Isomerization-Locked Alkene Analogues of Xaa–Pro Dipeptides in the Proteins Collagen and Bora

Paul Joseph Arcoria

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Chemistry

Felicia A. Etzkorn
Emily E. Mevers
Paul R. Carlier
Paul A. Deck

June 28, 2022
Blacksburg, Virginia

Keywords: collagen, triple helix, polyproline type II helix, peptide, fluoro-alkene, chloro-alkene, proteo-alkene, conformationally-locked isostere, stability, n→π*, Ser–*cis*-Pro, Ser–*trans*-Pro, aminoacylation, flexizyme, cell cycle, mitosis

Isomerization-Locked Alkene Analogues of Xaa–Pro Dipeptides in the Proteins Collagen and Bora

Paul Joseph Arcoria

Abstract

Collagen is one of the most abundant human proteins. It exists as a right-handed superhelix called the triple helix. The triple helix consists of three left-handed polyproline type II (PPII helices) that intertwine around a common axis. Each PPII helix has the repeating peptide sequence $(Gly–Xaa–Yaa)_n$ with a high content of (2*S*)-proline (Pro) in the Xaa position (ca. 28%) and (2*S*,4*R*)-hydroxyproline (Hyp) in the Yaa position (ca. 38%). Unique to the prolyl amide is the ease of cis-trans isomerization. Since the triple helix necessitates that all peptide bonds be in the trans conformation, isomerization is the rate-limiting step in collagen folding. However, eliminating isomerization with a trans-locked alkene isostere destabilizes collagen-like peptides. Collagen is stabilized by electronic interactions, namely the n→π* interaction. Halo-alkene isosteres may be used to recapture these electronic interactions and stabilize a collagen-like peptide.

An in-depth conformational analysis was conducted at the MP2/6-311+G(2d,p) level of theory to determine the viability of conformationally-locked halo-alkene isosteres. Fluoro-alkenes and chloro-alkenes were modeled at both the Gly–Pro and Pro–Pro (as a Pro–Hyp mimic) amide positions. Compared to the collagen crystal structure PDB ID: 1K6F, we found the fluoro-alkenes were closer geometric matches to both Gly–Pro and Pro–Pro than the corresponding chloro-alkenes. The chloro-alkene was predicted to have stronger n→π* interactions. The trans-locked proteo-alkene was also analyzed to

understand why it destabilized the triple helix. We found that these models had other local minima close to the desired PPII geometry, likely leading to enhanced backbone flexibility. This deleterious flexibility was not predicted for either fluoro-alkene or chloro-alkene models.

The conformationally-locked halo-alkene isostere Fmoc–Gly–Ψ[(*Z*)CF=C]-Pro–Hyp(^{t}Bu)–OH was designed and synthesized as a (*Z*)-fluoro-alkene Gly–Pro isostere. We used the chiral catalyst, L-Thr, for asymmetric aldol addition to cyclopentanone, which inadvertently enhanced the yield of the wrong enantiomer, in contrast with aldol addition to cyclohexanone. A Mg^{2+}-promoted Horner-Wadsworth-Emmons reaction afforded the (*Z*)-fluoro-alkene over the (*E*)-fluoro-alkene in about a 2:1 ratio. The two diastereomers, Fmoc–Gly–Ψ[(*Z*)CF=C]-L-Pro–Hyp(^{t}Bu)–OH and Fmoc–Gly–Ψ[(*Z*)CF=C]-D-Pro–Hyp(^{t}Bu)–OH were separated by supercritical CO_2 chromatography.

The collagen-like peptides Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C]-L-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2, Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C]-D-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2, and the control peptide Ac–(Gly–Pro–Hyp)$_8$–Gly–Gly–Tyr–NH_2 were synthesized on solid-phase resin. The CD spectra of all three peptides showed the characteristic collagen triple-helix signature. The folding stability was determined by thermal melting (T_m). The peptide with the fluoro-alkene guest, Gly–Ψ[(*Z*)CF=C]-L-Pro–Hyp, was found to have a T_m value of 42.2 °C. The T_m of the control peptide was found to be 49.0 °C, a difference in stability of only ΔT_m –6.8. Thus, the (*Z*)-fluoro-alkene as a Gly–Pro isostere forms a relatively stable triple helix. The peptide with the Gly–Ψ[(*Z*)CF=C]-D-Pro–Hyp guest was shown to have a linear relationship between ellipticity and temperature, indicating that a stable triple helix did not form. The enhanced

stability of the (*Z*)-fluoro-alkene compared to the (*E*)-alkene Gly–Pro isostere (T_m = 28.3 °C) may be due to a stabilizing n→π* interaction, as determined by NMR deshielding of the ^{19}F nucleus in the collagen-like peptide.

In biological systems, isomerization of the prolyl amide is catalyzed by enzymes called PPIases. The PPIase Pin1 specifically catalyzes isomerization of the pSer–Pro sequence from the cis-conformation to the trans-conformation. Pin1 plays a crucial role in the G2→M transition of the cell cycle, implying the importance of *cis-trans* isomerization. The dipeptides H–Ser–Ψ[(*Z*)CH=C]-Pro–OH, H–Ser–Ψ[(*E*)CH=C]-Pro–OH and native H–Ser–Pro–OH were synthesized by literature methods, and activated for aminoacylation of $tRNA^{CUA}$ for *in vitro* transcription-translation. Aminoacylation by chemical methods required the synthesis of a pdCpA dinucleotide. Formation of the dipeptide-dinucleotide complex was not completed because protection of the Ser side chain was problematic. On the other hand, conversion of the dipeptide into the 3,5-dinitrobenzyl ester conjugate allowed for enzymatic aminoacylation using the dFx flexizyme, an RNA enzyme. The native dipeptide was successfully coupled to $tRNA^{CUA}$ and is ready for incorporation into a full-length Bora protein by *in vitro* transcription-translation. Both cis- and trans-locked alkene mimics have been converted to their respective 3,5-dinitrobenzyl ester conjugates.

Isomerization-Locked Alkene Analogues of Xaa–Pro Dipeptides in the Proteins Collagen and Bora

Paul Joseph Arcoria

General Audience Abstract

The proline amide (Xaa–Pro) in peptides and proteins is unique in that it allows for *cis-trans* isomerization. The triple-helix region of human collagen consists mostly of the repeating sequence $(Gly–Pro–Hyp)_n$. Xaa–Pro amide-bond isomerization is rate-limiting for triple-helix formation. We eliminated isomerization at one position in a collagen-like peptide with a locked alkene mimic of Gly–Pro to attempt to stablize the triple-helix. Our computational results predicted that a fluoro-alkene Gly–Pro isostere would be a close geometric match for the native amide. Experimental results showed that a collagen-like peptide with a fluoro-alkene Gly–Pro isostere has an unfolding temperature that is 6.9 °C lower than the native control peptide. ^{19}F NMR data of the collagen-like peptide shows a surprising deshielding of the fluorine nucleus, suggesting its participation in a stabilizing $n\rightarrow\pi^*$ electronic interaction, similar to the native amide.

Isomerization also plays a key role in proper cell division. We followed established methods to synthesize the cis- and trans-locked alkene mimics of Boc–Ser–Pro–OH and converted them into the 3,5-dinitrobenzyl ester conjugates. The 3,5-dinitrobenzyl ester is recognized by the dinitrobenzyl flexizyme (dFx) for enzymatic aminoacylation of tRNA. Once the alkene isosteres are aminoacylated, they will be incorporated into a full-length cell cycle regulatory protein called Bora to determine whether the *cis*- or *trans*-Pro state is necessary for healthy human mitosis, and which results in cancerous human mitosis.

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Felicia Etzkorn, for her support, guidance, and for fostering a scientific environment in which I could succeed. Many doctorates have been earned through her mentorship, and I am proud to be among them. I would also like to express my gratitude to all faculty members who have served on my committee, Dr. Webster Santos, Dr. Emily Mevers, Dr. Paul Carlier, and Dr. Paul Deck, for their intellectual support and advice. I'd especially like to thank Dr. Emily Mevers for graciously stepping up and filling a vacant role when another member had to step down. She also helped tremendously in pushing my final project towards the finish line, for which I am extremely grateful.

I'd also like to thank all my current and former colleagues and friends who have not only provided research advice over the last five years, but who also made life more enjoyable in Blacksburg. Special thanks to (in no particular order): Johnathan Bowen, Daniel Foster, Karla Piedl, Connor Swetkowski, Justin Grams, and Russell Fritzemeier.

Finally, my family has my deepest gratitude. The only thing more difficult than earning a doctoral degree is being the family member of a Ph.D. candidate. Through all the late nights, missed calls, and working weekends, my family has been incredibly understanding. I especially give thanks to my wife, Casey, and my parents, Fran and Phil, for their encouragement and support through both the good times and the bad.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

Ac	Acetyl
Ala, A	Alanine
Asp, D	Aspartic acid
AurA	Aurora Kinase A
aq.	Aqueous
azGly	Aza-Glycine
Bn	Benzyl
Bora	Aurora Borealis
Boc	*tert*-Butoxycarbonyl
Bu_3N	Tributylamine
CD	Circular dichroism
CDK	Cyclin dependent kinase
Clp	(2*S*,4*R*)-4-Chloroproline
clp	(2*S*,4*S*)-4-Chloroproline
COSY	Correlation spectroscopy
D	Debye
DCC	*N*,*N*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
δ_{BD}	Distance between donating and accepting atoms in an n→π* interaction
dFx	3,5-Dinitro flexizyme
DIAD	Diisopropyl azodicarboxylate
DIEA	Diisopropylethylamine
DMAP	*N*,*N*-Dimethyl-4-amino pyridine
DMF	Dimethylformamide
DMTr	4,4′-Dimethoxytrityl
E2	Second-order perturbation energy
ee	Enantiomeric excess

Et	Ethyl
Et_3N, TEA	Triethylamine
Flp	(2*S*,4*R*)-4-Fluoroproline
flp	(2*S*,4*S*)-4-Fluoroproline
Fmoc	9-Fluorenylmethoxycarbonyl
FRET	Fluorescence resonance energy transfer
G2	Period of rapid cell growth during mitosis
Glu, E	Glutamic acid
Gly, G	Glycine
GSK3β	Glycogen synthase kinase 3β
h	Hour
HATU	2-(7-Aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium-hexafluorophosphate
HBTU	*O*-Benzotriazole-*N*,*N*,*N*',*N*',-tetramethyl-uronium-hexafluorophosphate
HILIC	Hydrophobic interaction liquid chromatography
HOAt	1-Hydroxy-7-azabenzotriazole
6-Cl-HOBt	*N*-Hydroxy-6-Cl-benzotriazole
HOBt	*N*-Hydroxy-benzotriazole
HPLC	High performance liquid chromatography
Hyp	(2*S*,4*R*)-4-Hydroxyproline
im	Imidazole
i-Pr	*iso*-Propyl
*i*PrOH	*iso*-Propanol
IR	Infrared spectroscopy
IVTT	*in vitro* Transcription-translation
$K_{cis/trans}$	Equilibrium constant between the *cis*- and *trans*- conformations
LDA	Lithium diisopropylamine
M	Mitosis

MALDI	Matrix-assisted laser desorption/ionization
MBHA resin	Methylbenzhydryl amine resin
Me	Methyl
Mep	(2*S*,4*S*)-4-Methylproline
mep	(2*S*,4*R*)-4-Methylproline
min	Minute
MP2	Møller–Plesset perturbation theory
mRNA	Messenger RNA
MS	Mass spectroscopy
n→π*	Non-bonding orbital overlapping with the π* orbital
NBO	Natural bond orbital
NIMA	Never in mitosis A kinase
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance spectroscopy
nOe	Nuclear Overhauser enhancement spectroscopy
Ns	*o*-Nitrobenzenesulfonyl
p	Phospho-
PBS	Phosphate buffered saline
pdCpA	Phosphodeoxycytosine-phosphoadenosine
Pin1	Protein interacting with NIMA #1
Plk1	Polo like kinase1
Pnt	4-Pentenoic acid
PP2A	Protein phosphatase type 2A
PPh_3	Triphenylphosphine
PPIases	Peptidyl-prolyl *cis-trans* isomerases
PPII	Polyproline type II
Pro, P	Proline
pS-P	phosphoSer-Pro

pT-P	phosphoThr-Pro
rt	room temperature
$scCO_2$	Supercritical carbon dioxide
$SCF^{\beta\text{-}TrCP}$	Skp, Cullin, F-box containing complex, β-transducin repeat-containing protein
Ser, S	Serine
$\sigma \rightarrow \pi^*$	Sigma orbital overlapping with the π* orbital
SPPS	Solid-phase peptide synthesis
Su	Succcinimide
τ_{BD}	Bürgi-Dunitz angle of nucleophilic approach
TBAF	Tetrabutylammonium fluoride
TBS	*tert*-Butyldimethylsilyl
*t*Bu	*tert*-Butyl
TFA	Trifluoroacetic acid
Θ_{BD}	Pyramidalization of accepting carbon in an n→π* interaction
THF	Tetrahydrofuran
Thr, T	Threonine
TLC	Thin layer chromatography
T_m	Melting temperature (mid-point)
TMS	Trimethylsilyl
TOF	Time-of-flight mass spectrometer
tRNA	Transfer RNA
Trp, W	Tryptophan
Tyr, Y	Tyrosine
UV	Ultraviolet spectroscopy
Xaa	Any amino acid in the X position
Yaa	Any amino acid in the Y position

Chapter 1. Introduction to Collagen

Collagen is one of the most abundant and important structural proteins in mammals.[1] In humans, this fibrous protein accounts for more than 30% of all protein, including more than 50% of skin protein and over 90% of the extracellular matrix.[2] Collagen is a structural protein that grants the human body with strength and elasticity while also providing structural integrity to tissues and organs.[3] Collagen is a superfamily with over 30 discovered members; each with its own structure and function.[4] The defining feature of all collagen types is a triple-helix region. Type I collagen is the most prevalent in humans, making up about 90% of all collagen.[5]

1.1 Collagen Structure

1.1.1 Primary Collagen Structure

Type I collagen exists as a polymer with the generic repeating sequence $(Gly–Xaa–Yaa)_n$. A wide variety of amino acids have been discovered in the Xaa and Yaa positions,[6] but certain amino acids are observed at a higher frequency than others. Approximately 28% of Xaa residues are (2*S*)-proline (Pro).[7] About 38% of all Yaa residues are (2*S*,4*R*)-4-hydroxyproline (Hyp),[7] an uncommon amino acid resulting from posttranslational hydroxylation of Pro by the enzyme, 4-prolyl hydroxylase.[8] Glycine (Gly) is the most important amino acid in the sequence. Due to its lack of side chain, Gly is the only amino acid small enough to fit into the compact core of the collagen triple helix. Substitution of Gly with the next smallest amino acid, alanine (Ala), results in osteogenesis imperfecta; a disease that causes low bone density with a high fracture rate.[9] (Gly–Pro–Hyp) is the most common repeating sequence.[7]

While Pro and Hyp are not required to form a triple helix, they are preferred. The tertiary structure of peptides and proteins can be described by their backbone angles, Ψ, Φ, and ω (Figure 1.1).

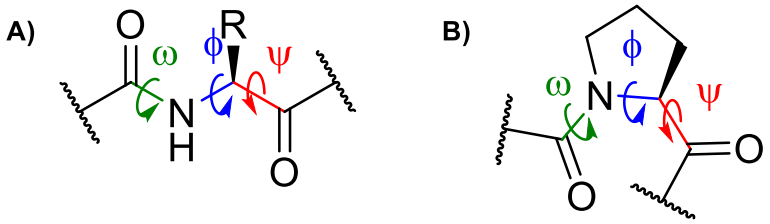


Figure 1.1. The Ψ, Φ, and ω backbone dihedral angles of amino acids. **A)** A generic amino acid in a peptide. The Φ dihedral angle is flexible for most amino acids. **B)** A proline residue in a peptide. The Φ dihedral angle is rigid due to the pyrrolidine ring. The Ψ dihedral angle is flexible in all sequences. The ω angle can either be 180° (trans) or 0° (cis, not shown).

For the collagen triple helix, the Xaa backbone angles Φ = –75°, Ψ = +164° are comparable to the preferred Pro angles Φ = –79°, Ψ = +177°.[10] Likewise, the Yaa backbone angles Φ = –60°, Ψ = +152° are complementary to the favored Hyp angles Φ = –57°, Ψ = +151°.[10] The high degree of similarity between the residues' natural dihedral angles and those necessary for helix formation reduces the entropic cost of protein folding. This phenomena is called the preorganization paradigm.[11]

Pro and Hyp are also favored in their respective Xaa and Yaa positions due to Cγ-ring puckering. (Figure 1.2).[12] In the absence of a 4-position substituent, Pro can adopt the sterically-restrictive C$^{\gamma}$-*endo* pucker that is preferred in the Xaa position (Figure 1.2, left). In the case of Hyp, the C$^{\gamma}$-*exo* pucker is more strongly favored due to the stereoelectronic gauche effect (Figure 1.2, right).[12-14] The gauche effect is the tendency for a molecule to organize so the maximum number of polar bonds are adjacent to one another.[15] The effect of other C$^{\gamma}$-position substitutions on ring puckering will be discussed.

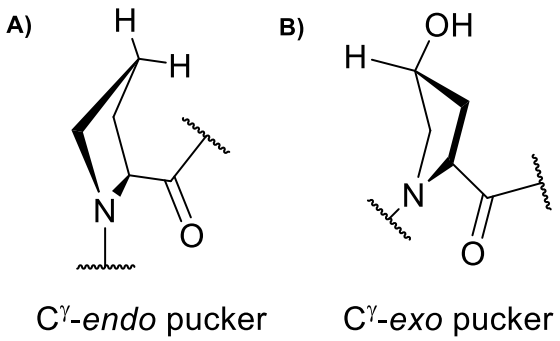


Figure 1.2. Puckering of the γ-carbon on a pyrrolidine ring. **A)** C$^{\gamma}$-*endo* puckering is favored by Pro. **B)** C$^{\gamma}$-*exo* puckering is favored by Hyp. Modified with permission of Annual Reviews, Inc. from reference [12], Copyright 2009.

1.1.2 Polyproline Type II Helix

The triple helix is comprised of three parallel polyproline Type II (PPII) helices. The PPII helix is one of the most abundant secondary structures[16] and is found in both folded[17, 18] and unfolded proteins.[19] They have proven to be key in biological signal transduction, transcription, cell motility, and immune response.[20, 21] PPII helices are shown to be one of the major secondary structures in unfolded proteins and are thought to aid in the energetics of protein folding.[22, 23] Due to their high rigidity and conformational stability, they are commonly used as "molecular rulers" and "molecular scaffolds".[24]

Polyproline peptides can adopt the PPII helix with chain lengths as short as six residues.[25] Early powder diffraction of polyproline peptides showed that all amide bonds are in the *trans* conformation and that every third residue is stacked on top of each other.[26] More information on their dynamic behavior was obtained through FRET, CD, IR, Raman, and NMR.[27-30] However, PPII helices have no characteristic backbone hydrogen bonds that are common in α- and β-sheets, so detecting them directly by NMR or X-ray crystallography can often be problematic.[31, 32]

The average dihedral angles for a PPII helix are Φ = –75°, Ψ = +145°, and ω = 180°.[17, 33] The average helical pitch is about 9.3 Å/turn and 3.0 residues/turn, therefore the rise per residue is approximately 3.1 Å.[26] A hexaproline PPII helix has been crystallized for X-ray diffraction with an *N*-terminal *p*-bromobenzoyl moiety and a *C*-terminal carboxylic acid.[34] This is the first recorded crystal structure of a PPII helix, and it showed an average helical pitch of 8.98 ± 0.14 Å, thus the rise per residue is 3.0 Å.[34] There is also a correlation between backbone dihedral angles and C^{γ} ring pucker. The closer the Φ and Ψ angles were to –65° and +140°, respectively, the ring adopted a stronger C^{γ}-*exo* pucker.[34] On the other hand, when the Φ and Ψ angles were around −73° and +155°, respectively, the Cγ-*endo* pucker became more pronounced.[34]

There has been much debate about stabilizing factors of the PPII helix, specifically the role of hydration and the n→π* interaction, both of which will be discussed in detail. The surface of a PPII helix is more exposed than most other structures.[17] It had previously been thought that coordinating water molecules were critical for the stability of the PPII helix.[16, 35, 36] However, the lack of water molecules in the crystal structure, despite using wet solvents for crystallization, indicates that hydration is not a prerequisite for PPII helicity.[34] The crystals were grown in organic solvents so these findings do not exclude the stabilizing effect of hydration in aqueous solutions; it merely eliminates hydration as a requirement for helix folding. The distances and angles between the carbonyl oxygen and carbonyl carbon of its respective neighboring residue ($O_{i-1}\cdots C_i$), as well as pyramidalization of the accepting carbon, all indicate that every PPII amide participates in an n→π* interaction.[34]

1.1.3 Collagen Triple Helix

The triple-helix model has been determined by X-ray diffraction of the collagen fibers in tendons.[37-40] It consists of three parallel PPII helices that coil around each other about a common

axis to form the right-handed triple helix (Figure 1.3).[41] The three PPII strands are polymers with the sequence $(Gly–Xaa–Yaa)_{333}$, as mentioned earlier. They are offset by a single residue allowing for a ladder of $N–H_{(Gly)} \cdots O=C_{(Xaa)}$ interstrand hydrogen bonding (Figure 1.3B).[41, 42] The hydrogen bonds are perpendicular to, and cross the helical axis. In the triple helix, Gly is the only residue capable of being a hydrogen bond donor and its N–H hydrogen atom points towards the center of the helix. This topology forces the necessary interstrand hydrogen bond and simultaneously prevents the formation of harmful aggregate structures, such as amyloid fibrils.[43]

The helical parameters of the triple helix have been reported from fiber diffraction patterns, ranging from 3.36 residues/turn (107° twist) to 3.27 residues/turn (110º twist).[44, 45] These values are averaged and the accepted parameters today are a 10/3 helix with 3.33 residues/turn (108º twist).[39] Each of the PPII strands adopts a left-landed helix with 7/2 symmetry and contain about 333 amino acids.[7] Each residue propagates the collagen helix by 2.86 Å for a total fibril length of about 300 nm.[7, 26]

Additional structural data was uncovered by studying (Gly–Pro–Pro) collagen-like peptides. X-ray diffraction at 1.0 Å resolution shows that the individual strands of a $(Gly–Pro–Pro)_9$ triple helix adopts similar helical parameters to the PPII helices of collagen.[46] The crystal structure of a $(Gly–Pro–Pro)_{10}$ triple helix has also demonstrated analogous $N–H_{Gly} \cdots O=C_{Pro}$ interstrand hydrogen bonding[47] demonstrate similar molecular conformation to the native $(Gly–Pro–Hyp)_{10}$ triple helix with similar carbonyl hydration patterns; however it has different crystal packing.[48] It is suspected that an extended hydration network involving the Hyp hydroxyl is the cause of the crystal packing difference. Such a hydration network is critical for the lateral assembly and supermolecular structure of collagen.[49]

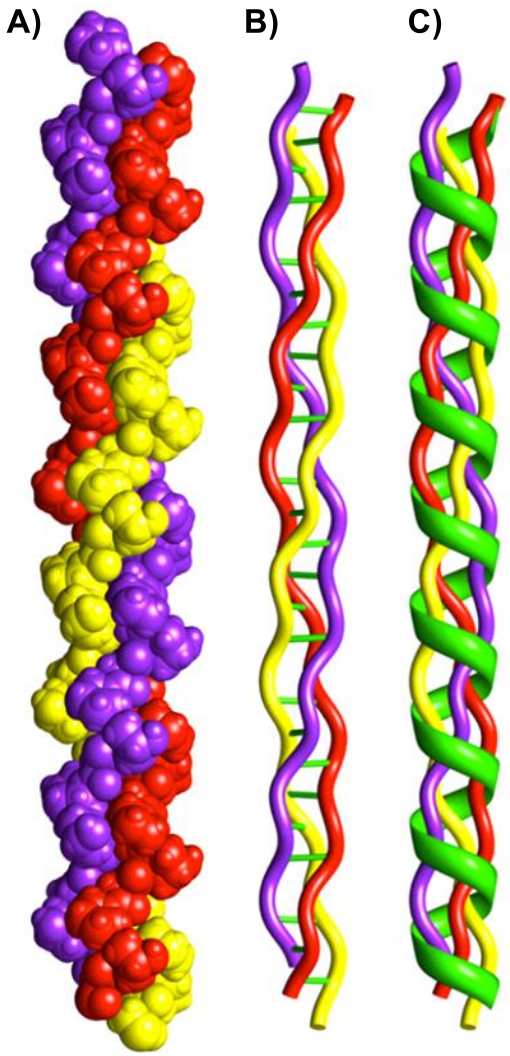


Figure 1.3. Different representations of a (Gly–Pro–Hyp)$_{10}$ collagen triple helix. **A)** The three PPII helix chains are colored purple, yellow, or red. Space-filling model. **B)** Ribbon diagram showing the interstrand hydrogen bonding in green. **C)** Ribbon diagram with a depiction of the superhelix in green. Modified with permission of Portland Press, Ltd. from reference [41], Copyright 2016.

1.2 Collagen Stability

1.2.1 Dipole-Dipole Interactions

The dipole-dipole interactions found in collagen can best be explained by reviewing the extensive research on the α-helix. The dipole moment of an α-helix is known to run parallel through the helix from the *N*-terminal (positively charged) to the *C*-terminal (negatively charged)

(Figure 1.4).[50-52] Every amino acid in the helix increases the strength of the dipole by about 3.5 D.[50]

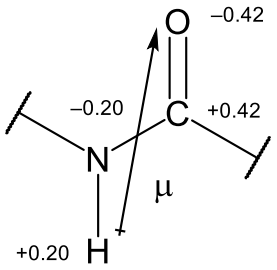


Figure 1.4. Dipole moment and partial charges of an amide α-helix.

Dipoles can be exploited to form a thermodynamically more stable helix through the addition of negatively charged residues (Asp, Glu) near the N-terminus.[51, 53] The electrostatic interaction between the side-chain and the positive dipole can increase helix stability by about 1.6 kcal/mol[51, 53] This stability is lost at higher pH due to negation of charge on the side chain.[53] Dipole-dipole interactions can be manipulated within the collagen triple helix as well. The collagen-like peptide $(Gly–Hyp–Hyp)_7$ has enhanced thermal stability compared to a $(Gly–Pro–Hyp)_7$ peptide by ΔT_m = 2.0 °C because of the extra dipole interactions induced by the additional hydroxyl group.[54] Although Hyp favors the C^γ-*exo* ring pucker which is unfavorable in the Xaa position, the interchain dipole-dipole interaction overcomes the backbone distortion to enhance stability.[54] This effect becomes even more pronounced with stronger electron withdrawing groups on the C^γ carbon. Using the fluorine derivative (2*S*,4*R*)-4-fluoroproline (Flp; μ = 6.0 D),[55] the collagen-like peptide $(Gly–Flp–Flp)_7$ is more stable than native $(Gly–Pro–Hyp)_7$ by ΔT_m = 16 °C.[54]

1.2.2 Hydrogen Bonding

The collagen triple helix is held together by a network of hydrogen bonds between neighboring strands.[1, 37, 38, 56, 57] Two hydrogen bonds are formed within each (Gly–Pro–Hyp) triplet, whereby $N–H_{Gly}$ is a H-bond donor and $C=O_{Pro}$ is a H-bond acceptor (Figure 1.5).[57, 58]

Each hydrogen bond is stabilizing by about 1.4 kcal/mol and is cumulative throughout the helix.[59] The strength of the collagen hydrogen bond is enough to overcome steric restrictions caused by unfavorable pairings between C^{γ}-*endo* and -*exo* puckering in the Yaa and Xaa positions, respectively.[60] Although it is more stable than the native sequence, $(Gly–Hyp–Hyp)_n$ is also considerably more flexible. The flexibility arises from weakened interstrand hydrogen bonding from the slight misalignment of residues caused by backbone distortion of Hyp in the Xaa position.[54]

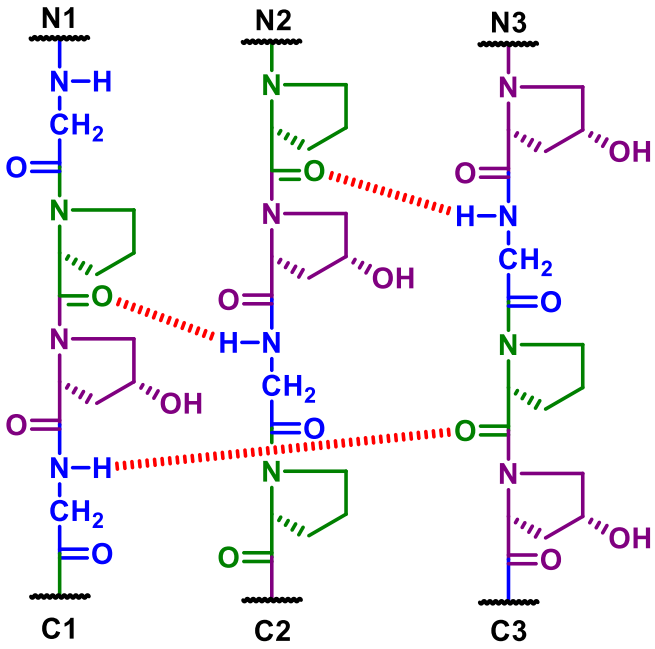


Figure 1.5. Three $N–H_{Gly}$ to $C=O_{Pro}$ interstrand hydrogen bonds among the three PPII helixes.

While hydration is not critical for PPII helicity,[34] it is abundant in a collagen fibril. High resolution diffraction data shows that water plays an integral part of helix stability through intrastrand, interstrand, and interfibrillar hydrogen bonding.[61, 62] The Hyp carbonyl, $C=O_{Hyp}$, is typically hydrogen bonded to two water molecules. The Gly carbonyl, $C=O_{Gly}$, is strictly hydrogen bonded to only one water molecule because its second site is sterically blocked by C_{α} of a neighboring strand.[61] Since $C=O_{Pro}$ is already involved in interstrand hydrogen bonding (Figure

1.5), it is never hydrated.[61] The hydroxyl groups on Hyp are also hydrogen bonded to an average of 1.7 water molecules.[61] Some of these hydrogen bonds form interstrand water bridges for enhanced helix stability while other hydrogen bonds form interfibrillar bridges that help in stabilizing collagen fibrils.[62]

1.2.3 n→π* Interaction

Similar to the stabilization of saturated heterocycles through the anomeric effect,[63] protein secondary structure is also stabilized through the delocalization of lone pairs into neighboring anti-bonding orbitals.[64] Termed the n→π* interaction (Figure 1.6), this is stabilizing to only residues in the trans conformation.[65]

The n→π* interaction has been located in about 72% of α-helices[66] resulting in cumulative stability of around 1.0 kcal/mol.[66, 67] In an n→π* interaction, the p-type lone pair of $C=O_{i-1}$ is delocalized into the π* orbital of $C=O_i$[64-66, 68-71] following the Bürgi-Dunitz trajectory for nucleophilic approach.[64, 66, 68] The Bürgi-Dunitz trajectory is where the angle, θ, between the approaching nucleophile and electrophilic carbonyl is about 109°.[72] The distance, d, between both moieties must be less than the sum of their respective van der Waals radii (3.22 Å in collagen), and some degree of out-of-plane pyramidalization, Δ, of the electrophilic carbonyl is usually

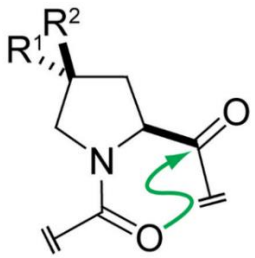


Figure 1.6. Prolyl n→π* interaction showing the electron lone pair of one oxygen delocalizing into the adjacent carbonyl. Used with permission of Annual Reviews, Inc. from reference[12], Copyright 2009.

present (Figure 1.7).[72-74] The n→π* interaction is a dominant force in collagen because of the restricted Φ dihedral angle of proline. The mean Φ dihedral angle of α-helices with n→π* delocalization is –63.3 ± 5.6° while the Φ dihedral in proline is –65.0 ± 10.6°, establishing it as a strong n→π* acceptor.[66] Interestingly, although their dihedral angles match closely, its lack of a hydrogen-bond donor causes Pro to destabilize α-helices.[75]

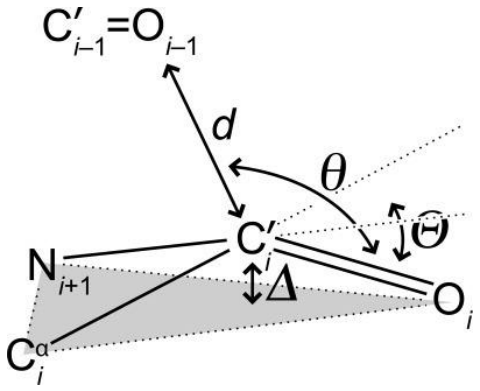


Figure 1.7. Bürgi-Dunitz trajectory for nucleophilic attack of a carbonyl lone pair as the nucleophile. To determine the presence of an n→π* interaction, the distance (d) between the donating oxygen and accepting carbon must be less the same sum of their van der Waal radii. Open access of The Protein Society from reference[68], Copyright 2011.

The Etzkorn lab has recently conducted a conformational analysis to determine the propensity for n→π* interactions to occur in the triple helix.[76] Using small models, they showed that every amide in the Gly–Pro–Pro sequence participated as an n→π* donor and n→π* acceptor, to some extent. At some geometries, the donating oxygen was seen interacting with the *re*-face of the accepting carbonyl, which is dominantly found in native collagen. Other geometries had the donating oxygen interacting with the *si*-face instead. NBO calculations estimate the strength of these interactions from < 0.1 – 1.29 kcal/mol.[76] At their respective collagen-like geometry, every amide participated simultaneously as an n→π* donor and n→π* acceptor.[76]

In general, the two lone electron pairs of a carbonyl are not degenerate One lone pair is rich in s-character (n_s = 60% s, 40% p) and lies along the carbonyl σ-bond and the other is rich in p-character (n_p = 100% p) and is located orthogonal to the π-bond (Figure 1.8).[66] A carbonyl's unsatisfied hydrogen bonding potential increases the energy of a system and allows for structural destabilization through undesired intermolecular hydrogen bonding.[68, 69] Therefore, most α-helix carbonyls form non-covalent interactions: n_s delocalizes into the σ* orbital of N–H_{i+4} to form a hydrogen bond and n_p delocalizes into C=O_{i+1} π* orbital to form an n→π* interaction.[12, 66, 69] It is also theorized that the n→π* interaction increases the strength of interstrand hydrogen bonding by further polarizing the π-bond of C=O_{i+1} and better aligning adjacent strands.[68] Like most non-covalent interactions, the strength of n→π* donation is based on geometry, donor/acceptor strength, and environment,[69] so not all π systems are appropriately situated. The chloro-alkene isostere, for example, has been predicted computationally to be an n→π* donor, but not an n→π* acceptor.[77] It has also been shown experimentally that the fluoro-alkene is not an n→π* acceptor due to steric repulsion from the π cloud, as well as having a less polarized alkene π* orbital.[71]

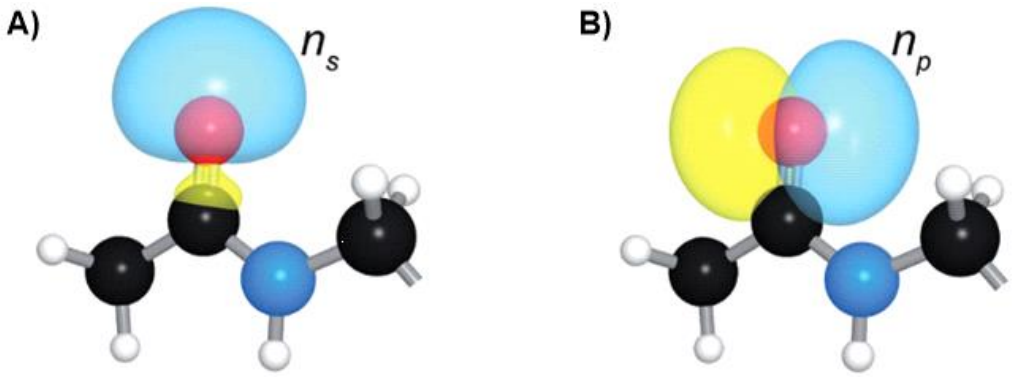


Figure 1.8. Non-degenerate electron lone pairs on a carbonyl oxygen. A) s-rich electrons located opposite of the C–O σ-bond and used for hydrogen bonding. B) p-rich electrons located orthogonal to the π-bond and used for n→π* donation. Used with permission of Springer Nature from reference [66], Copyright 2010.

1.2.4 Cγ-Substitutions

Substitution on C^{γ} of Pro is of considerable interest because of preorganization caused by the gauche effect.[15] Due to the gauche effect, the dihedral angles of Hyp are arranged in a conformation similar to that of the Yaa position of collagen (Figure 1.9).[13, 14]

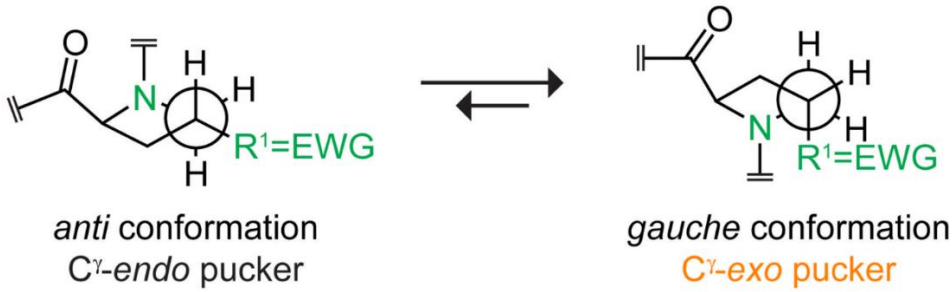


Figure 1.9. Electron withdrawing groups in the (4*R*)- position stabilize C^{γ}-*exo* puckering via the gauche effect. Used with permission of Annual Reviews, Inc. from reference [12], Copyright 2009.

Hyp is critical for structural stability by forming a network of inter- and intramolecular water bridges,[61, 62] dipole-dipole interactions,[54] and decreasing the entropic cost of folding.[78] To probe the stabilizing effect of different 4-position substitutions, the stability of several collagen-like peptides were tested in the Yaa position: (2*S*,4*R*)-4-fluoroproline (Flp), (2*S*,4*S*)-4-methylproline (Mep), and (2*S*,4*R*)-4-chloroproline (Clp).[78-84] All three models show stable triple helix formation via circular dichroism with maxima near 225 nm and large minima around 200 nm, both representative of triple-helix folding.[78-84] The native collagen sequence $(Gly–Pro–Hyp)_7$ has a melting temperature (T_m) of 36 °C.[64] The peptide $(Gly–Pro–Flp)_7$ exhibited hyperstability with T_m = 45 °C,[64] the methylated ring $(Gly–Pro–Mep)_7$ had T_m = 29 °C,[82] and $(Gly–Pro–Clp)_7$ had T_m = 23 °C (Table 1.1).[81] The additional stability of $(Gly–Pro–Flp)_7$ was surprising due to the poor hydration of fluorine. Entropic stabilization was expected to be overcome by enthalpic destabilization.[54] However, the Flp monomer has a more favorable gauche effect due to the higher electronegativity of fluorine with respect to oxygen ($\chi F = 3.98$, $\chi O = 3.44$), which likely led to a

more preorganized ring.[13, 84, 85] Additionally, fluorine's electronegativity creates a larger dipole on Flp that further stabilizes the helix through stronger interstrand dipole-dipole interactions.[54] Evidence shows that $(Gly–Pro–Hyp)_n$ helices are stabilized by the enthalpy of hydration but $(Gly–Pro–Flp)_n$ helices are stabilized by the entropy of preorganization.[86] Thermal stability decreases when a (Gly–Pro–Flp) guest is introduced into the middle of an Ac-Gly-$(Gly–Pro–Hyp)_8$-Gly-NH_2 host by disrupting the chain of hydration caused by sequential Hyp residues.[87]

Table 1.1. T_m values of Yaa-substituted peptides compared to a native collagen-like peptide.[64, 81, 82]

Yaa-substituted peptide	T_m (°C)	ΔT_m °C
$(Gly–Pro–Hyp)_7$	36	-
$(Gly–Pro–Flp)_7$	45	+9
$(Gly–Pro–Mep)_7$	29	–7
$(Gly–Pro–Clp)_7$	23	–13

The relative stability of $(Gly–Pro–Mep)_7$ also originates from the preorganization of Mep, but for steric reasons rather than electrostatic. The C^γ-methyl substituent protrudes radially outward from the helix (C^γ-*exo* puckering), assuming a pseudoequatorial position to avoid harmful steric interactions.[82] Density functional theory calculations suggest that Mep favors C^γ-*exo* puckering by 1.7 kcal/mol so it is well suited for the Yaa position.[82]

Chlorine provides the pyrrolidine ring with some C^γ-*exo* puckering, but it is less favorable than both Flp and Mep. Chlorine is less electronegative than both fluorine and oxygen ($\chi Cl = 3.16$), so Clp does not preorganize as well as either Flp or Hyp, resulting in a higher entropic cost of folding.[81] Since chlorine is smaller than a methyl substituent, Clp cannot preorganize as well as

Mep in terms of steric repulsion either.[81] The combination of these two factors causes (Gly–Pro–Clp)$_7$ to be the least stable 4-position substitution observed.[81]

Pro has a slight preference for C$^\gamma$-*endo* puckering which preorganizes the ring for the Xaa position.[13] To analyze stabilization afforded by Xaa C$^\gamma$-*endo* puckering, the diastereomers of the aforementioned residues (2*S*,4*R*)-4-fluoroproline (flp), (2*S*,4*S*)-4-chloroproline (clp) and (2*S*,4*R*)-4-methylproline (mep) were tested. The proto-collagen peptide (Gly–Pro–Pro)$_7$ does not form a triple helix at room temperature with T_m = 6 °C.[88] Both collagen-like peptides (Gly–flp–Pro)$_7$ and (Gly–mep–Pro)$_7$ were shown to be considerably more stable with T_m = 33 °C[80] and 13° C,[82] respectively (Table 1.2). The (Gly–clp–Pro)$_7$ peptide also does not form a stable helix at room temperature, but by extending the chain from 7 to 10 repeat units, (Gly–clp–Pro)$_{10}$ has T_m = 33 °C which is more stable than the corresponding (Gly–Pro–Pro)$_{10}$ by ΔT_m = 2 °C.[81]

Table 1.2. T_m values of Xaa-substituted peptides compared to a native collagen-like peptide.[64, 80-82, 88]

Xaa-substituted peptide	T_m (°C)	ΔT_m °C
(Gly–Pro–Hyp)$_7$	36	-
(Gly–Pro–Pro)$_7$	6	–30
(Gly–flp–Pro)$_7$	33	+27
(Gly–mep–Pro)$_7$	13	+7
(Gly–clp–Pro)$_7$	10	+4
(Gly–mep–Flp)$_7$	58	+52

Interestingly, all Xaa C$^\gamma$-*endo* puckered preferred diastereomers in collagen-like peptides are less stable than the Yaa C$^\gamma$-*exo* puckered diastereomers. This is possibly caused by the greater

extent of backbone preorganization in the Yaa position.[80] The isomers in the Yaa position significantly preorganize all backbone dihedral angles favorably, whereas the Xaa position isomers only favorably arrange ψ and φ angles but unfavorably affects the ω angle.[64, 80] A combination of both isomers can be employed to synthesize hyperstable collagen mimics. For example, (Gly–mep–Flp)$_7$, which is composed of 2/3 unnatural amino acids, is more stable than the native (Gly–Pro–Hyp)$_7$ by ΔT_m = 22 °C.[78]

1.2.5 Aza-Glycine

Glycine is a requisite part of the backbone in triple helical regions of all collagen types, found always at the first position of each (Gly–Xaa–Yaa) triplet.[6, 89] Glycine was once thought to be an irreplaceable cornerstone of the triple helix, but recent work suggests that a Gly mimic, Aza-Glycine (azGly), confers hyperstability on collagen-like peptides.[90, 91] Molecular dynamics simulations indicate that the α-N–H hydrogen may form a bifurcated hydrogen bond between α-N–H···O=C$_{Gly}$ (d = 2.2 Å) and α-N–H···O=C$_{Pro}$ (d = 3.2 Å) on a neighboring chain (Figure 1.10).[90] This agrees with experimental results that show when an (azGly–Pro–Hyp) guest is introduced to an Ac-(Gly–Pro–Hyp)$_7$-NH$_2$ host, T_m increases by 6-11 °C depending on its location, with a preference towards the middle of the chain.[90, 91] Furthermore, the additional hydrogen bond is thought to help preorganize azGly residues by restricting the Ψ dihedral, allowing peptides with just one azGly residue to refold faster.[90, 91]

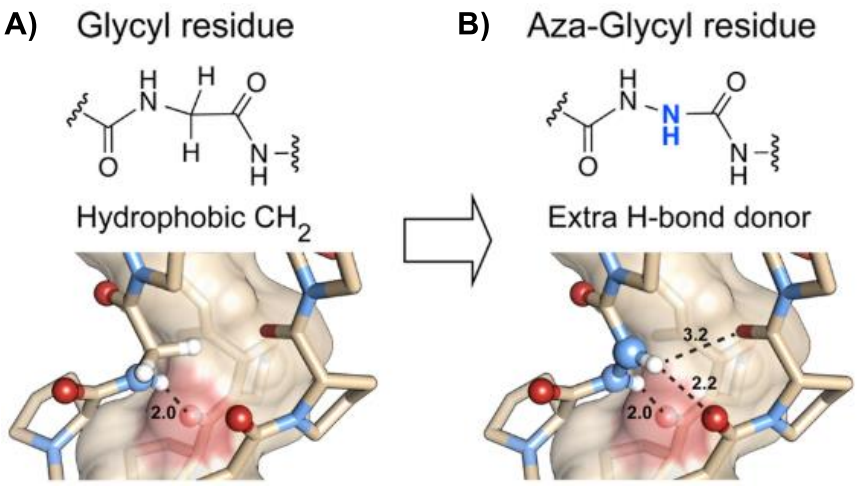


Figure 1.10. Interstrand hydrogen bonding of collagen and Aza-Glycine mimic. **A)** Endogenous collagen. **B)** The additional hydrogen bonding of Aza-Glycine. Reprinted with permission from Zhang, Y.; Malamakal, R. M.; Chenoweth, D. M., Aza-Glycine Induces Collagen Hyperstability. *J. Am. Chem. Soc.* **2015,** *137*, 12422-12425. Copyright 2015 American Chemical Society.[90]

1.2.6 Isomerization

Connecting the monomers of collagen, and of all peptides, is the peptide bond. Peptide bonds are planar amides that form a critical part of protein backbones.[57] Amide bond rotation leads to cis/trans isomerization of residues which affects the structure and function of proteins.[92] The barrier of rotation of a peptide bond is about 15-23 kcal/mol[93] due to the high degree of double bond character, as suggested by the resonance model.[93, 94] The resonance model is also used to explain the stability of the amide bond, which has a half-life of 500 years for uncatalyzed hydrolysis.[95] The lone pair of electrons on nitrogen sit in an unhybridized p-orbital and delocalize into the adjacent carbonyl π-orbital to significantly increase overall stability (Figure 1.11).[96]

Peptide bonds exist naturally in either the trans or cis conformation, but the trans conformation is thermodynamically more stable for most amides.[97] In the trans isomer, the negative end of the N–H dipole is in alignment with the positive end of the C=O dipole while the

cis isomer has both positive ends closer together (Figure 1.11C).[97] For most amides, the trans isomer also significantly reduces steric repulsion.

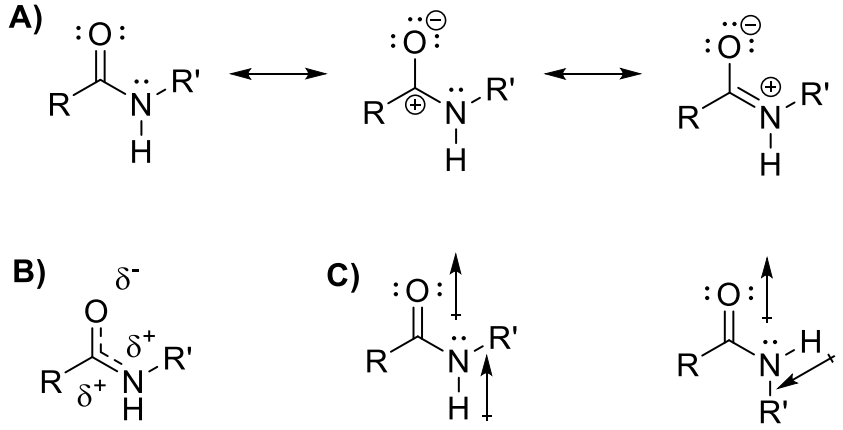


Figure 1.11. A) Resonance contributors of a common peptide bond. **B)** Amide resonance hybrid. **C)** Dipole moment of *trans*- and *cis*- amide.

Proline is unique among the 20 common amino acids in that it is the only species with an easily accessible cis conformation. As part of a peptide, proline forms a 3° amide that has nearly isoenergetic cis and trans conformations caused by steric destabilization of the trans conformer.[98] Whereas less than 1% of 2° amides exist in the cis conformation, as much as 20 ± 15% of peptidyl-prolyl bond are cis.[99] Computational analysis suggests that Gly–*cis*-Pro is about 50,000 times more likely to exist than Gly–*cis*-Gly (Figure 1.12**Error! Reference source not found.**).[67] For prolyl amides in polypeptides, the isomeric ratio is highly dependent on its local environment, such as the preceding residue, length of the peptide chain, and solvent polarity.[93, 100, 101]

Assembly of a PPII helix, and by extension the collagen triple helix, requires that all residues be in the trans conformation. Due to its high cis/trans ratio, proline isomerization is the rate-determining step in helix propagation.[102-104] The steric bulk imposed by the side chain of Pro causes a slightly higher rotation barrier than other peptides,[98] but this process can be catalyzed by

peptidyl-prolyl cis-trans isomerases (PPIases), lowering the barrier from the natural 19-20 kcal/mol to 5-6 kcal/mol.[93, 94, 101]

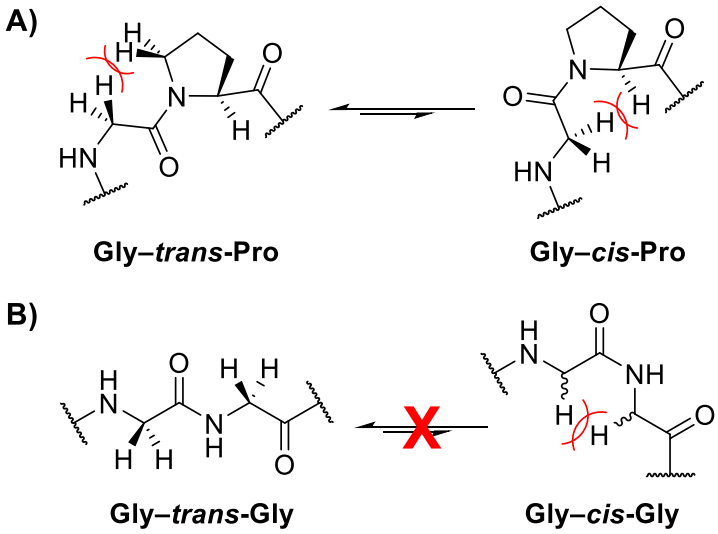


Figure 1.12. Isomerization of the Xaa-Pro motif. **A)** Steric hinderance from the pyrrolidine ring causes destabilization of the trans-conformation leading to an accessible cis conformation. **B)** The trans conformation of secondary peptides is unhindered, so isomerization is very limited.

1.3 Alkene mimics

The 20th century saw major technological advances such as better protein purification protocols and more accurate analysis tools which allowed for the emergence of peptidomimetics as a new class of therapeutics.[105] Peptides are strong drug candidates because of their high potency, selectivity, and diversity, and show lower toxicity with less accumulation in tissues, all without the constraints imposed on small molecule natural products by Lipinski's 'rule of five'.[105] Peptides have not become more dominant therapeutic agents because of their low membrane permeability, poor oral bioavailability (peptides are generally >500 Da), and most importantly, their poor metabolic stability.[105] Peptidases are a class of enzyme that hydrolyze peptide bonds and degrade proteins. Therefore a major focus of peptidomimetics is creating non-hydrolyzable amide isosteres to resist rapid degradation.[105, 106] The large degree of sp^2 character of an amide[93, 94] makes alkene

linkages the most similar functionality to natural peptide bonds in terms of rigidity, bond lengths, and angles.[105] X-ray diffraction shows the average bond length for a trans CH(sp^2)–CH(sp^2) bond is 1.31 Å and the average bond length of a secondary, acyclic amide is 1.33 Å.[107] Additionally, the H–C=C alkene bond angle is 120° while a peptidyl amide bond angle is around 125°.[108] Alkenes can be further modified to mimic other amide properties such as its steric and electronic interactions. Two additional peptide isosteres that show promise in the future of peptidomimetics are the fluoro-alkene and chloro-alkene (Figure 1.13).[105, 106, 108-110]

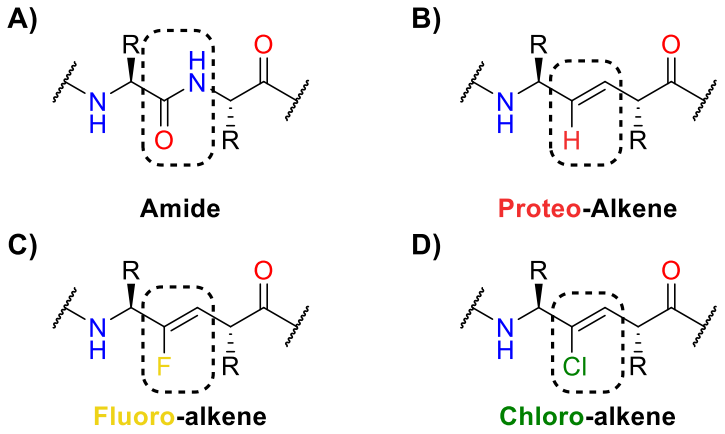


Figure 1.13. The peptide bond and common alkene isosteres. **A)** Native amide. **B)** (*E*)-Proteo-alkene isostere. **C)** Fluoro-alkene isostere. **D)** Chloro-alkene isostere.

1.3.1 Proteo-Alkene

The gauche effect that induces C^{γ}-*exo* pucker in Hyp, Flp, and Clp also contributes to the favorable trans ω angle and therefore a higher $K_{trans/cis}$.[79, 81, 82] The Etzkorn lab hypothesized that replacing an amide bond with a *trans*-alkene isostere to restrict the conformation would lead to a hyperstable triple helix.[111] Of the three repeating peptide bonds in the $(Gly–Pro–Hyp)_n$ sequence, the Gly–Pro amide was determined to be the most advantageous position to modify (Figure 1.14) because it does not participate in the essential interstrand hydrogen bond.[61, 62]

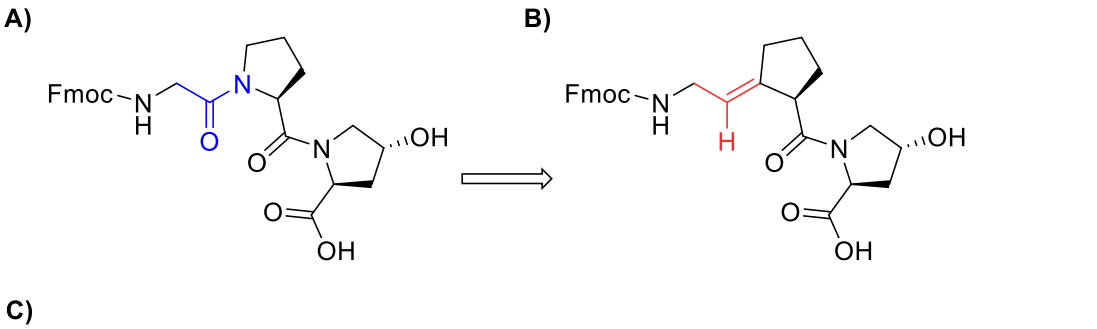


Figure 1.14. Fmoc–Gly–Ψ[(E)(CH=C)]–Pro–Hyp–OH trans-locked alkene mimic. **A)** Natural Gly–Pro–Hyp repeating triplet of collagen. **B)** Gly–Ψ[(*E*)(CH=C)]–Pro–Hyp trans-locked alkene mimic. **C)** (*E*)-alkene model peptide sequence. Reprinted with permission from Dai, N.; Wang, X. J.; Etzkorn, F. A., The Effect of a Trans-Locked Gly-Pro Alkene Isostere on Collagen Triple Helix Stability. *J. Am. Chem. Soc.* **2008,** *130*, 5396-5397. Copyright 2008 American Chemical Society.[111]

Altering the backbone in this way, however, was deleterious. The collagen-like peptide shown in Figure 1.13C had T_m = 28.3 °C while the control, Ac-(Gly-Pro-Hyp)$_8$-Gly-Gly-Tyr-NH$_2$ had T_m = 50.0 °C.[111] A comprehensive computational analysis to help explain these results are discussed in Chapter 2. The unexpected instability of the trans-locked Gly-Pro (*E*)-alkene led to the design and synthesis of two other trans-locked alkene collagen mimics to determine the effect that isomerization elimination has on triple-helix thermostability (Figure 1.15).[112] The two new mimics were synthesized by substituting either the Pro-Gly amide or the Pro-Pro amide with an (*E*)-alkene (Figure 1.15).[113] Through the loss of its hydrogen bond acceptor, isostere **1.3** was used to examine the cost of prolyl isomerization. This could be readily determined because the native Pro-Gly amide bond has a high $K_{trans/cis}$ (> 99%).[101] Therefore, isostere **1.2** has a negligible effect on folding ΔS but removes hydrogen bonding capability as does isostere **1.3**. As expected, interstrand hydrogen bonding has a significant effect on stability. Compared to the control peptide

(Figure 1.15A) with T_m = +31.6 °C, isostere **1.2** had T_m = –24.7 °C (ΔT_m = –57.5 °C).[112] Isostere **1.3**, which removes both prolyl isomerization as well as interstrand hydrogen bonding, reduced thermal degradation to T_m = –22.0 °C (ΔT_m = –53.6 °C).[112] Since both isosteres **1.2** and **1.3** eliminate hydrogen bonding, and **1.3** also removes prolyl isomerization, it can be concluded that the gain from eliminating cis/trans isomerization as the slow step in protein folding is $\Delta\Delta T_m$ = +3.9 °C, and its effects may be cumulatively significant.[112]

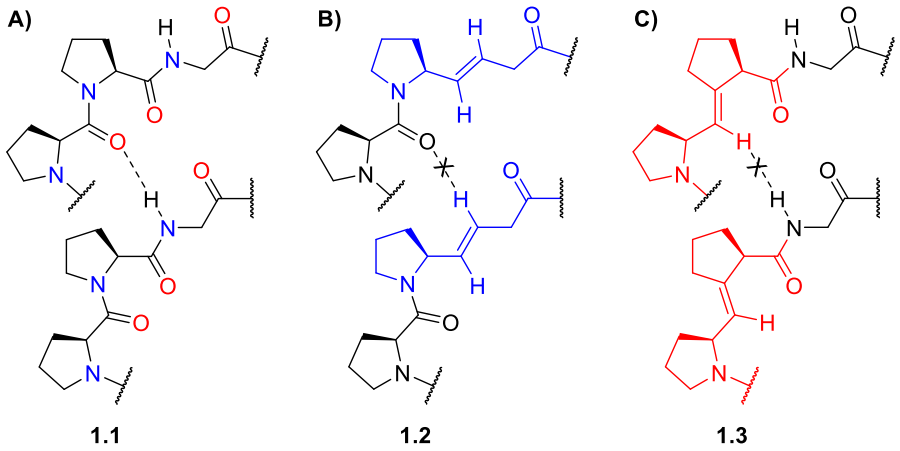


Figure 1.15. Interrupted hydrogen bonding by a Pro–Gly and Pro–Pro (E)-alkene. **A)** Depiction of the interstrand hydrogen bond found in the control peptide H–(Pro–Pro–Gly)$_{10}$–OH **1.1**. **B)** The interrupted N–H$_{Gly}$ hydrogen-bond donor caused by a Pro–Gly (*E*)-alkene in H–(Pro–Pro–Gly)$_4$-Pro–Pro-Ψ[(*E*)CH=C]–Gly–(Pro–Pro–Gly)$_5$–OH Pro–Gly trans isostere **1.2**. **C)** The interrupted C=O$_{Pro}$ hydrogen-bond acceptor caused by a Pro–Pro (*E*)-alkene in H–(Pro–Pro–Gly)$_4$–Pro-Ψ[(*E*)CH=C]–Pro–Gly–(Pro–Pro–Gly)$_5$–OH trans isostere **1.3**. Reprinted with permission from Dai, N.; Etzkorn, F. A., Cis-Trans Proline Isomerization Effects on Collagen Triple-Helix Stability are Limited. *J. Am. Chem. Soc.* **2009,** *131*, 13728-13732. Copyright 2009 American Chemical Society.[112]

1.3.2 Fluoro-Alkene

Fluoro-alkenes are a well-established amide isostere that have been used in peptidomimetics since the 1980s.[114] Computational and experimental data show that fluoro-

alkenes share many geometrical features with the peptide bond including size, bond length, and angle (Figure 1.16**Error! Reference source not found.**).[108, 114-116] The van der Waals radius of oxygen is 1.57 Å while that of fluorine is 1.47 Å.[117-119] Additionally, the C=O bond length of an amide is 1.23 Å compared to the C–F bond length of 1.37 Å of a fluoro-alkene,[108] which suggests that the fluoro-alkene should have similar steric effects as an amide.

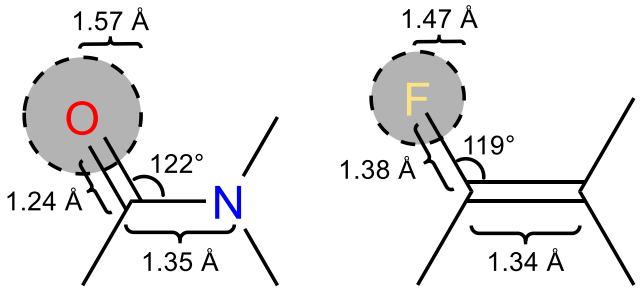


Figure 1.16. Key bond lengths (Å), angles (°) and van der Waals radii (Å) of *N*,*N*-dimethyl acetamide (left) and its fluoro-alkene isostere (right). Calculated with MP2/6-311++G(2d,p) level of theory and the PCM solvation model with water as the solvent.

