Biosynthetic Gene Clusters in (New) Natural Product Synthesis

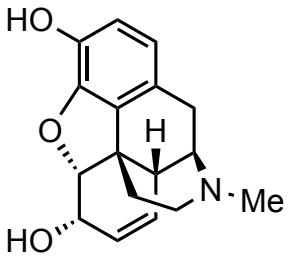
Autumn Flynn

Sigman Lab
March 31st, 2021

given time, nature creates solutions for all its past, current, and future problems

it's up to us to figure out what exactly it has done and benefit from it

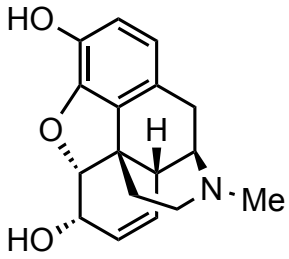
A familiar natural product



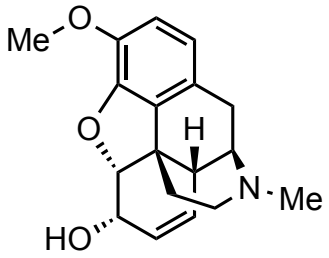
morphine

...is accompanied by many other natural products

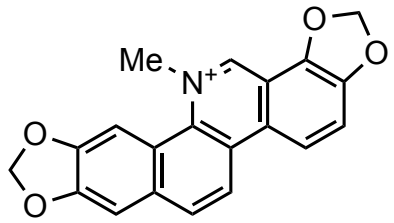
(some) Papaver somniferum natural products



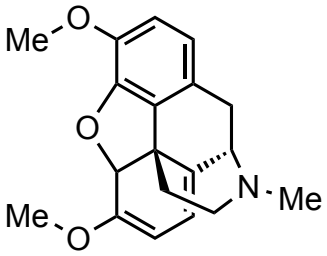
morphine



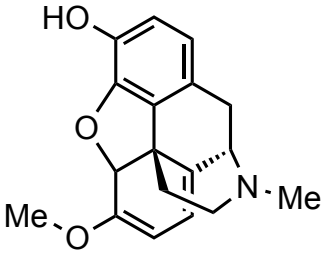
codeine



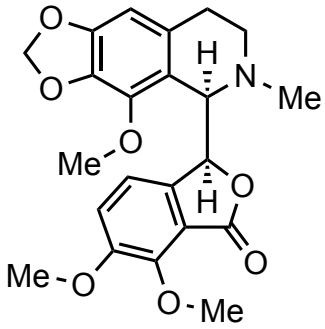
sanguinarine



thebaine



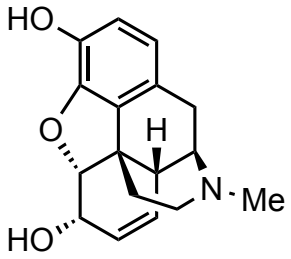
oripavine



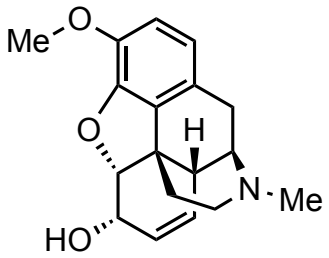
noscapine

different natural products are expressed under different environmental conditions

(some) Papaver somniferum natural products



morphine



codeine

sanguinarine

thebaine

oripavine

noscapine

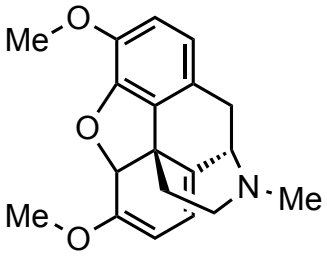
different natural products are expressed under different environmental conditions

(some) Papaver somniferum natural products

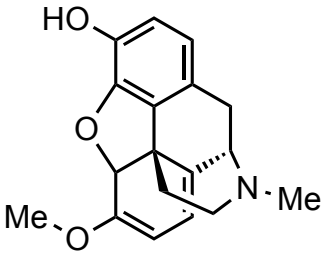
morphine

codeine

sanguinarine



thebaine



oripavine

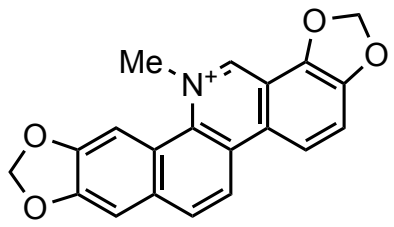
noscapine

different natural products are expressed under different environmental conditions

(some) Papaver somniferum natural products

morphine

codeine



sanguinarine

thebaine

oripavine

noscapine

different natural products are expressed under different environmental conditions

(some) Papaver somniferum natural products

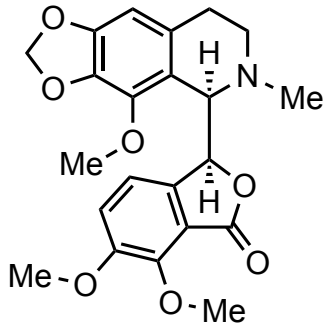
morphine

codeine

sanguinarine

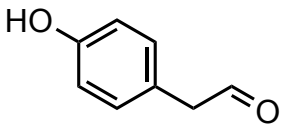
thebaine

oripavine



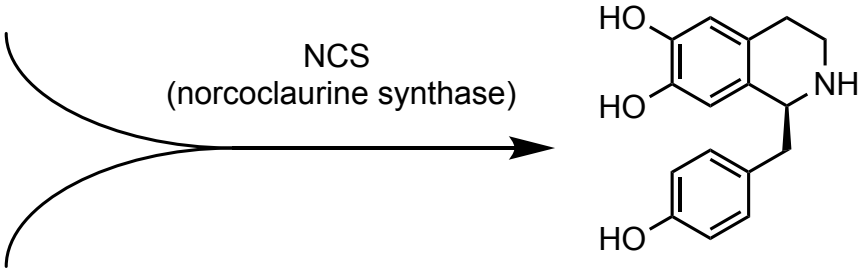
noscapine

First Question: How does the opium poppy synthesize these compounds?

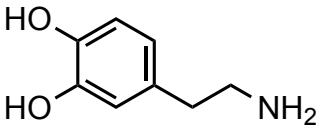


4-hydroxyphenylacetaldehyde

(from tyrosine —> tyramine)



(s)-norcoclaurine

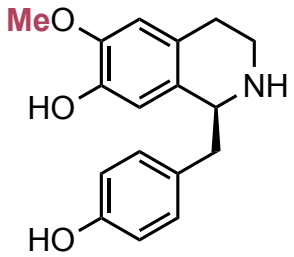


dopamine

(from tyrosine —> L-DOPA)

First Question: How does the opium poppy synthesize these compounds?

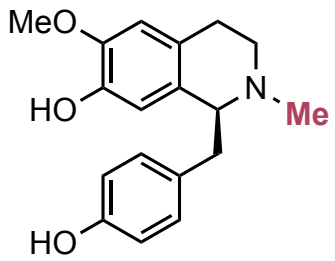
6OMT (6-O-methyltransferase)



(s)-coclaurine

First Question: How does the opium poppy synthesize these compounds?

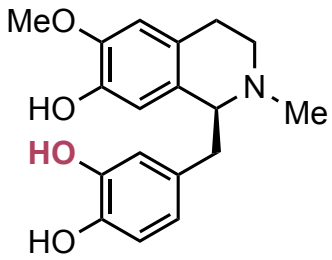
NCMT (N-coclaurine methyltransferase)



(s)-N-methylcoclaurine

First Question: How does the opium poppy synthesize these compounds?

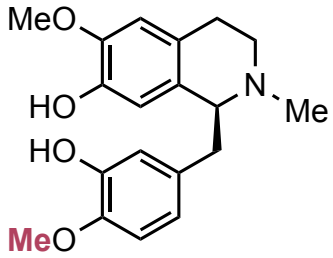
NMCH (N-methylcoclaurine 3' Hydroxylase)



(s)-3'Hydroxy N-methycoclaurine

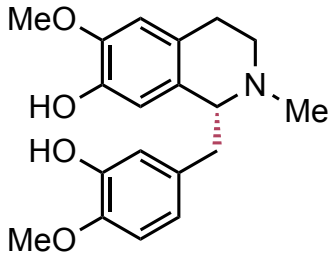
First Question: How does the opium poppy synthesize these compounds?

