

Engineering Microbial Cells as Chemical Factories I (b)

From Chassis organisms to metabolic pathway engineering strategies



Metabolic Engineering vs Pathway Design/Engineering

Metabolic engineering

Aims to optimise metabolic pathways in a producing organism by increasing precursors supply, tuning the activity of the intermediate reaction steps to eliminate bottlenecks, blocking competing pathways, etc.

e.g. increase the production of ethanol in yeast, citric acid in *Aspergillus niger*

Often a combination of both strategies are used!

Pathway design/engineering

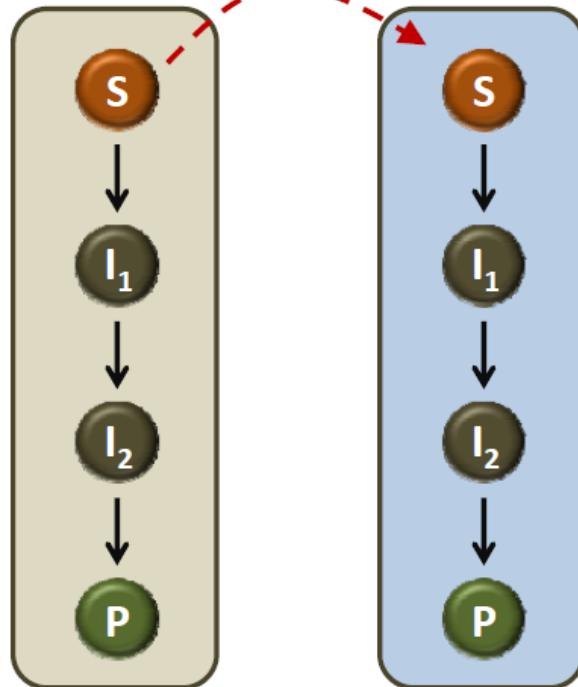
Uses gene-combinatorial methods to generate new pathways not in the organisms (chassis cells), comprising genes from different sources to find an optimal pathway configuration. Followed by debugging and optimisation steps (often use metabolic engineering principles).

e.g. *E. coli* that makes human insulin, yeast that makes antimalarial drug artemisinin



Strategies for Biosynthetic Pathway Design

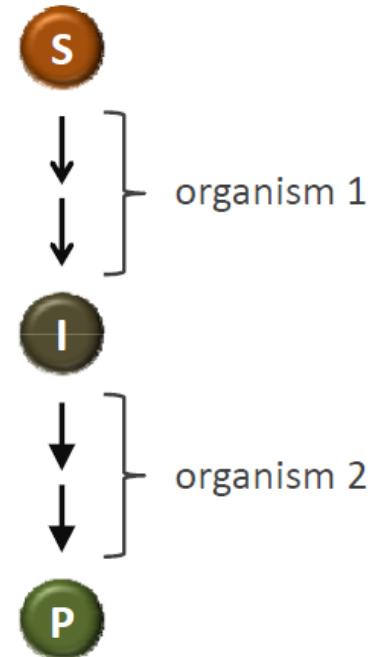
Transferring natural pathways to another host



- decoupling from native regulation
- expression in host that is easier to handle

e.g. genome mining of bioactive natural products

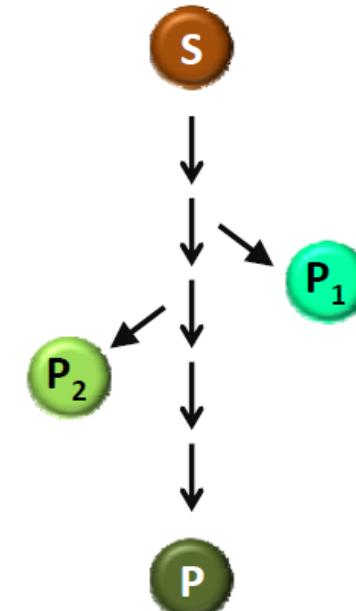
Combinations of existing pathways



pathway S to I is from one organism, I to P from another

e.g. artemisinin, opioids

Engineering of existing pathways



new products (P_i) are created from modifications or extensions of the pathway

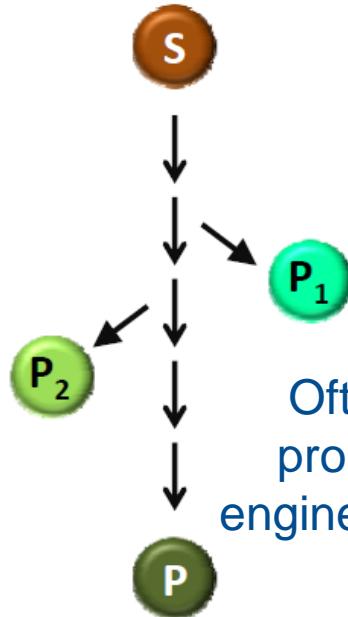
e.g. combinatorial biosynthesis of resorcylic acid lactone anticancers





Strategies for Biosynthetic Pathway Design

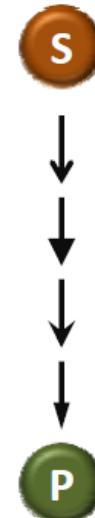
Engineering of existing pathways



Often involved promiscuous or engineered/enzymes

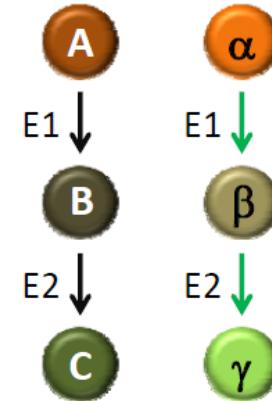
new products (P_i) are created from modifications or extensions of the pathway

de novo pathway design



each step is proposed independently

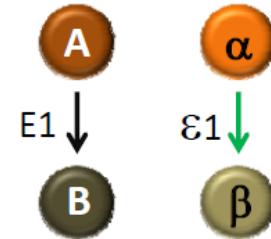
Exploiting enzymatic promiscuity to create new pathways and products



products from promiscuous enzymes may be from known or designed pathways

e.g. (S)-3-methyl-pentanol, 3-hydroxyvaleric acid

Creating new products and/or pathways using engineered enzymes

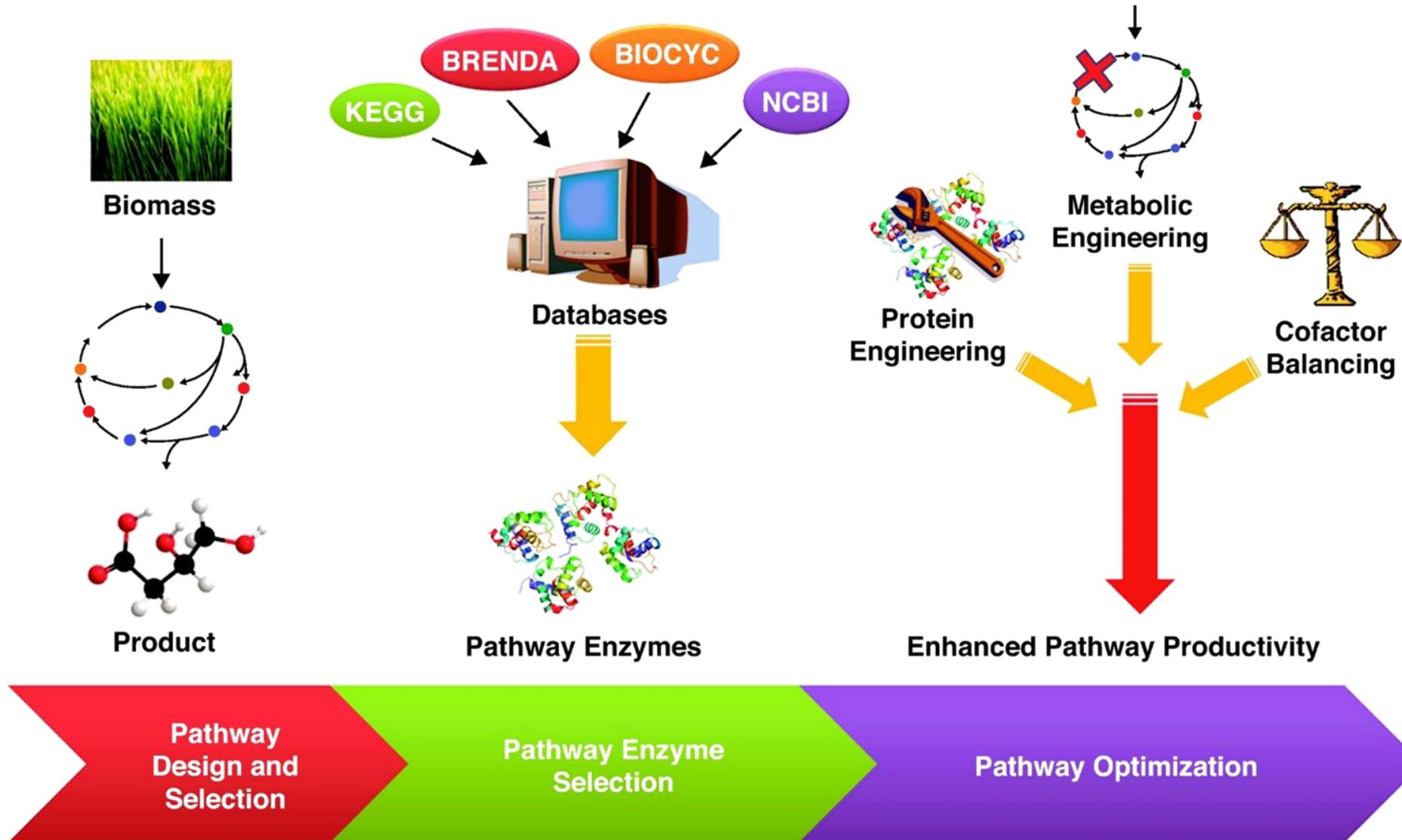


engineered enzymes are used to create an unnatural product or as more productive alternatives to existing enzymes

