

# (I) Novel Perfluorinated Aromatic Amino Acids: Synthesis and Applications (II) Thioflavin T Dimers as Novel Amyloid Ligands

Author: Luoheng Qin

Persistent link: <http://hdl.handle.net/2345/2620>

This work is posted on [eScholarship@BC](#),  
Boston College University Libraries.

---

Boston College Electronic Thesis or Dissertation, 2012

Copyright is held by the author, with all rights reserved, unless otherwise noted.

Boston College

The Graduate School of Arts and Sciences

Department of Chemistry

- (I) NOVEL PERFLUORINATED AROMATIC AMINO ACIDS:  
SYNTHESIS AND APPLICATIONS
- (II) THIOFLAVIN T DIMERS AS NOVEL AMYLOID LIGANDS

A dissertation

by

LUOHENG QIN

Submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

August 2012

© Copyright by LUOHENG QIN

2012

**(I) Novel Perfluorinated Aromatic Amino Acids: Synthesis and Applications**

**(II) Thioflavin T Dimers as Novel Amyloid Ligands**

**Luoheng Qin**

April 25, 2012

**Professor Jianmin Gao**

Research Advisor

**Abstract**

This thesis includes two projects: “Novel perfluorinated aromatic amino acids: synthesis and applications” and “Thioflavin T dimers as novel amyloid ligands”.

*I) Novel perfluorinated aromatic amino acids: synthesis and applications*

Fluorinated amino acids serve as powerful tools in protein chemistry. Using the commercially available Boc-protected pentafluorophenylalanine, we synthesized a series of para-substituted tetrafluorophenylalanines via the regioselective  $S_{\text{NAr}}$  reaction. These novel unnatural amino acids display useful and unique properties that can be applied to biological systems, including distinct  $^{19}\text{F}$  NMR signatures, pH-dependent amphiphilicity, lipid-binding selectivities, and halogen bonding capabilities.

*II) Thioflavin T dimers as novel amyloid ligands*

Fluorescent molecules that specifically target amyloid structures are highly desirable for Alzheimer’s disease research. We have designed a dimeric Thioflavin T

that, through a reduced entropic penalty, has an improved binding affinity to A $\beta$  amyloid by up to 70 fold. More importantly, the specificity and the “light-up” feature upon amyloid binding have not been sacrificed. Encouraged by the successful dimer design, we are further investigating the potential of amyloid-templated reactions (photoactive, click chemistry) to tailor-make ligands for amyloids.

## **Acknowledgements**

I would like to thank the members of my thesis defense committee, Professors James Morken, Marc Snapper and Eranthie Weerapana for taking time to critically read my thesis and to attend my defense.

I am grateful to Professor Jianmin Gao for accepting me into his lab and for his guidance over the past five years. From him I learned not only a broad scope of knowledge but also how to logically trouble-shoot problems with research. He is also a good friend.

I would like to thank my family and all my friends for the never-ending support to me.

I would like to thank Christopher Pace for carefully reviewing my thesis. Chris also contributed to the halogen bonding project presented in this thesis. I would also like to thank Fang Wang for her collaboration in the pH-dependent membrane-lysis and other projects. It has been a pleasure to work with both of you.

Finally, I would like to thank all the current and past group members of the Gao lab. We have helped each other in the lab, and have happily spent five years together. I also thank all the other people in the Boston College Chemistry Center. The five years I spent here will be unforgettable.

## Table of Contents

### Chapter 1. Novel perfluorinated aromatic amino acids: synthesis and applications

1.1	Introduction	2
	<i>I) Fluorination in medicinal chemistry</i>	2
	<i>II) Fluorination in peptide design and protein engineering</i>	6
	<i>III) Fluorinated aromatic amino acid derivatives</i>	8
	<i>IV) Synthesis of multi-fluorinated aromatic amino acids</i>	12
1.2	Synthesis of <i>para</i> -X-tetrafluorophenylalanine (ZpX) series	13
	<i>I) The regioselective <math>S_{NAr}</math> reaction on pentafluorophenyl moieties</i>	13
	<i>II) The “O” nucleophile and the subsequent cross-coupling reactions</i>	15
	<i>III) The “S” nucleophile</i>	24
	<i>IV) The “N” nucleophile and the Sandmeyer reactions</i>	26
1.3	Applications of ZpX series in biological systems	33
	<i>I) 2D <math>^{19}F</math> NMR cosy analysis of peptides</i>	33
	<i>II) Halogen bonding in peptides</i>	36
	<i>III) pH dependent membrane lytic peptides</i>	43
	<i>IV) Boronic acid based amino acids</i>	46
1.4	Conclusions	49
	<b>Materials and methods</b>	50

<b>References</b>	80
<b>Chapter 2. Thioflavin T dimers as novel amyloid ligands</b>	
2.1 Introduction	91
2.2 Synthesis of ThT dimer series	95
2.3 A $\beta$ 40 binding affinity measurements	98
2.4 A $\beta$ 40 binding specificity measurements	106
2.5 The investigation of the binding mechanism and the “head-to-tail” dimer	108
2.6 Conclusions and future work	112
<b>Materials and Methods</b>	118
<b>References</b>	147
<b>Appendix: <math>^1\text{H}</math> NMR and <math>^{13}\text{C}</math> NMR Spectra</b>	151
$^1\text{H}$ NMR	152
$^{13}\text{C}$ NMR	196



## Table of Abbreviations

A $\beta$	Amyloid $\beta$
AD	Alzheimer's disease
Asp	Aspartic acid
Boc	<i>tert</i> -Butoxycarbonyl
BTA	Benzylthiazole aniline
CD	Circular dichroism
cosy	Correlation spectroscopy
DCM	Dichloromethane
DIBAL	Diisobutylaluminium hydride
DIPEA	Diisopropylethylamine
DMAc	Dimethylacetamide
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
dppf	1,1'-Bis(diphenylphosphino)ferrocene
F	Phenylalanine
Flp	Fluoroproline
Fmoc	Fluorenylmethyloxycarbonyl
Glu	Glutamic acid
Gly	Glycine

HB	Hydrogen bonding
His	Histidine
HPLC	High-performance liquid chromatography
HSAB	Hard and soft acid and base (theory)
Hyp	Hydroxyproline
Leu	Leucine
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PEG	Poly-ethylene glycol
PET	Positron emission tomography
PG	Phosphatidylglycerol
Ph	Phenyl
Phe	Phenylalanine
pin	Pinacol
Pro	Proline
PS	Phosphatidylserine
SM	Sphingomyelin
SPPS	Solid-phase peptide synthesis
TBAF	Tetrabutylammonium fluoride

TBS	<i>tert</i> -Butyldimethylsilyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
ThT	Thioflavin T
TMS	Trimethylsilyl
Trp	Tryptophan
Tyr	Tyrosine
UV	Ultra violet
W	Tryptophan
XB	Halogen bonding
Y	Tyrosine
Z	Pentafluorophenylalanine
ZpX	<i>para</i> -X-Tetrafluorophenylalanine (X = functional groups)

# **Chapter 1:**

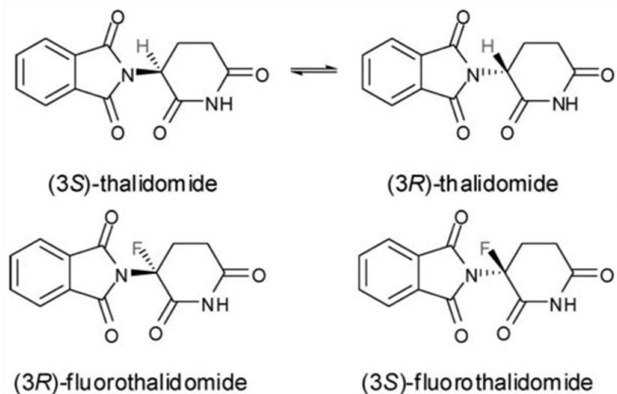
## **Novel Perfluorinated Aromatic Amino Acids: Synthesis and Applications**

## 1.1 Introduction

Fluorine has emerged as a “magic element” in medicinal chemistry and protein engineering<sup>[1-4]</sup>. Fluorination is now considered a popular and “standard” strategy for modulating the properties of small molecules and peptides. This strategy has been so successful that today 20-25% of all pharmaceuticals contain at least one fluorine atom,<sup>[3]</sup> despite the fact that organofluorine compounds are virtually absent in nature. The applications of fluorination are based on the unique stereoelectronic properties of fluorine. The “fluorine effects” in medicinal chemistry and protein engineering are listed and briefly introduced below.

### 1) Fluorination in medicinal chemistry

1) *Metabolic stability enhancement.* Fluorine serves as a strong  $\sigma$ -electron withdrawing group and can cause dramatic changes in metabolism. In fact, fluorination is primarily used to reduce the oxidation rate of the aromatic rings in drug molecules. In

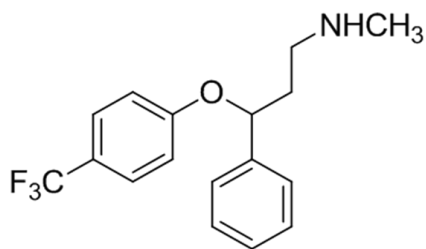


**Figure 1-1.** Thalidomides and the fluorinated derivatives

addition, fluorination can also suppress racemization mechanisms. For example, (*R*)-thalidomide, which was used to treat morning sickness, epimerizes at physiological pH to the (*S*)-isomer, leading to disastrous teratogenic

side effects. A replacement of the acidic proton with fluorine (Figure 1-1) enables the evaluation of each individual isomer<sup>[5]</sup>, and studies have shown the potential utility of these isomers in the treatment of leprosy and tumors<sup>[6-7]</sup>.

2) *Minimal steric hindrance*. Fluorine is the smallest in size among all elements except hydrogen. A C-F bond (1.30-1.43 Å) is only slightly longer than a C-H bond (1.09 Å)<sup>[8]</sup> so that fluorination introduces minimal steric perturbation to the parent structure. In the



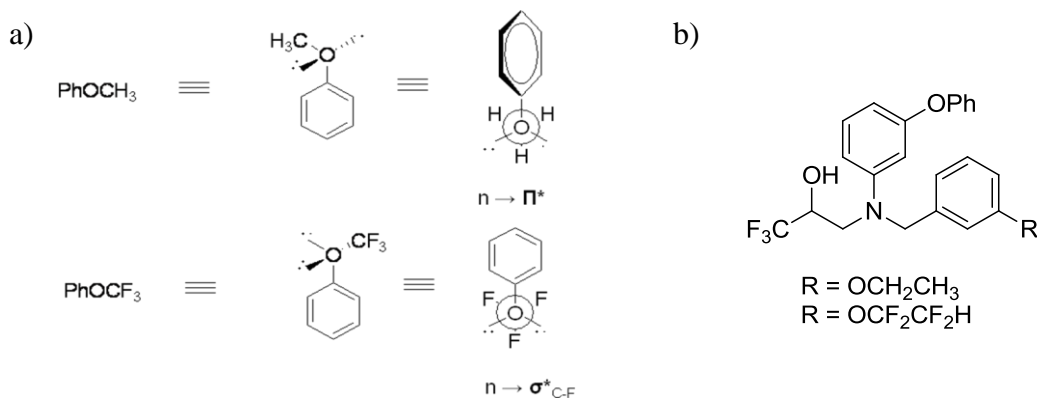
**Figure 1-2.** Fluoxetine

antidepressant Fluoxetine (sold as Prozac by Eli Lilly, see Figure 1-2), the trifluoromethyl group on the para-position of the phenyl ring increased the potency for inhibiting 5-hydroxytryptamine uptake by 6-fold

compared to the non-fluorinated compound<sup>[9]</sup>, illustrating that fluorination does not significantly disturb the steric properties of a molecule.<sup>[10]</sup>

3) *Altered hydrophobicity*. Due to its high electronegativity and small size, fluorine has the lowest polarizability among all common elements found in organic compounds<sup>[11]</sup>. In addition, a C-F bond can even cause a reduced overall molecular polarizability through the carbon framework of the whole molecule. Therefore, fluorination often increases the hydrophobicity of the compound<sup>[12]</sup>. It's important to note that fluorination does not always increase the hydrophobicity. Terminal mono-, di-, or trifluorination on an aliphatic chain typically decreases the hydrophobicity of the molecule<sup>[13]</sup>, possibly because in these cases the increased dipole moment upon fluorination overwhelms the decreased polarizability.

4) *Effects on conformational preorganization.* Fluorination causes conformational changes primarily via stereoelectronic effects. For example, methoxybenzene predominately adopts a planar conformation due to the  $n \rightarrow \Pi^*$  donation, while trifluoromethoxy-benzene has the  $-\text{OCF}_3$  group sticking out of the plane (dihedral angle for C-C-O-C close to  $90^\circ$ , see Figure 1-3a). This conformational difference was



**Figure 1-3.** a) The conformational preferences for methoxybenzene and trifluoromethoxybenzene; b) Cholesteryl ester transfer protein inhibitor.

successfully exploited to design superior inhibitors of the cholesteryl ester transfer protein<sup>[14]</sup> (Figure 1-3b), which transfers cholesteryl ester from high-density lipoprotein to low-density lipoprotein and has been implicated in coronary heart disease. The inhibition potency can be increased by 8-fold by modifying the R group from ethyl to tetrafluoroethyl (Figure 1-3b), which is rationalized by molecular modeling experiments as the out-of-plane orientation of the R group fits better into the target protein.

5) *pK<sub>a</sub> perturbation.* Due to the high electronegativity, a fluorine substitution can dramatically lower the pK<sub>a</sub> values of adjacent functional groups. For example, incorporation of a single fluorine atom onto the imidazole ring lowers the pK<sub>a</sub> value of a histidine residue by several pK<sub>a</sub> units (pK<sub>a, His</sub> = 6.0, pK<sub>a, (2F)-His</sub> = 1.22 and pK<sub>a, (4F)-His</sub> =

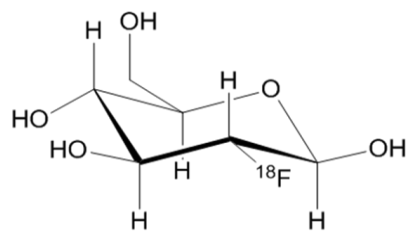
1.76)<sup>[15]</sup>. Jackson et al. incorporated (4F)-His into ribonuclease A and proved that His12 and His119 act as general acid/base catalysts, but the proton transfer is not the rate limiting step<sup>[16]</sup>.

6) <sup>19</sup>F NMR. Organic fluorine compounds do not exist in biology, so that fluorinated bio-molecules can be tracked in complex biological systems with zero background. The stable and naturally abundant isotope of fluorine, <sup>19</sup>F, is almost as sensitive (86%) as a proton in NMR spectroscopy. <sup>19</sup>F NMR has therefore evolved to be a sensitive and powerful tool in protein analysis, even in living cells<sup>[17]</sup>. One additional advantage of <sup>19</sup>F NMR is the relatively broad range of chemical shift values, which enables the detection of minor differences in the environment, for instance, protein folding into membranes<sup>[18]</sup>. Grage et al. incorporated 5-F-Trp into the membrane-associated ion-channel peptide gramicidin A to investigate the conformation and dynamics of tryptophan residues in lipid environments. With <sup>19</sup>F NMR, their results suggested that the dipole moment plays an important role in the ion-channel activity<sup>[19]</sup>.

7) <sup>18</sup>F PET imaging. The most important application of fluorination as diagnostic tools is the <sup>18</sup>F positron emission tomography (PET). PET is a nuclear medicine imaging technique that produces a three-dimensional image or movie of functional processes in the body. The system detects pairs of gamma rays emitted indirectly by a positron-emitting radionuclide, which is introduced into the body on a biologically active molecule<sup>[20]</sup>. Since the gamma ray pairs are generated strictly simultaneously and in opposite directions, PET has extraordinarily high sensitivity and resolution, and hence has become one of the best diagnostic tools for tumors. Among the



common positron-emitting isotopes in PET,  $^{18}\text{F}$  has the longest half-life time (half-life for  $^{11}\text{C}$ , 20 min;  $^{13}\text{N}$ , 10 min;  $^{15}\text{O}$ , 2 min; and  $^{18}\text{F}$ , 110 min). Therefore, fludeoxyglucose<sup>[21]</sup> (Figure 1-4) has become the most commonly used radiotracer in PET. More recently,



**Figure 1-4.** Fludeoxyglucose

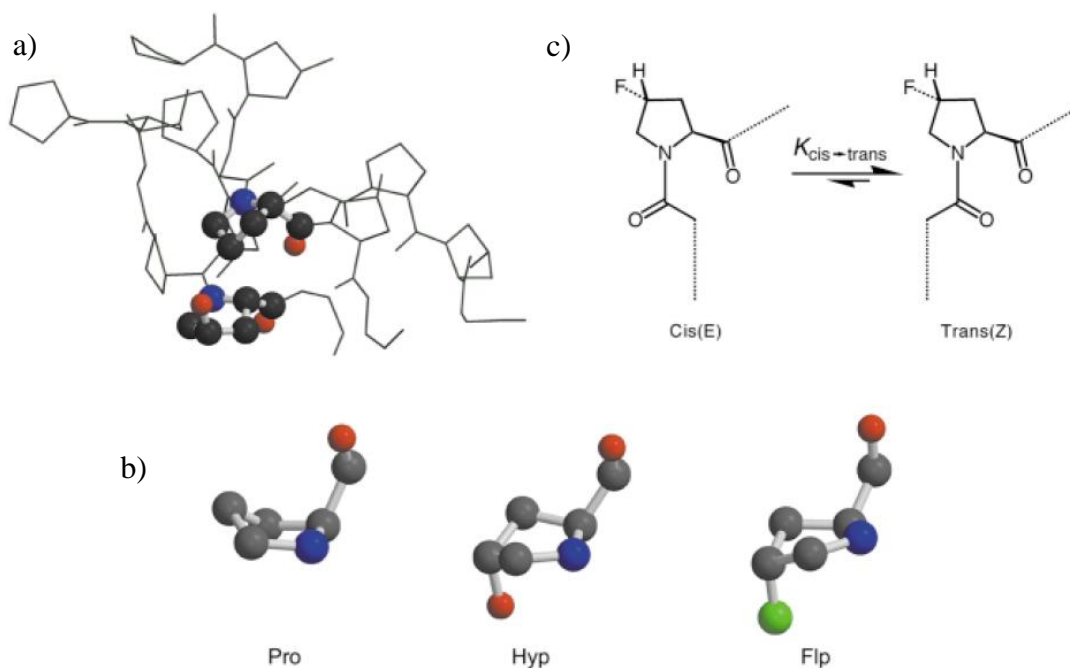
Ritter et al. developed methodologies to conveniently and quickly fluorinate aromatic rings in pre-drug small molecules with palladium reagents.<sup>[22-23]</sup> These new synthetic methodologies greatly expand the scope of drug molecules for  $^{18}\text{F}$  PET diagnosis.

## *II) Fluorination in peptide design and protein engineering*

Beyond small molecule medicinal chemistry, fluorination has also attracted a great deal of attention in peptide design and protein engineering. In peptides and proteins, multiple incorporations of fluorinated residues are feasible. Therefore, it is possible to obtain the emergent fluorine effects, which do not exist in small molecules.

1) *Conformational pre-organization by gauche effects.* A good example of fluorine-induced gauche effects is the investigation of the origins of collagen stability<sup>[24]</sup>. The triple helix of collagen is well known to be formed by the repeating trimers containing glycine (Gly), proline (Pro) and 4(*R*)-hydroxyproline (Hyp), which was long believed to be due to the additional hydrogen bonding provided by the Hyp residues. Raines et al. synthesized collagen-like peptides [Pro-Xaa-Gly]<sub>10</sub> with either Hyp or 4(*R*)-fluoroproline (Flp) as Xaa respectively and compared the stability of these two peptides. Surprisingly, the Flp peptide is much more stable, with a melting temperature of 91 °C (the melting

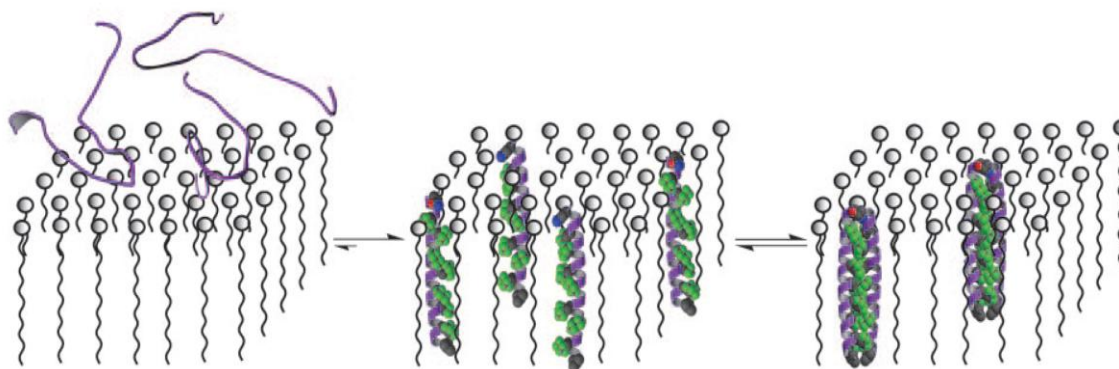
temperature for the Hyp peptide is 69 °C). NMR and computational studies suggested that the gauche effect stabilizes both the *exo*- conformation of the proline ring and the *trans*-conformation of the backbone amide bond (Figure 1-5), and the resulting predisposition of the substituted proline leads to the stabilization of the collagen helix<sup>[25-26]</sup>.



**Figure 1-5.** The gauche effect of fluorine substitution on proline rings and collagen mimicking peptides. a) Section of the structure of a collagen-mimetic peptide with one Pro and one Hyp residue, b) The favored conformation of Pro, Hyp and Flp residues, showing the substituted prolines adopt the *exo*- conformation. c) *cis*-*trans* Isomerization of the backbone amide bond in Flp residues, in which *trans*- is much more favorable. (This figure has been reproduced from Yoder, N. C.; Kumar, K., *Chem. Soc. Rev.* **2002**, 335-341)

2) “*Fluorous effect*”. Due to the low polarizability of fluorine atoms, the perfluorinated carbons are reluctant to engage in dispersive interactions with either water or hydrocarbons and tend to segregate from both aqueous and hydrocarbon-based environments. This so-called “fluorous effect” has been utilized to create Teflon-like

proteins. For example, Kumar et al. reported that multi-incorporation of hexafluoroleucine into a transmembrane peptide resulted in self-assembly in membranes presumably triggered by the “fluorous effect” (Figure 1-6)<sup>[27]</sup>.



**Figure 1-6.** Transmembrane helices incorporating multiple hexafluoroleucines self-assemble in membranes. Green: fluorine atoms; red: oxygen; blue: nitrogen; purple: peptide backbones. (This figure has been reproduced from Bilgicer, B.; Kumar, K.; *Proc. Natl. Acad. Sci. USA*, **2004**, 15324-15329)

Generally speaking, the effects of fluorination are based on both covalent and non-covalent properties of fluorine atoms and contribute widely to protein engineering and drug design. In the next section of the introduction, we will focus on the multi-fluorinated aromatic amino acids, phenylalanine and tyrosine derivatives, which are also what my research has focused on.

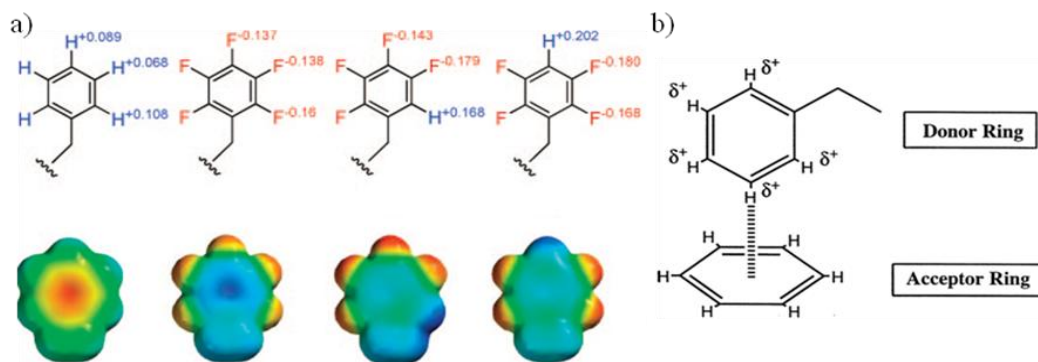
### *III) Fluorinated aromatic amino acids*

Although fluorinated peptides and proteins have attracted much attention, early work largely concentrated on the interactions of fluorinated aliphatic amino acids to produce “Teflon-like” proteins as introduced previously. However, the “Teflon-like” property

does not apply to fluorinated aromatic amino acids. Instead, the aromatic residues bear their own unique properties, interactions and related applications. These novel mechanisms include edge-face  $\pi$ - $\pi$  interactions, face-face  $\pi$ - $\pi$  interactions and CH- $\pi$  interactions<sup>[28]</sup>. Here, we briefly introduce the multi-fluorinated phenylalanine and tyrosine derivatives and their applications.

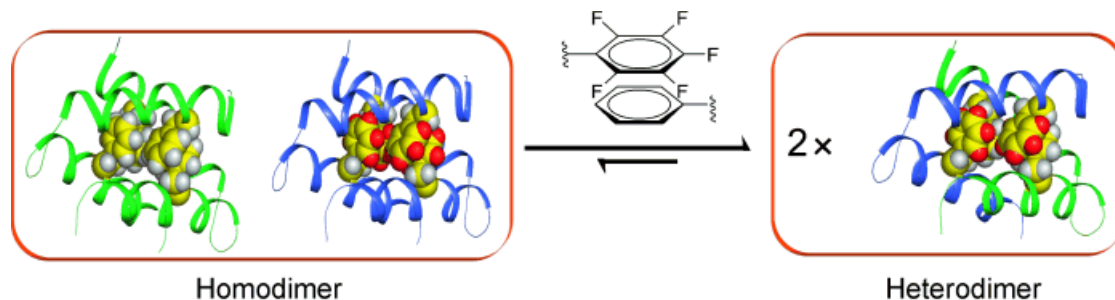
1) *Fluorinated phenylalanine (Phe) derivatives.*

The uneven electron distribution of aromatic rings results in the quadrupole moment even though the whole molecule does not have a net dipole. In benzene, the  $sp^2$  carbon is more electronegative than hydrogen (electronegativity in Pauling scale: 2.75 v.s. 2.48)<sup>[29]</sup>, so the center of the aromatic ring is partial negatively charged while the periphery hydrogen atoms are partial positively charged (Figure 1-7a). Therefore, electrostatic interactions, such as cation- $\pi$  interactions<sup>[30]</sup>, can be engaged by aromatic  $\pi$  faces. On the other hand, the periphery protons can interact with electron rich species to give, for example, the edge-face packing of benzene (Figure 1-7b).



**Figure 1-7.** a) Structure and space filling models (*Spartan*) showing partial charges and electrostatic potentials (red for negative and blue for positive) of non-fluorinated and some multi-fluorinated phenyl rings. b) The edge-face interaction between two phenyl rings.

However, in multi-fluorinated aromatic rings, the quadrupole moment is reversed due to the high electronegativity of fluorine. Therefore, the  $\pi$  face is electron-poor and tends to interact with electron-rich species (Figure 1-7a). Also, if there are any remaining hydrogen atoms on the multifluorinated arenes, they will be even more electron deficient than those on the non-fluorinated ones and more favorable for the edge-face interactions.



**Figure 1-8.** The face-face interaction between phenyl and pentafluorophenyl rings triggering the heterodimer formation of the wild type and fluorinated  $\alpha_2$ D helices.

Zheng et. al. strengthened the edge-face interactions and stabilized the villin headpiece subdomain (HP35) by mutating the Phe10 residue in the aromatic core with 2,3,4,5-tetrafluorophenylalanine<sup>[31]</sup>. They also demonstrated the significance of face-face electrostatic interactions between phenylalanine and pentafluorophenylalanine  $\pi$  clouds with the model peptide  $\alpha_2$ D, a *de novo* designed polypeptide that folds into a dimeric helix bundle<sup>[32]</sup>. In the folded  $\alpha_2$ D dimer, the F10 side chains of each monomer stack with those of F29 from the other in the face-to-face fashion. A double mutant where both phenylalanines are substituted by pentafluorophenylalanine also forms the similar dimer. When the two dimers are mixed, they transformed into heterodimers presumably driven by the enhanced face-face interaction (Figure 1-8). Also with the  $\alpha_2$ D system, Pace et. al. utilized pentafluorophenylalanine as the control molecule to demonstrate the energetic

significance of the CH- $\pi$  interactions between the aliphatic C-H from cyclohexylalanine and phenylalanine<sup>[33]</sup>.

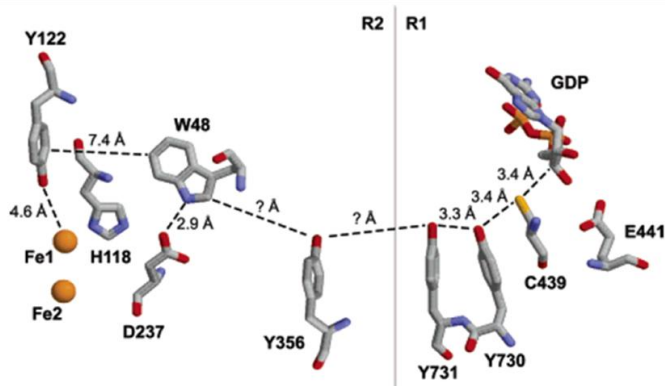
## 2) Fluorinated tyrosine (Tyr) derivatives.

Similar to the phenylalanine derivatives, fluorination can decrease the electron density on the aromatic ring of the tyrosine side chain, and cause the lone pair of the phenol hydroxyl group to be more delocalized to the phenyl ring. This effect directly results in the unique properties of multi-fluorinated tyrosines. The first one is the  $pK_a$  shift on the phenol group. The  $pK_a$  values for the two trifluorinated tyrosine isomers are 6.1 and 6.6<sup>[34-35]</sup>, and for tetrafluorotyrosine it is 5.6. Compared with the native tyrosine residue ( $pK_a = 10$ ), the multi-fluorinated ones are  $10^3$ - $10^4$  folds more acidic and are thus deprotonated even at physiological pH (7.2 - 7.4).

Thus, the fluorinated tyrosine analogues have been utilized as probes to elucidate the mechanisms of a number of tyrosine-tailoring enzymes, such as tyrosinases and kinases<sup>[35-36]</sup>.

Another important consequence of the tyrosine

fluorination is an increased redox potential, which ranges from 705 to 968 mV depending



**Figure 1-9.** Tyr356 serving as the redox active residue in *E. coli* ribonucleotide reductase. It is proposed to propagate the radical between Trp48 and Tyr731. (This figure has been reproduced from Seyedsayamdost, M. R.; Yee, C. S.; Stubbe, J., *Nat. Protoc.*, **2007**, 1225-1235)

on the extent and position of fluorination, while the native tyrosine has a peak potential of 642 mV<sup>[37]</sup>. Stubbe et al. incorporated a variety of di-, tri-, and tetrafluorotyrosines into *E. coli* ribonucleotide reductase. Analysis of the mutants showed that the residue Tyr356 acts as a redox active residue in this enzyme (Figure 1-9)<sup>[38]</sup>.

The effects of fluorination on the hydrophobicity of tyrosine are complicated. It is suggested that fluorinated tyrosine analogues are more hydrophilic<sup>[1]</sup>, because the phenol group is more polarized, more acidic and serves as a better hydrogen bonding donor. However, the neutral form of the tetrafluorotyrosine residue is more hydrophobic than native tyrosine based on our *Spartan* calculation results (the logP value for tetrafluorotyrosine is 2.76, for tyrosine is 2.13). It may still be true that some of the mono- or difluorotyrosines are more hydrophilic, if the fluorine-induced hydrophobicity enhancement is not enough to counteract the polarization effect on the phenol –OH group.

#### *IV) Synthesis of multi-fluorinated aromatic amino acids*

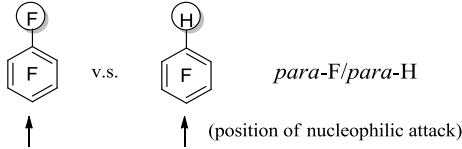
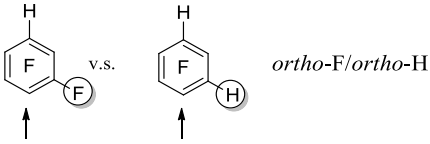
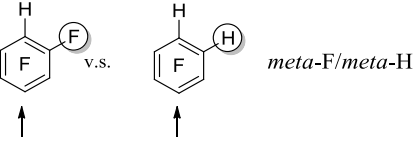
Although fluorination is proven to be a powerful tool in protein analysis and engineering, there are only a limited number of fluorinated amino acids available from commercial sources, as their synthesis is nontrivial. In the next section we will report a facile synthesis of *para*-X-tetrafluorophenylalanine derivatives via the regioselective S<sub>NAr</sub> chemistry of pentafluorophenylalanine, where X represents a range of functionalities, including hydroxyl, alkyl, aryl thiol, amino, and halogens. For convenience of the discussion, in later sections we name pentafluorophenylalanine as **Z** for brevity, and *para*-X-tetrafluorophenylalanine as **ZpX** with X specifying the *para* substituent (e.g.,

ZpOH stands for tetrafluorotyrosine). By using the starting material with desired stereochemistry, the D- or L- isomer of ZpX derivatives can be readily synthesized in gram quantities. Consistent with the native analogues in human body, all the amino acids described below are in the L-configuration; however, the synthetic protocols should enable the synthesis of the D-isomers as well.

## 1.2 Synthesis of *para*-X-tetrafluorophenylalanine (ZpX) series

### 1) The regioselective $S_{NAr}$ reaction on pentafluorophenyl moieties

Fluorine has the highest electronegativity in the periodic table, so that the C-F  $\sigma$ -bond is highly polarized and the carbon atom is partially positively charged. Thus,  $S_{NAr}$  reactions are favorable for highly fluorinated aromatic compounds. On the other hand, fluorine is in the second row, and the  $p$ -orbital overlap with carbon is perfect in terms of both size and geometry. Thus, in fluorinated aromatic compounds, the dipole moment from the C-F  $\sigma$ -bond can force fluorine to partially donate its

benzene derivatives compared	$k_F/k_H$
	0.43
	57
	106

**Table 1-1.** Activation effects of fluorine substitution on different positions in the  $S_{NAr}$  reaction: ratios of measured rate constants (MeO<sup>-</sup>/MeOH, 58 °C)<sup>[40]</sup>

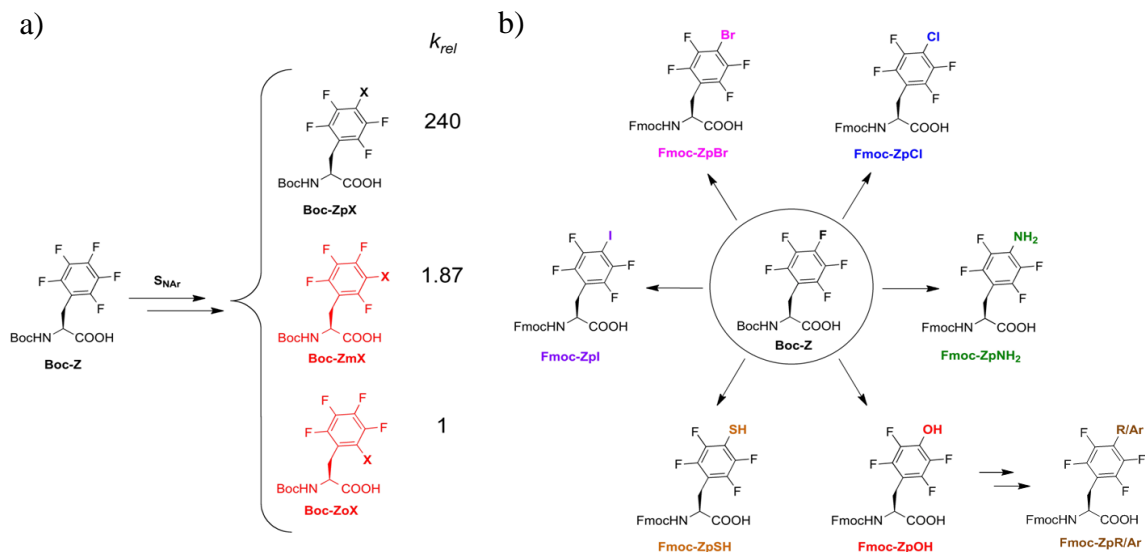


lone pair to the aromatic ring.

By combining the effects of fluorine as a  $\sigma$ -electronegative but  $\pi$ -electropositive group, we can conclude that as an electron withdrawing group,  $m\text{-F} > o\text{-F} > p\text{-F}$ . Firstly,  $m\text{-F}$  does not have the electron donating resonance effect and serves as the strongest one, while between  $o\text{-F}$  and  $p\text{-F}$ ,  $o\text{-F}$  has stronger electron withdrawing inductive effect due to the shorter distance. Therefore, for 2, 3, 4, 5, 6-pentafluorophenyl moieties, the  $S_{\text{NAr}}$  reaction will regioselectively occur at the 4-position, because the 4-position has two  $m\text{-F}$  and two  $o\text{-F}$  substitutions and all the other positions have one electron donating  $p\text{-F}$  substitution<sup>[39]</sup>.

Additionally, the empirical data of the activation effects of different positions of fluorine substitution in the  $S_{\text{NAr}}$  reaction are shown in Table 1-1<sup>[40]</sup>. Thus, by calculating the acceleration ratios for a pentafluorophenyl moiety (like the side chain of Z), we can see the regioselectivity is more than two orders of magnitude, with the *para*-fluorine substituted exclusively (Figure 1-10a).

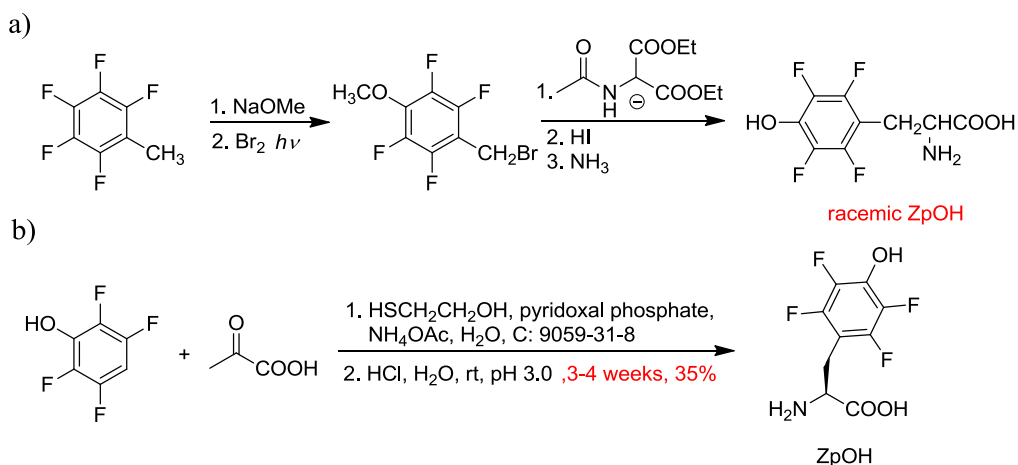
Based on this fact, we hypothesized that we can directly synthesize a series of ZpX amino acids from Boc-Z-OH (abbreviation for  $\alpha$ -amino Boc-protected,  $\alpha$ -COOH free, same below). The nucleophiles can be hydride, oxygen nucleophiles, nitrogen nucleophiles and sulfur nucleophiles<sup>[40]</sup>. Furthermore, from the readily accessible 4-substituted products, we can continue modifying the group and expand the scope of this class of unnatural amino acids (Figure 1-10b).



**Figure 1-10.** Expanding the repertoire of fluorinated aromatic amino acids *via*  $S_{\text{NAr}}$  chemistry. a) Regioselectivity of the  $S_{\text{NAr}}$  reaction on perfluorophenylalanine. b) Scope of the  $S_{\text{NAr}}$  reaction of Boc-Z to obtain ZpX derivatives with diverse functionality.

## II) The “O” nucleophile and the subsequent cross-coupling reactions

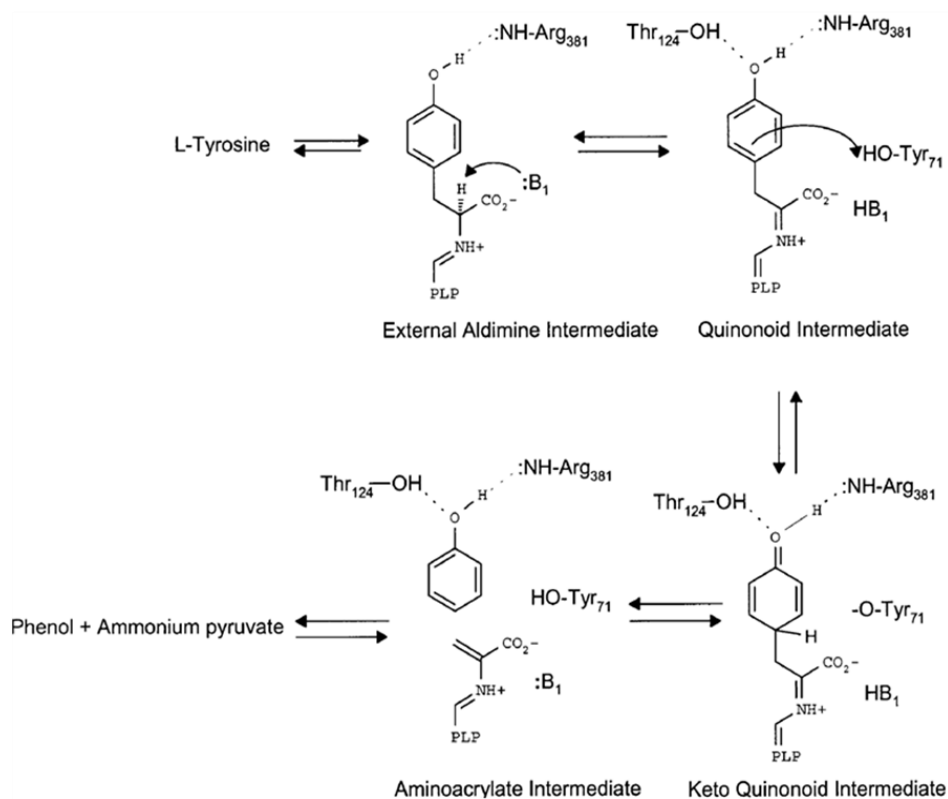
ZpOH (tetrafluorotyrosine) serves as a powerful probe to elucidate the mechanisms of a number of tyrosine-tailoring enzymes as we previously introduced. Therefore, it came



**Scheme 1-1.** Previously reported synthetic routes for ZpOH.

as a surprise to us that there was no reported chemical synthesis of ZpOH that is stereospecific. The only chemical synthesis of ZpOH was reported by Kang et. al. that couples 4-methoxy-tetrafluorobenzyl bromide to diethylacetamidomalonate to give the racemic mixture of O-methyl protected ZpOH (Scheme 1-1a)<sup>[41]</sup>.

A biosynthetic route for preparing fluorotyrosine analogues in the enantiomerically pure form has been developed by Cole et al. and Stubbe et al.<sup>[35, 37]</sup> This approach capitalizes on the enzymatic activity of tyrosine phenol lyase that converts fluorinated phenols to the corresponding L-fluorotyrosines in presence of pyruvate and ammonia (Scheme 1-1b). While this strategy works well with phenol (< 1 d, > 90% yield), the



**Scheme 1-2.** The enzymatic mechanism of tyrosine phenol lyase.

tetrafluorinated analogue appears to be a poor substrate for the enzyme. Consequently, the preparation of ZpOH takes 3-4 weeks and requires high concentrations of the enzyme. This phenomenon is explained by the enzymatic mechanism, in which a key step is the Michael addition between the aminoacrylate intermediate and phenolate (Scheme 1-2)<sup>[42]</sup>. This step is unfavorable both thermodynamically and kinetically when the phenol substrate is electron deficient. The experimental data also showed that catalysis efficiency drops with increasing number of fluorine substituents.

As we described previously, we chose Boc-Z-OH as the starting material to synthesize ZpOH via the S<sub>N</sub>Ar reaction. Our first attempt directly using hydroxide as the nucleophile

entry	nucleophile (solvent)	Yield <sup>b</sup> (%)
1	KOH ( <i>t</i> BuOH); KOH (DMSO); NaOH (H <sub>2</sub> O)	0
2	KOtBu (MeOH) → KOMe (MeOH) <sup>a</sup> , 48 h	>90
3	NaOMe(MeOH), 36 h	>90
4	Na (allyl alcohol) → NaOallyl (allyl alcohol) <sup>a</sup>	88

<sup>a</sup> The nucleophile shown on the right-hand side of the arrow was generated *in situ*. <sup>b</sup> 0% yield means no product was identified by crude <sup>19</sup>F NMR.

**Table 1-2.** Summary of the S<sub>N</sub>Ar conditions to make ZpOH

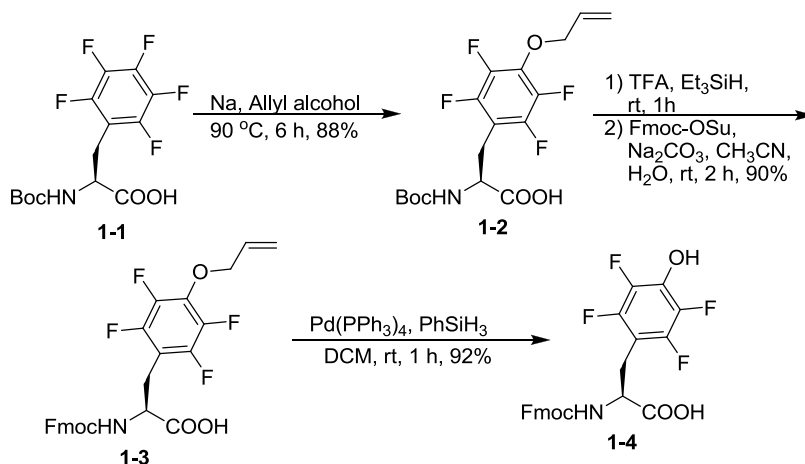
failed to give any product in organic or aqueous media (Table 1-2). One possible reason is that the hydroxide anion serves as both a hydrogen bonding donor and acceptor, and is solvated tightly through hydrogen bonding and strong dipole-dipole interactions in either water or polar organic solvents (it is insoluble in non-polar solvents). The thick layer of solvent molecules surrounding hydroxide reduces the reactivity. Hence we changed the nucleophile to methoxide, which only acts as a hydrogen bonding acceptor and so will not be solvated as strongly. The reaction proceeded nicely to completion with methoxide as we expected. Sodium as the counter-ion afforded a slightly

faster reaction than potassium presumably because sodium is a stronger Lewis acid (Table 1-2).

However, difficulty was encountered with the removal of the O-methyl protecting group:  $\text{BBr}_3$  resulted in a complex mixture; even the milder alternative  $\text{NaSC}_2\text{H}_5$ <sup>[43]</sup> caused the cleavage of the Boc- protecting group, which made the purification difficult. Finally, we used sodium allyloxide, which reacted as readily with Boc-Z-OH (Table 1-2) to give the O-allyl protection on the side chain. The advantage of the O-allyl protection is that the allyl ether can be cleaved specifically under mild conditions with a Pd catalyst<sup>[44]</sup>.

Our final synthetic route for Fmoc-ZpOH-OH is shown in Scheme 1-3. Sodium allyloxide was prepared *in situ* by dissolving sodium metal in allyl alcohol, and the  $\text{S}_{\text{NAr}}$  reaction was complete within 6 h at 90 °C to yield Boc-ZpOAllyl-OH (compound **1-2**). The protecting group on the  $\alpha$ -amino position was converted from Boc- to Fmoc- in one pot, and the allyl ether was then cleaved by using  $\text{Pd}(\text{PPh}_3)_4$  in presence of  $\text{PhSiH}_3$ <sup>[44]</sup> to yield the final product Fmoc-ZpOH-OH (compound **1-4**). Although the allyl deprotection step generates some greasy byproducts that are hard to separate on a silica gel column, they can be easily removed by ether extractions at pH 12 (Fmoc-ZpOH-OH will be double negatively charged and stay in the aqueous phase). We confirmed the absence of epimerization by conjugating Fmoc-ZpOH-OH to a chiral amine, the product of which elutes as a single peak during HPLC analysis. The lack of epimerization under such basic conditions is presumably because the negatively charged carboxylate prohibits the deprotonation on the  $\alpha$ -carbon. The synthetic route has an overall yield of 72% and can easily afford gram quantities of Fmoc-ZpOH-OH for peptide synthesis (ZpOH does not

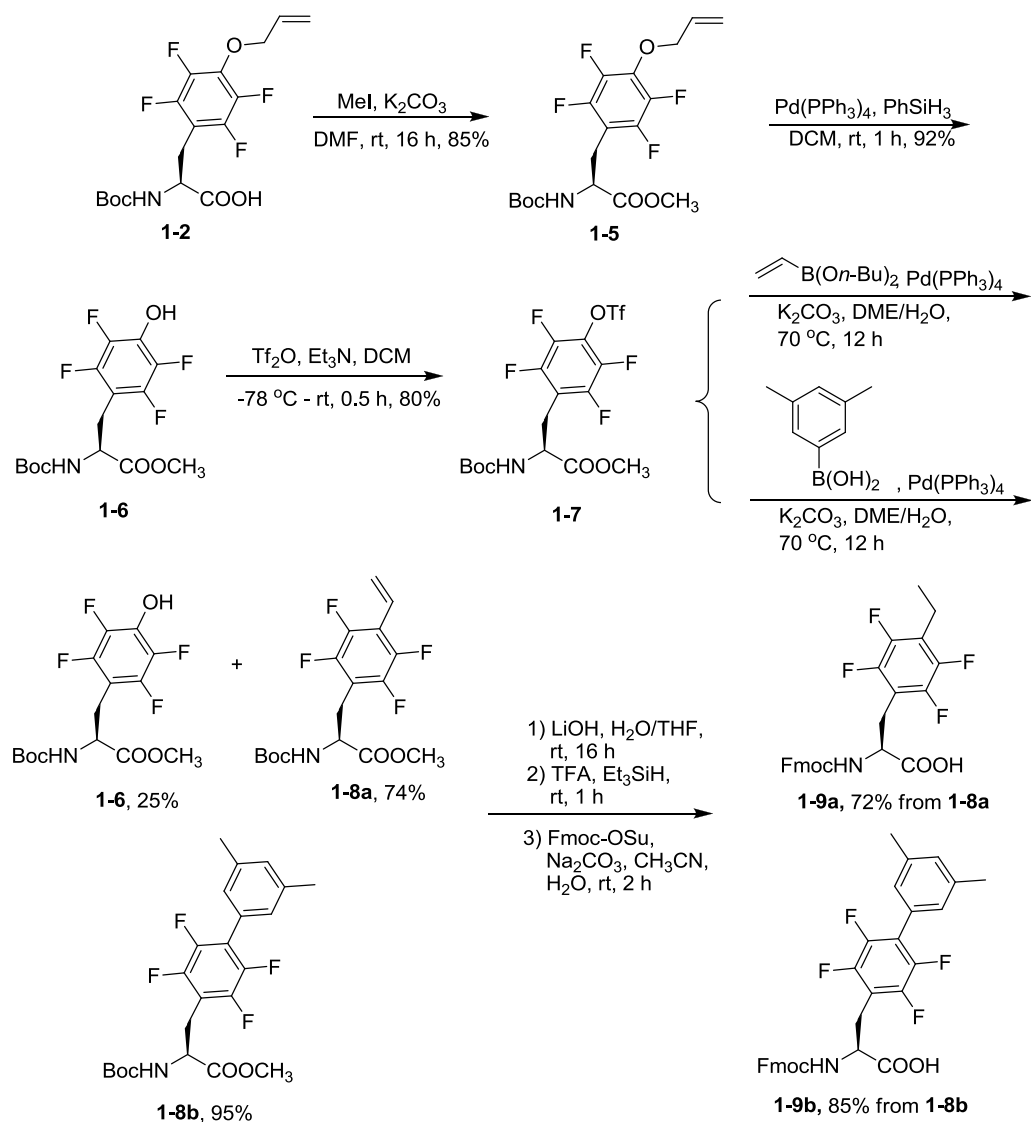
require side chain protection in regular Fmoc solid phase peptide synthesis (SPPS)<sup>[45]</sup>. If the side chain protection is necessary, the intermediate Fmoc-ZpOAllyl-OH (compound **1-3**) is compatible and the allyl deprotection can be conveniently carried out either in solution or on resin after the peptide synthesis.



**Scheme 1-3.** Synthesis of ZpOH.

By slightly modifying the synthetic route, we can obtain Boc-ZpOTf-OMe (compound **1-7**, Scheme 1-4), which can serve as a starting material for cross-coupling reactions in order to extend the carbon-carbon chain on the perfluorophenyl side chain. In this way we can generate different types of perfluoroaryl-aryl, aryl-vinyl and aryl-alkyl compounds, and utilize them to increase the hydrophobicity and investigate the strengthened  $\pi$ - $\pi$  interactions in peptides.

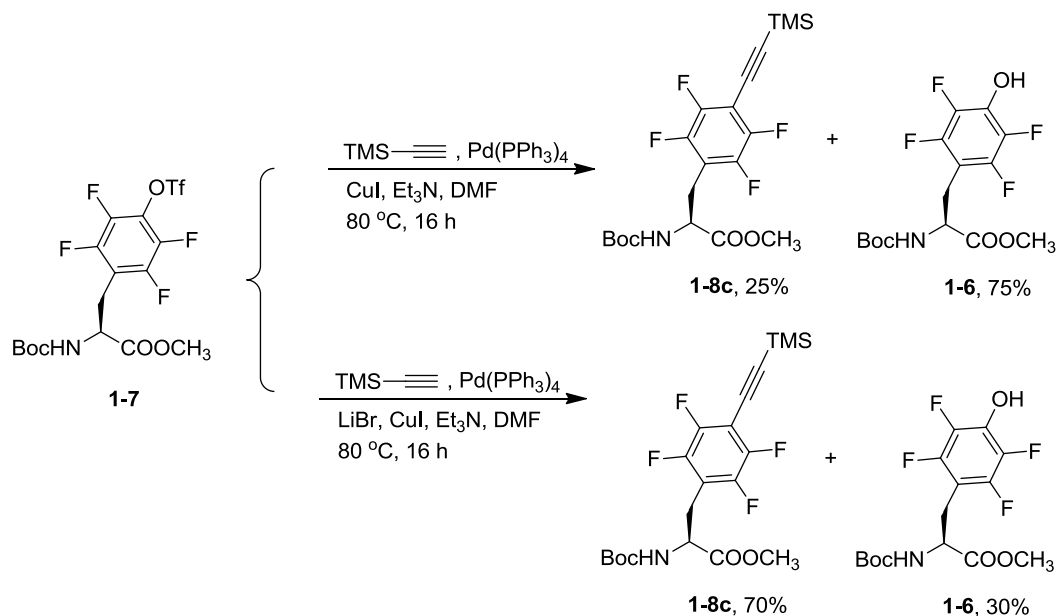
The synthesis of Boc-ZpOTf-OMe begins with the protection of the  $\alpha$ -carboxyl of Boc-ZpOAllyl-OH (compound **1-2**) using iodomethane. After methylation and the cleavage of the allyl ether, Boc-ZpOH-OMe (compound **1-5**) was treated with Tf<sub>2</sub>O to produce Boc-ZpOTf-COOMe with an overall yield of 55% over 4 steps.



**Scheme 1-4.** Synthesis of ZpOTf and the related Suzuki coupling reactions

The first cross-coupling reaction we attempted was the Suzuki coupling between Boc-ZpOTf-COOMe and 3, 5-dimethylphenylboronic acid<sup>[46-47]</sup> (Scheme 1-4). The reaction proceeded smoothly with 95% yield. The resulting ZpAr product **1-8b** was then converted to Fmoc-protected amino acid sequentially by the hydrolysis of the methyl ester with LiOH, deprotection of Boc with TFA and Fmoc protection on the  $\alpha$ -amino

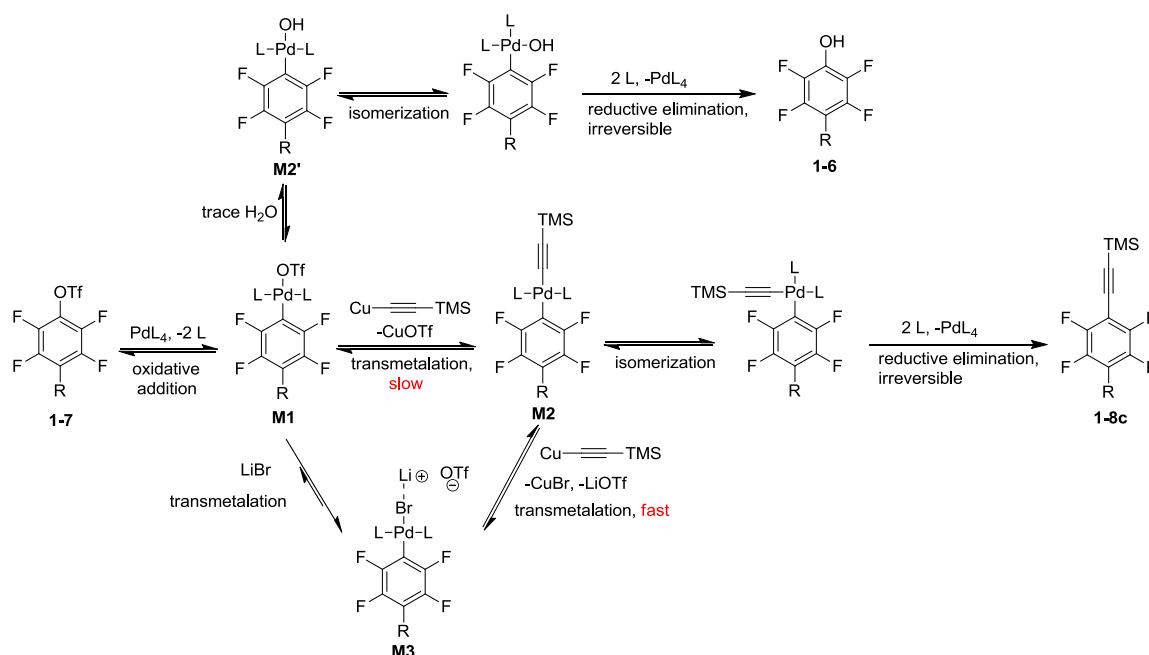
group in one pot. The overall yield for the final product **1-9b** was 81 % from Boc-ZpOTf-COOMe. Marfey's test proved that no epimerization happened<sup>[48]</sup>.



**Scheme 1-5.** The Sonogashira reaction with ZpOTf

The Suzuki coupling was not as perfect when we started with vinylboronic acid dibutyl ester, in which only 75% of ZpOTf was converted to ZpVinyl product **1-8a**, while the other 25% went back to ZpOH (**1-5**). The same problem became even more serious for Sonogashira coupling<sup>[49]</sup> (Scheme 1-5), in which the ratio of the desired product **1-8c** versus ZpOH was only 1/3. However, upon an addition of 2 equiv. of LiBr, the yield of **1-8c** can be increased to 70%. One more interesting phenomenon happens in the Boc-deprotection for **1-8a**, in which the vinyl group was 100% reduced by the carbocation scavenger triisopropylsilane, and the final product turned out to be Fmoc-ZpEt-OH (**1-9a**, Scheme 1-4). Alternative methods are then necessary to incorporate ZpVinyl into peptides.





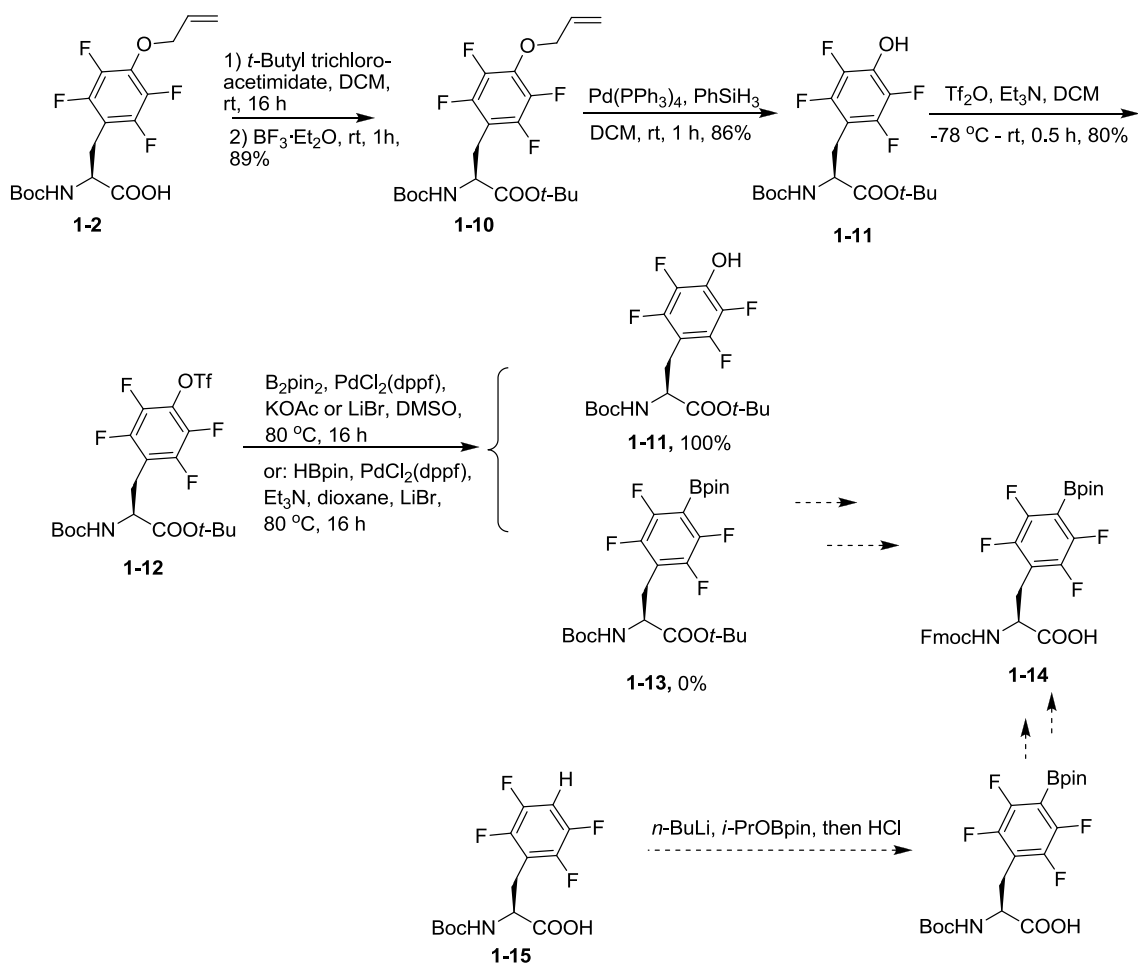
**Scheme 1-6.** A possible mechanism for the generation of ZpOH in Sonogashira reaction

The reason why ZpOH was generated is hypothesized with the Sonogashira reaction mechanism as an example (Scheme 1-6). Firstly the initial oxidative addition yields the Ar-Pd-OTf intermediate (**M1**), which is supposed to undergo transmetalation with alkyne cuprous (yielding **M2**), *trans-cis* isomerization and irreversible reductive elimination to yield the coupling product **1-8c**. However, if the transmetalation is not fast enough, a ligand exchange may take place between **M1** and the trace amount of water in the system in presence of base. This side reaction generates Ar-Pd-OH (**M2'**), which can also potentially undergo reductive elimination and generate ZpOH (**1-6**). Although this reductive elimination does not seem favorable on the electron deficient ring, it is irreversible and may not be the rate determining step. The generation of **M2'** competes with that of **M2**, which explains all the previous results. In Suzuki coupling, the boronic acid is very reactive and ZpOH was not observed, while the boronic ester is not as

reactive so that the generation of ZpOH became possible. In the Sonogashira reaction, the reactivity of the cuprous reagent is even poorer, so the byproduct was obtained in as much as 75% yield.

The addition of LiBr potentially leads to another transmetalation with **M1**, which should be favorable due to the hard-soft Lewis acid and base theory (HSAB). The hard pair  $\text{Li}^+$  and  $\text{TfO}^-$  associates and the soft pair Pd(II) and  $\text{Br}^-$  forms the intermediate Ar-Pd-Br (**M3**). Also based on HSAB, **M3** is favorable to transmetalate with alkynyl cuprous and yield **M2** due to the formation of the stable CuBr, which can also be additionally guided by the presence of  $\text{Li}^+$ . The summary of these two transmetalations is the acceleration of the **M1** to **M2** conversion. Since the generations of **M2** and **M2'** are competing processes, the side reaction is suppressed simultaneously.

Furthermore, we tried Miyaura borylation<sup>[50]</sup> with ZpOTf in order to obtain boronic acid-based amino acids (Scheme 1-7, the related application will be introduced in section **1.3**). One modification on the ZpOTf analogue is that *t*-butyl was used as the  $\alpha$ -carboxyl protecting group instead of methyl, because the pinacol ester in the potential product ZpBpin (compound **1-13**) is sensitive to aqueous acid or base, which is essential for the methyl ester cleavage. *t*-Butyl ester, which can potentially be cleaved together with Boc- under water-free conditions (details will be introduced in section **1.5**), is obviously a better choice. The *t*-butyl ester was formed with *t*-butyl trichloroacetimidate<sup>[51]</sup>, and Boc-ZpOTf-*Ot*Bu was synthesized efficiently. Unfortunately, Miyaura borylation failed to yield any desired product under various conditions (Scheme 1-7), and LiBr is not helpful here. This failure is not too surprising given our previous hypothesis on the mechanism.

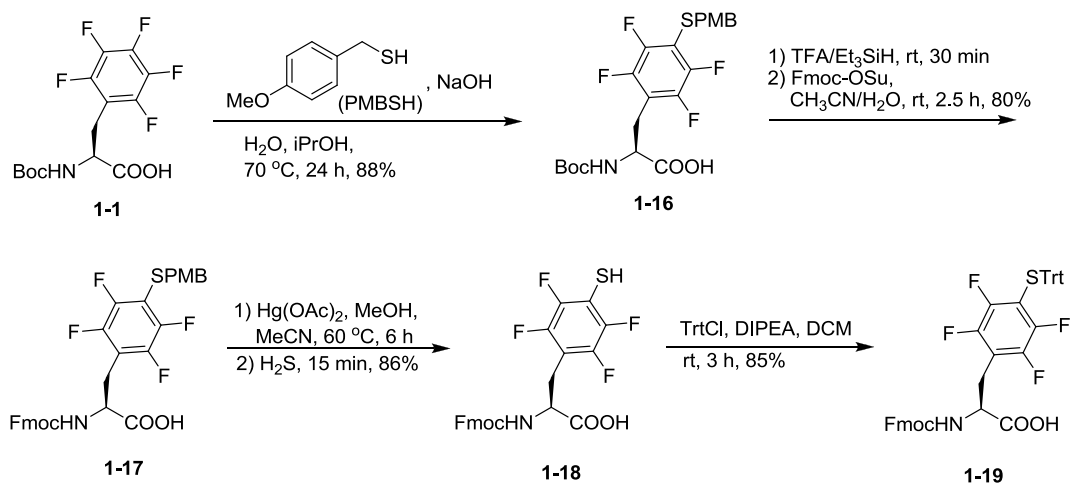


**Scheme 1-7.** Miyaura borylation with ZpOTf

Since the transmetalation acceleration based on the cuprous-bromide soft-soft match in Sonogashira coupling does not apply in the Miyaura reaction, in which  $\text{B}_2\text{pin}_2$  or  $\text{HBpin}$  takes the place of the cuprous reagent, and both of them are hard Lewis acids. An alternative synthesis for Fmoc-ZpBpin-OH (compound **1-14**) is proposed using Boc-ZpH-COOH (compound **1-15**) directly (Scheme 1-7)<sup>[52]</sup>.

### III) The “S” nucleophile

Similar to the ZpOH synthesis, the use of a sulfur nucleophile (e.g., thiolate) in the  $S_{\text{NAr}}$  reaction afforded the *para*-mercapto analogue ZpSH. Initially, we treated Boc-Z-OH with NaSH, which turned out to be too reactive and predominantly gave a thioether with two Z molecules crosslinked as a result of sequential  $S_{\text{NAr}}$  reactions within 1 h.



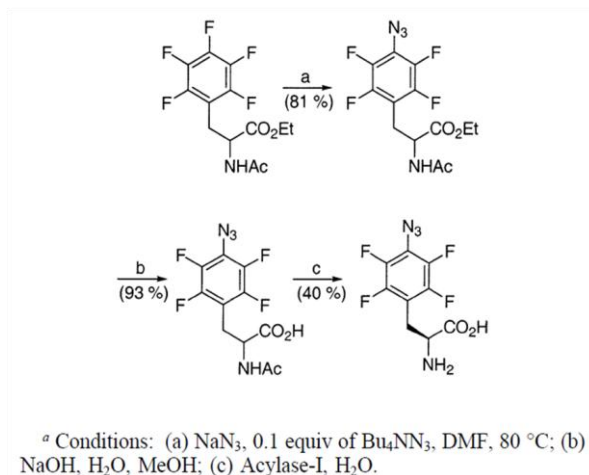
**Scheme 1-8.** Synthesis of ZpSH series

To prevent dimerization, we chose 4-methoxy benzenemethanethiol (PMBSH, Scheme 1-8) as an alternative nucleophile, which reacted readily with Boc-Z-OH in the presence of NaOH in *i*PrOH/H<sub>2</sub>O co-solvents to give Boc-ZpSPMB-OH (compound **1-16**) in 88% yield. Similar protecting group conversion yielded the product Fmoc-ZpSPMB-OH (compound **1-17**), which is compatible with SPPS via Fmoc/*t*Bu chemistry. No epimerization was observed by Marfey's test. The PMB protecting group was efficiently removed to give Fmoc-ZpSH-OH (compound **1-18**) by using Hg(OAc)<sub>2</sub> under mild conditions<sup>[53]</sup>. A similar protocol was also previously used for removing the PMB group from peptides in solution<sup>[54]</sup>. The side chain thiol of Fmoc-ZpSH-OH was then re-protected with trityl chloride, since trityl thioether can be directly deprotected along with

other regular protecting groups in Fmoc chemistry by TFA treatment. The overall yield for Fmoc-ZpSTrt-OH (compound **1-19**) is 51% over 4 steps.

#### IV) The “N” nucleophile and the Sandmeyer reactions

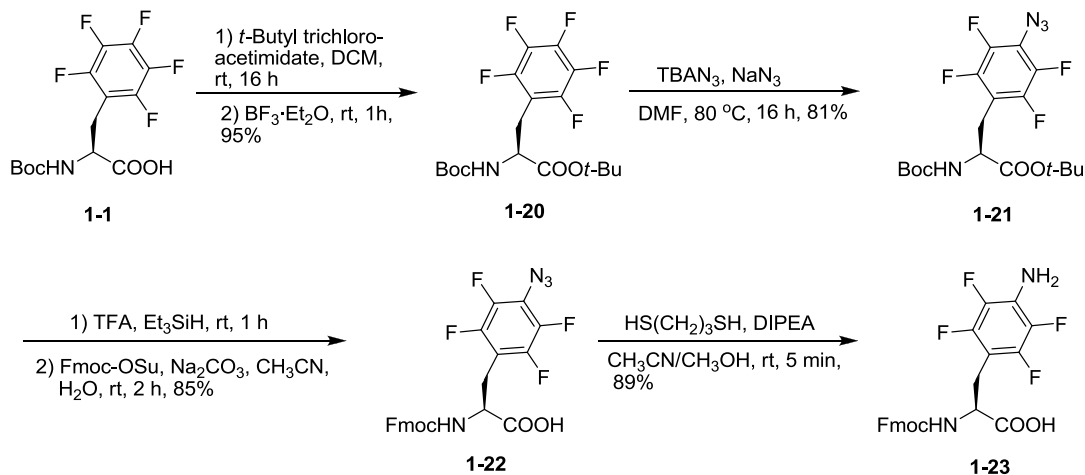
The  $S_{\text{NAr}}$  chemistry of Boc-Z-OH also applies to nitrogen nucleophiles. Ghadiri et. al. previously reported the synthesis of racemic ZpN<sub>3</sub> through the  $S_{\text{NAr}}$  chemistry (Scheme 1-9)<sup>[55]</sup>. In their work, the L-isomer of ZpN<sub>3</sub> was obtained through kinetic resolution with *Aspergillus melleus* acylase-I and incorporated into peptides as a photoactive group.



**Scheme 1-9.** Synthesis of ZpN<sub>3</sub> by Ghadiri et. al.<sup>[55]</sup>

We carried out a similar synthesis using the pure L-enantiomer of Boc-Z-OH (Scheme 1-10). The *t*-butyl protected substrate was treated with Bu<sub>4</sub>NN<sub>3</sub>/NaN<sub>3</sub> in DMF, and Boc-ZpN<sub>3</sub>-OtBu (compound **1-21**) was obtained in 81% yield. No epimerization was observed based on Marfey’s test<sup>[48]</sup>. The cleavage of Boc- and *t*-butyl ester was finished together with TFA, and the α-amino was re-protected with Fmoc-OSu in one pot. The product Fmoc-ZpN<sub>3</sub>-OH (compound **1-22**) is compatible with regular Fmoc SPPS chemistry. The three-step synthesis gave Fmoc-ZpN<sub>3</sub>-OH (compound **1-22**) with an overall yield of 65%. Reduction of Fmoc-ZpN<sub>3</sub>-OH was best accomplished by using propylene dithiol under mild conditions to afford Fmoc-ZpNH<sub>2</sub>-OH (compound **1-23**) in excellent yield<sup>[56]</sup>. The reaction only took 5 min to

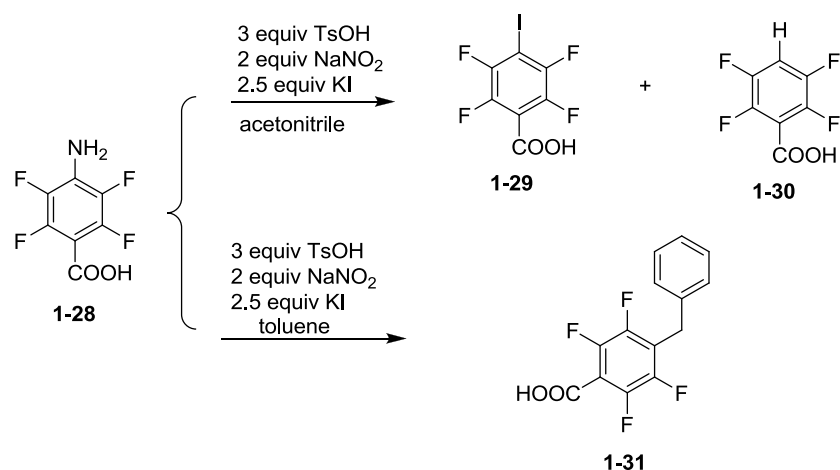
achieve completion due to the electron deficiency of the substrate. Propylene dithiol treatment effectively converted ZpN<sub>3</sub> in a peptide into ZpNH<sub>2</sub> as well; in other words, ZpN<sub>3</sub> conveniently serves as a precursor for the synthesis of peptides containing ZpNH<sub>2</sub>, the azido group can be considered as a “protected” amino group.



**Scheme 1-10.** Synthesis of ZpNH<sub>2</sub>

We then hypothesized that ZpNH<sub>2</sub> could be converted into a variety of fluorinated phenylalanine analogues through diazonium chemistry<sup>[57]</sup>. The Sandmeyer reaction is widely used to access halogenated aromatics<sup>[58]</sup>, and from this reaction we have developed synthetic protocols for the *para*-halogenated (Cl, Br, I) tetrafluorophenylalanines (ZpCl, ZpBr and ZpI). Unfortunately, since the heavily fluorinated anilines (ZpNH<sub>2</sub>) are less nucleophilic and more difficult to generate the diazonium species, the Sandmeyer chemistry is less straightforward. After a few initial failures, we decided to use the model compound **1-28** (*para*-carboxyl-tetrafluoroaniline, Table 1-3) to investigate the diazonium halogenation conditions. Indeed, treating **1-28** with the canonical condition of diazonium-iodination (TsOH·H<sub>2</sub>O/NaNO<sub>2</sub>/KI) only afforded less

than 60% conversion of the starting material, indicating inefficient formation of the diazonium species. In addition to the desired product **1-29**, the reaction yielded a major



solvent	1-29:1-30	solvent	1-29:1-30
H <sub>2</sub> O	<sup>b</sup>	DMSO	70:30
THF	0:100	CH <sub>3</sub> CN	75:25
DMAc	25:75	CH <sub>3</sub> CN <sup>a</sup>	95:5
DMF	30:70		

**Table 1-3.** Optimization of diazonium iodination conditions with the model compound.

<sup>a</sup>optimized condition: 0.05 M in CH<sub>3</sub>CN, 2 eq. I<sub>2</sub> and 1.5 eq. <sup>t</sup>BuONO. <sup>b</sup>neither **1-29** nor **1-30** was generated

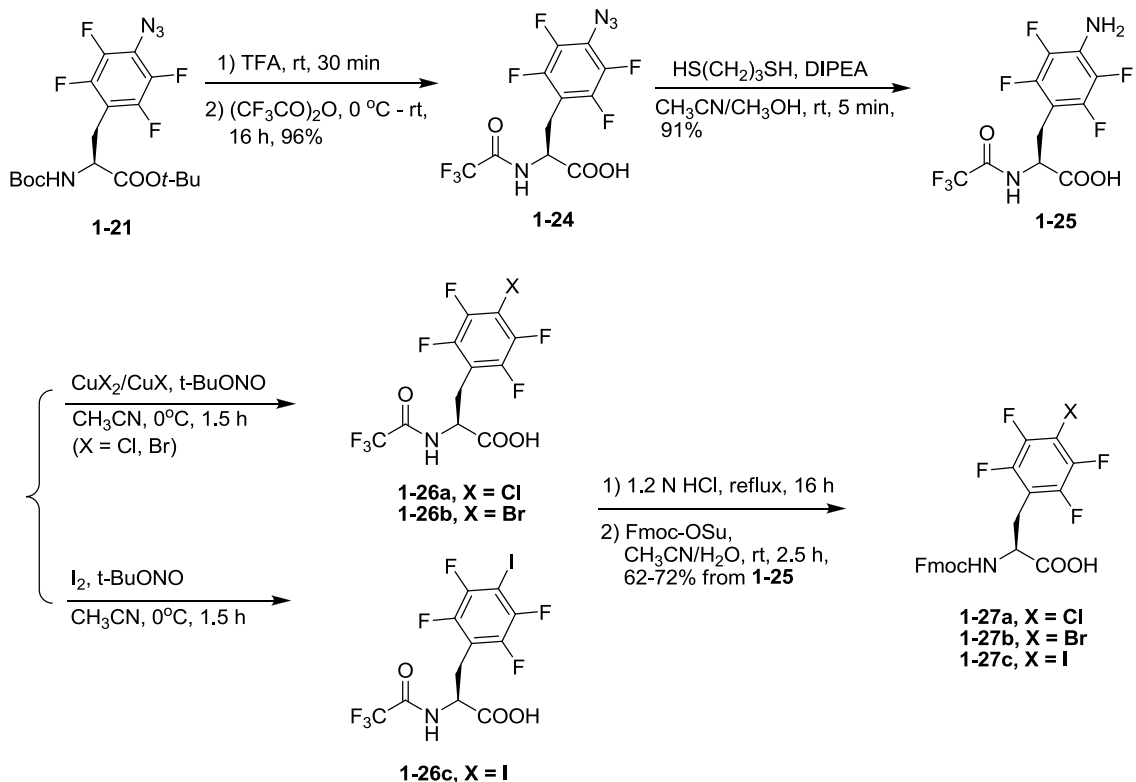
byproduct, the *para*-hydrogenated compound **1-30**. This byproduct is generated presumably because the diazonium intermediate and the subsequent phenyl radical are highly reactive and extract a hydrogen atom from the solvent molecules. This hypothesis is supported by the solvent dependence of byproduct formation: THF gives exclusively the hydrogenated byproduct, while CH<sub>3</sub>CN favors the desired iodo product (Table 1-3)

and affords a mixture of **1-29** and **1-30** in a ratio of 3:1. Interestingly compound **1-31** (with a *para*-benzyl substituent) is obtained as the sole product when toluene is used as a co-solvent. This is consistent with the radical mechanism of the diazonium chemistry, which leads to cross coupling between the fluorinated phenyl radical and toluene. This observation also led us to suspect that the benzylic position of TsOH might participate in the radical chemistry and therefore should be avoided in the diazonium chemistry. The alternative condition<sup>[59]</sup> (*t*BuONO/I<sub>2</sub>) reduced the byproduct formation to ~ 5%. Fortuitously, this condition also improved the diazonium formation to give complete conversion of the starting material.

