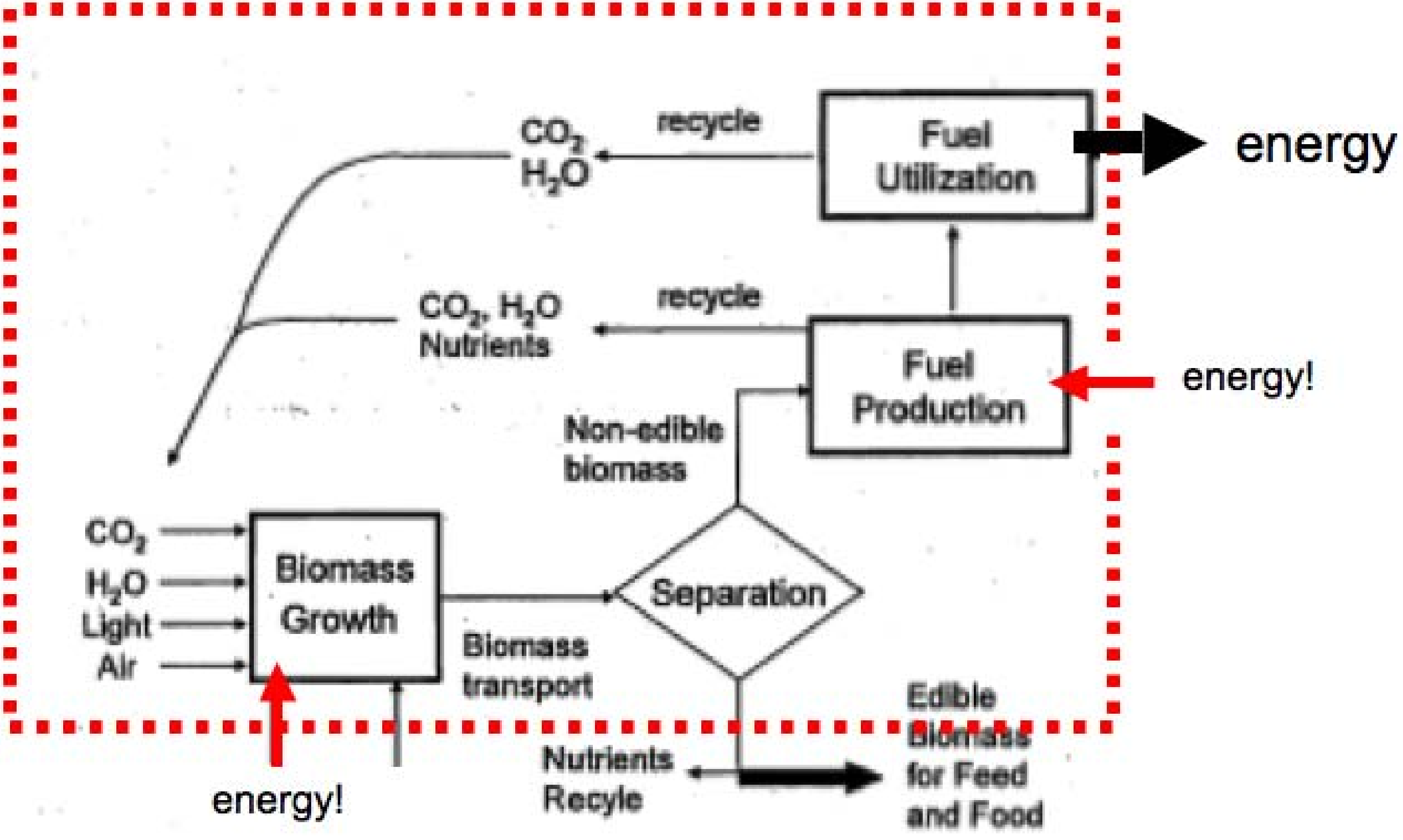
- Climate change

Biofuels are renewable
(solar energy is stored in the biomass).



To the extent that fossil fuels are not used in their production, biofuels are carbon-neutral: the emitted CO_2 gets taken up again by the biomass.

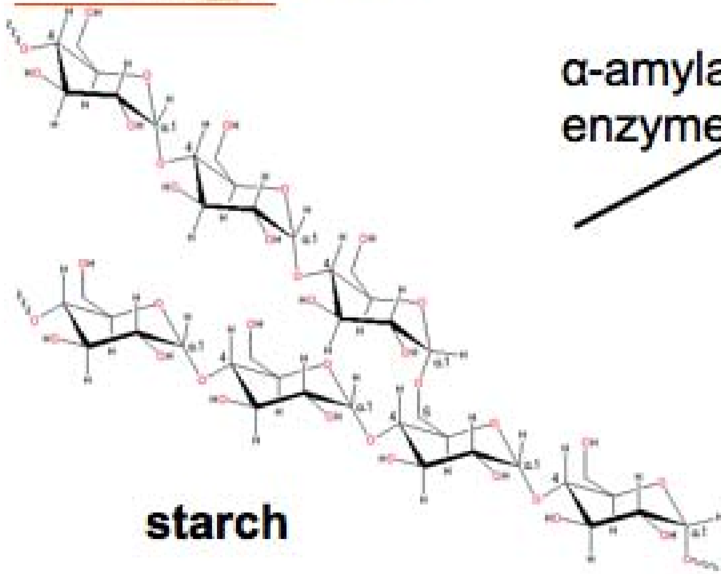
For the whole system, energy output should be greater than all the inputs



Corn ethanol creates more problems than it solves

72% starch
10% cellulose/hemicellulose

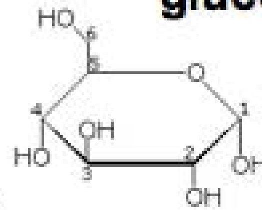
9% protein
4% oil
4% other



starch

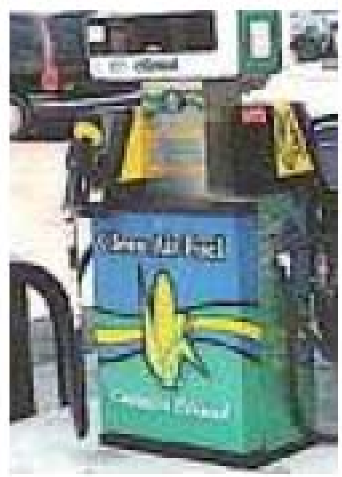
α-amylase
enzyme

glucose



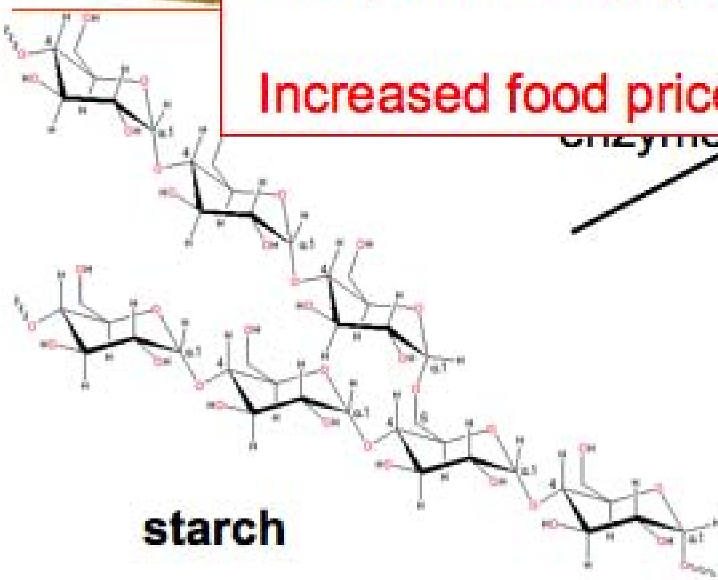
Brewer's yeast

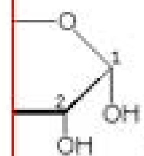
distillation

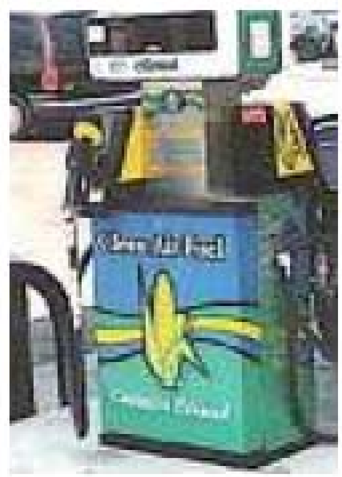


Corn ethanol creates more problems than it solves

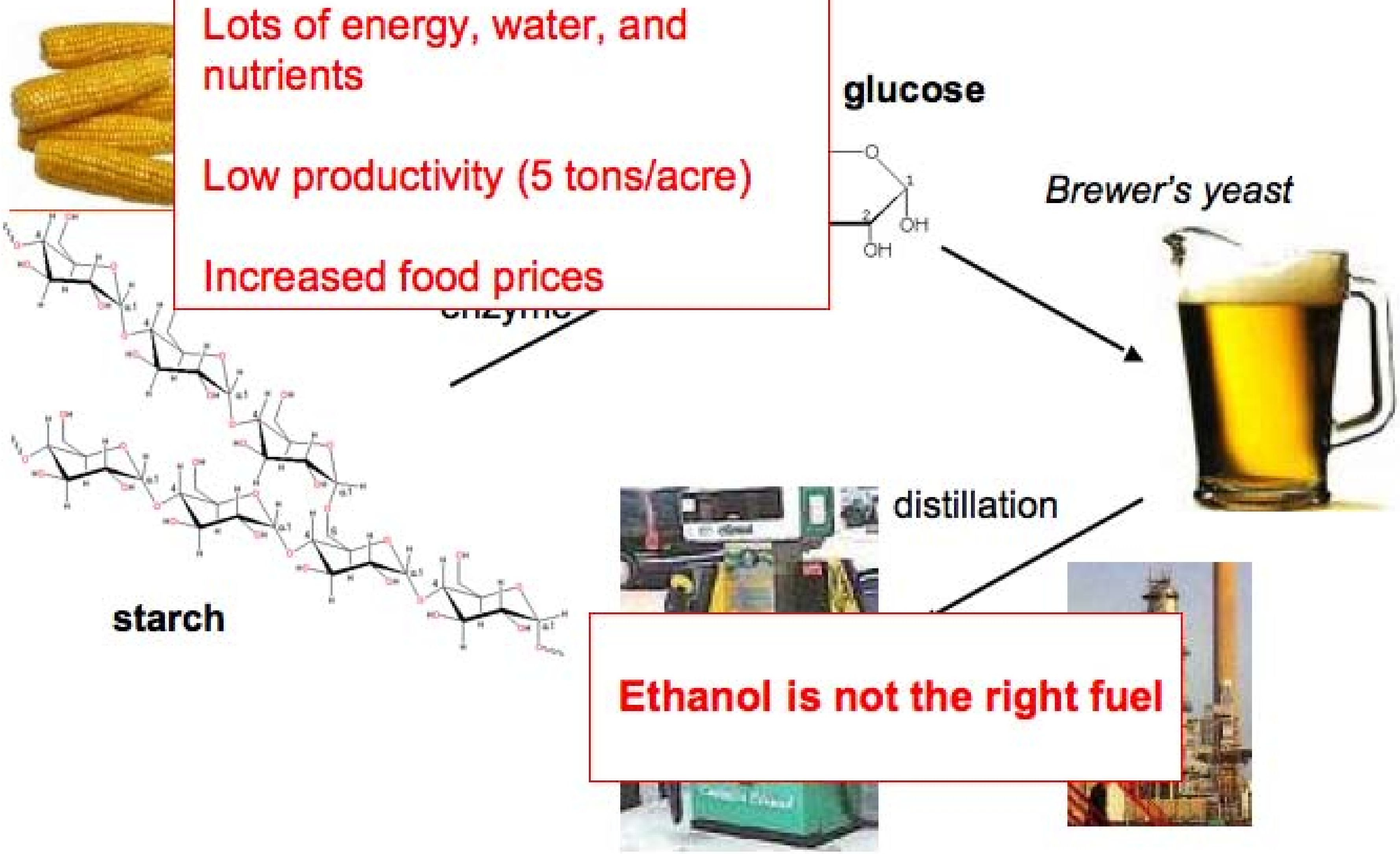
- Lots of energy, water, and nutrients
- Low productivity (5 tons/acre)
- Increased food prices







Corn ethanol creates more problems than it solves



Lots of energy, water, and nutrients

Low productivity (5 tons/acre)

Increased food prices

Ethanol is not the right fuel

Cellulosic feedstocks can be available on a large scale

- Short term
  - Wood, corn stover, rice straw, bagasse, corn fiber
- Longer term
  - High yield dedicated energy crops (switchgrass, **miscanthus**, willow, etc. 10-30 tons/acre)

1 year's growth without replanting

Cellulosic feedstocks can be available on a large scale

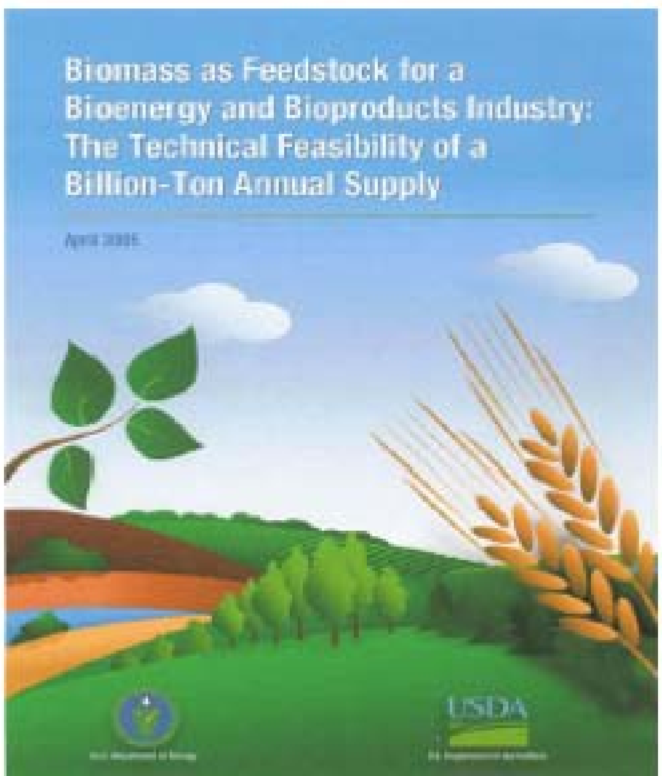
- Short term
  - Wood, corn stover, rice straw, bagasse, corn fiber
- Longer term
  - High yield dedicated energy crops (switchgrass, **miscanthus**, willow, etc. 10-30 tons/acre)

1 year's growth without replanting

Cellulosic feedstocks can be available on a large scale

- Short term
  - Wood, corn stover, rice straw, bagasse, corn fiber
- Longer term
  - High yield dedicated energy crops (switchgrass, **miscanthus**, willow, etc. 10-30 tons/acre)

1 year's growth without replanting



DOE and USDA estimate that 1.3 billion tons/year of cellulosic biomass could be made available

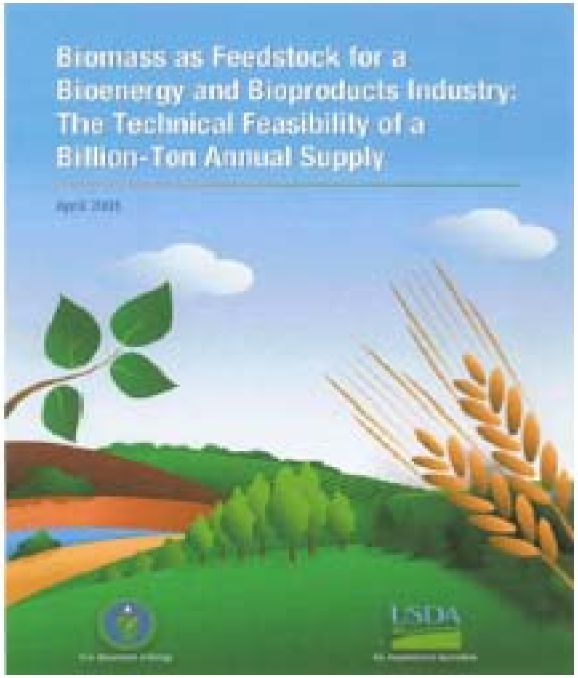
Energy content of 1.3×10^9 tons:

$= 1.9 \times 10^{19}$ J

$= 3.2 \times 10^9$ BOE (barrel of oil equivalent)

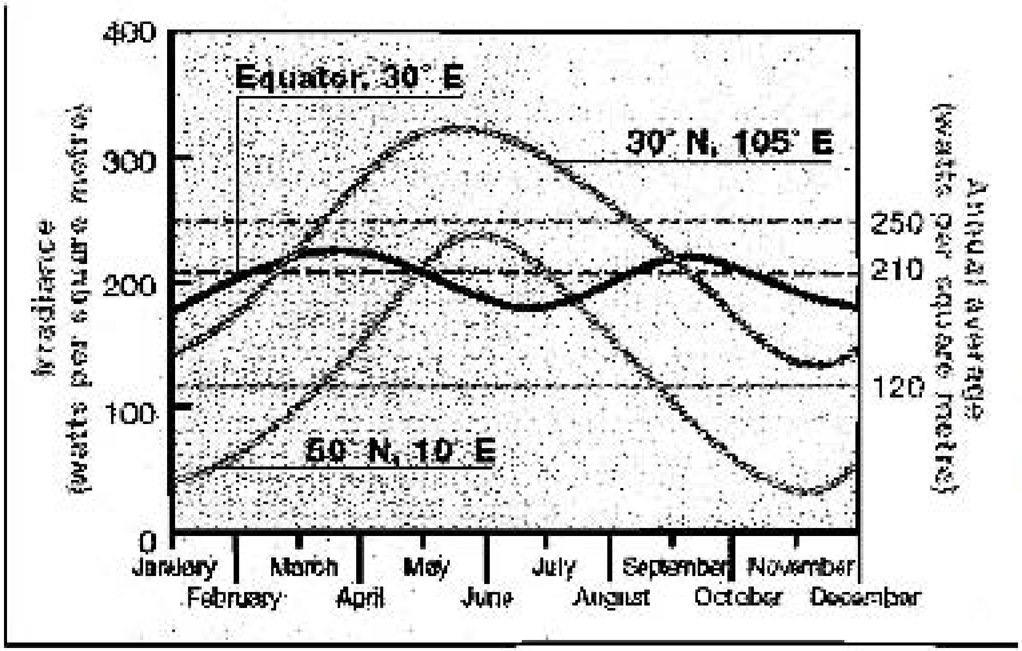
(cf. US consumption of $\sim 7 \times 10^9$ barrels of oil/year)

Substance	Energy Density (MJ/kg)
H_2	141.8
Crude Oil	44
Coal	30-40
Sugar cane	16
Wood	16
Ethanol	28



How efficient is biomass at collecting/storing the solar energy?

Time averaged insolation in mid-latitudes, averaged over the year (day/night) and weather patterns, is 200 $W/m^2 = 2.5 \times 10^{13}$ J/acre-year.



If (cellulosic) crop yield is 20 tons/acre-year, the energy content is ~ 2.9×10^{11} J/acre-year.

$$\frac{2.9 \times 10^{11}}{2.5 \times 10^{13}} = 0.012 \ (1.2\%)$$

If crop yield is only 5 tons (corn), efficiency drops to 0.29%.

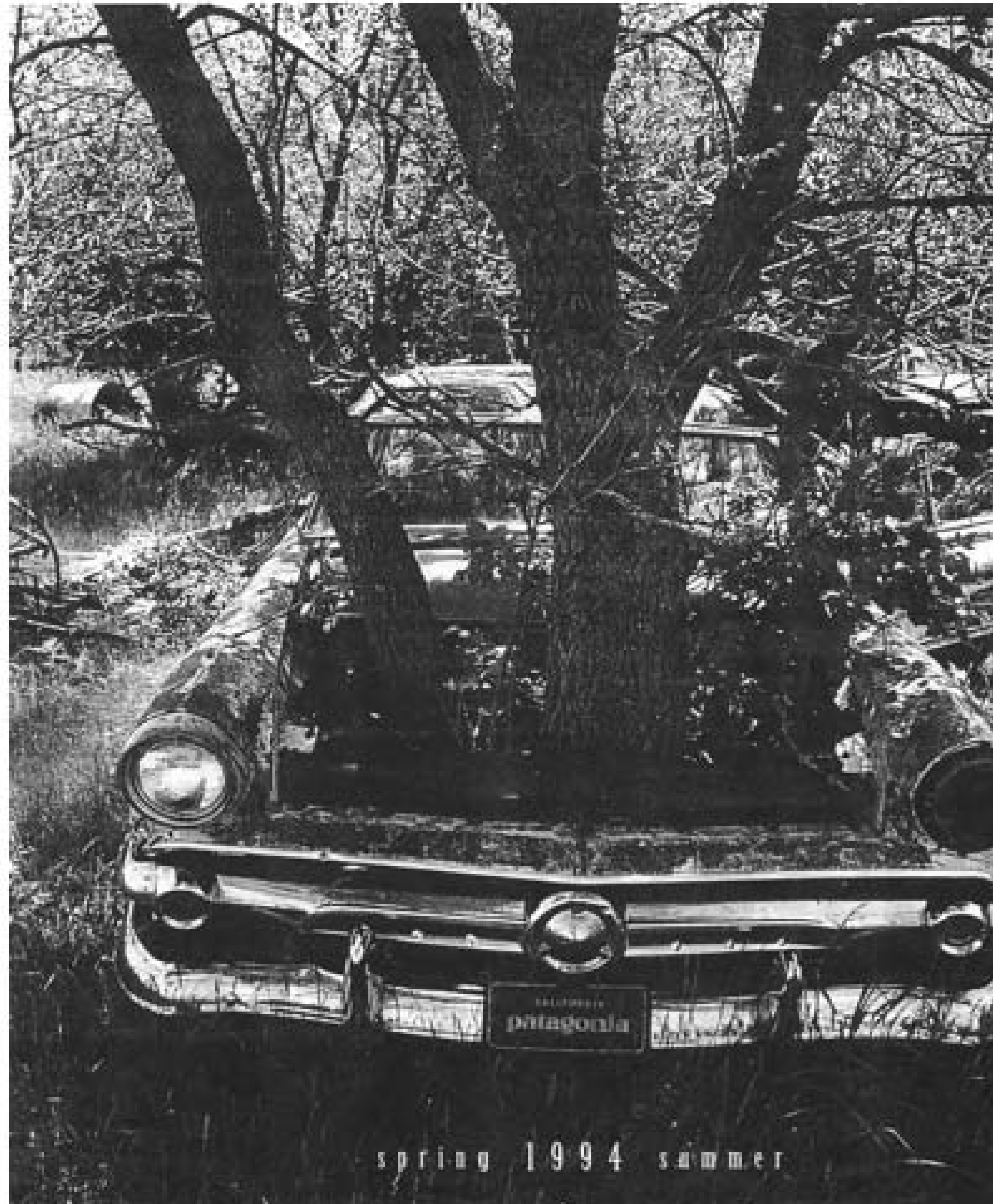
Good news:
There's a lot of energy stored in biomass.

Bad news:
The photosynthetic efficiency of converting sunlight into biomass is low.