e.g. simvastatin, L-homoalanine

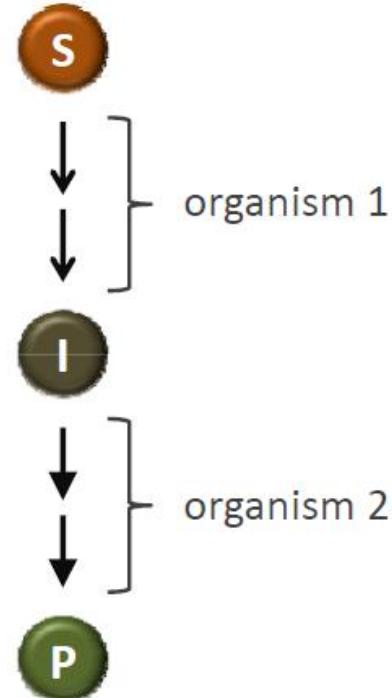
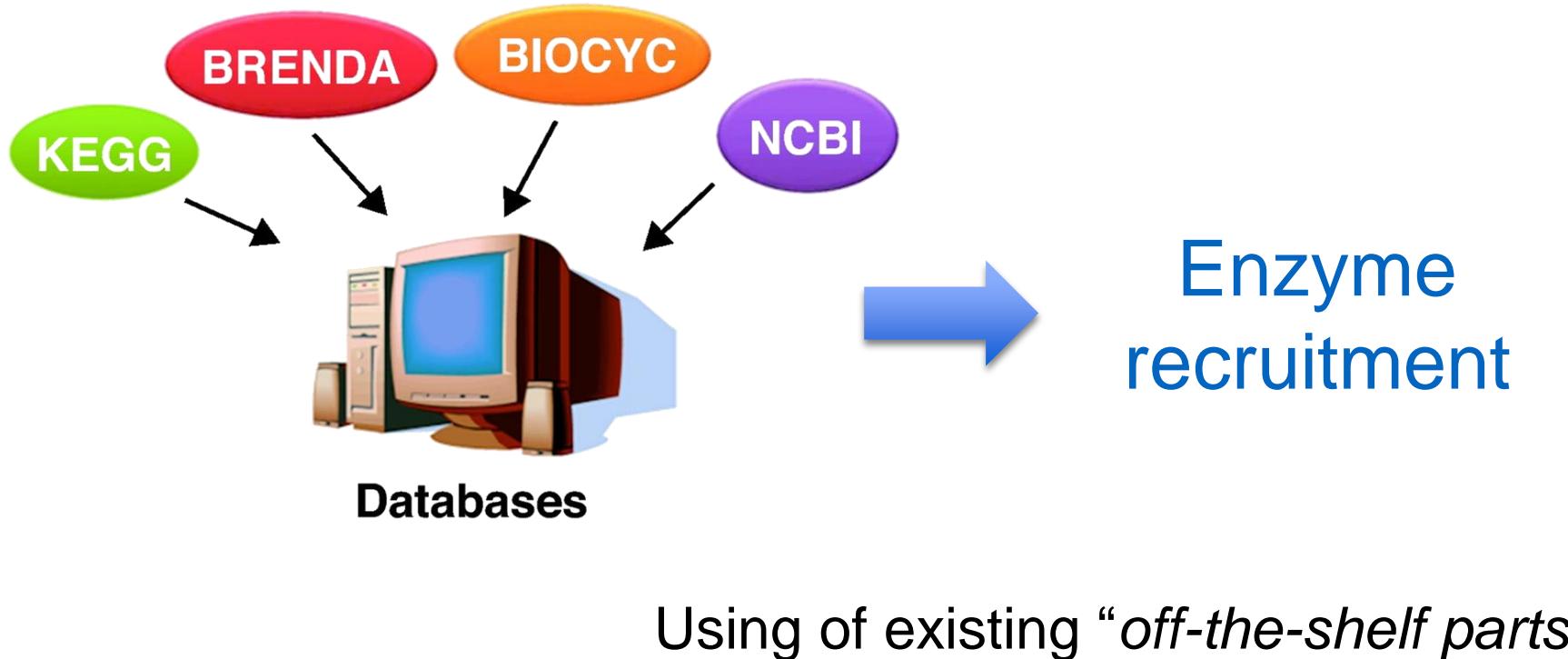


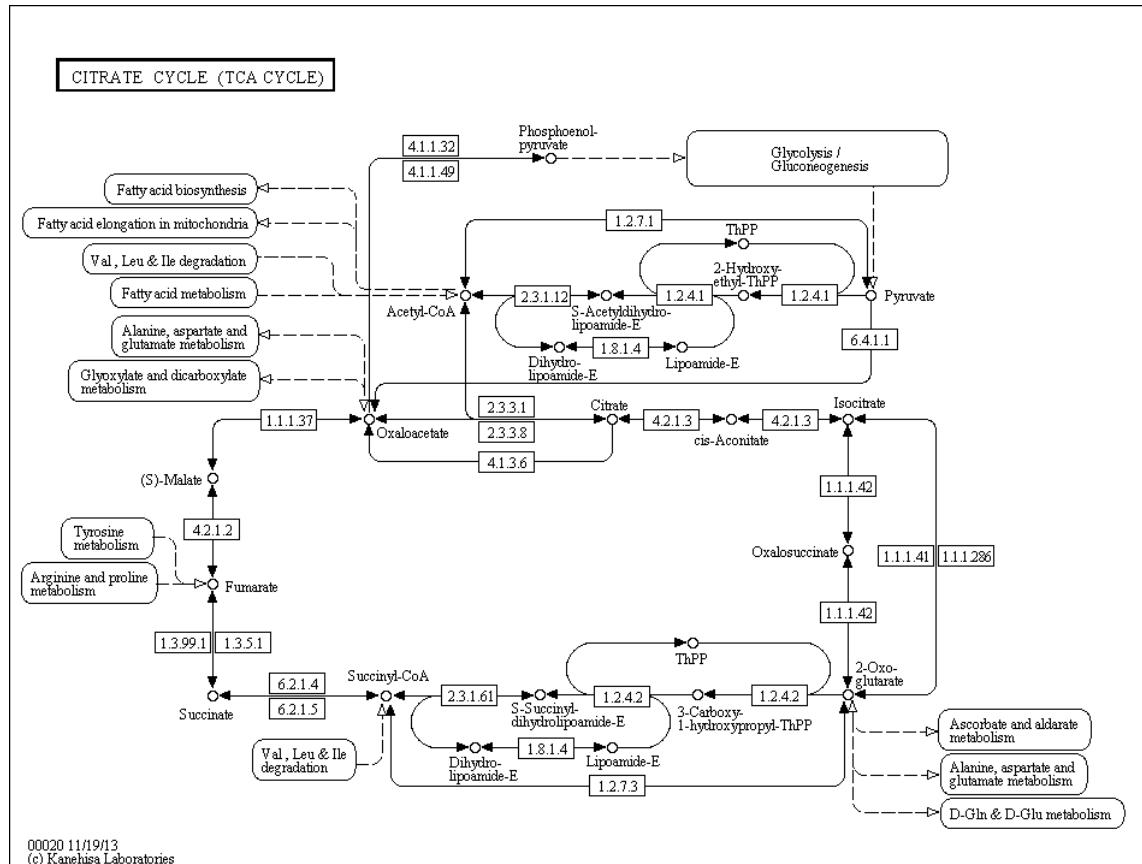
Design and Engineering of Pathway



Where to Get The parts?

Exploitation of Known Parts





TCA cycle from
KEGG (Kyoto
Encyclopedia of
Gene and Genomes)
<https://www.genome.jp/kegg/kegg2.html>

Emphasises
enzyme/EC numbers
and principal
products

KEGG and BioCyc (biocyc.org) are two places for finding information of metabolic pathways. Both are curated.

