Communicating Research to the General Public

The WISL Award for Communicating PhD Research to the Public launched in 2010, and since then over 100 Ph.D. degree recipients have successfully included a chapter in their Ph.D. thesis communicating their research to non-specialists. The goal is to explain the candidate's scholarly research and its significance—as well as their excitement for and journey through their area of study—to a wider audience that includes family members, friends, civic groups, newspaper reporters, program officers at appropriate funding agencies, state legislators, and members of the U.S. Congress.

WISL encourages the inclusion of such chapters in all Ph.D. theses everywhere, through the cooperation of PhD candidates, their mentors, and departments. WISL offers awards of $250 for UW-Madison Ph.D. candidates in science and engineering. Candidates from other institutions may participate, but are not eligible for the cash award. WISL strongly encourages other institutions to launch similar programs.

Wisconsin Initiative for Science Literacy

The dual mission of the Wisconsin Initiative for Science Literacy is to promote literacy in science, mathematics and technology among the general public and to attract future generations to careers in research, teaching and public service.

Contact: Prof. Bassam Z. Shakhashiri

UW-Madison Department of Chemistry

bassam@chem.wisc.edu

www.scifun.org

Characterization and Engineering of a Promiscuous L-Threonine Transaldolase to Access Novel Amino Acid Building Blocks

By

Prasanth Kumar

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biophysics)

at the

University of Wisconsin-Madison

2022

Date of final oral examination: 9th February 2022

The dissertation is approved by the following members of the Final Oral Committee:

Andrew R. Buller, Professor, Chemistry

Amy M. Weeks, Professor, Biochemistry

Philip A. Romero, Professor, Biochemistry

Srivatsan Raman, Professor, Biochemistry

Zachary K. Wickens, Professor, Chemistry

Chapter 5

How I used vitamins and proteins to synthesize new amino acids:

A chapter for the non-scientist

I believe that everyone will be able to enjoy and appreciate the fascinating world of chemical biology if the ideas are communicated in a simple and relatable language. This is why I wrote this chapter to describe the broader context and findings from my Ph.D. research to a wide, non-scientist audience. I want to thank the team at the Wisconsin Initiative for Science Literacy Program at the University of Wisconsin-Madison for their encouragement and support in the creation of this thesis chapter.

Chapter 5: How I used vitamins and proteins to synthesize new amino acids: A chapter for the non-scientist

5. 1. What did I do in my Ph.D.?

I am sure you have seen cereals at some point in your life, maybe even had them for breakfast this morning. If you read the nutrition facts on the side of the cereal box, you might have noticed they contain proteins and vitamins (see Figure 1). Today, I will tell a story of how we (my collaborators and I) used proteins and vitamins to build (synthesize) new amino acids. The first question that might come to your mind is: What are amino acids? Let's start with this question.

Figure 1. A picture of a multi-grain Cheerios cereal box. The red circle highlights the proteins and vitamin content.

5. 2. What are amino acids?

To build a sturdy house, we often use bricks as the building block. Similarly, cells in our bodies use 20 different amino acids as the building block to synthesize proteins (see Figure 2). When we eat foods that contain proteins, the protein gets broken down into individual amino acids, which are eventually used to build new proteins (similar to recycling plastic or paper). Our bodies can't make all 20 standard amino acids, so we need proteins in our diet.[1]

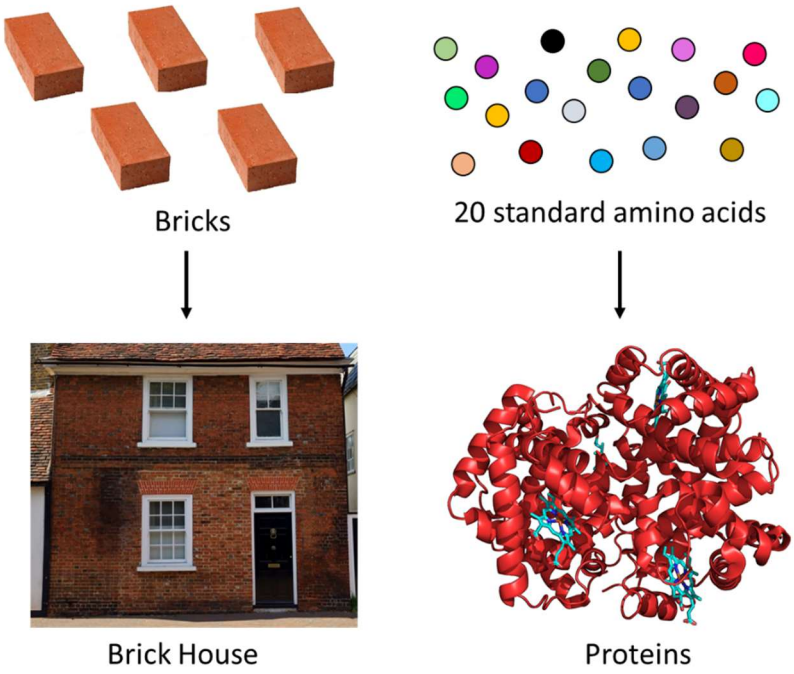


Figure 2. The emergence of complexity from simple building blocks around us and in Nature. The bottom right is a simplified representation of the Hemoglobin protein structure. The picture of the concrete home was downloaded with permission from pxhere.com.

