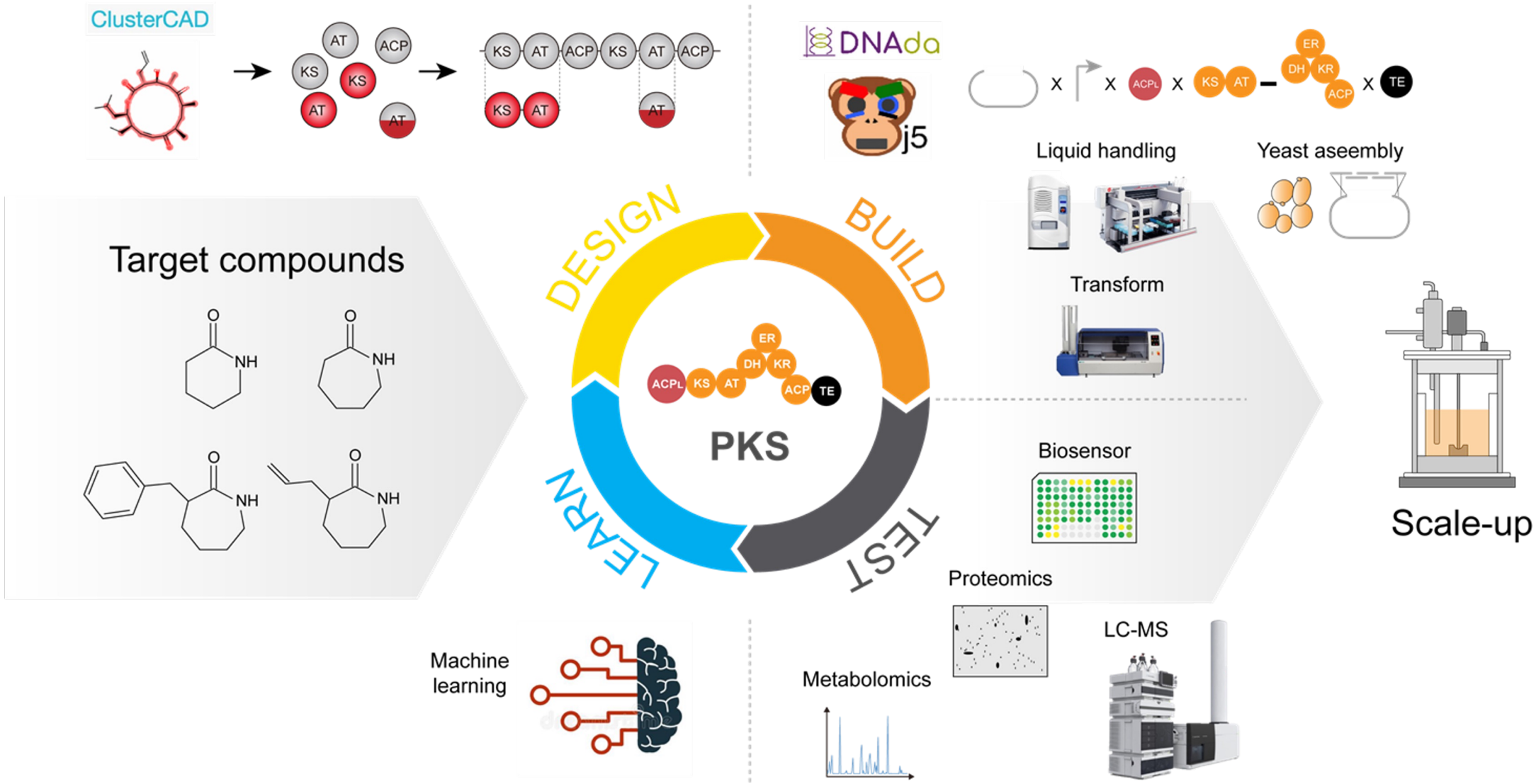
DOE Bioenergy Technologies Office (BETO)
2023 Project Peer Review

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

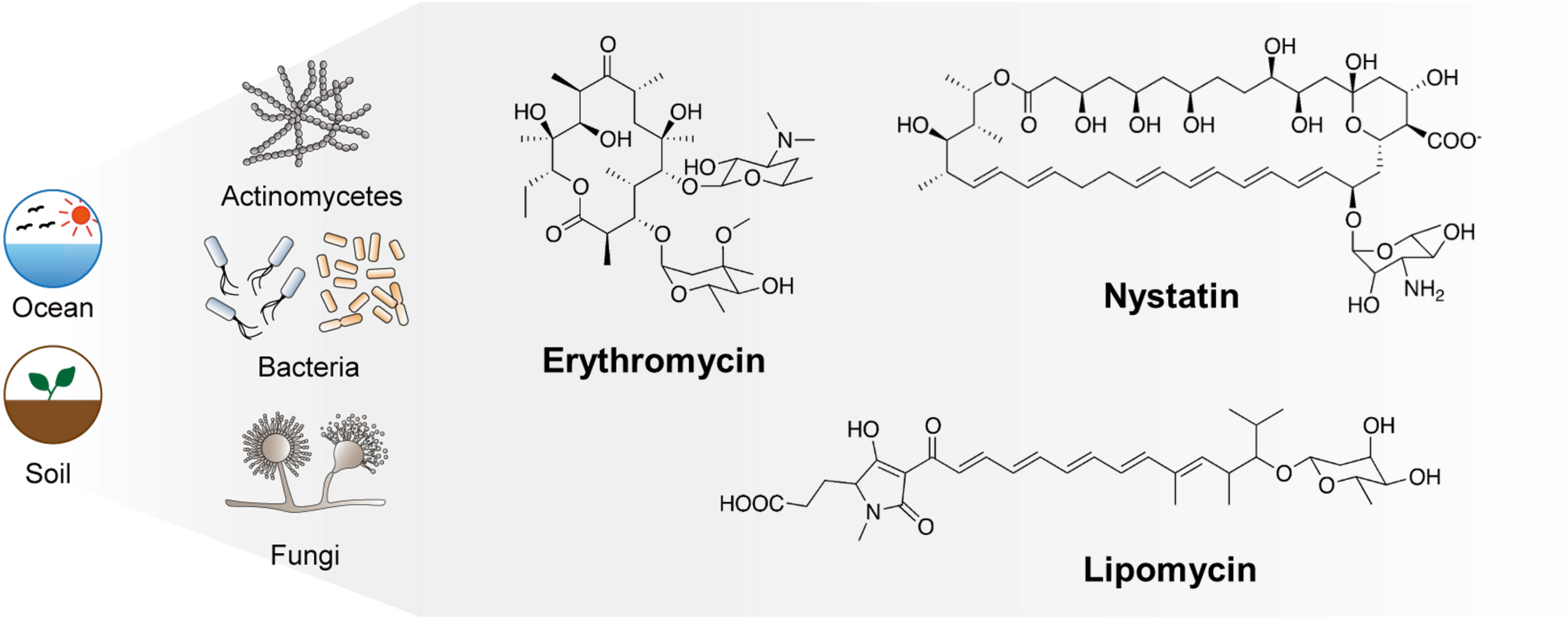
4/4/2023
Agile BioFoundry and Advanced Biofuels
Process Development Unit session

