Summer Scholar Report

Library development of D-enantiomer trifunctional chemical probes to assess the effect of probe chirality on protein selectivity in MCF7 breast cancer cells

Naomi Suminski, Kelly Harrison, Shalise Couvertier, Nancy Lee, Department of Chemistry and Physics, Simmons University, Boston, MA 02215

Abstract:

With the overwhelming rates of breast cancer and improved therapeutics needed, the development of small molecule chemical probes to use as tools to better understand the biology behind disease manifestation and the protein activities implicated in oncogenesis is essential. The goal of this project was to synthesize a probe library of D-enantiomer trifunctional chemical probes to study cysteine-mediated protein activities of estrogen receptor (ER) and progesterone receptor (PR) positive breast cancer cells. Distinct functional groups were installed onto the three sites of the probe scaffold, including an amino acid side chain directing group to target specific proteins in MCF7 ER/PR positive breast cancer cells. Specifically, focus was given to the incorporation of D-amino acid, a stereoisomeric form not innate to human proteins, to reveal the effect of chirality on probe design and protein specificity. At this time, the three synthesis steps in the development of these probes have been characterized for tyrosine and phenylalanine amino acid methyl esters, and multiple other amino acid probes are currently being synthesized. Techniques and protocols used for analysis of the probes-protein target interaction and activity have also been explored, with plans in the coming months to screen our synthesized probe library. This analysis could reveal unique insight into the mechanisms of breast cancer and may lead to future drug development for disease treatment.

Introduction and Background:

In 2018 alone, an estimated 266,120 new cases of invasive breast cancer will be diagnosed making it the second most commonly diagnosed cancer among American women.[1] Of these new cases, about 40,920 women are expected to die from the disease[1] A majority of these cases will be estrogen receptor (ER) or progesterone receptor (PR) positive, indicating that the cancer cells contain hormone receptors which receive hormone signals and encourage the cells to grow and proliferate. While hormone therapies exist to help treat the disease it is essential that the mechanisms that lead to this type of cancer continue to be explored in order to develop improved therapeutic options.[2]

The development of chemical probes is one way to explore the biological mechanisms behind the manifestation of breast cancer. Probes are small, drug-like molecules designed

with various functional chemical groups to identify protein-drug targets and are used as tools to reveal specific protein mechanisms in oncogenesis. Chemical probes are useful tools to study the reactive, functional amino acids responsible for catalysis and regulation in proteins implicated in diseases. These functional amino acids are nucleophilic residues that can be targeted by covalent modification with reactive electrophiles incorporated into probes by organic synthesis. Through various analyses the probe-protein targets can be investigated to gain a better understanding of functional site activity in specific proteins in cancer cells, specifically MCF7 ER/PR positive breast cancer cells.[3] One amino acid of particular interest in the human proteome is cysteine. Although it has a relatively low abundance it plays many unique and vital roles in protein catalysis and regulation. Therefore, designing chemical probes which can selectively target functional cysteine groups can relay important information about their role in protein activity within a cell of interest.

The chemical probe libraries created in this project are designed to covalently target and modify reactive cysteine residues of proteins within the cells of interest. Cysteines play a variety of diverse and functional roles within proteins and are therefore of great interest to be studied as a point of control in various biological processes. Specifically, cysteine is the most nucleophilic amino acid in proteins, and is highly reactive due to the polarizability and electron rich nature of the thiol moiety it contains.[4] Important functions of cysteine residues within proteins include nucleophilic and redox catalysis, metal binding and allosteric regulation.[5,6] As a result of these various roles, cysteine residues can be found on diverse protein types such as proteases, oxidoreductases and kinases.[5,6] These functional roles and proteins are of particular importance in oncogenesis when they become dysregulated. Knowing more about the location, reactivity, binding affinity and other characteristics of these cysteine residues within MCF7 cells via chemical probe interaction could reveal insight into the biochemical mechanisms at play not only in breast cancer oncogenesis but any other cell lines. Our chemical probe library has a cysteine-specific electrophile built into its molecular structure, which has been previously been used in probe libraries to target this amino acid.[6]

