Development of Bifunctional Lysophosphatidic Acid & Lysophosphatidylcholine Activity Probes to Characterize Their Specific Binding Protein Receptors

Ritu Nandal
rnandal@utk.edu

Recommended Citation
https://trace.tennessee.edu/utk_gradthes/1011
To the Graduate Council:

I am submitting herewith a thesis written by Ritu Nandal entitled "Development of Bifunctional Lysophosphatidic Acid & Lysophosphatidylcholine Activity Probes to Characterize Their Specific Binding Protein Receptors." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemistry.

Michael D. Best, Major Professor

We have read this thesis and recommend its acceptance:

George Kabalka, Ben Xue

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)
Development of Bifunctional Lysophosphatidic Acid and Lysophosphatidylcholine Activity Probes to Characterize Their Specific Binding Protein Receptors

A Thesis Presented for the Master of Science Degree
The University of Tennessee, Knoxville

Ritu Nandal
August 2011
Acknowledgements

Firstly, I would like to thank the Almighty for giving me the strength and support whenever things were not going well in life. I am grateful for being able to complete my Masters studies in two years. I would also like to thank my family and friends for their constant support and encouragement to achieve my goals in life. Last but not the least, I appreciate and thank my husband Rohit for his support and understanding when times were rough in life. All of this would not have happened without his patience.

I thank my advisor Dr. Michael Best for his guidance during these two years of graduate study. I express my thankfulness towards all the group members for their mentoring with the lab skills, and the senior lab members always do a great job at training novices like me. Working with you all was a learning experience, and I will miss you all.
Abstract

Lysophosphatidic acid (LPA) and lysophosphatidylcholine (LPC) activity has been found to be dysregulated in cancer cells and therefore is a crucial target for research. Only a few LPA receptors have been identified to date, namely LPA 1, LPA 2, LPA 3, LPA 4 and peroxisome proliferator-activated receptors (PPAR). In order to identify receptors, we are designing and synthesizing bifunctional LPA and LPC activity probes to characterize their protein targets using activity based protein profiling (ABPP) among other proteomic technologies. By synthesizing bifunctional signaling probes that can mimic the naturally occurring LPA and LPC molecules and selectively capture receptors by virtue of their binding properties, we can identify and study the different proteins that are aberrantly expressed in various pathophysiological states such as cancer.
# Table of Contents

**CHAPTER I Introduction** ...........................................................................1  
Background & Significance of Lipid Signaling and LPA Activity ...............2  
Biosynthesis of LPA ...............................................................................4  
LPA Receptors ..................................................................................6  
GPCRs and G-Protein Signaling .........................................................8  
Activity-Based Protein Profiling (ABPP) ..........................................10  

**CHAPTER II Research Design** ............................................................16  
Experimental Procedure ...................................................................29  
References .........................................................................................39  

Appendix  
NMR Spectra ..................................................................................42  
Vita ....................................................................................................56
List of Figures

Figure 1. LPA metabolic pathways.................................................................3
Figure 2. Major LPA signaling pathways involving G-coupled receptors.......9
Figure 3. Schematic representation of ABPP using probes bearing photoaffinity
tags...........................................................................................................13
Figure 4. Examples of photo-reactive groups.............................................14
Figure 5. Target LPA and LPC activity probes..........................................16
Figure 6. Synthesis of the head group.........................................................21
Figure 7. Synthesis of the lipid tail..............................................................22
Figure 8. Modular synthesis of the alkyne acid..........................................23
Figure 9. Synthetic strategy for introducing the bifunctional tag..............24
Figure 10. Alternative methods attempted for direct coupling of lipid tails...25
Figure 11. Attempt to Couple bifunctional tags directly onto diol 4...........26
Figure 12. Modular synthetic strategy of the LPA probe.........................27
Figure 13. Modular synthetic strategy of the LPC probe.........................28
List of Abbreviations

ABPP  Activity-Based Protein Profiling
ATX   Autotaxin
cAMP  Cyclic Adenosine Monophosphate
DCC   N,N-dicyclohexylcarbodiimide
DDQ   2,3-dichloro-5,6-dicyanobenzoquinone
EDCI  1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDG   Endothelial Gene Differentiation
EGF   Epidermal Growth Factor
FP    Fluorophosphonate
GPCR  G protein-coupled receptors
GDP   Guanosine Diphosphate
GTP   Guanosine Triphosphate
ICAT  Isotope Coded Affinity Tags
IP_3  Inositol-(1,4,5)-trisphosphate
LC-MS Liquid Chromatography Mass Spectrometry
LPA   Lysophosphatidic Acid
LPC   Lysophosphatidylcholine
LPP   Lipid Phosphate Phosphohydrolases
MAPK  Mitogen-Activated Protein Kinase
Mud-PIT Multidimensional Protein Identification Technology
PAL   Photo affinity labeling
PC    Phosphatidylcholine
PE    Phosphatidylethanolamine
PI3K  Phosphatidylinositol-3-kinase
PIP_2 Phosphatidylinositol-(4,5)-bisphosphate
PLA1  Phospholipase A1
PLC   Phospholipase C
PLD   Phospholipase D
PPAR  Peroxisome Proliferator-Activated Receptor
sPLA2  Secretory Phospholipase A2
TOP   Tandem Orthogonal Proteolysis
CHAPTER I
INTRODUCTION

Lipids are naturally occurring molecules that play important biological functions. They are integral structural components of cell membranes and are important signaling molecules. Eukaryotic cells are composed of a 2-molecule thick lipid bilayer, which primarily consists of phospholipids containing a hydrophilic head group and two hydrophobic tails each. The polar head groups are directed towards the inside (cytoplasm) and to the outside (extracellular matrix). For a long time, lipids were thought to have minimal biological roles that were limited to composition of cell membranes and as intermediates in lipid biosynthesis. However, later studies revealed how lipids play key roles in cell signaling that governs specific cellular responses. The most abundant glycerophospholipids in cells, phosphatidylcholine (PC) and phosphatidylethanolamine (PE), had been regarded as passive structural components of cell membranes for a long period of time. However, studies by Hokin and Hokin in 1950 described how phosphatidylinositol turnover is induced by acetylcholine. This, alongside with studies by Takai and Kishimoto, which revealed that their enzymatic activity as cyclic nucleotide-independent kinase is activated by the active phospholipids in membranes, implies that membrane-bound lipids play a crucial role in the transmembrane control of important biological processes and organelle biogenesis, thus underlying their active rather than passive role in biochemistry of eukaryotic cells. Studies have shown that lipid signaling is a vital component of cell signaling and often involves the binding
of lipids to protein targets, such as receptors, kinases or phosphatases, which in turn mediates the effects of these intermediate lipids in important biological pathways.

