7-2017

Generating Enzyme and Radical-Mediated Bisubstrates as Tools for Investigating Gcn5-Related N-Acetyltransferases

Cory T. Reidl
Loyola University Chicago, reidlchem@gmail.com

Karolina A. Majorek
University of Virginia

Joseph Dang
San Francisco State University

David Tran
San Francisco State University

Kristen Jew
San Francisco State University

See next page for additional authors

Recommended Citation

Reidl, Cory T.; Majorek, Karolina A.; Dang, Joseph; Tran, David; Jew, Kristen; Law, Melissa; Payne, Yasmine; Minor, Wladek; Becker, Daniel P.; and Kuhn, Misty L.. Generating Enzyme and Radical-Mediated Bisubstrates as Tools for Investigating Gcn5-Related N-Acetyltransferases. FEBS Letters, 591, 15: 2348-2361, 2017. Retrieved from Loyola eCommons, Chemistry: Faculty Publications and Other Works, http://dx.doi.org/10.1002/1873-3468.12753

Generating enzyme and radical-mediated bisubstrates as tools for investigating Gcn5-related *N*-acetyltransferases

Cory Reidl[2], Karolina A Majorek[3], Joseph Dang[1], David Tran[1], Kristen Jew[1], Melissa Law[1], Yasmine Payne[2], Wladek Minor[3], Daniel P. Becker[2], and Misty L. Kuhn[1*]

[1]Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, CA 94132, USA

[2]Loyola University Chicago, Department of Chemistry, 1032 W. Sheridan Rd., Chicago, IL 60660, USA

[3]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908, USA

[*]To whom correspondence may be addressed: Either Department of Chemistry and Biochemistry, San Francisco State University, 1600 Holloway Ave., San Francisco, CA 94132. Tel.: 415-405-2112; E-mail: mkuhn@sfsu.edu or Department of Chemistry and Biochemistry, Loyola University Chicago, 1032 W. Sheridan Rd., Chicago, IL 60660, Tel.: 773-508-3089; E-mail: dbecke3@luc.edu.

Running title: GNAT catalyzed and radical-mediated bisubstrates

Keywords: Gcn5-related N-acetyltransferase, GNAT, PA4794, bisubstrate inhibitors, radical-mediated bisubstrate, product-based transition-state modeling (PBTSM), acyltransferase, crystal structure, enzyme kinetics, Pseudomonas aeruginosa, peptide interaction

Abstract

Gcn5-related N-acetyltransferases (GNATs) are found in all kingdoms of life and catalyze important acyl transfer reactions in diverse cellular processes. Although many of their structures and/or functions are known, a significant number of GNATs remain functionally and structurally uncharacterized. Even when the 3D structures of these proteins have been determined, most do not contain acceptor substrates bound in their active sites. The underlying challenge with obtaining 3D structures of GNATs with functionally relevant ligands bound in their active sites involves discerning which ligands are potentially functionally relevant for a protein of unknown function, and coaxing the protein to bind and retain the ligand within the crystal for structure determination. To expand upon existing crystallographic strategies for improving ligand bound GNAT structures, we have synthesized peptide substrate analogs and reacted them with CoA in PA4794 protein crystals to form bisubstrates. We found two separate mechanisms for bisubstrate formation: 1) direct alkylation of CoA with a halogenated peptide, and 2) X-ray induced radical-mediated alkylation of CoA with an alkene peptide. Additionally, we present a computational approach for identifying potential key residues over the course of the enzymatic reaction using these structures in complex with bisubstrates. We term this method product-based transition-state modeling (PBTSM), as docking and molecular dynamics simulations of these interactions occur in the reverse direction, i.e. from products to substrates. Our approach is widely applicable across the GNAT superfamily and can be used to improve the success rate of obtaining liganded structures of other GNATs, including histone acetyltransferases.

Introduction

The three-dimensional structures of proteins with ligands bound in their active sites are valuable tools because they provide a framework for determining residues that bind ligands over the course of an enzymatic reaction.(1) In the present age where new sequencing information greatly exceeds the experimental functional annotation of genes, liganded structures yield important information that can be used to more accurately computationally predict protein function or modes of ligand binding. Using structural information to enhance our understanding of acetyltransferases is critically important, since transferring an acetyl group from one molecule to another is a fundamental biochemical process. Members of the very large Gcn5-related N-acetyltransferase (GNAT) superfamily are found in all kingdoms of life and play key roles in metabolic and cellular processes, including gene regulation, transcription, post-translational protein modification, detoxification and drug resistance.(2, 3) These proteins use an activated acyl-donor, typically acetyl-coenzyme A (AcCoA), to transfer an acyl group to an acceptor substrate. The identity of the acceptor substrate can range from small molecules, such as polyamines, antibiotics, or arylalkylamines to peptides and proteins.(2, 4, 5) Although many proteins belonging to this family have been investigated, many members remain uncharacterized and would benefit from additional structural and functional investigation.

GNATs have two adjacent but distinct substrate-binding sites: [1] the acyl donor substrate binding site, and [2] the acceptor substrate binding site. Hundreds of 3D structures of GNATs have been determined and deposited into the Protein Data Bank (PDB). While many of these structures have AcCoA/CoA bound in the acyl donor substrate binding site, it is less common to have ligands bound in the acceptor substrate binding site. Furthermore, the acyl donor site is highly conserved, making protein-CoA binding interactions in a newly-sequenced GNAT easier to predict. On the other hand, the acceptor substrate binding site, which is responsible for specificity, is quite diverse among GNATs and more difficult to predict ligand binding modes. Obtaining structures of GNATs with bound acceptor substrates is therefore crucial for understanding these enzymes. This has proven to be challenging, for even if the function of the protein is known or potential acceptor substrates have been identified the enzymes often do not bind the ligand in the crystal. Thus, new approaches and tools are needed to efficiently yield structures with acceptor ligands bound so we can more effectively understand key substrate recognition elements for catalysis across this diverse family of enzymes.

One strategy for obtaining liganded structures has been to crystallize proteins in the presence of bisubstrate inhibitors.(6, 7) These types of inhibitors have also been employed in medicinal chemistry in the design of more potent inhibitors, notably in the area of acetyltransferases and kinases to encompass both substrate binding pockets.(8, 9) The binding affinity of an enzyme for a bisubstrate can be significantly greater than that of the substrate alone; therefore, it is expected the enzyme will more efficiently bind the ligand in the crystal. Many GNAT structures have been determined in the presence of bisubstrates where the bisubstrate molecule is either synthesized by the enzyme itself or chemically prior to crystallization.(2, 4, 5, 10-15)

In our previous work, we used a broad-substrate screening assay as a method to identify acceptor substrates for an uncharacterized PA4794 GNAT enzyme from *Pseudomonas aeruginosa* and then structurally and kinetically characterized it in the presence of a dipeptide substrate.(16) We found that PA4794 acetylates the dipeptide N-phenylacetyl glycine lysine (NPAcGK) and determined the structure of the ternary complex with CoA and acetylated NPAcGK (PDB ID: 4L8A) (16). Here, we have used the PA4794 enzyme as a model enzyme from the GNAT family to develop a streamlined procedure for increasing liganded crystal structures of GNATs in the presence of bisubstrates. Our approach was to synthesize analogs of the PA4794 substrate and have the enzyme catalyze the formation of the bisubstrate in the presence of CoA. The advantages to this approach are: [1] the synthesis and purification procedures for acceptor ligands as opposed to full bisubstrates are much simpler, [2] it reduces the quantity of costly CoA needed for formation of the bisubstrate, [3] many different acceptor ligands can be tested using the same crystals, and [4] this approach can be used regardless of the size of the substrate to be bound. This final advantage is important because the donor and acceptor substrate binding sites of many GNATs are connected by a narrow tunnel, which may not allow a fully formed large bisubstrate analog to thread through their active sites. Allowing each substrate to enter its respective binding site separately facilitates formation of a bisubstrate even with a large or bulky acceptor substrate during crystal soaks or co-crystallization.

In this work, we describe two separate three-dimensional structures of PA4794 with NPAcGK analog-derived bisubstrates formed in the active site through two different mechanisms with CoA. The first bisubstrate was formed through direct alkylation with a reactive substrate, and the second bisubstrate was formed through an X-ray induced radical-mediated addition to the alkene moiety of the substrate analog. Additionally, we present a new approach we refer to as product-bound transition-state modeling (PBTSM) to simulate the reaction pathway of PA4794 catalysis in reverse chronological order beginning with products and proceeding through the tetrahedral intermediate/transition state toward substrate binding to explore key interactions within the active site along the reaction trajectory. We then interrogated some of the key active site residues identified through the bound bisubstrate structures and PBTSM molecular dynamics analyses by preparing point mutations and testing enzymatic activity. The approach outlined herein is widely applicable across the GNAT superfamily, especially those that show a preference for lysine acetylation, and can be used to advance the knowledge of modes of ligand binding in the acceptor site. This will ideally lead to more accurate computational methods of functional prediction of GNATs and drug design.

Results

Design of substrate analogs—We previously found that PA4794 acetylates the ε-amino group of NPAcGK and prefers to acetylate a C-terminal lysine residue of peptides(16), yet its native function remains unknown. We chose this enzyme as a model system for exploring formation of bisubstrates in GNATs because it is small and readily crystallizes, which makes it an ideal protein to study mechanisms of ligand binding. To begin, we analyzed the 3D structure of the ternary complex of PA4794 and its interactions with CoA and acetylated NPAcGK (NPAcGAcK) peptide (PDB ID: 4L8A) and designed and synthesized reactive analogs of the acceptor substrate that would react with CoA to form bisubstrates in the protein crystal. We reasoned that electrophilic

moieties installed near the ε-amino group of the acceptor substrate should be capable of alkylating the thiol of free CoA to produce a bisubstrate in the active site of the protein. To mimic the acetyl group transferred from CoA to the acceptor but with a reactive moiety, we installed an α-haloacetyl group on the ε-nitrogen of the lysine. The α-haloacetyl moiety has been utilized extensively as a group to alkylate active site residues in proteins(17), and in a similar approach, Hwang incorporated an α-haloacetyl moiety as a replacement for the acetyl moiety of AcCoA, which was able to react with active site residues of acetyltransferases that use AcCoA(18). Our approach for synthesizing haloacetyl substrates is summarized in Scheme 1 and was accomplished by reacting NPAcGK with chloroacetyl chloride or bromoacetyl chloride to obtain compounds **2a** and **2b**, respectively. The PA4794 enzyme-mediated formation of bisubstrate **3** occurred in the presence of CoA with these compounds in the active site.

Bisubstrate formation in crystals via an enzyme-mediated mechanism—We grew crystals of the PA4794 protein and soaked them with CoA, followed by soaking with the **2a** and **2b** chloroacetyl and bromoacetyl analogs (see Experimental Procedures for more details) and screened crystals for ligand binding. We found both analogs reacted with CoA to form bisubstrate **3** (Scheme 1, Figure 1A) and exhibited identical binding modes in the structures (Figure 1; PDB IDs: 5VDB, 5VD6). Although the bromoacetyl derivative **2b** yielded the same bisubstrate as the chloroacetyl derivative **2a**, the resolution of the structure from bromide **2b** was lower, perhaps due to the greater reactivity of the bromo compound leading to less selective and more random reactivity, or simply from greater radiation damage to the crystal. Hence, only the structure derived from chloroacetyl derivative **2a** was refined (PDB ID: 5VDB). The bisubstrate was formed via an enzyme-mediated mechanism whereby the enzyme active site placed the electrophilic alkylating moiety in close proximity for S_N2 displacement of the halide by the thiol of CoA.

Bisubstrate formation in crystals via a radical-mediated mechanism—When we soaked crystals of PA4794 with CoA and the alcohol/alkene mixture (**4**/**5**) (Scheme 2), our assumption was that the alcohol would bind preferentially in the active site of the enzyme, given the alcohol's ability to serve as both H-bond donor and H-bond acceptor, and the comparable steric volume to the amine substrate. We were surprised to observe formation of bisubstrate **6** in the structure (Scheme 2, Figure 1B, PDB ID: 5VD6). This was curious because both alcohol and alkene were expected to be chemically unreactive. To determine the mechanism of bisubstrate formation between the alcohol/alkene analog and CoA, we performed a series of tests to determine if it was produced enzymatically from the alcohol **4** or alkene **5**. Treatment of the **4**/**5** mixture with the PA4794 enzyme and CoA did not give rise to any detectable bisubstrate **6** by HPLC. However, it is well known that a thiol can add to an alkene under free radical conditions(24), which could have been initiated during X-ray data collection, so we tested whether or not the **4**/**5** mixture could produce a bisubstrate in the presence of a source of radicals. Treatment of **4**/**5** and CoA with AIBN at 65°C for 2 hours showed complete consumption of the alkene **5** from the mixture and corresponding reduction in the concentration of CoA (see Supplemental Figure S16). There was no detectable effect on the concentration of alcohol **6**, and the production of a new polar peak close in retention time to CoA was shown to be bisubstrate **6** by mass spectrometry (M^- observed = 1070.2), providing strong evidence that the formation of the bisubstrate **6** was due to a radical-mediated mechanism from X-ray irradiation during data collection.

Two distinct substrate analog binding modes reveal functionally important interactions—The bisubstrate analogs formed in the present study through either an enzyme- or radical-mediated mechanism occupy both the AcCoA and acceptor substrate binding sites and enable exploration of key interactions in both pockets. The CoA moieties of bisubstrates **3** and **6** are bound in similar conformations and through similar interactions as those of previously determined PA4794 structures (PDB IDs: 3PGP, 4KUB, 4L8A, Figure 1). Bisubstrate **6** binding is similar to that of the acylated substrate observed in the ternary complex (PDB ID: 4L8A) structure with NPAcGAcK (Figure 1C). In particular, the most

critical charge-charge interactions of the C-terminal carboxyl group of NPAcGAcK occur with R49 and the main chain nitrogen of N80 as well as several water-mediated interactions. The phenylacetyl moiety lies against a hydrophobic patch composed of residues F118, F27 and P31. The moiety corresponding to the C-terminal lysine side chain of the bisubstrate **6** (Figure 1B) is shorter by four atoms and in an extended conformation compared to the acetylated lysine side chain of the product NPAcGAcK. The bisubstrate compound **3** (Figure 1A) exhibits a different binding mode compared to the acylated product shown in Figure 1C, but this can be attributed to the longer-atom chain connecting the lysine mimetic to CoA. The modified acetyl-lysine moiety in bisubstrate **3** is extended and is not as distorted as in the ternary complex structure. Specifically, this is due to a movement of the C-terminal carboxyl group of the peptide that now interacts with more surface exposed R141 instead of R49 (Figure 1A). R141 is located closer to the binding pocket opening than R49 and allows a more extended conformation of the substrate analog moiety, which is slightly less buried inside of the protein as compared to bisubstrate **6** (Figure 1B). In this case, additional interactions are observed between the ε-amino group and the acetyl oxygen of the haloacetyl substrate analog derivative and the main chain oxygen of S116 and main chain nitrogens of N80 and M81. The two distinct conformations of the two bisubstrate analog moieties also induce significantly different conformations of Y68.

Importance of R49 and R141 residues of PA4794—The binding mode of the two bisubstrates revealed two arginine residues (R49 and R141) that are implicated as playing a potentially important role in binding substrate and/or product. To elucidate whether these two residues were critical for activity, we constructed alanine and glutamine mutants and compared their kinetic activity to the wild-type enzyme. We could not completely remove the polyhistidine tag from all mutants; therefore, we screened the kinetic activity of all mutants and wild-type proteins in the presence of the tag for this experiment. We previously determined the R49Q mutant displayed reduced PA4794 catalytic activity, but included it here since the tag was not present previously. (16) All five mutants (R49A, R49Q, R141A, R141Q, and R49QR141Q) showed a significant decrease in activity (between 3.5- and 23-fold decrease) compared to the wild-type enzyme, with the most severe decrease seen for the double mutant (Figure 2). Replacing the more surface-exposed arginine R141 with glutamine (R141Q) exhibited diminished activity compared to the wild-type protein, but its replacement with alanine (R141A) showed a greater decrease in activity (Figure 2). R49 mutations also led to decreased activity relative to wild-type, but the glutamine and alanine mutants showed opposite effects compared to the R141 mutants: the R49A retained greater activity relative to R49Q. Regardless, it is clear that both of these residues are important for binding peptide, and at least one needs to be present for catalysis to occur.

Product-Based Transition-State Modeling Approach—To complement our enzyme kinetics results and elucidate potential roles of residues during catalysis, we used a new *in silico* modeling method we refer to as Product-Based Transition-State Modeling (PBTSM) (See Experimental Procedures for details) to elucidate hypothetical substrate and tetrahedral intermediate binding modes for PA4794. This method proceeds in reverse chronological reaction order beginning with reaction products and progressing back through the tetrahedral intermediate toward the original substrates. The process is summarized in Figure 3. Crystallographic data of enzymatic reaction products (PDB ID: 4L8A) or bisubstrate **3** were used as a starting point to systematically backtrack through the catalytic intermediate steps to computationally guide the placement of atoms that are likely to occur during formation of the tetrahedral intermediate and substrate binding.

Active site regions important along the reaction trajectory—Using the bisubstrate bound crystal structures and the PBTSM protocol, we identified a set of important regions of the active site that appear to stabilize the tetrahedral intermediate and enable the catalytic reaction to proceed. For instance, we identified a potential oxyanion hole composed of the backbone amide hydrogen of M81 near the alpha carbon hydrogen of C29 and the backbone amide hydrogen of N80 through a water molecule that could stabilize the negatively

charged oxygen of the tetrahedral intermediate (Figure 4A, blue circle). The water observed in the crystal structure is present at a nearly identical site as the one observed using the PBTSM protocol, which suggests it plays an important role in the acetyl transfer mechanism when the system is solvated. We also observed that H-bonds were formed directly through two backbone carbonyl oxygens of S116 and L78 or indirectly through a water molecule to stabilize the nucleophilic approach of the ε-amino group of NPAcGK during formation of the tetrahedral intermediate (Figure 4A, red circle). Two important arginine residues stabilize the C-terminal carboxyl group of the NPAcGK substrate (Figure 4A, black square). Binding of the carbonyl oxygen atom in the oxyanion hole allows the accompanying elements involved in formation of the tetrahedral transition state intermediate leading to acetyl transfer to be recognized. A shallow hydrophobic pocket containing L78, I115, Y128 and M81 residues was also observed to accommodate the methyl of the acetyl group. Finally, the model indicates that the sulfur of CoA forms a stabilizing H-bond interaction with Y128 and an ordered water molecule. This catalytic tyrosine has been predicted to play a role as a general acid capable of reprotonating CoA after transfer of the acetyl group to the acceptor substrate has occurred.(16) The ordered water makes further interactions with the backbone carbonyl of Y28, while H-bond donation varies between the backbone amide of F118 and the side chain amide of N121. The PBTSM-generated transition state models are shown in Figures 4B and 4C.

Kinetic mechanism of PA4794 toward NPAcGK—GNAT enzymes typically use either a sequential (direct transfer) mechanism whereby both donor and acceptor substrates bind to the enzyme simultaneously and the acetyl group from AcCoA is transferred directly to the substrate, or they use a ping-pong or double-displacement mechanism with an indirect transfer of the acetyl group via an acetyl-enzyme intermediate. In the latter mechanism, AcCoA binds first and transfers the acetyl group to a residue in the enzyme active site and then CoA leaves. Then the acceptor substrate binds and the acetyl group is transferred from the enzyme to the acceptor substrate. Since our PBTSM modeling experiments predicted a tetrahedral transition-state intermediate, which is indicative of a direct transfer mechanism, we performed enzyme kinetic assays and tested kinetic models to experimentally determine which kinetic mechanism was preferred for PA4794. Similar to other GNATs, we found the PA4794 protein indeed uses a bisubstrate steady-state kinetic mechanism(25, 26) (see Supplemental Materials Fig. S1), which infers an overall direct transfer of the acetyl group to the acceptor substrate through a transition state intermediate. Thus, our *in vitro* and *in silico* results are compatible.

Mechanistic hypothesis of overall PA4794 reaction sequence—Based on our combined *in vitro* and *in silico* results, we propose the sequence of substrate approach, acetyl transfer and final product release in PA4794 occurs as follows. Our kinetic mechanism results indicate there is no specific order of binding of AcCoA and NPAcGK, thus both substrates may approach the active site in a random order. When the NPAcGK substrate binds, its C-terminal carboxylate interacts with R141 (located near the binding pocket opening) or R49 (located closer to the catalytic center), and the lysine moiety approaches the catalytic site as a neutral species, having been deprotonated by bulk solvent or an unidentified basic residue near the active site. (16) The nucleophilic ε-amino group of NPAcGK is guided into the correct conformation by residues that stabilize the incipient positive ammonium species while the acetyl oxygen takes its place in the oxyanion hole. Immediately following formation of the tetrahedral intermediate, the bond between the carbonyl carbon and the sulfur breaks while the negative charge on sulfur is stabilized and ultimately protonated by Y128. Reprotonation of Y128 may come from bulk solvent or from deprotonation of the positively charged ammonium species. The acetylated NPAcGAcK product and CoA would then exit the active site and the catalytic cycle continues. Our mechanistic hypothesis for acetyl transfer by PA4794 is summarized in Figure 5.

