

The Development and Application of Small Molecule Probes for the  
Identification and Characterization of Receptor Targets and the Total  
Synthesis of the Caeliferins, Elicitors of Plant Immune Response

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Inish Michael O'Doherty

January 2014

UMI Number: 3579133

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3579133

Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

© 2014 Inish Michael O'Doherty

The Development and Application of Small Molecule Probes for the Identification and Characterization of Receptor Targets and the Total Synthesis of the Caeliferins, Elicitors of Plant Immune Response

Inish O'Doherty, Ph.D.

Cornell University 2014

How a biologically active small molecule ultimately achieves its phenotype is of foremost interest to chemical biology and pharmaceutical research. The majority of endogenous, environmental, or drug-like signaling molecules elicit their effect through the binding to a proteinaceous receptor. Developing a detailed molecular understanding of ligand-receptor interactions is key to building our understanding of biological systems and furthering drug discovery.

Current bottom-up chemical genetic approaches are capable of indicating the identity of a small molecule receptor but frequently a deeper understanding of the pathway is required to fully describe the role the candidate receptor plays. To achieve this, *in vitro* analysis of the ligand-receptor interaction is required. However by abstracting the receptor from the biological system the relevance and recapitulation of the binding may be lost.

Advancements in mass spectrometry and bio-orthogonal reactions that can occur in biological system without perturbing the native biological processes have allowed the development of non-biased cell-based top-down, identification at the protein level, ligand receptor identification strategies of Activity Based Protein Profiling (ABPP) and Affinity Based Protein Profiling (AfBPP). Although extremely powerful in their application these top-down methods have largely excluded the analysis of lipophilic transmembrane receptors, due to their biochemical incompatibility with the purification techniques required.

The author of this dissertation describes the development of two chemical biology strategies that apply novel top-down receptor identification techniques capable of characterizing lipophilic membrane receptors. The first is a small-molecule candidate receptor screening strategy which demonstrates the direct binding of a bioactive photoaffinity probe of the *C. elegans* metabolite ascr#2 to its GPCR DAF-37, using a customized light amplification assay. The second is development of a methodology for the deconvolution and identification of ligand-receptor binding sites using cleavable stable-isotope click-chemistry tags (CSICT). This CSICT can be partnered with any clickable AfBPP or ABPP probe for binding site identification. These two new methods broaden the scope of ligand-receptor identification and characterization and allow the exploration of previously unexamined transmembrane receptors with top-down techniques.

The caeliferins are a family of insect-derived sulfoxy fatty acids that play an important role in plant defenses. In this dissertation the author described the total synthesis of this family of compounds using olefin cross metathesis (CM). Detailed NMR spectroscopic and mass spectrometric analyses of CM reaction mixtures revealed extensive isomerization and homologation of starting materials and products. It is shown that the degree of isomerization and homologation in CM strongly correlates with substrate chain length and lipophilicity. This observation has implications towards the use of CM in the total synthesis of natural products, as catalyst selection must be considered to ensure the synthesis of pure, isomer-free compound samples for accurate biological testing.

## BIOGRAPHICAL SKETCH

Inish Michael O'Doherty was born in Dublin, Ireland where he attended Newpark Comprehensive School, from which he received his Leaving Certificate. Inish then set out on a year-long world exploration prior to attending the University College Dublin (UCD). At UCD, Inish conducted research with Professors Donal O'Shea and Patrick Guiry, Department of Chemistry. In May of 2007, Inish graduated top of his class and received his Bachelor of Science degree in Chemistry from UCD. In the summer of 2007, Inish entered the Tri-Institutional Training Program in Chemical Biology (TPCB). After completing rotations in the labs of David Y. Gin and Derek Tan, Inish moved to Ithaca, New York and matriculated with the Department of Chemistry and Chemical Biology, Cornell University. Inish conducted his graduate research in the laboratory of Professor Frank C. Schroeder. During his time at Cornell, Inish received the O'Reilly Foundation Scholarship for his excellent academic record and desire to make significant contribution to the future of Ireland.

## ACKNOWLEDGMENTS

**Special Committee Chair and Adviser:** Professor Frank Schroeder.

**Collaborating Principal Investigators:** Professor Donald L. Riddle, Professor Ujendra Kuma, and Dr. Eric A. Schmelz.

**Collaborating Researchers:** Dr. Donha Park, Dr. Rishi K. Somvanshi, Dr. Axel Bethke, Josh A. Baccile, Joshua J. Yim, and Daniel Gentile

**Special Committee:** Professor Eric Richards, Professor Hening Lin, Professor Minkui Luo

**Facility Managers:** Dr. Ivan Keresztes, Dr. Sheng Zhang, and Dave Kiemle

**Organizations:** Department of Chemistry and Chemical Biology and The Boyce Thompson Institute for Plant Research, Cornell University; Michael Smith Laboratories, University of British Columbia; United States, and Center for Medical, Agricultural, and Veterinary Entomology, United States Department of Agriculture, Gainesville, FL.

**Funding:** Training Program in Chemical Biology Memorial Sloan-Kettering Cancer Center, Rockefeller University, and Cornell University; O'Reilly Foundation, Dublin, Ireland; Schroeder Laboratory Startup Funds, Cornell University.

## TABLE OF CONTENTS

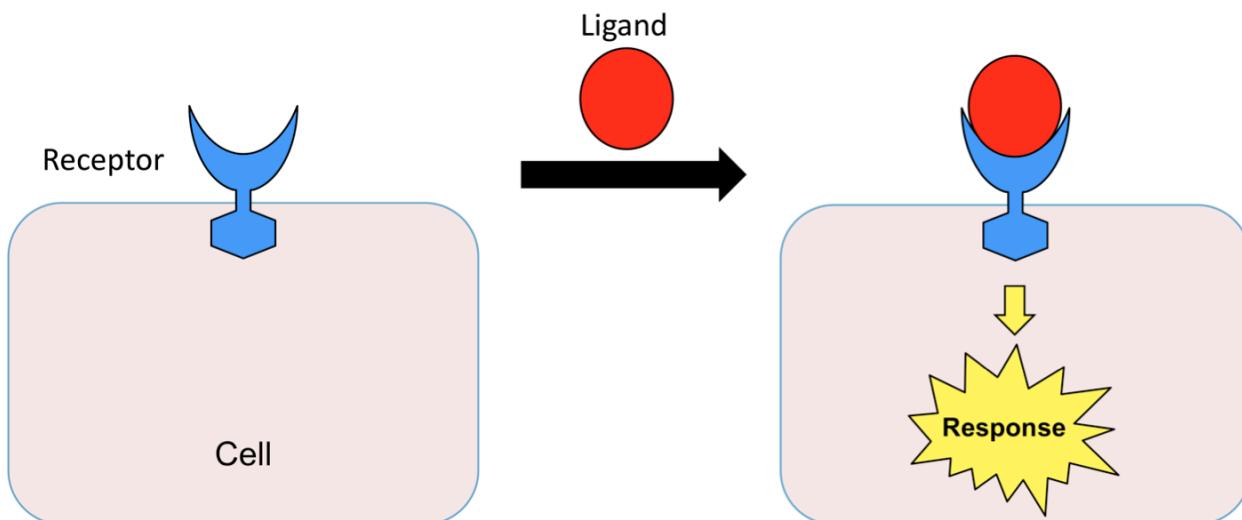
<u>Section</u>	<u>Page</u>
Biographical Sketch.....	iii
Acknowledgments.....	iv
Table of Contents.....	v
Preface.....	vi
Chapter 1; Interaction of structure-specific and promiscuous G-protein coupled receptors mediates small molecule signaling in <i>C.elegans</i> .....	1
Chapter 2; Advancements towards the differential labeling of small molecule binding sites using photoaffinity probes coupled to cleavable stable isotope click-chemistry tags (CSICT).....	28
Chapter 3; Synthesis of the caeliferins, elicitors of plant immune response: accessing lipophilic natural products via cross metathesis.....	61
Appendix A; Interaction of structure-specific and promiscuous G-protein coupled receptors mediates small molecule signaling in <i>C.elegans</i> .....	71
Appendix B; Advancements towards the differential labeling of small molecule binding sites using photoaffinity probes coupled to cleavable stable isotope click-chemistry tags (CSICT).....	93
Appendix C; Synthesis of the caeliferins, elicitors of plant immune response: accessing lipophilic natural products via cross metathesis.....	140

## PREFACE

**The Development of Techniques to Better Understand Small-Molecule Receptor Binding.** Understanding the interaction between signaling molecules and their receptors has long been the goal of both applied and basic chemical research. The desire to deconvolute<sup>1</sup> how a chemical signal connects different processes and locations within a single organism, or among organisms of the same or several different species has driven academic and pharmaceutical research<sup>2</sup> since their inception. The interest begins with a molecule and the observation of a particular phenotype or biologically phenomena that is correlated to the presence of this compound. Once the identity of the molecule has been established, through analytical characterization and synthetic replication of the chemical entity, the question quickly turns to how this molecule could be eliciting the phenotype<sup>2</sup>.

Proteins are the primary point of contact between an organism and biologically active small molecules. The initial discrete molecular binding event between these two entities sets in motion a signaling cascade that will ultimately result in the phenotype of interest (**Figure P1**). An investigation of the initial binding event and the downstream signaling cascade needed to propagate this signal is required to understand fully the pathways that mediates the compound dependant phenotype. Chemical genetics<sup>3-8</sup> is one area of research that has been used to investigate both the initial perception of a molecule and the downstream pathways that facilitate the biological response<sup>9</sup> and falls under the larger umbrella of reverse genetics. This form of chemical genetics is not to be confused with the technique of the same name, where genetic mutations are non-specifically introduced to organism via chemical mutagens in an effort to elucidate gene function. The discussion of chemical genetics in this thesis is exclusively referring to the former and not the latter mutagenizing technique. Chemical genetics research builds upon the advancements in genome sequencing, molecular biology, and gene silencing<sup>10,11</sup> that has allowed the specific targeting and removal of proteins from a

biological system, via genetic manipulation. With the protein(s) of interest removed or inhibited, the altered biological system can be examined for a differential or loss of phenotype in response to elicitor treatment. If a change in phenotype is observed in this augmented system, the protein is inferred to be part of the signaling pathway (**Figure P.2**). Knowing that



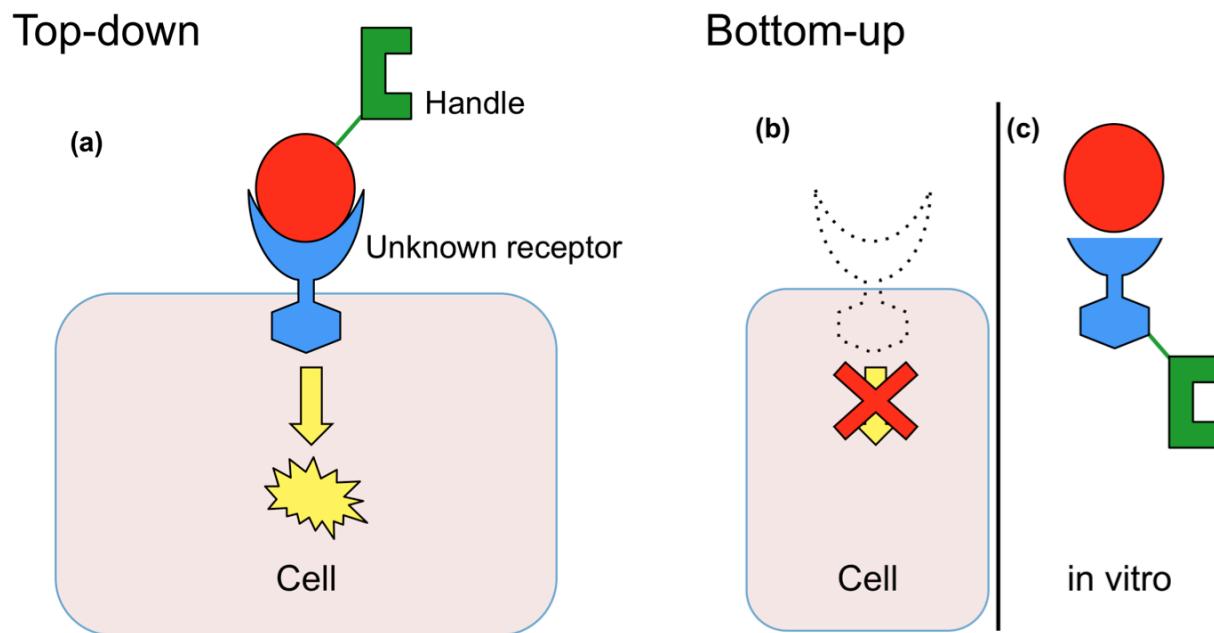
**Figure P.1: Cellular responses to ligands are mediated by cell surface receptor proteins.** The physical interaction between an extra-cellular ligand and a cell surface receptor initiates a signaling cascade within the cell that ultimately result in a change in cellular behavior.

the protein is involved in the signaling cascade is only the first step, and its relative position and function in the pathway must then be described next. By comparing the sequence of the identified protein against well-established databases<sup>12</sup>, a putative function and even potential placement in the biological pathway maybe attributed<sup>2-4</sup>. Although such bottom-up, genes to protein, approaches are capable of indicating the identity of a small molecule receptor involved in perception, frequently a deeper understanding of the pathway is required to describe fully the role of the protein. To achieve this deeper knowledge, it is necessary to carry out *in vitro* analysis along with

further genetic manipulations of other players in the pathway (**Figure P.2**). This analysis can be particularly challenging in novel or previously undescribed system<sup>13</sup>. The merits and limitations of these techniques are discussed in greater detail in Chapter 2. In short, the biggest challenge with a chemical genetic identification process is that binding of a small molecule to a putative receptor can only be inferred indirectly. When the candidate receptor is genetically removed from the biological system of interest and a loss of compound specific phenotype is observed, this receptor can be attributed to the perception of this small molecule, even without direct binding evidence. Chemical genetic analysis therefore depends on additional *in vitro* assays to show that there is a direct interaction between the binding partners. These assays are typically carried out with purified recombinant proteins or protein fragments that have been removed from a cellular environment or native system. As a result these purified protein elements may not have the correct regulatory domains, conformational folding, co-factors, or post-translational modifications required to recapitulate the binding event<sup>2,14</sup>. (**Figure P.2**).

To overcome these obstacles top-down, ligand-based identification strategies have been developed to analyze ligand-receptor binding *in vivo* or in cell-based models<sup>3,5</sup>. These top-down approaches use modified ligands, in the form of molecular probes that retain the ability to bind to the receptor but have additional functionality (**Figure P.2**). Such modified probes of a bioactive small-molecule are usually designed to be tri-functional, as they retain the ability to bind the cognate receptor<sup>15</sup> and once bound have the ability to covalently crosslink to the protein establishing a probe-receptor complex<sup>16,17</sup>. The third functionality, enabling pull-down of the probe-receptor complex, usually consists of a small chemical handle that can undergo a specific and selective bio-orthogonal reaction<sup>18-22</sup>. This reaction allows the installation of a purification or detection tag to the probe-receptor complex and is termed bio-orthogonal as it does not react with any endogenous chemical species in the biological system. If a purification tag such as biotin is selected, the probe-receptor complex can be fished out

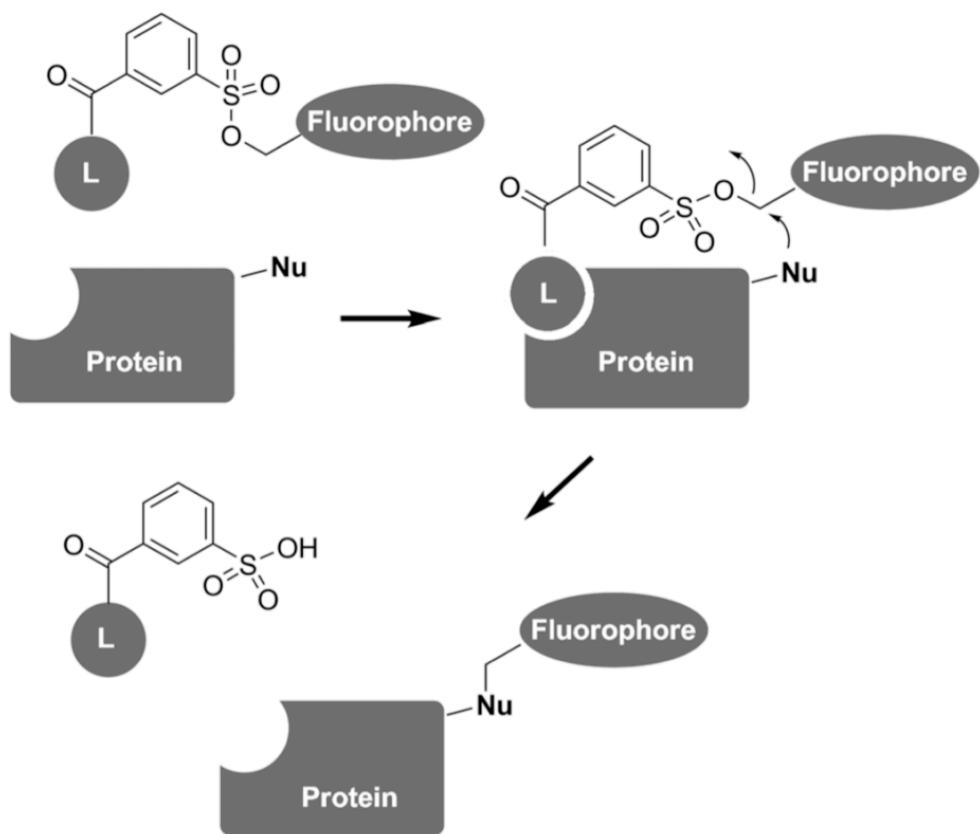
of a crude protein lysate using streptavidin beads<sup>23,24</sup>. This enriched probe receptor complex can then be eluted, trypsinized and identified and characterized by mass spectrometric methods<sup>23-25</sup>. Alternatively, a fluorogenic or epitope tag can be linked to the probe-receptor complex facilitating the detection of the labeled receptor with SDS PAGE/western blot analysis<sup>23,26</sup>. This top-down methodology provides a means for the investigation of ligand receptor interactions in more complex and biologically relevant environment than the *in vitro* methods used in chemical genetic screening<sup>2,14</sup> (**Figure P.2**).



**Figure P.2: Top-down versus bottom-up approaches for ligand-based receptor identification.** In a top-down scheme (a), a known ligand is derivatized into a probe such that it can covalently bind the receptor in *in vivo*/cell-based assay. The probe is also decorated with “handle” that contains moieties enabling the capture, purification and subsequent identification of the probe-receptor complex. In a bottom-up receptor identification strategy (b and c), proteins/receptors are first genetically silenced and screened for a loss of function in response to an elicitor (b). Candidate receptors or simplified protein elements from them are then recombinantly expressed and purified with a genetically encoded epitope purification handle. The purified receptors are then tested for *in vitro* binding assays (c). As demonstrated in (c), recombinant proteins may not be have the correct structure, binding partner or modifications to recapitulate *in vivo*/cell-based binding.

Two of the most prominent developments that have lead to the expansion of top-down receptor identification strategies from niche applications towards more general adoption, is the development of compact and reliable photocrosslinking groups<sup>27,28</sup> and the evolution of bio-orthogonal click-chemistry reactions, most notably the Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC)<sup>29-31</sup>. In the absence of a photocrosslinking group on a small molecule probe, the probe itself must be capable of binding the receptor covalently without any external stimulus. Typically in these cases, the protein that binds the small-molecule probe must be capable of presenting a nucleophilic amino acid at or proximal to the ligand binding site<sup>32</sup>. This amino acid side chain then undergoes a reaction with the electrophilic moiety on the probe, resulting in covalent inhibition of the protein. The establishment of this covalent linkage between the small molecule and the protein allows the click chemistry mediated labeling of the probe-receptor complex for subsequent detection. This covalent capture of receptor proteins, coupled to click chemistry based detection strategies is termed Activity Based Protein Profiling (ABPP)<sup>16,32,33</sup>. ABPP has been very successful in the identification of proteolytic enzymes that have nucleophilic residues at their active sites<sup>32</sup>. However, to apply ABPP beyond specific enzyme subclasses with reactive amino acids in their active site, requires the introduction of a highly reactive electrophilic functional group, such as a tosyl<sup>26</sup> to the probe. This electrophilic centre is then primed to react with any nucleophilic amino acid it comes in contact with. The goal is that the probe's molecular specificity will be sufficient to enhance its localization to the receptor of interest where it will be able to couple to a reactive nucleophilic amino acid side-chain next to the binding pocket<sup>3</sup> (**Figure P.3**).

This version of ABPP suffers significant drawbacks. The first is that there is a necessity for a nucleophilic residue being correctly positioned proximally to the small-molecule binding pocket of the receptor, this fortuitous positioning can't always be considered a certainty. The second is that by placing an electrophilic center on a small-molecule there will inherently be significant amounts of background labeling. If the protein/receptor one tries to

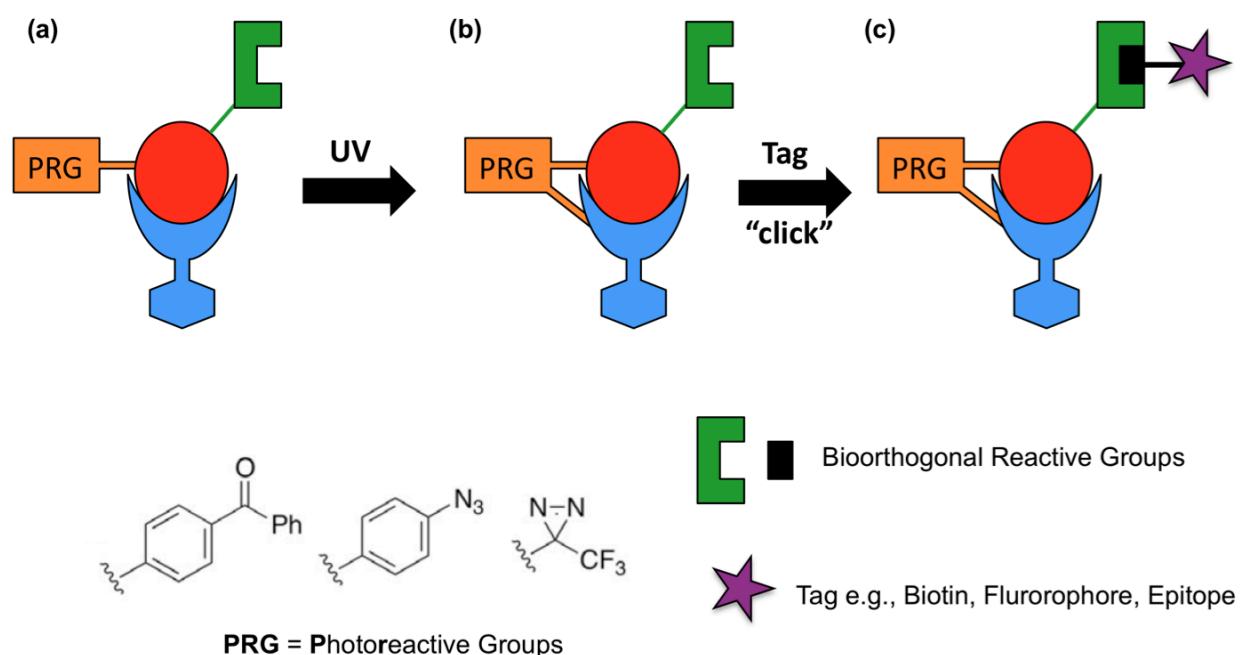


**Figure P.3: Ligand-directed protein labeling using tosyl chemistry, proximal binding site labeling.** Adapted from Ziegler *et al.* The bioactive tosyl containing ABPP ligand, L binds the receptor protein of interest. Subsequently a nucleophilic residue on the protein attacks the tosyl group resulting in the labeling of the receptor protein with a detection tag, in this case a fluorophore.

identify is present at high levels, this may not be a problem but if the receptor is present at medium to low levels, which many perception receptors are, the dynamic range of detection either via MS or in-gel methods maybe not be sufficient to detect the target protein from the non-specific background labeling. To overcome the issue of “always on” reactivity of generic electrophilic centers, photoreactive functional groups have been developed that upon specific UV irradiation trigger the generation of a highly reactive, shorted lived chemical species<sup>14,27,34</sup>. By incorporating a photoreactive group into a bioactive small molecule, it is possible to generate a chemical probe that will bind to the receptor of interest and then, upon UV treatment, generate a transient highly reactive species that can crosslink the probe to the receptor covalently<sup>15</sup> (**Figure P.4**). In the case where probe is not bound to protein, the transience of the reactive species ensures that the probe will react with its solvation shell before coming into contact with another protein, thereby minimizing non-specific crosslinking<sup>27</sup>. This photoaffinity labeling of receptors coupled to click chemistry based detection strategies is termed Affinity-Based Protein Profiling (AfBBP) and suffers from much lower non-specific labeling than ABPP<sup>16,32,33</sup>, previously discussed (**Figure P.4**).

Although multiple photoreactive groups (PGR) have been used in labeling experiments the trifluoromethylaryldiazirines<sup>27,34</sup> have distinguished themselves as the PRG of choice due their compact nature, chemical stability and biologically compatible photoirradiation profile (**Figure P.4**). While both ABPP and AfBBP probes have the ability to bind covalently to the candidate receptor, they must also provide a chemical handle that is capable of undergoing a highly efficient and selective bio-orthogonal reaction for the installation of a purification or detection tag. Suitable reactions for this include the Staudinger-Bertozzi ligation<sup>18</sup>, the copper catalyzed Huisgen 1,3-dipolar cycloaddition (click reaction)<sup>16,19,29,35</sup>, the copper-free click reaction variants<sup>20,21,36,37</sup> and the tetrazine ligation<sup>22</sup>. Although all of these reactions have proven merits, the Cu-

mediated click reaction has seen the greatest use due to the structural simplicity of the reactive partners, a terminal alkyne and an azide. The small structural size of these groups allows them to be incorporated into small-structurally sensitive signaling molecules without disrupting their binding activity<sup>15</sup>.



**Figure P.4 Schematic of photoaffinity labeling-bioorthogonal conjugation.** Adapted from Lapinsky *et. al.* Initially, the tri-functional probe possessing a photoreactive group and a ‘clickable’ handle/bioorthogonal-chemical reporter binds the receptor (**a**). With specific UV stimulus, the PRG covalently labels the target receptor, establishing the covalently bound probe-receptor complex (**b**). Using highly specific bioorthogonal click-chemistry, a detection or purification tags is installed to the probe-receptor complex (**c**), allowing further identification and analysis.

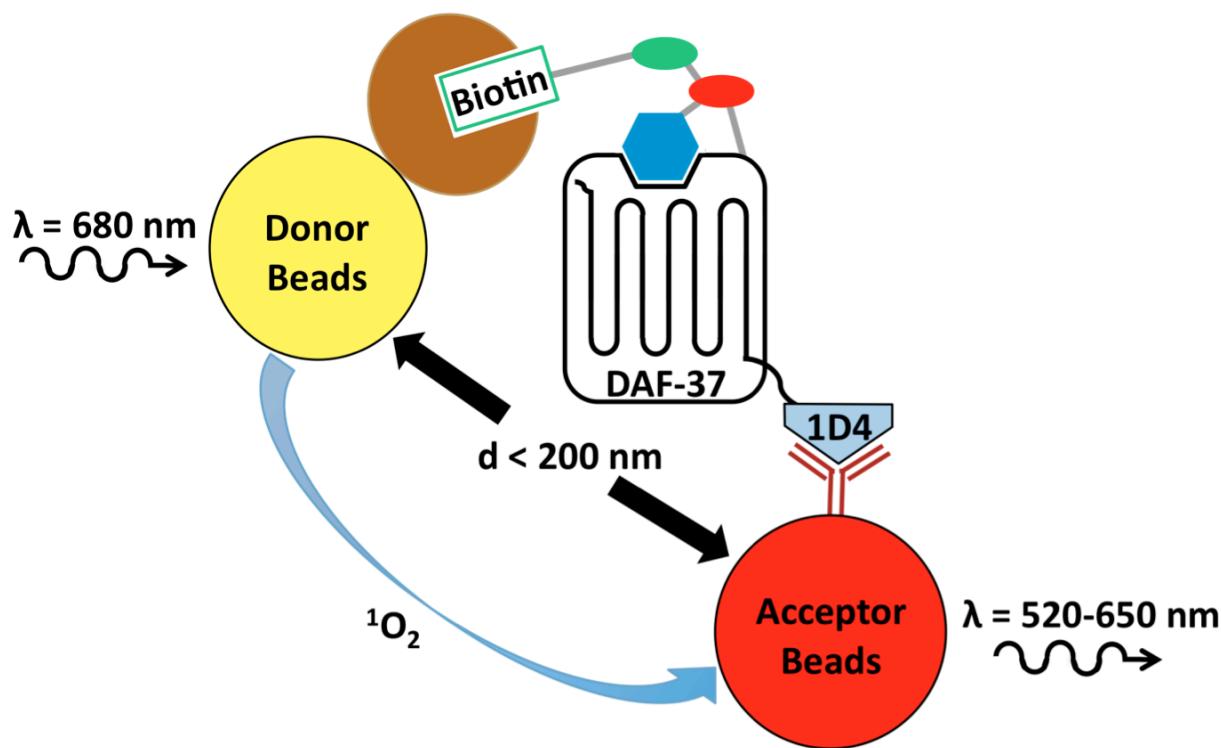
By placing one of the bio-orthogonal reactive partners, the alkyne, on the probe and the complimentary azide functionality on a purification handle, the probe-receptor complex can be labeled specifically with an affinity tag such as biotin in a Copper-catalyzed azide-alkyne cycloaddition (CuAAC)<sup>29-31</sup>. This biotinylated complex can then be enriched onto streptavidin beads, eluted and subjected to shotgun proteomics

analysis to identify the unknown receptor<sup>23,24</sup>. Recent literature has shown that although the alkyne and azide can both be considered bio-orthogonal, placement of the alkyne on the probe and the azide on the biotin yields better results than the reverse<sup>38</sup>. Click-chemistry provides a unique method of incorporating a small non-structurally disruptive chemical handle to the probe that can be readily derivatized into a much larger chemical entity for detection and purification of the receptor. Without this bio-orthogonal technique, the large detection functionality would have to be introduced to the probes structure, disrupting the biological activity many biologically active small molecules. These click-chemistry methods coupled to the strength and selectivity of the biotin-streptavidin interaction<sup>38-40</sup> have revolutionized the field of ligand receptor deconvolution<sup>23,24</sup>. The biotin-streptavidin coupling allows the stringent washing and elimination of unlabeled, non-specific background proteins; these washing steps streamline and enhance the downstream MS receptor identification and characterization process. However, the strength of the biotin-streptavidin interaction also is a major weakness of this method. Typically, elution from streptavidin beads is carried out at 95°C with strong ionic detergents<sup>38</sup>, chaotropic salts<sup>41</sup>, and reducing agents. These elevated temperatures cause solubility problems for lipophilic membrane proteins, including G-protein Coupled Receptors (GPCRs), which make up one of biggest classes of signaling receptors<sup>42</sup>. Multi-pass membrane proteins are particular hydrophobic and exposing them to elevated temperature causes these lipophilic receptors to aggregate, preventing further characterization<sup>43,44</sup>. Along with this elution challenge, biotinylation occurs as an endogenous post-translational modification of proteins in many organisms. Using streptavidin beads to enrich for the biotinylated probe-receptor complex will also cause an enrichment of the naturally biotinylated proteins, that are not involved with probe treatment. In this dissertation, two approaches are described to overcome this limitations; firstly, providing a method for the screening of ligand receptor candidates in a cell-based assay and secondly investigating the receptor biding site for biologically

active small molecules. Both of these approaches utilize photoaffinity labeling and CuAAC coupled to biotin-streptavidin enrichment of probe-receptor complexes. However, they circumvent the biotin elution problem via different means.

In Chapter 1, I describe the development of a small-molecule candidate receptor screening strategy to demonstrate the direct binding of bioactive photoaffinity probe to its GPCR partner using a customized light amplification assay<sup>15,45</sup>. Through a collaboration with the Riddle Lab, working in the model organism *Caenorhabditis elegans*, we discovered that the GPCR DAF-37 was a strong candidate receptor for the endogenous small-molecule signal ascr#2<sup>46</sup>. Ascr#2 is of particular interest in *C. elegans* as it plays pivotal roles in the worms' aging<sup>47</sup>, developmental progression<sup>46,48-50</sup>, and adult behaviours<sup>51,52</sup>, via the highly conserved downstream pathways of insulin, TGF-β, and the sirtuins. This discovery of the ascr#2, DAF-37 ligand-receptor pairing, provided a system to develop a cell-based candidate screening platform using a photoaffinity probe of ascr#2 and recombinantly expressed DAF-37 in human cell culture. Given the lipophilic nature of DAF-37, standard AfBPP methods were not capable of demonstrating direct binding of the probe to the receptor, due to the receptors incompatibility with the elevated temperatures necessary for biotin-streptavidin elution. Moreover, a systematic enrichment of DAF-37 from the cells prior to ligand photolabeling was ill advised, as the functional structure of GPCRs can be easily compromised upon removal from their native membrane environment<sup>43,44</sup>. To overcome these challenges, as described in Chapter 1, an in-cell photoaffinity labeling strategy coupled to a customized amplified luminescence assay<sup>45</sup> (alphascreen) was developed to demonstrate the first direct binding of an ascaroside to its receptor. This alphascreen method does not require the elution of the probe-receptor complex from the streptavidin beads. Instead, the photocrosslinking and subsequent biotinylation of the probe receptor complex can be detected via spatial colocalization of streptavidin coated acceptor beads and acceptor beads that are coated in a 1D4 antibody (DAF-37::1D4

expressed in this assay) as demonstrated in **Figure P.5**. This represents a novel method for screening small-molecule clickable photoaffinity probes and their lipophilic transmembrane receptors using customized biotin-streptavidin based alphasuren technology<sup>15</sup>.

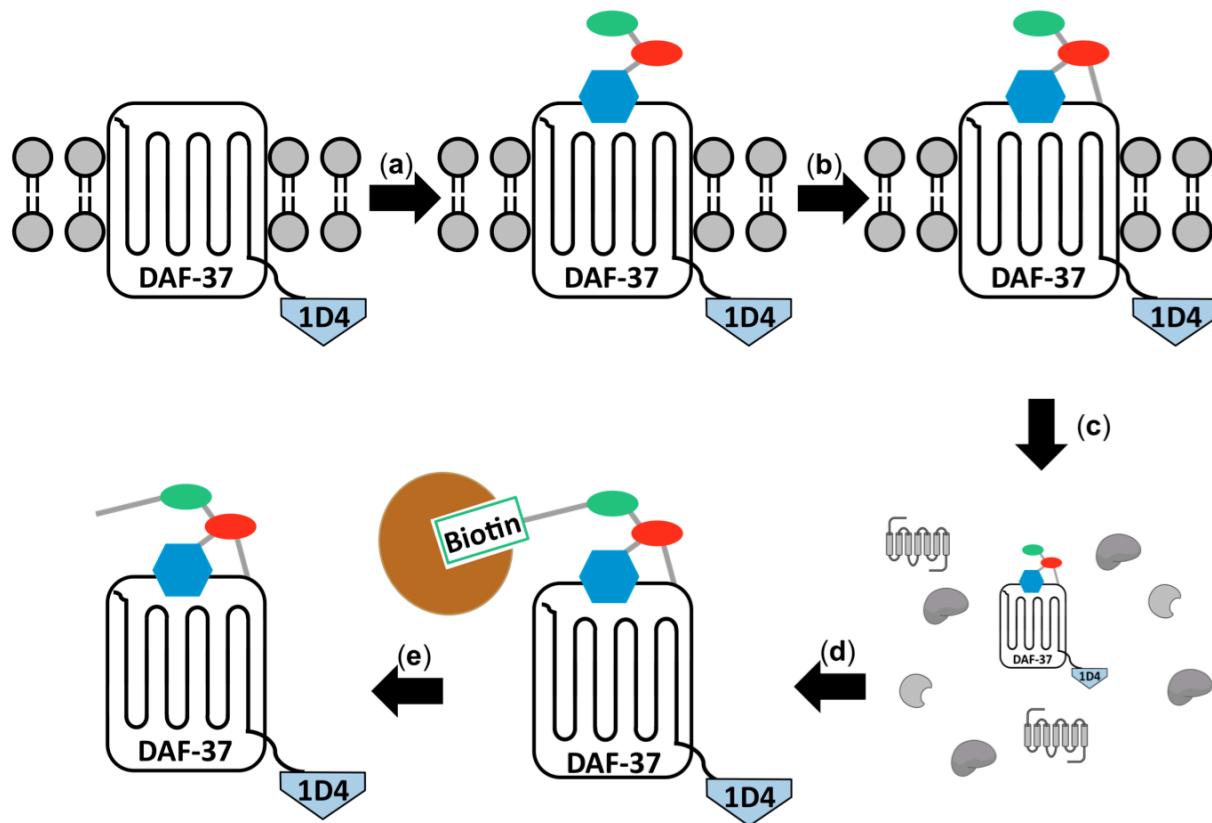


**Figure P.5: Alphascreen experimental set up and results for ascr#2 probe.** Schematic representing alphascreen of biotinylated ascr#2 probe covalently bound to DAF-37. Streptavidin-coated donor beads co-localize with 1D4 antibody-coated acceptor beads as a result of the ascr#2 probe being linked to DAF-37.

In Chapter 2 of this dissertation, a novel method for the deconvolution and identification of ligand-receptor binding sites using cleavable stable isotope click-chemistry tags (CSICT) is discussed. This method's use of a cleavable, clickable biotin purification tag eliminates the need for elevated temperatures to release the probe-receptor complex from the streptavidin bead. Instead, the elution is achieved through a mild chemical cleavage of diazobenzene group that is incorporated into the structure of the biotin

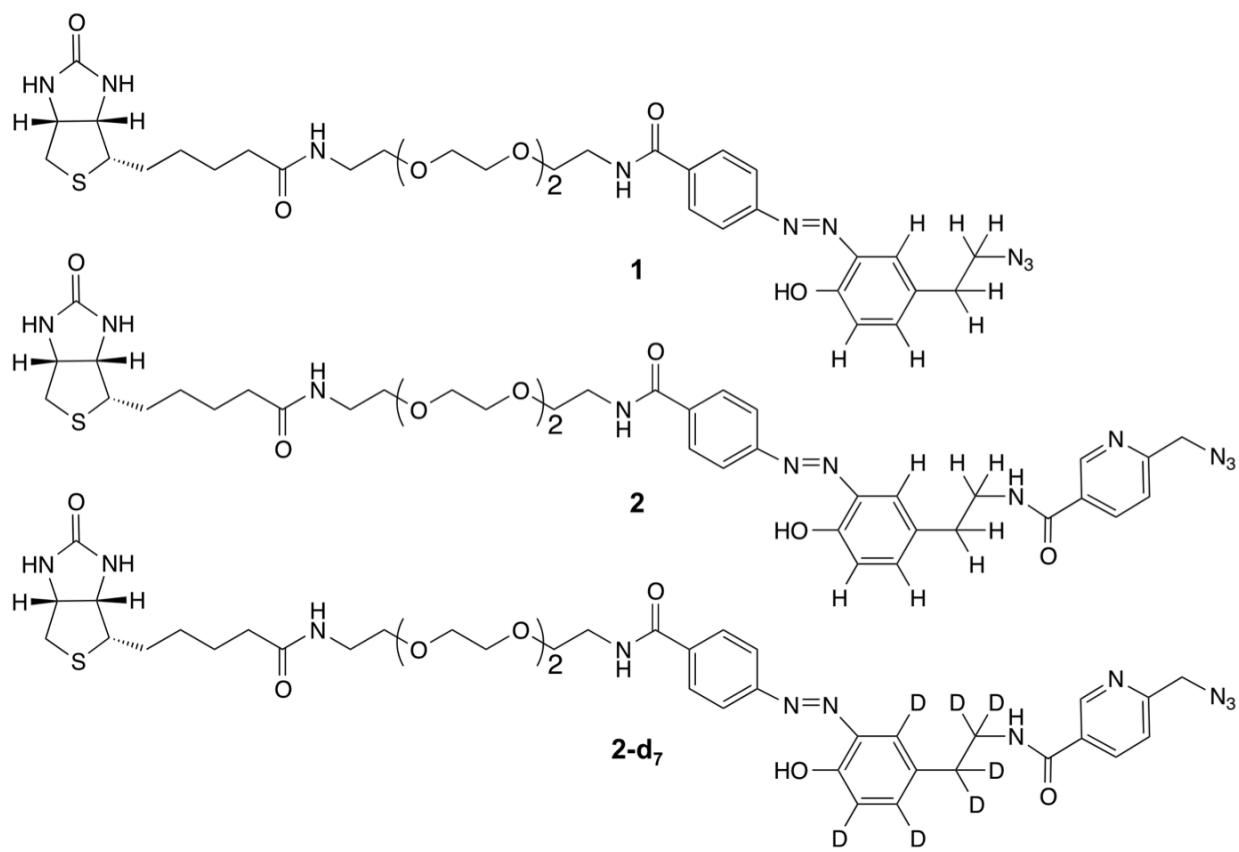
linker. Although a similar diazobenzene cleavable linker strategy has been previously described<sup>41,53</sup>, this is the first time that it has been coupled do a top-down photoaffinity labeling technique.

To demonstrate the cleavable linker's compatibility with the isolation of lipophilic membrane proteins, a first generation diazobenzene cleavable linker<sup>41</sup> (**1**, **Figure P.7**) was tested to see if it be utilized to release the of asc#2-modified DAF-37 GPCR. Following the methodology outlined in **Figure P.4**, an AfBPP approach was used and it was demonstrated that the probe-labeled GPCR could be efficiently eluted from streptavidin beads using this chemically cleavable linker. Having demonstrated the applicability of diazobenzene-based cleavable linkers to the isolation of membrane receptors, a second generation of the cleavable linkers were synthesized that incorporate a differential stable isotope labeling strategy.



**Figure P.6: Application of cleavable biotin linker to ascr#2 probe and DAF-37.** (a) Delivery and binding of ascr#2 probe to DAF-37::1D4 being expressed in HEK293T cells. (b) Generation of covalent probe-receptor complex with *in vivo* photocrosslinking of ascr#2 probe to DAF-37::1D4. (c) Solubilization of cells to create a whole cell protein lysate. (d) Biotinylation of probe-receptor complex via click chemistry, precipitation and enrichment of complex on to streptavidin beads. (e) Dithionite elution of probe-receptor complex from the streptavidin beads in preparation for western blot and silver stain analysis.

As outlined in **Figure P.7**, a heavy (**2**) and a light (**2-d<sub>7</sub>**) stable-isotope cleavable biotin azide were synthesized. The heavy azide was synthesized with seven deuteriums in place of hydrogens in the light azide, providing a robust chemical signature for differential characterization via mass spectrometry.



**Figure P.7. Showing the final structure of the first (1) and second-generation cleavable light (2) and heavy (2-d<sub>7</sub>) biotin azide linkers.**

Used in tandem, the heavy/light azides provide a means for the receptor-binding site identification of any clickable AfBPP or ABPP probe. By differentially labeling the crosslinked probe-receptor complex with heavy and light azides, a seven-amu difference is achieved on the probe-receptor complex. Once these heavy and light complexes are enriched, chemically eluted from the streptavidin beads and trypsinized, a differential heavy and light probe-peptide binding site fragment will be generated. When these heavy and light peptide fragments are subjected to LC-MS analysis they will elute from the LC at the same time but have a distinct and easily observable mass shift that allows the identification of peptide fragments around the probe-binding site<sup>54</sup>.

Without this distinctive chemical signature, detection and subsequent characterization of probe-peptide fragments is a significant challenge (frequently impossible) due to the number of receptor specific trypsin fragments, as well the presence of background peptides that occur from the trypsinization of contaminant proteins dragged along in the purification process. Literature examples for the recognition of probe-peptide fragments have used radioactive labeled ligands<sup>55-56</sup>, characteristic elemental isotopic abundancies<sup>41,57</sup>, photocrosslinking ligands coupled to stable isotope labeling in cell culture (SILAC)<sup>58</sup> and novel robust mass spectrometry methods<sup>59</sup>. However, all of these methods suffer from a limited scope of application due their own shortcomings that are outlined in detail in Chapter 2.

Lipophilic membrane proteins such as GPCRs serve as signaling conduits between the cell and its exterior environment and frequently act as the point-of-perception between a cell and endogenous, environmentally occurring or drug-like chemical entities. Although there have been significant advancements in top-down receptor identification strategies such as ABPP and AfBPP, multi-pass membrane receptors have previously not been accessible using these top-down techniques, due to the intrinsic biochemical nature of these proteins. To address this need, the work described in this dissertation describes two independent methods for the analysis of ligand-receptor interactions that are applicable to lipophilic membrane proteins. These studies provide a translatable methodology that has the ability to screen receptor candidates and also provide a technique to identify the binding sites of bioactive small molecule ligands, allowing the detailed molecular understanding of the initial binding event.

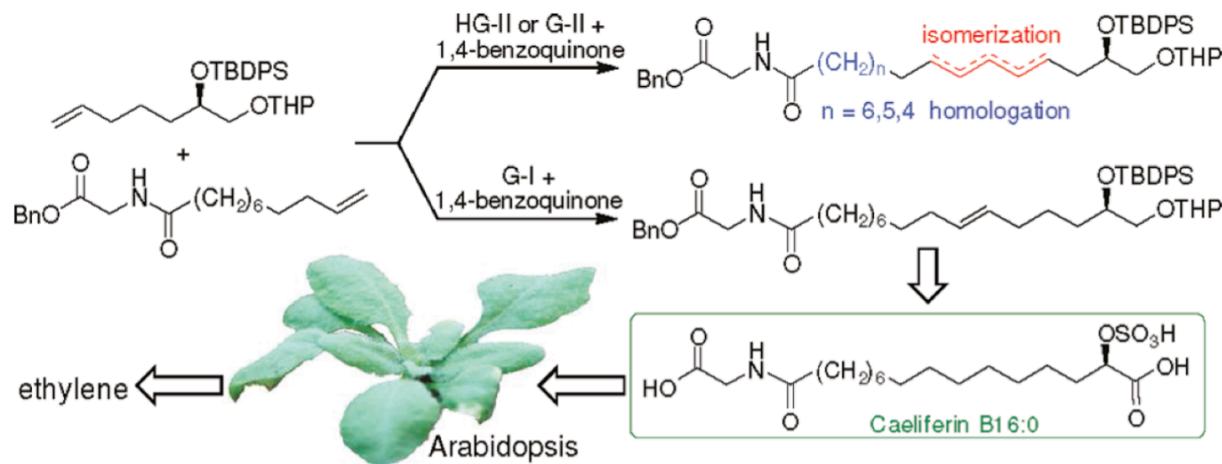
The research outlined in Chapter 1 contributed towards the publication<sup>15</sup>, “Interaction of structure-specific and promiscuous G-protein–coupled receptors mediates small-molecule signaling in *Caenorhabditis elegans*”. Authors’ contributions: Donha Park performed all of the *C.elegans* based research, generation of mutant

strains, phenotypic assays, worm microscopy, experimental design, data analysis and writing; Inish O'Doherty synthesis of ascr#2 photoaffinity probe, application in of probe in alphasure assay, experimental design, data analysis and writing; Rishi K Somvanshi receptor dimerization assay and data analysis; Axel Bethke experimental design and data analysis; Ujendra Kumar experimental design, data analysis and writing; Frank C. Schroeder experimental design, data analysis and writing; Donald Riddle experimental design, data analysis and writing.

The research outlined in Chapter 2 will contribute to a future publication, "Advancements towards the differential labeling of small-molecule binding sites using photoaffinity probes coupled to cleavable stable isotope click-chemistry tags". Authors' contributions: Inish O'Doherty, Patrick Boyle, Sarah Refi Hind, Frank C. Schroeder and Georgry Martin designed research and carried out data analysis. Inish O'Doherty synthesis, analytical characterization and *in vitro* testing of cleavable stable isotope click-chemistry tags and PAMP-probes, DAF-37 binding assay, experimental design, data analysis and writing; Daniel Gentile synthesis of cleavable stable isotope click-chemistry tags; Joshua A. Bacchie synthesis of PAMP-probes; Patrick Boyle PAMP-probe testing in *Styrax peruvianum*, experimental design, data analysis and writing; Sarah Refi Hind PAMP-probe testing in *Styrax peruvianum*, experimental design; Gregory B. Martin experimental design, data analysis and writing; Frank C. Schroeder experimental design, data analysis and writing.

**The Total Synthesis of the Caeliferins<sup>60</sup>.** A cross metathesis- (CM-) <sup>61-64</sup> based synthesis of the caeliferins<sup>65,66</sup>, a family of sulfoxy fatty acids that elicit plant immune responses is reported. Unexpectedly, detailed NMR-spectroscopic<sup>50,67</sup> and mass spectrometric analyses of CM reaction mixtures revealed extensive isomerization and homologation of starting materials and products<sup>68-71</sup>. In Chapter 3 it is shown that the degree of isomerization and homologation in CM strongly correlates with substrate

chain length and lipophilicity. Side-product suppression requires appropriate catalyst selection and use of 1,4-benzoquinone as hydride scavenger<sup>60</sup>.



**Figure P.8: Schematic outlining the metathesis based convergent synthesis of Caeliferin B16:0 and it's biological activity.**

The research outlined in Chapter 3 contributed towards the publication<sup>60</sup>, “Synthesis of Caeliferins, Elicitors of Plant Immune Responses: Accessing Lipophilic Natural Products via Cross Metathesis”. Authors’ contributions: Inish O’Doherty synthesis, experimental design, data analysis and writing; Joshua J. Yim synthesis, experimental design, data analysis and writing; Eric A. Schmelz *Arabidopsis thaliana* assay and data analysis; Frank C. Schroeder experimental design, data analysis and writing.

## REFERENCES

- (1) Terstappen, G. C.; Schlüpen, C.; Raggiaschi, R.; Gaviraghi, G. *Nat Rev Drug Discov* **2007**, *6*, 891–903.
- (2) Swinney, D. C.; Anthony, J. *Nat Rev Drug Discov* **2011**, *10*, 507–519.
- (3) Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. *Angew. Chem. Int. Ed.* **2013**, *52*, 2744–2792.
- (4) Bantscheff, M.; Drewes, G. *Bioorg. Med. Chem.* **2012**, *20*, 1973–1978.
- (5) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. *Nat Chem Biol* **2013**, *9*, 232–240.
- (6) Spring, D. R. *Chem Soc Rev* **2005**, *34*, 472–482.
- (7) Mitchison, T. J. *Chemistry & Biology* **1994**.
- (8) Stockwell, B. R. *Nat. Rev. Genet.* **2000**.
- (9) Pickart, M. A.; Klee, E. W.; Nielsen, A. L.; Sivasubbu, S.; Mendenhall, E. M.; Bill, B. R.; Chen, E.; Eckfeldt, C. E.; Knowlton, M.; Robu, M. E.; Larson, J. D.; Deng, Y.; Schimmenti, L. A.; Ellis, L. B. M.; Verfaillie, C. M.; Hammerschmidt, M.; Farber, S. A.; Ekker, S. C. *PLoS ONE* **2006**, *1*, e104.
- (10) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998**, *391*, 806–811.
- (11) Boutros, M.; Ahringer, J. *Nat. Rev. Genet.* **2008**, *9*, 554–566.
- (12) Clamp, M.; Fry, B.; Kamal, M.; Xie, X.; Cuff, J.; Lin, M. F.; Kellis, M.; Lindblad-Toh, K.; Lander, E. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19428–19433.
- (13) Hopkins, A. L. *Nat Chem Biol* **2008**.
- (14) Lapinsky, D. J. *Bioorg. Med. Chem.* **2012**, *20*, 6237–6247.
- (15) Park, D.; O'Doherty, I.; Somvanshi, R. K.; Bethke, A.; Schroeder, F. C.; Kumar, U.; Riddle, D. L. *pnas.org*.
- (16) Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686–4687.
- (17) Geurink, P. P.; Prely, L. M.; Marel, G. A.; Bischoff, R.; Overkleft, H. S. In *link.springer.com*; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; Vol. 324, pp. 85–113.
- (18) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.
- (19) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192–3193.
- (20) Chenoweth, K.; Chenoweth, D.; Goddard, W. A., III *Org. Biomol. Chem.* **2009**, *7*, 5255–5258.
- (21) Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 3688–3690.
- (22) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
- (23) Wirth, T.; Schmuck, K.; Tietze, L. F.; Sieber, S. A. *Angew. Chem. Int. Ed.* **2012**, n/a–n/a.
- (24) Eirich, J.; Orth, R.; Sieber, S. A. *J. Am. Chem. Soc.* **2011**, *133*, 12144–12153.
- (25) Lenz, T.; Fischer, J. J.; Dreger, M. *J Proteomics* **2011**, *75*, 100–115.
- (26) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. *Nat Chem Biol* **2009**, *5*, 341–343.
- (27) Brunner, J.; Senn, H.; Richards, F. M. *J. Biol. Chem.* **1980**, *255*, 3313–3318.
- (28) Das, J. *Chem. Rev.* **2011**, *111*, 4405–4417.
- (29) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.;

- Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.
- (30) Uttamapinant, C.; Tangpeerachaikul, A. *Angewandte Chemie*.
- (31) Hein, J. E.; Fokin, V. V. *Chem Soc Rev* **2010**, *39*, 1302–1315.
- (32) Nodwell, M. B.; Sieber, S. A. In *link.springer.com*; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; Vol. 324, pp. 1–41.
- (33) Li, N.; Overkleft, H. S.; Florea, B. I. *Current Opinion in Chemical Biology* **2012**, *16*, 227–233.
- (34) Smith, D. P.; Anderson, J.; Plante, J.; Ashcroft, A. E.; Radford, S. E.; Wilson, A. J.; Parker, M. J. *Chem. Commun.* **2008**, 5728–5730.
- (35) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596–2599.
- (36) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proceedings of the National Academy of Sciences* **2007**, *104*, 16721–16722.
- (37) de Almeida, G.; Sletten, E. M.; Nakamura, H.; Palaniappan, K. K.; Bertozzi, C. R. *Angew. Chem. Int. Ed. Engl.* **2012**.
- (38) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. *J. Am. Chem. Soc.* **2009**, *131*, 4967–4975.
- (39) Sohn, C. H.; Lee, J. E.; Sweredoski, M. J.; Graham, R. L. J.; Smith, G. T.; Hess, S.; Czerwieniec, G.; Loo, J. A.; Deshaies, R. J.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2012**, *134*, 2672–2680.
- (40) Lallana, E.; Riguera, R.; Fernandez-Megia, E. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 8794–8804.
- (41) Yang, Y. Y.; Grammel, M.; Raghavan, A. S.; Charron, G.; Hang, H. C. *Chemistry & Biology* **2010**, *17*, 1212–1222.
- (42) Gomes, I.; Gupta, A.; Devi, L. A. *G-Protein-Coupled Heteromers: Regulation in Disease*; 1st ed. Elsevier Inc., 2013; Vol. 521, pp. 219–238.
- (43) Ren, H.; Yu, D.; Ge, B.; Cook, B.; Xu, Z.; Zhang, S. *PLoS ONE* **2009**, *4*, e4509.
- (44) Corin, K.; Baaske, P.; Ravel, D. B.; Song, J.; Brown, E.; Wang, X.; Geissler, S.; Wienken, C. J.; Jerabek-Willemsen, M.; Duhr, S.; Braun, D.; Zhang, S. *PLoS ONE* **2011**, *6*, e23036.
- (45) Ullman, E. F.; Kirakossian, H.; Singh, S.; Wu, Z. P.; Irvin, B. R.; Pease, J. S.; Switchenko, A. C.; Irvine, J. D.; Dafforn, A.; Skold, C. N. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5426–5430.
- (46) Butcher, R. A.; Fujita, M.; Schroeder, F. C.; Clardy, J. *Nat Chem Biol* **2007**, *3*, 420–422.
- (47) Ludewig, A. H.; Izrayelit, Y.; Park, D.; Malik, R. U.; Zimmermann, A.; Mahanti, P.; Fox, B. W.; Bethke, A.; Doering, F.; Riddle, D. L.; Schroeder, F. C. *Proceedings of the National Academy of Sciences* **2013**, *110*, 5522–5527.
- (48) Jeong, P.-Y.; Jung, M.; Yim, Y.-H.; Kim, H.; Park, M.; Hong, E.; Lee, W.; Kim, Y. H.; Kim, K.; Paik, Y.-K. *Nature* **2005**, *433*, 541–545.
- (49) Butcher, R. A.; Ragains, J. R.; Kim, E.; Clardy, J. *PNAS* **2008**, *105*, 14288–14292.
- (50) Pungaliya, C.; Srinivasan, J.; Fox, B. W.; Malik, R. U.; Ludewig, A. H.; Sternberg, P. W.; Schroeder, F. C. *PNAS* **2009**, *106*, 7708–7713.
- (51) Srinivasan, J.; Kaplan, F.; Ajredini, R.; Zachariah, C.; Alborn, H. T.; Teal, P. E. A.; Malik, R. U.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *Nature* **2008**, *454*, 1115–

- (52) Srinivasan, J.; Reuss, von, S. H.; Bose, N.; Zaslaver, A.; Mahanti, P.; Ho, M. C.; O'Doherty, O. G.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *PLoS Biol* **2012**, *10*, e1001237.
- (53) Yang, Y.-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641.
- (54) Ross, P. L. *Molecular & Cellular Proteomics* **2004**, *3*, 1154–1169.
- (55) Couvinau, A.; Tan, Y.-V.; Ceraudo, E.; Laburthe, M. In *Methods in Enzymology*; Methods in Enzymology; Elsevier, 2013; Vol. 520, pp. 219–237.
- (56) Grunbeck, A.; Huber, T.; Sakmar, T. P. In *Methods in Enzymology*; Methods in Enzymology; Elsevier, 2013; Vol. 520, pp. 307–322.
- (57) Palaniappan, K. K.; Pitcher, A. A.; Smart, B. P.; Spicciarich, D. R.; Iavarone, A. T.; Bertozzi, C. R. *ACS Chem. Biol.* **2011**, *6*, 829–836.
- (58) Hulce, J. J.; Cognetta, A. B.; Niphakis, M. J.; Tully, S. E.; Cravatt, B. F. *Nature Methods* **2013**, *10*, 259–264.
- (59) Frei, A. P.; Jeon, O.-Y.; Kilcher, S.; Moest, H.; Henning, L. M.; Jost, C.; Plückthun, A.; Mercer, J.; Aebersold, R.; Carreira, E. M.; Wollscheid, B. *Nature Reviews Immunology* **2012**, *30*, 997–1001.
- (60) O'Doherty, I.; Yim, J. J.; Schmelz, E. A.; Schroeder, F. C. *Org. Lett.* **2011**, *13*, 5900–5903.
- (61) Grubbs, R. H.; Burk, P. L.; Carr, D. D. *J. Am. Chem. Soc.* **1975**, *97*, 3265–3267.
- (62) Connon, S. J.; Blechert, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 1900–1923.
- (63) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 11360–11370.
- (64) Hoveyda, A. H.; Zhugralin, A. R. *Nature* **2007**, *450*, 243–251.
- (65) Alborn, H. T.; Hansen, T. V.; Jones, T. H.; Bennett, D. C.; Tumlinson, J. H.; Schmelz, E. A.; Teal, P. E. A. **2007**, *104*, 12976.
- (66) Schmelz, E. A.; Engelberth, J.; Alborn, H. T.; Tumlinson, J. H.; Teal, P. E. A. *Proceedings of the National Academy of Sciences* **2009**, *106*, 653–657.
- (67) Taggi, A. E.; Meinwald, J.; Schroeder, F. C. *J. Am. Chem. Soc.* **2004**, *126*, 10364–10369.
- (68) Schmidt, B. *Eur. J. Org. Chem.* **2004**, *2004*, 1865–1880.
- (69) Courchay, F. C.; Sworen, J. C.; Ghiviriga, I.; Abboud, K. A.; Wagener, K. B. *Organometallics* **2006**, *25*, 6074–6086.
- (70) Fokou, P. A.; Meier, M. A. R. *J. Am. Chem. Soc.* **2009**, *131*, 1664–1665.
- (71) Mutlu, H.; de Espinosa, L. M.; Meier, M. A. R. *Chem Soc Rev* **2011**, *40*, 1404–1445.

## CHAPTER 1

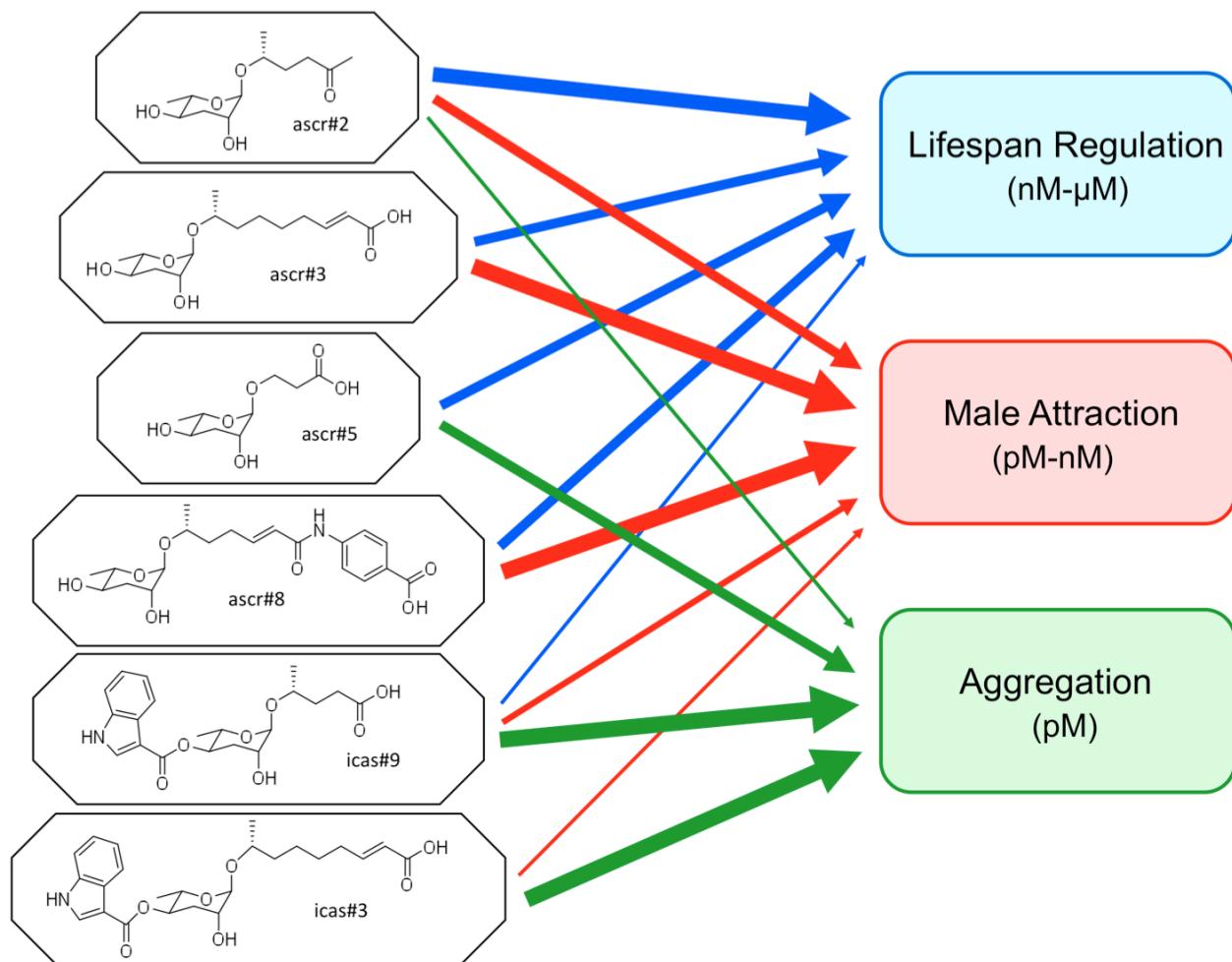
# INTERACTION OF STRUCTURE-SPECIFIC AND PROMISCUOUS G-PROTEIN-COUPLED RECEPTORS MEDIATES SMALL MOLECULE SIGNALLING IN *C. ELEGANS*

**Introduction:** Several different aspects of the life history of the model organism *Caenorhabditis elegans* are under the control of a chemically diverse family of small-molecule signals, the ascarosides. Ascarosides function as regulators of developmental timing<sup>1-4</sup>, mate attraction<sup>4,5</sup>, aggregation behavior<sup>6,7</sup>, and olfactory learning<sup>8</sup>. Chemically, the ascarosides form a modular library of signaling molecules based on the dideoxysugar ascarylose, which is linked to fatty acid-like side chains of varying lengths and decorated further with additional building blocks derived from amino acid metabolism and other pathways (**Figure 1.1**)<sup>7</sup>.

Ascarosides were first identified as the constituents of the dauer pheromone, a signal controlling entry into and exit from the dauer diapause, an alternate larval stage that is non-feeding, long-lived, and highly stress resistant. *C. elegans* larvae interrupt normal development and enter the dauer diapause when sensing unfavorable conditions, such as high population density, limited food, high temperature or microbial pathogenesis<sup>9-11</sup>. Formation of dauer larvae requires exposure of the developing L1 larvae to dauer pheromone<sup>12</sup>, the most important components of which are the ascarosides ascr#2, ascr#3, ascr#5, and ascr#8<sup>2-4</sup>.

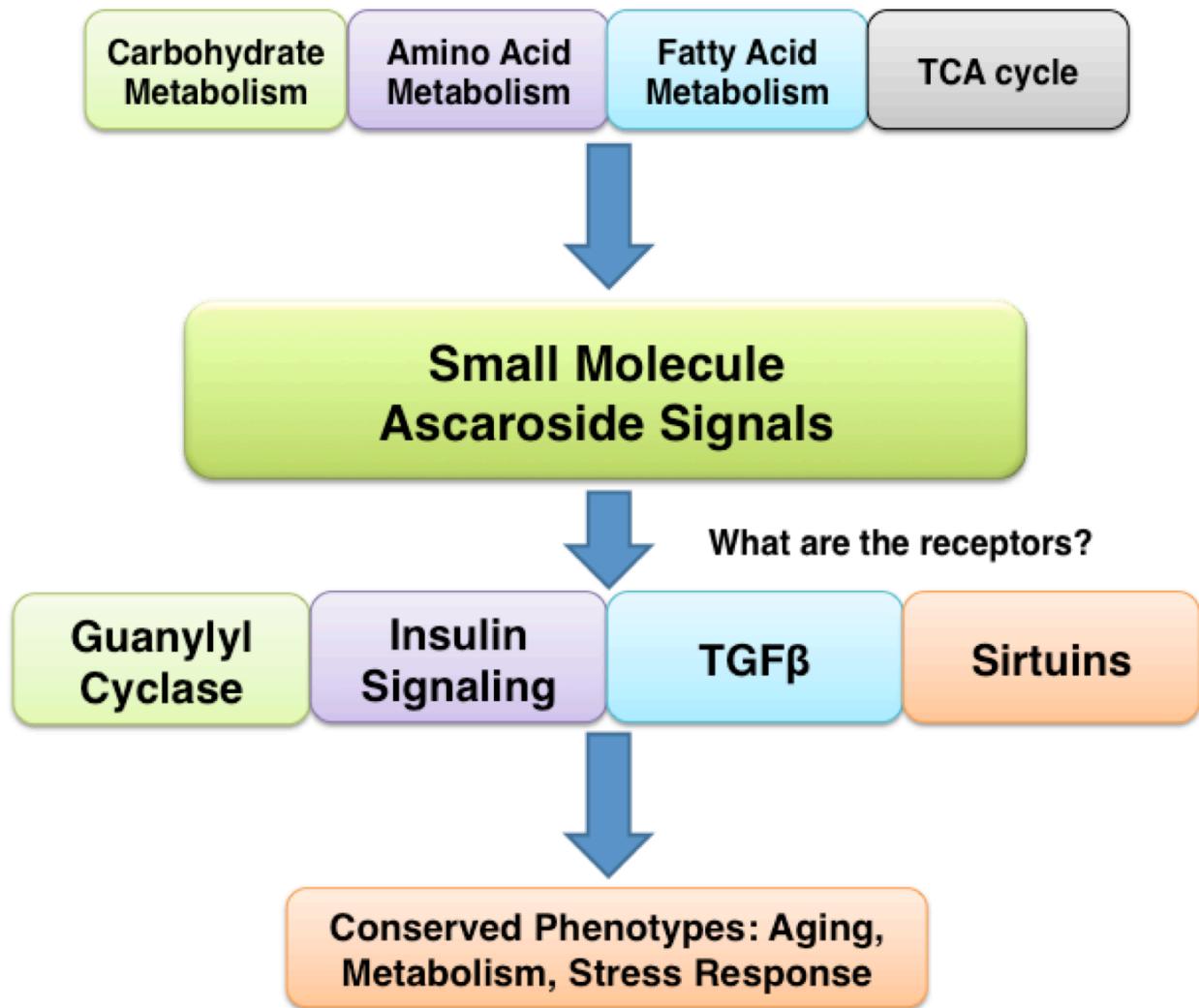
Ascaroside signaling is complex. More than 100 different ascaroside structures have been identified in *C. elegans*, and there is evidence that even small changes in ascaroside structure are associated with altered activity profiles<sup>6,7</sup>. Most ascaroside-mediated phenotypes involve synergistic action of two or more compounds which must be present in specific proportions and concentrations<sup>2-5</sup>. Furthermore, individual

ascarosides may have more than one function; for example, ascr#3 is involved in dauer formation, male attraction, and hermaphrodite repulsion<sup>2,5,6,13</sup>.



**Figure 1.1: Simplified scheme illustrating synergy and concentration dependence of ascaroside signaling.** For example, ascr#2, #3 and #5 additively promote longevity at nM- $\mu$ M concentrations, whereas ascr#2, #3 and #8 signal synergistically for male attraction pM-nM. Furthermore, some the ascarosides can signal independently without the assistance of other ascarosides, demonstrated by icas#3 acting as a potent hermaphrodite aggregant at pM concentrations

Therefore, ascaroside signaling in *C. elegans* constitutes a unique model for the study of small molecule perception. However, the mechanisms that underlie the diverse, partially overlapping and synergistic activities of the ascarosides are largely unknown. Molecular genetic analysis of dauer-constitutive (Daf-c) and dauer-defective (Daf-d) mutants revealed that dauer pheromone perception is upstream of several conserved pathways, including the *daf-11*/guanylyl cyclase, insulin/IGF, and TGF- $\beta$  pathways<sup>14-18</sup>. These studies suggested that the dauer pheromone signal is detected by G protein-coupled receptors (GPCRs) coupled to the *daf-11*/guanylyl cyclase pathway (**Figure 1.2**)<sup>14</sup>. Furthermore, previous work suggested that the dauer pheromone signal is detected by G protein-coupled receptors (GPCRs) coupled to the *daf-11*/guanylyl cyclase pathway (**Figure 1.2**)<sup>14</sup>. Recently, the GPCR-encoding *srbc-64* and *srbc-66* genes were shown to participate in ascaroside-mediated dauer induction; however, dauer-inducing activity of ascarosides is only partially abolished in *srbc-64;srbc-66* double mutants<sup>19</sup>. Subsequently, two functionally redundant GPCR-encoding genes, *srg-36* and *srg-37*, were shown to be required specifically for dauer induction by ascr#5<sup>20</sup>. Here we describe the synthesis and biologically activity of two ascaroside photoaffinity probes targeting receptors of ascr#8 and ascr#2. We show that *daf-37* and *daf-38* encode a heterodimeric pair of ascaroside-sensing GPCRs that mediate both



**Figure 1.2: Schematic outline of ascaroside signaling pathways.** The ascarosides are a modular library of small molecules integrating building blocks from several primary metabolic pathways in *C. elegans*. Perception of these compounds triggers signaling via highly conserved pathways, including *daf-11*/guanylyl cyclase, insulin/IGF, TGF- $\beta$  signaling and sirtuin-dependent pathways. The mechanisms of signal perception and activation of downstream pathways are largely unknown.

dauer formation and behavioral phenotypes. Whereas DAF-37 is required specifically for perception of the potent dauer inducer ascr#2, DAF-38 appears to play a cooperative role in the perception of ascr#2 as well as other ascarosides. Using our photoaffinity labeled ascr#2 probe, we demonstrate direct binding of ascr#2 to DAF-37. Via cell-specific expression we show that one small molecule (ascr#2) can elicit different phenotypes via binding to the same GPCR (DAF-37) in two different neurons.

The data in this Chapter contributed towards the publication, “*Interaction of structure-specific and promiscuous G-protein–coupled receptors mediates small-molecule signaling in Caenorhabditis elegans*”. Author contributions to publication: Inish O’Doherty, alphascreen design and data collection, synthesis of chemical probes, interpretation of data and writing; Donha Park, initial identification of daf-37, *C.elegans* experiments, interpretation of data and writing; Rishi K. Somvanshi, photobleaching experiment for receptor dimerization, interpretation of data and writing data; Axel Bethke, alphascreen design and data collection; Frank C. Schroeder, experimental design, interpretation of data and writing; Ujendra Kumard, experimental design, interpretation of data and writing; Donald L. Riddle, experimental design, interpretation of data and writing.

## **Results**

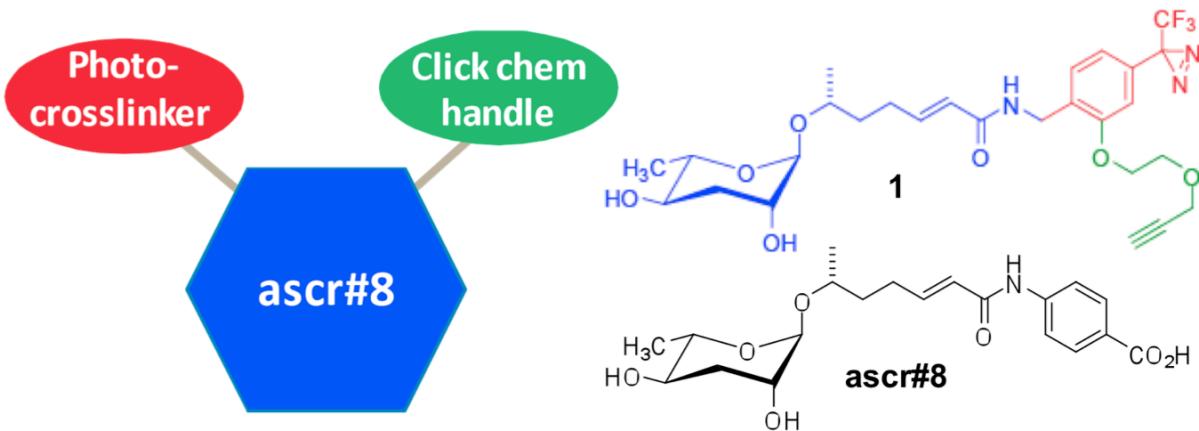
**Ascaroside Photoaffinity Probe Design:** Syntheses of a great variety of photocrosslinkable derivatives of small molecules have been reported, however, in many cases the ability to bind to their respective receptor was lost<sup>21</sup>. Thus, it was vital to couple chemical and biological intuition to make informed choices in the design of a functional probe. The most critical aspect of ascaroside probe design was possible interference with ligand binding. Analysis of ascaroside activity showed that small changes in structure can have profound effects on the signaling capabilities of these molecules. Specifically, the subtle but distinct changes in the 4’ hydroxy and the acyl

chain ultimately impart different phenotypic outcomes<sup>22</sup>. This was carefully considered during probe design. Otherwise, it is entirely likely that a probe designed for investigating the receptor binding profile for one ascaroside would mimic that of another. The probe design thus aims to: 1) retain signaling capabilities; 2) incorporate a photocrosslinking moiety; 3) include a chemical tag for purification purposes.

We evaluated the structures of known ascarosides (in 2009 only about one dozen different ascarosides were known, compared to over 200 today) for suitable probe candidates and we were immediately drawn to ascr#8, due to specific aspects of its chemical structure and its potent and experimentally robust signaling profile in phenotypic assays. The bacterially derived *para*-amino benzoic acid (PABA) moiety located at the terminal of ascr#8's alky side chain provided us with the unique ability to incorporate a photocrosslinking and click chemistry handle directly into the structure of the ascaroside without the need to greatly perturb the basic skeleton of the molecule. Furthermore, ascr#8 had been shown to elicit strong phenotypic responses in fast and reliable male mating attraction<sup>23</sup> and thermotolerance assays<sup>24</sup> that allowed for rapid testing of functionalized derivatives. These unique features made ascr#8 an excellent candidate for our first venture into the design of ascaroside-based photoaffinity probes.

Several different photocrosslinking species were considered for design of the ascr#8 probe: benzophenone, azides, and trifluorodiazirines (Tfd). Given the relatively small size of ascr#8, incorporation of a bulky benzophenone moiety was not pursued. Both aryl-attached azides and Tfds have been used extensively in the literature, and their attachment in place of the carboxy group of ascr#8 would yield a probe that has a similar overall size and shape as ascr#8, although hydrogen bonding capabilities of the acid group in ascr#8 would be abolished in azides and Tfds derivatives. Given ample literature precedents and given the fact that Tfd requires lower energy UV activation compared to the aryl-azides, which reduce the risk of photoactivating endogenous proteins<sup>25,26</sup>, we ultimately selected Tfd for our ascr#8 design<sup>21,27-29</sup>.

The additional installation of an alkyne would allow the utilization of established “click” chemistry techniques for protein purification<sup>30-32</sup>. The complete design of our trifunctional ascr#8 photoaffinity probe as shown in **Figure 1.3** thus closely mimics the structure of ascr#8, although we decided to insert an additional CH<sub>2</sub> group between the aromatic ring and the 7-carbon side chain, because of concern that N-substitution of the aromatic ring may interfere with photochemical efficacy of the probe.



**Figure 1.3: Structure of ascr#8 and ascr#8 probe (1).**

For synthesis of the crosslinking moiety, we started with *m*-bromoanisole, adapting a previous synthesis by Mayer et al.<sup>27</sup> Key steps include introduction of the diazirine moiety via a corresponding hydroxyimine tosylate and selective carbonylation of the aromatic ring in *para*-position to the introduced Tfd moiety (**Figure 1.4**). The ascaroside portion of the ascr#8 probe, equivalent to the 7-carbon side-chain ascaroside ascr#7, was prepared as described previously,<sup>23</sup> and subsequently coupled to the aromatic crosslinking moiety (**Figure 1.5**). With the final ascr#8 probe in hand, it was vital to test whether signaling activity had been retained. This was accomplished using both the male mating attraction<sup>33</sup> and aging<sup>24</sup> assays. Both assays (**Figure 1.5**) showed substantial activity of the ascr#8 probe. In the male mating attraction assay, the mean time spent in the scoring region was ~125 s for the ascr#8 probe, compared 180-200s observed for ascr#8. In the aging assay, the probe offered a mean lifespan increase of 13%, compared to an average 17% increase attributed to ascr#8. These results

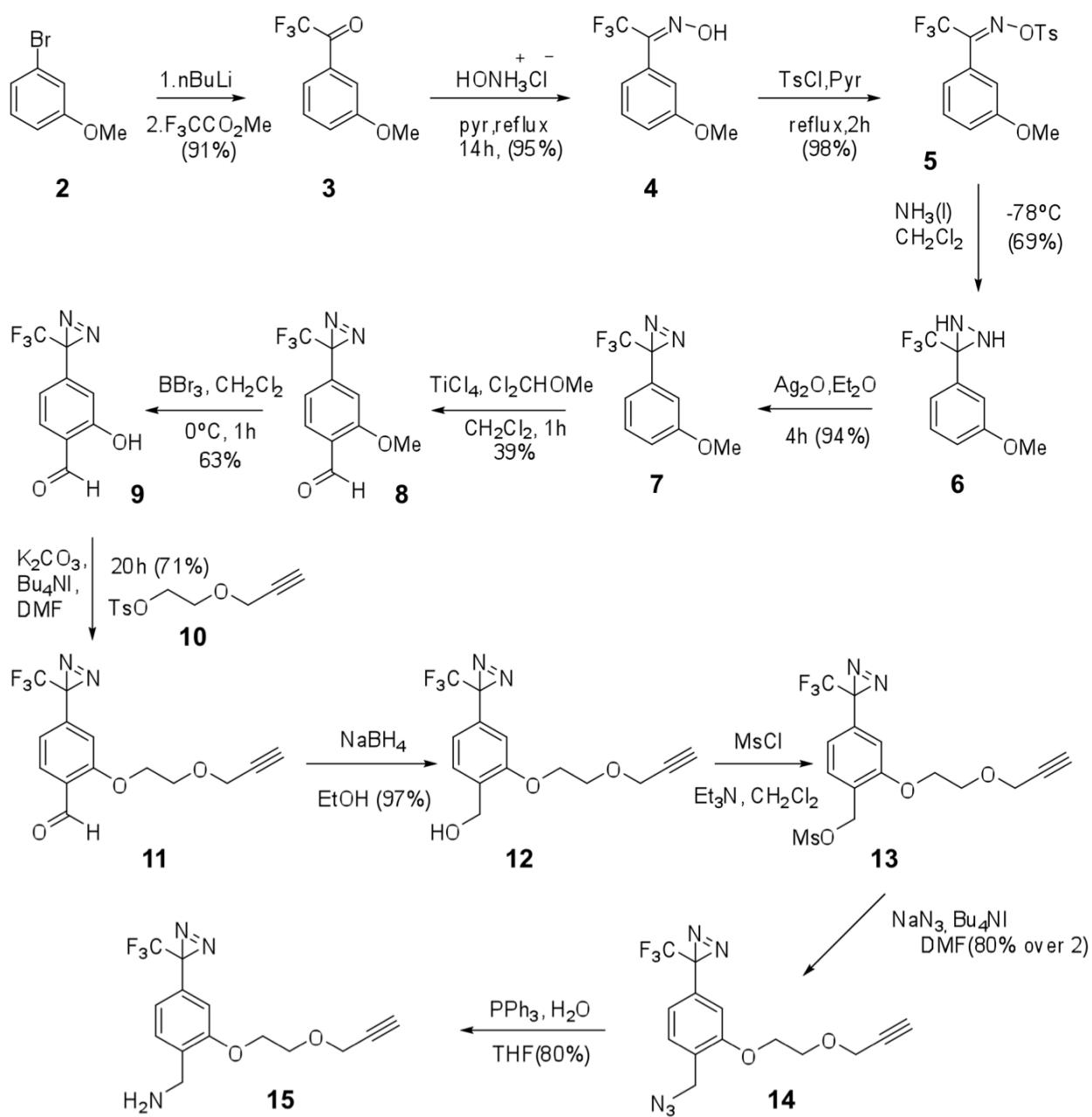
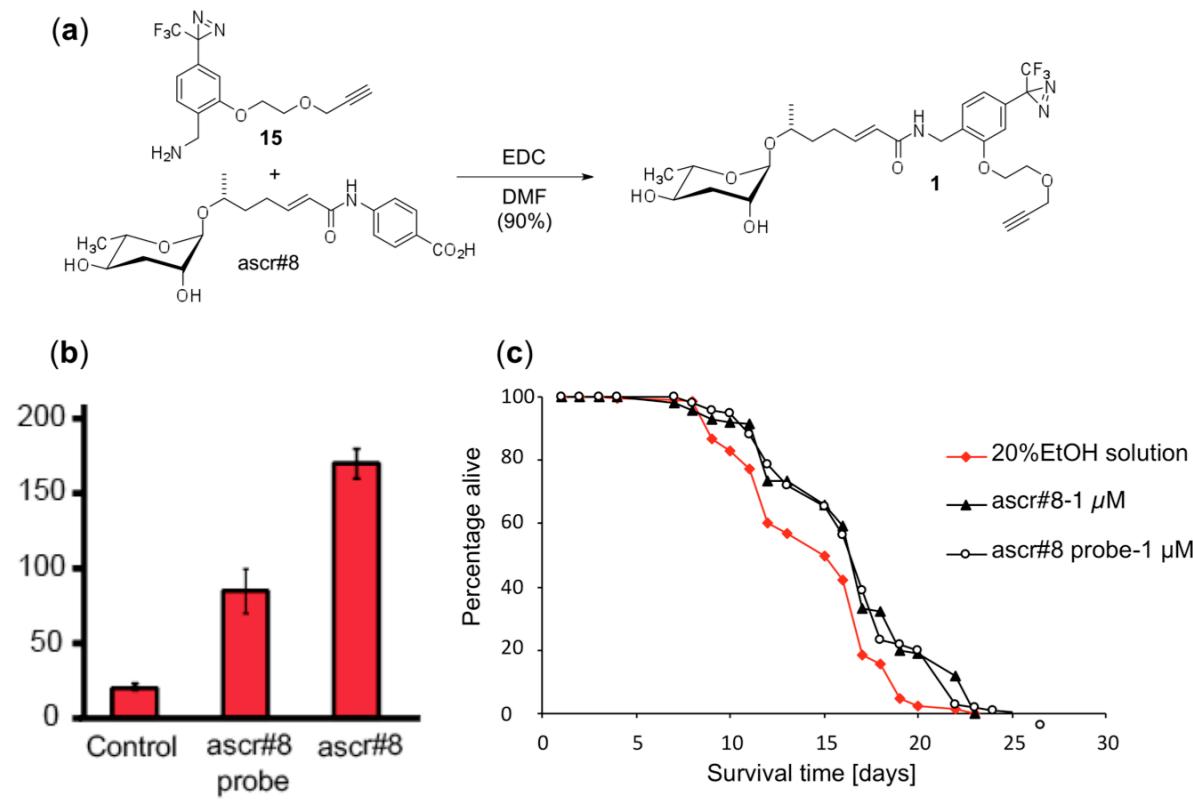


Figure 1.4: Synthetic route employed for aromatic portion of the probe (15).

demonstrate that the ascr#8 probe is still capable of eliciting ascr#8-specific phenotypic responses at relevant biological concentrations, suggesting that the probe still binds to the relevant ascr#8 receptor(s).