The optimized condition for the model reaction was then implemented for the synthesis of the halogenated tetrafluorophenylanines. We initially subjected Fmoc-ZpNH<sub>2</sub>-OH to the iodination condition, which generated a messy mixture. NMR analysis suggested possible degradation of the Fmoc group, presumably because the benzylic substructure of Fmoc was attacked by the radical species as we discussed previously. Alternatively, we used the trifluoroacetyl to protect the  $\alpha$ -amino group (CF<sub>3</sub>CO-ZpNH<sub>2</sub>-OH, compound **1-25**, Scheme 1-11) in the diazonium chemistry, which does not bring additional hydrogen atoms to the substrate and is unfavorable to form radicals. The iodination with CF<sub>3</sub>CO-ZpNH<sub>2</sub>-OH afforded ZpI in excellent yield (89%). Similar to the model reaction, the byproduct ZpH was found to be ~5%. Use of CuX<sub>2</sub>/CuX (X = Br or Cl) in the diazonium chemistry gave ZpBr and ZpCl respectively in good yields. In comparison to iodination, the byproduct (ZpH) percentage was found to be similar for bromination and slightly higher for chlorination (~12%), consistent with the relatively



low reactivity of chlorine in radical chemistry. The protecting group conversion was accomplished through a one-pot process to give Fmoc-ZpX-OH (X = Cl, Br, I, compounds **1-27a-c**), the form ready for peptide synthesis. The enantiopurity of these amino acids was confirmed by the Marfey's test. HPLC purification for **1-27a-c** can efficiently remove ZpH byproduct if high purity is needed.

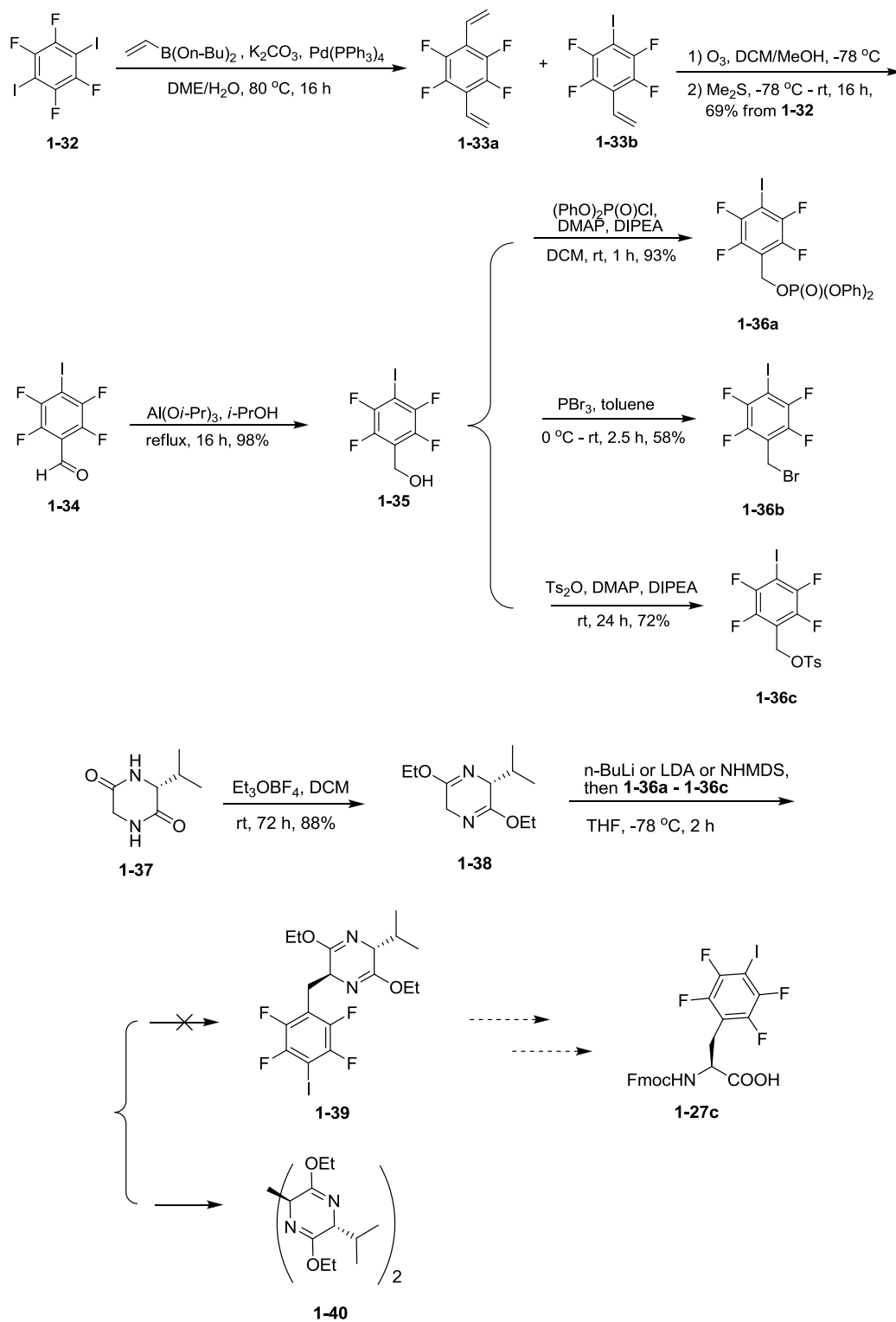


**Scheme 1-11.** Synthesis of ZpCl, ZpBr and ZpI

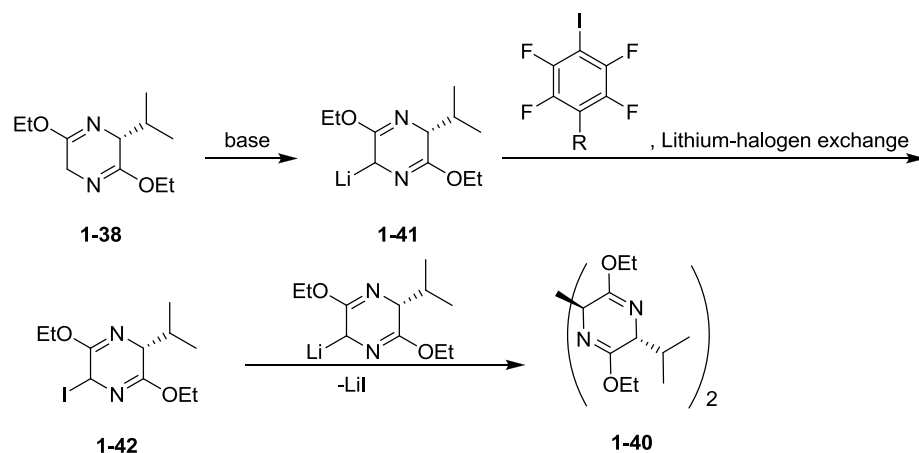
Here we want to highlight the advantage of our synthetic route for ZpXs by introducing another failed ZpI synthetic route (Scheme 1-12) we had attempted previously. Generally speaking, to synthesize an unnatural amino acid, one can either build up the backbone first and then modify the side chain functional groups, as we did in our ZpX series, or prepare the “side chain” first then couple it onto an amino acid

backbone precursor (usually also a chiral auxiliary). As shown in Scheme 1-12, we synthesized a series of para-iodotetrafluorobenzyl derivatives (compounds **1-36a-c**) with different leaving groups, and tried to couple them onto Schöllkopf's chiral auxiliary<sup>[60]</sup> (compound **1-38**, obtained from the double O-ethylation of **1-37** with triethyloxonium tetrafluoroborate<sup>[61]</sup>). The synthesis started from the Suzuki coupling between 1,4-diodotetrafluorobenzene and vinylboronic acid dibutyl ester. After ozonolysis, the mono-alkylated product **1-34** was isolated by silica gel chromatography as an aldehyde. **1-34** was reduced via Meerwein-Ponndorf-Verley reduction<sup>[62]</sup>, and the obtained alcohol **1-35** was then converted to **1-36a-c** with bromide, diphenyl phosphate or tosylate as the leaving group. Unfortunately, under the few attempted conditions, the coupling between any one of **1-36a-c** with **1-38** failed to yield the desired product, **1-39**, which can be converted to Fmoc-ZpI-OH (compound **1-27c**), while the major product turned out to be the dimerized chiral auxiliary (compound **1-40**).

One plausible mechanism for the generation of **1-40** is proposed (Scheme 1-13). The deprotonated **1-38** with  $\text{Li}^+$  as the counter-ion (compound **1-41**) can potentially undergo halogen-lithium exchange with any of **1-36a-c** to generate iodinated chiral auxiliary (compound **1-42**), which readily reacts with another molecule of **1-41** to yield the dimer **1-40**. The failure of these attempts indicates the reactivity of the ZpI side chain and the necessity to leave the iodination for a relatively later stage of the synthesis.



**Scheme 1-12.** A failed strategy to synthesize ZpI



**Scheme 1-13.** A plausible mechanism for the chiral auxiliary dimerization

### 1.3 Applications of ZpX series in biological systems

#### 1) 2D $^{19}\text{F}$ NMR analysis of peptides

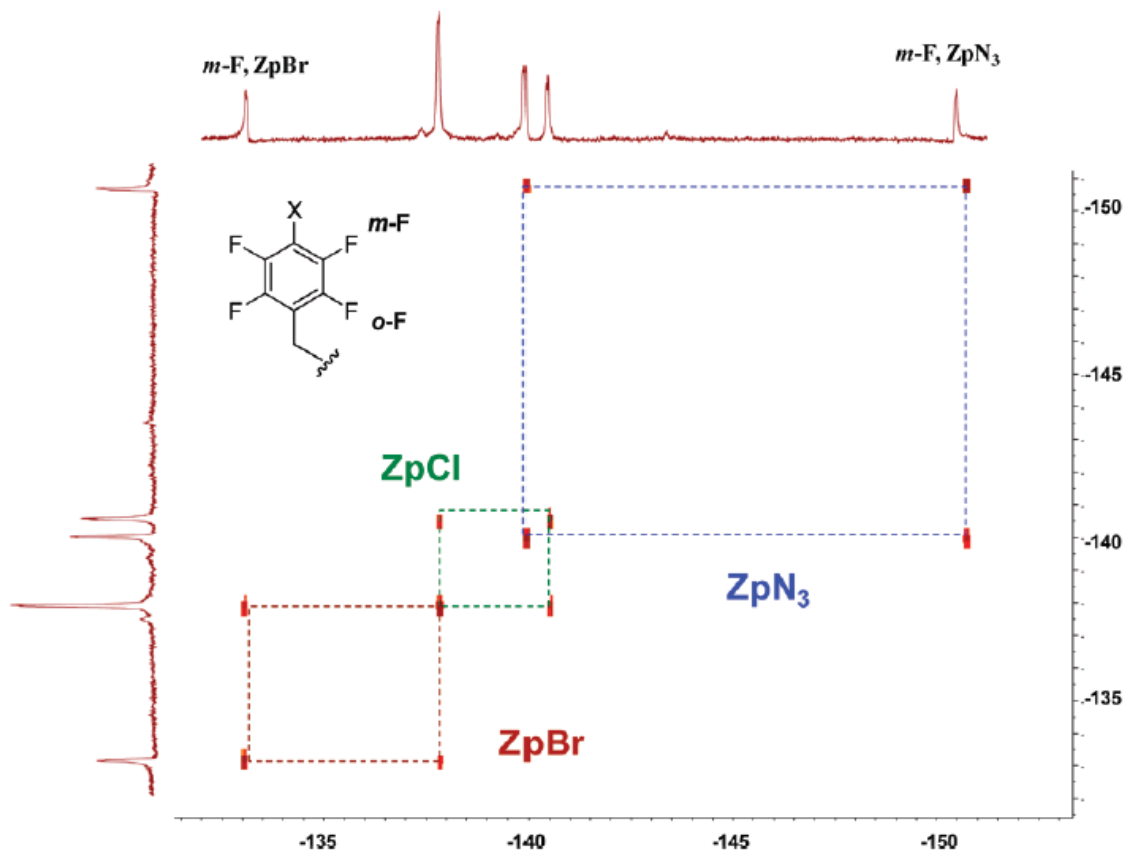
The application of  $^{19}\text{F}$  NMR in protein analysis was previously discussed in section 1.1. In order to utilize our novel amino acids as protein probes, we characterized their signatures in  $^{19}\text{F}$  NMR. Despite their structural similarity, the tetrafluorinated phenylalanine analogues display a wide distribution of their  $^{19}\text{F}$  resonances, ranging from -120 to -165 ppm (Table 1-4). The  $\pi$ -electron donating substituents (-F, -OH, -NH<sub>2</sub>) on the *para*- position (*para*- towards the  $\gamma$ -carbon of the amino acid backbone, same below) elicit significant upfield shift of the *meta*- fluorine resonances (Table 1-4, in Italic), which appear around -163 ppm. On the contrary, the bromo and iodo substituents on the *para*- position cause downfield shift of the *meta*- fluorines, with ZpBr and ZpI displaying a resonance around -133 and -120 ppm respectively.

	CDCl <sub>3</sub>			CD <sub>3</sub> OD		
<b>Z</b>	<i>-161.5</i>	<i>-154.4</i>	<i>-142.0</i>	<i>-165.0</i>	<i>-159.0</i>	<i>143.6</i>
<b>ZpH</b>	<i>-138.7</i>	<i>-142.6</i>		<i>-141.2</i>		<i>-144.4</i>
<b>ZpCl</b>	<i>-140.7</i>	<i>-141.3</i>		<i>-142.7</i>		<i>-143.7</i>
<b>ZpBr</b>	<i>-133.0</i>	<i>-140.8</i>		<i>-135.9</i>		<i>-142.3</i>
<b>ZpI</b>	<i>-120.1</i>	<i>-140.2</i>		<i>-123.1</i>		<i>-142.2</i>
<b>ZpN<sub>3</sub></b>	<i>-151.3</i>	<i>-142.5</i>		<i>-154.3</i>		<i>-144.6</i>
<b>ZpNH<sub>2</sub></b>	<i>-161.5</i>	<i>-145.3</i>		<i>-164.0</i>		<i>-148.0</i>
<b>ZpOH</b>	<i>-163.4</i>	<i>-144.8</i>		<i>-164.9</i>		<i>-147.4</i>
<b>ZpO<sup>-</sup></b>	-	-		<i>-168.2</i>		<i>-152.1</i>
<b>ZpSH</b>	<i>-137.3</i>	<i>-142.2</i>		<i>-139.8</i>		<i>-144.4</i>
<b>ZpS<sup>-</sup></b>	-	-		<i>-138.4</i>		<i>-148.7</i>

**Table 1-4.** <sup>19</sup>F NMR chemical shifts (ppm) of ZpX series in CDCl<sub>3</sub> and CD<sub>3</sub>OD, with the data for *meta*- fluorines in Italic (reference: fluorobenzene, -113.15 ppm).

We further examined the environment sensitivity of <sup>19</sup>F chemical shifts by using chloroform and methanol as solvents respectively (Table 1-4). Consistent with previous reports, the <sup>19</sup>F resonances are highly sensitive to the local environment: a polar solvent like methanol gives resonances that are upfield shifted by as much as 3 ppm in comparison to chloroform. On the other hand, the ionizable residues (ZpOH and ZpSH)

show dramatic chemical shift changes in response to the environment pH, as a result of protonation or deprotonation of their side chains.



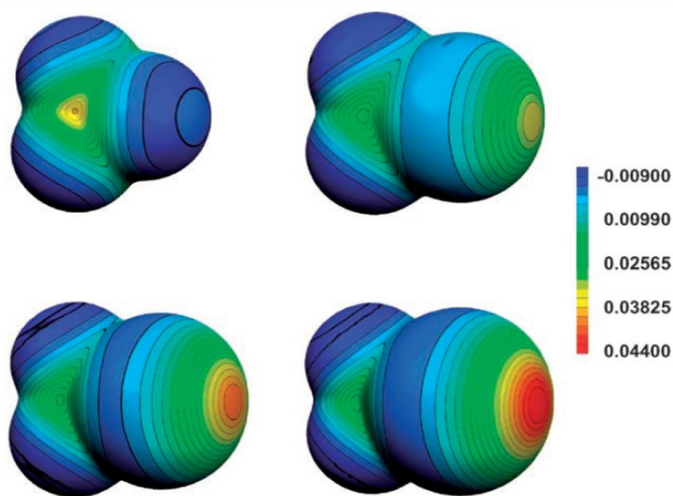
**Figure 1-11.**  $^{19}\text{F}$ - $^{19}\text{F}$  COSY spectrum of a magainin 2 mutant incorporating three ZpX (ZpBr, ZpCl, ZpN<sub>3</sub>) residues. The highly dispersed  $^{19}\text{F}$  resonances enabled facile assignment of all peaks in  $^{19}\text{F}$  NMR.

These characteristic chemical shifts should allow facile assignment of  $^{19}\text{F}$  resonances for peptides that incorporate multiple fluorinated residues. As a proof-of-concept experiment, we have incorporated three ZpX residues (ZpCl, ZpBr, and ZpN<sub>3</sub>) into magainin 2 (mag2), a well known membrane-lytic peptide<sup>[63]</sup>. Due to their distinct

chemical shifts, the meta-fluorine resonances of ZpBr and ZpN<sub>3</sub> were readily identified, which in conjunction with the <sup>19</sup>F-<sup>19</sup>F COSY<sup>[64]</sup> data enabled unambiguous assignment of all <sup>19</sup>F peaks (Figure 1-11).

## II) Halogen bonding in peptides

Halogens (Cl, Br and I) are known to feature anisotropic electron density around the atoms in organic halides, and a region of positive electrostatic potential is present along the covalent bond (Figure 1-12). Thus, the electrophilic pole on the halogen atom, which is known as the “σ-hole”, can accept *n* or *π* electrons of nucleophiles and hence form the halogen bonding (XB) interactions<sup>[65-66]</sup>. Main properties of halogen-bonding are listed



**Figure 1-12.** The molecular electrostatic potential: CF<sub>4</sub> (topleft), CF<sub>3</sub>Cl (topright), CF<sub>3</sub>Br (bottomleft) and CF<sub>3</sub>I (bottomright) (This figure has been reproduced from Mentrangolo, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G., *Angew. Chem. Int. Ed.*, **2008**, 6114-6127)

below:

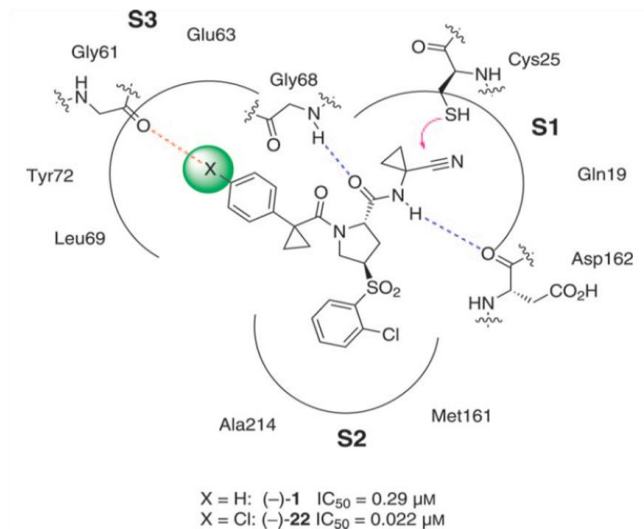
- 1) The strength of XB donor increases in the order of Cl < Br < I, C(sp<sup>3</sup>)-X < C(sp<sup>2</sup>)-X < C(sp)-X.
- 2) Neighboring electron withdrawing groups to the donor halogen enhance the XB strength.
- 3) The strength of XBs is comparable with that of

hydrogen bonds (HBs), in the range of 10-200 kJ/mol.

4) XBs are much more sensitive to steric hindrance, which is probably the biggest difference from HBs.

With these properties, XB has been utilized in liquid crystals, magnetic and conducting materials. As a unique mechanism of interaction, it is also attractive in the studies of biological systems. In 2011, Diederich et al. reported a systematic study of XB interactions in human proteinase Cathepsin L (hCatL) (Figure 1-13)<sup>[67]</sup>, in which the XB between Gly61 carbonyl and Cl, Br or I enhanced the protein-ligand binding affinity in the order of  $\text{Cl} < \text{Br} < \text{I}$ .

Our interest is to probe the effects of XB(s) on the secondary structure of biological molecules, and also the similarity and difference between XB(s) and other non-covalent interactions. Since haloarenes are usually good XB donors and adjacent electron-withdrawing groups can enhance XB<sup>[65]</sup>, ZpI is obviously excellent as a candidate to incorporate XB into peptides. The first model system we chose is the PinWW domain

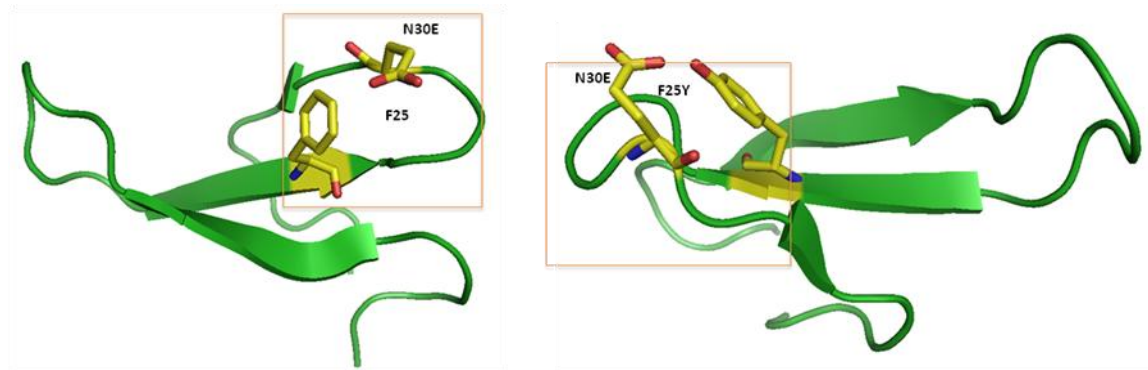


**Figure 1-13.** Halogen bonding interaction with amide carbonyl in the binding pockets of hCatL. (This figure has been reproduced from Hardegger, L. A.; Kuhn, B.; Spinner, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F., *Angew. Chem. Int. Ed.*, **2011**, 314-318)



(Figure 1-14), a 34-residue  $\beta$ -sheet peptide<sup>[68]</sup>, with which we were able directly compare the energetics of XB and HB in a solvent-exposed loop of a biologically relevant system.

Collaborating with C. J. Pace, we made a series of mutants of PinWW and compared their stability (Table 1-5). The control mutant is F25 N30E, in which Asn was substituted



**Figure 1-14.** Cartoon showing the 3D structures of two PinWW domain mutations, left: F25 N30E; right: F25YN30E.

PinWW mutant	$T_m$ (°C) <sup>a</sup>	$\Delta G_f$ (kcal/mol) <sup>b</sup>	$\Delta\Delta G_f$ (kcal/mol) <sup>c</sup>
F25 N30E	46.09 $\pm$ 0.14	-2.16 $\pm$ 0.04	--
F25Y N30E	52.32 $\pm$ 0.12	-2.91 $\pm$ 0.06	-0.75
F25Z N30E	52.82 $\pm$ 0.21	-3.25 $\pm$ 0.08	-1.09
F25ZpI N30E	52.23 $\pm$ 0.12	-3.23 $\pm$ 0.05	-1.07

<sup>a</sup>Calculated by fitting thermal denaturation curves

<sup>b</sup>Calculated by fitting chemical denaturation curves

$$^c\Delta\Delta G_f = \Delta G_f \text{ mutant} - \Delta G_f \text{ F25 N30E}$$

**Table 1-5.** PinWW mutants and the comparison of the stabilities.

by Glu to set a net negative charge to enhance the electron richness and flexibility on the HB/XB acceptor residue. By comparing F25Y N30E to F25 N30E, we can see the additional hydrogen bonding between Tyr and Glu can enhance the stability by 0.75 kcal/mol. However, a similar comparison between F25Z N30E and F25ZpI N30E did not show any significant difference.

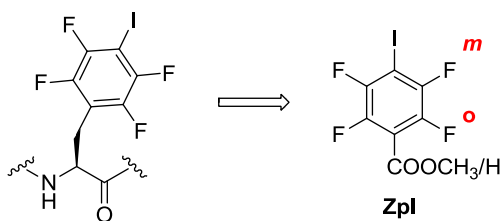
The failure of PinWW domain studies indicate that it is essential to systematically learn the unique characters of XB before incorporating XB into biological systems. There are a few unknown questions to answer:

- 1) What is the best XB acceptor that in peptides?
- 2) How strong can XB potentially be?
- 3) Can  $\pi$  clouds serve as XB acceptors<sup>[69]</sup> as well as lone pairs?
- 4) Can XB be significant in both hydrophobic and aqueous environments?
- 5) How much stronger is XB than the interaction between an electron-poor C-H and electron donors (this is argued as C-H hydrogen bonding or dipole-dipole interaction)?

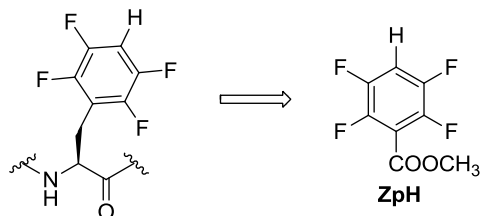
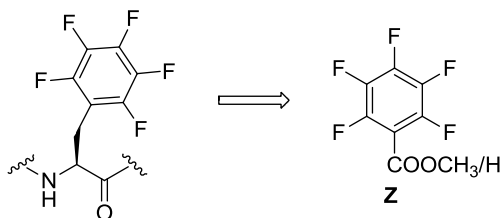
In order to find out the answers, <sup>19</sup>F NMR titration experiments were carried out to expose the nature of the interactions and obtain the association constants by calculating the chemical shifts upon interacting with electron donors. To simplify the system, we used *para*-iodotetrafluorobenzoic acid methyl ester in cyclohexane solvent to mimic ZpI residue in membrane environments, while *para*-iodotetrafluorobenzoic acid in water as ZpI residue exposed in aqueous environments (Figure 1-15). For convenience and consistency, we still call these small molecules “ZpI”. Similarly, some other small molecules are utilized as pseudo-“Z” and “ZpH” as controls (Figure 1-15).

In terms of the XB acceptors, the nitrogen lone pair is considered the best<sup>[70]</sup>, which can be explained by HSAB theory. Since aliphatic amines are protonated at physiological pH, aromatic nitrogens with lower  $pK_a$  values (pyridine and imidazole) were chosen in our experiments to stand for the relative amino acid residues (some are unnatural) in biological systems. We also tested electron-rich  $\pi$  systems (indole and pyrrole) to figure out if  $\pi$  clouds are also capable XB acceptors (Figure 1-15).

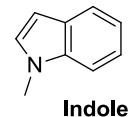
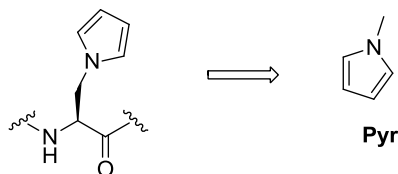
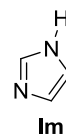
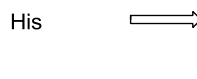
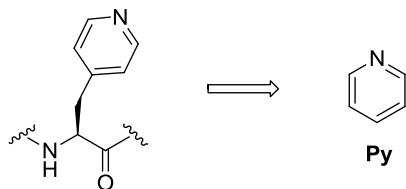
**Donors:**



**Controls:**

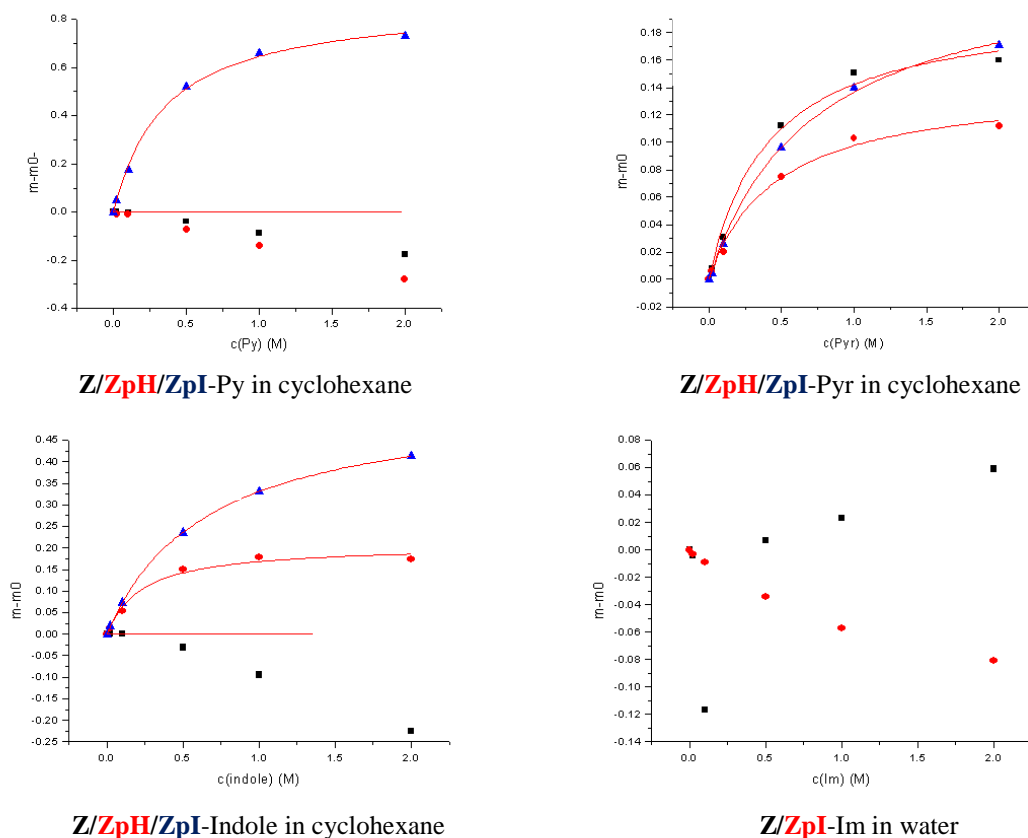


**Acceptors:**



**Figure 1-15.** Small molecules as amino acid mimics in the  $^{19}\text{F}$  NMR titration experiments

As we discussed previously, the chemical shifts of *meta*-fluorine (Figure 1-15, labeled on ZpI) reflect the environment changes most significantly. Hence we plotted the chemical shift changes of *meta*-fluorine (Figure 1-16, as  $m-m_0$ ) against the concentration of the XB acceptor and calculated the association energy based on the fitted hyperbolic curves (when the  $m-m_0$  value is half its maximum,  $K_a = 1/[\text{acceptor}]$ ).

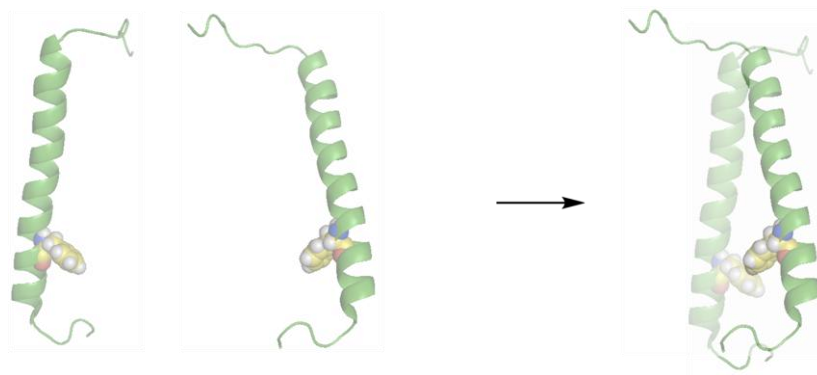


**Figure 1-16.** Titration curves of small molecule halogen bondings.

The results showed that in cyclohexane, XB interaction between ZpI and Py is significant, with  $K_d = 0.35 \text{ M}^{-1}$ ,  $\Delta G = -1 \text{ kcal/mol}$ . The *meta*-fluorine signal is shifted by close to 1 ppm. Under the same conditions, Z or ZpH do not interact with Py, which further indicated the interaction is specific. In contrast, Z/ZpH/ZpI are not significantly

different in terms of interacting with  $\pi$  systems (indole or Pyr). Due to the complexity of multiple possible weak interactions ( $\pi$ - $\pi$  face-face, edge-face, etc.), it's hard to interpret the halogen- $\pi$  interactions. On the other hand, neither Z nor ZpI showed any interactions with Im in water, even though Im bears more electronegative nitrogen lone pairs than Py. This fact indicates that in water (generally, protic solvents) the XB interaction is completely overwhelmed by strong hydrogen bonding with solvent molecules, which perfectly explains the failure of our PinWW system.

From our  $^{19}\text{F}$  NMR titration experiments, it is concluded that, in order to utilize XB in a biological system, we have to place it in hydrophobic environments (Diederich's XB is in water-free binding pockets<sup>[67]</sup>). Additionally, the aromatic nitrogen lone pair is the best candidate as the acceptor<sup>[65]</sup>. The high steric flexibility requirement should also be taken into account<sup>[65]</sup>.



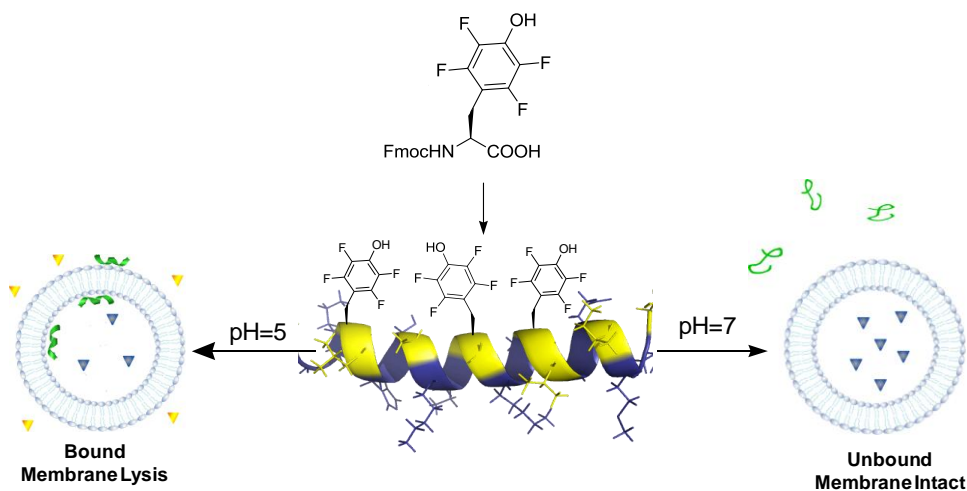
**Figure 1-17.** ErbB2 as the model system for halogen bonding

Combining all these considerations, we propose that XB may enable the control of protein-association in lipid membranes. Towards this end, we will use ErbB2 as the model system (Figure 1-17). ErbB2 is a transmembrane helix which dimerizes in membranes<sup>[71]</sup>. The ErbB2 system can potentially provide the essential hydrophobic

environment and also the geometry of XB interactions. Work along this direction is currently underway.

### *III) pH dependent membrane lytic peptides (in collaboration with F. Wang)*

Membrane-lytic peptides have attracted a great deal of attention in the search for novel antibiotics that are less prone to drug resistance<sup>[72-75]</sup>. For the same reason, membrane-lytic peptides also serve as potential anti-tumor drug candidates. However, the toxicity of these peptides to healthy mammalian cells limits their application as drug molecules. The design of pH-sensitive membrane-active peptides, which suppresses the membrane-lytic activity at physiological pH and increases the selectivity, has been largely relying on the protonation of Glu or Asp side chains, which display pK<sub>a</sub> values of 4.3 and 3.9 respectively. Residues with higher pK<sub>a</sub> values are desirable considering the pH values of (patho)physiological relevance are generally above 5.5<sup>[76]</sup>. Considering the pK<sub>a</sub> value of



**Figure 1-17.** ZpOH mutated magainin 2 serves pH dependent membrane-lytic peptide.

ZpOH is 5.6, it is a good candidate to provide novel pH sensing elements in the design of membrane-lytic peptides.

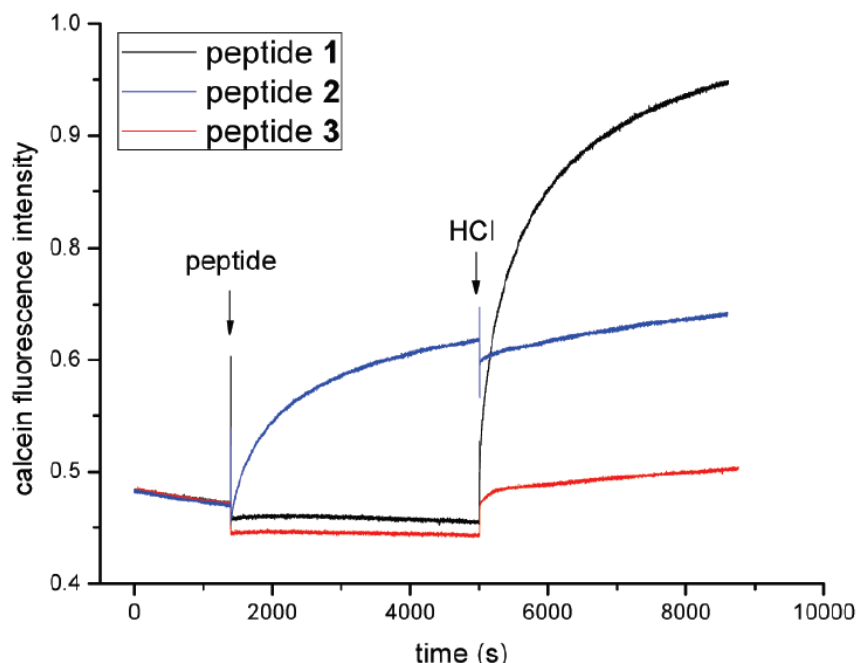
Again we used the well characterized membrane-lytic peptide magainin 2 (mag2) as the model system to test the potential of ZpOH in pH sensing (Figure 1-18)<sup>[63]</sup>. Mag2 is a twenty-three-residue peptide that folds into an amphiphilic helix upon membrane binding. Three mag2 variants were synthesized on the Rink-MBHA amide resin using the standard Fmoc/HBTU chemistry. A dansyl group was installed at the N-termini of all peptides to facilitate binding affinity measurement and concentration calibration. As shown in Table 1-6, peptide **2** is the corresponding Tyr control for the designed pH-responsive peptide **1**. For comparison, we also synthesized peptide **3** incorporating three Glu residues, which have been widely used as pH sensing elements in previous reports.

WT Mag2	GIGKFLHSAKKFGKAFVGEIMNS
<b>1</b>	dansyl-GIGK( <b>ZpOH</b> )LHSAKK( <b>ZpOH</b> )GKA( <b>ZpOH</b> )VGEIMNS
<b>2</b>	dansyl-GIGKYLHSAKKY GKAYVGEIMNS
<b>3</b>	dansyl-GIGKELHSAKKEGKAEVGEIMNS

**Table 1-6.** The Sequences of mag2 mutants, residues of interest are shown in bold: **ZpOH** (negative at pH7.0, neutral at pH5.0); **Y** (neutral in both); **E** ( negative in both)

The membrane-lytic activity of the mag2 mutants were tested using a calcein leakage assay<sup>[77]</sup> (Figure 1-18). At neutral pH (7.0), no membrane lysis was observed upon

mixing the peptides **1** or **3** with liposomes. The tyrosine variant **2** afforded ~18% of membrane damage. Acidification to pH 5.0 upon HCl addition caused little change for the control peptides **2** and **3**. In sharp contrast, the ZpOH mutant **1** displayed a dramatic increase in calcein fluorescence, showing that the ZpOH side chains indeed serve as effective pH sensors to turn on the membrane-lytic activity of mag2. Under neutral conditions (pH 7.0), no membrane damage was observed even at 10  $\mu$ M peptide concentration.



**Figure 1-18.** Kinetics of peptide induced cargo (calcein) leakage. The graph shows observed change in calcein fluorescence at 520 nm ( $\lambda_{\text{ex}}=485$  nm) upon addition of peptide (3  $\mu$ M) into LUVs of POPC (500  $\mu$ M lipids, pH 7.0), followed by the addition of 6 N HCl (12  $\mu$ l) to reach pH 5.0.

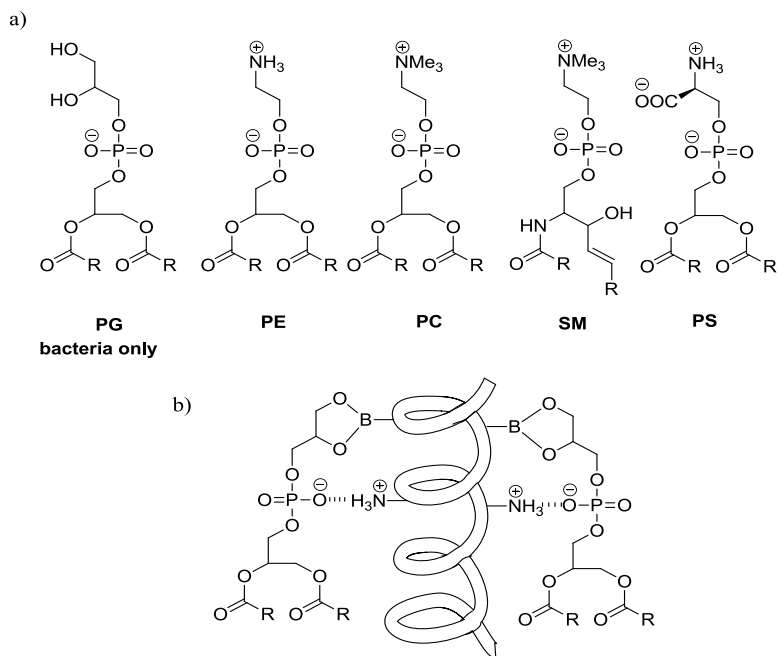
The leakage assay demonstrated that ZpOH functions as effective pH sensors in the design of “smart” membrane-lytic peptides. The mag2 mutant harboring ZpOH residues



responds quickly to the pH drop from 7.0 to 5.0. No pH activation was observed for the corresponding mutant that incorporates Glu residues as pH sensors. This is presumably because the pKa value of ZpOH falls into this physiologically relevant region, while that of Glu does not. In addition, the hydrophobic nature of the neutral ZpOH residue allows their incorporation onto the membrane-binding face of a peptide, whereas all previously reported pH sensing elements reside on the hydrophilic face of an amphipathic structure.

#### IV) Targeting bacterial membrane lipids

The design of the pH-sensitive membrane-lytic peptide increases selectivity based on the environmental difference. In this section we will introduce an alternative strategy: increasing the selectivity based on the membrane composition difference.



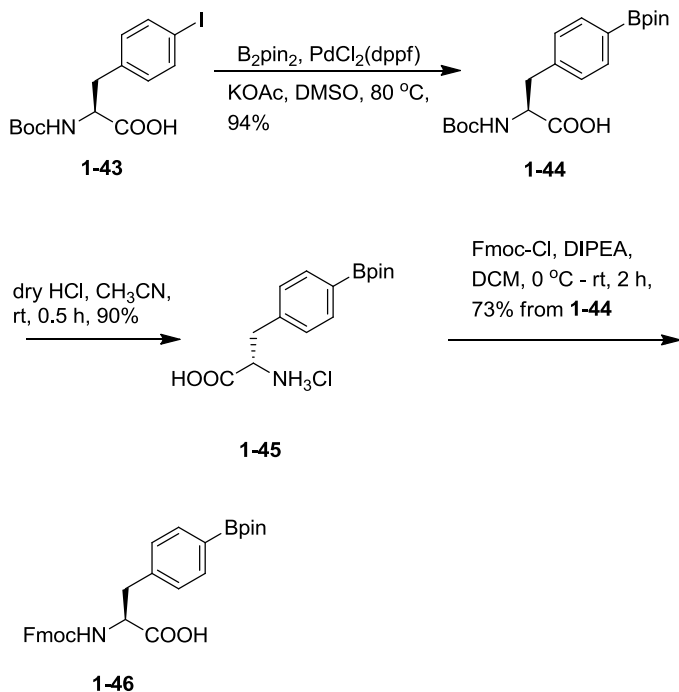
**Figure 1-19.** a) Structures of phospholipids. b) boronic acid based transmembrane helices chelating with PG

One significant difference between bacterial membranes and mammalian membranes is that bacterial membranes contain phosphatidyl-glycerol (PG) lipids (Figure 1-19a), which do not exist in mammalian cell membranes. Since among all the membrane lipids PG exclusively bears a 1,2-diol structure on its head, we can potentially incorporate functional groups that selectively bind diols into membrane-lytic peptides in order to increase the selectivity.

Boronic acids are known to chelate with 1,2-*syn*-diols<sup>[78]</sup>, and the boronic acid containing amino acid Fmoc-BpinPhe-OH (compound **1-55**, Scheme 1-14) was reportedly synthesized from Tyr and tested to chelate with the chromophoric diol alizarin<sup>[79]</sup>. Therefore, we hypothesized that magainin 2 mutants containing boronic acid can selectively chelate PG and bind bacterial membranes better than wild type. In other words, boronic acids can potentially enhance the membrane-lytic selectivity towards bacterial cells. Furthermore, tetrafluorinated BpinPhe (or ZpBpin, based on our naming system of ZpX series) is believed to chelate diols even stronger due to the electron deficiency<sup>[78]</sup>. Although the synthesis of Fmoc-ZpBpin-OH (compound **1-14**, see Section **1.2**) is still ongoing as we previously stated, we already synthesized the non-fluorinated Fmoc-BpinPhe-OH (compound **1-55**) as the control molecule.

Our synthesis of Fmoc-BpinPhe-OH was carried out with a different starting material, Boc-protected iodophenylalanine (compound **1-43**). The Miyaura borylation reaction between **1-43** and B<sub>2</sub>pin<sub>2</sub> was catalyzed by PdCl<sub>2</sub>(dppf) with KOAc as base in DMSO, and yielded Boc-BpinPhe-OH (compound **1-44**) in excellent yield (Scheme 1-14).

Since boronic acid pinacol ester is labile to hard Lewis bases, the Boc-Fmoc conversion was achieved under water-free conditions. Firstly, the Boc-group was cleaved by neat HCl in acetonitrile<sup>[79]</sup>, since the regular TFA condition also cleaved Bpin. After the cleavage is complete, the excess HCl was removed under vacuum, and the free amino acid H-BpinPhe-OH



**Scheme 1-14.** Synthesis of Fmoc-BpinPhe-OH.

(compound **1-45**) was isolated by filtration. The Fmoc-protection was carried out in DCM with Fmoc-Cl as the reagent, because the more often-used reagent Fmoc-OSu usually is only used in aqueous solutions. Since H-BpinPhe-OH is insoluble in DCM, the key point of this reaction is to suspend fine powders of H-BpinPhe-OH in DCM as even as possible, or the elimination of Fmoc-Cl will happen as a major side reaction and decrease the yield. The overall yield for the final product Fmoc-BpinPhe-OH (compound **1-46**) was 62% over 3 steps.

We already have incorporated BpinPhe into magainin 2 successfully, and the assay to demonstrate the PG-membrane selectivity is currently underway.

## 1.4 Conclusions

In summary, a series of novel fluoroaromatic amino acids has been synthesized that can be readily utilized in the chemical synthesis of proteins. Our synthetic protocol is fundamentally built on the regioselective  $S_{\text{NAr}}$  reaction of pentafluorophenylalanine Z and also associated with the cross-coupling reactions of ZpOTf and the diazonium chemistry of ZpNH<sub>2</sub>. the scope of the functional groups on the product side chains includes hydroxyl, alkyl/aryl, mercapto, amino and halogens. All the synthetic routes are short and efficient, showing no epimerizations.

We also incorporated ZpX series amino acids into peptides and demonstrated the potential applications of ZpXs as indicators to detect the local environment by a significant shift of their <sup>19</sup>F NMR signal, as halogen bonding agents to investigate novel non-covalent interactions, and as pH sensors to trigger highly selective membrane leakage. We expect the easy synthesis and unique physical properties of the ZpX series will make it suitable for an even wider array of applications in protein biochemistry and engineering.

## **Materials and Methods**

### **General Methods**

Chemicals were obtained from Fisher and Aldrich, unless otherwise indicated. Fmoc-OSu, Rink-Amide-MBHA resin, and O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate(HBTU) were purchased from Novabiochem (San Diego, CA). All natural Fmoc-protected amino acids were purchased from Advanced Chemtech (Louisville, KY). Boc-pentafluorophenylalanine was purchased from Peptech (Burlington, MA). 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was purchased from Avanti Polar Lipids (Alabaster, Alabama). Peptide synthesis was carried out on a Tribute peptide synthesizer (Protein Technologies, Tucson, AZ). <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data were collected on a Varian Gemini 400MHz, Varian 500 MHz and Varian 600 MHz NMR spectrometer. HR-MS data were generated by the Boston College Mass-Spec facility. The protein concentration of all samples used in this study was determined by measuring their absorption at 330nm on a Lambda 25 UV-Vis spectrometer (PerkinElmer, Waltham, MA) considering the extinction coefficient  $\epsilon$  (330nm) as 4,300 M<sup>-1</sup> cm<sup>-1</sup>. The liposomes were prepared by using Liposofast Mini Extrusion system (Avanti Polar Lipids, Alabaster, Alabama). All fluorescence data were recorded on a Fluorolog spectrometer (Horiba Jobin Yvon FL3-22).

### **Synthesis of ZpX series**

### **General procedure I: Boc- deprotection**

Boc- protected amino acid (0.1 mmol) was dissolved in a mixture of 95/2.5/2.5 TFA/<sup>i</sup>Pr<sub>3</sub>SiH/H<sub>2</sub>O (1 mL), and stirred at r.t. for 1 h. The solvents were then removed under vacuum, and the residue was directly used in general procedure II.

### **General procedure II: Fmoc- protection**

The crude free amino acid (0.1 mmol) from General procedure I or III was dissolved in 10% Na<sub>2</sub>CO<sub>3</sub> aq. (1 mL), followed by an addition of CH<sub>3</sub>CN (0.5 mL) and Fmoc-OSu (32 mg, 0.095 mmol). The resulting mixture was stirred vigorously at r.t. for 2.5 h, then poured into 1 N HCl (20 mL) and extracted with DCM (3 × 10 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (9:1 toluene/HOAc).

### **(S)-3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino) propanoic acid (1-2)**

Sodium (583 mg, 25.4 mmol) was dissolved in allyl alcohol (30 mL) in portions at 0 °C, followed by addition of (S)-2-(tert-butoxycarbonylamino)-3-(perfluorophenyl) propanoic acid (**1-1**, 1.0 g, 2.82 mmol). The resulting solution was sealed in a pressure bottle under N<sub>2</sub> protection and stirred at 90 °C for 6 h. Upon cooling to room temperature, the reaction mixture was poured into 1N HCl (60 mL), swirled thoroughly, and extracted with DCM (3 × 40 ml). The combined organic layers were washed with a 10:1 mixture of

brine and 1N HCl and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (2:1 Et<sub>2</sub>O/hexane first, to remove less-polar byproducts; followed by 2:1 Et<sub>2</sub>O/hexane with 2.5% HOAc). The product was obtained as a white crystal. Yield: 975 mg (88%). <sup>1</sup>H NMR (500 MHz, Acetone -d<sub>6</sub>) δ: 6.27 (d, *J* = 8.5 Hz, 1H), 6.07 (m, 1H), 5.42 (d, *J* = 17.5 Hz, 1H), 5.29 (d, *J* = 5.5 Hz, 1H), 4.75 (d, *J* = 6.0 Hz, 2H), 4.47, (m, 1H), 3.30 (m, 1H), 3.16 (m, 1H), 1.34 (s, 9H). <sup>13</sup>C NMR (125 MHz, Acetic acid-d<sub>4</sub>) rotamers δ: 175.9, 156.9, 147.8, 145.9, 143.3, 141.4, 137.0, 133.6, 120.0, 110.4, 81.4, 76.3, 53.6, 28.4, 26.2. <sup>19</sup>F NMR (470 MHz, CDCl<sub>3</sub>) δ: -146.7 (2F), -159.1 (2F). HRMS (ESI+): *m/z* calculated for C<sub>17</sub>H<sub>20</sub>F<sub>4</sub>NO<sub>5</sub> [M]<sup>+</sup>, 394.12776; found 394.12653.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)propanoic acid (1-3)**

(S)-3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoic acid (1-2) (975 mg, 2.48 mmol) was dissolved in a mixture of 95% trifluoroacetic acid, 2.5% water and 2.5% triethylsilane (30 mL), and stirred at r.t. for 1 h. The solvents were then removed under vacuum, and the residue was treated with a mixture of 10% Na<sub>2</sub>CO<sub>3</sub> aq (30 mL), water (15 mL) and acetonitrile (25 mL) to pH~10, followed by an addition of Fmoc-OSu (794 mg, 2.36 mmol). The resulting solution was stirred at r.t. for 2.5 h, then poured into 2N HCl (60 mL), and extracted with DCM (3 × 40 ml). The combined organic layers were washed with a 10:1 mixture of brine and 1N HCl and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with

silica gel column (2:1 Et<sub>2</sub>O/hexane first, to remove less-polar byproducts; then 9:1 toluene/HOAc). The desired product was obtained as a white solid. Yield: 1.15 g (90%). <sup>1</sup>H NMR (500 MHz, Acetone -d<sub>6</sub>) δ: 7.81 (d, *J* = 7.0 Hz, 2H), 7.67 (m, 2H), 7.37 (m, 2H), 7.29 (m, 2H), 6.94 (d, *J* = 9.0 Hz, 2H), 5.98 (m, 1H), 5.35 (d, *J* = 17.5 Hz, 1H), 5.20, (d, *J* = 10.0 Hz, 2H), 4.64, (m, 3H), 4.14-4.33, (m, 3H), 3.38 (m, 1H), 3.24 (m, 1H). <sup>13</sup>C NMR (125 MHz, Acetonitrile-d<sub>3</sub>) δ: 172.4, 157.0, 147.8, 145.9, 145.0, 143.3, 142.1, 141.3, 136.8, 133.7, 128.8, 128.2, 126.2, 121.0, 120.3, 118.4, 110.8, 76.4, 67.7, 53.9, 47.9, 26.0. <sup>19</sup>F NMR (470 MHz, Acetone-d<sub>6</sub>) δ: -144.0 (2F), -156.4 (2F). HRMS (ESI+): *m/z* calculated for C<sub>27</sub>H<sub>22</sub>F<sub>4</sub>NO<sub>5</sub> [M]<sup>+</sup>, 516.14341; found 516.14489.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-hydroxyphenyl)propanoic acid (1-4)**

Under N<sub>2</sub> protection, (S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-allyloxy)-2,3,5,6-tetrafluorophenyl)propanoic acid (**1-3**) (1.15 g, 2.25 mmol) and Pd(PPh<sub>3</sub>)<sub>4</sub> (130 mg, 0.11 mmol) were dissolved in DCM (30 mL) at r.t., followed by addition of phenylsilane (560 μL, 4.50 mmol). The resulting solution was stirred at r.t. under N<sub>2</sub> protection for 1.5 h. The reaction was quenched by 1N HCl (10 mL), and DCM was evaporated under vacuum. The residue was treated with 5% Na<sub>2</sub>CO<sub>3</sub> aq to pH~8-9, diluted with water to a final volume of 120 mL, and then extracted with Et<sub>2</sub>O (2 × 25 ml). The ether layers were discarded. The pH value of the aqueous layer was tuned to 0-1 with 2N HCl, and then extracted with EtOAc (3 × 50 ml). The combined EtOAc layers were dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was



isolated with silica gel column (7:1 toluene/HOAc) as a white solid. Yield: 975 mg (92%).  
 $^1\text{H}$  NMR (500 MHz, Acetone- $d_6$ )  $\delta$ : 7.85 (d,  $J = 9.5$  Hz, 2H), 7.67 (d,  $J = 9.0$  Hz, 2H), 7.41 (m, 2H), 7.31 (m, 2H), 6.91 (d,  $J = 11.5$  Hz, 1H), 4.56 (m, 1H), 4.16-4.33 (m, 3H), 3.32 (m, 1H), 3.20 (m, 1H).  $^{19}\text{F}$  NMR (470 MHz, Acetone- $d_6$ )  $\delta$ : -146.6 (2F), -164.3 (2F).

**(S)-methyl 3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoate (1-5)**

(S)-3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoic acid (**1-2**) (975 mg, 2.48 mmol) was dissolved in 10 mL DMF, followed by an addition of  $\text{K}_2\text{CO}_3$  (685 mg, 4.96 mmol) and  $\text{CH}_3\text{I}$  (310  $\mu\text{L}$ , 4.96 mmol). The resulting suspension was stirred at r.t. for 16 h. Then the reaction mixture was poured into 100 mL water, extracted with DCM (3  $\times$  50 mL). The combined organic layers were washed with brine and dried over  $\text{Na}_2\text{SO}_4$ . After the solvent was evaporated under vacuum, the product was isolated with silica gel column (4:1 hexanes/EtOAc). Yield: 859 mg (85%).

**(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-hydroxyphenyl)propanoate (1-6)**

Under  $\text{N}_2$  protection, (S)-methyl 3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoate (**1-5**) (859 mg, 2.11 mmol) and  $\text{Pd}(\text{PPh}_3)_4$  (130 mg, 0.11 mmol) were dissolved in DCM (30 mL) at r.t., followed by addition of phenylsilane (560  $\mu\text{L}$ , 4.50 mmol). The resulting solution was stirred at r.t. under  $\text{N}_2$  protection for 1.5 h. The reaction was quenched by 1N HCl (10 mL), and DCM was evaporated under

vacuum. The residue was treated with 5% Na<sub>2</sub>CO<sub>3</sub> aq to pH~8-9, diluted with water to a final volume of 120 mL, and then extracted with Et<sub>2</sub>O (2 × 25 ml). The ether layers were discarded. The pH value of the aqueous layer was tuned to 0-1 with 2N HCl, and then extracted with EtOAc (3 × 50 ml). The combined EtOAc layers were dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (1:1 hexanes/EtOAc) as a white solid. Yield: 774 mg (92%). <sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ: 6.91 (d, *J* = 8.8 Hz, 1H), 4.27 (m, 1H), 3.63 (s, 3H), 3.07 (m, 1H), 2.92 (m, 1H), 1.27 (s, 9H). <sup>13</sup>C NMR (100 MHz, Methanol-*d*<sub>4</sub>) δ: 171.6, 156.1, 146.9, 144.5, 139.2, 136.7, 135.5, 104.1, 79.3, 52.9, 51.5, 27.1, 24.5. <sup>19</sup>F NMR (470 MHz, Acetone-*d*<sub>6</sub>) δ: -146.9 (2F), -164.4 (2F). HRMS (ESI+): *m/z* calculated for C<sub>15</sub>H<sub>18</sub>F<sub>4</sub>NO<sub>5</sub> [M]<sup>+</sup>, 368.11211; found 368.11171.

**(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-(trifluoromethylsulfonyloxy)phenyl) propanoate (1-7)**

Under N<sub>2</sub> protection, (S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-hydroxyphenyl)propanoate (**1-6**) (135 mg, 0.37 mmol) and Et<sub>3</sub>N (0.205 mL, 1.47 mmol) were dissolved in anhydrous DCM (2 mL). The mixture was chilled to -78 °C and Tf<sub>2</sub>O (0.14 mL, 0.81 mmol) was added dropwise. The reaction was stirred at -78 °C for 10 min and then at room temperature for another 30 min. The reaction was quenched by saturated NaHCO<sub>3</sub> aq. (30 mL) and diluted with DCM (30 mL). The organic layer was separated, sequentially washed with 1 N HCl and brine, and dried over Na<sub>2</sub>SO<sub>4</sub>. After DCM was removed under vacuum, the product was isolated with silica gel column (5:1

hexanes/EtOAc) as colorless crystal. Yield: 146 mg (80%). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ: 4.59 (m, 1H), 3.75 (s, 3H), 3.33 (m, 1H), 3.08 (m, 1H), 1.33 (s, 9H). <sup>13</sup>C NMR (125 MHz, Chloroform-*d*) δ: 171.2, 155.1, 146.8, 144.8, 141.7, 139.7, 126.4, 120.0, 116.4, 80.5, 52.9, 52.6, 28.1, 26.9. <sup>19</sup>F NMR (470 MHz, Chloroform-*d*) δ: -73.0 (3F), -140.7 (2F), -152.0 (2F). HRMS (ESI+): *m/z* calculated for C<sub>16</sub>H<sub>20</sub>F<sub>7</sub>N<sub>2</sub>O<sub>7</sub>S [M + NH<sub>4</sub>]<sup>+</sup>, 517.08794; found 517.08934.

**(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-3',5'-dimethylbiphenyl-4-yl) propanoate (1-8b)**

(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-(trifluoromethylsulfonyloxy)phenyl)propanoate (**1-7**) (50 mg, 0.10 mmol), 3,5-dimethylphenylboronic acid (16.5 mg, 0.11 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (11.6 mg, 0.01 mmol) and K<sub>2</sub>CO<sub>3</sub> (27.6 mg, 0.20 mmol) were dissolved in a mixture of 1,2-dimethoxyethane/water (6:1, 1 mL totally) and sealed in a pressure vessel under N<sub>2</sub> protection. The reaction mixture was stirred at 70 °C for 12 h. Upon cooling to r.t., the reaction mixture was poured into 1 N HCl (15 mL) and extracted with DCM (3 × 15 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (5:1 hexanes/EtOAc) as a white solid. Yield: 44 mg (95%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ: 7.08 (s, 1H), 7.04 (s, 2H), 5.18 (d, *J* = 8.5 Hz, 1H), 4.66 (m, 1H), 3.80 (s, 3H), 3.39 (m, 1H), 3.14 (m, 1H), 2.37 (s, 6H), 1.41 (s, 9H). <sup>13</sup>C NMR (100 MHz, Chloroform-*d*) δ: 171.7, 155.1, 146.9, 144.9, 144.6, 142.6, 139.6, 138.4, 131.0, 128.0, 120.3, 113.8, 80.4,

52.9, 28.4, 26.6, 21.4.  $^{19}\text{F}$  NMR (470 MHz, Chloroform-*d*)  $\delta$ : -143.8 (2F), -144.1 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{23}\text{H}_{26}\text{F}_4\text{NO}_4$   $[\text{M}]^+$ , 456.17980; found 456.18015.

**(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-vinylphenyl)propanoate (1-8a)**

(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-(trifluoromethylsulfonyloxy)phenyl)propanoate (**1-7**) (50 mg, 0.10 mmol), vinylboronic acid dibutyl ester (26.5  $\mu\text{L}$ , 0.12 mmol),  $\text{Pd}(\text{PPh}_3)_4$  (11.6 mg, 0.01 mmol) and  $\text{K}_2\text{CO}_3$  (27.6 mg, 0.20 mmol) were dissolved in a mixture of 1,2-dimethoxyethane/water (6:1, 1 mL totally) and sealed in a pressure vessel under  $\text{N}_2$  protection. The reaction mixture was stirred at 70  $^\circ\text{C}$  for 12 h. Upon cooling to r.t., the reaction mixture was poured into 1 N HCl (15 mL) and extracted with DCM (3  $\times$  15 mL). The combined organic layers were washed with brine and dried over  $\text{Na}_2\text{SO}_4$ . After the solvent was evaporated under vacuum, the product was isolated with silica gel column (5:1 hexanes/EtOAc) as a white solid. Yield: 28 mg (74%).  $^1\text{H}$  NMR (500 MHz, Chloroform-*d*)  $\delta$ : 6.67 (dd,  $J_1 = 18.0$  Hz,  $J_2 = 12.0$  Hz, 1H), 6.08 (d,  $J = 18.0$  Hz, 1H), 5.69 (d,  $J = 12.0$  Hz, 1H), 5.12 (d,  $J = 8.0$  Hz, 1H), 4.61 (m, 1H), 3.78 (s, 3H), 3.34 (m, 1H), 3.10 (m, 1H), 1.40 (s, 9H).  $^{13}\text{C}$  NMR (100 MHz, Chloroform-*d*)  $\delta$ : 171.4, 154.9, 146.6, 145.6, 144.1, 143.0, 123.5, 122.4, 115.9, 113.5, 80.2, 52.7, 29.7, 28.1, 26.4.  $^{19}\text{F}$  NMR (470 MHz, Chloroform-*d*)  $\delta$ : -144.5 (4F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{17}\text{H}_{20}\text{F}_4\text{NO}_4$   $[\text{M}]^+$ , 378.13285; found 378.13354.

**(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-  
((trimethylsilyl)ethynyl)phenyl)propanoate (1-8c)**

(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-(trifluoromethylsulfonyloxy)phenyl)propanoate (**1-7**) (50 mg, 0.10 mmol), triethylamine (60  $\mu$ L), Pd(PPh<sub>3</sub>)<sub>4</sub> (11.6 mg, 0.01 mmol), cuprous iodide (3.8 mg, 0.02 mmol) and lithium bromide (8.7 mg, 0.10 mmol) were dissolved in anhydrous THF (1 mL totally), followed by an addition of trimethylsilylacetylene (42  $\mu$ L, 0.30 mmol). The resulting mixture was sealed in a pressure vessel under N<sub>2</sub> protection. The reaction mixture was stirred at 80 °C for 12 h. Upon cooling to r.t., the reaction mixture was poured into 1 N HCl (15 mL) and extracted with DCM (3  $\times$  15 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (5:1 hexanes/EtOAc) as a white solid. Yield: 31 mg (70%). <sup>1</sup>H NMR (400 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : 6.43 (d, *J* = 8.4 Hz, 1H), 4.50 (m, 1H), 3.72 (s, 3H), 3.33 (m, 1H), 3.24 (m, 1H), 1.34 (s, 9H) 0.29 (s, 9H). <sup>13</sup>C NMR (100 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : 172.2, 156.6, 149.4, 148.0, 146.8, 145.5, 119.5, 109.8, 103.9, 89.8, 80.2, 54.0, 53.2, 28.8, 27.3, 0.1. <sup>19</sup>F NMR (470 MHz, Chloroform-*d*)  $\delta$ : -140.1 (2F), -144.3 (2F). HRMS (ESI+): *m/z* calculated for C<sub>20</sub>H<sub>26</sub>F<sub>4</sub>NO<sub>4</sub>Si [M]<sup>+</sup>, 448.15672; found 448.15521.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-3',5'-  
dimethylbiphenyl-4-yl)propanoic acid (1-9b)**

(S)-methyl 2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-3',5'-dimethylbiphenyl-4-yl) propanoate (**1-8b**) (43 mg, 0.096 mmol) and LiOH·H<sub>2</sub>O (16 mg, 0.38 mmol) were mixed with H<sub>2</sub>O (2 mL) and THF (1 mL). The resulting suspension was stirred at r.t. for 16 h. Then the pH was tuned to 4-5 1 N HCl, and all the solvents were removed on rot. vap. The residue was directly treated sequentially as General Procedure I and II. Yield: 46 mg (85%). <sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ: 7.78 (d, *J* = 7.6 Hz, 2H), 7.58 (m, 2H), 7.34 (m, 2H), 7.24 (m, 2H), 7.02 (s, 1H), 6.81 (s, 2H), 4.56 (m, 1H), 4.40 (m, 1H), 4.06-3.97 (m, 2H), 3.40 (m, 1H), 3.24 (m, 1H), 2.24 (s, 6H). <sup>13</sup>C NMR (100 MHz, Methanol-*d*<sub>4</sub>) δ: 176.0, 158.4, 148.4, 146.0, 145.6, 145.0, 143.8, 142.7, 142.6, 139.4, 131.6, 128.9, 128.6, 128.4, 126.6, 126.2, 121.0, 117.0, 68.5, 55.4, 48.4, 27.5, 21.4. <sup>19</sup>F NMR (376 MHz, Methanol-*d*<sub>4</sub>) δ: -145.4 (2F), -146.4 (2F). HRMS (ESI+): *m/z* calculated for C<sub>32</sub>H<sub>26</sub>F<sub>4</sub>NO<sub>4</sub> [M]<sup>+</sup>, 564.17980; found 564.18131.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-ethyl-2,3,5,6-tetrafluorophenyl)propanoic acid (1-9a)**

The procedure is the same as for **1-9b**. Yield: 72%. <sup>1</sup>H NMR (400 MHz, Acetonitrile-*d*<sub>3</sub>) δ: 7.82 (d, *J* = 7.6 Hz, 2H), 7.60 (d, *J* = 7.6 Hz, 2H), 7.41 (m, 2H), 7.33 (m, 2H), 6.17 (d, *J* = 7.2 Hz, 1H), 4.48 (m, 1H), 4.30 (m, 1H), 4.25-4.10 (m, 2H), 3.28 (m, 1H), 3.16 (m, 1H), 2.69 (q, *J* = 7.2 Hz, 2H), 1.15 (t, *J* = 7.2 Hz, 3H). <sup>13</sup>C NMR (100 MHz, Acetonitrile-*d*<sub>3</sub>) δ: 171.9, 156.1, 146.9, 146.1, 144.3, 144.2, 143.9, 141.5, 128.2, 127.6, 125.8, 125.7, 121.5, 120.2, 113.9, 66.8, 53.1, 47.4, 25.7, 16.3, 13.9. <sup>19</sup>F NMR (376 MHz, Methanol-*d*<sub>4</sub>)

$\delta$ : -145.5 (2F), -148.0 (2F). HRMS (ESI+):  $m/z$  calculated for  $C_{26}H_{22}F_4NO_4$   $[M]^+$ , 488.14850; found 488.14941.

**(S)-tert-butyl 3-(4-(allyloxy)-2,3,5,6-tetrafluorophenyl)-2-((tert-butoxycarbonyl)amino)propanoate (1-10)**

The procedure is the same as for **1-20**. Yield: 89%.  $^1H$  NMR (500 MHz, Chloroform-*d*)  $\delta$ : 5.94 (m, 1H), 5.27 (m, 1H), 5.18 (m, 1H), 5.16 (m, 1H), 4.60 (d,  $J = 5.5$  Hz, 2H), 4.35 (m, 1H), 3.08 (m, 1H), 2.94 (m, 1H), 1.35 (s, 9H), 1.30 (s, 9H).

**(S)-tert-butyl 2-((tert-butoxycarbonyl)amino)-3-(2,3,5,6-tetrafluoro-4-hydroxyphenyl)propanoate (1-11)**

The procedure is the same as for **1-6**. Yield: 86%.  $^1H$  NMR (500 MHz, Chloroform-*d*)  $\delta$ : 5.23 (d,  $J = 8.0$  Hz, 1H), 4.41 (m, 1H), 3.15 (m, 1H), 2.99 (m, 1H), 1.42 (s, 9H), 1.37 (s, 9H).

**(S)-tert-butyl 2-((tert-butoxycarbonyl)amino)-3-(2,3,5,6-tetrafluoro-4-(((trifluoromethyl)sulfonyl)oxy)phenyl)propanoate (1-12)**

The procedure is the same as for **1-7**. Yield: 80%.  $^1H$  NMR (500 MHz, Benzene-*d*<sub>6</sub>)  $\delta$ : 5.22 (d,  $J = 7.5$  Hz, 1H), 4.42 (m, 1H), 2.89 (m, 1H), 2.77 (m, 1H), 1.28 (s, 9H), 1.18 (s, 9H).

**(S)-2-(tert-butoxycarbonylamino)-3-(2,3,5,6-tetrafluoro-4-(4-methoxybenzylthio)phenyl)propanoic acid (1-16)**

(S)-2-(tert-butoxycarbonylamino)-3-(perfluorophenyl) propanoic acid (**1-1**) (200 mg, 0.56 mmol) and (4-methoxyphenyl)methanethiol (125  $\mu$ L, 0.90 mmol) were dissolved in isopropanol (2 mL), followed by an addition of 1 N NaOH aq. (6 mL). The resulting solution was sealed in a pressure bottle under N<sub>2</sub> protection and stirred at 70 °C for 24 h. Upon cooling to room temperature, the reaction mixture was poured into a mixture of 1 N HCl (50 mL) and 1 M CuSO<sub>4</sub> aq. (10 mL), swirled thoroughly, and extracted with DCM (3  $\times$  30 mL). The combined organic layers were washed with a 10:1 mixture of brine and 1N HCl and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (3:1 hexanes/EtOAc with 3% HOAc). Yield: 243 mg (88%). <sup>1</sup>H NMR (500 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : 7.18 (d, *J* = 8.5 Hz, 2H), 6.83 (d, *J* = 8.5 Hz, 2H), 6.25 (d, *J* = 9.0 Hz, 1H), 4.48 (m, 1H), 4.13 (s, 2H), 3.76 (s, 3H), 3.34 (m, 1H), 3.21 (m, 1H), 1.34 (s, 9H). <sup>13</sup>C NMR (125 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : 172.8, 160.6, 156.5, 149.0, 147.8, 147.1, 145.8, 131.3, 130.1, 118.7, 115.2, 80.0, 55.9, 53.8, 39.5, 28.8, 27.2. <sup>19</sup>F NMR (470 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : -131.2 (2F), -138.7 (2F). HRMS (ESI+): *m/z* calculated for C<sub>22</sub>H<sub>24</sub>F<sub>4</sub>NO<sub>5</sub>S [M]<sup>+</sup>, 490.13113; found 490.13237.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-(4-methoxybenzylthio)phenyl)propanoic acid (1-17)**

Following General Procedure I and II. The product was purified with silica gel column (10:1 toluene/HOAc). Yield: 80%, over 2 steps. <sup>1</sup>H NMR (500 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ :



7.83 (d,  $J = 7.5$  Hz, 2H), 7.66 (d,  $J = 7.5$  Hz, 2H), 7.40 (m, 2H), 7.31 (m, 2H), 7.09 (d,  $J = 8.5$  Hz, 2H), 6.92 (d,  $J = 9.0$  Hz, 1H), 6.78 (d,  $J = 8.5$  Hz, 2H), 4.61 (m, 1H), 4.17-4.33 (m, 3H), 4.05 (s, 2H), 3.71 (s, 3H), 3.40 (m, 1H), 3.30 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Acetone- $d_6$ )  $\delta$ : 172.0, 160.2, 156.8, 149.0, 147.4, 146.6, 145.0, 144.9, 142.1, 130.9, 129.6, 128.6, 128.0, 126.1, 120.8, 118.1, 114.7, 113.2, 67.5, 55.5, 53.8, 47.9, 39.1, 26.5.  $^{19}\text{F}$  NMR (470 MHz, Chloroform- $d$ )  $\delta$ : -130.8 (2F), -138.5 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{32}\text{H}_{26}\text{F}_4\text{NO}_5\text{S} [\text{M}]^+$ , 612.14678; found 612.14419.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-mercaptophenyl)propanoic acid (1-18)**

(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-(4-methoxybenzylthio)phenyl)propanoic acid (**1-17**) (312 mg, 0.51 mmol) was dissolved in a mixture of methanol (8 mL) and acetonitrile (8 mL).  $\text{Hg}(\text{OAc})_2$  (195 mg, 0.61 mmol) was added. The resulting solution was stirred at 60 °C for 6 h in a sealed vial. Upon cooling to r.t., the reaction was quenched by bubbling  $\text{H}_2\text{S}$  into the reaction solution over 20 min. EtOAc (20 mL) was then added, and  $\text{HgS}$  was removed by vacuum filtration. 1N HCl (20 mL) was added to the filtrate, swirled thoroughly, and the organic layer was separated. The aqueous layer was extracted with EtOAc (2  $\times$  30 mL), and the combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . After the solvent was evaporated under vacuum, the product was isolated with silica gel column (1:1  $\text{Et}_2\text{O}$ /hexane first, to remove less-polar byproducts; followed by 9:1 toluene/HOAc). Yield: 215 mg (86%).  $^1\text{H}$  NMR (500 MHz, Acetone  $-d_6$ )  $\delta$ : 7.81 (d,  $J = 7.5$  Hz, 2H), 7.64 (d,  $J = 8.0$  Hz, 2H), 7.39 (m, 2H),

7.29 (m, 2H), 6.95 (d,  $J = 8.5$  Hz, 1H), 4.63 (m, 1H), 4.14-4.33 (m, 3H), 3.39 (m, 1H), 3.28 (m, 1H).  $^{13}\text{C}$  NMR (125 MHz, Acetone  $-d_6$ )  $\delta$ : 172.3, 156.9, 147.5, 145.3, 145.1, 144.9, 143.3, 142.0, 128.5, 127.9, 126.1, 114.1, 111.5, 67.5, 54.0, 47.9, 26.3.  $^{19}\text{F}$  NMR (470 MHz, Acetone- $d_6$ )  $\delta$ : -139.7 (2F), -143.6 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{24}\text{H}_{18}\text{F}_4\text{NO}_4\text{S} [\text{M}]^+$ , 492.08927; found 492.08680.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-(tritylthio)phenyl)propanoic acid (1-19)**

(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-mercaptophenyl)propanoic acid (**1-18**) (182 mg, 0.37 mmol) and trityl chloride (228 mg, 0.82 mmol) were dissolved in DCM (15 mL), followed by an addition of DIPEA (135  $\mu\text{L}$ , 0.82 mmol) slowly. The result mixture was stirred at r.t. for 3 h, then poured into 1%  $\text{H}_3\text{PO}_4$  aq. (20 mL) and swirled thoroughly. The organic layer was separated, and the aqueous layer was extracted with DCM (20 mL). The combined organic layers were washed with a 10:1 mixture of brine and 1%  $\text{H}_3\text{PO}_4$ , and dried over  $\text{Na}_2\text{SO}_4$ . After the solvent was evaporated under vacuum, the product was isolated with silica gel column (2:1 hexanes/EtOAc with 2% HOAc). Yield: 230 mg (85%).  $^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$ : 7.71 (m, 2H), 7.12-7.55 (m, 21H), 5.17 (d,  $J = 10.5$  Hz, 1H), 4.59 (m, 1H), 4.40 (m, 2H), 4.17 (m, 1H), 3.27 (m, 1H), 3.07 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Chloroform- $d$ )  $\delta$ : 174.9, 155.7, 149.3, 146.7, 145.9, 143.7, 143.5, 143.4, 141.2, 129.7, 127.8, 127.3, 127.0, 124.9, 120.0, 116.3, 112.8, 73.9, 67.3, 52.7, 47.0, 25.8.  $^{19}\text{F}$  NMR

(470 MHz, Acetone-*d*<sub>6</sub>)  $\delta$ : -130.2 (2F), -143.6 (2F). HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>43</sub>H<sub>31</sub>F<sub>4</sub>NO<sub>4</sub>SNa [M + Na]<sup>+</sup>, 756.1808; found 756.1824.