Discussion

We have designed two distinct types of tools that utilize different mechanisms to form bound bisubstrates in GNAT protein crystals: [1] chloroacetyl/bromoacetyl derivatives of an

acceptor substrate using an enzyme-mediated mechanism and [2] addition of a thiol to an alkene derivative of acceptor substrate via a radical-mediated mechanism. We chose to use the PA4794 protein of unknown native function from *P. aeruginosa* as a model system to test our bisubstrate tools since its crystallization is highly reproducible, with fast growth of crystals that typically give high-resolution structural data. Upon soaking crystals of PA4794 with the chloroacetyl derivative of NPAcGK **2a** in the presence of CoA followed by structure determination, we observed that the protein enzymatically formed a covalent bisubstrate (Figure 1) occupying both AcCoA and acceptor substrate sites. Incubation of PA4794 crystal with bromoacetyl substrate **2b** yielded an identical bisubstrate-bound complex but of lower resolution.

The formation of bisubstrate **6** was unexpected, but can be attributed to radiation-induced free-radical-mediated addition of the thiol to the alkene moiety of **5**. We hypothesize that bisubstrate **6** was not produced enzymatically, but rather through exposure to free radicals produced by X-ray irradiation during data collection (Scheme 3). X-Ray induced chemical modification of specific moieties is well documented, and sulfur-containing residues are particularly susceptible.(27) Scheme 3 presents a mechanistic hypothesis for the formation of bisubstrate **6** in the active site mediated by hydroxyl radicals produced from the interaction of radiation with ubiquitous water molecules in and around the active site. H-atom abstraction from CoA by a hydroxyl radical would yield the thiol radical, which could rapidly add to the carbon-carbon double bond of alkene **5** yielding a secondary radical. This radical could then abstract a hydrogen atom from water or a nearby residue to obtain bisubstrate **6**. The same overall mechanism is expected in the AIBN-mediated process with H-atom abstraction from CoA by the isobutyronitrile radical produced from thermal loss of nitrogen from AIBN. The radical-mediated addition of the thiol moiety of CoA has not been previously employed to prepare S-alkylated derivatives of CoA, and as demonstrated, this method may be used to prepare bisubstrates of acyltransfer enzymes that employ CoA. There are several reports of the conjugate Michael addition of the thiol moiety of CoA to enones(28-30) but 1,4-addition requires an activating group, typically a carbonyl.

The results of the enzyme-mediated and radical-mediated bisubstrates combined with PBTSM and enzyme kinetics studies provide new information regarding key residues that are important for PA4794 substrate binding and tetrahedral intermediate stabilization. The two different binding modes of the bisubstrates showed two arginine residues can interact with the carboxylate of the NPAcGK substrate: R49 and R141. The alkyl side chain of the C-terminal lysine of the substrate is less distorted in the bisubstrate **3** and bisubstrate **6** structures compared to the ternary complex (PDB ID: 4L8A) structure. The carboxylate of bisubstrate **3** interacts with R141 instead of R49 because the bisubstrate is formed by sulfur attacking the alpha-CH_2 of the chloroacetyl group, which provides an extra methylene and lengthens the molecule versus the sulfur attacking the e-carbon of the alkene for formation of bisubstrate **6**. In absence of the complex structure with bisubstrate **3**, we would not have identified R141 as a potentially important residue for ligand binding. It appears that either arginine residue can contribute to ligand stabilization for catalysis, but turnover is dramatically decreased when both are not present. Since the native substrate of this enzyme is not yet known, these results provide new information about the architecture of the active site that is critical for catalysis.

Our study has expanded upon the tools for forming bisubstrates for GNAT structure determination and can be applied to other acyltransferases, which indicates their broad applicability across enzyme families and types of acceptor substrates. In particular, our serendipitous radical-mediated approach for producing bisubstrates could enhance the quantity of acyltransferase crystal structures or other enzymes that use CoA obtained in complex with specific ligands through employment of unactivated alkene derivatives. Since PA4794 is specific for acylating a C-terminal lysine residue, we consider it a model system for identifying the best chemical approaches for studying and designing specific inhibitors for other lysine

acetyltransferases, such as histone acetyltransferases. Furthermore, substrates armed with alkylating agents can serve as covalent inhibitors with therapeutic application, especially given the recent (re)acceptance of covalent inhibitors as drugs(31) including those targeting protein kinases.(32)Finally, our PBTSM approach for modeling the mechanism of an enzyme in reverse chronological order offers a new approach to utilizing product- or bisubstrate-bound structures in predicting key enzyme interactions involved in catalysis, and in identifying specific amino acid residues for further study.

Experimental Procedures

Materials - All solvents were distilled prior to use, and all reagents were used without further purification unless otherwise noted. All synthetic reactions were conducted under a nitrogen atmosphere. Silica gel 60A, 40−75 μm (200 × 400 mesh), was used for column chromatography. Aluminum-backed silica gel 200 μm plates were used for TLC. ^{1}H NMR spectra were obtained using either a 300 MHz spectrometer or a 500 MHz spectrometer with trimethylsilane (TMS) as the internal standard. ^{13}C NMR spectra were obtained using a 75 MHz spectrometer or a 125 MHz spectrometer. HRMS spectra were measured on a TOF instrument by electrospray ionization (ESI). Acetyl coenzyme A (AcCoA) and coenzyme A (CoA) trilithium salts and N-phenylacetylglycine lysine (NPAcGK) were purchased from Sigma-Aldrich. All other reagents for biochemical assays were purchased at the highest quality available.

Synthesis of compounds to form bisubstrates — NPAcGK (**1**) was reacted with chloroacetyl chloride, and also with bromoacetyl bromide, affording electrophilic derivatives **2a** and **2b**, respectively (Scheme 1). As detailed below, **2a** and **2b** both yielded the same bisubstrate **3** in the active site of PA4794. We employed a classic nitrous acid conversion of the primary ε-amine of the lysine moiety of NPAcGK to obtain a mixture of the alcohol **4** and alkene **5**, produced by reaction with water, or by elimination, respectively (Scheme 2). We found that the diazonium reaction utilizing sodium nitroprusside(23) was milder and higher yielding than the traditional sodium nitrite and hydrochloric acid. The mixture of **4** and **5** proved to be inseparable with normal-phase chromatography, nevertheless this provided the mixture to test as a potential substrate for the enzyme, as well as providing material for cocrystallization.

Synthesis of N6-(2-chloroacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (2a)— K_2CO_3 (46.0 mg, 0.33 mmol) was added to a solution of NPAcGK **1** (53.0 mg, 0.165 mmol) in deionized water (0.55 mL). The reaction mixture was then cooled in an ice bath and chloroacetyl chloride (19.8 μL, 28.0 mg, 0.248 mmol) was added under nitrogen slowly over several minutes via micro syringe, and the reaction was allowed to warm to RT over 4 hours. The pH of the reaction was adjusted to 4 by adding 4N HCl before liquid/liquid extraction with three portions of ethylacetate to recover the desired product. The combined organic fractions were washed with brine and dried over magnesium sulfate before being concentrated under reduced pressure to afford the desired chloroacetyl derivative **2a** as a colorless glass (34.3 mg, 52%): ^{1}H NMR (500 MHz, 1:3 $CDCl_3$ + d_6-DMSO) δ 12.7 (bs, 1H), 7.91 (t, J=5.5 Hz, 1H), 7.78 (t, J= 5.0, 1H), 7.65 (d, J=8 Hz, 1H), 7.32-7.21 (m, 5H), 4.37 (dt, J=8.5, 5 Hz, 1H), 3.96 (s, 2H), 3.86 (d, J=5.5 Hz, 2H), 3.56 (d, J=2 Hz, 2H), 3.16 (q, J=6.5 Hz, 2H), 1.84-1.77 (m, 1H), 1.62 (m, 1H), 1.47 (n, J=7 Hz, 2H), 1.32 (m, 2H); HRMS $(MH)^+$ calc for $C_{18}H_{25}ClN_3O_5$, 398.1483; found, 398.1477.

Synthesis of N6-(2-bromoacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (2b)— In a manner analagous to **2a**, bromoacetyl derivative **2b** was prepared as a colorless glass (34 mg, 16%): ^{1}H NMR (500 MHz, d_6-DMSO) δ 13.03 (bs, 1H), 8.27-8.23 (m, 2H), 8.08 (d, J=8Hz, 1H), 7.31-7.19 (m, 5H), 4.17 (dt, J=8.5, 5Hz, 1H), 3.81 (s, 2H), 3.76 (dd, J=29, 5.5Hz, 1H), 3.73 (dd, J=29, 6Hz, 1H), 3.48 (s, 2H), 3.04 (q, J=6Hz, 2H), 1.69 (m, 1H), 1.56 (m, 1H), 1.41-1.24 (m, 4H).

Synthesis of (S)-6-hydroxy-2-(2-(2-phenylacetamido)acetamido)hexanoic acid (4) and (S)-2-(2-(2-phenylacetamido)acetamido)hex-5-enoic acid (5). NPAcGK (1) (100 mg, 0.312 mmol) was dissolved in nano-pure H_2O (4mL) followed by 4N NaOH (60 μL). The mixture was

then heated overnight at 60°C. Sodium nitroprusside (115.1 mg, 0.386 mmol) was added to a separate vial and dissolved in H_2O (200 µL) with gentle heating. The nitroprusside was then added to the NPAcGK basic solution slowly with alternating 30 µL portions of nitroprusside followed by 30µL of 4N NaOH over 35 min, total volume equaling 150 uL of iron catalyst solution and 170 uL NaOH. The sample was then allowed to stir at 60°C overnight. The reaction was adjusted to pH 4 by adding 4N HCl before liquid/liquid extraction with three portions of ethyl acetate to recover the desired product. The combined organic fractions were washed with brine and dried over magnesium sulfate before being concentrated under reduced pressure to afford the desired alkene/alcohol product mixture **4**/**5** as a colorless oil (30.1 mg, 30%). RP LC-MS analysis on a Surveyor Liquid Chromatography system equipped with a Symmetry Shield RP18 column (5.0 µm, 100 Å) and coupled to an LCQ Advantage Mass Spectrum System (Elution with a 5 % CH_3CN (0.1 v/v TFA)-H_2O [A]/ 100 % CH_3CN (0.1 v/v TFA) [B] at 1mL/min flow rate and T_{total} = 20 min. Solvent gradient of $T_{0\ min}$ = 90 % A / 10 % B to $T_{15\ min}$ = 10 % A / 90 % B until gradient switch at $T_{15.1\ min}$ back to 90 % A / 10 % B until $T_{20\ min}$) was performed. The alcohol **4** retention time was 5.067 min (MS m/z calc for $C_{16}H_{22}N_2O_5$ obs 322 and alkene **5** retention time was 7.414 min (MS m/z calc for $C_{16}H_{20}N_2O_4$ 304, obs 304). The product mixture was used without further purification.

Clones and site-directed mutagenesis—The PA4794 gene was cloned into the ampicillin resistant p11 pET-derived vector as described previously(16). Site-directed mutagenesis was performed using the Phusion site-directed mutagenesis kit from Thermo Fisher Scientific, PA4794 WT template DNA, and 5'-phosphorylated primers to create R49A, R49Q, R141A, R141Q, and R49QR141Q mutants. Primers included: R49A forward atcgccgaag*ca*cgcggcagc, R49Q forward atcgccgaa*cag*cgcggca, R49 reverse ggcggcggccagctggg, R141A forward atcgccgaag*cg*cacgatccg, R141Q forward atcgccgaa*cag*cacgatccg, and R141 reverse ggcgcgcggctggtagcc. Mutations are underlined. The R49QR141Q mutant was constructed using R49Q DNA as template and primers for R141Q. See Supplemental Tables S1 and S2 for details of mutagenesis conditions. After PCR, the products were ligated as described in the kit, transformed into *E. coli* DH5alpha competent cells prepared using the Zymo Mix & Go transformation kit, and plated on LB-Agar plates containing 100 µg/mL ampicillin. Bacterial colonies were submitted to Genewiz for DNA sequence confirmation. Afterwards, colonies containing plasmids with correct sequences were grown and DNA was purified and transformed into BL21(DE3) competent cells prepared using the Zymo Mix & Go transformation kit for protein expression.

Protein expression and purification—The WT PA4794 protein for crystallization was expressed and purified as described previously(16). All WT and mutant plasmids for kinetic studies were transformed into BL21(DE3) cells and grown in 100 mL of terrific broth (TB) in 500 mL shaker flasks on a benchtop shaker at 37°C until the OD_{600nm} reached 0.6-0.8. Cultures were chilled on ice and then 0.5 mM IPTG was added to induce protein expression overnight at RT on a benchtop shaker. The next day cells were harvested by centrifugation, resuspended in 15 mL of lysis buffer (10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5% glycerol, 5 mM beta-mercaptoethanol (BME), and 5 mM imidazole), sonicated, and centrifuged. The supernatant (crude extract) was loaded onto a 1 mL Hi-Trap FF Ni-affinity column (GE Healthcare) equilibrated with buffer A (10 mM Tris-HCl pH 8.3, 500 mM NaCl, 5 mM BME) using an AKTA Start FPLC system (GE Healthcare) at a flow rate of 1 mL/min. Unbound sample was washed from the column (5 CVs of buffer A plus 25 mM imidazole) and purified protein was eluted from the column with buffer B (5 CVs of buffer A plus 500 mM imidazole). Proteins were desalted using a PD-10 column (GE Healthcare) by gravity flow into buffer A and concentrated to <0.5 mL using Sartorius 6 mL concentrators with 10K MWCO. The polyhistidine tag on the WT protein was cleaved using TEV protease (purified as described previously(16)) in a 20:1 ratio of purified protein to TEV protease in Slide-a-lyzer dialysis tubes (Thermo Fisher Scientific). The cleavage buffer (50 mM Tris-HCl pH 8.3, 300 mM NaCl, 5% glycerol, 1 mM DTT) was added to the bottom portion of the dialysis

tube and incubated at 37°C on an oscillating shaker for 2 hrs. The buffer in the bottom of the dialysis tube was discarded, fresh cleavage buffer was added to the tube, and the sample was at incubated at 4°C on an oscillating shaker in a cold cabinet overnight. The sample was centrifuged to remove any precipitate and the soluble fraction was loaded onto a 1 mL HiTrap FF column (GE Healthcare). Protein was desalted using a PD-10 column into buffer A and concentrated as described above and stored in aliquots at -80°C. All proteins were purified to near homogeneity as assessed by SDS-PAGE. We typically remove the polyhistidine tag from this protein prior to performing enzyme kinetics because it binds the tag in its active site (33); however, the efficiency of TEV cleavage was not consistent for all mutants and we could not separate the cleaved from non-cleaved proteins. Therefore, we prepared a set of proteins, including all mutant and WT, with the tag remaining for kinetic analyses in order to compare all proteins consistently.

Colorimetric Ellman's enzyme assays—*Kinetic characterization of WT and mutant proteins:* Substrate saturation curves of NPAcGK for the WT and mutant enzymes were performed using the Ellman's colorimetric assay and analyzed as described previously(16, 34) with the following modifications. Since the polyhistidine tag could not be completely removed for all mutants, we decided to compare the kinetics of all proteins used for the kinetic analysis with the tags remaining at their N-termini for consistency. The enzyme concentration for WT, R49A, R49Q, R141A, R141Q, or R49QR141Q in the enzyme assay was 2.5 (WT) and 11.2 μM (all mutants). *Determination of kinetic mechanism:* The most compatible model for the WT PA4794 kinetic mechanism was determined as described previously(25) where a series of four NPAcGK substrate saturation curves at set concentrations of AcCoA (0.1, 0.25, 0.5, 1 mM) were produced using cleaved 2.1 μM WT protein while varying NPAcGK concentrations from 0-25mM. *Substrate analog test:* To determine if the PA4794 enzyme could acetylate the NPAcGK alcohol analog, we performed the following assay using the **4**/**5** mixture since the two compounds could not be separated from each other. A 50 mM stock of the **4**/**5** mixture was prepared by dissolving the compounds in 50 mM Bicine pH 9.0 buffer. Substrate saturation curves were produced in triplicate using reactions (50 μL volume) containing 50 mM Bicine pH 9.0 buffer, 0.5 mM AcCoA, and varying concentrations of the alcohol/alkene mixture (from 0-25 mM). The reactions were initiated with 5.1 μM WT cleaved enzyme and were processed as described previously(16). *Enzyme-mediated formation of bisubstrate with alkene:* To determine if the bisubstrate formation was enzyme-mediated, we prepared a 25 mM stock solution of the **4**/**5** mixture in 50 mM Bicine pH 9.0. Reactions (50 μL) contained 50 mM Bicine pH 9.0, 5 mM CoA, and 10 mM **4**/**5** mixture, and were initiated using either 1.1 or 11 μM cleaved WT enzyme and incubated at RT for 19 hrs. Control reactions included: 1) no enzyme, 2) no enzyme and no **4**/**5** mixture, 3) and no enzyme, no **4**/**5** mixture, and no CoA. The next day 150 μL of MeOH was added to each sample, centrifuged for 5 min at 21,000 x *g,* and transferred to an HPLC vial for analysis. No precipitate was observed.

HPLC analysis of enzyme-mediated formation of bisubstrate with alkene: To analyze possible PA4794 enzyme-mediated formation of the bisubstrate, high-performance liquid chromatography (HPLC) analysis was performed according to a previously described HPLC method(35) with the following modifications. Samples were analyzed on a Hewlett Packard 1050 series HPLC system with diode array detector (DAD) set to monitor 205 nm using a Regis Little Champ II ODS guard column coupled to a SymmetryShield RP18, 3.3 μm, 100A column. Filtered samples were loaded into standard size VWR autosampler HPLC vials fitted with 300μL conical glass inserts. Samples were eluted with a 5% acetonitrile (0.1% v/v TFA)-H_2O [A]/ 5% acetonitrile (0.1% v/v TFA) [B] at 1 mL/min flow rate. A solvent gradient from 90% A / 10% B at time 0 to 10% A / 90% B at time 15 min was used. At time 15.1 min the gradient was switched to 90% A / 10% B for a re-equilibration period until the end of the method at time 25 min. Samples were additionally analyzed on a Surveyor Liquid Chromatography system equipped with a SymmetryShield RP18, 5.0 μm, 100A column and a UV-Vis unit set to monitor absorbance between 200-275 nm. The LC system was coupled to an

LCQ Advantage Mass Spectrometer System monitoring both positive and negative modes under full scan.

HPLC analysis of radical-mediated AIBN formation of bisubstrate with alkene*:* The radical-mediated AIBN reaction was used to test for formation of the bisubstrate with the **4**/**5** mixture in the presence of coenzyme A. A sample of the **4**/**5** mixture (2.2 mg, ~6.8 μmol) was dissolved in HPLC grade methanol (35 μl) followed by addition of AIBN (300 mg, 0.34 μmol) and CoA (5.3 mg, 6.8 μmol). The reaction was stirred at 60°C over four hours and then diluted with 2.0 mL of methanol. HPLC analysis was performed as described above.

AIBN mediated bisubstrate sample purification & LC MS/MS analysis: Solid-phase extraction (SPE) was performed using an AASP Vac-Elut Model AI 6000 vacuum manifold from Analytichem International and was regulated to maintain a flowrate of 3-5 mL/min. The desalting procedure was performed using a 200 mg Oasis HLB resin (Waters) extraction cartridge loaded with a 3 mg sample of concentrated bisubstrate as follows: The SPE cartridge was conditioned with 1 mL acetonitrile followed by 1 mL of 0.1M triethylammonium acetate buffer (TEAA) at pH 7. The crude sample was dissolved in 500 μL 0.1 M TEAA and loaded onto the SPE cartridge. The cartridge was washed with 2 x 5 mL of 0.1 M TEAA followed by 5 mL DI H_2O to remove salts from the sample. The sample was recovered using three consecutive 500 μL elutions of 70% acetonitrile-H_2O. The purified analyte was then analyzed with direct inject mass spectroscopy on a LCQ Advantage Mass Spectrometer System monitoring negative mode under full scan. Mass calc for $C_{37}H_{56}N_9O_20P_3S$ 1071.26, found 1070.21. A MS^2 fragmentation ion spectrum was also obtained for the ion peak at 1070 with a 30% fragmentation power.(36)

Protein Crystallization—PA4794 protein crystals were grown using vapor diffusion and hanging drop setups. The crystallization drops were a 1 μl : 1 μl mixture of protein solution at 9 mg/mL and the precipitant composed of 2M ammonium sulfate and 100 mM Bis-Tris pH 6.5. To obtain structures of the PA4794 protein in the presence of bisubstrates, *apo*-form crystals were soaked first with 5mM CoA, followed by adding 10mM of the *N6-(2-chloroacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine* ***2a*** *or* mixture of *(S)-6-hydroxy-2-(2-(2-phenylacetamido)acetamido)hexanoic acid (****4****) and (S)-2-(2-(2-phenylacetamido)acetamido)hex-5-enoic acid (****5****)*. Prior to data collection, each crystal was transferred to a solution containing a 2:1 mixture of precipitant solution and ethylene glycol and immediately cryo-cooled in liquid nitrogen. Tracking and analysis of the experiments were performed with the LabDB(37) and XtalDB(38) data management system.

