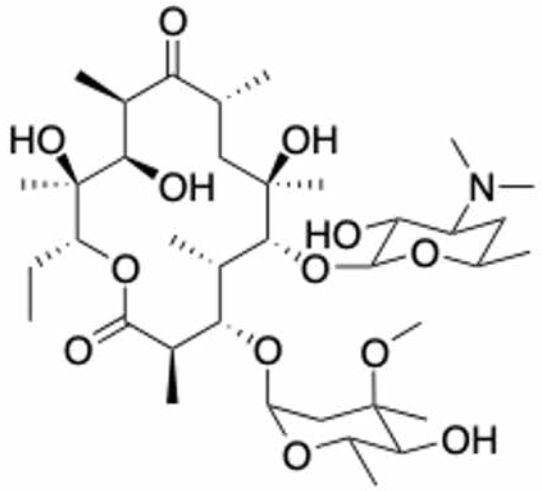
DOE Bioenergy Technologies Office (BETO)
2021 Project Peer Review

Accelerating polyketide synthase engineering for high TRY production of biofuels and bioproducts

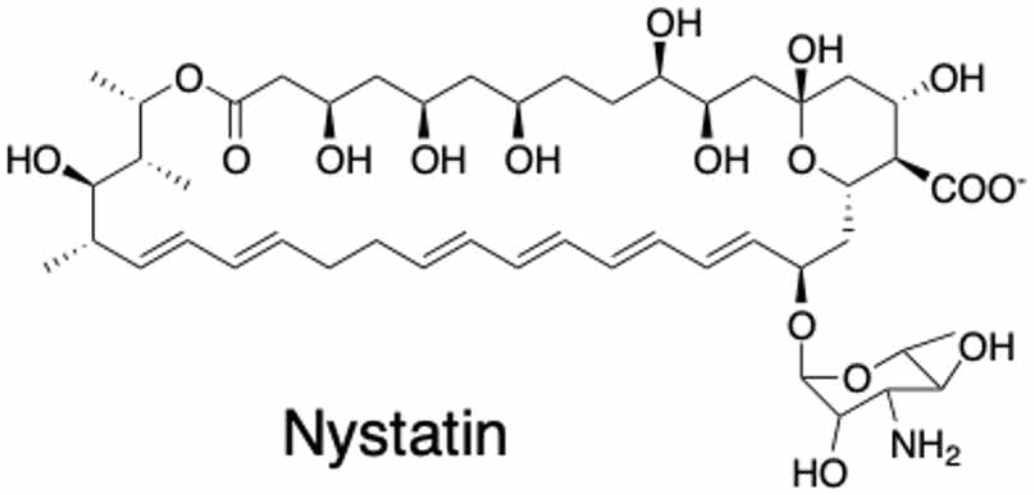
3/11/2021
Technology Area Session

Jay Keasling
University of California, Berkeley

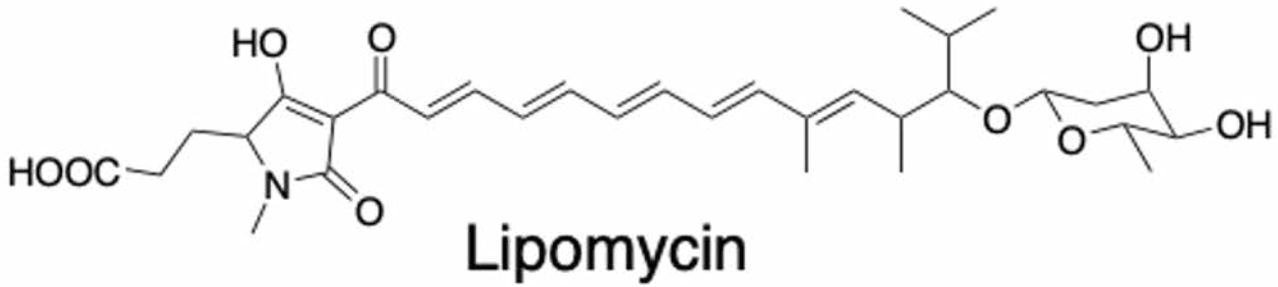
Project overview: What is a polyketide?