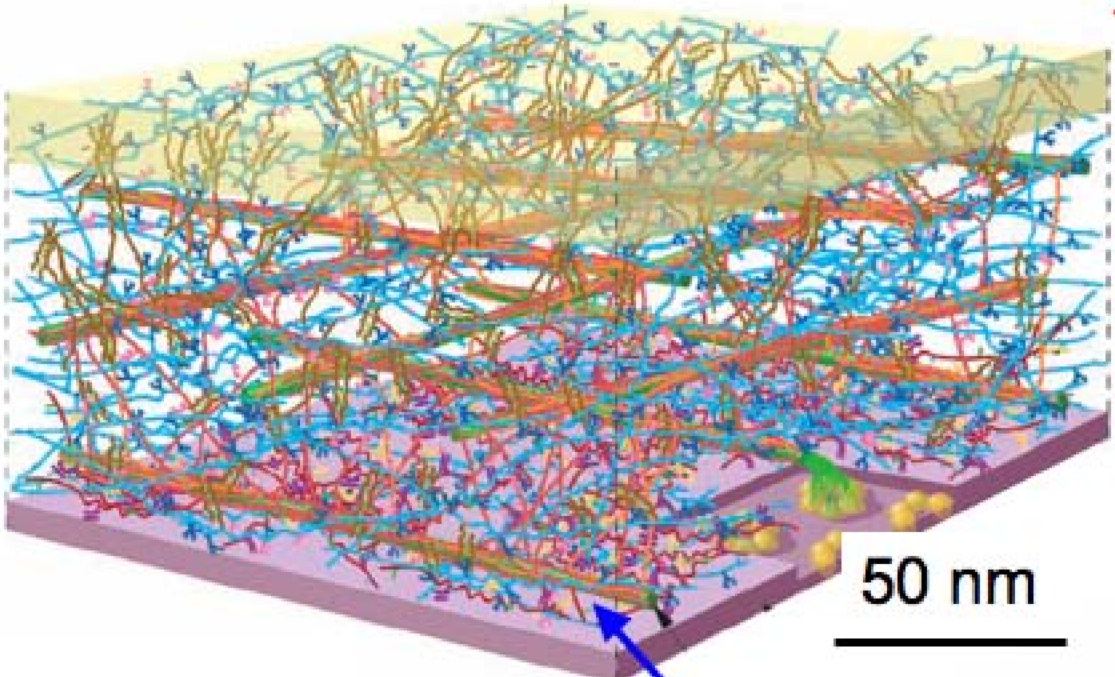
Land requirement is very large.

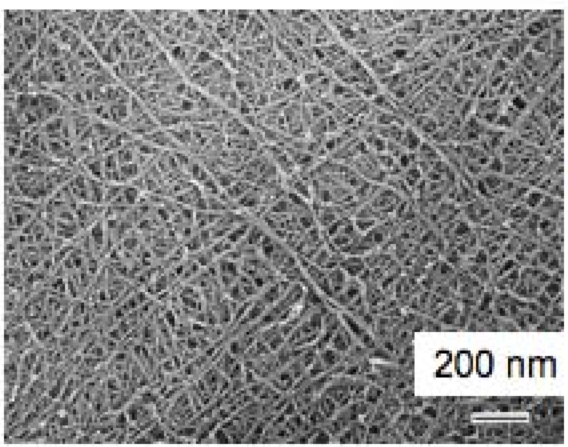
And…

…biomass is not useful for powering cars. It needs to be converted to a useful form (with further energy loss).



What's so hard about converting biomass to fuel?



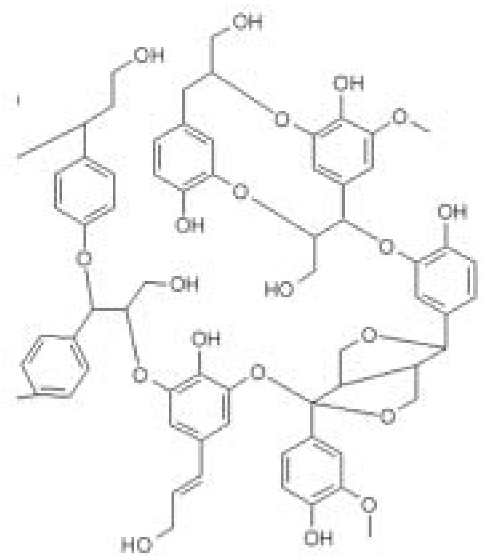


← **Cellulose microfibrils**

Somerville et al. *Science* 306:2206 (2004)

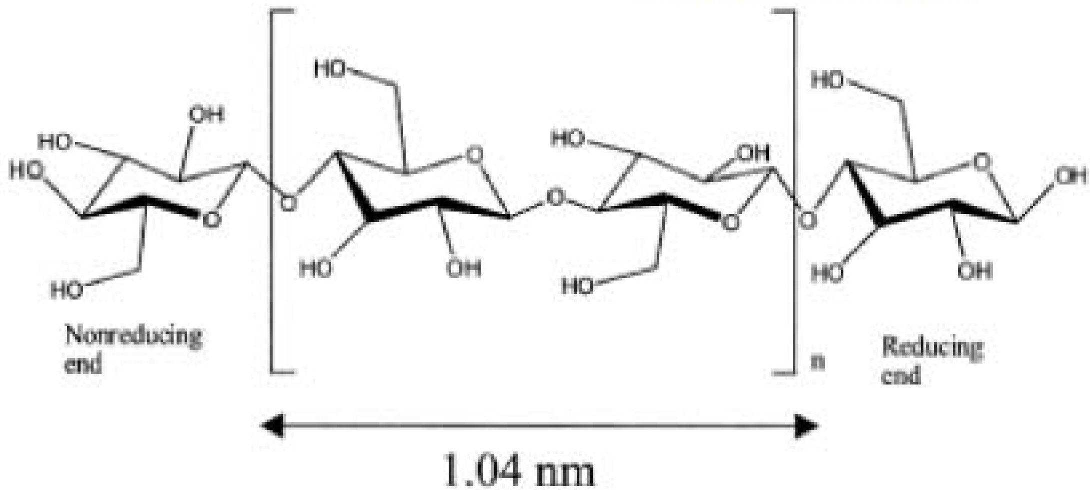
These produce usable sugars:

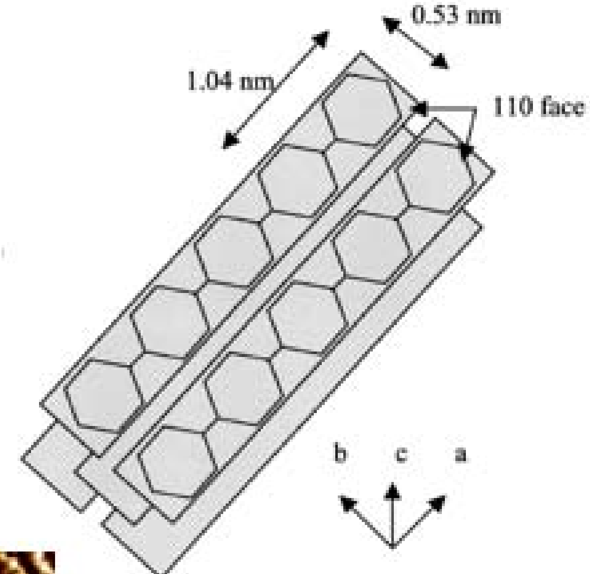
~45% Cellulose
~25% Hemicellulose
~25% Lignin
~ 5% Other

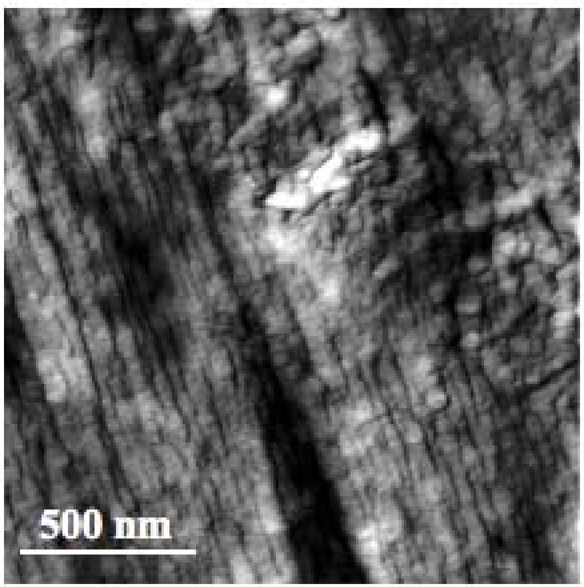


Lignin

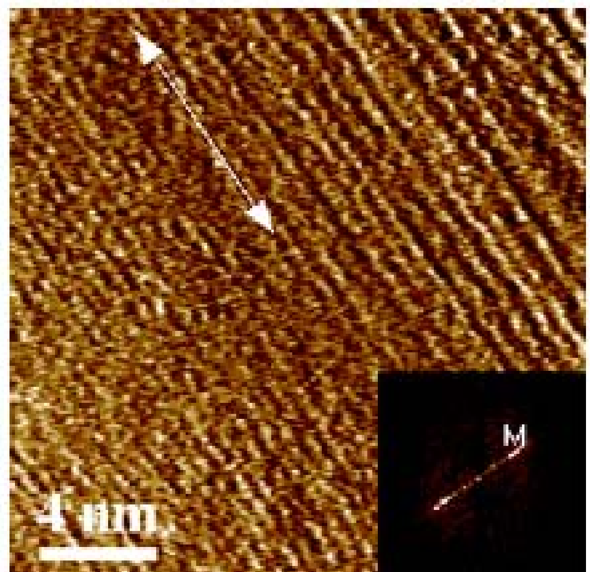
Cellulose is ‘recalcitrant’ because it is physically inaccessible



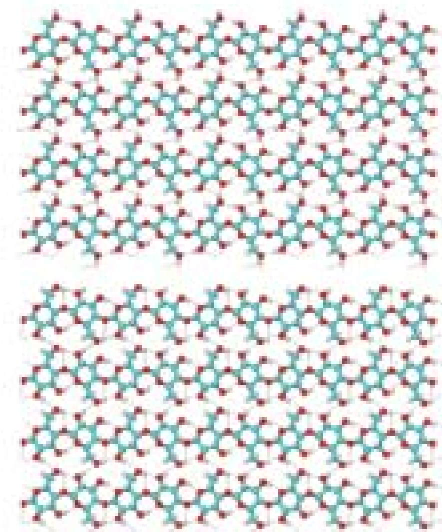




http://hvsee.asu.edu/imagegallery/RealLifeSciences/Spm/cellulose/CELLULOZ.jpg

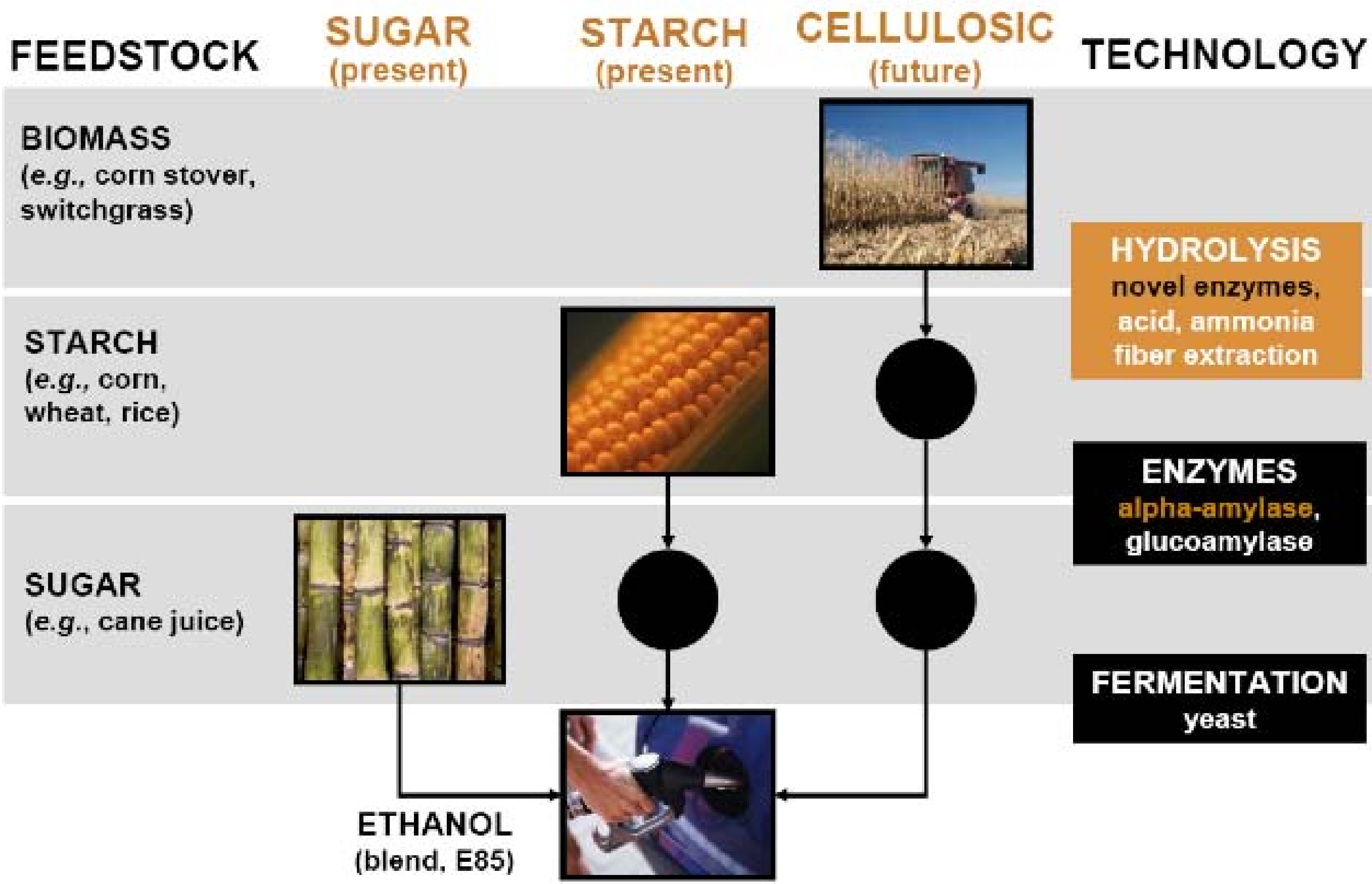


http://www.wellesley.edu/Chemistry/chem227/sugars/cellulose.gi



Zhang Y. H.-P., Lynd L. R. Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. 88, 797-824, 2004.

Cellulose utilization requires additional processing steps ($$)



Do we really want ethanol?

Low fuel quality

- Lower energy content than gasoline (poor mileage)
- Problematic when takes up water (e.g., in pipelines, storage tanks); high vapor pressure
- Can only blend up to c. 10% with gasoline without engine modification

What alternatives are there?

- Biodiesel (not made from lignocellulosics)
- Hydrocarbons

  technology not demonstrated yet
- **Butanol**

  butanol has been made by fermentation for 90 years;

  we can do more and better now

("Ethanol Is Not The Only Green In Town", Business Week, April 30)

Ethanol

High water content
Low energy content

Butanol

Low water content
High energy content

	Gasoline	Ethanol	Butanol
Energy content (BTU / gallon)	115 K	84 K	110 K

- Butanol can be distributed in existing pipelines, storage tanks (*not* hygroscopic; non-corrosive)
- Butanol burns cleanly in unmodified gasoline engines
- Butanol can be blended with gasoline at any ratio
- Good intermediate to other fuels and chemicals

Brief history of butanol

1916: Industrial ABE (Acetone, Butanol, Ethanol) fermentation process uses *Clostridium acetobutylicum* to produce acetone from starch (Chaim Weizmann patent)

1927: Butanol recognized as good motor fuel, solvent

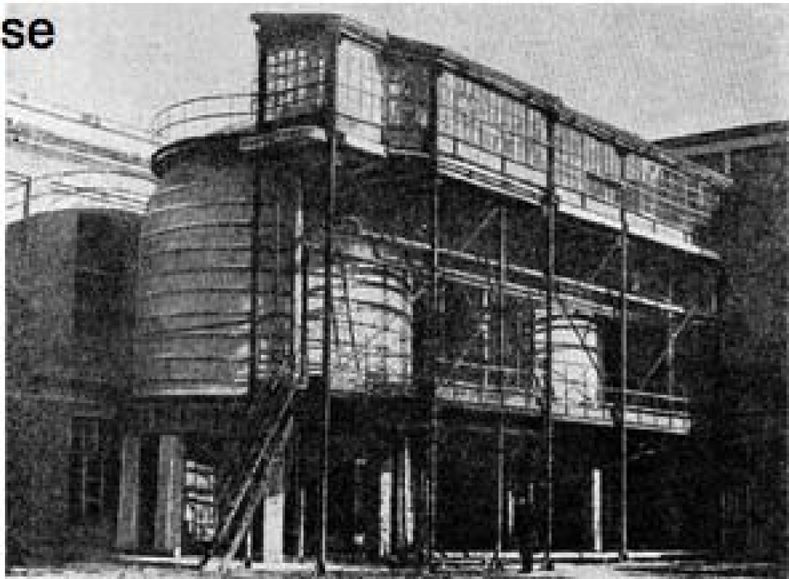
1930's: Large ABE fermentation plants produce butanol from molasses

1950's: Petroleum refining becomes more economical source of chemical & fuel feedstocks than molasses fermentation

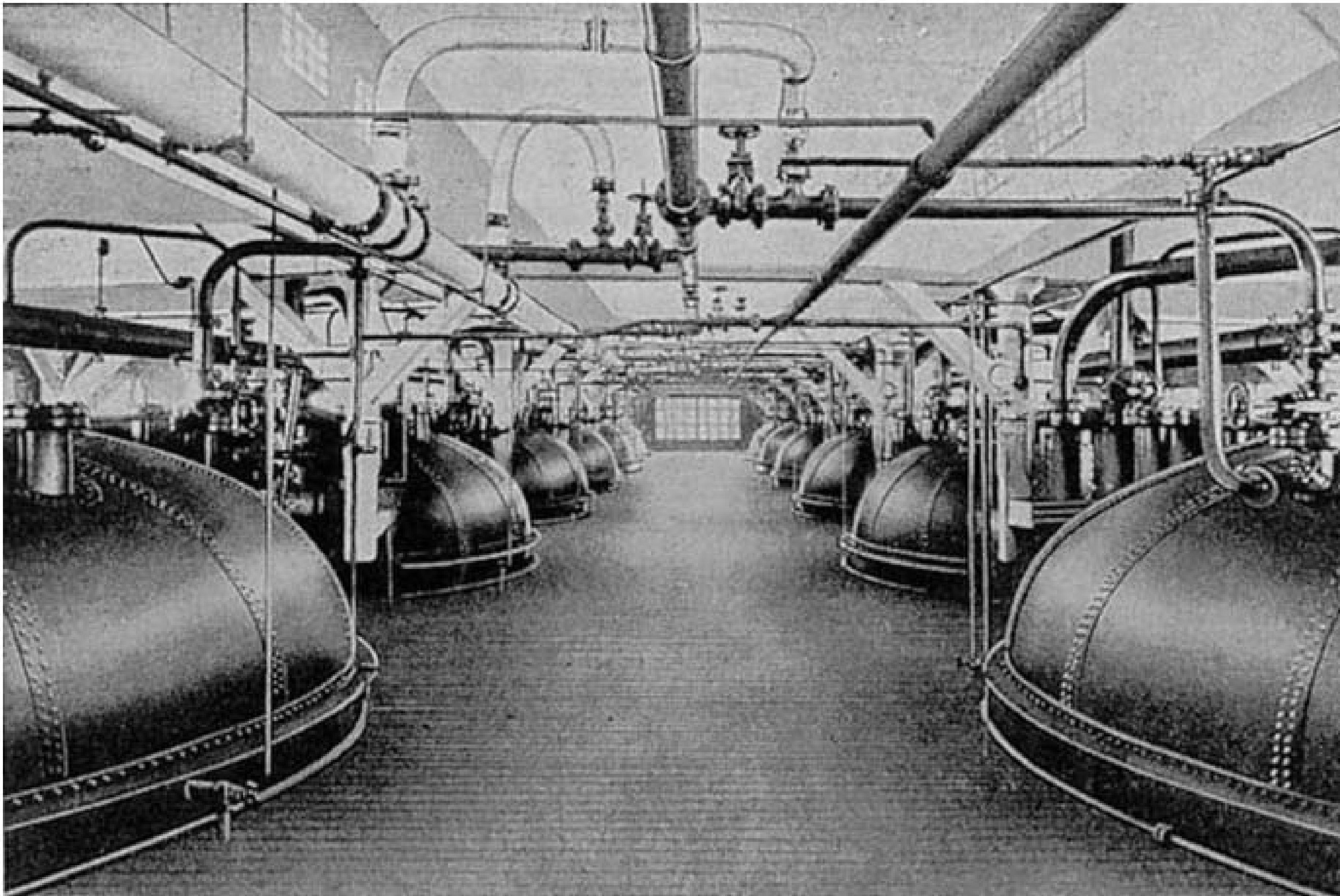
1960's: Closure of Western molasses ABE fermentation plants
Soviet agricultural waste fermentation plants continue operation, as do plants in South Africa

1980's: Soviet and South African plants close

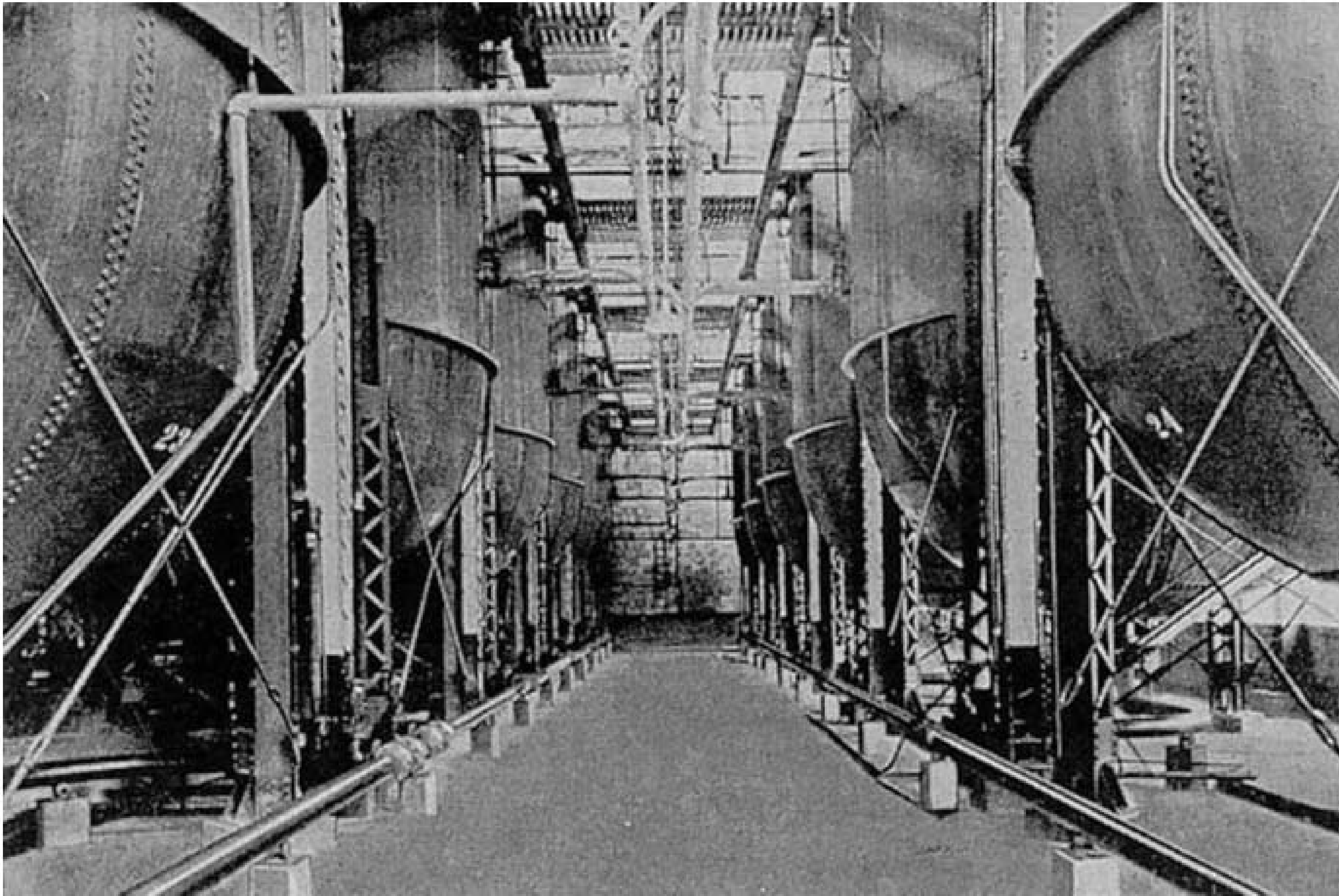
2000's: High fossil fuel costs rekindle interest in butanol