Data collection, structure determination and refinement—Data collection was performed at the 21-ID-G beam line of the Life Sciences Collaborative Access Team at the Advanced Photon Source (APS). Data were collected at a temperature of 100 K and processed with HKL-2000(39). The PA4794 structure (PDB ID: 4L8A) was used as a model to solve the structure of PA4794 in complex with bisubstrates by molecular replacement (MR) using HKL-3000 (40) coupled with MOLREP(41). Refinement was performed using HKL-3000 coupled with REFMAC5(42), COOT(43, 44) and selected programs from the CCP4 package (Collaborative Computional Project, 1994). The atom B-factors were refined using anisotropic refinement. Structure validation was performed using MOLPROBITY(45) and ADIT(46). Structures of PA4794 in complex with two bisubstrates were determined: (A) (3R,20S)-1-(((((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-4-hydroxy-3-(phosphonooxy)tetrahydrofuran-2-yl)methoxy)(hydroxy)phosphoryl)oxy)(hydroxy)phosphoryl)oxy)-3-hydroxy-2,2-dimethyl-4,8,14-trioxo-20-(2-(2-phenylacetamido)acetamido)-12-thia-5,9,15-triazahenicosan-21-oic acid (**3**) and (B) (2R)-5-((2-(3-((2R)-4-(((((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-4-hydroxy-3-(phosphonooxy)tetrahydrofuran-2-yl)methoxy)(hydroxy)phosphoryl)oxy)(hydroxy)phosphoryl)oxy)-2-hydroxy-3,3-dimethylbutanamido)propanamido)ethyl)thio)-2-(2-(2-phenylacetamido)acetamido)pentanoic acid (**6**). The coordinates, structure factors, and intensities were deposited in the PDB (PDB IDs: 5VD6, 5VDB). Statistics describing crystallographic data collection and refinement are summarized in Table 1.

In-Silico Modeling—Modeling and molecular dynamics simulations were performed using Chemical Computing Group's MOE (Molecular Operating Environment).[Molecular Operating Environment (MOE), 2013.08;] (47) To understand the critical components of the PA4794 acetyl transfer reaction mechanism, the crystal structure of the PA4794 ternary complex (PDB ID: 4L8A) and the bisubstrate co-crystal structure obtained from soaking chloroacetyl derivative **2a** and CoA (PDB ID: 5VDB, Figure 1A) were used to model the tetrahedral intermediate expected to be generated during the acetyl transfer reaction of AcCoA (donor) and NPAcGK (acceptor). Two separate models were prepared as starting points using MOE's utility Structure-Prep. (47) First, the products of the enzymatic reaction were modeled based on the conformation of bisubstrate **3** in 5VDB using MOE's utility 3D Builder by first removing the bond between the bisubstrate sulfur of CoA and the acetyl methyl carbon of NPAcGAcK. This left NPAcGK-εNHAc and HSCoA ligands bound to the active site of the model. The 4L8A structure(16) already possessed these two products so no remodeling of the second model was necessary. Next, the models were solvated in a simple water box at pH of 7.4 and treated with NaCl counter ions. Periodic boundary conditions were enabled, and the hydrogen bonding network of the model was optimized by automatic sampling of different tautomer/protomer states using Protonate3D(48). Atoms were then optimized with a short, localized molecular minimization (MM) using the MOE utility QuickPrep with receptor atom tether constraint strength set to 10 and atoms further then 8 Å from the ligand fixed. Refinement proceeded until a RMS Gradient of 0.1 kcal/mol/Å was reached. Molecular Dynamics parameters were set to globally minimize the protein, ligand and solvent atoms of the models. To minimize variables during in the simulation, the atoms involved in H-bonding between the beta-alanine to beta-cysteamine amide bond were tethered to maintain the occupancy of the CoA in the active site during the annealing process. Simulated annealing was then performed on the product bound models. Molecular dynamics experiments were performed using a NPA algorithm with a Amber12:EHT force field, an initial warming from 0-311 K over 100 ps with equilibration at 311 K over 100 ps, and a 200 ps production stage at constant temperature followed by a 100 ps cooling stage from 300-0 K.

The two equilibrated and optimized NPAcGK-εNAc + HS-CoA product models were then remodeled to produce tetrahedral intermediate models. Another local minimization using QuickPrep was performed to correct the distorted atomic bonds resulting from manual ligand modification. Molecular dynamics simulations were performed to optimize the residue interactions responsible for stabilizing the tetrahedral intermediate. A second simulated annealing simulation was performed as described above except a 700 ps production stage at constant temperature followed by a 100 ps cooling stage from 300-0 K was used. Binding affinities for the molecular dynamics results were calculated based on tetrahedral intermediates (Figure 6). Similar modeling was performed using the MOE 3D Build utility to systematically transform the tetrahedral intermediates for each model into corresponding substrate NPAcGK + AcCoA models. The bond between N6 and the carbon of the acetyl was deleted and the double bond of the carbonyl was reestablished, and the carbonyl oxygen and the nitrogen atoms were given a neutral charge. Then, a third and final molecular dynamics simulation was performed to validate the model using the same parameters as described above except with a 200 ps equilibration stage and 700 ps production stage at constant temperature and a 100 ps cooling stage from 300-0 K. A short molecular mechanic minimization was performed and binding affinities were calculated (Figure 8). The differences in binding modes between the two respective models for products, tetrahedral intermediates and substrates were compared to determine the likely residues critical for acetyl transfer and the most reasonable substrate approach to the active site.

Acknowledgements

Paul Chiarelli is gratefully acknowledged for instruction in using the LC-MS instrument and in performing the solid phase extraction technique. This project was funded in part with Startup Funds from San Francisco State University (to MLK). Additional funding for this project included Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), U.S. Department of

Health and Human Services, under Contracts No. HHSN272200700058C and HHSN272201200026C and by NIH grant U54 GM094585. Some results shown in this report are derived from work performed at the Structural Biology Center Sector 19 at the Advanced Photon Source. The Advanced Photon Source, an Office of Science User Facility, is operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory. Argonne is operated by UChicago Argonne, LLC, for the U.S. DOE Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions

CR, KAM, JD, DPB, and MLK were involved in the concept and design of the experiments. CR, KAM, JD, KJ, DT, ML, YP, and MLK acquired the data. CR, KAM, JD, DT, WM, DPB, and MLK analyzed and interpreted the data. CR, KAM, DPB, and MLK wrote the manuscript and JD, DT, KJ, ML, YP, and WM critically revised the manuscript. All authors approved the final version of the manuscript to be published and agree to be accountable for all aspects of the work.

References

1. Fetrow, J.S., Siew, N., Di Gennaro, J.A., Martinez-Yamout, M., Dyson, H.J., and Skolnick, J. (2001) Genomic-scale comparison of sequence- and structure-based methods of function prediction: does structure provide additional insight?. *Protein Sci.* **10,** 1005-1014

2. Vetting, M.W., de Carvalho, Luiz Pedro S, Yu, M., Hegde, S.S., Magnet, S., Roderick, S.L., and Blanchard, J.S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch.Biochem.Biophys.* **433,** 212-226

3. Neuwald, A.F., and Landsman, D. (1997) GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem.Sci.* **22,** 154-155

4. Yu, M., Sorio de Carvalho, L.P., Sun, G., and Blanchard, J.S. (2006) Activity-Based Substrate Profiling for Gcn5-Related N-Acetyltransferases: The Use of Chloroacetyl-Coenzyme A To Identify Protein Substrates. *J. Am. Chem. Soc.* **128,** 15356-15357

5. Vetting, M.W., Magnet, S., Nieves, E., Roderick, S.L., and Blanchard, J.S. (2004) A Bacterial Acetyltransferase Capable of Regioselective N-Acetylation of Antibiotics and Histones. *Chem. Biol.* **11,** 565-573

6. Gao, F., Yan, X., Baettig, O.M., Berghuis, A.M., and Auclair, K. (2005) Regio-and Chemoselective 6′-N-Derivatization of Aminoglycosides: Bisubstrate Inhibitors as Probes To Study Aminoglycoside 6′-N-Acetyltransferases. *Angewandte Chemie.* **117,** 7019-7022

7. Poux, A.N., Cebrat, M., Kim, C.M., Cole, P.A., and Marmorstein, R. (2002) Structure of the GCN5 histone acetyltransferase bound to a bisubstrate inhibitor. *Proc.Natl.Acad.Sci.U.S.A.* **99,** 14065-14070

8. Lavogina, D., Enkvist, E., and Uri, A. (2010) Bisubstrate Inhibitors of Protein Kinases: from Principle to Practical Applications. *ChemMedChem.* **5,** 23-34

9. Dancy, B.M., and Cole, P.A. (2015) Protein Lysine Acetylation by p300/CBP. *Chem.Rev.* **115,** 2419-2452

10. Favrot, L., Blanchard, J.S., and Vergnolle, O. (2016) Bacterial GCN5-Related N-Acetyltransferases: From Resistance to Regulation. *Biochemistry (N.Y.).* **55,** 989-1002

11. Cole, P.A. (2008) Chemical probes for histone-modifying enzymes. *Nat Chem Biol.* **4,** 590-597

12. Wolf, E., De Angelis, J., Khalil, E.M., Cole, P.A., and Burley, S.K. (2002) X-ray crystallographic studies of serotonin N-acetyltransferase catalysis and inhibition1. *J.Mol.Biol.* **317,** 215-224

13. Lau, O.D., Kundu, T.K., Soccio, R.E., Ait-Si-Ali, S., Khalil, E.M., Vassilev, A., Wolffe, A.P., Nakatani, Y., Roeder, R.G., and Cole, P.A. (2000) HATs off: Selective Synthetic Inhibitors of the Histone Acetyltransferases p300 and PCAF. *Mol.Cell.* **5,** 589-595

14. Wolf, E., De Angelis, J., Khalil, E.M., Cole, P.A., and Burley, S.K. (2002) X-ray crystallographic studies of serotonin N-acetyltransferase catalysis and inhibition1. *J.Mol.Biol.* **317,** 215-224

15. Pegg, A.E. (2008) Spermidine/spermine-N^1-acetyltransferase: a key metabolic regulator. *Am.J.Physiol.Endocrinol.Metab.* **294,** E995-E1010

16. Majorek, K.A., Kuhn, M.L., Chruszcz, M., Anderson, W.F., and Minor, W. (2013) Structural, Functional, and Inhibition Studies of a Gcn5-related N-Acetyltransferase (GNAT) Superfamily Protein PA4794. *J. Biol. Chem.* **288,** 30223-30235

17. Joest, C., Nitsche, C., Scholz, T., Roux, L., and Klein, C.D. (2014) Promiscuity and Selectivity in Covalent Enzyme Inhibition: A Systematic Study of Electrophilic Fragments. *J. Med. Chem.* **57,** 7590-7599

18. Hwang, Y., Thompson, P.R., Wang, L., Jiang, L., Kelleher, N.L., and Cole, P.A. (2007) A selective chemical probe for coenzyme A-requiring enzymes. *Angew. Chem., Int. Ed.* **46,** 7621-7624

19. Hamano, Y., Hoshino, Y., Nakamori, S., and Takagi, H. (2004) Overexpression and Characterization of an Aminoglycoside 6′-N-Acetyltransferase with Broad Specificity from an ε-Poly-l-lysine Producer, Streptomyces albulus IFO14147. *The Journal of Biochemistry.* **136,** 517-524

20. Vetting, M.W., Hegde, S.S., Javid-Majd, F., Blanchard, J.S., and Roderick, S.L. (2002) Aminoglycoside 2prime]-N-acetyltransferase from Mycobacterium tuberculosis in complex with coenzyme A and aminoglycoside substrates. *Nat Struct Mol Biol.* **9,** 653-658

21. Lau, E.Y., Felton, J.S., and Lightstone, F.C. (2006) Insights into the o-acetylation reaction of hydroxylated heterocyclic amines by human arylamine N-acetyltransferases: a computational study. *Chem.Res.Toxicol.* **19,** 1182-1190

22. Liu, Y., Zhang, W., and Sayre, L.M. (2011) An alternative total synthesis of pentosidine. *J. Heterocycl. Chem.* **48,** 426-433

23. McGarvey, G.J., and Kimura, M. (1986) Sodium nitroprusside-mediated substitution of oxygen for nitrogen at saturated carbon centers. *J. Org. Chem.* **51,** 3913-3915

24. Hoyle, C.E., and Bowman, C.N. (2010) Thiol-Ene Click Chemistry. *Angew. Chem., Int. Ed.* **49,** 1540-1573

25. Filippova, E.V., Kuhn, M.L., Osipiuk, J., Kiryukhina, O., Joachimiak, A., Ballicora, M.A., and Anderson, W.F. (2015) A Novel Polyamine Allosteric Site of SpeG from Vibrio cholerae Is Revealed by Its Dodecameric Structure. *J.Mol.Biol.* **427,** 1316-1334

26. Majorek, K.A., Osinski, T., Tran, D.T., Revilla, A., Anderson, W.F., Minor, W., and Kuhn, M.L. (2017) Insight into the 3D structure and substrate specificity of previously uncharacterized GNAT superfamily acetyltransferases from pathogenic bacteria. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics.* **1865,** 55-64

27. Burmeister, W.P. (2000) Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr D Biol Crystallogr.* **56,** 328-341

28. Makowski, K., Mera, P., Paredes, D., Herrero, L., Ariza, X., Asins, G., Hegardt, F.G., Garcia, J., and Serra, D. (2013) Differential Pharmacologic Properties of the Two C75 Enantiomers: (+)-C75 Is a Strong Anorectic Drug; (-)-C75 Has Antitumor Activity. *Chirality.* **25,** 281-287

29. Li, X., Liu, N., Zhang, H., Knudson, S.E., Li, H., Lai, C., Simmerling, C., Slayden, R.A., and Tonge, P.J. (2011) CoA Adducts of 4-Oxo-4-phenylbut-2-enoates: Inhibitors of MenB from the M. tuberculosis Menaquinone Biosynthesis Pathway. *ACS Med. Chem. Lett.* **2,** 818-823

30. Leggans, E.K., Akey, D.L., Smith, J.L., and Fecik, R.A. (2010) A general scheme for synthesis of substrate-based polyketide labels for acyl carrier proteins. *Bioorg. Med. Chem. Lett.* **20,** 5939-5942

31. Singh, J., Petter, R.C., Baillie, T.A., and Whitty, A. (2011) The resurgence of covalent drugs. *Nat. Rev. Drug Discovery.* **10,** 307-317

32. Liu, Q., Sabnis, Y., Zhao, Z., Zhang, T., Buhrlage, S.J., Jones, L.H., and Gray, N.S. (2013) Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chem. Biol. (Oxford, U. K.).* **20,** 146-159

33. Majorek, K.A., Kuhn, M.L., Chruszcz, M., Anderson, W.F., and and Minor, W. (2014) Double trouble—Buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments. - *Protein Science : A Publication of the Protein Society.* **10,** 1359-1368

34. Kuhn, M.L., Majorek, K.A., Minor, W., and Anderson, W.F. (2013) Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity. *Protein Science.* **22,** 222-230

35. Wang, W., Morohoshi, T., Ikenoya, M., Someya, N., and Ikeda, T. (2010) AiiM, a novel class of N-acylhomoserine lactonase from the leaf-associated bacterium Microbacterium testaceum. *Appl. Environ. Microbiol.* **76,** 2524-2530

36. Gilar, M., Belenky, A., and Wang, B.H. (2001) High-throughput biopolymer desalting by solid-phase extraction prior to mass spectrometric analysis. *J. Chromatogr. A.* **921,** 3-13

37. Zimmerman, M.D., Grabowski, M., Domagalski, M.J., MacLean, E.M., Chruszcz, M., and Minor, W. (2014) Data management in the modern structural biology and biomedical research environment. *Structural Genomics and Drug Discovery: Methods and Protocols.* 1-25

38. Zimmerman, M., Chruszcz, M., Koclega, K., Otwinowski, Z., and Minor, W. (2005) The Xtaldb system for project salvaging in high-throughput crystallization. *Acta Cryst A.* **61,** c178-c179

39. Otwinowski, Z., Minor, W., and W Jr, C.C. (1997) Processing of X-ray diffraction data collected in oscillation mode

40. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. *Acta Crystallographica Section D: Biological Crystallography.* **62,** 859-866

41. Vagin, A., and Teplyakov, A. (2009) Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography.* **66,** 22-25

42. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography.* **67,** 355-367

43. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography.* **66,** 486-501

44. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography.* **60,** 2126-2132

45. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B.,3rd, Snoeyink, J., Richardson, J.S., and Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35,** W375-83

46. Yang, H., Guranovic, V., Dutta, S., Feng, Z., Berman, H.M., and Westbrook, J.D. (2004) Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **D60,** 1833-1839

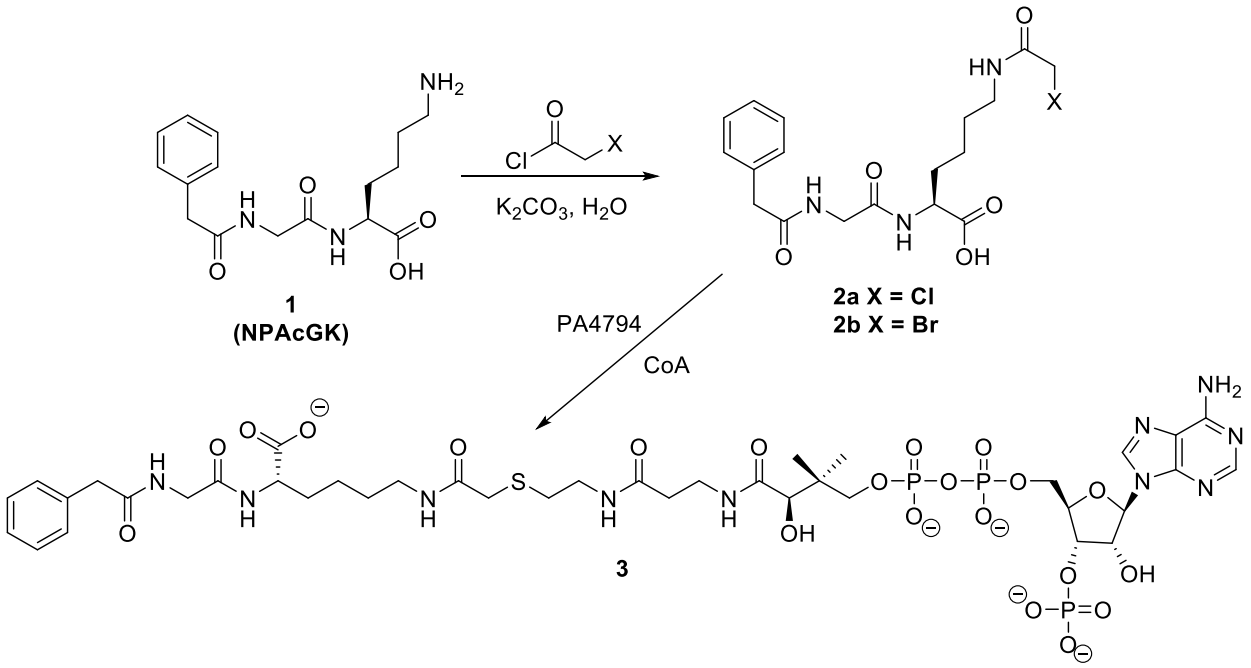
47. Molecular Operating Environment (MOE), 2013.08, and Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016

48. Labute, P. (2008) The generalized Born/volume integral implicit solvent model: Estimation of the free energy of hydration using London dispersion instead of atomic surface area. *Journal of Computational Chemistry.* **29,** 1693-1698

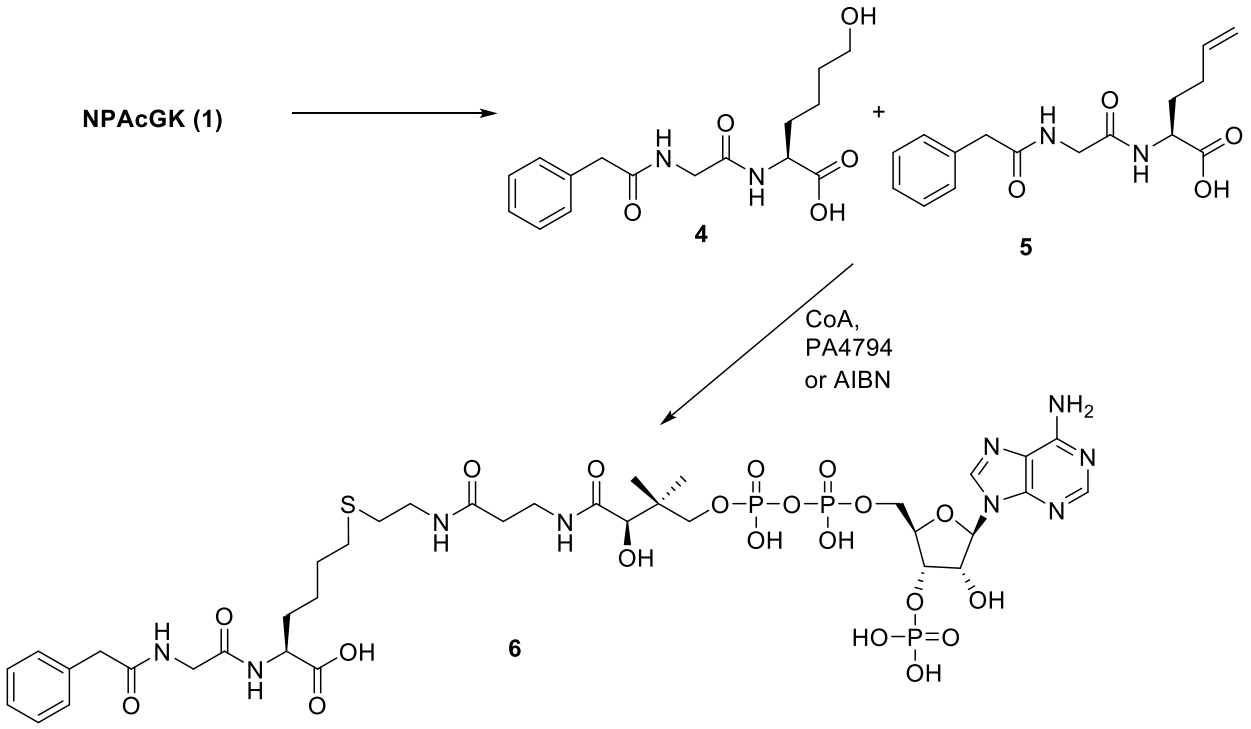
TABLE 1. Crystallographic data collection and refinement statistics. * Highest resolution shell values are shown in parentheses. ** 5% of reflections were randomly assigned to the Rfree set.

