Design and Synthesis of Small Molecules for Specific Targeting of Proteins by Non-Covalent Interactions

Avid Hassanpour

A Thesis in the

Department of Chemistry & Biochemistry

Presented in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Concordia University

Montréal, Québec, Canada

May, 2014

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Avid Hassanpour

Entitled: Design and Synthesis of Small Molecules for Specific Targeting of Proteins by Non-Covalent Interactions

and submitted in partial fulfillment of the requirements for the degree of

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Michael Sacher — Chair

Dr. Eric Marsault — External Examiner

Dr. Andreas Bergdahl — External to Program

DR. Xavier Otthenwaelder — Examiner

Dr. Christopher Wilds — Examiner

Dr. Pat Forgione — Thesis Supervisor

Approved by

Chair of Department or Graduate Program Director

Dr. Heidi Muchall

Dean of Faculty

Abstract

Design and Synthesis of Small Molecules for Specific Targeting of Proteins by Non-Covalent Interactions

Avid Hassanpour, Ph.D.

Concordia University, 2014

Hepatitis C virus (HCV) is a small, enveloped virus with a positive stranded RNA genome that encodes a polyprotein of approximately 3000 amino acids. HCV polyprotein requires two distinct proteases for viral replication, the NS2 and the NS3. HCV NS2/3 protease is a cysteine protease that features a highly unusual active site where NS2 forms a dimer with the active site cysteine on one monomer and histidine and glutamate on the other. Initial *in vitro* studies for identification of site-derived cleavage products as inhibitors of NS2 protease showed that the decapeptide from the N-terminal cleavage inhibits the cleavage between NS2 and NS3 with an IC_{50} of 90 μM thus providing a useful starting point for the development of other inhibitors in the current studies.[1] Our approach involves the synthesis of truncated NS2 site-derived peptide cleavage products in order to evaluate the importance of hydrogen bonding interactions that are essential for host recognition and to aid in the reduction of the size of the polypeptide chain. Therefore, a number of truncated peptides were synthesized through solid-phase peptide synthesis. Furthermore, immunoblotting and UPLC-MS/MS methods were used for characterization of the NS2/3 protease and quantification of the peptide inhibitors. Our results demonstrate that a hexapeptide

has encouraging potency towards inactivation of the NS2/3 auto-cleavage process.

The second project involves synthesis of islet amyloid polypeptide (IAPP) helical mimetics. Misfolding and aggregation of islet amyloid fibrils lead to the conversion of their secondary structure into cytotoxic β-sheet aggregates. Deposition of islet amyloid fibrils is related to the development and progression of type II diabetes, since their aggregation causes the impairment and death of the pancreatic beta cells.[2] In order to prevent the formation of amyloid fibrils, we have designed aryl-substituted heteroaromatic core scaffolds to direct the secondary structure of pro-amyloidogenic peptides into non-amyloidogenic conformers to mimic and induce/stabilize the IAPP helical state. A range of 2,5-diarylated thiophenes were synthesized as small molecule mimetics of the α-helix to modulate the amyloidogenesis and cytotoxic effect of islet amyloid polypeptide. 3-Substituted thiophene-2-carboxylic acids were used as key intermediates and functionalized by palladium decarboxylative cross-coupling and direct CH activation successively with overall yields ranging from 23 to 95%. The effect of the ligands on IAPP amyloid fibril formation was evaluated with the thioflavin T (ThT) fluorescence-based assay. Furthermore, the capacity of these compounds to inhibit the cytotoxic effect of IAPP was assessed using β-pancreatic cells.

Acknowledgements

First of all, I would like to express my gratitude to Dr. Pat Forgione for allowing me to work in his laboratory on such an interesting and challenging project, and for his invaluable support and guidance. Pat provided a good balance of direction and freedom to explore a very new research topic to me. I thank him for his numerous helpful discussions and encouragement that motivated me to grow as a scientist. I also thank him for providing a fun environment outside of the school by introducing us to the Shat Collione barbecues. It has been memorable being a member of his research group.

I would like to thank my committee members, Dr. Xavier Ottenwaelder and Dr. Christopher Wilds for their help and useful advice over the years of my studies. The annual committee meetings helped me to improve my ideas and presentation skills. I am very grateful to Dr. Eric Marsault and Dr. Andreas Bergdahl for kindly agreeing to be my examiners.

My special thanks to Dr. Simon Woo for proofreading this document and providing invaluable suggestions for its improvement. I was lucky to have a chance to know him as a great scientist and individual.

During my work, I encountered numerous people who provided essential help and supported me along the work in various ways. I would like to thank Dr. Vladimir Titorenko for allowing me to work in his lab and use his instruments/materials. I would also like to thank two great Ph.D. students in his lab, Vincent Richard and

Adam Beach from whom I learned how to run gels and western blot. I always felt welcome in their lab. I would also like to thank Dr. Steve Bourgault, Dr. Lekha Sleno, Makan Golizeh, Dr. Arnim Pause, Dr. Sarah Welbourn, and Dr. Guillaume Lamoureux for useful discussions and collaborations. I am grateful for the scientific advice of Marc Ouellet during these years. Also I would like to thank Dr. Alexey Denisov and Alain Tessier for the technical help with the NMR and mass spectrometry instruments.

My thanks to all past and present fORGione lab members, many of who have become great friends. Thanks to Dr. Dirk Ortgies for all his help (also in proof-reading parts of my thesis) and always listening to me. Stephane and Alexandre (Sikhounak), Fei, Arison, Daniel and Kris, you provided a pleasant environment in which to work. I was lucky to supervise some great undergraduate students, Amy Wan, Roger Chakkal, Joyce Zaftis, and Daniel Davis. You guys were great individuals to work with.

I am very thankful to all my friends at or outside of Concordia. Paknoosh, Behnoush, Shaghayegh, Shirin, Fatemeh (Fatool), Solmaz, Parisa, Meena, Derek, Rasha, Manal, Nathalie, Samaneh, Nooshin, Nassim, Mohammad, Samuel, Marica, Joanne, and Ellie - - without you I would not be where I am today.

My biggest thanks goes to my parents who sacrificed so much for me to have a better education and gave me endless love and encouragement. Above all, my eternal gratitude goes to my husband, Peyman, for all his love, patience, continuous support, and encouragement. I would like to dedicate this dissertation to the two loves of my life, my mum and Peyman.

Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

g	gravity
3' NTR	3' non-translated region
5' NTR	5' non-translated region
®	registered trademark
ACN	acetonitrile
Arg	arginine
Asp	aspartic acid
Bcl-2	B-cell lymphoma 2
Bcl-xL	B-cell lymphoma-extra large
BINAP	(2,2'-bis(diphenylphosphino)-1,1'-binaphthyl)
BQ	1,4-benzoquinone
Calmodulin, CaM	calcium-modulated protein
Cdc42	cell division cycle 42
CHO	formyl group
CMD	concerted metallation deprotonation

CN	nitrile
Congo red	disodium 4-amino-3-[4-[4-(1-amino-4-sulfonato-naphthalen-2-yl)diazenylphenyl]phenyl]diazenyl-naphthalene-1-sulfonate
Cys	cysteine
DavePhos	2-(dicyclohexylphosphino)-2'-(*N*,*N*-dimethylamino)-biphenyl
DIC	*N*,*N*'-diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropylethylamine
DM	*n*-dodecyl-β-D-maltoside
DM-2	diabetes mellitus type 2
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
E64	*N*-(*trans*-epoxysuccinyl)-L-leucine 4-guanidinobutylamide
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
Fmoc	fluorenylmethoxycarbonyl
GAGs	glycosaminoglycans
Gln	glutamine
Glu	glutamic acid

Gly	glycine
HBTU	*O*-(Benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate
HCV	hepatitis C virus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFIP	hexafluoro-2-propanol
His	histidine
IAM	iodoacetamide
IAPP	islet amyloid polypeptide
Ile	isoleucine
IPA	isopropyl alcohol
IRES	internal ribosome entry site
kDa	kilodalton
LDAO	lauryldimethylamine-oxide
Leu	leucine
Lys	lysine
Me	methyl
mesylate	methanesulfonate
MS	mass spectrometry
NS2/3	hepatitis C virus non-structural protein 2 and 3

NTPs	natural nucleoside triphosphates
OMe	methoxy
PBS	phosphate buffered saline
PEG-IFN	pegylated interferon
Phe	phenylalanine
pivalate	dimethylpropanoate
PVDF	polyvinylidene fluoride
QTOF	quadrupole-time-of-flight
rpm	revolutions per minute
S-Phos	sodium 2'-dicyclohexylphosphino-2,6-dimethoxy-1,1'-biphenyl-3-sulfonate hydrate
SDS	Sodium dodecyl sulfate
Ser	serine
SMDs	small-molecule drugs
smMLCK	smooth muscle myosin light chain kinase
SPE	solid phase extraction
SPP	signal peptide peptidase
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid

Thioflavine T	4-(3,6-dimethyl-1,3-benzothiazol-3 -ium-2-yl)-*N*,*N*-dimethylaniline chloride
ThT	thioflavin T
TIC	total ion current
TIS	triisopropylsilane
TLCK	tosyl lysine chloromethyl ketone
tosylate	*p*-toluenesulfonate
TPCK	tosyl phenylalanyl chloromethyl ketone
Tris	tris(hydroxymethyl)aminomethane
UHPLC	ultra high performance liquid chromatography
Val	valine

Chapter 1

General Introduction

1.1 Small Molecule Inhibitors of Proteins

Many drugs act by modulating the misregulation of a receptor/enzyme either by decreasing or increasing the activity of their target. Small molecule drugs (SMDs) as antagonists against the aberrant function of proteins are a major subset of drug molecules. SMDs are of great interest to the pharmaceutical industry and many advances have been developed in drug discovery in order to design and synthesize selective small-molecule inhibitors of a desired target with orally bioavailable properties. For example, compounds that target specific polo-like kinase (PLK1) activity (**a**)[4] or MDM2-p53 protein-protein interaction (**b**)[5,6] for cancer treatment, inhibit HIV-1 enzyme (**c**)[7–9] for the treatment of AIDS, and disrupt protein aggregation for the treatment of Alzheimer disease (**d**)[10] are shown in Figure 1.1.

Several guidelines exist to estimate the key parameters of developing drugs with proper solubility, permeation and absorption. For example in the well-known Lipinski's rule[11] a molecule with less than 5 hydrogen bond donors and 10 hydrogen bond acceptors, Log P less than 5 and a molecular weight of less than 500 is likely to be considered as a bioavailable drug.

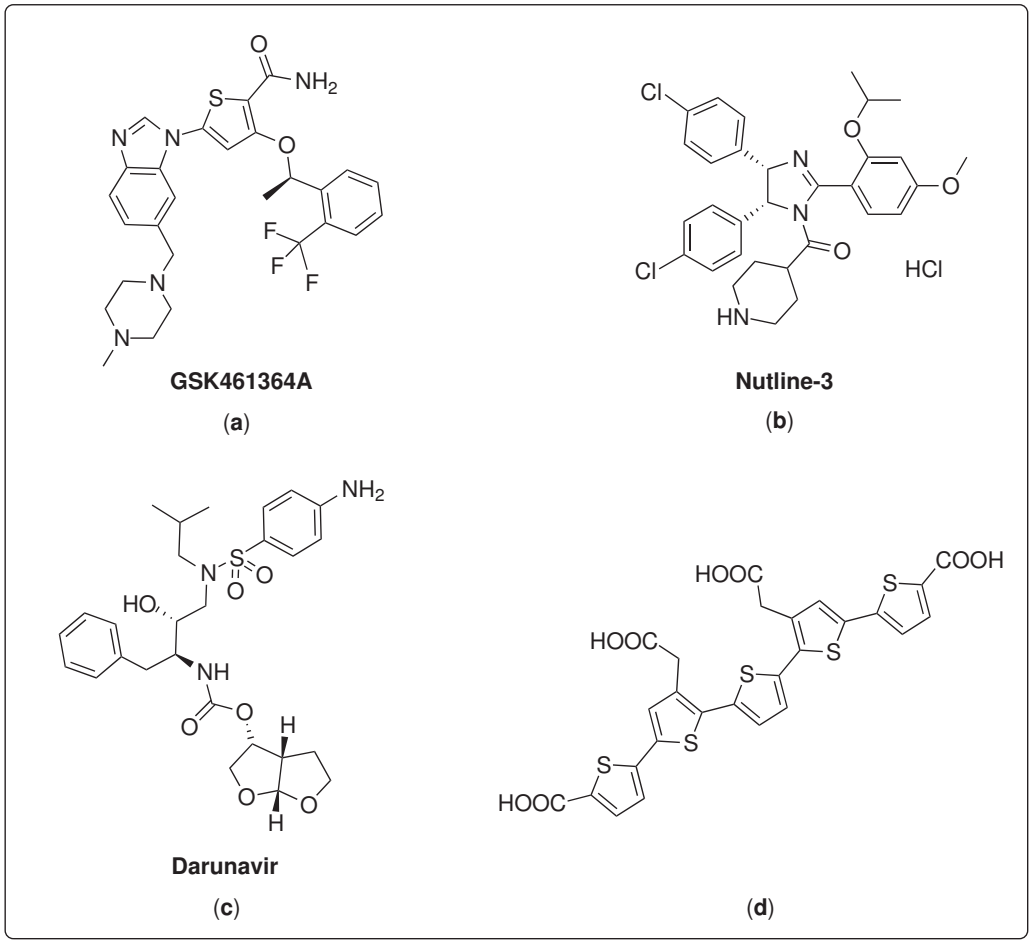


Figure 1.1: Examples of small-molecule inhibitors of enzyme activity and protein-protein interactions

In order to inhibit the irregular functionality of their target, small-molecule inhibitors are further divided into reversible and irreversible inhibitors. Irreversible inhibition usually results from covalent binding of the inhibitor to the active or allosteric site of the enzyme/receptor. The early examples of irreversible inhibitors possess reactive functional groups such as α-halo ketones, diazomethyl ketones and epoxides, which can covalently modify the target. The starting strategy in designing these inhibitors is often to attach such "warheads" to the natural substrate-based peptides in order to selectively deliver the electrophilic "warhead" of the inhibitor to the target protein. A substantial number of relatively safe and successful covalently bound inhibitors have been marketed as effective medicines[12–18] (Figure 1.2). Despite

their elevated potency, usually such "warhead"s lack specificity, possibly resulting in binding to diverse proteins and DNA which, can cause side effects. For example, peptidic halomethyl ketones bind to both serine and cysteine proteases through alkylating the active site of the enzyme similar to the alkylation observed by iodoacetamide.

Therefore, the use of covalent inhibitors is usually limited to study the mechanism of a specific target's action or for acute diseases when a low concentration of drug can be used for a short period of time.

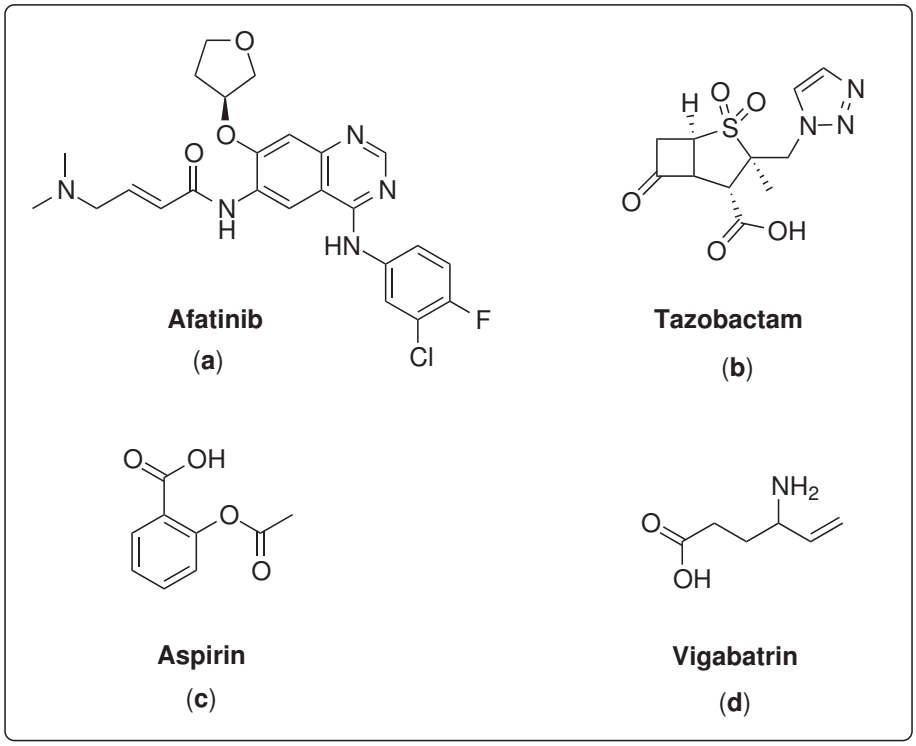


Figure 1.2: Covalent, small-molecule irreversible inhibitors. Target enzyme: (**a**, Afatinib): Epidermal growth factor receptor (EGFR) kinases and human epidermal growth factor receptor-2 (HER-2),[15] (**b**, Tazobactam): β-Lactamase;[17] (**c**, Aspirin): Cyclooxygenases COX-1 and COX-2[12] (**d**, Vigabatrin): GABA transaminase.[18] The S-enantiomer is active

Reversible inhibitors have overcome some of the issues of irreversible inhibitors. Reversible inhibitors generally bind to the enzyme/receptor via non-covalent interactions, including van der Waals forces, hydrophobic interactions and ionic and hydrogen bonds (Figure 1.3 **a** and **b**). However, some reversible inhibitors may also operate through a labile covalent but reversible bond with the target protein. Molecules with an aldehyde, nitrile and α-ketoamide "warhead" are examples of this group of

inhibitors (Figure 1.3 **c** and **d**).

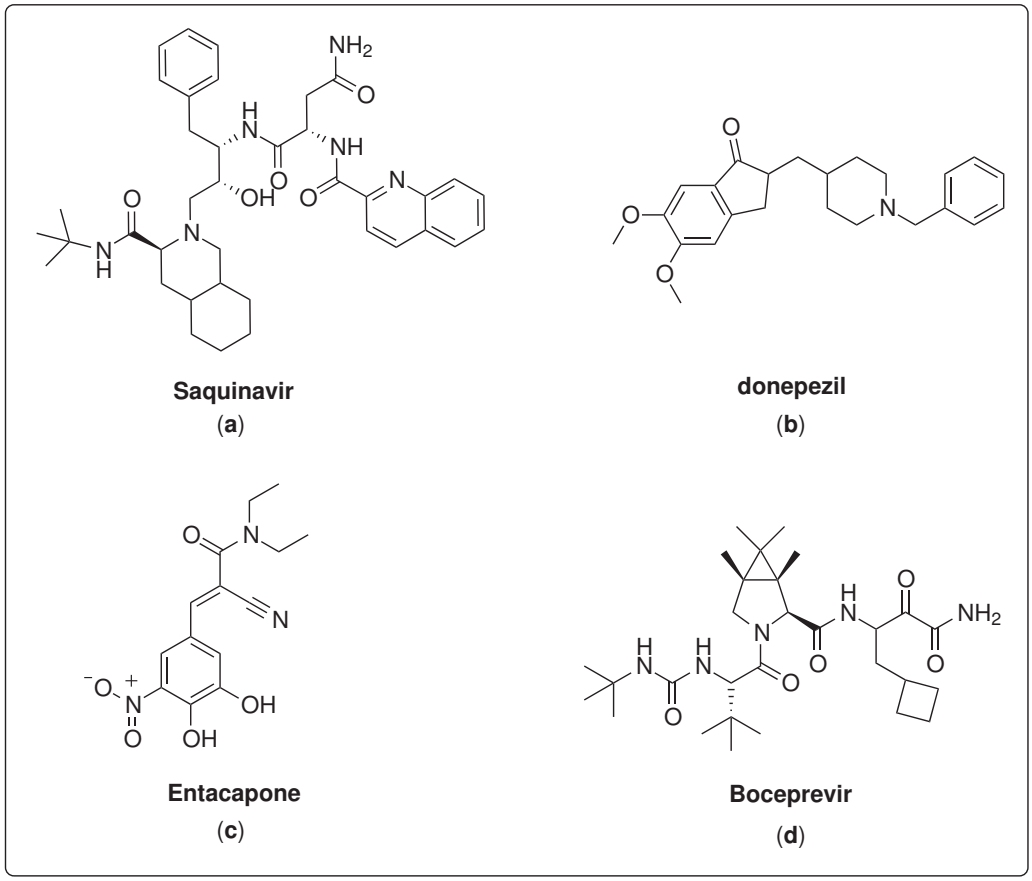


Figure 1.3: Covalent and non-covalent small-molecule reversible inhibitors. Target: (**a**, Entacapone): Catechol-O-methyltransferase (COMT) in treatment of Parkinson's disease;[19,20] (**b**, Boceprevir): NS3/4A serine protease in treatment of hepatitis C;[21,22] (**c**, Boceprevir): HIV protease in treatment of HIV;[23,24] (**d**, Donepezil): Acetyl cholinesterase in treatment of Azheimer disease[25,26]

Many parameters are involved in designing a small-molecule inhibitor with drug-like properties, so designing such a molecule demands a lot of time, effort, investment and the engagement of scientists in several areas. Small molecule inhibitors for a given target can be designed by several approaches. One traditional method is to screen libraries of compounds against a target to find a relatively potent inhibitor as a starting point, and then modify the selected compounds towards improved potency, selectivity, and ADMET properties (absorption, distribution, metabolism, excretion and toxicology). This method has been mostly used in pharmaceutical companies when libraries of thousands of compounds are accessible.

1.2 Thesis Perspective

Towards the design and synthesis of non-covalent small-molecule inhibitors, we have investigated mechanistically and structurally diverse targets: the NS2/3 protease of hepatitis C virus (HCV) and the α-helix of islet amyloid polypeptide.

The first project involves the study of the NS2/3 enzyme that is one of the two non-structural proteases of the hepatitis C virus. HCV NS2/3 protease is a cysteine protease that features a highly unusual active site where NS2 forms a dimer with the active site Cys on one monomer and the His and the Glu on the other. NS2/3 protease participates in the intramolecular cleavage of the enzyme such that replication of the virus occurs. Although NS2/3 protease processes a single cleavage between NS2 and NS3, it also has a significant role in viral assembly and RNA replication. This was shown by infection of chimpanzees with HCV containing a mutated NS2/3 protease.[27] Synthesis of small-molecule inhibitors of NS2/3 protease activity was initiated using the substrate-based peptide synthesis approach. However, the auto-cleavage activity of the NS2/3 protease hinders the ability of the inhibitors to compete with the substrate since commonly an intramolecular reaction is kinetically favored compared to an intermolecular reaction. The focus of this project was on the assay optimization for the NS2/3 auto-cleavage reaction by means of LC-MS and western blot techniques. Moreover, the synthesis and evaluation of some substrate-based peptides were investigated.

The second project involves synthesizing small molecules to target specific conformational states of islet amyloid polypeptide (IAPP) through non-covalent interactions. IAPP is an aggregation-prone peptide hormone that can undergo a secondary structural conversion into partially folded β-sheet intermediates, en route to the formation of amyloid fibrils. The misfolding and aggregation of IAPP in the pancreas lead to degeneration of the islets of Langerhans.[28] The strategy of this research was to target and trap the pro-amyloidogenic peptides and direct their secondary structure

into non-aggregating isomers through the development of small molecules capable of interacting (mainly hydrogen bond and π-π stacking interactions) with one side of the transient helical conformer of IAPP. By targeting this specific conformational state with small molecules, the equilibrium will be shifted from the pathogenic to the functional folded non-aggregating isoform. The other main focus of this project was to develop an efficient and modular synthetic pathway to allow for rapid synthesis of small molecules for exploring their structure-activity relationships to optimize various parameters including potency. Through the application of two palladium-catalyzed cross-coupling reactions, namely palladium-catalyzed decarboxylative cross-coupling and C–H activation reactions, this second goal was achieved. Finally, the results of evaluated compounds in the bio-assays are presented.

1.3 Contribution of the Author and Thesis Organization

Chapter 1 presents a general introduction related to both projects by introducing small-molecule inhibitors of proteins and considerations in design and development of small-molecule inhibitors.

Chapter 2 presents the introduction of the first project about the HCV NS2/3 protease. It introduces the background in HCV genome, life cycle, translation, and catalytic triad of the NS2/3 protease. It further introduces the two virally proteases and advances in development of inhibitors of these two proteases.

Chapter 3 discusses the results of the HCV NS2/3 project towards assay optimization, preliminary kinetic studies of NS2/3 proteases activity, site-derived peptide synthesis, and evaluation of the peptides. The discussion of this research is included in context.

In this project, the synthesis of all peptides, their purification and characterization

by mass spectrometry were performed solely by the author. Gel electrophoresis and western blot optimization and analysis were conducted by the author solely. Mass spectrometry experiments by Q-TOF were conducted at Concordia University by Jean-Pierre Falgueyret at Centre for Biological Applications of Mass Spectrometry (CBAMS). UHPLC-MS/MS experiments were conducted at Université du Québec à Montréal (UQÀM) in collaboration with Dr. Lekha Sleno's research group. Author participated in trypsin digestion experiments. Sample preparations were either conducted by the author solely or with Makan Golizeh and Dr. Lekha Sleno at UQÀM.

Chapter 4 presents the detail experimental of the NS2/3 protease project. The methodology, material and instruments employed in the first project is explained in detail.

Chapter 5 introduces the background of the second project about amyloid fibril formation and protein-protein interaction. The chemistry background of the project is followed by introducing palladium coupling, decarboxylative and C–H arylation reactions. In this project, the synthesis of all compounds was conducted by the author solely. The biological assays of this project such as thioflavine T and cytotoxicity assays were performed at Université du Québec à Montréal (UQÀM) through collaboration with Dr. Steve Bourgault and Carole Anne De Carufel. The X-ray crystallography of one compound was performed under supervision of Dr. Xavier Ottenwaelder by Dylan Mclaughlin and Mohammad Sharif Askari.

Chapter 6 presents the synthesis of 2,5-diaryl substituted thiophenes for modulation of islet amyloid polypeptide (IAPP) amyloid fibril formation and cytotoxicity. The discussion of the research is included in the context. The work resulted in publication of "Synthesis of 2,5-Diaryl Substituted Thiophenes as Helical Mimetics: Towards the Modulation of Islet Amyloid Polypeptide (IAPP) Amyloid Fibril Formation and Cytotoxicity" in Chemistry - A European Journal.[29]

Chapter 2

HCV NS2/3 Protease

2.1 HCV Epidemiology

Hepatitis C virus (HCV) infection is the major cause of chronic liver disease that can lead to hepatic fibrosis, liver cirrhosis and hepatocellular carcinoma (HCC).[30] Based on the World Health Organization's estimation, more than 170 million people (3% of the world population) have been chronically infected by this virus worldwide and this number is annually increasing by 3-4 million.[31–34] Hepatitis C is a blood-borne disease and is mainly transmitted through contaminated blood transfusion and drug injection or injury with unsterilized syringes or needles.[35–37] Various surveys reveal that Africa (mainly Egypt and Cameroon) and the Middle East have the highest prevalence of HCV infection while Western Europe, Northern Europe and North America have the lowest (Figure 2.1).[38–40]

2.2 Therapy and Challenges

The prevailing therapy for chronic HCV infection is a combination of Peginterferon alpha (PEG-IFN-α) and ribavirin;[41–46] however, not only is this treatment highly genotype-, and age-dependent,[47,48] but it is also only effective in 50% of the patient

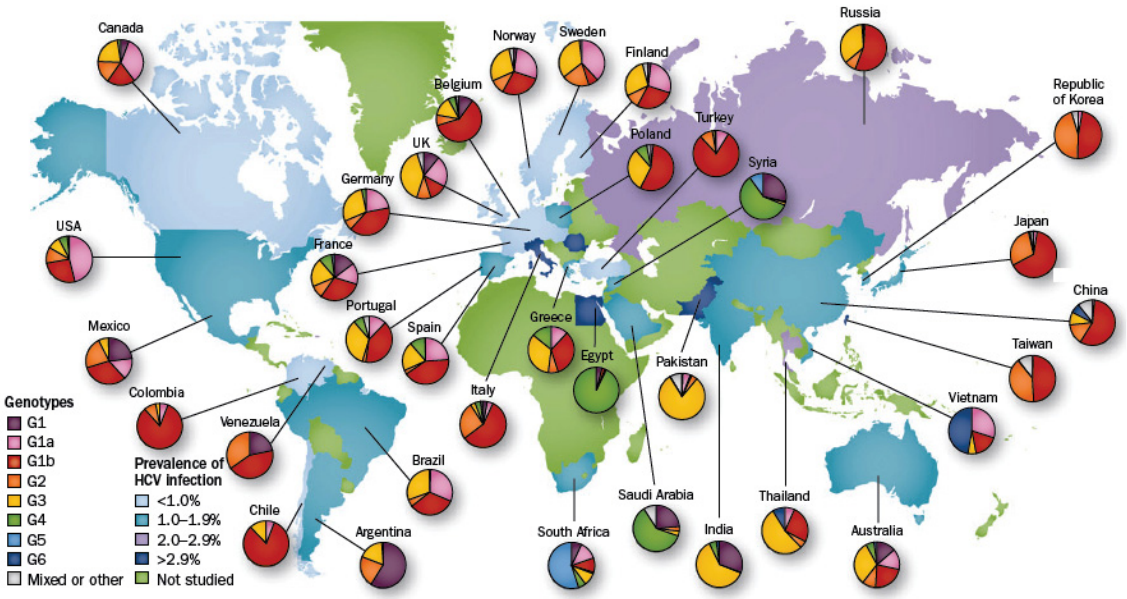


Figure 2.1: Worldwide HCV prevalence and genotype distribution. (Figure from reference [40])

population. Moreover, this medication suffers from many side effects such as exhaustion, depression, neutropenia, hemolytic anemia, anorexia and weight loss, dermatitis, pruritus, insomnia and flue-like symptoms that cause many patients to terminate the therapy.[49–52] As a result many attempts have been devoted to developing new therapies in recent years, and a number of compounds targeting NS3/4A protease, NS5A protein and NS5B polymerase (RNA-dependent RNA polymerase) have reached clinical trials.[53–60] For instance Boceprevir[21,22] (Victrelis®, Figure 1.3, **b**, Schering-Plough) and Telaprevir[56] (Incivek®, Figure 2.9, **d**, Vertex Pharmaceuticals) have been approved by the FDA for the inhibition of NS3/4A protease in 2011. In 2013 Sofosbuvir[61] (Gilead Sciences, Figure 2.2, **a**) was licensed by the FDA for inhibition of HCV NS5B polymerase. The drug is administrated with other antiviral drugs such as PEG-IFN-α and ribavirin. Other recently developed drugs for inhibition of NS5A protein include Ledipasvir[62] (Gilead Sciences, Figure 2.2, **b**) and Daclatasvir[63] (Bristol-Myers Squibb). These drugs are also used with standard PEG-IFN-α and ribavirin antiviral drugs as well as inhibitors of NS5B polymerase such as Sofosbuvir.

This highlights that more efficient approach for the battle against HCV disease relies on a combinational therapy where a patient is subjected to more than one drug simultaneously.

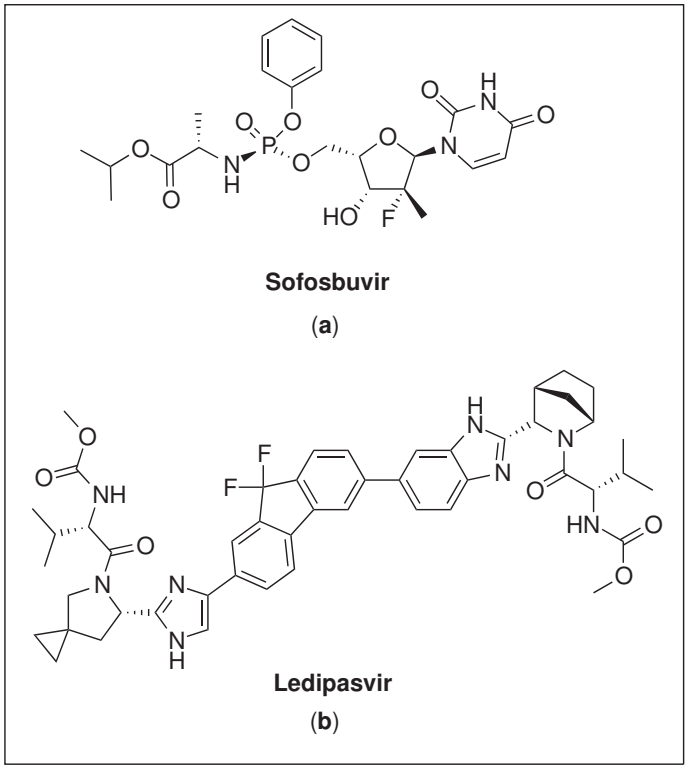


Figure 2.2: Inhibitors of NS5B polymerase and NS5A protein for the treatment of HCV. (**a**): Sofosbuvir was marketed by Gilead Sciences as inhibitor of NS5B polymerase; (**a**): Ledipasvir was developed by Gilead Sciences as inhibitor of NS5A protein

HCV has 6 general genotypes with each being classified into one or several subtypes.[40,47,64–66] The genotype distribution is very geographically dependent; for instance genotypes 1a, 1b, 2a and 2b are more prevalent in North America and Europe whereas genotypes 4 and 5 are more frequent in Africa. Variability between the genotypes is an issue since different genotypes contain varying genetic sequences and one specific inhibitor may not be effective for every genotype. Most pharmaceutical efforts have been focused on the most widespread genotypes (1a, 1b, 2b) in the western countries while other genotypes such as 4, 5, and 6 found exclusively in the third world countries have not been explored yet.

2.3 HCV Genome and Life Cycle

HCV is a positive, single stranded RNA virus. HCV is a member of hepacivirus *Flaviviridae* family similar to GB virus B (GBV-B) and tamarin virus.[67,68] The HCV IRES (internal ribosome entry site) at the 5' non-translated region (5' NTR) acts as one of the essential elements by which polyprotein translation initiates. The 3' non-translated region (3' NTR) contains essential factors for the viral genome replication[69] (Figure 2.3).

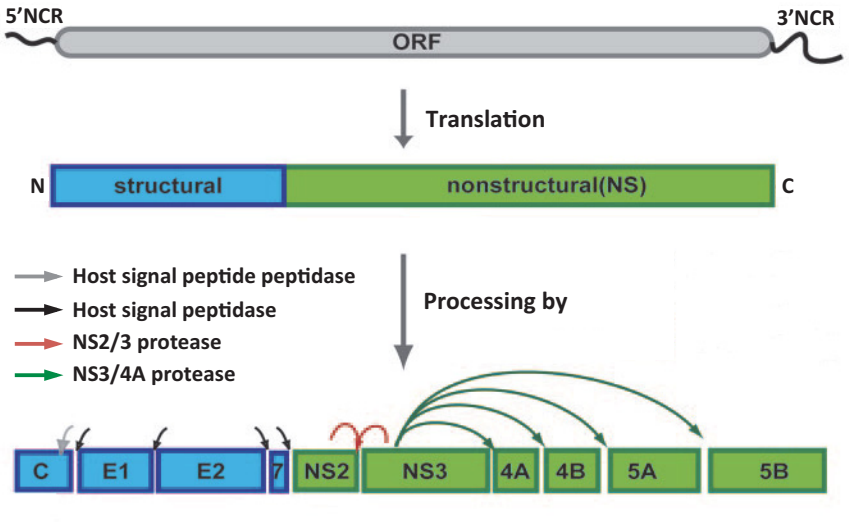


Figure 2.3: HCV genome translation and processing by host and viral proteases. (Figure from reference[70])

The polyprotein encompassing an open reading frame of approximately 3000 amino acids (9.6-kb) is processed into structural and non-structural proteins by cellular and viral proteins respectively.[71,72] Host cell signal peptide peptidase (SPP) releases the structural proteins (core, E1, E2, p7), whereas viral NS2/3 and NS3/4A proteases release the non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, NS5B) (Figure 2.3).

The HCV life cycle involves several steps depicted in Figure 2.4. After attachment of the virus through interaction of enveloped glycoproteins, E1 and E2, to the hepatic

cell receptors (specifically CD81,[73] SR-BI,[74] claudin-1[75] and occludin[76]) (Figure 2.4, **a**) the virus particle enters the host cell (endocytosis).[77] Due to its dependency on the pH of the cell environment, decapsidation of the viral particle in the cytoplasm takes place (Figure 2.4, **b**).[78] The released positive and single-stranded RNA functions as m-RNA and is translated to produce the HCV polyprotein that is then processed by cellular and viral proteases to structural and non-structural proteins respectively (Figure 2.4, **c**). Also the positive-strand RNA makes a negative-strand RNA that subsequently acts as a template to produce numerous other positive RNAs for translation and replication (Figure 2.4, **d**). The mechanism of virion assembly and release is not well understood mainly because of a lack of proper experimental models. Presumably, the structural core protein and RNA genome interaction results in the RNA delivery into the nucleocapsid while with the other viral components the virus assembly takes place (Figure 2.4, **e**).[77,79–81] Finally the mature viral particles are released from the liver cell and this cycle is repeated to generate numerous virions (Figure 2.4, **f**).

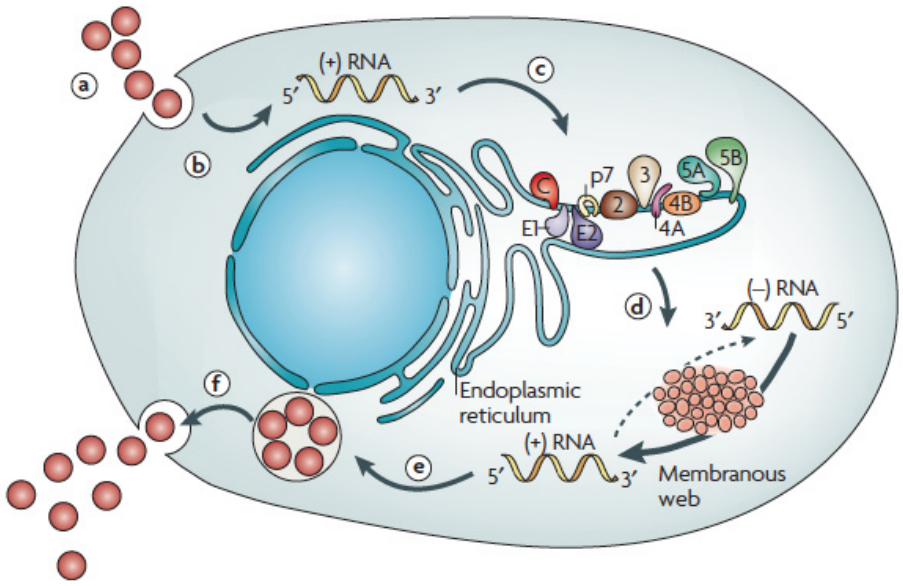


Figure 2.4: HCV life cycle. Life cycle process: (**a**): Viral particles attachment and entry; (**b**): Decapsidation and RNA release; (**c**): RNA translation and polyprotein production; (**d**): Replication of HCV RNA; (**e**): Virus particles maturation and assembly; (**f**): Virion release. (Figure from reference[82])

2.4 HCV Non-Structural Proteases

2.4.1 NS3/4A Protease

Two virally encoded proteases are responsible for processing non-structural proteins that are necessary for virus replication. NS3 protease is a 631 amino acid polypeptide known as a serine protease.[83] NS3 protease has been widely studied compared to the other HCV protease due to its multi functional behavior. The N-terminus one-third of the NS3 protein together with the NS4A protein form the active chymotrypsin-like serine protease that mediates the cleavage at the NS3/4A junction and all three downstream sites (NS4A/NS4B, NS4B/NS5A and NS5A/NS5B). Beside other functions in the viral life cycle of HCV, NS4A serves as a cofactor that stabilizes and modulates the activity of NS3.[84–87]

The catalytic triad of NS3 protease consists of His 57, Asp 81 and Ser 139 in which the NH's Ser 195 and Gly 195 in the backbone of residues serve as oxyanion holes[88] (numbering based on the sequence of NS3 protein alone). Mutation of any of the residues involved in the catalytic triad blocks the cleavage of the four sites without any effect on the processing of other HCV polyprotein positions. NS3/4A protease cleaves the junction of NS3 and NS4A through a rapid, co-translational and *cis* mechanism while the other sites are processed by this enzyme in a slower (except the NS5A/5B site) and *trans* manner.[89–93]

In addition to the proteinase domain, at the C-terminal two-thirds of NS3 protein is located the NTPase/helicase domain. Helicases are commonly responsible for DNA helix unwinding; however, because no DNA is involved in the HCV formation process, the exact role of NS3 helicase has remained controversial. Through hydrolysis of natural nucleoside triphosphates (NTPs) the required energy is provided for the helicases to presumably open the duplex HCV RNA.[94–97] NS3 helicase is essential in HCV replication, probably by contribution to the viral assembly process, since its

mutation stopped the viral replication in chimpanzee models.[98]

2.4.2 NS2/3 Protease

The NS2 protein is a 23-kDa hydrophobic transmembrane protein (amino acids 810 to 1026) which is a primary viral-translated protein of HCV.[99,100] The host enzyme cleaves the junction between p7 and NS2 which liberates the N-terminus of NS2. The NS2 protein is not capable of any enzymatic activity without the NS3 protein. In fact, the C-terminus of the NS2 protein along with one-third of the NS3 protein constitute the active NS2/3 protease (amino acids 810 to 1206).[1,82,100–102] Although the catalytic active site of NS2/3 protease is located in the NS2 region, a minimum of 180 amino acids placed in the N-terminus of NS3 is required for the NS2/3 protease activity (residues 1026 to 1206). The NS3 protein has no role in the catalytic activity of NS2/3 protease, but instead plays an important role in the proper folding of the enzyme which is essential for the efficacy of NS2/3 processing.[83,100,102–104] The very hydrophobic N-terminus region of the NS2 protein is not essential for the proteinase activity of this enzyme, and mutation studies demonstrated that this region can be truncated for better expression and purification of the protein. Therefore the minimum sequence necessary for the activity of the enzyme spans amino acids 907 to 1206 (Figure 2.5).[105]

In the earlier studies of NS2/3 protease, it was presumed that this enzyme is a zinc-dependent metalloprotease mainly because of its inhibition by zinc chelators such as EDTA.[103,104,106] However, later investigations demonstrated that this is due to a zinc-binding residue in the NS3 region that is essential for stabilizing the conformation and proper folding of the NS3 protein. Consequently, since proper folding of NS3 is essential for NS2/3 protease activity, zinc is also essential for proper NS2/3 protease function.[88,107–111] A zinc cation coordinates directly to three cysteine residues (1123, 1125, 1171) and, through a molecule of water, to a histidine residue (1175) of the NS3 protein (Figure 2.5).[93,112] The critical role of the cysteine residues was demonstrated

by removing any of the three binding cysteine residues in the NS3 region that abolished the activity of NS2 and NS3 proteases.[93] Consequently, zinc is also essential for binding to the cysteine residues. Some studies demonstrated that removing zinc from the protein and replacing other divalent coordinating metals such as cobalt or cadmium retains the activity of the protein (Figure 2.5).[93,112,113]

Although the direct role of NS2/3 protease in viral replication has not been determined, it has been demonstrated that this protein has several effects on the HCV life cycle. For example, an *in vivo* study of the HCV genome lacking the NS2 protein abolishes virus replication in chimpanzee models.[98] Similar to all non-structural proteins, the NS2 protein is involved in the assembly and release of viral particles although it is proposed that its enzymatic activity is not essential for this purpose.[114] This could be a consequence of several interactions between the NS2 protein and other structural and non-structural proteins and subsequently the roles that each of these proteins complexes have in the HCV life cycle (NS2:p7-E1-E2 complex,[115,116] NS2:NS3,[117] NS2:NS4A[118]). For instance, the liberated N-terminus of the NS3 protein has a direct effect on HCV replication through production of the replicase; therefore, inhibition of the junction cleavage between NS2 and NS3 abolishes virus replication.[114]

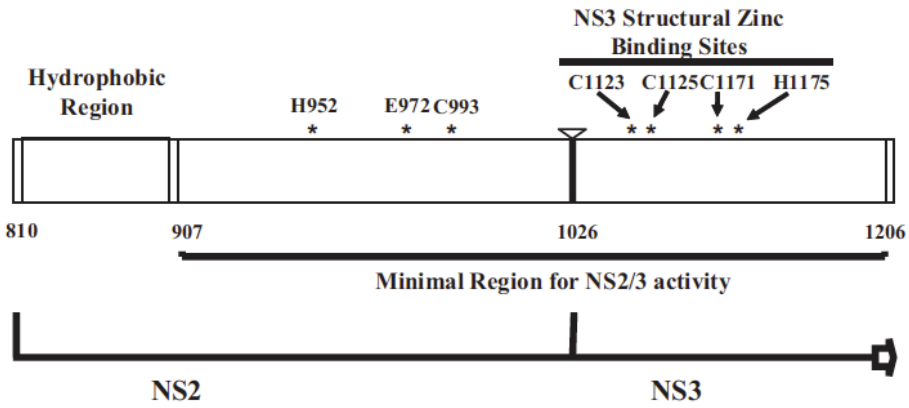


Figure 2.5: Required domain for catalytic activity of NS2/3 protease. Minimum sequence for NS2/3 processing begins from the C-terminus of NS2 (aa 907) to the N-terminus of the NS3 (aa 1206); H952, E972 and C993 form the NS2/3 protease catalytic triad; Three cysteine residues (C1123, C1125, C1171) and one histidine residue (H1175) form the NS3 structural zinc binding sites. (Figure from reference[83])

Based on the numbering of the full length of the HCV polyprotein, a combination of the three residues Cys 993, His 952 and Glu 972 creates the catalytic triad of NS2/3 protease, which is discussed in the next section.

2.4.3 Catalytic Triad of NS2/3 Protease

NS2/3 protease has been proposed as a cysteine protease. The function of NS2/3 protease was not completely clear prior to solving its crystal structure.[111] In several studies processing of NS2/3 cleavage was suggested to be through a unimolecular *cis* mechanism which is mediated by a second HCV-encoded protease as a viral enzyme or an unknown host.[100] In 2006 the crystal structure of the catalytic domain of the NS2 protein exposed a highly unusual active site where the three amino acid residues participating in the activity of this enzyme are not located on a single monomer of NS2. Rather, NS2 forms a dimer (Figure 2.6, **a**) with the active site Cys 184 on one monomer and His 143 and Glu 163 on the other (residue numbering starts from the C-terminus of the NS2 protein as opposed to the previous section that was based on the whole polyprotein) (Figure 2.6, **b**). Interaction of the N-terminus of one monomer with the C-terminus of the other monomer and vice versa constitutes the dimeric

active site. The auto-cleavage takes place when these two monomers are at the proper distance and geometry. This leads to the cleavage of the amide bond in one monomer by the active cysteine which is located in the same monomer.

Mutation studies on the NS2/3 protein indicate that substitution of residues His 143, Glu 163 and Cys 184 with alanine eliminates the cleavage at the NS2/3 cleavage site. Other mutations do not have any effect on the enzymatic reaction of NS2/3 protease.[100,105,111] In addition to the catalytic triad residues, several other significant amino acids in the NS2 protein affect the assembly and release of the infectious virus and as a result the overall replication of the virus.[114] For example, it has been identified that solvent-exposed Ser 168 is necessary for virus production as its mutation to Gly or Ala abolishes or reduces the production of infectious virus.[119] The other notable residue located in the C-terminus of the NS2 protease is Leu 217 which affects the virus production through coordination to the side chains of His 143, Cys 184 and the nitrogen of Cys 184. Both Ser 168 and Leu 217 are solvent-exposed residues and it is presumed that their mutation could disrupt protein-protein interaction(s) essential in the viral particles maturation when they are in the process of assembly.[119] Furthermore, the dimer is stabilized by Pro 164 in the *cis* conformation which bends the backbone of Glu 163 in the catalytic domain of the protein in order to form the required geometry of that domain. Recognition of such residues helps in the fundamental understanding of this significant protein as well as in designing potential inhibitors.

As opposed to previous studies, it was determined that since the protease consists of two monomer units, the NS2/3 cleavage is sensitive to the concentration of the respective monomers.[111] Discovering that dimerization is needed for the NS2/3 cleavage provides an explanation for the fact that having a certain concentration of NS2 is necessary for this cleavage. A low concentration of NS2, which would delay the

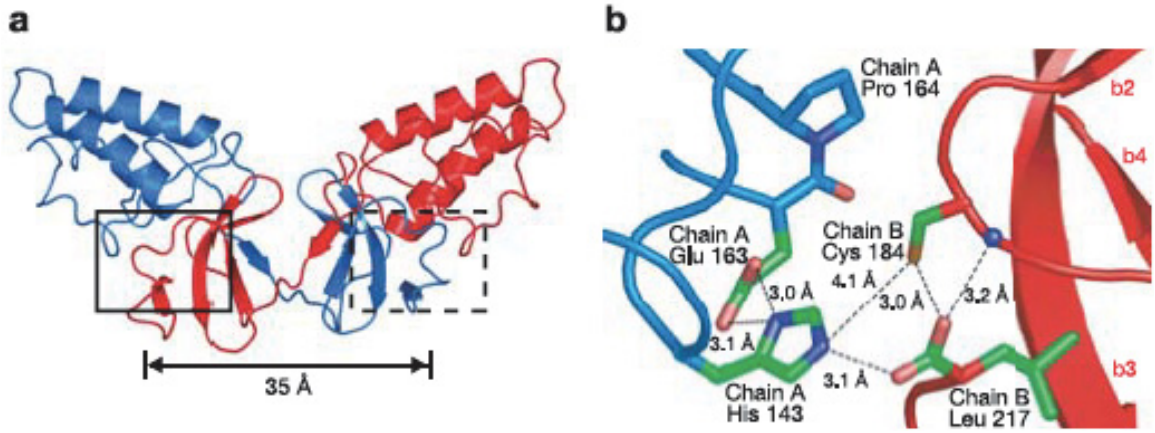


Figure 2.6: Catalytic domain of NS2 protease active site. (**a**): Dimer of the NS2 protease; (**b**): Catalytic triad of NS2 protease. (Figure from reference [111])

dimerization of NS2, may postpone the N-terminal liberation of NS3, which is required for viral replication.[70,120]

2.4.4 Mechanism of NS2/3 Proteolysis

The proposed mechanism of the hydrolysis of the viral polyprotein catalyzed by cysteine protease is shown in Figure 2.7. Initially, polarization of the thiol group in Cys 184 by His 143, which itself is activated by Glu 163, takes place (**1**). Accordingly, the nucleophilic thiolate attacks the carbonyl group of the amide and forms the first tetrahedral intermediate (**2**). It is postulated that the backbone nitrogen of Cys 184 interacts with the backbone carboxylic acid of Leu 217, which may act as an oxyanion hole to facilitate the hydrolysis.[111] Proton transfer from the acidic imidazolium ion to the NH of the leaving group forms the corresponding acyl enzyme.

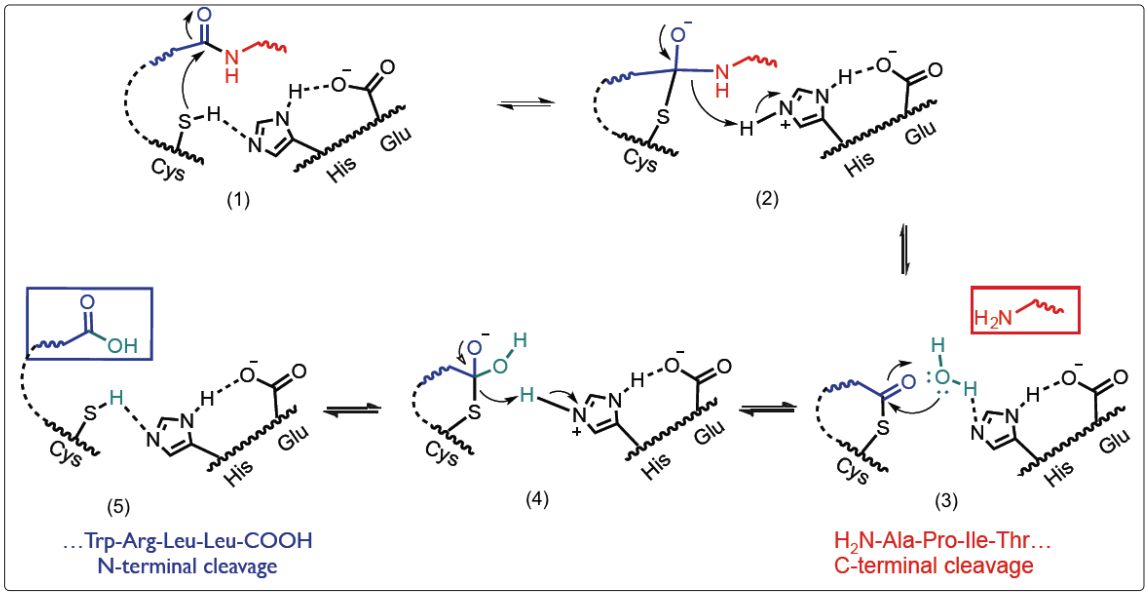


Figure 2.7: Mechanism of proteolysis of NS2/3 protease

After production of the acyl enzyme a water molecule that is polarized by His 143 through a hydrogen bond attacks the carbonyl group (**3**) and generates the second tetrahedral intermediate (**4**). This intermediate collapses and results in the formation of free acid as the N-terminal cleavage product and the catalytic triad in its initial state (**5**).[121,122]

2.5 Assay Developments and Characterization of NS2/3 Protease

The auto-cleavage character of NS2/3 protease raises several challenges for scientists in order to characterize the functionality of this protease and further detect and examine the inhibitors of this protease. Moreover, the highly hydrophobic N-terminus and membrane bound nature of the NS2 protein make the expression and purification of the protein challenging and results in the requirement of very specific assay conditions.

Some of the early studies of NS2/3 protease were focused on the optimization of assay conditions in order to increase the efficiency of the auto-cleavage reaction. It has been illustrated that microsomal membrane components are essential in the activity of NS2/3 protease by providing a hydrophobic ambiance in order to aid proper folding of the protein.[105,123] In the *in vitro* studies, various detergents were examined to be substituted by the microsomal membranes by providing artificial hydrophobic environment. It was discovered that some detergents such as Triton X-100, Nikkol, Tween 20, CHAPS and *n*-dodecyl-β-D-maltoside are able to promote the auto-cleavage activity of the NS2/3 protease although with lower proficiency.[1,123]

Adding up to 50% glycerol has also been shown to assist the detergent in initiating the cleavage, presumably by inducing proper folding of the protein.[1] Moreover, the effect of temperature was examined in separate studies and temperatures between 20 - 23 ℃ were found to be optimal, whereas temperatures below 20 ℃ or higher than 30 ℃ were detrimental to the process of auto-cleavage.[1,123]

As explained above, the very specific assay conditions require high glycerol and detergent concentrations and these additives hinder the utilization of several analysis techniques such as mass spectrometry, circular dichroism (CD), UV spectroscopy, NMR spectroscopy that are available for the characterization of many other proteins. Nevertheless, through assay optimizations as well as protein mutations and/or modifications, techniques are available for studying this protein. For example, by removing the hydrophobic sequence of NS2 at the N-terminus and introducing a solubilizing agent (ASKKKK) at the C-terminus, fluorescent and mass spectrometry characterization of NS2/3 protease has been accomplished.[102,124] However, the drawback is the high dependency of each technique on the protein's construct and buffer conditions.

Further characterization of NS2/3 protease has been done through both *in silico* and *in vitro* experiments. For instance, a recent computational modeling study combined with further *in vitro* studies explored the essential residues for NS2 dimerization

through alanine mutation of residues present at the dimer interface.[125] By ranking all calculations of three different *in silico* approaches (DCOMPLEX, EMPIRE, FastContact) to obtain $\Delta\Delta G$ values, a few residues were revealed to be potentially crucial for dimer formation of the NS2 protein. (Figure 2.8 demonstrates one example from FastContact v2.0).

Based on the obtained results five alanine mutated NS2 constructs (V162A, M170A, I175A, D186A, I201A, where numbering is based on the NS2 protease crystal structure, 2HD0)[111] were subjected to western blot. The results indicate that mutation of these residues decreases the formation of NS2 dimer (Figure 2.8). Quantification of the western blot reveals that the NS2 monomer to dimer ratio increases for M170A, I175A, I201A, D186A and V162A to 4.0, 3.2, 3.0, 2.8 and 1.5, respectively, when compared to the wild type (wt) NS2 protein. Also, the evaluated effect of mutants in the HCV life cycle using site-directed mutant HCV constructs (pJFH1-Rluc2A) illustrates that two mutated residues with a high monomer to dimer ratio (M170A, I201A) decrease the HCV genome replication 100 fold as opposed to I175A and D186A which only reduces the RNA replication 10 fold.[125] Although the mechanism by which NS2 dimer production, and accordingly HCV genome replication, is decreased through these mutants is not completely clear, NS2/3 cleavage deficiency could be responsible for hampering RNA replication.

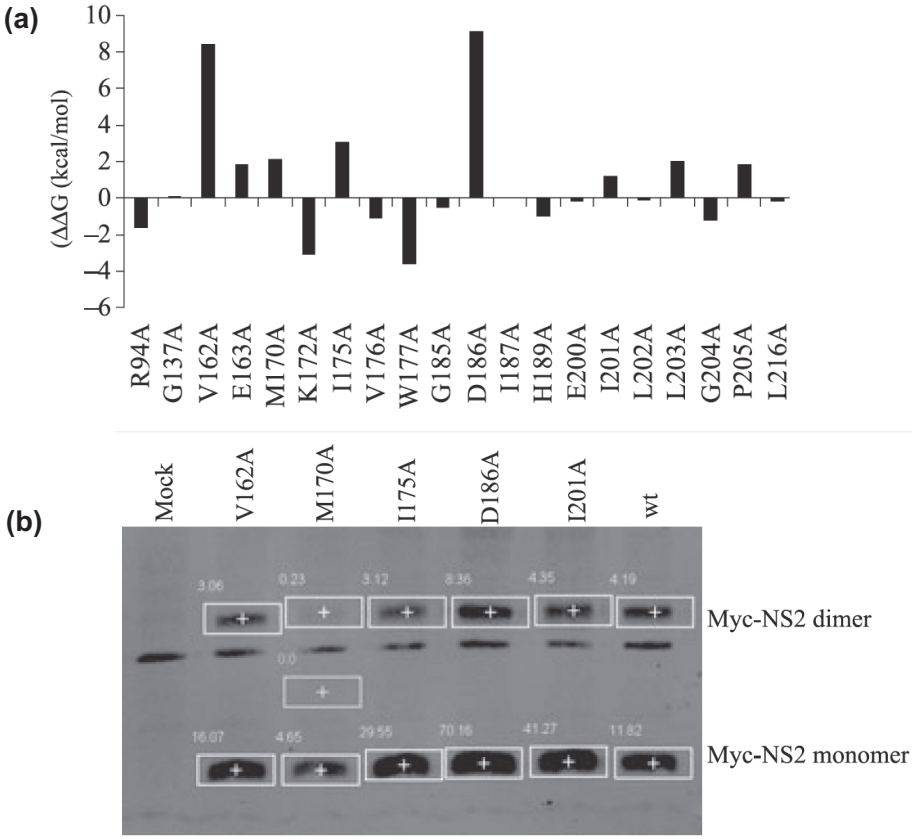


Figure 2.8: Evaluation of residues at the NS2 dimer interface. (**a**) $\Delta\Delta G$ evaluation of residues at the NS2 dimer interface;[125] (**b**) Reduction in the formation of NS2 dimer by alanine mutation[125] (**a**) The calculation was performed using FastContact v2.0 where $\Delta\Delta G = \Delta\Delta G_{wt} - \Delta\Delta G_{Mut}$; The protein is labeled with Myc-tag. Monomer to dimer ratio increase for M170A, I175A, I201A, D186A, V162A: 4.0, 3.2, 3.0, 2.8, 3.0, 1.5 respectively compared to the wildtype(wt)

2.6 General Approaches and Considerations of Synthesizing HCV Protease Inhibitors

Different stages of the HCV life cycle are potential targets for the development of drugs against this infectious virus. For example, the early stage of HCV entry into the cell is directed by host cell factors such as CD81, SR-B1, claudin 1 and occludin, and these have been targeted and several inhibitors have been developed. Two examples

are ITX-506, which inhibits the interaction of HCV E2 glycoprotein and SR-B1,[126] and, for previously untreated patients, Alisporivir (Debio-025) used in combination with PEG-IFN-α-2a, which inhibits cyclophilin.[127,128]

In addition, many attempts have been made to develop drugs which directly target protease activity. The goal of designing direct-acting antivirals (DAAs) is to achieve sustained virological response (SVR) in patients, where infectious HCV RNA is not observed.[54,129] Because of their exclusive function and well characterized role in viral replication, some of the HCV non-structural proteins provide leads to potential drug targets. Also, the compelling need for the development of alternatives for pegylated interferon alpha and ribavirin for the treatment of HCV infections resulted in the evolution of HCV non-structural protein inhibitors such as NS3/4A protease, NS5A phosphoprotein[130,131] and NS5B polymerase inhibitors.

Various strategies have been employed for the development of protease inhibitors. Commonly, inhibitor discovery starts with the identification of a hit compound. This can be found by screening natural products or peptide analogs of the natural substrate to identify compounds that could have some, often non-optimal, potency towards inactivation of the target. Later, various structural modifications are applied in order to improve the ADMET properties as well as to increase the selectivity and potency of the compounds towards a given target. Structural modification of the lead compounds is less complicated by acquiring information of the protease's active site. A number of techniques are available to help a medicinal chemist better understand the mechanism of the enzyme activity in order to discover an inhibitor, including NMR studies, X-ray crystallography and molecular modeling. These techniques not only aid in understanding the overall enzymatic activity itself, but also they provide insight to the enzyme-inhibitor interactions.[108,109,132–134]

One of the most essential modifications involves altering any peptidic nature of the lead molecules to peptidomimetics or non-peptidic small molecules. This is due

to the poor drug-like properties of most peptides. Modifications are applied to the molecules by replacing/altering some groups in the structure of the lead molecule to generate new compounds with, hopefully, greater potency, often by taking advantage of hydrogen bonding or hydrophobic interactions. Another practical modification is the introduction of bioisosteres, which refers to replacing groups with the ones having similar physical properties, such as size, shape, or polarity in order to improve ADMET properties of the molecule.[135,136]

Despite all these efforts to improve the properties of the lead compounds very few of the new analogs reach clinical application. This is due to the numerous and sometimes complex parameters that need to be taken into consideration for an inhibitor to be used as a drug. Interaction of structural properties (reactivity, hydrogen bonds, pKa, molecular weight, lipophilicity,..) with the protein and environment cause the biochemical and physicochemical properties, respectively.[137] Biochemical properties include metabolism, binding, target affinity, etc. whereas physicochemical properties incorporate solubility, chemical stability and permeability. Interaction of both these properties with the living system determines the pharmacokinetics (PK; bioavailability, half-life, clearance,..) and toxicity (LD_{50}).[137]

2.6.1 Advances in Development of NS3/4A Protease Inhibitors

The NS3 protease has been extensively studied in terms of both structure and activity, and to date several drug candidates that target NS3 protease have reached clinical trials.[138–140] Since NS3/4A protease mediates three cleavages in the HCV polyprotein, many efforts have focused on designing inhibitors which interfere with the interaction of NS3 and NS4A.[141,142] The first generation of NS3/4A inhibitors were substrate-derived peptides constructed from the N-terminal cleavage products of NS4A/4B, NS4B/5A and NS5A/5B (Table 2.1).[143–147] Some of the natural N-terminus hexapeptide product of the cleaved sites acted as non-covalent competitive inhibitors since they interacted

with the same active site as the substrate and thereby decreased the ability of the substrate to bind with the active site.

Compound[a]	**Peptide**[b]	**IC**$_{50}$ (μM)[c] **K**$_i$ (μM)[d]
NS4A/4B site-derived peptides[e, 145] (IRBM)		
1	Ac-DEMEEC-OH	1[c]
2	Ac-EMEEC-OH	21[c]
3	Ac-MEEC-OH	150[c]
4	Ac-DEME-Cha-C-OH	0.35[c]
5	Ac-D*E*LI-Cha-C-OH[f]	0.015[c]
NS4B/5A site-derived peptides[e, 146] (IRBM)		
6	DCSTPC-OH	180[d]
7	SGSWLADVWDKK-NH_2	>300[d]
NS5A/5B site-derived peptides[e, 143] (BI)		
8	DDIVPC-OH	71[c]
9	Ac-DDIVPC-OH	28[c]
10	Ac-D*D*IVPC-OH[f]	4[c]

Table 2.1: Substrate based peptide inhibiotors of NS3 protease. [a]Abbreviations: IRBM: Instituto di Ricerche di Biologia Moleculare; BI: Boehringer Ingelheim; [b]Abbreviations: Ac: Acetyl; Cha: β-cyclohexyl-L-alanine. [c,d]Definition: IC_{50}: Concentration of inhibitor required to reduce the target's activity by 50%; K_i:Binding affinities of the enzyme-inhibitor; Unlike IC_{50}, K_i is independent of substrate concentration. [e]Natural N-terminus cleavage products of NS4A/4B, NS4B-5A, NS5A-5B: DEMEEC, DCSTPC, DDIVPC respectively; [f] Italic letters indicate D-amino acids.

Because of their poor pharmacokinetic profile, peptides are not desirable as drugs. As a result, structural modifications were made to transform the peptides to peptidomimetics or small, non-peptidic molecules to improve their ADMET properties and possibly increase the potency and selectivity of the inhibitors towards the target at the same time. This approach demands structure-activity relationship (SAR) and structure-property relationship (SPR) studies that evaluate the effects of structure alterations on the compound's activity and properties, respectively.

Macrocyclic compounds and α-ketoamides, as linear inhibitors, are two types of compounds that have been developed as NS3/4A inhibitors. Ciluprevir[27,148](BILN 2061, Figure 2.9, **a**), one of the early examples of NS3/4A protease inhibitors, is a substrate-based macrocyclic compound that interacts in a non-covalent and reversible

manner with the substrate.[27] Structural modifications of hexapeptide **8** in Table 2.1 led to the development of this compound.

Since BILN 2061 showed signs of potential cardiotoxicity, additional modifications were made to its structure to generate improved inhibitors. Examples include improved macrocyclic inhibitors such as Danoprevir[149] (ITMN-191, Figure 2.9, **b**) and Vaniprevir[150] (MK-7009, Figure 2.9, **c**) or linear α-ketoamides like Telaprevir[56] (VX-950, Figure 2.9, **d**) and Boceprevir[21,22] (SCH-503034, Figure 1.3, **b**). The last two compounds are now in the marketplace for the treatment of genotype 1 HCV infections, still in combination with pegylated interferon-α and ribavirin.[151]

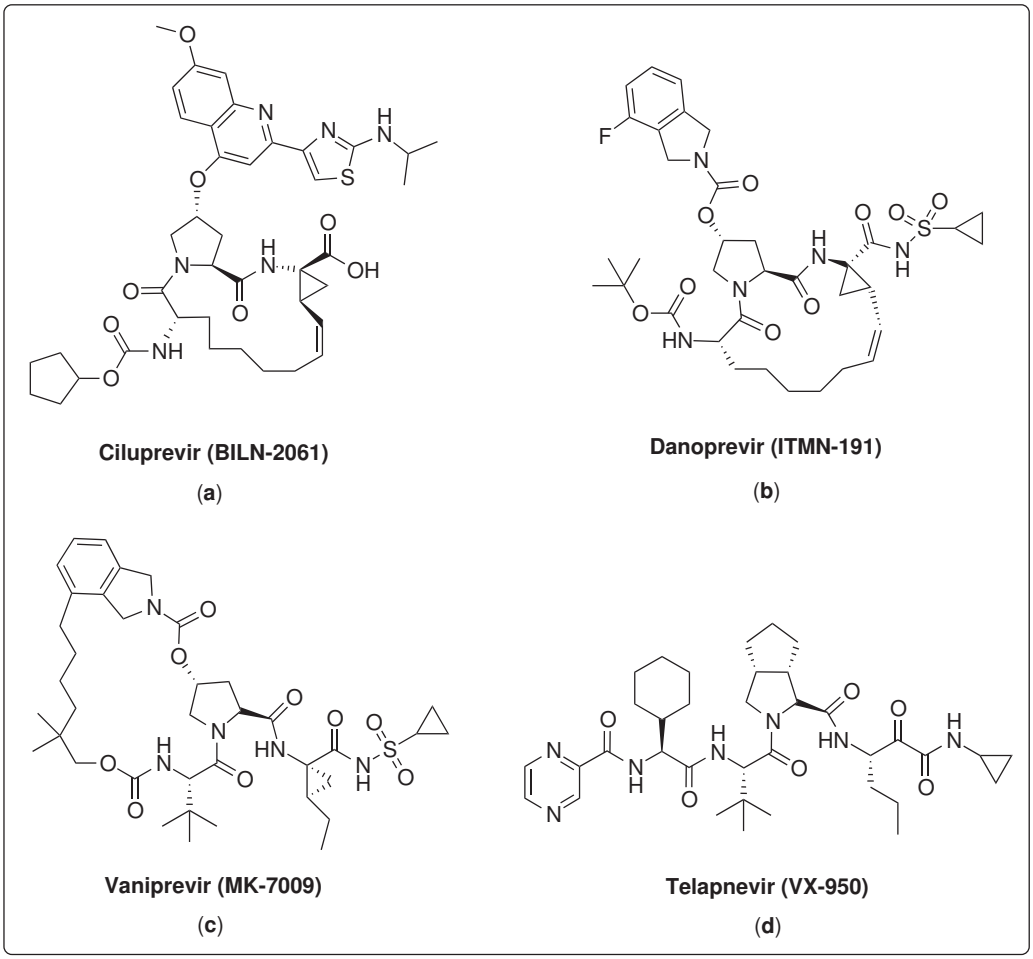


Figure 2.9: Examples of small-molecule inhibitors of NS3/4A protease

Despite all these therapeutic advances, the development of drug-resistant variants of the virus to one or more of the existing drugs is another major issue of concern. Thus, a more effective approach for the battle against HCV will rely on combination therapy, where patients are subjected to more than one drug simultaneously.

2.6.2 Advances in Development of NS2/3 Protease Inhibitors

To date no small-molecule inhibitors of NS2/3 protease have been reported despite the fact that it has been confirmed to be essential for viral replication.[100,152] In the often used sequence nomenclature[153] ..-P3-P2-P1-P1'-P2'-P3'-.., the bond between P1 and P1' is defined as the cleavage site between amino acid 1026 and 1027 by cysteine protease. Initial studies on the function of NS2/3 protease and the inhibitory activity of peptides on NS2/3 cleavage indicate that the enzyme has a potential to react easily with specific peptides.[1,152] These peptides encompass the central sequence of NS4A that is responsible for binding to NS3.[93,154] Cleavage of NS2/3 is influenced by the NS4A co-factor. This effect is related to the interaction of 12 amino acids of NS4A with the N-terminus of NS3.[88,155,156] As a result, the protein is appropriately folded and stabilized. Since NS3 is binding to NS2, logically NS4A has the same effect on NS2 as well.[31] The results by Darke *et al.* in Table 2.2 show that this 12 amino acids peptide of the NS4A site demonstrates inhibitory activity on NS2/3 protease.[152] A random combination of these 12 amino acids or peptides with less than 12 amino acid did not illustrate any effect on the processing of NS2/3 (compounds **13** and **14** respectively).

At the time, this inhibition of NS2/3 by NS4A confirmed the hypothesis of the temporal order of NS protein cleavages in which NS2/3 cleavage takes place before processing of NS3 to release of NS4A; otherwise, no NS2/3 cleavage occurs.[152]

In addition to NS4A peptides, the inhibitory ability of site-derived NS2/3 peptides

Compound	**Peptide**[a]	**IC**$_{50}$ (μM) or % **inhibition***
NS4A site-derived peptides[b]		
11	KGSVVIVGRIILSGRK	5.7
12	Ac-RGGSVVIVGRIILSGRK	3.4
13	VRLGSISVIGIVRGKK	-17*
14	Ac-RIILSGRK	-21*
15	Ac-KGSVVIV-NH_2	8*

Table 2.2: Inhibition of NS2/3 protease by NS4A site-derived peptides
[a]Abbreviation: Ac: Acetyl. [b]Natural 12 amino acids of NS4A influencing NS3 binding: VVIVGRIILSGR; Compound **11** includes residues 21-34 of NS4A in addition to two lysine residues to increase the solubility; In case of % inhibition all peptides were used in a final concentration of 50 μM.

were examined and several NS2/3 cleavage site derived peptides demonstrated no effect on the NS2/3 processing (Table 2.3).[102,152] The peptides spanned the sequence of P to P' (compounds 16-20) and P or P' only (compounds **21**-**23**). Since the results were accomplished prior to the determination of the crystal structure of NS2 protease, common opinion hypothesized an intra-molecular cleavage of the NS2/3 protease since the competing substrates did not have any effect upon the reaction rate.[102,152]

Compound	**Peptide**[a]	**[C]**[b]**(mM)**	% **inhibition**
16	DSFGEQGWRRLL*APITAYSQQTR	0.1	<5
17	EQGWRRLL*APITAYS	0.1	<5
18	GWRRLL*APITA	0.1	<5
19	EQGWRLL*APITAYS	0.62	15
20	GWRLL*APITA	0.92	20
21	EQGWRLL	1.1	14
22	APITAYS	1.3	-9
23	GRGLRLL	1.2	2

Table 2.3: Inhibition of NS2/3 protease by NS2/3 site-derived peptides (series 1)
[a]Abbreviation: Peptide sequences are the natural sequence of NS2/3 around the cleavage site; Asterisks indicate the cleavage site. [b]Abbreviation: [C] is the final concentration of the peptide in the reaction assay.

Despite these results, other *in vitro* studies for the identification of site-derived cleavage products as inhibitors of NS2/3 protease show that a decapeptide from the N-terminal cleavage with the sequence of SFEGQGWRLL inhibits the auto-cleavage reaction of NS2/3 with an IC_{50} value of 90 μM (Table 2.4).[1]

Compound	Peptide[a]	IC_{50} (μM)
24	SFEGQGWRLL*APITAYSQQT	270
25	KGWRLL*APITAY	630
26	SFEGQGWRLL	90
27	APITAYSQQT	> 1000
28	APITAY	> 1000

Table 2.4: Inhibition of NS2/3 protease by NS2/3 site-derived peptides (series 2) [a]Abbreviation: Peptide sequences are the natural sequence of NS2/3 around the cleavage site; Asterisks indicate the cleavage site.

Peptide **26** has been characterized as the most potent NS2/3 substrate-based inhibitor. These results raised another hypothesis that even though the mechanism of NS2/3 protease cleavage is known to be an intramolecular reaction, there is potential for developing inhibitors based on the NS2/3 substrate. Although the mechanism of NS2/3 enzyme inhibition by these inhibitors has not been studied, it is believed that they are reversible competitive inhibitors. The rationale is that these types of inhibitors resemble the substrate in terms of shape and chemical structure, and therefore compete with the substrate to interact with the same active site. Reversible competitive inhibitors mimic the features of the substrate; however, the interaction is not strong enough to sustain the inhibitor in the active site permanently.

2.7 Aims

The overall goals of this research are:

1- Characterization and assay optimization of NS2/3 protease cleavage through mass spectrometry and western blot techniques and subsequently optimization of these assays.

2- Synthesis of the natural substrate of NS2/3 protease encompassing the N-terminal cleavage product of the enzyme.

3- Synthesis, purification, and characterization, including evaluation of binding efficiencies, of analogs of the natural substrate to carry out rational studies for the identification of important hydrogen bonds in the binding interaction between the

inhibitor and the protease.

Chapter 3

HCV NS2/3 Protease: Results and Discussion

3.1 NS2/3 Protease Characterization Through Mass Spectrometry

Mass spectrometry is one of the most well-known and remarkable tools for the identification and quantification of proteins. Mass spectrometric analysis is applied for both qualitative and quantitative purposes. Through this approach the detection and relative or absolute quantification of modified proteins, such as processed proteins and posttranslationally modified proteins, without employing an antibody are possible.[157] The accuracy and sensitivity of mass spectrometry techniques explain their high acceptance compared to many immunological techniques.[157]

In spite of the broad advantages and apparent simplicity of mass spectrometry techniques, few studies have been devoted to their implementation for the identification and quantification of HCV NS2/3 protease. In one study by Orsatti *et al.*, a quantitative analysis of the NS2/3 protease through electrospray ionization (ESI) was accomplished, and the obtained mass of 32869 daltons was in good agreement with the theoretical

mass of 32871 daltons for the NS2/3 protein.[124] Also, the fragmentation pattern of the ionized peptides illustrated that during regular solubilization of the protein in a buffer containing β-mercaptoethanol, molecular mass of the protein increased by a molecule of β-mercaptoethanol.[124] This residue modification was shown to be due to the reaction of five out of nine cysteine residues in the NS2/3 protein sequence with β-mercaptoethanol.

In this research, the NS2/3 protease (904-1206) was obtained from Boehringer Ingelheim as a purified protein (the protein sequence differed from the one employed by Orsatti *et al*). Therefore, initially it was necessary to characterize and determine the functionality and enzymatic activity of the protease. The characterization and efficiency evaluation of the purified NS2/3 protease were performed by means of liquid chromatography (LC) coupled with an electrospray ionization quadrupole time-of-flight (Q-TOF) mass spectrometer. This method was applied as a facile and accurate tool for two purposes: 1) To observe the presence of the correct molecular mass of NS2/3 as the intact protein prior to activation of the protease, 2) To detect the enzymatic functionality of the protease for a specific incubation time after which the observation of the molecular masses corresponding to the cleaved products as well as that of the intact protein was expected. A detergent (n-dodecyl-β-D-maltoside) was added to each sample to obtain the enzymatic functionality of the enzyme since some detergents initiate the enzyme activity through providing a proper folding environment for the enzyme. To achieve the first objective, a 0.54 μM solution of NS2/3 protein in the cleavage buffer (containing 0.5% n-dodecyl-β-D-maltoside as detergent), quenched with formic acid, was directly injected into the LC-MS. However, in the chromatogram, no peak representing the intact protein was observed, only a peak corresponding to the detergent was identified. Considering that the presence of detergent and salt are not compatible with mass spectrometric analysis because of the production of intense ions, it was proposed that the protein's peak was suppressed by the detergent's peak.

To examine this hypothesis, the standard acetone precipitation protocol was applied in order to remove these interfering components. In this method, only the protein was precipitated from the solution, leaving the interfering compounds in acetone to facilitate their removal (Chapter 4). Two iterations of this procedure were generally sufficient for complete removal of the unwanted substances. This acetone precipitation protocol was applied to a protein sample that had not been initiated for enzymatic activity (sample had been quenched with formic acid immediately after the addition of detergent) to generate a "zero time" sample.

Following the protein precipitation and removal of the acetone, samples were dissolved in 5% acetonitrile/0.1% TFA injected onto a reversed-phase C4 column and eluted into a Q-TOF Ultima API mass spectrometer. From the sequence provided by Boehringer Ingelheim, the NS2/3 parent protein has a theoretical mass of 35979 daltons. Analysis of the NS2/3 protease sample by LC-MS provided a peak with a mass spectrum showing a deconvoluted mass of 36043 daltons. The additional 64 daltons in the experimental mass is due to acetonitrile and sodium adducts (Figure 3.1). The mass spectrum also showed a mass of 36118 daltons which did not match with the mass of any normally expected fragments. This mass would correspond to a modified form of the protein. It would therefore be instructive to further investigate the components of this peak by separation and collection of these two peaks by HPLC and performing a tryptic digestion on both proteins. This would provide the peptide fragments of the unknown protein (36118 daltons) to be compared to the peptide fragments of the main protein (36043 daltons). The differences in the amino acid sequences would provide information of any possible protein modification.

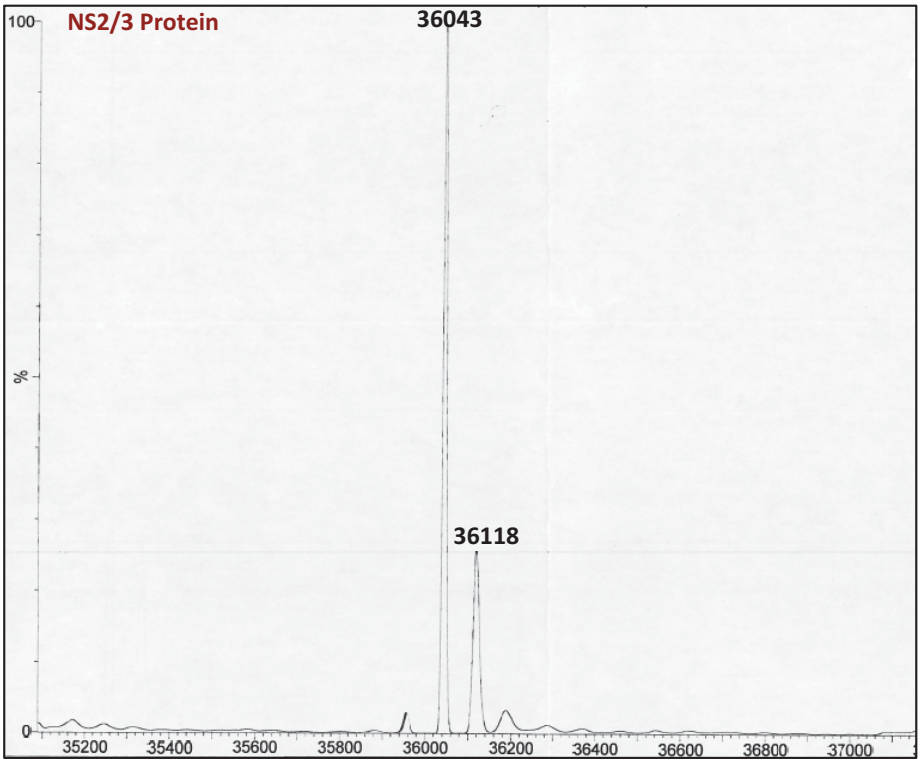


Figure 3.1: Deconvoluted mass of NS2/3 protease at zero time

In pursuit of the second objective of characterization of the enzyme's functionality, the NS2/3 protein was incubated for 4 hours in the presence of detergent, followed by the acetone precipitation protocol. Analysis of this sample by mass spectrometry provided a mass spectrum showing two masses of 15241 and 20810 daltons. The masses were compared with the theoretical masses of 15229 and 20767 daltons for NS2 and NS3 protein fragments respectively (Figure 3.2).

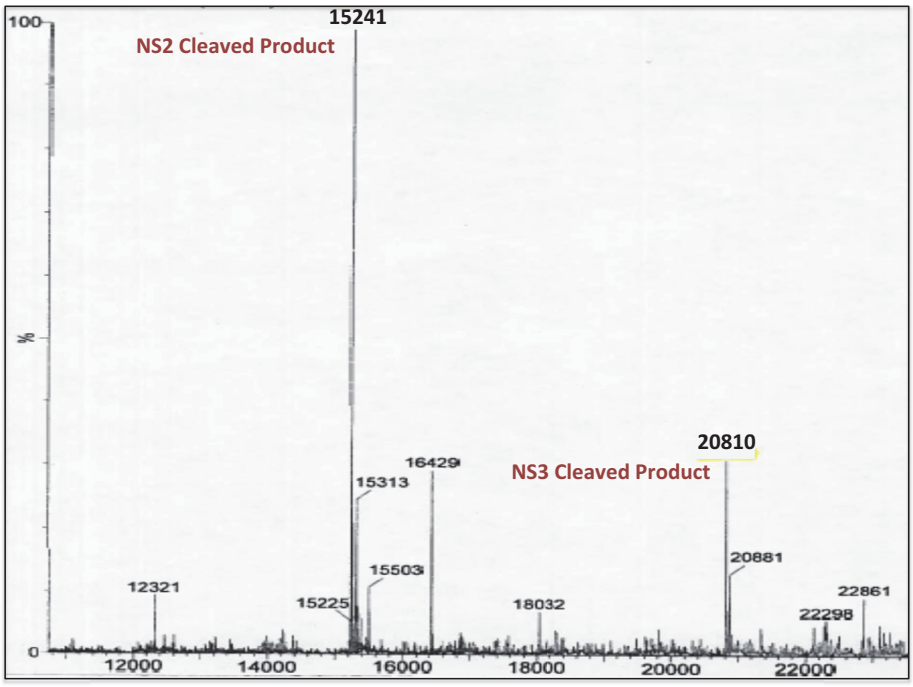


Figure 3.2: Deconvoluted masses of NS2 and NS3 cleaved products of NS2/3 protease after 4 hours

Overall, acetone precipitation was implemented as a sample preparation method to remove the interfering matrix material mainly because it was observed that the presence of detergent is detrimental to the characterization of the protein by mass spectrometry. The correct molecular masses of NS2/3 protease and the NS2 and NS3 fragments were obtained, thus validating that the protein was legitimate and functional to use. Some differences between the theoretical and experimental masses can refer to the instrument's calibration, however, the protein's sequence was validated through a second methodology employing trypsin digestion and LC-MS as will be explained in the next sections.

3.2 NS2/3 Protease Characterization Through Trypsin Digestion and LC-MS

Although acetone precipitation followed by mass spectrometry provided a means to carry out the initial characterization of the NS2/3 protein, the possibility of protein denaturation, resulting, in most instances, in difficulties with resolubilization of the protein, is a major drawback of this technique. Moreover, the quantification of the protein is not accurate without standards for measuring the response factor of the instrument.

To circumvent these issues, protein digestion, which is a key step in sample preparation prior to analysis by mass spectrometry, was employed. Trypsin protease is known for specifically cleaving peptide bonds that are followed by arginine or lysine residues in the C-terminus of a given polyprotein, except when either of these amino acids are followed by proline.[158]

Tryptic digestion followed by mass spectrometry has been used in the characterization and identification of the NS2/3 enzyme with cysteines modified by β-mercaptoethanol, mentioned above (Section 3.1), in two studies by one group.[102,124] However, this technique has been applied in the study of HCV NS3 enzyme several times.[159,160]

Protein digestion by trypsin produces different size peptide fragments that are beneficial for the analysis of proteins.[158] Since peptides are smaller than proteins their quantification provides higher sensitivity, as well as improved separation through liquid chromatography. In addition, since small peptides can be synthesized or purchased conveniently, they can be used as standards in those experiments that demand protein quantification. Due to these advantages, we have incorporated this method prior to mass spectrometry for quantification of enzymatic activity of NS2/3 protease.

3.2.1 Assay Optimization

External Calibration Curve Development

To generate an accurate and quantitative analysis when employing any analytical technique such as HPLC or mass spectrometry, it is essential to calibrate the response of the instrument to the compounds of interest using calibration standards. An external calibration curve is one of the most widely used calibrations, because of its simplicity and applicability to various methods. Therefore, in order to develop a quantitative method for the preliminary kinetics studies of NS2/3 protease, an external calibration curve was established based on the protease's cleavage products.

NS2/3 protease cleaves the junction between amino acids 1026 -1027 (Figure 3.3, **a**) and produces two cleaved peptide fragments: NS2 and NS3. In the NS2/3 protease cleavage studies, after a certain incubation time the reaction process was stopped by treating the samples with formic acid to produce a mixture of NS2 and NS3, with a substantial amount of intact NS2/3 protein remaining as well (Figure 3.3, **b**).

Treatment of this protein cocktail (**b** and **c**) with trypsin produces several other peptide fragments from digestion of the NS2, NS3 and NS2/3 polyproteins. In the NS2/3 parent protein, trypsin cleaves the peptide bonds right after arginine residues at two different sites (Figure 3.3, **b**) and produces a 13 amino acid peptide referred to as LLAPI (Figure 3.3, **d**). Thus, this peptide is the tryptic digestion fragment from unprocessed parent NS2/3 protease. Trypsin also cleaves the amide bond immediately after the arginine residue in the NS3 protein produced from cleavage of the NS2/3 enzyme to give an 11 amino acid peptide referred to as API (Figure 3.3, **e**). These two tryptic digestion peptides represent the cleaved (NS2, NS3) and un-cleaved proteins (NS2/3 protease). It is noteworthy to mention that the NS2 protein also gives rise to a trypsin digest product that has not been demonstrated in Figure 3.3, since monitoring the API peptide resulted from the NS3 product was sufficient. Furthermore, the

tryptic digestion fragment resulted from the NS2 protein did not interfere with the product from the NS3 protein.