**Background and Significance of Lipid Signaling and LPA activity**

Monoacyl-sn-glycero-3-phosphate, commonly known as lysophosphatidic acid (LPA, figure 1) has aroused substantial research interest in the recent past, as its growth-factor-like activity has been found to be upregulated in cancer cells. The causal role of LPA in cancer is substantiated by studies showing that metastasis is characterized by increased activity of the enzyme autotaxin, which produces LPA. Today LPA is known to contribute towards several physiological and pathophysiological processes in vivo. The discovery of the mitogenic nature of LPA and its signaling via specific G-protein coupled receptors (GPCRs) in the 1990s stimulated its emergence as a crucial target for research.\(^5\) LPA is today not only known to be active in several and diverse G-protein-mediated signaling pathways\(^6\), but also activates Ras and Rho/Rac family GTPases to control cell multiplication, migration and morphogenesis, respectively, and therefore also plays a role in wound healing and atherosclerosis among other processes. LPA has been found to be produced in diverse cell types including Schwann cells, leukocytes, T-lymphoma cells, skin, myofibroblasts, cortical cells, neuronal and non-neuronal cells.\(^7\) In addition, LPA levels are increased in a number of pathophysiological states including immunodeficiency diseases, ovarian and other cancer cells and atherosclerosis. This implicates LPA as an important mediator in the onset of cancer.\(^8\)
Therefore, LPA synthesis, signaling, receptors, bioactivity and degradation are crucial targets for research and therapy of cancer and other diseases.

Lysophosphocholine (LPC, figure 1), is a critical glycerophospholipid because it acts as a substrate for the enzyme autotaxin (ATX), which performs phosphate hydrolysis to produce lysophosphatidic acid (LPA). Consequently, it is an important therapeutic target in medicinal chemistry. However, few inhibitors of autotaxin are known so far to have drug like properties. The first ATX-directed fluorescent activity probe has been synthesized and characterized by Cavalli et al. in 2010, that detects and quantifies ATX activity under normal and pathological conditions. This probe contains a trapping moiety that becomes reactive after phosphate ester hydrolysis, and then, specifically and covalently binds to the
active site of ATX. Their study highlights the potential of such ATX-specific ABPs to be used as diagnostic reagents.⁹

**Biosynthesis of LPA**

LPA is synthesized in living systems both intracellularly and extracellularly. Inside the cell, LPA is produced and used as an intermediate for the synthesis of phospholipids. There is no evidence that this intracellular LPA is then released into the extracellular environment. LPA is present in saliva, seminal fluid, incubated plasma, follicular fluids and serum, but not in plasma, as is produced only during serum formation. Ovarian cancer cells and platelets produce LPA upon stimulation.⁷

LPA is produced by two enzymatic mechanisms: a) hydrolysis of the fatty acyl chain of phosphatidic acid (PA) at the sn-2 position by phospholipase A2 (PLA2), or at the sn-1 position by phospholipase A1 (PLA1) (mainly operative intracellularly); and b) Hydrolysis of membrane LPC (Lysophosphatidylcholine) by autotaxin, which removes the hydrophilic choline headgroup (in extracellular body fluids like serum) (Figure 1).⁸ Lysophospholipase D (lysoPLD), which catalyses the production of LPA from LPC, is identical to autotoxin(ATX).¹⁰ Both LPA and LPC contain a single fatty acyl chain. When PLA1 cleaves the sn-1 fatty acid, it produces sn-2 and polyunsaturated LPA isoforms, which are active on LPA receptor 3.¹¹ Type-II secretory phospholipase A2 (sPLA2) cleaves the sn-2 acyl chain, but is incapable of hydrolyzing lipids in intact cell membranes, resulting in low LPA levels in the plasma. However, sPLA2 can selectively
hydrolyze lipids in damaged micro vesicles and cells that are produced in malignant fluids leading to abnormally high LPA levels in cancer cells.\textsuperscript{8}

The protein ATX was first isolated in the early 1990s as an ‘autocrine/tumor motility factor’ produced by human melanoma cells and later characterized as a member of the ecto-nucleotide phosphodiesterase/pyrophosphatase (NPPs) family of enzymes that hydrolyze phosphodiester and pyrophosphate bonds in ATP and ADP.\textsuperscript{12} ATX has been found to be upregulated in several cancers (including lung, renal-cell, mammary, hepato-cellular carcinoma) and is expressed in the highest levels in brain, ovary, kidney, lung, intestine, and testis.\textsuperscript{13} ATX is a type II transmembrane protein with a short amino-terminal region, a single transmembrane domain, two cysteine rich somatomedin-B-like domains and a large catalytic domain. Studies have shown that the motility-stimulating activity of ATX is attributed to production of LPA from LPC, and the detection of elevated ATX expression in malignant tumor cells provides enough evidence for the role of LPA in the onset of cancer.\textsuperscript{14} Therefore, ATX converts extracellular LPC into receptor-active LPA, which stimulates malignant cell proliferation and differentiation of adipose tissue. The mechanism that has been proposed for cancer cell motility is through the activation of G-protein-coupled receptor LPA1. ATX is also upregulated in genetically obese mice. In addition, LPA exists in extracellular fluid of adipocytes and is released by adipocytes in vitro. These studies suggest ATX mediating LPA paracrine control of adipose tissue.\textsuperscript{15}
LPA has been termed as a double-edged sword. Studies have elucidated physiological roles of LPA in the reproductive system, embryonic development, wound healing, tissue repair, blood clotting and to maintain homeostasis of the body against stress. However, it simultaneously plays a key role in the initiation and propagation of various human malignancies and pathological processes. Studies suggest both a protective (under ischemia) and destructive (as causes atherosclerosis) role of LPA in different aspects of cardiovascular disorders. LPA has been identified as a growth factor for cancer (ovarian, prostate, breast, lung, gastric, pancreatic, colorectal cancer) cells by inducing cell proliferation, growth, survival, migration, invasion and angiogenesis. Also, recent studies indicate a significant contribution of LPA signaling to neuropsychiatric disorders and neuropathic pain. In addition, LPA’s involvement in not only the male and female reproductive systems, but also in embryonic development, can help develop new assisted reproductive technologies and birth control. This functional switch of LPA from physiological to pathophysiological mechanisms underscores its signaling and identification of its receptors as a crucial therapeutic target for prevention and treatment of various diseases like cancer. As illustrated by these examples, LPA shows both a positive and a negative role in pathological processes and has a contribution in normal physiological processes, and therefore is the primary focus of my research.