Not only can chemical probes be tuned to target specific amino acids, but can also have an innate chirality, which may affect their protein selectivity. Chirality is essential in organisms because organic molecules often exhibit a specific “handedness” or stereoisomeric conformation (Figure 1). This conformation of the molecule is important to its function in the organism as certain enzymes and receptors only respond to a specific chirality. In humans, natural amino acids are in the L conformation. However, the opposite D-enantiomer corresponds to naturally occurring amino acids in bacteria and are

inherent to the bacterial immune response.[7] Using D-amino acids in probes will potentially relate the unique specificities of proteins in humans that recognize to opposite "handed" proteins in bacteria. Comparing the effects of the L-enantiomer library, which is being developed by Kelly Harrison a senior student at Simmons University, to a D-amino acid library, will reveal these unique protein selectivities within the MCF7 ER/PR positive breast cancer cells.

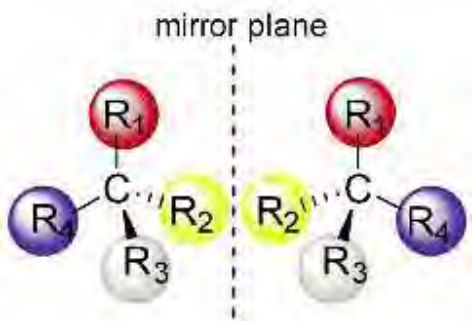

Figure 1. Depiction of carbon chiral centers and chirality A chiral molecule is non-superimposable on its mirror image. In amino acids, the presence of an asymmetric carbon center endows the the property of chirality on the molecule.

The construction of this D-enantiomer chemical probe library follows a three step synthesis in which three functional groups are installed onto a probe scaffold to encourage specificity and target selectivity of the probes. These three functional groups are: (1) a directing group of D-amino acid side chains; (2) an alkyne handle for visualization/enrichment of protein targets; and (3) a cysteine-reactive electrophile for activity-based protein profiling of proteins possessing reactive cysteine residues. Each of these moieties endow important functionality to the probe design. The amino acid directing group (1) allows for specific interaction with protein targets based on directing group structure and chirality. The alkyne handle (2) is the site at which 'click' chemistry can be performed, appending a fluorescent molecule for visualization or a biotin moiety for enrichment of protein targets. Lastly, the the third synthesis step involves attachment of a cysteine reactive electrophile by which the probe covalently modifies reactive cysteine residues in the MCF7 proteome.

The goal of the project was to create two complementary D- and L-enantiomer libraries of 5-7 probes, with each probe consisting of a different amino acid directing group. Thus far, D-phenylalanine and L-tyrosine probes have been created, with other amino acid probes (alanine, glycine, aspartic acid, and methionine) in the process of being synthesized. Following the synthesis of the library we will test the protein selectivity of

the library members via ‘click’ chemistry at the alkyne functional group of each probe and screen in MCF7 lystates and whole cells by SDS-PAGE gel imaging.

Originally, a seven-step synthesis had been proposed at the start of the summer research period, which proved less efficient than we had hoped. The three-step process we used in this project was created from this initial seven-step process in which a diketopiperazine probe scaffold was employed. In the initial process, the first step, a dipeptide coupling proved challenging as it was difficult to characterize our product by NMR, a instrumental constraint. After several attempts, we decided to investigate an alternate synthesis process in which we could still incorporate the three functional groups stated above, in addition to introducing chirality into the probe design.