In order to quantify the enzymatic reaction of the NS2/3 protease, these two trypsin-digested peptide fragments were monitored at two different times: when the enzyme had not undergone any self-processing (zero time) and after a certain incubation time, when the enzyme had undergone some self-processing.

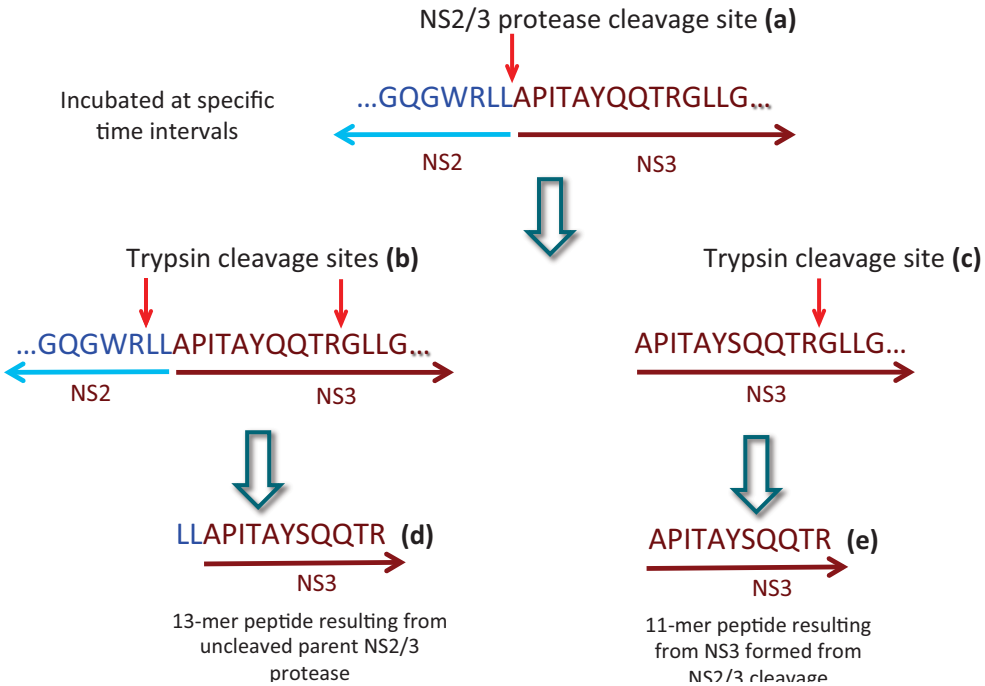


Figure 3.3: Schematic representation of NS2/3 protease cleavage and tryptic digestion products

To establish an accurate quantification method and to consider the respective response factors of the two tryptic digestion peptides, an external calibration curve was developed using synthetic standard samples of API and LLAPI (the 11 and 13 amino acid tryptic digest peptides, respectively). Hence, a series of API and LLAPI peptide mixtures with various concentrations were prepared. In these mixtures the quantity

of LLAPI peptide (represents the un-cleaved peptide or *substrate*) was kept constant (1 mg) and the quantity of API peptide (representing the cleaved peptide or *product*) varied (Table 3.1). External standards were prepared in the same buffer solution that was used for the NS2/3 protease cleavage. Standard samples were subjected to reversed-phase ultra high performance liquid chromatography coupled to mass spectrometry (UHPLC-MS/MS) with a hybrid quadrupole-time-of-flight (Q-TOF) MS instrument.

API/LLAPI [a] *(mg/mg)*	**API/LLAPI** *(area)*	**API/LLAPI** *(mmol/mmol)*	**% Cleavage**	**SD**[b]
1	0.56	0.85	45.8	2.7×10^{-02}
0.5	0.3	0.42	29.7	2.3×10^{-02}
0.2	0.13	0.17	14.5	8.9×10^{-03}
0.1	0.06	0.08	7.8	3.5×10^{-03}
0.05	0.03	0.04	4.1	1.9×10^{-03}
0.02	0.01	0.02	1.7	2.0×10^{-03}

Table 3.1: External calibration for quantification of NS2/3 protease. [a]API/LLAPI: 11-mer peptide over 13-mer peptide (representative of cleaved over un-cleaved peptide), [b]SD: Standard deviation of API/LLAPI (*area*). (Number of replicates: 3)

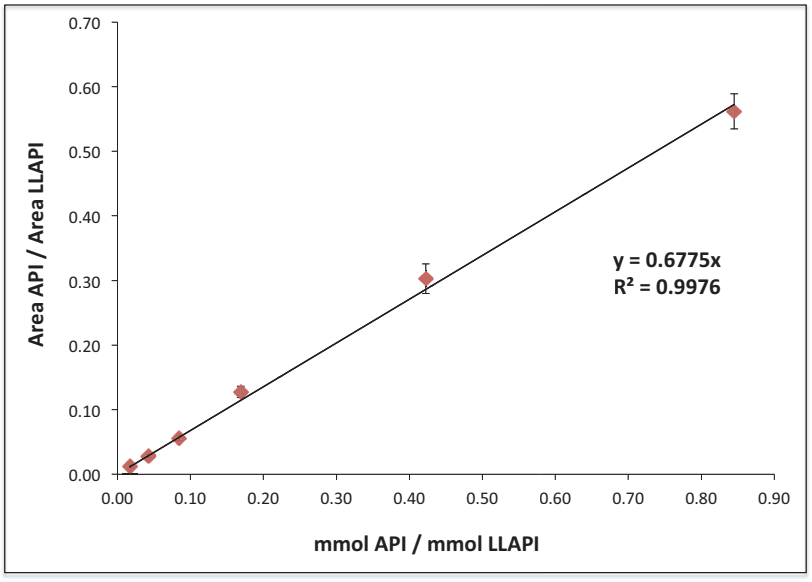


Figure 3.4: External calibration curve using API and LLAPI synthetic peptides. (Number of replicates: 3)

From the analysis of each of the API-LLAPI peptide standard mixtures by UHPLC-MS/MS, a corresponding area ratio of the peaks resulting from each component was obtained using total ion current (TIC) as the detector response. Since the molar ratios of the API-LLAPI peptides are known in these standard samples, a calibration curve could be generated. Figure 3.4 illustrates variations of the area ratio versus molar ratio of API over LLAPI peptide which resulted in a linear calibration curve for employment in the NS2/3 protease assay system. The molar ratio of API to LLAPI can also be related to the percent cleavages of the NS2/3 protease by re-expressing the quantities as $\frac{API}{(API+LLAPI)}$, representing the amount of cleaved (NS3) peptide over the initial amount of NS2/3 protease.

This calibration curve will permit samples from the NS2/3 protease assays to be easily quantified following trypsin digestion.

Solid Phase Extraction Optimization

Several parameters can have a direct effect on LC-MS analysis, of which matrix composition and analyte concentration are noteworthy. Solid-phase extraction (SPE) has been an extensively employed approach for sample preparation. With SPE, the analyte of interest is retained on a solid phase while undesired components are washed away with an appropriate mobile phase. The analyte is then released from the solid phase by changing the mobile phase. In this way, components that suppress electrospray ionization of the target sample or damage/shorten the column and MS system lifetime can be removed. Moreover, enrichment of analyte concentration is another advantage of utilizing this method.[161,162] The application of SPE in the characterization and quantification of the NS2/3 protease experiments, immediately after trypsin digestion and prior to mass spectrometry analysis, would be advantageous in several ways. These advantages included the removal of detergent and glycerol as well as concentration of the protein sample for better overall sensitivity of the protocol.

Moreover, via this technique the sample buffer could be easily exchanged to solvents with better compatibility with the LC-MS.

Depending on the analyte of interest and the composition of impurities, various stationary phases will perform differently for analyte retention and impurity separation.[163] Therefore, method optimization studies were started by examining two different types of SPE cartridges: Oasis® HLB, a hydrophilic-lipophilic reversed-phase cartridge, and Strata™-X, a polymer-based cation exchange cartridge.

The HLB cartridge retains polar analytes and the cation exchange polymer-based cartridge was expected to retain the target peptides due to the presence of a positively charged arginine residues in the peptide sequence.

The percent recovery of the representative trypsin-digested cleaved (API) and uncleaved (LLAPI) peptides was calculated using the synthetic standard peptides as controls. A solution containing a 1:1 ratio of the standard peptides (API and LLAPI) was prepared in the buffer used for the NS2/3 experiments and samples were applied on both cartridges. After removing the interfering compounds, such as detergent and salts, by washing the cartridge with water and methanol, the retained analytes were recovered from the column, employing single or mixed solvents (Chapter 4). The recovered peptides were subjected to LC-MS analysis and the percent recovery of each peptide fragment was measured. The results are shown in Figure 3.5.

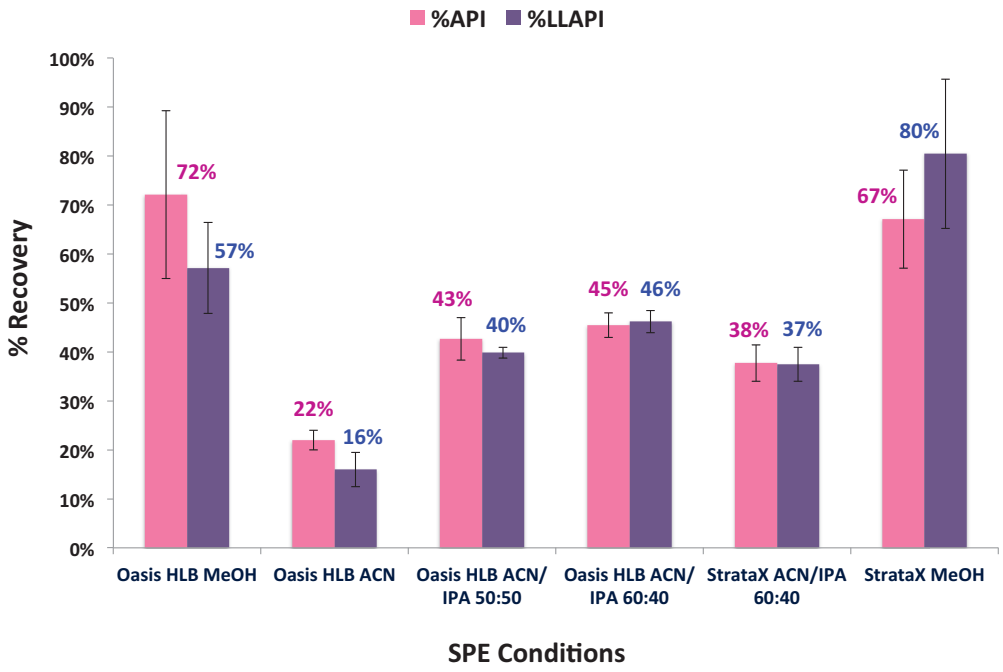


Figure 3.5: Solid-phase extraction of control synthetic peptides. API and LLAPI peptides were subjected on two different columns. Y-axis shows the percentage recovery of each peptide; X-axis represents the column and solvent conditions; HLB: Hydrophilic-Lipophilic-Balanced reversed-phase sorbent; ACN: Acetonitrile, IPA: isopropyl alcohol. (Number of replicates: 3)

Methanol, acetonitrile (ACN) and mixtures of isopropyl alcohol (IPA)/ACN were examined as the eluting solvents for peptide recovery. MS directed quantification of these recoveries demonstrated that a combination of ACN: IPA (60:40) provided an equal recovery of both peptides from either the Oasis® HLB or the Strata™-X cartridges, with slightly higher overall recovery from the Oasis® HLB cartridge. Although with this mixture of solvents elution of the peptides from the cartridge was incomplete (about half of each analyte was retained in the cartridge), this loss was inconsequential since the equivalent degree of recovery of each peptide would be sufficient for accurate quantification of the API/LLAPI ratio.

From these studies, a sample preparation method was established, based on solid-phase extraction, to remove the analytical sample incompatible mass spectrometry

interferences such as glycerol, detergent and salt present in the NS2/3 protease matrix buffer. After removing the impurities, a 60:40 ACN:IPA solvent mixture was found to elute the peptides to an equal degree from an Oasis® HLB column.

3.2.2 Time Course Studies of Auto-Cleavage of the NS2/3 Protease

Obtaining kinetic information of a protease reaction is of great interest in order to understand the mechanism of the enzyme reaction, to generate efficient enzymatic assays and further to develop inhibitors for such an enzyme.

The productivity and rate of an enzymatic reaction are highly affected by the protein's construct as well as the type and concentration of detergent in the case of membrane-bound, hydrophobic proteins. The NS2/3 protease is even more influenced by these parameters because of its intra-molecular, auto-cleavage process. Several groups have shown the time course study of the auto-cleavage reaction to estimate the overall rate of reaction.[1,102,123,152] Darke *et al.*[152] demonstrated that a 3 hour incubation of NS2/3 (810-1615, from HCV BK strain) resulted in 60% processing of the enzyme. Therefore the observed rate constant of the first order reaction (k_{obs}) was estimated as 0.04 min^{-1}. This result was comparable to the reaction rate of the identical protein construct measured by Pieroni *et al.*[123] under optimized conditions (detergent and temperature) which revealed 75% cleavage of the enzyme after a 4 hour incubation time. Two later studies showed the time course assay of this protease with different protein constructs. Pallaoro *et al.*[102] obtained a rate constant of k_{obs}= 0.05 min^{-1} for the truncated NS2/3 protease (907-1206-ASK_4) while Thibeault *et al.*[1] obtained a maximum of 50% cleavage after a 5 hour incubation time for the truncated NS2/3 protease (904-1206). The latter's enzyme construct, genotype and assay conditions are very similar to the ones of the present work. To ensure the functionality of the NS2/3 protease and further to compare the preliminary kinetics of

this protease with the available data in the literature, a time course study was carried out.

The experiment was followed at different time intervals with a 0.54 μM final concentration of protein, since this concentration was utilized previously by Thibeault *et al.*[1] To initiate the reaction, *n*-dodecyl-β-D-maltoside (DM) was added to each sample as detergent. Samples were incubated at 23 °C and the enzymatic reaction was monitored up to four hours. After each time course interval, reactions were quenched with formic acid and treated with trypsin following the trypsin digestion procedure (Chapter 4). Subsequently, samples were extracted through solid-phase extraction and then subjected to UHPLC-MS/MS.

Peptides were identified based on the MS/MS data using ProteinPilot™ software, and quantified through peak integration using MultiQuant™ software (Chapter 4). Figure 3.6 shows six analyzed chromatograms obtained from UHPLC-MS/MS analysis where each plot presents the amounts of cleaved (API or product) and un-cleaved (LLAPI or substrate) peptides versus time. In Figure 3.6 chromatogram 1 shows the analysis of the control reaction at zero incubation time, in which no cleavage happened. During the first hour of incubation, samples were collected at 30 minute intervals. Analysis of these samples showed the formation of 4 and 7% cleaved products after 30 and 60 minutes, respectively (chromatograms 2 and 3). After the first hour, data were collected at one hour time intervals until a total of 4 hours had elapsed. The amount of cleaved product slowly increased during this time, reaching 19% after 4 hours (chromatogram 6).

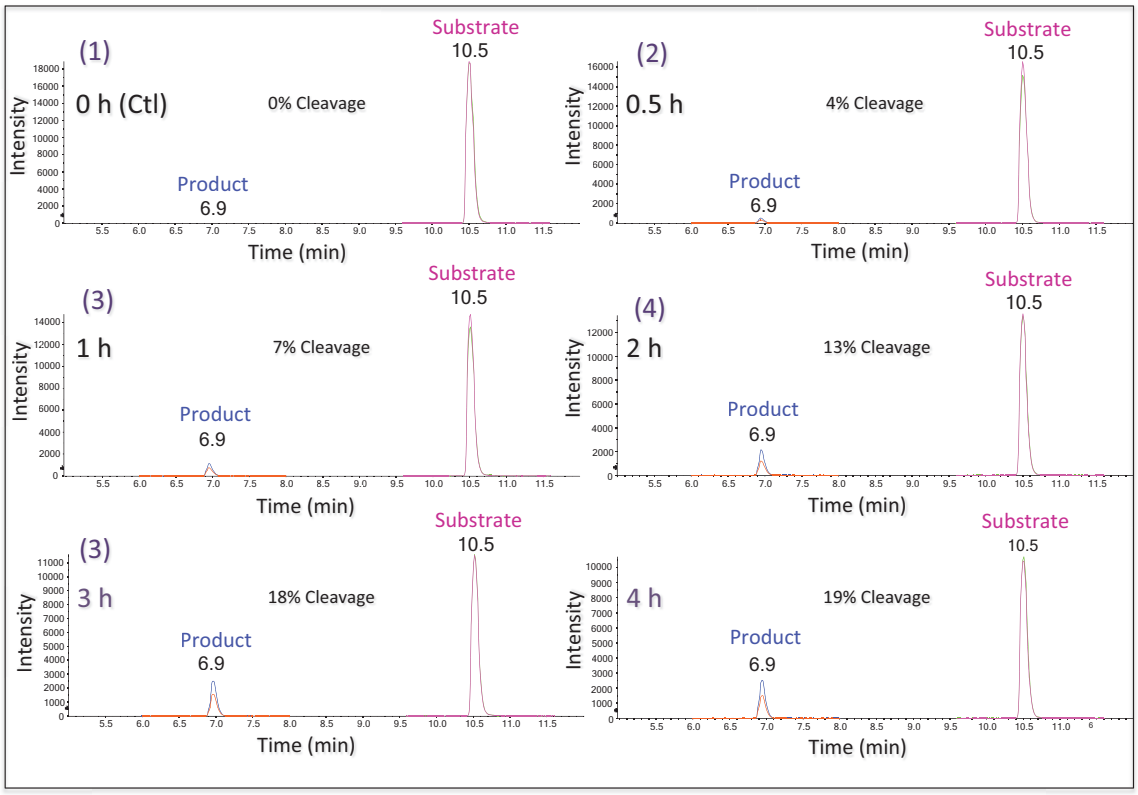


Figure 3.6: UHPLC-MS/MS profile of NS2/3 protease time course experiment. Y-axis shows the total ion current of cleaved or un-cleaved peptides. (Number of replicates: 3)

The external calibration curve (Figure 3.4) was used to convert the peak areas of the cleaved and un-cleaved peptides in each chromatogram to the molar ratios (Table 3.2). From these molar ratios, the percent cleavage of the enzyme as a function of time could be calculated (Figure 3.7).

Incubation time (*min*)	**API/LLAPI** [a] (*area*)	**API/LLAPI** (*molarity*)	**% Cleavage**	**SD**[b]
0	2.0×10^{-03}	2.9×10^{-03}	0.3	0.1
15	1.0×10^{-02}	1.5×10^{-02}	1.5	0.2
30	2.5×10^{-02}	3.7×10^{-02}	3.5	0.5
60	5.1×10^{-02}	7.5×10^{-02}	7.0	0.2
90	7.8×10^{-02}	1.1×10^{-01}	10.3	1.0
120	9.8×10^{-02}	1.4×10^{-01}	12.6	1.0
180	1.4×10^{-01}	2.1×10^{-01}	17.6	0.6
240	1.6×10^{-01}	2.4×10^{-01}	19.1	0.2

Table 3.2: NS2/3 time course cleavage data. [a]API/LLAPI: Cleaved product over un-cleaved substrate; [b]SD: Standard deviation of API/LLAPI (*area*). (Number of replicates: 3)

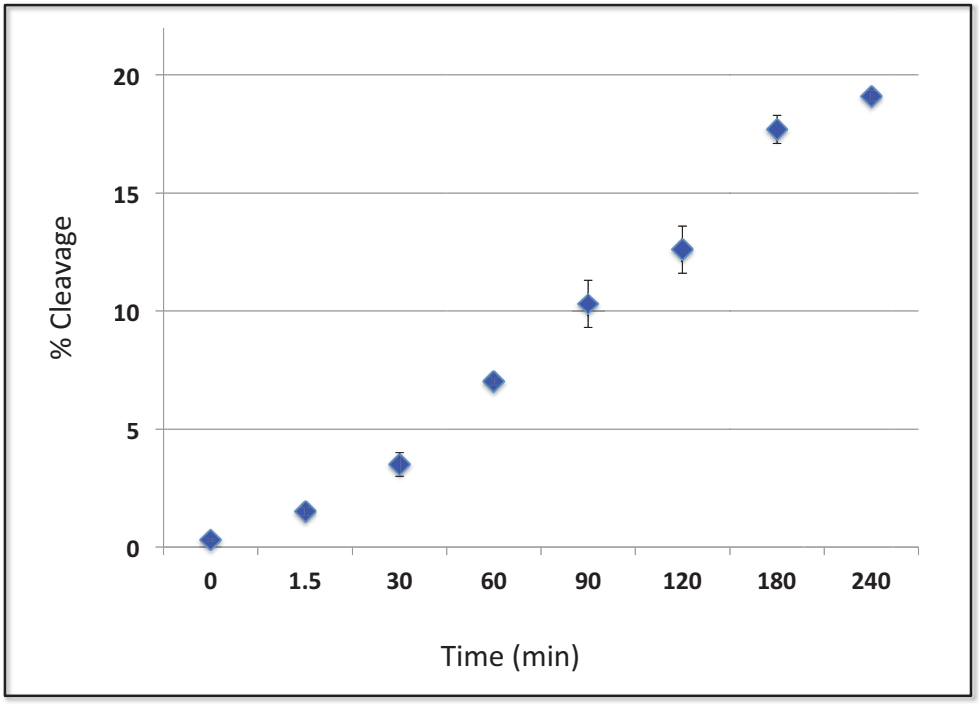


Figure 3.7: NS2/3 time course cleavage by UHPLC-MS/MS. (Number of replicates: 3)

Since the enzymatic reaction of the NS2/3 protease is considered a pseudo-first order reaction, the percentage of substrate remaining over time follows an exponential rate law described by the following equation,

$$\frac{S}{S+P} = e^{-k_{obs}t}, \tag{3.1}$$

where S represents the amount of substrate or un-cleaved NS2/3, P is the amount of cleaved product, k_{obs} is the observed rate constant, and t corresponds to the elapsed time. The observed rate constant (k_{obs}) based on the 3 replicates of the experiments was calculated as 1.0×10^{-05} s^{-1} (Figure 3.8). For a similar construct under similar assay conditions for the NS2/3 protease, Thibeault *et al.*[1] obtained 50% cleavage after 5 hours enzyme activity using the western blot technique for quantification, corresponding to a k_{obs} of 3.8×10^{-05} s^{-1}. Although the quantification techniques were different, this data demonstrates that a slower reaction process was observed in the case of our protein. Nonetheless, our data obtained from the mass spectrometry experiments is close to the value from the literature. Overall, these experiments were important as controls for establishing the baseline activity for auto-cleavage of our NS2/3 protease construct under our assay conditions in order to eventually measure the effects of the developed inhibitors.

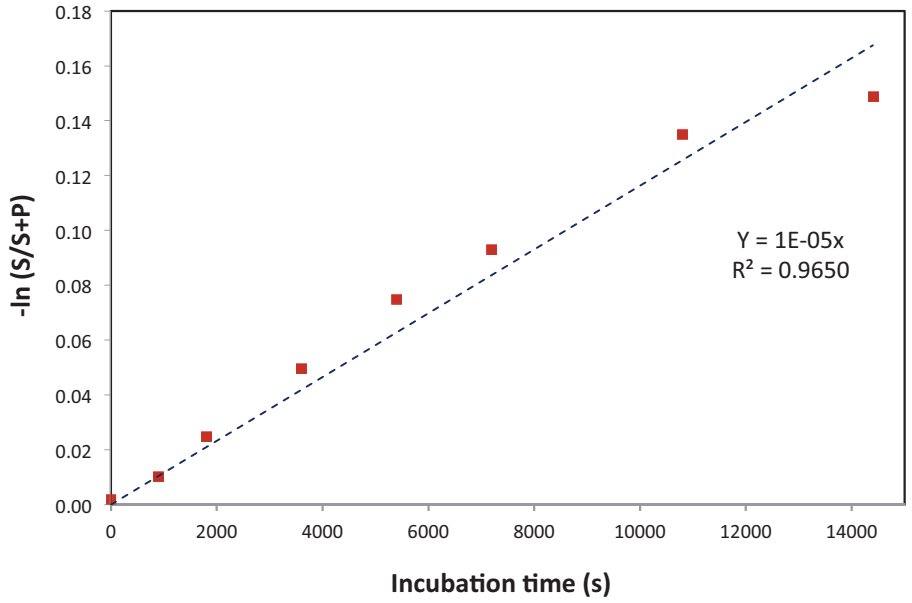


Figure 3.8: Determination of rate constant of NS2/3 processing from UHPLC-MS/MS time course data. (Number of replicates: 3)

In summary, trypsin digestion followed by tandem mass spectrometry was carried out to obtain an accurate and robust assay for quantification of the NS2/3 protease. This involved characterization of the cleavage reaction over time in order to achieve initial kinetic information of the protease reaction. Since this method was applied for quantification purpose, sample preparation and calibration curve development were performed prior to subjecting protein samples to mass spectrometric analysis. For this reason, samples were treated with trypsin (Chapter 4) followed by their solid-phase extraction to remove salt and detergent and then subjected onto reversed-phase UPLC-MS/MS using a hybrid quadrupole-time-of-flight (Q-TOF) MS equipment.

3.2.3 NS2/3 Protease Sequence Alignment

In order to validate the obtained masses of the intact and cleaved NS2/3 protease from the acetone precipitation followed by mass spectrometry, the referenced theoretical masses of NS2/3 protease from the literature (provided by Boehringer Ingelheim in a patent)[3] was compared to the peptide fragment sequences obtained from trypsin digestion and HPLC-MS/MS. The alignment of the two sequences is shown in Figure 3.9 where the highlighted regions identify the peptide sequences from the trypsin digestion and HPLC-MS/MS that are identical to the sequence according to the Boehringer Ingelheim patent. The sequence alignment provided a very good sequence coverage (97% coverage based on the Blast® software) proving that the masses obtained from the mass spectrometry after acetone precipitation reflect the actual masses of this protease and its cleavage products.

Figure 3.9: Sequence alignment of the peptide fragments from trypsin digestion and UHPLC-MS/MS compared to the literature[3]

3.2.4 NS2/3 Protease Inhibition by Classical Inhibitors

To better understand the binding sites of proteases, several types of classical protease-inhibitors have been evaluated. Some classical cysteine, cysteine/serine and metallo-protease inhibitors are able to inhibit NS2/3 protease processing *in vitro* as shown in Table 3.3.[1] These results provide important insights into the mechanism of this enzyme.

For instance, NS2/3 protease inhibition by cysteine/serine protease inhibitors such as tosyl lysine chloromethyl ketone (TLCK) and tosyl phenylalanyl chloromethyl ketone (TPCK) confirm the presence of an active site histidine, while the inhibition by alkylating agents such as iodoacetamide specifies an active site cysteine. The inability of 1,7-phenanthroline to cause inhibition, taken together with the inhibitory activity of 1,10-phenanthroline, indicate that the enzyme is inhibited through a chelation mechanism by this metalloprotease inhibitor (Table 3.3).[1]

Compound ** Target Protease*	**Inhibitor Concentration**	**Inhibition of NS2/3**
** Cysteine protease*		
N-Ethylmaleimide	0.1 mM	100% inhibition
Iodoacetamide	1 mM	100% inhibition
E64	0.2 mg/mL	No inhibition
** Serine protease*		
Aprotinin	1 mg/mL	No inhibition
Pefabloc	1 mg/mL	No inhibition
** Cysteine/Serine protease*		
TLCK	0.5 mM	100% inhibition
TPCK	0.5 mM	100% inhibition
Leupeptin	0.1 mg/mL	No inhibition
** Metalloprotease*		
EDTA	2 mM	100% inhibition
1,10-Phenanthroline	1 mM	80% inhibition
1,7-Phenanthroline	1 mM	No inhibition
** Aspartic acid protease*		
Pepstatin	0.01 mg/mL	No inhibition

Table 3.3: Effect of classical protease inhibitors on NS2/3 protease inhibition.

Table information was adopted from Thibeault *et al*;[1] 0.8 μM of the NS2/3 protease was used in the assays; TLCK: Tosyl Lysine Chloromethyl Ketone; TPCK: Tosyl Phenylalanyl Chloromethyl Ketone

In addition to the analysis of the enzyme's activity, we have examined one classical NS2/3 enzyme inhibitor in order to evaluate the optimized sample preparation protocol and mass spectrometry assay prior to measuring the biological activity of the substrate-based inhibitors. Iodoacetamide was selected for evaluation since it is known

as potent inhibitor of cysteine proteases.[1] A dose-response curve was obtained using eight different concentrations of iodoacetamide (Figure 3.10) employing an NS2/3 protease concentration of 0.54 μM and an incubation time of 1 hour.

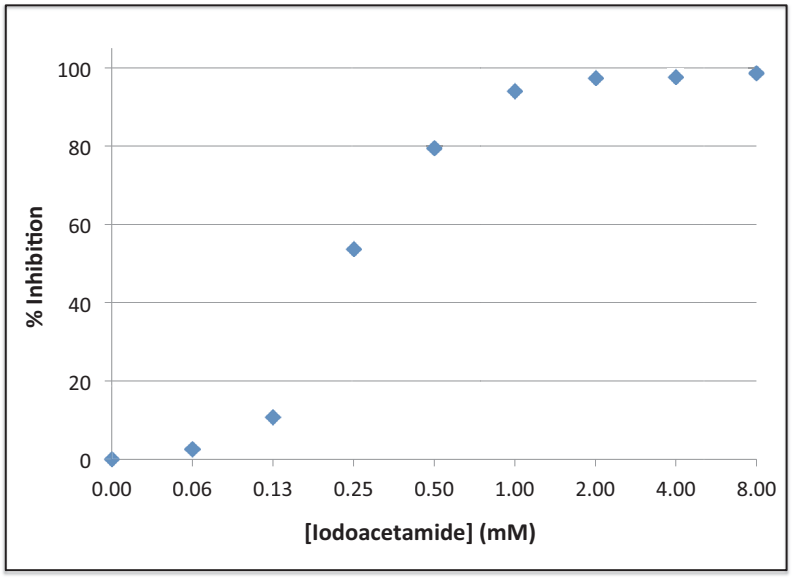


Figure 3.10: Dose-response curve of NS2/3 inhibition by iodoacetamide

The percent cleavage of the NS2/3 was quantified by mass spectrometry, as previously described at each concentration of inhibitor. Control samples without inhibitors were quantified by the same method. Subsequently, the percent inhibition was obtained based on the following equation,

$$\frac{\%C_{\text{ctl}} - \%C_{\text{subt}}}{\%C_{\text{ctl}}} \times 100, \tag{3.2}$$

where $\%C_{\text{ctl}}$ is the percent cleavage of the NS2/3 observed in the control samples without inhibitor and $\%C_{\text{subt}}$ is the percent cleavage of the NS2/3 at a particular concentration of inhibitor. The percent inhibition was plotted as a function of the inhibitor concentration, generating the typical sigmoidal dose-response curve shown in Figure 3.10. The result is comparable with the literature values of the percent

inhibition reported in Table 3.3 for iodoacetamide (95% vs. 100% inhibition at 1 mM inhibitor concentration respectively).[1]

Overall, the purpose of accomplishing this experiment was to attain a dose-response inhibition curve as a proof of concept through our optimized mass spectrometry method, in which the enzymatic activity of the protease can be followed.

3.3 NS2/3 Protease Characterization Through Immunoblotting

Although trypsin digestion followed by mass spectrometry usually provides a very efficient and accurate analysis of proteins, it was essential to analyze and detect the protein by a second method to cross-validate the two approaches. Moreover, having a second technique available provides options, and one of the two techniques may prove to be better suited for the evaluation of substrate-based inhibitors. Towards this aim, western blotting (immunoblotting) as a major and practical technique for the identification and quantification of the desired protein from a complex mixture was selected. Most of the NS2/3 protease quantification and mechanistic studies have utilized western blot strategies. Particularly, in the early characterization, mutation and cleavage studies of this protein several groups took advantage of this technique.[99,100,103,123,164] Because of the complexity of the NS2/3 protease auto-cleavage mechanism, in which the substrate is part of the enzyme, the application of several protein identification/characterization strategies is challenging. In the western blotting technique proteins separated by gel electrophoresis are transfered from the sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) to the binding supports such as nitrocellulose or polyvinylidene fluoride (PVDF) membranes. This technique is particularly practical for the NS2/3 protease, since the antibody for the detection of the NS2 or NS3 cleaved fragments can be applied. Hence, the individual protein is

initially detected by this antibody (primary antibody), and thereupon a secondary antibody detects the primary antibody for further visualization of the protein.

3.3.1 Determination of the Optimal NS2/3 Protease Concentration

One parameter to be optimized in the NS2/3 cleavage western blot assay was the acceptable concentration range of protein, in which both the substrate and the product could be quantified in a reliable manner. One general rule for obtaining such a reliable quantification method is to avoid the production of over-saturated immunoreactive bands. This is particularly essential in inhibition assays since saturated bands can introduce errors to the quantification assays and therefore alter the inhibitor's efficiency results by orders of magnitude. Producing quantifiable NS2/3 protein signals is even more challenging since the substrate is part of the enzyme. This means that starting with a high concentration of the enzyme can result in saturated substrate bands but un-saturated, quantifiable product bands at certain incubation times. On the other hand, starting with a low enzyme concentration can produce un-saturated, quantifiable bands of the substrate but weak signals of the product that can not be quantified. Both cases are not reliable for quantification purposes.

Despite such a wide application of western blotting on the study of NS2/3 protease activity, few have considered the effect of the protein concentration on their quantification techniques and often the assays were carried out within a specific range of protease concentrations without specifying the reason of selecting such a concentration.[1,102,152] To determine the suitable concentration for the NS2/3 protease western blotting assays, four protein concentrations, each differing two folds from each other were used. Previously, for quantification of NS2/3 protease inhibitors by western blot, 0.54 and 0.8 μM of final protein (904-1206) concentrations were employed in the literature.[1] Therefore a concentration range of 0.1 to 1 μM of the protein was employed in this

work. Samples with these concentrations were subjected to gel electrophoresis and western blot assays after one and four hours incubation time and the results were compared to those of the control zero-time incubation sample in each set of reaction. The blots were visualized on the film and quantified by ImageJ software (Chapter 4).

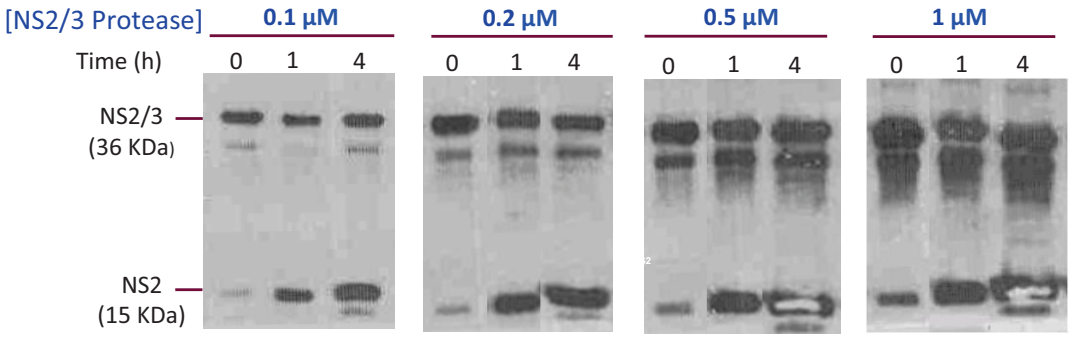


Figure 3.11: Evaluation of the effect of enzyme concentration on immunoblotting assays

In the gels shown in Figure 3.11 the top band is the NS2/3 parent protein (unprocessed or substrate) and the bottom band is the NS2 cleaved product. Quantification of 0.1 and 0.2 μM protein concentration samples and plotting the percent cleavage product over time demonstrated linearity over the time periods used in these experiments. Increasing the protein concentration to 0.5 and 1.0 μM resulted in saturation of both the substrate and cleaved product bands. Over-saturation is more evident from the formation of white spots on the NS2 cleaved product bands. Although some of these NS2/3 concentrations have been used in literature in the western blot experiments for quantification purposes, the saturation of bands that we observed at these concentrations could hamper their reliable quantification.

After repeating the experiments, some inconsistent results were observed at a concentration of 0.1 μM NS2/3 protein. Presuming that a concentration of 0.1 μM was

near the detection limit for quantification, 0.2 μM was selected as the concentration to use with the western blot quantification experiments.

3.3.2 Determination of the Optimal NS2/3 Protease Incubation Time

Evaluation of the enzymatic time course reaction is one of the essential factors in the optimization of any technique being employed. As demonstrated in the mass spectrometry approach, NS2/3 protease cleaves itself over time and this cleavage is initiated by using detergents *in vitro*.

With the goal of using this assay to aid in designing reversible substrate-based inhibitors, the objective was to perform the NS2/3 cleavage assay during the time period when the enzyme kinetics are linear. As in the case with the enzyme concentration, measuring an inhibitor's binding efficiency out of the linear range can change the efficiency results by orders of magnitude mainly through underestimating the inhibitor's activity. This is simply demonstrated in Figure 3.12 where, in the early stage of the enzyme reaction, the percent cleavage increases linearly over time. In this stage (linear phase) the forward enzyme-catalyzed reaction convert the substrate to the product.[165] Later, as the substrate becomes depleted, the plot of product formation versus time reaches the plateau phase. Evaluation of the efficiency of reversible inhibitors should be performed in the linear range of protein incubation time. This is because in the plateau phase the enzyme-inhibitor equilibrium has been reached and therefore alteration of inhibitors concentrations does not reflect their potency.

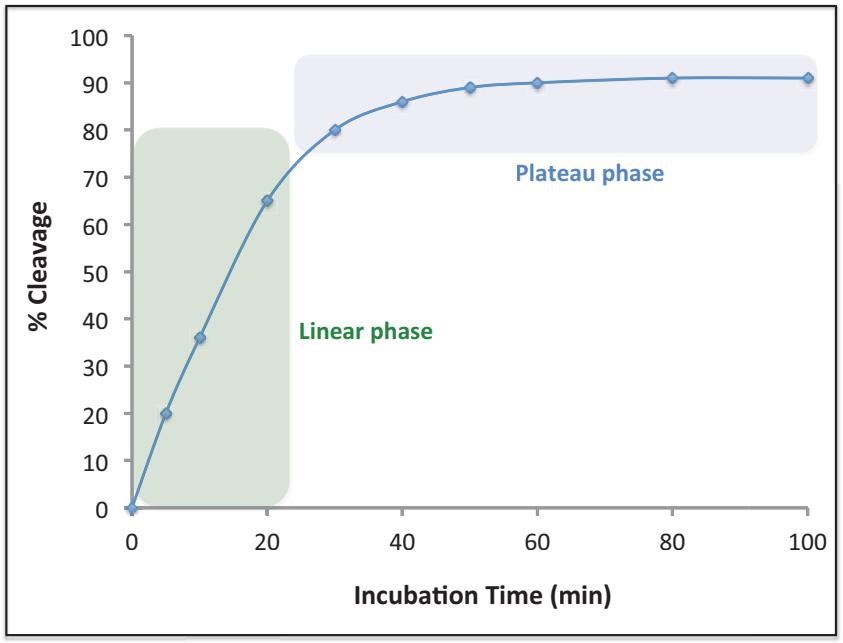


Figure 3.12: Schematic representation of product formation over time. Example curve; percent cleavages and incubation time parameters are not related to any specific enzyme reaction

Therefore, a time course reaction using a concentration of 0.2 μM of the NS2/3 protein, selected based on the experiments described in the previous section, was performed. Following incubation times of 15, 30, 60 and 180 minutes, it was observed that the percent cleavage increased over time (Figure 3.13).

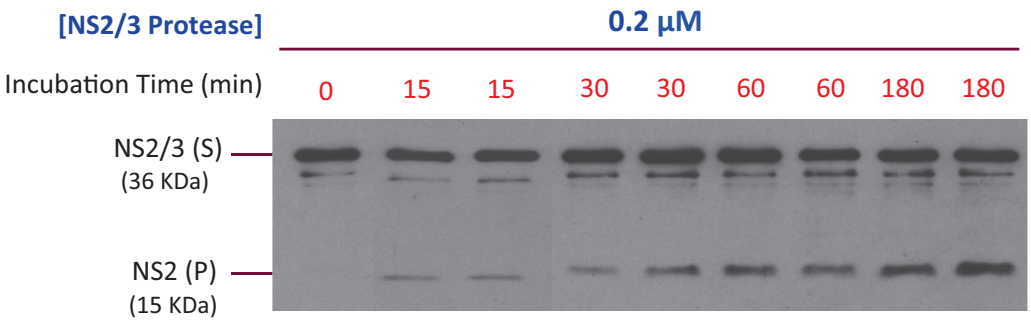


Figure 3.13: Evaluation of the effect of enzyme incubation time on immunoblotting assays (Number of replicates: 2)

The amount of protein in each band was quantified by comparing to those of the control zero incubation time and the plot of percent cleavage versus time was obtained (Figure 3.14). This plot showed that the percent cleavage as a function of time was linear up to an incubation time of approximately 30 minutes. Since it is important to stay in this linear phase when using immunoblotting assay to evaluate inhibitors, an incubation time of 15 minutes, well within this linear phase, was selected as the protein incubation time.

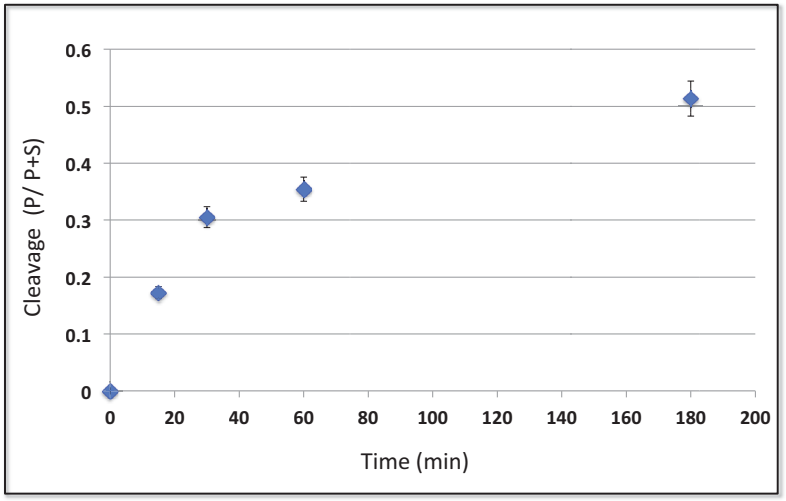


Figure 3.14: Plot of cleavage product versus time obtained from time course experiment immunoblot (Number of replicates: 2)

In addition to determining the linear phase of the cleavage process, some preliminary information about the enzyme kinetics was obtained from this time course study. Based on equation 3.1, by plotting the natural logarithm of ratio $\left(\frac{S}{S+P}\right)$ as a function of time (Figure 3.15), the observed rate constant (k_{obs}) of the pseudo-first order NS2/3 cleavage reaction can be extracted from the slope. The experimentally determined value of k_{obs} of $3.0{\times}10^{-05}$ s^{-1} by western blot assay is comparable with the k_{obs} of 1.0×10^{-05} s^{-1} from our study by mass spectrometry, which demonstrated a 3-fold difference. Yet, our data demonstrates a close value of NS2/3 protease rate of reaction

when compared to the study by Thibeault *et al.*[1] for similar NS2/3 protease construct and assay condition (3.8×10^{-05} s^{-1}).

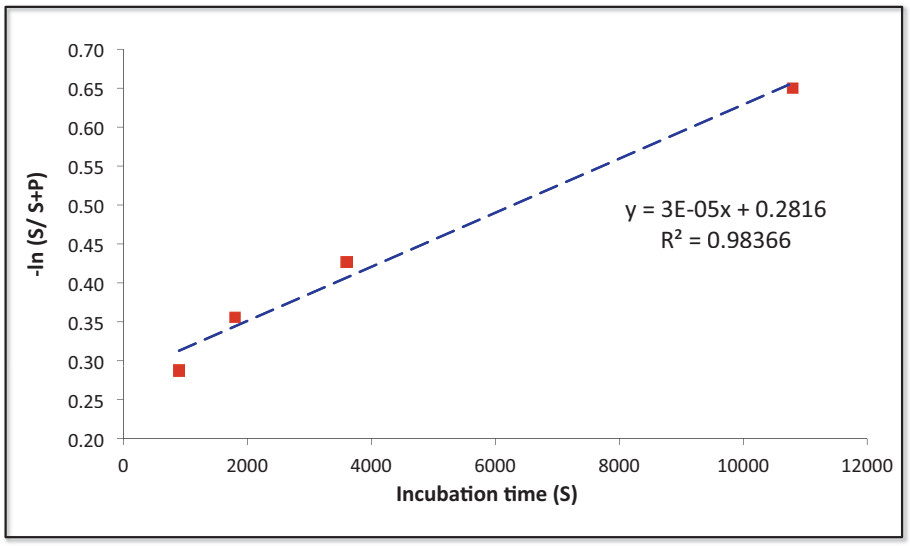


Figure 3.15: Plot of natural logarithm of the ratio $\left(\frac{S}{S+P}\right)$ as a function of incubation time. (Number of replicates: 2)

Overall, a time course study of NS2/3 protease through a western blot assay was performed in order to determine the linear phase time parameters. Based on two replicates of the experiments, this linear phase was observed over the first 30 minutes of the enzyme reaction.

3.4 Rational Design of NS2/3 Substrate-Based Inhibitors

Employing substrate-based inhibitors has been a useful approach towards the development of potent inhibitors of various enzymes. Initial studies on the function of NS2/3 protease demonstrated that a decapeptide from the N-teminus of the substrate (SFEGQGWRLL, P10-P1) inhibited the enzyme with an IC_{50} of 90 μM.[1] Therefore this decapeptide provided a starting point for the development of NS2/3 protease

inhibitors.

Although the natural substrate-based decamer provided some inhibition of the NS2/3 protease, large peptide inhibitors generally do not have drug-like properties. Some of the obstacles of large molecules are their poor solubility, metabolism by proteases, low absorption through membranes and as a result low entry into the circulatory system.[137] Although large molecules can be administrated through intravenous injection to overcome some of these deficiencies, this is not an ideal route of administration for patients compared to oral administration. Determination of the minimum peptidic sequence that is recognized by the enzyme aids in development of small molecule inhibitors of this enzyme. Since each residue in the peptide backbone acts as a potential recognition element for the enzyme; (Figure 3.16) therefore, removing or substituting each amino acid sequentially provides a means to probe the importance of various interactions of the peptide backbone and side chains with the enzyme.

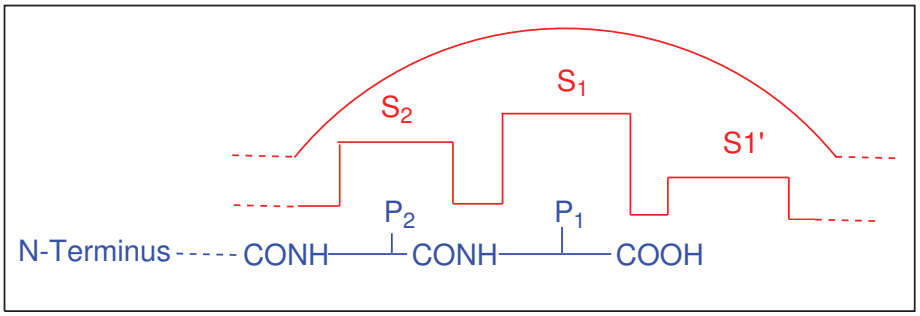


Figure 3.16: Substrate nomenclature and inhibitor binding to the active site. The blue sequence of amino acids represents the designed inhibitor; The cleavage site of protein is between S_1 and S_1'

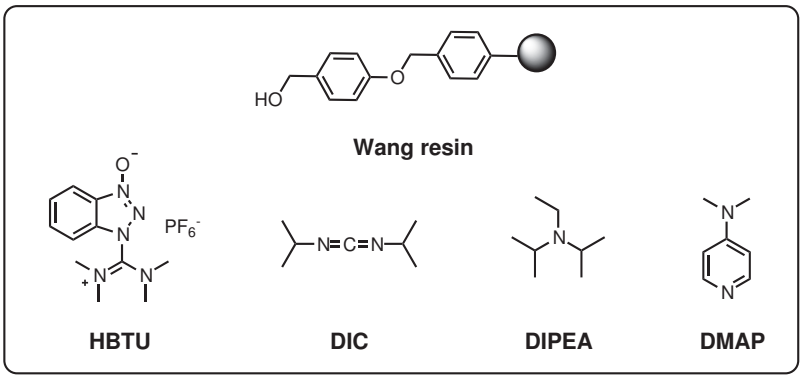
As part of the process of designing and synthesizing small molecule inhibitors of the HCV NS2/3 protease, the potency of each peptide was determined after systematic removal of each residue (truncation studies) from the N-terminus of the decapeptide. The backbone and side chains of the amino acids starting from the P1 position of each peptide (C-terminus) potentially provide more effective interactions with the NS2/3 cysteine protease compared to those that are farther from this position. Therefore,

each amino acid removal was started from the N-terminus of the peptide to gradually probe the importance of each amino acid interaction with the enzyme. Further, the substrate-based decapeptide could be used as a comparative reference for the *in vitro* evaluation of the new peptides as NS2/3 protease inhibitors.

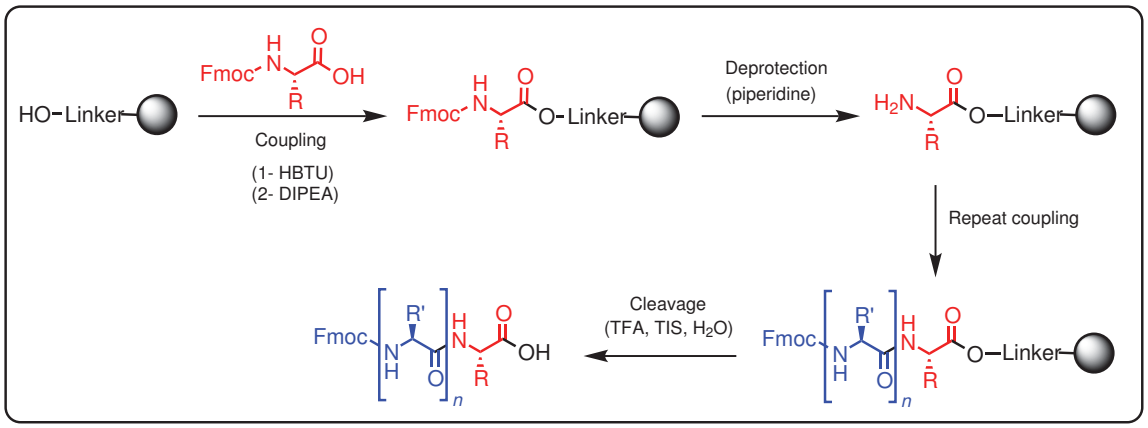
3.4.1 Peptide Synthesis

A general synthetic protocol for the preparation of the P10-P1 sequence of the natural NS2/3 substrate and N-terminus truncated peptides (P9-P1, P8-P1, etc.) was developed. Peptides were manually synthesized on the Wang resin[166] based on standard Fmoc solid-phase peptide chemistry[167,168] (Chapter 4). Solid-phase peptide synthesis has several advantages over traditional solution-phase peptide synthesis. Among these are the convenient removal of unreacted reagents in each step, allowing for the use of excess amounts of reagents in order to drive each step towards completion, and simplified isolation/purification of the intermediates in each step as a consequence of their attachment to the resin.[168]

Fmoc-protected amino acids were attached to the Wang resin (Scheme 3.1) as the solid support through their carboxyl groups. Reactive groups in the side chains of the amino acid residues were protected with protecting groups that are stable under basic conditions. However, the α-amino group was masked with a temporary protecting group (Fmoc) that was removed under basic condition and before the introduction of subsequent amino acids (Scheme 3.2). The next Fmoc protected amino acids were attached to the free amine using coupling reagents such as HBTU and DIC and bases such as DIPEA and DMAP (Scheme 3.1). Treating the peptides with a mixture of TFA/H_2O/TIS released the peptides from the resin. In addition, all the side-chain protecting groups were removed under these conditions (Scheme 3.2).



Scheme 3.1: Chemical structures of resin, coupling reagents and bases employed in the solid-phase peptide synthesis



Scheme 3.2: Solid-phase peptide synthesis on the Wang resin employing the Fmoc-protected amino acids

Based on this procedure, a range of peptides, listed in Table 3.4 were synthesized and purified through reversed-phase high performance liquid chromatography (RP-HPLC). Initially, the Fmoc-protecting group was retained and was not deprotected in the final stage of the synthesis of the substrate-based decapeptide in order to avoid zwitterion formation and complications in the purification. Based on the promising results of this Fmoc-protected decapeptide (vide infra) compared to the reference decapeptide from the literature, however, the Fmoc protecting group was maintained in all final peptides synthesized. The purities of the peptides were confirmed by

analytical RP-HPLC and the molecular masses of all compounds were obtained by Q-TOF mass spectrometry (Table 3.4).

Peptide Sequence	**Theoretical** $[M+H]^+$	**Obtained** $[M+H]^+$
Fmoc-S-F-E-G-Q-G-W-R-L-L	1414.67	1414.60
Fmoc-F-E-G-Q-G-W-R-L-L	1327.64	1327.30
Fmoc-E-G-Q-G-W-R-L-L	1180.57	1180.91
Fmoc-G-Q-G-W-R-L-L	1051.53	1051.50
Fmoc-Q-G-W-R-L-L	994.51	994.56
Fmoc-G-W-R-L-L	866.45	866.54
Fmoc-W-R-L-L	809.43	809.47
Fmoc-R-L-L	623.35	623.38
Fmoc-L-L	467.25	467.28

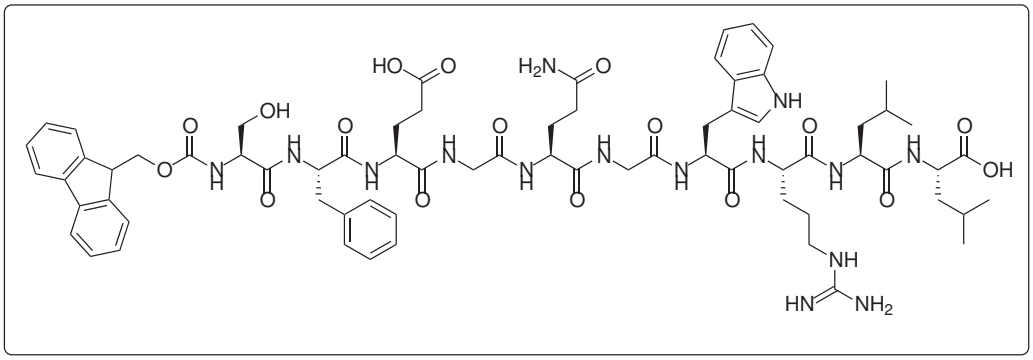
Table 3.4: Synthesized peptides from truncation approach using solid phase peptide synthesis. Peptides were purified using semi-preparative HPLC with a C-18 column; Full names of amino acids available in abbreviations

3.4.2 *In Vitro* Evaluation of Substrate-Based Peptides

The peptides listed above (Table 3.4) were evaluated for NS2/3 inhibitory activity by means of the western blot technique. Stock solutions of peptides were prepared in DMSO to a range of concentrations and pre-incubated with the NS2/3 protease for 15 minutes. After the pre-incubation period the detergent was added to the protein-inhibitor mixture and each sample was then incubated for 15 minutes (Chapter 4).

Up to this point, a decapeptide (SFEGQGWRLL) that is the N-terminal cleavage product of the NS2/3 protease had been identified as the most potent substrate-based peptide inhibitor of NS2/3 protease cleavage.[1] Initially, the synthesized Fmoc-decapeptide (Scheme 3.3) was evaluated by the western blot technique to provide a comparison to the deprotected decapeptide in the literature. Using our optimized con-

ditions (protein concentration: 0.2 μM, incubation time: 15 minutes) a series of assays were performed with various concentrations (0.37 to 120 μM) of the Fmoc-decapeptide.



Scheme 3.3: Structure of Fmoc-decapeptide

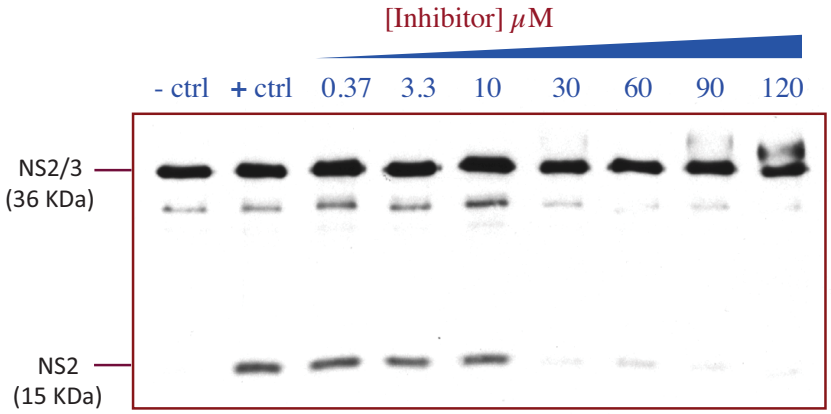


Figure 3.17: Dose-response NS2/3 inhibitory activity of Fmoc-decapeptide by immunoblotting

The resulting western blot from this experiment is shown in Figure 3.17. The first band from the left (−ctrl) is the zero time control in which the reaction was stopped immediately after the addition of detergent. The second band from the left (+ctrl) is the positive control protein that was performed in the absence of an inhibitor, and represents the maximum cleavage of the NS2/3 to give NS2 under these assay conditions. The blot displays a dose-response relationship whereby an increase in the

concentration of inhibitor results in a decrease in the amount of cleaved NS2 product formed. Since substantial amounts of the substrate were still present after 15 minutes incubation time, the changes in substrate concentration were not visually obvious on the gel; however, quantification of the signals by ImageJ software provided data that could be used to generate an IC_{50} value. The IC_{50} indicates the concentration of the inhibitor or drug at which the target's activity is 50% inhibited. The IC_{50} was determined from the dose-response curve (Figure 3.18), generated by plotting the percent inhibition against the common logarithm of the concentration of the inhibitor for the series of assays conducted. Origin software was used to fit a curve to the data.

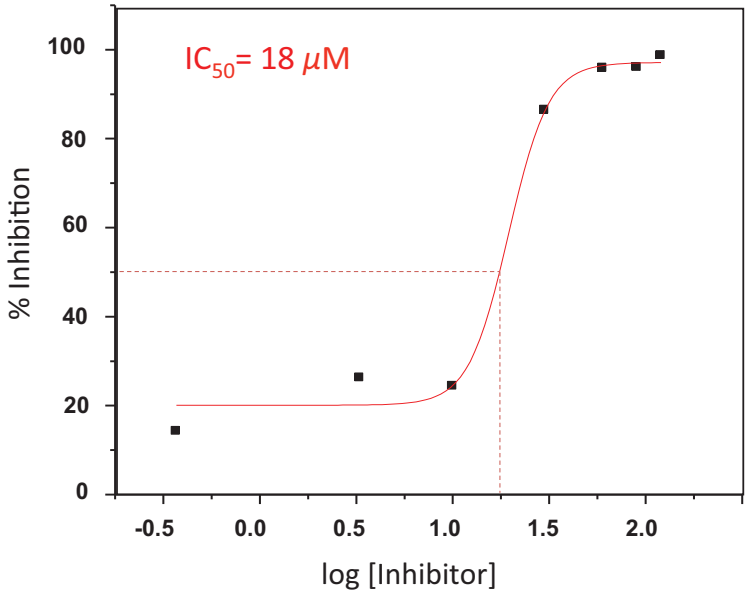
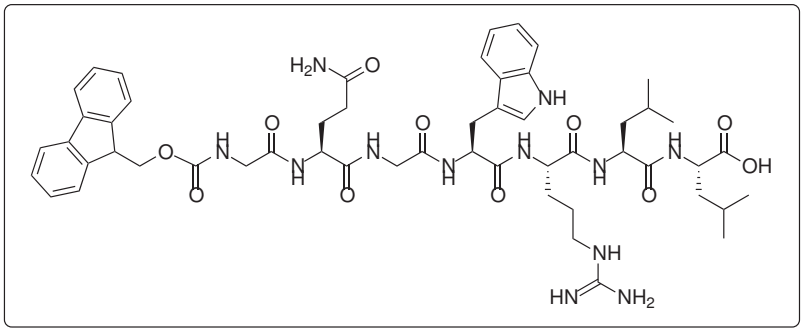


Figure 3.18: Dose-response curve of Fmoc-decapeptide. Standard deviation: 0.58 (Two experiments)

The half maximal inhibitory concentration of the NS2/3 protease by Fmoc-decapeptide was determined to be 18 μM. This value shows that the Fmoc-decapeptide is five-fold more potent than the deprotected decapeptide (without the Fmoc protecting group) which has an IC_{50} value of 90 μM. This increase in potency suggests that this

protecting group may be involved in some additional interactions with the enzyme that stabilize the inactive conformer and thereby decrease the cleavage activity. Regardless, the results suggested that maintaining the Fmoc protecting group in the structure of the other synthesized site-derived peptides may be beneficial for inhibitory activity.

In vitro evaluation of the 9- and 8-mer peptides was hindered due to solubility issues. Therefore the 7-mer peptide (Scheme 3.4) was evaluated next for NS2/3 inhibition through the same procedure used for the decapeptide.



Scheme 3.4: Structure of Fmoc-heptapeptide

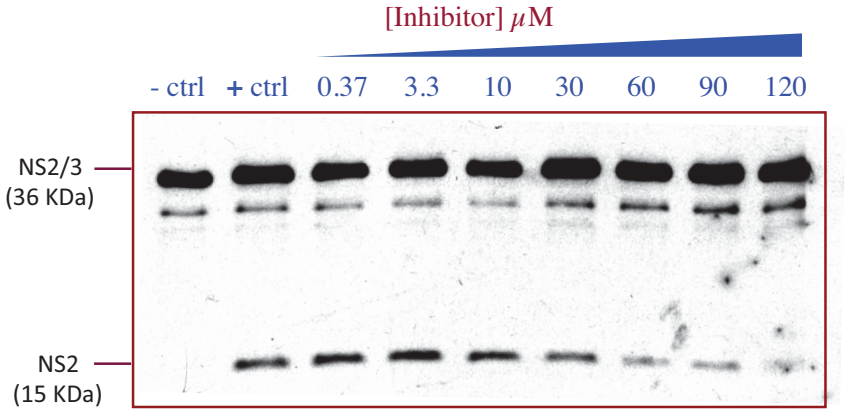


Figure 3.19: Dose-response NS2/3 inhibitory activity of Fmoc-heptapeptide by immunoblotting

The removal of an amino acid may lead to the loss of some enzyme-inhibitor

interactions since each residue in the peptide backbone functions as a potential recognition element for the enzyme. Thus, as a general principle, it was expected that truncation of residues from the C-terminus of the decapeptide would result in some loss of potency if the residues were involved in important binding interactions.

The western blot for the Fmoc-heptapeptide is shown in Figure 3.19, and the resulting dose-response curve is shown in Figure 3.20. From the dose-response curve, the IC_{50} value of the Fmoc-heptapeptide was measured as 32 μM, which corresponds to about a two-fold loss in potency compared to the Fmoc-decapeptide. This amount of difference can be attributed to errors in the reaction assays/quantification methods, and should not necessarily lead to the conclusion that the heptapeptide is less potent than the decapeptide.

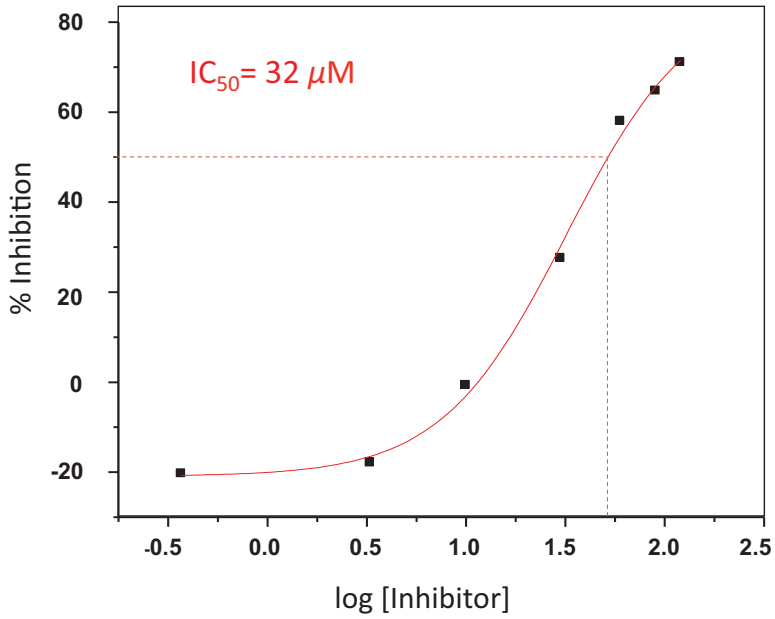
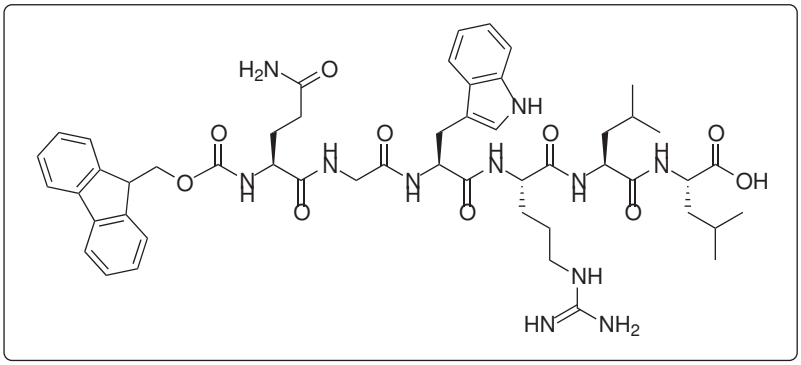


Figure 3.20: Dose-response curve of Fmoc-heptapeptide. Standard deviation: 0.30 (Two experiments)

Finally, Fmoc-hexapeptide (Scheme 3.5) was subjected to the western blot assay.

For the same reasons outlined previously, a higher or equal value of the IC_{50} was expected for this peptide in comparison to the Fmoc-heptapeptide. The potency of Fmoc-hexapeptide to inactivate the NS2/3 protease was evaluated using the same procedure described for the other truncated peptides, except that a maximum concentration of 60 μM was used due to solubility issues.



Scheme 3.5: Structure of Fmoc-hexapeptide

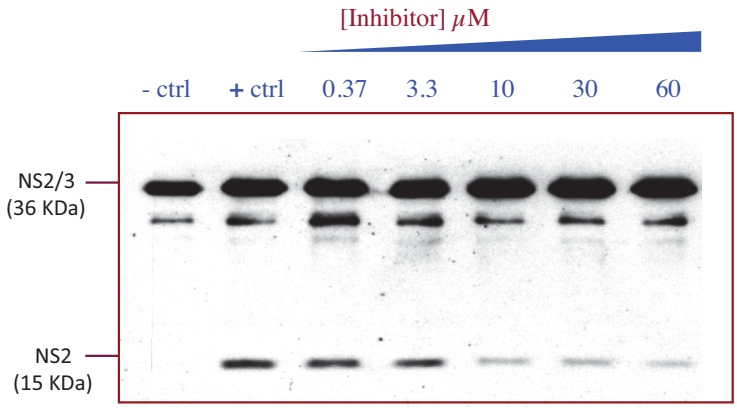


Figure 3.21: Dose-response NS2/3 inhibitory activity of Fmoc-hexapeptide by immunoblotting

The western blot for the Fmoc-hexapeptide is shown in Figure 3.21, and the resulting dose-response curve is shown in Figure 3.22. Interestingly, from the dose-response curve, an IC_{50} value of 4.8 μM was measured for the Fmoc-hexapeptide.

Thus, removal of amino acid residues in this case resulted in 6- and 4-fold gains in potency compared to the Fmoc-heptapeptide and Fmoc-decapeptide, respectively.

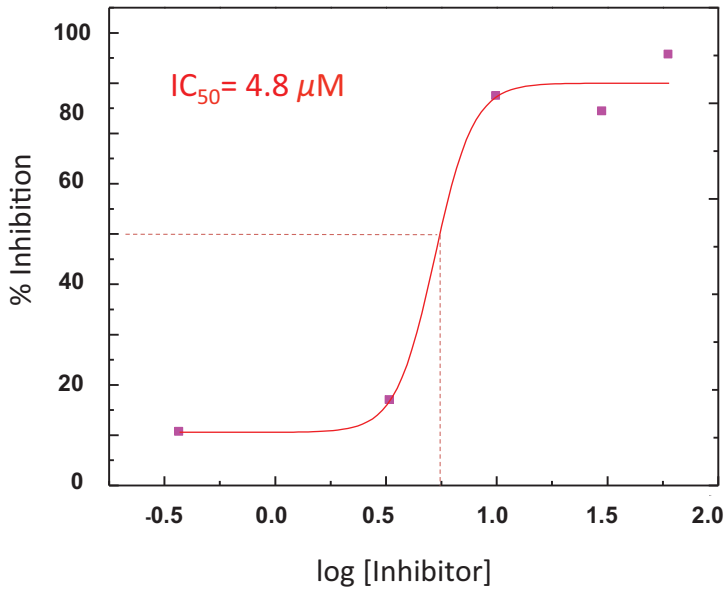


Figure 3.22: Dose-response curve of Fmoc-hexapeptide. Standard deviation: 0.63 (Two experiments)

Other NS2/3 site-derived peptides containing a similar hexapeptide sequences have been previously evaluated by other groups.[1,102] As illustrated in Table 3.5, compound (**18**) encompassing a hexapeptide in the P site and a pentapeptide in the P' site did not have any inhibitory effect on the enzyme activity, whereas compound (**25**), which differs in possessing a hexapeptide in the P' site, inhibited the enzyme with the IC_{50} value of 630 μM. Employing the N-terminus acetylated hexapeptide aldehyde as well as the acetylated hexapeptide hydroxamate did not improve the potency for inactivation of the NS2/3 protease.

Compound	**Peptide**	**[C]^a(mM)**	**IC$_{50}$ (μM) or % inhibition***
NS2/3 site-derived peptides			
18[102] (Table 2.3)	GWRRLL*APITA	0.1	<5*
25[1] (Table 2.4)	KGWRLL*APITAY	-	630
29[102]	Ac-GWRRLL-CHO	0.1	<5*
30[102]	Ac-GWRRLL-CONHOH	0.1	<5*

Table 3.5: Effect of various hexapeptides as part of a larger peptide sequence on NS2/3 protease inhibition

^{a}Abbreviation: [C] is the final concentration of the peptide in the reaction assay; Ac: Acetyl

While some studies examining these types of site-derived peptides as NS2/3 protease inhibitors have been reported in the literature, none have examined the simple P-site derived hexapeptide (P6-P1) as we did in this study. This Fmoc-hexapeptide has been the most potent NS2/3 site-derived peptide inhibitor reported so far. These results strengthen the hypothesis that the Fmoc protecting group is able to interact specifically with the enzyme, presumably through $\pi - \pi$ stacking interactions with the aromatic residues of the enzyme. Moreover, it is likely that removing the amino acids alleviates some steric interactions between the inhibitor and the enzyme, and therefore allows for improved potency. In the absence of an enzyme-inhibitor crystal structure, NMR studies or a molecular modeling studies to provide supporting evidence, however, this explanation remains conjecture at this point. Overall, the encouraging results obtained from the Fmoc-hexapeptide would open the research area for developing small molecule inhibitors of the NS2/3 protease based on this peptide.

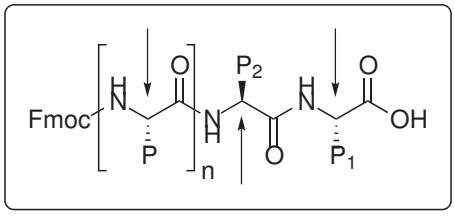
3.5 Future Directions

In order to prepare the NS2/3 protease inhibitors, different strategies and synthetic approaches can be investigated. In our study, the Fmoc-hexapeptide was demonstrated to be the most potent site-derived peptide reported against the inhibition of the NS2/3 protease so far. Truncation of amino acids enables the determination of the important interactions of the inhibitor's backbone and side chain residues with the enzyme.

Another approach in the development of protease inhibitors involves the identification of important interactions of the inhibitor's side chains solely, with the enzyme through alanine scanning. Following the determination of the critical interactions, modification of site-derived peptides at P1 position will be explored to prepare potent cysteine protease traps. The following approaches will be taken to increase the binding efficiency of the synthesized peptides.

3.5.1 Evaluation of the Side-Chain Binding Affinity

The importance of the binding affinity of each amino acid residue (such as hydrogen binding), which is a key interaction in the recognition of natural substrate, will be explored. Therefore, systematic replacement of each amino acid in the N-terminus cleavage product of the enzyme with the amino acid alanine has been employed (Scheme 3.6).



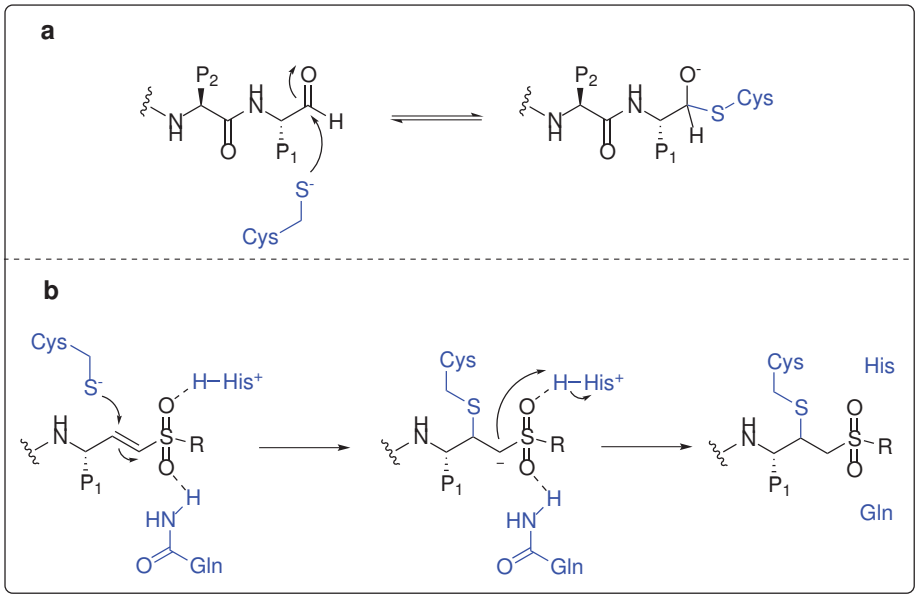
Scheme 3.6: Evaluation of hydrogen bonding by alanine scanning. Arrows show the position that alanine will be replaced

Alanine has a small, hydrophobic moiety and its backbone conformation and

flexibility resemble the replaced residues. By this approach functional information of each amino acid and its importance in the recognition by the active site will be achieved. In our study, Fmoc-hexapeptide was subjected to these replacements to identify the essential amino acids required for recognition by the enzyme. Each replacement will then be evaluated with *in vitro* assays such as western blot and mass spectrometry.

3.5.2 Increasing the Electrophilicity of P1 Anchor

Most inhibitors of the cysteine proteases have exploited the mechanism of amide bond hydrolysis and contained an electrophilic functionality that reacts with the active site cysteine residue. Thereby, the minimum fragments of the peptides that have shown enzymatic activity will be coupled to better cysteine electrophiles. Both reversible and irreversible inhibitors of cysteine targets will be employed at the P1 position. Therefore, we expect to have high-affinity active-site ligands in this phase of the project by employing Michael acceptors (Scheme 3.7, **b**) and aziridines as irreversible inhibitors of the NS2/3 protease since they will bind to the enzyme covalently.



Scheme 3.7: Increasing the electrophilicity of potential peptide inhibitors. **a**: A peptide aldehyde as reversible inhibitor of cysteine protease; **b**: A peptide Michael acceptor as irreversible inhibitor of cysteine protease

Alternatively, reversible inhibitors bind to the active site of the enzyme non-covalently through hydrogen bonds, ionic bonds or van der Waals interactions. However our approach is to synthesize two classes of reversible inhibitors which bind to the enzyme covalently. Such P1 cysteine protease traps are typically aldehydes (Scheme 3.7, **a**) and nitriles that we will enhance the design of future generations of the reversible NS2/3 protease inhibitors.

3.6 Conclusion

NS2/3 protease has been one of the most challenging HCV proteins to study. This is evident by the number of marketed dugs to inhibit NS3/4A protease, NS5A protein and NS5B polymerase but non for the inhibition of the NS2/3 protease. To date neither a small-molecule inhibitor nor an effective drug target of NS2 protease has been reported. Despite the fact that designing inhibitors for an enzyme with intra-molecular enzymatic reaction appears as an obstacle, a rational design assisted by various methods such as

molecular modeling, mass spectrometry and NMR studies can provide such molecules. In this work, tryptic digestion followed by tandem mass spectrometry were carried out to obtain initial enzymatic information of the NS2/3 protease. Tandem mass spectrometry was established as a precise method for the kinetics studies of the enzyme, however, it would be a starting point to employ this method for evaluation of potential inhibitors. Furthermore, gel electrophoresis and western blot techniques were optimized for this enzyme and the obtained kinetics data were compared to similar studies. Rational design of the NS2/3 protease inhibitors initiated with systematic truncation of the NS2/3 protease site-derived peptides implicating peptide synthesis. An Fmoc-hexapeptide was discovered as the most potent peptidic inhibitor of this enzyme. This would be a starting point to modify and develop more potent and smaller molecule inhibitors towards inhibition of the NS2/3 protease.

Chapter 4

Experimental

4.1 NS2/3 Protease

In the present work the purified NS2/3 protein was kindly provided by Boehringer Ingelheim (Laval, Canada, Ltd.). The NS2/3 protein (904-1206) contained four lysine residues, followed by a histidine tag at its N-terminus and another four lysines at the C-terminus. After purification by a chelating column containing Ni^{+2}, the inactive NS2/3 protein was stored in the refolding buffer containing 50 mM Tris pH 8.0, 0.5 M arginine HCl, 5 mM TCEP, 1% LDAO.[1] The stock aliquots of the protein were stored at -80 °C until their activation for auto-cleavage.

4.1.1 Materials

n-dodecyl-β-D-maltoside, HEPES, Tween® 20, Tris HCl and glycerol were obtained from BioShop® in biotechnology grades. TCEP was purchased from Thermo Fisher Scientific.

Materials for western blot assay: Amersham ECL western blotting reagent from GE Healthcare Life Sciences or Pierce ECL western blotting reagent from Thermo Fisher Scientific were purchased. Protein bands were visualized on Carestream®

Kodak® BioMax® MS films (20×25 cm) using a radiography instrument. A rabbit polyclonal anti-NS2 antibody raised against NS2 (residues 904-1026) for probing NS2 protein in western blot experiments. The NS2 antibody was donated by the McGill cancer center (Dr. Arnim Pause laboratory). Goat polyclonal secondary antibody to rabbit (horseradish peroxidase conjugated secondary antibody) was purchased from Abcam®. Nitrocellulose membranes (pore size 0.2 μm) and prestained protein ladder (all blue, 10-250 kDa) were obtained from Bio-Rad.

4.1.2 Enzyme Auto-Cleavage Activity

Based on the procedure reported by Thibeault *et al.*[1] the "cleavage buffer" including 50 mM Hepes pH 7.0, 50% glycerol (w/v), 1 mM TCEP and *n*-dodecyl-β-D-maltoside (DM) was employed. The concentration of detergent varied depending on the 0.5% final detergent concentration for the 0.54 μM protein concentration, however it never exceeded 0.5% in the final reaction mixture. All samples contained a 2-5% final concentration of DMSO depending on the experiment. Also the concentration of DMSO did not exceed 5% in the final reaction mixture.

Protein samples were prepared for the enzymatic reaction in the following manner (protein concentration varied in some experiments): To the cleavage buffer were added the inhibitor or the same volume of DMSO as the vehicle control. Protein was added and the mixture was pre-incubated for 15 minutes at 23 °C (mixtures were stirred at 400 rpm in the mass spectrometry experiments). *n*-Dodecyl-β-D-maltoside was added to initiate the cleavage reaction and the incubation time was measured from this time point. The reaction was stopped by addition of SDS (sodium dodecyl sulfate), Laemmli buffer in the western blot assays and by addition of formic acid in the mass spectrometry experiments. These were added to the samples right after the addition of the detergent (DM) in case of zero incubation time.

4.2 Acetone Precipitation

The procedure provided by Thermo Scientific[169] was followed for acetone precipitation with slight modifications. To the mixture of the NS2/3 protein, cleavage buffer and 2% DMSO was added 0.5% of *n*-dodecyl-β-D-maltoside when the final concentration of the protein was 0.54 μM. After specific incubation time, pre-cooled acetone at -20 °C was added to the four times volume of this reaction mixture. The sample was mixed well and incubated overnight at -20 °C. The sample was centrifuged at 13000-15000 xg for 10 minutes at room temperature and the supernatant was carefully disposed. This cycle was repeated twice. Subsequently, residual acetone in the sample was evaporated at room temperature and the sample was prepared for the mass spectrometry analysis.

4.3 Mass Spectrometry Measurement

Following the acetone precipitation, protein pellets were dissolved in 5% acetonitrile/0.1% TFA. Samples were injected onto a reversed-phase VYDAC® column (5μ, 100 mm) that was equilibrated with 5% aqueous acetonitrile/0.1% formic acid using an Agilent 1100 HPLC. A flow rate of 0.2 ml/min was used. Solvent gradients were as the following: 5-95% acetonitrile in 5 minutes, constant acetonitrile in 95% for 3 minutes and 95-5% acetonitrile in 3 minutes. Samples were eluted into the electrospray (ESI) source of a Q-TOF Ultima API Mass Spectrometer (Waters). Mass calibration was applied by employing horse heart myoglobin as a standard (average mass = 16951.49 u, $C_{769}H_{1212}N_{210}O_{218}S_2$). Other employed mass spectrometry parameters were as the following: source temperature 80 °C and desolvation temperature 300 °C. Capillary voltage 3.5 kV and cone voltage 35 V. MaxEnt1 algorithm was employed for deconvolution of protein envelopes.

4.4 Trypsin Digestion and Sample Preparation

For the final volume of 150 μL reaction mixture, the NS2/3 protease (0.54 μM) was added to the cleavage buffer (100 μL) containing 5% DMSO and the mixture was pre-incubated at 23 °C for 20 minutes with shaking at 750 rpm. n-Dodecyl-β-D-maltoside (0.5%) was added and protein samples were incubated at specific time intervals. The enzymatic reaction was quenched by addition of 1% formic acid (30 μL). 100 mM ammonium bicarbonate pH 8.5 was added subsequently. The sample was incubated with 50 mM dithiothreitol (DTT) for 10 minutes at 25 °C to reduce the disulfide bonds (750 rpm) and was incubated with 50 mM iodoacetamide (IAM) in the dark for 30 minutes at 37 °C (750 rpm) to alkylate the reduced bonds. 2.4 μg trypsin were added to the sample and it was incubated for 18 hours at 37 °C (750 rpm). After digestion process, 500 μL water was added to the sample and it was loaded onto an OASIS® HLB column (30 mg) which was pre-washed with 1 mL methanol and 1 mL water. The sample tube was washed with another 500 μL water and loaded onto the column and 1 mL water was added to the column too. Sample on the column was washed with optimized organic solvent mixture, ACN/IPA 60:40 for two times (500 μL). Through this mixture of solvent, sample was collected from the column and was dried under vacuum (Thermo Fisher Scientific Universal Vacuum System, Asheville, NC) for 3 hours. 10% ACN (100 μL) was added to the protein sample for the LC-MS/MS analysis.

4.5 Reverse-Phase UHPLC-MS/MS

NS2/3 protein samples in 10% ACN or synthetic standard peptides in NS2/3 cleavage buffer were loaded (20 μL) onto a 2.1 × 100 mm Kinetex® XB-C18 column with 1.7 μm, 100 Å, solid core particles (Phenomenex, Torrance, CA), by employing a Nexera® UHPLC (Shimadzu, Columbia, MD). The column was equilibrated with 5%

aqueous acetonitrile-0.1% formic acid (B). As it is illustrated in Figure 4.1, the column was maintained at 5% (B) at a flow rate of 300 μL/min and a gradient from 5-24% acetonitrile over 9 minutes, 24-80% acetonitrile over 30 seconds, 80% acetonitrile over 4.5 minutes, 80-5% acetonitrile in 17 minutes and 5% acetonitrile over 8 minutes was applied.

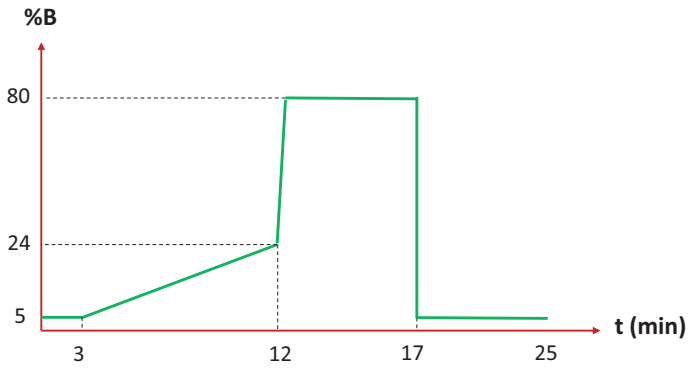


Figure 4.1: Solvent gradient in UHPLC-MS/MS

A Rheodyne switch valve (IDEX Health and Science, Oak Harbor, WA) was employed to avoid the entering of any salt into the ion source. Therefore, elution between 0-2 minutes and 17.5-25 minutes were sent to the waste.

A high-resolution hybrid quadrupole-time-of-flight (Q-TOF) TripleTOF® 5600mass spectrometer (AB Sciex, Concord, ON, Canada) combined with a DuoSpray™ ion source (positive ion mode) was used and the total ion current (TIC), MS and MS/MS data were visualized by PeakView® software version 1.2. The peptides were identified based on the MS/MS data by ProteinPilot™ software version 4.1. Quantification of peptides and peak integration were performed by MultiQuant™ software version 2.1.

4.6 SDS-PAGE and Western Blot

Following the addition of SDS Laemmli buffer, protein samples were boiled at 95 °C for 5 minutes. In case of using 0.21 μM of protein concentration, which was employed

in most of the experiments, 52 μg of the protein was loaded on a 15% sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis applying the voltage of 100-130. The separated proteins were transferred to the supported nitrocellulose membrane using transfer buffer (5.8 g tris base, 2.9 g glycine, 0.37 g SDS, 200 mL MeOH, 800 mL dH_2O) by applying 80 voltage for 80 minutes. The membrane was blocked for one hour in 5% w/v dried milk dissolved in TBS (tris-buffered saline) with 0.5% tween® 20. The membrane was blocked with 1:5000 dilution of anti-NS2 antibody in TBS buffer for one hour. After 4 times wash, the membrane was blocked in a 1:10000 dilution of horseradish peroxidase (HRP) conjugated secondary antibody (anti-rabbit). After 7 times washing steps, the membrane was incubated in the ECL reagent and the protein signals were detected on a Kodak film using a radiography instrument. Protein densitometry was carried out using ImageJ analysis software.

4.7 Peptide Synthesis

4.7.1 Materials

All Fmoc-L-amino acids, Wang resin (1.0-1.5 mmol/g OH loading, 1% cross-linked with divinylbenzene, 70-90 mesh), O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU), N,N'-Diisopropylcarbodiimide (DIC), 4-Dimethylaminopyridine (DMAP), pyridine, trifluoroacetic acid (TFA), triisopropylsilane (TIS) were purchased from Sigma-Aldrich. DMF was dried and stored over activated 3 Å molecular sieves in a flame-dried Schlenk flask. Peptides were synthesized in a cylindrical reaction vessel equipped with fritted disc (coarse porosity) and two valves for introducing inert gas and vacuum. Peptide identification was performed on an Agilent 1200 reversed-phase HPLC. The preparative RP-HPLC was performed on semi-preparative C-18 column (Waters, 19×50 mm) or C-18 (ACE 5, 250×10 mm). The analytical RP-HPLC was carried out on a C-18 column (Agilent, Elipse XDB, 4.6×150 mm).