Jay Keasling
University of California, Berkeley

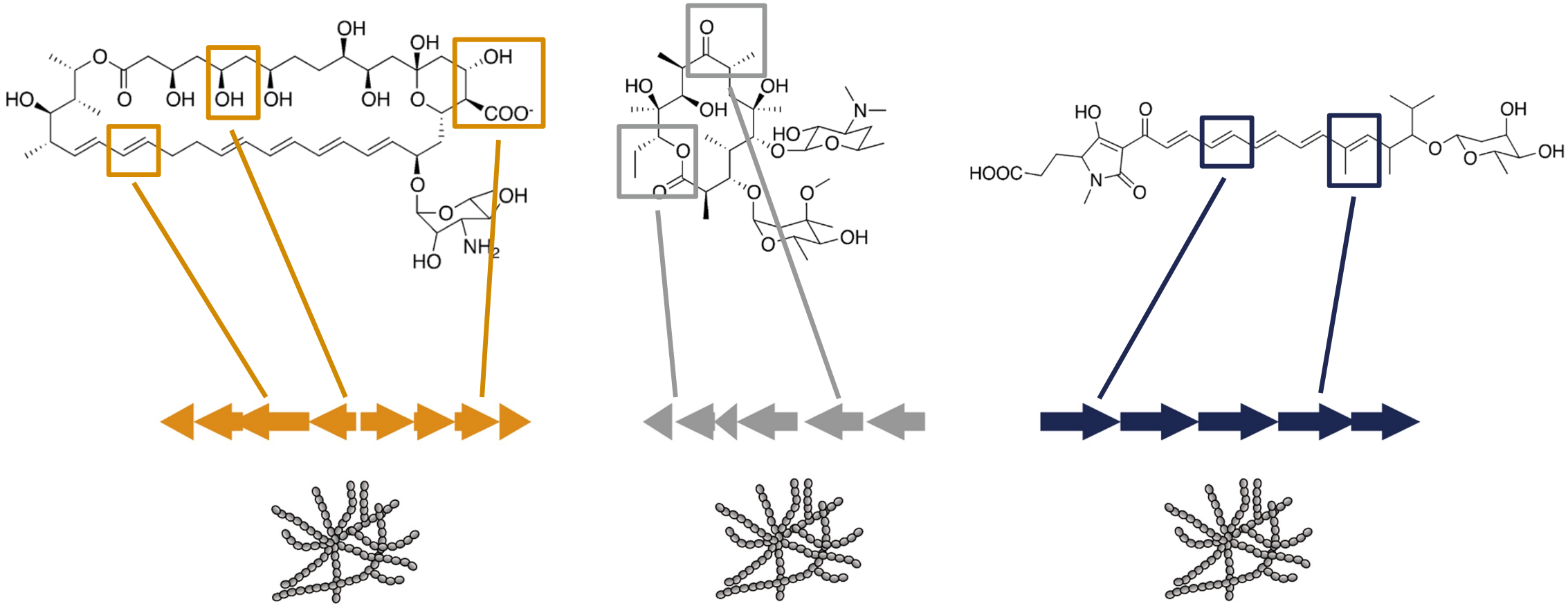
Project Overview



What is a polyketide?