The next question is: What are proteins? Proteins are the molecular machines in our body that carry out almost all bodily functions. For example, Hemoglobin is the protein that carries oxygen to our lungs, and Insulin is the protein that regulates sugar levels in the blood. Structurally, proteins are polymers (a molecule composed of many simple repeating units) of the 20 standard

amino acids, like beads on a necklace (see Figure 3). These amino acids differ in their side chain or R group, which brings unique properties to them.

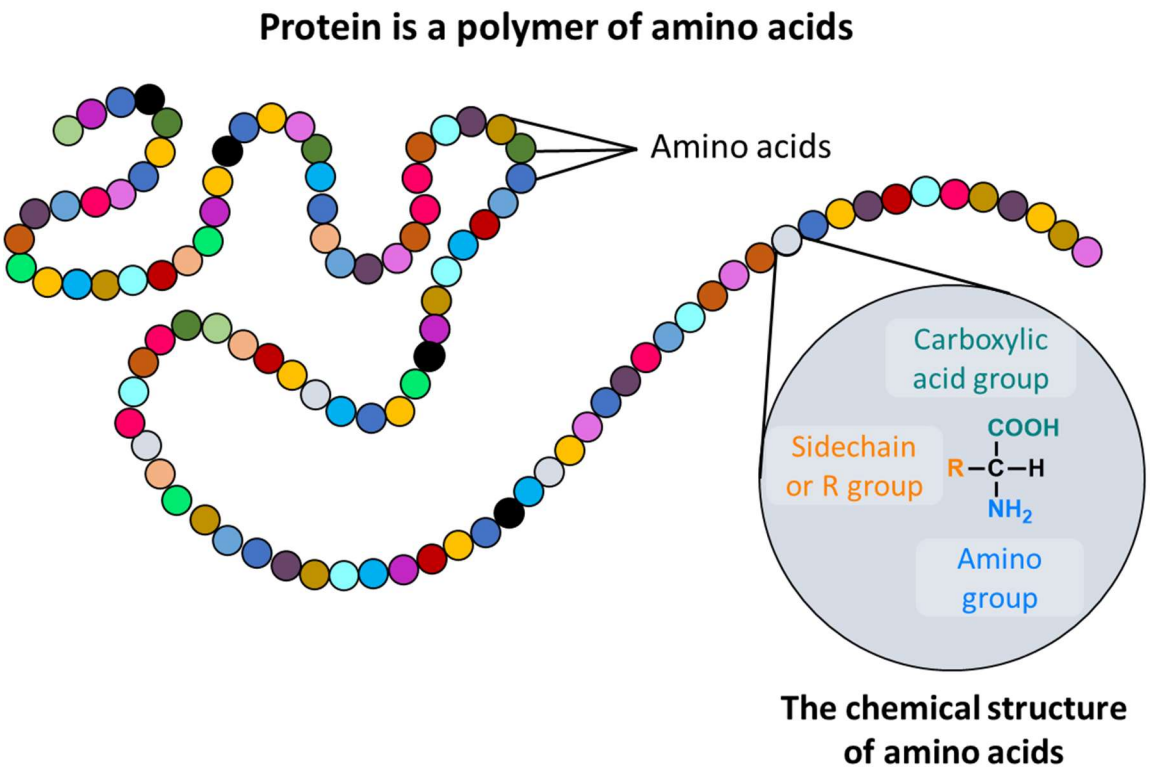


Figure 3. Primary structure of proteins. Proteins are polymers of the 20 standard amino acids. The general chemical structure of amino acids is shown in the grey circle.

5. 3. Why do we need new amino acids?

The next question that might come to your mind is: If cells only need 20 amino acids to make proteins, why do we need *new* amino acids? That's an excellent question! The answer is cells need more than 20 amino acids to build several biologically valuable compounds (see Figure 4).[2] For example, the first isolated antibiotic, Penicillin, is built from three different amino acids. Out of the three, two of the amino acids are not found in proteins. These amino acids are called non-standard or non-proteogenic amino acids. The hormone thyroxine (produced by the thyroid gland) and neurotransmitter serotonin are also built from such non-standard amino acids.