**LPA Receptors**

So far, five distinct LPA receptors (LPA1-LPA5) have been identified that mediate LPA activity. Due to their sequence homology, LPA 1, LPA 2 and LPA 3
receptors belong to the endothelial differentiation gene (EDG) sub-family of the GPCR superfamily (therefore also called EDG 2, EDG 4 and EDG 7, respectively) which traverses the membrane seven times.\textsuperscript{7} LPA 4 was identified recently, which is more closely related to the Purinergic family, and shares no significant similarity with the EDG family.\textsuperscript{7} LPA 1, the first discovered receptor, is most widely expressed and best characterized, whereas LPA2 and LPA 3 have a more restricted distribution, and expression of LPA 4 is very low in most human tissues, but significant in the ovary.\textsuperscript{8} Although LPA\textsubscript{5} is structurally different from LPA\textsubscript{1-3}, but it shares 35% homology with LPA\textsubscript{4}.\textsuperscript{16} A probable sixth LPA receptor, LPA\textsubscript{6}, has been reported recently. LPA\textsubscript{6} is closely related to LPA\textsubscript{4}, and is essential for human hair growth.\textsuperscript{17}

LPA 1 is most widely expressed with high mRNA levels in the colon, small intestine, placenta, brain, heart, and significantly high levels in cerebral cortical ventricular zone during neurogenesis, adult oligodendrocytes and Schwann cells, in addition to moderate expression in pancreas, ovary and prostate.\textsuperscript{8} Expression of LPA 2 and LPA 3 receptors is elevated in cancer cells, especially in ovarian cancer cells, thereby suggesting a potential role in ovarian cancer progression. LPA1, LPA2 and LPA3 receptors of the EDG family exhibit 50% sequence homology in amino acid sequences in all areas except that of the carboxy (C)-terminal tail region of these receptors is much less homologous with only 27% between LPA1 and LPA2, and 17% between LPA2 and LPA3 receptors. This difference in sequences of amino acids only in the C-terminal tail domain could be a possible reason for the specificity of the receptors in mediating LPA
Experiments were conducted in a yeast two-hybrid screening by Xu et al. to understand the mechanisms and specificity by which these LPA receptors mediate LPA activity, and found that TRIP 6 (a focal adhesion molecule) specifically binds to carboxyl-terminal tail of the LPA 2 receptor through its LIM domains and not other LPA 1 or LPA 3 receptors. This specific interaction between TRIP 6 and LPA2 alone might help to study the specific roles of LPA2 in the pathophysiology of ovarian cancer cells.

**G-protein-coupled receptors (GPCRs) and G-protein signaling**

The reason why LPA signals are involved in diverse biological and physiological processes, such as tissue remodeling, neurogenesis, immune, reproductive functions, wound healing, cancer, atherosclerosis, is because LPA receptors can couple to at least 3 different heterotrimeric proteins - $G_q$, $G_i$ and $G_{12/13}$, which further feed into multiple effector pathways. The major downstream effectors are shown in figure 2. LPA activates $G_q$, which stimulates Phospholipase C (PLC), stimulating the hydrolysis of phosphatidylinositol-(4,5)-bisphosphate (PIP$_2$) to diacylglycerol (DAG) and Inositol-(1,4,5)-trisphosphate (IP$_3$).
Figure 2. Major LPA signaling pathways involving G-coupled receptors

LPA also activates G\textsubscript{i}, which feeds into 3 different signaling routes a) inhibition of adenyl cyclase b) G-mediated stimulation of mitogenic RAS-MAP kinase (mitogen-activated protein kinase) cascade, thus triggering downstream events, and c) activation of phosphatidylinositol-3-kinase (PI3K), which leads to the activation of the guanosine diphosphate/guanosine triphosphate (GDP/GTP) exchange factor TIAM 1 and the downstream RAC GTPase, in addition to
activation of AKT’s ‘survival’ pathway to promote cell movement and prevent cell death or apoptosis. Lastly, LPA activates G_{12/13}, which stimulates GTPase RHO A, plays a role in cytoskeletal contraction and cell rounding.\textsuperscript{7} Therefore, LPA, in addition to activating second messengers such as cAMP, calcium and diacylglycerol, also stimulates RAS and RHO family GTPases, which are master switches that control cell proliferation, migration and morphogenesis\textsuperscript{8} (Figure 2).

LPA induced activation of G\textsubscript{i} in turn stimulates the RAS-mediated downstream MAP kinase cascade and plays a role in DNA synthesis, cell proliferation and cell survival. The mechanism of G\textsubscript{i} coupling to Ras is complex and not yet clear. In order to maintain proper cell functionality, the accumulated LPA must be counterbalanced by inactivation processes. It has long been known that extracellular LPA in mammalian cells is degraded by removal of a phosphate group to produce biologically inactive monoacylglycerol.