**(S)-tert-butyl 2-(tert-butoxycarbonylamino)-3-(perfluorophenyl)propanoate (1-20)**

t-Butyl trichloroacetimidate (4.04 mL, 22.5 mmol) was slowly added to a solution of (S)-2-(tert-butoxycarbonylamino)-3-(perfluorophenyl) propanoic acid (**1-1**) (4.0 g, 11.3 mmol) in 120 mL DCM/cyclohexane (1:2). The resulting solution was stirred at r.t. for 16 h. White crystal precipitated out. BF<sub>3</sub>·H<sub>2</sub>O (0.22 mL) was then added, and the reaction was stirred at r.t. for another 1 h. The reaction was quenched with saturated NaHCO<sub>3</sub> aq. (100 mL). The organic layer was separated, and the aqueous layer was extracted with DCM (100 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (9:1 hexanes/EtOAc) as white crystal. Yield: 4.4 g (95%). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$ : 5.12 (d, *J* = 8.8 Hz, 1H), 4.46 (m, 1H), 3.21 (m, 1H), 3.05 (m, 1H), 1.45 (s, 9H), 1.40 (s, 9H). <sup>13</sup>C NMR (100 MHz, Chloroform-*d*)  $\delta$ : 169.9, 155.0, 146.9, 144.5, 140.9, 138.7, 136.2, 110.9, 82.9, 79.9, 53.3, 28.1, 27.8, 26.3. <sup>19</sup>F NMR (400 MHz, Chloroform-*d*)  $\delta$ : -142.3 (2F), -156.5 (1F), -163.1 (2F). HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>18</sub>H<sub>23</sub>F<sub>5</sub>NO<sub>4</sub> [M]<sup>+</sup>, 412.15472; found 412.15470.

**(S)-tert-butyl 3-(4-azido-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoate (1-21)**

A mixture of (S)-tert-butyl 2-(tert-butoxycarbonylamino)-3-(perfluorophenyl)propanoate (**1-20**) (400 mg, 0.97 mmol), tetrabutylammonium azide (28 mg, 0.097 mmol) and NaN<sub>3</sub> (126 mg, 1.94 mmol) in anhydrous DMF (4 mL) was sealed in a pressure vessel and stirred at 80 °C for 16 h. Upon cooling to room temperature, the reaction mixture was poured into saturated NaHCO<sub>3</sub> aq. (40 mL), swirled thoroughly, and extracted with DCM (40 mL) and chloroform (2 × 40 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. After the solvent was evaporated under vacuum, the product was isolated with silica gel column (6:1 hexanes/EtOAc) as white solid. Yield: 342 mg (81%). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ: 5.11 (d, *J* = 7.6 Hz, 1H), 4.44 (m, 1H), 3.22 (m, 1H), 3.03 (m, 1H), 1.45 (s, 9H), 1.40 (s, 9H). <sup>13</sup>C NMR (100 MHz, Chloroform-*d*) δ: 170.0, 155.1, 146.9, 144.5, 141.8, 139.3, 119.5, 111.5, 83.0, 80.2, 53.5, 28.4, 28.1, 26.6. <sup>19</sup>F NMR (400 MHz, Chloroform-*d*) δ: -142.6 (2F), -153.0 (2F). HRMS (ESI+): *m/z* calculated for C<sub>18</sub>H<sub>23</sub>F<sub>4</sub>N<sub>4</sub>O<sub>4</sub> [M]<sup>+</sup>, 435.16554; found 435.16337.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-azido-2,3,5,6-tetrafluorophenyl)propanoic acid (1-22)**

(S)-tert-butyl 3-(4-azido-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoate (**1-21**) was treated following General Procedure I and II. Yield: 85%. Known compound. <sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ: 7.77 (d, *J* = 7.2 Hz, 2H), 7.60 (m, 2H), 7.40 (m, 2H), 7.30 (m, 2H), 4.53 (m, 1H), 4.12-4.34, (m, 3H), 3.33 (m, 1H), 3.19 (m, 1H).

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-amino-2,3,5,6-tetrafluorophenyl)propanoic acid (1-23)**

1,3-propanedithiol (138  $\mu\text{L}$ , 1.38 mmol) was added into a solution of (S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-azido-2,3,5,6-tetrafluorophenyl)propanoic acid (**1-22**) (345 mg, 0.69 mmol) and DIPEA (240  $\mu\text{L}$ , 1.38 mmol) in 5 mL methanol/acetonitrile (1:1). The resulting solution was stirred at r.t. for 5 min, then poured into 1N HCl (20 mL) and extracted with DCM (5  $\times$  20 mL). The combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . DCM was evaporated under vacuum, and the product was isolated with silica gel column (3:2 hexanes/EtOAc, 2% HOAc) as white crystal. Yield: 314 mg (96%).  $^1\text{H}$  NMR (400 MHz, Methanol- $d_4$ )  $\delta$ : 7.77 (d,  $J = 7.6$  Hz, 2H), 7.60 (d,  $J = 7.6$  Hz, 2H), 7.36 (m, 2H), 7.28 (m, 2H), 4.41 (m, 1H), 4.12-4.32 (m, 3H), 3.20 (m, 1H), 3.05 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Methanol- $d_4$ )  $\delta$ : 172.9, 156.9, 146.7, 144.4, 143.7, 141.0, 137.5, 134.9, 127.3, 126.7, 124.8, 119.4, 100.6, 66.8, 53.6, 24.4.  $^{19}\text{F}$  NMR (400 MHz, Acetone- $d_6$ )  $\delta$ : -148.1 (2F), -164.4 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{24}\text{H}_{19}\text{F}_4\text{N}_2\text{O}_4$   $[\text{M}]^+$ , 475.12809; found 475.12717.

**(S)-3-(4-azido-2,3,5,6-tetrafluorophenyl)-2-(2,2,2-trifluoroacetamido)propanoic acid (1-24)**

(S)-tert-butyl 3-(4-azido-2,3,5,6-tetrafluorophenyl)-2-(tert-butoxycarbonylamino)propanoate (**1-21**) (300 mg, 0.69 mmol) was dissolved in TFA (2.5 mL) at 0  $^\circ\text{C}$ . After stirred 0.5 h at 0  $^\circ\text{C}$ , trifluoroacetic anhydride (125  $\mu\text{L}$ , 0.90 mmol) was added dropwise. The resulting solution was stirred at 0  $^\circ\text{C}$  for 2 h and at r.t. for

another 16 h. TFA was removed under vacuum, and the residue was directly purified with silica gel column (9:1 toluene/HOAc). Yield: 258 mg (100%).  $^1\text{H}$  NMR (400 MHz, Methanol- $d_4$ )  $\delta$ : 4.68 (m, 1H), 3.42 (m, 1H), 3.20 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Methanol- $d_4$ )  $\delta$ : 170.5, 157.3, 146.7, 144.3, 141.6, 139.1, 118.9, 114.4, 111.1, 51.7, 24.1.  $^{19}\text{F}$  NMR (400 MHz, Acetone- $d_6$ )  $\delta$ : -77.0 (3F), -144.9 (2F), -154.8 (2F). HRMS (ES-):  $m/z$  calculated for  $\text{C}_{11}\text{H}_4\text{F}_7\text{N}_4\text{O}_3$   $[\text{M}]^-$ , 373.0172; found 373.0163.

**(S)-3-(4-amino-2,3,5,6-tetrafluorophenyl)-2-(2,2,2-trifluoroacetamido)propanoic acid (1-25)**

1,3-propanedithiol (138  $\mu\text{L}$ , 1.38 mmol) was added into a solution of (S)-3-(4-azido-2,3,5,6-tetrafluorophenyl)-2-(2,2,2-trifluoroacetamido)propanoic acid (**1-24**) (258 mg, 0.69 mmol) and DIPEA (240  $\mu\text{L}$ , 1.38 mmol) in 5 mL methanol/acetonitrile (1:1). The resulting solution was stirred at r.t. for 5 min, then poured into 1N HCl (20 mL) and extracted with DCM ( $5 \times 20$  mL). The combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . DCM was evaporated under vacuum, and the product was isolated with silica gel column (3:2 hexanes/EtOAc, 2% HOAc) as white crystal. Yield: 228 mg (95%).  $^1\text{H}$  NMR (400 MHz, Acetone- $d_6$ )  $\delta$ : 8.70 (d,  $J = 8.4$  Hz, 1H), 4.77 (m, 1H), 3.38 (m, 1H), 3.22 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Acetone- $d_6$ )  $\delta$ : 169.6, 156.5, 146.8, 144.5, 137.3, 134.9, 127.1, 114.5, 100.0, 52.0, 23.7.  $^{19}\text{F}$  NMR (400 MHz, Acetone- $d_6$ )  $\delta$ : -77.0 (3F), -148.4 (2F), -164.4 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{11}\text{H}_8\text{F}_7\text{N}_2\text{O}_3$   $[\text{M}]^+$ , 349.04231; found 349.04386.

### General procedure III: CF<sub>3</sub>CO- deprotection

CF<sub>3</sub>CO- protected amino acid (0.1 mmol) was suspended in 1.2 N HCl (3 mL) and refluxed for 16 h. HCl was removed on rot. vap. And the residue was directly used in general procedure II.

### (S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-iodophenyl)propanoic acid (1-27c)

*t*-Butyl nitrite (57  $\mu$ L, 0.43 mmol) was added dropwise into a mixture of (S)-3-(4-amino-2,3,5,6-tetrafluorophenyl)-2-(2,2,2-trifluoroacetamido)propanoic acid (**1-25**) (100 mg, 0.287 mmol) and iodine (146 mg, 0.576 mmol) in acetonitrile (5.8 mL) at 0 °C. The resulting solution was stirred at 0 °C for 1.5 h. The solvent was removed on rot. vap., and the residue was treated with 2 M NaHSO<sub>3</sub> (30 mL) then extracted with DCM (3  $\times$  30 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>. DCM was evaporated under vacuum, and the crude product was isolated with silica gel column (9:1 toluene/HOAc) as a mixture of ZpI/ZpH (20:1). The ZpI/ZpH mixture was treated sequentially as General Procedure III and II, and finally purified with RP-HPLC (Waters Prep LC, Jupiter 10u C18 300A Column). Alluent condition: 70%-100% B over 20 min (Buffer A: 95% water, 5% acetonitrile and 0.1% TFA; Buffer B: 5% water, 95% acetonitrile and 0.1% TFA). Yield: 114 mg (68%, over 3 steps). <sup>1</sup>H NMR (400 MHz, Acetonitrile-*d*<sub>3</sub>)  $\delta$ : 7.79 (d, *J* = 7.8 Hz, 2H), 7.58 (d, *J* = 7.8 Hz, 2H), 7.40 (m, 2H), 7.30 (m, 2H), 6.18 (d, *J* = 7.8 Hz, 1H), 4.49 (m, 1H), 4.10-4.28, (m, 3H), 3.29 (m, 1H), 3.19 (m, 1H). <sup>13</sup>C NMR (100 MHz, Acetonitrile-*d*<sub>3</sub>)  $\delta$ : 171.5, 156.2, 148.3, 146.4, 145.9, 144.2, 141.3, 128.0, 127.5, 127.4,

125.4, 120.2, 71.3, 66.9, 52.9, 47.1, 25.9.  $^{19}\text{F}$  NMR (400 MHz, Chloroform-*d*)  $\delta$ : -120.7 (2F), -140.7 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{24}\text{H}_{17}\text{F}_4\text{INO}_4$   $[\text{M}]^+$ , 586.01384; found 586.01431.

#### **General procedure IV: Sandmeyer reaction**

*t*-Butyl nitrite (57  $\mu\text{L}$ , 0.43 mmol) was added dropwise into a mixture of (S)-3-(4-amino-2,3,5,6-tetrafluorophenyl)-2-(2,2,2-trifluoroacetamido)propanoic acid (**1-25**) (100 mg, 0.287 mmol),  $\text{CuX}$  ( $\text{X} = \text{Cl}, \text{Br}$ ) (0.576 mmol) and  $\text{CuX}_2$  (0.576 mmol) in acetonitrile (5.8 mL) at 0  $^\circ\text{C}$ . The resulting solution was stirred at 0  $^\circ\text{C}$  for 1.5 h. The solvent was removed on rot. vap., and the residue was treated with 1 M HCl (30 mL) then extracted with DCM (3  $\times$  30 mL). The combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . DCM was evaporated under vacuum, and the crude product was isolated with silica gel column (9:1 toluene/HOAc) as a mixture of  $\text{ZpX}/\text{ZpH}$ , which was directly used in General procedure III.

#### **(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-bromo-2,3,5,6-tetrafluorophenyl)propanoic acid (1-27b)**

Following General Procedure IV, III and II. The final product was purified with RP-HPLC (Waters Prep LC, Jupiter 10u C18 300A Column). Alluent condition: 70%-100% B over 20 min (Buffer A: 95% water, 5% acetonitrile and 0.1% TFA; Buffer B: 5% water, 95% acetonitrile and 0.1% TFA). Yield: 71%, over 3 steps.  $^1\text{H}$  NMR (500 MHz, Acetonitrile- $d_3$ )  $\delta$ : 7.81 (d,  $J = 7.5$  Hz, 2H), 7.60 (d,  $J = 8.0$  Hz, 2H), 7.40 (m, 2H), 7.32



(m, 2H), 6.13 (d,  $J = 8.0$  Hz, 1H), 4.64 (m, 1H), 4.16-4.30, (m, 2H), 4.16 (m, 1H), 3.30 (m, 1H), 3.20 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Methanol- $d_4$ )  $\delta$ : 172.0, 156.9, 146.7, 145.8, 144.2, 143.7, 143.3, 141.0, 127.3, 126.5, 124.7, 119.3, 116.1, 97.8, 66.8, 52.8, 46.8, 25.3.  $^{19}\text{F}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$ : -133.5 (2F), -139.9 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{24}\text{H}_{17}\text{BrF}_4\text{NO}_4$   $[\text{M}]^+$ , 538.02771; found 538.02805.

**(S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-chloro-2,3,5,6-tetrafluorophenyl)propanoic acid (1-27a)**

Following General Procedure IV, III and II. The final product was purified with RP-HPLC (Waters Prep LC, Jupiter 10u C18 300A Column). Alluent condition: 70%-100% B over 20 min (Buffer A: 95% water, 5% acetonitrile and 0.1% TFA; Buffer B: 5% water, 95% acetonitrile and 0.1% TFA). Yield: 62%, over 3 steps.  $^1\text{H}$  NMR (500 MHz, Acetonitrile- $d_3$ )  $\delta$ : 7.82 (d,  $J = 7.5$  Hz, 2H), 7.60 (d,  $J = 8.0$  Hz, 2H), 7.40 (m, 2H), 7.32 (m, 2H), 6.13 (d,  $J = 8.0$  Hz, 1H), 4.64 (m, 1H), 4.14-4.30, (m, 3H), 3.30 (m, 1H), 3.20 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz, Acetone- $d_6$ )  $\delta$ : 171.2, 155.9, 147.0, 145.0, 144.5, 144.0, 142.6, 141.1, 127.6, 127.1, 125.1, 119.9, 115.9, 110.0, 66.5, 53.0, 47.0, 25.5.  $^{19}\text{F}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$ : -141.2 (2F), -141.8 (2F). HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{24}\text{H}_{17}\text{ClF}_4\text{NO}_4$   $[\text{M}]^+$ , 494.07822; found 494.07935.

**(S)-2-(((tert-butoxycarbonyl)amino)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid (1-44)**

(S)-2-((tert-butoxycarbonyl)amino)-3-(4-iodophenyl)propanoic acid (**1-43**) (2.0 g, 5.11 mmol), B<sub>2</sub>pin<sub>2</sub> (1.56 g, 6.14 mmol), Pd(dppf)Cl<sub>2</sub> (112 mg, 0.15 mmol) and KOAc (2.0 g, 20.4 mmol) were mixed with anhydrous DMSO (25 ml) in a pressure bottle under N<sub>2</sub> protection. The resulting mixture was stirred at 80 °C for 7 h, then poured into 1 N HCl (150 ml), and extracted with EtOAc (100 ml). The aqueous layer was dispensed, and the organic layer was washed with water (2 × 50 ml) and brine, then dried over Na<sub>2</sub>SO<sub>4</sub>. The product was isolated by silica gel column (hexane/EtOAc = 1:1, with 1% HOAc). Yield: 1.89 g (94%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ: 7.77 (d, *J* = 8.0 Hz, 2H), 7.20 (d, *J* = 8.0 Hz, 2H), 4.90 (d, *J* = 8.5 Hz, 1H), 4.58 (m, 1H), 3.22 (m, 1H), 3.15 (m, 1H), 1.42 (s, 9H), 1.36 (s, 12H).

**(S)-2-amino-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid (1-45)**

Neat HCl gas was bubbled into a solution of (S)-2-((tert-butoxycarbonyl)amino)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid (**1-44**) (1044 mg, 2.67 mmol) in anhydrous acetonitrile (20 ml) at r.t. for 2 min. The resulting solution was stirred at r.t. for another 30 min. The excess HCl was removed under vacuum. The product was isolated by filtration as white precipitate, washed with acetonitrile (2 ×), and dried under vacuum. Yield: 724 mg (83%). <sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O) δ: 7.63 (d, *J* = 8.0 Hz, 2H), 7.23 (d, *J* = 8.0 Hz, 2H), 4.19 (t, *J* = 6.0 Hz, 1H), 3.23 (m, 1H), 3.13 (m, 1H), 1.08 (s, 12H).

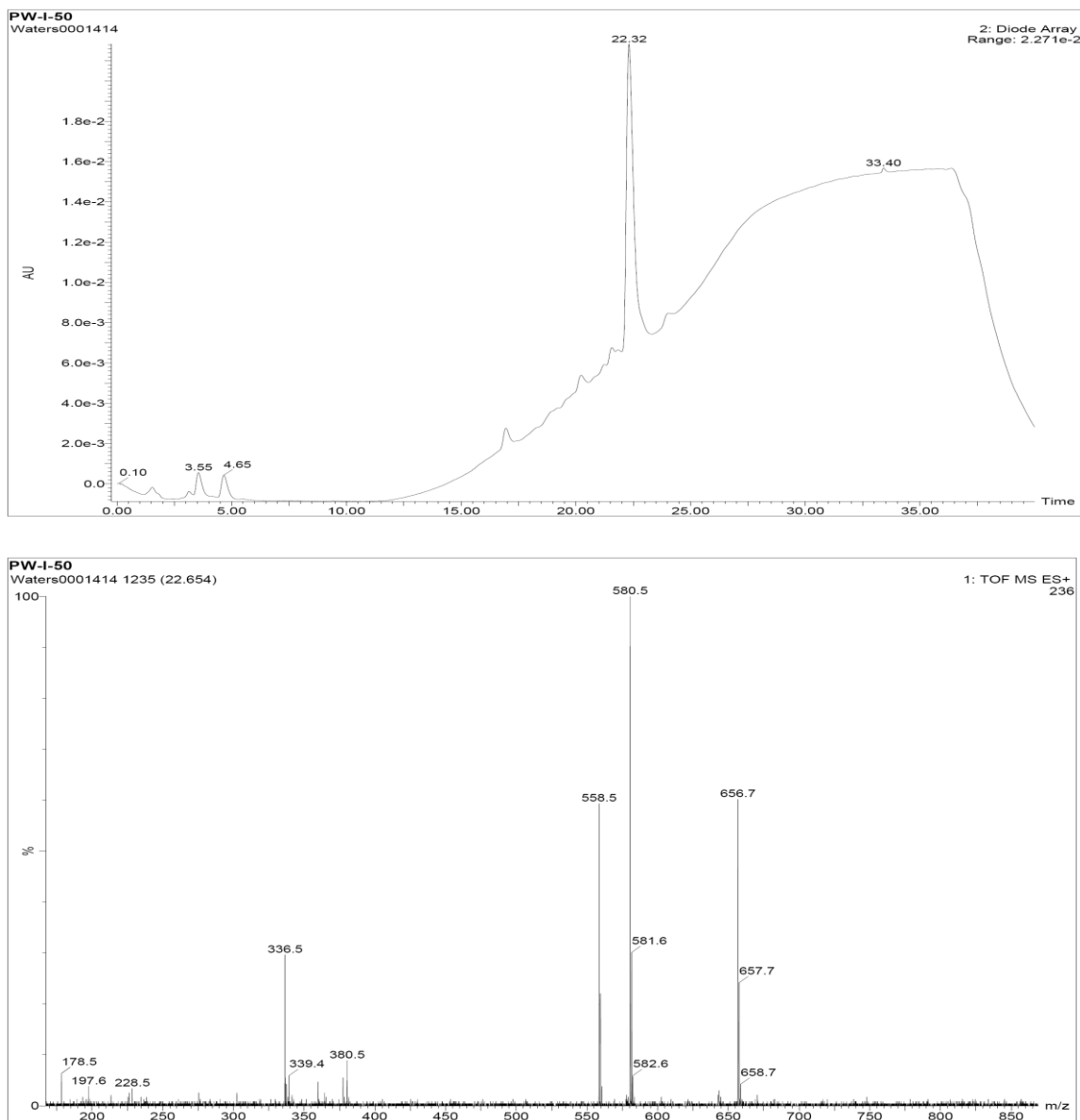
**(S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid (1-46)**

(S)-2-amino-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid (**1-45**) (624 mg, 1.90 mmol) was suspended in a mixture of DCM (50 ml) and DIPEA (0.99 ml, 5.7 mmol). The suspension was chilled to 0 °C and added with Fmoc-Cl (492 mg, 1.90 mmol) in portions. The reaction mixture was stirred vigorously at 0 °C for 0.5 h and at r.t. for another 1.5 h. the resulting solution was washed with 1 N HCl and brine, and then dried over Na<sub>2</sub>SO<sub>4</sub>. The product was isolated with silica gel column (hexanes/EtOAc = 3:2, with 1% HOAc). Yield: 769 mg (79%). <sup>1</sup>H NMR (500 MHz, Acetonitrile-*d*<sub>3</sub>) δ: 7.82 (d, *J* = 8.0 Hz, 2H), 7.66 (d, *J* = 8.0 Hz, 2H), 7.58 (m, 2H), 7.42 (m, 2H), 7.26-7.37 (m, 4H), 6.03 (d, *J* = 8.0 Hz, 1H), 4.57 (m, 1H), 4.16-4.38 (m, 3H), 3.25 (m, 1H), 3.00 (m, 1H) 1.35 (s, 12H).

**Conjugation of Fmoc-ZpOH-OH with (R)-(-)-3,3-Dimethyl-2-butylamine**

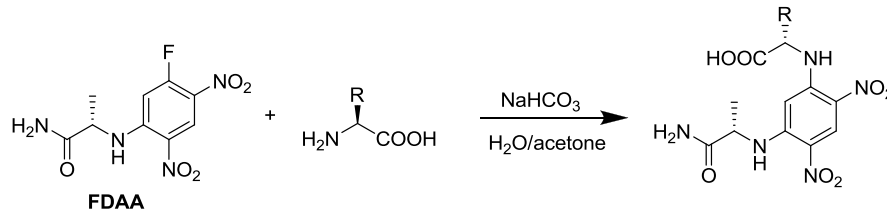
Under N<sub>2</sub> protection, (S)-2-((((9H-fluoren-9-yl) methoxy)carbonylamino)-3-(2,3,5,6-tetrafluoro-4-hydroxyphenyl)propanoic acid (**1-4**) (81.4mg, 0.17mmol) and HBTU (71.5mg 0.17mmol) were dissolved in 0.4M DIPEA in anhydrous THF (1ml). The mixture was allowed to stir for 2 minutes at room temperature before the addition of (R)-3,3-dimethyl-2-butylamine (0.27 ml, 2.0 mmol). The reaction was stirred at room temperature for overnight. The reaction mixture was quenched by 1N HCl (3ml) and extracted with DCM (3×5ml). The organic layers were combined, washed with a 10:1 mixture of brine and 1N HCl and dried over Na<sub>2</sub>SO<sub>4</sub>. After DCM was removed under

vacuum, the residues were dissolved in methanol for LC\_MS analysis (Figure 1-20). One single product was identified.  $m/z$  cal: 558.2 ( $M^+$ ) $m/z$  obs:558.5( $M^+$ ); 580.5 ( $M+Na^+$ ). The result suggested that no epimerization happened in our synthesis of Fmoc-ZpOH-OH.

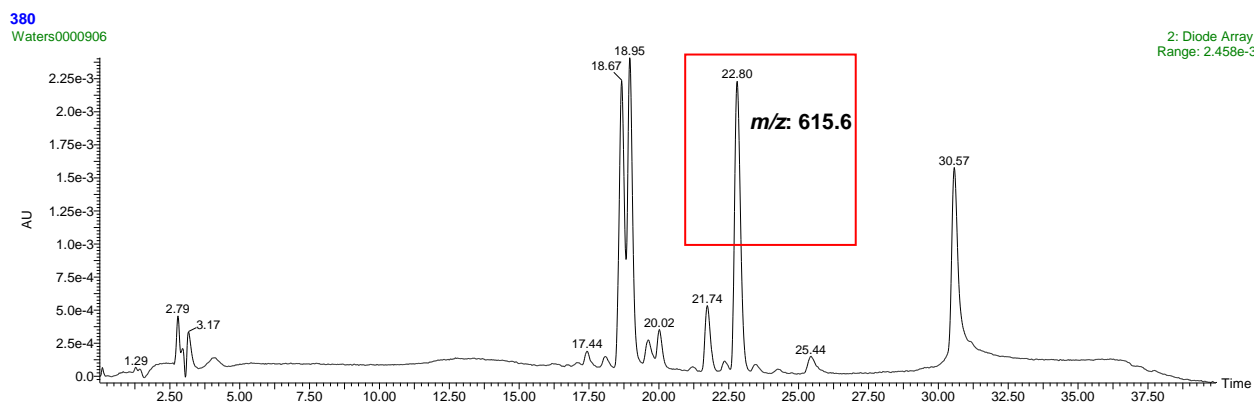


**Figure 1-20.** HPLC trace and the Mass spectrum of Conjugation product of FmocZpOH-OH with (R)-(-)-3,3-Dimethyl-2-butylamine

## Marfey's test to confirm on the stereochemistry of ZpXs



Fmoc- ZpXs were deprotected with piperidine in DMF, and the free amino acids were isolated by Prep HPLC. The amino acids were dissolved in water to give stock solutions of 50 mM. Then 50  $\mu$ l of an amino acid stock solution was mixed with 50  $\mu$ l water, 20  $\mu$ l of 1 M NaHCO<sub>3</sub> aq., and 100  $\mu$ l of the Marfey's reagent (1-Fluoro-2,4-dinitrophenyl-5-L-alanine amide, FDAA) stock solution (1% FDAA in acetone). The resulting mixture was incubated with gentle shaking at 35 °C for 1 h. The reaction was quenched by addition of 10  $\mu$ l 2 N HCl. The acidified solution was diluted to 1 ml with methanol, and examined by LC-MS by monitoring absorption at 340 nm. For all amino acids examined, a single peak was observed confirming lack of epimerization. Shown below (Figure 1-21) is the result of ZpI as an example, only one peak was identified with the molecular weight of the FDAA-coupled product.



**Figure 1-21.** HPLC trace of the Marfey's test for Fomc-ZpI-OH.

## Peptide synthesis and characterization

All peptides were synthesized on Rink-Amide-MBHA resin (90mg, 0.56mmol/g 0.05mmol) using the standard Fmoc/HBTU chemistry. Five equivalents of the commercially available amino acids were used for the coupling reaction.

For the ZpOH containing peptide, the incorporation of the ZpOH was accomplished by using three equivalents of the Fmoc-ZpOAllyl-OH and extended coupling time (1 hour). After cleavage of the N-terminal Fmoc group from the fully protected peptide on the resin, a dansyl residue was incorporated by reaction of the N-terminal amino group with dansyl chloride (40mg, 0.15mmol) in the presence of DIPEA (52 $\mu$ l, 0.30mmol) overnight. The orthogonal allyl protecting group in peptide **1** was removed by treating the resin-bound peptide with Pd(PPh<sub>3</sub>)<sub>4</sub> (100mg, 0.09mmol), and PhSiH<sub>3</sub>(0.4ml, 3mmol) in 2ml of DCM under argon protection for 3 hours, and repeated. Then the peptide was cleaved off the resin and deprotected fully with reagent K (80% TFA, 5% H<sub>2</sub>O, 2.5% EDT, 5% Thioanisole and 7.5% Phenol) at room temperature for two hours. The mixture was triturated three times with Et<sub>2</sub>O. The precipitate was collected and dried under vacuum. The crude peptides were purified by RP-HPLC (Waters Prep LC, Jupiter 10u C4 300A Column). The identity of each peptide was confirmed by MALDI mass spectrometric analysis. **Peptide 1**, m/z cal 2962(M<sup>+</sup>), found 2962 (M<sup>+</sup>); **peptide 2**, m/z cal 2746(M<sup>+</sup>), found, 2746(M<sup>+</sup>); **peptide 3**, m/z cal 2644(M<sup>+</sup>), found 2644(M<sup>+</sup>).

For the ZpN<sub>3</sub>/ZpCl/ZpBr containing peptide **4** (for the <sup>19</sup>F-<sup>19</sup>F cosy experiment , sequence: GIGKFL(ZpN<sub>3</sub>)SAKK(ZpBr)GKA(ZpCl)VGEIMNS), the incorporation of unnatural residues was accomplished by using three equivalents of the amino acids and extended coupling time (1 hour). Then the peptide was cleaved off the resin and deprotected fully with reagent K (80% TFA, 5% H<sub>2</sub>O, 2.5% EDT, 5% Thioanisole and 7.5% Phenol) at room temperature for two hours. The mixture was triturated three times with Et<sub>2</sub>O. The precipitate was collected and dried under vacuum. The crude peptides were purified by RP-HPLC (Waters Prep LC, Jupiter 10u C4 300A Column). The identity of the peptides was confirmed by ESI mass spectrometric analysis. Peptide **4**: for [M+3H]<sup>3+</sup>, *m/z* calculated 950.3, found 951.3; for [M+4H]<sup>4+</sup>, *m/z* calculated 713.0, found 713.2.

## **2D <sup>19</sup>F-<sup>19</sup>F COSY experiment for a fluorinated magainin.**

The magainin mutant (peptide **4** incorporating ZpBr, ZpCl and ZpN<sub>3</sub>) was dissolved in 90% DMSO-*d*<sub>6</sub> and 10% TRIS buffer (pH = 7.4) at the final concentration of 1 mM. The NMR data were collected on Varian 600 MHz NMR spectrometer and shown in Figure 1-11.

## **<sup>19</sup>F NMR Titration experiments for halogen bonding measurements**

The stock solutions of pI-4F-PhCOOMe (ZpI), pH-4F-PhCOOMe (Zp) and 5F-PhCOOMe (Z) were made as 0.05 M in anhy. cyclohexane, while pI-4F-PhCOOH (ZpI) and 5F-PhCOOH (Z) were dissolved to 0.05 M in water (10% D<sub>2</sub>O) at pH~10.

The following XB pairs were tested: ZpI, Zp, Z with pyridine (Py), 1-methyl pyrrole (Pyr) and 1-methyl indole (indole) in cyclohexane (9 totally); and ZpI, Z with imidazole (Im) in water (2 totally).

For each XB pair, 6 NMR samples with different XB acceptor concentrations (0, 0.02, 0.1, 0.5, 1 and 2 M) were made. Each sample contains 0.025 M ZpI (or Z, Zp), various concentration of the XB acceptor as shown above and the total volume is 0.6 ml. For samples made in cyclohexanes, 1% fluorobenzene (PhF) was added as <sup>19</sup>F NMR standard (-113.15 ppm), and cyclohexane-*d*<sub>12</sub> sealed in capillary tubes was used as pseudo-internal deuterium lock signals. For samples made in water (10% D<sub>2</sub>O), 1% fluoroacetonitrile was added as NMR standard (-224.7 ppm).

Each NMR sample was scanned 24 ct on VNMR 500 spectrometer. The chemical shifts of the *meta*-fluorine atoms were recorded and fitted as hyperbola in *Origin* (Figure 1-16).

## Liposome preparation

*Buffers:*

**Buffer A:** 10mM TES (pH=7.0 or pH=5.0), 100mM sodium citrate;

**Buffer B:** 10mM TES (pH=7.0), 100mM sodium citrate, 20mM calcein;



**Buffer C:** 10mM TES (pH=7.0), 3M guanidinium chloride, 100mM sodium citrate.

#### *Preparation of POPC LUVs*

A chloroform solution containing 152mg POPC was dried under reduced pressure to form a thin film that was mixed with 2ml **Buffer A** or **Buffer B**. After 20 cycles of freezing and thawing, the liposome suspension was extruded 21 times through a 100nm polycarbonate membrane (Liposofast Mini Extrusion system, Avanti Polar Lipid; Whatman Nuclepore Track-Etch Membrane) at room temperature. Free calcein was removed from liposome by gel filtration (ÄKTA FPLC with HiPrep<sup>TM</sup> 16/60 Sephacryl TM-S-500HR column). The liposome hydrolyzed by **Buffer A**, gel filtration was used to separate small liposome out, making the liposome more uniform in size. The size of liposome was confirmed by Dynamic Light Scattering (DLS) (DynaPro nanostar, Wyatt, Santa Barbara, CA). All liposomes we made has a radii around 70nm and the polydispersity falls into the range of 15%-30%.

#### **Calcein leakage assay**

Peptides **1-3** were dissolved in **Buffer C** and concentration of the stock solutions were determined based on dansyl absorbance at 330nm ( $\epsilon = 4,300 \text{ M}^{-1} \text{ cm}^{-1}$ ). The calcein containing POPC LUVs were diluted to a final lipid concentration of 500 $\mu\text{M}$  in **Buffer A** and placed into a fluorescence cell with magnetic stirring. The membrane lysis caused by mag2 peptides were monitored by the calcein fluorescence at 520nm ( $\lambda_{\text{ex}}=485\text{nm}$ ) on a fluorescence spectrometer (Horiba Jobin Yvon Fluorolog FL 3-22). During the

experiment, peptide stock was added into the POPC( $\geq 40$  mM calcein) LUVs suspension with a final concentration of  $3 \mu\text{M}$  at 1,400 s, and  $12 \mu\text{l}$  of 6N HCl was added at 5,000 s to tune the solution pH to 5.0, finally,  $30 \mu\text{l}$  10% w/v TritonX-100 was added to cause the total leakage at 8,600 s (Figure 1-18).

## References

1. Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Kokschi, B., *Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions*. Chem. Soc. Rev., 2012. **41**: p. 2135-2171.
2. Jackel, C.; Kokschi, B., *Fluorine in Peptide Design and Protein Engineering*. Euro. J. Org. Chem., 2005: p. 4483-4503.
3. Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V., *Fluorine in medicinal chemistry*. Chem. Soc. Rev., 2008. **37**: p. 320-330.
4. Muller, K.; Faeh, C.; Diederich, F., *Fluorine in pharmaceuticals: looking beyond intuition* Science, 2007. **317**: p. 1881-1886.
5. Takeuchi, Y.; Shiragami, T.; Kimura, K.; Suzuki, E.; Shibata, N., *(R)- and (S)-3-Fluorothalidomides: isosteric analogues of Thalidomide*. Org. Lett., 1999. **1**: p. 1571-1573.
6. Silverman, W. A., *The schizophrenic career of a "monster drug"*. Pediatrics, 2002. **110**: p. 404-406.
7. *The approval for Thalidomide*. National Cancer Institute, 2012.
8. Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G., *Tables of bond lengths determined by X-Ray and neutron diffraction*. J. Chem. Soc. Perkin Trans. II, 1987: p. S1-S19.

9. Wong, D. T.; Bymaster, F. P.; Engleman, E. A., *Prozac (fluoxetine, Lilly 110140), the first selective serotonin uptake inhibitor and an antidepressant drug: twenty years since its first publication*. Life Sci., 1995. **57**: p. 411-441.
10. Roman, D. L.; Walline, C. C.; Rodriguez, G. J.; Barker, E. L., *Interactions of antidepressants with the serotonin transporter: a contemporary molecular analysis*. Eur. J. Pharmacol., 2003. **479**: p. 53-63.
11. Dunitz, J. D., *Organic Fluorine: Odd Man Out*. ChemBioChem, 2004. **5**: p. 614-621.
12. Biffinger, J. C.; Kim, H.; DiMagno, S. G., *The polar hydrophobicity of fluorinated compounds*. ChemBioChem, 2004. **5**: p. 622-627.
13. Smart, B. E., *Fluorine substituent effects (on bioactivity)*. J. Fluorine Chem., 2001. **109**: p. 3-11.
14. Massa, M. A.; Spangler, D. P.; Durley, R. C.; Hickory, B. S.; Connolly, D. T.; Witherbee, B. J.; Smith, M. E.; Sikorski, J. A., *Novel heteroaryl replacements of aromatic 3-tetrafluoroethoxy substituents in trifluoro-3-(tertiaryamino)-2-propanols as potent inhibitors of cholesteryl ester transfer protein*. Bioorg. Med. Chem. Lett., 2001. **11**: p. 1625-1628.
15. Yeh, H. J. C.; Kirk, K. L.; Cohen, L. A.; Cohen, J. S., *<sup>19</sup>F and <sup>1</sup>H nuclear magnetic resonance studies of ring-fluorinated imidazoles and histidines*. J. Pharm. Soc., Perkin II, 1975: p. 928.

16. Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A., *A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues*. *Science*, 1994. **266**: p. 243-247.
17. Li, C.; Wang, G.; Wang, Y.; Creager-Allen, R.; Lutz, E. A.; Scronce, H.; Slade, K. M.; Ruf, R. A. S.; Mehl, R. A.; Pielak, G. J., *Protein <sup>19</sup>F NMR in Escherichia coli*. *J. Am. Chem. Soc.*, 2010. **132**: p. 321-327.
18. Sun, Z.-Y.; Pratt, E. A.; Ho, C., *Biomedical frontiers of fluorine chemistry*, 1996. **ACS Symposium Series 639**: p. 296-310.
19. Grage, S. L.; Wang, J.; Cross, T. A.; Ulrich, A. S., *Solid-state <sup>19</sup>F-NMR analysis of <sup>19</sup>F-labeled tryptophan in gramicidin A in oriented membranes*. *Biophys. J.*, 2002. **83**: p. 3336-3350.
20. Bailey, D. L.; Townsend, D. W.; Valk P. E.; Maisey, M. N., *Positron Emission Tomography: Basic Sciences*. 2005.
21. *Fludeoxyglucose drug information*. 2009.
22. Lee, E.; Kamlet, A. S.; Powers, D. C.; Neumann, C. N.; Boursalian, G. B.; Furuya, T.; Choi, D. C.; Hooker, J. M.; Ritter, T., *A Fluoride-Derived Electrophilic Late-Stage Fluorination Reagent for PET Imaging*. *Science*, 2011. **334**: p. 639-642.
23. Ritter, T., *Catalysis: Fluorination made easier*. *Nature*, 2010. **466**: p. 447-448.
24. Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T., *Code for collagen's stability deciphered*. *Nature*, 1998. **392**: p. 666-667.
25. DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines R. T.; Markley, J. L., *Collagen Stability: Insights from NMR*

- Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations.* J. Am. Chem. Soc., 2002. **124**: p. 2497-2505.
26. Renner, C.; Alefelder, S. A.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L., *Fluoroprolines as Tools for Protein Design and Engineering.* Angew. Chem. Int. Ed., 2001. **40**: p. 923-925.
27. Kumar, K.; Bilgicer, B., *De novo design of defined helical bundles in membrane environments.* Proc. Natl. Acad. Sci. USA, 2004. **101**: p. 15324-15329.
28. Pace, C. J.; Gao, J., *Exploring and exploiting polar- $\pi$  interactions with fluorinated aromatic amino acids.* 2012.
29. Politzer, P.; Huheeyb, J. E.; Murraya, J. S.; Grodzicki, M., *Electronegativity and the concept of charge capacity.* Journal of Molecular Structure: THEOCHEM, 1992. **259**: p. 99-120.
30. Gallivan, J. P.; Dougherty, D. A., *Cation- $\pi$  interactions in structural biology.* Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 9459-9464.
31. Zheng, H.; Comeforo, K.; Gao, J., *Expanding the Fluorous Arsenal: Tetrafluorinated Phenylalanines for Protein Design.* J. Am. Chem. Soc., 2009. **131**: p. 18-19.
32. Zheng, J.; Gao, J., *Highly Specific Heterodimerization Mediated by Quadrupole Interactions.* Angew. Chem. Int. Ed., 2010. **49**: p. 8635-8639.
33. Pace, C.; Kim, D; Gao, J., *Experimental Evaluation of CH- $\pi$  Interactions in a Protein Core.* Chem. Eur. J., 2012.

34. Kim, K.; Sondhi, D.; Cole, P. A., *Chemical approaches to the study of protein kinases and their implications for mechanism and inhibitor design*. *Pharmacol. Ther.*, 1999. **82**: p. 219-229.
35. Kim, K.; Cole, P. A., *Kinetic analysis of a protein tyrosine kinase transition state in the forward and reverse directions*. *J. Am. Chem. Soc.*, 1998. **120**: p. 6851-6858.
36. Wang, D.; Cole, P. A., *Protein tyrosine kinase Csk-catalyzed phosphorylation of Src containing unnatural tyrosine analogs*. *J. Am. Chem. Soc.*, 2001. **123**: p. 8883-8886.
37. Seyedsayamdost, M. R.; Reece, S. Y.; Nocera, D. G.; Stubbe, J., *Mono-, di-, tri-, and tetra-substituted fluorotyrosines: new probes for enzymes that use tyrosyl radicals in catalysis*. *J. Am. Chem. Soc.*, 2006. **128**: p. 1569-1579.
38. Seyedsayamdost, M. R.; Yee, C. S.; Stubbe, J., *Site-specific incorporation of fluorotyrosines into the R2 subunit of E. coli ribonucleotide reductase by expressed protein ligation*. *Nat. Protoc.*, 2007. **2**: p. 1225-1235.
39. Rodionov, P. P.; Furin, G. G., *Kinetics of nucleophilic substitution reactions of polyfluoroaromatic compounds*. *J. Fluorine Chem.*, 1990. **47**: p. 361-434.
40. Chambers, R. D., *Fluorine in Organic Chemistry*.
41. Filler, R.; Ayyangar, N. R.; Gustowski, W.; Kang, H. H., *New reactions of polyfluoroaromatic compounds. Pentafluorophenylalanine and tetrafluorotyrosine*. *J. Org. Chem.*, 1969. **34**: p. 534-538.

42. Demidkina, T. V.; Barbolina, M. V.; Faleev, N. G.; Sundararaju, B.; Gollnick, P. D.; Phillips, R. S., *Threonine-124 and phenylalanine-448 in Citrobacter freundii tyrosine phenol-lyase are necessary for activity with L-tyrosine*. *Biochem. J.*, 2002. **363**: p. 745-752.
43. Feutrill, G. I.; Mirrington, R. N., *Demethylation of aryl methyl ethers with thioethoxide ion in dimethyl formamide* *Tetrahedron Lett.*, 1970. **11**: p. 1327-1328.
44. Dessolin, M.; Guillerez, M. G.; Thieriet, N.; Guibe, F.; Loffet, A., *New allyl group acceptors for palladium catalyzed removal of allylic protections and transacylation of allyl carbamates* *Tetrahedron Lett.*, 1995. **36**: p. 5741-5744.
45. Coin, I.; Beyermann, M.; Bienert, M., *Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences*. *Nat. Protoc.*, 2007. **2**: p. 3247-3256.
46. Suzuki, A., *Organoboron compounds in new synthetic reactions*. *Pure Appl. Chem.*, 1985. **57**: p. 1749-1758.
47. Charrier, N.; Demont, E.; Dunsdon, R.; Maile, G.; Naylor, A.; O'Brien, A.; Redshaw, S.; Theobald, P.; Vesey, D.; Walter, D., *Synthesis of indoles: efficient functionalization of the 7-position*. *Synthesis*, 2006. **20**: p. 3467-3477.
48. Bhushan, R.; Brückner, H., *Marfey's reagent for chiral amino acid analysis: A review*. *Amino Acids*, 2004. **27**: p. 231-247.
49. Sonogashira, K., *Coupling reactions between  $sp^2$  and  $sp$  carbon centers*. *Comp. Org. Synth.*, 1991. **3**: p. 521-549.



50. Ishiyama, T.; Murata, M.; Miyaura, N., *Palladium(0)-Catalyzed Cross-Coupling Reaction of Alkoxydiboron with Haloarenes: A Direct Procedure for Arylboronic Esters*. *J. Org. Chem.*, 1995. **60**: p. 7508-7510.
51. Armstrong, A.; Brackenridge, I.; Jackson, R. F. W.; Kirk, J. M., *A new method for the preparation of tertiary butyl ethers and esters*. *Tetrahedron Lett.*, 1988. **29**: p. 2483-2486.
52. Adonin, N. Y.; Bardin, V. V.; Flörke, U.; Frohn, H., *Polyfluoroorganoboron-Oxygen Compounds. 4 [1] Lithium Pentafluorophenyltrimethoxyborate, Li[C6F5B(OMe)3], Reactions with Selected Electrophiles and Nucleophiles* *Zeitschrift für anorganische und allgemeine Chemie*, 2005. **631**: p. 2638-2646.
53. Nishimura, O.; Kitada, C.; Fujino, M., *New Method for Removing the S-p-Methoxybenzyl and S-t-Butyl Groups of Cysteine Residues with Mercuric Trifluoroacetate*. *Chem. Pharm. Bull.*, 1978. **26**: p. 1576-1585.
54. Macmillan, D.; Anderson D.W., *Rapid synthesis of acyltransfer auxiliaries for cysteine-free Native Glycopeptide Ligation*. *Org. Lett.*, 2004. **6**: p. 4659-4662.
55. Redman, J. E., and Ghadiri, M. R., *Synthesis of photoactive p-azidotetrafluorophenylalanine containing peptide by solid-phase Fmoc methodology*. *Org. Lett.*, 2002. **4**: p. 4467-4469.
56. Bayley, H.; Standring, D. N.; Knowles, J. R., *A selective reagent for the efficient reduction of alkyl and aryl azides to amines*. *Tetrahedron Lett.*, 1978. **39**: p. 3633-3634.
57. Hodgson, H. H., *The Sandmeyer reaction*. *Chem. Rev.*, 1947. **40**: p. 251-277.

58. Cowdrey, W. A.; Davies, D. S., *Quart. Rev.*, 1952. **6**: p. 358-379.
59. Ek, F.; Wistrand, L. G.; Frejd, T., *Aromatic Allylation via Diazotization: Variation of the Allylic Moiety and a Short Route to a Benzazepine Derivative*. *J. Org. Chem.*, 2003. **68**: p. 1911-1918.
60. Schölkopf, U.; Groth, U.; Deng, C., *Enantioselective Syntheses of (R)-Amino Acids Using L-Valine as Chiral Agent*. *Angew. Chem. Int. Ed.*, 2003. **20**: p. 798-799.
61. Chen, J.; Corbin, S. P.; Holman, N. J., *An Improved Large Scale Synthesis of the Schölkopf Chiral Auxiliaries: (2R)- and (2S)-2,5-Dihydro-3,6-dimethoxy-2-isopropylpyrazine*. *Org. Process Res. Dev.*, 2005. **9**: p. 185-187.
62. Meerwein, H.; Schmidt, R., *New method for the reduction of aldehydes and ketones*. *Ann.*, 1925. **444**: p. 221-238.
63. Matsuzaki, K., *Magainins as paradigm for the mode of action of pore forming polypeptides*. *Biochim. Biophys. Acta.*, 1998. **1376**: p. 391-400.
64. Yonemori, S.; Sasakura, H., *Structural determination of fluoro-organic mixtures by <sup>19</sup>F COSY NMR spectroscopy*. *J. Fluorine Chem.*, 1995. **75**: p. 151-156.
65. Metrangolo, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G., *Halogen bonding in supramolecular chemistry*. *Angew. Chem. Int. Ed.*, 2008. **47**: p. 6114-6127.
66. Xu, L.; Zou, J.; Yu, Q.; Sang, P., *Computational insights into halogen bonding between P-Cl contact and several electron donors*. *Int. J. Quant. Chem.*, 2009. **110**: p. 1245-1251.

67. Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F., *Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions*. *Angew. Chem. Int. Ed.*, 2011. **50**: p. 314-318.
68. Luo, Z.; Ding, J.; Zhou, Y., *Temperature-Dependent Folding Pathways of Pin1 WW Domain: An All-Atom Molecular Dynamics Simulation of a Go Model*. *Biophysical J.*, 2007. **93**: p. 2152-2161.
69. Lu, Y.; Zou, J.; Wang, Y.; Yu, Q., *Theoretical investigations of the C-X/ $\pi$  interactions between benzene and some model halocarbons*. *Chemical Physics*, 2007. **334**: p. 1-7.
70. Sarwar, M. G.; Dragisic, B.; Salsberg, L. J.; Gouliaras, C.; Taylor, M. S., *Thermodynamics of halogen bonding in solution substituent structural and solvent effects*. *J. Am. Chem. Soc.*, 2010. **132**: p. 1646-1653.
71. Bocharov, E. V.; Mineev, K. S.; Volynsky, P. E.; Ermolyuk, Y. S.; Tkach, E. N.; Sobol, A. G.; Chupin, V. V.; Kirpichnikov, M. P.; Efremov, R. G.; Arseniev, A. S., *Spatial Structure of the Dimeric Transmembrane Domain of the Growth Factor Receptor ErbB2 Presumably Corresponding to the Receptor Active State*. *J. Bio. Chem.*, 2008. **283**: p. 6950-6956.
72. Scott, R. W.; DeGrado, W. F.; Tew, G. N., *De novo designed synthetic mimics of antimicrobial peptides*. *Curr Opin Biotechnol.*, 2008. **19**: p. 620-627.

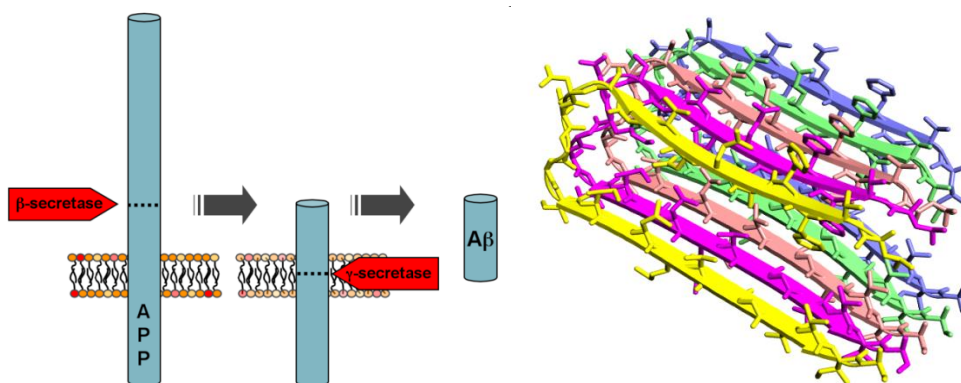
73. Porter, E. A.; Weisblum, B.; Gellman, S. H., *Mimicry of host-defense peptides by unnatural oligomers: antimicrobial beta-peptides*. J. Am. Chem. Soc., 2002. **124**: p. 7324-7330.
74. Fernandez-Lopez, S., Kim, H.-S., Choi, E. C., Delgado, M., Granja, J. R., Khasanov, A., Kraehenbuehl, K., Long, G., Weinberger, D. A., Wilcoxon, K. M., and Ghadiri, M. R., *Antibacterial agents based on the cyclic D,L-alpha-peptide architecture*. Nature, 2001. **412**: p. 452-456.
75. Chongsiriwatana, N. P.; Patch, J. A.; Czyzewski, A. M.; Dohm, M. T.; Ivankin, A.; Gidalevitz, D.; Zuckermann, R. N.; Barron, A. E., *Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides*. Proc. Natl. Acad. Sci. USA, 2008. **105**: p. 2794-2799.
76. Thevenin, D.; An, M.; Engelman, D. M., *pHLIP-Mediated Translocation of Membrane-Impermeable Molecules into Cells*. Chem. Bio., 2009. **16**: p. 754-762.
77. Lee, H. M.; Chmielewski, J., *Liposomal Cargo Unloading Induced By pH-Sensitive Peptides*. J. Pept. Res., 2005. **65**: p. 355-363.
78. Yan, J.; Springsteen, G.; Deeter, S.; Wang, B., *The relationship among  $pK_a$ , pH, and binding constants in the interactions between boronic acids and diols: it is not as simple as it appears*. Tetrahedron, 2004. **60**: p. 11205-11209.
79. Duggan, P. J.; Offermann, D. A., *The preparation of solid-supported peptide boronic acids derived from 4-borono-L-phenylalanine and their affinity for alizarin*. Aust. J. Chem., 2007. **60**: p. 829-834.

## **Chapter 2:**

### Thioflavin T Dimers as Novel Amyloid Ligands

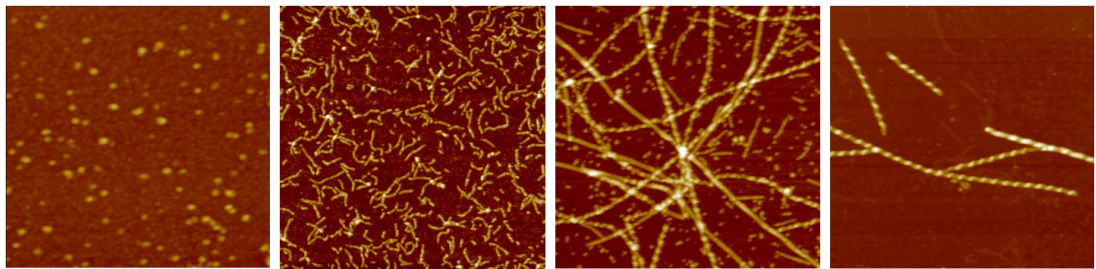
## 2.1 Introduction

Protein aggregation into fibrillar beta-sheet rich structures, termed amyloid, has been associated with an increasing number of human diseases. As a prominent example of this group, Alzheimer's disease (AD) affects 4.5 million people in the United States and over 14 million worldwide, with most of the sufferers being people over 65 years of age<sup>[1]</sup>. The symptoms of AD include confusion, mood swings, troubles with language and memory loss<sup>[2]</sup>. Gradually, body functions are weakened and lost, which ultimately leads to death<sup>[3]</sup>. While many details of the AD pathology are still unknown, it is well recognized that aggregation of the amyloid beta peptide ( $A\beta$ , see Figure 2-1) is a hallmark event<sup>[4]</sup>.  $A\beta$  is processed from the amyloid precursor protein with a sequential cleavage by  $\beta$ -secretase and  $\gamma$ -secretase<sup>[5]</sup>, but their original functions are still poorly understood.  $A\beta$  ranges from 39 to 43 amino acids, and the 40-residue variant ( $A\beta_{40}$ ) is the major protein component of the plaques observed in the brains of AD patients<sup>[6]</sup>.



**Figure 2-1.** Generation of  $A\beta$  (left) and the formation (right) of the aggregates.

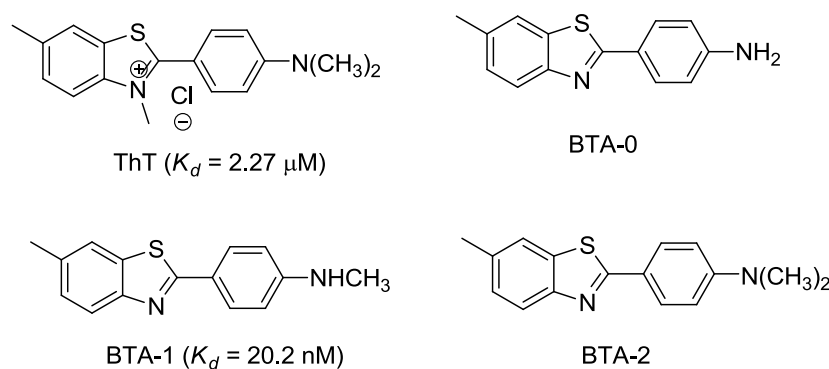
A $\beta$  aggregates can be prepared *in vitro* using chemically synthesized peptides. The monomeric A $\beta$  peptides assemble into oligomers, and finally turn into an extended fibril structure (Figure 2-2). They exhibit low nM toxicity towards neuronal cells, while the monomeric peptide is harmless. However, A $\beta$ 40 aggregation usually yields a heterogeneous mixture of species with varying size and morphology<sup>[7]</sup>, and to date the structure-toxicity relationship of the aggregated A $\beta$  still remains controversial. Nevertheless, the collective amount of A $\beta$  aggregates presents an important parameter in all branches of research related to AD.



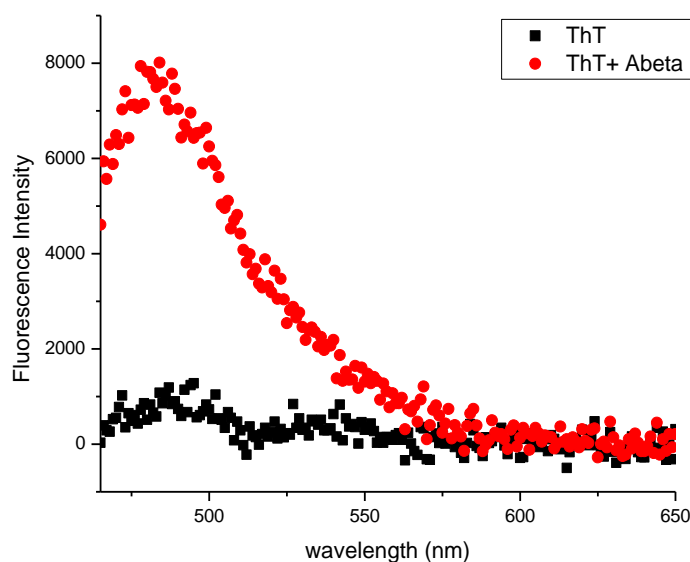
**Figure 2-2.** The aggregation process of A $\beta$ : from monomers to the oligomers and finally the fibril structures. Each picture is on the scale of 1  $\mu\text{m} \times 1 \mu\text{m}$

The ability of quantifying A $\beta$  aggregates is highly desirable for the early diagnosis of AD in a clinical setup, monitoring disease progression and evaluating the efficacy of the therapeutic drugs. However, there is no efficient *in vivo* method yet, while the most commonly used *in vitro* method for amyloid detection is the thioflavin T (ThT, Figure 2-3) assay<sup>[8]</sup>. This cationic dye binds to the aggregated A $\beta$ 40 with an apparent dissociation constant of  $\sim 1 \mu\text{M}$ , while exhibiting no detectable binding affinity to A $\beta$  monomers. When free in aqueous solution, ThT exhibits a poor fluorescence emission.

Binding onto amyloid species elicits a strong fluorescence peaked at 485 nm (Figure 2-4). The fluorescence intensity has been demonstrated to be proportional to the A $\beta$  amyloid concentration<sup>[8]</sup>. Though widely used in lab research, the relatively low amyloid binding affinity of ThT prohibits its application in monitoring A $\beta$  aggregates in the nanomolar range, which has been shown to be the relevant concentration of A $\beta$



**Figure 2-3.** Structures of ThT and its neutral analogues



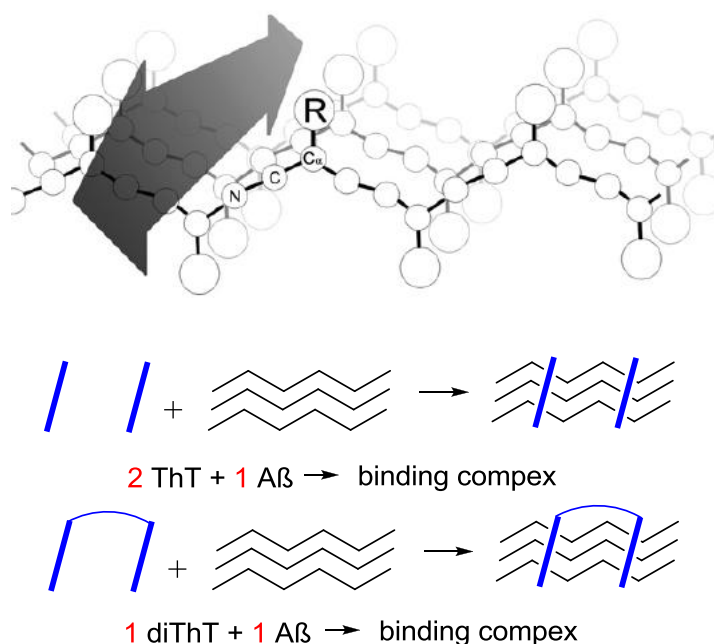
**Figure 2-4.** The fluorescence emission spectra of ThT: 1  $\mu\text{M}$  ThT with/without A $\beta$  40 in 1  $\times$  buffer (0.3 M NaCl, 0.05M phosphates, pH 7.4), excitation at 440 nm, slits 6 nm, emission scan 465-600 nm in microplate reader (50  $\mu\text{L}$  each sample).



toxicity<sup>[9]</sup>. Eliminating the positive charge of ThT results in neutral analogues, named as the BTA series<sup>[10]</sup> (Figure 2-3), which display significantly enhanced binding affinities to A $\beta$  amyloid. For example, the compound BTA-1 (Figure 2-3) binds to A $\beta$  amyloid tighter than ThT by two orders of magnitude, with an apparent  $K_d$  value of 20 nM<sup>[11]</sup>. However, these molecules display no fluorescence change upon amyloid binding. In addition, the BTA molecules are highly hydrophobic, resulting in poor solubility in aqueous solutions and non-specific binding to biomolecules other than amyloids<sup>[12]</sup>. In terms of the binding mechanism, to date there is no atomic-level structure available for the amyloid-ThT complexes. Biochemical and spectroscopic data suggest that ThT and BTA-1 molecules bind A $\beta$ 40 amyloid with its long dimension parallel to the amyloid fibril axis (perpendicular to the beta-strands) by inserting into the grooves of beta-sheet between  $i$  and  $i+2$  side chains<sup>[13-14]</sup> (Figure 2-5). The binding is driven by hydrophobicity interactions and  $\pi$ -stacking. Due to the differences in terms of the hydrophobicity and conformations, BTA-1 binds onto slightly different positions of A $\beta$  compared with ThT, but the difference is minor<sup>[15]</sup>.

Considering the binding mechanism and the highly repetitive nature of amyloid structures<sup>[16]</sup>, we hypothesized that an oligomeric construct of known amyloid ligands will elicit stronger association by reducing the entropic cost of binding (Figure 2-5). To test the hypothesis, we have synthesized a group of ThT dimers and examined their potentials as amyloid staining ligands. The results show that the dimeric design affords significantly enhanced binding to A $\beta$  amyloid, with the  $K_d$  values reduced by

as many as 70 folds. While comparable to the BTA series in amyloid-binding affinity, the dimeric ThTs display dramatically enhanced fluorescence emission upon binding; i.e. the ThT dimers combine the attractive features of both BTA and ThT as fluorescent reporters of amyloid.

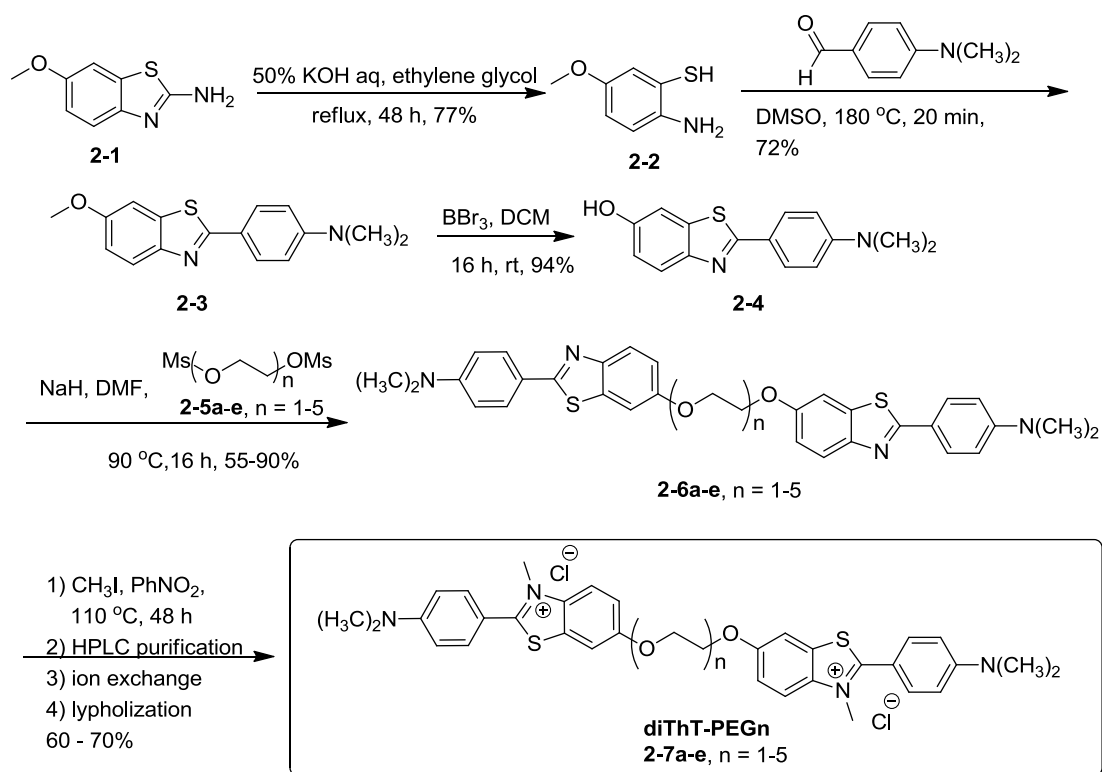


**Figure 2-5.** The binding mode of ThT and BTA-1 with Aβ aggregates (top) and the hypothesis of the ThT dimer design (bottom)

## 2.2 Synthesis of ThT dimer series

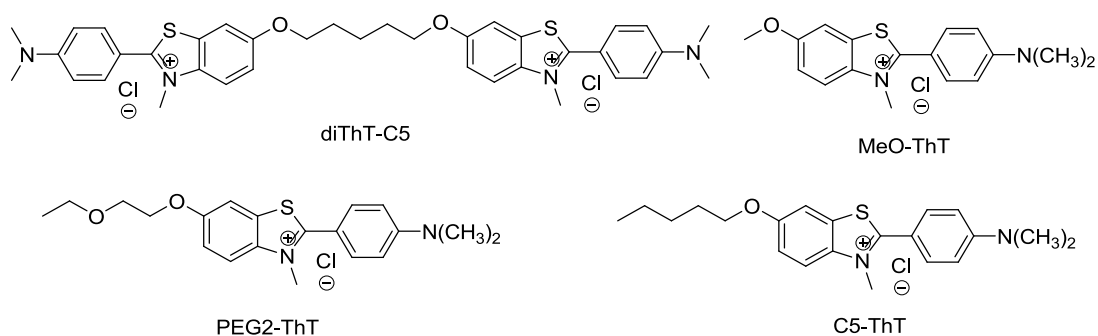
Based on our hypothesis about the binding entropy, we designed and synthesized a series of ThT dimers (Scheme 2-1). The two ThT units are linked in a head-to-head fashion via polyethylene glycols of varying lengths. The synthesis was started with the hydrolysis of 2-amino-6-methoxybenzothiazole (compound **2-1**) in 50% KOH aq.

with ethylene glycol as the co-solvent. Then compound **2-2** was condensed with 4-dimethylaminobenzaldehyde then oxidized in DMSO to regenerate the thiazole heterocycle<sup>[17]</sup> (compound **2-3**). Compound **2-3** was efficiently demethoxylated by using boron tribromide in DCM to yield the 6-hydroxy derivative of BTA-2 (compound **2-4**). Coupling of the deprotonated compound **2-4** to the dimesylated polyethylene glycols in DMF proceeded smoothly to give the dimers of BTA-2 (compounds **2-5a – 2-5e**). Fortuitously, the dimers of BTA-2 were poorly soluble in DMF and precipitated out of solution once they formed. Thus, a simple filtration yielded the BTA-2 dimers in the pure form. Some alternative methods were also tested to bond the linker with compound **2-4**, such as the Mitsunobu reaction, but the conversion was not complete due to the chemical nature of polyethylene glycol.



**Scheme 2-1.** Synthesis of ThT dimers

BTA-2 dimers are poorly soluble in all kinds of common organic solvents, especially for **2-6a** and **2-6b**, which have relatively shorter linkers. After screening a series of solvents, nitrobenzene was found to be the only viable one for the final methylation step. Interestingly, methylation with excess amount of CH<sub>3</sub>I at 70 °C yielded a messy mixture of products, yet at 120 °C afforded the desired ThT dimers (**2-7a** – **2-7e**) in good yields. Presumably at the lower temperature, S-methylation and N-methylation of the benzothiazole moiety kinetically compete and the S-methylated product further degrades to yield multiple side products. At the elevated temperature, N-methylation wins out; the resulting cationic benzothiazole prevents further alkylation on the sulfur atom and thus remains intact.

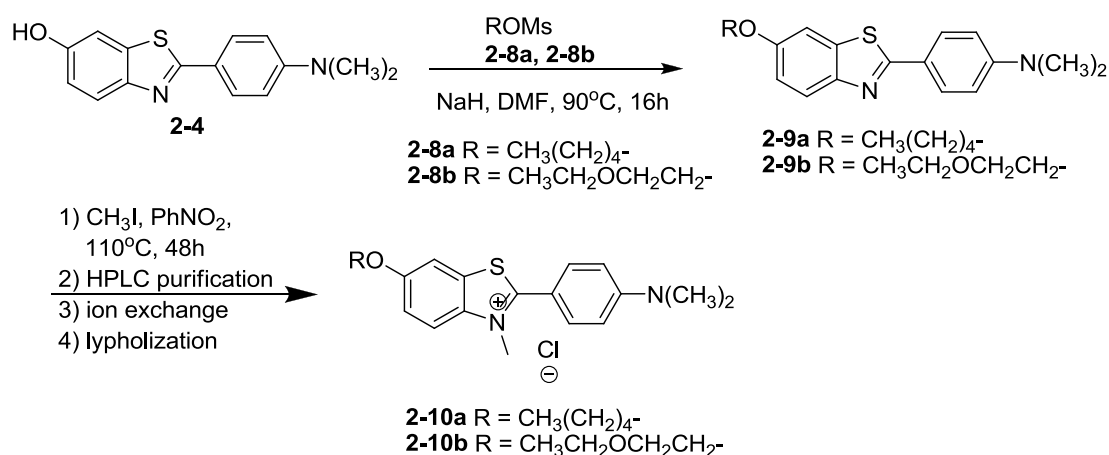


**Figure 2-6.** Other ThT derivatives as controls

The crude products were purified through reverse phase HPLC using a C18 column with water/acetonitrile (with 0.1% TFA) as eluents. The HPLC purified ThT dimers have trifluoroacetate as the counter-ion. Coevaporation five times with dilute HCl solution allowed the complete conversion of the counter-ion to chloride. This counter-ion exchange was confirmed by <sup>19</sup>F NMR. The overall yields were 17-35% over 5 steps. For convenience of discussion, we named the dimers as diThT-PEG<sub>n</sub>,

with n referring to the number of ethylene glycol units within the linker.

With the same strategies, we also synthesized a series of control molecules, including an isostere of the diThT-PEG2 (diThT-C5, **compound 2-7f**, Figure 2-6) with the pentylene group as the linker instead of the PEG2 moiety, and three monomeric ThT derivatives with different linkers on the 6-position (MeO-ThT, PEG2-ThT and C5-ThT, Figure 2-6). MeO-ThT was obtained by directly methylating compound 2-3 and the synthesis of C5-ThT and PEG2-ThT is shown in Scheme 2-2 (Compound **2-10a** and **2-10b**). The related assays about these control molecules will be introduced in the following sections.



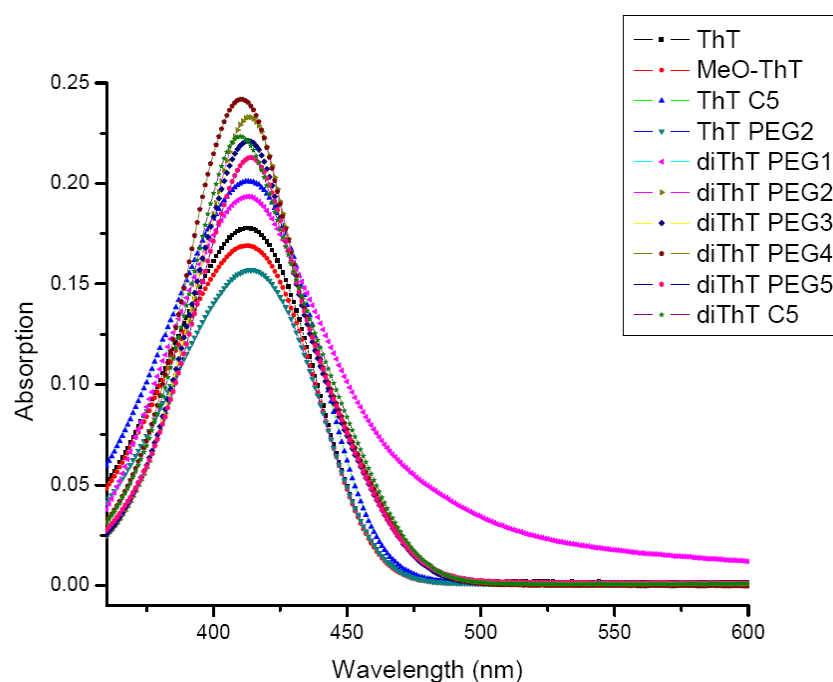
**Scheme 2-2.** Synthesis of C5-ThT and PEG2-ThT

### 2.3 A $\beta$ 40 binding affinity measurements

In order to fairly compare the A $\beta$  binding capability between the monomeric ThT and the dimers, we firstly tested several physical and chemical properties of the

dimers to make sure they behave similarly.

Firstly we tested the solubility of all the ThT derivatives in aqueous solution. According to the UV-Vis absorption data (Figure 2-7), all ThT derivatives display no tailing in their absorption spectra at 5  $\mu\text{M}$  (dimers at 2.5  $\mu\text{M}$ ) except diThT PEG1, indicating that they are well-soluble at this concentration. Nevertheless, all the ThT dimers and monomers exhibit similar absorption curves (Figure 2-7) and extinction co-efficients values (Table 2-1). The dimerization does not affect the general optical properties of the ThT core.

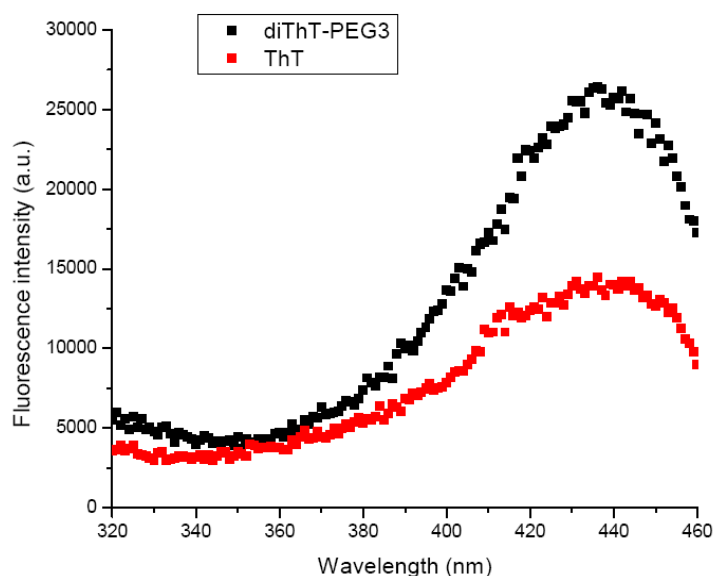


**Figure 2-7.** UV/Vis absorption spectra of ThT derivatives: 5  $\mu\text{M}$  (dimers at 2.5  $\mu\text{M}$ ) of the ThT derivatives in 1  $\times$  buffer (0.3 M NaCl, 0.05M phosphates, pH 7.4)

ThT derivatives	Extinction coefficient (at 410 nm, $\mu\text{M}^{-1}\text{cm}^{-1}$ )
ThT	0.0353
diThT-PEG2	0.0458
diThT-PEG3	0.0437
diThT-PEG4	0.0483
diThT-PEG5	0.0417
diThT-PEG2-HT	0.0449
diThT-C5	0.0446
MeO-ThT	0.0337
PEG2-ThT	0.0310
C5-ThT	0.0400

**Table 2-1.** Extinction coefficients of ThT derivatives

We then compared the fluorescence characteristics of the monomers and dimers. The results showed that the solutions of any of the ThT monomers or dimers alone are essentially non-fluorescent, while upon the addition of A $\beta$ 40, the monomers and dimers share similar fluorescence excitation curves (Figure 2-8), which indicates that the chromophores within the dimer design also behave in a similar fashion as the ThT monomers.



**Figure 2-8.** The fluorescence excitation spectra of ThT and diThT-PEG3: 1  $\mu$ M ThT/diThT-PEG3 + 2  $\mu$ M A $\beta$ 40 in 1  $\times$  buffer, excitation scan 320-460 nm in microplates (50  $\mu$ L/sample), slits 5 nm.