Fluoro-alkenes exhibit some weak interactions that are analogous to peptidyl hydrogen bonding,[108] but high-level ab initio theory suggests little hydrogen bonding with solvent, and they may not adopt typical peptide conformations.[109] The strength of a $C_{(sp2)}$–F···H–O hydrogen bond has been calculated to be 1.48 kcal/mol while $O_{(sp2)}$···H–O hydrogen bonding is conservatively estimated to be 5 kcal/mol.[117] Fluoro-alkenes do, however, participate in intramolecular hydrogen bonding.[117] When an amide was replaced by a fluoro-alkene in an anti-HIV peptide, the stability of the α-helix remained relatively unchanged when substitution occurred at the N-terminus (ΔT_m = –2.6 °C) and decreased with substitution towards the middle of the chain.[120] The deleterious effect on stability is attributed to the fluoro-alkene having only a hydrogen-bond acceptor and no hydrogen-bond donor.[120]

Reexamination of previously-characterized structures such as the serine protease-inhibitor complex show that a fluoro-alkene engages in n→π* delocalization.[121, 122] A survey of small molecule crystal structures from the Cambridge Structural Database containing short C–X···C=O contacts suggests F and Cl n→π* donation in a variety of different compounds.[123] The data shows no preference of C–F···C=O or C–Cl···C=O dihedral angles, which is indicative of n→π* donation for C-X···C=O compounds.[123] Additionally, X-ray structures with halide anions were analyzed to confirm if halides were capable of n→π* donation or strictly dipole-dipole interactions. The data shows significant clustering around the Bürgi-Dunitz angle of nucleophilic approach as well as carbonyl puckering when the X^{-}···C=O distance was less than the sum of halide and carbon van der Waals radii. Structures in which the X^{-}···C=O distances were greater than the halide and carbon van der Waals radii showed no preference for angle with the C=O and no significant puckering.[123] A computational analysis on the effect of a fluoro-alkene on triple helix stability is described in Chapter 2.[77]

1.3.3 Chloro-Alkene

Compared to the fluoro-alkene, a chloro-alkene may be a stronger n→π* donor due to its size (Cl radius = 1.75 Å.) [118, 119] A larger atom, however, may be deleterious in terms of the compact triple-helix packing. Computational studies show noticeably more overlap between the chloro-alkene non-bonding orbitals and its adjacent carbonyl π* orbital, which is discussed in detail in Chapter 2.[77]

A chloro-alkene is also better able to mimic the dipole of an amide. A natural peptide bond has a dipole moment equal to μ = 3.6 D and vector around 45°.[114] A similar chloro-alkene has μ = 2.0 D and dipole vector around 45° whereas a fluoro-alkene has μ = 1.4 D and dipole vector around 69° (Figure 1.17).[106, 110, 124] Chloro-alkenes have also been seen to act as intramolecular hydrogen-

bond acceptors, so they are considered by some to be a superior isostere.[125-127] However, there is not a lot of literature on the topic, possibly due to the greater 1,3-allylic strain when chloro-alkenes are included in peptides and their synthetic difficulty.[106]

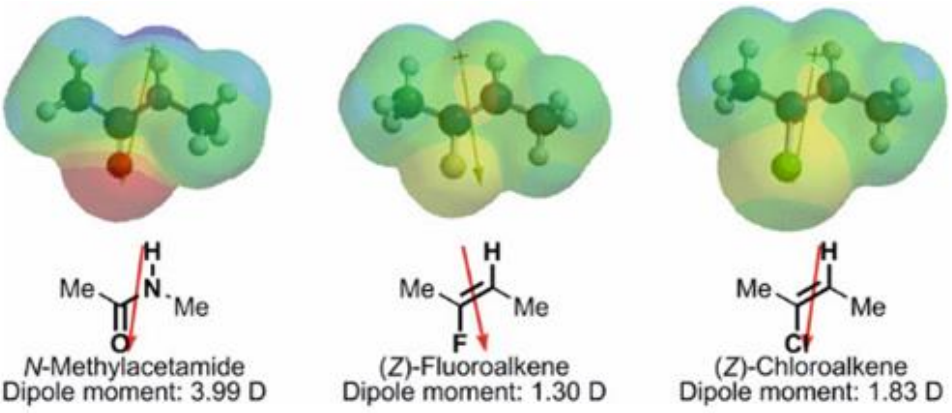


Figure 1.17. Dipole vector, dipole moment and electrostatic potential map of *N*-methyl acetamide (left) with its fluoro-alkene (center) and chloro-alkene (right) mimics. Used with permission of Springer Nature from reference[106], Copyright 2018.

1.4 References

1. Ramachandran, G. N.; Mitra, A. K., Biochemistry of Collagen. *J. Mol. Biol.* **1976,** *107*, 85–92.

2. Kadler, K., Collagen Fibril Formation. *Biochem. J.* **1995,** *2*, 491-619.

3. Birk, D. E.; Silver, F. H.; Trelstad, R. L., *Cell Biology of Extracellular Matrix*. 2nd ed.; Hay, E. D. ed.; Plenum Press: New York, 1991.

4. Yalovac, A.; Ulusu, N. N., Collagen and Collagen Disorders. *J. Pharm. Sci.* **2007,** *32*, 139-144.

5. Di Lullo, G. A.; Sweeney, S. M.; Korkko, J.; Ala-Kokko, L.; San Antonio, J. D., Mapping the Ligand-Binding Sites and Disease-Associated Mutations on the Most Abundant Protein in the Human, Type I Collagen. *J. Biol. Chem.* **2002,** *277*, 4223-4231.

6. Fietzek, P. P.; Kuhn, K., Information Contained in the Amino Acid Sequence of the α1-Chain of Collagen and its Consequences Upon the Formation of the Triple Helix, of Fibrils and Crosslinks. *Mol. Cell. Biochem.* **1975,** *8*, 141-157.

7. Brodsky, B.; Ramshaw, J. A. M., The Collagen Triple-Helix Structure. *Matrix Biol.* **1997,** *15*, 545-554.

8. Kucharz, E. J., *The Collagens: Biochemistry and Pathophysiology*. Springer-Verlag: Berlin, 1992.

9. Gajko-Galicka, A., Mutations in Type I Collagen Genes Resulting in Osteogenesis Imperfecta in Humans. *Acta Biochim. Pol.* **2002,** *49*, 433-441.

10. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal Structure of the Collagen Triple Helix Model $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$. *Protein Sci.* **2002,** *11*, 262-270.

11. Ramshaw, J. A.; Shah, N. K.; Brodsky, B., Gly-X-Y Tripeptide Frequencies in Collagen A Context for Host-Guest Triple-Helical Peptides. *J. Struct. Biol.* **1998,** *122*, 86-91.

12. Shoulders, M. D.; Raines, R. T., Collagen Structure and Stability. *Annu Rev. Biochem.* **2009,** *78*, 929-958.

13. DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L., Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **2002,** *124*, 2497-2505.

14. Improta, R.; Benzi, C.; Barone, V., Understanding the Role of Stereoelectronic Effects in Determining Collagen Stability. 1. A Quantum Mechanical Study of Proline, Hydroxyproline, and Fluoroproline Dipeptide Analogues in Aqueous Solution. *J. Am. Chem. Soc.* **2001,** *123*, 12568-12577.

15. Brunck, T. K.; Weinhold, F., Quantum-Mechanical Studies on the Origin of Barriers to Internal Rotation about Single Bonds. *J. Am. Chem. Soc.* **1979,** *101*, 1700-1709.

16. Stapley, B. J.; Creamer, T. P., A Survey of Left-Handed Polyproline II Helices. *Protein Sci.* **1999,** *8* (3), 587-595.

17. Adzhubei, A. A.; Sternberg, M. J., Left-Handed Polyproline II Helices Commonly Occur in Globular Proteins. *J. Mol. Biol.* **1993,** *229* (2), 472-93.

18. Siligardi, G.; Drake, A. F., The Importance of Extended Conformations and, in Particular, the PII Conformation for the Molecular Recognition of Peptides. *Biopolymers* **1995,** *37* (4), 281-92.

19. Shi, Z.; Woody, R. W.; Kallenbach, N. R., Is Polyproline II a Major Backbone Conformation in Unfolded Proteins? *Adv. Protein Chem.* **2002,** *62*, 163-240.

20. Kelly, M. A.; Chellgren, B. W.; Rucker, A. L.; Troutman, J. M.; Fried, M. G.; Miller, A.-F.; Creamer, T. P., Host−Guest Study of Left-Handed Polyproline II Helix Formation. *Biochemistry* **2001,** *40* (48), 14376-14383.

21. Rath, A.; Davidson, A. R.; Deber, C. M., The Structure of "Unstructured" Regions in Peptides and Proteins: Role of the Polyproline II Helix in Protein Folding and Recognition. *Biopolymers* **2005,** *80* (2-3), 179-85.

22. Hamburger, J. B.; Ferreon, J. C.; Whitten, S. T.; Hilser, V. J., Thermodynamic Mechanism and Consequences of the Polyproline II (PII) Structural Bias in the Denatured States of Proteins. *Biochemistry* **2004,** *43 30*, 9790-9.

23. Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P., Urea Promotes Polyproline II Helix Formation: Implications for Protein Denatured States. *Biochemistry* **2005,** *44* (16), 6269-75.

24. Horng, J.-C.; Raines, R. T., Stereoelectronic Effects on Polyproline Conformation. *Protein Sci.* **2006,** *15* (1), 74-83.

25. Kakinoki, S.; Hirano, Y.; Oka, M., On the Stability of Polyproline-I and II Structures of Proline Oligopeptides. *Polymer Bulletin* **2005,** *53* (2), 109-115.

26. Cowan, P. M.; McGavin, S.; North, A. C. T., The Polypeptide Chain Configuration of Collagen. *Nature* **1955,** *176*, 1062-1064.

27. Walton, A. G.; Rippon, W. B.; Koenig, J. L., Raman Spectroscopy of Proline Oligomers and Poly-L-Proline. *J. Am. Chem. Soc.* **1970,** *92* (25), 7455-7459.

28. Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A., Polyproline and the “Spectroscopic Ruler” Revisited with Single-Molecule Fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (8), 2754-2759.

29. Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A., Effect of Flexibility and *cis* Residues in Single-Molecule FRET Studies of Polyproline. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (48), 18964-18969.

30. Sreerama, N.; Woody, R. W., Molecular Dynamics Simulations of Polypeptide Conformations in Water: A Comparison of α, β, and Poly(pro)II Conformations. *Proteins* **1999,** *36* (4), 400-406.

31. Bochicchio, B.; Tamburro, A. M., Polyproline II Structure in Proteins: Identification by Chiroptical Spectroscopies, Stability, and Functions. *Chirality* **2002,** *14* (10), 782-792.

32. Lam, S. L.; Hsu, V. L., NMR Identification of Left-Handed Polyproline Type II Helices. *Biopolymers* **2003,** *69* (2), 270-81.

33. Kleywegt, G. J.; Jones, T. A., Phi/psi-chology: Ramachandran revisited. *Structure* **1996,** *4* (12), 1395-400.

34. Wilhelm, P.; Lewandowski, B.; Trapp, N.; Wennemers, H., A Crystal Structure of an Oligoproline PPII-Helix, At Last. *J. Am. Chem. Soc.* **2014,** *136*, 15829-15832.

35. Adzhubei, A. A.; Sternberg, M. J.; Makarov, A. A., Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol.* **2013,** *425* (12), 2100-32.

36. Creamer, T. P.; Campbell, M. N., Determinants of the Polyproline II Helix from Modeling Studies. *Adv. Protein Chem.* **2002,** *62*, 263-82.

37. Ramachandran, G. N.; Kartha, G., Structure of Collagen. *Nature* **1955,** *176*, 593-595.

38. Ramachandran, G. N.; Kartha, G., Structure of Collagen. *Nature* **1954,** *174*, 269-270.

39. Rich, A.; Crick, F. H. C., The Molecular Structure of Collagen. *J. Mol. Biol.* **1961,** *3*, 483-506.

40. Rich, A.; Crick, F. H. C., The Structure of Collagen. *Nature* **1955,** *176*, 915-916.

41. Bella, J., Collagen Structure: New Tricks from a Very Old Dog. *Biochem. J.* **2016,** *473*, 1001-1025.

42. Engel, J.; Bächinger, H. P., Structure, Stability and Folding of the Collagen Triple Helix. In *Collagen: Primer in Structure, Processing and Assembly*, Brinckmann, J.; Notbohm, H.; Müller, P. K., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 7-33.

43. Parmar, A. S.; Nunes, A. M.; Baum, J.; Brodsky, B., A Peptide Study of the Relationship Between the Collagen Triple-Helix and Amyloid. *Biopolymers* **2012,** *97* (10), 795-806.

44. Fraser, R. D. B.; MacRae, T. P.; Suzuki, E., Chain Conformation in the Collagen Molecule. *J. Mol. Biol.* **1979,** *129* (3), 463-481.

45. Fraser, R. D.; MacRae, T. P.; Miller, A.; Suzuki, E., Molecular Conformation and Packing in Collagen Fibrils. *J. Mol. Biol.* **1983,** *167* (2), 497-521.

46. Hongo, C.; Nagarajan, V.; Noguchi, K.; Kamitori, S.; Okuyama, K.; Tanaka, Y.; Nishino, N., Average Crystal Structure of $(Pro\text{-}Pro\text{-}Gly)_9$ at 1.0 Å Resolution. *Polym. J.* **2001,** *33* (10), 812-818.

47. Nagarajan, V.; Kamitori, S.; Okuyama, K., Crystal Structure Analysis of Collagen Model Peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$. *J. Biochem.* **1998,** *124* (6), 1117-1123.

48. Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. M., X-ray Crystallographic Determination of a Collagen-Like Peptide with the Repeating Sequence (Pro-Pro-Gly). *J. Mol. Biol.* **1998,** *280* (4), 623-38.

49. Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y. H.; Brodsky, B., Self-Association of Collagen Triple Helic Peptides into Higher Order Structures. *J. Biol. Chem.* **2006,** *281* (44), 33283-90.

50. Hol, W. G. J.; Hali, L. M.; Sander, C., Dipole of the α-helix and β-sheet and Their Role in Protein Folding. *Nature* **1981,** *294*, 532-536.

51. Serrano, L.; Fersht, A. R., Capping and α-helix Stability. *Nature* **1989,** *342*, 296–299.

52. Hol, W. G. J.; van Duijnen, P. T.; Berendsen, H. J. C., The α-helix Dipole and the Properties of Proteins. *Nature* **1978,** *273*, 443-446.

53. Nicholson, H.; Becktel, W. J.; Matthews, B. W., Enhanced Protein Thermostability from Designed Mutations that Interact with α-helix Dipoles. *Nature* **1988,** *336*, 651-656.

54. Shoulders, M. D.; Raines, R. T., Interstrand Dipole-Dipole Interactions Can Stabilize the Collagen Triple Helix. *J. Biol. Chem.* **2011,** *286*, 22905-22912.

55. Siebler, C.; Maryasin, B.; Kuemin, M.; Erdmann, R. S.; Rigling, C.; Grünenfelder, C.; Ochsenfeld, C.; Wennemers, H., Importance of Dipole Moments and Ambient Polarity for the Conformation of Xaa–Pro Moieties – A Combined Experimental and Theoretical Study. *Chem. Sci.* **2015,** *6*, 6725-6730.

56. Mirsky, A. E.; Pauling, L., On the structure of Native, Denatured, and Coagulated Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1936,** *22*, 439–447.

57. Pauling, L.; Corey, R. B.; Branson, H. R., The Structure of Proteins Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U.S.A.* **1951,** *37*, 205-211.

58. Ramachandran, G. N.; Bansal, M.; Bhatnagar, R. S., A Hypothesis on the Role of Hydroxyproline in Stabilizing Collagen Structure. *Biochim. Biophys. Acta* **1973,** *322*, 166-171.

59. Boryskina, O. P.; Bolbukh, T. V.; Semenov, M. A.; Gasan, A. I.; Maleev, V. Y., Energies of Peptide–Peptide and Peptide–Water Hydrogen Bonds in Collagen: Evidences from Infrared

Spectroscopy, Quartz Piezogravimetry and Differential Scanning Calorimetry. *J. Mol. Struct.* **2007,** *827*, 1-10.

60. Erdmann, R. S.; Wennemers, H., Importance of Ring Puckering Versus Interstrand Hydrogen Bonds for the Conformational Stability of Collagen. *Angew. Chem., Int. Ed.* **2011,** *50*, 6835-6838.

61. Bella, J.; Brodsky, B.; Berman, H. M., Hydration Structure of a Collagen Peptide. *Structure* **1995,** *3*, 893-906.

62. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal Structure of a Collagen-Like Polypeptide with Repeating Sequence Pro-Hyp-Gly at 1.4 A Resoultion Implications for Collagen Hydration. *Biopolymers* **2001,** *56*, 8-13.

63. Lemieux, R. U., Effects of Unshared Pairs of Electrons and their Solvation on Conformational Equilibria. *Pure Appl. Chem.* **1971,** *25*, 527-548.

64. Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T., Conformational Stability of Collagen Relies on a Stereoelectronic Effect. *J. Am. Chem. Soc.* **2001,** *123*, 777-778.

65. Hodges, J. A.; Raines, R. T., Energetic of an n-to-π* Interaction that Impacts Protein Structure. *Org. Lett.* **2006,** *8*, 4695-4697.

66. Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N., n-->π* Interactions in Proteins. *Nat. Chem. Biol.* **2010,** *6*, 615-620.

67. Zimmerman, S. S.; Scheraga, H. A., Stability of Cis, Trans, and Nonplanar Peptide Groups. *Macromolecules* **1976,** *9*, 408-416.

68. Choudhary, A.; Raines, R. T., Signature of n-->π* Interactions in α-Helices. *Protein Sci.* **2011,** *20*, 1077-1081.

69. Bartlett, G. J.; Woolfson, D. N., On the Satisfaction of Backbone-Carbonyl Lone Pairs of Electrons in Protein Structures. *Protein Sci.* **2016,** *25*, 887-897.

70. Gauba, V.; Hartgerink, J. D., Self-Assembled Heterotrimeric Collagen Triple Helices Directed through Electrostatic Interactions. *J. Am. Chem. Soc.* **2007,** *129*, 2683-2690.

71. Jakobsche, C. E.; Choudhary, A.; Miller, S. J.; Raines, R. T., n -> π* Interaction and n)(π Pauli Repulsion Are Antagonistic for Protein Stability. *J. Am. Chem. Soc.* **2010,** *132*, 6651–6653.

72. Bürgi, H. B.; Dunitz, J. D.; Shefter, E., Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **1973,** *95*, 5065-5067.

73. Bürgi, H. B.; Dunitz, J. D.; Wipff, G., Ab Initio Study of Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **1974,** *96*, 1956-1957.

74. Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G., Stereochemistry of Reaction Paths at Carbonyl Centres. *Tetrahedron* **1974,** *30*, 1563-1572.

75. Kim, M. K.; Kang, Y. K., Positional Preference of Proline in Alpha-Helices. *Protein Sci.* **1999,** *8* (7), 1492-1499.

76. Etzkorn, F. A.; Ware, R. I.; Pester, A. M.; Troya, D., Conformational Analysis of n→π* Interactions in Collagen Triple Helix Models. *J. Phys. Chem. B.* **2019,** *123*, 496-503.

77. Arcoria, P. J.; Ware, R. I.; Makwana, S. V.; Troya, D.; Etzkorn, F. A., Conformational Analysis of Fluoro-, Chloro-, and Proteo-Alkene Gly–Pro and Pro–Pro Isosteres to Mimic Collagen. *J. Phys. Chem. B* **2022,** *126* (1), 217-228.

78. Shoulders, M. D.; Satyshur, K. A.; Forest, K. T.; Raines, R. T., Stereoelectronic and Steric Efects in Side Chains Preorganize a Protein Main Chain. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 559-564.

79. Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T., Code for Collagen's Stability Deciphered. *Nature* **1998,** *392*, 666-667.

80. Hodges, J. A.; Raines, R. T., Stereoelectronic Effects on Collagen Stability The Dichotomy of 4-Fluoroproline Diastereomers. *J. Am. Chem. Soc.* **2003,** *125*, 9262-9263.

81. Shoulders, M. D.; Guzei, I. A.; Raines, R. T., 4-Chloroprolines: Synthesis, Conformational Analysis, and Effect on the Collagen Triple Helix. *Biopolymers* **2008,** *89*, 443-454.

82. Shoulders, M. D.; Hodges, J. A.; Raines, R. T., Reciprocity of Steric and Stereoelectronic Effects in the Collagen Triple Helix. *J. Am. Chem. Soc.* **2006,** *128*, 8112-8113.

83. Shoulders, M. D.; Raines, R. T., Modulating Collagen Triple-Helix Stability with 4-Chloro, 4-Fluoro, and 4-Methylprolines. *Adv. Exp. Med. Biol.* **2009,** *611*, 251-256.

84. Shoulders, M. D.; Kamer, K. J.; Raines, R. T., Origin of the Stability Conferred Upon Collagen by Fluorination. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 3859-3862.

85. Eberhardt, E. S.; Panisik, N.; Raines, R. T., Inductive Effects on the Energetics of Prolyl Peptide Bond Isomerization: Implications for Collagen Folding and Stability. *J. Am. Chem. Soc.* **1996,** *118*, 12261-12266.

86. Nishi, Y.; Uchiyama, S.; Doi, M.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y., Different Effects of 4-Hydroxyproline and 4-Fluoroproline on the Stability of Collagen Triple Helix. *Biochemistry* **2005,** *44*, 6034-6042.

87. V., P. A.; M., R. J. A.; A., K.; B., B., Triple-Helix Propensity of Hydroxyproline and Fluoroproline. Comparison of Host-Guest and Repeating Tripeptide Collagen Models. *J. Am. Chem. Soc.* **2003,** *125*, 11500–11501.

88. Hodges, J. A.; Raines, R. T., Stereoelectronic and Steric Effects in the Collagen Triple Helix: Toward a Code for Strand Association. *J. Am. Chem. Soc.* **2005,** *127*, 15923-15932.

89. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 A Resolution. *Science* **1994,** *226*, 75-81.

90. Zhang, Y.; Malamakal, R. M.; Chenoweth, D. M., Aza-Glycine Induces Collagen Hyperstability. *J. Am. Chem. Soc.* **2015,** *137*, 12422-12425.

91. Zhang, Y.; Herling, M.; Chenoweth, D. M., General Solution for Stabilizing Triple Helical Collagen. *J. Am. Chem. Soc.* **2016,** *138*, 9751-9754.

92. Baum, N.; Schiene-Fischer, C.; Frost, M.; Schumann, M.; Sabapathy, K.; Ohlenschläger, O.; Grosse, F.; Schlott, B., The Prolyl cis/trans Isomerase Cyclophilin 18 Interacts with the Tumor Suppressor p53 and Modifies its Functions in Cell Cycle Regulation and Apoptosis. *Oncogene* **2009,** *28* (44), 3915-3925.

93. Kang, Y. K.; Park, H. S., Internal Rotation About the C–N bond of Amides. *J. Mol. Struct.* **2004,** *676*, 171-176.

94. Fischer, S.; Dunbrack, R. L. J.; Karplus, M., Cis-Trans Imide Isomerization of Proline Dipeptide. *J. Am. Chem. Soc.* **1994,** *116*, 11931-11937.

95. Radzicka, A.; Wolfenden, R., Rates of Uncatalyzed Peptide Bond Hydrolysis in Neutral Solution and the Transition State Affinities of Proteases. *J. Am. Chem. Soc.* **1996,** *118*, 6105-6109.

96. Berg, J.; Tymoczko, J. L.; Stryer, L., *Biochemistry*. 6 ed.; W. H. Freeman: San Francisco, 2006.

97. Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. University Science: Sausalito, California, 2004.

98. Micheau, J.; Zhao, J., Cis/Trans Configurations of the Peptide C-N bonds: Isomerization and Photoswitching. *J. Phys. Org. Chem.* **2007,** *20*, 810-820.

99. Stellwagen, E., Proline Peptide Isomerization and the Reactivation of Denatured Enzymes. *J. Mol. Biol.* **1979,** *135*, 217-229.

100. Grathwohl, C.; Wuthrich, K., NMR Studies of the Rates of Proline Cis-Trans Isomerization in Oligopeptides. *Biopolymers* **1981,** *20*, 2623-2633.

101. Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G., Side-Chain Effects on Peptidyl-Prolyl cis/trans Isomerisation. *J. Mol. Biol.* **1998,** *279*, 449-460.

102. Bachinger, H., P.; Engel, J.; Bruckner, P.; Timpl, R., The Role of Cis-Trans Isomerization of Peptide Bonds in the Coil ⇄ Triple Helix Conversion of Collagen. *Eur. J. Biochem.* **1978,** *90* (3), 605-613.

103. Davis, J. M.; Boswell, B. A.; Bächinger, H. P., Thermal Stability and Folding of Type IV Procollagen and Effect of Peptidyl-Prolyl cis-trans-Isomerase on the Folding of the Triple Helix. *J. Biol. Chem.* **1989,** *264* (15), 8956-62.

104. Bächinger, H. P.; Bruckner, P.; Timpl, R.; Prockop, D. J.; Engel, J., Folding Mechanism of the Triple Helix in Type-III Collagen and Type-III pN-Collagen. Role of Disulfide Bridges and Peptide Bond Isomerization. *Eur J Biochem.* **1980,** *106* (2), 619-32.

105. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The Future of Peptide-Based Drugs. *Chem. Biol. Drug Des.* **2013,** *81*, 136-147.

106. Tamamura, H.; Kobayakawa, T.; Ohashi, N., Introduction to Mid-Size Drugs and Peptidomimetics. In *Mid-size Drugs Based on Peptides and Peptidomimetics*, Springer: Singapore, 2018; pp 1-16.

107. Allen, F., H.; Kennard, O.; Watson, D., G.; Brammer, L.; Orpen, G., A.; Taylor, R., Tables of Bond Lengths Determined by X-ray and Neutron Diffraction. Part 1. Bond Lengths in Organic Compounds. *J. Chem. Soc., Perkin Trans. II* **1987**, S1-S19.

108. Urban, J. J.; Tillman, B. G.; Cronin, W. A., Fluoroolefins as Peptide Mimetics: A Computational Study of Structure, Charge Distribution, Hydration, and Hydrogen Bonding. *J. Phys. Chem. A* **2006,** *110*, 11120-11129.

109. McKinney, B. E.; Urban, J. J., Fluoroolefins as Peptide Mimetics 2. A Computational Study of the Conformational Ramifications of Peptide Bond Replacement. *J. Phys. Chem. A* **2010,** *114*, 1123−1133.

110. Kobayakawa, T.; Matsuzaki, Y.; Hozumi, K.; Nomura, W.; Nomizu, M.; Tamamura, H., Synthesis of a Chloroalkene Dipeptide Isostere-Containing Peptidomimetic and Its Biological Application. *ACS Med. Chem. Lett.* **2018,** *9*, 6-10.

111. Dai, N.; Wang, X. J.; Etzkorn, F. A., The Effect of a Trans-Locked Gly-Pro Alkene Isostere on Collagen Triple Helix Stability. *J. Am. Chem. Soc.* **2008,** *130*, 5396-5397.

112. Dai, N.; Etzkorn, F. A., Cis-Trans Proline Isomerization Effects on Collagen Triple-Helix Stability are Limited. *J. Am. Chem. Soc.* **2009,** *131*, 13728-13732.

113. Jenkins, C. L.; Vasbinder, M. M.; Miller, S. J.; Raines, R. T., Peptide Bond Isosteres: Ester or (E)-Alkene in the Backbone of the Collagen Triple Helix. *Org. Lett.* **2005,** *7*, 2619-2622.

114. Abraham, R. J.; Ellison, S. L. R.; Schonholzer, P.; Thomas, W. A., A Theoretical and Crystallographic Study of the Geometries and Conformations of Fluoro-Olefins as Peptide Analogues. *Tetrahedron* **1986,** *42*, 2101-2110.

115. Van der Veken, P.; Senten, K.; Kertesz, I.; De Meester, I.; Lambeir, A.; Maes, M.; Scharpe, S.; Haemers, A.; Augustyns, K., Fluoro-Olefins as Peptidomimetic Inhibitors of Dipeptidyl Peptidases. *J. Med. Chem.* **2005,** *48*, 1768-1780.

116. Allmendinger, T.; Furet, P.; Hungerbühler, E., Fluoroolefin Dipeptide Isosteres: The Synthesis of Glyψ(CF=CH)Gly and Racemic Pheψ(CF=CH)Gly. *Tetrahedron Lett.* **1990,** *31*, 7297-7300.

117. O'Hagan, D.; Rzepa, H. S., Some Influences of Fluorine in Bioorganic Chemistry. *Chem. Commun.* **1997**, 645-652.

118. Bondi, A., Van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68*, 441-451.

119. Batsanov, S. S., Van der Waals Radii of Elements. *Inorg. Mater.* **2001,** *37*, 871−885.

120. Oishi, S.; Kamitani, H.; Kodera, Y.; Watanabe, K.; Kobayashi, K.; Narumi, T.; Tomita, K.; Ohno, H.; Naito, T.; Kodama, E.; Matsuoka, M.; Fujii, N., Peptide Bond Mimicry by (*E*)-Alkene and (*Z*)-Fluoroalkene Peptide Isosteres: Synthesis and Bioevaluation of α-helical Anti-HIV Peptide Analogues. *Org. Biomol. Chem.* **2009,** *7*, 2872-2877.

121. Adler, M.; Davey, D. D.; Phillips, G. B.; Kim, S.; Jancarik, J.; Rumennik, G.; Light, D. R.; Whitlow, M., Preparation, Characterization, and the Crystal Structure of the Inhibitor ZK-807834 (CI-1031) Complexed with Factor Xa. *Biochemistry* **2000,** *39*, 12534-12542.

122. Hof, F.; Scofield, D. M.; Schweizer, W. B.; Diederich, F., A Weak Attractive Interaction Between Organic Fluorine and an Amide Group. *Angew. Chem., Int. Ed.* **2004,** *43*, 5056-5059.

123. Kamer, K. J.; Choudhary, A.; Raines, R. T., Intimate Interactions with Carbonyl Groups: Dipole-Dipole or n-->π*? *J. Org. Chem.* **2013,** *78*, 2099-2103.

124. Bauer, P. J.; Exner, O.; Ruzziconi, R.; An, T. D.; Tarchini, C.; Schlosser, M., Identification of Stereoisomers Based on Dielectric Studies: Dipole Moments of Chloroalkenes and Chlorocumulenes. *Tetrahedron* **1994,** *50*, 1707-1716.

125. Metz, A. E.; Podlesny, E. E.; Carroll, P. J.; Klinghoffer, A. N.; Kozlowski, M. C., Axial Chiral Bisbenzophenazines: Solid-State Self-Assembly via Halide Hydrogen Bonds Triggered by Linear Alkanes. *J. Am. Chem. Soc.* **2014,** *136*, 10601-10604.

126. Zhu, Y.; Yi, H.-P.; Li, C.; Jiang, X.; Li, Z., The NH-X Hydrogen Bonding Pattern in Aromatic Amides A Crystallographic and H NMR Study. *Cryst. Growth Des.* **2008,** *8*, 1294−1300.

127. Jelsch, C.; Soudani, S.; Ben Nasr, C., Likelihood of Atom–Atom Contacts in Crystal Structures of Halogenated Organic Compounds. *IUCrJ* **2015,** *2*, 327-340.

Chapter 2. Conformational Analysis of Fluoro-, Chloro-, and Proteo-alkene Gly–Pro and Pro–Pro Isosteres to Mimic Collagen

Paul J. Arcoria, Rachel I. Ware, Sunny V. Makwana, Diego Troya, Felicia A. Etzkorn

2.1 Abstract

Collagen is the most abundant human protein, with the canonical sequence $(Gly\text{-}Pro\text{-}Hyp)_n$ in its triple helix region. Cis-trans isomerization of the Xaa-Pro amide has made two of these amide bonds the target of alkene replacement: the Gly-Pro and the Pro-Hyp positions. The conformations of Gly-Pro and Pro-Pro (as a Pro-Hyp model) fluoro-, chloro- and proteo-alkene mimic models were investigated computationally to determine whether these alkenes can stabilize the polyproline type II (PPII) conformation of collagen. MP2 calculations with various basis sets were used to perform the conformational analyses and locate stationary points. The calculation results predict that fluoro- and chloro-alkene mimics of Gly-Pro and Pro-Pro can participate in $n\rightarrow\pi^*$ donation to stabilize PPII conformations, yet they are poor $n\rightarrow\pi^*$ acceptors, shifting the global minima away from PPII conformations. For the proteo-alkene mimics, the lack of significant $n\rightarrow\pi^*$ interactions and unstable PPII-like geometries explains their known destabilization of the triple helix in collagen-like peptides.

2.2 Author Contribution

Paul Arcoria and Rachel Ware conducted the scans and optimization of all compounds reported in this study with support from Prof. Diego Troya. Paul Arcoria and Sunny Makwana analyzed the parameters of optimized geometries and collagen X-ray crystal structures. The final manuscript was prepared in collaboration between Paul Arcoria and Prof. Felicia Etzkorn. Rachel Ware, Sunny Makwana, and Prof. Diego Troya edited the final manuscript. Prof. Felicia Etzkorn served as an advisor throughout this study and contributed to the preparation, editing submission,

and revision of the final manuscript. This work was published in the Journal of Physical Chemistry B and is available online. Reprinted with permission from Arcoria, P.J., Ware, R.I., Makwana, S.V., Troya, D., Etzkorn, F.A. Conformational Analysis of Fluoro-, Chloro-, and Proteo-Alkene Gly-Pro and Pro-Pro Isosteres to Mimic Collagen. *J. Phys. Chem. B.* **2022**, *126*, 217-228.

2.3 Introduction

Collagen is one of the most important structural proteins in vertebrates, accounting for 30% of all vertebrate protein.[1] It consists of three left-handed poly-proline type II (PPII) helices that intertwine around a common screw axis to form the right-handed triple helix.[2, 3] PPII helices are a frequently occurring protein secondary structure that are prevalent in fibrillar proteins.[4] They serve key roles in signal transduction and protein complex assembly.[4] The three PPII helices of collagen exist as repeating polymers of $(Gly–Xaa–Yaa)_n$, in which 25% of all Xaa residues are (2*S*)-proline (Pro) and 38% of all Yaa residues are (2*S*,4*R*)-hydroxyproline (Hyp).[1] This makes Gly–Pro–Hyp the most common sequence of the triple helix.[5] Gly–Pro–Pro collagen-like peptides have been frequently used in studies of the triple helix.[6]

The tertiary amide of the Xaa-Pro motif has an unusually high proportion of the cis isomer due to steric destabilization of the trans isomer.[7] While over 99% of secondary amides adopt the trans conformation, 10–30% of Xaa-Pro tertiary amides in peptides and unfolded proteins appear as the cis conformation.[8] Propagation of the collagen triple helix is rate-limited by the isomerization of the Xaa-Pro motif.[9]

Due to the high degree of sp^2-character of amides, alkenes have proven to be excellent mimics of prolyl amides.[10-12] However, we have previously shown that converting an Xaa-Pro amide to an alkene decreases the stability of the triple helix.[13, 14] Substitution of a single Gly–Pro

peptide per strand with an alkene in a 27-residue collagen-like peptide leads to a change in the folding temperature of ΔT_m = –21.7 °C compared to the native collagen-like peptide control.[13] Replacement of a Pro-Pro amide, with concomitant loss of an interstrand hydrogen bond, was even more disruptive to triple helix folding with a change in folding temperature of ΔT_m = –53.6 °C compared to the native control.[14] This showed that the entropic reduction afforded by the olefin mimic does not compensate for some missing interactions that are intrinsic to the peptide bond. The destabilization with the proteo-alkene mimics could be caused by the loss of noncovalent interactions that may be present with the native amide in the collagen triple helix.

Raines and coworkers have shown, computationally and experimentally, that collagen is stabilized by a series of electronic n→π* interactions.[7, 15] In the trans conformation of Xaa-Pro models, a lone pair of Xaa C_i=O delocalizes into the Pro C_{i+1}=O π* anti-bonding orbital.[7, 16] Electron delocalization follows the Bürgi-Dunitz trajectory of nucleophilic addition to a carbonyl, in which the angle of approach (τ_{BD}: $\angle O_i \cdots C_{i+1}$=O) is approximately 109°, and the distance between the donor oxygen and acceptor carbon (δ_{BD}) is within the sum of their respective van der Waals radii (Figure 2.1).[17] Donation into the π* orbital also produces slight sp^3 character in the acceptor C_{i+1} atom resulting in pyramidalization, as measured by the angle Θ_{BD}.[18] Carbonyl pyramidalization is a signature of the n→π* interaction, and the mean Θ_{BD} angle in α-helices has been reported as +4° ± 1.1°.[18] In α-helices, nearly every carbonyl acts as both an n→π* donor and acceptor.[15] A density functional theory calculation estimated that the n→π* interaction of Ac-Pro-NMe_2 contributes approximately 0.14 (C^γ-endo) or 0.53 kcal/mol (C^γ-exo) of stability, and the weighted average was estimated at 0.27 kcal/mol.[19] However, evidence of the n→π* interaction is lacking in the ^{13}C=O NMR chemical shifts of proteins.[20]

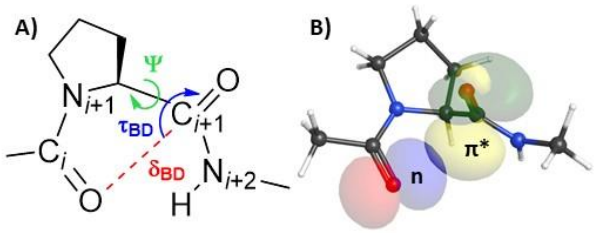


Figure 2.1. Bürgi-Dunitz parameters for an n→π* interaction. **A)** Bürgi-Dunitz trajectory angle (τ_{BD}) and distance (δ_{BD}) of the oxygen lone-pair donor (n) with the carbonyl acceptor (π*) of the subsequent amide bond describes an n→π* interaction found in collagen. The torsion angle varied for the calculations is Ψ. **B)** Calculated NBOs showing the overlap of oxygen non-bonding orbital (blue/red) and C'$_{i+1}$=O antibonding π* orbital (yellow/green).[16]

In a recent paper, we reported the conformational variation in stability of small PPII models of the amide-amide interactions. We found forward and reverse, re- and si-face, and reciprocal n→π* interactions,[21] and hydrogen bonds that stabilize PPII or competing conformations.[16] We calculated the relative energies of these minima and the barriers between them, as well as the second-order perturbation energies (E2PERT) from Natural Bond Orbital (NBO) calculations of the n→π* and hydrogen-bonded interactions.[16] For the Gly–Pro–Pro (GPP) model, the global minimum shows an n→π* interaction that stabilizes the PPII conformation.[16] For the Pro–Pro–Gly (PPG) model, the global minimum involves a hydrogen bond between the Pro$_{i+1}$ N and the Gly$_{i+2}$ NH that is conformationally distinct from the PPII conformation, though only 1.0 kcal/mole more stable.[16] The Pro–Gly–Pro (PGP) model was shown to be quite conformationally flexible with multiple minima due to the lack of a restricted Φ-bond in Gly.[16]

Due to their rigidity and hydrolytic stability, fluoro- and chloro-alkenes have found use in medicinal chemistry as peptide-bond bioisosteres.[22, 23] The similar sizes of fluorine and oxygen atoms (van der Waals radii = 1.47 Å and 1.52 Å, respectively),[24] and the similar bond lengths of C=O (1.2 Å) and C–F (1.4 Å), allow for analogous pseudo-1,3-allylic strain.[25] However, since

fluorine is smaller and more electronegative than oxygen (EN = 3.98 and 3.44, respectively),[26] n→π* donation by the fluoro-alkene may be less prevalent. Kamer et al. found crystallographic evidence of n→π* interactions of halides (F, Cl, Br, and I) with amide carbonyls.[27] Jakobsche et al. have shown experimentally that a Pro-Gly alkene isostere is a poor n→π* acceptor due to Pauli repulsion with the π-cloud, while a similar fluoro-alkene reduces n)(π Pauli repulsion.[28] Replacement of the OH of Hyp with the analogous stereoisomer, 4(*R*)-fluoroproline, significantly stabilizes the collagen triple helix.[29] We are curious to see if the Gly-Pro and Pro-Pro fluoroalkene isosteres might similarly enhance the stability of the collagen triple helix.

The atomic radius of chlorine (1.75 Å) and the C–Cl bond length (1.72 Å) [24] are expected to have a deleterious effect on the pseudo-1,3-allylic strain and interstrand packing of a chloro-alkene mimic. Chloro-alkenes may also suffer as n→π* acceptors due to Pauli repulsion.[28] However, the larger 3rd row lone pairs and lower electronegativity of chlorine (EN = 3.16)[26] may provide more overlap and stronger n→π* donation. It would be interesting to see how a chloro-alkene mimic might be accommodated in a collagen triple-helix, as chlorine was as 4(*R*)-chloroproline (Clp), in which Cl is substituted for the side chain OH in Hyp.[30]

In this work, we calculate the impacts of fluoro-, chloro-, and proteo-alkene mimics of Gly-Pro and Pro-Pro on the conformation of the PPII helix found in collagen. While the relatively small models of this work cannot capture the extent of all non-bonded interactions within a full triple helix, our intention was to closely examine very local interactions that contribute to the PPII-like conformation with high-level *ab initio* calculations. The calculations herein help us to predict the potential of fluoro- and chloro-alkenes to stabilize collagen-like peptides, and to understand the failure of the simple proteo-alkene mimics to stabilize the collagen conformation.[13, 14] Substituting halogenated alkenes, i.e. a fluoro- or chloro-alkene, may provide non-covalent interactions that

would allow them to adopt the PPII-like conformation,[15, 31, 32] while still reducing the entropic cost of cis-trans proline isomerization in folding.[9]

2.4 Computational Methods

Preliminary geometries were obtained from residues Hyp20–Gly21–Pro22–Hyp23–Gly24 of the collagen triple helix high-resolution crystal structure 1CAG of Bella et al.[33] and modified to the highlighted portions shown in Figure 2.2. (4*S*)-Hydroxyproline was changed to L-proline to reduce the number of heavy atoms in the calculations. Initial optimizations were conducted using Gaussian 09 in WebMO at the second-order Møller-Plesset (MP2) level of theory with the 6-31+G(d) basis set, including the solvent effects of water by the polarizable continuum model.[34, 35] Diffuse functions on heavy atoms were incorporated to capture delocalization of electron density that is a signature of n→π* interactions.[36] Coordinate scans allowed for optimization of all degrees of freedom except for the Ψ dihedral angle being rotated.

For the G(–X)=PP n→π* donor and GP(–X)=P n→π* acceptor models, where X = F, Cl, or H, scans of the Ψ dihedral angle were conducted from +170° to –180° in 10° increments. For the P(–X)=PG n→π* donor and PG(–X)=P n→π* acceptor models, scans of the Ψ dihedral angle were performed from –170° to +180° in 10° increments. For the PG(–X)=P scans, the Φ dihedral angle (C'_i–N_{i+1}–$C\alpha_{i+1}$–C'_{i+1}) was also fixed to –72°, the Φ angle found in collagen-like peptides,[33] and the remaining coordinates were fully optimized. The energy of the global minimum structure in each model series was normalized to 0.0 kcal/mol.

The two lowest-energy minima found in each of the 12 models were fully optimized at the MP2/6-31+G(d) level, then optimized further at the MP2/6-311+G(2d,p) level (Tables 2.1, 2.2, S2.1, S2.2). Single point energies of the unrestrained MP2/6-31+G(d) geometries were also calculated at the MP2/6-311+G(2d,p) level. In all scans, the maxima found closest to the ideal

collagen angles were optimized unrestrained at the MP2/6-31+G(d) level, and single point energies of those geometries were calculated at the MP2/6-311+G(2d,p) level. The energy difference between the single point calculations of minima and maxima at MP2/6-311+G(2d,p) was used to estimate the energy of conformational barriers relative to minima optimized at that level (Tables S2.1, S2.2).

For minima optimized at the MP2/6-311+G(2d,p) level, the distances between C_i–X···C_{i+1}=O or C_i=O···C_{i+1}–X (δ_{BD}), the angles between C_i–X···C_{i+1}=O or C_i=O···C_{i+1}–X (τ_{BD}), and the pyramidalization Θ_{BD} angle formed between the $C\alpha_i$–C'_i–C/N_{i+1} plane and the C'_i=O vector were determined (Tables 2.1, 2.2, S2.1, S2.2).[16] In this work, Θ_{BD} is positive when the accepting carbon is puckered toward the donating atom. The accepting *re* or *si* π* face of the carbonyl for PPII-like geometries of Gly-Pro models (Table 2.1) and Pro-Pro models (Table 2.2) that were within the Bürgi-Dunitz limits were noted. The energy of orbital interactions was quantified by Natural Bond Order (NBO) second-order perturbation analysis (E2).[37] We consider an absence of orbital overlap when the NBO E2 energy is less than a 0.1 kcal/mol threshold. NBOs were used to visualize orbital interactions.[37] All NBO images have an MO isosurface value of 0.050 with 64,000 grid points. The accepting π* face and pyramidalization, Θ_{BD}, are not reported for those models that do not show significant n→π* overlap. An n→π* donation is expected to occur in donor models when the donating atom is fluorine, δ_{BD} < 3.17 Å; when the donating atom is chlorine, δ_{BD} < 3.45; when the donating atom is hydrogen, δ_{BD} < 2.90 Å; and when the donating atom is oxygen, δ_{BD} < 3.22 Å. The accepting atom is carbon in all cases. Geometric parameters of all minima and maxima of Gly-Pro and Pro-Pro mimic models are given in Tables S2.1 and S2.2 respectively.

The average Ψ and Φ dihedral angles (Table S2.3), were measured using the 'phi_psi' PyMOL function, and the average distance (δ_{BD}), and angles (τ_{BD}) (Table S2.4) of the high-

resolution (1.3 Å) crystal structure of the collagen-like peptide, $(Gly\text{-}Pro\text{-}Pro)_{10}$ (PDB ID: 1K6F) were measured using PyMOL.[6, 38] Amino acids ± 6 from both ends of each peptide chain were excluded to limit end-effects on our data. The 1K6F crystal structure includes two independent triple helices giving six measurements at each residue position for more robust averages.[6]

All alkene PPII-like minima were superimposed on the appropriate segment of the 1.3 Å collagen-like peptide crystal structure (PDB ID: 1K6F residues Pro13–Gly14–Pro15–Pro16–Gly17)[6] using Maestro software from Schrödinger, Inc.[39] The root-mean-square deviation (rmsd) between calculated and experimental models were acquired by manual superposition of all heavy atoms, plus the alkene hydrogen for proteo-alkene models (Tables 2.1, 2.2).

2.5 Results and Discussion

We designed our models to capture both n→π* donor and n→π* acceptor interactions of the alkenes that would be present at *one* position of a PPII helix using high-level conformational analysis of separate donor and acceptor models for a total of 12 unique models. In a full-length peptide, these mimics would be subject to both types of non-bonded interactions. For each alkene, the n→π* donor and acceptor models must be considered together to predict how that alkene will behave in a triple-helix peptide. Initial models of Gly-Pro and Pro-Pro fluoro-, chloro-, and proteo-alkene isosteres were created as shown in the colored portions of Figure 2.2.

Replacement of the Gly-Pro amide bond led to the G(–X)=PP model with the alkene X group as an n→π* donor (Figure 2.2A), and the PG(–X)=P model with the C=C π* orbital as an n→π* acceptor (Figure 2.2B). Replacement of the Pro-Pro amide bond led to the P(–X)=PG model with the alkene X group as an n→π* donor (Figure 2.2C) and the GP(–X)=P model with the C=C π* orbital as an n→π* acceptor (Figure 2.2D). In our model numbering scheme, *D* = potential n→π* donor and *A* = potential n→π* acceptor (Table 2.1). The GPP amide is both the Gly-Pro

n→π* donor and Pro-Pro n→π* acceptor, so it is labeled as *D*/*A* in both cases. Models with a Pro–Gly amide substitution were omitted because they naturally adopt the trans conformation, and collagen would not benefit from this substitution.[40] Substitution of an aza-Gly mimic into collagen, however, was stabilizing.[41]

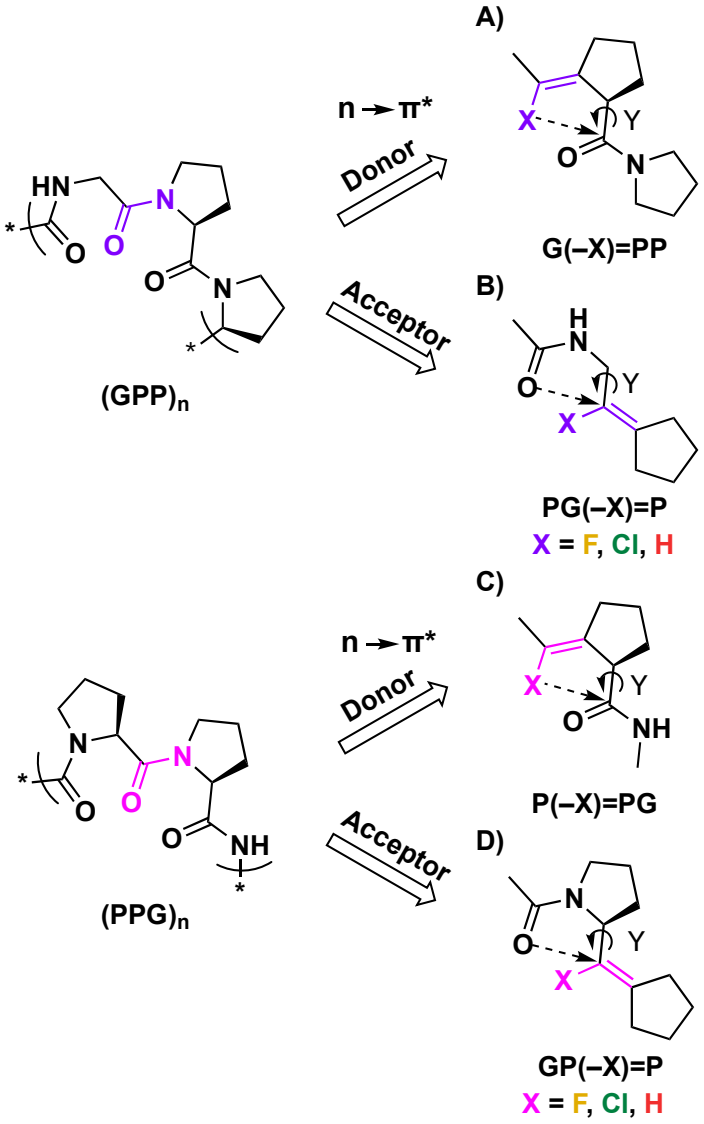


Figure 2.2. Two Xaa-Pro amide bond substitution models: Gly-Pro (top) and Pro-Pro (bottom). **A)** Gly-Pro alkene substitutions are potential n→π* donors to the Pro-Pro amide. **B)** Gly-Pro alkene substitutions are potential n→π* acceptors from the Pro-Gly amide. **C)** Pro-Pro alkene substitutions are potential n→π* donors to the Pro-Gly amide. **D)** Pro-Pro alkene substitutions are potential n→π* acceptors from the Gly-Pro amide.

Only the torsion angles of the bonds that mimic the Ψ-bonds of Gly-Pro or Pro-Pro models were scanned because the Φ-bonds of these models are conformationally constrained by the 5-membered rings. The Φ-bonds of PG(–X)=P mimics are conformationally mobile, so they were restricted to the collagen-like peptide angle (72°).[16] The relative energies of all conformational minima, maxima, and NBO[37] interactions are compared to our previous work on the conformational energy landscape of collagen PPII model amides GPP, PGP, and PPG.[16] The Bürgi–Dunitz parameters of PPII-like geometries are directly compared with the average conformations found in the high-resolution (1.3 Å) crystal structure of the collagen-like peptide $(Gly\text{-}Pro\text{-}Pro)_{10}$ (PDB ID: 1K6F).[6] Each model's PPII-like minimum conformation was superimposed on PDB ID: 1K6F.[6]

2.5.1 Gly-Pro Alkene Isosteres

The Ψ dihedral angle scans for G(–X)=PP donor models (Figure 2.3A) and PG(–X)=P acceptor models (Figure 2.3B) are overlaid with their respective GPP and PGP amide coordinate scans at the MP2/6-31+G(d) level of theory.[16] A complete list of geometric and Bürgi-Dunitz parameters for all minima and maxima calculated at the MP2/6-311+G(2d,p) level can be found in Supplemental Information Table S2.1. The closest minimum to the PPII-like Ψ dihedral angle in each series was analyzed further to determine conformational compatibility with collagen, and the presence or absence of a stabilizing n→π* interaction.

For the G(–X)=PP donor series, all models have their global minima near the GPP *D*/*A***1** global minimum at Ψ = +160° (Figure 2.3A). These will be discussed in detail below. Both fluoro- and chloro-alkene donor models have local minima near Ψ = –20° that are about 2 kcal/mol less stable than their respective global minima, as was found for the GPP amide *D*/*A***1** (Table S2.1).[16] The G(–F)=PP local minimum at Ψ = –18° has no discernible n→π* or hydrogen bonding

interactions (Figure S2.1A). The G(–Cl)=PP local minimum at Ψ = –19° is stabilized by an n→*si*-π* interaction (Figure S2.1B). A similar n→*si*-π* interaction is seen with the amide GPP *D/A***1** at Ψ = –17° (Table S2.1).[16] The only comparable local minimum for the proteo-alkene G(–H)=PP is found near Ψ = –60° where a stabilizing σ→π* interaction is seen (Figure S2.1C).

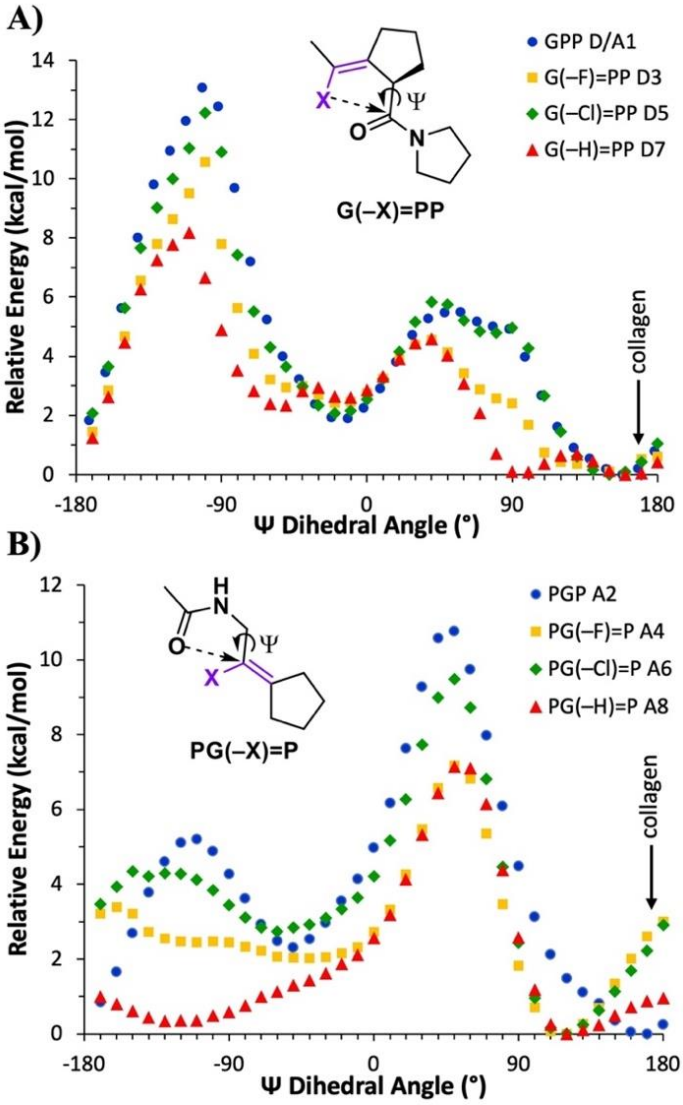


Figure 2.3. Potential energy scans of the Ψ dihedral angle in models of Gly-Pro amide substitutions where X = F, Cl, or H. **A)** G(–X)=PP n→π* donor, and **B)** PG(–X)=P n→π* acceptor models overlaid with their respective PPII amide model energies as a function of Ψ.[16] The global minimum for each model was normalized to 0.0 kcal/mol. The average collagen Ψ dihedral angle from PDB ID: 1K6F in Table S2.3 is labeled.[6]

The local maxima between the global and local minima are all about 4 – 5 kcal/mol higher in energy than the global minima (Figure 2.3A). These maxima result from steric clashing between the X group (or O in the case of the amide) and protons on the 5-membered ring. The global maxima for all models that are located around Ψ = +100° are between 8 – 12 kcal/mol above the global minima (Figure 2.3A). These maxima are caused by steric clashing between the X group and the C_{i+1}=O oxygen.

For the PG(–X)=P acceptor series, all models have their global minima around Ψ = +120°, which are within the same conformational energy well as the PPII-like PGP *A***2** amide global minimum (Figure 2.3B). These are discussed in detail below. The PG(–F)=P and PG(–Cl)=P models have local minima around Ψ = –60°, which are 2 – 3 kcal/mol higher in energy than their global minima (Table S2.1). An n→π* interaction is found at both halo-alkene local minima but missing for PG(–H)=P (Figure S2.2).

The local maxima for both halo-alkenes around Ψ = –160° are about 3 – 4 kcal/mol above the global minima (Figure 2.3B). These low-energy maxima are stabilized by a reverse n→π* interaction (Figure S2.3A, B). The analogous local maximum for the PG(–H)=P model is missing this n→π* interaction (Figure S2.3C). The global maxima for all models located around Ψ = +50° are about 7 – 9 kcal/mol higher in energy than the global minima (Figure 2.3B). These maxima correspond to eclipsing interactions between the X atom and one Gly H_α, as well as the C_i=O oxygen and the other Gly H_α.

2.5.1A Gly-Pro Fluoro-Alkene

The G(–F)=PP *D***3** donor model has great alignment with collagen-like peptide GPP from 1CAG (Figure 2.4A),[6] and *D***3** is the most geometrically similar mimic to GPP *D***1** (Table 2.1). Unrestrained optimization of *D***3** (global minimum) shows Ψ = +162º, which is nearly identical to

the GPP *D***1** angle Ψ = +160º, and superimposes very well with PDB ID: 1K6F (Table 2.1).[6] This suggests the G(–F)=PP isostere is a reasonably good mimic of the PPII conformation. This fluoro-alkene *D***3** global minimum has the Bürgi-Dunitz angle τ_{BD} = 95°, which is favorable for n→π* donation, although the distance δ_{BD} = 3.3 Å is slightly above the 3.17 Å sum of van der Waals radii (Table 2.1). Indeed, the NBO E2 energy of <0.1 kcal/mol and NBO images (Figure 2.4B) suggest the lack of a significant n→π* interaction. Nevertheless, the n→π* interaction does not appear to be necessary to stabilize the PPII-like conformation of the G(–F)=PP model. Other non-bonding interactions beyond n→π* may play a greater role than expected to make the PPII-like conformation the global minimum.

Table 2.1. Key geometric parameters of average collagen structure, and the PPII-like global minima of the Gly-Pro amide and alkene models.

Model	Ψ (°)	π* face	δ_{BD} (Å)	τ_{BD} (°)	Θ_{BD} (°)	NBO E2	rmsd (Å) 1K6F [c]
1K6F GPP[a]	+165		2.9	90			
1K6F PGP[a]	+176		3.1	90			
GPP *D*/*A***1**[b]	+160	*re*	3.1	93	2.9	0.41	
PGP *A***2**[b]	+167	-	3.2	89	0.2	0.19	
G(–F)=PP *D***3**	+162	-	3.3	95	-	< 0.1	0.14
PG(–F)=P *A***4**	+117	-	3.3	124	-	< 0.1	0.77
G(–Cl)=PP *D***5**	+155	*re*	3.3	99	2.9	0.39	0.19
PG(–Cl)=P *A***6**	+118	-	3.2	122	-	< 0.1	0.78
G(–H)=PP *D***7**	+163	-	3.4	97	-	< 0.1	0.21
PG(–H)=P *A***8**	+120	-	3.3	122	-	< 0.1	0.78

[a] Average Ψ angle, δ_{BD} and τ_{BD} from collagen crystal structure (PDB ID: 1K6F).[6] [b] Data for PPG *D***9** and GPP *D*/*A***1** from reference [16]. [c] Superposition rmsd of model minima with PDB ID: 1K6F.[6]

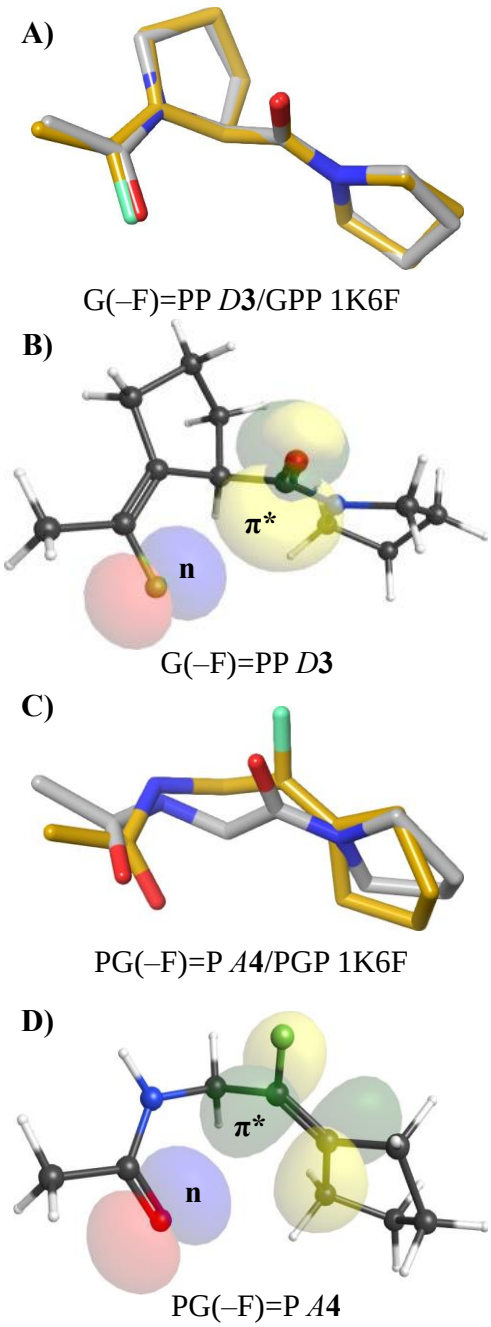


Figure 2.4. Gly-Pro fluoro-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** G(–F)=PP *D***3** (gold) superimposed on PDB ID: 1K6F (grey). **B)** G(–F)=PP *D***3** NBOs show no overlap between fluorine n orbital (blue/red) and C=O π* orbital (green/yellow). **C)** PG(–F)=P *A***4** (gold) superimposed on PDB ID: 1K6F (grey). **D)** PG(–F)=P *A***4** NBOs show no overlap between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

The PG(–F)=P *A***4** acceptor model global minimum is notably shifted from the PGP *A***2** PPII-like geometry (Figure 2.3B). Unrestrained optimization of *A***4** gives Ψ = +117º, which deviates significantly from the amide PGP *A***2** at Ψ = +167º (Table 2.1). This conformation does not superimpose well with the PGP 1K6F fragment,[6] which suggests that *A***2** is not a good fit for the PPII conformation (Figure 2.4C). However, since the global minimum is only about 2 kcal/mol lower in energy than the PPII conformation, and within the same energy well, the PPII conformation still appears to be thermally accessible. The global minimum of *A***4** has δ_{BD} = 3.3 Å and τ_{BD} = 124°, and an NBO E2 energy < 0.1 kcal/mol, all of which point to the absence of an n→π* interaction (Table 2.1, Figure 2.4D). Since the donor conformation is PPII-like, and the acceptor deviates, some compromise is expected to occur in the collagen triple helix.

2.5.1B Gly-Pro Chloro-Alkene

The Ψ-conformational energy profile of donor model G(–Cl)=PP *D***5** closely resembles that of the amide GPP *D/A***1** (Figure 2.3A). The unrestrained optimized geometry of *D***5** at its global minimum has Ψ = +155º (Table 2.1). The superposition of *D***5** with 1K6F shows very favorable alignment,[6] suggesting that the chloro-alkene could mimic the PPII conformation well (Figure 2.5A). *D***5** has Bürgi-Dunitz parameters δ_{BD} = 3.3 Å and τ_{BD} = 99°, which are within the geometric limits for a sizable n→π* interaction (Table 2.1). The presence of n→π* donation is supported by pyramidalization of C_{i+1}=O carbon with Θ_{BD} = 2.9° (Table 2.1). The NBO E2 energy of 0.39 kcal/mol (Table 2.1) supports the presence of an n→π* interaction, depicted with NBOs in Figure 2.5B.