**Activity Based Protein Profiling (ABPP)**

Advances in genome sequencing techniques have provided useful insights into the crucial aspects of the proteome encoded by genomes. Techniques such as chromosomal translocation, transcriptional profiling, and RNA interference based gene silencing have enlightened us about the physiological and pathological processes of genome sequences, which are controlled by proteins and RNA molecules.\textsuperscript{19} Nevertheless, they do not furnish information about the post-translational events that regulate protein function and activity. To overcome this problem, several proteomic strategies have been developed including LC-MS, gel electrophoresis, multidimensional protein
identification technology (MudPIT) and isotope coded affinity tags (ICAT), yeast
two–hybrid systems, and protein microarrays that enrich the knowledge of the
functions of proteins in vitro. However, they fail to assess protein’s functional
state and activity inside cells and tissues. Therefore, as a complement to
conventional genomic techniques, ABPP has been developed as a functional
proteomic technology to identify and understand the functional state of novel
proteins, and characterize proteins on the basis of activity rather than
abundance.

The chemical proteomic approach of ABPP uses active-site directed
probes to collectively analyze proteins based on activities. ABPP probes use a
range of mechanism-based inhibitors, protein-reactive natural products and
general electrophiles to target and react with specific classes of enzymes sharing
a similar catalytic activity.

The basic unit of ABPP is a probe that typically consists of two elements:
(1) a reactive group by which the probe selectively and covalently cross-links to
residues in the active site of an active enzyme. (2) a secondary tag that may be
either a reporter group such as fluorophore, biotin, or a latent analytical handle
such as alkyne or azide that can undergo derivatization via click chemistry to
visualize and purify target proteins. These groups are incorporated into an active
molecule such that it mimics the natural compound. Therefore, ABPP involves
the synthesis of small-sized probes that can covalently label active site residues
of target proteins, as shown in figure 3. Subsequently, the labeled proteins are
visualized and purified using the secondary reporter tag, commonly alkyne or
azide tags attached on the probe are reacted with complementary “clickable” fluorophores or biotin bearing azide or alkyne, respectively, forming a 1,4-disubstituted triazole product which covalently introduces these tags on to the probe-labeled protein facilitating analysis. This is performed by gel fluorescence scanning for fluorescent probes or by purification on a streptavidin column for biotinylated probes, which is followed by mass spectrometry analysis for identification of the probe-labeled enzymes.19

Today, ABPP is being used in multiple biological experiments including the discovery of protein targets using comparative ABPP, as well as inhibitor identification and enzyme active site characterization. ABPP has many advantages over other conventional methods for target discovery as it can account for the post-translational mechanisms that regulate enzyme activity and labels even low-abundance proteins. The inhibitor discovery involves competitive ABPP to identify inhibitors which are able to block probe labeling of enzymes. The advantage of this method is that it does not require recombinant expression and purification as the enzymes are tested within the proteome. Also, ABPP is a powerful tool to assign an uncharacterized enzymes to a specific mechanistic class and recognize new enzyme active sites.19
Figure 3. Schematic representation of Activity Based Protein Profiling (ABPP) using probes bearing photoaffinity tags
Photo affinity labeling (PAL) involves a specific covalent binding between a light-sensitive chemical tag and the active site of a receptor protein molecule upon irradiation, in a specifically and reversibly bound state.\textsuperscript{22} This enables specific cross-linking of the ligand to the specific protein targets to which it preferentially binds. This identification of drug-receptor pairs helps to design drug molecules that show selective recognition of specific targets associated with a specific pathology. A photo reactive group should be stable to ambient light, and its excited state should have a lifetime long enough to abstract a hydrogen on carbon atom to form a cross link with the target protein. Also, its UV excitation wavelength should be higher than the absorption range of most proteins. Commonly used groups are shown in Figure 4. Phenyl azides are usually the first choice to use as they are easy to prepare. Diazirines are very stable before UV irradiation and need wavelengths higher than that for proteins, but require multistep synthesis to form the strained 3-membered heterocyclic rings. The benzophenone probe requires prolonged irradiation resulting in cell death, but the excited carbonyl diradical is less reactive towards water than a nitrene or

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{photo_reactive_groups.png}
\caption{Examples of photo-reactive groups}
\end{figure}
Therefore, benzophenone is being used as a photoprobe for the synthesis of my analogues since it is easily activated with long wavelength UV light, and is stable during the chemical reactions used for the probe synthesis. It also has been found to covalently modify a single protein target.\textsuperscript{23}

Photoactivable lipid analogues have been used earlier to determine structure of membrane-bound proteins.\textsuperscript{24} Studies published by Li \textit{et al.}, report the synthesis of a benzophenone photoactivable analogue of acyl-LPA, with an oleoyl (ester) chain.\textsuperscript{22} Previous structure-activity studies have found that the oleoyl chain most closely represents the fatty acyl chain of naturally occurring LPAs.\textsuperscript{25} This probe was found to covalently modify protein targets in rat plasma.\textsuperscript{22}
Chapter II: Research Design

The biological properties of LPA and LPC have been studied widely but their roles and activity in many pathophysiological conditions are not yet completely understood. Therefore, we are designing photoreactive analogs of LPA and LPC in order to track and elucidate their activities in the biological systems. This research project targets the development of bifunctional LPA and LPC analog probes, as shown in figure 5. These probe designs consist of LPA and LPC analogs. A Y-shaped lysine linker bearing both a photoaffinity benzophenone group as well as a clickable reactive azide or alkyne reporter tag, is used as a secondary handle to append a fluorophore or biotin tag. These probes will allow for selective labeling of target proteins, followed by visualization and purification on a streptavidin column for the selective enrichment and identification of crosslinked proteins.

Figure 5. Target LPA & LPC activity probes
As shown in figure 6, the head group synthesis started with the protection of the free alcohol of (S)-glycerolacetonide 2 with a \( p \)-methoxybenzyl protecting group to 3, followed by acetal deprotection to produce diol 4. Next, the primary alcohol of 4 was selectively protected by monomethoxytrityl chloride (MMTrCl) to 5, followed by secondary alcohol protection using \( tert \)-butyldiphenylsilyl chloride to yield 6. Next, the MMTr group was deprotected using camphorsulfonic acid producing 7 with a free alcohol at the \( sn \)-1 position.