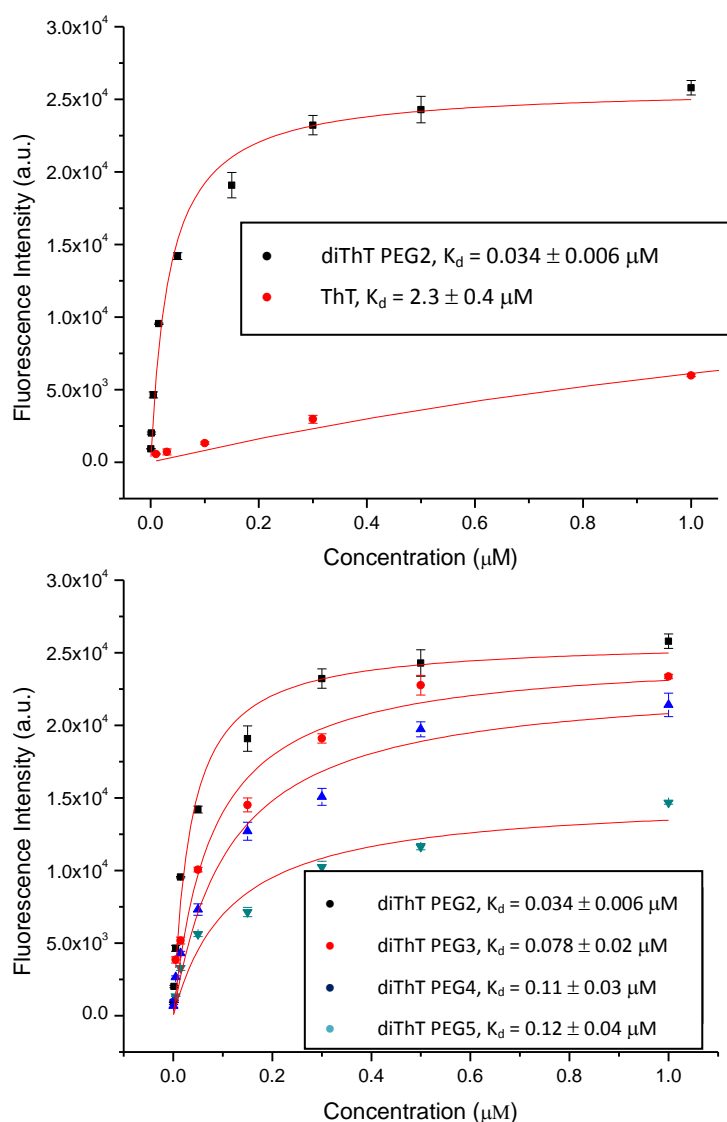
Based on the fact that the dimers and the monomers behave similarly, the amyloid-binding experiments of the diThTs were performed by using A $\beta$ 40 aggregates prepared *in vitro*. The A $\beta$ 40 peptide with the wild-type sequence was obtained through solid phase peptide synthesis. To ensure the purity of the peptide, double couplings were performed for the residues after beta-branched amino acids. The peptide was purified through reverse phase HPLC using a C18 column. LC-MS analysis confirmed the integrity of the peptide and showed the purity to be greater than 90%. For amyloid preparation, the purified peptide was treated following a reported protocol to yield fresh monomers, which were allowed to aggregate for seven days to completion. The resulting amyloid sample displays the characteristic beta-sheet signature in its circular dichroism spectrum. The amyloid sample was



further validated using the Congo Red binding assay<sup>[18]</sup>. As expected, amyloid binding induced a red shift in the absorption spectrum of Congo Red. The spectral shift allowed an estimation of the amyloid concentration, which agreed well with the A $\beta$ 40 monomer concentration used for amyloid preparation, assuring that all the A $\beta$ 40 peptides aggregated into amyloid fibrils. The details of the A $\beta$ 40 preparation and characterization are introduced in the experimental section.

The A $\beta$ 40 binding affinities of the ThT dimers were conveniently measured through titration experiments, in which an increasing amount of ThT derivatives were titrated into solutions of A $\beta$ 40 aggregates and the fluorescence emission spectra recorded. Plotting the fluorescence intensities at 485 nm against the diThT concentrations gave the binding curves shown in Figure 2-9. The  $K_d$  values were extracted by fitting these curves into a hyperbolic equation. For the monomeric ThT, a  $K_d$  value of 2.3  $\mu$ M was obtained, which agrees well with literature data<sup>[19]</sup>. All ThT dimers bind to the A $\beta$ 40 aggregates more tightly than the monomeric ThT: the  $K_d$  values ranges from 0.11 to 0.034  $\mu$ M, corresponding to an affinity improvement of 20 to 70 fold. Among the diThT-PEGn series, the amyloid binding affinities show a small (less than four fold) but reproducible variation following the order of PEG2 > PEG3 > PEG4  $\cong$  PEG5 (PEG1 was not tested due to its poor solubility). The most efficient dimer, diThT-PEG2, affords a  $K_d$  value of 34 nM, which is comparable to that of BTA-1 ( $K_d$  = 20 nM). Hence, diThT-PEG2 can be used as a potential substitute for BTA-1, with the added advantage of fluorescence “turn-on” upon binding onto A $\beta$ 40

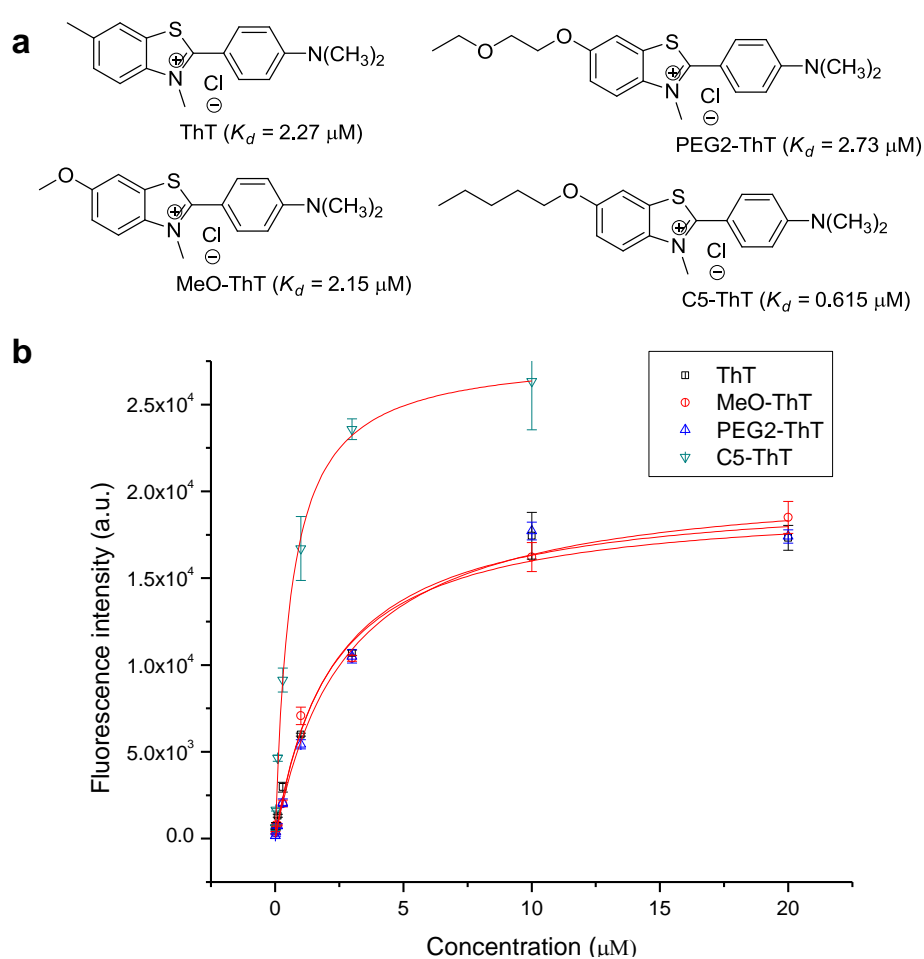
aggregates.



**Figure 2-9.** Binding curves of the diThTs to A $\beta$ 40 amyloid. Top: Comparison between the binding curves of diThT-PEG2 and ThT demonstrating the greatly enhanced binding affinity of the ThT dimer. Bottom: Binding curves of the diThT-PEGn series showing the slight variation of the  $K_d$  values according to the linker length.

In order to further investigate whether this observation can be explained by the increased entropy of longer linkers, we carried out the control experiment by

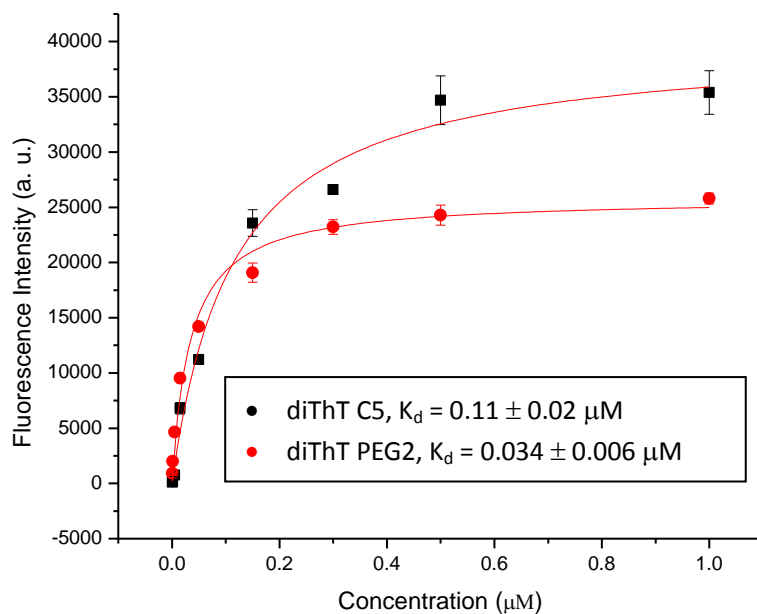
measuring the A $\beta$ 40 binding affinities of the monomeric ThT derivatives featuring different linkers (Figure 2-10). Firstly, 6-methoxy ThT (short as MeO-ThT) yields an essentially identical  $K_d$  value to that of ThT, consistent with previous reports that ThT binding to amyloid is insensitive to the substituents on 6-position. Another derivative PEG2-ThT installs a PEG2 mimic, the ethoxyethoxyl group, onto the 6-position of ThT. The A $\beta$  amyloid binding test affords overlapping titration curves for PEG2-ThT



**Figure 2-10.** (a) The control monomeric ThTs and the relative  $K_d$  values. (b) Binding curves of the monomeric ThTs to A $\beta$ 40 amyloid.

and ThT, indicating that the PEG2 moiety does not directly facilitate the amyloid

binding capabilities of the ThT monomer (Figure 2-10). Therefore, the improved binding affinity of diThT-PEG2 should be attributed to multivalency, which reduces the entropic cost of the binding event.



**Figure 2-11.** A $\beta$ 40 amyloid binding curves of diThT-C5 and diThT-PEG2

Interestingly, C5-ThT, made as a PEG2-ThT isostere, displays an improved binding affinity by 3 fold, presumably due to the hydrophobic nature of the pentyl substituent. In addition, amyloid bound C5-ThT also yields twice as strong fluorescence emission (Figure 2-10). Based on this observation, we also measured the A $\beta$ 40 binding affinity of diThT-C5, hoping that it may bind the amyloid even better than diThT-PEG2. However, it is disappointing that diThT-C5 failed to give improved binding affinity in comparison to diThT-PEG2. In fact, the  $K_d$  value of diThT-C5 increased by 3 fold (Figure 2-11). One possible explanation is the hydrophobic C5 linker tends to be buried into proteins, which introduces distortion to the binding

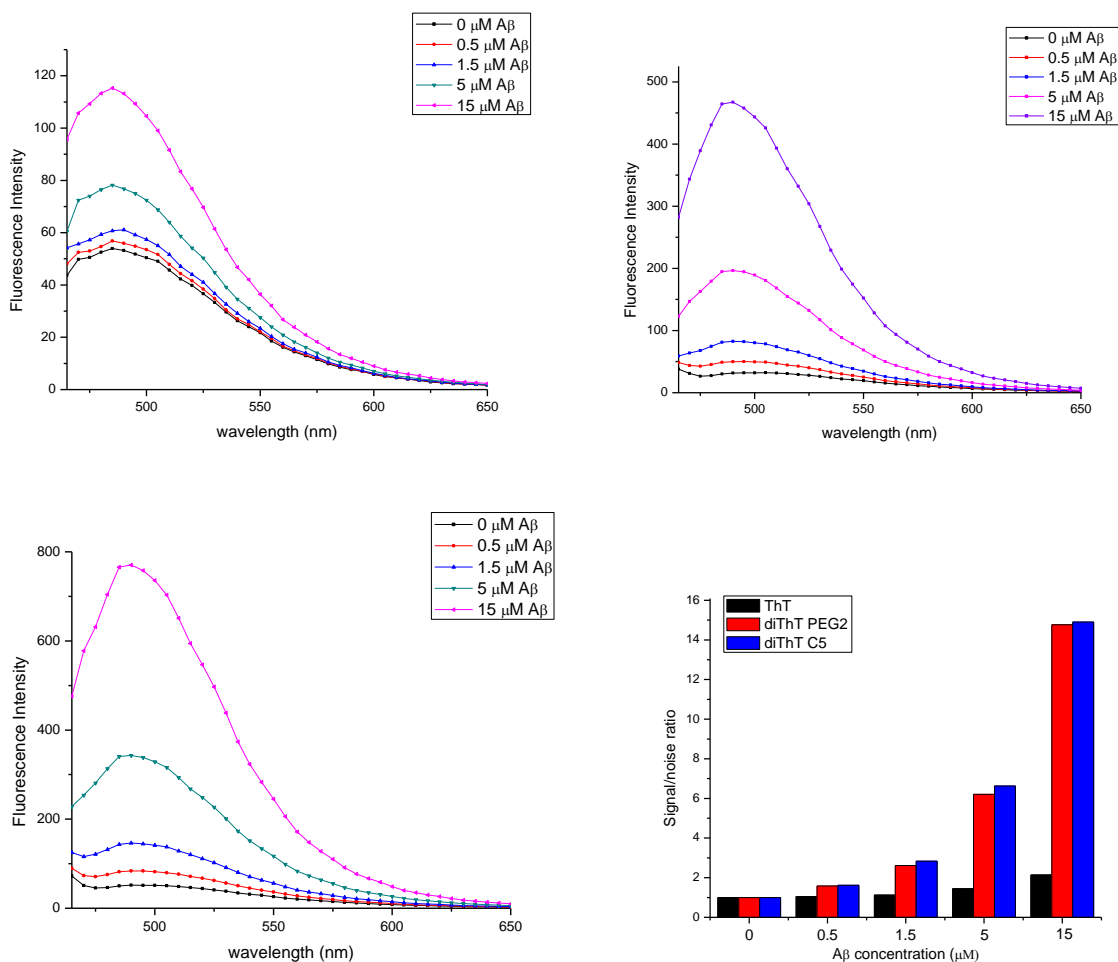
geometry of the ThT units. Although the C5 linkage does not give an improved amyloid binder, the result shows that the chemical nature of the linker is important. It remains to be seen if linkers of other chemical compositions may yield more favorable amyloid binders.

## **2.4 A $\beta$ 40 binding specificity measurements**

In addition to the binding affinity, the binding specificity toward A $\beta$  aggregates is essential for amyloid detection and quantification, because the samples for the diagnosis are usually taken from the patients' blood or tissues, and a good detector ought to be able to tease out the target from the complicated background noises. Based on our previous conclusion, the dimers increase the binding affinity by inserting into the adjacent  $\beta$ -sheet grooves, while the extended  $\beta$ -sheet structure is rare in regular proteins. Therefore, the ThT dimers should only have enhanced binding affinity to amyloid, but not to other proteins.

In order to prove this hypothesis, we evaluated the amyloid binding specificity of diThT-PEG2 and diThT-C5, and the results were compared with that of the monomeric ThT. Bovine serum, purchased from Sigma, was diluted 10 times (v/v) into a phosphate buffer and used as a blank control to evaluate the non-specific binding of the ThT derivatives. Small molecule solutions were prepared at 1  $\mu$ M ThT monomer concentration and the fluorescence emission spectra were collected, the

intensity of which are referred to as noise. We also recorded a series of fluorescence emission data with varying amount of A $\beta$ 40 aggregates added, which we have referred to as signals.



**Figure 2-12.** A $\beta$  amyloid binding specificity tests for ThT (top left), dimer PEG2 (top right), dimer C5 (bottom left) and the signal-to-noise ratio bar graph (bottom right) . Experimental conditions: 1  $\mu$ M ThT derivatives was mixed with various concentration of A $\beta$ 40 in 10% bovine serum 1  $\times$  buffer solution, excitation at 440 nm, slits 10 nm, emission scan 465-600 nm in microplate reader (150  $\mu$ L each sample).

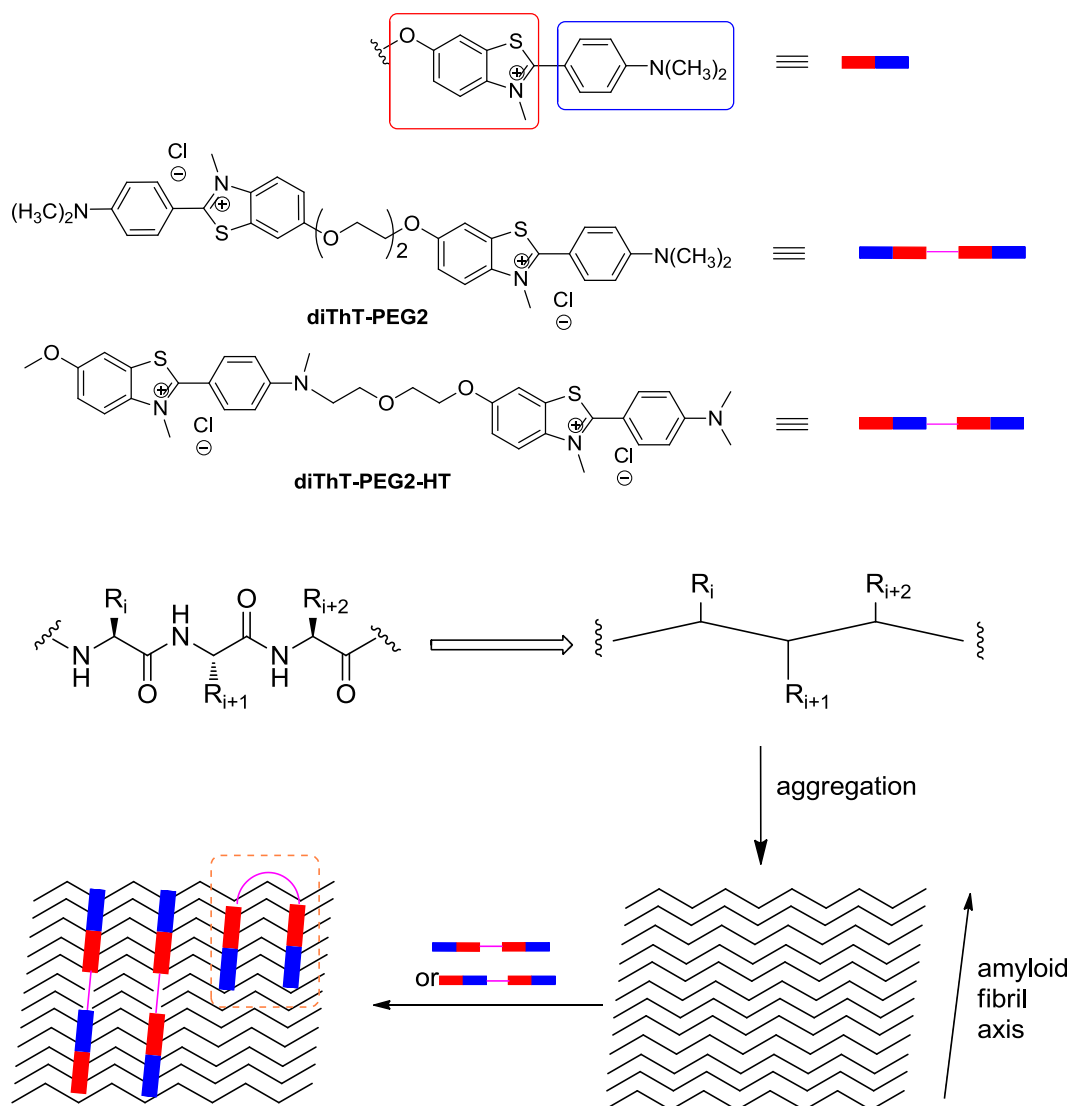
Shown in Figure 2-12 are the fluorescence emission curves and the signal-to-noise

ratios for each ThT derivatives. It is evident that the noise levels for diThT-PEG2 and diThT-C5 are similar to that of the ThT monomer, i.e. the dimer design does not render the molecules stickier to random proteins. For monomeric ThT, the signal-to-noise ratio reaches 2 when 15  $\mu\text{M}$  A $\beta$ 40 aggregates are added, while both dimers only need less than 1.5  $\mu\text{M}$  A $\beta$  40 aggregates to provide the same ratio. In other words, the ThT dimers are more sensitive for amyloid detection and effectively report the presence of amyloid at concentrations at least 10 times lower than what monomeric ThT is able to. Also, this result further validates our previous conclusion that the A $\beta$ 40 binding affinity enhancement of the ThT dimers is due to the reduced entropic penalty.

## **2.5 The investigation of the binding mechanism and the “head-to-tail” dimer**

So far we already demonstrated that our ThT dimers are capable of serving as more efficient and specific A $\beta$ 40 ligands. Now if we go back and look deeper into the proposed binding mechanism, in which the ThT core inserts into the grooves of  $\beta$ -sheet between  $i$  and  $i+2$  side chains (Figure 2-5), we discover that the linker of PEG2 spans a maximum distance of 8 Å when adopting a completely extended conformation, and this is not long enough to position the two ThT units in two adjacent grooves respectively. As a consequence, the two ThT units have to reside in the same groove with the opposite orientation pre-determined by the dimer design

(Figure 2-13).



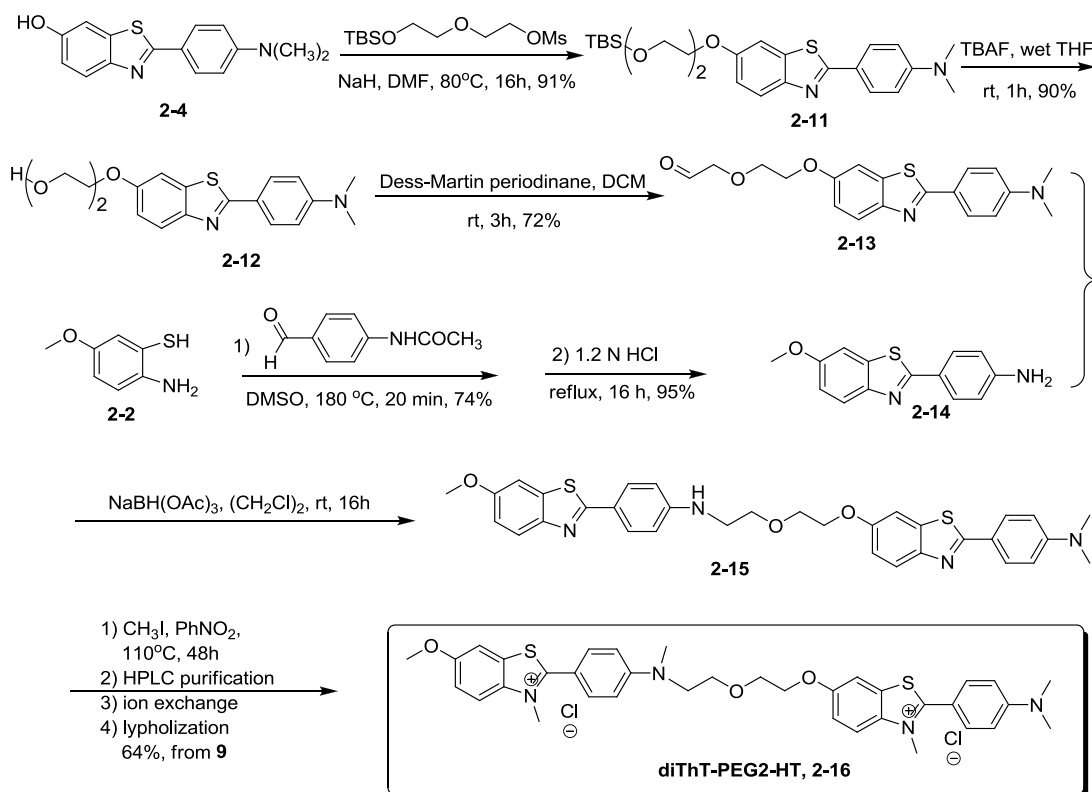
**Figure 2-13.** The possible binding modes for “head-to-head” and “head-to-tail” ThT dimers to amyloid aggregates

If we consider the benzothiazole heterocycle as the “head” of a ThT unit, while the dimethylaniline part as the “tail”, then all of our previous dimers are “head-to-head” dimers, and the two ThT units also reside in the “head-to-head” fashion. Now it is very natural to imagine if the ThT unit binds to the  $\beta$ -sheet groove with an orientation



preference, there has to be at least one ThT unit in the “head-to-head” dimers not being satisfied in terms of the orientation. In other words, a “head-to-tail” dimer, which can satisfy both the ThT units’ orientation preferences simultaneously, would have a better binding affinity compared with the “head-to-head” dimer bearing the same linker length (Figure 2-13).

In order to investigate whether this orientation preference actually exists, a head-to-tail dimer (diThT-PEG2-HT) was synthesized harboring a six-atom linker connecting the aniline N atom of one monomer to the 6-C of the other (Scheme 2-3).

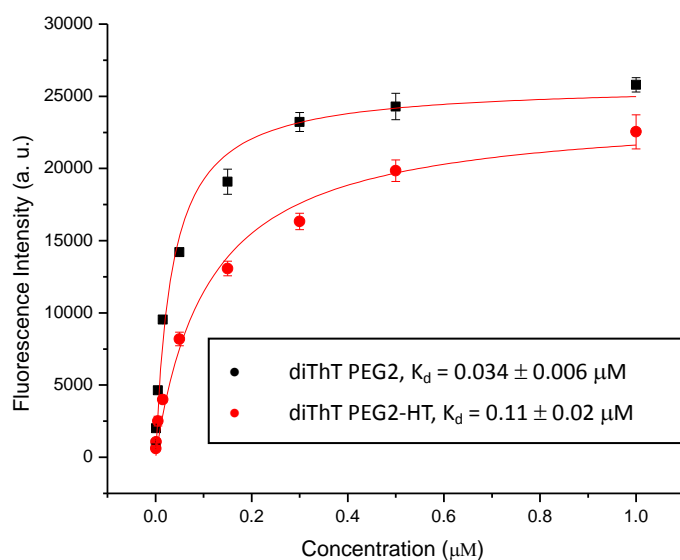


**Scheme 2-3.** Synthesis of the “head-to-tail” dimer, diThT-PEG2-HT

Starting from linking compound 2-4 with a mono TBS-protected diethylene glycol mesylate, the PEG2 linker was firstly coupled to the “head” of one unit (compound

**2-11**). After the deprotection of TBS-ether with TBAF in wet THF and subsequent oxidation with Dess-Martin periodinane, the un-linked the end of the linker was converted to an aldehyde (compound **2-13**). On the other side, 6-methoxyl-BTA-0 (compound **2-14**) was synthesized with the previous strategy from compound **2-2** and 4-acetylaminobenzaldehyde, followed by the hydrolysis of the acetanilide in 1.2 N HCl. Compounds **2-13** and **2-14** were linked together by reductive amination<sup>[20]</sup>, so the linker was coupled to the “tail” of another unit (compound **2-15**). Similar to the previous synthesis, the final step is the methylation with methyl iodide. During the methylation, the secondary aniline generated from the reductive amination step was also methylated to tertiary aniline, making the structure of diThT-PEG2-HT more consistent with the “head-to-head” dimers. The whole synthesis was finished in five steps with an overall yield of 37%.

However, fluorescence-based binding measurement shows that the head-to-tail dimer displays a slightly reduced (~ 3 fold) A $\beta$  amyloid binding affinity than diThT-PEG2 (Figure 2-14), which is not consistent with our hypothesis. Nevertheless, the head-to-tail dimer is still relatively favorable (by ~20 fold) compared to the ThT monomer in their amyloid binding affinities. This result indicates that ThT binding to A $\beta$ 40 amyloid is largely determined by the shape of the flat aromatic ligand. The binding site does not appear to discriminate the two different orientations of the ThT molecules; otherwise, the head-to-tail dimer should yield much improved binding affinities to A $\beta$  amyloid as previously mentioned.



**Figure 2-14.** A $\beta$ 40 amyloid binding curves of diThT-PEG2 and diThT-PEG2-HT

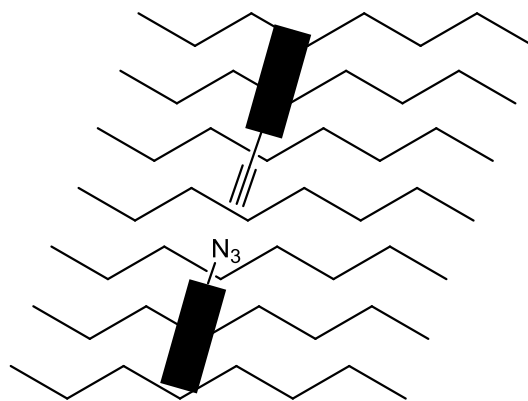
## 2.6 Conclusions and future work

In summary, we designed and evaluated a series of dimeric ThT derivatives as amyloid binding ligands. The dimer design affords tighter amyloid binding and retains the “light-up” feature in fluorescence emission upon associating with A $\beta$  aggregates, in the order of PEG2 > PEG3 > PEG4  $\cong$  PEG5, and the best dimer, diThT-PEG2, features the advantages of both ThT and BTA. The affinity gain is attributed to the reduced entropic cost of binding by the multivalency design. Nevertheless, the dimeric ThT derivatives are more specific in terms of detecting amyloid species out of complicated background components in the samples.

We also designed and synthesized the “head-to-tail” ThT dimer and investigated

the binding mode of ThT in the amyloid  $\beta$ -sheet grooves. The result shows that the “head-to-tail” dimer does not behave differently from the “head-to-head” dimer, indicating that there is no orientation preference in the binding process.

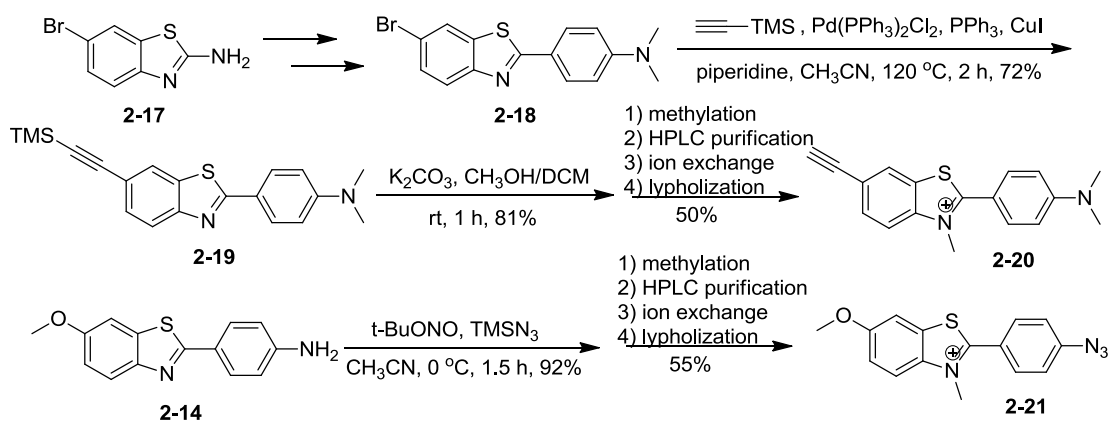
Our future work is currently focusing on two directions. The first one we are working on is the amyloid-templated dimerization. When two ThT/BTA monomers bind onto the same  $\beta$ -sheet groove of the amyloid, they can be considered to be “dragged”



**Figure 2-15.** Amyloid-templated reaction

together by the amyloid and their conformation is rigidified. Hence, if the two monomers bear functional groups reactive with each other but suffering from high activation energy barrier, the reaction could be templated and then accelerated by the amyloid (e.g. “click chemistry between azide and alkyne<sup>[23]</sup>”, Figure 2-15). Since the ThT/BTA-amyloid binding is not covalent and reversible, the templated reactions can turn over and over again with very low concentration of amyloid. By measuring the kinetics of the templated reactions, we can potentially develop highly sensitive amyloid detection methods.

We synthesized two groups of the monomers to carry out the templated reaction assay. One is the “click” pair, alkynyl-ThT (compound **2-20**) and ThT-azide (compound **2-21**). The synthesis is shown in Scheme 2-4.

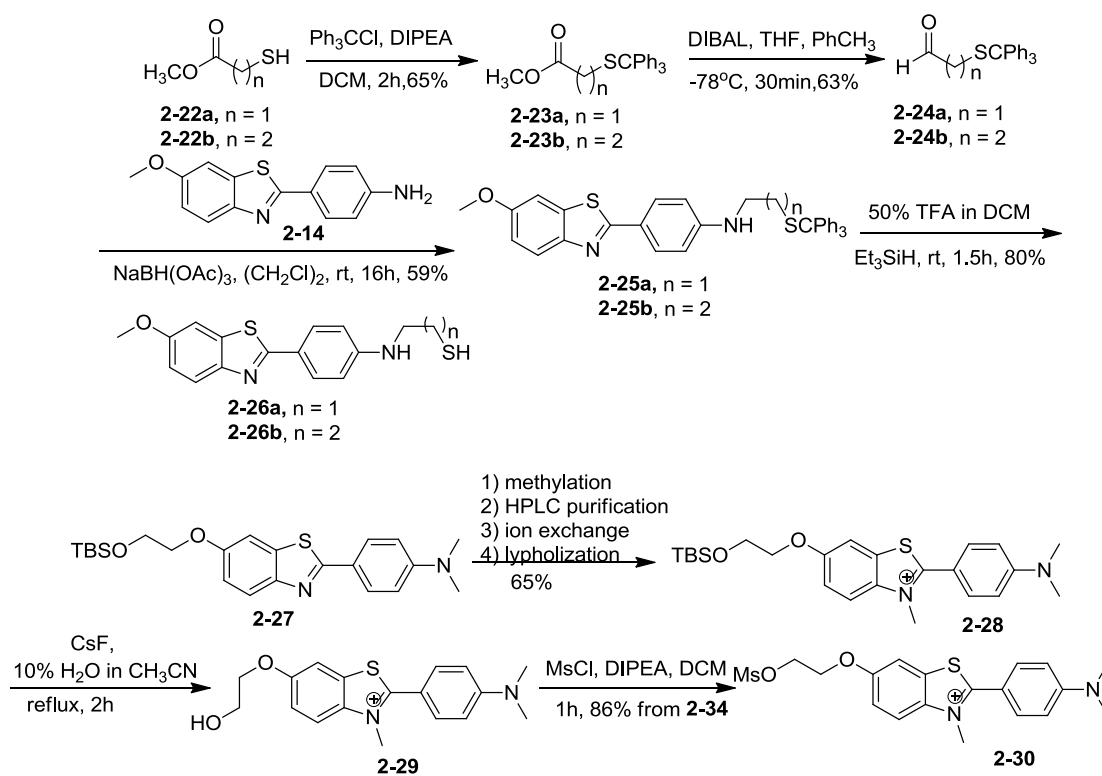


**Scheme 2-4.** Synthesis of alkynyl-ThT and ThT-azide

The key step for synthesizing alkynyl-ThT is the Sonogashira coupling reaction<sup>[24]</sup> between Br-ThT (compound **2-18**) and TMS-acetylene, followed by a TMS cleavage using potassium carbonate in methanol and DCM.

The other group for the templated reaction is the  $S_N2$  pair, BTA-1-SH (compound **2-26a**)/BTA-2-SH (compound **2-26b**) and MsO-ThT (compound **2-30**). As shown in Scheme 2-6, we synthesized the BTA-SH series from mercapto-acetic/propionic acid methyl ester (compound **2-22a, b**). The thiol was protected with the trityl group, and the ester was reduced to the aldehyde by DIBAL-H. The aldehyde (compound **2-24a, b**) was then linked onto **2-14** by reductive amination, and the trityl group was finally removed by TFA to yield compounds **2-26a, b**. We did not methylate these two compounds to ThT analogues in order to have some control compounds to figure out if the templated reaction will be more efficient with ThT or BTA derivatives. Importantly, we have kept the other site, MsO-ThT (compound **2-30**, Scheme 2-5) positively charged because, if both sites are neutral, the dimer generated from them

will be very poorly soluble in both water and organic solvents based on our previous experiences, and would not come off from the amyloid, hence stopping the turnover. The synthesis of MsO-ThT is slightly different, as we did the methylation first while leaving the mesylation as the final step. Another key point is the TBS cleavage in the ThT analogue, compound **2-28**. It was done by refluxing the substrate with cesium fluoride in acetonitrile/water 9:1. The advantage of this condition is that the solvents can be easily evaporated, and the residue is directly used in the following mesylation step without further purification.



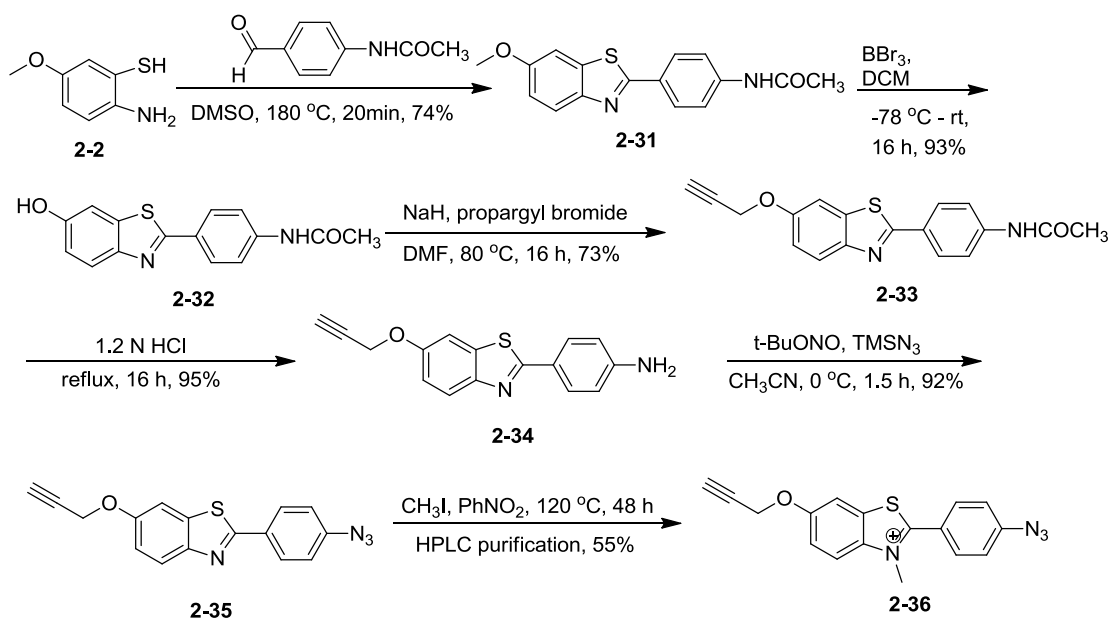
**Scheme 2-5.** Synthesis of BTA-1-SH/BTA-2-SH and MsO-ThT

Now we are working on the assays, trying to use LC-MS to detect the templated product and study the kinetics of the reaction. The optimization of these conditions is

currently underway.

The other direction is to elucidate the role of ThT in extending the nematodes lifespan, and we are collaborating with Dr. Weerapana on this project. In 2011, Lithgow and co-workers published that ThT can dramatically increase the lifespan of *Caenorhabditis elegans*, a family of nematodes<sup>[21]</sup>. Based on this interesting discovery, we plan to investigate the endogenous protein targets of ThT and the mechanism to extend lifespan. Generally speaking, identifying the cellular protein targets of ThT in *C. elegans* will determine if this bioactivity is governed by a specific aggregating protein target or a particular subcellular localization, thereby serving to partially illuminate the role of protein aggregation in the aging process and clarify the further development of similar amyloid-binding small molecules as anti-aging therapeutics.

A permanent linkage between ThT and its target is essential to increase the resolution of the analysis. To achieve this, our strategy is to synthesize ThT analogs with photo-crosslinking groups that will serve to covalently bind ThT analogues to their cellular protein targets in *C. elegans*. Also, incorporation of a bio-orthogonal click chemistry handle will enable enrichment for subsequent identification by mass spectrometry. Therefore, the synthesis will incorporate two key elements into the ThT scaffold: 1) a photoreactive group (e.g. azide), which upon brief exposure of *C. elegans* to UV light (5-10 min) promotes covalent labeling of proximal proteins; and 2) an alkyne functionality to facilitate click chemistry-mediated conjugation of reporter tags for detection (rhodamine) and affinity enrichment (biotin) of labeled proteins.



**Scheme 2-6.** Synthesis of propargoxy-ThT-azide

Propargoxy-ThT-azide (compound **2-36**, Scheme 2-6) was hence obtained over 6 steps with the similar strategies. The propargyl group was induced with propargyl bromide and the azido group was built up via diazonium chemistry with *tert*-butyl nitrite and TMS azide<sup>[22]</sup>. The *C. elegans* assay of propargoxy-ThT-azide and the following analyses are currently ongoing.

## Materials and Methods



## General Methods

All Fmoc-protected amino acids were purchased from Advanced Chemtech (Louisville, KY). All other chemical reagents and the bovine serum stock were purchased from Sigma-Aldrich (Milwaukee, WI). 96 well and 384 well microplates were purchased from Corning (Corning, NY). Peptide synthesis was carried out on a Tribute peptide synthesizer (Protein Technologies, Tucson, AZ).  $^1\text{H-NMR}$  and  $^{13}\text{C-NMR}$  data were collected on a Varian Gemini 400MHz NMR spectrometer. HR-MS data were generated by Boston College Mass-Spec facilities. Circular dichroism measurements were performed on an AVIV CD spectrometer (Aviv Biomedical Inc. Lakewood, NJ). The protein concentration of all samples used in this study was determined by measuring their absorption at 280 nm on a Lambda 25 UV-Vis spectrometer (PerkinElmer, Waltham, MA). The extinction coefficient  $\epsilon$  (280 nm) of A $\beta$ 40 was calculated using ExPASy ProtParam Tool (<http://ca.expasy.org/tools/protparam.html>) to be  $1280 \text{ M}^{-1}\text{cm}^{-1}$ . Fluorescence excitation and emission measurements in binding affinity tests were performed on a Fluorolog-3 fluorometer and a MicroMax 384 microwell plate reader (Jobin Yvon Inc. Edison, NJ). Fluorescence emission measurements in binding specificity tests were performed on a Spectra Max M5 microwell plate reader (Molecular Devices, Sunnyvale, CA).

## Synthesis of ThT derivatives

### 2-amino-5-methoxythiophenol (2-2)

2-amino-6-methoxybenzothiazole (10.0 g, 55.5 mmol) was suspended in 50% KOH aq (70 ml), and then ethylene glycol (13.5 ml) was added. The mixture was heated to reflux for 48h. Upon cooling to room temperature, toluene (100 ml) was added and then the mixture was neutralized with acetic acid (70 ml). The organic layer was separated, and the aqueous layer was extracted with toluene (2 ×100 ml). The combined organic layers were washed with water and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was then evaporated under vacuum, and the residue was washed with hexanes to give 6.6g (77%) of 2-amino-5-methoxythiophenol as a yellow solid. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ: 6.81 (d, *J* = 2.8 Hz, 1H), 6.70 (d, *J* = 8.8 Hz, 1H), 6.49 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.8 Hz, 1H), 5.60 (b, 3H), 3.62 (s, 3H).

### 2-(4'-(Dimethylamino)phenyl)-6-methoxybenzothiazole (2-3)

2-amino-5-methoxythiophenol (2-2) (9.85 g, 63.5 mmol) and 4-(dimethylamino)benzaldehyde (9.60 g, 64.4 mmol) were dissolved in DMSO (80 ml). The mixture was heated to 180 °C for 20 min. After cooled to room temperature, the reaction mixture was poured into 500 ml water. Ethyl acetate (100 ml) was then added and the whole mixture was stirred thoroughly.

2-(4'-(Dimethylamino)phenyl)-6-methoxybenzothiazole as pale-yellowish precipitate was isolated by vacuum filtration, washed with ethyl acetate (3 × 10 ml) and anhydrous ethanol (3 × 10 ml), and dried under vacuum. Yield: 6.75g (72%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.88 (d, *J* = 8.8 Hz, 2H), 7.84 (d, *J* = 8.8 Hz, 1H), 7.30 (d, *J* = 2.8 Hz, 1H), 7.01 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.8 Hz, 1H), 6.72 (d, *J* = 8.8 Hz, 2H), 3.86 (s, 3H), 3.03 (s, 6H).

### **2-(4'-(Dimethylamino)phenyl)-6-hydroxybenzothiazole (2-4)**

To a solution of 2-(4'-(Dimethylamino)phenyl)-6-methoxybenzothiazole (**2-3**) (1.0g, 3.5mmol) in anhydrous DCM (100ml) was added neat BBr<sub>3</sub> (4.6ml, 49mmol) dropwise at -78 °C under Ar protection. The reaction mixture was warmed to room temperature slowly and then stirred for 16 h. The reaction was quenched with water (100 ml) and the pH was adjusted to 4-7 with NaOH aq. 2-(4'-(Dimethylamino)phenyl)-6-hydroxybenzothiazole as yellowish precipitate was isolated by vacuum filtration, washed with water (3 × 10 ml), methanol (10 ml), DCM (10ml) and anhydrous ethanol (3 × 10 ml), and dried under vacuum. Yield: 0.89g (94%). <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ: 9.71 (s, 1H) 7.80 (d, *J* = 8.8 Hz, 2H), 7.72 (d, *J* = 8.8 Hz, 1H), 7.33 (d, *J* = 2.4 Hz, 1H), 6.91 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.4 Hz, 1H), 6.80 (d, *J* = 8.8 Hz, 2H), 3.01 (s, 6H).

### **General procedure 2A: BTA-linker coupling**

**4,4'-(6,6'-(3,6,9,12-tetraoxatetradecane-1,14-diylbis(oxy))bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6e; diBTA-PEG5)**

To a suspension of NaH (10.5 mg, 0.44 mmol) and anhydrous DMF (4 ml) in a pressure vial was added 2-(4'-(Dimethylamino)phenyl)-6-hydroxybenzothiazole (**3**) (120 mg, 0.44 mmol) in one portion. The mixture was stirred for 5 min, and hydrogen gas was released. Then 2,2'-(ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl)dimethanesulfonate (**2-5e**) (80 mg, 0.20 mmol) was added, and the resulting mixture was stirred for 16 h at 90 °C in the screw-capped vial. The reaction mixture was then quenched with saturated NH<sub>4</sub>Cl aq. (5 ml), poured to water (30 ml), and extracted with DCM (3 × 20 ml). After the solvent was evaporated, the product was isolated with silica gel column (EtOAc/hexane=3:1). Yield: 123mg (82%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.88 (d, *J* = 8.8 Hz, 4H), 7.84 (d, *J* = 8.8 Hz, 2H), 7.31 (d, *J* = 2.4 Hz, 2H) 7.04 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.72 (d, *J* = 8.8 Hz, 4H), 4.17 (t, *J* = 4.8 Hz, 4H), 3.87 (t, *J* = 4.8 Hz, 4H), 3.68-3.74 (m, 8H), 3.67 (s, 4H), 3.04 (s, 12H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ: 166.7, 156.4, 152.1, 149.3, 135.9, 128.7, 122.9, 121.9, 115.7, 111.9, 105.7, 71.1, 70.9, 70.0, 68.4, 40.4. HRMS (ESI+): *m/z* calculated for C<sub>40</sub>H<sub>47</sub>N<sub>4</sub>O<sub>6</sub>S<sub>2</sub> [M]<sup>+</sup>, 743.29370; found 743.29457.

**4,4'-(6,6'-(ethane-1,2-diylbis(oxy))bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6a; diBTA-PEG1)**

The product was directly isolated by vacuum filtration after the reaction was

quenched, washed with DMF (10 ml), water (3 × 10 ml), ethanol (3 × 10 ml) and ether (3 × 10 ml), dried under vacuum, and directly used in the followed step. Yield: 87%. The product was insoluble in all common solvents and improper for NMR analysis. HRMS (ESI+):  $m/z$  calculated for  $C_{32}H_{31}N_4O_2S_2 [M]^+$ , 567.18884; found 567.18879.

**4,4'-(6,6'-(2,2'-oxybis(ethane-2,1-diyl)bis(oxy))bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6b; diBTA-PEG2)**

The product was directly isolated by vacuum filtration after the reaction was quenched, washed with DMF (10 ml), water (3 × 10 ml), ethanol (3 × 10 ml) and ether (3 × 10 ml), dried under vacuum, and directly used in the followed step. Yield: 76%. The product was only slightly soluble in  $CDCl_3$  and improper for  $^{13}C$  NMR analysis.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$ : 7.87 (d,  $J = 8.8$  Hz, 4H), 7.84 (d,  $J = 8.8$  Hz, 2H), 7.31 (d,  $J = 2.8$  Hz, 2H) 7.04 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.4$  Hz, 2H), 6.70 (d,  $J = 8.8$  Hz, 4H), 4.22 (t,  $J = 4.8$  Hz, 4H), 3.97 (t,  $J = 4.8$  Hz, 4H), 3.02 (s, 12H). HRMS (ESI+):  $m/z$  calculated for  $C_{32}H_{31}N_4O_2S_2 [M]^+$ , 611.2151; found 611.2158.

**4,4'-(6,6'-(2,2'-(ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(oxy)bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6c; diBTA-PEG3)**

The product was directly isolated by vacuum filtration after the reaction was quenched, washed with DMF (10 ml), water (3 × 10 ml), ethanol (3 × 10 ml) and

ether (3 × 10 ml), dried under vacuum, and directly used in the followed step. Yield: 70%. The product was only slightly soluble in CDCl<sub>3</sub> and improper for <sup>13</sup>C NMR analysis. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ: 7.86 (d, *J* = 9.0 Hz, 4H), 7.82 (d, *J* = 8.4 Hz, 2H), 7.29 (d, *J* = 2.4 Hz, 2H) 7.03 (dd, *J*<sub>1</sub> = 9.0 Hz, *J*<sub>2</sub> = 2.7 Hz, 2H), 6.70 (d, *J* = 8.7 Hz, 4H), 4.17 (t, *J* = 4.8 Hz, 4H), 3.89 (t, *J* = 4.8 Hz, 4H), 3.77 (s, 4H), 3.02 (s, 12H). HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>36</sub>H<sub>39</sub>N<sub>4</sub>O<sub>4</sub>S<sub>2</sub> [M]<sup>+</sup>, 655.24127; found 655.23766.

**4,4'-(6,6'-(2,2'-(2,2'-oxybis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))bis(oxy)bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6d; diBTA-PEG4)**

The product was directly isolated by vacuum filtration after the reaction was quenched, recrystallized with EtOAc, and directly used in the followed step. Yield: 56%. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.88 (d, *J* = 8.8 Hz, 4H), 7.84 (d, *J* = 8.8 Hz, 2H), 7.30 (d, *J* = 2.4 Hz, 2H) 7.05 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.72 (d, *J* = 8.8 Hz, 4H), 4.17 (t, *J* = 4.8 Hz, 4H), 3.89 (t, *J* = 4.8 Hz, 4H), 3.71-3.76 (m, 8H), 3.04 (s, 12H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ: 156.6, 152.1, 128.9, 122.7, 115.6, 112.0, 105.7, 71.1, 70.0, 68.4, 40.5. HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>40</sub>H<sub>47</sub>N<sub>4</sub>O<sub>6</sub>S<sub>2</sub> [M]<sup>+</sup>, 699.2675; found 699.2656.

**4,4'-(6,6'-(pentane-1,5-diylbis(oxy))bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (2-6f; diBTA-C5)**

The product was directly isolated by vacuum filtration after the reaction was

quenched, washed with DMF (10 ml), water (3 × 10 ml), ethanol (3 × 10 ml) and ether (3 × 10 ml), dried under vacuum, and directly used in the followed step. Yield: 91%. The product was insoluble in all common solvents and improper for NMR analysis. HRMS (ESI+):  $m/z$  calculated for  $C_{38}H_{43}N_4O_2S_2$   $[M]^+$ , 651.28274; found 651.28372.

**N,N-dimethyl-4-(6-(pentyloxy)benzo[d]thiazol-2-yl)aniline (2-9a; C5-BTA)**

The product was isolated with silica gel column (EtOAc/hexane=1:4). Yield: 75%.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$ : 7.81 (d,  $J = 9.2$  Hz, 2H), 7.77 (d,  $J = 8.8$  Hz, 1H), 7.20 (d,  $J = 2.8$  Hz, 1H) 6.93 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.8$  Hz, 1H), 6.63 (d,  $J = 9.2$  Hz, 2H), 3.91 (t,  $J = 4.8$  Hz, 2H), 2.93 (s, 6H), 1.73 (m, 2H), 1.30-1.37 (m, 4H), 0.86 (t,  $J = 4.8$  Hz, 3H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$ : 166.4, 156.7, 152.0, 148.9, 135.9, 128.6, 122.8, 121.8, 115.5, 111.9, 105.2, 68.8, 40.4, 29.2, 28.4, 22.7, 14.3. HRMS (ESI+):  $m/z$  calculated for  $C_{20}H_{25}N_2OS$   $[M]^+$ , 341.1687; found 341.1689.

**4-(6-(2-ethoxyethoxy)benzo[d]thiazol-2-yl)-N,N-dimethylaniline (2-9b;**

**PEG2-BTA)**

The product was isolated with silica gel column (EtOAc/hexane=1:3). Yield: 76%.  $^1H$  NMR (400 MHz,  $CD_2Cl_2$ )  $\delta$ : 7.81 (d,  $J = 9.2$  Hz, 2H), 7.74 (d,  $J = 8.8$  Hz, 1H), 7.27 (d,  $J = 2.4$  Hz, 1H) 6.97 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.8$  Hz, 1H), 6.66 (d,  $J = 9.2$  Hz, 2H), 4.07 (m, 2H), 3.71 (m, 2H), 3.51 (q,  $J = 7.2$  Hz, 2H), 2.95 (s, 6H), 1.16 (t,  $J = 7.2$

Hz, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$ : 166.3, 156.5, 152.1, 149.2, 135.9, 128.4, 122.7, 121.6, 115.4, 111.8, 105.4, 69.1, 68.4, 66.9, 40.1, 15.2. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{19}\text{H}_{23}\text{N}_2\text{O}_2\text{S}$   $[\text{M}]^+$ , 343.1480; found 343.1483.

**4-(6-(2-(2-(tert-butyldimethylsilyloxy)ethoxy)ethoxy)benzo[d]thiazol-2-yl)-N,N-dimethylaniline (2-11)**

The product was isolated with silica gel column (EtOAc/hexane=1:3). Yield: 91%.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.87 (d,  $J = 8.8$  Hz, 2H), 7.84 (d,  $J = 9.2$  Hz, 1H), 7.29 (d,  $J = 2.8$  Hz, 1H) 7.04 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.4$  Hz, 1H), 6.68 (d,  $J = 9.2$  Hz, 2H), 4.14 (t,  $J = 4.8$  Hz, 2H), 3.86 (t,  $J = 4.8$  Hz, 2H), 3.79 (t,  $J = 5.2$  Hz, 2H), 3.63 (t,  $J = 5.2$  Hz, 2H), 2.98 (s, 6H), 0.88 (s, 9H), 0.07 (s, 6H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 166.5, 156.3, 151.9, 149.1, 135.8, 128.5, 122.7, 121.6, 115.5, 111.8, 105.4, 73.0, 69.9, 68.3, 62.9, 40.2, 26.1, 18.5, -5.1. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{25}\text{H}_{37}\text{N}_2\text{O}_3\text{SiS}$   $[\text{M}]^+$ , 473.2294; found 473.2295.

**2-(2-(2-(4-(dimethylamino)phenyl)benzo[d]thiazol-6-yloxy)ethoxy)ethanol (2-12)**

4-(6-(2-(2-(tert-butyldimethylsilyloxy)ethoxy)ethoxy)benzo[d]thiazol-2-yl)-N,N-dimethylaniline (**2-11**) (302 mg, 0.64 mmol) was dissolved in 1 M TBAF in THF (5% water) (1 ml, 1.0 mmol) and stirred for 1 h at rt. The reaction was quenched with water, extracted with DCM (3  $\times$  15 ml), and the combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . The product was isolated with silica gel column (EtOAc/hexane=1:5).



Yield: 206 mg (90%).  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-}d_6$ )  $\delta$ : 7.83 (d,  $J = 9.2$  Hz, 2H), 7.81 (d,  $J = 9.2$  Hz, 1H), 7.64 (d,  $J = 2.4$  Hz, 1H), 7.07 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 1H), 6.81 (d,  $J = 8.8$  Hz, 2H), 4.64 (t,  $J = 9.2$  Hz, 1H), 4.17 (m, 2H), 3.78 (m, 2H), 3.52 (m, 4H) 3.01 (s, 6H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{DMSO-}d_6$ )  $\delta$ : 165.3, 155.9, 151.8, 148.2, 135.1, 128.0, 122.3, 120.4, 115.5, 111.8, 105.6, 72.4, 68.8, 67.8, 60.2. High-resolution mass spectrometry (HRMS) (ESI+):  $m/z$  calculated for  $\text{C}_{19}\text{H}_{23}\text{N}_2\text{O}_3\text{S}_1$   $[\text{M}]^+$ , 359.14294; found 359.14300.

**2-(2-(2-(4-(dimethylamino)phenyl)benzo[d]thiazol-6-yloxy)ethoxy)acetaldehyde (2-13)**

To a solution of 2-(2-(2-(4-(dimethylamino)phenyl)benzo[d]thiazol-6-yloxy)ethoxy)ethanol (**2-12**) (88 mg, 0.24 mmol) in DCM (5 ml) was added Dess-Martin periodinane (126 mg, 0.29 mmol) in one portion. The mixture was stirred at rt for 2 h. The reaction was quenched with sat.  $\text{NaHCO}_3$  and  $\text{Na}_2\text{S}_2\text{O}_3$ , the organic layer was separated, and the aqueous layer was extracted with DCM (2  $\times$  15 ml). The combined organic layers were washed with brine and dried over  $\text{Na}_2\text{SO}_4$ . After the solvent was evaporated, the product was isolated with silica gel column (EtOAc/hexane=3:1). Yield: 63 mg (72%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 9.75 (d,  $J = 0.4$  Hz, 1H), 7.89 (d,  $J = 8.8$  Hz, 2H), 7.86 (d,  $J = 8.8$  Hz, 1H), 7.32 (d,  $J = 2.4$  Hz, 1H), 7.04 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.4$  Hz, 1H), 6.72 (d,  $J = 8.8$  Hz, 2H), 4.25 (s, 2H), 4.22 (t,  $J = 4.8$  Hz, 2H), 3.95 (t,  $J = 4.8$  Hz, 2H),

3.03 (s, 6H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 200.5, 156.1, 152.1, 149.4, 135.9, 128.7, 122.3, 121.7, 115.4, 111.9, 105.6, 76.9, 70.6, 68.4, 40.4, 29.9. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{19}\text{H}_{21}\text{N}_2\text{O}_3\text{S}_1$   $[\text{M}]^+$ , 357.12729; found 357.12885.

**4-(6-(2-(2-(4-(6-methoxybenzo[d]thiazol-2-yl)phenylamino)ethoxy)ethoxy)benzo[d]thiazol-2-yl)-N,N-dimethylaniline (2-15)**

To a solution of 2-(2-(2-(4-(dimethylamino)phenyl)benzo[d]thiazol-6-yloxy)ethoxy)acetaldehyde (**2-13**) (44mg, 0.13 mmol) and 4-(6-methoxybenzo[d]thiazol-2-yl)aniline (**2-14**) (40 mg, 0.16 mmol) in anhydrous 1,2-dichloroethane (10 ml) was added  $\text{NaBH}(\text{OAc})_3$  (40 mg, 0.19 mmol). The mixture was stirred at rt for 24 h. Precipitate was generated gradually. The reaction was quenched with sat.  $\text{NaHCO}_3$ , and the whole mixture was filtered. The filtrate (product) was washed with water, MeOH and ether, dried under vacuum, and directly used in the followed step without further purification.

**General procedure 2B: methylation of BTA derivatives.**

**6,6'-(3,6,9,12-tetraoxatetradecane-1,14-diylbis(oxy))bis(2-(4-(dimethylamino)phenyl)-3-methylbenzo[d]thiazol-3-ium) chloride (2-7e; diThT-PEG5)**

4,4'-(6,6'-(3,6,9,12-tetraoxatetradecane-1,14-diylbis(oxy))bis(benzo[d]thiazole-6,2-diyl))bis(N,N-dimethylaniline) (**2-6e**) (56 mg, 0.075 mmol), MeI (94  $\mu\text{l}$ , 1.5 mmol) and nitrobenzene (2ml) were mixed in a pressure vial and stirred 48 h at 110  $^\circ\text{C}$ . The

solvent was evaporated under vacuum, and the residue was dissolved in DMSO (5 ml), filtered, and purified on HPLC. The isolated product was anion-exchanged with 1 M HCl in methanol (5 × 20 ml) and lyophilized. Yellow solid was finally obtained. Yield: 44 mg (69%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN) δ: 7.88 (d, *J* = 9.6 Hz, 2H), 7.69 (d, *J* = 9.2 Hz, 4H), 7.64 (d, *J* = 2.4 Hz, 2H), 7.39 (dd, *J*<sub>1</sub> = 9.2 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.90 (d, *J* = 9.2 Hz, 4H), 4.22 (t, *J* = 4.8 Hz, 4H), 4.13 (s, 6H), 3.84 (t, *J* = 4.8 Hz, 4H), 3.58-3.66 (m, 8H), 3.57 (s, 4H), 3.12 (s, 12H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>CN) δ: 173.5, 160.1, 155.5, 138.5, 133.4, 131.3, 120.2, 113.3, 112.2, 108.2, 71.8, 71.6, 70.4, 70.1, 40.8, 39.6. HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>42</sub>H<sub>52</sub>N<sub>4</sub>O<sub>6</sub>S<sub>2</sub> [M]<sup>+</sup>, 386.1658; found 386.1668.

**6,6'-(ethane-1,2-diylbis(oxy))bis(2-(4-(dimethylamino)phenyl)-3-methylbenzo[d]thiazol-3-ium) chloride (2-7a; diThT-PEG1)**

Yield: 63%. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN) δ: 7.93 (d, *J* = 9.2 Hz, 2H), 7.74 (d, *J* = 2.4 Hz, 2H), 7.71 (d, *J* = 8.8 Hz, 4H), 7.45 (dd, *J*<sub>1</sub> = 9.2 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.93 (d, *J* = 8.8 Hz, 4H), 4.54 (s, 4H), 4.16 (s, 6H), 3.13 (s, 12H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>CN) δ: 173.8, 159.8, 155.5, 138.9, 133.5, 131.4, 120.2, 113.4, 112.3, 108.4, 69.0, 40.8, 39.6. HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>34</sub>H<sub>36</sub>N<sub>4</sub>O<sub>2</sub>S<sub>2</sub> [M]<sup>+</sup>, 298.1134; found 298.1140.

**6,6'-(2,2'-oxybis(ethane-2,1-diyl)bis(oxy))bis(2-(4-(dimethylamino)phenyl)-3-met**

**hylbenzo[d]thiazol-3-ium) chloride (2-7b; diThT-PEG2)**

Yield: 63%. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN) δ: 7.88 (d, *J* = 9.2 Hz, 2H), 7.67 (d, *J* = 9.2 Hz, 4H), 7.64 (d, *J* = 2.4 Hz, 2H), 7.40 (dd, *J*<sub>1</sub> = 9.2 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.90 (d, *J* = 9.6 Hz, 4H), 4.28 (t, *J* = 4.4 Hz, 4H), 4.13 (s, 6H), 3.94 (t, *J* = 4.4 Hz, 4H), 3.12 (s, 12H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>CN) δ: 173.5, 160.1, 155.4, 138.6, 133.4, 131.2, 120.1, 113.3, 112.2, 108.2, 70.7, 70.1, 40.8, 39.6. HRMS (ESI+): *m/z* calculated for C<sub>36</sub>H<sub>40</sub>N<sub>4</sub>O<sub>3</sub>S<sub>2</sub> [M]<sup>+</sup>, 320.1265; found 320.1275.

**6,6'-(2,2'-(ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(oxy)bis(2-(4-(dimethyl amino)phenyl)-3-methylbenzo[d]thiazol-3-ium) chloride (2-7c; diThT-PEG3)**

Yield: 59%. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN) δ: 7.88 (d, *J* = 9.2 Hz, 2H), 7.66 (d, *J* = 8.8 Hz, 4H), 7.64 (d, *J* = 2.4 Hz, 2H), 7.39 (dd, *J*<sub>1</sub> = 9.2 Hz, *J*<sub>2</sub> = 2.4 Hz, 2H), 6.89 (d, *J* = 9.2 Hz, 4H), 4.22 (t, *J* = 4.4 Hz, 4H), 4.13 (s, 6H), 3.86 (t, *J* = 4.8 Hz, 4H), 3.70 (s, 4H), 3.11 (s, 12H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>CN) δ: 172.2, 158.9, 154.2, 137.3, 132.2, 130.0, 119.0, 112.1, 111.0, 106.9, 70.6, 69.2, 68.9, 39.6, 38.3. HRMS (ESI+): *m/z* calculated for C<sub>38</sub>H<sub>44</sub>N<sub>4</sub>O<sub>4</sub>S<sub>2</sub> [M]<sup>+</sup>, 342.1396; found 342.1405.

**6,6'-(2,2'-(2,2'-oxybis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))bis(oxy)bis(2-(4-(dimethylamino)phenyl)-3-methylbenzo[d]thiazol-3-ium) chloride (2-7d; diThT-PEG4)**

Yield: 62%. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ: 8.14 (d, *J* = 9.2 Hz, 2H), 7.97 (d, *J*

= 2.4 Hz, 2H), 7.77 (d,  $J = 9.2$  Hz, 4H), 7.47 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 2H), 6.96 (d,  $J = 9.2$  Hz, 4H), 4.22 (m, 4H), 4.20 (s, 6H), 3.82 (m, 4H), 3.58-3.63 (m, 8H), 3.12 (s, 12H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$ : 171.6, 158.4, 154.0, 137.3, 132.4, 130.0, 119.0, 118.1, 112.4, 111.2, 107.3, 70.3, 69.1, 68.8, 38.5. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{40}\text{H}_{48}\text{N}_4\text{O}_5\text{S}_2$   $[\text{M}]^+$ , 364.1527; found 364.1534.

**6,6'-(pentane-1,5-diylbis(oxy))bis(2-(4-(dimethylamino)phenyl)-3-methylbenzo[d]thiazol-3-ium) chloride (2-7f; diThT-C5)**

Yield: 67%.  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$ : 8.15 (d,  $J = 9.6$  Hz, 2H), 7.97 (d,  $J = 2.8$  Hz, 2H), 7.79 (d,  $J = 8.8$  Hz, 4H), 7.47 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.4$  Hz, 2H), 6.97 (d,  $J = 9.2$  Hz, 4H), 4.21 (s, 6H), 4.17 (t,  $J = 6.4$  Hz, 4H), 3.12 (s, 12H), 1.90 (m, 4H), 1.65 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{CN}$ )  $\delta$ : 160.3, 155.4, 138.4, 133.4, 131.4, 120.2, 113.3, 112.3, 107.9, 70.4, 40.8, 39.5, 29.8, 23.6. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{37}\text{H}_{42}\text{N}_4\text{O}_2\text{S}_2$   $[\text{M}]^+$ , 319.1369; found 319.1378.

**2-(4-(dimethylamino)phenyl)-3-methyl-6-(pentyloxy)benzo[d]thiazol-3-ium chloride (2-10a; C5-ThT)**

Yield: 71%.  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$ : 8.14 (d,  $J = 9.6$  Hz, 1H), 7.97 (d,  $J = 2.4$  Hz, 1H), 7.79 (d,  $J = 8.8$  Hz, 2H), 7.44 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 1H), 6.96 (d,  $J = 9.2$  Hz, 2H), 4.21 (s, 3H), 4.11 (t,  $J = 6.4$  Hz, 2H), 3.11 (s, 6H), 1.79 (m, 2H), 1.40 (m, 4H), 0.91 (t,  $J = 6.8$  Hz, 3H).  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$ : 171.1,

158.1, 153.4, 136.7, 131.9, 129.7, 118.4, 117.7, 111.9, 111.0, 107.0, 68.6, 38.2, 28.1, 27.6, 21.9, 13.9. HRMS (ESI+):  $m/z$  calculated for  $C_{21}H_{27}N_2OS$   $[M]^+$ , 355.1838; found 355.1837.

**2-(4-(dimethylamino)phenyl)-6-(2-ethoxyethoxy)-3-methylbenzo[d]thiazol-3-ium chloride (2-10b; PEG2-ThT)**

Yield: 70%.  $^1H$  NMR (400 MHz,  $CD_3CN$ )  $\delta$ : 7.89 (d,  $J = 9.2$  Hz, 1H), 7.70 (d,  $J = 9.2$  Hz, 2H), 7.66 (d,  $J = 2.0$  Hz, 1H), 7.41 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 1H), 6.91 (d,  $J = 9.6$  Hz, 2H), 4.23 (m, 2H), 4.15 (s, 3H), 3.80 (m, 2H), 3.56 (q,  $J = 8.8$  Hz, 2H), 3.12 (s, 6H), 1.18 (t,  $J = 6.8$  Hz, 3H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$ : 171.6, 159.3, 154.2, 136.9, 132.1, 129.7, 119.7, 117.3, 112.4, 110.8, 106.7, 68.9, 68.7, 67.0, 58.8, 40.2, 38.6, 23.9. HRMS (ESI+):  $m/z$  calculated for  $C_{20}H_{25}N_2O_2S$   $[M]^+$ , 357.1631; found 357.1642.

**2-(4-(dimethylamino)phenyl)-6-(2-(2-((4-(6-methoxy-3-methylbenzo[d]thiazol-3-ium-2-yl)phenyl)(methyl)amino)ethoxy)ethoxy)-3-methylbenzo[d]thiazol-3-ium chloride (2-16, diThT-PEG2-HT)**

Yield: 64% from (2-14).  $^1H$  NMR (400 MHz,  $DMSO-d_6$ )  $\delta$ : 8.15 (m, 2H), 7.94 (m, 2H), 7.78 (m, 4H), 7.46 (m, 2H), 6.95 (m, 4H), 4.20-4.24 (m, 7H), 3.90-3.91 (m, 3H), 3.85 (m, 2H), 3.74 (s, 3H), 3.11 (m, 11H).  $^{13}C$  NMR (100 MHz,  $DMSO-d_6$ )  $\delta$ : 171.1, 171.0, 158.8, 158.0, 153.4, 152.7, 136.9, 136.8, 131.9, 129.7, 129.6, 118.3, 118.2,

117.8, 112.0, 111.9, 111.0, 110.9, 107.2, 106.4, 68.8, 68.2, 67.9, 56.2, 56.1, 51.1, 38.2.

HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>40</sub>H<sub>47</sub>N<sub>4</sub>O<sub>6</sub>S<sub>2</sub> [M]<sup>+</sup>, 320.1265; found 320.1259.

### **N,N-dimethyl-4-(6-((trimethylsilyl)ethynyl)benzo[d]thiazol-2-yl)aniline (2-19)**

A mixture of 4-(6-bromobenzo[d]thiazol-2-yl)-N,N-dimethylaniline (53 mg, 0.16 mmol), Pd(PPh<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub> (5.6 mg, 8 μmol), CuI (3.0 mg, 16 μmol), PPh<sub>3</sub> (8.3 mg, 32 μmol), trimethylsilylacetylene (27 μl, 0.19 mmol) and piperidine (22 μl, 0.22 mmol) in anhydrous DMF (2 ml) was stirred in a pressure vial under N<sub>2</sub> protection at 120 °C for 2 h. After cooled to r.t., the resulting solution was treated with Et<sub>2</sub>O (20 ml). Grey precipitate was formed and removed by filtration. The supernatant was sequentially washed with 2 N HCl, sat. NaHCO<sub>3</sub> aq. and brine, and then dried over Na<sub>2</sub>SO<sub>4</sub>. The product was purified with silica gel column. Yield: 40 mg (72%). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ: 7.93 (d, *J* = 1.6 Hz, 1H), 7.91 (d, *J* = 9.2 Hz, 2H), 7.84 (d, *J* = 8.4 Hz, 1H), 7.49 (dd, *J*<sub>1</sub> = 8.4 Hz, *J*<sub>2</sub> = 1.6 Hz, 1H), 6.71 (d, *J* = 9.2 Hz, 2H), 3.04 (s, 6H), 0.24 (s, 9H).

### **2-(4-(dimethylamino)phenyl)-6-ethynyl-3-methylbenzo[d]thiazol-3-ium (2-20)**

A suspension of N,N-dimethyl-4-(6-((trimethylsilyl)ethynyl)benzo[d]thiazol-2-yl)aniline (**2-19**) (52 mg, 0.15 mmol) and K<sub>2</sub>CO<sub>3</sub> (205 mg, 1.5 mmol) in DCM (1.5 ml) and methanol (6 ml) was stirred vigorously at r.t. for 1 h. The solvents were then removed under vacuum.

The product was isolated with silica gel column (hexane/EtOAc = 6:1), and then treated as general procedure **2B**. Yield: 41% from **2-19**.  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$ : 8.34 (d,  $J = 1.5$  Hz, 1H), 8.11 (d,  $J = 9.0$  Hz, 2H), 7.94 (d,  $J = 9.0$  Hz, 1H), 7.87 (dd,  $J_1 = 9.0$  Hz,  $J_2 = 1.5$  Hz, 1H), 7.05 (d,  $J = 9.2$  Hz, 2H), 4.33 (s, 3H), 3.22 (s, 6H).

### **2-(4-azidophenyl)-6-methoxy-3-methylbenzo[d]thiazol-3-ium (2-21)**

4-(6-methoxybenzo[d]thiazol-2-yl)aniline (**2-14**) was treated with the same procedure for **2-35**, then followed general procedure **2B**. Yield: 51%.  $^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$ : 8.02 (d,  $J = 9.2$  Hz, 1H), 7.84 (d,  $J = 8.8$  Hz, 2H), 7.55 (d,  $J = 2.4$  Hz, 1H), 7.44 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 1H), 7.32 (d,  $J = 8.8$  Hz, 2H), 4.36 (m, 3H), 3.97 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 170.3, 160.9, 146.9, 136.6, 132.2, 131.4, 120.8, 118.2, 117.5, 114.6, 105.5, 56.6, 38.4.

### **methyl 2-(tritylthio)acetate (2-23a)**

DIPEA (1.04 ml, 6 mmol) was added slowly to a solution of methyl 2-mercaptoacetate (**2-22a**) (0.40 ml, 4.5 mmol) and trityl chloride (0.84 g, 3.0 mmol) in DCM (20 ml). The resulting solution was stirred at r.t. for 2 h. 1 M NaOH (20 ml) was added to the reaction mixture. The organic layer was separated, washed with 1 M NaOH (10 ml) and dried over  $\text{Na}_2\text{SO}_4$ . The product was isolated with silica gel column (hexanes/EtOAc = 20:1). Yield: 1.05 g (67%).  $^1\text{H}$  NMR (400 MHz,



Chloroform-*d*)  $\delta$ : 7.1-7.4 (m, 15H), 3.57 (s, 3H), 2.98 (s, 2H).

### **methyl 3-(tritylthio)propanoate (2-23b)**

The procedure is the same as for compound **2-23a**. Yield: 65%.  $^1\text{H}$  NMR (300 MHz, Chloroform-*d*)  $\delta$ : 7.1-7.5 (m, 15H), 3.63 (s, 3H), 2.46 (t,  $J = 7.2$  Hz, 2H), 2.24 (t,  $J = 7.2$  Hz, 2H).

### **2-(tritylthio)acetaldehyde (2-24a)**

DIBAL (1 ml, 1.5 M in toluene) was added dropwise to a stirred solution of **2-23a** (240 mg, 0.69 mmol) in anhydrous DMF (5 ml) under  $\text{N}_2$  protection at  $-78$  °C. The reaction was stirred for another 30 min at  $-78$  °C, quenched with methanol (5 ml), and added with 1 M Rochelle's salt solution (10 ml). After warmed to r.t., the resulting mixture was added with ether (20 ml) and shaken vigorously. The organic layer was separated, and the aqueous layer was extracted with ether ( $2 \times 15$  ml). The combined organic layers were washed with brine and dried over  $\text{Na}_2\text{SO}_4$ . The product was isolated with silica gel column (hexanes/EtOAc = 20:1). Yield: 140 mg (63%).  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$ : 8.84 (t,  $J = 2.8$  Hz, 1H), 7.2-7.5 (m, 15H), 3.09 (d,  $J = 2.8$  Hz, 2H).

### **3-(tritylthio)propanal (2-24b)**

The procedure is the same as for compound **2-24a**. Yield: 53%.  $^1\text{H}$  NMR (300

MHz, Chloroform-*d*)  $\delta$ : 9.55 (s, 1H) 7.2-7.5 (m, 15H), 2.3-2.5 (m, 4H).

#### **4-(6-methoxybenzo[d]thiazol-2-yl)-N-(2-(tritylthio)ethyl)aniline (2-25a)**

Compounds **2-14** (98 mg, 0.37 mmol) and **2-24a** (112 mg, 0.35 mmol) were dissolved in 1,2-dichloroethane (7 ml), and treated with NaBH(OAc)<sub>3</sub> (112 mg, 0.53 mmol). The resulting solution was stirred at r.t. for 4 h. The reaction was quenched with sat. NaHCO<sub>3</sub> (10 ml). the organic layer was separated, and the aqueous layer was extracted with DCM (2  $\times$  15 ml). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The product was isolated with silica gel column (hexanes/EtOAc = 4:1). Yield: 92 mg (47%). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$ : 7.86 (d, *J* = 8.8 Hz, 1H), 7.79 (d, *J* = 8.4 Hz, 2H), 7.2-7.5 (m, 16H), 7.04 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.0 Hz, 1H), 6.45 (d, *J* = 8.4 Hz, 2H), 3.88 (s, 3H), 3.07 (m, 2H), 2.05 (t, *J* = 4.8 Hz, 2H).