4'OMT (4' O-Methyltransferase))



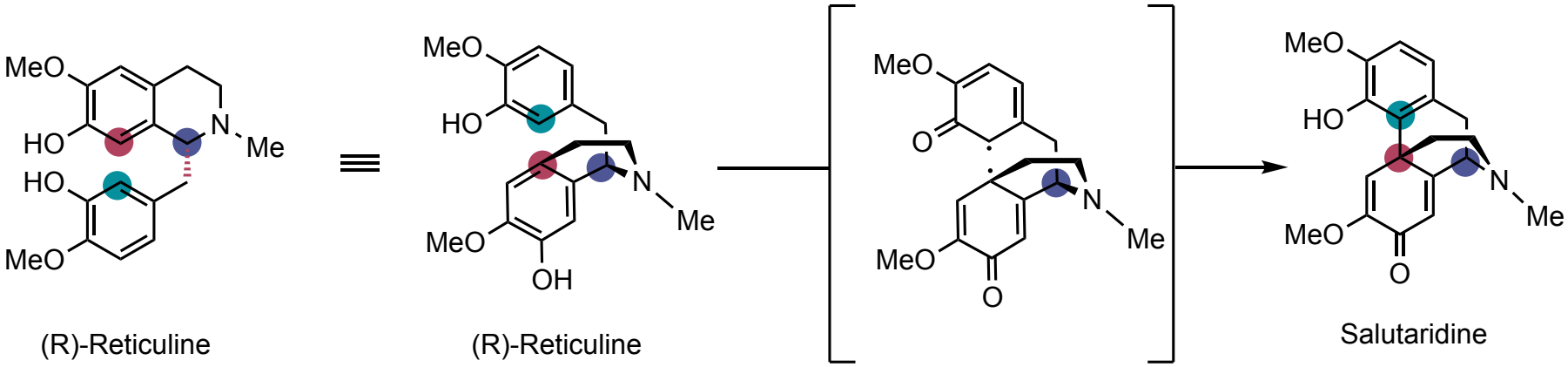
(s)-Reticuline

STORR (S-to-R Reticuline)



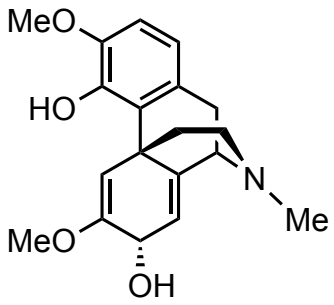
(R)-Reticuline

SalSyn (salutaridine synthase)



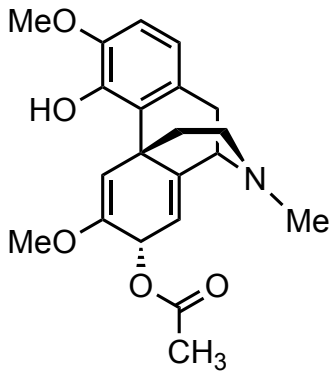
(R)-Reticuline (R)-Reticuline Salutaridine

SalR (salutaridine reductase)

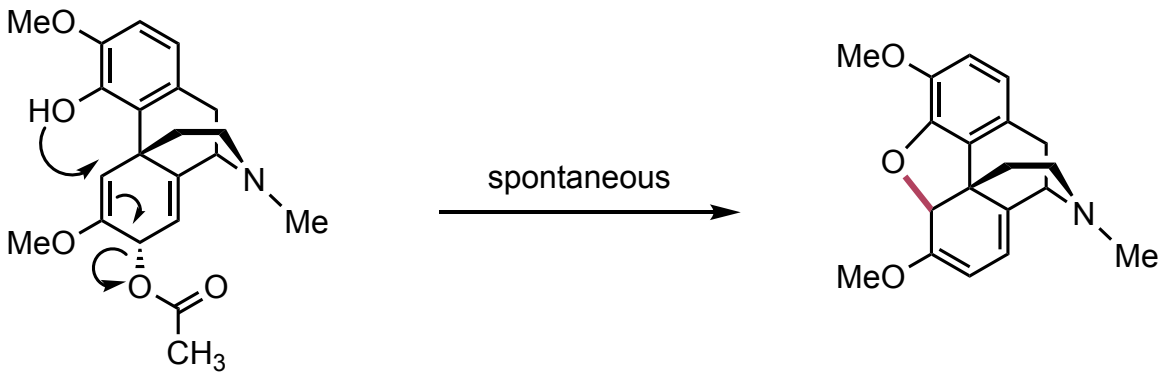


Salutaridinol

SalAT (salutaridinol acetyltransferase)

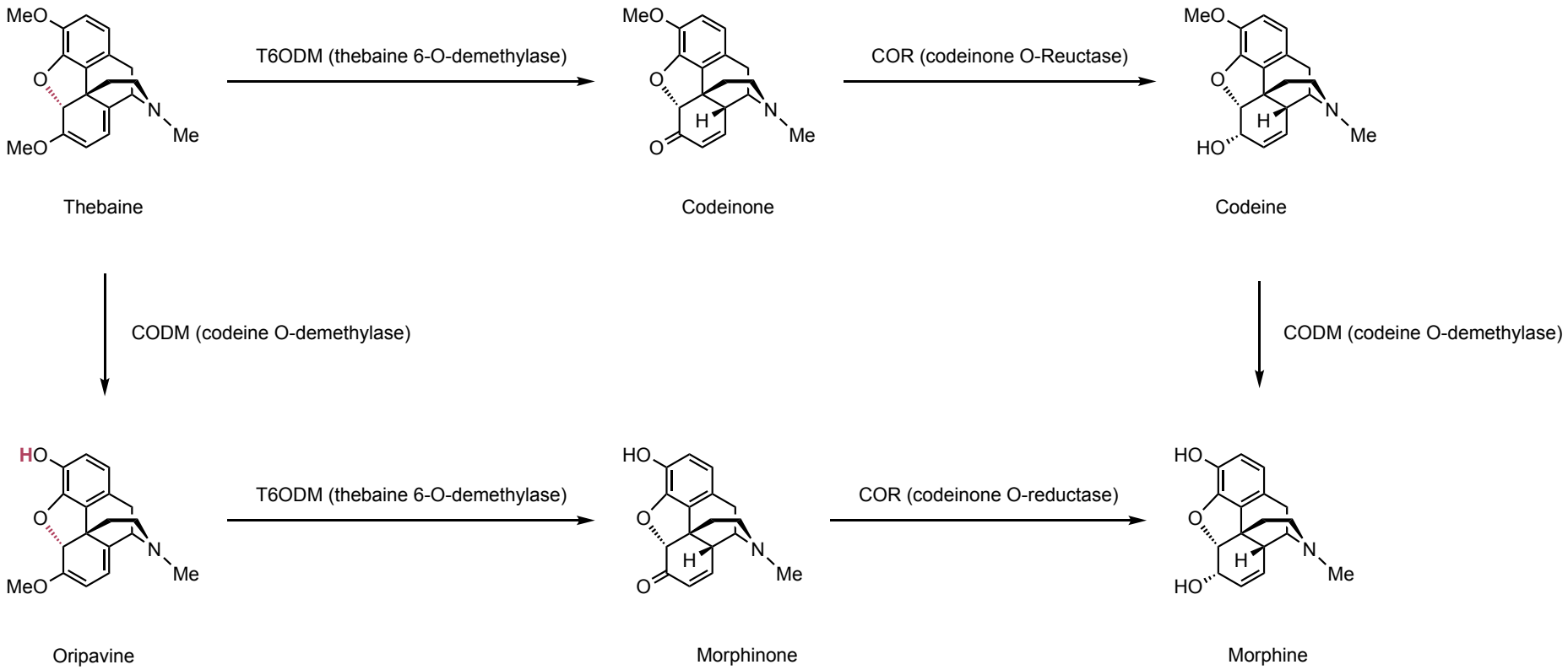


Salutaridinol 7-O-acetate



Salutaridinol 7-O-acetate

Thebaine


MeO
O
MeO
N
Me
Thebaine
T6ODM (thebaine 6-O-demethylase)
MeO
O
H
O
N
Me
Codeinone
COR (codeinone O-Reuctase)
MeO
O
H
HO
N
Me
Codeine
CODM (codeine O-demethylase)
HO
O
MeO
N
Me
Oripavine
T6ODM (thebaine 6-O-demethylase)
HO
O
H
O
N
Me
Morphinone
COR (codeinone O-reductase)
HO
O
H
HO
N
Me
Morphine
CODM (codeine O-demethylase)

The ENZYMES are doing the transformations

6OMT

CNMT

NCS

NMCH

CODM

SalSyn

4'OMT

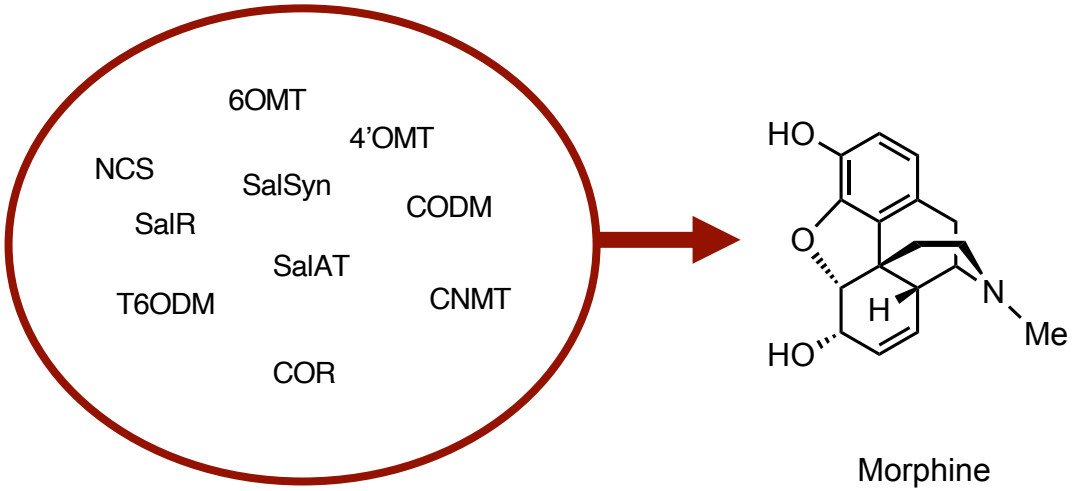
T6ODM

SalR

SalAT

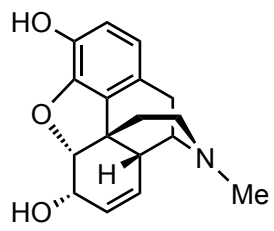
COR

*Second Question: How does the opium poppy know **how/when** to make the enzymes for morphine?*