The synthesis of the lipid tail containing the benzophenone cross-linking group and azide tag (Figure 7) began with the coupling of 4-benzoylbenzoic acid to a Y-shaped protected lysine linker 8 to produce 9. The Boc protecting group was removed using trifluoroacetic acid to produce the free amine of 10, and in order to keep the synthetic strategy modular (figure 8), 10 was converted into both an azide 12a and alkyne 12b, the latter by reacting with alkyne acid 11. After the azide/alkyne methyl esters 12a and 12b were purified, they were hydrolyzed to yield acids 13a and 13b, respectively.

**Coupling of Head group and Lipid tail**

Once synthesized, we attempted to couple head group 7 and lipid tail 13a/b to produce 21, as shown in Figure 9. The coupling of alcohol 7 with the azide and alkyne acids 13a and 13b was attempted several times using different coupling reagents, including \( N,N \)-dicyclohexylcarbodiimide (DCC), 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) and thionyl chloride (\( SOCl_2 \)), but the reaction never gave very good yields, and the product was difficult to characterize. It may be attributed to the steric hindrance at the alcohol reaction
center, caused by the two bulky protecting groups. Therefore, to get around this problem, another synthetic strategy was followed in which we attempted to couple the acid directly with the diol, and the mixture of 1 and 2 substituted products would be separated to obtain the latter, and this would be followed by the protection of the secondary alcohol. Therefore, a set of trial reactions were run in order to find out the best reaction conditions for the coupling of the diol with lauric acid as it is similar to the lipid tail, as shown in figure 10. The best results were obtained with EDCI and DCC coupling reagents as they yielded the highest yields of the desirable regioisomeric ester, and the two regioisomers were conveniently separated. Next, the diol and the lysine acid were coupled under these reaction conditions, as shown in figure 11, but again the yields were not very good.

Once they have been coupled, the $p$-methoxybenzyl group of 21a,b would then be deprotected with DDQ to yield the free alcohol 14a,b (Figure 12). This would be followed by the installation of protected phosphoramidites 19 or 20, as shown in figure 10 and 11. Finally the dibenzyl and ethylcyano groups of 17a/b and 15 a/b respectively, will be deprotected producing the free phosphate, followed by the deprotection of TBDPS group to produce the target LPA activity probes 18a/b and LPC activity probes 16a/b, that can be utilized to identify and characterize new protein targets.

Once the bifunctional probes have been synthesized, they will first be tested for their effective activity as LPA and LPC mimics by the means of TLC (exploiting the differential adsorption of LPA and LPC on a TLC plate), MS
studies and fluorescence resonance energy transfer (FRET) through our ongoing collaborations with Dr. Daniel L. Baker and Dr. Abby L. Parrill at the University of Memphis. After it is validated as an effective mimic, it will be then be subjected to ABPP and other analytical studies, again in collaboration, where the incubation of the proteome with our synthesized probes will be followed by UV irradiation and fluorophore or biotin–labeling of the probe–bound proteins. Next, this mixture of proteins will be separated using electrophoresis (1D or 2D polyacrylamide gel electrophoresis PAGE), and the labeled proteins will be detected by fluorescence scanning of the gels. Proteins labeled with biotinylated probes will be separated by affinity chromatography using a streptavidin column. Though "gel-free" methods have emerged, gel-based methods are still widely used because of their robustness and efficiency. Liquid chromatography-mass spectrometry (LC–MS) methods provide better resolution than the gel–based methods, and is classified into two types. The first type involves the analysis of protein targets of probes, and the second one specifically analyses probe-labeled peptides derived from these targets. The first method is a combination of ABPP and MudPIT (multidimensional protein identification technology), in which biotinylated probe-labeled enzymes are enriched on streptavidin column, followed by enzymatic digestion (with trypsin) of the protein into peptide fragments that are then analyzed by LC-MS using triple-quadruple and ion trap mass spectrometers. The disadvantage of this method is that it does not offer a straightforward way to identify probe labeled peptides of enzyme targets. Therefore, active–site peptide profiling (ASPP) is used, which reverses the order of enzymatic digestion
and incubation, followed by the use of SEQUEST algorithm to identify protein receptors and sites of probe labeling. Speers and Cravatt developed tandem orthogonal proteolysis–ABPP (TOP–ABPP), which combined the advantages of ABPP–MudPIT and ASPP to complement each other. In TOP, the probe-labeled proteins are clicked with a biotin tag bearing a tobacco etch virus protease (TEV) cleavage site, and then enriched. Following the on-bead tryptic digestion of the tagged proteins, the probe-labeled peptides are eluted from the beads by incubation with TEV. Then sequential MudPIT experiments are conducted to analyze the trypsin and TEV digests, for characterizing the probe-labeled proteins and sites of probe labeling. 26 Though LC-MS offers high resolution, but requires larger sample volume each time and is slower, and thus not a desirable method to analyze hundreds of samples. 19

These analytical methods will help us to identify the inhibitors of LPA production and action.
Figure 6. Synthesis of the head group
Figure 7. Synthesis of the lipid tail
Figure 8. Modular synthesis of the alkyne acid
Figure 9. Synthetic strategy for introducing the bifunctional tag
Figure 10. Alternative methods attempted for direct coupling of lipid tails

<table>
<thead>
<tr>
<th>Coupling Reagent</th>
<th>Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCC, DMAP</td>
<td>% a 30</td>
</tr>
<tr>
<td>EDCI, NMM, DMAP</td>
<td>% b 40 % c 10</td>
</tr>
<tr>
<td>HBTU, HOBT, DIEA</td>
<td>(one spot, didn't stain in KMnO₄)</td>
</tr>
<tr>
<td>SOCl₂, CH₂Cl₂</td>
<td>% a 40 % c 5</td>
</tr>
<tr>
<td>NHS, EDCI, CH₂Cl₂</td>
<td>(single spot, only tail)</td>
</tr>
<tr>
<td>CDI, DMAP, DMF</td>
<td></td>
</tr>
</tbody>
</table>