**Figure 1.5: Ascr#8 probe (1) and its biological activity.** (a) Final coupling of ascr#7 to benzyl amine (15), forming active probe. Ascr#8 probe activity in (b) male attraction, (c) aging assay

**Ascr#2 Photoaffinity Probe Design.** Based on the success of designing a biologically active the ascr#8 probe, we decided to test if we could devise probes for other, structurally even simpler ascarosides. We focused on ascr#2, because of its strong activity in the dauer formation assay and because it is the only ascaroside that includes an alpha-methyl ketone at the terminus of its alkyl side chain. This ketone functionality plays an important role for ascr#2's biological activity. For example, if this

carbonyl is reduced to the corresponding alcohols, ascr#6.1/6.2, or oxidized to the carboxylic acid, ascr#9, the dauer-inducing activity of ascr#2 is lost almost completely, demonstrating the chemical significance of the methyl ketone for signaling. We postulated that if the biological activity of ascr#2 is directly tied to the retention of the methyl ketone, then perhaps derivatization (for incorporation of photoaffinity and click chemistry handle) at the distal 4' hydroxy would be tolerated, without abolishing biological activity. Our selection of 4' over the 2' derivatization for probe design was further based on the observation that 2'-glucosylation of ascr#2, as in the also naturally occurring ascr#4, completely abolished activity in ascr#2 phenotypic assays.

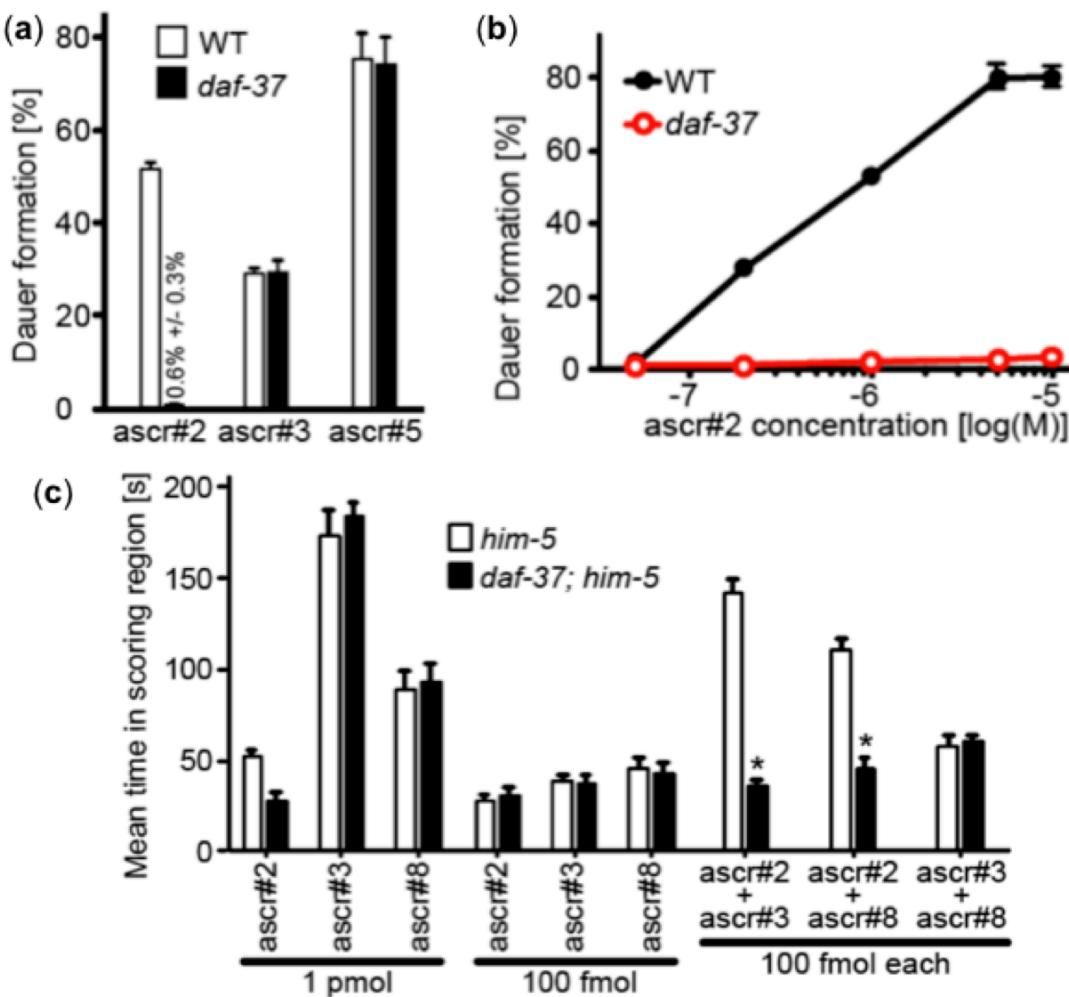
Using the benzaldehyde intermediate (**11**, **Figure 1.4**) we generated the carboxylic acid (**17**), which was coupled to ascr#2, generating both the 2' and 4' linked ascr#2 derivatives, which were separated via HPLC<sup>4,34</sup>. The 4' linked ascr#2 probe was then tested alongside ascr#2 in a dauer formation assay. These bioassays showed that the ascr#2 probe retained some dauer-inducing activity, though dauer induction was much weaker than for unmodified ascr#2 (**Figure 1.6**), likely due to the relative bulk of the added moieties. However, the partial retention of biological activity of our ascr#2 probe demonstrated that targeted derivatization, preserving characteristic structural features, can lead to biologically active photoaffinity probes even for very small signaling molecules.

During synthesis and initial biological characterization of the ascr#8 and ascr#2 probes, we identified in collaboration with lab of Don Riddle, University of British Columbia, two previously undescribed GPCRs named *daf-37* and *daf-38* that appeared to be involved with ascr#2 perception. This development allowed us to use our ascr#2 probe in a candidate approach aiming to demonstrate direct binding of an ascaroside to its receptor.

**The GPCR DAF-37 Specifically Mediates ascr#2 Perception.** To identify interacting proteins of the *C. elegans* Similar Mothers Against Decapentaplegic (SMAD) DAF-8, a

component of TGF- $\beta$  signaling <sup>35</sup>, DAF-8 was immunoprecipitated and analyzed the pulled down proteins by MS. This procedure produced large quantities of DAF-21, a member of the Hsp90 family of molecular chaperones, as well as two previously uncharacterized GPCR's, which were named DAF-37 and DAF-38. Their association with components of the TGF- $\beta$  pathway suggested the possibility that DAF-37 and DAF-38 may partake in dauer pheromone perception. *daf-37* encodes a 465-amino acid GPCR belonging to the serpentine receptor class w (srw) family of chemoreceptors in *C. elegans*, one of several chemoreceptor families that have undergone recent expansion in *Caenorhabditis* <sup>36</sup> (**Figure A.1**). It was found that dauer formation in response to ascr#2 was absent or greatly reduced in *daf-37(0)* mutants (**Figure 1.6**), whereas dauer induction by ascr#3 and ascr#5 in the *daf-37* mutant was similar to wild type (**Figure 1.6**) <sup>2</sup>. Next, it was then investigated whether *daf-37* is required for ascaroside-induced adult behaviors. Wild-type adult males are attracted to ascr#2, ascr#3, and ascr#8, whereas hermaphrodites are repulsed by all three compounds <sup>5,13</sup>. *daf-37* males were not attracted to ascr#2, whereas their attraction to ascr#3 or ascr#8 was comparable to that of wild-type males (**Figure 1.6**).

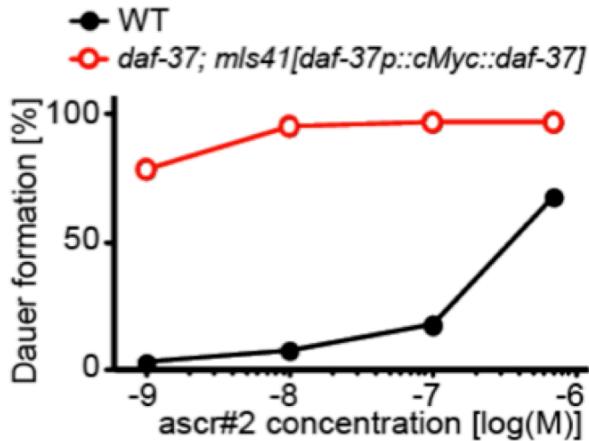
As previously reported <sup>4,5</sup>, mixtures of ascr#2 with ascr#3 or ascr#8 showed synergistic male attraction in wild type, whereas no synergy was detected in the *daf-37* mutant (**Figure 1.6**). *daf-37* hermaphrodites were not repelled by 10 nM of ascr#2, whereas they responded normally to ascr#3 or ascr#5. Furthermore, equimolar addition of ascr#2 to either ascr#3 or ascr#5 did not show any effect (**Figure A.3**), indicating that the additive and synergistic effects of ascr#2 are absent in *daf-37* mutants. These results indicate that *daf-37* mutants are defective in larval and adult behaviors mediated specifically by ascr#2.



**Figure 1.6: Ascaroside-dependent phenotypes in *daf-37* mutant worms.** (a) Dauer induction by ascr#2 (700 nM), ascr#3 (700 nM), or ascr#5 (500 nM) in wild type and *daf-37* mutants. Error bars in (A-D) represent SEM. (b) Dose response curve for ascr#2-dependent dauer formation in wild type and *daf-37* mutants. (c) Attraction of *him-5*(e1467) and *daf-37; him-5*(e1467) males to ascarosides. Individual ascarosides or combinations of two ascarosides were used. Error bars represent SEM (\*, p < 0.01, unpaired t-test, compared to *him-5*).

Next, we asked whether over-expression of DAF-37 rescues the strong *daf-37* loss-of-function phenotype. Worms carrying the *mls41[daf-37p::cMyc::daf-37 cDNA]* transgene in *daf-37* mutant background showed more than 50% dauer formation even at 1 nM ascr#2, a concentration at which wild-type worms form less than 3% dauer (**Figure 1.7**). Under favorable conditions, DAF-37 overexpressing worms develop normally and do not form dauers. This indicates that over-expression of DAF-37 not

only rescues the Daf-d phenotype of the *daf-37* mutant but also confers hyper-sensitive dauer formation in response to ascr#2.



**Figure 1.7: Dose response curve for ascr#2.** Ascr#2 dependent dauer formation in wild type and hypersensitive *daf-37; mls41[daf-37p::cMyc::daf-37]* worms. The hypersensitive worms display a greatly increased sensitivity to ascr#2 with maximal dauer formation being reached at 10 nM ascr#2.

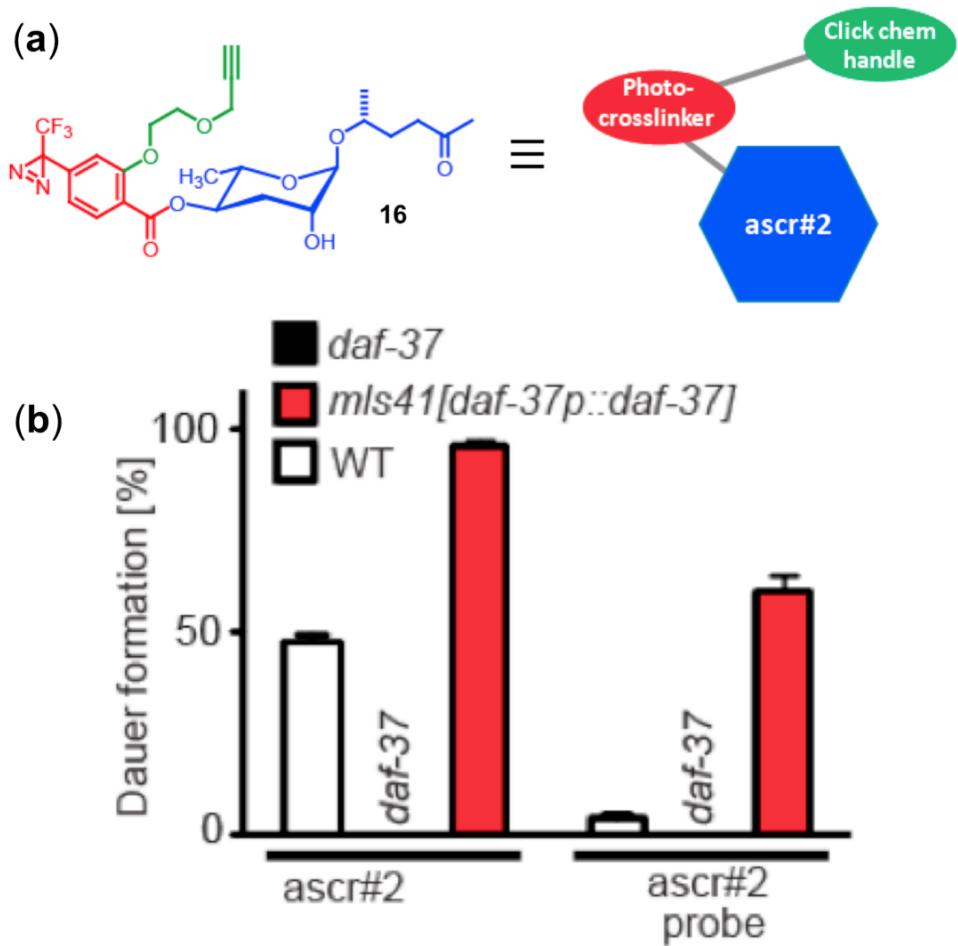
In contrast, DAF-38, a 465 amino acid GPCR with homology to human gonadotropin releasing hormone receptor, participates in perception of ascr#2 as well as other ascarosides. This GPCR was found to not be directly responsible for any signaling by an individual ascaroside as removal of the receptor did not fully abolish ascaroside mediated phenotypes (Figure A.7). Instead, DAF-38 appeared to play a role in enhancing the activity of various ascarosides, and from genetic epistasis studies of DAF-38, it was found to act in parallel or downstream of DAF-37.

Using standard dye-filling assays, it was assessed that the abolition in perception of ascr#2 in *daf-37* mutants was due to the loss of the active receptor and was not caused by generally sensory or structural defects in the nerve cells (Figure A.4). Through epistasis analysis, DAF-37 was placed up-stream of specific dauer network regulators and found to be expressed in the ciliated sensory neurons. This supports the

hypothesis that DAF-37 acts as the first point of contact between the worm and the signal ascr#2.

**Photoaffinity Labeled ascr#2 Binds to DAF-37:** Based on the observation that *daf-37* is required for all ascr#2-dependent phenotypes, we asked whether ascr#2 directly binds to *daf-37* as its bona-fide receptor. Given the challenges inherent to applying photoaffinity-labeling (PAL) approaches to live worms or whole-organism proteomes, we decided to investigate potential binding of ascr#2 to DAF-37 using recombinant DAF-37 expression in human cell culture. For this purpose we generated several constructs for the recombinant expression of *daf-37*. In addition to a C-terminal GFP fusion, we created a *daf-37* construct with a C-terminal 1D4 tag (TETSQVAPA)<sub>1D4</sub>. The small 1D4 epitope is derived from the C-terminal sequence of the optical GPCR rhodopsin and serves as an excellent and unobtrusive tag for GPCRs.

Initially we attempted to carry out a classical PAL-based demonstration of binding of the ascr#2 probe to DAF-37::1D4. First we verified *daf-37* expression in human embryonic kidney cells (HEK293) by observing GFP fluorescence in cells transfected with *daf-37::gfp*. The ascr#2 probe was then delivered to cells expressing DAF-37::1D4. After a brief incubation period, the probe was photocrosslinked to the receptor, generating a covalently bound probe-receptor complex. Using strong denaturing detergents, a whole-cell protein lysate was prepared and the probe-receptor complex biotinylated via click-chemistry. The click chemistry conditions were optimized *in vitro*, using BTTP ligands and conditions described by Wang *et. al.* Following click chemistry, the protein lysate was precipitated and washed to remove the excess click reagents. The washed protein pellet was then resolubilized with standard methods using high amounts of SDS (2-6%) and heating. The detergent was then diluted to allow use of streptavidin beads for the enrichment of the biotinylated probe-receptor complex.

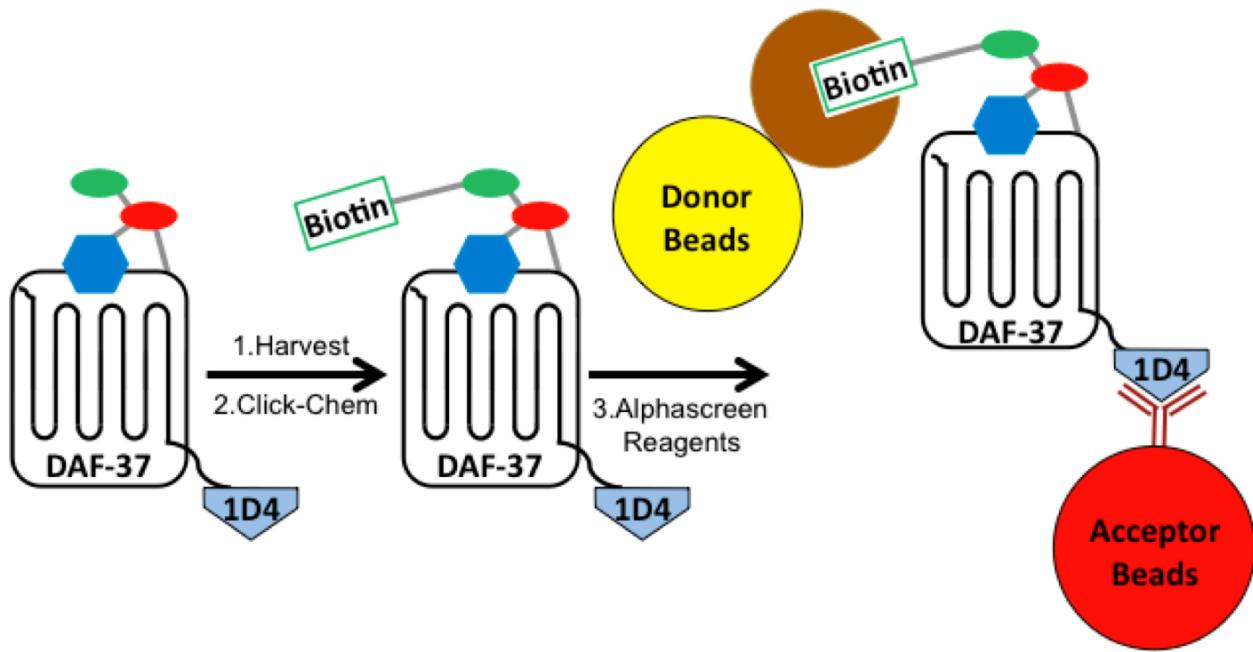


**Figure 1.8: Ascr#2 probe (16) and it's dauer activity.** (a) Structure of ascr#2 probe. (b) Dauer induction by ascr#2 (500 nM) or ascr#2 probe (500 nM). ascr#2 causes significant dauer formation (48%), while the derivatized version of ascr#2, ascr#2 probe, displays a reduced but appreciable ability for dauer induction (4%). Error bars represent S.E.M.

In the final step, the probe-DAF-37:1D4 complex must be eluted from the streptavidin beads for western blot detection. As a result, one would expect to see a band in the western blot corresponding to DAF-37::1D4 in the eluant from a probe treatment and no DAF-37::1D4 where the probe was not given. However, no bands potentially representing DAF-37 were observed in any of the conducted experiments, despite extensive optimization of expression and click chemistry conditions. Likely, the failure of this approach is a direct result of incompatibility of DAF-37, a lipophilic membrane

protein, with the conditions required for elution of biotinylated proteins from streptavidin beads, which involves boiling of the beads in the presence of high amount of SDS (2-8%) and reducing agents. Both heating and strong detergents may lead to aggregation and precipitation of lipophilic proteins, removing them from further biochemical analysis.

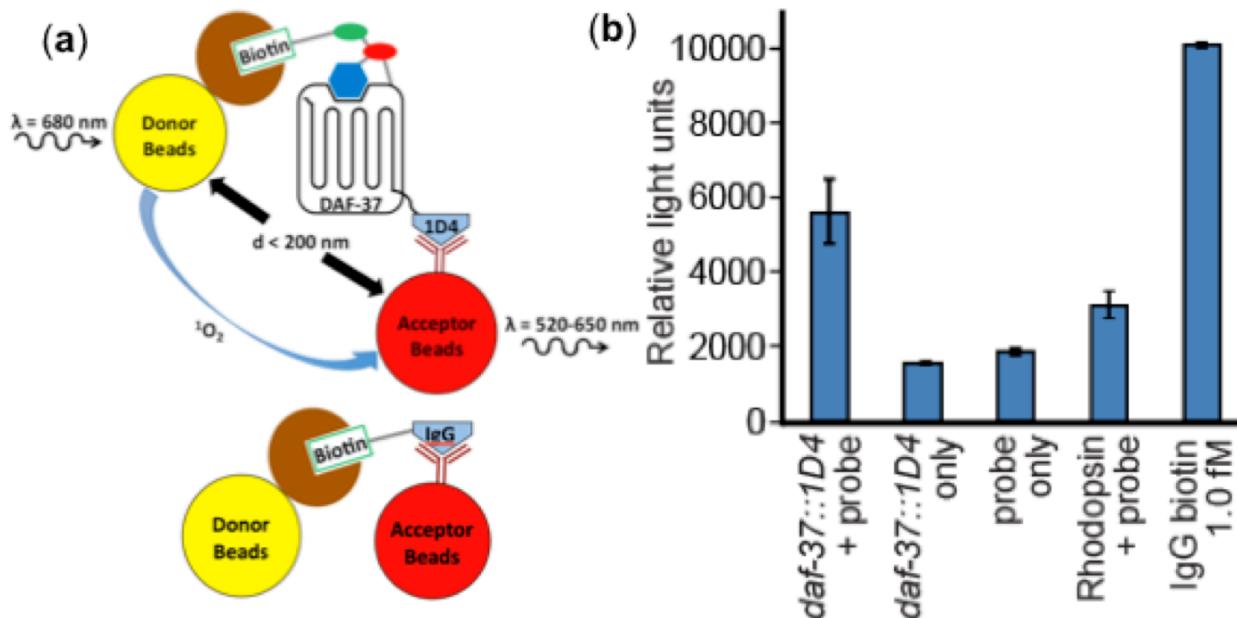
To overcome these limitations, we next attempted to demonstrate the binding of the ascr#2 probe to DAF-37 using amplified luminescence assays (alphascreen<sup>37</sup>) (**Figure 1.9**). Alphascreens usually involve observation of luminescence resulting from close spatial proximity of two different photo-active beads ("donor" and "acceptor" beads, **Figure 1.10**) coated with different antibodies that, for example, bind to two epitopes on one protein, or two different proteins that bind to each other, thereby bringing the two different bead types together to produce a signal. In order to produce a signal, the donor beads must come within roughly 200 nm of acceptor beads<sup>37</sup>. Using the alphascreen approach, the probe-receptor complex was generated via photocrosslinking of ascr#2 probe to DAF-37::1D4 in HEK293 cells as before; however, the method did not require a post click-chemistry precipitation step or streptavidin-based enrichment. Instead of solubilizing the cells with strong detergents, the cells were fixed with paraformaldehyde and disrupted by modest agitation resulting in a cell slurry containing intact membrane fragments, that was pelleted via centrifugation. The click-chemistry step was carried out on this slurry. Subsequently, excess reagents were simply removed through iterative washes, in which the cells were pelleted and the supernatant was carefully removed.



**Figure 1.9: Alphascreen overview.** Ascr#2 probe is photocrosslinked to form a covalently bound complex with DAF-37::1D4. The cells are then harvested and the probe-receptor complex is biotinylated via click-chemistry, washed and prepped for alphascreen analysis.

Once sufficiently washed, 1D4-antibody coated acceptor beads and streptavidin coated donor beads were added to cell slurry containing the labeled probe-receptor complex (**Figure 1.10**). Any crosslinked ascr#2 probe/DAF-37::1D4 complex should induce co-localization of donor and acceptor beads, resulting in emission of light at 520-620 nm (**Figure 1.10**). As shown (**Figure 1.10**), addition of ascr#2 probe to *daf-37::1D4*-expressing cells resulted in an increase of luminescence compared to control conditions lacking either *daf-37::1D4* or ascr#2 probe.

Another version of this alphascreen assay can be carried out with the solubilization of the probe labeled cells with non-denaturing detergents. The click-chemistry step is then performed on this solution and the probe-DAF-37::1D4 complex is immunoprecipitated from the click reaction using the acceptor beads coated in 1D4-antibody.



**Figure 1.10: Alphascreen experimental set up and results for ascr#2 probe. (a)** Schematic representing alphascreen of biotinylated ascr#2 probe covalently bound to DAF-37. Streptavidin-coated donor beads co-localize with 1D4 antibody-coated acceptor beads as a result of the ascr#2 probe being linked to DAF-37. Below, the positive control for the alphascreen is represented in the form of IgG-biotin (b) Alphascreen shows binding of ascr#2 probe to DAF-37-1D4. Incubation of ascr#2 probe (25  $\mu\text{M}$ ) with DAF-37-1D4 expressing cells generates significantly higher luminescence than control experiments without probe, without DAF-37-1D4 transfection, or with cells expressing rhodopsin. IgG-biotin at 1 fmol demonstrates a robust positive control signal. As a positive control 1 fmol of IgG-biotin was used. Error bars represent SEM.

Co-transfection of *daf-38* did not further increase luminescence. As an additional control, we transfected cells with human rhodopsin, a related class-A GPCR (*daf-37* belongs to rhodopsin-like class-A GPCRs) that naturally contains a 1D4 epitope. Ascr#2 probe addition to rhodopsin-transfected cells resulted in a small increase of light emission compared to non-transfected cells (Figure 1.10); however, this increase was significantly lower than observed for cells transfected with *daf-37::1D4*, despite the fact that rhodopsin expression levels were dramatically higher than those of *daf-37* (Figure A.2). These results show that the ascr#2 probe specifically binds to DAF-37.

***daf-37* is Expressed in ASI and ASK Chemosensory Neurons with Distinct Roles in Each.** Additional studies in collaboration with the Riddle lab showed that expression of *daf-37* in different neurons was responsible for the observed life stage dependant phenotypic responses. In a *daf-37* mutant background, ASI-specific expression resulted in L2 larvae hyper-sensitive to dauer formation (**Figure 1.8**) whereas expression in the ASK neurons of adult hermaphrodites caused hyper-sensitivity to ascr#2 in the repulsion assay (**Figure A.6**)

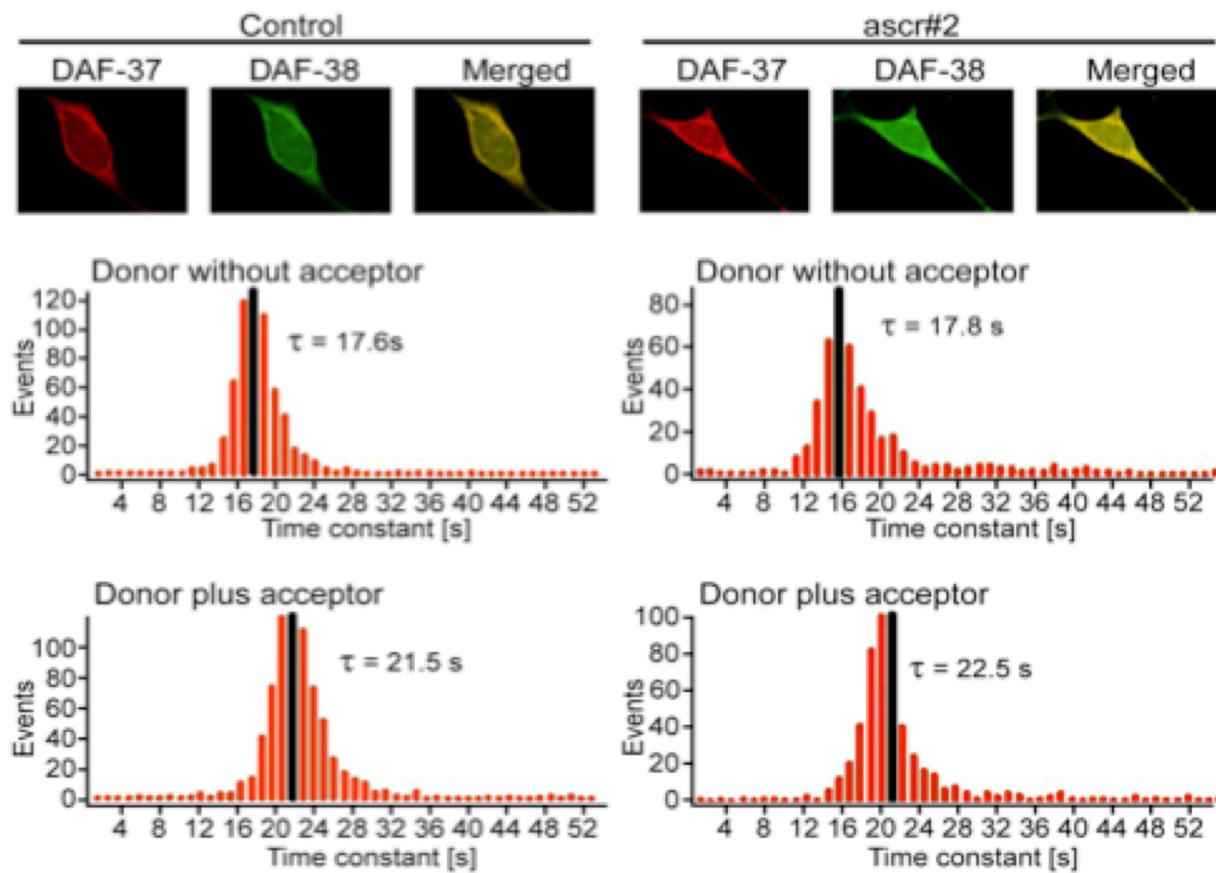
**DAF-37 Forms Homodimers and Heterodimers with DAF-38.** Based on Fluorescence Resonance Energy Transfer (FRET) studies carried out in collaboration with the Kumar lab (University of British Columbia), it was shown that *daf-37* exists as homodimer and as a heterodimer with *daf-38*. Using the Photobleaching (Pb)-FRET assay, it is possible to determine if a GPCR exists as a monomer, homodimer or heterodimer as well as investigating what role the ligand plays in the stability of these states <sup>38-40</sup>. In the Pb-FRET assay, the receptor(s) of interest is expressed recombinantly with an epitope tag on the extracellular N-terminal of the receptor(s). Using immunohistochemical conditions, antibodies with FRET compatible donor and acceptor fluorophores are specifically localized to the N-terminal epitope of the receptor(s). The sample is then exposed to light of a specific wavelength that is capable of causing excitation and subsequent photobleaching of the donor fluorophore present on the antibody. This photobleaching light is delivered at a low enough intensity that allows the photobleaching process to occur a measurable time period of 1-2 minutes. The time taken for the donor fluorophore to photobleach is measured in the absence ( $\tau_{D-A}$ ) and the presence of the acceptor ( $\tau_{D+A}$ ). In the case where both the donor fluorophore and acceptor fluorophore are present and specifically localized within a 100 Å radius of each other, a FRET will occur prolonging the time taken to photobleach the donor fluorophore ( $\tau_{D+A}$ ). The increase in the Pb constant from ( $\tau_{D-A}$ ) to ( $\tau_{D+A}$ ) can be represented as the FRET efficiency  $E$ , where  $E = 1 - ((\tau_{D-A}) / (\tau_{D+A})) \times 100$ . Significant  $E$

values can be used to determine if homo/heterodimerization has occurred, with  $E$  values of 11 and 15% corresponding to in cell receptor dimerization levels of 24 and 30%, respectively<sup>38</sup>.

A N-terminally c-myc tagged DAF-37 construct, *c-myc::daf-37* was monotransfected into HEK293 cells and subjected to Pb-FRET assay using a anti-c-myc antibody conjugated to fluorescein isothiocyanate (FITC) as the donor and an anti-c-myc antibody conjugated to Cyanine Dye 3 (Cy3) as the acceptor. In the absence of the acceptor Cy3 antibody, an average photobleaching time constant ( $\tau_{D-A}$ ) of 18.4 s was observed. When both the donor and acceptor antibodies were delivered a photobleaching time constant ( $\tau_{D+A}$ ) of 21.6 s was observed. The increase photobleaching results in FRET efficiency of  $15.4 \pm 0.5\%$  (**Table A.1**) and this indicates that DAF-37 is present as a homodimer in these cells. Upon treatment with ascr#2, the relative Pb-FRET efficiency increased slightly to  $19.7 \pm 0.7\%$  (**Table A.1**), suggesting agonist-induced stabilization of a receptor complex. HEK293 cells cotransfected with DAF-37/DAF-38 displayed a relative Pb-FRET efficiency of  $18.3 \pm 0.8\%$  (**Table A.1**), indicating that the receptors exist as heterodimers. Again, a small but significant increase in relative Pb-FRET efficiency was observed upon ascr#2 treatment,  $21.5 \pm 0.0\%$  (**Figure 1.11**). The specificity of heterodimerization was confirmed in control experiments based on Pb-FRET analysis of cells cotransfected with DAF-37 and rhodopsin or DAF-37 and the chimeric somatostatin receptor SST5CR1, which are both incapable of forming dimers (**Figure A.8, A.9**).

Another supporting factor that demonstrated that heterodimerization of DAF-37 and DAF-38 is required for ascr#2 signaling, was provided by monitoring the decrease in adenyl cyclase formation in DAF-37/DAF-38 mono- and bi-transfected HEK293T cells. Only weak inhibition of FSK-stimulated cAMP levels was observed upon treatment with ascr#2 in monotransfected cells. In contrast, DAF-37/DAF-38 cotransfected cells displayed significant inhibition of cAMP levels at 100 nM of ascr#2 (**Figure A.10**). Higher concentrations of ascr#2 (500 nM) resulted in lower inhibition. These data

suggest that presence of both DAF-37 and DAF-38 is required for a fully functional receptor complex.



**Figure 1.11 Microscopic Pb-FRET analysis in HEK-293 cells coexpressing DAF-37 and DAF-38 reveals heterodimerization of these two GPCRs.** Representative photomicrographs illustrating DAF-37 (red), DAF-38 (green) and colocalization (yellow) in HEK-293 cells. Pb-FRET microscopy on control (A) and 100 nM ascr#2 treated cotransfected HEK-293 cells (B) were performed as described in Methods. Histograms represents pixel by pixel analysis of time constant of donor in absence or presence of acceptor. Gaussian mean time constants ( $\tau$ ) are shown in black. Note the increased time constant ( $\tau$ ) of donor in presence of acceptor, indicating strong interaction between DAF-37 and DAF-38.

**Discussion:** Ascarosides affect nearly every aspect of *C. elegans* life history, and elucidation of the molecular mechanisms of ascaroside perception will form an important component of advancing this model organism's biology. The fact that ascaroside signals consist of mixtures of several different compounds could suggest that each ascaroside is sensed by one or more specific GPCRs. Alternatively, multiple ascarosides may bind to one or several receptors in a (partially) redundant manner. The nature of ascaroside signaling is characterized by a significant amount of molecular redundancy<sup>3</sup>, synergy<sup>4,5,13</sup>, and plasticity<sup>3-5</sup>. This complexity makes the specific characterization of ligand-receptor pairings' significantly challenging using standard chemical genetic approaches. Our development of ascaroside-based PAL probes demonstrated that careful design probe allows to retain characteristic biological activity of their parent ascaroside, and that ascaroside probes can be used to unambiguously show direct binding of ascaroside probes to their cognate receptors.

We describe the chemical synthesis and biological evaluation of two structurally distinct ascaroside photoaffinity probes, based on ascr#8 and ascr#2. Although both of these probes utilize the same aromatic probe core, the manner in which they incorporate this aryl moiety is strikingly different. In the case of ascr#8 the probe is directly nested into the carbon back-bone of ascr#8, utilizing the PABA group as structural base for the aromatic probe. In the case of ascr#2, we used structural insights from ascaroside signaling that allowed us to identify the alpha-methyl ketone as a key chemical moiety required for ascr#2's signaling activity. By placing the probe on the 4' hydroxy group of the ascarylose ring we able to retain a small but relevant amount of biological activity, showing that even the more structurally simple ascarosides can be derivatized if relevant structural features are preserved.

Our photoaffinity labeling studies show that specific recognition of ascr#2 by DAF-37 is associated with direct binding of the ascaroside to this GPCR, whereas direct molecular interactions between other ascaroside receptors, *srg*- and *srbc*-family GPCRs and ascarosides have not been demonstrated. Extension of this labeling approach to

other ascarosides may facilitate identification of additional specific receptors. Because of the small size of ascr#2, attachment of the photosensitive sidechains likely affected the binding properties of the synthesized ascr#2 probe. The structures of many of the more recently identified ascarosides are significantly larger<sup>7</sup>, and thus may enable design of probes with better crosslinking efficacy and identification of binding sites via MS-based proteomics. The PAL alphascreen approach highlighted here can be used as an efficient way to screen future ascaroside receptor pairings.

Our finding that DAF-37 expression in the ASI neurons mediates ascr#2-dependent dauer formation, whereas DAF-37 expression in the ASK neurons regulates ascr#2-dependent hermaphrodite repulsion demonstrates that perception of one ascaroside by the same receptor expressed in two different neurons can mediate two different phenotypes. Whether *srg-36/37* contribute similarly to ascr#5-dependent hermaphrodite repulsion is unclear, but is suggested by the fact that transgenic expression of *srg-36* and *srg-37* in neurons associated with adult social behaviour, ASH neurons, where these receptors are not normally expressed, resulted in ascr#5-specific avoidance behavior<sup>20</sup>.

Our results indicate that DAF-37 is required for specific recognition of ascr#2, and that its heterodimerization with DAF-38 is required to form a functional complex for signal transduction. Many recent studies suggest that GPCRs associate as dimers or even higher order oligomers<sup>41-43</sup>. For example, the GABA-b receptor forms a heterodimer in the endoplasmic reticulum and is targeted to the cell surface as a preformed dimer without any agonist dependent regulation<sup>44</sup>. It appears that the complex signaling properties of the ascarosides may in part result from the interaction of several different types of ascaroside receptors, including highly structure specific GPCRs that directly bind to ascarosides as well as more promiscuous GPCRs that either bind to several different ascarosides or form heterodimers with several specific ascaroside-binding GPCRs. The potential to use PAL probes to provide a further means

of comprehension for the GPCRs is a yet unexplored area, but could provide valuable insights into the signaling capabilities of these higher order receptor structures.

## REFERENCES

- (1) Jeong, P. Y.; Jung, M.; Yim, Y. H.; Kim, H.; Park, M.; Hong, E.; Lee, W.; Kim, Y. H.; Kim, K.; Paik, Y. K. *Nature* **2005**, *433*, 541.
- (2) Butcher, R. A.; Fujita, M.; Schroeder, F. C.; Clardy, J. *Nat Chem Biol* **2007**, *3*, 420.
- (3) Butcher, R. A.; Ragains, J. R.; Kim, E.; Clardy, J. *PNAS* **2008**, *105*, 14288.
- (4) Pungaliya, C.; Srinivasan, J.; Fox, B. W.; Malik, R. U.; Ludewig, A. H.; Sternberg, P. W.; Schroeder, F. C. *Proc Natl Acad Sci U S A* **2009**, *106*, 7708.
- (5) Srinivasan, J.; Kaplan, F.; Ajredini, R.; Zachariah, C.; Alborn, H. T.; Teal, P. E.; Malik, R. U.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *Nature* **2008**, *454*, 1115.
- (6) Srinivasan, J.; von Reuss, S. H.; Bose, N.; Zaslaver, A.; Mahanti, P.; Ho, M. C.; O'Doherty, O. G.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *PLoS Biol* **2012**, *10*, e1001237.
- (7) von Reuss, S. H.; Bose, N.; Srinivasan, J.; Yim, J. J.; Judkins, J. C.; Sternberg, P. W.; Schroeder, F. C. *J Am Chem Soc* **2012**, *134*, 1817.
- (8) Yamada, K.; Hirotsu, T.; Matsuki, M.; Butcher, R. A.; Tomioka, M.; Ishihara, T.; Clardy, J.; Kunitomo, H.; Iino, Y. *Science* **2010**, *329*, 1647. (9) Riddle DL, Albert PS. *C. elegans II 2nd Ed* **1997**. Chap 26.
- (10) Hu, P. J. *WormBook* **2007**, <http://www.wormbook.org>.
- (11) Jensen, V. L.; Simonsen, K. T.; Lee, Y.-H.; Park, D.; Riddle, D. L. *PloS one* **2010**, *5*, e15902.
- (12) Golden, J. W.; Riddle, D. L. *Molecular & General Genetics* **1985**, *198*, 534.
- (13) Macosko, E. Z.; Pokala, N.; Feinberg, E. H.; Chalasani, S. H.; Butcher, R. A.; Clardy, J.; Bargmann, C. I. *Nature* **2009**, *458*, 1171.
- (14) Birnby, D. A.; Link, E. M.; Vowels, J. J.; Tian, H.; Colacurcio, P. L.; Thomas, J. H. *Genetics* **2000**, *155*, 85.
- (15) Kimura, K.; Tissenbaum, H.; Liu, Y.; Ruvkun, G. *Science* **1997**, *277*, 942.
- (16) Riddle, D.; Swanson, M.; Albert, P. *Nature* **1981**, *290*, 668.
- (17) Ludewig, A. H.; Izrayelit, Y.; Park, D.; Malik, R. U.; Zimmermann, A.; Mahanti, P.; Fox, B. W.; Bethke, A.; Doering, F.; Riddle, D. L.; Schroeder, F. C. *Proc Natl Acad Sci U S A* **2013**, *110*, 5522.
- (18) Ludewig, A. H.; Schroeder, F. C. *WormBook* **2013**, 1.
- (19) Kim, K.; Sato, K.; Shibuya, M.; Zeiger, D. M.; Butcher, R. A.; Ragains, J. R.; Clardy, J.; Touhara, K.; Sengupta, P. *Science* **2009**, *326*, 994.
- (20) McGrath, P. T.; Xu, Y.; Ailion, M.; Garrison, J. L.; Butcher, R. A.; Bargmann, C. I. *Nature* **2011**, *477*, 321.
- (21) Li, X.; Cao, J.-H.; Li, Y.; Rondard, P.; Zhang, Y.; Yi, P.; Liu, J.-F.; Nan, F.-J. *J. Med. Chem.* **2008**, *51*, 3057.
- (22) Edison, A. S. *Curr Opin Neurobiol* **2009**, *19*, 378.
- (23) Pungaliya, C.; Srinivasan, J.; Fox, B. W.; Malik, R. U.; Ludewig, A. H.; Sternberg, P. W.; Schroeder, F. C. *Proc Natl Acad Sci USA* **2009**, *106*, 7708.
- (24) Ludewig, A. H.; Malik, R. U.; Schroeder, F. C. 2009.
- (25) Vodovozova, E. L. *Biochemistry (Moscow)* **2007**, *72*, 1.
- (26) Brunner, J.; Senn, H.; Richards, F. M. *J Biol Chem* **1980**, *255*, 3313.

- (27) Mayer, T.; Maier, M. E. *Eur. J. Org. Chem.* **2007**, 2007, 4711.
- (28) MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J. *J Am Chem Soc* **2007**, 129, 14560.
- (29) Nakashima, H.; Hashimoto, M.; Sadakane, Y.; Tomohiro, T.; Hatanaka, Y. *J Am Chem Soc* **2006**, 128, 15092.
- (30) Park, K. D.; Liu, R.; Kohn, H. *Chemistry & Biology* **2009**, 16, 763.
- (31) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. *J Am Chem Soc* **2009**, 131, 4967.
- (32) Breinbauer, R.; Kohn, M. *Chembiochem* **2003**, 4, 1147.
- (33) Srinivasan, J.; Kaplan, F.; Ajredini, R.; Zachariah, C.; Alborn, H. T.; Teal, P. E. A.; Malik, R. U.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *Nature* **2008**, 5.
- (34) Mayer, T. *European Journal of Organic Chemistry* **2007**, 4711.
- (35) Park, D.; Estevez, A.; Riddle, D. L. *Development* **2010**, 137, 477.
- (36) Thomas, J. H.; Robertson, H. M. *Bmc Biology* **2008**, 6, 42.
- (37) Ullman, E. F.; Kirakossian, H.; Singh, S.; Wu, Z. P.; Irvin, B. R.; Pease, J. S.; Switchenko, A. C.; Irvine, J. D.; Dafforn, A.; Skold, C. N. *Proceedings of the National Academy of Sciences* **1994**, 91, 5426.
- (38) Rocheville, M.; Lange, D. C.; Kumar, U.; Sasi, R.; Patel, R., C.; Patel, Y., C. *J Biol Chem* **2000**, 278, 7862.
- (39) Grant, M.; Patel, R. C. *J Biol Chem* **2004**, 279, 38636.
- (40) Somvanshi, R. K.; Billova, S; Kharmate, G.; Rajput, P. S.; Kumar, U. *Cell Signal* **2009**, 21, 1396.
- (41) Bai, M. *Cellular signalling* **2004**, 16, 175.
- (42) Terrillon, S.; Bouvier, M. *EMBO reports* **2004**, 5, 30.
- (43) Milligan, G. *Current opinion in pharmacology* **2010**, 10, 23.
- (44) White, J. H.; Wise, A.; Main, M. J.; Green, A.; Fraser, N. J.; Disney, G. H.; Barnes, A. A.; Emson, P.; Foord, S. M.; Marshall, F. H. *Nature* **1998**, 396, 679.

## CHAPTER 2

### ADVANCEMENTS TOWARDS THE DIFFERENTIAL LABELING OF SMALL MOLECULE BINDING SITES USING PHOTOAFFINITY PROBES COUPLED TO CLEAVABLE STABLE ISOTOPE CLICK-CHEMISTRY TAGS (CSICT).

***Introduction:*** Understanding the interaction between signaling molecules and their receptors has long been the goal of both applied and basic chemical research. The desire to deconvolute how a chemical signal connects different processes and locations within a single organism, or among organisms of the same or several different species has driven academic and pharmaceutical research<sup>1</sup> since their inception.

The interest begins with a molecule and the observation of a particular phenotype or biologically phenomena that is correlated to the presence of this compound. Once the identity of the molecule has been established, through analytical characterization and synthetic replication of the chemical entity, the question quickly turns to how this molecule could be eliciting the phenotype.

This question can be broken into two related but distinct areas. One is deciphering the initial point of contact between the molecule and organism. The other is the investigation of downstream signaling pathways required for the biological response. The former hinges upon the molecular interaction or binding of the compound to a macromolecule, be it a protein, DNA or RNA in the organism. This chapter will focus on developing approaches for the deconvolution<sup>2</sup> and characterization of the initial interaction between the biologically active small molecule and its receptor protein. Other examples of proteins that interact with small molecules are enzymes and transcription factors. Whether it is investigating how a biologically relevant signaling molecule or potential drug candidate interacts with a receptor protein, the same basic methods are used. These investigatory methods fall into two broad categories, chemical genetics<sup>3-5</sup> and top-down ligand receptor identification<sup>3-6</sup>.

**Chemical Genetics.** Chemical genetics relies on advancements in genome sequencing, molecular biology and gene silencing that have been pioneered over the last two decades. This method is termed a bottom-up approach, in which the potential binding protein is removed from the biological system via genetic manipulation; e.g. gene knock-outs, loss of function mutants or gene silencing with RNAi<sup>7</sup>. If the phenotype of interest is no longer observed upon compound treatments, when the gene for the candidate receptor has been removed from the biological system, it indicates that either the protein is the receptor for the molecule or that it is a signaling protein, acting downstream of the binding event. To characterize accurately the protein's placement in the signaling pathway requires deeper knowledge of the system. For an undescribed pathway this can only be achieved through the further investigation of other pathway components, via similar genetic manipulations and biochemical *in vitro* techniques. Although this bottom-up method is capable of identifying small-molecule receptors, it is subject to several limiting factors. One, is that the generation of genetic mutants can be a time consuming task and is not always trivial<sup>8</sup>. Two, by removing proteins from a signaling pathway, the biological system can be perturbed in unforeseen ways making the direct comparison of the mutant phenotype to that of the native system difficult to interpret<sup>1,4</sup>. Three, the production of certain proteins for *in vitro* assays is not always easily achievable, with some proteins proving particularly challenging to express in sufficient quantities recombinantly. Four, even if enough of a candidate protein can be generated for binding assays, the interaction of the small molecule may be mediated by the presence of another protein, co-factor or co-receptor. If this other entity facilitates the docking of the compound, it may also need to be included in the *in vitro* binding assays<sup>1,4</sup>. The final but potentially largest limitation of bottom-up approaches is that simply focusing on a single protein, will frequently ignore lower level auxiliary or "off-target" binding events that might still be of biological relevance. This has been highlighted by the frequent failure of pharmaceuticals for various indications in late

stage clinical trials due to an insufficient understanding of such “off target” effects of candidate drugs<sup>9,10</sup>.

**Bottom-up and Top-down.** To tackle the shortcomings of bottom-up chemical genetics, novel top-down identification methods have been developed. Top-down characterization techniques rely on molecular probes that are structurally based on the ligand of interest but are capable of binding the receptor covalently, while providing a purification handle to isolate the ligand-receptor complex for characterization via mass spectrometric (MS) methods<sup>3,5,11,12</sup>. These probes are considered tri-functional as they maintain their ability to bind the receptor of interest, establish an irreversible covalent linkage to the protein and can ultimately be used to purify the probe-receptor complex from the proteome for MS identification. This top-down labeling strategy is only capable of identifying proteins that directly bind with the ligand and not those involved in downstream signaling. Importantly, top-down approaches provide an inherently less biased method for receptor identification than bottom-up methods as the probe labeling occurs in a native biological context with all of the relevant potential players present. The candidate receptors “hits” generated from top-down approaches are then verified and validated with directed chemical genetic analyses<sup>4,5</sup>.

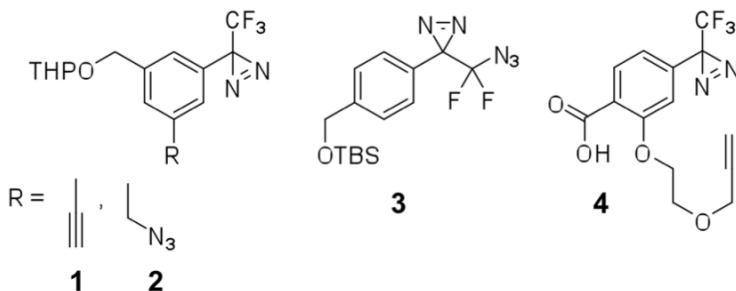
**Affinity Based Protein Profiling.** Small molecules in the form of peptides, secondary metabolites, or drug-like synthetic compounds typically interact with their binding protein(s) in a non-covalent, reversible fashion. Exceptions to this have been used in a top-down receptor identification strategy called Activity Based Protein Profiling (ABPP)<sup>2,7,13-15</sup>. This method has proved fruitful but has been primarily limited to the identification of enzymes that have nucleophilic residues present in their active sites, capable of reacting with the electrophilic probes. The more broadly applicable approach of Affinity Based Protein Profiling (AfBPP) does not require a receptor to have a specific chemical property to stimulate covalent binding<sup>3,7,8,12,16</sup>. Instead, a photoreactive

functional group is incorporated into the small molecule of interest, allowing the irreversible conjugation of the compound to the biological target upon specific UV-treatment. The trifluorodiazirines (TFD) have become the photocrosslinking group of choice due to their compact nature, chemical stability and biologically compatible photoirradiation profile<sup>1,4,7,16,17</sup>.

In addition to the ability to bind and covalently photo-crosslink the candidate receptor, a functional AfBPP probe must provide chemical functionality that is capable of undergoing a bio-orthogonal reaction. This reaction will then be used to specifically couple the photocrosslinked probe-receptor complex to a purification handle. Suitable reactions for this include the Staudinger-Bertozzi ligation<sup>1,4,8,18</sup>, the copper catalyzed Huisgen 1,3-dipolar cycloaddition (click reaction)<sup>13,18-21</sup>, the copper-free click reaction variants<sup>22-24</sup>, and the tetrazine ligation<sup>18,25,26</sup>. Although all of these reactions have proven merits, the Cu-mediated click reaction has seen the greatest use due to the structural simplicity of the reactive partners, a terminal alkyne and an azide. By placing one of the bio-orthogonal reactive partners, the alkyne, on the probe and the complimentary azide functionality on a purification handle<sup>27</sup>, the probe-receptor complex can be labeled specifically with an affinity tag such as biotin in a Copper-Catalyzed azide-alkyne cycloaddition (CuAAC)<sup>21</sup>. This biotinylated complex can then be enriched onto streptavidin beads, eluted and subjected to shotgun proteomics analysis to identify the unknown receptor<sup>3,13</sup>. Recent literature has shown that although the alkyne and azide can both be considered bio-orthogonal, placement of the alkyne on the probe and the azide on the biotin yields better results than the reverse<sup>25-28</sup>.

Robust synthetic methods have been developed for the generation of compact aryl tags that contain both the TFD for photocrosslinking and alkyne or azide for click chemistry. These aromatic cores (**Figure 2.1, 1-4**) have an additional benzyl alcohol that can be specifically tuned for the attachment to various biologically active molecules<sup>12,26,29-32</sup>. Although these tags can be easily coupled to a multitude of active compounds with unknown receptors, an understanding of the structure-activity

relationships of the active compound is still essential to choose an effective ligation site. Once the derivatized small molecule is generated, phenotypic data showing biologically activity of the probe compound is still required before proceeding with a labeling and receptor identification efforts.



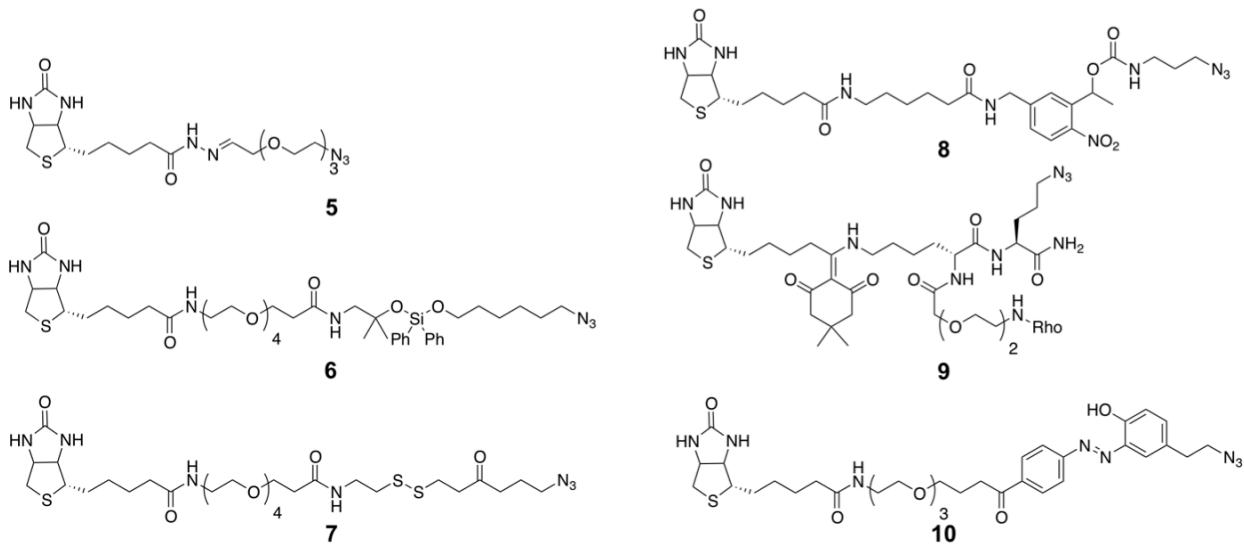
**Figure 2.1: Compact photocrosslinking and clickable probes.**

**Cleavable Linkers.** The strength and selectivity of the biotin-streptavidin interaction underpins the enrichment strategy of AfBPP, ABPP and similar methods<sup>27,33,34</sup>. This strong interaction allows the use of rigorous wash steps that efficiently remove nonspecifically bound proteins from streptavidin beads. However, this strength can cause problems when the enriched biotinylated probe-receptor complex needs to be released from the beads. Typically, elution from streptavidin beads is carried out at 95°C with strong ionic detergents, chaotropic salts and reducing agents. These elevated temperatures cause solubility problems for lipophilic membrane proteins, including G-protein Coupled Receptors (GPCRs), which make up one of biggest classes of signaling receptors<sup>27,28,35,36</sup>. Multi-pass membrane proteins are particular greasy and exposing them to elevated temperature causes these lipophilic receptors to aggregate, preventing further characterization. Along with this elution challenge, biotinylation occurs as an endogenous post-translational modification of proteins. Using streptavidin beads to enrich for the biotinylated probe-receptor complex will also cause an enrichment of the naturally biotinylated proteins, that are not involved with probe treatment. Although such endogenously biotinylated proteins can be identified with the appropriate negative controls, they can still cause problems if present at high levels. The severity of the

problem depends in part on the dynamic range of the MS instrument and the difference in abundance of the peptides associated with the endogenously biotinylated proteins and the “hits” from the probe treatment<sup>11,35,36</sup>. Given that the receptors at the start of signaling cascades are typically present only at very low levels in the cell, circumventing this potential pitfall is of great importance<sup>11,28</sup>.

Significant efforts have been made to improve the elution step of the biotinylated probe-receptor complex from streptavidin beads. They have focused on the use of selectively cleavable biotin linkers that are capable of eluting the desired proteins at low temperatures, while keeping the endogenously biotinylated protein bound to the beads. These mild elution techniques either use an enzymatically cleavable or chemically cleavable linker that can be hydrolyzed under mild and specific chemical conditions. With an enzymatic strategy<sup>28,37</sup>, a peptide cleavage site that is enzyme accessible must be incorporated into the linker; this can introduce significant bulk with the potential to complicate MS identification. The amount of proteolytic enzyme required to efficiently cleave the linker will likely be multiple orders higher than the amount isolated receptor, again raising the issue of dynamic range. More recently, chemical cleavage strategies have superseded enzymatic methods as both of the highlighted problems do not apply to a small molecule reaction.

Several different chemically cleavable linker groups (**Figure 2.2**) have been used in selective release strategies, including imines (**5**)<sup>37,38</sup> and dialkoxydiphenylsilanes (**6**)<sup>21,38,39</sup> that are labile under hydrazine and strong acidic conditions, respectively, disulfide (**6**)<sup>21,39,40</sup> and diazobenzene (**10**)<sup>40,41</sup> groups that are cleaved with reducing agents, as well as the 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl (Dde) protecting group (**9**)<sup>41,42</sup> that is sensitive to hydrazine and photocleavable linkers (**8**)<sup>42,43</sup>. Due to ease of application and efficiency of cleavage, we selected the diazobenzene cleavable linker for development of our Cleavable Stable Isotope Click-Chemistry Tags (CSICT) method.



**Figure 2.2: Previous examples of cleavable biotin linkers.** In compound **9** Rho= rhodamine fluorophore

**Binding Site Identification Strategies.** When a biologically active TFD photoaffinity probe is crosslinked to its cognate receptor, the position of the linkage will occur proximal to the binding interface of the small molecule. This photo-derived modification can then be used to identify the binding site of the small-molecule probe to the protein,

once the probe-receptor complex has been enriched and eluted from the streptavidin beads<sup>12,43</sup>. The isolated receptor complex can be further purified with an SDS-PAGE step or directly trypsinized into peptide fragments for shotgun proteomics analysis to identify the receptor<sup>3,5,6,13,30,44</sup>. In the population of trypsinized peptide fragments, there will be a small amount, ideally one or two unique peptide fragments, that will have the cleaved linker-probe moiety attached to it. Finding the identity and characterization of the precise probe-peptide fragments is a significant challenge due to the number of receptor specific trypsin fragments, as well the presence of background peptides that occur from the trypsinization of contaminant proteins dragged along in the purification process. Various techniques have been employed to impart a specific chemical signature to the probe to facilitate this<sup>6,40,44-48</sup>. This specific chemical signature is then used to identify those probe-peptide fragments that should be further investigated and fully characterized. Previous literature examples for methods facilitating the recognition of probe-peptide conjugates include the incorporation of hot isotopes to the small molecule probes<sup>41-43,45,46</sup>, the utilization of elements with characteristic isotopic abundances<sup>12,47,48</sup> and Stable Isotope Labeling by Amino acids in Cell Culture (SILAC) . Although useful, these methods have significant drawbacks and limitations, including the difficulty in handling radioactive isotopes and the small relative mass difference of stable isotopes that are naturally abundant in elements such as bromine. Incorporating bromine into PAL probes or purification linkers to identify the specific probe-peptide fragments, as described recently, has the particular limitation that the characteristic difference of 2 amu between <sup>79</sup>Br and <sup>81</sup>Br, present on a brominated probe-peptide fragment is difficult to detect by ESI-MS, given the naturally complex isotopic patterns of multiply-charged peptide ions. Unfortunately it is not possible to compare non-brominated and brominated PAL probe/linker experiments via LC-MS, as the non-brominated probes or linkers generate non-isobarically equivalent peptide fragments that will have different ionization behavior, fragmentation patterns, and retention times in

LC-MS analysis. These limitations have hindered the development of stable-isotope labeling methods for PAL-based receptor characterization.

One novel and noteworthy method outlined by Hucle *et. al*,utilizes the isobaric quantitative proteomic strategy of SILAC coupled to competition experiments between photoaffinity probes and native ligands<sup>12,44</sup>. In its initial application, this technology proved useful in the identification of multiple known and unknown cholesterol binding proteins, highlighting the power of stable isotope strategies coupled to high-resolution MS. Using a cholesterol photoaffinity probe (Capable of click-chemistry) coupled to a SILAC strategy, the authors were able to identify a variety of previously described and novel cholesterol binding proteins. In their method a cholesterol derived photoaffinity probe was delivered and crosslinked to the heavy isotope labeled human cells, while the same sterol probe was crosslinked to the light labeled cells in the presence of a ten fold excess of unmodified cholesterol. The completion assay that takes place in the light cell experiment results in a significant decrease in sterol probe phootoaffinty labeling. As result, when the probe-labeled proteins from both heavy/light experiments are biotinylated via click-chemistry, enriched and trypsinized, the amount of probe-protein fragments in the light sample will have significantly decreased. As the SILAC method allows the quantitative comparative analysis of the heavy and light trypsinized peptide fragments, the decrease in light probe-peptides can be observed and used to identify fragments that are specific and required for cholesterol binding. Theses identified probe-peptide fragments of differential intensity can then be fully characterized, yielding the identity of the cholesterol binding protein. This technique marks a significant development in ligand-receptor identification strategies. However, expanding the use of SILAC-based methods beyond cell culture to easily manipulated, simple organisms such as *C. elegans* has been non-trivial<sup>25,44</sup>, and the further expansion of SILAC methods to high organisms is so far an unattainable possibility.

A recent development in the field of receptor identification using top-down ligand-based receptor capture methods is the tri-functional chemical proteomics (TriCEPS)

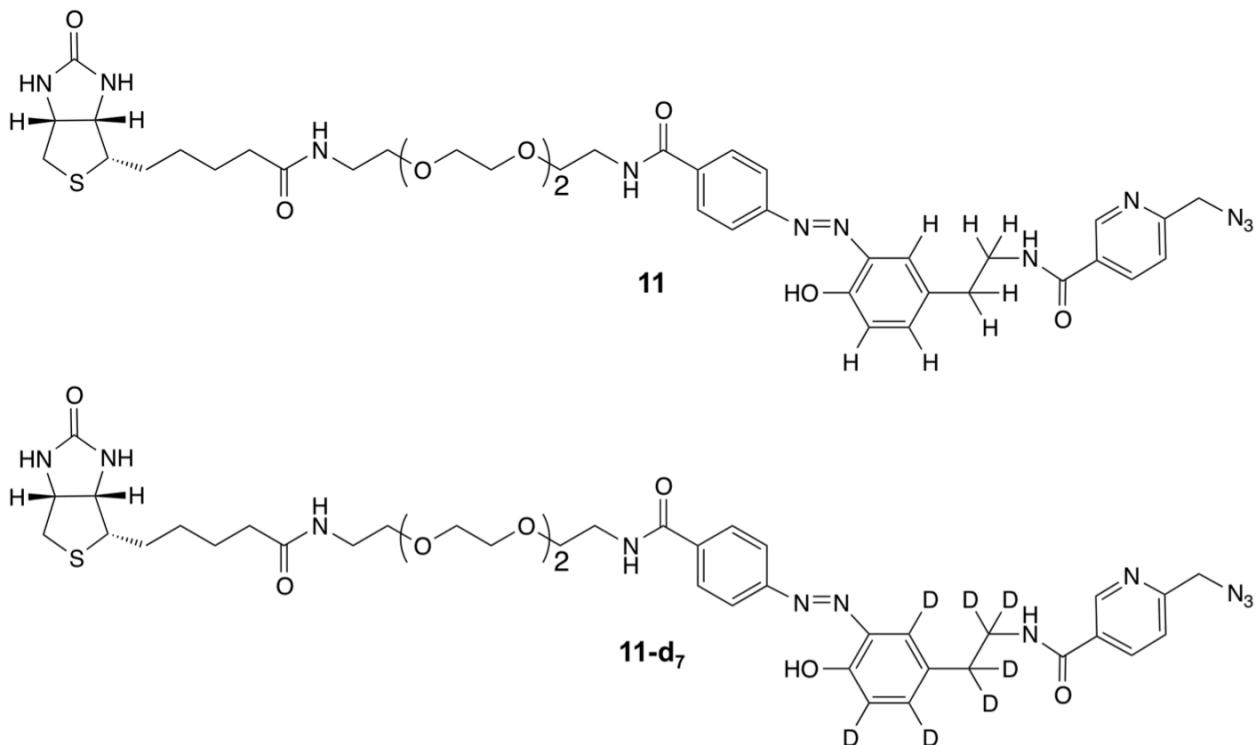
strategy developed by Frei *et. al*<sup>25,28</sup>. In its initial application, this technique was used for the identification of glycosylated membrane receptors using a chemical ligation step to covalently label the receptor, in place of a photo-crosslink approach. Although powerful in this application, the technique is limited to the identification of glycosylated membrane receptors because the chemical ligation of the probe to the receptor relies on treatment of the cells with a mild oxidant to create electrophilic aldehydes on the glycosylated receptor. The probe of the biologically active compound has nucleophilic hydrazine incorporated to it that are capable of attacking the aldehyde of the glycosylated receptor, establishing a covalent linkage. The identification of the labeled receptors is then achieved using statistical comparisons of MS data generated from probe and control experiments. Although TriCEPS is currently limited to identification of glycosylated membrane receptors, there is a lot of interest in this technique with the reagents necessary for probe derivatization recently becoming commercially available.

Although there has been significant developments in the area of ligand-receptor binding site elucidation, each of the methods described above has their own strengths but also their own limitations, namely; reliance on radioactive isotopes, small mass differences between stable-isotopes, limited application to specific biological systems or receptor subtypes and continuous challenge of dynamic range. Drawing from the strengths of stable isotope labeling and cleavable-clickable biotin linkers, we here describe development of a novel method for the characterization and identification of small-molecule binding sites with their receptors, using stable-isotope labeling.

**Stable Isotope-Labeling Strategy.** We sought to develop an isobaric stable isotopic labeling strategy that imparts a robust multi-amu difference to the probe-peptide fragments associated with ligand binding. We selected deuterium and hydrogen as the heavy and light isotope pairing due the ease of synthetic incorporation, non-radioactive nature, and large relative mass difference achievable by incorporation of many deuteriums. Incorporation of deuterium in place of the hydrogen will generally not cause

a drift in the chromatographic retention times of peptides in LC-MS analysis or change their fragmentation patterns<sup>28,35</sup>. As result, the heavy and light probe peptide fragments will elute at the same time point but with a distinct and easily identifiable mass shift<sup>35</sup>. Using this distinct mass-shift difference one can specifically hone in on the specific probe-peptide species, selecting this particular probe-peptide from the complex milieu of other non-probe fragments, where it would otherwise go unnoticed. Once identified, this specific probe-peptide fragment is used to identify the receptor protein it is derived from. More detailed mass spectrometric characterization can then be used to hone in on the site of ligand binding.

**Design of Cleavable Heavy/Light Azido Biotin Click-Chemistry Tags.** Once we had selected the deuterium and hydrogen as the heavy and light isotope pairing, we assessed how these might be incorporated into a PAL click-chemistry based strategy. We chose a standard PAL approach, coupled to a cleavable diazobenzene biotin azide purification method, since diazobenzene seems to be particularly well suited as a cleavable functional group for biotin azide linkers<sup>35,47</sup>. We envisioned that including a cleavable diazobenzene linker could be used to circumvent the problems previously discussed with the release of probe-labeled lipophilic membrane proteins from streptavidin beads. We also chose to use an enhanced click-chemistry azide with proximal copper-chelating pyridine moiety that has recently been shown to assist in the efficiency of the CuAAC reactions<sup>47,49</sup>. The incorporation of copper-chelating click-activating groups adjacent to the azide has been shown to enhance the rate and yield of the triazole click product in CuAAC reactions. With the isotope pairing, cleavable linker and enhanced click-chemistry azide selected the basic structure of the linker was defined (**Figure 2.3, 11**)



**Figure 2.3: Showing the final structure of the cleavable light (11) and heavy (11-d<sub>7</sub>) biotin azide linkers**

The final choice was to decide where and how many deuteriums could be incorporated in **11** to generate a stable isotopic chemical signature, keeping in mind relative ease of synthesis and the need to generate a multi-amu difference between the heavy and the light probe-peptide fragments. Our chosen design (**Figure 2.3, 11-d<sub>7</sub>**) includes seven deuteriums on the portion of the linker that will be left on the probe-receptor complex after streptavidin enrichment and reductive elution. By placing seven deuteriums in place of seven hydrogens a robust chemical signature of seven amu difference will be present in all probe-peptide fragments, addressing the challenges presented by charge dilution of relative mass differences associated with ESI-MS of peptides. "Charge dilution" refers to the problem that for a three-fold positively charged ion, a mass difference of one amu results in a m/z difference of only one third of one amu, making it much more difficult to detect, especially given the complex isotope pattern of large peptide fragments. Using the significant mass difference of seven deuteriums, we were confident that even peptide fragments carrying multiple charges in ESI-MS would still be

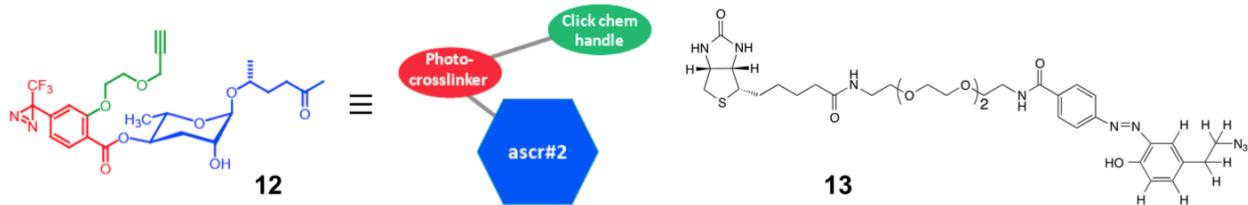
readily distinguishable from their isobarically labeled hydrogen counterparts<sup>35,49</sup>. Moreover, by carrying out the heavy and light click-chemistry on protein lysates from the sample biological replicate, it should be possible to avoid problems associated with variation of receptor expression and differential probe labeling between samples. Following LC-MS-based detection of the specific probe-peptide fragment(s) and bioinformatic identification of the parent protein (or several protein candidates, in the case of short peptides that cannot be unambiguously mapped to the proteome) further analysis may set the stage for characterization of ligand binding sites.

Of great importance for efficient CuAAC is the selection of an appropriate copper-ligand. Previous examples showed that use of the highly water soluble 3-[4-({bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]propanol (BTTP)<sup>21</sup> allows the CuAAC to occur in high concentration protein lysates without the precipitation of proteins, reducing the amount of azide required.

The proposed method should impart a robust differential isotopic signature to the probe-receptor complex in a late-stage click-chemistry step. This late labeling allows the use of small amounts of the deuterated azide (**11-d<sub>7</sub>**) that is used to couple to the alkyne of probe. Comparatively, in the SILAC method, cells must be passaged multiple times to develop a stable isotope signature using large amounts of expensive deuterated media<sup>6</sup>. Coupling in this stable isotopic signature via a click chemistry also imparts a flexibility to the CSICT approach, as the heavy and light azides can be applied to protein lysates derived from any biological origin in which a clickable chemical probe has been used and is not limited to a cell type of specific organism.

**Proof of Concept for Cleavable Linker (**daf-37**).** Before undertaking the synthesis of both the heavy and light cleavable diazobenzene biotin azide (**11**, **11-d7**), we first decided to test if the cleavable diazobenzene biotin linker was compatible with the isolation of lipophilic membrane receptors. To investigate this, we used a simplified first-

generation diazobenzene (**Figure 2.4, 13**) that was readily synthesized using previously described methods<sup>35,47,50</sup>. To test this diazobenzene cleavable linker strategy, we chose the known ligand-receptor pair of ascr#2 and the GPCR DAF-37<sup>32,47,50</sup>.

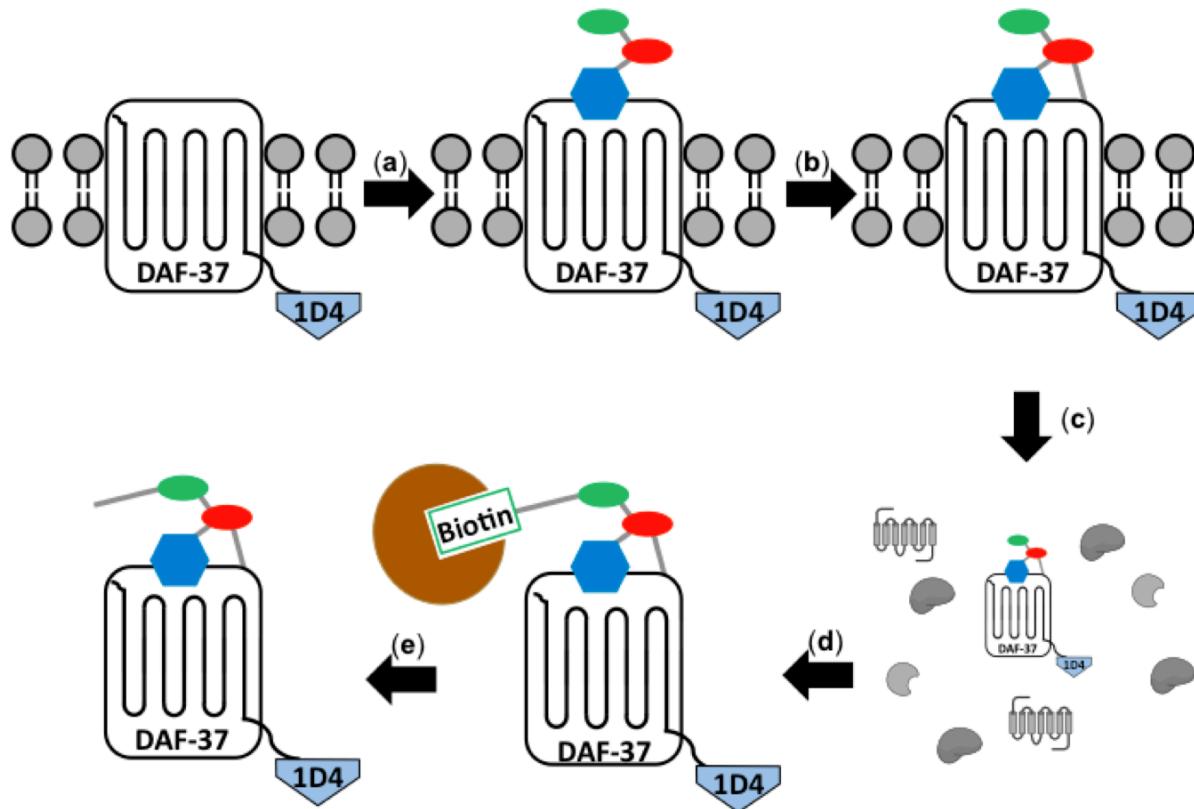


**Figure 2.4: Structure of ascr#2 probe (12) and first generation biotin cleavable linker (13)**

Previously, we have shown that the photoaffinity-probe of ascr#2 (**Figure 2.4, 12**) is capable of binding the recombinantly expressed *Caenorhabditis elegans* GPCR, DAF-37 in an alphascreen-binding assay<sup>32,47,50</sup>. The alphascreen approach was developed because of the incompatibility of the click-chemistry biotinylated DAF-37 probe-receptor complex to be eluted from the streptavidin beads under standard thermal disruption.. Replacing the heating step required for the release of the DAF-37 from the streptavidin beads with mild reductive elution of the diazobenzene cleavable linker (**13**) should circumvent the aggregation of lipophilic GPCRs.

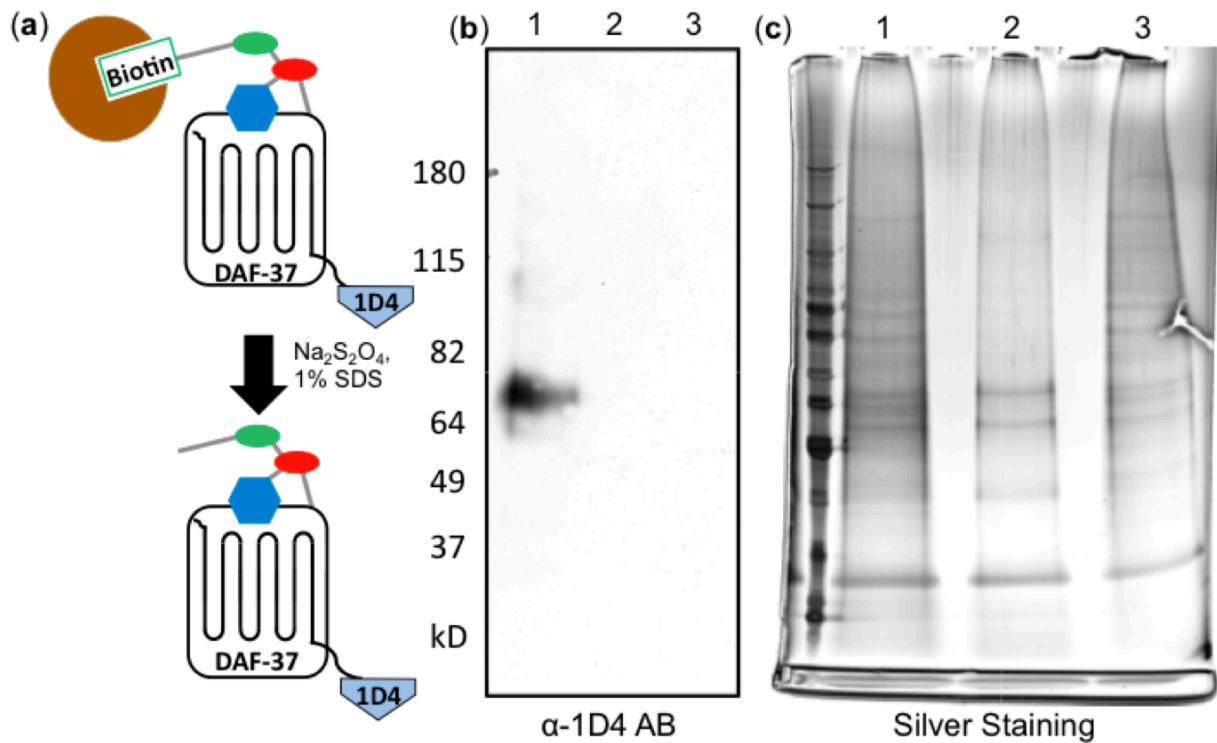
To test this strategy, we used the same DAF-37::1D4 expression system of HEK-293T cells as previously described (**Chapter 2**). We assayed if the diazobenzene (**13**) was capable of biotinyling the ascr#2 probe-DAF-37::1D4 complex via click-chemistry and subsequently allowing its elution under low temperature, mild dithionite reduction conditions<sup>32,47,51</sup> (**Figure 2.5**). Taking three confluent 10 cm plates expressing DAF-37::1D4, 10 μM ascr#2 probe was incubated and photocrosslinked to DAF-37. The cells were then processed and then clicked with the cleavable biotin azide (**13**). After the click chemistry reaction, the excess reagents were removed by precipitating the proteins with

chloroform. This protein pellet would typically be resolubilized using heat in the presence of strong detergents or chaotropic salts<sup>47,51,52</sup>. However, this method is not compatible with lipophilic membrane proteins. Fortunately, we discovered that simply by leaving the protein pellet overnight at room temperature in a suitable resolubilizing buffer (6M urea, 2M thiourea 10 mM HEPES pH 7.4 or 4% SDS, 150 mM NaCl, PBS pH 7.4) is sufficient to coax most of the precipitated proteins back into solution. Once resolubilized, the detergents and chaotropic salts were diluted to ensure compatibility with streptavidin beads. Following incubation with the solubilized protein, the beads were washed and then exposed to dithionite reduction conditions<sup>47,52,53</sup> to release the probe-DAF-37::1D4. The eluants were then desalting and prepped for SDS-PAGE with subsequent western blot and silver staining analysis. The presence of a band in lane one (**Figure 2.6**) and not in lane two of the 1D4-antibody western blot clearly demonstrates the ability of this diazobenzene cleavable biotin azide strategy to specifically enrich ascr#2 probe DAF-37::1D4. As evident from comparing lanes one (probe and DAF\_37::1D4) and 3 (Probe but no DAF-37) in the silver stain, representing the total protein content of the elutions, the non-specific labeling of ascr#2 probe can also be seen by comparing lanes one and three. Lane three, representing a high concentration of ascr#2 probe photocrosslinked to cells transected with GFP (but not DAF-37) serves as a negative control.