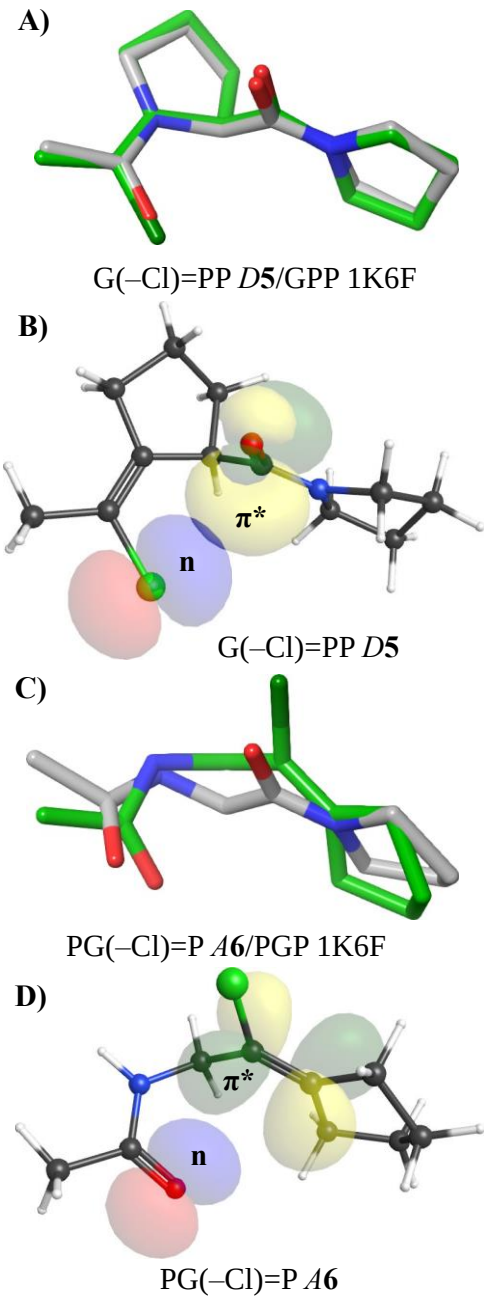


Figure 2.5. Gly-Pro chloro-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** G(–Cl)=PP *D***5** (green) superimposed on PDB ID: 1K6F (grey). **B)** G(–Cl)=PP *D***5** NBOs show overlap between chlorine n orbital (blue/red) and C=O π* orbital (green/yellow). **C)** PG(–Cl)=P *A***6** (green) superimposed on PDB ID: 1K6F (grey). **D)** PG(–Cl)=P *A***6** NBOs show no overlap between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

The corresponding acceptor model, PG(–Cl)=P *A***6**, does not have favorable PPII geometry when overlayed with the PGP portion of 1K6F (Figure 2.5C).[6] The global minimum of *A***6** exhibits Ψ = +118°, which diverges considerably from the PPII global minimum conformation of PGP *A***2** (Ψ = +167°, Table 2.1). Again, since the potential energy surface around the global minimum is very isotropic, and the *A***6** conformation is only about 2 kcal/mol lower in energy than that of the PPII conformation, the PPII conformation is still accessible to PG(–Cl)=P. Interestingly, *A***6** has δ_{BD} = 3.2 Å, which is within the expected distance for an n→π* interaction with a chlorine, and the Bürgi-Dunitz angle of τ_{BD} = 122° is only 13° higher than the optimal 109°, yet it appears to lack an n→π* interaction (Table 2.1). The NBO E2 energy is less than 0.1 kcal/mol (Table 2.1) suggesting no significant orbital interactions, as can be appreciated in the NBO images (Figure 2.5D). Besides the lack of favorable conformation, the Gly-Pro chloro-alkene mimic is expected to disrupt triple helix formation due to the long C–Cl bond, and the size of Cl.

2.5.1C Gly-Pro Proteo-Alkene

The torsional energy profile of donor model G(–H)=PP *D***7** has two low-energy minima at Ψ = +160° and Ψ = +100° (Figure 2.3A). Full optimizations starting from each geometry show the true global minimum lies at Ψ = +96°, while the PPII-like Ψ = +163° is a local minimum that is only 0.20 kcal/mol higher in energy, and with an estimated barrier between them of ΔE = 0.82 kcal/mol (Table S2.1). The PPII-like minimum of G(–H)=PP *D***7** can mimic the PPII conformation well, as evidenced by its superposition on GPP from 1K6F (Figure 2.6A).[6] However, the broad, shallow conformational energy profile around the global minimum helps to explain the known destabilization of the collagen-like triple helix we reported with a Gly-Pro proteo-alkene.[13] The alkene *D***7** minimum at Ψ = +163° has the Bürgi-Dunitz angle τ_{BD} = 97°, which is favorable for n→π* donation, although the distance δ_{BD} = 3.4 Å is highly unfavorable for a C–H---C=O

interaction (Table 2.1). The NBO image also shows no C–H σ→π* overlap (Figure 2.6B), which is verified by the lack of NBO E2 energy for the interaction (Table 2.1).

The energy profile of acceptor model PG(–H)=P *A***8** has minima at Ψ = +120º and Ψ = –120° (Figure 2.3B). Unrestrained optimization of both minima puts the global minimum at Ψ = +120º (Table 2.1) and a broad, shallow local minimum at Ψ = –119° only 0.2 kcal/mol higher in energy, with a low estimated barrier of ΔE = 1.2 kcal/mol between them (Table S2.1). The global minimum of *A***8** does not superimpose well with the PPII conformation of 1K6F,[6] again helping to explain the destabilization of collagen-like peptides by a Gly-Pro proteo-alkene mimic (Figure 2.6C).[13] The high flexibility of this model is not stabilized by an n→π* interaction based on its Bürgi-Dunitz parameters, NBO E2 energy, and NBO images (Table 2.1, Figure 2.6D).

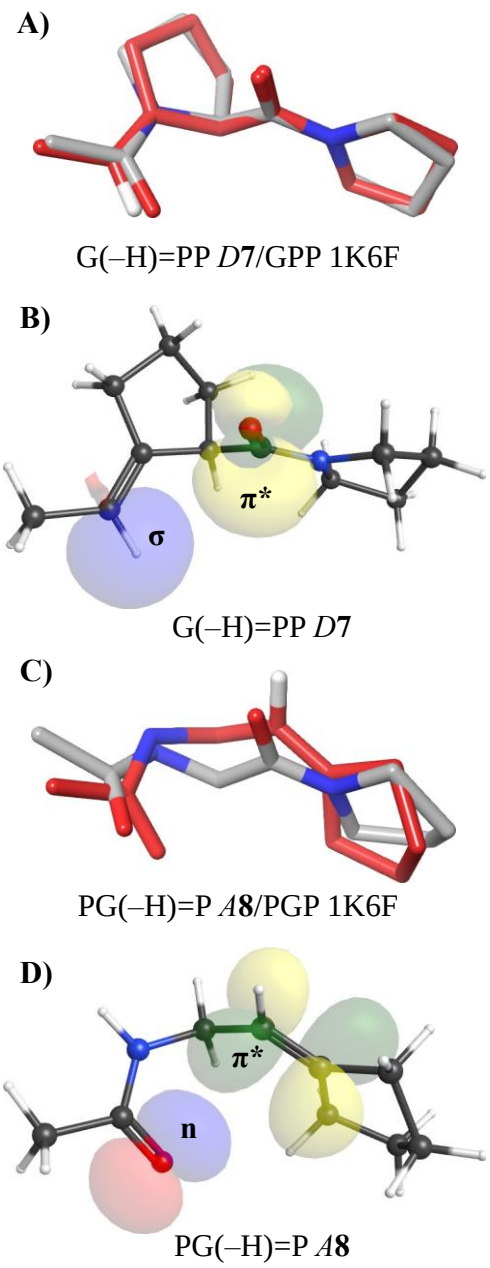


Figure 2.6. Gly-Pro proteo-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** G(–H)=PP *D***7** (red) superimposed on PDB ID: 1K6F (grey). **B)** G(–H)=PP *D***7** NBOs show no overlap between the C–H σ orbital (blue/red) and C=O π* orbital (green/yellow). **C)** PG(–H)=P *A***8** (red) superimposed on PDB ID: 1K6F (grey). **D)** PG(–H)=P *A***8** NBOs show no overlap between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

2.5.2 Pro-Pro Alkene Isosteres

Although fluorine and chlorine atoms have been known to establish hydrogen bonding interactions,[42, 43] substitution of the Pro-Pro amide may prove problematic due to weaker C–X···H–N compared to C=O···H–N interstrand hydrogen bonds that strengthens the interactions between the three strands of the collagen triple helix (Figure S2.4).[33] We include the following Pro-Pro model conformational analyses to see how well the fluoro- or chloro-alkenes might mimic the PPII conformation, and to help explain the failure of the Pro-Pro proteo-alkene isostere to stabilize the collagen-like triple helix beyond the missing interstrand hydrogen bond.[14] The Ψ dihedral angle coordinate scans for P(–X)=PG donor models (Figure 2.7A) and GP(–X)=P acceptor models (Figure 2.7B) are overlaid with their respective PPG *D***9** and GPP *D*/*A***1** amide coordinate scans at the MP2/6-31+G(d) level of theory.[16] A complete list of geometric and Bürgi-Dunitz parameters for all minima and maxima calculated at the MP2/6-311+G(2d,p) level can be found in Supplemental Information Table S2.2.

The PPII-like amide minimum of PPG *D***9** is not the global minimum, which indicates that other long-range forces, such as Pro C=O···H–N Gly interstrand hydrogen bonds, contribute to the stabilization of the collagen triple helix, even for native amides.[16] The PPII-like minima for all models in the P(–X)=PG donor series are also local, not global, minima that are approximately 1 – 2 kcal/mol higher in energy than their respective global minima (Figure 2.7A). These PPII-like minima are discussed in detail below.

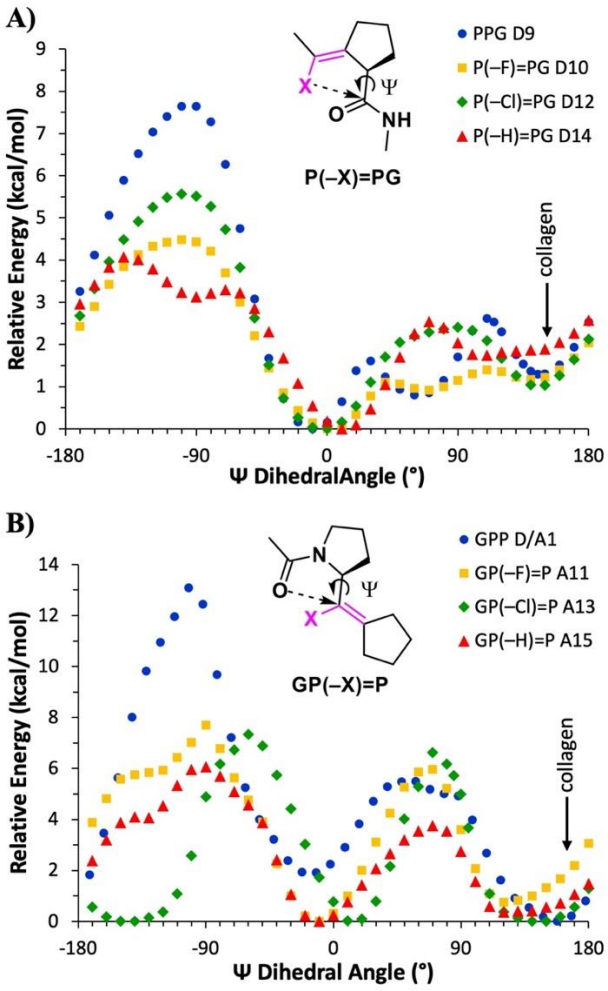


Figure 2.7. Potential energy scans of the Ψ dihedral angle in models of Pro-Pro amide substitutions where X = F, Cl, or H. **A)** P(–X)=PG n→π* donor, and **B)** GP(–X)=P n→π* acceptor models overlaid with their respective PPII-like amide model energies as a function of Ψ.[16] The global minimum for each model was normalized to 0.0 kcal/mol. The average collagen Ψ dihedral angle from PDB ID: 1K6F in Table S2.3 is labeled.[6]

All three alkene models were found to have their global minima between Ψ = –15° and Ψ = +15° (Figure 2.7A). This geometry includes an alkene C_i=C_{i+1} π···H–N_{i+2} intramolecular hydrogen bond (Table S2.2, Figure S2.5). Intramolecular alkene π···H–N hydrogen bonding has been found experimentally.[44] This minimum is also reported for the amide PPG ***D*9** as an N_{i+1}···H–N_{i+2} hydrogen bond (Table S2.2, Figure S2.5D).[16]

Of the alkene models, only P(–F)=PG *D***10** has a local minimum at Ψ = +64°, which is about 1 kcal/mol higher in energy than the global minimum (Figure 2.7A). This conformation features an inverse γ-turn with an intramolecular 7-membered ring closed by a C_i–F···H–N_{i+1} hydrogen bond (Figure S2.6A). A similar C_i=O···H–N_{i+1} hydrogen bond is found for PPG *D***9** at Ψ = +62° (Figure S2.6B).[16] Such an interaction is unique to the fluoro-alkene mimic; of course, analogous hydrogen bonds are not seen in either chloro- or proteo-alkene models.

For these Pro-Pro donor alkene models, the global maxima at Ψ = –100°, resulting from an eclipsing interaction between the X atom (or oxygen for the amide) and a proton on the 5-membered ring, are 4 to 6 kcal/mol higher in energy than the global minima. The alkene hydrogen of P(–H)=PG D**14** is too small to be sterically hindered in this way, and instead has an eclipsing interaction between the N_{i+1}–H and the 5-membered ring around Ψ = –140° (Figure 2.7A).

The GP(–X)=P acceptor model energy scans are comparable to the native amide GPP *D*/*A***1** scan, which has an energy difference of 1.7 kcal/mol between its PPII-like global minimum at Ψ= +160° and a local minimum around Ψ = –20°, with a large 5.1 kcal/mol barrier between them (Table S2.2).[16] The PPII-like local minima for all GP(–X)=P acceptor models are between 0.5 – 1 kcal/mol less stable than their respective global minimum (Figure 2.7B).

The energy scans of the GP(–F)=P *A***11** and GP(–H)=P *A***15** acceptor models are conformationally similar (Figure 2.7B). The global minima of both models are sterically favored conformations around Ψ = –10° with no H-bond or n→π* interactions (Figures 5.7A, 5.7B). In a collagen-like peptide, the GP(–F)=P alkene mimic is expected to have similar challenges as the previously synthesized GP(–H)=P mimic.[14] The GP(–Cl)=P *A***13** global minimum is located around Ψ = +10° and is stabilized by a weak n→π* interaction (Figure S2.7C). For all GP(–X)=P

acceptor models, the maxima around Ψ = +70° are due to eclipsing interactions between the X atom (or oxygen for amide) and the 5-membered ring (Figure 2.7B).

2.5.2A Pro-Pro Fluoro-Alkene

Unrestrained optimization of the most PPII-like minimum of donor model P(–F)=GP *D***10** gives Ψ = +142º, which is 1.4 kcal/mol above the global minimum at Ψ = +1º (Table S2.2). This torsion angle is relatively close to that of the PPG *D***9** PPII-like minimum at Ψ = +150º, which is similarly 1.0 kcal/mol above the global minimum (Table S2.2). The *D***10** conformation superimposes reasonably well with the PPG segment of 1K6F,[6] which suggests that a Pro-Pro fluoro-alkene mimic could be geometrically compatible with the triple helix (Figure 2.8A). Additionally, *D***10** has ideal Bürgi-Dunitz parameters with δ_{BD} = 3.1 Å and τ_{BD} = 104° (Table S2.2). Our calculations show pyramidalization of the C_{i+1}=O carbon by Θ_{BD} = 2.2°, and an NBO E2 energy of 0.16 kcal/mol. The weak n→π* interaction can be visualized in Figure 2.8B.

The fluoro-alkene acceptor model, GP(–F)=P *A***11,** has an unfavorable energy landscape for a PPII mimic (Figure 2.7B). Unrestrained optimization of its most PPII-like minima gives Ψ = +123° at 1.0 kcal/mol above the global minimum (Table 2.2). The significant difference compared to the collagen GPP *D*/*A***1** geometry at Ψ = +160º, and its inadequate superposition with 1K6F, may preclude it as a PPII mimic (Figure 2.8C). However, the *A***11** conformation is within the same energy well, only 1.3 kcal/mol lower in energy, as the PPII-like conformation at Ψ = +160° (Figure 2.7B). Interestingly, *A***11** has good Bürgi-Dunitz parameters with δ_{BD} = 3.2 Å and τ_{BD} = 118° (Table 2.2); however, NBO images and E2 energies agree on the absence of strong orbital overlap (Figure 2.8D, Table 2.2). Whether a Pro-Pro fluoro-alkene mimic would provide a sufficient interstrand H-bond acceptor remains an open question.

Table 2.2. Key geometric parameters of the collagen structure, and the PPII-like minima of Pro-Pro amide and alkene models.

Model	Ψ (°)	ΔE[c]	π* face	δ_{BD} (Å)	τ_{BD} (°)	Θ_{BD} (°)	NBO E2	rmsd (Å) 1K6F[d]
1K6F PPG[a]	+152			3.1	81			
1K6F GPP[a]	+165			2.9	90			
PPG *D***9**[b]	+150	1.0	*re*	3.0	98	3.0	0.46	
GPP *D*/*A***1**[b]	+160	0.0	*re*	3.1	93	2.9	0.41	
P(–F)=PG *D***10**	+142	1.4	*re*	3.1	104	2.2	0.16	0.32
GP(–F)=P *A***11**	+123	1.0	-	3.2	118	-	< 0.1	0.54
P(–Cl)=PG *D***12**	+147	1.2	*re*	3.3	104	2.6	0.45	0.35
GP(–Cl)=P *A***13**	+145	0.5	*re*	3.1	99	1.6	0.20	0.25
P(–H)=PG *D***14**	+101	1.8	-	3.0	130	-	< 0.1	0.56
GP(–H)=P *A***15**	+120	0.6	-	3.3	123	-	< 0.1	0.53

[a] Average Ψ angle, δ_{BD} and τ_{BD} from collagen crystal structure (PDB ID: 1K6F).[6] [b] Data for PPG *D***9** and GPP *D*/*A***1** from reference [16]. [c] Energy relative to the respective global minimum (kcal/mol). [d] Superposition rmsd of model minima with PDB ID: 1K6F.[6]

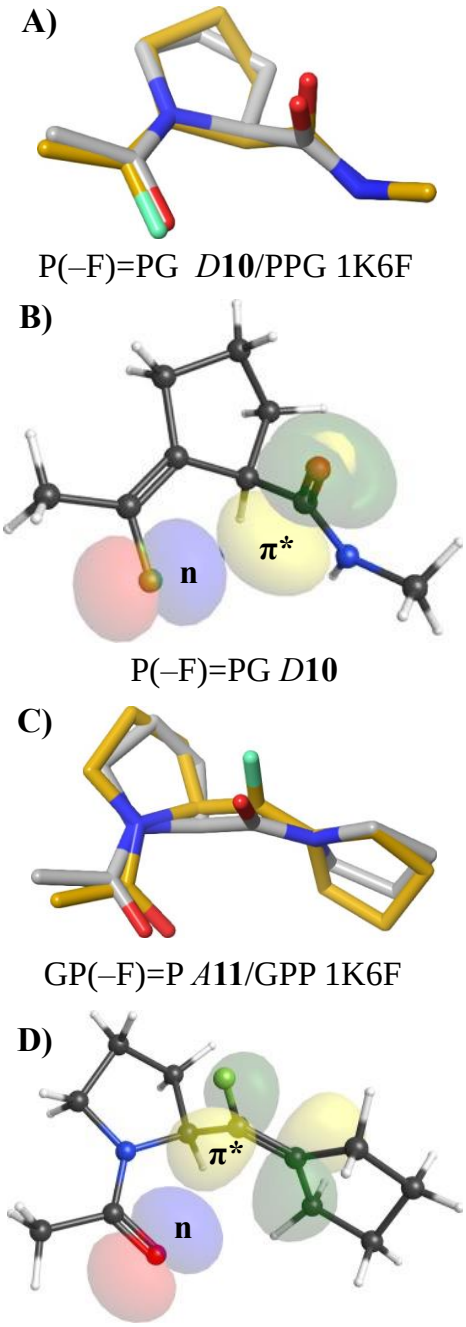


Figure 2.8. Pro-Pro fluoro-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** P(–F)=PG *D***10** (gold) superimposed on PDB ID: 1K6F (grey). **B)** P(–F)=PG *D***10** NBOs show limited overlap between fluorine n orbital (blue/red) and C=O π* orbital (green/yellow). **C)** GP(–F)=P *A***11** (gold) superimposed on PDB ID: 1K6F (grey). **D)** GP(–F)=P *A***11** NBOs show no overlap between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

2.5.2B Pro-Pro Chloro-Alkene

Unrestrained optimization of the PPII-like minimum of the chloro donor model P(–Cl)=PG *D***12** converges to Ψ = +147° (Table 2.2). This structure is geometrically very similar to amide model PPG *D***9** with Ψ = +150°, and it shows reasonably good superposition on the PPII segment of 1K6F (Figure 2.9A).[6] *D***12** also has good Bürgi-Dunitz parameters with δ_{BD} = 3.3 Å and τ_{BD} = 104° (Table 2.2). The NBO E2 energy (0.45 kcal/mol) confirms an n→π* interaction, which can be visualized in Figure 2.9B. The presence of n→π* donation is further supported by the pyramidalization of the C_{i+1}=O carbon, Θ_{BD} = 2.6° (Table 2.2).

The coordinate scan of the GP(–Cl)=P *A***13** acceptor model has two minima at Ψ = +140° and Ψ = –140° (Figure 2.7B). Unrestrained optimization of both structures converged to the same point at Ψ = +145°, which is not far from the PPII-like GPP *D/A***1** conformation at Ψ = +160° (Table 2.2). Model *A***13** also shows good superposition with the PPII in 1K6F, so this mimic is expected to mimic the PPII conformation well (Figure 2.9C).[6] The fully optimized model has δ_{BD} = 3.1 Å and τ_{BD} = 99°, which are ideal parameters for n→π* donation to occur (Table 2.2). Remarkably, an NBO E2 energy of 0.20 kcal/mol and carbon pyramidalization Θ_{BD} = 1.65° suggest this chloroalkene can also act as an n→π* acceptor (Table 2.2). This is further supported by the NBO image (Figure 2.9D). What is perhaps most interesting about this n→π* interaction is that it occurs despite repulsion between the oxygen lone pair and the π-bonding orbital (Figure S2.7D).

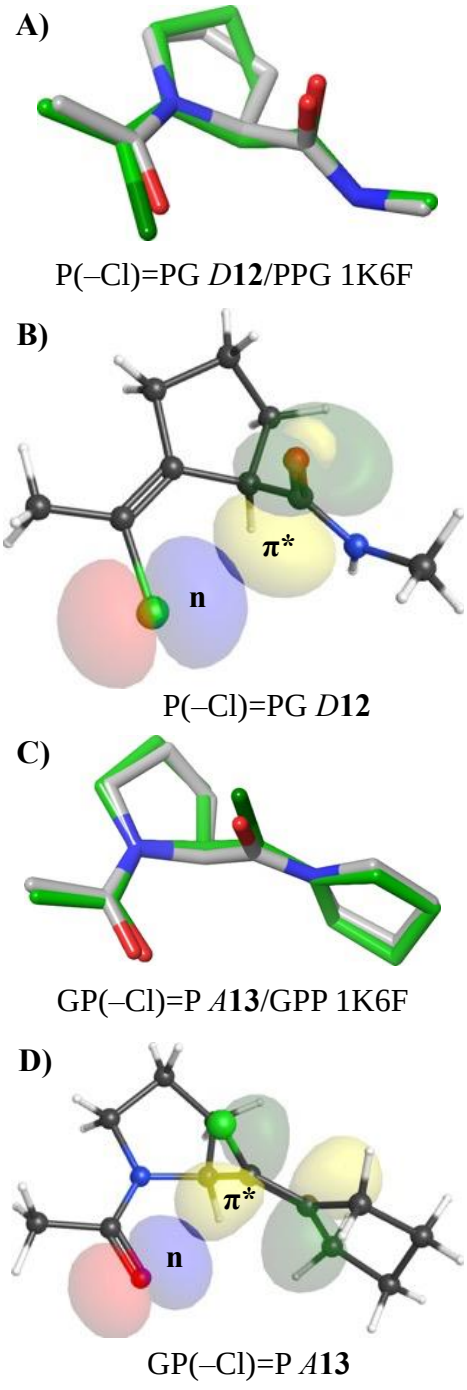


Figure 2.9. Pro-Pro chloro-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** P(–Cl)=PG *D***12** (green) superimposed on PDB ID: 1K6F (grey). **B)** P(–Cl)=PG *D***12** NBOs show strong overlap between chlorine n orbital (blue/red) and C=O π* orbital (green/yellow). **C)** GP(–Cl)=P *A***13** (green) superimposed on PDB ID: 1K6F (grey). **D)** GP(–Cl)=P *A***13** NBOs show an n→π* interaction between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

Although the Pro-Pro chloro-alkene models mimic the PPII conformation very well, we expect that the chloroalkene would destabilize collagen-like peptides because of the long C–Cl bond, the size of the Cl, and the need for a triple-helix interstrand hydrogen-bond acceptor atom at this position. However, this chloroalkene might serve to stabilize a PPII helix that is not part of a triple helix.

2.5.2C Pro-Pro Proteo-Alkene

Unrestrained optimization of donor model P(–H)=PG *D***14** results in a broad minimum at Ψ = +101°, which is 1.8 kcal/mol above its global minimum at Ψ = +12° (Table 2.2). This differs considerably from the PPII-like minimum of PPG *D***9** at Ψ = +150° that is 1.0 kcal/mol above the global minimum (Table 2.2). Model *D***14** does not overlap well with the PPII segment in 1K6F (Figure 2.10A).[6] The geometric parameters, δ_{BD} = 3.0 Å and τ_{BD} = 130°, are outside the ideal Bürgi-Dunitz threshold for hydrogen (Table 2.2). Both NBO E2 energy and NBO images indicate no significant overlap of the C-H σ and C=O π* orbitals (Figure 2.10B).

Acceptor model GP(–H)=P *A***15** also exhibits an energetic landscape unfavorable to PPII stabilization (Figure 2.7B). The unrestrained optimized structure of its global minimum gives Ψ = +120° compared to the amide GPP *D/A***1** with Ψ = +160° (Table 2). Model *A***15** also shows very poor superposition on GPP from 1K6F (Figure 2.10C).[6] The Bürgi-Dunitz parameters of δ_{BD} = 3.3 Å and τ_{BD} = 123° are inadequate for a strong n→π* interaction (Table 2.2). The NBO E2 energy of less than 0.1 kcal/mol and the NBO images show no significant n→π* overlap (Figure 2.10D).

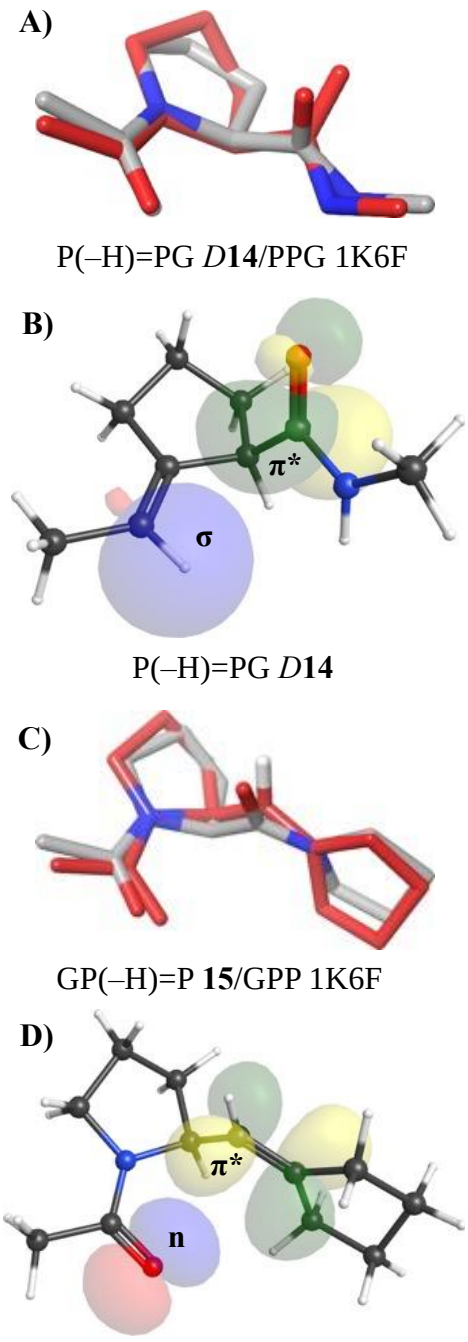


Figure 2.10. Pro-Pro proteo-alkene models calculated at the MP2/6-311+G(2d,p) level. **A)** P(–H)=PG *D***14** (red) superimposed on PDB ID: 1K6F (grey). **B)** P(–H)=PG *D***14** NBOs show no interaction between the C–H σ orbital (blue/red) and C=O π* orbital (green/yellow). **C)** GP(–H)=P *A***15** (red) superimposed on PDB ID: 1K6F (grey). **D)** GP(–H)=P *A***15** NBOs show no interaction between oxygen n orbital (blue/red) and C=C π* orbital (green/yellow).

The shallow conformational energy surfaces for both donor and acceptor Pro-Pro proteo-alkene models, the large differences in Ψ torsion angles from the PPII-helix, poor overlap with a collagen-like peptide, and the lack of stabilizing non-covalent interactions, help to explain the experimental destabilizing effect of our Pro-Pro proteo-alkene mimic in a collagen-like peptide.[14]

2.6 Conclusions

Halo-alkene isosteres are interesting as potential collagen mimics, since their global minima or low-lying local minima are near the PPII-like conformation in our models. Both Gly-Pro fluoro- and chloro-alkene n→π* donor models are stable in the conformation that is most similar to the PPII conformation. An n→π* interaction is found in the Gly-Pro donor model G(–Cl)=PP *D***5**, as well as Pro-Pro donor models P(–F)=PG *D***10**, P(–Cl)=PG *D***11**, and acceptor model GP(–Cl)=P *A***13**. The halo-alkenes serve well as n→π* donors, but the conformations are altered in the models in which the halo-alkenes serve as n→π* acceptors, shifting the lowest energy conformation to about Ψ = +120°, instead of the PPII-like Ψ = +145° to +160°. The Pro-Pro chloro-alkene mimic is an exception because it captures the native amide conformations in both the donor and acceptor models. In situations where acceptor minima are different from PPII-like minima, the donor interactions may provide counterbalance, suggesting they could serve as conformationally restricted isosteres for collagen. The lack of n→π* interactions in some of the PPII-like minima of the models that actually mimic PPII well suggests that other non-bonding interactions may be greater forces than expected.

All n→π* interactions are missing entirely from models of our previously reported Gly-Pro and Pro-Pro proteo-alkene mimics in collagen-like peptides.[13, 14] In addition, interactions with other strands within a triple helix are not captured in our models, which may provide additional

stabilizing or destabilizing forces. Further computational and/or experimental studies are necessary to fully answer these questions.

If the polarity and size of the amide bond dominates triple-helix folding, then we expect the fluoro-alkene to be the better mimic. However, if an n→π* interaction is critical for folding, we expect the chloro-alkene to be a better mimic. Further, we predict that substitution at the Gly-Pro position will provide more stability than substitution at the Pro-Pro position, because chlorine and fluorine are unlikely to be good interstrand hydrogen-bond acceptors in the collagen triple helix. The Gly-Pro haloalkene mimics are currently under experimental investigation.

2.7 Supplementary Tables

Table S2.1. Relative energies and key geometric parameters of alkene Gly-Pro models. Minima (blue), collagen-like minima (green) and maxima (white). Models labeled with D are potential n→π* donors, and A are potential n→π* acceptors.

Model	Ψ [a] (°)	Φ [a] (°)	ΔE[a] (kcal/mol)	π* face/ H-bond	δ_{BD} /d_{HB} (Å)	τ_{BD}/∠$_{HB}$ (°)	Θ_{BD} (°)	NBO E2[c] (kcal/mol)
G(–F)=PP *D***3**								
min *D***3a**	–18	–71	2.0	-	3.23	121		
TS *D***3b**	+39	–111	4.3	-	3.84	151		
min *D***3c**	+162	–79	0.0	-	3.26	95		< 0.1
PG(–F)=P *A***4**								
max *A***4a**	–175	–78	3.1	-	3.30	74		
"	"	"	"	*si*[b]	3.02	88	0.50	0.26[f]
min *A***4b**	–59	–63	2.0	*si*	3.00	86	1.7	0.48
min *A***4c**	+117	–80	0.0	-	3.29	124		< 0.1
G(–Cl)=PP *D***5**								
min *D***5a**	–19	–69	1.6	*si*	3.34	116	0.46	0.65
max *D***5b**	+44	–117	5.2	-	4.19	151		
min *D***5c**	+155	–75	0.0	*re*	3.36	99	2.9	0.39
PG(–Cl)=P *A***6**								
max *A***6a**	–171	–84	2.8	-	3.45	79		
"	"	"	"	*si*[b]	3.29	92	0.85	0.56[f]
min *A***6b**	–61	–68	2.7	*si*	3.08	87	2.7	0.38
min *A***6c**	+118	–75	0.0	-	3.23	122		< 0.1
G(–H)=PP *D***7**								
min *D***7a**	–57	–25	1.9	*si*	2.65	110	0.51	0.55
max *D***7b**	+38	–113	5.2	-	3.77	151		
min *D***7c**	+96	–42	0.0	-	2.70	118		
min *D***7d**	+163	–90	0.2	-	3.40	97		< 0.1
PG(–H)=P *A***8**								
max *A***8a**	–180	–77	1.2	-	3.28	77		
min *A***8b**	–119	–85	0.2	-	3.41	62		
min *A***8c**	+120	–80	0.0	-	3.32	122		< 0.1

[a]Unconstrained Ψ and Φ angles, and minima (min) or maxima (max) energies relative to the lowest minimum. Min structures were optimized at MP2/6-311+G(2d,p). Max energies are the difference between MP2/6-311+G(2d,p) single point energies of global min and max structures optimized at MP2/6-31+G(d). [b]Reverse $O_{i+1}\cdots C_i$=O n→π* interaction. [c]Sum of n→π* or n→σ* NBO E2 orbital interactions.

Table S2.2. Relative energies and key geometric parameters of alkene Pro-Pro models. Minima (blue), PPII-like minima (green) and maxima (white). Models labeled with D are potential n→π* donors. Models labeled with A are potential n→π* acceptors.

Model	Ψ [a] (°)	Φ [a] (°)	ΔE[a] (kcal/mol)	π* face/ H-bond	δ_{BD} /d_{HB} (Å)	τ_{BD}/$\angle_{HB}$ (°)	Θ_{BD} (°)	NBO E2[c] (kcal/mol)
P(–F)=PG ***D*10**								
min ***D*10a**	+1	–81	0.0	C=C$_\pi$···H–N$_{i+1}$	3.37	136		1.34
max ***D*10b**	+44	–86	1.1	-	3.46	169		
min ***D*10c**	+64	–84	1.1	F···H–N$_{i+1}$	3.41	172		3.67
min ***D*10d**	+142	–70	1.4	*re*	3.11	104	2.2	0.16
GP(–F)=P ***A*11**								
min ***A*11a**	–12	–79	0.0	*si*	3.21	121		
max ***A*11b**	+67	–115	6.1	-	3.77	154		
min ***A*11c**	+123	–77	1.0	*re*	3.16	118	1.2	< 0.1
P(–Cl)=PG ***D*12**								
min ***D*12a**	–5	–78	0.0	C=C$_\pi$···H–N$_{i+1}$	3.50	128		0.97
max ***D*12b**	+90	–88	2.2	-	3.70	154		
min ***D*12c**	+147	–72	1.2	*re*	3.32	104	2.6	0.45
GP(–Cl)=P ***A*13**								
min ***A*13a**	–14	–78	0.0	*si*	3.20	120	0.24	< 0.1
max ***A*13b**	+58	–123	7.1	-	3.86	154		
min ***A*13c**	+145	–70	0.5	*re*	3.08	99	1.7	0.20
P(–H)=PG **14**								
min ***D*14a**	+12	–87	0.0	C=C$_\pi$···H–N$_{i+1}$	3.42	147		2.05
max ***D*14b**	+73	–87	2.4	-	3.40	162		
min ***D*14c**	+101	–68	1.8	-	3.04	130		< 0.1
GP(–H)=P ***A*15**								
min ***A*15a**	–12	–80	0.0	-	3.29	124		
max ***A*15b**	+72	–117	3.9	-	3.83	155		
min ***A*15c**	+120	–80	0.6	-	3.25	123		< 0.1

[a]Unconstrained Ψ and Φ angles, and minima (min) or maxima (max) energies relative to the lowest minimum. Min structures were optimized at MP2/6-311+G(2d,p). Max energies are the difference between MP2/6-311+G(2d,p) single point energies of global min and max structures optimized at MP2/6-31+G(d). [c]Sum of n→π* or n→σ* NBO E2 orbital interactions.

Table S2.3. Average phi (Φ) and psi (Ψ) angles for PDB ID: 1K6F collagen crystal structure divided by residue and residue position. Pro-Gly-Pro (left), Gly-Pro-Pro (middle), and Pro-Pro-Gly (right).[45]

Gly (PGP)	Ψ (°)	Φ (°)	Pro1 (GPP)	Ψ (°)	Φ (°)	Pro2 (PPG)	Ψ (°)	Φ (°)
Residue Position			Residue Position			Residue Position		
9	179 ± 2	–74 ± 3	10	164 ± 3	–76 ± 2	8	153 ± 2	–60 ± 2
12	175 ± 4	–70 ± 4	13	165 ± 3	–76 ± 3	11	152 ± 2	–60 ± 1
15	177 ± 2	–73 ± 2	16	165 ± 4	–74 ± 2	14	152 ± 3	–60 ± 3
18	174 ± 5	–71 ± 4	19	164 ± 2	–74 ± 2	17	153 ± 1	–61 ± 2
21	176 ± 1	–73 ± 3	22	166 ± 3	–75 ± 2	20	152 ± 4	–60 ± 2
Average			Average			Average		
	176 ± 4	–72 ± 4		165 ± 3	–75 ± 3		152 ± 3	–60 ± 2

Table S2.4. Average Bürgi-Dunitz $C_i{=}O\cdots C_{i+1}{=}O$ distance (δBD) and $\angle C_i{=}O\cdots C_{i+1}{=}O$ (τBD) for PDB ID: 1K6F collagen crystal structure divided by residue and residue position. Pro-Gly-Pro (left), Gly-Pro-Pro (middle), and Pro-Pro-Gly (right).[45]

Gly (PGP)	δ_{BD} (Å)	τ_{BD} (°)	Pro1 (GPP)	δ_{BD} (Å)	τ_{BD} (°)	Pro2 (PPG)	δ_{BD} (Å)	τ_{BD} (°)
Residue Position			Residue Position			Residue Position		
9	3.15 ± 0.04	90 ± 1	10	2.87 ± 0.04	91 ± 1	8	3.12 ± 0.03	82 ± 2
12	3.14 ± 0.04	90 ± 2	13	2.90 ± 0.04	90 ± 1	11	3.16 ± 0.05	81 ± 1
15	3.12 ± 0.03	90 ± 1	16	2.90 ± 0.04	90 ± 2	14	3.12 ± 0.03	81 ± 2
18	3.14 ± 0.05	90 ± 1	19	2.91 ± 0.04	90 ± 1	17	3.15 ± 0.03	80 ± 2
21	3.11 ± 0.03	89 ± 1	22	2.92 ± 0.02	91 ± 3	20	3.13 ± 0.06	82 ± 2
Average			Average			Average		
	3.13 ± 0.04	90 ± 1		2.90 ± 0.04	90 ± 2		3.14 ± 0.04	81 ± 2

2.8 Supplementary Figures

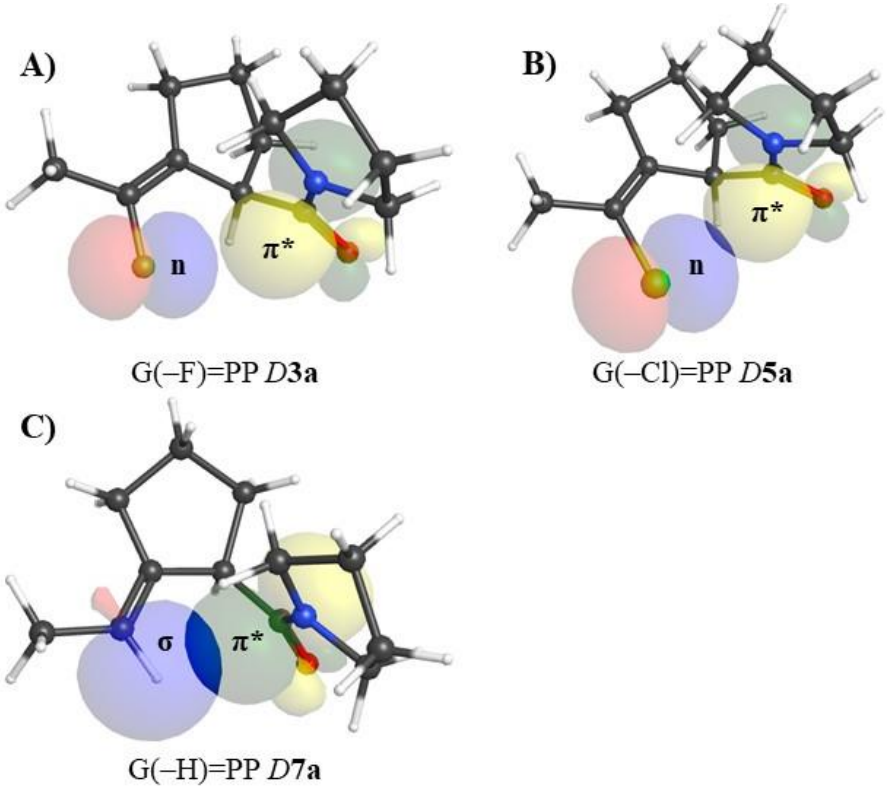


Figure S2.1. NBO images of representative orbitals depicting local minima of the G(–X)=PP donor series calculated at the MP2/6-311+G(2d,p) level. **A)** G(–F)=PP min *D***3a** at Ψ = –18° shows no orbital overlap of fluorine non-bonding orbital (blue/red) and C=O antibonding *si*-π* orbital (green/yellow). **B)** G(–Cl)=PP min *D***5a** at Ψ = –19° shows an n→*si*-π* interaction between chlorine non-bonding orbital (blue/red) and C=O antibonding π* orbital (green/yellow). **C)** G(–H)=PP min *D***7a** at Ψ = –57° shows a σ→*si*-π* interaction between the C–H hydrogen σ orbital (blue/red) and C=O π* orbital (green/yellow).

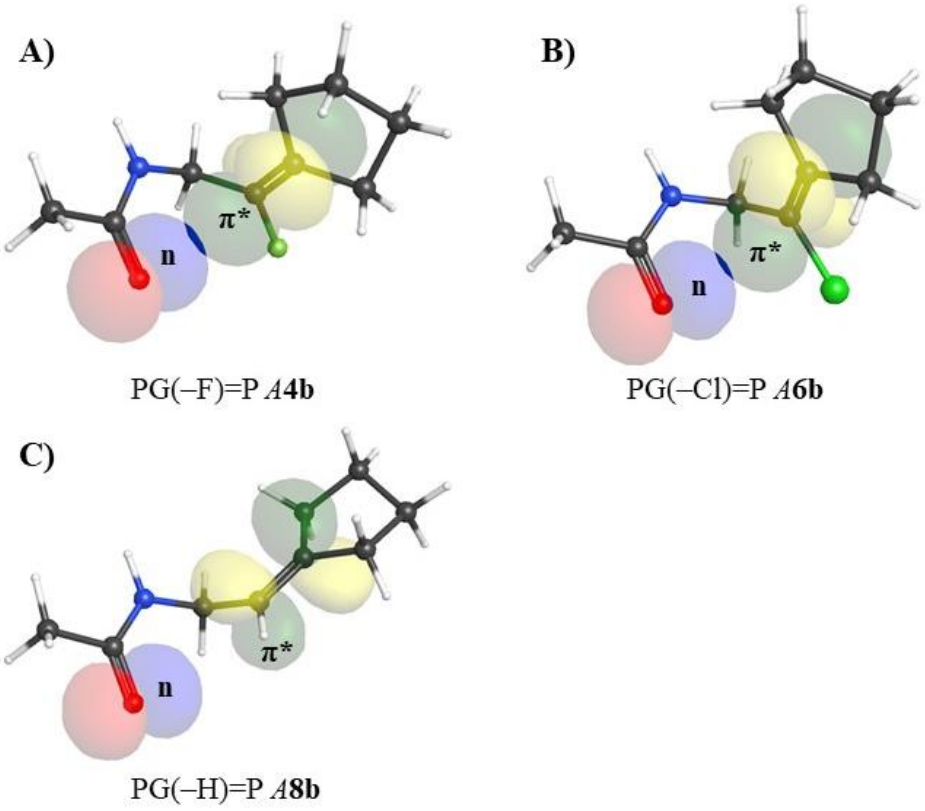


Figure S2.2. NBO images of representative orbitals depicting local minima of the PG(–X)=P acceptor series calculated at the MP2/6-311+G(2d,p) level. **A)** PG(–F)=P min *A***4b** at Ψ = –59° shows an n→*si*-π* interaction between oxygen non-bonding orbital (blue/red) and C=C antibonding π* orbital (green/yellow). **B)** PG(–Cl)=P min *A***6b** at Ψ = –61° shows an n→*si*-π* interaction between oxygen non-bonding orbital (blue/red) and C=C antibonding π* orbital (green/yellow). **C)** PG(–H)=P min *A***8b** at Ψ = –121° shows no significant orbital interactions.

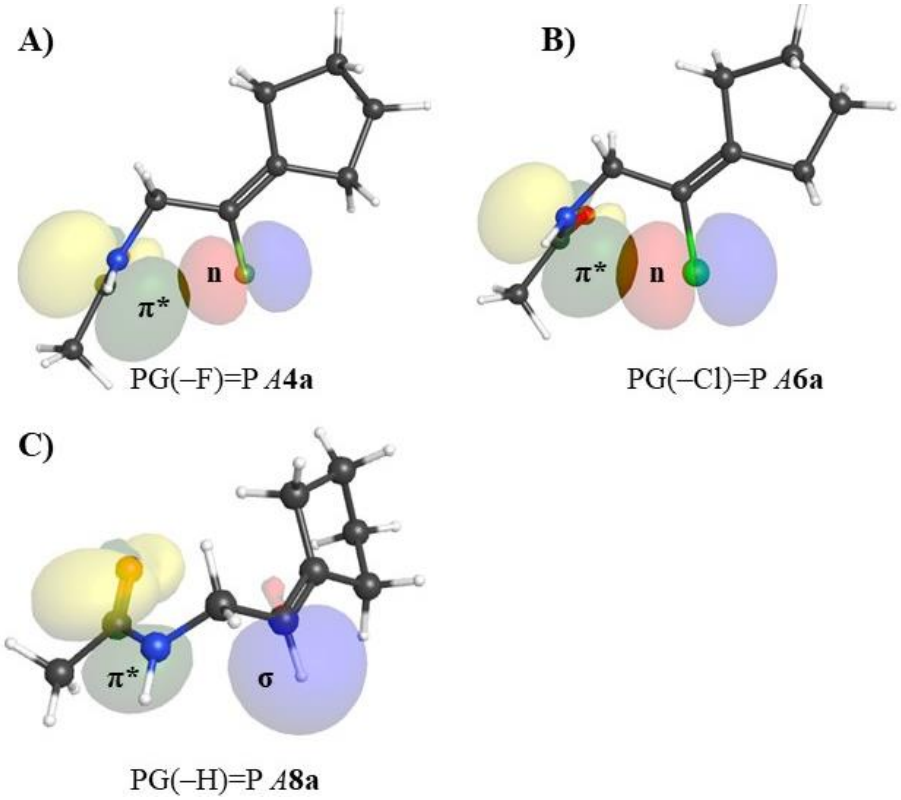


Figure S2.3. NBO images of representative orbitals depicting local minima of the PG(–X)=P acceptor series calculated at the MP2/6-311+G(2d,p) level. **A)** PG(–F)=P TS *A***4a** at Ψ = –175° shows a reverse n→*si*-π* interaction between fluorine non-bonding orbital (blue/red) and C_i=O antibonding π* orbital (green/yellow). **B)** PG(–Cl)=P TS *A***6a** at Ψ = –171° shows a reverse n→*si*-π* interaction between oxygen non-bonding orbital (blue/red) and C=C antibonding π* orbital (green/yellow). **C)** PG(–H)=P TS *A***8a** at Ψ = –180° shows no significant non-covalent interactions.

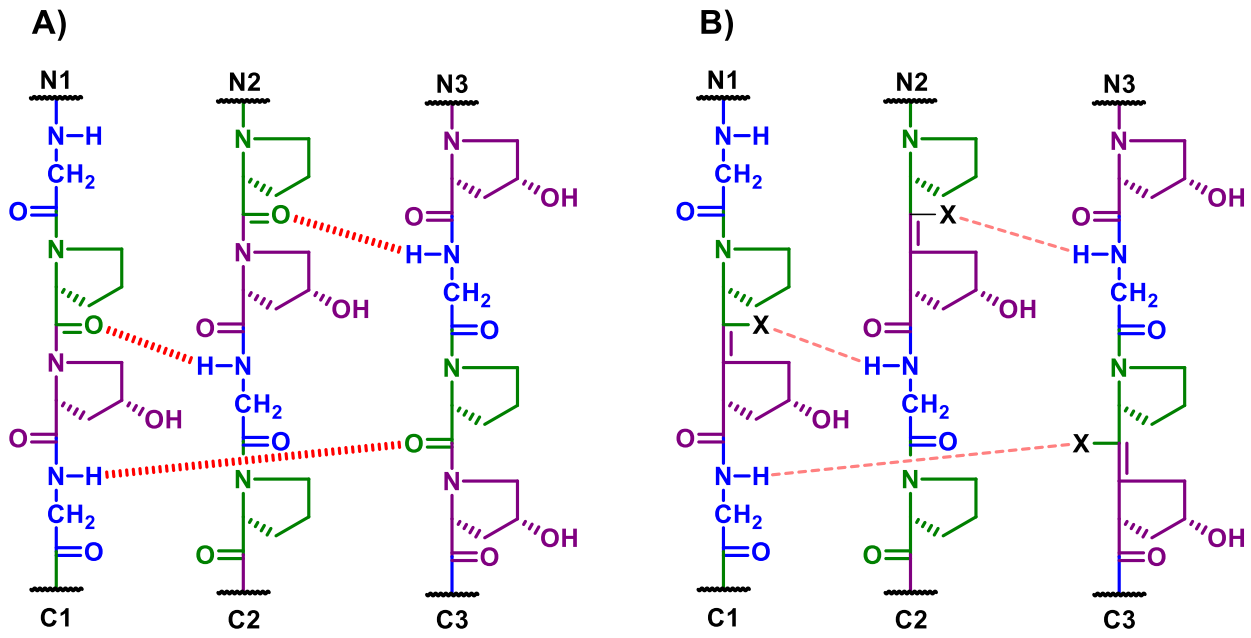


Figure S2.4. Interstrand hydrogen bonding between the three PPII helices of collagen.[46, 47] **A)** Three Gly N–H to Pro C=O interstrand hydrogen bonds. **B)** Replacement of a Pro-Hyp amide with a halo-alkene isostere is expected to significantly weaken interstrand Gly N–H to pseudo-Pro C–X hydrogen bonding.

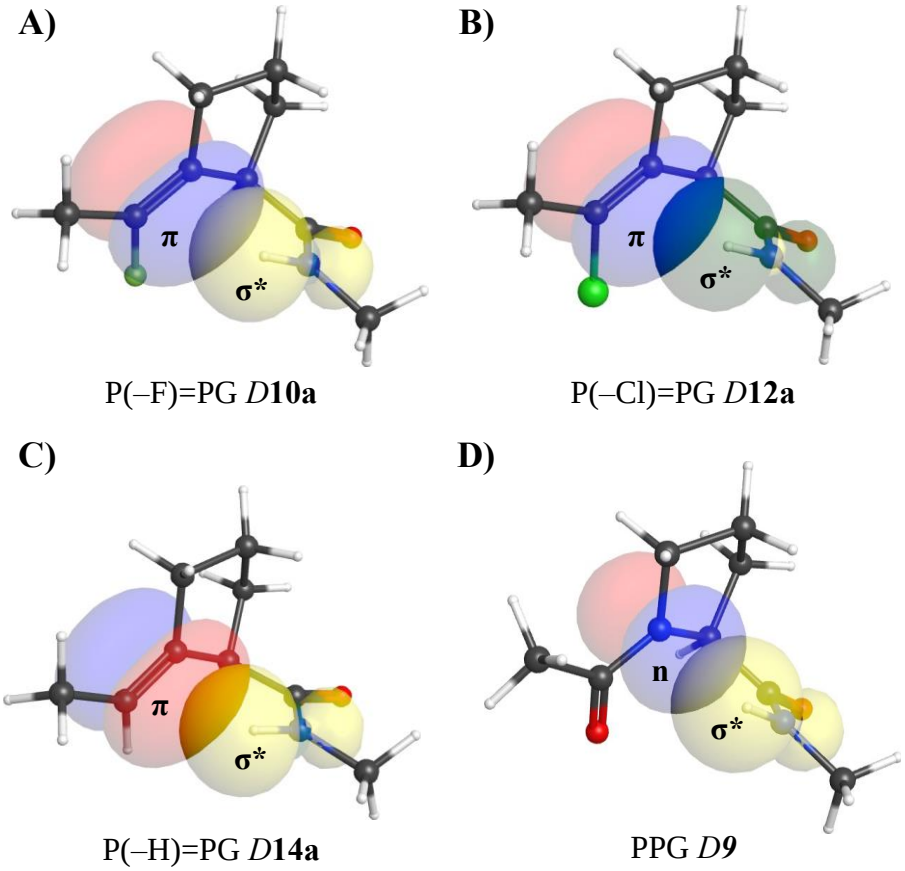


Figure S2.5. NBO pictures of representative orbitals depicting intramolecular C_i=C_{i+1} π ⋯ H–N_{i+2} hydrogen bonding of the P(–X)=PG donor series calculated at the MP2/6-311+G(2d,p) level. **A)** P(–F)=PG min *D***10a** at Ψ = +1° showing overlap between C_i=C_{i+1} bonding π orbital (red/blue) and H–N antibonding σ* orbital (green/yellow). **B)** P(–Cl)=PG min *D***12a** at Ψ = –5° showing overlap between C_i=C_{i+1} bonding π orbital (red/blue) and H–N antibonding σ* orbital (green/yellow). **C)** P(–H)=PG min *D***14a** Ψ = +12° showing overlap between C_i=C_{i+1} bonding π orbital (red/blue) and H–N antibonding σ* orbital (green/yellow). **D)** PPG min *D***9** at Ψ = –12° with N_{i+1} ⋯ H–N_{i+2} hydrogen bond showing overlap between N_{i+1} non-bonding orbital (red/blue) and H–N antibonding σ* orbital (green/yellow).[48]

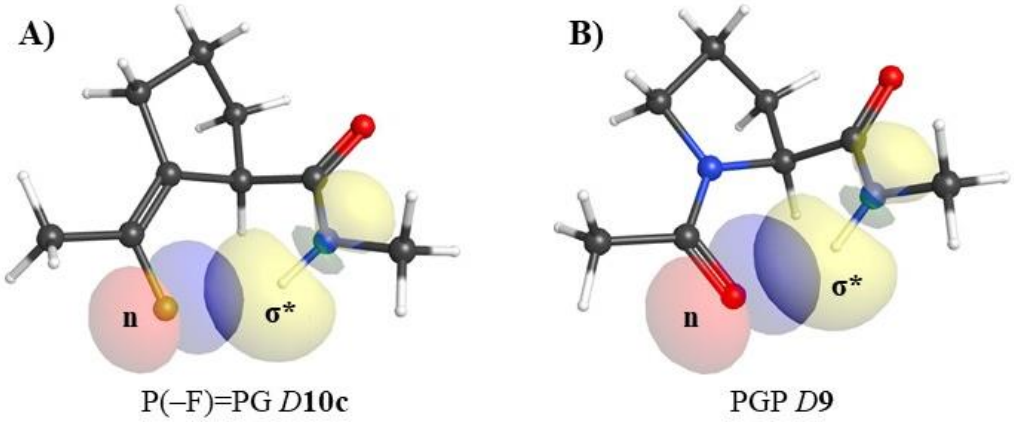


Figure S2.6. NBO images of representative orbitals depicting hydrogen-bonded inverse γ-turns of the P(–X)=PG donor series calculated at the MP2/6-311+G(2d,p) level. **A)** P(–F)=PG min ***D*10c** at Ψ = +64° shows overlap between fluorine non-bonding orbital (red/blue) and H–N antibonding σ* orbital (green/yellow). **B)** PPG min ***D*9c** at Ψ = +62° shows overlap between oxygen non-bonding orbital (red/blue) and H–N antibonding σ* orbital.[48]

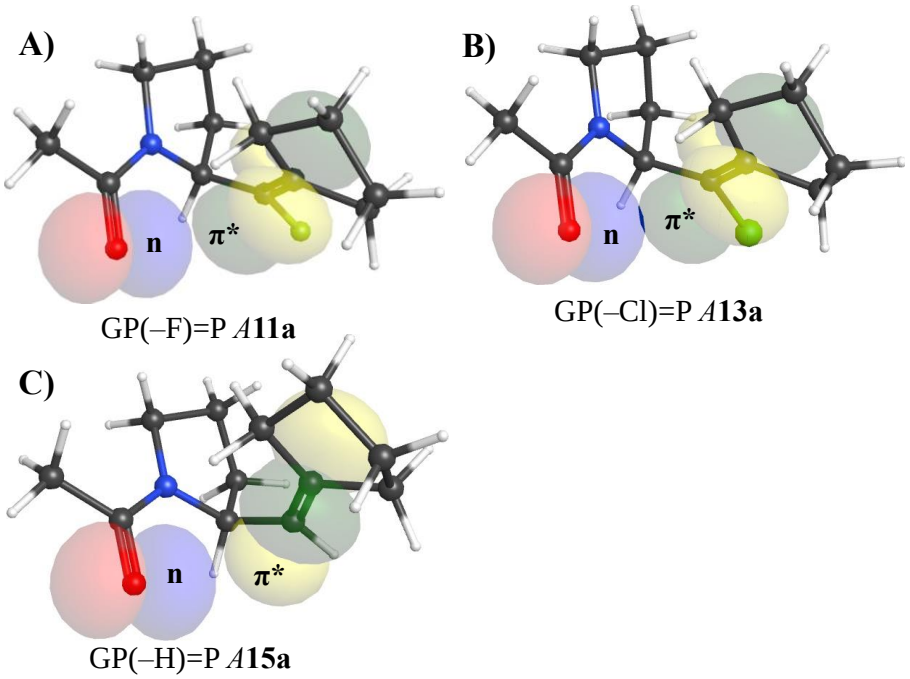


Figure S2.7. NBO images of representative orbitals depicting global minima of the GP(–X)=P acceptor series calculated at the MP2/6-311+G(2d,p) level. **A)** GP(–F)=P min *A***11a** at Ψ = –12° shows no n→*si*-π* overlap between oxygen non-bonding orbital (blue/red) and C=C antibonding π* orbital (green/yellow). **B)** GP(–Cl)=P min *A***13a** at Ψ = –14° shows a possible weak n→*si*-π* interaction between oxygen non-bonding orbital (blue/red) and C=C antibonding π* star orbital. **C)** GP(–H)=P min *A***15b** at Ψ = –12° shows no n→π* interaction between oxygen non-bonding orbital (blue/red) and C=C antibonding π* orbital (green/yellow).

2.11 References

1. Shoulders, M. D.; Raines, R. T., Collagen Structure and Stability. *Annu. Rev. Biochem.* **2009,** *78*, 929-958.

2. Rich, A.; Crick, F. H. C., The Molecular Structure of Collagen. *J. Mol. Biol.* **1961,** *3*, 483-506.

3. Brodsky, B.; Shah, N. K., The Triple-Helix Motif in Proteins. *FASEB J.* **1995,** *9* (15), 1537-1546.

4. Adzhubei, A. A.; Sternberg, M. J. E.; Makarov, A. A., Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol.* **2013,** *425* (12), 2100-2132.

5. Ramshaw, J. A.; Shah, N. K.; Brodsky, B., Gly-X-Y Tripeptide Frequencies in Collagen A Context for Host-Guest Triple-Helical Peptides. *J. Struct. Biol.* **1998,** *122*, 86-91.

6. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal Structure of the Collagen Triple Helix Model $[(Pro\text{-}Pro\text{-}Gly)_{(10)}]_{(3)}$. *Protein Sci.* **2002,** *11* (2), 262-70.

7. DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L., Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **2002,** *124* (11), 2497-2505.

8. Brandts, J. F.; Halvorson, H. R.; Brennan, M., Consideration of the Possibility That the Slow Step in Protein Denaturation Reactions is Due to Cis-Trans Isomerism of Proline Residues. *Biochemistry* **1975,** *14* (22), 4953-63.

9. Bachinger, H., P.; Engel, J.; Bruckner, P.; Timpl, R., The Role of Cis-Trans Isomerization of Peptide Bonds in the Coil ⇄ Triple Helix Conversion of Collagen. *Eur. J. Biochem.* **1978,** *90* (3), 605-613.

10. Etzkorn, F. A.; Zhao, S., Stereospecific Phosphorylation by the Central Mitotic Kinase Cdk1-Cyclin B. *ACS Chem. Biol.* **2015,** *10* (4), 952-956.

11. Mayfield, J. E.; Fan, S.; Wei, S.; Zhang, M.; Li, B.; Ellington, A. D.; Etzkorn, F. A.; Zhang, Y. J., Chemical Tools to Decipher Regulation of Phosphatases by Proline Isomerization on Eukaryotic RNA Polymerase II. *ACS Chem. Biol.* **2015,** *10* (10), 2405-2414.

12. Wang, X. J.; Xu, B.; Mullins, A. B.; Neiler, F. K.; Etzkorn, F. A., Conformationally Locked Isostere of PhosphoSer-cis-Pro Inhibits Pin1 23-Fold Better than PhosphoSer-trans-Pro Isostere. *J. Am. Chem. Soc.* **2004,** *126*, 15533-15542.

13. Dai, N.; Wang, X. J.; Etzkorn, F. A., The Effect of a Trans-Locked Gly-Pro Alkene Isostere on Collagen Triple Helix Stability. *J. Am. Chem. Soc.* **2008,** *130*, 5396-5397.

14. Dai, N.; Etzkorn, F. A., Cis-Trans Proline Isomerization Effects on Collagen Triple-Helix Stability are Limited. *J. Am. Chem. Soc.* **2009,** *131*, 13728-13732.

15. Newberry, R. W.; Raines, R. T., The n-->pi* Interaction. *Acc. Chem. Res.* **2017,** *50* (8), 1838-1846.

16. Etzkorn, F. A.; Ware, R. I.; Pester, A. M.; Troya, D., Conformational Analysis of n→π* Interactions in Collagen Triple Helix Models. *J. Phys. Chem. B* **2019,** *123*, 496-503.