Overall, we successfully characterized each of our three synthesis steps and will now continue to expand the probe library. Following synthesis, the probes will be tested in MCF7 ER/PR positive breast cancer cell lysates in order to determine their viability and specificity. MCF7 cells will be treated with the probes, exposed to click-chemistry conditions to append a fluorescent molecule to probe modified-proteins and then visualized after SDS-PAGE for fluorescent protein bands. This type of analysis will achieve an understanding of whether the probe was able to enter the cell, reach a protein target, and covalently modify the protein target. Based on preliminary data, the next steps will be to modify the most promising probes and optimize their interactions with the protein target including tuning their chemical structure via synthesis techniques. Lead probes will be further studied to identify the specific protein target using mass spectrometry at Boston College. After we identify the protein targets, we will perform biochemical assays to determine the effect of probe modification on protein activity.

Results and Progress:

The overall goal of this project is to create a probe library consisting of 5-7 D-enantiomer amino acid derived probes which can selectively target cysteine mediated proteins in MCF7, breast cancer, cells. This probe library will be compared to an identical L-enantiomer probe library to determine the selectivity that chirality contributes to the probe design. Thus far, two major developments have been investigated, the first a synthesis process for the probe library, and the second, a process to assess the selectivity of the probe’s developed and their biological targets in MCF7 breast cancer cells.

The first major stage of this project was developing a synthesis experimental protocol, which includes a three-step process by which three functional moieties are installed onto a probe scaffold. These functional groups, as discussed in the previous section, are critical to probe cysteine residue selectivity, covalent modification of the protein, and probe visualization by fluorescence. This stage included a number of optimization

experiments which were carried out to increase the economy, yield, and safety of each step. The second stage of project experimentation was the utilization of a 'click' chemistry protocol by which the synthesized probes are analyzed. This process includes attachment of a fluorescent moiety, a rhodamine (TAMRA) azide tag, to an alkyne handle incorporated into the probe scaffold. Visualization of this fluorescent tag via SDS-page and fluorescent gel imaging techniques allows qualitative assessment of probe selectivity. Finally, comparison of our findings to the complementary library of L-enantiomer probes will help to determine the effect of chirality on target selectivity.

Initially a seven-step synthesis was proposed in which a diketopiperazine (DKP) probes scaffold would be used, onto which the three functional groups described prior would be incorporated. DKP's are a natural small molecule and secondary metabolite generated by fungi, and are useful building blocks for probes.[8] The original DKP probe design was chosen for its ability to have innate chirality and to easily permeate the cell membrane. However, despite the promising design and attempts to follow the initial seven-step process, difficulties to complete and characterize one of the intermediate products resulted in a re-design of the probe scaffold and synthesis steps. In particular, a dipeptide intermediate (the second intermediate compound in the synthesis schema) was difficult to fully characterize by our limited HNMR and IR instrumentation. This synthesis step consisted on a peptide coupling under EDC catalysis conditions in an attempt to yield a linear dipeptide from two amino acid methyl esters and generate the DKP scaffold. This synthesis was attempted with multiple amino acid methyl esters of both enantiomers but the same challenges to characterization occurred with each reaction iteration. Due to the efficiency and time constraints of the summer research period, we decided to implement a simpler probe scaffold.

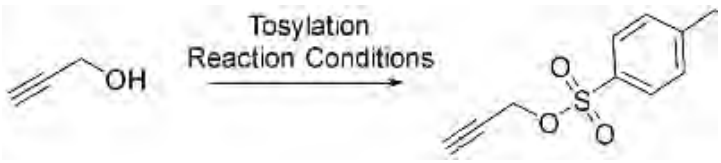

Figure 2. Tosylated Alkyne Handle Synthesis Reaction of p-toluenesulfonyl chloride by propargyl alcohol yields a product essential to installing an 'alkyne' handle onto the probe scaffold in the next synthesis step (Conditions for reaction: triethylamine, tosyl chloride, DCM, 0 °C, 12 hrs).