25
Figure 11. Attempt to couple bifunctional tags directly onto diol 4
Figure 12. Modular synthetic strategy of the LPA probe
Figure 13. Modular synthetic strategy of the LPC probe
Experimental Procedure

(S)-4-(((4-methoxybenzyl)oxy)methyl)-2,2-dimethyl-1,3-dioxolane (3)

Sodium hydride (518 mg, 12.1 mmol) was dissolved in 40 mL of N,N-dimethylformamide at 0 °C under Nitrogen gas. S-Glycerol acetonide 2 (1 mL, 8.1 mmol) was then dissolved in 40 mL of N,N-dimethylformamide and slowly added to the Sodium hydride solution over 10 min. This mixture was then stirred at 0 °C for 1 hour, after which 4-methoxybenzyl chloride (1.64 mL, 12.1 mmol) was added at room temperature and stirred overnight at room temperature. This was quenched by water (50 mL) the next day. This was followed by extraction with dichloromethane and saturated sodium chloride (2×50 mL), the organic layer was collected and dried using magnesium sulphate, then filtered and the solvent was removed. Column chromatography on silica gel and a gradient solvent system of 10-50% ethyl acetate/hexanes gave the product as a clear oil (1.23 g, 61%).

$^1$H NMR (300 MHz, CDCl$_3$) δ 7.33 – 7.21 (m, 4H), 6.95 – 6.83 (m, 2H), 4.59 – 4.43 (m, 2H), 4.27 (dd, $J = 12.1, 5.9$ Hz, 1H), 4.05 (dd, $J = 8.2, 6.4$ Hz, 1H), 3.81 (s, 3H), 3.73 (dd, $J = 8.2, 6.4$ Hz, 1H), 3.48 (ddd, $J = 27.0, 9.8, 5.6$ Hz, 2H), 1.61 – 1.51 (m, 4H), 1.39 (d, $J = 17.8$ Hz, 5H), 0.23 – -0.25 (m, 38H).

NMR data matched that in the literature.$^{27}$
(R)-3-((4-methoxybenzyl)oxy)propane-1,2-diol (4)

Acetal 3 (1.63 g, 6.5 mmol) was dissolved in methanol/dichloromethane (1:1, 40 mL) and p-toluenesulphonic acid (570 mg, 3 mmol) was added, then it was allowed to stir at room temperature for 24 hours. The reaction mixture was then quenched with sodium bicarbonate (505 mg, 6 mmol) followed by extraction with water and dichloromethane (2x50mL), the organic layer was collected and dried using MgSO₄, then filtered, and the solvent was removed. Column chromatography on silica gel and a gradient solvent system of 50% ethyl acetate/hexanes to 5% methanol/dichloromethane gave the diol 4 (682 mg, 50%).

¹H NMR (300 MHz, CDCl₃) δ 7.31–7.21 (m, 1H), 6.90 (dd, J = 9.0, 2.3 Hz, 1H), 4.50 (s, 1H), 3.96–3.84 (m, 1H), 3.82 (d, J = 0.5 Hz, 1H), 3.77–3.72 (m, 1H), 3.72–3.67 (m, 1H), 3.63 (dd, J = 11.0, 5.0 Hz, 1H), 3.55 (dd, J = 5.8, 5.1 Hz, 1H), 3.51 (d, J = 6.1 Hz, 1H).

(S)-1-((4-methoxybenzyl)oxy)-3-((4-methoxyphenyl)diphenylmethoxy)propan-2-ol (5)
The \( p \)-methoxybenzyl protected diol 4 (305 mg, 1.4 mmol) was dissolved in pyridine (20 mL) under Nitrogen, followed by the addition of monomethoxytrityl chloride (889.4 mg, 2.88 mmol), 4-dimethylaminopyridine (17.59 mg, 0.14 mmol) and some molecular sieves. The reaction mixture was then stirred overnight at room temperature. Next day, it was quenched by methanol (1 mL) and then 1 drop of water added and stirred, followed by the removal of solvent. Column chromatography on silica gel and a gradient solvent system of 20-35% ethyl acetate/hexanes gave the product 5 (524.4 mg, 75%).

\(^1\)H NMR (300 MHz, CDCl\(_3\)) \( \delta \) 7.39 (dt, \( J = 9.1, 8.0 \) Hz, 4H), 7.34 – 7.12 (m, 11H), 6.93 – 6.73 (m, 4H), 4.60 (s, 1H), 4.46 (s, 2H), 3.98 (dt, \( J = 10.3, 4.2 \) Hz, 1H), 3.82 – 3.73 (m, 6H), 3.62 – 3.45 (m, 2H), 3.20 (pd, \( J = 9.4, 5.8 \) Hz, 2H), 2.47 (d, \( J = 4.8 \) Hz, 1H), 1.64 (s, 1H).

NMR data matched that in literature.

\((S)\)-tert-butyldiphenyl((1-(\((4\)-methoxybenzyl\)oxy)-3-\((4\)-methoxyphenyl\)diphenylmethoxy)propan-2-yl\)oxy)diphenylsilane (6)

Alcohol 5 (204 mg, 0.42 mg) was dissolved in pyridine (5 mL) under nitrogen, followed by the addition of imidazole (57.3 mg, 0.84 mmol), \( t \)-butyldiphenylchlorosilane (0.218 mL, 0.84 mmol) and molecular sieves. This reaction mixture was then stirred at room temperature overnight. Once the reaction was complete, the molecular sieves were removed by filtering the
mixture over celite. The mixture was then quenched with 10 mL of methanol and then a few drops of water added. This was followed by removal of solvent by rotary evaporation, and column chromatography on silica gel and a gradient solvent system of 3-10% ethyl acetate/hexanes gave product 6 (302 mg, 99%).