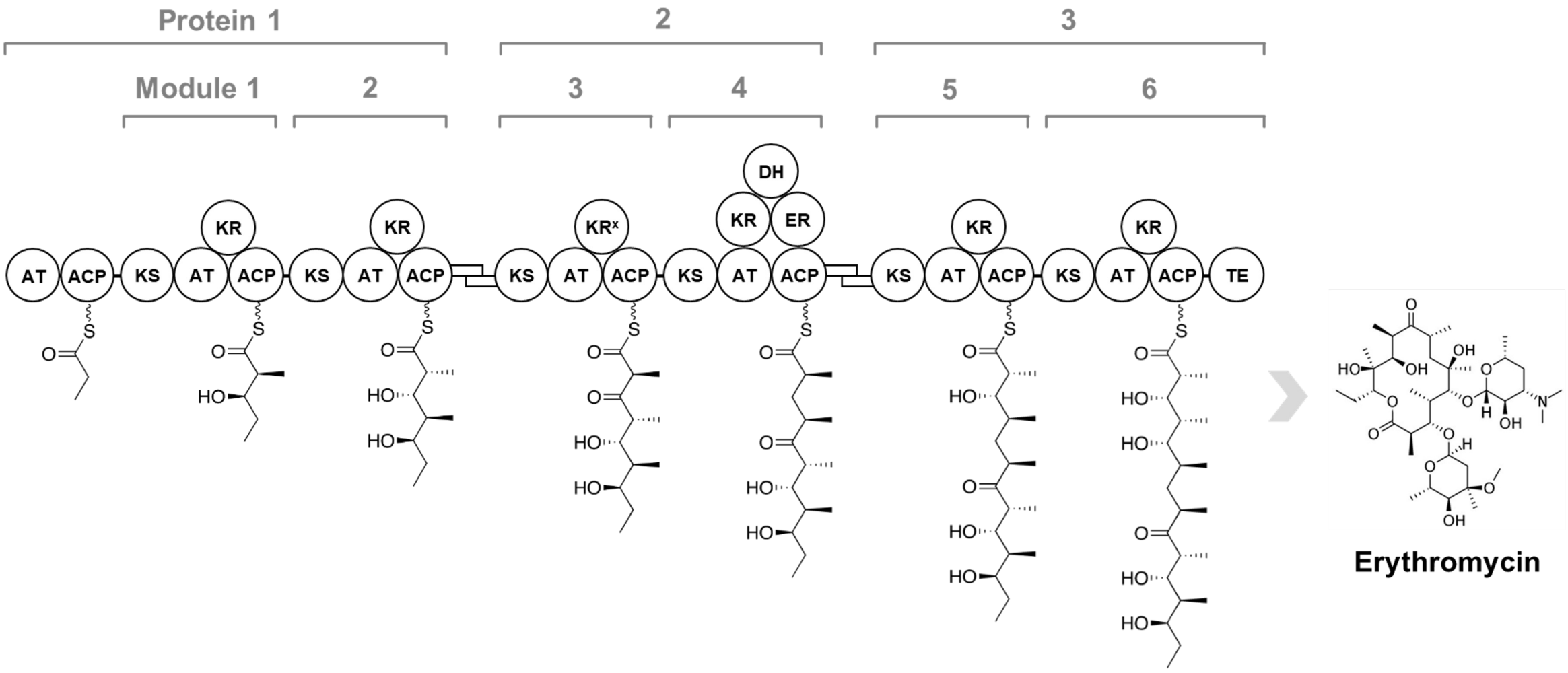


Polyketides are produced using modular enzymes

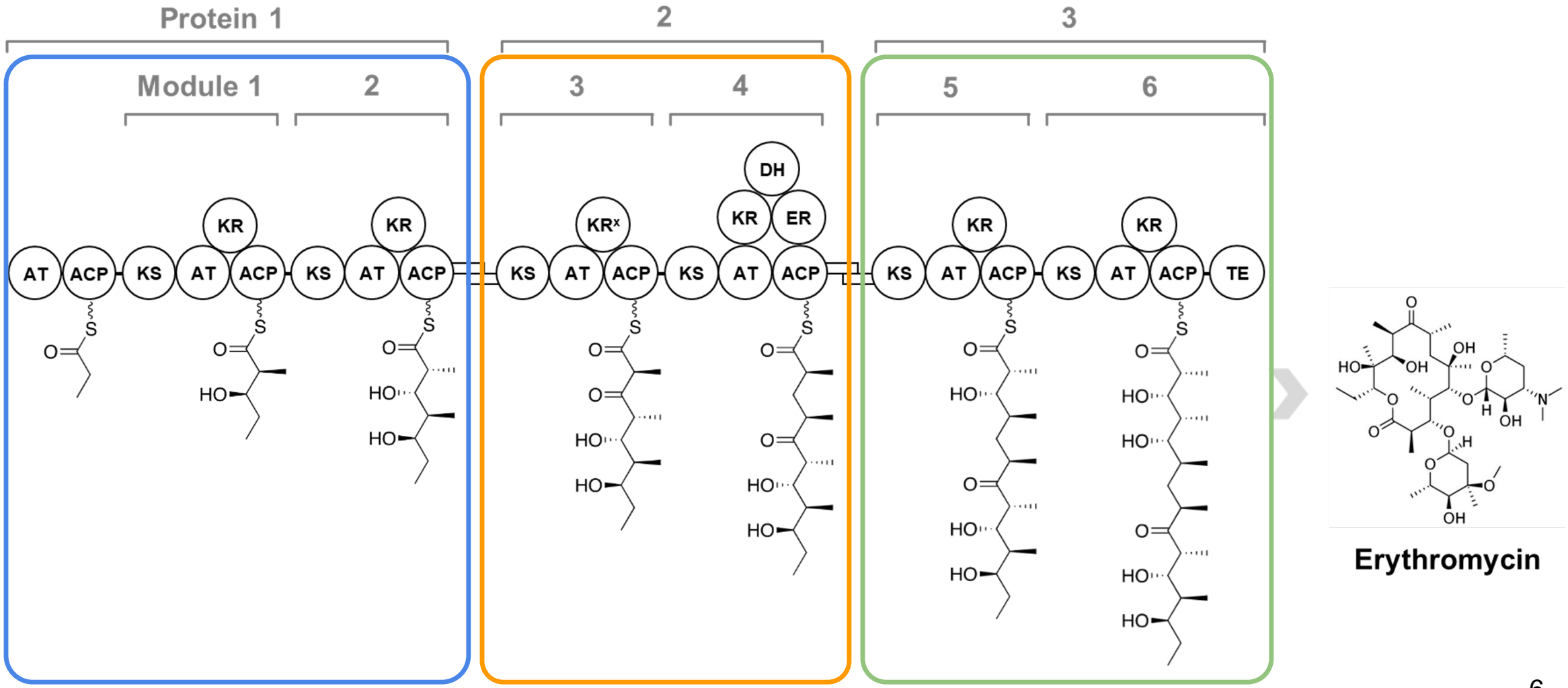


Modular parts - Each module is responsible for some part of the complex chemical

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

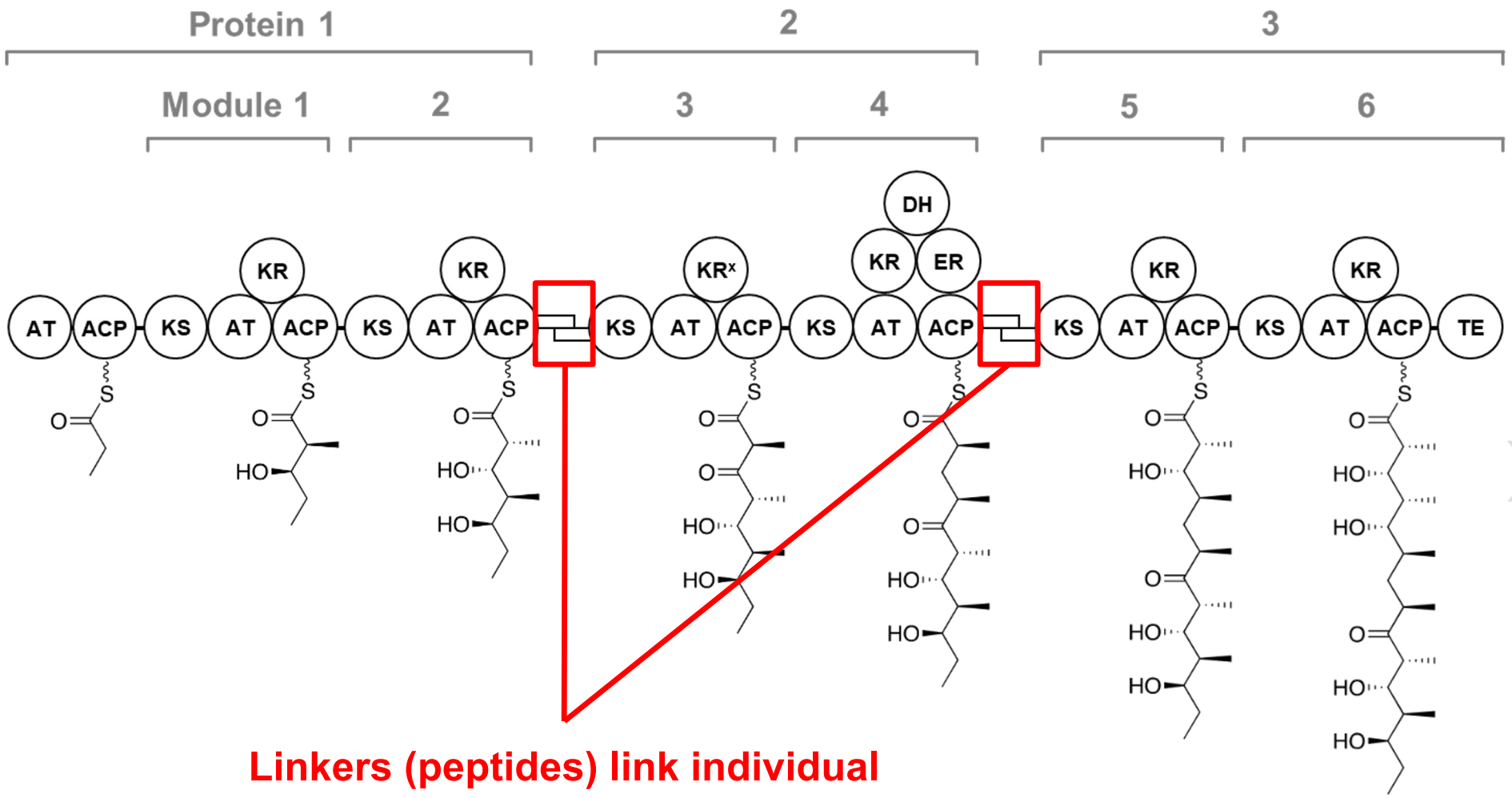


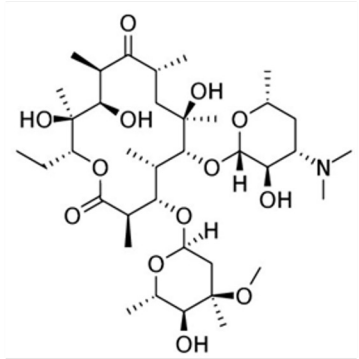
Polyketide synthases (PKSs) are large, multi-activity proteins