Erythromycin

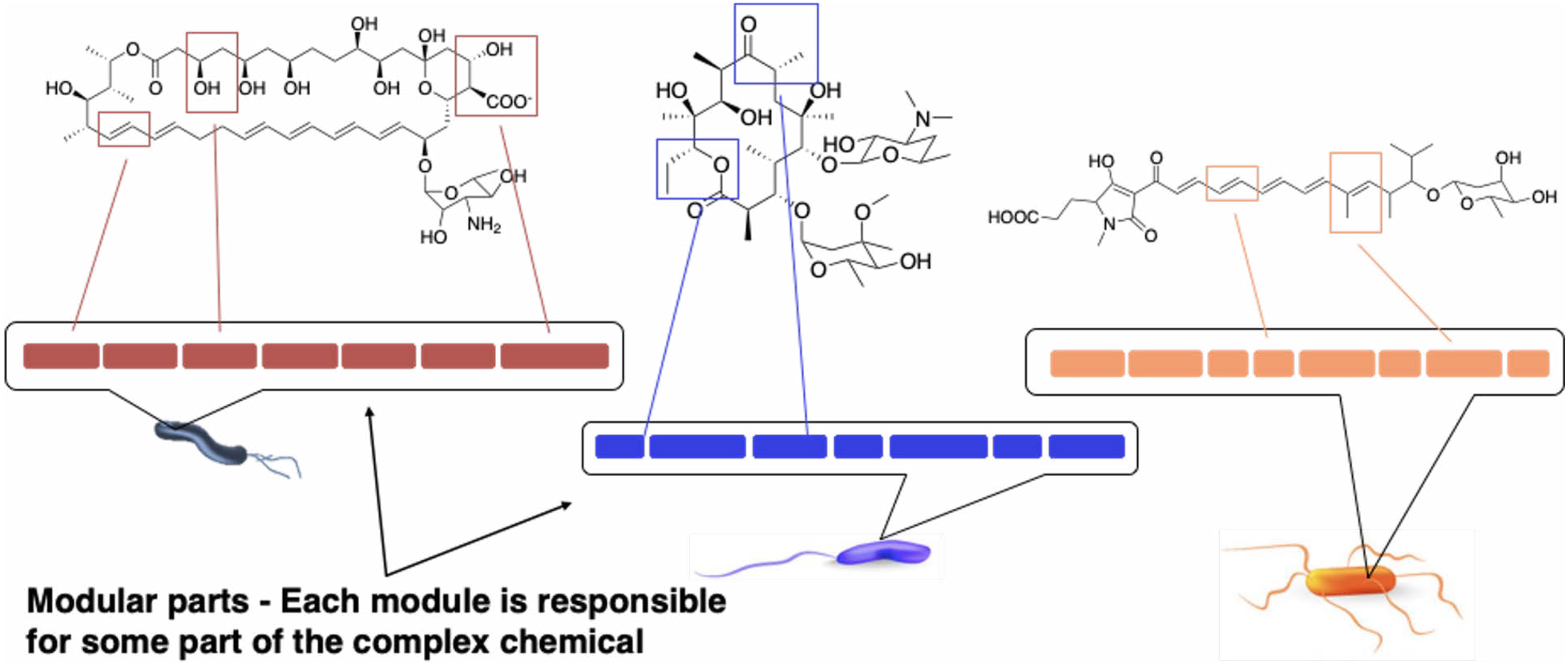


Nystatin



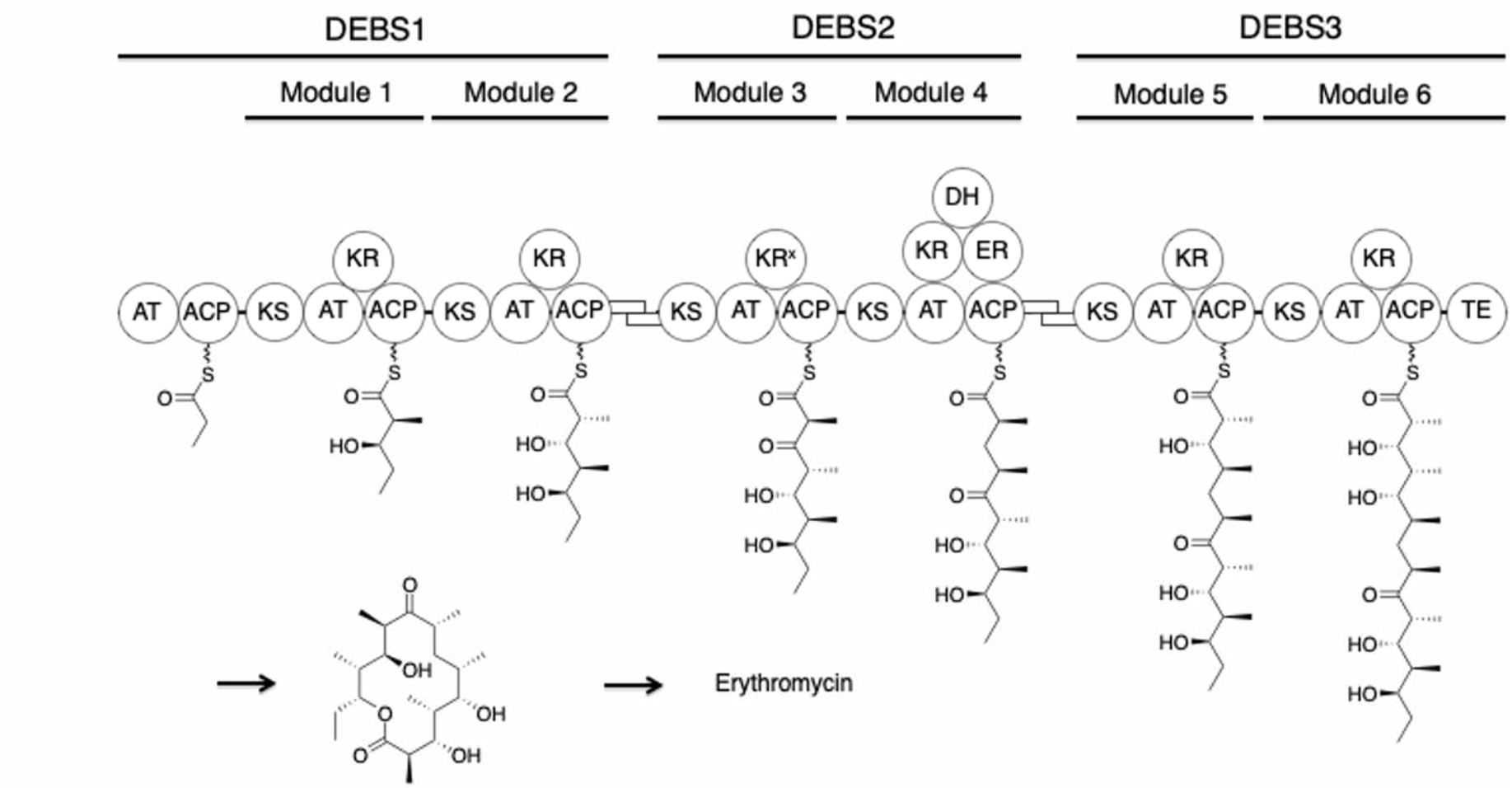
Lipomycin

Polyketides are produced using modular enzymes: each module is responsible for part of the molecule



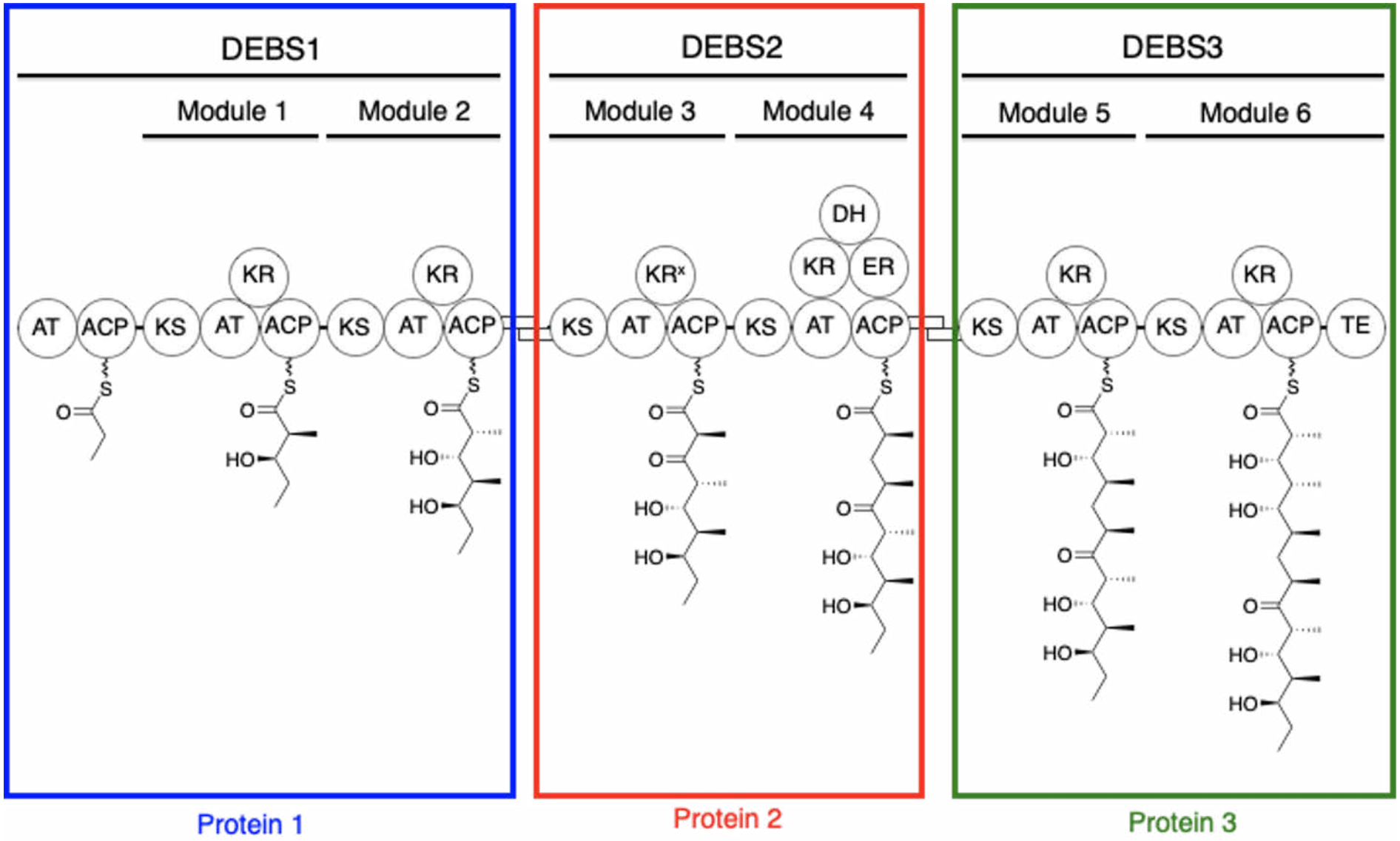
Modular parts - Each module is responsible for some part of the complex chemical

Erythromycin synthase is a well-studied polyketide synthase (PKS)