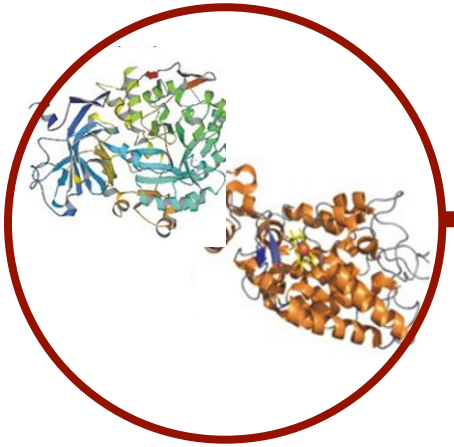


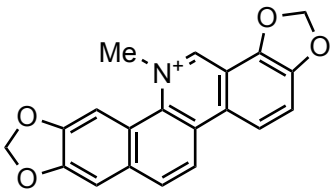
Morphine

And when/how to make the enzymes that make other compounds?

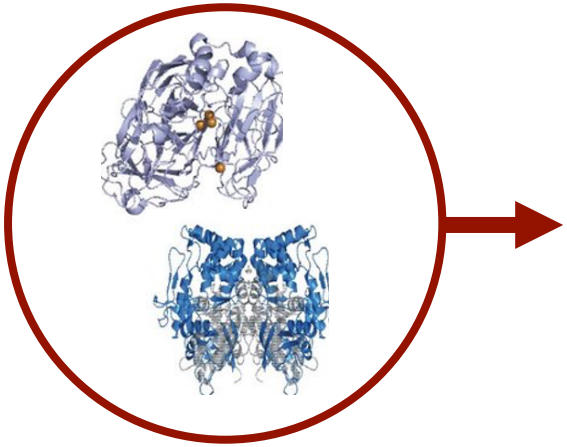


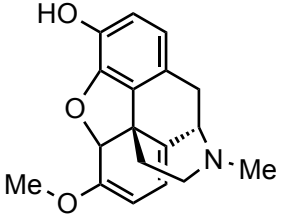
Morphine



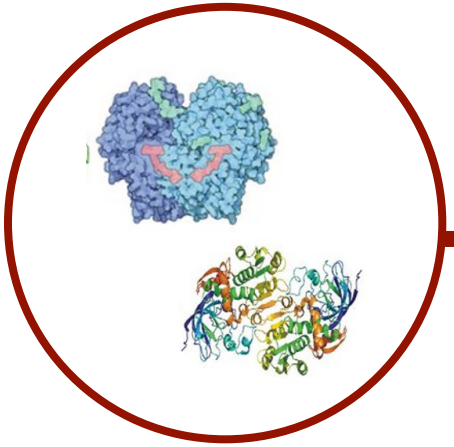


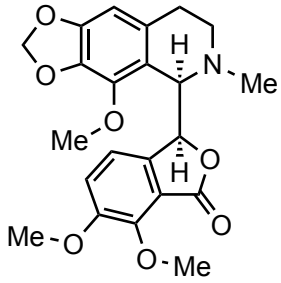
sanguinarine





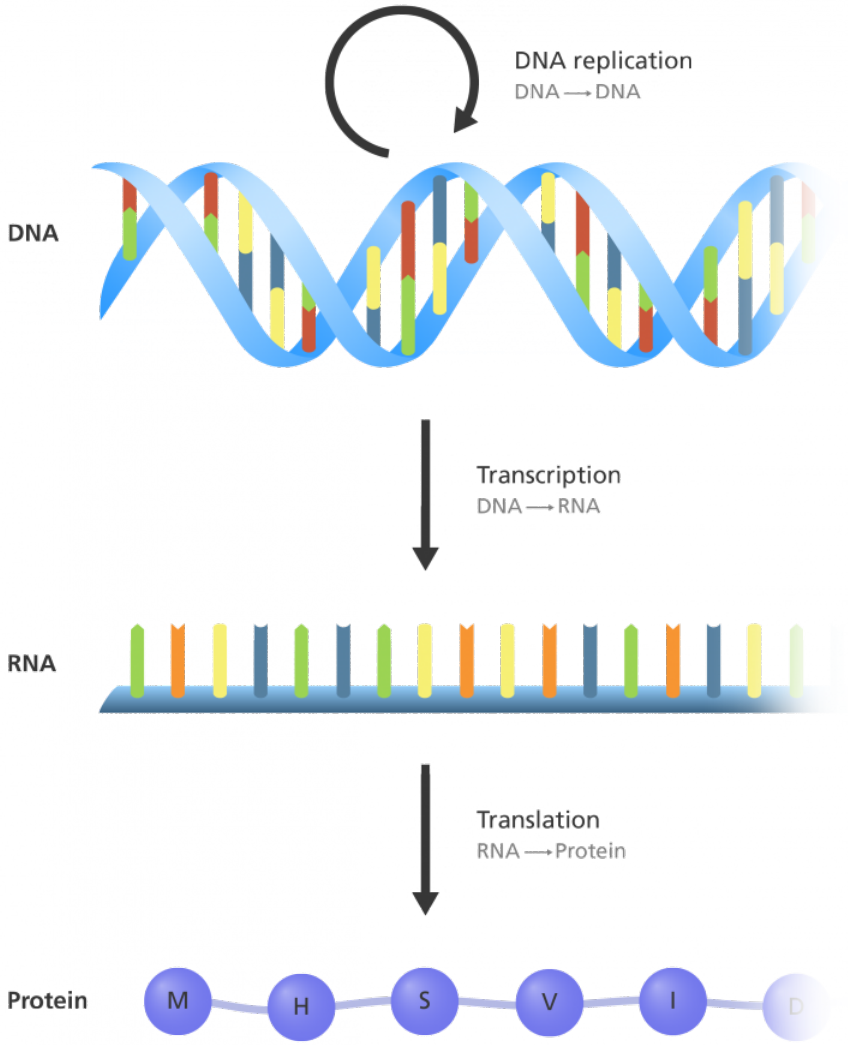
oripavine



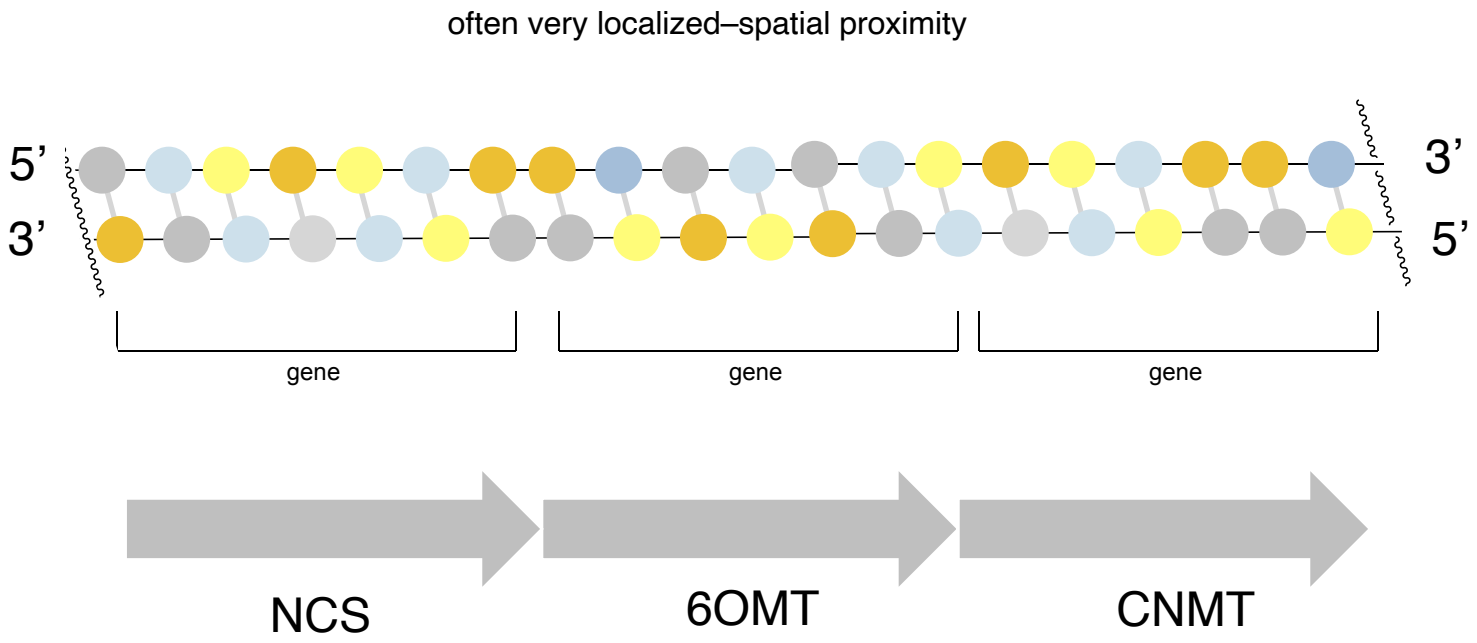


noscapine

Central Dogma Recap



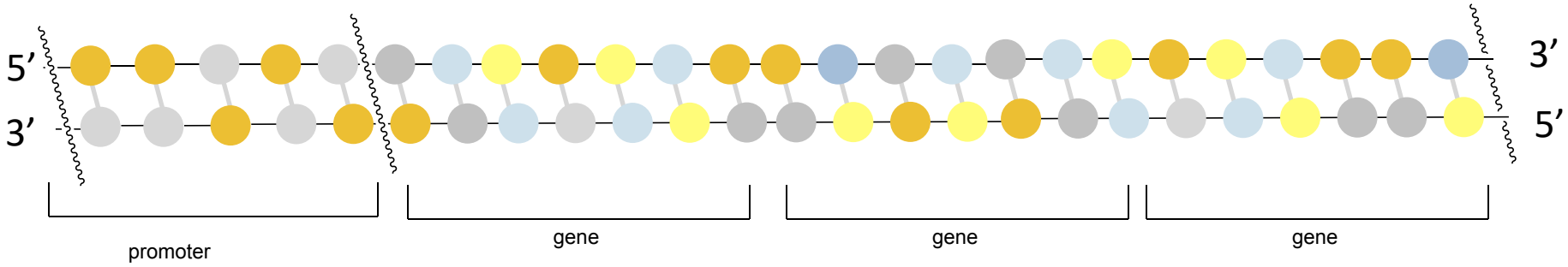
*biosynthetic gene **clusters** code for proteins that work together*