This alternate probe design was hypothesized to be as effective at successfully covalently modifying a cysteine protein target with a more efficient and cost-effective synthesis protocol. This simpler probe, an amino acid methyl ester backbone with the three functional groups installed, was then synthesized and characterized starting with a D-

phenylalanine methyl ester hydrochloride. The three synthesis steps are as follows: (1) preparation of tosylated alkyne (alkyne handle) for attachment in the following step (Figure 2), (2) monoalkylation of an amino acid methyl ester, with the R-group of the amino acid acting as the probe's directing group, and (3) chloroacetyl chloride addition serving as the reactive electrophile for the probe (Figure 3). Thus far a D-phenylalanine methyl ester probe has been synthesized and characterized. In the development pipeline are D-tyrosine, D-methionine, Daspartic acid, and D-glycine methyl ester probes. Complementary L-amino acid probes are also being synthesized, with an L-tyrosine methyl ester probe fully characterized during the summer research period.

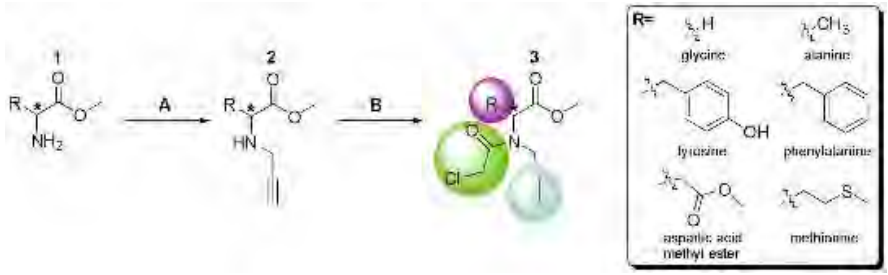

Figure 3. Synthesis of Amino Acid Methyl Ester Probe Reaction A demonstrates the monoalkylation of an amino acid methyl ester using propargyl tosylate. (Conditions for Reaction A: Sodium iodide, potassium carbonate, acetonitrile, 90 °C, 12hrs.). Reaction B shows the synthesis of the final probe from the monoalkylated intermediate via a chloroacetyl chloride coupling step (Conditions for Reaction B: secondary amine intermediate, chloroacetyl chloride, triethylamine, and dichloromethane, 0 °C,, 12 hrs.) The final probe includes the three essential functional moieties describes above: amino acid directing group (pink), reactive cysteine electrophile (green), and alkyne handle for 'click' chemistry (blue).

Considerable measures were taken in the development of each reaction step including reaction conditions, TLC and purification conditions, and proton nuclear magnetic resonance (1H NMR) for each intermediate step. For each intermediate and the final reaction product the general workflow was as follows: reaction set-up, extraction, purification by column chromatography, and finally characterization.

Developing each of the three synthesis steps, and characterizing each intermediate product, came with its own set of challenges. In particular, optimizing yields due to water and humidity conditions during summer months was essential. This was particularly true in conducting the first synthesis step (Figure 2) in which a tosylated alkyne was synthesized. A considerable portion of the summer research period was focused on optimizing this reaction. A number of iterations of this synthesis step were conducted in which reaction conditions were manipulated. Initial attempts at producing the tosylated product resulted in low yields (6% yield) and excess tosyl chloride, one of the starting materials, after aqueous work-up of the crude reaction mixture. Because of this excess of

starting material and concerns that humidity was affecting the reaction a number of experiments were performed to increase our product yield.

First, we increased the equivalence of tosyl chloride by 36% which slightly increased the overall reaction yield (9% yield). In the next iterations of the reaction we attempted to eliminate the effect of humidity in all aspects of the reaction. We were scrupulous in our anhydrous technique, including drying all of our reagents with sodium sulfate, using molecular sieves, and using sure-sealed solvent. These attempts, however, did little to improve our yields. However, by scaling the reaction to produce a few hundred milligrams of the tosylated alkyne product, enough material was synthesized and purified to confidently characterize by 1H NMR and proceed to the next step.