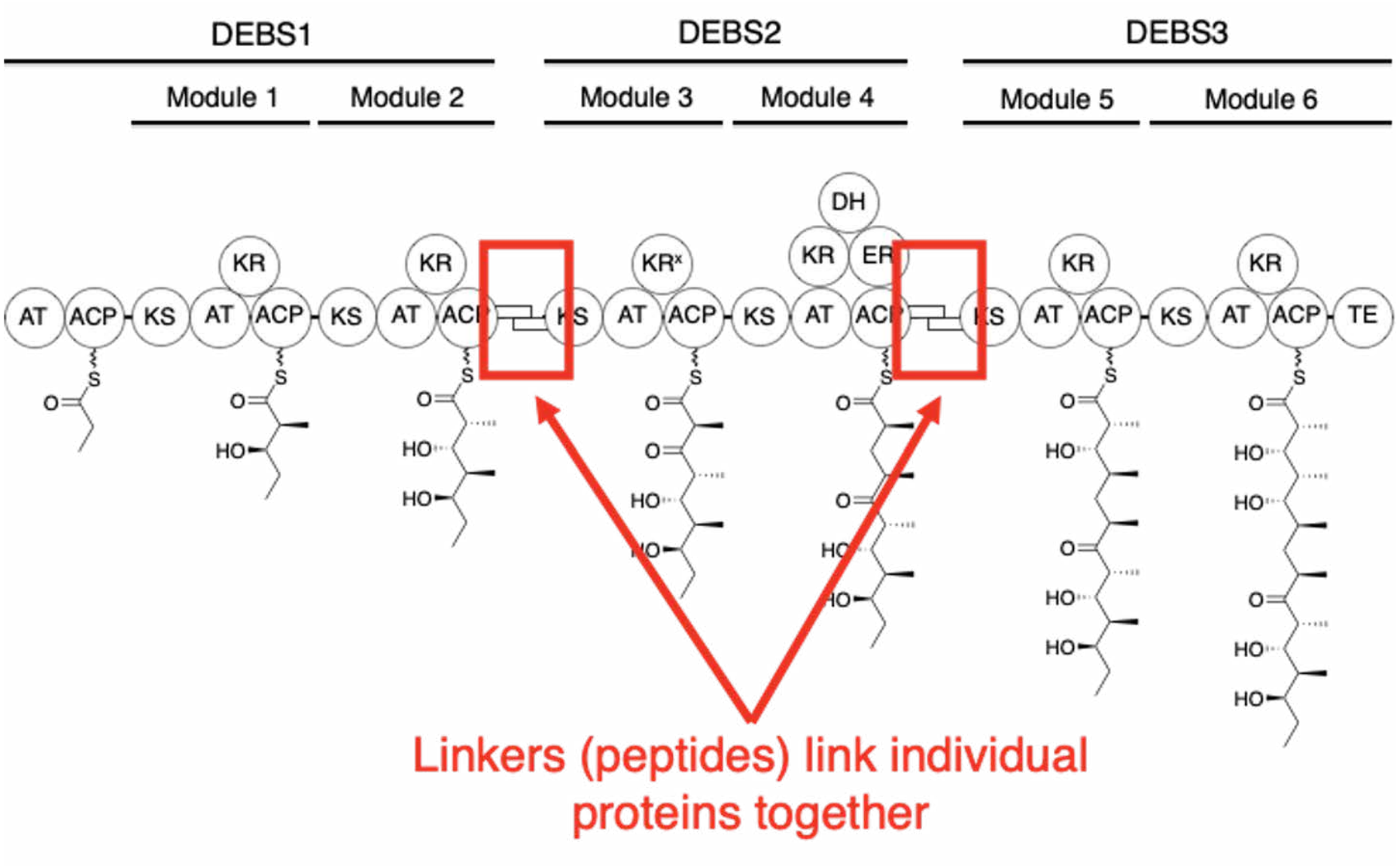


Donadio 1991 *Science*

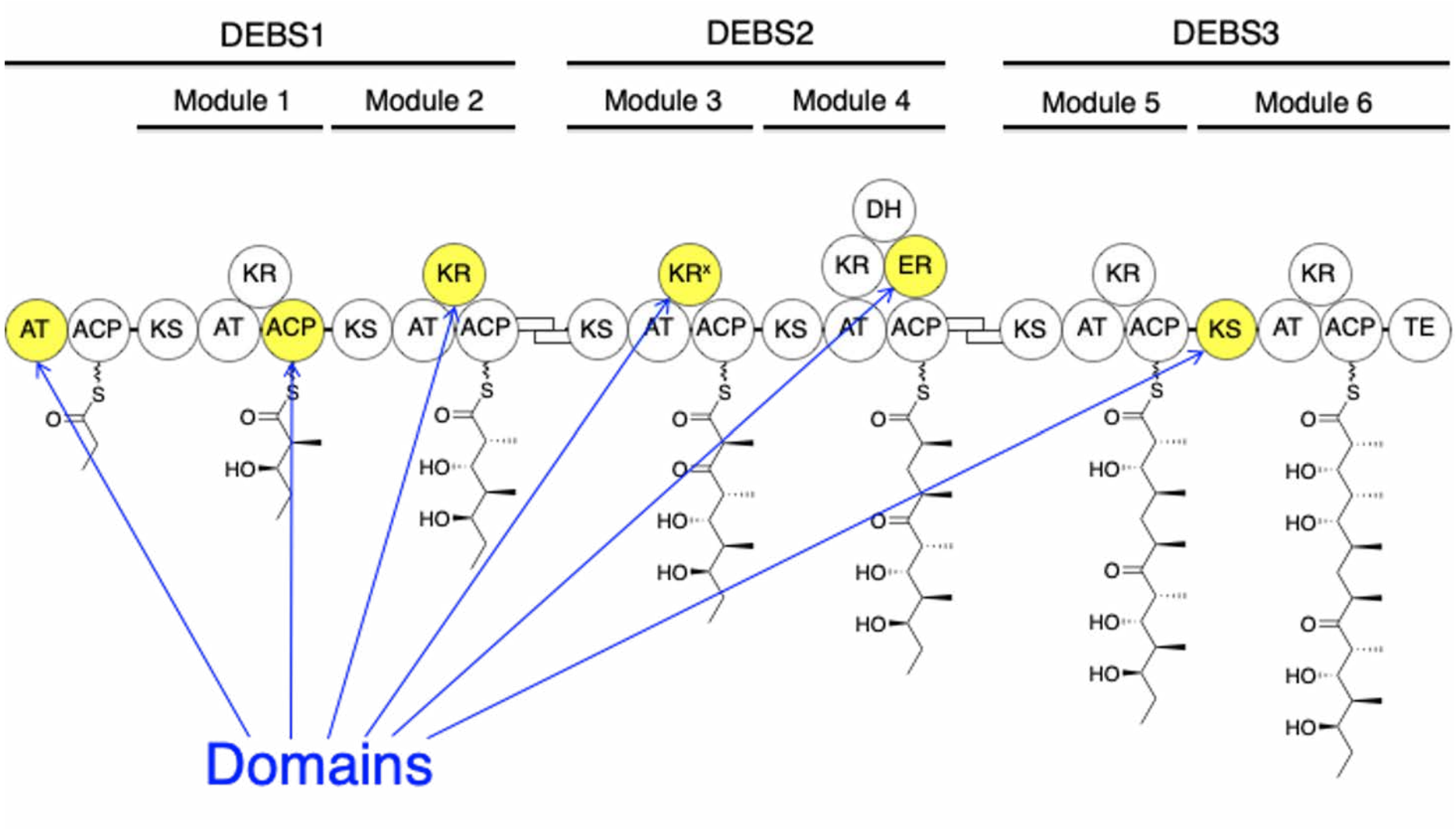
Polyketide synthases (PKSs) are large, multi-activity proteins



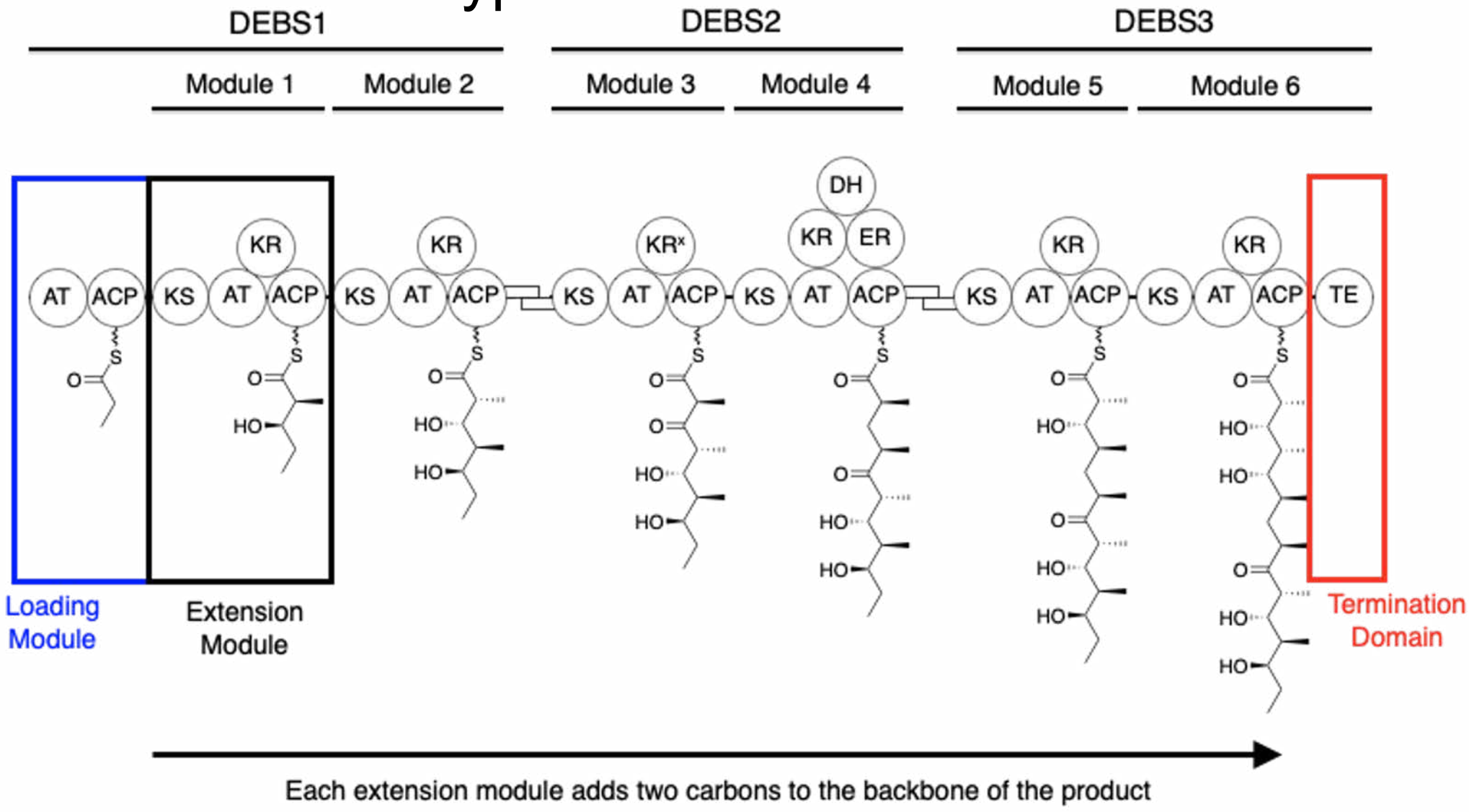
Polyketide synthases (PKSs) are large, multi-activity proteins