4.7.2 Fmoc-Solid Phase Peptide Synthesis

Peptides were manually synthesized on the Wang resin. Initially resin (1 equiv) was soaked in DMF for 30 minutes to swell. To a flame dried flask was added the first Fmoc amino acid (6.6 equiv) dissolved in DMF (2 mL). DIC (7 equiv) was added and the mixture was stirred at 0 °C for 20 minutes. DMAP (0.1 equiv) was added and after a few minutes stirring, the solution was added to the resin in a fritted reaction cylinder and agitated under argon for 1 hour. This step was repeated 2 more times. Beads were washed 5 times with DMF in order to remove the excess reagents and the Fmoc was removed by treating the beads with 20% solution of piperidine/DMF for 10 minutes (3 times). After more washing steps (5 times with DMF) the next couplings of the Fmoc amino acids (3 equiv) were performed with HBTU (3 equiv) as the coupling reagent and DIPEA (3 equiv) as the base in DMF for 1 hour. Ninhydrin (Kaiser test) was used for the identification of residual free amine in each step. Several coupling and deprotection steps were performed until the desired length of peptide sequence was obtained. Finally beads were washed 3 times with DMF, DCM and methanol separately. In the last step, a solution of TFA/TIS/H_2O (95:2.5:2.5 v/v) was added to the resin to cleave the peptide from the resin and to deprotect the side chain protecting groups from the peptide. TFA was removed from the samples by rotary evaporation and samples were lyophilized to obtain the crude peptide.

Mass spectrometry and HPLC were performed on the crude samples and later semi-preparative HPLC was carried out to purify the samples. The solvent system for both analytical and semi-preparative HPLC were water-0.05% TFA (A) and acetonitrile-0.05% TFA (B). Analytical HPLC was carried out using a linear gradient of 5-95% solvent B over 16 minutes with the flow rate of 1 ml/min. A linear gradient of 30-95% B was used for peptides purification over 30 minutes with the flow rate of 5 ml/min. Fractions were characterized by a Q-TOF Ultima API Mass Spectrometer (Waters) and analytical HPLC was performed to ensure the purity of collected fractions.

Combined fractions were dissolved in water/acetonitrile and lyophilized.

Chapter 5

Synthesis of 2,5-Diaryl Substituted Thiophenes as Helical Mimetics

5.1 Protein Structures

A great part of protein properties is governed by their dynamic characteristics including their folding, conformational and structural features. Four levels of protein structures are defined for the structural organization of proteins;[170] the primary structure of proteins is the linear arrangement of amino acids connected through covalent bonds to make the polypeptide chains (Figure 5.1, **a**). Secondary structure results from the backbone hydrogen bond interactions within a segment of a polypeptide chain. Two main motifs of secondary structure are α-helices (Figure 5.1, **b**) and β-sheets (β-pleated sheets, Figure 5.1, **c**). α-Helices are formed by the hydrogen bonds between the oxygen atom of a carbonyl group in the backbone (i) with the hydrogen atom of the amide group four residues further (i+4). The repetition of these hydrogen bonds establishes a right-handed coil motif engaging up to 40 amino acids. Each complete α-helix turn consists of 3.6 amino acids and spans a length of 5.4 Å.[170] β-Sheets are the parallel or anti-parallel shapes of the polypeptide segment resulting from the

hydrogen bonds between the NH of the amino acid in one chain and the oxygen atom of the carbonyl group in the adjacent chain. The tertiary structure of the protein refers to the three dimensional shape of the whole polypeptide chain that is formed as a consequence of hydrogen bonds, hydrophobic interactions and disulfide bonds. The tertiary structure demonstrates the folding of all secondary structures (α-helices and β-sheets) into a globular polyprotein. Proteins associated with more than one polypeptide chain can be stabilized by interactions similar to those found in tertiary structures and form quaternary structures (Figure 5.1, **e**).[170] Not every protein has a quaternary structure. Hemoglobin is an example of a protein possessing a quaternary structure.

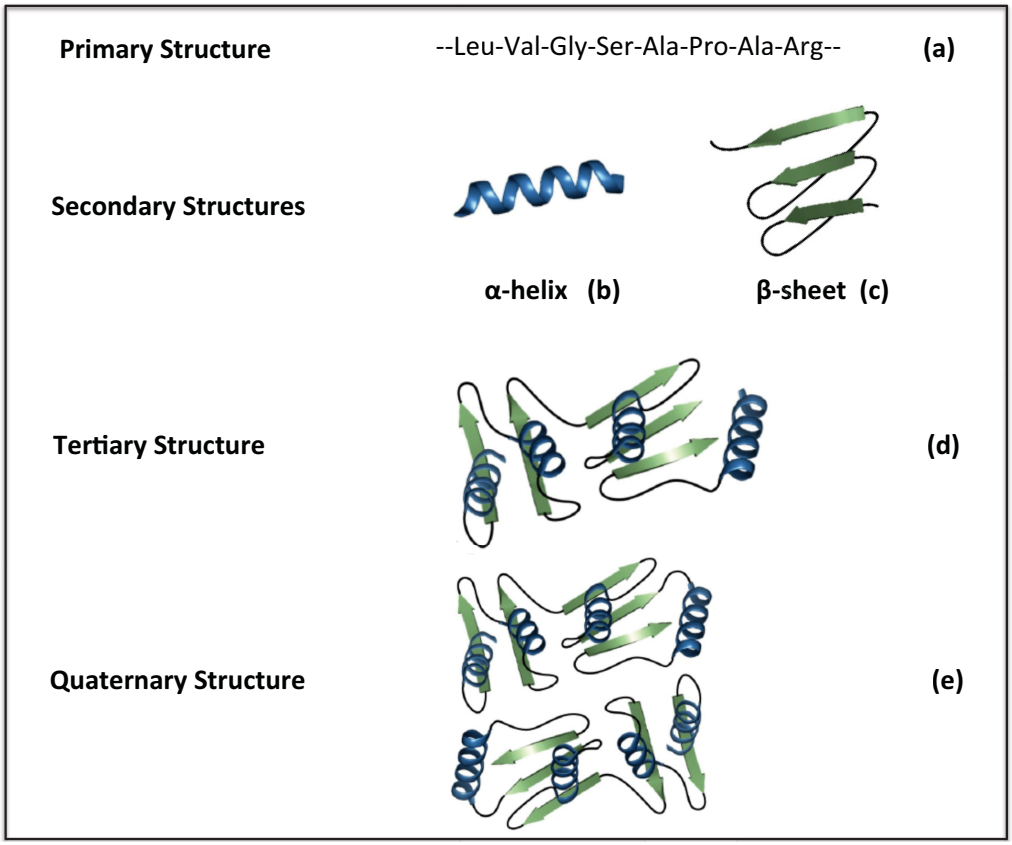


Figure 5.1: Levels of protein structures. (a) Primary structure; (b) α-Helix of the secondary structure; (c) β-Sheet of the secondary structure; (d) Tertiary structure; (e) Quaternary structure

5.2 Protein Misfolding

Protein folding refers to the proper distribution and unique arrangement of α-helices, β-sheets and random coils to establish the tertiary structure of the protein in its functional and native form. The aberrant folding of a specific protein or the failure to adopt and/or maintain its native conformation is attributed to protein misfolding.[28] Several pathological conditions arising from protein misfolding are specified as conformational or misfolding diseases.[171] Molecular chaperones are responsible for directing the proper and efficient folding of proteins and therefore also preventing their aggregation.[172,173] Ineffective or low levels of chaperone proteins that can result from aging leads to protein misfolding. Environmental factors including pH, temperature, oxidation, and glycation can promote misfolding of the proteins as well.[174] Moreover, genetic factors that can affect amino acid composition and gene mutations also influence protein misfolding.[175,176]

The accumulation of misfolded proteins leads to the formation of protein aggregates. These aggregates can be cytotoxic to various types of cells through the loss of the functional structure of the native protein and/or their interaction with other components in the cell. The generation of these aggregates can lead to the production of either disordered amorphous aggregates or highly organized fibrils.[174] The highly ordered and insoluble fibrils are referred to as amyloid fibrils or plaques[28] when they accumulate in the extracellular space, as opposed to their intracellular counterparts which are referred to as intracellular inclusions.[177]

Figure 5.2 illustrates a general proposed mechanism of amyloid formation.[178] Assisted by folding enzymes and molecular chaperones, the synthesized protein folds in the lumen of the endoplasmic reticulum (ER). The properly folded protein is liberated from the cell into the extracellular space. The functional and native protein (N) is converted to a partially folded intermediate (I) or completely unfolded (U) protein.[178] During this process several α-helical structures unfold and β-sheets accumulate. The

formation of unfolded protein from the native protein is not thermodynamically favored and this intermediate has a high propensity to aggregate. Protein oligomerization through the intermolecular interaction of β-sheets is a step towards the formation of aggregates.[174]

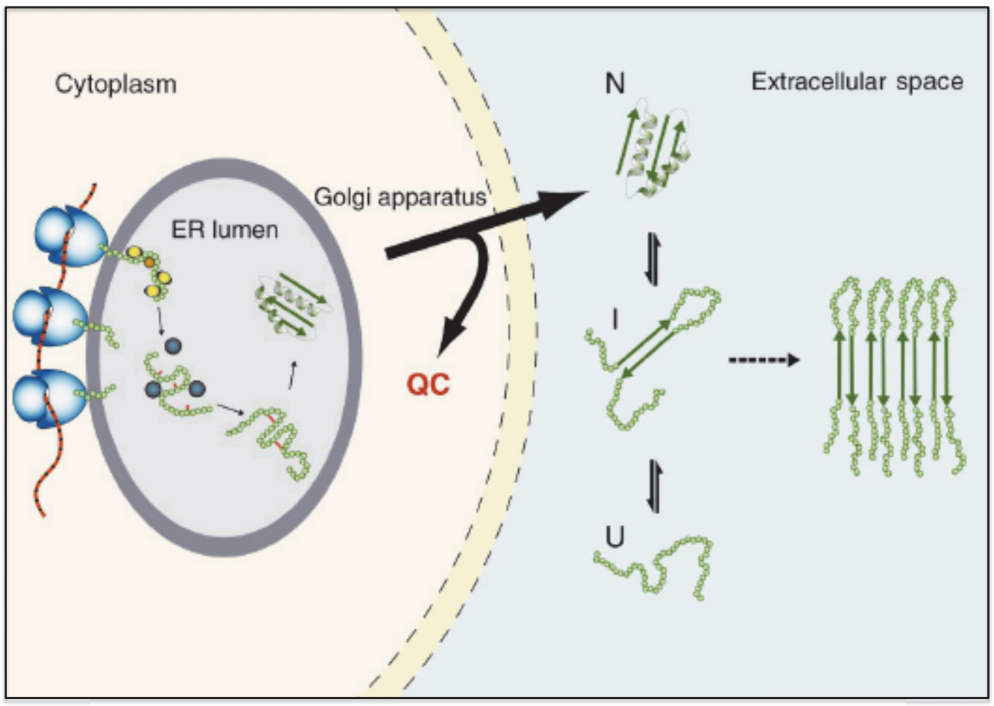


Figure 5.2: General representation of protein misfolding and aggregation. Abbreviations: ER: endoplasmic reticulum; N: native protein; I: partially folded intermediate; U: unfolded protein; QC: quality control factors which hinder protein misfolding (Figure from reference[178])

Since under denaturing conditions several non-disease related proteins are able to form the amyloids *in vitro*, it is known that amyloid fibril formation is an intrinsic property of many proteins;[179–181] however several genetic and environmental factors, such as those outlined earlier, as well as the protein's charge and hydrophobicity, which is affected by the identity of the side chain residues, promote the formation of amyloid fibrils. Amyloid fibril formation is related to about 20 progressively degenerative diseases.[182] Although all amyloid fibrils are similar in their characteristic cross-β-sheet structures, the specific protein misfolding/aggregation disorders caused depend

on the identity of the misassembling protein and the tissue subjected to protein deposition and cellular degeneration.[183,184] For example, the fibrils in Alzheimer's disease, Parkinson's disease, spongiform encephalopathy disease (mad cow disease), systemic amyloidosis and Huntington's disease are generated from problems with the folding of amyloid β-peptide (Aβ), α-synuclein, prion protein (PrP), lysozyme or transthyretin and huntingtin, respectively.[182,185,186] Diabetes mellitus type 2 (DM-2 or type II diabetes) is the target of this study and involves islet amyloid polyeptide (IAPP) for the formation of fibril aggregates.

5.2.1 Islet Amyloid Polypeptide (IAPP)

IAPP, also known as amylin, is a 37-amino acid, C-terminally α-amidated peptide that is co-secreted with insulin by the pancreatic β-cells. IAPP is an unusual aggregation-prone peptidic hormone that readily forms amyloid fibrils.[187] The IAPP amyloidogenic process observed in the pancreas is believed to accelerate type II diabetes pathogenesis by exacerbating β-cell degeneration, ultimately compromising insulin secretion. In patients afflicted by type II diabetes, the islet amyloid polypeptide (IAPP) deposits in the pancreas, leading to the degeneration of the islets of Langerhans.[187]

Previous studies have proposed a model for the formation of the fibrillar aggregates.[188–191] The model relies on the nucleation polymerization mechanism and contains two phases. The initial phase is the lag phase with the formation of unfolded peptides from the soluble monomeric peptides as illustrated schematically in Figure 5.3. A fragment of the unfolded peptide aggregate acts as a nucleation site to provide a template for the formation of other oligomeric fibrillar intermediates (protofibrils). The lag phase is followed by the elongation phase that leads to the rapid growth of the protofibril species. Subsequently the mature and insoluble fibrils are formed as the plateau phase is reached.

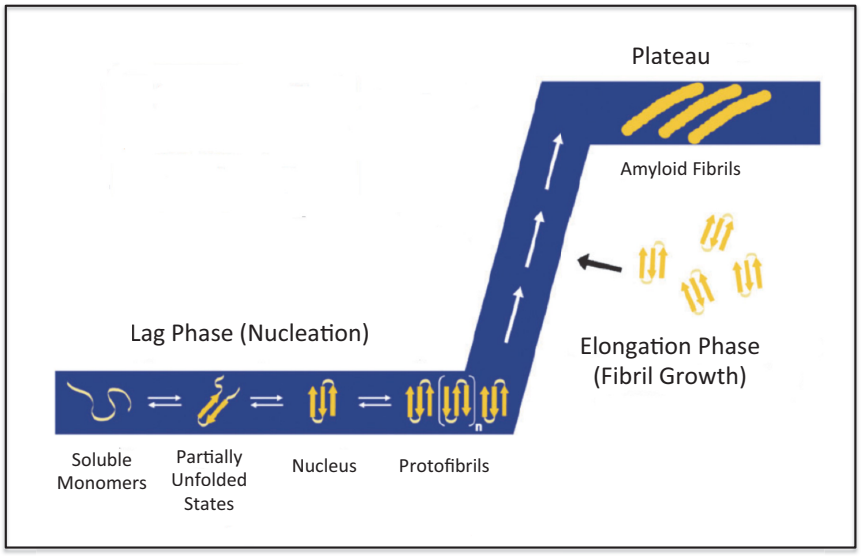


Figure 5.3: General representation of amyloid aggregate formation. (Figure from reference[188])

5.2.2 General Therapeutic Approaches to Prevent Protein Aggregation

Similar to many other disease targets, the formation of amyloid fibril aggregates can be prevented at several stages. These approaches have been presented by Soto *et al.*[171] and are illustrated in Figure 5.4. The first step towards protein aggregation is protein unfolding and misfolding, making it one of the essential target stages to be inhibited. This can be achieved through stabilization of the protein conformation in the native state (Figure 5.4, **a**). The application of synthetic small molecules in Alzheimer's disease[192] or protein mutation in transthyretin amyloidosis[193] are examples of stabilizing the native conformation of protein in order to prevent protein misfolding.

The second strategy is the employment of compounds, mainly small peptides, in order to inhibit unfolding of and/or re-constitute the native form of protein (Figure 5.4, **b**). Peptides that specifically interact with the protein and either stabilize the native conformation or destabilize the unfolded protein were shown to inhibit the

aggregation of the prion protein and amyloid beta.[194–196]

Inhibitors can also be used to compete with the protein-protein interactions involved in the misfolding process. These competitive inhibitors discourage protein aggregation through the interaction with either protein monomers or oligomers (Figure 5.4, **c**, **d**). These inhibitors can be small molecules that interact with the nuclei to prevent their growth towards the formation of oligomeric fibrils. These compounds also can interact with the already formed oligomers and interfere with the intermolecular forces, either to other oligomers or to other proteins, that favor the formation of aggregates. An example involves blocking a key protein (protein X) for prion protein propagation to prevent its interaction with the prion protein.[197,198]

Finally, several approaches have been proposed to enhance the removal of amyloid aggregates or misfolded proteins. For example, elimination of the interaction of some accessory components with the amyloid aggregates decreases the stability of the plaques and their further accumulation and insolubility. Anionic sulphates or sulphonates have been demonstrated to prevent the stabilization and improve the clearance of amyloid aggregates in Alzheimer's disease.[198,199]

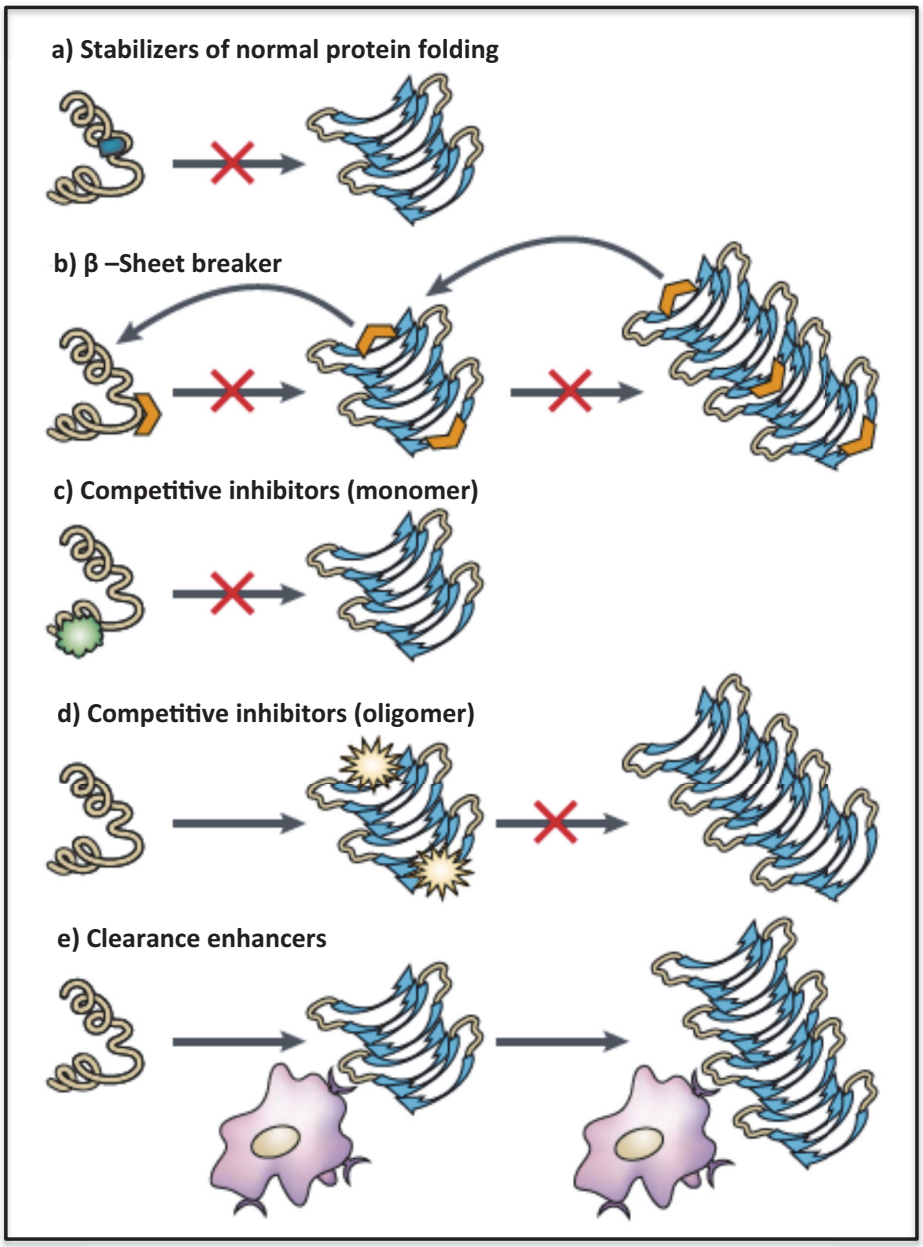


Figure 5.4: General representation of therapeutic approaches towards preventing protein misfolding and aggregate formation. (Figure from reference [171])

5.2.3 Approaches Towards Inhibition of Islet Amyloid Fibril Formation

Peptides as IAPP Receptor Agonists and IAPP Fibril Formation Inhibitors

Since IAPP is co-secreted with insulin and its deficiency is related to the emergence of type II diabetes, theoretically it is expected that the employment of synthesized IAPP peptide would compensate the deficiency of natural IAPP hormone in body. However, the high propensity of IAPP to aggregate limits its use in the treatment of type II diabetes. Therefore other analogs with low or no aggregation properties have been developed as IAPP receptor agonists. Pramlintide (Symlin) has been approved as a drug for both diabetes type I and II as an IAPP replacement for IAPP hormone deficiency (Figure 5.5, **b**).[200] Replacement of the residues of the IAPP sequence involved in the cross β-sheet structure of aggregates with proline provides a non-aggregating and soluble polypeptide hormone with similar properties to the IAPP hormone. The replacement of the three amino acid residues at positions 25, 28 and 29 of human IAPP with proline residues was inspired by the primary sequence of rat IAPP, which does not aggregate like human IAPP.

Over the last two decades, several compounds have been reported to inhibit IAPP aggregation *in vitro* by interfering with the later stages of fibrillogenesis through destabilization of the ordered cross β-sheet quaternary structure of the amyloid fibrils.[201,202] One example of these compounds is generated from modification of the natural sequence of IAPP through N-methylation of the two residues at positions 24 and 26 (Figure 5.5, **c**).[202] This polypeptide has improved solubility and lowered cytotoxicity compared to the IAPP as a result of these residue modifications.

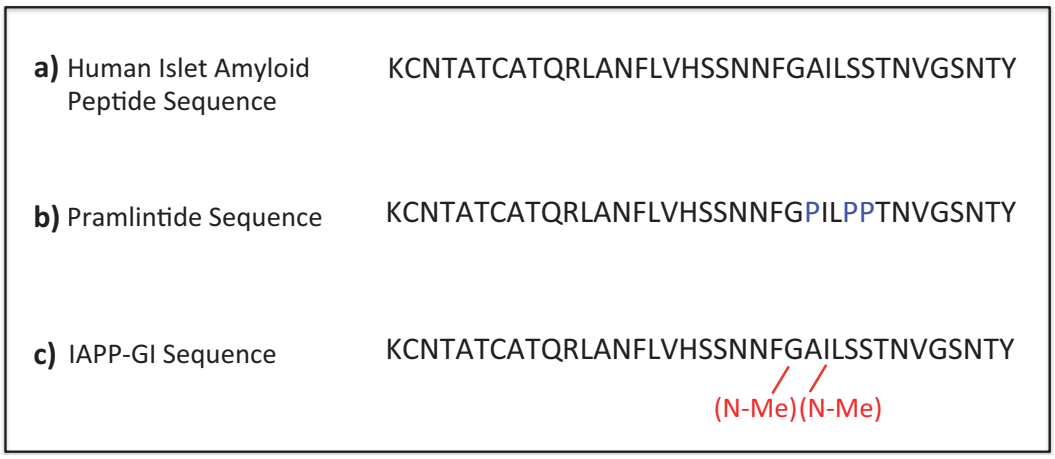


Figure 5.5: (**a**): Primary structure of IAPP (amylin); (**b**): An IAPP receptor agonist;[200] (**c**): An IAPP aggregation inhibitor[202]

Small-Molecule Inhibitors of IAPP Fibril Formation

Some of the obstacles of using peptide molecules as therapeutic drugs are poor bioavailability, low absorption through membranes and low stability in the biological circulatory system.[137] Although large molecules can be dosed through intravenous injection, oral dosing still remains the preferred route of administration for patients.

To overcome the drawback of peptides as therapeutic agents, small-molecule inhibitors have emerged. However, despite advances in the development of small-molecule inhibitors of fibril formation of the relevant proteins for some neurodegenerative diseases, the development of small-molecule inhibitors of islet amyloid fibril formation has seen less progress. The observed inhibition of amyloid fibril formation by small molecules such as the dye Congo red initiated the prospect of developing small-molecule inhibitors for IAPP fibrils.[203–205] For example, the co-crystallization of Congo red and insulin demonstrated an interaction between aromatic residues of Congo red and phenylalanine 24 of insulin that inhibited the formation of insulin fibrils.[206,207] Based on this observation, several natural and synthetic polyphenols were evaluated for the

inhibition of IAPP fibril formation and a synthesized phenolsulfonphthalein exhibited an IC_{50} of 1μM.[208] In a separate study some substituted rhodanine heterocyclics that were characterized as tau aggregation inhibitors[209] were evaluated against IAPP fibril inhibition. Two derivatives of this class of compounds inhibited IAPP fibril formation (Figure 5.6, **a**, **b**).[210] Although the exact mechanism of this inhibition is not clear, π-stacking interactions of the aromatic compounds with the monomer or oligomer of IAPP, preventing the formation of the aggregated fibrils, constitutes a reasonable proposal.[210]

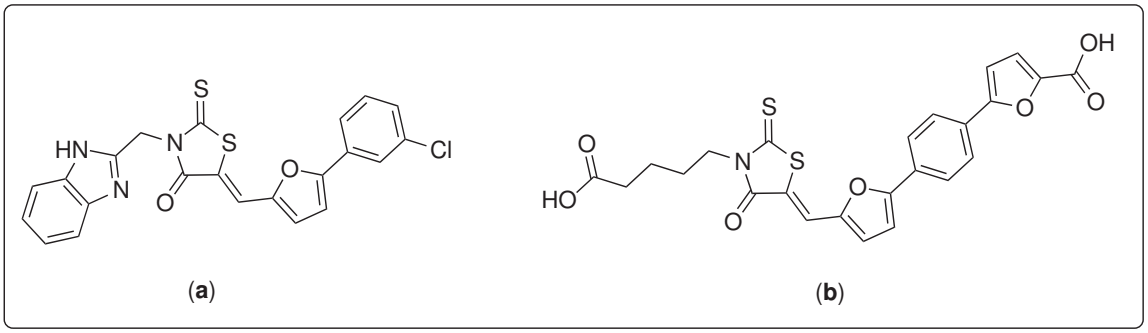


Figure 5.6: Chemical structure of rhodanine heterocyclic core scaffolds. IC_{50} values for compounds (**a**) and (**b**) were obtained as 1.23 and 0.45 μM respectively[210]

Interfering with the later stages of fibrillogenesis does not prevent the formation of the prefibrillar oligomers that were recently recognized as the most cytotoxic proteospecies of the amyloidogenic cascade,[211–214] suggesting a potential drawback to many of these therapeutic approaches.

5.2.4 Small-Molecule Mimetics of the α-Helices

As the most encountered secondary structure of proteins ($\sim$40%), α-helices are involved in several biological functions of proteins. α-Helices mediate protein-protein interactions (PPIs) that are important in cellular processes and regulatory pathways

and their aberrant function leads to the development of many diseases. Examples are PPIs involved in HIV fusion[215] or the mis-regulation of anti-apoptosis proteins such as Bcl-2 (B-cell lymphoma 2) that leads to tumor induction and several cancer diseases.[216] Consequently, targeting α-helices is arising as an attractive strategy in the development of therapeutic approaches and drug discovery. Developing the ligands that can chemically and spatially interact with the partial α-helix state in order to prevent some protein-protein interactions or to restrict the conformational ensemble of a specific protein into non-aggregating helical conformations has become a promising alternative strategy.[217,218]

Stabilizing α-helices of proteins has been accomplished through various approaches such as covalent linkages (disulfide, lactam, olefin, azobenzene)[219–222] and non-covalent interactions (π-π interactions, salt bridges, cation-π interactions, metal chelates,)[223–226] using peptidic molecules (Figure 5.7).

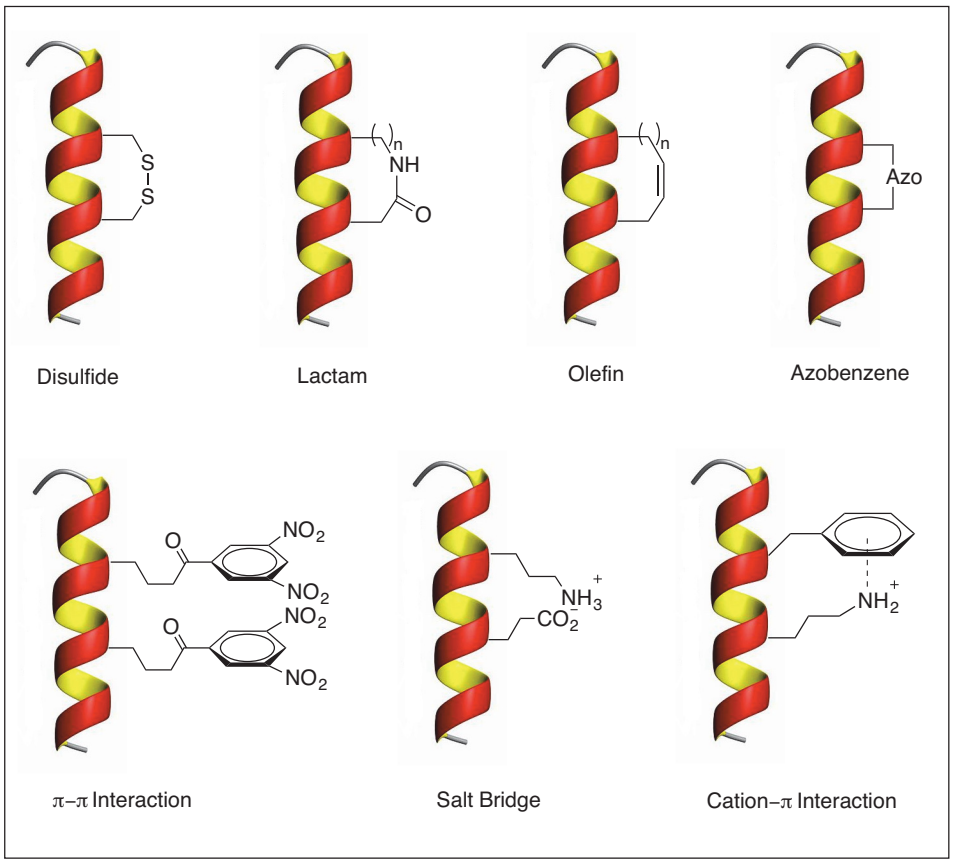


Figure 5.7: Approaches towards stabilization of helical state of proteins. Covalent linkages: disulfide, lactam, olefin, azobenzene; Non-covalent linkages: π-π interaction, salt bridge, cation-π interaction;[218] Helical ribbon structure was obtained from `http://en.citizendium.org/wiki/Protein_structure`

Although designing non-peptidic small molecules to interact with the large and usually buried interface area of proteins[227] is challenging, strategies have been developed to mimic the structural and conformational characters of protein secondary structures. Trisubstituted indanes (Figure 5.8) were one of the early achievements in mimicking partial areas (i-1, i, i+1 residues) of the α-helix as a proof of concept for interacting with neuropeptide receptors (tachykinin receptors)[228,229] to inhibit protein-protein interactions. In later attempts, small molecules capable of interacting with and stabilizing larger areas of the α-helices were designed. Functionalized terphenyls[217,230,231] (Figure 5.8) represent one such scaffold with a reported application

as a mimic of the i, i+4 and i+7 positions of the α-helix side chain of smooth muscle myosin light chain kinase (smMLCK) to disrupt its interaction with calmodulin (CaM) and consequently to prevent the hydrolysis of 3'-5'-cyclic nucleotide phosphodiesterase (PDE) enzyme.[232]

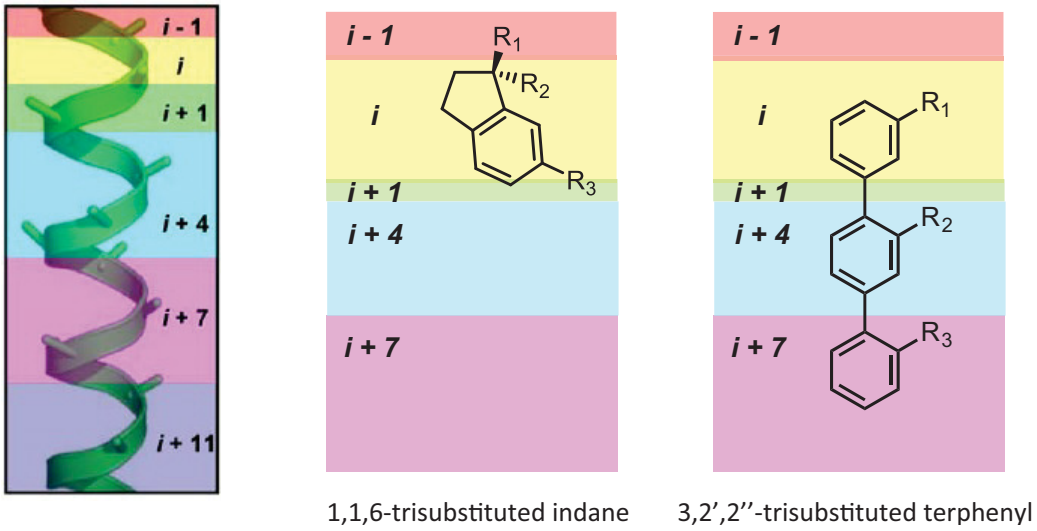


Figure 5.8: Early examples of small-molecule mimetics of α-helices

Alternatively, other synthetic templates have been shown to be attractive scaffolds for interacting/stabilizing the α-helical region of proteins. Some drawbacks of the terphenyl compounds such as long synthetic pathways and low aqueous solubility (high log P, the octanol/water partition coefficient) were improved by substituting the phenyl rings with heterocyclic scaffolds. Terpyridine[233] and pyridazine[234] α-helix mimetics were developed through substitution of the phenyl rings with the heterocyclic scaffolds. These compounds showed improved solubility compared to the terphenyl scaffolds, but were not as potent for inhibiting the interaction of Bcl-xL (B-cell lymphoma-extra large) and Bak peptides (Figure 5.9, **a** and **b**). A series of imidazole-phenyl-thiazole compounds demonstrated improved solubility compared to the terphenyl compounds, and were used for inhibition of the interaction between Dbs (a guanine nucleotide ex-

change factor) and Cdc42 (cell division cycle 42) proteins. The compound shown below in Figure 5.9, **c**, inhibits the Dbs/Cdc42 interaction with an IC_{50} value of 67 μM.[235] The Miranker and Hamilton groups recently developed small molecules targeting the transient helical state of IAPP in order to inhibit lipid-catalyzed aggregation.[236,237] These polycarboxylate ligands were developed on pyridyl (Figure 5.9, **d**), quinoline or peptoid scaffolds and were shown to inhibit lipid-induced IAPP aggregation, but to strongly accelerate IAPP fibrillogenesis under lipid-free conditions.

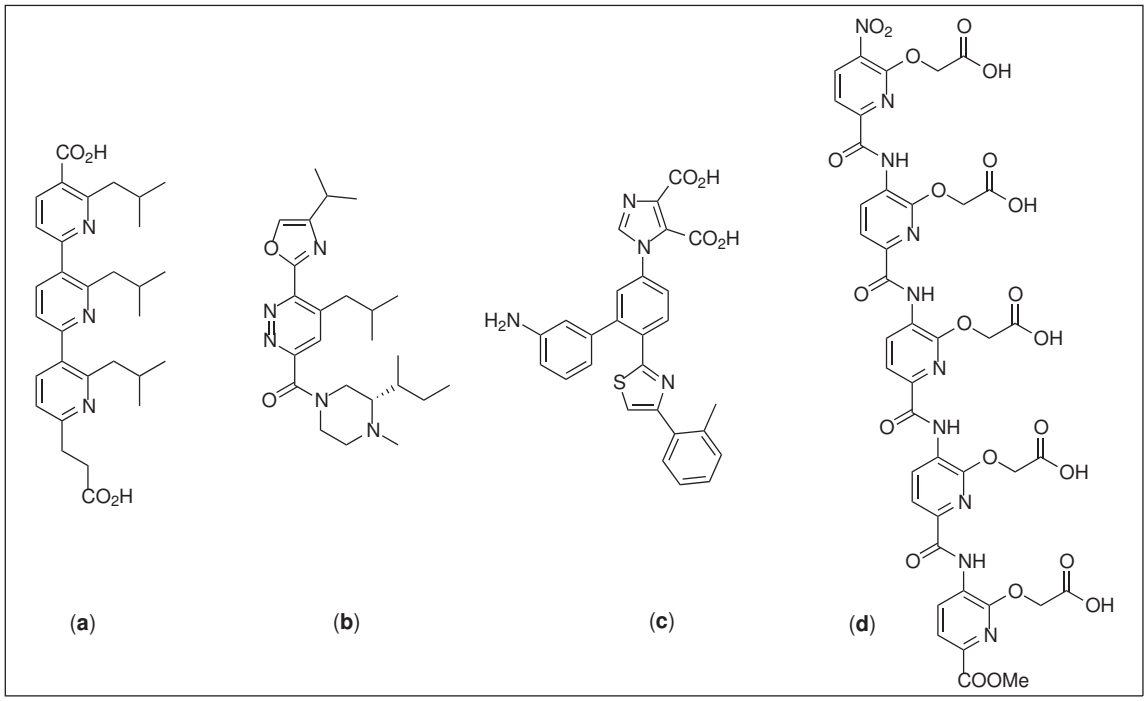


Figure 5.9: Small-molecule mimetics of the α-helices. (**a**): Functionalized terpyridine to mimic Val74, Leu78, Ile81 and Ile85 of the α-helix of Bak protein;[233] (**b**): Oxazole-pyridazine-piperazine scaffold to mimic Val74, Leu78, Ile81, and Ile85 of the α-helix of Bak protein;[234] (**c**): Imidazole-phenyl-thiazole scaffold to mimic Gln770, Lys774 and Leu777 of Dbs protein;[235] (**d**) Pentapyridyl scaffold to inhibit lipid-catalyzed aggregation of IAPP[236,237]

5.3 Palladium-Catalyzed Cross-Coupling Reactions

Small-molecule mimetics of α-helices have been synthesized through various approaches. During the pioneering work in the development of α-helix mimetics, the main objective was obtaining an active compound capable of interacting with the α-helix of the target protein, with less consideration given to the modularity of the synthetic pathways. As the field has matured and the amount of research with α-helix mimetics has grown, the synthetic accessibility of these mimetic compounds has gained importance to allow for their rapid synthesis for exploring structure-activity relationships to optimize various parameters, including potency.

Among chemical transformations, carbon-carbon bond formation through transition metal-catalyzed reactions stands out as one of the most valuable synthetic transformations. In a survey of 1039 transformations used for the synthesis of 128 potential drug molecules in 2006, 11% of the total reactions performed by three companies (AstraZeneca, GlaxoSmithKline, Pfizer) were carbon-carbon bond forming reactions.[238] Palladium-catalyzed coupling reactions emerged as a versatile method for the formation of C–C bonds, accounting for 22% of these carbon-carbon bond forming reactions.[238] Similar findings were presented in another study, which stated that of 7315 reactions reported in 139 articles in three journals, 11.5% were carbon-carbon bond forming reactions, and 62.3% of these were catalyzed by palladium.[239]

5.3.1 Classical Palladium Catalyzed Cross-Coupling Reactions

Generally palladium-catalyzed coupling reactions are classified into two types: classical and modern palladium cross-coupling reactions. Palladium-mediated coupling reactions between olefins and aryl- or alkylmercuric halides were pioneered by Heck in 1968 using stoichiometric amounts of palladium catalyst.[240,241] Improved protocols were later developed by Heck, Mizoroki and Fitton through a key discovery of replacing

organomercury compound by an aryl halide as well as altering this reaction into a catalytic process.[242–247] The traditional Suzuki and Negishi coupling reactions employ arylboronic acids/esters and organozincs respectively and a wide range of aryl halides and, along with the Heck reaction, earned the 2010 Nobel Prize in Chemistry for their discoveries (Figure 5.10).[248]

Other useful classical palladium-mediated reactions employ organotins (Stille), organosilanes (Hiyama) and Grignard reagents (Kumada) (Figure 5.10). The Sonogashira coupling is a mild palladium-catalyzed reaction for the formation of sp-sp2 carbon-carbon bond between alkynes and aryl halides, but also requires a copper co-catalyst. Finally, in the Tsuji-Trost coupling, carbon-carbon bond formation takes place between nucleophiles and compounds bearing an allylic leaving group (Figure 5.10).

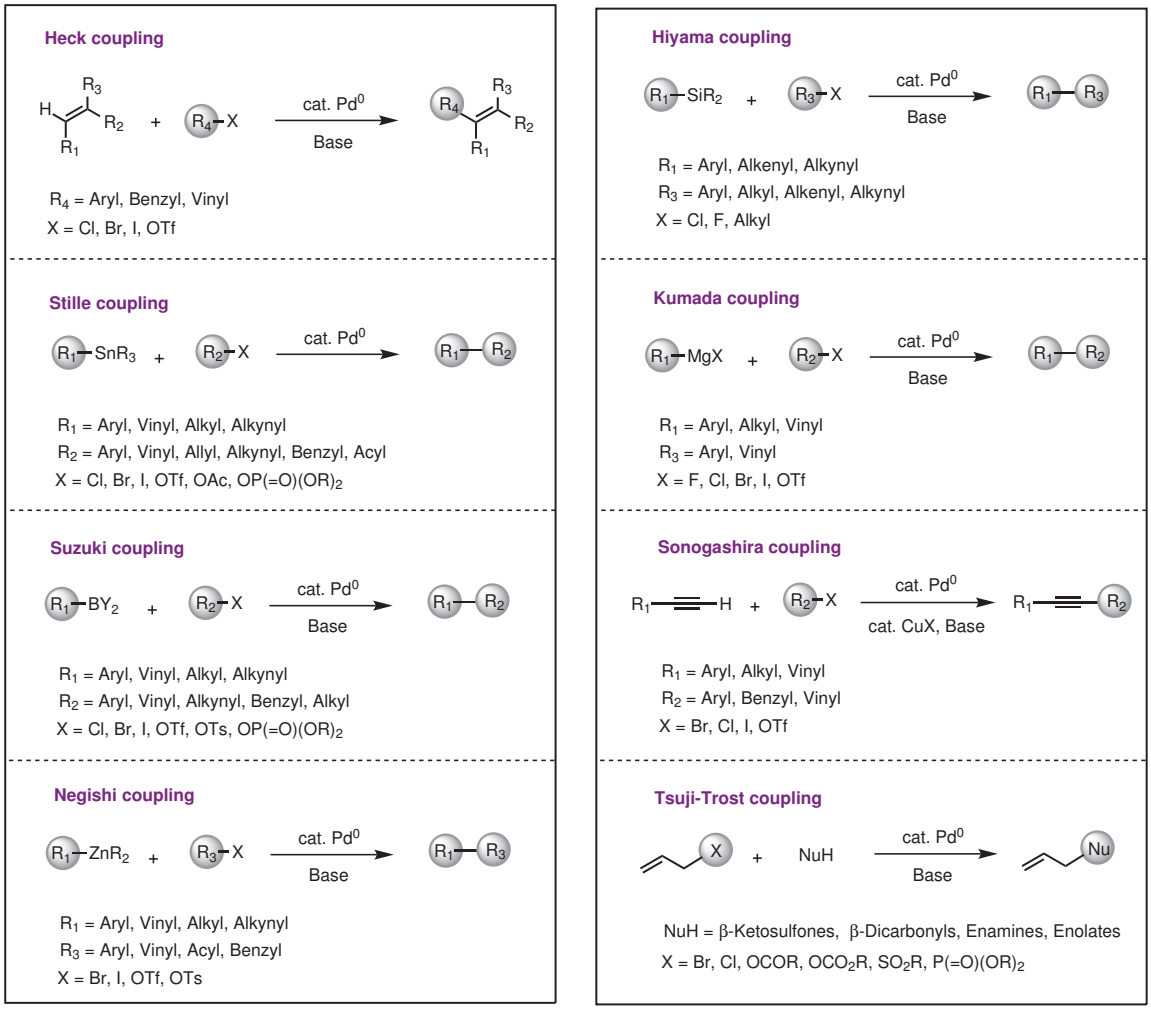
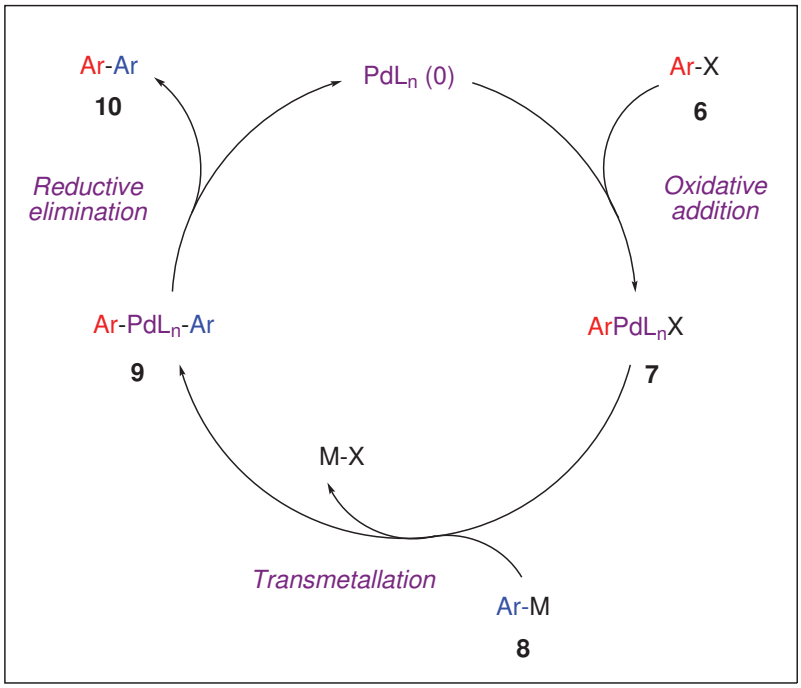


Figure 5.10: Most utilized traditional palladium-catalyzed cross-coupling reactions

The general catalytic cycle of cross-coupling reactions is illustrated in Scheme 5.1. The first step of the catalytic cycle involves the oxidative addition of a palladium(0) complex to the aryl halide (**6**) to provide an aryl-substituted palladium(II) complex (**7**). The transmetallation of the organometallic coupling partner (**8**) with its aryl group with the Pd(II) species (**7**) forms the bisarylated palladium complex (**9**) and the metal salt byproduct. In the last step reductive elimination of biaryl (**10**) from complex (**9**) furnishes the final product and regenerates the Pd(0) catalyst (Figure 5.1). In this catalytic cycle Pd(II) pre-catalyst sources can also be used, however, they

need to be reduced to Pd(0) in situ using a number of different methods. Ligands such as phosphine ligands, solvents or various reagents can reduce Pd(II) to Pd(0).[249,250]



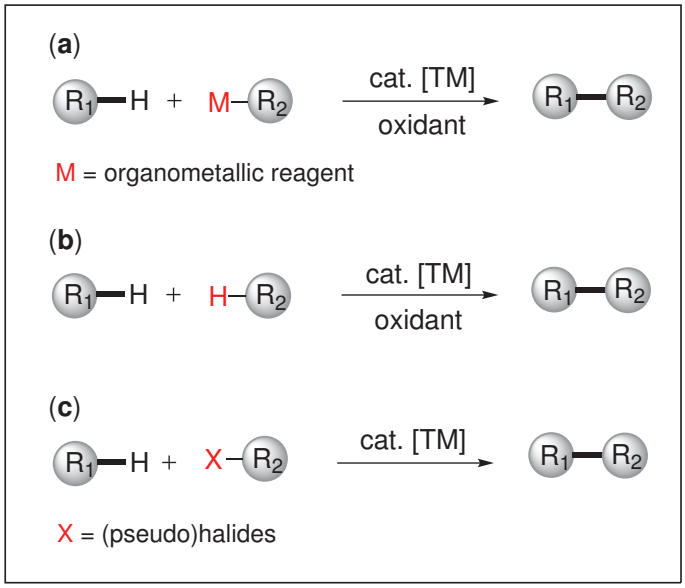
Scheme 5.1: General catalytic cycle of cross-coupling reactions

Although these traditional cross-coupling reactions are robust and well-established methods for the formation of carbon-carbon bonds in medicinal chemistry, materials science, total synthesis and industrial chemistry, they suffer from several drawbacks. For example some of the organometallic reagents require special precautions since they are either toxic (organotin reagents)[251] and/or sensitive to air (organotin, organozinc and Grignard reagents). Also, the generation of stoichiometric amounts of byproducts such as metal salts results in poor atom-economy. Moreover, to prepare the organometallic coupling reagents and functionalize them, several synthetic steps are required which is not favorable in terms of cost, energy consumption and waste production. Lastly, many organometallic coupling reagents can no be carried through other synthetic steps, or are not compatible with a number of other functional groups. To overcome

these limitations, methods have been developed to circumvent the requirement of stoichiometric amounts of an organometallic coupling partner while maintaining the efficiency and selectivity of the conventional coupling reactions.

5.3.2 C–H Arylations

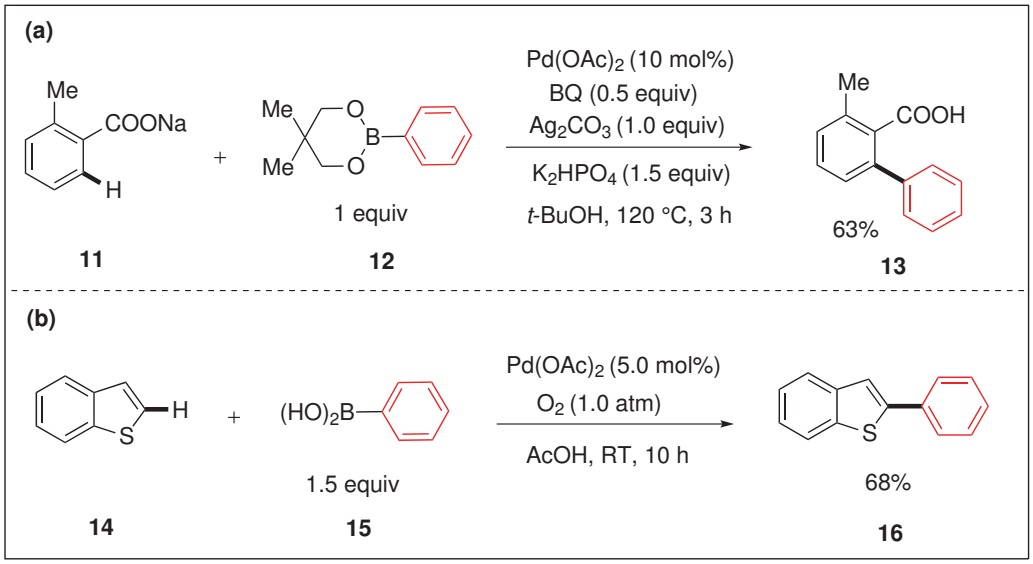
Direct C–H arylation coupling reactions were developed to address the previously outlined issues with conventional cross-coupling reactions. This transformation can take place through the reaction of an unactivated C–H bond via direct oxidative arylation employing organometallic (Scheme 5.2, **a**) or simple (hetero)arene coupling partners (Scheme 5.2, **b**) or a via direct arylation reaction using aryl (pseudo)halide coupling reagents (Scheme 5.2, **c**).[252–254]



Scheme 5.2: Classification of transition metal-catalyzed direct arylations of (hetero)arenes. R1 and R2: (hetero)arenes; TM: transition-metal catalyst

The oxidative arylation reaction with stoichiometric amounts of organometallic reagents was developed using various additives, and molecular oxygen, solvents or metal salts have served as oxidants in these reactions.[255] This method can be considered as an improved halogen-free version of conventional coupling reactions for the

formation of C–C bonds.[256–258] An example of a direct arylation of benzoic acids (**11**) with aryl boronates (**12**) developed by Yu *et al.* is shown in Scheme 5.3, (**a**).[259,260] A substantially improved oxidative arylation of a broad range of (hetero)arenes with aryl boronic acids (**15**) was accomplished under considerably milder conditions using oxygen as an oxidant instead of metal salts (Scheme 5.3, **b**).[261]

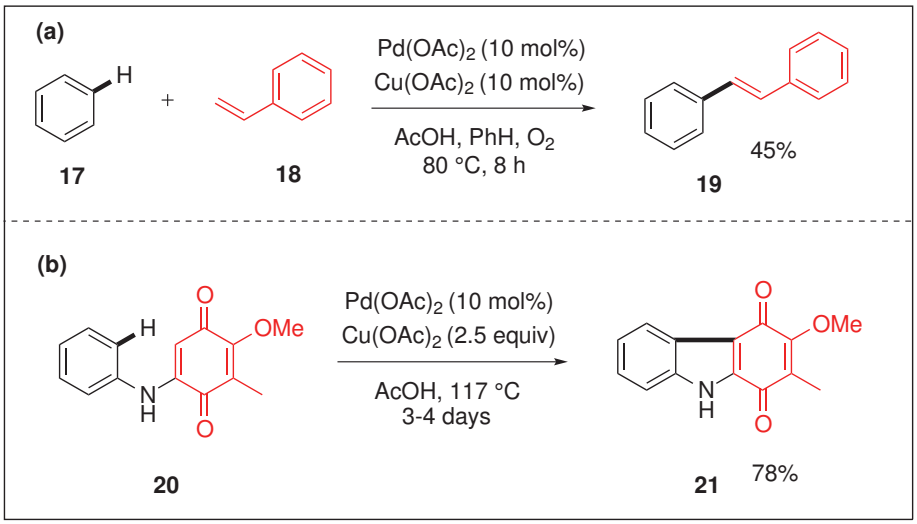


Scheme 5.3: Oxidative direct arylation of arenes and heteroarenes with organoboronic coupling partners

Subsequently, cross-coupling reactions have been developed employing other organometallic compounds such as organotin[262] and organomercury reagents. Although more advanced methods use less toxic coupling reagents such as organosilanes,[263,264] the main drawbacks of these methods remain the need for pre-activation of the coupling partners, the sensitivity of the organometallic reagents, which often can not be carried through synthetic steps, and the production of stoichiometric amounts of metallic waste.

Dehydrogenative arylations also utilize oxidants for the formation of C–C bonds via C–H functionalization; however, they take advantage of the reaction of two distinct

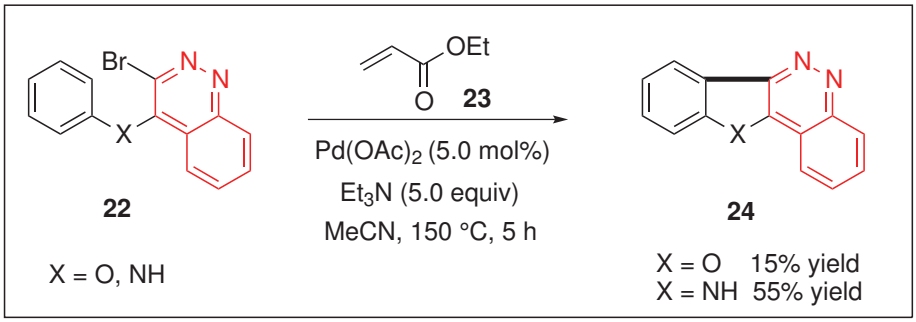
C–H bonds and eliminate the need for an organometallic coupling partner. This method was pioneered by Moritani and Fujiwara[265–267] for the direct arylation of olefins (Scheme 5.4, **a**). This method has also been expanded to intramolecular oxidative arylations of biphenyl compounds, which were particularly advantageous for the preparation of key intermediates towards the synthesis of naturally occurring compounds[268,269] (Scheme 5.4, **b**). The challenges of cross-coupling two simple arenes include the issues of regioselectivity (because of the presence of several non-symmetry related unactivated C–H bonds), chemoselectivity (both simple arenes can react with the catalyst at various parts of the catalytic cycle to generate homocoupled products as well as the cross-coupled product) and finally reactivity (both C-H bonds have relatively inert electronic properties).[270] A route to homocoupled products via a dehydrogenative arylation of functionalized (hetero)arenes has also been developed, with improved regioselectivity achieved through the installation of directing groups; however, stoichiometric amounts of oxidants were needed in these reactions.[271–274]



Scheme 5.4: Inter- (**a**) and intramolecular (**b**) dehydrogenative arylation reactions

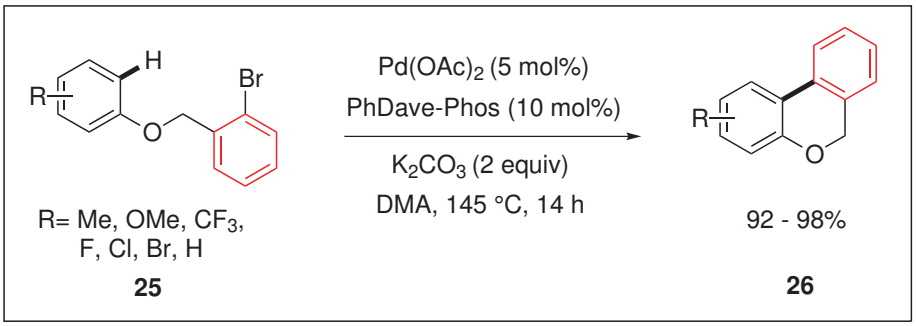
The first example of a direct arylation reaction using aryl halide coupling reagents was disclosed by Ames *et al.*[275–277] During an attempt to perform a Heck reaction

between aryl bromide (**22**) and alkene (**23**), product (**24**), resulting from the intramolecular cyclization of aryl bromide (**22**) via a direct arylation reaction, was obtained (Scheme 5.5). Further experiments revealed that alkene (**23**) was not involved in the reaction. The conditions were further optimized, and a new route to several related heterocycles was developed.



Scheme 5.5: An intramolecular direct arylation of simple arenes and aryl bromides by Ames *et al.*[275–277]

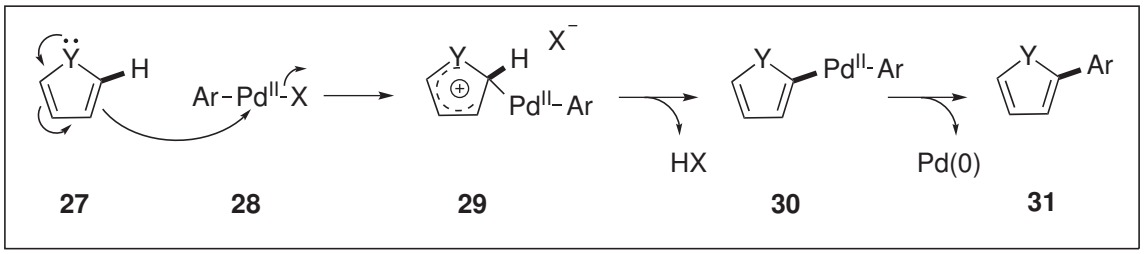
The intramolecular reaction of arene C–H bonds with aryl halides was greatly studied by Fagnou *et al.* and an early example involved an intramolecular direct arylation of the C–H bond of substituted benzenes (**25**) with aryl bromides/chlorides to form tricyclic biaryls (**26**) (Scheme 5.6).[278,279]



Scheme 5.6: An intramolecular synthesis of biaryls via direct arylation by Fagnou *et al.*[278]

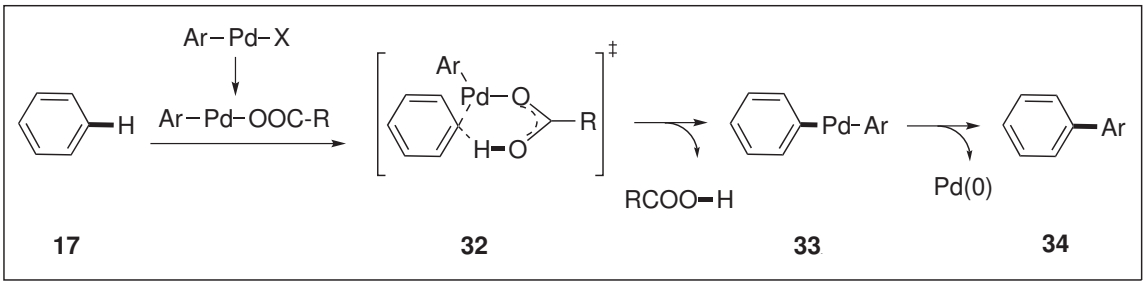
Commonly, C–H palladation is more favored with electron-rich (hetero)arenes and although debated, the proposed mechanism for many of these reactions involves elec-

trophilic aromatic substitution (S_EAr). For instance five membered heteroaromatics (**27**) undergo a C–H palladation reaction due to their high nucleophilicity (Scheme 5.7).



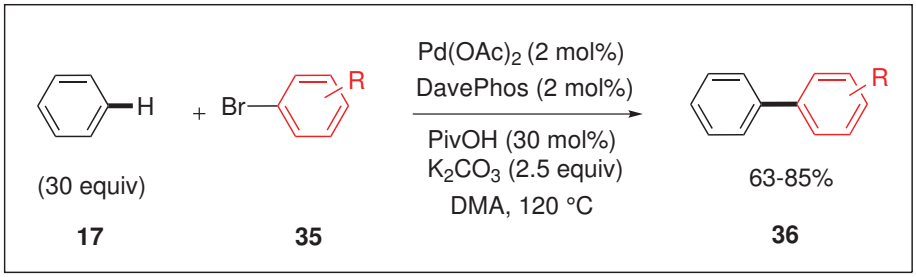
Scheme 5.7: General electrophilic aromatic substitution (S_EAr) for direct arylation mechanism of heteroarenes

Electrondeficient or electron-neutral aromatics are not prone to undergo direct arylation through the above electrophilic substitution mechanism.[280] Fagnou *et al.* established a complementary catalytic system for these compounds. In these reactions polyfluorinated biaryl compounds were synthesized through the arylation of polyfluoroaromatic compounds with various aryl halides employing a $Pd(OAc)_2$ and S-Phos ligand ligand catalyst system with K_2CO_3 as a base.[281] Kinetic isotope effect studies supported the postulate that these reactions proceed through a concerted metallation deprotonation (CMD) mechanism in which carbon-metal bond formation and carbon-hydrogen bond breakage take place simultaneously[282] (Scheme 5.8).



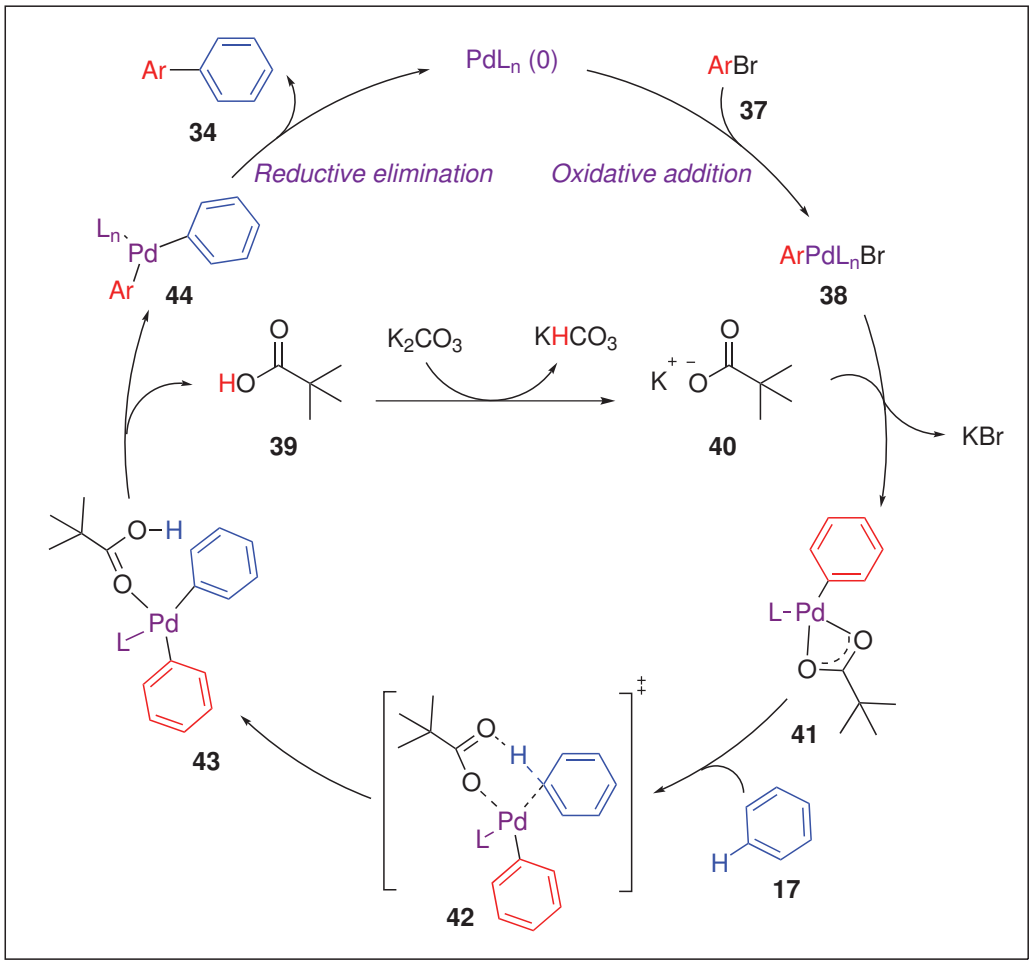
Scheme 5.8: General concerted metallation deprotonation (CMD) mechanism

Fagnou *et al.* also described the intermolecular direct arylation of electron-neutral benzene (**17**), used in excess, by aryl bromides (**35**) with a catalyst system consisting of $Pd(OAc)_2$, the DavePhos ligand and pivalic acid as an additive and K_2CO_3 as a base. The broad scope of this reaction was demonstrated by the successful reaction of a variety of aryl bromides with diverse electronic and steric properties to generate high yields of the biaryl products (**36**); however, the reaction was not efficient with aryl chloride and aryl iodide substrates.[283] Pivalic acid was shown to operate as a proton shuttle in the catalytic cycle.[283]



Scheme 5.9: Intermolecular direct arylation of unactivated benzene

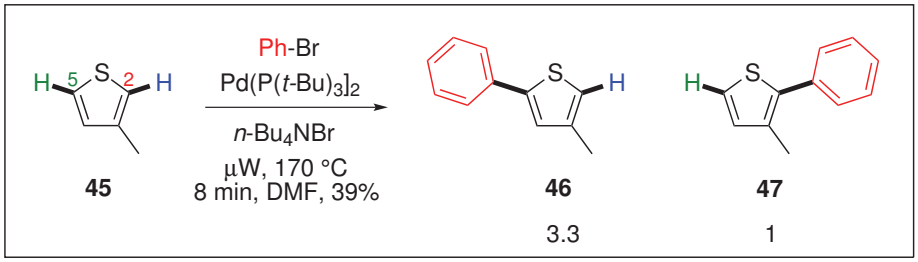
The onset of the catalytic cycle is the oxidative addition of the palladium(0) complex to the aryl bromide (**37**) to provide the aryl-substituted Pd(II) species (**38**) (Scheme 5.10). Deprotonation of pivalic acid (**39**) by the carbonate base forms the pivalate anion (**40**) that can undergo ligand exchange on complex (**38**) to generate KBr and species (**41**). Coordination of the benzene ring (**17**) and concerted proton transfer from the benzene and metallation occurs in the next step (transition state **42**).[284,285] Dissociation of the pivalic acid from (**43**) generates (**44**), followed by reductive elimination to form the biaryl product (**34**) and regenerate the Pd(0) catalyst.



Scheme 5.10: Catalytic cycle of palladium-catalyzed direct arylation of benzene

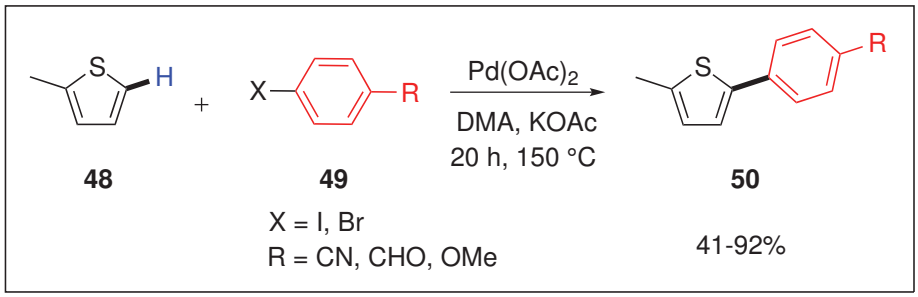
Despite many advances in C–H activation reactions, chemo- and regioselectivity continue to be a hurdle. For instance the C–H activation of substituted, unsymmetrical arenes is challenging because of the presence of non-equivalent hydrogen atoms in the molecule. Although these hydrogen atoms will have different acidities and reactivities, unless a reaction can be done exclusively with only one of them, mixtures of products will result. Likewise, similar considerations apply for unsymmetrically substituted heteroaromatics. For example, for 3-methylthiophene (**45**), while there is a sufficient difference in reactivity so that none of the C4 arylated product is formed, the reactivity of the protons at the C2 and C5 positions is similar enough that a mixture of regioisomers (**46** and **47**) is produced in the direct arylation reaction with aryl

bromide substrates[286] (Scheme 5.11).



Scheme 5.11: Regioselectivity in C–H arylation of 3-methylthiophene

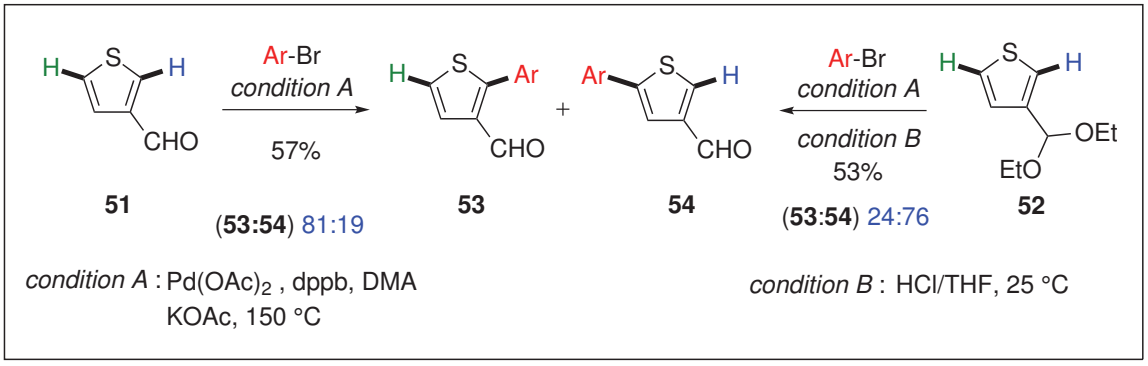
Various strategies have been employed to overcome these hurdles and provide some control over the regio- and chemoselectivity. The utilization of blocking groups at the position with competitive reactivity or the employment of steric bulk at one position have resulted in improved regioselectivity in direct arylation reactions. For example blocking one of the α-positions of a thiophene (C2 of **48**) with a methyl group results in the formation of only one regioisomer (**50**) in a direct arylation reaction with aryl halides[287] (Scheme 5.12). Other functional groups, such as acetyl, nitrile, n-butyl and sulfonyls at the C2 position of thiophenes and furans can be used for the same purpose and have been reported by Doucet *et al.*[287,288]



Scheme 5.12: Regioselectivity in C–H arylation of 2-methylthiophene

An example of the use of steric bulk to control regioselectivity is shown in Scheme 5.13. In these reactions, a direct arylation reaction with 3-thiophenecarboxaldehyde (**51**) produced a mixture of regioisomers (**53**) and (**54**), with a 4:1 preference for C2

arylation. When the C3 aldehyde was converted to the bulkier diethyl acetal group in compound (**52**), direct arylation at the less sterically encumbered C5 position was favored, giving a 1:3 ratio of products (**53**):(**54**) following deprotection of the acetal.[289]



Scheme 5.13: The effect of steric bulk on the C–H arylation of 3-substituted thiophene

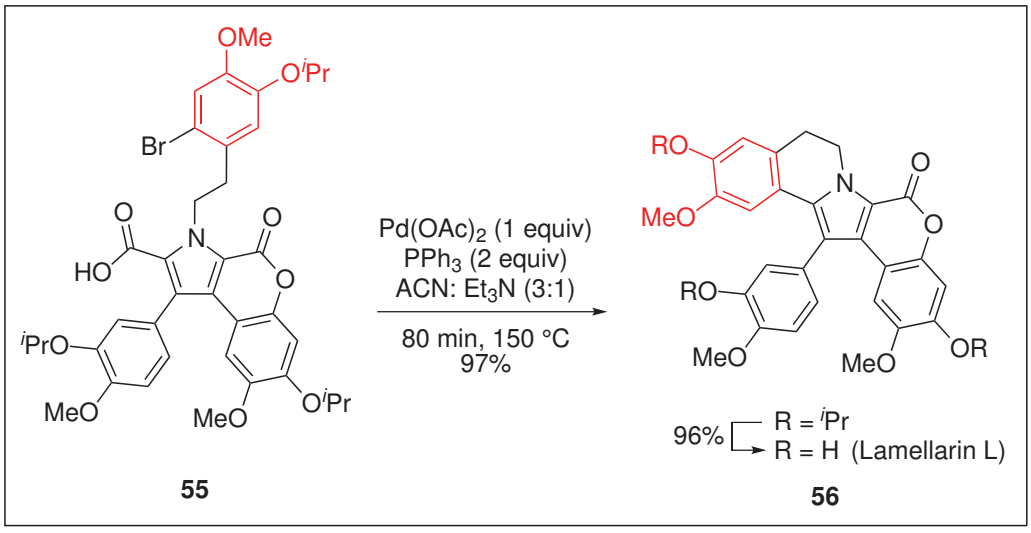
Although advances have been made in controlling the regio- and chemoselectivity of direct arylation reactions, examples with complete control are somewhat rare. These limitations highlight the need for alternative methods that show better selectivity yet still do not rely on organometallic coupling partners.

5.3.3 Decarboxylative Cross-Coupling Reactions

In light of the regioselectivity and chemoselectivity limitations of direct arylation reactions and the issues stemming from the need for organometallic coupling partners in classical palladium-catalyzed cross-coupling reactions, decarboxylative cross-coupling reactions are an attractive alternative. Decarboxylative cross-coupling reactions were developed based on the substitution of carboxylic acids for either the organometallic reagents of traditional cross-coupling reactions or the aryl halides of Heck-type couplings. Carboxylic acids and their related salts are stable to storage, readily available, and easy to handle, unlike some of their organometallic counterparts. The byproduct from a decarboxylative cross-coupling reaction is carbon dioxide gas, rather than the

metal salt byproducts formed from traditional cross-coupling reactions.

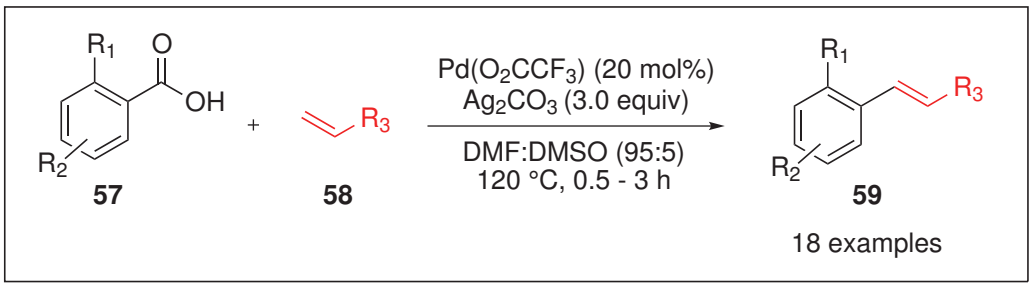
The first example of a decarboxylative coupling reaction was established by Nilsson *et al.* in 1966 with a super-stoichiometric copper-catalyzed reaction of iodoarenes and nitrobenzoic acids in low yield and limited substrate scope.[290] These reactions remained unstudied for about 30 years until Steglich *et al.* reported an intramolecular decarboxylative coupling towards the synthesis of the alkaloid heterocycle Lamellarin G[291] and later with higher yield and reduced reaction time for the synthesis of Lamellarin L[292] (Scheme 5.14, **56**). Although the reaction did not require a metal additive as a co-catalyst, a stoichiometric amount of palladium was employed.



Scheme 5.14: An early example of palladium-catalyzed decarboxylative cross-coupling reaction

A breakthrough towards palladium-catalyzed decarboxylative cross-couplings was reported by Myers *et al.* in a Heck-type reaction.[293] Decarboxylative coupling occurred between *ortho*-substituted carboxylic acids (**57**) and acrylates/styrenes (**58**) utilizing a stoichiometric amount of silver carbonate that served both as an oxidant and base. Heteroaromatic and electron-poor and -rich aromatic acids were tolerated in the reactions. The olefin substrate scope was expanded by the same group to 5-7

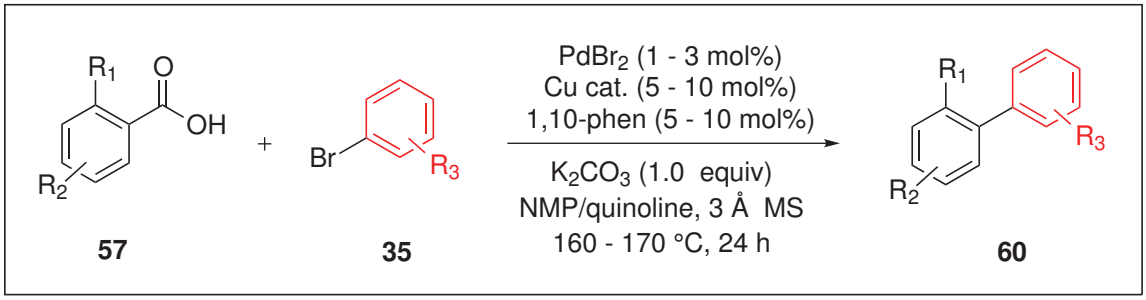
membered cyclic α,β-unsaturated ketones and *ortho*-substituted benzoic acids that are more challenging in traditional Heck-coupling reactions.[294]



Scheme 5.15: Intermolecular decarboxylative coupling of *ortho*-substituted benzoic acids and olefins by Myers *et al.*

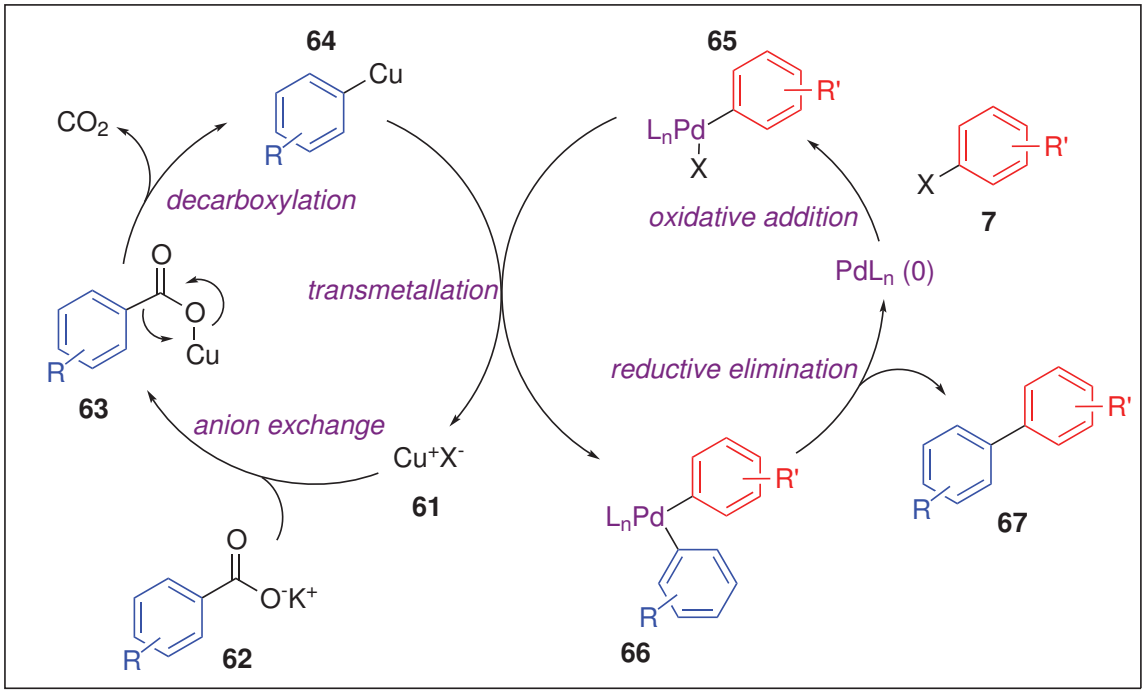
There are several successful protocols for decarboxylative cross-couplings differing in the substrate scope and utilization of mono- or bimetallic catalytic systems.

Gooßen *et al.* established the first intermolecular decarboxylative cross-coupling reaction of a range of *ortho*-substituted benzoic acids and aryl halides using a Pd/Cu catalytic system for the formation of biaryls.[295] The method was envisioned based on the limitations of Nilsson's protocol where super-stoichiometric amounts of copper were required and limited substrates were applicable.[290] The early example employed a stoichiometric amounts of copper (1.5 equiv) and catalytic palladium(II) acetylacetonate (2 mol%) for the decarboxylative coupling of *ortho*-nitrobenzoic acids with a range of aryl bromides at a reaction temperature of 120 °C.[295] Gooßen's group also developed modified reaction conditions that employed catalytic amounts of both palladium and copper (a more stable and less active copper complex) at temperatures[295] as shown in Scheme 5.16. The requirement for *ortho*-substituted benzoic acids still remained the limitation of this method. Further optimization expanded the reaction scope to other aryl coupling partners such as aryl chorides,[296] tosylates[297] and triflates.[298] The use of aryl triflate coupling partners circumvents the need for *ortho*-substitution of the benzoic acids.[298,299]



Scheme 5.16: Decarboxylative cross-coupling using copper co-catalyst by Gooßen *et al.*

The mechanism proposed by Gooßen for the bimetallic reactions discussed above is illustrated in Scheme 5.17. Initially copper carboxylate (**63**) is formed through anion exchange between copper salt (**61**) and carboxylate anion (**62**). This leads to the extrusion of CO_2 and the generation of organocuprate species (**64**). The other part of the catalytic cycle involves a typical palladium-catalyzed process where the oxidative addition of a Pd(0) complex into the aryl halide (**7**) forms intermediate (**65**). Through the coordination of the two catalytic cycles, transmetallation of the organocuprate species (**64**) with the aryl-substituted Pd(II) intermediate (**65**) generates the biarylated palladium(II) species (**66**). In the last step, reductive elimination of intermediate (**66**) results in the formation of cross-coupled biaryl product (**67**) and regenerates the Pd(0) catalyst.

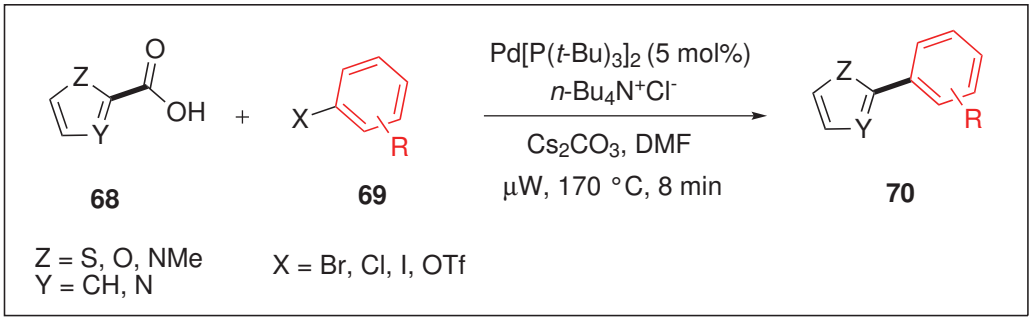


Scheme 5.17: Mechanism of co-catalyzed decarboxylative cross-coupling proposed by Gooßen

Another useful protocol was developed by Becht *et al.* where no copper catalyst was used but instead silver carbonates were applied. The decarboxylative cross-coupling occurred between electron-rich *ortho*-arene carboxylic acids and aryl iodides[300] or diaryliodonium salts.[301] Similar to Gooßen's original protocol, this method also required stoichiometric amounts of silver carbonate. It was proposed that the silver salt serves as a base and possibly as a co-catalyst through coordination to the carboxylic acid. Similar reactions were developed by Wu *et al.* through the use of a $PdCl_2$ catalyst and the BINAP ligand and stoichiometric amounts of silver carbonate.[302]

In the same year that Gooßen disclosed the bimetallic decarboxylative cross-coupling, Forgione *et al.* revealed a decarboxylative cross-coupling of aryl bromides and heteroaromatics while attempting direct arylations of heterocycles utilizing carboxylic acids as blocking groups at the C2 position of heteroaromatics.[286] The reaction

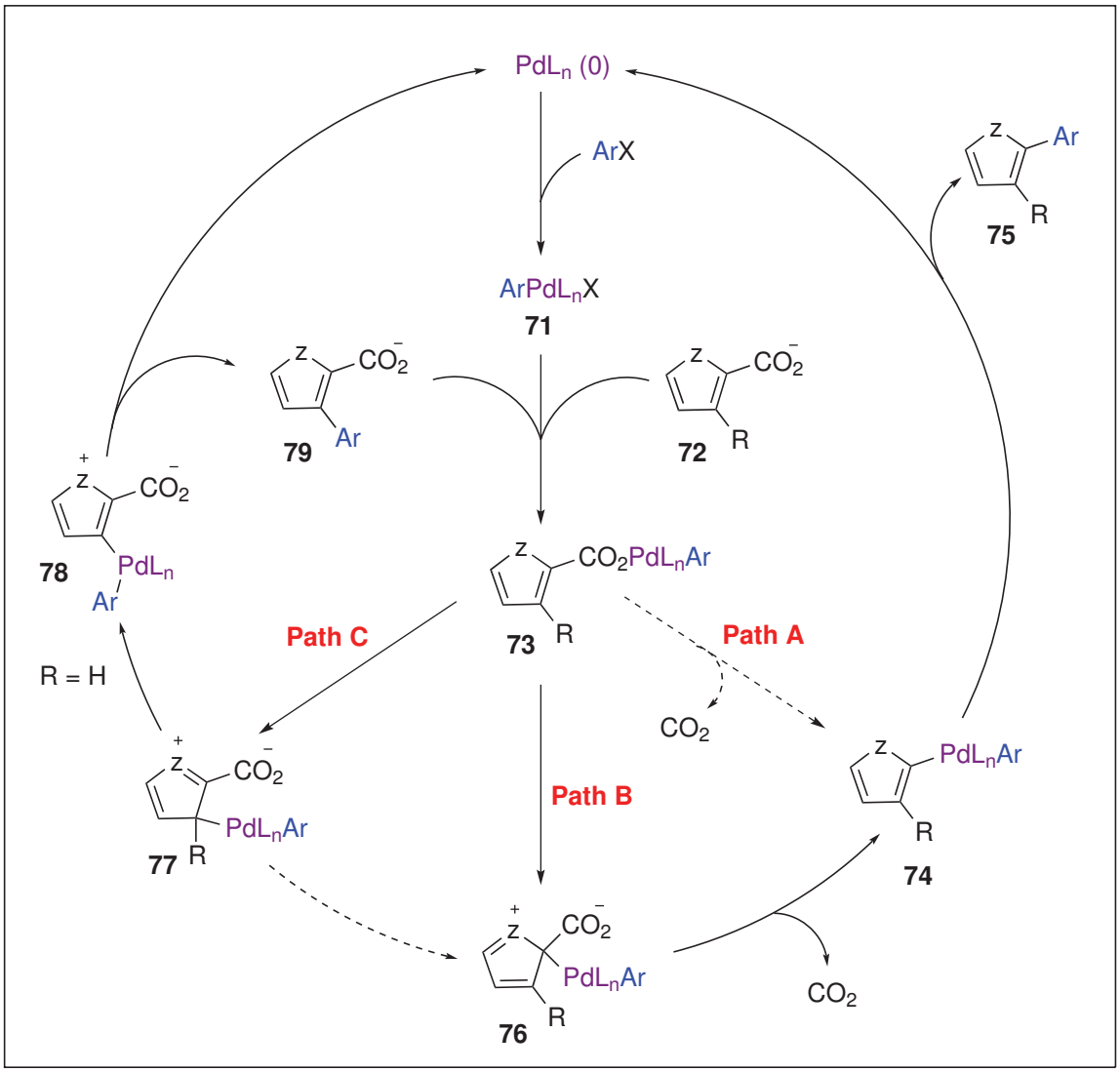
differed from previous decarboxylative cross-couplings in several aspects; only palladium was employed as the catalyst and tetrabutylammonium chloride was used presumably as a phase-transfer catalyst. One great advantage of this methodology involved the short reaction time (8 minutes) through microwave irradiation (Scheme 5.18). In a follow-up study by the same group, the substrate scope was expanded to aryl chlorides, iodides and triflates.[303]



Scheme 5.18: Decarboxylative cross-coupling of heteroaromatics using mono catalyst

The proposed mechanism (Scheme 5.19) starts with the typical oxidative addition of the palladium(0) complex into the aryl halide to provide the aryl palladium(II) intermediate (**71**). A ligand exchange forms intermediate (**73**) through displacement of the halide with heteroaryl carboxylate (**72**). Three mechanistic pathways were postulated from this step. In path A direct decarboxylation and extrusion of CO_2 result in generating C2 palladated species (**74**). This intermediate would undergo reductive elimination to form the desired heteroaryl product (**75**). Benzoic acids or heteroaromatic-3-carboxylic acids failed to cross-couple under these conditions and therefore path A would not be a plausible mechanism. Instead, the electron richness of the heteroaromatic and the carboxylate as the directing group in path B assist the electrophilic palladation of the heteroaromatic to form intermediate (**76**). To re-obtain the aromaticity of the heteroaromatic ring, CO_2 extrusion occurs to provide bisarylated palladium species (**74**) that can undergo reductive elimination to form the

desired product (**75**). The formation of trace amounts of C3 palladation by-product, if C3 of the heteroaromatic is hydrogen, suggested path C as a possible mechanism. Path C still utilizes the π-nucleophilicity of the heteroaromatic in which intermediate (**77**) is formed through C3 electrophilic palladation. A C3 to C2 migration forms the more stable intermediate (**76**) that goes through the same CO_2 extrusion and reductive elimination such as path A. In case, where R is hydrogen, a deprotonation at C3 provides the rearomatized complex (**78**) that undergoes reductive elimination to generate (**79**). This intermediate can re-renter the catalytic cycle to form the 2,3-biarylated compound. Several parameters such as base and catalyst were evaluated and the condition demonstrated on scheme 5.18 were utilized as the optimized reaction condition. In addition both electron-poor and -rich aryl halides were found to be tolerated.