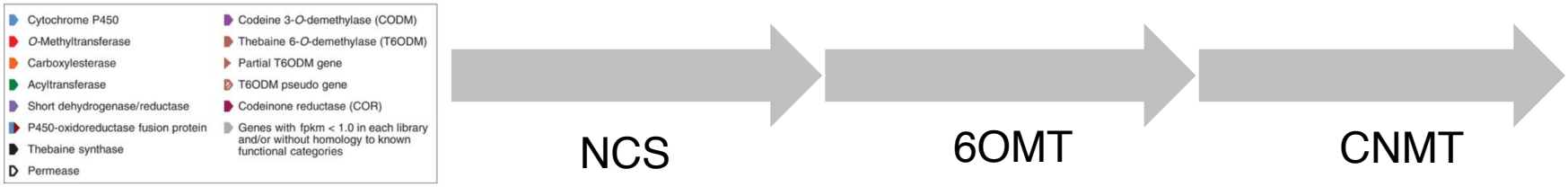


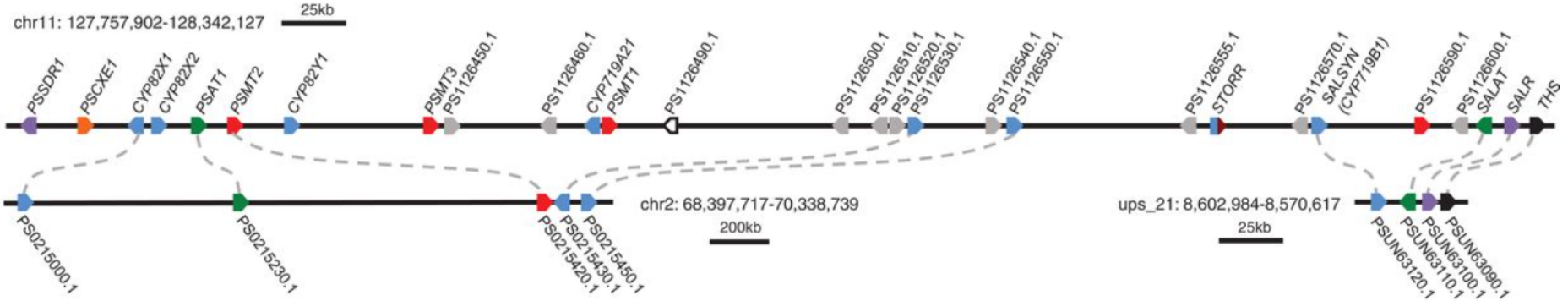
Cluster-specific ”promoter regions” regulate cluster expression

often only one promoter for multiple genes

often very localized–spatial proximity







A biosynthetic gene cluster:

a group of DNA bases that code for enzymes that will make a natural product ("secondary metabolite")

1. Expression of cryptic gene clusters

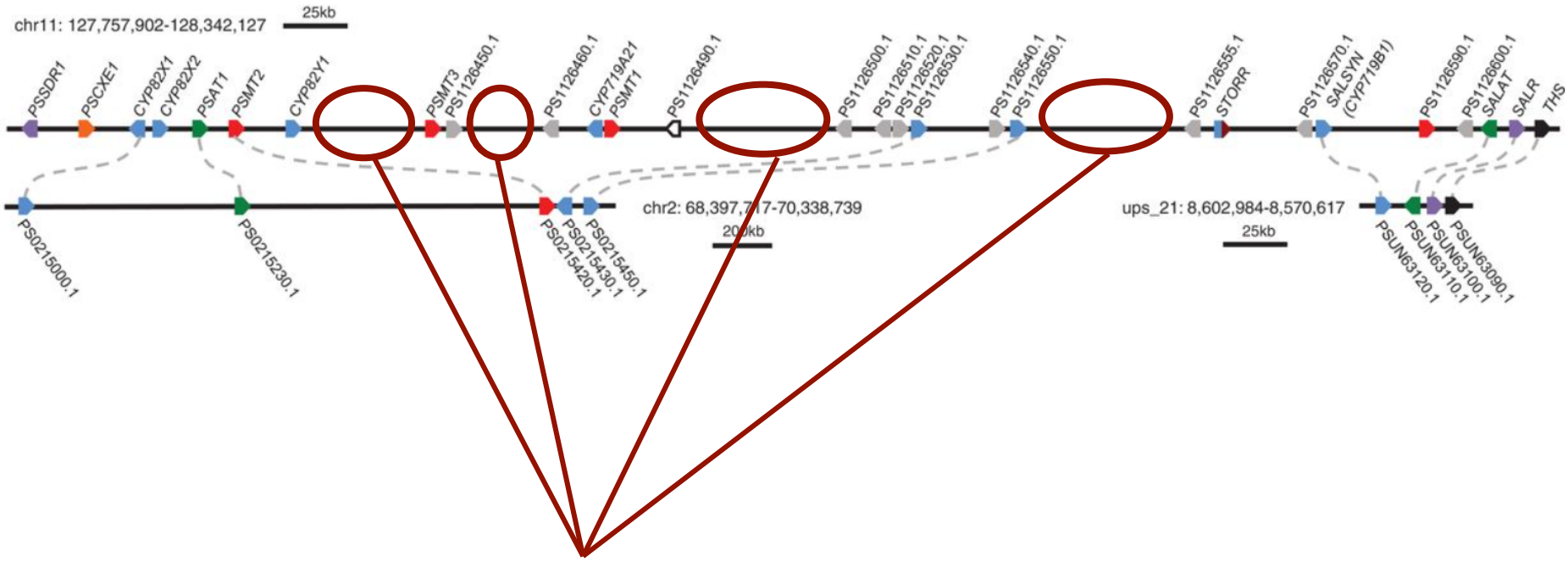
2. Expression of natural product intermediates for semisynthesis

3. Engineering gene clusters

Are these all the natural products can papaver somniferum possibly produce?