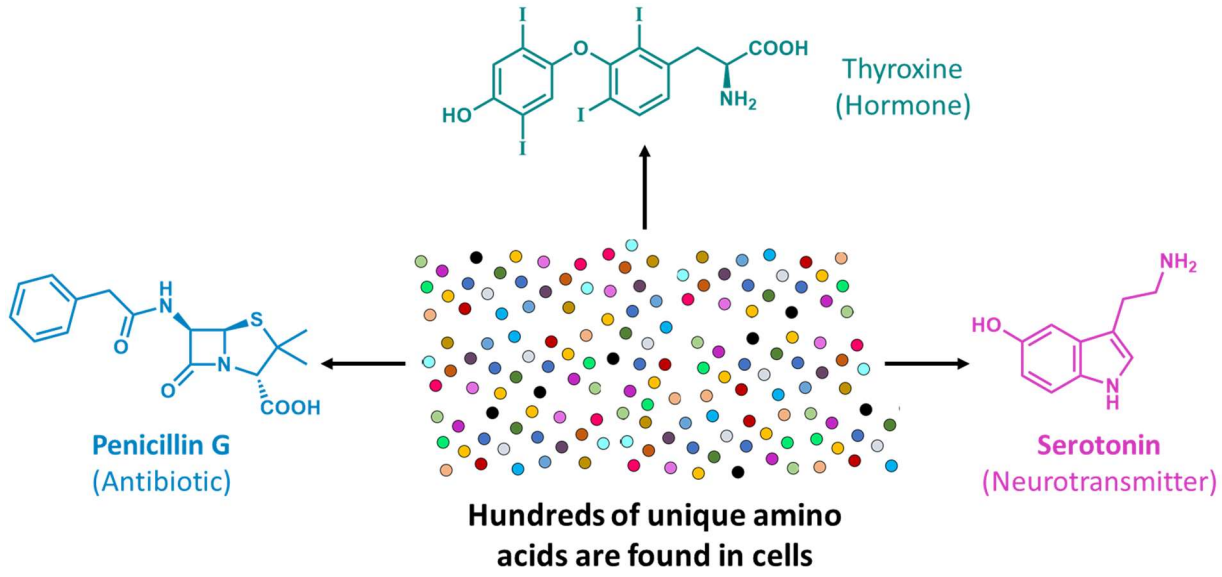


Figure 4. Several biologically valuable compounds are built from non-standard amino acids. Penicillin, thyroxine, and serotonin are shown in a format that is commonly used by chemists to depict chemical structures.

Inspired by the biological properties of non-standard amino acids, medicinal chemists (a chemist who develops and tests new drugs) have used non-standard amino acids to build several therapeutic drugs. These include drugs such as droxidopa (used in the treatment of Parkinson's disease) and sitagliptin (antidiabetic drug), which have benefited millions of lives. With the

emergence of new diseases and antibiotic resistance (where bacteria develop the ability to defeat the drugs designed to kill them), it is vital to build new drugs and improve existing drugs.

Similar to how distinct Lego building blocks help us build complex structures, access to new amino acids will help us create new molecules, which could be evaluated as therapeutics (see Figure 5). This is the reason we need new amino acids, which can be further developed into new therapeutics. So, how do we make them?

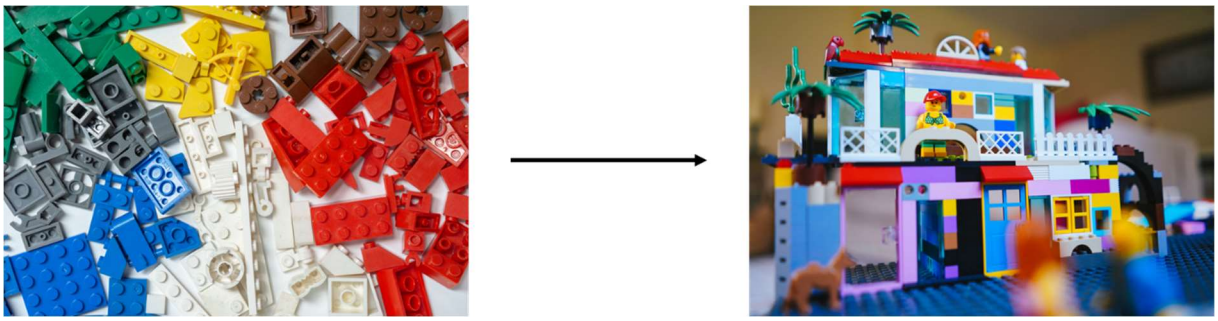


Figure 5. Building complex structures from simple LEGO building blocks. These pictures were downloaded with permission from unsplash.com.

There are several chemical methods to make amino acids. These methods usually require expensive starting materials (substrates), toxic solvents (a liquid used to dissolve substrates), and a catalyst (a substance that increases the rate of a chemical reaction) to build amino acids. Additionally, these methods sometimes require extreme temperatures or oxygen-free chambers. These challenges have motivated scientists to copy the strategy cells have been using to make amino acids for millions of years. Cells use enzymes as catalysts to build amino acids. What are enzymes, and what's special about them?

5. 4. What are enzymes?

Enzymes are a class of proteins that can perform chemical reactions. They are part of our everyday life. For example, the enzymes in the bacteria *Lactobacillus* (naturally found in the human gut) break down lactose in milk into lactic acid during the fermentation of milk into yogurt (see Figure 6). Similarly, the yeast enzymes break down glucose to ethanol and carbon dioxide in the bread-making process. The released carbon dioxide makes the bread rise and leads to soft fluffy bread.