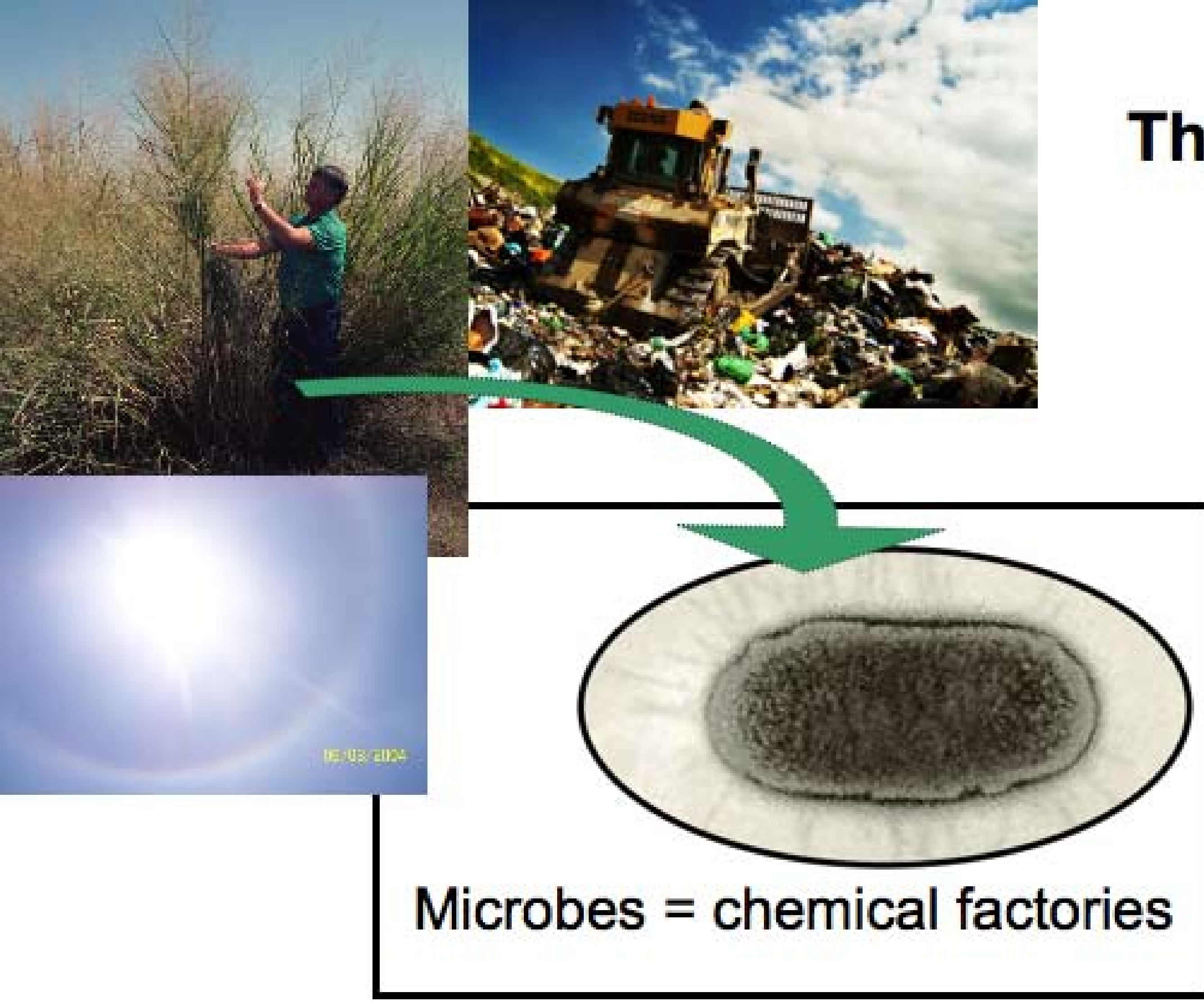
CSC 50,000 GAL TANKS UPPER



CSC 50,000 GAL TANKS LOWER



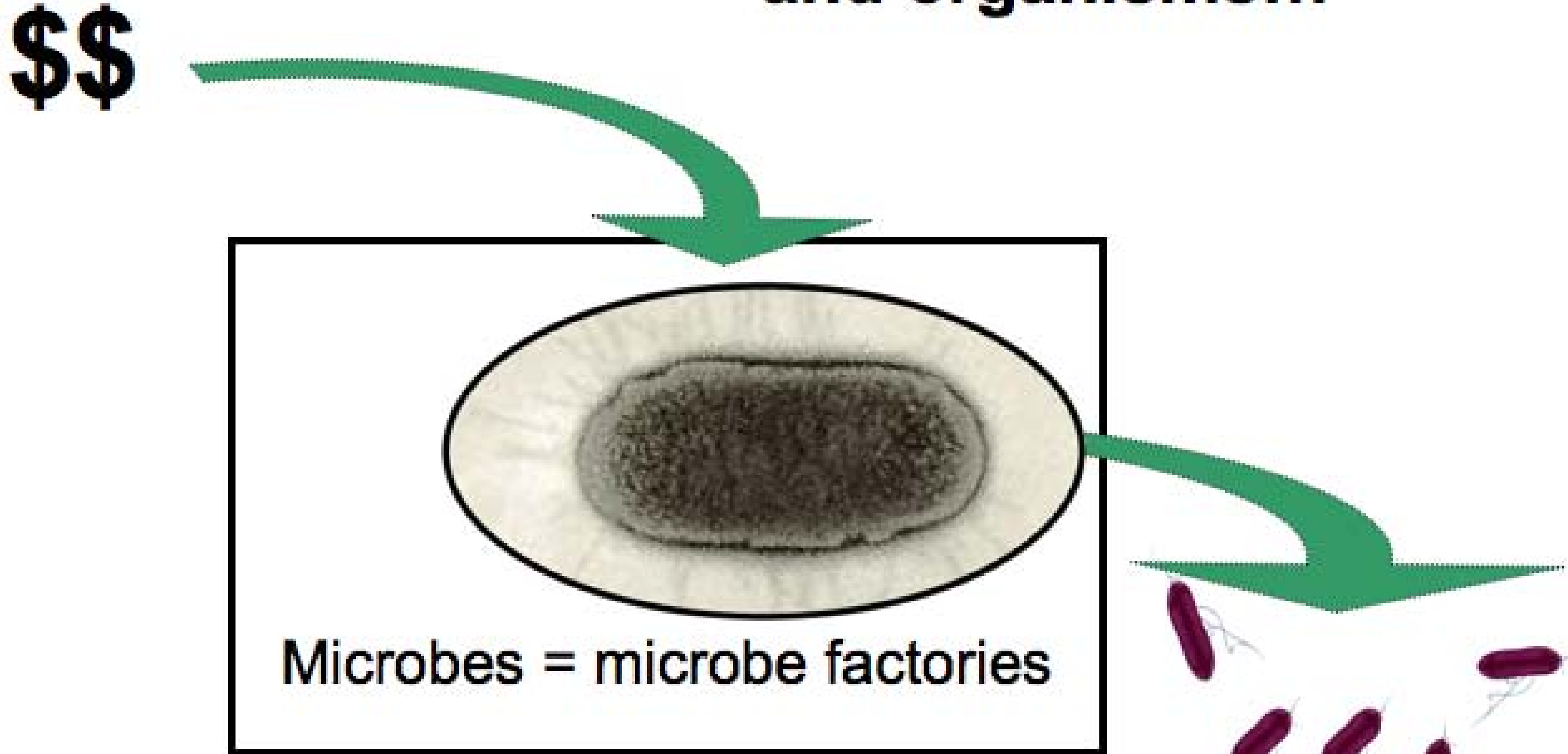
The dream…



The dream…

Microbes = chemical factories

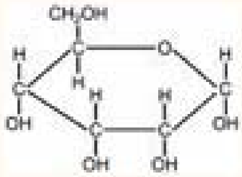
The reality for butanol with the old processes and organisms…



Product is a mixture (acetone and ethanol are produced along with butanol) and yields are low

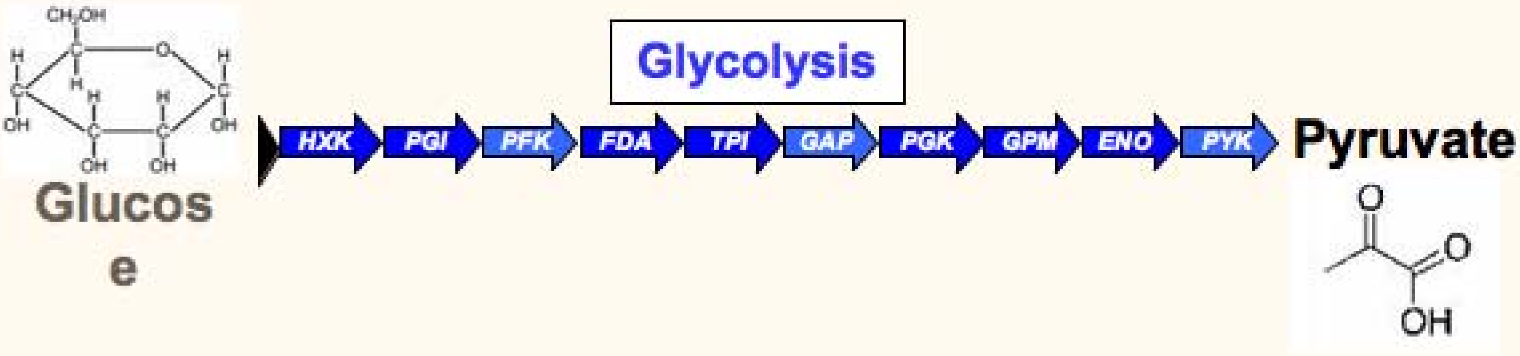
A series of enzymes catalyzes the multi-step conversion of sugar to butanol inside the bacterium

A series of <u>enzymes</u> catalyzes the multi-step conversion of sugar to butanol inside the bacterium

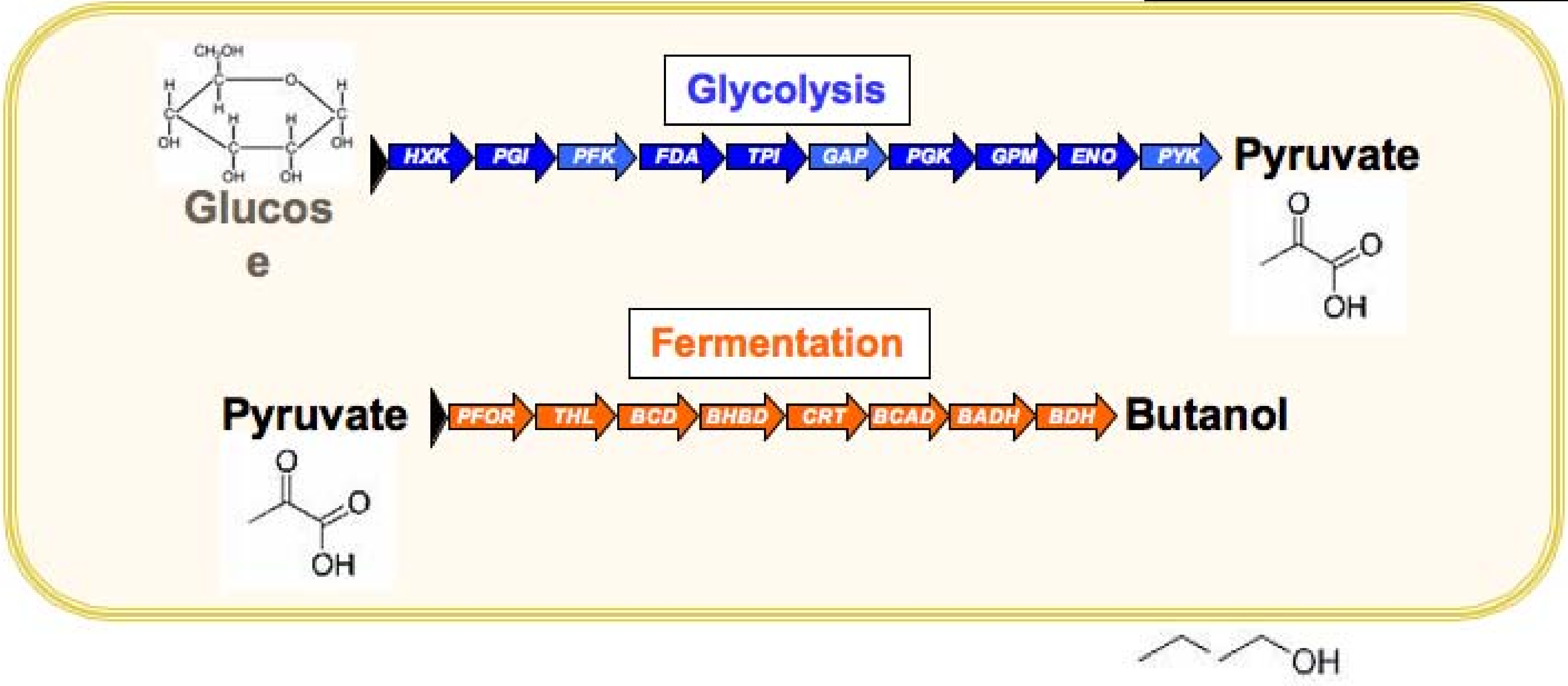


Glucos e

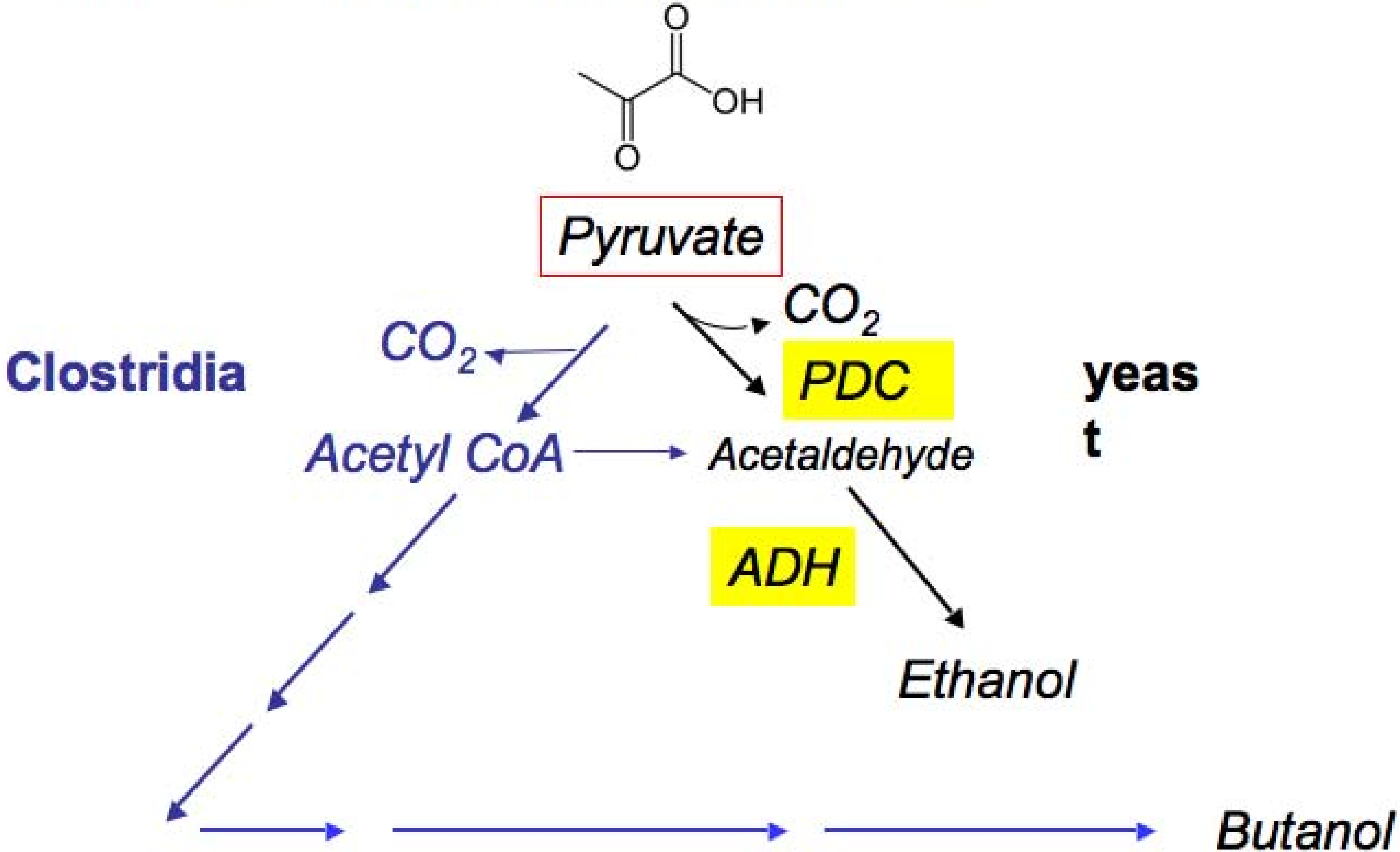
A series of enzymes catalyzes the multi-step conversion of sugar to butanol inside the bacterium



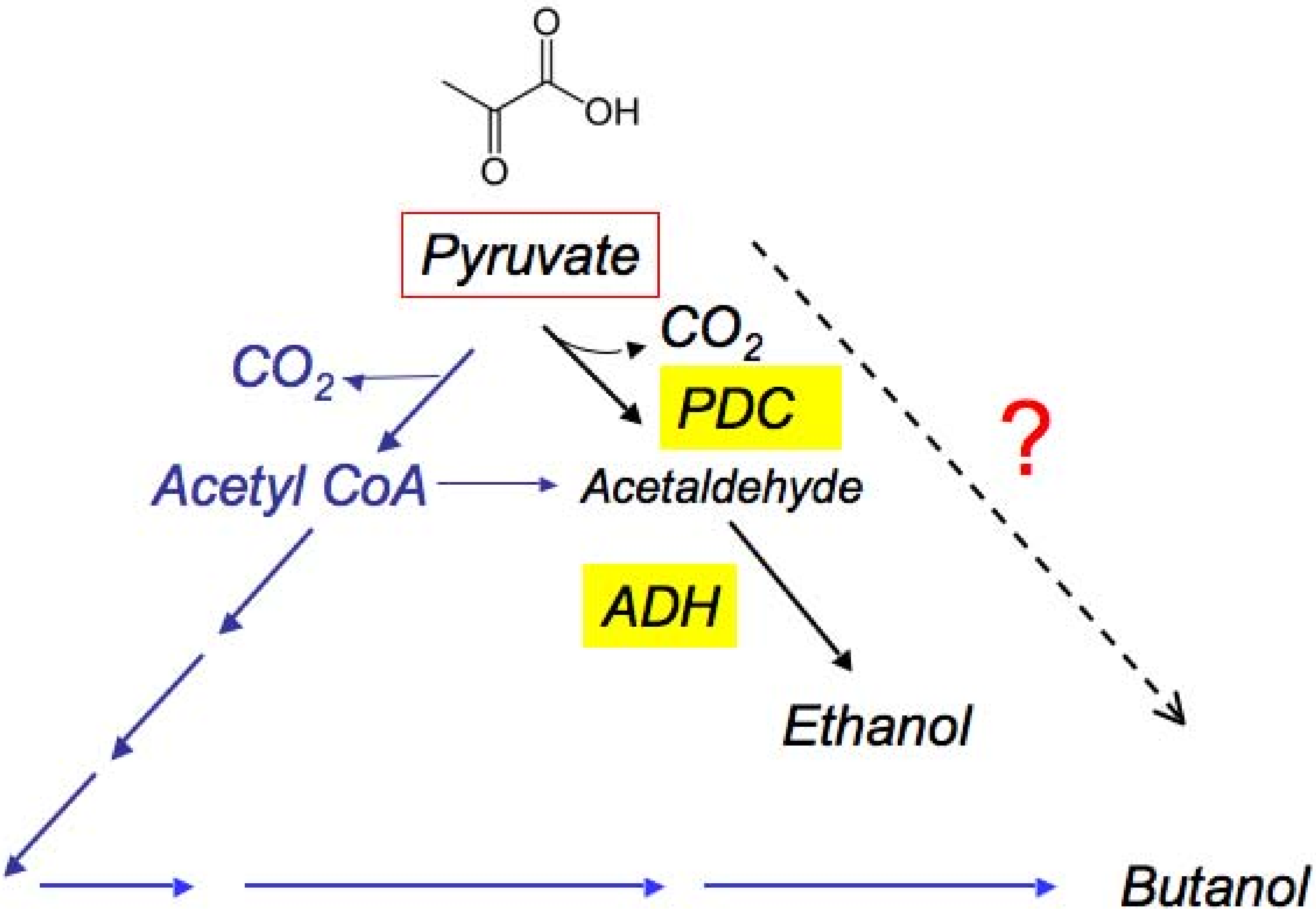
A series of enzymes catalyzes the multi-step conversion of sugar to butanol inside the bacterium



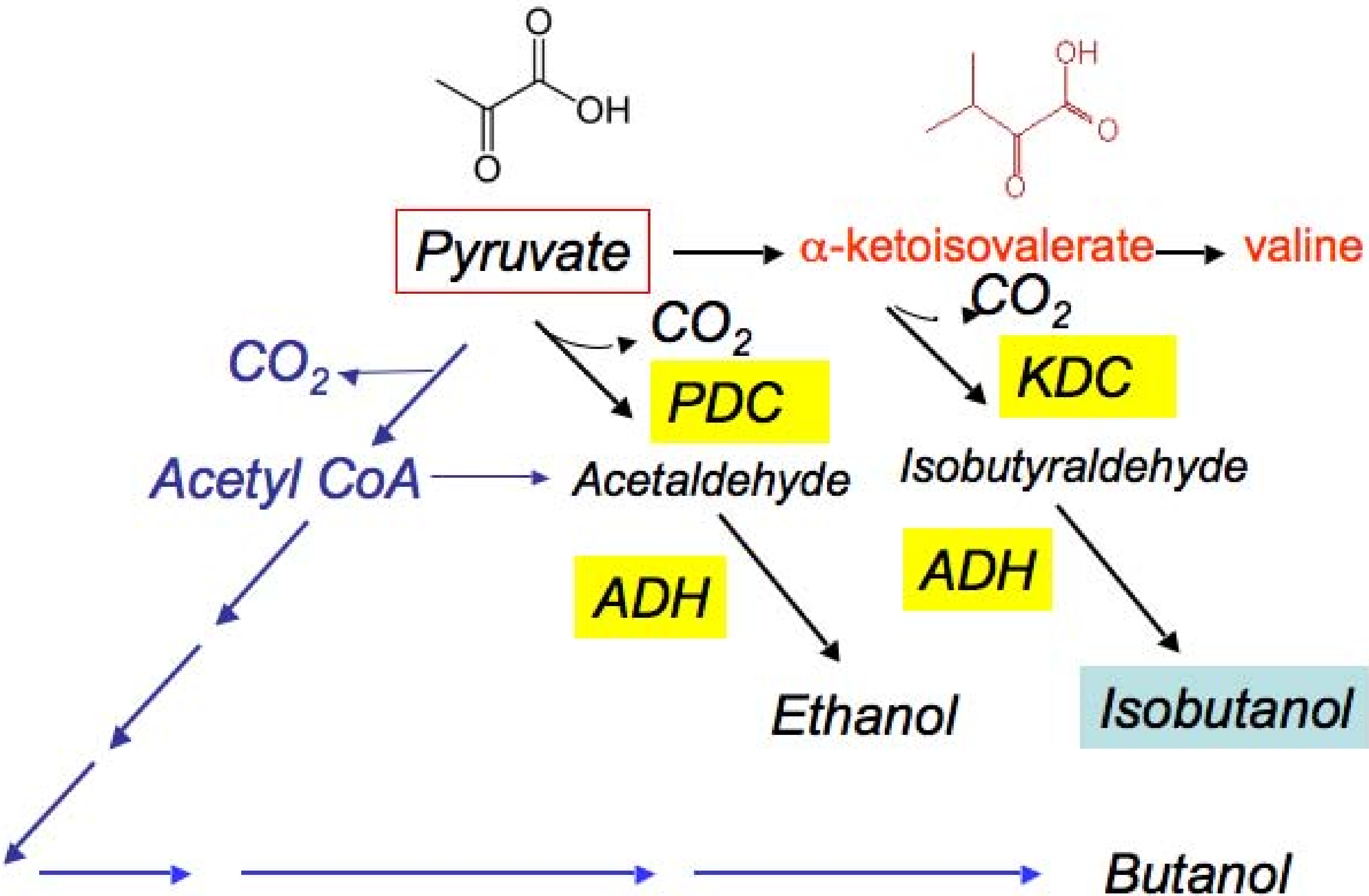
Pyruvate is an intermediate common to ethanol and butanol production. Butanol requires many more steps, and its CoA intermediates limit the flux.