**Figure 2.5: Application of cleavable biotin linker (13) to ascr#2 probe (12) and DAF-37::1D4.** (a) Delivery and binding of ascr#2 probe to DAF-37::1D4 being expressed in HEK293T cells. (b) Generation of covalent probe-receptor complex with *in vivo* photocrosslinking of ascr#2 probe to DAF-37::1D4. (c) Solubilization of cells to create a whole cell protein lysate. (d) Biotinylation of probe-receptor complex via click chemistry, precipitation and enrichment of complex on to streptavidin beads. (e) Dithionite elution of probe-receptor complex from the streptavidin beads in preparation for western blot and silver stain analysis

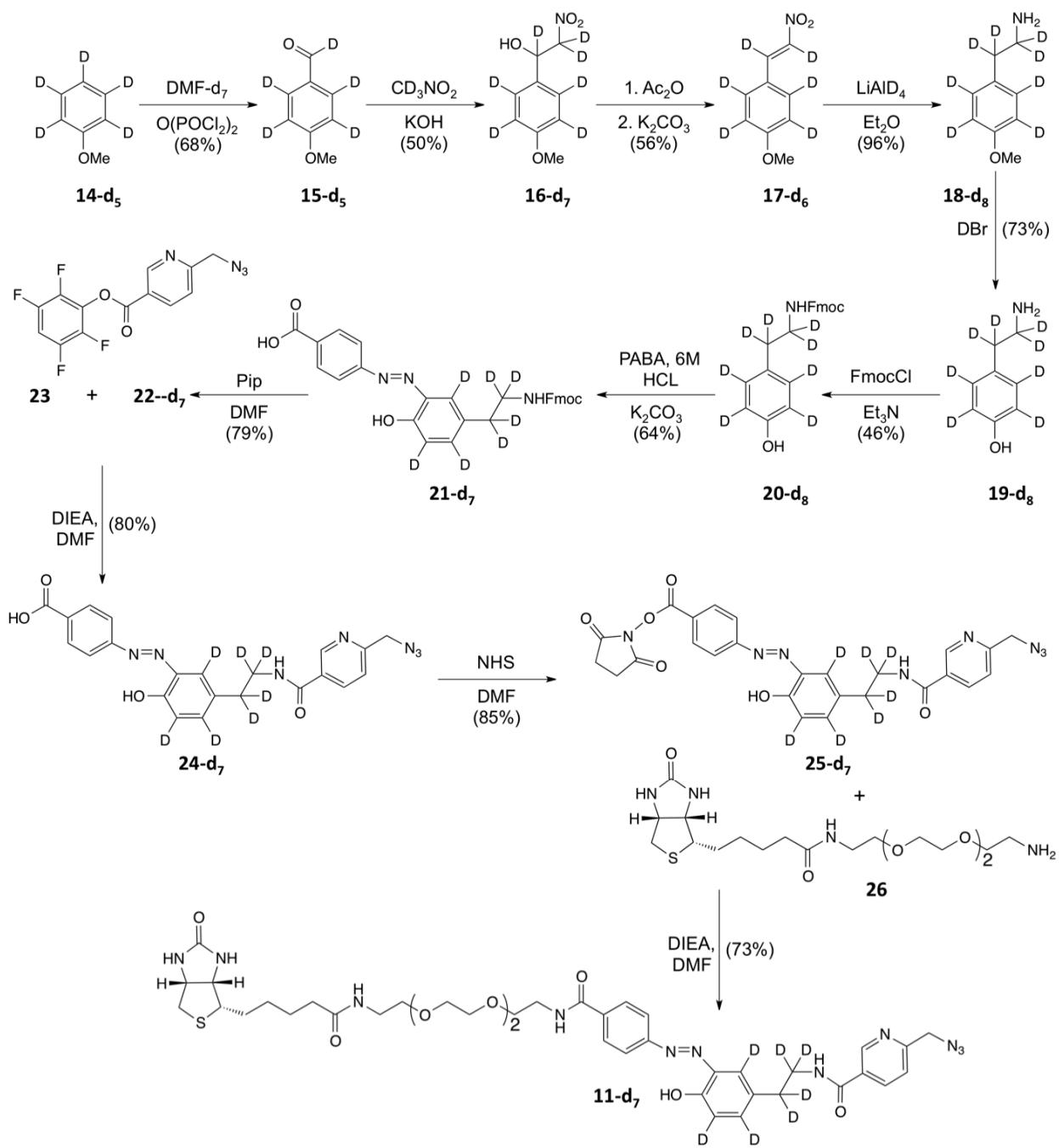
Importantly, background non-specific enrichment resulting from the cleavable azido biotin is low as demonstrated by the relatively small number of bands present in lane two compared to lanes one and three. This shows that the cleavable biotin azide strategy can be used to selective purify alkyne labeled lipophilic membrane receptors and that the diazobenzene moiety is a suitable choice for our heavy and light cleavable stable isotope click-chemistry tags.



**Figure 2.6: Gel analysis of cleavable linker (13) trial.** (a) Scheme showing the generation of the DAF-37 enriched dithionite eluant that was loaded to both gels. (b) Anti-1D4 western blot analysis of dithionite elution of streptavidin beads from ascr#2 probe labeling experiment. The lanes correspond to the dithionite elutions of three experimental conditions; Lane 1 is from a  $10 \mu\text{M}$  treatment of ascr#2 probe (12) on HEK293T cells expressing DAF-37::1D4, lane 2 was generated from cells that were expressing DAF-37::1D4 but underwent no probe treatment and lane 3 serves as a negative control where cells expressing GFP were treated with  $10 \mu\text{M}$  of ascr#2 probe. Strong band for DAF-37::1D4 in lane 1 show specific enrichment of probe labeled DAF-37::1D4 (c) Silver staining analysis of total protein content from the same dithionite elution, showing specific enrichment of proteins with molecular weights similar to that of DAF-37. The similarities between lanes 1 and 3 of the silver staining can be attributed to the non-specific labeling of the ascr#2 probe and are not an artifact of the CSICT purification system as they are not present in lane 2. A comparison of lanes 1 and 3 with 2 shows the non specific labeling as result of probe exposure

**Synthesis of Deuterated and Non-Deuterated Linkers.** The novel heavy (**11-d<sub>7</sub>**) and light (**11**) CSICTs were synthesized in a modular fashion by adapting established synthetic methods (**Figure 2.7**). Whereas the non-deuterated version of the linker could be synthesized from commercially available tyramine, the corresponding deuterated version had to be synthesized from labeled tyramine-d<sub>7</sub> (**19-d<sub>8</sub>**), which, in turn, was prepared from anisole-d<sub>6</sub> in five steps<sup>47,50,53,54</sup>. This tyramine synthesis was first piloted with non-deuterated reagents .

The synthesis starts with conversion of anisole-d<sub>6</sub> into *p*-anisaldehyde-d<sub>6</sub> using the Vilsmeier–Haack and deuterated DMF, following a previously described procedure<sup>50</sup>. However, subsequent Michael addition of deuterated nitromethane did not proceed as expected based on literature precedent<sup>50,54,55</sup>. Following addition of nitromethane, the reaction product did not undergo subsequent elimination required to form the desired nitro-styrene (**17-d<sub>6</sub>**). Instead the benzylic alcohol (**16-d<sub>7</sub>**) was generated as the main product. However, following activation via acetylation, the alcohol quantitatively eliminated to form the desired styrene (**17-d<sub>6</sub>**). The styrene was then reduced with LiAlD<sub>4</sub>, yielding O-methyl tyramine-d<sub>7</sub> (**18-d<sub>8</sub>**). Prior to demethylation of the O-methyl tyramine, the exchangeable hydrogens of the amine were exchanged to deuteriums using repeated treatments with deuterated methanol, in order to prevent D-to-H exchange of the aromatic protons during the harshly acidic conditions of the subsequent elimination. The O-methyl tyramine-d<sub>9</sub> was then exposed to DBr/D<sub>2</sub>O under reflux followed by extraction with deuterated methanol, recrystallization and reverse phase purification. After conversion into the corresponding DBr salt, the tyramine was directly protected with Fmoc-Cl.

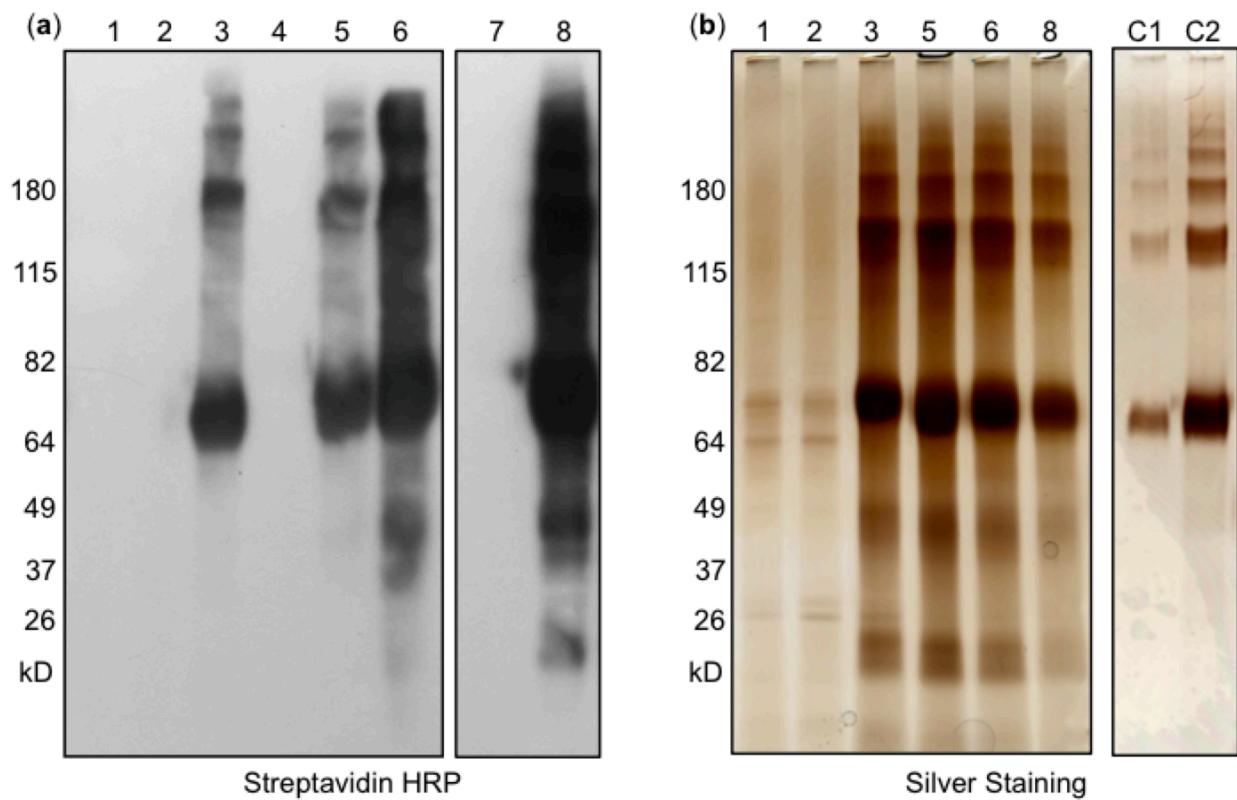


**Figure 2.7: Synthetic scheme for the synthesis of the heavy (11-d<sub>7</sub>) linker.** The protocol was first piloted for the synthesis of the light azide (11)

With the deuterated Fmoc tyramine (**20-d<sub>8</sub>**) in hand, we started with a non-deuterated pilot synthesis to generate the light diazobenzene cleavable biotin azide linker, followed by optimization of conditions for preparation of the deuterated version. Although literature precedent exists for most of the reactions employed<sup>47,50,55</sup>, we found that reported conditions for the coupling of the diazonium salt of *para*-amino benzoic acid and Fmoc tyramine (**20**, **20-d<sub>8</sub>**) do not result appreciable amounts of the desired diazobenzene. We found that particular attention had to be paid to the amount of potassium carbonate that was added to the aqueous-THF tyramine solution, before the addition of the acidic diazonium salt solution. We found that it was necessary to have at least 1.1 equivalents K<sub>2</sub>CO<sub>3</sub>, relative to the 6 M HCl in the diazonium solution, present in the tyramine solution prior to the addition of the diazonium solution. The diazobenzene was then readily Fmoc-deprotected using piperidine and coupled to the activated tetrafluorophenol ester of the picolyl azide (**23**). The acid moiety of the resulting diazobenzene (**24-d<sub>7</sub>**) was then activated by conversion into the corresponding NHS ester (**25-d<sub>7</sub>**) and coupled to the amine of a polyethylene glycol-linked biotin amine (**26**), finally yielding the heavy (**11-d<sub>7</sub>**) and light (**11**) cleavable biotin azide linkers.

**Comparison of First and Second-Generation Cleavable Linkers.** To test if incorporation of the picolyl azide disrupted the dithionite induced reductive cleavage from the streptavidin beads, experiments were carried out on alkyne-labeled Bovine Serum Albumin (BSA-alkyne). BSA contains nine cysteine residues, eight of which participate in disulfide-bonds, with one (Cys34) remaining as the free thiol. This reduced cysteine can specifically couple to propargyl maleimide<sup>47,51</sup>. The resulting BSA-alkyne was then used for the optimization of labeling and elution conditions, in place of probe-receptor complexes. Using BSA-alkyne, we investigated the efficiency of the CuAAC and reductive elution with our novel CSICT linkers (**Figure 2.8**). We directly compared the amount of first (**13**) and second (**11**) generation biotin azide that was coupled to

BSA-alkyne across various copper and ligand concentrations (**Table 2.1**). We compared the amounts of biotinylation via western blot analysis with streptavidin HRP (**Figure 2.8**). The second generation (**11**) clearly shows greater amounts of biotinylation in lanes 6 and 8 compared to the first generation cleavable biotin azide (**13**) in lanes 3 and 5. Neither of the lanes showed any background labeling on unmodified (non-alkynylated) BSA. Also, neither the first (**13**) or second (**11**) generation linker was able to label significant amounts of BSA alkyne at very low copper and ligand levels, as shown in lanes 4 and 7. The first and (**11**) second (**13**) generation labeling experiments that underwent significant biotinylation, lanes 3 and 5, and 6 and 8 respectively, were selected to test whether the presence of the pyridine moiety affected the dithionite mediated reductive elution. For this, biotinylated BSA-alkyne was isolated onto streptavidin beads, washed, and reductively eluted. The eluants were desalted and run on SDS-PAGE for analysis of total protein via silver staining. By comparing the elutions from the first (**13**) (lanes 3 and 5) and second (**11**) (lanes 6 and 8) generation we saw no significant difference in the amount of BSA released, indicating that the pyridine moiety has the ability to enhance the click reaction without disrupting the reductive elution. As hoped, these results show that incorporation of the pyridine and deuterium labels into the cleavable azide tag made no significant difference for labeling and purifying the BSA-alkyne.



**Figure 2.8: Comparative click chemistry and elution of first (13) and second (11) generation cleavable biotin azides.** The concentration of azides (13) and (11) were kept constant at 100  $\mu\text{M}$ . (a) Streptavidin HRP treatment of western blot for click-chemistry conditions. Lanes 1 and 2 contained native BSA exposed to azide (13) and (11), respectively, in the presences of; 500  $\mu\text{M}$  BTTP, 100  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O. Lanes 3-5 contained BSA-alkyne that was exposed to azide (13) with; 500  $\mu\text{M}$  BTTP, 100  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O; 50  $\mu\text{M}$  BTTP, 10  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O; 500  $\mu\text{M}$  BTTP, 250  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O, respectively. Lanes 6-8 contained BSA-alkyne that was exposed to azide (11) with; 500  $\mu\text{M}$  BTTP, 100  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O; 50  $\mu\text{M}$  BTTP, 10  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O; 500  $\mu\text{M}$  BTTP, 250  $\mu\text{M}$  CuSO<sub>4</sub> $\cdot$ 5H<sub>2</sub>O, respectively (b) Samples that revealed significant amounts of biotinylation (lanes 3, 5, 6 and 8) were selected for streptavidin bead-based enrichment. Following reductive elution samples were subjected to silver stain analysis. C1 and C2 are controls of known amount of BSA, 33 ng and 330 ng respectfully

**Table 2.1: Click-chemistry conditions for the comparative click analysis of first (13) and second (11) generation cleavable biotin azide linkers**

Sample	BSA Type	BTTP	CuSO <sub>4</sub> ·5H <sub>2</sub> O	Azide
1	BSA	500 µM	100 µM	<b>13</b>
2	BSA	500 µM	100 µM	<b>11</b>
3	BSA-alkyne	500 µM	100 µM	<b>13</b>
4	BSA-alkyne	50 µM	10 µM	<b>13</b>
5	BSA-alkyne	500 µM	250 µM	<b>13</b>
6	BSA-alkyne	500 µM	100 µM	<b>11</b>
7	BSA-alkyne	50 µM	10 µM	<b>11</b>
8	BSA-alkyne	500 µM	250 µM	<b>11</b>
C1-Control 1	BSA 33 ng	n/a	n/a	n/a
C2-Control 2	BSA 330ng	n/a	n/a	n/a

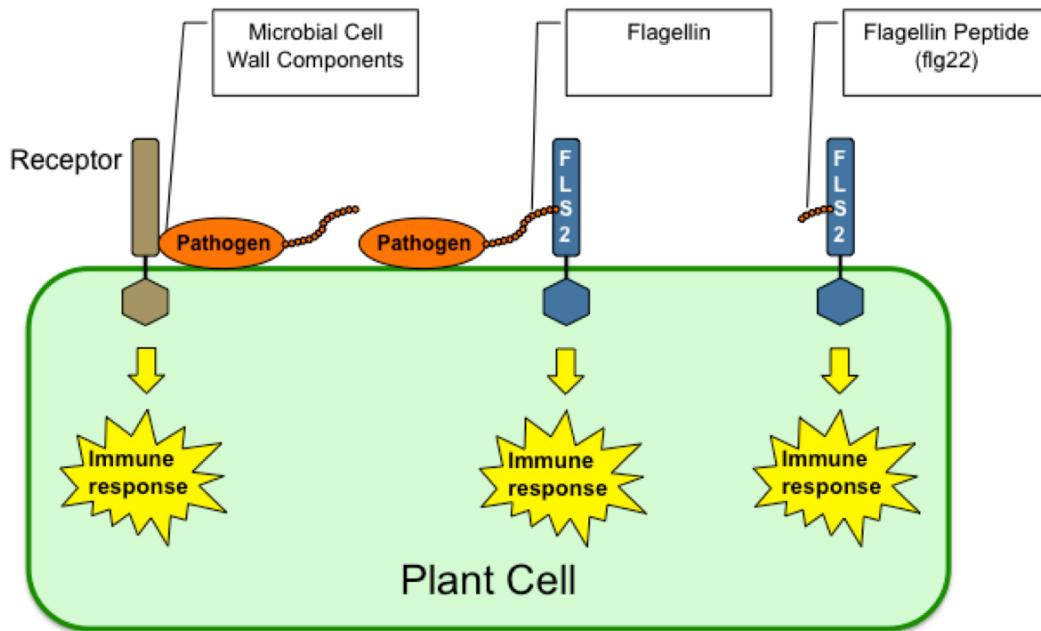
Using this BSA-alkyne CSICT method we also attempted to identify the heavy (**11-d<sub>7</sub>**) and light (**11**) twenty-two amino acid fragments that would be generated after the trypsinization of the BSA-alkyne protein gel-bands generated after the elution from the streptavidin beads. Unexpectedly, we found that the standard method of excising the BSA specific bands and subjecting them to an in-gel dithiotheritol (DTT) treatment, to reduce the disulfide bonds, followed by iodoacetamide treatment to cap the thiols in preparation for MS, resulted in loss of the cysteine-linked maleimide-tethered sidechain. In all our streptavidin-purified samples, using both the heavy and light linkers, we exclusively observed the capped carbamide version of the target amino-acid fragment and never the cleaved heavy (**11-d<sub>7</sub>**) or light (**11**) modifications. This suggested that the thioether group of the cysteine-maleimide is susceptible to the reducing capacity of DDT. By changing the reduction step from an in-gel DDT treatment, to a pre-gel in-solution tris(2-carboxyethyl)phosphine (TCEP) reduction. We were able to eliminate the undesired carbamide capped twenty-two amino acid fragment from the isotopic labeling experiments. However, the proposed heavy and light modified peptide fragments were

still not detectable and upon direct comparison of the LC-MS runs from the heavy and light experiments no other candidates were found.

In reviewing the literature, Wu *et al*, demonstrated that maleimide linkages on the Cys34 of BSA are capable of undergoing aminolysis with the N-terminal amine<sup>52</sup>. This cyclization would generate a 34 amino acid macrocycle that upon trypsinization would create a peptide fragment of 41 amino acids. Large peptide fragments of this size are challenging to identify using the standard LC-MS methods due to ionization difficulties. Searching for the corresponding large heavy and light peptides did not result in a positive identification. Given the challenges associated with BSA characterization, we decided to demonstrate the ability of heavy and light cleavable stable isotope click-chemistry tags in a different system using flagellin-derived signaling peptide flgII-28 (*vide infra*).

**A First Application: Plant-Recognized Bacterial Peptides.** Bacteria feature a number of distinct molecular signatures that are recognized by their plant and animal hosts to trigger a wide variety of specific immune responses<sup>52,53</sup>. Such bacterially-derived immune system triggers are generally referred to as Pathogen Associated Molecular Patterns (PAMPs). PAMPs can be derived from a variety of sources including the flagellin, cell wall, and nucleus of bacterial pathogens (**Figure 2.9**). The best characterized PAMPs are derived from flagellin, the appendage promoting bacterial motility, more precisely the peptide monomer forming the flagella<sup>53,54</sup>. In humans, Toll-like receptor 5 (TLR5) binds to the flgII-28 peptide in flagellin, thereby activating NF- $\kappa$ B-based immunity. In plants, the well-studied FLS2 receptor binds a different PAMP in flagellin, referred to as flg-22<sup>54,55</sup>. Recently, tomato has been found to also recognize flgII-28, via an as-yet unknown receptor we provisionally refer to as FLS3<sup>56,57</sup>. Here we present the synthesis and biological activity for a flgII-28 photoaffinity probe (flgII-28\*). The goal is to use this flgII-28\* probe with the CSICT method to first identify the FLS3 receptor and ultimately map the binding site for flgII-28. Identification of the FLS3

receptor may enable new strategies for the protection of agriculturally important crops from bacterial and other pathogens.

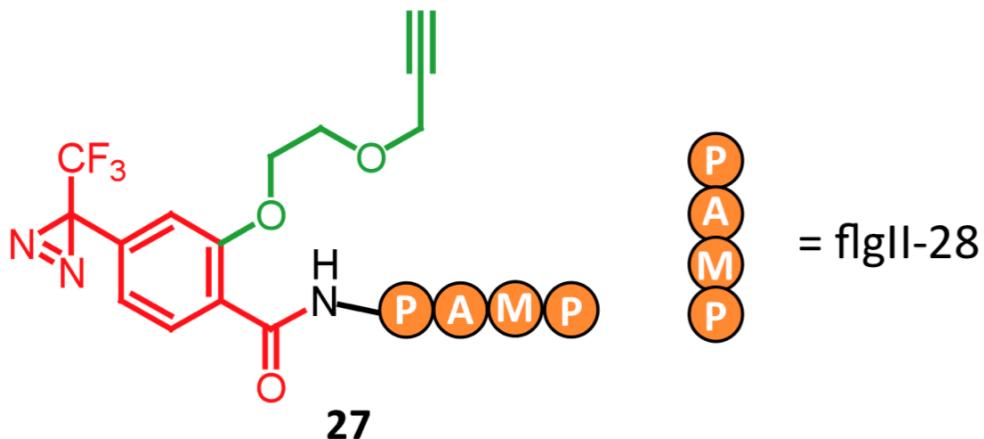


**Figure 2.9: Pathogen Associated Molecular Patterns PAMPs.** PAMPs are based on a variety of pathogen derived chemical entities that when isolated from the pathogen still elicits an immune response.

#### flgII-28 Photoaffinity Probe (flgII-28\*, 27)

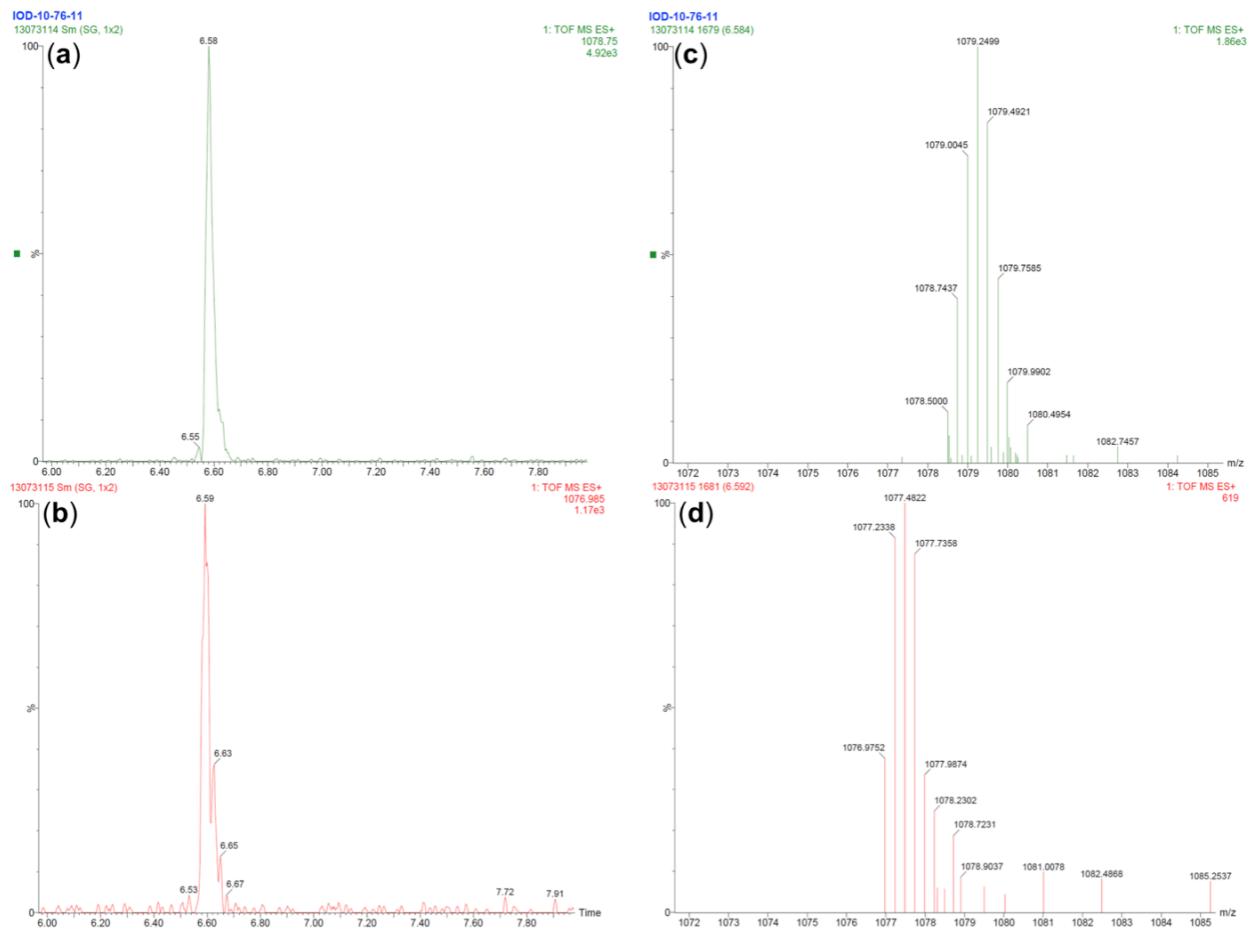
flgII-28 is a twenty eight amino acid peptide monomer repeat (ESTNILQRMRELAVQSRNDSNSSTDRA), that is found in the flagellin of bacterial pathogens. Its structure is notable for the absence of any nucleophilic lysine residues and thus has no primary amines, except for the N-terminus of the free peptide<sup>56,57</sup>. The absence of any additional nucleophilic sites enabled a very short and straight-forward synthesis of an N-terminally-linked probe version of flgII-28\* (**Figure 2.10, 27**) using the activated NHS ester of the aryl-probe (**28**) and commercially available, high-purity flgII-28. We hypothesized that by incorporating the probe to the N-terminal of the PAMP it would mimic the amide bond that would typically be found in full-length flagellin, minimizing any disruption resulting from attachment of the label. Locating the TFD and

alkyne group on the same amino acid prevents the isotopic signature from being lost when the probe-receptor complex is trypsinized for MS analysis.



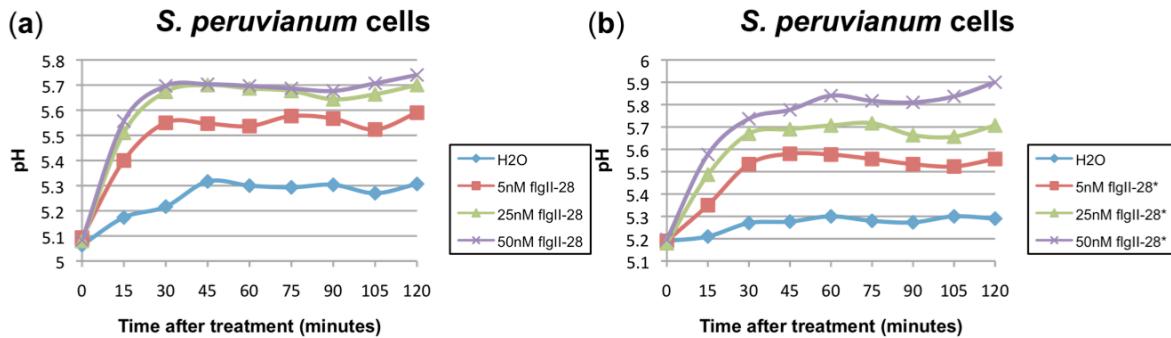
**Figure 2.10: N-terminally tagged flgII-28 probe (flgII-28\*, 27).** In green is the alkyne for CSICT coupling. In red is the Tfd group that will allow photocrosslinking to the FLS3 receptor

Featuring a terminal alkyne, the flgII-28\* probe provided an excellent candidate for a proof of concept experiment, to demonstrate peptide labeling with the heavy and light cleavable stable isotope click-chemistry azide tags. The flgII-28\* probe was clicked with the heavy and light azides, and the resulting triazole products were characterized by LC-MS and purified by preparative HPLC (**Figure 2.11**). The isobarically labeled peptide fragments were then subject to dithionite reduction conditions and analyzed by LC-MS. In both cases, the heavy and light yielded the desired aniline products with no other substantial entities being generated, as determined by LC-MS and LC-UV analysis. These results show that the reduced heavy and light peptide conjugates are reduced by dithionite and compatible with LC-MS analysis and detection (**Figure 2.11, data being recollected panels not yet present in figure.**).



**Figure 2.11: Mass spectrometric analysis of heavy/light flgII-28\* probe.** Demonstrated in two independent LC-MS runs, the isobaric heavy and light flgII-28\* probe elute at the same from the LC. In (a) the ion peak corresponding to the heavy-flgII-28\* + 4H<sup>+</sup> (1078.750 m/z) eluting at 6.58 min is shown, while in (b) light-flgII-28\* + 4H<sup>+</sup> (1076.985 m/z) can be seen to elute at the same time. Comparing the MS spectrum of each of the ion peaks in (a) and (b); heavy-flgII-28\* (c) and light-flgII-28\* (d), demonstrates a robust difference m/z mass difference, even in this 4<sup>+</sup> state. Data corresponding to the cleaved heavy/light-flgII-28\* probe is currently being reacquired.

**Biological Activity of flgII-28 Probe.** Treatment of plant cells with pathogens or isolated PAMPs causes a host of physiological immune responses that can be reliably measured and are used to indicate activity of individual PAMPs or any particular pathogen-derived materials. One of the most rapid and best well characterized plant responses used for measuring PAMP activity is the medium alkalization assay, which measures changes in the pH of plant cell culture media as a consequence of changing ion fluxes across the plasma membrane in test sample-treated plant cells<sup>55,58</sup>.



**Figure 2.12: Biological activity of flgII-28 and flgII-28 probe (flgII-28\*) in an alkalization assay with *S. peruvianum* cells.** Increasing concentration profile of (a) flgII-28 and (b) flgII-28\* on *S. peruvianum* cells shows that flgII-28\* has an similar activity profile to the native flgII-28 peptide.

Cell suspensions of *Solanum peruvianum* (a tomato relative) are capable of perceiving native flgII-28 and as result shift the pH of their growth medium within 30 min. Testing the modified flgII-28 photoaffinity probe we found that it retains most of the activity to the native peptide (Figure 2.12), indicating that incorporation of the additional N-terminal functionality does not interfere with binding to the receptor FLS3.

## Next Steps Toward Identification of FLS3

In collaboration with Martin lab, we propose to identify and characterize the PAMP binding site of the unknown flgII-28 receptor, FLS3. The identification of this receptor will provide a deeper understand of plants' innate immunity to bacterial pathogens and has the potential to provide a means to protect agriculturally important crops from bacterial and other pathogens without the use of additional agri-chemical treatments. To carry out the identification and characterization of flgII-28 FLS3 ligand-receptor pairing, we envisage utilizing the optimized photoaffinity labeling CSICT techniques that have been outlined in this Chapter.

The flgII-28\* photoaffinity probe that we have synthesized, displays a potent biological activity with almost paralleled efficacy to the unmodified flgII-28 peptide. This leads us to believe that the flgII-28\* probe is still capable of binding the FLS3 receptor. We plan to proceed to with a photoaffinity labeling stagey in unpigmented plant suspension cells, to circumvent the issues associated with UV crosslinking in the presence of chlorophyll. These plant specific photolabeling techniques have been pioneered and developed in collaboration with the Martin Lab. PAMP receptors such as characterized flg-22 receptor FLS2 are a lipophilic multi-pass transmembrane protein that can have challenging biochemical properties, making incompatible with standard AfBPP biotin-streptavidin enrichment techniques. We have demonstrated that (**Figure 2.5, 2.6**) the biotin-streptavidin can be utilized inconjucion with lipophilic membrane receptors. Based on this precedent, we believe that the heavy (**11-d<sub>7</sub>**) and light (**11**) CSICT azides will be capable efficiently coupling to the probe-receptor complex providing a means for the enrichment fgl-28\*-FLS3 complex. Once these heavy and light samples have been eluted from the beads and subjected to a trypsinization, they will analyzed by LC-MS to identify the FLS3 receptor. Finally, the ample mass difference between the deuterated and non-deuterated linkers will provide a sufficient isobaric mass difference that the specific heavy and light probe-peptide fragments will be

identified by through manually comparing the MS data, allowing the full characterization of the binding site peptide. The identification of FLS3 along with the binding site for flgII-28 binding site with our novel CSICT approach, will demonstrate the first application of a robust, generally applicable method for the deconvolution ligand receptor interactions that can coupled to any clickable ABPP or AfBPP.

## REFERENCES

- (1) Swinney, D. C.; Anthony, J. *Nat Rev Drug Discov* **2011**, *10*, 507–519.
- (2) Terstappen, G. C.; Schlüpen, C.; Raggiaschi, R.; Gaviraghi, G. *Nat Rev Drug Discov* **2007**, *6*, 891–903.
- (3) Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. *Angew. Chem. Int. Ed.* **2013**, *52*, 2744–2792.
- (4) Bantscheff, M.; Drewes, G. *Bioorganic & medicinal chemistry* **2012**, *20*, 1973–1978.
- (5) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. *Nat Chem Biol* **2013**, *9*, 232–240.
- (6) Ong, S.-E.; Schenone, M.; Margolin, A. A.; Li, X.; Do, K.; Doud, M. K.; Mani, D. R.; Kuai, L.; Wang, X.; Wood, J. L.; Tolliday, N. J.; Koehler, A. N.; Marcaurelle, L. A.; Golub, T. R.; Gould, R. J.; Schreiber, S. L.; Carr, S. A. *Proceedings of the ...* **2009**.
- (7) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998**, *391*, 806–811.
- (8) Boutros, M.; Ahringer, J. *Nat. Rev. Genet.* **2008**, *9*, 554–566.
- (9) Lounkine, E.; Keiser, M. J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J. L.; Lavan, P.; Weber, E.; Doak, A. K.; Côté, S.; Shoichet, B. K.; Urban, L. *Nature* **2012**, *486*, 361–367.
- (10) Fliri, A. F.; Loging, W. T.; Thadeio, P. F.; Volkmann, R. A. *Nat Chem Biol* **2005**, *1*, 389–397.
- (11) Domon, B.; Aebersold, R. *Nature Reviews Immunology* **2010**, *28*, 710–721.
- (12) Lapinsky, D. J. *Bioorganic & medicinal chemistry* **2012**, *20*, 6237–6247.
- (13) Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686–4687.
- (14) Li, N.; Overkleft, H. S.; Florea, B. I. *Current Opinion in Chemical Biology* **2012**, *16*, 227–233.
- (15) Nodwell, M. B.; Sieber, S. A. In *link.springer.com*; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; Vol. 324, pp. 1–41.
- (16) Dubinsky, L.; Krom, B. P.; Meijler, M. M. *Bioorganic & medicinal chemistry* **2012**, *20*, 554–570.
- (17) Brunner, J.; Senn, H.; Richards, F. M. *J. Biol. Chem.* **1980**, *255*, 3313–3318.
- (18) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.
- (19) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192–3193.
- (20) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596–2599.
- (21) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.
- (22) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proceedings of the National Academy of Sciences* **2007**, *104*, 16721–16722.
- (23) Chenoweth, K.; Chenoweth, D.; Goddard, W. A., III *Org. Biomol. Chem.* **2009**, *7*, 5255–5258.
- (24) Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 3688–3690.

- (25) Fredens, J.; Engholm-Keller, K.; Giessing, A.; Pultz, D.; Larsen, M. R.; Højrup, P.; Møller-Jensen, J.; Færgeman, N. J. *Nature Methods* **2011**, *8*, 845–847.
- (26) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
- (27) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. *J. Am. Chem. Soc.* **2009**, *131*, 4967–4975.
- (28) Frei, A. P.; Jeon, O.-Y.; Kilcher, S.; Moest, H.; Henning, L. M.; Jost, C.; Plückthun, A.; Mercer, J.; Aebersold, R.; Carreira, E. M.; Wollscheid, B. *Nature Reviews Immunology* **2012**, *30*, 997–1001.
- (29) Hosoya, T.; Hiramatsu, T.; Ikemoto, T.; Nakanishi, M.; Aoyama, H.; Hosoya, A.; Iwata, T.; Maruyama, K.; Endo, M.; Suzuki, M. *Org. Biomol. Chem.* **2004**, *2*, 637–641.
- (30) Speers, A. E.; Cravatt, B. F. *Chemistry & Biology* **2004**, *11*, 535–546.
- (31) Mayer, T. *Eur. J. Org. Chem.* **2007**.
- (32) Park, D.; O'Doherty, I.; Somvanshi, R. K.; Bethke, A.; Schroeder, F. C.; Kumar, U.; Riddle, D. L. *pnas.org*.
- (33) Sohn, C. H.; Lee, J. E.; Sweredoski, M. J.; Graham, R. L. J.; Smith, G. T.; Hess, S.; Czerwieniec, G.; Loo, J. A.; Deshaies, R. J.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2012**, *134*, 2672–2680.
- (34) Lallana, E.; Riguera, R.; Fernandez-Megia, E. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 8794–8804.
- (35) Ross, P. L. *Molecular & Cellular Proteomics* **2004**, *3*, 1154–1169.
- (36) Gomes, I.; Gupta, A.; Devi, L. A. *G-Protein-Coupled Heteromers: Regulation in Disease*; 1st ed. Elsevier Inc., 2013; Vol. 521, pp. 219–238.
- (37) Anna E Speers, B. F. C. *J. Am. Chem. Soc.* **2005**, *127*, 10018.
- (38) Park, K. D.; Liu, R.; Kohn, H. *Chemistry & Biology* **2009**, *16*, 763–772.
- (39) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. *J. Am. Chem. Soc.* **2010**, *132*, 18351–18360.
- (40) Nessen, M. A.; Kramer, G.; Back, J.; Baskin, J. M.; Smeenk, L. E. J.; de Koning, L. J.; van Maarseveen, J. H.; de Jong, L.; Bertozzi, C. R.; Hiemstra, H.; de Koster, C. G. *J Proteome Res* **2009**, *8*, 3702–3711.
- (41) Yang, Y.-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641.
- (42) Yang, Y.; Verhelst, S. H. L. *Chem. Commun.* **2013**, *49*, 5366.
- (43) Kim, H. Y. H.; Tallman, K. A.; Liebler, D. C.; Porter, N. A. *Molecular & Cellular Proteomics* **2009**, *8*, 2080–2089.
- (44) Hulce, J. J.; Cognetta, A. B.; Niphakis, M. J.; Tully, S. E.; Cravatt, B. F. *Nature Methods* **2013**, *10*, 259–264.
- (45) Couvineau, A.; Tan, Y.-V.; Ceraudo, E.; Laburthe, M. In *Methods in Enzymology*; Methods in Enzymology; Elsevier, 2013; Vol. 520, pp. 219–237.
- (46) Grunbeck, A.; Huber, T.; Sakmar, T. P. In *Methods in Enzymology*; Methods in Enzymology; Elsevier, 2013; Vol. 520, pp. 307–322.
- (47) Yang, Y. Y.; Grammel, M.; Raghavan, A. S.; Charron, G.; Hang, H. C. *Chemistry & Biology* **2010**, *17*, 1212–1222.
- (48) Palaniappan, K. K.; Pitcher, A. A.; Smart, B. P.; Spicciarich, D. R.; Iavarone, A. T.; Bertozzi, C. R. *ACS Chem. Biol.* **2011**, *6*, 829–836.
- (49) Uttamapinant, C.; Tangpeerachaikul, A. *Angewandte Chemie*.
- (50) Watson, D. G.; Baines, R. A.; Midgley, J. M.; Bacon, J. P. *Methods Mol. Biol.* **1997**, *72*, 225–237.

- (51) Dirks, A. {. J.; Cornelissen, J. J. L. M.; Nolte, R. J. M.
- (52) Wu, C.-W.; Yarbrough, L. R.; Wu, F. Y. *Biochemistry* **1976**, *15*, 2863–2868.
- (53) Ronald, P. C.; Beutler, B. *Science* **2010**.
- (54) Gómez-Gómez, L.; Boller, T. *Trends Plant Sci.* **2002**, *7*, 251–256.
- (55) Gómez-Gómez, L.; Boller, T. *Molecular Cell* **2000**, *5*, 1003–1011.
- (56) Cai, R.; Lewis, J.; Yan, S.; Liu, H.; Clarke, C. R.; Campanile, F.; Almeida, N. F.; Studholme, D. J.; Lindeberg, M.; Schneider, D.; Zaccardelli, M.; Setubal, J. C.; Morales-Lizcano, N. P.; Bernal, A.; Coaker, G.; Baker, C.; Bender, C. L.; Leman, S.; Vinatzer, B. A. *PLoS Pathog.* **2011**, *7*, e1002130.
- (57) Clarke, C. R.; Chinchilla, D.; Hind, S. R.; Taguchi, F.; Miki, R.; Ichinose, Y.; Martin, G. B.; Leman, S.; Felix, G.; Vinatzer, B. A. *New Phytol* **2013**, n/a–n/a.
- (58) Atkinson, M. M.; Huang, J. S.; Knopp, J. A. *Plant Physiol.* **1985**, *79*, 843–847.

## CHAPTER 3

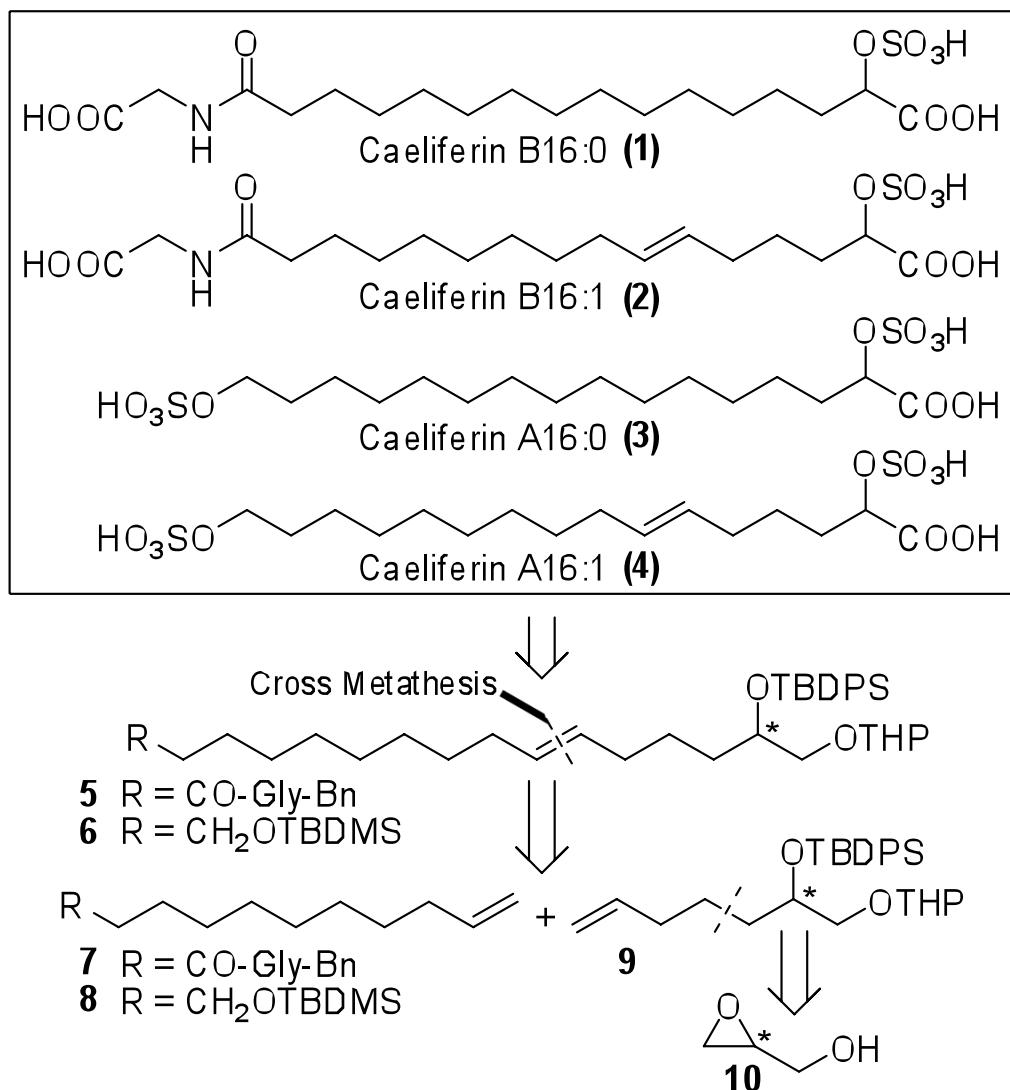
### SYNTHESIS OF THE CAELIFERINS, ELICTORS OF PLANT IMMUNE RESPONSES: ACCESSING LIPOPHILIC NATURAL PRODUCTS VIA CROSS METATHESIS

**Introduction:** The caeliferins (**1-4**) are a family of insect-derived small molecule signals that play an important role in plant defenses. Originally isolated from grasshopper saliva, these molecules elicit immune responses in corn and Arabidopsis that attract natural predators of the feeding herbivores<sup>1,2</sup>.

Despite their interesting role in insect-plant interactions, a lack of synthetic caeliferin samples has prevented further study of their mode of action, including the identification of specific receptors in plants that activate induced immune responses. We envisioned that a cross metathesis (CM)-based approach would provide us with the flexibility to create all known caeliferins **1-4** from chiral synthon **9** via variation of its long-chained metathesis partners, **7** and **8** (Figure 3.1). Using a large excess of inexpensive **7** or **8** in the metathesis reaction should allow suppression of competing dimerization of **9**.

The enantiomers of **9** were obtained from THP-protected glycidol via a Cu(I) catalyzed Grignard reaction with 4-bromo-1-butene (Figure 3.2)<sup>3,4</sup>. Alkenes **7** and **8** were each prepared in one step from commercially available precursors<sup>5</sup>.

Next we investigated conditions for CM of **9** with **7** or **8**<sup>6-9</sup>. Strikingly, CM on substrates **9** and **8** using Grubbs 1<sup>st</sup> generation (G-I), Grubbs 2<sup>nd</sup> generation (G-II), or Hoveyda-Grubbs 2<sup>nd</sup> generation (HG-II) catalysts resulted in complex product mixtures, containing only 20-50% of desired **6** in addition to significant quantities of side products.



**Figure 3.1: Caeliferin Structures and Retrosynthesis**

ESI<sup>+</sup>-MS analysis of CM reactions revealed CH<sub>2</sub>-insertions and deletions for substrate **9** as well as CH<sub>2</sub>-deletions for product **6** (Figures 3.3, C.4). Similar results were obtained in CM of **9** with **7** (Figure C.5). Using the more active G-II or HG-II catalysts, significant amounts of chain shortened starting materials and products were already observed after 2 h, at which time less than 40% of starting material **9** had been consumed.

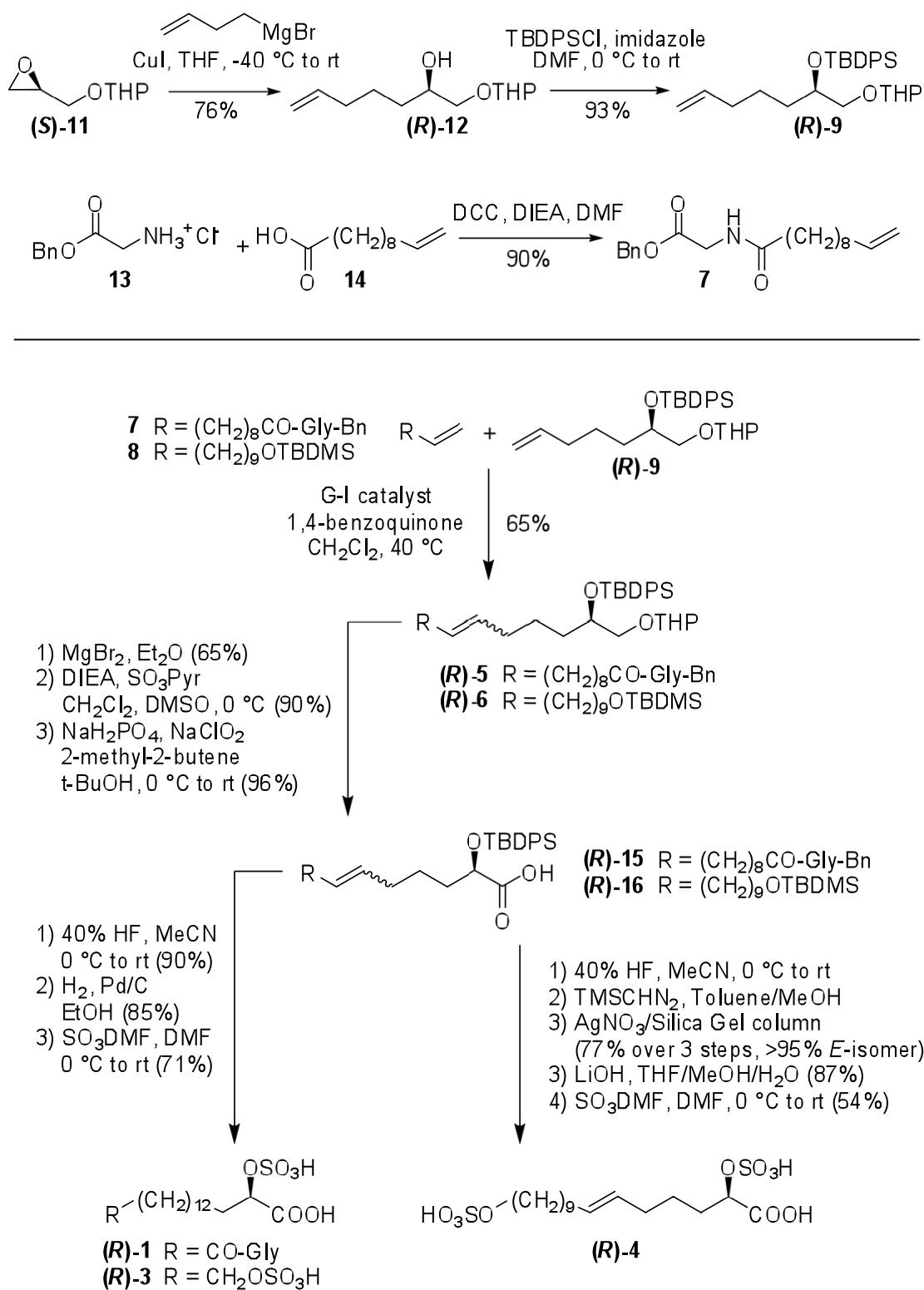
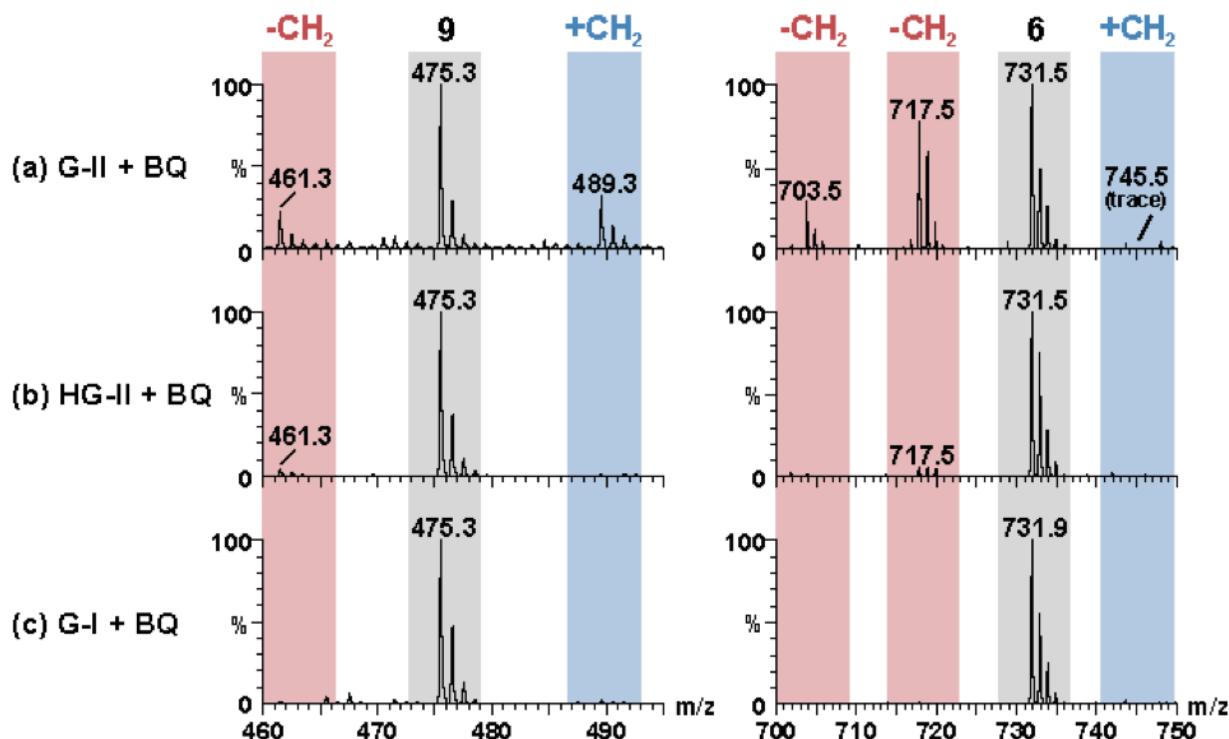


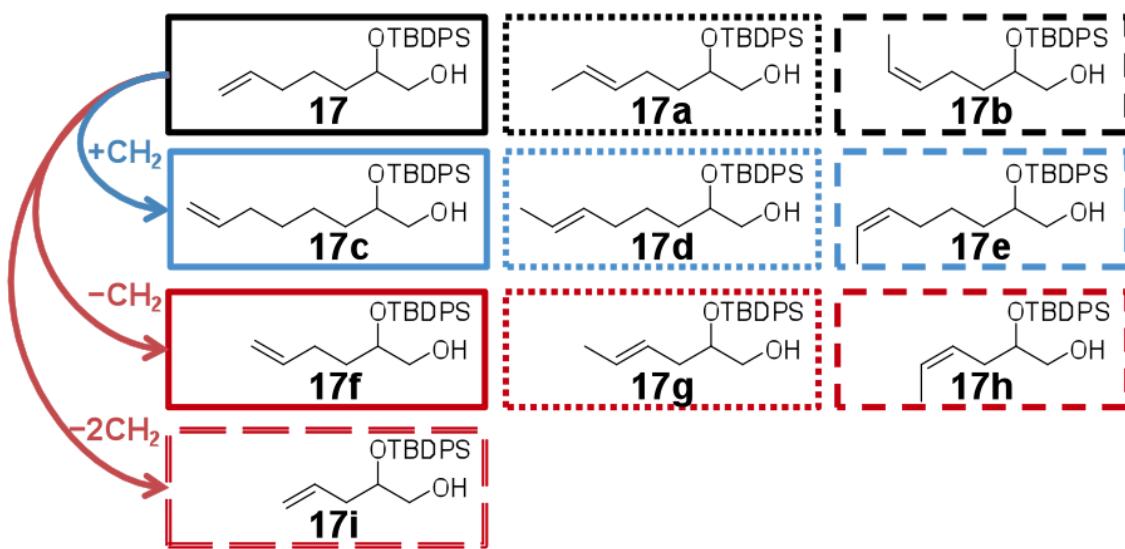
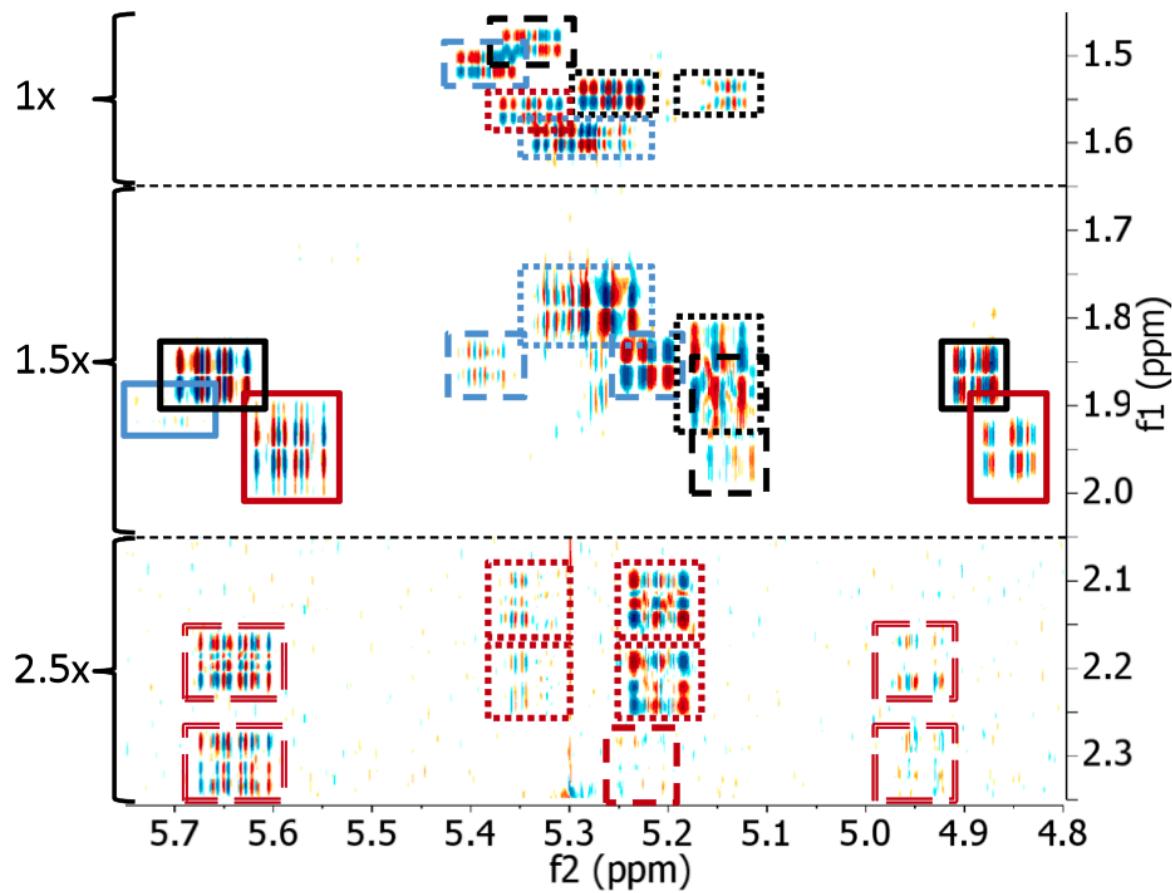
Figure 3.2: Synthesis of Caeliferins 1, 3, and 4

Previous studies have shown that isomerization as well as CH<sub>2</sub>-insertion and deletion during metathesis reactions are likely the result of ruthenium hydride formation<sup>10-13</sup>, which can be suppressed by addition of hydride scavengers, e.g. 1,4-benzoquinone (BQ)<sup>14</sup>. We found that even with the addition of BQ, the use of either G-II or HG-II resulted in formation of significant quantities of side products (**Figures C.4, C.5**).



**Figure 3.3: ESI<sup>+</sup>-MS analysis of CM reaction mixtures.** Showing [M+Na<sup>+</sup>] for starting material 9 and product 6. Ion signals corresponding to chain shortened and chain extended homologues are shown in red and blue, respectively. Conditions: (a) G-II (0.05 equiv), BQ (0.1 equiv), 40 °C, 20 h; (b) HG-II (0.05 equiv), BQ (0.1 equiv), 40 °C, 20 h; (c) G-I (0.05 equiv), BQ (0.1 equiv), 40 °C, 20 h

Our ESI<sup>+</sup>-MS-based analyses revealed significant homologation; however, the extent of CM-induced product and starting material isomerization remained unclear. Building on recent experience using 2D NMR spectroscopy for the analysis of mixtures<sup>15,16</sup>, we used high-resolution dqcOSY spectra to further characterize CM reaction outcomes.

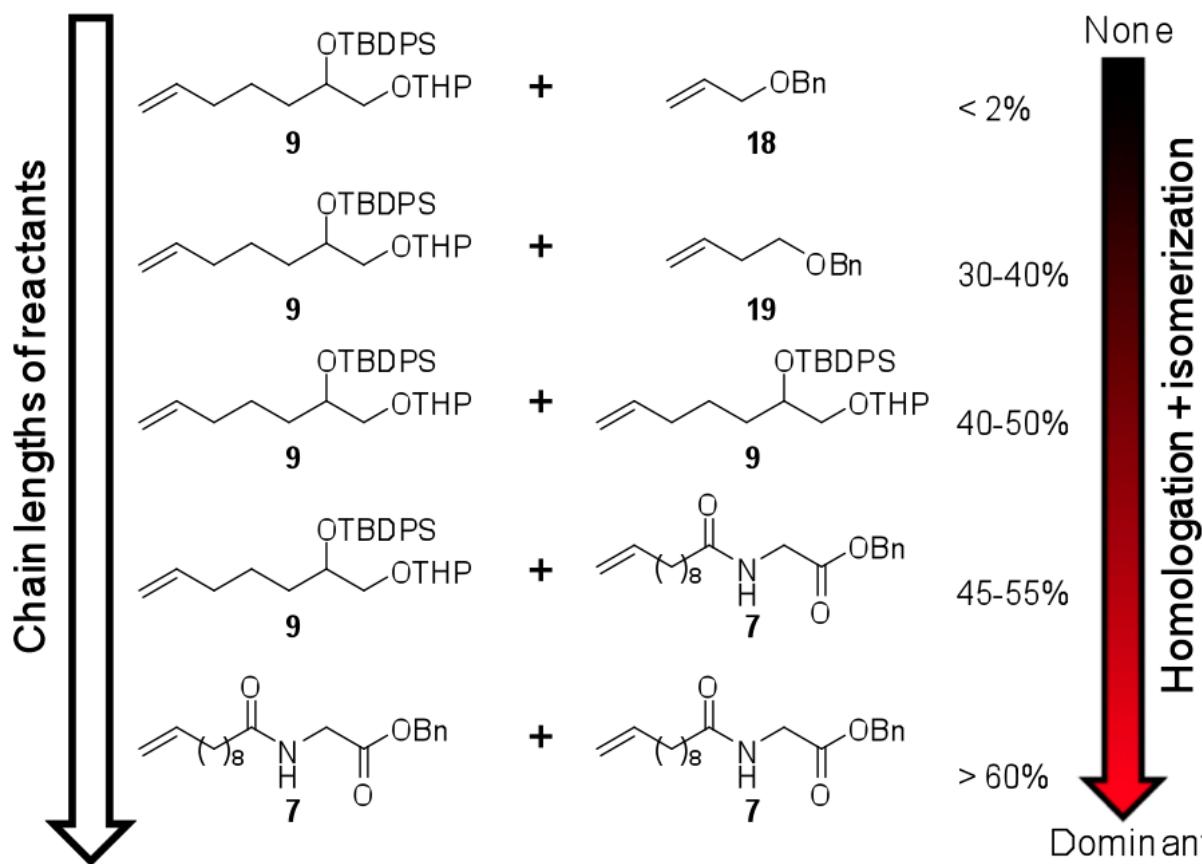


**Figure 3.4: dqcOSY analysis of isomerization.** Section of dqcOSY spectrum of the mixture of **17** and its isomers and homologues, as derived from the corresponding mixture of isomers and homologues of **9** isolated from CM of **8** with **9** using G-II and BQ (600 MHz,  $\text{CDCl}_3$ , see **Figure C.1** for full spectrum). Intensity of parts of the spectrum was scaled (1.5x, 2.5x)

To simplify NMR-spectroscopic analysis, we isolated the mixture of residual **9**, its isomers and homologues from G-II-catalyzed CM of **8** and **9** and subsequently removed the stereogenic THP group, producing a corresponding mixture of **17** and derivatives (**Figure 3.4**). dqcCOSY spectra<sup>15,16</sup>, of this mixture enabled identification of all significant isomerization and homologation products (**17a-i**), revealing extensive double-bond isomerization and confirming CH<sub>2</sub>-deletions in **9** during CM, which explains the formation of nor-homologues of metathesis product **6**. Analysis of the dqcCOSY spectra further revealed large quantities of the chain-extended compounds **17d** and **17e**. However, only trace amounts of the chain-extended terminal alkene **17c** were found. This is consistent with the hypothesis that CH<sub>2</sub>-extended variants of **9** form mostly via ethyldene-transfer during CM of isomerized starting materials, without significant contribution from isomerized product **6** (isomerization and homologation pathways are shown in **Figure C.3**). The low abundance of the CH<sub>2</sub>-extended terminal alkene **17c** also explains that only trace amounts of CH<sub>2</sub>-insertion products of **6** were observed (m/z 745.5 in **Figure 3.3**), as their formation in the absence of chain-extended terminal alkenes would require multiple isomerization and metathesis steps (see **Figure C.3**). In contrast to G-II and HG-II, use of G-I catalyst in the presence of BQ did not result in homologation (**Figure 3.3**) or isomerization (as confirmed by dqcCOSY), even when using as much as 0.1 equivalents of catalyst and reaction times of up to 48 h. Using G-I catalyst, **6** was obtained in 80% yield (E/Z = 4:1), based on consumed **9**.

Given that both G-II and HG-II have been employed successfully in many CM reactions<sup>17-19</sup>, we investigated whether the degree of homologation and isomerization depends on specific properties of the reactants **7-9** (**Figure 3.5**). Using G-II catalyst without BQ, we found that CM of **9** with allyl benzyl ether proceeded cleanly without formation of isomerized products or substrates, whereas CM of **9** and homoallyl benzyl ether and homodimerization of **9** consistently yielded significant amounts of isomerized and homologated products (**Figure C.7**). Even larger quantities of isomerized and

homologized products were obtained in G-II-catalyzed homodimerization reactions of **7** and **8** (**Figures C.8, C.9**), similar to amounts of side products produced in G-II-catalyzed CM reactions of **9** with **7** or **8** (**Figures C.4, C.5**). Thus it appears that the severity of isomerization in our metathesis reactions generally correlates with the chain-lengths of the reaction partners and products (**Figures 3.5, C.4, C.5, C.7-C.9**).

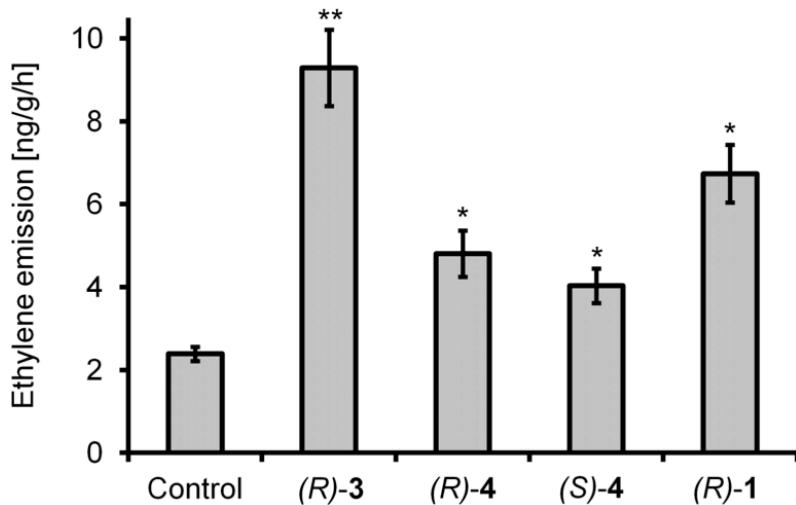


**Figure 3.5: Isomerization trends with substrate chain length.** Amounts of detected homologation and isomerization products correlate with chain lengths of CM reactants. Percentages refer to amounts of homologated products formed in G-II-catalyzed CM without BQ, as determined by ESI<sup>+</sup>-MS (see **Appendix C** for conditions). Additional side product formation occurs due to product isomerization.

That long-chain, lipophilic substrates are particularly susceptible to homologation and isomerization during CM is also suggested by recent acyclic diene metathesis (ADMET) studies<sup>12,13,20</sup>. Avoiding homologation and isomerization is of importance for both

ADMET and natural product synthesis, as physical or biological properties may be affected greatly by small amounts of structurally similar impurities. Our results show that among tested catalysts, G-I with added BQ is most suitable for CM involving long-chain substrates.

To complete synthesis of the caeliferins, CM products **5** and **6** were deprotected<sup>21</sup> and converted into acids **15** and **16** via sequential Swern and Pinnick oxidation<sup>22</sup>. Desilylation followed by hydrogenation and sulfation produced **1** and **3**. For the preparation of pure (*E*)-**4**, desilylated **16** was methylated and freed from contaminating (*Z*)-isomer using AgNO<sub>3</sub>-impregnated silica gel<sup>23</sup>. Samples of synthetic caeliferins were tested in *Arabidopsis*<sup>1,2</sup> for their activity to elicit immune responses and thus mimic the effects of a feeding herbivore's saliva.



**Figure 3.6: Bioassays of Caeliferins.** Ethylene emission of *Arabidopsis* seedlings after treatment with synthetic caeliferin solutions; errors bars, s.d.; two-tailed Student's t-test; \*P<0.05, \*\*P<0.005. Caeliferins were applied to mechanically scratched seedlings at a concentration of 22 nmol  $\mu\text{l}^{-1}$ , with control seedlings undergoing only mock buffer addition to damaged tissue.

When delivered at concentrations corresponding to those found in grasshopper saliva (22 nmol  $\mu\text{l}^{-1}$ )<sup>1,2</sup>, the tested caeliferins strongly induced ethylene production (**Figure**

**3.6**), revealing differences in the relative potency of different compounds<sup>1,2</sup>. Ethylene-inducing activity of synthetic caeliferin A16:0 (**3**) is similar to that previously reported.

In summary, we report the first enantioselective syntheses of *(R)*-**1**, *(R)*-**3**, *(R)*-**4** and *(S)*-**4**, using CM conditions suitable for long-chained substrates. Our CM-based caeliferin synthesis is flexible and provides access to related compounds, which will enable studies aimed at identifying molecular targets of these elicitors of plant immune responses. Better understanding of the caeliferins' biological mode of action may facilitate development of sustainable pesticides that take advantage of natural plant defense responses.

## REFERENCES

- (1) Alborn, H. T.; Hansen, T. V.; Jones, T. H.; Bennett, D. C.; Tumlinson, J. H.; Schmelz, E. A.; Teal, P. E. A. *Proceedings of the National Academy of Sciences* **2007**, *104*, 12976–12981.
- (2) Schmelz, E. A.; Engelberth, J.; Alborn, H. T.; Tumlinson, J. H.; Teal, P. E. A. *Proceedings of the National Academy of Sciences* **2009**, *106*, 653–657.
- (3) Cahiez, G.; Chaboche, C.; Jézéquel, M. *Tetrahedron* **2000**, *56*, 2733–2737.
- (4) Díez, E.; Dixon, D. J.; Ley, S. V.; Polara, A.; Rodríguez, F. *HCA* **2003**, *86*, 3717–3729.
- (5) Xu, P.; Lin, W.; Zou, X. *Synthesis* **2002**, *2002*, 1017–1026.
- (6) Grubbs, R. H.; Burk, P. L.; Carr, D. D. *J. Am. Chem. Soc.* **1975**, *97*, 3265–3267.
- (7) Connon, S. J.; Blechert, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 1900–1923.
- (8) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 11360–11370.
- (9) Hoveyda, A. H.; Zhugralin, A. R. *Nature* **2007**, *450*, 243–251.
- (10) Schmidt, B. *Eur. J. Org. Chem.* **2004**, *2004*, 1865–1880.
- (11) Courchay, F. C.; Sworen, J. C.; Ghiviriga, I.; Abboud, K. A.; Wagener, K. B. *Organometallics* **2006**, *25*, 6074–6086.
- (12) Fokou, P. A.; Meier, M. A. R. *J. Am. Chem. Soc.* **2009**, *131*, 1664–1665.
- (13) Mutlu, H.; de Espinosa, L. M.; Meier, M. A. R. *Chem Soc Rev* **2011**, *40*, 1404–1445.
- (14) Hong, S. H.; Sanders, D. P.; Lee, C. W.; Grubbs, R. H. *J. Am. Chem. Soc.* **2005**, *127*, 17160–17161.
- (15) Taggi, A. E.; Meinwald, J.; Schroeder, F. C. *J. Am. Chem. Soc.* **2004**, *126*, 10364–10369.
- (16) Pungaliya, C.; Srinivasan, J.; Fox, B. W.; Malik, R. U.; Ludewig, A. H.; Sternberg, P. W.; Schroeder, F. C. *Proceedings of the National Academy of Sciences* **2009**, *106*, 7708–7713.
- (17) Canova, S.; Bellosta, V.; Cossy, J. *Synlett* **2004**, 1811–1813.
- (18) Rai, A. N.; Basu, A. *Org. Lett.* **2004**, *6*, 2861–2863.
- (19) Sheddan, N. A.; Mulzer, J. *Org. Lett.* **2006**, *8*, 3101–3104.
- (20) Qin, H.; Chakulski, B. J.; Rousseau, I. A.; Chen, J.; Xie, X.-Q.; Mather, P. T.; Constable, G. S.; Coughlin, E. B. *Macromolecules* **2004**, *37*, 5239–5249.
- (21) Kim, S.; Park, J. H. *Tetrahedron Letters* **1987**, *28*, 439–440.
- (22) Smith, A. B. *Proceedings of the National Academy of Sciences* **2004**, *101*, 12042–12047.
- (23) Williams, C. M.; Mander, L. N. *Tetrahedron* **2001**, *57*, 425–447.

## APPENDIX A

### INTERACTION OF STRUCTURE-SPECIFIC AND PROMISCUOUS G-PROTEIN-COUPLED RECEPTORS MEDIATES SMALL MOLECULE SIGNALLING IN *C. ELEGANS*

**Antibodies and Reagents.** Monoclonal and polyclonal antibodies against HA- and cMyc- were procured from Sigma-Aldrich, Inc., St. Louis, MO. Somatostatin-14 was obtained from Bachem, Torrance, CA. Fluorescein and Cy3-conjugated goat-anti-mouse and goat-anti-rabbit secondary antibodies were purchased from Jackson Immuno Research ON. Protein A/G-Agarose beads were procured from Calbiochem, EMD Biosciences, Darmstadt, Germany. Reagents for electrophoresis were purchased from BIO-RAD Laboratories Mississauga ON, Canada and EMD Millipore, Darmstadt, Germany. Reagents for cell culture were purchased from GIBCO, Invitrogen, Burlington, ON, Canada.

**Immunohistochemistry and Microscopy.** The antibody staining was performed as previously described<sup>1</sup> by using anti-cMyc antibody (Abcam, ab39688) and FITC labeled goat polyclonal secondary antibody (Abcam, ab6717) to detect cMyc-DAF-37. A Zeiss axioscope equipped with a QImaging camera (RETIGA 2000R) was used for differential interference contrast (DIC) microscopy, dsRed and GFP expression analysis. For Dil staining, cultures were synchronized by hatching purified eggs into M9 Buffer, grown on NG agar plates until the L2 stage, washed in M9 buffer and Dil stained as described<sup>2</sup>.

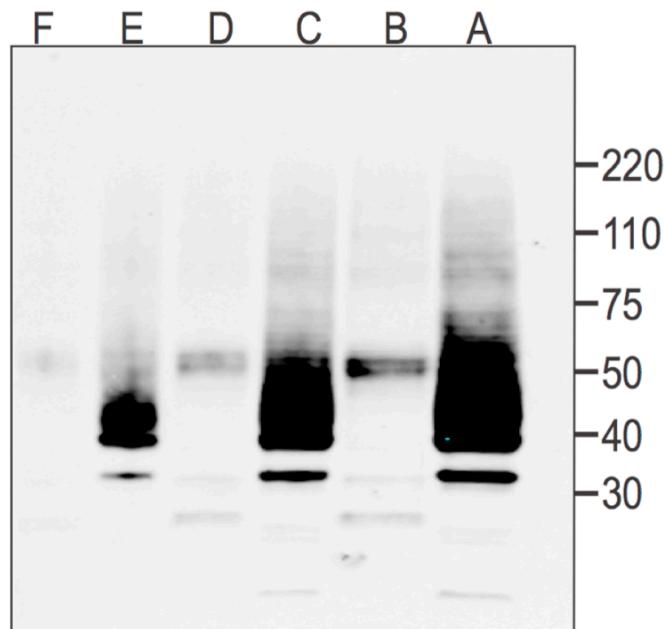
1 MDVIGNITDL SPTVSGIPDE CGLEPHDFLE VKFFLISVVG TLIGLFGLFG  
51 NATTALILTR PSMRNPNNLF LTALAVFDSC LLITAFFIYA MEYIIETYAA  
101 FDLYVAWLTY LRFAFALSHI SQTGSVYITV AVTIERYLAV CHPKSSKNMC  
151 GPGGAAWTIL GVTTFAVVFN CTKFFELOVT VNPSCPDGSN WQSYILLPSA  
201 MASNPIYQQV YSLWVTNFVM VFFPFLTLLL FNAIIAYTIR QSLEKYDFHN  
251 QKISSRNELK EKSREATLVL VIIVFIFLGC NFWGFVLTLL ERIMGOETLM  
301 VEHHIFYTFS REAINFLAII NSSINFVIYP LFGKDFRKEL VVYGCIRG  
351 ISLRLPVQDK FVIWRHWKRT KSRISMNTTN RTRHKISLPQ TLVEHANLER  
401 LEETRFLAHH EDGVQTQVSP IHALRNGSTP KIDTLQDLTS NGRPCKTSII  
451 DDNGTVVCTV TEFP\*

**Figure A.1: The predicted amino acid sequence of DAF-37.** The predicted transmembrane domain is underlined. The mos1 insertion is at T267mos1

**Cell Culture and Transfections.** The complementary DNA for *daf-37* and *daf-38* was isolated using 5' primers containing c-Myc or HA sequences at the 5' end. cMyc-DAF-37 (hygromycin resistance) and HA-DAF-38 (neomycin resistance) were constructed using the pCDNA3.1 vector as previously described<sup>3,4</sup>. HA-Rhodopsin complementary DNA was generated using a 5'-primer containing a HA sequence and the wild-type rhodopsin plasmid<sup>5</sup>, before being constructed into a pCDNA3.1 vector<sup>6</sup>. Stable transfections of HEK293 cells expressing cMyc-DAF-37 were prepared by Lipofectamine transfection reagent as previously described (Invitrogen). After selection, clones were maintained in Dulbecco's MEM supplemented with 10% fetal bovine serum (FBS) and 400 µg/ml hygromycin. HEK-293 cells expressing cMyc-DAF-37 were transiently transfected with HA-DAF-38 cDNA (0.2 µg/ml) for 36 h in DMEM using Lipofectamine transfection reagent before use in experiments.