Erythromycin

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

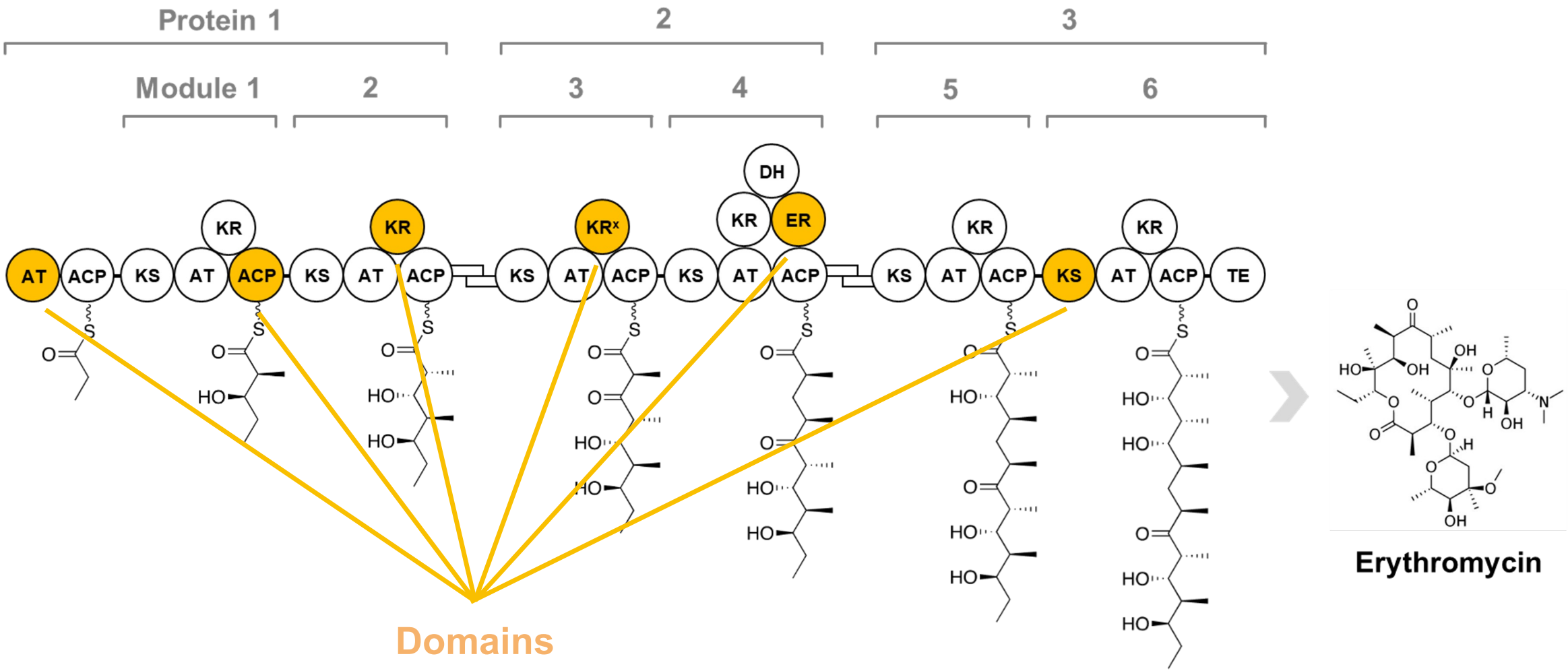




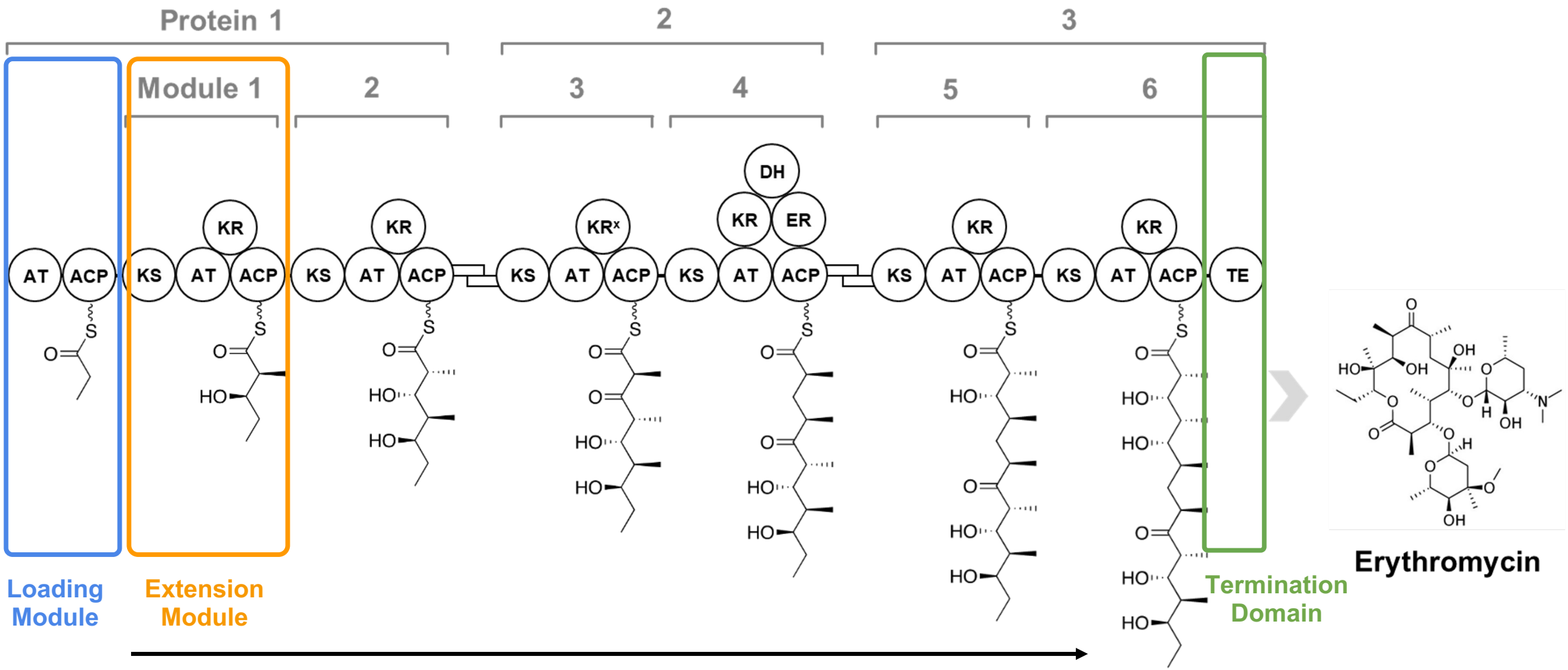
Erythromycin

Linkers (peptides) link individual proteins together

Each domain in a PKS has an individual enzyme activity



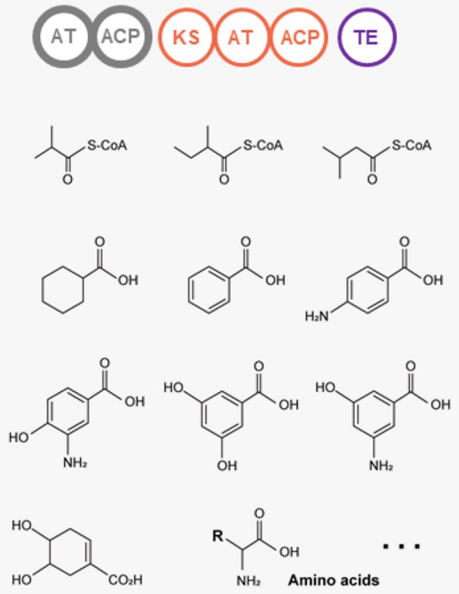
Polyketide synthases generally have Load and Extension Modules and some type of Termination



Enormous potential of PKS based retrobiosynthesis

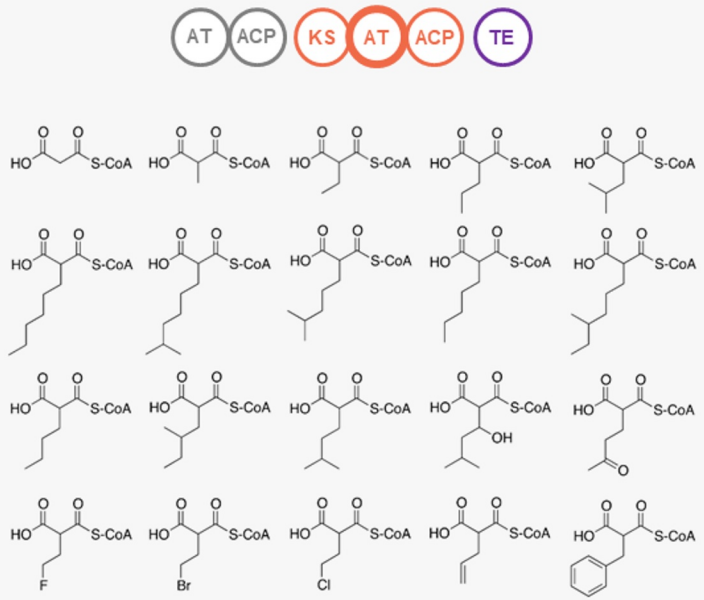
Hybrid unnatural PKS could synthesize millions of different molecules