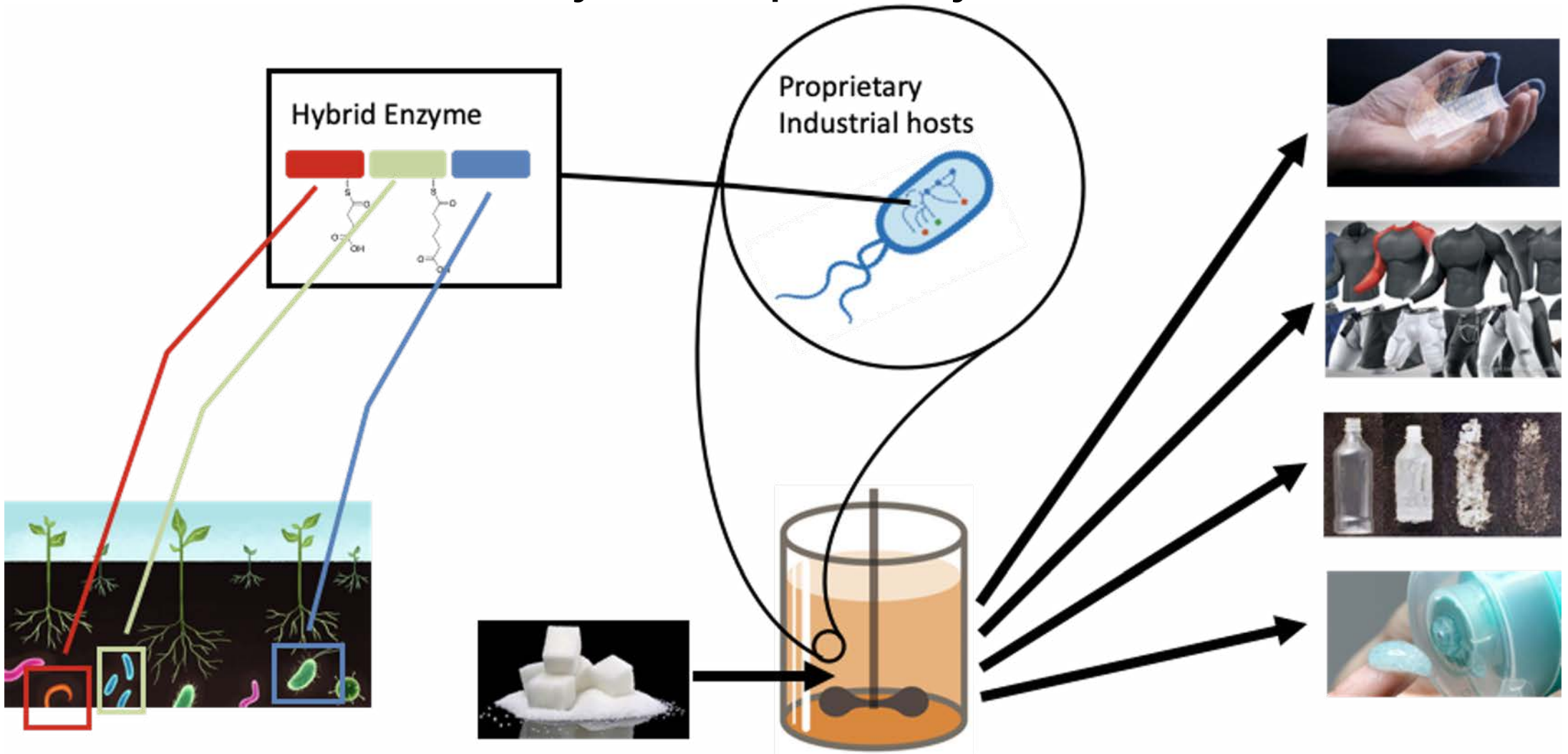
Each domain in a PKS has an individual enzyme activity



Polyketide synthases generally have Load and Extension Modules and some type of Termination



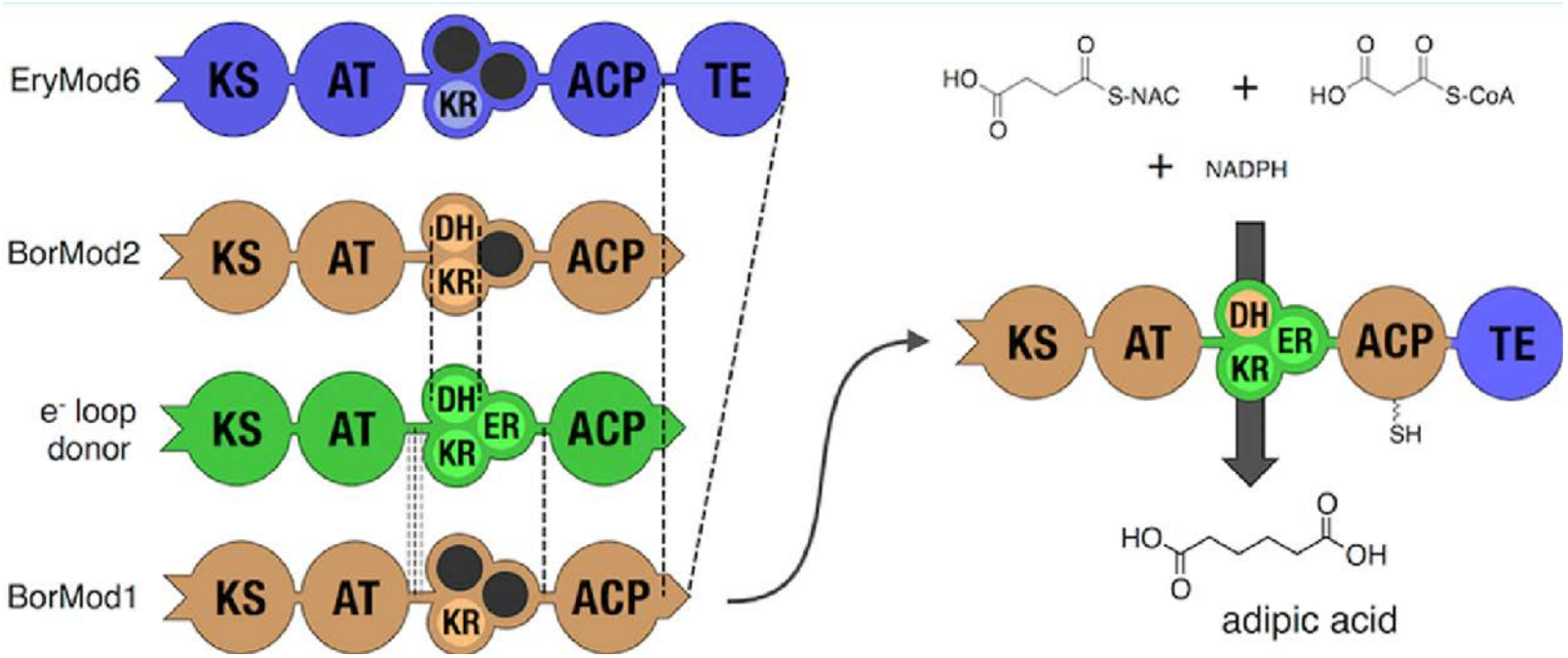
What we are trying to do: engineer PKSs for production of commodity and specialty chemicals



Develop a high throughput design-build-test-learn cycle for polyketide synthase (PKS) engineering, and test it by building a PKS to make the Nylon-6 monomer caprolactam, and novel caprolactam derivatives.