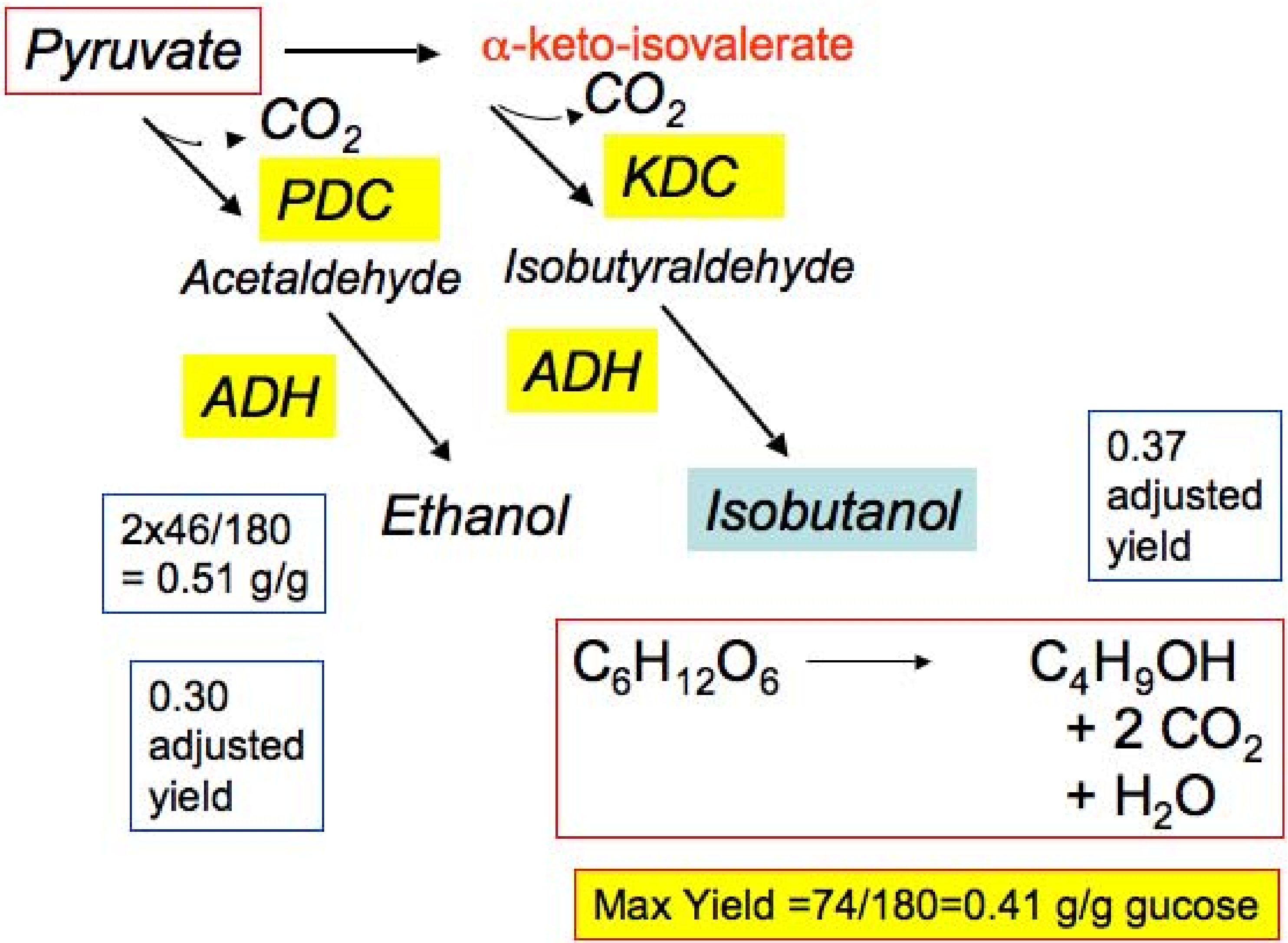
Can we engineer a new, more efficient pathway to butanol?

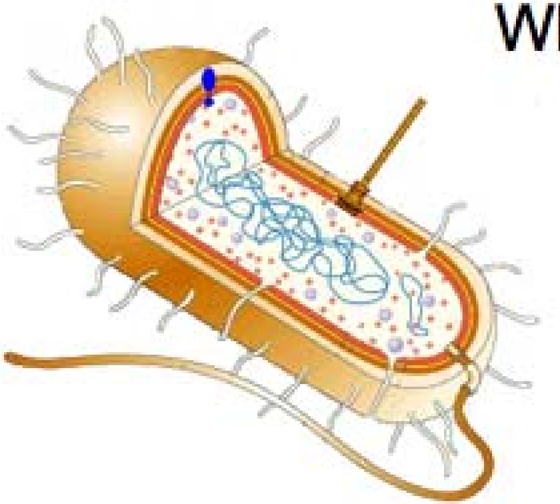


Jim Liao (UCLA): make isobutanol (branched-chain isomer) via the valine pathway with only two added enzymes



Jim Liao (UCLA): make isobutanol (branched-chain isomer) via the valine pathway with only two added enzymes

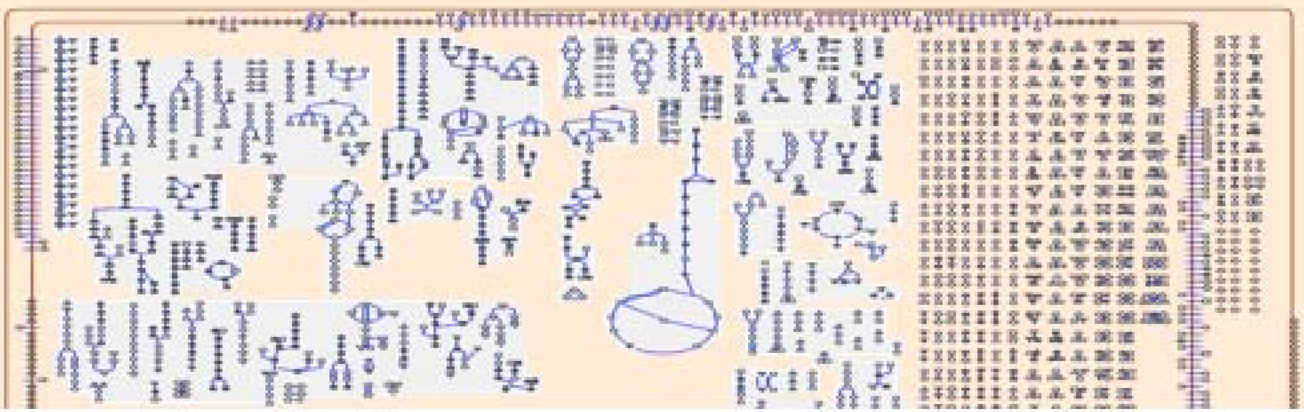




What organism?

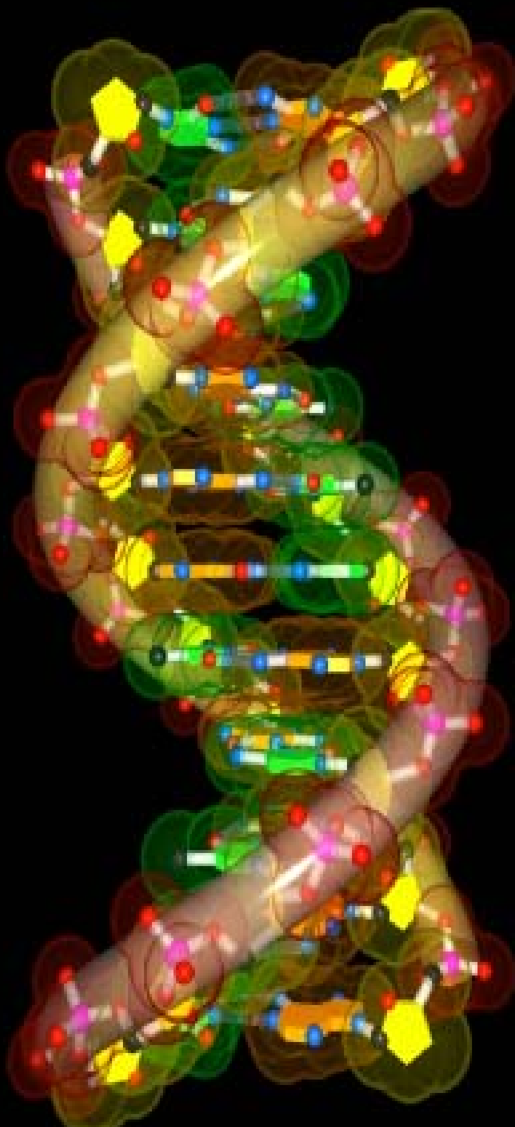
versus

?



The *Escherichia coli* ‘chassis’
(we think we understand how this one is programmed)

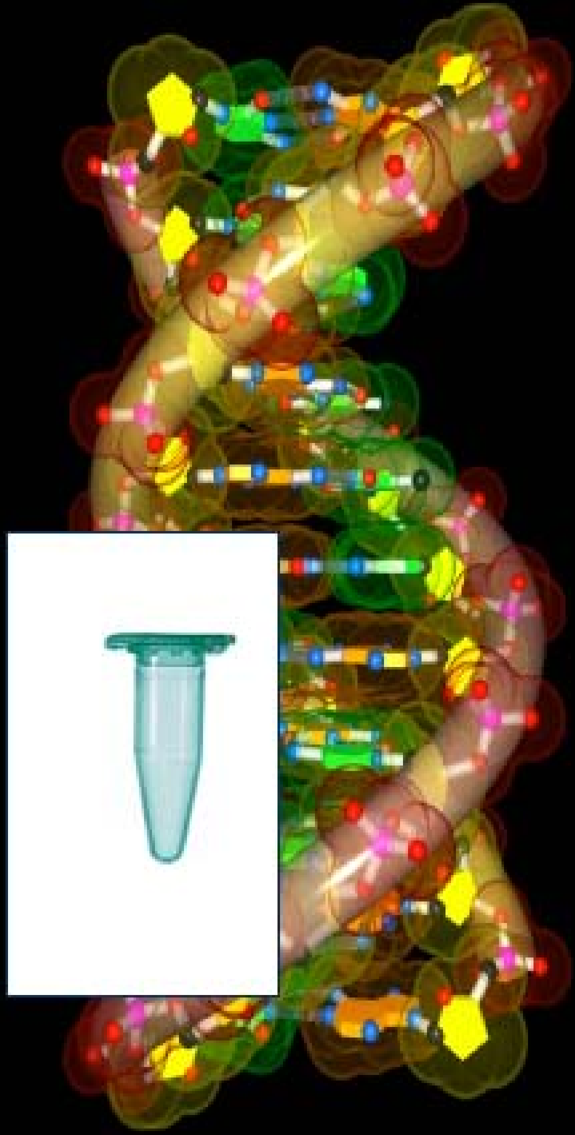
…all of it encoded in DNA:



…AATAGCCGTTATTTCCGGATG
TGCATAGCTGATTTGACCCATCC
GGTACACCAATGGGTCGGACAA
ATCCCGATTTGATCGTGTGCGC
GACATGTCTTCCGGCGACACAT
GTGTCTCTCACTCCGAGAGATC
GGTTAGAGTCTCGGTTAACCAC
ACGTCCCGGATATATTTAATTGG
CCGGAGAGTCTCCCGCGCGACA
TAAGGAGTCCTCGTTTCGAGAT
ACGTACGGCATGGTGACACCAG
TTGCCCTCTGATTCCCGGAGCC
TCTTTGAAAACGTCGAGTCGAAT
CGAAGTTCGAACCCCGGATCGG
GTCCACCAACTTAGAGATGTGT
GTGCGCTGACTCAGTCAT…

Synthesizing DNA used to be expensive and error-prone

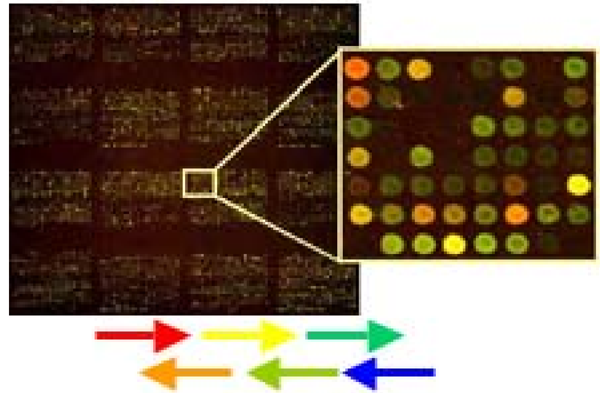
...AATAGCCGTTATTTCCGGATG
TGCATAGCTGATTTGACCCATCC
GGT...



RKM.COM.AU

...AATAGCCGTTATTTCCGGAT

Today

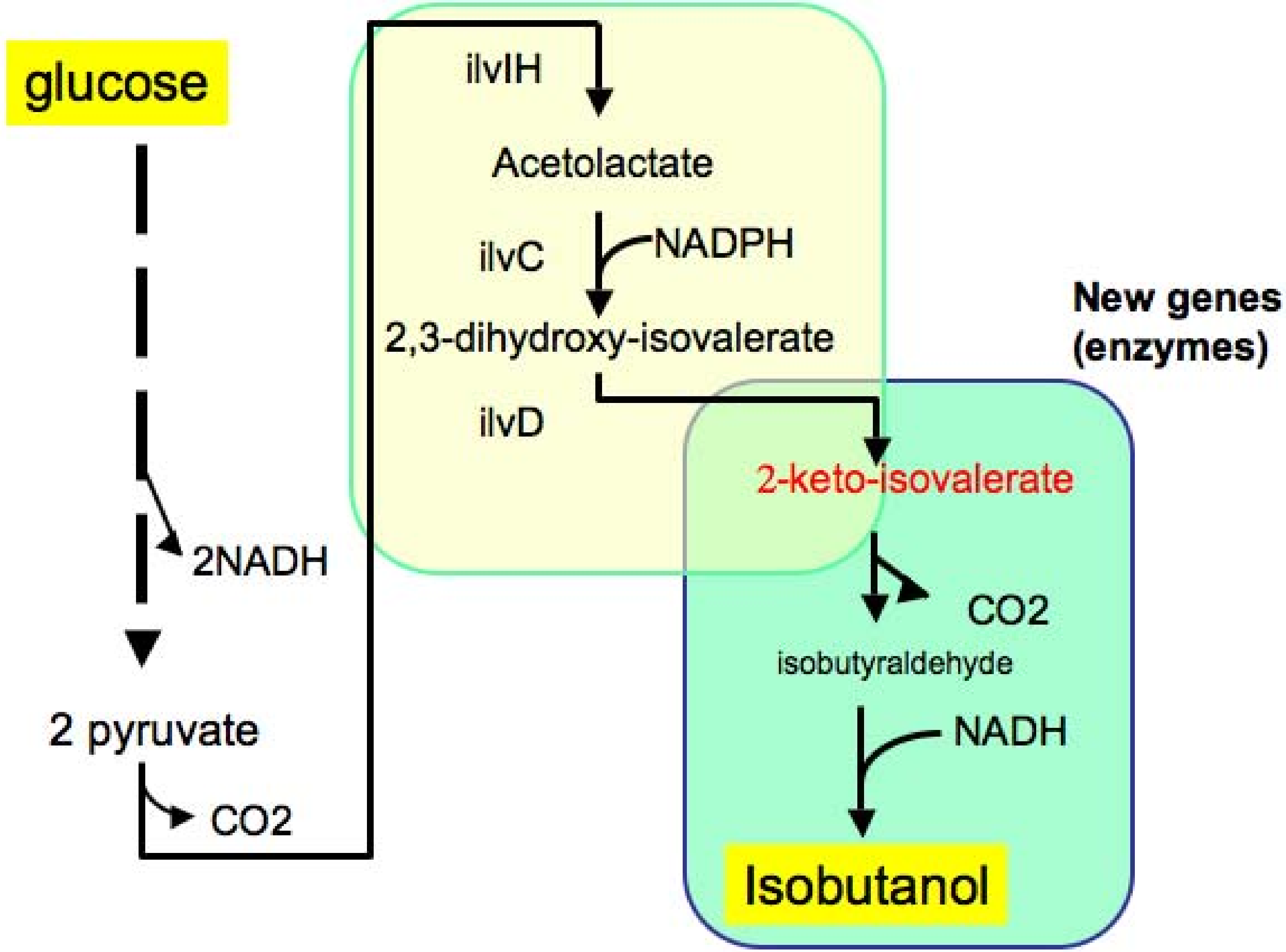


Massively parallel DNA synthesis

De novo synthesis of desired parts, to > 100,000 base pairs

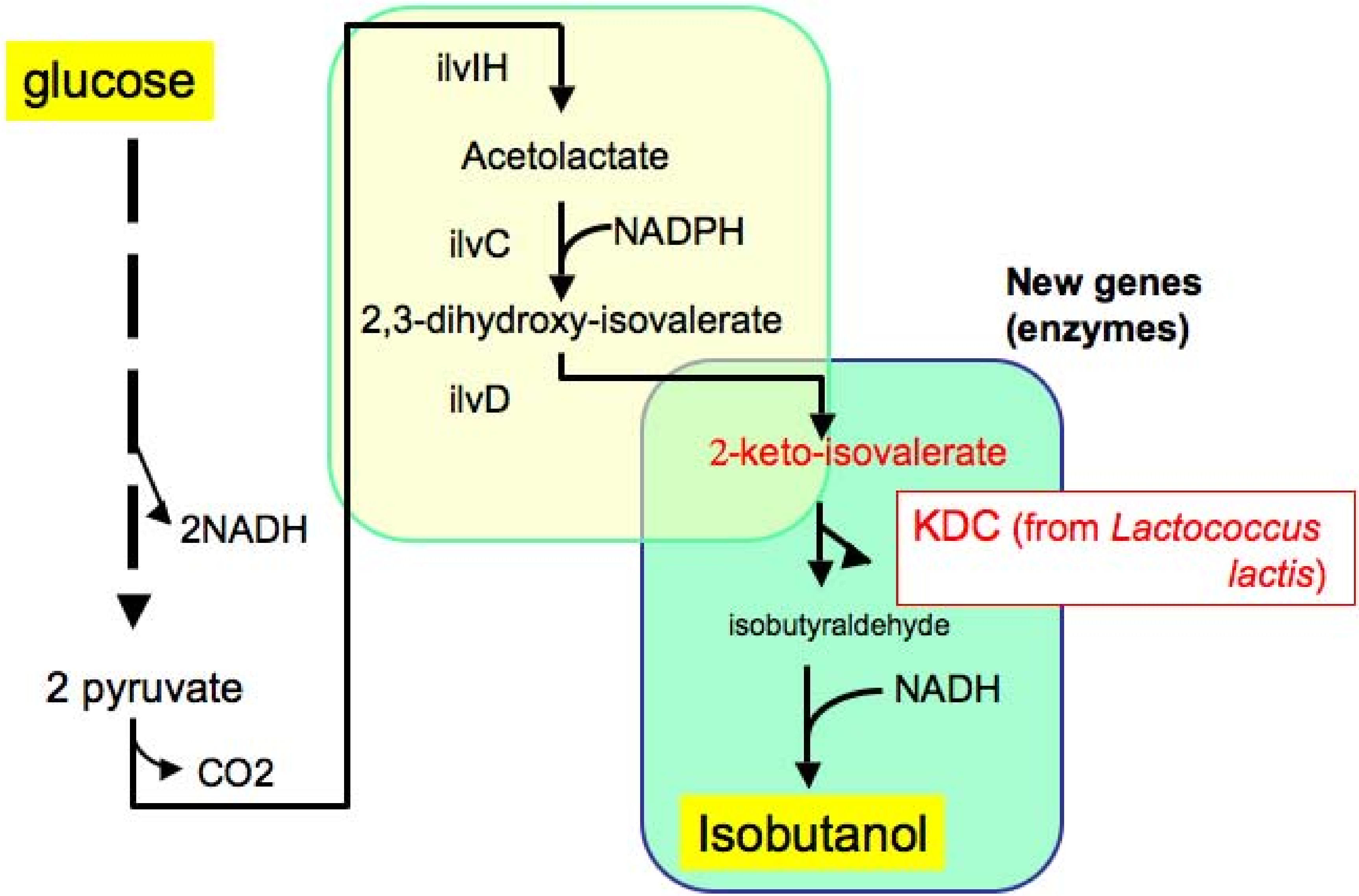
```
FCGGSLINSQWVVSAAHC
FCGGSLVNENWVVSAAHC
FCGGSLINENWVVTAAHC
TCGGTLIRQNWVMTAAHC
FCGGTLIHPSFVLTAAHC
WCGGSLLNANTVLTAAHC
LCGGVLVAEQWVLSAAHC
```