17. Bürgi, H. B.; Dunitz, J. D.; Shefter, E., Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **1973,** *95*, 5065-5067.

18. Choudhary, A.; Gandla, D.; Krow, G. R.; Raines, R. T., Nature of Amide Carbonyl--Carbonyl Interactions in Proteins. *J. Am. Chem. Soc.* **2009,** *131* (21), 7244-7246.

19. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T., n→π* Interactions of Amides and Thioamides: Implications for Protein Stability. *J. Am. Chem. Soc.* **2013,** *135*, 7843-7846.

20. Worley, B.; Richard, G.; Harbison, G. S.; Powers, R., ^{13}C NMR Reveals No Evidence of n-π* Interactions in Proteins. *PLoS One* **2012,** *7* (8), e42075.

21. Rahim, A.; Saha, P.; Jha, K. K.; Sukumar, N.; Sarma, B. K., Reciprocal Carbonyl–Carbonyl Interactions in Small Molecules and Proteins. *Nat. Comm.* **2017,** *8*, 78.

22. Drouin, M.; Paquin, J., Recent Progress in the Racemic and Enantioselective Synthesis of Monofluoroalkene-Based Dipeptide Isosteres. *Beilstein J Org Chem* **2017,** *13*, 2637-2658.

23. Tamamura, H.; Kobayakawa, T.; Ohashi, N., Chloroalkene Dipeptide Isosteres as Peptidomimetics. In *Mid-size Drugs Based on Peptides and Peptidomimetics*, Springer: Singapore, 2018; pp 17-47.

24. Bondi, A., Van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68*, 441-451.

25. Abraham, R. J.; Ellison, S. L. R.; Schonholzer, P.; Thomas, W. A., A Theoretical and Crystallographic Study of the Geometries and Conformations of Fluoro-Olefins as Peptide Analogues. *Tetrahedron* **1986,** *42*, 2101-2110.

26. Pauling, L., The nature of the chemical bond. IV. The energy of single bonds and the relative electroonegativity of atoms. *J. Am. Chem. Soc.* **1932,** *54*, 3570-3582.

27. Kamer, K. J.; Choudhary, A.; Raines, R. T., Intimate Interactions with Carbonyl Groups: Dipole-Dipole or n—>pi*? *J. Org. Chem.* **2013,** *78* (5), 2099-2103.

28. Jakobsche, C. E.; Choudhary, A.; Miller, S. J.; Raines, R. T., n -> π* Interaction and n)(π Pauli Repulsion Are Antagonistic for Protein Stability. *J. Am. Chem. Soc.* **2010,** *132*, 6651–6653.

29. Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T., Code for Collagen's Stability Deciphered. *Nature* **1998,** *392* (6677), 666-7.

30. Shoulders, M. D.; Guzei, I. A.; Raines, R. T., 4-Chloroprolines: Synthesis, Conformational Analysis, and Effect on the Collagen Triple Helix. *Biopolymers* **2007**.

31. Improta, R.; Berisio, R.; Vitagliano, L., Contribution of Dipole-Dipole Interactions to The Stability of the Collagen Triple Helix. *Protein Sci.* **2008,** *17* (5), 955-961.

32. Shoulders, M. D.; Raines, R. T., Interstrand Dipole-Dipole Interactions Can Stabilize the Collagen Triple Helix. *J. Biol. Chem.* **2011,** *286* (26), 22905-22912.

33. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 A Resolution. *Science* **1994,** *226*, 75-81.

34. Schmidt, J. R.; Polik, W. F. *WebMO Enterprise*, version 17.0.012e; WebMO LLC: Holland, MI, USA, 2019.

35. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., *Gaussian 09* (Gaussian, Inc., Wallingford CT, 2009).

36. Kang, Y. K.; Park, H. S., Internal Rotation About the C–N Bond of Amides. *J. Mol. Str.: THEOCHEM* **2004,** *676* (1), 171-176.

37. Weinhold, F.; Landis, C., R., Natural Bond Orbitals and Extensions of Localized Bonding Concepts. *Chem. Educ. Res. Pract.* **2001,** *2*, 91-104.

38. The PyMOL Molecular Graphics System 2.0; Schrödinger, LLC.

39. Schrödinger Release 2021-3: Maestro, S., LLC, New York, NY, 2021.

40. Ramachandran, G. N.; Mitra, A. K., An Explanation for the Rare Occurrence of Cis Peptide Units in Proteins and Polypeptides. *J. Mol. Biol.* **1976,** *107*, 85–92.

41. Harris, T.; Chenoweth, D. M., Sterics and Stereoelectronics in Aza-Glycine: Impact of Aza-Glycine Preorganization in Triple Helical Collagen. *J. Am. Chem. Soc.* **2019,** *141* (45), 18021-18029.

42. von R. Schleyer, P.; West, R., Comparison of Covalently Bonded Electro-Negative Atoms As Proton Acceptor Groups In Hydrogen Bonding. *J. Am. Chem. Soc.* **1959,** *81* (12), 3164-3165.

43. Mishra, S. K.; Suryaprakash, N., Intramolecular Hydrogen Bonding Involving Organic Fluorine: NMR Investigations Corroborated by DFT-Based Theoretical Calculations. *Molecules* **2017,** *22*, 423-467.

44. Gallo, E. A.; Gellman, S. H., N-H-π Hydrogen Bonding in a Norbornenyl Diamide. *Tetrahedron Lett.* **1992,** *33* (49), 7485-7488.

45. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3). *Protein Sci* **2002,** *11* (2), 262-70.

46. Pauling, L.; Corey, R. B.; Branson, H. R., The Structure of Proteins Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U.S.A.* **1951,** *37*, 205-211.

47. Ramachandran, G. N.; Kartha, G., Structure of Collagen. *Nature* **1955,** *176*, 593-595.

48. Etzkorn, F. A.; Ware, R. I.; Pester, A. M.; Troya, D., Conformational Analysis of n→π* Interactions in Collagen Triple Helix Models. *J. Phys. Chem. B.* **2019,** *123*, 496-503.

Chapter 3. A (*Z*)-Fluoro-Alkene Mimic of Gly–*trans*-Pro Produces a Relatively Stable Collagen Triple Helix

Arcoria, P. J., Etzkorn, F. A.

3.1 Abstract

We report the design and asymmetric synthesis, of a (*Z*)-fluoro-alkene isostere of Gly–*trans*-Pro, and its effect on the thermostability of a collagen triple helix. The mimic Fmoc–Gly–Ψ-[(*Z*)CF=C]-L-Pro–Hyp(^{t}Bu)–OH **3.1-(*R,Z*)** was made in 12 steps with 3.5% overall yield. The diastereomer Fmoc–Gly–Ψ-[(*Z*)CF=C]-D-Pro–Hyp(^{t}Bu)–OH **3.1-(*S,Z*)**, was made in 12 steps with 8.7% overall yield. Our attempt to promote asymmetric introduction of the primary chiral center to mimic L-Pro in an aldol reaction inadvertently produced a 7:3 ratio of the undesired D-Pro isostere. Both isosteres were incorporated into the collagen-like peptides Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ-[(*Z*)CF=C]-L/D-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH$_2$. Each peptide was subjected to thermal melting, and circular dichroism was used to measure the loss of ellipticity at 226 nm with increasing temperature. The peptide with the fluoro-alkene guest, Gly–Ψ[(*Z*)CF=C]-L-Pro–Hyp, was found to have a T_m value of 42.2 °C, and the T_m of the control peptide was found to be 49.0 °C, a difference in stability of only ΔT_m –6.8 °C. Thus, the (*Z*)-fluoro-alkene as a Gly–Pro isostere forms a relatively stable triple helix. The enhanced stability of the Gly–Pro (*Z*)-fluoro-alkene isostere compared to the Gly–Pro (*E*)-proteo-alkene (T_m = 28.3 °C) confirms the importance of n→π* electronic interactions in triple-helix stability, as predicted by our modeling studies.

3.2 Introduction

Type I collagen accounts for over 25% of all mammalian protein, making it one of the most abundant structural proteins.[1] Its primary structure can be described as a "helix of helices" where three individual polyproline type II (PPII) helices intertwine around a common screw axis to form

the collagen triple helix.[2, 3] Each PPII helix is a polymer with the repeating sequence (Gly–Xaa–Yaa)$_n$ with a high content of proline in the Xaa position (ca. 28%) and (4R)-hydroxyproline in the Yaa position (ca. 38%).[4] The glycine (Gly) residue is mandatory at every third position because its lack of side-chain allows it to fit into the compact core where the three PPII helices come together.[5]

The peptide bond has a large degree of sp^2 character due to resonance.[6] The prolyl tertiary amide bond has the propensity to isomerize, leading to an unusually high cis:trans ratio.[7] In fact, while over 99% of secondary peptide bonds naturally assume the trans conformation, as much as 30% of tertiary peptide bonds in folded and unfolded proteins may be in the cis conformation.[7] All amide bonds in collagen must adopt the trans conformation.[4] Collagen triple-helix folding is therefore rate-limited by the isomerization of Xaa–Pro sequence.[8] We hypothesized that preventing isomerization by locking it in the trans conformation could lead to a more stable triple helix.

Alkenes are common isosteres of a peptide bond.[9, 10] However, we have previously shown that replacing either one of the tertiary amides in collagen, Gly–Pro or Pro–Pro, with a *trans*-alkene destabilizes the triple helix.[11, 12] Replacement of a single Gly–Pro peptide bond in a 27-residue collagen-like peptide led to ΔT_m = –21.7 ºC compared to the native collagen-like peptide control.[11] Replacement of a single Pro–Pro amide was even more deleterious, with ΔT_m = –53.6 ºC compared to the native control, due to the absence of an interstrand C=O···H–N hydrogen bond.[12] These results suggest that eliminating isomerization is not sufficient to compensate for the loss of noncovalent interactions, such as dipole-dipole interactions or n→π* delocalization, that may be present within the native peptide bond.

Collagen is stabilized by a series of electronic n→π* interactions.[13, 14] In the n→π*interaction, electron density from C_i=O delocalizes into the adjacent C_{i+1}=O π* antibonding orbital.[13, 15] This interaction typically follows the Bürgi-Dunitz trajectory for nucleophilic approach[16] where the energy of approach is strongest around an angle of 108º, and within a distance of < 3.22 Å; the sum of their respective van der Waals radii.[17] The energy afforded by the n→π* interaction of an Ac–Pro–NMe_2 model was estimated to be 0.27 kcal/mol.[18] Our modeling suggests that the Gly–Pro proteo-alkene has two low-energy minima, suggesting flexibility that destabilizes the triple helix.[11, 19] Decreased stabilities of the trans-locked collagen-model peptides were caused, at least in part, by the missing n→π* interaction from the replaced amide.[19] We predict that this effect may be recaptured by using a haloalkene, instead.

We designed the Gly–Pro–Hyp collagen-like isostere **3.1** (Figure 3.1), in which the Gly–Pro peptide bond is replaced with a (*Z*)-fluoro-alkene. Fluoro-alkenes are exceptionally good isosteres of peptide bonds, with many advantages over a proteo-alkene isostere.[20] For example, the van der Waals radius of fluorine (1.47 Å) is much closer to oxygen (1.57 Å) than is hydrogen (1.2 Å).[21] The fluoro-alkene is calculated to have similar orientation and dipole moment to an amide.[22] Fluorine also has lone pairs that can participate in n→π* interactions.[19]

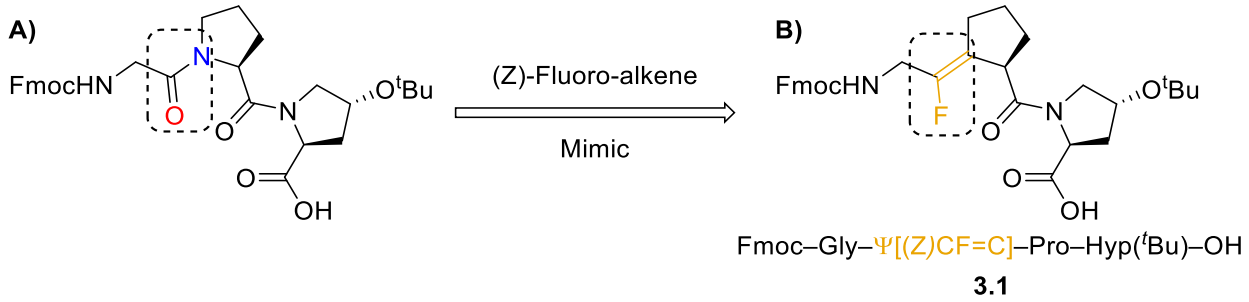


Figure 3.1. Native collagen tripeptide repeat and a (Z)-fluoroalkene Gly–Pro isostere. **A)** Fmoc–Gly–Pro–Hyp(^{t}Bu)–OH tripeptide repeat of native collagen. **B)** Synthesized (*Z*)-fluoro-alkene mimic 3.1. **C)** Host–guest collagen-like peptide sequence 3.2 with trimer mimic 3.1 in the center of the triple helix. **D)** Native collagen-like peptide control 3.3.

We have shown computationally that the global minimum of the Gly–Pro fluoro-alkene as an n→π* donor is at Ψ = +162°,[19] whereas the native peptide bond has its global minimum at Ψ = +160°,[15] suggesting this mimic would be geometrically compatible with the collagen triple helix (Figure 3.2A). Our small models also suggest that the fluorine-carbonyl interaction is just outside of the Bürgi-Dunitz parameters for a significant n→π* interaction to occur. Fortunately, the fluoro-alkene does not have local minima near the collagen geometry that could cause similar conformational instability as the proteo-alkene.[19]

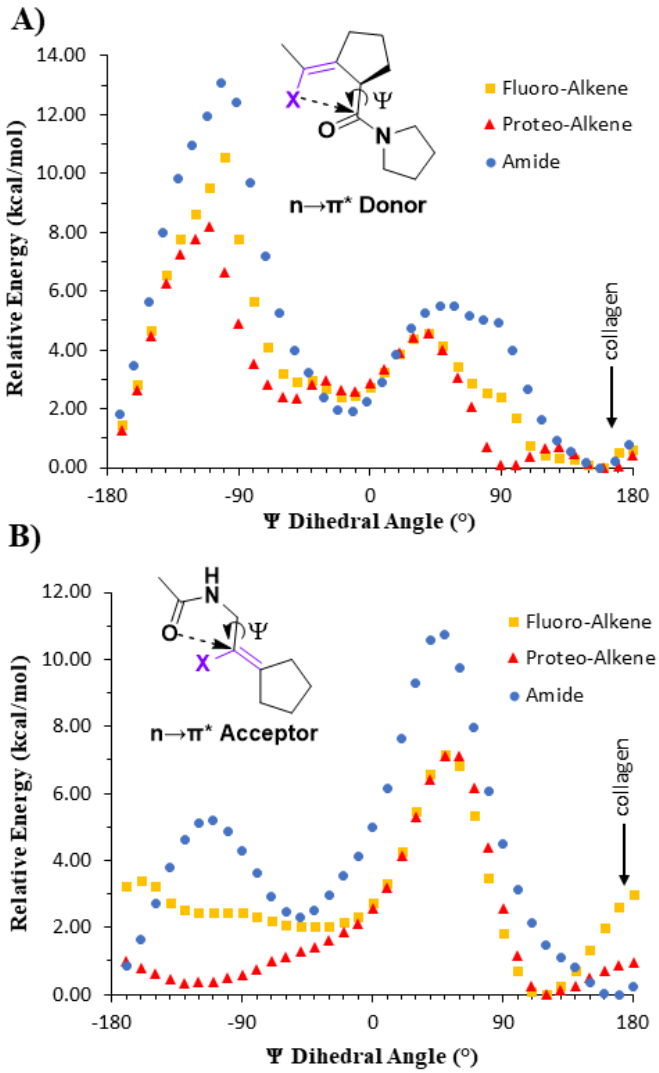


Figure 3.2. Potential energy scans of the Ψ dihedral angle in the fluoro-alkene, proteo-alkene[19] and amide Gly–Pro models.[15] **A)** The X atom acting as an n→π* donor. **B)** The sp^2 carbon acting as an n→π* acceptor models. The average Ψ dihedral angle from the collagen crystal structure PDB ID: 1K6F is labeled.[23] Modified with permission from Arcoria, P. J.; Ware, R. I.; Makwana, S. V.; Troya, D.; Etzkorn, F. A.; Conformational Analysis of Fluoro-, Chloro-, and Proteo-Alkene Gly–Pro and Pro–Pro Isosteres to Mimic Collagen. *The Journal of Physical Chemistry B* 2021. Copyright 2021 American Chemical Society.

The global minimum of the Gly–Pro model fluoro-alkene as an n→π* acceptor is Ψ = +117°,[19] which deviates significantly from the desired minimum at Ψ = +167° (Figure 3.2B).[15]

The difference in Ψ-angle may be caused by the fluoro-alkene being a demonstrably poor n→π* acceptor due to n)(π Pauli repulsion from the alkene π-cloud.[24] This is concerning because the global minimum for the proteo-alkene as an n→π* was about the same Ψ = +120° (Figure 3.2B).[19] We undertook this experimental study of the fluoro-alkene to see if it would maintain PPII-like geometry while also reintroducing noncovalent interactions, such as dipole-dipole and n→π* donation leading to a more stable triple helix.

We now report the design and asymmetric synthesis of Fmoc–Gly–Ψ[(*Z*)CF=C)]–Pro–Hyp(ᵗBu)–OH fluoroalkene mimic **3.**1, and its effect on collagen triple helix stability.

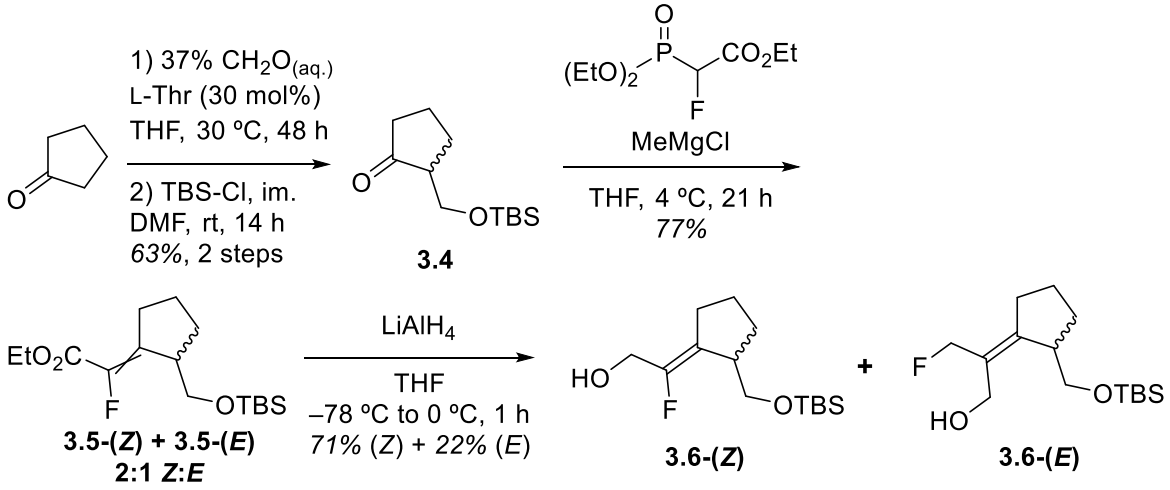
3.3 Results and Discussion

3.3.1 Synthesis of Fluoro-Alkene Isosteres

Enolic addition of formaldehyde to cyclopentanone was asymmetrically catalyzed by L-Thr to give a scalemic mixture (Scheme 3.1). L-Thr has been shown to give the cyclohexanone α-hydroxymethyl product in 90% ee of the desired (*S*)–enantiomer.[25] However, with the 5-membered ring substrate, the undesired (*R*)-enantiomer was produced in high yield, as we learned upon incorporation of the separated diastereomers into the final collagen-like peptides. The resulting primary alcohol (not shown) could not be isolated in satisfactory yields, possibly due to retro-aldol degradation on silica gel. To circumvent this issue, the crude reaction mixture was dried, concentrated, and the alcohol was protected with TBS–Cl in the presence of imidazole to give ketone **3.4** in 63% yield over two steps (Scheme 3.1). Optical rotation was used to confirm scalemic α-hydroxymethyl addition. To ensure racemization did not occur at any point, optical rotation was also obtained on all intermediates throughout the synthesis. The ylide of the commercially available fluoromethylphosphonate ethyl ester was employed to install the fluoro-alkene motif. In our hands, the standard phosphate-ketone coupling[26] resulted in low yield of **3.5**

(39%) with an unfavorable *Z*:*E* ratio (1:1.3 by ^{1}H NMR). The Mg(II)-promoted phosphonate ylide coupling is known to enhance *Z*-selectivity.[27, 28] Both yield (77%) and selectivity (ca. 2.1:1 *Z*:*E* by ^{1}H NMR) were dramatically improved by using MeMgCl as the base in lieu of NaH (Scheme 3.1). This transformation changes the stereochemical designation of the desired enantiomer to (*R*).[29] The mixture of esters **3.5-(*Z*)** and **3.5-(*E*)** was then regioselectively reduced with $LiAlH_4$ to the corresponding 1,2-reduced allylic alcohols **3.6-(*Z*)** and **3.6-(*E*)**. Reduction of **3.5** was achieved with $LiAlH_4$ without 1,4-addition.

Scheme 3.1. Asymmetric synthesis of alcohol **3.6-(*Z*)**.



Although isolation of the desired **3.5-(*Z*)** isomer by flash chromatography was possible, separation of the alcohols was much easier (Scheme 3.1). The allylic protons of **3.6-(*Z*)** and **3.6-(*E*)** were assigned by COSY (Supporting Information), and the configurations of the isolated stereoisomers were confirmed by 1D nOe NMR (Appendix B). Irradiation of the new allylic protons showed correlation either to the prolyl pseudo-δ protons of the (*Z*)-isomer, or to the methylene protons of the CH_2–OTBS group and the pseudo-α-CH proton of the (*E*)-isomer (Figure 3.3)

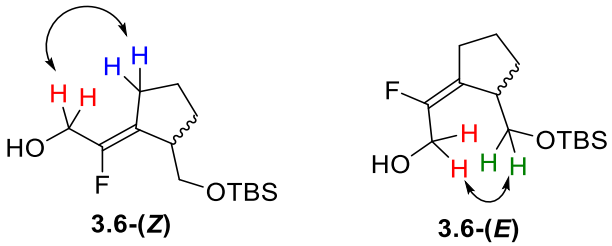
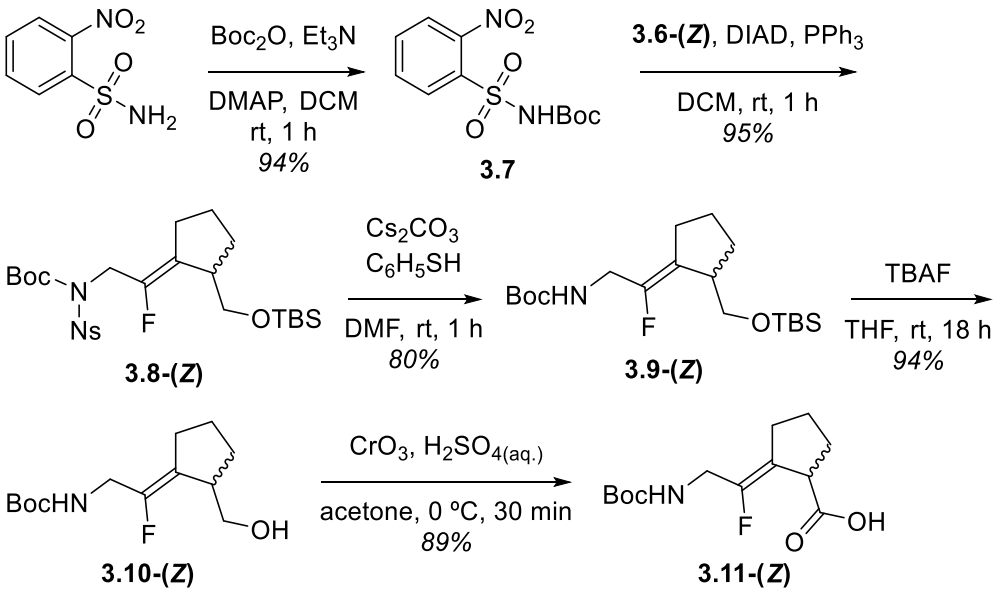


Figure 3.3. Observed 1D-nOe (Appendix B) correlations of allylic alcohols **3.6-(*Z*)** and **3.6-(*E*)**.

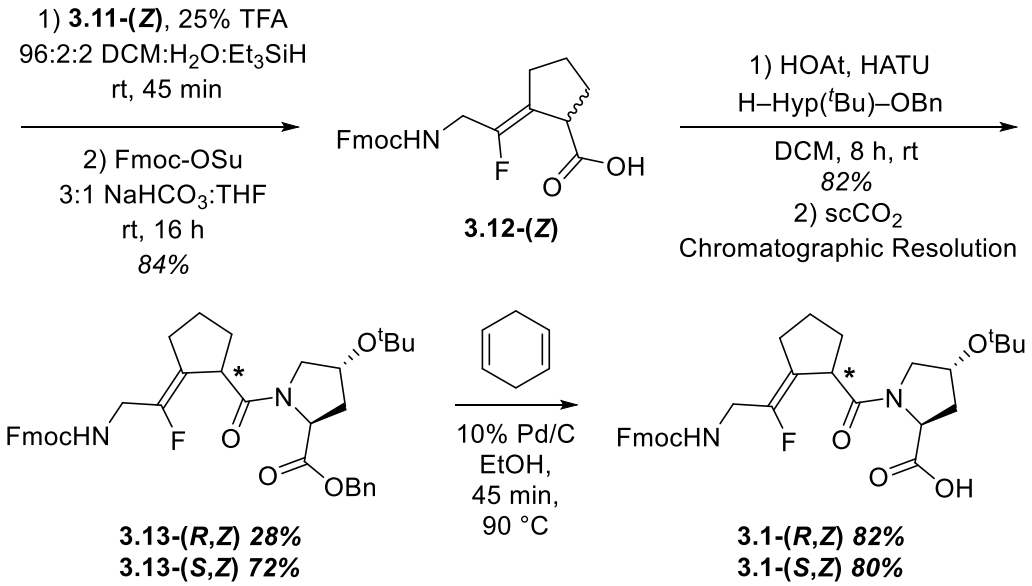
Substitution of the allylic alcohol with the amine was necessary to mimic the native Gly residue, which was accomplished by conversion to the *N*-carboalkoxy-2-nitrobenzenesulfonamide (Scheme 3.2).[28, 30] The *N*-Boc-2-nitrobenzenesulfonamide reagent **3.7** was synthesized from commercially available 2-nitrobenzenesulfonamide and Boc_2O (Scheme 3.2). Subsequent displacement of the allylic alcohol with sulfonamide **3.7** using diisopropyl azodicarboxylate (DIAD) and PPh_3 gave sulfonamide **3.8-(*Z*)** in 95% yield (Scheme 3.2).[28, 30] Deprotection of *o*-nitrobenzenesulfonyl (Ns) was accomplished using thioanisole and Cs_2CO_3 in DMF to quickly give carbamate **3.9-(*Z*)** in 80% yield (Scheme 3.2). The silyl protecting group was then removed by stirring overnight with tetra-*n*-butyl ammonium fluoride (TBAF) to prepare alcohol **3.10-(*Z*)** in 94% yield (Scheme 3.2). Oxidation of **3.10-(*Z*)** in acetone with CrO_3/H_2SO_4 afforded acid **3.11-(*Z*)** in 89% yield (Scheme 3.2). The reaction duration was kept short (30 min), and the temperature was kept cold (0 °C) to help prevent premature cleavage of the Boc group because the resulting ammonium salt would be difficult to isolate.

Scheme 3.2 Synthesis of the (*Z*)-fluoro-alkene Gly–*trans*-Pro mimic Fmoc–Gly–Ψ[(*Z*)CF=C)]–Pro–OH **3.11-(*Z*)**.



Following oxidation, the Boc group was removed with 25% TFA in DCM using 2% each of Et_3SiH and H_2O as radical scavengers. The TFA salt (not isolated) was protected with Fmoc–OSu in $NaHCO_3$:THF giving Fmoc amine **3.12-(*Z*)** in 84% yield (Scheme 3.3). H–Hyp(^{t}Bu)–OBn, which was made from Fmoc–Hyp(^{t}Bu)–OH in 83% yield over two steps, was coupled to Fmoc amine **3.12-(*Z*)** using HATU and HOAt to give benzyl ester **3.13-(*Z*)** in 82% yield as a mixture of diastereomers in a 72:28 *S*:*R* ratio, as determined by HPLC (Scheme 3.3). The diastereomers were separated by HPLC using 15% *i*PrOH in $scCO_2$ on a normal-phase pyridine column. Transfer hydrogenation with 1,4-cyclohexa-diene on 10% Pd/C removed the benzyl protecting group, without reduction of the fluoro-alkene (Scheme 3.3).[11] In 12 steps, the desired Fmoc–Gly–Ψ[(*Z*)CF=C)]–L-Pro–Hyp(^{t}Bu)–OH **3.1-(*R*,*Z*)** was obtained in 3.5% overall yield, and Fmoc–Gly–Ψ[(*Z*)CF=C)]–D-Pro–Hyp(^{t}Bu)–OH **3.1-(*S*,*Z*)** in 8.7% overall yield.

Scheme 3.3. Synthesis of the tripeptide isosteres, Fmoc–Gly–Ψ[(*Z*)CF=C)]–L-Pro–Hyp(ᵗBu)–OH **3.1-(*R*,*Z*)** and Fmoc–Gly–Ψ[(*Z*)CF=C)]–D-Pro–Hyp(ᵗBu)–OH **3.1-(*S*,*Z*)**.



All peptides were synthesized by coupling Fmoc-protected tripeptide units with either HOAt/HATU or 6–Cl–HOBt/HBTU on solid-phase Rink amide MBHA resin. The Tyr residue was included to determine peptide concentration by UV. The two Gly residues were included as a spacer to facilitate peptide synthesis and to decrease interference of Tyr with triple-helix folding.[11] The Fmoc–Gly–Pro–Hyp(ᵗBu)–OH trimer units were synthesized by the method of Ottl et al..[31] The crude peptides were purified by HPLC on a C18 column, and their identity was confirmed by MALDI-TOF (Table S3.1).

3.3.2 Triple-Helix Stability

The stability of the triple helix was determined using circular dichroism (CD) to conduct a thermal melt from 4 °C to 85 ºC (Figure 3.4). The measured loss of ellipticity of the signature triple helix peak at 226 nm was used to determine triple-helix stability. The data was fit into a two-state model to calculate T_m. The control peptide Ac–(Gly–Pro–Hyp)$_8$–Gly–Gly–Tyr–NH_2 **3.3** was found to have a T_m of 49.0 °C (Figure 3.5). This closely matches the literature melting temperature of a

collagen-like peptide with the same sequence (T_m = 50.0).[11] Our fluoro-alkene model, Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C)]–L-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH$_2$ **3.2-(*R*,*Z*)**, has T_m = 42.2 ºC (Figure 3.5). The diastereomer Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C)]–D-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH$_2$ **3.2-(*S*,*Z*)** showed a linear decrease in ellipticity, meaning that it did not form a triple helix (Figure 3.6).

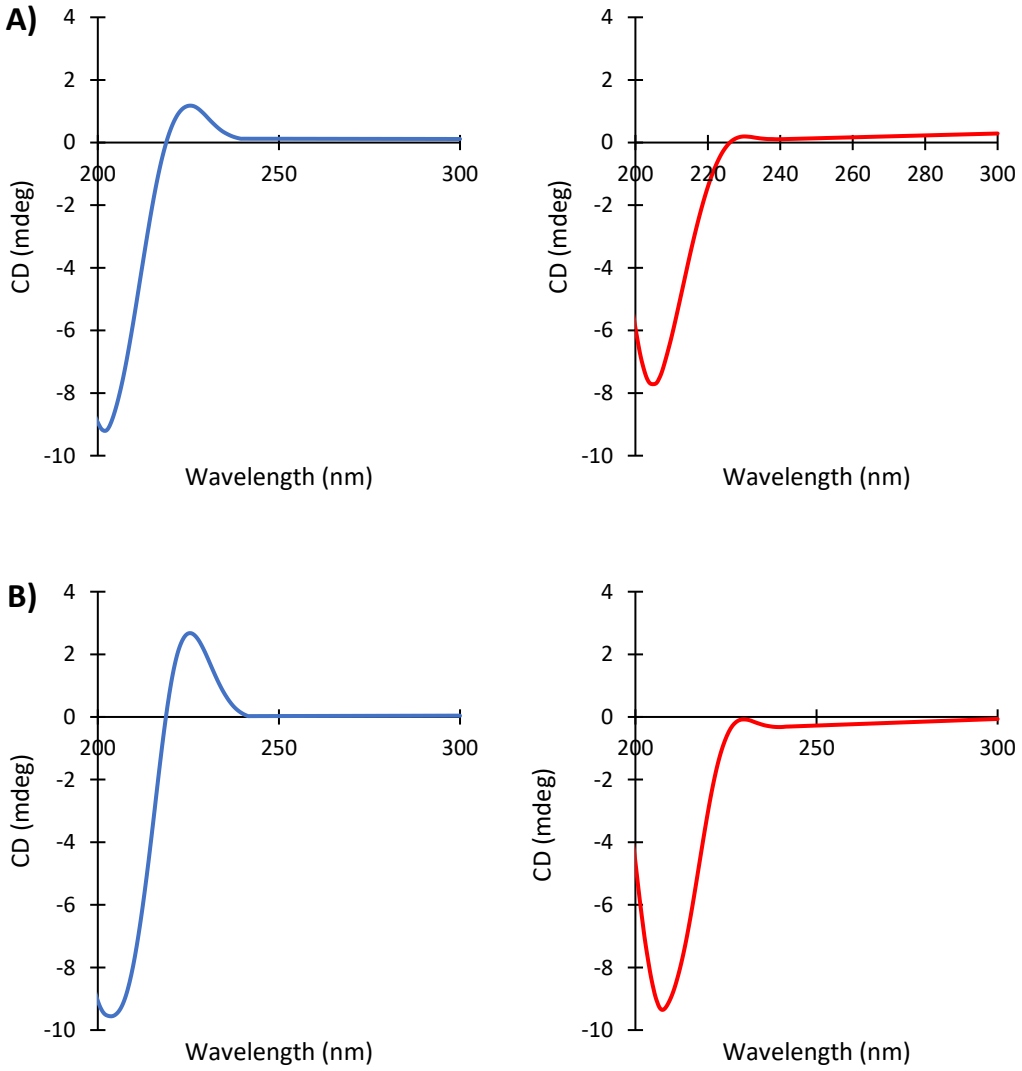


Figure 3.4. Full-length CD spectra. Temperature at 5 °C (blue) and 85 °C (red) of **A)** peptide **3.2-(*R*,*Z*)** (0.097 mM) and **B)** control **3.3** (0.53 mM).

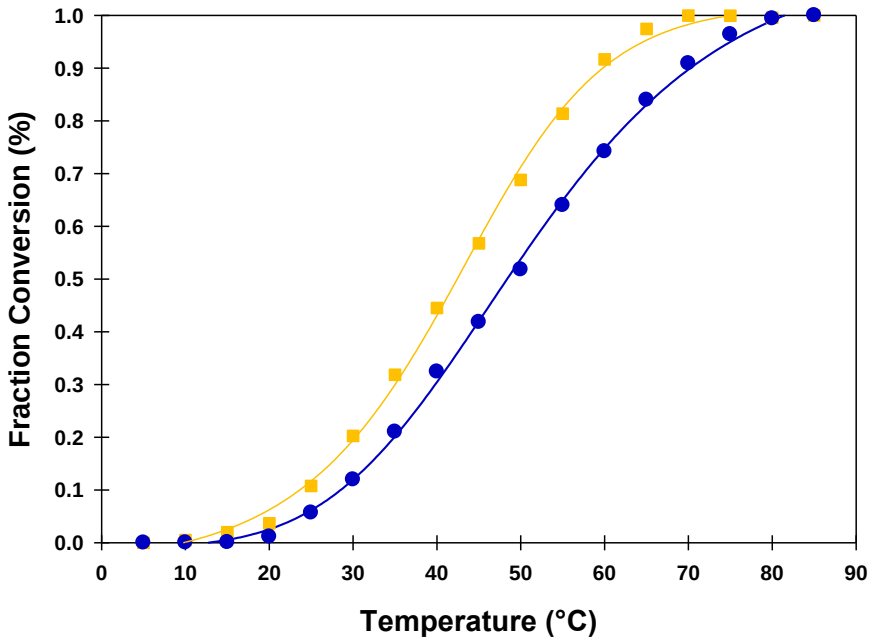


Figure 3.5. Thermal melting curves of the fluoro-alkene mimic **3.2-(*R,Z*)** (■) and control **3.3** (●).

The absolute stereochemistry of the chiral center introduced into the aldol intermediate **3.4** (Scheme 3.1) was inferred from the thermal stability of their respective peptides (Figure 3.5, Figure 3.6). Since the native triple helix contains all L-amino acids, a single D-amino acid in the sequence disrupts its ability to fold in phosphate buffered saline solutions.[32] The diastereomer of **3.4** that ultimately provided a sigmoidal melting curve was assigned the (*R*) configuration, a mimic of L-Pro mimic. The diastereomer with a linear correlation between temperature and ellipticity was assumed to be the D-Pro mimic. We intended to promote the formation of the L-Pro mimic by using L-Thr as a chiral catalyst.[25] Our results, however, show that the major diastereomer formed did not have a cooperative melting curve (Figure 3.6) and is therefore the undesired D-Pro mimic. We propose that the envelope conformation of the cyclopentanone starting material caused L-Thr to introduce the chiral center in the opposite stereochemistry from that induced by the chair conformation of cyclohexanone.[25] Further study of this phenomenon is needed. Nevertheless, the

minor diastereomer was sufficient to determine the effect of a (*Z*)-fluoro-alkene Gly–Pro isostere on triple helix stability.

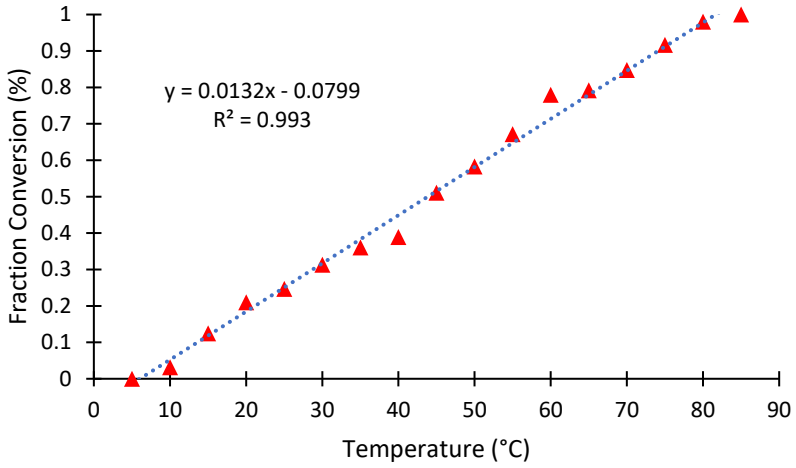


Figure 3.6. Linear ellipticity of the major diastereomer, **3.2-(S,Z)**.

The T_m of peptide **3.2-(*R,Z*)** was found to be 42.2 °C, a decrease (ΔT_m) of 6.8 °C compared to control peptide **3.3**. Previous results from Dai et al.[11] showed that replacement of the Gly–Pro amide with the proteo-alkene isostere reduced T_m to 28.3 ºC (ΔT_m = –21.7). In our modeling, we found very similar potential energy landscapes of the Pro–Pro Ψ dihedral angle for the fluoro-alkene and proteo-alkene (Figure 3.2A).[19] The key difference between the two is a nearby local minimum found with the proteo-alkene that is not observed for the fluoro-alkene (Figure 3.2A). Our results here imply that the extra flexibility conferred by that local minimum significantly impacts the stability of the triple helix. Our Gly–Pro fluoro-alkene has ΔT_m +13.3 °C higher than the Gly–Pro proteo-alkene[11] suggesting significant stability is afforded by electronic interactions.

Our theoretical results predicted that the fluoro-alkene might not behave as an n→π* donor because the fluorine atom rests just outside of the necessary Bürgi-Dunitz distance and angle to the acceptor carbonyl.[19] Furthermore, experimental and computational results predict that the fluoro-alkene would be a poor n→π* acceptor.[19, 24] However, in this work, ^{19}F NMR data showed

significant deshielding of the fluorine nucleus (Figure 3.7). A shift from –111.6 ppm to –76.0 ppm occurs between intermediates **3.12-(*Z*)** and **3.13-(*R*,*Z*)**, which may be caused by the loss of electron density through n→π* donation. After coupling of Hyp(^{t}Bu)–OBn, the newly-formed Pro–*trans*–Hyp amide of **3.13-(*R*,*Z*)** may preorganized to be an n→π* acceptor, and the ^{19}F NMR peak at –76.0 ppm reflects that deshielding (Figure 3.7). Removal of the benzyl group permits easier *cis*-*trans* isomerization of the Pro–Hyp amide, as seen with the two rotamer ^{19}F peaks at –115.1 and –112.6 ppm for **3.1-(*R*,*Z*)** (Figure 3.7).

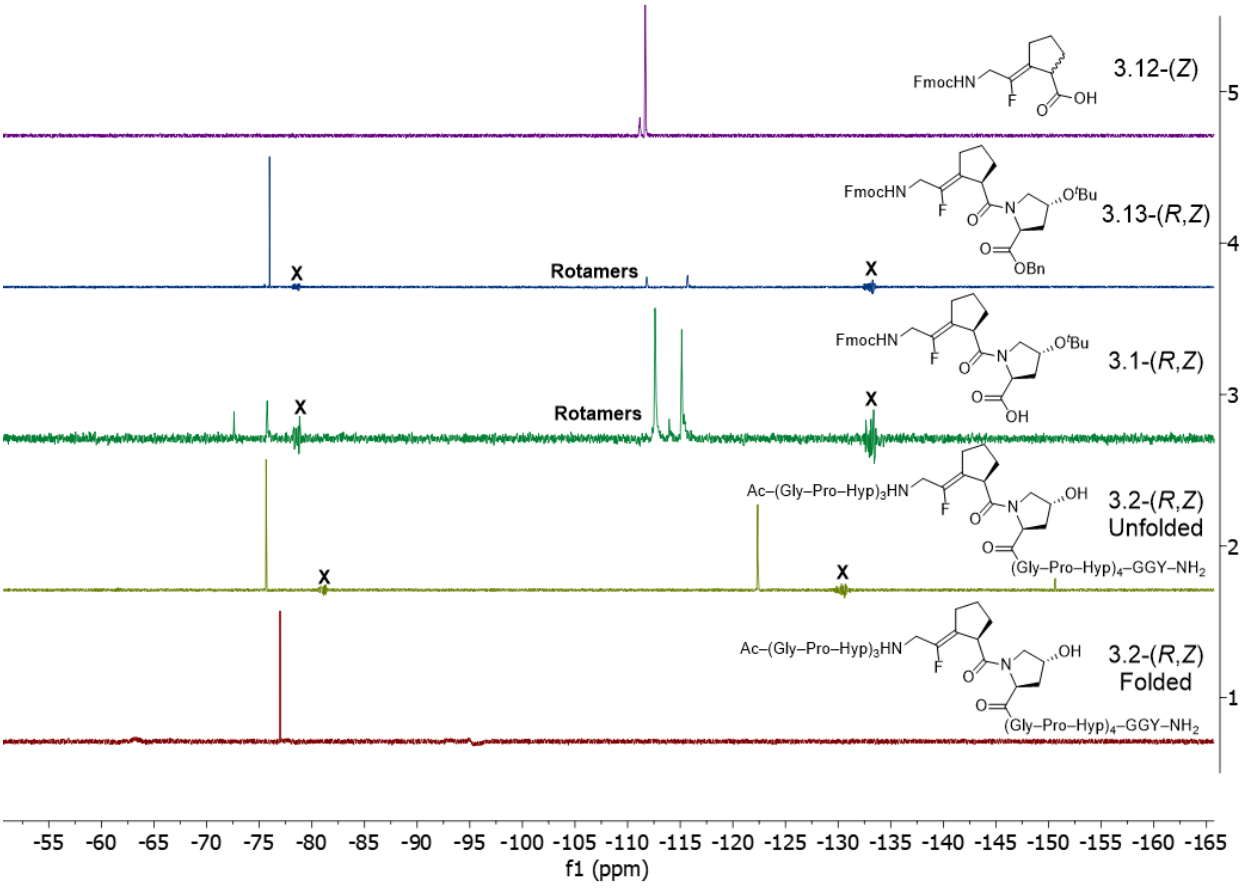


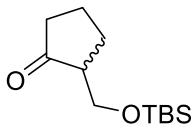
Figure 3.7. Stacked ^{19}F NMR spectra of the final five compounds of the synthesis showing the increasingly deshielded fluorine nucleus. Instrument noise is marked with an X. Rotamer peaks are labeled. The downfield peaks between –70 to –80 ppm represent fluorine atoms acting as an n→π* donor. The upfield peaks between –110 to –120 ppm represent fluorine atoms that are not engaged in n→π* donation.

Increased flexibility in the carboxylic acid **3.1-(*R,Z*)**, indicated by the higher cis/trans ratio, would make n→π* donation more difficult, as suggested by the ^{19}F peaks at –75.7 and –72.5 ppm that are much less intense than the upfield rotamer peaks. Following HPLC purification, the unfolded collagen-like peptide **3.2-(*R,Z*)** showed only two ^{19}F NMR signals at –75.6 ppm and –122.4 ppm (Figure 3.7). We interpret these signals to correspond to part of the peptide that is already folded (–75.6 ppm, n→π* interaction) and part that is unfolded (–122.4 ppm, no n→π* interaction). Following incubation at 4 °C for 72 h, the folded collagen-like peptide **3.2-(*R,Z*)** had only a single peak at –77.0 ppm, suggesting that all fluorine atoms are deshielded by their participation in n→π* donation (Figure 3.7). An n→π* interaction would explain the enhanced stability of the fluoro-alkene (T_m = 42.2 °C) over the proteo-alkene (T_m = 28.3 °C).[11]

3.4 Experimental

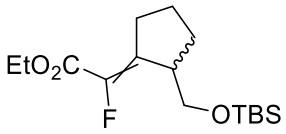
General Methods. All reactions were performed under inert gas in oven-dried glassware. All reagents and resins were obtained from commercial suppliers and used as received, unless otherwise stated. All amine bases were stirred over CaH_2 overnight and distilled prior to use. Solvents were dried using an Innovative Technology Pure Solv-MD solvent purification system. Brine (NaCl), $NaHCO_3$, and NH_4Cl refer to saturated aqueous solutions, unless otherwise stated. SiO_2 chromatography was performed with SiliaFlash P60 silica gel (230-400 mesh) provided by Silicycle with HPLC-grade solvents. ^{1}H (400 MHz), ^{13}C (100 MHz), and ^{19}F (376 MHz) NMR spectra were recorded on an Agilent MR-400 MHz NMR spectrometer. Chemical shifts for ^{1}H

NMR (unless otherwise noted) are reported in ppm with reference to residual $CHCl_3$ at 7.26 ppm or TMS at 0.00 ppm. Proton-decoupled carbon chemical shifts are reported in ppm with reference to $CDCl_3$ at 77.0 ppm; the chemical shifts of minor rotamers are in parenthesis. ^{19}F NMR spectra were unlocked. NMR data are described as follows: chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, dd = doublet of doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), and integration. The gCOSY for **3.6-(*Z*)** and **3.6-(*E*)** were obtained at 25 °C in $CDCl_3$ with 8 scans and a 1 s relaxation delay. The 1D-nOe spectra for **3.6-(*Z*)** and **3.6-(*E*)** were obtained at 25 °C in $CDCl_3$ with 256 scans and a 1 s relaxation delay. ^1H NMR spectra of collagen-like peptides **3.2-(*R*,*Z*)**, **3.2-(*S*,*Z*)**, and **3.3** were obtained on an Avance III 600 MHz spectrophotometer equipped with a liquid nitrogen prodigy cryoprobe. Spectra were acquired using the standard one-dimensional zgpr pulse sequence with water presaturation during the recycle delay of 2 s. Each dataset was averaged over 256 scans using 64k time domain points. High resolution mass spectra (HRMS) were acquired on an Agilent 6220 using electrospray ionization (ESI^+) and time of flight analyzer (TOF). Organic solutions were concentrated under reduced pressure using a Büchi RE-111 rotary evaporator. Optical rotation was measured on a Jasco P-2000 Polarimeter at the sodium D-line (589 nm); S aperture Φ 1.8, L Aperture Φ 1.0. Circular dichroism was performed with a Jasco model 815 spectropolarimeter.



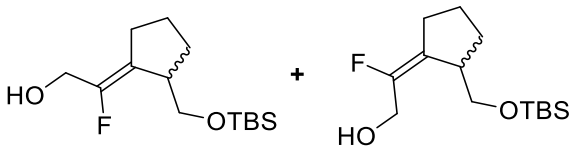
Ketone 3.4. Cyclopentanone (48 mL, 540 mmol) and L-Thr (3.6 g, 30. mmol) were stirred in THF (100 mL) at 30 °C for 1 h. Then a 37% formalin solution (7.5 mL, 100 mmol) was added dropwise. The reaction was stirred at 30 °C for 48 h, then diluted with EtOAc (50 mL), dried over Na_2SO_4, filtered, and concentrated. The resultant brown oil was dissolved in DMF (36 mL) with imidazole

(10 g, 150 mmol) and TBS-Cl (11 g, 75 mmol) at 0 °C. After mixing for 14 h at rt, the reaction mixture was diluted with EtOAc (100 mL), washed with water (3 × 100 mL), dried over Na_2SO_4, and concentrated. Purification by SiO_2 chromatography (5 cm × 17 cm, 17% EtOAc/hexanes; R_f = 0.42) gave ketone **3.4** as a colorless oil (7.2 g, 63%). The ^{1}H NMR data matches the literature.[33] ($CDCl_3$, 400 MHz): δ 3.85 (dd, *J* = 9.8, 4.8 Hz, 1H), 3.73 (dd, *J* = 9.9, 3.3 Hz, 1H), 2.30 – 1.90 (m, 6H), 1.84 – 1.71 (m, 1H), 0.85 (s, 9H), 0.2 (s, 3H), 0.1 (s, 3H); ^{13}C{^{1}H} NMR ($CDCl_3$, 125MHz): δ 220.4, 62.2, 51.1, 39.3, 26.5, 26.0, 21.1, 18.4, –5.4, –5.5. HRMS (ESI$^+$/TOF) *m/z*: [M + H]$^+$ Calcd for $C_{12}H_{25}O_2Si^+$ 229.1618; Found 229.1608. $[\alpha]_D^{21}$ –31.2° (*c* 4, MeOH).



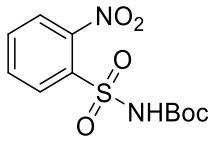
Esters 3.5-(*Z*) and 3.5-(*E*). A 3.0 M solution of MeMgCl in THF (2.4 mL, 7.1 mmol) was added dropwise to a stirring solution of triethyl 2-fluoro-2-phosphonoacetate (1.7 g, 7.0 mmol) in THF (30 mL) at –78 °C and stirred for 30 min. Ketone **3.4** (1.7 g, 7.2 mmol) in THF (20 mL) was added dropwise and the reaction was stirred for 17 h at 4 °C. It was then quenched with water (150 mL) and extracted with DCM (4 × 100 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 cm × 18 cm, 1:10 EtOAc:hexanes; R_f = 0.40) provided esters **3.5-(*Z*)** and **3.5-(*E*)** as a colorless oil (1.8 g, 77%) in a mixture of diastereomers with ca. 1:0.49 *Z*:*E* ratio, as determined by ^{1}H NMR by integration of the diastereomeric proton: ($CDCl_3$, 400 MHz): **3.5-(*Z*)**: δ 3.73 (dd, *J* = 9.9 Hz, 4.3 Hz, 1H) and **3.5-(*E*)** δ 3.64 (dd, *J* = 9.2 Hz, 4.1 Hz, 1H). **3.5-(*Z*)**. The ^{1}H NMR data matches literature.[28, 34] ($CDCl_3$, 400 MHz): δ 4.26 (q, *J* = 7.1 Hz, 2H), 3.74 (dd, *J* = 9.9, 4.3 Hz, 1H), 3.50 (dd, *J* = 9.7, 8.3 Hz, 1H), 3.19 – 3.08 (m, 1H), 2.72 – 2.65 (m, 2H), 1.91 – 1.65 (m, 4H), 1.32 (t, *J* = 7.1 Hz, 3H), 0.87 (s, 9H), 0.03 (s, 3H), 0.02 (s, 3H); ^{13}C{^{1}H} NMR ($CDCl_3$, 125 MHz): δ 161.4 (d, $^2J_{C,F}$ = 35.1 Hz),

142.4 (d, $^{1}J_{C,F}$ = 247.2 Hz), 141.4 (d, $^{2}J_{C,F}$ = 12.6 Hz), 63.2 (d, $^{4}J_{C,F}$ = 4.2 Hz), 61.1, 46.2, 31.2 (d, $^{3}J_{C,F}$ = 1.9 Hz), 28.7, 26.0, 24.8, 18.4, 14.4, –5.3, – 5.4; ^{19}F NMR ($CDCl_3$, 376 MHz): δ –127.2 (s). $[\alpha]_D^{21}$ –12.3° (*c* 4, MeOH). **3.5-(*E*)**: ^{1}H NMR ($CDCl_3$, 400 MHz): δ 4.28 (q, *J* = 7.1 Hz, 2H), 3.65 (dd, *J* = 9.2, 4.1 Hz, 1H), 3.48 (dd, *J* = 9.2, 8.1 Hz, 1H), 3.44 – 3.35 (m, 1H), 2.56 – 2.49 (m, 2H), 2.08 – 1.99 (m, 1H), 1.84 – 1.64 (m, 3H), 1.33 (t, *J* = 7.1 Hz, 3H), 0.88 (s, 9H), 0.04 (s, 3H), 0.02 (s, 3H).

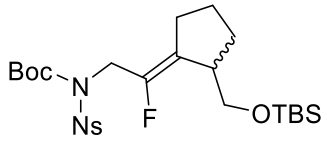


Allylic Alcohols 3.6-(*Z*) and 3.6-(*E*). To a stirring solution of esters **3.6-(*Z*)** and **3.6-(*E*)** (1.4 g, 4.5 mmol) in THF (20 mL) at –78 °C, $LiAlH_4$ (1.0 M in THF, 14 mL, 14 mmol) was added dropwise and the mixture was stirred at 0 °C for 1 h. The reaction was quenched with EtOAc (100 mL), washed with water (3 × 50 mL) and brine (1 × 30 mL), dried over Na_2SO_4, and concentrated. Purification by flash chromatography (5 cm × 16 cm, 20% EtOAc:hexanes; R_f = 0.22) gave allylic alcohol **3.6-(*Z*)** as a colorless oil (870 mg, 71%) and allylic alcohol **3.6-(*E*)** as a colorless oil (270 mg, 22%). **Allylic alcohol 3.6-(*Z*)** ^{1}H NMR ($CDCl_3$, 400 MHz): δ 4.27 – 4.07 (m, 2H), 3.75 (dd, *J* = 9.8, 4.4 Hz, 1H), 3.42 (dd, *J* = 9.8, 8.6 Hz, 1H), 3.02 – 2.93 (m, 1H), 2.32 – 2.22 (m, 2H), 1.90 – 1.55 (m, 5H), 0.88 (s, 9H), 0.04 (s, 3H), 0.03 (s, 3H); ^{13}C{^{1}H} NMR ($CDCl_3$, 100 MHz): δ 151.5 (d, $^{1}J_{C,F}$ = 245.3 Hz), 123.5 (d, $^{2}J_{C,F}$ = 14.8 Hz), 64.0 (d, $^{4}J_{C,F}$ = 3.8 Hz), 59.7 (d, $^{2}J_{C,F}$ = 30.5 Hz), 43.6, 29.41, 28.7 (d, $^{3}J_{C,F}$ = 5.1 Hz), 26.1, 24.9, 18.5, –5.2, –5.3; ^{19}F NMR ($CDCl_3$, 376 MHz) δ –119.8 (dd, $^{3}J_{H,F}$ = 21, 21 Hz). HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{14}H_{28}FO_2Si^+$ 275.1837; Found 275.1842. $[\alpha]_D^{21}$ –11.3° (*c* 4, MeOH). **Allylic alcohol 3.6-(*E*)** ^{1}H NMR ($CDCl_3$, 400 MHz): δ 4.27 – 4.09 (m, 2H), 3.52 (dd, *J* = 9.8, 5.9 Hz, 1H), 3.35 (dd, *J* = 9.8, 9.2 Hz, 1H), 2.99 – 2.83

(m, 1H), 2.55 – 2.41 (m, 1H), 2.40 – 2.28 (m, 1H), 1.86 – 1.47 (m, 4H), 0.90 (s, 9H), 0.08 (s, 3H), 0.08 (s, 3H).

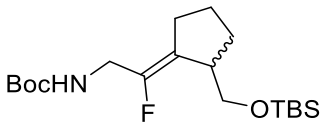


***N*-Boc-2-nitrobenzenesulfonamide 3.7.** 2-Nitrobenzenesulfonamide (510 mg, 2.5 mmol), Boc_2O (700 mg, 3.2 mmol), Et_3N (630 μL, 4.5 mmol), and a catalytic amount of DMAP (53 mg, 0.43 mmol) were dissolved in DCM (5.0 mL) and stirred at rt for 1 h. The reaction was quenched with 1 M HCl (12 mL), and the product was extracted with Et_2O (4 × 15 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. Trituration with 40% Et_2O in hexanes (5 × 10 mL) gave **7** as a white solid (710 mg, 94%): ^{1}H NMR data matches the literature.[35] (CD_2Cl_2, 400 MHz): δ 8.31 – 8.27 (m, 1H), 7.88 – 7.77 (m, 3H), 1.42 (s, 9H); $^{13}C\{^1H\}$ NMR data matches the literature.[35] (CD_2Cl_2, 100 MHz): δ 149.2, 148.6, 135.5, 133.6, 133.0, 132.2, 125.6, 85.4, 28.1.

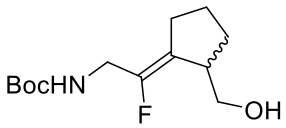


Sulfonamide 3.8-(*Z*). DIAD (570 mg, 2.8 mmol) in DCM (5 mL) was added dropwise to a solution of allylic alcohol **3.6-(*Z*)** (590 mg, 2.2 mmol), *N*-Boc-2-nitrobenzenesulfonamide **3.7** (830 mg, 2.8 mmol), and PPh_3 (720 mg, 2.8 mmol) in DCM (15 mL) at 0 °C and stirred for 1 h at rt. The solution was then diluted with EtOAc (75 mL), washed with water (3 × 50 mL) and brine (1 × 50 mL), dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 cm × 17 cm, 25% EtOAc/hexanes; R_f = 0.36) afforded sulfonamide **3.8-(*Z*)** as a colorless oil (1.1 g, 95%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.31 – 8.25 (m, 1H), 7.78 – 7.70 (m, 3H), 4.53 (d, *J* = 17.2 Hz, 2H), 3.77 (dd, 9.8 Hz, 4.4 Hz, 1H), 3.39 (dd, 9.8 Hz, 8.9 Hz, 1H) 3.05 – 2.95 (m, 1H), 2.47 – 2.25 (m, 2H), 1.86 – 1.54 (m, 4H), 1.33 (s, 9H), 0.87 (s, 9H), 0.04 (s, 3H), 0.03 (s, 3H); $^{13}C\{^1H\}$ NMR ($CDCl_3$,

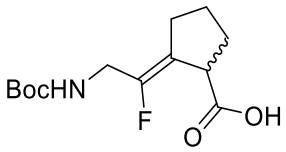
100 MHz): δ 150.1, 148.4, 147.8, 146.0, 133.6, 133.3, 133.1 (d, $^{1}J_{C,F}$ = 236.6 Hz), 124.6, 124.2 (d, $^{2}J_{C,F}$ = 14.2 Hz), 85.4, 63.8 (d, $^{4}J_{C,F}$ = 3.8 Hz), 45.5 (d, $^{2}J_{C,F}$ = 28.6 Hz), 43.9, 29.3, 28.6 (d, $^{3}J_{C,F}$ = 4.7 Hz), 27.9, 26.0, 24.8, 18.4, –5.2, –5.3; ^{19}F NMR ($CDCl_3$, 376 MHz): δ –118.5 (dd, $^{3}J_{H,F}$ = 17, 17 Hz). HRMS (ESI$^+$/TOF) *m/z*: [M + H]$^+$ Calcd for $C_{25}H_{40}FN_2O_7SSi^+$ 559.2304; Found 559.2301. $[\alpha]_D^{21}$ –6.7° (*c* 4, MeOH).



Carbamate 3.9-(*Z*). Thiophenol (280 mg, 2.5 mmol) and Cs_2CO_3 (1.3 g, 4.0 mmol) were stirred in DMF (100 mL) at 0 °C for 15 min. A solution of sulfonamide **3.8-(*Z*)** (780 mg, 1.4 mmol) in DMF (75 mL) was then added dropwise and stirred for 1 h at rt. It was then diluted with EtOAc (100 mL), washed with water (4 × 100 mL) and brine (1 × 50 mL), dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 cm × 17 cm, 19% EtOAc/hexanes; R_f = 0.48) provided carbamate **3.9-(*Z*)** as a colorless oil (420 mg, 80%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 4.78 (bs, 1H), 3.93 – 3.75 (m, 2H), 3.70 (dd, *J* = 9.8 Hz, 4.4 Hz, 1H), 3.38 (dd, *J* = 9.8 Hz, 8.6 Hz, 1H), 2.97 – 2.87 (m, 1H), 2.34 – 2.20 (m, 2H), 1.87 – 1.48 (m, 4H), 1.42 (s, 9H), 0.86 (s, 9H), 0.02 (s, 3H), 0.01 (s, 3H); ^{13}C{^{1}H} NMR ($CDCl_3$, 100 MHz): δ 155.7, 149.4 (d, $^{1}J_{C,F}$ = 245.7 Hz), 122.5 (d, $^{2}J_{C,F}$ = 15.0 Hz), 79.9, 64.0 (d, $^{4}J_{C,F}$ = 3.8 Hz), 43.6, 40.0 (d $^{2}J_{C,F}$ = 29.4 Hz), 29.5, 28.8 (d, $^{3}J_{C,F}$ = 5.0 Hz), 28.5, 26.1, 24.8, 18.4, –5.2, –5.3; ^{19}F NMR ($CDCl_3$, 376 MHz): δ –117.2 (dd, $^{3}J_{H,F}$ = 20, 20 Hz). HRMS (ESI$^+$/TOF) *m/z*: [M + H]$^+$ Calcd for $C_{19}H_{37}FNO_3Si^+$ 374.2521; Found 374.2518. $[\alpha]_D^{21}$ –8.9° (*c* 4, MeOH).