### **Comparison of DAF-37 and Rhodopsin Expression in Transfected HEK293T Cells.**

In a single well of a six well plate, cells were transfected with either *daf-37* or rhodopsin, using previously described methods<sup>7</sup>. The cells were then harvested and suspended in 250 µl of 4% SDS, sonicated and spun down at 20,000g. The resulting supernatant was measured for total protein levels (Invitrogen, Quan-iT), before being prepped for SDS-PAGE, with the addition of sample buffer (4X) and reducing reagent (10X) (Invitrogen, NuPAGE). The samples loaded and run on 4-12% bis-tris gel (Invitrogen, NuPAGE) at 200V for 1h. The gel was then blotted on to a PVDF membrane (Millipore, Immobilon-FL) and blocked with a 5% milk solution in TBS-T. The blot was then treated with an anti-1D4 antibody ([www.flintbox.com](http://www.flintbox.com), 1:4,000, 5% milk in TBS-T), washed and treated with Goat anti-Mouse IRDye® 800CW (Li-Cor, 1:10,000, 5% milk in TBS-T). The blot was then washed and detected by infrared imaging using an Odyssey imager (Li-Cor).



**Figure A.2: Western blot comparing abundance of DAF-37 and rhodopsin expression in transfected HEK293T cells.** Lanes A,C,E contain lysate from rhodopsin expressing cells, with lanes B,D,F originating from *daf-37:1D4* transfected cells. The total protein content for lanes A and B is 48 µg, 24 µg for C and D and 8 µg for E and F

**Photocrosslinking of ascr#2 Probe.**  $1.3 \times 10^6$  HEK293T cells were split into a 10 cm plate with Dulbecco's Modification of Eagle's Medium (10% fetal bovine serum, 1% Penicillin/Streptomycin). The next morning they were transected using standard protocols<sup>7</sup>. Two days after transfection, the growth media was aspirated and the cells were washed with PBS buffer. After washing, PBS (4.5 ml) was pipetted onto the cells, and subsequently ascr#2 probe (17  $\mu$ L, 6.5 mM in DMSO) was added with gentle mixing. The cells were incubated with the probe at 37 °C for 15 min and then cooled to 4 °C. Next, for photocrosslinking, the cells were held for 40 min at 4 °C under three General Electric F40BLB 40 W bulbs (4 cm working distance). The media was then removed and the cells were harvested using PBS supplemented with EDTA-free Roche complete mini protease cocktail inhibitor. Cells were pelleted at 4 °C under 1,000 g for 3 min and the supernatant was removed. Fixing solution (3.7% PFA in PBS, 500  $\mu$ L) was added and the cells were mixed by pipetting before being placed on a rotary shaker for 10 min. Cells were pelleted under the above conditions and the fixing media was decanted. Next, the cells were washed with PBS (three times 5 min on a rotary shaker) followed by treatment with blocking solution for 30 min (10% FBS, 50 mg/ml sucrose, 20 mg/ml BSA, 500  $\mu$ L). The cells were washed again with PBS (three times 5 min). Next, cells were subjected to click chemistry conditions <sup>8</sup> with freshly prepared reaction cocktail (PBS (236.5  $\mu$ L), azido-biotin (azide-PEG4-biotin conjugate [http://www.clickchemistrytools.com], 1  $\mu$ L, 5 mM stock in dimethylformamide), tris(2-carboxyethyl)phosphine (TCEP, 5  $\mu$ L, 50 mM fresh stock in water), tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, 2.5  $\mu$ L, 10 mM stock in DMSO), CuSO<sub>4</sub>·5H<sub>2</sub>O (5  $\mu$ L, 50 mM fresh stock in water) and placed on a rotary shaker for 1 h. The cells were washed (PBS with 500  $\mu$ M EDTA, three times 5 min followed by PBS three times 5 min) and resuspended in antibody binding buffer (PBS with 0.1% BSA) in preparation for the ALPHAscreen assay.

**ALPHAscreen Assay<sup>9</sup>.** Co-localization of the biotinylated ascr#2 probe and 1D4-tagged GPCR (DAF-37) was measured using the PerkinElmer ALPHAscreen Mouse IgG detection kit #6760606C. Mouse IgG acceptor beads (1.5 µg) were incubated for 30 min on ice with 1D4 antibody (3 µg of antibody, <http://ubc.flintbox.com>) in antibody binding buffer, followed by washing (3 x 20,000g for 5 min, PBS with 0.1% BSA). Next, 1.5 µg of 1D4 antibody-loaded acceptor beads and 1.5 µg of donor beads were added to the cells such that the final volume was 120 µL. This mixture was then incubated for 20 min at rt and loaded into three wells of a 384 OPTIplate (PerkinElmer) and assayed in a BIOTEK Synergy H2 plate reader. Samples were first excited by light from a tungsten bulb with a 680/30 nm filter and luminescence was measured directly afterwards with a 570/100 nm filter

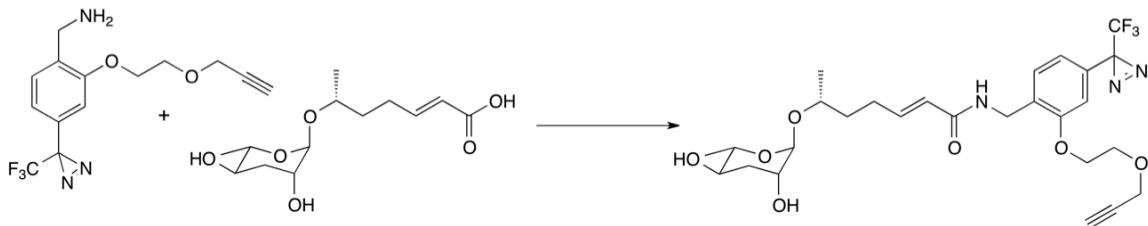
## Instrumentation and General Procedures

NMR spectra were recorded on Varian INOVA 600 (600 MHz), spectrometers in Cornell University's NMR facility.  $^1\text{H}$  NMR chemical shifts are reported in ppm ( $\delta$ ) relative to residual solvent peaks ( $\delta$  2.05 ppm for acetone-d<sub>6</sub>). NMR-spectroscopic data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), and integration. Thin-layer chromatography (TLC) was performed using J. T. Baker Silica Gel IB2-F. Flash chromatography was performed using Teledyne Isco CombiFlash systems and Teledyne Isco RediSep Rf silica columns. Unless stated otherwise, reagents were purchased from Sigma-Aldrich and used without further purification. *N,N*-dimethylformamide (DMF), dichloromethane (DCM), dimethyl sulfoxide (DMSO), and tetrahydrofuran (THF) were dried over 4 Å molecular sieves prior to use.

**Ascarosides.** Ascarosides were synthesized as described<sup>10,11</sup>. Ascarosides are named using their four letter SMIDs (Small Molecule IDentifiers, see [www.smid-db.org](http://www.smid-db.org)).

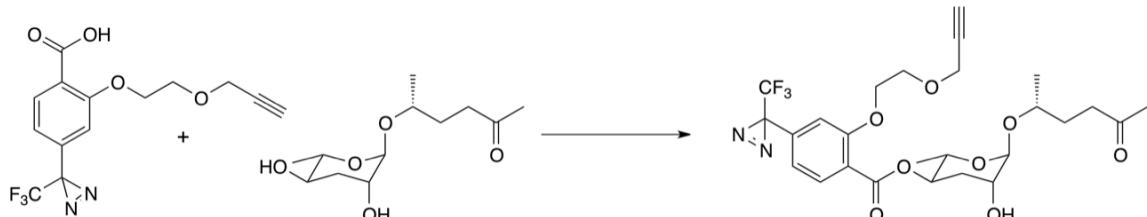
## Synthesis of Ascaroside Probes

**N-(6'R-[3''R,5''R-dihydroxy-6''S-methyl-(2H)-tetrahydropyran-2-yloxy]-2'Eheptenoyl)-methyl-2-[2-(Prop-2-ynyl)ethoxy]-4-(3-trifluoromethyl-3H-diazirin-3-yl)-benzene (1)**



To a stirred solution of DMF (400  $\mu$ L) and ascr#7<sup>11</sup> (12 mg, 43.80  $\mu$ mol) at 0°C under argon, a solution of the benzyl amine probe<sup>12,13</sup> in DMF (72  $\mu$ L, 700 mM, **15**) was added. This was followed by a 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide solution in DCM (230  $\mu$ L, 209 mM), 4-dimethylaminopyridine in DMF (43  $\mu$ L, 112 mM), DIEA (8  $\mu$ L, 45.99  $\mu$ mol). This was allowed to warm up to room temp over and stir for 16 h before being concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:10) yielding light yellow (22 mg, 39.42  $\mu$  mol, 90% yield, **1**) **1H NMR (600 MHz, CD<sub>3</sub>OD)**:  $\delta$  7.28 (d,  $J$  = 8.0 Hz, 1H, 6-H benzylamide), 6.84–6.82 (m, 1H, 5-H benzylamide), 6.80 (dt,  $J$  = 15.4, 6.8 Hz, 1H,  $\beta$ -of  $\alpha\beta$ ), 6.74 (br. s, 1H, 3H-benzylamide), 5.97 (dt,  $J$  = 15.4, 1.7 Hz, 1H,  $\alpha$ -of  $\alpha\beta$ ), 4.63 (br. s, 2'-H ascarylose), 4.43 (br. s, 2H, benzyl), 4.24 (d,  $J$  = 2.4 Hz, 2H, CH<sub>2</sub>-C≡CH) 4.19–4.17 (m, 2H, ArOCH<sub>2</sub>CH<sub>2</sub>) , 3.91–3.89 (m, 2H, ArOCH<sub>2</sub>CH<sub>2</sub>), 3.82–3.77 (m, 1H, 5'-H ascarylose), 3.70–3.68 (m, 1H, 7'-H sidechain), 3.57 (dt,  $J$  = 9.4, 6.2Hz 1H, 6'-H ascarylose), 3.48 (ddd,  $J$  = 16.4, 9.6, 4.6 Hz 1H, 3'-H ascarylose), 2.85 (t,  $J$  = 2.4 Hz, 1-H, CH<sub>2</sub>- C≡CH), 2.37–2.24 (m, 2H, 4' ascarylose), 1.92 (dt,  $J$  = 13.0, 3.9 Hz, 2H, 3'' sidechain), 1.74 (ddd,  $J$  = 15.4, 13.0, 3.9 Hz, 2H, 4' sidechain), 1.69–1.59 (m, 2H, 5',6'-H side chain) , 1.17 (d,  $J$  = 6.2 Hz, 3-H, 6'-CH<sub>3</sub> ascarylose), 1.13 (d,  $J$  = 6.1 Hz, 8'-H side chain).

**5R-(3'R,5'R-O-(4-(2-[Prop-2-ynyloxy)ethoxy]-4-[3-(trifluoromethyl)-3H-di-aziren-3-yl]benzoyl)-6'S-methyl-(2H)-tetrahydropyran-2'-yloxy)-2-hexanone (16)**



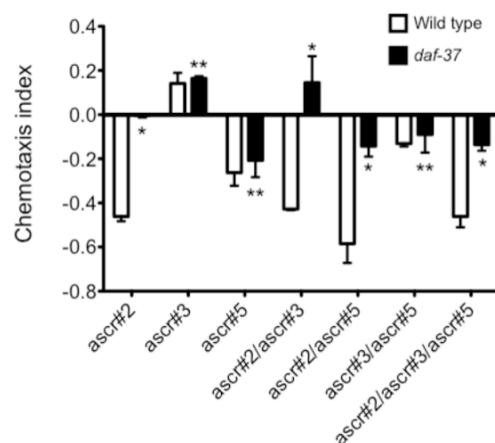
2-[2-(Prop-2-ynyloxy)ethoxy]-4-[3-(tri-fluoromethyl)-3H-diazirin-3-yl]benzoic acid (**17**) and ascr#2 were prepared as previously described<sup>13</sup> and combined as follows: An argon-charged dry dram vial containing a solution of ascr#2<sup>10</sup> (3.8 mg, 15.43 µmol) in dimethylformamide (300 µL), was cooled to 0 °C with stirring. Once the solution was cooled, N-(3-dime- thylaminopropyl)-N'-ethylcarbodiimide hydrochloride (6.3 mg; 33.0 µmol in 129 µL dichloromethane), the benzoic acid probe (10 mg; 30.86 µmol in 80 µL dimethylformamide), and 4-dimethylaminopyridine (4 mg, 30.86 µmol) were added. The solution was allowed to warm to 22 °C and was stirred for 16 h. Progress of the reaction was monitored by TLC (methanol:dichloromethane 1:10). The formation of the bis-acylated ascr#2 occurred as a competing reaction. After 16 h, the reaction mixture was evaporated to dryness in vacuo and was purified by silica gel flash chromatography [0–30% methanol (vol/vol) in dichloromethane with 0.1% acetic acid], yielding a mixture of the 5'- and 3'-linked probe (~2.5 mg). These isomers then were separated by HPLC using an Agilent Eclipse XDP-C18 preparative column [5–50% acetonitrile (vol/vol) in water with 0.1% acetic acid]. The final product, 5'-substituted ascr#2 probe (1.9 mg, 3.41 µmol, 23% yield, **16**). **<sup>1</sup>H NMR (600 MHz, Acetone-d6).** δ 7.77 (d, J = 8.2 Hz, 1H, 6-H benzoate), 7.04–7.01 (m, 1H, 5-H benzoate), 6.91 (d, J = 1.3 Hz, 1H, 3-H benzoate), 5.12 (ddd, J = 10.8 Hz, J = 9.5 Hz, J = 4.7 Hz, 1H, 5'-H ascaridole), 4.72 (br. s, 1H, 2'-H ascaridole), 4.32–4.29 (m, 2H, ArOCH<sub>2</sub>CH<sub>2</sub>), 4.27 (d, J = 2.4 Hz, 2H, CH<sub>2</sub>C≡CH), 3.92–3.90 (m, 3H, ArOCH<sub>2</sub>CH<sub>2</sub> and 6'-H ascaridole), 3.87–3.79 (m, 2H, 3'- H

ascarylose and 5"-H side chain), 2.98 (t, J = 2.4 Hz, 1-H, CH<sub>2</sub>- C≡CH), 2.60 (br. t, J = 7.3 Hz, 2-H, CH<sub>2</sub>C = OCH<sub>3</sub>), 2.17 (ddd, J<sub>4'ax,4'eq</sub> = 12.9, J<sub>4'eq,5'</sub> = 4.7, J<sub>3',4'eq</sub> = 3.2, 4-Heq ascarylose), 1.94 (ddd, J<sub>4'ax,4'eq</sub> = 12.9, J<sub>4'ax,5'</sub> = 11.2, J<sub>3',4'ax</sub> = 2.9, 4-Hax ascarylose), 1.83–1.70 (m, 2H, 4"-H side chain), 2.10 (br. s, 3-H, CH<sub>2</sub>C = OCH<sub>3</sub>), 1.21 (d, J = 6.3 Hz, 3-H, 6'-CH<sub>3</sub> ascarylose), 1.13 (d, J = 6.1 Hz, 6"-H side chain).

**Nematode Strains.** *C. elegans* strains were cultured according to standard techniques<sup>14</sup> unless otherwise noted. Worm strains and alleles used are LG I: *daf-16(m26)*; LG II: *daf-37(ttTi3058)*; LG III: *daf-7(e1372)*, *daf-2(e1370 unc-119(e2498))*; LG IV: *daf-38(ok2765)*, *daf-38(gk220535)*, *daf-10(e1387)*; LG V: *daf-11(m47)*, *him-5(e1476)*; LG X: *daf-3(mgDf90)*. *mEx182[daf-38p::gfp, rol-6(su1006)]*, *Ex184[daf-37p::dsRED, rol-6(su1006)]*, *mEx187[gpa-4p::daf-37, unc-119(+)]*, *mEx188[srbc-64p::daf-37, unc-119(+)]*, *mls41[daf-37p::daf-37, rol-6(su1006)]*, *otEx2503[gcy-27p::gfp, rol-6(su1006)]*, *otEx2310[gcy-19p::gfp, unc-122::gfp]*, *mls7[daf-7p::gfp::daf-7 3'UTR]*, *smls23[pdk-2::gfp]*, *daf-37; him-5, daf-37; unc-119, daf-37; mls7, daf-37; mls41, daf-37; daf-38, mls41; daf-38, daf-38; mls4, 1mEx182; mEx184, mEx184; otEx2503, mEx184; otEx2310, mEx184; smls23, mls41; daf-3, mls41; daf-16, mls41; daf-10*.

**PCR fusions and Transgenic Animals.** PCR fusion technique<sup>15</sup> was used for constructing transcriptional or translational fusion of *daf-37* gene with *dsRED* (or *gfp*). For transcriptional fusion, *dh\_233* and *dh\_235* primer pairs was used to amplify 2.5 kbp upstream region, first exon and first intron of *daf-37* gene. The *pdsRED::unc-54* (gift from John Tyson) was used for amplifying open reading frame of *dsRED* gene fused with *unc-54* 3'-UTR by using *C\_dsRED* and *D\_dsRED* primer pair. The second round PCR fusion was performed with *dh\_234* and *D\*\_dsRED* nested primer set. For translational fusion, *dh\_233* and *dh\_236* pair was used to amplify whole genomic region of *daf-37* gene including 2.5 kbp upstream region. And the second round PCR fusion was performed with *dh\_234* and *D\*\_dsRED* primer pair by using *C\_dsRED/D\_dsRED* PCR products described above. Three independent purified PCR products from each second round PCR were mixed for microinjection. The detailed PCR conditions will be provided upon request. Transgenic animals were generated by microinjection of PCR products (ca 200 ng) with *pRF4 [rol-6 (su1006)]* as a transgene marker (10ng/ul). The primer sequences used in this study will be provided upon request.

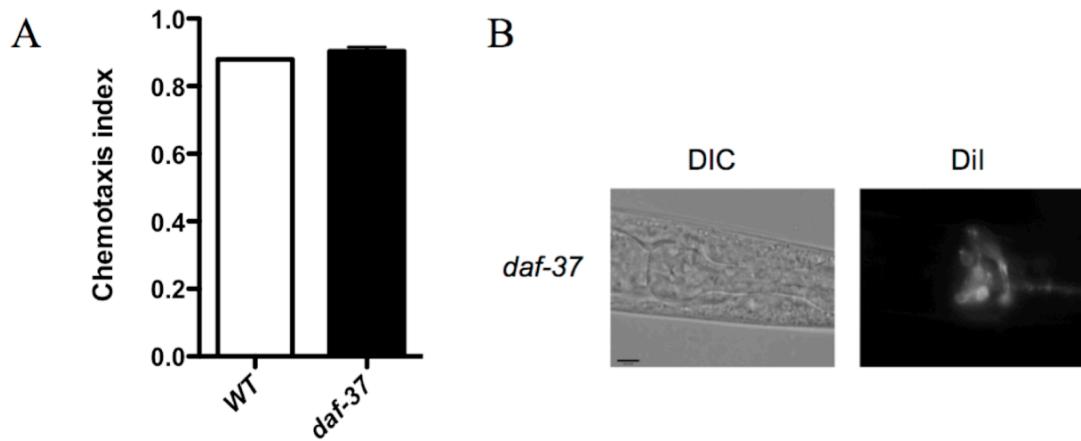
**Dauer Formation, Male Attraction and Hermaphrodite Repulsion.** To assay dauer formation, different concentrations of ascarosides were added to NG agar plates without peptone. Stock solutions of ascarosides were prepared by dissolving synthetic ascarosides in ethanol, which were diluted further as needed with water or into agar. The plates were then dried for 24 h at room temperature. A 2-ml overnight culture of OP50 bacteria was heat killed at 95 °C for 30min. The concentration was adjusted to 8µg/µl by adding sterile D.W. and 20 µl was spotted on each assay plate. Three or four adults were put on each plate to lay 40-50 eggs and removed. The number of dauer larvae was scored after 48-72 hrs at 25 °C. The male attraction assays and the hermaphrodite repulsion assays were performed as described<sup>6,16</sup>.



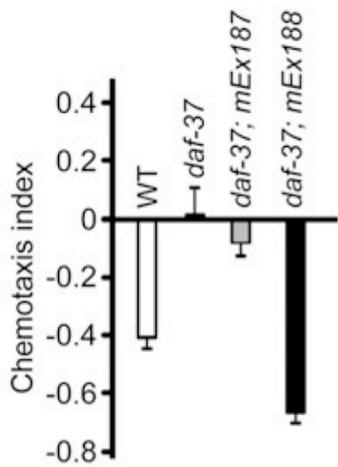
**Figure A.3: Ascaroside induced hermaphrodite repulsion in *daf-37* mutant worms.** Individual and combinations of equimolar concentrations of ascarosides (10 nM) were tested for hermaphrodite repulsion behavior. Hermaphrodite adults were synchronized (24 hrs after L4). Error bars represent SEM. (\*, p < 0.01; \*\*, p > 0.05 compared to wild type)

**Chemotaxis analysis.** Synchronized L4 stage of animals were washed 2 times with S-basal, 1 time with sterile D.W. and were put on S-basal agar assay plates containing (5 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM CaCl<sub>2</sub> and 1 mM MgSO<sub>4</sub>, 1.6% agar). The assay plate was marked with two spots 180 degrees apart each other 1 cm away from the edge. The two marks were spotted with 1  $\mu$ l of benzaldehyde (1:100 dilution in ethanol) or ethanol. The chemotaxis index was scored as described

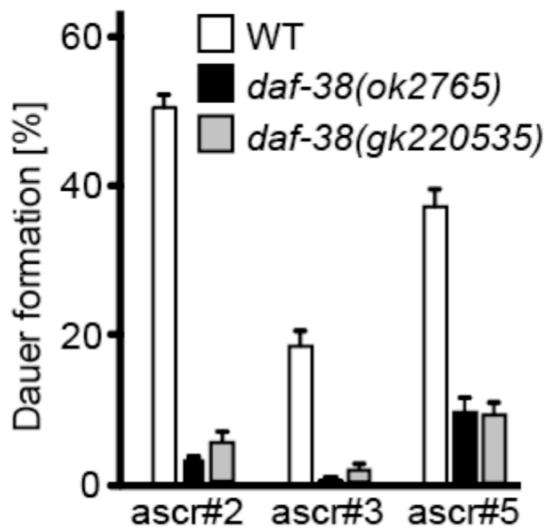
**Dye Filling Assay.** The age synchronized animals were washed in M9 buffer, incubated with the fluorescent dyes such as Dil (1,1'-dioctadecyl-3,3,3'3'-tetramethylindocarbocyanine perchlorate) at a final concentration of 10 ug/ml for 2 hrs, washed with M9 buffer. The animals were cultured on NG agar plates seeded with OP50 bacteria for an additional 1 hr then subjected for microscopy.



**Figure A.4: Chemotaxis analysis and dye filling assay in *daf-37* mutants.** (A) Bars represent the chemotaxis index of wild type and *daf-37* mutants upon benzaldehyde. ( $p > 0.05$ ). The chemotaxis assay was performed as described in supplemental experimental procedures. (B) The age synchronized L4 larva were collected and incubated with The fluorescent dyes Dil (1,1'-dioctadecyl-3,3,3'3'-tetramethylindocarbocyanine perchlorate) for dye filling assay



**Figure A.6: Rescue of *daf-37* phenotypes by *daf-37* overexpression and tissue-specific roles of DAF-37.** Hermaphrodite repulsion behavior in adult worms in wildtype, *daf-37*, *daf-37*; mEx187[gpa-4p::*daf-37*] (ASI-specific *daf-37*-expression) and *daf-37*; mEx188[srbc-64p::*daf-37*] (ASK-specific *daf-37* expression).. Error bars represent SEM



**Figure A.7: Dauer induction by ascr#2 (700 nM), ascr#3 (700 nM), or ascr#5 (500 nM) in wild type and two *daf-38* mutant alleles.** DAF-38 appears to play in mediating role, enhancing ascr perception

**Pb-FRET Microscopic Analysis.** HEK293 cells expressing HA-DAF-38 and/or cMyc-DAF-37 were grown on coverslips to 60-70% confluence and treated with 100 nM ascr#2 for 15 min at 37 °C. Cells were fixed with 4% paraformaldehyde for 20 min on ice and processed for immunocytochemistry as previously described<sup>3,4</sup>. In order to create a donor-acceptor pair, monoclonal anti-HA and polyclonal cMyc primary antibodies were used, followed by incubation with FITC- and rhodamine-conjugated secondary antibodies, respectively. The plasma membrane region was used to analyze the photobleaching decay as previously described<sup>3,4</sup>. The FRET efficiency (E) was calculated based upon the photo bleaching (Pb) time constants of the donor taken in the absence (D-A) and presence (D+A) of acceptor according to  $E=1-(D-A/D+A) \times 100$ .

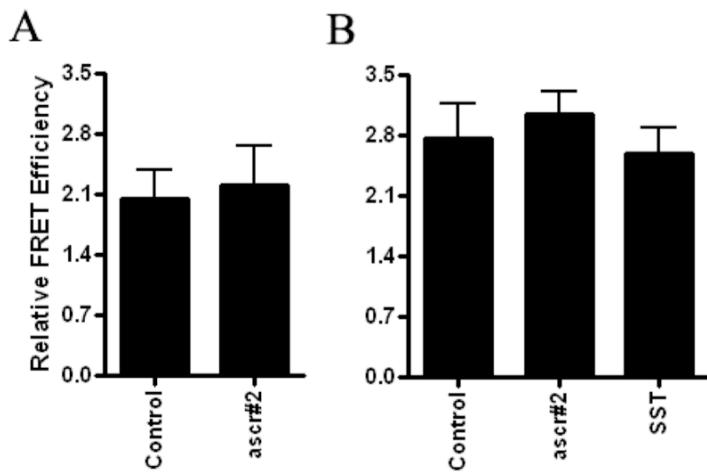
Receptor	Condition	Time constant ( $\tau_{avg}$ )	Average FRET efficiency (%)
<b>DAF-37 without ascr#2</b>	Donor	18.4 s	
	Donor + Acceptor	21.6 s	15.4 ± 0.5
<b>DAF-37 with ascr#2</b>	Donor	18 s	
	Donor + Acceptor	22.4s	19.7 ± 0.7*
<b>DAF-37/38 without ascr#2</b>	Donor	17.6 s	
	Donor + Acceptor	21.5s	18.3 ± 0.8†
<b>DAF-37/38 without ascr#2</b>	Donor	17.8 s	
	Donor + Acceptor	22.5s	21.5 ± 0†,‡

**Table A.1:** Photobleaching FRET analysis of HEK293 cells expressing DAF-37 or DAF-37/DAF-38 in the presence or absence of ascr#2. Data represent mean ± SE of three independent experiments performed in duplicate. Data were analyzed using ANOVA followed by a Newman–Keuls Multiple Comparison Test. Donor and Donor + Acceptor correspond to donor in the absence and presence of acceptor, respectively.  $\tau_{avg}$ , mean of photobleaching time constants.

\*P < 0.001 (comparison within a group).

†P < 0.001 (comparison between groups).

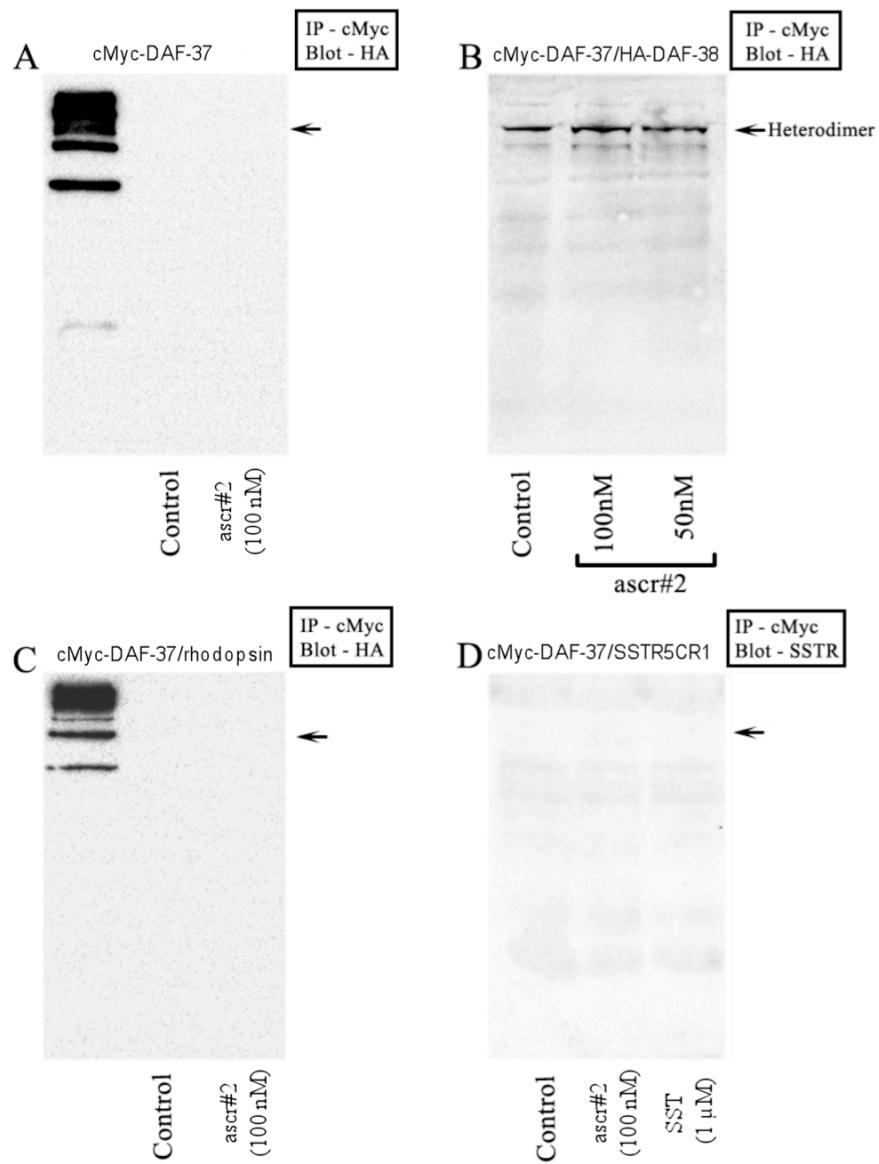
‡P < 0.01 (comparison within a group).



**Figure A.8:** Relative FRET efficiencies in cells cotransfected with cMyc-DAF-37 and HA-rhodopsin or cMyc-DAF-37/SST5CR1. (A) cMyc-DAF-37/HA-Rhodopsin displayed a relative FRET efficiency of  $2.05 \pm 0.57\%$  in basal condition with no significant effect of ascr#2 treatment ( $2.20 \pm 0.81\%$ ). (B) DAF-37/SST5CR1 displayed a relative FRET efficiency of  $2.76 \pm 0.71\%$  in basal condition with no significant effect of ascr#2 treatment ( $3.04 \pm 0.47\%$ ) or somatostatin (SST) treatment ( $2.59 \pm 0.51\%$ ), indicating the absence of receptor complex formation. Mean  $\pm$  S.E. are representative of three independent experiments. Data analysis was done by using student *t*-test and one way ANOVA (post hoc Dunnett's Multiple Comparison test) to compare with control and treated conditions

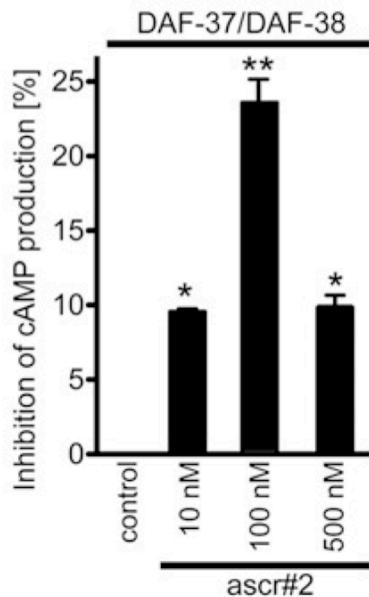
**Co-immunoprecipitation (Co-IP).** Cotransfected HEK-293 cells [cMyc-DAF-37/HA-DAF-38 or cMyc-DAF-37/HA-Rhodopsin or cMyc-DAF-37/SST5CR1] were treated with ascr#2 (50 nM, 100 nM) and/or SST (1  $\mu$ M) for 30 min at 37 °C. Cells were lysed and membrane protein (250  $\mu$ g) was solubilized in 1 ml of radioimmune precipitation assay buffer (RIPA Buffer, 150 mM NaCl, 50 mM Tris-HCL, 1% Nonidet P-40, 0.1% SDS, 0.5% sodium deoxycholate, pH 8.0) for 1 h at 4 °C and Co-IP was performed as described earlier<sup>17</sup>. Briefly, samples were incubated with 1  $\mu$ g antibody (anti-cMyc) overnight at 4 °C and 25  $\mu$ l of protein A/G-agarose beads were added to immunoprecipitate antibody for 2 h at 4 °C. Agarose beads were then washed and solubilized in Laemmli sample buffer (Bio-Rad). Samples were fractionated by electrophoresis on a 10% SDS-polyacrylamide gel and transferred to a 0.2  $\mu$ m

nitrocellulose membrane. Membranes were blotted with anti-HA antibody (dilution 1:500) for the expression of HA-DAF-38 or HA-Rhodopsin and anti-SSTR specific antibody (dilution 1:500) for the expression of SST5CR1. Bands were detected by chemiluminescence using ECL western blotting detection reagent from Millipore according to the manufacturer's instructions. Images were captured using an Alpha Innotech FluorChem 8800 gel box imager (Alpha Innotech Co., San Leandro, CA). It should be noted that many studies have reported receptor homo-/heterodimers by western blotting under reducing condition<sup>18-23</sup>. Importantly, monomer to dimer ratios for several receptors including dopamine receptor 2 remained mostly unchanged under reducing or non-reducing condition in co-immunoprecipitation experiments<sup>19-22</sup>.

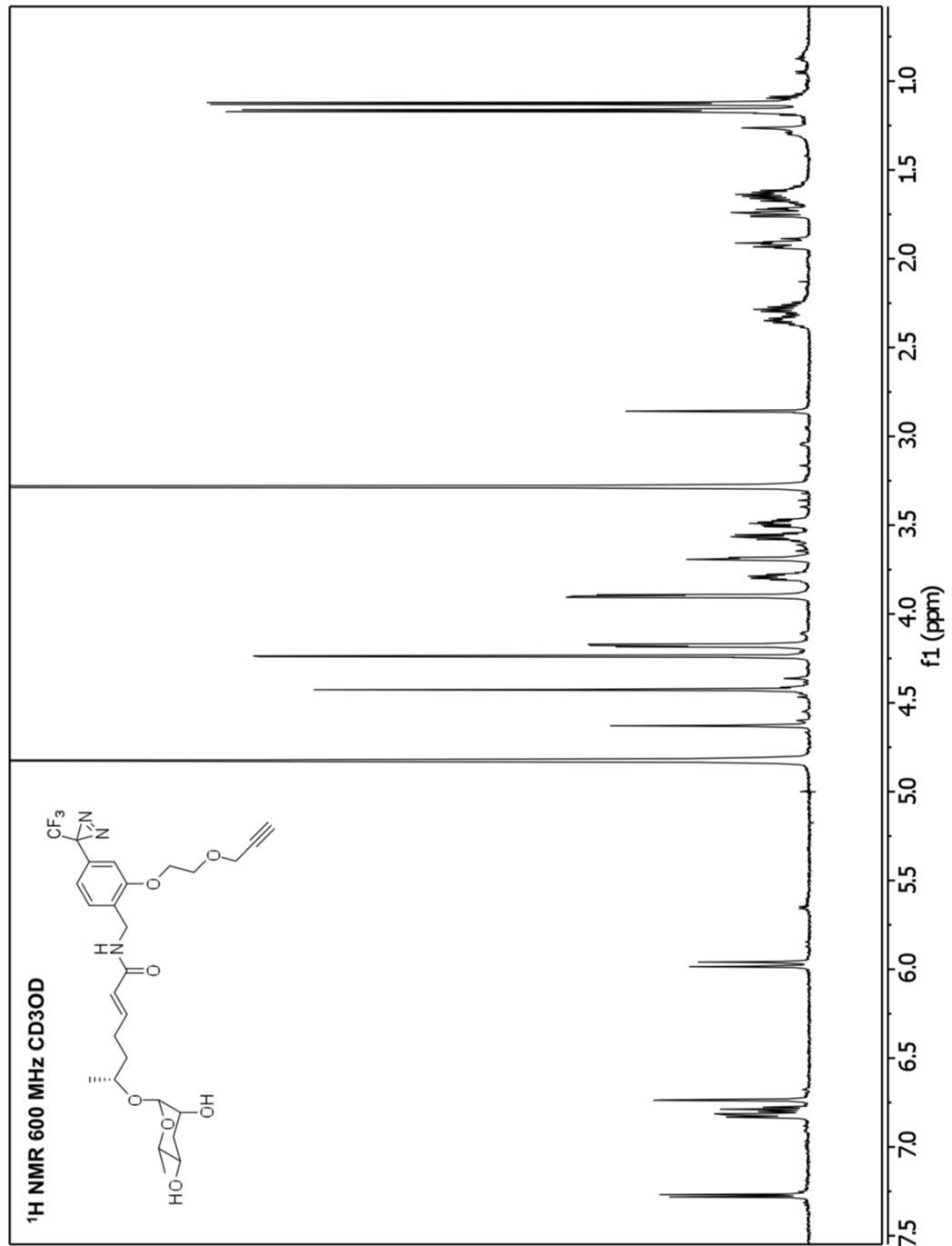


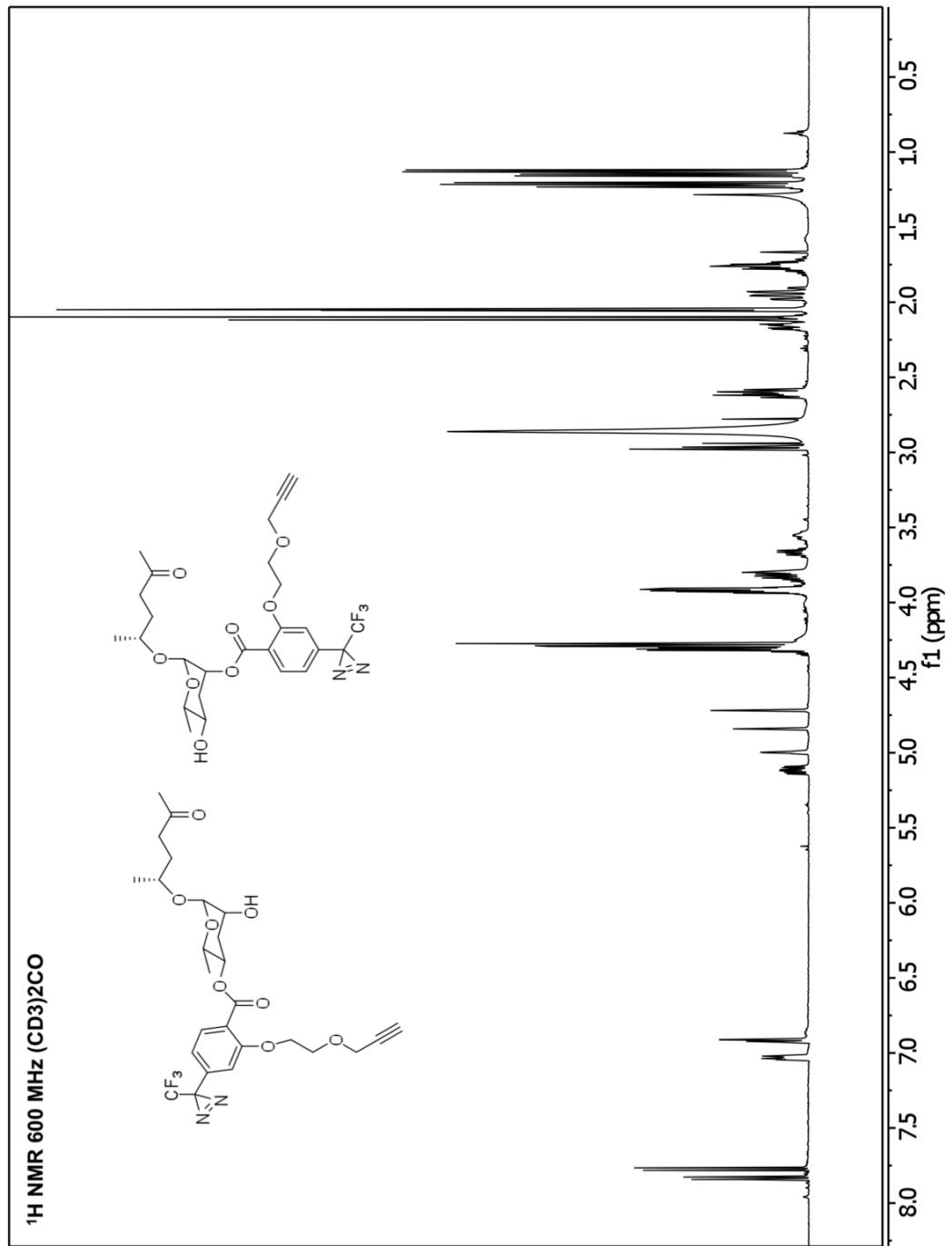
**Figure A.9:** Heterodimerization of cMyc-DAF-37/HA-DAF-38. (A) Co-IP of cells transfected only with cMyc-DAF-37 in the absence and presence of ascr#2 produced no band. (B) Co-IP of cells co-transfected with cMyc-DAF-37 and HA-DAF-38 resulted in a single band at the expected size of ~110 kDa, indicated the presence of cMyc-DAF-37/HA-DAF-38 heterodimers in DAF-37 immunoprecipitate (arrow). (C and D) To determine the specificity of heterodimerization, cells cotransfected with cMyc-DAF-37/HA-rhodopsin or cMyc-DAF-37/SST5CR1 were probed for the formation of heterodimers, in the absence and presence of ascr#2 and somatostatin (SST). No specific bands were observed in these experiments, indicating the absence of heterodimerization. Data are representative of three independent experiments

**Coupling to Adenylyl Cyclase.** To determine the receptor coupling to AC, mono- and cotransfected cells were grown in 6-well culture plates (>70% cell confluence) and processed for cAMP assay as described earlier<sup>4</sup>. Briefly, cells were incubated with 20  $\mu$ M forskolin (FSK) and 0.5 mM 3-isobutyl-1-methylxanthine for 30 min at 37 °C in the presence or absence of different concentrations of ascr#2. Cells were then scraped in 0.1 N HCl and cAMP was determined by immunoassay using a cAMP Kit from BioVision, Inc. CA, USA (39).



**Figure A.10. Effect of ascr#2 on forskolin-induced cAMP production in HEK293 cells stably transfected with DAF-37 and DAF-38.** Data are presented as the percent inhibition of camp production in the absence of ascr#2. Error bars represent SE. Data analysis used ANOVA and post hoc Dunnett's test for comparison against basal level.\*P < 0.05; \*\*P < 0.01.





## REFERENCES

- (1) Finney, M.; Ruvkun, G. *Cell* **1990**, *63*, 895–905.
- (2) Blacque, O. E.; Reardon, M. J.; Li, C.; McCarthy, J.; Mahjoub, M. R.; Ansley, S. J.; Badano, J. L.; Mah, A. K.; Beales, P. L.; Davidson, W. S.; Johnsen, R. C.; Audeh, M.; Plasterk, R. H. A.; Baillie, D. L.; Katsanis, N.; Quarmby, L. M.; Wicks, S. R.; Leroux, M. R. *Genes & Development* **2004**, *18*, 1630–1642.
- (3) Grant, M.; Patel, R. C.; Kumar, U. *J. Biol. Chem.* **2004**, *279*, 38636–38643.
- (4) Somvanshi, R. K.; Billova, S.; Kharmate, G.; Rajput, P. S. *Cellular signalling* **2009**.
- (5) Min, K. C.; Zvyaga, T. A.; Cypress, A. M.; Sakmar, T. P. *Journal of Biological Chemistry* **1993**, *268*, 9400.
- (6) Macosko, E. Z.; Pokala, N.; Feinberg, E. H.; Chalasani, S. H.; Butcher, R. A.; Clardy, J.; Bargmann, C. I. *Nature* **2009**, *458*, 1171–1175.
- (7) Watzl, C. *Current Protocols in Protein Science*.
- (8) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. *J. Am. Chem. Soc.* **2009**, *131*, 4967–4975.
- (9) Ullman, E. F.; Kirakossian, H.; Singh, S.; Wu, Z. P.; Irvin, B. R.; Pease, J. S.; Switchenko, A. C.; Irvine, J. D.; Dafforn, A.; Skold, C. N. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5426–5430.
- (10) Butcher, R. A.; Fujita, M.; Schroeder, F. C.; Clardy, J. *Nat Chem Biol* **2007**, *3*, 420–422.
- (11) Pungaliya, C.; Srinivasan, J.; Fox, B. W.; Malik, R. U.; Ludewig, A. H.; Sternberg, P. W.; Schroeder, F. C. *Proceedings of the National Academy of Sciences* **2009**, *106*, 7708–7713.
- (12) Kan, T.; Kita, Y.; Morohashi, Y.; Tominari, Y.; Hosoda, S.; Tomita, T.; Natsugari, H.; Iwatubo, T.; Fukuyama, T. *Org. Lett.* **2007**, *9*, 2055–2058.
- (13) Mayer, T. *Eur. J. Org. Chem.* **2007**.
- (14) Brenner, S. *Genetics* **1974**.
- (15) Hobert, O. *BioTechniques* **2002**, *32*, 728–730.
- (16) Srinivasan, J.; Kaplan, F.; Ajredini, R.; Zachariah, C.; Alborn, H. T.; Teal, P. E. A.; Malik, R. U.; Edison, A. S.; Sternberg, P. W.; Schroeder, F. C. *Nature* **2008**, *454*, 1115–1118.
- (17) Somvanshi, R. K.; Chaudhari, N.; Qiu, X.; Kumar, U. **2011**.
- (18) Lee, S. P.; O'Dowd, B. F.; George, S. R. *Life sciences* **2003**.
- (19) George, S. R. *Journal of Biological Chemistry* **2000**, *275*, 26128–26135.
- (20) Hebert, T. E.; Moffett, S.; Morello, J. P.; Loisel, T. P.; Bichet, D. G.; Barret, C.; Bouvier, M. *J. Biol. Chem.* **1996**, *271*, 16384–16392.
- (21) Devost, D. *Journal of Molecular Endocrinology* **2003**, *31*, 461–471.
- (22) Lee, S. P.; O'Dowd, B. F.; Rajaram, R. D.; Nguyen, T.; George, S. R. *Biochemistry* **2003**, *42*, 11023–11031.
- (23) Zhu, W. Z. *Circulation Research* **2005**, *97*, 244–251.

## APPENDIX B

### ADVANCEMENTS TOWARDS THE DIFFERENTIAL LABELING OF SMALL MOLECULE BINDING SITES USING PHOTOAFFINITY PROBES COUPLED TO CLEAVABLE STABLE ISOTOPE CLICK-CHEMISTRY TAGS (CSICT).

#### **Cleavable Linker Use with ascr#2 Probe and DAF-37**

**Cell Tranfection and Photo-crosslinking.** HEK293T cells ( $1.3 \times 10^6$ ) were split into a 10-cm plate with DMEM [10% FBS (vol/vol), 1% penicillin/streptomycin] and transfected using standard methods<sup>1</sup>. This was repeated twice for each condition, resulting in three 10 cm plates per-condition, with a total of nine 10 cm plates. The cDNA constructs *daf-37* and *gfp* that were used in the previous transfection had been previously generated<sup>2</sup>. Two days after transfection, the growth medium was aspirated, and the cells were washed with PBS buffer. After washing, PBS (4.5 mL, pH 7.4) was pipetted onto the cells, and subsequently the ascr#2 probe (17  $\mu$ L, 6.5 mM in DMSO) was added with gentle mixing. The cells were incubated with the probe at 37 °C for 15 min and then were cooled to 4 °C. Next, for photocrosslinking, the cells were held for 15 min at 4 °C under a 100-W bulb (Blak-Ray-B100AP; UVP) at a 4-cm working distance. Then the medium was removed, and the cells were harvested using PBS supplemented with EDTA-free Roche Complete Mini Protease Mixture Inhibitor. The cells were pelleted at 4 °C at 1,000  $\times$  g for 3 min, and the supernatant was removed.

**Cell Lysis and Click-chemistry.** The pelleted cells were lysized with 4% SDS in PBS (1.5 mL, pH 7.4) for 20 min on a nurator. The cells were then sonicated with a probe sonicator (Bason Sonifer) for ten blasts at 10% before being spun down at 20,000g for 20 min at RT. The supernatant was removed and the lysates volume was adjusted to 1.619 mL with additional lysis solution added if necessary. This solution was then subjected to click-chemistry with addition of BTTP<sup>3</sup> (35  $\mu$ L, final conc 500  $\mu$ M [frozen

stock 25 mM in 9:1, water:ethanol]) vortexed, CuSO<sub>4</sub>·5H<sub>2</sub>O (8.8 µL, final conc 250 µM (50 mM fresh stock in water) vortexed, sodium ascorbate (52.5 µL, final conc 3 mM [100 mM in water freshly prepared]) vortexed, cleavable diazobenzene-azido-biotin (**14**) (35 µL, final conc 100 µM [5 mM frozen stock in DMSO]) vortexed, before being placed on a rotary shaker for 1 h. After the 1h the reaction was split into seven aliquots of 250 µL in individual 2 mL eppendorfs. These were each precipitated using chloroform methanol techniques<sup>4</sup>. One washed, the wafer like protein pellets were carefully transferred to same 2 mL eppendorf and 1.5 mL of 6 M urea, 2 M thiourea, 10 mM HEPES pH 7.4 was added. This solution was then repeatedly vortexed before being allowed to stand overnight at RT, resolubilizations were also carried out successfully with 4% SDS, 150 mM NaCl in PBS pH 7.4.

**Streptavidin Binding and Elution.** After standing at RT overnight the resolubilization solution was vigorously vortexed before being centrifuged at 20,000 g for 15 min. The supernatant was removed and transferred to a new eppendorf. Streptavidin beads (100 µL, high capacity streptavidin agarose resin; Thermoscientific). Note, do not use magnetic streptavidin beads as they are not compatible with the dithionite reduction) that had been were washed in the resolubilization buffer (3 x 400 µL, 1,000 g for 2 min) were added to the solution and incubated for 1.5 h at RT on an end over shaker. After binding, the beads were transferred to a 1.5 mL eppendorf and washed with resolubilization solution (3 x 400 µL) followed by 1 % SDS in PBS (6 x 400 µL, pH 7.4). The beads were then subjected to reductive elution conditions with 25 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> in 1 % SDS<sup>5</sup> in PBS (3 x 165 µL, pH 7.4). These elutions were pooled and added to 3 k size exclusion spin column (Nanosep 3k omega; Pall). The spin column was centrifuged at 14,000 g for 10 min and the reservoir was washed with 1 % SDS in PBS (300 µL, pH 7.4), this was repeated twice. The contents of the reservoir were transfer to a new eppendorf and speed-vaced till dryness. The pellets are then resolubilized and prepped for SDS/PAGE, with the addition of 60 µL sample buffer (NuPAGE; Invitrogen). The

samples are allowed to stand for at least 16 h at RT with occasional vortexing, without further heating. Once the protein pellets have gone back into solution, the samples were loaded and run on 4–12% bis-Tris gel (NuPAGE; Invitrogen) at 200 V for 1 h. The gel then was blotted onto a PVDF membrane (Immobilon-FL; Millipore) and was blocked with a 5% skim milk powder solution (wt/vol) in Tris-buffered saline and Tween 20 (TBST). The blot then was treated with an anti-1D4 antibody ([www.flintbox.com](http://www.flintbox.com)) (diluted 1:4,000 into 5% skim milk powder in TBST [wt/vol]), washed, and treated with goat anti-mouse-HRP (Vector Laboratories) (diluted 1:10,000 into 5% skim milk powder in TBST [wt/vol]). The blot then was washed and detected by developing (SRX-101A; Konica Minolta) with luminol (SuperSignal® femto west; Thermoscientific).

### Instrumentation and general procedures

NMR spectra were recorded on Varian INOVA 600 (600 MHz), spectrometers in Cornell University's NMR facility.  $^1\text{H}$  NMR chemical shifts are reported in ppm ( $\delta$ ) relative to residual solvent peaks ( $\delta$  2.05 ppm for acetone-d<sub>6</sub>, 7.26 ppm for CDCl<sub>3</sub>, 3.31 ppm for CD<sub>3</sub>OD, 4.79 ppm for D<sub>2</sub>O, 2.50 ppm (CD<sub>3</sub>)<sub>2</sub>SO ). NMR-spectroscopic data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), and integration.  $^{13}\text{C}$  NMR chemical shifts are reported in ppm ( $\delta$ ) relative to CHCl<sub>3</sub> ( $\delta$  77.2) in CDCl<sub>3</sub>, CH<sub>3</sub>OH ( $\delta$  49.0) in CD<sub>3</sub>OD, CH<sub>3</sub>OH ( $\delta$  49.0) in D<sub>2</sub>O and (CH<sub>3</sub>)<sub>2</sub>SO ( $\delta$  39.5) in (CD<sub>3</sub>)<sub>2</sub>SO).  $^2\text{H}$  NMR chemical shifts are reported in ppm ( $\delta$ ) relative to the natural deuterium abundance of solvent peaks( $\delta$  2.50 ppm for (CD<sub>3</sub>)<sub>2</sub>SO, 2.05 ppm for acetone-d<sub>6</sub> and 3.31 ppm for CD<sub>3</sub>OD. Positive-ion electrospray ionization mass spectra (ESI<sup>+</sup>-MS) were obtained on a Micromass Quattro II mass spectrometer using MassLynx software. High-resolution mass spectrometry unless otherwise stated was performed on a Waters nanoACQUITY UPLC system equipped with a Waters Acquity UPLC HSS C-18 column (2.1 x 100 mm, 1.8  $\mu\text{m}$  particle diameter) connected to a Xevo G2 QToF Mass Spectrometer operated in electrospray positive (ESI+) ionization mode. Low-resolution mass spectrometer was

performed on an HPLC-MS system equipped with a diode array detector and connected to a Quattro II spectrometer (Micromass/Waters) operated in positive electrospray ionization (ESI+) mode. Data acquisition and processing for the HPLC-MS was controlled by Water MassLynx software. Thin-layer chromatography (TLC) was performed using J. T. Baker Silica Gel IB2-F. Flash chromatography was performed using Teledyne Isco CombiFlash systems and Teledyne Isco RediSep Rf silica columns. Unless stated otherwise, reagents were purchased from Sigma-Aldrich and used without further purification. *N,N*-dimethylformamide (DMF), dichloromethane (DCM), dimethyl sulfoxide (DMSO), and tetrahydrofuran (THF) were dried over 4 Å molecular sieves prior to use.

## Application of CSICT Azides

**Click-chemistry of CSICT Azides with BSA-alkyne; Click and Cleavage.** To demonstrate the ability of the second generation cleavable pyridine azido-biotin linker (**11**) to efficiently perform in click reactions with less copper compared to first generation cleavable azido-biotin linker (**14**). Also being demonstrated was that the presence of the pyridine moiety did not disrupt the dithionite reductive elution from streptavidin beads. BSA-alkyne<sup>6</sup> (5 µg, at 1 µg/µL) was added to a 4% SDS protein lysate from HEK293T cells (185 µL, 1.44 µg/µL). To this a click-chemistry reaction cocktail was added to bring the final volume to 200 µL and carried out for 1 h as described above. The final concentrations of sodium ascorbate and azide were kept constant at 2.5 mM and 100 µM respectively, while the final concentration of BTTP and CuSO<sub>4</sub>·5H<sub>2</sub>O were varied as shown in (Table 2.1). The reactions were then precipitated with a standard chloroform methanol approach. The resulting protein pellet was resolubilized in 4% SDS in PBS (60 µL, pH 7.4), vortexed repeatedly and allowed to sit overnight at RT. The next morning the solution was vortex and centrifuged at 20,000 g and the supernatant was removed. An aliquot (10 µL) of each was taken for western blot analysis with the remainder of the sample stored at RT. The aliquot was prepped with sample buffer and run on SDS PAGE, blotted on PVDF and blocked as previously described. The blot then was treated with an streptavidin-HRP (Thermoscientific) (diluted 1:2,000 into 5% skim milk powder in TBST [wt/vol]) for 1 h. The blot then was washed and detected by developing with luminol.

To assay the effect of the pyridine moiety on linker cleavage, clicked BSA-alkyne samples were selected using the results of streptavidin-HRP western blot. Samples, **3** and **5** and samples **6**, and **8** from azide treatments **14** and **11** respectively, had undergone significant biotinylation via click-chemistry were directly compared for dithionite mediated release of BSA, with samples **1** and **2** serving as negative controls. **C1** and **C2** are 33 ng and 330 ng standard of unmodified BSA used for quantification purposes. The remainder of samples **5**, **6**, **7** and **8**, that was not used in the western

blot (**Figure 2.8**) was subjected to the same streptavidin binding, washing and elution method as previously described. The dried BSA pellets from the elution were prepped and run on SDS-PAGE. The gel was then silver stained (Sliver Quest™ Kit; Invitrogen) with a 3 min developing period (**Figure 2.8**).

**Preparation of Gel Bands for MS.** The excised gel bands were carefully cut into 1 mm cubes and hydrated with water (100 µL) for 5 min. The water was removed and a 100 mM ammonium bicarbonate:water solution (100 µL, 50:50) was added and let sit for 10 min before being removed. Acetonitrile (50 µL) was added and let sit for 5 min before being removed, the gel pieces appeared shrunken and opaque. The gel pieces were then dried, by placing in a ventilated fume hood for 5-10 minutes. A trypsin solution was prepared on ice 10 ng/µL trypsin in 4 °C 50 mM ammonium bicarbonate with 10% acetonitrile. The gel pieces were rehydrated with 15 µL of the trypsin solution per tube and incubated on ice for 30 mins. Then 50 mM ammonium bicarbonate (10 µL) was added to immerse the gel pieces and the tubes were incubated at 30 °C in water bath overnight. Formic acid was added to each sample to achieve a final formic acid concentration 1% to stop enzymatic reaction. The supernatant was removed and saved in another tube. A solution of 50 acetonitrile and 5% formic acid (30 µL) was then added and allowed to sit for 45 min. the tube was then placed in a bath sonicator for 5 min and the supernatant was removed and added to the other tube that contains the quenched enzyme solution. This wash and sonication step was repeated. The gel pieces were then washed with 90% acetonitrile with 5% formic acid (30 µL) for 5 min and the supernatant was removed and placed with the other washes. The combined washes were then placed in a speed vacuum until dry.

**Protein mapping by nanoLC/MS/MS analyses.** The tryptic digest was reconstituted in 20 µL of 2% ACN with 0.5% FA for nanoLC-ESI-MS/MS analysis, which was carried out by a LTQ-Orbitrap Velos mass spectrometer (Thermo-Fisher Scientific, San Jose, CA)

equipped with a “CorConneX” nano ion source device (CorSolutions LLC, Ithaca, NY). The Orbitrap was interfaced with a nano HPLC carried out by Dionex UltiMate3000 MDLC system (Dionex, Sunnyvale, CA). The gel extracted peptide samples (2-4 $\mu$ L) were injected onto a PepMap C18 trap column (5  $\mu$ m, 300  $\mu$ m  $\times$  5 mm, Dionex) at 20  $\mu$ L/min flow rate for on-line desalting and then separated on a PepMap C18 RP nano column (3  $\mu$ m, 75  $\mu$ m  $\times$  15 cm, Dionex) which was installed in the “CorConneX” device with a 10- $\mu$ m spray emitter (NewObjective, Woburn, MA). The peptides were then eluted in a 90 min gradient of 5% to 38% acetonitrile (ACN) in 0.1% formic acid at a flow rate of 300 nL/min. The Orbitrap Velos was operated in positive ion mode with nano spray voltage set at 1.6 kV and source temperature at 275 °C. Internal calibration was performed with the background ion signal at m/z 445.120025 as the lock mass. The instrument was performed at data-dependent acquisition (DDA) mode by one precursor ions MS survey scan from m/z 300 to 1800 at resolution 60,000 using FT mass analyzer followed by up to 10 MS/MS scans at unit resolution on top 10 most intensity peaks with multiple charged ions above an ion intensity threshold of 10000 in FT survey scan mode. Dynamic exclusion parameters were set at repeat count 1 with a 15s repeat duration, exclusion list size of 500, 30 s exclusion duration, and  $\pm 10$  ppm exclusion mass width. Collision induced dissociation (CID) parameters were set at the following values: isolation width 2.0 m/z, normalized collision energy 35 %, activation Q at 0.25, and activation time 10 ms. All data are acquired under Xcalibur 2.1 operation software (Thermo-Fisher Scientific).

**Protein Mapping by nanoLC/MS/MS Analyses.** The tryptic digest was reconstituted in 20  $\mu$ L of 2% ACN with 0.5% FA for nanoLC-ESI-MS/MS analysis, which was carried out by a LTQ-Orbitrap Velos mass spectrometer (Thermo-Fisher Scientific, San Jose, CA) equipped with a “CorConneX” nano ion source device (CorSolutions LLC, Ithaca, NY). The Orbitrap was interfaced with a nano HPLC carried out by Dionex UltiMate3000 MDLC system (Dionex, Sunnyvale, CA). The gel extracted peptide samples (2-4 $\mu$ L)

were injected onto a PepMap C18 trap column (5 µm, 300 µm × 5 mm, Dionex) at 20 µL/min flow rate for on-line desalting and then separated on a PepMap C18 RP nano column (3 µm, 75 µm × 15 cm, Dionex) which was installed in the “CorConneX” device with a 10-µm spray emitter (NewObjective, Woburn, MA). The peptides were then eluted in a 90 min gradient of 5% to 38% acetonitrile (ACN) in 0.1% formic acid at a flow rate of 300 nL/min. The Orbitrap Velos was operated in positive ion mode with nano spray voltage set at 1.6 kV and source temperature at 275 °C. Internal calibration was performed with the background ion signal at m/z 445.120025 as the lock mass. The instrument was performed at data-dependent acquisition (DDA) mode by one precursor ions MS survey scan from m/z 300 to 1800 at resolution 60,000 using FT mass analyzer followed by up to 10 MS/MS scans at unit resolution on top 10 most intensity peaks with multiple charged ions above an ion intensity threshold of 10000 in FT survey scan mode. Dynamic exclusion parameters were set at repeat count 1 with a 15s repeat duration, exclusion list size of 500, 30 s exclusion duration, and ±10 ppm exclusion mass width. Collision induced dissociation (CID) parameters were set at the following values: isolation width 2.0 m/z, normalized collision energy 35 %, activation Q at 0.25, and activation time 10 ms. All data are acquired under Xcalibur 2.1 operation software (Thermo-Fisher Scientific).

### Synthesis of flgII-28 Probe (flgII-28\*)

The lyophilized flg-28 peptide (ESTNILQRMRELAVQSRNDSNSSTD RD, 2mg, 640 nmol, from Genscript) was solubilized in PBS (250 µL, pH 8) and placed in a HPLC vial with a stir bar. DMSO (125 µL) was added and the solution was allowed to cool before addition of the NHS-probe<sup>7</sup> (**41**, 125 µL of 7.0 mM in DMSO). The reaction was stirred for 3.5 h before the water was removed in vacuo, with the remaining DMSO being lyophilized off. Once dry, the sample was resolubilized in water (3 x 200 µL) and separated via preparative HPLC as previously described, unreacted flg-28 peptide was also recovered (0.5 mg, 160 nmol). flgII-28\* probe (1.2 mg, 384 nmol, 60% yield, **27**).

**ESI-MS (*m/z*): [M+4H<sup>+</sup>] calcd. for C<sub>139</sub>H<sub>226</sub>F<sub>3</sub>N<sub>47</sub>O<sub>54</sub>S: 877.1673; found 877.1500, 20 ppm diff.**

The purified flgII-28\* probe peptide was then subjected to click-chemistry conditions with the heavy (**11-d<sub>7</sub>**) and light (**11**) azide in a 200 μL reaction volume. The same concentrations of reagents were used as described above and these click reactions were allowed to proceed for 3 h before being concentrated down in vacuo and purified via HPLC to yield the heavy-flg-II28 probe (**ESI-MS (*m/z*): [M+4H<sup>+</sup>] calcd. for C<sub>177</sub>H<sub>266</sub>D<sub>7</sub>F<sub>3</sub>N<sub>58</sub>O<sub>61</sub>S<sub>2</sub>: 1078.7469; found 1078.7445**) and the light-flg-II28 probe (**ESI-MS (*m/z*): [M+4H<sup>+</sup>] calcd. for C<sub>177</sub>H<sub>273</sub>F<sub>3</sub>N<sub>58</sub>O<sub>61</sub>S<sub>2</sub>: 1076.9859; found 1076.9850**).

The purified modified peptides were lyophilized and brought up in acetonitrile:water (100 μL, 90:10). A aliquot (25 μL) of the sample was transferred to an HPLC vial insert and 10 x PBS (3 μL, pH 7.4) was added followed by a dithionite solution (1.25 μL, freshly prepared 2.5 M) before being allowed to stand at RT for 15 min. The reduction reaction desalting was then directly injected in to the LC-MS. Cleaved-heavy-flgII-28\* probe (**ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. C<sub>154</sub>H<sub>231</sub>D<sub>7</sub>F<sub>3</sub>N<sub>53</sub>O<sub>56</sub>S: data currently being recollected**) and cleaved-light-flgII-28\* probe (**ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. C<sub>154</sub>H<sub>238</sub>F<sub>3</sub>N<sub>53</sub>O<sub>56</sub>S data currently being recollected**).

### **Bioassay of flgII-28\***

**Cell Suspension Culture Maintenance.** A wild tomato (*S. peruvianum*, formerly *Lycopersicon. peruvianum*) suspension-cultured cell line<sup>8</sup> was kindly provided by Jürg Felix, Universität Tübingen. Cells were cultivated under constant light with shaking at 110-125 rpm. Cells were subcultured weekly by transferring 5ml of cell suspension to 45ml of sterile media as previously described<sup>8</sup> into 125-ml Erlenmeyer flasks. Cells were used for experiments 7-8 days after subculturing.

**Medium Alkalization Response (AR) Assays.** For the experiments, 1.5ml cells were transferred to each well of 12 multi-well plates (Corning, Tewksbury, MA, USA) that were shaken at 125rpm on an orbital shaker under ambient room light and temperature conditions. After a brief equilibration period (approximately 1 hour), the initial pH of the medium was measured using a standard pH probe (Mettler Toledo, Columbus, OH, USA) before adding the appropriate peptides at the indicated concentrations. The pH of the medium of each well was measured at the indicated times. Three replicate wells were used for each treatment, and the pH values for each treatment represent the average of the three wells.

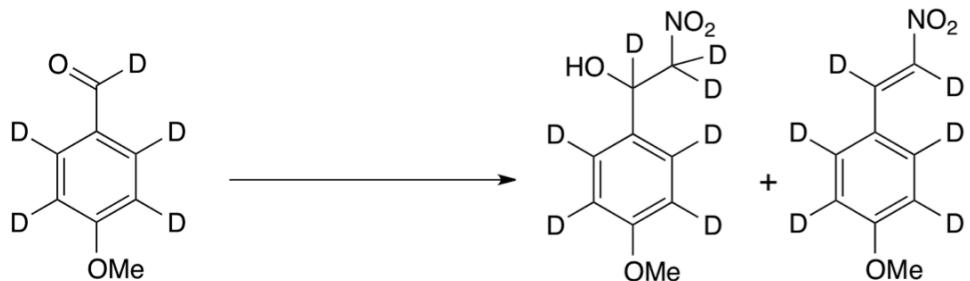
## Synthesis of the CSICT Azides<sup>5,9</sup>

### 4-methoxy-benzaldehyde-formyl,2,3,5,6-d<sub>5</sub> (15-d<sub>5</sub>)



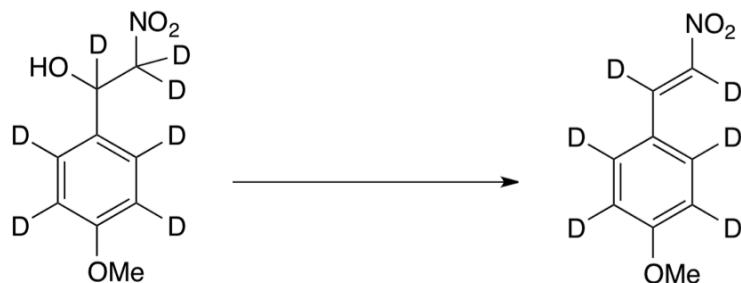
A stirred solution of N,N-dimethylformamide-d<sub>7</sub> (5.8 mL, 80.14 mmol) and anisole-2,3,4,5,6-d<sub>5</sub> (5.4 mL, 49.91 mmol, **14-d<sub>5</sub>**) was cooled to 0°C. Diphosphoryl chloride (10.3 mL, 74.49 mmol) was then added slowly over the course 30 minutes. The reaction was then heated to 105°C and stirred for 24 h. The cooled reaction mixture was then poured into a vigorously stirring bath of ice (120 g). The pH was adjusted to pH 10 through the addition of 2M NaOH (270 mL) and extracted with DCM (6 x 100 mL). The combined organic solutions were then dried with Mg<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:10) yielding two light yellow oils, *ortho* (350 mg, 2.48 mmol, 5% yield) and *para* (4.8g, 34.04mmol, 68% yield, **15-d<sub>5</sub>**).

**1-(4-methoxyphenyl 2,3,5,6-d<sub>4</sub>)-2-nitroethan-d<sub>3</sub>-1-ol (16-d<sub>7</sub>)**



A stirred solution of 4-methoxy-benzaldehyde-*formyl*,2,3,5,6-d<sub>5</sub> (4.7 g, 33.3 mmol, **15-d<sub>5</sub>**), nitromethane-d<sub>3</sub> (5.4 mL, 49.91 mmol) and methan(ol-d) (68 mL) was cooled to -10°C . Once chilled, a solution of KOH in methan(ol-d) (20 mL at 1.8 M) was added drop wise to the reaction while maintaining the temperature below 0°C. The reaction was then allowed to stir for 1.5 h at 0°C before the addition of ice cold 0.1 M HCl (140 mL), generating a yellow precipitate. The methan(ol-d) is then carefully removed *in vacuo* and the organic precipitate is dissolved and extracted with chloroform (4 x 50 mL). The combined organic solutions were then dried with Mg<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:3) yielding two yellow solids, 4-methoxy-β-nitrostyrene-d<sub>6</sub> (890 mg, 5.00 mmol, R<sub>f</sub> 0.40, 15% yield) and 1-(4-methoxyphenyl 2,3,5,6-d<sub>4</sub>)-2-nitroethan-d<sub>3</sub>-1-ol (3.4 g, 16.55 mmol, R<sub>f</sub> 0.27, 50% yield, **16-d<sub>7</sub>**). Multiple attempts were made to increase the elimination reaction including elevated temperatures, more acidic work ups and longer reaction times. However these were not fruitful. **16-d<sub>7</sub>** <sup>2</sup>H NMR (92 MHz, CH<sub>3</sub>OH): Benzyl alcohol δ 7.33 (2D), 6.92 (2D), 6.30, 4.85, 4.72 ppm.

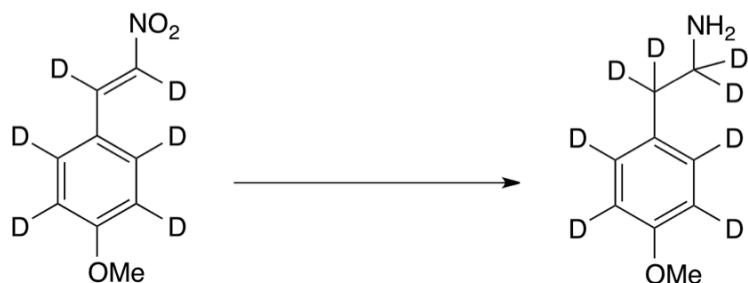
**4-methoxy- $\beta$ -nitrostyrene- $d_6$  (17- $d_6$ )**



1-(4-methoxyphenyl) 2,3,5,6-*d*<sub>4</sub>)-2-nitroethan-*d*<sub>3</sub>-1-ol (3.4 g, 16.55 mmol, **16-d<sub>7</sub>**) was solubilized in methan(ol-d) (2 mL) and concentrated *in vacuo*. This process was then repeated with chloroform-d. The deuterated starting material was then dissolved in pyridine (4ml) and cooled to 0°C, with agitation. Acetic anhydride (3.13 mL, 33.11 mmol) was added to the reaction and it was allowed to warm to RT for 30 min. The reaction was cooled back to 0°C and quenched with 1M HCl (12 mL) and extracted with diethyl ether (5 x 10 mL). The combined organic layers were washed with brine (10 mL) before being dried with Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude acetylated alcohol was dissolved in dry toluene (10 mL) and azeotroped three times to remove any residual pyridine and water. This oil was then dissolved in benzene (50 mL) with vigorous stirring. K<sub>2</sub>CO<sub>3</sub> (7.3 g, 52.96 mmol) was added and the reaction was warmed to 40°C for 6 h. The reaction was cooled and filtered with the filtrate being washed with water (2 x 10 mL) and brine (10 mL). The organic layer was then dried with MgSO<sub>4</sub> and concentrated *in vacuo* and purified by silica gel flash chromatography (EtOAc:hexanes, 1:3) yielding a yellow solid (2.4 g, 13.57 mmol, 82% yield, **17-d<sub>6</sub>**). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): Hydrogenated compound 17** δ 7.97 (d, *J* = 13.9 Hz, 1H), 7.52 (d, *J* = 13.2 Hz, 1H), 7.50 (d, *J* = 8.9 Hz, 2H), 6.95 (d, *J* = 8.9 Hz, 2H), 3.87 (s, 3H) ppm. **<sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD):** δ 164.5, 140.0, 136.5, 132.5, 124.1, 115.8, 56.1 ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>9</sub>H<sub>10</sub>NO<sub>3</sub>: 180.0661; found 180.0661. **<sup>2</sup>H NMR (92 MHz, CH<sub>3</sub>OH):**

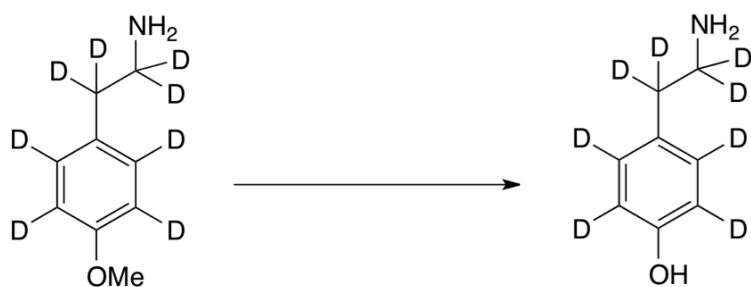
**Deuterated compound 17-d<sub>6</sub>** δ 7.99, 7.76, 7.63 (2D), 6.99 (2D) ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>9</sub>H<sub>4</sub>D<sub>6</sub>NO<sub>3</sub>: 186.1037; found 186.1034.

### 4-methoxy-phenylethylamine-d<sub>8</sub> (**18-d<sub>8</sub>**)



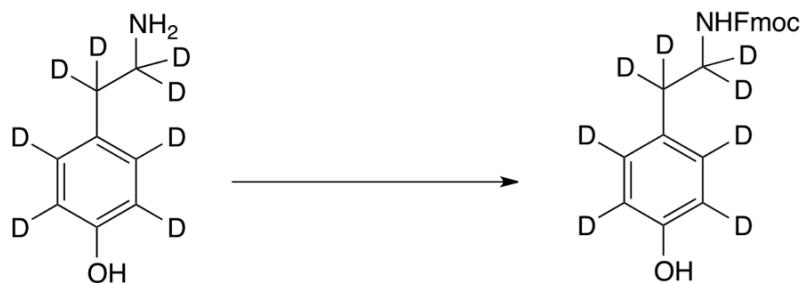
4-methoxy- $\beta$ -nitrostyrene- $d_6$  (2.1 g, 11.78 mmol, **17-d<sub>6</sub>**) was loaded in into a Soxhelet extraction thimble and placed into a Soxhelet extraction vessel, which was connected to a round bottom flask containing a stirred solution of LiAlD<sub>4</sub> (1.9g, 45.26 mmol) in diethyl ether (60 mL). The ether solution was refluxed and after 2 h the Soxhelet vessel was removed as all the styrene had been added. The reaction was maintained under reflux for a further 22 h. The flask was cooled to 0°C and ice cold water (50 mL) was carefully added drop wise to the reaction. The organic layer was decanted from the flask and the oxide slurry was filter through celite. This solid slurry was placed back in the reaction flask and washed vigorously in diethyl ether (10 x 30 mL). The combined organic solutions were then dried with Mg<sub>2</sub>SO<sub>4</sub> and added dried down *in vacuo* onto celite (13.5g) and purified via C18 reverse phase chromatography (water (0.1% acetic acid) :acetonitrile, 9:1) yielding an off yellow solid (1.8g, 11.31 mmol, 96% yield, **18-d<sub>8</sub>**). <sup>2</sup>H NMR (92 MHz, CH<sub>3</sub>OH): Deuterated compound **18-d<sub>8</sub>** δ 7.16 (2D), 6.88 (2D), 3.02 (2D), 2.78 (2D) ppm.

### 4-hydroxyphenylethylamine-d<sub>8</sub> (**19-d<sub>8</sub>**)



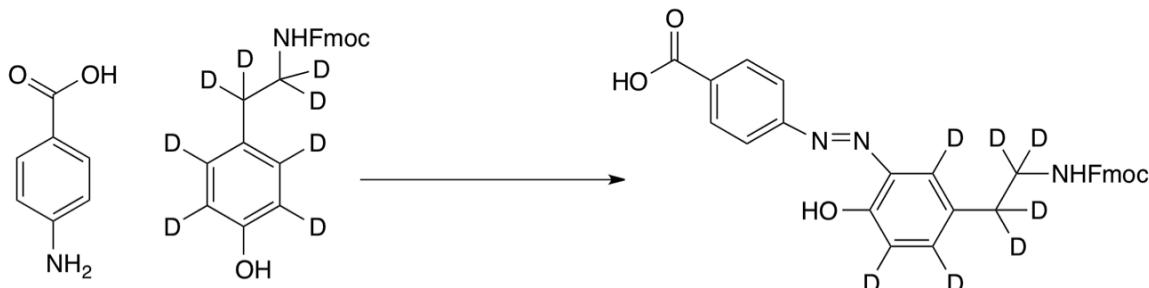
4-methoxy-phenylethylamine-d<sub>8</sub> (1.0g, 6.28 mmol, **18-d<sub>8</sub>**) was solubilized in a flask with methan(ol-d) (5mL) and concentrated in vacuo, this step was then repeated and the amine was stored under argon. In a separate flask argon was lightly bubbled through a solution of DBr (4 mL, 47 wt. % in D<sub>2</sub>O )to the amine and the reaction was stirred and warmed to 100°C for 2 h. The reaction was cooled and concentrated in vacuo. The resulting yellow spate was extracted with deutereated methanol (3 x 5 mL). Th combined washes were concerntrated onto celite (5g) before being purified via C18 reverse phase chromatography (water (0.1% acetic acid):acetonitrile, 18:1) to yield a yellow red solid which was recrystallized from methanol with diethyl ether to give a pale yellowish solid, which was deemed to to be the acetate salt of the amine (666 mg, 4.59 mmol, 73% yield, **19-d<sub>8</sub>**). **<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD): Hydrogenated compound 19 δ** 7.02 (d, *J* = 8.5 Hz, 2H), 6.71 (d, *J* = 8.52 Hz, 2H), 2.81 (t, *J* = 7.5 Hz, 2H), 2.65 (d, *J* = 7.5 Hz, 2H), ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>8</sub>H<sub>12</sub>NO: 138.0919; found 138.0922. **<sup>2</sup>H NMR (92 MHz, CH<sub>3</sub>OH): Deuterated compound 19-d<sub>8</sub> δ** 7.20 (2D), 6.93 (2D), 3.05 (2D), 2.82 (2D) ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>8</sub>H<sub>4</sub>D<sub>8</sub>NO: 146.1422; found XXX.XXXX.

### N-fluorenyloxycarbonyl-4-hydroxyphenylethylamine-d<sub>8</sub> (**20-d<sub>8</sub>**)



The acetate salt of 4-hydroxyphenylethylamine-d<sub>8</sub> (92 mg, 0.44 mmol, **19-d<sub>8</sub>**) was dissolved in D<sub>2</sub>O (2 mL). To this a solution of DBr in D<sub>2</sub>O (170 μL, 3.5 mM) was added and stirred for 5 min before being concentrated in vacuo. The bromide was then dissolved in a mixed solution of THF (6 mL), DMF (2 mL) and triethylamine (98 μL, 1.32 mmol) before being stirred for 15 min at room temp. A freshly prepared solution of 9-fluorenylmethoxycarbonyl chloride (455 mg, 1.76 mmol) in THF (1.5 mL) was then added to the amine and reaction was allowed to stir for 1h before being concentrated in vacuo, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:1) yielding a pale solid (74 mg, 0.21 mmol, 46% yield, **20-d<sub>8</sub>**). **1H NMR (600 MHz, (CD3)2CO):** Deuterated compound **20-d<sub>8</sub>** δ 7.86 (d, J = 7.9 Hz, 2H), 7.68 (d, J= 7.34 Hz, 2H), 7.41 (t, J = 7.6 Hz, 2H), 7.32 (t, J = 7.9 Hz, 2H), 6.46 (s, 1H), 4.32 (d, J = 7.1 Hz, 2H), 4.21 (t, J = 7.6 Hz, 1H), ppm. **2H NMR (92 MHz, (CH3)2CO):** δ 7.06 (2D), 6.78 (2D), 3.27 (2D), 2.66 (2D) ppm.