Programming *E. coli* to produce butanol



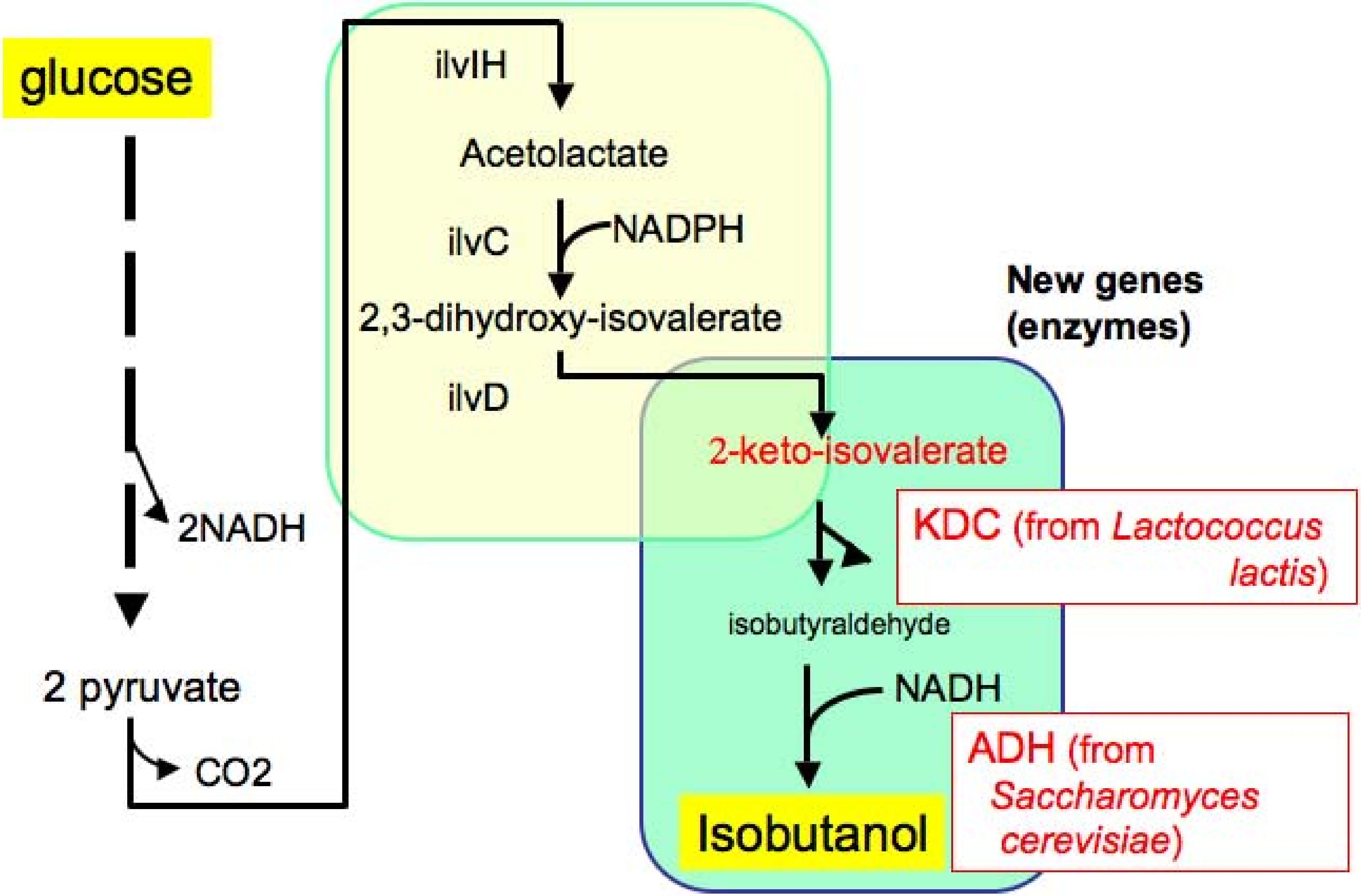
Atsumi et al. *Nature*, in press

Programming *E. coli* to produce butanol



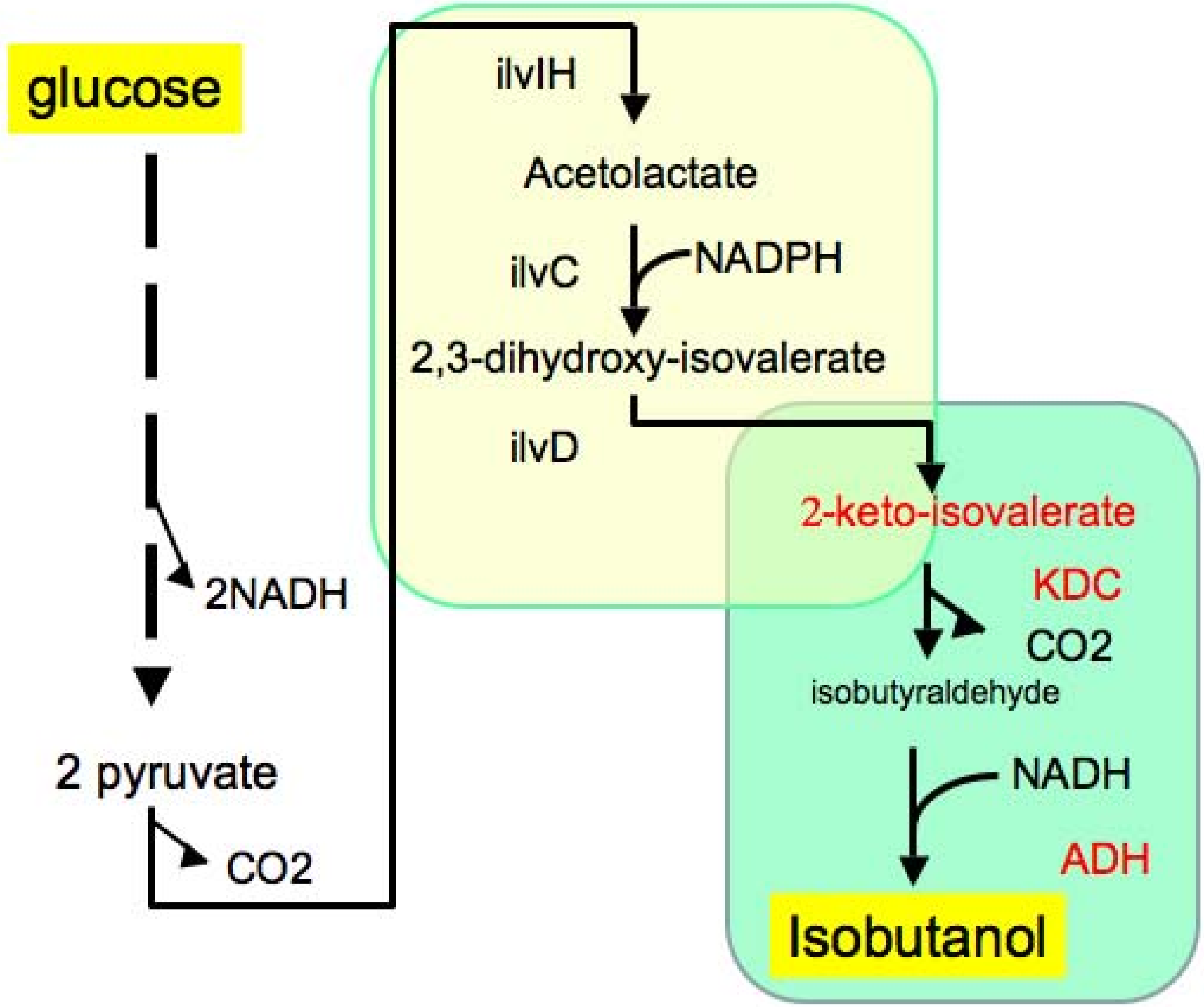
Atsumi et al. *Nature*, in press

Programming *E. coli* to produce butanol



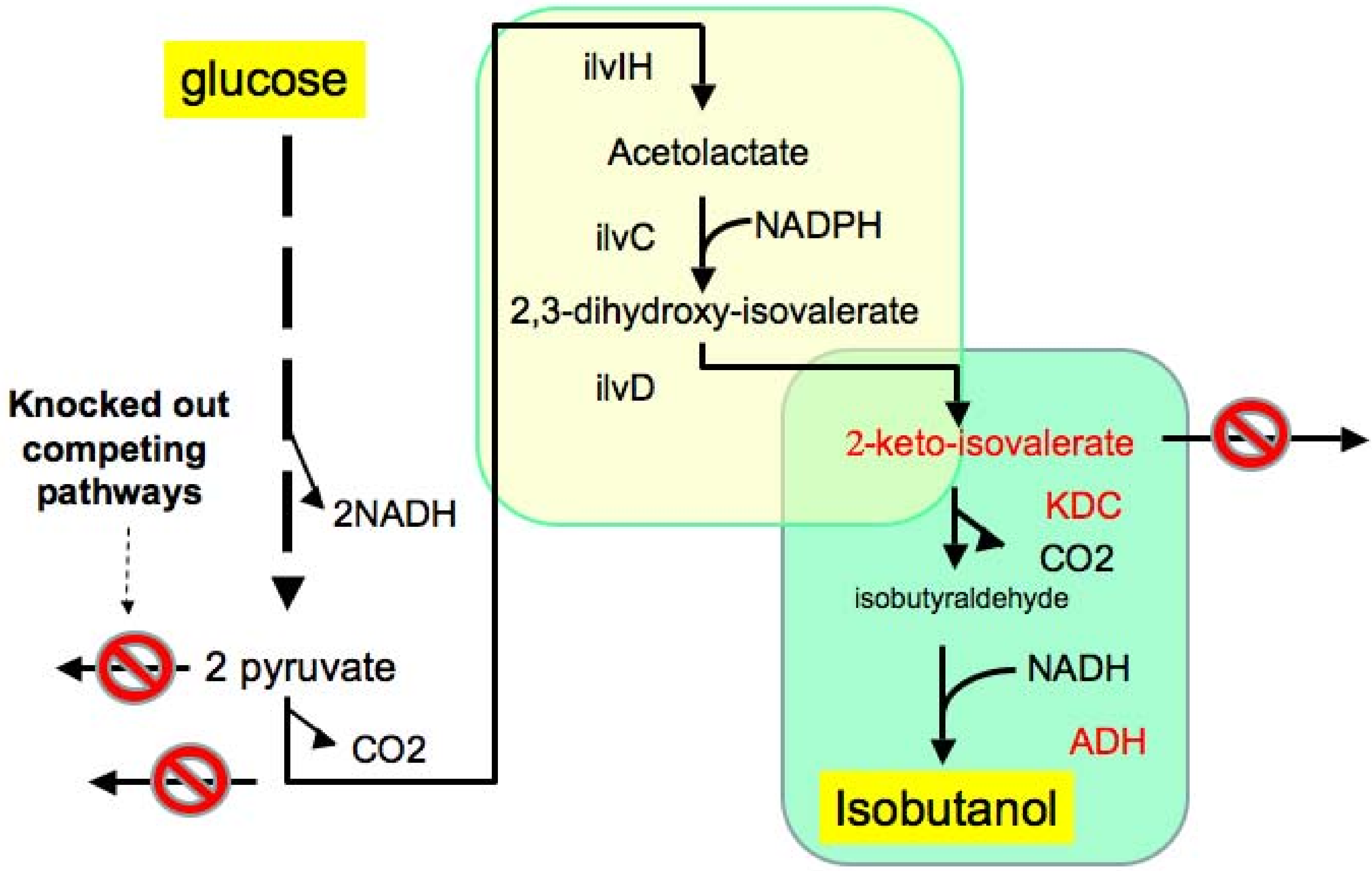
Atsumi et al. *Nature*, in press

Optimization of the new isobutanol pathway in *E. coli*



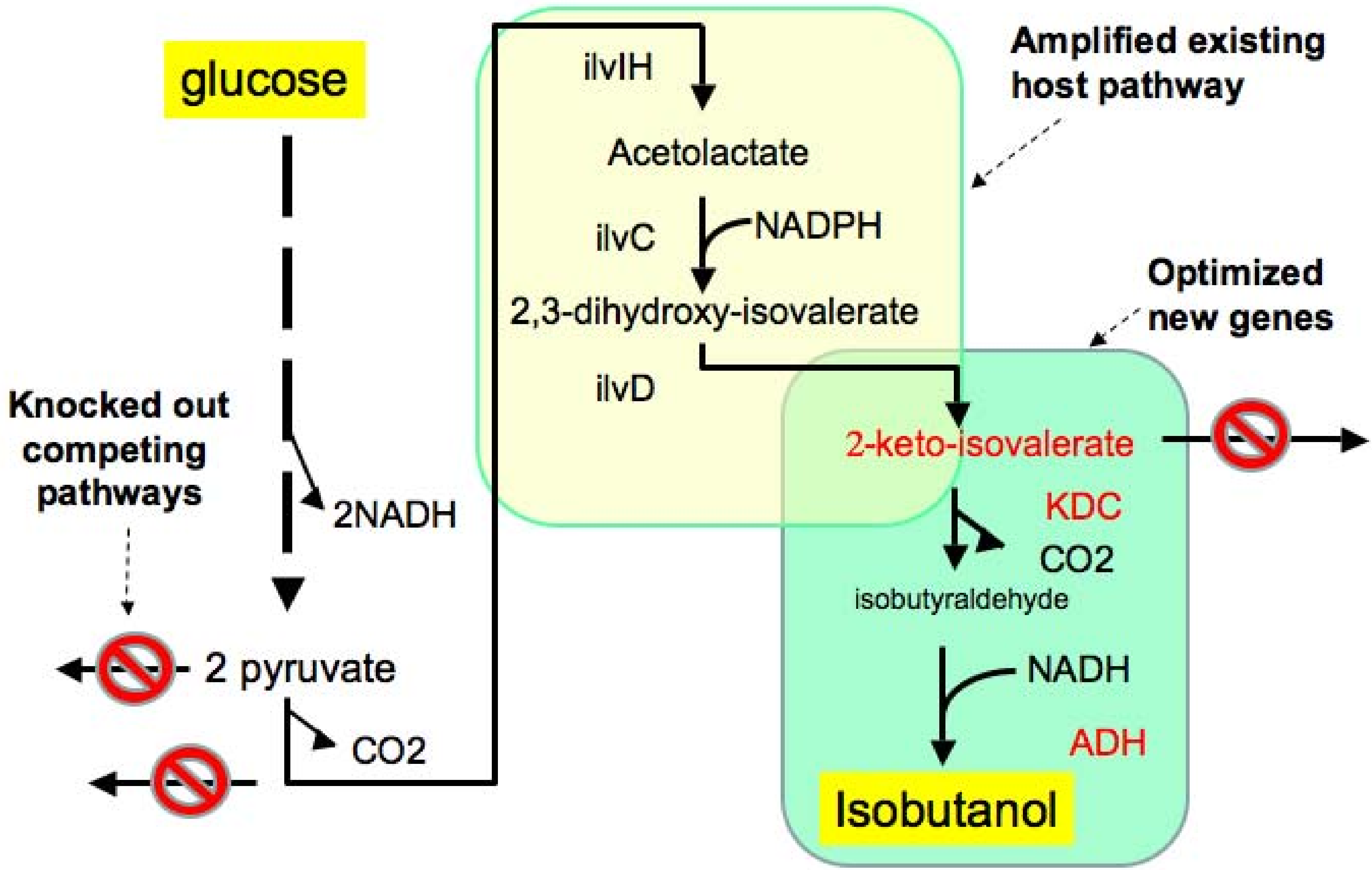
Atsumi et al. *Nature*, in press

Optimization of the new isobutanol pathway in *E. coli*



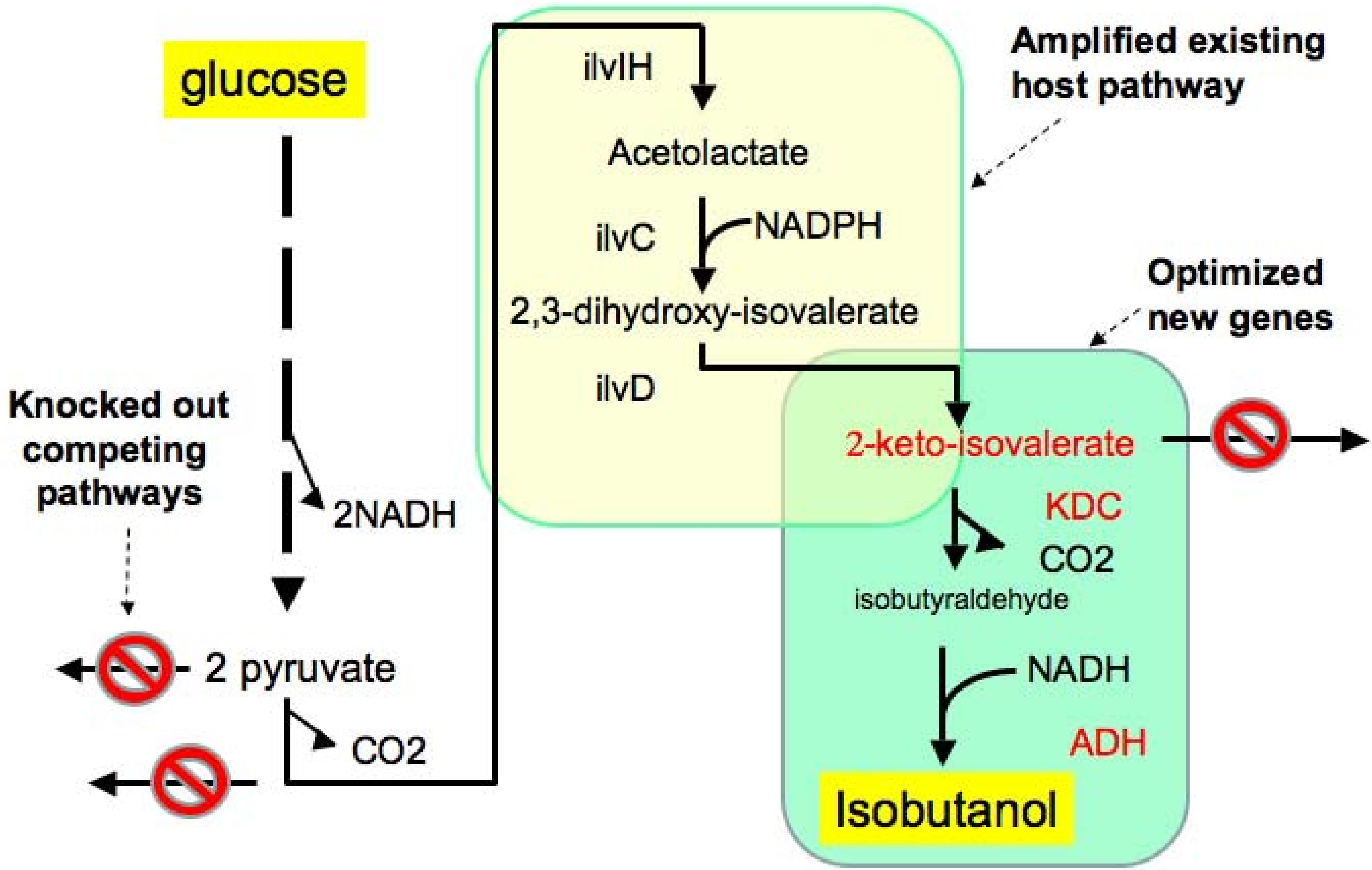
Atsumi et al. *Nature*, in press

Optimization of the new isobutanol pathway in *E. coli*



Atsumi et al. *Nature*, in press

Optimization of the new isobutanol pathway in *E. coli*

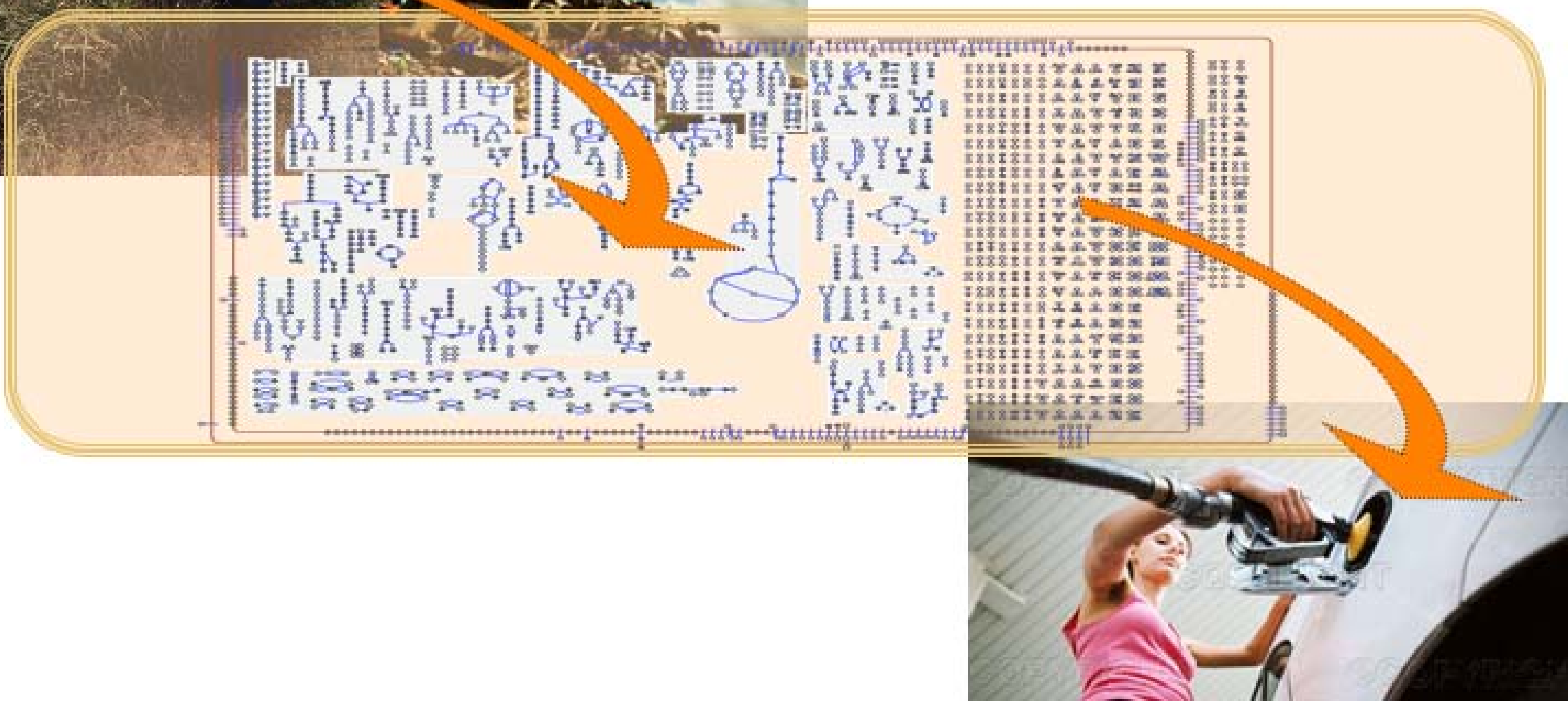


Atsumi et al. *Nature*, in press

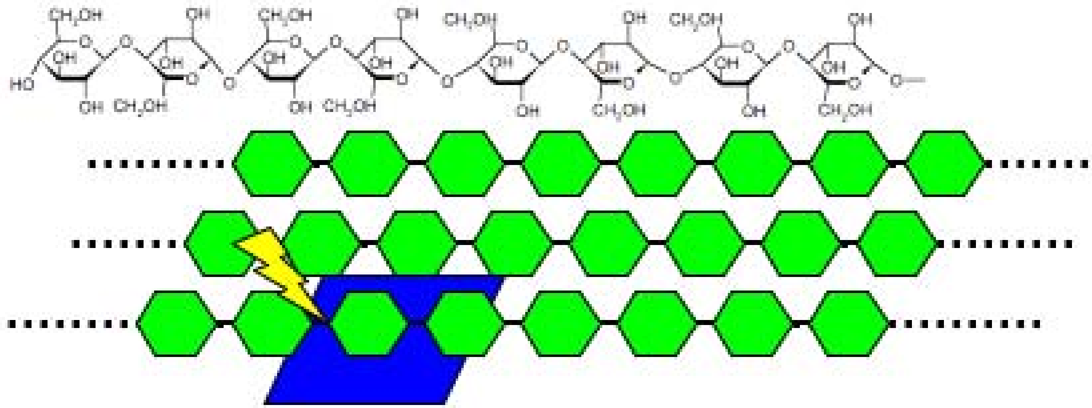
20 g/L and 86% of theoretical yield

Can we make a microbe that produces butanol directly from <u>cellulose</u>?

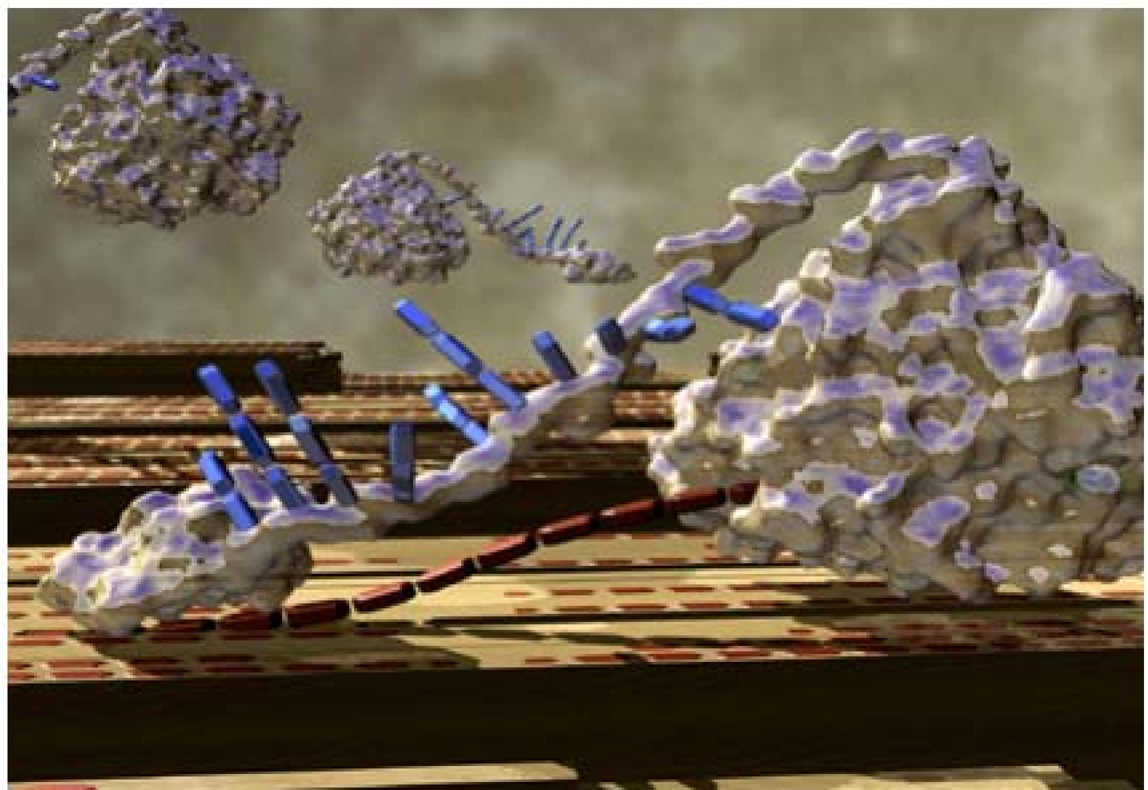
Now we have to add cellulose degradation capability.