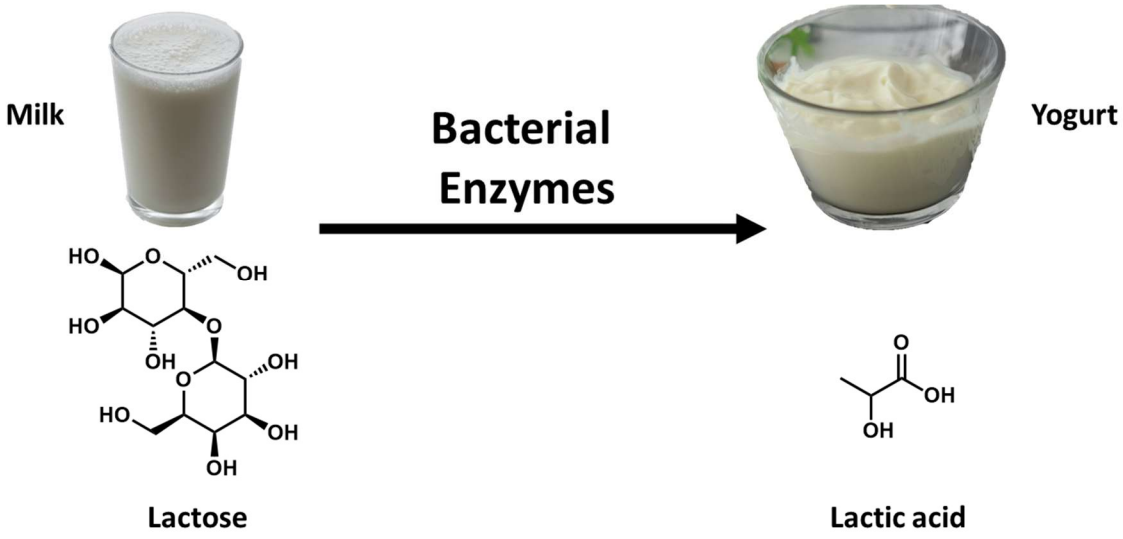


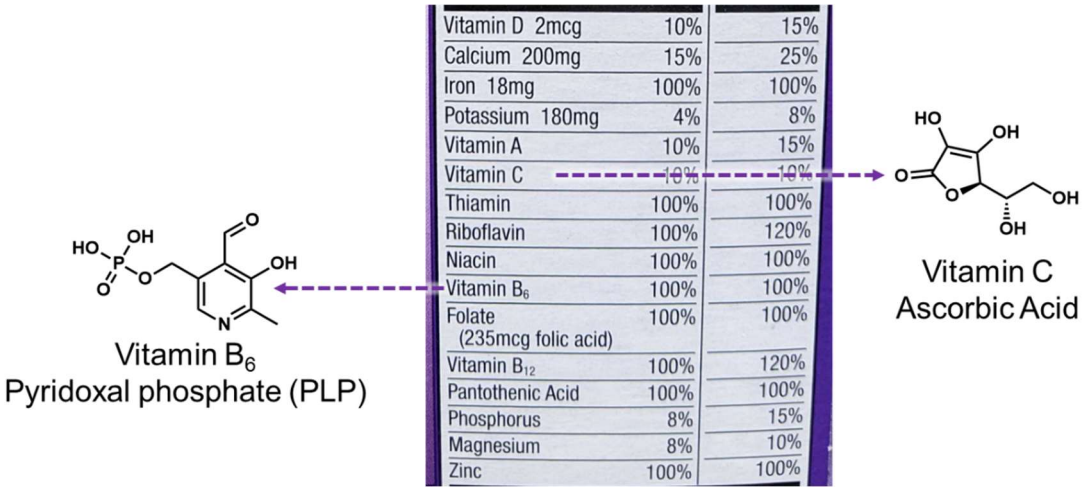
Figure 6. *Lactobacillus* enzymes break down lactose in milk to lactic acid during yogurt formation. Pictures of milk and yogurt were downloaded with permission from unsplash.com.

Several people, including me, have lost the ability to digest milk, a condition called lactose intolerance.[3] This is because as some of us age, our body stops producing lactase, the enzyme that breaks down lactose into glucose and galactose. This indigestion of lactose causes abdominal cramps, bloating, and diarrhea. Fortunately, I can buy lactase pills from a pharmacy. Once consumed, the lactase enzyme breaks down lactose in the digestive tract, and I can still enjoy dairy products.

Enzymes have been used by scientists for several applications due to their ability to work well outside our bodies. For example, lipases that break down fat stains are found in laundry detergents.[4] Enzymes are used to produce bioethanol (a commercially used biofuel) from the degradation of starchy plant materials.[5] Enzymes are also used in the textile, paper, and leather industries. Recently, enzymes that can break down plastics have been discovered.[6] Hopefully, in the coming years, plastic that was thought to take millions of years to break down could be broken down by enzymes in a few weeks.

5. 5. What are vitamins?

Some of the enzymes need helper molecules to perform their functions. These helper molecules are called cofactors. These could be inorganic metal ions such as Magnesium or organic compounds such as vitamins. There are 13 essential vitamins, and they also have roles beyond helping enzymes, such as preventing infections and maintaining strong bones. Our body cannot synthesize vitamins like vitamin B_6 and vitamin C, and we must acquire them through dietary sources. Like vitamins, minerals like iodine and fluoride also help our body function well. This is why we need to consume a variety of fruits and vegetables to make sure our body gets sufficient vitamins and minerals. Vitamins and minerals are also added to foods such as milk, cereals, fruit juices, and even salt to boost their nutritional value (see Figure 7).



Vitamin D 2mcg	10%	15%
Calcium 200mg	15%	25%
Iron 18mg	100%	100%
Potassium 180mg	4%	8%
Vitamin A	10%	15%
Vitamin C	10%	10%
Thiamin	100%	100%
Riboflavin	100%	120%
Niacin	100%	100%
Vitamin B_6	100%	100%
Folate (235mcg folic acid)	100%	100%
Vitamin B_{12}	100%	120%
Pantothenic Acid	100%	100%
Phosphorus	8%	15%
Magnesium	8%	10%
Zinc	100%	100%

Figure 7. Chemical structure of select vitamins found in the Cheerios cereal box.

With the help of cofactors, enzymes can catalyze (accelerate the rate of) several different chemical reactions in our bodies. Chemists have harnessed the power of naturally occurring enzymes from humans and microbes to convert simple starting materials into biologically valuable molecules. This field of research is called ‘biocatalysis’ and has revolutionized the synthesis of chemical building blocks.

5. 6. Why did we choose to synthesize β-hydroxy amino acids?

We chose to synthesize a class of amino acids called β-hydroxy amino acids (β is pronounced as beta). These β-hydroxy amino acids contain a hydroxy group (-OH group) at the β-carbon of the amino acid backbone. The R-group, or the side chain, determines the properties of these amino acids. For example, if the R-group is a hydrogen atom, the resulting amino acid is serine. Similarly, if the R-group is a methyl (-CH_3) group, the resulting amino acid is threonine (see Figure 8). Serine and threonine are two of the twenty amino acids found in proteins.