Scheme 5.19: Mechanism of decarboxylative cross-coupling proposed by Forgione, Bilodeau *et al.*

Decarboxylative cross-coupling reactions were established as convenient and alternative methods for the formation of carbon-carbon bonds. Decarboxylative coupling circumvent the regioselectivity issues arised from some direct arylation reactions and take advantage of easily available carboxylic acids. However, these methods require extend optimizations to overcome their own limitations such as proto-decarboxylation.

Chapter 6

Synthesis of IAPP α-Helix Mimetics: Results and Discussion

6.1 Project Perspective

IAPP exhibits a conformational ensemble mainly populated by the disordered conformations in the non-aggregated soluble state, although it diverges from an absolute random coil by the presence of local and transient ordered structure.[304] Recent mechanistic studies have suggested that the pro-amyloidogenic peptide undergoes a random coil to α-helix conformational conversion during the initial phase of self-assembly where the helical intermediates could be on-pathway to amyloid formation.[304,305] According to this model, α-helix formation and self-association of helical segments are linked and accelerate self-assembly,[305] with similar driving forces to those of coiled coil motif formation. Consequently, the accelerated self-assembly generates a high local concentration of the amyloidogenic domain of IAPP (segment 20-29, Figure 6.1), which has a high propensity to adopt a β-structure, favoring the formation of cross-β-sheet assemblies en route to amyloid formation. Consistently, IAPP was shown to adopt a helical structure spanning approximately residues 8 to 19 (Figure 6.1) when

the peptide was bound to model membranes[306,307] or glycosaminoglycans (GAGs),[308] and the interactions accelerated the rate of IAPP amyloid fibril formation.

Figure 6.1: The primary sequence of IAPP representing the helix region and random coils

According to the helical intermediates hypothesis described above, an alternative strategy to control the formation of IAPP amyloid fibrils would be to design molecules that target and stabilize the transient helical segment 8-19 of IAPP, modulating the helix-assembly process. This approach could inhibit the formation of oligomeric and fibrillar aggregates by over-stabilizing the helical intermediates, not allowing the propagation of the β-sheet conformation from the 20-29 domain of IAPP. Recent studies using membrane models have shown that, indeed, IAPP can be trapped in a non-amyloid prone helical conformation.[305,307,309] Alternatively, as reported for lipids and GAGs, helical targeting ligands could potently accelerate the self-assembly of IAPP into β-sheet-rich amyloid fibrils by initially shifting the conformational equilibrium towards the α-helix, without overly stabilizing the helical motif. Considering that oligomers are the most potent cytotoxic proteospecies,[213] both pathways will decrease the toxicity induced by the amyloidogenic process of IAPP, either by blocking the formation of pre-fibrillar assemblies (α-helix over-stabilization) or by accelerating the structural conversion of oligomers into less cytotoxic amyloid fibrils.

In the present work, small molecules that are chemically and spatially paired with the residue side chains confined to one side of the helical conformer of IAPP have been developed. Consequently, the helical motif of IAPP can be stabilized or the conversion

of toxic oligomers into less toxic amyloid fibrils can be accelerated. Eventually, the synthesized scaffolds can be densely substituted with a wide range of functional groups through modular synthetic routes. In the current work, the molecules produced by our synthetic approach were tested as modulators of the formation of IAPP fibrils as a proof-of-concept. However, the general synthetic route can be used for the preparation of molecules tailored with different side-chain residues to stabilize and/or interact with the α-helix of other proteins for various applications.

6.2 Design of the Small-Molecule Mimetics

The synthesis of α-helical templates of proteins was established and widely applied by Hamilton *et al.* and has proven to be an attractive approach towards interacting/stabilizing the α-helices of proteins. Functionalized terphenyls[217,230,231] represent one such scaffold with a reported application as a mimic of the α-helix side chain of smooth muscle myosin light chain kinase (smMLCK) to disrupt its interaction with calmodulin (CaM) (Figure 6.2, **b**).[310]

Binding of ligands to the helical motif largely results from the interaction of the ligand with the amino acid side chains projecting on one face of the α-helix and spaced three or four residues away from each other, referred to as i, i+4 and i+7. In the transient helical conformation of IAPP comprising residues 8-19, residues Arg 11, Phe 15 and His 18 are oriented on one face of the α-helix[306] and represent the key motif that will be targeted to stabilize the transient α-helix of IAPP (Figure 6.2, **a**). As hypothesized from coiled coil motifs formation, the presence of hydrophilic side chains (Arg, His) will provide the specificity of the interaction whereas the hydrophobic residue Phe will contribute to the thermodynamic stability of the interaction by hydrophobic core packing.

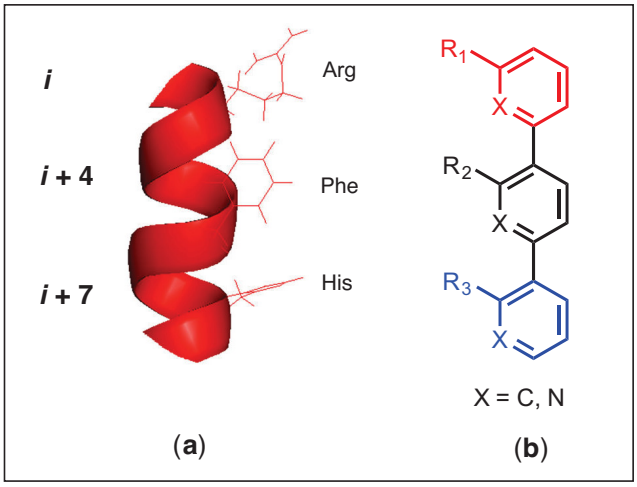
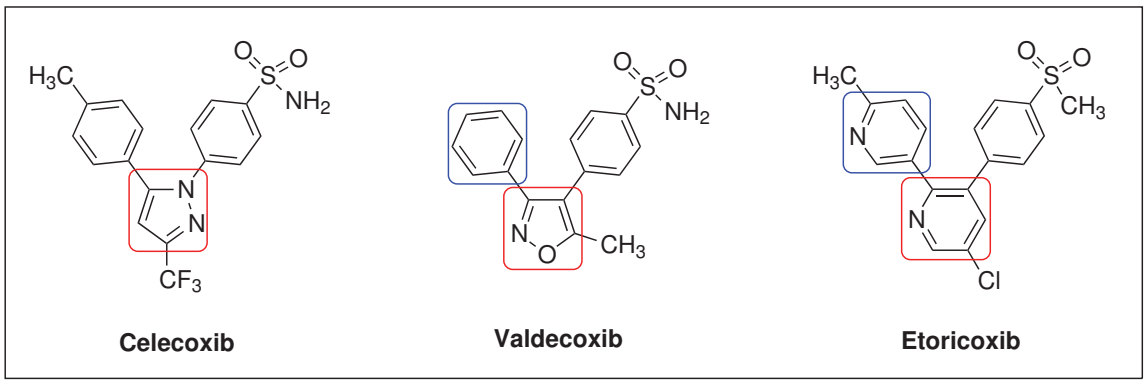


Figure 6.2: Representation of the side chain helical motifs. (**a**): Ribbon representation of IAPP α-helix (PDB ID: 2KB8);[311] (**b**): 3,2',2"-Tris-substituted terphenyl template

Although used as starting points for the synthesis of small-molecule mimetics of α-helices, terphenyls and several of their analogs suffer from long synthetic pathways (10 sequential and 7 linear steps for the synthesis of 3,2',2"-tris-substituted terphenyl templates).[310] Facile modifications of the side-chain R groups are crucial in order to optimize the peptide-ligand interactions in an efficient manner. To avoid the long synthetic routes required to prepare the terphenyls yet to take advantage of the ability of such templates to interact with the side chain α-helix of IAPP, we attempted to design a new template that would be amenable to a flexible synthetic approach to quickly construct the ligands in an efficient and modular manner.

Bioisosterism is a strategy that has been widely applied in medicinal chemistry for the rational design of potential drugs. This strategy involves the substitution of atoms or groups in a molecule with others possessing similar physicochemical properties to give new molecules with similar or improved biological properties. Bioisosteric replacement allows for modifications to improve a compound's solubility, potency, bioavailability and safety. Parameters such as size, electronic distribution, shape, functional group reactivity, and lipid and water solubility play a role in the group's or

atom's ability to function as a bioisostere.[135,136] Ring equivalents are one of the classes of bioisosteres with the most extensive use in drug discovery and development. Under this classification, various heterocycles often function as bioisosteres for benzene rings or other heterocycles. Examples of ring bioisosterism in various anti-inflammatory drugs are illustrated in Scheme 6.1.[135]



Scheme 6.1: Ring equivalent bioisosterism in three classes of anti-inflammatory drugs. Same colored squares indicate ring equivalents[135]

Relying on the bioisosterism of thiophene and benzene,[135,312] a template molecule related to the terphenyl scaffold was designed by maintaining the two terminal benzene rings and replacing the central benzene ring with a thiophene (Figure 6.3, **b**).

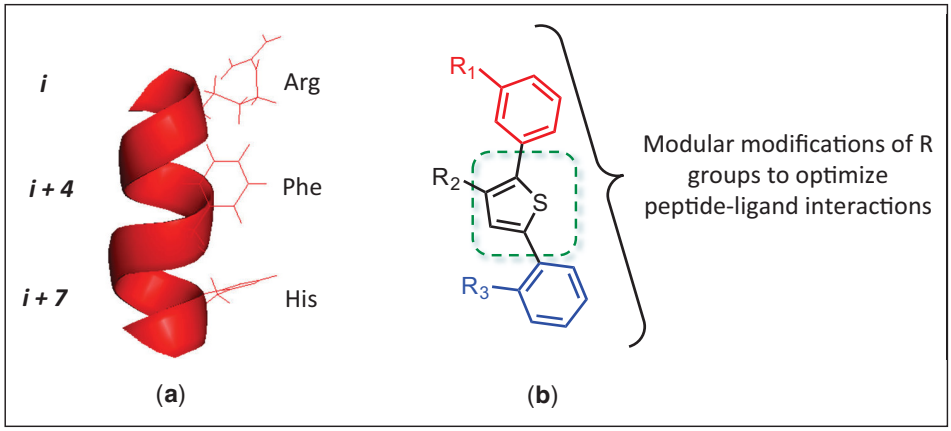


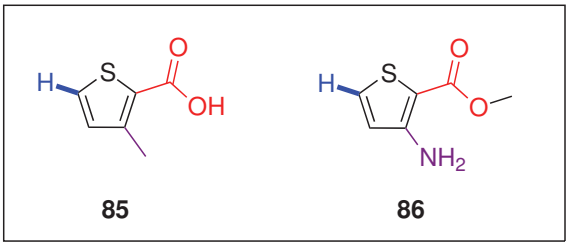
Figure 6.3: Representation of the proposed thiophene template helical mimetics. (**a**): Ribbon representation of IAPP α-helix (PDB ID: 2KB8);[311] (**b**): 2,5- Diaryl substituted thiophene template

The replacement of benzene with thiophene allows for several significant synthetic advantages. A five-membered heteroaromatic core scaffold (Figure 6.3, **b**) affords a flexible synthetic approach in which substituent modifications can be made in a modular manner while avoiding the long synthetic routes that have been used previously for the synthesis of terphenyls.[310] The presence of the heteroatom introduces changes in reactivity that allow convenient chemo- and regioselective pathways that are unavailable in the synthesis of the terphenyl compounds.

6.2.1 Synthesis of the Key Intermediate

Ideally, the side chain positions of the synthetic scaffolds would interact with the amino acid side chains of i, i+4 and i+7 in order to stabilize the α-helix of the protein. The i+4 position of the IAPP is occupied with a phenylalanine residue. To increase the versatility of the synthesized molecules, two commercially available 3-substituted thiophenes, 3-methyl-2-thiophenecarboxylic acid (**85**) and methyl 3-amino-2-thiophenecarboxylate (**86**), were employed. The methyl group at the C3 position of thiophene carboxylic acid (**85**) would provide hydrophobic interactions with the phenylalanine residue of the IAPP, while the amino group of the thiophenecarboxylate

(**86**) could be further functionalized to examine the interaction of bulkier groups at this position as well as other possible interactions such as hydrogen bonding.

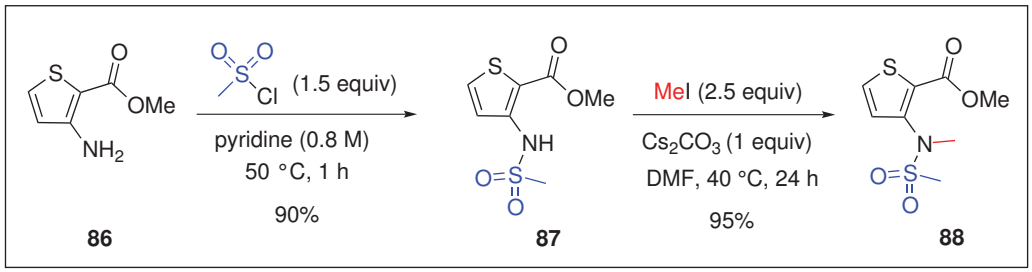


Scheme 6.2: Two commercially available substituted thiophenes

From a synthetic aspect, these commercially available thiophenes provide readily accessible starting materials from which highly regioselective synthetic routes can be established. Moreover, thiophene is more electron-rich than a typical benzene ring and therefore by using properly designed reaction conditions, a regioselective synthesis would be achieved. Functionalization of the amino group at the C3 position of methyl 3-amino-2-thiophenecarboxylate (**86**) provides one of the key intermediates for the synthesis of diaryl substituted thiophenes and serves as a protecting group to allow further cross-coupling reactions to be carried out.

The primary amine of the thiophene carboxylate (**86**) was initially functionalized with a mesyl group. Previous attempts in our group to functionalize the primary amine of the same thiophene with a tosyl group using strong bases such as sodium hydride (1 equiv) generated a mixture of products due to decomposition of the starting material. Therefore, an excess of pyridine, which also served as the base, was used as the reaction solvent along with an optimized 1.5 equivalents of mesyl chloride to obtain the mesylated amine (**87**) in 90% yield (Scheme 6.3). The reaction could also be performed on a multi-gram scale and the product was purified via recrystallization. Methylation of the secondary sulfonamide (**87**) was then performed based on the protocol reported by Taddei *et al.*[313] The methylated sulfonamide (**88**) obtained from this method, which employs Cs_2CO_3 as the base and methyl iodide, did not require

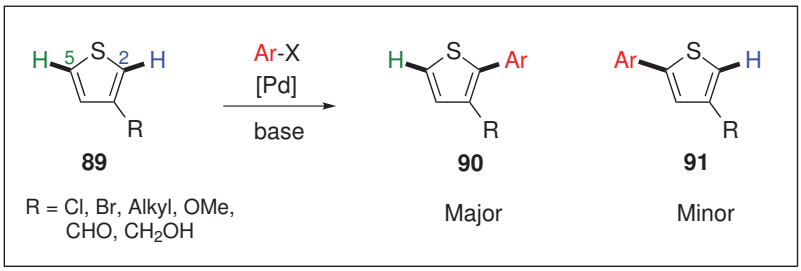
any chromatographic purification and was produced in excellent yield (95%) (Scheme 6.3).



Scheme 6.3: Mesylation and methylation of methyl 3-amino-2-thiophenecarboxylate

6.2.2 Synthesis of 2,5-Diaryl Substituted Thiophenes

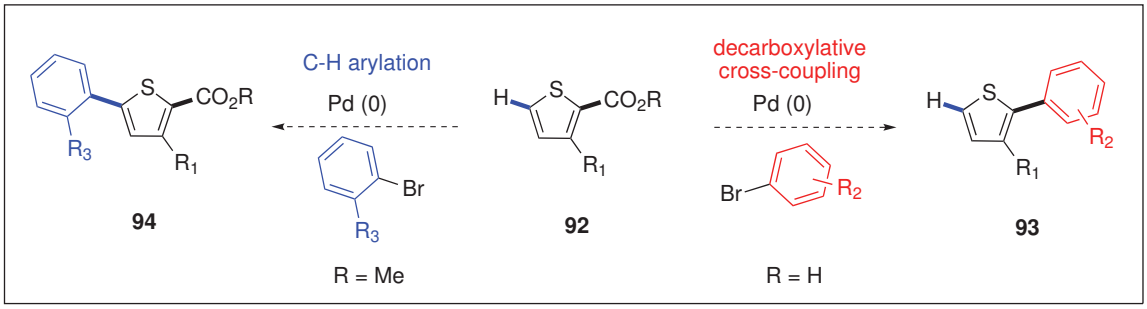
The synthesis of diaryl substituted heteroaromatics has been previously accomplished through various palladium-catalyzed cross-coupling reactions. The predominant strategies involve the utilization of organometallic precursors and/or result in the formation of symmetrically substituted heteroaromatics.[272,312,314–317] For example, palladium-catalyzed C–H activation reactions have emerged as attractive methods for the formation of carbon-carbon bonds between heteroarenes and aryl halides without the use of organometallic derivatives.[318–335]



Scheme 6.4: Regioselectivity in the C–H activation reaction of 3-substituted thiophenes[286,289,336]

However, the main limitation of the C–H functionalization of 3-substituted thiophenes is the formation of mixed arylated products at the C2 and C5 positions (Scheme 6.4).[286,336]

To avoid this limitation yet still take advantage of the C–H arylation strategy while controlling the regioselectivity of the products, the previously outlined commercially available 3-substituted thiophene 2-carboxylic acids have been used in combination with a decarboxylative cross-coupling reaction (Scheme 6.5). Decarboxylative cross-coupling reactions have been developed as a powerful method for the formation of carbon-carbon bonds between aliphatic and aromatic carboxylic acids and aryl or vinyl substrates.[337–342] Decarboxylative arylation processes circumvent the requirement of organometallic building blocks[343–345]offering readily available, inexpensive and easy to use coupling partners. In this view, we performed palladium-catalyzed decarboxylative cross-coupling reactions of thiophene carboxylic acids and various aryl bromides.[29,286]



Scheme 6.5: Two methods of palladium-catalyzed cross-coupling reactions[286,346]

The combination of both the C–H arylation and decarboxylative cross-coupling reactions allow for a short and modular synthetic pathway through which a large library of α-helix mimetic compounds can be readily synthesized.

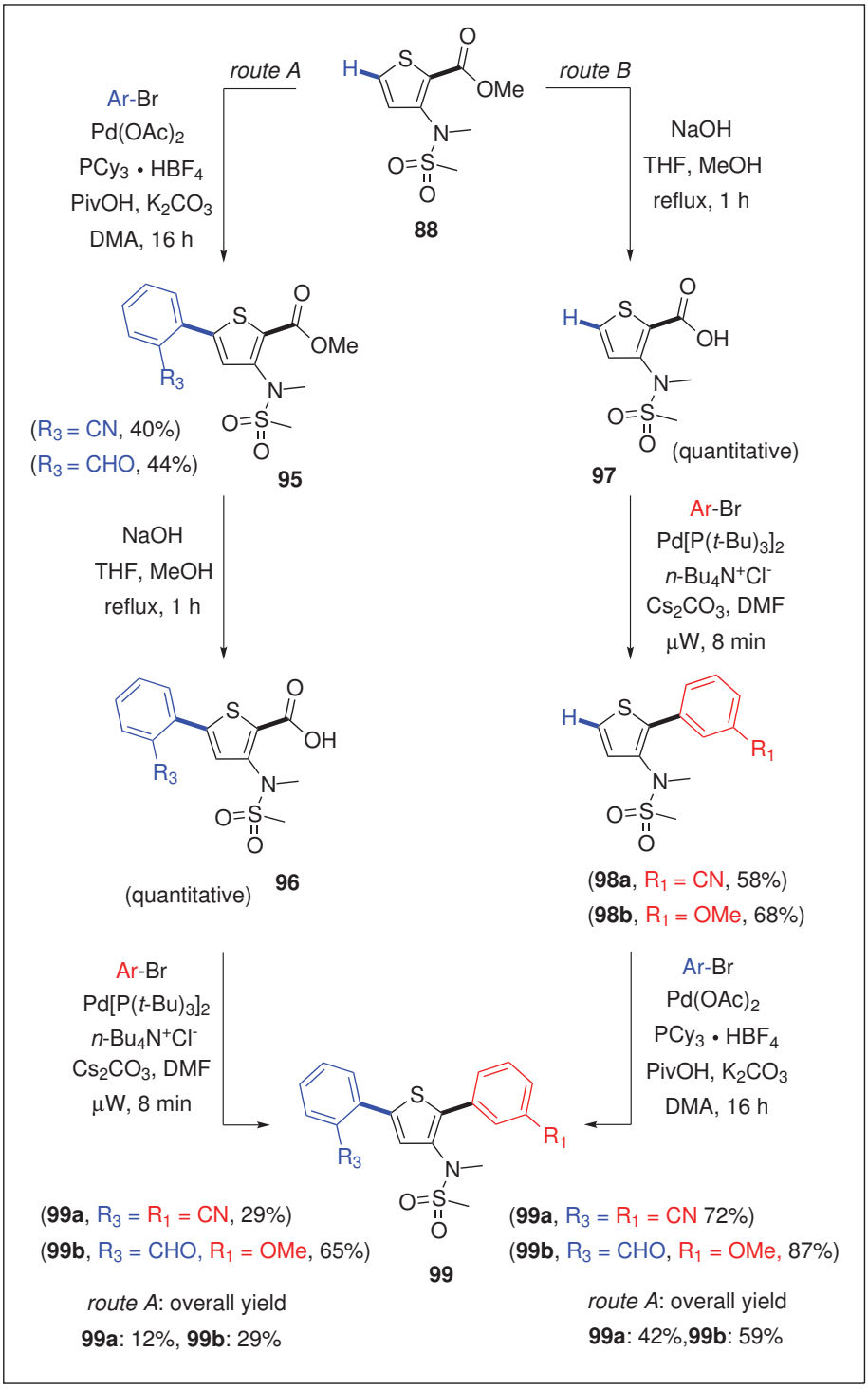
6.2.3 Synthetic Pathways

Two pathways have been envisaged for the preparation of 2,5-diarylated thiophenes (**99**), differing only in the order of the two different coupling reactions.

As illustrated in Scheme 6.6, route A utilizes C5-arylation of the substituted thiophene methyl ester (**88**) resulting in aryl-thiophene intermediate (**95**) followed

by saponification to provide carboxylic acid (**96**). Decarboxylative cross-coupling of acid (**96**) results in the formation of the desired 2,5-diaryl substituted thiophene (**99**). Alternatively, initial saponification of ester (**86**) in pathway B provides the thiophene carboxylic acid intermediate (**97**) that can undergo decarboxylative cross-coupling to afford aryl-thiophene (**98**). This is followed by C5-arylation to provide the desired 2,5-diaryl substituted thiophene (**99**).

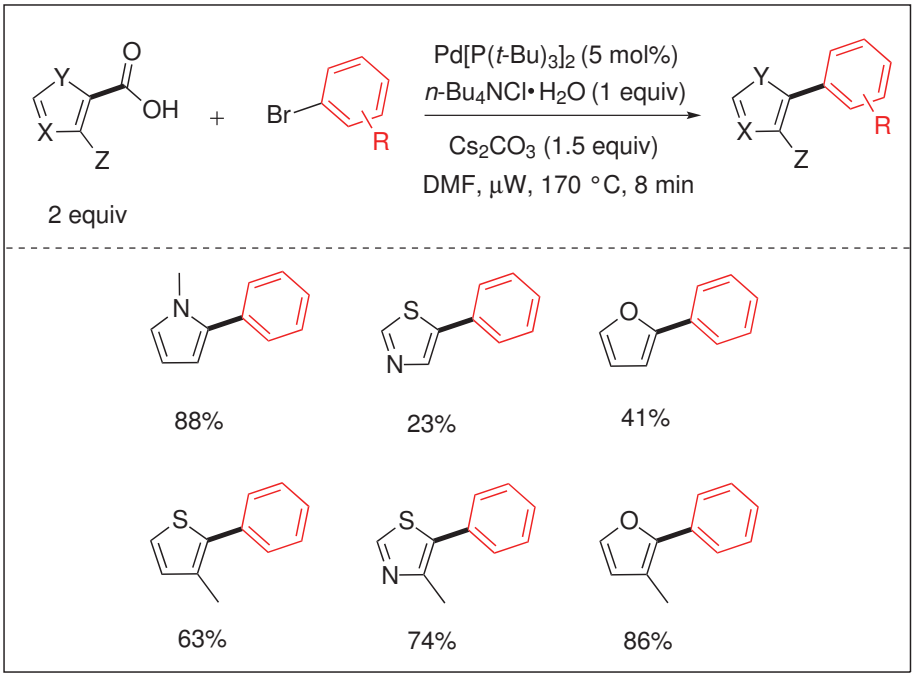
Initially, in order to compare the efficiency of each pathway, both routes were carried out using the same substituted arylbromides (2- and 3-bromobenzonitrile). Interestingly, both the C5-arylation and decarboxylative cross-coupling steps in route A resulted in lower yields compared to route B, giving overall yields of 2,5-diaryl substituted thiophene (**99a**) of 12 and 42% (R_1=CN, R_3=CN), respectively. In order to examine whether the superior efficiency of route B was general, other functionalized aryl halides were also employed in both pathways. Scheme 6.6 shows one other example using 2-bromobenzaldehyde and 3-bromoanisole in which, once again, a higher overall yield was observed with route B compared to route A (59 vs. 29%, respectively, R_1=OMe, R_3=CHO). Route B was therefore chosen for the preparation of the remaining analogues.



Scheme 6.6: Comparison of the two synthetic pathways

6.2.4 Decarboxylative Cross-Coupling Reaction of Thiophene

Forgione *et al.* have reported optimized conditions for the decarboxylative cross-coupling of various heteroaromatics.[286,303] The reaction optimization was developed using *N*-substituted pyrrole carboxylic acids and altering various parameters such as base, solvent, catalyst, aryl substrate and heating system (thermal vs. microwave heating). The optimized reaction conditions employ microwave heating with a reaction time of only 8 minutes and offers a strong control of regioselectivity. Some examples of the molecules generated with this methodology are illustrated in Scheme 6.7.

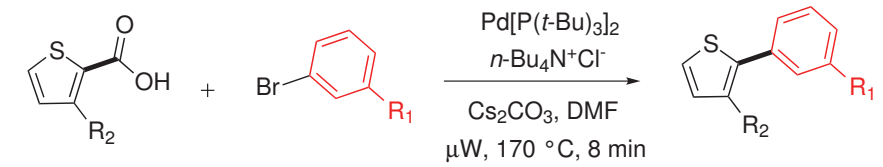


Scheme 6.7: Examples of decarboxylative cross-coupling reaction products by Forgione *et al.*[286,303]

These optimized conditions for decarboxylative cross-coupling reactions were employed to generate the intermediates towards the synthesis of 2,5-diaryl substituted thiophenes. The required thiophene carboxylic acids underwent decarboxylative cross-coupling with a variety of aryl-bromides to produce the corresponding arylated

thiophenes (Table 6.1).

The reaction of both 3-methyl-2-thiophenecarboxylic acid (**85**) and 3-(*N*-methyl methylsulfonamido)thiophene-2-carboxylic acid (**97**) with electron-poor and -rich aryl bromides were performed and the results illustrate that both types of aryl bromide substituents are well tolerated in the reaction.



R_2 = Me **85**
R_2 = *N*-MeMs **97**

100

98a - f

Entry	**R_2**	**R_1**	**Product**	Yield (%)[a]
1	*N*-MeMs	CN	**98a**	58
2	*N*-MeMs	OMe	**98b**	68
3	*N*-MeMs	CO_2Et	**98c**	66
4	Me	CN	**98d**	69
5	Me	OMe	**98e**	75
6	Me	CO_2Et	**98f**	61

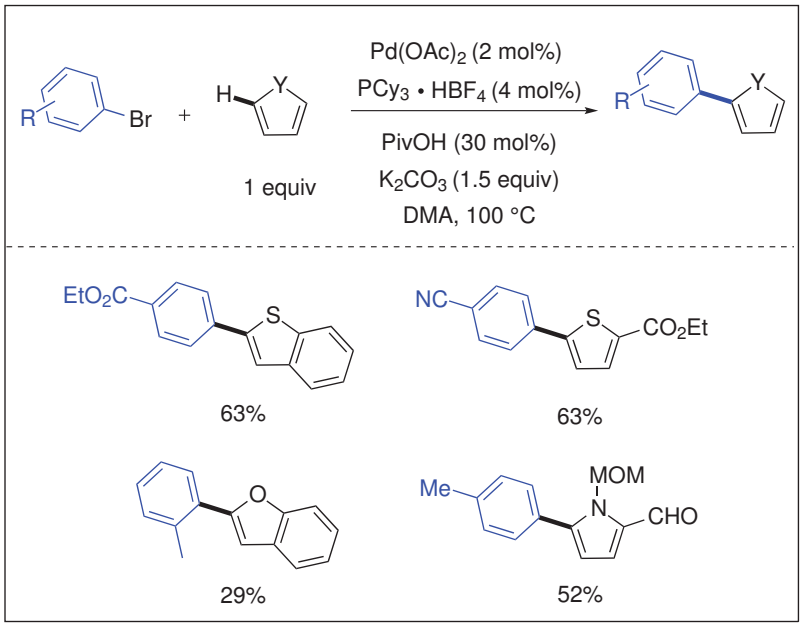
Table 6.1: Synthesis of monoaryl substituted thiophenes. [a]Isolated yields. Condition: thiophene carboxylic acid (2 equiv), aryl bromide (1 equiv), Pd[P(*t*-Bu)$_3$]$_2$ (0.05 equiv), *n*-Bu$_4$N$^+$Cl$^-$ (1 equiv), Cs_2CO_3 (1.5 equiv), anhydrous DMF, 8 min microwave irradiation at 170 °C

These intermediate scaffolds provide one terminal and the central part of the final desired scaffolds mimicking i and i+4. The variety of R_1 substituents would help probe the requirements for effective interaction with the i position of the α-helix of IAPP while the R_2 functional groups provide potential hydrophobic or hydrogen bond interactions with the i+4 position of the target peptide.

6.2.5 C–H Activation Reaction of Aryl Thiophenes

To obtain the desired 2,5-diaryl substituted thiophene templates, the C5 positions of the intermediate compounds resulting from the decarboxylative cross-coupling reactions (**98a-f**) were subjected to a C–H arylation reaction. Although the C5 position of these intermediates is more reactive than the C4 position due to the higher acidity of the proton on C5, complete regioselectivity can not be attained without an appropriate choice of reaction conditions. A survey of the literature revealed that a direct arylation reaction reported by Fagnou *et al.* would meet our requirements.

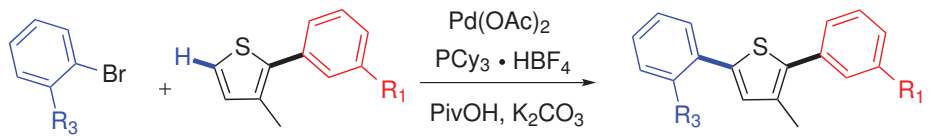
Fagnou and co-workers have developed C–H activation reaction conditions utilizing a wide range of heteroaromatics and aryl bromides.[346] The reaction was specifically effective when the C2 position of the heteroaromatic was blocked, providing regioselective products. Some examples produced from this methodology are illustrated in Scheme 6.8.



Scheme 6.8: Examples of the C–H activation reaction reported by Fagnou *et al.*[346]

Direct arylation reactions previously reported from our group utilized Fagnou's

conditions with slight modifications for the efficient synthesis of thienoisoquinolines.[347] The modified conditions involved a higher loading of the palladium source and ligand (5 mol% and 10 mol%, respectively). The products from the decarboxylative cross-coupling reaction were subjected to the optimized C–H activation conditions to effect a regioselective C–H activation at the C5 position of the 2-arylthiophene to generate 2,5-diaryl substituted thiophenes. A diverse range of products was generated using both the 3-sulfonamide-2-arylthiophenes and 3-methyl-2-arylthiophenes. The results of reactions of 3-methyl-2-arylthiophenes with a range of arylbromides are shown in Table 6.2.



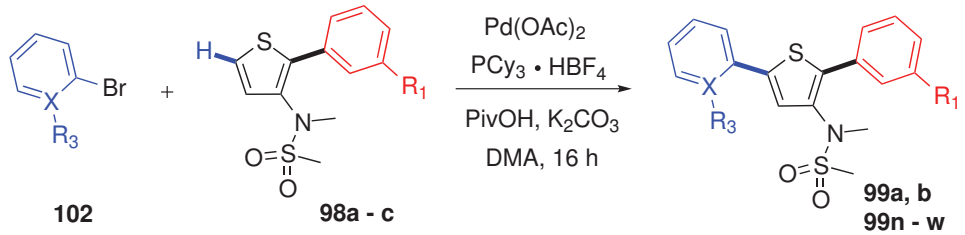
101 **98d - f** **99c - m** **99x**

Entry	R_3	R_1	Product	Yield (%)[a]
7	CN	OMe	**99c**	23
8	CHO	OMe	**99d**	79
9	CF_3	OMe	**99e**	81
10	CO_2Et	OMe	**99f**	73
11	CO_2Et	CO_2Et	**99g**	94
12	CN	CO_2Et	**99h**	81
13	CF_3	CO_2Et	**99i**	72
14	CHO	CO_2Et	**99j**	65
15	CF_3	CN	**99k**	47
16	CN	CN	**99l**	50
17	H	CN	**99m**	62
18	OMe	CO_2Et	**99x**	9

Table 6.2: Synthesis of 2,5-diaryl substituted thiophenes (Series 1) [a]Isolated yields. Conditions: 2-aryl thiophene (1 equiv), aryl bromide (2 equiv), $Pd(OAc)_2$ (0.05 equiv), $PCy_3.HBF_4$ (0.1 equiv), PivOH (0.3 equiv), K_2CO_3 (1.5 equiv), anhydrous DMA, 16 h thermal heating at 100 °C

From these results, it became apparent that the reaction worked well with electron-deficient aryl bromides. However, electron-rich aryl bromide partners exhibited diminished or no reactivity. For example, the above reaction provided 9% yield with 2-bromoanisole (**99x**). These results differ from those observed in Fagnou's work since they observed decreased reactivity for electron-deficient aryl bromides with some coupling partners.[346] More investigation is required to rationalize this change in reactivity.

The results from the reactions of 3-sulfonamide-2-arylthiophenes with a range of aryl bromides under the same reaction conditions are summarized in Table 6.3.



102 **98a - c**

Entry	**X**	**R_3**	**R_1**	**Product**	Yield (%)[a]
19	C	CHO	OMe	**99b**	87
20	C	CN	OMe	**99n**	75
21	C	CF_3	OMe	**99o**	78
22	C	CO_2Et	CN	**99p**	77
23	C	CHO	CN	**99q**	64
24	C	CF_3	CN	**99r**	95
25	C	CN	CN	**99a**	72
26	C	CHO	CO_2Et	**99s**	65
27	C	CF_3	CO_2Et	**99t**	75
28	C	CN	CO_2Et	**99u**	93
29	N	-	CO_2Et	**99v**	10
30	C	OMe	CO_2Et	**99w**	14

Table 6.3: Synthesis of 2,5-diaryl substituted thiophenes (Series 2) [a]Isolated yields; Conditions: 2-aryl thiophene (1 equiv), aryl bromide (2 equiv), $Pd(OAc)_2$ (0.05 equiv), $PCy_3.HBF_4$ (0.1 equiv), PivOH (0.3 equiv), K_2CO_3 (1.5 equiv), anhydrous DMA, 16 h thermal heating at 100 °C

Once again the electron-rich aryl bromides provided lower yields than electron-poor ones. Examples are the reaction of 3-sulfonamide-2-arylthiophene with 2-bromopyrdine and 2-bromoanisole, which generated products (**99v**) and (**99w**) in only 10% and 14% yield respectively.

The results from the above reactions demonstrated that the electron density of the

existing 2-aryl group on the thiophene did not affect the yields of the C5 arylation reaction. However, the identity of the C3 substituent of the thiophene had a small effect on the yields, giving slightly higher yields for the 3-sulfonamide-2-arylthiophene products.

All products were characterized by NMR and high resolution mass spectrometry. For additional confirmation of the regioselectivity of the C–H activation reaction, X-ray crystallography was performed on product (**99n**). The X-ray crystal structure provided clear evidence that the decarboxylative cross-coupling reaction had occurred at the C2 position and the C–H arylation reaction at the C5 position of the thiophene (Figure 6.4).

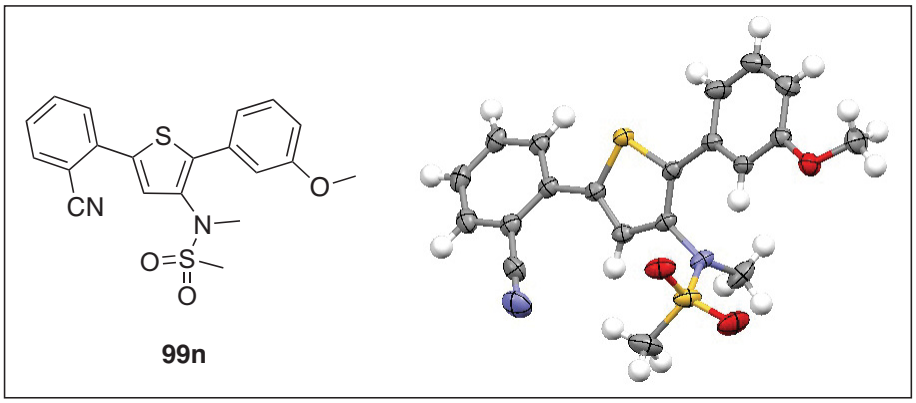


Figure 6.4: X-ray crystal structure of a 2,5-diaryl substituted thiophene

6.3 Evaluation of Islet Amyloid Polypeptide Modulation and Cytotoxicity

6.3.1 ThT Fluorescence Assay

All synthesized 2,5-diaryl substituted thiophenes were initially investigated for their capacity of modulating IAPP amyloid fibril formation by means of the Thioflavin T (ThT) fluorescence assay. Thioflavin T is a dye containing both hydrophobic and polar segments that forms micelles in aqueous media[348] (Figure 6.5, **a**).

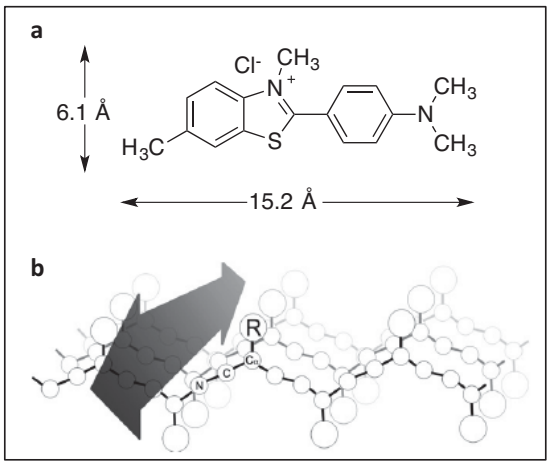


Figure 6.5: Chemical structure of Thioflavin T and β-sheet diagram **a**: Structure and dimensions of Thioflavin T; **b**: Schematic representation of the β-sheet, R represents the side chain residue and C_α, C and N represent the backbone of the β-sheet. Hydrogen bonds are not shown[349]

Among several proposed mechanisms for the binding of ThT to the fibrils, Krebs *et al.*[349] introduced the β-sheet of the proteins responsible for the formation of the binding channels (6.5 to 6.9 Å) where ThT (6.1 Å short axis, 4.3 Å thick) can bind perpendicularly through its short axis (Figure 6.5, **b**). ThT-fibril interactions maintain the ThT structure in a flat and excited conformation. Therefore, ThT dye fluoresces upon binding to protein aggregates with a cross β-sheet structure, mostly fibrillar in morphology.[350]

IAPP amyloidogenesis is a nucleation-dependent polymerization process that is characterized by a ThT-negative phase (lag-phase; around 6 h), in which the high-energy nucleus is formed, followed by a thermodynamically favorable elongation phase that is characterized by the rapid growth of ThT-positive fibrils (Figure 5.3). According to the described helical intermediates, the random coil α-helix conformational conversion occurred during the initial stage of the lag phase. Analysis of the aggregation kinetics obtained by ThT fluorescence gave us early mechanistic insights about the effects of these substituted thiophenes on IAPP amyloidogenic pathway.

6.3.2 ThT Assay Results

IAPP was synthesized by solid phase peptide synthesis on a Rink amide polystyrene resin based on Fmoc chemistry and was purified by preparative scale reversed-phase high performance liquid chromatography (RP-HPLC) (Appendix). The ThT fluorescence assay was performed using aliquots of IAPP dissolved in hexafluoro-2-propanol (HFIP). Following several steps including filtering the solution, lyophilizing the peptide and dissolving it in the assay buffer, ThT fluorescence was measured after addition of each substituted thiophene, and the results were compared to a control sample run without the substituted thiophene. Measurements were obtained every 10 minutes over the course of 25 hours with excitation at 440 nm and emission at 485 nm (Appendix).

Several 2,5-diaryl substituted thiophenes tested by the ThT assay had little or no effect on the kinetics of IAPP amyloid formation (Figure 6.6). Also, several of these compounds (**99i**, **99j**, **99n**, **99o**, **99p** and **99t**) increased the final ThT fluorescence without affecting the lag phase or the rate of amyloid fibrils formation (Appendix).

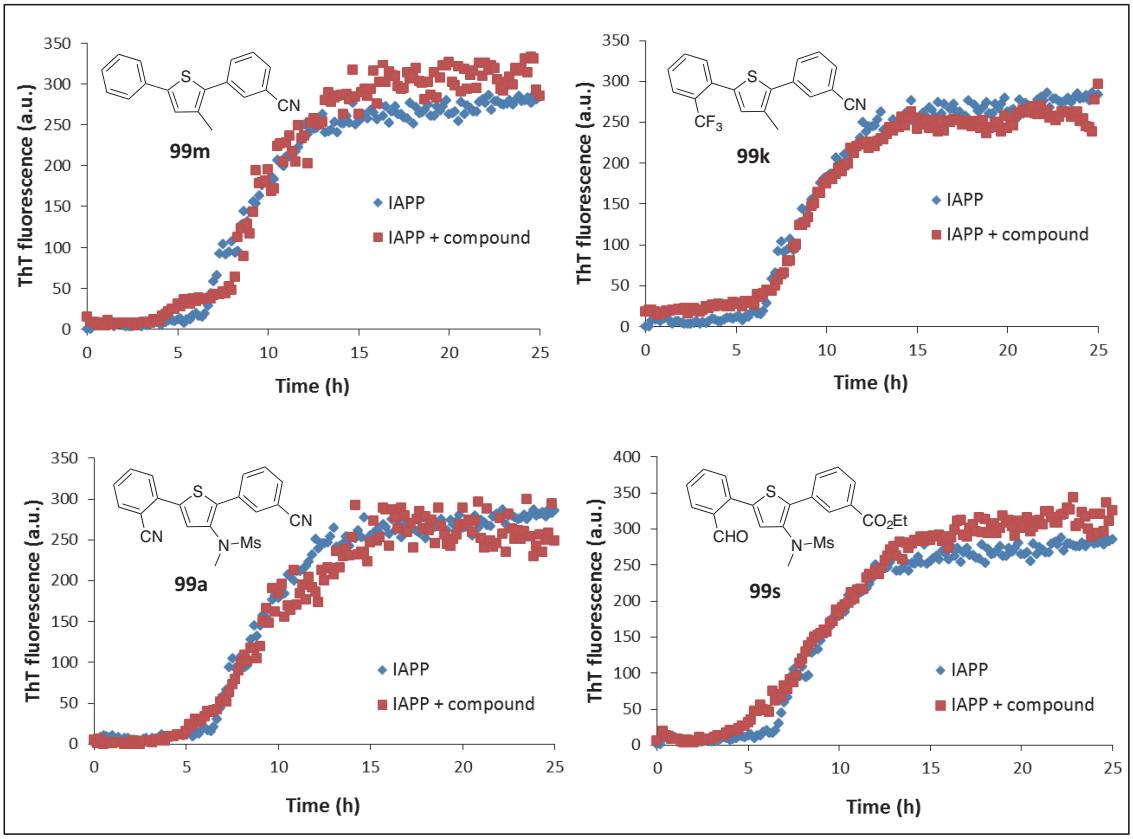


Figure 6.6: Effects of 2,5-diaryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence (Series 1). IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (diamond, blue) or in the presence of 12.5 μM of compound (square, red). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

Among all the compounds prepared in the course of this study, only compound (**99l**) (Table 6.3, entry 16) slowed the formation of ThT-positive aggregates, as observed by the increase of the lag phase period (Figure 6.7, **A**) when the compound was used at an equimolar ratio to IAPP. Moreover, compound (**99l**) showed concentration-dependent inhibition of the formation of IAPP ThT-positive aggregates, with a lag phase of 15 h at 8 molar equivalents (100 μM; Figure 6.7, **B**).

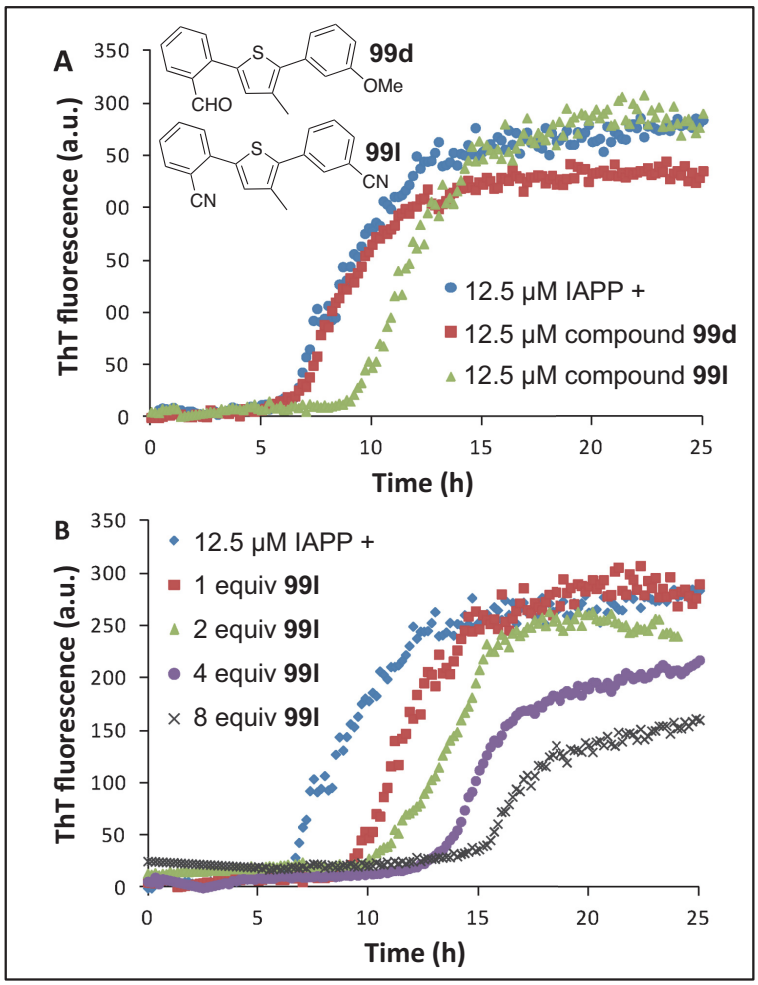


Figure 6.7: Effects of 2,5-diaryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence (Series 2). IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (**A** and **B**; circle, blue) or in the presence of 12.5 μM of compound (**99d**) (**A**; square, red), μM compound **99l** (**A**; triangle, green), increasing molar equivalents of compound **99l** (**B**). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

At 50 and 100 μM (4 and 8 equivalents, respectively) compound (**99l**) also decreased the final ThT fluorescence, suggesting that a lower amount of IAPP amyloid fibrils were formed and/or that these aggregates showed a less defined cross-β-sheets quaternary structure. We also varied the concentration of ThT fluorescent dye to confirm that this inhibitory effect was not the result of a displacement of ThT binding to fibrillar aggregates by compound (**99l**). Our results showed that in the presence of 10, 40 or 100 μM ThT, the increase of the lag-phase period observed with 12.5 μM of compound

(**99l**) was very similar, strongly suggesting that this molecule was, indeed, slowing the amyloidogenic process.

The mechanism by which this 2,5-diaryl substituted thiophene decelerates and partially inhibits IAPP amyloid formation warrants more investigation based on these interesting preliminary results.

6.3.3 Mono- and Di-Carboxylic Acid Substituted Aryl Thiophenes

IAPP is a charged peptide that displays three positive charges at physiological pH, thus favoring electrostatic interactions with negatively charged molecules (Figure 6.8). As a consequence, we designed several mono- and di-carboxylic acid substituted aryl thiophenes to target one side of the transient IAPP helix that exhibits a hydrophobic region (Phe15) surrounded by polar and/or charged residues (Arg11 and His18; Table 3). In this way, the potential interaction of positively charged Arg and partially positively charged His side chains at i and i+7 positions of IAPP with the negatively charged side chains of the ligands could be probed (Figure 6.8).

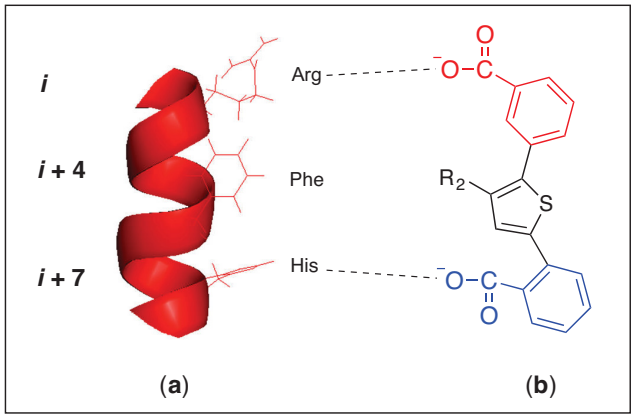
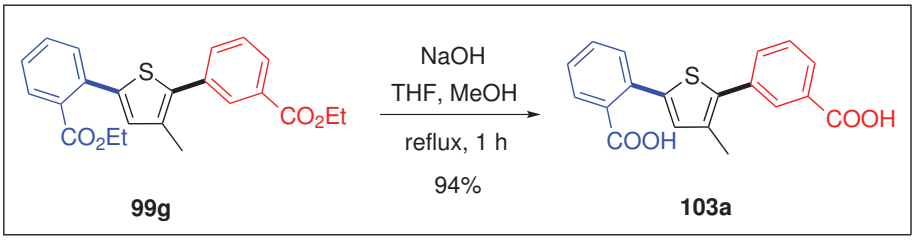


Figure 6.8: Modifications of the side chain functional groups to carboxylic acids towards improved interaction (**a**): Ribbon representation of IAPP α-helix (PDB ID: 2KB8);[311] (**b**): Di-carboxylic acid substituted aryl thiophene template

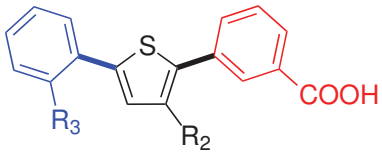
2,5-Diaryl substituted thiophenes encompassing one or two diethyl esters on the

aryl groups were hydrolyzed to provide the mono- or di-carboxylic acid substituted aryl thiophenes, respectively. Scheme 6.9 exemplifies the reaction conditions for the generation of a dicarboxylic acid substituted aryl thiophene (**103a**).



Scheme 6.9: Hydrolysis of diester substituted aryl thiophenes

Subsequently, other monoester aryl substituted thiophenes were hydrolyzed using the above conditions (Table 6.4).



103b - e

Entry	**R_3**	**R_2**	**Product**
31	CN	Me	**103b**
32	CF_3	Me	**103c**
33	CHO	Me	**103d**
34	CF_3	*N*-MeMs	**103e**

Table 6.4: Monoacid aryl substituted thiophenes

The mono- and di-carboxylic acid substituted aryl thiophenes were then evaluated for any effects of electrostatic interactions with the positively charged residues of IAPP at physiological pH.

As suspected, carboxylic acid-functionalized thiophenes showed profound effects on IAPP amyloidogenesis at a 1:1 molar ratio. Interestingly, monocarboxylic acid substituted thiophenes with a methyl group at position R_2 (compounds **103b**, **103c** and **103d**) virtually abolished the lag phase without significantly affecting the final ThT fluorescence. This type of aggregation kinetics suggests that these compounds induce the formation of IAPP aggregates with lower ThT-binding capacities, indicative of non-fibrillar structure (Figure 6.9).

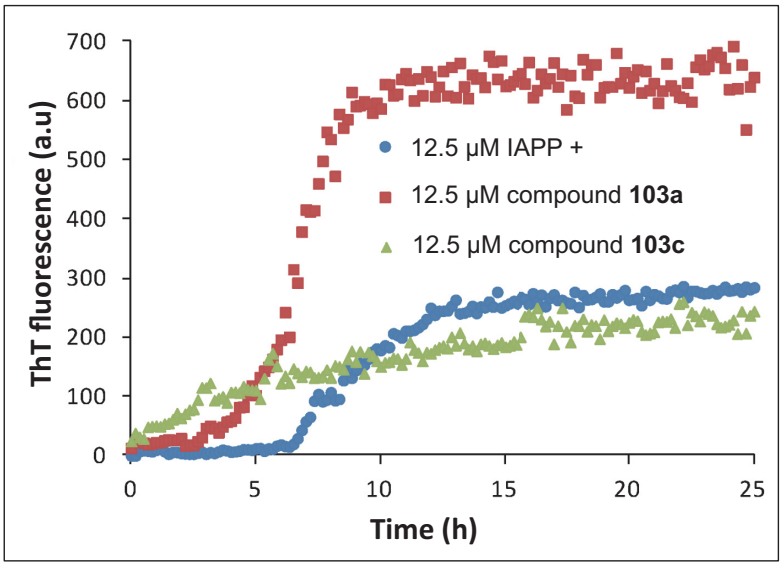


Figure 6.9: Effects of mono- and di-carboxylic acid aryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (circle, blue) or in the presence of 12.5 μM of compound **103a** (square, red) or 12.5 μM of compound **103c** (triangle, green). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

In sharp contrast, the dicarboxylic acid analogue (compound **103a**) reduced the lag phase and led to a significant increase of the final ThT fluorescence (Figure 6.9). This suggests that a larger amount of amyloids was formed in the presence of one equivalent of compound (**103a**) and/or that these amyloid fibrils exhibit a better-defined cross-β-sheet quaternary structure. These possibilities should be investigated more to better understand the mechanism by which these molecules interact with

IAPP.

To probe if the accelerating effects of the mono- and di-carboxylic acid aryl substituted thiophenes on IAPP amyloidogenesis were simply a result of non-specific charge neutralization effects, benzoic acid was used as a control compound. The kinetic data for amyloid formation obtained in the presence of 12.5 μM (1 equivalent) and 125 μM (10 equivalents) of benzoic acid are very similar to the control (Figure 6.10). Together, these data indicated that the negative charge(s) on the thiophene scaffold are crucial for the modulating activity and that the nature and/or the position of other substituents also play a key role, suggesting specific interactions. Again, further investigation would be necessary to determine the exact mechanism by which these derivatives modulate the formation of amyloid fibrils.

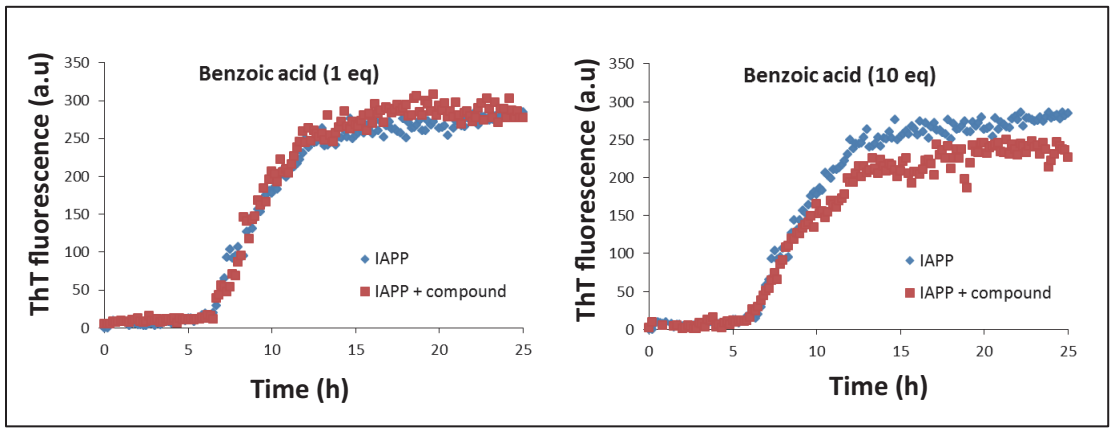
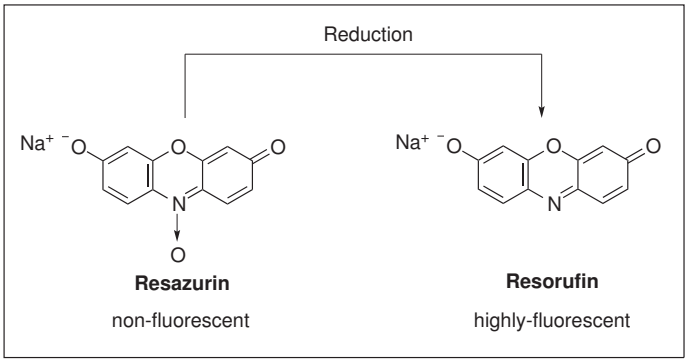


Figure 6.10: Effects of benzoic acid on kinetics of IAPP amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (circle, blue) or in the presence of 1 equivalent and 10 equivalents of benzoic acid (square, red). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

6.3.4 Cell Viability Assays

The cytotoxicity of IAPP species that had been pre-incubated for 20 h in the absence or presence of selected 2,5-diarylthiophene derivatives were analyzed using rat INS-1 (β-pancreatic cell line) cells (Appendix). Cell viability was measured by the resazurin reduction assay. Resazurin can measure the viability of bacterial and mammalian cells. Resazurin is a non-fluorescent dye that can form highly fluorescent resorufin when reduced by living cells (Scheme 6.10). The cell viability was calculated from the ratio of the fluorescence of the treated sample to the control cells (non-treated).



Scheme 6.10: Reduction of Resazurin to Resorufin.

Other groups have previously reported that IAPP induces death of pancreatic cells when the amyloidogenic peptide is directly added to the cell culture medium.[202,308] In fact, IAPP pre-incubated for 20 h without compounds decreased pancreatic β-cells viability in a concentration-dependent manner (Figure 6.11, **A**). When IAPP was pre-incubated with 1 molar equivalent of either compound (**99d**), (**99l**) or (**103c**), no changes in the proteotoxic effects induced by 50 μM IAPP were observed (Figure 6.11, **B**). However, pre-incubation of IAPP with the dicarboxylic acid substituted aryl thiophene (compound **103a**) before cell treatment abolished the cytotoxic effects of IAPP. This result suggests that this compound stimulates the formation of poorly toxic IAPP quaternary species, mostly fibrillar, according to the high ThT fluorescence

observed (Figure 6.11, **B**). It is noteworthy that all tested compounds were not toxic on β-pancreatic cells when used at a concentration of 50 μM.

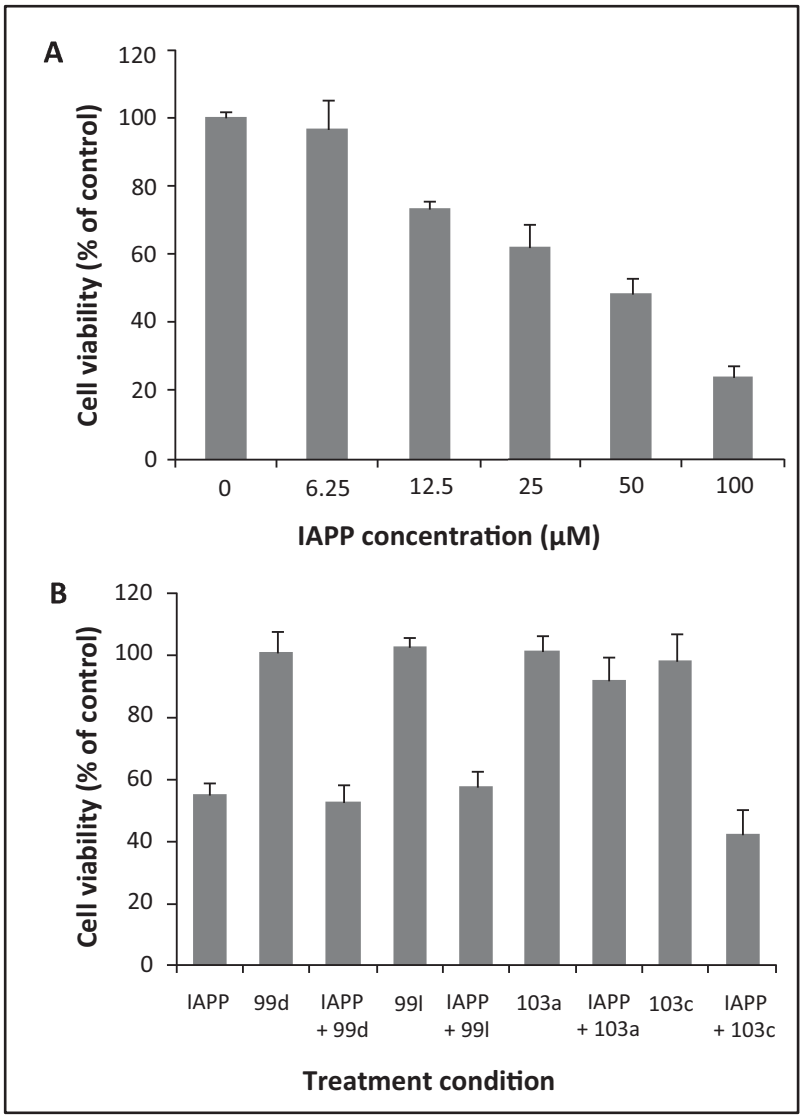


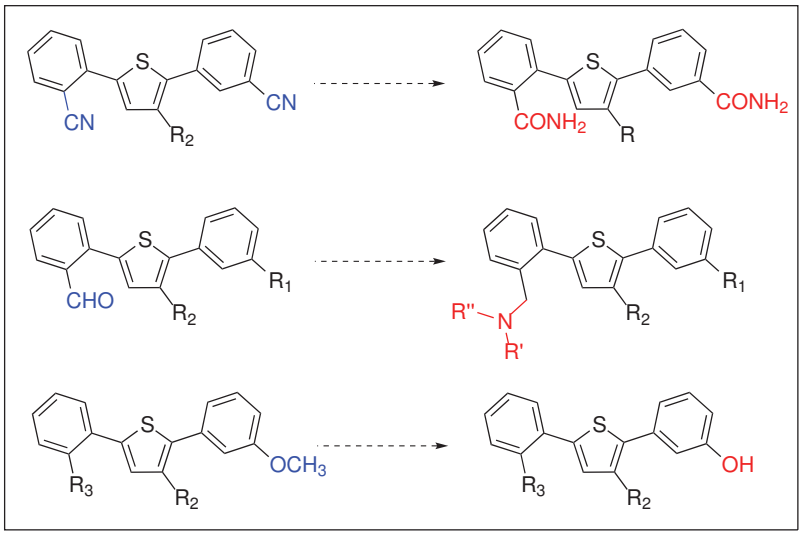
Figure 6.11: Effects of 2,5-diaryl substituted thiophenes on IAPP-induced toxicity on pancreatic β-cells. (**A**): INS-1 cells were treated with concentrations of IAPP ranging from 0 to 100 μM for 24 h and cell viability was measured by the resazurin reduction assay and compared to cells treated with vehicle only (100% cell viability). (**B**): INS-1 cells were treated with 50 μM IAPP that had been pre-incubated for 20 h in 20 mM Tris, pH 7.4, 25 °C, in the absence or presence of one molar equivalent of the thiophene derivatives. After 24 h incubation, cell viability was measured.

It is worthwhile to perform biophysical investigations to delineate the mechanisms by which these molecules interfere with IAPP amyloidogenic process. This study

demonstrates that we can modulate not only the kinetics of amyloid fibril formation of an amyloidogenic peptide, but also its cytotoxicity with small molecules that were designed to mimic/target the transient helical motif.

6.4 Future Directions

The modular and short synthetic pathway for the synthesis of 2,5-diaryl substituted thiophenes allows for rapid installation of the aryl substituted groups on the thiophene. Some of these side chain substituents can also be easily converted to other functional groups. This modular synthesis is particularly advantageous in structure-activity relationship (SAR) studies where compound modifications can lead to improved biological activity. With this small library of 2,5-diaryl substituted thiophenes, further diversification of the side chain substituents will be investigated for improving the ligand-IAPP interactions based on the biological assays. For instance, transformations of the nitriles to amides, aldehydes to amines and methoxy groups to alcohols can be used to access side chain functionalities that are not compatible with our chosen conditions of palladium-catalyzed cross-coupling reactions.



Scheme 6.11: Diversification of the side chain substituents of the aryl groups

Despite developing a successful methodology for the arylation of thiophenes, the current conditions are unable to efficiently effect a C–H arylation reaction between thiophene and electron-rich aryl bromides, which places some limitations on the generation of various aryl substituted thiophenes. Therefore, the development of catalytic systems capable of tolerating both electron-deficient and -rich aryl halides would be beneficial.

6.5 Conclusion

There is a great interest in designing non-peptidic small molecules to interact with proteins. Due to their essential roles in mediating protein-protein interactions, α-helices have become attractive targets. Several classes of compounds have been developed to mimic these ordered secondary structures and subsequently stabilize their conformation. These developed molecules were mainly used to disrupt protein-protein interactions. Inspired by this, we targeted IAPP, which is a peptidic hormone that forms amyloid fibrils. We have developed a modular approach using palladium-mediated cross-coupling reactions for the synthesis of highly functionalized small molecules based on a 2,5-diaryl substituted thiophene scaffold. This strategy allows us to quickly construct ligands to screen for interaction with and stabilization of α-helices in an efficient manner. In this effort, the ligands were assessed for their capacity to modulate IAPP amyloidogenesis and influence the cytotoxicity of the species generated from this process on β-pancreatic cells. The results demonstrated that some of the molecules could act as modulators of IAPP amyloidogenesis by increasing or decreasing the lag-phase period of IAPP amyloid fibril formation. This would be a potential research area to better understand the mechanism by which these molecules interact with IAPP. As several amyloidogenic natively disordered (poly)peptides, including the amyloid-b peptide, calcitonin and α-synuclein, populate helical intermediates during

the initial phase of fibril formation, these 2,5-diaryl substituted thiophenes could ultimately lead to the development of novel therapeutics for protein amyloid-related diseases.

Bibliography

[1] Thibeault, D.; Maurice, R.; Pilote, L.; Lamarre, D.; Pause, A. *J. Biol. Chem.* **2001**, *276*, 46678–46684.

[2] Marzban, L.; Verchere, C. B. *Can. J. Diabetes* **2004**, *28*, 39–47.

[3] Lamarre, D.; Pilote, L. Purified active HCV NS2/3 protease. Patent US7264811B2, **2007**.

[4] Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28–39.

[5] Shangary, S.; Wang, S. *Annu. Rev. Pharmacol. Toxicol.* **2009**, *49*, 223.

[6] Khoury, K.; Domling, A. *Curr. Pharm. Des.* **2012**, *18*, 4668–4678.

[7] De Meyer, S.; Azijn, H.; Surleraux, D.; Jochmans, D.; Tahri, A.; Pauwels, R.; Wigerinck, P.; de Béthune, M.-P. *Antimicrob. Agents Chemother.* **2005**, *49*, 2314–2321.

[8] Ghosh, A. K.; Dawson, Z. L.; Mitsuya, H. *Bioorganic Med. Chem.* **2007**, *15*, 7576–7580.

[9] Molina, J.-M.; Hill, A. *Expert Opin. Pharmacother.* **2007**, *8*, 1951–1964.

[10] Belluti, F.; Rampa, A.; Gobbi, S.; Bisi, A. *Expert Opin. Ther. Pat.* **2013**, *23*, 581–596.

[11] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev* **2001**, *46*, 1–3.

[12] Vane, J. R. *Nature* **1971**, *231*, 232–235.

[13] Swinney, D. C. *Nat. Rev. Drug Discov.* **2004**, *3*, 801–808.

[14] Swinney, D. C. *Pharmaceut. Med.* **2008**, *22*, 23–34.

[15] Yang, J. C.-H.; Shih, J.-Y.; Su, W.-C.; Hsia, T.-C.; Tsai, C.-M.; Ou, S.-H. I.; Yu, C.-J.; Chang, G.-C.; Ho, C.-L.; Sequist, L. V.; Dudek, A. Z.; Shahidi, M.; Cong, X. J.; Lorence, R. M.; Yang, P.-C.; Miller, V. A. *Lancet Oncol.* **2012**, *13*, 539–548.

[16] Johnson, D. S.; Weerapana, E.; Cravatt, B. F. *Future Med. Chem.* **2010**, *2*, 949–964.

[17] Stiefel, U.; Harmoinen, J.; Koski, P.; Kääriäinen, S.; Wickstrand, N.; Lindevall, K.; Pultz, N. J.; Bonomo, R. A.; Helfand, M. S.; Donskey, C. J. *Antimicrob. Agents Chemother.* **2005**, *49*, 5190–5191.

[18] Grant, S. M.; Heel, R. C. *Drugs* **1991**, *41*, 889–926.

[19] Najib, J. *Clin. Ther.* **2001**, *23*, 802–832.

[20] Schrag, A. *Lancet Neurol.* **2005**, *4*, 366–370.

[21] Venkatraman, S. *Trends Pharmacol. Sci.* **2012**, *33*, 289–294.

[22] Rotella, D. P. *Expert Opin. Drug Discov.* **2013**, *8*, 1439–1447.

[23] Collier, A. C.; Coombs, R. W.; Schoenfeld, D. A.; Bassett, R. L.; Timpone, J.; Baruch, A.; Jones, M.; Facey, K.; Whitacre, C.; McAuliffe, V. J.; Friedman, H. M.; Merigan, T. C.; Reichman, R. C.; Hooper, C.; Corey, L. *N. Engl. J. Med.* **1996**, *334*, 1011–1018.

[24] Kim, A. E.; Dintaman, J. M.; Waddell, D. S.; Silverman, J. A. *J. Pharmacol. Exp. Ther.* **1998**, *286*, 1439–1445.

[25] Rogers, S. L.; Farlow, M. R.; Doody, R. S.; Mohs, R.; Friedhoff, L. T. *Neurology* **1998**, *50*, 136–145.

[26] Sugimoto, H. *Chem. Rec.* **2001**, *1*, 63–73.

[27] Lamarre, D.; Anderson, P. C.; Bailey, M.; Beaulieu, P.; Bolger, G.; Bonneau, P.; Bös, M.; Cameron, D. R.; Cartier, M.; Cordingley, M. G.; Faucher, A.-M.; Goudreau, N.; Kawai, S. H.; Kukolj, G.; Lagace, L.; LaPlante, S. R.; Narjes, H.; Poupart, M.-A.; Rancourt, J.; Sentjens, R. E.; St George, R.; Simoneau, B.; Steinmann, G.; Thibeault, D.; Tsantrizos, Y. S.; Weldon, S. M.; Yong, C.-L.; Llinas-Brunet, M. *Nature* **2003**, *426*, 186–189.

[28] Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.

[29] Hassanpour, A.; De Carufel, C. A.; Bourgault, S.; Forgione, P. *Chem. Eur. J.* **2014**, *20*, 2522–2528.

[30] Binder, M.; Bartenschlager, R. *Cell Res.* **2013**, *23*, 1–2.

[31] Gordon, C. P.; Keller, P. A. *J. Med. Chem.* **2005**, *48*, 1–20.

[32] Sy, T.; Jamal, M. M. *Int. J. Med. Sci.* **2006**, *3*, 41–46.

[33] Alter, M. J. *World J. Gastroenterol.* **2007**, *13*, 2436–2441.

[34] Hanafiah, K. M.; Groeger, J.; Flaxman, A. D.; Wiersma, S. T. *Hepatology* **2013**, *57*, 1333–1342.

[35] Shepard, C. W.; Finelli, L.; Alter, M. J. *Lancet Infect. Dis.* **2005**, *5*, 558–567.

[36] Lavanchy, D.; Purcell, R.; Hollinger, F. B.; Howard, C.; Alberti, A.; Kew, M.; Dusheiko, G.; Alter, M.; Ayoola, E.; Beutels, P. *J. Viral Hepat.* **1999**, *6*, 35–47.

[37] Rustgi, V. K. *J. Gastroenterol.* **2007**, *42*, 513–521.

[38] Frank, C.; Mohamed, M. K.; Strickland, G. T.; Lavanchy, D.; Arthur, R. R.; Magder, L. S.; Khoby, T. E.; Abdel-Wahab, Y.; Ohn, E. S. A.; Anwar, W. *Lancet* **2000**, *355*, 887–891.

[39] Abdel-Aziz, F.; Habib, M.; Mohamed, M. K.; Abdel-Hamid, M.; Gamil, F.; Madkour, S.; Mikhail, N. N.; Thomas, D.; Fix, A. D.; Strickland, G. T. *Hepatology* **2000**, *32*, 111–115.

[40] Hajarizadeh, B.; Grebely, J.; Dore, G. J. *Nat. Rev. Gastroenterol. Hepatol.* **2013**, *10*, 553–562.

[41] Bronowicki, J.-P.; Ouzan, D.; Asselah, T.; Desmorat, H.; Zarski, J.-P.; Foucher, J.; Bourlière, M.; Renou, C.; Tran, A.; Melin, P. *Gastroenterology* **2006**, *131*, 1040–1048.

[42] McHutchison, J. G.; Gordon, S. C.; Schiff, E. R.; Shiffman, M. L.; Lee, W. M.; Rustgi, V. K.; Goodman, Z. D.; Ling, M.-H.; Cort, S.; Albrecht, J. K. *N. Engl. J. Med.* **1998**, *339*, 1485–1492.

[43] Manns, M. P.; McHutchison, J. G.; Gordon, S. C.; Rustgi, V. K.; Shiffman, M.; Reindollar, R.; Goodman, Z. D.; Koury, K.; Ling, M.-H.; Albrecht, J. K. *Lancet* **2001**, *358*, 958–965.

[44] Davis, G. L.; Wong, J. B.; McHutchison, J. G.; Manns, M. P.; Harvey, J.; Albrecht, J. *Hepatology* **2003**, *38*, 645–652.

[45] Wohl, B. M.; Smith, A. A.; Kryger, M. B.; Zelikin, A. N. *Biomacromolecules* **2013**, *14*, 3916–3926.

[46] van der Meer, A. J.; Wedemeyer, H.; Feld, J. J.; Hansen, B. E.; Manns, M. P.; Zeuzem, S.; Janssen, H. L. A. *J. Hepatol.* **2013**,

[47] Poynard, T.; Marcellin, P.; Lee, S. S.; Niederau, C.; Minuk, G. S.; Ideo, G.; Bain, V.; Heathcote, J.; Zeuzem, S.; Trepo, C. *Lancet* **1998**, *352*, 1426–1432.

[48] Zeuzem, S.; Hultcrantz, R.; Bourliere, M.; Goeser, T.; Marcellin, P.; Sanchez-Tapias, J.; Sarrazin, C.; Harvey, J.; Brass, C.; Albrecht, J. *J. Hepatol.* **2004**, *40*, 993–999.

[49] Fried, M. W. *Hepatology* **2002**, *36*, S237–S244.

[50] Manns, M. P.; Wedemeyer, H.; Cornberg, M. *Gut* **2006**, *55*, 1350–1359.

[51] Schaefer, M.; Schmidt, F.; Folwaczny, C.; Lorenz, R.; Martin, G.; Schindlbeck, N.; Heldwein, W.; Soyka, M.; Grunze, H.; Koenig, A. *Hepatology* **2003**, *37*, 443–451.

[52] Ghany, M. G.; Strader, D. B.; Thomas, D. L.; Seeff, L. B. *Hepatology* **2009**, *49*, 1335–1374.

[53] Gane, E. J.; Roberts, S. K.; Stedman, C. A. M.; Angus, P. W.; Ritchie, B.; Elston, R.; Ipe, D.; Morcos, P. N.; Baher, L.; Najera, I.; Chu, T.; Lopatin, U.; Berrey, M. M.; Bradford, W.; Laughlin, M.; Shulman, N. S.; Smith, P. F. *Lancet* **2010**, *376*, 1467–1475.

[54] Asselah, T.; Marcellin, P. *Liver Int.* **2011**, *31*, 68–77.

[55] Sherman, K. E.; Flamm, S. L.; Afdhal, N. H.; Nelson, D. R.; Sulkowski, M. S.; Everson, G. T.; Fried, M. W.; Adler, M.; Reesink, H. W.; Martin, M.; Sankoh, A. J.; Adda, N.; Kauffman, R. S.; George, S.; Wright, C. I.; Poordad, F. *N. Engl. J. Med.* **2011**, *365*, 1014–1024.

[56] Jacobson, I. M.; McHutchison, J. G.; Dusheiko, G.; Di Bisceglie, A. M.; Reddy, K. R.; Bzowej, N. H.; Marcellin, P.; Muir, A. J.; Ferenci, P.; Flisiak, R.; George, J.; Rizzetto, M.; Shouval, D.; Sola, R.; Terg, R. A.; Yoshida, E. M.;

Adda, N.; Bengtsson, L.; Sankoh, A. J.; Kieffer, T. L.; George, S.; Kauffman, R. S.; Zeuzem, S. *N. Engl. J. Med.* **2011**, *364*, 2405–2416.

[57] Sofia, M. J.; Chang, W.; Furman, P. A.; Mosley, R. T.; Ross, B. S. *J. Med. Chem.* **2012**, *55*, 2481–2531.

[58] Lok, A. S.; Gardiner, D. F.; Lawitz, E.; Martorell, C.; Everson, G. T.; Ghalib, R.; Reindollar, R.; Rustgi, V.; McPhee, F.; Wind-Rotolo, M.; Persson, A.; Zhu, K.; Dimitrova, D. I.; Eley, T.; Guo, T.; Grasela, D. M.; Pasquinelli, C. *N. Engl. J. Med.* **2012**, *366*, 216–224.

[59] Haudecoeur, R.; Peuchmaur, M.; Ahmed-Belkacem, A.; Pawlotsky, J.-M.; Boumendjel, A. *Med. Res. Rev.* **2012**, *33*, 934–984.

[60] De Clercq, E. *Med. Res. Rev.* **2013**, *33*, 1249–1277.

[61] Sofia, M. J.; Bao, D.; Chang, W.; Du, J.; Nagarathnam, D.; Rachakonda, S.; Reddy, P. G.; Ross, B. S.; Wang, P.; Zhang, H.-R.; Bansal, S.; Espiritu, C.; Keilman, M.; Lam, A. M.; Micolochick Steuer, H. M.; Niu, C.; Otto, M. J.; Furman, P. A. *J. Med. Chem.* **2010**, *53*, 7202–7218.

[62] Afdhal, N.; Zeuzem, S.; Kwo, P.; Chojkier, M.; Gitlin, N.; Puoti, M.; Romero-Gomez, M.; Zarski, J.-P.; Agarwal, K.; Buggisch, P.; Foster, G. R.; Brau, N.; Buti, M.; Jacobson, I. M.; Subramanian, G. M.; Ding, X.; Mo, H.; Yang, J. C.; Pang, P. S.; Symonds, W. T.; McHutchison, J. G.; Muir, A. J.; Mangia, A.; Marcellin, P. *N. Engl. J. Med.* **2014**, *370*, 1889–1898.

[63] Gao, M.; Nettles, R. E.; Belema, M.; Snyder, L. B.; Nguyen, V. N.; Fridell, R. A.; Serrano-Wu, M. H.; Langley, D. R.; Sun, J.-H.; O'Boyle II, D. R.; Lemm, J. A.; Wang, C.; Knipe, J. O.; Chien, C.; Colonno, R. J.; Grasela, D. M.; Meanwell, N. A.; Hamann, L. G. *Nature* **2010**, *465*, 96–100.

[64] Dusheiko, G.; Schmilovitz-Weiss, H.; Brown, D.; McOmish, F.; Yap, P.-L.; Sherlock, S.; McIntyre, N.; Simmonds, P. *Hepatology* **1994**, *19*, 13–18.

[65] Szabó, E.; Lotz, G.; Páska, C.; Kiss, A.; Schaff, Z. *Pathol. Oncol. Res.* **2003**, *9*, 215–221.

[66] Simmonds, P.; Bukh, J.; Combet, C.; Deléage, G.; Enomoto, N.; Feinstone, S.; Halfon, P.; Inchauspé, G.; Kuiken, C.; Maertens, G. *Hepatology* **2005**, *42*, 962–973.

[67] Choo, Q. L.; Richman, K. H.; Han, J. H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coit, D.; Medina-Selby, R.; Barr, P. J. *Proc. Natl. Acad. Sci.* **1991**, *88*, 2451–2455.

[68] Borowski, P.; Niebuhr, A.; Schmitz, H.; Hosmane, R. S.; Bretner, M.; Siwecka, M. A.; Kulikowski, T. *Acta Biochim. Pol.* **2002**, *49*, 597–614.

[69] Romero-López, C.; Berzal-Herranz, A. *RNA* **2009**, *15*, 1740–1752.

[70] Tellinghuisen, T. L.; Evans, M. J.; Von Hahn, T.; You, S.; Rice, C. M. *J. Virol.* **2007**, *81*, 8853–8867.

[71] Suzuki, R.; Suzuki, T.; Ishii, K.; Matsuura, Y.; Miyamura, T. *Intervirology* **1999**, *42*, 145–152.

[72] Lindenbach, B. D.; Evans, M. J.; Syder, A. J.; Wölk, B.; Tellinghuisen, T. L.; Liu, C. C.; Maruyama, T.; Hynes, R. O.; Burton, D. R.; McKeating, J. A.; Rice, C. M. *Science* **2005**, *309*, 623–626.

[73] Pileri, P.; Uematsu, Y.; Campagnoli, S.; Galli, G.; Falugi, F.; Petracca, R.; Weiner, A. J.; Houghton, M.; Rosa, D.; Grandi, G.; Abrignani, S. *Science* **1998**, *282*, 938–941.

[74] Scarselli, E.; Ansuini, H.; Cerino, R.; Roccasecca, R. M.; Acali, S.; Filocamo, G.; Traboni, C.; Nicosia, A.; Cortese, R.; Vitelli, A. *EMBO J.* **2002**, *21*, 5017–5025.

[75] Evans, M. J.; von Hahn, T.; Tscherne, D. M.; Syder, A. J.; Panis, M.; Wölk, B.; Hatziioannou, T.; McKeating, J. A.; Bieniasz, P. D.; Rice, C. M. *Nature* **2007**, *446*, 801–805.

[76] Ploss, A.; Evans, M. J.; Gaysinskaya, V. A.; Panis, M.; You, H.; de Jong, Y. P.; Rice, C. M. *Nature* **2009**, *457*, 882–886.

[77] Pawlotsky, J.-M.; Chevaliez, S.; McHutchison, J. G. *Gastroenterology* **2007**, *132*, 1979–1998.

[78] Bartosch, B.; Vitelli, A.; Granier, C.; Goujon, C.; Dubuisson, J.; Pascale, S.; Scarselli, E.; Cortese, R.; Nicosia, A.; Cosset, F.-L. *J. Biol. Chem.* **2003**, *278*, 41624–41630.

[79] Rehermann, B.; Nascimbeni, M. *Nat. Rev. Immunol.* **2005**, *5*, 215–229.

[80] Chevaliez, S.; Pawlotsky, J.-M. *Hepatitis C Viruses: Genomes and Molecular Biology*; Horizon Scientific Press, 2006; pp 5–47.

[81] Scheel, T. K. H.; Rice, C. M. *Nat. Med.* **2013**, *19*, 837–849.

[82] Moradpour, D.; Penin, F.; Rice, C. M. *Nat. Rev. Microbiol.* **2007**, *5*, 453–463.

[83] Welbourn, S.; Pause, A. *Curr. Issues Mol. Biol.* **2007**, *9*, 63–70.

[84] Satoh, S.; Tanji, Y.; Hijikata, M.; Kimura, K.; Shimotohno, K. *J. Virol.* **1995**, *69*, 4255–4260.

[85] Tanji, Y.; Hijikata, M.; Satoh, S.; Kaneko, T.; Shimotohno, K. *J. Virol.* **1995**, *69*, 1575–1581.

[86] Lin, C.; Rice, C. M. *Proc. Natl. Acad. Sci.* **1995**, *92*, 7622–7626.

[87] Tomei, L.; Failla, C.; Vitale, R. L.; Bianchi, E.; De Francesco, R. *J. Gen. Virol.* **1996**, *77*, 1065–1070.

[88] Yan, Y.; Li, Y.; Munshi, S.; Sardana, V.; Cole, J. L.; Sardana, M.; Kuo, L. C.; Chen, Z.; Steinkuehler, C.; Tomei, L.; Others, *Protein Sci.* **1998**, *7*, 837–847.

[89] Tomei, L.; Failla, C.; Santolini, E.; De Francesco, R.; La Monica, N. *J. Virol.* **1993**, *67*, 4017–4026.

[90] Bartenschlager, R.; Ahlborn-Laake, L.; Mous, J.; Jacobsen, H. *J. Virol.* **1994**, *68*, 5045–5055.

[91] Lin, C.; Pragai, B. M.; Grakoui, A.; Xu, J.; Rice, C. M. *J. Virol.* **1994**, *68*, 8147–8157.

[92] Yasunori, T.; Makoto, H.; Yuji, H.; Kunitada, S. *Gene* **1994**, *145*, 215–219.

[93] Bartenschlager, R. *J. Viral Hepat.* **1999**, *6*, 165–181.