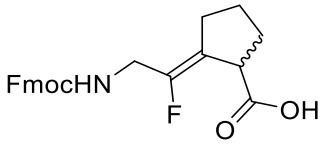


Alcohol 3.10-(*Z*). Carbamate **3.9-(*Z*)** (1.1 g, 2.9 mmol) and Bu_4NF (4.9 g, 17.5 mmol) were stirred in THF (25 mL) for 21 h at rt. The solution was then quenched with NH_4Cl (50 mL), and the product was extracted with DCM (50 × 3 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 cm × 16 cm, 40% EtOAc/hexanes; R_f = 0.28) gave alcohol **3.10-(*Z*)** as a colorless oil (700 mg, 94%): ^{1}H NMR ($CDCl_3$, 400 MHz): δ 4.96 (bs, 1H), 3.84 (d, *J* = 20.0 Hz, 2H), 3.63 (dd, *J* = 10.7 Hz, 5.1 Hz, 1H), 3.50 (dd, *J* = 10.7 Hz, 7.4 Hz, 1H), 3.00 – 2.89 (m, 1H), 2.37 – 2.22 (m, 2H), 2.18 (bs, 1H), 1.84 – 1.52 (m, 4H), 1.40 (s); $^{13}C\{^1H\}$ NMR ($CDCl_3$, 100 MHz): δ 155.8, 149.7 (d, $^1J_{C,F}$ = 259.3 Hz), 122.2 (d, $^2J_{C,F}$ = 14.3 Hz), 79.9, 64.4 (d, $^4J_{C,F}$ = 3.8 Hz), 43.7, 40.0 (d, $^2J_{C,F}$ = 29.3 Hz), 29.7, 28.7 (d, $^3J_{C,F}$ = 5.1 Hz), 28.5, 25.0; ^{19}F NMR ($CDCl_3$, 376 MHz): δ –116.8 (dd, $^3J_{H,F}$ = 20, 20 Hz). HRMS (ESI$^+$/TOF) m/z: $[M + H]^+$ Calcd for $C_{19}H_{37}FNO_3Si^+$ 260.1656; Found 260.1644. $[\alpha]_D^{21}$ – 8.83° (*c* 4, MeOH).

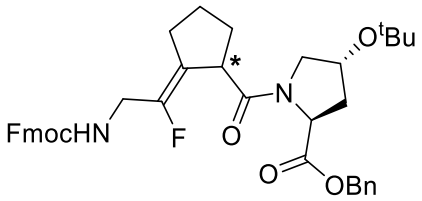


Acid 3.11-(*Z*). Jones reagent (2M, 2 mL, 4 mmol) was added dropwise to a solution of alcohol **3.10-(*Z*)** (300 mg, 1.2 mmol) in acetone (60 mL) at 0 °C and stirred at 0 °C for 30 min. The reaction was quenched with *i*-PrOH (50 mL) and stirred at rt for an additional 10 min, then filtered through Celite and concentrated. The residue was diluted with EtOAc (25 mL) and washed with water (4 × 10 mL) and brine (1 × 10 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (3 cm × 16 cm, 5% MeOH/DCM; R_f = 0.46) to give acid **3.11-(*Z*)** as a pale-yellow oil (280 mg, 89%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 10.85 (br s, 1H), 4.81 (br s, 1H), 4.06 – 3.76 (m, 2H), 3.60 – 3.52 (m, 1H), 2.50 – 2.31 (m, 2H), 2.10 – 1.83 (m, 3H), 1.76 – 1.63 (m, 1H), 1.44 (s, 9H). $^{13}C\{^1H\}$ ($CDCl_3$, 100 MHz): δ 179.7, 155.8, 150.6 (d, $^1J_{C,F}$ = 250 Hz),

120.2 (d, $^{2}J_{C,F}$ = 15 Hz), 80.0, 45.5, 39.7 (d, $^{2}J_{C,F}$ = 28 Hz), 31.8, 28.5, 28.4, 25.9. ^{19}F NMR ($CDCl_3$, 376 MHz): δ –111.4 (dd, $^{3}J_{H,F}$ = 20, 20 Hz). HRMS (ESI^{+}/TOF) m/z: [M + Na]$^{+}$ Calcd for $C_{13}H_{20}FNO_4Na^{+}$ 296.1274; Found 296.1277. $[\alpha]_D^{24}$ –2.54° (*c* 18, $CHCl_3$).



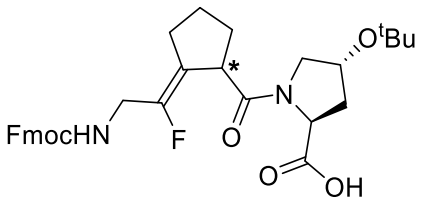
Fmoc Amine 3.12-(*Z*). To a solution of acid **3.11-(*Z*)** (270 mg, 0.98 mmol) in 96:2:2 DCM:H_2O:Et_3SiH (11 mL) was added TFA (4 mL) and stirred at rt for 40 min. The reaction was concentrated and residual TFA was removed by high vacuum overnight. The resultant dark yellow oil was dissolved in $NaHCO_3$ (20 mL) and stirred for 5 min, then a solution of Fmoc–OSu (500 mg, 1.5 mmol) in THF (10 mL) was added and stirred at rt for 16 h. The solution was acidified to pH = 2 with 1 M HCl and the product was extracted with DCM (3 × 30 mL). The combined organic layers were washed with water (2 × 30 mL) and brine (1 × 30 mL), dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 cm × 15 cm, 2% AcOH/DCM; R_f = 0.14) gave Fmoc amine **3.12-(*Z*)** as a white solid (330 mg, 84% yield), mp 68 – 70 ºC. ^{1}H NMR ($CDCl_3$, 400 MHz): 10.51 (br s, 1H), 7.76 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 8.0 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.5 Hz, 2H), 5.24 – 5.16 (m, 1H), 4.55 – 4.34 (m, 2H), 4.28 – 4.16 (m, 1H), 4.15 – 4.00 (m, 1H), 3.96 – 3.81 (m, 1H), 3.63 – 3.42 (m, 1H), 2.62 – 2.22 (m, 2H), 2.18 – 1.79 (m, 3H), 1.77 – 1.55 (m, 1H). ^{13}C{^{1}H} NMR ($CDCl_3$, 100 MHz): 179.7, 156.4, 150.2 (d, $^{1}J_{C,F}$ = 250 Hz), 143.9, 141.4, 127.8, 127.1, 125.1, 120.6 (d, $^{2}J_{C,F}$ = 15 Hz), 120.7, 67.1, 47.2, 45.5, 40.1 (d, $^{2}J_{C,F}$ = 29 Hz), 31.8, 28.5, 25.9. ^{19}F NMR ($CDCl_3$, 376 MHz): –111.7 (dd, $^{3}J_{H,F}$ = 20, 20 Hz). HRMS (ESI^{-}/TOF) m/z: [2M – H]$^{-}$ Calcd for $C_{46}H_{43}F_2N_2O_8^{-}$ 789.2993; Found 789.2988. $[\alpha]_D^{24}$ – 1.61° (*c* 0.14, $CHCl_3$).



3.13-(*R*,*Z*), 3.13-(*S*,*Z*)

Benzyl Ester 3.13-(*Z*). HBTU (75 mg, 200 μmol), 6-Cl-HOBT (35 mg, 200 μmol), and Fmoc amine **3.12-(*Z*)** (52 mg, 130 μmol) were stirred in DCM (2.5 mL) at 0 °C for 10 min. A solution of H–Hyp(^{t}Bu)–OBn **3.15** (55 mg, 200 μmol) in DCM (0.5 mL) was added and the reaction was stirred at rt for 8 h. It was then filtered through Celite and concentrated, diluted with EtOAc (5 mL), washed with 0.01 M HCl (2 × 2 mL), $NaHCO_3$ (2 × 2 mL), and brine (2 × 2 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (2 cm × 15 cm, 2.5% MeOH in DCM, R_f = 0.18) to give a diastereotopic mixture of **3.13-(*Z*)** as a pale-yellow oil (71 mg, 82%). The mixture of diastereomers (100 mg) was separated by HPLC by injecting 20 μL aliquots in *i*-PrOH on a PrincetonSFC's pyridine column (60 Å, 5μ, 250 × 4.6 mm) with TharSFC Fluid Delivery Module with eluent isocratic 15% *i*-PrOH/$scCO_2$ at 3.0 mL/min, 120 bar system back pressure, column oven temperature at 40 °C. The diastereomers were obtained as off-white solids: **3.13-(*S*,*Z*)** (retention time 11.4 min, 10 mg, 14% recovery) and **3.13-(*R*,*Z*)** (retention time 13.9 min, 4 mg, 14% recovery) in a 28:72 *R*:*S* ratio (Figure B.1). **Benzyl Ester 3.13-(*R*,*Z*):** ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.58 (d, *J* = 7.12 Hz, 2H), 7.40 (t, *J* = 7.4 Hz, 2H), 7.37 – 7.27 (m, 7H), 5.21 (dd, *J* = 12.4, 10.3 Hz, 1H), 5.08 (dd, *J* = 14.9, 12.4 Hz, 1H), 5.03 – 4.93 (m, 0.77 H), 4.75 (t, *J* = 5.5 Hz, 0.55H), 4.51 (t, *J* = 7.8 Hz, 0.45H), 4.46 – 4.32 (m, 2H), 4.28 – 3.96 (m, 3H), 3.87 – 3.62 (m, 2.4H), 3.58 (dd, *J* = 10.2, 6.4 Hz, 0.45H), 3.44 – 3.32 (m, 1.55H), 2.50 – 2.29 (m, 3.8H), 2.08 – 1.52 (m, 4H), 1.17 (s, 4H), 1.12 (s, 5H). ^{13}C{^{1}H} NMR ($CDCl_3$, 125 MHz): δ 174.0, (172.8), 172.1, (171.5), 156.4, 149.0 (d, $^1J_{C,F}$ = 246 Hz), (144.03), 143.99, 141.5, 136.1, 135.5, 128.8, (128.6), 128.53, (128.48), 128.2, (128.1), 127.8, 127.2, 125.2,

123.8 (d, $^{2}J_{C,F}$ = 14 Hz), (122.0 (d, $^{2}J_{C,F}$ = 16 Hz)), 120.1, (74.24), 74.18, 69.7, 68.2, (67.1), (67.0), 66.7, (58.80), 58.76, 54.4, (54.0), 47.3, 44.3, (44.2), 40.3 (d, $^{2}J_{C,F}$ = 6 Hz), (40.2, (d, $^{2}J_{C,F}$ = 6 Hz)), 39.5, 37.3, 32.2, (32.0), 31.3, (30.8), 29.9, (29.7), 29.4, (28.9), (28.4), 28.3, 26.3, (26.1). ^{19}F NMR ($CDCl_3$, 376 MHz): δ (–76.0 (s)), (–111.8 (dd, $^{3}J_{H,F}$ = 20, 20 Hz)), –115.7 (dd, $^{3}J_{H,F}$ = 20, 20 Hz). HRMS (ESI$^+$/TOF) m/z: [M + H]$^+$ Calcd for $C_{39}H_{44}FN_2O_6^+$ 655.3178; Found 655.3184. $[\alpha]_D^{24}$ + 0.077° (*c* 0.32, $CHCl_3$). **Benzyl Ester 3.13-(*S*,*Z*)**: ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.76 (d, *J* = 7.4 Hz, 2H), 7.59 (d, *J* = 7.4 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.37 – 7.27 (m, 7H), 5.24 – 5.01 (m, 3H), 4.81 – 4.72 (m, 0.1H), 4.64 (dd, *J* = 8.7, 4.2 Hz, 0.5H), 4.61 – 4.52 (m, 0.4H), 4.45 – 3.72 (m, 7H), 3.69 – 3.59 (m, 0.9H), 3.47 (dd, *J* = 9.7, 6.0 Hz, 0.5H), 3.42 – 3.32 (m, 1.1H), 2.51 – 2.23 (m, 2.5H), 2.23 – 2.19 (m, 0.15H), 2.18 – 1.49 (m, 6.5H), 1.15 (s, 9H). ^{13}C{^{1}H} NMR ($CDCl_3$, 125 MHz): δ 173.1, 172.3, 156.4, 149.0 (d, $^{1}J_{C,F}$ = 250 Hz), 144.03, 141.43, (135.92), 135.85, (128.9), 128.7, 128.61, 128.57, 128.4, 128.3, 127.8, 127.2, 122.4 (d, $^{2}J_{C,F}$ = 13 Hz), 120.1, (74.3), 74.2, 69.7, (69.44), 69.41, 67.09, (67.05), 66.94, (66.91), 57.9, 57.6, 57.3, 54.2, 54.0, 53.8, 47.3, 44.4 (d, $^{2}J_{C,F}$ = 6 Hz), 44.2, 40.3, 40.1, 37.1, 36.8, (31.32), 31.28, 30.8, 29.9, 29.1, 28.4, 36.1, (26.0), 25.5. ^{19}F NMR ($CDCl_3$, 376 MHz): δ (–76.0 (s)), –113.1 (dd, $^{3}J_{H,F}$, = 20, 20 Hz), (–114.2 (t, *J* = 20 Hz)). HRMS (ESI$^+$/TOF) m/z: [M + H]$^+$ Calcd for $C_{39}H_{44}FN_2O_6^+$ 655.3178; Found 655.3190. $[\alpha]_D^{24}$ − 0.081° (*c* 0.80, $CHCl_3$).

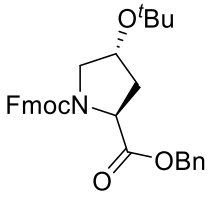


3.1-(*R*,*Z*), 3.1-(*S*,*Z*)

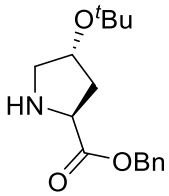
Fmoc–Gly–Ψ[(*Z*)CF=C)]–Pro–Hyp(ᵗBu)–OH, 3.1-(*R*,*Z*) and 3.1-(*S*,*Z*). 1,4-cyclohexadiene (7 μL,70 μmol) was added to a suspension of the tripeptide isostere (4 mg, 6 μmol) and 10% Pd/C (1

mg) in anhydrous EtOH (250 μL) and the solution was heated to reflux at 90 °C for 45 min. After cooling to rt, it was filtered through Celite and diluted with Et_2O (1 mL). The organic phase was washed with 5% $NaHCO_3$ (3 × 0.5 mL) and brine (0.5 mL). The organic layer was dried over Na_2SO_4 and concentrated to recover deprotected tripeptide. The combined aqueous phases were acidified to pH = 2 with 0.1 M HCl, and the product was extracted with EtOAc (4 ×2 mL). The combined organic layers were washed with brine (2 mL), dried over Na_2SO_4, and concentrated to give a white solid (3 mg, 82%). **Acid 3.1-(*R*,*Z*):** ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.58 (d, *J* = 7.5 Hz, 2H), 7.40 (t, J = 7.5 Hz, 2H), 7.31 (t, *J* = 7.4 Hz, 2H), 5.38 – 5.05 (m, 0.64H), 4.70 –3.27 (m, 9.26H), 2.59 – 1.77 (m, 11H), 1.20 (s, 9H). $^{13}C\{^1H\}$ NMR ($CDCl_3$, 125 MHz): δ 176.7, 172.7, 156.5, (156.6), 149.7 (d, $^1J_{C,F}$ = 247 Hz) 143.98, (143.95), 141.44, (141.43), 127.9, (127.23), 127.21, (125.3), 125.1, 121.8 (d, $^2J_{C,F}$ = 15 Hz), 120.13, (120.10), (74.6), 74.5, 69.3, (69.2), (67.2), 67.1, (58.99), 58.94, 54.0, 49.54, (49.49), 47.3, (47.2), 40.1 (d, $^2J_{C,F}$ = 30 Hz), (33.90), 33.88, (29.9), 28.4, (25.7), 25.0, 22.8. ^{19}F NMR ($CDCl_3$, 376 MHz): δ (–75.8 (s)), –112.6 (dd, $^3J_{H,F}$ = 19, 19 Hz), (–115.1 (dd, $^3J_{H,F}$ = 20, 20 Hz)). HRMS (ESI^+/TOF) m/z: $[M + H]^+$ Calcd for $C_{39}H_{44}FN_2O_6^+$ 565.2708; Found 565.2709. **Acid 3.1-(*S*,*Z*)** (6 mg, 9 μmol) was prepared with 1,4-cyclohexadiene (10 μL, 110 μmol) and 10% Pd/C (1 mg) using the same method to give a white solid (6 mg, 80%): ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.76 (d, *J* = 7.4 Hz, 2H), 7.58 (d, *J* = 7.2 Hz, 2H), 7.40 (t, *J* = 7.4 Hz, 2H), 7.31 (t, *J* = 7.0 Hz, 2H), 5.33 – 5.05 (m, 0.4H), 4.70 – 4.46 (m, 0.6H), 4.47 – 4.29 (m, 2.4H), 4.27 – 4.18 (m, 1H), 4.16 – 4.00 (m, 0.4H), 3.72 (dd, *J* = 9.8, 6.8 Hz, 0.38H), 3.67 – 3.59 (m, 0.8H), 3.50 – 3.33 (m, 1.4H), 2.59 – 2.25 (m, 3.4H), 2.14 – 1.81 (m, 4H), 1.74 – 1.49 (m, 3.7H), 1.2 (s, 9H). $^{13}C\{^1H\}$ NMR ($CDCl_3$, 125 MHz): δ 176.7, 172.7, 156.5, (156.6), 149.7 (d, $^1J_{C,F}$ = 247 Hz) 144.0, 141.4, 127.9, (127.23), 127.21, (125.3), 125.1, 121.8 (d, $^2J_{C,F}$ = 15 Hz), 120.13, (120.10), (74.6), 74.5, 69.3, (69.2), (67.2), 67.1, (58.97), 58.94, 54.0, 49.54,

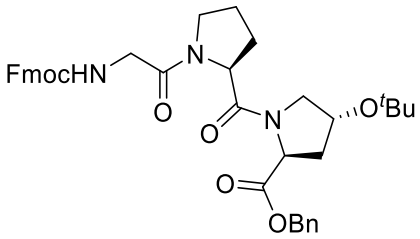
(49.49), 47.3, (47.2), 40.1 (d, $^{2}J_{C,F}$ = 30 Hz), (33.90), 33.88, 32.1, (31.6), (29.9), 28.4, 28.3, (26.3), 26.2, (26.1), 25.7, 25.0, (24.9). ^{19}F NMR ($CDCl_3$, 376 MHz): δ –75.9 (s), (–112.6 (dd, $^{3}J_{H,F}$ = 19, 19 Hz)), (–113.9 (dd, $^{3}J_{H,F}$ = 18, 18 Hz)). HRMS (ESI$^+$/TOF) m/z: $[M + H]^+$ Calcd for $C_{39}H_{44}FN_2O_6^+$ 565.2708; Found 565.2708.



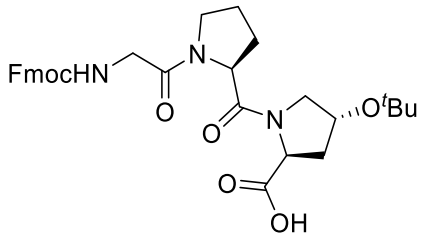
Fmoc–Hyp(ᵗBu)–OBn 3.14. DCC (2.0 g, 9.7 mmol), 6–Cl–HOBt (1.7 g, 9.7 mmol), DMAP (120 mg, 0.97 mmol) and DIEA (3.4 mL, 19 mmol) were added to a solution of Fmoc–Hyp(ᵗBu)–OH (2.0 g, 4.9 mmol) in DCM (90 mL) at 0 ºC and stirred for 10 min. BnOH (1.0 mL, 9.7 mmol) was added and the reaction was stirred at rt for 16 h. The reaction was filtered through Celite and concentrated. The residue was diluted with EtOAc (40 mL), washed with 1 M HCl (2 × 20 mL), $NaHCO_3$ (2 × 20 mL), and brine (20 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 20% EtOAc/hexanes) to give **3.14** as a colorless oil (2.4 g, 99%). ^{1}H NMR matches previous reports[11] ($CDCl_3$, 400 MHz): δ 7.79 – 7.72 (m, 1.85H), 7.63 – 7.50 (m, 2.15H), 7.42 – 7.28 (m, 9H), 5.29 – 5.15 (m, 1H), 5.15 – 5.04 (m, 1H), 4.71 (dd, *J* = 6.0, 1.6 Hz, 1H), 4.55 (dd, *J* = 9.8, 5.5 Hz, 0.5H) 4.49 (dd, *J* = 8.8, 4.6 Hz, 0.5H), 4.43 – 4.21 (m, 3.5H), 4.02 (t, *J* = 7.0 Hz, 0.40H), 3.80 (td, *J* = 11.4, 6.0 Hz, 1H), 3.36 (dd, *J* = 10.7, 5.1 Hz, 1H), 2.31 – 2.00 (m, 2H), 1.70 (t, *J* = 1.3 Hz, 0.5H), 1.18 (s, 5H), 1.16 (s, 4H).



H–Hyp(ᵗBu)–OBn 3.15. Piperidine (9.8 mL) was added to a solution of Fmoc–Hyp(ᵗBu)–OBn **3.14** (2.4 g, 4.9 mmol) in DCM (40 mL) at rt and stirred for 40 min. The reaction was then concentrated and purified by SiO_2 chromatography (5 × 15 cm, 1:2 EtOAc:hexanes 500 mL, then 5% MeOH/$CHCl_3$ 1.2 L) to give **3.15** as a pale-yellow oil (1.1 g, 81%). The ^{1}H NMR matches the literature.[11] ($CDCl_3$, 400 MHz): δ 7.38 – 7.30 (m, 5H), 5.15 (d, *J* = 1.2 Hz, 2H), 4.17 – 4.1 (m, 1H), 3.97 (t, *J* = 7.7 Hz, 1H), 3.19 (dd, *J* = 11.2, 5.7 Hz, 1H), 2.77 (dd, *J* = 11.2, 4.3 Hz, 1H), 2.3 (br s, 1H), 2.03 (dd, *J* = 7.6, 5.5 Hz, 2H), 1.14 (s, 9H).



Fmoc–Gly–Pro–Hyp(ᵗBu)–OBn 3.16. The Fmoc–Gly–Pro–OSu unit was prepared by the method developed by Ottl.[36] H–Hyp(ᵗBu)–OBn **3.15** (730 mg, 2.6 mmol) in DMF (5 mL) was added to a solution of Fmoc–Gly–Pro–OSu (1.3 g, 2.6 mmol) and DIEA (225 μL, 1.3 mmol) in DMF (5 mL) and stirred at rt for 8 h. The reaction was then diluted with EtOAc (25 mL) and washed with 5% $NaHCO_3$ (3 × 10 mL), 0.1 M HCl (3 × 10 mL), H_2O (3 × 10 mL) and brine (10 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 cm × 15 cm, 2% MeOH/DCM) to give **3.16** as a pale-yellow oil (1.1 g, 59%). The ^{1}H NMR matches the literature.[11] ($CDCl_3$, 400 MHz): δ 7.68 –7.49 (m, 4H), 7.32 – 7.18 (m, 9H), 5.93 – 5.77 (m, 1H), 5.16 – 4.96 (m, 3H), 4.62 (dd, *J* = 8.7, 4.2 Hz, 0.7H), 4.58 – 4.52 (m, 1H), 4.30 – 4.20 (m, 2.5H), 4.17 – 3.84 (m, 4.3H), 3.70 – 3.59 (m, 1.2H), 3.54 – 3.44 (m, 1.5H), 3.18 (br s, 1.2H), 3.11 (dd, *J* = 11.2, 5.6 Hz, 0.6H), 2.70 – 2.68 (m, 1H), 2.03 – 1.95 (m, 2.8H), 1.87 – 1.77 (m, 1.7H), 1.08 (s, 4.5H), 1.08 (s, 4.5H).



Fmoc–Gly–Pro–Hyp(ᵗBu)–OH 3.17. A flask containing 10% Pd/C (100 mg) in MeOH (40 mL) was charged with H_2. A solution of tripeptide **3.16** (1.0 g, 1.5 mmol) in MeOH (14 mL) was added slowly and the reaction was allowed to stir for 16 h at rt. The reaction mixture was filtered through Celite and concentrated to give **3.17** as a white solid (870 mg, > 99%). The ^{1}H NMR matches the literature.[11] ($CDCl_3$, 400 MHz): δ 10.01 (br s, 1H), 7.73 (d, *J* = 7.5 Hz, 2H), 7.59 (t, *J* = 6.3 Hz, 2H), 7.36 (t, *J* = 7.5 Hz, 2H), 7.28 (t, *J* = 7.2 Hz, 2H), 6.40 – 5.90 (m, 1H), 4.68 – 4.48 (m, 2H), 4.48 – 4.15 (m, 4.4H), 4.15 – 4.00 (m, 1H), 4.00 – 3.83 (m, 1H), 3.82 – 3.13 (m, 4.6H), 2.26 – 1.83 (m, 6H), 1.15 (s, 6.8H), 1.14 (s, 2.2H).

Solid-Phase Peptide Synthesis. Ac–(Gly–Pro–Hyp)$_3$–Gly-Ψ[(*Z*)CF=C]–L-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2 **3.2-(*R*,*Z*)**, Ac–(Gly–Pro–Hyp)$_3$–Gly-Ψ[(*Z*)CF=C]–D-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2 **3.2-(*S*,*Z*)**, and Ac–(Gly–Pro–Hyp)$_8$–Gly–Gly–Tyr–NH_2 **3.3**. All reactions were done by shaking at 30 °C. Rink amide MBHA resin (100 mg, 0.37 mmol/g) was placed in a 10 mL polypropylene tube and swollen with DCM (5 mL) for 1 h. The resin was filtered and washed with NMP (3 × 5 mL) between every step. The Fmoc group was then removed by shaking with 20% piperidine in NMP (5 mL) for 40 min. The deprotection step was repeated. To couple single amino acids (Gly, Gly, and Tyr) 6–Cl–HOBt (20 mg, 0.12 mmol), HBTU (44 mg, 0.12 mmol), DIEA (40 μL, 0.23mmol), and the Fmoc-amino acid (0.12 mmol) were dissolved in NMP (5 mL) and shaken with the resin for 30 min. Each coupling was repeated a second time. The first four repeats of Fmoc–Gly–Pro–Hyp(ᵗBu)–OH (65 mg, 0.12 mmol) dissolved in NMP (5 mL) were coupled to the resin with 6–Cl–HOBt (20 mg, 0.12 mmol), HBTU (44 mg, 0.12 mmol), and

DIEA (40 μL, 0.23 mmol), by shaking for 30 min. Each coupling was repeated a second time. After every coupling, the peptide was capped by shaking with 10% Ac_2O and 10% DIEA in DCM (5 mL) for 20 min once. Each Fmoc group was removed by shaking with 20% piperidine in NMP (5 mL) for 10 min, then again for 20 min. At the end of every day, the resin was shrunken with MeOH, dried *in vacuo*, stored at 4 °C overnight, and swollen with DCM in the morning. The resin was then separated into three 10 mL polypropylene tubes.

3.2-(*R,Z*). Fluoro-alkene isostere **3.13-(*R,Z*)** (3 mg, 0.005 mmol), HOAt (2 mg, 0.016 mmol), HATU (6 mg, 0.016 mmol) and 2,4,6-collidine (4 μL, 0.032 mmol) were dissolved in NMP (0.4 mL) and shaken with a portion of the H–(Gly–Pro–Hyp(^{t}Bu))$_4$–Gly–Gly–Tyr–● resin (13 mg) for 2 h once. The Fmoc group was removed by shaking with 10% piperidine in NMP (0.4 mL) for 10 min, then again for 20 min. The final three repeats of Fmoc–Gly–Pro–Hyp(^{t}Bu)–OH (9 mg, 0.016 mmol) dissolved in NMP (0.4 mL) were coupled to the resin with HOAt (2 mg, 0.016 mmol), HATU (6 mg, 0.016 mmol), and DIEA (5 μL, 0.029 mmol), by shaking for 30 min. Each coupling was repeated. Each Fmoc group was removed by shaking with 10% piperidine in NMP (0.4 mL) for 10 min, then again for 20 min. The peptide was capped by shaking with 10% Ac_2O and 10% DIEA in DCM (0.4 mL) for 20 min. The peptide was then cleaved from the resin by shaking with 2% H_2O and 3% triethyl silane in TFA (1 mL) for 3.5 h.

3.2-(*S,Z*). Fluoro-alkene isostere **3.13-(*S,Z*)** (8 mg, 0.014 mmol), HOAt (6 mg, 0.043 mmol), HATU (16 mg, 0.043 mmol), and 2,4,6-collidine (11 μL, 0.085 mmol) were dissolved in NMP (1 mL) and shaken with a portion of the H–(Gly–Pro–Hyp(^{t}Bu))$_4$–Gly–Gly–Tyr–● resin (24 mg) for 2 h once. The Fmoc group was removed by shaking with 10% piperidine in NMP (1 mL) for 10 min, then again for 20 min. The final three repeats of Fmoc–Gly–Pro–Hyp(^{t}Bu)–OH (24 mg, 0.43 mmol) in NMP (1 mL) were coupled to the resin with HOAt (6 mg, 0.043 mmol), HATU (16 mg,

0.043 mmol), and DIEA (10 μL, 0.053 mmol), by shaking for 30 min. Each coupling was repeated. Each Fmoc group was removed by shaking with 10% piperidine in NMP (1 mL) for 10 min, then again for 20 min. The peptide was capped by shaking with 10% Ac_2O and 10% DIEA in DCM (1 mL) for 20 min. The peptide was then cleaved from the resin by shaking with 2% H_2O and 3% triethyl silane in TFA (2.5 mL) for 3.5 h.

Native 3.3. The final five repeats of Fmoc–Gly–Pro–Hyp(^{t}Bu)–OH (27 mg, 0.016 mmol) in NMP (3 mL) were coupled to a portion of the H–(Gly–Pro–Hyp(^{t}Bu))$_4$–Gly–Gly–Tyr–● resin (30 mg) with HOAt (7 mg, 0.048 mmol), HATU (18 mg, 0.048 mmol), and DIEA (17 μL, 0.0097 mmol), by shaking for 30 min. Each coupling was repeated. Each Fmoc group was removed by shaking with 10% piperidine in NMP (3 mL) for 10 min, then again for 20 min. The peptide was capped by shaking with 10% Ac_2O and 10% DIEA in DCM (3 mL) for 20 min. The peptide was then cleaved from the resin by shaking with 2% H_2O and 3% triethyl silane in TFA (7.5 mL) for 3.5 h.

Purification and Characterization of Peptides

All peptides were filtered and concentrated, then precipitated with cold Et_2O (ca. 3 mL) from a minimal amount of MeOH (ca. 1 mL). Each peptide was purified by reverse phase HPLC by injecting 20 μL aliquots in MeOH on a Kinetex 5 μm C18 column (250 × 10 mm) with an Agilent 1200 Series Gradient HPLC System using solvents A: 0.1% F.A. H_2O and B: 0.1% F.A. in CH_3CN with 10% B for 5 min, then 10% – 54% B over 15 min at 3 mL/min, and UV detection at 280 nm. The diastereomers were obtained as white solids: **3.2-(*R,Z*)** (retention time 12.2 min, 1 mg, 11% recovery), **3.2-(*S,Z*)** (retention time 12.2 min, 4 mg, 8% recovery), and **3.3** (retention time 12.1 min, 0.5 mg, 3% recovery) (Figure B.2). ^{19}F NMR data was obtained on both fluoro-alkene collagen-like peptides before and after incubation at 4 °C for 72 h. **3.2-(*R,Z*)** (Pre-Incubation) ^{19}F NMR ($CDCl_3$, 376 MHz): δ (–75.6 (s)), (–75.7 (br s)), –122.4 (s), (–150.58 (s)), (–150.63 (s)); **3.2-**

(*R,Z*) (Post-Incubation) ^{19}F NMR ($CDCl_3$, 376 MHz): δ –77.0 (s); **3.2-(*S,Z*)** (Pre-Incubation) ^{19}F NMR ($CDCl_3$, 376 MHz): δ –77.0 (s); **3.2-(*S,Z*)** (Post-Incubation) ^{19}F NMR ($CDCl_3$, 376 MHz): δ –75.7 (d, J = 1.3 Hz).

Table 3.1. MALDI-TOF results of peptides **3.2-(*R,Z*)**, **3.2-(*S,Z*)**, and **3.3**.

	$[M + H]^+$		$[M + Na]^+$		$[M + K]^+$	
	Calcd[a]	Found	Calcd[a]	Found	Calcd[a]	Found
3.2-(R,Z)	2474.1190	N.D.	2497.1088	2497.0881	2513.0827	2514.0687
3.2-(S,Z)	2474.1190	2475.1099	2497.1088	2497.0918	2513.0827	2513.0670
3.3	2474.1259	2474.1081	2496.1084	2496.0927	2512.0823	2512.0674

[a]The calculated mass is calculated for the most abundant isotopes. N.D. = not detected.

Circular Dichroism

The concentration of the peptide in PBS (10 mM Na_xPO_4, 150 mM NaCl, pH 7.0) was determined by the UV absorption of the Tyr residue (ε = 1490 $M^{-1}\cdot cm^{-1}$) at 280 nm. The concentration of **3.2-(*R,Z*)** was 0.097 mM , **3.2-(*S,Z*)** was 1.5 mM, and **3.3** was 0.53 mM. The peptides were incubated at 4 °C for 72 h. The CD spectra were obtained in 0.5 nm increments, 1 nm bandwidth, and 1 mm pathlength at a scan speed of 100 nm/min. The spectra were averaged over three consecutive scans, and blank buffer scans were subtracted from the baseline. Full scan spectra from 190 to 300 nm were obtained from 5 °C to 85 °C (Figure 3.4), heated in 5 °C increments with a 5 min equilibrium time at each temperature change. The ellipticity at 226 nm was monitored at each temperature. The T_m of **2-(R,Z)** and control **3** were calculated by fitting the data to the following four-parameter Hill equation using SigmaPlot 10 (Figure 5):

$$F = \frac{\mathrm{F_0} + (\mathrm{F}_{max} - \mathrm{F}_{min})}{(1 + \exp(-(T - T_m)/b))} \quad (1)$$

$$F = \frac{[\theta] - [\theta]_u}{[\theta]_u - [\theta]_f} \quad (2)$$

where [θ] is the measured ellipticity at 226 nm and temperature T, $[\theta]_u$ is the ellipticity at 85 °C, $[\theta]_f$ is the ellipticity at 5 °C, F is the fraction of unfolded peptide, b is the inflection point, and T_m is the calculated melting temperature.

3.5 Conclusions

A (*Z*)-fluoro-alkene isostere of Gly–Pro was designed and synthesized, and its stabilization of a collagen-like triple helix peptide was measured by CD. The Fmoc–Gly–Ψ[(*Z*)CF=C)]–L-Pro–Hyp–OH fluoro-alkene isostere **3.13-(*R,Z*)** was made in 12 steps and 3.5% overall yield, as well as the diastereomer Fmoc–Gly–Ψ[(*Z*)CF=C)]–D-Pro–Hyp–OH fluoroalkene isostere **3.13-(*S,Z*)** in 12 steps and 8.7% overall yield. The asymmetric aldol addition to introduce the first stereocenter using L-Thr as a chiral catalyst[25] gave the undesired D-Pro mimic as the major diastereomer. Formation of the (*Z*)-fluoro-alkene moiety was enhanced over the (*E*)-fluoro-alkene by using MeMgCl as the base, instead of NaH, in a fluorophosphonate ylide addition to the ketone.[28] Diastereomer separation was accomplished by HPLC using 15% *i*PrOH in $scCO_2$ on a pyridine column.

Collagen-like peptides Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C)]–L-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2 **3.2-(*R,Z*)**, Ac–(Gly–Pro–Hyp)$_3$–Gly–Ψ[(*Z*)CF=C)]–D-Pro–Hyp–(Gly–Pro–Hyp)$_4$–Gly–Gly–Tyr–NH_2 **3.2-(*S,Z*)**, and Ac–(Gly–Pro–Hyp)$_8$–Gly–Gly–Tyr–NH_2 **3.3** were made by solid-phase peptide synthesis. All peptides were subjected to thermal melting and CD was used to measure the signature triple-helix maximum at 226 nm. Peptide **3.2-(*R,Z*)** was found to have T_m = 42.2 °C compared to **3.3** with T_m = 49.1 °C, while **3.2-(*S,Z*)** had a linear relationship between ellipticity and temperature. Stereochemistry was assigned based on the lack of a sigmoidal melting curve for the D-Pro mimic isomer. The enhanced thermostability of the fluoro-alkene peptide compared to the proteo-alkene peptide[11] confirms our theoretical results[19]

suggesting the flexible Ψ angle of the proteoalkene destabilized the collagen triple helix; the fluoro-alkene showed much less destabilization. The Gly–Pro (*Z*)-fluoro-alkene is not an n→π* acceptor,[24] and our computational results had suggested it would not be an n→π* donor.[19] However, the significant deshielding of the fluorine nucleus as detected by ^{19}F NMR may be indicative of n→π* donation that confers a degree of stability on the collagen triple helix by mimic **3.2-(*R*,*Z*).**

3.6 References

1. Ramachandran, G. N.; Mitra, A. K., Biochemistry of Collagen. *J. Mol. Biol.* **1976,** *107*, 85–92.

2. Rich, A.; Crick, F. H. C., The Molecular Structure of Collagen. *J. Mol. Biol.* **1961,** *3*, 483-506.

3. Brodsky, B.; Shah, N. K., The Triple-Helix Motif in Proteins. *FASEB* **1995,** *9* (15), 1537-1546.

4. Shoulders, M. D.; Raines, R. T., Collagen Structure and Stability. *Annu Rev. Biochem.* **2009,** *78*, 929-958.

5. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 A Resolution. *Science* **1994,** *226*, 75-81.

6. Pauling, L.; Corey, R. B.; Branson, H. R., The Structure of Proteins Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U.S.A.* **1951,** *37*, 205-211.

7. Brandts, J. F.; Halvorson, H. R.; Brennan, M., Consideration of the Possibility that the Slow Step in Protein Denaturation Reactions is Due to Cis-Trans Isomerism of Proline Residues. *Biochemistry* **1975,** *14*, 4953-4963.

8. Bachinger, H., P.; Engel, J.; Bruckner, P.; Timpl, R., The Role of Cis-Trans Isomerization of Peptide Bonds in the Coil ⇄ Triple Helix Conversion of Collagen. *Eur. J. Biochem.* **1978,** *90* (3), 605-613.

9. Etzkorn, F. A.; Zhao, S., Stereospecific Phosphorylation by the Central Mitotic Kinase Cdk1-Cyclin B. *ACS Chem. Biol.* **2015,** *10* (4), 952-956.

10. Mayfield, J. E.; Fan, S.; Wei, S.; Zhang, M.; Li, B.; Ellington, A. D.; Etzkorn, F. A.; Zhang, Y. J., Chemical Tools To Decipher Regulation of Phosphatases by Proline Isomerization on Eukaryotic RNA Polymerase II. *ACS Chem Biol* **2015,** *10* (10), 2405-14.

11. Dai, N.; Wang, X. J.; Etzkorn, F. A., The Effect of a Trans-Locked Gly-Pro Alkene Isostere on Collagen Triple Helix Stability. *J. Am. Chem. Soc.* **2008,** *130*, 5396-5397.

12. Dai, N.; Etzkorn, F. A., Cis-Trans Proline Isomerization Effects on Collagen Triple-Helix Stability are Limited. *J. Am. Chem. Soc.* **2009,** *131*, 13728-13732.

13. DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L., Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **2002,** *124*, 2497-2505.

14. Newberry, R. W.; Raines, R. T., The n→pi* Interaction. *Acc. Chem. Res.* **2017,** *50* (8), 1838-1846.

15. Etzkorn, F. A.; Ware, R. I.; Pester, A. M.; Troya, D., Conformational Analysis of n→π* Interactions in Collagen Triple Helix Models. *J. Phys. Chem. B.* **2019,** *123*, 496-503.

16. Hinderaker, M. P.; Raines, R. T., An Electronic Effect on Protein Structure. *Protein Sci.* **2003,** *12*, 1188-1194.

17. Bürgi, H. B.; Dunitz, J. D.; Shefter, E., Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group. *J. Am. Chem. Soc.* **1973,** *95*, 5065-5067.

18. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T., n→π* Interactions of Amides and Thioamides: Implications for Protein Stability. *J. Am. Chem. Soc.* **2013,** *135*, 7843-7846.

19. Arcoria, P. J.; Ware, R. I.; Makwana, S. V.; Troya, D.; Etzkorn, F. A., Conformational Analysis of Fluoro-, Chloro-, and Proteo-Alkene Gly–Pro and Pro–Pro Isosteres to Mimic Collagen. *The Journal of Physical Chemistry B* **2021**.

20. Drouin, M.; Paquin, J., Recent Progress in the Racemic and Enantioselective Synthesis of Monofluoroalkene-Based Dipeptide Isosteres. *Beilstein J Org Chem* **2017,** *13*, 2637-2658.

21. Bondi, A., Van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68*, 441-451.

22. Urban, J. J.; Tillman, B. G.; Cronin, W. A., Fluoroolefins as Peptide Mimetics: A Computational Study of Structure, Charge Distribution, Hydration, and Hydrogen Bonding. *J. Phys. Chem. A* **2006,** *110*, 11120-11129.

23. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A., Crystal Structure of the Collagen Triple Helix Model $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$. *Protein Sci.* **2002,** *11*, 262-270.

24. Jakobsche, C. E.; Choudhary, A.; Miller, S. J.; Raines, R. T., n -> π* Interaction and n)(π Pauli Repulsion Are Antagonistic for Protein Stability. *J. Am. Chem. Soc.* **2010,** *132*, 6651–6653.

25. Chen, A.; Xu, J.; Chiang, W.; Chai, C. L. L., L-Threonine-Catalysed Asymmetric α-Hydroxymethylation of Cyclohexanone: Application to the Synthesis of Pharmaceutical Compounds and Natural Products. *Tetrahedron* **2010,** *66*, 1489-1495.

26. Van der Veken, P.; Kertèsz, I.; Senten, K.; Haemers, A.; Augustyns, K., Synthesis of (E)- and (Z)-Fluoro-Olefin Analogues of Potent Dipeptidyl Peptidase IV Inhibitors. *Tetrahedron Lett.* **2003,** *44*, 6231-6234.

27. Sano, S.; Ando, T.; Yokoyama, K.; Nagao, Y., New Reaction Mode of the Horner-Wadsworth-Emmons Reaction for the Preparation of α-Fluoro-α,β-unsaturated Esters. *Synlett* **1998,** *1998* (07), 777-779.

28. Sano, S.; Matsumoto, T.; Nanataki, H.; Tempaku, S.; Nakao, M., Z-Selective Horner–Wadsworth–Emmons reaction of 2-TOM-cyclopentanone for the synthesis of rac-N-Cbz-Gly-Ψ[(Z)-CFC]-Pro-OH dipeptide isostere. *Tetrahedron Lett.* **2014,** *55* (45), 6248-6251.

29. Cahn, R. S.; Ingold, C.; Prelog, V., Specification of Molecular Chirality. *Angew. Chem., Int. Ed. Engl.* **1966,** *5* (4), 385-415.

30. Fukuyama, T.; Cheung, M.; Kana, T., N-Carboalkoxy-2-Nitrobenzenesulfonamides A Practical Preparation of N-Boc-, N-Alloc, and N-Cbz-Protected Primary Amines. *Syn. Lett.* **1999,** *8*, 1301-1303.

31. Ottl, J.; Jürgen Musiol, H.; Moroder, L., Heterotrimeric Collagen Peptides Containing Functional Epitopes. Synthesis of Single-Stranded Collagen Type I Peptides Related to the Collagenase Cleavage Site. *Journal of Peptide Science* **1999,** *5* (2), 103-110.

32. Shah, N. K.; Brodsky, B.; Kirkpatrick, A.; Ramshaw, J. A., Structural Consequences of D-Amino Acids in Collagen Triple-Helical Peptides. *Biopolymers* **1999,** *49* (4), 297-302.

33. Bellenie, B., R.; Carter, M., K.; Cheung, K., M., J.; Davis, O., A.; Hoelder, S.; Lloyd, M., G.; Varela, R., A. Benzimidazolone Derived Inhibitors of BCL6. WO 2018/215801 A1, 2018.

34. Van der Veken, P.; Senten, K.; Kertesz, I.; De Meester, I.; Lambeir, A.; Maes, M.; Scharpe, S.; Haemers, A.; Augustyns, K., Fluoro-Olefins as Peptidomimetic Inhibitors of Dipeptidyl Peptidases. *J. Med. Chem.* **2005,** *48*, 1768-1780.

35. Barrett, S.; O'Brien, P.; Steffens, C. H.; Towers, T. D.; Voith, M., New Route to 4-aminocyclopent-2-en-1-ols Synthesis andEnantioselective Rearrangement of 4-Amino-substitutedCyclopentene Oxides. *Tetrahedron* **2000,** *56*, 9633-9640.

36. Ottl, J.; Musiol, H. J.; Moroder, L., Heterotrimeric Collagen Peptides Containing Functional Epitopes. Synthesis of Single-Stranded Collagen Type I Peptides Related to the Collagenase Cleavage Site. *J. Pept. Sci.* **1999,** *5* (2), 103-10.

Chapter 4. Isomerization-Locked Alkene Analogues of Ser–Pro for Incorporation into Full-Length Bora Protein

4.1 Introduction

4.1.1 Peptidyl-Prolyl Isomerase, Pin1

Peptidyl-prolyl isomerases (PPIases) are a type of enzyme that catalyzes the cis-trans isomerization of certain Xaa–Pro peptide bonds in both folded and unfolded proteins.[1-3] This unique class of enzyme has evolved due to the nearly isoenergeic cis- and trans- conformations of the prolyl amide, as well as the relatively low barrier between them (ca. 20 kcal/mol).[4] Xaa–Pro peptide isomerization occurs on the order of sec^{-1} at physiological temperature, which can be catalyzed to increase the rate by 10^6 over thermal background with PPIases.[5] Isomerization adds an additional regulatory step that is implicated in many biological processes.[6, 7]

The PPIase Pin1 (Protein Interacting with NIMA #1) belongs to the parvulin family and was discovered in 1996.[3] It was found through its interaction with the kinase NIMA (Never In Mitosis A), which is an essential interaction for the progression of *Aspergillus nidulans* through mitosis.[3] Pin1 is a small enzyme (168 amino acids; ca. 18 kDa) with two domains: An *N*-terminal WW binding domain and a *C*-terminal isomerization domain.[8] The WW domain is thought to be involved with substrate recognition, interaction with anchoring proteins for subcellular localization, and facilitation of nuclear import.[3, 9] The primary substrates of Pin1 are the pSer/pThr–Pro motif in proteins.[10] While it can recognize and isomerize the non-phosphorylated Ser/Thr–Pro sequence, it is highly phosphorylation dependent.[11] This small PPIase is an interesting target because it is also overexpressed in a variety of cancer cell lines.[12, 13] It can also act as a tumor suppressor and has been implicated in other diseases such as Alzheimer's.[12, 14] Pin1 is most active during mitosis where it interacts with, and regulates several critical cell cycle proteins, including

NIMA kinase,[3] Cdc25 phosphatase,[15] Cdc2 kinase,[16] Plk1 phosphatase,[15, 17] and Aurora Borealis (Bora).[18]

4.1.2 The Aurora Borealis Protein

The Bora protein is the partner protein of a key mitotic regulator, Aurora Kinase A (AurA). AurA[19] is a cytoplasmic protein that is responsible for, among other things, chromosome separation and spindle assembly at the G2→M transition of mitosis (Figure 4.1).[20] These important processes require the complexation of AurA with Bora to proceed.

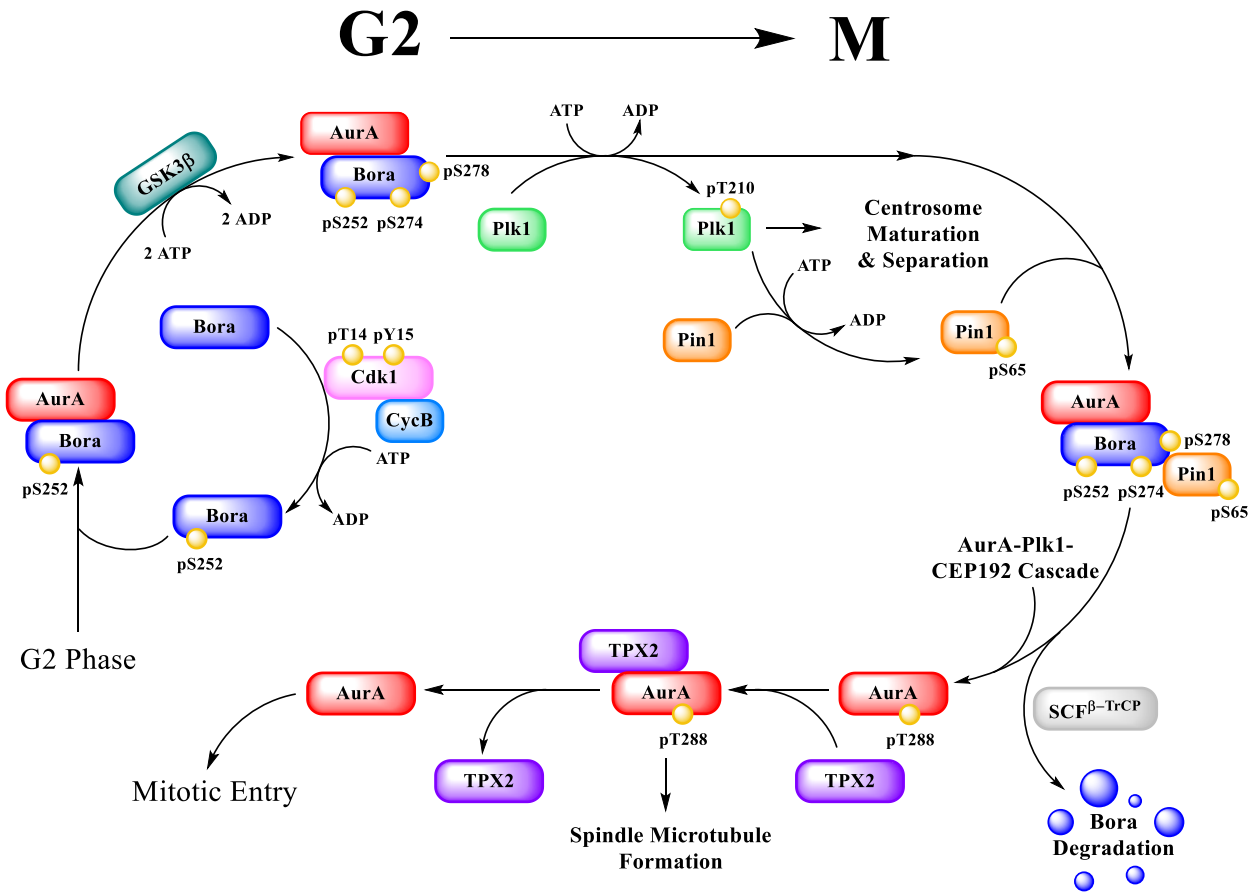


Figure 4.1. Life cycle of the Bora protein during the G2→M transition of the cell cycle. Image was made with Dr. Christina Kim and Karla Piedl.

For most of the cell cycle, Bora is in an inactive state inside the nucleus. During late G2 phase, the Cdk1-CycB complex is activated, phosphorylating Bora on ^{252}Ser and causing it to exit the nucleus into the cytoplasm.[21] There, it complexes with AurA, forming the AurA-Bora complex, which enhances the kinase activity of AurA. Two phosphorylation events occur on the AurA-Bora complex by glycogen synthase kinase 3β (GSK3β) on Bora at positions ^{274}Ser and ^{278}Ser.[22] Both Ser residues are followed by Pro, and they are both pSer–Pro Pin1 binding sites.[18]

The now active AurA-Bora complex phosphorylates the T-loop of Polo Like Kinase1 (Plk1) on ^{210}Thr.[23] AurA can phosphorylate Plk1 in the absence of Bora but is much slower.[23, 24] This event is essential for the maturation and separation of centrosomes, as well as cytokinesis later during the cell cycle. It also serves as a recovery checkpoint allowing the cell to restart the cell cycle following DNA damage.[25] Activated Plk1 is responsible for the phosphorylation of the Cdc25c phosphatase; a crucial step for mitotic entry.[26]

Up until this point, Pin1 is inhibited by a previous phosphorylation on ^{16}Ser by AurA.[18] This phosphate is removed by the phosphatase Protein Phosphatase type 2A (PP2A),[27] allowing it to be activated by Plk1 by phosphorylation on ^{65}Ser.[28, 29] Once activated, Pin1 interacts with Bora at one of the two pSer–Pro binding sites (^{274}Ser–Pro or ^{278}Ser–Pro).[18] This interaction presumably induces a conformational change in Bora through Pin1-mediated isomerization, leading to recognition of Bora by the ubiquitin ligase SCF$^{\beta\text{-TrCP}}$ (Skp, Cullin, F-box containing complex, β-transducin repeat-containing protein).[30] SCF$^{\beta\text{-TrCP}}$ tags Bora with ubiquitin, marking it for degradation by cytoplasmic proteases.[30] Overexpression of Pin1 inhibits the G2→M transition, while its depletion leads to premature entry into mitosis, followed by apoptosis.[3]

The degradation of Bora allows the free AurA kinase to interact with its other protein partners. Specifically, those involved in the AurA-Plk1-Cep192 cascade, which is responsible for

bipolar centrosome placement,[31] and formation of the AurA-TPX2 complex – which is responsible for spindle fiber formation.[32] We hypothesize that without Bora degradation, AurA will be unable to function properly leading to multipolar spindles (Figure 4.2).[33] Abnormal AurA activity has been linked to many different types of cancers.[20]

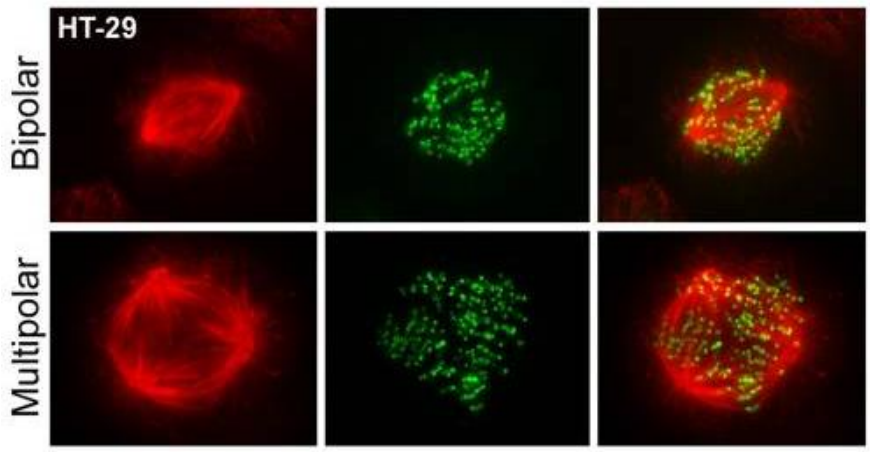


Figure 4.2. Colorectal cancer cells HT-29 prometaphase cells. Immunostaining shows **Left:** α-tubulin **Middle:** CREST. **Right:** Merged images. **Top:** A cell undergoing bipolar spindle formation **Bottom:** A cell undergoing cancerous multipolar spindle formation. Reprinted with permission of PLos One reference [33] under the terms of the CC BY-SA 4.0 license.

4.1.3 Isomerization-Locked Ser–Pro Dipeptides

To elucidate the role that Ser–Pro isomerization has on healthy mitosis, we plan to incorporate cis- and trans-locked Ser–Pro alkene derivatives (Figure 4.3) into a full-length Bora protein. Since the degradation of Bora occurs through Pin1-mediated ubiquitin tagging, by controlling and locking the isomeric state of the Ser–Pro Pin1 binding sites, we can determine the necessary conformation required for healthy bipolar spindle formation in live cells. Our work will pave the way to studying the regulatory role of Xaa–Pro isomerization in other key pathways.

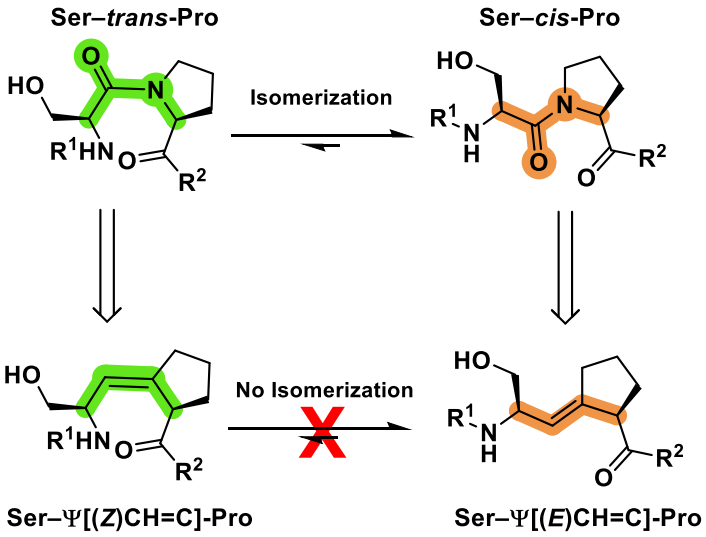


Figure 4.3. Locked alkene analogues are not isomerized under physiological conditions. **(Top left):** Native Ser–Pro in the cis conformation. **(Bottom left):** cis-locked (*Z*)-alkene isostere Ser–Ψ[(*Z*)CH=C]-Pro. **(Top right):** Native Ser–Pro in the trans conformation. **(Bottom right):** trans-locked (*E*)-alkene isostere Ser–Ψ[(*E*)CH=C]-Pro.

Previously, cis- and trans-locked Ser–Pro alkenes have been designed and used to determine the necessary conformation of various peptides.[34] Ser–Pro alkene analogues were used to determine the isomeric state necessary for Pin1 binding. The peptide Ac–FFS(*E*/*Z*)=PR–NH_2 was synthesized based on the affinity of Pin1 for aromatic amino acids at the *N*-terminus and basic residues at the *C*-terminus.[10] It was concluded that Pin1 binding of the cis-locked alkene was about 23 times more potent than that of the trans-locked alkene.[11]

To determine the isomeric state required for the phosphorylation of ^{168}Ser–Pro on Cdc25x by the Cdk1-cycB complex, the 11-mer peptide with the sequence Ac–MKYLGS(*E*/*Z*)=PITTV–NH_2, as well as the native Ac–MKYLGSPITTV–NH_2, were synthesized.[35] The *in vitro* assay showed that Cdk1-cycB preferentially phosphorylated the Ser-*trans*-Pro isomer over the Ser-*cis*-Pro isomer.[35]

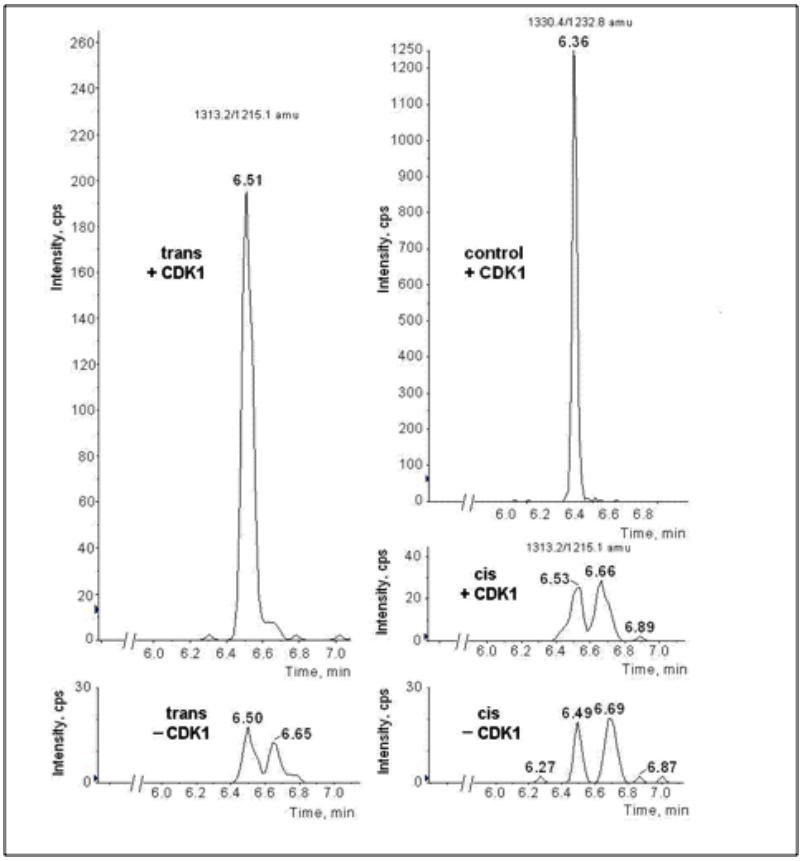


Figure 4.4. Phosphorylation assay of Ac–MKYLG(E/Z)S=PITTV–NH2 by Cdk1-cyclin B complex**.** Loss of H_3PO_4 as monitored by LC-MSMS. Both cis- and trans-locked peptides were measured in the presence (top) and absence (bottom) of Cdk1-cycB.[35] Reprinted with permission from Etzkorn, F. A.; Zhao, S., Stereospecific Phosphorylation by the Central Mitotic Kinase Cdk1-Cyclin B. *ACS Chem. Biol.* **2015,** *10* (4), 952-956. Copyright 2015 American Chemical Society.

4.1.4 i*n vitro* Transcription-Translation

Once the cis- and trans-locked Ser–Pro alkene mimics are synthesized (discussed below), they will be incorporated into a full-length Bora protein through *in vitro* transcription-translation (IVTT, Figure 4.5).

The plan for IVTT starts with modified Bora DNA. This DNA has been altered to include a TAG site at either the ^{274}Ser–Pro or ^{278}Ser–Pro positions. The TAG site is critical, as it allows for our dipeptides to be introduced at that specific location. The DNA is then transcribed by T7

RNA Polymerase, forming the mRNA strand. This mRNA is translated by an S-30 modified ribosome, obtained from Sidney Hecht (Arizona State U.).[36] During translation, the ribosome reads the mRNA and recruits an endogenous tRNA attached to a single amino acid. This modified ribosome has an enlarged pocket allowing for tRNA to incorporate our dipeptide analogues.[36] This is a necessary modification, as our dipeptide alkenes would not fit into the pocket of a wild-type ribosome. The protein is built one amino acid at a time until the ribosome reaches the TAG site, at which point our dipeptide would be incorporated. The transcription-translation process would then continue as normal until the full-length Bora protein is produced. Unfortunately, many attempts to incorporate native Ser–Pro into Bora protein by IVTT have failed to produce any protein. This is likely due to a poor fit of $tRNA^{CUA}$–Ser–Pro into the active site of the modified ribosome. The next step is to find a collaborator who can evolve a ribosome specifically for the Ser–Pro dipeptide.