Choose starter unit



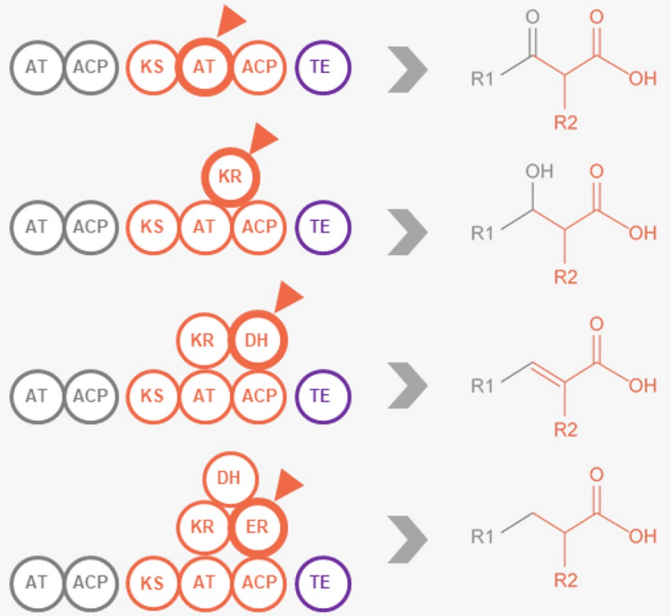
Wide variety of starter units

Choose extender unit



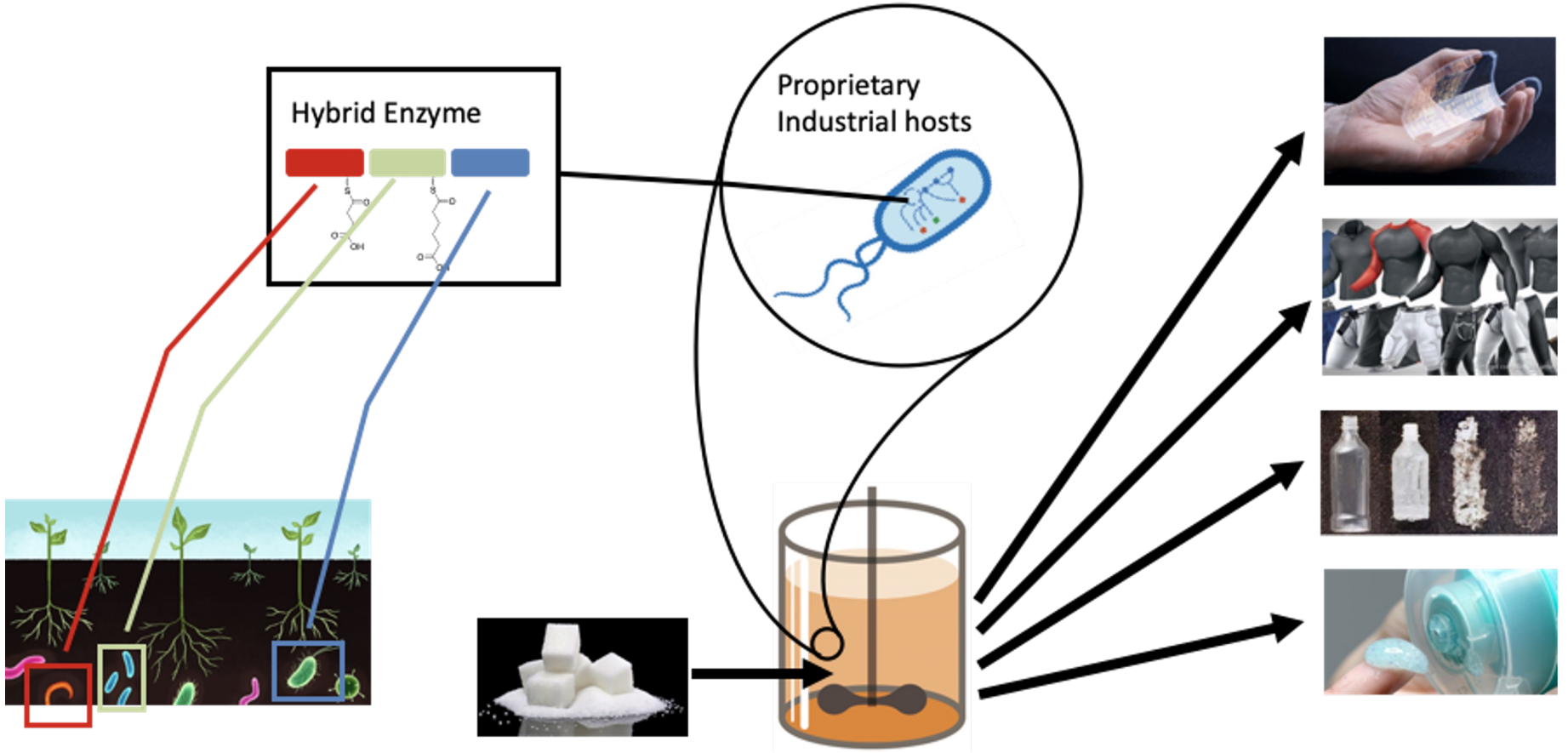
~20 malonyl-CoA derivatives

Choose reduction level



Four levels

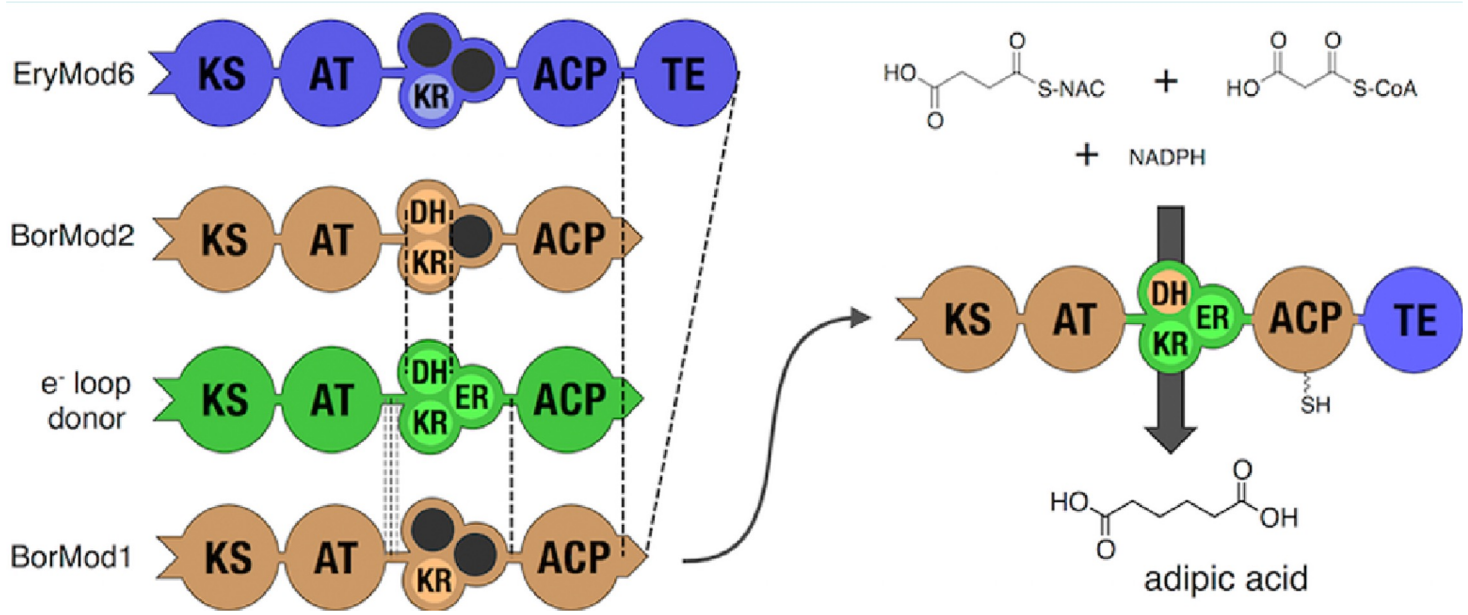
What we are trying to do: engineer PKSs for production of commodity and specialty chemicals



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 monomers caprolactam, valerolactam, and novel 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.