[94] Lam, A. M. I.; Keeney, D.; Eckert, P. Q.; Frick, D. N. *J. Virol.* **2003**, *77*, 3950–3961.

[95] Frick, D. N. *Hepatitis C Viruses: Genomes and Molecular Biology*; Horizon Scientific Press, 2006; pp 207–244.

[96] Lam, A. M. I.; Frick, D. N. *J. Virol.* **2006**, *80*, 404–411.

[97] Beran, R. K. F.; Pyle, A. M. *J. Biol. Chem.* **2008**, *283*, 29929–29937.

[98] Kolykhalov, A. A.; Mihalik, K.; Feinstone, S. M.; Rice, C. M. *J. Virol.* **2000**, *74*, 2046–2051.

[99] Grakoui, A.; McCourt, D. W.; Wychowski, C.; Feinstone, S. M.; Rice, C. M. *J. Virol.* **1993**, *67*, 2832–2843.

[100] Grakoui, A.; McCourt, D. W.; Wychowski, C.; Feinstone, S. M.; Rice, C. M. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10583–10587.

[101] Hijikata, M.; Mizushima, H.; Tanji, Y.; Komoda, Y.; Hirowatari, Y.; Akagi, T.; Kato, N.; Kimura, K.; Shimotohno, K. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10773–10777.

[102] Pallaoro, M.; Lahm, A.; Biasiol, G.; Brunetti, M.; Nardella, C.; Orsatti, L.; Bonelli, F.; Orrù, S.; Narjes, F.; Steinkühler, C. *J. Virol.* **2001**, *75*, 9939–9946.

[103] Hijikata, M.; Mizushima, H.; Akagi, T.; Mori, S.; Kakiuchi, N.; Kato, N.; Tanaka, T.; Kimura, K.; Shimotohno, K. *J. Virol.* **1993**, *67*, 4665–4675.

[104] Welbourn, S.; Green, R.; Gamache, I.; Dandache, S.; Lohmann, V.; Bartenschlager, R.; Meerovitch, K.; Pause, A. *J. Biol. Chem.* **2005**, *280*, 29604–29611.

[105] Santolini, E.; Pacini, L.; Fipaldini, C.; Migliaccio, G.; Monica, N. *J. Virol.* **1995**, *69*, 7461–7471.

[106] Ryan, M. D.; Monaghan, S.; Flint, M. *J. Gen. Virol.* **1998**, *79*, 947–960.

[107] Kim, J. L.; Morgenstern, K. A.; Lin, C.; Fox, T.; Dwyer, M. D.; Landro, J. A.; Chambers, S. P.; Markland, W.; Lepre, C. A.; O'malley, E. T.; Harbeson, S. L.; Rice, C. M.; Murcko, M. A.; Caron, P. R.; Thomson, J. A. *Cell* **1996**, *87*, 343–355.

[108] Love, R. A.; Parge, H. E.; Wickersham, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Hostomska, Z. *Cell* **1996**, *87*, 331–342.

[109] Love, R. A.; Parge, H. E.; Wickersham, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Margosiak, S.; Dagostino, E.; Hostomska, Z. *Clin. Diagn. Virol.* **1998**, *10*, 151–156.

[110] Lackner, T.; Müller, A.; Pankraz, A.; Becher, P.; Thiel, H.-J.; Gorbalenya, A. E.; Tautz, N. *J. Virol.* **2004**, *78*, 10765–10775.

[111] Lorenz, I. C.; Marcotrigiano, J.; Dentzer, T. G.; Rice, C. M. *Nature* **2006**, *442*, 831–835.

[112] Stempniak, M.; Hostomska, Z.; Nodes, B. R.; Hostomsky, Z. *J. Virol.* **1997**, *71*, 2881–2886.

[113] De Francesco, R.; Urbani, A.; Nardi, M. C.; Tomei, L.; Steinkühler, C.; Tramontano, A. *Biochemistry* **1996**, *35*, 13282–13287.

[114] Jirasko, V.; Montserret, R.; Appel, N.; Janvier, A.; Eustachi, L.; Brohm, C.; Steinmann, E.; Pietschmann, T.; Penin, F.; Bartenschlager, R. *J. Biol. Chem.* **2008**, *283*, 28546–28562.

[115] Selby, M. J.; Glazer, E.; Masiarz, F.; Houghton, M. *Virology* **1994**, *204*, 114–122.

[116] Steinmann, E.; Penin, F.; Kallis, S.; Patel, A. H.; Bartenschlager, R.; Pietschmann, T. *PLoS Pathog.* **2007**, *3*, 962–971.

[117] Kiiver, K.; Merits, A.; Ustav, M.; Žusinaite, E. *Virus Res.* **2006**, *117*, 264–272.

[118] Flajolet, M.; Rotondo, G.; Daviet, L.; Bergametti, F.; Inchauspé, G.; Tiollais, P.; Transy, C.; Legrain, P. *Gene* **2000**, *242*, 369–379.

[119] Yi, M.; Ma, Y.; Yates, J.; Lemon, S. M. *PLoS Pathog.* **2009**, *5*, 1–15.

[120] Lemon, S. M.; McKeating, J. A.; Pietschmann, T.; Frick, D. N.; Glenn, J. S.; Tellinghuisen, T. L.; Symons, J.; Furman, P. A. *Antiviral Res.* **2010**, *86*, 79–92.

[121] Otto, H.-H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133–172.

[122] Patrick, G. L. *An Introduction to Medicinal Chemistry*; Oxford University Press, 2006.

[123] Pieroni, L.; Santolini, E.; Fipaldini, C.; Pacini, L.; Migliaccio, G.; La Monica, N. *J. Virol.* **1997**, *71*, 6373–6380.

[124] Orsatti, L.; Pallaoro, M.; Steinküler, C.; Orru, S.; Bonelli, F. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1919–1927.

[125] Gorzin, A. A.; Ramsland, P. A.; Tachedjian, G.; Gowans, E. J. *J. Viral Hepat.* **2012**, *19*, 189–198.

[126] Syder, A. J.; Lee, H.; Zeisel, M. B.; Grove, J.; Soulier, E.; Macdonald, J.; Chow, S.; Chang, J.; Baumert, T. F.; McKeating, J. A.; Others, *J. Hepatol.* **2011**, *54*, 48–55.

[127] Flisiak, R.; Feinman, S. V.; Jablkowski, M.; Horban, A.; Kryczka, W.; Pawlowska, M.; Heathcote, J. E.; Mazzella, G.; Vandelli, C.; Nicolas-Métral, V.; Grosgurin, P.; Liz, J. S.; Scalfaro, P.; Porchet, H.; Crabbe, R. *Hepatology* **2009**, *49*, 1460–1468.

[128] Fischer, G.; Gallay, P.; Hopkins, S. *Curr. Opin. Invest. Drugs* **2010**, *11*, 911–918.

[129] Jazwinski, A. B.; Muir, A. J. *Gastroenterol. Hepatol.* **2011**, *7*, 154–162.

[130] Belda, O.; Targett-Adams, P. *Virus Res.* **2012**, *170*, 1–14.

[131] Lok, A. S. *Clin Liver Dis* **2013**, *17*, 111–121.

[132] Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359–1472.

[133] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.

[134] Leung, D.; Abbenante, G.; Fairlie, D. P. *J. Med. Chem.* **2000**, *43*, 305–341.

[135] Lima, L. M.; Barreiro, E. J. *Curr. Med. Chem.* **2005**, *12*, 23–49.

[136] Brown, N. *Bioisosterism in Medicinal Chemistry*; John Wiley & Sons, 2012; pp 1–14.

[137] Kerns, E.; Di, L. *Drug-Like Properties: Concepts, Structure Design and Methods: From ADME to Toxicity Optimization*; Academic Press, 2010.

[138] Kim, J. L.; Morgenstern, K. A.; Griffith, J. P.; Dwyer, M. D.; Thomson, J. A.; Murcko, M. A.; Lin, C.; Caron, P. R. *Structure* **1998**, *6*, 89–100.

[139] Lin, K.; Perni, R. B.; Kwong, A. D.; Lin, C. *Antimicrob. Agents Chemother.* **2006**, *50*, 1813–1822.

[140] Hezode, C.; Forestier, N.; Dusheiko, G.; Ferenci, P.; Pol, S.; Goeser, T.; Bronowicki, J.-P.; Bourliere, M.; Gharakhanian, S.; Bengtsson, L.; McNair, L.; George, S.; Kieffer, T.; Kwong, A.; Kauffman, R. S.; Alam, J.; Pawlotsky, J.-M.; Zeuzem, S. *N. Engl. J. Med.* **2009**, *360*, 1839–1850.

[141] Lin, C.; Thomson, J. A.; Rice, C. M. *J. Virol.* **1995**, *69*, 4373–4380.

[142] Foy, E.; Li, K.; Wang, C.; Sumpter, R.; Ikeda, M.; Lemon, S. M.; Gale, M. *Science* **2003**, *300*, 1145–1148.

[143] Llinas-Brunet, M.; Bailey, M.; Fazal, G.; Goulet, S.; Halmos, T.; Laplante, S.; Maurice, R.; Poirier, M.; Poupart, M.-A.; Thibeault, D.; Wernic, D.; Lamarre, D. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1713–1718.

[144] Llinàs-Brunet, M.; Bailey, M.; Déziel, R.; Fazal, G.; Gorys, V.; Goulet, S.; Halmos, T.; Maurice, R.; Poirier, M.; Poupart, M.-A.; Rancourt, J.; Thibeault, D.; Wernic, D.; Lamarre, D. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2719–2724.

[145] Ingallinella, P.; Altamura, S.; Bianchi, E.; Taliani, M.; Ingenito, R.; Cortese, R.; De Francesco, R.; Steinkühler, C.; Pessi, A. *Biochemistry* **1998**, *37*, 8906–8914.

[146] Steinkühler, C.; Biasiol, G.; Brunetti, M.; Urbani, A.; Koch, U.; Cortese, R.; Pessi, A.; De Francesco, R. *Biochemistry* **1998**, *37*, 8899–8905.

[147] Llinàs-Brunet, M.; Bailey, M.; Fazal, G.; Ghiro, E.; Gorys, V.; Goulet, S.; Halmos, T.; Maurice, R.; Poirier, M.; Poupart, M.-A.; Rancourt, J.; Thibeault, D.; Wernic, D.; Lamarre, D. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2267–2270.

[148] Hinrichsen, H.; Benhamou, Y.; Wedemeyer, H.; Reiser, M.; Sentjens, R. E.; Calleja, J. L.; Forns, X.; Erhardt, A.; Crönlein, J.; Chaves, R. L.; Yong, C.-L.; Nehmiz, G.; Steinmann, G. G. *Gastroenterology* **2004**, *127*, 1347–1355.

[149] Jiang, Y.; Andrews, S. W.; Condroski, K.; Buckman, B.; Serebryany, V.; Wenglowsky, S.; Kennedey, A.; Madduru, M.; Wang, B.; Lyon, M.; Doherty, G. A.; Woodard, B. T.; Lemieux, C.; Geck Do, M.; Zhang, H.; Ballard, J.; Vigers, G.; Brandhuber, B. J.; Stengel, P.; Josey, J. A.; Beigelman, L.; Blatt, L.; Seiwert, S. D. *J. Med. Chem.* **2014**, *57*, 1753–1769.

[150] McCauley, J. A.; McIntyre, C. J.; Rudd, M. T.; Nguyen, K. T.; Romano, J. J.; Butcher, J. W.; Gilbert, K. F.; Bush, K. J.; Holloway, M. K.; Swestock, J.; Others, *J. Med. Chem.* **2010**, *53*, 2443–2463.

[151] Kiser, J. J.; Flexner, C. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 427–449.

[152] Darke, P. L.; Jacobs, A. R.; Waxman, L.; Kuo, L. C. *J. Biol. Chem.* **1999**, *274*, 34511–34514.

[153] Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.

[154] Shimizu, Y.; Yamaji, K.; Masuho, Y.; Yokota, T.; Inoue, H.; Sudo, K.; Satoh, S.; Shimotohno, K. *J. Virol.* **1996**, *70*, 127–132.

[155] Bartenschlager, R.; Lohmann, V.; Wilkinson, T.; Koch, J. O. *J. Virol.* **1995**, *69*, 7519–7528.

[156] Failla, C.; Tomei, L.; De Francesco, R. *J. Virol.* **1995**, *69*, 1769–1777.

[157] Hale, J. E. *Int. J. Proteomics* **2013**, 1–6.

[158] Hustoft, H. K.; Malerod, H.; Wilson, S. R.; Reubsaet, L.; Lundanes, E.; Greibrokk, T. *Proteomics* **2012**, 73–92.

[159] Kurokohchi, K.; Akatsuka, T.; Pendleton, C. D.; Takamizawa, A.; Nishioka, M.; Battegay, M.; Feinstone, S. M.; Berzofsky, J. A. *J. Virol.* **1996**, *70*, 232–240.

[160] Liefhebber, J. M. P.; Hensbergen, P. J.; Deelder, A. M.; Spaan, W. J. M.; Van Leeuwen, H. C. *J. Gen. Virol.* **2010**, *91*, 1013–1018.

[161] Hennion, M.-C. *J. Chromatogr. A* **1999**, *856*, 3–54.

[162] Poole, C. F. *Trends Anal. Chem.* **2003**, *22*, 362–373.

[163] Kushnir, M. M.; Rockwood, A. L.; Bergquist, J. *Mass Spectrom. Rev.* **2010**, *29*, 480–502.

[164] Grakoui, A.; Wychowski, C.; Lin, C.; Feinstone, S. M.; Rice, C. M. *J. Virol.* **1993**, *67*, 1385–1395.

[165] Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*; John Wiley & Sons, 2013.

[166] Wang, S.-S. *J. Am. Chem. Soc.* **1973**, *95*, 1328–1333.

[167] Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161–214.

[168] Merrifield, R. B. *Solid-phase Peptide Synthesis*; Springer, 1973; pp 335–361.

[169] Accessed May 27, 2014;. `http://www.piercenet.com/files/TR0049-Acetone-precipitation.pdf`.

[170] Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell, 1994*; Garland, New York, 1994; pp 139–194.

[171] Soto, C. *Nat. Rev. Neurosci.* **2003**, *4*, 49–60.

[172] Hartl, F. U.; Hayer-Hartl, M. *Science* **2002**, *295*, 1852–1858.

[173] Young, J. C.; Agashe, V. R.; Siegers, K.; Hartl, F. U. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 781–791.

[174] Herczenik, E.; Gebbink, M. F. B. G. *FASEB J.* **2008**, *22*, 2115–2133.

[175] Tsubuki, S.; Takai, Y.; Saido, T. C. *Lancet* **2003**, *361*, 1957–1958.

[176] Greenbaum, E. A.; Graves, C. L.; Mishizen-Eberz, A. J.; Lupoli, M. A.; Lynch, D. R.; Englander, S. W.; Axelsen, P. H.; Giasson, B. I. *J. Biol. Chem.* **2005**, *280*, 7800–7807.

[177] Westermark, P.; Benson, M. D.; Buxbaum, J. N.; Cohen, A. S.; Frangione, B.; Ikeda, S.-I.; Masters, C. L.; Merlini, G.; Saraiva, M. J.; Sipe, J. D. *Amyloid* **2005**, *12*, 1–4.

[178] Dobson, C. M. *Trends Biochem. Sci.* **1999**, *24*, 329–332.

[179] Dobson, C. M. *Nature* **2003**, *426*, 884–890.

[180] Mališauskas, M.; Zamotin, V.; Jass, J.; Noppe, W.; Dobson, C. M.; Morozova-Roche, L. A. *J. Mol. Biol.* **2003**, *330*, 879–890.

[181] Bemporad, F.; Taddei, N.; Stefani, M.; Chiti, F. *Protein Sci.* **2006**, *15*, 862–870.

[182] Johansson, J. *Swiss Med. Wkly.* **2003**, *133*, 275–282.

[183] Kelly, J. W. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101–106.

[184] Cohen, F. E.; Kelly, J. W. *Nature* **2003**, *426*, 905–909.

[185] Kelly, J. W. *Curr. Opin. Struct. Biol.* **1996**, *6*, 11–17.

[186] Rochet, J.-C.; Lansbury Jr, P. T. *Curr. Opin. Struct. Biol.* **2000**, *10*, 60–68.

[187] Westermark, P. *Ups. J. Med. Sci.* **2011**, *116*, 81–89.

[188] DeToma, A. S.; Salamekh, S.; Ramamoorthy, A.; Lim, M. H. *Chem. Soc. Rev.* **2012**, *41*, 608–621.

[189] Butterfield, S. M.; Lashuel, H. A. *Angew. Chem. Int. Ed.* **2010**, *49*, 5628–5654.

[190] Wilson, M. R.; Yerbury, J. J.; Poon, S. *Mol. Biosyst.* **2008**, *4*, 42–52.

[191] Harrison, R. S.; Sharpe, P. C.; Singh, Y.; Fairlie, D. P. *Reviews of Physiology, Biochemistry and Pharmacology*; Springer, 2007; pp 1–77.

[192] Salomon, A. R.; Marcinowski, K. J.; Friedland, R. P.; Zagorski, M. G. *Biochemistry* **1996**, *35*, 13568–13578.

[193] Longo Alves, I.; Hays, M. T.; Saraiva, M. J. M. *Eur. J. Biochem.* **1997**, *249*, 662–668.

[194] Soto, C.; Sigurdsson, E. M.; Morelli, L.; Kumar, R. A.; Castaño, E. M.; Frangione, B. *Nat. Med.* **1998**, *4*, 822–826.

[195] Permanne, B.; Adessi, C.; Saborio, G. P.; Fraga, S.; Frossard, M.-J.; Van Dorpe, J.; Dewachter, I.; Banks, W. A.; Van Leuven, F.; Soto, C. *FASEB J.* **2002**, *16*, 860–862.

[196] Soto, C.; Kascsak, R. J.; Saborío, G. P.; Aucouturier, P.; Wisniewski, T.; Prelli, F.; Kascsak, R.; Mendez, E.; Harris, D. A.; Ironside, J. *Lancet* **2000**, *355*, 192–197.

[197] Perrier, V.; Wallace, A. C.; Kaneko, K.; Safar, J.; Prusiner, S. B.; Cohen, F. E. *Proc. Natl. Acad. Sci.* **2000**, *97*, 6073–6078.

[198] Sacchettini, J. C.; Kelly, J. W. *Nat. Rev. Drug Discov.* **2002**, *1*, 267–275.

[199] Kisilevsky, R.; Lemieux, L. J.; Fraser, P. E.; Kong, X.; Hultin, P. G.; Szarek, W. A. *Nat. Med.* **1995**, *1*, 143–148.

[200] Schmitz, O.; Brock, B.; Rungby, J. *Diabetes* **2004**, *53*, 233–238.

[201] Kapurniotu, A.; Schmauder, A.; Tenidis, K. *J. Mol. Biol.* **2002**, *315*, 339–350.

[202] Yan, L.-M.; Tatarek-Nossol, M.; Velkova, A.; Kazantzis, A.; Kapurniotu, A. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 2046–2051.

[203] Lorenzo, A.; Yankner, B. A. *Proc. Natl. Acad. Sci.* **1994**, *91*, 12243–12247.

[204] Lee, V. M.-Y. *Neurobiol. Aging* **2002**, *23*, 1039–1042.

[205] Poli, G.; Ponti, W.; Carcassola, G.; Ceciliani, F.; Colombo, L.; Dall'Ara, P.; Gervasoni, M.; Giannino, M. L.; Martino, P. A.; Poller, C. *Arzneim.-Forsch.* **2003**, *53*, 875–888.

[206] Porat, Y.; Abramowitz, A.; Gazit, E. *Chem. Biol. Drug Des.* **2006**, *67*, 27–37.

[207] Turnell, W. G.; Finch, J. T. *J. Mol. Biol.* **1992**, *227*, 1205–1223.

[208] Porat, Y.; Mazor, Y.; Efrat, S.; Gazit, E. *Biochemistry* **2004**, *43*, 14454–14462.

[209] Bulic, B.; Pickhardt, M.; Khlistunova, I.; Biernat, J.; Mandelkow, E.-M.; Mandelkow, E.; Waldmann, H. *Angew. Chemie* **2007**, *119*, 9375–9379.

[210] Mishra, R.; Bulic, B.; Sellin, D.; Jha, S.; Waldmann, H.; Winter, R. *Angew. Chem. Int. Ed.* **2008**, *47*, 4679–4682.

[211] Kayed, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. *Science* **2003**, *300*, 486–489.

[212] Bourgault, S.; Choi, S.; Buxbaum, J. N.; Kelly, J. W.; Price, J. L.; Reixach, N. *Biochem. Biophys. Res. Commun.* **2011**, *410*, 707–713.

[213] Selkoe, D. J. *Nat. Med.* **2011**, *17*, 1060–1065.

[214] Bemporad, F.; Chiti, F. *Chem. Biol.* **2012**, *19*, 315–327.

[215] Sia, S. K.; Carr, P. A.; Cochran, A. G.; Malashkevich, V. N.; Kim, P. S. *Proc. Natl. Acad. Sci.* **2002**, *99*, 14664–14669.

[216] Oltersdorf, T. et al. *Nature* **2005**, *435*, 677–681.

[217] Yin, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163.

[218] Yin, H.; Lee, G.-I.; Hamilton, A. D. *Alpha-Helix Mimetics in Drug Discovery*; John Wiley & Sons, 2007; pp 281–299.

[219] Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 9391–9392.

[220] Yu, C.; Taylor, J. W. *Bioorg. Med. Chem.* **1999**, *7*, 161–175.

[221] Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 6966–6973.

[222] Flint, D. G.; Kumita, J. R.; Smart, O. S.; Woolley, G. A. *Chem. Biol.* **2002**, *9*, 391–397.

[223] Albert, J. S.; Goodman, M. S.; Hamilton, A. D. *J. Am. Chem. Soc.* **1995**, *117*, 1143–1144.

[224] Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479–2494.

[225] Olson, C. A.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2001**, *123*, 6451–6452.

[226] Tsou, L. K.; Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 14917–14921.

[227] Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233–1250.

[228] Horwell, D. C.; Howson, W.; Ratcliffe, G.; Willems, H. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2825–2830.

[229] Horwell, D. C.; Howson, W.; Ratcliffe, G. S.; Willems, H. M. G. *Bioorg. Med. Chem.* **1996**, *4*, 33–42.

[230] Yin, H.; Lee, G.-i.; Sedey, K. A.; Kutzki, O.; Park, H. S.; Orner, B. P.; Ernst, J. T.; Wang, H.-G.; Sebti, S. M.; Hamilton, A. D. *J. Am. Chem. Soc.* **2005**, *127*, 10191–10196.

[231] Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc.* **2002**, *124*, 11838–11839.

[232] Cummings, C. G.; Hamilton, A. D. *Curr. Opin. Chem. Biol.* **2010**, *14*, 341–346.

[233] Davis, J. M.; Truong, A.; Hamilton, A. D. *Org. Lett.* **2005**, *7*, 5405–5408.

[234] Biros, S. M.; Moisan, L.; Mann, E.; Carella, A.; Zhai, D.; Reed, J. C.; Rebek Jr, J. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4641–4645.

[235] Cummings, C. G.; Ross, N. T.; Katt, W. P.; Hamilton, A. D. *Org. Lett.* **2008**, *11*, 25–28.

[236] Saraogi, I.; Hebda, J. A.; Becerril, J.; Estroff, L. A.; Miranker, A. D.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2010**, *49*, 736–739.

[237] Kumar, S.; Miranker, A. D. *Chem. Commun.* **2013**, *49*, 4749–4751.

[238] Carey, J. S.; Laffan, D.; Thomson, C.; Williams, M. T. *Org. Biomol. Chem.* **2006**, *4*, 2337–2347.

[239] Roughley, S. D.; Jordan, A. M. *J. Med. Chem.* **2011**, *54*, 3451–3479.

[240] Heck, R. F. *J. Am. Chem. Soc.* **1968**, *90*, 5518–5526.

[241] Heck, R. F. *J. Am. Chem. Soc.* **1968**, *90*, 5526–5531.

[242] Fitton, P.; McKeon, J. E. *Chem. Commun.* **1968**, 4–6.

[243] Fitton, P.; Johnson, M. P.; McKeon, J. E. *Chem. Commun.* **1968**, 6–7.

[244] Mizoroki, T.; Mori, K.; Ozaki, A. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 581.

[245] Fitton, P.; Rick, E. A. *J. Organomet. Chem.* **1971**, *28*, 287–291.

[246] Heck, R. F.; Nolley Jr, J. P. *J. Org. Chem.* **1972**, *37*, 2320–2322.

[247] Mori, K.; Mizoroki, T.; Ozaki, A. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 1505–1508.

[248] AB, N. M. The Nobel Prize in Chemistry 2010 - Press Release. `http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2010/press.html`.

[249] Negishi, E.-i.; De Meijere, A.; Wiley, J. *Handbook of Organopalladium Chemistry for Organic Synthesis*; John Wiley & Sons, 2002.

[250] Wei, C. S.; Davies, G. H. M.; Soltani, O.; Albrecht, J.; Gao, Q.; Pathirana, C.; Hsiao, Y.; Tummala, S.; Eastgate, M. D. *Angew. Chem. Int. Ed.* **2013**, *52*, 5822–5826.

[251] Kimbrough, R. D. *Environ. Health Perspect.* **1976**, *14*, 51–56.

[252] Ackermann, L.; Vicente, R.; Kapdi, A. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 9792–9826.

[253] Ackermann, L. *Modern Arylation Methods*; Wiley-VCH Verlag, 2009.

[254] Lessene, G.; Feldman, K. S. *Oxidative Aryl-Coupling Reactions in Synthesis*; Wiley-VCH Verlag, 2002; pp 479–538.

[255] Cho, S. H.; Kim, J. Y.; Kwak, J.; Chang, S. *Chem. Soc. Rev.* **2011**, *40*, 5068–5083.

[256] Shi, Z.; Li, B.; Wan, X.; Cheng, J.; Fang, Z.; Cao, B.; Qin, C.; Wang, Y. *Angew. Chem. Int. Ed.* **2007**, *46*, 5554–5558.

[257] Wang, D.-H.; Mei, T.-S.; Yu, J.-Q. *J. Am. Chem. Soc.* **2008**, *130*, 17676–17677.

[258] Yang, S.; Li, B.; Wan, X.; Shi, Z. *J. Am. Chem. Soc.* **2007**, *129*, 6066–6067.

[259] Chen, X.; Goodhue, C. E.; Yu, J.-Q. *J. Am. Chem. Soc.* **2006**, *128*, 12634–12635.

[260] Giri, R.; Maugel, N.; Li, J.-J.; Wang, D.-H.; Breazzano, S. P.; Saunders, L. B.; Yu, J.-Q. *J. Am. Chem. Soc.* **2007**, *129*, 3510–3511.

[261] Yang, S.-D.; Sun, C.-L.; Fang, Z.; Li, B.-J.; Li, Y.-Z.; Shi, Z.-J. *Angew. Chem. Int. Ed.* **2008**, *120*, 1495–1498.

[262] Kawai, H.; Kobayashi, Y.; Oi, S.; Inoue, Y. *Chem. Commun.* **2008**, 1464–1466.

[263] Li, B.-J.; Yang, S.-D.; Shi, Z.-J. *Synlett* **2008**, *2008*, 949–957.

[264] Li, W.; Yin, Z.; Jiang, X.; Sun, P. *J. Org. Chem.* **2011**, *76*, 8543–8548.

[265] Moritanl, I.; Fujiwara, Y. *Tetrahedron Lett.* **1967**, *8*, 1119–1122.

[266] Fujiwara, Y.; Noritani, I.; Danno, S.; Asano, R.; Teranishi, S. *J. Am. Chem. Soc.* **1969**, *91*, 7166–7169.

[267] Jia, C.; Kitamura, T.; Fujiwara, Y. *Acc. Chem. Res.* **2001**, *34*, 633–639.

[268] Akermark, B.; Eberson, L.; Jonsson, E.; Pettersson, E. *J. Org. Chem.* **1975**, *40*, 1365–1367.

[269] Knölker, H.-J.; O'Sullivan, N. *Tetrahedron* **1994**, *50*, 10893–10908.

[270] Stuart, D. R.; Fagnou, K. *Inventing Reactions*; Springer, 2013; pp 91–119.

[271] Masui, K.; Ikegami, H.; Mori, A. *J. Am. Chem. Soc.* **2004**, *126*, 5074–5075.

[272] Takahashi, M.; Masui, K.; Sekiguchi, H.; Kobayashi, N.; Mori, A.; Funahashi, M.; Tamaoki, N. *J. Am. Chem. Soc.* **2006**, *128*, 10930–10933.

[273] Masuda, N.; Tanba, S.; Sugie, A.; Monguchi, D.; Koumura, N.; Hara, K.; Mori, A. *Org. Lett.* **2009**, *11*, 2297–2300.

[274] Hull, K. L.; Lanni, E. L.; Sanford, M. S. *J. Am. Chem. Soc.* **2006**, *128*, 14047–14049.

[275] Ames, D. E.; Bull, D. *Tetrahedron* **1982**, *38*, 383–387.

[276] Ames, D. E.; Opalko, A. *Synthesis* **1983**, *1983*, 234–235.

[277] Ames, D. E.; Opalko, A. *Tetrahedron* **1984**, *40*, 1919–1925.

[278] Campeau, L.-C.; Parisien, M.; Leblanc, M.; Fagnou, K. *J. Am. Chem. Soc.* **2004**, *126*, 9186–9187.

[279] Campeau, L.-C.; Thansandote, P.; Fagnou, K. *Org. Lett.* **2005**, *7*, 1857–1860.

[280] Ryabov, A. D. *Chem. Rev.* **1990**, *90*, 403–424.

[281] Lafrance, M.; Shore, D.; Fagnou, K. *Org. Lett.* **2006**, *8*, 5097–5100.

[282] Davies, D. L.; Donald, S. M. A.; Macgregor, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 13754–13755.

[283] Lafrance, M.; Fagnou, K. *J. Am. Chem. Soc.* **2006**, *128*, 16496–16497.

[284] Lafrance, M.; Gorelsky, S. I.; Fagnou, K. *J. Am. Chem. Soc.* **2007**, *129*, 14570–14571.

[285] Gorelsky, S. I.; Lapointe, D.; Fagnou, K. *J. Am. Chem. Soc.* **2008**, *130*, 10848–10849.

[286] Forgione, P.; Brochu, M.-C.; St-Onge, M.; Thesen, K. H.; Bailey, M. D.; Bilodeau, F. *J. Am. Chem. Soc.* **2006**, *128*, 11350–11351.

[287] Roger, J.; Požgan, F.; Doucet, H. *Green Chem.* **2009**, *11*, 425–432.

[288] Bheeter, C. B.; Bera, J. K.; Doucet, H. *J. Org. Chem.* **2011**, *76*, 6407–6413.

[289] Dong, J. J.; Doucet, H. *Eur. J. Org. Chem.* **2010**, *2010*, 611–615.

[290] Nilsson, M. *Acta Chem. Scand.* **1966**, *20*, 423–426.

[291] Heim, A.; Terpin, A.; Steglich, W. *Angew. Chemie Int. Ed.* **1997**, *36*, 155–156.

[292] Peschko, C.; Winklhofer, C.; Steglich, W. *Chem. Eur. J.* **2000**, *6*, 1147–1152.

[293] Myers, A. G.; Tanaka, D.; Mannion, M. R. *J. Am. Chem. Soc.* **2002**, *124*, 11250–11251.

[294] Tanaka, D.; Myers, A. G. *Org. Lett.* **2004**, *6*, 433–436.

[295] Gooßen, L. J.; Deng, G.; Levy, L. M. *Science* **2006**, *313*, 662–664.

[296] Gooßen, L. J.; Zimmermann, B.; Knauber, T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7103–7106.

[297] Gooßen, L. J.; Rodríguez, N.; Lange, P. P.; Linder, C. *Angew. Chem. Int. Ed.* **2010**, *49*, 1111–1114.

[298] Gooßen, L. J.; Rodriguez, N.; Linder, C. *J. Am. Chem. Soc.* **2008**, *130*, 15248–15249.

[299] Cornella, J.; Larrosa, I. *Synthesis* **2012**, *2012*, 653–676.

[300] Becht, J.-M.; Catala, C.; Le Drian, C.; Wagner, A. *Org. Lett.* **2007**, *9*, 1781–1783.

[301] Becht, J.-M.; Drian, C. L. *Org. Lett.* **2008**, *10*, 3161–3164.

[302] Wang, Z.; Ding, Q.; He, X.; Wu, J. *Tetrahedron* **2009**, *65*, 4635–4638.

[303] Bilodeau, F.; Brochu, M.-C.; Guimond, N.; Thesen, K. H.; Forgione, P. *J. Org. Chem.* **2010**, *75*, 1550–1560.

[304] Abedini, A.; Raleigh, D. P. *Phys. Biol.* **2009**, *6*, 1–6.

[305] Abedini, A.; Raleigh, D. P. *Protein Eng. Des. Sel.* **2009**, *22*, 453–459.

[306] Knight, J. D.; Hebda, J. A.; Miranker, A. D. *Biochemistry* **2006**, *45*, 9496–9508.

[307] Apostolidou, M.; Jayasinghe, S. A.; Langen, R. *J. Biol. Chem.* **2008**, *283*, 17205–17210.

[308] Carufel, C. A.; Nguyen, P. T.; Sahnouni, S.; Bourgault, S. *Pept. Sci.* **2013**, *100*, 645–655.

[309] Jayasinghe, S. A.; Langen, R. *Biochemistry* **2005**, *44*, 12113–12119.

[310] Orner, B. P.; Ernst, J. T.; Hamilton, A. D. *J. Am. Chem. Soc.* **2001**, *123*, 5382–5383.

[311] Patil, S. M.; Xu, S.; Sheftic, S. R.; Alexandrescu, A. T. *J. Biol. Chem.* **2009**, *284*, 11982–11991.

[312] Bey, E.; Marchais-Oberwinkler, S.; Negri, M.; Kruchten, P.; Oster, A.; Klein, T.; Spadaro, A.; Werth, R.; Frotscher, M.; Birk, B.; Others, *J. Med. Chem.* **2009**, *52*, 6724–6743.

[313] Cardullo, F.; Donati, D.; Fusillo, V.; Merlo, G.; Paio, A.; Salaris, M.; Solinas, A.; Taddei, M. *J. Comb. Chem.* **2006**, *8*, 834–840.

[314] Amaladass, P.; Clement, J. A.; Mohanakrishnan, A. K. *Tetrahedron* **2007**, *63*, 10363–10371.

[315] Apperloo, J. J.; Groenendaal, L.; Verheyen, H.; Jayakannan, M.; Janssen, R. A. J.; Dkhissi, A.; Beljonne, D.; Lazzaroni, R.; Brédas, J.-L.; Others, *Chem. Eur. J.* **2002**, *8*, 2384–2396.

[316] Dondoni, A.; Fogagnolo, M.; Medici, A.; Negrini, E. *Synthesis* **1987**, *1987*, 185–186.

[317] Oster, A.; Hinsberger, S.; Werth, R.; Marchais-Oberwinkler, S.; Frotscher, M.; Hartmann, R. W. *J. Med. Chem.* **2010**, *53*, 8176–8186.

[318] Wesch, T.; Berthelot-Brehier, A.; Leroux, F. R.; Colobert, F. *Org. Lett.* **2013**, *15*, 2490–2493.

[319] Beydoun, K.; Boixel, J.; Guerchais, V.; Doucet, H. *Catal. Sci. Technol.* **2012**, *2*, 1242–1248.

[320] Fu, H. Y.; Chen, L.; Doucet, H. *J. Org. Chem.* **2012**, *77*, 4473–4478.

[321] Della Ca, N.; Maestri, G.; Catellani, M. *Chem. Eur. J.* **2009**, *15*, 7850–7853.

[322] Goikhman, R.; Jacques, T. L.; Sames, D. *J. Am. Chem. Soc.* **2009**, *131*, 3042–3048.

[323] Bouazizi, Y.; Beydoun, K.; Romdhane, A.; Ben Jannet, H.; Doucet, H. *Tetrahedron Lett.* **2012**, *53*, 6801–6805.

[324] Chen, X.; Engle, K. M.; Wang, D.-H.; Yu, J.-Q. *Angew. Chem. Int. Ed.* **2009**, *48*, 5094–5115.

[325] Chu, J.-H.; Wu, C.-C.; Chang, D.-H.; Lee, Y.-M.; Wu, M.-J. *Organometallics* **2012**, *32*, 272–282.

[326] Mei, T.-S.; Kou, L.; Ma, S.; Engle, K. M.; Yu, J.-Q. *Synthesis* **2012**, *44*, 1778–1791.

[327] Peng, J.; Liu, L.; Hu, Z.; Huang, J.; Zhu, Q. *Chem. Commun.* **2012**, *48*, 3772–3774.

[328] Derridj, F.; Gottumukkala, A. L.; Djebbar, S.; Doucet, H. *Eur. J. Inorg. Chem.* **2008**, *2008*, 2550–2559.

[329] Gottumukkala, A. L.; Doucet, H. *Eur. J. Inorg. Chem.* **2007**, *2007*, 3629–3632.

[330] Chiong, H. A.; Daugulis, O. *Org. Lett.* **2007**, *9*, 1449–1451.

[331] Campeau, L.-C.; Schipper, D. J.; Fagnou, K. *J. Am. Chem. Soc.* **2008**, *130*, 3266–3267.

[332] Liégault, B.; Petrov, I.; Gorelsky, S. I.; Fagnou, K. *J. Org. Chem.* **2010**, *75*, 1047–1060.

[333] Gorelsky, S. I.; Lapointe, D.; Fagnou, K. *J. Org. Chem.* **2011**, *77*, 658–668.

[334] Schipper, D. J.; Fagnou, K. *Chem. Mater.* **2011**, *23*, 1594–1600.

[335] Fu, H. Y.; Doucet, H. *Eur. J. Org. Chem.* **2011**, *2011*, 7163–7173.

[336] Bheeter, C. B.; Bera, J. K.; Doucet, H. *RSC Adv.* **2012**, *2*, 7197–7206.

[337] Yamashita, M.; Hirano, K.; Satoh, T.; Miura, M. *Org. Lett.* **2009**, *12*, 592–595.

[338] Jafarpour, F.; Zarei, S.; Barzegar Amiri Olia, M.; Jalalimanesh, N.; Rahiminejadan, S. *J. Org. Chem.* **2013**, *78*, 2957–2964.

[339] Song, B.; Knauber, T.; Gooßen, L. J. *Angew. Chem. Int. Ed.* **2013**, *52*, 2954–2958.

[340] Kim, M.; Park, J.; Sharma, S.; Kim, A.; Park, E.; Kwak, J. H.; Jung, Y. H.; Kim, I. S. *Chem. Commun.* **2013**, *49*, 925–927.

[341] Hu, P.; Shang, Y.; Su, W. *Angew. Chem. Int. Ed.* **2012**, *51*, 5945–5949.

[342] Messaoudi, S.; Brion, J.-D.; Alami, M. *Org. Lett.* **2012**, *14*, 1496–1499.

[343] Haley, C. K.; Gilmore, C. D.; Stoltz, B. M. *Tetrahedron* **2013**, *69*, 5732–5736.

[344] Mitchell, D.; Coppert, D. M.; Moynihan, H. A.; Lorenz, K. T.; Kissane, M.; McNamara, O. A.; Maguire, A. R. *Org. Process Res. Dev.* **2011**, *15*, 981–985.

[345] Collet, F.; Song, B.; Rudolphi, F.; Gooßen, L. J. *Eur. J. Org. Chem.* **2011**, *2011*, 6486–6501.

[346] Liégault, B.; Lapointe, D.; Caron, L.; Vlassova, A.; Fagnou, K. *J. Org. Chem.* **2009**, *74*, 1826–1834.

[347] Wong, N. W. Y.; Forgione, P. *Org. Lett.* **2012**, *14*, 2738–2741.

[348] Khurana, R.; Coleman, C.; Ionescu-Zanetti, C.; Carter, S. A.; Krishna, V.; Grover, R. K.; Roy, R.; Singh, S. *J. Struct. Biol.* **2005**, *151*, 229–238.

[349] Krebs, M. R. H.; Bromley, E. H. C.; Donald, A. M. *J. Struct. Biol.* **2005**, *149*, 30–37.

[350] Naiki, H.; Higuchi, K.; Hosokawa, M.; Takeda, T. *Anal. Biochem.* **1989**, *177*, 244–249.

Appendix A

Supporting Information

Synthesis of 2,5-Diaryl Substituted Thiophenes as Helical Mimetics: Towards the Modulation of Islet Amyloid Polypeptide (IAPP) Amyloid Fibril Formation and Cytotoxicity

General Conditions and Instrumentation

Reactions were carried out in flame-dried glassware under an argon atmosphere unless otherwise noted. Commercially available chemicals were purchased from Aldrich and Alfa Aesar and used without further purification. Bis(tri-*tert*-butylphosphine)palladium (0) and palladium (II) acetate were stored under inert gas. All solvents were purchased from Fisher Scientific or JT Baker as ACS grade. Unless stated otherwise, solvents were dried using and stored over activated 3 Å molecular sieves in a flame-dried Schlenk flask. Distilled water was obtained from an in-house distillery. Compounds were purified using column chromatography on silica-gel (Zeoprep 60 Eco, 40 – 63 μm, Zeochem AG).
Microwave assisted reactions were carried out using the Biotage Initiator™ 2.3 build 6250 microwave system with a 400 W magnetron. ^{1}H and ^{13}C NMR data were measured on a Varian VNMRS-500 (500 MHz ^{1}H NMR and 125 MHz ^{13}C NMR) in chloroform-d or dimethyl sulfoxide-d_6. ^{1}H and ^{13}C NMR spectra were referred to residual solvent peaks. The chemical shifts are reported in parts per million (ppm), followed in parentheses by the multiplicity of the signals (s = singlet, d = doublet, dd = doublet of doublets, ddd = doublet of doublet of doublets, t = triplet, q = quartet and m = multiplet), followed by the number of protons and coupling constants *J* (Hz). High-resolution mass spectral data (HRMS) were collected using a LC-TOF ESI mass spectrometer operated in positive ion mode (unless stated otherwise).

Experimental Procedures

Procedure for the synthesis of methyl 3-(methylsulfonamido) thiophene-2-carboxylate (compound 87)

The procedure employed by Tondi and co-workers was used with some modifications.[1] Methyl 3-amino-thiophene-2-carboxylate (1 equiv) was dissolved in pyridine (0.8M of heterocycle solution) and methanesulfonyl chloride (1.5 equiv) was added to the stirred solution. The reaction was heated for 1 hour at 50 °C. The reaction mixture was then diluted with EtOAc and washed with water. The aqueous phase was extracted with EtOAc and the combined organic phases were washed 5 times with water. The organic phase was dried over anhydrous sodium sulphate, filtered and concentrated. The crude solid was recrystallized from EtOAc and hexanes to provide the title compound as light brown crystals in 90% yield.

Procedure for the synthesis of methyl 3-(N-methylmethylsulfonamido) thiophene-2-carboxylate (compound 88)

The procedure employed by Cardullo and co-workers was used for the methylation of the secondary amine with some modifications.[2] The sulfonamide prepared above (1 equiv) was dissolved in anhydrous DMF (0.4 M of sulfonamide solution) and Cs_2CO_3 (1 equiv) and MeI (2.5 equiv) were added to the solution. The mixture was stirred for 24 hours at 40 °C. Subsequent to filtering the solution, the solvent was evaporated under reduced pressure. The crude material was dissolved in anhydrous chloroform and stirred for 1 hour at room temperature. The mixture was filtered and the solvent was evaporated to give the title compound as a white solid in 95% yield. The compound was used without further purification.

General procedure for saponification

The methyl ester (1 equiv) was dissolved in a 1:1:2 mixture of 2 M $NaOH_{(aq.)}$ (5 equiv):MeOH:THF, and the mixture was refluxed for 1 hour at 80 °C. The solution was diluted with EtOAc and acidified with HCl (1M) to bring the pH to 3 or 4. The aqueous

phase was extracted with additional EtOAc and the combined organic phases were washed with water (3x). The solution was dried over anhydrous sodium sulphate, filtered, and the solvent was evaporated. The compound was used without further purification.

General procedure for decarboxylative cross-coupling

The procedure employed by Forgione and co-workers was used with slight modifications.[3] In a 2-5 mL, open to air, oven dried microwave vial were added the heterocyclic carboxylic acid (2 equiv), aryl bromide (1 equiv), tetra-*n*-butylammonium chloride (1 equiv), cesium carbonate (1.5 equiv), Bis(tri-*tert*-butylphosphine)palladium (0) (0.05 equiv) and anhydrous DMF (0.1 M of the aryl bromide solution). The vial was capped with a septum and the mixture was pre-stirred for 30 seconds at 23 °C and submitted to microwave heating at 170 °C for 8 min with stirring and the high absorption setting. The crude mixture was cooled to 23 °C and was filtered over Celite®. The solution was then diluted with EtOAc and the organic layer was washed with a saturated NaCl aqueous solution (3x), saturated $NaHCO_3$ aqueous solution (2x), water (1x), and saturated NaCl aqueous solution (1x). The aqueous phases were combined and extracted with EtOAc. The combined organic phases were dried over sodium sulfate, and after filtration the solvent was evaporated to provide the crude compound.

General procedure for C-H activation

A procedure employed by Fagnou and co-workers was used with slight modifications.[4] An oven dried vial equipped with a magnetic stir bar was charged with heterocycle (1 equiv), aryl bromide (2 equiv), $PCy_3.HBF_4$ (0.1 equiv), PivOH (0.3 equiv), K_2CO_3 (1.5 equiv), and palladium (II) acetate (0.05 equiv). Anhydrous DMA (0.08 M of the heterocycle solution) was added. Liquid aryl bromides were added after the addition of solvent. The mixture was heated for 16 hours at 100 °C. After cooling to 23 °C, the reaction mixture was diluted with EtOAc and filtered through a pad of Celite®. The filtrate was washed with a saturated NaCl aqueous solution (3x), saturated $NaHCO_3$ aqueous solution (2x) (unless otherwise stated), water (1x), and saturated NaCl aqueous solution (1x). The aqueous phases were combined and extracted with EtOAc. The

combined organic phases were dried over sodium sulfate, and after filtration the solvent was evaporated to provide the crude compound.

X-ray crystallography of compound 99n

A colorless rhomb-like specimen of $C_{20}H_{18}N_2O_3S_2$, approximate dimensions 0.356 mm x 0.371 mm x 0.438 mm, was used for the X-ray crystallographic analysis. The X-ray intensity data were measured.

A total of 1464 frames were collected. The total exposure time was 4.07 hours. The frames were integrated with the Bruker SAINT software package using a narrow-frame algorithm. The integration of the data using an orthorhombic unit cell yielded a total of 18642 reflections to a maximum θ angle of 28.97° (0.73 Å resolution), of which 4493 were independent (average redundancy 4.149, completeness = 93.3%, R_{int} = 4.56%) and 3794 (84.44%) were greater than $2\sigma(F^2)$. The final cell constants of $\underline{a}$ = 10.862(3) Å, $\underline{b}$ = 9.289(3) Å, $\underline{c}$ = 18.357(5) Å, volume = 1852.2(9) $Å^3$, are based upon the refinement of the XYZ-centroids of 5458 reflections above 20 σ(I) with 4.437° < 2θ < 52.96°. Data were corrected for absorption effects using the multi-scan method (SADABS). The ratio of minimum to maximum apparent transmission was 0.802. The calculated minimum and maximum transmission coefficients (based on crystal size) are 0.8760 and 0.8970.

The structure was solved and refined using the Bruker SHELXTL Software Package, using the space group P n a 2_1, with Z = 4 for the formula unit, $C_{20}H_{18}N_2O_3S_2$. The final anisotropic full-matrix least-squares refinement on F^2 with 247 variables converged at R1 = 3.67%, for the observed data and wR_2 = 8.54% for all data. The goodness-of-fit was 1.043. The largest peak in the final difference electron density synthesis was 0.201 $e^-/Å^3$ and the largest hole was -0.315 $e^-/Å^3$ with an RMS deviation of 0.048 $e^-/Å^3$. On the basis of the final model, the calculated density was 1.429g/cm^3 and F(000), 832 e^-.

Preparation and Characterization of Compounds

_N_-(2-(3-cyanophenyl)thiophen-3-yl)-_N_-methylmethanesulfonamide (98a; Table 6.1, Entry 1)

General procedure 2 was followed using 3-bromobenzonitrile (38.5 mg, 0.21 mmol) and compound **97** to yield the title compound as a yellow solid (36 mg, 58%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 (ddd, J = 7.8, 1.9, 1.2 Hz, 1H), 7.87 - 8.86 (m, 1H), 7.64 (ddd, J = 7.8, 1.4, 1.2 Hz, 1H), 7.54 (dd, J = 7.8, 7.8 Hz, 1H), 7.39 (d, J = 5.4 Hz, 1H), 7.12 (d, J = 5.4 Hz, 1H), 3.21 (s, 3H), 2.87 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 138.1, 135.2, 133.8, 132.7, 131.8, 131.6, 129.7, 125.5, 125.3, 118.4, 113.1, 38.8, 36.8.

HRMS (EI): calculated for $C_{13}H_{12}N_2O_2S_2$ $[M + H]^+$: 293.0413, found: 293.0414.

_N_-(2-(3-methoxyphenyl) thiophen-3-yl)-_N_-methylmethanesulfonamide (98b; Table 6.1, Entry 2)

General procedure 2 was followed using 1-bromo-3-methoxybenzene (26.5 μL, 0.21 mmol) and compound **97** to yield the title compound as a light yellow solid (43 mg, 68%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32 (dd, J = 8.0, 8.0 Hz, 1H), 7.27 (d, J = 5.4 Hz, 1H), 7.22 (dd, J = 2.6, 1.7 Hz, 1H), 7.17 (ddd, J = 8.0, 1.7, 0.9 Hz, 1H), 7.08 (d, J = 5.4 Hz, 1H), 6.90 (ddd, J = 8.0, 2.6, 0.9 Hz, 1H), 3.85 (s, 3H), 3.19 (s, 3H), 2.81 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 159.8, 140.1, 134.2, 133.5, 129.8, 126.4, 123.9, 120.8, 114.2, 113.8, 55.4, 38.5, 38.1.

HRMS (EI): calculated for $C_{13}H_{15}NO_3S_2$ $[M + H]^+$: 298.0566, found: 298.0561.

Ethyl 3-(3-(_N_-methylmethylsulfonamido)thiophen-2-yl)benzoate (98c; Table 6.1, Entry 3)

General procedure 2 was followed using ethyl 3-bromobenzoate (33.6 μL, 0.21 mmol) and compound **97** to yield the title compound as a brown solid (47 mg, 66%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.27 (dd, *J* = 2.2, 1.6 Hz, 1H), 8.04 - 8.02 (m, 1H), 7.84 (ddd, *J* = 7.8, 1.8, 1.1 Hz, 1H), 7.51 (dd, *J* = 7.8, 7.8 Hz, 1H), 7.33 (d, *J* = 5.5 Hz, 1H), 7.12 (d, *J* = 5.5 Hz, 1H), 4.40 (q, *J* = 7.1 Hz, 2H), 3.20 (s, 3H), 2.84 (s, 3H), 1.41 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.1, 139.1, 134.7, 132.7, 132.6, 131.1, 129.4, 129.3, 129.0, 126.4, 124.4, 61.2, 38.6, 37.8, 14.3.

HRMS (EI): calculated for $C_{15}H_{17}NO_4S_2$ $[M + H]^+$: 340.0672, found: 340.0664.

3-(3-methylthiophen-2-yl)benzonitrile (98d; Table 6.1, Entry 4)

General procedure 2 was followed using 2-bromobenzonitrile (63.7 mg, 0.35 mmol) and 3-methylthiophene-2-carboxylic acid to yield the title compound as a light yellow oil (48 mg, 69%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.74 - 7.73 (m, 1H), 7.69 (ddd, *J* = 7.8, 1.9, 1.2 Hz, 1H), 7.59 (ddd, *J* = 8.0, 1.4, 1.4 Hz, 1H), 7.51 (ddd, *J* = 8.0, 7.5, 0.5 Hz, 1H), 7.27 (d, *J* = 5.2 Hz, 1H), 6.95 (d, *J* = 5.2 Hz, 1H), 2.33 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 136.2, 135.1, 134.5, 133.2, 132.2, 131.4, 130.5, 129.4, 124.7, 118.6, 112.8, 14.9.

HRMS (EI): calculated for $C_{12}H_9NS$ $[M + H]^+$: 200.0528, found: 200.0525.

2-(3-methoxyphenyl)-3-methylthiophene (98e; Table 6.1, Entry 5)

General procedure 2 was followed using 1-bromo-3-methoxybenzene (44.2 μL, 0.35 mmol) and 3-methylthiophene-2-carboxylic acid to yield the title compound as a colorless oil (54 mg, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34 - 7.30 (m, 1H), 7.20 (d, *J* = 5.1 Hz, 1H), 7.07 - 7.04 (m, 1H), 7.02 - 7.00 (m, 1H), 6.92 (d, *J* = 5.1 Hz, 1H), 6.88 - 6.86 (m, 1H), 3.84 (s, 3H), 2.34 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 159.6, 137.7, 136.1, 133.3, 131.1, 129.5, 123.4, 121.5, 114.6, 112.7, 55.3, 15.0.

HRMS (EI): calculated for $C_{12}H_{12}OS$ $[M + H]^+$: 205.0682, found: 205.0682.

Ethyl 3-(3-methylthiophen-2-yl) benzoate (98f; Table 6.1, Entry 6)

General procedure 2 was followed using 1-bromo-3-methoxybenzene (56.1 μL, 0.35 mmol) and 3-methylthiophene-2-carboxylic acid to yield the title compound as a colorless oil (53 mg, 61%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.15 - 8.14 (m, 1H), 8.01 - 7.98 (dm, *J* =7.8 Hz, 1H), 7.64 (ddd, *J* = 7.8, 1.9, 1.3 Hz, 1H), 7.48 (dd, *J* = 7.8, 7.8 Hz, 1H), 7.24 (d, *J* = 5.1 Hz, 1H), 6.94 (d, *J* = 5.1 Hz, 1H), 4.40 (q, *J* = 7.1 Hz, 2H), 2.33 (s, 3H), 1.40 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.4, 136.7, 135.1, 133.8, 133.2, 131.1, 130.8, 130.0, 128.5, 128.2, 123.9, 61.1, 14.9, 14.3.

HRMS (EI): calculated for $C_{14}H_{14}O_2S$ $[M + H]^+$: 247.0787, found: 247.0785.

2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzonitrile (99c; Table 6.2, Entry 7)

General procedure 3 was followed using compound **98e** (120 mg, 0.58 mmol) and 2-bromobenzonitrile to yield the title compound as a yellow oil (41 mg, 23%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.74 - 7.74 (m, 1H), 7.64 - 7.62 (m, 1H), 7.58 (ddd, *J* = 8.0, 7.4, 1.5 Hz, 1H), 7.50 (s, 1H), 7.38 - 7.33 (m, 2H), 7.10 (ddd, *J* = 8.0, 1.5, 0.9 Hz, 1H), 7.04 (dd, *J* = 2.5, 1.5 Hz, 1H), 6.90 (ddd, *J* = 8.0, 2.5, 0.9 Hz, 1H), 3.86 (s, 3H), 2.38 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 159.7, 140.0, 137.5, 136.9, 135.3, 134.6, 134.4, 133.0, 131.4, 129.6, 129.3, 127.3, 121.5, 119.0, 114.6, 113.3, 109.5, 55.3, 15.2.

HRMS (EI): calculated for $C_{19}H_{15}NOS$ $[M + H]^+$: 306.0947, found: 306.0948.

2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzaldehyde (99d; Table 6.2, Entry 8)

General procedure 3 was followed using compound **98e** (40 mg, 0.2 mmol) and 2-bromobenzaldehyde to yield the title compound as an orange solid (49 mg, 79%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.31 (s, 1H), 8.01 (ddd, *J* = 7.5, 1.5, 0.6 Hz, 1H), 7.64 - 7.60 (m, 1H), 7.59 -7.57 (m, 1H), 7.49 - 7.46 (m, 1H), 7.35 (dd, *J* = 8.0, 8.0 Hz, 1H), 7.10

(ddd, J = 7.8, 1.6, 0.9 Hz, 1H), 7.05 (dd, J = 2.5, 1.6 Hz, 1H), 6.91 (ddd, J = 7.8, 2.5, 0.9 Hz, 1H), 6.89 (s, 1H), 3.86 (s, 3H), 2.38 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 192.2, 159.7, 140.1, 138.0, 136.5, 135.3, 134.1, 134.0, 133.7, 133.6, 131.0, 129.7, 128.0, 127.9, 121.4, 114.5, 113.2, 55.3, 15.1.

HRMS (EI): calculated for $C_{19}H_{16}O_2S$ $[M + H]^+$: 309.0944, found: 309. 0937.

2-(3-methoxyphenyl)-3-methyl-5-(2-(trifluoromethyl)phenyl)thiophene (99e; Table 6.2, Entry 9)

General procedure 3 was followed using compound **98e** (40 mg, 0.2 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a colorless oil (57 mg, 81%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.75 (d, J = 7.5 Hz, 1H), 7.55 - 7.54 (m, 2H), 7.47 - 7.54 (m, 1H), 7.33 (dd, J = 8.0, 8.0 Hz, 1H), 7.10 (ddd, J = 7.5, 1.6, 0.9 Hz, 1H), 7.05 (dd, J = 2.5, 1.6 Hz, 1H), 6.95 (s, 1H), 6.88 (ddd, J = 8.0, 2.5, 0.9 Hz, 1H), 3.86 (s, 3H), 2.36 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 159.6, 138.7, 137.5, 135.6, 133.7, 133.2, 133.0, 132.8, 131.8, 131.4, 129.8 (q, J_{CF} = 30.0 Hz), 127.8, 126.5 (q, J_{CF} = 5.0 Hz), 124.0 (q, J_{CF} = 274.0 Hz), 121.5, 114.5, 113.0, 55.3, 15.2.

HRMS (EI): calculated for $C_{19}H_{15}F_3OS$ $[M + H]^+$: 349.0868, found: 349.0875.

Ethyl 2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzoate (99f; Table 6.2, Entry 10)

General procedure 3 was followed using compound **98e** (11 mg, 0.05 mmol) and ethyl 2-bromobenzoate to yield the title compound as a yellow oil (13.6 mg, 73%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.71 (d, J = 7.5 Hz, 1H), 7.52 - 7.46 (m, 2H), 7.40 - 7.36 (m, 1H), 7.33 (dd, J = 8.0, 8.0 Hz, 1H), 7.08 (ddd, J = 7.6, 1.5, 0.9 Hz, 1H), 7.03 (dd, J = 2.6, 1.5 Hz, 1H), 6.89 - 6.86 (m, 2H), 4.25 (q, J = 7.0 Hz, 2H), 3.85 (s, 3H), 2.34 (s, 3H), 1.20 (t, J = 7.0 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 168.9, 159.6, 139.9, 138.3, 135.8, 134.0, 133.4, 132.1, 130.9, 130.8, 130.5, 129.5, 129.4, 127.6, 121.4, 114.5, 112.8, 61.3, 55.3, 15.1, 13.9.

HRMS (EI): calculated for $C_{21}H_{20}O_3S$ $[M + H]^+$: 353.1206, found: 353.1207.

Ethyl 2-(5-(3-(ethoxycarbonyl)phenyl)-4-methylthiophen-2-yl)benzoate (99g; Table 6.2, Entry 11)

General procedure 3 was followed using compound **98f** (38 mg, 0.15 mmol) and ethyl 2-bromobenzoate to yield the title compound as a colorless oil (55 mg, 94%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.17 (dd, *J* = 1.6, 1.6 Hz, 1H), 8.00 (ddd, *J* = 7.8, 1.4, 1.3 Hz, 1H), 7.73 (d, *J* = 7.5 Hz, 1H), 7.66 (ddd, *J* = 7.8, 1.5, 1.4 Hz, 1H), 7.52 - 7.47 (m, 3H), 7.41 - 7.38 (m, 1H), 6.88 (s, 1H), 4.40 (q, *J* = 7.1, 2H), 4.27 (q, *J* = 7.1, 2H), 2.34 (s, 3H), 1.41 (t, *J* = 7.1, 3H), 1.22 (t, *J* = 7.1, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 168.8, 166.4, 140.4, 137.3, 134.8, 133.9, 133.8, 133.1, 132.0, 130.92, 130.89 (2C), 130.5, 129.9, 129.4, 128.6, 128.2, 127.8, 61.3, 61.1, 15.0, 14.3, 13.9.

HRMS (EI): calculated for $C_{23}H_{22}O_4S$ $[M + H]^+$: 395.1311, found: 395.1307.

Ethyl 3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoate (99h; Table 6.2, Entry 12)

General procedure 3 was followed using compound **98f** (45 mg, 0.18 mmol) and 2-bromobenzonitrile to yield the title compound as a white solid (52 mg, 81%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.19 - 8.18 (m, 1H), 8.03 (ddd, *J* = 7.8, 1.2, 0.5 Hz, 1H), 7.74 (ddd, *J* = 7.8, 1.6, 0.5 Hz, 1H), 7.69 (ddd, *J* = 7.5, 2.0, 1.0 Hz, 1H), 7.64 (ddd, *J* = 7.5, 2.0, 0.5 Hz, 1H), 7.61 - 7.58 (m, 1H), 7.53 - 7.49 (m, 2H), 7.40 - 7.36 (ddd, *J* = 7.6, 7.6, 1.4 Hz, 1H), 4.41 (q, *J* = 7.0 Hz, 2H), 2.38 (s, 3H), 1.42 (t, *J* = 7.0 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.2, 138.9, 137.4, 137.3, 135.1, 134.4, 134.3, 133.2, 133.0, 131.5, 131.0, 130.0, 129.3, 128.72, 128.69, 127.5, 118.9, 109.6, 61.2, 15.1, 14.4.

HRMS (EI): calculated for $C_{21}H_{17}NO_2S$ $[M + H]^+$: 348.1053, found: 348.1049.

Ethyl 3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99i; Table 6.2, Entry 13)

General procedure 3 was followed using compound **98f** (44 mg, 0.18 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a colorless oil (50 mg, 72%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.19 (dd, J = 1.6, 1.6 Hz, 1H), 8.01 (ddd, J = 7.8, 1.4, 1.4 Hz, 1H), 7.77 (d, J = 8.2 Hz, 1H), 7.69 (ddd, J = 7.8, 1.8, 1.2 Hz, 1H), 7.56 - 7.54 (m, 2H), 7.51 - 7.45 (m, 2H), 6.97 (s, 1H), 4.41 (q, J = 7.1 Hz, 2H), 2.36 (s, 3H), 1.41 (t, J = 7.1 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.4, 138.0, 137.7, 134.6, 133.8, 133.5, 133.2, 133.0, 131.9, 131.4, 130.9, 129.7 (q, J_{CF} = 30.0 Hz), 128.6, 128.3, 128.0, 126.5 (q, J_{CF} = 5.0 Hz), 125.1, 124.0 (q, J_{CF} = 274.0 Hz), 61.1, 15.1, 14.4.

HRMS (EI): calculated for $C_{21}H_{17}F_3O_2S$ $[M + H]^+$: 391.0974, found: 391.0963.

Ethyl 3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoate (99j; Table 6.2, Entry 14)

General procedure 3 was followed using compound **98f** (75 mg, 0.3 mmol) and 2-bromobenzaldehyde to yield the title compound as a yellow solid (68 mg, 65%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.26 (s, 1H), 8.15 (dd, J = 1.6, 1.6 Hz, 1H), 7.98 (m, 2H), 7.66 - 7.64 (m, 1H), 7.61 - 7.53 (m, 2H), 7.50 - 7.43 (m, 2H), 6.88 (s, 1H), 4.40 (q, J = 7.0 Hz, 2H), 2.36 (s, 3H), 1.41 (t, J = 7.0 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 192.1, 166.2, 139.0, 137.8, 137.0, 134.6, 134.3, 134.1, 133.7, 133.6, 133.1, 131.06, 131.05, 129.9, 128.7, 128.6, 128.2, 127.9, 61.2, 15.0, 14.4.

HRMS (EI): calculated for $C_{21}H_{18}O_3S$ $[M + H]^+$: 351.1049, found: 351.1049.

3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzonitrile (99k; Table 6.2, Entry 15)

General procedure 3 was followed using compound **98d** (40 mg, 0.2 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a white solid (32 mg, 47%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 - 7.77 (m, 2H), 7.75 - 7.72 (dm, J = 7.8 Hz, 1H), 7.62 - 7.60 (dm, J = 7.8 Hz, 1H), 7.57 (dd, J = 7.2, 7.2 Hz, 1H), 7.55 - 7.52 (m, 2H), 7.49 (dd, J = 7.8, 7.2 Hz, 1H), 6.97 (s, 1H), 2.35 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 138.9, 136.0, 135.7, 134.4, 133.15, 133.12, 133.0, 132.2, 132.1, 131.5, 130.6, 129.5, 128.9 (q, J_{CF} = 30.0 Hz), 128.2, 126.5 (q, J_{CF} = 5.0 Hz), 123.9 (q, J_{CF} = 274.0 Hz), 118.6, 112.9, 15.1.
HRMS (EI): calculated for $C_{19}H_{12}F_3NS$ $[M + H]^+$: 344.0715, found: 344.0710.

2-(5-(3-cyanophenyl)-4-methylthiophen-2-yl)benzonitrile (99l; Table 6.2, Entry 16)

General procedure 3 was followed using compound **98d** (80 mg, 0.4 mmol) and 2-bromobenzonitrile to yield the title compound as a yellow solid (60 mg, 50%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 - 7.73 (m, 3H), 7.65 - 7.61 (m, 3H), 7.55 (dd, J = 7.8, 7.8 Hz, 1H), 7.50 (s, 1H), 7.43 - 7.39 (m, 1H), 2.37 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 138.3, 137.2, 137.0, 135.7, 135.4, 134.5, 133.2, 133.1, 132.2, 131.6, 131.0, 129.6, 129.4, 127.8, 118.8, 118.5, 113.0, 109.8, 15.2.
HRMS (EI): calculated for $C_{19}H_{12}N_2S$ $[M + H]^+$: 301.0794, found: 301.0778.

3-(3-methyl-5-phenylthiophen-2-yl)benzonitrile (98m; Table 6.2, Entry 17)
General procedure 3 was followed using compound **98d** (40 mg, 0.2 mmol) and bromobenzene to yield the title compound as a yellow solid (38 mg, 69%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.78 - 7.77 (m, 1H), 7.73 (ddd, J = 7.8, 1.9, 1.2 Hz, 1H), 7.61 - 7.59 (m, 3H), 7.53 (ddd, J = 8.0, 7.8, 0.6 Hz, 1H), 7.41 - 7.38 (m, 2H), 7.32 - 7.29 (m, 1H), 7.17 (s, 1H), 3.35 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 143.3, 136.0, 135.5, 134.4, 133.8, 133.0, 132.9, 132.0, 130.5, 130.3, 129.5, 129.0, 127.8, 127.4, 125.6, 118.6, 112.9, 15.2.
HRMS (EI): calculated for $C_{18}H_{13}NS$ $[M + H]^+$: 276.0841, found: 276.0835.

***N*-(5-(2-formylphenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99b; Table 6.2, Entry 19)**

General procedure 3 was followed using compound **98b** (32 mg, 0.11 mmol) and 2-bromobenzaldehyde to yield the title compound as a light yellow oil (37 mg, 87%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.30 (s, 1H), 8.04 - 8.02 (m, 1H), 7.65 (ddd, 7.5, 7.3, 1.5 Hz, 1H), 7.59 - 7.57 (m, 1H), 7.55 - 7.51 (m, 1H), 7.38 - 7.35 (dd, *J* = 8.0, 7.5 Hz, 1H), 7.29 - 7.27 (m, 1H), 7.23 (m, 1H), 7.08 (s, 1H), 6.95 (ddd, *J* = 8.0, 2.5, 1.0 Hz, 1H), 3.87 (s, 3H), 3.23 (s, 3H), 2.87 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 191.5, 159.9, 142.1, 137.0, 136.5, 134.3, 134.1, 133.8, 132.8, 131.0, 130.0, 128.8, 128.6, 128.3, 120.7, 114.8, 113.7, 55.4, 38.6, 38.3.

HRMS (EI): calculated for $C_{20}H_{19}NO_4S_2$ $[M + H]^+$: 402.0828, found: 402.0833.

***N*-(5-(2-cyanophenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99n; Table 6.3, Entry 20)**

General procedure 3 was followed using compound **98b** (26 mg, 0.08 mmol) and 2-bromobenzonitrile to yield the title compound as a colorless oil (26 mg, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 - 7.75 (m, 1H), 7.68 (s, 1H), 7.65 - 7.63 (m, 2H), 7.45 - 7.41 (m, 1H), 7.37 - 7.33 (m, 2H), 7.28 - 7.25 (m, 1H), 6.94 (ddd, *J* = 8.0, 2.5, 0.7 Hz, 1H), 3.86 (s, 3H), 3.24 (s, 3H), 3.96 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 159.8, 142.6, 136.8, 136.5, 134.6, 134.4, 133.3, 132.8, 129.9, 128.9, 128.1, 126.3, 120.8, 118.8, 114.9, 113.6, 109.6, 55.4, 38.7, 37.5.

HRMS (EI): calculated for $C_{20}H_{18}N_2O_3S_2$ $[M + H]^+$: 399.0832, found: 399.0831.

***N*-(2-(3-methoxyphenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-N-methylmethanesulfonamide (99o; Table 6.3, Entry 21)**

General procedure 3 was followed using compound **98b** (32 mg, 0.11 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a colorless oil (37 mg, 78%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.78 (d, J = 8.0 Hz, 1H), 7.62 - 7.55 (m, 2H), 7.52 - 7.49 (m, 1H), 7.34 (d, J = 7.5 Hz, 1H), 7.32 - 7.31 (m, 1H), 7.26 - 7.24 (m, 1H), 7.14 (s, 1H), 6.92 (ddd, J = 8.2, 2.5, 1.0 Hz, 1H), 3.86 (s, 3H), 3.23 (s, 3H), 2.87 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 159.8, 141.3, 137.5, 133.6, 133.1, 132.8, 132.7, 131.7, 129.8, 128.9 (q, J_{CF} = 30.0 Hz), 128.5, 126.8, 126.7 (q, J_{CF} = 5.0 Hz), 123.9 (q, J_{CF} = 274.0 Hz), 120.7, 114.6, 113.6, 55.4, 38.6, 37.5.
HRMS (EI): calculated for $C_{20}H_{18}F_3NO_3S_2$ $[M + H]^+$: 442.0753, found: 442.0752.

Ethyl 2-(5-(3-cyanophenyl)-4-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99p; Table 6.2, Entry 22)

General procedure 3 was followed using compound **98a** (17.5 mg, 0.06 mmol) and ethyl 2-bromobenzoate to yield the title compound as a colorless oil (20 mg, 77%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 (ddd, J = 8.0, 2.0, 1.5 Hz, 1H), 7.92 - 7.89 (m, 1H), 7.82 (ddd, J = 8.0, 1.5, 0.5 Hz, 1H), 7.64 (ddd, J = 8.0, 1.5, 1.5 Hz, 1H), 7.57 - 7.74 (m, 2H), 7.51 - 7.45 (m, 2H), 7.08 (s, 1H), 4.25 (q, J = 7.1 Hz, 2H), 3.23 (s, 3H), 2.91 (s, 3H), 1.24 (t, J = 7.1 Hz, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 168.0, 141.5, 138.3, 134.6, 133.7, 132.9, 132.5, 131.9, 131.65, 131.61, 131.3, 131.0, 129.9, 129.8, 128.7, 124.9, 118.4, 113.1, 61.5, 38.8, 36.7, 14.0.
HRMS (EI): calculated for $C_{22}H_{20}N_2O_4S_2$ $[M + H]^+$: 441.0937, found: 441.0937.

***N*-(2-(3-cyanophenyl)-5-(2-formylphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99q; Table 6.3, Entry 23)**

General procedure 3 was followed using compound **98a** (46 mg, 0.16 mmol) and 2-bromobenzaldehyde to yield the title compound as a yellow solid (40 mg, 64%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 10.22 (s, 1H), 8.02 - 7.98 (m, 1H), 7.95 (ddd, J = 8.0, 2.5, 1.0 Hz, 1H), 7.89 - 7.88 (m, 1H), 7.67 - 7.63 (m, 2H), 7.57 - 7.52 (m, 3H), 7.07 (s, 1H), 3.24 (s, 3H), 2.90 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 191.1, 139.8, 138.2, 136.2, 135.2, 134.1, 133.9, 133.2, 132.6, 132.0, 131.7, 131.1, 129.9, 129.2, 128.8, 127.7, 118.2, 113.3, 38.8, 37.1.
HRMS (EI): calculated for $C_{20}H_{16}N_2O_3S_2$ $[M + H]^+$: 397.0675, found: 397.0665.

***N*-(2-(3-cyanophenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99r; Table 6.3, Entry 24)**

General procedure 3 was followed using compound **98a** (30 mg, 0.1 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a yellow oil (42 mg, 95%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.01 (ddd, J = 8.0, 2.0, 1.5 Hz, 1H), 7.93 - 9.92 (m, 1H), 7.81 (d, J = 7.5 Hz, 1H), 7.66 - 7.64 (m, 1H), 7.63 - 7.61 (m, 1H), 7.58 - 7.54 (m, 3H), 7.15 (s, 1H), 3.27 (s, 3H), 2.90 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 139.1, 139.0, 134.5, 133.5, 132.7, 132.6, 132.1, 131.9, 131.8, 131.7, 129.8, 129.0, 128.9 (q, J_{CF} = 30.0 Hz), 126.8 (q, J_{CF} = 5.0 Hz), 125.9, 121.7 (q, J_{CF} = 274.0 Hz), 118.4, 113.1, 38.8, 36.3.
HRMS (EI): calculated for $C_{20}H_{15}F_3N_2O_2S_2$ $[M + H]^+$: 437.0600, found: 437.0606.

***N*-(5-(2-cyanophenyl)-2-(3-cyanophenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99a; Table 6.3, Entry 25)**

General procedure 3 was followed using compound **98a** (22 mg, 0.08 mmol) and 2-bromobenzonitrile to yield the title compound as a white solid (21 mg, 72%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.05 (ddd, J = 7.9, 1.8, 1.2 Hz, 1H), 7.94 - 7.93 (m, 1H), 7.80 (ddd, J = 7.9, 1.2, 0.6 Hz, 1H), 7.70 (s, 1H), 7.69 - 7.68 (m, 1H), 7.67 - 7.64 (m, 2H), 7.58 (ddd, J = 7.9, 7.9, 0.6 Hz, 1H), 7.48 (ddd, J = 7.8, 7.0, 1.7 Hz, 1H), 3.28 (s, 3H), 2.99 (s, 3H).
^{13}C NMR (125 MHz, $CDCl_3$) δ 140.2, 138.2, 136.0, 135.6, 134.5, 133.5, 133.2, 132.7, 132.0, 131.8, 129.9, 129.0, 128.6, 125.6, 118.7, 118.3, 113.2, 109.9, 38.9, 36.4.
HRMS (EI): calculated for $C_{20}H_{15}N_3O_2S_2$ $[M + H]^+$: 394.0678, found: 394.0676.

Ethyl 3-(5-(2-formylphenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99s; Table 6.3, Entry 26)

General procedure 3 was followed using compound **98c** (34 mg, 0.1 mmol) and 2-bromobenzaldehyde to yield the title compound as a yellow solid (29 mg, 65%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 10.31 (s, 1H), 8.33 - 8.32 (m, 1H), 8.09 - 8.03 (dm, J = 8 Hz, 1H), 8.05 - 8.03 (dm, J = 8 Hz, 1H), 7.91 - 7.89 (m, 1H), 7.69 - 7.65 (m, 1H), 7.60 - 7.58 (m, 1H), 7.57 - 7.53 (m, 2H), 7.12 (s, 1H), 4.42 (q, J = 7.5, Hz, 2H), 3.25 (s, 3H), 2.89 (s, 3H), 1.42 (t, J = 7.5 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 191.3, 166.0, 141.0, 137.1, 136.8, 134.8, 134.1, 133.9, 132.5, 132.0, 131.3, 131.0, 129.7, 129.4, 129.2, 128.9, 128.6, 128.4, 61.3, 38.7, 38.0, 14.3.

HRMS (EI): calculated for $C_{22}H_{21}NO_5S_2$ $[M + H]^+$: 444.0934, found: 444.0936.

Ethyl 3-(3-(*N*-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99t; Table 6.2, Entry 27)

General procedure 3 was followed using compound **98c** (30 mg, 0.09 mmol) and 1-bromo-2-(trifluoromethyl)benzene to yield the title compound as a colorless oil (32 mg, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.32 (dd, J = 1.8, 1.6 Hz, 1H), 8.06 - 8.04 (m, 1H), 7.93 (ddd, J = 7.8, 1.9, 1.2 Hz, 1H), 7.79 (d, J = 7.5 Hz, 1H), 7.63 - 7.56 (m, 2H), 7.55 - 7.50 (m, 2H), 7.17 (s, 1H), 4.41 (q, J = 7.13 Hz, 2H), 3.25 (s, 3H), 2.28 (s, 3H), 1.42 (t, J = 7.13 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.1, 140.3, 138.0, 134.0, 132.8, 132.53, 132.50, 132.3, 131.7, 131.1, 129.5, 129.4, 129.0, 128.9 (q, J_{CF} = 30.0 Hz), 128.7, 126.72 (q, J_{CF} = 5.0 Hz), 126.67, 123.9 (q, J_{CF} = 274.0 Hz), 61.2, 38.7, 37.3, 14.3.

HRMS (EI): calculated for $C_{22}H_{20}F_3NO_4S_2$ $[M + H]^+$: 484.0859, found: 484.0854.

Ethyl 3-(5-(2-cyanophenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99u; Table 6.3, Entry 28)

General procedure 3 was followed using compound **98c** (40 mg, 0.12 mmol) and 2-bromobenzonitrile to yield the title compound as a yellow oil (49 mg, 93%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.34 - 8.33 (m, 1H), 8.08 - 8.059 (m, 1H), 7.98 - 7.96 (m, 1H), 7.79 - 7.77 (m, 1H), 7.71 (s, 1H), 7.67 - 7.65 (m, 2H), 7.56 - 7.52 (m, 1H), 7.47 - 7.44 (m, 1H), 4.41 (q, *J* = 7.0 Hz, 2H), 3.27 (s, 3H), 2.96 (s, 3H), 1.42 (t, *J* = 7.0 Hz, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 166.0, 141.6, 137.3, 136.4, 135.1, 134.4, 133.3, 132.5, 132.0, 131.2, 129.8, 129.5, 129.1, 129.0, 128.3, 126.2, 118.7, 109.8, 61.3, 38.8, 37.2, 14.3.

HRMS (EI): calculated for $C_{22}H_{20}N_2O_4S_2$ $[M + H]^+$: 441.0937, found: 441.0930.

2-(5-(3-carboxyphenyl)-4-methylthiophen-2-yl)benzoic acid (103a)

General procedure 1 was followed using compound **99g** (41 mg, 0.1 mmol) to yield the title compound as a white solid (35 mg, 94%).

^{1}H NMR (500 MHz, d_6-DMSO) δ 8.02 (dd, *J* = 1.7, 1.6 Hz, 1H), 7.92 (ddd, *J* = 8.0, 1.4, 1.4 Hz, 1H), 7.77 (ddd, *J* = 7.7, 1.1, 0.7 Hz, 1H), 7.64 - 7.59 (m, 2H), 7.56 - 7.53 (m, 2H), 7.47- 7.44 (m, 1H), 7.06 (s, 1H), 2.32 (s, 3H).

^{13}C NMR (125 MHz, d_6-DMSO) δ 170.3, 167.4, 140.2, 136.8, 134.6, 134.5, 133.4, 132.9, 132.3, 131.9, 131.2, 131.1, 130.6, 129.8, 129.2, 129.1, 128.6, 128.5, 15.4.

HRMS (EI): calculated for $C_{19}H_{14}O_4S$ $[M + H]^+$: 339.0686, found: 339.0686.

3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoic acid (103b; Table 6.4, Entry 31)

General procedure 1 was followed using compound **99h** (17 mg, 0.05 mmol) to yield the title compound as a white solid in quantitative yield.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.27 (dd, *J* = 2.0, 1.7 Hz, 1H), 8.10 (ddd, *J* = 7.8, 1.4, 1.2 Hz, 1H), 7.77 - 7.74 (m, 2H), 7.67 - 7.64 (m, 1H), 7.61 (ddd, *J* = 7.4, 7.4, 1.4 Hz, 1H),

7.56 (dd, J = 7.7, 7.7 Hz, 1H), 7.52 (s, 1H), 7.39 (ddd, J = 7.5, 7.5, 1.4 Hz, 1H), 2.40, (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 170.4, 138.6, 137.6, 137.3, 135.3, 134.6, 134.5 134.1, 133.0, 131.5, 130.5, 129.7, 129.4, 129.3, 128.9, 127.6, 118.9, 109.6, 15.17.

HRMS (EI): calculated for $C_{19}H_{13}NO_2S$ [M - H]$^-$: 318.0594, found: 318.0600.

3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103c; Table 6.4, Entry 32)

General procedure 1 was followed using compound **99i** (33 mg, 0.08 mmol) to yield the title compound as a white solid in quantitative yield.

^{1}H NMR (500 MHz, d_6-DMSO) δ 8.03 (dd, J = 1.7, 1.6 Hz, 1H), 7.95 - 7.93 (ddd, J = 8.0, 1.4, 1.4 Hz, 1H), 7.87 (d, J = 8.0 Hz, 1H), 7.79 (ddd, J = 7.7, 1.5, 1.0 Hz, 1H), 7.74 (dd, J = 7.6, 7.5 Hz, 1H), 7.66 - 7.60 (m, 3H), 7.07 (s, 1H), 2.34 (s, 3H).

^{13}C NMR (125 MHz, d_6-DMSO) δ 167.4, 137.6, 137.5, 134.3, 134.2, 133.6, 133.04, 133.02, 132.9, 132.8, 131.9, 129.9, 129.45, 129.42 (q, J_{CF} = 30.0 Hz), 129.3, 128.8, 127.0 (q, J_{CF} = 5.0 Hz), 124.4 (q, J_{CF} = 274.0 Hz), 15.3.

HRMS (EI): calculated for $C_{19}H_{13}F_3O_2S$ [M - H]$^-$: 361.0516, found: 361.0521.

3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoic acid (103d; Table 6.4, Entry 33)

General procedure 1 was followed using compound **99j** (38 mg, 0.11 mmol) to yield the title compound as a white solid in quantitative yield.

^{1}H NMR (500 MHz, d_6-DMSO) δ 10.17 (s, 1H), 8.01 (s, 1H), 7.89 (d, J = 7.7 Hz, 1H), 7.85 (d, J = 7.6 Hz, 1H), 7.75 (d, J = 7.7 Hz, 1H), 7.70 (dd, J = 7.2, 7.2 Hz, 1H), 7.61 - 7.51 (m, 3H), 7.14 (s, 1H), 2.3 (s, 3H).

^{13}C NMR (125 MHz, d_6-DMSO) δ 192.0, 167.4, 138.4, 137.2, 136.8, 135.3, 134.7, 134.6, 134.2, 133.90, 133.1, 131.9, 131.4, 129.9, 129.3, 129.0, 128.9, 128.2, 15.3.

HRMS (EI): calculated for $C_{19}H_{14}O_3S$ [M - H]$^-$: 321.0591, found: 321.0597.

3-(3-(N-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103e; Table 6.4, Entry 34)

General procedure 1 was followed using compound **99t** (16 mg, 0.03 mmol) to yield the title compound as a white solid in quantitative yield.

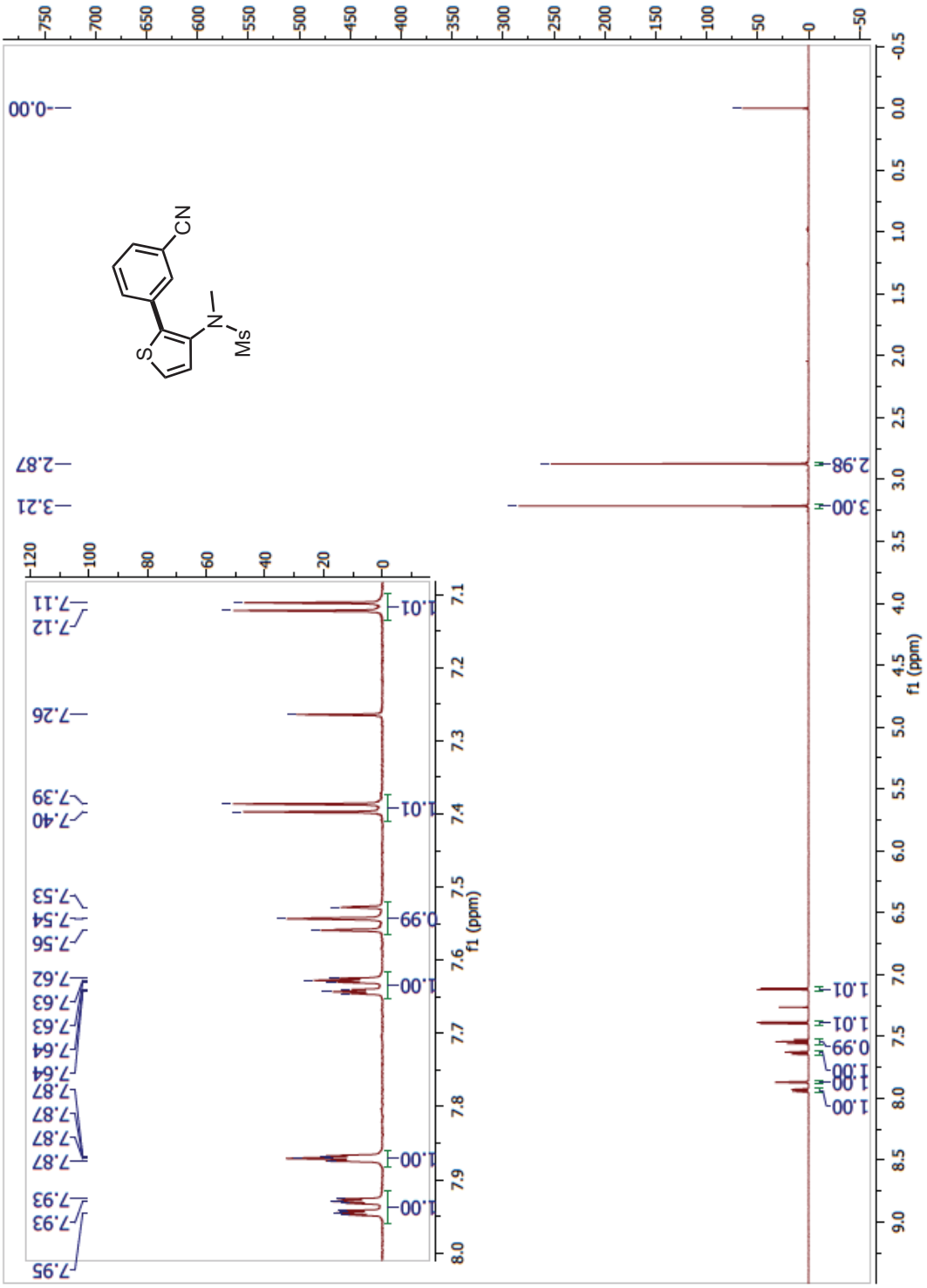
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.39 (s, 1H), 8.12 (d, J = 7.8 Hz, 1H), 8.01 (d, J = 7.9 Hz, 1H), 7.80 (d, J = 7.9 Hz, 1H), 7.63 - 7.56 (m, 3H), 7.53 (dd, J = 7.9, 7.3 Hz, 1H), 7.18 (s, 1H), 3.27 (s, 3H), 2.89 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 170.6, 140.1, 138.2, 134.1, 133.5, 132.8 (2C), 132.6, 132.5, 131.8, 130.1, 130.0, 129.2, 129.0 (q, J_{CF} = 30.0 Hz), 128.7, 126.7, 126.6 (q, J_{CF} = 5.0 Hz), 123.9 (q, J_{CF} = 274.0 Hz), 38.8, 37.2.