Complex	bisubstrate analog **3**	bisubstrate analog **6**
PDB code	**5VDB**	**5VD6**
Data collection		
Space group	$P2_12_12$	$P2_12_12$
Unit cell		
a, b, c (Å)	57.3, 76.1, 39.2	57.1, 76.2, 39.0
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)*	1.40 (1.40 - 1.42)	1.20 (1.20 - 1.22)
Observed reflections	237096	325622
Rmerge	0.053 (0.800)	0.067 (0.664)
$\langle I \rangle / \langle \sigma(I) \rangle$*	39.9 (2.0)	41.3 (2.2)
Completeness (%)*	99.9 (100.0)	99.0 (97.8)
Redundancy*	6.9 (6.8)	6.1 (5.5)
Refinement		
Rwork (%)/Rfree (%)**	14.5 / 17.0	13.2 / 16.3
No. atoms		
Protein	1279	1302
Ligand/Ion	131	100
Water	188	253
B factors		
Wilson B factor ($Å^2$)	17.7	13.6
Protein ($Å^2$)	19	16
Ligand/Ion ($Å^2$)	29	26
Water ($Å^2$)	36	32
Structure quality		
R.m.s. deviations		
bond length (Å)	0.01	0.01
bond angles (°)	1.2	1.5
Statistics		
Ramachandran Favored (%)	99.4	98.1

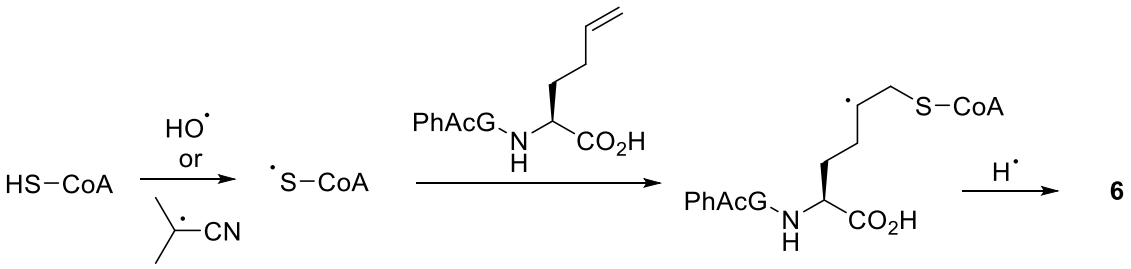
FIGURES



Scheme 1. Synthesis of a-haloacetyl derivatives of NPAcGK **2a** (X = Cl) and **2b** (X = Br) and enzyme-mediated formation of bisubstrate **3**.



Scheme 2. Formation of the alcohol/alkene mixture (**4** and **5**) from NPAcGK, and radical-mediated reaction of CoA with alkene **5** to form the covalent bisubstrate **6.**



Scheme 3. Mechanistic hypothesis for the formation of bisubstrate **6** from CoA and alkene **5.**

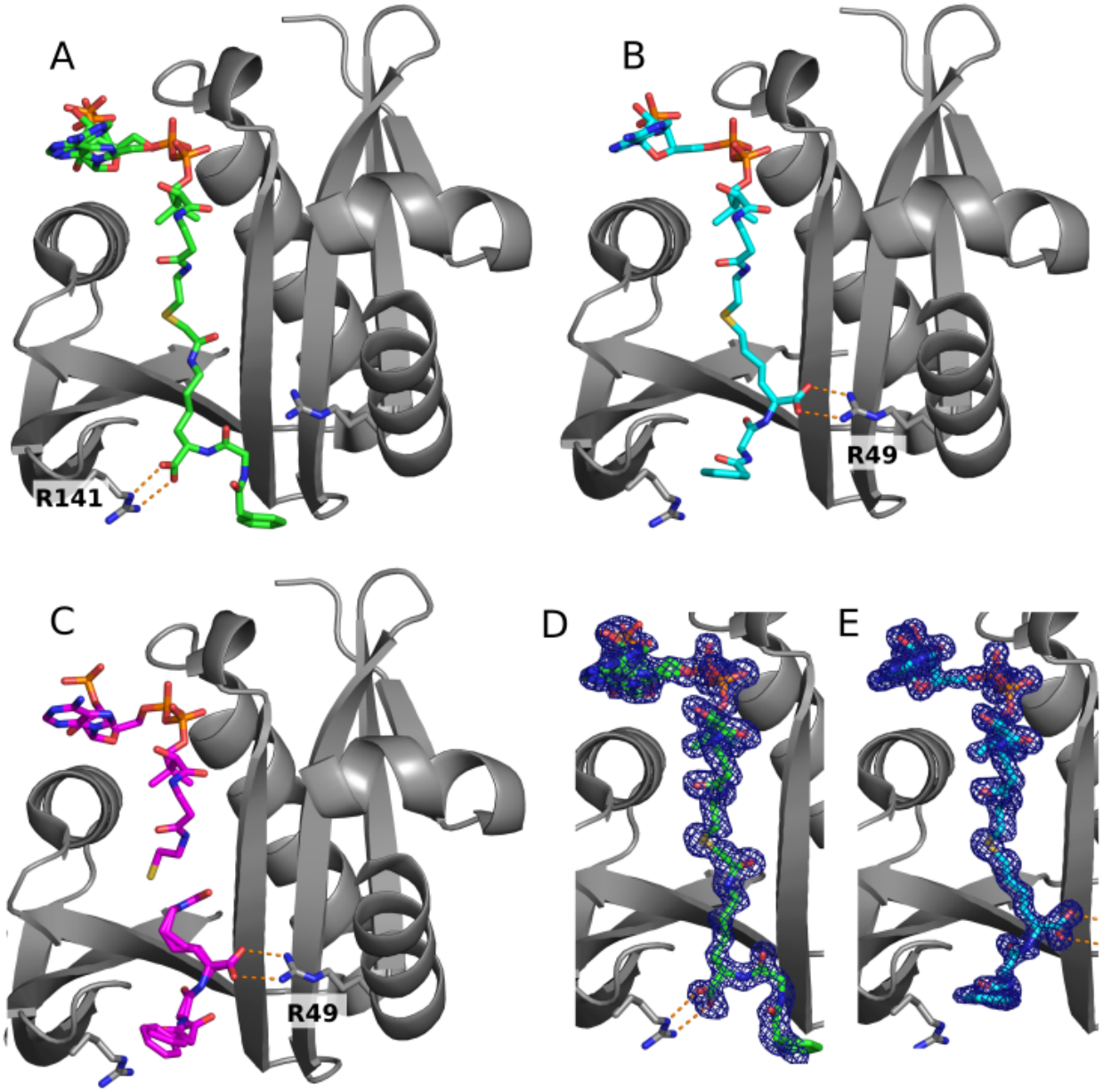


Figure 1. Comparison of binding products and bisubstrate 3 and 6 analogs. PA4794 is shown as a gray ribbon diagram. Residues 19-37 are not shown for clarity. **(A)** PA4794 in complex with bisubstrate analog **3** synthesized from CoA and chloroacetyl derivative **2a** determined at 1.4 Å resolution. (**B)** PA4794 in complex with bisubstrate analog **6** synthesized from CoA and alcohol/alkene mixture **4/5** determined at 1.2 Å resolution. **(C)** Ternary complex structure of PA4794 with the bound reaction products CoA and N6-acetylated NPAcGAcK. **(D)** 2Fo-Fc electron density map of the bound bisubstrate analog **3** (σ=1.0). **(E)** 2Fo-Fc electron density map of the bound bisubstrate analog **6** (σ=1.0).

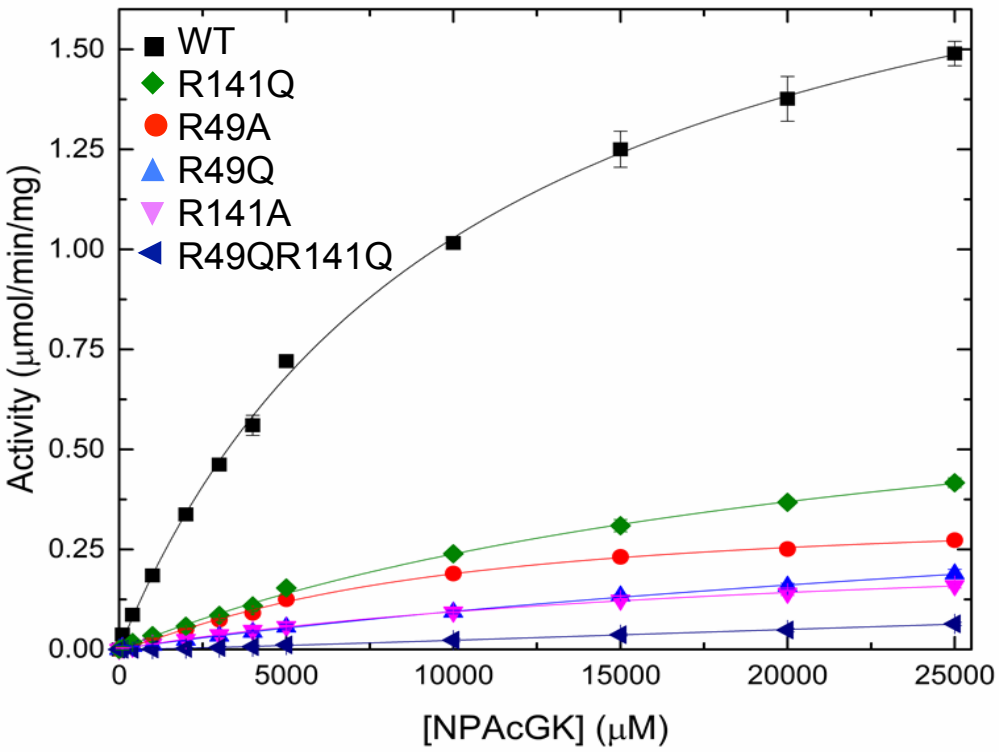


Figure 2. Substrate saturation curves for wild-type and R49 and R141 mutants. Steady-state enzyme kinetic experiments were performed by holding AcCoA concentration constant at 0.5 mM and varying NPAcGK peptide concentration. Standard deviations for triplicate reactions for all proteins are shown. See Experimental procedures for more details.

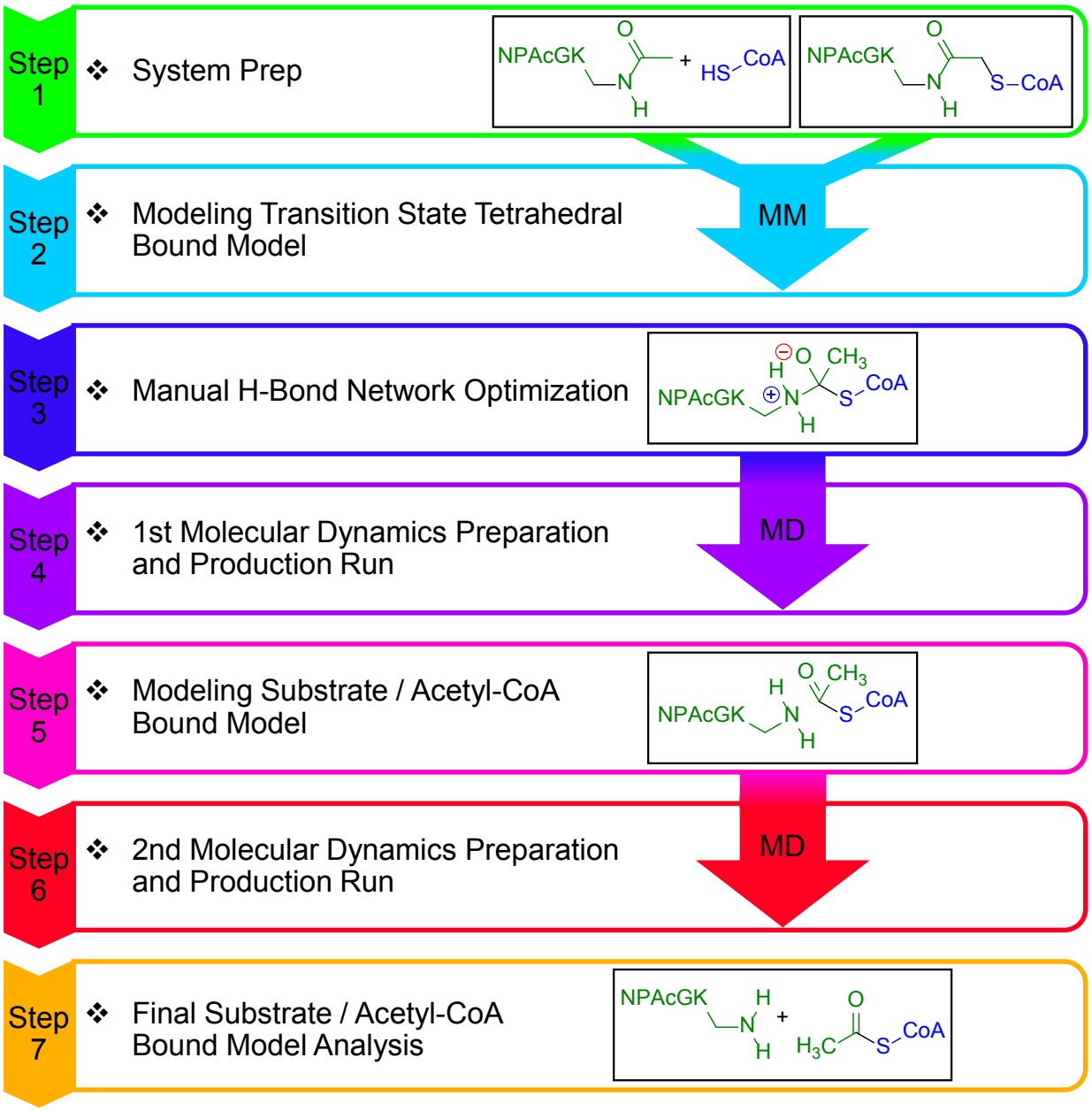


Figure 3. Product-Based Transition-State Modeling (PBTSM). Protocol for systematic remodeling of tetrahedral intermediate and substrate binding modes from product or bisubstrate bound crystal structures. Molecular minimization (MM) and molecular dynamics (MD) stages in the protocol are shown. Further details are presented in Experimental Procedures.

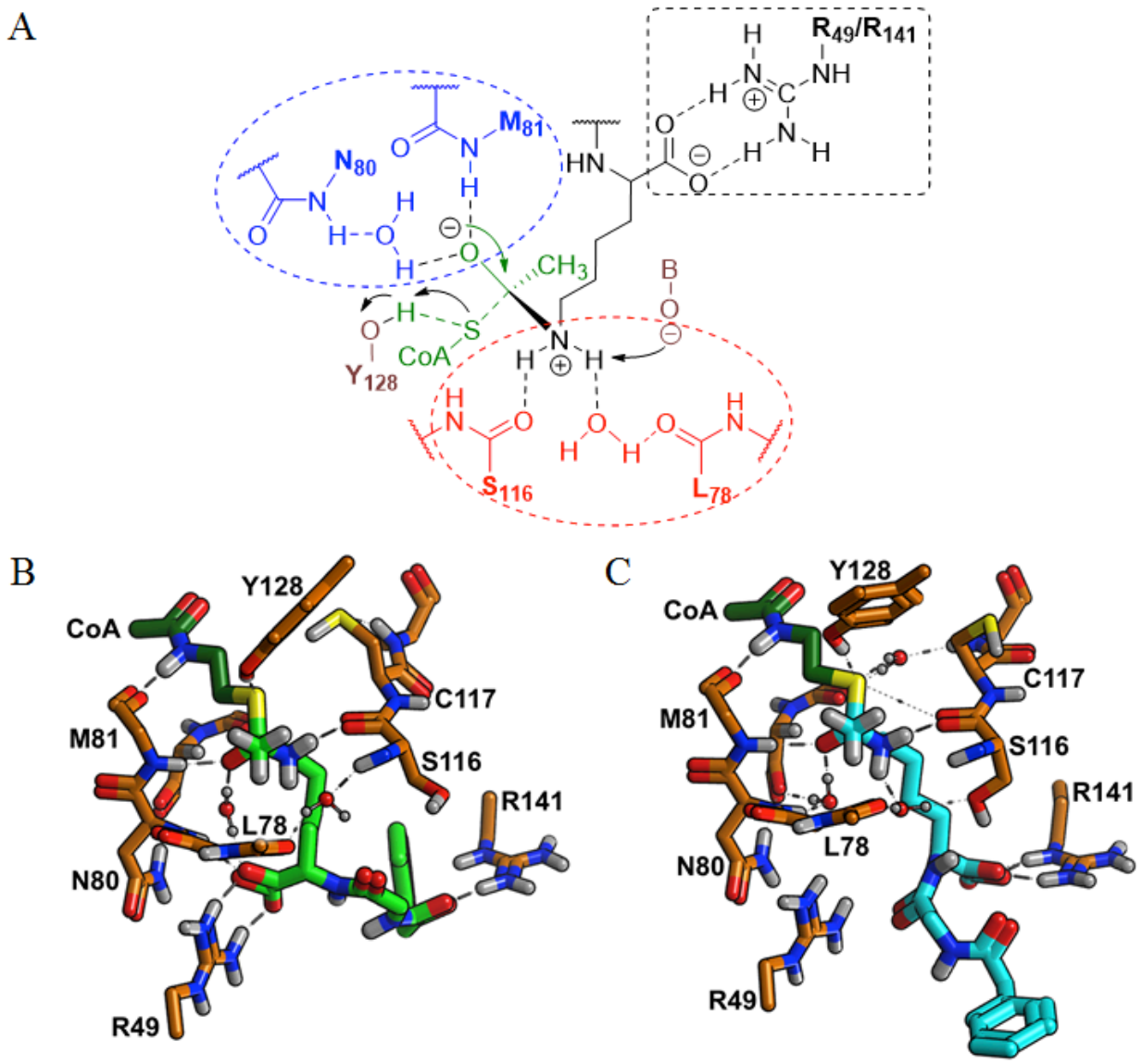


Figure 4. Product-Based Transition-State Modeling (PBTSM) Analysis. A) Schematic diagram of tetrahedral transition state intermediate indicating important regions of substrate-enzyme interactions in the active site, including interactions that anchor the substrate C-terminal carboxylate (black square), the oxyanion hole (blue circle), and stabilizing interactions for substrate amine approach (red circle). B) PBTSM-generated transition state model based on PA4794 co-crystal structure with products bound (PDB ID: 4L8A) shown in light green. C) PBTSM-generated transition state model based on PA4794 structure in complex with bisubstrate **3** shown in cyan. The CoA moiety of the tetrahedral intermediate is shown in dark green and protein side chains are in orange. Amino groups are in dark blue, oxygens in red, sulfurs in yellow, and waters as red and white spheres.