4.2 Results and Discussion

The IVTT process requires that a $tRNA^{CUA}$ that has not been loaded with an amino acid be aminoacylated with our dipeptide isostere. To do this, we pursued two aminoacylation methods: 1) chemical aminoacylation, which requires the coupling of our dipeptide to a phospho-deoxycytosine-phospho-adenosine (pdCpA) dinucleotide prior to tRNA coupling (Figure 4.5A),[37] and 2) enzymatic aminoacylation, which uses the 3,5-dinitro dFx Flexizyme (dFx) to attach the dipeptide directly to the tRNA (Figure 4.5B).[38, 39]

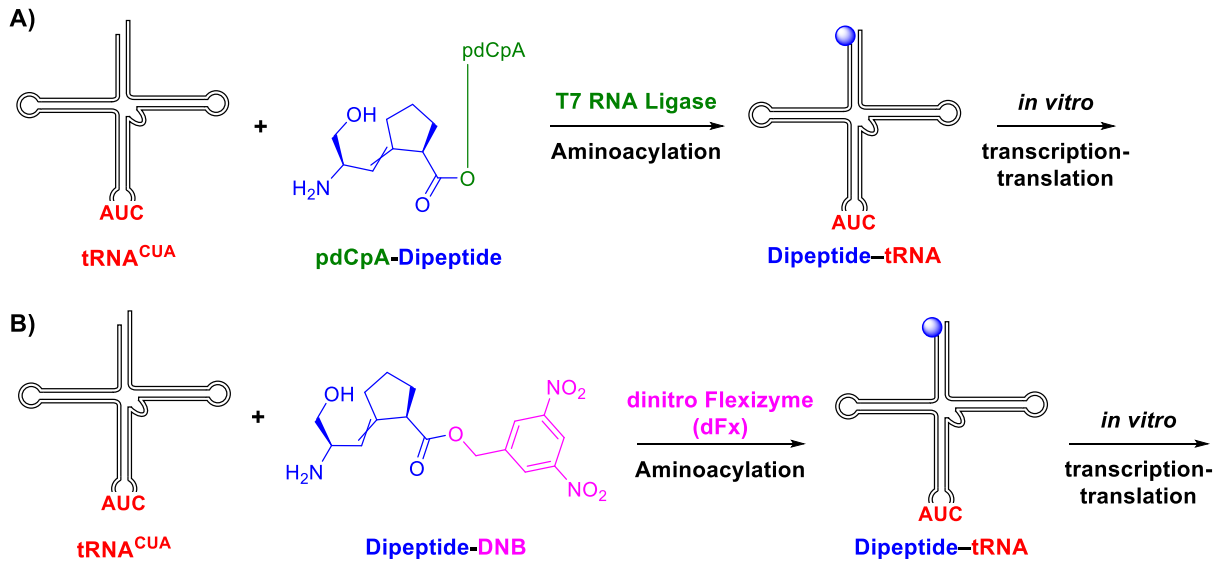


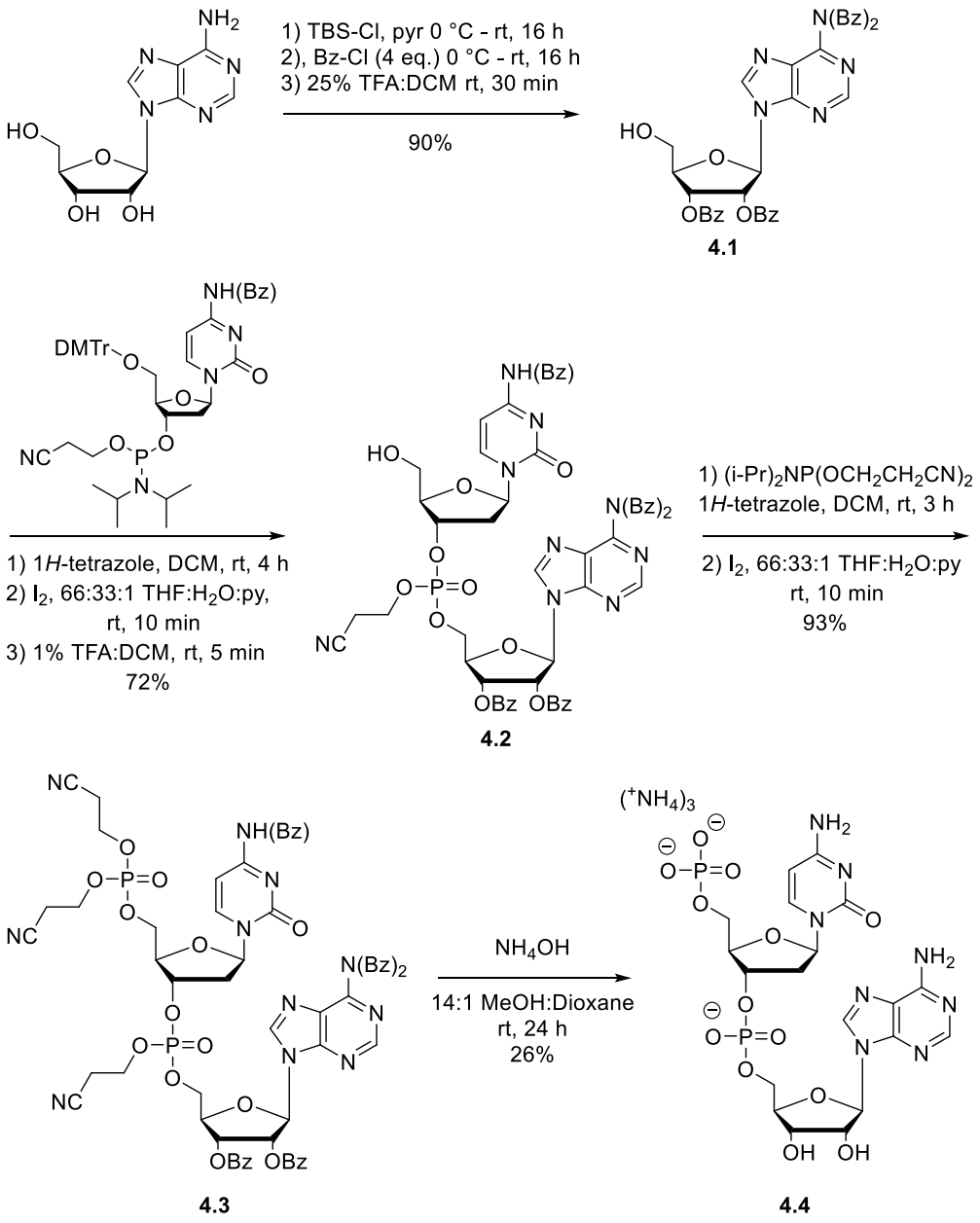
Figure 4.5. Overview of the two aminoacylation methods. Preparation for IVTT by coupling the dipeptide to $tRNA^{CUA}$ through **A)** chemical aminoacylation or **B)** enzymatic aminoacylation.

4.2.1 Chemical Aminoacylation

Nucleotide coupling has been standard procedure since the 1970s.[40] It has evolved since then by improving on the protecting groups, such as the phosphoramidite and cyanoethyl phosphoether. The inclusion of these protecting groups has streamlined the synthesis of pdCpA, a common dinucleotide used for tRNA aminoacylation.[41] We synthesized pdCpA by the method of Zhu et al. with minor modifications.[41]

Adenosine was dissolved in pyridine and the 5' hydroxyl was protected with TBS–Cl (Scheme 4.1). For this reaction to work, it is critically important for the pyridine to be dry. Best results were obtained by freshly distilling it from CaH_2 prior to use. After stirring overnight, four equivalents of benzoyl chloride were added to protect the 2' hydroxyl, 3' hydroxyl, and the adenine amine. Following another 16 h of stirring, the 5' hydroxyl was deprotected using 25% TFA in DCM to give **4.1**, which was isolated on silica gel in 90% overall yield.

Scheme 4.1. Synthesis of pdCpA **4.4**.



The deoxycytosine nucleotide was commercially purchased as *N*4-benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxycytidine-3′-*O*-[*O*-(2-cyanoethyl)-*N*,*N*′-diisopropylphosphoramidite. It was coupled to the adenosine 5′ hydroxyl using 1*H*-tetrazole as a nucleophilic catalyst. The resulting phosphotriester was then quickly oxidized with an aqueous solution of I_2 and pyridine in THF, and the dimethoxytrityl protecting group was removed in 5 min with 1% TFA in DCM.

Compound **4.2** was isolated by silica gel chromatography in 72% overall yield. Intermediate **4.2** was then phosphorylated on the 5′ hydroxyl in a similar manner using bis(2-cyanoethyl)-*N*,*N*-diisopropyl phosphoramidite in the presence of the nucleophilic catalyst 1*H*-tetrazole, followed by oxidation with an aqueous I_2 and pyridine. Purification by silica gel chromatography gave the fully protected dinucleotide **4.3** in 93% yield.

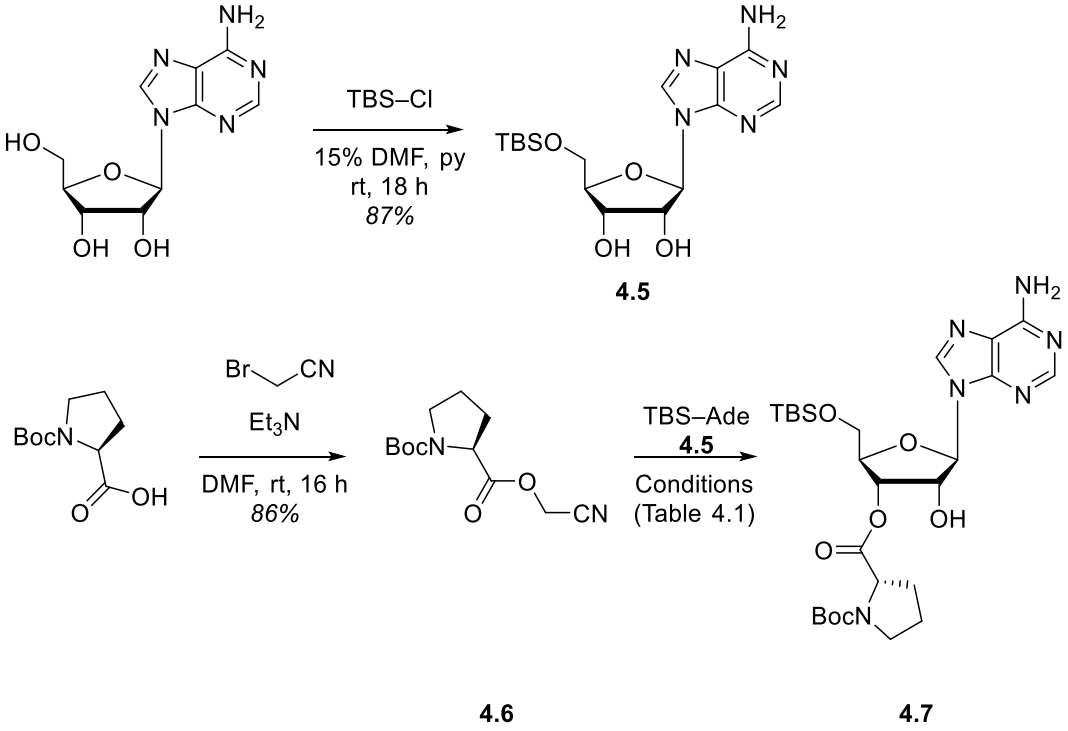
The final step in the synthesis was a 24 h deprotection using conc. NH_4OH (Scheme 4.1). This reaction cleanly removes all three cyanoethyl protecting groups through β-elimination, the byproduct of which is volatile and hazardous. At the same time, five equivalents of benzamide were formed through amidation of the benzoyl esters. Literature procedure calls for HPLC purification of the pdCpA **4.4** salt.[41] To bypass this technical issue, I instead devised a method for purification by silica gel. By implementing Hydrophilic Interaction Liquid Chromatography (HILIC)[42] on silica gel that was deactivated with conc. NH_4OH, the salt was isolated using 5% H_2O:MeCN. Unsurprisingly, **4.4** is much more polar than benzamide, so removing the impurity was simple. The low isolated yield (26%) stems from **4.4** streaking from the baseline. Theoretically, yield may have increased with a higher concentration of polar solvent, but I was hesitant to use more than 5% H_2O to prevent dissolving the silica gel. Regardless, ca. 150 mg of pdCpA was isolated and ready for aminoacylation.

4.2.1A Dipeptide–Dinucleotide Model Coupling

While synthesizing target pdCpA **4.4**, work began on a model coupling reaction (Scheme 4.2). TBS–adenosine was used as a model for dinucleotide **4.4**, which first had to be synthesized. Pyridine was freshly distilled aftert stirring over CaH_2 overnight. Adenosine and TBS–Cl were added to the dry pyridine at 0 °C. Overnight stirring and purification on silica gel gave TBS–Ade **4.5** in 87% yield. Separately, the carboxylic acid Boc–Pro–OH, was activated with

bromoacetonitrile and Et_3N to give the activated ester in 86% yield after liquid-liquid extraction. The cyanomethyl ester is a common moiety to active amino acids for nucleotide coupling due to the small size and high dielectric constant of the cyano group.

Scheme 4.2: Model coupling reaction of Boc–Pro–OCH_2CN and TBS–adenosine.



It quickly became evident that literature protocol[36] did not work with our substrates. This is likely due to steric hinderance of the pyrrolidine ring preventing nucleophilic approach. Additionally, there was no literature precedent for coupling a prolyl carboxylic acid to a nucleotide, so an effective method had to be found (Table 4.1).

Table 4.1 Dipeptide–dinucleotide coupling conditions.

Entry	Solvent[a] (9:1, *v*:*v*)	Temp (°C)	Time (h)					Results
1	DMF:Et_3N	25 Sonicated[36]	2	4	6	24	—	NP[b]
2	DMF:Et_3N	40	2	4	6	24	—	NP[b]
3	DMF:Et_3N	70	2	5	6	24	48	9%
4	DMF:Et_3N	Reflux (90)	2	4	6	24	48	9%
5	DMF:Bu_3N	100	2	4	6	24	48	NP[b]
6	DMF:Bu_3N	Reflux (160)	2	4	6	24	48	NP[b] Decomposition (48 h)
7	9:1 molar DMF:im	100	2	4	6	24	36	12% (4 h), 19% (6 h), 48% (24 h), Decomposition (36 h)
8	9:1 molar DMF:im	160	2	4	6	24	—	11% (4 h) 20% (6 h) Decomposition (24 h)

[a]Imidazole (im) was used at 10% molar equivalent. All other bases were used at 10% volume equivalent. [b]NP = no product detected by TLC.

The product **4.7** was easily visualized on TLC by UV. Literature protocol called for a ratio of 1:7 nucleotide:peptide.[36] In our case, the alkene-locked dipeptides were more precious than the dinucleotide, so this ratio was unfavorable. Fortunately, this is a second order reaction that relies on the concentration of both species. Therefore, we could invert the ratio without consequence.

The coupling did not work with Et_3N as the base while sonicating at room temperature or stirring at 40 °C (Table 4.1, Entries 1 and 2).[37] When heated to 70 °C, **4.7** was observed by TLC. After stirring for 48 h, product was isolated in 9% yield (Table 4.1, Entry 3). Again, **4.7** was obtained in 9% yield after refluxing (90 °C) for 48 h (Table 4.1, Entry 4). Overcoming the steric hindrance of the prolyl activated ester seemed to require additional energy. To this end, the base Bu_3N was used for its higher boiling point (Table 4.1, Entries 5-6). Unfortunately, product was not

formed at any point, and reactants decomposed after 48 h at reflux (160 °C). It was assumed that the size of Bu_3N was preventing product formation, so the much smaller imidazole was used. Product formation was observed after 2 h at 100 °C. The reaction was allowed to continue, and stirring for 4 h at 100 °C gave **4.7** in 12% yield (Table 4.1, Entry 7). Optimization of this reaction gave 19% and 48% yield after 6 h and 24 h, respectively (Table 4.1, Entry 7). The reaction decomposed after 36 h at 100 °C and no product could be isolated (Table 4.1, Entry 7). The temperature dependence of this reaction was tested by increasing the heat to 160 °C again using 9:1 molar equivalents of DMF:im (Table 4.1, Entry 8). Only 11% of **4.7** could be isolated after 4 h, and 20% after 6 h (Table 4.1, Entry 8). The reaction decomposed after 24 h (Table 4.1, Entry 8). Optimal coupling conditions were 9:1 (*m*:*m*) DMF:im at 100 °C for 24 h (Table 4.1, Entry 7).

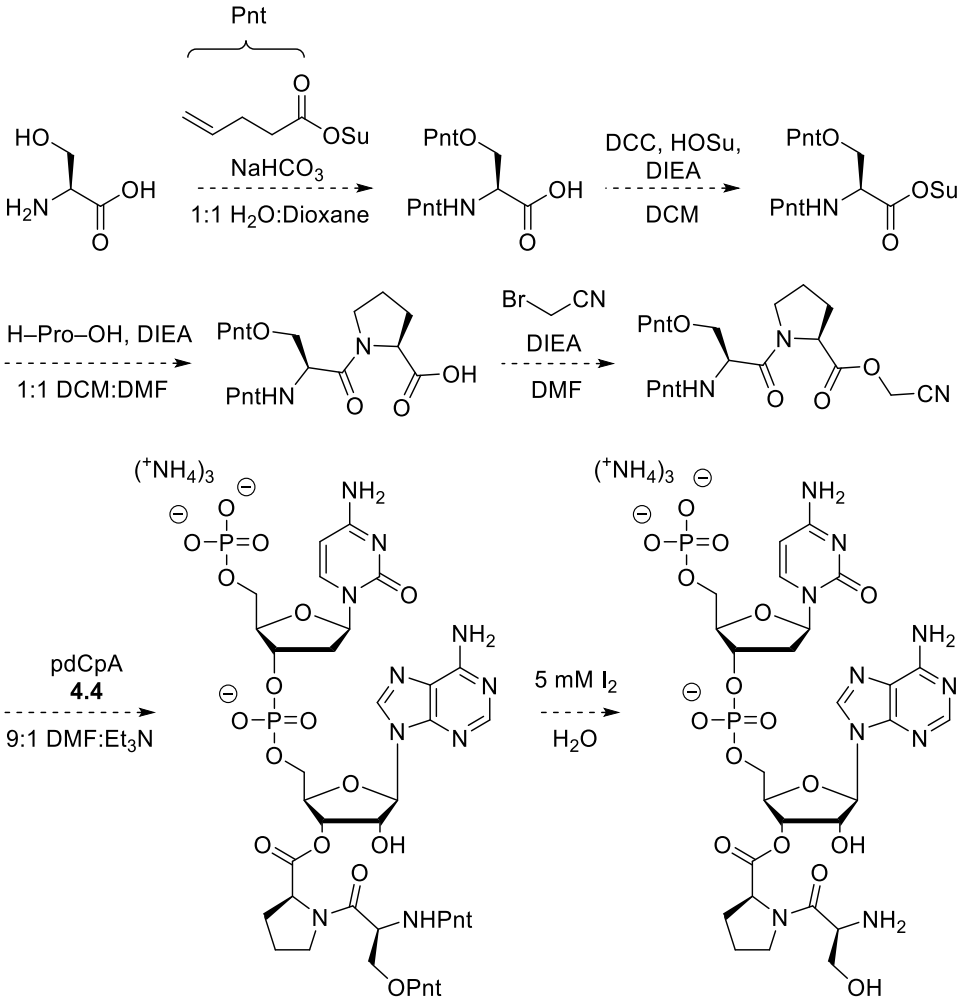
4.2.1B Native Dipeptide for Chemical Misacylation

As a model for coupling and protein production optimization, and as a control for the bioassays, a Ser–Pro dipeptide was needed. The carboxylic acid of an *N*-protected amino acid can be activated with a number of different carbodiimides (DCC, DIC, etc.) or aminium/uronium and phosphate salts (HATU, HBTU, HCTU, etc.) allowing for nucleophilic addition of a second amino acid.[43] Preparation of a Ser–Pro dipeptide was unexpectedly difficult because it necessitated the protection of the Ser side chain hydroxyl. Without side chain protection, any attempt to link the dipeptide to pdCpA **4.4** would have resulted in polymerization.

Our initial plan towards the dinucleotide-dipeptide complex involved activating the 4-pentenoyl (Pnt) carboxylate with *N*-hydroxy succinimide and using it to protect the *N*-terminus and side chain of Ser simultaneously (Scheme 4.3). The Pnt moiety is an established protecting group for amino acid coupling to pdCpA due to the facile conditions for its removal which do not compromise phosphate linkages.[37, 44, 45] While this easily protected the Ser amine, the side chain

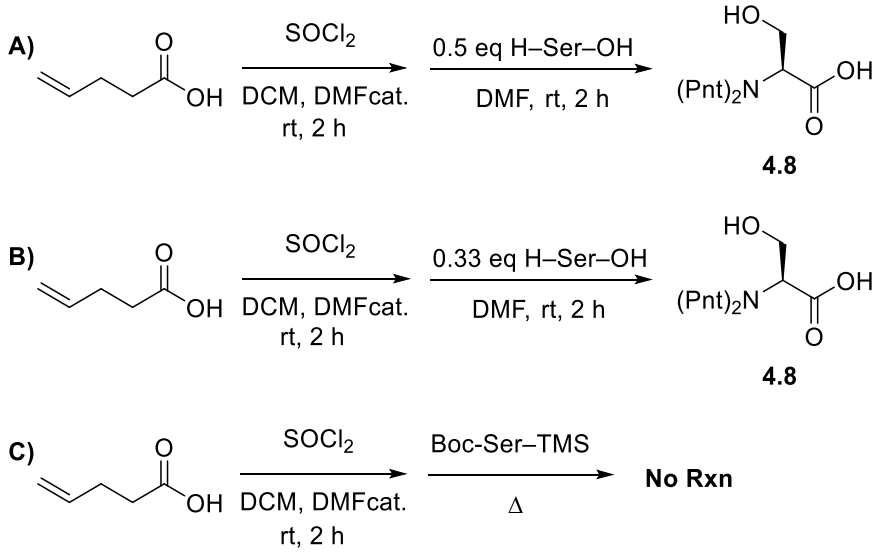
hydroxyl remained unreacted regardless of time (4 h, 8 h, 12 h, 24 h, 28 h), temperature (rt, 45 °C, 60 °C, 80 °C), equivalents (2 eq., 4 eq., 6 eq.), or activating group (HOSu, HOBt, 6-Cl-HOBt, HATU).

Scheme 4.3. Proposed synthesis of the dinucleotide-dipeptide complex pdCpA–Pro–Ser–H.



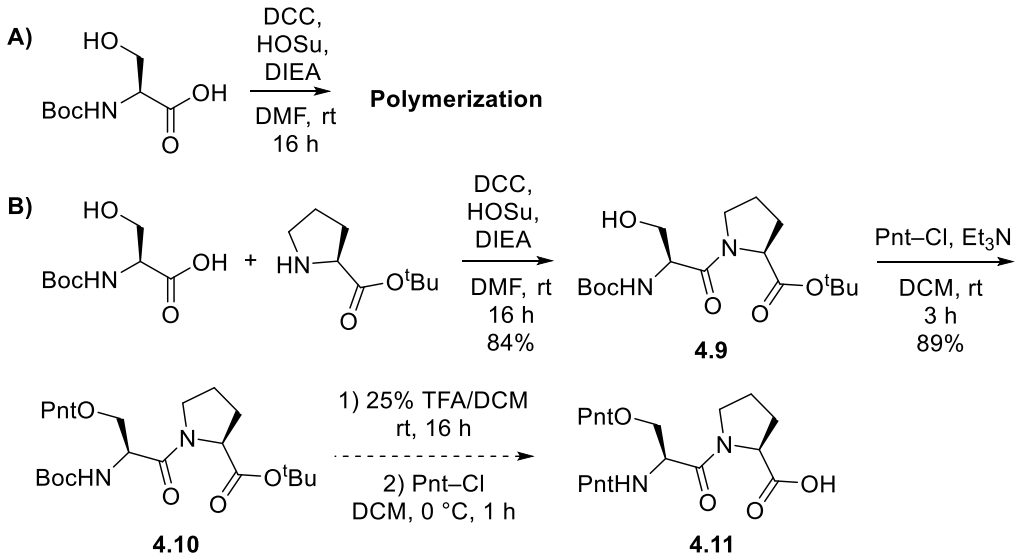
Pnt–Cl, a more potent electrophile, was made *in situ* by stirring Pnt–OH with freshly-distilled thionyl chloride (Scheme 4.4). The resulting acyl chloride could not be isolated, so L-serine was added after two hours of acyl chloride formation. Using two equivalents of Pnt–Cl, bis(Pnt)–Ser–OH **4.8** was formed (Scheme 4.4A) instead of the desired Pnt–Ser(Pnt)–OH.

Scheme 4.4. Pnt–Cl to protect the Ser hydroxyl of Boc–Ser–OH.



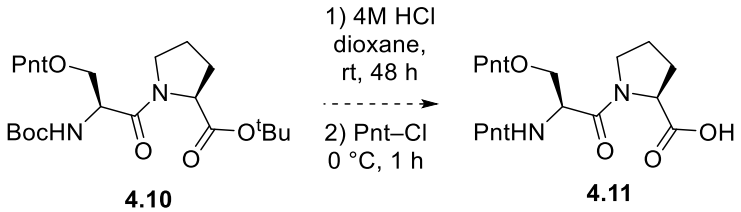
Adding a third equivalent (Scheme 4.4B) saw no additional Pnt addition, presumably due to the formation and subsequent hydrolysis of the bis(Pnt)–Ser-4-pentenoic anhydride. To prevent the formation of the anhydride, TMS–Cl was used to protect the carboxylic acid of Boc–Ser–OH. Once isolated, this waxy compound was not soluble in DCM and did not react with Pnt–Cl even with overnight stirring and gentle heating (Scheme 4.4C).

Scheme 4.5. Pnt-Cl to protect the Ser hydroxyl of Boc–Ser–Pro–O^{t}Bu.



To isolate the Ser hydroxyl as the only available nucleophile, the Boc-Ser-Pro-O^{t}Bu dipeptide **4.9** (Scheme 4.5) was synthesized. However, activation of the Boc–Ser–OH carboxyl with HOSu caused polymerization to occur with the side chain hydroxyl (Scheme 4.5A). To resolve this issue, Boc–Ser–OH and H–Pro–O^{t}Bu were coupled with HOSu and DIEA. By keeping the solution dilute, the nucleophilicity of the secondary amine outperformed the alcohol and the dipeptide was isolated in 84% yield (Scheme 4.5B). From there, the hydroxyl was protected with commercial Pnt–Cl for 3 h at rt to give the dipeptide Pnt ester **4.10** in 89% yield (Scheme 4.5B). Unexpectedly, removal of the *t*-butyl protecting group required overnight stirring with 25% TFA/DCM. The resulting crude product was neutralized with $NaHCO_3$ and subsequent addition of 1.0 equivalent of Pnt–Cl resulted in three Pnt signals by ^{1}H NMR after product isolation. A strong signal from ^{19}F NMR was observed that was distinct from TFA ($CDCl_3$ δ –77.0 ppm). The fluorinated byproduct could not be separated by column chromatography, nor identified, due to an R_f value that is nearly identical to **4.11**.

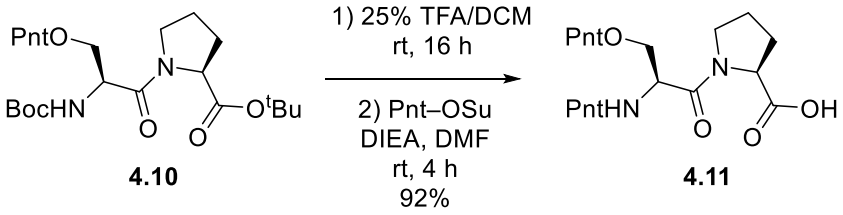
Scheme 4.6. HCl deprotection of Boc and *t*-butyl ester with Pnt–Cl protection.



4.10 **4.11**

To bypass the potential of a trifluoromethyl 4-pentenoic anhydride, or trifluoroacetamide, the Boc and *t*-butyl groups were removed with 4 M HCl in dioxane (Scheme 4.6). While the Boc group was removed in approximately 1 h (as determined by TLC), the *t*-butyl ester took 48 h for 99% removal (as determined by ^{1}H NMR of the reaction mixture). The resultant HCl salt was dissolved in DMF and stirred with Pnt–Cl at 0 °C for 1 h. The reaction was purified, but the isolated spot compound found by TOF mass spectroscopy not to be the desired **4.11**.

Scheme 4.7. TFA deprotection of Boc and *t*-butyl ester with Pnt–OSu protection.



The evidence has shown that Pnt–Cl is too strong of an electrophile to use for *N*-terminal protection, therefore preactivation of Pnt–OH to form Pnt–OSu became the more viable option. After overnight deprotection of **4.10** in 25% TFA/DCM, the crude product was neutralized with $NaHCO_3$, dissolved in DMF, and successfully protected to give **4.11** in 92% yield (Scheme 4.7). Protecting the dipeptide in this manner, instead of attaching both Pnt groups to L-serine as initially intended, serves as a much better model protection of the Ser–Pro *cis*- and *trans*-locked alkene dipeptides, as well. It was at this time that we realized the Pnt protecting group was unsuitable for our compounds, and for the same reason that they are commonly used: iodolactonization (Figure 4.6).

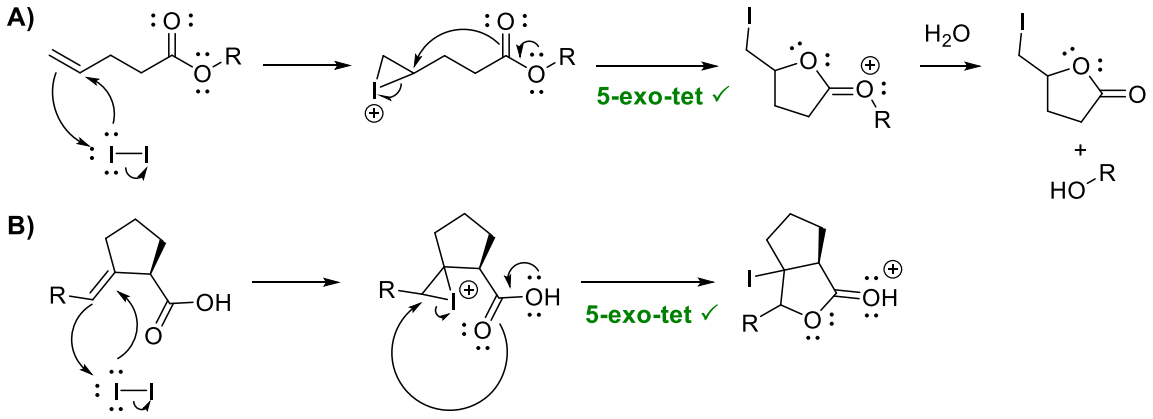
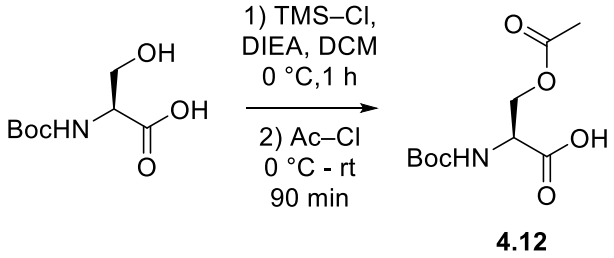


Figure 4.6. A) Iodolactonization of the Pnt group of a generic ester and **B)** potential lactonization of our alkene dipeptides.

The Pnt protecting group is frequently used for tRNA aminoacylation because its removal through iodolactonization with aqueous I_2 (Figure 4.6A) does not interfere with the pdCpA

phosphate backbone. In our case, the *cis*- and *trans*- alkene mimics can undergo the same favorable 5-exo-tet intramolecular cyclization (Figure 4.6B), so they will likely decompose in the presence of I_2. Fortunately, Fahmi et al. previously utilized the acetyl group for hydroxyl protection in the glycosylation of pdCpA, also for IVTT.[44] The acetyl ester was removed from a pdCpA-glycoside using 50% $HBF_4 \cdot Et_2O$.[44] Since the Boc group is also acid-labile, we decided that Boc was a suitable *N*-terminal protecting group for the tRNA coupling.

Scheme 4.8. Acetylation of Boc–Ser–OH.



4.12

To acetylate the Ser side chain, the carboxyl of Boc–Ser–OH was first protected with TMS–Cl in DCM (Scheme 4.8). After 1 h of stirring at 0 °C, acetyl chloride was added, and the reaction was stirred at rt for 90 min. By TLC, the product seemed to form in quantitative yield; however, it decomposed on silica gel. Intermediate **4.12** was then purchased commercially as a viscous oil. The oil was dissolved in EtOAc, portioned into separate 6-dram vials, and dried *in vacuo*. After drying on high vacuum overnight, ^{1}H NMR showed ca. 30% of the acetyl group had hydrolyzed. Hydrolysis of the Boc–Ser(Ac)–OH ester **4.12** may be facilitated by an intramolecular hydrogen bond (Figure 4.7). The carboxyl may act as a hydrogen-bond donor and the ester oxygen as the hydrogen-bond acceptor to decrease electron density on the carboncyl, which would facilitate hydrolysis.

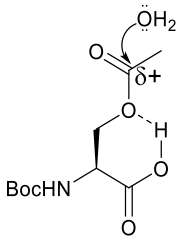
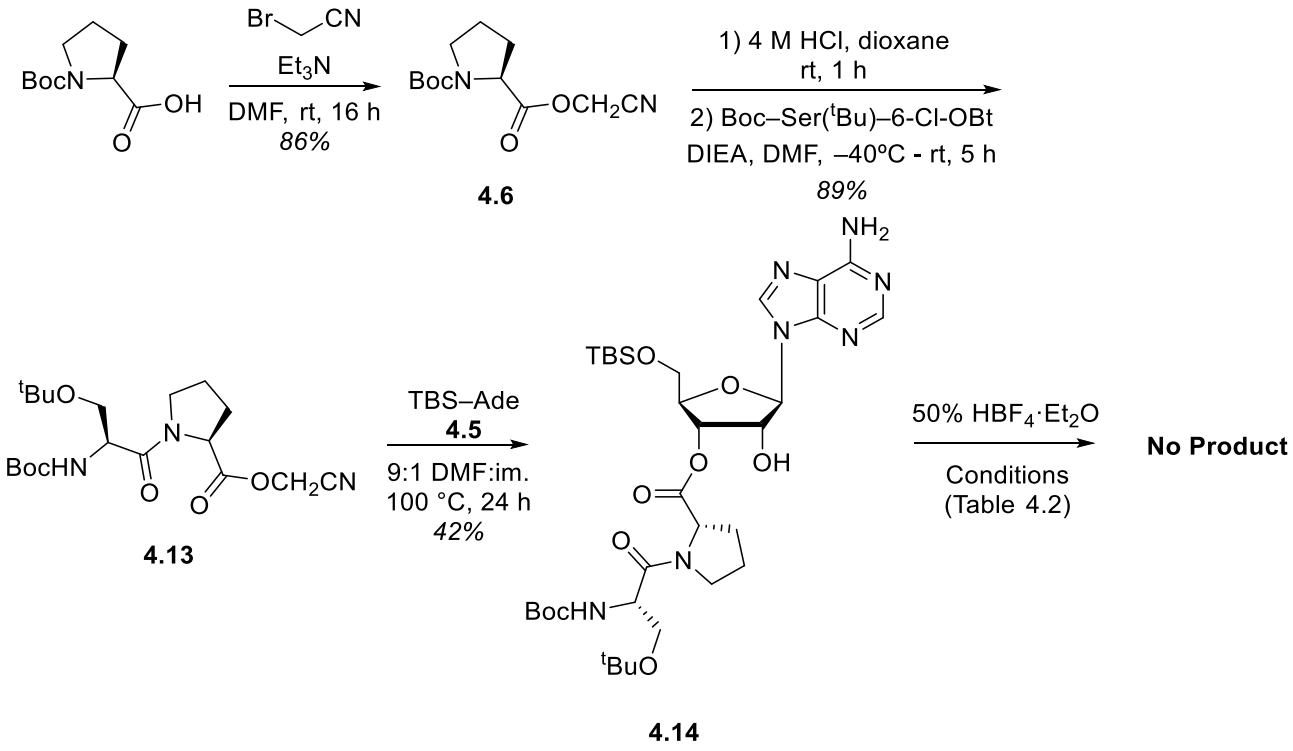


Figure 4.7. Proposed facile hydrolysis of the Boc–Ser(Ac)–OH ester **4.12**.

If intramolecular hydrogen bonding was indeed enabling ester hydrolysis, then the acetyl would be far more stable on the Ser–Pro dipeptide due to the conformationally-locked Φ dihedral angle of Pro. Unfortunately, intermediate **4.9** (Scheme 4.5) could not be used because removal of the *t*-butyl ester (for subsequent activation as the cyanomethyl ester) would also remove our *N*-terminal Boc group. Thus, a new method was devised.

This unusual synthesis relies on "protecting" the prolyl carboxyl of Boc–Pro–OH as the activated cyanomethyl ester, which was easily obtained in 86% yield after a series of liquid-liquid extractions (Scheme 4.9). The Boc group was then removed with 4 M HCl. After drying overnight, was coupled with the activated 6-Cl-HOBt ester of Boc–Ser–OH. Literature precedent for reacting two activated esters could not be found, but polymerization of the H–Pro–OCH_2CN intermediate was minimized by keeping the solution dilute and cold. Presumably, decreasing temperature encourages the displacement of 6-Cl-HOBt (p*K*a = 4.62)[46] instead of $HOCH_2CN$ (estimated p*K*a = 11.31).

Scheme 4.9. Synthesis of the activated dipeptide Boc–Ser(OAc)–Pro–OCH_2CN **4.13**, coupling to the pdCpA model, TBS-adenosine **4.5** and subsequent deprotection with 50% $HBF_4 \cdot Et_2O$.



The dipeptide **4.13** has only acid-labile protecting groups that were both cleaved with $HBF_4 \cdot Et_2O$ in model reactions. To model the reaction with pdCpA, activated dipeptide **4.13** was successfully coupled to model nucleotide **4.5** in 42% yield using the optimized conditions from Table 4.1.

Table 4.2. Deprotection optimization for **4.14**.

Entry	$HBF_4 \cdot Et_2O$ Eq	Time (h)	Temp (°C)	Results
1[45]	377	3	rt	Decomposition
2	5	3, 6, 12, 24	rt	NP[a]
3	5	3, 6, 12, 24	35	NP[a]
4	5	3, 6	50	Decomposition

[a]NP = no product detected by TLC

Attempted deprotection of acetyl and Boc protecting groups with 50% $HBF_4 \cdot Et_2O$ was unsuccessful under several conditions (Table 4.2). All reactions were performed in MeOH and monitored by TLC using UV light. Following the Fahmi protocol,[45] the first trial used 377 eq of $HBF_4 \cdot Et_2O$ and stirred at rt for 3 h (Table 4.2, Entry 1). The reaction mixture quickly turned black and ^1H NMR showed multiple sets of adenine peaks, suggesting the purine had disconnected from the sugar ring to some extent. The number of $HBF_4 \cdot Et_2O$ equivalents was then scaled back to 5, and the reaction was repeated (Table 4.2, Entry 2). Monitoring this reaction for 24 revealed no product formation. It was then heated to 35 ºC for an additional 24 h, with no product formation at any time (Table 4.2, Entry 3). Heating to 50 °C saw a new, faint, UV active spot on TLC after 3 h; however, after 6 h the solution had turned black and ^1H NMR again showed dissociation of the purine.

Around this time, we turned to an alternative aminoacylation technique: The dFx flexizyme.

4.2.2 Enzymatic Aminoacylation

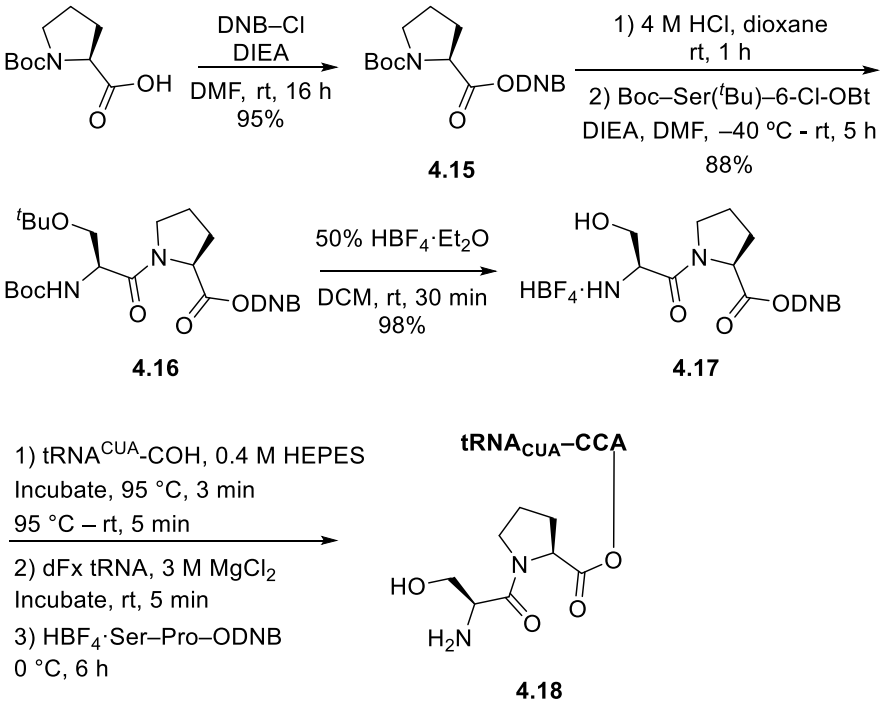
Ribozymes are catalytic pieces of RNA that have previously been used for aminoacylation.[47] The issue with ribozymes, however, are their specificity for the amino acid and the tRNA. A more functional ribozyme (r24) was discovered that preferentially links aromatic amino acids, activated with the cyanomethyl ester group, to specific kinds of tRNA.[39] This ribozyme was then mutated to decrease its tRNA specificity.[39] This new enzyme was called a flexizyme due to its flexibility of aminoacylation of different types of tRNA. This flexizyme was modified further to eliminate amino acid specificity, as well. The newly formed dinitro flexizyme (dFx) will aminoacylate nearly any tRNA with nearly any amino acid that is activated on the *C*-terminal with the 3,5-dinitrobenzyl ester moiety.[48] Enzymatic aminoacylation is beautiful in its

simplicity because it requires no protecting groups, only the 3,5-dinitrobenzyl (DNB) ester leaving group.

4.2.2A Native HBF_4·H–Ser–Pro–ODNB for Enzymatic Aminoacylation

Using our chemical misacylation work, Scheme 4.9 was modified to include the DNB ester (Scheme 4.10). For the dipeptide, coupling with Boc–Pro–ODNB was more efficient than first coupling Boc–Ser(*t*Bu)–OH to H–Pro–OBn, followed by benzyl deprotection and activation with DNB.

Scheme 4.10. Enzymatic misacylation of HBF_4·H–Ser–Pro–ODNB **4.17**.



The prolyl carboxyl was again "protected" as the activated DNB ester prior to coupling with Boc–Ser(*t*Bu)–OH. Fortunately, the p*K*a of DNB–OH is estimated to be close to $HOCH_2CN$ so the same coupling conditions were applied to attaching the Boc–Ser(*t*Bu) moiety. Both TFA and HCl deprotections of the Boc group resulted in hydrolysis of the DNB ester. It was found that one equivalent of a 50% HBF_4·Et_2O solution will cleanly remove both the Boc and *t*-butyl ether

groups in 30 min in excellent yield without any hydrolysis of the ester. The HBF_4 salt was then given to our molecular biologist, Karla Piedl, who performed the tRNA aminoacylation (Figure 4.8).

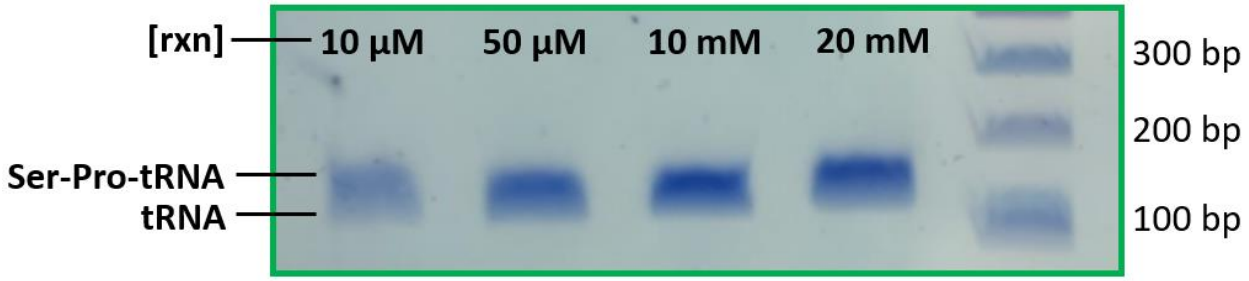


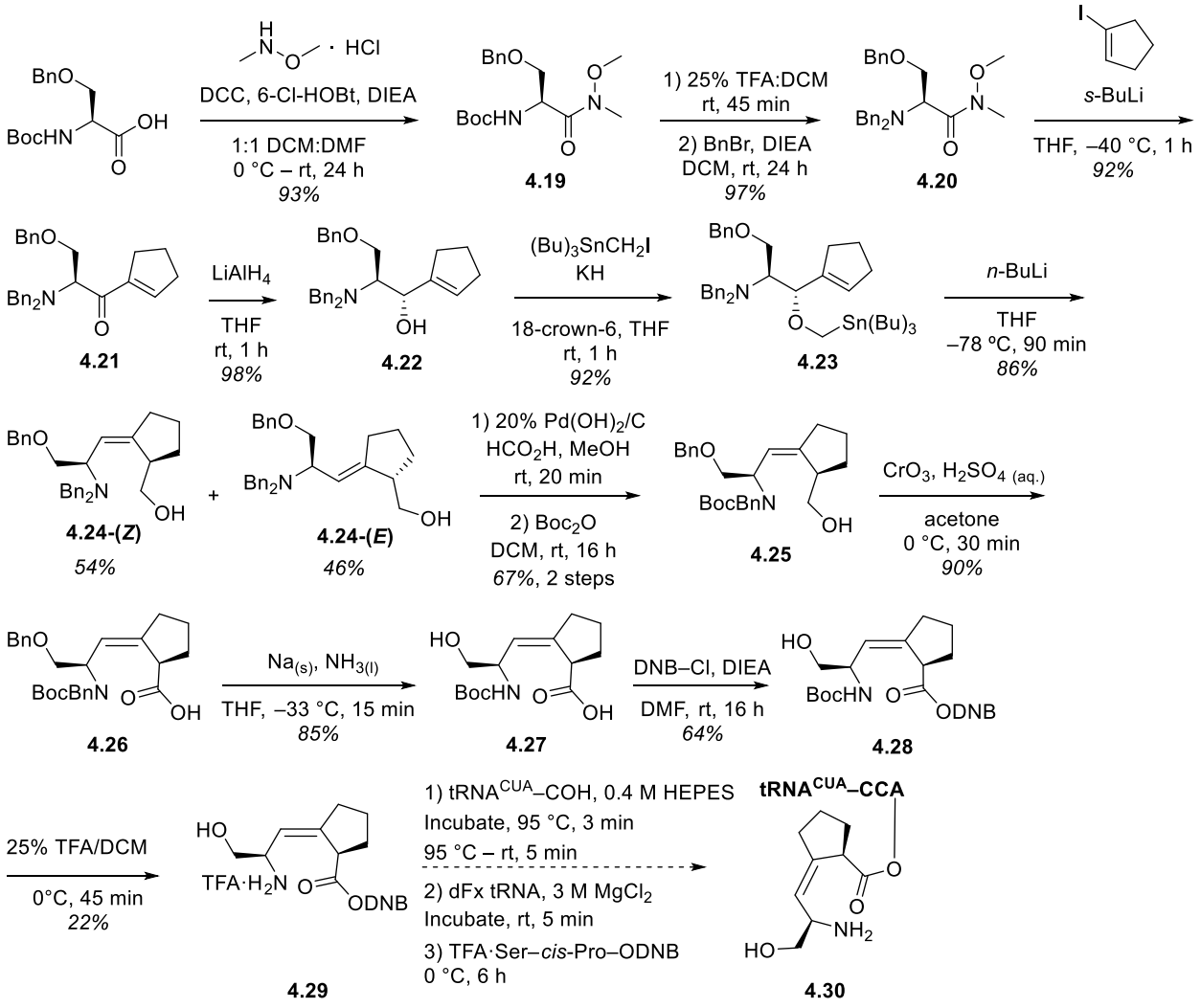
Figure 4.8. SDS-Page Gel of the aminoacylation reaction. The top band shows the Ser–Pro dipeptide linked to the tRNA. The bottom band shows the empty tRNA. In each lane, the concentration of the reaction mixture varied. The concentration of empty tRNA was kept constant. (Credit: Karla N. Piedl, unpublished).

The gel above shows that the concentration of the Ser–Pro-loaded tRNA band increases as the concentration of the loading reaction mixture increases. At the same time, as the concentration of the loading reaction mixture increases, the concentration of the empty tRNA band remains unchanged. This shows that there is very little unloaded tRNA left in the reaction mixture and intermediate **4.18** was ready for IVTT. The native dipeptide was made in 82% yield over 3 steps and was shown capable of aminoacylating tRNA by dFx.

4.2.2B TFA·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB for Enzymatic Aminoacylation

The target cis-locked alkene, TFA·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB **4.29** was synthesized from Boc–Ser–Ψ[(*Z*)CH=C)]–Pro–OH that was resynthesized by the published method (Scheme 4.11).[34] The acid was then converted into the DNB ester conjugate for tRNA aminoacylation[38] (Scheme 4.11).

Scheme 4.11. Synthesis of the cis-locked alkene TFA·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB **4.29** and enzymatic aminoacylation.[34]



The published synthesis of the cis-locked alkene isostere **4.27** was followed closely with few deviation.[34] The process for the iodocyclopentene intermediate was adjusted slightly. It is made in three consecutive steps from cyclopentanone hydrazone, the last of which is refluxing in Et_2O. It was instead stirred in Et_2O at rt for the same amount of time. This modification did not affect yield or purity. The cyclopentanone hydrazone intermediate procedure was modified as well. During liquid-liquid extraction, $CHCl_3$ was used instead of CH_2Cl_2. This serendipitously increased yield from 80%[34] to 96%. It was also discovered that if the cyclopentanone hydrazone is not freshly

prepared, it has a propensity to irreversibly dimerize. Dimerization also occurs by drying on high vacuum overnight. All other steps were performed as reported.[34]

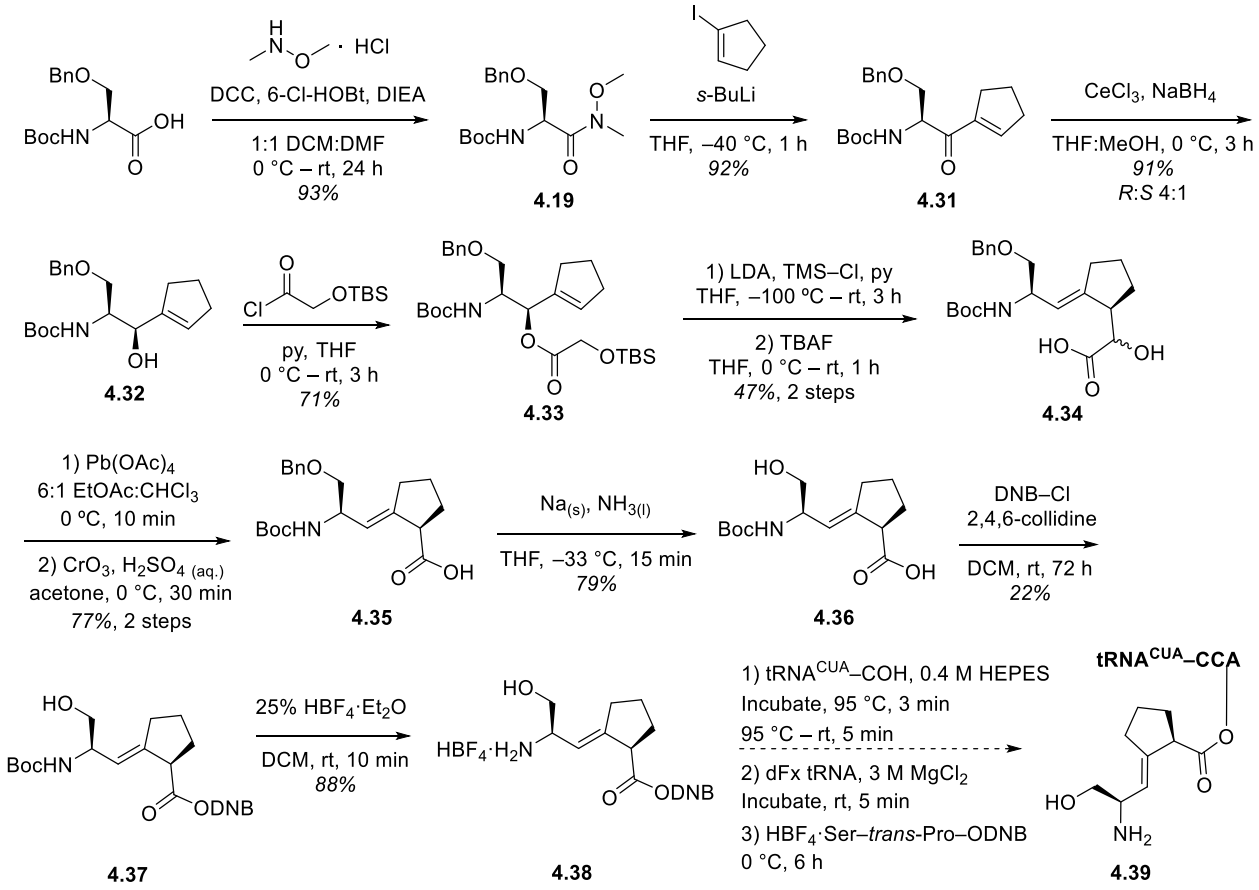
Boc–Ser–Ψ[(*Z*)CH=C)]–Pro–OH **4.27** was converted to the DNB ester **4.28** by using one equivalent of DNB–Cl and excess DIEA. Only one equivalent of DNB–Cl was used to prevent benzylation of the Ser hydroxyl. It was later discovered that hydrolysis of the alkyl chloride is unexpectedly difficult, taking up to three weeks at reflux.[49] The yield for this esterification reaction can likely be improved by increasing the equivalents of DNB–Cl.

Deprotection of the Boc group was accomplished using a solution 25% TFA/DCM at 0 °C for 45 min to give the target compound TFA·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB **4.29** in 3% overall yield over 11 steps.

4.2.2C HBF_4·H–Ser–Ψ[(*E*)CH=C)]–Pro–ODNB for Enzymatic Aminoacylation

The target trans-locked alkene, HBF_4·H–Ser–Ψ[(*E*)CH=C)]–Pro–ODNB **4.38** was synthesized from Boc–Ser–Ψ[(*E*)CH=C)]–Pro–OH **4.36** that was resynthesized by the published method with no deviation (Scheme 4.12).[34] The acid was then carefully converted to the DNB ester conjugate (Scheme 4.12).

Scheme 4.12. Synthesis of the trans-locked alkene HBF_4·H–Ser–Ψ[(*E*)CH=C)]–Pro–ODNB **4.38** and enzymatic aminoacylation.[34]



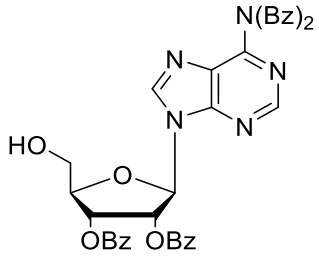
Esterification of the Boc–Ser–Ψ[(*E*)CH=C)]–Pro–OH **4.36** acid is a delicate process. The alkene is known to isomerize to the α,β-unsaturated carbonyl in the presence of an activated ester and a base.[11] Indeed, even one equivalent of DNB–Cl and one equivalent of DIEA isomerized the (*E*)-alkene. It has been shown that 2,4,6-collidine limits isomerization of the (*E*)-alkene.[11] However, isomerization still occurred while using one equivalent of DNB–Cl and two equivalents of collidine. An umpolung approach was next envisioned with DNB–OH as the Nu: and activated Pro as E+. Hydrolysis of the DNB–Cl was attempted to make DNB–OH for standard carboxylic acid coupling with HOAt/HATU. The difficulty of DNB–Cl hydrolysis was then realized.[49] A test reaction confirmed that DNB–Cl will not hydrolyze in 1:1 acetone:H_2O at reflux over 24 h. This

means that an excess of DNB–Cl could likely be used without benzylation of the Ser hydroxyl. Therefore, five equivalents of DNB–Cl and one equivalent of 2,4,6-collidine were used and stirred for 72 h to give the activated DNB ester **4.37** in 22% yield with no isomerization or hydroxyl benzylation. The Boc group was then deprotected by stirring for 15 min with $HBF_4 \cdot Et_2O$ in DCM to give target $HBF_4 \cdot$H–Ser–Ψ[(*E*)CH=C)]–Pro–ODNB **4.38** in 2% yield over nine steps. This gave a sufficient quantity for chemical misacylation and several IVTT attempts.

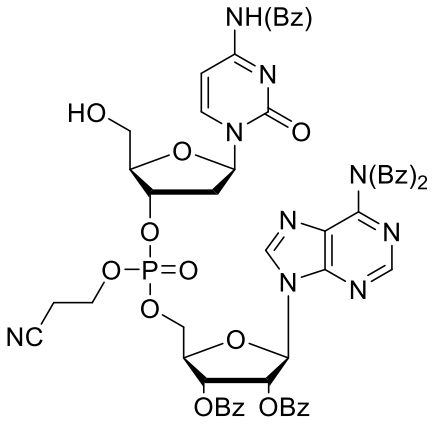
4.3 Experimental

General Methods. All reagents were obtained from commercial suppliers and used as received, unless otherwise stated. Solvents were dried and deoxygenated under Ar by filtration through a column of molecular sieves and a column of activated alumina using Innovative Technology's PureSolv solvent purification system. Amine bases were stirred over CaH_2 overnight and distilled prior to use. Disappearance of starting material was monitored by analytical thin layer chromatography, performed on aluminum-backed SiliaPlate (200 μm thickness) silica gel plates provided by Silicycle. Visualization was accomplished using UV light or phosphomolybdic acid stain followed by gentle heating, unless otherwise stated. Brine (NaCl) and NH_4Cl refer to saturated aqueous solutions, unless otherwise stated. Organic solutions were concentrated under reduced pressure using a Büchi RE-111 rotary evaporator. Column chromatography was performed with SiliaFlash P60 silica gel (230-400 mesh) provided by Silicycle. Medium pressure chromatography was performed on a Teledyne Isco Rf200 CombiFlash system. ^{1}H NMR and ^{13}C NMR spectra were recorded on an Agilent Varian model U4-DD2 NMR spectrometer (^{1}H NMR at 400 MHz, ^{13}C NMR at 100 MHz). Chemical shifts for protons are reported in ppm with reference to $CHCl_3$ at 7.26 ppm. Chemical shifts for carbons are reported in ppm with reference to $CHCl_3$ at 77.0 ppm and are proton decoupled. NMR data are described as follows: chemical shift,

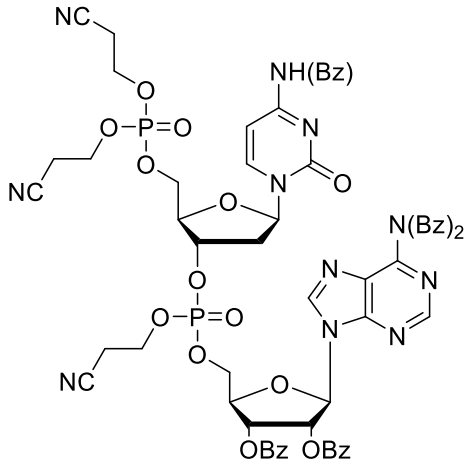
multiplicity (s = singlet, bs = broad singlet, d = doublet, br d = broad doublet, dd = doublet of doublets, t = triplet, dt = doublet of triplets, m = multiplet), coupling constants (Hz), and integration. Rotamer peaks are in parenthesis where they could be determined. High-resolution mass spectra (HRMS) were acquired on an Agilent 6220 using electrospray ionization (ESI^+) and time of flight analyzer (TOF).



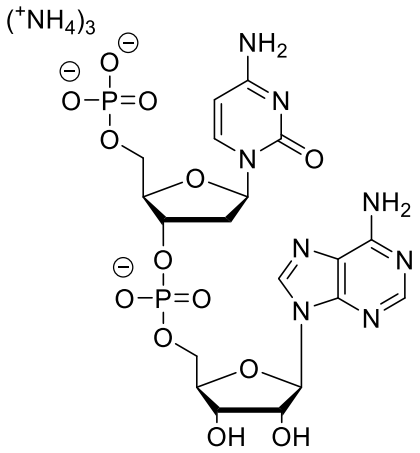
Tetrabenzoyl-adenosine 4.1. By the method of Zhu et al. with minor modifications.[41] TBS-Cl (370 mg, 2.4 mmol) in DMF (1 mL) was added to adenosine (Ade) (580 mg, 2.2 mmol) in pyridine (15 mL) at 0 °C and stirred at rt for 16 h. The TBS-Ade was then cooled to 0 °C and benzoyl chloride (1.2 mL, 10 mmol) was added dropwise. The mixture was allowed to warm to rt and stirred for 16 h, then it was diluted with DCM (100 mL), washed with brine (40 mL), dried over Na_2SO_4 and concentrated. The crude residue was dissolved in DCM (36 mL), TFA (9 mL) was added, and the mixture was stirred for 30 min. The solution was neutralized with solid $NaHCO_3$ at 0 °C until bubble formation ceased, and the product was extracted with DCM (4 × 50 mL). The combined organic layers were dried with Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 5% MeOH/DCM) to give **4.1** as a white solid (1.3 g, 90%). The ^1H NMR matched the literature.[41] ($CDCl_3$, 400 MHz): δ 8.69 (s, 1H), 8.22 (s, 1H), 8.05 – 8.04 (m, 1H), 8.03 – 8.02 (m, 1H), 7.88 – 7.84 (m, 6H), 7.52 – 7.31 (m, 13H), 6.38 – 6.32 (m, 2H), 6.05 (dd, *J* = 5.2, 1.6 Hz, 1H), 5.59 – 5.48 (m, 1H), 4.62 (d, *J* = 1.7 Hz, 1H), 4.11 – 4.55 (m, 1H), 4.04 – 3.96 (m, 1H).



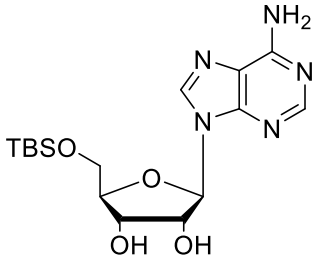
Cyanoethyl (Bz)$_5$dCpA 4.2.[41] Tetrabenzoyl adenosine **4.1** (175 mg, 257 μmol), DMT-dC(Bz) phosphoramidite (176 mg, 211 μmol), and 1*H*-tetrazole (0.45 M in MeCN, 700 μL, 0.32 μmol) were dissolved in dry DCM (750 μL). After stirring for 4 h at rt, iodine (200 mg, 790 μmol) in 66:33:1 THF:H_2O:pyridine (2.1 mL) was added. The resultant solution was stirred at rt for 10 min. Then it was diluted with EtOAc (37 mL), washed with 0.2 M $NaHSO_3$ (2 × 7.5 mL), and brine (7.5 mL), dried over Na_2SO_4, and concentrated. The residue was dissolved in 1% TFA/DCM (18.5 mL) and stirred at rt for 5 min. Then it was quenched with $NaHCO_3$ (3.75 mL). The organic layer was washed with brine (3.75 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by medium-pressure SiO_2 chromatography (3 × 15 cm, 40 g SiO_2, 0–4% MeOH/DCM over 40 column volumes). The product peak, which eluted between 12 and 15 min, gave **4.2** as a white solid (177 mg, 72%, mixture of diastereomers). The ^{1}H NMR matched the literature.[41] ($CDCl_3$, 400 MHz): δ 8.69 (s, 1H), 8.67 (s, 1H), 8.50 (s, 1H), 8.49 (s, 1H), 8.24 (d, *J* = 7.5 Hz, 1H), 8.21 (d, *J* = 7.5 Hz, 1H), 7.93 (d, *J* = 7.4 Hz, 4H), 7.91 – 7.89 (m, 8H), 7.82 (d, *J* = 7.6 Hz, 8H), 7.58 – 7.45 (m, 16H), 7.38 – 7.33 (m, 16H), 6.57 (d, *J* = 4.9 Hz, 1H), 6.55 (d, *J* = 4.9 Hz, 1H), 6.31 – 6.24 (m, 2H), 6.24 – 6.11 (m, 4H), 5.21 – 5.15 (m, 1H), 5.15 – 5.11 (m, 1H), 4.72 – 4.68 (m, 2H), 4.60 – 4.51 (m, 4H), 4.31 – 4.17 (m, 6H), 3.83 – 3.72 (m, 4H), 3.15 – 3.07 (m, 2H), 2.80 – 2.71 (m, 6H), 2.43 (m, 1H), 2.34 (m, 2H).