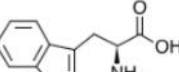


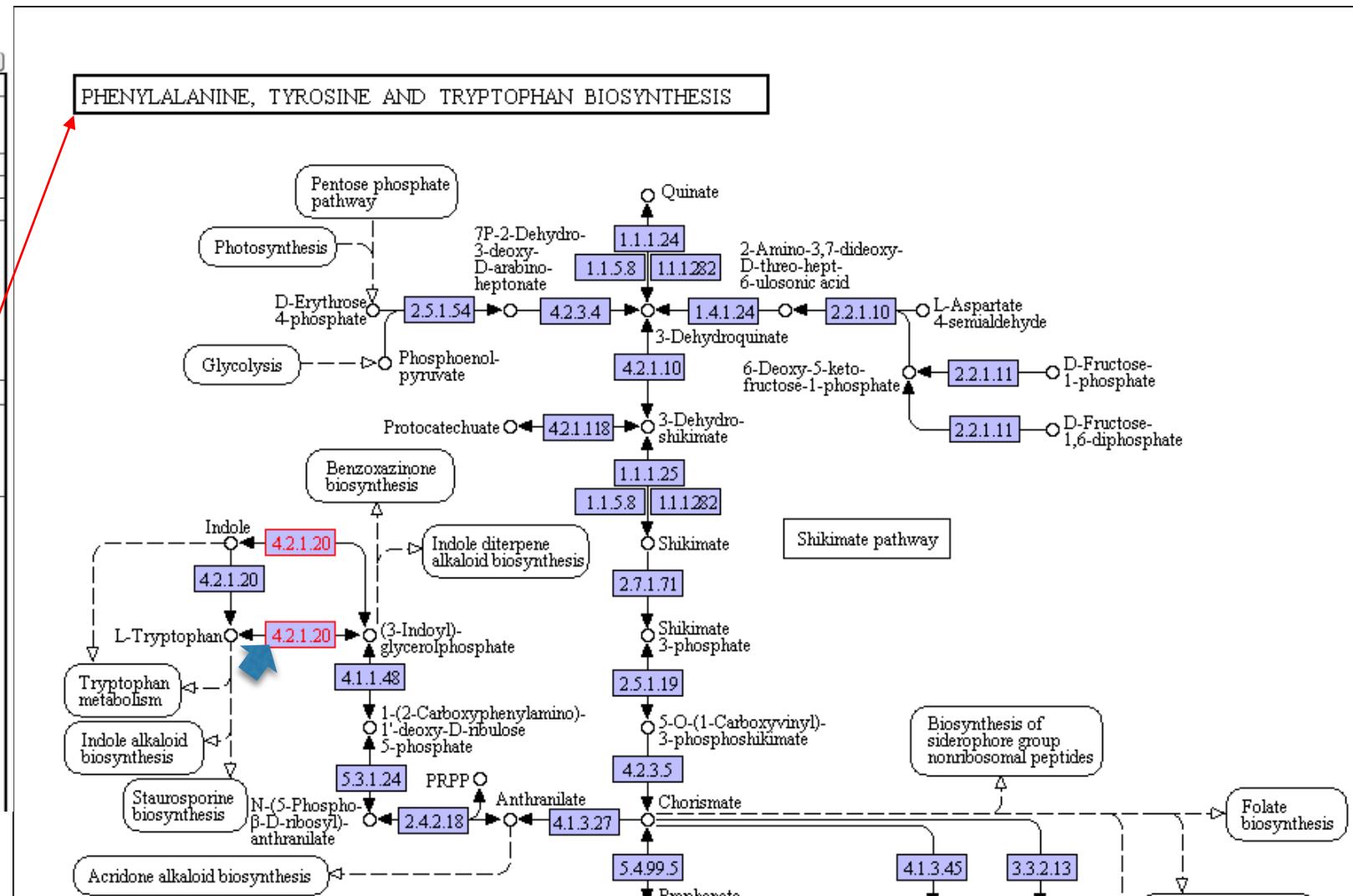
Metabolic Pathways – KEGG

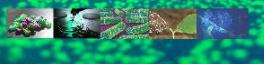
Kegg

COMPOUND: C00078

Help

Entry	C00078	Compound
Name	L-Tryptophan; Tryptophan; (S)-alpha-Amino-beta-(3-indolyl)-propionic acid	
Formula	C11H12N2O2	
Exact mass	204.0899	
Mol weight	204.2252	
Structure	 C00078	
	Mol file KCF file DB search	
Remark	Same as: D00020	
Reaction	R00673 R00674 R00675 R00676 R00677 R00678 R00679 R00681 R00682 R00683 R00684 R00685 R01376 R01657 R01814 R02722 R03664 R05317 R07213 R08160 R08547 R09570 R09583 R09638 R10180 R10661 R11119 R11328 R12031 R12032 R12055 R12151 R12413 R12417 R12436 R12542	
Pathway	map00260 Glycine, serine and threonine metabolism map00380 Tryptophan metabolism map00400 Phenylalanine, tyrosine and tryptophan biosynthesis map00404 Staurosporine biosynthesis map00901 Indole alkaloid biosynthesis map00966 Glucosinolate biosynthesis map00970 Aminoacyl-tRNA biosynthesis map00998 Biosynthesis of various secondary metabolites - part 2 map01060 Biosynthesis of plant secondary metabolites map01061 Biosynthesis of phenylpropanoids map01063 Biosynthesis of alkaloids derived from shikimate pathway map01070 Biosynthesis of plant hormones map01100 Metabolic pathways map01110 Biosynthesis of secondary metabolites map01210 2-Oxocarboxylic acid metabolism map01230 Biosynthesis of amino acids	





Metabolic Pathways – KEGG



ORTHOLOGY: K01695

Help

Entry	K01695	KO
Name	trpA	
Definition	tryptophan synthase alpha chain [EC:4.2.1.20]	
Pathway	ko00260 Glycine, serine and threonine metabolism ko00400 Phenylalanine, tyrosine and tryptophan biosynthesis ko01100 Metabolic pathways ko01110 Biosynthesis of secondary metabolites ko01130 Biosynthesis of antibiotics ko01230 Biosynthesis of amino acids	
Module	M00023 Tryptophan biosynthesis, chorismate => tryptophan	
Brite	KEGG Orthology (KO) [BR: ko00001] 09100 Metabolism 09105 Amino acid metabolism 00260 Glycine, serine and threonine metabolism K01695 trpA; tryptophan synthase alpha chain 00400 Phenylalanine, tyrosine and tryptophan biosynthesis K01695 trpA; tryptophan synthase alpha chain Enzymes [BR: ko01000] 4. Lyases 4.2 Carbon-oxygen lyases 4.2.1 Hydro-lyases 4.2.1.20 tryptophan synthase K01695 trpA; tryptophan synthase alpha chain BRITE hierarchy	
Other DBs	RN: R00674 R02340 R02722 COG: COG0159 GO: 0004834	
Genes	NVE: NEMVE_v1g152188 ATH: AT3G54640(TSA1) AT4G02610 ALY: ARALYDRAFT_485868 ARALYDRAFT_911953 CRB: 17884150 17884716	





GenPept

At3g54640 [Arabidopsis thaliana]

GenBank: ABF85777.1

[Identical Proteins](#) [FASTA](#) [Graphics](#)

Go to:

LOCUS ABF85777 312 aa linear PLN 07-JUN-2006

DEFINITION At3g54640 [Arabidopsis thaliana].

ACCESSION ABF85777

VERSION ABF85777.1

DBSOURCE accession BT025875.1

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM [Arabidopsis thaliana](#)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Gunneridae; Pentapetalae; rosids; malvids; Brassicales; Brassicaceae; Camelinae; Arabidopsis.

REFERENCE 1 (residues 1 to 312)

AUTHORS Quinto,C., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

REFERENCE 2 (residues 1 to 312)

AUTHORS Quinto,C., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.

TITLE Direct Submission

JOURNAL Submitted (07-JUN-2006) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10810 N. Torrey Pines Road, La Jolla, CA 92087, USA

COMMENT Method: conceptual translation.