In step two, seen as the first reaction in Figure 2, we utilized our first step product to alkylate an amino acid methyl ester. This reaction required refluxing over high heat for 12 hours, extracting and concentrating in vacuo after column purification. Initial attempts at this reaction produced moderate yields with D- and L- phenylalanine methyl esters (54% yield) as well as the L-tyrosine methyl ester (63% yield) and enough product to purify and then confidently characterize by 1H NMR. Loss of yield was due to the generation of a dialkylated product in addition to the desired monoalkylated product.

In an effort to increase the efficiency and cost-effectiveness of our process even more, we investigated if we could synthesize the same probe in two steps instead of three. An alternate method for monoalkylation of amino acid methyl esters was utilized to increase production of the monoalkylated amino methyl ester and decrease the production of the alkylated product (see Cho et al., 2002).[8] This procedure allowed us to side-step the tosylation step and directly alkylate our amino acids of interest via propargyl bromide and lithium hydroxide. In reference to Figure 3, this process allowed us to synthesize compound 2 with a more efficient workflow. It allowed for more manageable reaction conditions, without the need to reflux with high heat. This reaction was performed with both D- and L-phenylalanine methyl esters with plans to apply it to other amino esters of interest.

The third and final synthesis step began with compound 2, the monoalkylated amino methyl ester and attached a cysteine reactive electrophile in the form of a chloroacetyl chloride addition. This step was completed and fully characterized thus far for L-tyrosine and is the final stages of characterization for D-phenylalanine.

In addition to our efforts towards building our probe library, we also developed a protocol to assess the cell permeability, protein selectivity and binding affinity of each probe after

it's addition to the library. The foundation of this 'click' chemistry protocol is to append a fluorescent tag to the alkyne handle of each probe (essentially "clicking" the fluorescent tag onto the probe's alkyne handle), expose the probe to cell lysates at various concentrations, analyze these sample by SDS-PAGE, and visualize the fluorescence of each probe with gel imaging (Figure 4).[3]

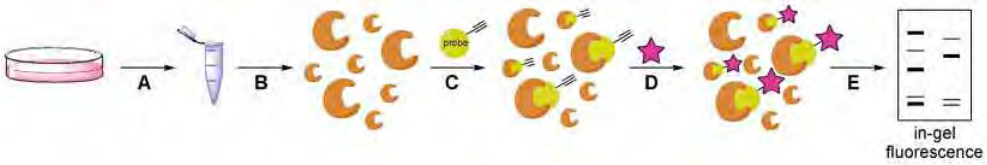

Figure 4. 'Click' Chemistry Work-Flow Performing the protocol utilized the following experimental schema: (A) lyse the cells of interest (B) normalize protein concentration to 1 mg/ML (C) label the cells with probes (D) perform copper click chemistry with rhodamine azide fluorescent tag and (E) SDSPAGE gel analysis for protein target visualization.

The relative fluorescence seen in each sample is a direct indication of protein targeting by the probes and their binding affinity. Because our probe library was still in the developing stages, this 'click' chemistry protocol was developed using a different, but similar probe library, which also incorporated an alkyne handle into its design. A more promiscuous control was used, an iodoacetamide (IA) probe, which is used for global profiling of cysteines.

Discussion:

In summary, the progress of this research project has been the development and characterization of a three step synthesis to create a library of D-enantiomer amino acid methyl ester chemical probes and a complementary L-enantiomer library to assess the effect of chirality on probe target. So far, this synthesis protocol has been used to create a D-phenylalanine methyl ester probe and L-tyrosine methyl ester probes. Additional D- and L- amino acid probes are in various synthesis steps and will soon be fully synthesized, characterized and added to their respective libraries. In the meantime, a 'click' chemistry protocol has also been exercised and will be used in future experiments to assess each probe's biological activity and protein selectivity in cell lysates and eventually, whole cells.