Tricyanoethyl (Bz)$_5$pdCpA 4.3.[41] To a solution of cyanoethyl (Bz)$_5$dCpA **4.2** (62 mg, 55 μmol) in DCM (375 μL) was added 1*H*-tetrazole in 0.45 M in MeCN (9.6 μL, 110 μmol) followed by bis(2-cyanoethyl)-*N*,*N*-diisopropyl phosphoramidite (29 μL, 110 μmol) and the solution was stirred at rt for 4 h. Iodine (56 mg, 220 μmol) in 66:33:1 THF:H_2O:pyr (560 μL) was added and the resulting solution was stirred for 10 min. The reaction was then was diluted with EtOAc (12 mL), washed with 0.2 M $NaHSO_3$ (2 × 2 mL) and brine (2 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by medium-pressure SiO_2 chromatography (2 × 15 cm, 10 g SiO_2, 0-2% MeOH/DCM for 1 min, 2-4% MeOH/DCM for 10 min, 4-8% MeOH/DCM for 10 min). The product peak, which eluted between 6 and 11 min, gave **4.3** as a white solid (47 mg, 93%, mixture of diastereomers). The ^{1}H NMR matched the literature.[41] ($CDCl_3$, 400 MHz): δ 8.68 (s, 2H), 8.52 (s, 1H), 8.50 (s, 1H), 8.03 – 7.86 (m, 14H), 7.85 (s, 4H), 7.83 (s, 4H), 7.61 – 7.53 (m, 7H), 7.51 – 7.46 (m, 9H), 7.41 – 7.34 (m, 17H), 6.58 (d, *J* = 4.9 Hz, 1H), 6.56 (d, *J* = 4.8 Hz, 1H), 6.29 (ddd, *J* = 13.7, 5.8, 4.8 Hz, 2H), 6.25 – 6.14 (m, 4H), 5.19 – 5.11 (m, 2H), 4.75 – 4.69 (m, 2H), 4.61 – 4.53 (m, 4H), 4.47 – 4.07 (m, 21H), 3.39 – 3.34 (m, 1.5H), 2.94 – 2.63 (m, 18.5H).

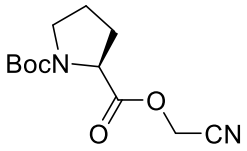


pdCpA 4.4.[41] Concentrated NH_4OH (490 μL) was added to a solution of protected pdCpA **4.3** (42 mg, 32 μmol) in dioxane (63 μL) and MeOH (420 μL), and the mixture was stirred for 24 h, then lyophilized. A solution of 4% NH_4OH/5% H_2O/MeCN in three column volumes was used to deactivate SiO_2, followed by washing with three column volumes of a solution of 5% H_2O/MeCN. The crude product was purified by NH_4OH-deactivated SiO_2 chromatography (2 × 15 cm, 5% H_2O/MeCN) to give **4.4** as a white solid (6 mg, 26%). The ^1H NMR matched the literature.[41] (D_2O, 400 MHz): δ 8.46 – 8.38 (m, 1H), 8.17 – 8.11 (m, 1H), 7.65 (dd, *J* = 33.4, 7.6 Hz, 1H), 6.09 (dd, *J* = 8.3, 5.8 Hz, 1H), 6.04 (d, *J* = 5.7 Hz, 1H), 5.99 – 5.91 (m, 1H), 4.77 – 4.71 (m, 3H), 4.57 – 4.52 (m, 1H), 4.37 – 4.32 (m, 1H), 4.26 – 4.20 (m, 1H), 4.06 – 3.88 (m, 3H), 2.81 (t, *J* = 5.9 Hz, 1H), 2.44 – 2.28 (m, 1H), 1.94 – 1.75 (m, 1H), 1.51 – 0.91 (m, 1H).

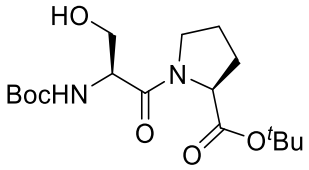


TBS–adenosine 4.5. TBS–Cl (430 mg, 2.9 mmol) in DMF (3 mL) was added to a suspension of adenosine (640 mg, 2.4 mmol) in pyridine (17 mL) at 0 °C, and stirred for 18 h at rt. It was then concentrated by SiO_2 chromatography (5 × 15 cm, 10% MeOH/DCM, R_f = 0.14) to yield **4.5** as a

white solid (795 mg, 87%). The ^{1}H NMR matched the literature.[50] (CD_3OD, 400 MHz): δ 8.38 (s, 1H), 8.20 (s, 1H), 6.05 (d, *J* = 4.2 Hz, 1H), 4.55 (t, *J* = 4.6 Hz, 1H), 4.36 (t, *J* = 5.0 Hz, 1H), 4.13 (dd, *J* = 5.2, 2.8 Hz, 1H), 4.01 (dd, *J* = 11.6, 3.0 Hz, 1H), 3.88 (dd, *J* = 11.6, 3.0 Hz, 1H), 3.30 (p, *J* = 1.6 Hz, 1H), 0.92 (s, 9H), 0.11 (d, *J* = 1.7 Hz, 6H).

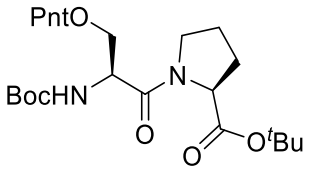


Boc–Pro–OCH$_2$CN 4.6. Bromoacetonitrile (390 mg, 3.3 mmol) in DMF (6 mL) was added to a solution of Boc-Pro-OH (550 mg, 2.6 mmol) and Et_3N (0.75 mL, 5.4 mmol) in DMF (6 mL) at 0 °C, and the mixture was stirred at rt for 16 hr. The reaction mixture was diluted with EtOAc (20 mL), washed with 1 M HCl (3 × 10 mL), $NaHCO_3$ (3 × 10 mL), H_2O (3 × 10 mL), and brine (10 mL). The organic layer was dried over Na_2SO_4 and concentrated to obtain **4.6** as a colorless oil (530 mg, 86%). The ^{1}H NMR matched the literature.[51] ($CDCl_3$, 400 MHz): δ 4.90 – 4.66 (m, 2H), 4.37 (dd, *J* = 8.5 Hz, 3.5 Hz, 0.45H), 4.32 (dd, *J* = 8.5 Hz, 4.0 Hz, 0.55H) 3.60 – 3.37 (m, 2H), 2.35 – 2.20 (m, 1H), 2.06 – 1.85 (m, 4H), 1.46 (s, 4H), 1.43 (s, 5H). The ^{13}C{^{1}H} NMR matched the literature.[51] ($CDCl_3$, 100 MHz, rotamer peaks are shown in parenthesis): δ 171.8, (171.6), (154.4), 153.4 , (114.2), 114.0, 80.5, (80.3), 58.6, (58.4), (48.8), 48.5, (46.5), 46.3, 30.9, (29.8), (28.4), 28.3, (24.5), 23.6.



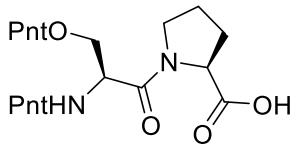
Boc–Ser–Pro–O^{t}Bu 4.9. A solution of DCC (310 mg, 1.5 mmol) and HOSu (170 mg, 1.5 mmol) in DMF (3.5 mL) was prepared and added to a mixture of Boc–Ser–OH (210 mg, 1.0 mmol), H–Pro–O^{t}Bu (200 mg, 1.1 mmol) and DIEA (0.73 mL, 4.3 mmol) in DMF (2.5 mL) at 0 °C. The

solution was allowed to warm to rt, and it was stirred for 16 h. The reaction mixture was then diluted with EtOAc (15 mL), washed with 1 M HCl (3 × 5 mL), $NaHCO_3$ (3 × 5 mL), H_2O (3 × 5 mL), and brine (2 × 5 mL). The organic layer was dried over Na_2SO_4 and concentrated. Dipeptide **4.9** was purified via SiO_2 chromatography (2 × 15 cm, 1% AcOH/55% EtOAc/hexanes) to give a white solid (323 mg 84%). The ^{1}H NMR matched the literature.[52] ($CDCl_3$, 400 MHz): δ 5.52 (br d, *J* = 8.3 Hz, 1H), 4.64 (br s, 1H), 4.58 (dd, *J* = 4.5, 4.6 Hz, 1H), 4.16 (dd, *J* = 12.0, 17.6 Hz, 1H), 3.92 (dd, *J* = 4.6, 11.3, 1H), 3.85 – 3.75 (m, 2H), 2.41 –2.25 (m, 2H), 2.16 –1.90 (m, 3H), 1.45 (s, 9H), 1.42 (s, 9H).

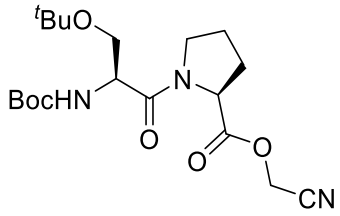


Boc–Ser(Pnt)–Pro–OH 4.10. To a solution of dipeptide **4.9** (178 mg, 0.50 mmol) and Et_3N (0.21 mL, 1.50 mmol) in DCM (3 mL) was added 4-pentenoyl chloride (71 mg, 0.60 mmol) in DCM (2 mL) at 0 °C. After 10 min the mixture was warmed to rt and stirred for 3 h. The solution was then diluted with EtOAc (20 mL), washed with 1 M HCl (3 × 10 mL), Na_2CO_3 (3 × 10 mL), and brine (1 × 10 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (3 × 15 cm, 35% EtOAc/hexanes, R_f = 0.31) to give pentenoyl ester **4.10** as a colorless oil (215 mg, 98%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 5.88 – 5.78 (m, 1H), 5.47 (br d, *J* = 8.78, 1H), 5.09 – 4.97 (m, 2H), 4.84 – 4.74 (m, 0.85H), 4.64 – 4.56 (m, 0.32H), 4.47 (dd, *J* = 11.37 Hz, 4.8 Hz, 0.92H), 4.40 (dd, *J* = 8.79 Hz, 4.44 Hz, 0.92 H), 4.20 (dd, *J* = 10.88 Hz, 5.89 Hz, 0.17H), 4.15 – 4.09 (m, 0.17H), 4.00 (dd, *J* = 11.4 Hz, 8.1 Hz, 0.89H), 3.86 – 3.62 (m, 2H), 3.53 – 3.42 (m, 0.16 H), 2.50 – 2.32 (m, 4H), 2.30 – 1.84 (m, 4H), 1.46 – 1.42 (m, 18H); ^{13}C {^{1}H} NMR ($CDCl_3$, 100 MHz, rotamer peaks are shown in parenthesis): δ 172.9, (172.5), 170.6, (170.5), (168.3), 167.5, 155.4, (154.7), 136.7, (136.6), (115.6), 115.4, (82.9), 81.4, 79.8,

(79.6), (64.8), 64.1, (60.0), 59.8, 51.4, (51.2), 47.1, (46.4), 33.3, (31.1), 29.0, 28.6, (28.6), 28.3, (28.3), 28.0, (27.8), 24.8, (22.2); HRMS (ESI$^+$/TOF) *m/z*: [M + H]$^+$ Calcd for $C_{22}H_{37}N_2O_7^+$ 441.2595; Found 441.2578.

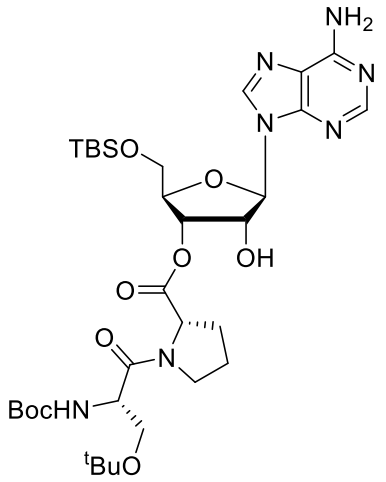


Pnt–Ser(Pnt)–Pro–OH 4.11. TFA (375 μL) was added to a solution of dipeptide **4.10** (220 mg, 0.5 mmol) in DCM (1.1 mL) and stirred at rt for 16 h. The reaction was then concentrated and neutralized with $NaHCO_3$. The crude mixture was taken up in DMF (2.5 mL), a solution of Pnt–OSu (150 mg, 0.75 mmol) and DIEA (520 μL, 3 mmol) in DMF (0.5 mL) was added, and the mixture was stirred at rt for 4 h. The solution was acidified with 1 M HCl to pH 2 and extracted with EtOAc (4 × 15 mL). The combined organic layers were washed with H_2O (4 × 20 mL) and brine (20 mL), then dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (2 × 15 cm, 1% MeOH/20% acetone/DCM; R_f = 0.29) to give **4.11** as a pale-yellow oil (168 mg, 92%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 5.89 – 5.63 (m, 2H), 5.23 – 4.87 (m, 5H), 4.50 – 3.50 (m, 6H), 2.55 – 1.93 (m, 13H).



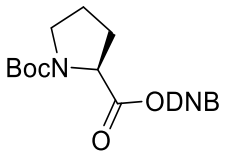
Boc-Ser(*t*Bu)-Pro-OCH$_2$CN 4.13. Boc-Pro-OCH$_2$CN **4.6** (1.2 g, 4.6 mmol) was stirred in 4 M HCl in dioxane (11 mL) for 1 h and then concentrated to dryness overnight give the HCl·H–Pro–OCH$_2$CN salt in quantitative yield. Activated Boc–Ser(*t*Bu)–6-ClOBt was prepared by adding DCC (1.0 g, 5.0 mmol) in DCM (7.2 mL) to a solution of Boc-Ser(*t*Bu)-OH (1.3 g, 5.0 mmol) and

6-Cl-HOBt (860 mg, 5.0 mmol) in DMF (7.2 mL) at 0 °C. The solution was stirred at rt for 16 h, filtered through Celite, rinsed with DCM, and concentrated to remove DCM. The resultant solution was cooled to –40 °C, and the prolyl HCl salt (870 mg, 4.6 mmol) in DMF (11 mL) was added, followed by slow addition of DIEA (1.6 mL, 9.2 mmol). The mixture was stirred at –40 °C for 3 h, then at rt for 2 h. The mixture was then diluted with EtOAc (40 mL), washed with 1 M HCl (3 × 15 mL), sat. $NaHCO_3$ (3 × 15 mL), and brine (15 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 10% acetone/DCM; R_f = 0.59) to give **4.13** as a colorless oil (1.5 g, 89%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 5.28 – 5.15 (m, 0.8H), 5.07 –5.96 (m, 0.2H), 4.84 – 4.62 (m, 2H), 4.60 – 4.24 (m, 2H), 3.94 – 3.84 (m, 0.7H), 3.75 – 3.33 (m, 3.4H), 2.30 – 1.8 (m, 4H), 1.40 (s, 6.6H), 1.38 (s, 2.4H), 1.15 (s, 7H), 1.09 (s, 2H). ^{13}C {^{1}H} NMR ($CDCl_3$, 100 MHz, rotamer peaks are shown in parenthesis): δ (171.1), 171.0, (170.5), 170.4, 155.3, (155.0), 114.3, (114.2), (79.8), 79.8, 73.5, (71.3), (64.0), 63.0, (58.7), 58.6, 52.3, (51.8), (49.4), 48.8, 47.2, (46.4), (30.9), 29.0, 28.4, (28.4), (27.4), 27.3, 25.0, (22.7). HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{19}H_{32}N_3O_6^+$ 398.2286; Found 398.2277.



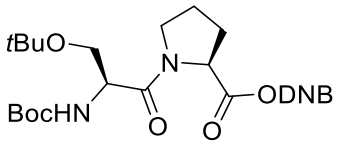
TBS–Dinucleotide-dipeptide 4.14. The activated dipeptide **4.13** (69 mg, 170 μmol) and TBS–adenosine **4.5** were stirred at 100 °C in 9:1 DMF:im (*m*:*m*) (1 mL) for 24 h. The solution was then

concentrated, and the crude residue was purified by SiO_2 chromatography (2 × 15 cm, 3% MeOH/ DCM; R_f = 0.21) to give **4.14** as a white solid (53 mg, 42%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.32 – 8.22 (m, 1H), 8.17 – 8.06 (m, 1H), 6.35 – 6.12 (m, 2H), 6.07 (dd, *J* = 16.3, 4.7 Hz, 1H), 5.80 (dd, *J* = 21.0, 8.9 Hz, 1H), 5.53 – 5.36 (m, 1H), 5.15 – 5.00 (m, 1H), 4.85 – 4.76 (m, 1H), 4.64 – 4.53 (m, 1H), 4.49 – 4.37 (m, 1H), 4.36 – 4.26 (m, 1H), 3.99 – 3.29 (m, 6H), 2.37 – 1.75 (m, 4H), 1.41 – 1.28 (m, 9H), 1.20 – 1.04 (m, 9H), 0.89 – 0.79 (m, 9H), 0.08 – 0.00 (m, 6H). ^{13}C {^{1}H} NMR ($CDCl_3$, 400 MHz): δ 171.9, 171.4 171.03, 171.94, 155.71, 155.69, 155.6, 155.5, 153.1, 153.0, 149.7, 138.82, 138.75, 119.9, 89.5, 88.9, 86.0, 85.3, 82.9, 82.4, 80.3, 79.8, 74.4, 74.2, 74.0, 73.72, 73.68, 73.54, 73.46, 70.4, 63.9, 62.8, 62.7, 62.5, 59.8, 59.4, 59.1, 53.0, 52.4, 51.9, 47.7, 46.4, 30.9, 29.8, 29.2, 29.1, 28.43, 28.35, 27.42, 27.37, 27.2, 26.01, 25.99, 25.97, 25.93, 25.91, 25.3, 25.0, 22.7, 18.5, 18.4, –5.3, –5.37, –5.43, – 5.48, –5.54. HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{33}H_{56}N_7O_9Si^+$ 722.3903; Found 722.3897.



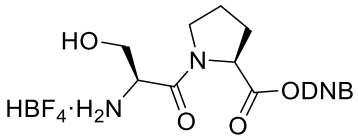
Boc–Pro–ODNB 4.15. 3,5-Dinitrobenzyl chloride (500 mg, 2.3 mmol) in DMF (1.1 mL) was added to Boc–Pro–OH (460 mg, 2.1 mmol) and DIEA (570 μL, 3.3 mmol) in DMF (1 mL) and stirred at rt for 16 h. The mixture was diluted with EtOAc (15 mL), washed with 1 M HCl (3 × 5 mL), $NaHCO_3$ (3 × 5 mL), and brine (5 mL). The organic phase was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (3 × 15 cm, 25% EtOAc/hexanes; R_f = 0.24) giving **4.15** as a pale-yellow oil (800 mg, 95%). ^{1}H NMR ($CDCl_3$, 400 MHz, mixture of rotamers in 6:4 ratio): δ 9.02 (s, 0.4H), 8.99 (s, 1.2H), 8.58 (s, 0.8H), 8.56 (s, 0.6H), 5.42 – 5.28 (m, 2H), 4.42 (dd, *J* = 8.3 Hz, 3.9 Hz, 0.6H), 4.36 (dd, *J* = 8.4 Hz, 3.7 Hz, 0.4 H), 3.63 – 3.38 (m, 2H), 2.38 – 2.20 (m, 1H), 2.07 – 1.85 (m, 3H), 1.44 (s, 5.4H), 1.36 (s, 3.6H).

^{13}C {1H} NMR ($CDCl_3$, 100 MHz, rotamer peaks are shown in parenthesis): δ (172.7), 172.6, 154.5, (153.5), (148.7), 148.7, 140.6, (140.2), (127.8), 127.6, (118.6), 118.4, 80.2, (80.1), 64.1, (64.0), (58.9), 58.8, 46.6, (46.4), (31.0), 30.0, 28.4, (28.3), 24.6, (23.7).

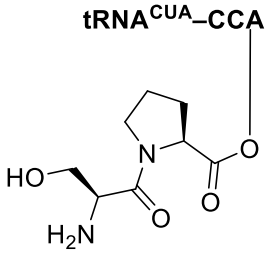


Boc–Ser(ᵗBu)–Pro–ODNB 4.16. Boc–Pro–DBE **4.15** (450 mg, 1.1 mmol) was dissolved in 4 M HCl in dioxane (3 mL), and after 1 h the solution was concentrated to give the HCl·H–Pro–ODBE salt (not shown) as a white solid in quantitative yield (370 mg). Activated Boc–Ser(ᵗBu) –6–Cl–OBt was prepared by adding DCC (360 mg, 1.7 mmol) in DCM (3 mL) was added to a solution of Boc–Ser(ᵗBu)–OH (450 mg, 1.7 mmol) and 6-Cl-HOBt (290 mg, 1.7 mmol) in DMF (2.5 mL) at 0 °C. The solution was stirred at rt for 16 h, filtered through Celite, and concentrated to remove DCM. The resultant solution was cooled to –40 °C, the HCl·H–Pro–ODBE salt (370 mg, 1.1 mmol) in DMF (3 mL) was added, followed by slow addition of DIEA (280 μL, 1.7 mmol), and the mixture was stirred at –40 °C for 5 h. The mixture was then diluted with EtOAc (20 mL), washed with 1 M HCl (3 × 5 mL), $NaHCO_3$ (3 × 5 mL), and brine (5 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (3 × 15 cm, 10% EtOAc/DCM, 2.5% acetone; R_f = 0.42) to give **4.16** as a white solid (525 mg, 88%). ^{1}H NMR ($CDCl_3$, 400 MHz, mixture of rotamers in 8:2 ratio): δ 9.02 (br t, *J* = 2.1 Hz, 0.2H), 8.99 (br t, *J* = 2.1 Hz, 0.8H), 8.58 (br d, *J* = 2.1 Hz, 2H), 5.50 – 5.06 (m, 3H), 4.60 (dd, *J* = 14.0 Hz, 8.1 Hz, 0.8H) 4.55 (dd, *J* = 8.7 Hz, 4.7 Hz, 0.8 Hz), 4.49 – 4.39 (m, 0.2H), 3.98 – 3.88 (m, 0.8H), 3.77 – 3.64 (m, 1H), 3.61 – 3.44 (m, 2H), 3.36 (dd, *J* = 10.4 Hz, 8.1 Hz, 0.2H) 2.34 – 2.18 (m, 1H), 2.16 – 1.86 (m, 3H), 1.43 (s, 7.2H), 1.35 (s, 1.8H), 1.13 (s, 9H). ^{13}C {^{1}H} NMR ($CDCl_3$, 100 MHz, rotamer peaks are shown in parenthesis): δ (171.8), 171.6, (170.9), 170.5, 155.3, (154.8), (148.8),

148.7, 140.6, (140.1), (128.0), 127.7, (118.6), 118.4, 79.8, (79.6), (73.5), 73.4, (64.9), (64.1), 64.1, 63.0, 59.0, (58.9), 52.2, (51.8), 47.2, (46.4), (30.8), 29.2, 28.4, (28.2), (27.3), 27.2, 25.1, (22.7). HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{24}H_{35}N_4O_{10}^+$ 539.2348; Found 539.2346.

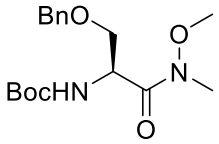


HBF_4·H–Ser–Pro–ODNB 4.17. A solution of 50% HBF_4·Et_2O (11 µL, 39 µmol) was added to a solution of Boc–Ser(^{t}Bu)–Pro–ODNB **4.16** (21 mg, 39 µmol) in DCM (78 µmol) at rt. After stirring for 30 min, the solution was concentrated, diluted with H_2O (0.5 mL), and washed with Et_2O (3 × 0.25 mL). The aqueous phase was lyophilized to give **4.17** as a white solid (23 mg, 98%). ^{1}H NMR (D_2O, 400 MHz, mixture of rotamers in 8:2 ratio) δ 8.83 – 8.87 (m, 1H), 8.70 (br d, *J* = 2.1 Hz, 0.2H), 8.65 (br d, *J* = 2.0 Hz, 1.8H), 8.11 (br s, 2H), 5.48 – 5.35 (m, 3.3 H), 4.55 – 4.48 (m, 1H), 4.19 (dd, *J* = 7.8 Hz, 4.0 Hz, 0.9H), 4.07 (dd, *J* = 9.1 Hz, 3.4 Hz, 0.1H), 3.90 (t, *J* = 5.7 Hz, 0.15H), 3.83 – 3.38 (m, 4.85H), 2.36 – 2.20 (m, 1H), 2.04 – 1.89 (m, 3H); ^{13}C {^{1}H} NMR (DMSO-*d*$_6$, 100 MHz, rotamer peaks are shown in parenthesis): δ 169.2, (169.1), 164.5, (164.4), (148.1), 148.8, 147.5, (147.4), (128.0), 126.3, (116.9), 116.8, 61.2, (61.1), 60.0, (59.9), 58.4, (58.3), 56.8, (58.7), (44.5), 44.6, (30.7), 28.8, 22.1, (22.0). HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{15}H_{19}N_4O_8^+$ 383.1195; Found 383.1197.

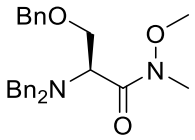


$tRNA^{CUA}$–Dipeptide 4.18. Procedure done by Karla Piedl following the method from Murakami et al with minor modification.[38] $tRNA^{CUA}$ (1 µL, 50 µM) and HK Buffer (1.5 µL, 400 mM HEPES-

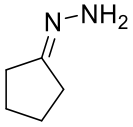
K, 400 mM KCl, pH 8.0) were incubated at 95 °C for 3 min in an RNase-free 1.5 mL Eppendorf tube. The reaction was allowed to cool to rt for 5 min, then HBF_4·H–Ser–Pro–ODBE (1 µL, 25 µM in DMSO) was added and the mixture was incubated at 0 °C for 6 h. The reaction was stopped by addition of aqueous NaOAc (100 µL, 250 mM, pH 5.0). The RNA was recovered by addition of aqueous NH_4OHOAc (0.2 volumes, 7.5 M) and cold EtOH (≥95%, 2.5 volumes) and incubation at –20 °C for 16 h. RNA was recovered by centrifugation at 9,000 × g at 4 °C for 30 min. The supernatant was poured off and the remaining EtOH was removed by pipette. The pellet was rinsed with EtOH (70%, 300 mL) and incubated at rt for 2 min, then centrifuged at 9,000 × g at 4 °C for 15 min. The supernatant was carefully poured off and the remaining EtOH was removed by pipette. The pellet was incubated at rt until all traces of EtOH were evaporated (ca. 5 min). The aminoacylated tRNA pellet was dissolved in RNase-free water (20 µL) and stored at –80 °C.



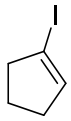
Weinreb Amide 4.19. 6-Cl-HOBt (2.0 g, 12 mmol), DCC (2.5 g, 12 mmol) and DMAP (30 mg, 0.25 mmol) were added to a solution of Boc–Ser(Bn)–OH (3.0 g, 10 mmol), *N*,*O*-dimethylhydroxylamine·HCl (2.0 g, 20 mmol), and DIEA (7 mL, 40 mmol) in 1:1 DCM:DMF (100 mL) at 0 °C. The reaction was allowed to stir at rt for 24 h, after which time it was filtered through Celite and concentrated. The resulting residue was diluted with EtOAc (150 mL), washed with NH_4Cl (2 × 50 mL), $NaHCO_3$ (2 × 50 mL), and brine (50 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 50% EtOAc/hexanes) to give **4.19** as a colorless oil (3.2 g, 93%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.35-7.23 (m, 5H), 5.42 (d, *J* = 8.5 Hz, 1H), 4.87 (br s, 1H), 4.56 (d, *J* = 12.5 Hz, 1H), 4.49 (d, *J* = 12.5 Hz, 1H), 3.71 (s, 3H), 3.66 (m, 2H), 3.17 (s, 3H), 1.43 (s, 9H).



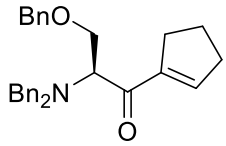
***N,N,O*-Tribenzyl Serine Weinreb Amide 4.20**. Weinreb amide **4.19** (880 mg, 2.6 mmol) was dissolved in 25% TFA/DCM (20 mL) and stirred at rt for 45 min. It was then concentrated, neutralized with $NaHCO_3$, and extracted with DCM (4 × 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. To the crude oil in DCM (4.5 mL) was added BnBr (1.0 mL, 8.4 mmol) and DIEA (2.2 mL, 13 mmol), and the reaction was stirred for 24 h. It was then diluted with EtOAc (50 mL), washed with NH_4Cl (4 × 25 mL), and brine (25 mL), dried over Na_2SO_4 and concentrated. The product was isolated by SiO_2 chromatography (5 × 15 cm, 25% EtOAc/hexanes) to give **4.20** as a colorless oil (1.0 gg, 97%). The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.40-7.17 (m, 15H), 4.56 (d, *J* = 11.9 Hz, 1H), 4.48 (d, *J* = 11.9 Hz, 1H), 4.13 (m, 1H), 3.98 – 3.84 (m, 4H), 3.76 (d, *J* = 14.1 Hz, 2H), 3.28 (br s, 3H), 3.20 (br s, 3H).



Cyclopentanone Hydrazone 4.40. Cyclopentanone (14 mL, 160 mmol) was added slowly to a flask containing 65% aqueous hydrazine (22 mL, 180 mmol) at rt, then it was refluxed at 125 °C for 16 h. It was then quenched with H_2O (100 mL) and extracted with $CHCl_3$ (4 × 50 mL). The combined organic layers were washed with brine (50 mL), dried over Na_2SO_4, and concentrated to give **4.40** as a yellow liquid (15 g, 96%). The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 4.80 (br s, 2H), 2.33 – 2.30 (m, 2H), 2.16 – 2.12 (m, 2H), 1.85 – 1.67 (m, 4H).

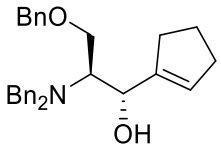


Iodocyclopentene 4.41. Tetramethylguanidine (TMG) (76 g, 660 mmol) in Et_2O (120 mL) was added to a stirring solution of I_2 (38 g, 150 mmol) in Et_2O (200 mL) at 0 °C over 60 min. After addition, hydrazone **4.40** (7.2 g, 73 mmol) in Et_2O (78 mL) was added over 75 min at 0 °C then the reaction was stirred at rt for 16 h. The mixture was filtered and concentrated, then heated at 85 °C for 2 h, diluted with Et_2O (250 mL), washed with 1 M HCl (4 × 55 mL), $Na_2S_2O_3$ (4 × 55 mL), $NaHCO_3$ (4 × 55 mL) and brine (55 mL). The organic layer was dried over Na_2SO_4 and concentrated to give **4.41** as a pale-yellow liquid (8.7 g, 62%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 6.12 – 6.10 (m, 1H), 2.64 – 2.58 (m, 2H), 2.36 – 2.30 (m, 2H), 1.98 – 1.90 (m, 2H).

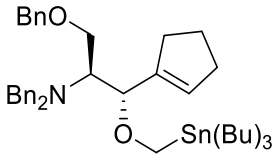


Ketone 4.21. *sec*-BuLi (1.4 M in cyclohexane; 29 mL, 41 mmol) was added to a solution of iodocyclopentene **4.41** (6.5 g, 33 mmol) in THF (65 mL) at –40 °C and stirred for 70 min. Weinreb amide **4.20** (4.7 g, 11.0 mmol) in THF (19 mL) at –40 °C was added slowly via cannula, and the solution was stirred at –40 °C for 1 h. The reaction was quenched with NH_4Cl (15 mL), diluted with EtOAc (375 mL), washed with NH_4Cl (3 × 50 mL) and brine (50 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 5% EtOAc/hexanes) to give **4.21** as a yellow oil (4.3 g, 92%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.39 – 7.20 (m, 15H), 6.11 (m, 1H), 4.55 (d, J = 12.3 Hz, 1H),

4.48 (d, *J* = 12.3 Hz, 1H), 4.24 (app t, *J* = 6.6 Hz, 1H), 3.90 (d, *J* = 6.6 Hz, 2H), 3.79 (d, *J* =13.6 Hz, 2H), 3.71 (d, *J* = 14.1 Hz, 2H), 2.59 – 2.39 (m, 4H), 1.98 – 1.84 (m, 2H).

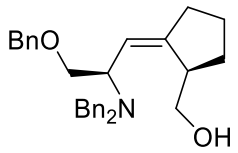


Alcohol 4.22. $LiAlH_4$ (3.7 g, 98 mmol) was added to a solution of ketone **4.21** (4.2 g, 9.8 mmol) in THF (160 mL) and stirred at rt for 1 hr. The reaction was cooled to 0 °C and quenched with MeOH (40 mL) then NH_4Cl (40 mL), diluted with EtOAc (320 mL), washed with NH_4Cl (1 × 95 mL) and 1 M potassium sodium tartrate (2 × 95 mL). The combined aqueous layers were extracted with DCM (3 × 130 mL). The combined organic layers were washed with brine (95 mL), dried over Na_2SO_4, and concentrated to give alcohol **4.22** as a yellow oil (4.6 g, 98%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.49 – 7.24 (m, 15H), 5.65 (m, 1H), 4.62 (d, *J* = 11.9 Hz, 1H), 4.53 (d, *J* = 11.9 Hz, 1H), 4.48 (s, 1H), 4.26 (d, *J* = 10.1 Hz, 1H), 4.02 (d, *J* = 13.2 Hz, 2H), 3.80 – 3.70 (m, 3H), 3.58 (dd, *J* = 10.6 Hz, 3.1 Hz, 1H), 3.07 (m, 1H), 2.43 – 2.17 (m, 3H), 2.00 – 1.75 (m, 3H).

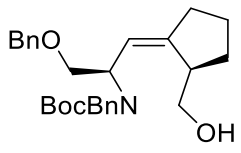


Stannane Ether 4.23. Potassium hydride (560 mg, 6.9 mmol) was added to a solution of 18-crown-6 (3.6 g, 14 mmol) and alcohol **4.22** (2.0 g, 4.6 mmol) in THF (30 mL) and stirred for 30 min. Then $(Bu)_3SnCH_2I$ (3.0 g, 6.9 mmol) in THF (10 mL) was added dropwise and the reaction was stirred for 1 h. It was then quenched with MeOH (50 mL), diluted with EtOAc (400 mL), washed with NH_4Cl (2 × 100 mL) and brine (100 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 3% EtOAc/hexanes) to give **4.23**

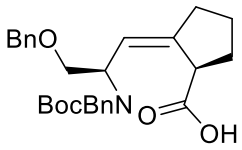
as a pale-yellow oil (3.1 g, 92%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.40 – 7.26 (m, 15H), 5.60 (br s, 1H), 4.45 (d, *J* = 12.0 Hz, 1H), 4.37 (d, *J* = 12.0 Hz, 1H), 4.05 (d, *J* = 7.8 Hz, 1H), 3.99 (d, *J* = 13.7 Hz, 2H), 3.83 (d, *J* = 13.7 Hz, 2H), 3.74 (d, *J* = 9.9 Hz, 1H), 3.60 (dd, *J* = 9.6, 5.7 Hz, 1H), 3.53 (dd, *J* = 9.6, 4.6 Hz, 1H), 3.41 (d, *J* = 9.6 Hz, 1H), 2.99 – 2.94 (m, 1H), 2.40 – 2.28 (m, 2H), 2.00 –1.94 (br s, 2H), 1.83 – 1.74 (m, 2H), 1.58 – 1.47 (m, 6H), 1.36 – 1.22 (m, 6H), 1.01 – 0.83 (m, 15H).



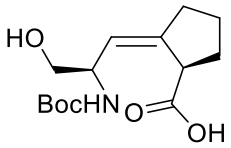
(*Z*)-Alkene 4.24. *n*-BuLi (2.5 M in hexanes, 3.7 mL, 9.2 mmol) in THF (10 mL) at –78 °C was added to a solution of stannane ether **4.23** in THF (25 mL) at –78 °C by cannula and stirred at –78 °C for 90 min. The reaction was then quenched with MeOH (25 mL) and concentrated. The residue was diluted with a 1:1 solution of MeOH:NH_4Cl (100 mL) and the product was extracted with $CHCl_3$ (3 × 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 15% EtOAc/hexanes) to give **4.24** as a colorless oil (669 mg, 46%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.38 – 7.26 (m, 15H), 5.55 (br d, *J* = 8.7 Hz, 1H), 4.57 (d, *J* = 12.2 Hz, 1H), 4.53 (d, *J* = 12.2 Hz, 1H), 4.12 (br s, 1H), 3.89 (d, *J* = 13.3, 2H), 3.79 (m, 1H), 3.67 (m, 4H), 3.31 (dd, *J* = 10.6, 6.3 Hz, 1H), 3.23 (dd, *J* = 10.7, 8.3 Hz, 1H), 2.56 – 2.47 (m, 1H), 2.31 – 2.18 (m, 2H), 1.71 – 1.47 (m, 4H).



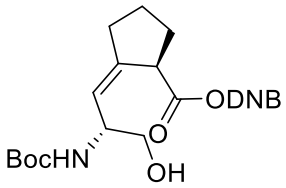
Boc-Benzylamine 4.25. Formic acid (6.5 mL, 170 mmol) was added to a suspension of (*Z*)-alkene **4.24** (800 mg, 1.8 mmol) and 20% $Pd(OH)_2/C$ (110 mg) in MeOH (20 mL) and stirred for 20 min. It was then filtered through Celite and concentrated. The residue was neutralized with $NaHCO_3$ and the product was extracted with $CHCl_3$ (4 × 75 mL) to give 489 mg of a pale-yellow oil (77% crude yield). The combined organic layers were dried over Na_2SO_4 and concentrated. The crude product was dissolved in DCM (35 mL) and Boc_2O (760 mg, 3.5 mmol) was added and stirred for 16 h. The reaction was concentrated and purified by SiO_2 chromatography (5 × 15 cm, 20% EtOAc/hexanes) to give **4.25** as a colorless oil (546 mg, 67%, 2 steps). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.36 – 7.16 (m, 10H), 5.36 (br d, *J* = 8.9 Hz, 1H), 5.18 (br s, 1H), 4.47 – 4.37 (m, 4H), 3.48 – 3.46 (m, 4H), 2.87 (br s, 1H), 2.30 – 2.11 (m, 2H), 1.77 – 1.25 (m, 15H).



Acid 4.26. Freshly-prepared Jones reagent (2.5 M, 1.5 L, 3.8 mmol) was added slowly to a solution of Boc-benzylamine **4.25** (480 mg, 1.1 mmol) in acetone (60 mL) at 0 °C. The reaction was stirred for 30 min at 0 °C then it was quenched with *i*-PrOH (45 mL) and stirred for an additional 10 min. The mixture was filtered through Celite, diluted with H_2O (75 mL), and the product was extracted with DCM (4 × 30 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 20% EtOAc/hexanes) to give **4.26** as a pale-yellow oil (423 mg, 90%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 10.40 (br s, 1H), 7.34 – 7.16 (m, 10H), 5.59 (br s, 1H), 4.92 – 4.55 (m, 1H), 4.50 – 4.17 (m, 4H), 3.69 –3.24 (m, 3H), 2.54 –2.38 – 2.22 (m, 1H), 2.21 – 1.99 (m, 1H), 1.98 – 1.77 (m, 2H), 1.66 – 1.55 (m, 1H), 1.38 (br s, 9H).

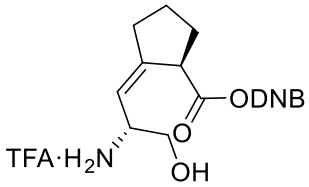


Boc–Ser–Ψ[(*Z*)CH=C]-Pro–OH 4.27. NH_3 (ca. 30 mL) was distilled into a suspension of Na (ca. 390 mg) in THF (2 mL) at –33 ºC and stirred until a deep blue color was maintained. A solution of acid **4.26** (380 mg, 0.82 mmol) in THF (5 mL) was added slowly and the reaction was stirred at –33 °C for 15 min, after which time it was warmed to rt (caution! NH_3 evolved) until ca. 10 mL NH_3 remained. It was then diluted with NH_4Cl (15 mL), acidified to pH 7 with 1 M HCl, and extracted with $CHCl_3$ (4 × 10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated to give **4.27** as a pale-yellow oil (199 mg, 85%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 6.80 (br s, 1H), 5.41 (br s, 1H), 5.21 (br s, 1H), 4.28 (br s, 1H), 3.60 (d, *J* = 6.2 Hz, 2H), 3.50 (br s, 1H), 2.48 – 2.37 (m, 1H), 2.35 – 2.23 (m, 1H), 2.11 – 1.91 (m, 2H), 1.89 – 1.74 (m, 1H), 1.64 – 1.51 (m, 1H), 1.40 (s, 9H).

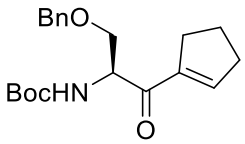


Boc–Ser–Ψ[(*Z*)CH=C]-Pro–ODNB 4.28. 3,5-dinitrobenzyl chloride (65 mg, 300 μmol) was added to a solution of Boc–Ser–Ψ[(*Z*)CH=C]-Pro–OH **4.27** (83 mg, 290 μL) and DIEA (55 μL, 320 μL) in DMF and stirred at rt for 16 h. The reaction was diluted with EtOAc (2.5 mL), washed with 1 M HCl (2 × 1 mL), $NaHCO_3$ (2 × 1 mL), and brine (1 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (2 × 15 cm, 2.5% MeOH/DCM; R_f = 0.28) to give Boc–Ser–Ψ[(*Z*)CH=C]-Pro–ODNB **4.28** as a yellow oil (86 mg, 64%). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.99 (dd, *J* = 2.1, 2.1 Hz, 1H), 8.59 (d, *J* = 2.1 Hz, 2H), 5.47 (br d, *J* = 9.4 Hz, 1H), 5.42 – 5.26 (m, 2H), 4..78 (br d, *J* = 7.3 Hz, 1H), 4.37 (d, *J* =

8.3 Hz, 1H), 3.67 (ddt, *J* = 8.7, 5.2, 1.9 Hz, 1H), 3.60 (d, *J* = 5.4 Hz, 2H) 2.60 (br s, 1H), 2.55 – 2.45 (m, 1H), 2.43 – 2.32 (m, 1H), 2.17 – 1.99 (m, 2H), 1.94 – 1.81 (m, 1H), 1.74 – 1.61 (m, 1H), 1.34 (s, 9H). ^{13}C {^{1}H} NMR ($CDCl_3$, 100 MHz): δ 173.9, 156.0, 148.7, 145.3, 141.0, 127.8, 151.6, 118.4, 79.8, 66.2, 64.3, 52.7, 456.0, 34.2, 31.8, 28.4, 24.9. $[M + H]^+$ Calcd for $C_{21}H_{28}N_3O_9^+$ 466.1820; Found 466.1829.

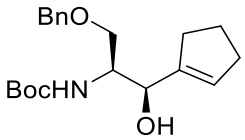


TFA·H–Ser–Ψ[(*Z*)CH=C]-Pro–ODNB 4.29. Boc–Ser–Ψ[(*Z*)CH=C]-Pro–ODNB **4.28** (15 mg, 32 μmol) was dissolved in 25% TFA/DCM (1.1 mL) at 0 °C, stirred at 0 ºC for 45 min, and concentrated under reduced pressure. The crude product was purified by SiO_2 chromatography (0.5 × 3 cm, 0% – 5% MeOH/DCM, hold 5% MeOH/DCM R_f = 0.32) to give **4.29** as a brown oil (4 mg, 22%). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 9.02 (d, *J* = 2.2, 2.2 Hz, 1H), 8.58, *J* = 2.1 Hz, 1H), 6.71, (d, *J* = 5.9 Hz, 1H), 5.61 (dq, *J* = 9.6, 2.1 Hz, 1H), 5.43 (br s, 1H), 5.40 – 5.20 (m, 2H), 4.75 – 4.67 (m, 1H), 3.85 – 3.72 (m, 2H), 3.69 (ddt, *J* = 8.7, 4.7, 1.9 Hz, 1H), 2.59 – 2.51 (m, 2H), 2.46 – 2.39 (m, 2H), 2.26 – 2.01 (m, 4H). ^{13}C NMR ($CDCl_3$, 150 MHz): δ 173.3, 156.8, 156.6, 148.6, 147.1, 140.55, 127.8, 119.4, 118.4, 64.5, 64.4, 51.6, 46.0, 34.2, 31.7, 24.8. $[M + CF_3CO_2Na – H_2O]^+$ Calcd for $C_{16}H_{18}N_3O_6NaCF_3CO_2$ 484.0948; Found 484.0942

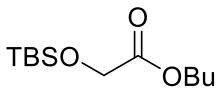


Ketone 4.31. *sec*-BuLi (6.6 mL, 9.26 mmol) was added to a solution of iodocyclopentene **4.41** (1.4 mg, 7.3 mmol) in THF (14 mL) at –40 °C and stirred for 75 min. Weinreb amide **4.19** (856 mg, 2.53 mmol) in THF (4.3 mL) at –40 °C was added slowly via cannula and the reaction was

stirred at –40 ºC for 90 min. The mixture was quenched with NH_4Cl (5 mL), diluted with EtOAc (75 mL), washed with 1 M HCl (3 × 10 mL) and brine (10 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 10% EtOAc/hexanes) to give **4.31** as a colorless oil (800 mg, 92%). The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.34 – 7.22 (m, 5H), 6.79 (m, 1H), 5.57 (d, *J* =10.5 Hz, 1H), 5.00 (m, 1H), 4.54 (d, *J* = 12.4 Hz, 1H), 4.43 (d, *J* = 12.0 Hz, 1H), 3.71 (d, *J* = 4.4 Hz, 2H), 2.62 (m, 1H), 2.54 (m, 3H), 2.00 – 1.82 (m, 2H), 1.44 (s, 9H).

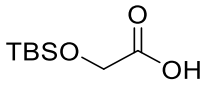


Alcohol 4.32. $NaBH_4$ (630 mg, 17 mmol) was added to a solution of ketone **4.31** (2.8 g, 8.2 mmol) and $CeCl_3$ (2.5 g, 10.2 mmol) in 2.5:1 THF:MeOH (93 mL) at 0 °C and stirred at 0 °C for 2 h. The reaction was quenched with NH_4Cl (35 mL), diluted with EtOAc (150 mL), washed with NH_4Cl (2 × 75 mL) and brine (75 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 10% EtOAc/hexanes) to give colorless oil **4.32** as a pair of diastereomers (2.6 g, 91%). The major diastereomer was isolated by EtOAc/*n*-heptane precipitation at –20 °C. The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.36 – 7.28 (m, 5H), 5.65 (m, 1H), 5.35 (d, *J* = 8.4 Hz, 1H), 4.51 (d, *J* = 11.6 Hz, 1H), 4.42 (d, *J* = 12.0 Hz, 1H), 4.33 (br s, 1H), 3.84 (br s, 1H), 3.71 – 3.68 (dd, *J* = 13.4 Hz, 3.4 Hz, 1H), 3.60 – 3.55 (dd, *J* = 9.4 Hz, 2.6 Hz, 1H), 3.18 (d, *J* = 8.4 Hz, 1H), 2.35 – 2.20 (m, 4H), 1.87 (m, 2H), 1.44 (s, 9H).

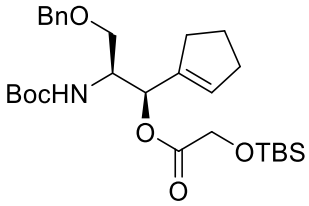


Butyl Ester 4.42. TBS-Cl (1.3 g, 8.4 mmol) in DCM (1 mL) was added to a solution of butyl glycolate (1.0 g, 7.6 mmol) and imidazole (1.6 g, 23 mmol) in DMF (1 mL) at 0 °C. The reaction

was allowed to warm to rt and stirred for 16 h. The mixture was diluted with EtOAc (30 mL), washed with 1 M HCl (2 × 5 mL), and brine (5 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 10% EtOAc/hexanes, R_f = 0.46) to give **4.42** as a colorless oil (1.7 g, 90%). The ^{1}H NMR matched the literature.[34, 53] ($CDCl_3$, 400 MHz): δ 4.23 (s, 2H), 4.14 (t, *J* = 6.7 Hz, 2H), 1.63 (p, *J* = 6.4 Hz, 2H), 1.38 (sextet, *J* = 7.5 Hz, 2H) 0.93 (t, *J* = 7.4 Hz, 3H), 0.92 (s, 9H), 0.11 (s, 6H).

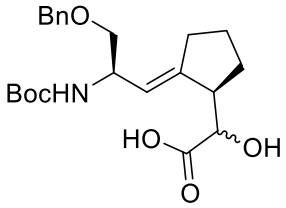


TBS Acid 4.43. KOH (380 mg, 6.7 mmol) in 2:1 H_2O:MeOH (2.4 mL) was added to butyl ester **4.42** (1.7 g, 6.7 mmol) in THF (3.5 mL) at 0 °C. After 30 min, the reaction was allowed to warm to rt, and stirred for 6 h. Then it was diluted with H_2O (20 mL) and washed with Et_2O (2 × 10 mL). The aqueous phase was acidified to pH 1 with 1 M HCl and the product was extracted with Et_2O (3 × 25 mL). The combined organic phases were dried over Na_2SO_4 and concentrated, giving **4.43** as a colorless oil that solidified upon refrigeration (1.1 g, 86%). The ^{1}H NMR matched the literature.[53] ($CDCl_3$, 400 MHz): δ 4.23 (s, 2H), 0.93 (s, 9H), 0.14 (s, 6H).

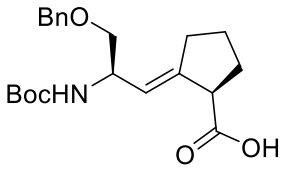


Ester 4.33. Oxalyl chloride (1.3 mL, 15 mmol) was added to a solution of acid **4.43** (1.9 g, 9.9 mmol) in benzene at rt and stirred at reflux for 30 min. The acyl chloride was isolated by distillation, dissolved in THF (4 mL) and added dropwise to a 0 ºC solution of alcohol **4.32** (3.4 g, 9.8 mmol) in THF (4 mL) and pyridine (2.5 mL, 25 mmol) and stirred at rt for 3 h. The reaction mixture was diluted with Et_2O (40 mL), washed with 1 M HCl (2 × 20 mL), $NaHCO_3$ (10 mL),

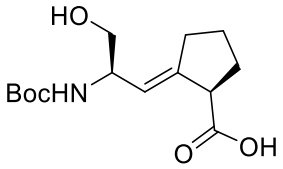
and brine (10 mL), dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 4% EtOAc/hexanes to give **4.33** as a colorless oil (3.6 g, 71%). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.35-7.28 (m, 5H), 5.67 (s, 1H), 5.58 (d, *J* = 8.0 Hz, 1H), 4.83 (d, *J* = 9.4 Hz, 1H), 4.51 (d, *J* = 11.9 Hz, 1H), 4.42 (d, *J* = 11.9 Hz, 1H), 4.16 (s, 2H), 4.04 (m, 1H), 3.55 (dd, *J* = 3.5, 9.4 Hz, 1H), 3.48 (dd, *J* = 3.3, 9.5 Hz, 1H), 2.48 – 2.18 (m, 4H), 1.91 – 175 (m, 2H), 1.40 (s, 9H), 0.90 (s, 9H), 0.07 (s, 6H).



***α*-Hydroxy acid 4.34.** A solution of **e**ster **4.33** (2.9 g, 5.6 mmol) in THF (19 mL) at –100 ºC was added by cannula to a –100 ºC solution of pyridine (5.4 mL, 67 mmol), TMS–Cl (7.8 mL, 62 mmol), and LDA (2.0 M in THF/heptane/ethylbenzene, 11 mL, 22 mmol) in THF (16 mL). The reaction was warmed to rt after 30 min and stirred for an additional 3 h. It was then cooled to 0 °C and quenched with 1 M HCl (70 mL). The product was extracted with Et_2O (3 × 150 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. A solution of TBAF (3.3 g, 12 mmol) in THF (10 mL) at 0 °C was added to the crude mixture in THF (10 mL) at 0 °C and stirred at rt for 1 h. The reaction was then quenched with 1 M HCl (50 mL) and the product was extracted with Et_2O (2 × 100 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. Purification by SiO_2 chromatography (5 × 15 cm, 20% EtOAc/hexanes) gave **4.34** as a colorless oil (1.1 g, 47%, 2 steps). The ^{1}H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.34 – 7.25 (m, 5H), 6.93 (br s, 1H), 5.24 – 4.91 (m, 1H), 4.52 (d, *J* = 12.0 Hz, 1H), 4.47 (d, *J* = 11.9 Hz, 1H), 4.33 (m, 1H), 4.15 (br s, 1H), 3.49 – 3.32 (m, 2H), 2.97 (br s, 1H), 2.53 – 2.41 (m, 1H), 2.30 – 2.06 (m, 1H), 1.99 – 1.71 (m, 3H), 1.61 – 1.45 (m, 1H), 1.41 (s, 9H).

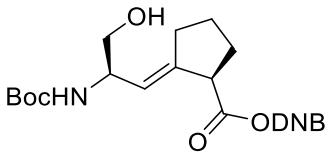


Acid 4.35. A solution of $Pb(OAc)_4$ (1.0 g, 2.3 mmol) in $CHCl_3$ (5 mL) was added to α-hydroxy acid **4.34** (860 mg, 2.1 mmol) in EtOAc (31 mL) at 0 °C and stirred for 10 min. It was then quenched with ethylene glycol (3 mL), diluted with EtOAc (50 mL) and washed with H_2O (4 × 30 mL) and brine (30 mL). The organic layer was dried over Na_2SO_4 and concentrated to give the aldehyde as a yellow oil (765 mg, 100% crude yield). The crude mixture was dissolved in acetone (57 mL) and Jones reagent (2.5 M, 1.7 mL, 4.3 mmol) was added at 0 °C. After stirring for 30 min, the reaction was quenched with *i*-PrOH (5 mL), stirred for 10 min, filtered through Celite, and concentrated. The residue was extracted with EtOAc (4 × 80 mL), and the combined organic layers were washed with H_2O (4 × 30 mL) and brine (30 mL), then dried over Na_2SO_4 and concentrated. The crude product was purified by SiO_2 chromatography (5 × 15 cm, 30% EtOAc/hexanes) to give **4.35** as a pale-yellow oil (620 mg, 77%). The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 10.20 (br s, 1H), 7.34 – 7.19 (m, 5H), 5.52 (d, *J* = 8.4 Hz, 1H), 4.96 (br s, 1H), 4.51 (d, *J* = 12.1 Hz, 1H), 4.44 (d, *J* = 12.1 Hz, 1H), 4.37 (br s, 1H), 3.51 – 3.25 (m, 3H), 2.58 – 2.43 (m, 1H), 2.34 – 2.20 (m, 1H), 2.00 – 1.81 (m, 3H), 1.67 – 1.56 (m, 1H), 1.40 (s, 9H).

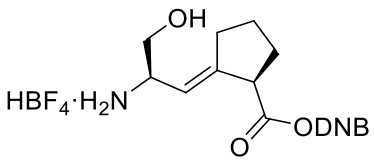


Boc–Ser–Ψ[(*E*)CH=C]-Pro–OH 4.36. NH_3 (ca. 40 mL) was distilled into a suspension of Na (ca. 460 mg) in THF (3 mL) at –33 ºC and stirred until a deep blue color was maintained. Then a solution of acid **4.35** (380 mg, 1.0 mmol) in THF (10 mL) was added slowly and the reaction was stirred at –33 °C for 15 min, after which time it was warmed to rt (caution! NH_3 evolved) until ca.

10 mL NH_3 remained. It was then diluted with NH_4Cl (15 mL), acidified to pH 7 with 1 M HCl, and extracted with $CHCl_3$ (4 × 10 mL). The combined organic layers were dried over Na_2SO_4 and concentrated to give **4.36** as a pale-yellow oil (225 mg 79%). The ^1H NMR matched the literature.[34] ($CDCl_3$, 400 MHz): δ 7.59 (br s, 1H), 5.39 (d, *J* = 6.6 Hz, 1H), 5.22 (br s, 1H), 4.25 (br s, 1H), 3.53 (d, *J* = 5.3 Hz, 2H), 3.29 (td, *J* = 7.5, 3.7 Hz, 1H), 2.51 – 2.40 (m, 1H), 2.35 – 2.20 (m, 1H), 1.99 – 1.79 (m, 3H), 1.67 – 1.57 (m, 1H), 1.38 (s, 9H).



Boc–Ser–Ψ[(*E*)CH=C]-Pro–ODNB 4.37. 2,4,6-Collidine (14 μL, 110 μmol) was added to a solution of **4.36** (30 mg, 110 μmol) in DMF (350 μL) and stirred for 15 min. Then, 3,5-dinitrobenzyl chloride (110 mg, 530 μmol) was added in one portion and the reaction was stirred for 72 h. The reaction was diluted with EtOAc (2 mL) and washed with 5% $NaHCO_3$ (2 × 1 mL), 1 M HCl (2 × 1 mL), H_2O (2 × 1 mL), brine (1 mL), dried over Na_2SO_4, and concentrated. The crude product was purified by SiO_2 chromatography (2 × 15 cm, 3% MeOH/DCM; R_f = 0.23) to give **4.37** as a pale-yellow oil (10 mg, 22% yield). Starting material can be recovered by combining the 5% $NaHCO_3$ washes, acidifying to pH 2 with 1 M HCl, and extracting with EtOAc. ^1H NMR ($CDCl_3$, 400 MHz): δ 9.00 (s, 1H), 8.53 (s, 2H), 5.40 (br d, *J* = 8.8 Hz, 1H), 5.34 – 5.27 (m, 2H), 4.75 (s, 1H), 4.30 (s, 1H), 3.59 –3.44 (m, 3H), 2.60 – 2.50 (m, 1H), 2.43 – 2.30 (m, 1H), 2.05 – 1.86 (m, 3H) 1.77 – 1.66 (m, 1H), 1.43 (s, 9H); ^{13}C {^1H} NMR ($CDCl_3$, 125 MHz): δ 173.7, 156.1, 148.8, 145.2, 140.7, 127.9, 121.9, 118.7, 80.0, 65.7, 64.2, 52.7, 49.5, 30.3, 29.9, 28.5, 25.2. HRMS (ESI^+/TOF) *m/z*: $[M + Na]^+$ Calcd for $C_{21}H_{27}N_3O_9Na^+$ 488.1645; Found 488.1654.



HBF_4·H–Ser–Ψ[(*E*)CH=C]-Pro–ODNB 4.38. A solution of 50% HBF_4·Et_2O (1.5 μL 10 μmol) was added to DNB ester **4.37** (10 mg, 21 μmol) in DCM (200 μL) and stirred at rt for 10 min. The solution was then concentrated, diluted with 1:1 MeOH:H_2O (3 mL) and washed with DCM (3 × 2 mL). The aqueous phase was lyophilized to give **4.38** as a white solid (5 mg, 88% yield). ^{1}H NMR (D_2O, 400 MHz): δ 9.06 (dd, *J* = 2.2, 2.2 Hz, 1H), 8.65 (d, *J* = 2.1 Hz, 2H), 5.42 – 5.36 (m, 3H), 4.07 – 4.00 (m, 1H), 3.64 (dd, *J* = 12.2, 4.2 Hz, 1H), 3.63 – 3.60 (m, 1H), 3.50 (dd, *J* = 11.8, 7.8 Hz, 1H), 2.54 – 2.46 (m, 1H), 2.43 – 2.34 (m, 1H), 2.08 – 1.95 (m, 2H), 1.92 – 1.83 (m, 1H), 1.76 – 1.67 (m, 1H). ^{13}C {^{1}H} NMR (D_2O, 150 MHz): δ 175.7, 150.8, 148.3, 139.8, 128.5, 118.8, 115.3, 64.9, 61.1, 52.3, 49.6, 29.8, 29.2, 24.5. HRMS (ESI$^+$/TOF) *m/z*: $[M + H]^+$ Calcd for $C_{16}H_{20}N_3O_7^+$ 366.1296; Found 366.1300.

4.4 Conclusions

The native dipeptide Boc–Ser–Pro–OH and the locked alkene mimics Boc–Ser–Ψ[(*Z*)CH=C)]–Pro–OH and Boc–Ser–Ψ[(*E*)CH=C)]–Pro–OH were made through literature procedure.[34] Aminoacylation through chemical means was unsuccessfully attempted. While the pdCpA dinucleotide was successfully made through literature procedure,[41] synthesis of the dipeptide-dinucleotide complex proved problematic. Using TBS–adenosine and Boc–Ser(^{t}Bu)–Pro–OH as a model, the Ser hydroxyl could not be unprotected without destruction of the purine. Other side chain protecting groups were attempted, including Ac and Pnt, which were either too labile (Ac) or could not be deprotected on the alkene isosteres (Pnt).

Enzymatic aminoacylation using the dFx flexizyme was found to be a much simpler approach. All dipeptides were converted from the acid to the 3,5-dinitrobenzyl ester conjugate using 3,5-dinitrobenzyl chloride. The Boc–Ser–Ψ[(*E*)CH=C)]–Pro–OH isostere required the use of 2,4,6-collidine as the base to prevent isomerization of the double bond. Deprotection of the Boc group in the presence of the activated DNB ester was most successful with 0.5 equivalent of a 50% $HBF_4 \cdot Et_2O$ solution. Other deprotection methods (HCl, TFA) resulted in ester hydrolysis. The dipeptide HBF_4·H–Ser–Pro–ODNB **4.17** was made in three steps and 82% yield. The cis-locked alkene mimic TFA·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB **4.29** was made in 11 steps and 3% yield. The trans-locked alkene mimic HBF_4·H–Ser–Ψ[(*Z*)CH=C)]–Pro–ODNB **4.38** was made in nine steps and 2% yield.

The dFx ribozyme has shown its ability to aminoacylate dipeptides with the successful coupling of the native HBF_4·H–Ser–Pro–ODBE dipeptide **4.17** to $tRNA^{CUA}$. Aminoacylation of both locked-alkene mimics **4.29** and **4.38** are currently underway.

4.5 References

1. Fischer, G.; Wittmann-Liebold, B.; Lang, K.; Kiefhaber, T.; Schmid, F. X., Cyclophilin and Peptidyl-Prolyl cis-trans Isomerase are Probably Identical Proteins. *Nature* **1989,** *337* (6206), 476-478.

2. Harding, M. W.; Galat, A.; Uehling, D. E.; Schreiber, S. L., A Receptor for the Immunosuppressant FK506 is a cis-trans Peptidyl-Prolyl Isomerase. *Nature* **1989,** *341* (6244), 758-60.

3. Ping Lu, K.; Hanes, S. D.; Hunter, T., A Human Peptidyl–Prolyl Isomerase Essential for Regulation of Mitosis. *Nature* **1996,** *380* (6574), 544-547.

4. Kang, Y. K.; Park, H. S., Internal Rotation About the C–N bond of Amides. *J. Mol. Struct.* **2004,** *676*, 171-176.

5. Fischer, G., Peptidyl-Prolyl cis/trans Isomerases and Their Effectors. *Angew. Chem. Int. Ed.* **1994,** *33* (14), 1415-1436.

6. Göthel, S. F.; Marahiel, M. A., Peptidyl-Prolyl cis-trans Isomerases, a Superfamily of Ubiquitous Folding Catalysts. *Cell. Mol. Life Sci.* **1999,** *55* (3), 423-36.

7. Joseph, J. D.; Yeh, E. S.; Swenson, K. I.; Means, A. R.; Winkler, The Peptidyl-Prolyl Isomerase Pin1. *Prog Cell Cycle Res* **2003,** *5*, 477-87.

8. Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P., Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggests Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997,** *89*, 875-886.

9. Verdecia, M. A.; Bowman, M. E.; Lu, K. P.; Hunter, T.; Noel, J. P., Structural Basis for PhosphoSerine-Proline Recognition by Group IV WW Domains. *Nat. Struct. Biol.* **2000,** *7* (8), 639-43.