How is it done today and what are the limits

Currently new PKSs are constructed via trial and error, and can take years to build successfully.

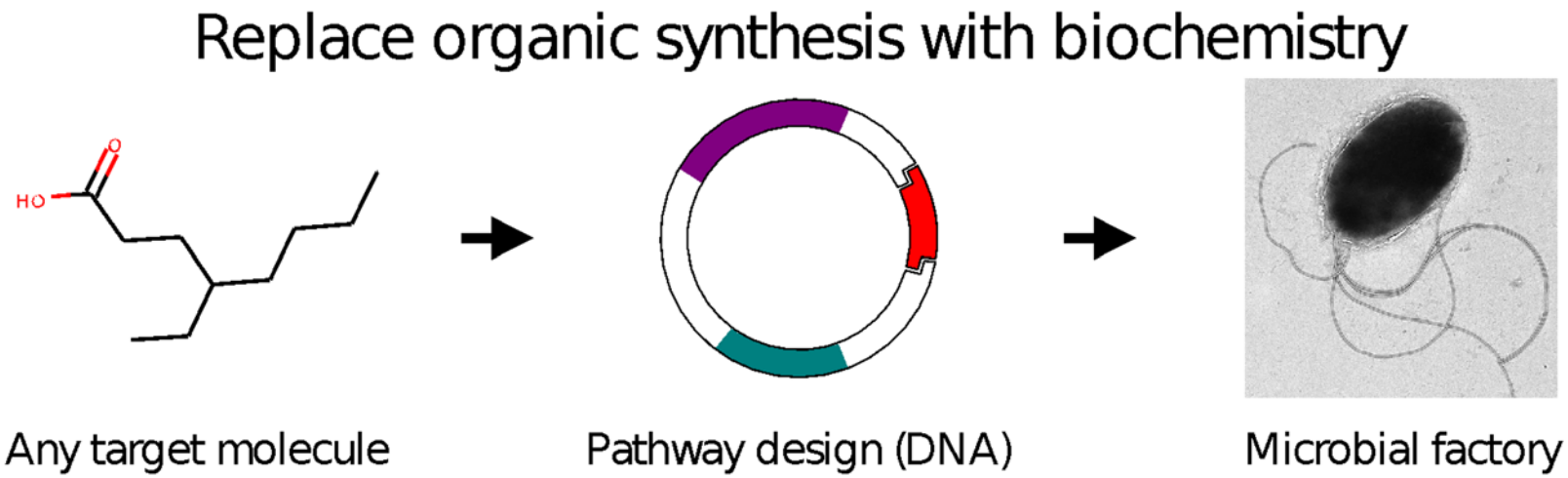


A novel PKS producing adipic acid built via extensive trial and error

(Hagen *et al.* ACS Synthetic Biology, 2015)

Why it is important

- PKSs provide the potential to access a massive chemical space of useful small molecules, and produce these from renewable carbon sources
- Potential to access new molecules with drastically improved properties



What are the risks

- Failure of protein folding or catalytic activity for unknown reasons
- Synthesizing and constructing DNA for large enzymes
- Expressing the engineered PKS *in vivo*
- Completing the Design-Build-Test-Learn cycle fast enough

Later in the presentation we will discuss how these risks are mitigated by our project plan.