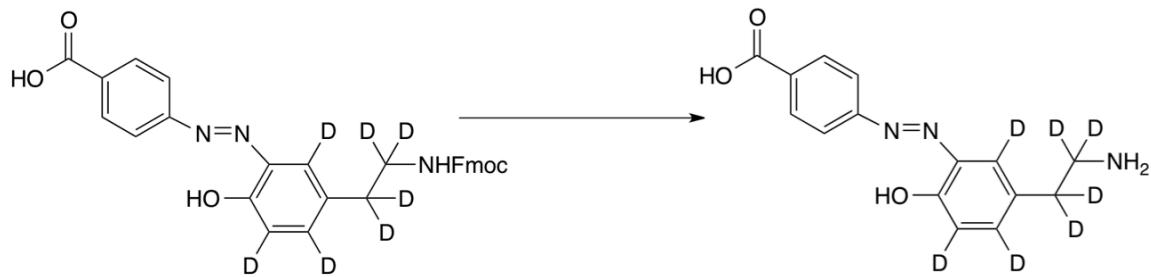
**4-[5-(N-fluorenyloxycarbonyl-2-Amino-ethyl-d<sub>4</sub>)-2-hydroxy-phenylazo-d<sub>3</sub>]-benzoic acid (21-d<sub>7</sub>)**



4-[5-(N-fluorenyloxycarbonyl-2-Amino-ethyl-d<sub>4</sub>)-2-hydroxy-phenylazo-d<sub>3</sub>]-benzoic acid **21-d<sub>7</sub>** was prepared from 4-carboxy-benzenediazonium chloride. Benzoic acid (390 mg, 2.84 mmol) was dissolved with stirring in 6 M HCl (6 mL) and cooled to 0°C. A 2.5 M of NaNO<sub>2</sub> in H<sub>2</sub>O (1.47 mL) was added to the acid and the temperature was maintained at 0°C for 30 min. In another flask a 795 mM solution of K<sub>2</sub>CO<sub>3</sub> (50 mL) was added to a solution of deuterated Fmoc tyramine (520 mg, 1.42 mmol, **20-d<sub>8</sub>**) in THF (15 mL) with vigorous stirring, this was then cooled to 0°C. The 6M solution of HCl and benzoic acid was then added very slowly to the basic solution of deuterated Fmoc tyramine while being sure to maintain the reaction temperature at 0°C with a basic pH. After the acid has been added the reaction is allowed to warm up to RT and stir for 16 h. Ethyl acetate (50 mL) was then added to the reaction and the pH was adjusted to pH 3 with 2M HCl. The organic layer was separated and the aqueous was washed with ethyl acetate (2 x 50 mL). The combined organic solutions were then dried with Mg<sub>2</sub>SO<sub>4</sub>, concentrated in vacuo, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:1) yielding a bright orange solid (470 mg, 910 µmol, 64% yield). **1H NMR (600 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): Hydrogenated compound 21 δ** 13.19 (bs, 1H), 10.93 (bs, 1H), 8.11 (d, J = 8.4 Hz, 2H), 8.02 (d, J = 8.4 Hz, 2H), 7.86 (d, J = 7.5 Hz, 2H), 7.64 (d, J = 7.6 Hz, 2H), 7.60 (s, 1H), 7.38 (t, J = 7.5 Hz, 2H), 7.31-7.27 (m, 3H), 7.01 (d, J = 8.4 Hz, 1H), 4.28 (d, J = 7.0 Hz, 2H), 4.18 (t, J = 7.0 Hz, 1H), 3.25 (q, J = 6.5 Hz, 2H), 2.73 (t, J = 6.5 Hz,

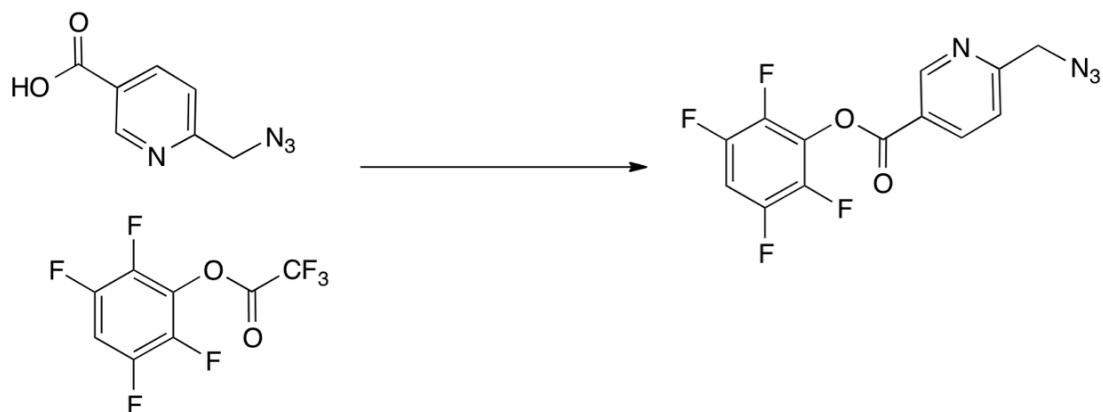
1H) ppm. **13C NMR** (151 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): δ 166.5, 155.7, 153.6, 154.3, 143.5, 141.3, 141.1, 138.7, 134.6, 132.1, 127.1, 121.3, 118.3, 126.2, 124.7, 122.2, 121.4, 119.8, 64.7, 46.4, 41.3, 33.9 ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>30</sub>H<sub>26</sub>N<sub>3</sub>O<sub>5</sub>: 508.1872; found 508.1869. **2H NMR** (92 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): **Deuterated compound 21-d<sub>7</sub>**, δ 7.94-6.68 (m, 3D), 3.15 (s, 2D), 2.65 (s, 2D) ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>30</sub>H<sub>19</sub>D<sub>7</sub>N<sub>3</sub>O<sub>5</sub>: 515.2312; found 515.2308.

**4-[5-(2-aminoethyl-d<sub>4</sub>)-2-hydroxy-phenylazo-d<sub>3</sub>]benzoic acid (22-d<sub>7</sub>)**



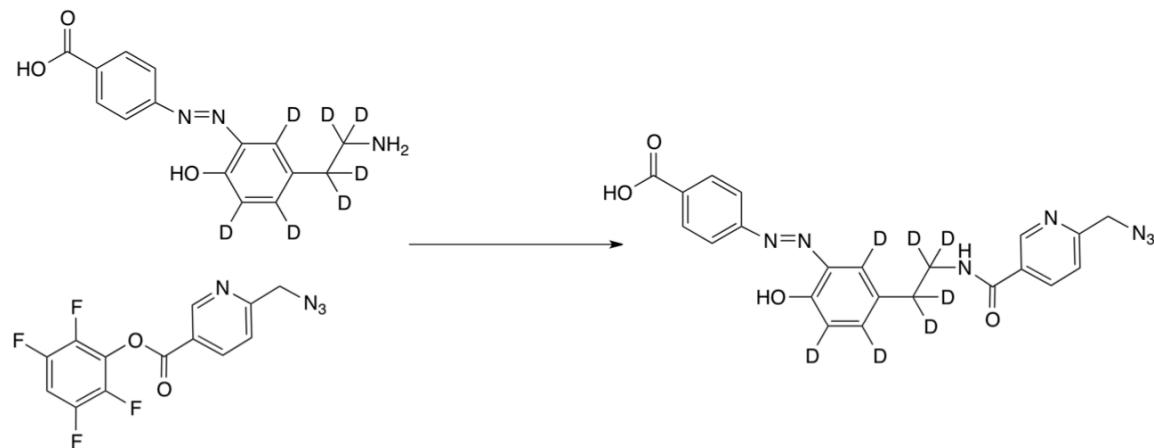
**21-d<sub>7</sub>** (520 mg, 1.01 mmol) was dissolved in a solution of 20% pipridine in DMF (5 mL) and the reaction was stirred for 1 h before being loaded onto celite (5 g), *in vacuo* and purified via C18 reverse phase chromatography (water (0.1% acetic acid):acetonitrile, 1:1) yielding a bright orange solid (253 mg, 869  $\mu$ mol, 79%). **<sup>1</sup>H NMR (600 MHz, (CD<sub>3</sub>OD, D<sub>2</sub>O, CD<sub>3</sub>COOD): Deuterated compound 22-d<sub>7</sub>,  $\delta$  8.16 (d,  $J$  = 8.4 Hz, 2H), 7.97 (d,  $J$  = 8.5 Hz, 2H) ppm.** **<sup>2</sup>H NMR (92 MHz, (CH<sub>3</sub>OH, H<sub>2</sub>O, CH<sub>3</sub>COOH)):**  $\delta$  7.86 (1D), 7.36 (1D), 7.05 (1D), 3.18 (2D), 2.95 (2D) ppm.

### 2,3,5,6-tetrafluorophenyl 5-(azidomethyl)nicotinate (23)



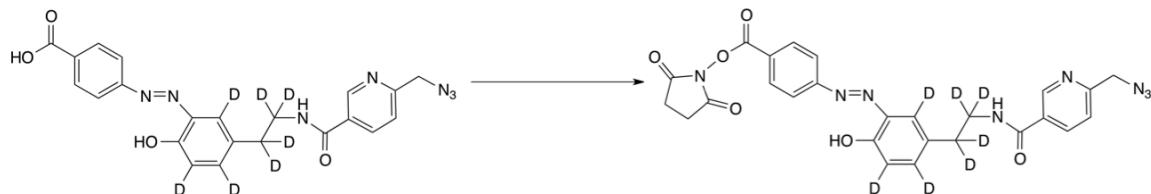
N,N-Diisopropylethylamine (152  $\mu$ L, 875  $\mu$ mol) was added to a stirred solution of 6-azidomethylnicotinic acid<sup>10</sup> (228mg, 795  $\mu$ mol) in DMF (4 mL). The solution was then cooled to 0°C before the addition of acetic acid, trifluoro-, 2,3,5,6-tetrafluorophenyl ester (190  $\mu$ L, 1.19 mmol). The reaction was then allowed to warm to RT and stir for 1 h before being concentrated in vacuo, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:3) yielding a pale yellow solid (218 mg, 675  $\mu$ mol, 85% yield, **23**). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)**:  $\delta$  9.36 (d, J = 2.3 Hz, 1H), 8.50 (dd, J = 8.3, 2.3 Hz, 1H), 7.59 (d, J = 8.3 Hz, 1H), 7.12-7.05 (m, 1H), 4.65 (s, 1H) ppm.

**4-[5-(2-(6-(azidomethyl)nicotinamido) ethyl-d4)-2-hydroxy-phenylazo-d3]-benzoic acid (24-d<sub>7</sub>).**



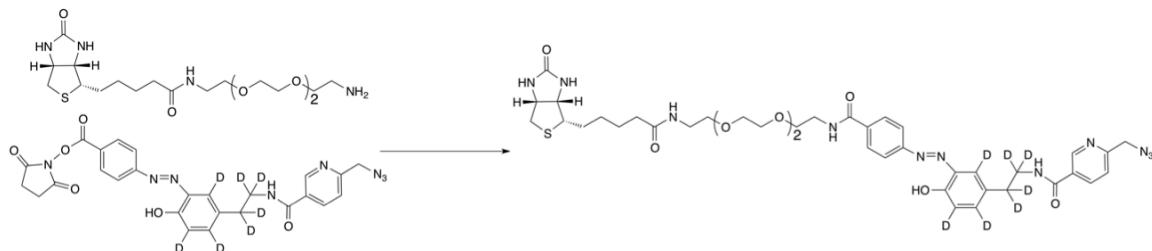
Under argon, N,N-Diisopropylethylamine (111  $\mu$ L, 673  $\mu$ mol) was added to a stirred solution of amine (160mg, 561  $\mu$ mol, **23-d<sub>7</sub>**) in DMF (5 mL) and 1,4-dioxane (2 mL). Azide (**23**) was then added to the reaction and it allowed to stir for 2 h before being concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:1) bright orange solid (203 mg, 449  $\mu$ mol, 80% yield, **24-d<sub>7</sub>**). **<sup>2</sup>H NMR (92 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): Deuterated compound 24-d<sub>7</sub> δ** 7.50 (bs, 3D), 3.47 (s, 2D), 2.87 (s, 2D) ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>22</sub>H<sub>13</sub>D<sub>7</sub>N<sub>7</sub>O<sub>4</sub>: 453.2016; found 453.2006.

**4-(2,5-dioxopyrrolidin-1-yl)-[5-(2-(6-(azidomethyl)nicotinamido)hydroxy-phenylazo-d<sub>3</sub>]-benzoate (25-d<sub>7</sub>)      ethyl-d<sub>4</sub>-2-**



Under argon, the benzoic acid (46 mg, 100 µmol, **24-d<sub>7</sub>**) was dissolved in THF (4 mL), stirred and cooled to 0°C. *N,N'*-dicyclohexylcarbodiimide (23 mg, 110 µmol) was then added to the solution and the reaction was allowed to warm to RT. Once at RT, 1-Hydroxy-2,5-pyrrolidinedione (13 mg, 113 µmol) was added to the reaction and it was stirred 2 h before **being** concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:2) bright orange solid (46 mg, 84 µmol, 85% yield, **25-d<sub>7</sub>**). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): Deuterated compound 25-d<sub>7</sub> δ 12.54 (s, 1H), 8.88 (d, J = 2.0 Hz, 1H), 8.28 (d, J = 8.4 Hz, 2H), 8.28 (dd, J = 8.3, 2.0 Hz, 1H), 7.96 (d, J = 8.4 Hz, 2H), 7.42 (d, J = 8.3 Hz, 1H), 6.28 (s, 1H), 4.53 (s, 2H), 2.93 (s, 4H) ppm.**

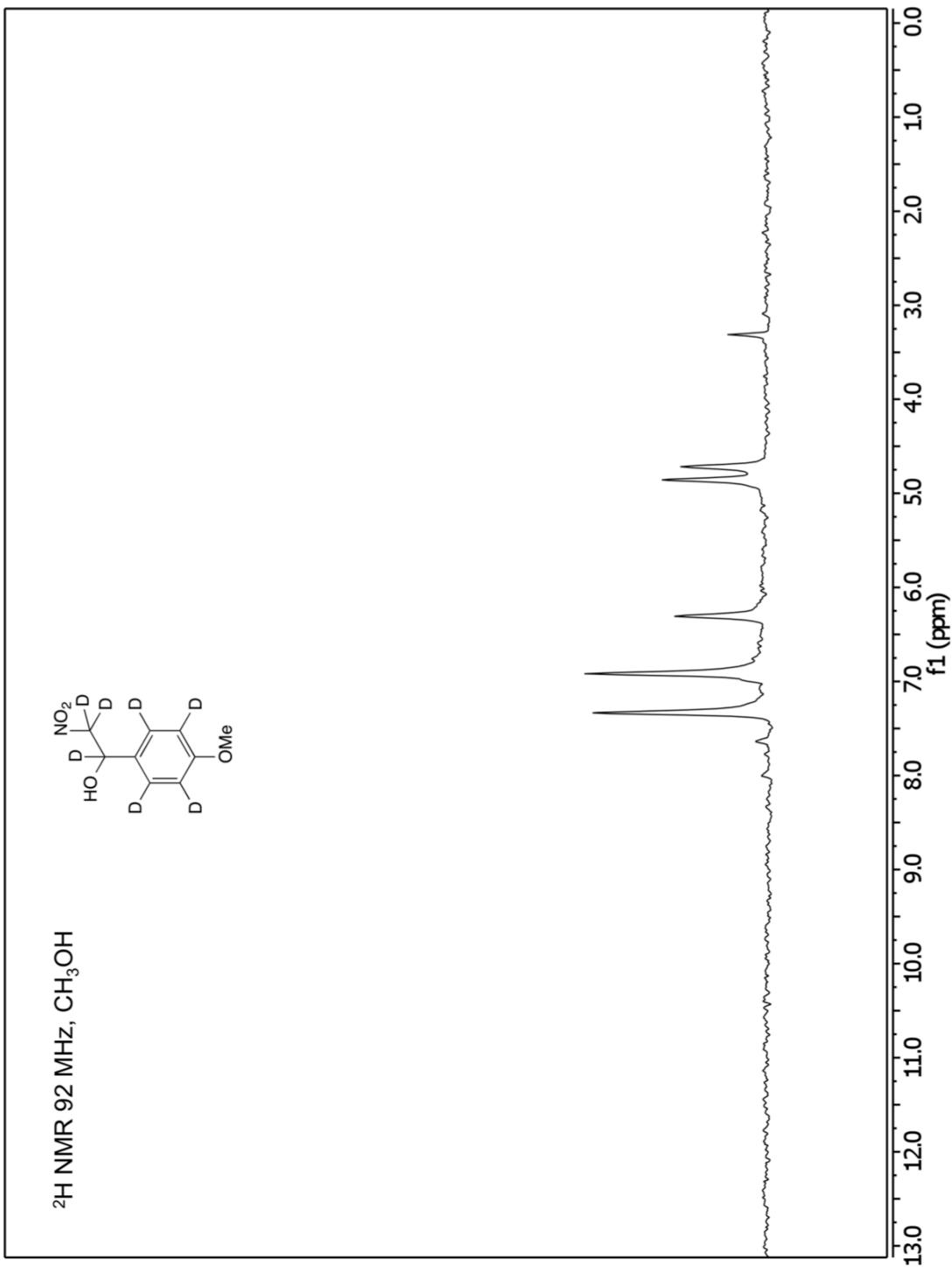
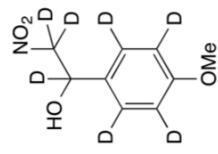
**4-[5-(2-(azidomethyl)nicotinamido) ethyl-d4]-2-hydroxy-phenylazo-d3-N-(15-oxo-19-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-4,7,10-trioxa-14-azanonadecyl) benzamide (11-d<sub>7</sub>)**



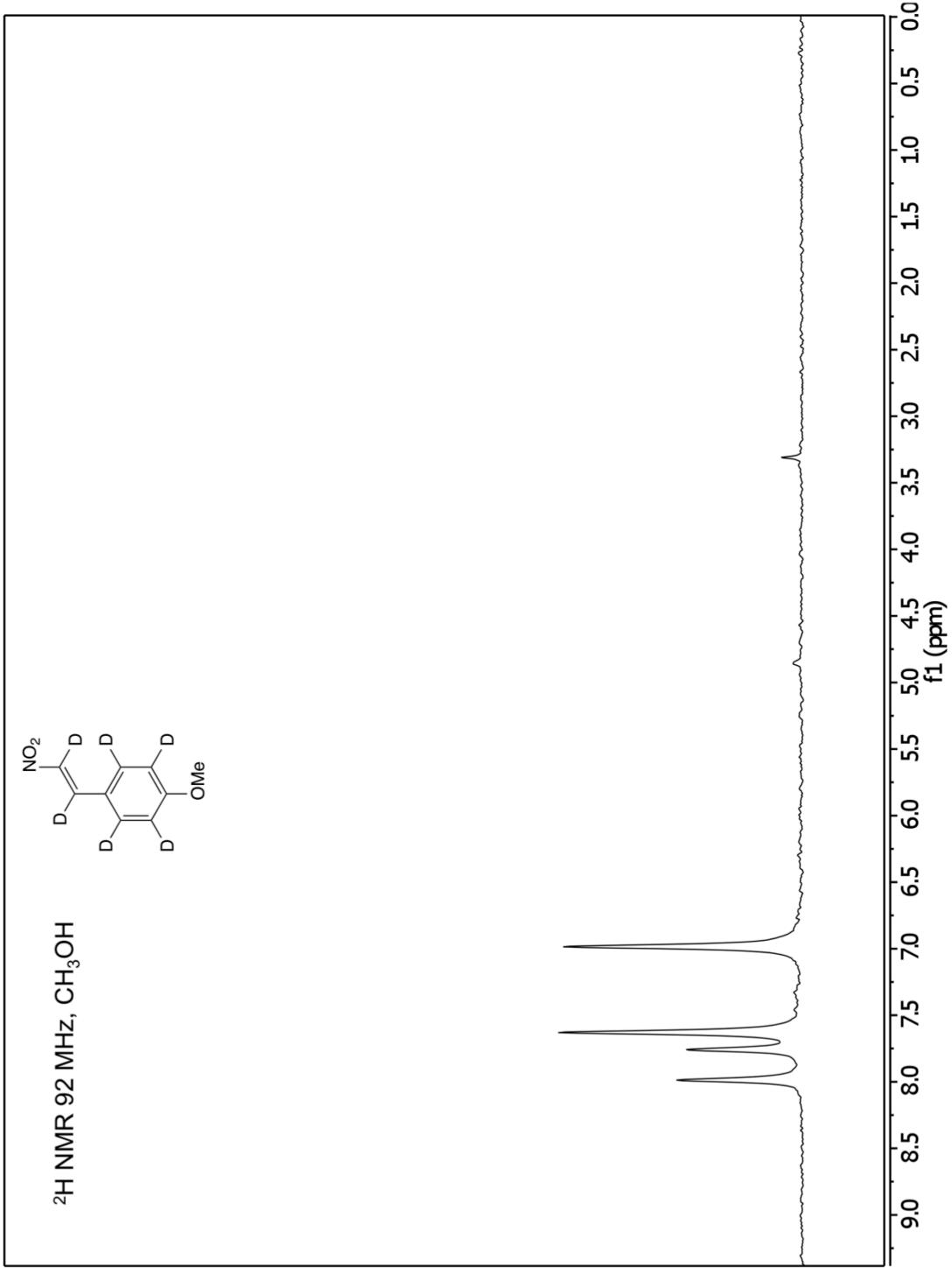
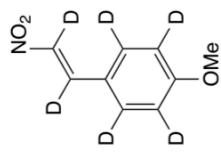
A stirred solution of the NHS ester (18 mg, 33 μmol, **25-d<sub>7</sub>**) in DMF (2 mL) was prepared under argon in a small vial. In a separate vial, a solution of the biotin amine<sup>5</sup> (13 mg, 33 μmol) in DMF (500 μL) was made. The biotin solution was then added to the reaction vial and it was stirred for 16 h. The reaction was dried down in vacuo onto celite (1g) before being purified via C18 reverse phase chromatography (water (0.1% acetic acid):acetonitrile, 18:1) to yield a bright orange solid (19 mg, 24 μmol, 73% yield, **11-d<sub>7</sub>**) **<sup>1</sup>H NMR (600 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): Hydrogenated compound 11 δ** 10.85 (bs, 1H), 8.96 (d, *J* = 2.3 Hz, 1H), 8.78 (t, *J* = 5.6 Hz, 1H), 8.69 (t, *J* = 5.6 Hz, 1H), 8.17 (dd, *J* = 8.1, 2.3 Hz, 1H), 8.04 (d, *J* = 9.0 Hz 2H), 8.02 (d, *J* = 9.0 Hz, 2H), 7.81 (t, *J* = 5.8 Hz, 1H), 7.63 (d, *J* = 2.1 Hz, 1H), 7.51 (d, *J* = 8.1 Hz, 1H), 7.34 (dd, *J* = 8.4, 2.3 Hz, 1H), 7.02 (d, *J* = 8.4 Hz, 1H), 6.40 (s, 1H), 6.34 (s, 1H), 4.57 (s, 2H), 4.30-4.27 (m, 1H), 4.12-4.09 (m, 1H), 3.58-3.55 (m, 4H), 3.54-3.49 (m, 4H), 3.45 (q, *J* = 5.6 Hz, 2H), 3.40 (t, *J* = 5.8 Hz, 2H), 3.18 (q, *J* = 5.8 Hz, 2H), 3.09-3.05 (m, 1H), 2.85 (t, *J* = 7.1 Hz, 2H), 2.80 (dd, *J* = 12.5, 5.1 Hz, 1H), 2.56 (d, *J* = 12.5 Hz, 1H), 2.05 (t, *J* = 7.5, 2H) 1.63-1.56 (m, 1H), 1.53-1.41 (m, 3H), 1.33-1.22 (m, 2H) ppm. **<sup>13</sup>C NMR (151 MHz, (CD<sub>3</sub>)<sub>2</sub>SO): δ** 171.8, 165.2, 164.1, 162.3, 157.8, 153.5, 152.7, 147.8, 144.7, 138.0, 135.7, 134.5, 130.4, 129.0, 128.0, 122.1, 1, 121.1, 69.2, 68.9, 68.7, 68.5, 65.9, 60.7, 58.9, 55.0, 53.7, 53.7, 40.45, 40.1, 39.5 (2C), 39.0, 38.1, 34.8, 33.6, 27.9, 27.7, 24.9 ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>38</sub>H<sub>48</sub>N<sub>11</sub>O<sub>7</sub>S: 802.3459; found 802.3458. **<sup>2</sup>H NMR (92 MHz,**

**(CD<sub>3</sub>)<sub>2</sub>SO): Deuterated compound 11-d<sub>7</sub>** δ 8.03-6.68 (m, 3D), 3.41 (s, 2D), 2.80 (s, 2D) ppm. ESI-MS (*m/z*): [M+H<sup>+</sup>] calcd. for C<sub>38</sub>H<sub>41</sub>D<sub>7</sub>N<sub>11</sub>O<sub>7</sub>S: 809.3098; found 809.3902.

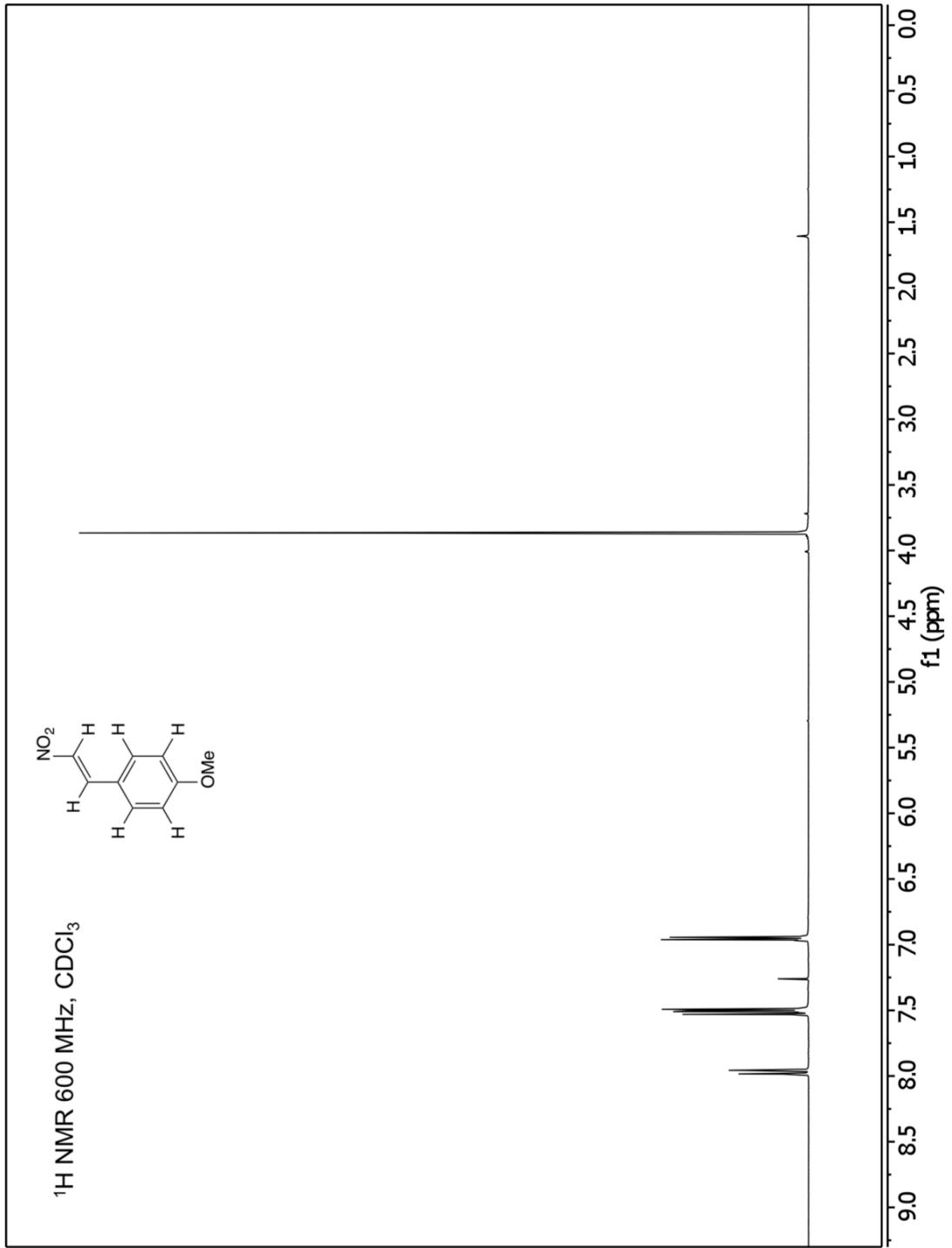
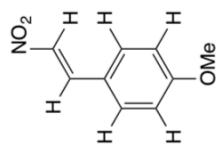
$^2\text{H}$  NMR 92 MHz,  $\text{CH}_3\text{OH}$



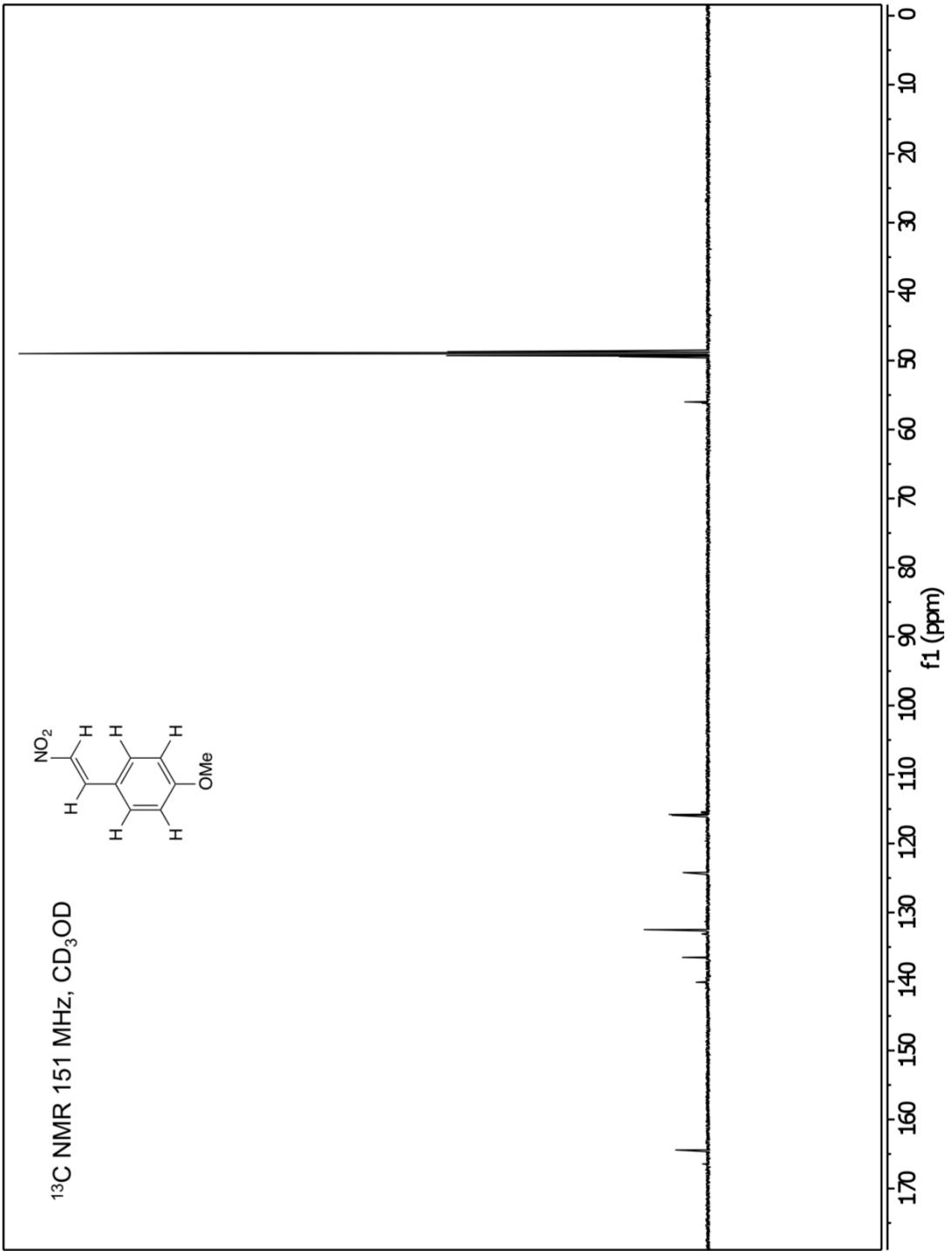
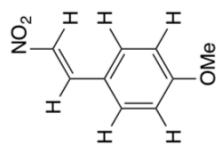
$^2\text{H}$  NMR 92 MHz,  $\text{CH}_3\text{OH}$



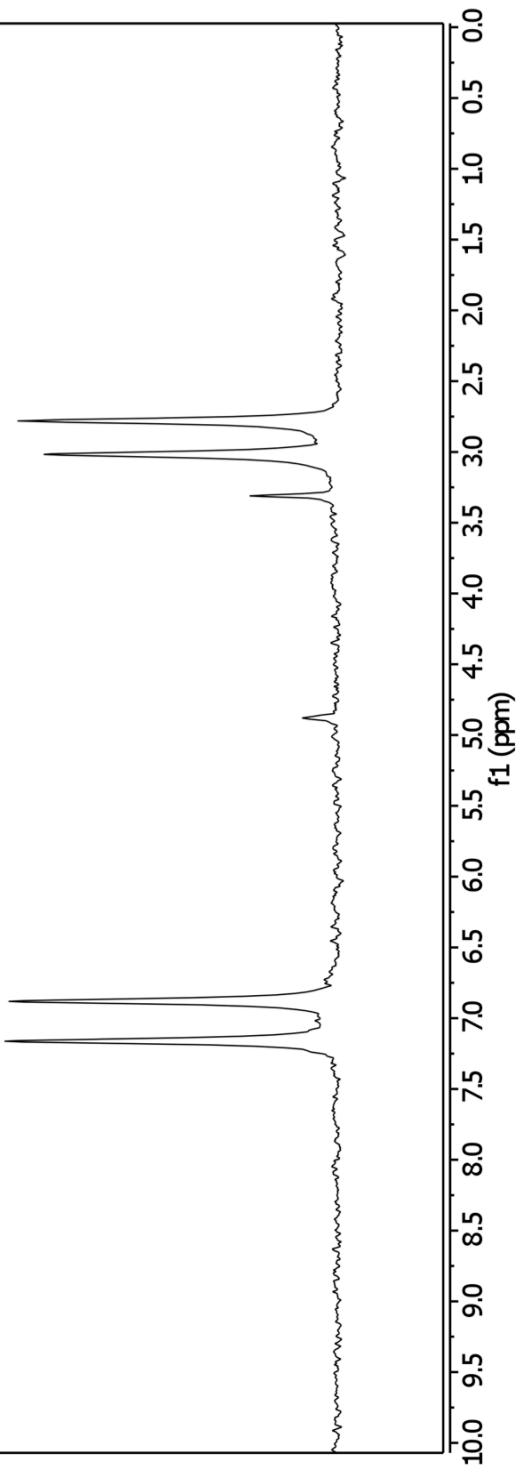
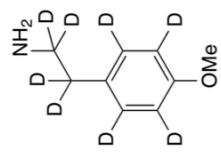
$^1\text{H}$  NMR 600 MHz,  $\text{CDCl}_3$



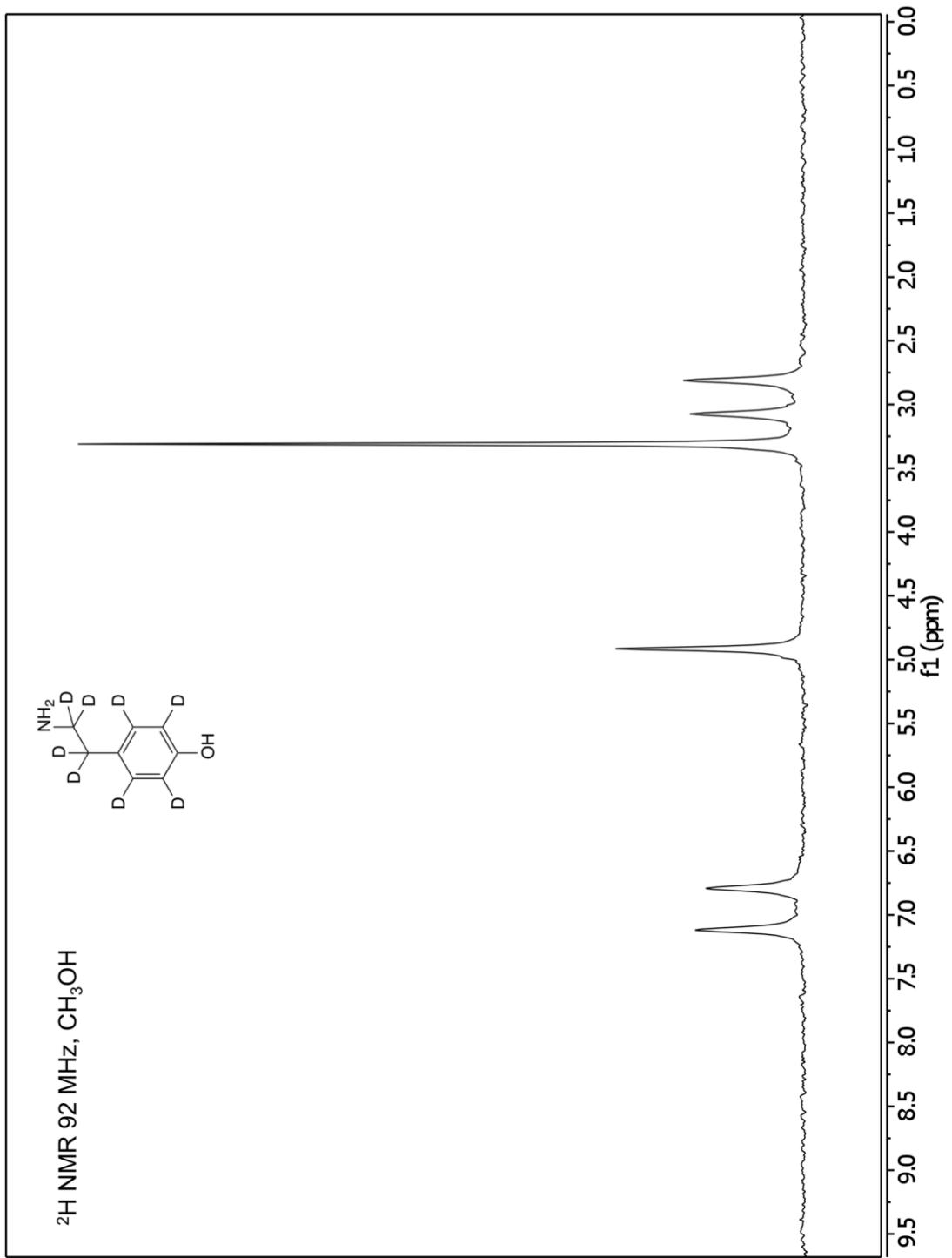
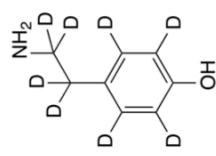
$^{13}\text{C}$  NMR 151 MHz,  $\text{CD}_3\text{OD}$

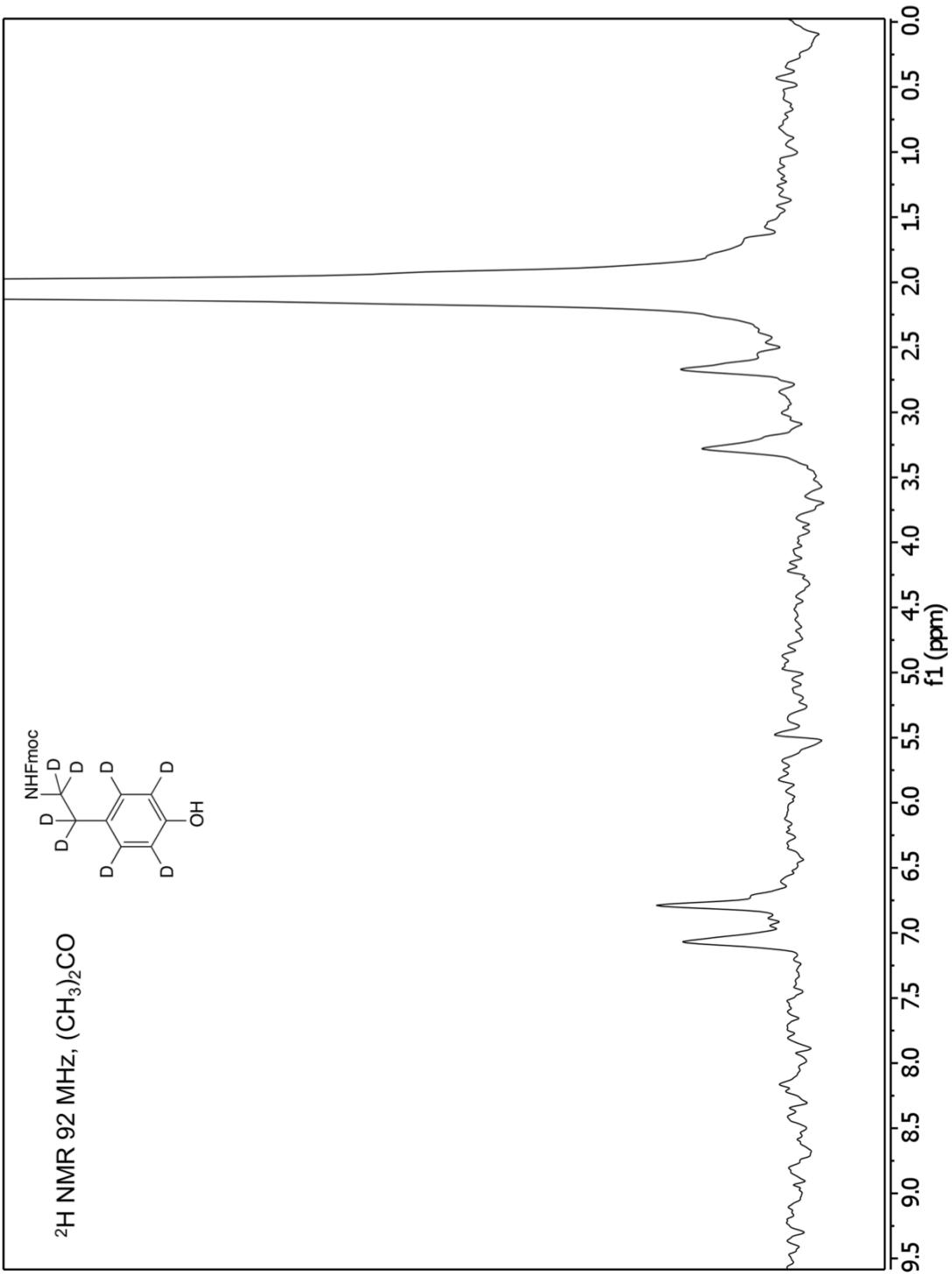


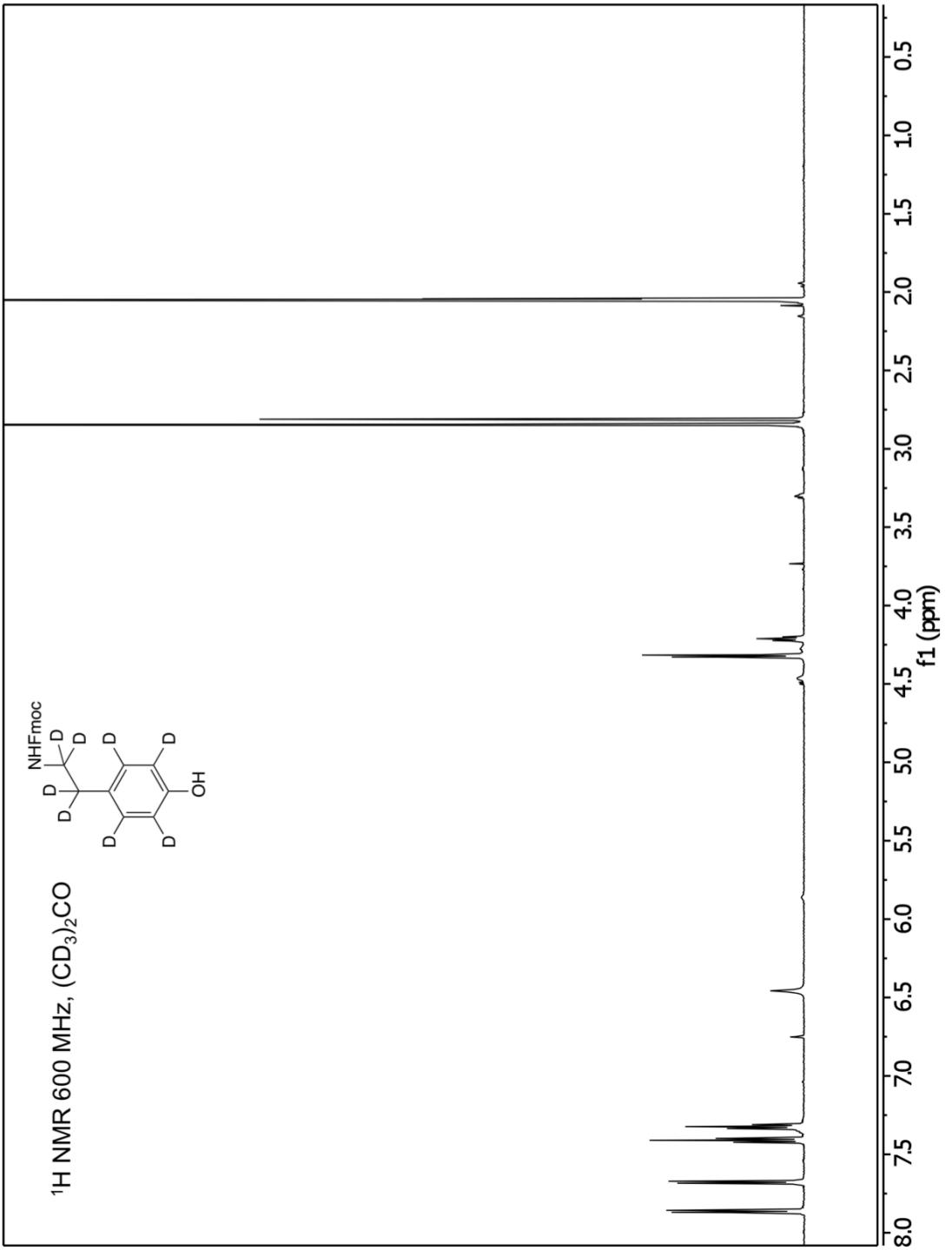
$^2\text{H}$  NMR 92 MHz,  $\text{CH}_3\text{OH}$



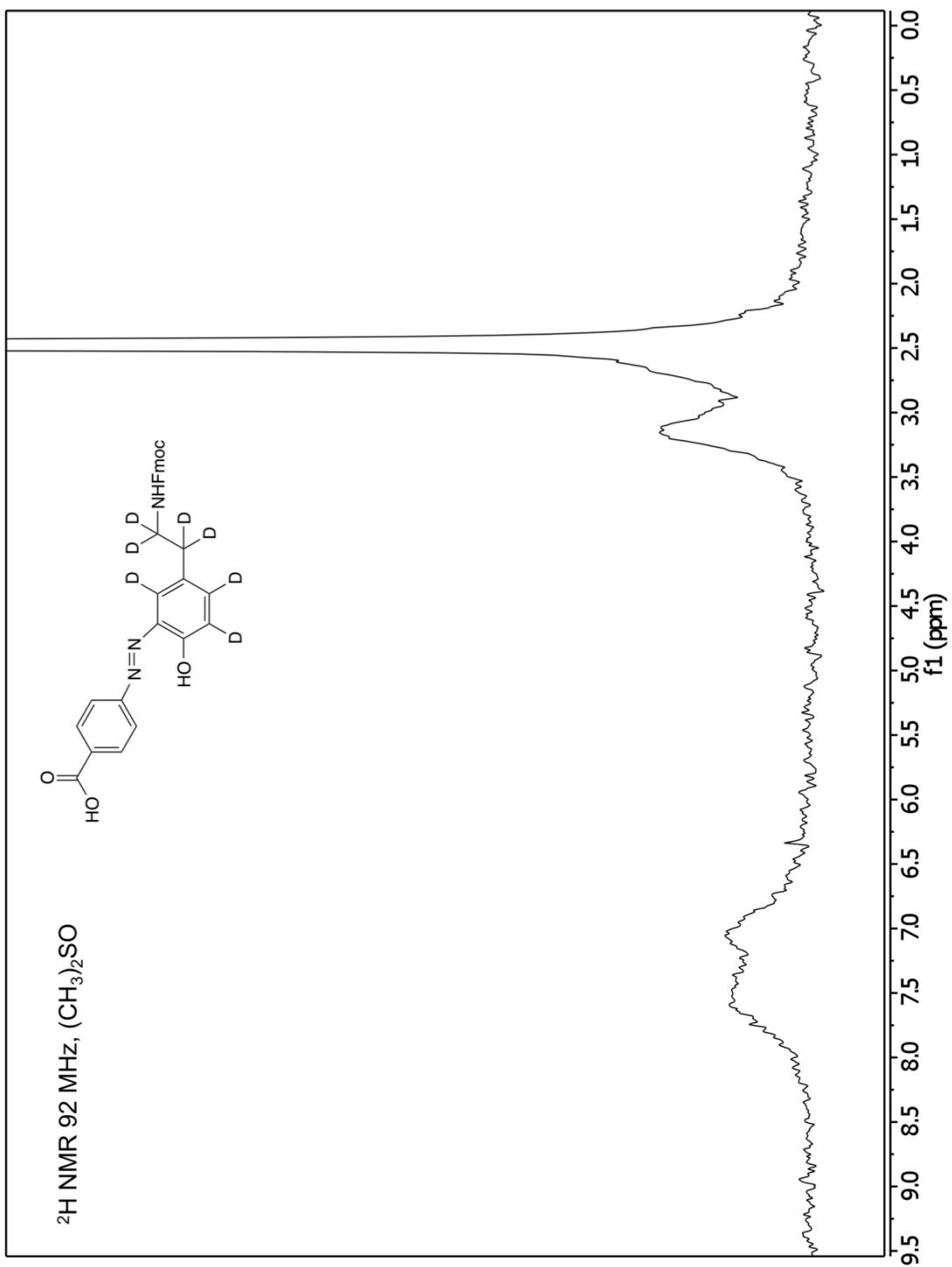
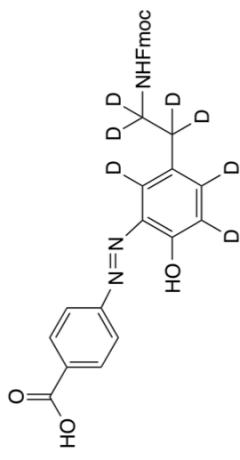
$^2\text{H}$  NMR 92 MHz,  $\text{CH}_3\text{OH}$



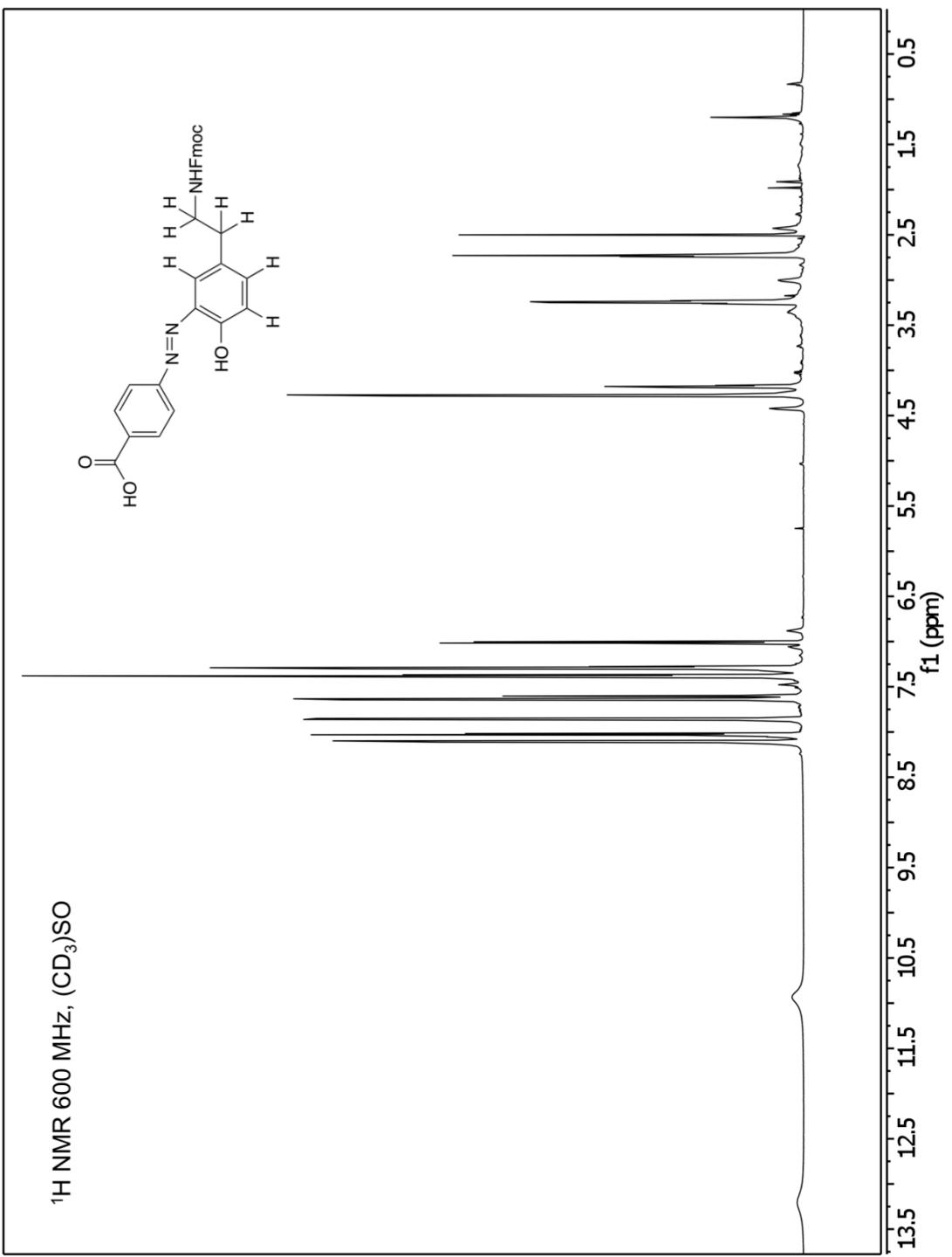
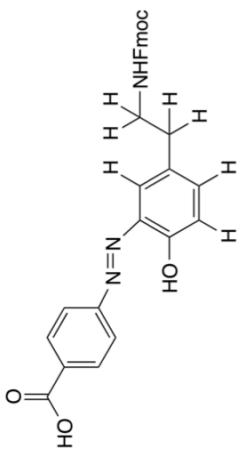


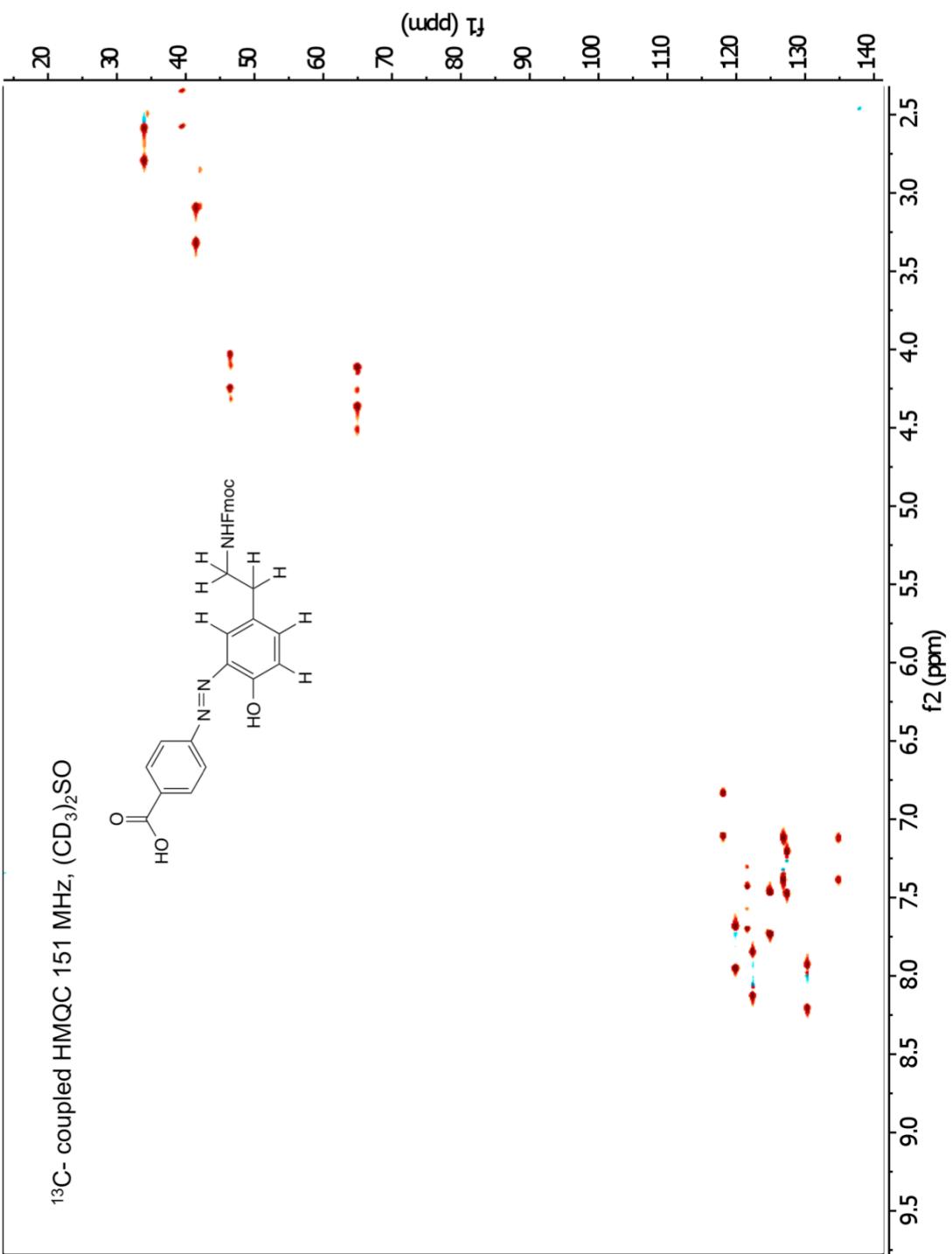


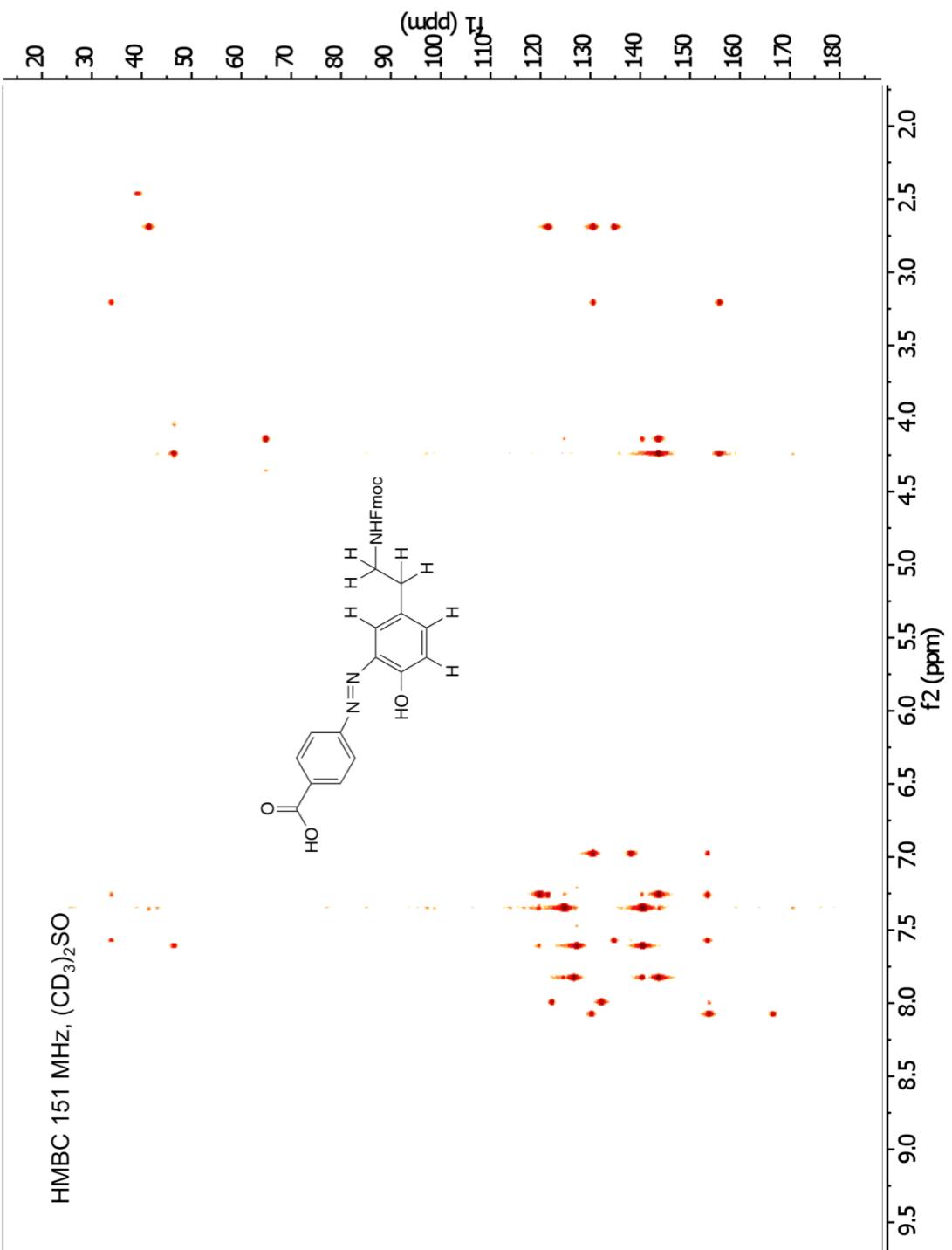
<sup>2</sup>H NMR 92 MHz, (CH<sub>3</sub>)<sub>2</sub>SO



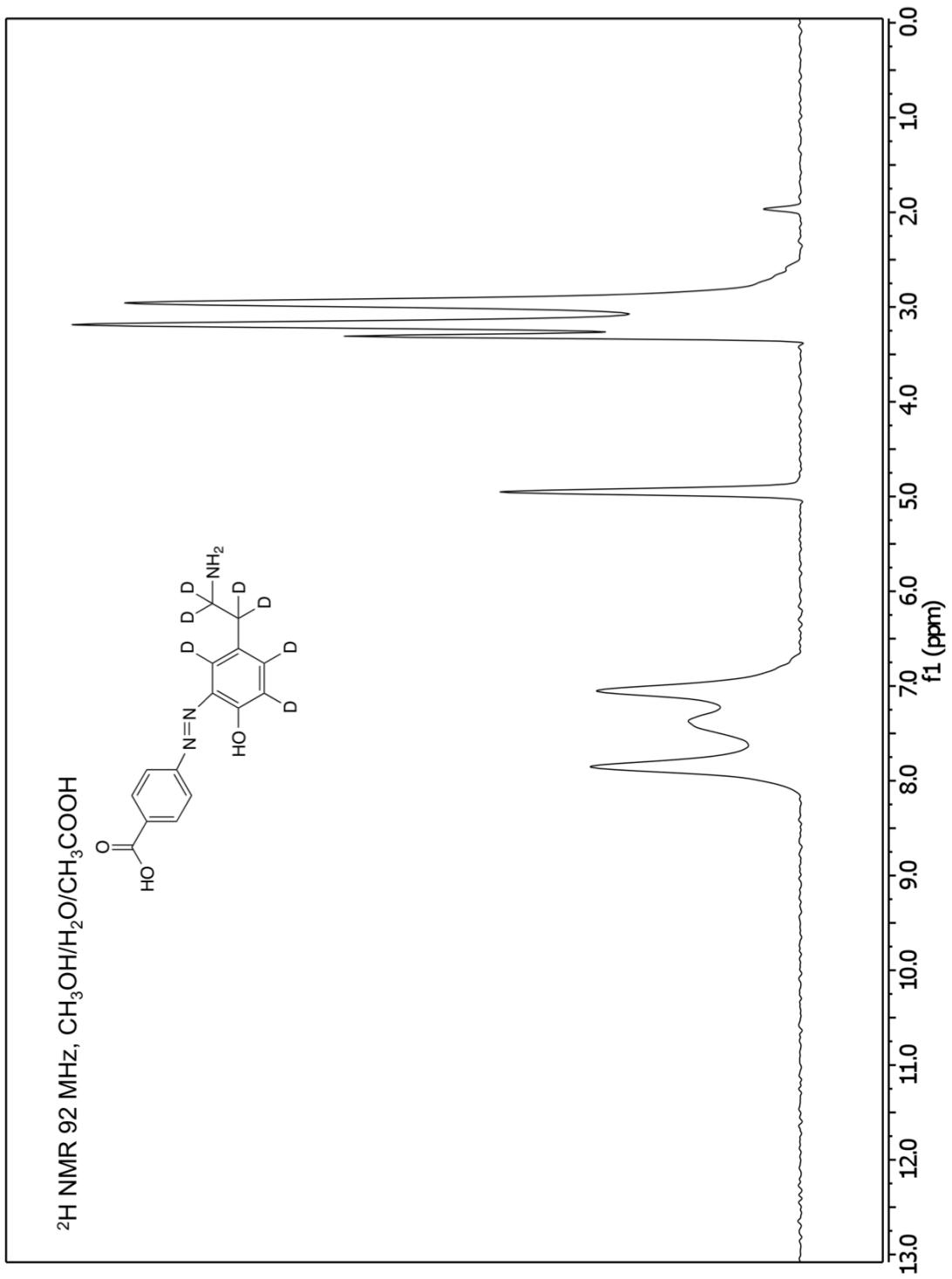
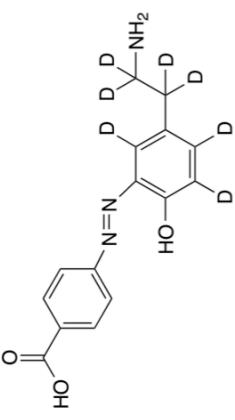
$^1\text{H}$  NMR 600 MHz,  $(\text{CD}_3)\text{SO}$



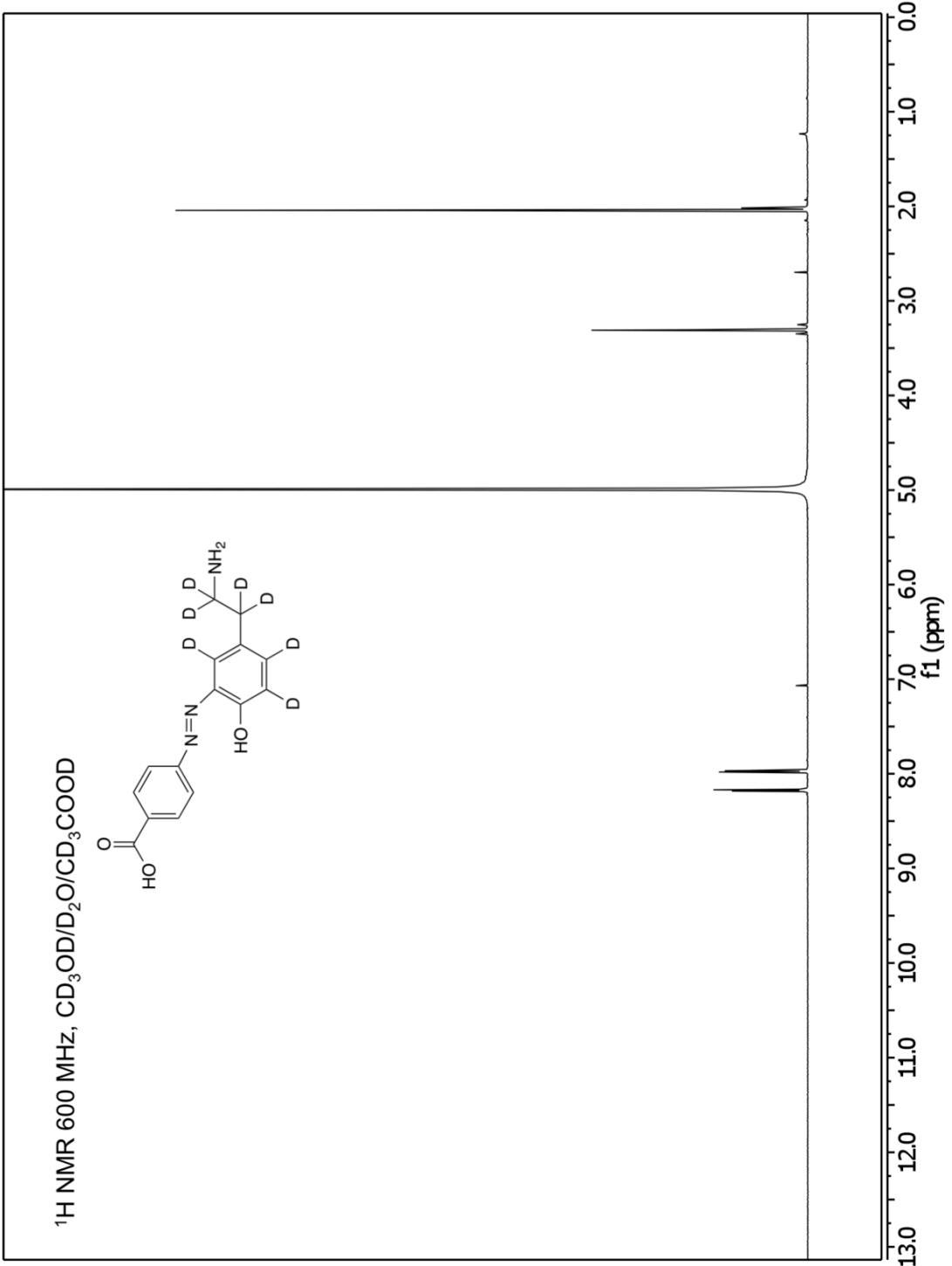
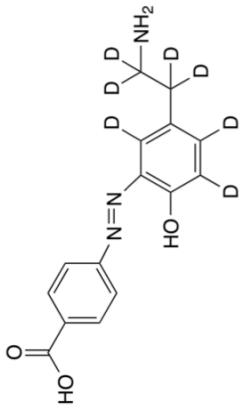




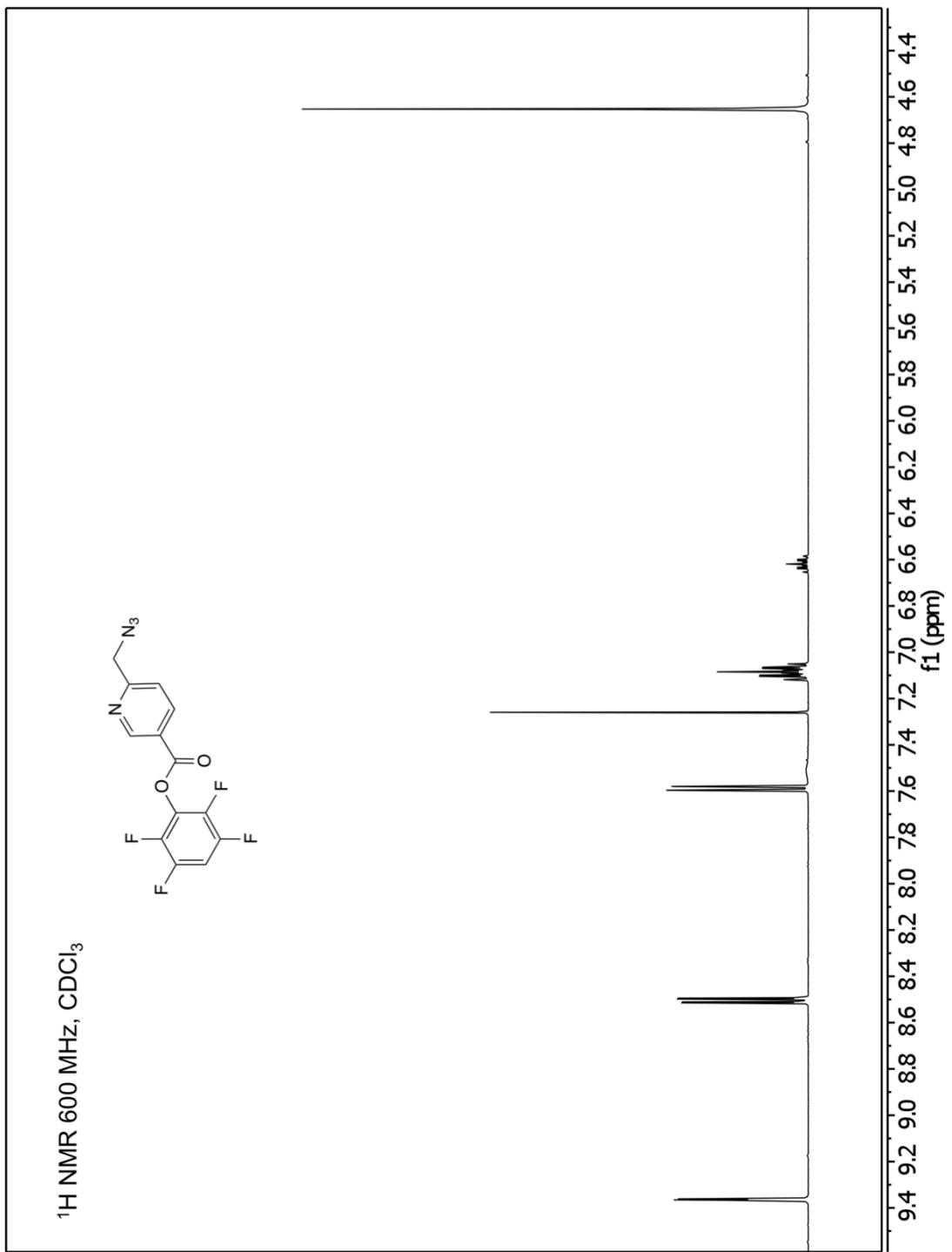
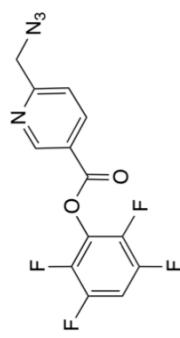
$^2\text{H}$  NMR 92 MHz,  $\text{CH}_3\text{OH}/\text{H}_2\text{O}/\text{CH}_3\text{COOH}$



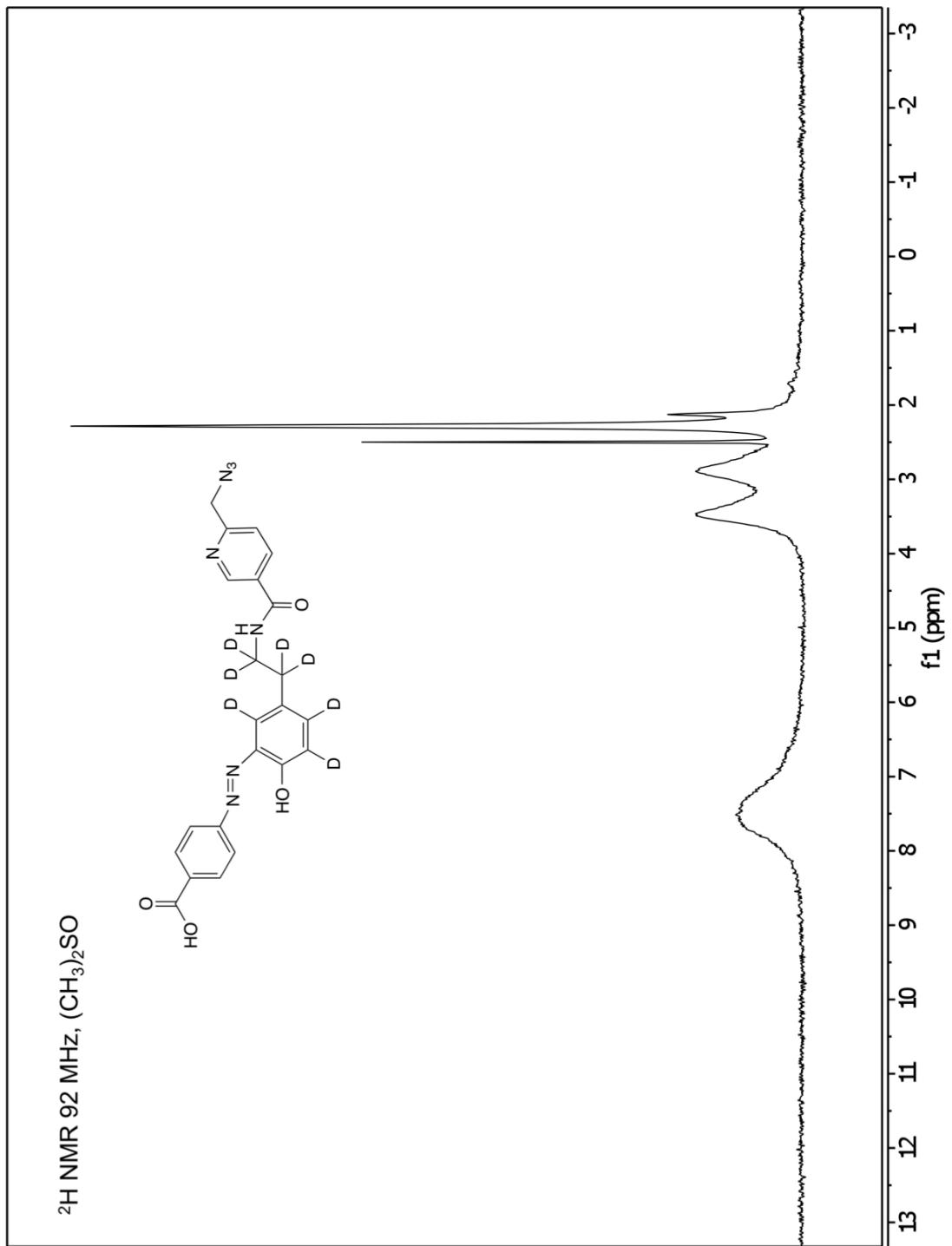
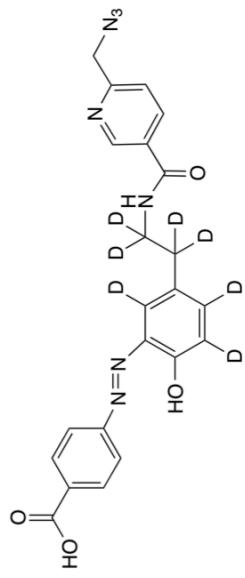
<sup>1</sup>H NMR 600 MHz, CD<sub>3</sub>OD/D<sub>2</sub>O/CD<sub>3</sub>COOD



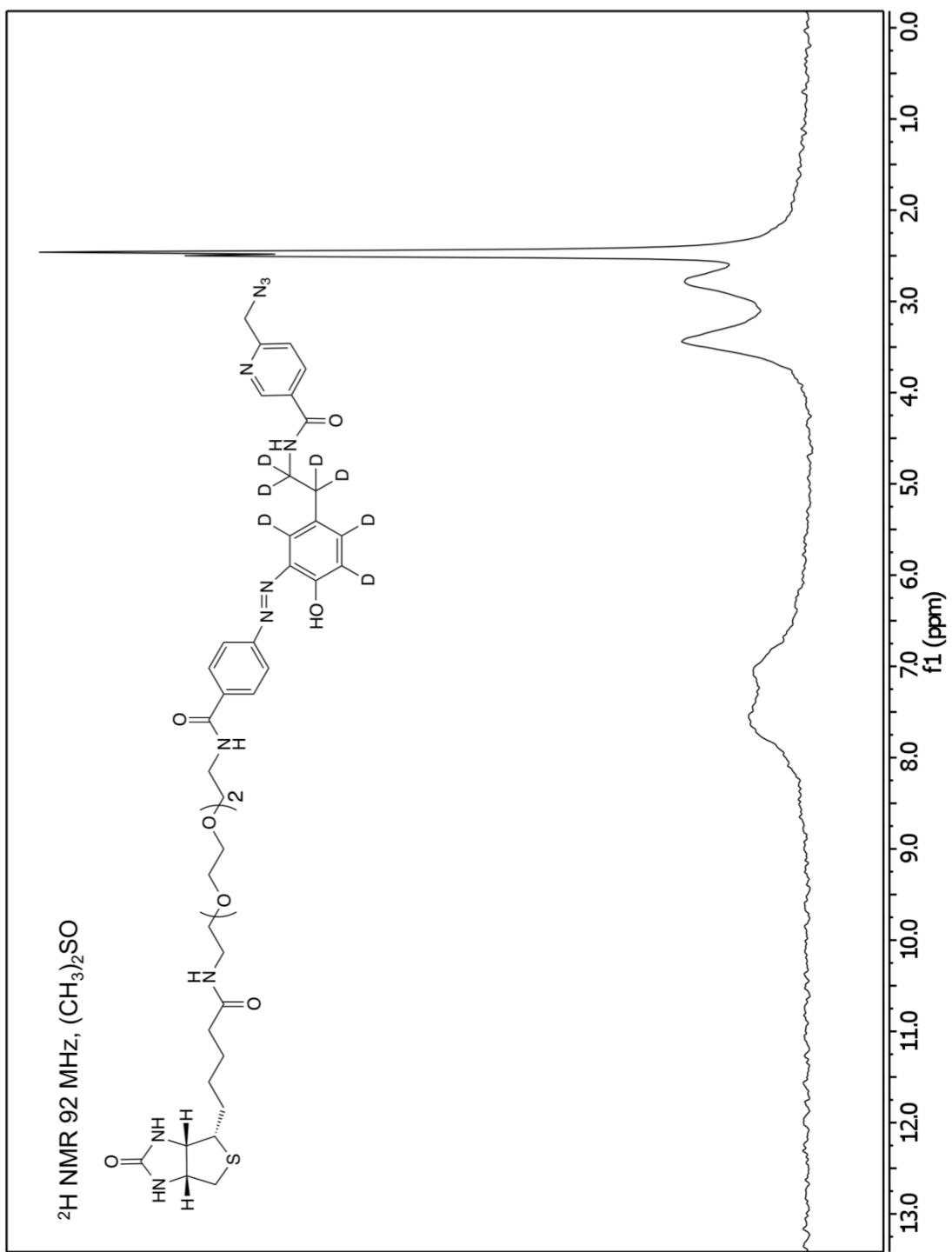
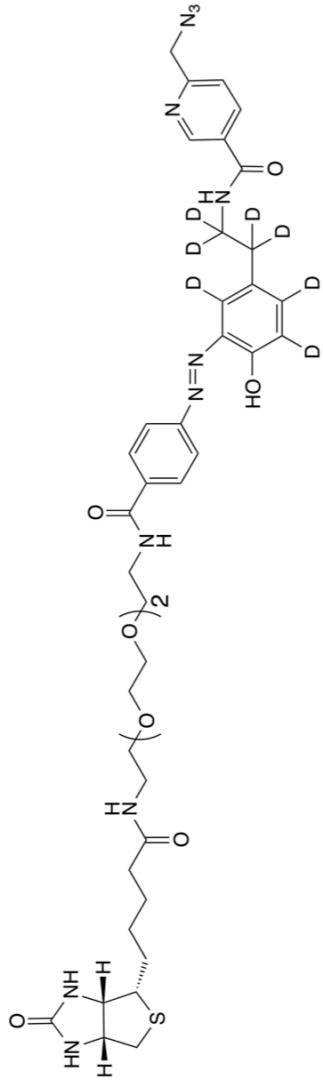
$^1\text{H}$  NMR 600 MHz,  $\text{CDCl}_3$