No ☺

Cryptic Gene Clusters for new natural products



silent or *silenc**ed**?*

Theoretically many many more NPs that could be highly beneficial than we even know about

(in too small quantities to be isolated under laboratory or wild conditions)

Our goal: turn on these clusters

Our problem: we can't know how to turn them on if we don't know where or what they are

The Hunt for New Natural Products (Genome Mining)

Uses ClusterFinder algorithm

Kai Blin; Victòria Pascal Andreu; Emmanuel L. C. de los Santos; Francesco Del Carratore; Sang Yup Lee; Marnix H. Medema; Tilmann Weber
Nucleic Acids Research 2018

1. Expression of cryptic gene clusters

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

Cite This: *J. Am. Chem. Soc.* 2019, 141, 769–773

pubs.acs.org/JACS

Genome-Mined Diels–Alderase Catalyzes Formation of the *cis*-Octahydrodecalins of Varicidin A and B

Dan Tan,[†,‡,⊥] Cooper S. Jamieson,[§,⊥] Masao Ohashi,[‡] Man-Cheng Tang,[*,‡] K. N. Houk,[*,‡,§] and Yi Tang[*,‡,§]

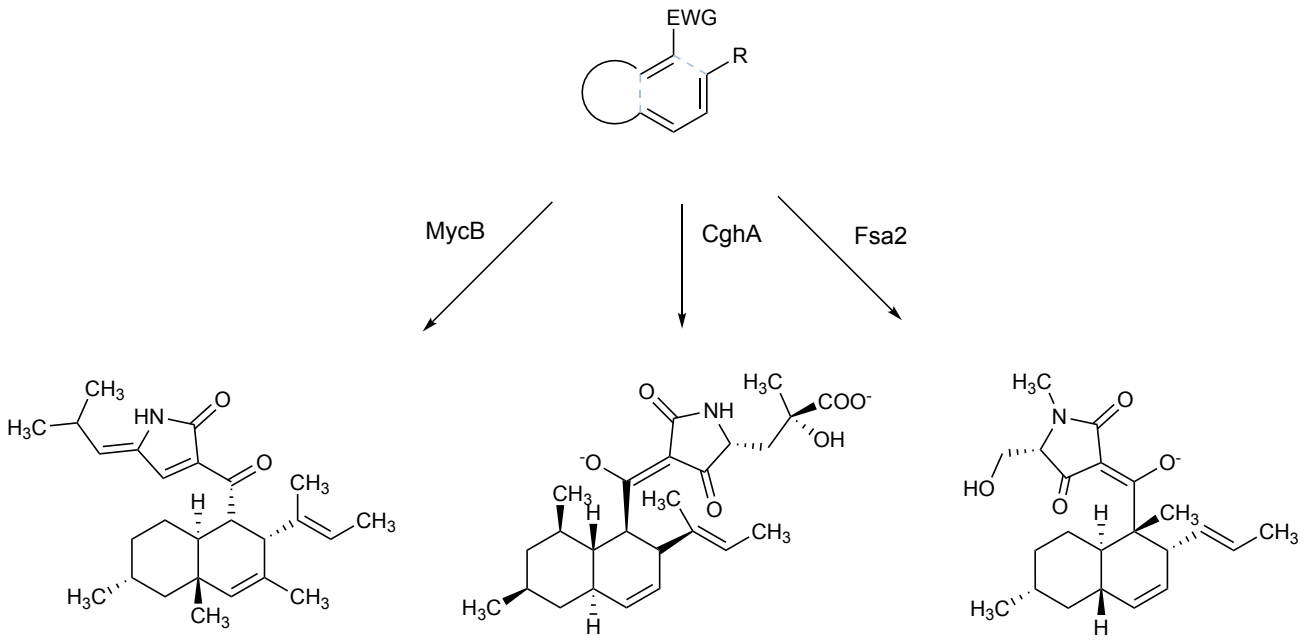
[†]Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

[‡]Department of Chemical and Biomolecular Engineering and [§]Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States

genome-mining enzymes to produce new natural products

motivation: we need new antibiotics!!

trans-decalins : *NP antibiotic class (encoded by DAses)*



myceliothermophin E

Sch201972

Equisetin

Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

motivation: we need new antibiotics!!

trans-decalins : *NP antibiotic class (encoded by DAses)*

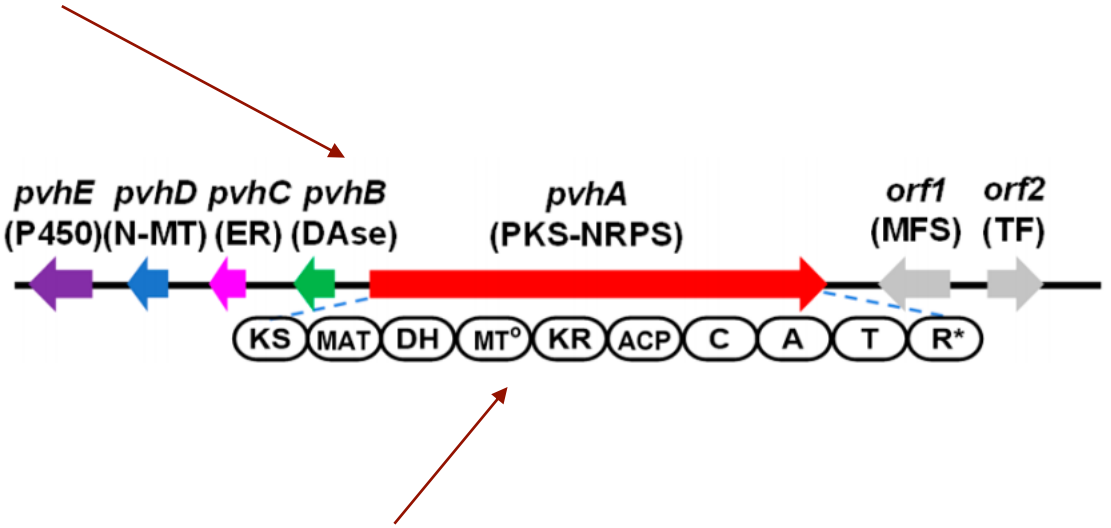
What about cis-decalins?

Not observed

The hunt for new natural products

Penicillium variabile has ***a cryptic gene cluster*** *(identified through sequence analysis)*

Diels-Alderase looking gene

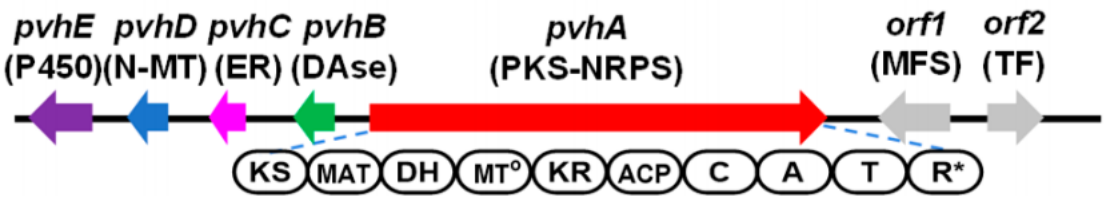


Pericyclase

polyketide synthase-nonribosomal peptide synthetase

Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

*We can match sequence to function **really well** through mining programs (homology)*



KS
ketosynthase

DH
dehydratase

DA
Diels-Alderase

MAT
Malonyl-CoA transferase

KR
ketoreductase

ACP
Acyl carrier protein

MT
methyltransferase

C
condensation

MFS
Major facilitator superfamily

R
reductase

A
adenylation

ER
enoylreductase

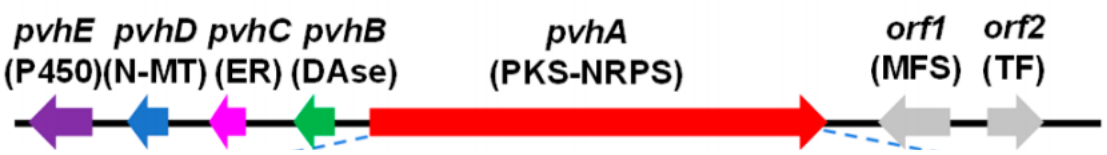
T
thiolation

TF
Transcription factor

N-MT
N-methyltransferase

Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

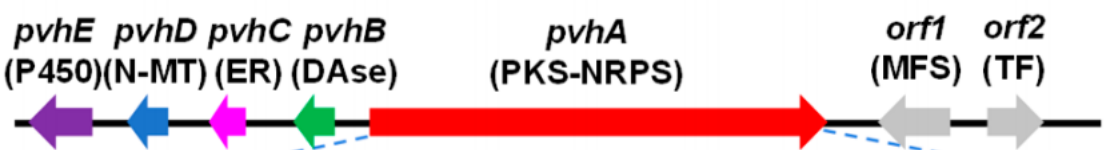
but that doesn't tell us exactly what it might make...



Problem : we don't know how to express this cluster in *penicillium variabile* (cryptic)

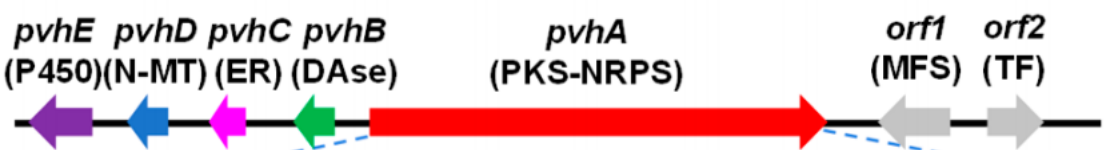
Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

but that doesn't tell us exactly what it might make...