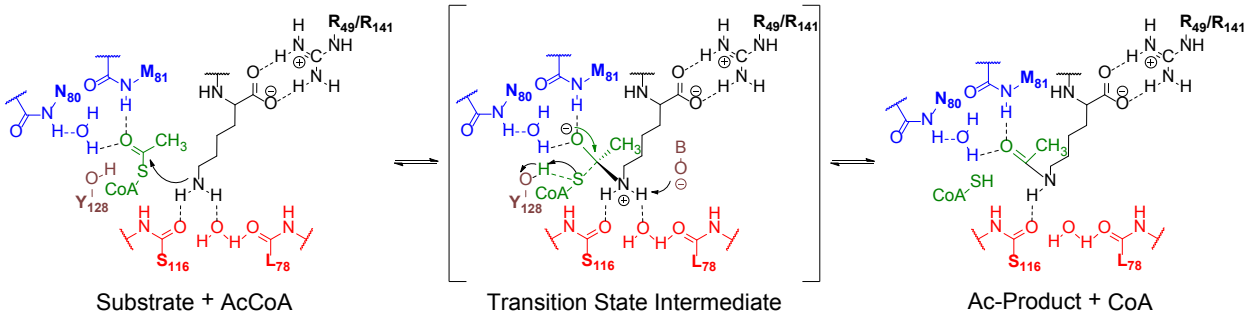


Figure 5. Proposed chemical mechanism for PA4794 acetyltransfer. Residues that occupy the oxyanion hole are in blue, AcCoA substrate is in green, Y128 that acts as the general acid is in brown, residues that stabilize the epsilon amino group are in red, and NPAcGK peptide with stabilizing arginine residues in black.

Supplemental Information

Generating enzyme and radical-mediated bisubstrates as tools for investigating Gcn5-related *N*-acetyltransferases

Cory Reidl, Karolina A Majorek, Joseph Dang, Kristen Jew, David Tran, Melissa Law, Yasmine Payne, Wladek Minor, Daniel P. Becker, and Misty L. Kuhn

Table S1. PA4794 Mutagenesis Conditions.

Mutants were constructed using the conditions outlined below. All mutants were made using WT DNA as template except R49QR141Q, which used R49Q DNA as template. Variations to the procedure included adding 2% DMSO to the reaction for the R49A mutant and changing the annealing temperature to 68°C for R49A, R141A, R141Q and 73°C for R49Q and R49QR141Q mutants.

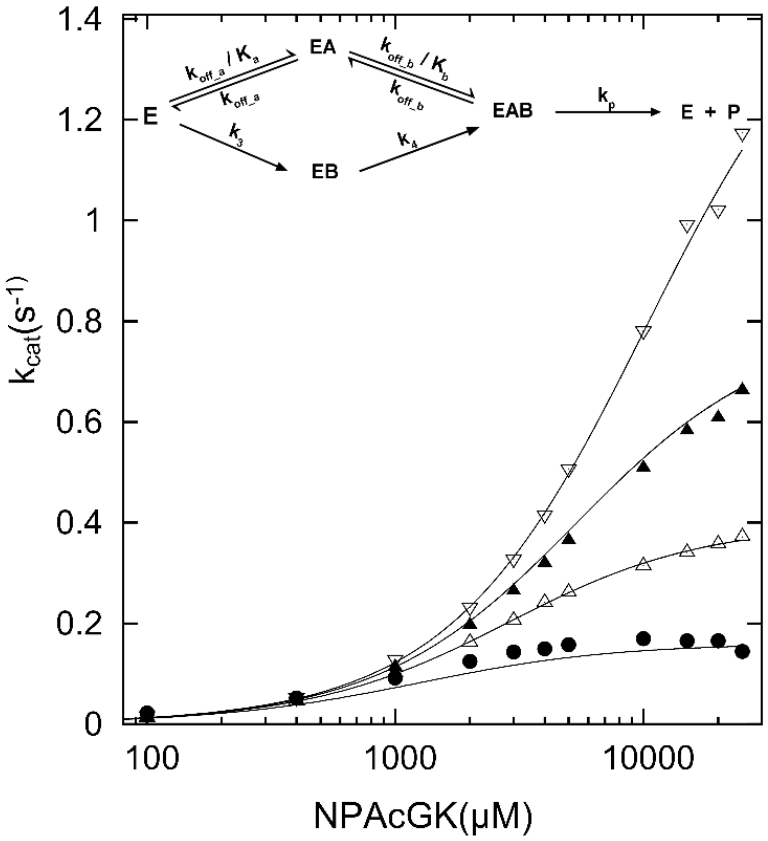
Original reagents and volumes from Phusion Site Directed Mutagenesis Protocol:

Component	**50 µL Reaction**	**Final Concentration**
Water	26 µL	-
5X Phusion HF Buffer	10 µL	1x
10 mM dNTPs	1 µL	200 µM
Forward Primer (10µM)	5 µL	1 µM
Reverse Primer (10µM)	5 µL	1 µM
Template DNA (10ng/µL)	2.5 µL	25 ng
Phusion Hot Start DNA Polymerase	0.5 µL	0.02 U/µL

Table S2. Original Thermocycler conditions from Phusion Site Directed Mutagenesis Protocol:

	Temperature (^{o} C)	Time	Number of Cycles
Initial Denaturation	98	30 seconds	1
Denaturation	98	10 seconds	25
Anneal	70	30 seconds	25
Extension	72	5 minutes	25
Final Extension	72	7 minutes	1
"Hold"	4	∞	-

Figure S1. Series of NPAcGK curves at varying concentrations of AcCoA. The fitting for the kinetic model that was most compatible with the data, i.e. gave the best AICc value, is shown. AICc values for each model are shown in the table beneath the figure. AcCoA concentrations were 0.1, 0.25, 0.5, and 1mM for filled circles, open triangles, filled triangles, and open upside down triangles, respectively. The most compatible model was the bireactant steady-state 3 model as described in [1]. A represents NPAcGK substrate and B is AcCoA.



Model	AICc Value
I : Bireactant Ordered A-First	-214
II : Bireactant Ordered B-First	-248
III : Bireactant Ping-Pong	-309
IV : Bireactant Cooperative Ping-Pong	-343
V : Bireactant Random Independent	-270
VI : Bireactant Random Cooperative	-329
VII : Bireactant Modified Hill	-335
VIII : Bireactant Random Steady-State	-348
IX : Bireactant Random Steady-State 2	-350
X : Bireactant Random Steady-State 3	-351

Figure S2. ^{1}H NMR for N6-(2-chloroacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2a**)

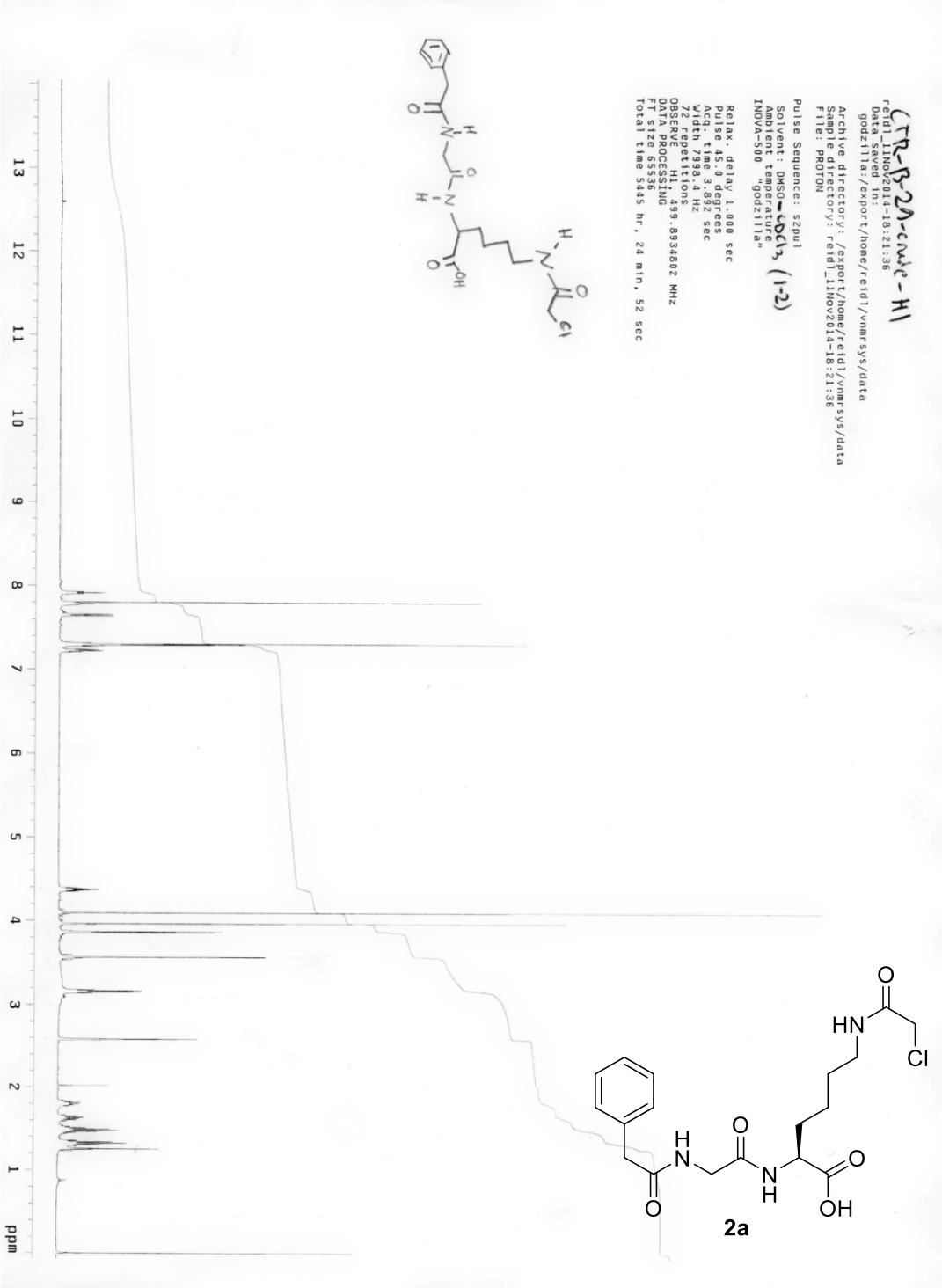


Figure S3. Homonuclear ^{1}H-^{1}H COSY NMR for N6-(2-chloroacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2a**)

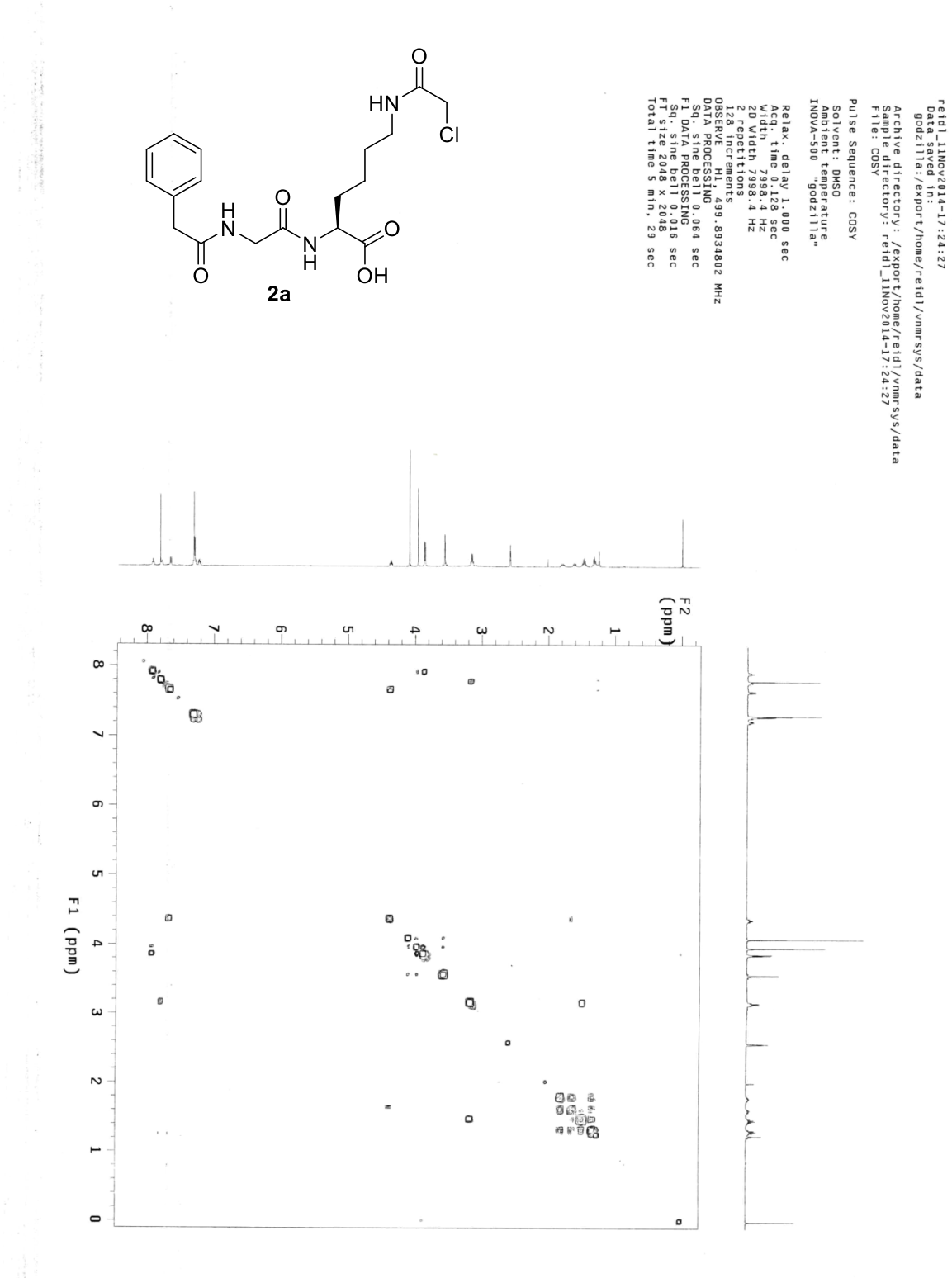


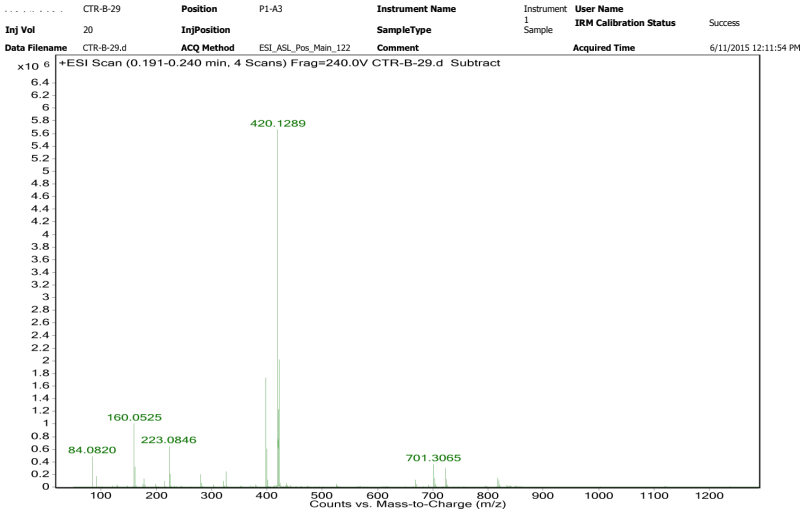
Figure S4. Mass Spectrum for N6-(2-chloroacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2a**)

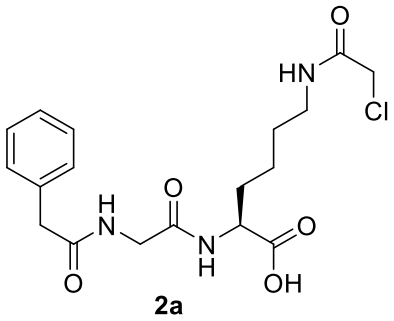
Spectrum Identification Results: + Scan (0.191-0.240 min) Sub (CTR-B-29. D **(2a)**

Best	Formula	Score	Mass	Mass (MFG)	Diff (ppm)	Diff (abs. ppm)	Diff (mDa)	ID Source	Score (MFG)
	C18 H24 Cl N3 O5	97.99	397.1398	397.1404	1.61	1.61	0.64	MFG	97.99

Species	Ion Formula	m/z	Height	Score (MFG)	Score (MFG, MS)	Score (MFG, MS/MS)	Score (MFG, mass)	Score (MFG, abund)	Score (MFG, iso. spacing)
(M+H)+	C18 H25 Cl N3 O5	398.1477	1730524.1	97.99	97.99		97.74	99.37	96.83

m/z	m/z (Calc)	Diff (ppm)	Diff (mDa)	Height	Height (Calc)	Height %	Height % (Calc)	Height Sum %	Height Sum% (Calc)
398.147	398.1477	1.71	0.7	1730593.6	1744058.1	100	100	60.4	60.8
399.1507	399.1508	0.13	0.1	368798.2	366989.9	21.3	21	12.9	12.8
400.1457	400.1455	-0.39	-0.2	609662.1	612684.6	35.2	35.1	21.3	21.4
401.148	401.1482	0.7	0.3	115070.5	123512.9	6.6	7.1	4	4.3
402.1541	402.1505	-8.94	-3.6	28906.1	18043.2	1.7	1	1	0.6
403.1594	403.153	-15.87	-6.4	5258.3	1991.8	0.3	0.1	0.2	0.1
404.1537	404.1554	4.21	1.7	9172.1	180.3	0.5	0	0.3	0





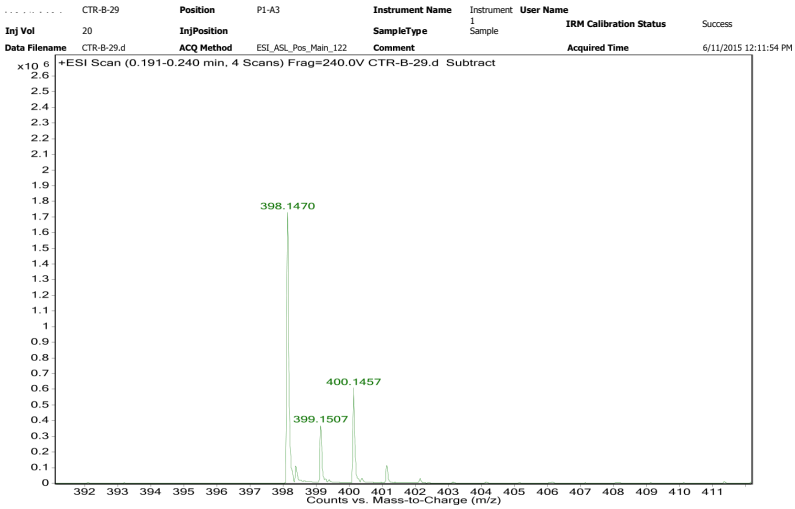


Figure S5. ^{1}H NMR for N6-(2-bromoacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2b**)

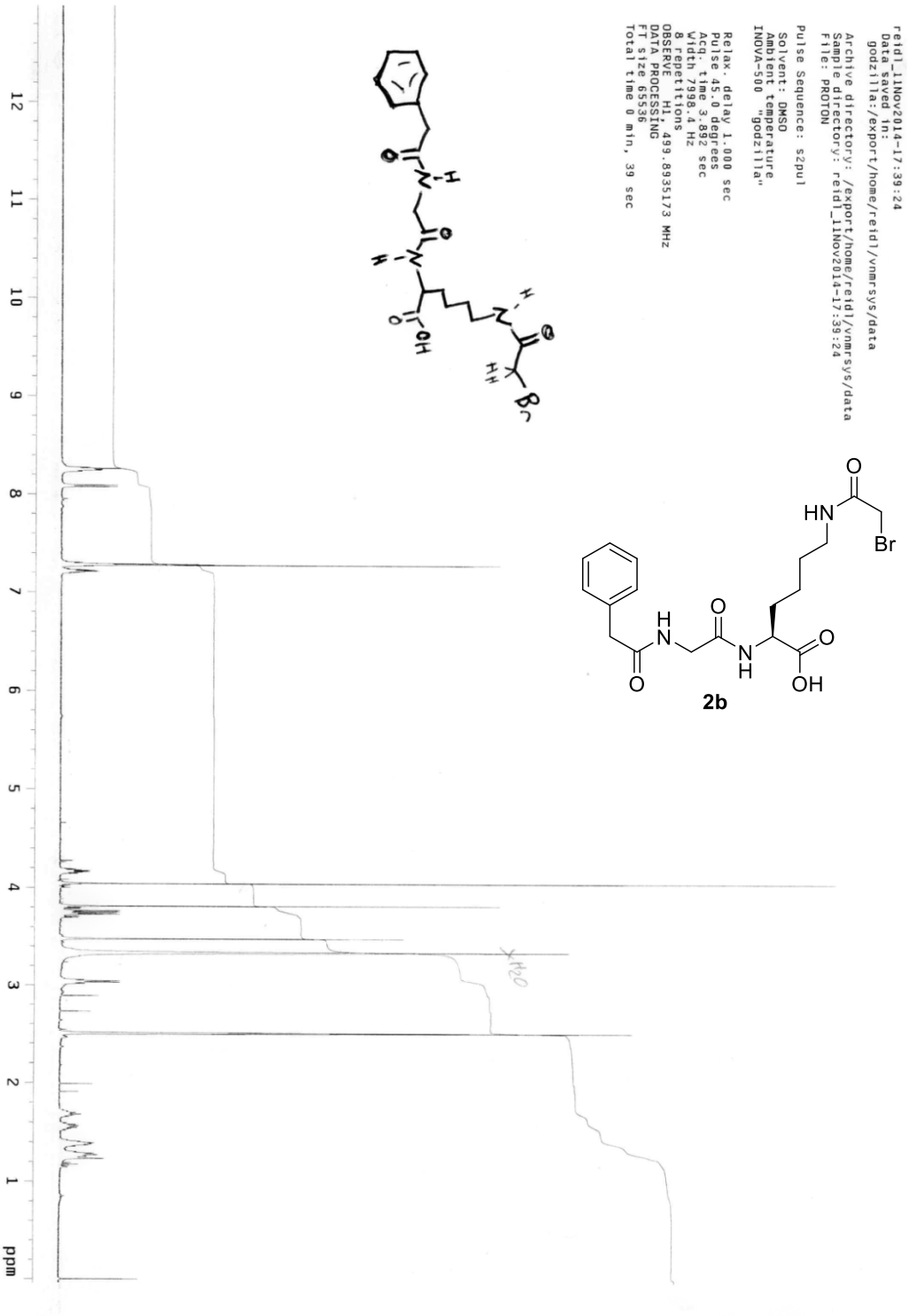


Figure S6. ^{13}C NMR for N6-(2-bromoacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2b**)