Moving forward, the focus of the project will be to expand the probe library. After additional D-enantiomer amino acid methyl esters, including glycine, methionine, aspartic acid, glutamic acid, and tyrosine, probes are synthesized, each which will undergo 'click' chemistry reactions to append a rhodamine azide fluorescent tag.

Subsequently they will be exposed to MCF7 cell lysates to assess if they were able to reach a protein target. Pending this interaction, we will also assess their ability to selectively target proteins in whole cells and evaluate their ability to permeate the cell membrane. The ultimate goal of this biological analysis of our probe library will be identify the specific cysteine residue targets of the probes.

One exciting potential of these libraries is their potential to modify not only the cysteine residues of MCF7 cells, but their ability to be screened against any cell line. Expanding our biological assessment of the libraries to other oncogenic cell lines could also garner information about important cysteine residues in other types of cancer. Even more, if a relationship between probe chirality and protein target is established, this concept can be applied to the design of other existing and novel probe libraries to generate more information about how stereochemistry plays a role in covalent protein modification.

Acknowledgements:

I would like to thank the Northeastern Section of the American Chemical Society for their generous support of this project and the opportunity to share my research with their members. I would also like to thank my fellow Simmons University student and researcher, Kelly Harrison, for her shared dedication and partnership on this project. Finally, thank you to my advisors: to Dr. Shalise Couvertier, without whom this project would not be possible, thank you for your mentorship and guidance. And to Dr. Nancy Lee, thank you for your tireless commitment and organic chemistry expertise.

Appendix:

1. **Synthesis of a tosylated alkyne (Compound 1):** To a round bottom flask, 3 mL (51.9 mMol) of propargyl alcohol and 93 mL of DCM was added under nitrogen atmosphere. The mixture was cooled to 0°C in an ice bath. Next, 8 mL (57 mMol) of triethylamine and 12 g (77.86 mMol) of tosyl chloride was added. The reaction flask was purged with nitrogen and allowed to slowly warm to room temperature. The reaction was allowed to run overnight. The reaction was quenched with water (100 mL) and extracted with DCM (3x80 mL). The resulting reaction mixture was purified by flash chromatography (2:10 Ethyl Acetate, Hexanes v/v). The product was concentrated in vacuo to yield as a clear oil (9% yield) 1H NMR 90 MHz (DMSO) δ 2.44 (s, 3H) δ 2.58 (t, 1H, J = 2.25 Hz) δ 4.701 (s, 2H) δ 7.309 (d, H2, J = 8.01 Hz) δ 7.752 (d, 2H, J = 8.28 Hz)
2. **Synthesis of monoalkylated amino acid methyl esters, (Compound 2):** To a double necked round bottom flask with a stir bar was added sodium iodide (0.58 mMol), potassium carbonate (3.52 mMol), and L-tyrosine methyl ester (1.17 mMol). This was dissolved in 1.5 mL of acetonitrile. The round bottom was connected to a reflux condenser and purged with nitrogen. The reaction was allowed to heat to

90°C , then the tosylated alkyne was dissolved in acetonitrile (1.5 mL) and added to the reaction vessel drop wise. The reaction was allowed to stir for 24 hours at 90°C . The reaction was then quenched with water and extracted with DCM (3x10 mL) and dried with sodium sulfate. The resulting reaction mixture was purified by flash chromatography (2:10 Ethyl Acetate, Hexanes v/v). The product was concentrated in vacuo to yield as a yellow, translucent oil (63% yield). 1H NMR 90 MHz (DMSO) δ 1.906 (s,1H) δ 2.698 (d, 2H, J = 6.57 Hz) δ 3.029 (m, 2H, J*) δ 3.52 (m, 1H, J*) δ 3.54 (s, 3H) δ 6.683 (d, 2H, J*) δ 6.897 (d, H2, J*) *Some limitations in 1H NMR resolution inhibited accurate J value determination.