FEATURES Location/Qualifiers

source 1..312
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U82429"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

Protein

1..312
/product="At3g54640"

Region

63..311
/region_name="PLN02591"
/note="tryptophan synthase"
/db_xref="CDD:178281"

Site

order(96,107,111,221,229..230,258..259,280..281)

Site

/site_type="other"
/note="substrate binding site [chemical binding]"
/db_xref="CDD:240075"
order(96,107,148,229)

Site

/site_type="active"
/db_xref="CDD:240075"
order(96,107,148)
/note="catalytic residues [active]"
/db_xref="CDD:240075"

Site

order(101..104,106,109,112..113,150,153,175..176,178,181,
201,203..205,208)
/site_type="other"
/note="heterodimer interface [polypeptide binding]"
/db_xref="CDD:240075"

CDS

1..312
/coded_by="BT025875.1:1..939"
/note="unknown protein"

ORIGIN

1 maiafksgvf flqspksqig frhssppdss lsfkrftpmal slstssptlg ladftqlkk
61 qgkvafipyi tagddplstt aeaikvlidac gsdiielgvp ysdpladgvp iqaaatrsle

Send to:

Change region shown

Customize view

Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

Articles about the TSA1 gene

Mutations in *Arabidopsis thaliana* genes involved in the tryptophan biosynthesis path [Plant J. 1996]

Arabidopsis indole synthase, a homolog of tryptophan synthase alc [J Integr Plant Biol. 2006]

Arabidopsis thaliana tryptophan synthase alpha: gene cloning, expression. [Mol Gen Genet. 1995]

See all...

Pathways for the TSA1 gene

Glycine, serine and threonine metabolism

Biosynthesis of amino acids

Tryptophan biosynthesis, chorismate => tryptophan

See all...

Reference sequence information

RefSeq protein

See the reference protein sequence for tryptophan synthase alpha chain (NP_567004.1).

More about the gene TSA1

Catalyzes the conversion of indole-3-glycerol phosphate to indole, the penultimate reaction in the biosynthesis of tryptophan.
Functions as a...

Also Known As: AT3G54640, TRP3, TRYPTO...

Related information

Nucleotide

Taxonomy

BioSystems

CDD Search Results

Conserved Domains (Concise)

THE UNIVERSITY OF
WESTERN
AUSTRALIA

KEGG links to NCBI GenBank database





Metabolic Pathways - BioCyc

EcoCyc
A member of the BioCyc database collection

Enter a gene, protein, metabolite or pathway
Searching *Escherichia coli* K-12 substr. MG1655

gene **trpA** enzyme **tryptophan synthase subunit α**
Escherichia coli K-12 substr. MG1655

Synonyms: try, trp, α subunit, TSase α, A protein

Accession IDs: EG11024 (EcoCyc)
b1260
ECK1254
P0A877 (UniProt)

Length: 807 bp / 268 aa
Map Position: [1,316,416 < 1,317,222] (28.36 centisomes, 102²) View in Genome Browser

Location: cytosol

Reactions: (1S,R)-1-C-(indol-3-yl)glycerol 3-phosphate + L-serine → L-tryptophan + D-glyceraldehyde 3-phosphate + H₂O (catalyzed by complex)
(1S,R)-1-C-(indol-3-yl)glycerol 3-phosphate ↔ indole + D-glyceraldehyde 3-phosphate

Pathway: L-tryptophan biosynthesis

Evidence: Assay of protein purified to homogeneity from its native host [Yutani87, Crawford80]

[Summary](#) [GO Terms \(8\)](#) [Essentiality](#) [Reactions \(2\)](#) [Protein Features](#) [Operons](#) [References](#) [Show All](#)

Regulation Summary Diagram



Summary

The TrpA polypeptide (TSase α) functions as the α subunit of the tetrameric (α₂β₂) tryptophan synthase complex [Miles77]. As a purified protein, the α subunit is a monomer. TSase α contains the binding site for indole-3-glycerol-phosphate (InGP) and can carry out the cleavage reaction of InGP to indole and glyceraldehyde-3-phosphate, also termed the α reaction. Within the physiological complex with the β subunit, the reaction rate is increased by 1-2 orders of magnitude (in [Kirschner91]).

TrpA has been shown to lack tryptophan residues [Henning62]. Numerous TrpA mutant studies have examined structure-function relationships in this protein. Mutations that affect catalytic activity [Hiraga96, Sarker95, Lim91, Milton86, Yutani87, Yee96, Yanofsky93, Lim91a], subunit interactions [Swift92, Lim91a], conformational stability [Hiraga96, Lim92] and folding [Kim01g, Lim91a] have been identified.

The crystal structure of the wild-type TrpA protein has been reported at 2.8 Å resolution [Jeong04c, Jeong04d], 2.5 Å resolution [Jeong05] and 2.3 Å resolution [Nishio05]. The crystal structure of a double mutant TrpA protein has been reported at 1.6 Å resolution [Jeong04d, Jeong05].

The TrpA protein has been structurally classified as a (βα)₈ TIM barrel protein, a member of the common TIM barrel superfamily. Nuclear magnetic resonance spectroscopic techniques have been used to investigate its equilibrium folding mechanism in order to obtain insights into the development of structure and stability [Vadrevu08]. Many previous studies of the folding mechanism of recombinant wild-type and mutant TrpA proteins using various biophysical techniques identified intermediates in the folding pathway, for example [Beasty86, Choi95, Gualfetti99, Jeong03, Wintrode05, Wu05c, Wu07g].

Additional Citations: [Yanofsky81, Yanofsky86, Imamoto85, Imamoto86]

Molecular Weight of Polypeptide	28,724 kD (from nucleotide sequence), 30.0 kD (experimental) [Gschwind79]
pI	5.54
Gene Product Copy Number	653 molecules/cell [T=37°C, pH=7.2, medium=Neidhardt EZ rich defined medium minus methionine, Li14] 2508 molecules/cell [T=37°C, pH=7.2, medium=MOPS medium with 2% glucose, Li14] 544 molecules/cell [T=37°C, pH=7.2, medium=Neidhardt EZ rich defined medium, Li14]
Component of	tryptophan synthase (summary available): [TrpA ₂ /TrpB ₂]

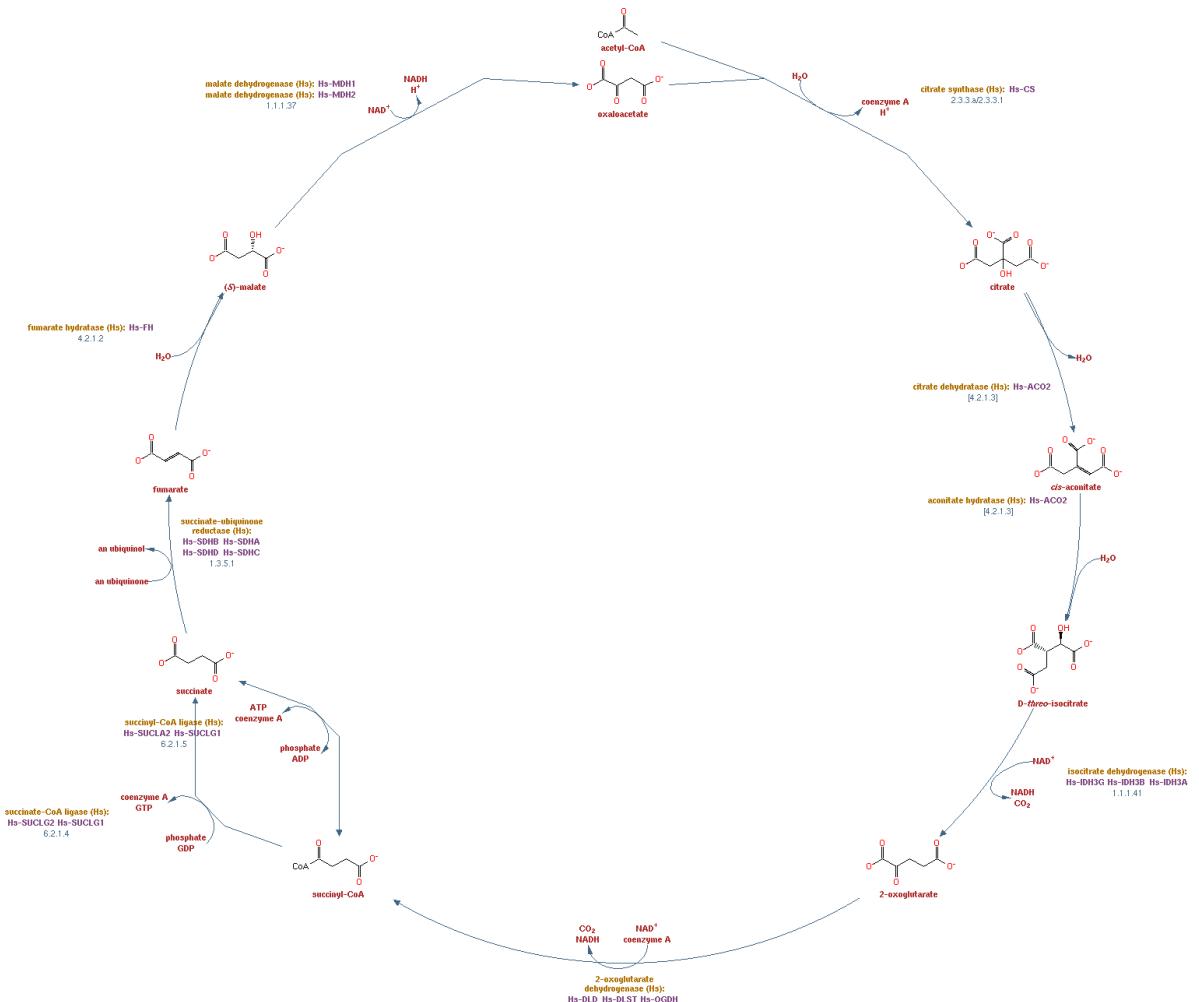
Unification Links

ASAP	ABE-0004232
CDS	74
DIP	DIP-35857N
DefProt	DP00252
EcoCyc	E9-1017
EcoPI	b1260
MetBase	P0A877
OU-Microarray	b1260
Proteo	trpA
PR	PRO_000024117
Pride	P0A877
Protein Model Portal	P0A877
RefSeq	NP_415776
RegulonDB	b1260
SMR	P0A877
String	51145,b1260
UniProt	P0A877

please checkout the databases on your web browser



Metabolic Pathways – BioCyc



Metazoan TCA cycle

BioCyc pathways are drawn on the fly (rather than being precomputed maps).