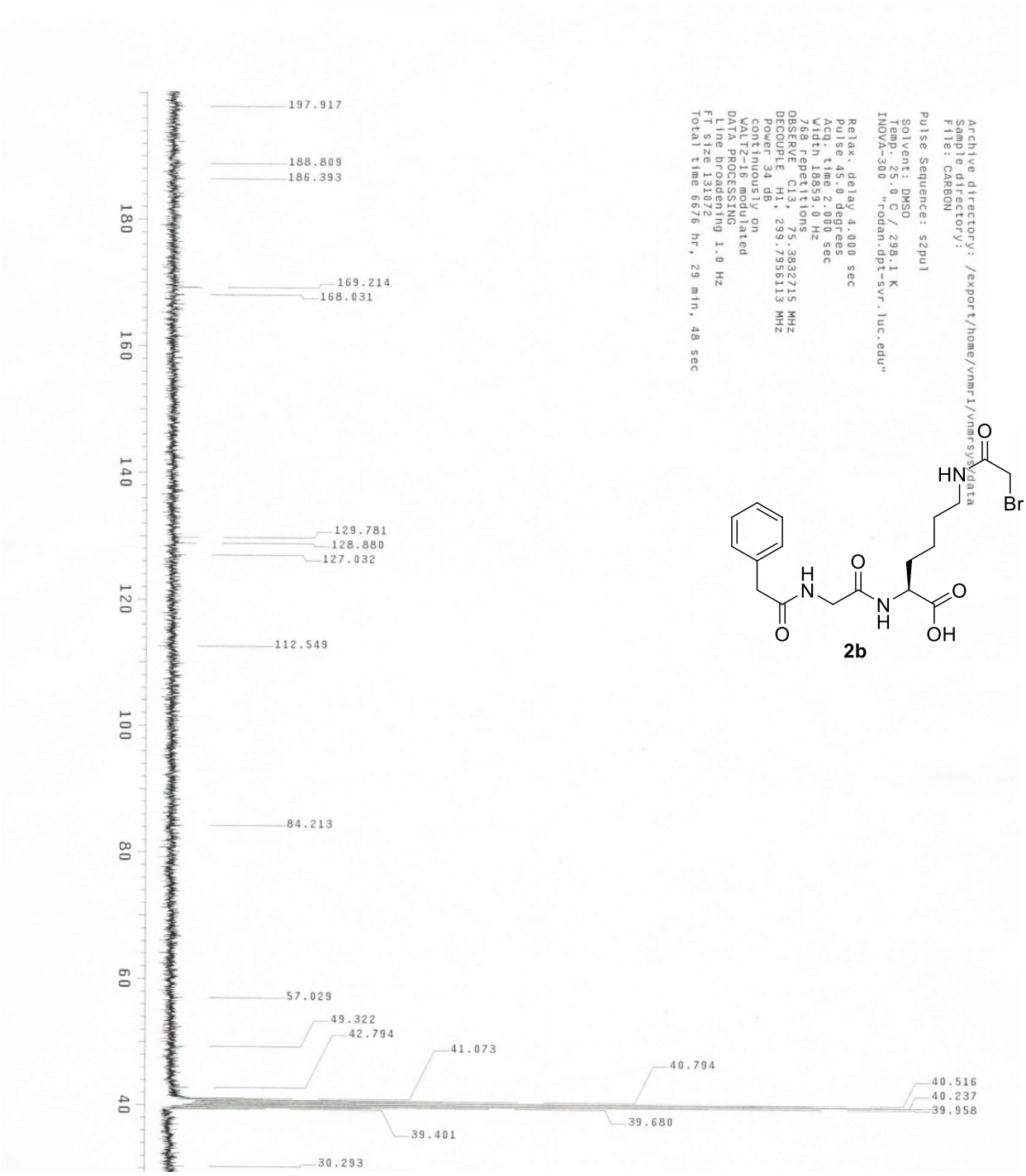


Figure S7. Homonuclear 1H-1H COSY NMR for N6-(2-bromoacetyl)-N2-((2-phenylacetyl)glycyl)-L-lysine (**2b**)

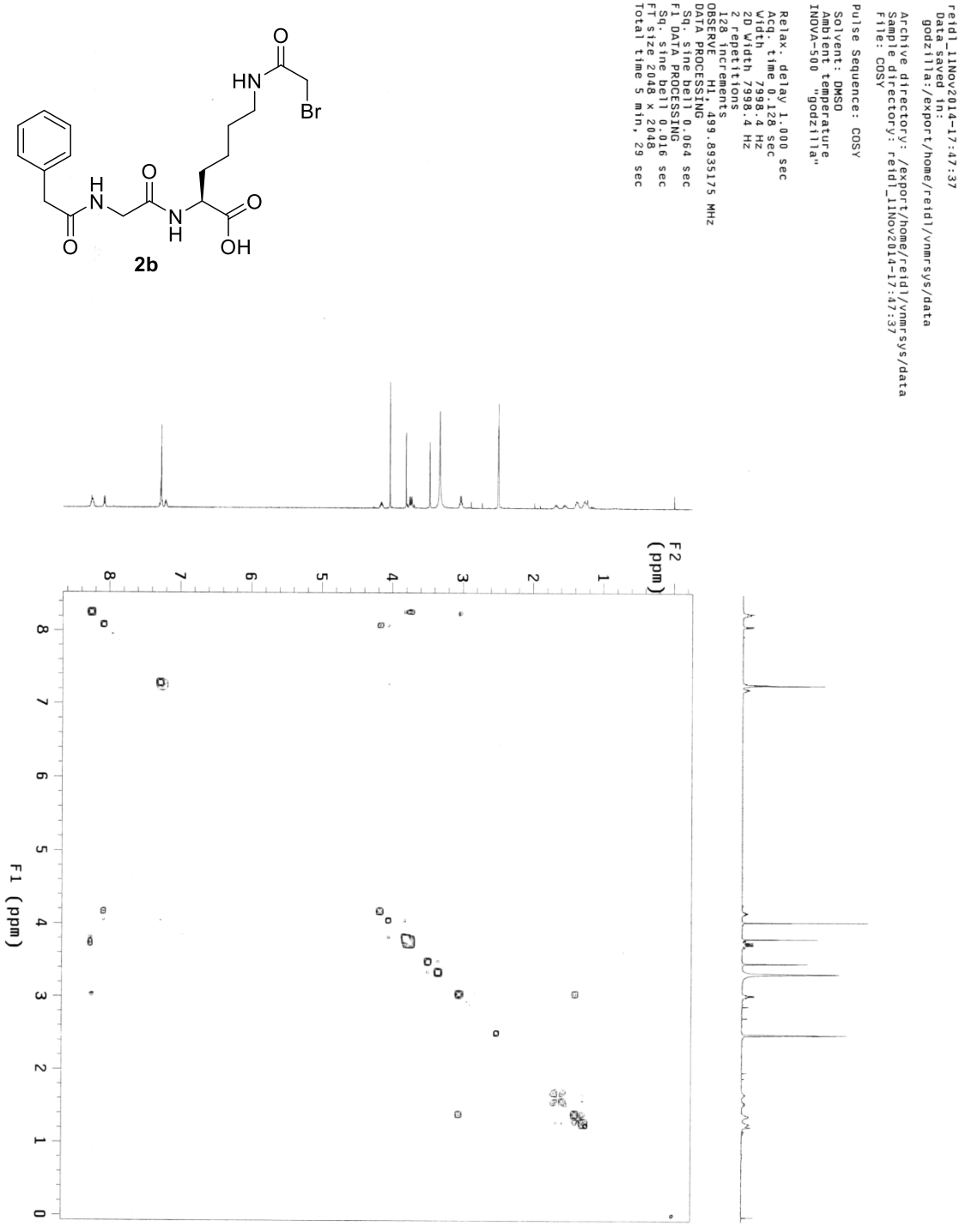
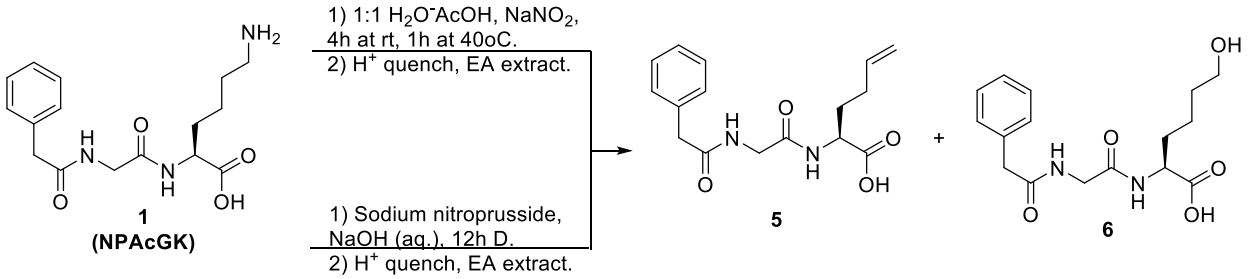
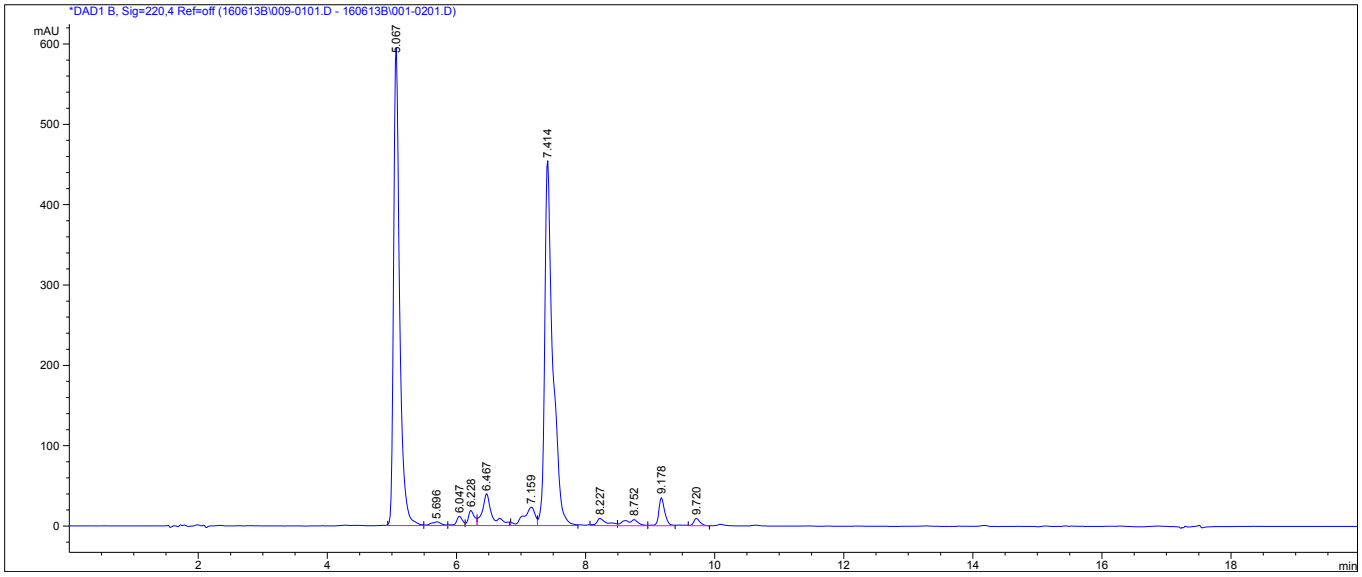


Figure S8. HPLC Data for **4**/**5** Alkene/Alcohol Mixture



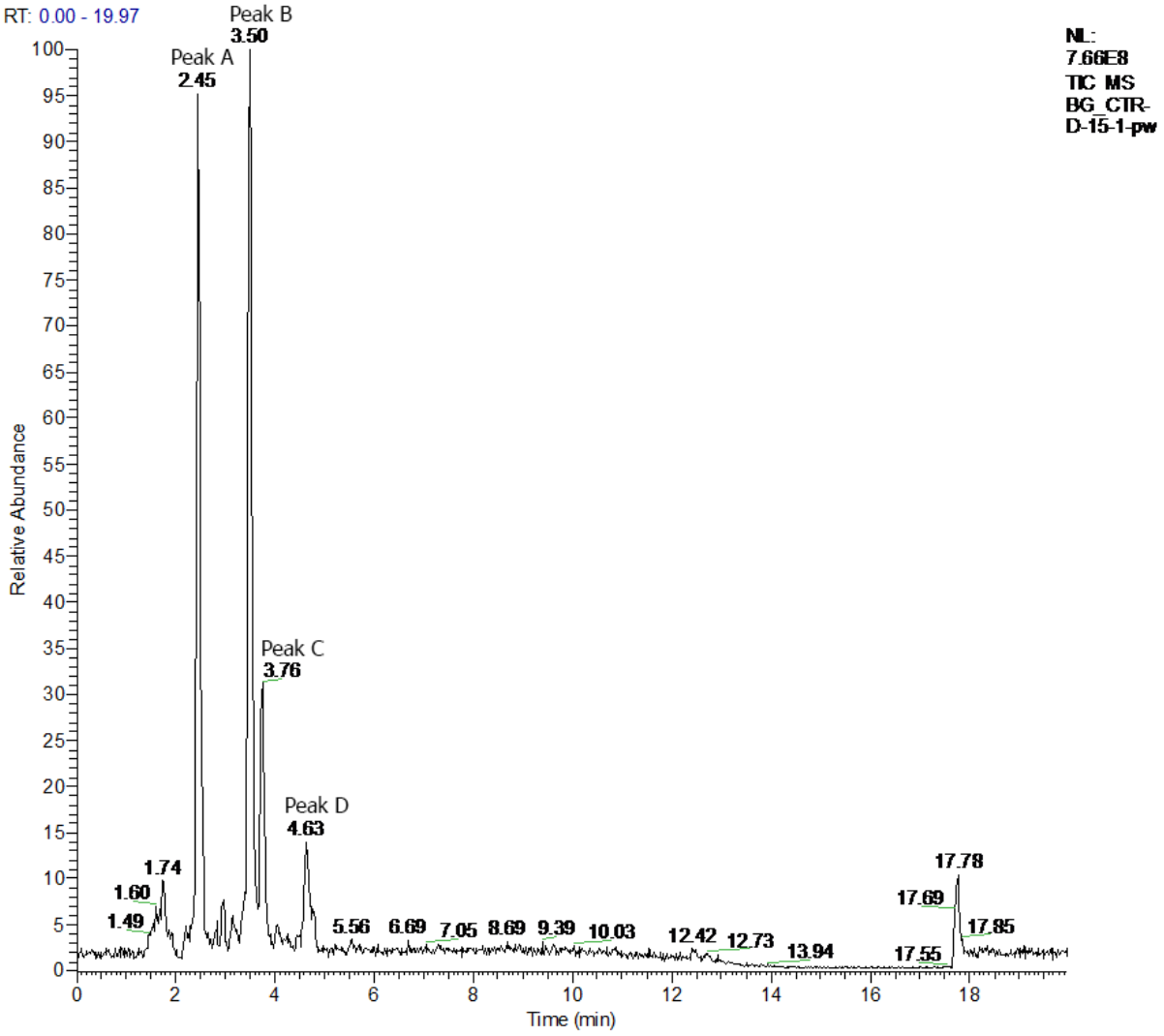
HPLC Data



File Information	
LC-File	009-0101.D
File Path	C:\HPCHEM\1\DATA\160613B\
Date	13-Jun-16, 14:18:04
Sample	CTR-D-15-1
Sample Info	
Barcode	
Operator	Cory
Method	CTVN.M
Analysis Time	19.96 min
Sampling Rate	0.0133 min (0.798 sec), 1498 datapoints

#	Time	Area	Height	Width	Area%	Symmetry
1	5.067	3861.1	594.2	0.1038	41.690	0.667
2	5.696	51.3	4.6	0.1792	0.554	1.242
3	6.047	75.1	11.4	0.1006	0.811	1.024
4	6.228	125.8	18.8	0.1021	1.358	0.796
5	6.467	419.8	39.7	0.1514	4.533	0.663
6	7.159	278.2	23	0.1691	3.004	2.145
7	7.414	3957.8	454	0.1297	42.734	0.528
8	8.227	100.6	9	0.1577	1.086	0.478
9	8.752	109.8	7.5	0.194	1.186	2.258
10	9.178	221.5	34.9	0.098	2.392	0.783
11	9.72	60.4	8.9	0.1074	0.652	0.781

Figure S9. LC-MS Data for **4**/**5** Alkene/Alcohol Mixture

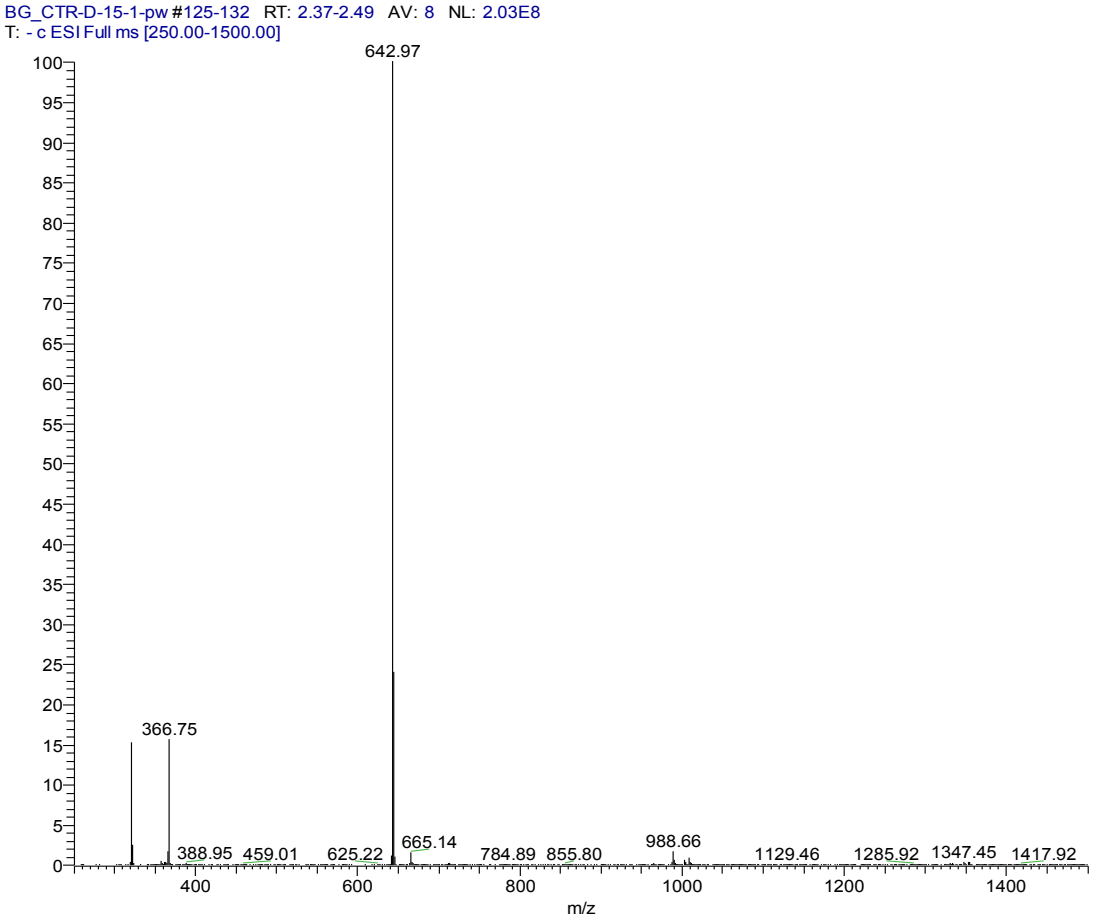


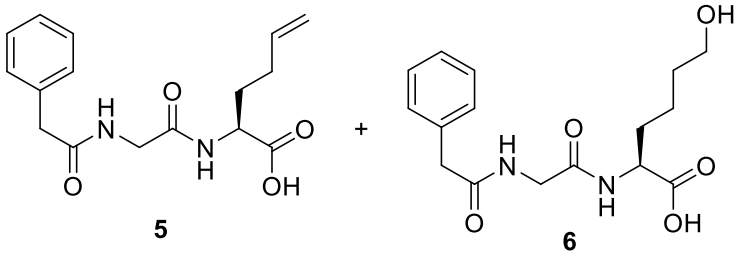
Peak **Figure S10.** LC-MS Data for **4**/**5** Alkene/Alcohol Mixture