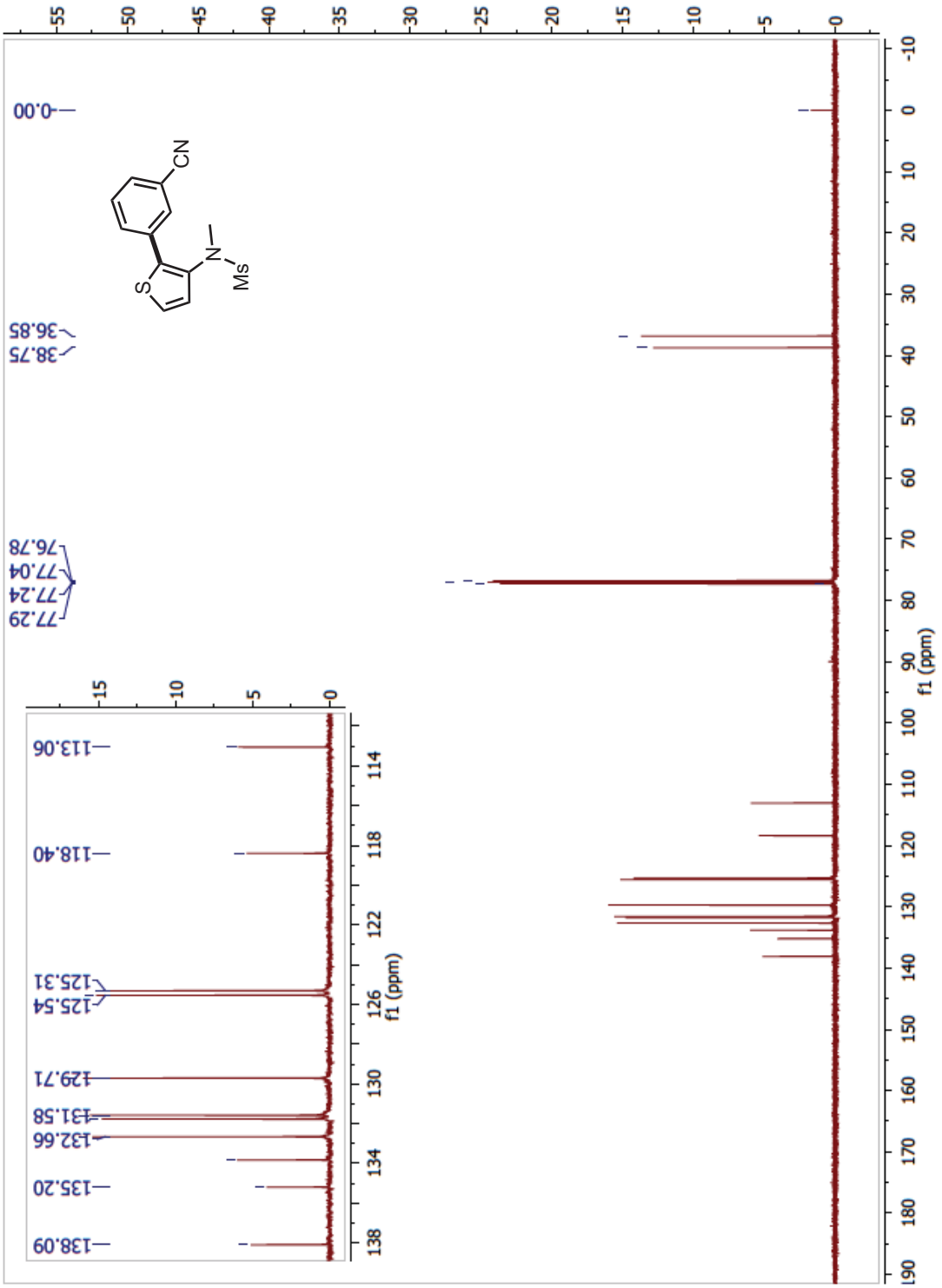
HRMS (EI): calculated for $C_{20}H_{16}F_3NO_4S_2$ $[M - H]^-$: 454.0400, found: 454.0404.

^{1}H and ^{13}C NMR Spectra

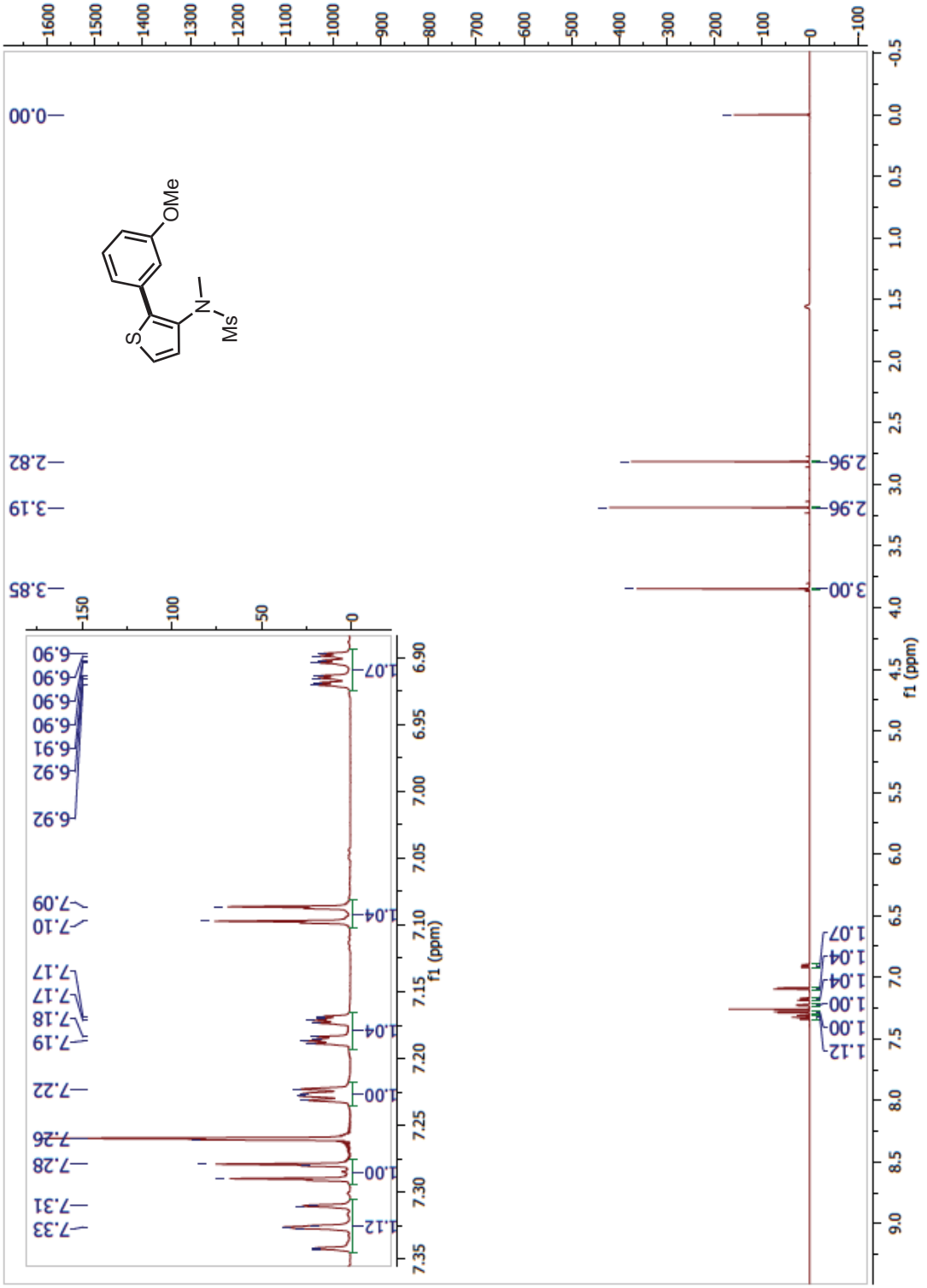
***N*-(2-(3-cyanophenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (98a; Table 6.1, Entry 1) ^{1}H**:



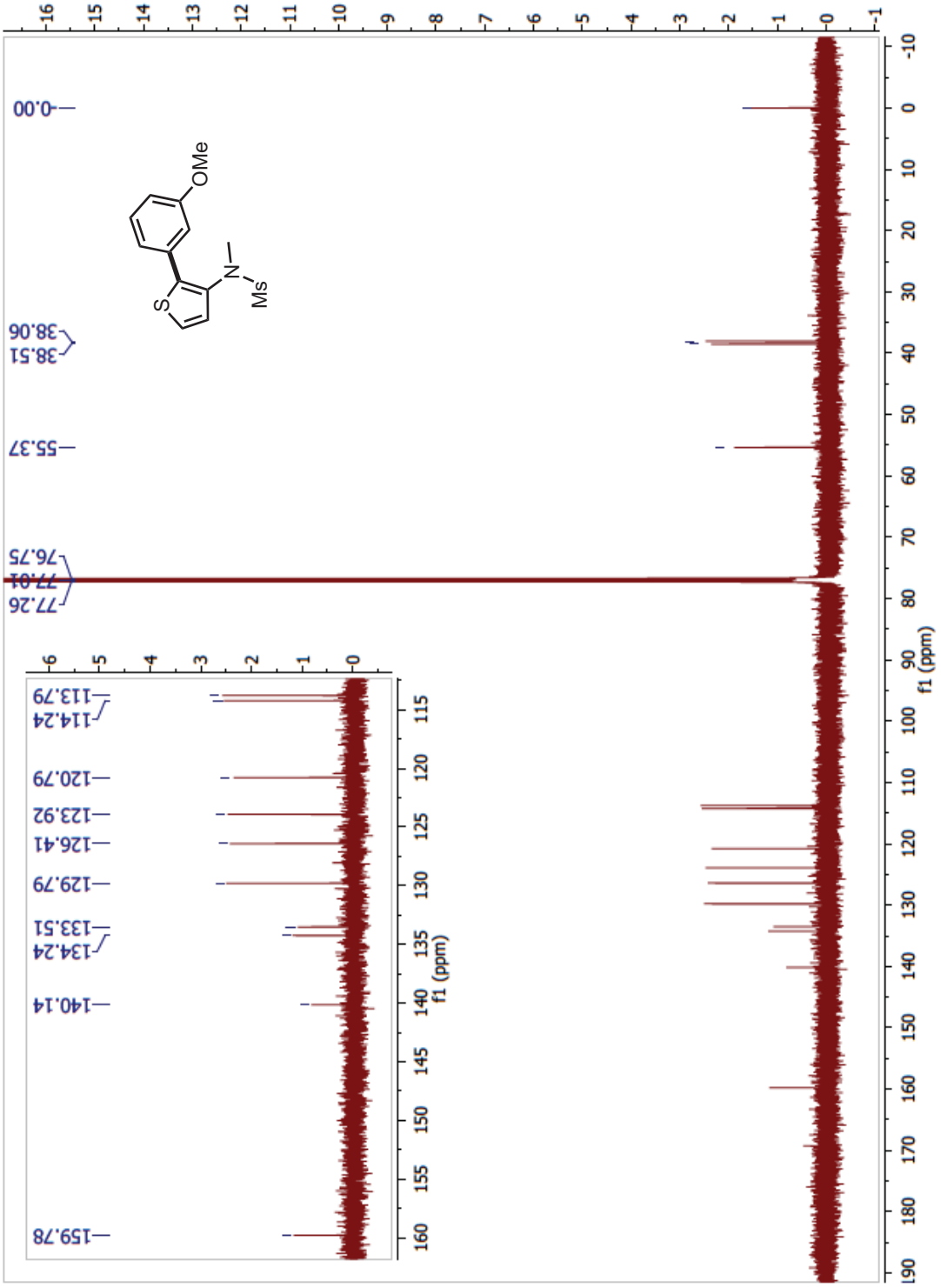
***N*-(2-(3-cyanophenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (98a; Table 6.1, Entry 1) ^{13}C**:



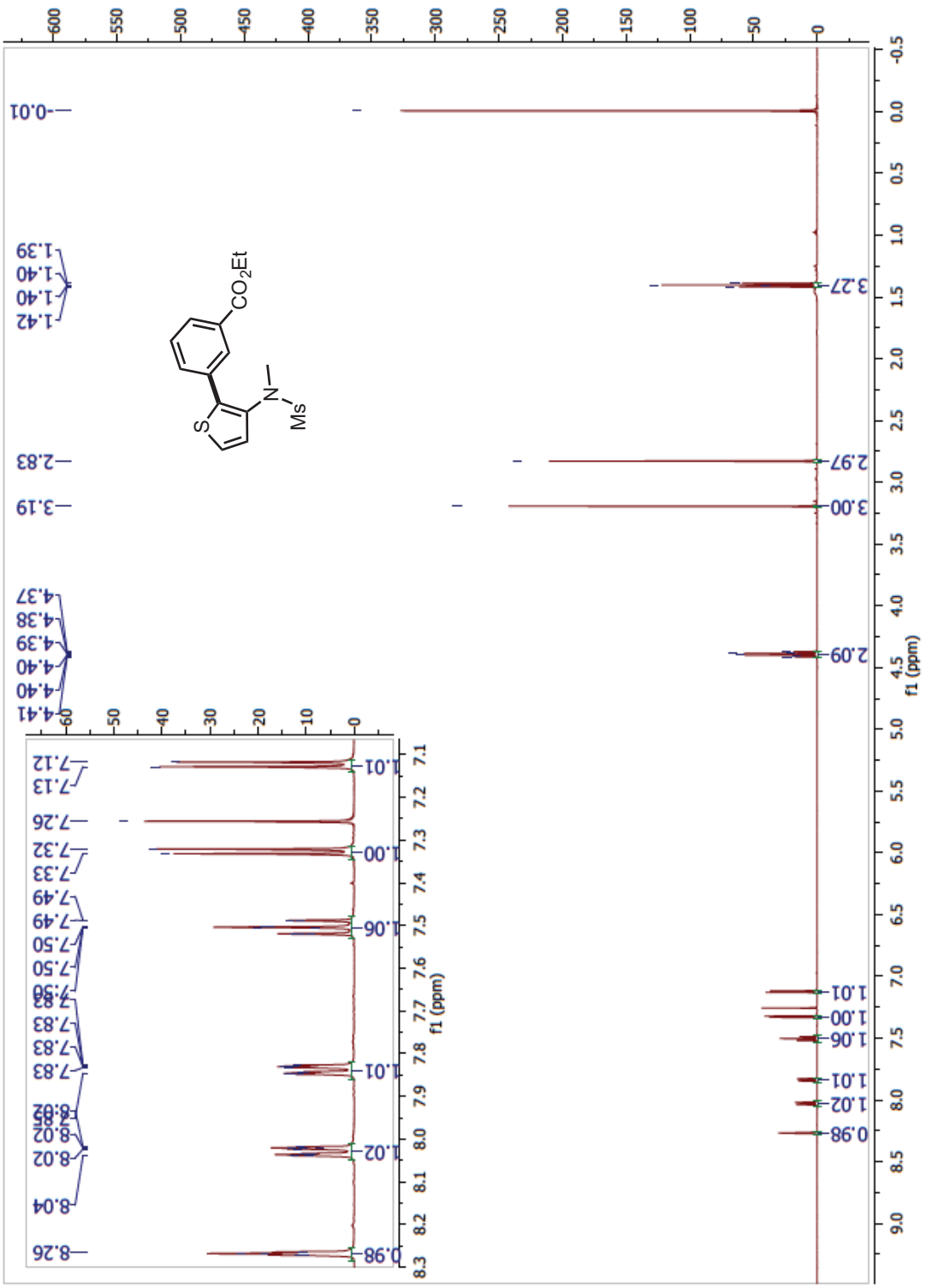
***N*-(2-(3-methoxyphenyl) thiophen-3-yl)-*N*-methylmethanesulfonamide (98b; Table 6.1, Entry 2) ^{1}H**:



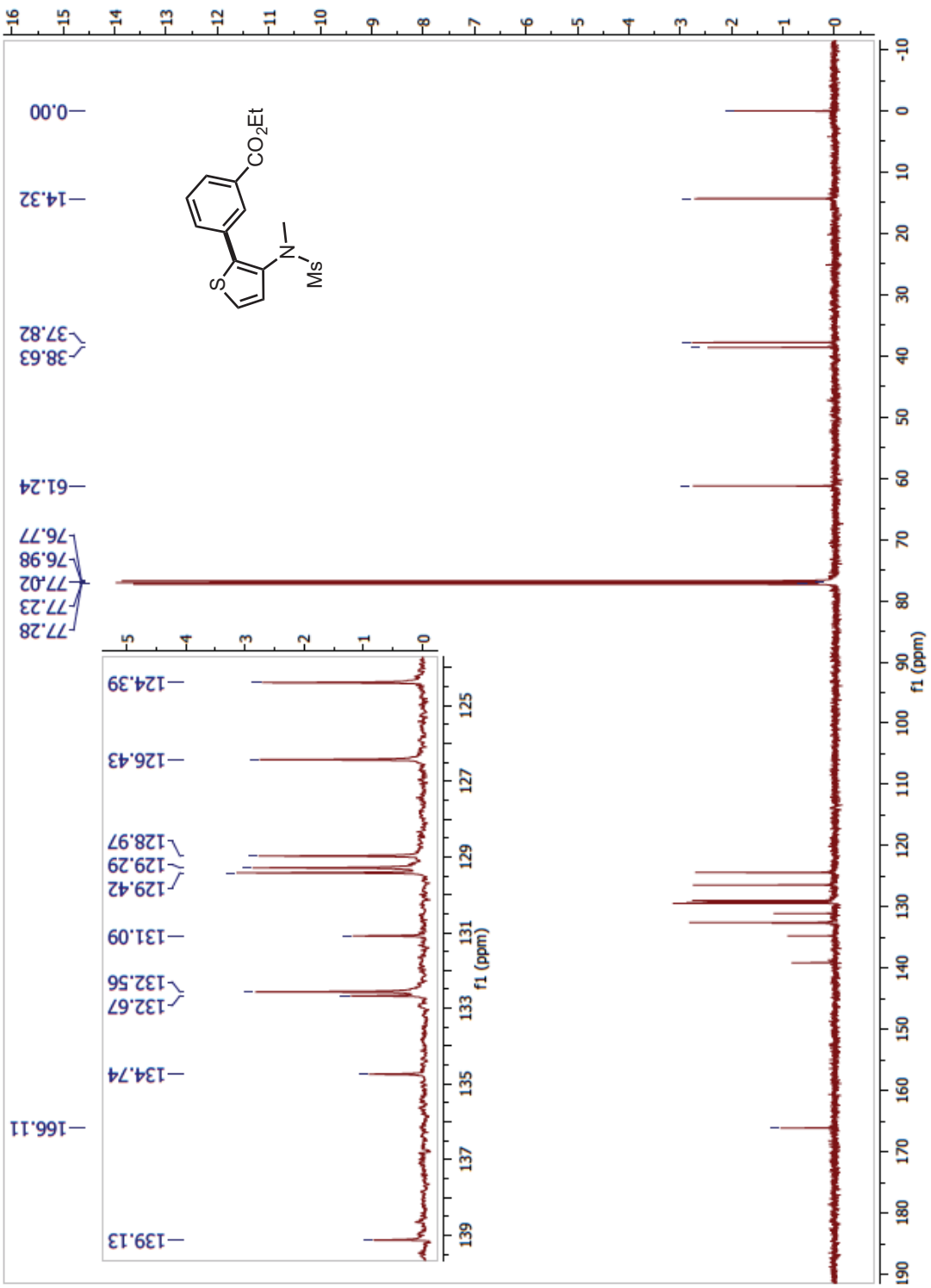
***N*-(2-(3-methoxyphenyl) thiophen-3-yl)-*N*-methylmethanesulfonamide (98b; Table 6.1, Entry 2) ^{13}C:**



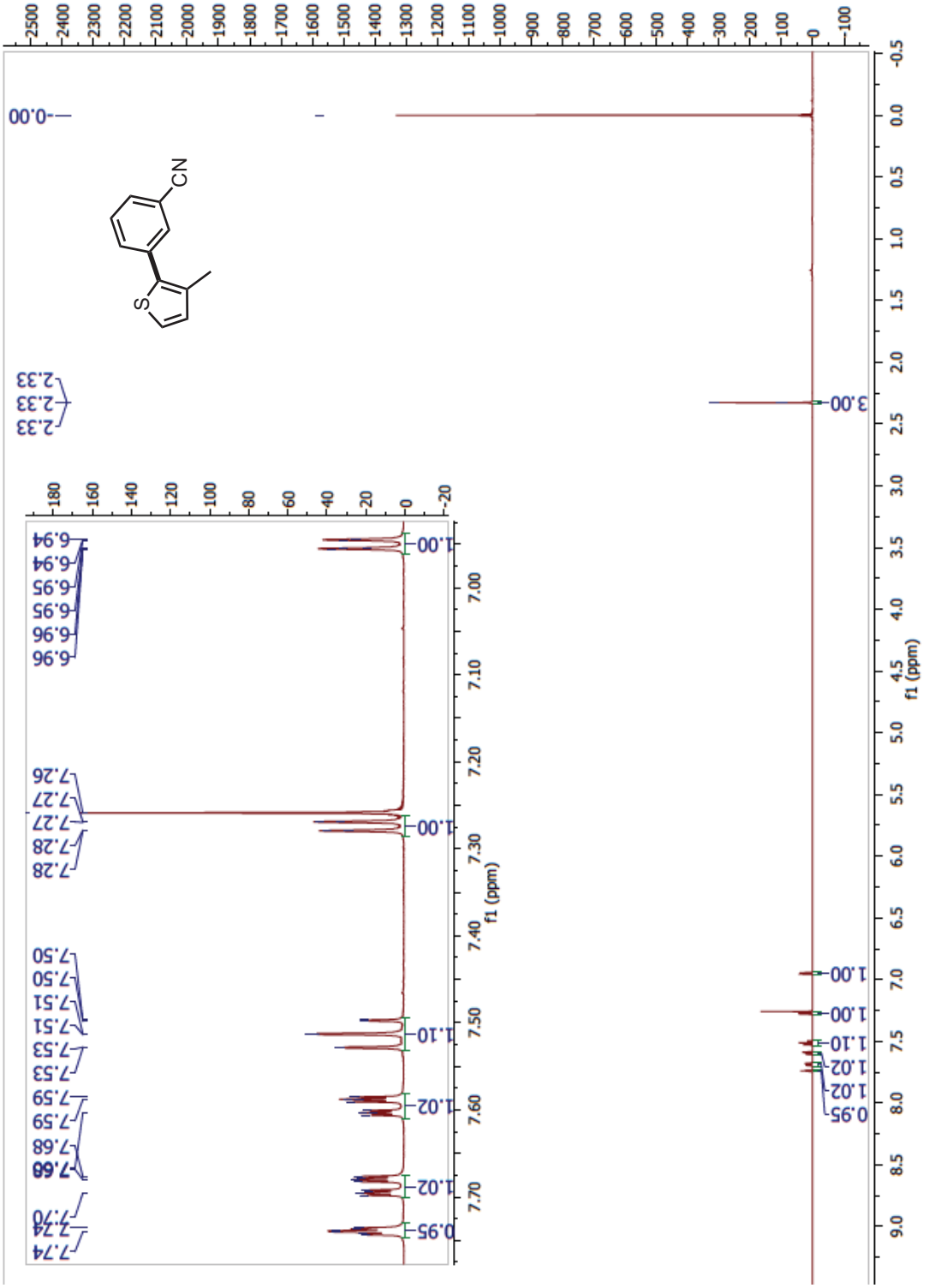
Ethyl 3-(3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (98c; Table 6.1, Entry 3) ^{1}H:



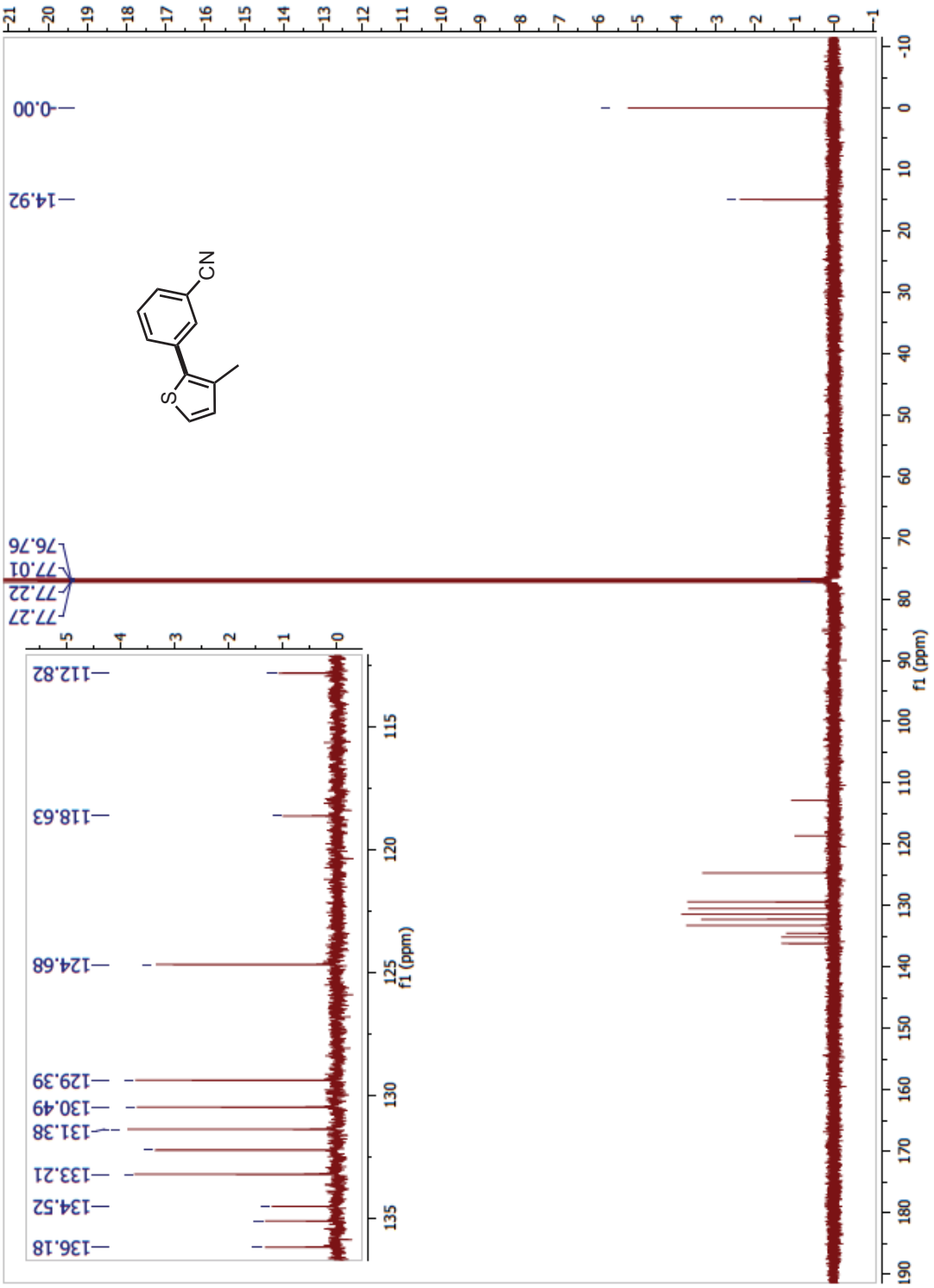
Ethyl 3-(3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (98c; Table 6.1, Entry 3) ^{13}C:



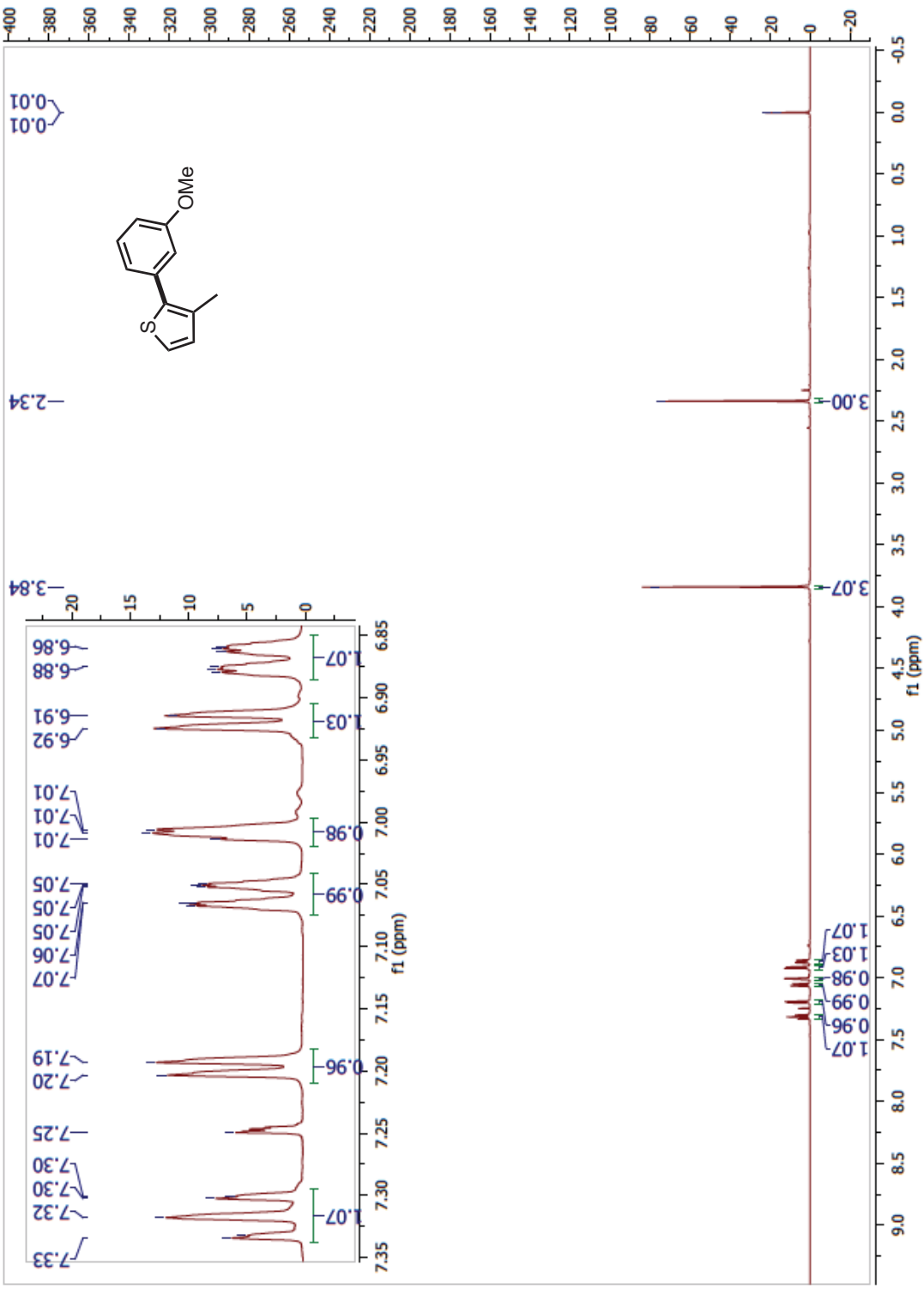
3-(3-methylthiophen-2-yl)benzonitrile (98d; Table 6.1, Entry 4) ^{1}H:



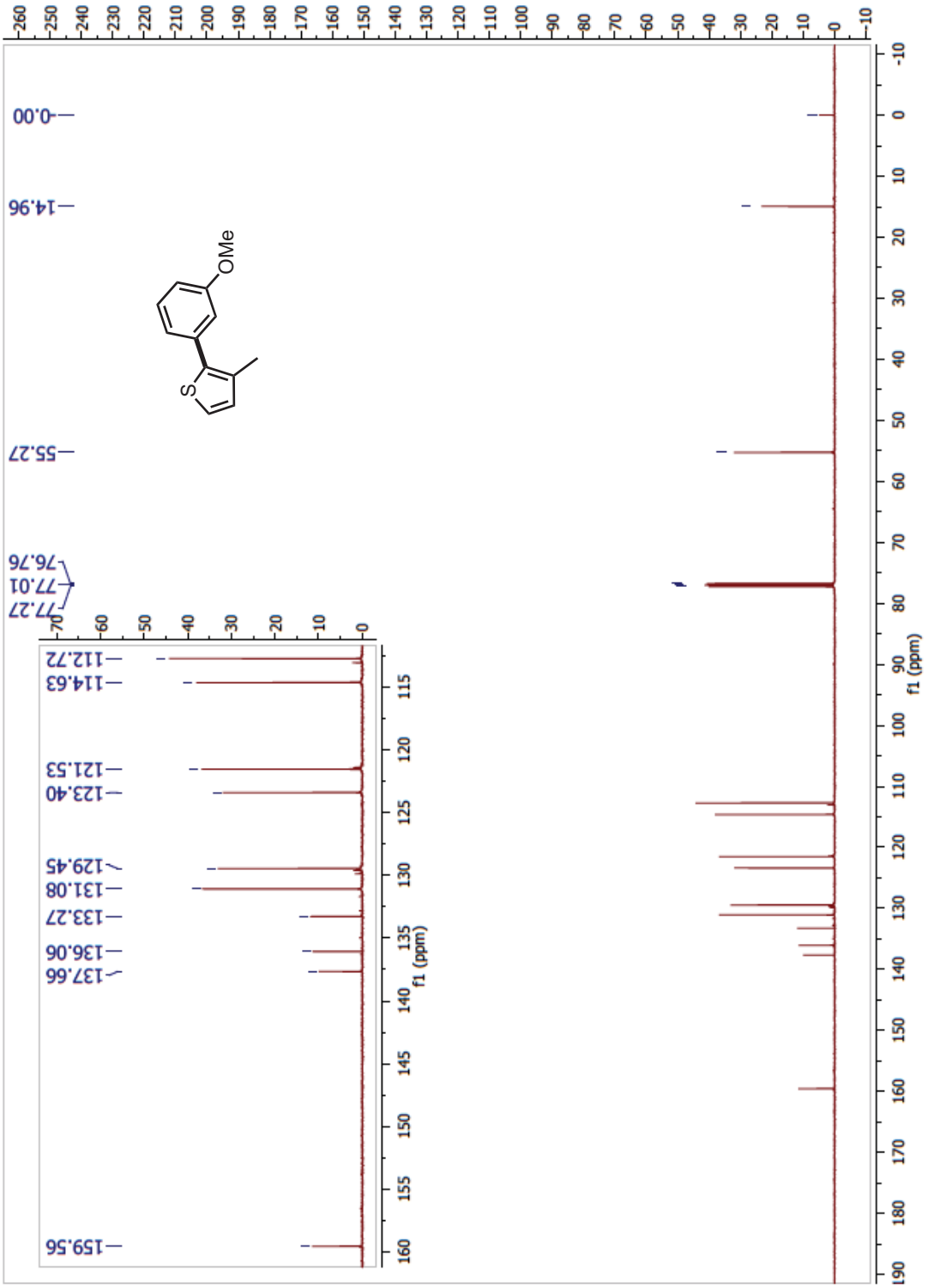
3-(3-methylthiophen-2-yl)benzonitrile (98d; Table 6.1, Entry 4) ^{13}C:



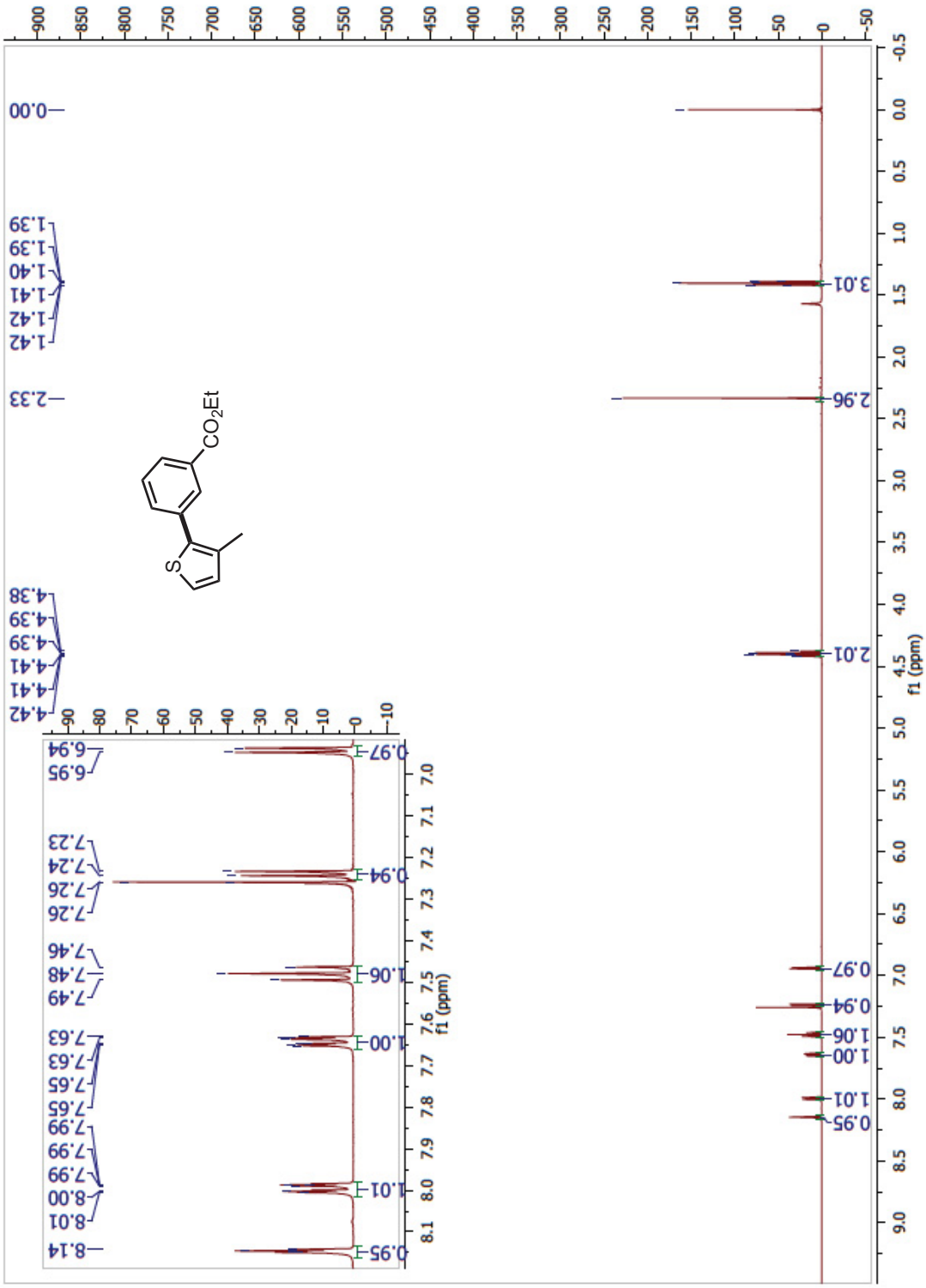
2-(3-methoxyphenyl)-3-methylthiophene (98e; Table 6.1, Entry 5) ^{1}H:



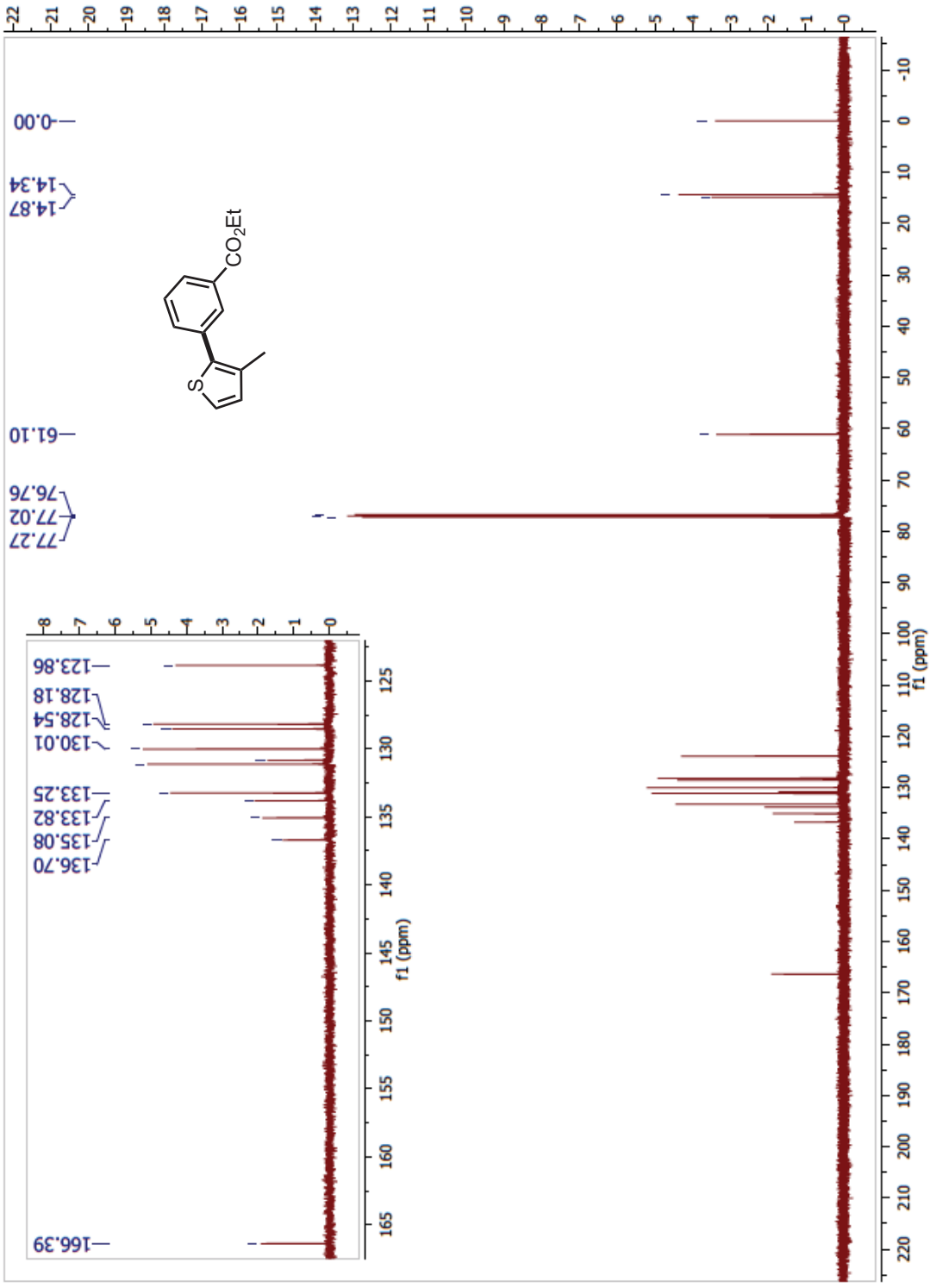
2-(3-methoxyphenyl)-3-methylthiophene (98e; Table 6.1, Entry 5) ^{13}C:



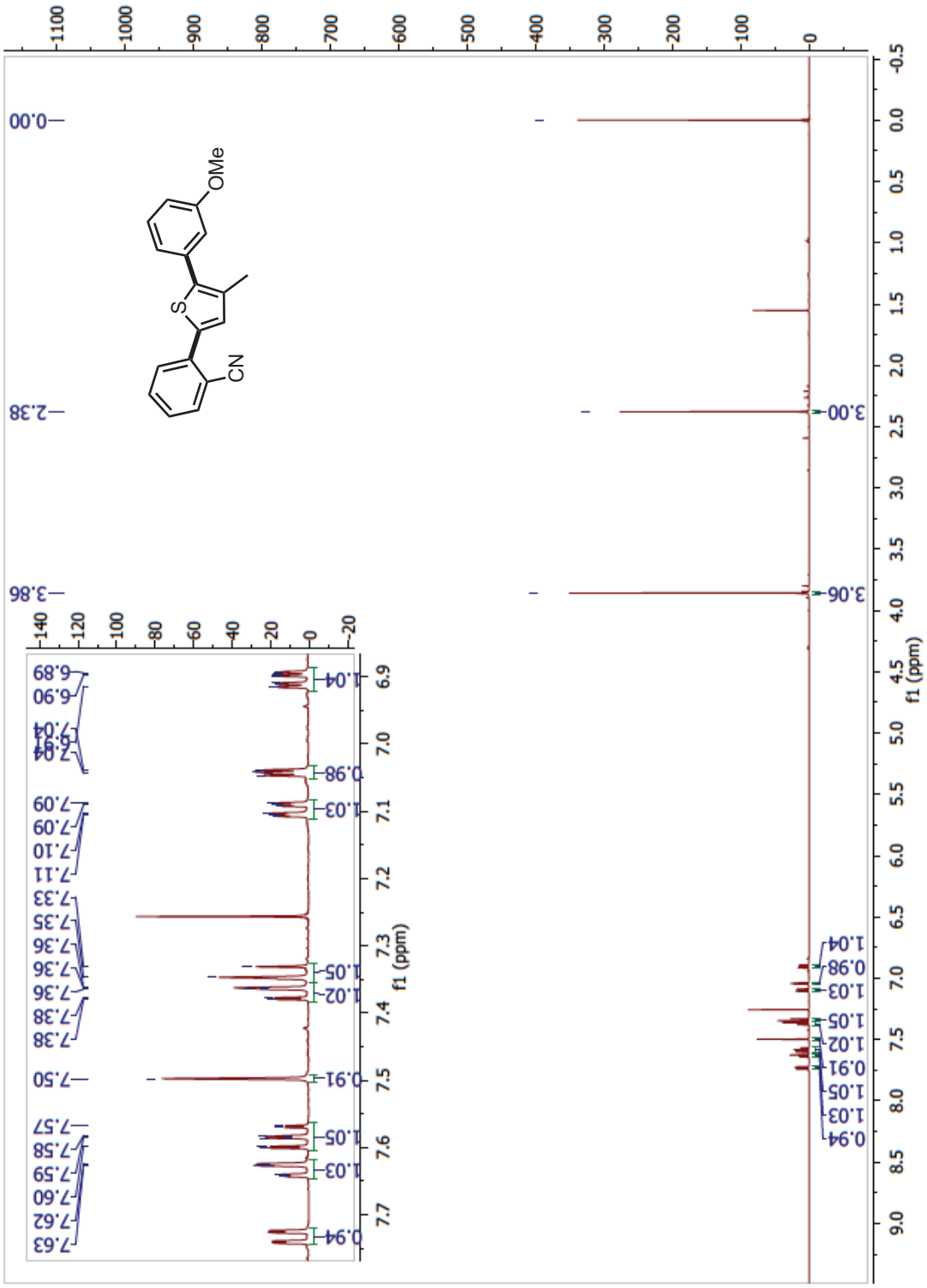
Ethyl 3-(3-methylthiophen-2-yl) benzoate (98f; Table 6.1, Entry 6) ^{1}H:



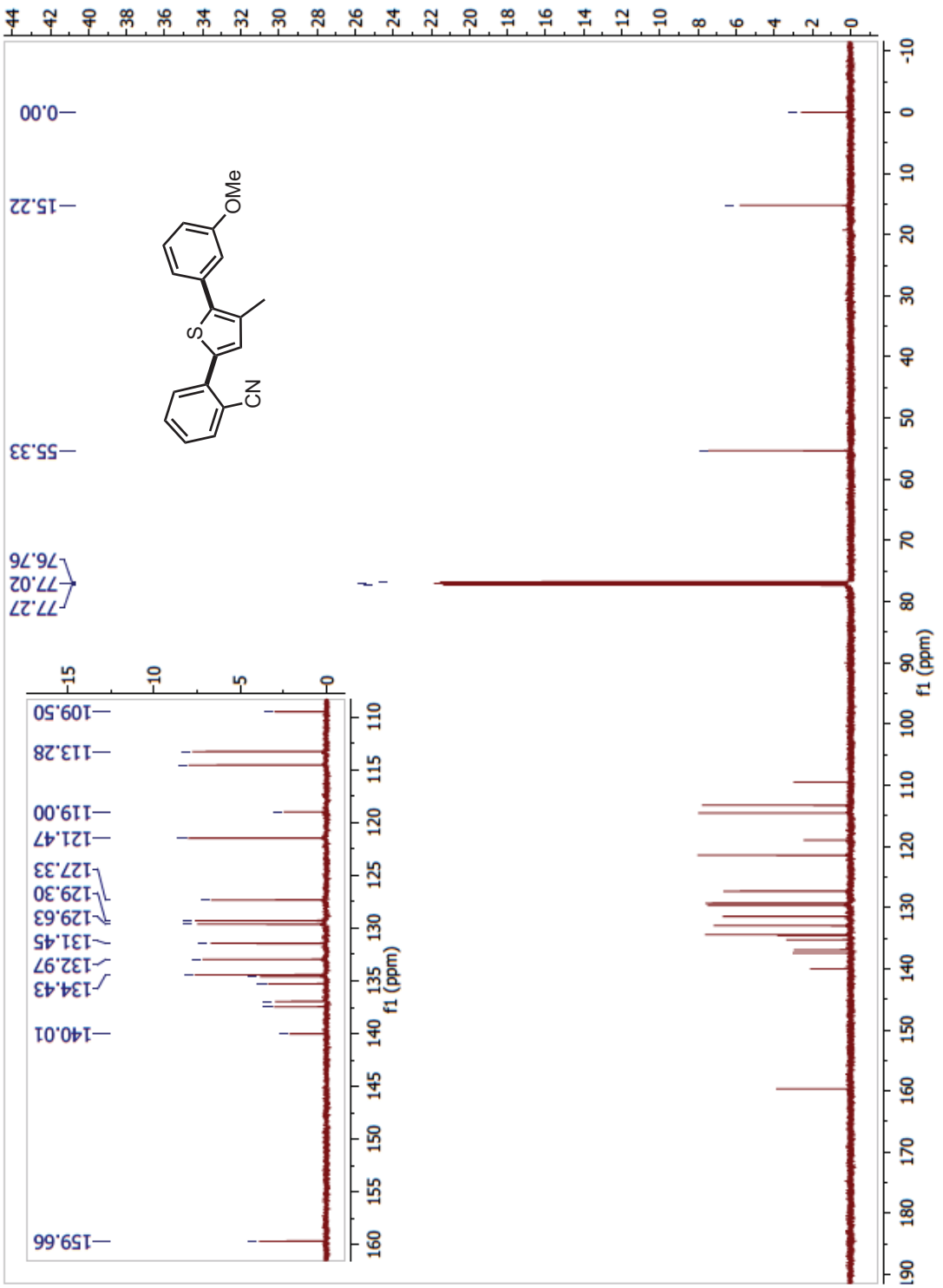
Ethyl 3-(3-methylthiophen-2-yl) benzoate (98f; Table 6.1, Entry 6) ^{13}C:



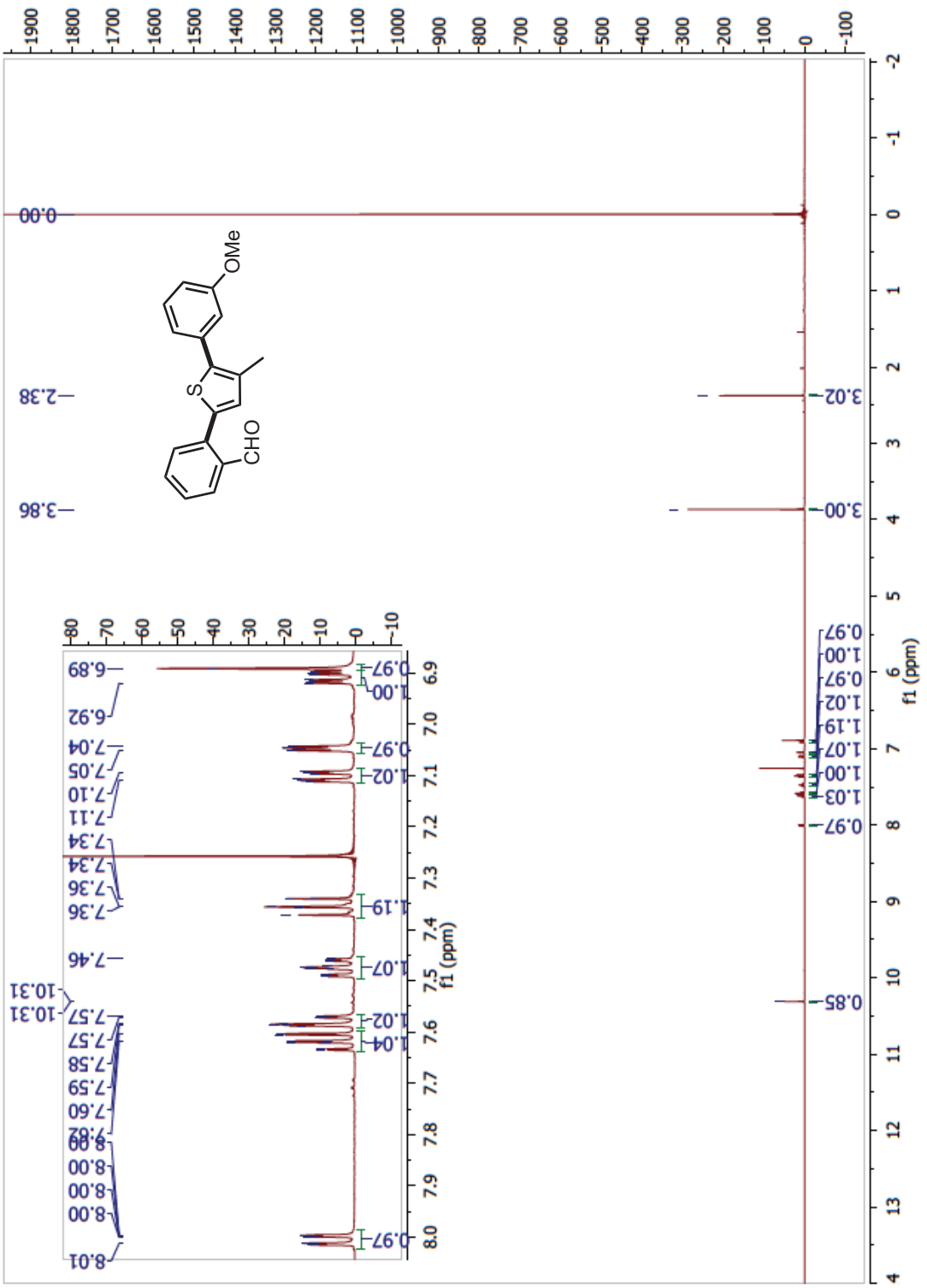
2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzonitrile (99c; Table 6.2, Entry 7)
^{1}H:



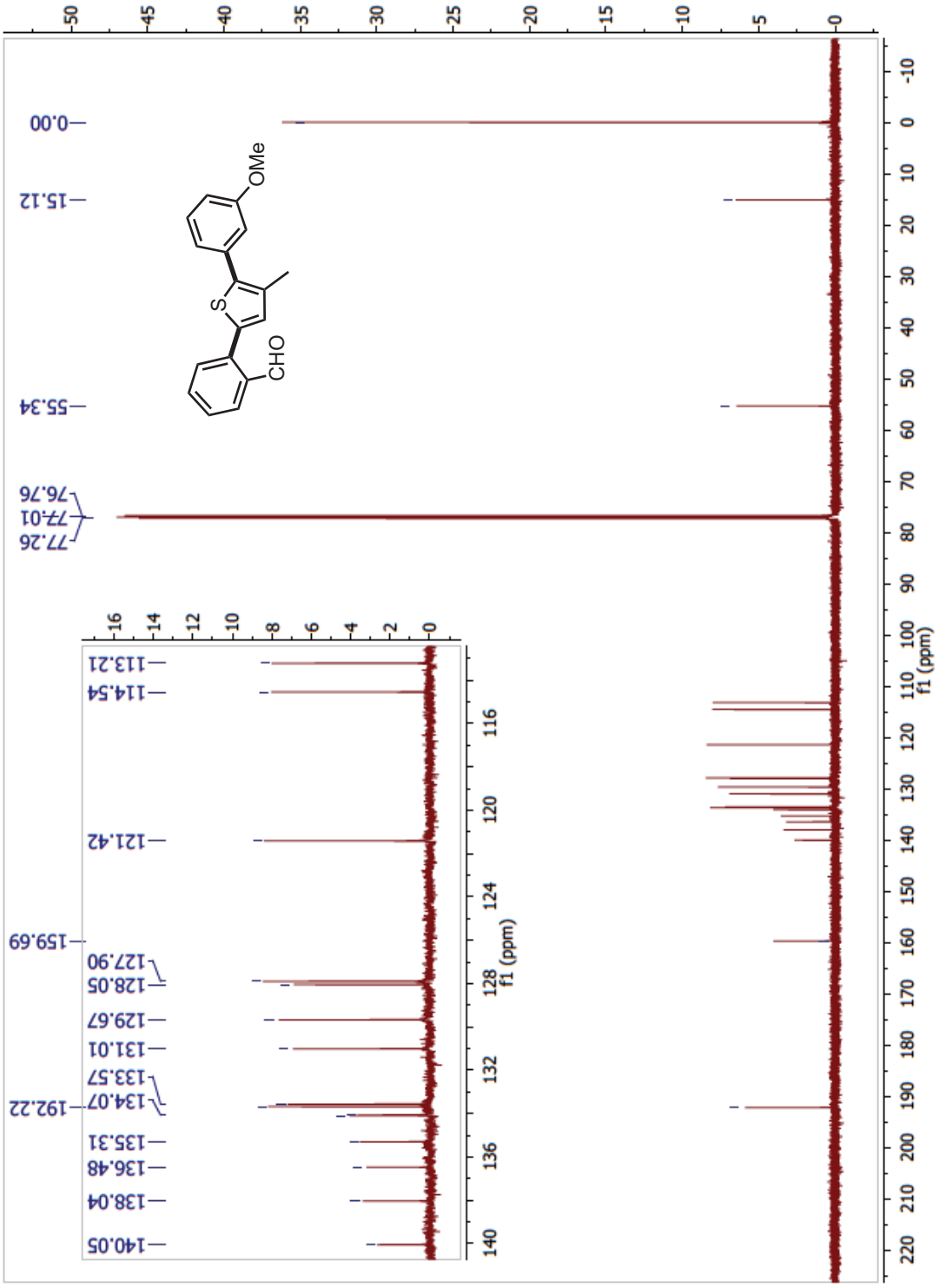
2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzonitrile (99c; Table 6.2, Entry 7)
^{13}C:



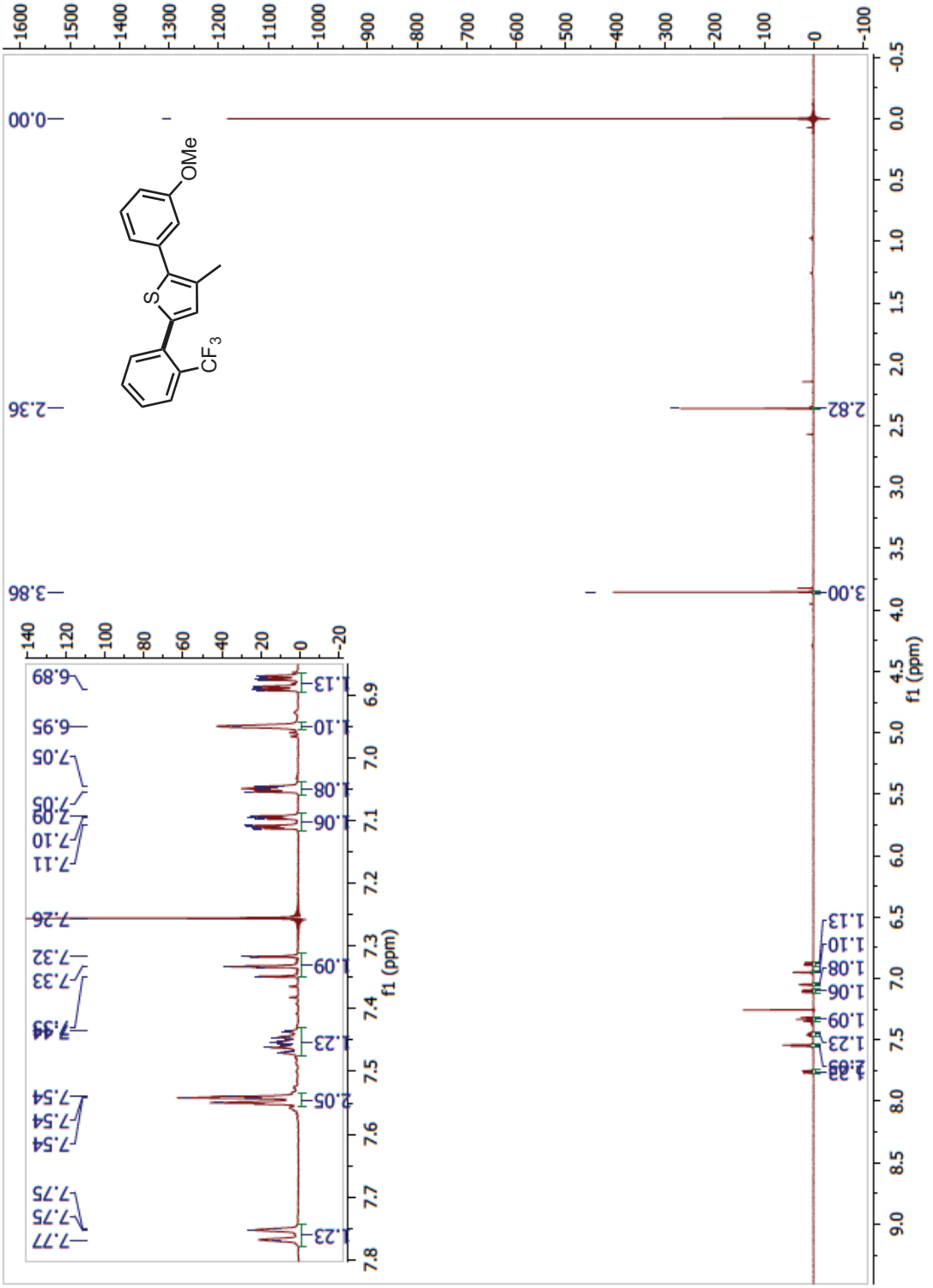
2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzaldehyde (99d; Table 6.2, Entry 8) ^{1}H:



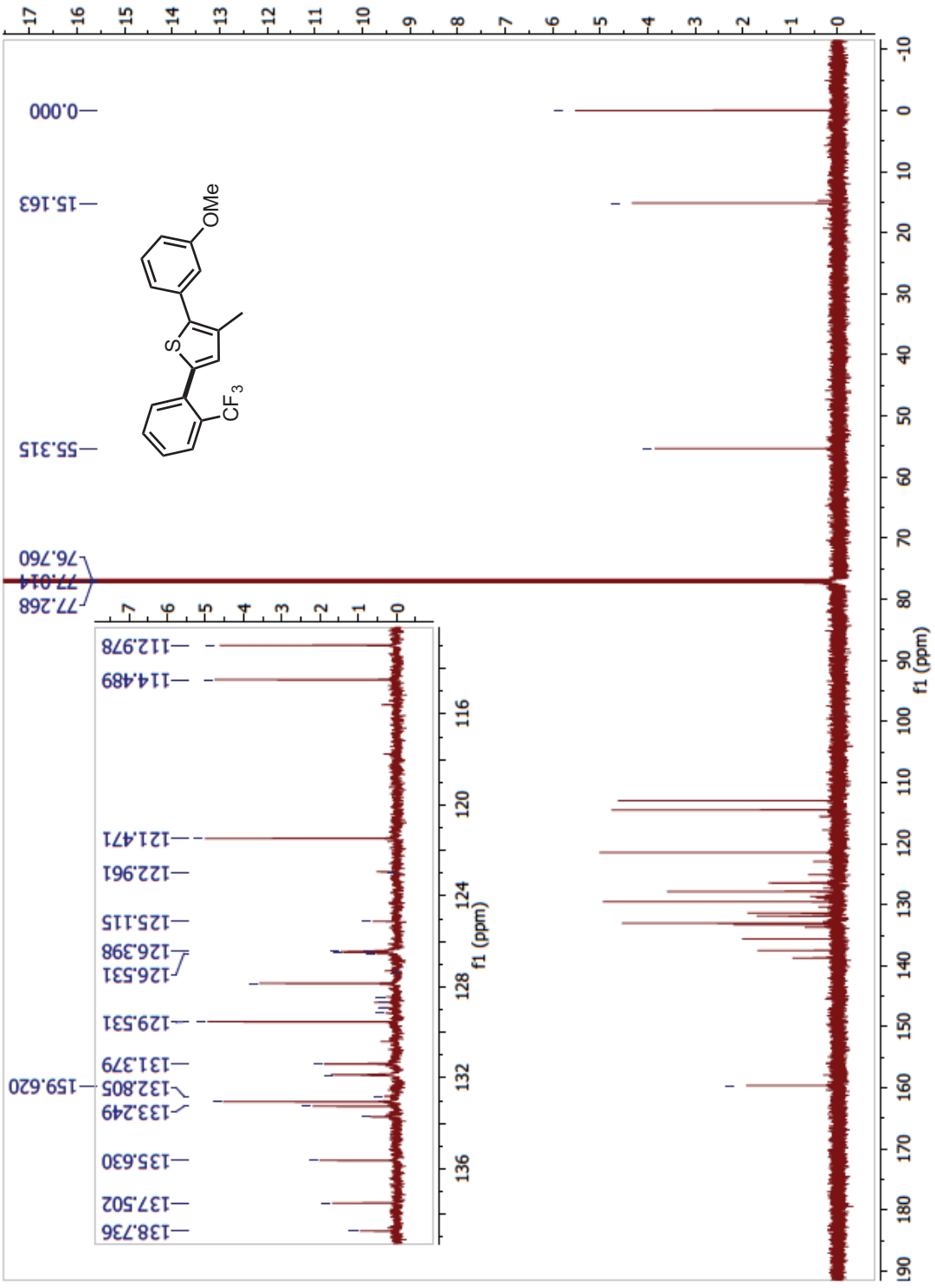
2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzaldehyde (99d; Table 6.2, Entry 8) ^{13}C:



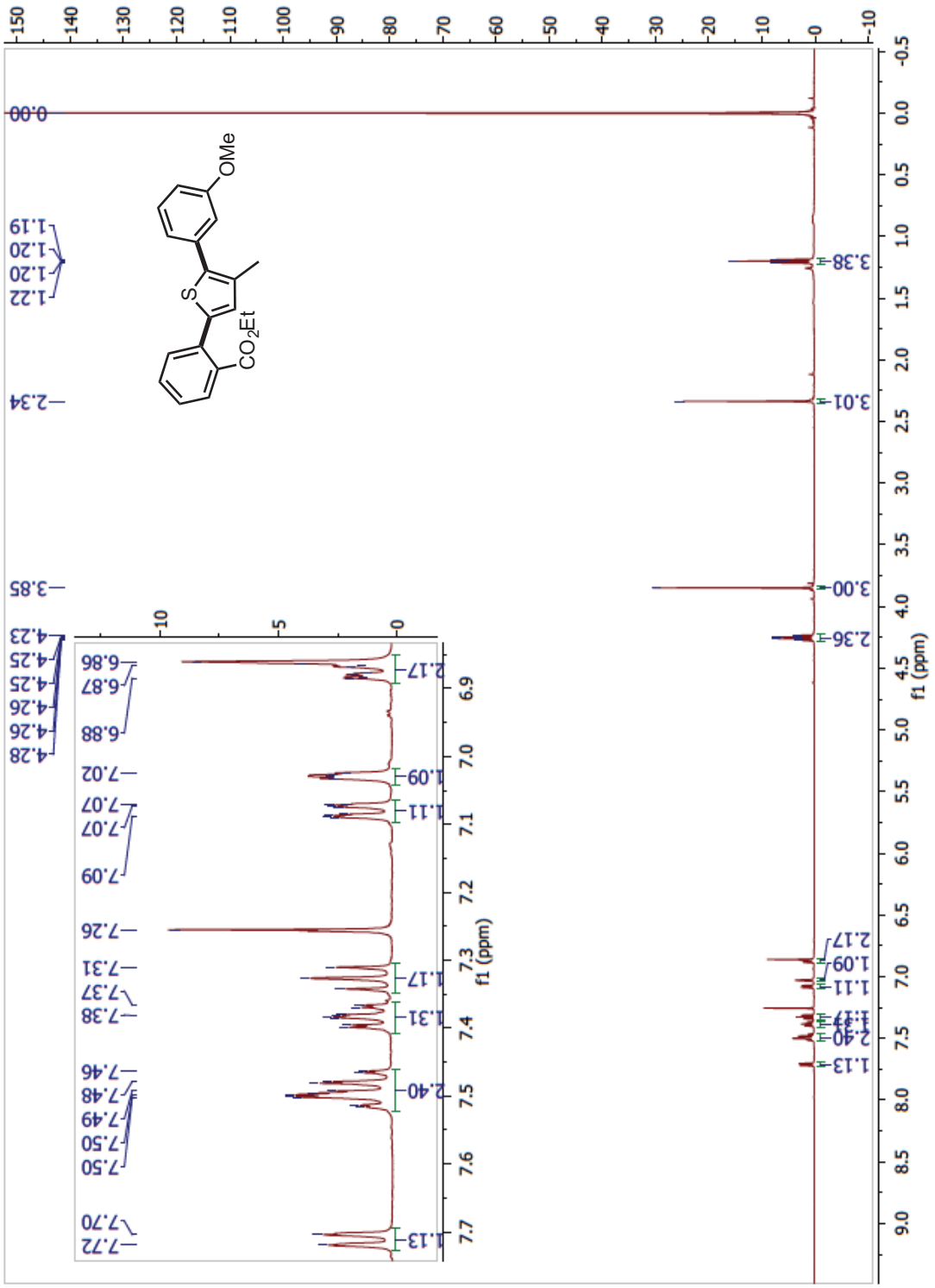
2-(3-methoxyphenyl)-3-methyl-5-(2-(trifluoromethyl)phenyl)thiophene (99e; Table 6.2, Entry 9) ^{1}H:



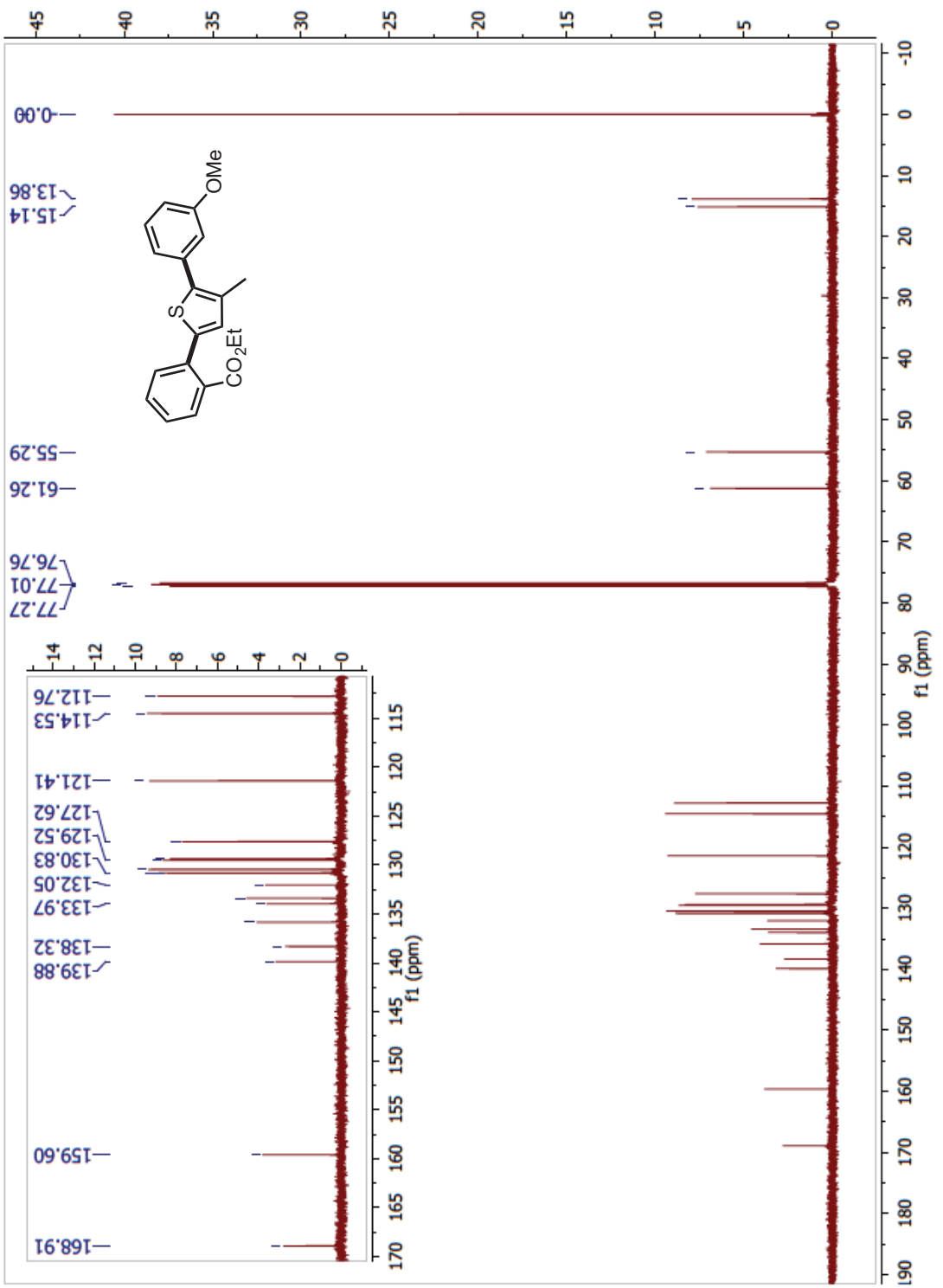
2-(3-methoxyphenyl)-3-methyl-5-(2-(trifluoromethyl)phenyl)thiophene (99e; Table 6.2, Entry 9) ^{13}C:



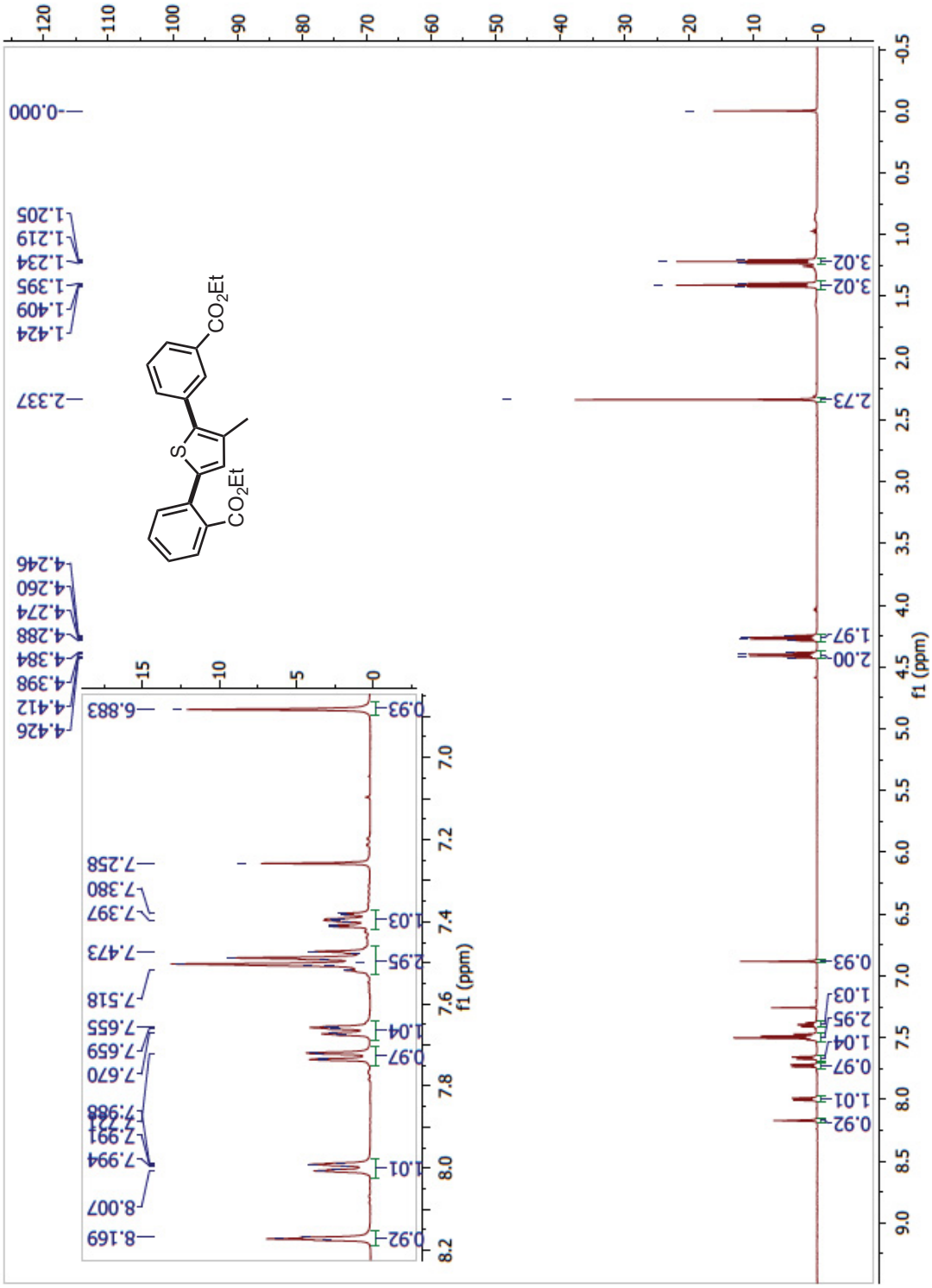
Ethyl 2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzoate (99f; Table 6.2, Entry 10) ^{1}H:



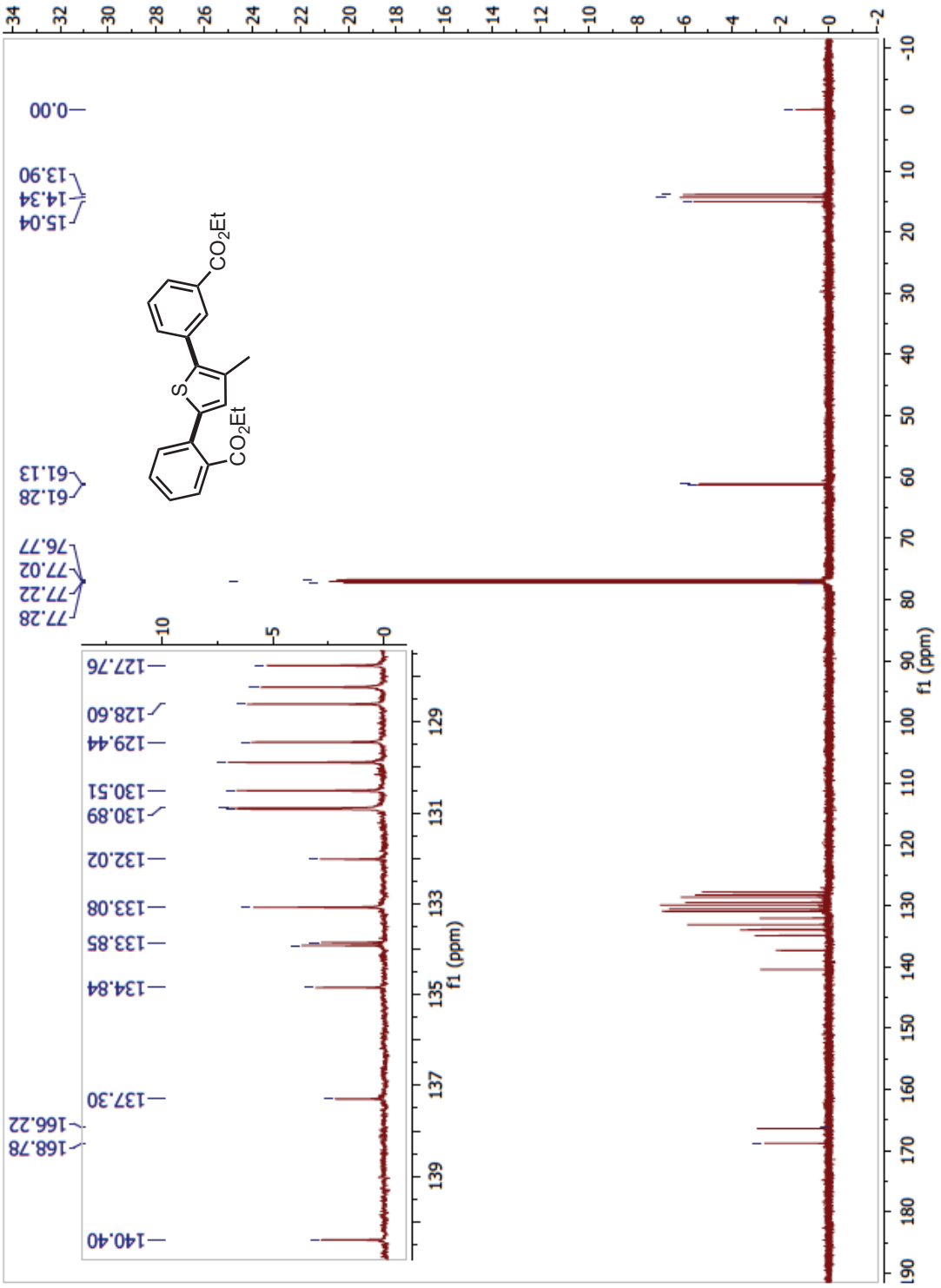
Ethyl 2-(5-(3-methoxyphenyl)-4-methylthiophen-2-yl)benzoate (99f; Table 6.2, Entry 10) ^{13}C:



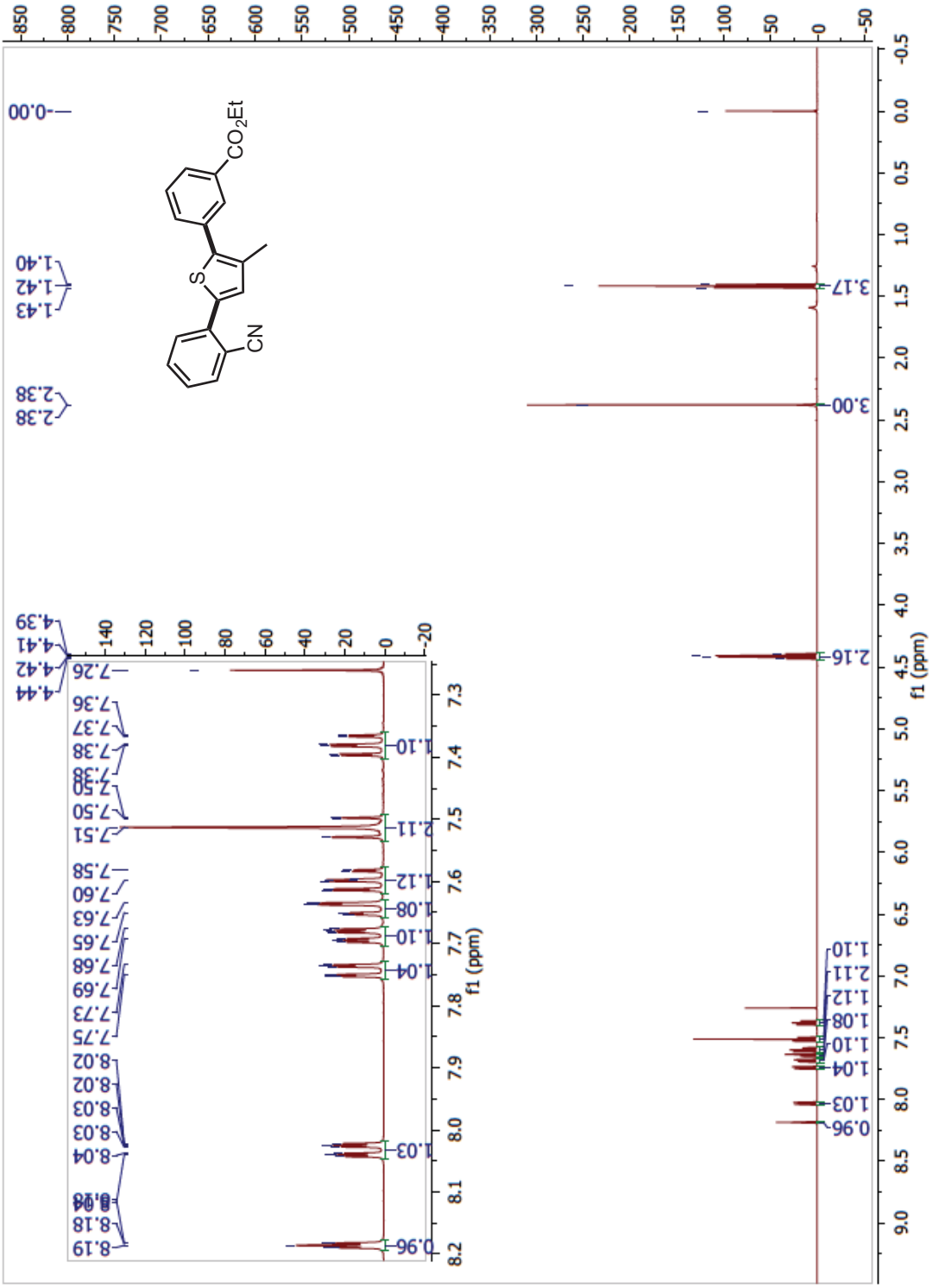
Ethyl 2-(5-(3-(ethoxycarbonyl)phenyl)-4-methylthiophen-2-yl)benzoate (99g; Table 6.2, Entry 11) ^{1}H:



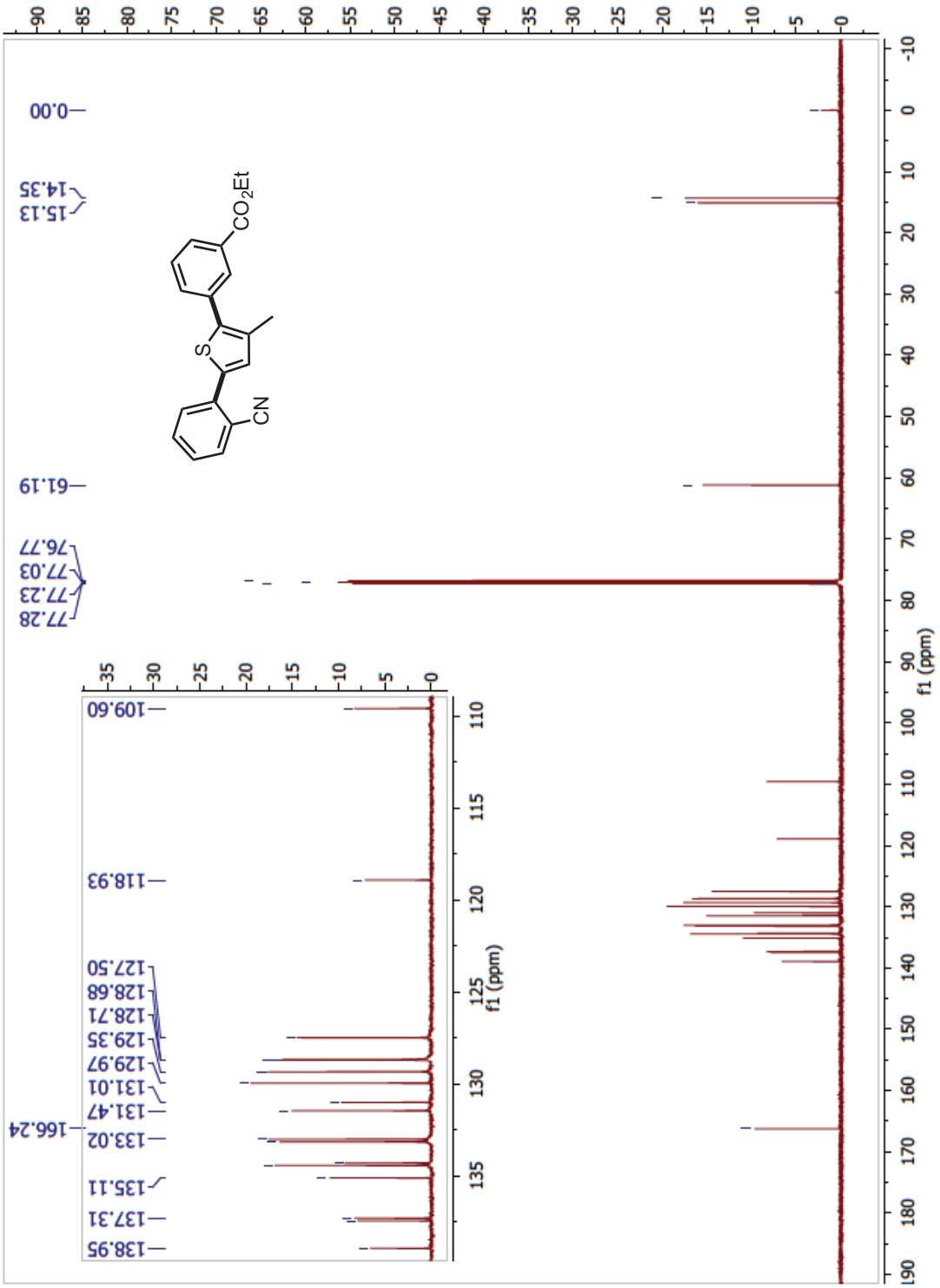
Ethyl 2-(5-(3-(ethoxycarbonyl)phenyl)-4-methylthiophen-2-yl)benzoate (99g; Table 6.2, Entry 11) ^{13}C:



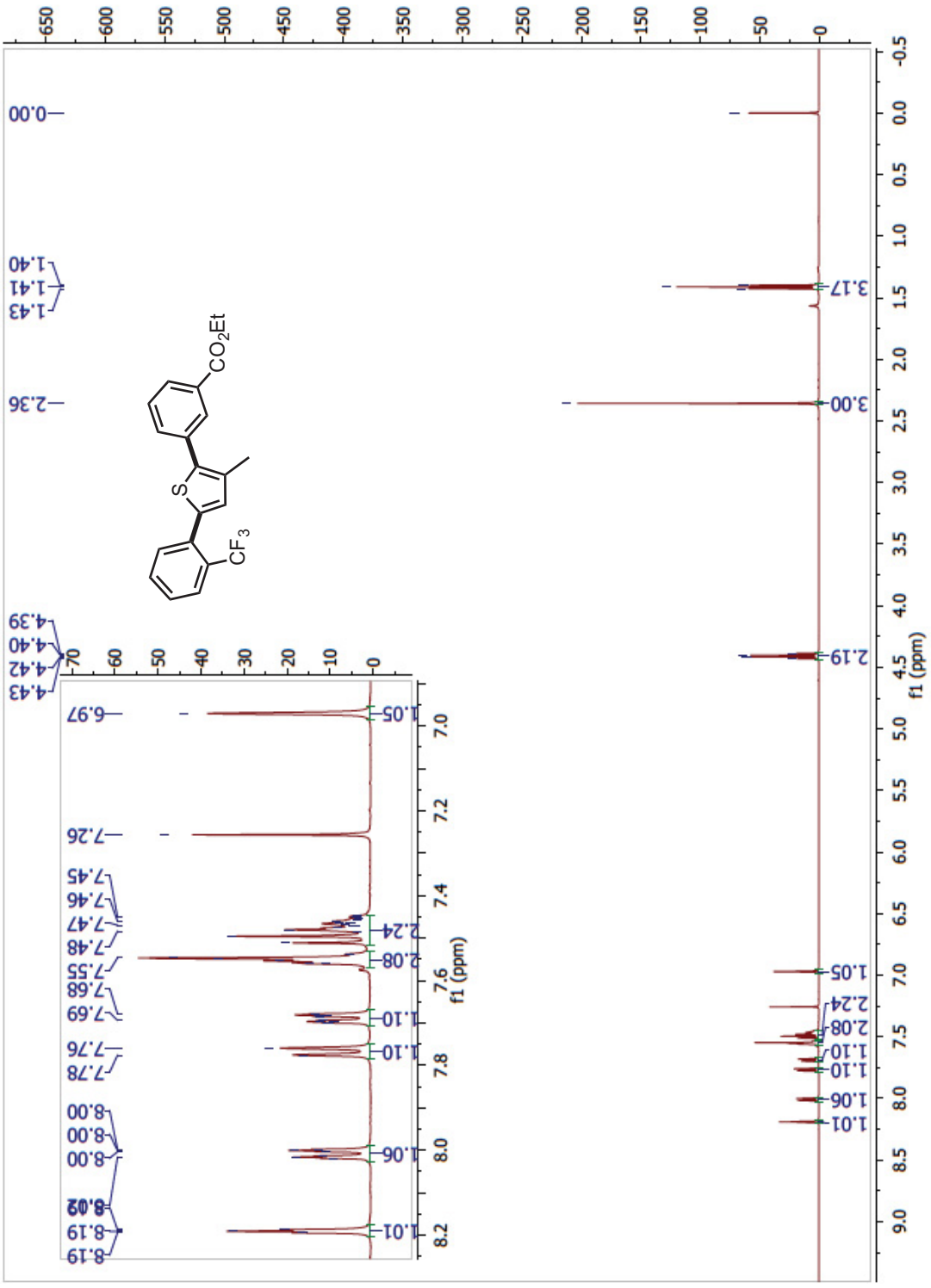
Ethyl 3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoate (99h; Table 6.2, Entry 12) ^{1}H:



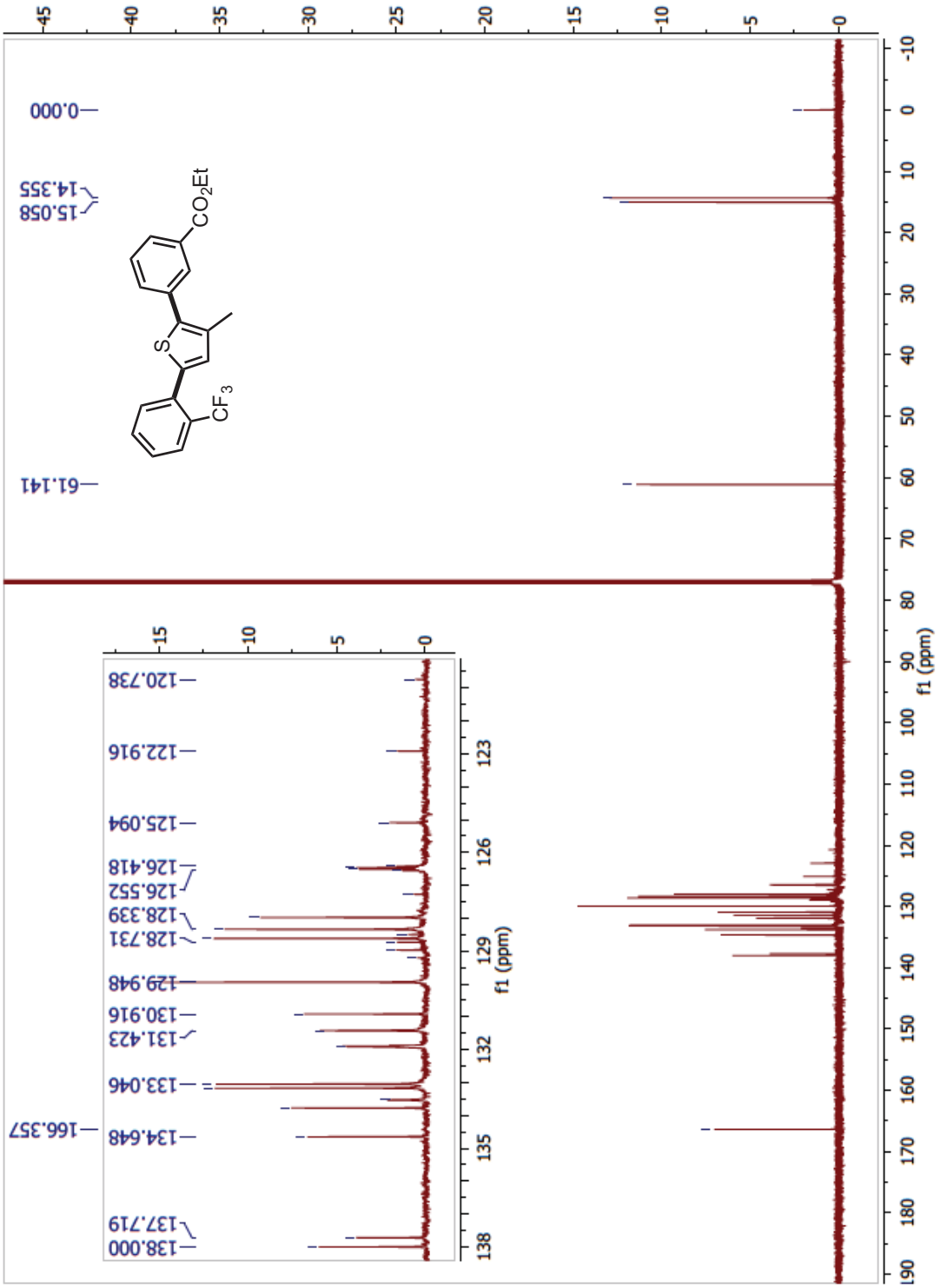
Ethyl 3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoate (99h; Table 6.2, Entry 12) ^{13}C:



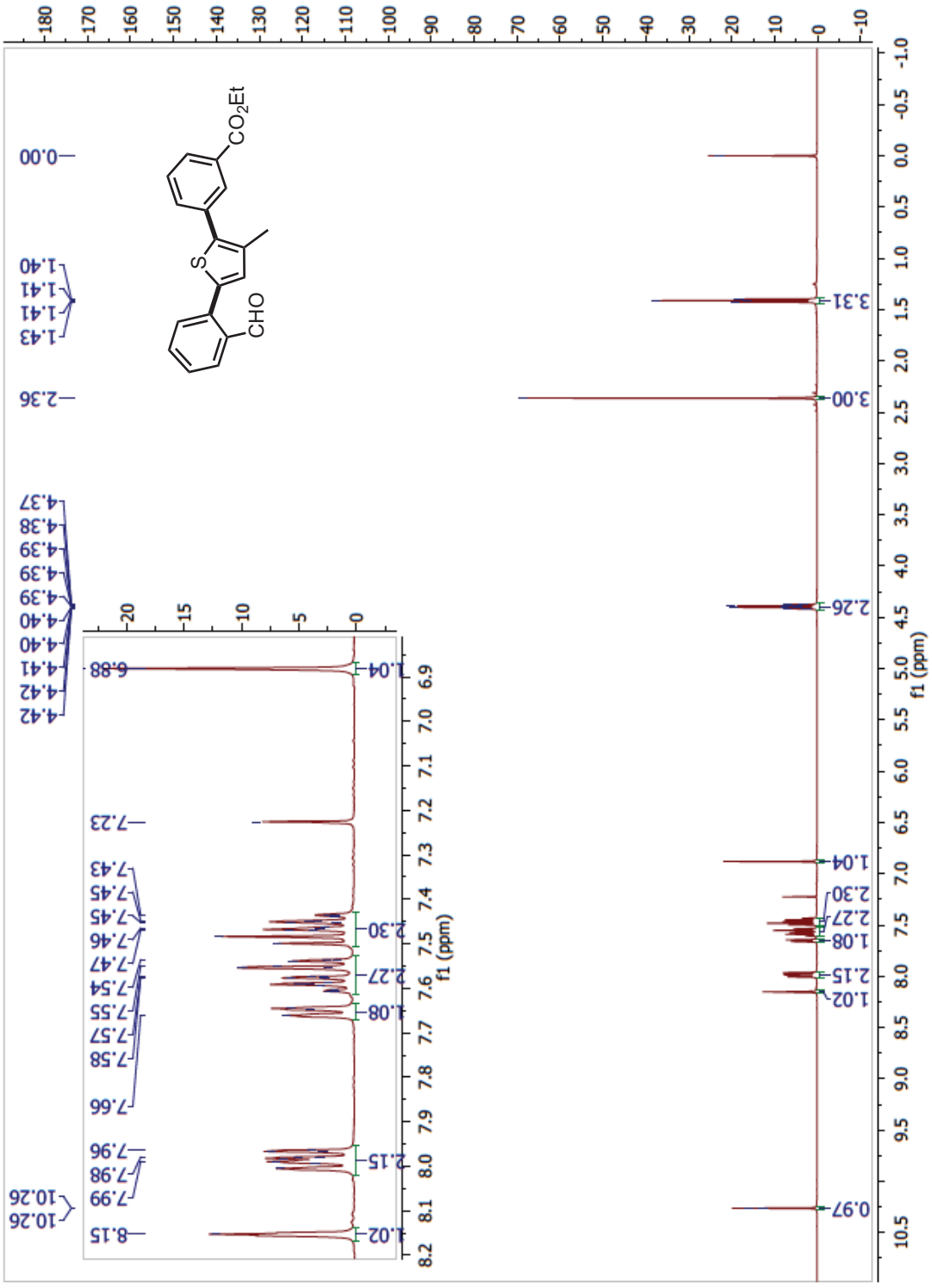
Ethyl 3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99i; Table 6.2, Entry 13) ^{1}H:



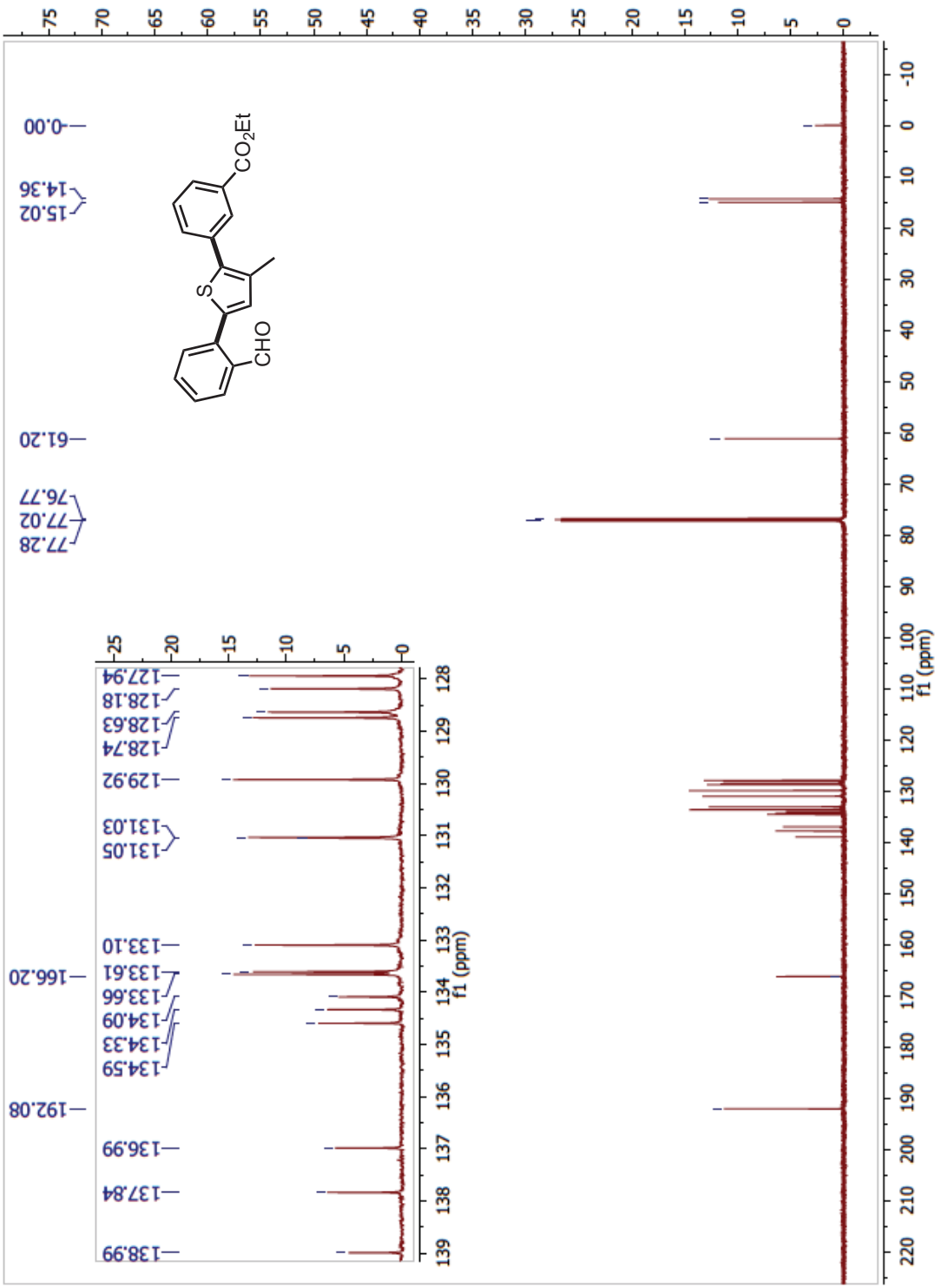
Ethyl 3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99i; Table 6.2, Entry 13) ^{13}C:



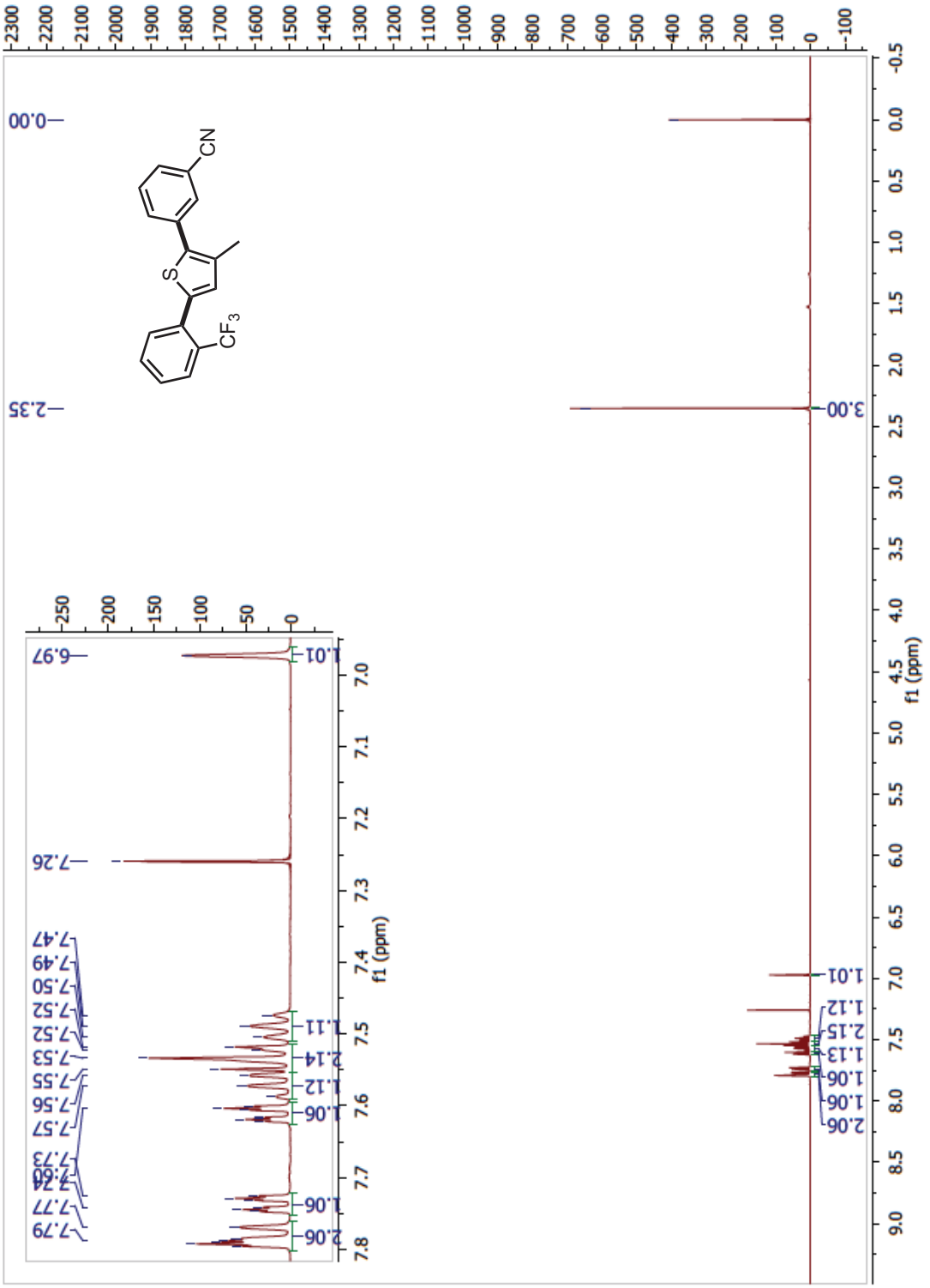
Ethyl 3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoate (99j; Table 6.2, Entry 14) ^{1}H:



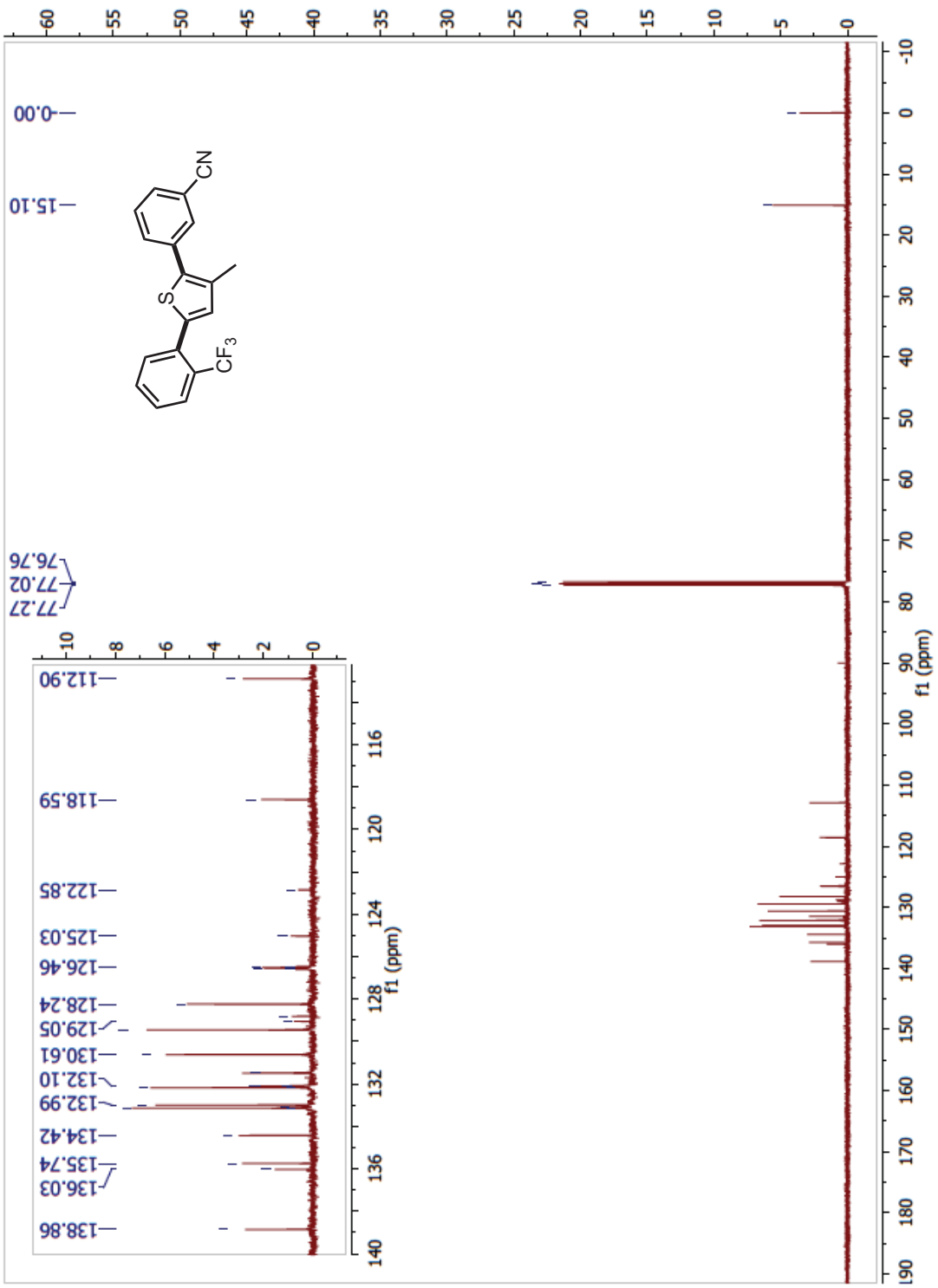
Ethyl 3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoate (99j; Table 6.2, Entry 14) ^{13}C:



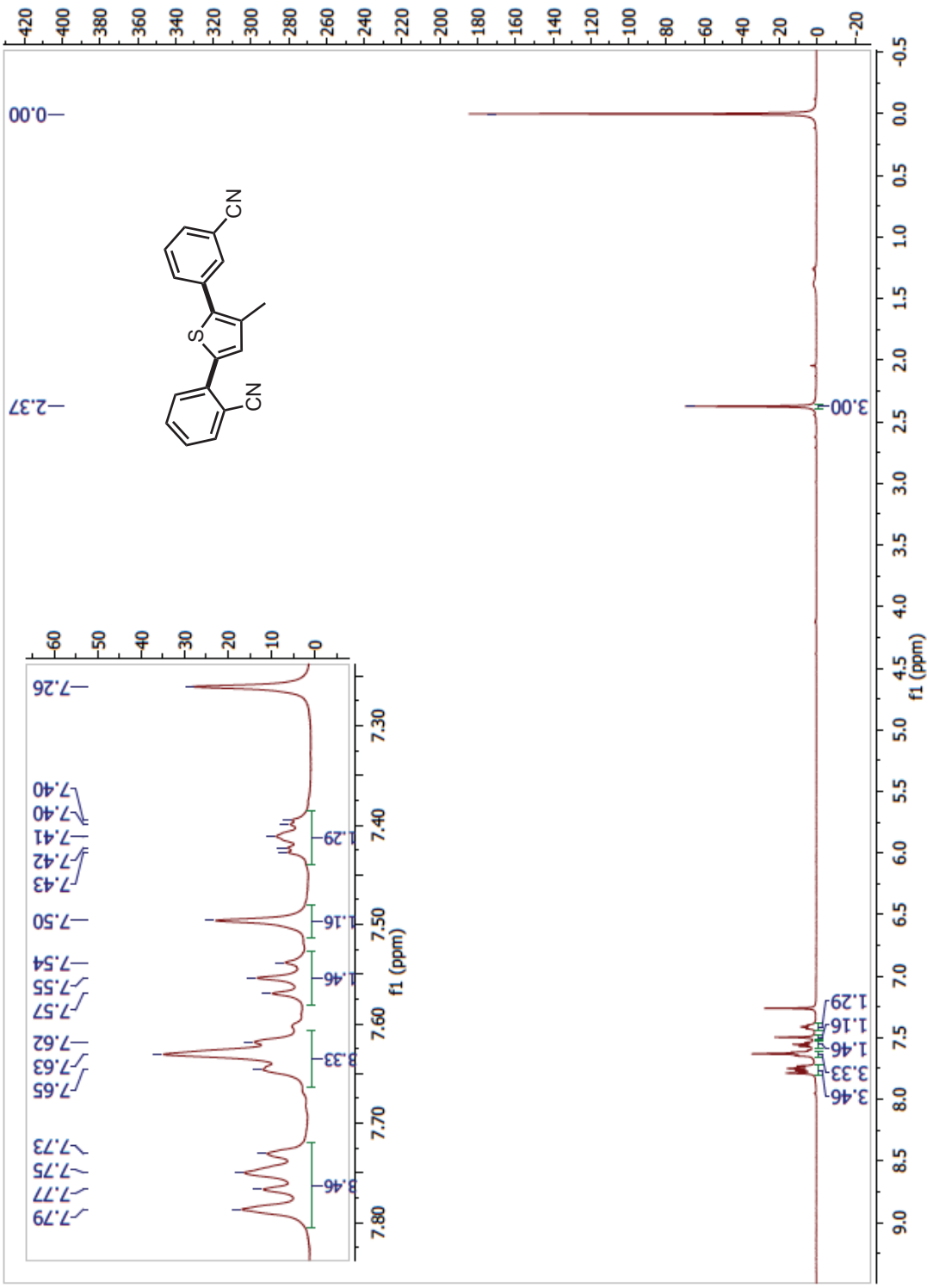
3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzonitrile (99k; Table 6.2, Entry 15) ^{1}H:



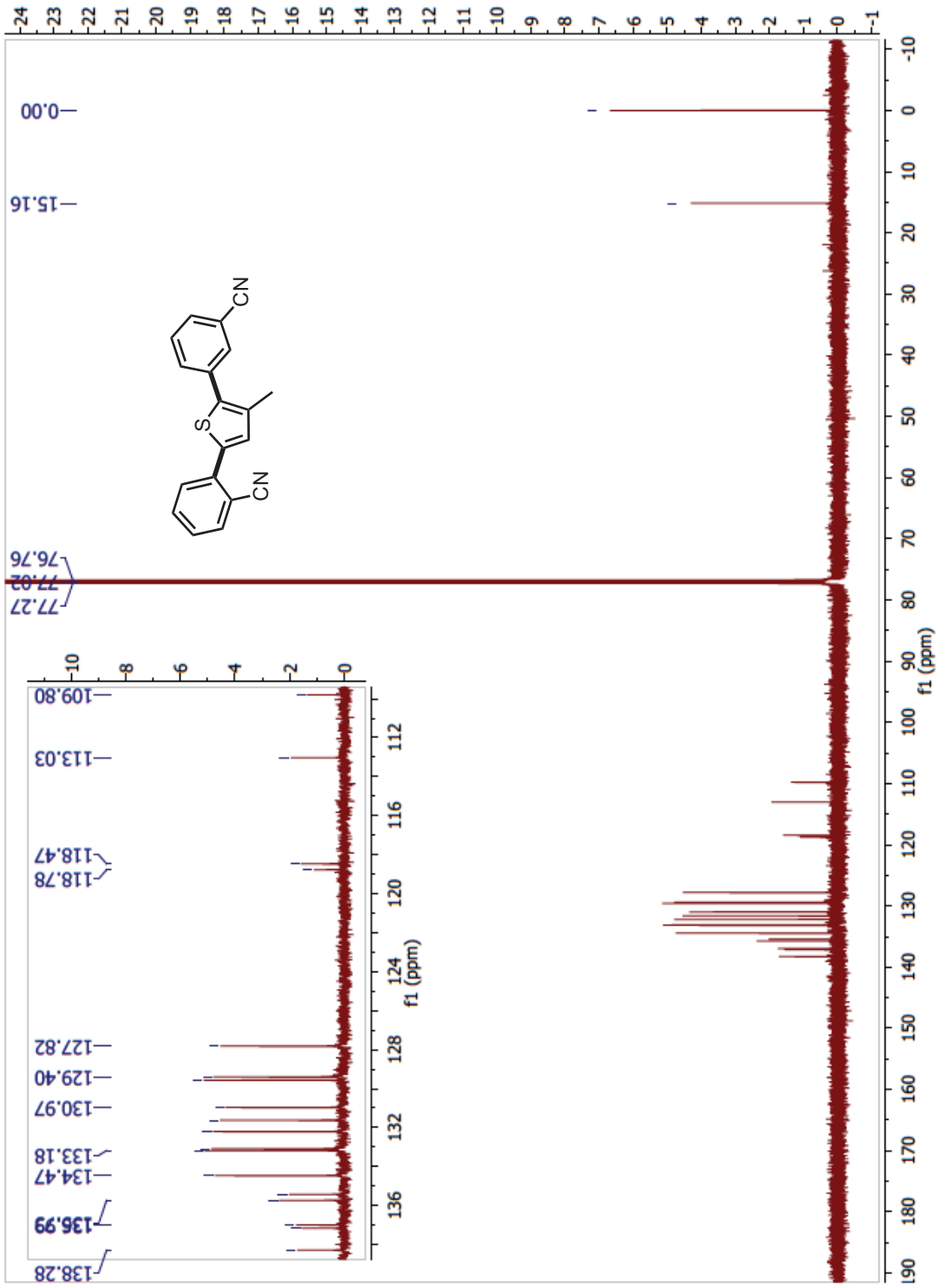
3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzonitrile (99k; Table 6.2, Entry 15) ^{13}C:



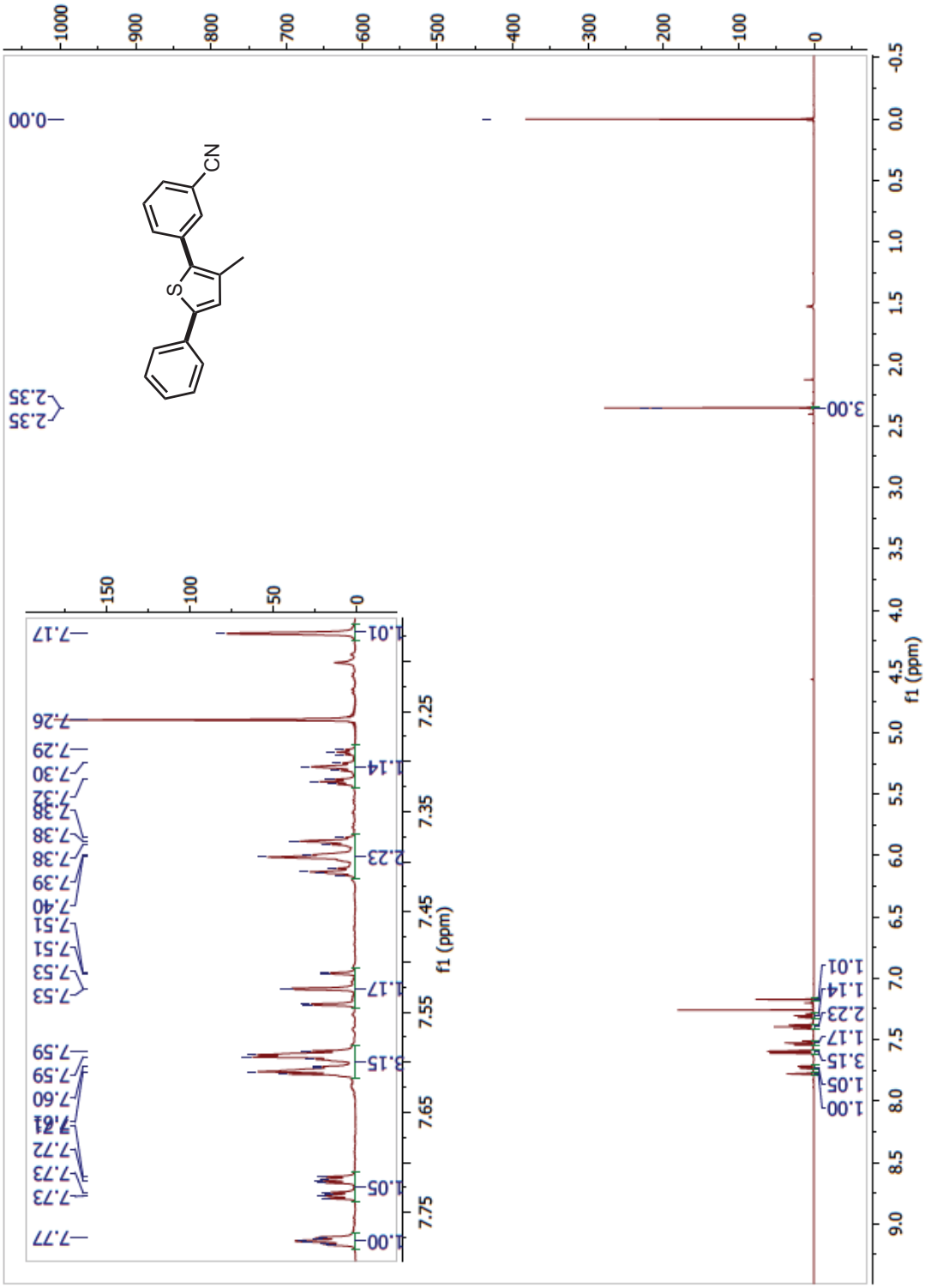
2-(5-(3-cyanophenyl)-4-methylthiophen-2-yl)benzonitrile (99l; Table 6.2, Entry 16)
^{1}H:



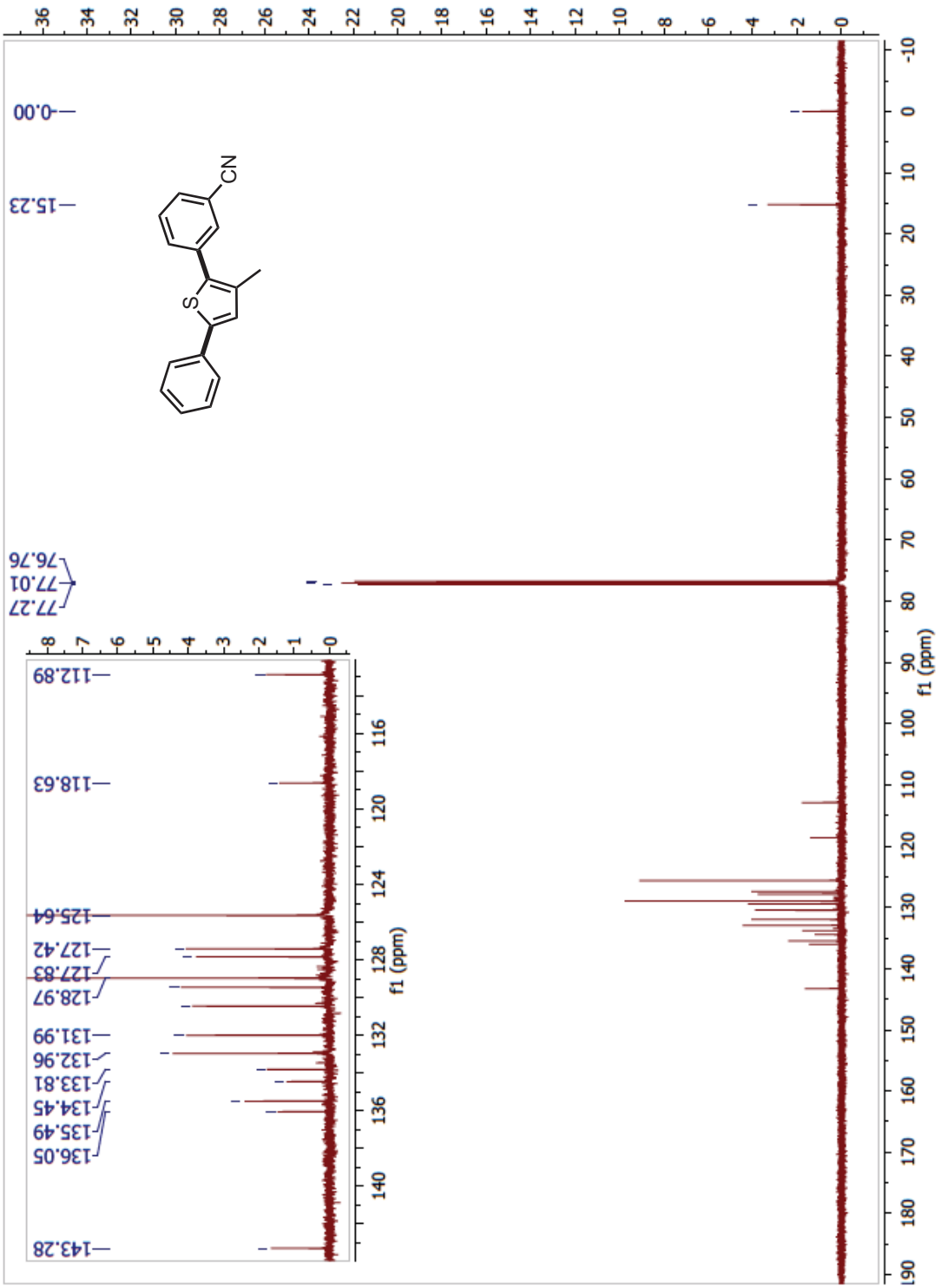
2-(5-(3-cyanophenyl)-4-methylthiophen-2-yl)benzonitrile (99l; Table 6.2, Entry 16)
^{13}C:



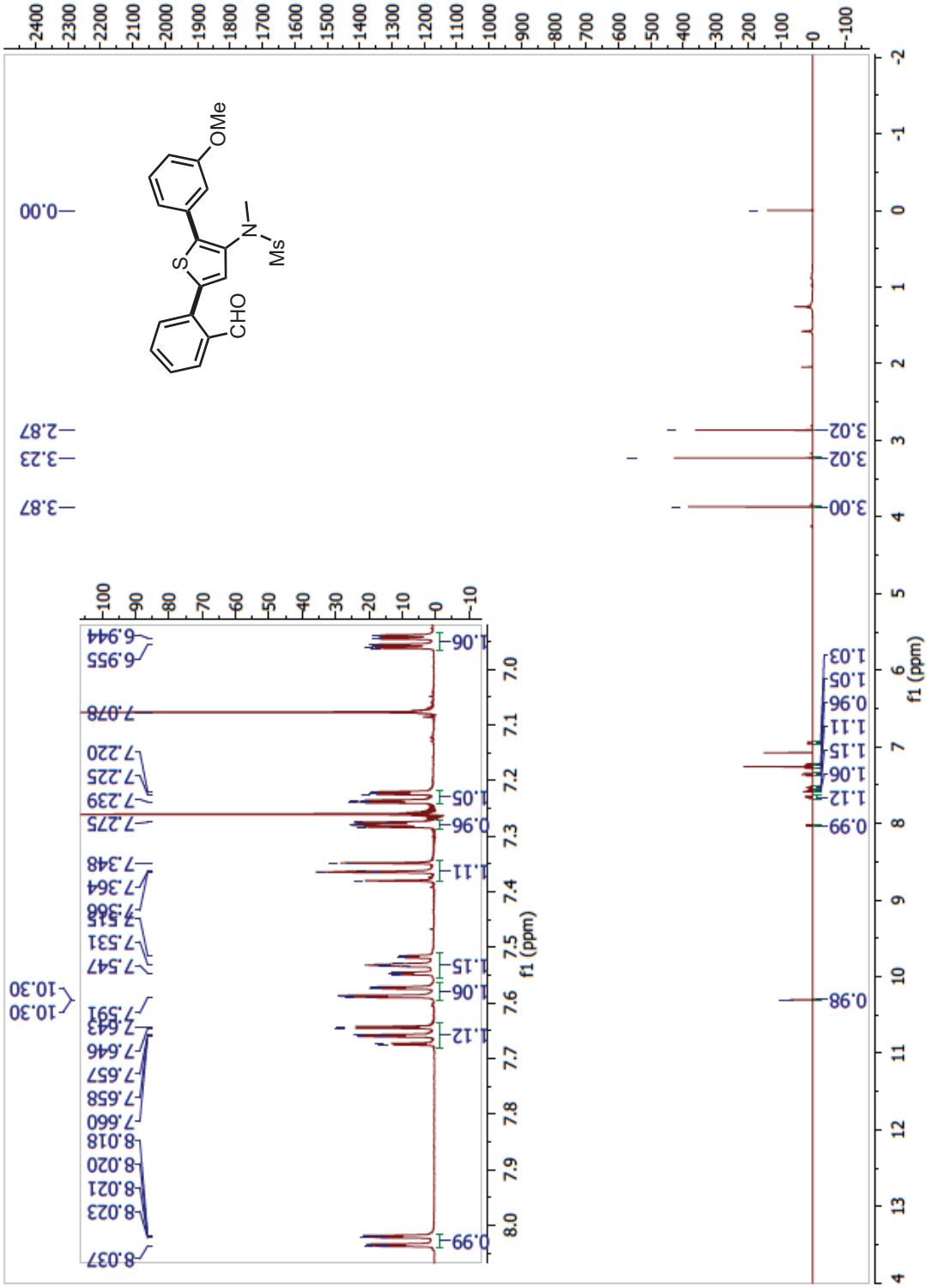
3-(3-methyl-5-phenylthiophen-2-yl)benzonitrile (99m; Table 6.2, Entry 17) ^{1}H:



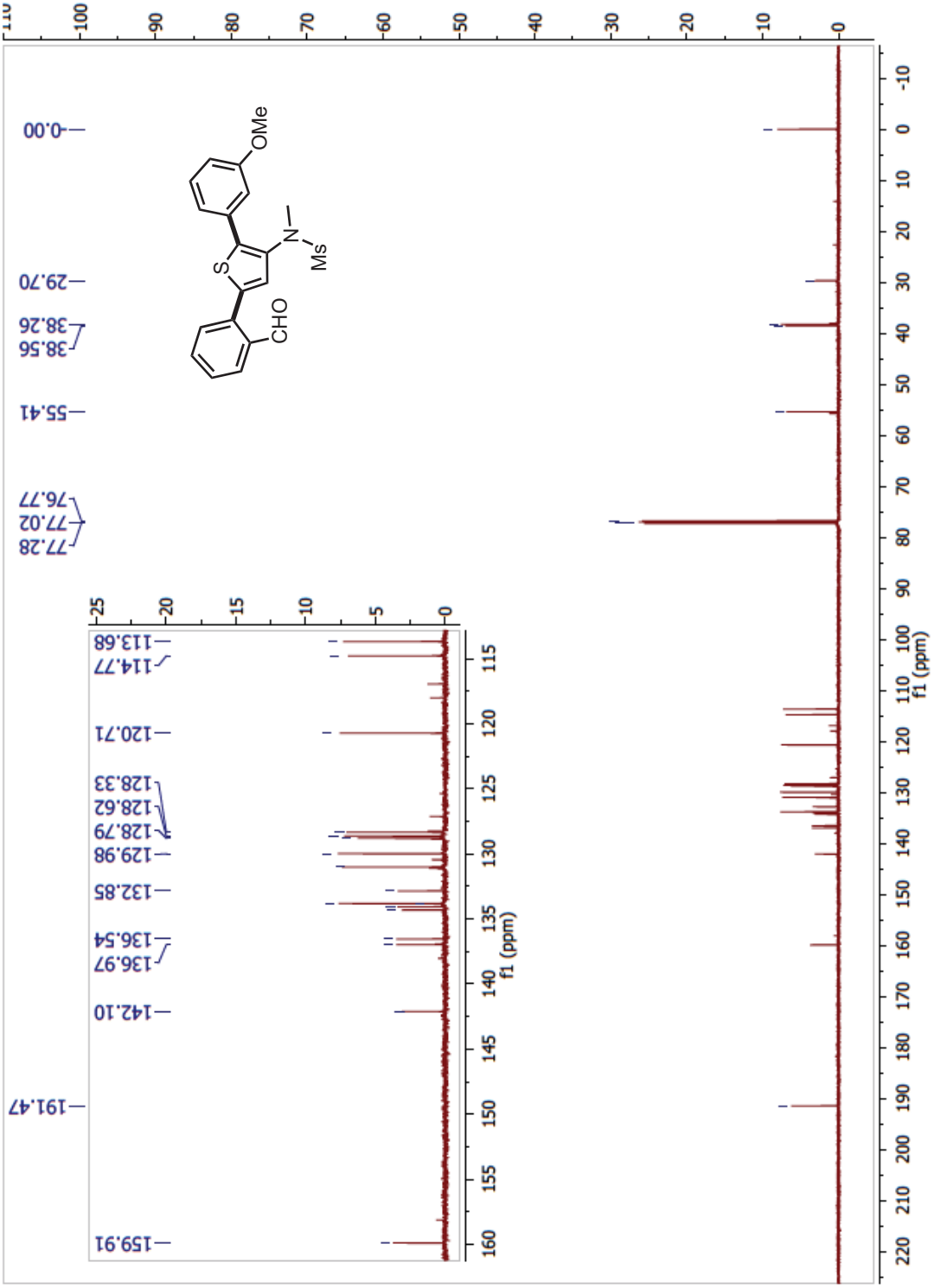
3-(3-methyl-5-phenylthiophen-2-yl)benzonitrile (99m; Table 6.2, Entry 17) ^{13}C:



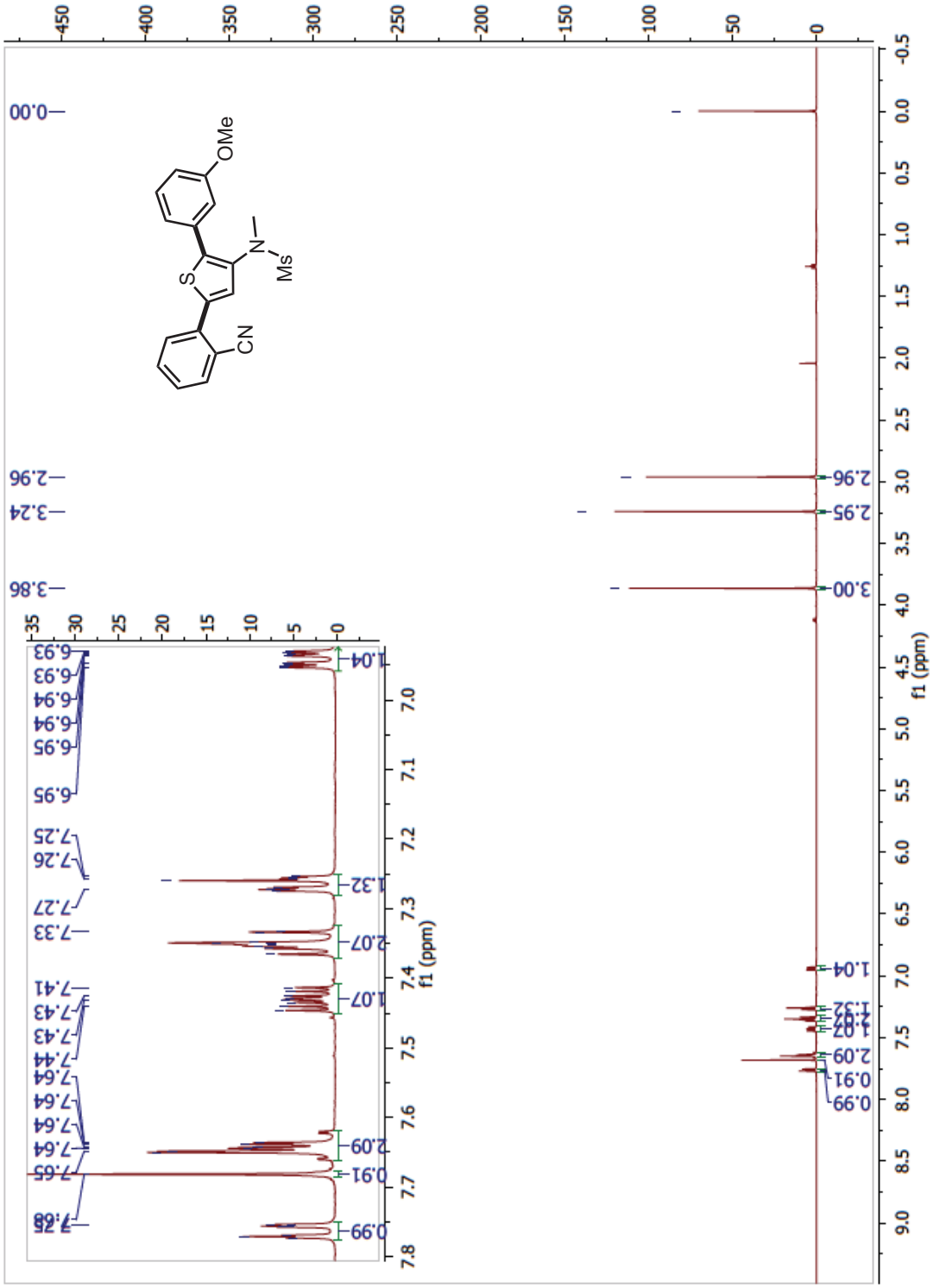
***N*-(5-(2-formylphenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N* methylmethanesulfonamide (99b; Table 6.3, Entry 19) ^{1}H:**



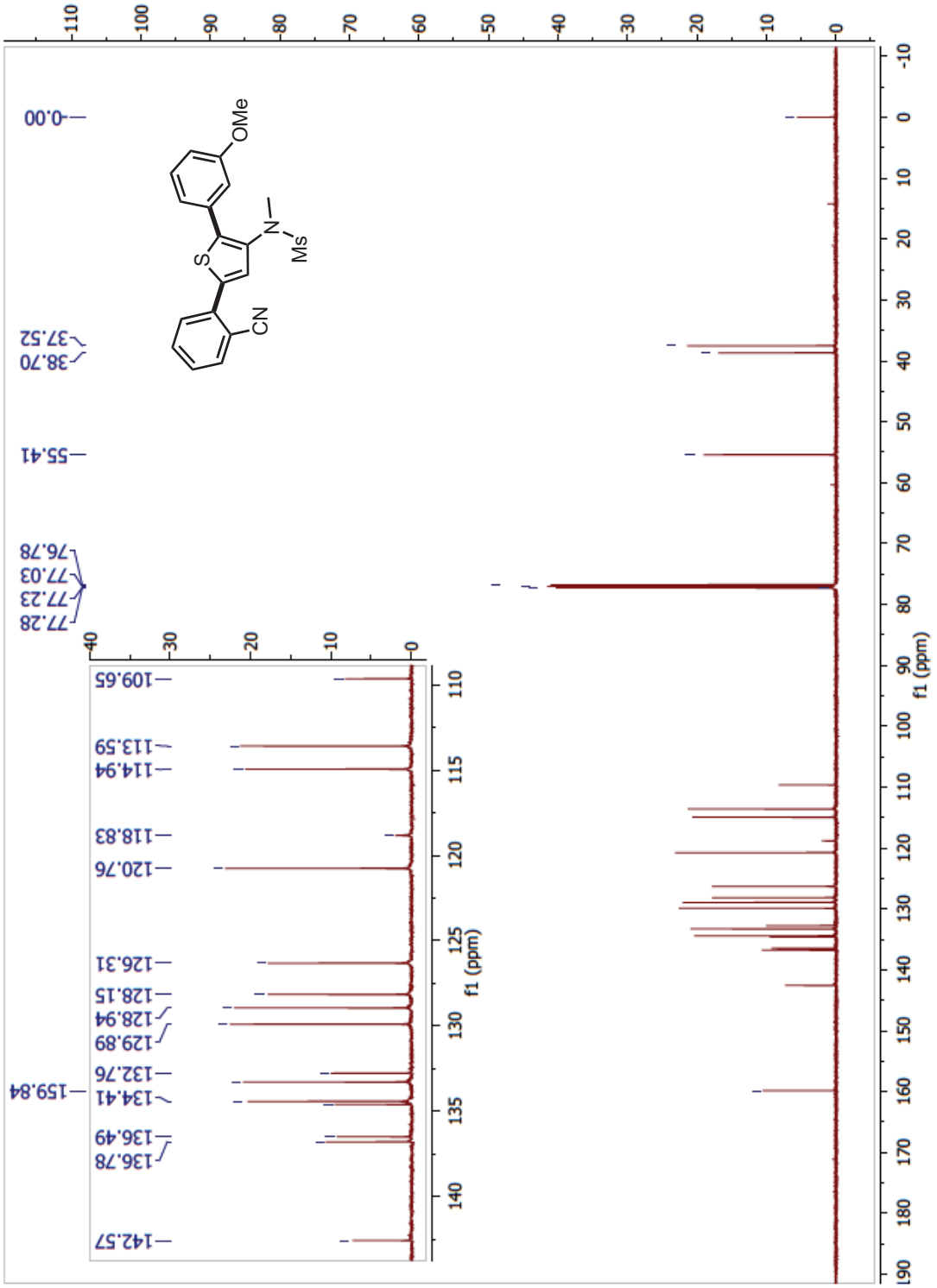
***N*-(5-(2-formylphenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99b; Table 6.3, Entry 19) ^{13}C:**



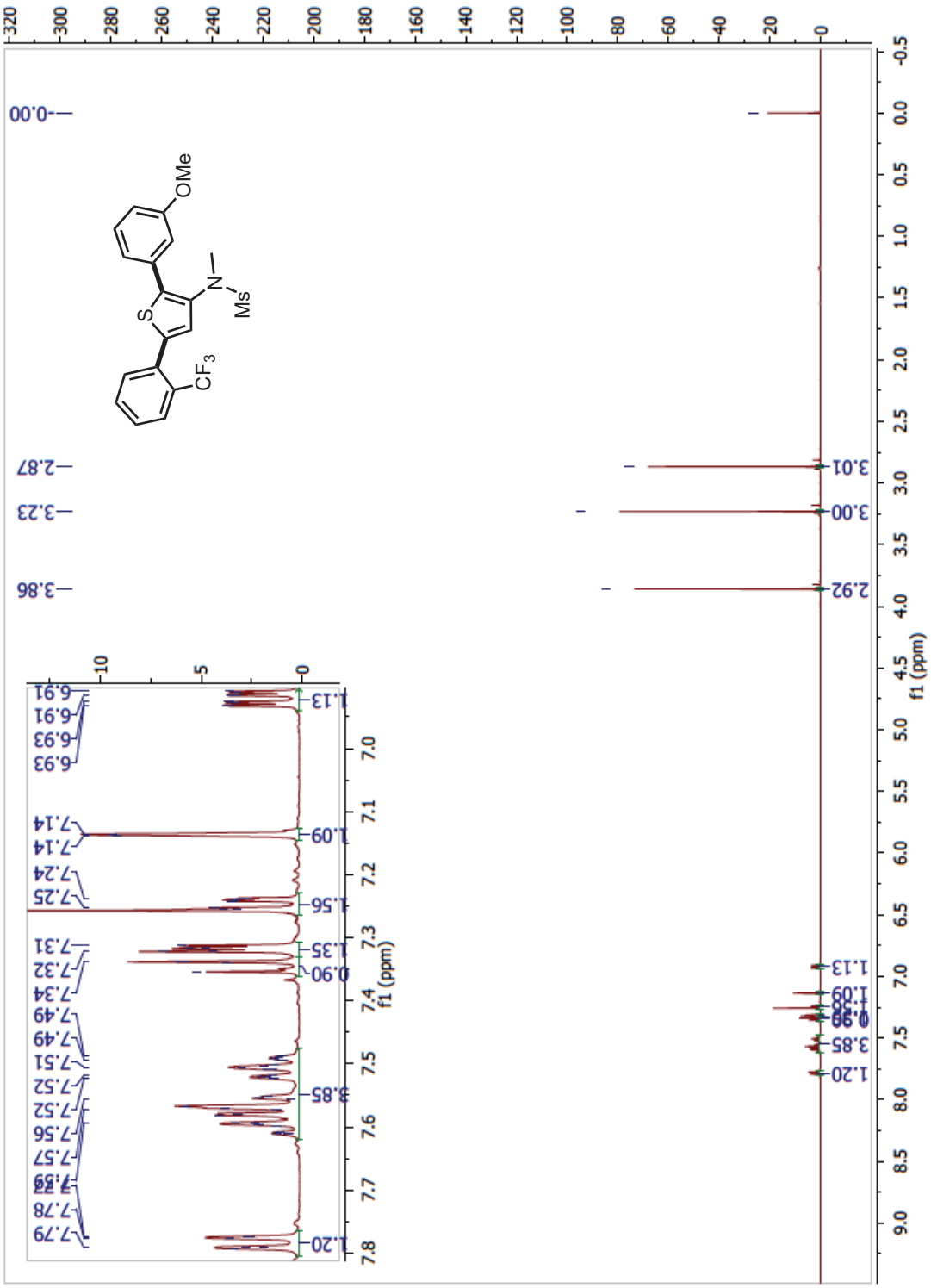
***N*-(5-(2-cyanophenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99n; Table 6.3, Entry 20) ^{1}H:**



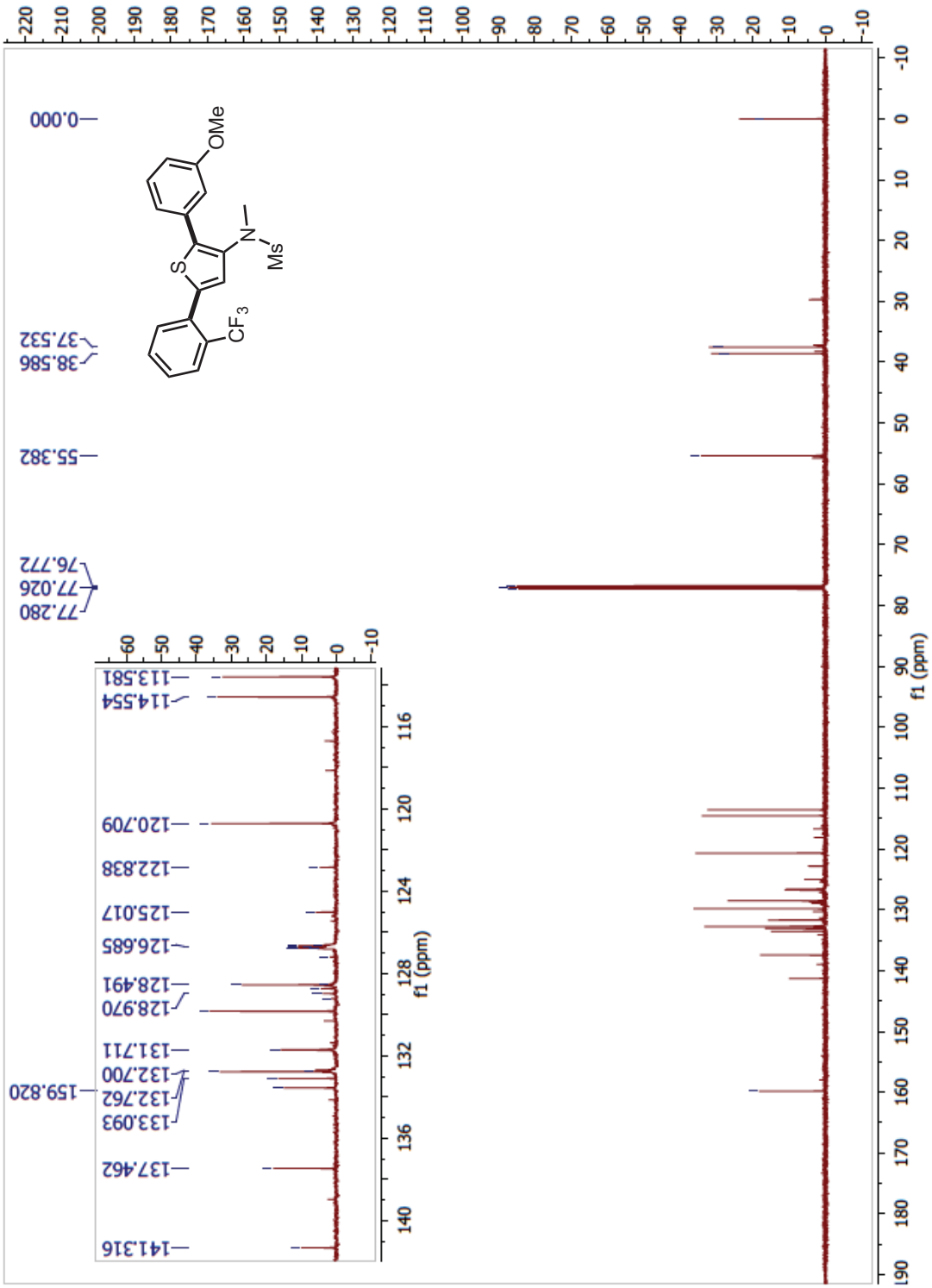
***N*-(5-(2-cyanophenyl)-2-(3-methoxyphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99n; Table 6.3, Entry 20) ^{13}C:**



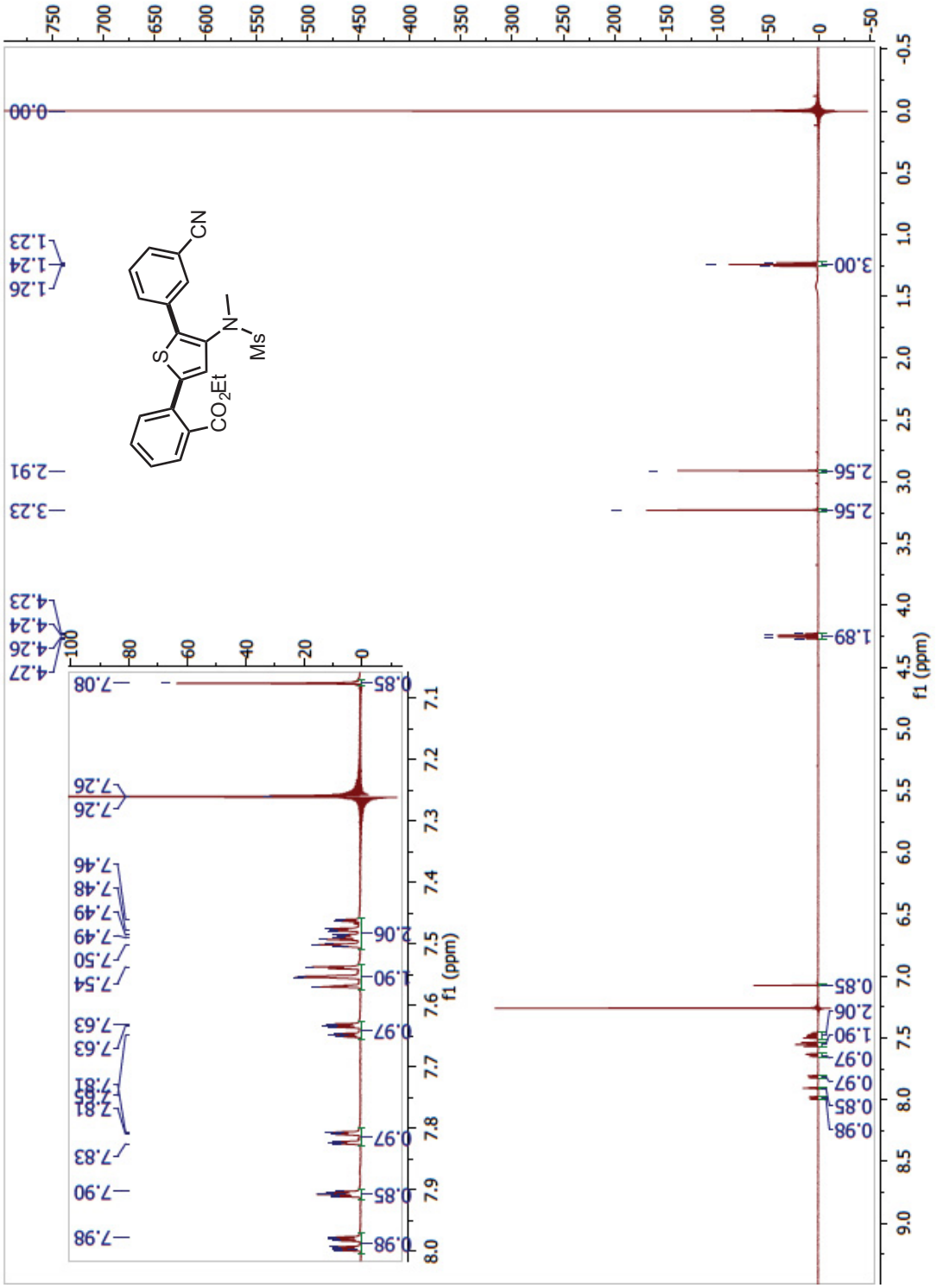
***N*-(2-(3-methoxyphenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-N-methylmethanesulfonamide (99o; Table 6.3, Entry 21) ^{1}H:**



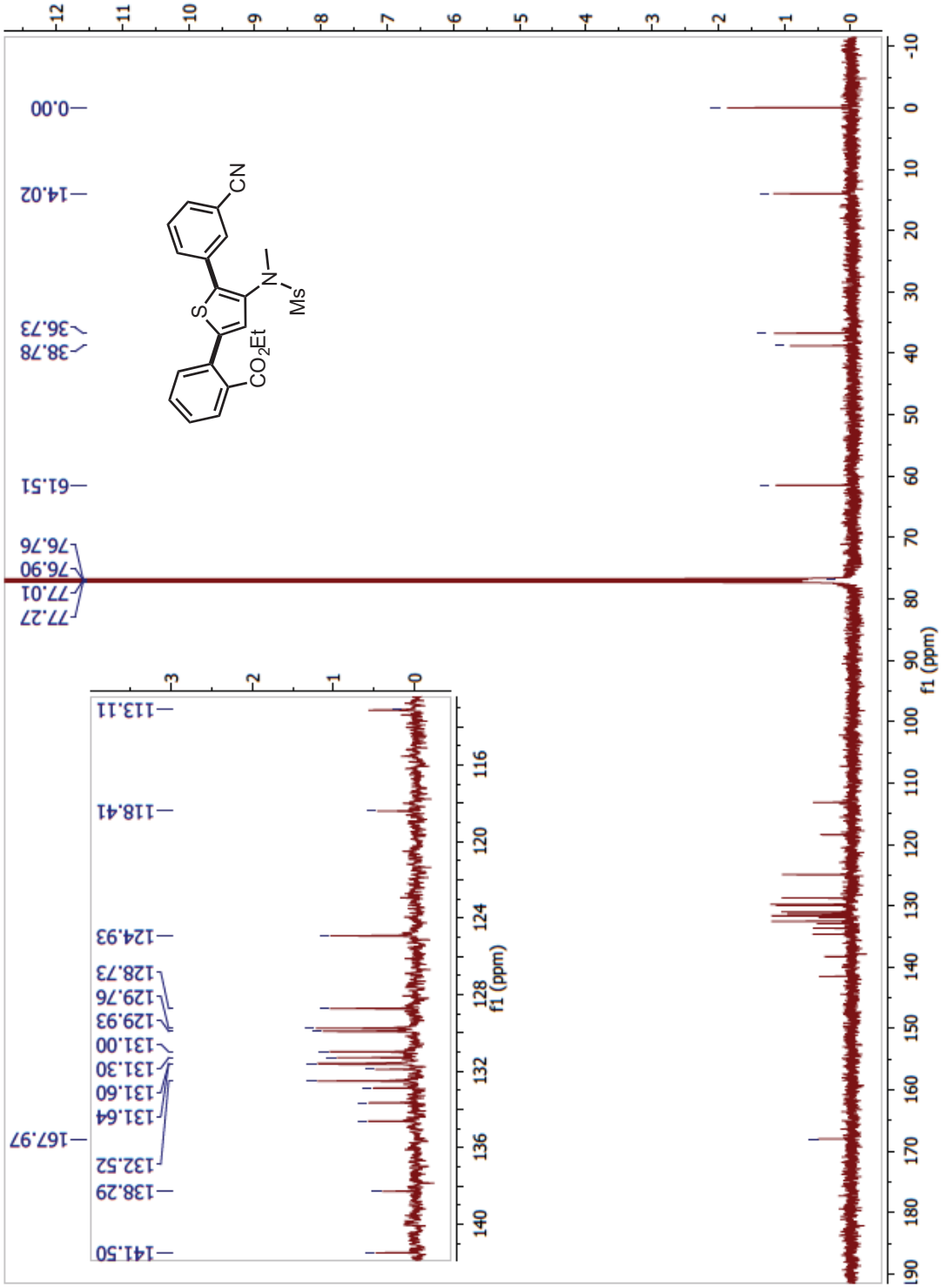
***N*-(2-(3-methoxyphenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-N-methylmethanesulfonamide (99o; Table 6.3, Entry 21) ^{13}C:**



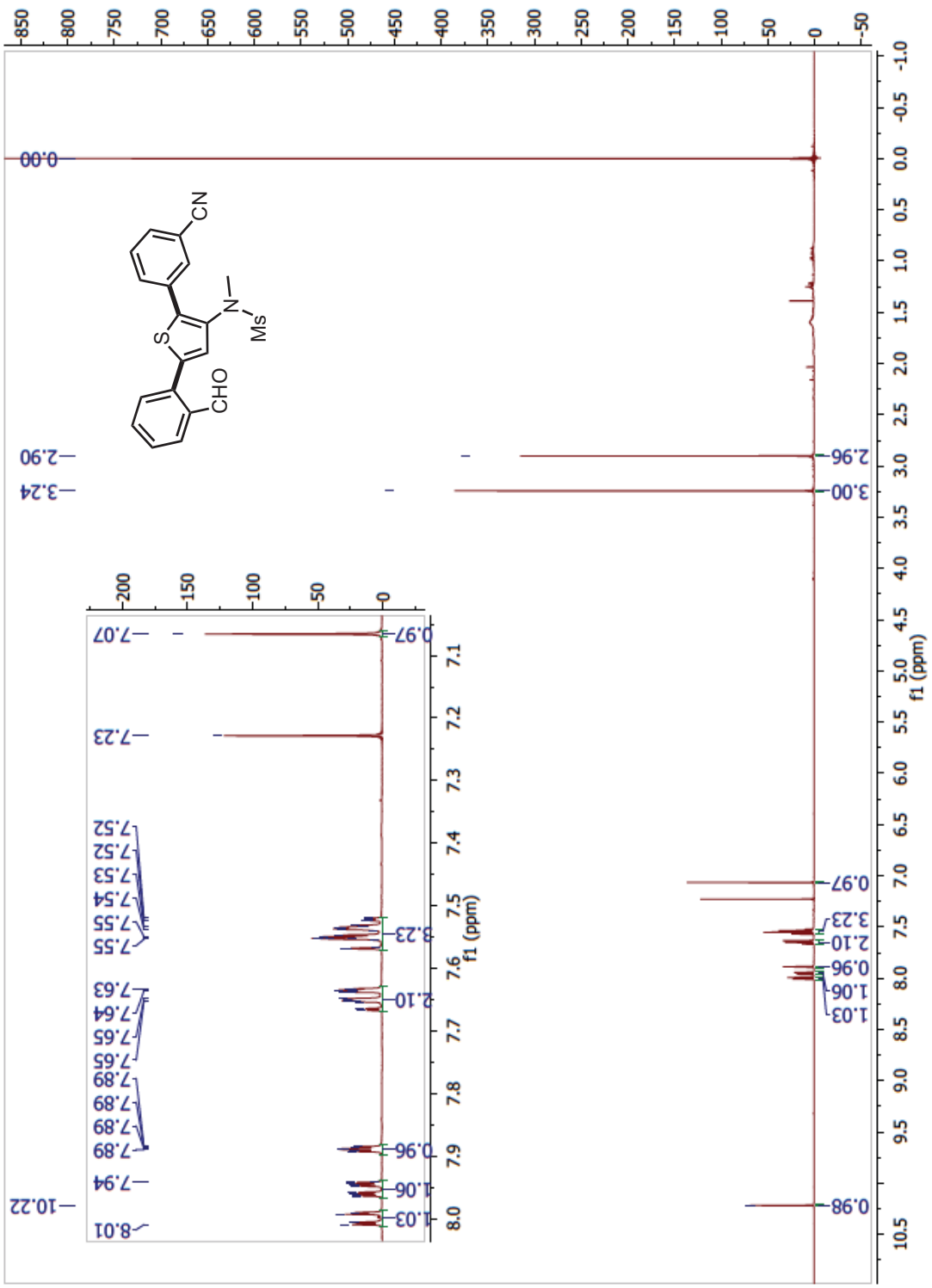
Ethyl 2-(5-(3-cyanophenyl)-4-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99p; Table 6.3, Entry 22) ^{1}H:



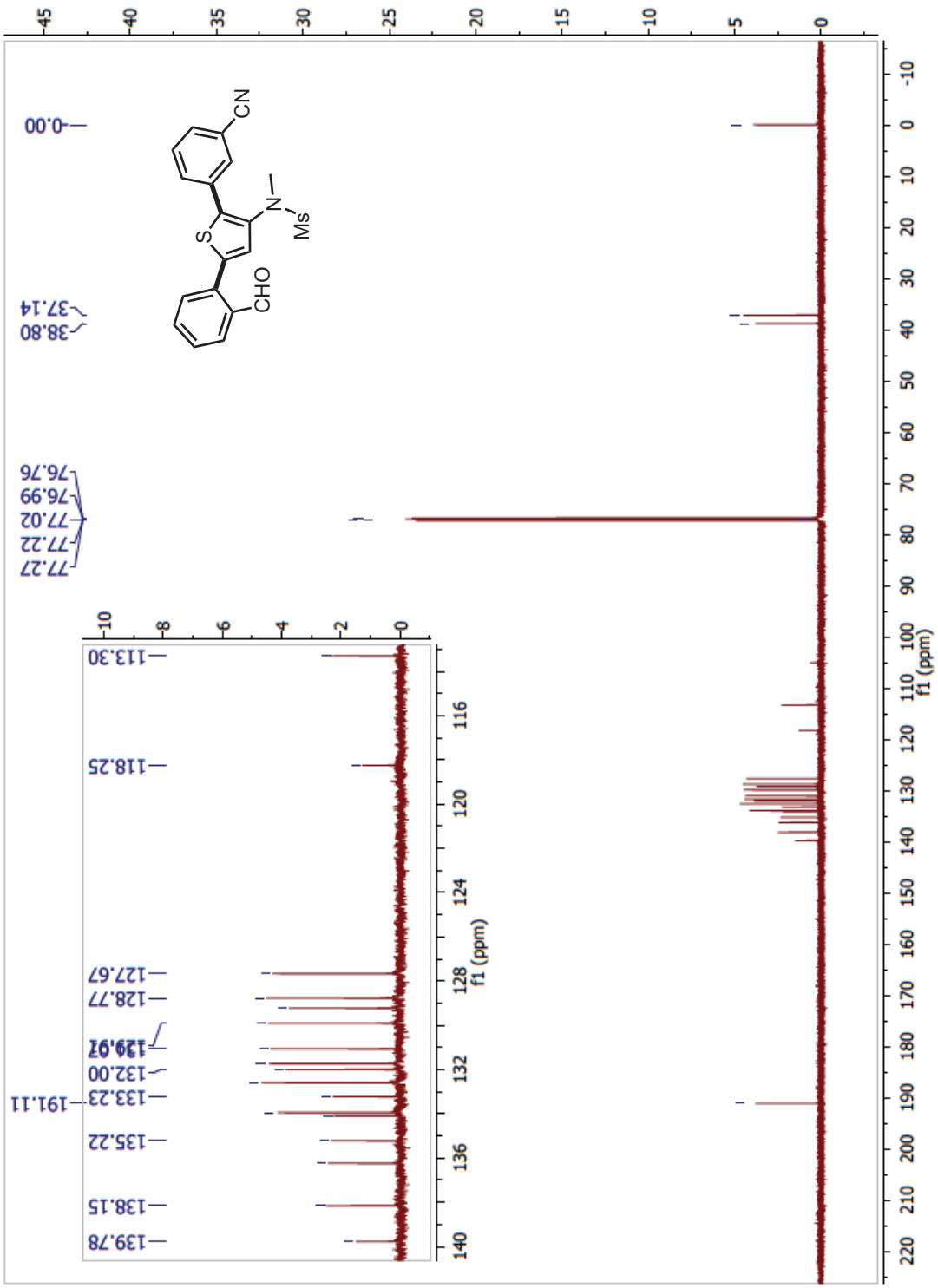
Ethyl 2-(5-(3-cyanophenyl)-4-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99p; Table 6.3, Entry 22) ^{13}C:



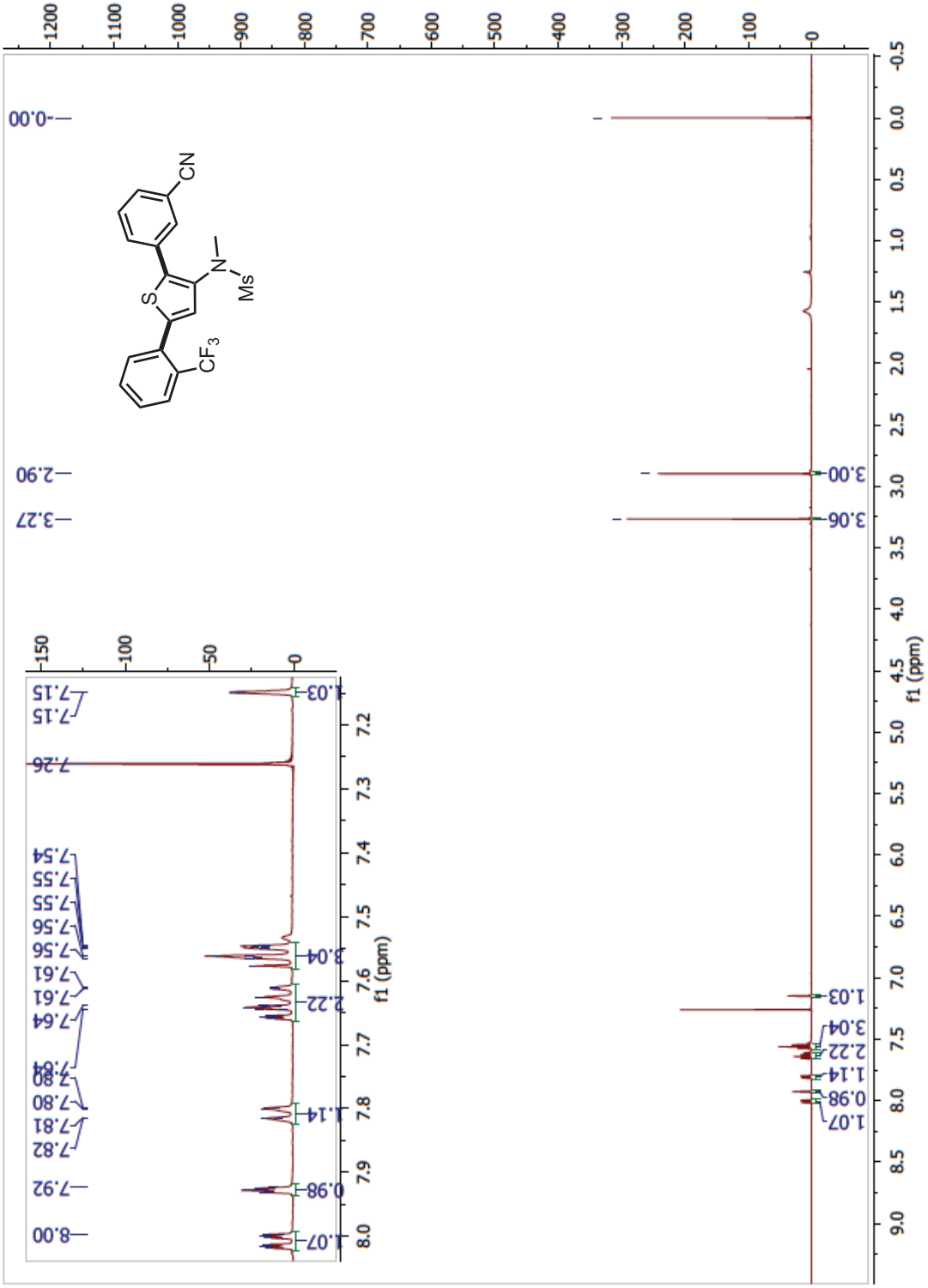
***N*-(2-(3-cyanophenyl)-5-(2-formylphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99q; Table 6.3, Entry 23) ^{1}H:**



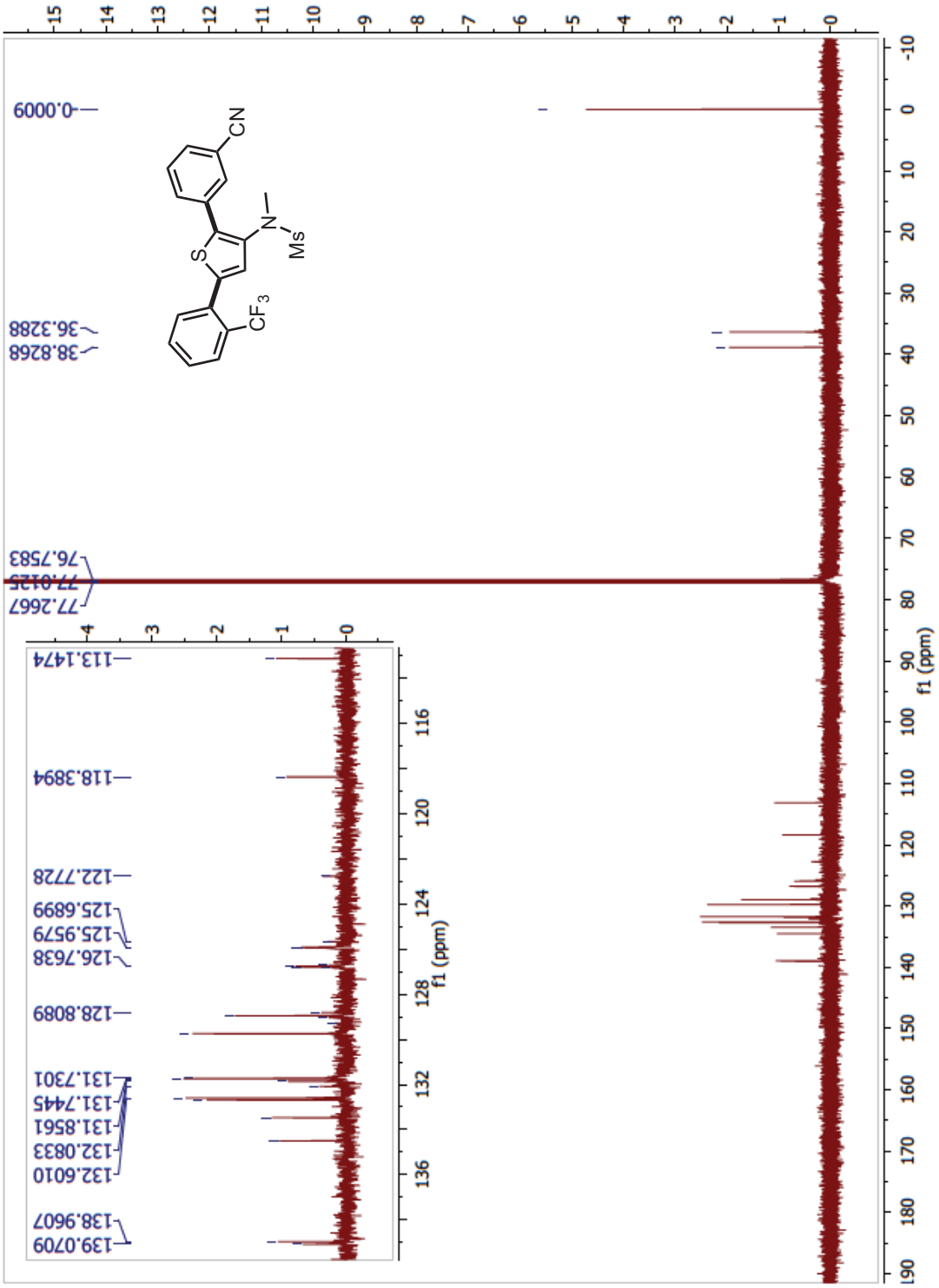
***N*-(2-(3-cyanophenyl)-5-(2-formylphenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99q; Table 6.3, Entry 23) ^{13}C:**



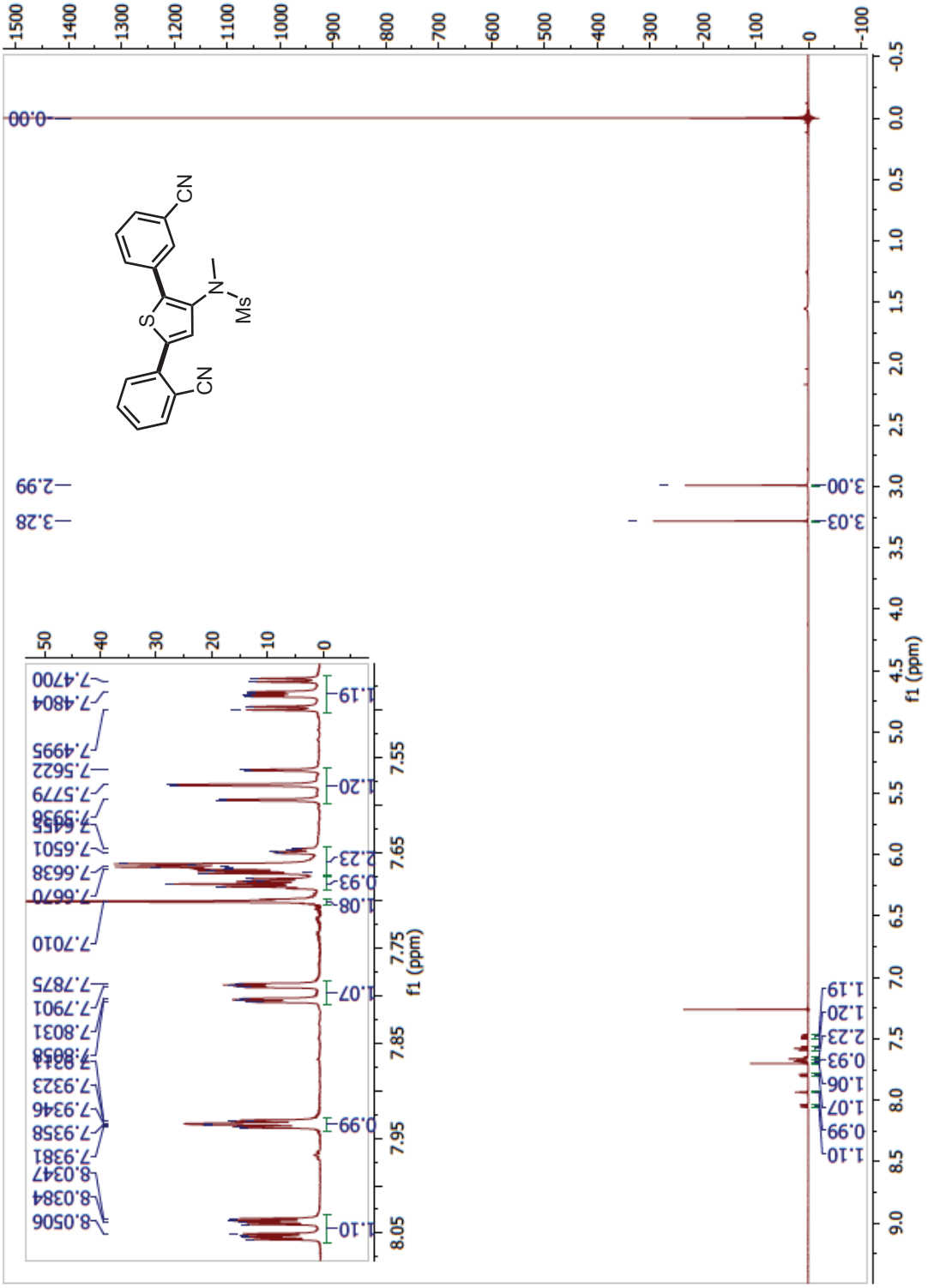
***N*-(2-(3-cyanophenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99r; Table 6.3, Entry 24) ^{1}H:**