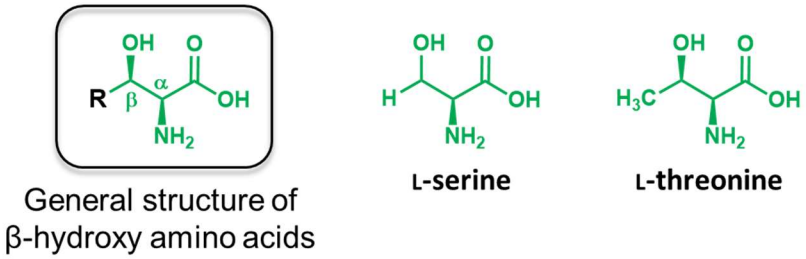


Figure 8. The chemical structure of β-hydroxy amino acids and two proteogenic β-hydroxy amino acids are shown on the right.

A number of microbes have enzymes that synthesize β-hydroxy amino acids with different R-groups, which are eventually used as building blocks for the synthesis of more complex biomolecules, also known as natural products.[7] These natural products include vancomycin (clinically used antibiotic) and cyclosporin (used as an immunosuppressant during organ transplantation). Inspired by these natural products, medicinal chemists have also designed several pharmaceutical drugs that contain β-hydroxy amino acid motifs, such as droxidopa (treatment of Parkinson's disease) and vilanterol (treatment of asthma). I hope you can see that β-hydroxy amino acids are biologically valuable compounds. However, existing chemical and enzymatic methods for the synthesis of β-hydroxy amino acids have limitations. Hence, we decided to search for a new enzyme to synthesize several β-hydroxy amino acids.

5. 7. How did we use enzymes and vitamins to build β-hydroxy amino acids?

We chose an enzyme called ObiH (pronounced as Oh-Bee-H), which is found in a soil bacteria called *Pseudomonas fluorescens*.[8] This bacterium uses the ObiH enzyme to make a β-hydroxy amino acid. This β-hydroxy amino acid is used as a building block to synthesize obafluorin, an antibiotic.

ObiH uses threonine and aldehyde as the two substrates (starting materials), with vitamin B_6 as the cofactor (see Figure 9, native reaction). Threonine is one of the 20 standard amino acids. Aldehydes are compounds in which a carbon atom shares a double bond with an oxygen atom, a single bond with a hydrogen atom, and a single bond with another atom or group of atoms. These two substrates are cheap and commercially available. Out of the two substrates, the aldehyde is the one that determines the side chain or the R-group of the resulting β-hydroxy amino acid. We asked the question: Could ObiH react with other aldehydes to make new β-hydroxy amino acids? (see Figure 9, desired reaction).

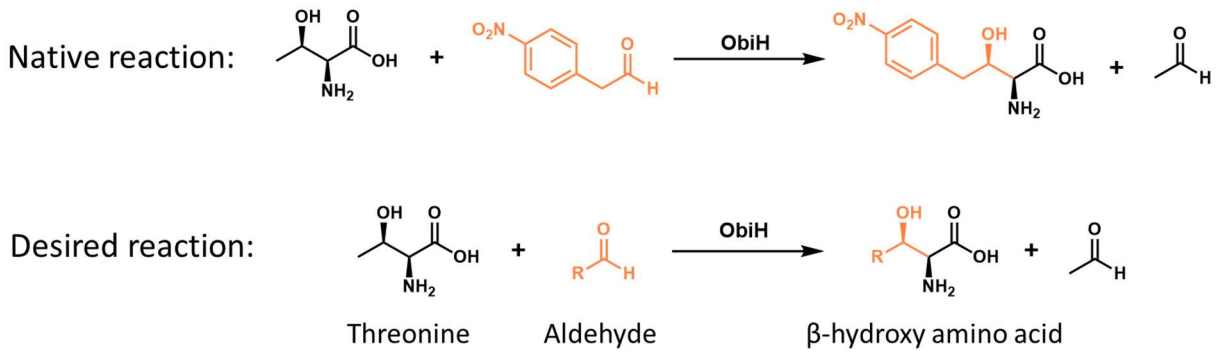


Figure 9. The native and the desired reaction catalyzed by ObiH. Aldehyde (orange) determines the side chain of the resulting β-hydroxy amino acid.

To answer our question, we bought a chemically synthesized ObiH gene (gene is the sequence of DNA alphabets that has the instructions for the cells to make a protein). We stitched this ObiH gene into a plasmid (a circular piece of DNA) and transferred it into *E. coli* (a bacteria

used in the labs to produce proteins). When we grow these *E. coli* bacteria in a liquid growth media (food for bacteria), they multiply rapidly and produce the ObiH enzyme inside them (see Figure 10). ObiH is a pink enzyme, so the bacteria turn pink in color.