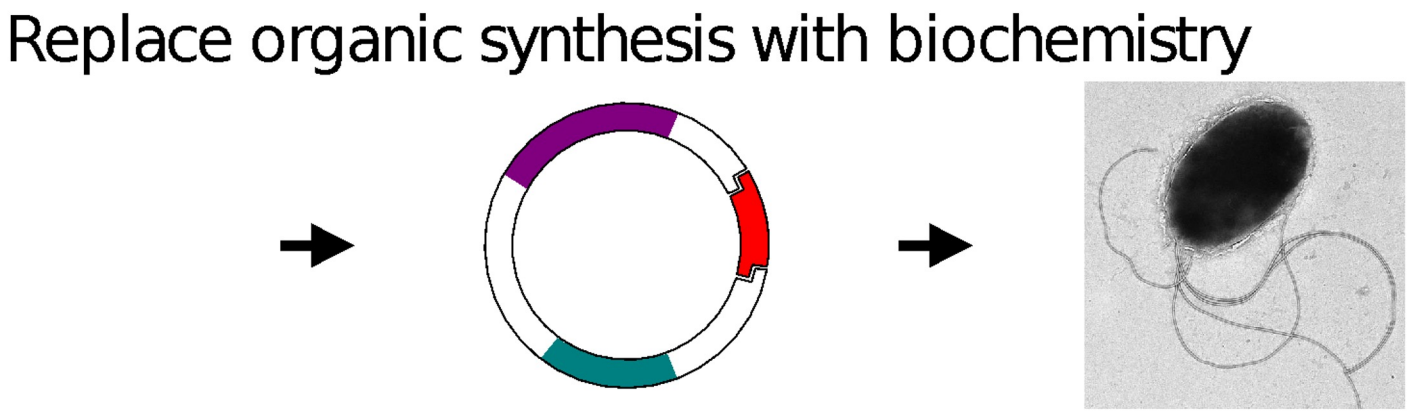


A novel PKS producing adipic acid built via extensive trial and error

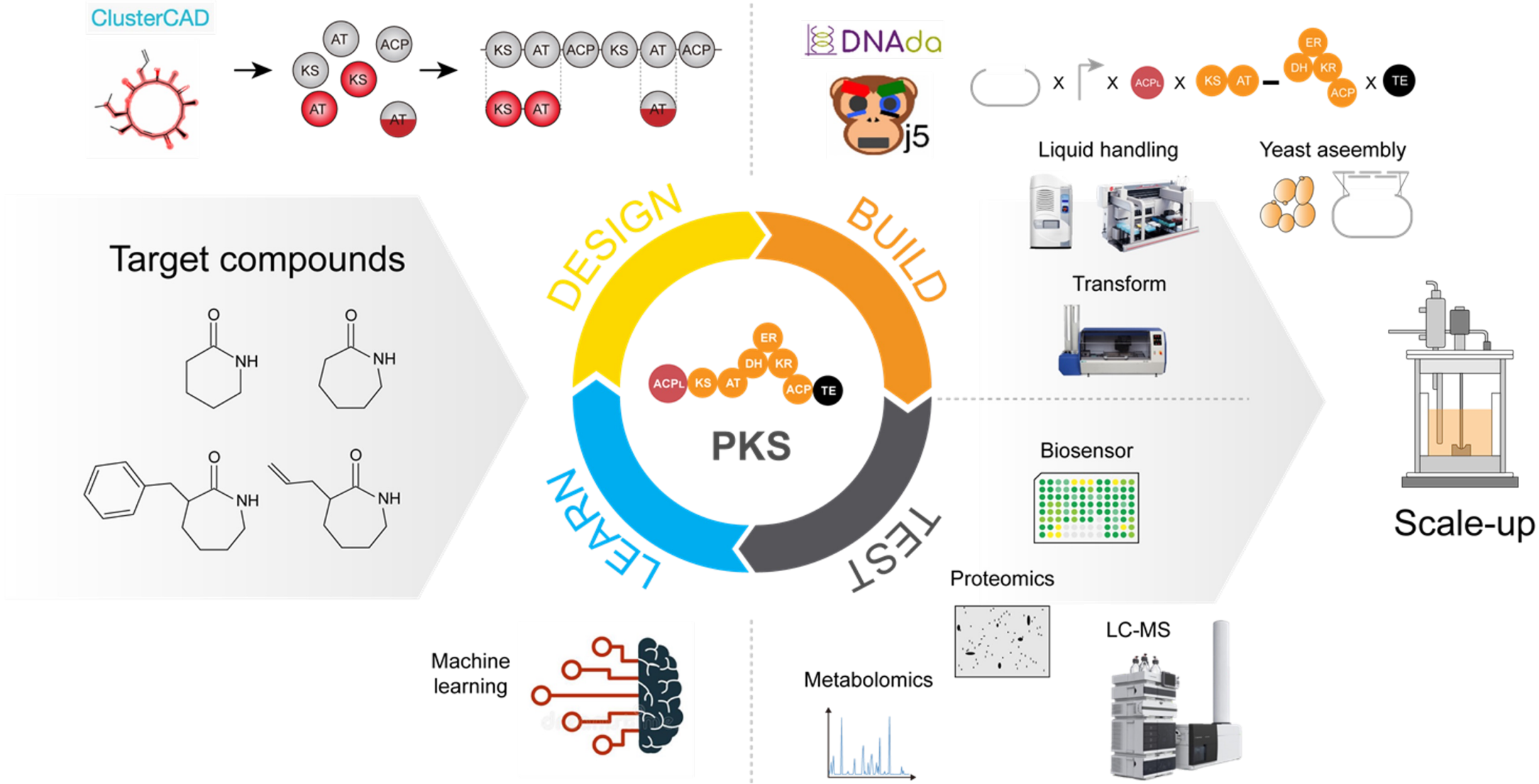
(Hagen *et al.* ACS Synthetic Biology, 2015)

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
- This project is making an array of valuable products including several novel plastic (nylon) monomers with potentially better properties
- We have engineered the microbe *P. putida* to produce PKSs, which enables faster growth and more bioengineering potential vs the natural microbes they occur in

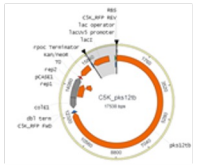


Approach: DBTL Cycle for PKS Engineering



PKS build pipeline

1 Choose genetic parts using Inventory of Composable Elements (**ICE**) and **ClusterCAD**



JBEI Joint BioEnergy Institute

ClusterCAD

2 Design primers for assembly fragments preparation using **J5**
Create files for downstream automation using **DNAda**

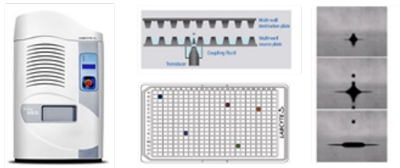
j5

DNAda

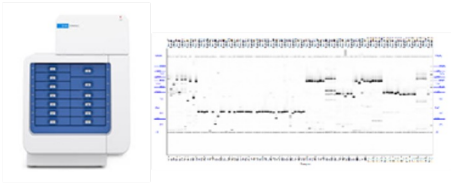
3 Order primers and synthetic templates using automatically generated sheet from **J5**

IDT INTEGRATED DNA TECHNOLOGIES

4 Mix primers and template for assembly fragment preparation PCR using **Echo555**

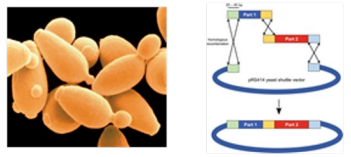


5 QC of PCR products using Zero Agarose Gel (**ZAG**) capillary electrophoresis

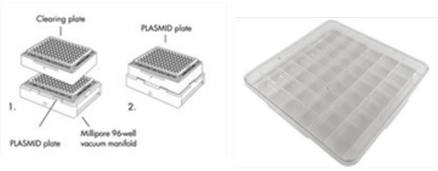


6 Purify the PCR product and mix the PCR product for yeast assembly using **Echo555**

7 Transform *S. cerevisiae* with linear fragments for yeast assisted homologous recombination



8 Purify circular plasmid DNA from yeast, transform into *E. coli,* and plate on Qtray



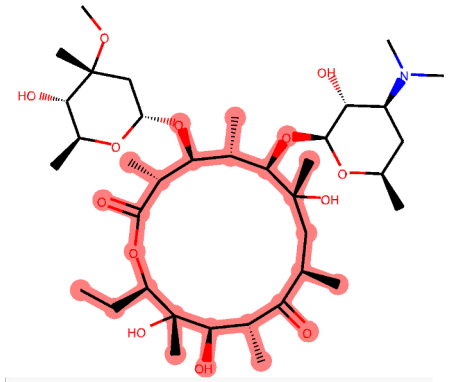
9 Pick and inoculate colonies into 96 deep well plates using Robotics (**Qpix**)

10 Boil prep the culture, do the diagnostic PCR for sample selection, and submit for NGS based sequence check

11 DIVA team performs NGS and alignments with target sequences to identify insertions/deletions/SNPs

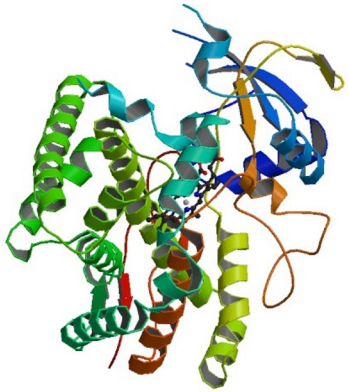
12 Samples free of mutations in PKS coding sequence are retrieved from storage and submitted to ICE

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



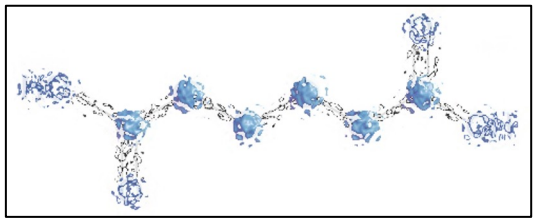
ClusterCAD PKS Designs

AI - 1



Predict well-formed PKS structure

AI - 2



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.