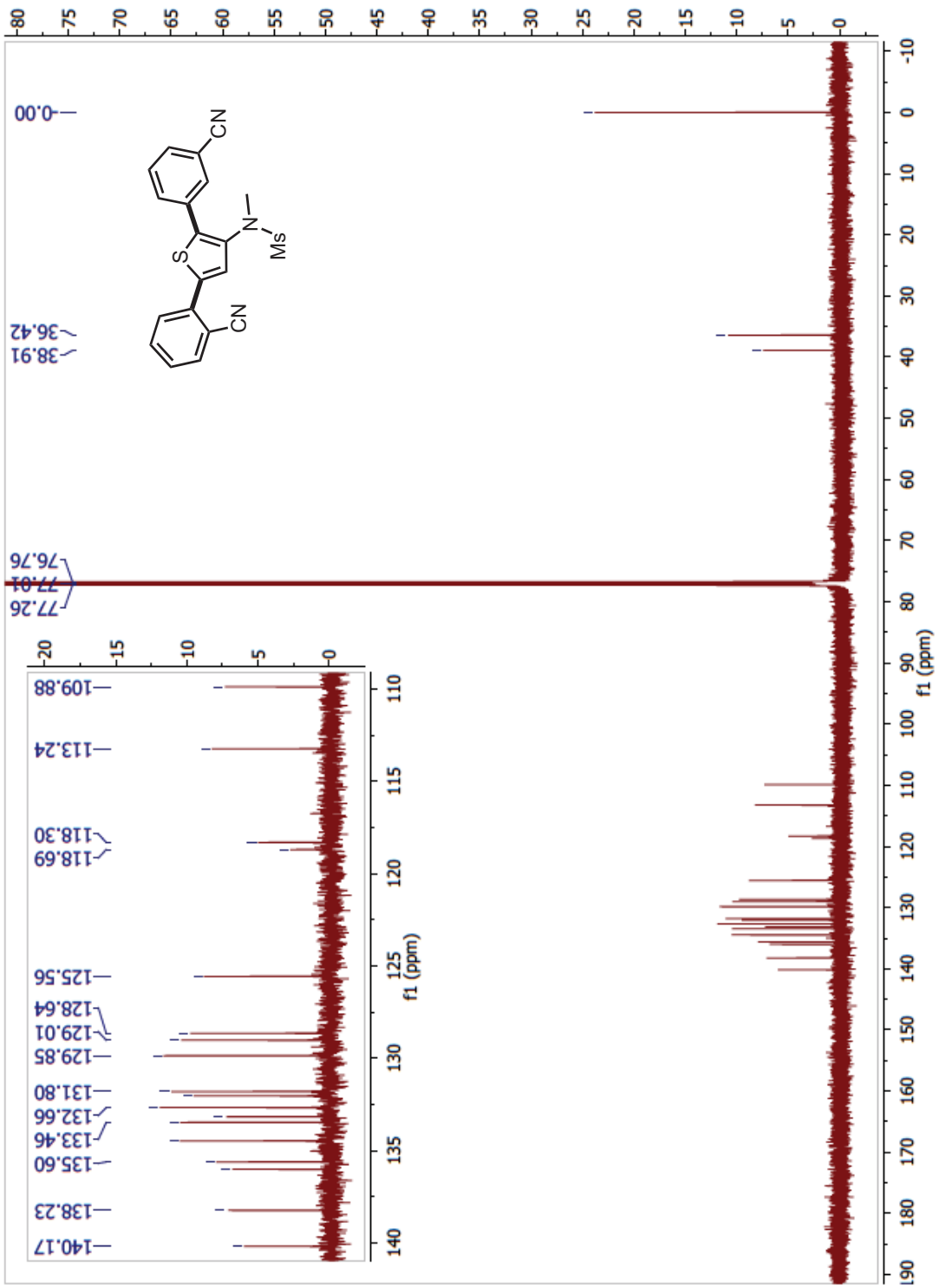
***N*-(2-(3-cyanophenyl)-5-(2-(trifluoromethyl)phenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99r; Table 6.3, Entry 24) ^{13}C:**



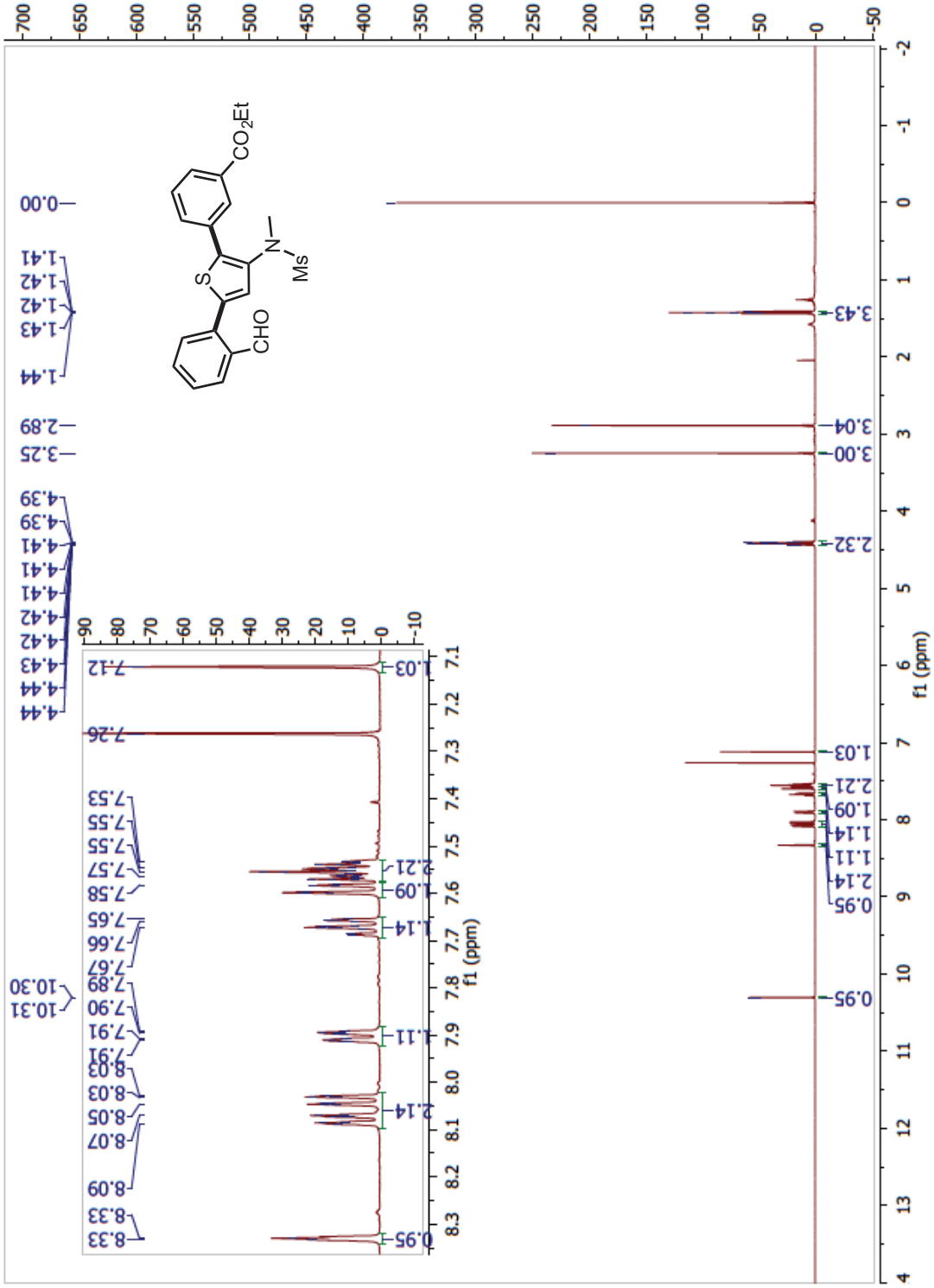
***N*-(5-(2-cyanophenyl)-2-(3-cyanophenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99a; Table 6.3, Entry 25) ^{1}H:**



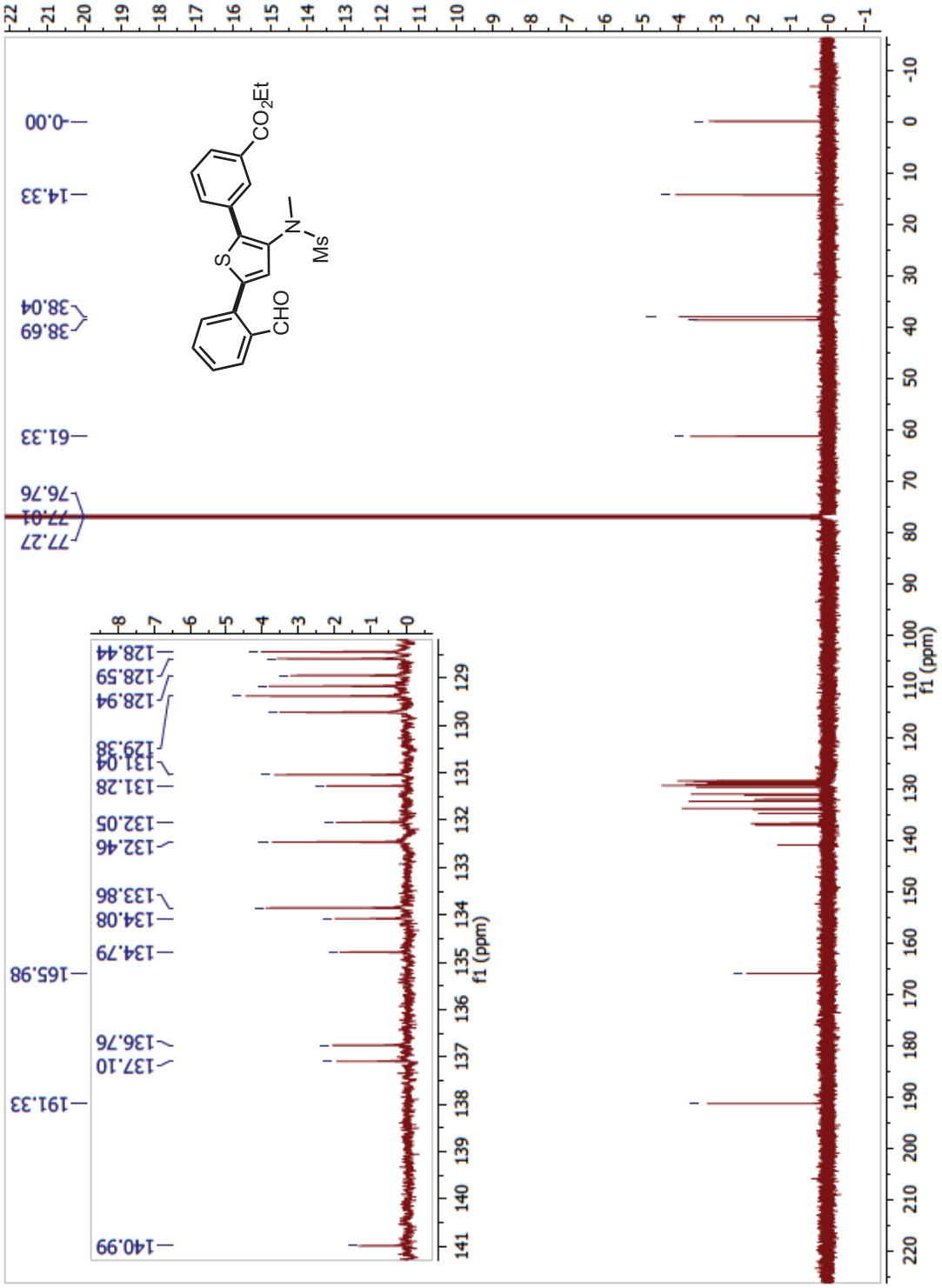
***N*-(5-(2-cyanophenyl)-2-(3-cyanophenyl)thiophen-3-yl)-*N*-methylmethanesulfonamide (99a; Table 6.3, Entry 25) ^{13}C:**



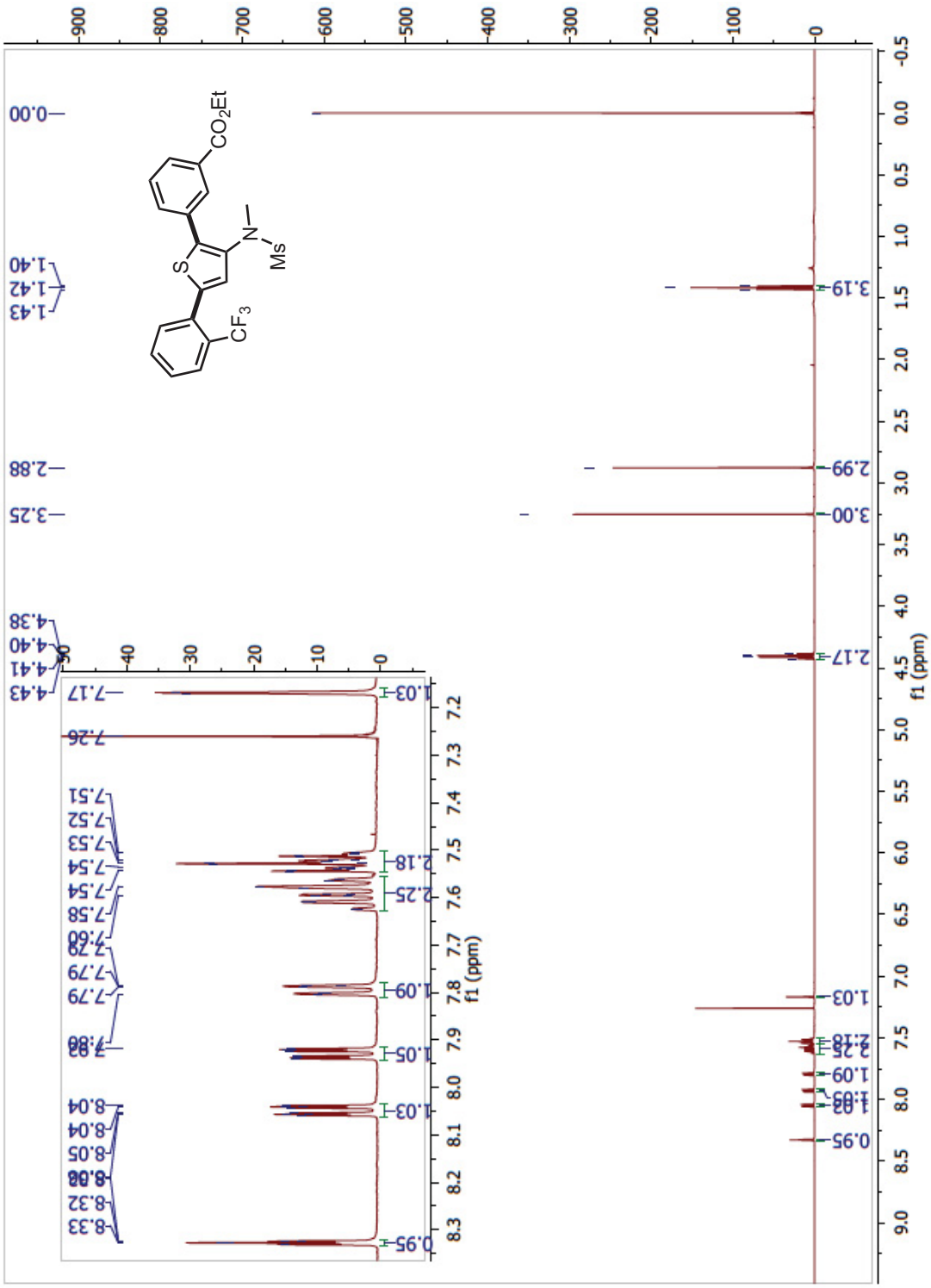
Ethyl 3-(5-(2-formylphenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99s; Table 6.3, Entry 26) ^{1}H:



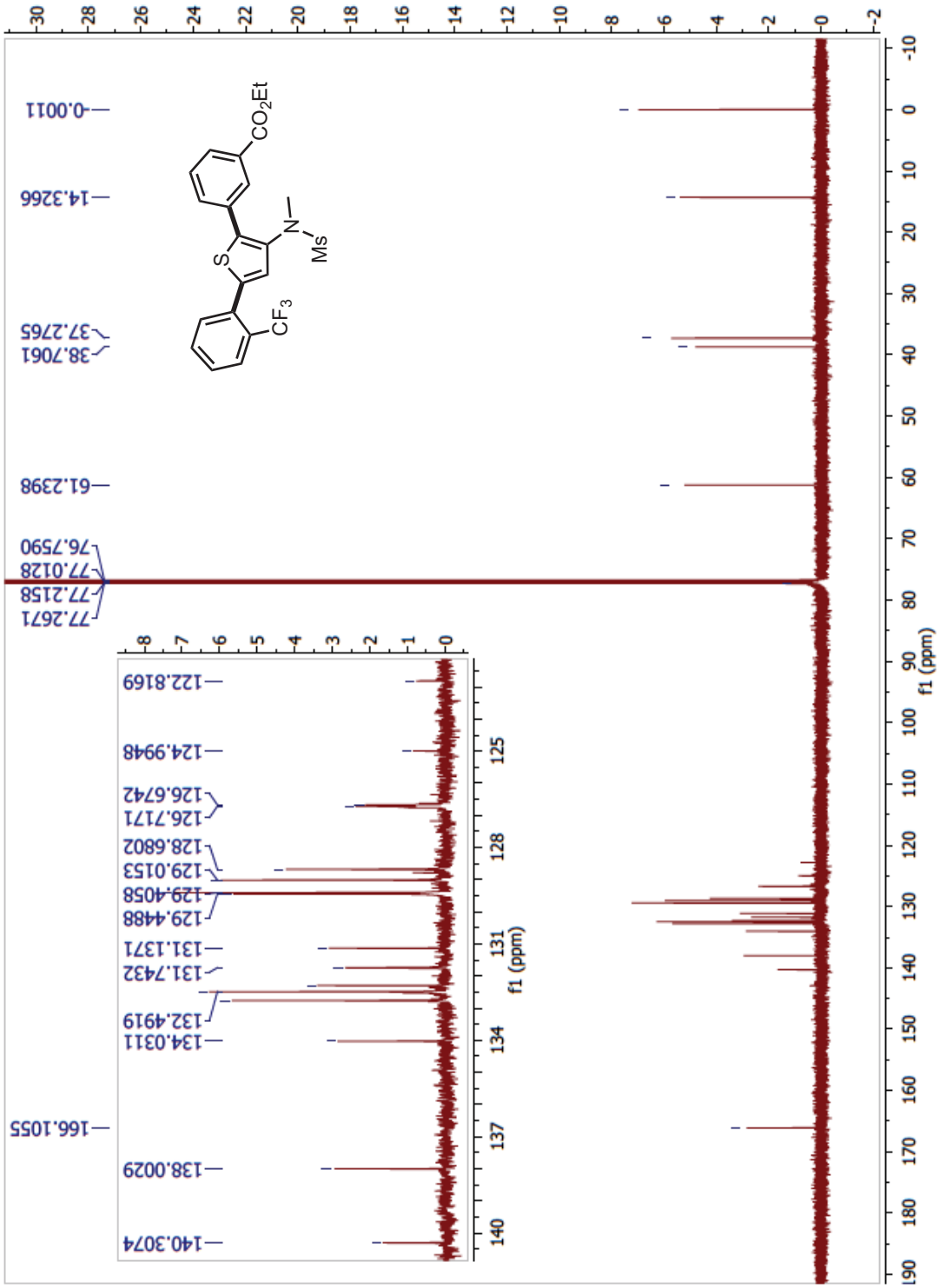
Ethyl 3-(5-(2-formylphenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99s; Table 6.3, Entry 26) ^{13}C:



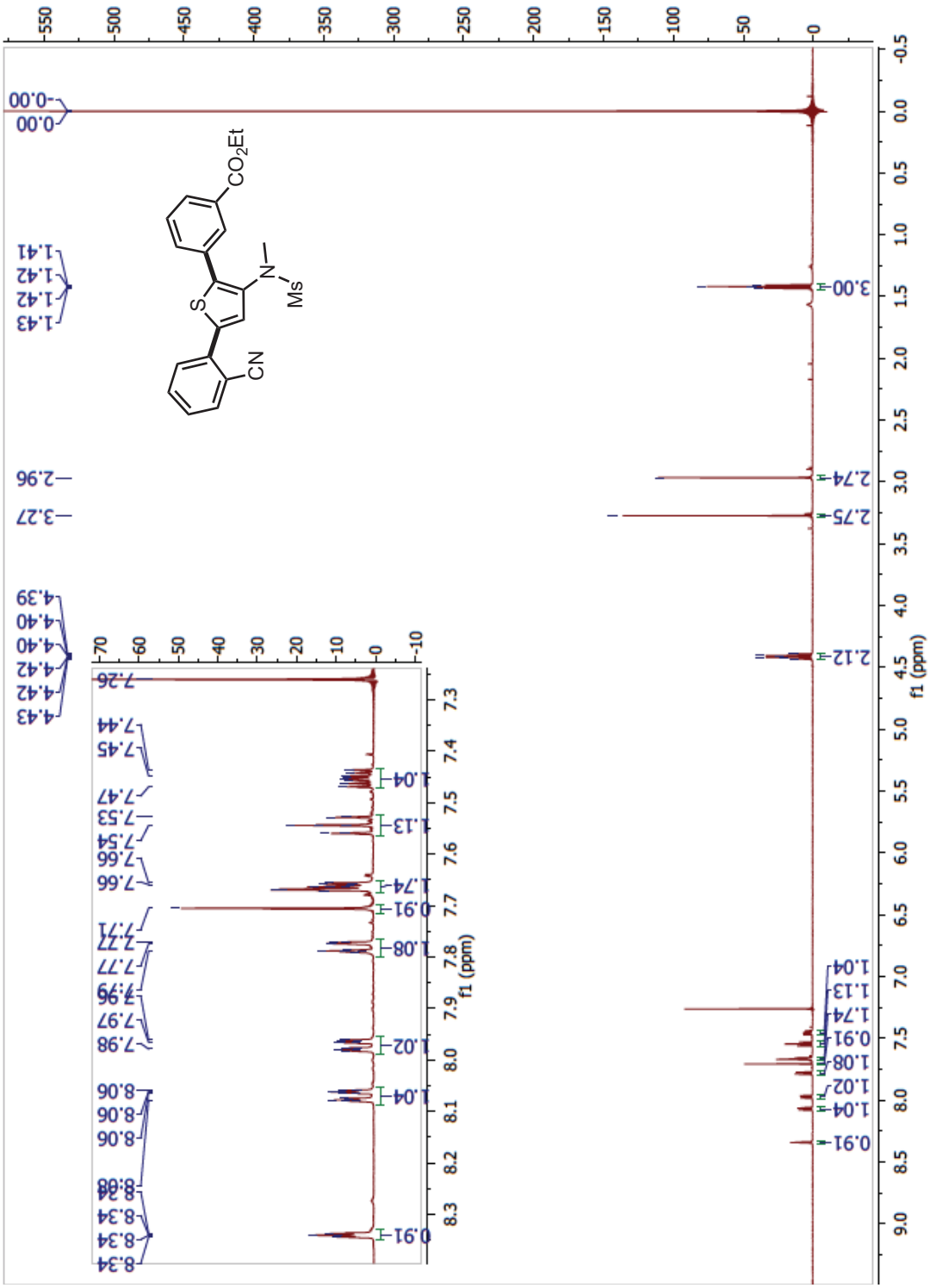
Ethyl 3-(3-(*N*-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99t; Table 6.3, Entry 27) ^{1}H:



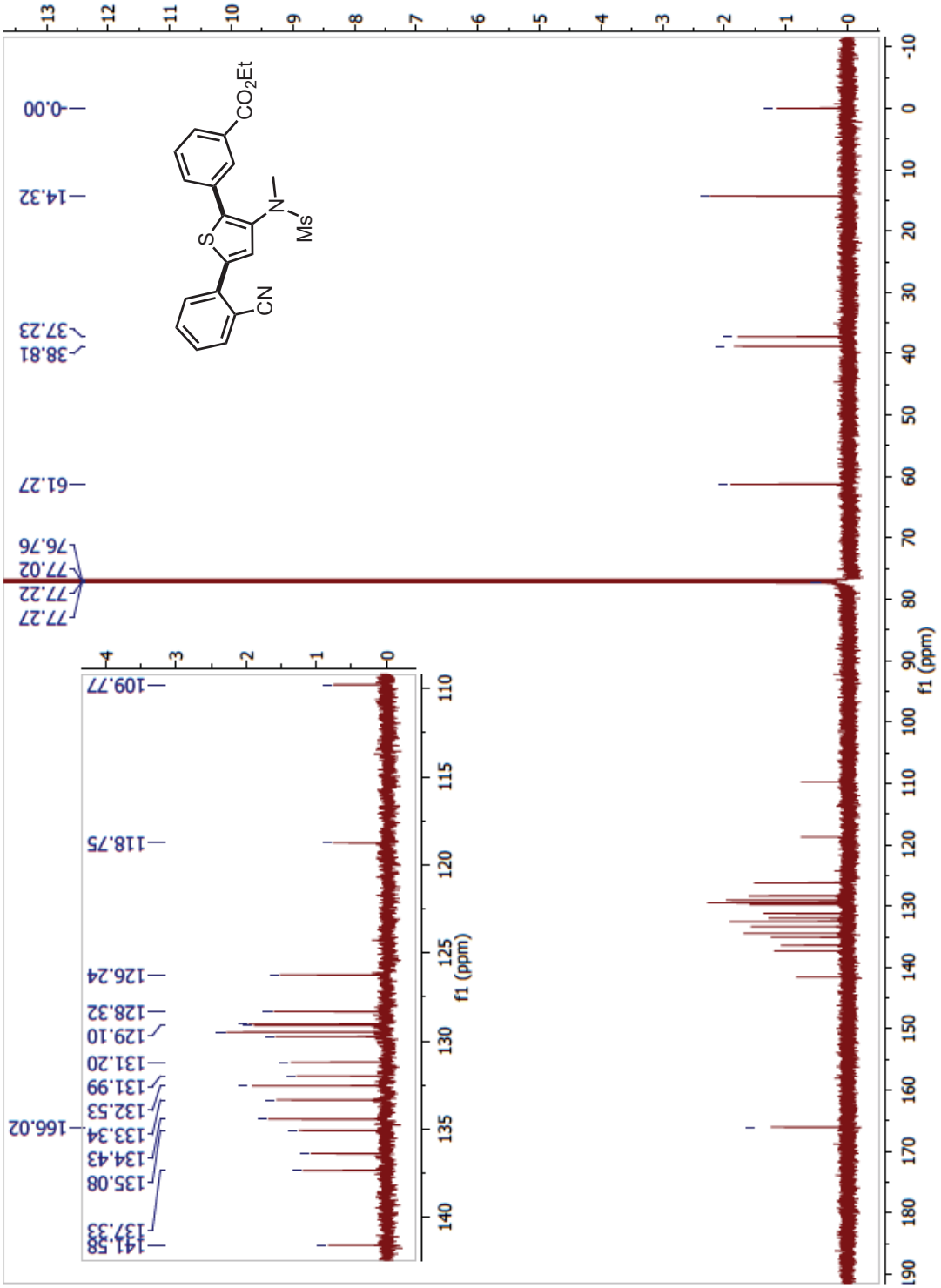
Ethyl 3-(3-(*N*-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoate (99t; Table 6.3, Entry 27) ^{13}C:



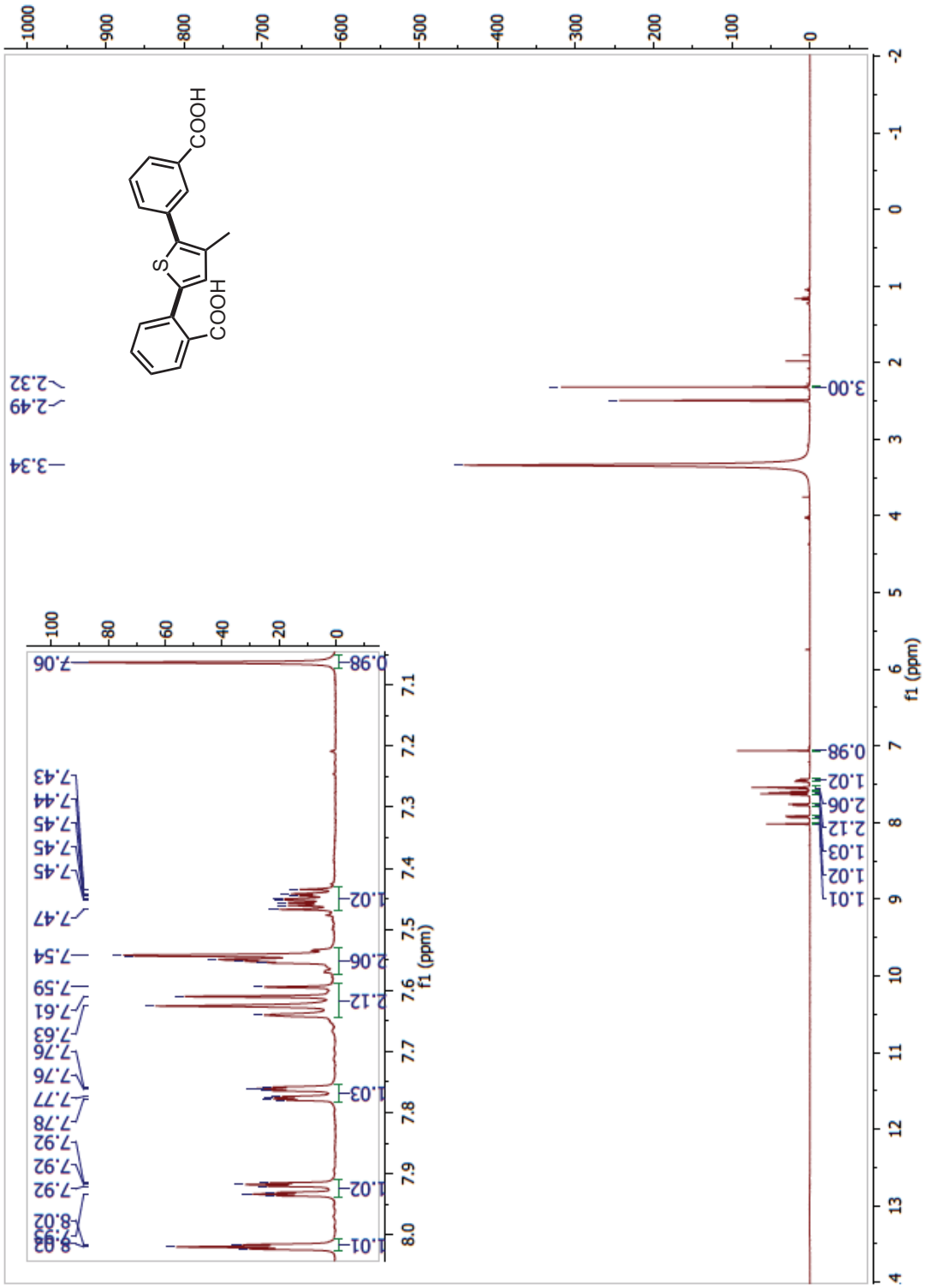
Ethyl 3-(5-(2-cyanophenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99u; Table 6.3, Entry 28) ^{1}H:



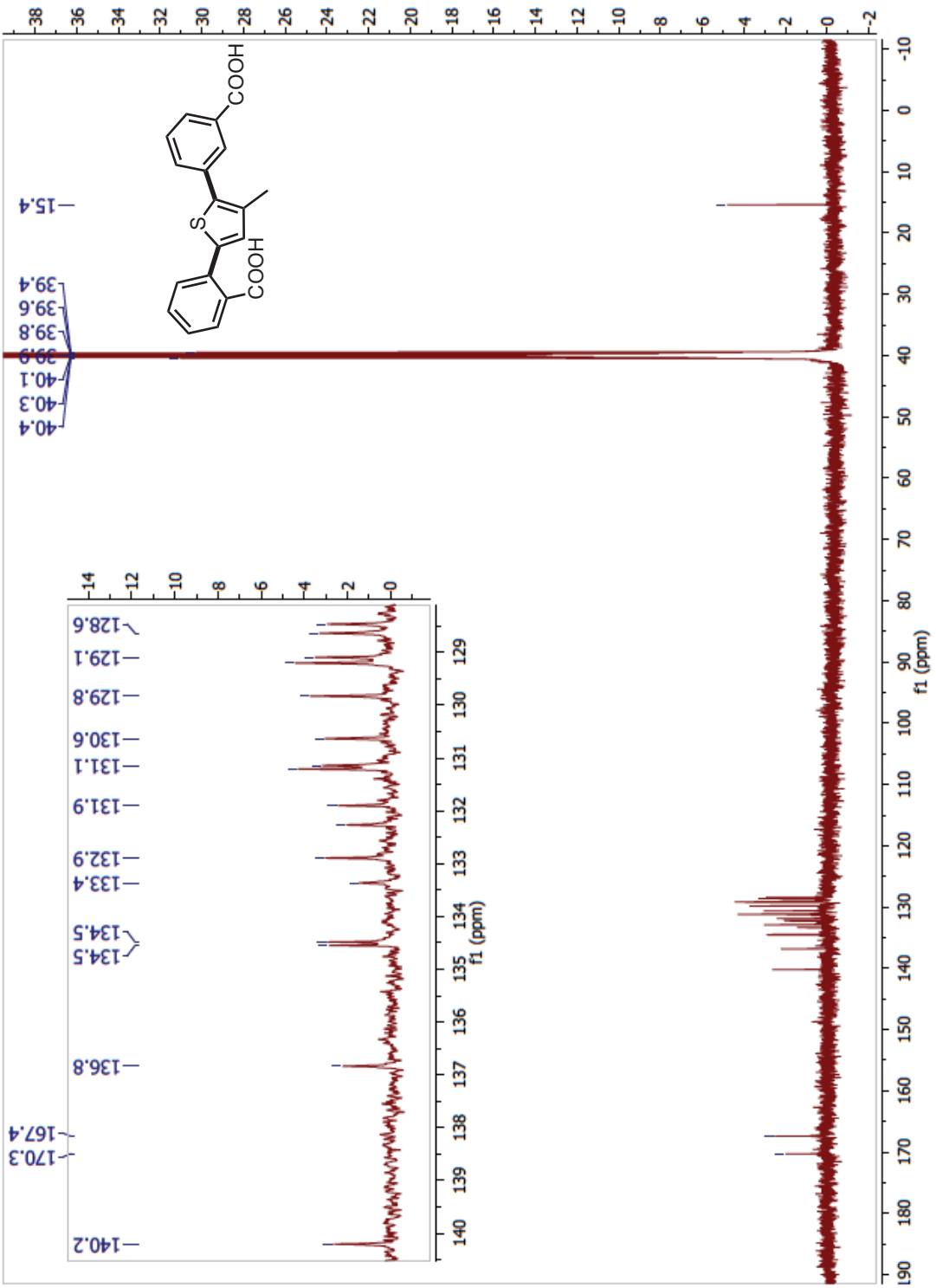
Ethyl 3-(5-(2-cyanophenyl)-3-(*N*-methylmethylsulfonamido)thiophen-2-yl)benzoate (99u; Table 6.3, Entry 28) ^{13}C:



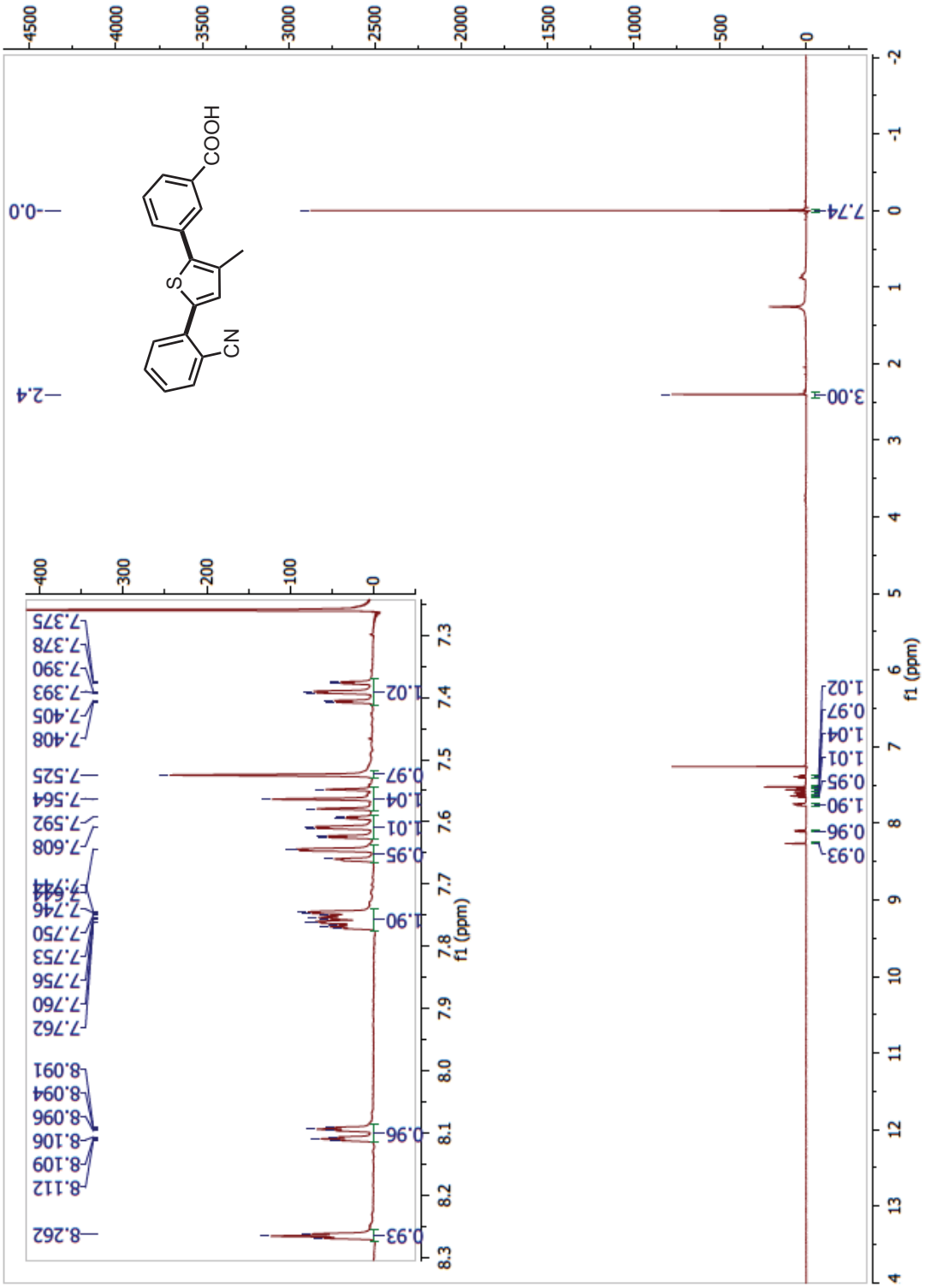
2-(5-(3-carboxyphenyl)-4-methylthiophen-2-yl)benzoic acid (103a) ^{1}H:



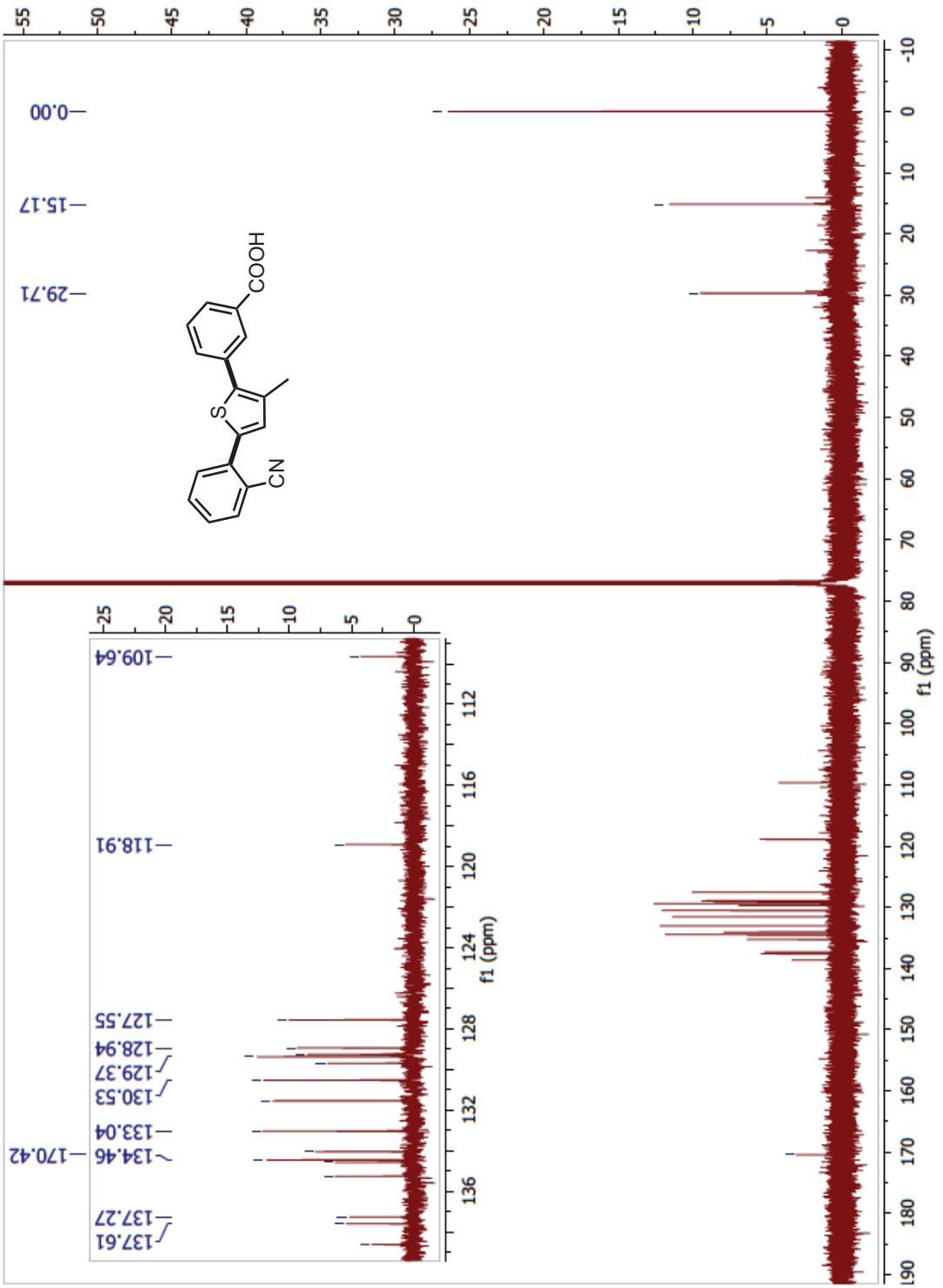
2-(5-(3-carboxyphenyl)-4-methylthiophen-2-yl)benzoic acid (103a) ^{13}C:



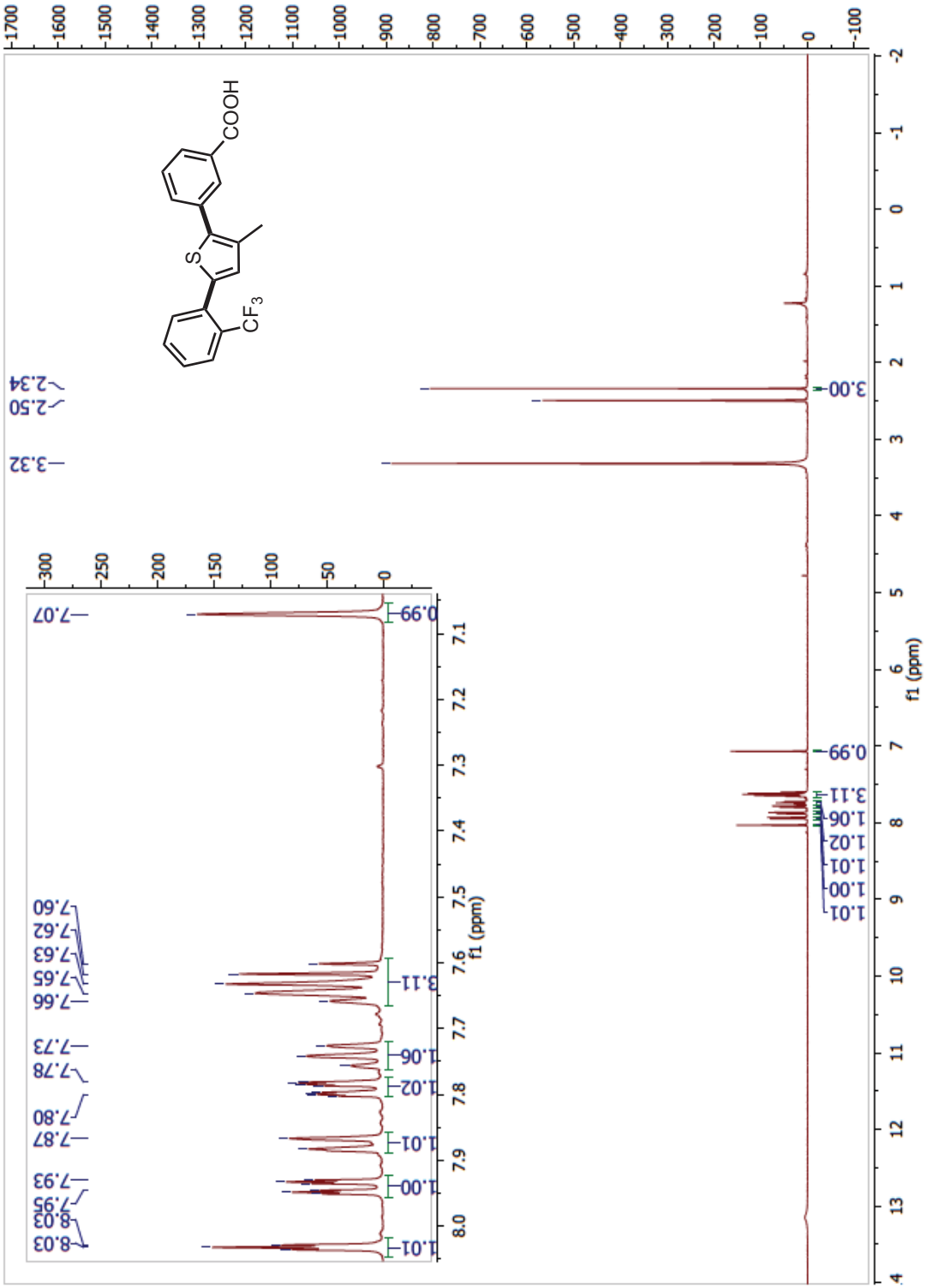
3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoic acid (103b; Table 6.4, Entry 31)
^{1}H:



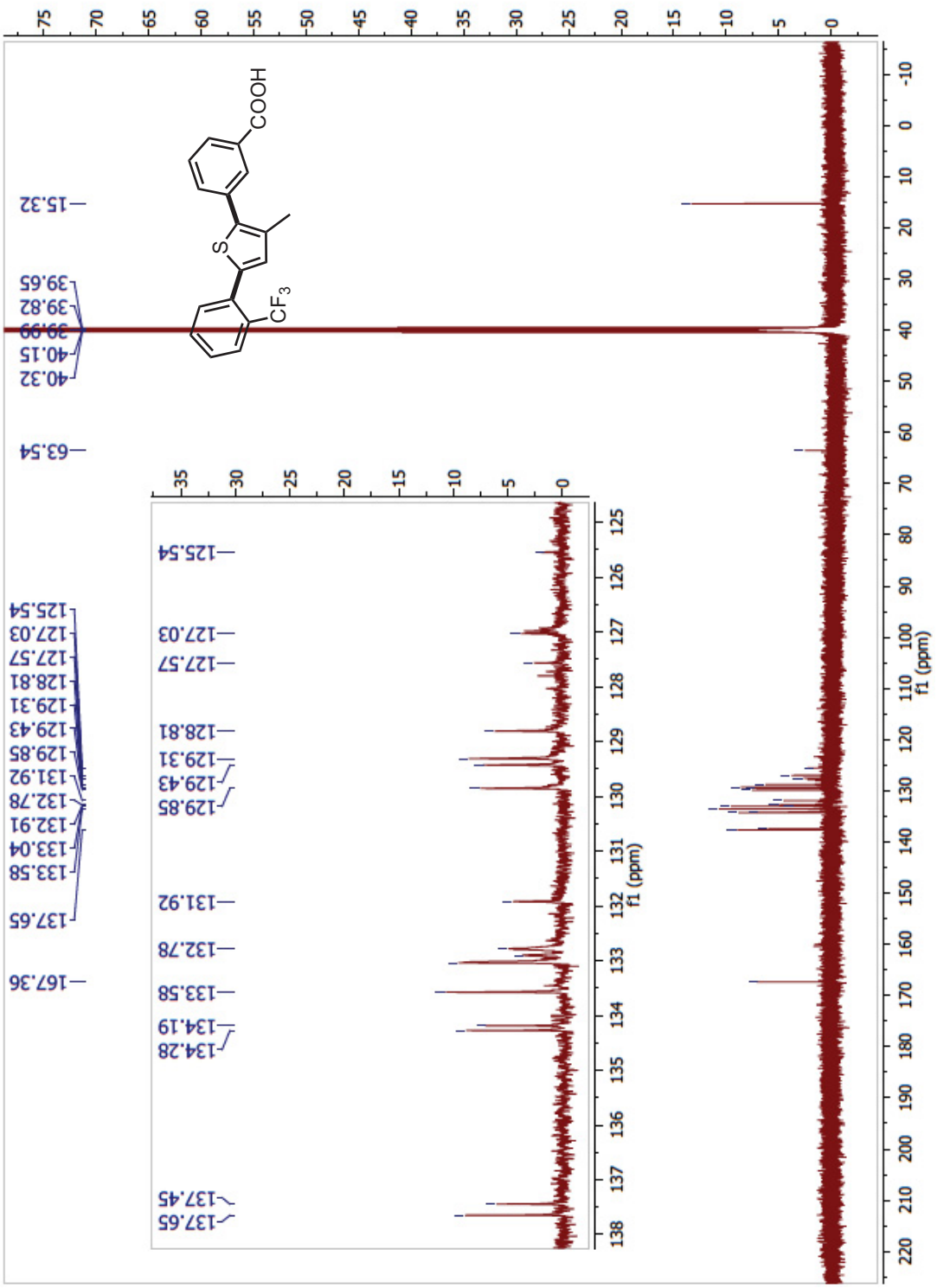
3-(5-(2-cyanophenyl)-3-methylthiophen-2-yl)benzoic acid (103b; Table 6.4, Entry 31)
^{13}C:



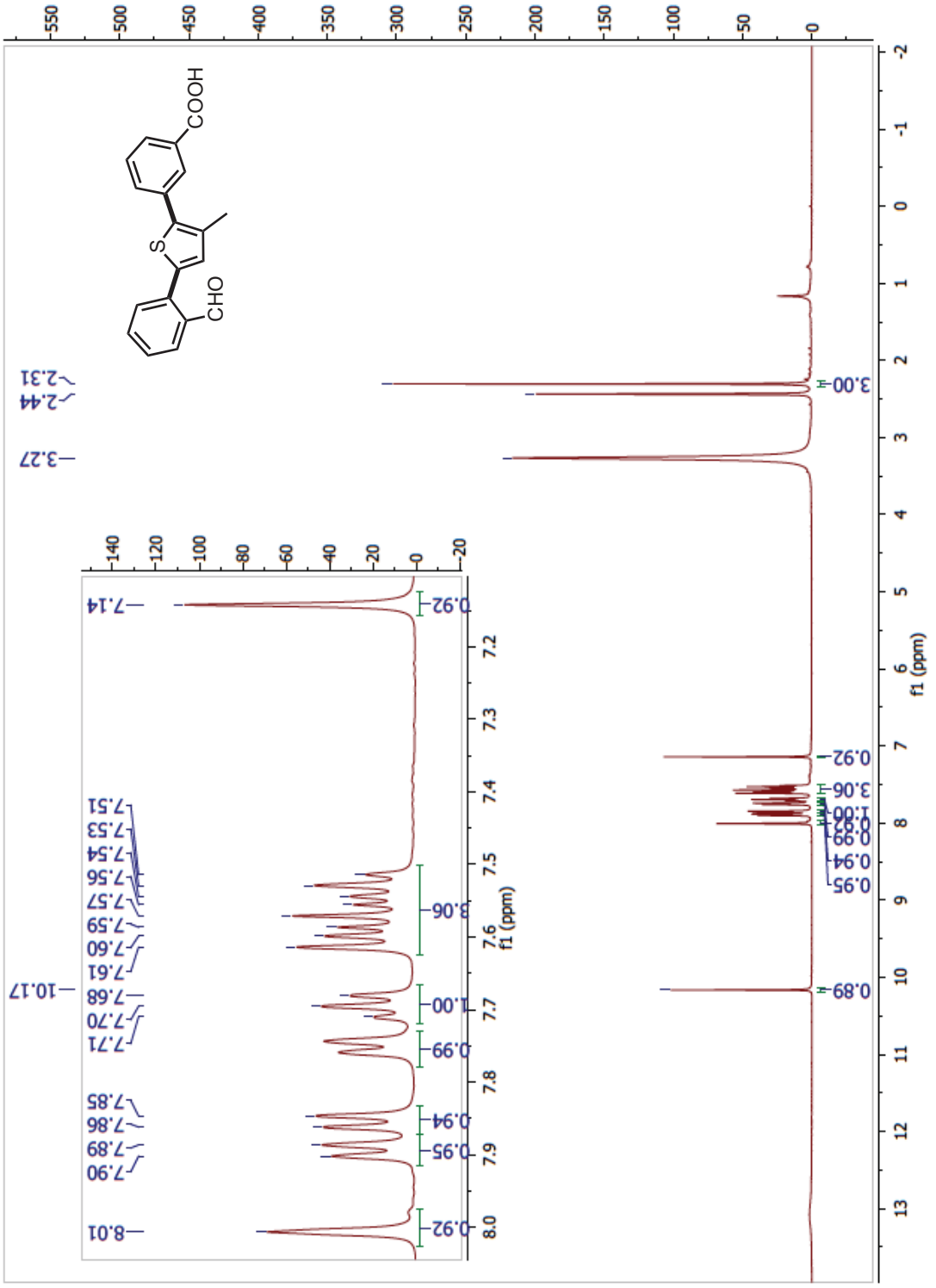
3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103c; Table 6.4, Entry 32) ^{1}H:



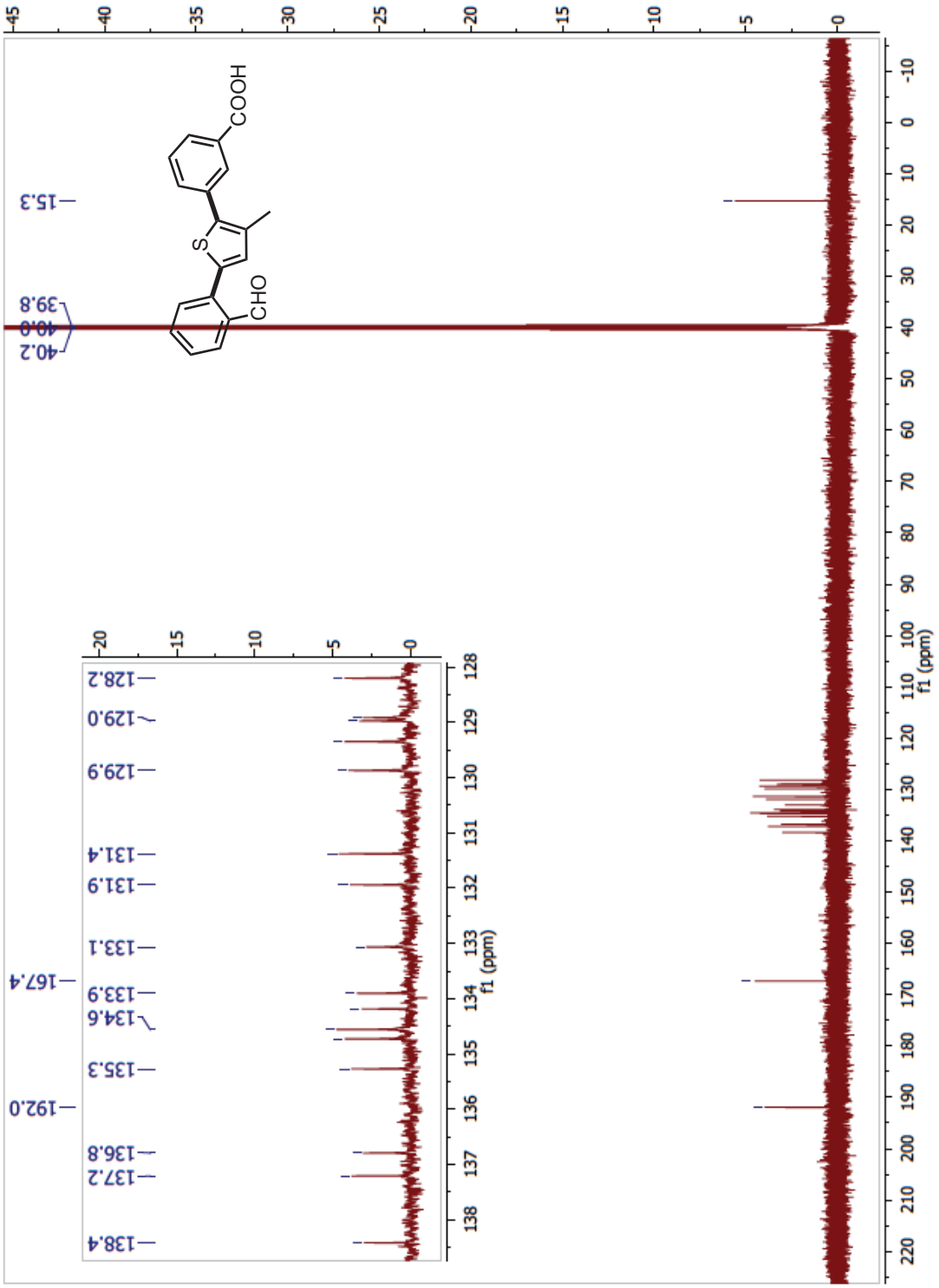
3-(3-methyl-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103c; Table 6.4, Entry 32) ^{13}C:



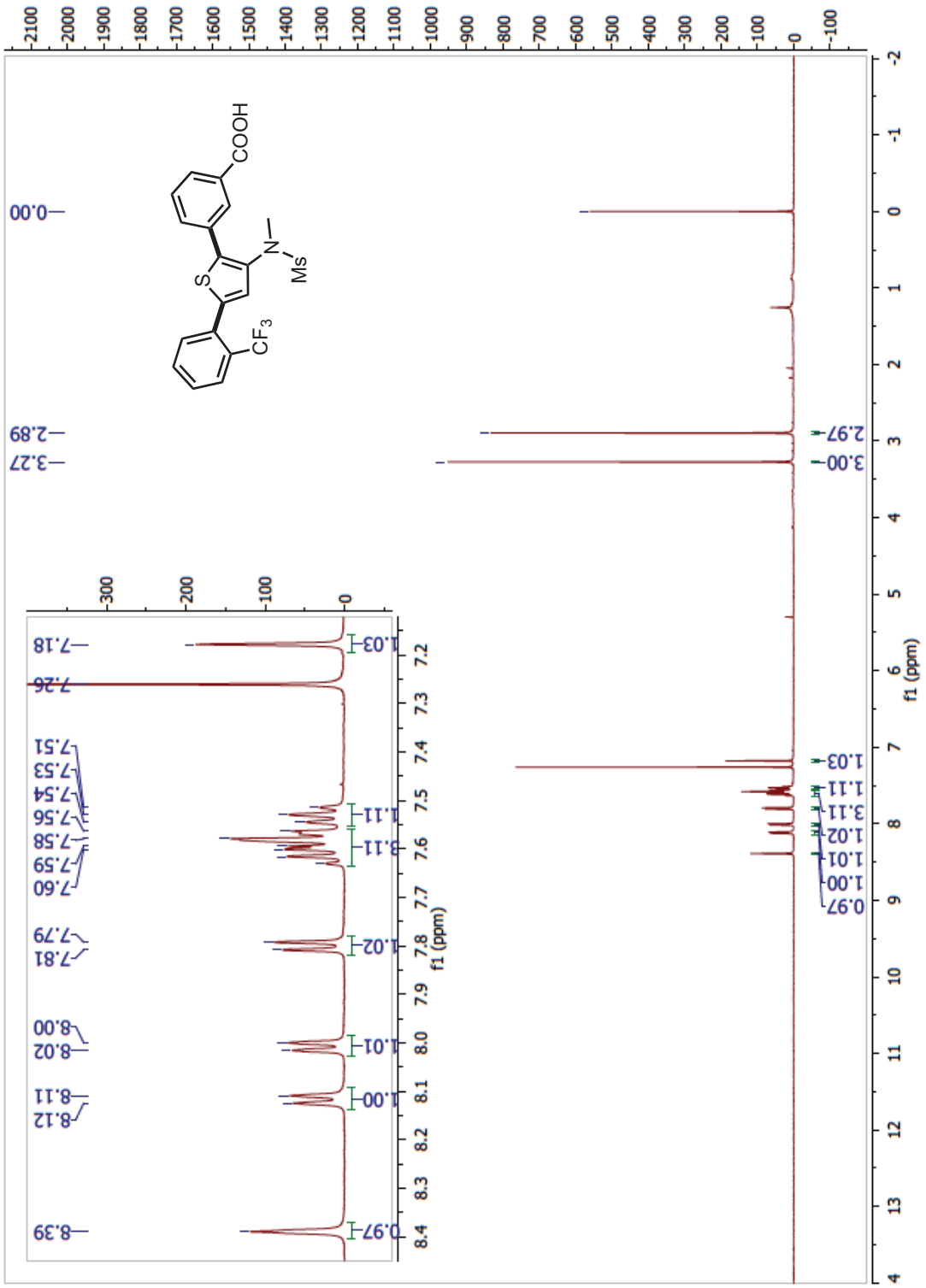
3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoic acid (103d; Table 6.4, Entry 33) ^{1}H:



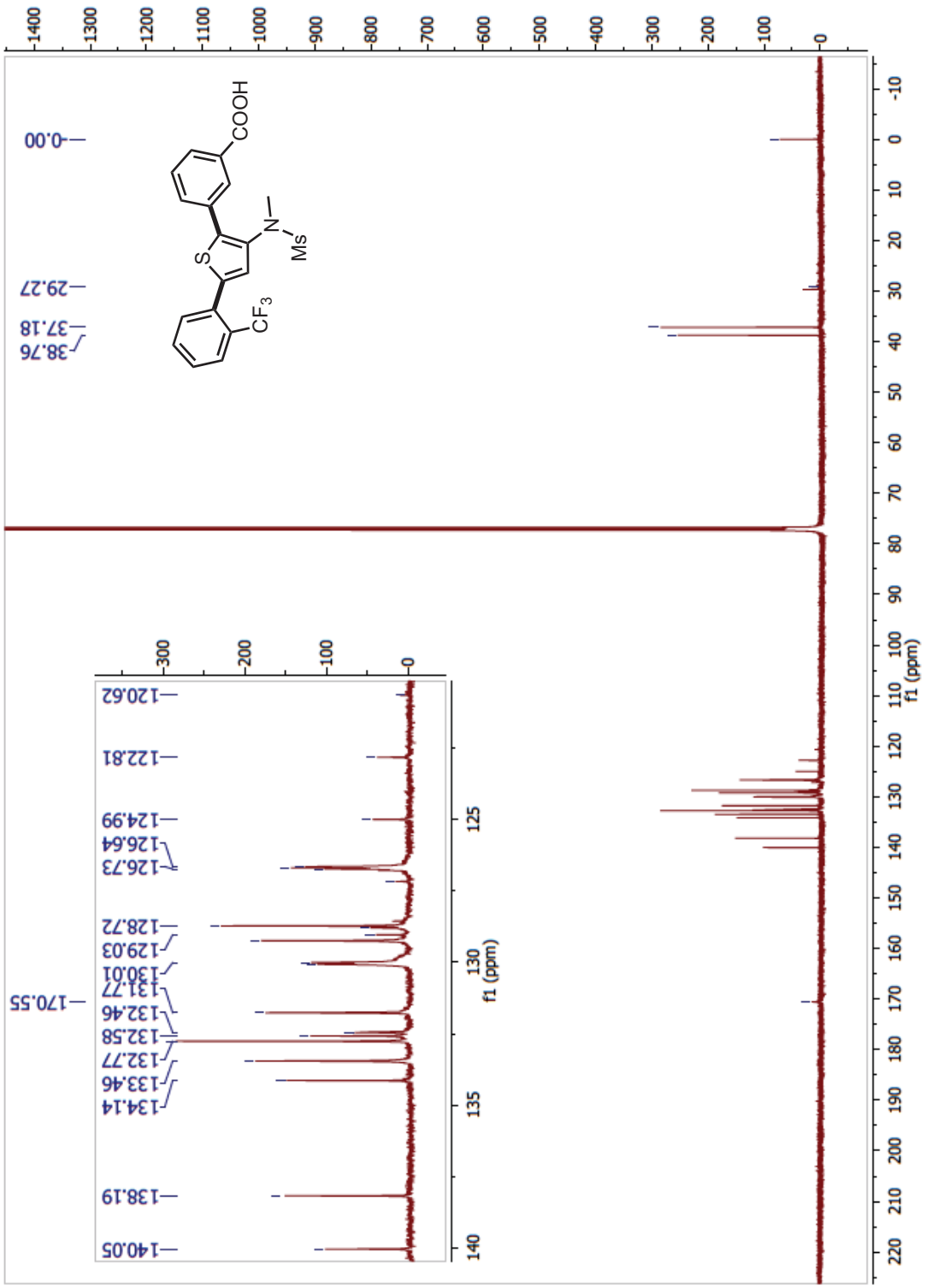
3-(5-(2-formylphenyl)-3-methylthiophen-2-yl)benzoic acid (103d; Table 6.4, Entry 33) ^{13}C:



3-(3-(N-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103e; Table 6.4, Entry 34) ^{1}H:



3-(3-(N-methylmethylsulfonamido)-5-(2-(trifluoromethyl)phenyl)thiophen-2-yl)benzoic acid (103e; Table 6.4, Entry 34) ^{13}C:



Materials and Methods

IAPP synthesis, purification and characterization

IAPP was synthesised by solid phase peptide synthesis on a Rink amide polystyrene resin based on Fmoc chemistry and 2-(6-chloro-1-H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium haxafluorophosphate (HCTU) coupling as previously described.[5] Cleavage from the resin was achieved with a mixture of TFA/ethanedithiol/phenol/water (92/2.5/3/2.5). Crude IAPP was purified by preparative scale reversed-phase high performance liquid chromatography (RP-HPLC) using a linear gradient of ACN in H_2O/TFA (0.06% v/v). Collected fractions were analyzed by analytical RP-HPLC using a C18 (3.6 μm) column (250 mm x 4.6 mm) and a linear gradient of ACN in H_2O/TFA (0.06% v/v). Fractions were also analysed by LC/MS-TOF to confirm the identity of the peptide. Disulfide bond formation between Cys-2 and Cys-7 was achieved by dimethyl sulfoxide (DMSO) oxidation according to the method developed by Abedini.[6] IAPP was re-purified by RP-HPLC as described above. Fractions corresponding to the desired product, as revealed by MS-TOF analysis, and with purity higher than 95%, confirmed by analytical RP-HPLC, were finally pooled and lyophilized.

Amyloid formation measured by thioflavin T fluorescence

Aliquots of monomerized IAPP were prepared by dissolving the lyophilized peptide in 100% hexafluoro-2-propanol (HFIP), and the solution was sonicated for 30 minutes and filtered through a 0.22 μm hydrophilic polypropylene filter before lyophilisation. Lyophilized IAPP was resolubilized in HFIP, sonicated for 30 minutes and the solution was aliquoted and lyophilized again to remove HFIP. Samples were kept dried at -80 °C until used. IAPP solutions were prepared by dissolving IAPP at a concentration of 25 μM (2x of final concentration) in 20 mM Tris, pH 7, 4, 40 μM ThT, immediately before final dilution and measurement. Substituted thiophene solutions were prepared at 100x (1.25 mM; final concentration of 12.5 μM, unless otherwise specified) in DMSO before being incorporated in the assay mixture. Assays were performed at room temperature (RT) without stirring in sealed black-wall, clear-bottom 96-well non-binding surface plates

with a volume of 100 μL per well. ThT fluorescence was measured from the bottom of the well every 10 min over the course of 25 h with excitation at 440 nm and emission at 485 nm. Data obtained from triplicate runs were averaged and corrected by subtracting the corresponding control reaction.

Cell toxicity assay

Rat INS-1 (β-pancreatic cell line) cells were seeded in black wall clear bottom 96-well plates at a density of 35 000 cells/well (100 μl/well) in RPMI-1640 complete medium.[5] After 24 h incubation at 37 °C in a 5% CO_2 incubator, cells were treated by directly adding 50 μl of IAPP solutions at 3x final concentrations in 20 mM Tris, pH 7,4 that had been pre-incubated for 20 h at room temperature in presence or in absence of 1 equivalent of substituted thiophene derivatives. These solutions were then incubated for an additional 24h, and cell viability was measured by the resazurin reduction assay. Control conditions were performed in presence of 50 μM of compounds. Cell viability (in %) was calculated from the ratio of the fluorescence of the treated sample to the control cells (non-treated).

Figures

Figure S.1: Effects of 2,5-diaryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (◆, blue) or in the presence of 12.5 μM of compound (■, red). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

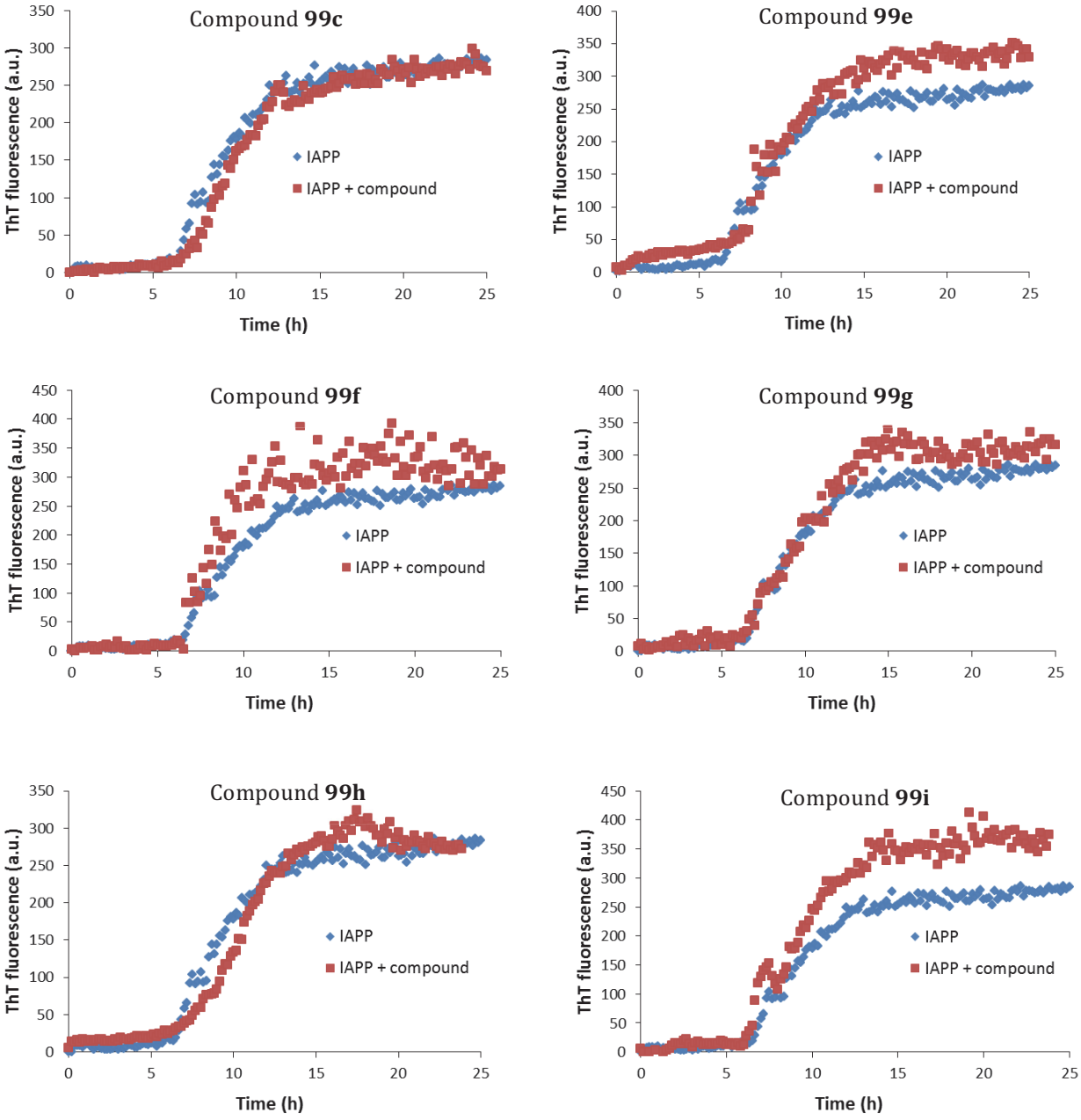


Figure S.2: Effects of 2,5-diaryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (♦, blue) or in the presence of 12.5 μM of compound (■, red). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

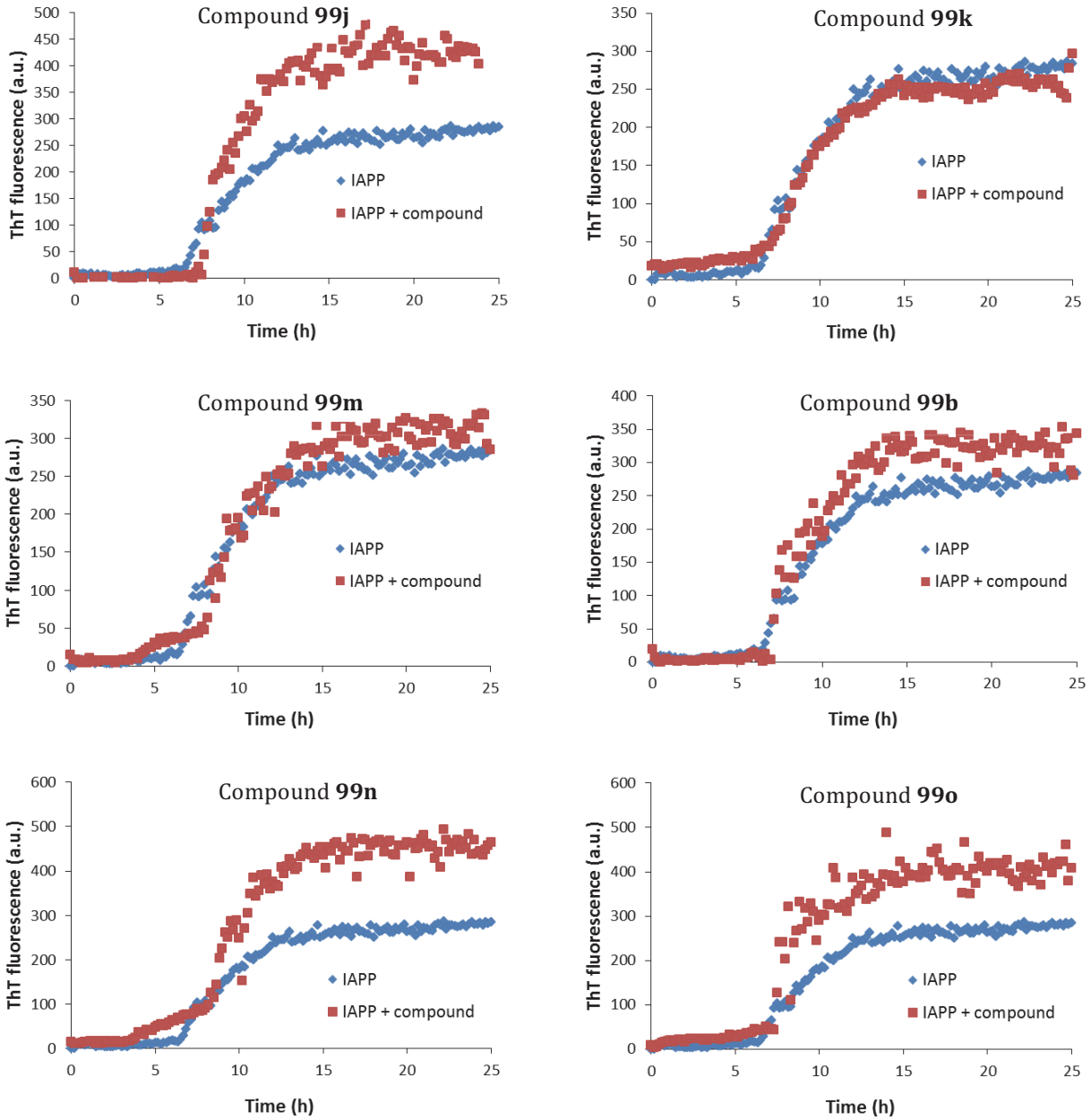


Figure S.3: Effects of 2,5-diaryl substituted thiophenes on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (♦, blue) or in the presence of 12.5 μM of compound (■, red). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.

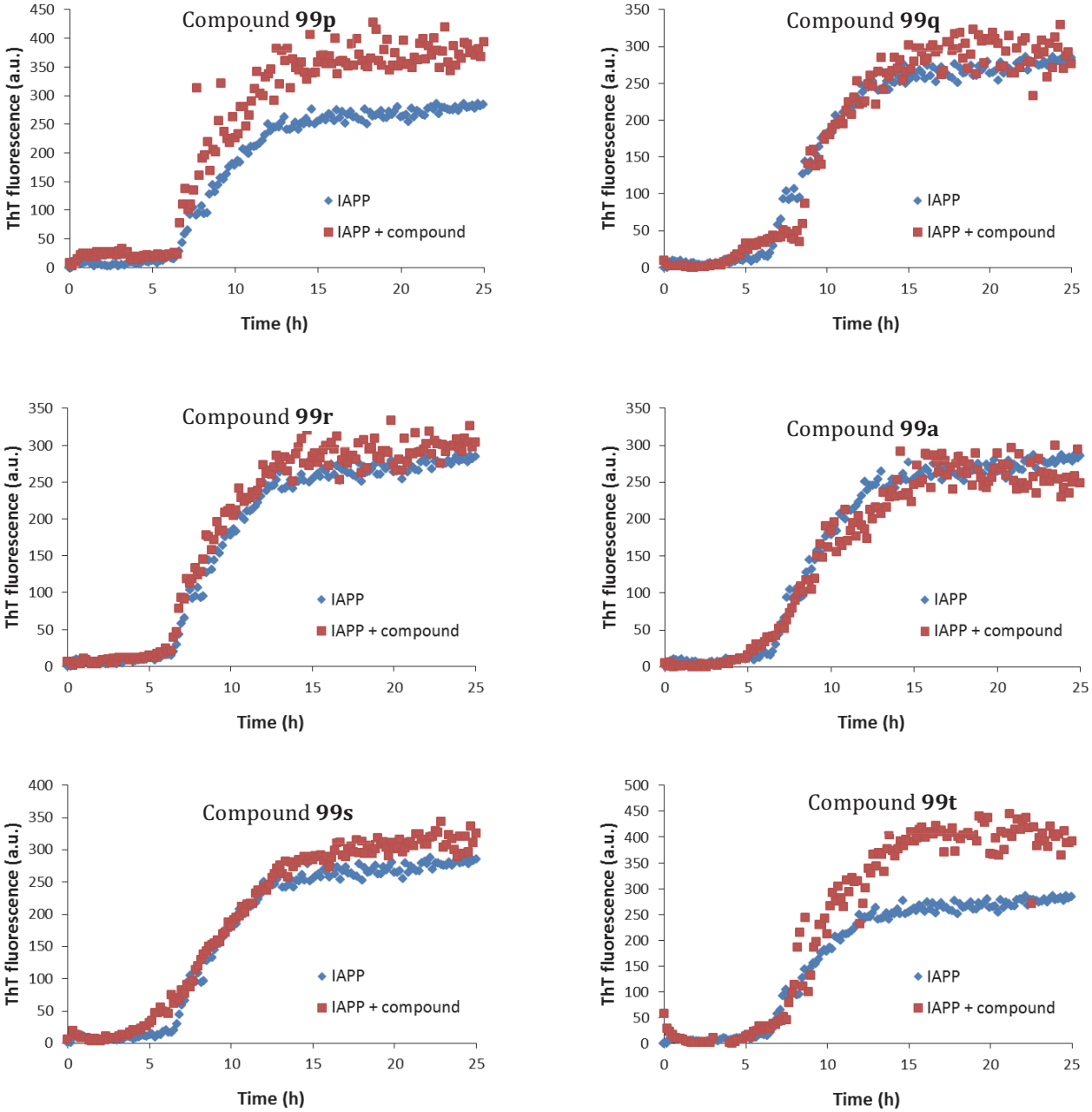
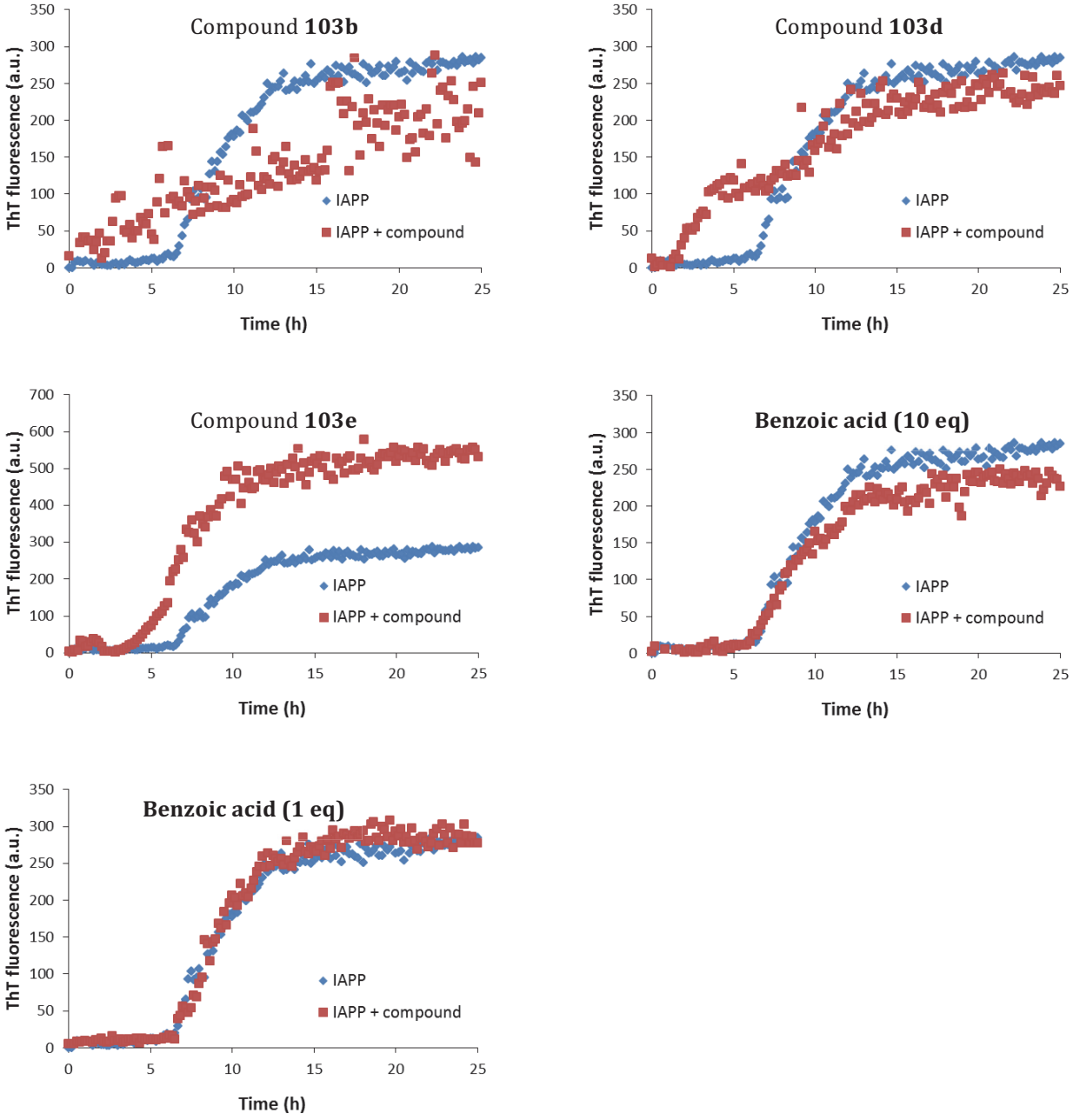


Figure S.4: Effects of 2,5-diaryl substituted thiophenes and benzoic acid on IAPP kinetics of amyloid fibril formation monitored by ThT fluorescence. IAPP (12.5 μM) was incubated in 20 mM Tris, pH 7.4, at 25 °C without agitation in the absence (♦, blue) or in the presence of 12.5 μM of compound (■, red). 12.5 μM (1 equiv) and 125 μM of benzoic acid was used (10 equiv). ThT fluorescence (40 μM) was measured every 10 min over the course of 25 h, with excitation at 440 nm and emission at 485 nm.



References

[1] D. Tondi, F. Morandi, R. Bonnet, M. P. Costi, B. K. Shoichet, *J. Am. Chem. Soc.* **2005**, *127*, 4632-4639.

[2] F. Cardullo, D. Donati, V. Fusillo, G. Merlo, A. Paio, M. Salaris, A. Solinas, M. Taddei, *J. Comb. Chem.* **2006**, *8*, 834-840.

[3] P. Forgione, M.-C. Brochu, M. St-Onge, K. H. Thesen, M. D. Bailey, F. Bilodeau, *J Am Chem Soc* **2006**, *128*, 11350-11351.

[4] B. t. Liégault, D. Lapointe, L. Caron, A. Vlassova, K. Fagnou, *J. Org. Chem.* **2009**, *74*, 1826-1834.

[5] C. A. De Carufel, P. T. Nguyen, S. Sahnouni, S. Bourgault, *Biopolymers* **2013**.

[6] A. Abedini, G. Singh, D. P. Raleigh, *Anal. Biochem.* **2006**, *351*, 181-186.