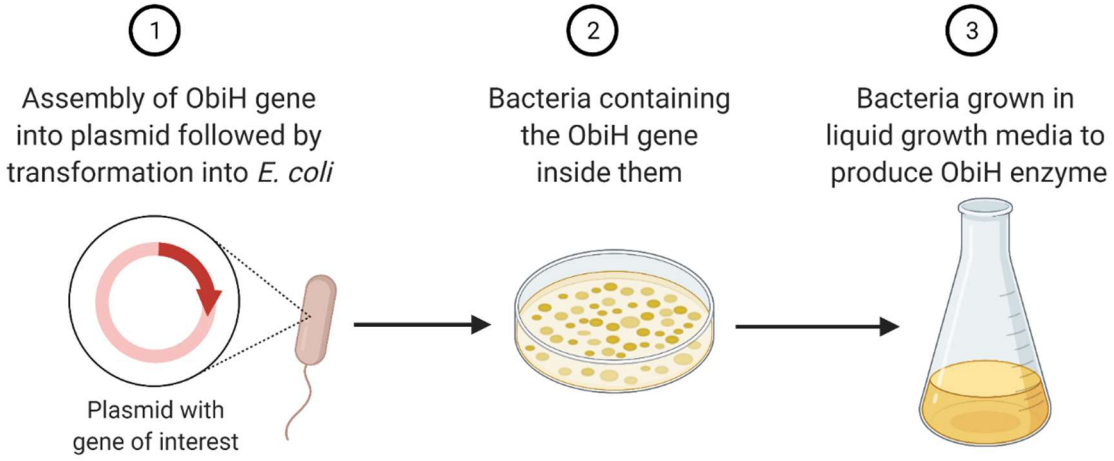


Figure 10. Production of the ObiH enzyme in *E. coli*. First, the commercially bought ObiH gene is assembled into a plasmid and transformed into *E. coli*. This *E. coli* can be grown in a liquid growth media to produce a large number of cells with the ObiH enzyme inside them. *This figure was created with BioRender.com from a pre-existing template.*

To isolate the cells from the growth media, we perform a centrifugation step. In centrifugation, liquid growth media containing bacteria is transferred to a specially designed tube and spun at high speeds using an instrument called a centrifuge. Spinning generates a centrifugal force that causes the heavy cells to settle down at the bottom of the tube. After centrifugation, we dump the liquid growth media into a biowaste container. Then, we transfer the remaining cells into a syringe and inject them into liquid nitrogen (see Figure 11). This causes cells to freeze rapidly, and these frozen cells can be stored in the freezer for a few months. Whenever we need to do experiments, we take a few of the cell noodles out of the freezer and add them to our reaction mixture.

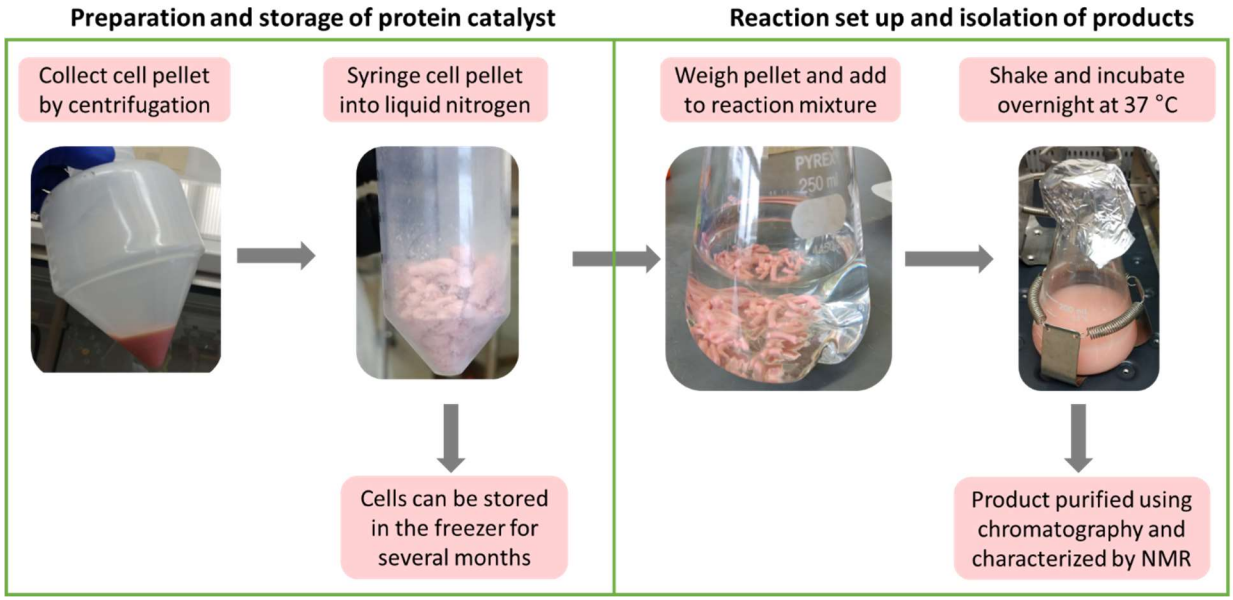


Figure 11. A flowchart describing the steps involved in the preparation of ObiH cells, the reaction set up, and isolation of the β-hydroxy amino acid.

The reaction mixture contains threonine, aldehyde, vitamin B_6, and ObiH cells dissolved in a buffer (a salt and water solution that keeps the cells happy). We shake this mixture in a flask at 37 °C overnight using an instrument called a shaking incubator. During this time, ObiH catalyzes the chemical reaction between threonine and aldehyde using vitamin B_6 as the cofactor to build the β-hydroxy amino acid. We use two techniques (chromatography in conjunction with Nuclear Magnetic Resonance or NMR) to isolate our desired β-hydroxy amino acid and determine its purity.