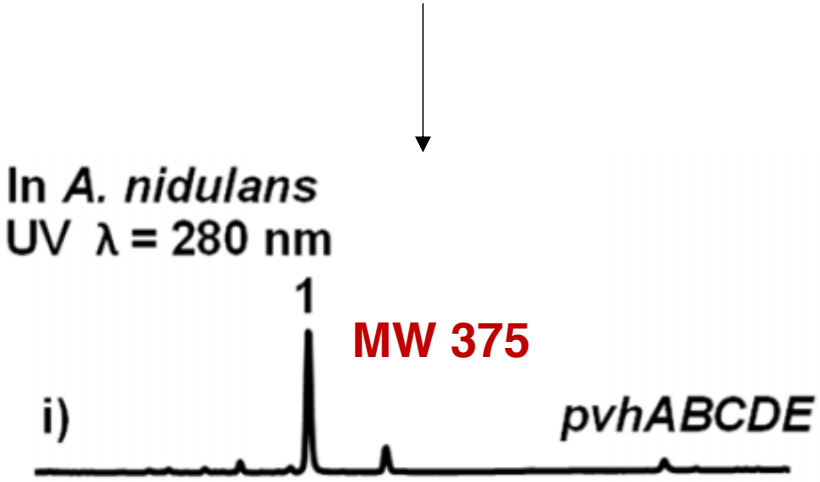


Solution trick: find a new *host*

Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

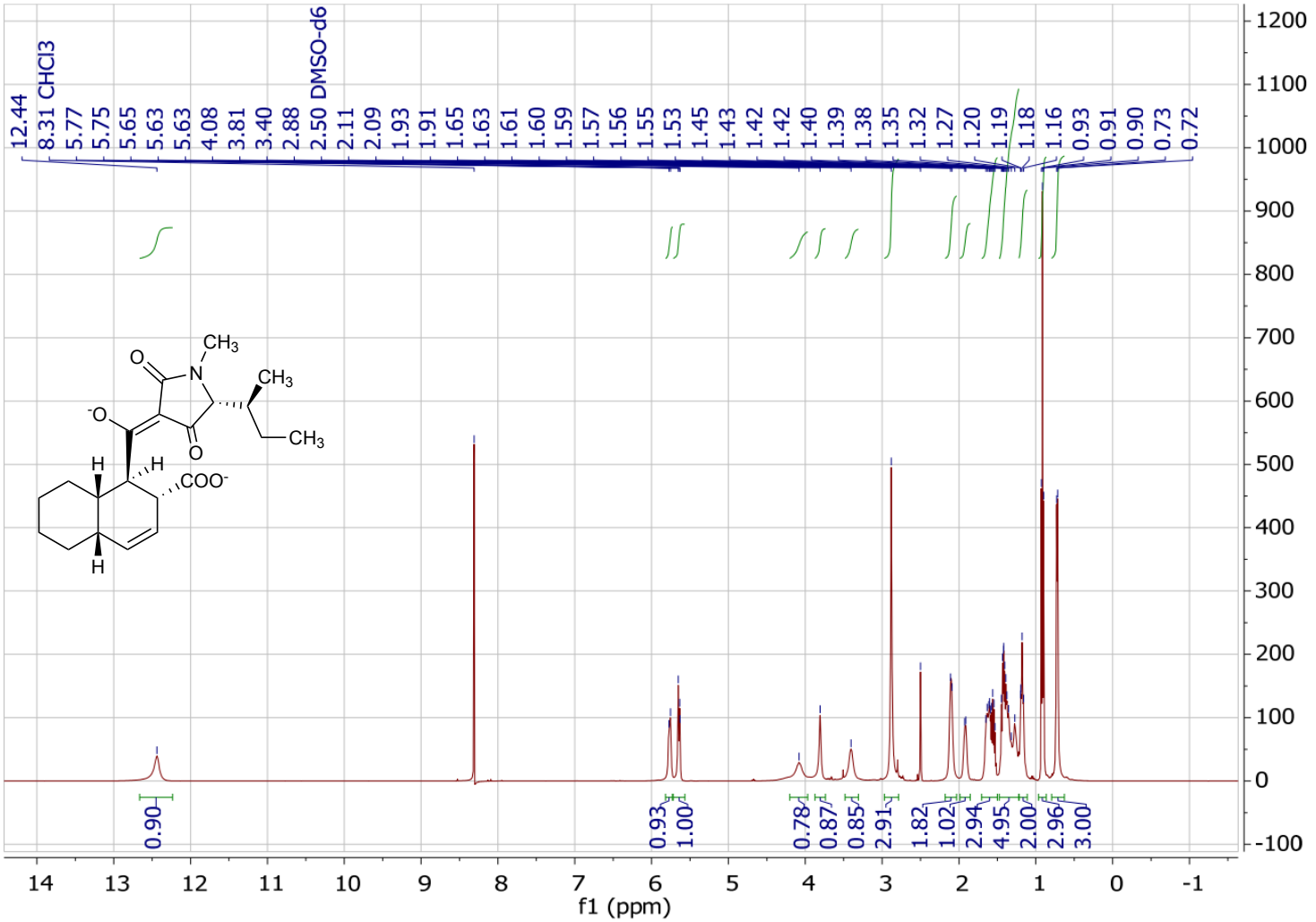


Expressed in *Aspergillus Nidulans*



Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

It's a trans-decalin!!!



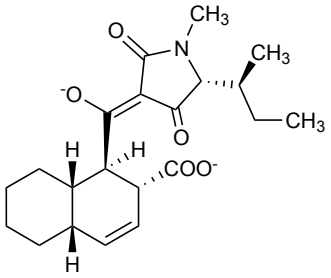
Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

Next Steps

Next part of paper was elucidating the exact biosynthesis of the cis-decalin (what each gene does and in what order...)

Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

Next Steps

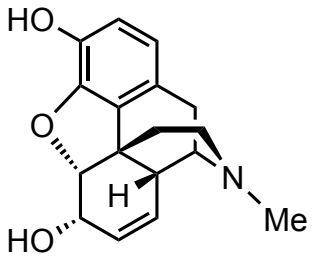


Next: biological evaluation for antibiotic activity

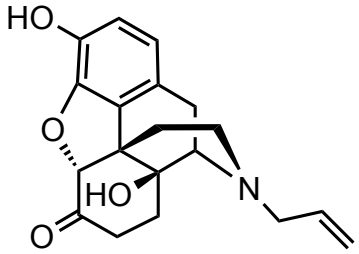
Dan Tan, Cooper Jamieson, Masao Ohashi, Man-Cheng Tang, K. Houk, Yi Tang. *JACS*. **2019**, *141*, 1796-1773