Management: Team members

Tyler Backman

Sarah LaFrance

Hector Garcia Martin

Tijana Radivojevic

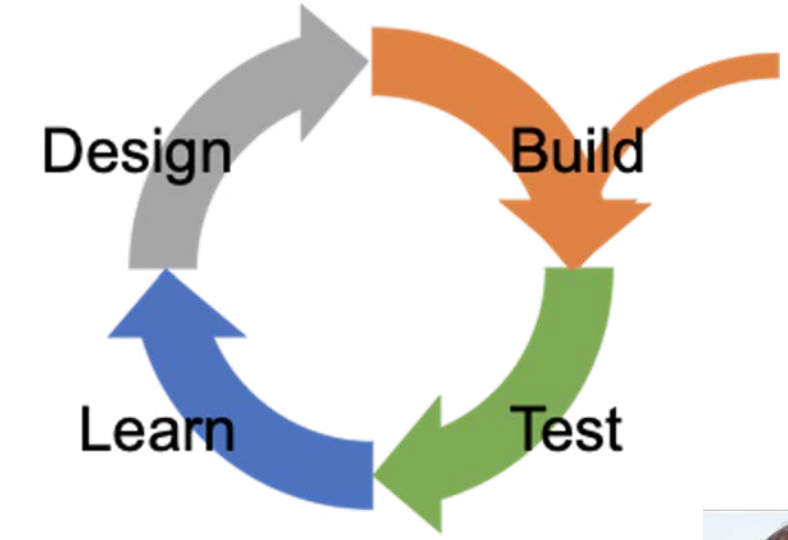
Christopher Johnson

Davinia Salvachua

Jay Keasling

Bob Haushalter

Eunice Kim



Matthias Schmidt

Anna Lisa Fear

Clem Fortman

Philip Laible

Peter Larsen

Nicholas Dylla

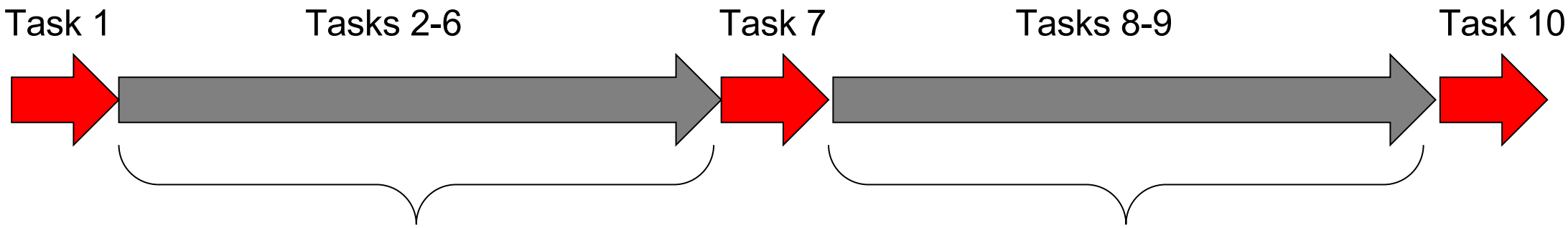
Christopher Petzold

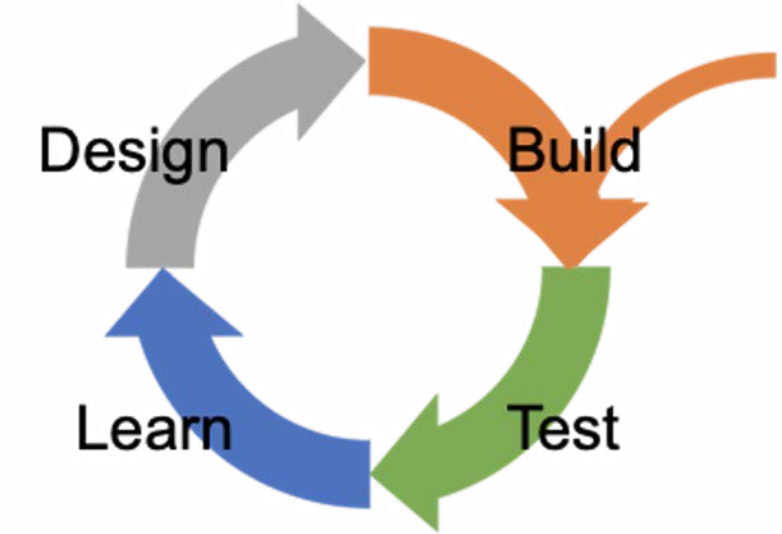
Jon Magnuson

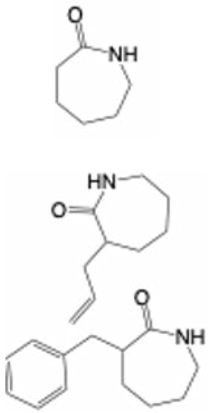
Kristin Burnum-Johnson

Young-Mo Kim

Program design







Communication and collaboration

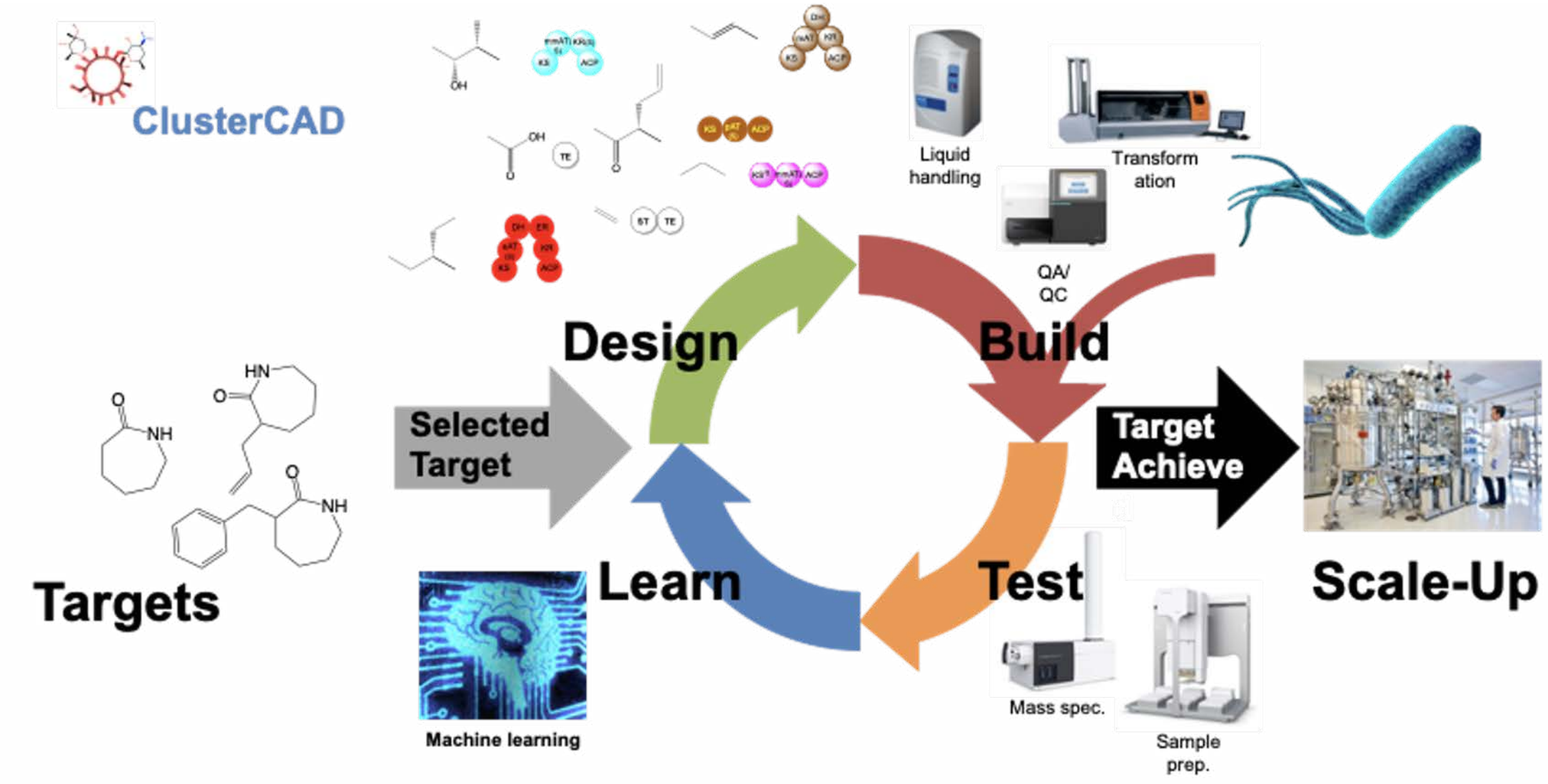
- Organized into Design, Build, Test, Learn teams
- We have bi-weekly meetings with all teams, and a rotating presentation schedule
- Teams meet individually as needed
- We provide quarterly reports and progress update presentations to the EERE

Management: we collectively identified key risks and developed a plan to mitigate each

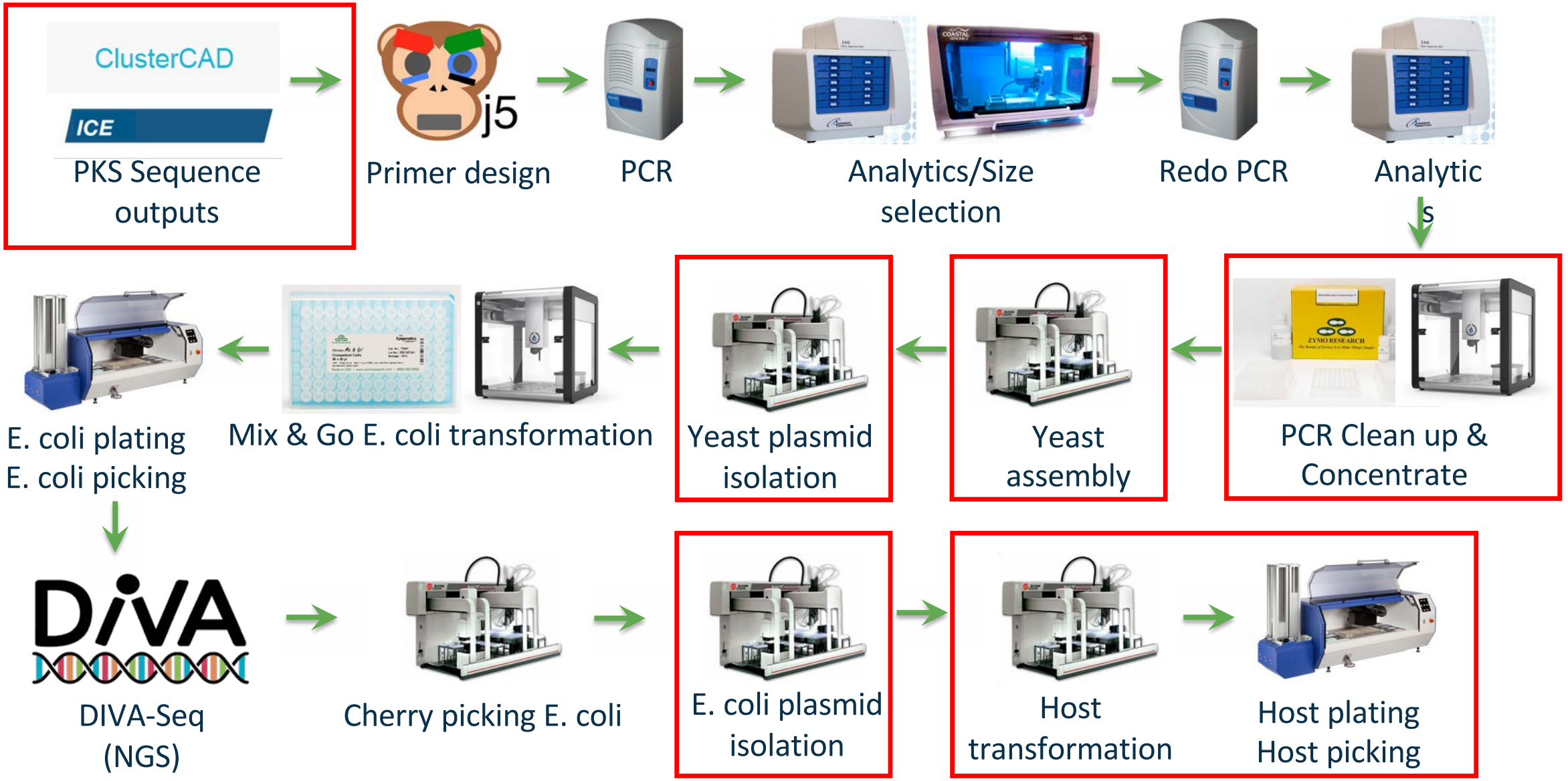
Key Risks

- Failure of protein folding or catalytic activity for unknown reasons
  Mitigation approach: With an extremely large volume of PKSs built, some will likely work, and can begin to inform models of why they sometimes fail.
- Synthesizing and constructing DNA for large enzymes
  Mitigation approach: A large number of diverse designs will ensure that some will be successfully constructed. Our approach is currently working for many different PKS designs.
- Expressing the engineered PKSs *in vivo*
  Mitigation approach: We have successfully expressed PKSs in our host organism thus far, so a high diversity and volume of designs can also overcome this risk.
- Completing the Design-Build-Test-Learn cycle fast enough
  Mitigation approach: We have robotically automated the slowest and most time consuming steps.