We found that ObiH could catalyze a reaction with almost every aldehyde we tested. We were able to successfully isolate 15 different β-hydroxy amino acids (in the range of 50 to 500 milligrams of the compound), a few of which were never made by anyone before (see Figure 12).[9] Our methodology used cheap and commercially available starting materials with a naturally occurring enzyme as the catalyst to isolate several β-hydroxy amino acids in an environmentally friendly fashion. This is an exciting result. We have been recently contacted by other researchers who have requested the compounds we made to do their experiments – even more exciting!

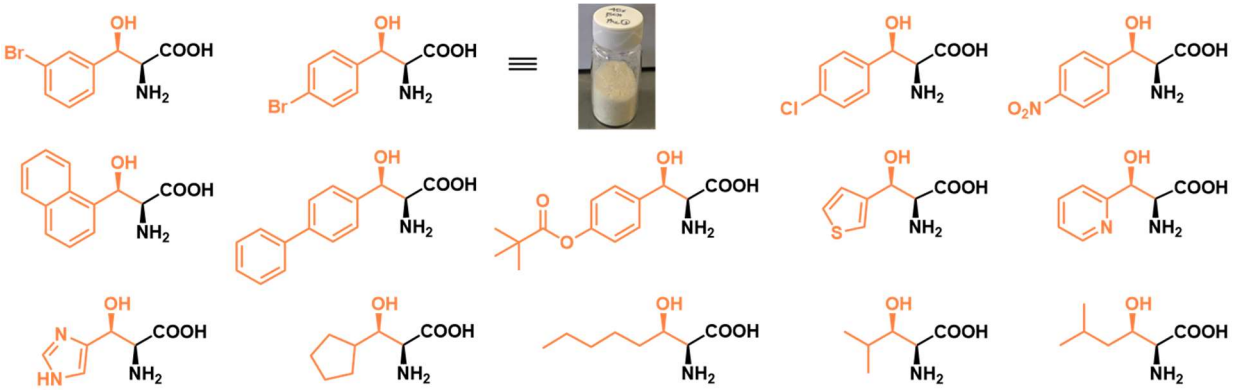


Figure 12. The structures of β-hydroxy amino acids isolated using ObiH as the catalyst are shown. A picture of the isolated compound is shown for one of the β-hydroxy amino acids.

5. 8. What else did we do with ObiH?

In addition to applying the ObiH enzyme for β-hydroxy amino acid synthesis, we were also interested in understanding how the enzyme interacts with the substrate and cofactor spatially to catalyze the chemical reaction. So, we decided to solve the three-dimensional (3D) structure of ObiH using a technique called X-ray diffraction. To do this, we crystallized the ObiH enzyme by slowly evaporating the buffer in which ObiH is stored (similar to making salt by evaporating seawater). We then shine powerful X-rays onto the crystals. When these X-rays pass through and exit the crystal, they produce a diffraction pattern that we can use to build a 3D model of the ObiH enzyme (see Figure 13). The ObiH structure gave us information about how the vitamin B_6 cofactor and substrates interact with the enzyme.[10] Most importantly, this structure gave us clues about how to modify the enzyme (also known as protein engineering) to enable the ObiH enzyme to react with more substrates.

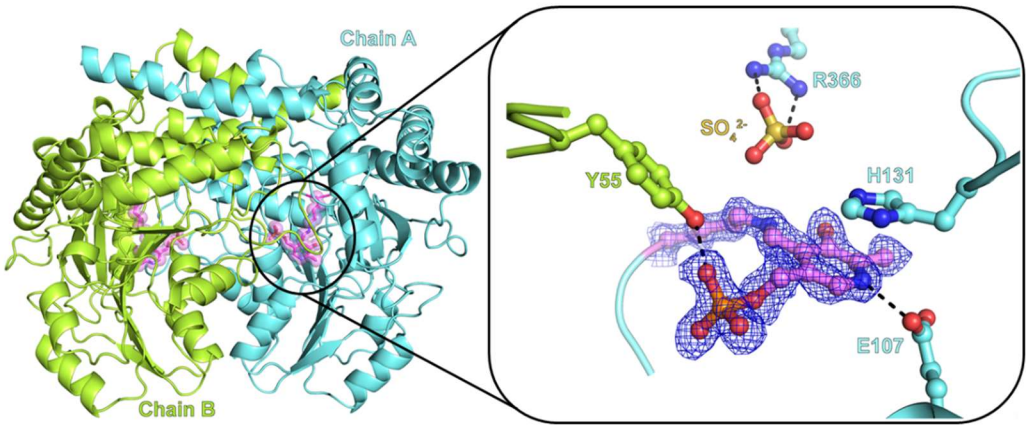


Figure 13. The 3D structure of ObiH. ObiH crystallized as a dimer (which means there are two copies of ObiH in contact with each other). Each of the monomers (one copy of ObiH) is shown in green and blue, respectively. The zoomed panel shows the enzyme's active site. The active site is the part of the enzyme where the starting materials and cofactor bind for the chemical reaction to take place.

5. 9. What if the ObiH enzyme does not accept specific substrates?

The ObiH enzyme accepts several different aldehydes as substrates to make β-hydroxy amino acids. But there are a few aldehydes that ObiH did not accept or accepted poorly as a substrate.[9] So, we decided to evolve the ObiH enzyme to accept even more aldehydes. What does it mean to evolve an enzyme?