#### **4-(6-methoxybenzo[d]thiazol-2-yl)-N-(3-(tritylthio)propyl)aniline (2-25b)**

The procedure is the same as for compound **2-25a**. Yield: 59%. <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$ : 7.85 (d, *J* = 8.8 Hz, 1H), 7.81 (d, *J* = 8.4 Hz, 2H), 7.2-7.5 (m, 16H), 7.02 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 2.0 Hz, 1H), 6.53 (d, *J* = 8.4 Hz, 2H), 3.88 (s, 3H), 3.13 (m, 2H), 2.30 (m, 2H), 1.68 (m, 2H).

#### **2-(4-(6-methoxybenzo[d]thiazol-2-yl)phenylamino)ethanethiol (2-26a)**

Compound **2-25a** (92 mg, 0.15 mmol) and triethylsilane (60  $\mu$ l, 0.375 mmol) were dissolved in 1:1 TFA/DCM mixture (4 ml). The reaction solution was stirred at r.t. for 1.5 h, and neutralized with sat.  $\text{NaHCO}_3$  to pH  $\sim$  6. The organic layer was separated, and the aqueous layer was extracted with DCM (2  $\times$  15 ml). The combined organic layers were washed with sat.  $\text{NH}_4\text{Cl}$  and dried over  $\text{Na}_2\text{SO}_4$ . The product was isolated with silica gel column (hexanes/EtOAc = 2:1). Yield: 42 mg (80%).  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$ : 7.86 (m, 3H), 7.32 (d,  $J$  = 2.4 Hz, 1H), 7.04 (dd,  $J_1$  = 8.8 Hz,  $J_2$  = 2.4 Hz, 1H), 6.68 (m, 2H), 3.88 (s, 3H), 3.43 (m, 2H), 2.81 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 167.6, 129.7, 128.5, 122.1, 112.9, 104.4, 55.8, 44.6, 33.4, 29.7, 22.5, 14.1. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{16}\text{H}_{17}\text{N}_2\text{OS}_2$   $[\text{M}]^+$ , 317.07902; found 317.07823.

### **3-(4-(6-methoxybenzo[d]thiazol-2-yl)phenylamino)propane-1-thiol (2-26b)**

The procedure is the same as for compound **2-25a**. Yield: 80%.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$ : 7.87 (m, 3H), 7.32 (d,  $J$  = 2.8 Hz, 1H), 7.03 (dd,  $J_1$  = 8.8 Hz,  $J_2$  = 2.4 Hz, 1H), 6.67 (d,  $J$  = 8.4 Hz, 2H), 3.88 (s, 3H), 3.35 (m, 2H), 2.67 (m, 2H), 1.96 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 126.4, 120.5, 113.0, 109.9, 101.8, 53.3, 39.3, 30.7, 30.5, 29.4, 27.1, 20.1, 19.6, 11.6. HRMS (ESI+):  $m/z$  calculated for  $\text{C}_{17}\text{H}_{19}\text{N}_2\text{OS}_2$   $[\text{M}]^+$ , 331.09393; found 331.09388.

### **2-(4-(dimethylamino)phenyl)-6-(2-hydroxyethoxy)-3-methylbenzo[d]thiazol-3-yl**

### **m (2-29)**

Compound **2-33** was treated as general procedure 2B (yield: 65%), and the product was mixed with 4:1 acetonitrile/water mixture (30 ml) and CsF (10 equiv.). The resulting solution was refluxed for 2 h, and the solvents were removed under vacuum. The product was isolated by HPLC purification. <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$ : 7.91 (d, *J* = 8.8 Hz, 1H), 7.64 (d, *J* = 8.8 Hz, 2H), 7.52 (m, 1H), 7.41 (m, 1H), 6.88 (d, *J* = 8.8 Hz, 2H), 4.31 (s, 3H), 4.25 (m, 2H), 4.06 (m, 2H), 3.17 (s, 6H).

### **2-(4-(dimethylamino)phenyl)-3-methyl-6-(2-(methylsulfonyloxy)ethoxy)benzo[d]thiazol-3-ium (2-30)**

Compound **2-29** (20 mg, 0.045 mmol) and DIPEA (9.3  $\mu$ l, 0.054 mmol) were dissolved in anhydrous DCM (1 ml), and added with mesyl chloride (2.6  $\mu$ l, 0.054 mmol) at 0 °C. The resulting solution was stirred at r.t. for 1 h. DCM was then removed under vacuum, and the product was isolated by HPLC purification. Yield: 86 from **2-28**. <sup>1</sup>H NMR (400 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$ : 7.94 (d, *J* = 9.2 Hz, 1H), 7.75 (s, 1H), 7.70 (d, *J* = 8.8 Hz, 2H), 7.44 (m, 1H), 6.88 (d, *J* = 9.2 Hz, 2H), 4.62 (m, 2H), 4.44 (m, 2H), 4.30 (s, 3H), 3.16 (s, 6H), 3.10 (s, 3H). <sup>13</sup>C NMR (100 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$ : 158.8, 154.8, 137.7, 132.6, 119.9, 117.8, 112.8, 111.1, 107.7, 68.4, 67.7, 40.6, 39.0, 38.2. HRMS (ESI<sup>+</sup>): *m/z* calculated for C<sub>19</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>S<sub>2</sub> [M]<sup>+</sup>, 407.109375; found 407.11230.

### **N-(4-(6-methoxybenzo[d]thiazol-2-yl)phenyl)acetamide (2-31)**

The procedure is the same as for compound **2-3**. Yield: 74%.  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$ : 10.25 (s, 1H), 7.97 (d,  $J = 8.0$  Hz, 2H), 7.90 (d,  $J = 8.8$  Hz, 1H), 7.76 (d,  $J = 8.0$  Hz, 2H), 7.69 (d,  $J = 2.0$  Hz, 1H), 7.11 (dd,  $J_1 = 8.8$  Hz,  $J_2 = 2.0$  Hz, 1H), 3.85 (s, 3H), 2.10 (s, 3H).

#### **N-(4-(6-hydroxybenzo[d]thiazol-2-yl)phenyl)acetamide (2-32)**

The procedure is the same as for compound **2-4**. Yield: 93%.  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$ : 10.23 (s, 1H), 7.95 (d,  $J = 9.0$  Hz, 2H), 7.81 (d,  $J = 8.5$  Hz, 1H), 7.75 (d,  $J = 9.0$  Hz, 2H), 7.39 (d,  $J = 2.5$  Hz, 1H), 6.98 (dd,  $J_1 = 8.5$  Hz,  $J_2 = 2.5$  Hz, 1H), 2.10 (s, 3H).

#### **N-(4-(6-(prop-2-ynyloxy)benzo[d]thiazol-2-yl)phenyl)acetamide (2-33)**

Followed general procedure 2A. Yield: 73%.  $^1\text{H}$  NMR (500 MHz, Acetone- $d_6$ )  $\delta$ : 8.02 (d,  $J = 8.8$  Hz, 2H), 7.94 (d,  $J = 9.2$  Hz, 1H), 7.83 (d,  $J = 8.8$  Hz, 2H), 7.70 (d,  $J = 2.4$  Hz, 1H), 7.20 (dd,  $J_1 = 9.2$  Hz,  $J_2 = 2.4$  Hz, 1H), 4.92 (d,  $J = 2.0$  Hz, 2H), 2.96 (m, 1H), 2.14 (s, 3H).

#### **4-(6-(prop-2-ynyloxy)benzo[d]thiazol-2-yl)aniline (2-34)**

N-(4-(6-(prop-2-ynyloxy)benzo[d]thiazol-2-yl)phenyl)acetamide (**2-33**) (100 mg, 0.32 mmol) was refluxed in 10 ml 1.2 N HCl for 16 h. After cooled to r.t., the reaction solution was neutralized to pH 7-8 with 1 M NaOH, and then extracted with DCM (3

× 10 ml). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The product was purified with silica gel column (hexane/EtOAc = 3:2). Yield: 87 mg (95%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ: 7.88 (d, *J* = 9.0 Hz, 1H), 7.84 (d, *J* = 9.0 Hz, 2H), 7.43 (d, *J* = 2.5 Hz, 1H), 7.11 (dd, *J*<sub>1</sub> = 9.0 Hz, *J*<sub>2</sub> = 2.5 Hz, 1H), 6.73 (d, *J* = 9.0 Hz, 1H), 4.76 (d, *J* = 2.5 Hz, 2H), 2.55 (m, 1H).

### **2-(4-azidophenyl)-6-(prop-2-ynyloxy)benzo[d]thiazole (2-35)**

To a solution of 4-(6-(prop-2-ynyloxy)benzo[d]thiazol-2-yl)aniline (**2-34**) (18 mg, 0.064 mmol) in 2 ml acetonitrile was added *t*-butyl nitrite (11.4 μl, 0.096 mmol) at 0 °C. After stirred 10 min at 0 °C, the resulting solution was added with azidotrimethylsilane (10 μl, 0.077 mmol) dropwise. The reaction was then stirred at r.t. for another 1.5 h. The solvent was removed under vacuum, and the product was isolated with silica gel column (5:1 hexane/EtOAc). Yield: 18 mg (92%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ: 8.03 (d, *J* = 8.5 Hz, 2H), 7.95 (d, *J* = 9.0 Hz, 1H), 7.46 (d, *J* = 2.0 Hz, 1H), 7.17 (dd, *J*<sub>1</sub> = 9.0 Hz, *J*<sub>2</sub> = 2.0 Hz, 1H), 7.12 (d, *J* = 8.5 Hz, 2H), 4.78 (d, *J* = 2.5 Hz, 2H), 2.57 (m, 1H).

### **2-(4-azidophenyl)-3-methyl-6-(prop-2-ynyloxy)benzo[d]thiazol-3-ium (2-36)**

Followed general procedure 2B. Yield: 55%. <sup>1</sup>H NMR (400 MHz, Acetone-*d*<sub>6</sub>) δ: 8.22 (d, *J* = 9.2 Hz, 1H), 8.02 (d, *J* = 2.4 Hz, 1H), 7.90 (d, *J* = 9.2 Hz, 2H), 7.57 (dd, *J*<sub>1</sub> = 9.2 Hz, *J*<sub>2</sub> = 2.4 Hz, 1H), 7.05 (d, *J* = 9.2 Hz, 2H), 5.02 (d, *J* = 2.4 Hz, 2H),

4.47 (s, 1H).

### **Spectroscopic characterization of ThT derivatives**

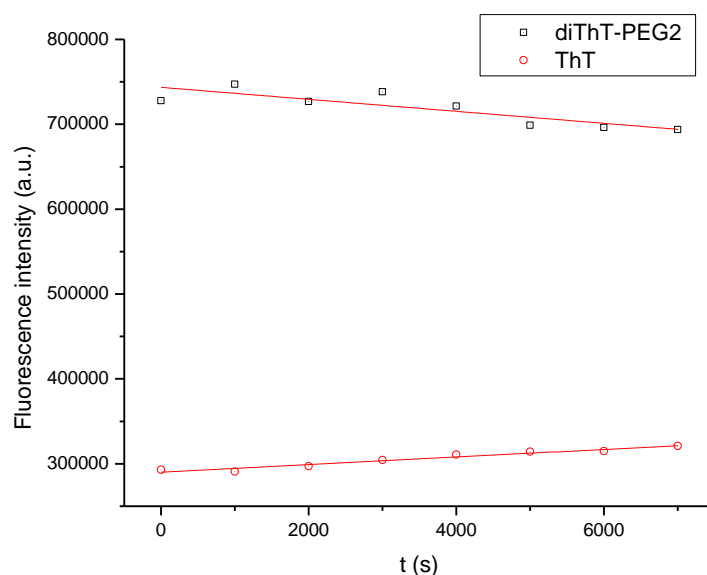
All ThT monomer derivatives were stored in DMSO stock solution at 2.5 mM (dimers at 1.25 mM) and diluted into a phosphate buffer (50 mM phosphates, 300 mM NaCl, pH 7.4) in all tests below. UV/Vis spectra were collected for each ThT monomer at 5  $\mu$ M (dimer at 2.5  $\mu$ M) in the phosphate buffer. The extinction coefficient values were obtained from the absorption spectra at 410 nm.

The fluorescence excitation spectra of ThT and diThT-PEG3 were collected in the phosphate buffer with 2 $\mu$ M A $\beta$ -(1-40) fibril on a Fluorolog-3 fluorimeter and a MicroMax 384 microwell plate reader (Jobin Yvon Inc. Edison, NJ). The small molecule concentration used was 2  $\mu$ M for ThT and 1  $\mu$ M for diThT PEG3. The result shows that the ThT monomer and dimer share the same excitation profile ( $\lambda_{\text{max}} = 440$  nm). The fluorescence intensity difference is due to the fact that diThT-PEG3 binds better than the ThT monomer under the experimental conditions.

Since only the portions of ThT derivatives bond to amyloid show fluorescence emission, the measurements of quantum yields were disrupted by the binding constants. Therefore, the quantum yield of ThT was measured at the plateau and set as a relative standard. The quantum yields for all other ThT monomer and dimer derivatives were also taken at their plateaus and compared with ThT. The result shows

that all ThT derivatives have similar quantum yields as ThT, and the difference is no more than 2 folds. All the other details of the results are shown in the section 2.3.

The photobleaching spectra of ThT and diThT-PEG2 were collected in the phosphate buffer with 2  $\mu\text{M}$  A $\beta$ -(1-40) fibril on a Fluorolog-3 fluorimeter (Jobin Yvon Inc. Edison, NJ). The small molecule concentration used was 1  $\mu\text{M}$  for ThT and 0.5  $\mu\text{M}$  for diThT PEG2. The samples were excited at 440 nm, and the emission fluorescence intensity was collected at 490 nm over 7,000 s (Figure 2-16). The result shows that photobleaching is not significant for either the monomer or the dimer.



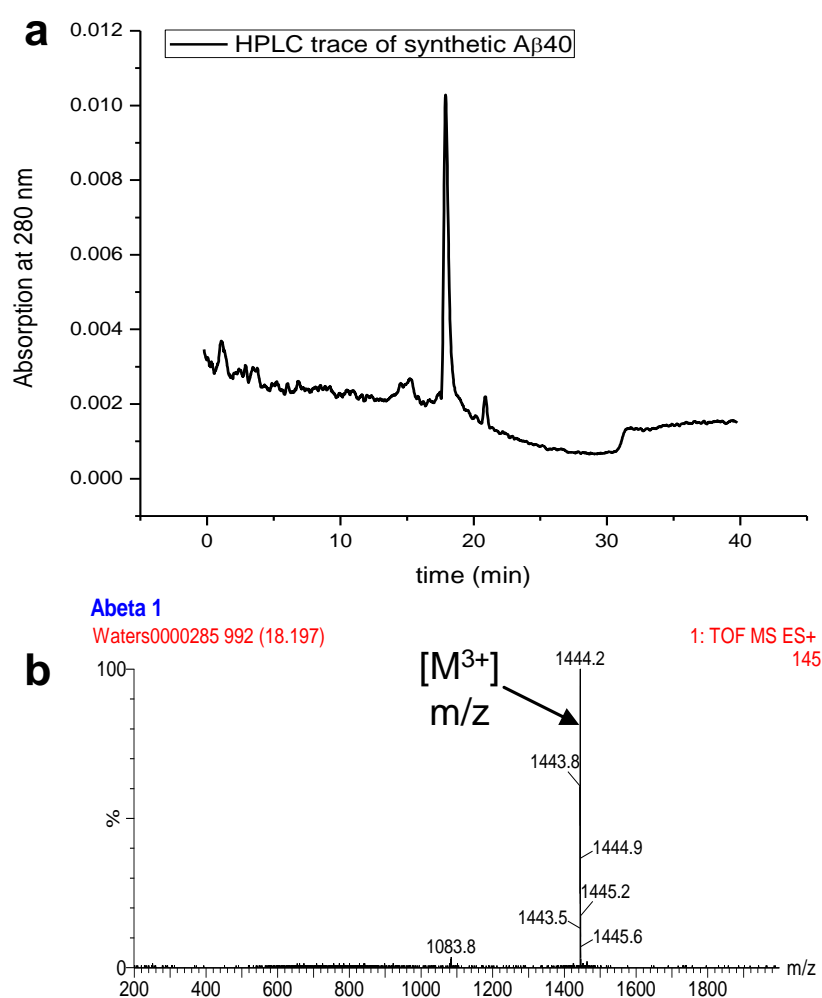
**Figure 2-16.** The photobleaching spectra of ThT and diThT-PEG2

### A $\beta$ 40 synthesis

A $\beta$ 40 was synthesized through the standard Fmoc/<sup>t</sup>Bu chemistry with the



Fmoc-Phe-Wang resin (Novabiochem) as the solid support. The synthesis was carried out on 0.1 mmole scale. Five equivalents of amino acids and HBTU were used for the coupling reaction. To ensure the quality of the peptide, double couplings were carried out for residues after all beta-branched amino acids. The peptides were cleaved off the resin and deprotected with TFA (95% TFA, 2.5% H<sub>2</sub>O, 2.5% triethylsilane). The crude products were purified by RP-HPLC (Waters Prep LC, Jupiter 10u C18 300A



**Figure 2-17.** Analytical HPLC trace of the synthetic A $\beta$ 40 exhibiting good purity (a) and ESI-MS result confirming the identity of the peptide (b).

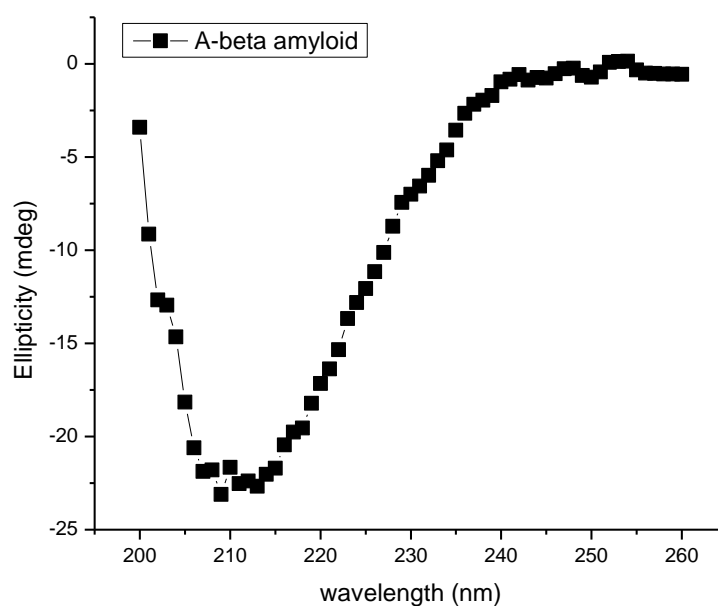
Column). The integrity of the peptide was conformed by LC-MS, which showed the purity is greater than 90% (Figure 2-17).

### **Preparation and characterization of A $\beta$ 40 fibrils**

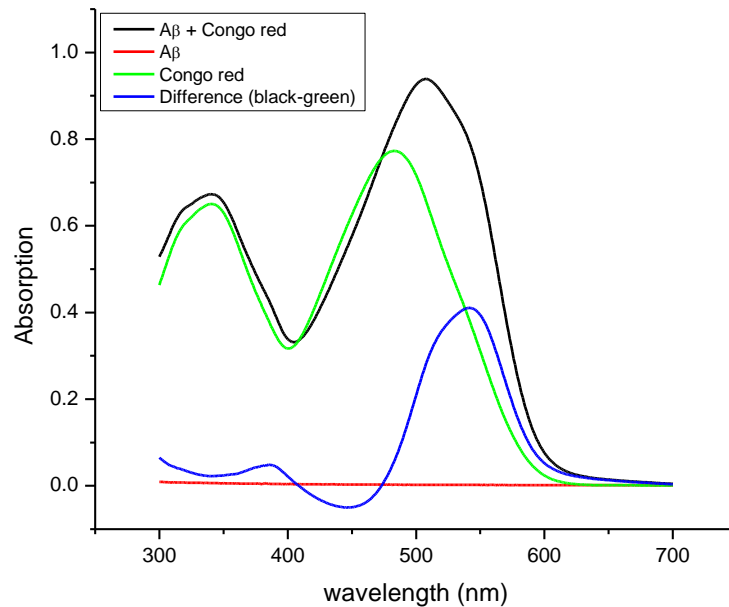
In order to prepare A $\beta$ 40 amyloid, the peptide was treated according to a published protocol to obtain fresh A $\beta$ 40 monomers. Briefly, the lyophilized A $\beta$ 40 peptide was dissolved in 1 mM NaOH solution, filtered through 0.22  $\mu$ M syringe filter (Millipore, Billerica, MA), and then through a Centricon filter with 10KDa molecular weight cutoff (Millipore, Billerica, MA). The peptide was diluted into the phosphate buffer (50 mM phosphates, 300 mM NaCl, pH 7.4); the concentration was determined by the tyrosine absorption at 280 nm right after the dilution. The sample was put on the rotating wheel at 30 rpm. The aggregation progress was monitored by circular dichroism spectroscopy, which showed complete aggregation after seven days. The formation of beta-sheet rich amyloids was confirmed by CD data (Figure 2-18) and the Congo red binding assay (Figure 2-19). The stock fibril solution was stored at 2 °C and used within 2 weeks.

*Congo Red assay:* <sup>[25]</sup> The Congo Red stock was prepared in 90% phosphate buffer saline (10 mM phosphate, 2.7 mM KCl, and 137 mM NaCl; pH 7.4) and 10% ethanol. The concentration was determined by measuring the absorbance of a diluted aliquot in a solution of sodium phosphate (1 mM, pH 7.0) and 40% ethanol at 505 nm. The

stock solution of A $\beta$ 40 amyloid was mixed with the Congo Red stock in phosphate buffer (50 mM phosphates, 300 mM NaCl, pH 7.4) to yield a final concentration of 20  $\mu$ M Congo Red and 20  $\mu$ M A $\beta$ -(1-40) fibril. Two control samples were prepared: (1) a solution of 20  $\mu$ M Congo Red alone and (2) a solution of 20  $\mu$ M A $\beta$ 40 amyloid alone. The UV/Vis absorption spectra of all three samples were collected and shown below. The spectral shift allowed us to back calculate the amyloid concentration, which nicely agreed with the concentration that we prepared it to be. This result indicates A $\beta$ 40 aggregation largely yielded beta-sheet structured amyloid, rather than non-specific aggregates.



**Figure 2-18.** Circular dichroism spectrum of the aggregated A $\beta$ 40.



**Figure 2-19.** UV/Vis absorption shift of Congo red binding to A $\beta$  amyloid.

### Measuring the amyloid binding affinities of the ThT derivatives

The A $\beta$ 40 amyloid stock solution was sonicated in ice-water bath for 1 h and added into solutions of a ThT derivative at various concentrations. The final concentration of A $\beta$ 40 amyloid was set at 2  $\mu$ M. The fluorescence emission was scanned immediately in a 384 well microplate (50  $\mu$ L each sample) at 465-650 nm with the excitation at 440 nm and slits as 6 nm on a Fluorolog-3 fluorometer and a MicroMax 384 microwell plate reader (Jobin Yvon Inc. Edison, NJ). Triplicates of each sample were prepared and measured.

The fluorescence intensity at 485nm was plotted against the small molecule

concentration, yielding the binding curves shown in the section **2.3**.

### **Evaluating the binding specificity of the ThT dimers to A $\beta$ 40 amyloid**

The stock A $\beta$ 40 amyloid fibril solution was sonicated for 1 h and diluted to various concentrations into solutions of 1  $\mu$ M ThT or 0.5  $\mu$ M diThT derivatives and 10% (v/v) bovine serum (protein concentration ~60-80 mg/ml) in phosphate buffer. The fluorescence emission was scanned immediately in a 96 well microplate (3  $\times$  150  $\mu$ L each sample) at 465-650 nm with the excitation at 440 nm on a Spectra Max M5 microwell plate reader (Molecular Devices, Sunnyvale, CA). The data are shown in the section **2.4**.

## References

1. Brookmeyer, R; Gray, S.; Kawas, C., *Projections of Alzheimer's Disease in the United States and the Public Health Impact of Delaying Disease Onset*. American Journal of Public Health, 1998. **88**: p. 1337-1342.
2. Waldemar, G., Dubois, B., Emre, M., Georges, J., McKeith, I. G., Rossor, M., Scheltens, P., Tariska, P., Winblad, B., *Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline*. European Journal of Neurology, 2007. **14**(1): p. e1-e26.
3. *About Alzheimer's Disease: Symptoms*. National Institute on Aging.
4. Schenk, D. J.; Schenk, D., *Alzheimer's disease: molecular understanding predicts amyloid-based therapeutics*. Annu. Rev. Pharmacol. Toxicol., 2003. **43**: p. 545-584.
5. Hartmann, T.; Bieger, S. C.; Brühl, B.; Tienari, P. J.; Ida, N.; Allsop, D.; Roberts, G. W.; Masters, C. L.; Dotti, C. G.; Unsicker, K.; Beyreuther, K., *Distinct sites of intracellular production for Alzheimer's disease A $\beta$ 40/42 amyloid peptides*. Nature Medicine, 1997. **3**: p. 1016-1020.
6. Yin, Y. I.; Bassit, B.; Zhu, L.; Yang, X.; Wang, C.; Li, Y.,  *$\gamma$ -Secretase Substrate Concentration Modulates the A $\beta$ 42/A $\beta$ 40 Ratio*. J. Bio. Chem. 2007. **282**(32): p. 23639-23644.

7. Caughey, B.; Lansbury, P. T., *Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders*. *Annu Rev Neurosci.*, 2003. **26**: p. 267-298.
8. LeVine, H. 3rd., *Quantification of beta-sheet amyloid fibril structures with Thioflavin T*. *Methods Enzymol.*, 1999. **309**: p. 274-284.
9. Lambert, M. P.; Barlow, A. K.; Chromy, B. A.; Edwards, C.; Freed, R.; Liosatos, M.; Morgan, T. E.; Rozovsky, I.; Trommer, B.; Viola, K. L.; Wals, P.; Zhang, C.; Finch, C. E.; Krafft, G. A.; Klein, W. L., *Diffusible, nonfibrillar ligands derived from A $\beta$  1-42 are potent central nervous system neurotoxins*. *Proc. Natl. Acad. Sci. USA*, 1998. **95**: p. 6448-6453.
10. Mathis, C. A.; Wang, Y.; Holt, D. P.; Huang, G. F.; Debnath, M. L.; Klunk, W. E., *Synthesis and evaluation of  $^{11}\text{C}$ -labeled 6-substituted 2-aryl benzothiazoles as amyloid imaging agents*. *J. Med. Chem.*, 2003. **46**: p. 2740-2754.
11. Klunk, W. E.; Wang, Y.; Huang, G. F.; Debnath, M. L.; Holt, D. P.; Mathis, C. A., *Uncharged thioflavin-T derivatives bind to amyloid-beta protein with high affinity and readily enter the brain*. *Life Sci.*, 2001. **69**: p. 1471-1484.
12. Klunk, W. E.; Engler, H.; Nordberg, A.; Wang, Y.; Blomqvist, G.; Holt, D. P.; Bergstrom, M.; Savitcheva, I.; Huang, G. F.; Estrada, S.; Ausen, B.; Debnath, M. L.; Barletta, J.; Price, J. C.; Sandell, J.; Lopresti, B. J.; Wall, A.; Koivisto, P.; Antoni, G.; Mathis, C. A.; Langstro, B., *Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B*. *Ann. Neurol.*, 2004. **55**: p.

306-319.

13. Krebs, M. R.; Bromley, E. H.; Donald, A. M., *The binding of thioflavin-T to amyloid fibrils: localisation and implications*. J. Struct. Biol., 2005. **149**: p. 30-37.
14. Biancalana, M.; Makabe, K.; Koide, A.; Koide, S., *Molecular Mechanism of Thioflavin-T Binding to the Surface of  $\beta$ -Rich Peptide Self-Assemblies* J. Mol. Biol., 2009. **385**: p. 1052-1063.
15. Wu, C.; Wang, Z. X.; Lei, H. X.; Duan, Y.; Bowers, M. T.; Shea, J.-E., *The binding of thioflavin T and its neutral analog BTA-1 to protofibrils of the Alzheimer's disease A $\beta$ 16-22 peptide probed by molecular dynamics simulations*. J. Mol. Biol., 2008. **384**: p. 718-729.
16. Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert D. and Riek, R., *3D structure of Alzheimer's amyloid- $\beta$ (1-42) fibrils*. Proc. Natl. Acad. Sci. U. S. A., 2005. **102**: p. 17342-17347.
17. Zhuang, Z.-P.; Kung, M.-P.; Hou, C.; Skovronsky, D. M.; Gur, T. L.; Plössl, K.; Trojanowski, J. Q.; Lee, V. M.-Y.; Kung, H. F., *Radioiodinated Styrylbenzenes and Thioflavins as Probes for Amyloid Aggregates*. J. Med. Chem., 2001. **44**: p. 1905-1914.
18. Klunk, W. E.; Jacob, R. F.; Mason, R. P., *Quantifying Amyloid by Congo Red Spectral Shift Assay*. Methods Enzymol., 1999. **309**: p. 285-305.
19. Yona, R. L.; Mazeret, S; Faller, P.; Gras, E., *Thioflavin Derivatives as Markers*

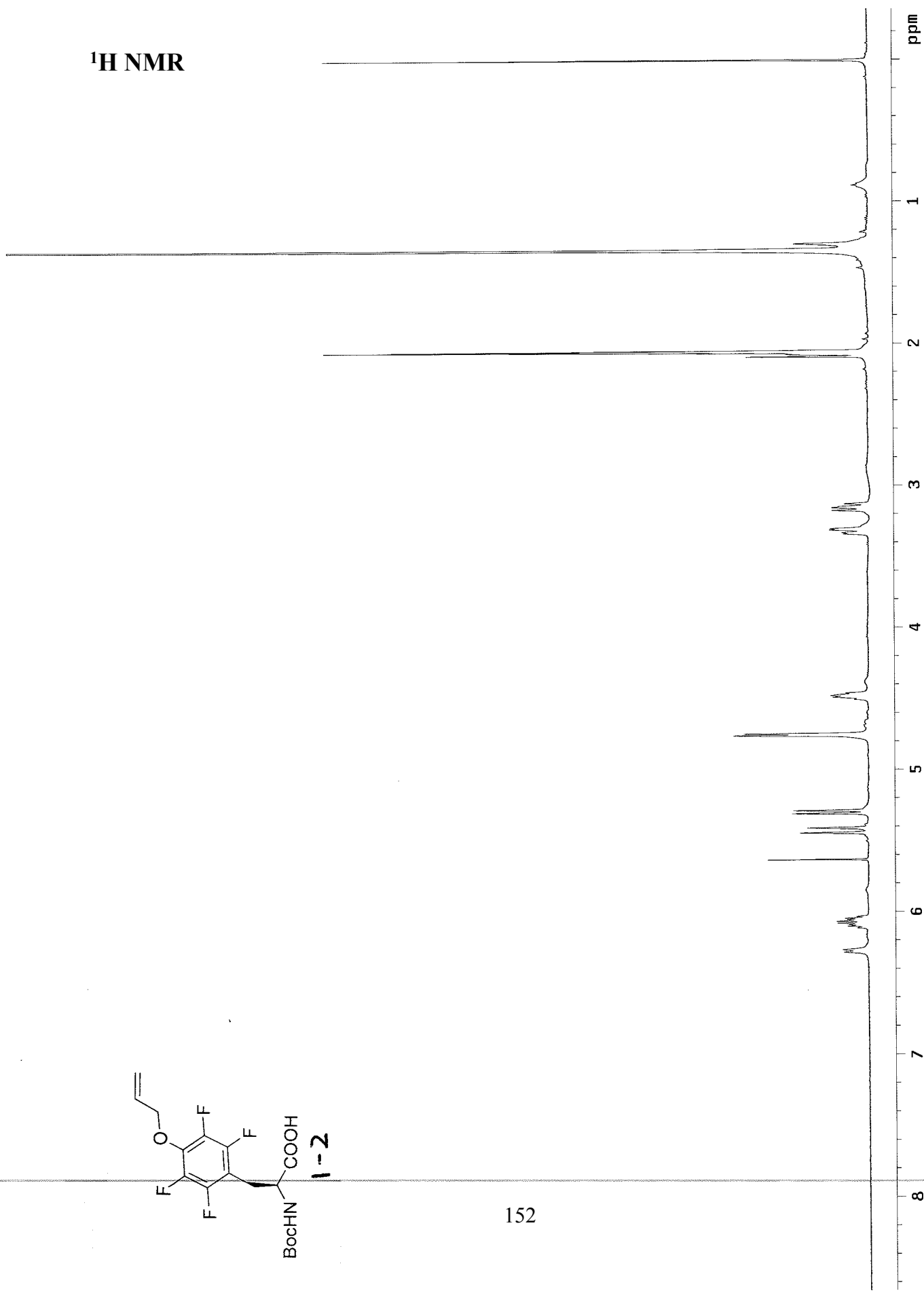
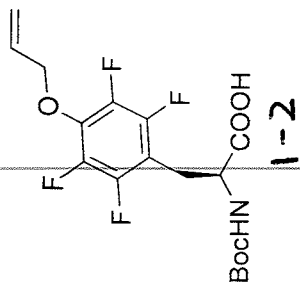


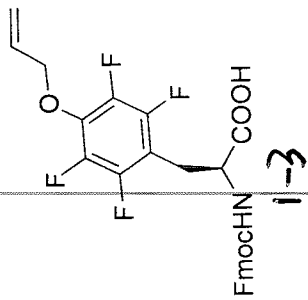
- for Amyloid- $\beta$  Fibrils: Insights into Structural Features Important for High-Affinity Binding*. ChemMedChem, 2008. **3**: p. 63-66.
20. Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D., *Reductive Amination of Aldehydes and Ketones with Sodium Triacetoxyborohydride. Studies on Direct and Indirect Reductive Amination Procedures*. J. Org. Chem., 1996. **61**: p. 3849-3862.
21. Alavez, S.; Vantipalli, M. C.; Zucker, D. J.; Klang, I. M.; Lithgow, G. J., *Amyloid-binding compounds maintain protein homeostasis during ageing and extend lifespan*. Nature, 2011. **472**: p. 226-229.
22. Barral, K.; Moorhouse, A. D.; Moses, J. E., *Efficient Conversion of Aromatic Amines into Azides: A One-Pot Synthesis of Triazole Linkages*. Org. Lett., 2007. **9**: p. 1809-1811.
23. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., *Click Chemistry: Diverse Chemical Function from a Few Good Reactions*. Angewandte Chemie International Edition, 2001. **40**(11): p. 2004-2021.
24. Sonogashira, K., *Development of Pd-Cu catalyzed cross-coupling of terminal acetylenes with sp<sup>2</sup>-carbon halides*. J. Organomet. Chem., 2002. **653**: p. 46-49.
25. Klunk, W. E.; Jacob, R. F.; Mason, R. P., *Quantifying amyloid beta-peptide (A $\beta$ ) aggregation using the Congo red-A $\beta$  (CR- $\beta$ ) spectrophotometric assay*. Anal Biochem, 1999. **266**(1): p. 66-76.

## **Appendix:**

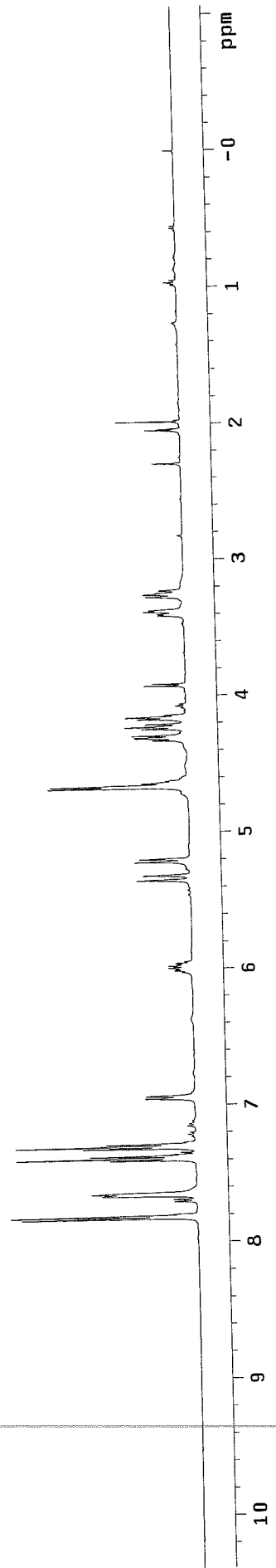
$^1\text{H}$  NMR and  $^{13}\text{C}$  NMR Spectra

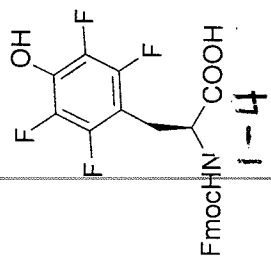
<sup>1</sup>H NMR



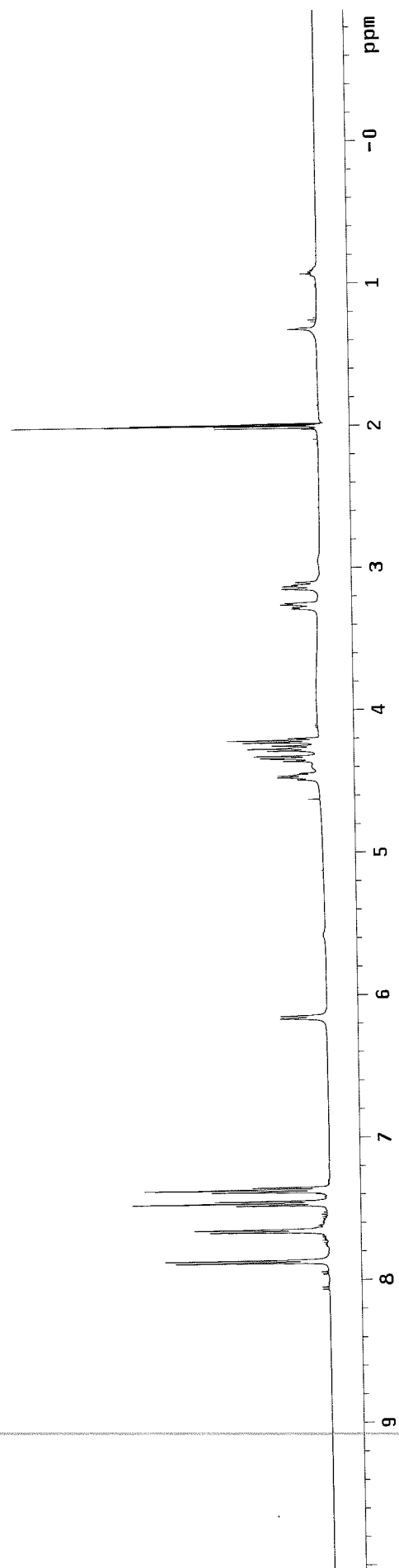


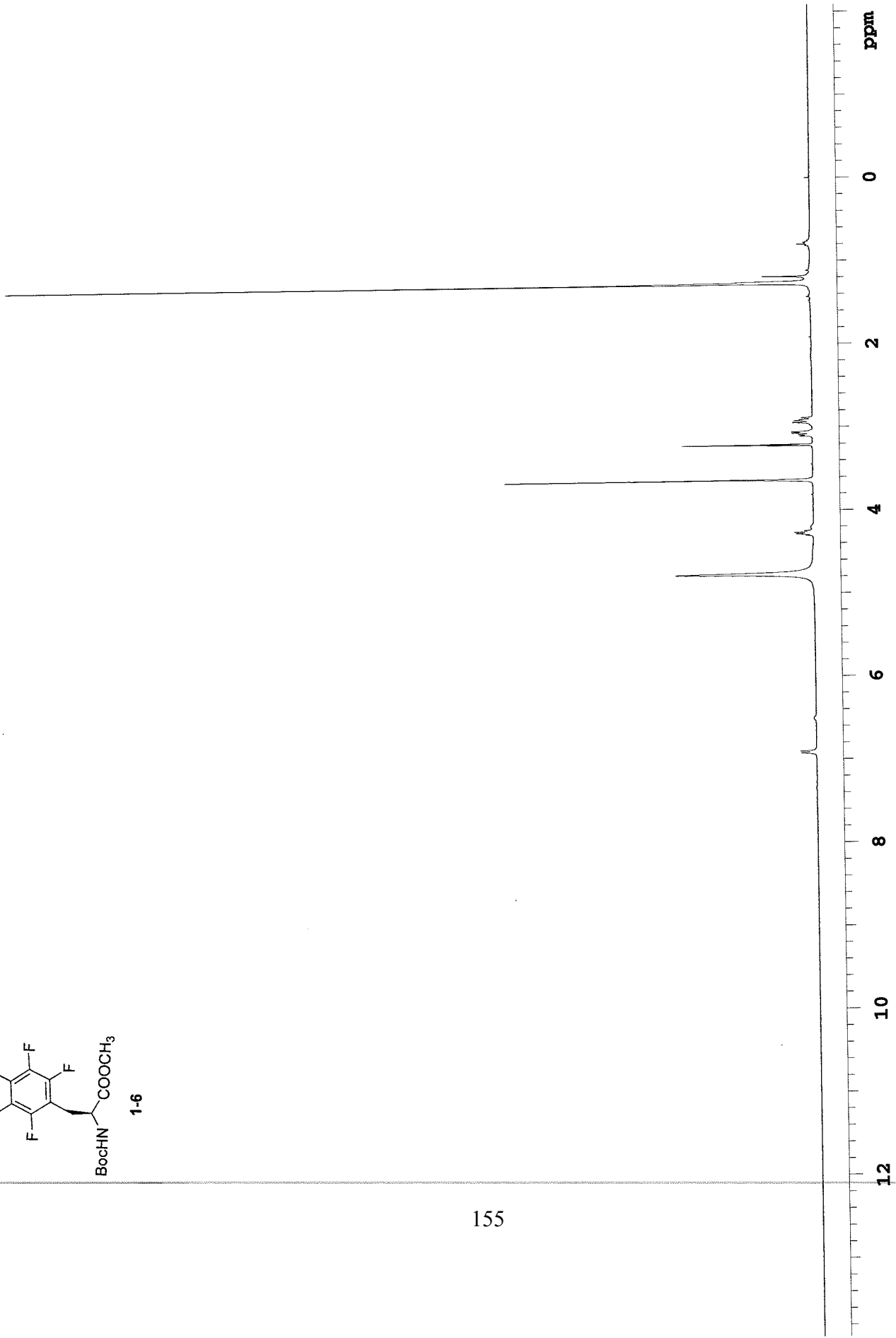
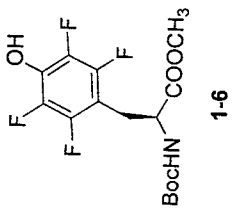
153

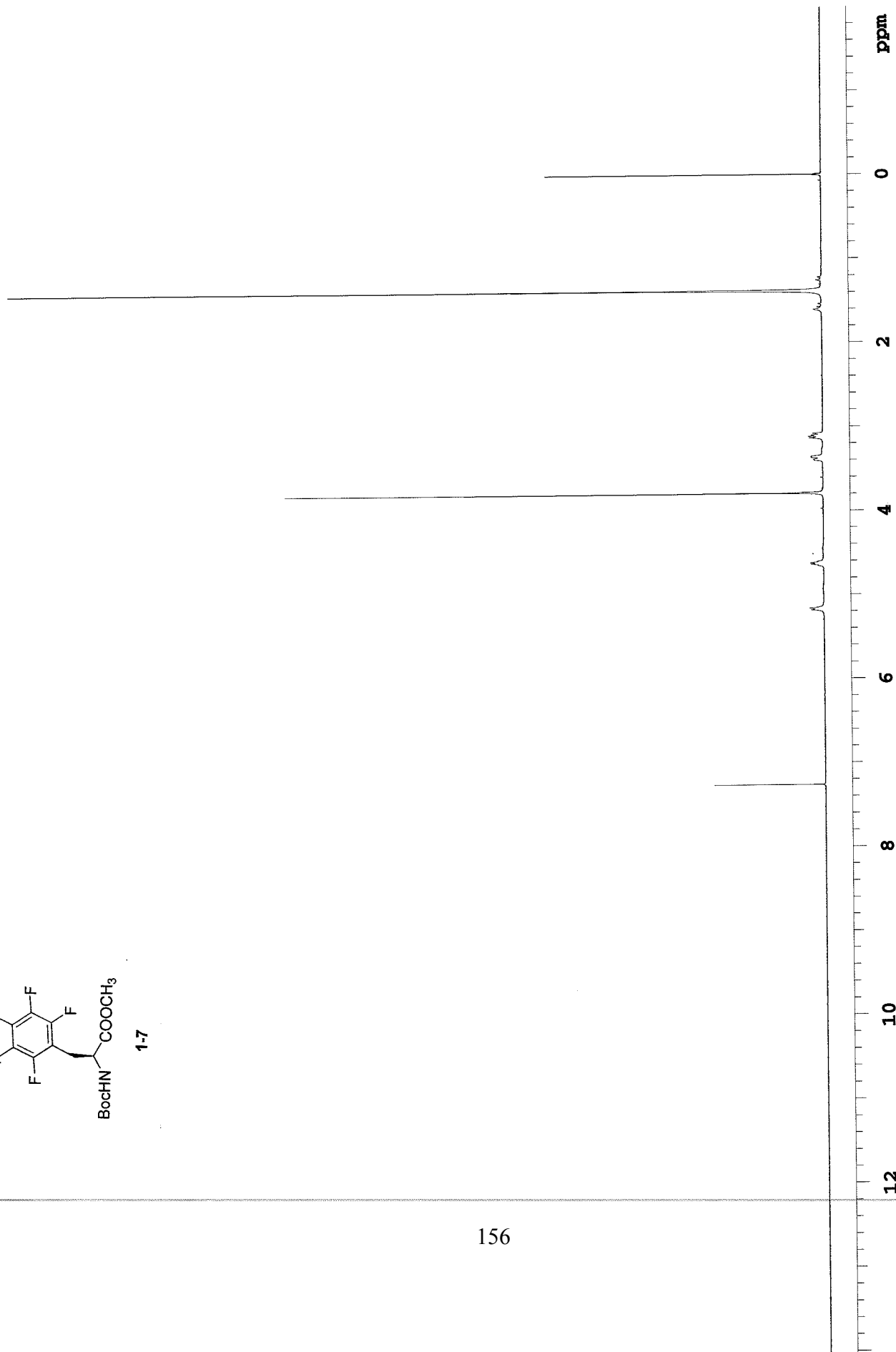
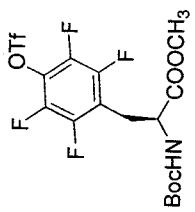


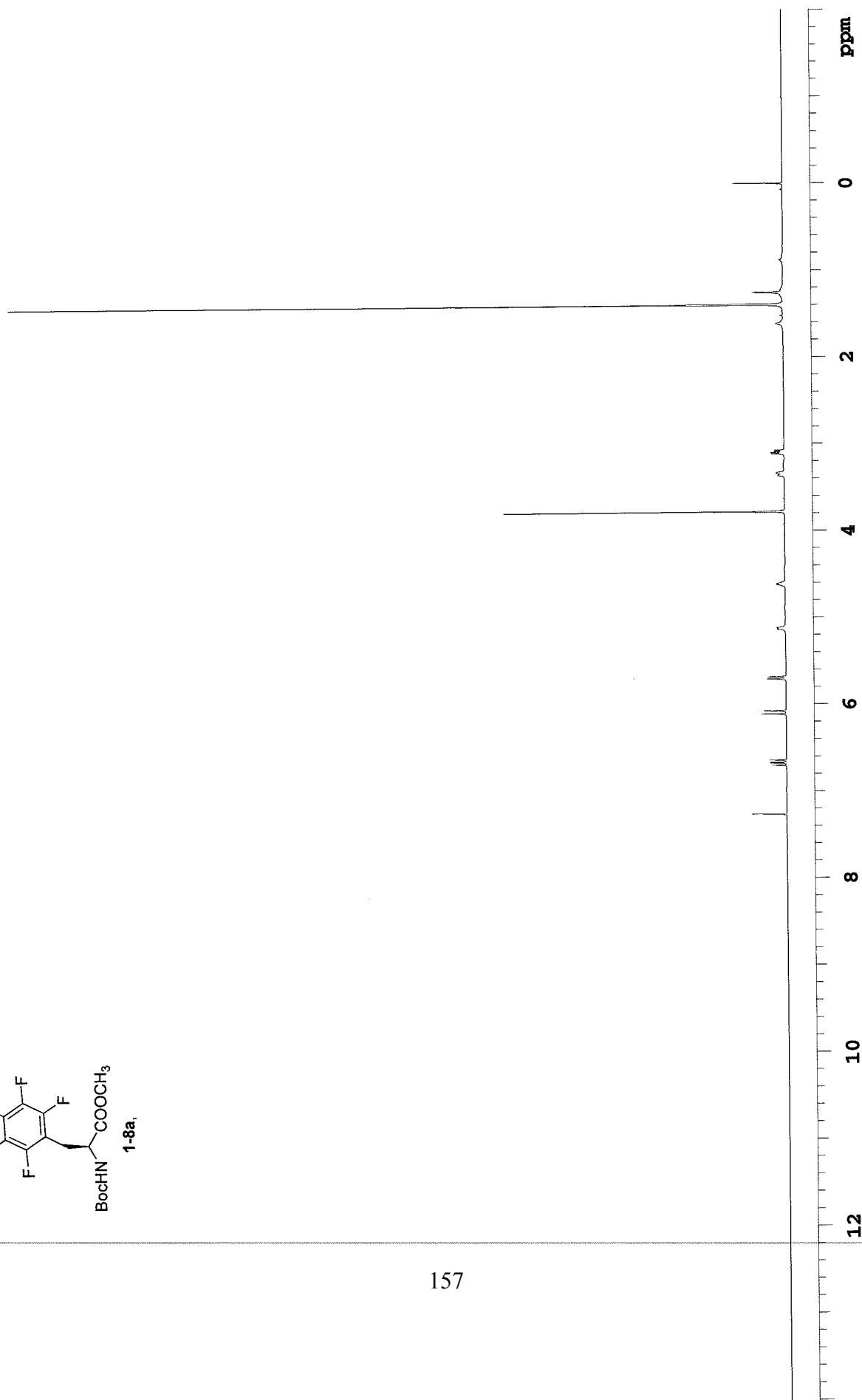
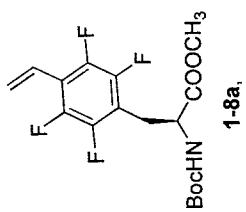


154

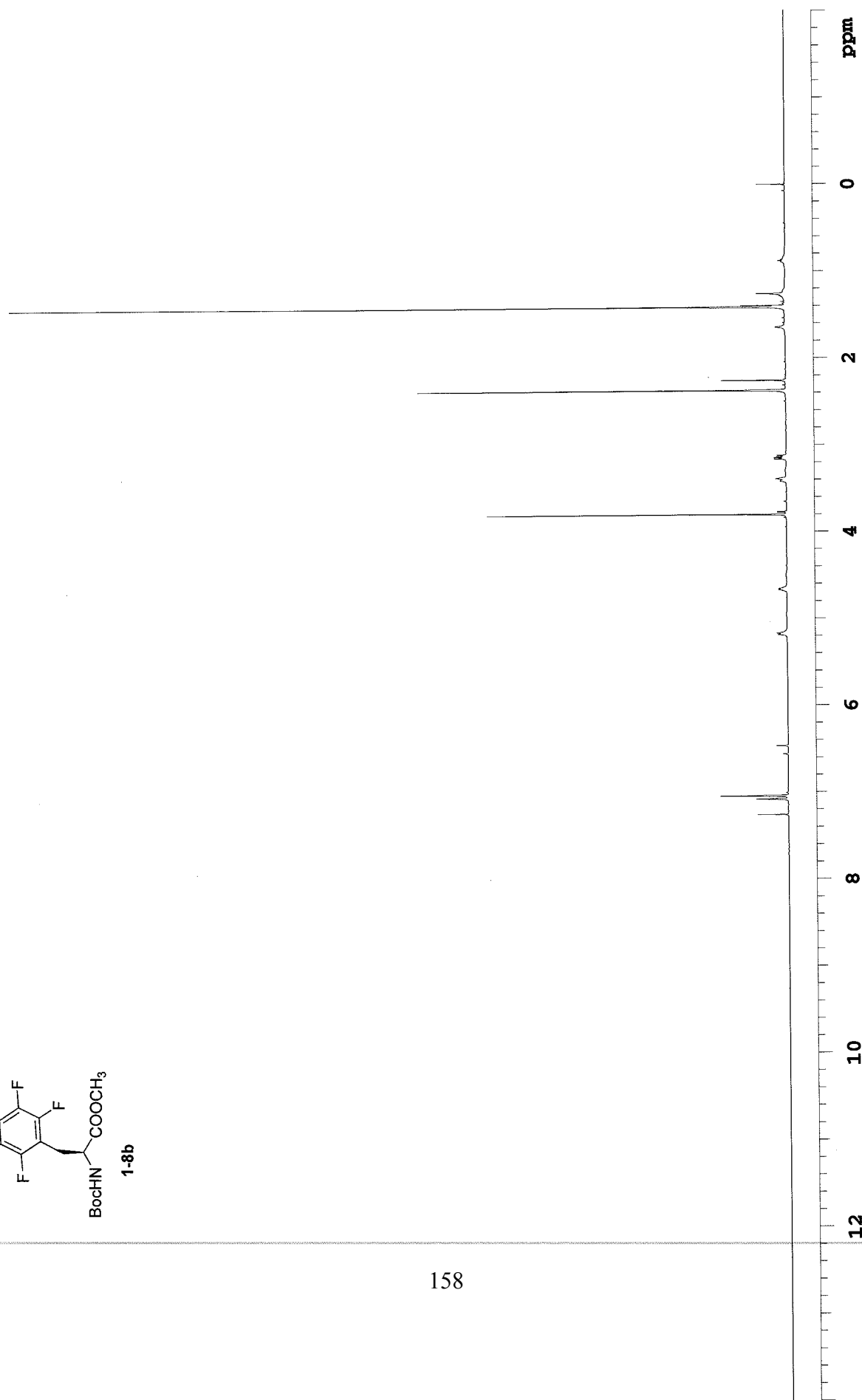
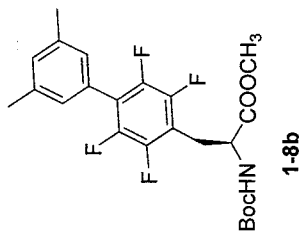


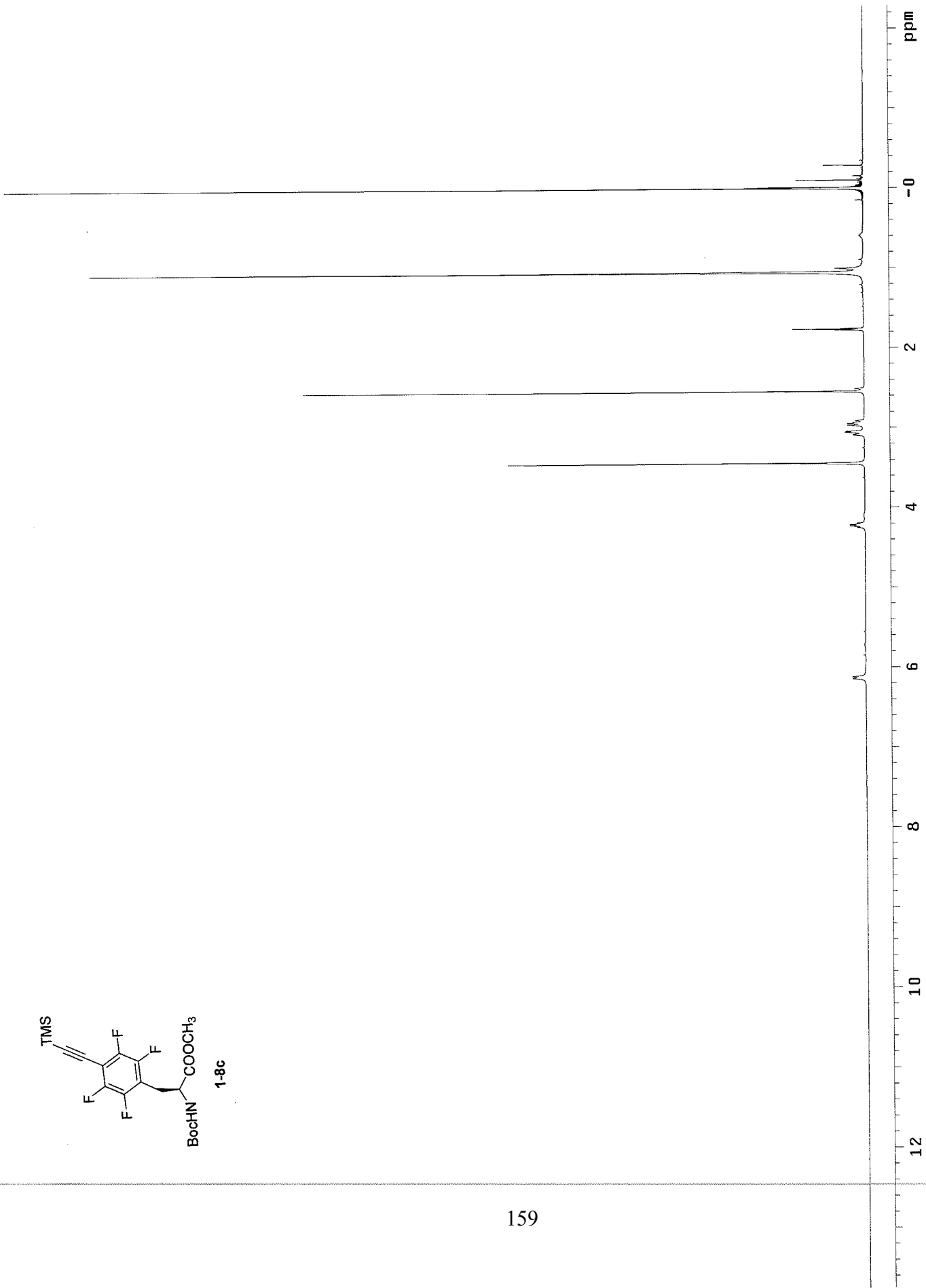
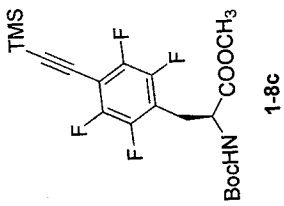


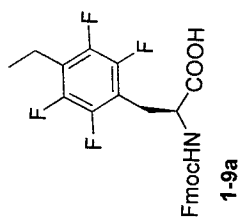




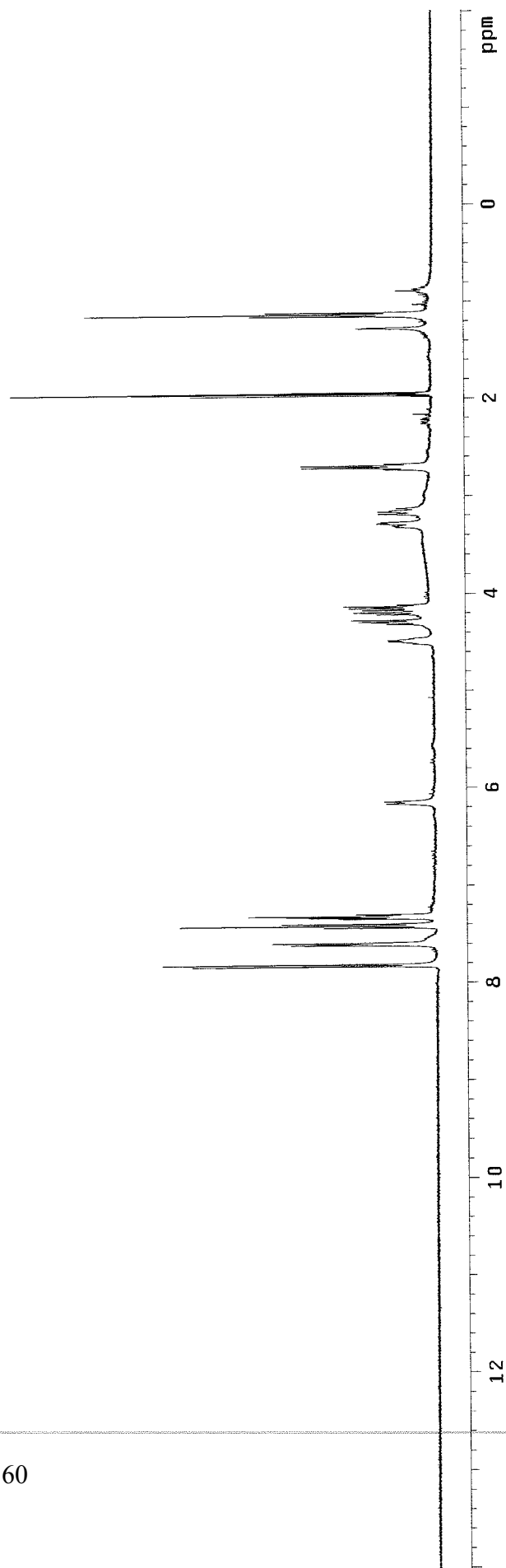


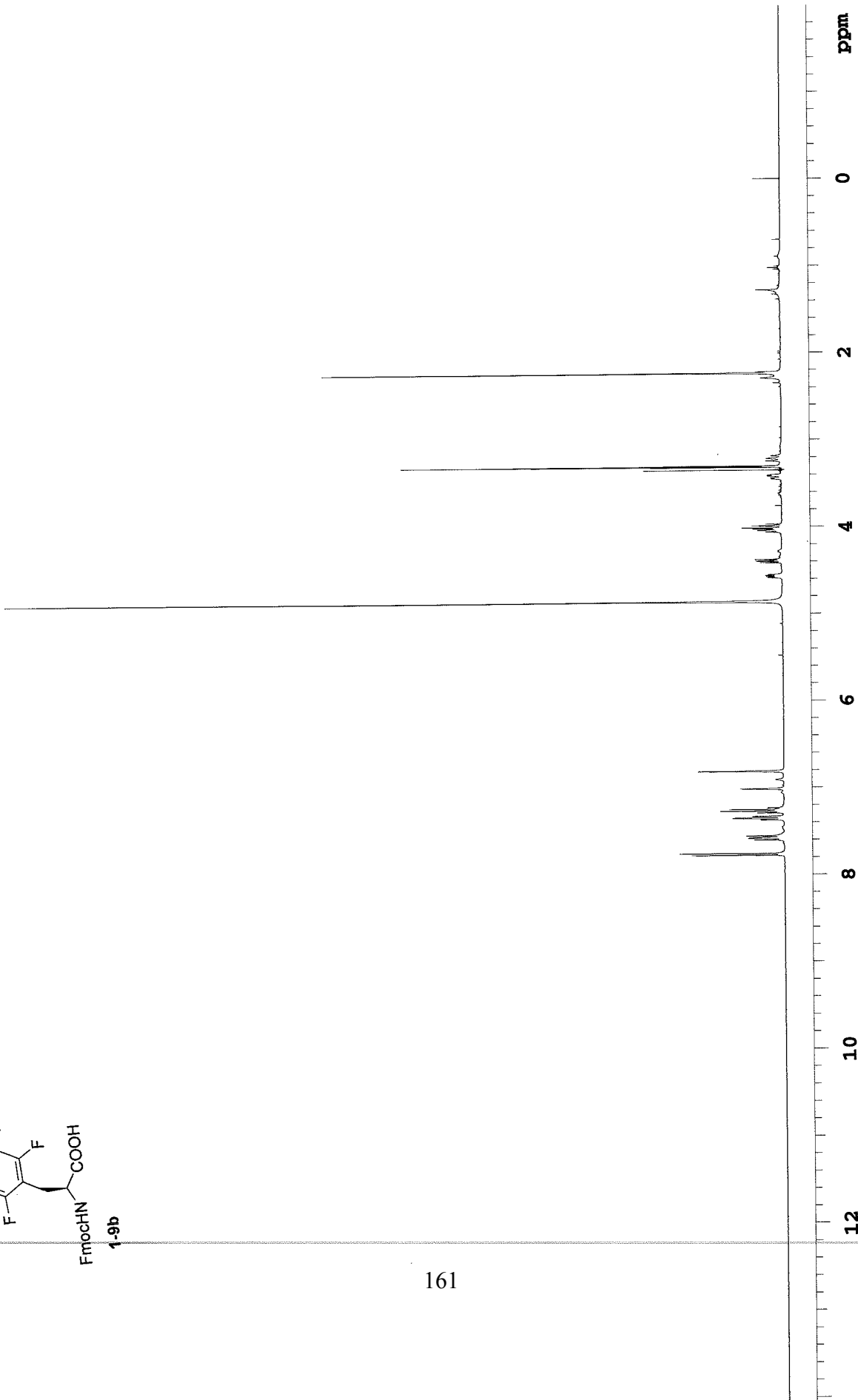
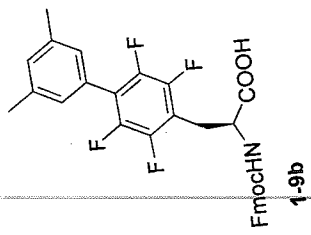


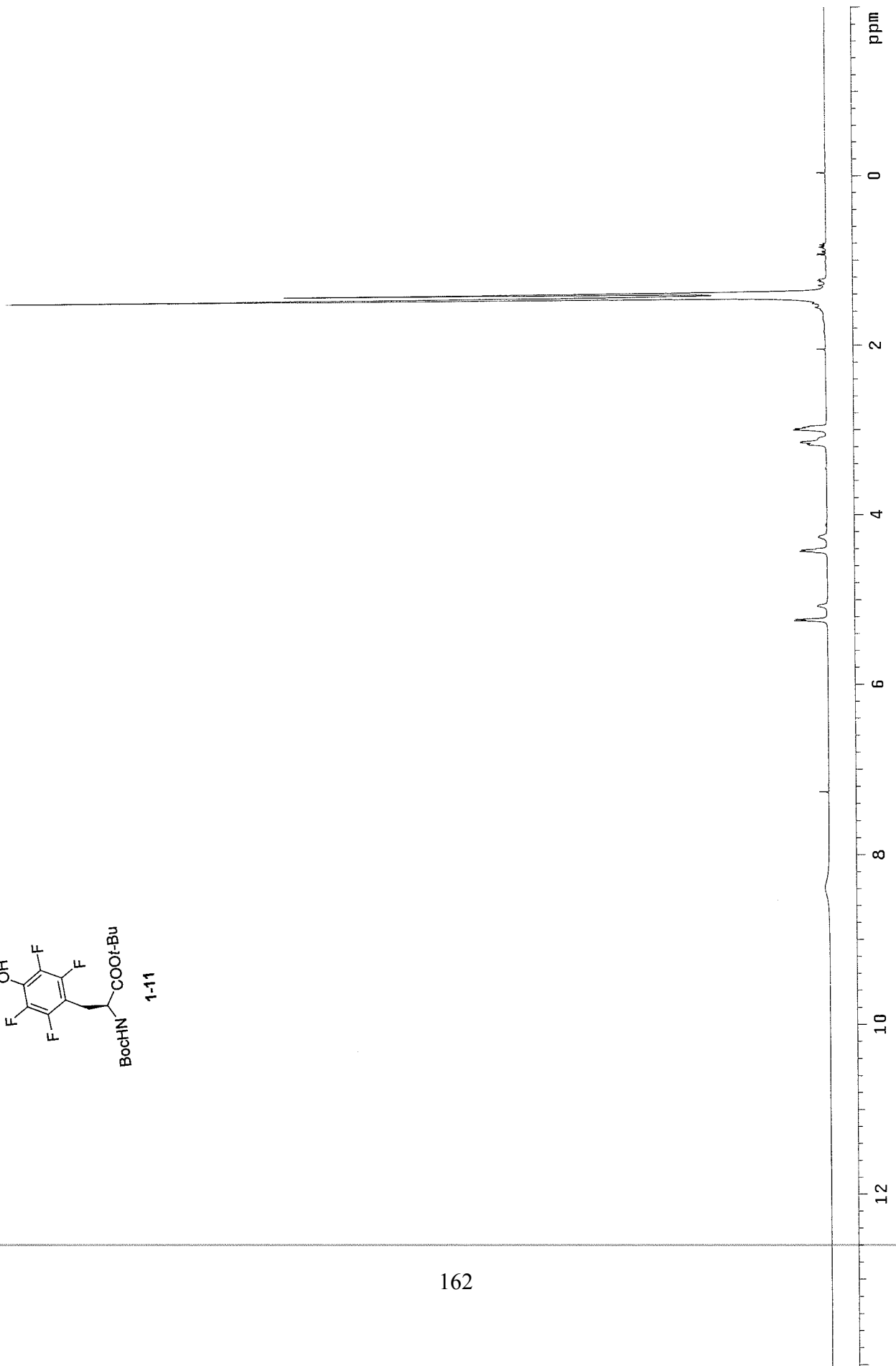
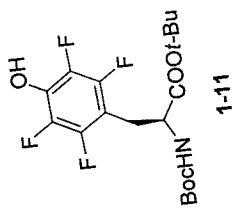


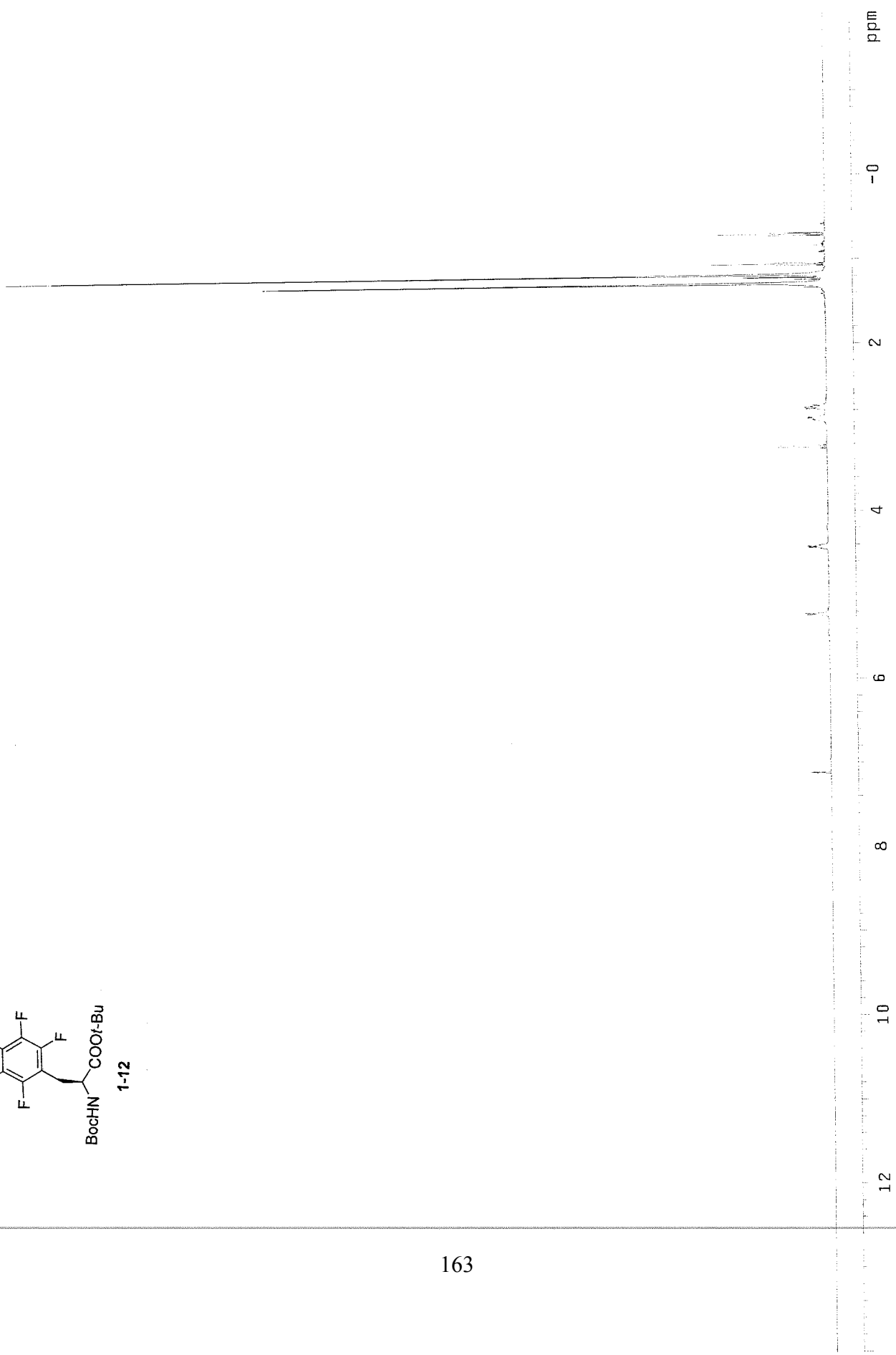
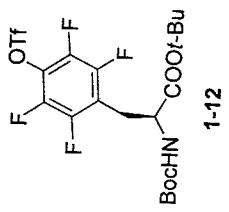


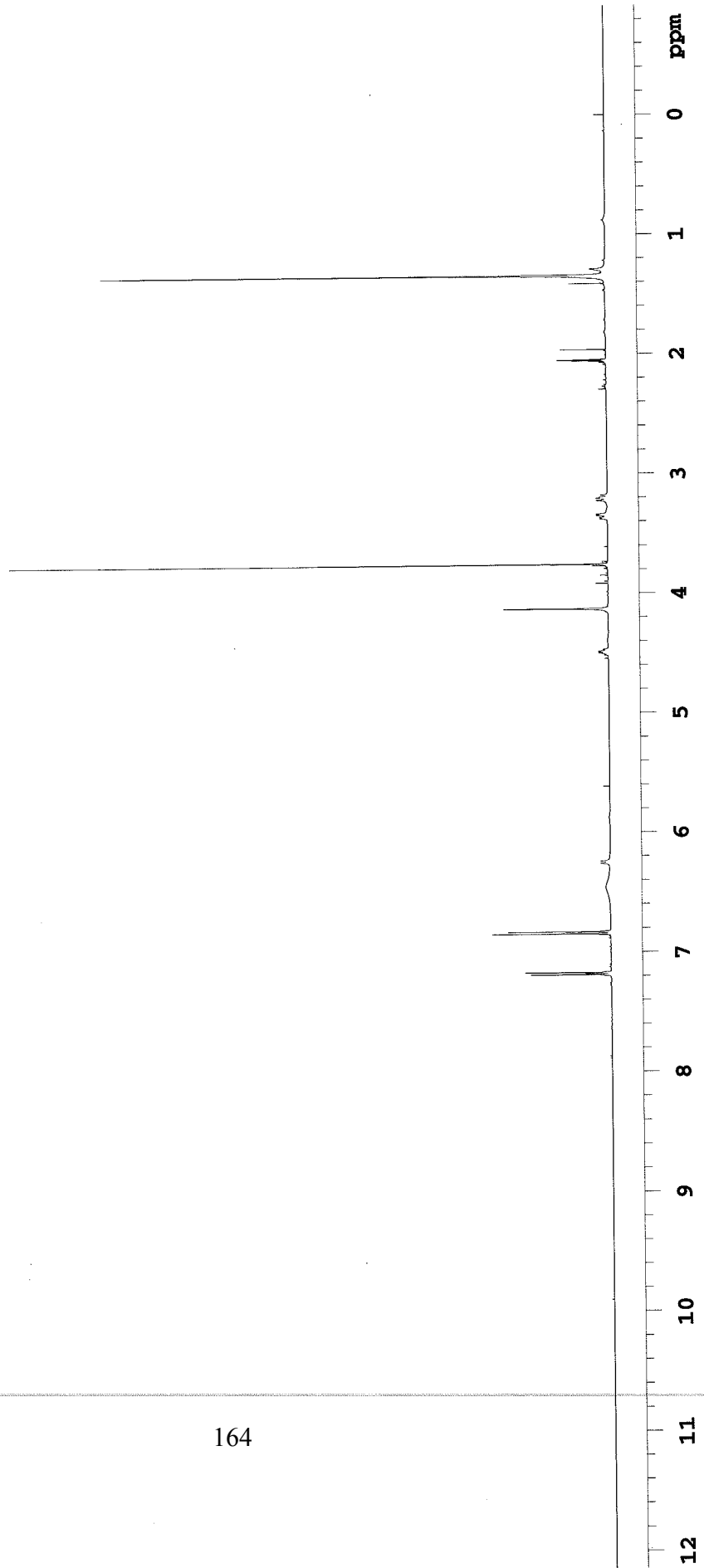
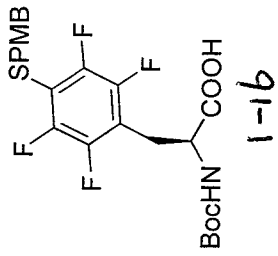
160

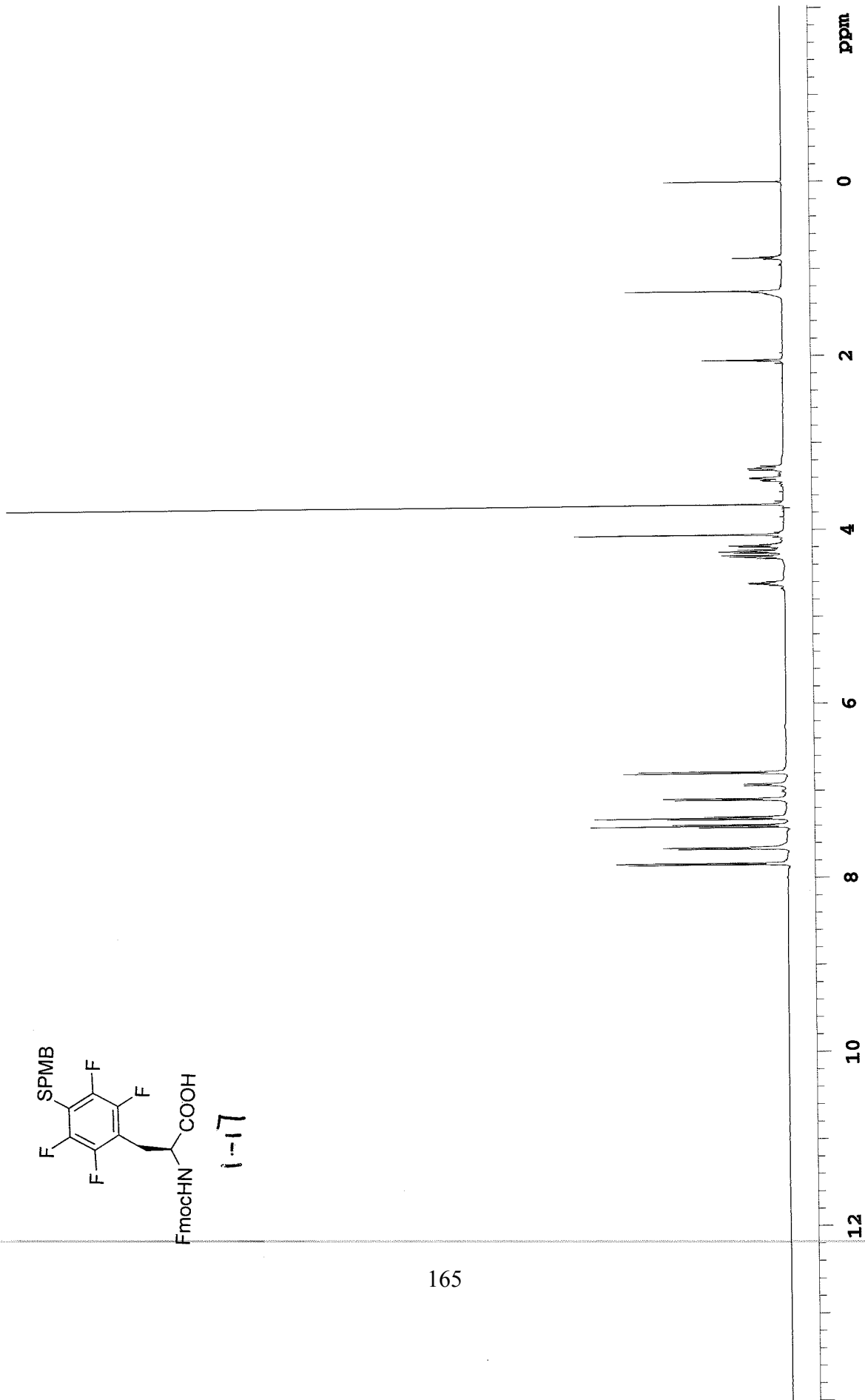
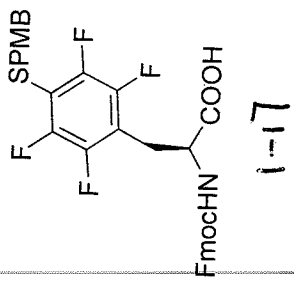