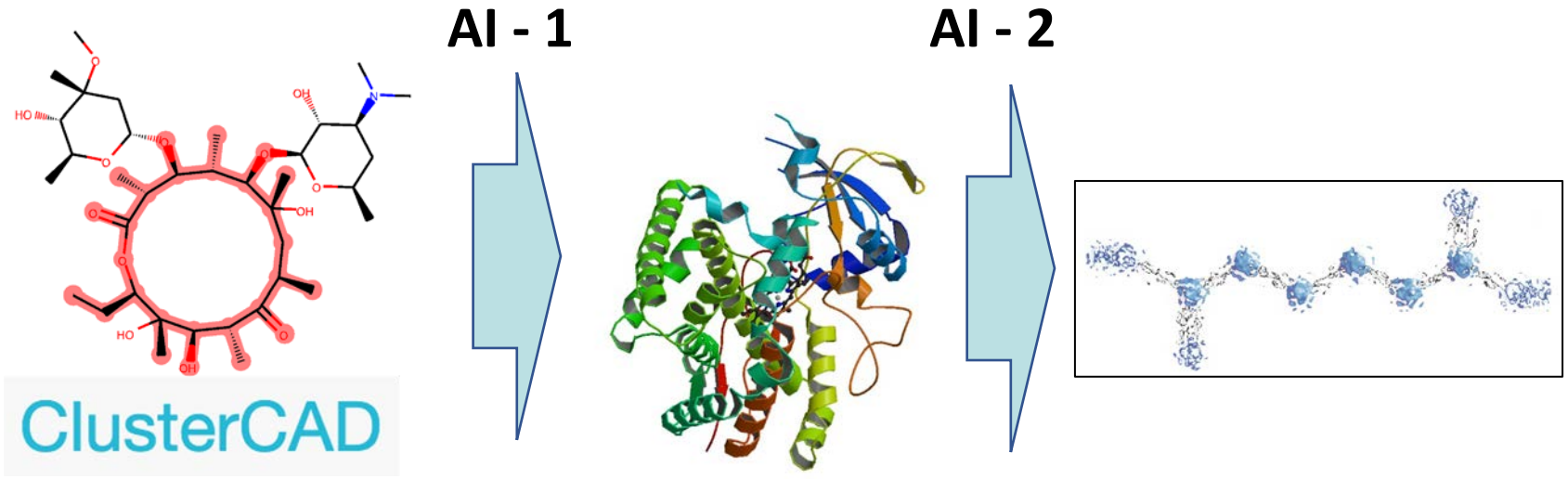
Approach: DBTL Cycle for PKS Engineering



PKS build pipeline



Machine learning models which can predict successful PKS designs will lower barriers to PKS engineering



ClusterCAD PKS Designs

Predict well-formed PKS structure

Predict PKS Activity

Top potential challenges facing the technical approach, and mitigation strategy

1. Failure of protein folding or catalytic activity for unknown reasons
   Mitigation approach: With an extremely large volume of PKSs built, some will likely work, and can begin to inform models of why they sometimes fail.
2. Synthesizing and constructing DNA for large enzymes
   Mitigation approach: A large number of diverse designs will ensure that some will be successfully constructed. Our approach is currently working for many different PKS designs.
3. Expressing the engineered PKSs *in vivo*
   Mitigation approach: We have successfully expressed PKSs in our host organism thus far, so a high diversity and volume of designs can also overcome this risk.

Key Go/No-Go decision points ensure progress

Go/No-Go #1: Verified status of ClusterCAD, throughput & speed of automatic PKS synthesis platform, & final product & targeted proteomic analyses. Designs for initial caprolactam PKS and acyl-CoA precursor pathways (M3)
Why it's important: This demonstrates our baseline capabilities, to show we are capable of starting the project as planned.

Go/No-Go #2: One round of DBTL cycle requires ≤9 mos to complete and have a throughput of 100 PKSs designed, built, tested, and learned from (100 PKSs constructed & transformed into P. putida, the engineered. P. putida grown & tested for PKS function, the data analyzed, and learnings input into design software) (M18) (UCB/LBNL, ABF)
Why it's important: This demonstrates that the speed and capacity of the PKS DBTL cycle are sufficient to constitute a powerful platform for high throughput PKS engineering.

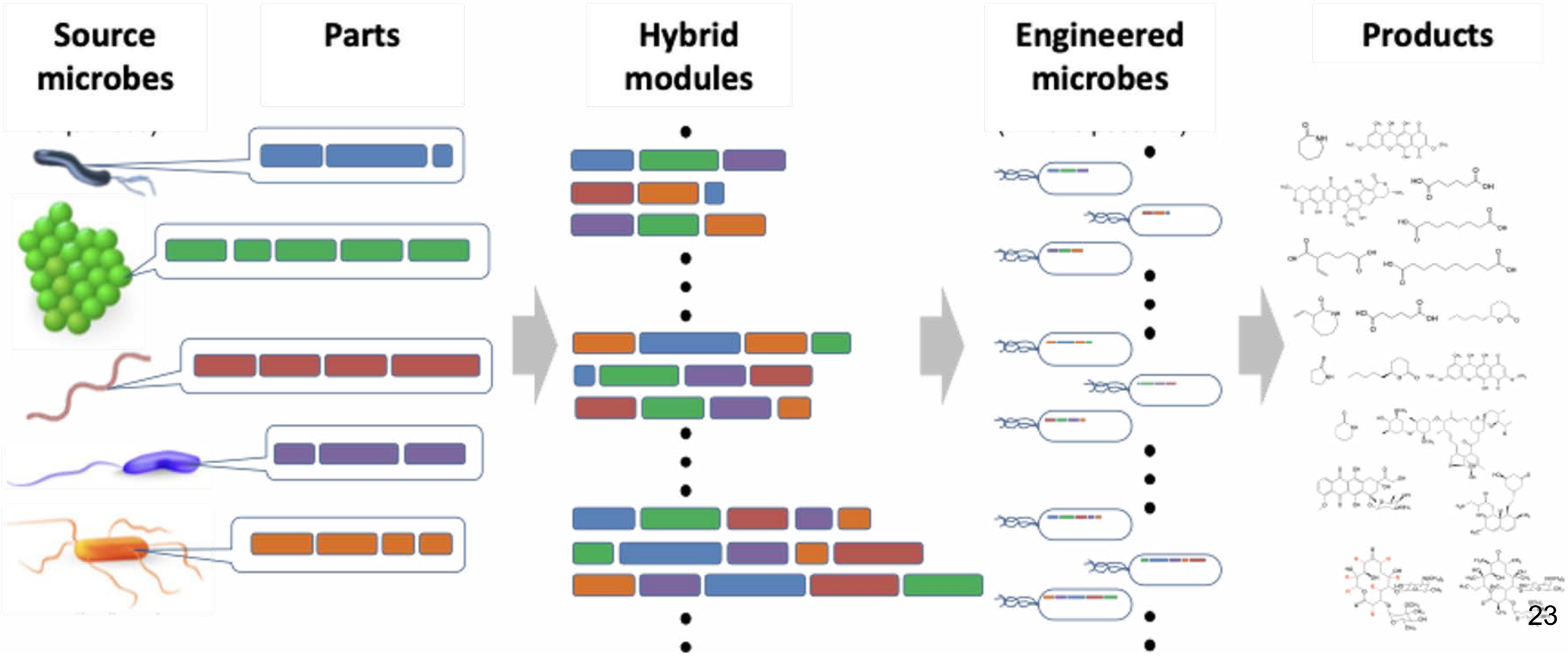
We will measure progress by rate of novel PKSs built and tested, and by production levels of the desired target molecules

Metric 1: Speed of Design-Build-Test-Learn (DBTL) cycle
Basis: For largely unkown reasons, only a small fraction of engineered PKSs function. Our metrics of increasing the speed and throughput of the DBTL cycle will allow us to improve the chances of finding functional PKSs, and building models to predict functioning PKSs.