$^1$H NMR (300 MHz, CDCl$_3$) δ 7.95 (d, $J = 8.2$ Hz, 2H), 7.90–7.76 (m, 4H), 7.63 (t, $J = 7.4$ Hz, 1H), 7.51 (t, $J = 7.5$ Hz, 2H), 6.88 (s, 1H), 4.83 (d, $J = 5.4$ Hz, 1H), 4.60 (s, 1H), 3.81 (s, 3H), 3.14 (d, $J = 6.4$ Hz, 2H), 2.04–1.93 (m, 1H), 1.84 (s, 1H), 1.58 (dd, $J = 3.7$, 3.0 Hz, 1H), 1.56–1.47 (m, 2H), 1.41 (s, 11H).

NMR data matched that in literature.$^{27}$

![Chemical structure](image)

**(R)-2-((tert-butyldiphenylsilyl)oxy)-3-((4-methoxybenzyl)oxy)propan-1-ol (7)**

Compound 6 (296 mg, 0.41 mmol) was dissolved in methanol:dichloromethane (15 mL, 2:1), followed by the addition of camphorsulfonic acid (47.55 mg, 0.20 mmol). The reaction mixture was then stirred at room temperature for 3.5 hours, followed by its neutralization by triethylamine and concentration. Column chromatography on silica gel and a gradient solvent system of 20–35% ethyl acetate/hexanes gave the product 7 (99 mg, 51%).

$^1$H NMR (300 MHz, CDCl$_3$) δ 7.72–7.59 (m, 4H), 7.50–7.29 (m, 6H), 7.18–7.05 (m, 2H), 6.89–6.75 (m, 2H), 4.37–4.19 (m, 2H), 3.99–3.84 (m, 1H), 3.79 (s, 3H), 3.64 (d, $J = 4.3$ Hz, 2H), 3.57–3.34 (m, 2H), 1.07 (d, $J = 2.9$ Hz, 9H).
(S)-methyl-2-(4-benzoylbenzamido)-6-((tert-butoxycarbonyl)amino)hexanoate (9)

Cbz protected lysine 8 (600 mg, 1.52 mmols) was dissolved in methanol (15 mL), followed by the addition of 10% palladium hydroxide (60 mg) by weight. The mixture was then stirred for 24 hours at room temperature under hydrogen, followed by filtration over celite. Then the filtrate was concentrated, and the residue (386.1 mg, 97%) was dissolved in N,N-dimethylformamide (10 mL) under nitrogen, followed by the addition of 4-benzoylbenzoic acid (186 mg, 0.82 mmol), EDCI (189 mg, 0.99 mmol) and 4-dimethylaminopyridine (120 mg, 0.99 mmol). The reaction mixture was stirred at room temperature for 5 hours, and then concentrated. Column chromatography on silica gel and a gradient solvent system of 50% ethyl acetate/hexanes gave the product 9 (240 mg, 60%).

$^1$H NMR (300 MHz, CDCl$_3$) δ 7.95 (d, $J$ = 8.2 Hz, 2H), 7.90–7.76 (m, 4H), 7.63 (t, $J$ = 7.4 Hz, 1H), 7.51 (t, $J$ = 7.5 Hz, 2H), 6.88 (s, 1H), 4.83 (d, $J$ = 5.4 Hz, 1H), 4.60 (s, 1H), 3.81 (s, 3H), 3.14 (d, $J$ = 6.4 Hz, 2H), 2.04–1.93 (m, 1H), 1.84 (s, 1H), 1.58 (dd, $J$ = 3.7, 3.0 Hz, 1H), 1.56–1.47 (m, 2H), 1.41 (s, 11H).
(S)-methyl 6-azido-2-(4-benzoylbenzamido)hexanoate (12a)

Boc-protected amine 9 (30.4 mg, 0.06 mmols) was dissolved in dichloromethane (4 mL), followed by the addition of trifluoroacetic acid (4 mL), and the reaction mixture was stirred at room temperature for 2 hours. Then it was concentrated, and the residue put under high vacuum overnight. The next day, the residue was dissolved in methanol (10 mL), followed by the addition of imidazole-1-sulfonyl azide hydrochloride (16.35 mg, 0.08 mmols), 0.804 M potassium carbonate solution (0.14 mL, 0.11 mmol) and 0.01 M copper (II) sulfate pentahydrate solution (0.65 mg, 0.006 mmols). The reaction mixture was stirred at room temperature for 3 hours, followed by extraction with ethyl acetate and 2M HCl. The organic phase was collected and dried with MgSO₄ followed by concentration. Column chromatography on silica gel and a gradient solvent system of 50% ethyl acetate/hexanes gave the product 12a (21 mg, 80%).

¹H NMR (300 MHz, CDCl₃) δ 7.91 (dd, J = 14.1, 6.6 Hz, 3H), 7.88–7.75 (m, 4H), 7.68–7.57 (m, 1H), 7.57–7.43 (m, 2H), 6.88 (d, J = 7.7 Hz, 1H), 4.87 (dd, J = 12.9, 7.4 Hz, 1H), 3.82 (s, 3H), 3.37–3.19 (m, 2H), 2.1 –1.95 (m, 1H), 1.86 (ddd, J = 20.8, 11.7, 6.4 Hz, 1H), 1.74–1.58 (m, 2H), 1.58–1.38 (m, 2H).
NMR data matched that in literature.\textsuperscript{28}

\chem{(S)-6-azido-2-(4-benzoylbenzamido)hexanoic acid (13a)}

Azide 12\textsubscript{b} (355 mg, 0.9 mmol) was dissolved in methanol (4 mL), followed by the addition of 2M NaOH (4 mL) was added, and then the reaction mixture stirred at room temperature for 30 minutes. Then the reaction was extracted with 2M HCl, and the organic layer collected and dried with magnesium sulfate and concentrated to give the product 13\textsubscript{a} (308 mg, 90%).