1. 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.

1. 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: Throughput of Design-Build-Test-Learn (DBTL) cycle
Basis: For largely unknown 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.

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

Approach to diversity, equity, and inclusion

- Mentoring young scientists from diverse backgrounds is central to this project
  - A large number of undergraduate and graduate students were introduced to research through this project
  - Let young scientists own a sub-project: The first author of our first publication from this project is a UCB undergraduate student
- Reimagined hiring and student recruitment to be more inclusive
  - Position requirements emphasize skills rather than credentials
  - Lower pressure interview process, discussing past projects vs high pressure whiteboard interviews
- Cultivate a work culture that is inclusive
  - Social and networking events designed to include people with different cultural or religious constraints, and family responsibilities
  - Manage meetings and discussions so everyone has a voice/input
  - Identify discrimination and related problems by soliciting feedback and take corrective action

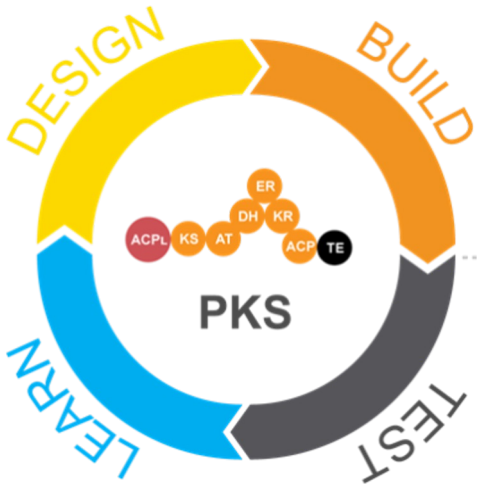
Program design

We are on Tasks 8-9

Task 1 | Tasks 2-6 | Task 7 | Tasks 8-9 | Task 10

Development of PKS automation pipeline

Demonstration of pipeline capabilities for lactam nylon precursors and derivatives



Key technical accomplishments to date and tasks that led to them

- Demonstrated a complete round of the PKS DBTL cycle in under 9 months, making over 100 PKSs per cycle
  Tasks: developed automated (robotic) assembly pipeline, design software, and machine learning algorithms

- Designed and built hundreds of lactam nylon precursors PKSs
  Tasks: developed software to design PKSs automatically, and fed into build pipeline

- Built precursor pathways for project targets
  Tasks: researched natural pathways, and obtained strains and/or synthesized DNA

- Modified *P. putida* metabolism to support PKSs
  Tasks: identified from literature genes affecting degradation of lactams, and other key genes

- Produced four nylon precursor lactam molecules with PKSs
  Tasks: Iterating our DBTL cycle successfully

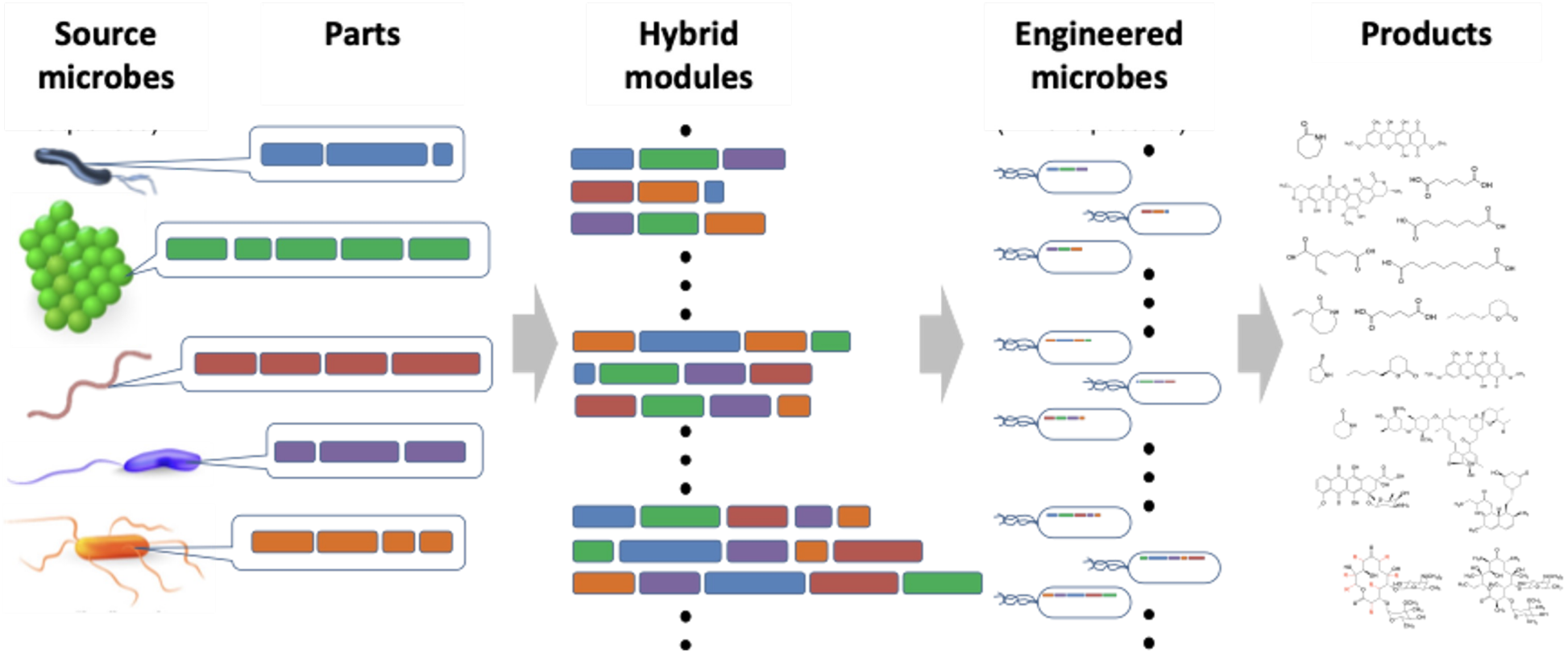
All milestones for previous phases were completed successfully

Subtask	% finished
Milestone 2.1.1: Automatically produce chimeric PKS designs (DNA) from a target chemical structure for 10 commercial chemicals and 10 specialty chemicals	100%
Milestone 3.2.1: No fewer than 100 successfully assembled and sequenced PKS constructs made in six weeks	100%
Milestone 4.1.1: Demonstrate production of >10 mg/L allylmalonyl-CoA/ACP in an engineered *P. putida* via direct detection of allylmalonate	100%
Milestone 5.3.1: High throughput P-pant assay developed: capable of analyzing ≥ 1000 *in vivo* PKS samples in < 10 days	100%
Milestone 6.2.1: Using machine learning, rank 10 potential PKS designs for each of 10 commodity chemicals, and 10 specialty chemicals (200 designs total) by probability of producing detectable levels of the desired target molecule	100%