Kegg and BioCyc
maps created by
linking individual data
points
(reactions/metabolite
s)



[BRENDA home](#)
[BACK](#)
[History of your search](#)
[Improve BRENDA!](#)

Enzyme Nomenclature
EC number
Recommended Name
Reaction
Reaction Type
Pathway
Systematic Name
Synonyms
CAS Registry Number
Enzyme-Ligand Interactions
Substrate/Product
Natural Substrates
Cofactor
Metals and Ions
Inhibitors
Activating Compound
Functional Parameters
KM Value
Turnover Number
kcat/KM Value
Ki Value



BRENDA

The Comprehensive Enzyme Information System
EC 2.7.1.1 - hexokinase



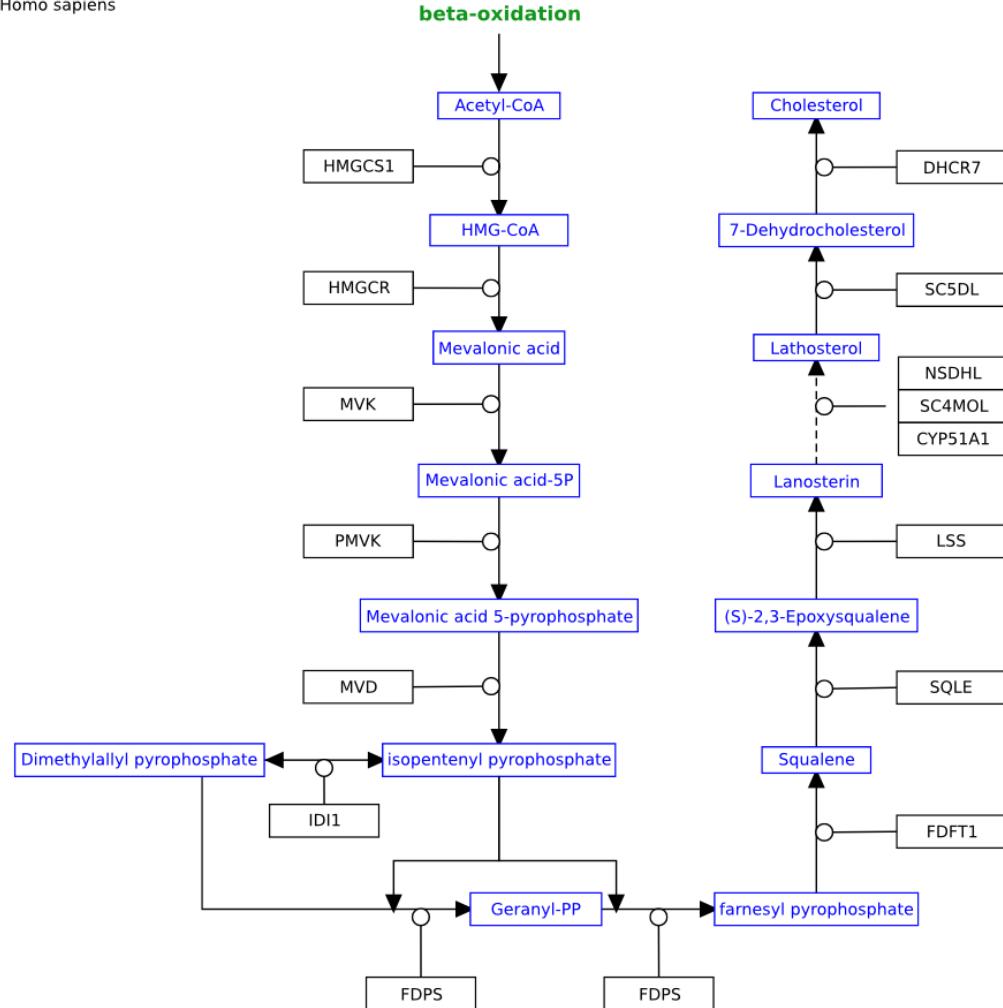
29	-	D-glucose	Homo sapiens	-	wild type enzyme	702320
34	-	D-glucose	Homo sapiens	-	mutant enzyme G72R, at 37°C, in 100 mM Tris and 125 mM KCl, pH 7.4, in the presence of 14 mM beta-mercaptoethanol or 5 mM dithiothreitol	702282
37	-	D-glucose	Homo sapiens	-	glucokinase with C-terminal 5 alanine addition	702320
38	-	D-glucose	Homo sapiens	-	wild-type enzyme	672188
38	-	D-glucose	Homo sapiens	-	glucokinase with C-terminal 10 alanine addition	702320
39	-	D-glucose	Homo sapiens	-	mutant enzyme Y214C	672188
42.1	-	D-glucose	Homo sapiens	-	mutant enzyme W257F, in 25 mM HEPES (pH 7.4), at 37°C	703624
46.8	-	D-glucose	Homo sapiens	-	mutant enzyme W99F, in 25 mM HEPES (pH 7.4), at 37°C	703624
54	-	D-glucose	Homo sapiens	-	mutant enzyme V62M, at 30°C, in 100 mM Tris and 125 mM KCl, pH 7.4, in the presence of 14 mM beta-mercaptoethanol or 5 mM dithiothreitol	702282
60.4	-	D-glucose	Homo sapiens	-	wild type enzyme, in 25 mM HEPES (pH 7.4), at 37°C	703624
61	-	D-alucose	Homo	-	mutant enzyme V62M, at 37°C, in 100 mM Tris and 125 mM KCl, pH 7.4, in the presence of 14 mM beta-mercaptoethanol or 5 mM dithiothreitol	702282

Brenda is an excellent database containing information about enzymes, their kinetic parameters KM, Kcat, Ki, etc



WikiPathways

Title: Cholesterol Biosynthesis
 Last modified: 2/22/2013
 Organism: Homo sapiens



- An open, collaborative platform built on MediaWiki dedicated to the curation of biological pathways.
- a new model for pathway databases that enhances and complements ongoing efforts, such as BioCyc, KEGG, Reactome etc..
- Include PPI and other biological networks



MIBiG (Secondary Metabolite Biosynthetic Pathway)



MIBiG Minimum Information about a Biosynthetic Gene cluster [Home](#) [Upload](#) [!](#) [Download](#)

Select Genomic Locus:
[Overview](#) 1

BGC0000070, griseofulvin biosynthetic gene cluster from *Penicillium aethiopicum*. Locus 1. Full MIBiG entry.

Gene cluster description
griseofulvin (BGC0000070). Gene Cluster 1. Biosynthetic class = Polyketide. GenBank GU574478, positions 9800-41750. Click on genes for more information. [Download cluster GenBank file](#)



Legend:
 biosynthetic genes
 transport-related genes
 regulatory genes
 other genes

Domain annotation
ADI24953.1



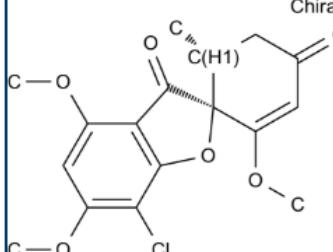
General MIBiG information on this cluster

Complete gene cluster sequence?	complete
Evidence for cluster-compound connection:	Knock-out studies, Enzymatic assays, Heterologous expression, Proven expression in natural host
MIxS-compliance:	Unknown
Contact for this cluster:	Yit-Heng Chooi (The Australian National University)