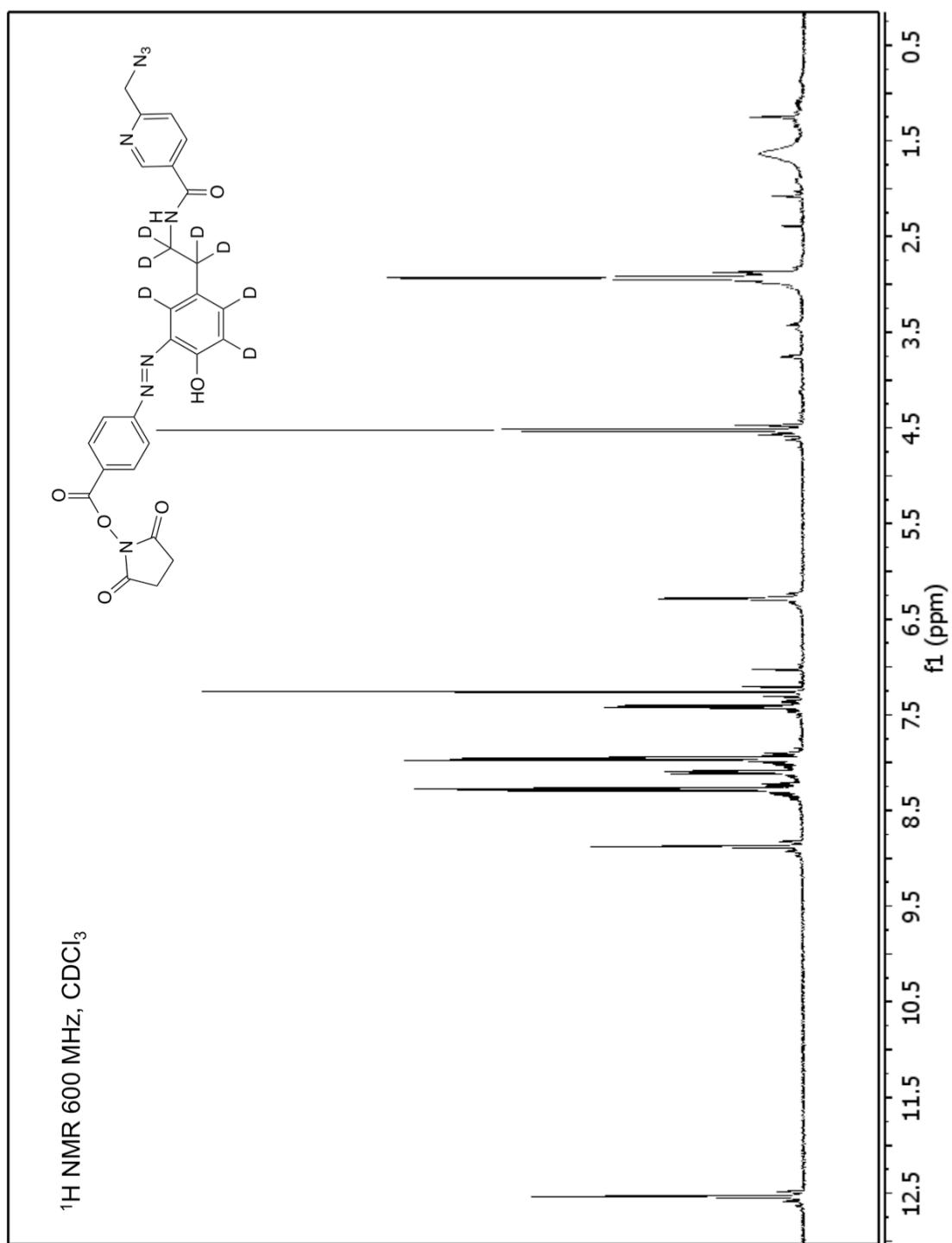


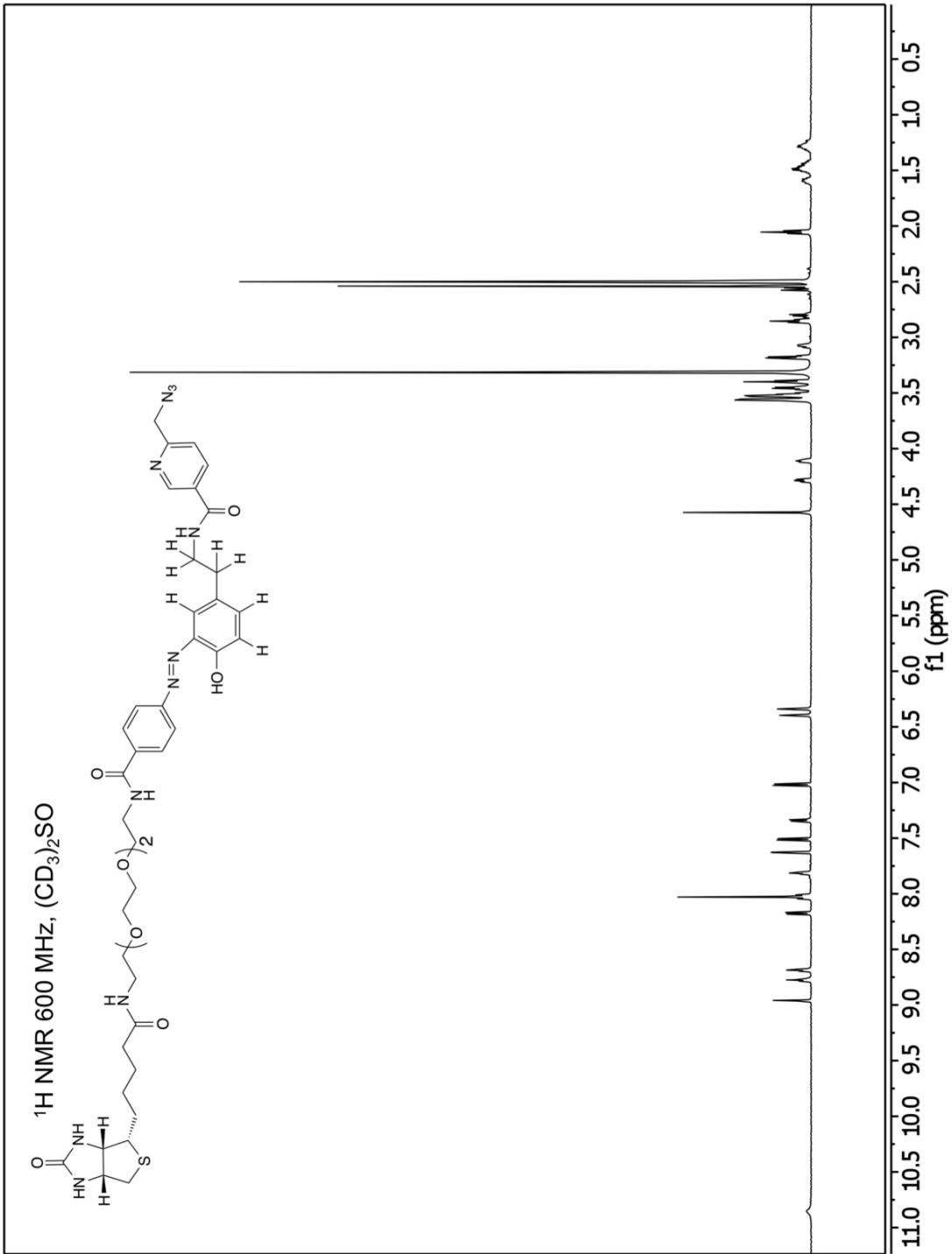
$^2\text{H}$  NMR 92 MHz,  $(\text{CH}_3)_2\text{SO}$

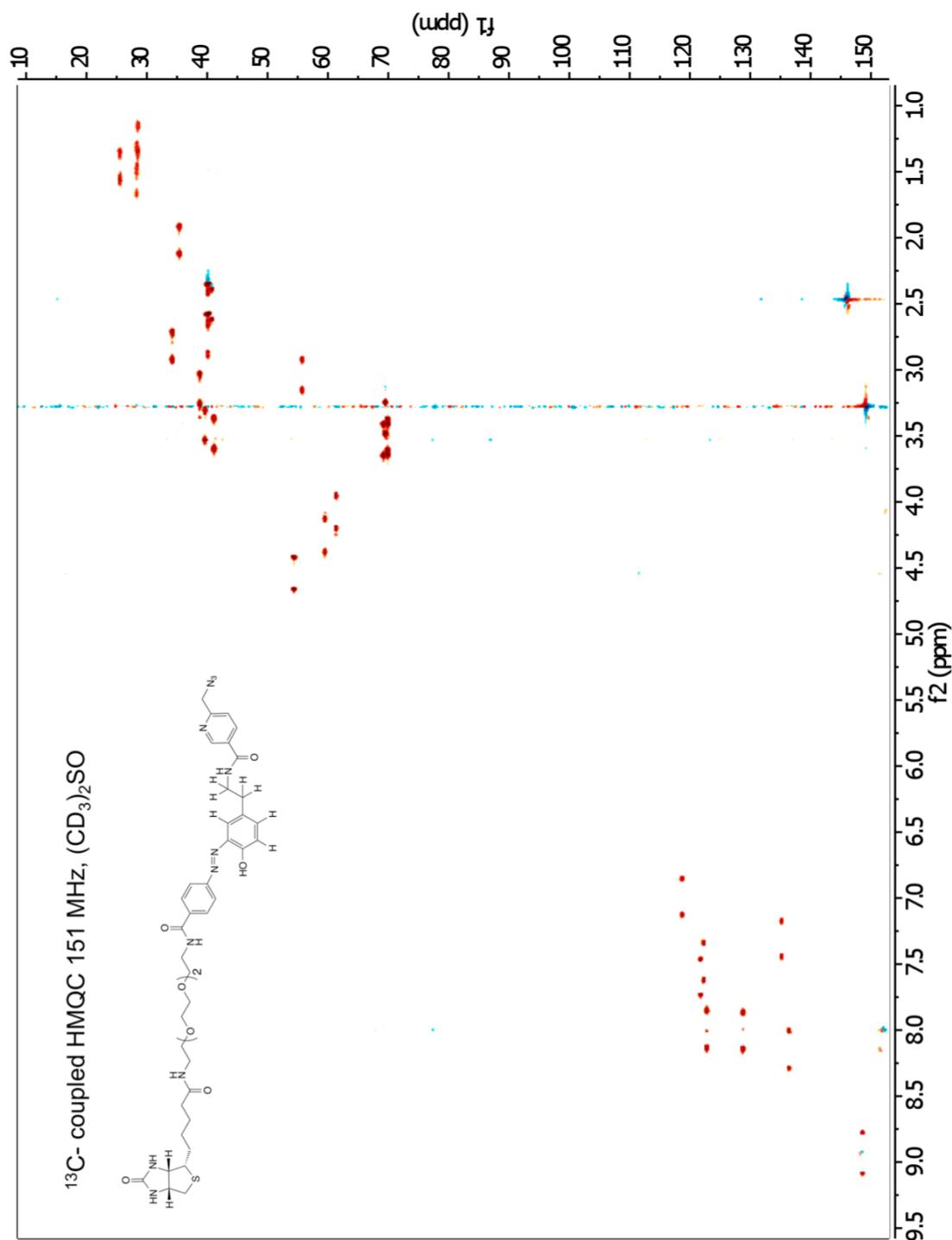


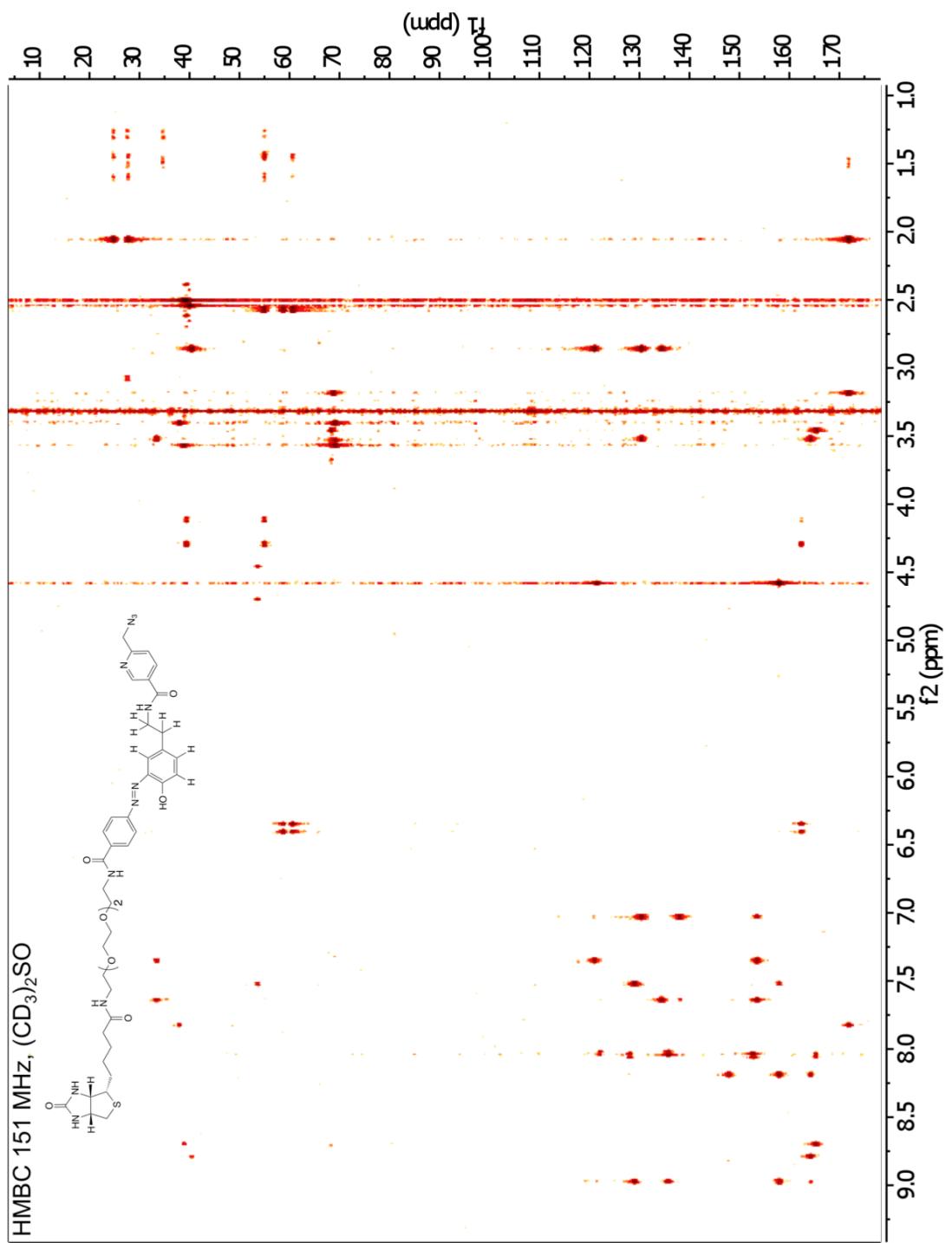
<sup>2</sup>H NMR 92 MHz, (CH<sub>3</sub>)<sub>2</sub>SO











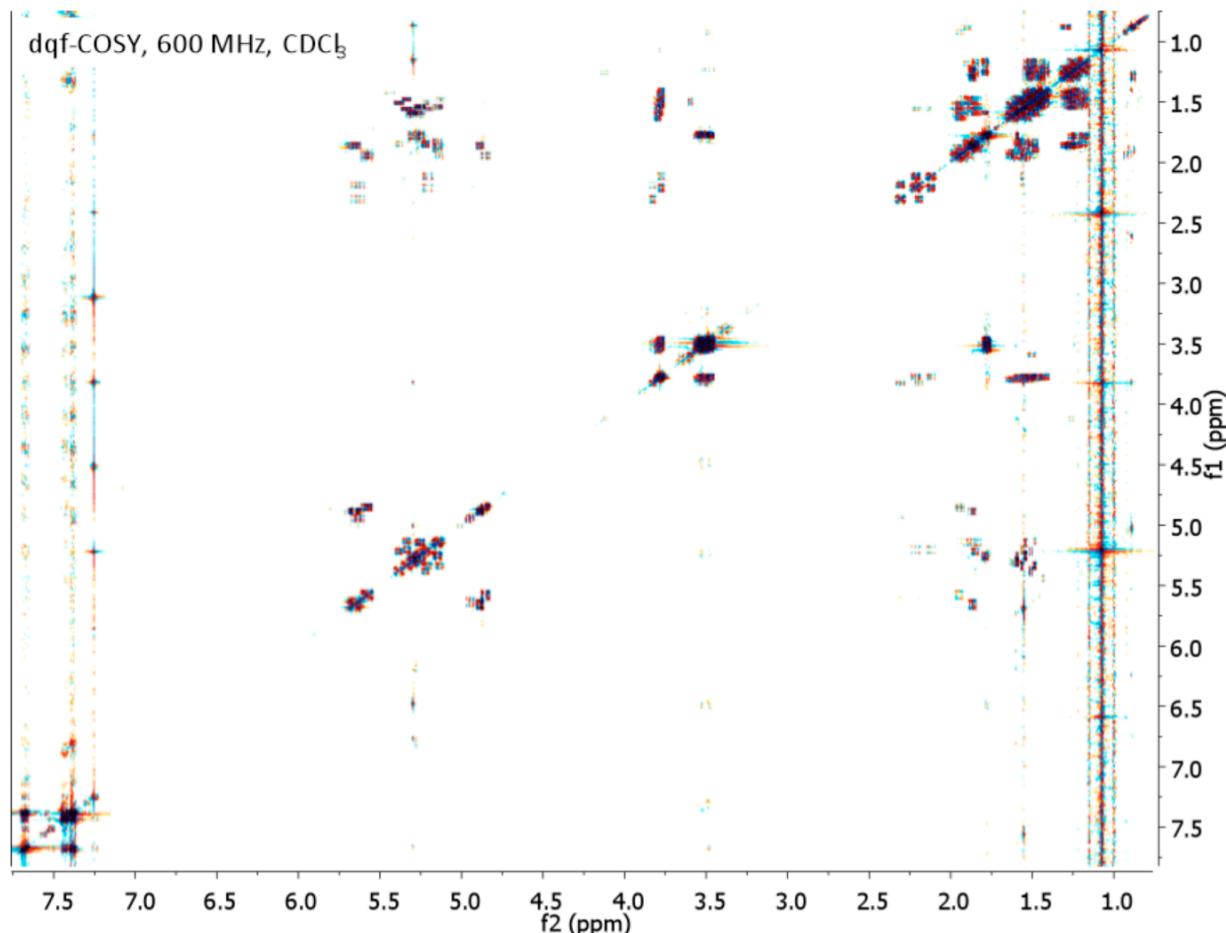
## REFERENCES

- (1) Watzl, C. *Current Protocols in Protein Science*.
- (2) Park, D.; O'Doherty, I.; Somvanshi, R. K.; Bethke, A.; Schroeder, F. C.; Kumar, U.; Riddle, D. L. *pnas.org*.
- (3) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.
- (4) Wessel, D.; Flügge, U. I. *Analytical Biochemistry* **1984**, *138*, 141–143.
- (5) Yang, Y. Y.; Grammel, M.; Raghavan, A. S.; Charron, G.; Hang, H. C. *Chemistry & Biology* **2010**, *17*, 1212–1222.
- (6) Dirks, A. {. J.; Cornelissen, J. J. L. M.; Nolte, R. J. M.
- (7) Mayer, T. *Eur. J. Org. Chem.* **2007**.
- (8) Scharf, K. D.; Nover, L. *Cell* **1982**.
- (9) Watson, D. G.; Baines, R. A.; Midgley, J. M.; Bacon, J. P. *Methods Mol. Biol.* **1997**, *72*, 225–237.
- (10) Uttamapinant, C.; Tangpeerachaikul, A. *Angewandte Chemie*.

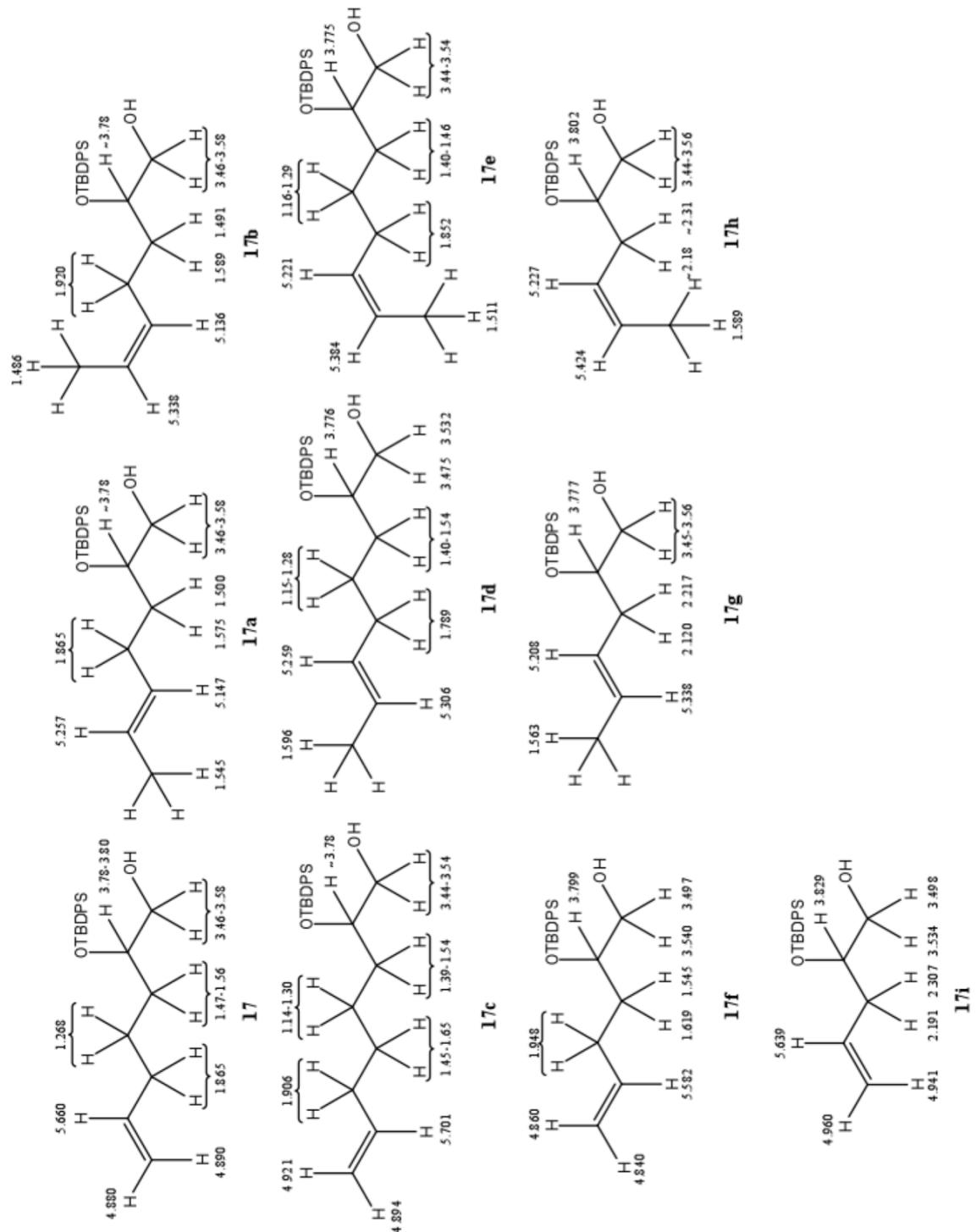
## APPENDIX C

### SYNTHESIS OF THE CAELIFERINS, ELICTORS OF PLANT IMMUNE RESPONSES: ACCESSING LIOPHILIC NATURAL PRODUCTS VIA CROSS METATHESIS

#### 1. 2D NMR Spectroscopic Analysis



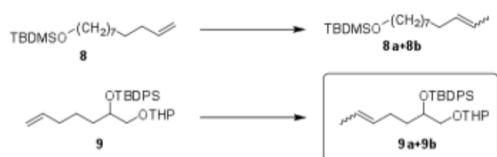
**Figure C.1:** dqc-COSY spectrum of the mixture of **17**, its isomers (**17a** and **17b**), and homologues (**17c-17i**) (600 MHz,  $\text{CDCl}_3$ ). This sample was obtained via THP-protection of **9** isolated from CM reaction of **8** and **9** using G-II (0.05 equiv) and 1,4-benzoquinone (0.10 equiv)



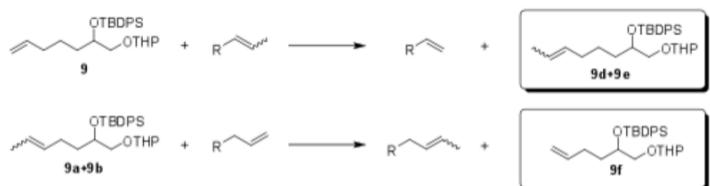
**Figure C.2:** Structures of compounds 17 and 17a-17i as obtained from CM of 8 with 9 with G-II (0.05 equiv) and 1,4-benzoquinone (0.10 equiv).

## 2. Proposed Isomerization and Homologation Pathways

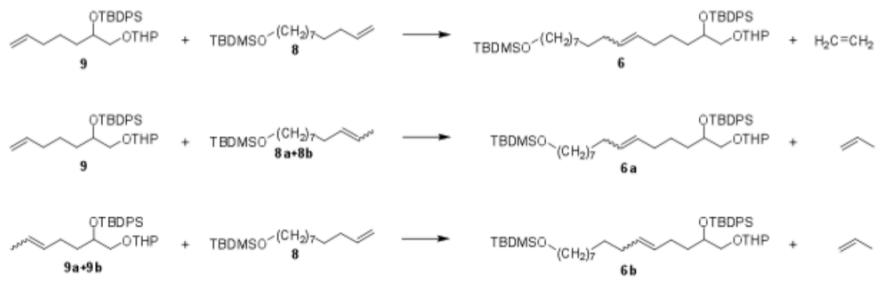
### A. Isomerization of Starting Materials



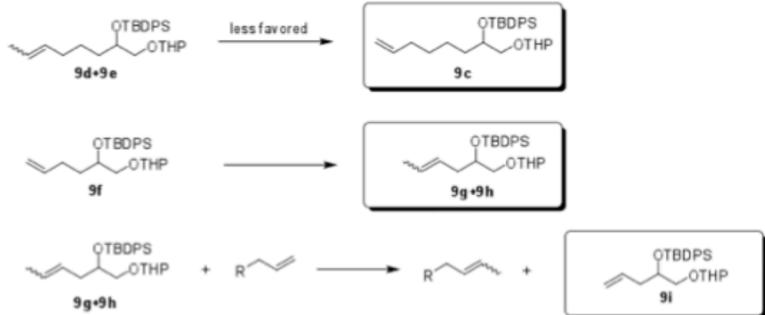
### B. Metathesis: Homologation of Isomerized Starting Materials



### C. Metathesis: Formation of Product and its Nor-Homologues



### D. Further Isomerization and Homologation



**Figure C.3: Proposed isomerization and homologation pathways in the cross metathesis of 8 and 9 with G-II (0.05 equiv) and 1,4-benzoquinone (0.10 equiv), based on analysis of dqFCOSY and ESI+-MS spectra of reaction mixtures. Boxed: isomerization products of 9 and its homologues as detected in the dqFCOSY analysis of their THP-deprotected derivatives, 17 (see Figure C.1)**

## **Experimental**

### **1. Instrumentation and General Procedures**

NMR spectra were recorded on Varian INOVA 600 (600 MHz), Varian INOVA 500 (500 MHz), or Varian INOVA 400 (400 MHz) spectrometers in Cornell University's NMR facility. <sup>1</sup>H NMR chemical shifts are reported in ppm ( $\delta$ ) relative to residual solvent peaks ( $\delta$  2.05 ppm for acetone-d<sub>6</sub>, 7.26 ppm for CDCl<sub>3</sub>, 3.31 ppm for CD<sub>3</sub>OD, 4.79 ppm for D<sub>2</sub>O). NMR-spectroscopic data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), and integration. <sup>13</sup>C NMR chemical shifts are reported in ppm ( $\delta$ ) relative to CHCl<sub>3</sub> ( $\delta$  77.2) in CDCl<sub>3</sub>, CH<sub>3</sub>OH ( $\delta$  49.0) in CD<sub>3</sub>OD, and CH<sub>3</sub>OH ( $\delta$  49.0) in D<sub>2</sub>O). Positive-ion electrospray ionization mass spectra (ESI<sup>+</sup>-MS) were obtained on a Micromass Quattro II mass spectrometer using MassLynx software. Optical rotations were measured on a Perkin Elmer 241 polarimeter. Solvents used for taking optical rotations (water, methanol, chloroform) were not further purified prior to use. Thin-layer chromatography (TLC) was performed using J. T. Baker Silica Gel IB2-F. Flash chromatography was performed using Teledyne Isco CombiFlash systems and Teledyne Isco RediSep Rf silica columns. Unless stated otherwise, reagents were purchased from Sigma-Aldrich and used without further purification. *N,N*-dimethylformamide (DMF), dichloromethane (DCM), dimethyl sulfoxide (DMSO), and tetrahydrofuran (THF) were dried over 4 Å molecular sieves prior to use.

### **2. Cross Metathesis Test Reactions:**

**(A) General Procedure for CM Test Reactions<sup>3</sup>.** A Schlenk flask was flame-dried under vacuum and charged with argon. Starting materials were added as solutions in dry DCM (2 ml of solvent per 0.1 mmol of compound **9**) and the resulting mixture was stirred for 10 min. Next, the catalyst was added and the reaction flask was brought to 40 °C with stirring. The reaction was stirred at 40 °C under argon and monitored by TLC and ESI<sup>+</sup>-MS.

MS samples were prepared by diluting a small aliquot of the CM reaction into methanol, followed by filtration over a pad of silica. The filtered sample was evaporated to dryness under reduced pressure (<0.1 Torr) and re-suspended in methanol. The sample was then directly injected into the MS, monitoring a mass range of *m/z* 250-1000.

**(B) The Effect of 1,4-benzoquinone<sup>4</sup>.** Test CM reactions with catalysts G-I, G-II, and HG-II were conducted as outlined in 3.1.1 above, using mixtures of starting materials 8 and 9 in ratios of 5:1 respectively. Generally, test reactions were performed using 0.2 mmol of 9 and 0.01 mmol of catalyst.

After 2 h, ESI<sup>+</sup>-MS analysis of the G-II and HG-II reaction mixtures revealed 30-35% conversion of starting material 9 and formation of corresponding amounts of product 6, as well as significant amounts of homologues of both reactant and product. Chain-shortened derivatives of 6 accounted for 5-15% of products at this time point. Using G-I, 15-20% of 9 was consumed after 2 h, and corresponding amounts of 6 were detected.

After 20 h, ESI<sup>+</sup>-MS analysis of the G-II and HG-II reaction mixtures revealed more than 65% conversion of 9. Remaining starting materials 8 and 9 as well as product 6 were accompanied by significant amounts of nor-homologues, in addition to trace amounts of chain-extended variants. To investigate whether the production of these unwanted nor-homologues could be suppressed by the inclusion of 1,4-benzoquinone (1,4-BQ), a second series of reactions was performed, with addition of a 0.02 mmol solution of BQ in DCM to the reaction mix prior to the addition of the catalyst (**Table C.1**).

**Table C.1: Reaction conditions for test CM of 8 and 9.** All reactions were carried out at least twice using two different batches of catalyst.

Catalyst	Amount of Catalyst	Amount of 9	Amount of 8	Amount of 1,4-benzoquinone	Solvent (DCM)	Conversion of 9
G-II	0.012 mmol	0.24 mmol	1.18 mmol	0.024 mmol	11 mL	65%
G-II	0.012 mmol	0.24 mmol	1.18 mmol	-	10 mL	90%
HG-II	0.016 mmol	0.32 mmol	1.60 mmol	0.032 mmol	11 mL	70%
HG-II	0.016 mmol	0.32 mmol	1.60 mmol	-	10 mL	90%
G-I	0.012 mmol	0.24 mmol	1.22 mmol	0.024 mmol	11 mL	50%
G-I	0.012 mmol	0.24 mmol	1.22 mmol	-	10 mL	60%

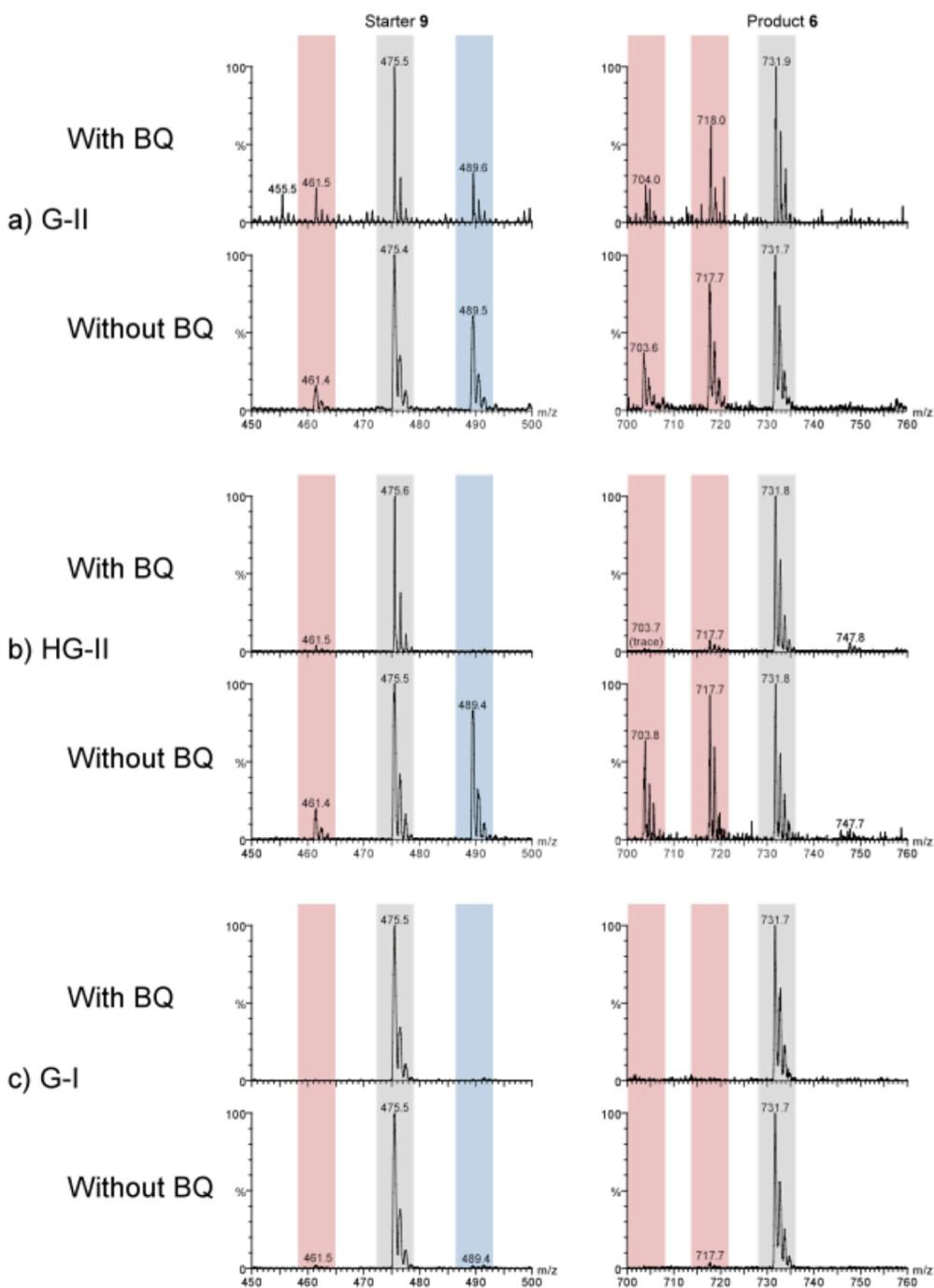
**Figure C.4** shows the ESI<sup>+</sup>-MS spectra of samples from the reaction of **8** and **9** after 20 h, using either G-II, HG-II or G-I, in the presence or absence of 1,4-benzoquinone. Using G-II without BQ resulted in the generation of up to 60% of isomerization, CH<sub>2</sub>-insertion and -deletion products. The inclusion of BQ with G-II in the CM reaction reduced the formation homologues to about 40%. In the case of HG-II, 40-50% of product **6** and starter **9** were converted to their respective homologues. Similarly to G-II, homologation process was suppressed to 10% by addition of BQ. Comparatively small amounts of nor-homologues were found in the BQ-free CM using G-I, and addition of 1,4-BQ to G-I reaction suppressed homolog formation below the detection limit (less than 0.5% homologation). It should be noted that the MS analyses only reveal the extent of homologation, but not double bond isomerization. Therefore, products and starting materials isolated from CM reactions with all three catalysts were additionally analyzed by <sup>1</sup>H NMR, which revealed 5-50% of double bond isomerization, except for the case of using G-I in combination with 1,4-BQ. 2D-NMR spectroscopic experiments are described in the following section.

Similar results were obtained test CM reaction of **7** and **9** (**Table C.2, Figure C5**).

**Table C.2: Reaction conditions for test CM of 7 and 9.** All reactions were carried out at least twice using two different batches of catalyst.

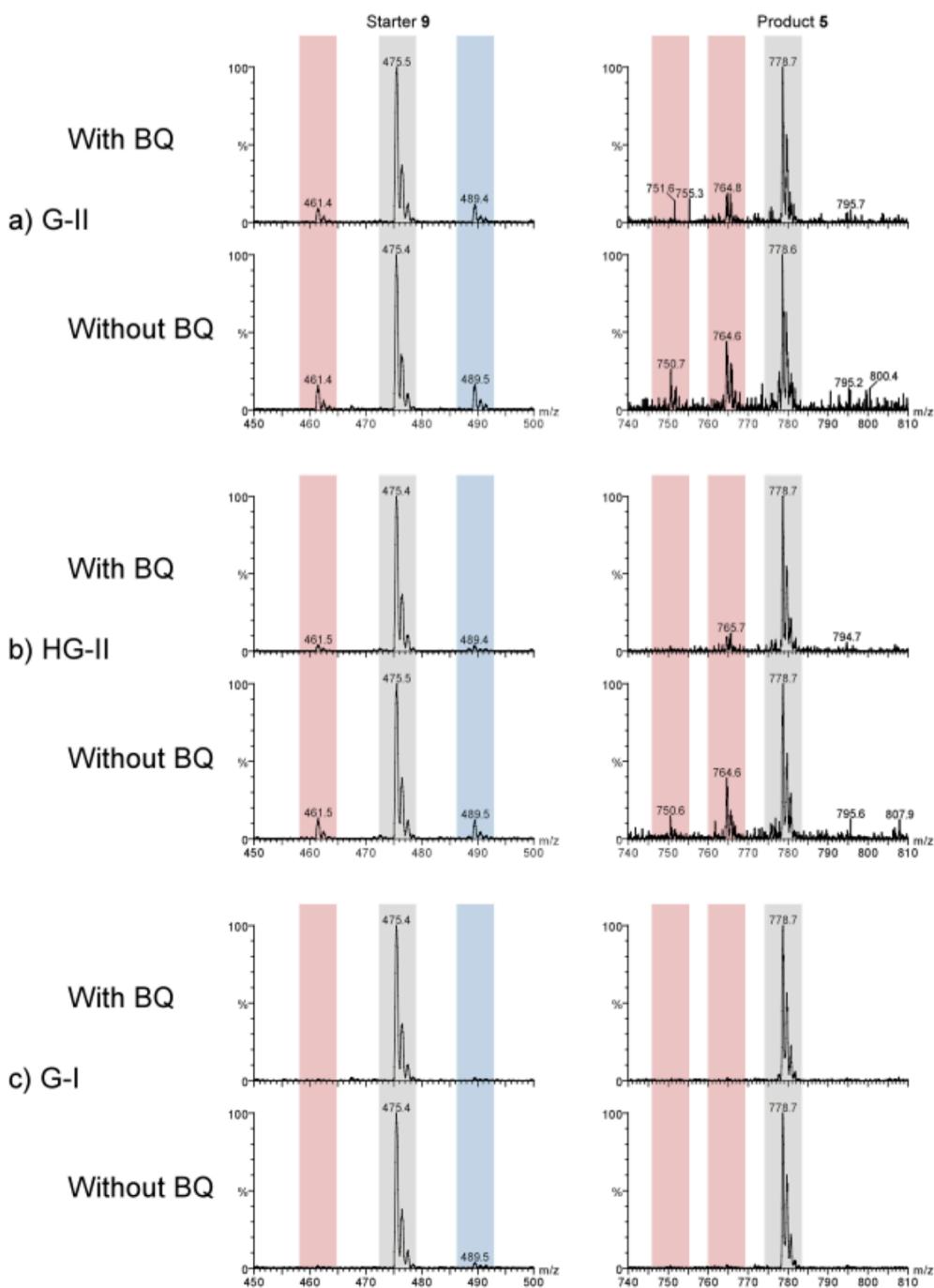
<u>Catalyst</u>	<u>Amount of Catalyst</u>	<u>Amount of 9</u>	<u>Amount of 7</u>	<u>Amount of 1,4-benzoquinone</u>	<u>Solvent (DCM)</u>	<u>Conversion of 9</u>
G-II	0.012 mmol	0.24 mmol	1.18 mmol	0.024 mmol	11 mL	70%
G-II	0.012 mmol	0.24 mmol	1.18 mmol	-	10 mL	65%
HG-II	0.016 mmol	0.32 mmol	1.60 mmol	0.032 mmol	11 mL	65%
HG-II	0.016 mmol	0.32 mmol	1.60 mmol	-	10 mL	65%
G-I	0.012 mmol	0.24 mmol	1.22 mmol	0.024 mmol	11 mL	65%
G-I	0.012 mmol	0..24 mmol	1.22 mmol	-	10 mL	75%

## CM of **8** and **9**



**Figure C.4: ESI<sup>+</sup>-MS spectra of CM reaction mixtures showing [M+Na<sup>+</sup>] for starter 9 ( $m/z$  475.3) and product 6 ( $m/z$  731.5).** Ion signals corresponding to chain shortened and chain extended homologues are shown in red and blue, respectively. All reactions were carried out with 0.05 equiv. of catalyst at 40 °C for 20 h. For reactions with BQ, 0.1 equiv of 1,4-benzoquinone was added to the reaction. Conditions: (a) G-II catalyst; (b) HG-II catalyst; (c) G-I catalyst. For specific reaction conditions see **Table C.1**.

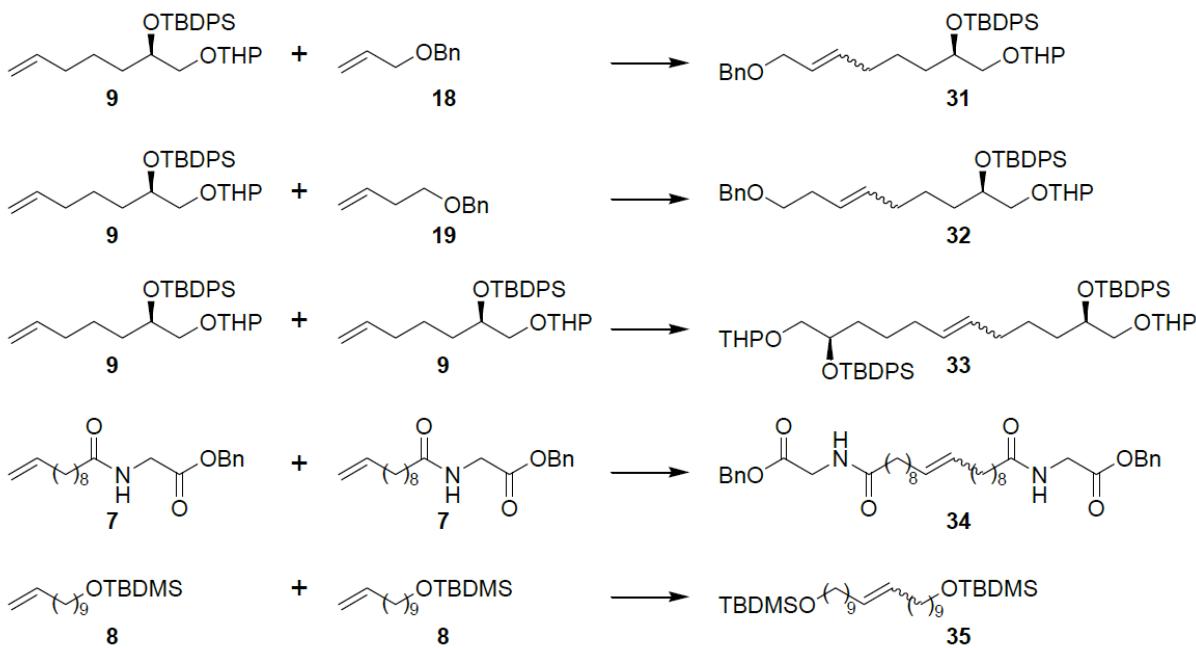
## CM of **7** and **9**



**Figure C.5: ESI<sup>+</sup>-MS spectra of CM reaction mixtures showing [M+Na<sup>+</sup>] for starter 9 ( $m/z$  475.3) and product 5 ( $m/z$  778.4). Ion signals corresponding to chain shortened and chain extended homologues are shown in red and blue, respectively. All reactions were carried out with 0.05 equiv. of catalyst at 40 °C for 20 h. For reactions with BQ, 0.1 equiv of 1,4-benzoquinone was added to the reaction. Conditions: (a) G-II catalyst; (b) HG-II catalyst; (c) G-I catalyst. For specific reaction conditions see Table C.2**

**(C) 2D NMR-spectroscopic Characterization of Isomerization Products.** High resolution dqfCOSY NMR analysis<sup>5</sup> of products from a G-II reaction (as above, reaction time 20 h) with **8**, **9** and 1,4-benzoquinone were conducted after THP-deprotecting the mixture of recovered **9** and its homologues and isomers. For this purpose, the CM reaction was filtered over a pad of silica and evaporated *in vacuo*. This crude mixture was then dissolved in Et<sub>2</sub>O and anhydrous MgBr<sub>2</sub> (5 equiv.) was added with stirring. After stirring overnight, the reaction was cooled to 0 °C and quenched by addition of 5% NaHCO<sub>3</sub> (10 equiv) in distilled H<sub>2</sub>O. Cold pentane was added and the mixture was stirred for 15 min, during which time a voluminous white precipitate of Mg(OH)<sub>2</sub> formed. Subsequently, the organic layer was decanted and the reaction flask rinsed with pentane. The combined organic solutions were dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:9). Fractions containing **17**, its isomers, and homologues were combined, evaporated *in vacuo*, and characterized via dqfCOSY spectra (**Figure 3.4** and **Figure C.1**).

**(D) Alkene Chain Lengths and Degree of Homologation.** CM test reactions using **7** and **9** or **8** and **9** showed that the degree of isomerization was dependent on the activity of the catalyst (Section C.2B, The Effect of Benzoquinone). We further observed that homologation and isomerization were generally less prevalent in the synthesis of **5** (via CM of **7** and **9**) than in the preparation of **6** (via CM of **8** and **9**). To investigate the effect of specific features of CM starting materials on isomerization and homologation, we conducted the additional test reactions shown in **Figure C.7** (for conditions, see **Tables C.3-C.5**).



**Figure C.6: Additional test metathesis reactions**

**Table C.3: Reaction conditions for CM of 9 and 18, 9 and 19, and homodimerization of 9 using G-II catalyst without 1,4-benzoquinone**

Catalyst	Amount of Catalyst	Amount of 9	Amount of second substrate	Solvent (DCM)	Conversion of 9
G-II	0.012 mmol	0.24 mmol	18, allyl benzyl ether 1.18 mmol	10 mL	40%
G-II	0.012 mmol	0.24 mmol	19, homoallyl benzyl ether 1.18 mmol	10 mL	40%
G-II	0.012 mmol	0.24 mmol	-	8 mL	80%

**Table C.4: Reaction conditions for test CM of homodimerization of 7**

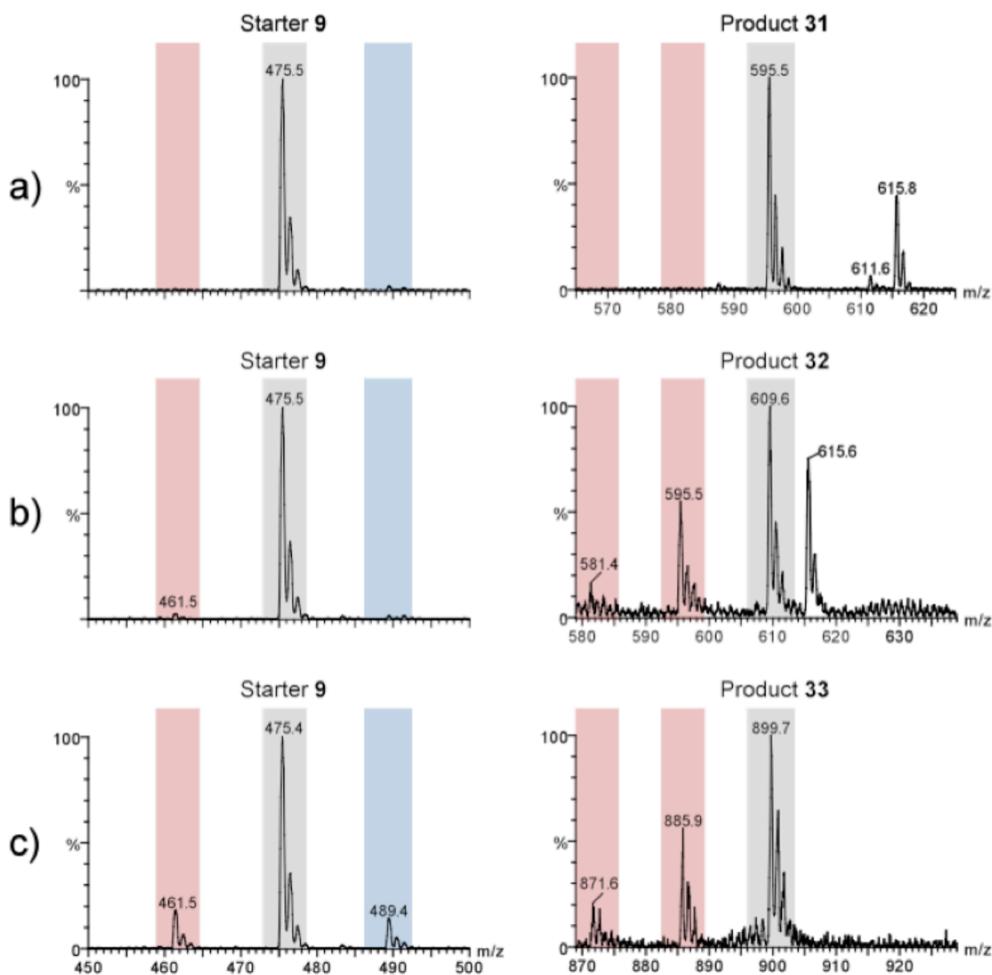
<u>Catalyst</u>	<u>Amount of Catalyst</u>	<u>Amount of 7</u>	<u>Amount of 1,4-benzoquinone</u>	<u>Solvent (DCM)</u>	<u>Conversion of 7</u>
G-II	0.024 mmol	0.24 mmol	0.024 mmol	8 mL	80%
G-II	0.024 mmol	0.24 mmol	-	7 mL	90%
HG-II	0.016 mmol	0.32 mmol	0.032 mmol	8 mL	70%
HG-II	0.016 mmol	0.32 mmol	-	7 mL	90%
G-I	0.012 mmol	0.24 mmol	0.024 mmol	6 mL	70%
G-I	0.012 mmol	0.24 mmol	-	5 mL	80%

**Table C.5: Reaction conditions for test CM of homodimerization of 8**

<u>Catalyst</u>	<u>Amount of Catalyst</u>	<u>Amount of 8</u>	<u>Amount of 1,4-benzoquinone</u>	<u>Solvent (DCM)</u>	<u>Conversion of 8</u>
G-II	0.012 mmol	0.24 mmol	0.024 mmol	8 mL	80%
G-II	0.012 mmol	0.24 mmol	-	7 mL	90%
HG-II	0.016 mmol	0.32 mmol	0.032 mmol	8 mL	80%
HG-II	0.016 mmol	0.32 mmol	-	7 mL	90%
G-I	0.012 mmol	0.24 mmol	0.024 mmol	8 mL	80%
G-I	0.012 mmol	0.24 mmol	-	7 mL	80%

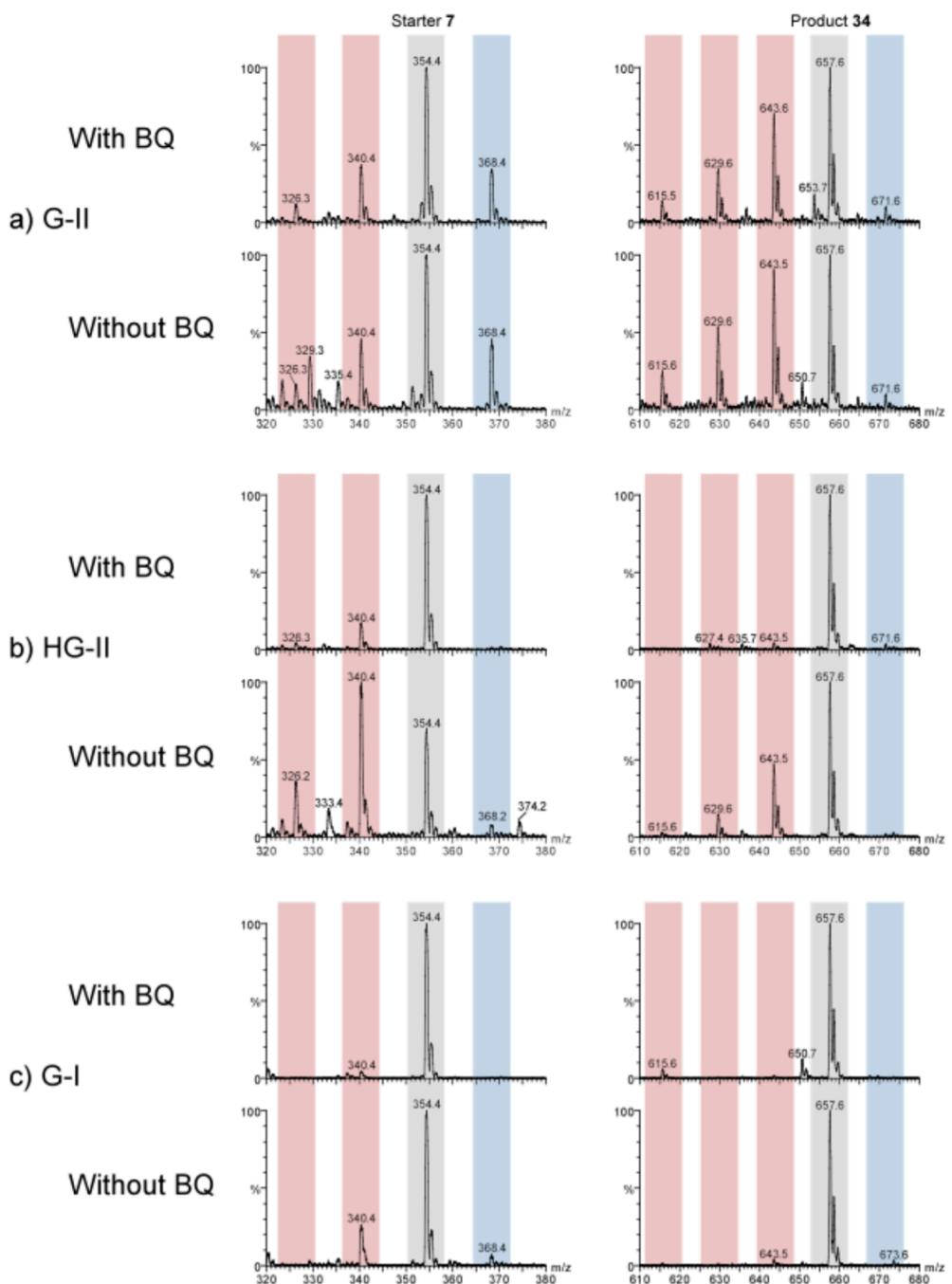
Homologation of remaining starting materials or products was detected in the homodimerization of **9** as well as in CM of **9** with **18** using G-II as catalyst (**Figure C.7**). Upon the addition of 1,4-BQ to these short chain CM reactions, full suppression of homolog formation was achieved, which was verified by ESI<sup>+</sup>-MS. In contrast, homodimerization of long-chain substrates **7** and **8** resulted in extensive homologation when using G-II catalyst. Smaller quantities of homologues were observed when using HG-II and G-I catalysts (**Figures C8, C9**). Addition of 1,4-benzoquinone reduced homologation extent, but in case of the homodimerization of **8** did not fully suppress homologation, even when using the least active G-I catalyst (**Figure C.9**). Comparing the reaction series that resulted in significant homologation and isomerization (**Figures C.4, C.5, C.7-C.9**), the degree of homologation strongly correlates with the chain

lengths of the reactants. Reaction times and conversion of starting materials was roughly similar for all reactions (see **Tables C.1-C.5**).



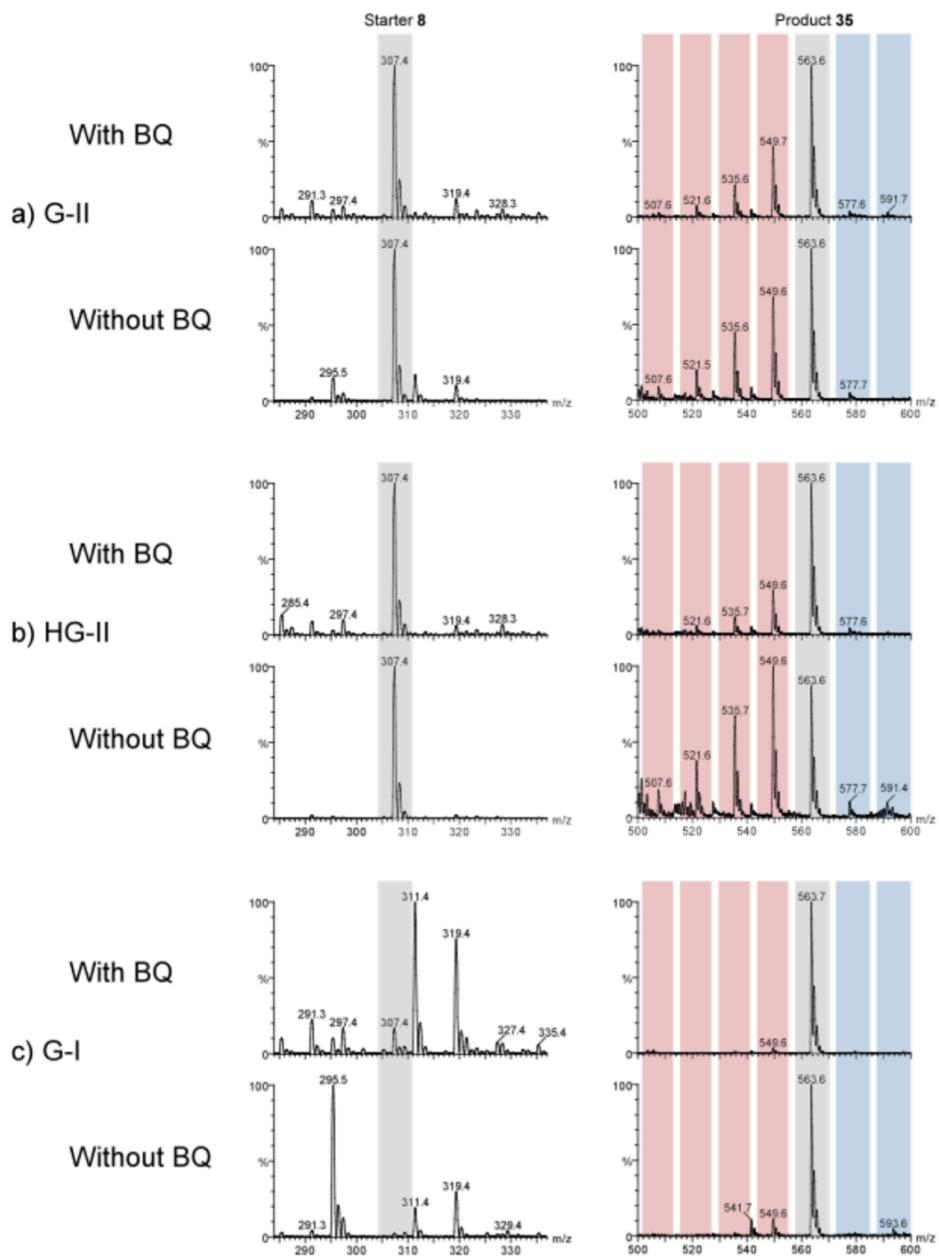
**Figure C.7: ESI<sup>+</sup>-MS analysis of CM test reactions of (a) 9 and 18, (b) 9 and 19, and (c) the homodimerization of 9, using G-II catalyst without benzoquinone.** For each reaction [M+Na<sup>+</sup>] for starter 9 ( $m/z$  475.3) and the products [31 ( $m/z$  595.3), 32 ( $m/z$  609.3), and 33 ( $m/z$  899.5)] are shown, along with signals representing varying quantities of homologues and norhomologues. All reactions were carried out with 0.05 equiv. of G-II catalyst at 40 °C for 20 h. For specific reaction conditions see **Table C.3**.

# Homodimerization of **7**



**Figure C.8: ESI<sup>+</sup>-MS spectra of CM reaction mixtures showing [M+Na<sup>+</sup>] for starter 7 ( $m/z$  354.2) and product 34 ( $m/z$  657.4).** Ion signals corresponding to chain shortened and chain extended homologues are shown in red and blue, respectively. All reactions were carried out with 0.05 equiv. of catalyst at 40 °C for 20 h. For reactions with BQ, 0.1 equiv of 1,4-benzoquinone was added to the reaction. Conditions: (a) G-II catalyst; (b) HG-II catalyst; (c) G-I catalyst. For specific reaction conditions see **Table C.4**

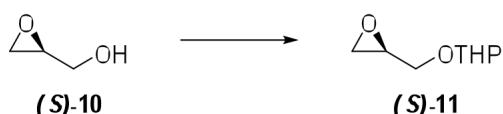
## Homodimerization of **8**



**Figure C.9: ESI<sup>+</sup>-MS spectra of CM reaction mixtures showing [M+Na<sup>+</sup>] for starter 8 ( $m/z$  307.2) and product 35 ( $m/z$  563.5).** Ion signals corresponding to chain shortened and chain extended homologues are shown in red and blue, respectively. All reactions were carried out with 0.05 equiv. of catalyst at 40 °C for 20h. For reactions with BQ, 0.1 equiv of 1,4-benzoquinone was added to the reaction. Conditions: (a) G-II catalyst; (b) HG-II catalyst; (c) G-I catalyst. For specific reaction conditions see **Table C.5**. Note: in contrast to all other substrates surveyed, remaining starting material **8** was *not* accompanied by significant amounts of chain-shortened or elongated derivatives, neither in this homodimerization reaction (see this Figure), nor in CM with compound **9**, despite the fact that reaction of **8** resulted in the largest percentages of homologated products

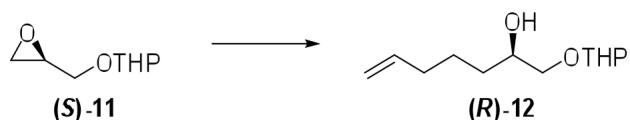
## Synthesis of (*R*)-1

### (A) 2-((*S*)-Oxiran-2-ylmethoxy)tetrahydro-2H-pyran ((*S*)-11):



To a 500 mL round bottom flask, (*S*)-(−)-glycidol (**S**-10) (10.0 g, 135 mmol, 97% ee) and DCM (450 mL) were added. To this mixture, 3,4-dihydropyran (56.8 g, 74.9 mmol, 5.0 equiv) and *p*-toluenesulfonic acid (260 mg, 1.34 mmol, 0.01 equiv) were added with stirring. The mixture turned a light purple color and stirred for 2 h. Subsequently, saturated aqueous NaHCO<sub>3</sub> solution (570 mL) was added to the reaction mixture and stirring was continued for 10 min. The organic layer was separated, and the aqueous layer extracted with two 50ml-portions of DCM. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:2) yielding (**S**-11 as a yellow oil (18.9 g, 120 mmol, 89% yield). **1H NMR** (400 MHz, acetone-d<sub>6</sub>): δ 4.65–4.60 (m, 1H), 3.90–3.85 (dd, *J* = 11.7, 3.0 Hz, 0.5H), 3.85–3.75 (m, 1H), 3.67–3.61 (dd, *J* = 11.6, 3.2 Hz, 0.5H), 3.59–3.53 (dd, *J* = 11.6, 5.6 Hz, 0.5H), 3.48–3.41 (m, 1H), 3.33–3.26 (dd, *J* = 11.7, 6.5 Hz, 0.5H), 3.13–3.07 (m, 1H), 2.73–2.69 (dd, *J* = 5.2, 4.1 Hz, 1H), 2.59–2.50 (m, 1H), 1.87–1.41 (m, 6H) ppm.

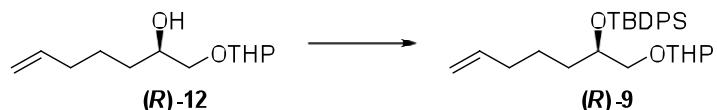
(B) (2*R*)-1-(Tetrahydro-2H-pyran-2-yloxy)hept-6-en-2-ol ((*R*)-12):



A multi-neck round bottom flask was flame-dried under vacuum and charged with argon. Magnesium (3.0 g, 123.9 mmol, 5.0 equiv) was placed into the round bottom flask and activated by heating the flask to 80 °C and stirring with a large stir-bar for 30 min. After cooling the flask with the activated magnesium to r.t., THF (36 mL) and a few iodine crystals were placed into the flask. A 100 mL dropping funnel was fitted to the multi-neck round bottom flask, and 4-bromobut-1-ene (10 g, 74.1 mmol, 3.0 equiv) and THF (36 mL) were placed into the dropping funnel. A few drops of the 4-bromobut-1-ene/THF solution were added to the Mg-containing suspension and stirred gently until the brown iodine color disappeared and the mixture turned clear. Subsequently, the remaining 4-bromobut-1-ene/THF solution was added dropwise to the reaction mixture over a period of 1 h. Following the completion of addition, the mixture stirred for an additional 25 min. Then, the mixture was cooled to -40 °C, copper iodide (0.36 g, 1.9 mmol, 0.07 equiv) was added, and the resulting mixture was stirred for 20 min<sup>6</sup>. Subsequently, the reaction was warmed to -10 °C and stirred for another 10 min. The resulting purple-black mixture was cooled again to -40 °C and stirred for 25 min at this temperature. Next, a solution of (**S**)-**11** (4.05 g, 25.6 mmol) in THF (36 mL) was added dropwise over a period of 20 min. The mixture was gradually warmed to r.t and allowed to stir overnight. Subsequently, the reaction mixture was cooled to 0 °C, and pre-cooled (0 °C) saturated aqueous NH<sub>4</sub>Cl solution (180 mL) was added, followed by stirring for 25 min. The resulting aquamarine-colored mixture was extracted with Et<sub>2</sub>O (3 × 100 mL), and the combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*.

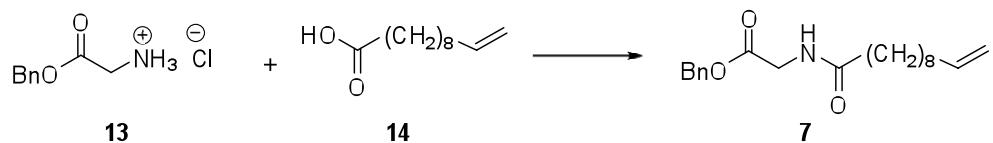
The residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:2) yielding (**R**)-**12** as a colorless oil (4.16 g, 19.4 mmol, 76% yield). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)**: δ 5.81 (ddtd, *J* = 0.8, 6.7, 10.3, 17.1 Hz, 1H), 5.05-4.97 (m, 1H), 4.97-4.92 (m, 1H), 4.59-4.54 (m, 1H), 3.95-3.85 (m, 1H), 3.81-3.72 (m, 1.5H), 3.64 (dd, *J* = 2.55, 10.89 Hz, 0.5H) 3.57-3.47 (m, 1.5H), 3.37-3.31 (dd, *J* = 8.2, 10.8 Hz, 0.5H), 3.18 (d, *J* = 2.8 Hz, 0.5H), 2.64 (d, *J* = 4.0 Hz, 0.5H), 2.15-2.02 (m, 2H), 1.90-1.69 (m, 2H), 21.64-1.36 (m, 8H) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>12</sub>H<sub>22</sub>O<sub>3</sub>Na: 237.15; found 237.1.

**(C) (*2R*)-(*t*-Butyldimethylsilyloxy)-1-(tetrahydro-2*H*-pyran-2-yloxy)hept-6-ene  
((*R*)-9):**



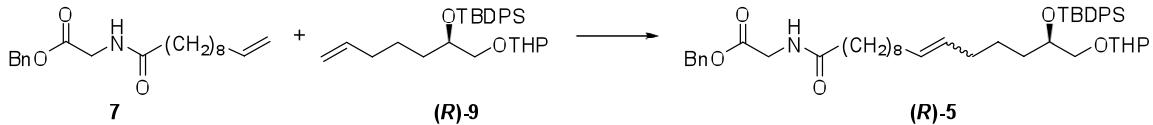
A round bottom flask was flame-dried under vacuum and charged with argon. Alcohol **(R)-12** (3.0 g, 14.0 mmol), imidazol (1.43 g, 21.0 mmol, 1.5 equiv) and dry DMF (20 mL) were added. The mixture was cooled to 0 °C and stirred for 10 min. To the cooled reaction flask, *tert*-butylchlorodiphenylsilane (TBDPSCl) (6.7 mL, 26.3 mmol, 1.9 equiv) was added. The reaction was warmed to r.t. and stirred overnight. The reaction was quenched with saturated aqueous NaCl solution (230 mL) and extracted with Et<sub>2</sub>O (3 × 100 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:2) yielding **(R)-9** as a colorless oil (5.91 g, 13.1 mmol, 93% yield). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)**: δ 7.74-7.67 (m, 4H), 7.44-7.32 (m, 6H), 5.71 (ddt, J=6.6, 10.3, 17.1 Hz, 1H), 4.95-4.87 (m, 2H), 4.49-4.46 (m, 0.5H), 4.36-4.34 (m, 0.5H), 3.94-3.86 (m, 1H), 3.75-3.67 (m, 1H), 3.65-3.61 (m, 1H), 3.44-3.37 (m, 1H), 3.32-3.26 (m, 1H), 1.95-1.88 (m, 2H), 1.77-1.60 (m, 1H), 1.56-1.35 (m, 8H), 1.31-1.24 (m, 1H), 1.04 (d, J = 6.7 Hz, 9H) ppm. **<sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)**: δ 138.8, 114.3, 98.8, 72.2, 70.9, 61.8, 33.8, 33.7, 30.1, 27.0 (3C), 25.4, 24.0, 19.4, 19.2 ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>28</sub>H<sub>40</sub>O<sub>3</sub>SiNa: 475.26; found 475.4. [α]<sup>22</sup><sub>D</sub> = +5.5 (c 0.01, chloroform).

**(D) *N*-(Undec-10-enoyl)glycine benzyl ester (7):**



This compound was prepared as described by Xu et al.<sup>1</sup>

(E) *N*-[15*R*-(*t*-Butyldiphenylsilyloxy)-16-(tetrahydro-2H-pyran-2-yloxy)hexadec-10-enoyl] glycine benzyl ester ((*R*)-5):

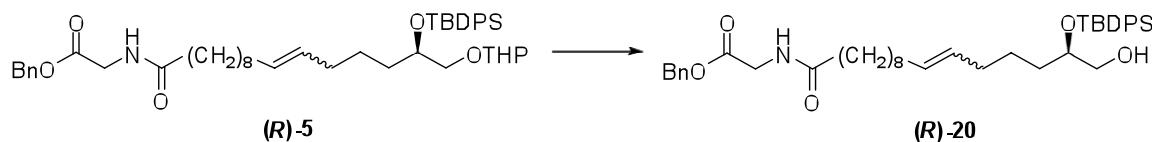


A Schlenk flask was flame-dried under vacuum and charged with argon. To the flask, (*R*)-**9** (1.99 g, 4.4 mmol) and dry DCM (45 mL) were added. This mixture was stirred for 10 min, at which time benzyl 2-undec-10-enamidoacetate **7** (7.28 g, 22.0 mmol, 5.0 equiv) was added. This mixture was stirred for 10 min. Subsequently, Grubb's 1<sup>st</sup> Generation (G-I) catalyst (180 mg, 0.22 mmol, 0.05 equiv) was added to the mixture. The reaction flask was brought to 40 °C and the purple mixture was allowed to stir for 12 h during which time the reaction was monitored by TLC (EtOAc:toluene, 1:2) and ESI<sup>+</sup>-MS. The reaction mixture was then filtered through silica. The resulting residue was concentrated *in vacuo* and was purified by silica gel flash chromatography (EtOAc:toluene, 1:4) yielding (*R*)-**5** as a colorless oil (950 mg, 1.25 mmol, 40% yield, 1:4 cis:trans ratio). 18% of unreacted (*R*)-**9** was recovered and could be used for subsequent reactions.

**<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>):** δ 7.73-7.67 (m, 4H), 7.43-7.32 (m, 9H), 7.27-7.23 (m, 1H), 7.19-7.14 (m, 1H), 5.96-5.88 (m, 1H), 5.34-5.22 (m, 2H), 5.19 (s, 2H), 4.49-4.46 (m, 0.5H), 4.37-4.33 (m, 0.5H), 4.09 (d, *J* = 5.2 Hz, 2H), 3.94-3.84 (m, 1H), 3.75-3.67 (m, 1H), 3.65-3.59 (m, 1H), 3.44-3.36 (m, 1H), 3.31-3.25 (m, 1H), 2.25-2.20 (m, 2H), 1.98-1.80 (m, 4H), 1.75-1.19 (m, 22H), 1.07-1.03 (m, 9H) ppm.

**ESI-MS (*m/z*):** [M+Na<sup>+</sup>] calcd. for C<sub>46</sub>H<sub>65</sub>NO<sub>6</sub>SiNa: 778.45; found 778.4.

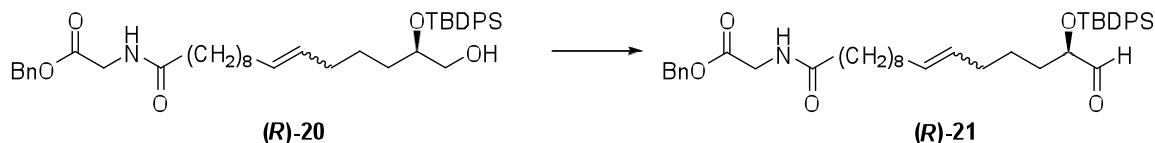
**(F) *N*-[15*R*-(*t*-Butyldiphenylsilyloxy)-16-hydroxyhexadec-10-enoyl]glycine benzyl ester (*R*)-20:**



To a round bottom flask, **(R)-5** (390 mg, 0.52 mmol) and Et<sub>2</sub>O (10 mL) were added. To this stirred mixture, anhydrous MgBr<sub>2</sub> (480 mg, 2.59 mmol, 5.0 equiv) was added and stirring was continued for 6 h. Subsequently, the reaction was cooled to 0 °C and quenched with NaHCO<sub>3</sub> (480 mg, 5.69 mmol, 11.0 equiv) in distilled H<sub>2</sub>O (20 mL). Pentane (12 mL) were added to the reaction mixture and stirring was continued for 15 min. While stirring, large quantities of Mg(OH)<sub>2</sub> precipitated from the aqueous layer. The organic layer was decanted and the reaction flask washed with pentane (2 × 20 mL). The combined organic solutions were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by silica gel flash chromatography (EtOAc:toluene, 1:2) yielding **(R)-20** as a colorless oil (210 mg, 0.32 mmol, 61% yield). **<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)**: δ 7.69-7.62 (m, 5H), 7.46-7.29 (m, 10H), 6.00-5.92 (m, 1H), 5.33-5.18 (m, 2H), 5.19 (s, 2H), 4.09 (d, *J*=5.2 Hz, 2H), 3.80-3.73 (m, 1H), 3.56-3.43 (m, 2H), 2.26-2.19 (m, 2H), 1.96-1.74 (m, 4H), 1.69-1.57 (m, 2H), 1.56-1.37 (m, 2H), 1.35-1.12 (m, 12H), 1.07 (s, 9H) ppm. **<sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)**: δ 173.2, 170.0, 135.84, 135.75 (4C), 135.71, 135.1, 133.9, 133.8, 129.7 (2C), 128.7, 128.5 (2C), 128.4 (2C), 127.7, 127.6, 130.7, 129.6, 74.0, 67.2, 66.0, 41.3, 36.39, 36.40, 33.0, 32.5, 32.4, 29.4, 29.3-29.2 (2C), 29.2, 27.0 (3C), 25.5, 25.0, 19.3 ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>41</sub>H<sub>57</sub>NO<sub>5</sub>SiNa: 694.39; found 694.0.

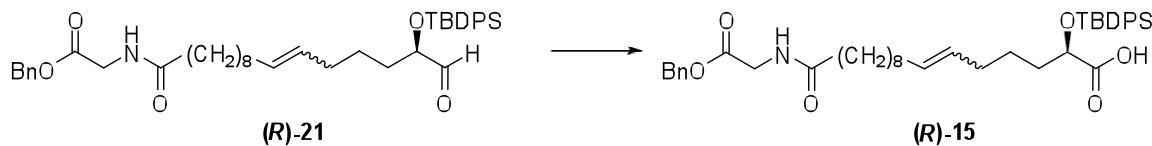
**(G) *N*-[15*R*-(*t*-Butyldiphenylsilyloxy)-16-oxohexadec-10-enoyl]glycine**

**benzyl ester ((*R*)-21):**



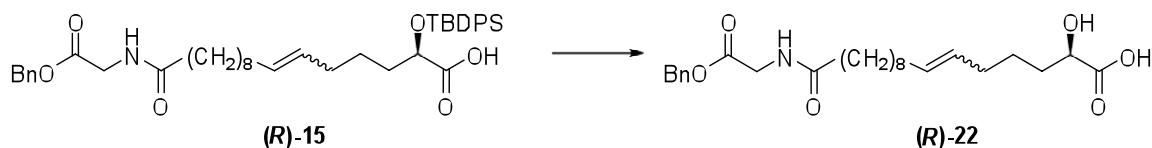
A Schlenk flask was flame-dried under vacuum and charged with argon. (*R*)-20 (240 mg, 0.36 mmol) and dry DCM (6.1 mL) were added under argon. This stirred mixture was cooled to 0 °C, *N,N*-diisopropylethylamine (150 µL, 1.07 mmol, 3.0 equiv) was added, and stirring was continued for 15 min. In a separate flame-dried and argon-charged Schlenk flask, a mixture of SO<sub>3</sub>-pyridine complex (860 mg) and dry DMSO (10 mL) was prepared. Consequently, a portion of the SO<sub>3</sub>-pyridine/DMSO mixture (2.2 mL, 1.18 mmol, 3.3 equiv) was added to the reaction flask at 0 °C. This reaction was allowed to stir for 1.5 h, at which time the reaction was quenched with saturated aqueous NaCl solution (10 mL) and extracted with DCM (2 × 10 mL) and then with EtOAc (2 × 10 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:toluene, 1:3) yielding (*R*)-21 as a colorless oil (200 mg, 0.29 mmol, 83% yield). <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 9.58-9.55 (m, 1H), 7.66-7.61 (m, 5H), 7.46-7.32 (m, 10H), 5.99-5.91 (m, 1H), 5.37-5.22 (m, 2H), 5.19 (s, 2H), 4.09 (d, *J* = 5.2 Hz, 2H), 4.03 (td, *J* = 5.9, 1.6 Hz, 1H), 2.34-2.21 (m, 2H), 1.97-1.84 (m, 4H), 1.65-1.19 (m, 16H), 1.11 (s, 9H) ppm. ESI-MS (*m/z*): [M+CH<sub>3</sub>OH+Na<sup>+</sup>] calcd. for C<sub>42</sub>H<sub>59</sub>NO<sub>6</sub>SiNa: 724.40; found 724.1.

**(H) *N*-[15*R*-(*t*-Butyldiphenylsilyloxy)-15-carboxypentadec-10-enoyl]glycine benzyl ester ((*R*)-15):**



In an Erlenmeyer flask, a buffer solution was prepared from NaH<sub>2</sub>PO<sub>4</sub> (8.938 g) in distilled-H<sub>2</sub>O (160 mL). To a round bottom flask, aldehyde (**R**)-21 (680 mg, 1.0 mmol), *t*-butanol (37.0 mL), and a freshly distilled 2-methylbut-2-ene (18.0 mL) were added, and the reaction flask was brought to 0 °C. To this mixture, a portion of the prepared buffer solution (34.0 mL) was added in one portion. In a separate flask, a fresh oxidant solution was prepared by mixing NaClO<sub>2</sub> (2.0 g) and buffer solution (34.0 mL). Immediately after preparation, the oxidant solution was added via Pasteur pipette to the reaction mixture stirred at 0 °C. The reaction was stirred for 10 min at 0 °C, followed by stirring for 3 h at r.t. The reaction was quenched by addition of dimethyl sulfide (DMS) (5.5 mL) and the remaining buffer solution and then brought to pH = 4 using citric acid solution (citric acid:H<sub>2</sub>O, 1:10, {w:v}). The resulting mixture was extracted with EtOAc (3 × 100 mL), and the combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:toluene, 1:3) yielding (**R**)-15 as a colorless oil (660 mg, 96% yield). **<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):** δ 7.68-7.59 (m, 5H), 7.49-7.32 (m, 10H), 6.11-6.04 (m, 1H), 5.35-5.19 (m, 2H), 5.19 (s, 2H), 4.34-4.29 (m, 1H), 4.14-4.04 (m, 2H), 2.27-2.20 (m, 2H), 1.97-1.81 (m, 4H), 1.72-1.41 (m, 6H), 1.34-1.19 (m, 10H), 1.12 (s, 9H) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>41</sub>H<sub>55</sub>NO<sub>6</sub>SiNa: 708.37; found 708.1.

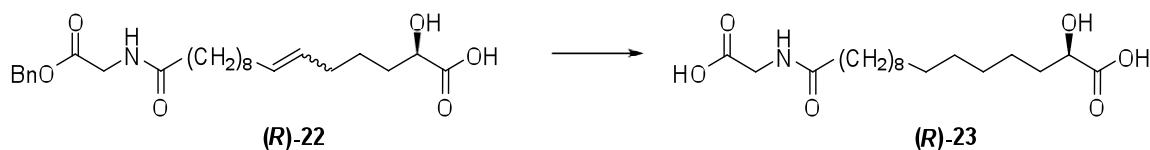
(I) ***N*-(15-Carboxy-15*R*-hydroxypentadec-10-enoyl)glycine benzyl ester ((*R*)-22):**



To a stirred solution of **(R)-15** (230 mg, 0.42 mmol) in acetonitrile (5 mL) at 0 °C was added 40% HF (2 drops). The reaction was allowed to warm to r.t. and stirred for 18 h. Subsequently, the reaction was cooled to 0 °C and quenched with saturated aqueous NaHCO<sub>3</sub> solution (20 drops). The resulting mixture was concentrated *in vacuo* and was purified by C18 reverse-phase flash chromatography, using a water/acetonitrile gradient as solvents to yield **(R)-22** (170 mg, 0.37 mmol, 90% yield). **<sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):** δ 7.39-7.29 (m, 5H), 5.46-5.31 (m, 2H), 5.17 (s, 2H), 4.09-4.00 (m, 1H), 3.96 (s, 2H), 2.27-2.21 (m, 2H), 2.11-1.94 (m, 4H), 1.82-1.70 (m, 1H), 1.67-1.55 (m, 3H), 1.53-1.43 (m, 2H), 1.38-1.23 (m, 10H) ppm.