All milestones for this phase of the project are on track

Subtask	% finished
Milestone 8.1.1: Design 190 variants of caprolactam, valerolactam, or derivatives thereof synthase	100%
Milestone 8.2.1: Build and sequence verify 90 variants of caprolactam, valerolactam, or derivatives thereof synthase (of the 190 designed in Milestone 8.1.1)	50%
Milestone 8.3.1: Production of caprolactam, valerolactam, or derivatives thereof at ≥ 0.5 g/L, 2.5% of theoretical yield, & 0.01 g/l/h	80%
Milestone 8.5.1: P-pant assay performed on strains that produced no caprolactam, valerolactam, or derivatives thereof	50%
Milestone 8.7.1: All successful/failed caprolactam, valerolactam, or derivatives thereof synthase designs analyzed in Leave One Out Cross Validation approach to finalize AI modules. Use for design of PKSs for additional molecular classes	85%
Milestone 9.4.1: Confirm production of 2 caprolactam or valerolactam derivatives at ≥ 0.05 g/L, 0.25% theoretical yield, & 0.001 g/l/h from cellulosic biomass	90%

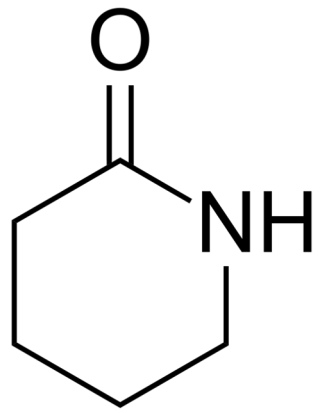
Impact: We have lowered the barrier to successfully producing millions of new compounds biologically



- Sustainable and carbon neutral
- Access novel chemicals previously inaccessible

Impact: We have successfully made a number of valuable plastic precursors, potentially with improved properties

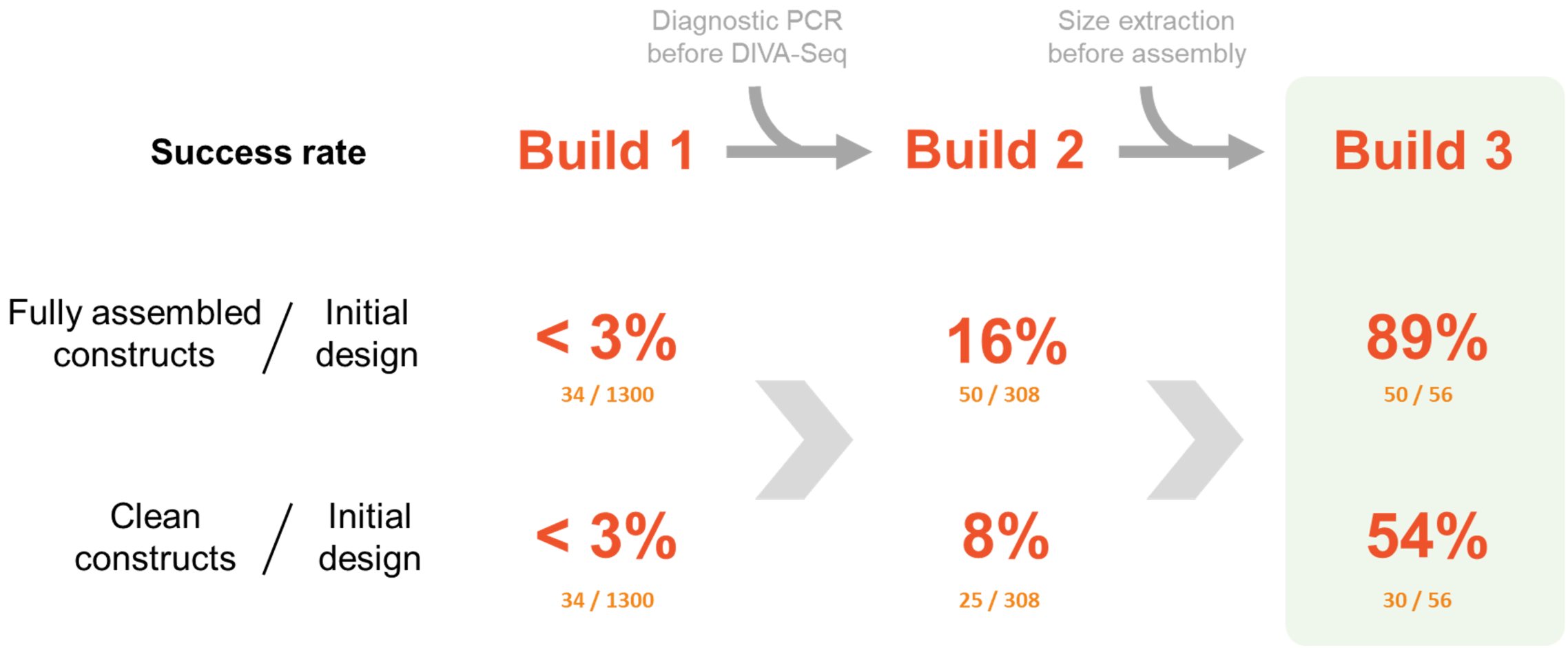
- We have confirmed production of valerolactam and three additional novel nylon precursors with PKSs with titers from 1 to 10 mg/L.
- We have demonstrated that our build pipeline can build hundreds of PKSs at once in high throughput.



Valerolactam

(Nylon figure from technologystudent.com)

Impact: Improved success ratio of assembly pipeline



Dissemination of results

- We have published one peer reviewed paper, and will prepare remaining results in additional peer reviewed papers
- We will make our platform available for the Agile BioFoundry and others to engineer PKSs rapidly
- We have released our software under open source licenses, for public use
- We intend to apply for a patent for production of novel molecules and materials we have now successfully generated

Summary

- We have built and demonstrated a reusable and flexible high-throughput DBTL cycle for PKSs that can build hundreds of PKSs in a short time period.
- We have developed a computational pipeline for high throughput PKS design.
- We have successfully made four valuable plastic precursors, including novel compounds that may have improved properties over existing plastics.

Quad Chart Overview

Timeline

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

	FY22 Costed	Total Award
DOE Funding	*(10/01/2021 – 9/30/2022)*	1,189,999.00 *(negotiated total federal share)*
Project Cost Share *	1,935,000.0	

TRL at Project Start: 2
TRL at Project End: 4

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 or valerolactam at ≥ 5 g/L, 25% of theoretical yield, & 0.1 g/l/hr from cellulosic biomass
- Production of 2 caprolactam or valerolactam 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

Funding Mechanism

FY19 FOA DE-FOA-0002029

Topic area: AOI 7- Advanced Bioprocessing and Agile BioFoundry

Project Partners

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

Additional Slides

Responses to Previous Reviewers' Comments from 2021 Peer Review

- **Comment:** [This project] incurs some risk because PKS can be difficult to find and to express in functional forms. In large part, the team is addressing these risks by brute-force, buy building 100's of candidate systems. Unfortunately, they report that so far, only a very small fraction of the engineered PKSs function, so this approach may be insufficient.

- **Response:** Indeed, in our initial design rounds, we found no active PKSs from among hundreds of different designs. We developed a new iterative design approach, which starts with natural PKSs, and verifies changes one at a time, slowly building to the desired design. We found this offered greater success, and led to a large number of successful designs.

Highlights from 2022 Intermediate Verification and Go/No-Go

- All milestones and objectives were achieved, the verification team recommends continuation of the project
- The verification team recommends the parallel addition of valerolactam as a product (and potential derivatives) to the project scope
- The verification team recommends continuing to actively utilize lessons learned in Task 3.0 to increase efficiency of PKS expression construct builds
- The verification team recommends that the team introduce more variety into the PKS constructs
- The verification team also made other technical recommendations, which subsequently proved useful to the project

Publications, Patents, Presentations, Awards, and Commercialization

Publications:

- Tao, XB, LaFrance, S, Xing, Y, Nava, AA, Martin, HG, Keasling, JD, Backman, TWH **(2023)**. ClusterCAD 2.0: an updated computational platform for chimeric type I polyketide synthase and nonribosomal peptide synthetase design. *Nucleic Acids Res*, 51, D1:D532-D538.