2. Expression of natural products/intermediates for semisynthesis

Natural Product SAR: Important but Challenging

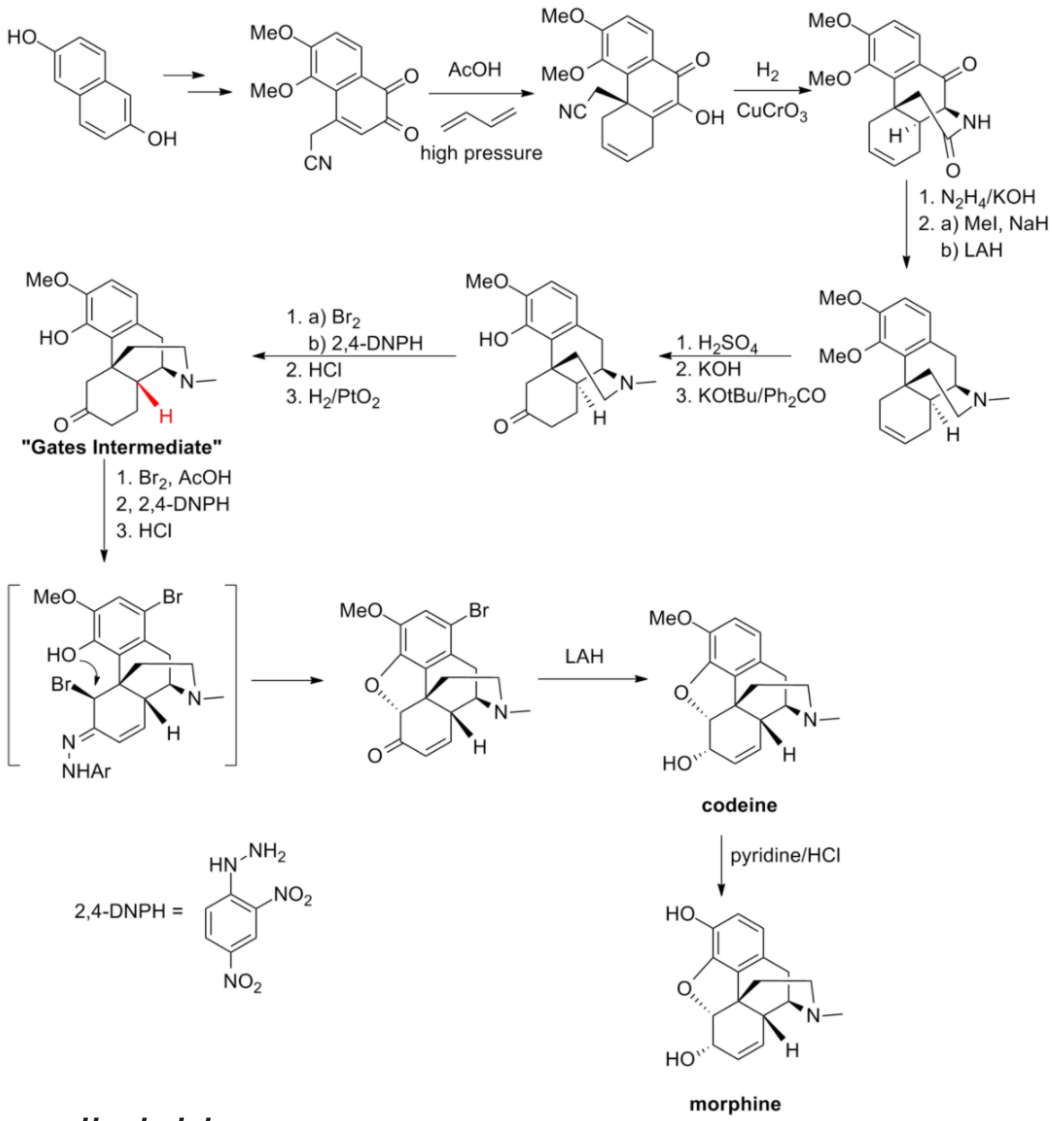


Morphine



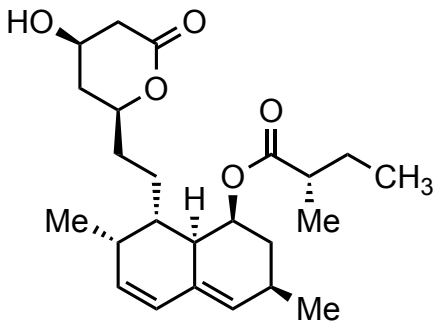
Naloxone

First Morphine Total Synthesis: Gates, 1952



31 steps and 0.06% overall yield

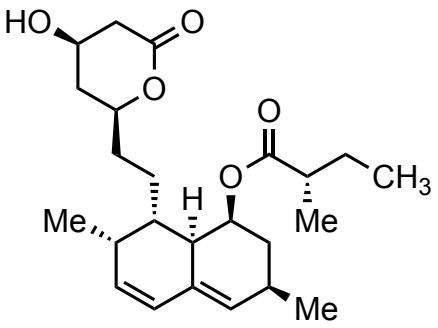
Lovastatin: A treatment for high cholesterol



lovastatin

treats high cholesterol

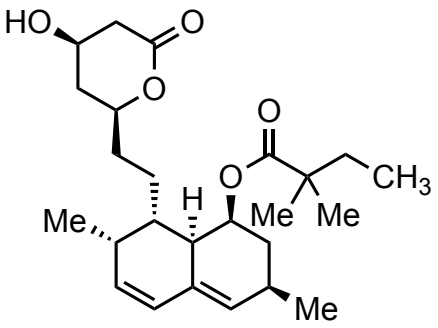
Lovastatin in a natural product



lovastatin

Biosynthetically produced *by fermentation from various fungi*

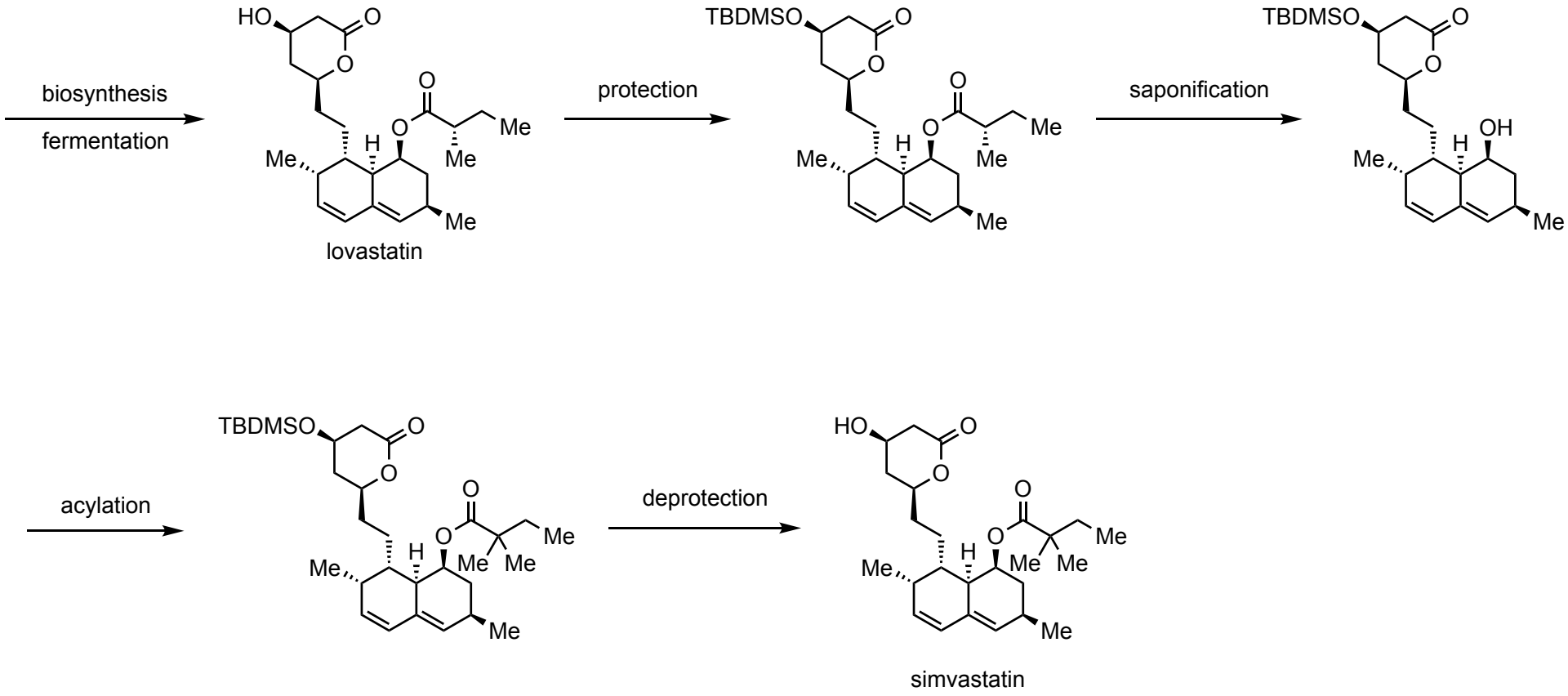
Simvastatin in a derivative of Lovastatin



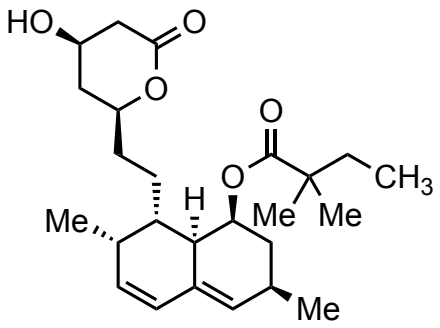
simvastatin

treats high cholesterol ***much more effectively***

Industrial synthesis of Simvastatin



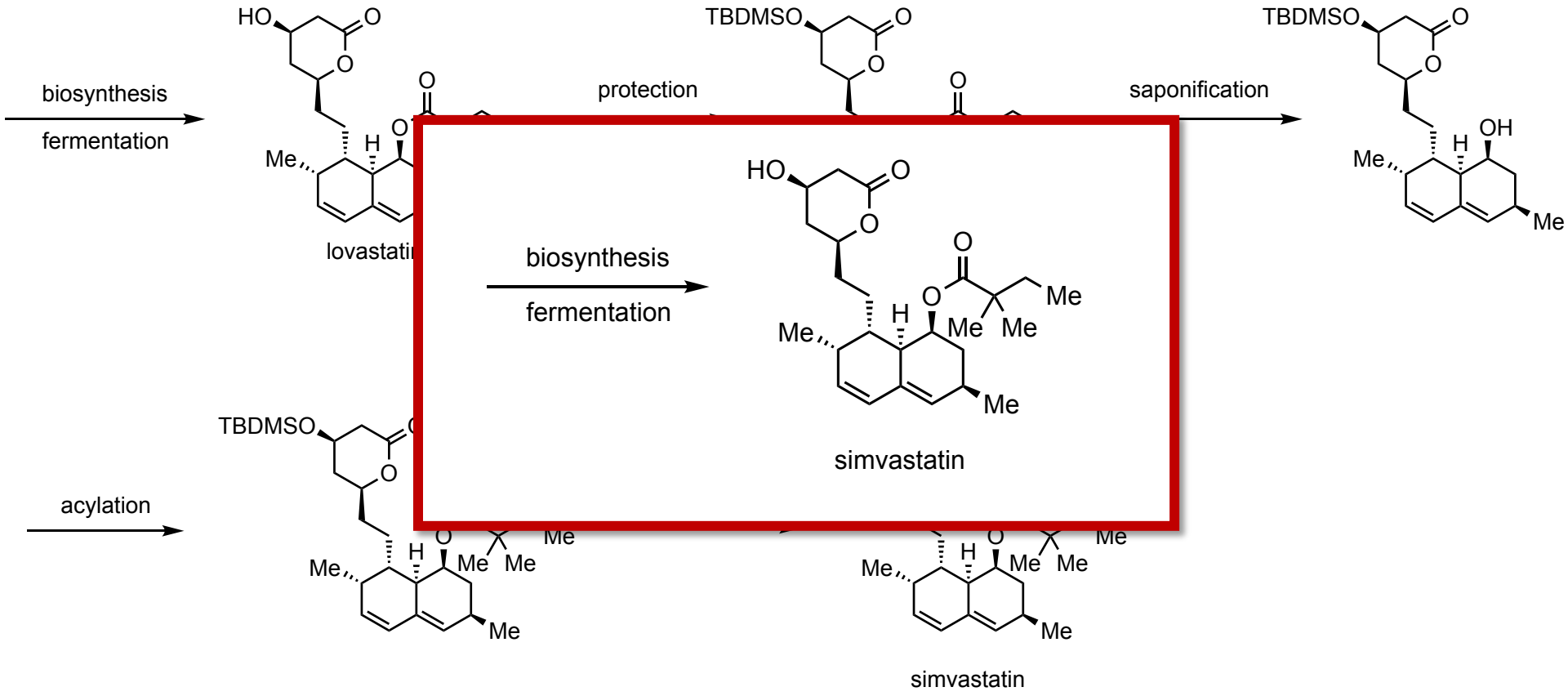
W. F. Hoffman, A. W. Alberts, P. S. Anderson, J. S. Chen, R. L. Smith, and A. K. Willard* *J. Med. Chem.* **1986**, 29, 849-852