Literature references

1. Chooi YH et al. (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* 17(5):483-94. doi: 10.1016/j.chembiol.2010.03.015.
2. Cacho RA et al. (2013) Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. *ACS Chem Biol* 8(10):2322-30. doi: 10.1021/cb400541z. Epub 2013 Sep
3. Raab MS et al. (2012) GF-15, a novel inhibitor of centrosomal clustering, suppresses tumor cell growth in vitro and in vivo. *Cancer Res* 72(20):5374-85. doi: 10.1158/0008-5472.CAN-12-2026. Epub
4. Ronnest MH et al. (2012) Disparate SAR data of griseofulvin analogues for the dermatophytes *Trichophyton mentagrophytes*, *T. rubrum*, and MDA-MB-231 cancer cells. *J Med Chem* 55(2):652-60. doi: 10.1021/jm200835c. Epub 2012 Jan 17.
5. Panda D et al. (2005) Kinetic suppression of microtubule dynamic instability by griseofulvin: implications for its possible use in the treatment of cancer. *Proc Natl Acad Sci U S A* 102(28):9878-83. Epub 2005 Jun 28.
6. Roobol A et al. (1977) Griseofulvin-induced aggregation of microtubule protein. *Biochem J* 167(1):39-43.
7. Wohland J et al. (1977) Interaction of griseofulvin with microtubules, microtubule protein and tubulin. *J Mol Biol* 111(2):329-42

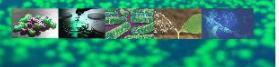
Chemical compounds


Compound: griseofulvin
PubChem ID: 441140
SMILES string: Copy to clipboard
Molecular formula: C17H17ClO6
Average molecular mass: 352.76628 Da
Molecular activity: Antifungal, Cytotoxic
Molecular target: microtubule protein

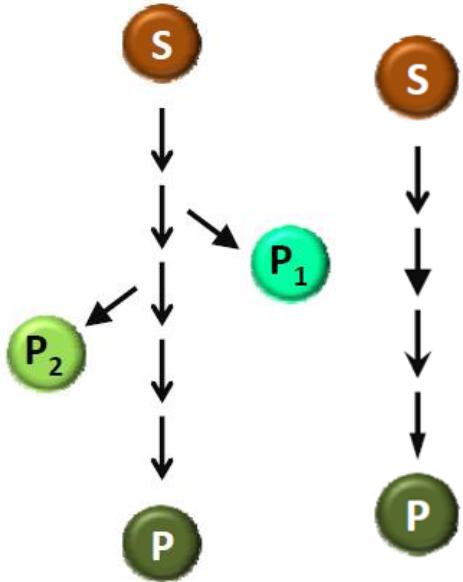
Class-specific details

Biosynthetic class(es):
Polyketide
Polyketide subclass:
Polyphenol (cyclic)
Polyketide synthase subclass:
Iterative type I
Starter unit:
Acetyl-CoA
Polyketide synthase / ketosynthase-encoding genes:
ADI24953
Iterative PKS subtype:
Non-reducing
Number of iterations:
6
Other
Thioesterase type:
None
Release / cyclization type:
Claisen condensation

Also provide links to NCBI GenBank and can link to BLAST search



Exploiting enzyme promiscuity for design and engineering of novel pathways

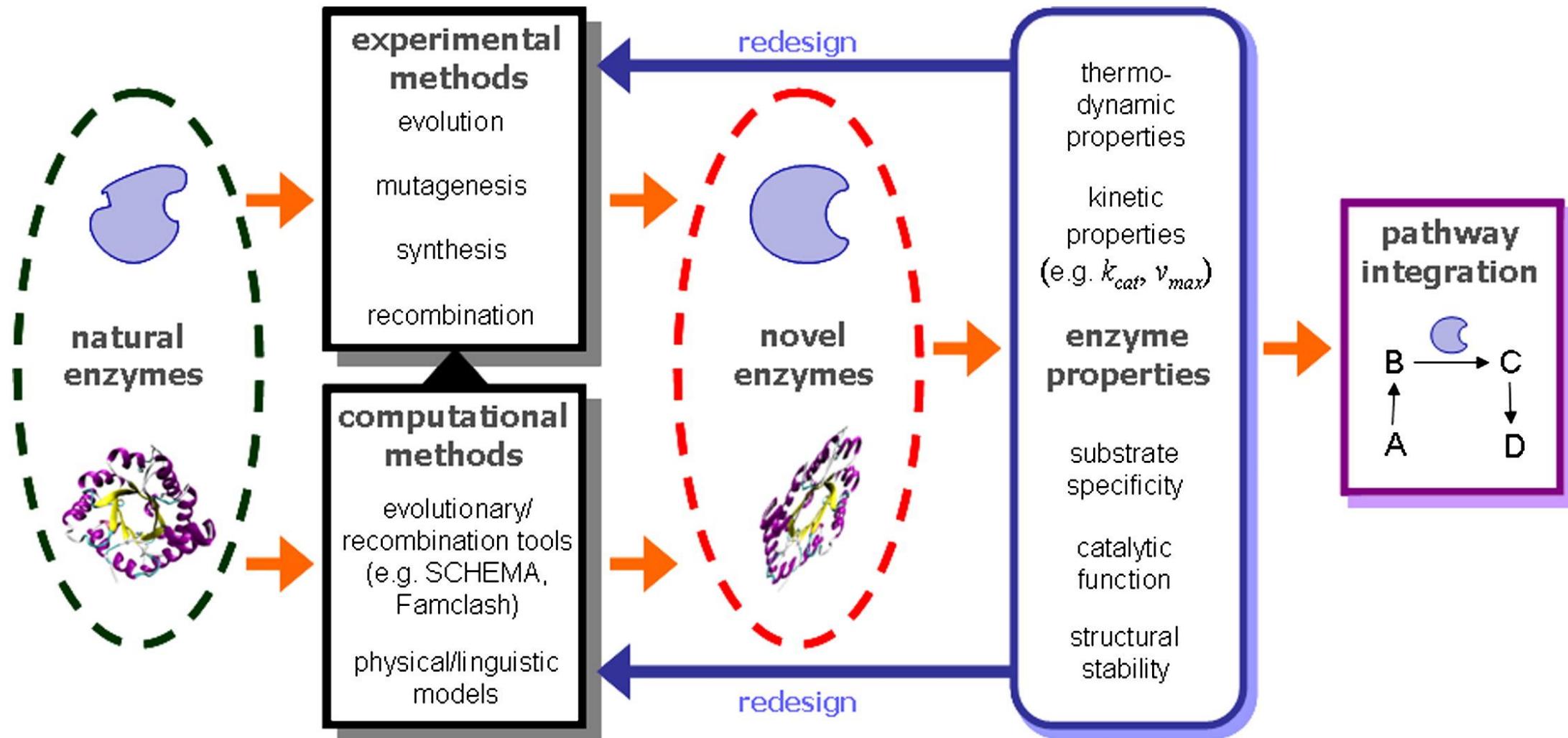


Enzyme promiscuity = ability to catalyze the same reactions with a variety of structurally and chemically similar substrates (**substrate promiscuity**) or multiple chemical reactions (**catalytic promiscuity**)

Enzyme promiscuity can serve as starting point for engineering novel biocatalysts with higher activity and specificity towards non-native substrates



Generation of novel enzymes and enhancement of activity through protein engineering





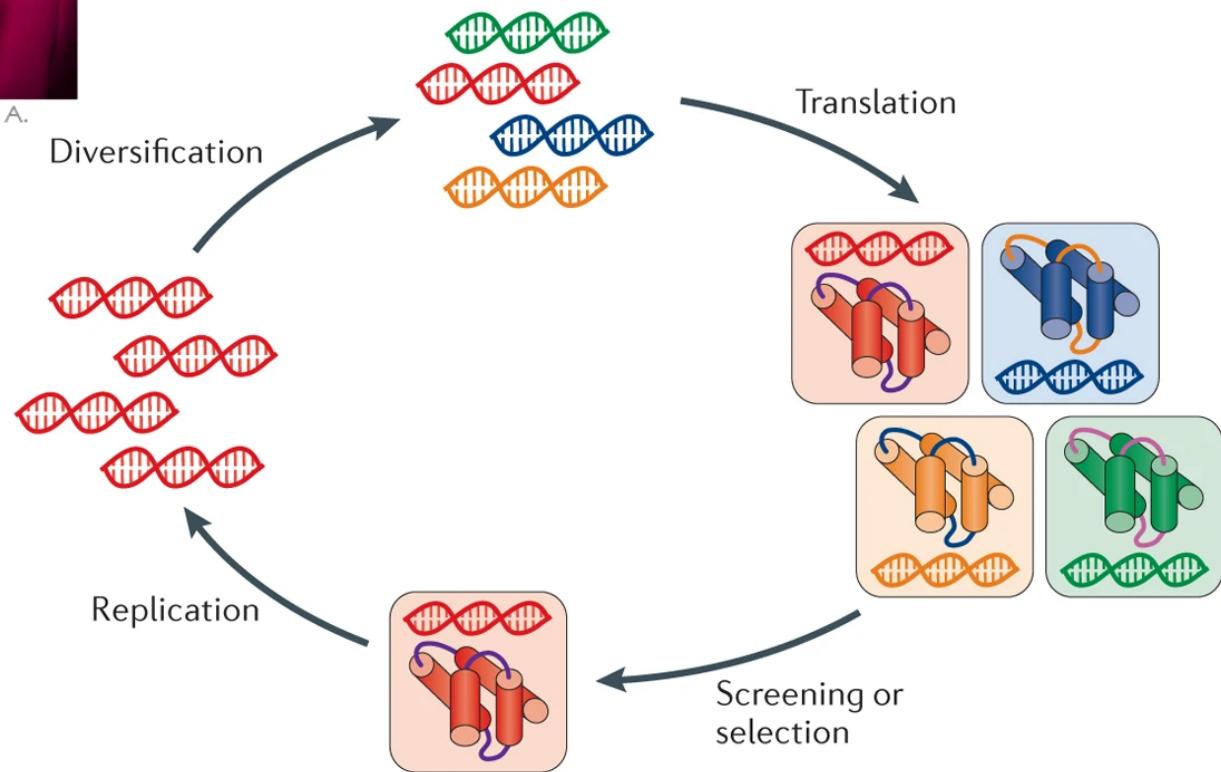
Frances H. Arnold
The Nobel Prize in Chemistry 2018