3. **Synthesis of final amino acid methyl ester probe (Compound 3):** To a flame dried vial equipped with stir bar was added the monoalkylated L-tyrosine methyl ester (0.4467 mMol) and DCM (2 mL). The reaction vessel was purged with N2 and cooled to 0˚C. Next, chloroacetyl chloride (0.594 mMol) was added drop-wise followed by triethylamine (0.594 mMol) added dropwise as well. The resulting mixture was allowed to warm to room temperature and stir for 8 hours. The reaction was quenched by the addition of sodium bicarbonate (10 mL), extracted with DCM (3x10 mL) and concentrated in vacuo to yield a crude oil. The resulting mixture was purified by column chromatography (2:10 Ethyl acetate, hexanes v/v). (39 % yield) 1H NMR 90 MHz (CDCl3) δ 2.291 (m, 1H, J = 7.2 Hz) δ 3.727 (s, 3H) δ 3.964 (m, 2H, J = 4.05 Hz) δ 4.153 (s, 2H) δ 4.286 (s, 2H) δ 4.8 (m, 1H, J = 9 Hz) δ 6.977 (d, 2H, J = 8.46 Hz) δ 7.217 (d, 2H, J = 8.19 Hz)

References:

1. Siegel, R. L., Miller, K. D., & Jemal, A. (2018). *Cancer Statistics*, 2018. CA: a cancer journal for clinicians, 68(1), 7-30.
2. DeVita, Vincent T.; Lawrence, Theodore S.; Rosenberg, Steven A.; Robert A. Weinberg; Ronald A. DePinho (2008-04-01). *DeVita, Hellman, and Rosenberg's cancer: principles & practice of oncology*. Lippincott Williams & Wilkins. pp. 1646–.
3. Couvertier, S. M. (2016). *Chemical-proteomic strategies to study cysteine posttranslational modifications* (Doctoral dissertation, Boston College).
4. Weerapana, Eranthie, et al. "Quantitative reactivity profiling predicts functional cysteines in proteomes." *Nature* 468.7325 (2010): 790.
5. Pace, Nicholas J., and Eranthie Weerapana. "Diverse functional roles of reactive cysteines." *ACS Chemical Biology* 8.2 (2012): 283-296.
6. Shannon, D. Alexander, and Eranthie Weerapana. "Covalent protein modification: the current landscape of residuespecific electrophiles." *Current Opinion in Chemical Biology* 24 (2015): 18-26.
7. Cava, F., Lam, H., De Pedro, M. A., & Waldor, M. K. (2011). Emerging knowledge of regulatory roles of Damino acids in bacteria. *Cellular and Molecular Life Sciences*, 68(5), 817-831.

8. Cho, Jong Hyun, and B. Moon Kim. “LiOH-mediated Nmonoalkylation of α-amino acid esters and a dipeptide ester using activated alkyl bromides.” *Tetrahedron Letters* 43.7 (2002): 1273-1276.
9. Heydari, Akbar, et al. “A general one-pot, three-component mono N-alkylation of amines and amine derivatives in lithium perchlorate/diethyl ether solution.” *Synthesis* 2005.04 (2005): 627-633.
10. Ikawa, Takashi, et al. “Selective N-alkylation of amines using nitriles under hydrogenation conditions: facile synthesis of secondary and tertiary amines.” *Organic & Biomolecular Chemistry* 10.2 (2012): 293-304.
11. Salvatore, R. N., Nagle, A. S., Schmidt, S. E., & Jung, K. W. (1999). Cesium hydroxide promoted chemoselective N-alkylation for the generally efficient synthesis of secondary amines. *Organic Letters* , 1(12), 1893-1896.
12. Paulsen, Candice E., and Kate S. Carroll. “Cysteine-mediated redox signaling: chemistry, biology, and tools for discovery.” *Chemical Reviews* 113.7 (2013): 4633-4679.