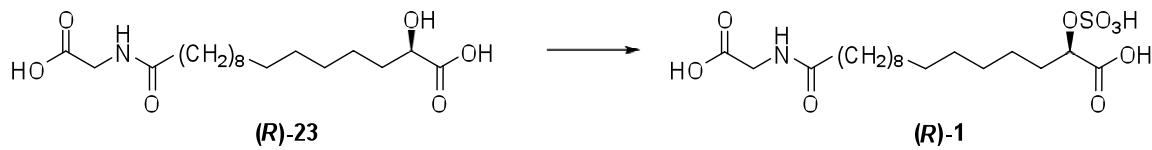
**ESI-MS (*m/z*):** [M+Na<sup>+</sup>] calcd. for C<sub>25</sub>H<sub>37</sub>NNO<sub>6</sub>Na: 470.25; found 470.5.

**(J) *N*-(15-Carboxy-15*R*-hydroxypentadecanoyl)glycine ((*R*)-23):**



A Schlenk flask was flame-dried under vacuum and charged with argon. To the Schlenk flask, 10 wt % Pd/C (17 mg) and EtOH (4 mL) were added. The reaction flask was flushed briefly with argon and subsequently with H<sub>2</sub> gas for 5 min. The H<sub>2</sub> flow was reduced and acid **(R)-22** (19 mg, 0.04 mmol) was added to the Pd/C/EtOH mixture. The reaction was stirred for 10 h under H<sub>2</sub>. Subsequently, the mixture was filtered through silica, using methanol to (50 ml) for elution. The combined filtrates were concentrated *in vacuo* and purified by C-18 reverse-phase flash chromatography, using a water/acetonitrile solvent gradient, to yield **(R)-23** (13 mg, 0.36 mmol, 85% yield). **<sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):** δ 4.09-4.04 (m, 1H), 3.86 (s, 2H), 2.28-2.21 (m, 2H), 1.80-1.70 (m, 1H), 1.68-1.57 (m, 3H), 1.48-1.21 (m, 20H) ppm. **<sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD):** δ 178.3, 176.3, 173.5, 71.4, 41.9, 36.5, 35.2, 30.4-30.3 (6C), 30.3, 30.1, 26.0, 26.5, 25.8 ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>18</sub>H<sub>33</sub>NO<sub>6</sub>Na: 382.22; found 382.3. [α]<sup>22</sup><sub>D</sub> = +5.5 (c 0.002, methanol).

**(K) *N*-(15-carboxy-15*R*-sulfooxypentadecanoyl) glycine ((*R*)-1):**



A Schlenk flask was flame-dried under vacuum and charged with argon. The acid (**R**)-23 (32 mg, 0.09 mmol) and dry DMF (6.6 mL) were added to the flask. In a separate Schlenk flask, a solution of SO<sub>3</sub>DMF complex (320 mg) in dry DMF (3 mL) was prepared. A portion of the SO<sub>3</sub>DMF solution (0.8 mL, 0.54 mmol, 6.0 equiv) was added to the alcohol/DMF mixture at 0 °C. The reaction was stirred for 30 min at 0 °C and then for 20 min at r.t. Subsequently, the reaction flask was cooled again to 0 °C and quenched with a solution of KHCO<sub>3</sub> (110 mg, 1.1 mmol, 12.0 equiv) in distilled H<sub>2</sub>O (3 mL), resulting in a yellow mixture with pH = 4. This mixture was concentrated under reduced pressure, and the residue purified by silica gel flash chromatography (solvents: DCM with 1% AcOH and acetonitrile containing 1% AcOH and 5% distilled H<sub>2</sub>O in a ratio of 2:1) to yield (**R**)-1 as a white solid (28 mg, 0.06 mmol, 71% yield). **<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O):** δ 4.52-4.47 (m, 1H), 3.70 (s, 2H), 2.27-2.22 (m, 2H), 1.79-1.65 (m, 2H), 1.60-1.49 (m, 2H), 1.40-1.00 (m, 20H) ppm. **<sup>13</sup>C NMR (151 MHz and 126 MHz, D<sub>2</sub>O):** δ 178.0, 176.9, 176.5, 77.9, 41.7, 35.4, 31.9, 28.6-28.2 (8C), 27.9, 25.1, 23.9 ppm. ESI-MS (*m/z*): [M-H] calcd. for C<sub>18</sub>H<sub>32</sub>NO<sub>9</sub>S: 358.22; found 358.4. [α]<sup>22</sup><sub>D</sub> = +15.6 (c 0.002, methanol).

## Synthesis of (*R*)-3, (*R*)-4, and (*S*)-4

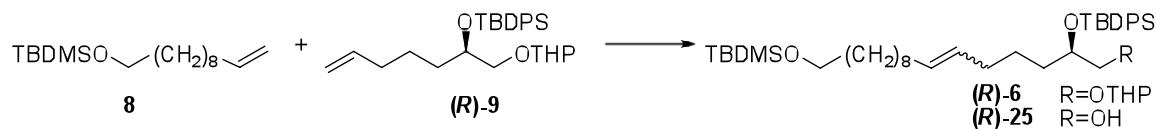
### (A) *O*-(*t*-Butyldimethylsilyl)undec-10-enol (8):



To a stirred solution of alcohol **24** (5.0 g, 29.4 mmol) and imidazol (3.0 g, 44.0 mmol, 1.5 equiv) in dry DMF (20 mL) at 0 °C was added *tert*-butylchlorodimethylsilane (TBDMSCl) (7.08 g, 47.0 mmol, 1.6 equiv), and the mixture was stirred overnight at r.t. The reaction was quenched with saturated aqueous NaCl solution (60 mL) and extracted with Et<sub>2</sub>O (3 × 100 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:2) yielding **8** as a colorless oil (7.64 g, 26.9 mmol, 92% yield). **1H NMR (600 MHz, CDCl<sub>3</sub>)**: δ 5.81 (ddt, *J* = 17.1, 10.2, 6.7 Hz, 1H), 5.01-4.97 (m, 1H), 4.94-4.91 (m, 1H), 3.595 (t, *J* = 6.7 Hz, 2H), 2.07-2.01 (m, 2H), 1.54-1.47 (m, 2H), 1.40-1.34 (m, 2H), 1.33-1.24 (m, 10H), 0.89 (s, 9H), 0.05 (s, 6H) ppm. **13C NMR (151 MHz, CDCl<sub>3</sub>)**: δ 139.2, 114.2, 63.3, 33.8, 32.8, 29.5-29.4 (2C), 29.4, 29.2, 28.9, 26.0 (3C), 25.8, 18.4, -5.3 (2C) ppm.

**(B) 16-(*t*-Butyldimethylsilyloxy)-2*R*-(*t*-butyldiphenylsilyloxy)hexadec-6-enol**

((*R*)-25):

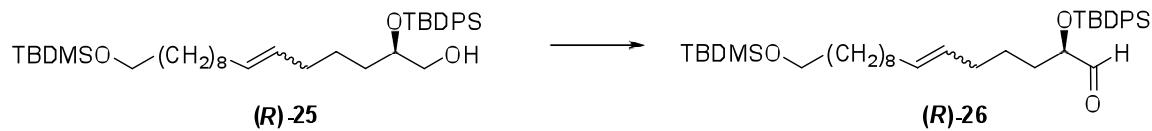


A Schlenk flask was flame-dried under vacuum and charged with argon. To this flask, a solution of (*R*)-9 (0.76 g, 1.68 mmol) in dry DCM (18 mL), a solution of benzoquinone (36 mg, 0.34 mmol, 0.2 equiv) in dry DCM (1.5 mL), and a solution of **8** (2.38 g, 8.39 mmol, 5.0 equiv) in DCM (2 ml) was added. All solutions were prepared under argon. The resulting mixture was stirred for 10 min and Grubb's 1<sup>st</sup> Generation (G-I) catalyst (138 mg, 0.17 mmol, 0.1 equiv) was added. Following the addition of the catalyst, the reaction was brought to 40 °C and the purple mixture was stirred for 20 h. The reaction was monitored by TLC (EtOAc:hexanes, 1:9) and ESI<sup>+</sup>-MS. Subsequently, the reaction mixture was cooled to r.t. and filtered over a small pad of silica. The filtrate was concentrated *in vacuo* and partially purified by silica gel flash chromatography (EtOAc:hexanes, 1:9), yielding a mixture (1.39 g) of desired (*R*)-6, starting material (*R*)-9, starting material **8**, and homodimer of **8**. The ratio of the THP group protons for (*R*)-6 and (*R*)-9 was 1.8:1 as determined by NMR-spectroscopic analysis, corresponding to 65% yield of (*R*)-6. This mixture was not purified further at this stage and instead directly subjected to THP-deprotection using MgBr<sub>2</sub>.

For this purpose, the crude mixture (1.39 g, 1.96 mmol) was dissolved in Et<sub>2</sub>O (24 mL) and anhydrous MgBr<sub>2</sub> (1.85 g, 10.1 mmol, 5.1 equiv) was added with stirring. After stirring overnight, the reaction was cooled to 0 °C and

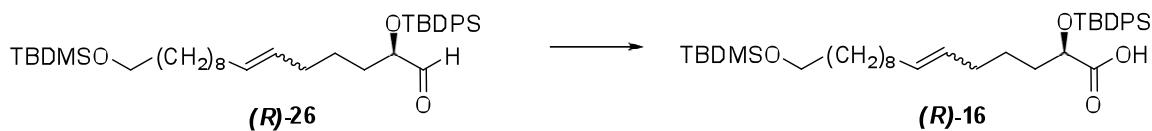
quenched by addition of a solution of NaHCO<sub>3</sub> (1.69 g, 20.1 mmol, 10.2 equiv) in distilled H<sub>2</sub>O (34 mL). Cold pentane (24 mL) were added and the mixture was stirred for 15 min, during which time a voluminous white precipitate of Mg(OH)<sub>2</sub> formed. Subsequently, the organic layer was decanted and the reaction flask rinsed with pentane (2 × 30 mL). The combined organic solutions were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:9) yielding (*R*)-25 as a colorless oil (476 mg, 0.73 mmol, 45% yield over two steps). 130 mg (0.35 mmol, 20%) of THP-deprotected **9** was recovered from the deprotection, giving (*R*)-25 in 55% yield over two steps based on consumed **9**. **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)**: δ 7.70-7.65 (m, 4H), 7.46-7.41. (m, 2H), 7.40-7.34 (m, 4H), 5.33-5.15 (m, 2H), 3.80-3.75 (m, 1H), 3.593 (t, *J* = 6.7 Hz, 2H), 3.56-3.44 (m, 2H), 1.94-1.88 (m, 2H), 1.86-1.76 (m, 2H), 1.54-1.39 (m, 4H), 1.33-1.17 (m, 14H), 1.07 (s, 9H), 0.89 (s, 9H), 0.05 (s, 6H) ppm. **<sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)**: δ 135.8 (4C), 133.9 (2C), 130.8, 129.7 (2C), 129.6, 127.7 (4C), 74.0, 66.0, 63.3, 33.0, 32.9, 32.6, 32.4, 29.6-29.5 (2C), 29.53, 29.45, 27.10, 27.05 (3C), 26.0 (3C), 25.8, 25.2, 19.5, 18.4, -5.3 (2C) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>38</sub>H<sub>64</sub>O<sub>3</sub>Si<sub>2</sub>Na: 647.43; found 647.2.

**(C) 16-(*t*-Butyldimethylsilyloxy)-2*R*-(*t*-butyldiphenylsilyloxy)hexadec-6-enal  
((R)-26):**



A Schlenk flask was flame-dried under vacuum and charged with argon. Alcohol **(R)-25** (0.8 g, 1.28 mmol) and dry DCM (15 mL) were placed into the flask under argon. The stirred reaction mixture was cooled to 0 °C, at which time *N,N*-diisopropylethylamine (0.6 mL, 4.1 mmol, 3.2 equiv) was added. The mixture was stirred for 15 min. In a separate flame-dried and argon-charged Schlenk flask, a mixture of SO<sub>3</sub>-pyridine complex (1.42 g, 8.9 mmol) and dry DMSO (8.8 mL) was prepared under argon atmosphere. Subsequently, a portion of the SO<sub>3</sub>-pyridine/DMSO mixture (4 mL, 4.09 mmol, 3.2 equiv) was added dropwise to the reaction flask at 0 °C. This mixture was stirred for another 1.5 h, at which time the reaction was quenched with saturated aqueous NaCl solution (20 mL), followed by extraction with DCM (2 × 10 mL) and EtOAc (3 × 20 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and the residue was purified by silica gel flash chromatography (EtOAc:hexanes, 1:9) yielding **(R)-26** as a colorless oil (746 mg, 1.2 mmol, 93% yield). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>):** δ 9.58-9.55 (m, 1H), 7.66-7.61 (m, 4H), 7.46-7.41 (m, 2H), 7.40-7.35 (m, 4H), 5.37-5.21 (m, 2H), 4.02 (dt, *J* = 1.6, 5.8 Hz, 1H), 3.60 (t, *J* = 6.7 Hz, 2H), 1.96-1.86 (m, 4H), 1.68-1.55 (m, 3H), 1.53-1.48 (m, 2H), 1.48-1.40 (m, 1H), 1.39-1.23 (m, 12H), 1.11 (s, 9H), 0.89 (s, 9H), 0.05 (s, 6H) ppm. ESI-MS (*m/z*): [M+CH<sub>3</sub>OH+Na<sup>+</sup>] calcd. for C<sub>39</sub>H<sub>66</sub>O<sub>4</sub>Si<sub>2</sub>Na: 677.44; found 677.1.

**(D) 16-(*t*-Butyldimethylsilyloxy)-2*R*-(*t*-butyldiphenylsilyloxy)hexadec-6-enoic acid ((*R*)-16):**



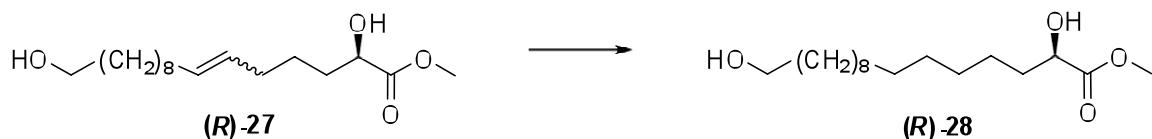
A buffer solution was prepared from NaH<sub>2</sub>PO<sub>4</sub> (8.17 g) and distilled H<sub>2</sub>O (150 mL). A stirred mixture of aldehyde (**R**-26 (547 mg, 0.88 mmol), *t*-BuOH (33.5 mL), and freshly distilled solution of 2-methylbut-2-ene (16.7 mL) was cooled to 0 °C, and 30 mL of the prepared buffer solution was added. Separately, a fresh oxidant solution was prepared by dissolving NaClO<sub>2</sub> (1.79 g) in buffer solution (34 mL). Immediately following preparation, the oxidant solution was added drop-wise to the reaction mixture. The reaction was allowed to stir for 10 min at 0 °C and then for 3 h at r.t. The reaction was quenched by addition of DMS (3 mL) and the remaining buffer solution. The reaction mixture was then brought to pH = 4 using a citric acid solution (citric acid:H<sub>2</sub>O, 1:10, {w:v}). The resulting mixture was extracted with EtOAc (3 × 50 mL) and the combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (EtOAc:hexanes, 1:9) yielding (**R**)-16 as a colorless oil (490 mg, 0.77 mmol, 88% yield). **<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>):** δ 7.67-7.58 (m, 4H), 7.49-7.36 (m, 6H), 5.45-5.15 (m, 2H), 4.35-4.29 (m, 1H), 3.60 (t, *J* = 6.7 Hz, 2H), 1.99-1.77 (m, 4H), 1.71-1.41 (m, 6H), 1.35-1.19 (m, 12H), 1.12 (s, 9H), 0.89 (s, 9H), 0.05 (s, 6H) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>38</sub>H<sub>62</sub>O<sub>4</sub>Si<sub>2</sub>Na: 661.41; found 661.3.

**(E) 16-(*t*-Butyldimethylsilyloxy)-2*R*-(*t*-butyldiphenylsilyloxy)hexadec-6-enoic acid ((*R*)-16):**



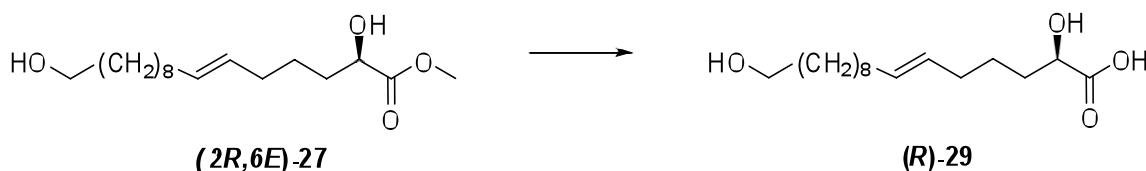
Acid (*R*)-16 (556 mg, 0.87 mmol) was dissolved in acetonitrile (10 mL) at 0 °C and 40% HF (28 drops, 0.2 mL) was added with stirring. The reaction was allowed to warm to r.t. and stirred for 40 h. Subsequently, the reaction was cooled to 0 °C and quenched with saturated aqueous NaHCO<sub>3</sub> solution (5 mL). The resulting mixture was extracted with ether (3 x 20 ml) and the combined ether extracts were concentrated *in vacuo*, filtered over a pad of silica (ethylacetate/hexanes) and used for the next step without further purification. The deprotected acid was methylated as follows: the acid was dissolved in a mixture of 3:2 toluene:MeOH, and trimethylsilyldiazomethane (2 M in ether, 0.65 mL, 1.31 mmol, 1.5 equiv) was added until a yellow color persisted. The reaction mixture was concentrated *in vacuo* and purified by silica gel flash chromatography (MeOH:DCM, 1:10) to yield (*R*)-27 (193 mg, 0.67 mmol, 77% yield over two steps). An AgNO<sub>3</sub> impregnated on silica column<sup>2</sup> was used to obtain (*2R,6E*)-27 with >95% 6*E* configuration. <sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD): δ 5.44-5.33 (m, 2H), 4.17-4.13 (m, 1H), 3.79 (s, 3H), 3.66-3.61 (m, 2H), 2.702 (d, *J* = 5.8 Hz, 1H), 2.04-1.94 (m, 4H), 1.82-1.75 (m, 1H), 1.67-1.59 (m, 1H), 1.54-1.39 (m, 2H), 1.38-1.22 (m, 14H) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>17</sub>H<sub>32</sub>O<sub>4</sub>Na: 323.22; found 323.3.

**(F) 16-(*t*-Butyldimethylsilyloxy)-2*R*-(*t*-butyldiphenylsilyloxy)hexadec-6-enoic acid ((*R*)-16):**



A Schlenk flask was flame-dried under vacuum and charged with argon, and Pd/C (34mg) and EtOH (15 mL) were added. The reaction flask was briefly flushed with argon and subsequently flushed with H<sub>2</sub> gas for 15 min. The H<sub>2</sub> flow was reduced and (*R*)-27 (38 mg, 0.13 mmol, mixture of cis/trans), dissolved in ethanol (5 mL) was added. The reaction was allowed to stirred for 30 h under H<sub>2</sub>. Subsequently, the mixture was filtered through a pad of silica with EtOH (50 mL). The resulting mixture was concentrated *in vacuo* to yield (*R*)-28 (35 mg, 0.12 mmol, 91% yield). **<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD):** δ 4.13 (dd, *J* = 4.5, 7.9 Hz, 1H), 3.72 (s, 3H), 3.54 (t, *J* = 6.7 Hz, 2H), 1.77-1.69 (m, 1H), 1.68-1.58 (m, 1H), 1.56-1.49 (m, 2H), 1.45-1.24 (m, 22H) ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>17</sub>H<sub>34</sub>O<sub>4</sub>Na: 325.24; found 325.1.

(G) 2*R*,16-Dihydroxyhexadec-(6*E*)-enoic acid ((*R*)-29):

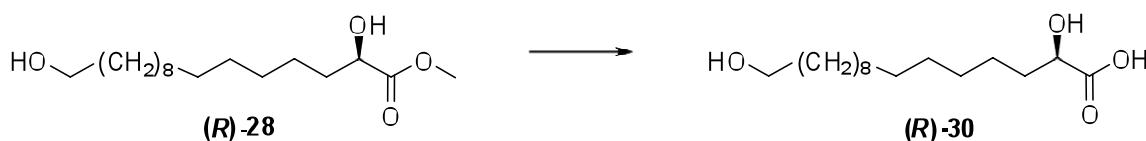


To a stirred solution of **(2R,6E)-27** (24 mg, 0.08 mmol) in a mixture of THF:MeOH:H<sub>2</sub>O (2:2:1, {v:v:v}, 3 mL) LiOH (48 mg, 2.0 mmol, 25 equiv) was added at r.t. After stirring for 5 h, the reaction was acidified to pH = 1 using 0.25 M aqueous HCl and extracted with EtOAc (3 × 6 mL). The combined organic extracts were washed with 0.25 M aqueous HCl (6 mL) and brine (6 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (MeOH:DCM with 0.2% AcOH, 1:9) to yield **(R)-29** as a white solid (20 mg, 0.07 mmol, 87% yield). **<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD):** δ 5.46-5.33 (m, 2H), 4.10 (dd, *J* = 4.3, 7.8 Hz, 1H), 3.54 (t, *J* = 6.7 Hz, 2H), 2.11-1.95 (m, 4H), 1.83-1.72 (m, 1H), 1.68-1.59 (m, 1H), 1.56-1.44 (m, 4H), 1.39-1.28 (m, 12H) ppm.

**<sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD):** δ 177.7, 131.8, 130.6, 71.1, 62.7, 34.6, 33.4, 33.1, 30.4, 30.4-30.2 (5C), 26.7, 25.9 ppm.

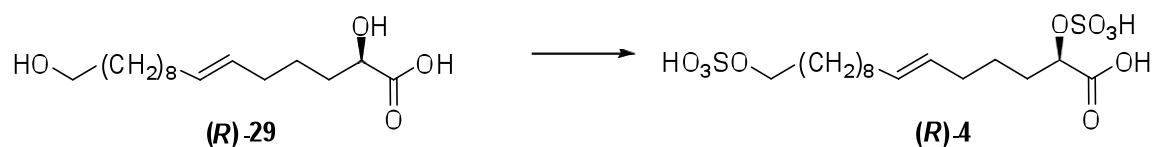
ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>16</sub>H<sub>30</sub>O<sub>4</sub>Na: 309.20; found 309.2. [α]<sub>D</sub><sup>22</sup> = -0.4 (c 0.02, methanol).

**(H) 2*R*,16-Dihydroxyhexadecanoic acid ((*R*)-30):**



To stirred solution of (*R*)-28 (35 mg, 0.116 mmol) in a mixture of THF:MeOH:H<sub>2</sub>O 2:2:1 {v:v:v} (4.5 mL) LiOH (71 mg, 2.96 mmol, 25.5 equiv) was added at r.t. After stirring for 5 h, the reaction was acidified to pH = 1 using 0.25 M aqueous HCl and extracted with EtOAc (3×10 mL). The combined organic extracts were washed with 0.25M HCl (10 mL) and brine (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated *in vacuo*, and purified by silica gel flash chromatography (MeOH:DCM with 0.2% AcOH, 1:9) to yield (*R*)-30 as a white solid (29 mg, 0.10 mmol, 87% yield). <sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD): δ 4.09 (dd, *J* = 4.4, 7.8 Hz, 1H), 3.54 (t, *J* = 6.7 Hz, 2H), 1.80-1.72 (m, 1H), 1.68-1.59 (m, 1H), 1.57-1.49 (m, 2H), 1.48-1.16 (m, 22H) ppm. <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD): δ 177.8, 71.2, 62.73, 35.2, 35.1, 33.4, 30.4-30.2 (5C), 30.31, 30.25, 30.20, 26.7, 25.9 ppm. ESI-MS (*m/z*): [M+Na<sup>+</sup>] calcd. for C<sub>16</sub>H<sub>32</sub>O<sub>4</sub>Na: 311.22; found 311.2. [α]<sup>22</sup><sub>D</sub> = -1.2 (c 0.01, methanol).

**(I) 2*R*,16-Disulfooxyhexadec-(6*E*)-enoic acid ((*R*)-4):**



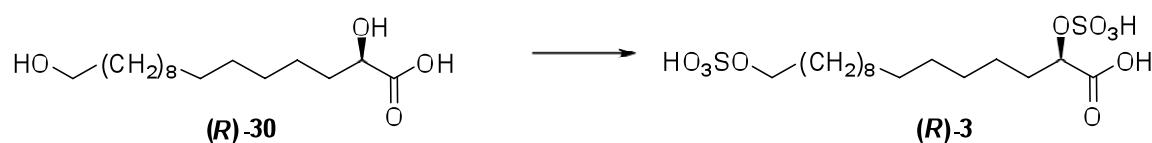
A Schlenk flask was flame-dried under vacuum and charged with argon. The acid (***R*-29**) (20 mg, 0.07 mmol) and dry DMF (3 mL) were added to the flask. In a separate Schlenk flask, a solution of SO<sub>3</sub>DMF complex (290 mg) in dry DMF (2 mL) was prepared. A portion of the SO<sub>3</sub>DMF solution (0.75 mL, 0.7 mmol, 10.0 equiv) was added to the alcohol/DMF mixture at 0 °C and the resulting mixture was stirred for 30 min. The reaction was then allowed to warm to r.t. and stirred for an additional 20 min. Subsequently, the reaction flask was brought back to 0 °C and quenched by addition of a solution of KHCO<sub>3</sub> (142 mg, 1.42 mmol, 20.3 equiv) in distilled H<sub>2</sub>O (1 mL). The resulting yellow mixture was checked to be at pH = 6, concentrated *in vacuo*, and purified by silica gel flash chromatography (DCM with 1% AcOH:MeOH, 1:1) to yield (***R*-4**) as a white solid (17 mg, 0.04 mmol, 54% yield). **<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD):** δ 5.45-5.34 (m, 2H), 4.69-4.58 (m, 1H), 3.99 (t, *J* = 6.6 Hz, 2H), 2.19-1.94 (m, 4H), 1.91-1.82 (m, 1H), 1.82-1.73 (m, 1H), 1.69-1.62 (m, 2H), 1.57-1.49 (m, 2H), 1.44-1.37 (m, 2H), 1.37-1.26 (m, 10H) ppm.

**<sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD):** δ 182.2, 131.3, 130.8, 80.4, 68.8, 33.6, 33.4, 33.1, 30.2, 30.4-30.1 (5C), 26.6, 26.1 ppm.

**ESI-MS (*m/z*):** [(M-2H)/2] calcd. for (C<sub>16</sub>H<sub>28</sub>O<sub>10</sub>S<sub>2</sub>)/2: 222.06; found 222.9. [α]<sup>22</sup><sub>D</sub> = +3.2 (c 0.004, water).

**(S)-4** was prepared in the same manner as (***R*-4**) starting with (***R*-(+)**-glycidol. [α]<sup>22</sup><sub>D</sub> = -1.2 (c 0.005, water)

**(J) 2*R*,16-Disulfooxyhexadecanoic acid ((*R*)-3):**

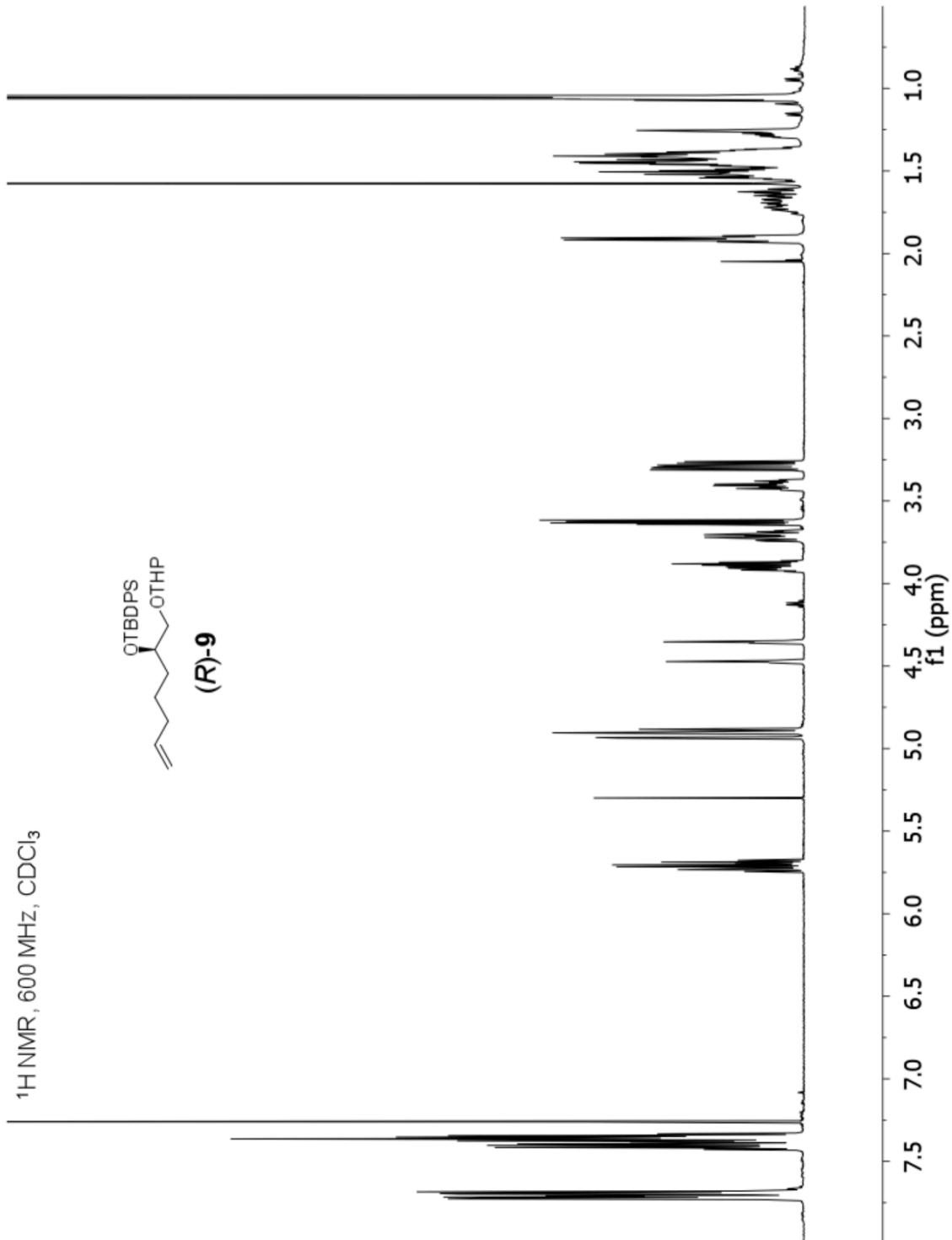
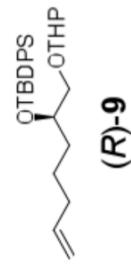


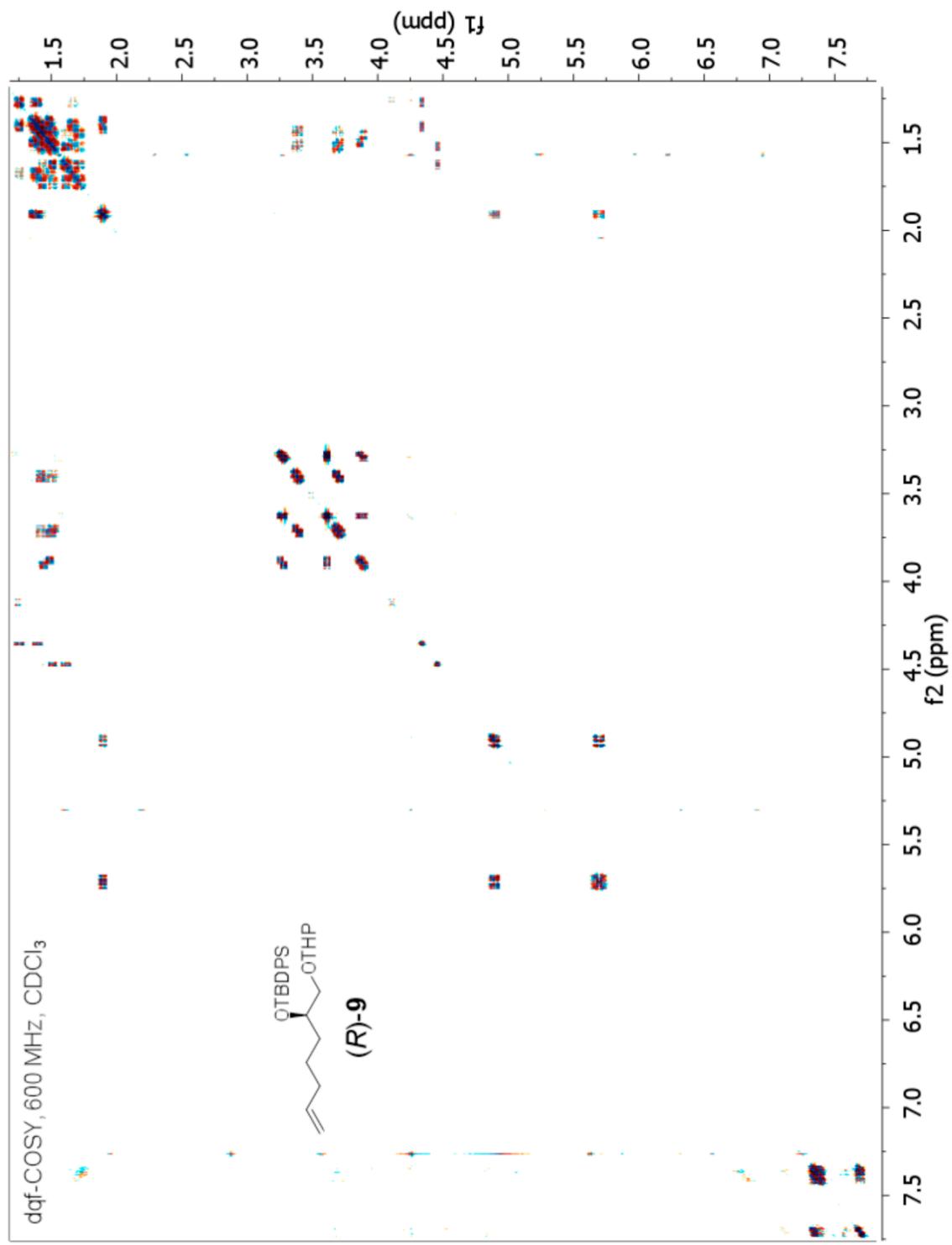
A Schlenk flask was flame-dried under vacuum and charged with argon. The acid (*R*)-30 (16 mg, 0.06 mmol) and dry DMF (3 mL) were added to the flask. In a separate Schlenk flask, a solution of SO<sub>3</sub>DMF complex (300 mg) in dry DMF (2 mL) was prepared. A portion of the SO<sub>3</sub>DMF solution (0.6 mL, 0.59 mmol, 10.6 equiv) was added to the alcohol/DMF mixture at 0 °C and the resulting mixture was stirred for 30 min. The reaction was allowed to warm to r.t. and stirred for an additional 20 min. Subsequently, the reaction flask was brought back to 0 °C and quenched by addition of a solution of KHCO<sub>3</sub> (112 mg, 1.12 mmol, 20.0 equiv) in distilled H<sub>2</sub>O (0.56 mL). The resulting yellow mixture was checked to be at pH = 6, concentrated *in vacuo*, and purified by silica gel flash chromatography (DCM with 1% AcOH : MeOH, 1:1) to yield (*R*)-3 as a white solid (17 mg, 0.04 mmol, 68% yield). **<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD):** δ 4.66-4.58 (m, 1H), 3.99 (t, *J* = 6.6 Hz, 2H), 1.89-1.81 (m, 1H), 1.81-1.73 (m, 1H), 1.69-1.62 (m, 2H), 1.51-1.26 (m, 22H) ppm. **<sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD):** δ 182.2, 80.2, 68.8, 33.9, 30.6-30.2 (7C), 30.3, 30.2, 30.1, 26.6, 26.0 ppm. ESI-MS (*m/z*): [(M-2H)/2] calcd. for (C<sub>16</sub>H<sub>30</sub>O<sub>10</sub>S<sub>2</sub>)/2: 223.07; found 223.1. [(M-3H+Na<sup>+</sup>)/2] calcd. for (C<sub>16</sub>H<sub>29</sub>NaO<sub>10</sub>S<sub>2</sub>)/2: 234.06; found 234.0. [α]<sup>22</sup><sub>D</sub> = +1.9 (c 0.008, water)

### Bioassays with synthetic Caeliferins in *Arabidopsis*<sup>7</sup>

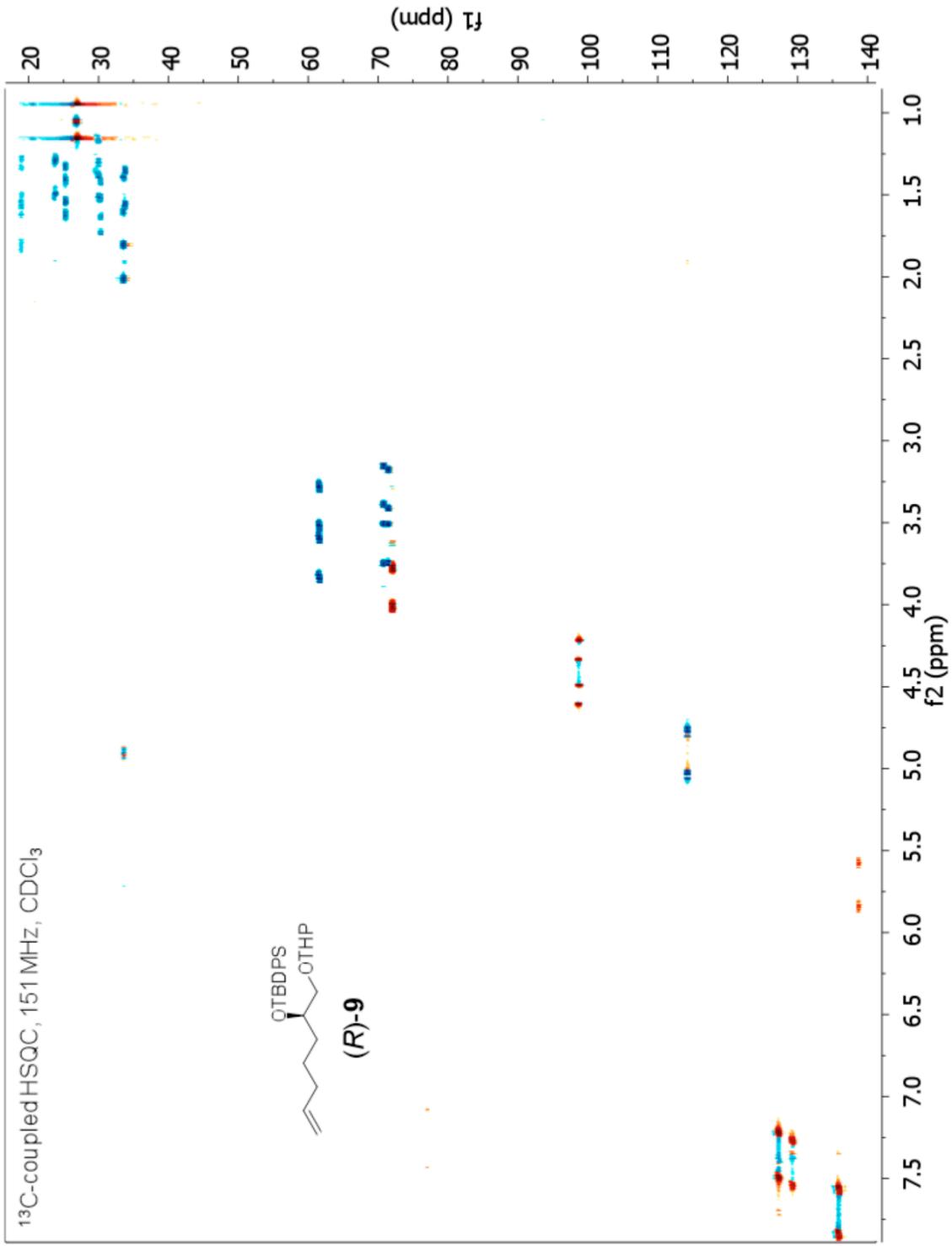
*Arabidopsis thaliana* Columbia (Col-O) were germinated in MetroMix® 200 (Sun Gro Horticulture Distribution, Inc) supplemented with 14-14-14 Osmocote (Scotts Miracle-Gro) . Plants were maintained in a walk-in growth chamber with a 12-h photoperiod, approximately  $130 \mu\text{mol}^{-2} \text{s}^{-1}$  of photosynthetically active radiation supplied by supplemental lighting, 65% relative humidity, and constant temperature of 21°C. One month old plants were treated with caeliferins at concentration of  $22 \text{ nmol } \mu\text{l}^{-1}$ . The adaxial sides of the 3 largest fully expanded leaves were scratched with a razor removing 5-10% of the waxy cuticle. The damage sites included the central leaf tip spanning both sides of the midrib and two midbasal sections on opposite sides of the midrib. Test solutions of  $5 \mu\text{l plant}^{-1}$  or 50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8) alone were immediately applied and dispersed over the damage sites. Leaves remained on the intact plants for 1 h before shoot excision, sampling and GC analyses. Excised shoots were sealed in 6 mL tubes for an additional 1 hr before headspace sampling as described<sup>7</sup>. All experiments were carried out in four independent replicates for each synthetic sample.

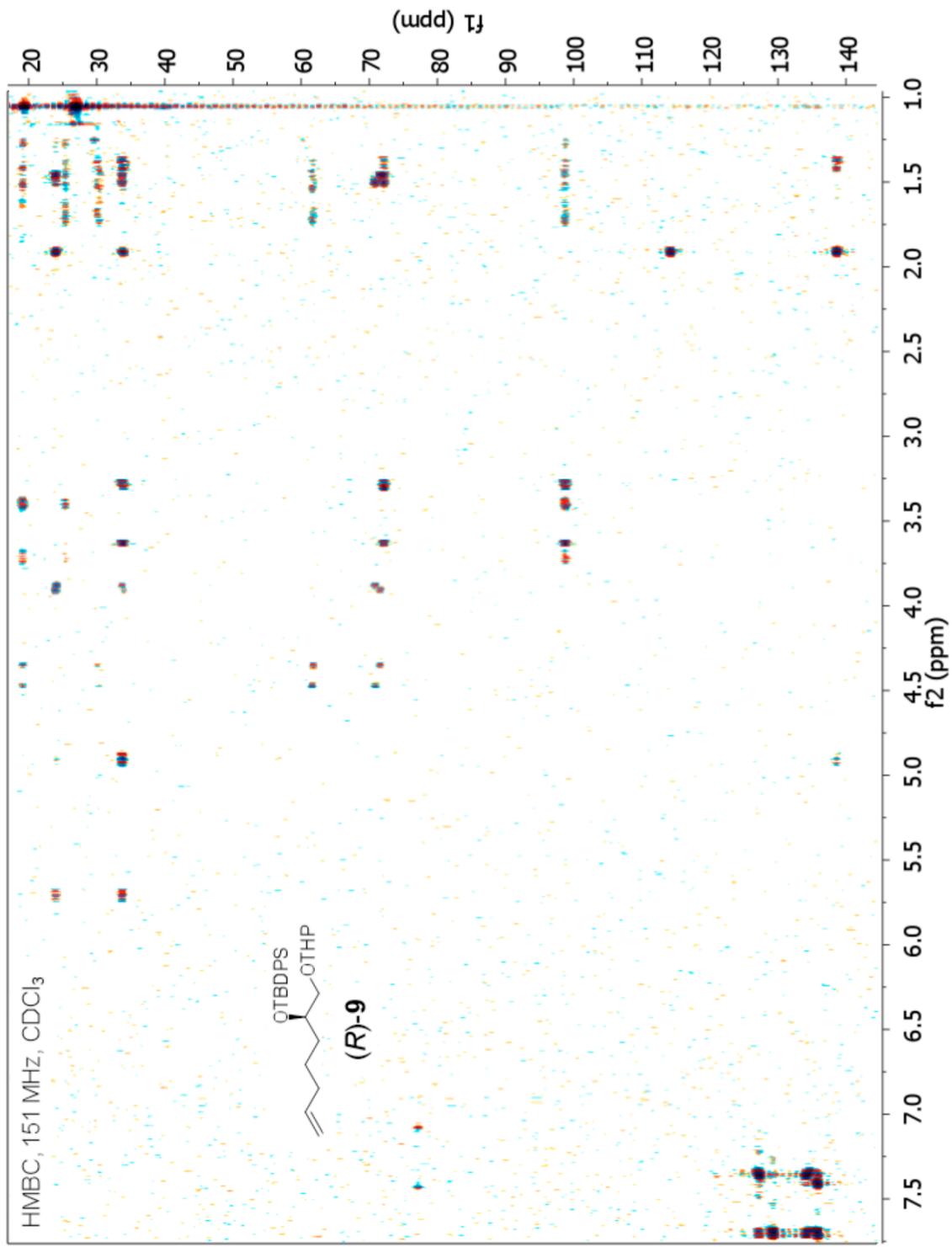
$^1\text{H}$  NMR, 600 MHz,  $\text{CDCl}_3$





<sup>13</sup>C-coupled HSQC, 151 MHz, CDCl<sub>3</sub>

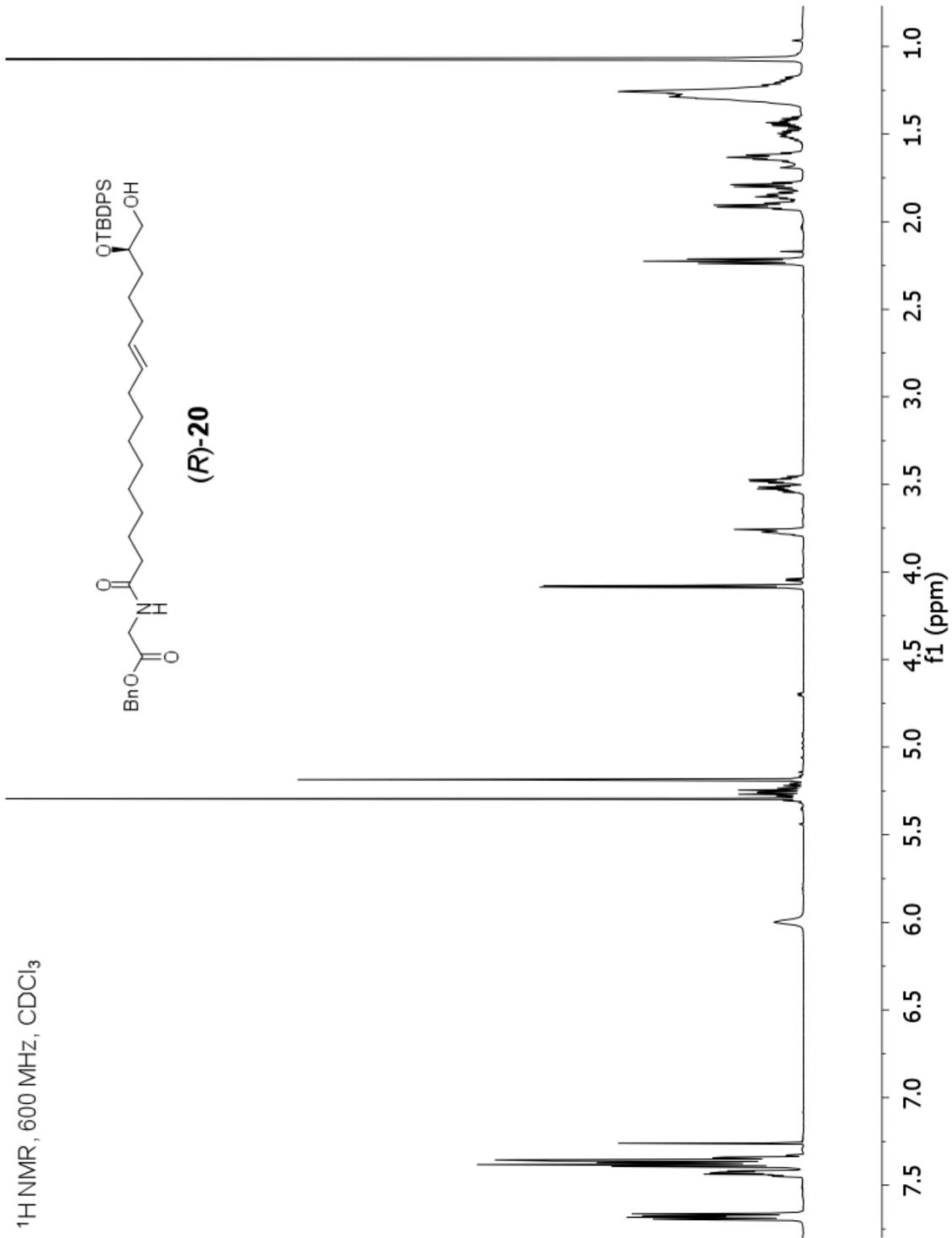


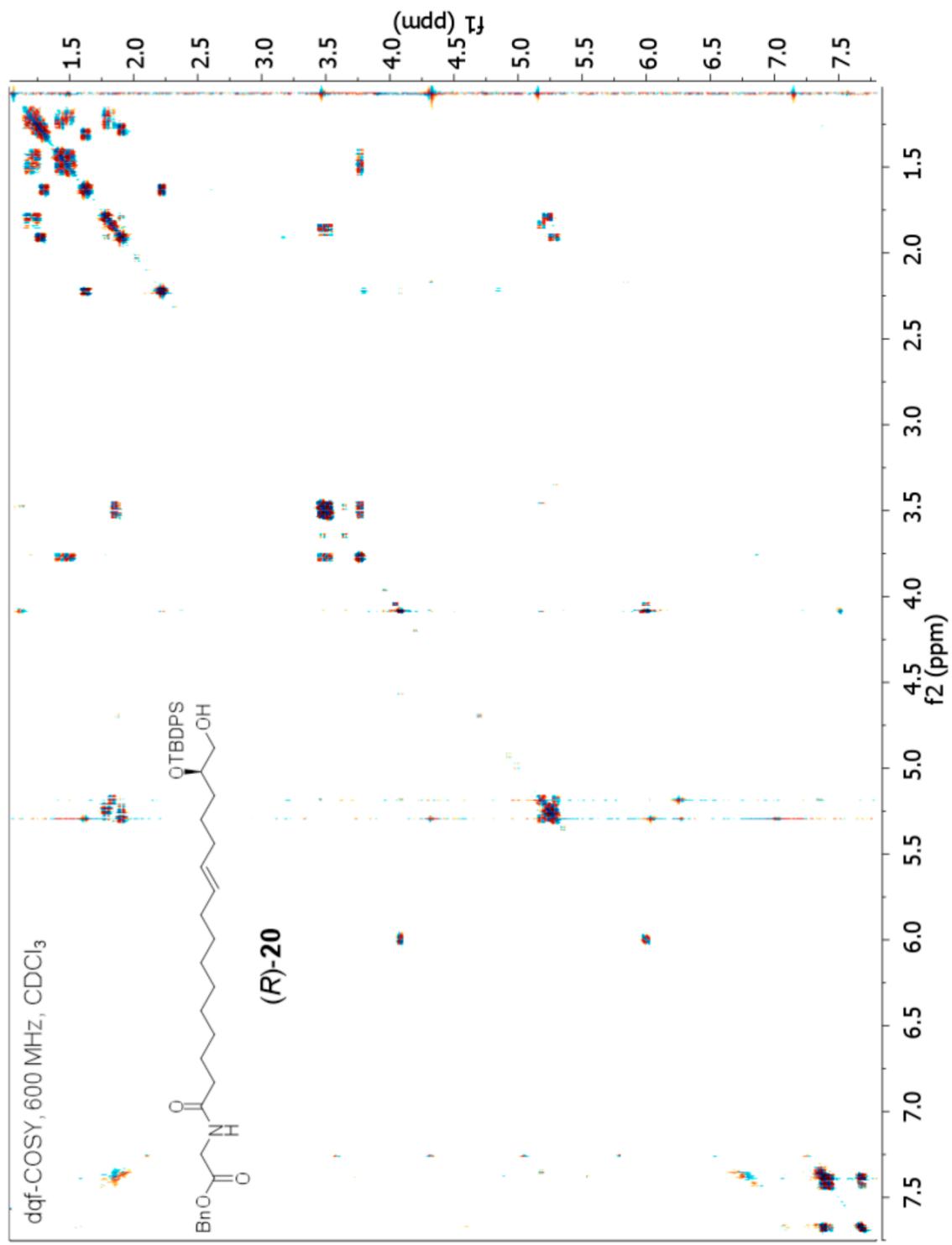


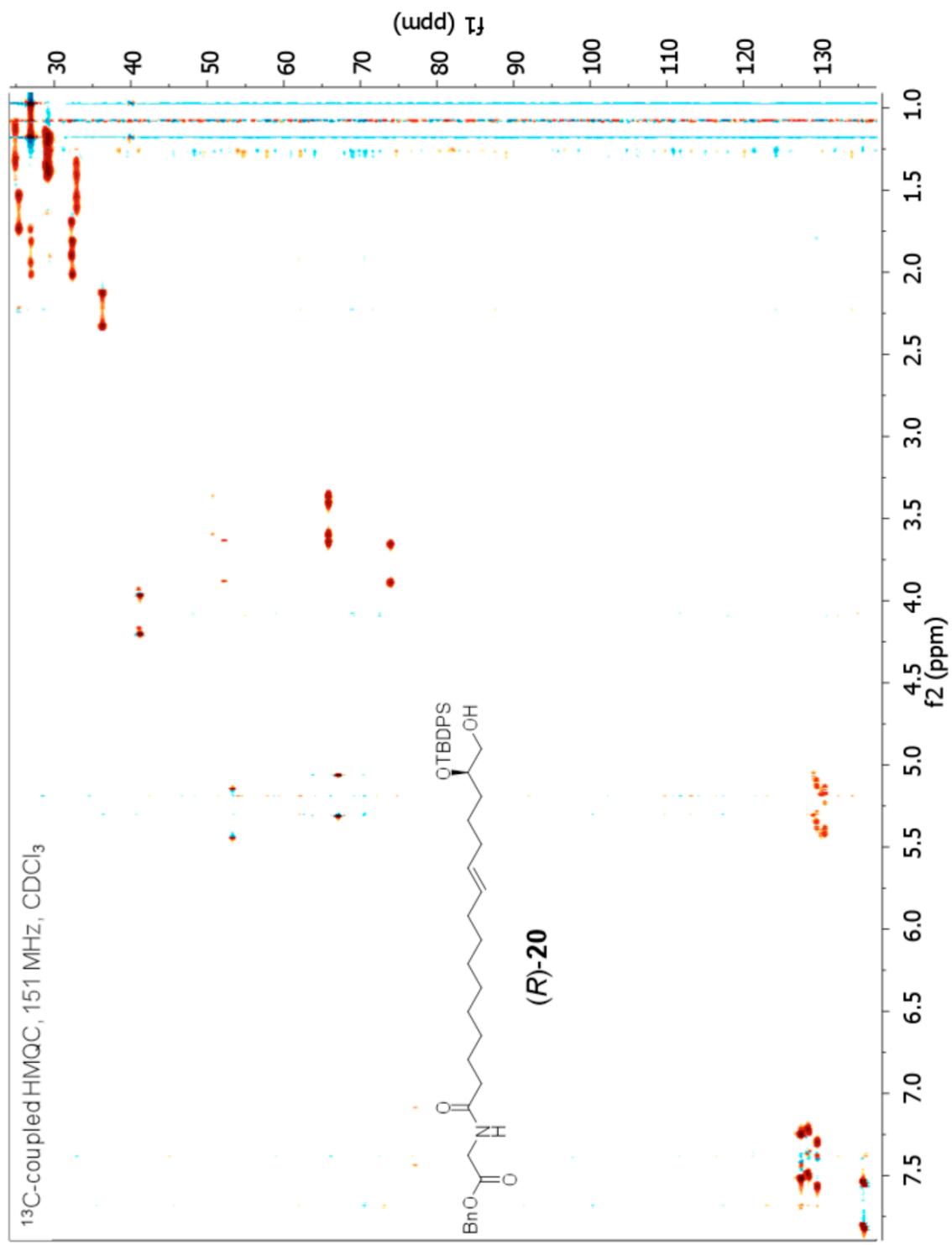
<sup>1</sup>H NMR, 600 MHz, CDCl<sub>3</sub>

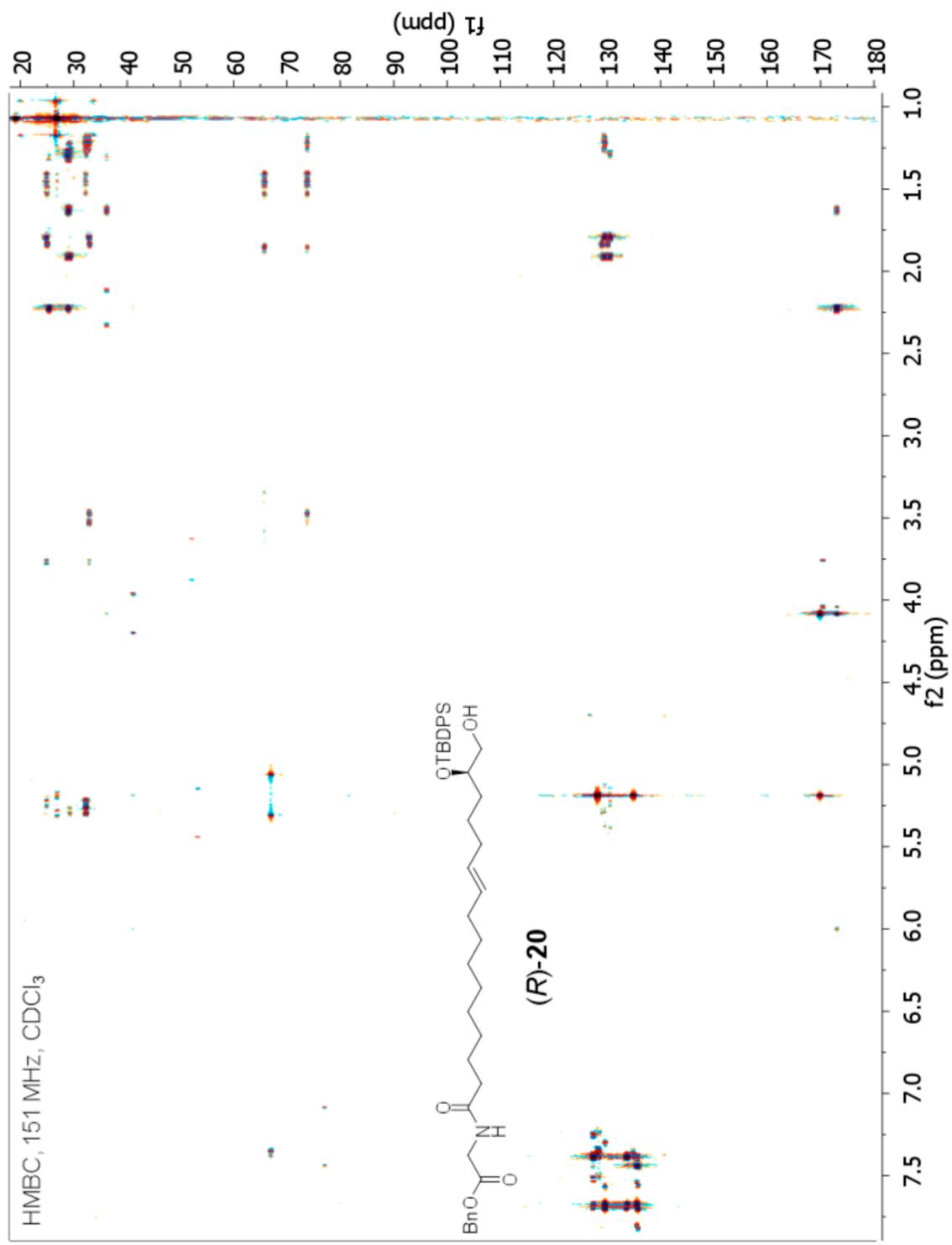


(R)-20

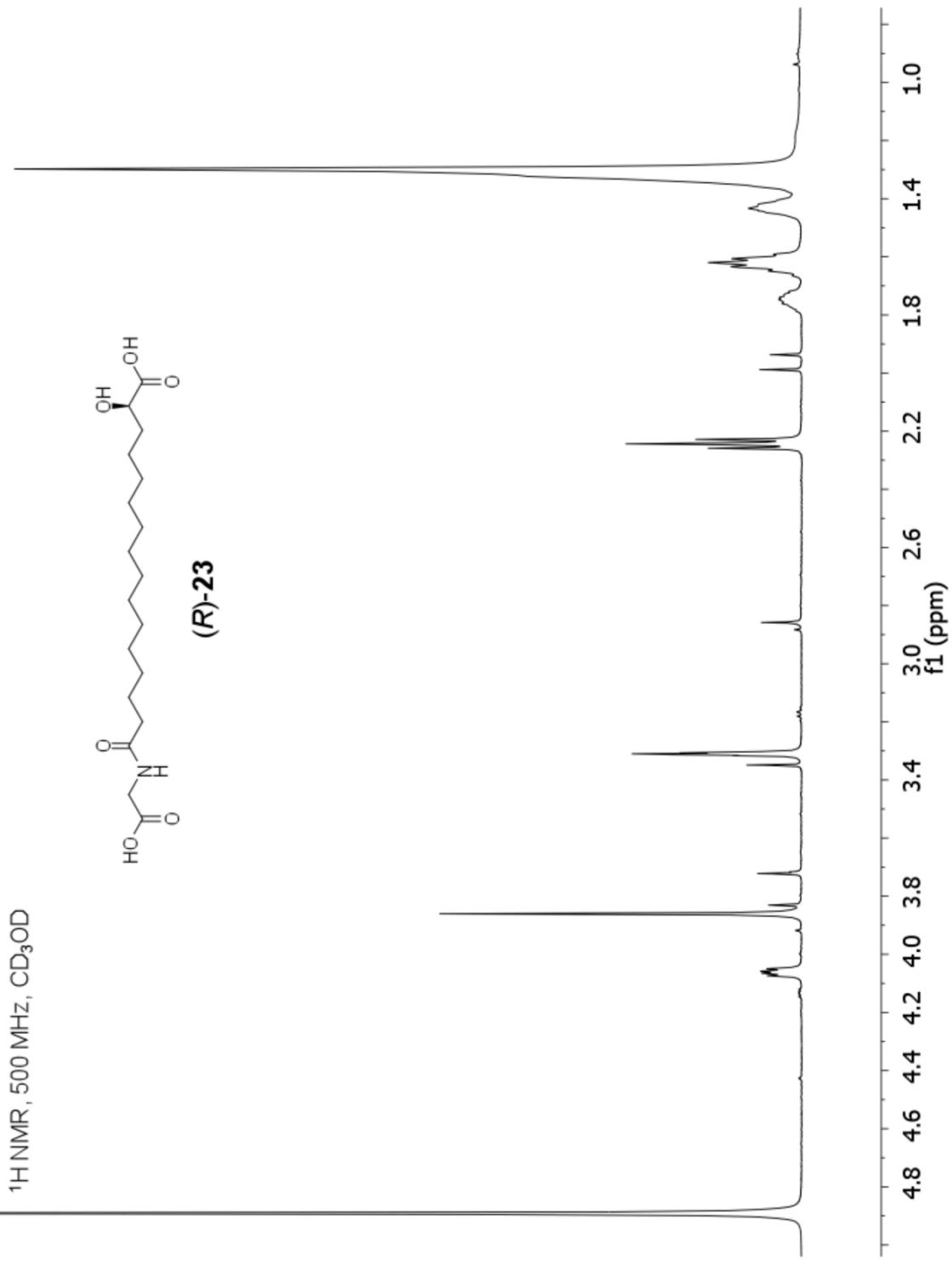
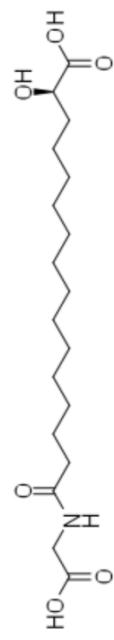


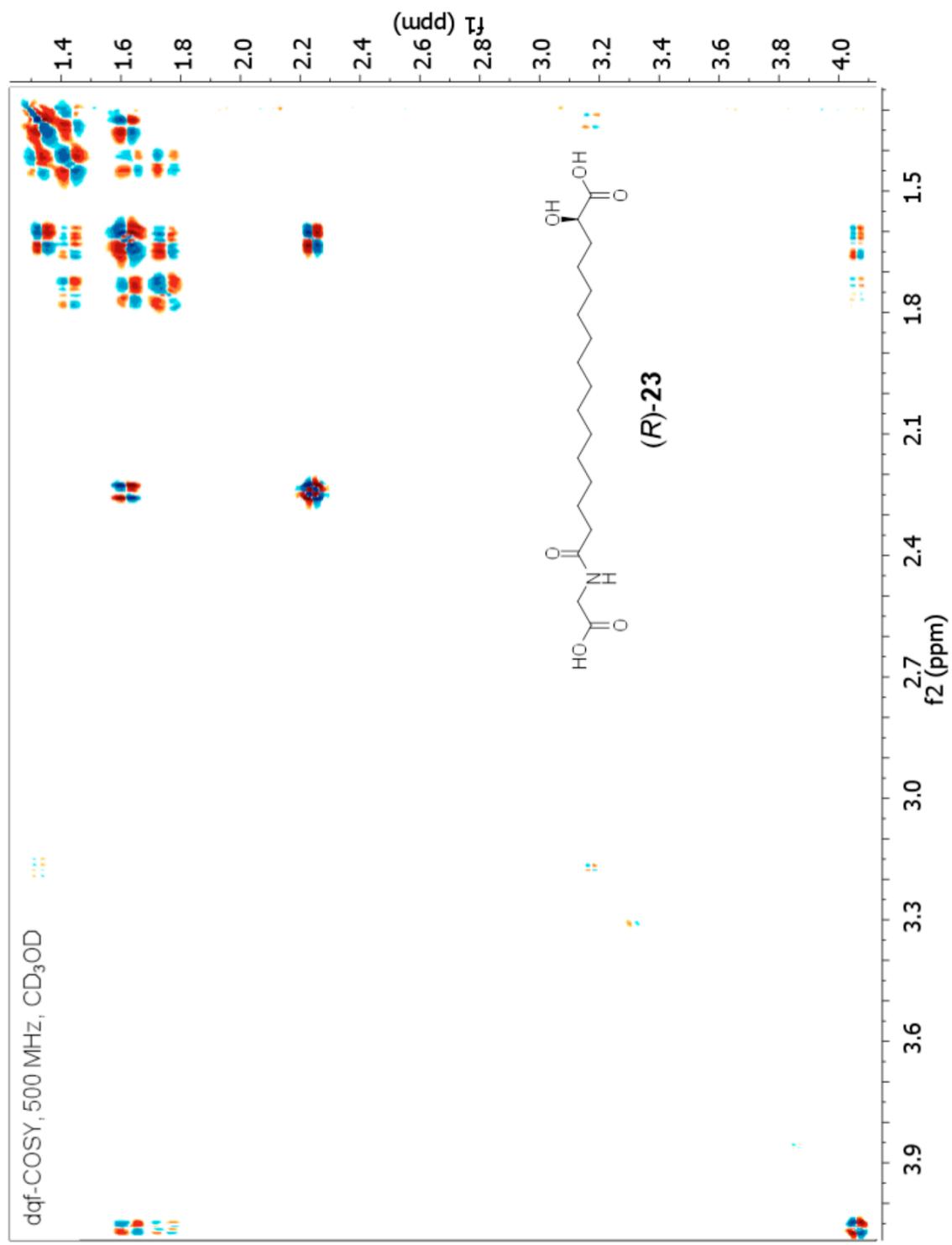


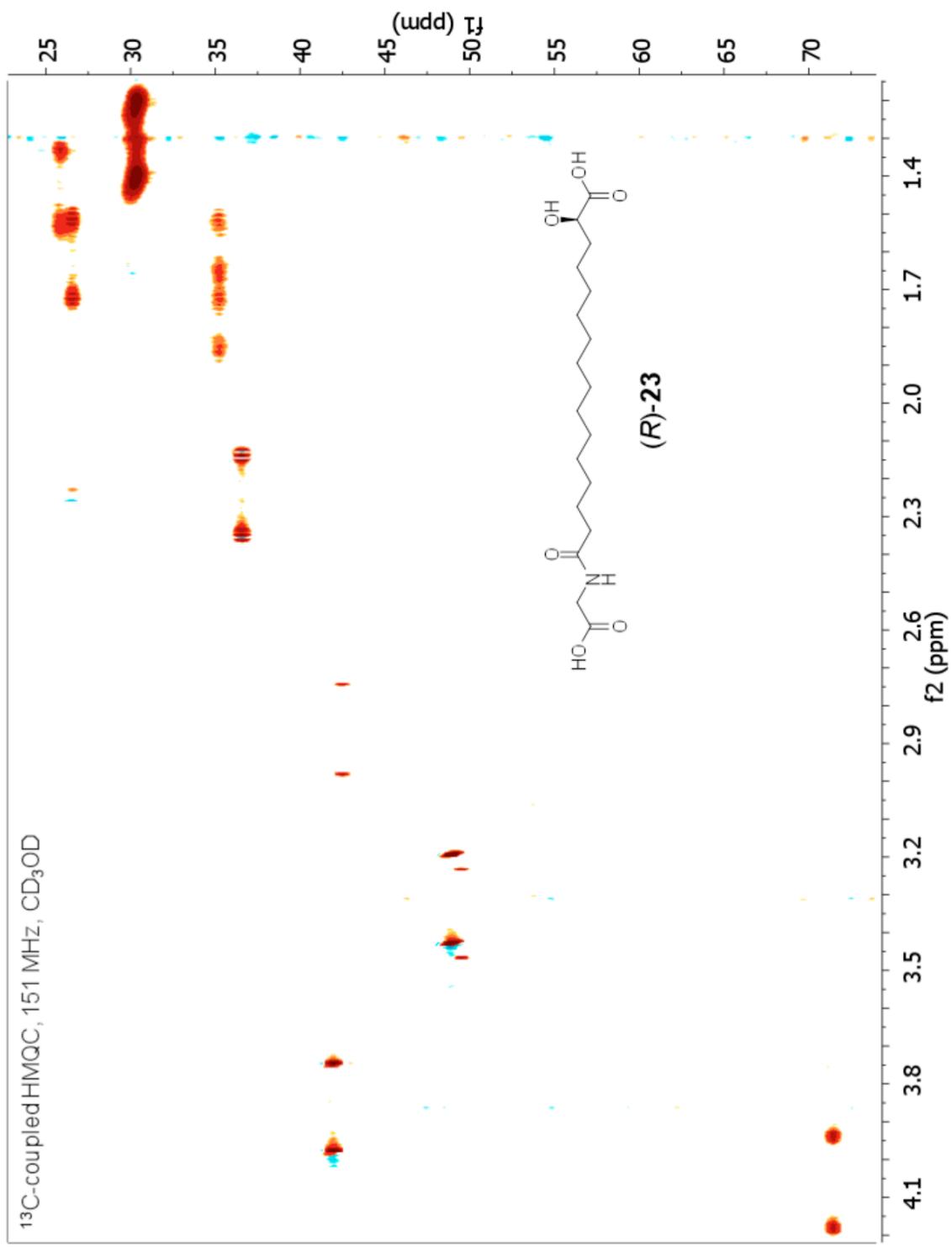


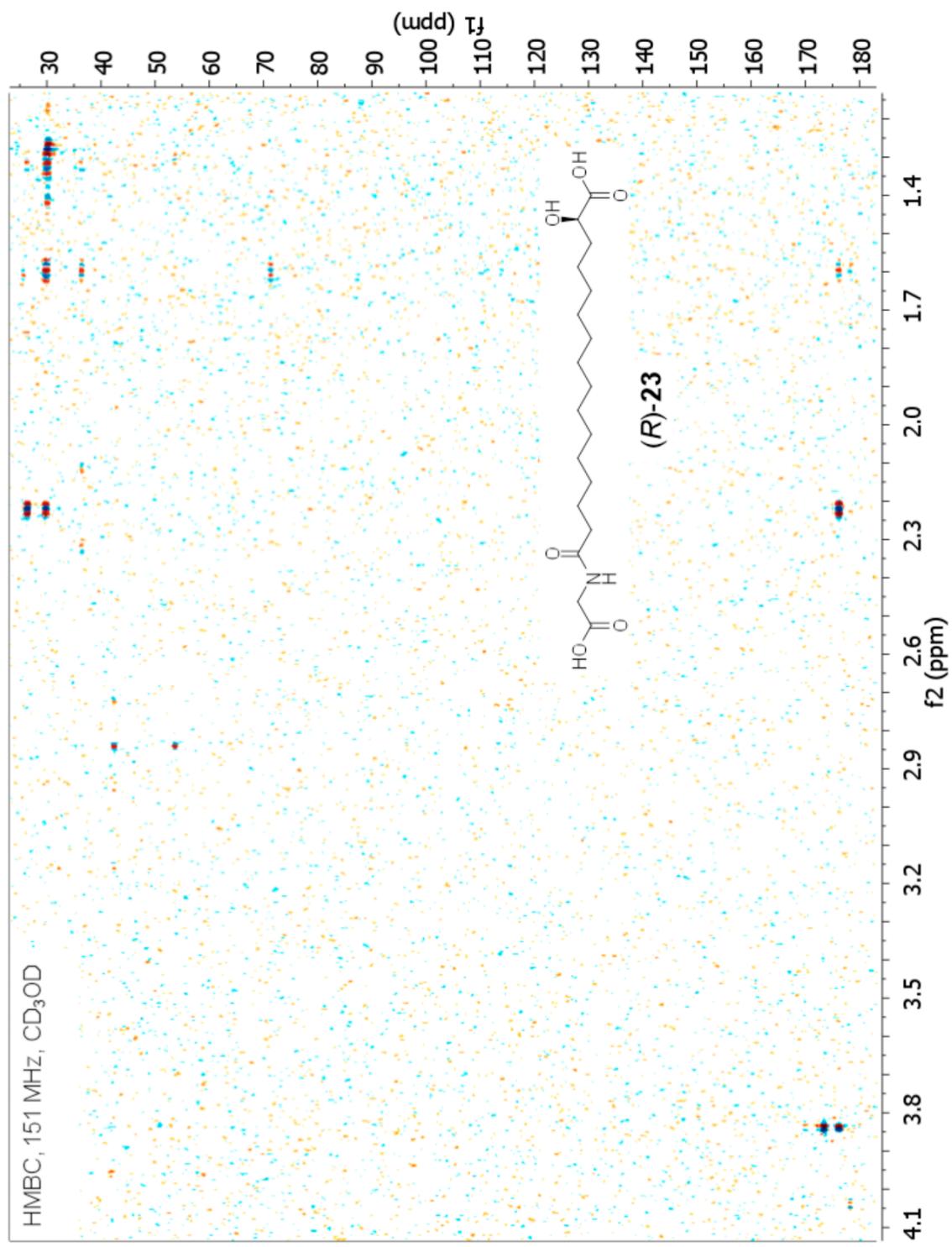


$^1\text{H}$  NMR, 500 MHz,  $\text{CD}_3\text{OD}$

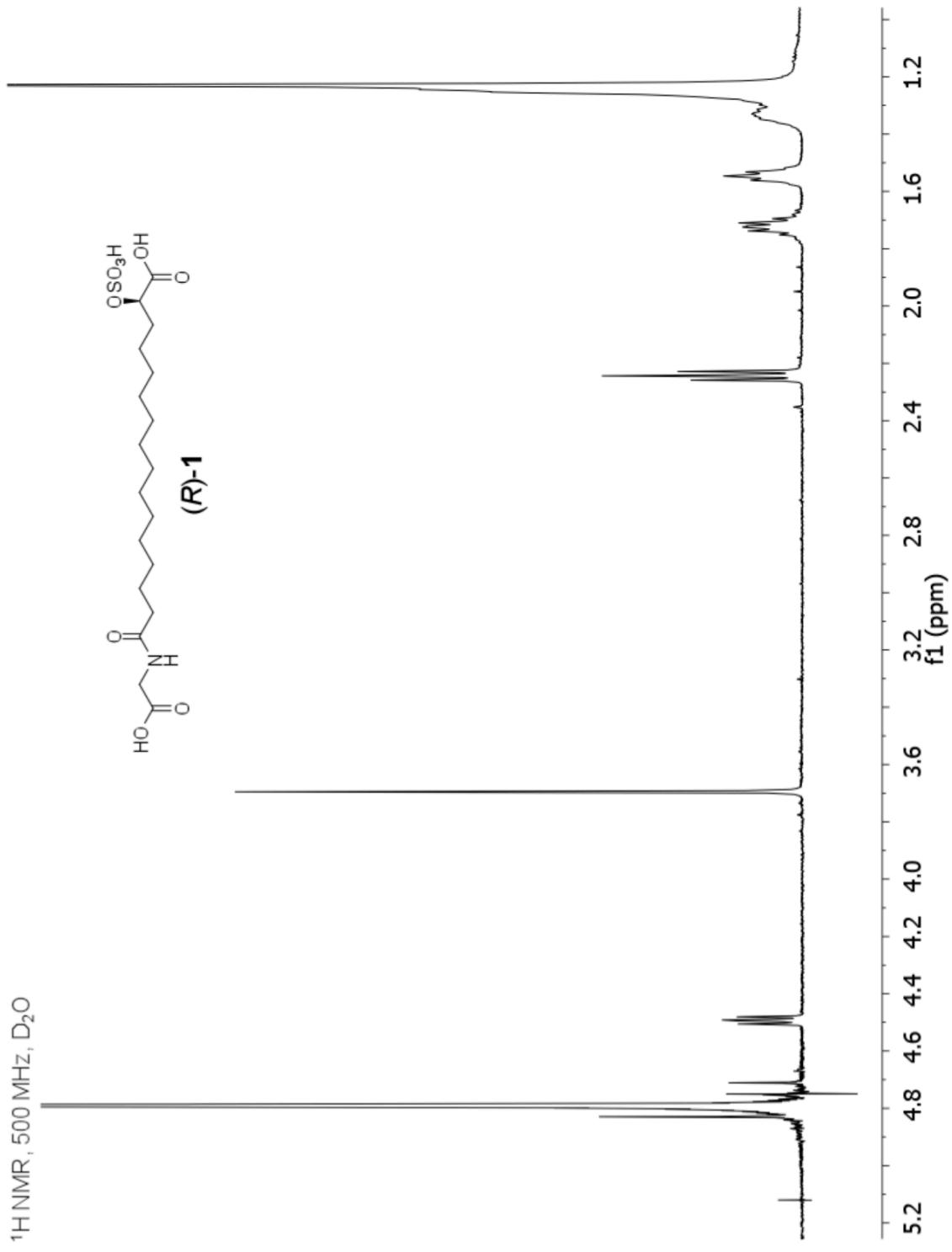


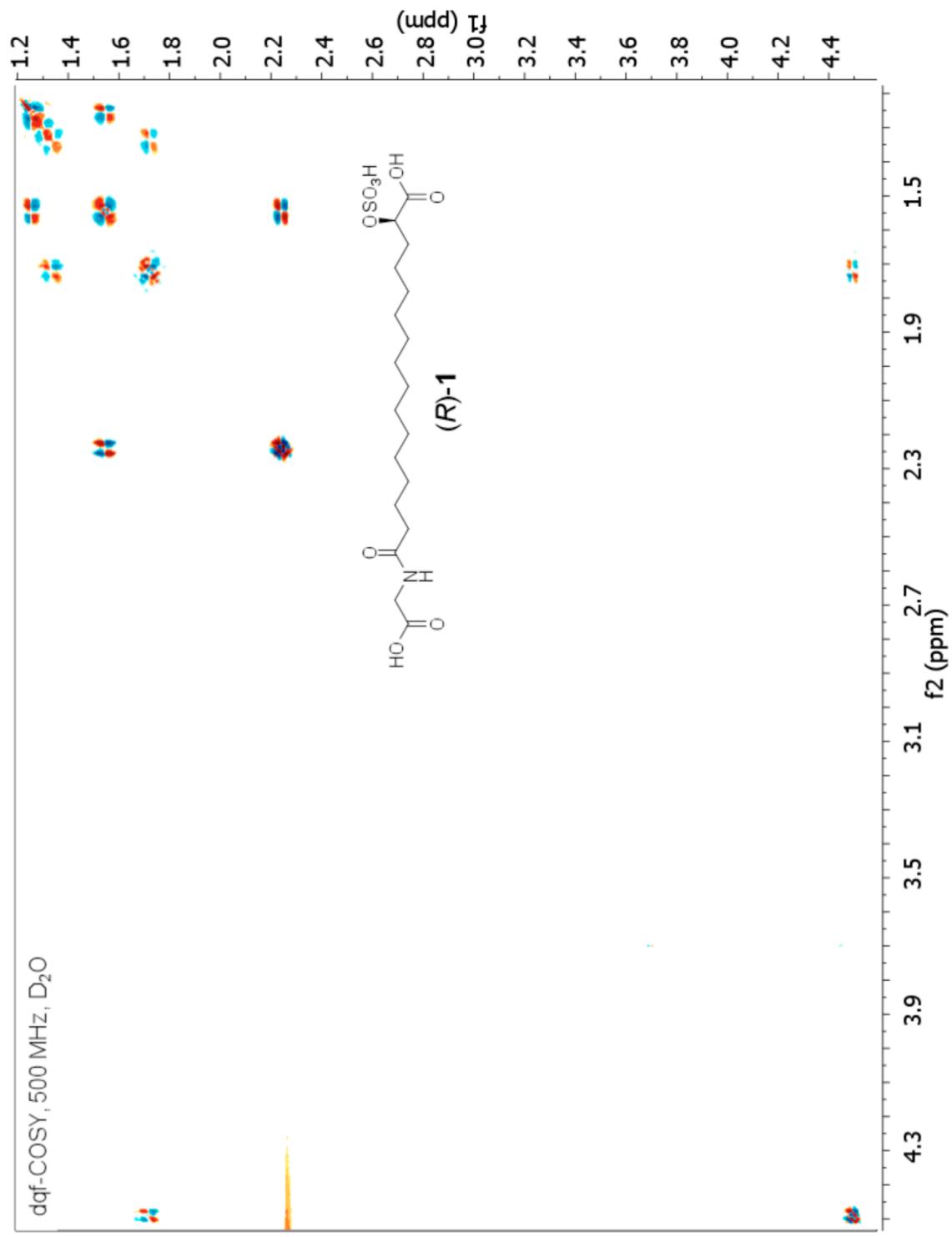




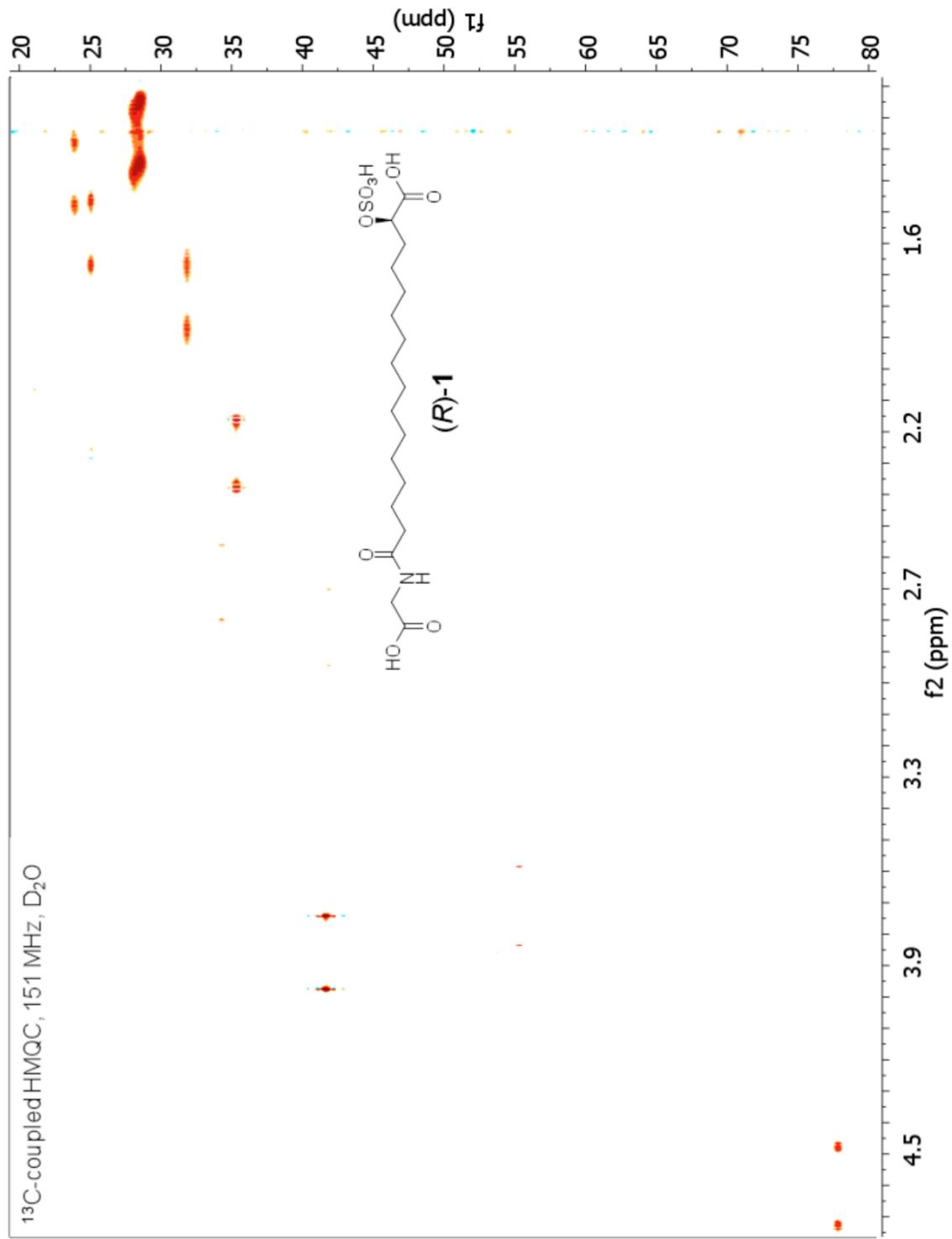


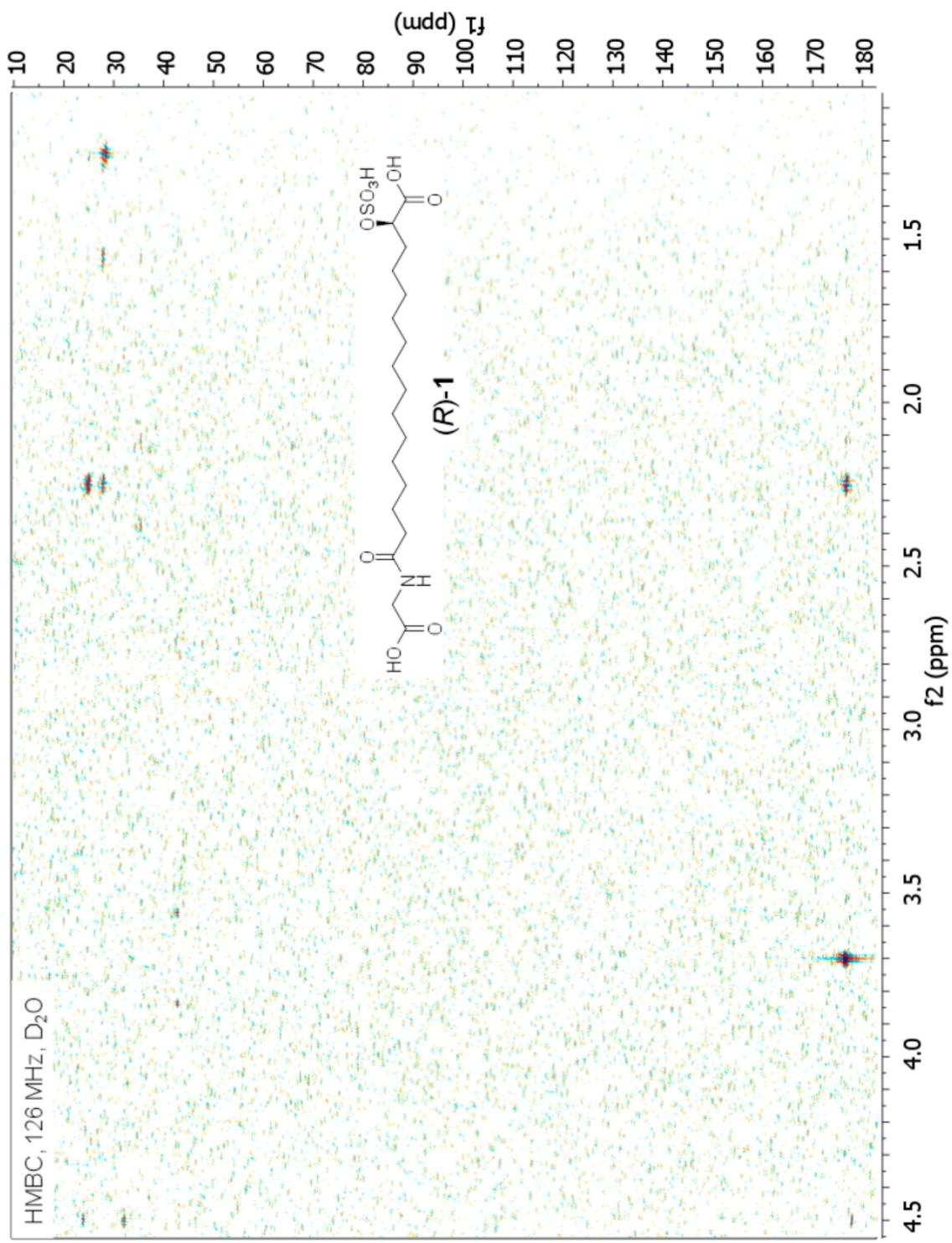
$^1\text{H}$  NMR, 500 MHz,  $\text{D}_2\text{O}$