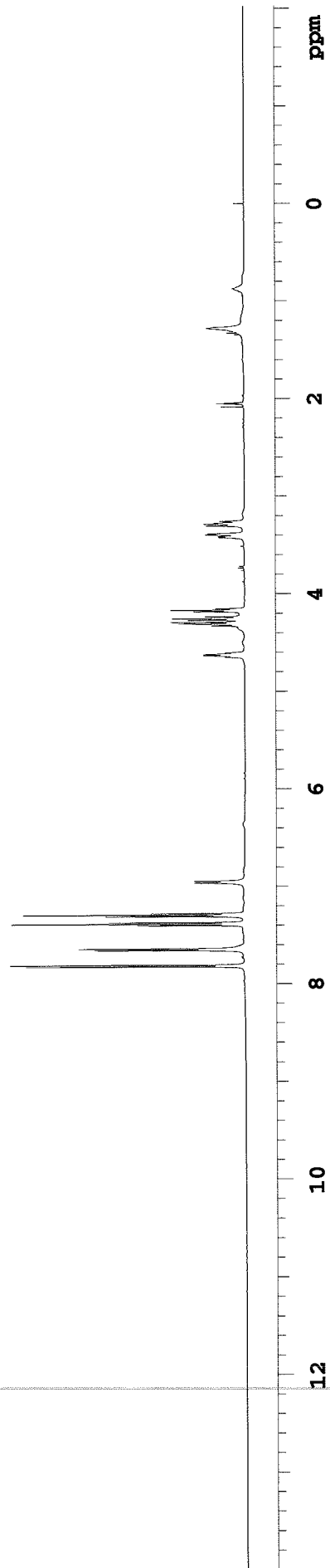
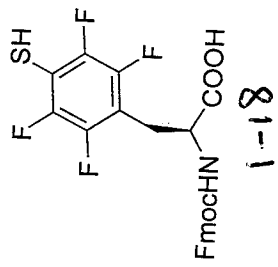


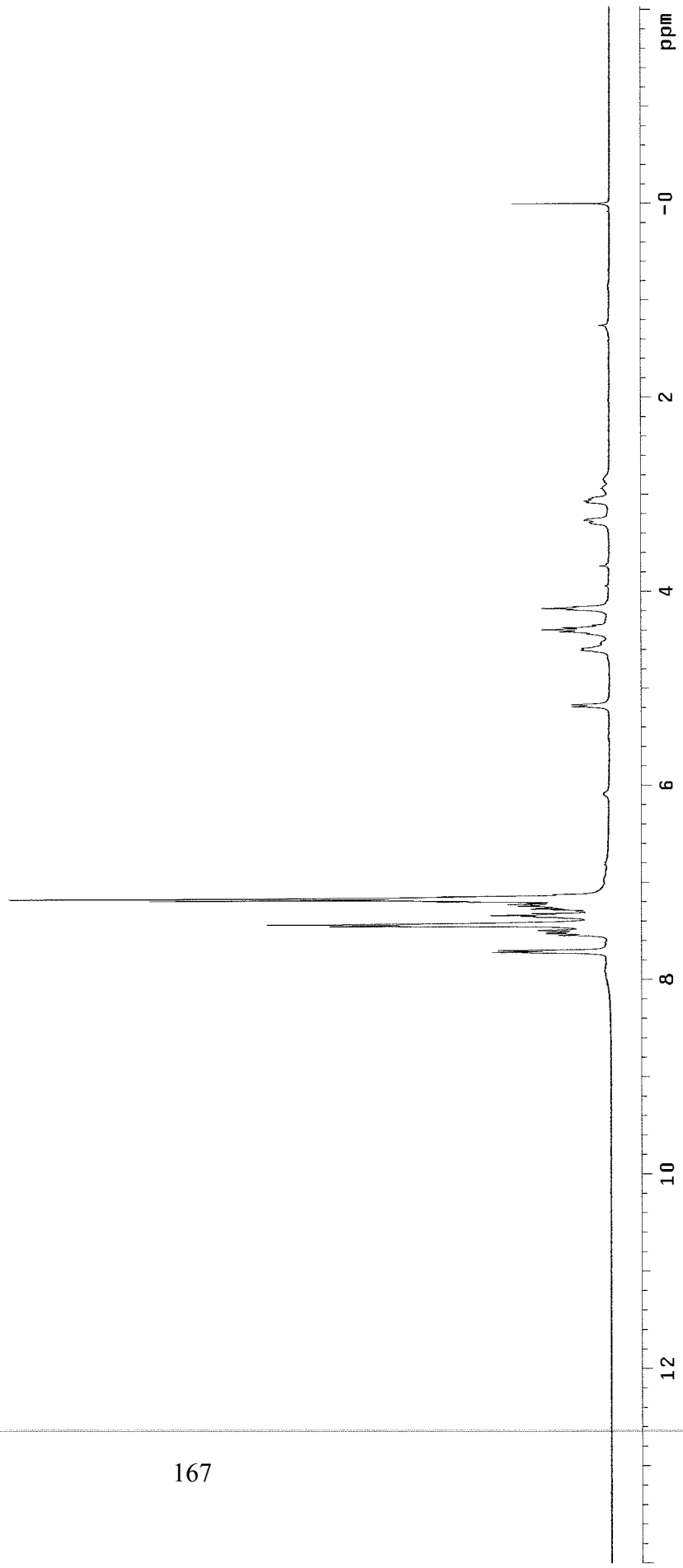
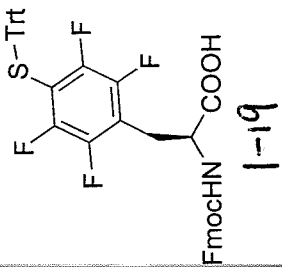


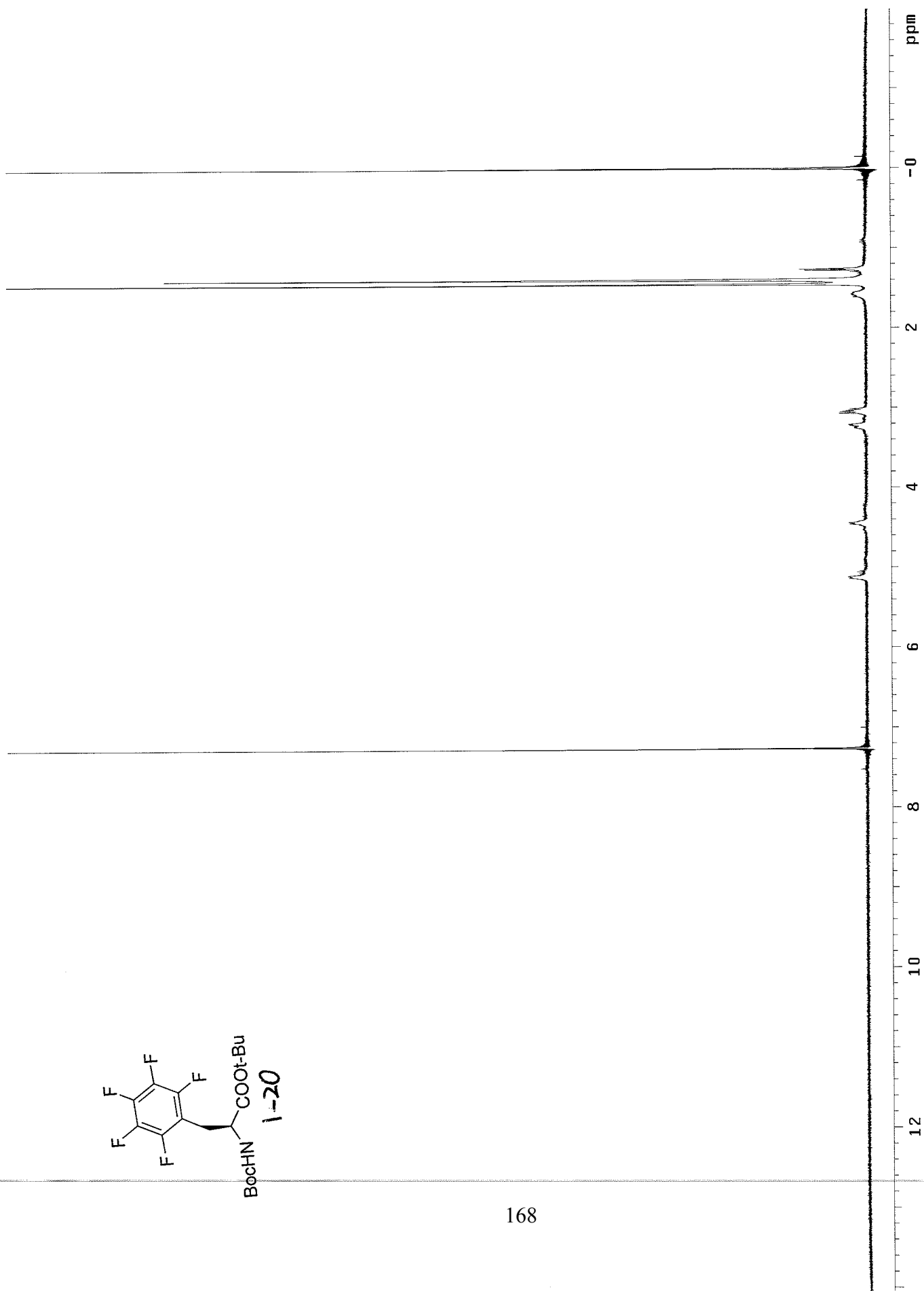
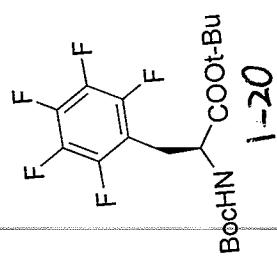


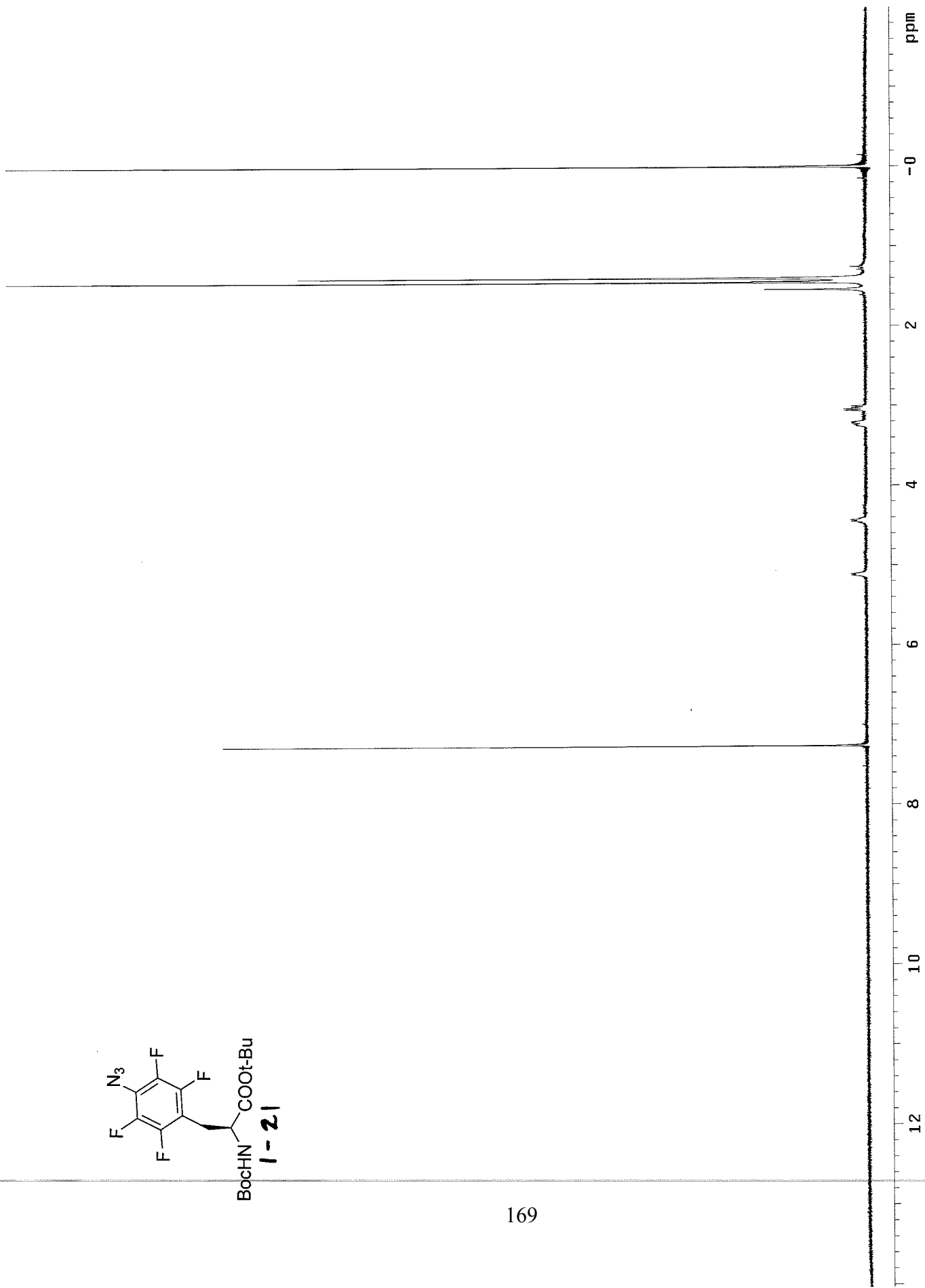
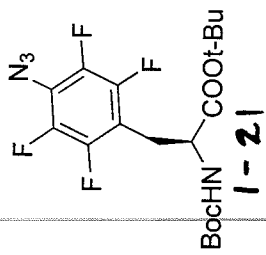


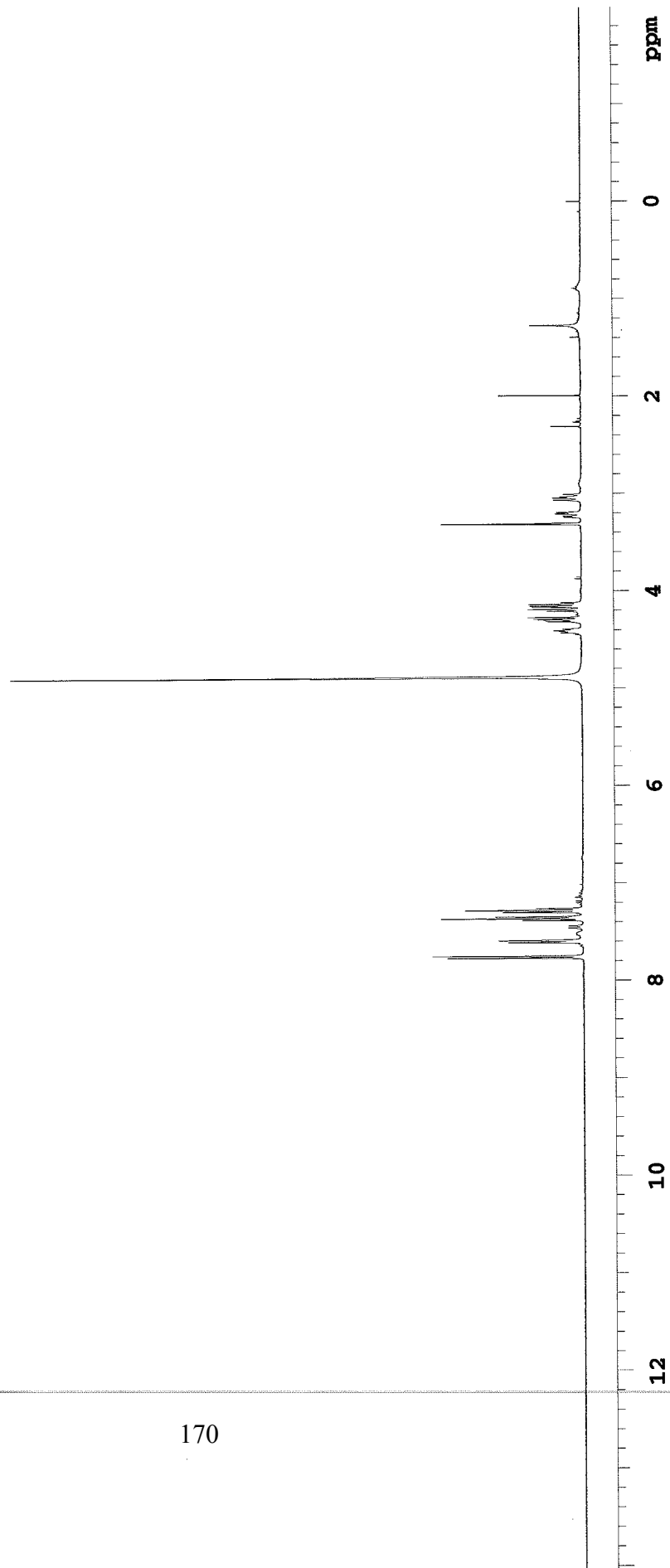
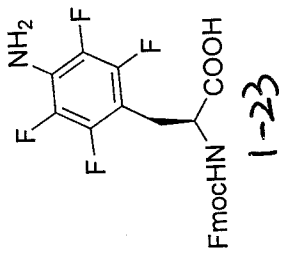


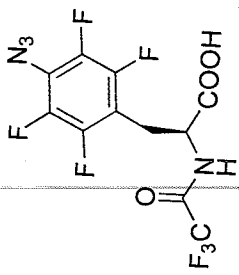






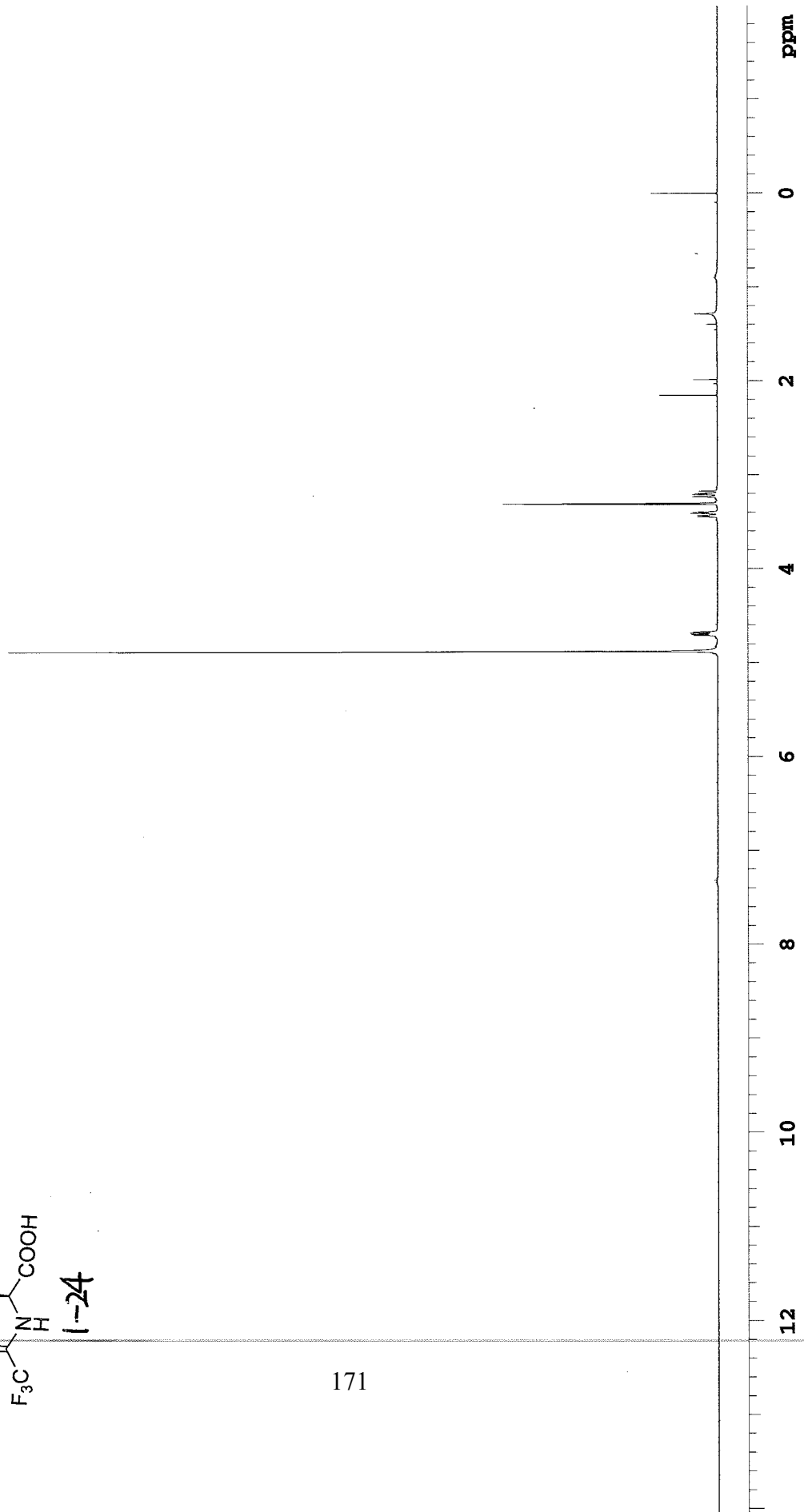


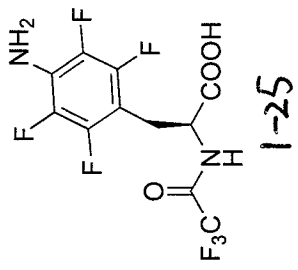




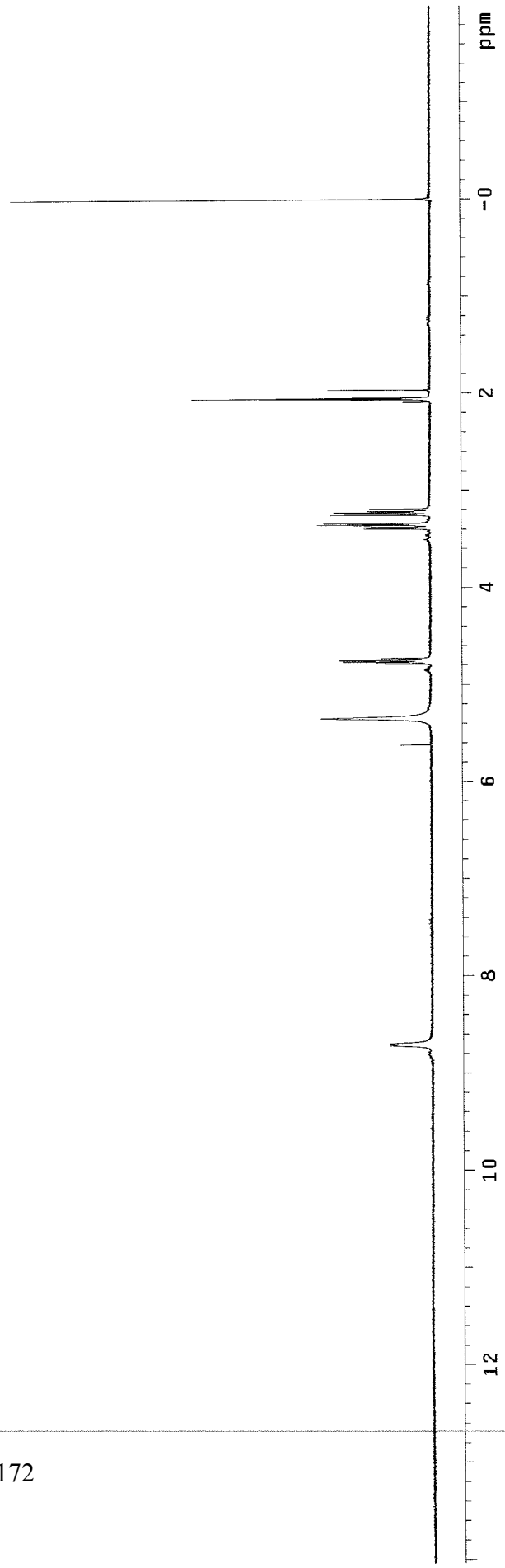
1-24

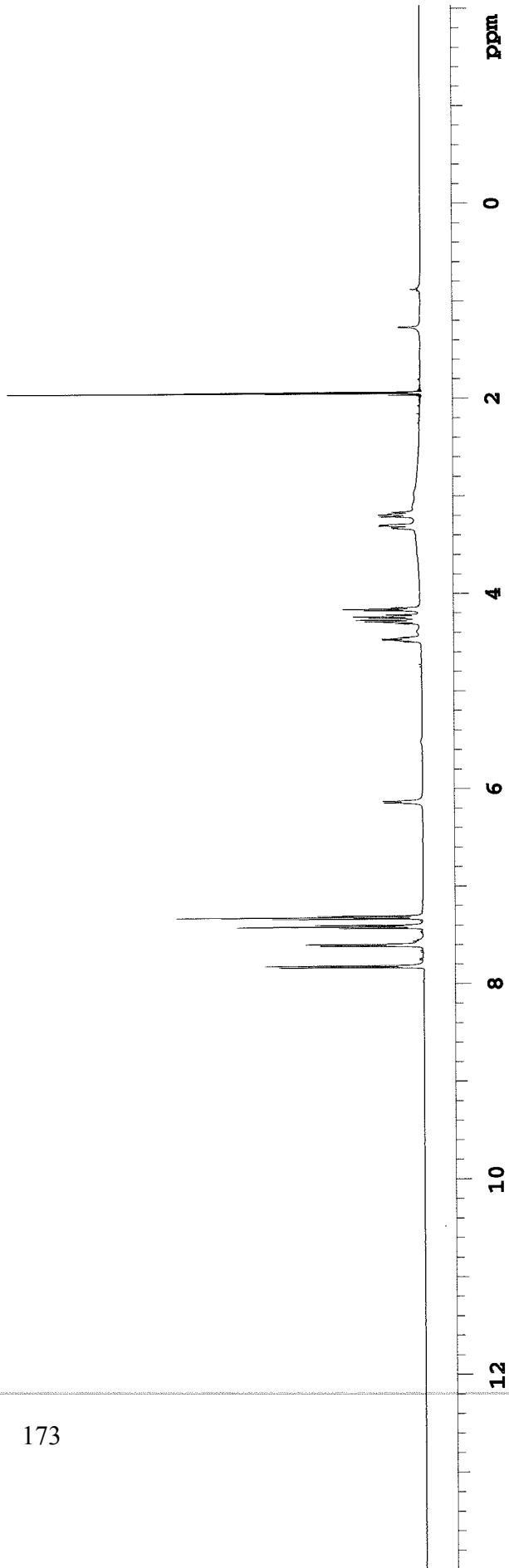
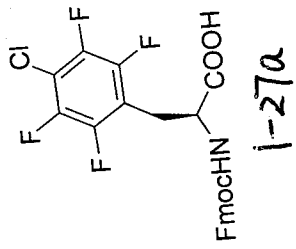
171



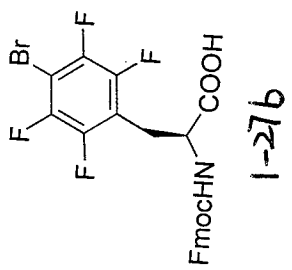


172

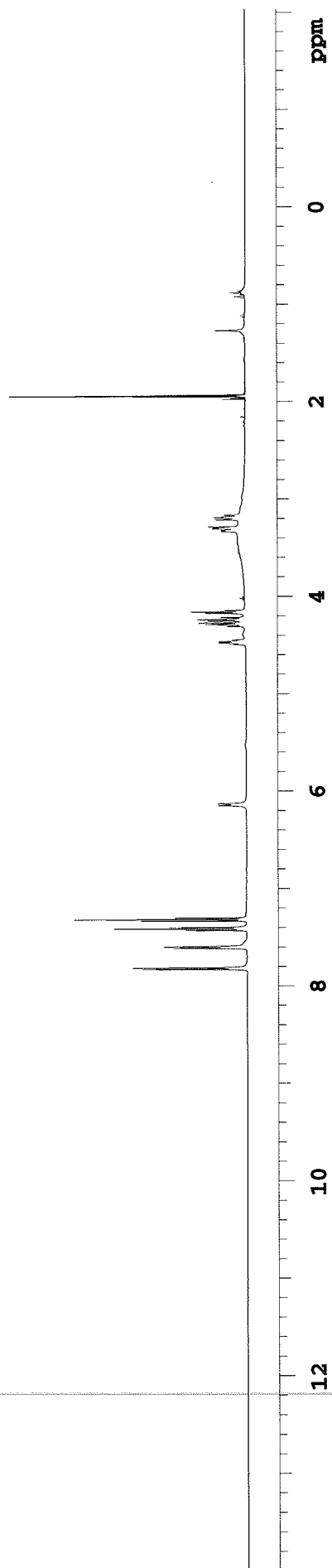


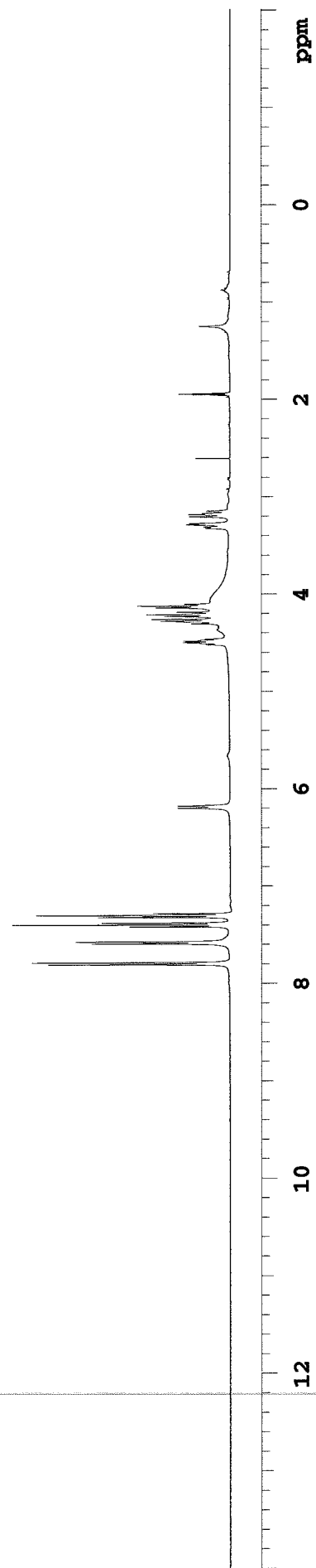
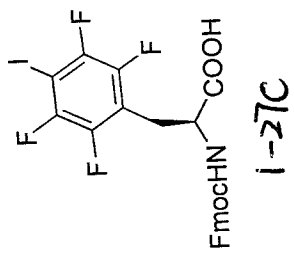


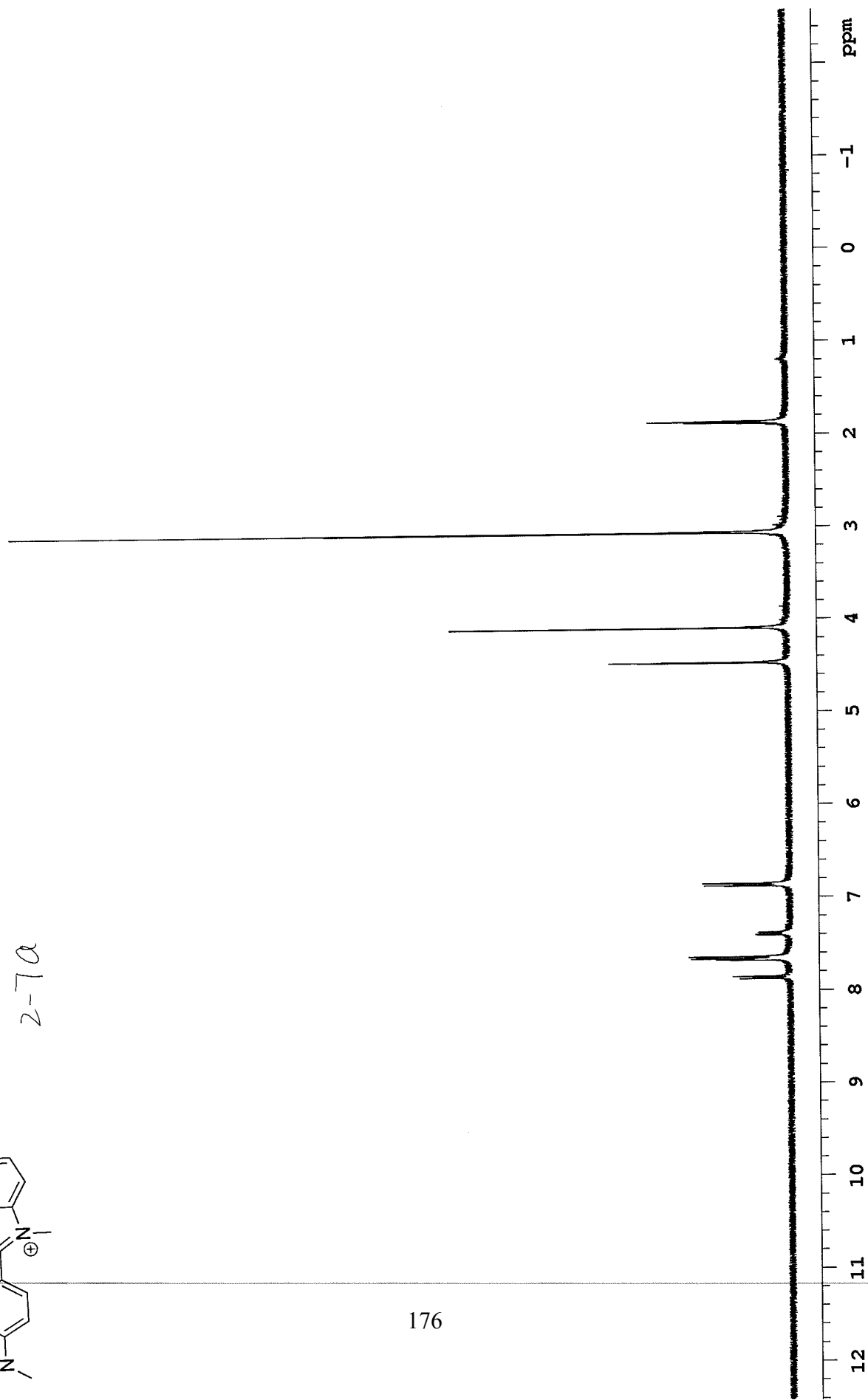
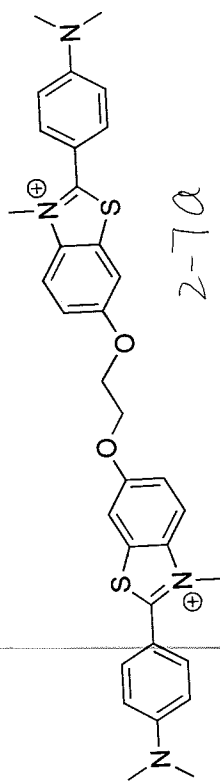


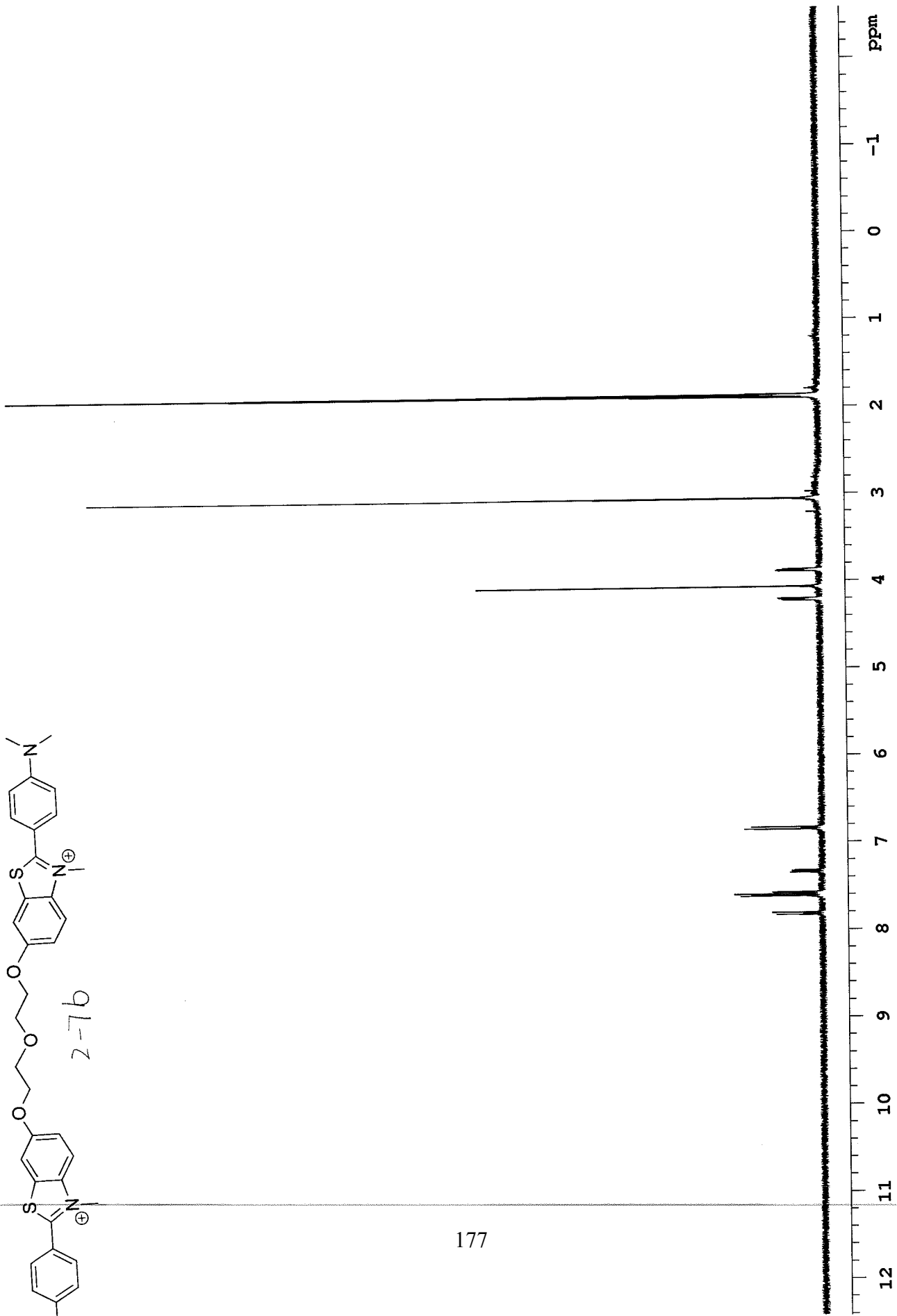
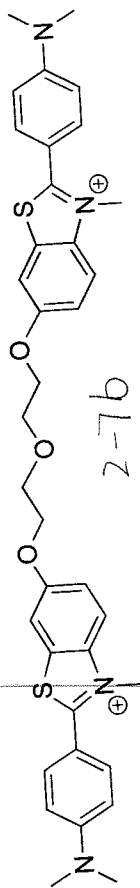


174





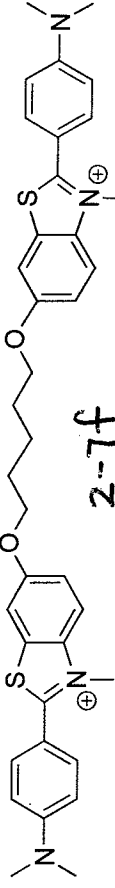












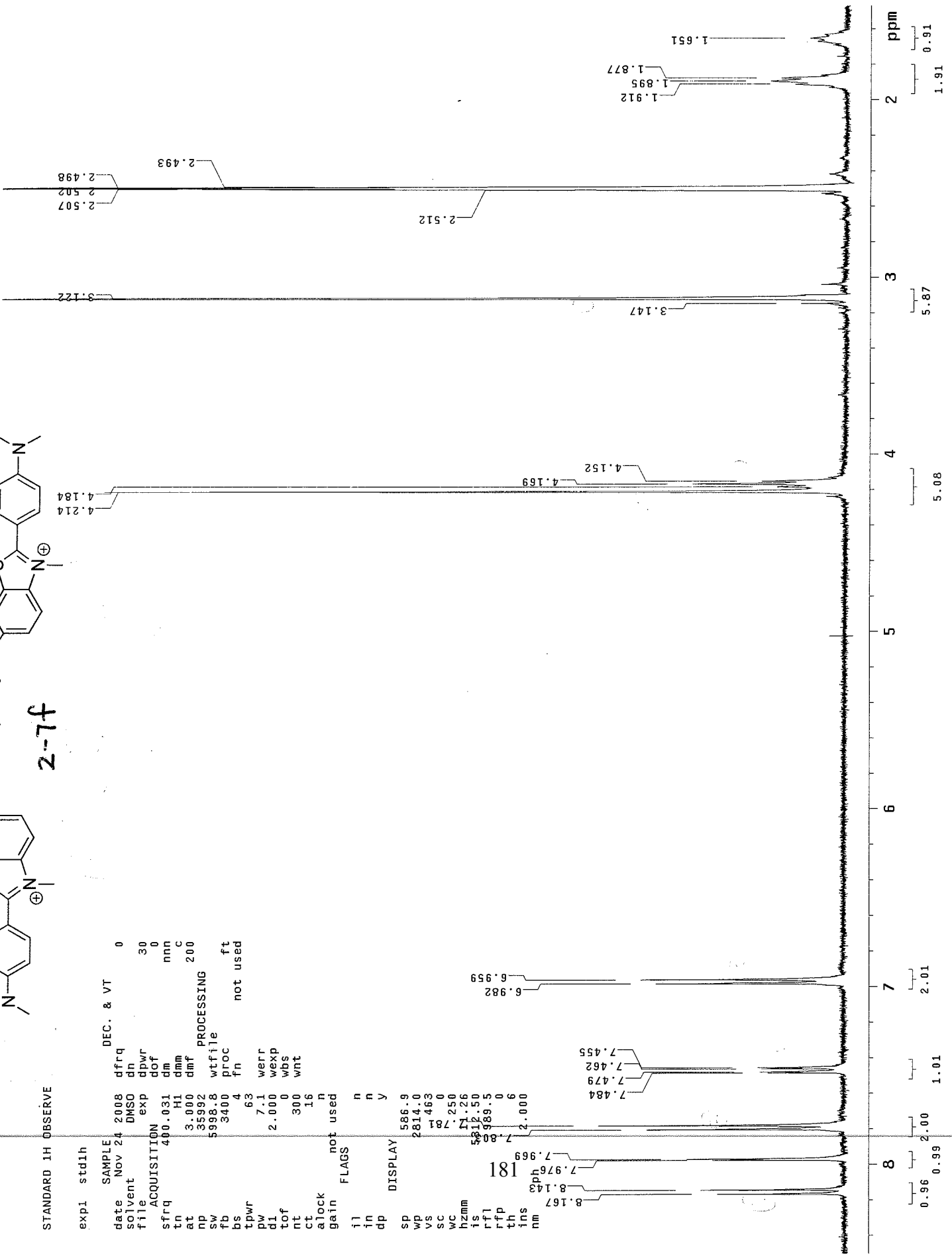
STANDARD 1H OBSERVE

```

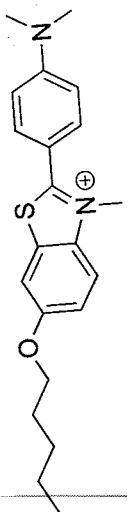
expl stdlh
SAMPLE
date Nov 24 2008
solvent DMSO
file
ACQUISITION
sfrq 400.031
in HI
at 3.000
np 35992
sw 5998.8
fb 3400
bs 4
tpwr 63
pw 7.1
d1 2.000
tof 0
nt 300
ct 16
gain not used
alock n
flags n
DISPLAY
sp 586.9
wp 2814.0
vs 463
sc 781
wc 250
hzmm 11.26
is 12.50
rfl 989.5
rff 0
th 6
ins 2.000
nm 181
  
```

```

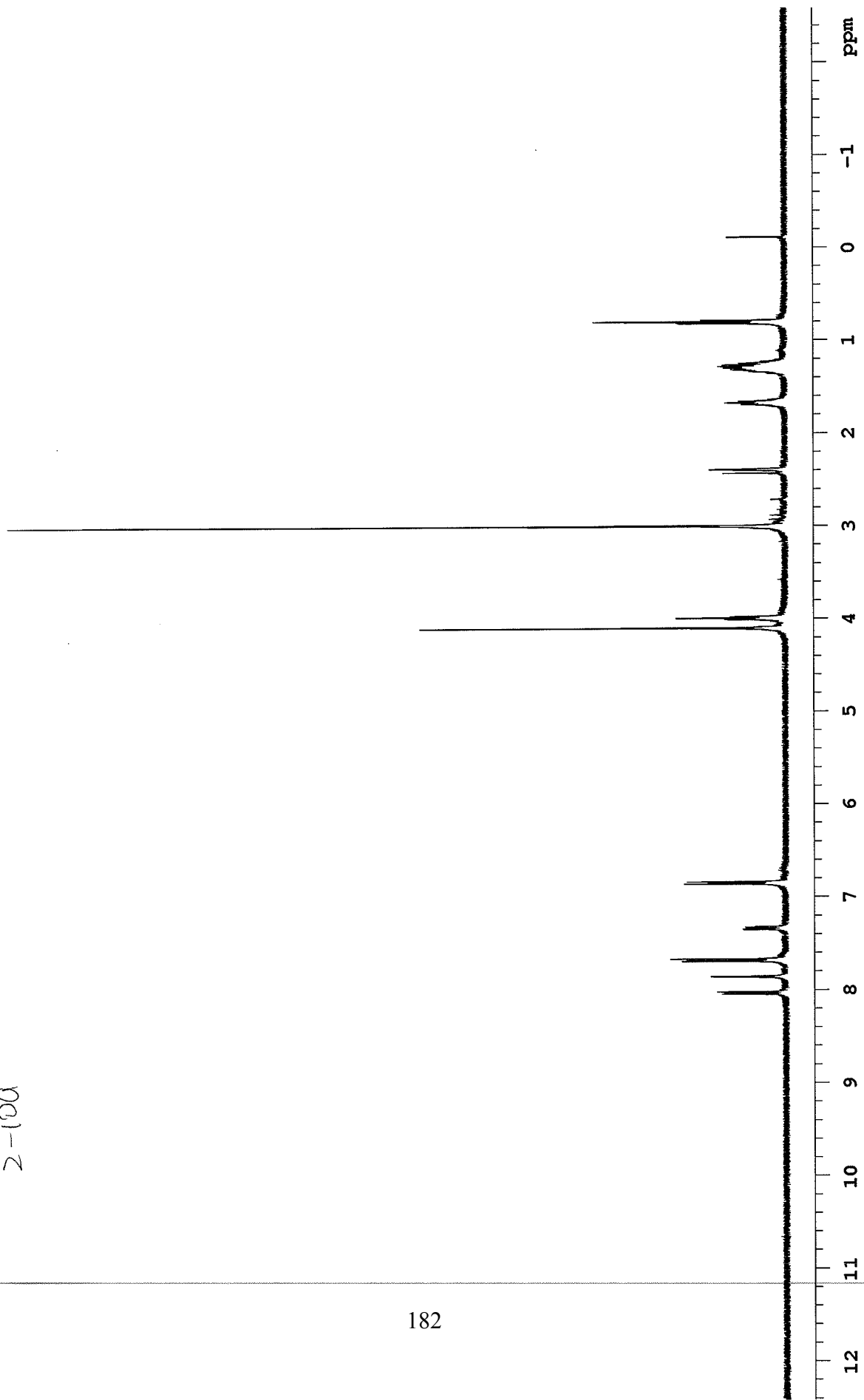
DEC. & VT 0
dfrq 0
dn 30
dpwr 0
dof 0
dm nnn
dmm C
dmf 200
wf file
proc ft
fn not used
  
```

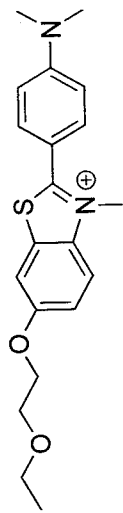




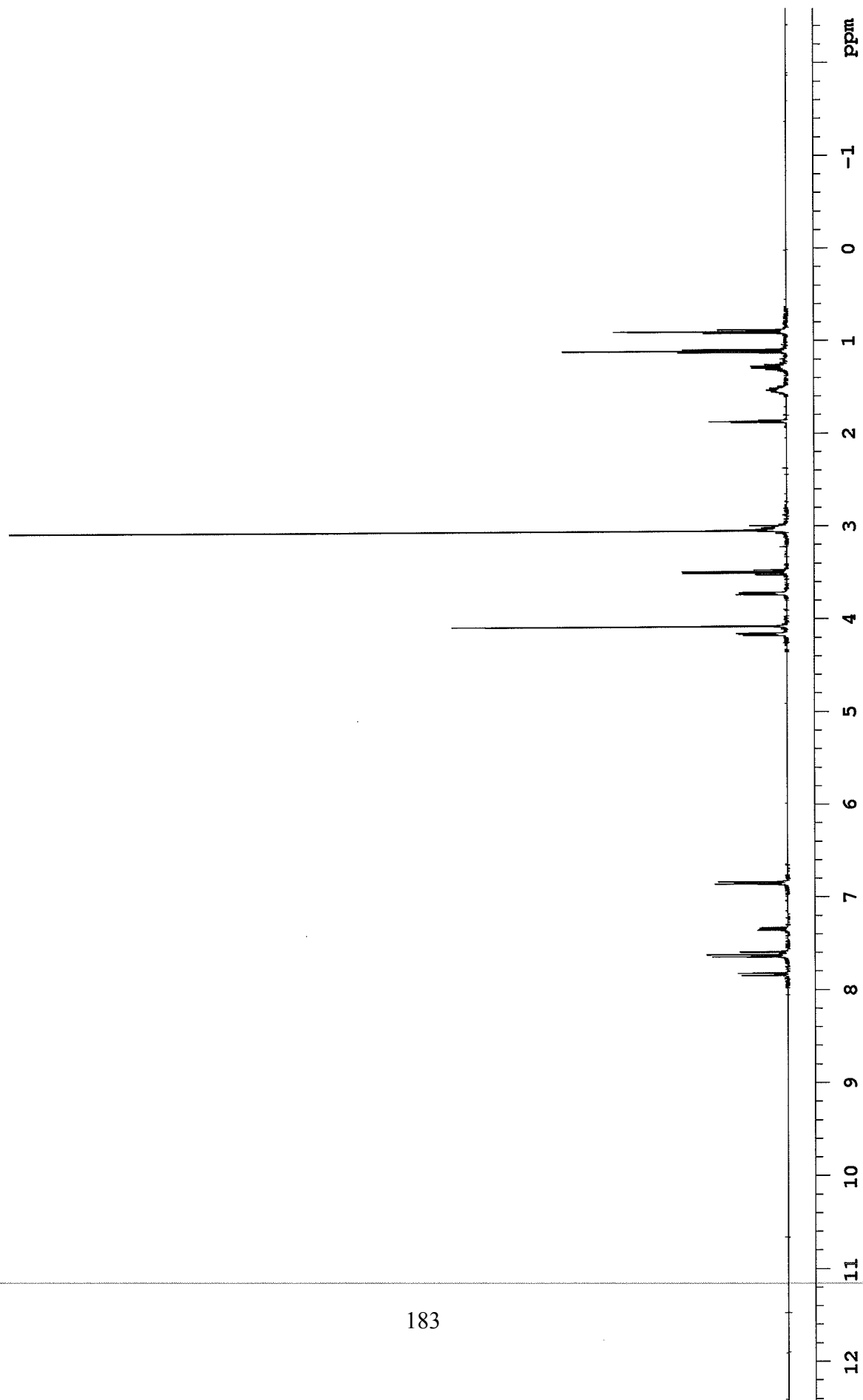


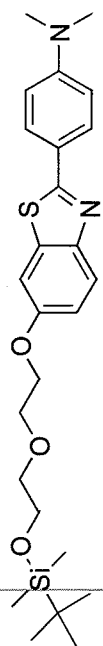
2-100





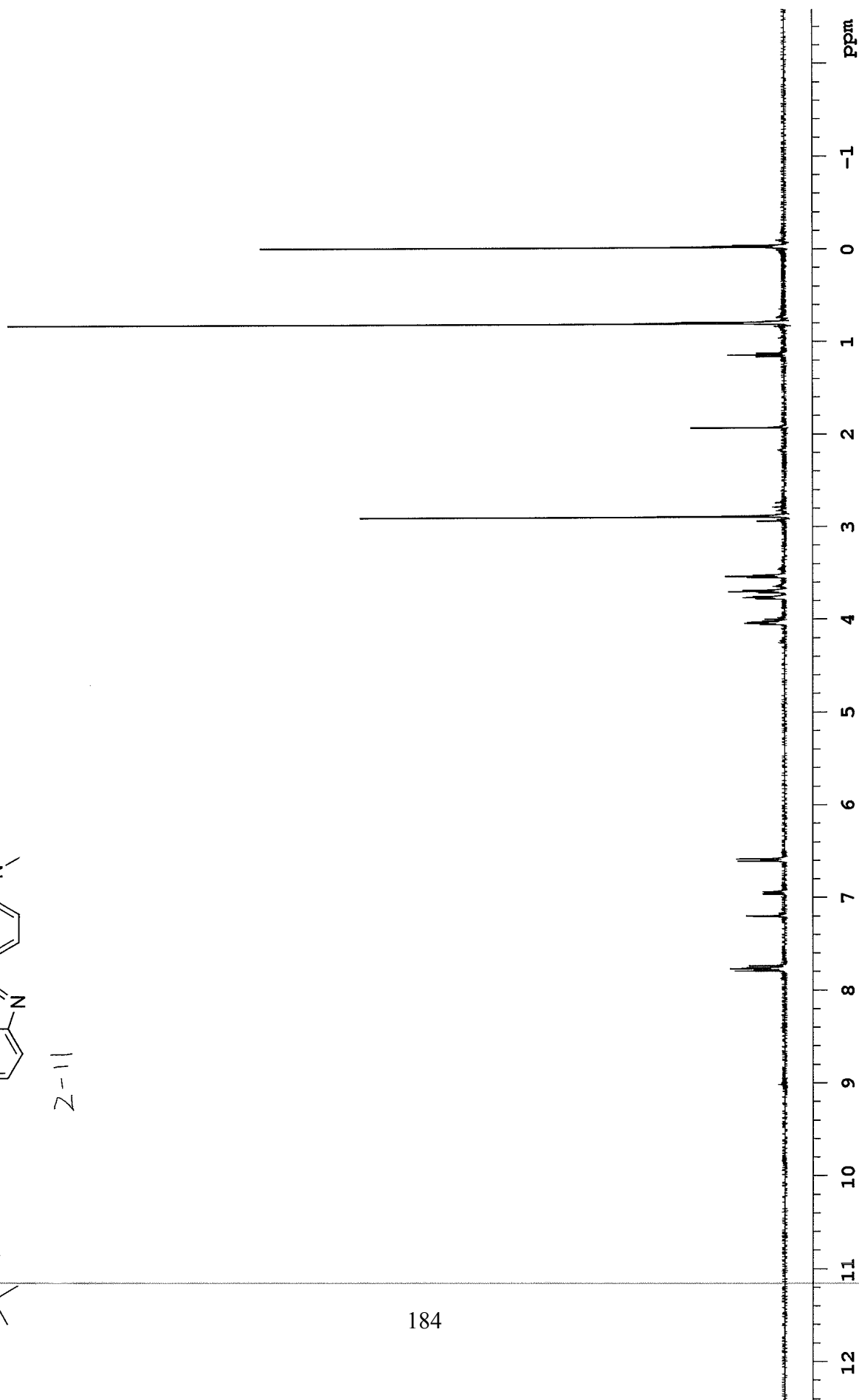
Z-10b

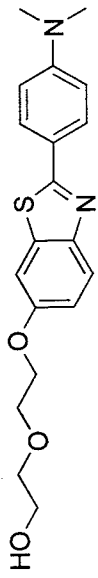




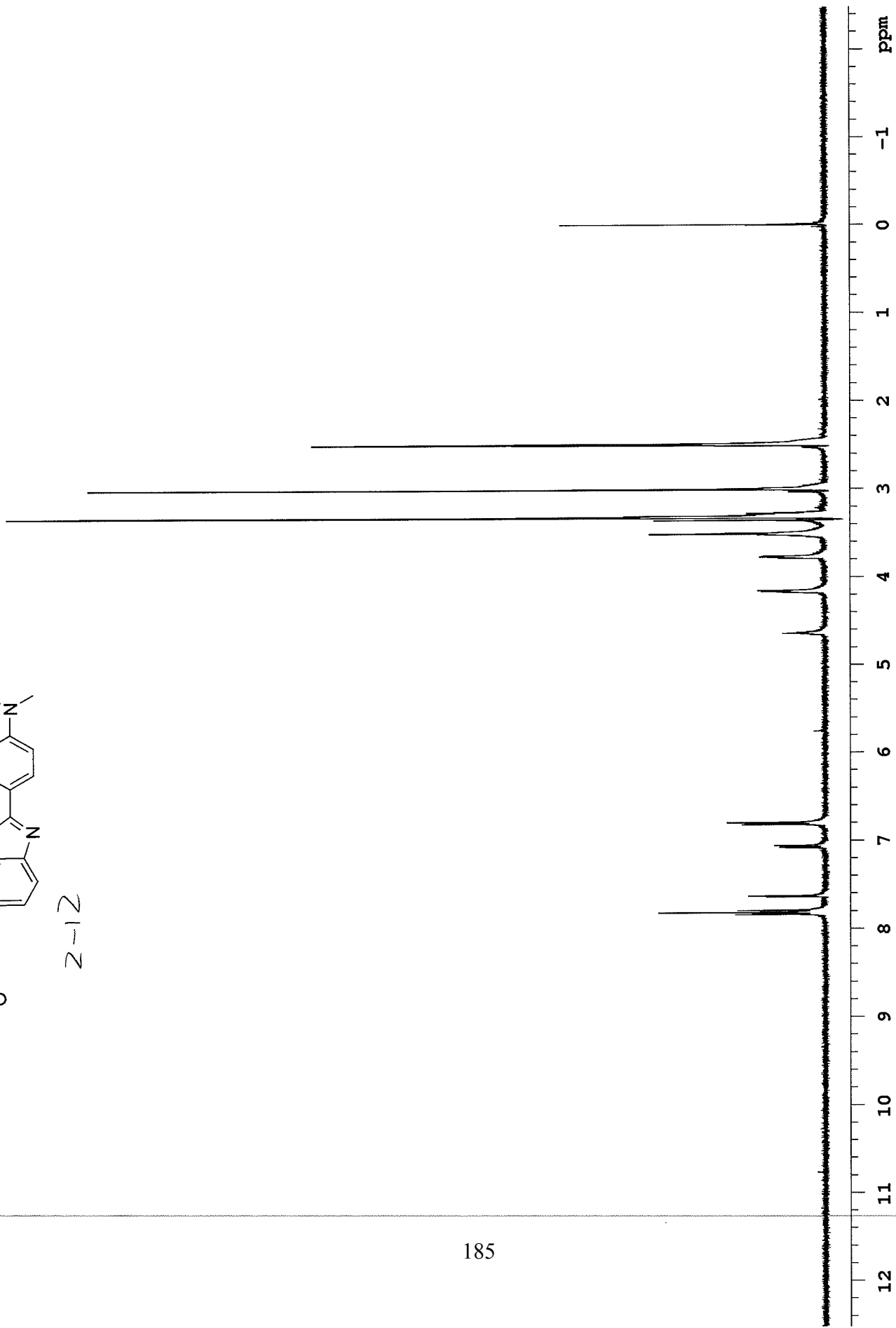
Z-11

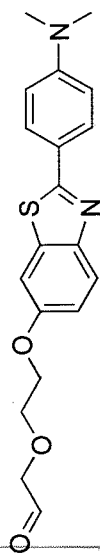
184



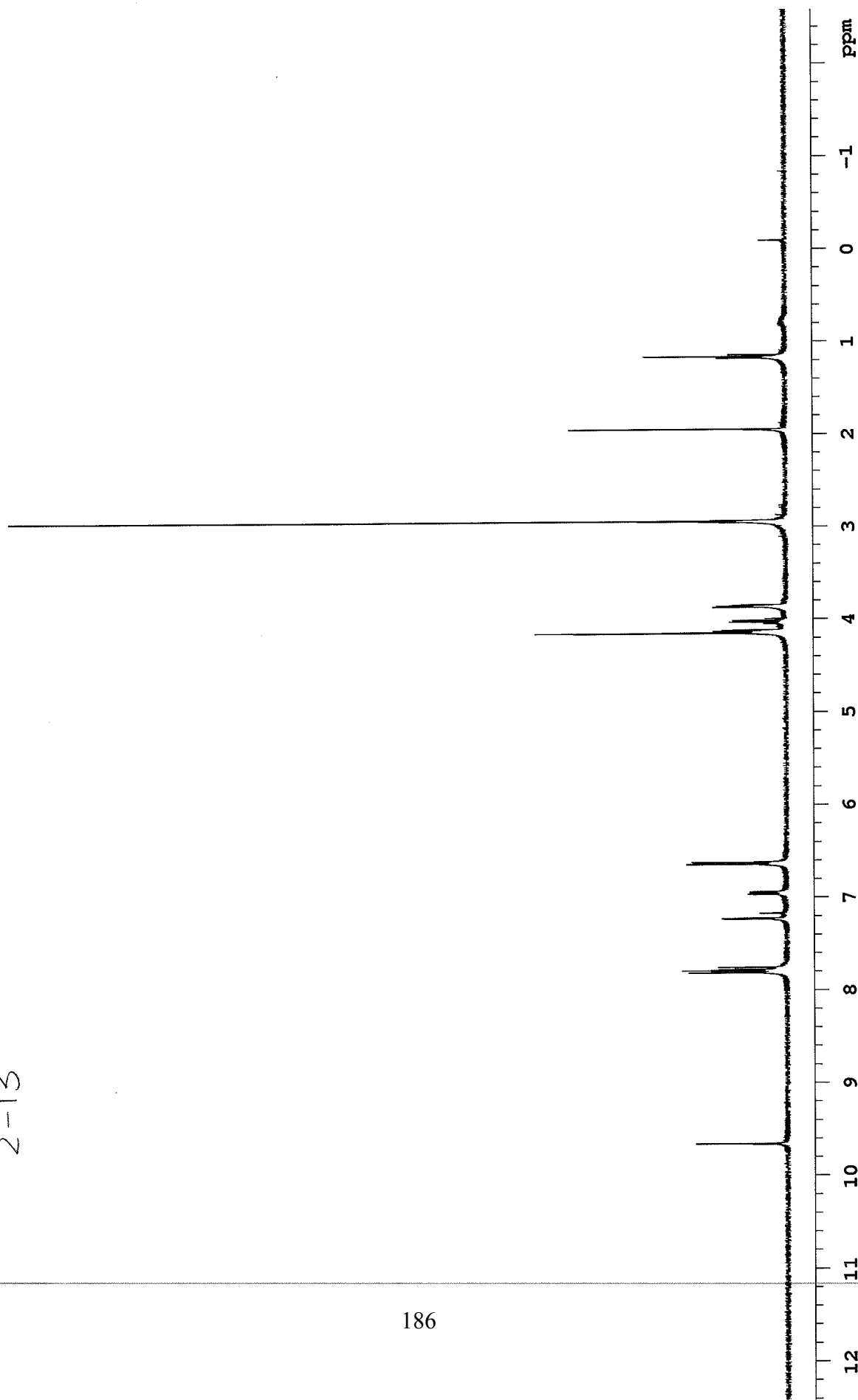


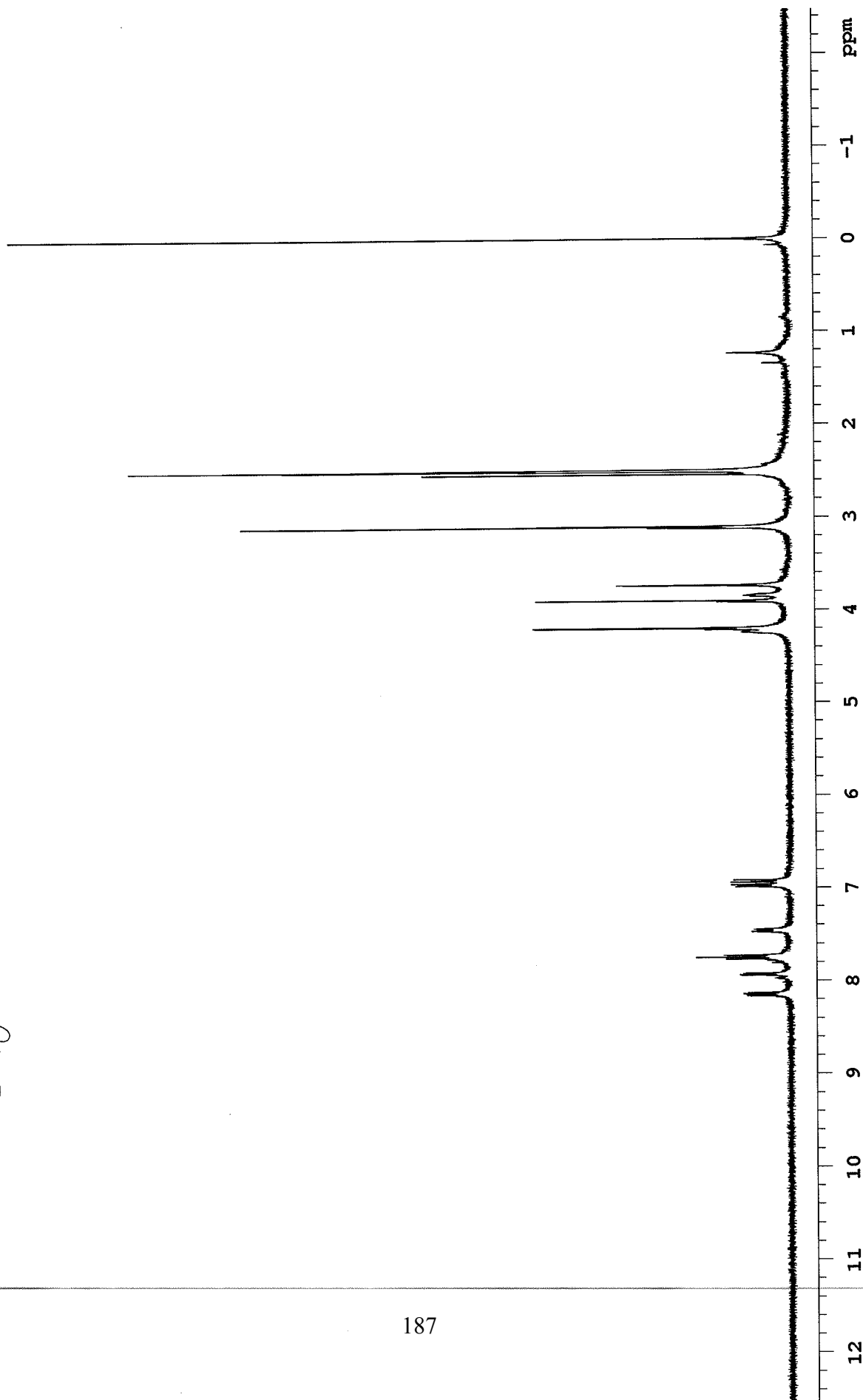
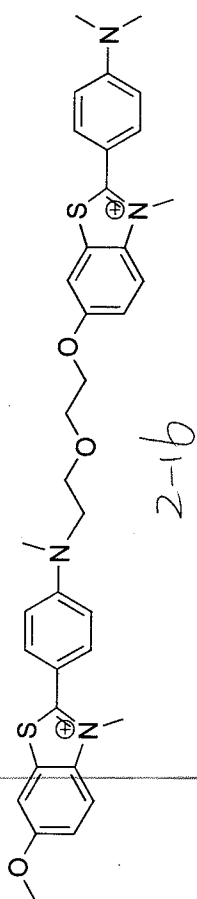
2-12

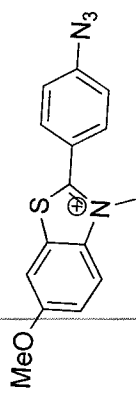




2-13



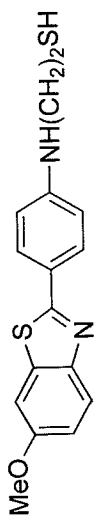




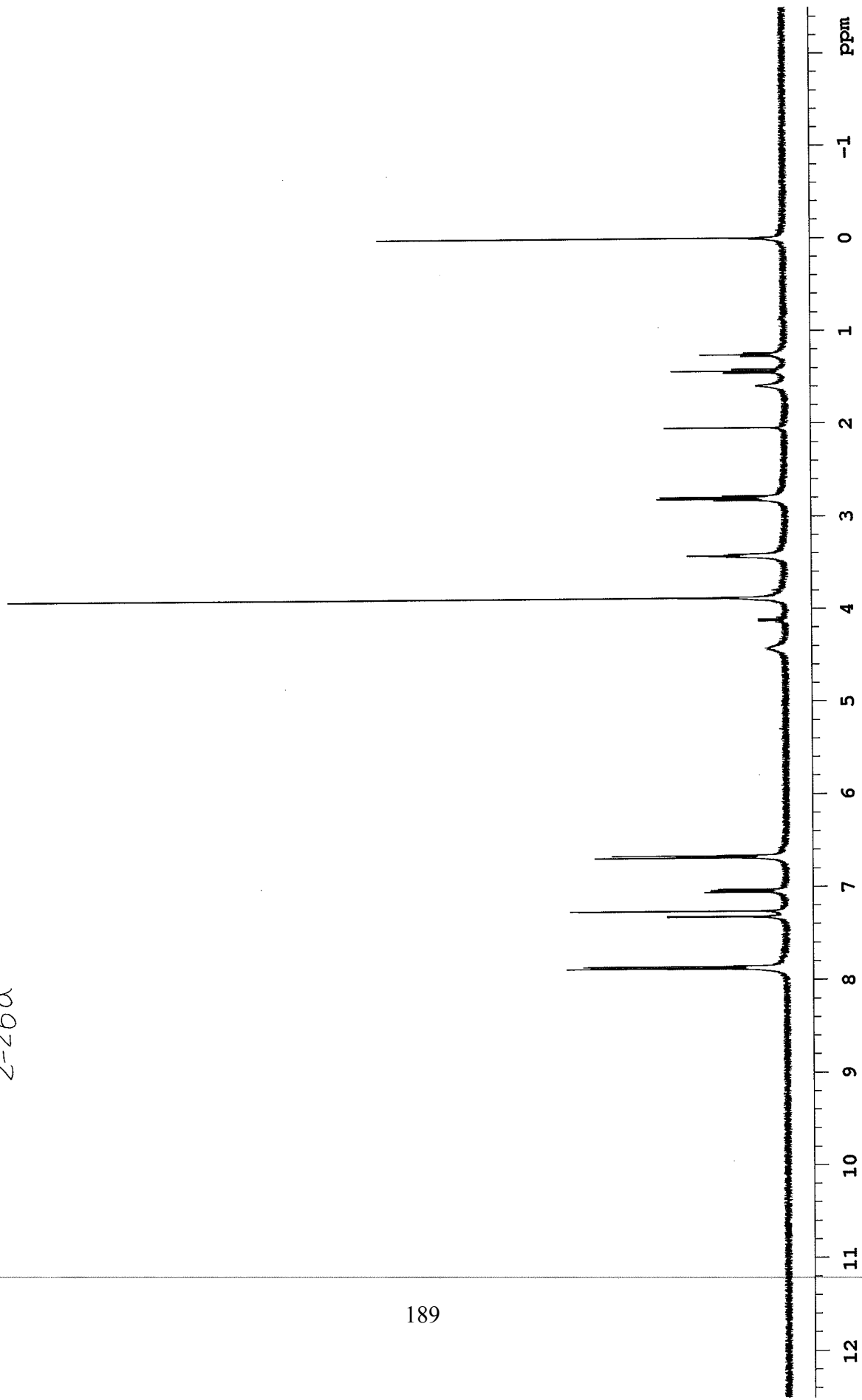
2-21

188

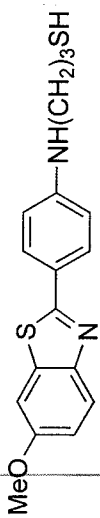




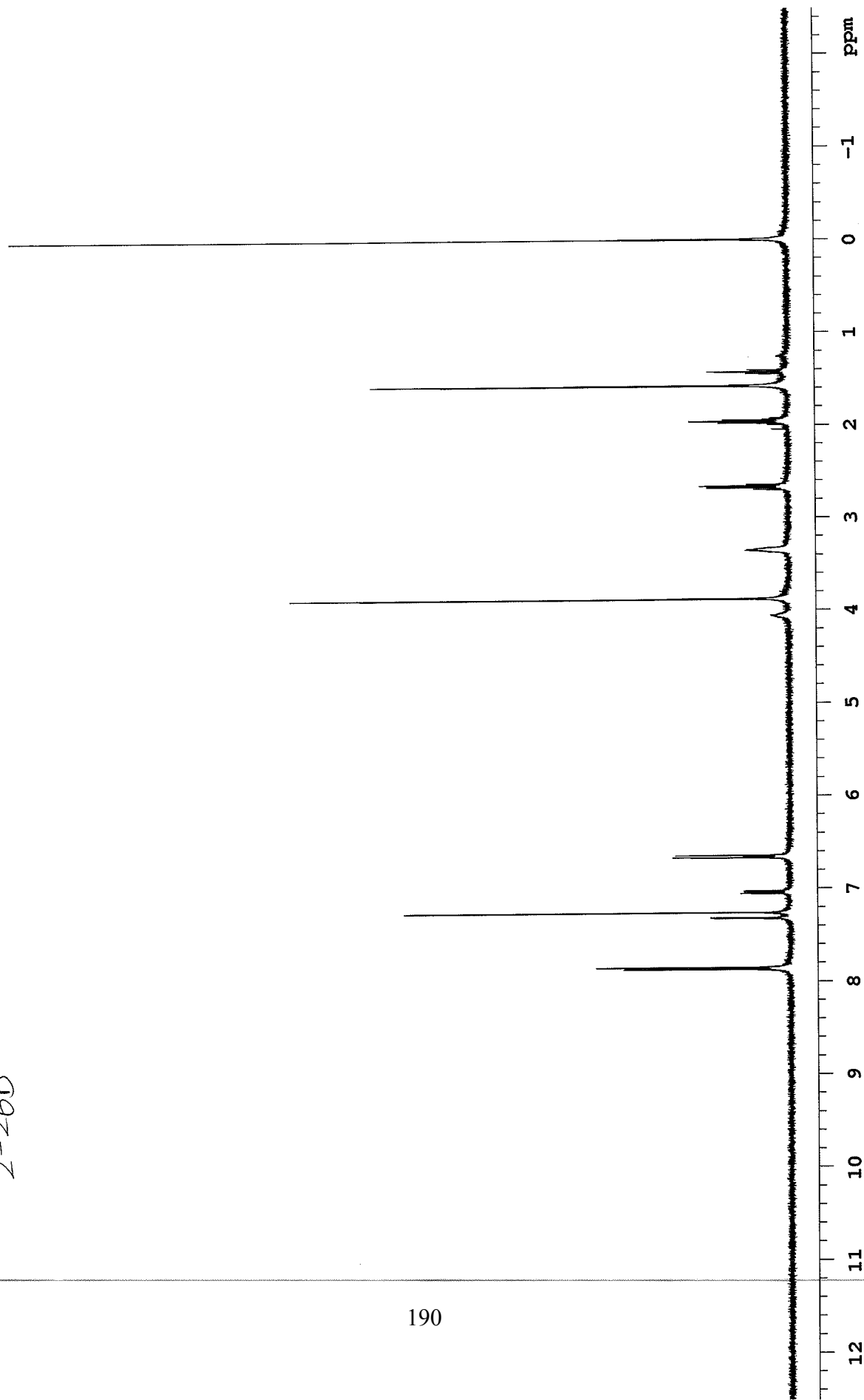
2-26a

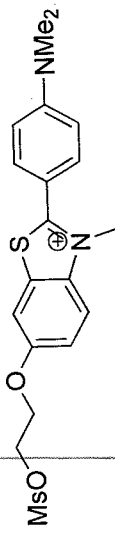






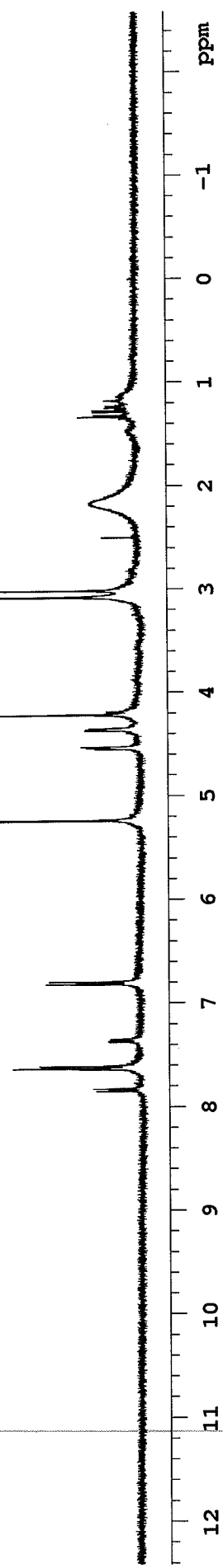
2-26b

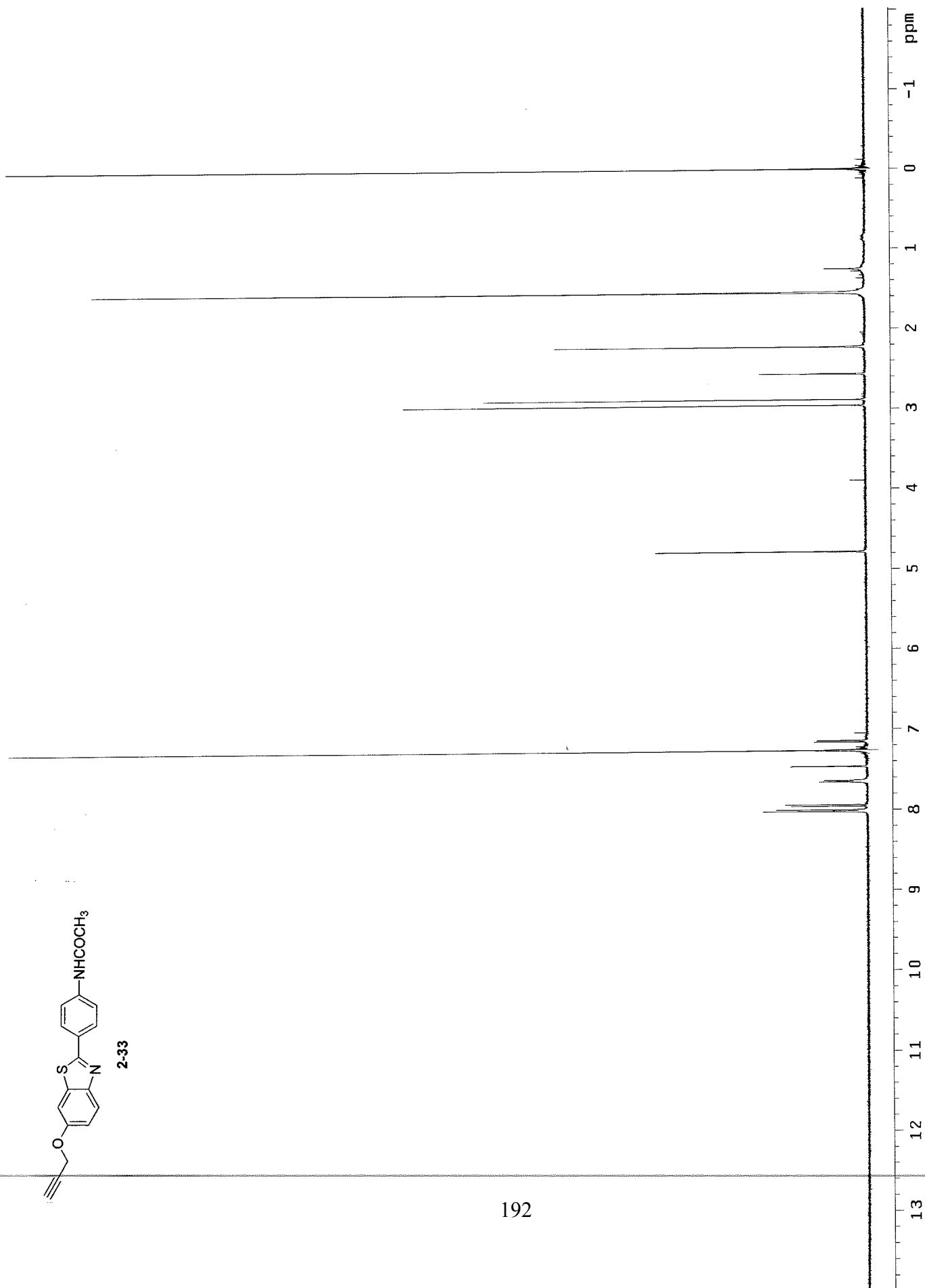
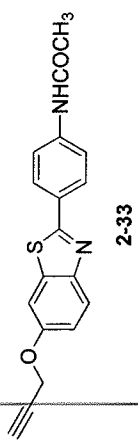