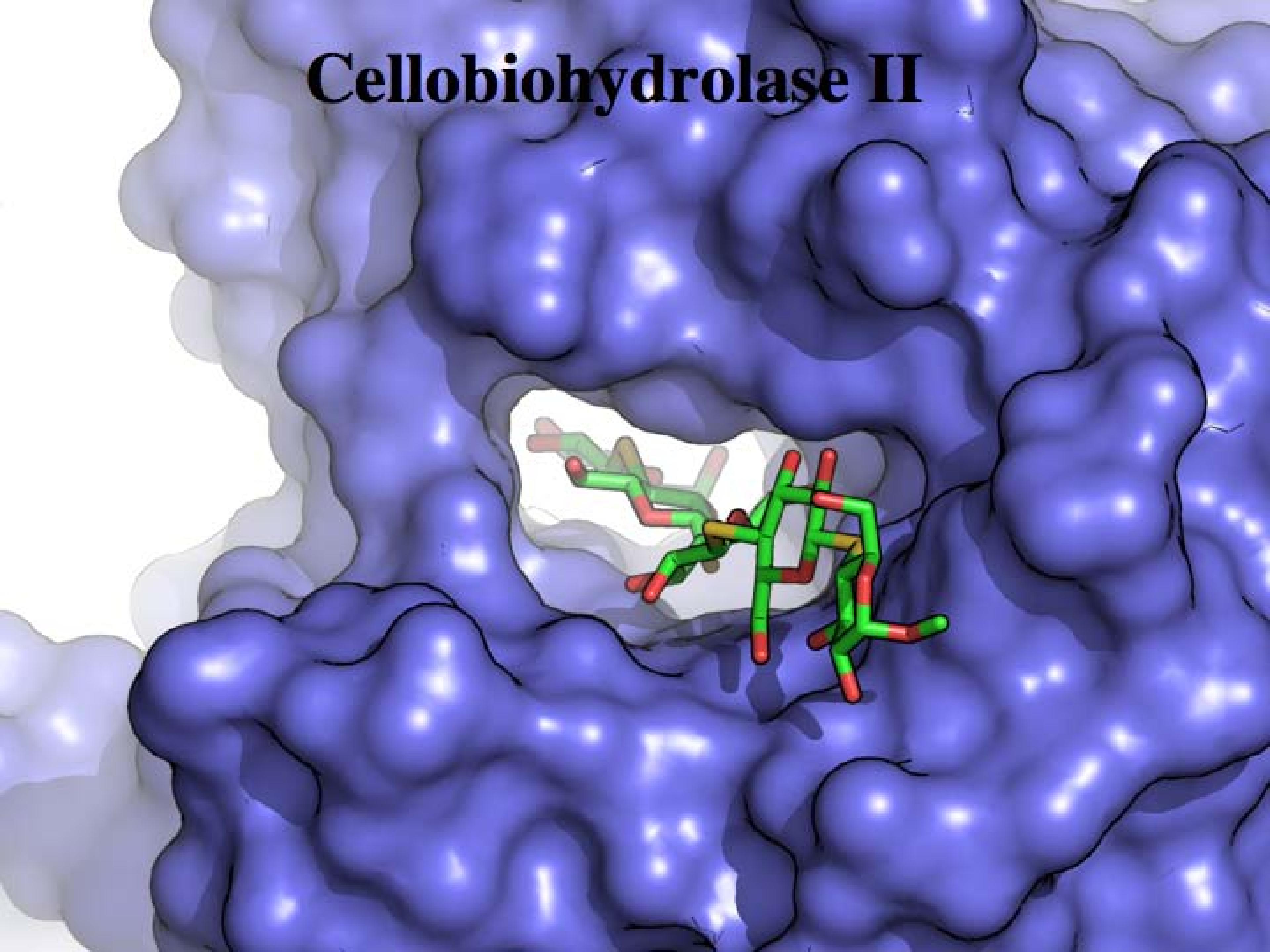
Complex molecular machines break down cellulose (white rot fungi figured it out, so did termites and many bacteria…)



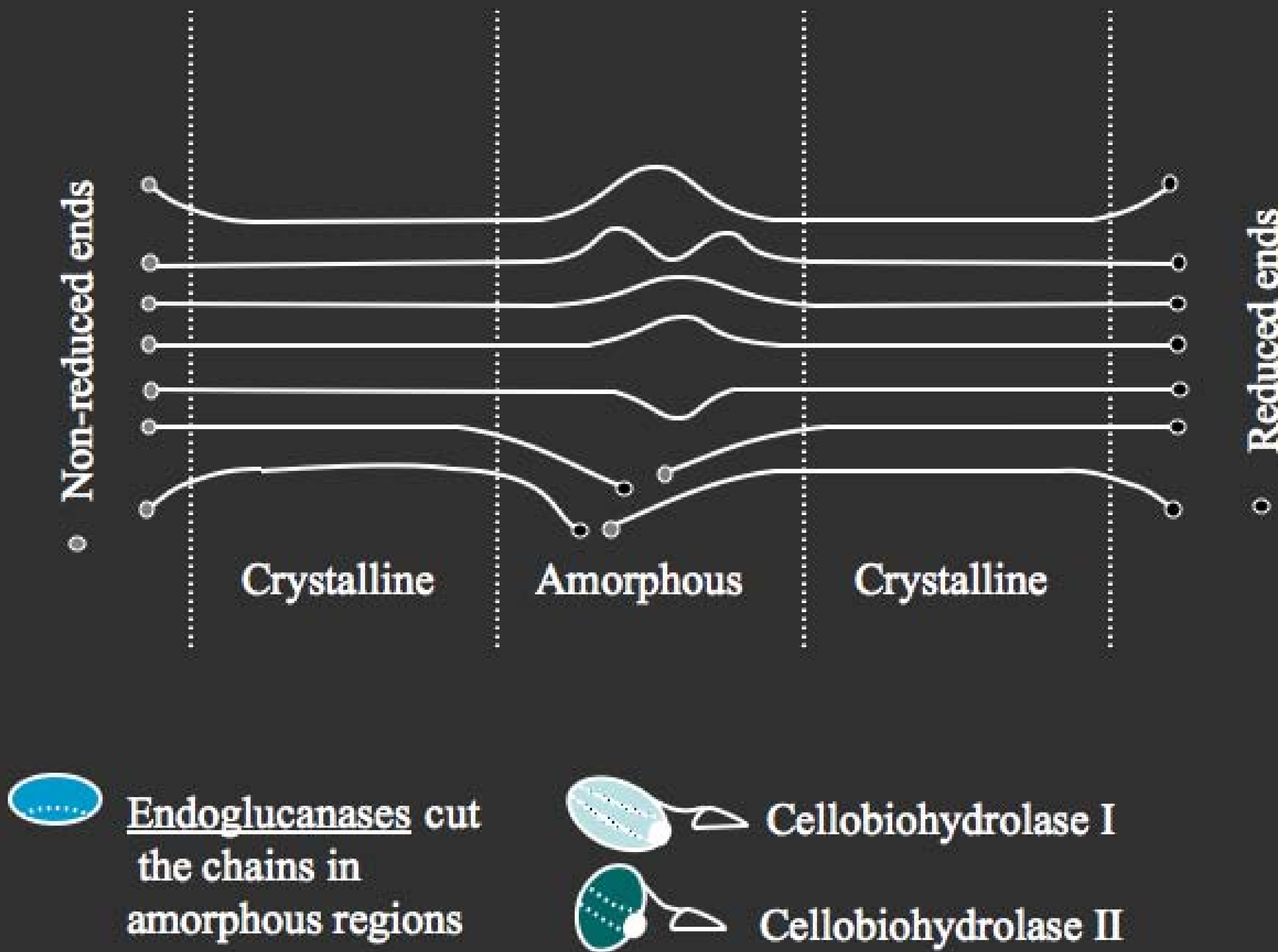
1/4 inch Milled Poplar



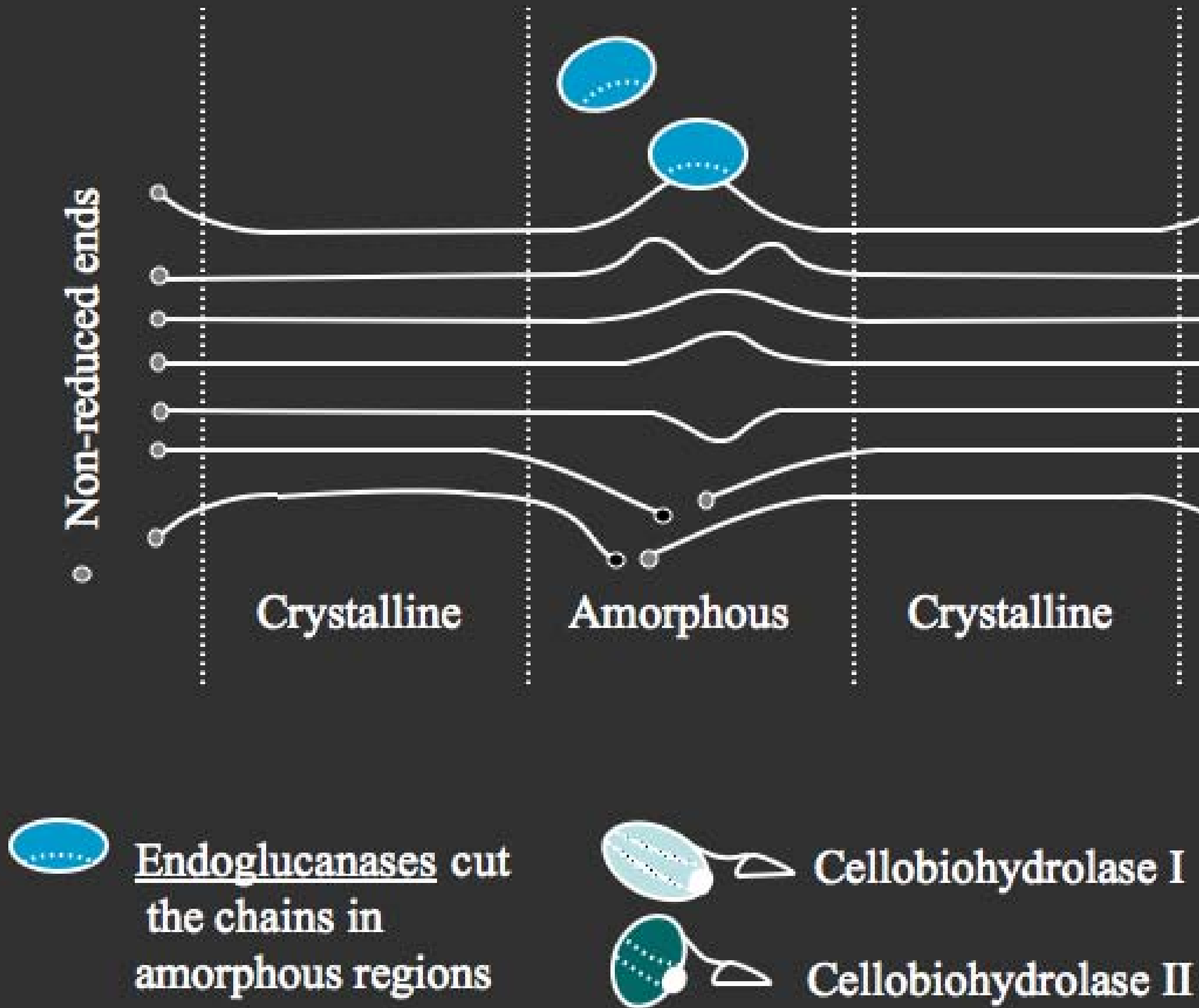
Cellobiohydrolase II



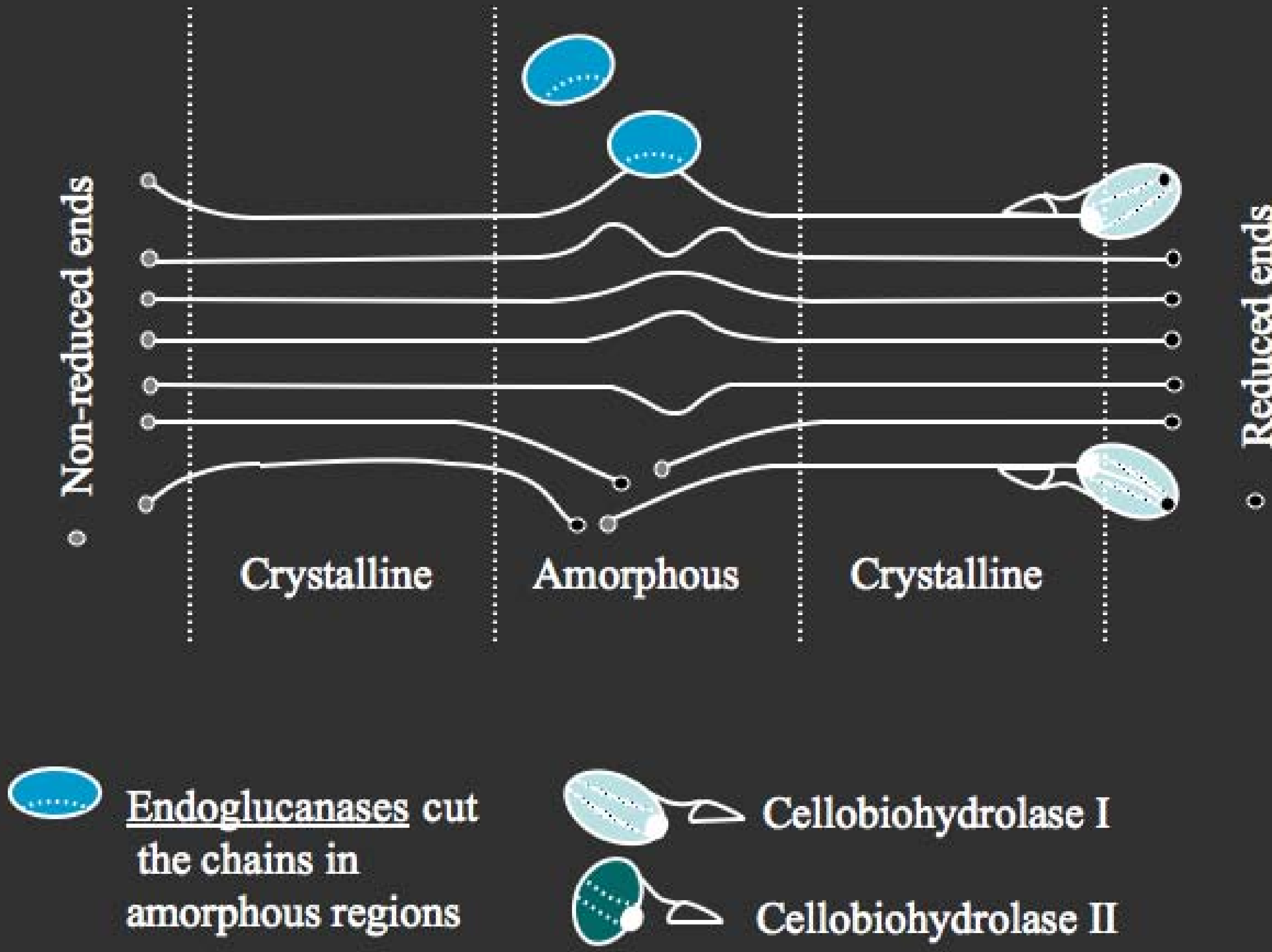
Cellulose degradation involves the simultaneous and synergistic action of multiple enzymes



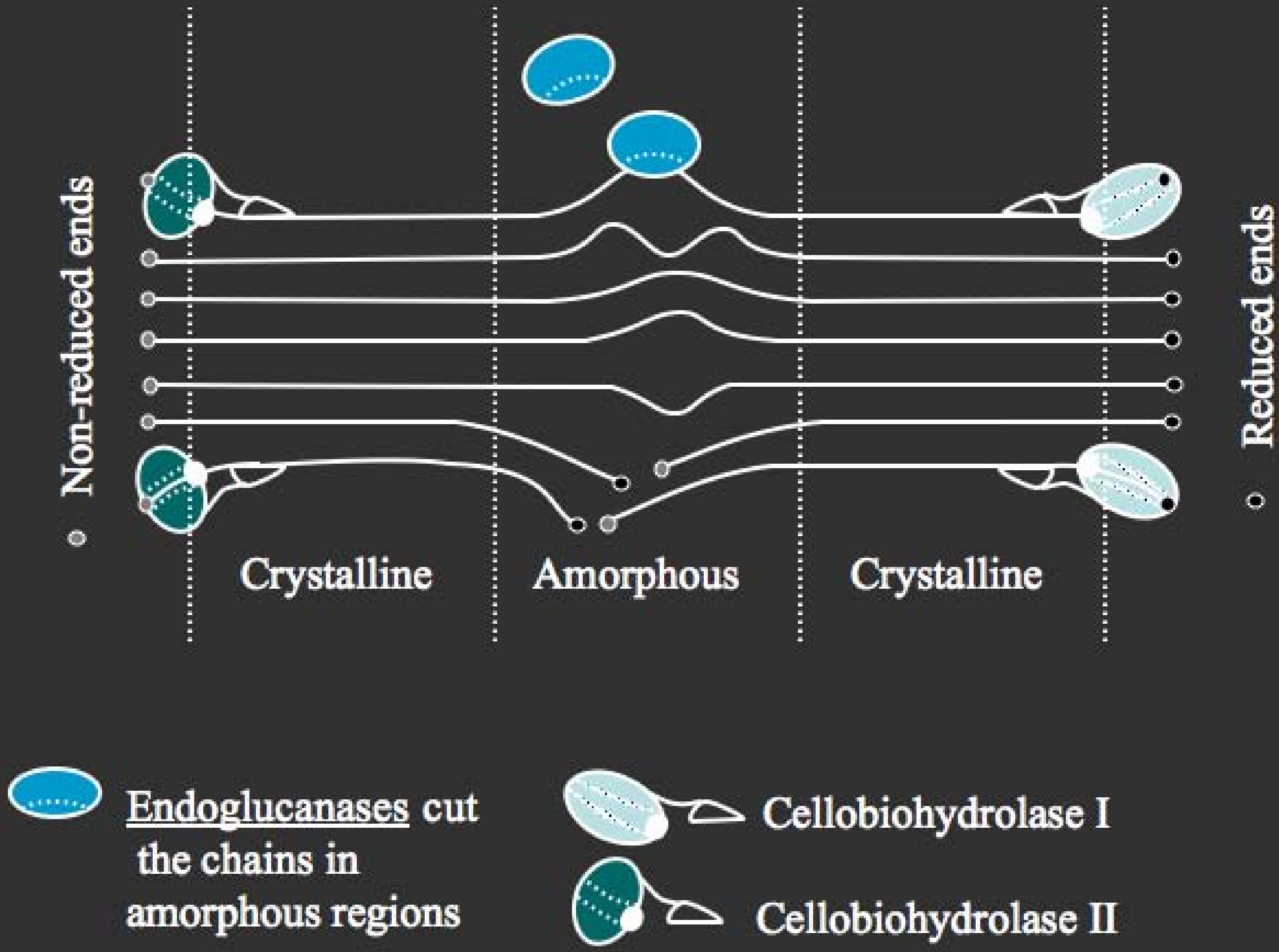
Cellulose degradation involves the simultaneous and synergistic action of multiple enzymes



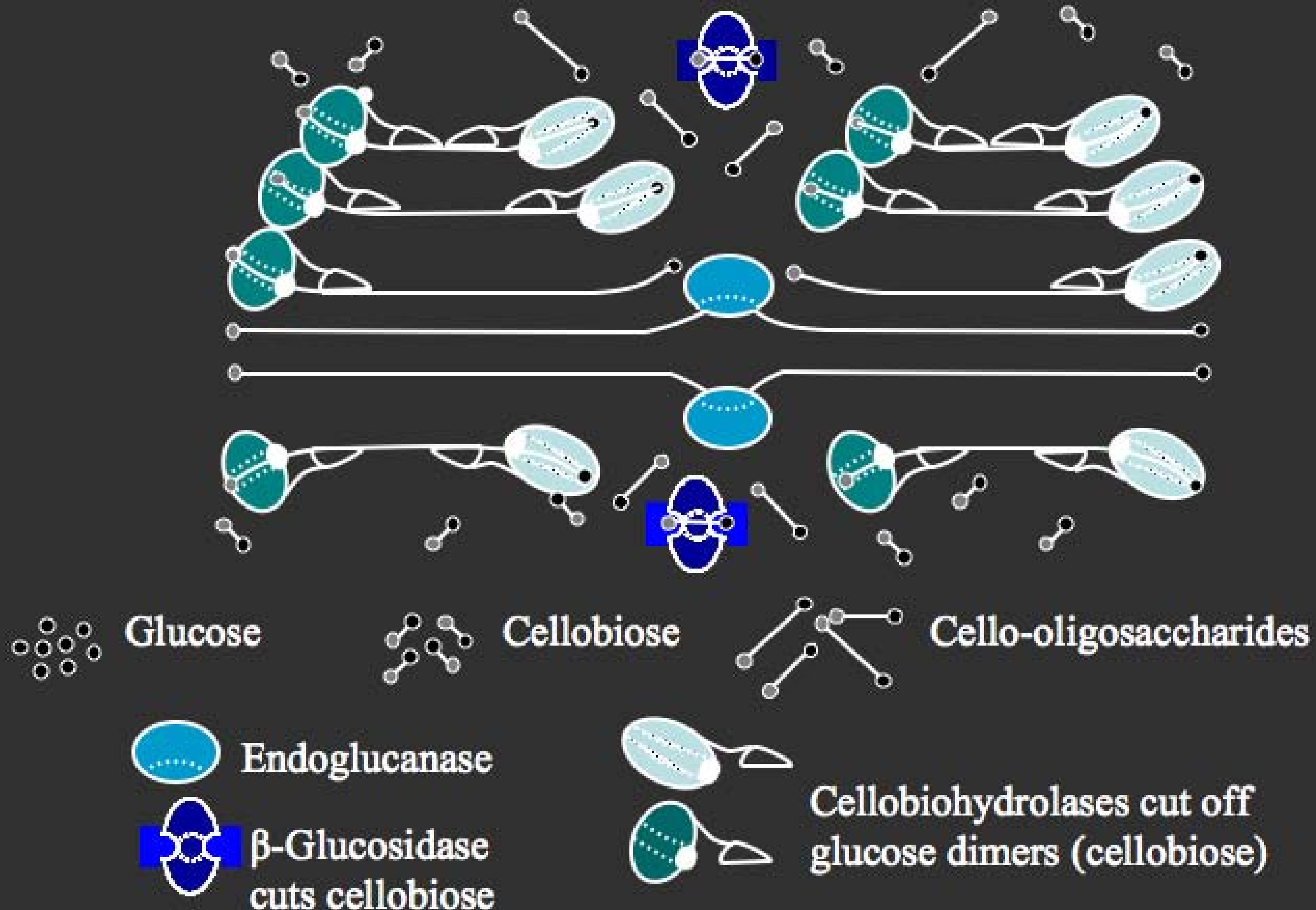
Cellulose degradation involves the simultaneous and synergistic action of multiple enzymes