A

RT=2.45m

Alcohol





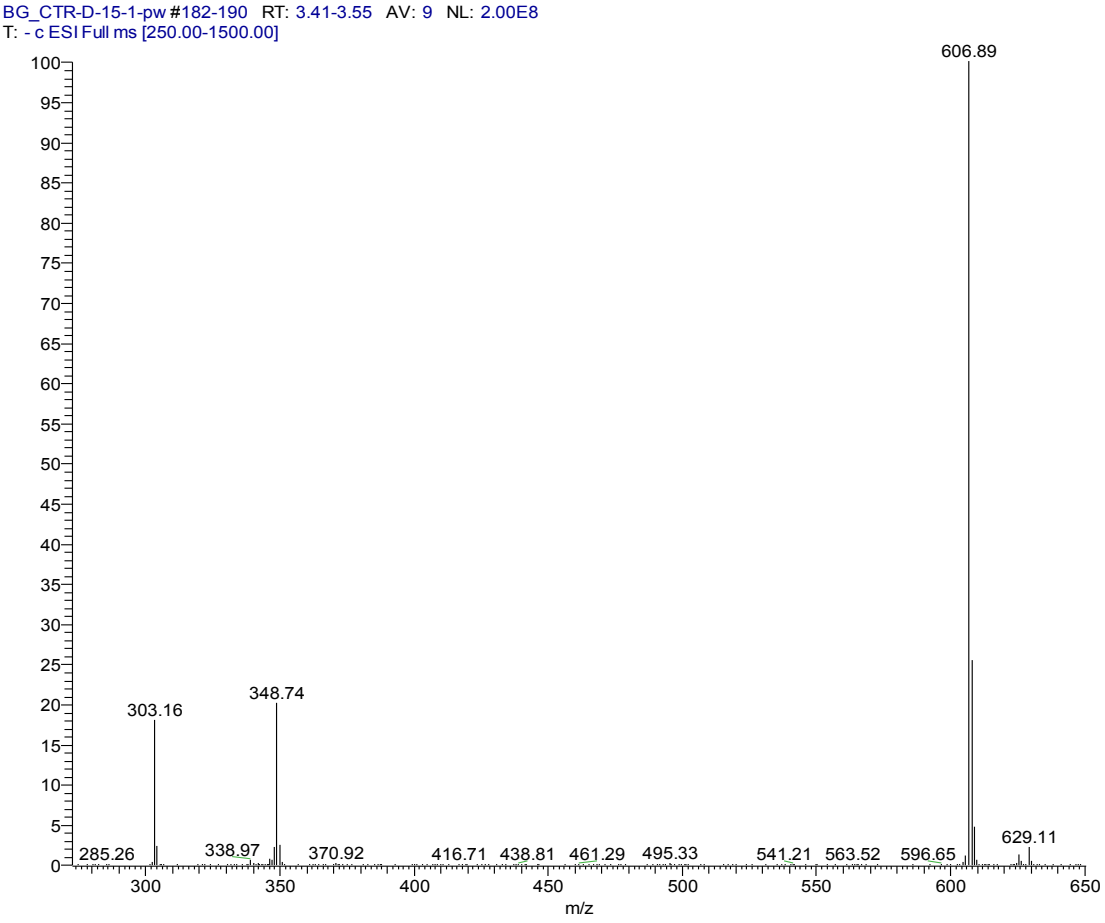
5 **6**

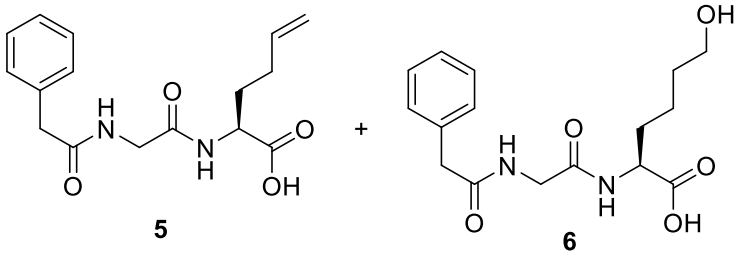
Peak **Figure S11.** LC-MS Data for **4**/5 Alkene/Alcohol Mixture

B

RT=3.50m

Alkene

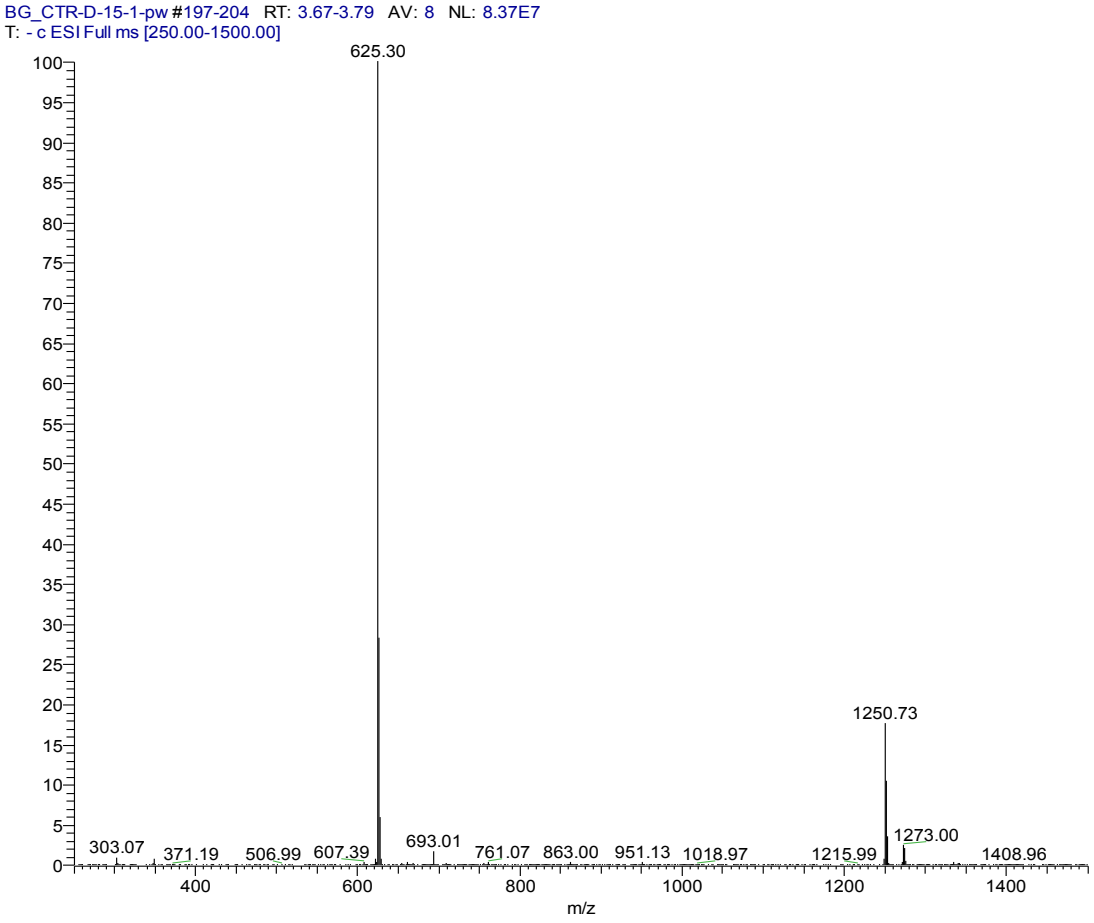




5 **6**

Peak **Figure S12.** LC-MS Data for **4**/**5** Alkene/Alcohol Mixture

C



Peak **Figure S13.** LC-MS Data for **4**/**5** Alkene/Alcohol Mixture

D

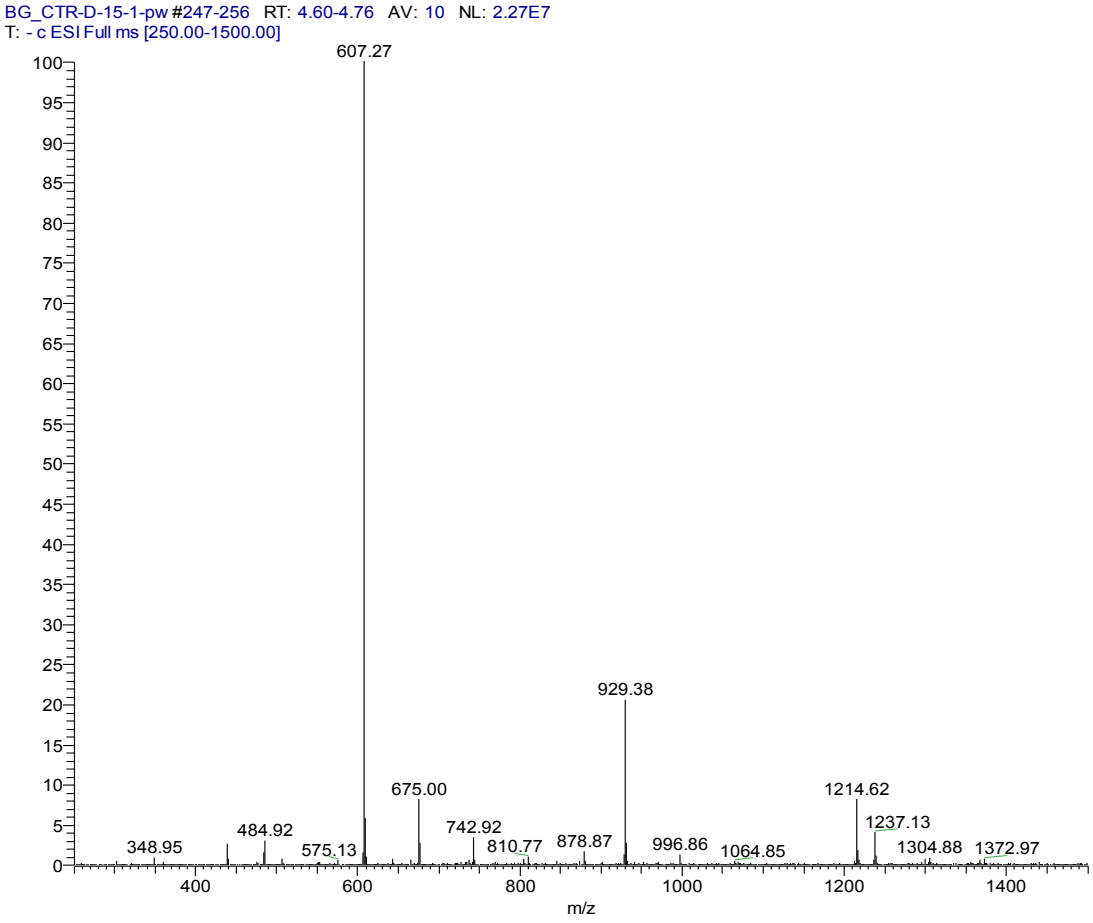
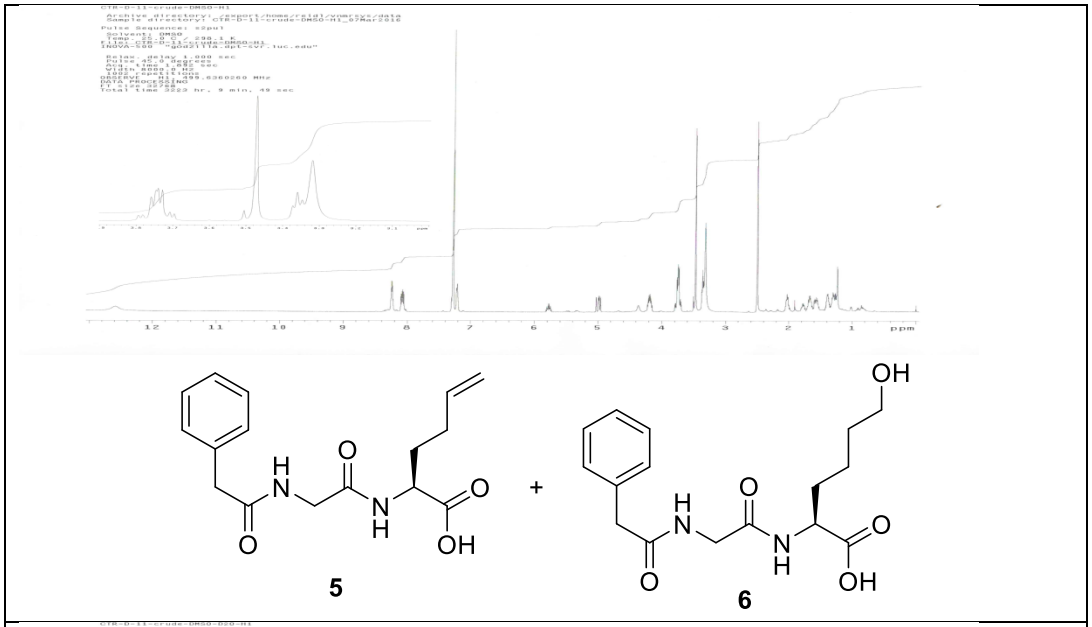


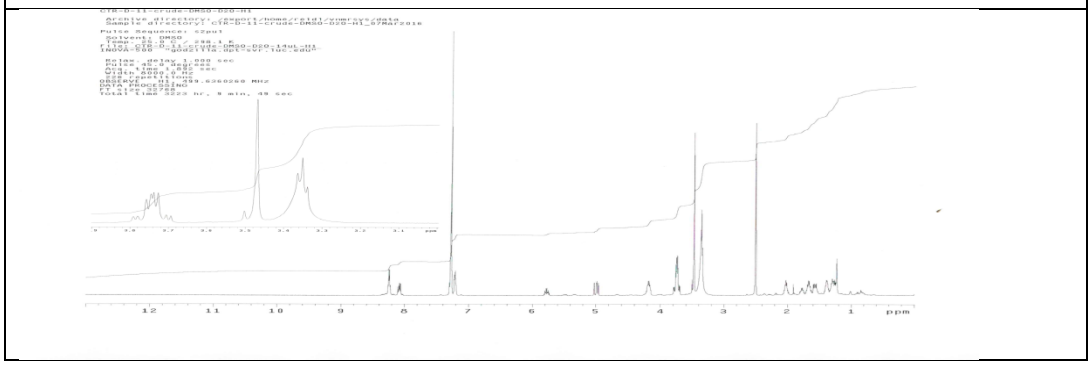
Figure S14. ^{1}H NMR of D_2O Exchange on **4/5** Mixture

^{1}H NMR:
7.5mg **4/5** in
750uL DMSO



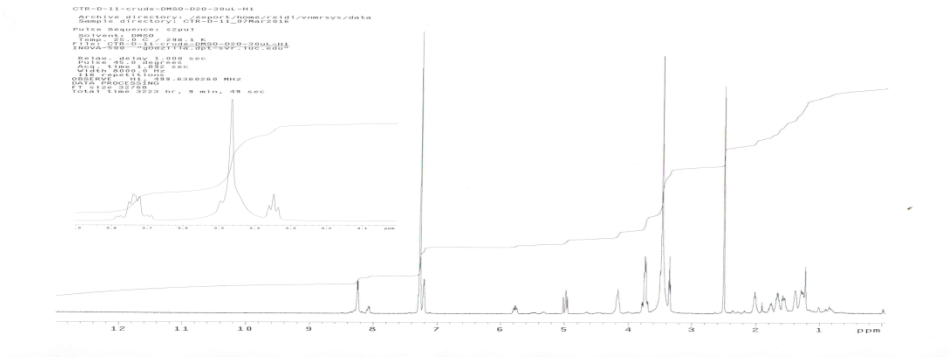
5 + **6**

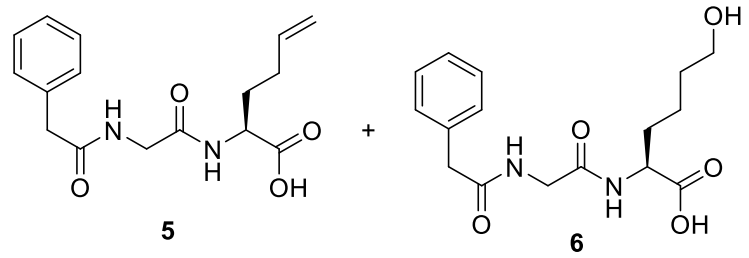
^{1}H NMR:
7.5mg **4/5** in
750uL DMSO
+ 14uL D2O



^{1}H NMR:
7.5mg **4/5** in
750uL DMSO
+ 30uL D2O

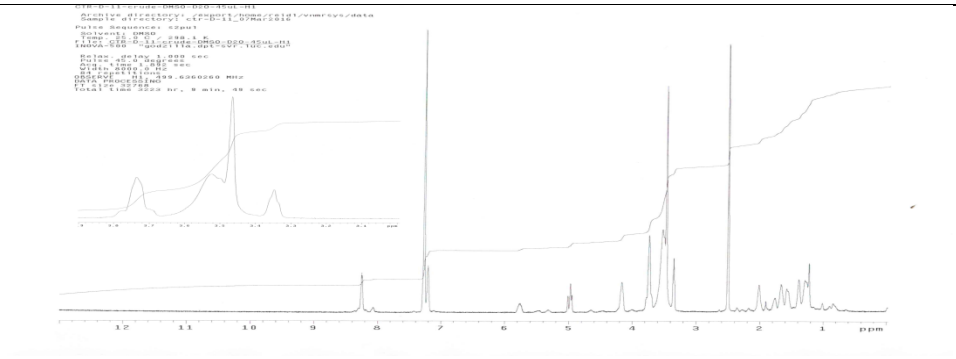
Figure S15. D_2O Exchange on **4**/**5** Mixture





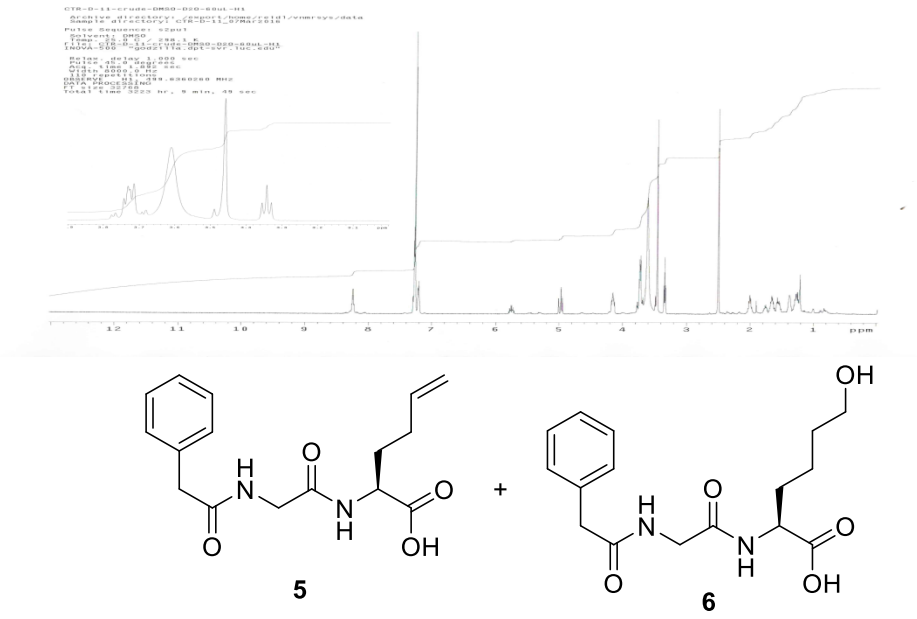
5 **6**

^{1}H NMR:
7.5mg **4/5** in
750uL DMSO
+ 45uL D2O



^{1}H NMR:
7.5mg **4/5** in
750uL DMSO
+ 60uL D2O

Figure S16. D_2O Exchange on **4**/**5** Mixture



5 + **6**

COSY NMR:
7.5mg **4/5** in
750uL DMSO

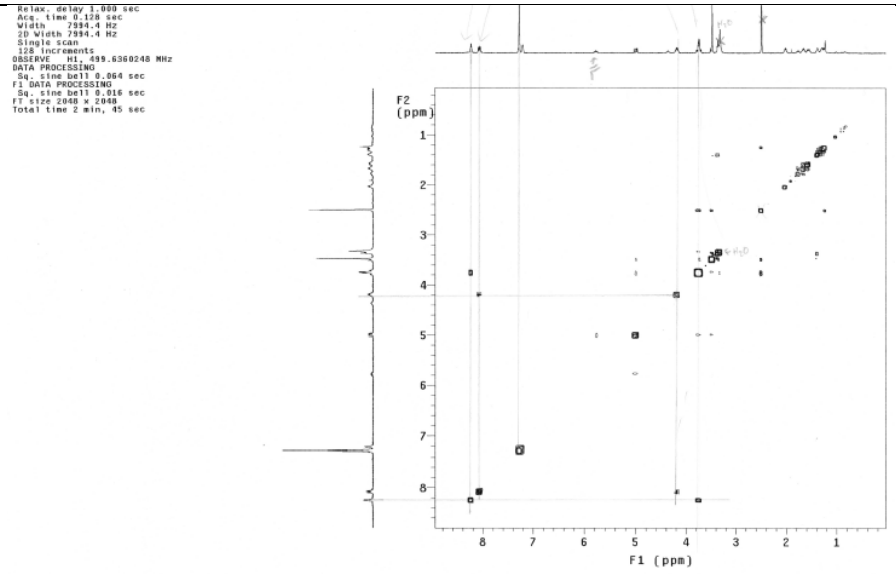
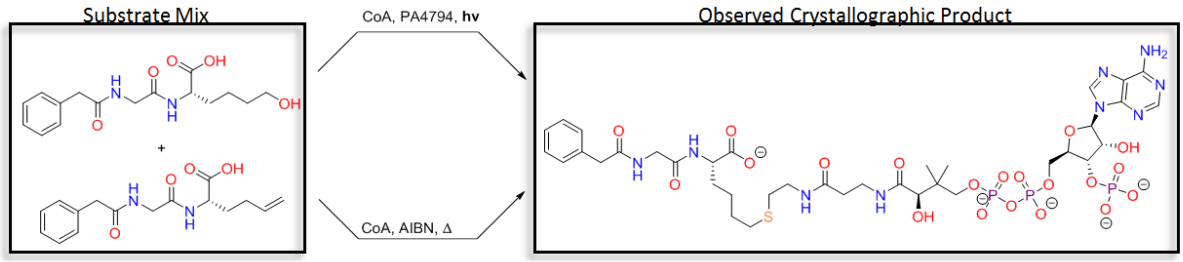