2-30

191



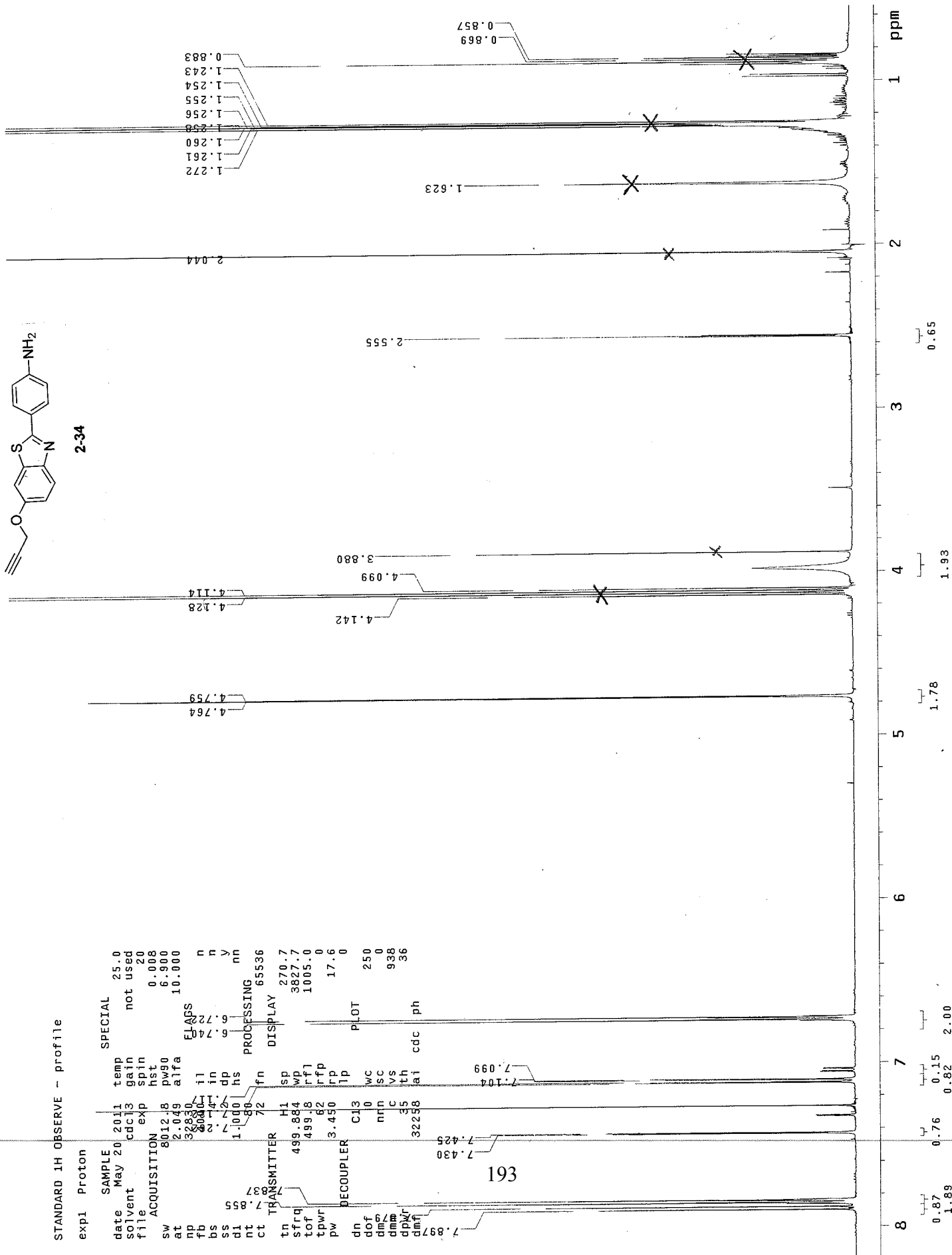


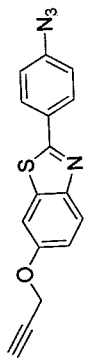
STANDARD 1H OBSERVE - profile

exp1 Proton

SAMPLE SPECIAL 25.0  
 date May 20 2011 temp not used  
 solvent cdc13 gain 20  
 file exp spin 0.008  
 ACQUISITION hst 6.900  
 sw 8012.8 pw90 10.000  
 at 2.049 alfa  
 np 32830 il  
 fb 22614 in  
 bs 1.174 dp  
 ss 7.240 hs  
 di 1.090 fn  
 nt 7.895  
 ct 837  
 TRANSMITTER H1 sp 270.7  
 tn 499.884 wp 3827.7  
 sfrq 499.8 rfl 1005.0  
 tof 62  
 tpwr 3.450 tp 17.6  
 pw DECOUPLER C13  
 up PLOT  
 dof 250  
 dms 0  
 dnm 938  
 dpr 35  
 dm1 32238 ai cdc ph  
 7.897  
 7.430  
 7.425  
 7.104  
 7.099

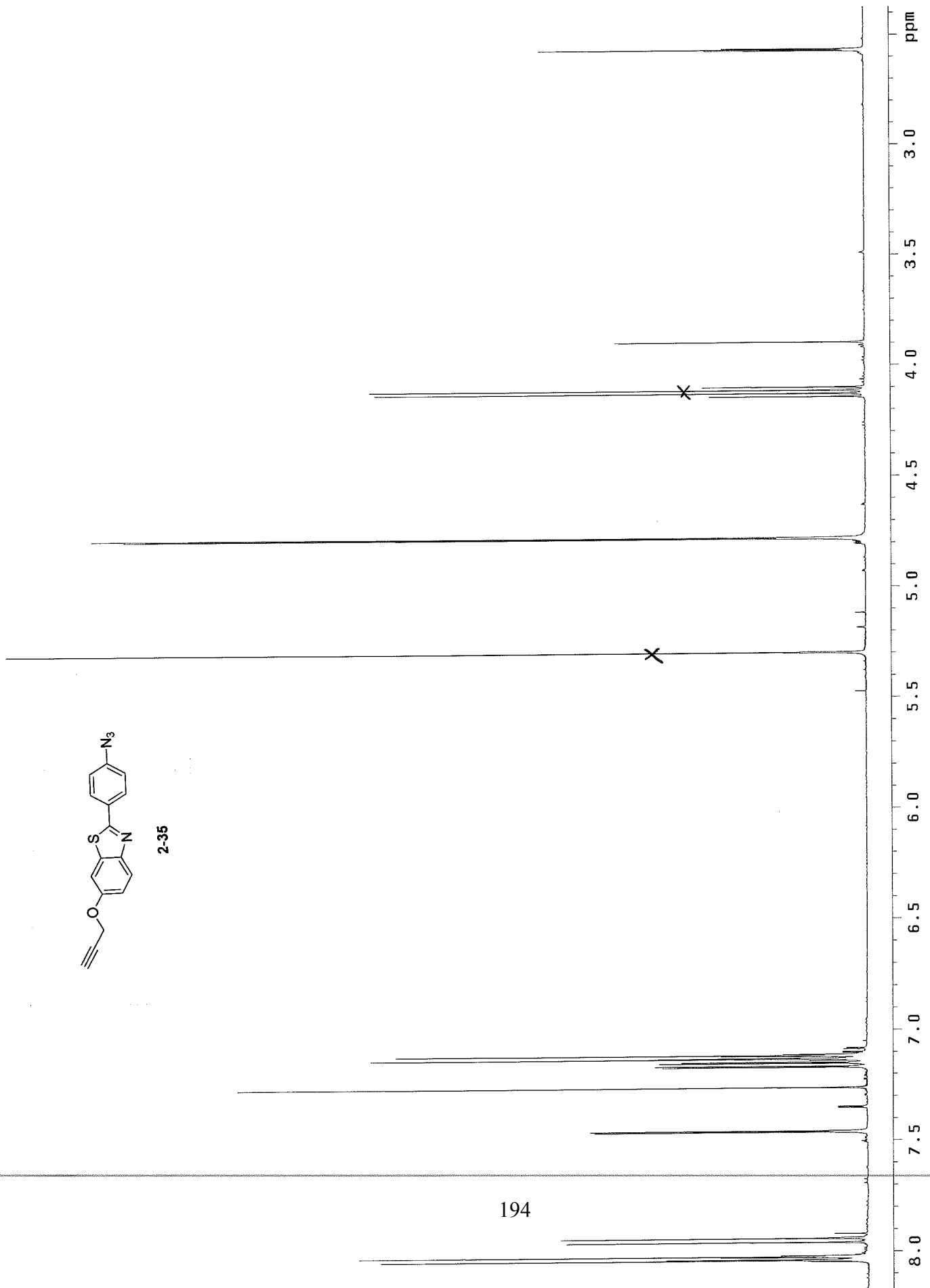
193

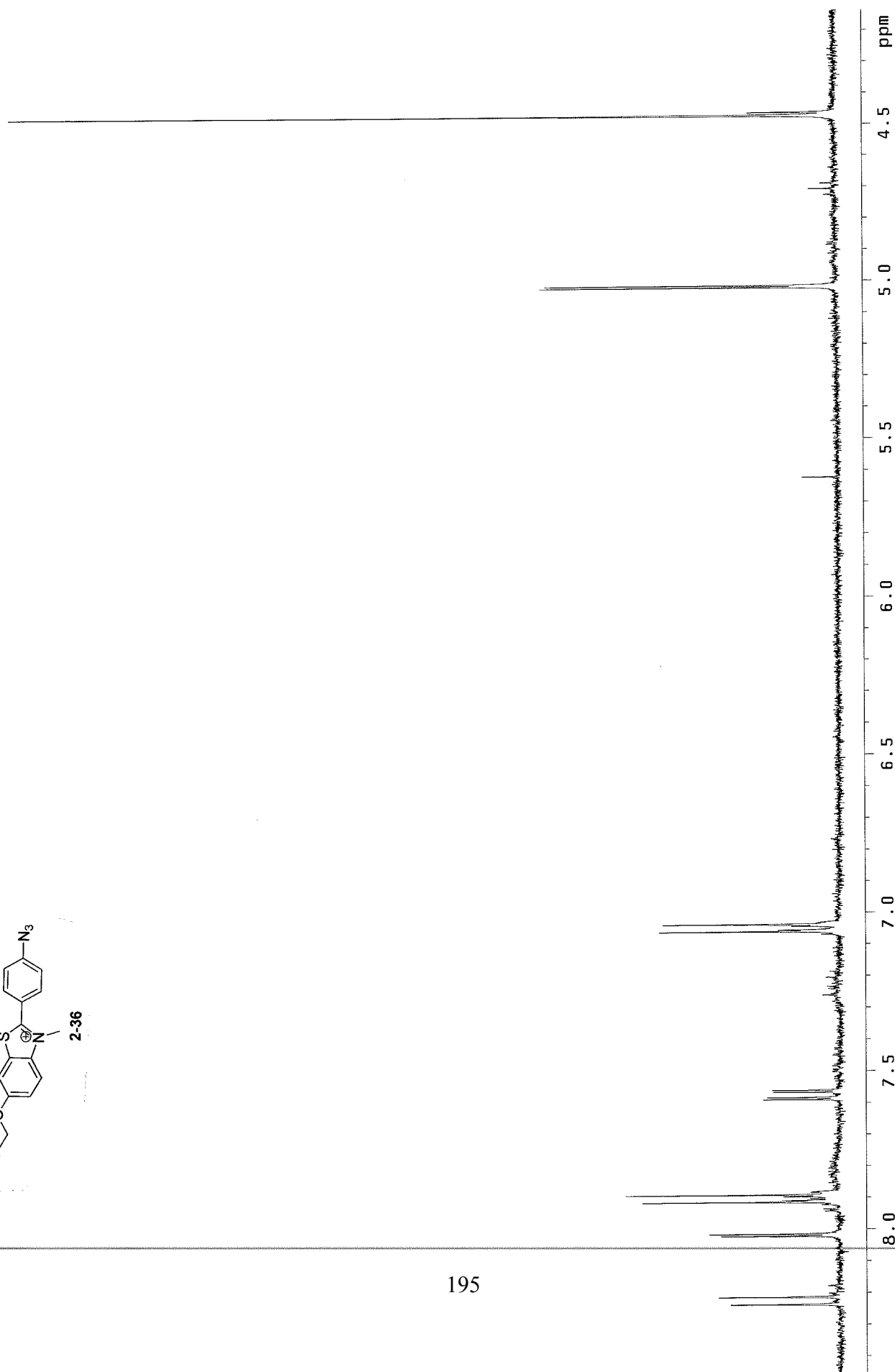
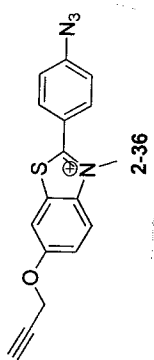




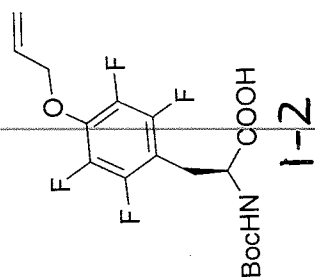
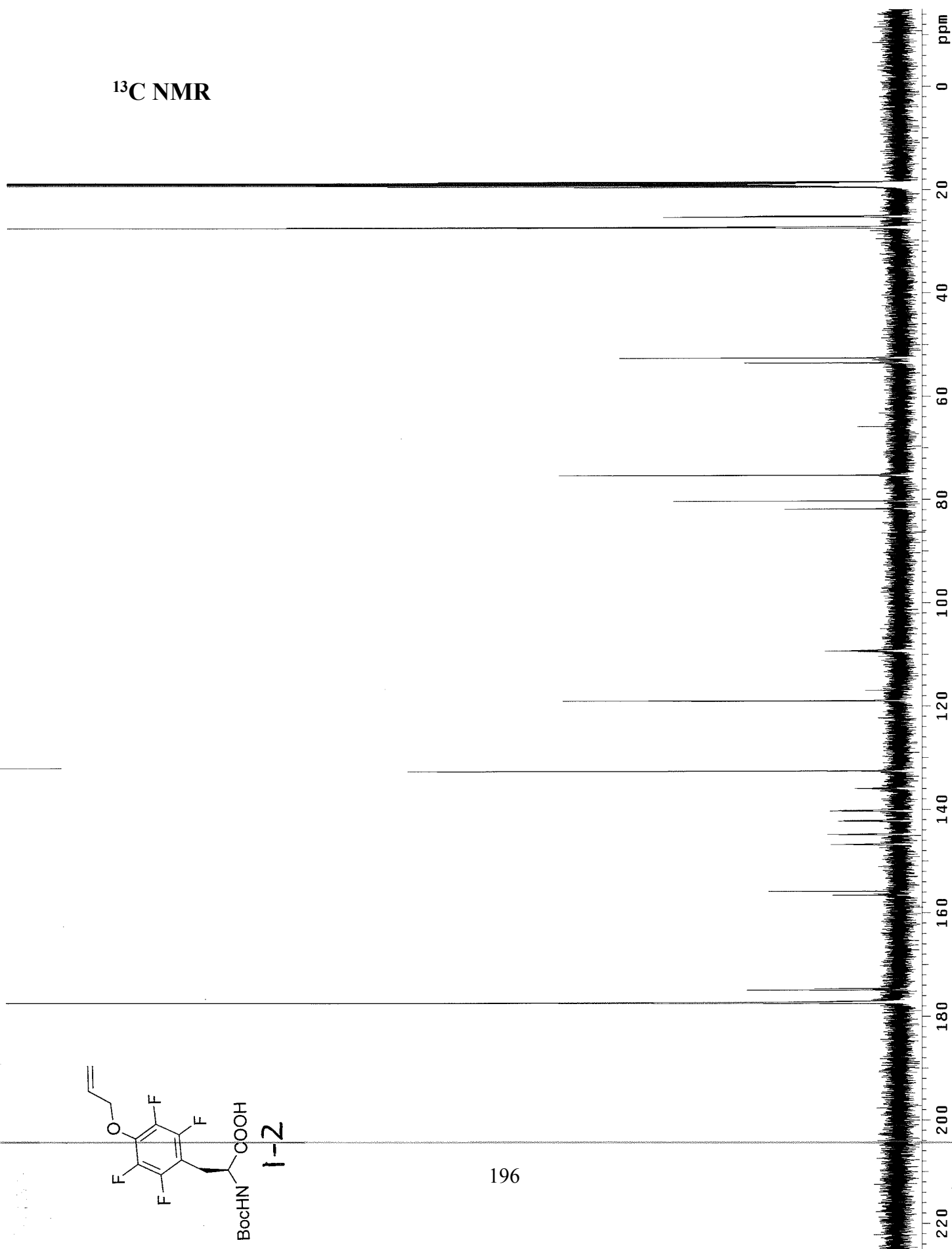
2-35

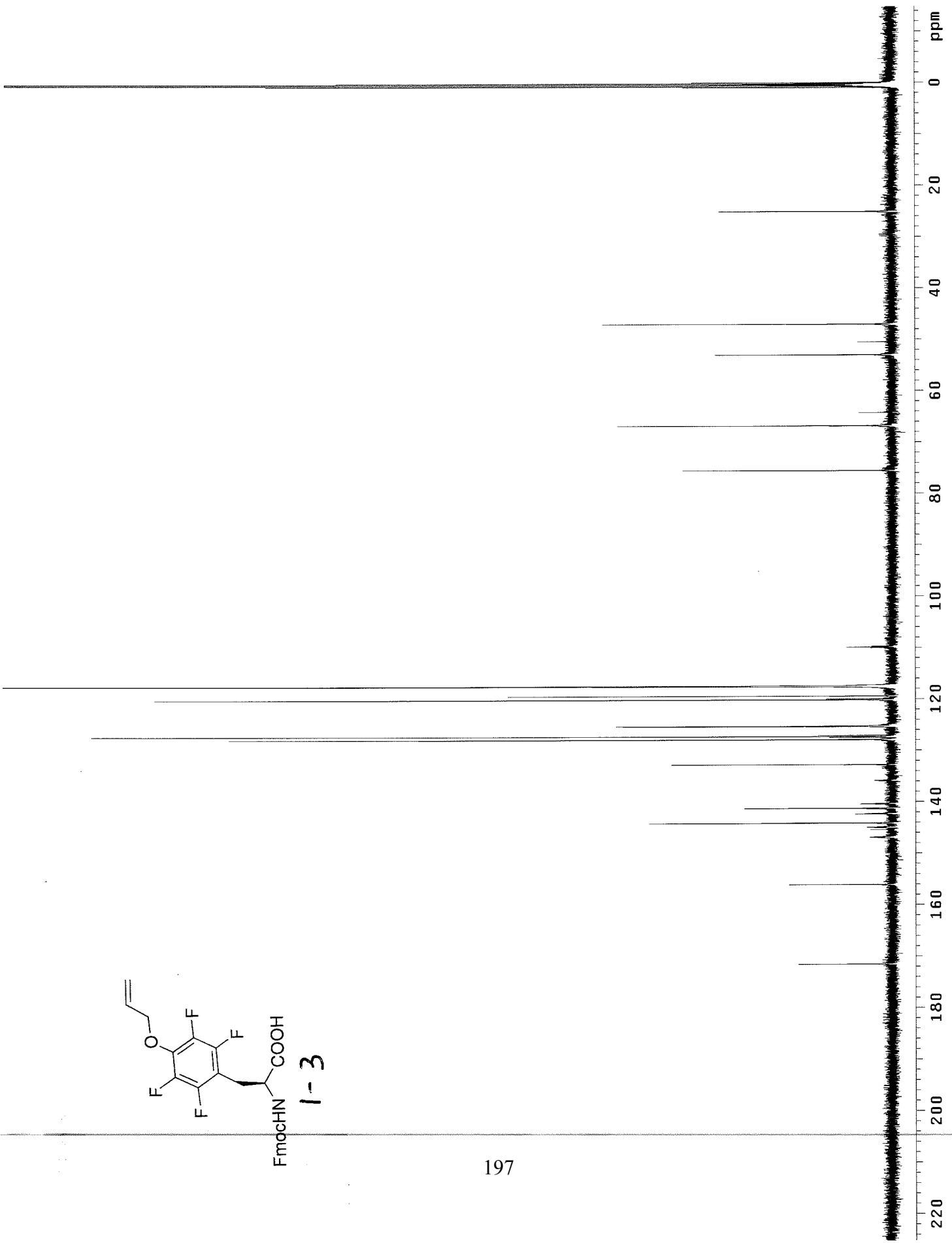
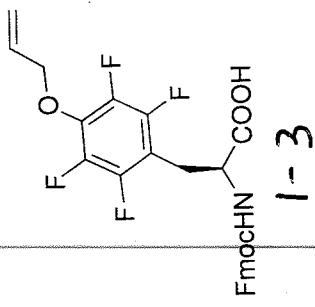
194



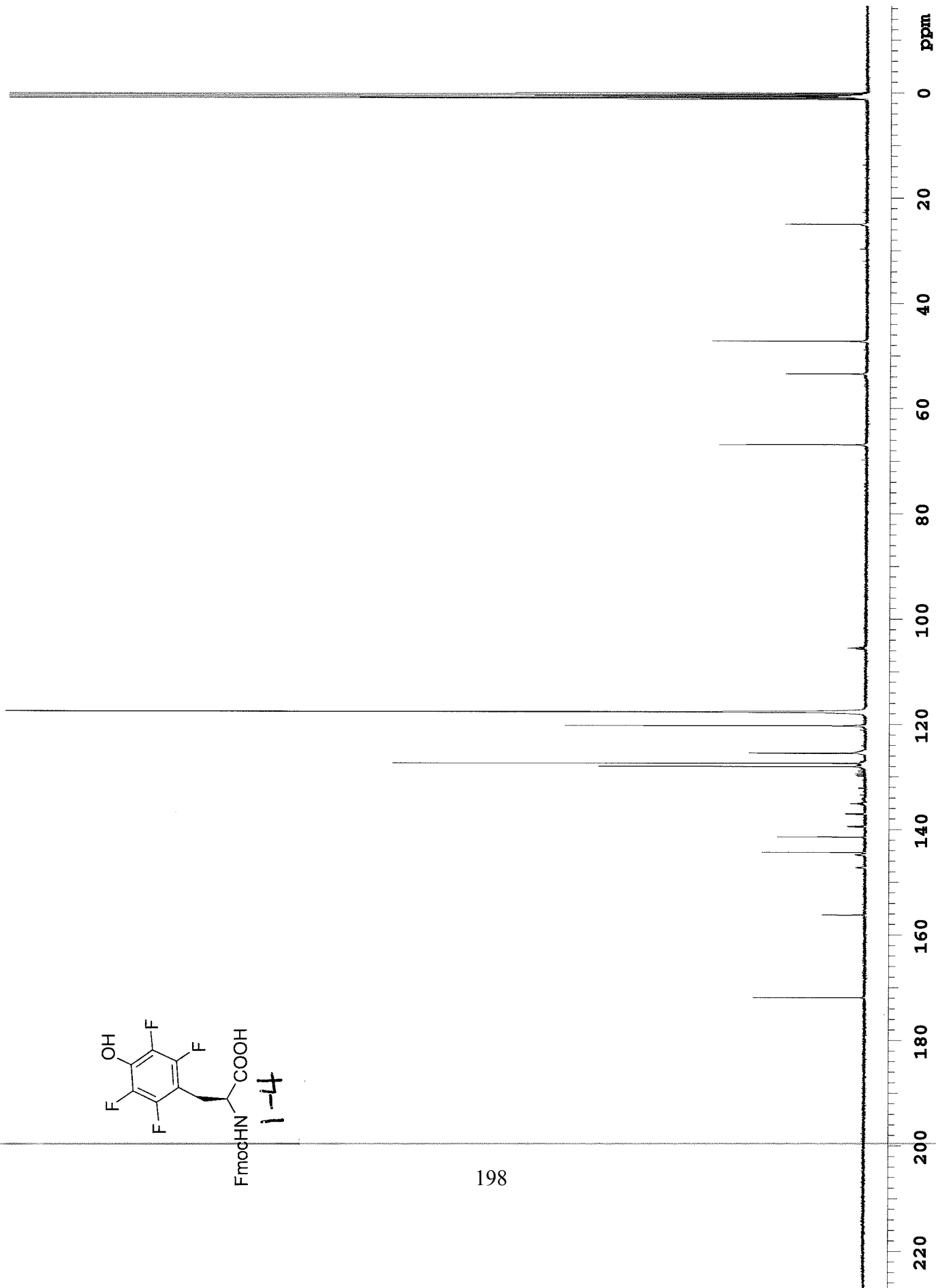
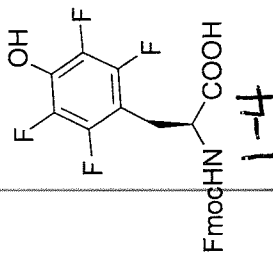


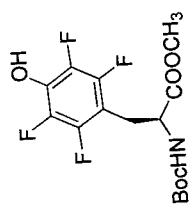
<sup>13</sup>C NMR



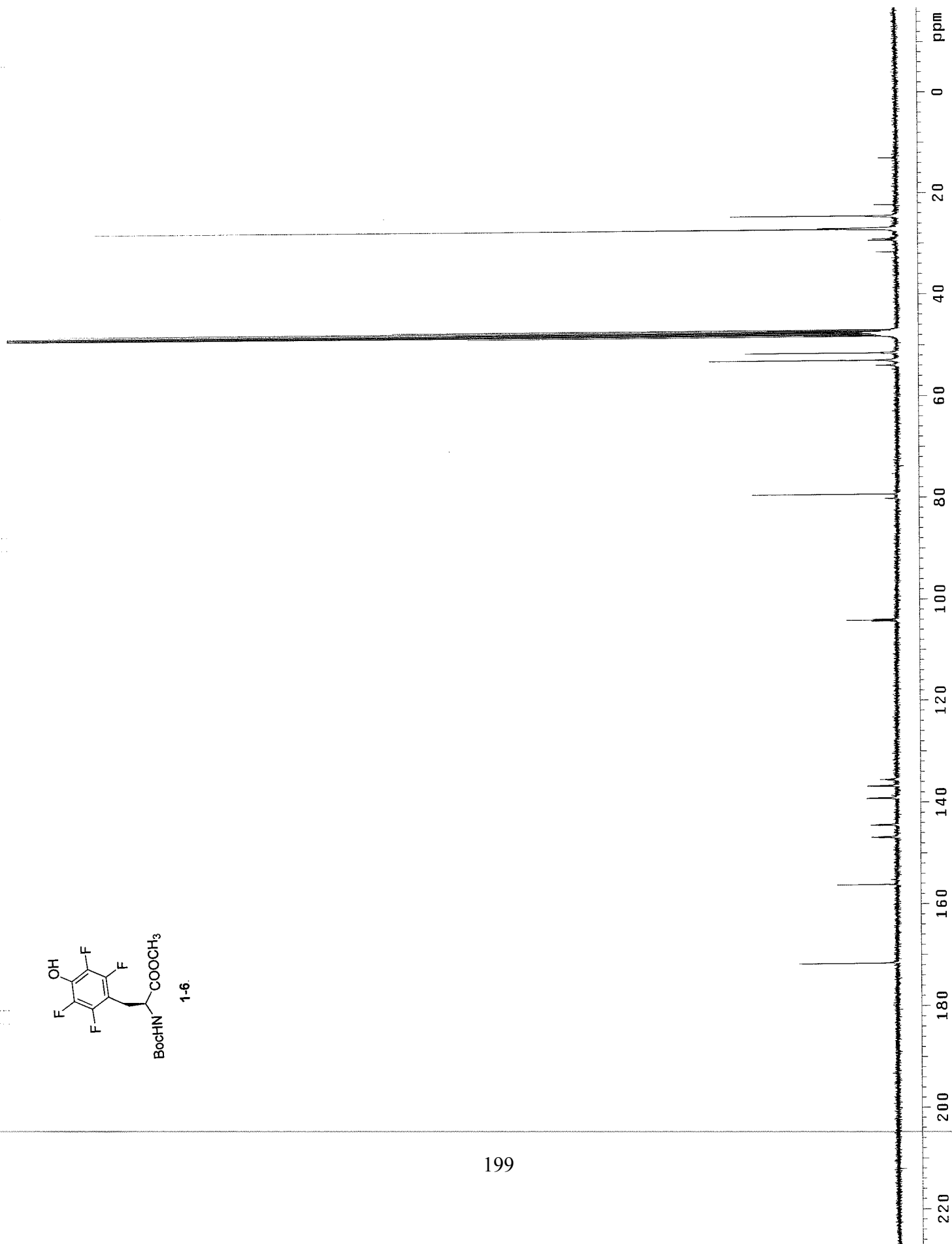


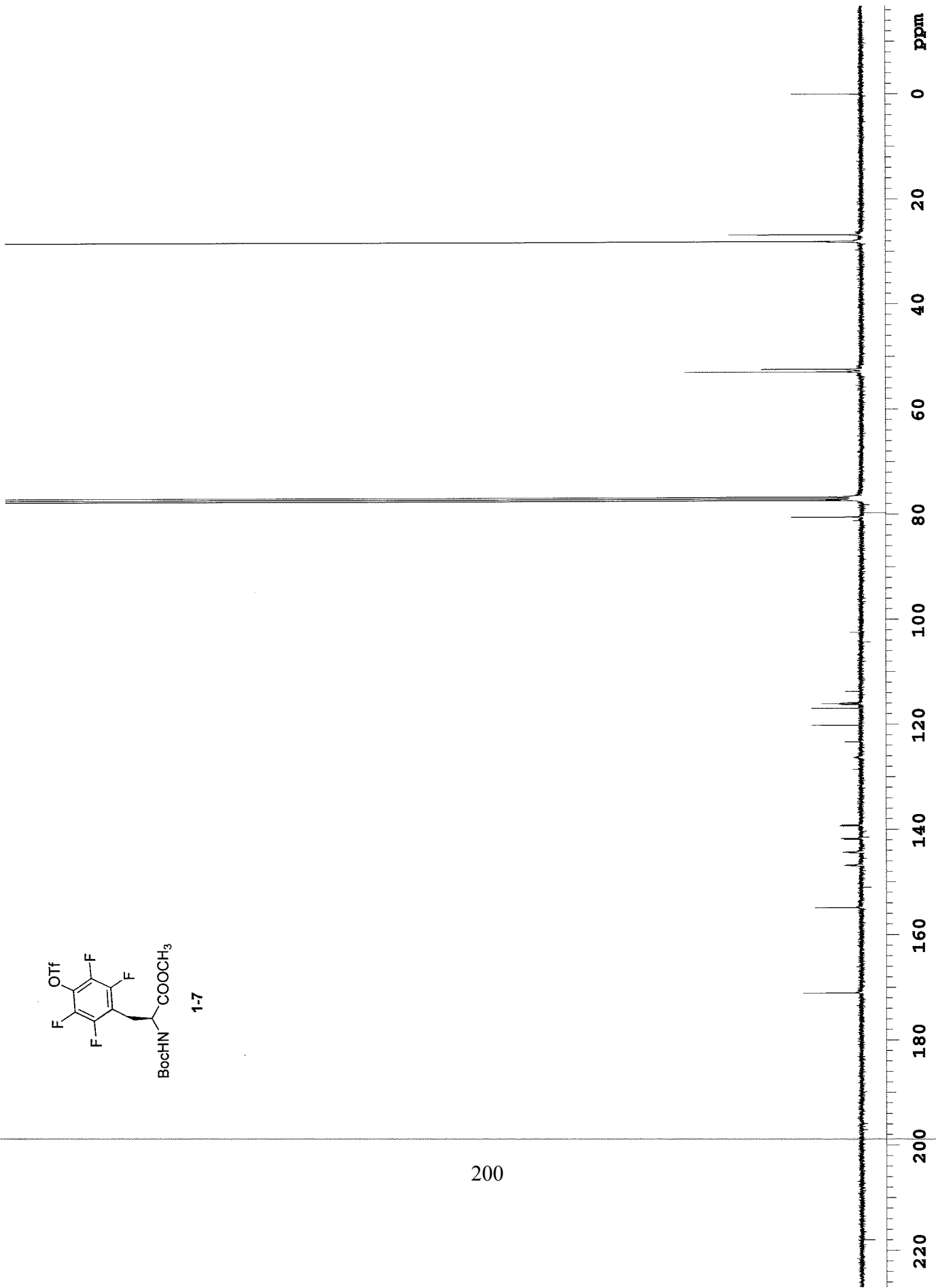
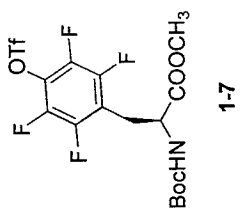


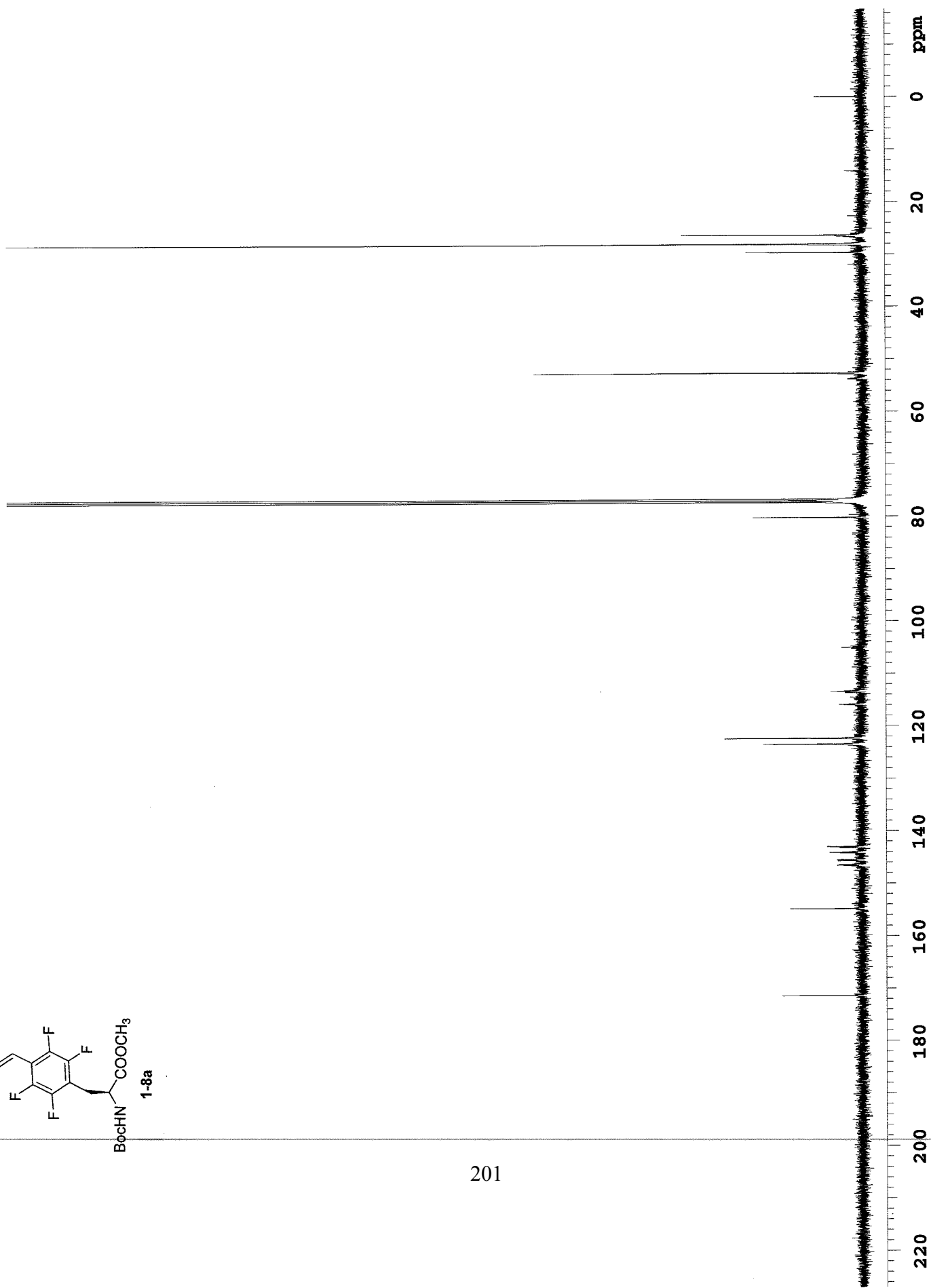
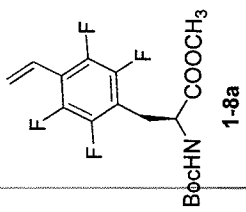


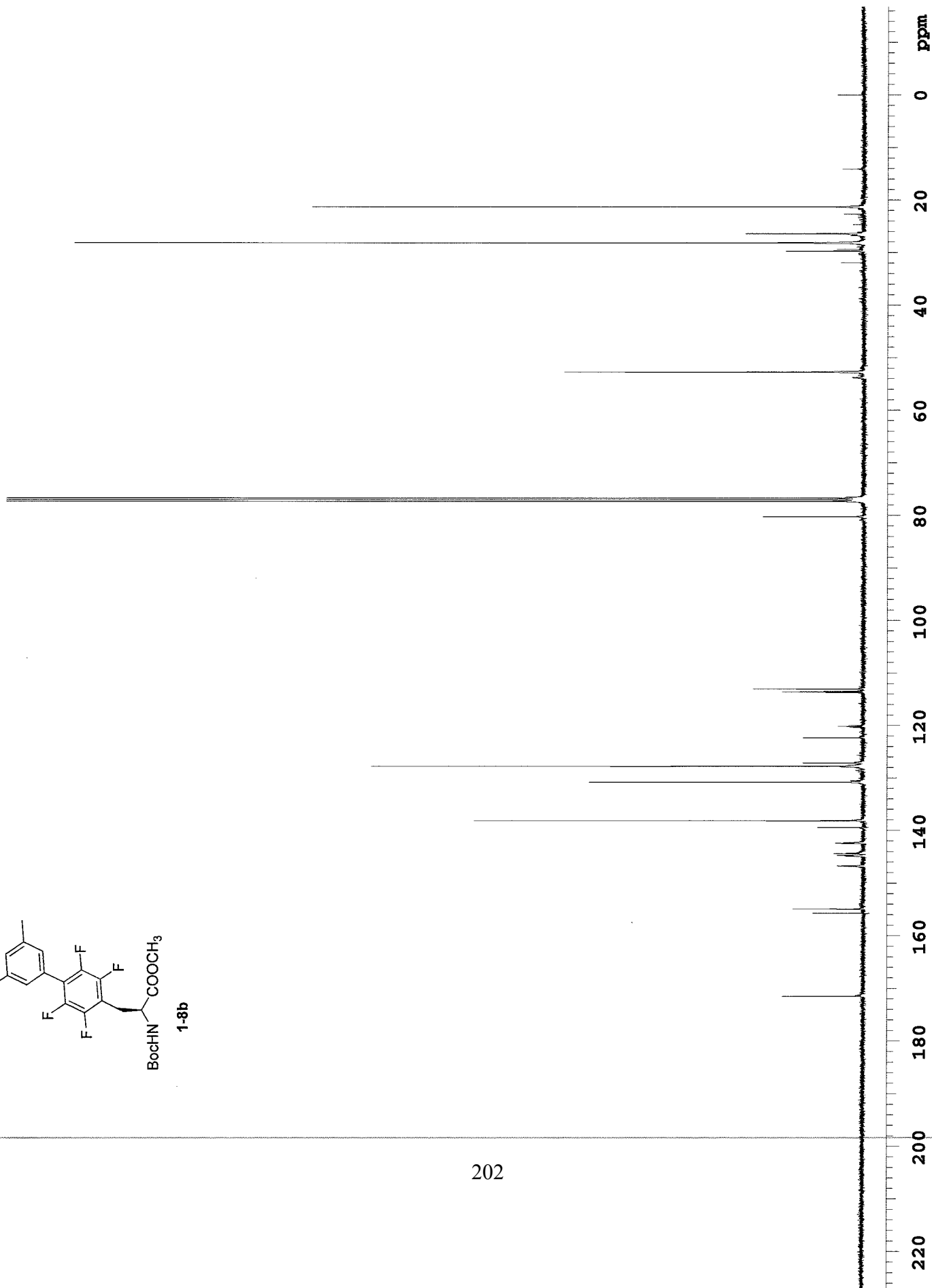
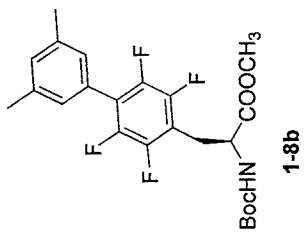


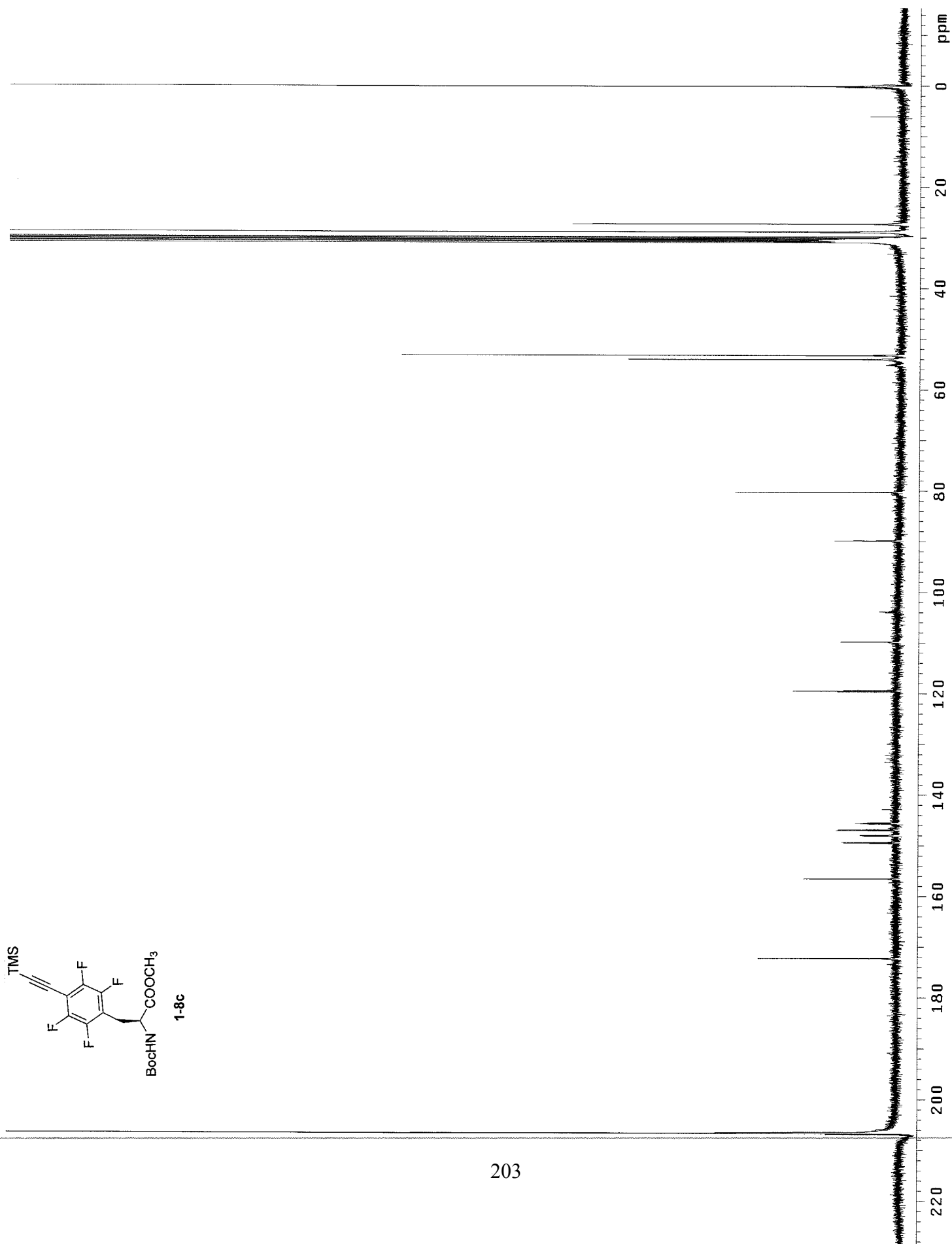
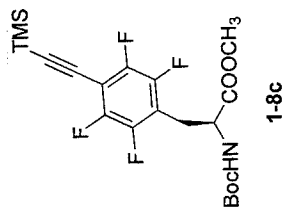
1-6.

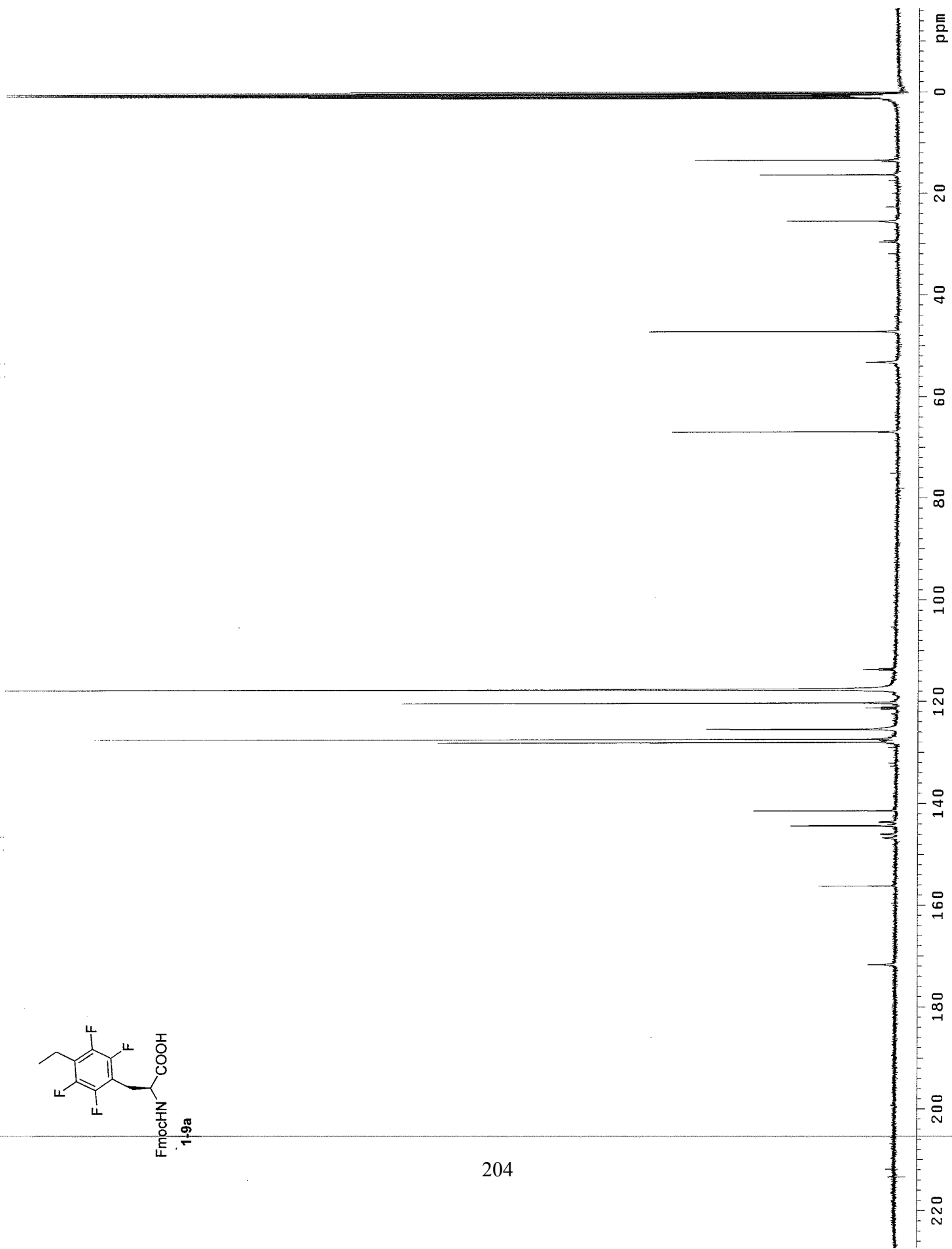
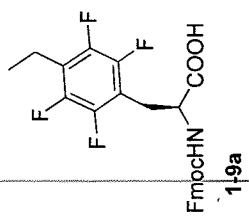


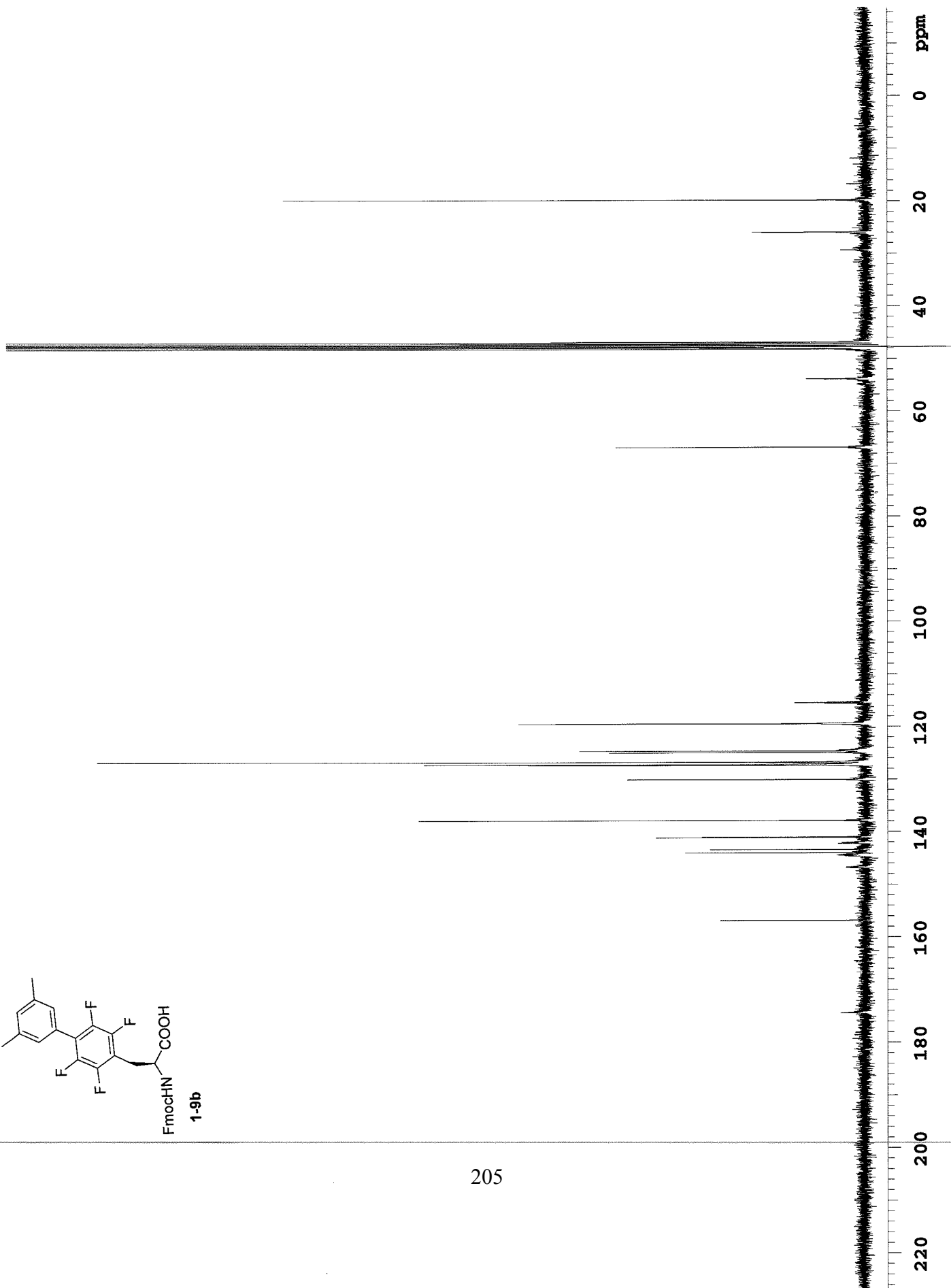
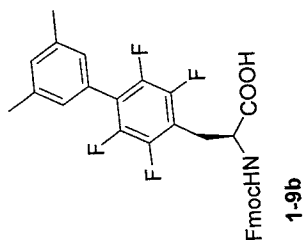




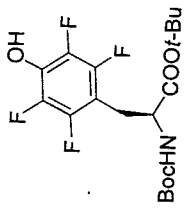




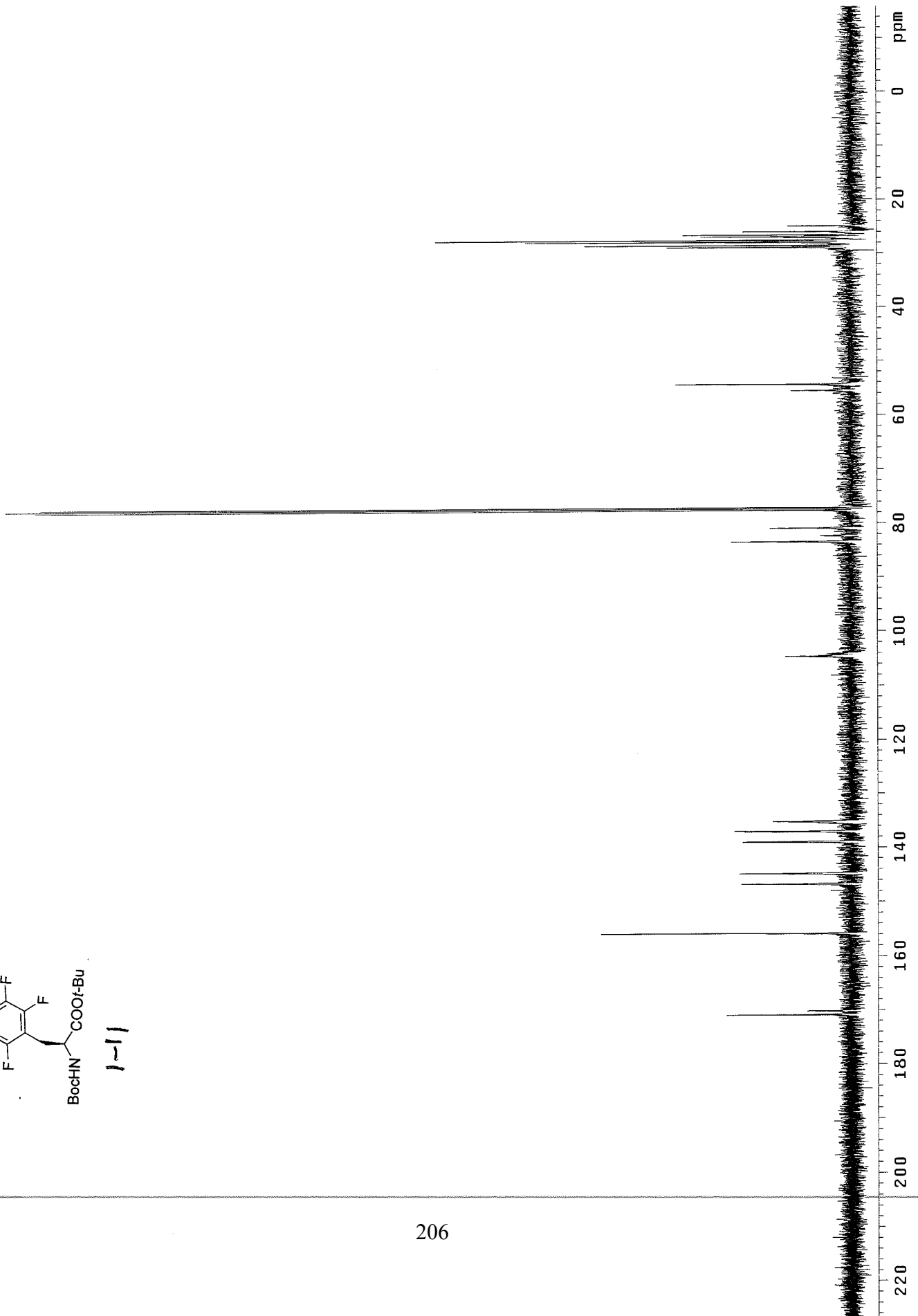


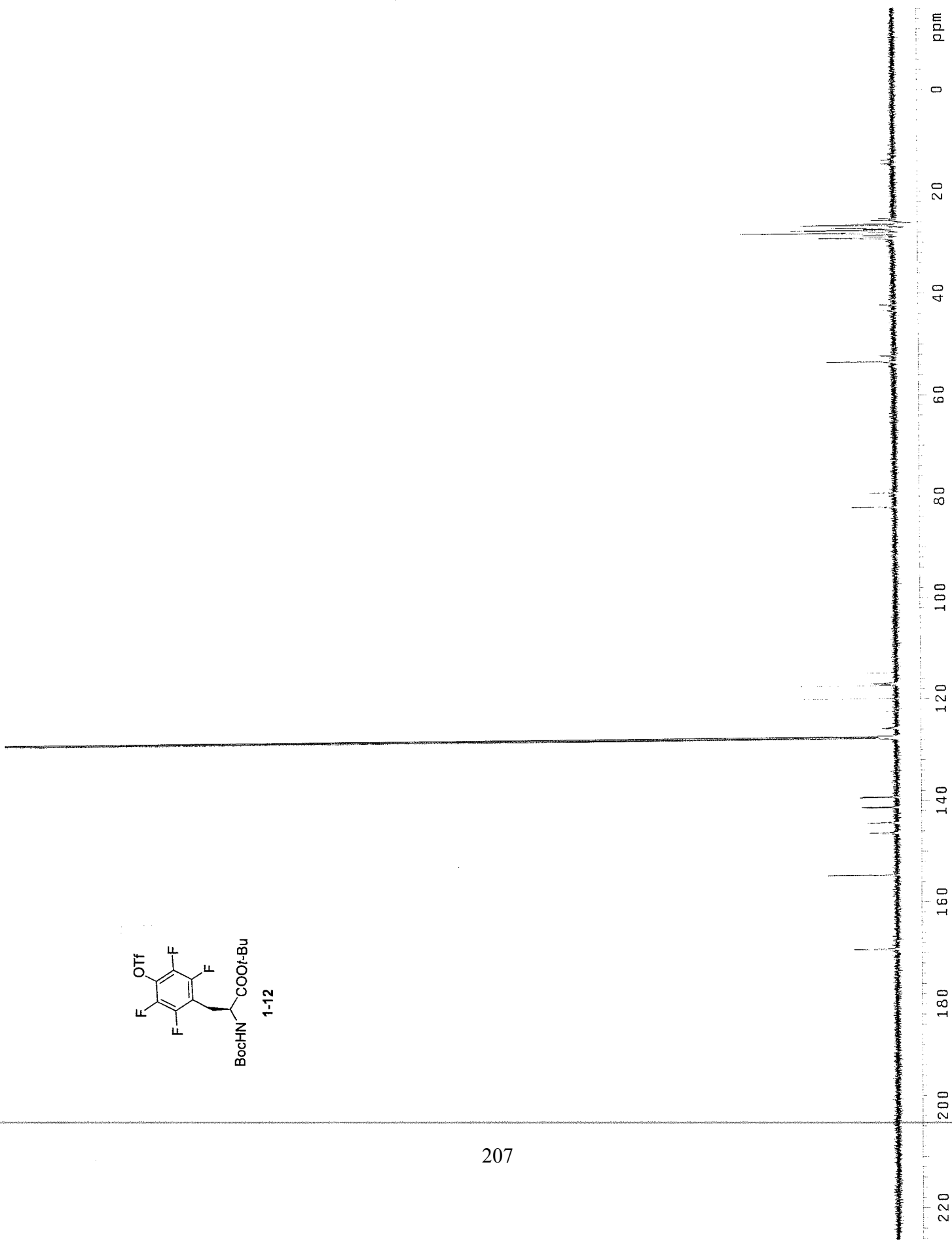
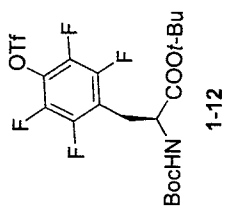


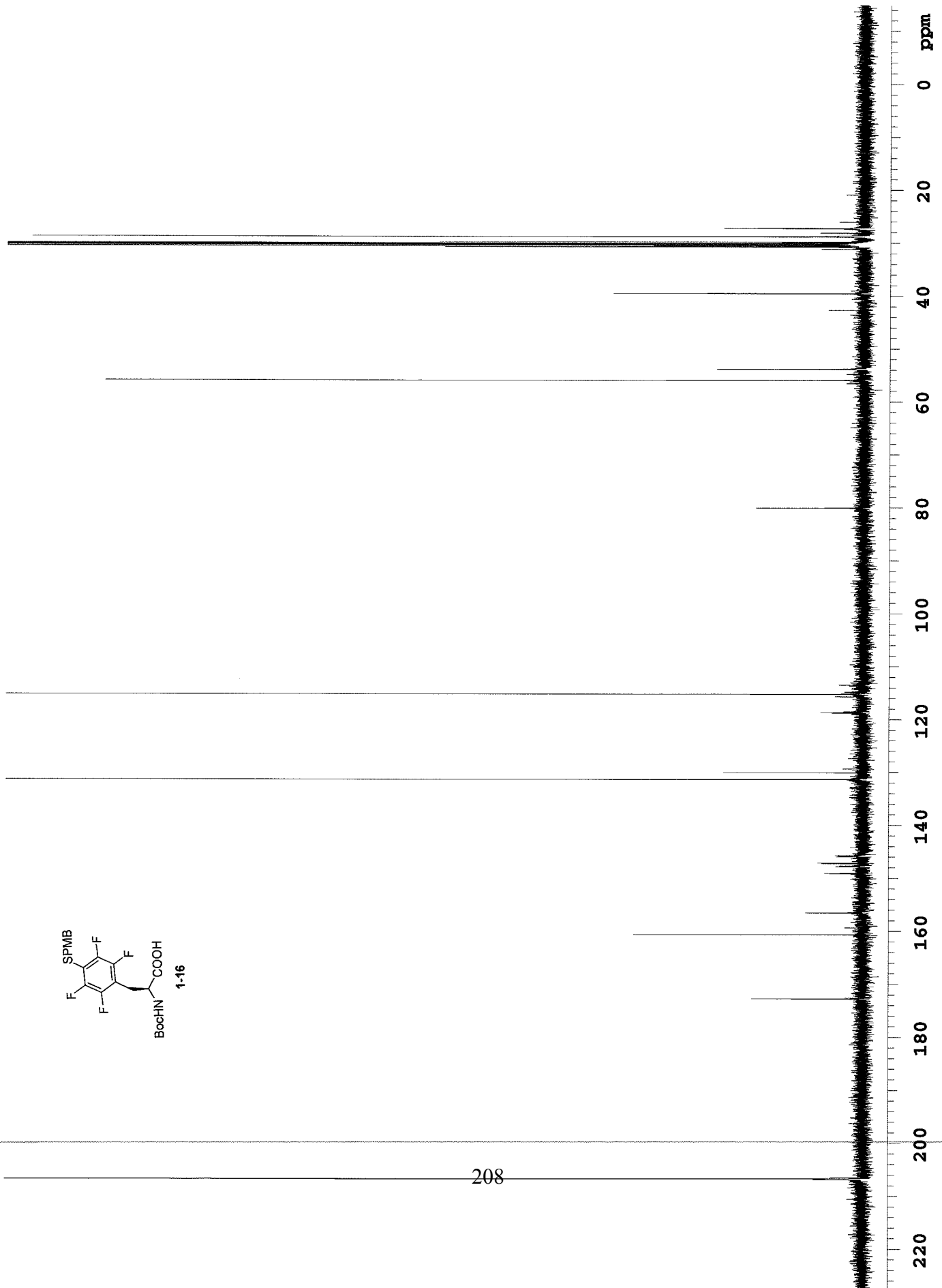
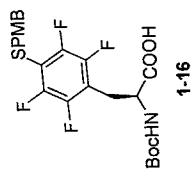


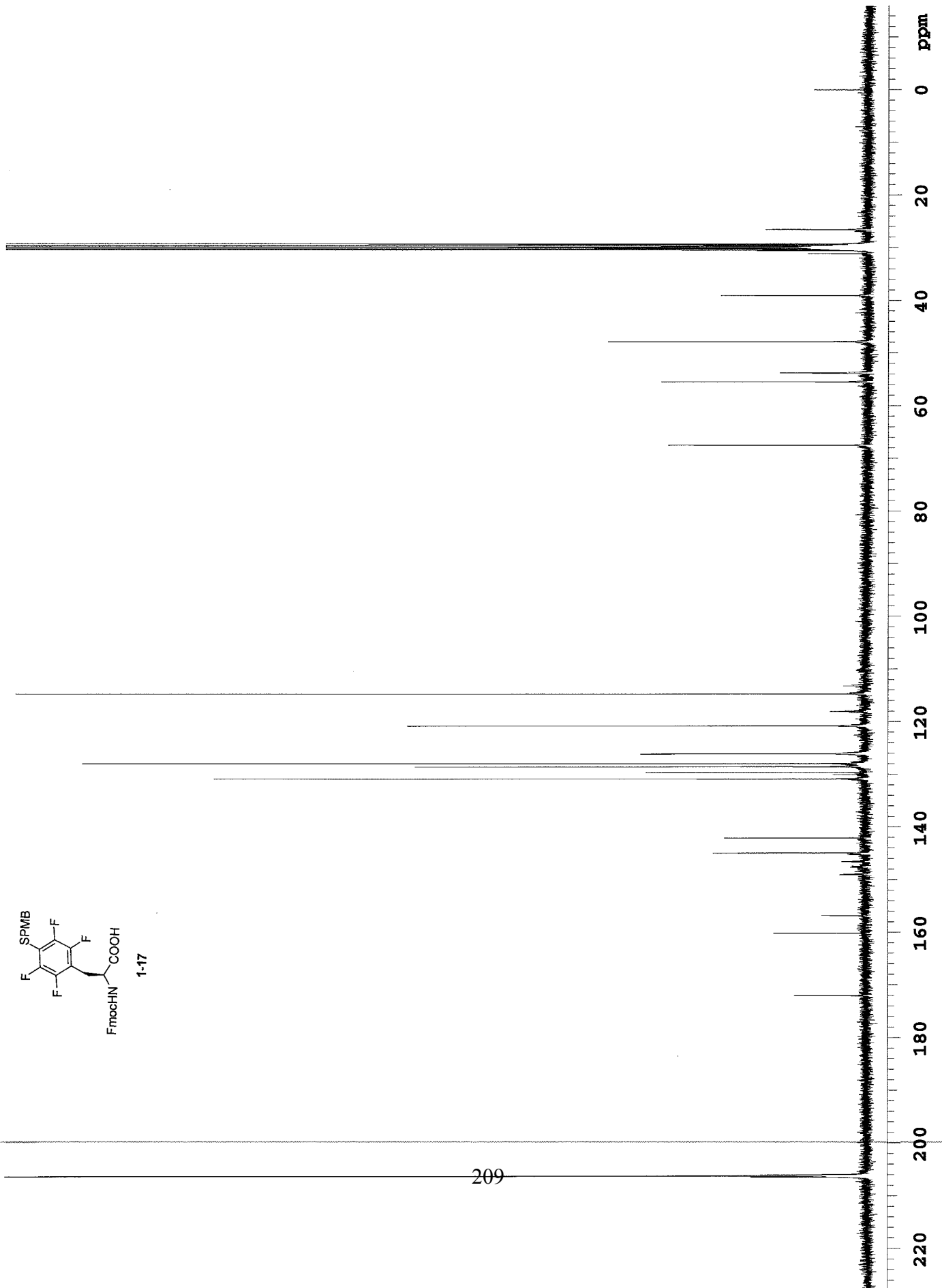
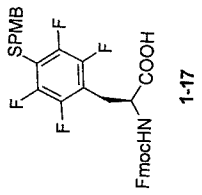


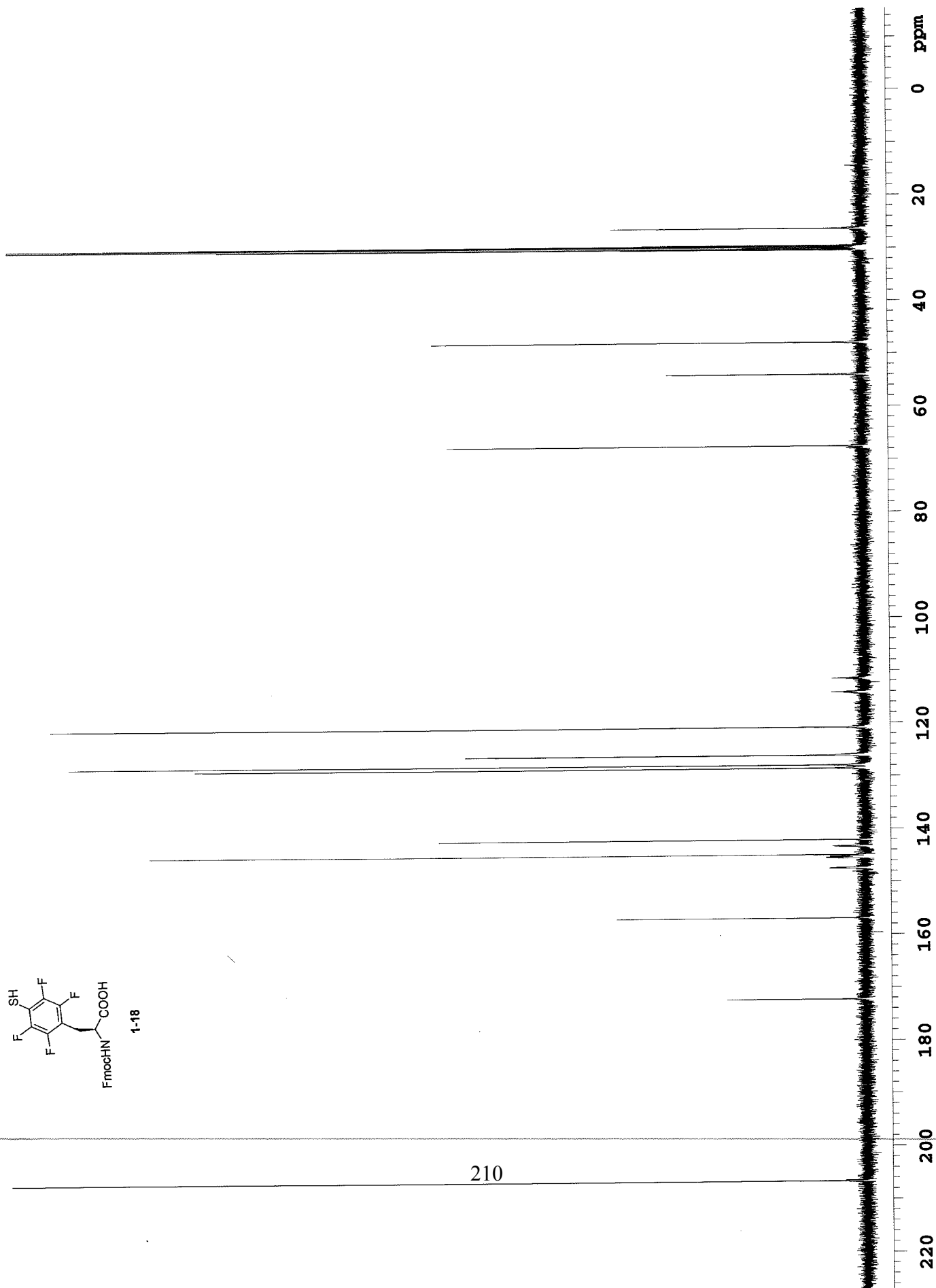
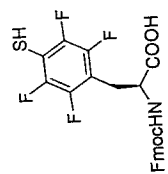
1-11