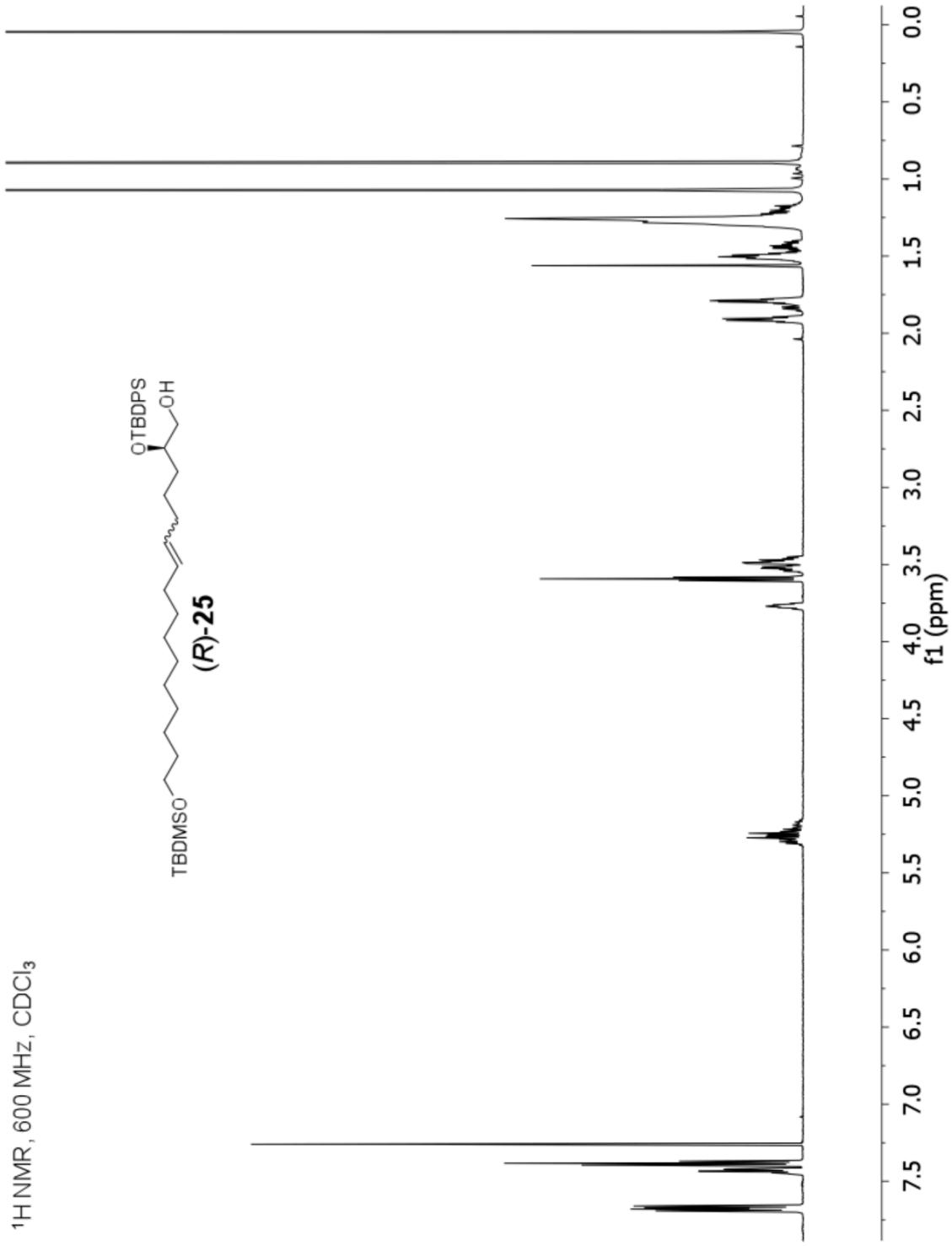


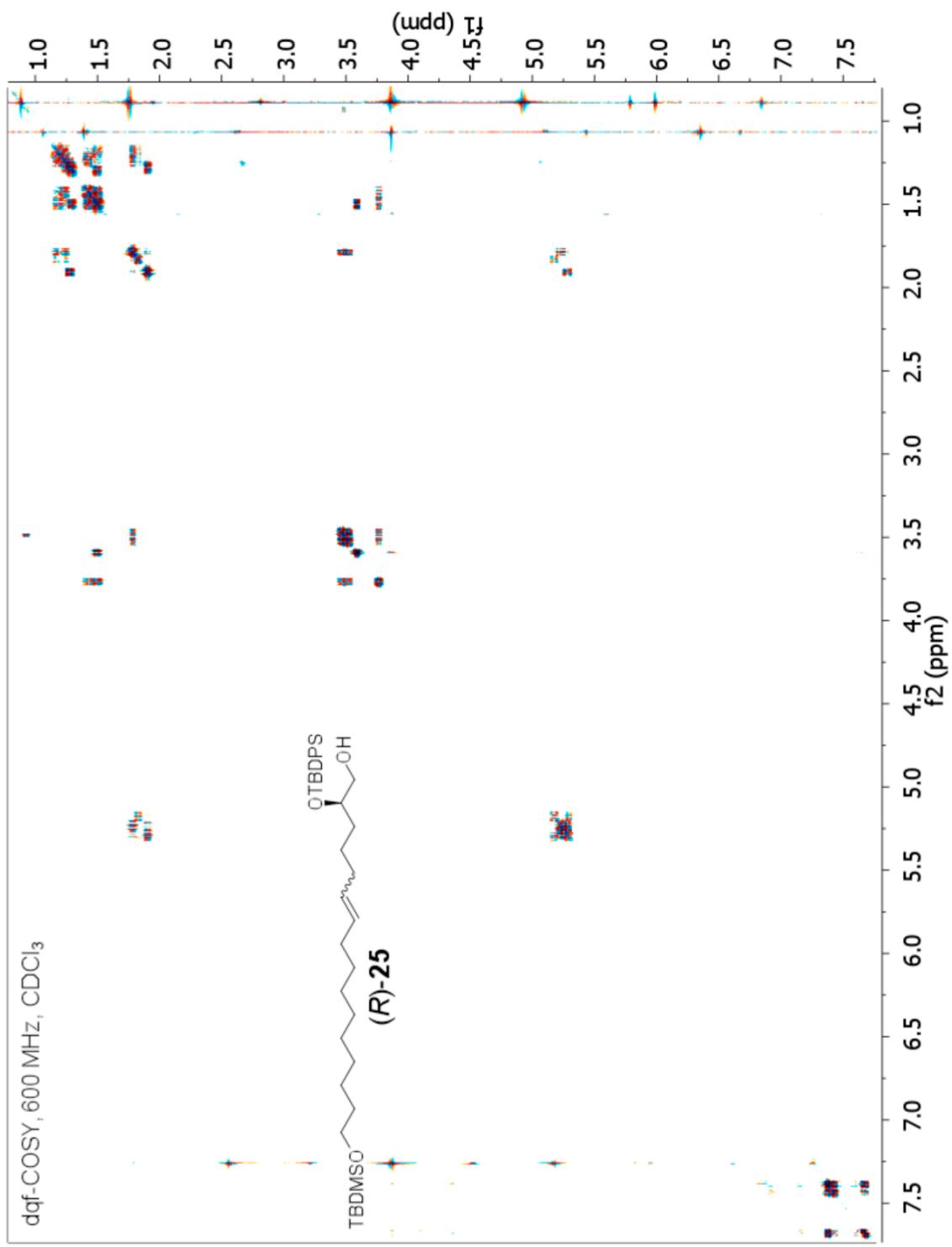
$^{13}\text{C}$ -coupled HMQC, 151 MHz,  $\text{D}_2\text{O}$

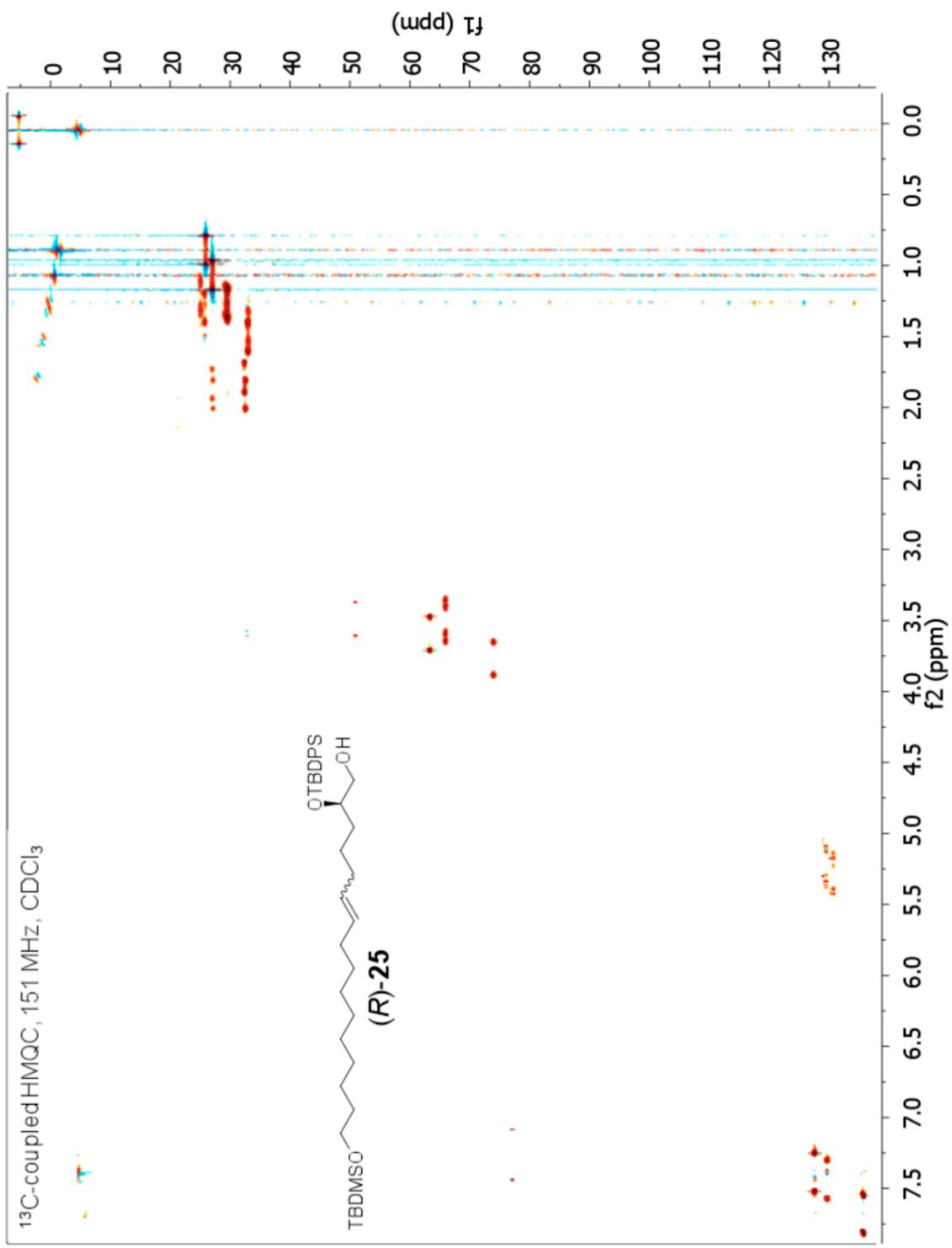


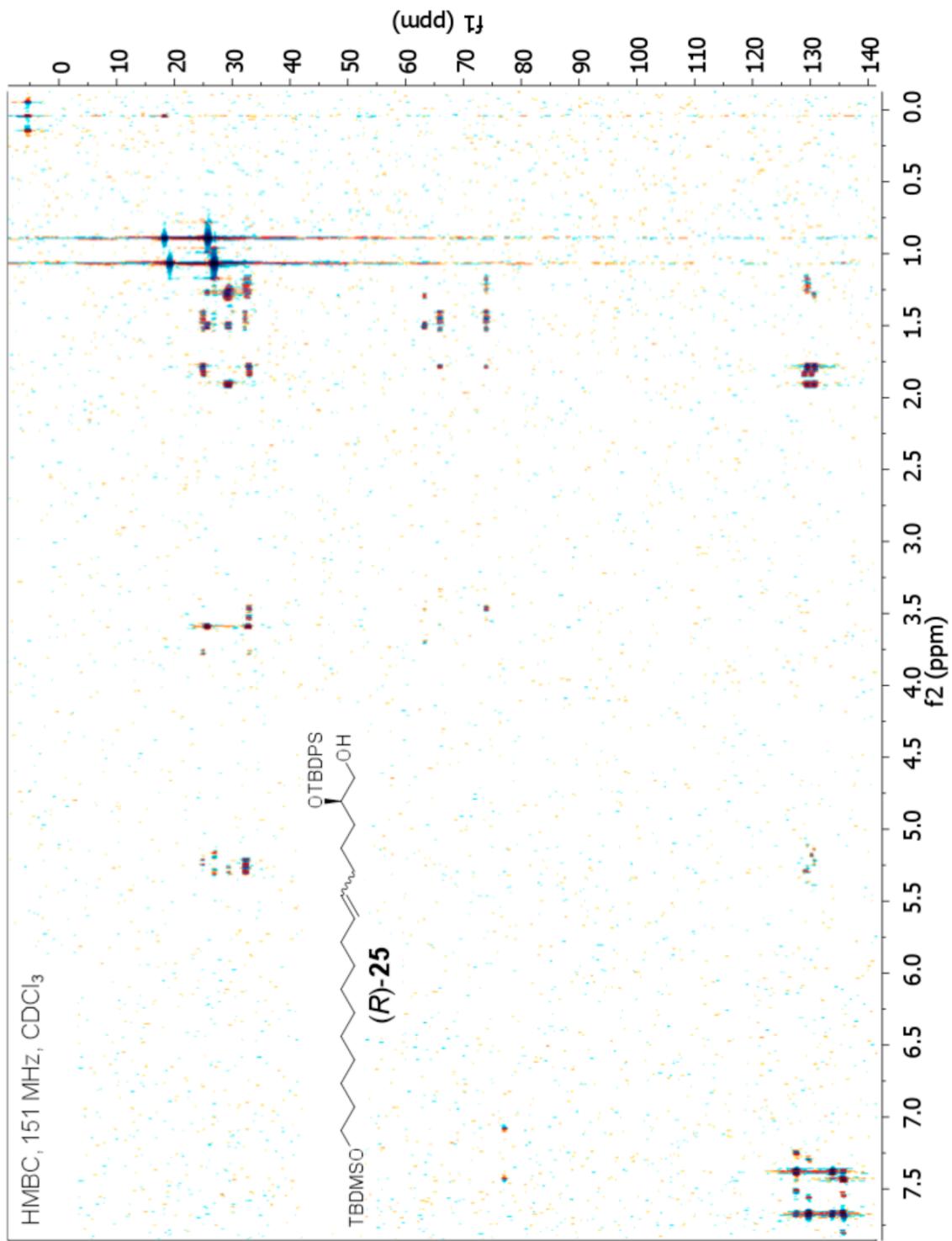


$^1\text{H}$  NMR, 600 MHz,  $\text{CDCl}_3$

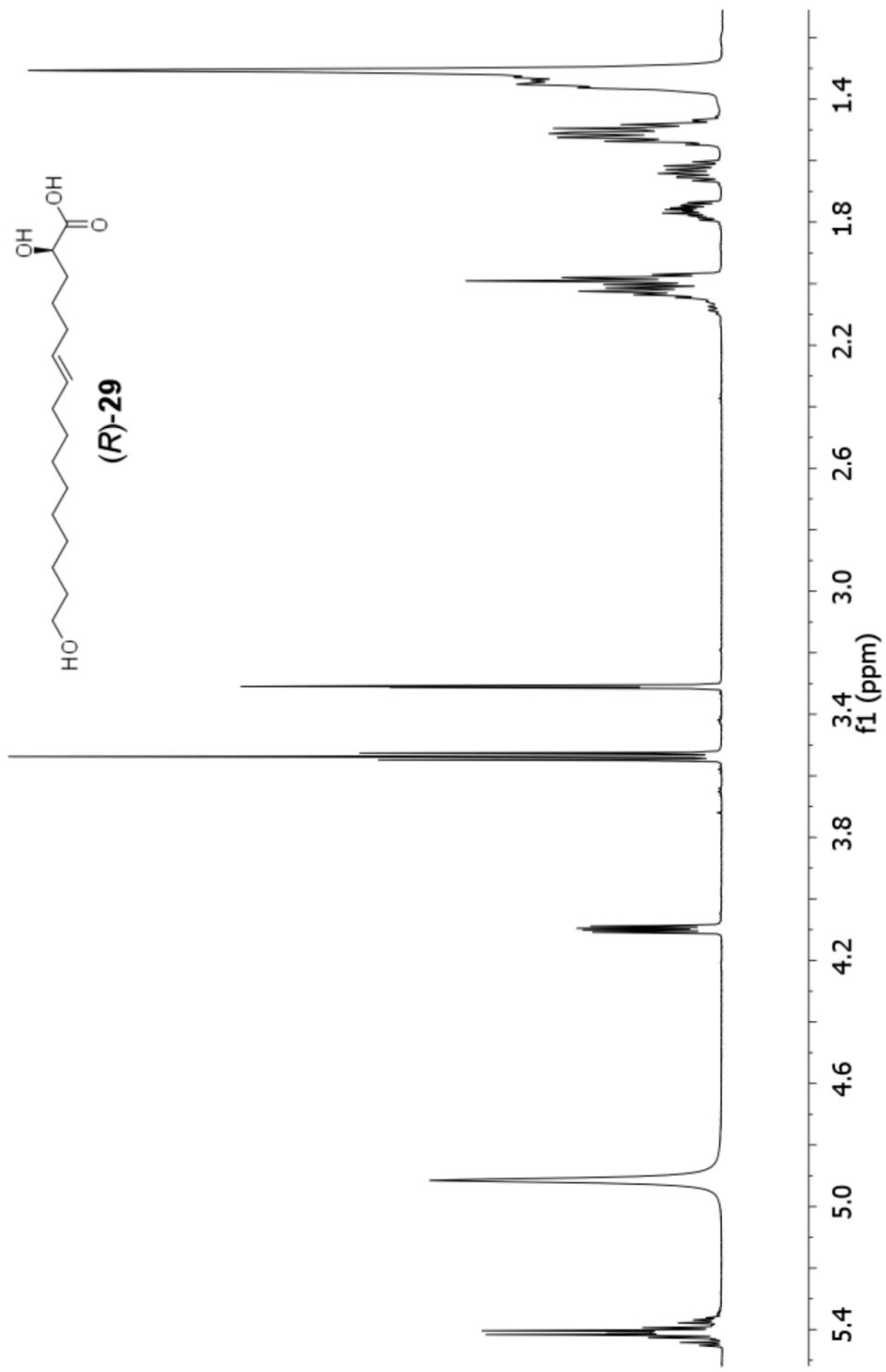


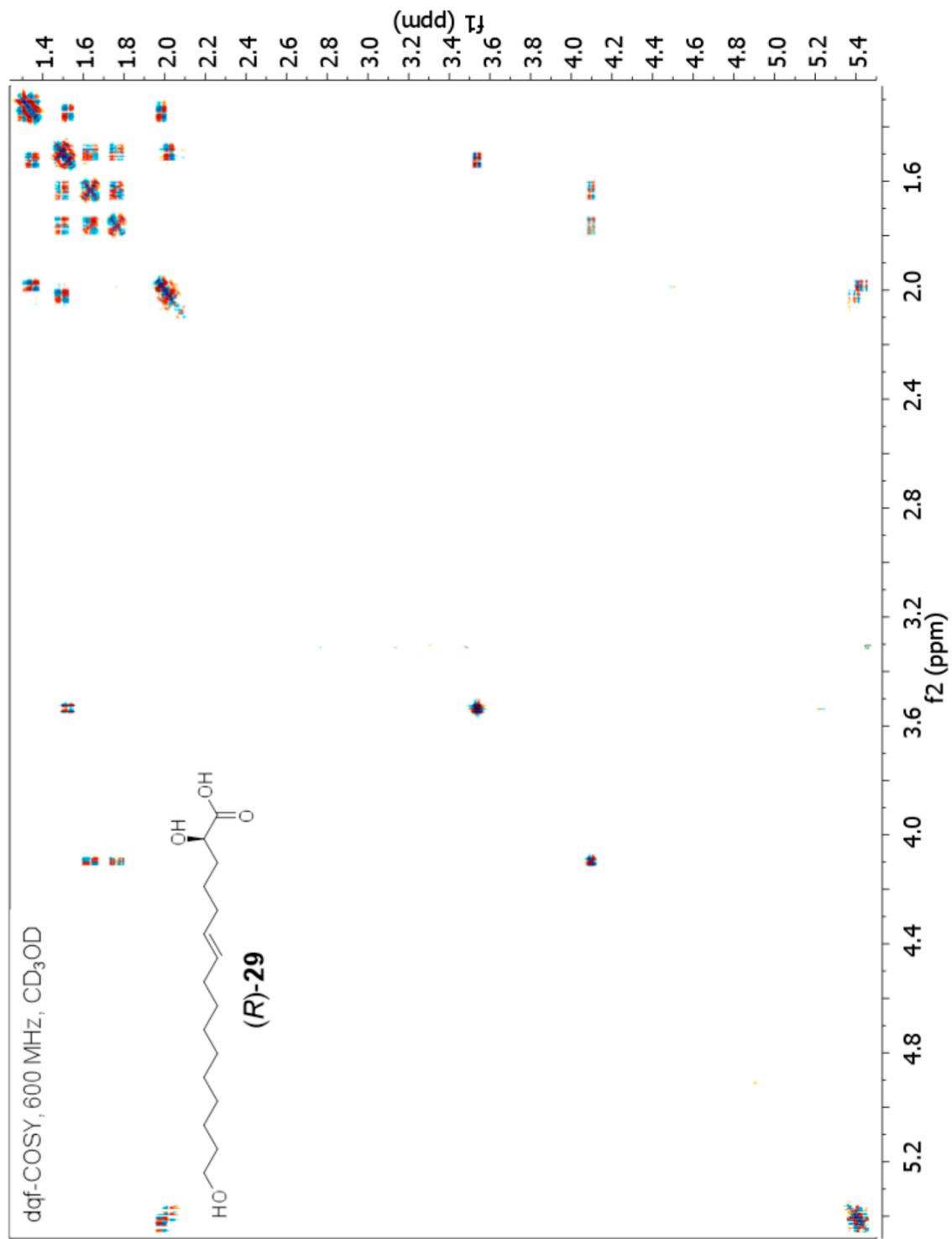


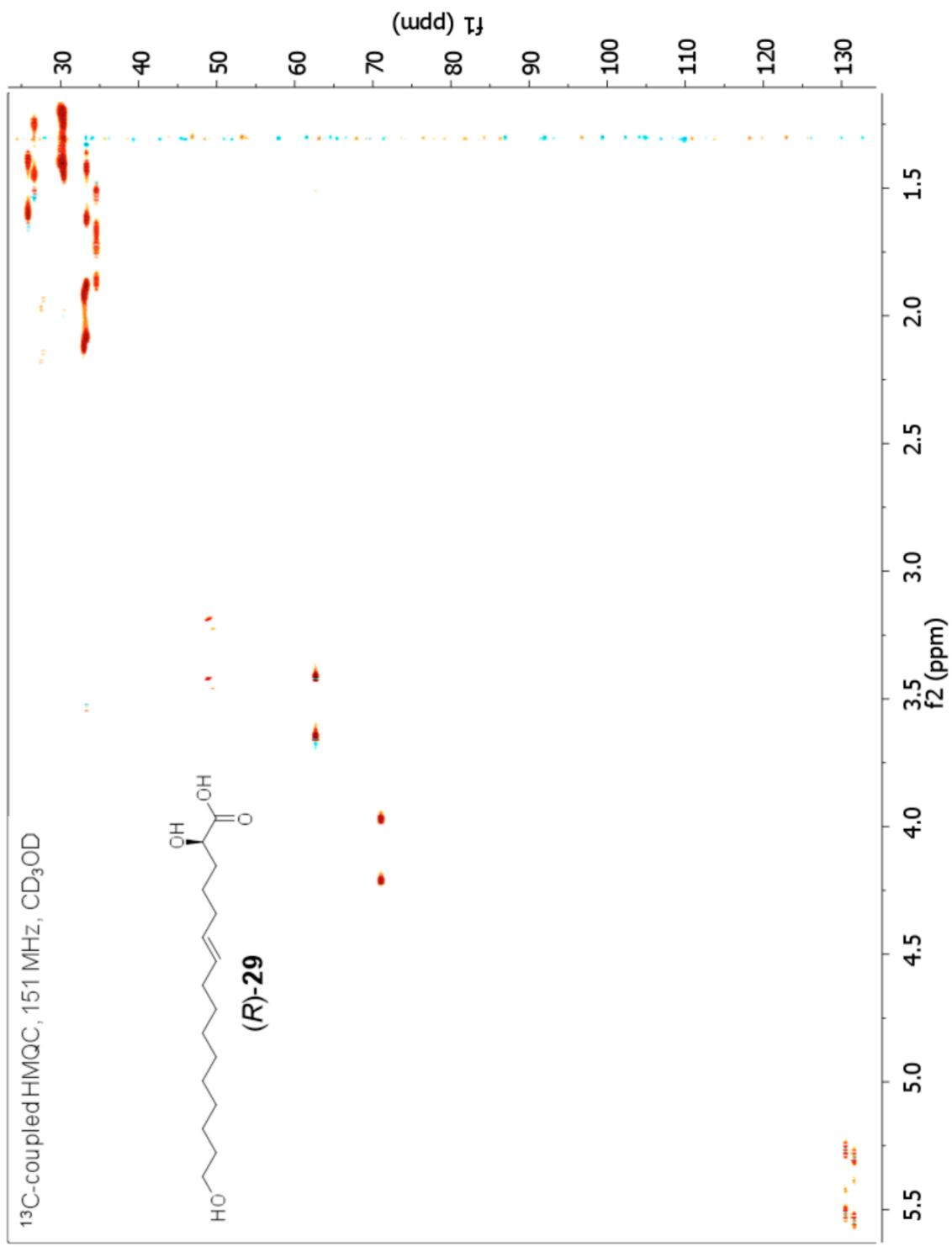


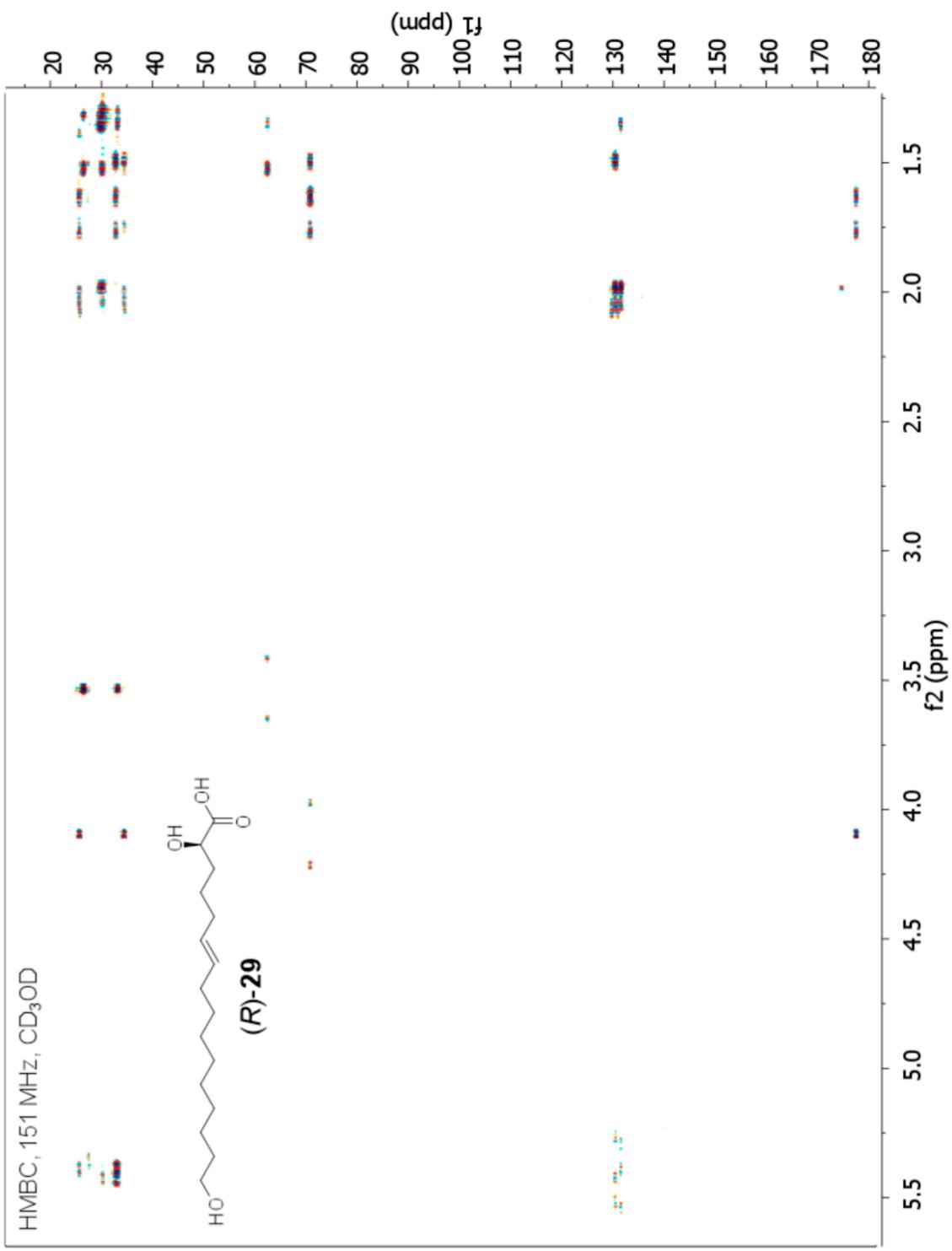


$^1\text{H}$  NMR, 600 MHz,  $\text{CD}_3\text{OD}$

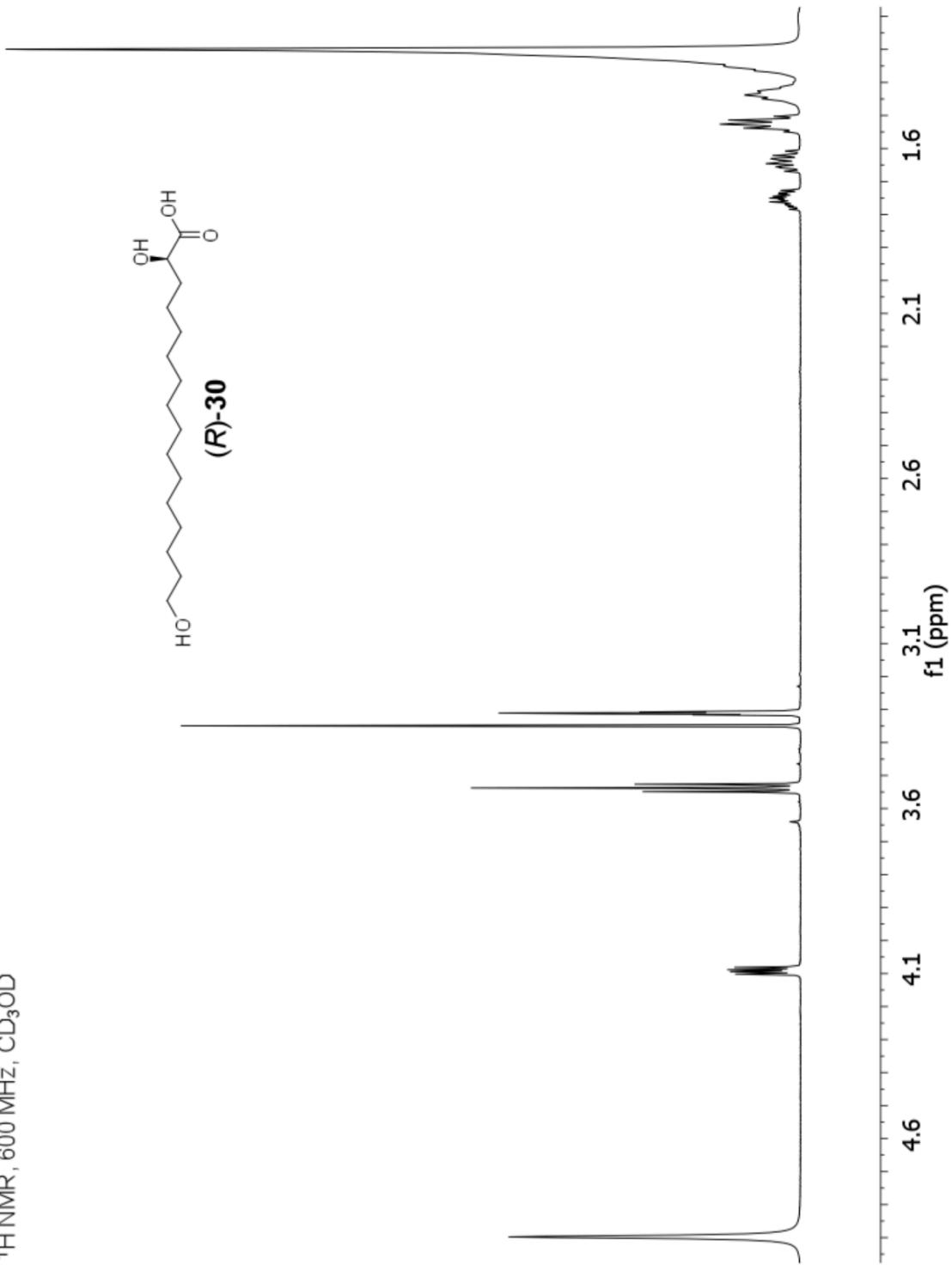


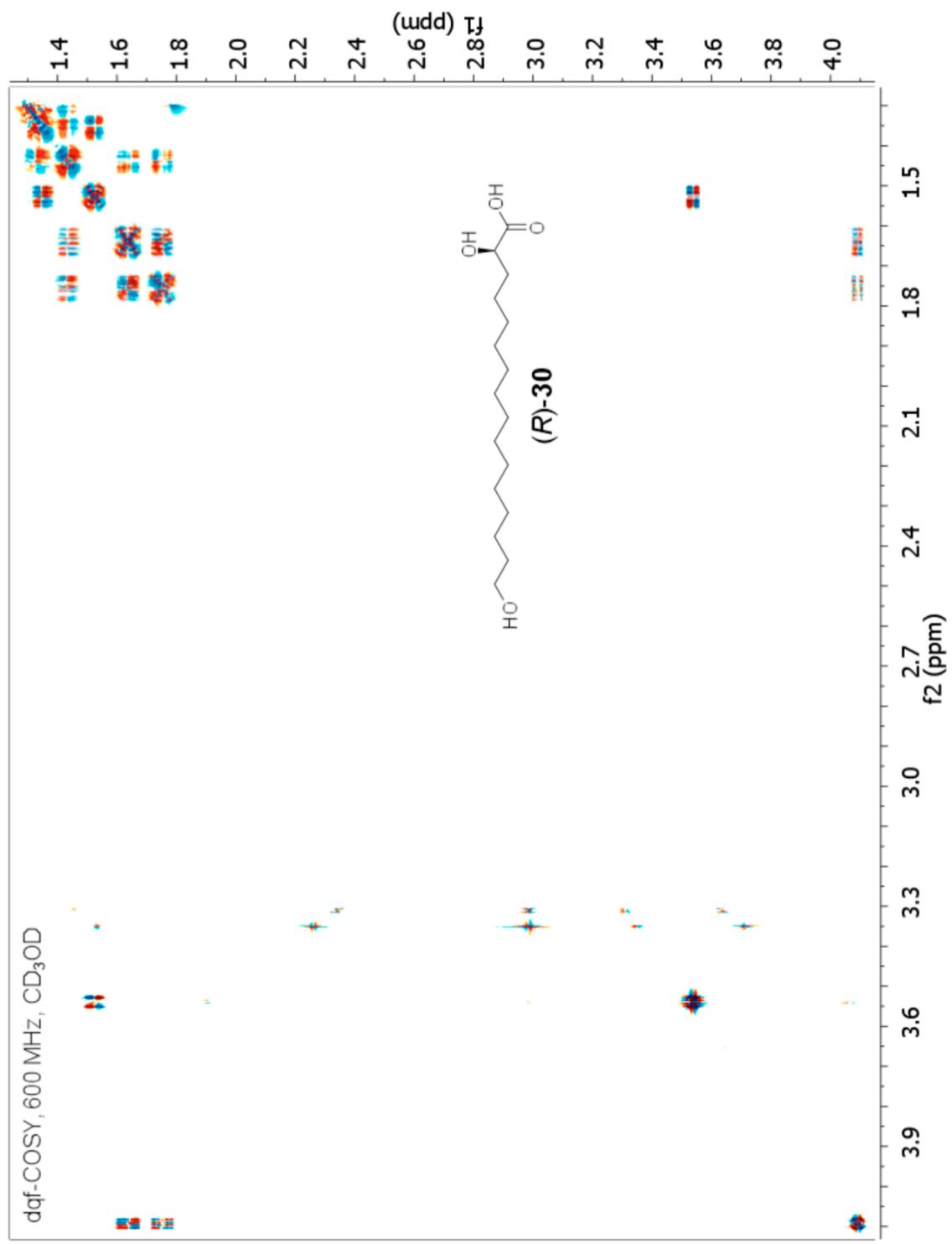


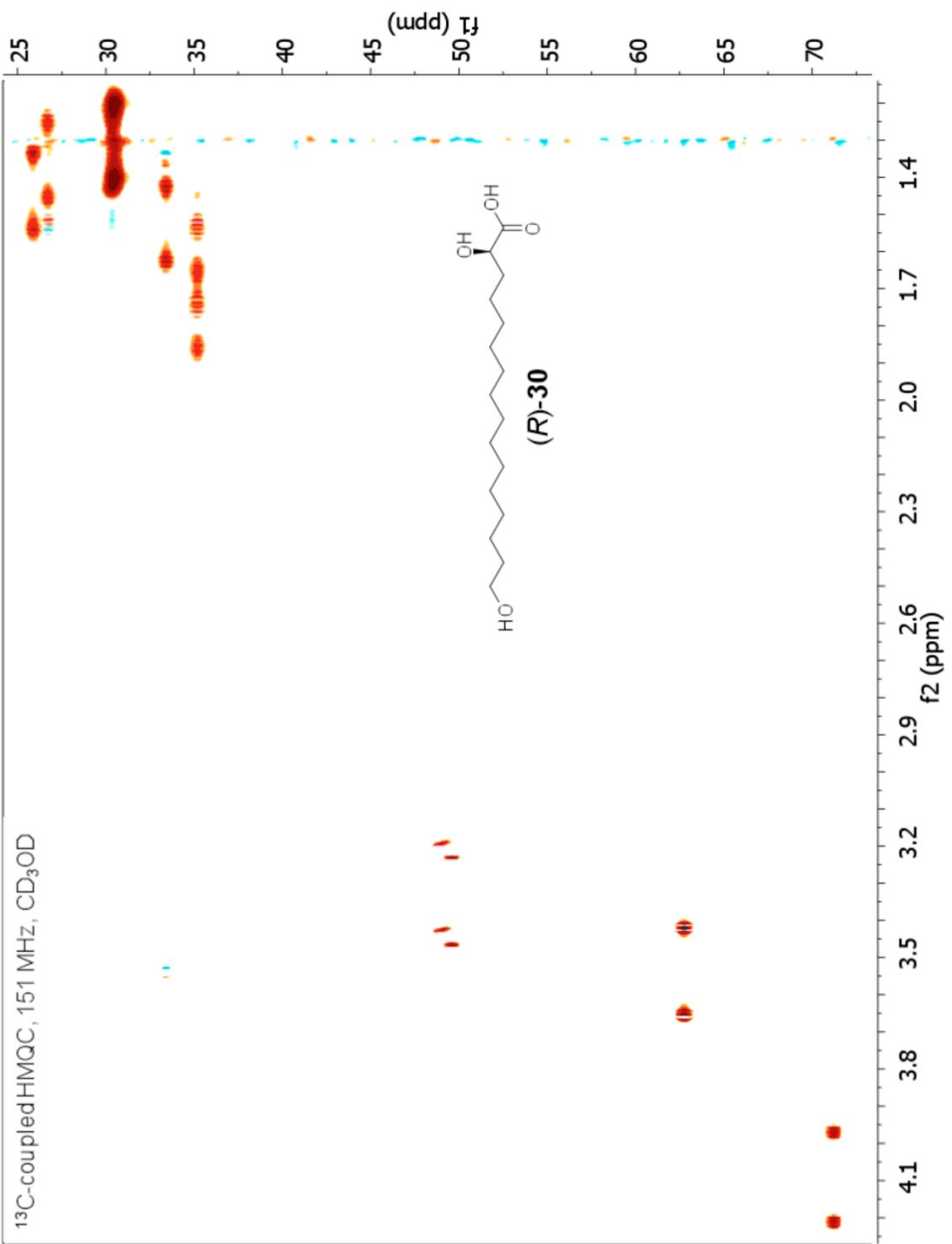


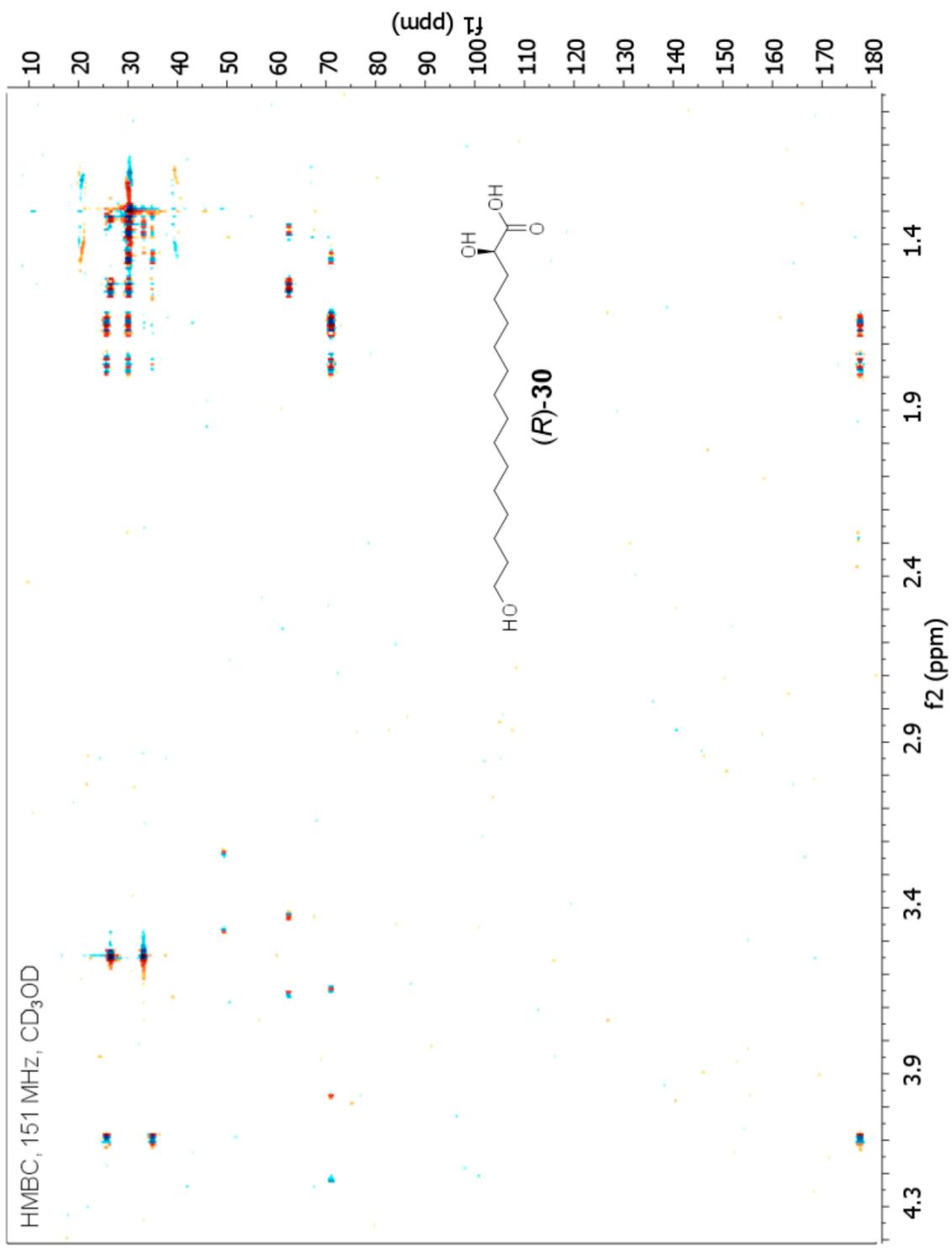


<sup>1</sup>H NMR, 600 MHz, CD<sub>3</sub>OD

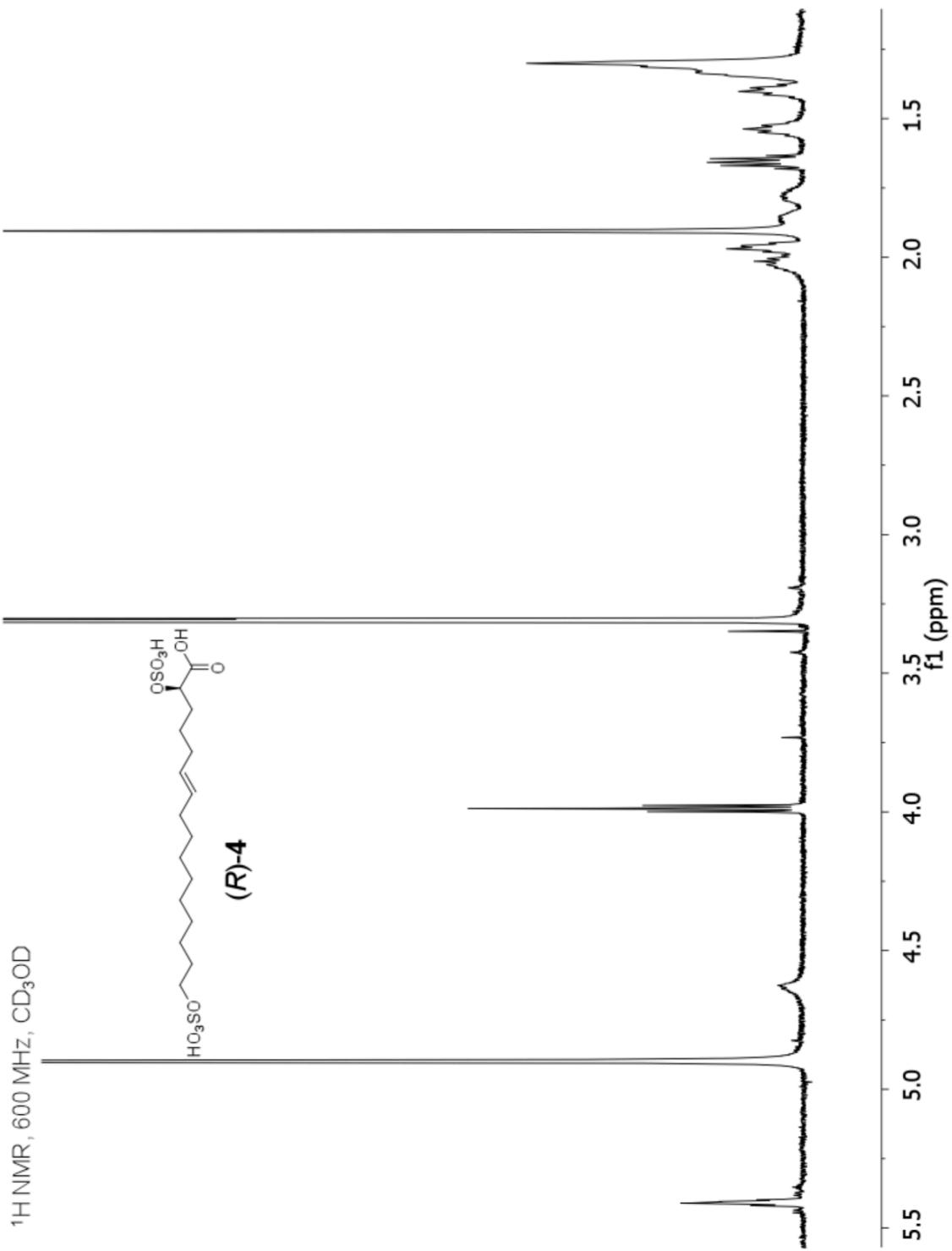


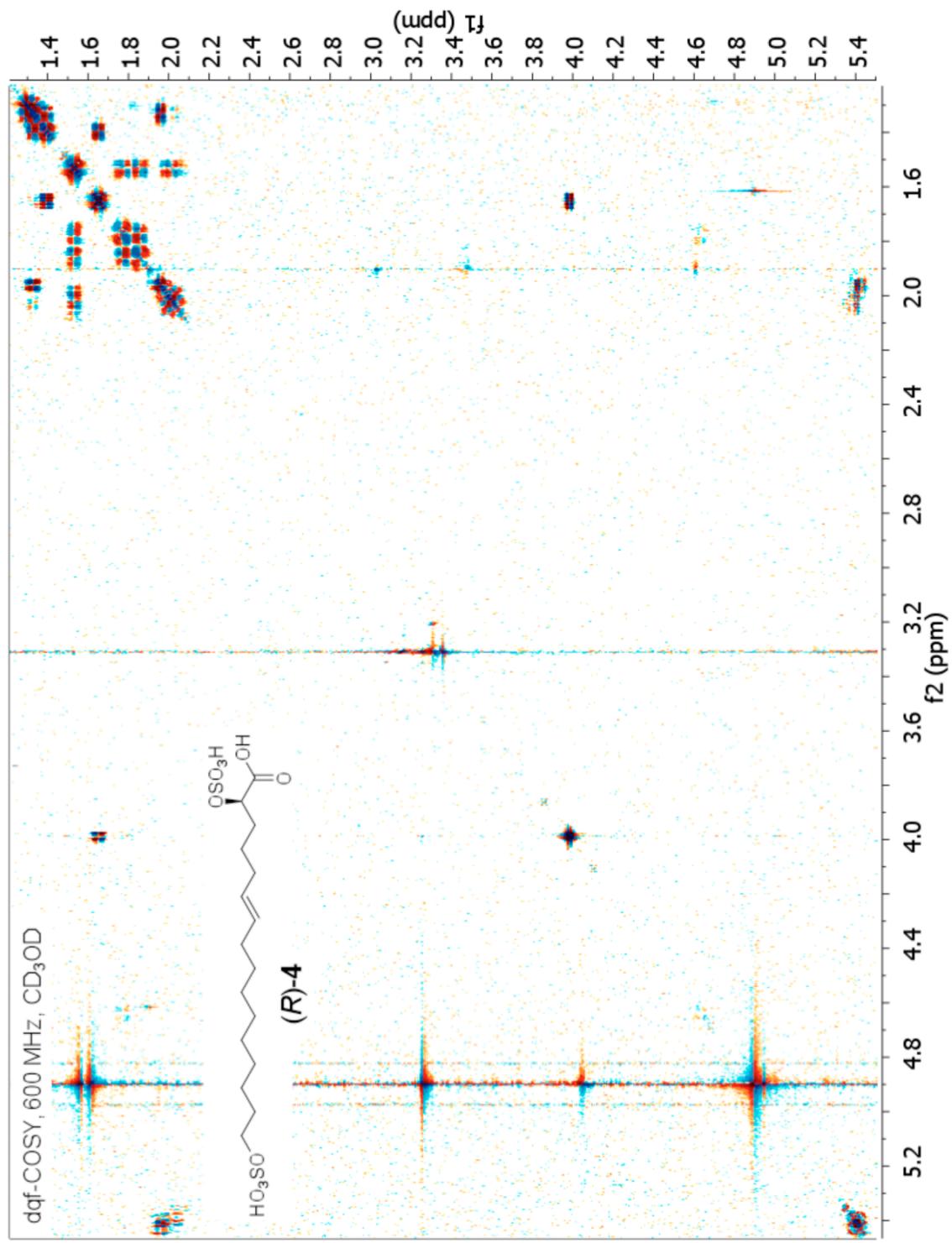


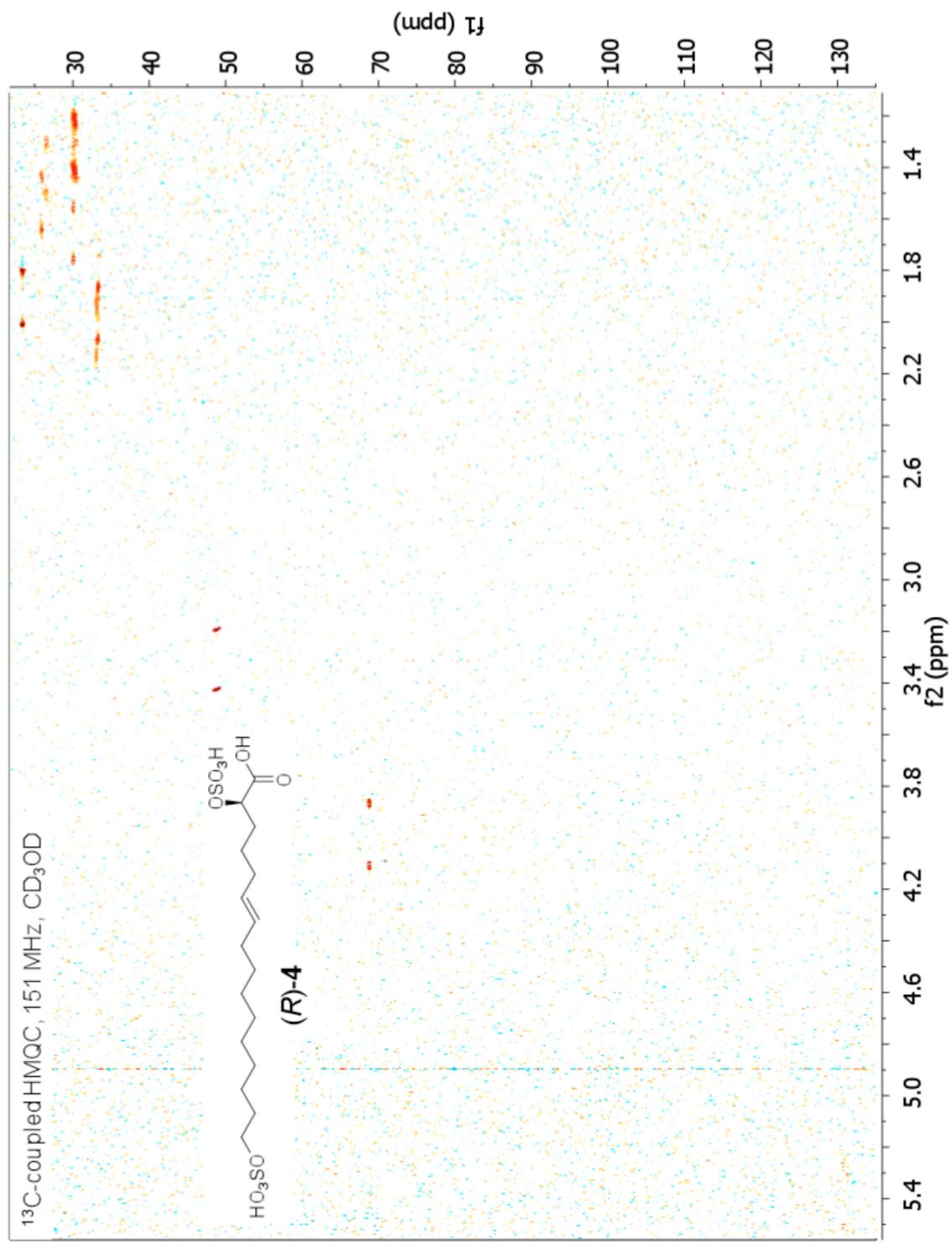


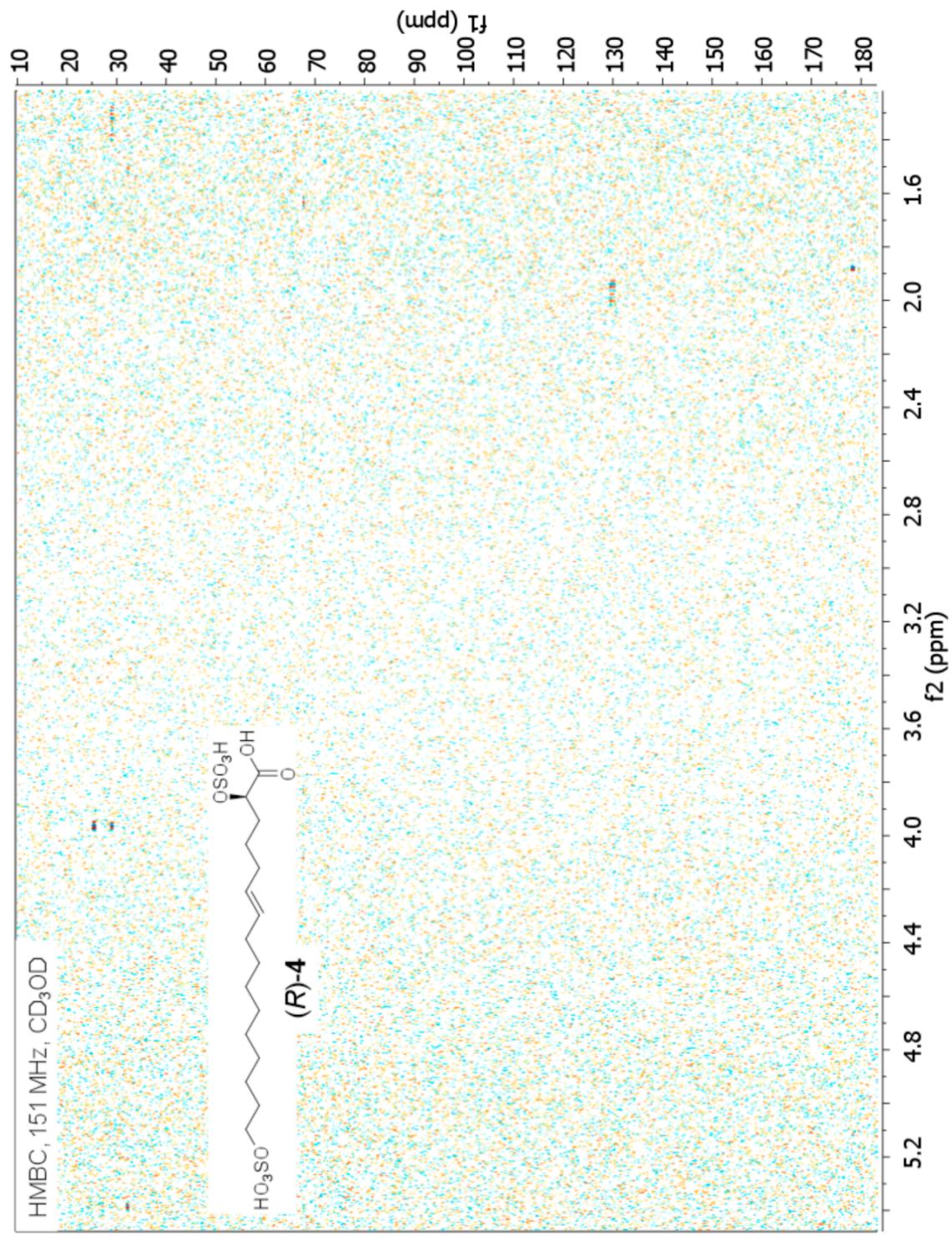


<sup>1</sup>H NMR, 600 MHz, CD<sub>3</sub>OD

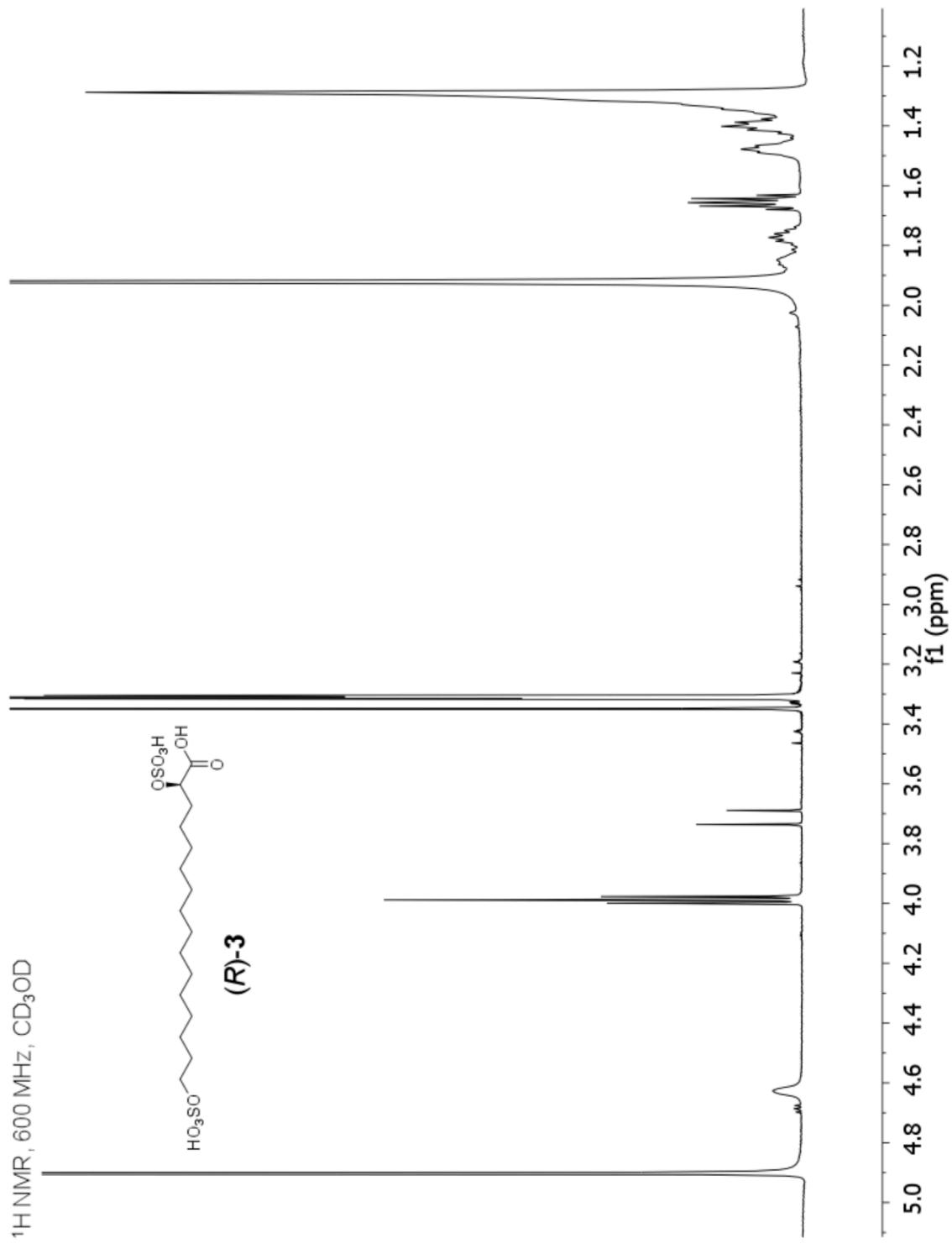


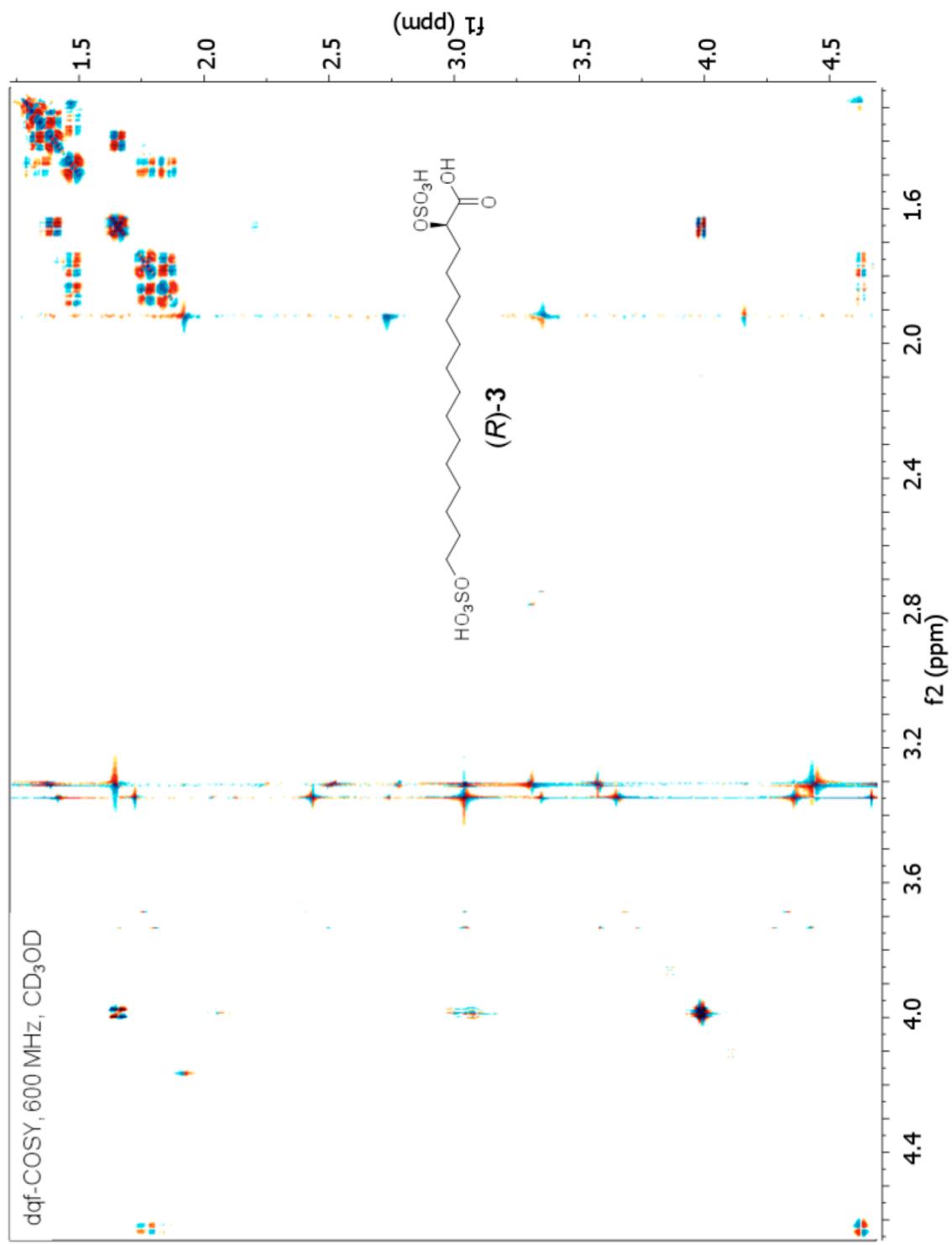


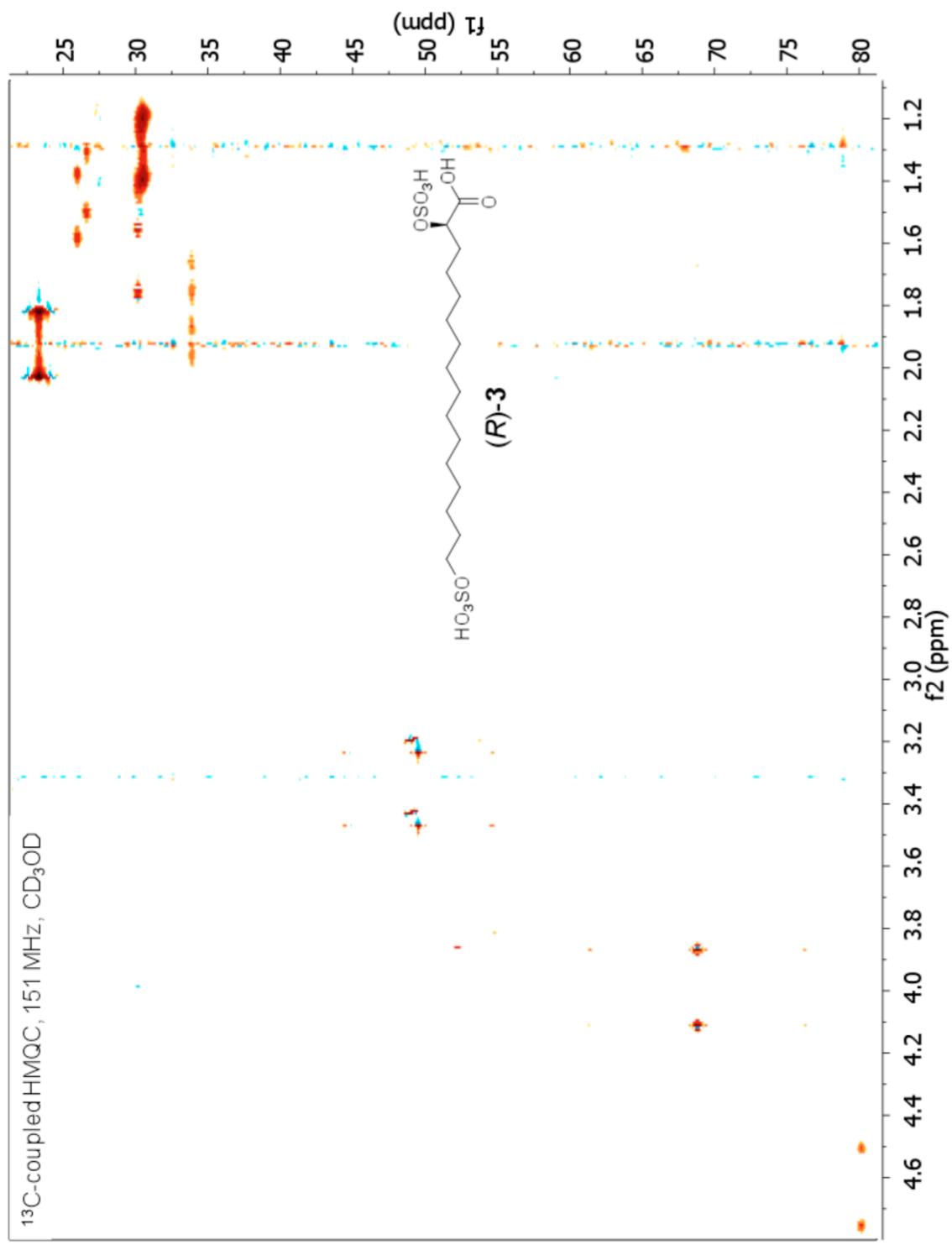


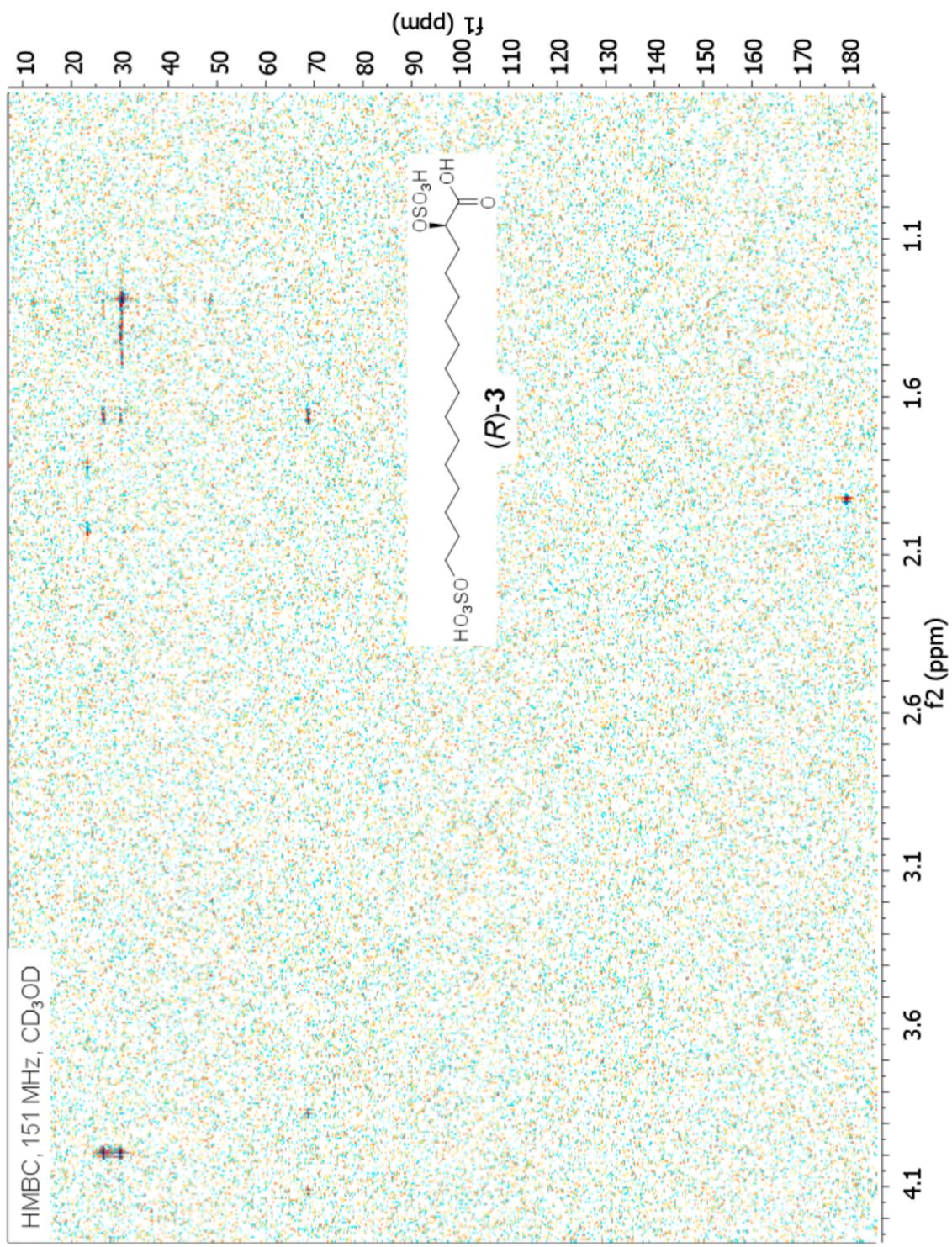


$^1\text{H}$  NMR, 600 MHz,  $\text{CD}_3\text{OD}$









## RERERENCE

1. Xu, P.; Lin, W.; Zou, X. *Synthesis* **2002**, 8, 1017-1026.
2. Williams, C. M.; Mander, L. N. *Tetrahedron* **2001**, 57, 425-447.
3. Grubbs, R. H.; Burk, P. L.; Carr, D. D. *J. Am. Chem. Soc.* **1975**, 97, 3265-3267.  
(b) Connon, S. J.; Blechert, S. *Angew. Chem., Int. Ed. Engl.* **2003**, 42, 1900-1923. (c) Hoveyda, A. H.; Zhugralin, A. R. *Nature* **2007**, 450, 243-251
4. Hong, S. H.; Sanders, D. P.; Lee, C. W.; Grubbs, R. H. *J. Am. Chem. Soc.* **2005**, 127, 17160-17161
5. Taggi, A. E.; Meinwald, J.; Schroeder, F. C. *J. Am. Chem. Soc.* **2004**, 126, 10364-10369.
6. Cahiez, G.; Chaboche, C.; Jezequel, M. *Tetrahedron* **2000**, 56, 2733-2737.
7. Schmelz, E. A.; Carroll, M. J.; LeClere, S.; Phipps, S. M.; Meredith, J.; Chourey, P. S.; Alborn, H. T.; Teal, P. E. *Proc. Natl. Acad. Sci. USA* **2006**, 103, 8894-8899.