Metric 2: Titre, rate, yield of target molecules
Basis: These are standard metrics for biological production of valuable small molecules, and help to set specific goals that can eventually make biological production of small molecules economically viable.

Impact: When successful, we will have lowered the barrier to successfully producing millions of new compounds biologically

- Sustainable and carbon neutral
- Access novel chemicals previously inaccessible



Dissemination of results

- No results yet to disseminate
- We will make our platform available for the Agile BioFoundry and others to engineer PKSs rapidly
- We will release software under open source licenses
- We will publish results in peer reviewed journals

Project goals are progressing according to the management plan

Subtask	Due	% finished
Go/No-Go #1: Verified status of ClusterCAD, throughput & speed of automatic PKS synthesis platform, & final product & targeted proteomic analyses. Designs for initial caprolactam PKS and acyl-CoA precursor pathways (M3)	09/30/20	100%
Subtask 2.1: Expand ClusterCAD to include many machine curated PKSs (M4–9)	03/31/21	20%
Milestone 2.1.1: Automatically produce chimeric PKS designs (DNA) from a target chemical structure for 10 commercial chemicals and 10 specialty chemicals (M6) (UCB/LBNL)	12/31/20	100%
Subtask 3.1: Integrate ClusterCAD with Device Editor and J5 software tools. (M4–12)	06/30/21	20%
Subtask 4.1: Construct pathways for acyl-CoAs/acyl-ACPs (M4–12)	06/30/21	10%
Subtask 5.1: Develop growth conditions and medium for PKS-engineered P. putida (M4-9	03/31/21	20%

Key technical accomplishments to date and tasks that led to them

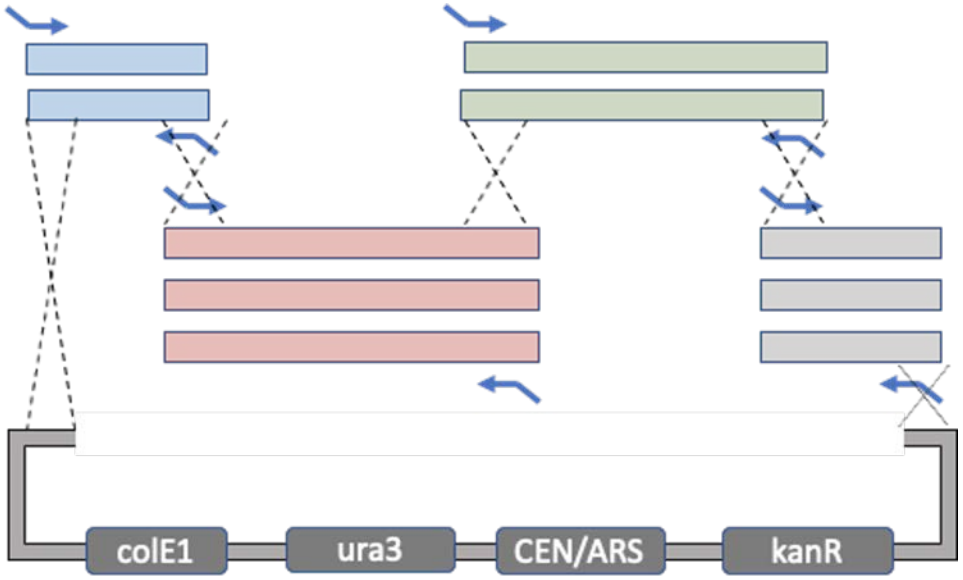
- Completed a test round of the PKS build process
  Tasks: developed automated (robotic) assembly pipeline
- Designed 384 caprolactam PKSs
  Tasks: developed software to design PKS junctions
- Designed precursor pathways for project targets
  Tasks: researched natural pathways, and obtained strains and/or synthesized DNA
- Planned the engineering of *P. putida* metabolism to support PKSs
  Tasks: identified from literature genes affecting degradation of lactams, and other key genes
- Demonstrated a full retrobiosynthesis pipeline producing DNA from chemical structures for 20 compounds
  Tasks: developed PKS design software tools

All milestones for this phase of the project are on track

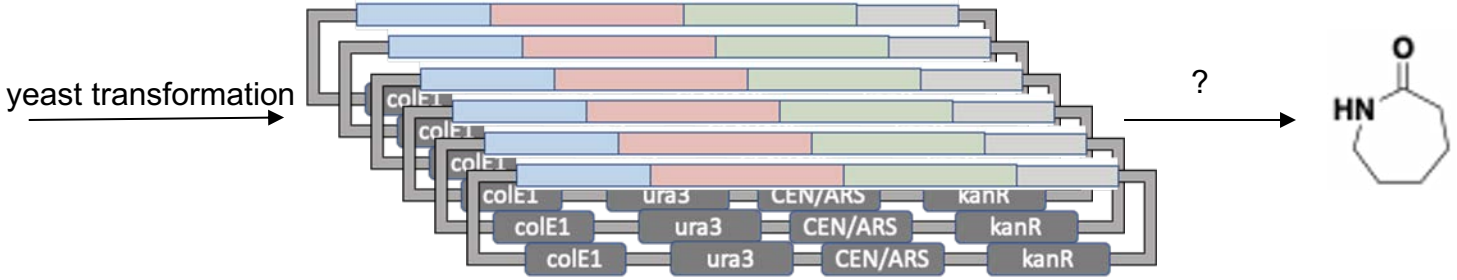
- **Milestone 2.1.1**: Automatically produce chimeric PKS designs (DNA) from a target chemical structure for 10 commercial chemicals and 10 specialty chemicals
  Due: 12/31/2020
  Status: Completed, example Python code produces DNA.
- **Milestone 3.2.1**: No fewer than 100 successfully assembled and sequenced PKS constructs made in six weeks
  Due: 3/31/2021
  Status: We have designed 384+ new PKSs and are in the process of constructing them

Results: We have completed designs and have completed PCR amplifications for assembling >100 caprolactam PKSs

Linear PCR fragments encoding putative caprolactam synthase domains



Linear PCR fragments encoding parts of putative caprolactam synthases are transformed into yeast to assemble a combinatorial library of plasmid vectors



Combinatorial library of PKS expression plasmids built by Q3

We do not know the ideal caprolactam synthase components *a priori*, so we will assemble >100 synthases and test their activities to find the best enzyme

We will construct >100 PKSs by performing combinatorial assemblies of hundreds of individual DNA parts into circular plasmid vectors that we can use to transform the caprolactam synthases into our host cell

PCRs were analyzed by agarose gel and were found to be 80-90% successful. We anticipate that we will reach our milestone of >100 unique circular constructs from the linear fragments in hand now.

Summary

Key Takeaways
We are building a reusable and flexible high-throughput DBTL cycle for PKSs.
ML algorithms + large PKS data may solve big outstanding problems in PKS engineering.
Publications, presentations, and IP
None yet
Commercialization potential
Novel caprolactam analogues may have improved biodegradability, flame resistance, or other properties.

Quad Chart Overview

Timeline

- Project start date: 2020.07.01
- Project end date: 2023.06.30

	FY20 Costed	Total Award
DOE Funding	(10/01/2019 – 9/30/2020)	1,189,999.00 (negotiated total federal share)
Project Cost Share	1,935,000.00	

Project Partners

- University of California, Berkeley
- National Labs: LBNL, NREL, PNNL, ANL

Project Goal

The goal of the proposed work is to develop a rapid, high throughput, Design-Build-Test-Learn (DBTL) cycle for polyketide synthases (PKSs) and demonstrate its utility for production of materials precursors.

End of Project Milestone

- Production of caprolactam at ≥ 5 g/L, 25% of theoretical yield, & 0.1 g/l/hr from cellulosic biomass
- Production of 2 caprolactam derivatives at ≥ 0.5 g/L, 2.5% theoretical yield, & 0.01 g/l/hr from cellulosic biomass
- One round of DBTL cycle requires ≤2 mos to complete and have a throughput of 500 PKSs designed, built, tested, and learned from per cycle (500 PKSs constructed & transformed into P. putida, the engineered. P. putida grown & tested for PKS function, the data analyzed, and learnings input into design software)

Funding Mechanism

FY19 FOA DE-FOA-0002029

Topic area: AOI 7- Advanced Bioprocessing and Agile BioFoundry

Additional Slides

Responses to Previous Reviewers' Comments

- No previous reviewers' comments (new project)

Publications, Patents, Presentations, Awards, and Commercialization

- None (new project)