Figure S17. [A] Discontinuous assay demonstrating C-S bond formation. Assay result HPLC Chromatogram for [B] Control: **4**/5 Mix + Buffer. [C] Control: HSCoA + **4**/**5** Mix + Buffer. [D] Trial 1: High conc. Enzyme + HSCoA + **4**/5 Mix [E] Trial 2: Low conc. Enzyme + HSCoA + **4**/**5** Mix + Buffer. [F] Trial 3: HSCoA + **4**/**5** Mix + AIBN + Buffer.[5]

[A]



[B]

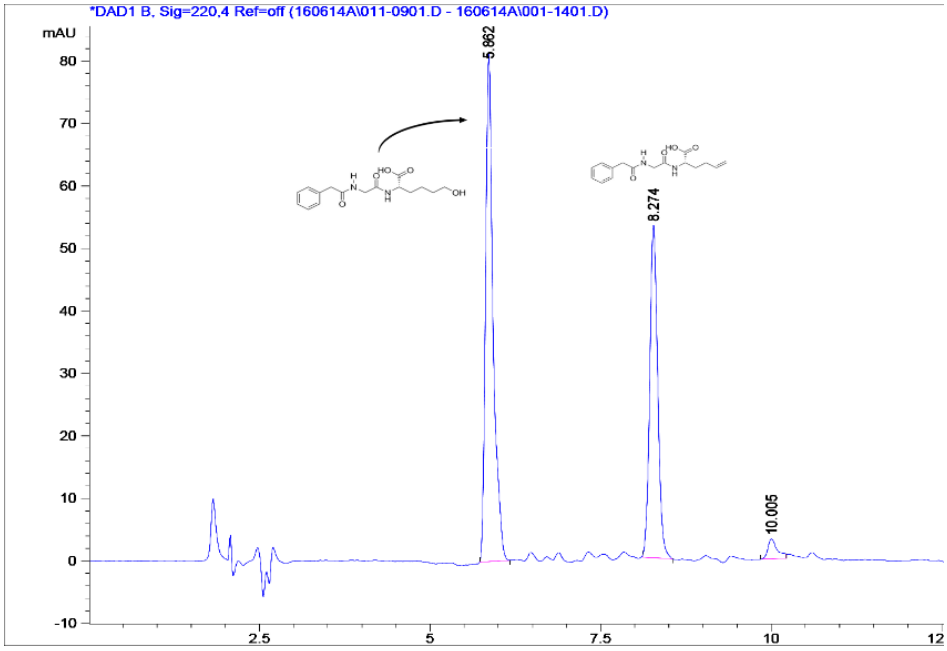
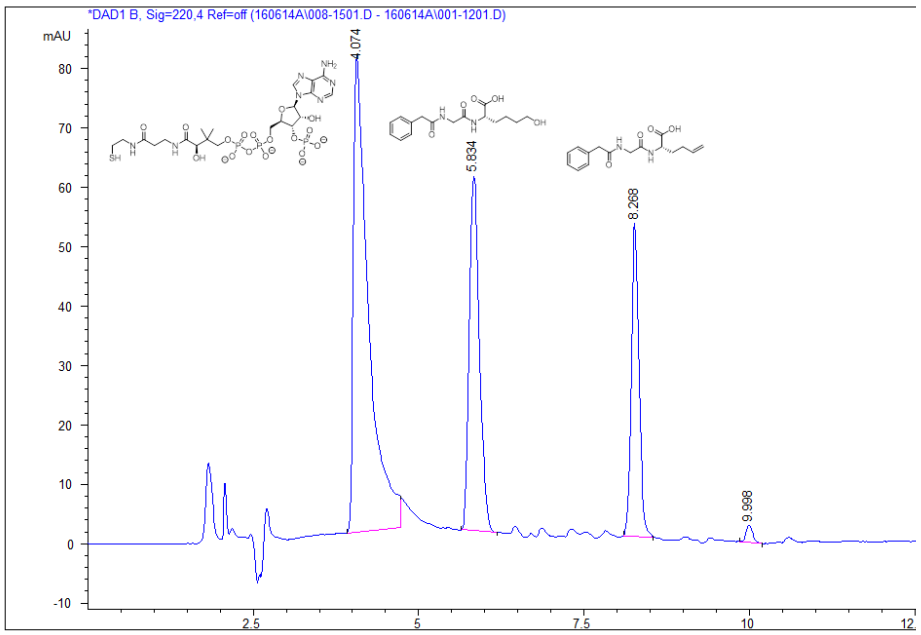


Figure S18. [A] Discontinuous assay demonstrating C-S bond formation of bisubstrate **6**.

[C]



[D]

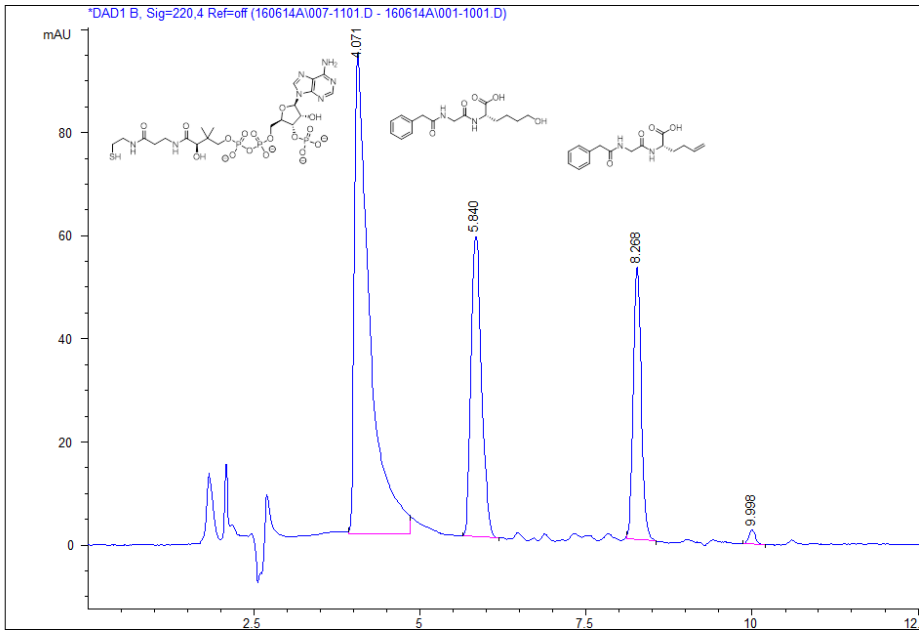
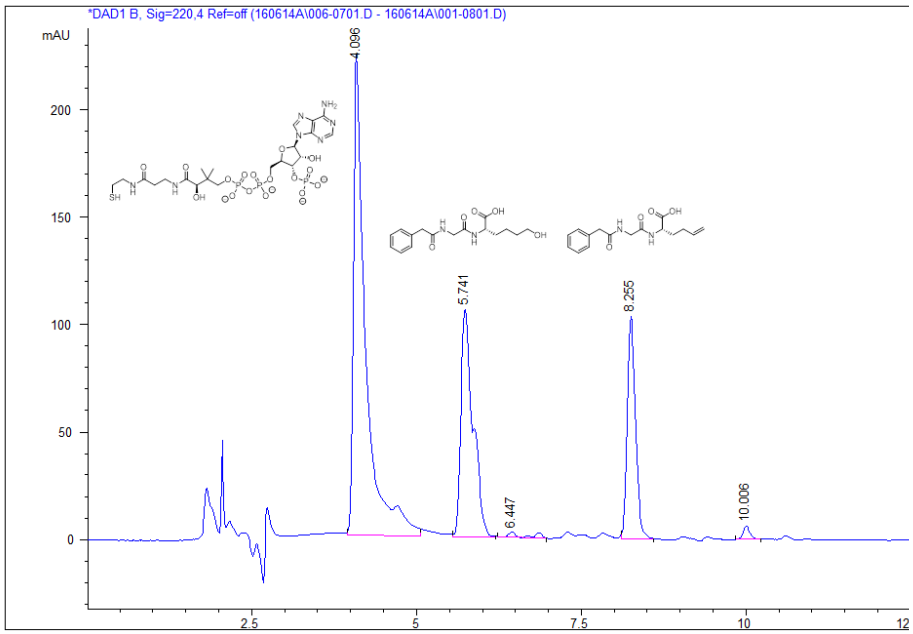


Figure S19. [A] Discontinuous assay demonstrating C-S bond formation producing bisubstrate **6**

[E]



[F]

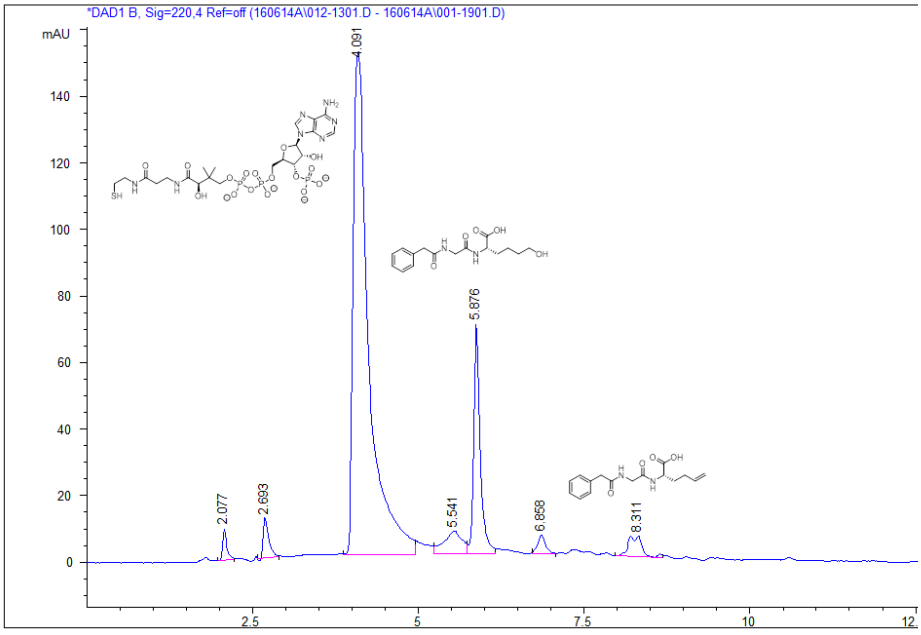
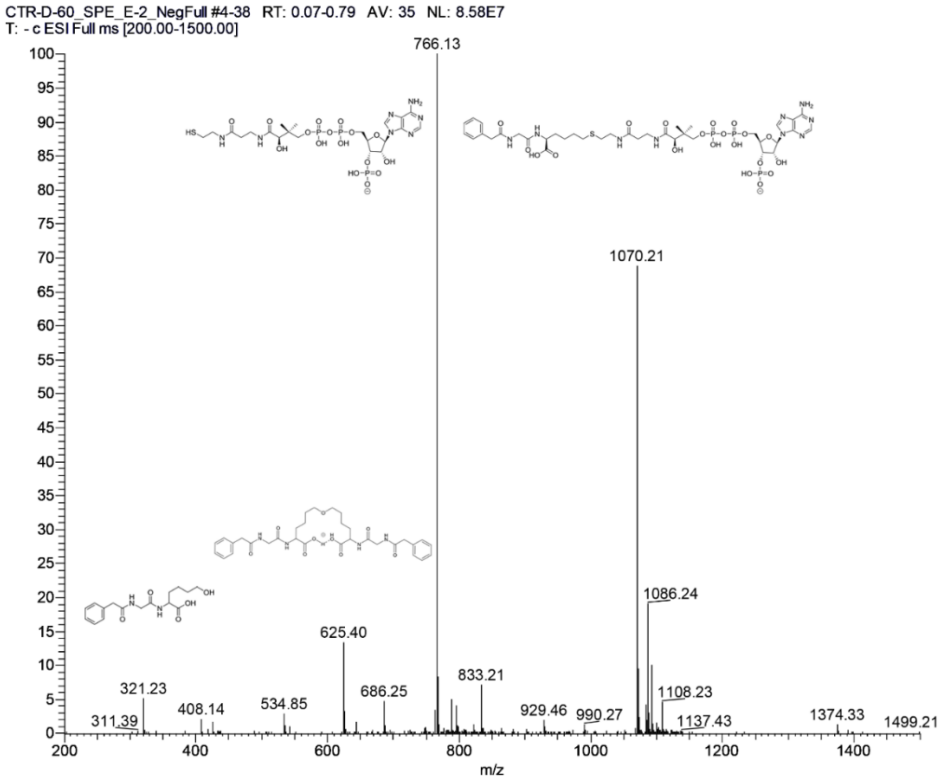
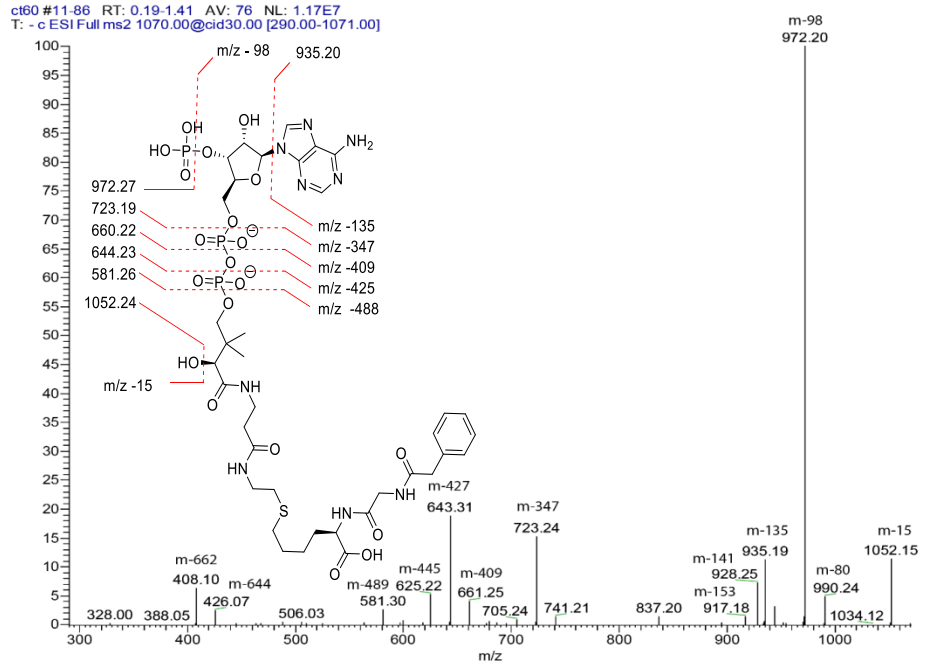


Figure S20. A) MS and B) MS-MS Fragmentation Analysis of bisubstrate **6** produced by radical addition to alkene, after Solid Phase Extraction (SPE) Preparation.

[A]



[B]



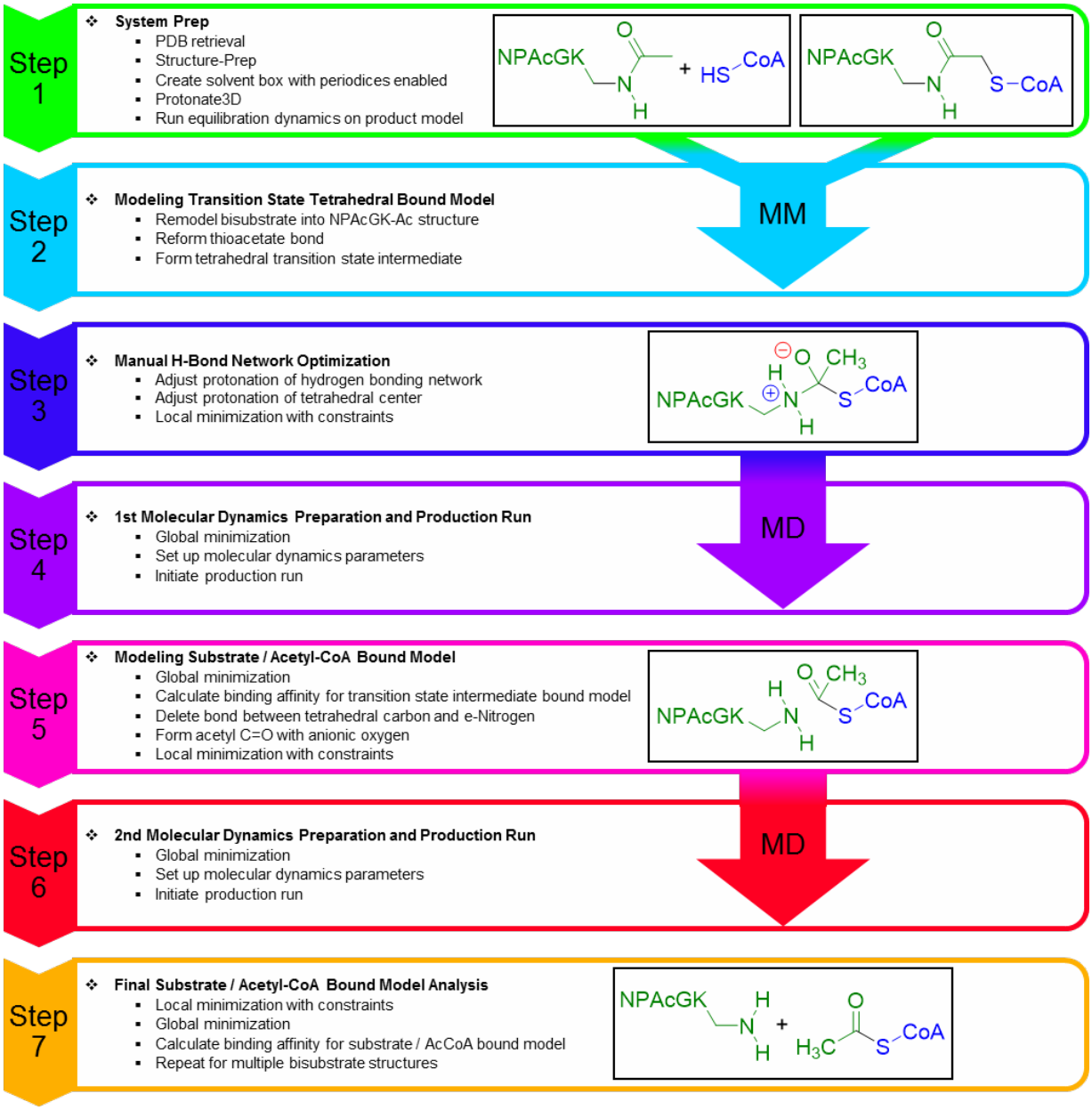


Figure S21. Product-Based Transition-State Modeling (PBTSM) Expanded Description. Protocol for systematic remodeling of tetrahedral intermediate and substrate binding modes from product or bisubstrate bound crystal structures. Molecular minimization (MM) and molecular dynamics (MD) stages in the protocol are shown. Further details are presented in Experimental Procedures.

Reference

1. Filippova, E. V.; Kuhn, M. L.; Osipiuk, J.; Kiryukhina, O.; Joachimiak, A.; Ballicora, M. A.; Anderson, W. F. A Novel Polyamine Allosteric Site of SpeG from Vibrio cholerae Is Revealed by Its Dodecameric Structure. *J. Mol. Biol.* **2015**, *427,* 1316-1334.