Born: 25 July 1956, Pittsburgh, PA, USA

Affiliation at the time of the award: California Institute of Technology (Caltech), Pasadena, CA, USA

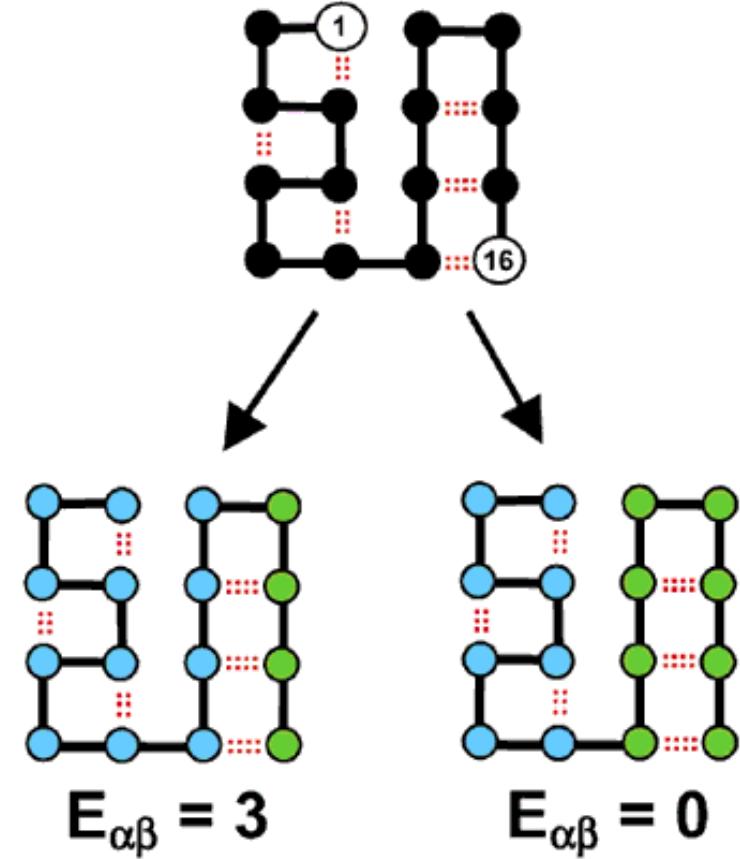
Prize motivation: "for the directed evolution of enzymes."

Prize share: 1/2

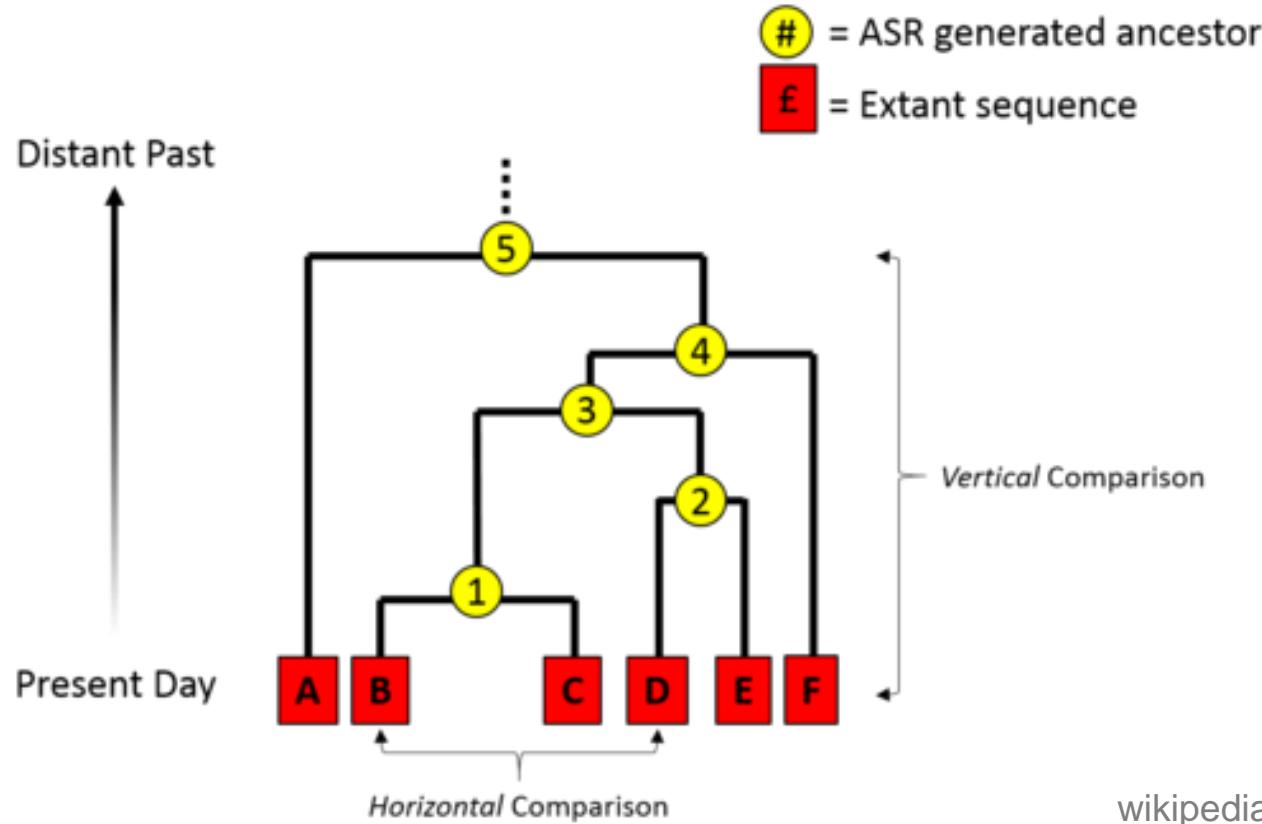


Enzyme Directed Evolution

Also developed the structure-guided protein recombination/shuffling method SCHEMA to accelerate directed evolution



Ancestral Protein Reconstruction



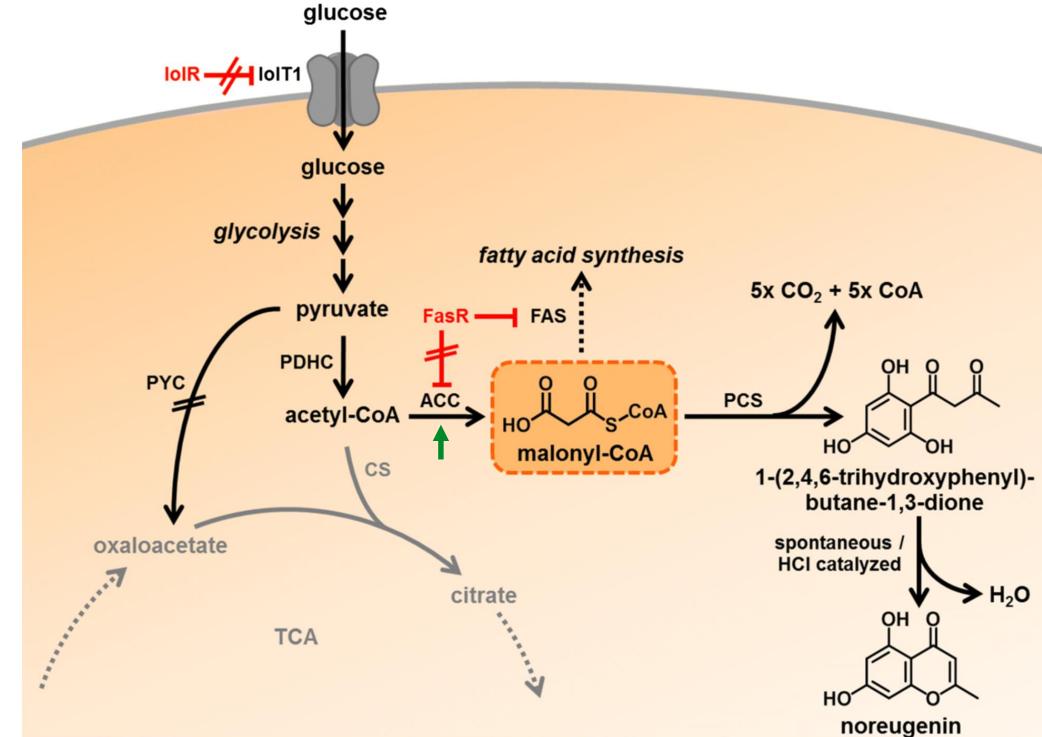
ASR poses as a valuable tool for protein engineers as protein derived from ASR has a tendency towards ancient thermostability and enzymatic promiscuity.