The *E. coli* cells we have in the lab contain the ObiH gene (sequence of DNA alphabets that has the instructions to make the ObiH enzyme). We use a technique called mutagenesis to make small changes to the ObiH gene. This change in the gene changes the structure and function of the enzyme (see Figure 14). Some changes could be beneficial (reactivity with more substrates), whereas others could be harmful (no reactivity or reactivity with fewer substrates). Since it's challenging to predict what changes would be beneficial, we evaluate as many mutated enzymes as possible to identify an improved enzyme. This methodology of improving protein activity is called protein engineering.[11]

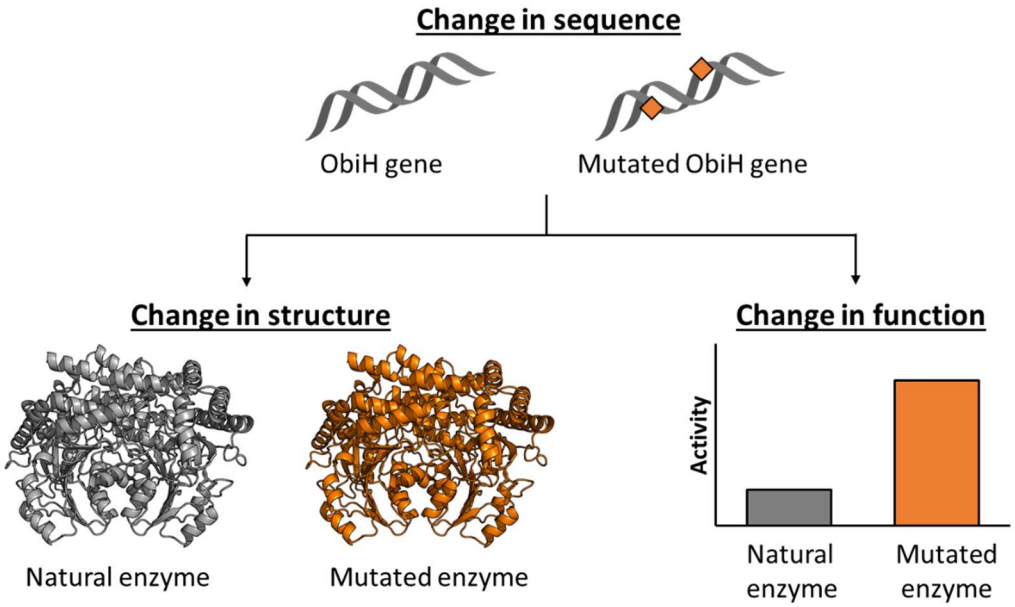


Figure 14. Protein engineering allows us to tailor enzymes for improved functionality in the lab.

Once we identify an improved enzyme, we subject it to additional changes to see if we can make the enzyme even better. We can repeat this process of small modifications as many times as we want. This process of gradually making changes to improve the enzyme activity is called directed evolution, and this methodology was awarded the Nobel Prize in Chemistry in 2018.[12] This process is similar to natural evolution and the concept of "survival of the fittest." In the natural world, evolution takes hundreds to millions of years. In the lab, we accelerate this evolution process over a few weeks.

We subjected the ObiH enzyme to directed evolution. The 3D structure of ObiH guided us on what changes to make. After evaluating hundreds of mutated enzymes, we have found a few mutated enzymes that accept more aldehydes as substrates. Now, we have evolved ObiH enzymes in the lab that are much better at producing a wide range of β-hydroxy amino acids than the naturally occurring ObiH enzyme.

5. 10. References

1. Watford, M. & Wu, G. Protein. *Adv. Nutr.* **9**, 651 (2018).

2. Blaskovich, M. A. T. Unusual Amino Acids in Medicinal Chemistry. *Journal of Medicinal Chemistry* **59**, 10807–10836 (2016).

3. Szilagyi, A. & Ishayek, N. Lactose Intolerance, Dairy Avoidance, and Treatment Options. *Nutrients* **10**, (2018).

4. Niyonzima, F. N. Detergent-compatible bacterial cellulases. *J. Basic Microbiol.* **59**, 134–147 (2019).

5. Vasić, K., Knez, Ž. & Leitgeb, M. Bioethanol Production by Enzymatic Hydrolysis from Different Lignocellulosic Sources. *Molecules* **26**, (2021).

6. Amobonye, A., Bhagwat, P., Singh, S. & Pillai, S. Plastic biodegradation: Frontline microbes and their enzymes. *Sci. Total Environ.* **759**, (2021).

7. Wang, S. & Deng, H. Peculiarities of promiscuous l-threonine transaldolases for enantioselective synthesis of β-hydroxy-α-amino acids. *Appl. Microbiol. Biotechnol.* **105**, 3507–3520 (2021).

8. Schaffer, J. E., Reck, M. R., Prasad, N. K. & Wencewicz, T. A. β-Lactone formation during product release from a nonribosomal peptide synthetase. *Nat. Chem. Biol.* **13**, 737–744 (2017).

9. Doyon, T. J. *et al.* Scalable and Selective β-Hydroxy-α-Amino Acid Synthesis Catalyzed by Promiscuous l-Threonine Transaldolase ObiH. *ChemBioChem* **23**, e202100577 (2022).

10. Kumar, P. *et al.* l-Threonine Transaldolase Activity Is Enabled by a Persistent Catalytic Intermediate. *ACS Chem. Biol.* **16**, 86–95 (2020).

11. Lutz, S. & Iamurri, S. M. Protein Engineering: Past, Present, and Future. *Methods Mol. Biol.* **1685**, 1–12 (2018).

12. Bloom, J. D. & Arnold, F. H. In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U. S. A.* **106 Suppl 1**, 9995–10000 (2009).