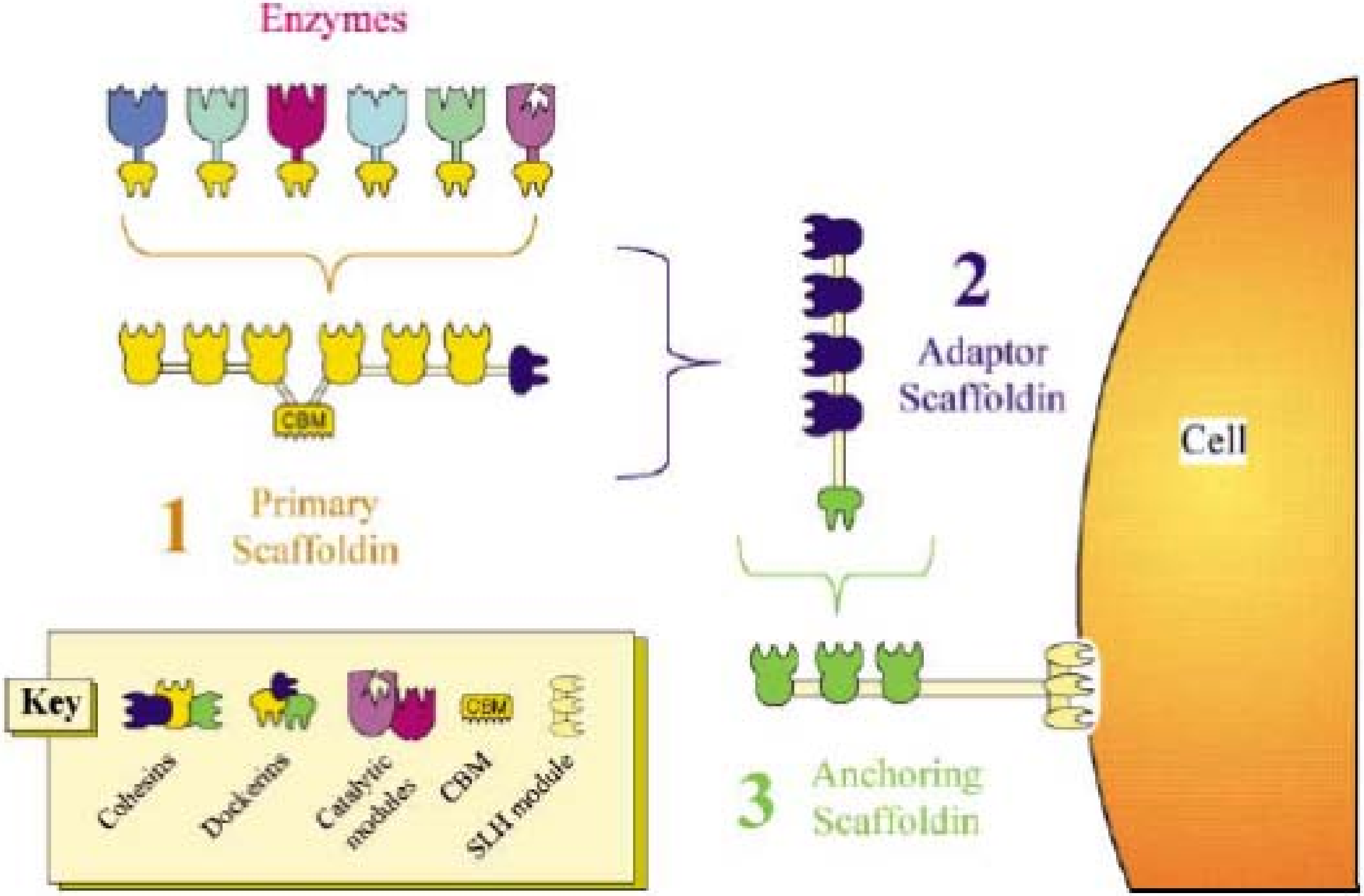
Cellulose degradation involves the simultaneous and synergistic action of multiple enzymes



Cellulose degradation involves the simultaneous and synergistic action of multiple enzymes

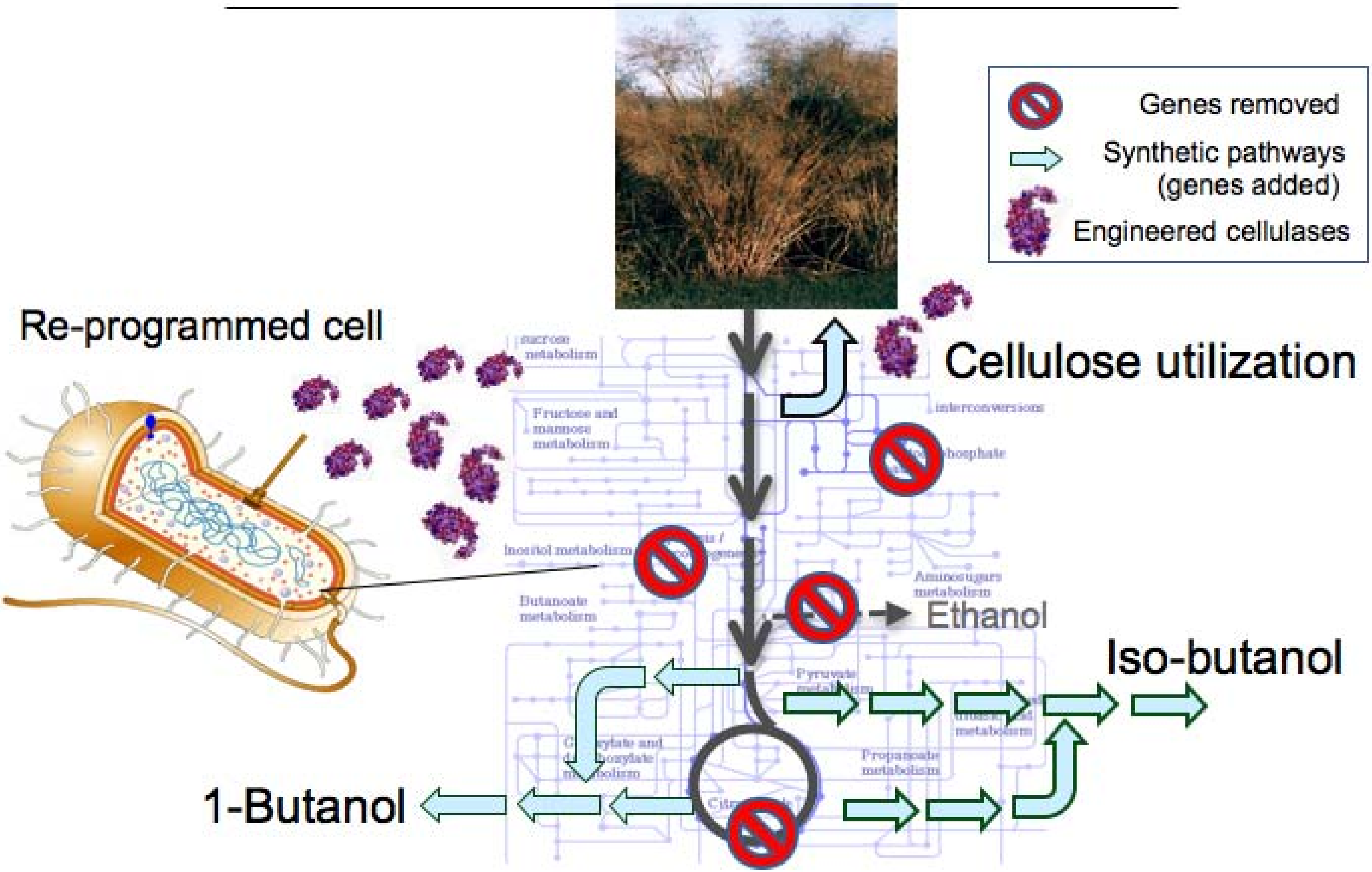


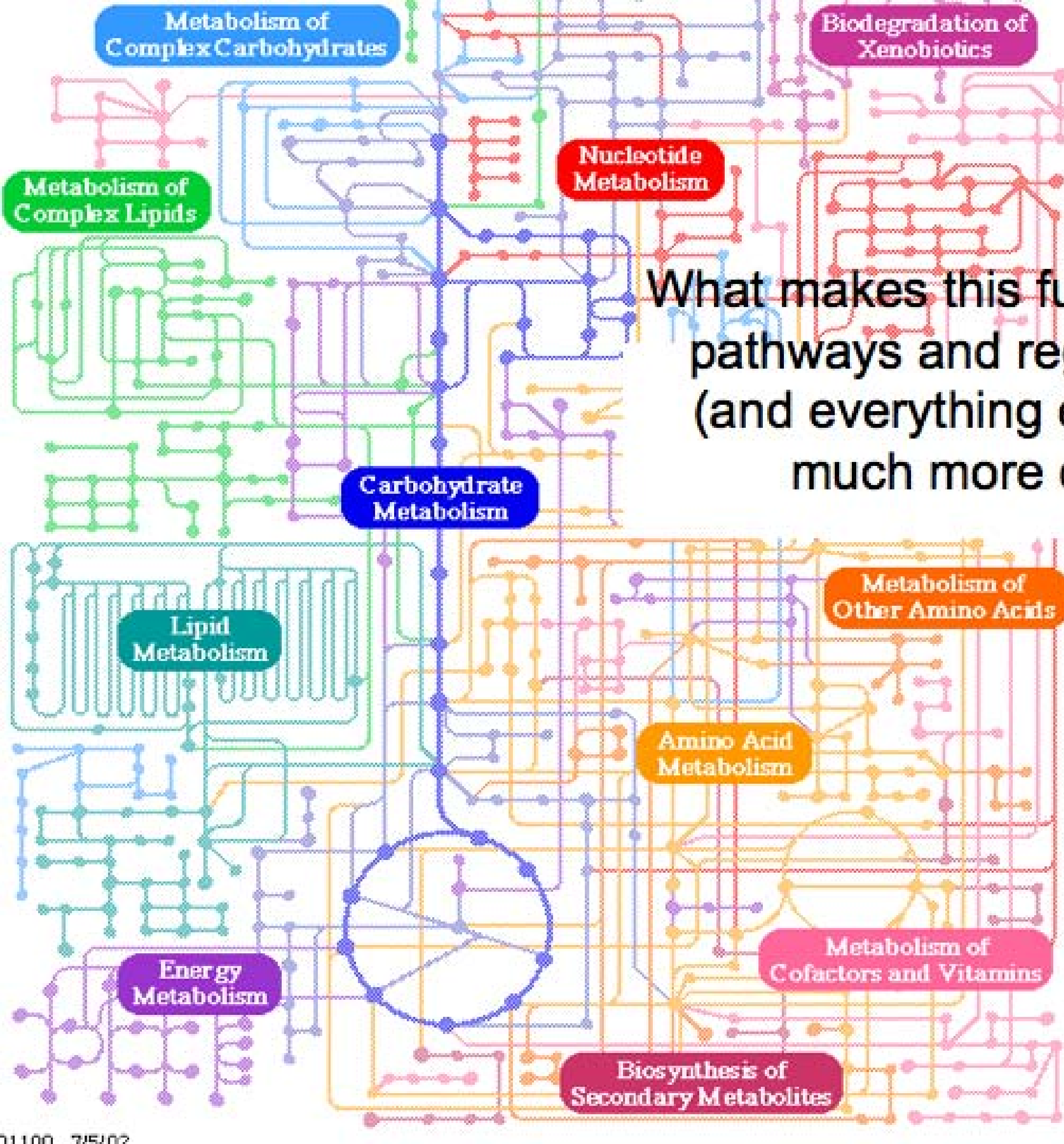
In some organisms all these components are assembled into ‘cellulosomes’ on the cell surface



Annu. Rev. Microbiol. 2004.58:521-554

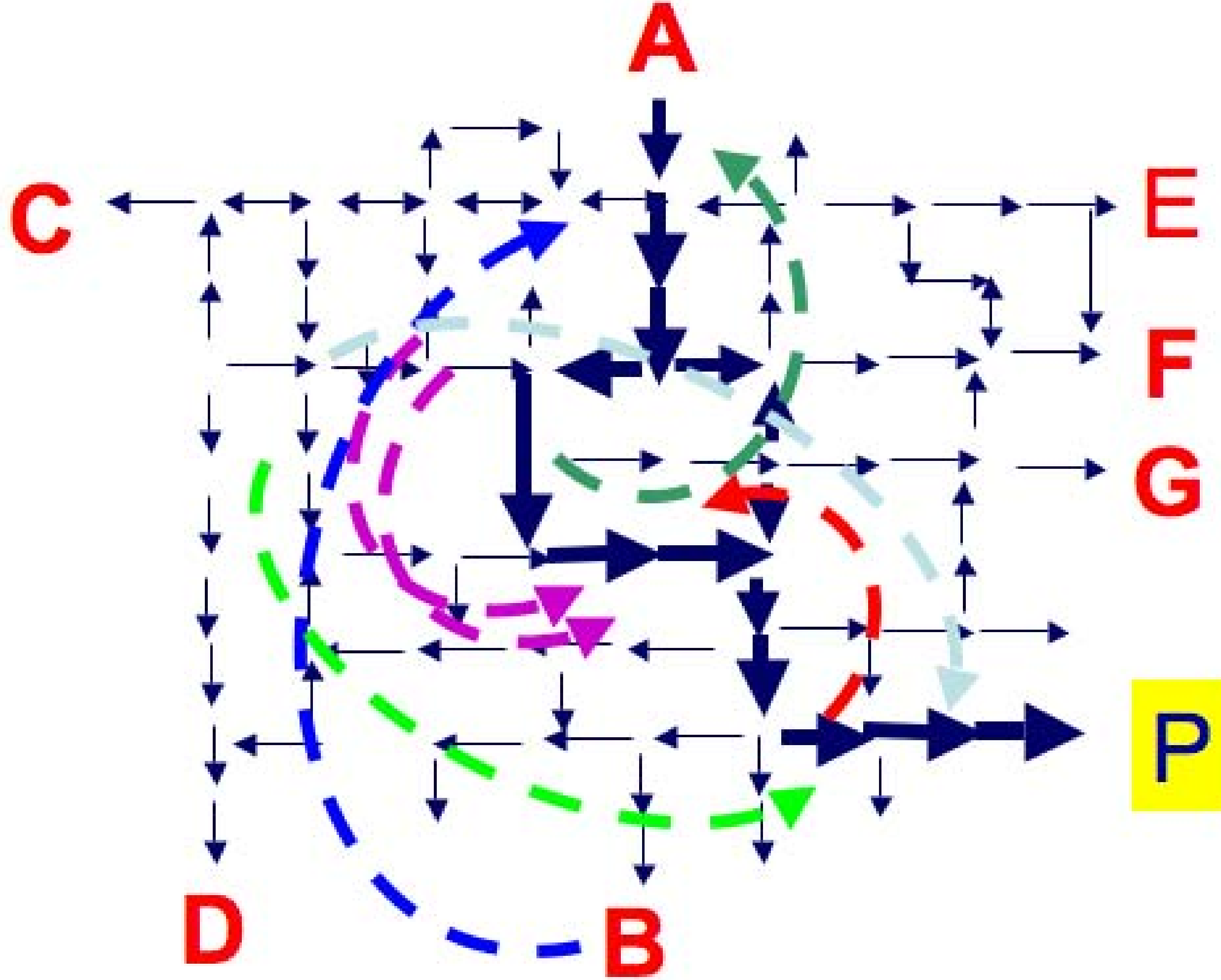
Caltech/UCLA synthetic biology challenge:
Construct a microbe that converts cellulose to a better biofuel





What makes this fun is that metabolic pathways and regulation in *E. coli* (and everything else) are actually much more complicated.

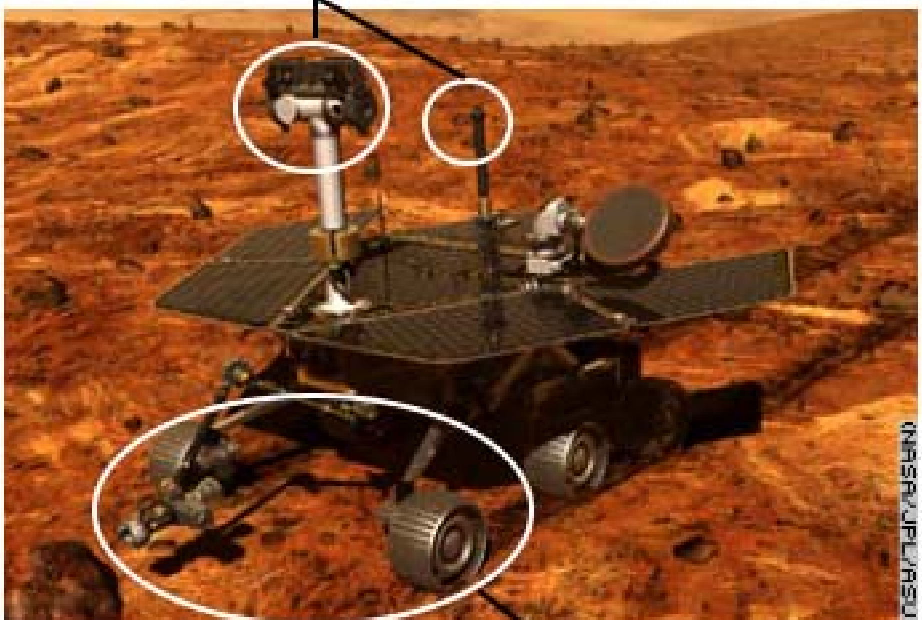
Intermediates and products feed back to regulate fluxes; multiple pathways interconnect

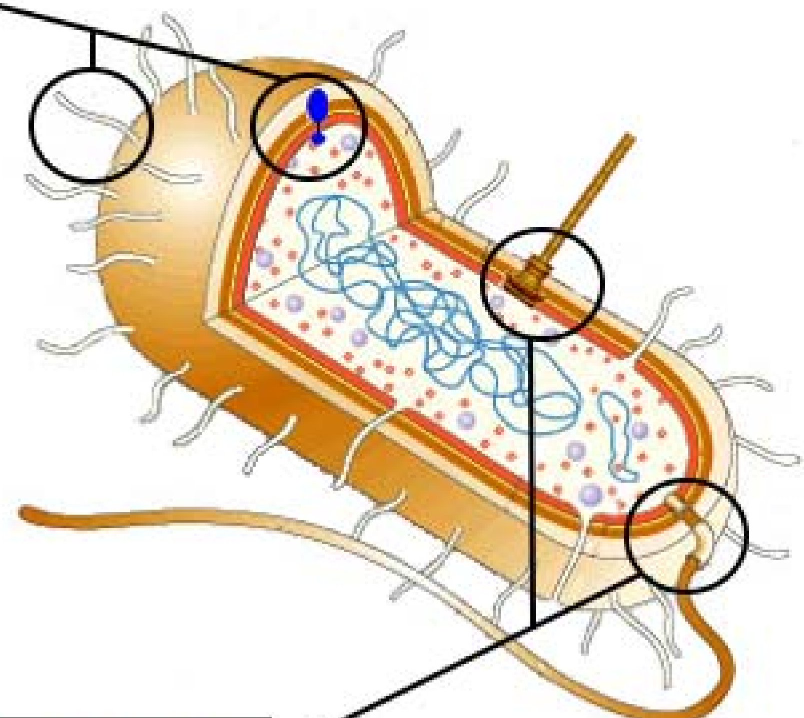


Bacteria as little programmable robots

Sensors (inputs)

- Temperature
- Light
- Remote control
- Chemicals
- Touch
- Communication





Actuators (outputs)

- Motility
- Secretion
- Self-organization
- Chemical production

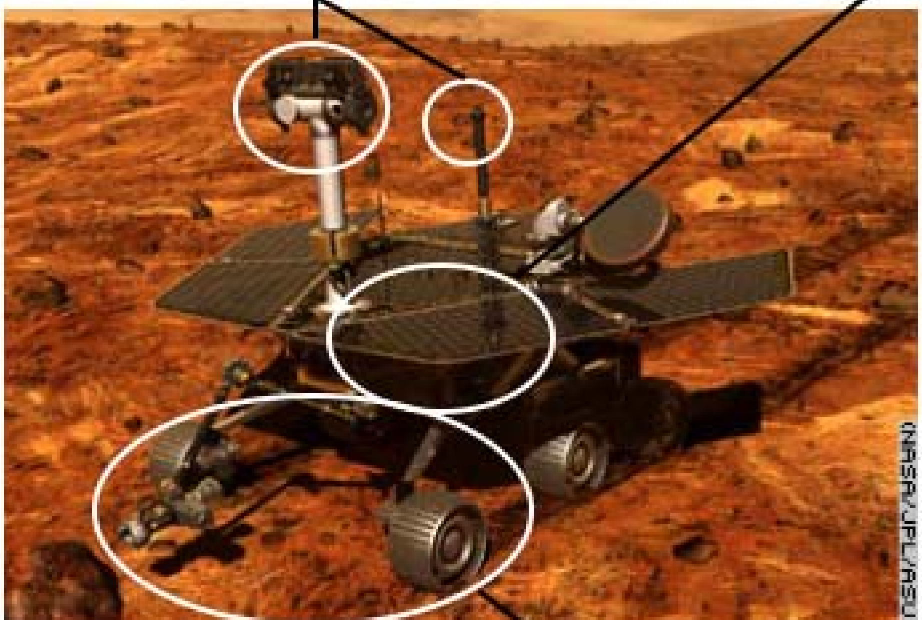
Bacteria as little programmable robots

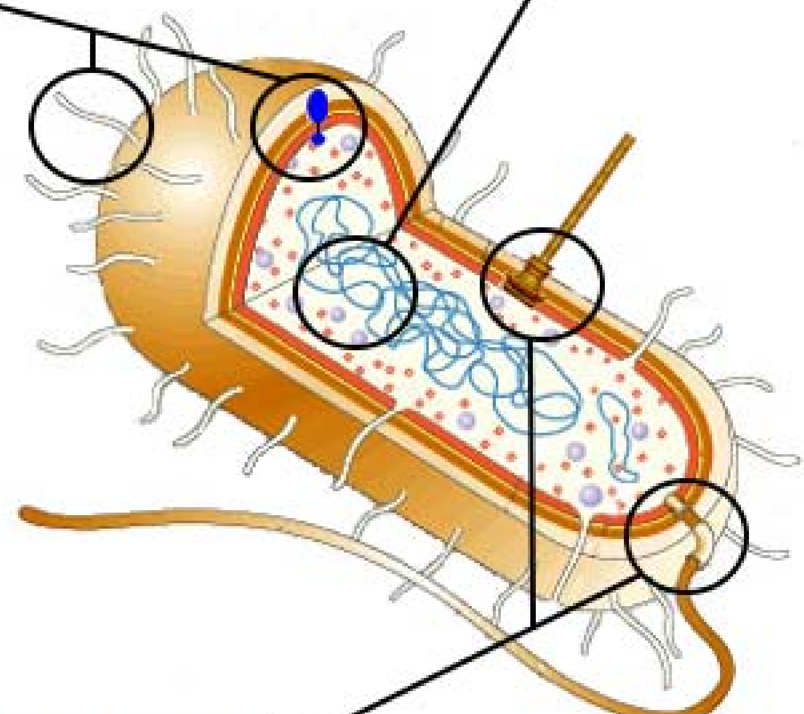
Sensors (inputs)