\begin{align*}
^{1}H \text{ NMR (300 MHz, CDCl}_3) &\delta 7.89 (dd, J = 19.8, 8.5 \text{ Hz}, 3\text{H}), 7.79 (dd, J = 5.3, 3.2 \text{ Hz}, 1\text{H}), 7.68–7.57 (m, 1\text{H}), 7.50 (t, J = 7.5 \text{ Hz}, 1\text{H}), 6.86 (d, J = 7.4 \text{ Hz}, 1\text{H}), 4.88 (dd, J = 12.9, 7.1 \text{ Hz}, 1\text{H}), 3.32 (t, J = 6.5 \text{ Hz}, 1\text{H}), 2.07 (s, 1\text{H}), 1.92 (d, J = 7.4 \text{ Hz}, 1\text{H}), 1.67 (dt, J = 12.2, 6.0 \text{ Hz}, 2\text{H}), 1.55 (d, J = 4.1 \text{ Hz}, 2\text{H}), 1.25 (s, 4\text{H}), 0.87 (d, J = 7.0 \text{ Hz}, 1\text{H}).
\end{align*}

NMR data matched that in literature.\textsuperscript{28}
Acid 13a (57 mg, 0.15 mmol) was dissolved in dichloromethane (5 mL) under nitrogen, followed by the addition of a few molecular sieves, 4-dimethylaminopyridine (25.5 mg, 0.20 mmol) and N,N-dicyclohexylcarbodiimide (43 mg, 0.20 mmol). This reaction mixture was stirred at room temperature for 30 minutes. Next, the alcohol 7 (70 mg, 0.15 mmol) was added and the reaction mixture was stirred for 24 hours. Then it was filtered over celite and concentrated. Column chromatography on silica gel and a gradient solvent system of 20-35% ethyl acetate/hexanes gave the product 21 (17 mg, 14%).

$^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 7.99 – 7.89 (m, 1H), 7.90 – 7.76 (m, 4H), 7.75 (s, 1H), 7.67 (d, $J = 8.2$ Hz, 5H), 7.53 (t, $J = 7.6$ Hz, 3H), 7.39 (d, $J = 7.4$ Hz, 9H), 7.11 (t, $J = 11.3$ Hz, 2H), 6.85 (d, $J = 8.4$ Hz, 2H), 4.70 (s, 1H), 4.48 – 4.40 (m, 1H), 4.39 – 4.22 (m, 3H), 4.08 (s, 1H), 3.80 (d, $J = 9.8$ Hz, 20H), 3.66 – 3.57 (m, 1H), 3.57 – 3.43 (m, 2H), 3.39 (s, 4H), 3.26 (s, 2H), 3.18 – 3.04 (m, 1H), 2.21 (s, 1H), 2.07 (s, 1H), 1.97 – 1.82 (m, 1H), 1.75 (s, 1H), 1.61 (s, 3H), 1.44 (s, 2H), 1.27 (s, 5H), 1.12 – 0.96 (m, 9H), 0.88 (d, $J = 7.6$ Hz, 2H), 0.18 – 0.19 (m, 64H).
Acid 13b (30 mg, 0.06 mmol) was dissolved in dichloromethane (5 mL) under nitrogen, followed by the addition of a few molecular sieves, 4-dimethylaminopyridine (10.4 mg, 0.08 mmol) and N,N-dicyclohexylcarbodiimide (23 mg, 0.12 mmols). This reaction mixture was stirred at room temperature for 30 minutes. Next, the alcohol 7 (14 mg, 0.07 mmols) was added and the reaction mixture was stirred for 24 hours. Then it was filtered over celite and concentrated. Column chromatography on silica gel and a gradient solvent system of 75% ethyl acetate/hexanes-10%methanol/dichloromethane gave the product 21b (5 mg, 12%).

$^1$H NMR (300 MHz, CDCl$_3$) δ 8.00 (d, $J = 7.9$ Hz, 2H), 7.93 – 7.72 (m, 4H), 7.63 (t, $J = 7.4$ Hz, 1H), 7.50 (t, $J = 7.6$ Hz, 2H), 7.35 – 7.17 (m, 5H), 6.97 – 6.81 (m, 2H), 6.30 (s, 1H), 6.26 – 6.13 (m, 1H), 4.75 (s, 1H), 4.49 (s, 2H), 4.43 – 4.16 (m, 1H), 4.11 (s, 1H), 3.97 (dt, $J = 5.3$, 2.6 Hz, 2H), 3.84 – 3.74 (m, 3H), 3.67 (s, 1H), 3.54 (dd, $J = 8.0$, 5.0 Hz, 2H), 3.26 (s, 2H), 2.44 (d, $J = 4.8$ Hz, 4H), 2.19 (t, $J = 2.5$ Hz, 1H), 2.00 (d, $J = 26.1$ Hz, 3H), 1.65 (s, 6H), 1.57 – 1.32 (m, 6H), 1.32 – 1.11 (m, 6H), 0.87 (d, $J = 7.0$ Hz, 2H).
Propargylamine (1.5 ml, 21.87 mmols) was dissolved in 10 ml of N,N-dimethylformamide/acetonitrile (1:1) at 0°C under Nitrogen. Next, succinic anhydride dissolved in 10 ml of acetonitrile was added to this solution, and stirred overnight at room temperature. This was followed by removal of the solvent, and the residue was then washed with 40 ml of hexane to obtain a yellow brown solid. Column chromatography on silica gel and a gradient solvent system of 20-50% ethyl acetate/hexanes/0.1% acetic acid-10% methanol/dichloromethane/0.1% acetic acid gave the product 11 (2502 mg, 38%).

$^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 3.93 (t, $J = 2.3$ Hz, 2H), 3.30 (dt, $J = 3.3$, 1.6 Hz, 8H), 2.64 – 2.52 (m, 3H), 2.52 – 2.41 (m, 2H).
List of References


15. Valet, P.; Pages, C.; Jeanneton, O.; Daviaud, D.; Barbe, P.; Record, M.; Saulnier-Blache, J. S.; Lafontan, M., Alpha(2)-adrenergic receptor-mediated release of


Appendix
Vita

Ritu Nandal was born in Haryana, India in 1983. She is the daughter of Jagwanti and Satpal Hooda. She completed her schooling from S.M Public School in 2000. After completing her undergraduate from A.I.J.H.M College in 2004, she attended M.D University and completed her Master of Science degree in 2007.