Enhancing pathway productivity through optimization of pathway and chassis metabolism

Optimizations to achieve

- selective diversion of the precursor metabolites and pathway intermediates from natural metabolism or competing side reactions without hampering growth and cell viability
- enhancing precursor metabolite pools
 - by knocking out host genes encoding enzymes catalyzing competing reactions
 - by controlled overproduction of pathway enzymes to enhance flux along the pathway and to reduce metabolic burden
 - continuous flux through pathway avoiding accumulation of intermediates
 - by fine-tuning of expression levels and activities of pathway components

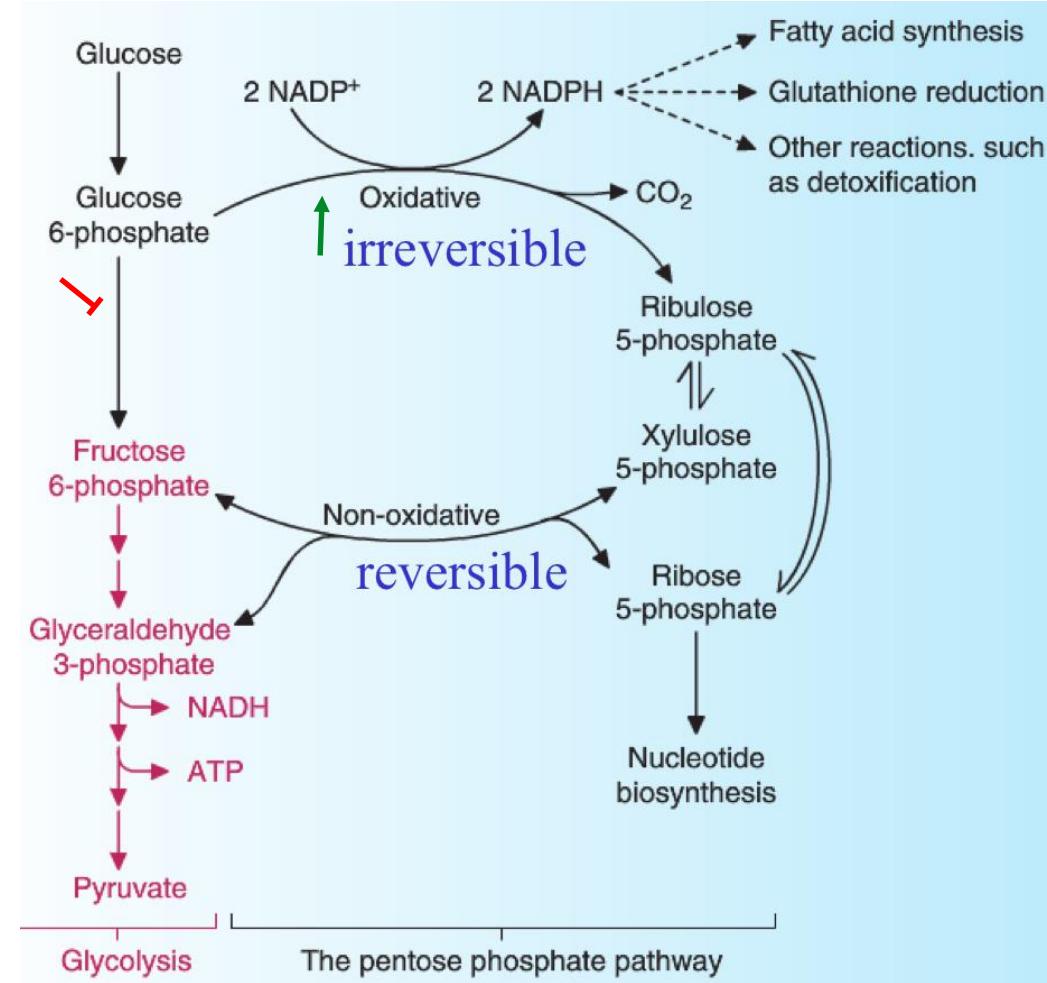


Milke et al. 2019 Microbial Cell Fact

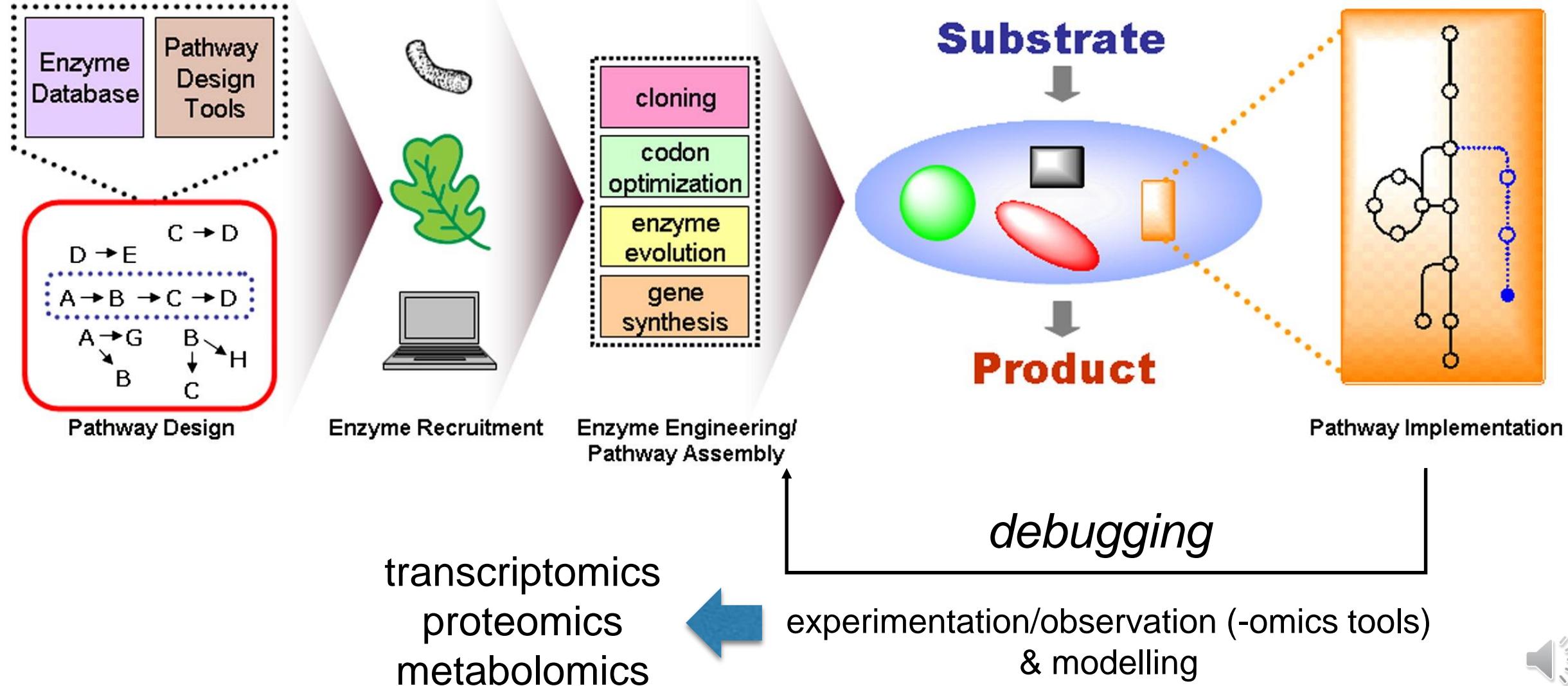


Enhancing pathway productivity through cofactor balancing

- Metabolic pathways often employ reactions catalyzed by enzymes using cofactors example: oxidation-reduction reactions using NAD(H) or NADP(H) as cofactors
- Cofactors need to be recycled back to their active form via natural cell metabolism or a subsequent pathway step
- Pathway enzymes often have to compete with natural cellular enzymes for essential cofactors
- Imbalance between pathway demands and natural cofactor recycling can pose serious limitations on pathway productivity



Scheme for pathway creation





Optimal synthetic pathway integrated into a cell

- no accumulation of intermediates
- enzymes operating at or near maximum capacity
- no formation of side products
- minimal interference with normal cellular functions
- achieve maximum possible yield