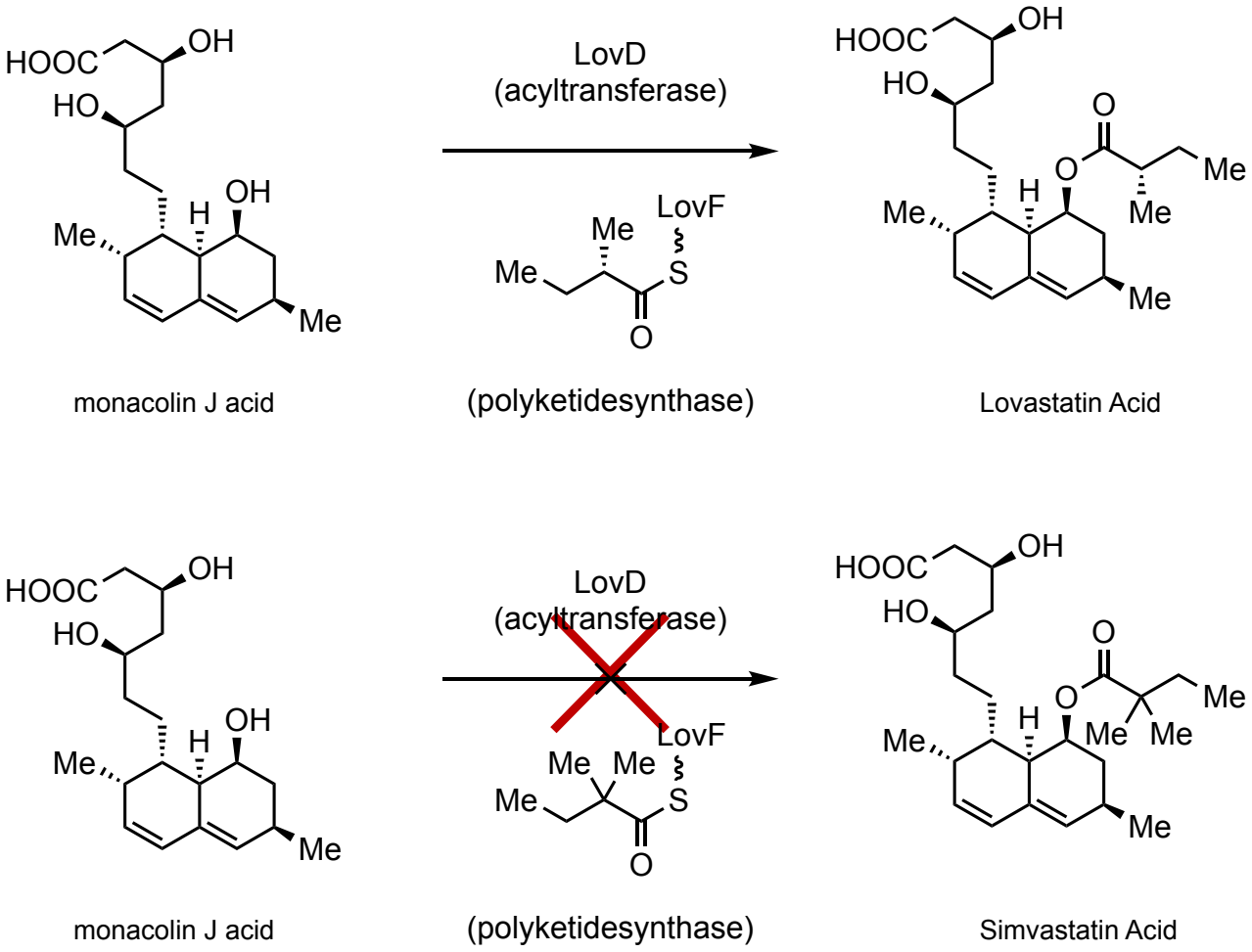
simvastatin

Semi-synthetically *produced from Lovastatin*

Completely biosynthetic route?



The problem: natural enzyme will not accept the substrate for Simvastatin



3. Engineering gene clusters

Can we use this strategy to produce designed therapeutics using biosynthesis?

Contents lists available at ScienceDirect

Metabolic Engineering

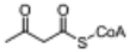
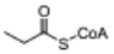
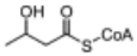
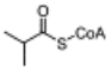
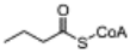
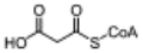
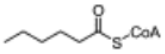
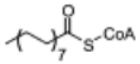
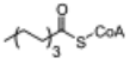
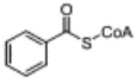
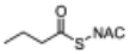
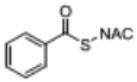
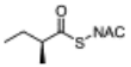
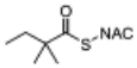
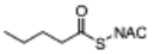
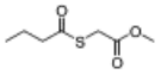
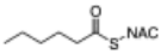
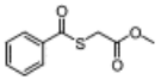
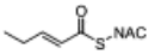
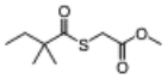
journal homepage: www.elsevier.com/locate/meteng

Engineering *Saccharomyces cerevisiae* for production of simvastatin

Carly M. Bond[a], Yi Tang[a,b,*]

[a] *Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095, United States*
[b] *Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, United States*

WT LovD only converts to 10% , but engineered LovD (Lov9) converts >99%

Acyl Thioester Substrate	Conversion (%)[b]/RT (min)[c]	Acyl Thioester Substrate	Conversion (%)/RT (min)
	7/5.1		89/4.6
	35/6.0		35/4.2
	52/6.8		6/6.5
	87/6.8		NR
	32/8.7		NR
	7/10.6		69/7.6
	50/6.8		58/7.6
	22/7.6		10/8.5
	52/7.8		92/6.8
	33/8.7		70/7.6
	2/7.5		17/8.5

Xinkai Xie, Kenji Wantanabe, Wladyslaw Wojcicki, Clay Wang, Yi Tang. **2006,** *13*, 1161 – 1169

New Problem: Native Host is resistant to engineering

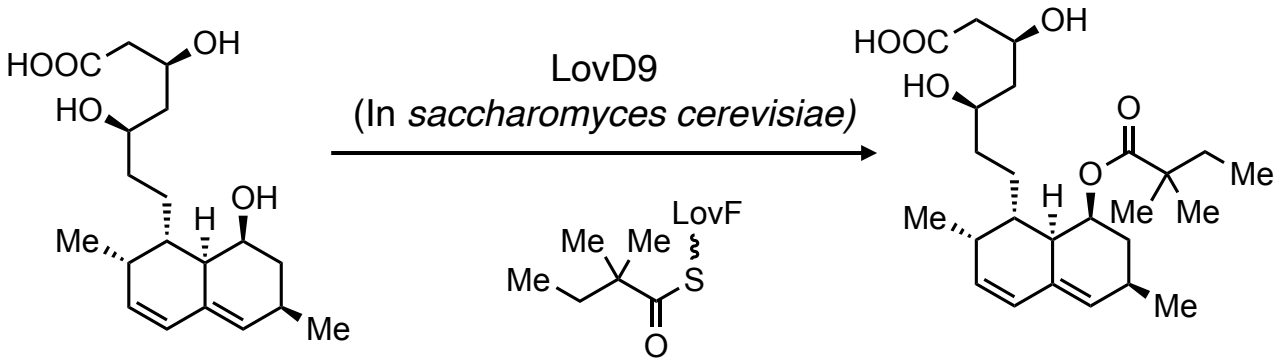
Natural host: *A. terreus.* Resists engineering (can't insert Lov9)

E-Coli: can produce LovD9, but not LovF

Co-cultures: not trivial

Saccharomyces cerevisiae: can express both

Engineering saccharomyces cerevisiae for production of simvastatin



monacolin J acid

Simvastatin Acid

95% yield, 55 mg/L

Carly Bond, Yi Tang. *Molecular Engineering,* **2019,** *51*, 1 – 8

Personal outlook/discussion/questions

---

1. Expression of cryptic gene clusters

2. Expression of natural product intermediates for semisynthesis

3. Engineering gene clusters

- Can you build from the ground up?

- Design a biosynthetic gene cluster from scratch to make exactly what you want?