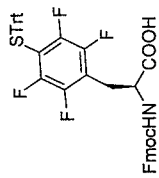






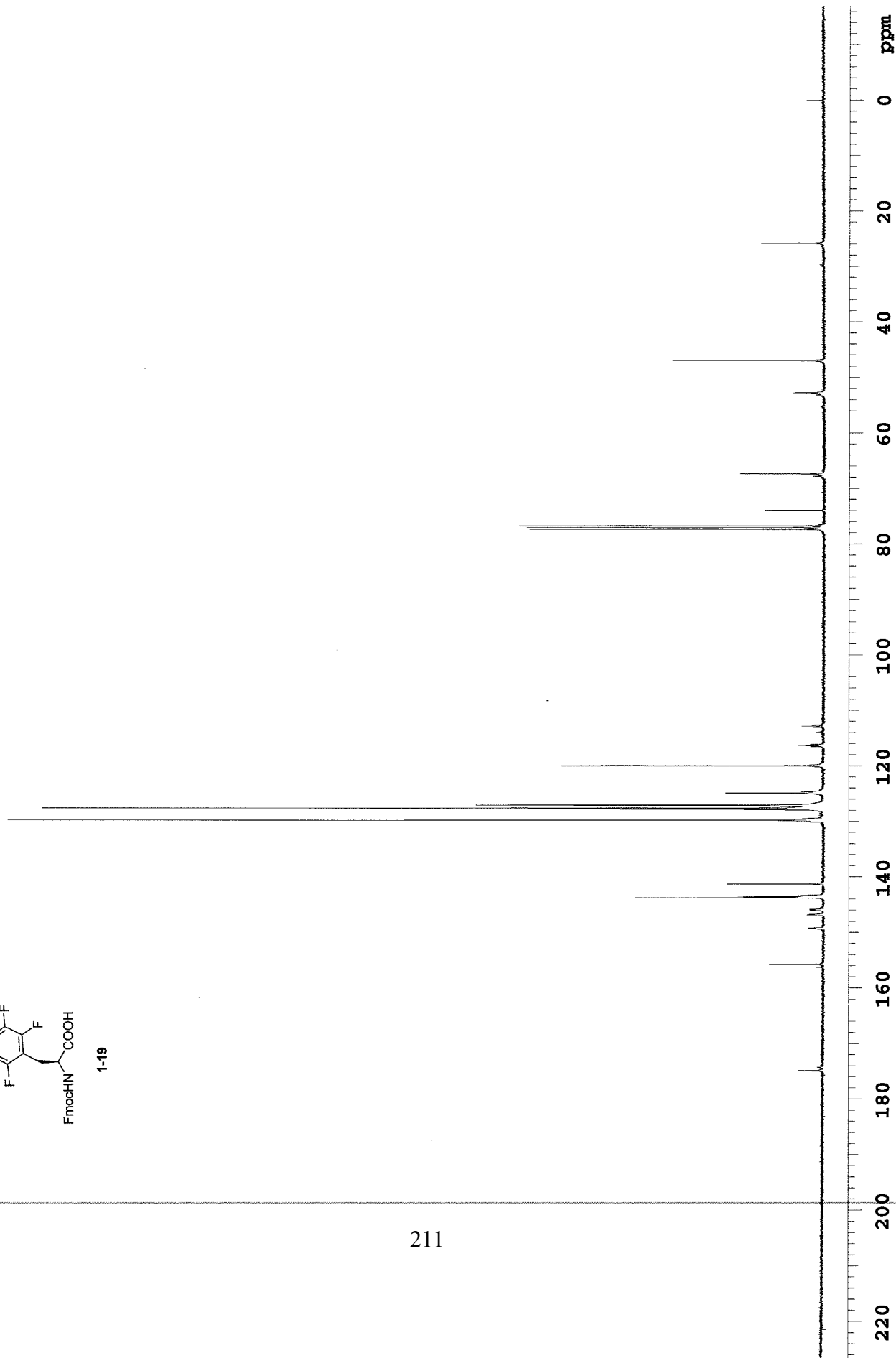


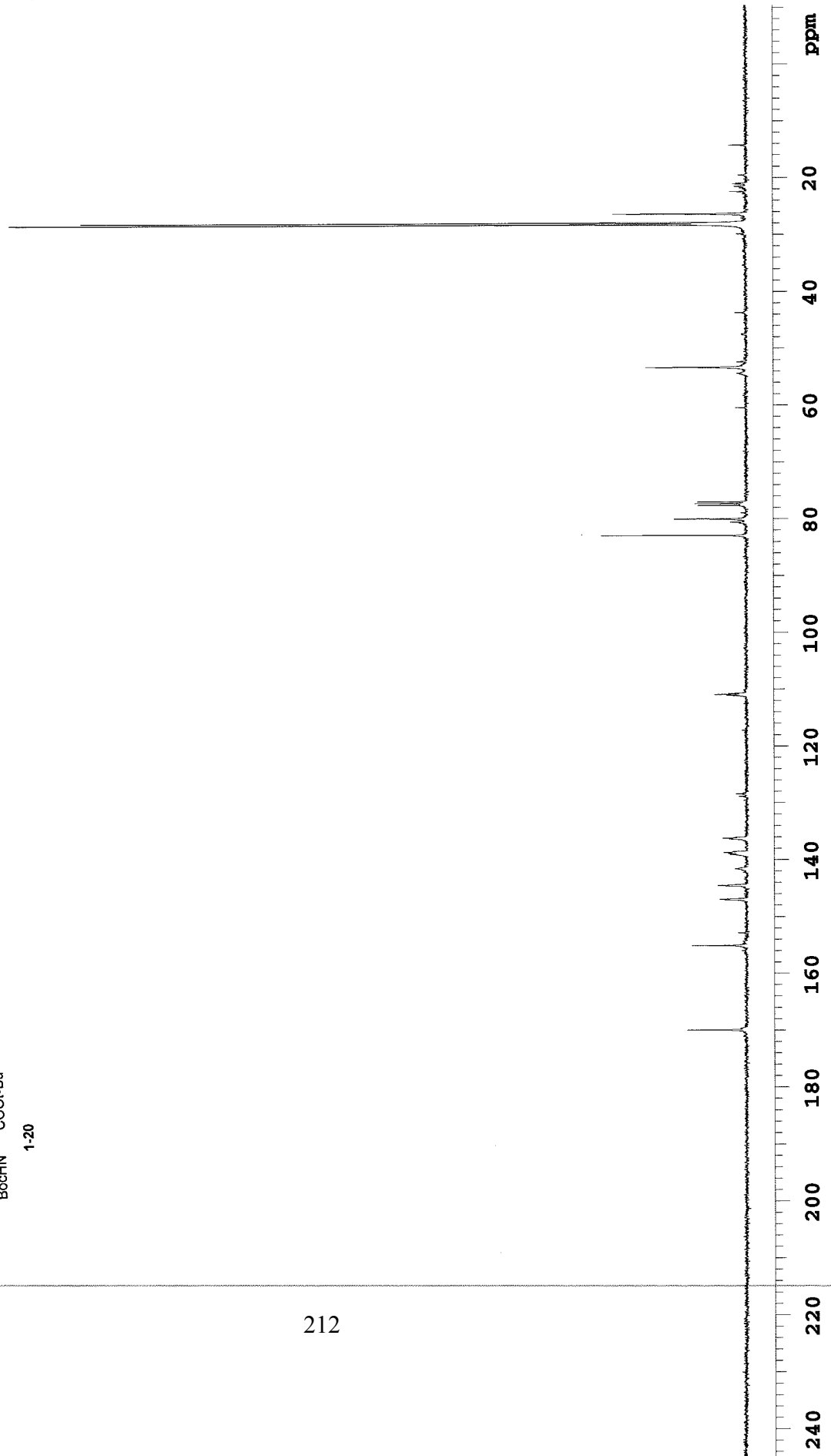
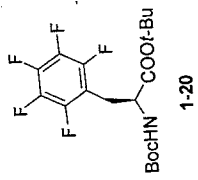


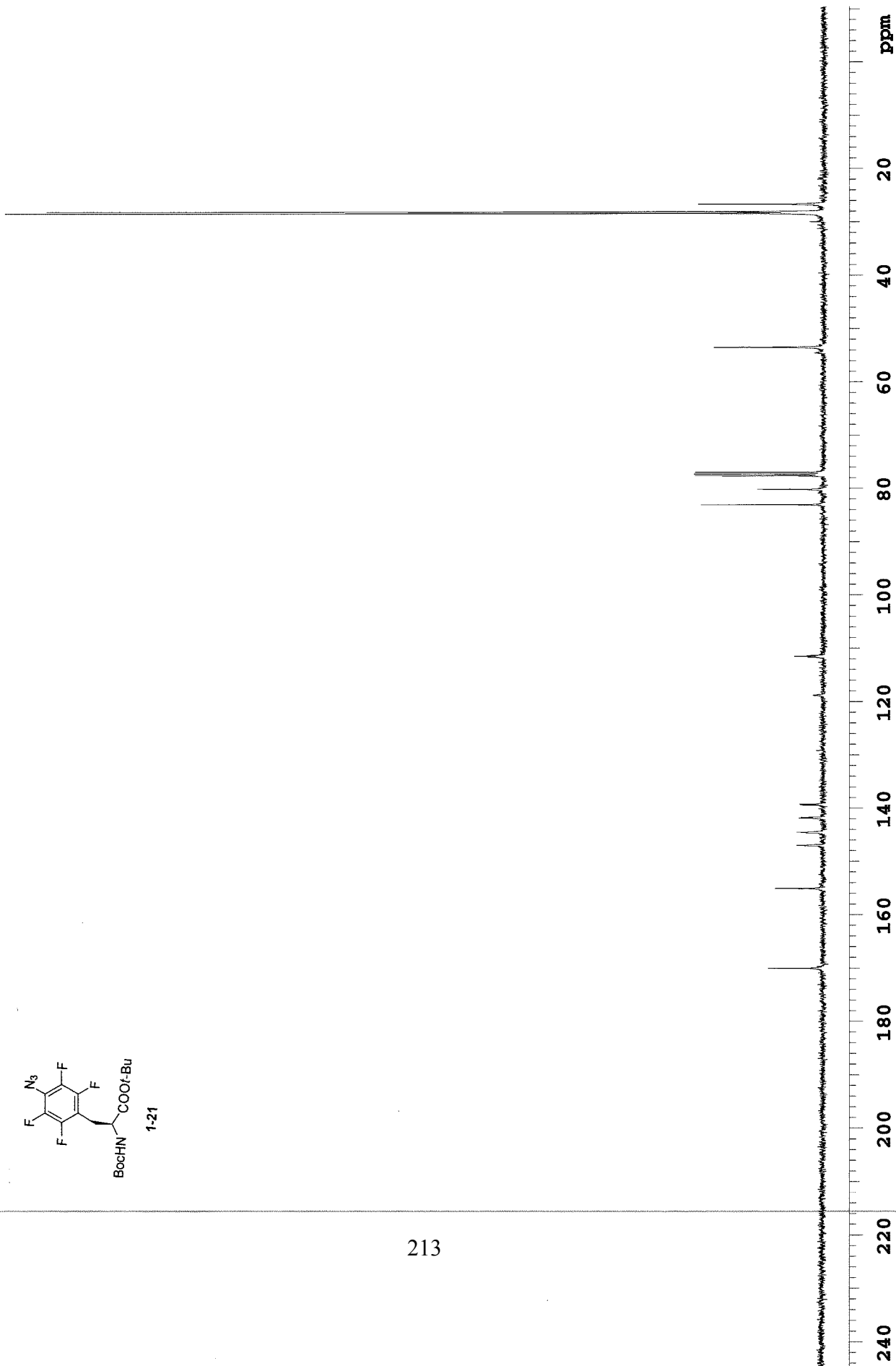
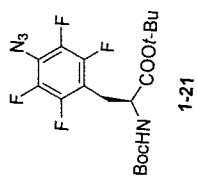


1-19

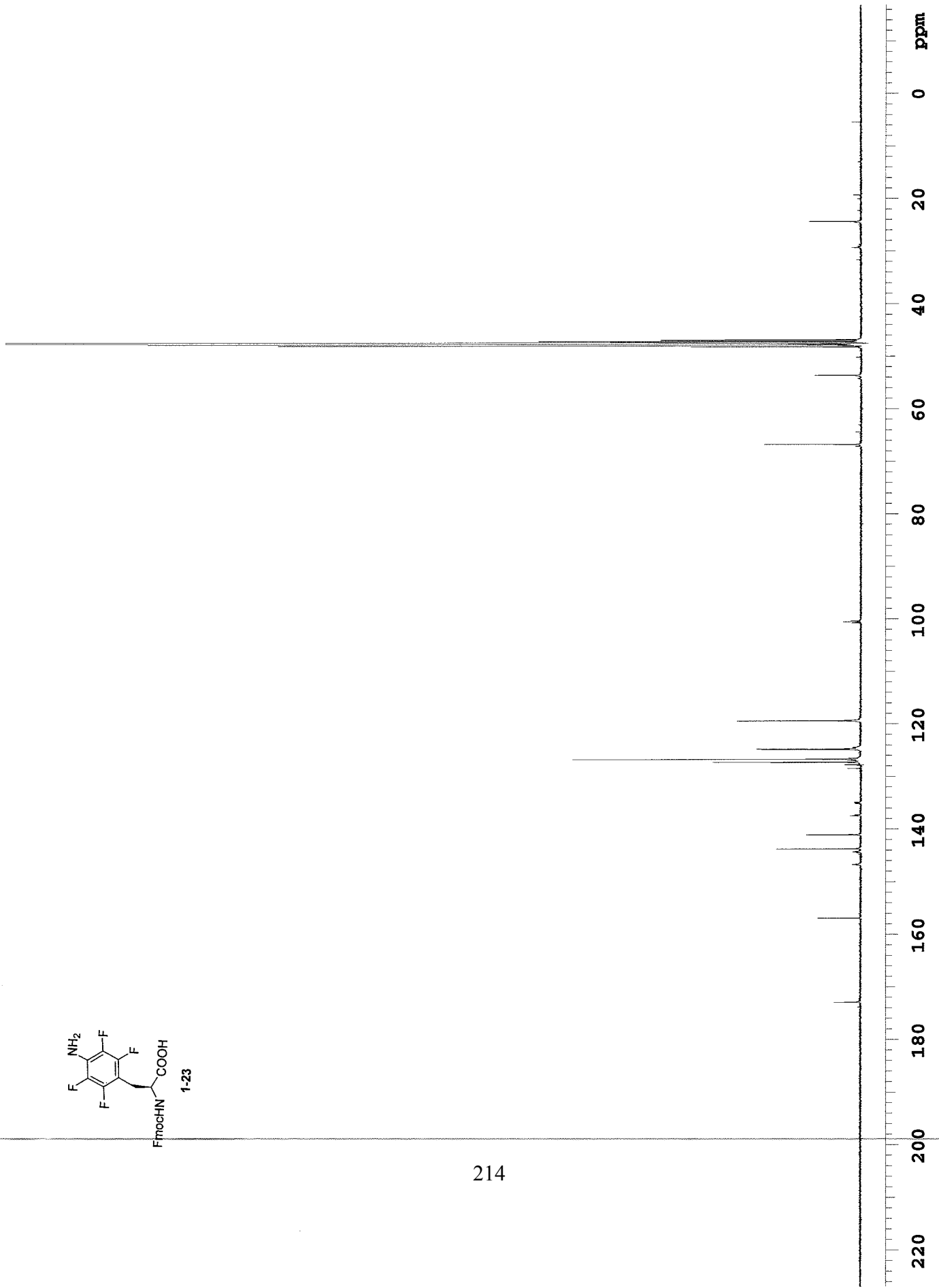
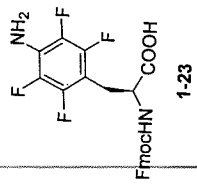
211

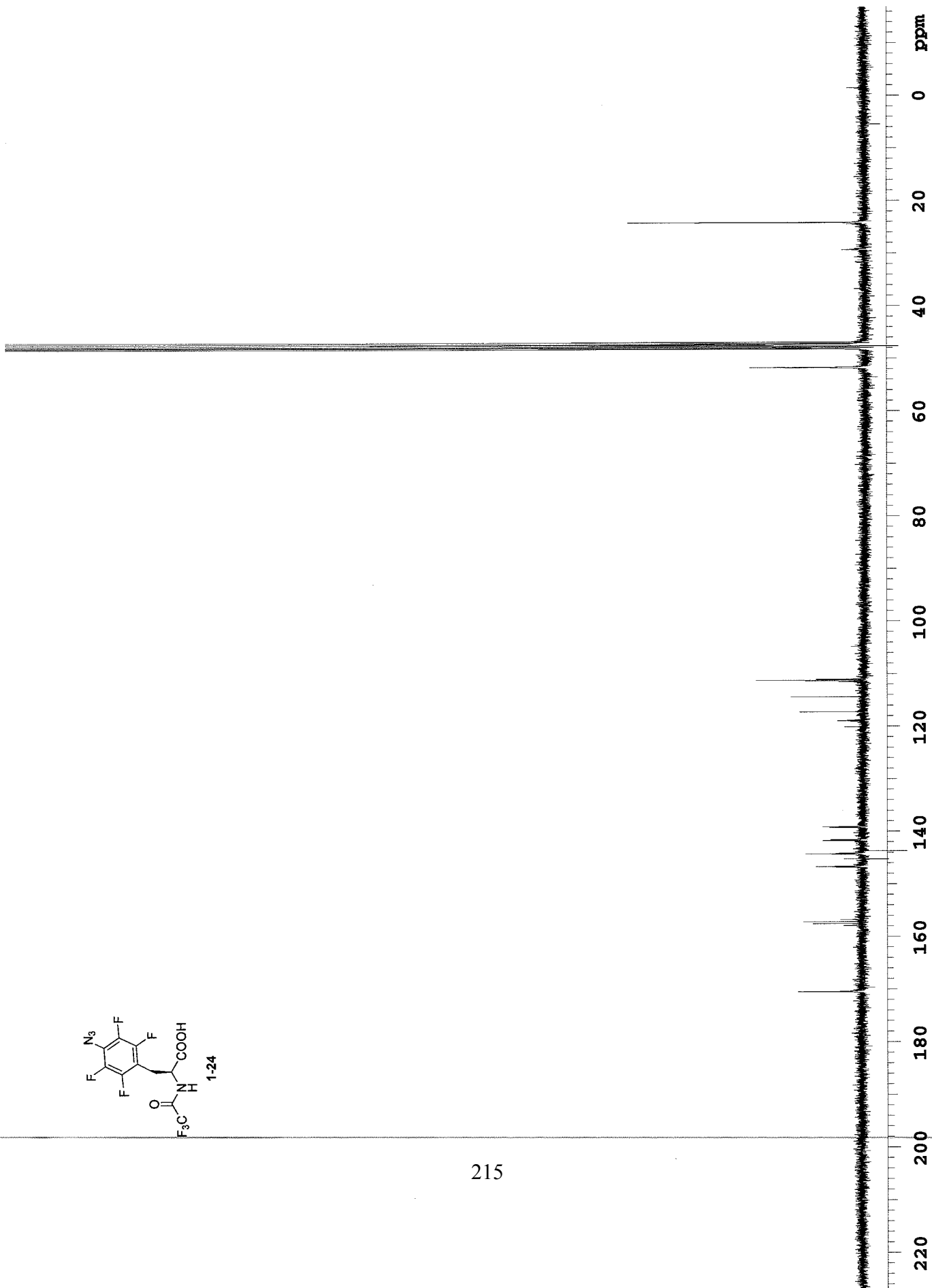
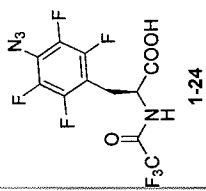


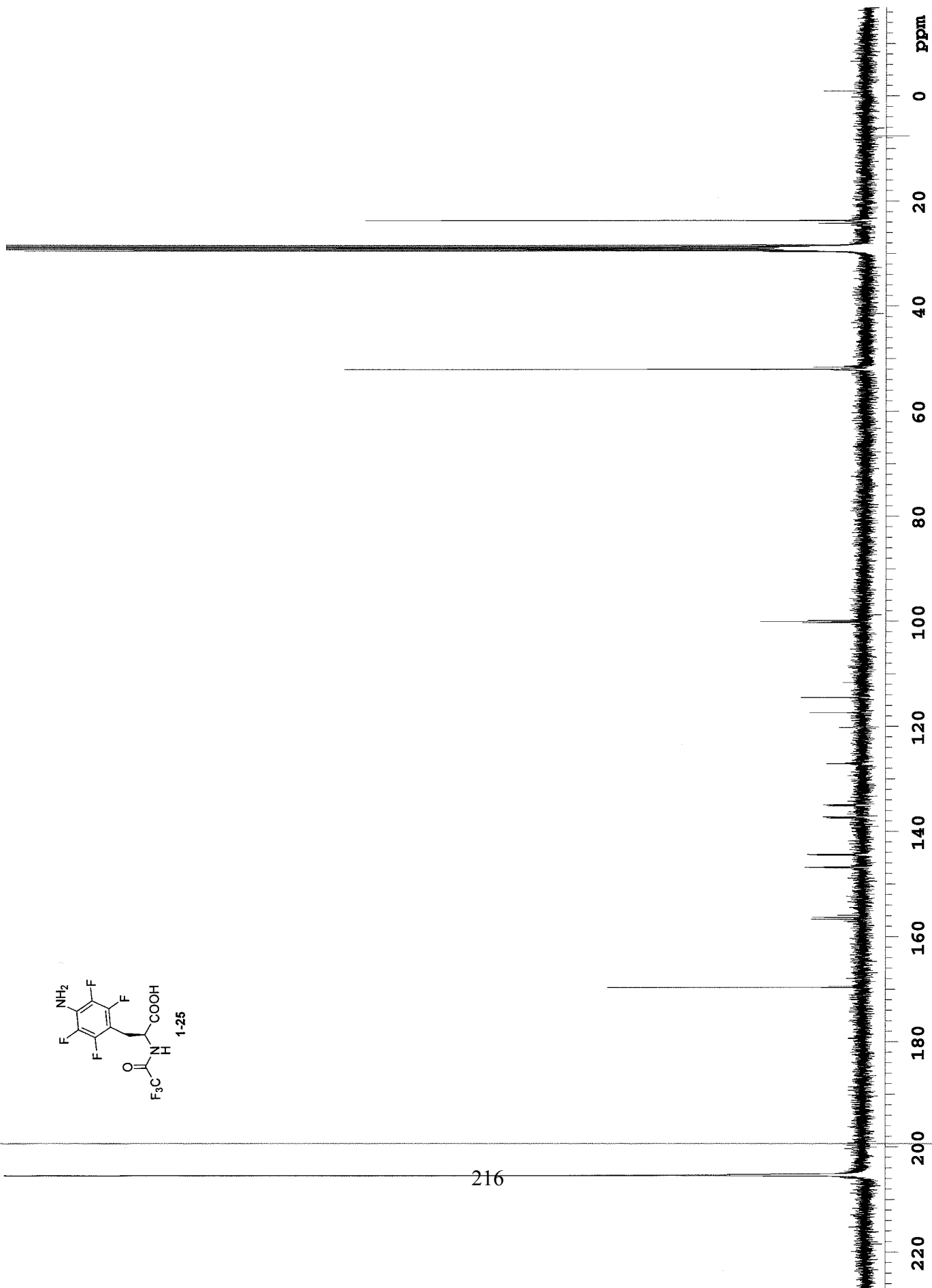
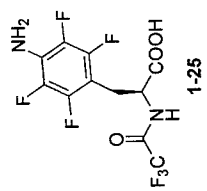


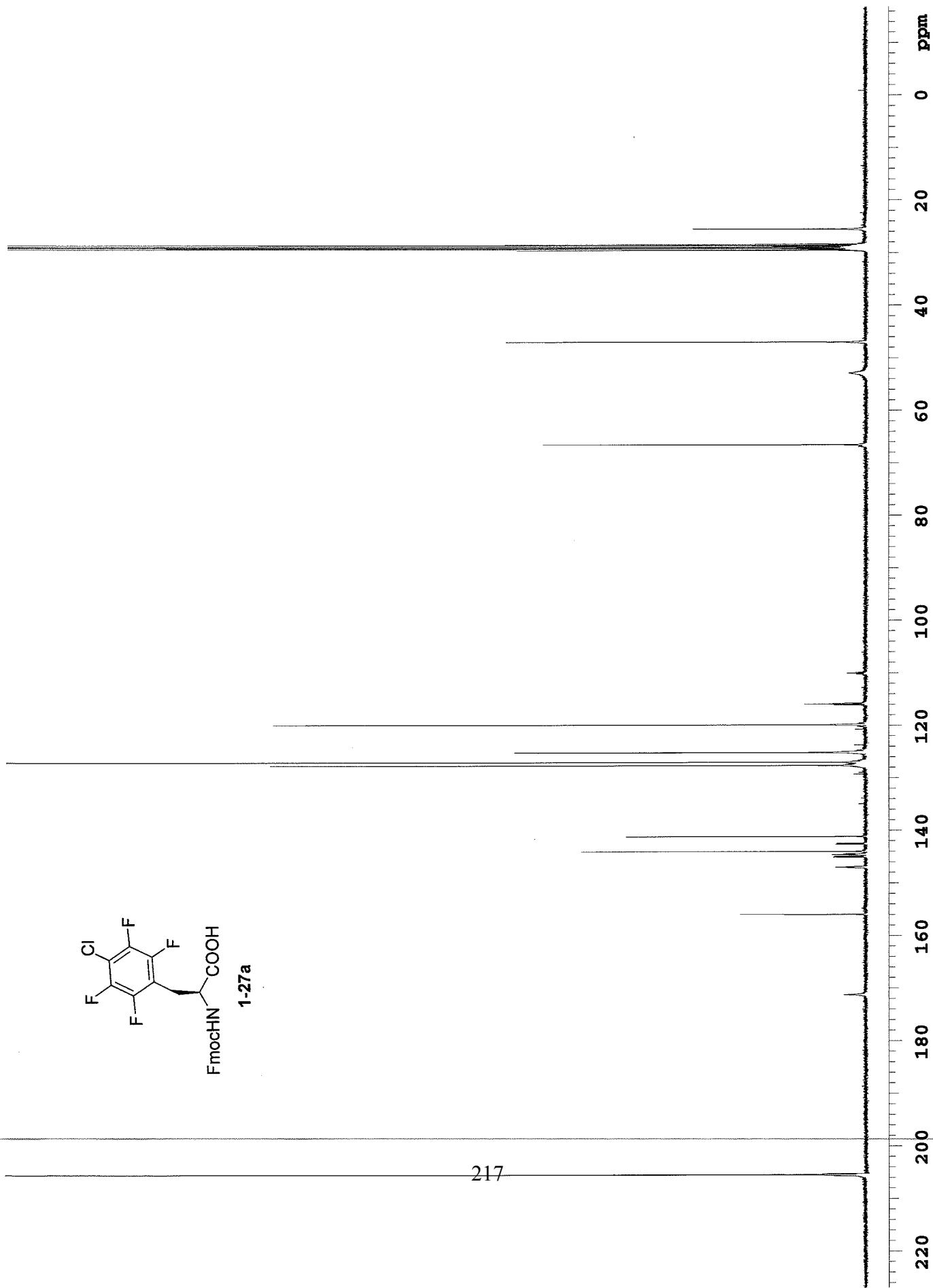
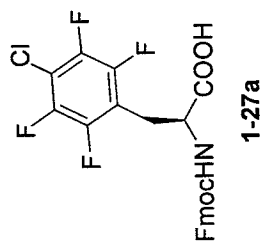


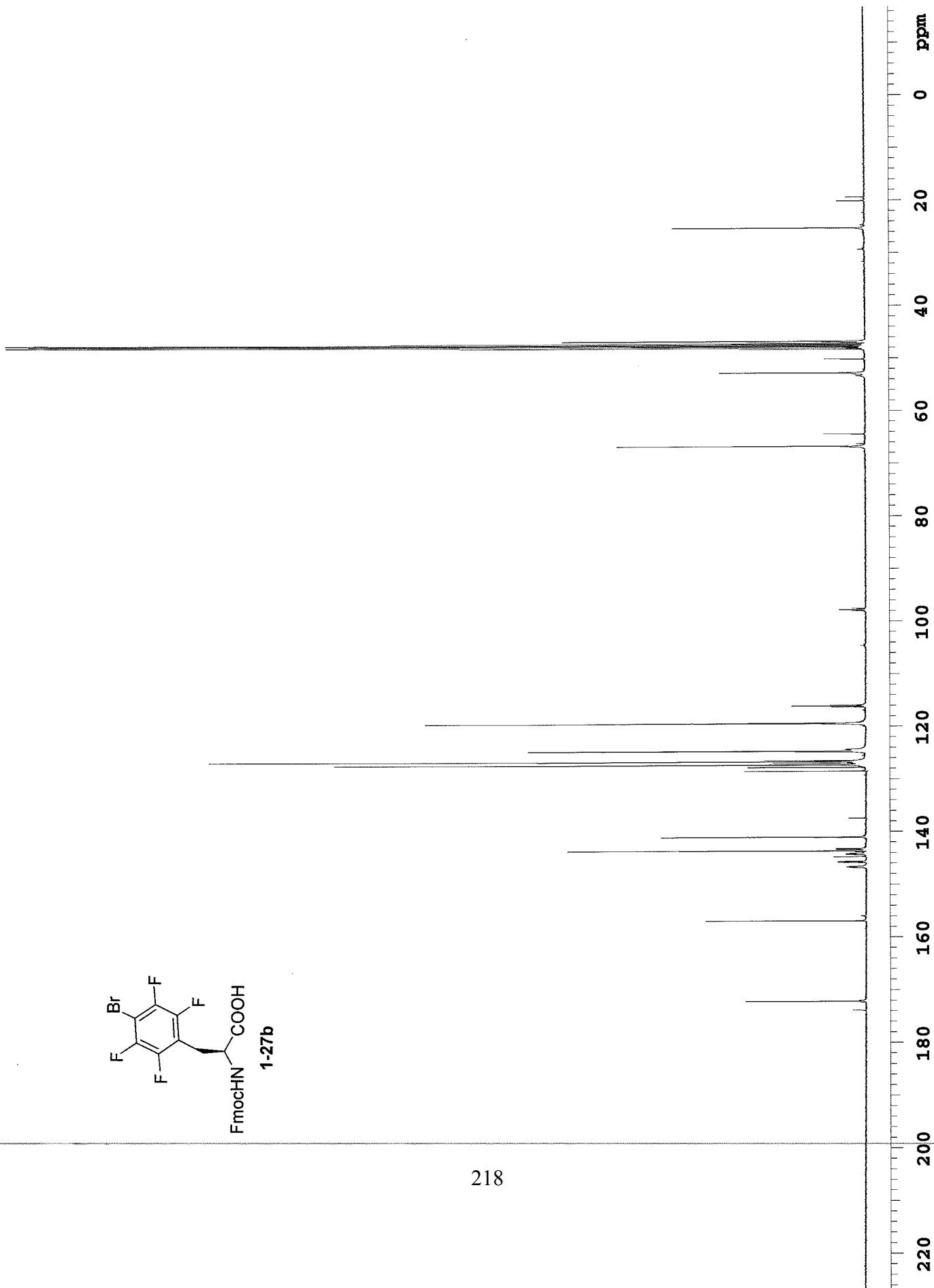
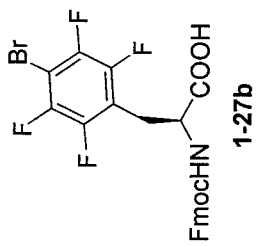


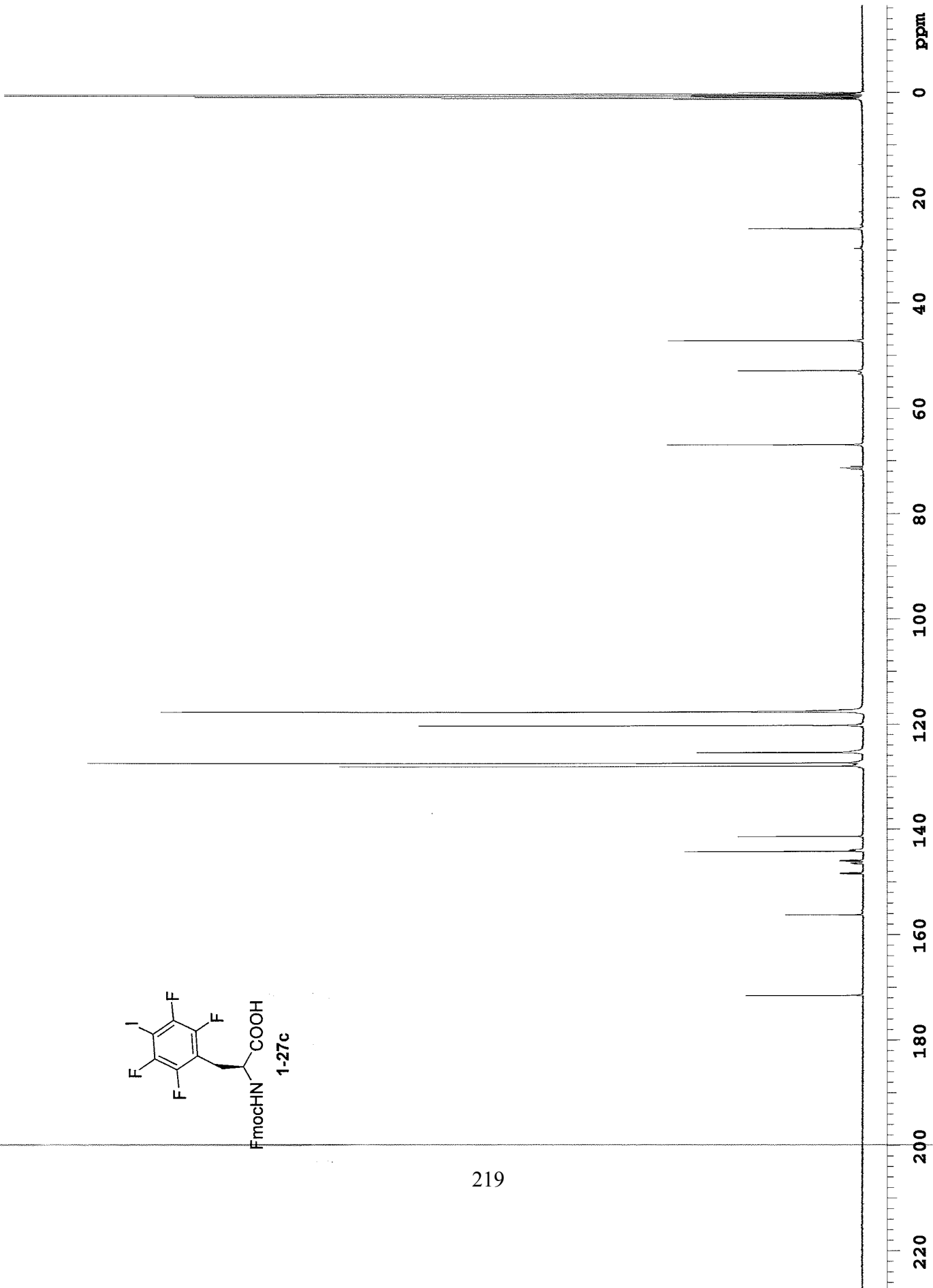
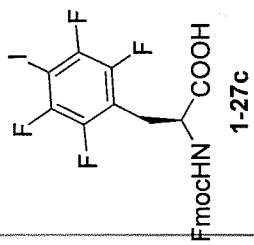


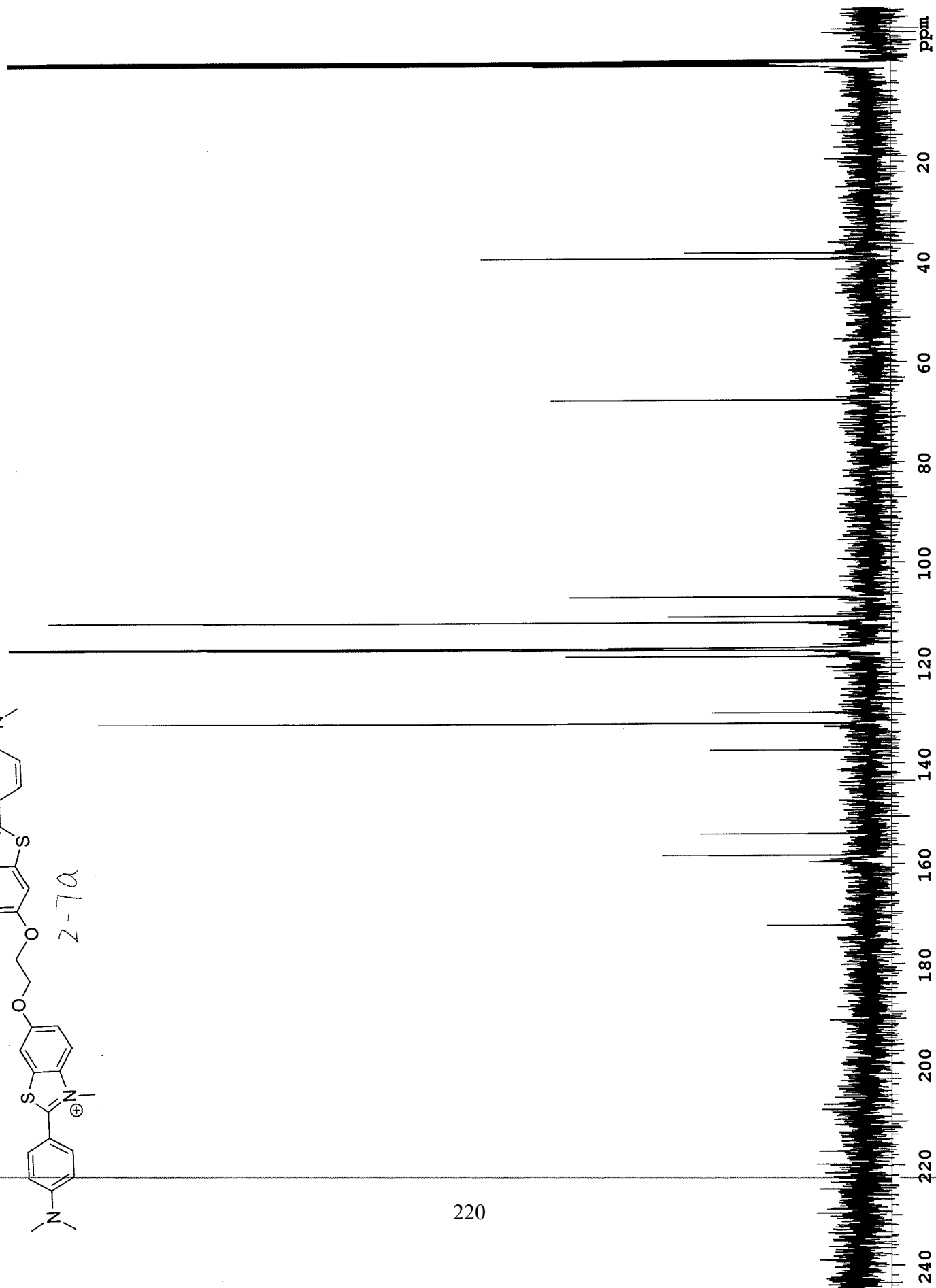
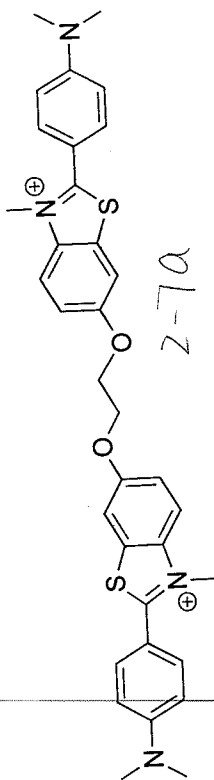


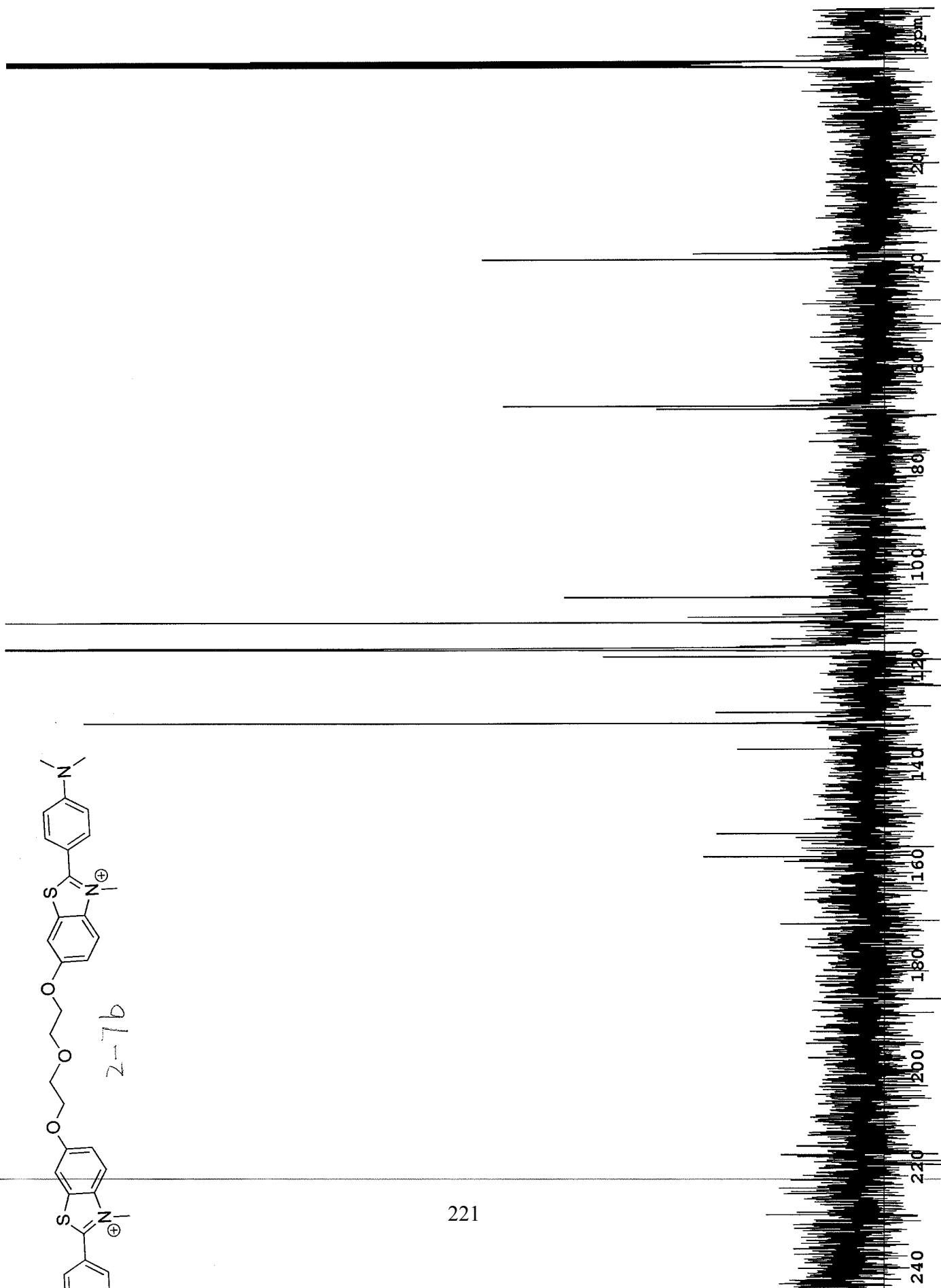
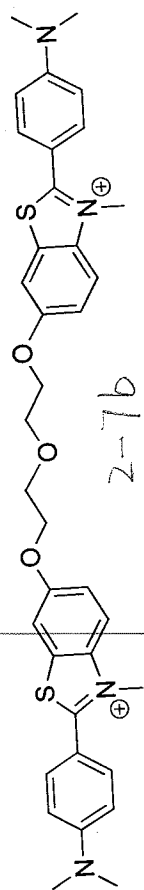




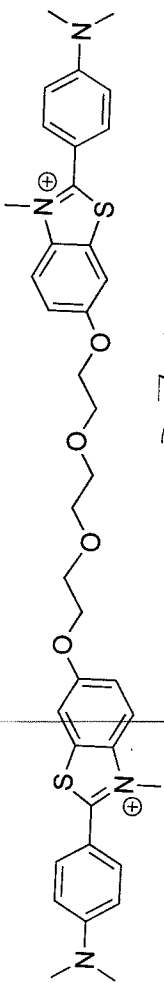






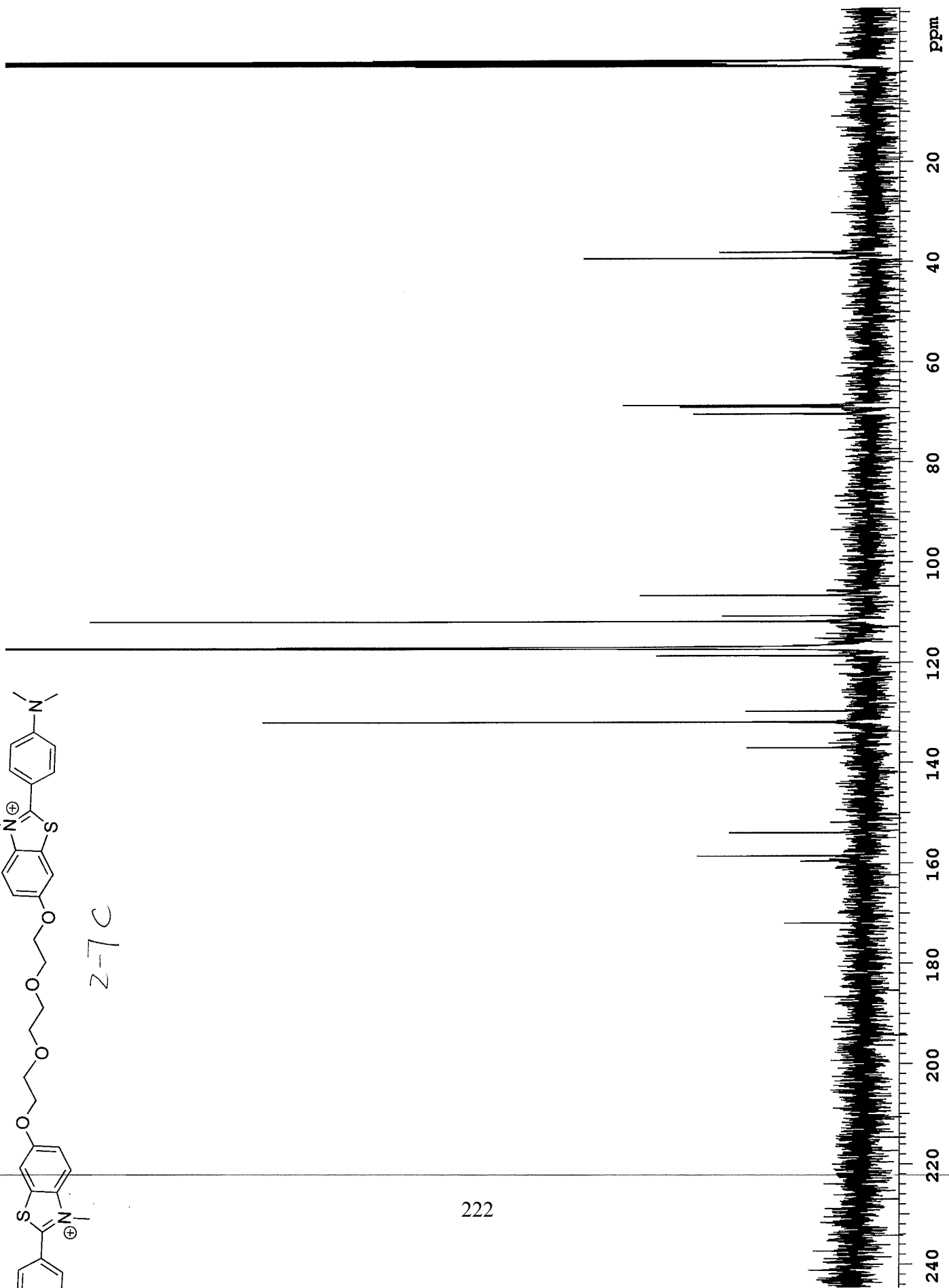






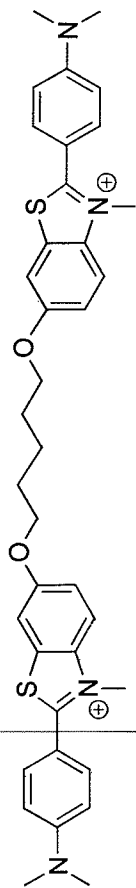
z-7c

222



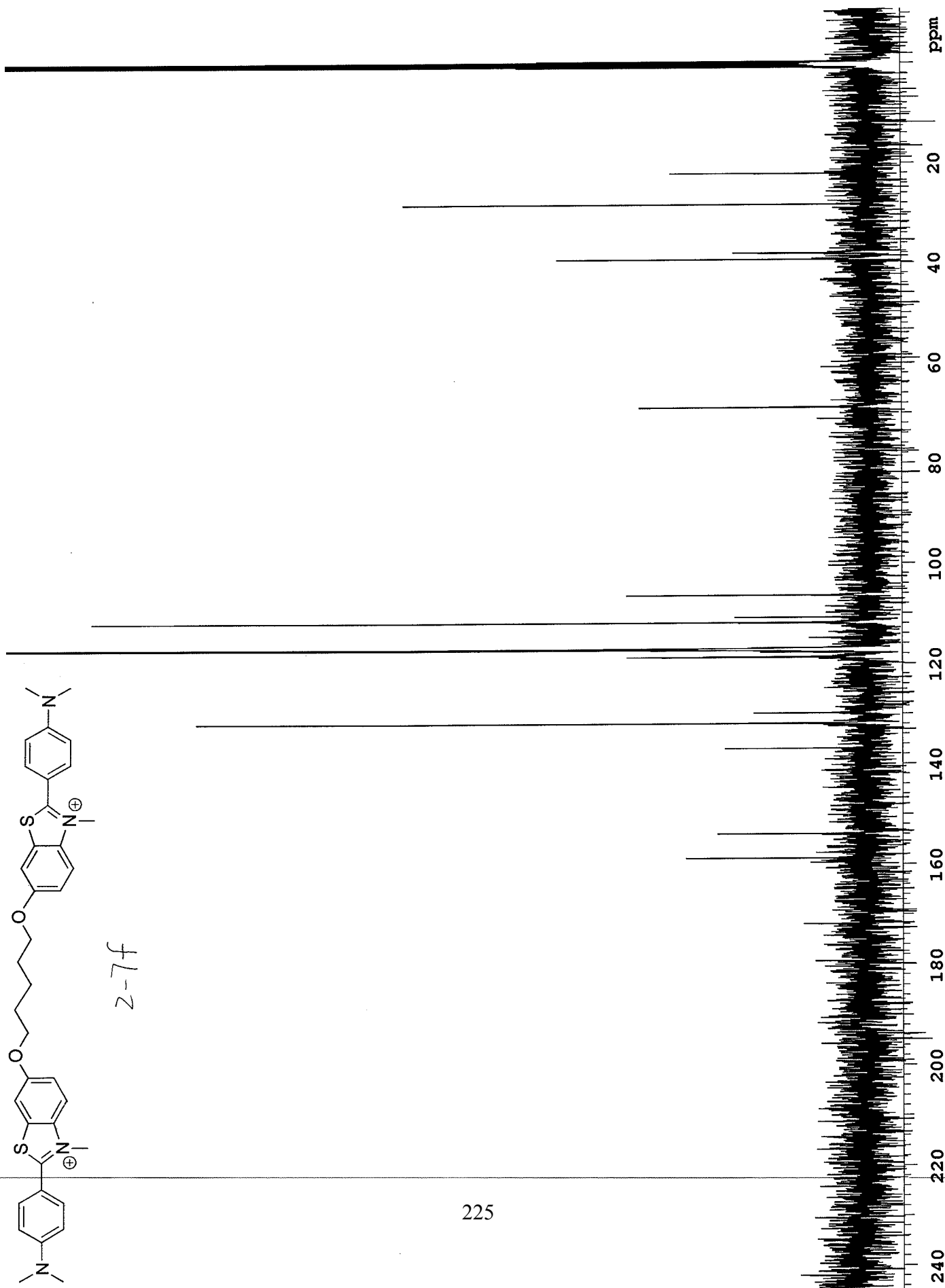


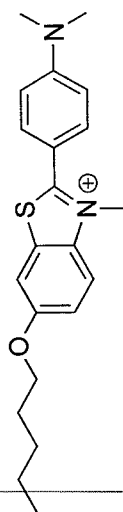




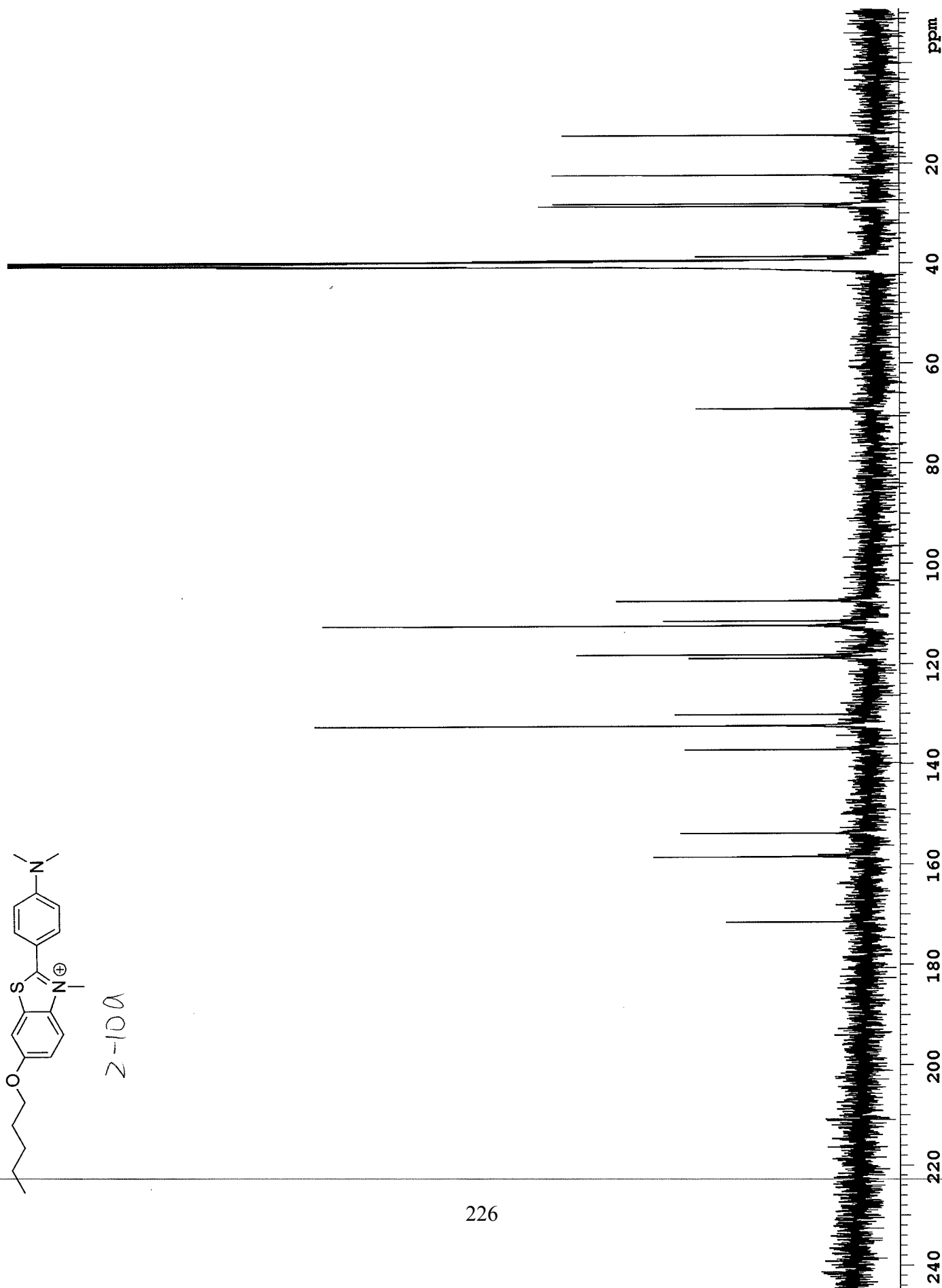
2-7f

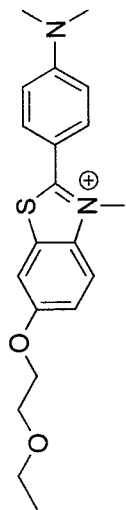
225



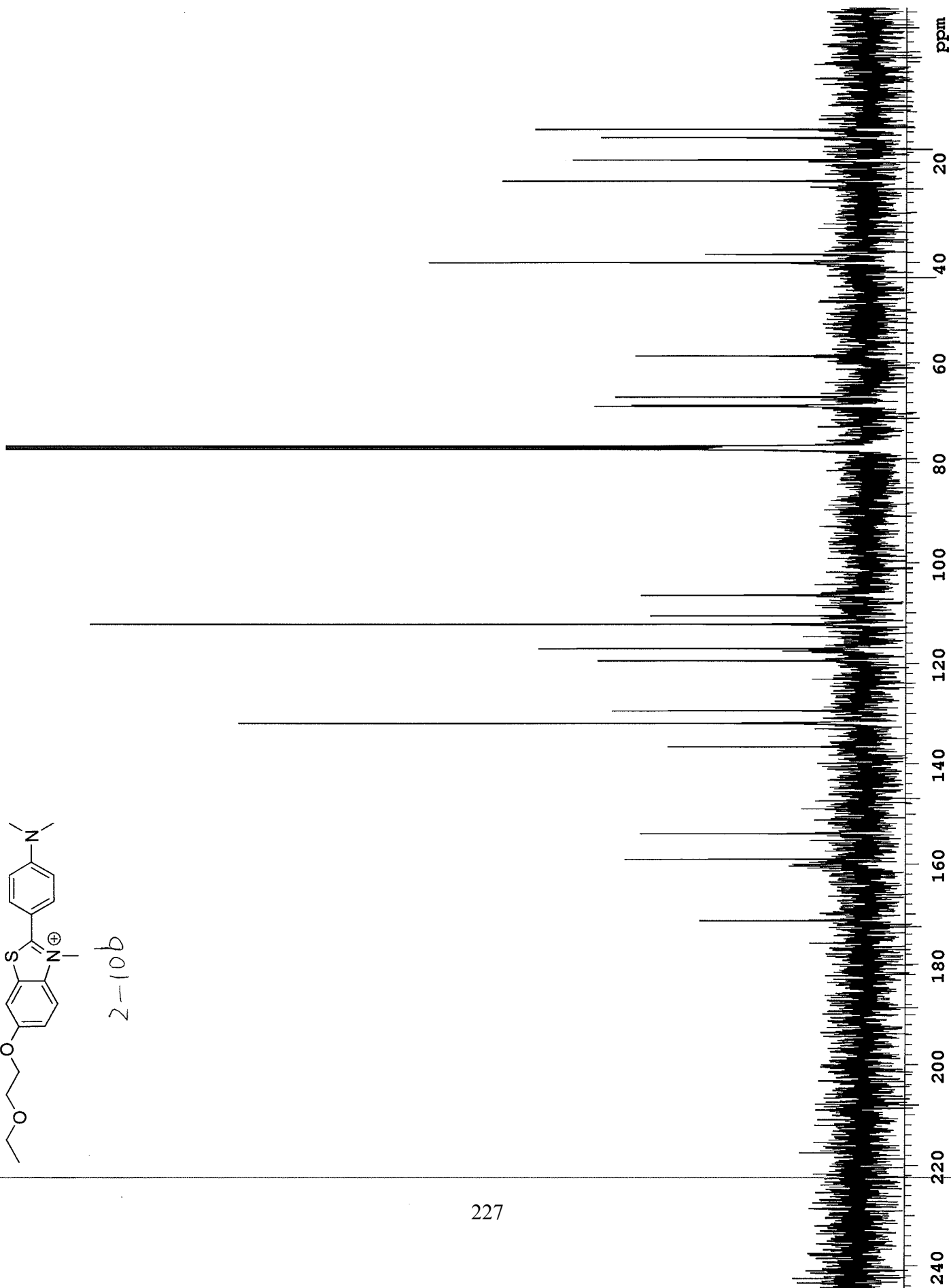


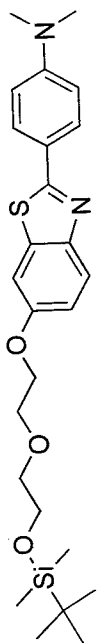
Z-100



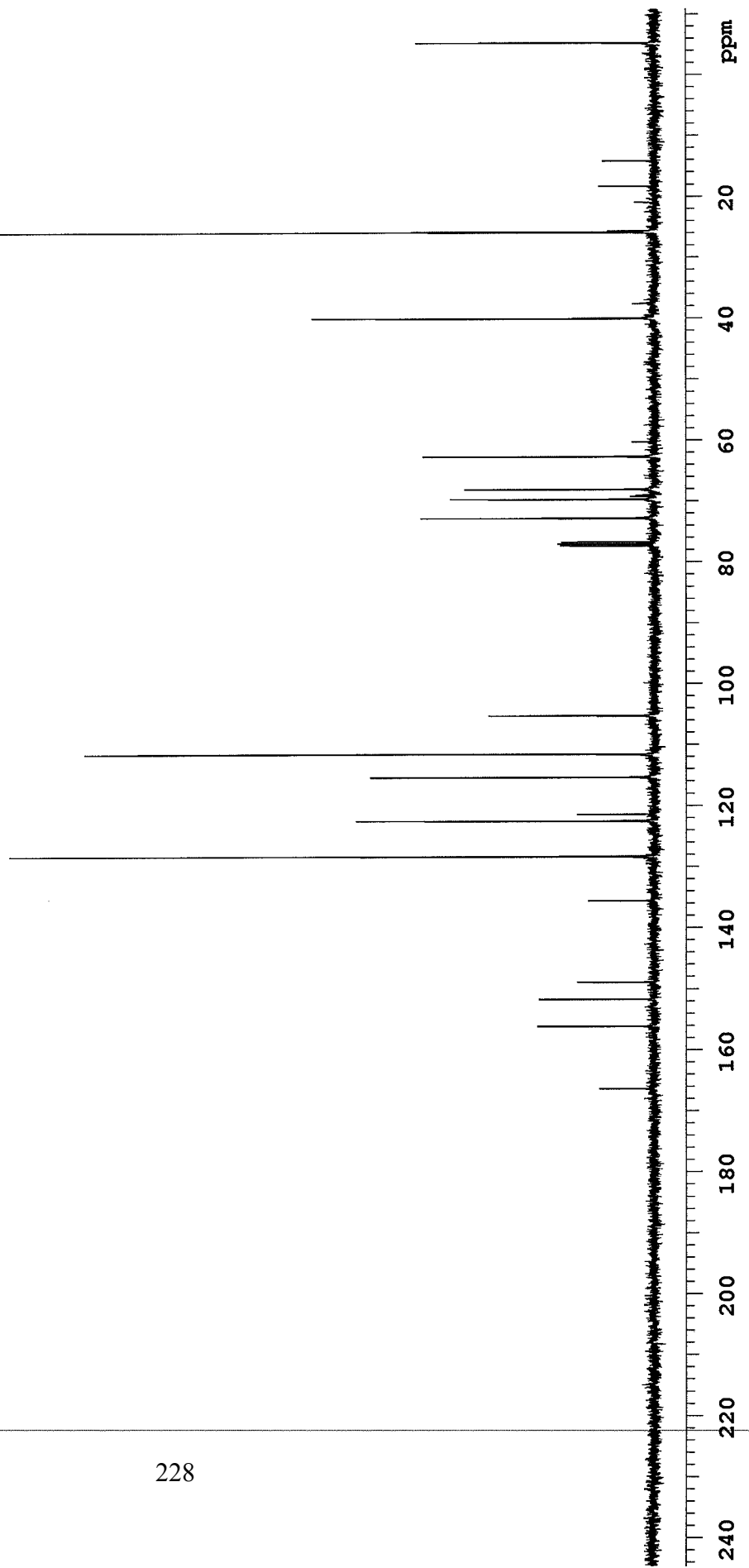


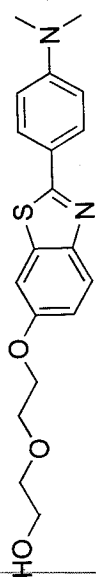
2-10b



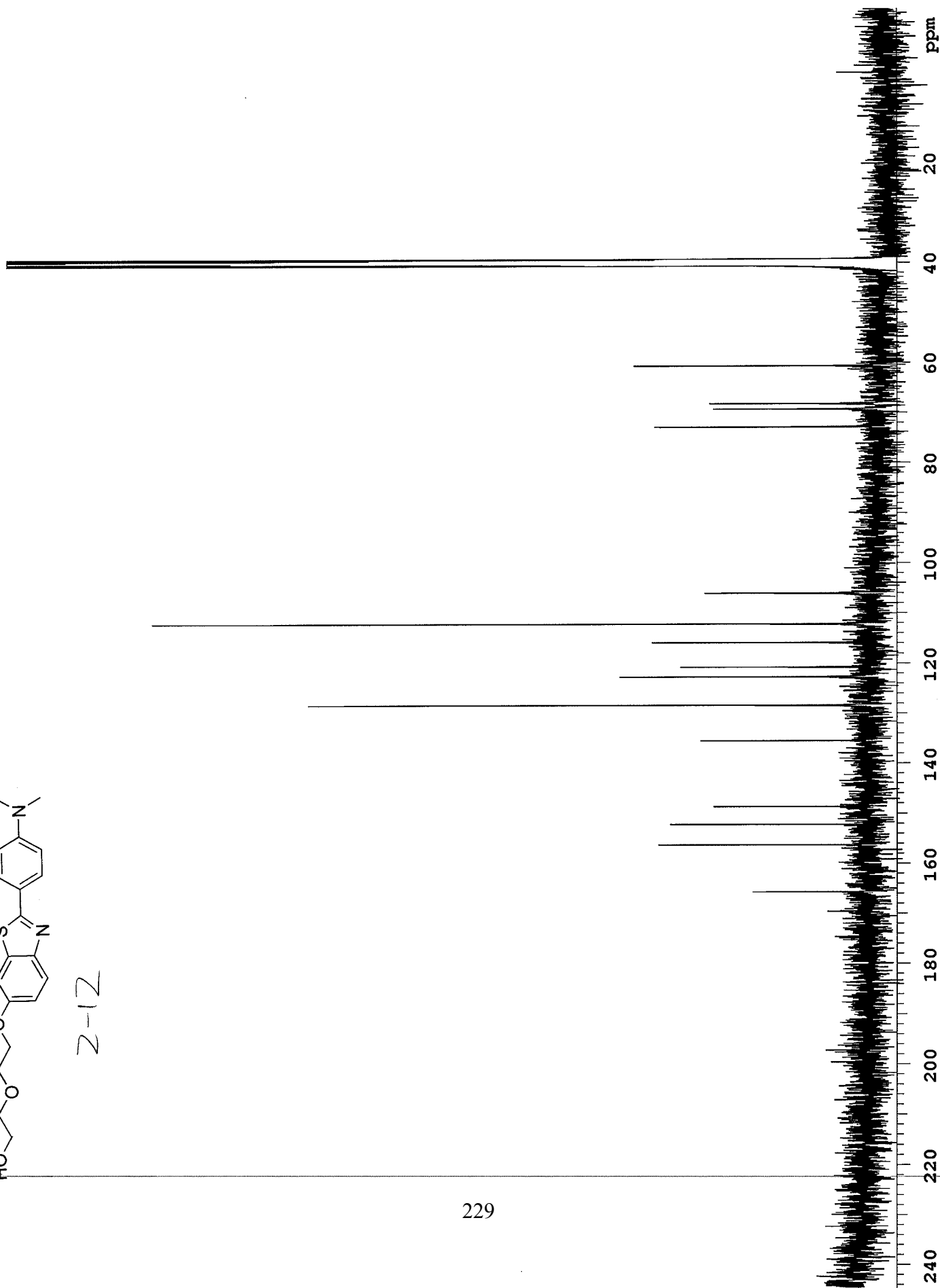


2-11

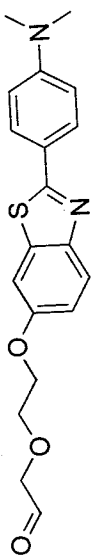




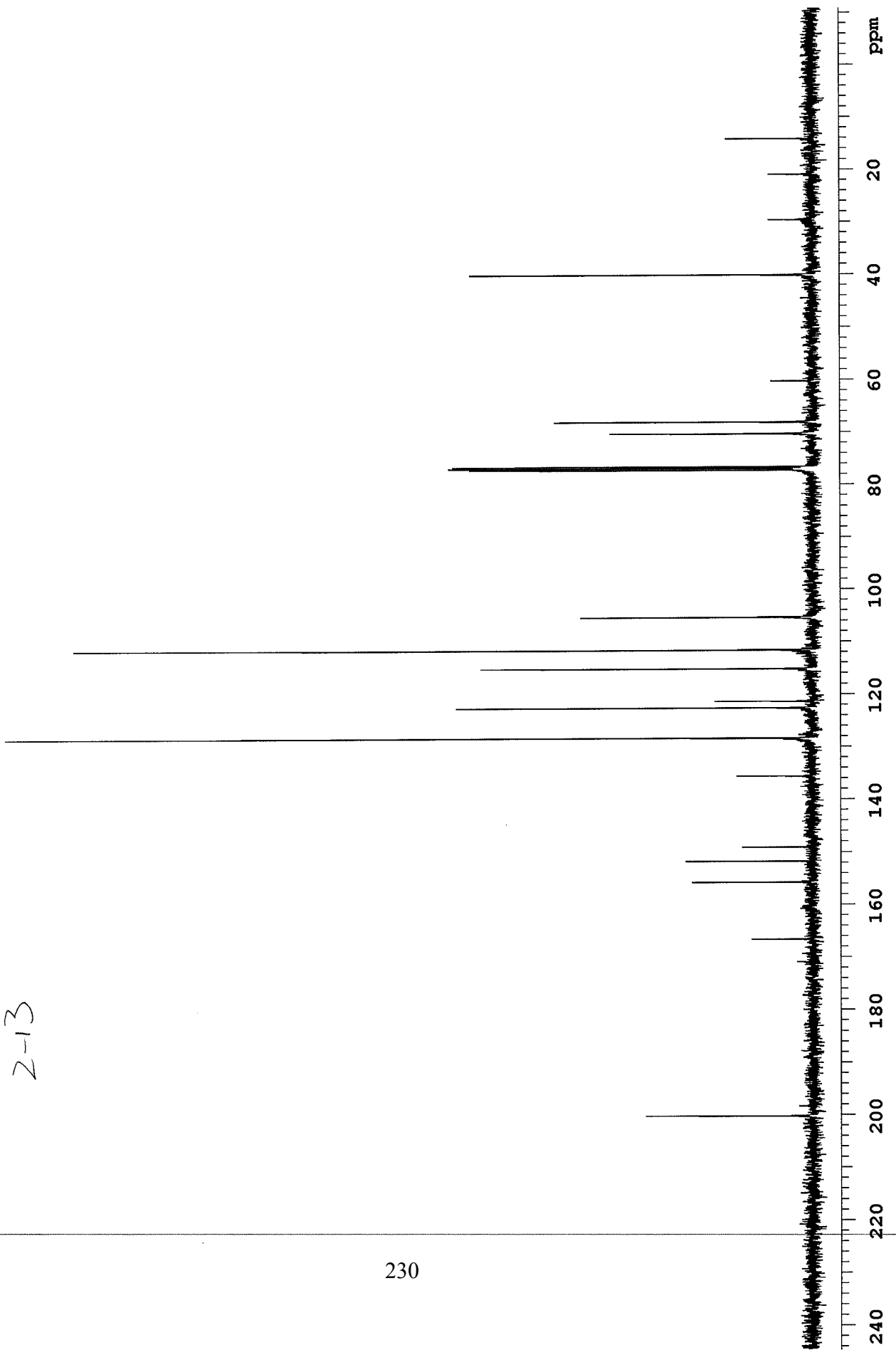
2-12

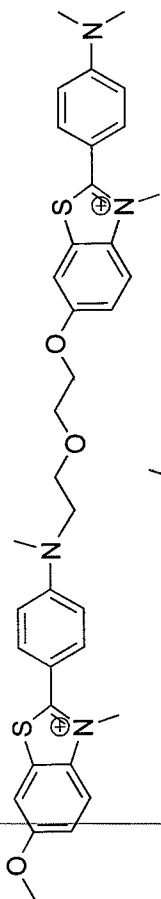




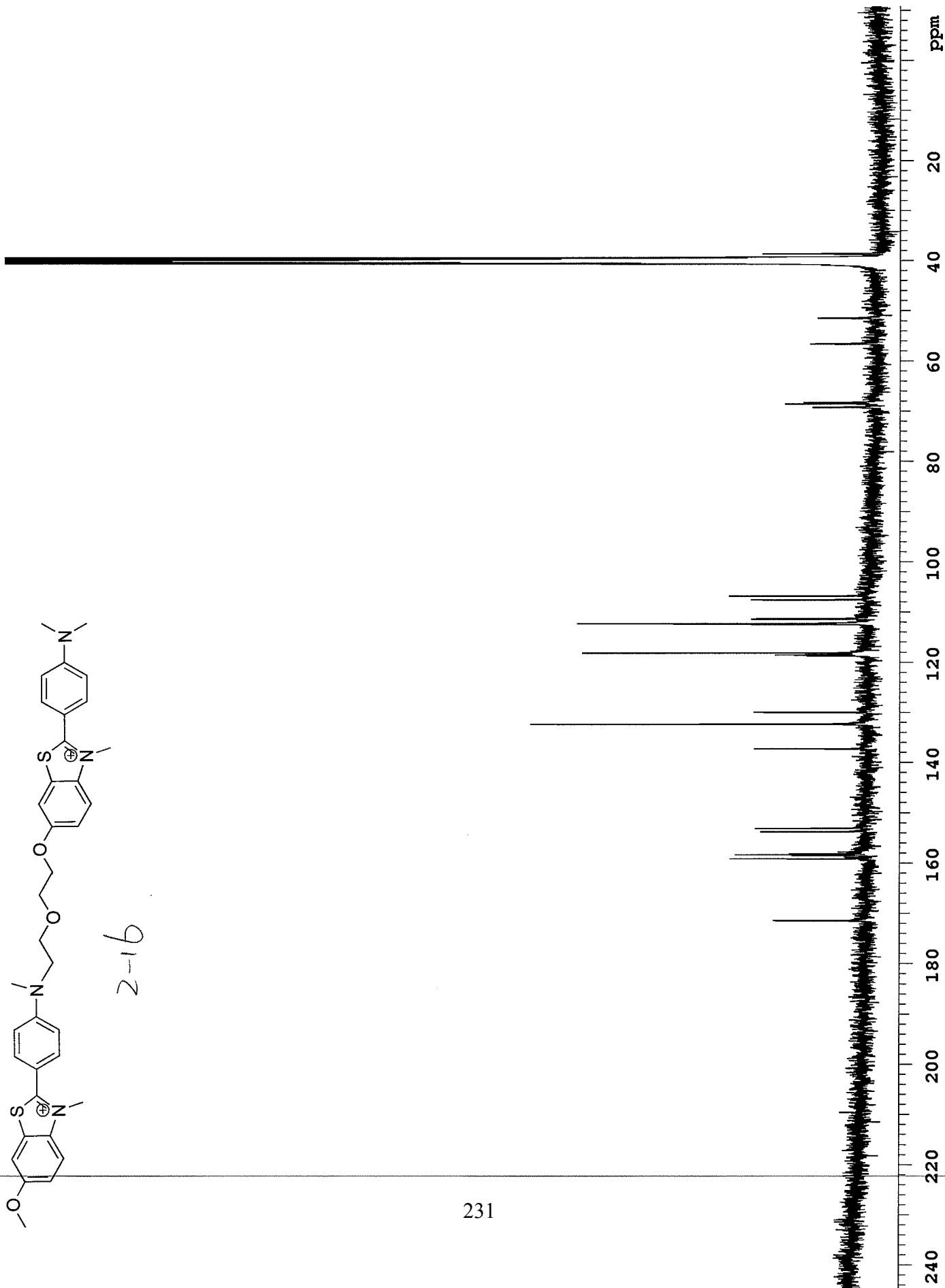


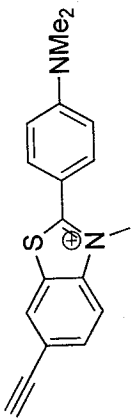
2-13



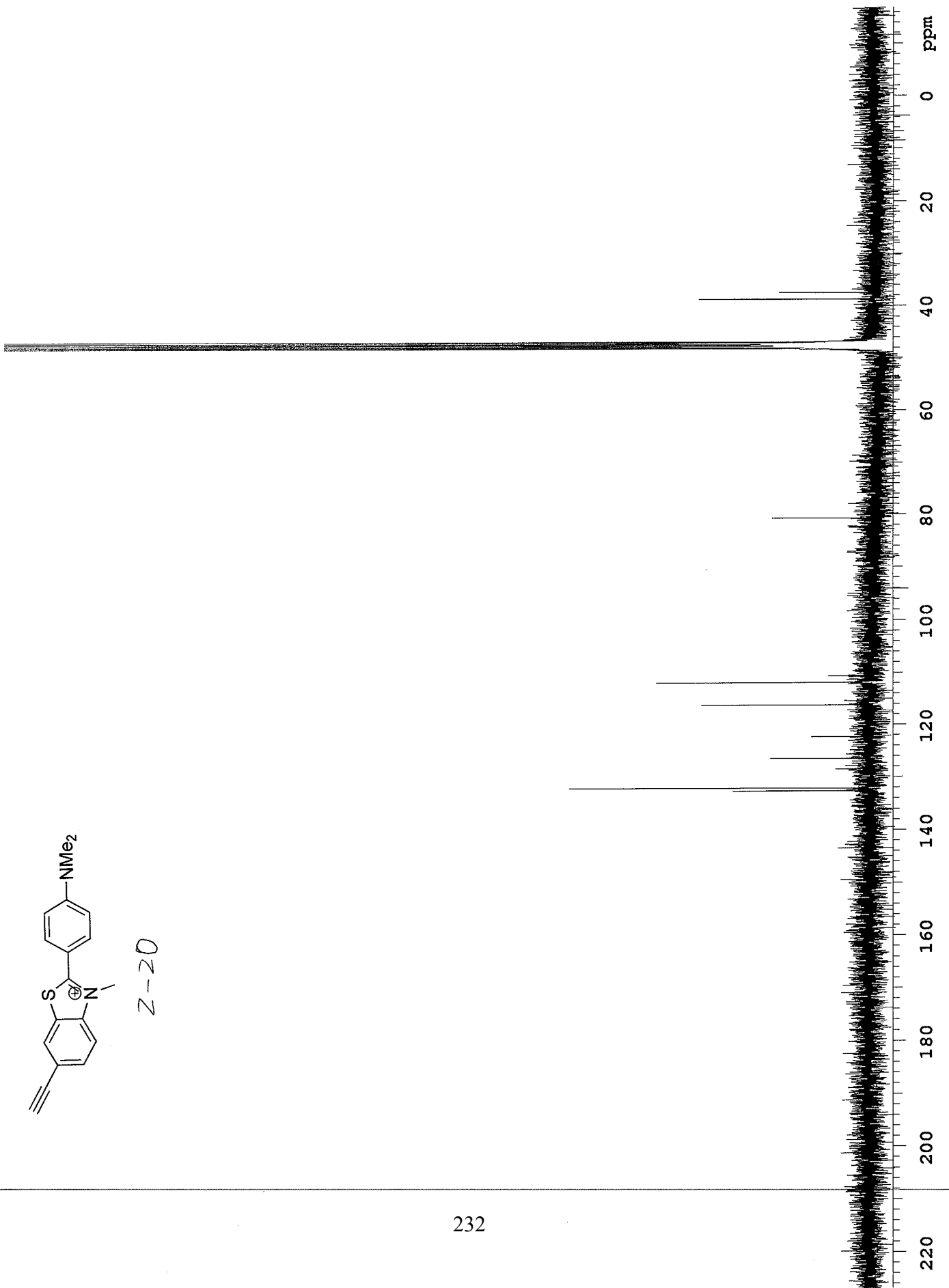


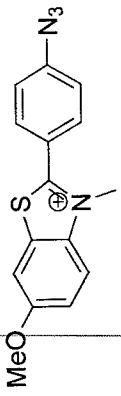
2-16





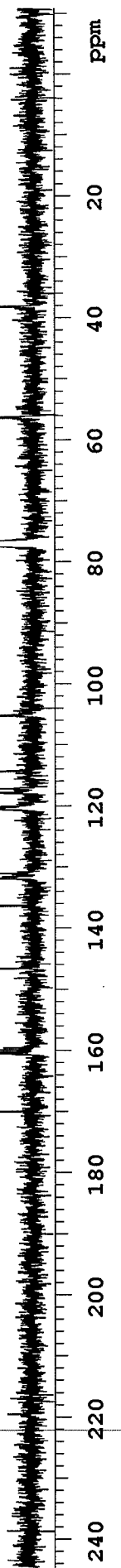
Z-20

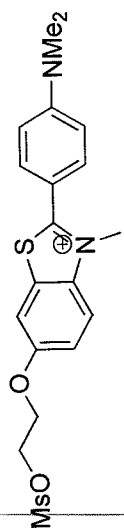




2-21

233





2-30