Temperature
Chemicals
Light
Touch
Remote control
Communication

Circuits

Signal processing
Logic gates
Dynamic Control
Memory

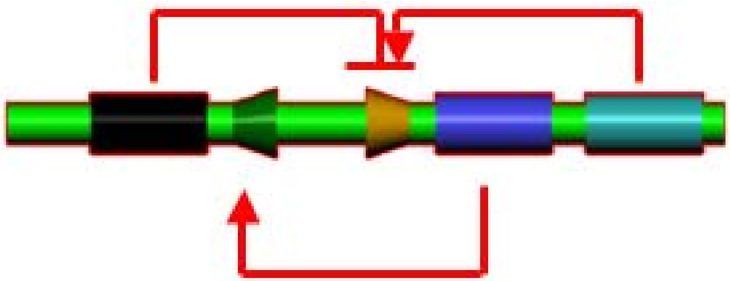




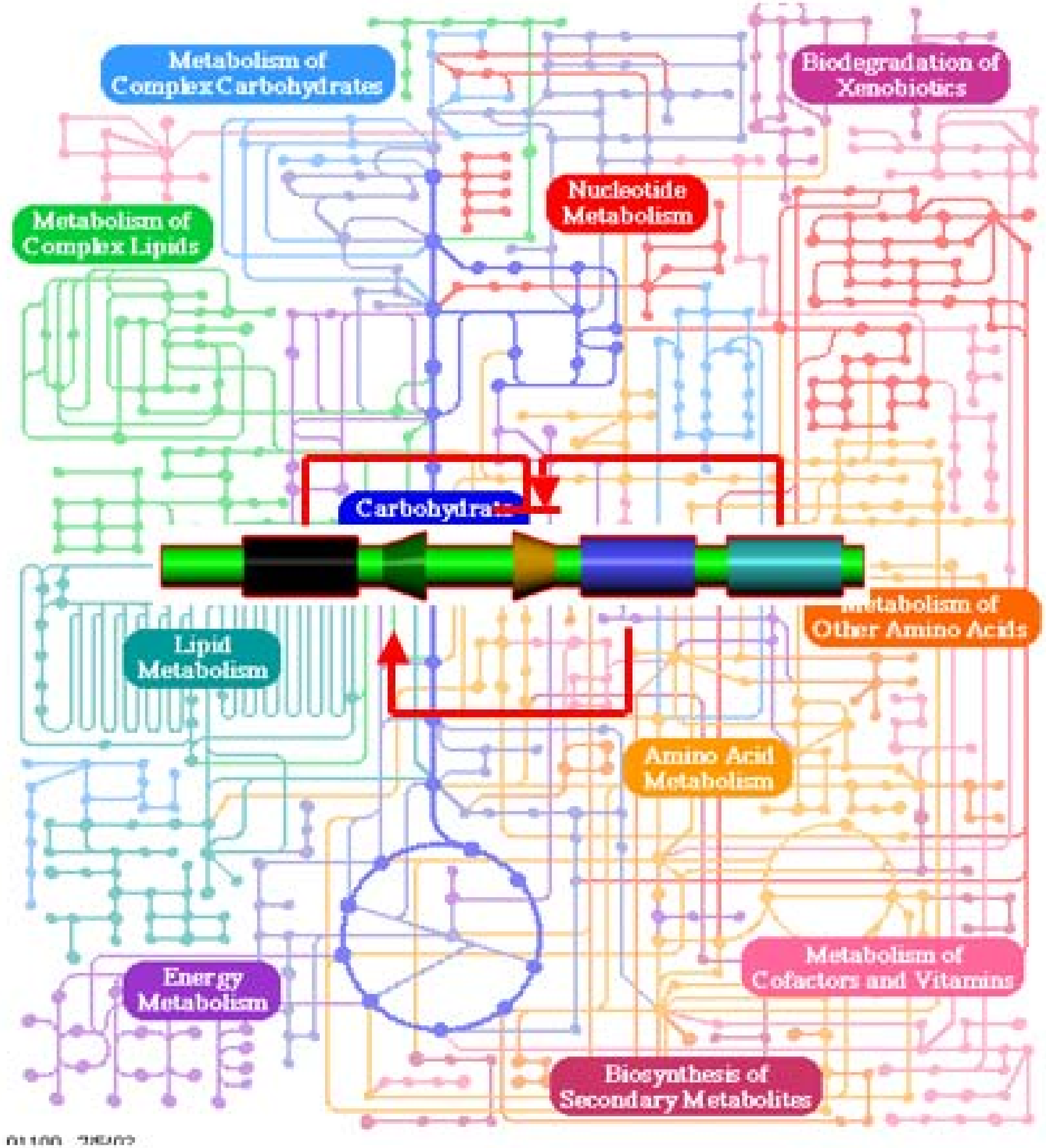
Actuators (outputs)

Motility
Secretion
Self-organization
Chemical production

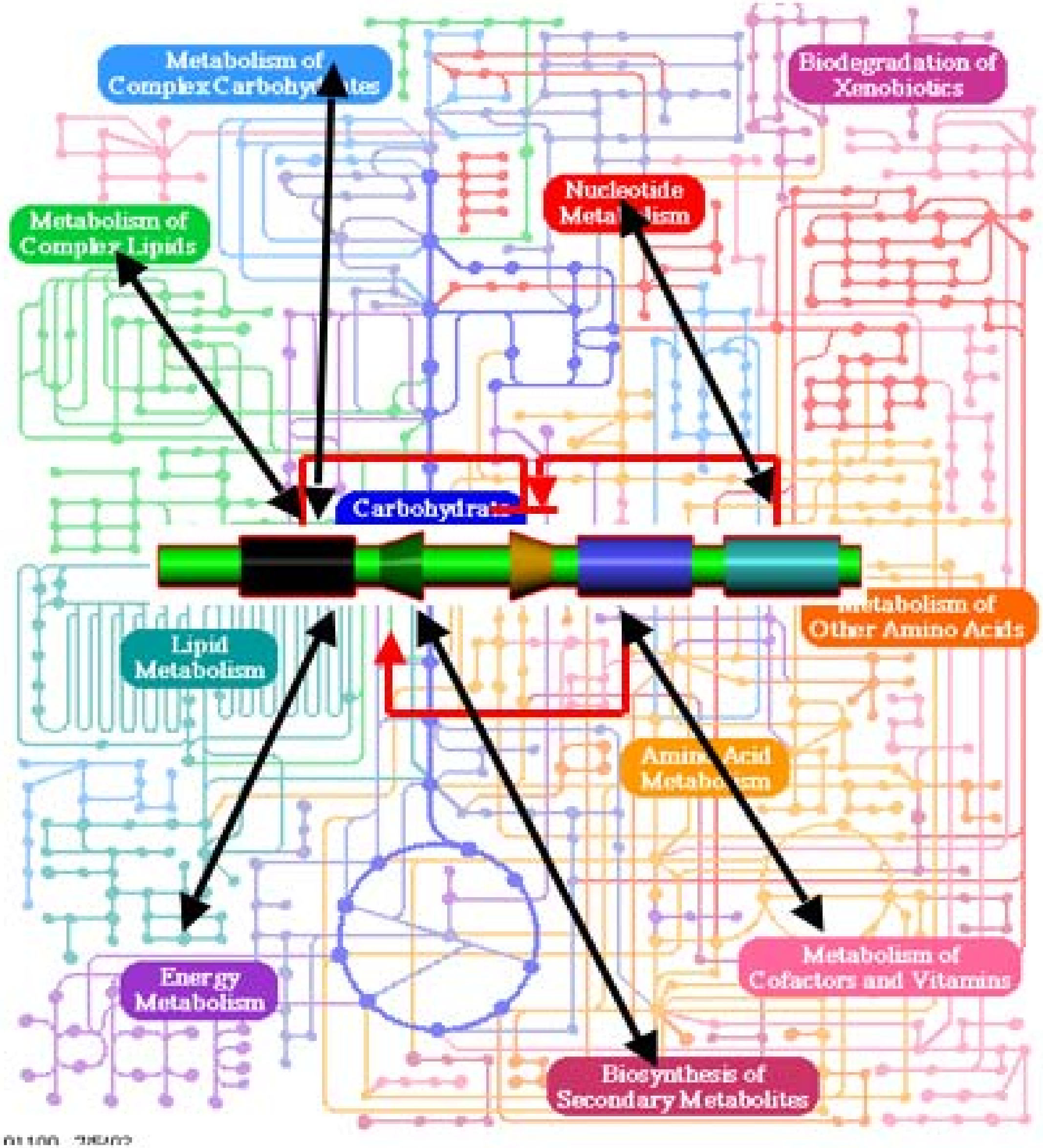
Integrated gene-metabolic circuits



Integrated gene-metabolic circuits



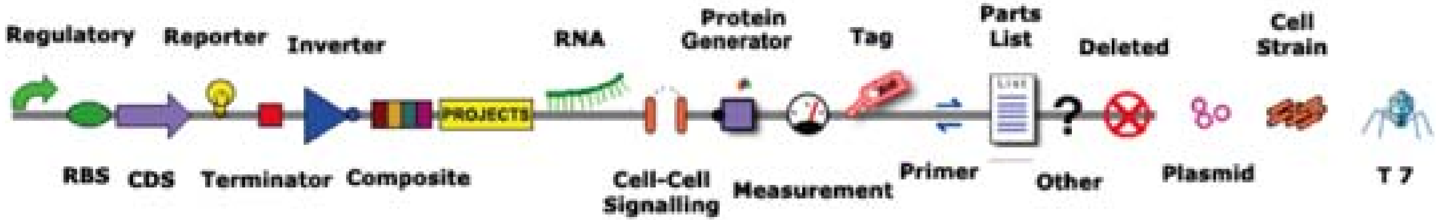
Integrated gene-metabolic circuits



Registry of Standard Biological Parts

MIT

Registry Parts Index



The Registry of Standard Biological Parts contains hundreds of parts. To see a list of all these parts (except deleted and intermediate) click on the Parts List icon on the Part Selection Bar above. It will normally be more convenient to view a selection table of similar parts. For example, clicking on the red Terminator icon will give you a table of transcriptional terminators, their part names, performance and status. To see details about a particular part, click on its part name in the table or anywhere you see a colored part name. If you know a part name and want to go directly to that part, just type it into the 'Jump to Part' box in the menu on the left of these pages.

The physical Repository currently contains 227 basic parts and 457 composite parts. 267 parts are currently being synthesized or assembled. 922 parts are specified and may or may not ever become available. The database contains a total of 2056 part entries.

Part Types

Basic parts, devices, and systems are all stored as parts in the registry. Parts can be combined into more complex parts, devices can be combined into more complex devices and one designer's system may be a device in another designer's system. However, there is an important distinction between basic parts and all composite parts. Basic parts have their own sequence information while the sequence information for all other parts is derived from their component parts. Given this definition, a part is not allowed to contain both its own sequence and other parts.

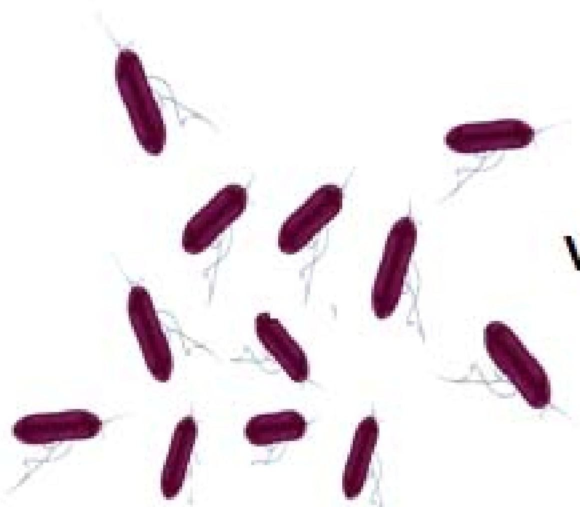
Each part has a part type field matching one of the entries below. Normally, the part name contains the letter associated with the part's type. Confusion is possible when a part fits into multiple categories. We deal with this on a case-by-case basis. The special part type 'Intermediate' is generated by the automatic assembly program and rarely created by a user.

Basic Part Types			Composite Part Types		
Type		Description	Type		Description
R	Regulatory	Operator region	E	Reporter	Compound reporter devices
B	RBS	Ribosome binding site	Q	Inverter	Inverter and logic
C	CDS	Protein coding sequence		Composite	Other composite parts
B	Terminator	Transcriptional terminator	I	Project	Student projects
	RNA	RNA binding sites and coding	G	Generator	TIPS-to-Protein converter
F	Signalling	Cell-cell signalling		Measurement	Performance measurement constructs
E	Reporter	Basic reporter CDS	T	Temporary	Temporary and trial parts

We can insert new code into cells, and they will 'read' it…

<u>But</u>…we often don't know what to write.

(DNA sequence $\xrightarrow{?}$ behavior)



versus

…AATAGCCGTTATTTCC
ATATATTTAATTGGCCG
GAGAGTCTCCCGCGCG
ACATAAGGAGTCCTCG
TTTCGAGATACGTACG
GCATGGTGACACCAGT
TGCCCTCTGATTCCCG
GAGCCTCTTTGAAAAC
GTCGAGTCGAATCGAA
GTTCGAACCCCGGATC
GGGTCCACCAACTTAG
AGATGTGTGTGCGCTG
ACTCAGTCAT…

IgNobel abstract: clear summary that anyone can understand, in seven words or less

(E. Lander) Human Genome Project:

Genome. Bought the book. Hard to read.

IgNobel abstract: clear summary that anyone can understand, in seven words or less

(F. Arnold) Synthetic Biology

Genome. Great story! Hard to write...

IgNobel abstract: clear summary that anyone can understand, in seven words or less

(F. Arnold) Synthetic Biology

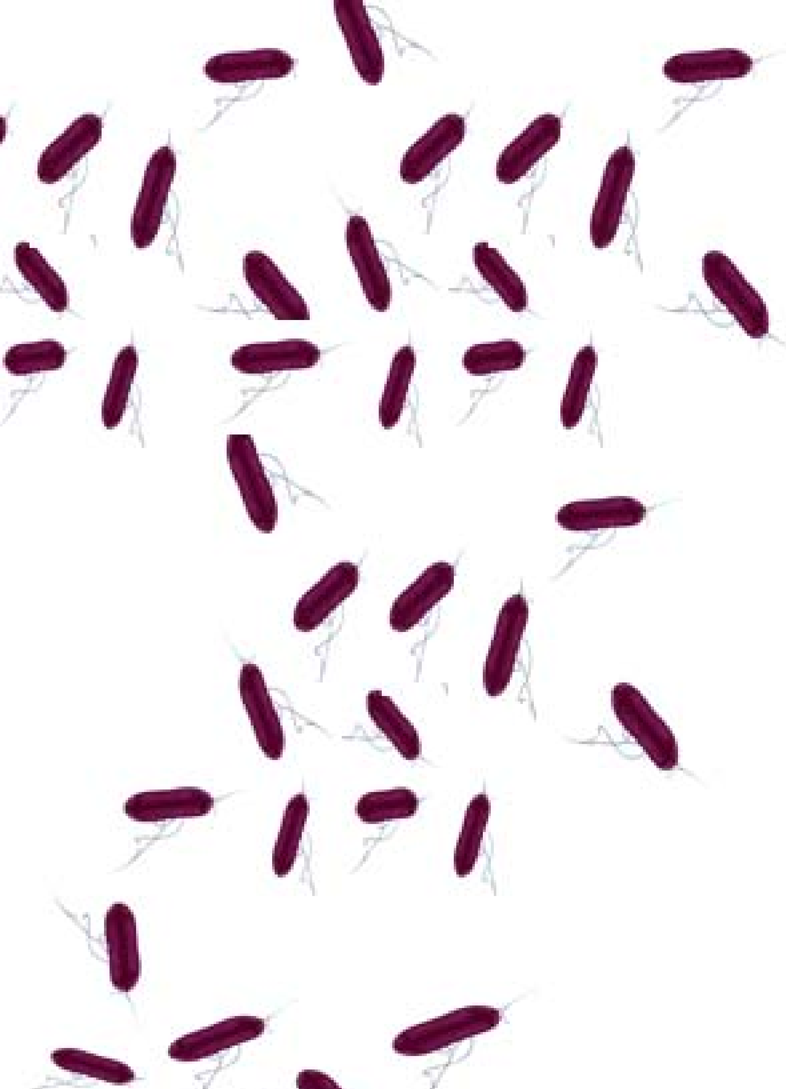
Genome. Great story! Hard to write...

IgNobel abstract: clear summary that anyone can understand, in seven words or less

(F. Arnold) Synthetic Biology

Genome. Great story! Hard to write...

Get a good editor!

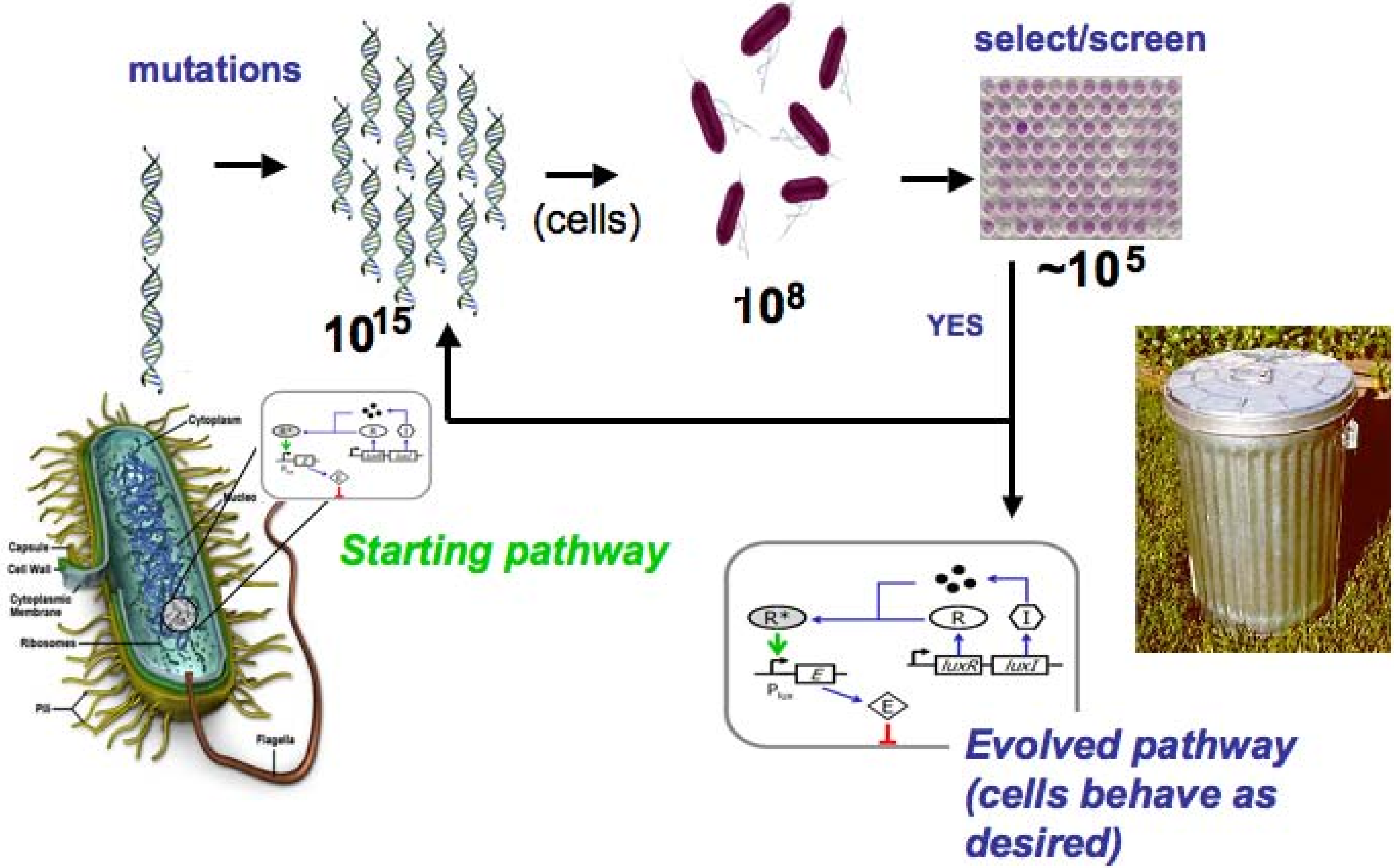


Biology is massively parallel: billions of these in 1/1000th of a liter.

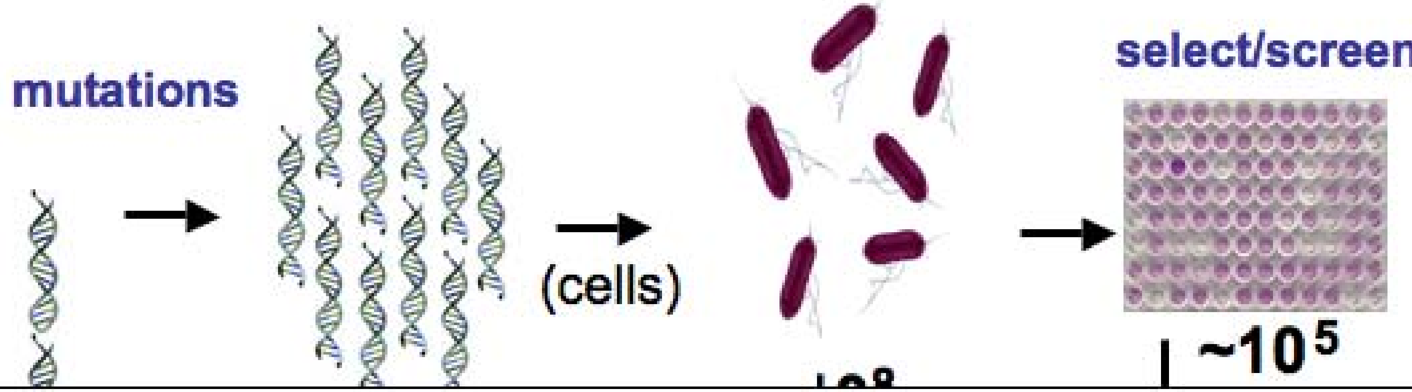
Sometimes we can even set it up so that the one that solves the problem gets to reproduce (‘directed evolution’).

The best editor: evolution

Biological optimization by directed evolution:



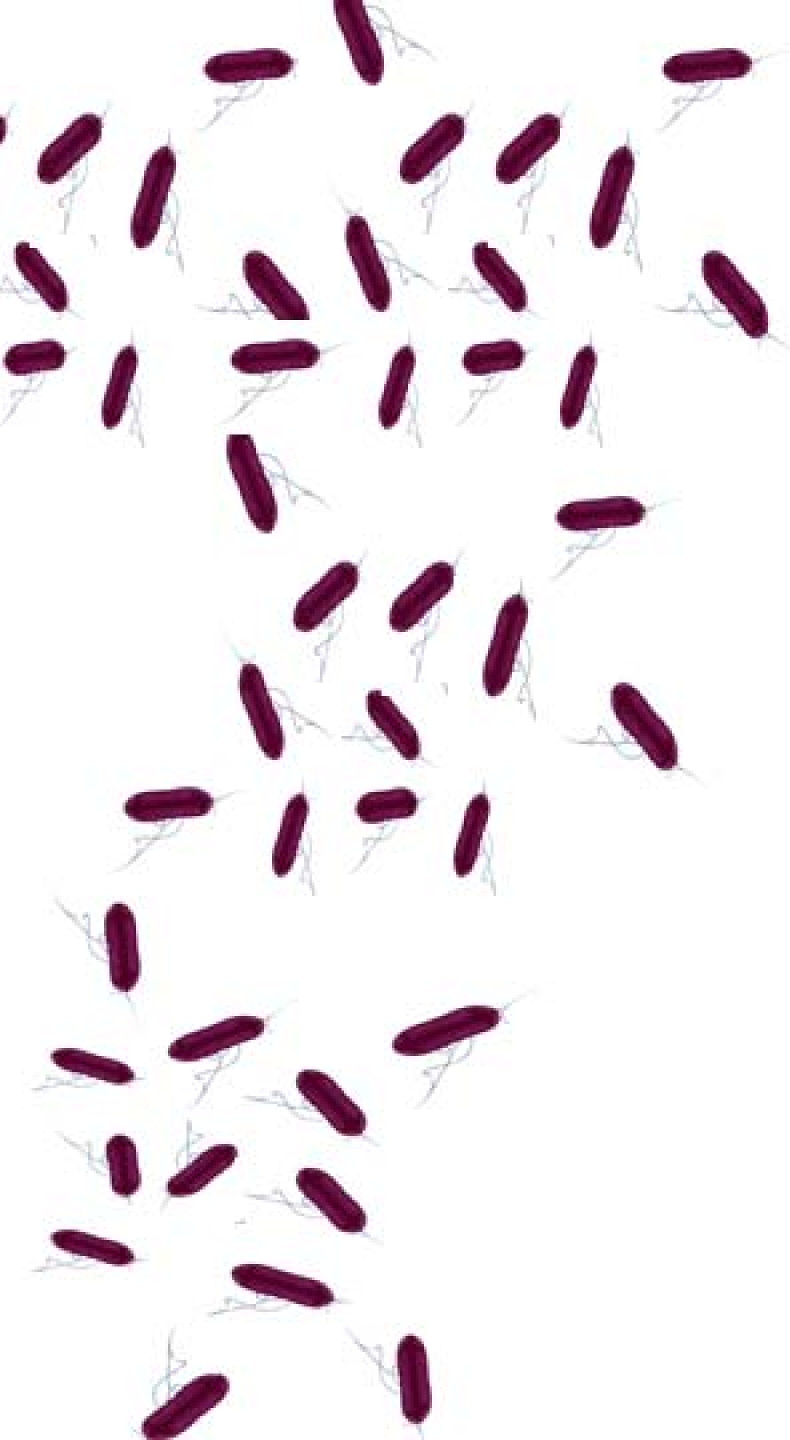
Biological optimization by directed evolution:



Easy mapping from code (DNA) to function (cells do the work and remember the answer).

One design algorithm for all levels of biological complexity!

Biology is unique among engineered systems for its capacity to evolve and for the ease with which evolution can be exploited for forward engineering.



Jim Liao (UCLA)

Pete Heinzelman
Kevin Boulware