10. Yaffe, M. B.; Schutkowski, M.; Shen, M.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J. U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; Cantley, L. C.; Lu, K. P., Sequence-Specific and Phosphorylation-Dependent Proline Isomerization: A Potential Mitotic Regulatory Mechanism. *Science* **1997,** *278* (5345), 1957-60.

11. Wang, X. J.; Xu, B.; Mullins, A. B.; Neiler, F. K.; Etzkorn, F. A., Conformationally Locked Isostere of PhosphoSer-cis-Pro Inhibits Pin1 23-Fold Better than PhosphoSer-trans-Pro Isostere. *J. Am. Chem. Soc.* **2004,** *126*, 15533-15542.

12. El Boustani, M.; De Stefano, L.; Caligiuri, I.; Mouawad, N.; Granchi, C.; Canzonieri, V.; Tuccinardi, T.; Giordano, A.; Rizzolio, F., A Guide to PIN1 Function and Mutations Across Cancers. *Front. Pharmacol.* **2019,** *9*.

13. Bao, L.; Kimzey, A.; Sauter, G.; Sowadski, J. M.; Lu, K. P.; Wang, D. G., Prevalent Overexpression of Prolyl Isomerase Pin1 in Human Cancers. *Am. J. Pathol.* **2004,** *164* (5), 1727-37.

14. Wang, L.; Zhou, Y.; Chen, D.; Lee, T. H., Peptidyl-Prolyl Cis/Trans Isomerase Pin1 and Alzheimer's Disease. *Front. Cell Dev. Biol.* **2020,** *8*.

15. Crenshaw, D. G.; Yang, J.; Means, A. R.; Kornbluth, S., The Mitotic Peptidyl-Prolyl Isomerase, Pin1, Interacts with Cdc25 and Plx1. *EMBO J.* **1998,** *17* (5), 1315-1327.

16. Messenger, M. M.; Saulnier, R. B.; Gilchrist, A. D.; Diamond, P.; Gorbsky, G. J.; Litchfield, D. W., Interactions Between Protein Kinase CK2 and Pin1. Evidence for Phosphorylation-Dependent Interactions. *J. Biol. Chem.* **2002,** *277* (25), 23054-64.

17. Shen, M.; Stukenberg, P. T.; Kirschner, M. W.; Lu, K. P., The Essential Mitotic Peptidyl-Prolyl Isomerase Pin1 Binds and Regulates Mitosis-Specific Phosphoproteins. *Genes Dev.* **1998,** *12* (5), 706-720.

18. Lee, Y. C.; Que, J.; Chen, Y. C.; Lin, J. T.; Liou, Y. C.; Liao, P. C.; Liu, Y. P.; Lee, K. H.; Lin, L. C.; Hsiao, M.; Hung, L. Y.; Huang, C. Y.; Lu, P. J., Pin1 Acts as a Negative Regulator of the G2/M Transition by Interacting with the Aurora-A-Bora Complex. *J. Cell Sci.* **2013,** *126* (Pt 21), 4862-72.

19. Hutterer, A.; Berdnik, D.; Wirtz-Peitz, F.; Zigman, M.; Schleiffer, A.; Knoblich, J. A., Mitotic Activation of the Kinase Aurora-A Requires its Binding Partner Bora. *Dev Cell.* **2006,** *11* (2), 147-57.

20. Barr, A. R.; Gergely, F., Aurora-A: The Maker and Breaker of Spindle Poles. *J. Cell. Sci.* **2007,** *120* (Pt 17), 2987-96.

21. Feine, O.; Hukasova, E.; Bruinsma, W.; Freire, R.; Fainsod, A.; Gannon, J.; Mahbubani, H. M.; Lindqvist, A.; Brandeis, M., Phosphorylation-Mediated Stabilization of Bora in Mitosis Coordinates Plx1/Plk1 and Cdk1 Oscillations. *Cell Cycle* **2014,** *13* (11), 1727-36.

22. Lee, Y. C.; Liao, P. C.; Liou, Y. C.; Hsiao, M.; Huang, C. Y.; Lu, P. J., Glycogen Synthase Kinase 3β Activity is Required for hBora/Aurora A-Mediated Mitotic Entry. *Cell Cycle* **2013,** *12* (6), 953-60.

23. Macurek, L.; Lindqvist, A.; Lim, D.; Lampson, M. A.; Klompmaker, R.; Freire, R.; Clouin, C.; Taylor, S. S.; Yaffe, M. B.; Medema, R. H., Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature* **2008,** *455* (7209), 119-23.

24. Seki, A.; Coppinger, J. A.; Jang, C.-Y.; Yates, J. R. I.; Fang, G., Bora and the Kinase Aurora A Cooperatively Activate the Kinase Plk1 and Control Mitotic Entry. *Science* **2008,** *320*, 1655-1658.

25. Bruinsma, W.; Aprelia, M.; Garcia-Santisteban, I.; Kool, J.; Xu, Y. J.; Medema, R. H., Inhibition of Polo-like kinase 1 during the DNA damage response is mediated through loss of Aurora A recruitment by Bora. *Oncogene* **2017,** *36* (13), 1840-1848.

26. Joukov, V.; De Nicolo, A., Aurora-PLK1 Cascades as Key Signaling Modules in the Regulation of Mitosis. *Sci. Signal* **2018,** *11* (543).

27. Makinwa, Y.; Cartwright, B. M.; Musich, P. R.; Li, Z.; Biswas, H.; Zou, Y., PP2A Regulates Phosphorylation-Dependent Isomerization of Cytoplasmic and Mitochondrial-Associated ATR by Pin1 in DNA Damage Responses. *Front. Cell Dev. Biol.* **2020,** *8*.

28. Eckerdt, F.; Yuan, J.; Saxena, K.; Martin, B.; Kappel, S.; Lindenau, C.; Kramer, A.; Naumann, S.; Daum, S.; Fischer, G.; Dikic, I.; Kaufmann, M.; Strebhardt, K., Polo-Like Kinase 1-Mediated Phosphorylation Stabilizes Pin1 by Inhibiting its Ubiquitination in Human Cells. *J. Biol. Chem.* **2005,** *280* (44), 36575-83.

29. Lin, C. H.; Li, H. Y.; Lee, Y. C.; Calkins, M. J.; Lee, K. H.; Yang, C. N.; Lu, P. J., Landscape of Pin1 in the Cell Cycle. *Exp. Biol. Med. (Maywood)* **2015,** *240* (3), 403-8.

30. Seki, A.; Coppinger, J. A.; Du, H.; Jang, C.-Y.; Yates, J. R., 3rd; Fang, G., Plk1- and beta-TrCP-Dependent Degradation of Bora Controls Mitotic Progression. *J. Cell Biol.* **2008,** *181* (1), 65-78.

31. Joukov, V.; Walter, J. C.; De Nicolo, A., The Cep192-Organized Aurora A-Plk1 Cascade is Essential for Centrosome Cycle and Bipolar Spindle Assembly. *Mol. Cell* **2014,** *55* (4), 578-591.

32. Bayliss, R.; Sardon, T.; Vernos, I.; Conti, E., Structural Basis of Aurora-A Activation by TPX2 at the Mitotic Spindle. *Mol. Cell* **2003,** *12*, 851-862.

33. Silkworth, W. T.; Nardi, I. K.; Scholl, L. M.; Cimini, D., Multipolar Spindle Pole Coalescence Is a Major Source of Kinetochore Mis-Attachment and Chromosome Mis-Segregation in Cancer Cells. *PLOS ONE* **2009,** *4* (8), e6564.

34. Wang, X. J.; Hart, S. A.; Xu, B.; Mason, M. D.; Goodell, J. R.; Etzkorn, F. A., Serine-cis-proline and Serine-trans-proline Isosteres Stereoselective Synthesis of Z- and E-Alkene Mimics by Still-Wittig and Ireland-Claisen Rearrangements. *J. Org. Chem.* **2003,** *68*, 2343-2349.

35. Etzkorn, F. A.; Zhao, S., Stereospecific Phosphorylation by the Central Mitotic Kinase Cdk1-Cyclin B. *ACS Chem. Biol.* **2015,** *10* (4), 952-956.

36. Maini, R. D., L., M.; Paul, R.; Madathil, M., M.; Chowdhury, S., R.; Chen, S.; Hecht, S., M., Ribosome-Mediated Incorporation of Dipeptides and Dipeptide Analogues and Proteins in Vitro. *J. Am. Chem. Soc.* **2015,** *137*, 11206-11209.

37. Roesser, J. R.; Xu, C.; Payne, R. C.; Surratt, C. K.; Hecht, S. M., Preparation of misacylated aminoacyl-tRNAPhe's useful as probes of the ribosomal acceptor site. *Biochemistry* **1989,** *28* (12), 5185-5195.

38. Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H., A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* **2006,** *3* (5), 357-359.

39. Morimoto, J.; Hayashi, Y.; Iwasaki, K.; Suga, H., Flexizymes: Their Evolutionary History and the Origin of Catalytic Function. *Acc. Chem. Res.* **2011,** *44* (12), 1359-1368.

40. Letsinger, R. L.; Finnan, J. L.; Heavner, G. A.; Lunsford, W. B., Nucleotide Chemistry. XX. Phosphite Coupling Procedure for Generating Internucleotide Links. *J. Am. Chem. Soc.* **1975,** *97* (11), 3278-3279.

41. Zhu, X.-F.; Scott, A. I., An Improved Synthesis Of The Dinucleotides pdCpA And pdCpdA. *Nucleosides, Nucleotides Nucleic Acids* **2001,** *20* (3), 197-211.

42. Buszewski, B.; Noga, S., Hydrophilic Interaction Liquid Chromatography (HILIC)--A Powerful Separation Technique. *Anal Bioanal Chem* **2012,** *402* (1), 231-247.

43. El-Faham, A.; Albericio, F., Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **2011,** *111* (11), 6557-6602.

44. Fahmi, N. E.; Dedkova, L.; Wang, B.; Golovine, S.; Hecht, S. M., Site-Specific Incorporation of Glycosylated Serine and Tyrosine Derivatives into Proteins. *J. Am. Chem. Soc.* **2007,** *129* (12), 3586-3597.

45. Fahmi, N. E.; Golovine, S.; Wang, B.; Hecht, S. M., Studies Toward the Site Specific Incorporation of Sugars into Proteins: Synthesis of Glycosylated Aminoacyl–tRNAs. *Carbohydrate Research* **2001,** *330* (2), 149-164.

46. Magda Fouad Fathalla; Khattab, S. N., Spectrophotometric Determination of pKa's of 1-Hydroxybenzotriazole and Oxime Derivative in 95% Acetonitrile-Water. *J. Chem. Soc. Pak.* **2011,** *33*, 324-332.

47. Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H., Ribozyme-Catalyzed tRNA Aminoacylation. *Nat. Struct. Biol.* **2000,** *7* (1), 28-33.

48. Niwa, N.; Yamagishi, Y.; Murakami, H.; Suga, H., A Flexizyme that Selectively Charges Amino Acids Activated by a Water-Friendly Leaving Group. *Bioorg. Med. Chem. Lett.* **2009,** *19* (14), 3892-4.

49. Swain, C. G.; Langsdorf, W. P., Concerted Displacement Reactions. VI. m- and p-Substituent Effects as Evidence for a Unity of Mechanism in Organic Halide Reactions. *J. Am. Chem. Soc.* **1951,** *73* (6), 2813-2819.

50. Cho, J. H.; Coats, S. J.; Schinazi, R. F., Efficient Synthesis of Exo-N-carbamoyl Nucleosides: Application to the Synthesis of Phosphoramidate Prodrugs. *Organic Letters* **2012,** *14* (10), 2488-2491.

51. Caillé, J.; Pantin, M.; Boeda, F.; Pearson-Long, M. S. M.; Bertus, P., Zinc-Mediated Double Addition on Functionalized Nitriles. *Synthesis* **2019,** *51* (06), 1329-1341.

52. Birchall, A., C.; Bush, S., M.; North, M., Copolymerization of Peptide Derived Monomers and Methyl Methacrylate. *Polymer* **2001,** *42*, 375-389.

53. Bischofberger, N.; Waldmann, H.; Saito, T.; Simon, E. S.; Lees, W.; Bednarski, M. D.; Whitesides, G. M., Synthesis of Analogues of 1,3-Dihydroxyacetone Phosphate and Glyceraldehyde 3-Phosphate for Use in Studies of Fructose-1,6-diphosphate Aldolase. *J. Org. Chem.* **1988,** (53), 3457-3465.

Appendix A

Supplementary Information for Chapter 2

A.1 Cartesian (xyz) coordinates from Table S2.1.

(Molecular formula, basis set, Ψ angle)

G(–F)=PP D3a

$C_{12}H_{18}FNO$, MP2/6-311+G(2d,p), –17.6°

```
 7 1.390599 0.145290 -0.211385
 6 0.610866 1.010359 -0.890014
 6 -0.898286 1.095029 -0.673004
 6 -1.609081 -0.047610 0.009002
 6 -2.491069 -2.441400 -0.028551
 6 -1.798919 -1.240249 -0.552414
 9 -1.259646 -1.453680 -1.801107
 8 1.115035 1.782085 -1.729876
 6 2.836874 0.140795 -0.503404
 6 3.371515 -1.007286 0.347502
 6 2.424350 -1.005318 1.546423
 6 1.065993 -0.781844 0.887592
 1 0.334497 -0.357406 1.568725
 1 0.673582 -1.722187 0.488654
 1 2.448615 -1.927633 2.128084
 1 2.657079 -0.166455 2.208843
 1 3.279650 -1.951793 -0.195602
 1 4.418302 -0.864917 0.617190
 1 3.002269 0.010523 -1.573538
 1 3.269581 1.101430 -0.207929
 6 -1.269969 2.310362 0.209279
 6 -1.418860 1.721967 1.614726
 6 -2.124333 0.376716 1.370679
 1 -3.208796 0.528662 1.333991
 1 -1.924064 -0.359222 2.154124
 1 -0.430321 1.561906 2.056303
 1 -1.978961 2.372010 2.289670
 1 -0.535533 3.115256 0.142203
 1 -2.234687 2.702689 -0.126628
 1 -1.280247 1.234288 -1.687739
 1 -1.790133 -3.276202 0.062950
 1 -2.922008 -2.233064 0.950265
 1 -3.290972 -2.747601 -0.707807
```

G(–F)=PP *D*3b

$C_{12}H_{18}FNO$, MP2/6-31+G(d), +38.6°

```
7 1.519821 0.287791 0.019150
6 0.646434 1.224039 -0.418083
6 -0.804991 0.853994 -0.783013
6 -1.504404 -0.162345 0.084311
6 -2.610868 -2.458388 0.405059
6 -1.839913 -1.400564 -0.296089
9 -1.355147 -1.862589 -1.518684
8 1.025712 2.409212 -0.589699
6 2.947510 0.651723 0.141905
6 3.658601 -0.691272 0.027825
6 2.681044 -1.644313 0.711612
6 1.314151 -1.164742 0.219588
1 0.529012 -1.348995 0.952897
1 1.036197 -1.642176 -0.726357
1 2.850982 -2.697762 0.469592
1 2.746191 -1.529185 1.800051
1 3.777685 -0.961974 -1.028085
1 4.648905 -0.677184 0.492770
1 3.209730 1.363847 -0.643381
1 3.130481 1.129600 1.113349
6 -1.716786 2.070341 -0.533244
6 -1.929531 2.032339 0.986788
6 -2.020297 0.525020 1.335559
1 -3.049784 0.225696 1.560583
1 -1.411458 0.274967 2.214747
1 -1.067080 2.487727 1.487095
1 -2.820286 2.584885 1.303186
1 -1.278044 3.005280 -0.888741
1 -2.666384 1.904252 -1.058060
1 -0.793702 0.547342 -1.837759
1 -1.956939 -3.301382 0.657771
1 -3.046577 -2.068760 1.327019
1 -3.412741 -2.836114 -0.238277
```

G(–F)=PP *D*3c

$C_{12}H_{18}FNO$, MP2/6-311+G(2d,p), +161.9°

```
7 1.739571 -0.146375 0.186416
6 0.451188 0.030173 0.541594
6 -0.518238 0.360029 -0.584713
6 -1.952155 0.111161 -0.186037
6 -3.896721 -1.499771 0.178775
6 -2.509440 -1.096897 -0.156416
9 -1.715778 -2.179314 -0.469856
8 0.091024 -0.015635 1.731164
6 2.785105 -0.313410 1.205166
6 4.070982 -0.079624 0.416891
6 3.722395 -0.628449 -0.967668
6 2.289340 -0.140293 -1.178564
1 1.719567 -0.798022 -1.835709
1 2.273995 0.871946 -1.595257
1 4.392595 -0.285126 -1.756206
1 3.739011 -1.721524 -0.949247
```

```
1 4.275234 0.992663 0.352398
1 4.932035 -0.569094 0.872808
1 2.629392 0.397567 2.017100
1 2.744740 -1.325287 1.621808
6 -0.500754 1.885569 -0.852095
6 -1.506972 2.435940 0.163144
6 -2.656914 1.416614 0.120889
1 -3.354196 1.672239 -0.685117
1 -3.225763 1.383692 1.053002
1 -1.056310 2.441688 1.158936
1 -1.832131 3.451476 -0.071466
1 0.495995 2.323089 -0.756602
1 -0.858264 2.069378 -1.870211
1 -0.249880 -0.206986 -1.477503
1 -3.903475 -2.170278 1.042253
1 -4.498895 -0.621490 0.409439
1 -4.353728 -2.028656 -0.661742
```

PG(–F)=P *A*4a

C9H14FNO, MP2/6-31+G(d), –175.5°

```
6 4.331404 -0.351251 -0.193094
6 2.987032 0.115698 0.307502
8 2.628780 -0.042130 1.491658
7 2.179938 0.708945 -0.609215
6 0.815680 1.072076 -0.289871
6 -0.141025 -0.100687 -0.303626
9 0.470600 -1.323292 -0.521361
6 -1.464128 -0.060662 -0.132978
6 -2.400833 -1.250015 -0.131586
6 -3.693946 -0.674127 0.467362
6 -3.704724 0.772594 -0.037472
6 -2.240893 1.217586 0.119334
1 -2.064160 1.580910 1.142127
1 -1.964475 2.029558 -0.564153
1 -3.988486 0.791444 -1.097703
1 -4.402493 1.417341 0.507375
1 -4.584862 -1.240586 0.175862
1 -3.634463 -0.683124 1.563482
1 -2.581637 -1.595306 -1.159530
1 -1.998998 -2.100085 0.430388
1 0.810098 1.541959 0.699501
1 0.488725 1.826013 -1.012978
1 2.468860 0.720232 -1.581545
1 4.283507 -1.429807 -0.374668
1 4.631476 0.146677 -1.118029
1 5.080308 -0.168439 0.579930
```

PG(–F)=P *A*4b

C9H14FNO, MP2/6-311+G(2d,p), –59.5°

6 -3.219269 -1.618969 -0.585983
6 -2.466697 -0.345434 -0.291081
8 -2.428349 0.598596 -1.089805
7 -1.820895 -0.290135 0.894883
6 -1.028619 0.870518 1.267812
6 0.148487 1.106359 0.366176
9 0.132675 2.349015 -0.216202
6 1.169501 0.291386 0.114633
6 2.340719 0.604575 -0.789251
6 3.389709 -0.434967 -0.367792
6 2.542643 -1.666188 -0.033865
6 1.344320 -1.081651 0.733894
1 1.588780 -0.984734 1.798732
1 0.449779 -1.702449 0.656632
1 2.192032 -2.128124 -0.962496
1 3.081945 -2.423708 0.537875
1 4.138885 -0.621218 -1.139714
1 3.907192 -0.089854 0.533353
1 2.057396 0.444229 -1.836504
1 2.681259 1.637579 -0.694477
1 -1.653812 1.766105 1.249256
1 -0.693119 0.712642 2.294959
1 -1.867750 -1.078932 1.522612
1 -2.761303 -2.099393 -1.452349
1 -3.219082 -2.317939 0.250214
1 -4.247980 -1.366935 -0.845265

PG(–F)=P *A*4c

$C_9H_{14}FNO$, MP2/6-311+G(2d,p), +116.9°

6 3.805304 -0.691287 -0.928273
6 2.522750 -0.640780 -0.134442
8 1.770059 -1.617544 -0.042900
7 2.247314 0.534924 0.478586
6 1.011151 0.736904 1.221507
6 -0.141923 1.038275 0.329456
9 0.043217 2.215276 -0.360366
6 -1.259884 0.344917 0.131289
6 -2.371379 0.744976 -0.815048
6 -3.206915 -0.538270 -0.937438
6 -3.087417 -1.163920 0.455138
6 -1.602595 -0.970480 0.803179
1 -1.001820 -1.774075 0.364405
1 -1.415780 -0.967001 1.878981
1 -3.708681 -0.602700 1.160833
1 -3.396227 -2.210399 0.489771
1 -4.237890 -0.347052 -1.241095
1 -2.751883 -1.204253 -1.677885
1 -2.974065 1.542046 -0.364002
1 -1.995474 1.118200 -1.769922
1 0.794613 -0.163449 1.796060
1 1.174487 1.564610 1.915618
1 2.864463 1.320432 0.331905
1 3.559410 -0.893105 -1.972002

```
1 4.381314 0.231934 -0.869030
1 4.412360 -1.518799 -0.559459
```

G(–Cl)=PP *D*5a

C12H18ClNO, MP2/6-311+G(2d,p), –19.0°

```
6 0.783184 1.024632 0.883287
6 -0.697623 0.728855 1.114110
7 -1.457341 0.102695 0.194672
6 -2.874997 -0.144884 0.519224
6 -3.378033 -0.962165 -0.667425
6 -2.511902 -0.452633 -1.818695
6 -1.134992 -0.352611 -1.168561
1 -0.478539 0.344817 -1.680928
1 -0.653255 -1.335086 -1.136739
1 -2.505448 -1.110301 -2.688768
1 -2.846699 0.540636 -2.131779
1 -3.183804 -2.025120 -0.500422
1 -4.447429 -0.829110 -0.833120
1 -2.954478 -0.662980 1.475900
1 -3.399070 0.811617 0.607004
8 -1.196574 1.099091 2.195042
6 1.524592 0.253489 -0.178924
6 1.820433 -1.048803 -0.115160
6 2.573511 -1.854334 -1.116722
1 1.936190 -2.635293 -1.541983
1 2.933959 -1.216087 -1.924009
1 3.430962 -2.345640 -0.649538
17 1.241510 -1.984188 1.265782
6 1.934907 1.181216 -1.309439
6 1.117768 2.456544 -1.047126
6 1.005337 2.500306 0.478482
1 0.214501 3.157152 0.845635
1 1.953695 2.840668 0.905231
1 0.121190 2.370821 -1.491149
1 1.588454 3.348762 -1.464564
1 3.005756 1.404194 -1.237044
1 1.757047 0.752607 -2.299176
1 1.222173 0.830273 1.865815
```

G(–Cl)=PP *D*5b

C12H18ClNO, MP2/6-31+G(d), +44.1°

```
6 -0.577382 0.951461 -0.781133
6 0.923768 1.173190 -0.501812
7 1.688587 0.176804 -0.000435
6 3.151797 0.373877 0.077708
6 3.695197 -1.050418 0.101839
6 2.609257 -1.811182 0.859886
6 1.314597 -1.221716 0.300620
```

1 0.496628 -1.261648 1.019917
1 1.006741 -1.735670 -0.616813
1 2.653623 -2.895702 0.721750
1 2.681630 -1.599409 1.933359
1 3.781913 -1.434863 -0.921427
1 4.679688 -1.110116 0.575351
1 3.486408 0.964117 -0.778339
1 3.404803 0.924694 0.993058
8 1.439594 2.282018 -0.786135
6 -1.356708 0.221791 0.284760
6 -1.982826 -0.963619 0.164826
6 -2.857681 -1.633127 1.177098
1 -2.410590 -2.580811 1.499216
1 -2.988141 -1.000483 2.056811
1 -3.840915 -1.859629 0.750437
17 -1.776448 -1.933580 -1.290298
6 -1.583439 1.193427 1.431394
6 -1.405446 2.588979 0.776071
6 -1.300368 2.309215 -0.731808
1 -0.768698 3.089919 -1.280457
1 -2.297209 2.184360 -1.173709
1 -0.482872 3.063560 1.126551
1 -2.233566 3.262567 1.019560
1 -2.578906 1.078926 1.870357
1 -0.854341 1.030085 2.237040
1 -0.658041 0.465921 -1.762173

G(–Cl)=PP *D*5c

$C_{12}H_{18}ClNO$, MP2/6-311+G(2d,p), +155.3°

6 -0.421362 0.581864 -0.600726
6 0.541987 0.316320 0.546163
7 1.805534 -0.003558 0.206948
6 2.855204 -0.112015 1.228903
6 4.139124 -0.048495 0.406332
6 3.738633 -0.741319 -0.897350
6 2.329164 -0.207137 -1.153386
1 1.713578 -0.915493 -1.709264
1 2.356946 0.741257 -1.699789
1 4.412853 -0.533399 -1.728591
1 3.698617 -1.823338 -0.745562
1 4.400286 0.994969 0.209704
1 4.979485 -0.526720 0.910298
1 2.752651 0.697060 1.952614
1 2.763592 -1.064366 1.761876
8 0.197222 0.451658 1.733695
6 -1.851348 0.337400 -0.186668
6 -2.397758 -0.870290 -0.022106
6 -3.804627 -1.185256 0.355635
1 -3.839676 -1.738432 1.298024
1 -4.380813 -0.265998 0.465390
1 -4.278669 -1.805208 -0.409861
17 -1.382119 -2.299451 -0.254819

```
6 -2.577272 1.656732 -0.007774
6 -1.442513 2.692107 -0.001573
6 -0.405333 2.089003 -0.953233
1 0.587742 2.535319 -0.856917
1 -0.734426 2.211605 -1.989874
1 -1.019606 2.767862 1.002958
1 -1.774346 3.685152 -0.311509
1 -3.244165 1.835672 -0.859694
1 -3.184378 1.685639 0.899756
1 -0.153559 -0.032758 -1.461141
```

PG(–Cl)=P *A*6a

$C_9H_{14}ClNO$, MP2/6-31+G(d), –171.1°

```
6 -4.323827 0.055408 -0.223952
6 -2.990929 -0.444549 0.274481
8 -2.728856 -0.534795 1.490160
7 -2.082814 -0.771586 -0.682680
6 -0.714298 -1.099704 -0.353166
6 0.214988 0.096069 -0.199870
17 -0.525173 1.681170 -0.206562
6 1.545609 -0.024565 -0.068221
6 2.545387 1.096841 0.110849
6 3.833703 0.371726 0.528206
6 3.746145 -0.964083 -0.215704
6 2.269162 -1.358916 -0.062250
1 2.111940 -1.874186 0.897089
1 1.921886 -2.036836 -0.851133
1 3.982298 -0.813255 -1.277012
1 4.429990 -1.725195 0.174848
1 4.737571 0.942036 0.289095
1 3.827459 0.192539 1.611153
1 2.695342 1.624980 -0.842471
1 2.208915 1.842007 0.840120
1 -0.728895 -1.672908 0.579708
1 -0.327128 -1.753106 -1.141487
1 -2.310056 -0.577893 -1.652667
1 -4.307171 1.150346 -0.233199
1 -4.553704 -0.298225 -1.231899
1 -5.105589 -0.270098 0.464640
```

PG(–Cl)=P *A*6b

$C_9H_{14}ClNO$, MP2/6-311+G(2d,p), –60.7°

```
6 -3.297593 -1.766485 -0.677750
6 -2.528930 -0.531805 -0.275766
8 -2.561818 0.511183 -0.939451
7 -1.788260 -0.629457 0.851146
6 -0.999676 0.488353 1.342469
6 0.168705 0.845454 0.455008
17 0.177093 2.503152 -0.138128
6 1.184931 0.031726 0.151707
```

```
6 2.386713 0.374436 -0.697170
6 3.385926 -0.746623 -0.379138
6 2.484299 -1.963396 -0.158234
6 1.313181 -1.393818 0.657721
1 1.565844 -1.381729 1.725871
1 0.395434 -1.971826 0.542485
1 2.115118 -2.324262 -1.123765
1 2.985755 -2.792865 0.343971
1 4.126329 -0.892572 -1.167863
1 3.918435 -0.511514 0.548214
1 2.108050 0.342516 -1.758223
1 2.769716 1.376159 -0.491539
1 -1.643760 1.361423 1.458691
1 -0.630662 0.206659 2.332114
1 -1.807602 -1.489803 1.378242
1 -2.947846 -2.088450 -1.659912
1 -3.184455 -2.588654 0.028809
1 -4.353907 -1.508989 -0.764415
```

PG(–Cl)=P *A*6c

$C_9H_{14}ClNO$, MP2/6-311+G(2d,p), +117.9°

```
6 3.753341 -0.901553 -0.947439
6 2.449235 -0.909374 -0.188004
8 1.619664 -1.817092 -0.317016
7 2.239898 0.135130 0.648676
6 0.982984 0.294910 1.361686
6 -0.124411 0.777560 0.476972
17 0.182486 2.366415 -0.227783
6 -1.257487 0.129677 0.193159
6 -2.357145 0.622496 -0.719293
6 -3.221663 -0.624382 -0.953838
6 -3.116246 -1.373812 0.376608
6 -1.629669 -1.237357 0.738943
1 -1.037300 -2.001871 0.224011
1 -1.442392 -1.338427 1.809791
1 -3.728256 -0.867707 1.130404
1 -3.442423 -2.413976 0.317912
1 -4.247742 -0.382899 -1.237347
1 -2.782200 -1.231637 -1.751890
1 -2.940805 1.396021 -0.204600
1 -1.967058 1.067340 -1.637343
1 0.700392 -0.667877 1.788506
1 1.155334 1.001210 2.176900
1 2.913475 0.887030 0.669611
1 3.533203 -0.810395 -2.012532
1 4.415875 -0.088801 -0.650240
1 4.257408 -1.855220 -0.788246
```

G(–H)=PP *D*7a

$C_{12}H_{19}NO$, MP2/6-311+G(2d,p), –56.6°

```
6 -0.943183 -0.991164 0.689887
6 0.499204 -0.789216 1.106597
7 1.341575 -0.157359 0.267307
6 2.755825 0.007512 0.638142
6 3.332556 0.825854 -0.516945
6 2.453872 0.418720 -1.701637
6 1.065166 0.378162 -1.071271
1 0.364284 -0.252369 -1.612736
1 0.640386 1.384198 -0.996854
1 2.505551 1.112268 -2.541737
1 2.731605 -0.579370 -2.052541
1 3.208495 1.893504 -0.316800
1 4.393153 0.626210 -0.671680
1 2.839155 0.501967 1.607292
1 3.227644 -0.976750 0.718134
8 0.900416 -1.214192 2.209635
6 -1.726491 0.257295 0.326044
6 -1.449113 1.490108 0.772169
6 -2.214783 2.740777 0.459096
1 -1.561462 3.491541 0.004023
1 -3.047908 2.552823 -0.218783
1 -2.618282 3.184760 1.374295
1 -0.581580 1.608694 1.422628
6 -2.821216 -0.177677 -0.614893
6 -2.140341 -1.260114 -1.465762
6 -1.166246 -1.972943 -0.501469
1 -0.227141 -2.242405 -0.990580
1 -1.604289 -2.899441 -0.125336
1 -1.591618 -0.782349 -2.281742
1 -2.853645 -1.953336 -1.915883
1 -3.639076 -0.625323 -0.035319
1 -3.239387 0.634761 -1.211146
1 -1.411737 -1.428212 1.576789
```

G(–H)=PP *D*7b

C12H19NO, MP2/6-31+G(d), +38.2°

```
6 0.936245 -0.750479 -0.845122
6 -0.470302 -1.176791 -0.381094
7 -1.406017 -0.260451 -0.039974
6 -2.791824 -0.718191 0.196035
6 -3.616972 0.547285 -0.006559
6 -2.682871 1.640540 0.507355
6 -1.315537 1.216484 -0.029559
1 -0.497045 1.554107 0.605091
1 -1.148435 1.594247 -1.045468
1 -2.955742 2.646070 0.173111
1 -2.668563 1.636755 1.603838
1 -3.821060 0.694135 -1.073897
1 -4.573439 0.507485 0.523161
1 -3.032737 -1.524147 -0.500773
1 -2.887573 -1.109945 1.217232
8 -0.763204 -2.398993 -0.379726
```

```
6 1.582077 0.431994 -0.159132
6 1.837452 1.609347 -0.760751
6 2.556424 2.774026 -0.147486
1 1.856656 3.593621 0.062940
1 3.045317 2.504622 0.792474
1 3.314490 3.171622 -0.832046
1 1.439848 1.765667 -1.765754
6 2.103114 -0.045196 1.183752
6 2.175903 -1.587701 1.034599
6 1.942604 -1.850673 -0.462516
1 1.575814 -2.857720 -0.673686
1 2.867279 -1.685997 -1.030429
1 1.381330 -2.067794 1.617349
1 3.129240 -1.994295 1.388429
1 3.083557 0.388536 1.409116
1 1.430204 0.248364 2.002020
1 0.870723 -0.594102 -1.932214
```

G(–H)=PP *D*7c

$C_{12}H_{19}NO$, MP2/6-311+G(2d,p), +96.0°

```
6 0.830531 -0.658363 -0.645386
6 -0.284128 -0.899201 0.349454
7 -1.494691 -0.393032 0.044175
6 -2.625167 -0.538584 0.972421
6 -3.818015 -0.060213 0.146857
6 -3.194708 0.985706 -0.779860
6 -1.873760 0.332842 -1.179440
1 -1.111734 1.059693 -1.460791
1 -2.019032 -0.367785 -2.008568
1 -3.812874 1.227313 -1.645073
1 -2.996164 1.906817 -0.224451
1 -4.216513 -0.888866 -0.444679
1 -4.620098 0.334865 0.770733
1 -2.708183 -1.574087 1.304562
1 -2.461767 0.091201 1.853066
8 -0.080981 -1.502867 1.418489
6 1.652700 0.567566 -0.229574
6 1.127111 1.713754 0.228192
6 1.893496 2.931866 0.652019
1 1.666205 3.192126 1.690210
1 2.970850 2.793194 0.559505
1 1.610966 3.793900 0.040110
1 0.042681 1.783710 0.310985
6 3.113182 0.195586 -0.317227
6 3.091710 -1.292470 0.049992
6 1.855424 -1.815594 -0.690441
1 1.452690 -2.730996 -0.255364
1 2.114181 -2.018428 -1.733505
1 2.950911 -1.395749 1.129807
1 4.007468 -1.818213 -0.229598
1 3.471426 0.312530 -1.348387
1 3.751331 0.796223 0.332653
```

```
1 0.408086 -0.500653 -1.640400
```

G(–H)=PP *D*7d

$C_{12}H_{19}NO$, MP2/6-311+G(2d,p), +163.5°

```
6 -0.653928 0.149704 -0.622553
6 0.331191 -0.153386 0.495695
7 1.633421 -0.216066 0.150488
6 2.674892 -0.380092 1.174048
6 3.958199 -0.056245 0.413453
6 3.647089 -0.548641 -1.001000
6 2.197853 -0.107938 -1.203642
1 1.662408 -0.746775 -1.906290
1 2.147160 0.925698 -1.560780
1 4.310968 -0.138949 -1.762789
1 3.707854 -1.639736 -1.037394
1 4.119902 1.025136 0.404072
1 4.833519 -0.532828 0.855645
1 2.480352 0.287194 2.014311
1 2.674100 -1.409788 1.546723
8 -0.033501 -0.286434 1.678099
6 -2.073062 -0.194004 -0.227713
6 -2.636857 -1.390931 -0.438355
6 -4.052523 -1.748870 -0.095986
1 -4.084135 -2.548712 0.650431
1 -4.598390 -0.892207 0.302347
1 -4.586743 -2.115785 -0.977559
1 -2.025486 -2.174332 -0.886137
6 -2.755675 1.044606 0.327044
6 -1.671914 2.137147 0.285944
6 -0.725968 1.680335 -0.831383
1 0.250761 2.169274 -0.797408
1 -1.174588 1.879957 -1.809952
1 -1.130880 2.159033 1.234930
1 -2.080813 3.133695 0.107410
1 -3.602589 1.320452 -0.311077
1 -3.148075 0.886639 1.335174
1 -0.357045 -0.374146 -1.533174
```

PG(–H)=P *A*8a

$C_9H_{15}NO$, MP2/6-31+G(d), –179.6°

```
6 4.394675 -0.407655 -0.255195
6 3.002347 -0.118316 0.252246
8 2.582224 -0.605091 1.324181
7 2.233298 0.687418 -0.517388
6 0.845973 0.982959 -0.185691
6 -0.086025 -0.156097 -0.524454
1 0.369896 -1.043528 -0.965814
6 -1.412476 -0.103611 -0.312180
6 -2.408536 -1.204852 -0.616049
```

```
6 -3.591724 -0.871832 0.302945
6 -3.629279 0.660888 0.265134
6 -2.145070 1.060368 0.333576
1 -1.825324 1.171295 1.380413
1 -1.935487 2.016649 -0.162957
1 -4.063470 0.991267 -0.687421
1 -4.225754 1.100315 1.071778
1 -4.530564 -1.334832 -0.020149
1 -3.381404 -1.215106 1.324624
1 -2.735901 -1.143387 -1.664386
1 -1.989665 -2.205186 -0.459857
1 0.798850 1.220427 0.883967
1 0.575380 1.894106 -0.730884
1 2.602780 1.019036 -1.401752
1 4.444874 -1.454499 -0.569414
1 4.681753 0.228209 -1.095739
1 5.103837 -0.267074 0.563865
```

PG(–H)=P *A*8b

$C_9H_{15}NO$, MP2/6-311+G(2d,p), –121.3°

```
6 4.114491 -0.164148 1.097142
6 3.081962 0.067510 0.019494
8 3.236071 0.953058 -0.849761
7 1.990124 -0.733023 0.049571
6 0.856781 -0.556863 -0.859274
6 -0.115032 0.463317 -0.339429
1 0.291870 1.461852 -0.169304
6 -1.414057 0.224030 -0.079854
6 -2.393271 1.259002 0.446213
6 -3.767528 0.652866 0.133846
6 -3.533064 -0.844788 0.358793
6 -2.156795 -1.089686 -0.281387
1 -2.275977 -1.297525 -1.353935
1 -1.629961 -1.946125 0.155221
1 -3.488870 -1.050378 1.436350
1 -4.317937 -1.477820 -0.068865
1 -4.569671 1.064611 0.755898
1 -4.029819 0.834860 -0.916841
1 -2.280066 1.361950 1.535326
1 -2.232639 2.248926 0.005151
1 1.268084 -0.253370 -1.827278
1 0.381079 -1.532703 -0.985702
1 1.876416 -1.370458 0.831174
1 4.134101 0.706989 1.758653
1 3.915202 -1.057860 1.692832
1 5.097827 -0.253691 0.629151
```

PG(–H)=P *A*8c

$C_9H_{15}NO$, MP2/6-311+G(2d,p), +120.0°

```
6 -3.773751 -0.794208 0.745808
6 -2.488145 -0.468785 0.022878
8 -1.711394 -1.356110 -0.354762
7 -2.239749 0.842669 -0.189962
6 -0.993981 1.305335 -0.796406
6 0.135475 1.343200 0.191400
1 -0.006534 1.982442 1.062339
6 1.271609 0.643645 0.069895
6 2.399691 0.638196 1.078422
6 3.058700 -0.728011 0.852099
6 3.006307 -0.869572 -0.673561
6 1.610107 -0.335532 -1.043772
1 0.868920 -1.141583 -1.045333
1 1.589864 0.127825 -2.033489
1 3.778105 -0.233515 -1.118736
1 3.174315 -1.891063 -1.020201
1 4.070493 -0.792817 1.257725
1 2.451894 -1.511500 1.318750
1 3.121577 1.426391 0.829741
1 2.052595 0.810392 2.099689
1 -0.755842 0.640077 -1.627263
1 -1.189745 2.300548 -1.203525
1 -2.884901 1.523956 0.183631
1 -3.525979 -1.263824 1.699375
1 -4.394287 0.082349 0.932019
1 -4.337155 -1.515091 0.152536
```

A.2 Cartesian (xyz) coordinates from Table S2.2.

(Molecular formula, basis set, Ψ angle)

P(–F)=PG *D*10a

$C_9H_{14}FNO$, MP2/6-311+G(2d,p), +1.1°

```
6 -3.094332 -1.066367 0.459247
1 -3.006767 -2.039416 0.954538
1 -3.908725 -1.131088 -0.270372
1 -3.347983 -0.311326 1.206162
6 -1.823437 -0.711486 -0.221682
9 -1.487455 -1.645820 -1.197802
6 -0.990270 0.320594 -0.036861
6 0.246663 0.578114 -0.867104
1 0.097326 0.312209 -1.920860
6 1.500176 -0.173071 -0.415343
8 2.587892 -0.000748 -1.009330
7 1.385693 -0.995618 0.645394
6 2.514703 -1.764871 1.149422
1 2.901106 -2.433870 0.375826
1 3.318658 -1.096335 1.467433
1 2.173283 -2.352753 2.001368
1 0.466703 -1.129473 1.054349
6 0.452736 2.096699 -0.682211
```

```
1 1.475880 2.408142 -0.913920
1 -0.229178 2.632589 -1.354690
6 0.043142 2.330935 0.775142
1 -0.143811 3.385474 1.003107
1 0.838779 1.982868 1.446025
6 -1.214156 1.459431 0.946042
1 -1.350040 1.112402 1.976724
1 -2.112477 2.028404 0.671497
```

P(–F)=PG *D*10b

$C_9H_{14}FNO$, MP2/6-31+G(d), +44.5°

```
6 2.983204 -1.294368 -0.492364
1 2.818587 -2.192082 -1.098219
1 3.737705 -1.531395 0.265426
1 3.367394 -0.500415 -1.136117
6 1.720834 -0.862783 0.158420
9 1.230151 -1.850953 1.021026
6 1.008980 0.263073 0.036123
6 -0.240929 0.610471 0.822085
1 -0.190368 0.214132 1.845954
6 -1.534575 0.055533 0.202459
8 -2.581992 0.733190 0.136971
7 -1.486515 -1.222869 -0.224439
6 -2.668132 -1.906249 -0.731699
1 -3.449466 -1.952241 0.032352
1 -3.064521 -1.382849 -1.605011
1 -2.379476 -2.917672 -1.017716
1 -0.631572 -1.750087 -0.089362
6 -0.225288 2.144817 0.790326
1 -1.206576 2.572702 1.010336
1 0.489431 2.508420 1.539847
6 0.284263 2.460342 -0.618818
1 0.624666 3.494985 -0.731819
1 -0.519348 2.292964 -1.346560
6 1.417780 1.441794 -0.834290
1 1.542812 1.164165 -1.886982
1 2.374092 1.856733 -0.489512
```

P(–F)=PG *D*10c

$C_9H_{14}FNO$, MP2/6-311+G(2d,p), +64.2°

```
6 2.808097 -1.548822 -0.572897
1 2.557840 -2.396635 -1.215907
1 3.538002 -1.889199 0.166185
1 3.261427 -0.766866 -1.181096
6 1.593701 -1.032993 0.102467
9 0.998988 -1.996389 0.898004
6 1.031912 0.173599 0.049568
6 -0.191367 0.609166 0.834073
1 -0.201599 0.156473 1.830594
```

```
6 -1.496091 0.201645 0.136413
8 -2.320264 1.023970 -0.282446
7 -1.692728 -1.126676 0.026386
6 -2.876975 -1.669346 -0.616993
1 -3.780950 -1.353453 -0.093906
1 -2.942910 -1.335085 -1.653867
1 -2.814472 -2.755573 -0.595704
1 -0.978358 -1.751697 0.371914
6 -0.030870 2.132130 0.876790
1 -0.972422 2.643251 1.078019
1 0.684860 2.388204 1.664552
6 0.562654 2.466032 -0.495244
1 1.007094 3.462003 -0.540061
1 -0.224855 2.408477 -1.250409
6 1.595266 1.349937 -0.728407
1 1.741631 1.125275 -1.787476
1 2.567074 1.639204 -0.314630
```

P(–F)=PG *D*10d

$C_9H_{14}FNO$, MP2/6-311+G(2d,p), +142.5°

```
6 -2.915887 -1.649656 0.392489
1 -2.748645 -2.350236 1.214837
1 -3.448606 -2.181668 -0.400037
1 -3.541879 -0.831385 0.746425
6 -1.623799 -1.127904 -0.114661
9 -0.801105 -2.134927 -0.570318
6 -1.166714 0.120462 -0.183973
6 0.191368 0.480233 -0.746104
1 0.339466 0.052198 -1.741567
6 1.316611 0.018756 0.164524
8 1.254329 0.102903 1.399034
7 2.417456 -0.441894 -0.463937
6 3.599068 -0.856953 0.273065
1 4.003200 -0.026150 0.853437
1 3.359358 -1.673082 0.956404
1 4.349674 -1.194925 -0.438468
1 2.411840 -0.507785 -1.470827
6 0.147749 2.026939 -0.800954
1 1.140768 2.477279 -0.730134
1 -0.293302 2.328009 -1.756085
6 -0.786685 2.417072 0.347816
1 -1.164532 3.437457 0.256138
1 -0.257097 2.327906 1.298871
6 -1.900412 1.361700 0.275146
1 -2.418658 1.223982 1.226480
1 -2.646263 1.658872 -0.471412
```

GP(–F)=P *A*11a

$C_{12}H_{18}FNO$, MP2/6-311+G(2d,p), –12.5º

6 -2.328699 -2.443461 -0.224276
6 -1.669503 -1.235798 -0.843035
8 -1.232635 -1.238133 -2.004973
7 -1.563306 -0.147525 -0.048718
6 -1.000540 1.100443 -0.550841
6 0.493899 1.082339 -0.606813
9 0.932433 2.172250 -1.328831
6 1.413289 0.259012 -0.105024
6 2.902418 0.461167 -0.317061
6 3.508521 -0.883469 0.107618
6 2.589427 -1.329943 1.245950
6 1.184613 -0.995336 0.717589
1 0.457504 -0.868725 1.520171
1 0.832790 -1.807885 0.071066
1 2.698717 -2.383825 1.509548
1 2.798116 -0.731895 2.139399
1 3.439750 -1.597165 -0.720004
1 4.557034 -0.803024 0.400887
1 3.143313 0.746216 -1.343213
1 3.266798 1.259601 0.340141
6 -1.531876 2.137506 0.450383
6 -1.574987 1.342668 1.756731
6 -2.112617 -0.015164 1.308626
1 -3.206937 -0.014699 1.276869
1 -1.783777 -0.837206 1.947148
1 -0.564166 1.231035 2.157986
1 -2.202246 1.800637 2.521819
1 -0.900335 3.025216 0.496482
1 -2.540872 2.439386 0.156411
1 -1.362708 1.286367 -1.566597
1 -1.722209 -2.825384 0.601261
1 -3.314246 -2.194114 0.174840
1 -2.427872 -3.216819 -0.983010

GP(–F)=P *A*11b

$C_{12}H_{18}FNO$, MP2/6-31+G(d), +67.4º

6 -2.658677 2.346230 -0.759892
6 -1.761992 1.581895 0.187639
8 -1.276891 2.130967 1.202132
7 -1.502231 0.283684 -0.134284
6 -0.803138 -0.606994 0.811171
6 0.546542 -1.034857 0.281988
9 0.632475 -2.398515 0.001373
6 1.647453 -0.304492 0.071251
6 2.960905 -0.854050 -0.447841
6 3.955333 0.281478 -0.162602
6 3.108412 1.543616 -0.351058
6 1.772256 1.185159 0.321342
1 0.925197 1.755566 -0.068232
1 1.824553 1.388094 1.400932
1 3.566390 2.442923 0.074794
1 2.948965 1.724126 -1.422351

1 4.304447 0.220789 0.876533
1 4.835292 0.247398 -0.813952
1 3.231864 -1.805098 0.023123
1 2.894077 -1.035652 -1.530393
6 -1.803899 -1.758423 0.935506
6 -2.299099 -1.934934 -0.505079
6 -2.332430 -0.504606 -1.075915
1 -3.348657 -0.103705 -1.090101
1 -1.924312 -0.451630 -2.090777
1 -1.600016 -2.556980 -1.069872
1 -3.285385 -2.405457 -0.555845
1 -1.371529 -2.667550 1.359641
1 -2.620365 -1.423589 1.585970
1 -0.666541 -0.051499 1.744314
1 -2.460361 2.102091 -1.806593
1 -3.710442 2.120929 -0.549921
1 -2.494698 3.412229 -0.595763

GP(–F)=P *A*11c

$C_{12}H_{18}FNO$, MP2/6-311+G(2d,p), +122.6º

6 -2.964357 2.240399 -0.666800
6 -1.737601 1.586877 -0.077119
8 -0.721457 2.234196 0.224701
7 -1.803223 0.248095 0.098459
6 -0.693775 -0.483115 0.729249
6 0.448213 -0.667939 -0.207673
9 0.093826 -1.353158 -1.348221
6 1.715451 -0.286736 -0.063342
6 2.818123 -0.548123 -1.068365
6 4.089988 -0.115256 -0.321836
6 3.600375 1.029021 0.570839
6 2.259005 0.504617 1.107352
1 1.578834 1.296686 1.426578
1 2.440069 -0.156200 1.963289
1 4.301931 1.291789 1.365100
1 3.427287 1.921331 -0.040088
1 4.447508 -0.938449 0.305337
1 4.899707 0.173836 -0.994580
1 2.841664 -1.586802 -1.405808
1 2.668062 0.079974 -1.954319
6 -1.365549 -1.798587 1.159964
6 -2.489153 -1.998279 0.138815
6 -3.002607 -0.579143 -0.095653
1 -3.767673 -0.311238 0.640783
1 -3.415825 -0.439493 -1.095478
1 -2.090296 -2.408810 -0.789149
1 -3.277697 -2.658546 0.501346
1 -0.656366 -2.626645 1.193434
1 -1.788376 -1.664731 2.159153
1 -0.324423 0.084513 1.585582
1 -3.147633 1.869649 -1.678042
1 -3.850931 2.022942 -0.067053

```
1 -2.804735 3.315975 -0.701514
```

P(–Cl)=PG *D*12a

C9H14ClNO, MP2/6-311+G(2d,p), –5.0°

```
6 -2.997479 0.087606 0.911459
1 -3.104762 -0.726933 1.632516
1 -3.889323 0.089529 0.279357
1 -2.949145 1.033394 1.451525
6 -1.773704 -0.100109 0.081343
17 -1.770001 -1.597507 -0.853704
6 -0.728302 0.731422 -0.014008
6 0.477344 0.525579 -0.897941
1 0.201937 0.155839 -1.889199
6 1.506679 -0.467900 -0.361445
8 2.507356 -0.761797 -1.031346
7 1.290445 -0.969142 0.866765
6 2.193068 -1.938290 1.463905
1 2.221273 -2.858619 0.877151
1 3.202810 -1.530378 1.518531
1 1.842919 -2.164017 2.469149
1 0.411872 -0.752626 1.318209
6 1.099935 1.935460 -0.964390
1 2.154989 1.909487 -1.241378
1 0.560997 2.524830 -1.712588
6 0.832510 2.501759 0.432117
1 0.956587 3.584721 0.487363
1 1.517483 2.045944 1.153670
6 -0.608540 2.054763 0.724634
1 -0.820483 1.965823 1.792577
1 -1.318163 2.777244 0.305458
```

P(–Cl)=PG *D*12b

C9H14ClNO, MP2/6-31+G(d), +91.6°

```
6 2.980702 -0.347297 -0.975595
1 2.974787 -1.112120 -1.756553
1 3.847296 -0.534060 -0.335722
1 3.093799 0.629940 -1.444923
6 1.720768 -0.400547 -0.178992
17 1.486189 -1.951107 0.638680
6 0.831198 0.588095 -0.028317
6 -0.422547 0.558408 0.822456
1 -0.277465 -0.058864 1.713229
6 -1.592378 -0.007423 0.018649
8 -2.326162 0.698142 -0.685988
7 -1.769633 -1.340092 0.120376
6 -2.774126 -2.034622 -0.667835
1 -3.769723 -1.653638 -0.438023
```

```
1 -2.591075 -1.904578 -1.736291
1 -2.733135 -3.094090 -0.422881
1 -1.099146 -1.876370 0.651248
6 -0.633844 2.039530 1.165030
1 -1.666546 2.259209 1.442102
1 0.012198 2.291066 2.012110
6 -0.173863 2.793597 -0.088799
1 0.093838 3.830722 0.122566
1 -0.974699 2.787764 -0.827451
6 1.026333 1.981307 -0.608518
1 1.079663 1.972866 -1.700226
1 1.968270 2.400204 -0.239951
```

P(–Cl)=PG *D*12c

$C_9H_{14}ClNO$, MP2/6-311+G(2d,p), +147.2°

```
6 -3.100604 -0.858845 0.599051
1 -3.073240 -1.481684 1.497107
1 -3.753648 -1.351811 -0.125844
1 -3.532227 0.109230 0.854294
6 -1.732133 -0.683921 0.034719
17 -0.958262 -2.200372 -0.439676
6 -1.064226 0.460992 -0.134286
6 0.315486 0.576054 -0.739831
1 0.381247 0.031947 -1.686597
6 1.385990 0.054022 0.205537
8 1.320864 0.198943 1.434091
7 2.447559 -0.524561 -0.391094
6 3.575580 -1.024033 0.376667
1 4.044799 -0.216340 0.940210
1 3.254016 -1.795668 1.078316
1 4.302739 -1.447961 -0.312918
1 2.430610 -0.656898 -1.391417
6 0.487078 2.099117 -0.949763
1 1.535360 2.406996 -0.953618
1 0.050996 2.367760 -1.916812
6 -0.333650 2.722260 0.181938
1 -0.570111 3.773265 0.003910
1 0.217954 2.646999 1.121713
6 -1.584886 1.834504 0.243550
1 -2.069062 1.848772 1.222283
1 -2.320611 2.171205 -0.496698
```

GP(–Cl)=P *A*13a

$C_{12}H_{18}ClNO$, MP2/6-311+G(2d,p), –13.9º

```
6 -2.760655 -2.060529 -0.873182
6 -1.880085 -0.861350 -1.123013
8 -1.369042 -0.630334 -2.230432
7 -1.659967 -0.055414 -0.060092
```

6 -0.884709 1.172203 -0.197308
6 0.601092 0.934623 -0.248288
17 1.437787 2.392116 -0.800474
6 1.327763 -0.149251 0.049819
6 2.837307 -0.220113 -0.080390
6 3.157550 -1.712050 0.076166
6 2.086747 -2.191537 1.055782
6 0.816368 -1.489556 0.552871
1 0.052114 -1.407546 1.324488
1 0.391385 -2.057689 -0.283689
1 1.978995 -3.277481 1.092472
1 2.326892 -1.838142 2.064230
1 3.028438 -2.220635 -0.884829
1 4.177273 -1.892516 0.421423
1 3.190310 0.206345 -1.021874
1 3.302365 0.354505 0.730294
6 -1.315339 1.978485 1.038171
6 -1.566984 0.886292 2.078888
6 -2.277685 -0.194006 1.266694
1 -3.353105 0.002407 1.206954
1 -2.130265 -1.199068 1.666676
1 -0.614526 0.506619 2.458611
1 -2.165311 1.224984 2.925058
1 -0.561835 2.706580 1.338663
1 -2.245593 2.508370 0.814918
1 -1.167785 1.672597 -1.127834
1 -2.284960 -2.740422 -0.161333
1 -3.724157 -1.761774 -0.454381
1 -2.918730 -2.581803 -1.814945

GP(–Cl)=P *A*13b

$C_{12}H_{18}ClNO$, MP2/6-31+G(d), +57.6º

6 -2.855135 -2.310503 0.676277
6 -1.991434 -1.537264 -0.295005
8 -1.724128 -1.998366 -1.427683
7 -1.491120 -0.346618 0.136973
6 -0.813639 0.572737 -0.789285
6 0.599276 0.871491 -0.301420
17 1.031279 2.568430 -0.105505
6 1.556978 -0.047301 -0.068914
6 2.980920 0.254002 0.353795
6 3.724549 -1.067080 0.115681
6 2.662341 -2.126177 0.416409
6 1.388816 -1.548541 -0.220741
1 0.474438 -1.920053 0.246995
1 1.337405 -1.816401 -1.286472
1 2.913864 -3.119701 0.029708
1 2.525521 -2.214405 1.502288
1 4.031903 -1.137569 -0.935878
1 4.623519 -1.164217 0.733821
1 3.408151 1.098544 -0.197171
1 3.006643 0.522011 1.420937

```
6 -1.794575 1.749493 -0.772182
6 -2.195573 1.862177 0.707815
6 -2.020079 0.431993 1.273122
1 -2.960504 0.002268 1.618656
1 -1.304271 0.407197 2.102510
1 -1.544578 2.563255 1.236468
1 -3.225216 2.212681 0.824277
1 -1.392275 2.675660 -1.187127
1 -2.657513 1.452095 -1.379238
1 -0.759664 0.086703 -1.768578
1 -2.440410 -2.308327 1.687816
1 -3.862954 -1.882768 0.718508
1 -2.929602 -3.337188 0.315416
```

GP(–Cl)=P *A*13c

$C_{12}H_{18}ClNO$, MP2/6-311+G(2d,p), +145.0º

```
6 3.410372 1.907746 0.374393
6 2.038010 1.494340 -0.097992
8 1.104152 2.305487 -0.217023
7 1.872618 0.183921 -0.376565
6 0.583586 -0.325808 -0.836580
6 -0.474311 -0.302307 0.231909
17 0.077972 -0.677601 1.862395
6 -1.771412 -0.056907 0.019342
6 -2.867217 -0.089713 1.060766
6 -4.160301 -0.012088 0.235792
6 -3.757664 0.846403 -0.965856
6 -2.365653 0.307498 -1.328337
1 -1.756941 1.026706 -1.882226
1 -2.463911 -0.590244 -1.950837
1 -4.463051 0.791409 -1.797264
1 -3.675467 1.893690 -0.657108
1 -4.438056 -1.013729 -0.108145
1 -5.000346 0.397248 0.799983
1 -2.807893 -0.974077 1.699302
1 -2.778172 0.788032 1.713304
6 0.942883 -1.733962 -1.341493
6 2.128382 -2.153297 -0.467422
6 2.911465 -0.852189 -0.289164
1 3.644337 -0.718258 -1.091298
1 3.429028 -0.802720 0.670217
1 1.779556 -2.515688 0.500030
1 2.737560 -2.931736 -0.928107
1 0.093062 -2.415127 -1.282411
1 1.257072 -1.660273 -2.386401
1 0.219680 0.303157 -1.652343
1 3.616301 1.477308 1.357784
1 4.186990 1.561040 -0.310496
1 3.444354 2.992841 0.445286
```

P(–H)=PG *D*14a

$C_9H_{15}NO$, MP2/6-311+G(2d,p), +11.6°

```
6 -2.990559 -1.684352 -0.103329
6 -1.760556 -1.077983 -0.709243
6 -1.128070 0.011192 -0.245793
6 0.099644 0.624385 -0.884391
6 1.425377 0.020620 -0.414736
8 2.508168 0.557705 -0.692915
7 1.349881 -1.114591 0.300308
6 2.539608 -1.804894 0.766741
1 2.232656 -2.672635 1.347099
1 3.153774 -2.133253 -0.073717
1 3.138931 -1.145811 1.396128
1 0.432101 -1.514580 0.447588
6 0.017751 2.091196 -0.430097
6 -0.467295 1.961642 1.017855
6 -1.545713 0.866271 0.940831
1 -2.526962 1.312717 0.745420
1 -1.631290 0.289708 1.865633
1 0.361279 1.626855 1.650502
1 -0.843747 2.899178 1.430901
1 0.969113 2.614067 -0.534103
1 -0.735302 2.610447 -1.031343
1 0.091875 0.522991 -1.974179
1 -1.358491 -1.564308 -1.598182
1 -2.800310 -2.712672 0.218736
1 -3.337505 -1.114807 0.759898
1 -3.801757 -1.725852 -0.836240
```

P(–H)=PG *D*14b

$C_9H_{15}NO$, MP2/6-31+G(d), +72.8°

```
6 -2.582110 2.239191 -0.239752
6 -1.441393 1.512769 0.411437
6 -1.101071 0.230280 0.179018
6 0.042329 -0.499683 0.874192
6 1.368003 -0.276377 0.140233
8 1.938011 -1.165667 -0.525759
7 1.902840 0.958069 0.276199
6 3.121383 1.350713 -0.417670
1 3.373874 2.366609 -0.113092
1 3.940469 0.680047 -0.148652
1 2.984943 1.319523 -1.502833
1 1.365423 1.672476 0.752947
6 -0.401944 -1.965466 0.813468
6 -1.090750 -2.083967 -0.552201
6 -1.846527 -0.752186 -0.713162
1 -2.884640 -0.851821 -0.368698
1 -1.889184 -0.419824 -1.757443
1 -0.330459 -2.187875 -1.332422
1 -1.756687 -2.951528 -0.614835
```

1 0.430472 -2.664200 0.939835
1 -1.126420 -2.143437 1.619038
1 0.163350 -0.140367 1.905232
1 -0.864127 2.088686 1.138455
1 -2.219754 3.094207 -0.823569
1 -3.148192 1.586596 -0.909425
1 -3.270557 2.638112 0.514766

P(–H)=PG *D*14c

$C_9H_{15}NO$, MP2/6-311+G(2d,p), +100.5°

6 -2.211048 2.534971 -0.263168
6 -1.113741 1.662549 0.269889
6 -1.104286 0.322430 0.245801
6 0.033336 -0.530281 0.813313
6 1.279820 -0.332706 -0.027634
8 1.438776 -0.876793 -1.129590
7 2.198377 0.512206 0.486220
6 3.410831 0.853870 -0.238599
1 3.973948 1.575354 0.349959
1 4.026832 -0.031200 -0.405447
1 3.166080 1.293420 -1.206571
1 2.046844 0.897293 1.406276
6 -0.501781 -1.970273 0.726928
6 -1.479648 -1.949622 -0.453056
6 -2.176261 -0.589860 -0.316500
1 -3.004531 -0.664505 0.399046
1 -2.589924 -0.224650 -1.259057
1 -0.920106 -1.985617 -1.389478
1 -2.180197 -2.787355 -0.434836
1 0.297434 -2.705058 0.611521
1 -1.044291 -2.196990 1.650134
1 0.250335 -0.243535 1.846328
1 -0.261697 2.174176 0.717433
1 -1.829498 3.218553 -1.027646
1 -3.019852 1.948243 -0.700238
1 -2.633557 3.154311 0.533908

GP(–H)=P *A*15a

$C_{12}H_{19}NO$, MP2/6-311+G(2d,p), –11.8º

6 1.021298 -0.982922 -0.939733
7 1.482102 0.114808 -0.085504
6 1.541571 1.383285 -0.536724
6 2.091610 2.404642 0.429684
1 1.429740 2.503308 1.294430
1 3.077054 2.111501 0.798427
1 2.163497 3.364890 -0.076812
8 1.147893 1.696173 -1.674338
6 1.988493 -0.351187 1.213133

```
6 1.557414 -1.816437 1.235429
6 1.633362 -2.208537 -0.241068
1 1.095134 -3.127133 -0.478047
1 2.677780 -2.323023 -0.546428
1 0.525760 -1.897391 1.588313
1 2.193960 -2.422460 1.880883
1 3.079314 -0.262339 1.249968
1 1.566021 0.232137 2.033780
6 -0.472748 -1.095101 -1.064176
6 -1.437860 -0.488872 -0.356520
6 -2.916364 -0.739671 -0.607347
6 -3.601106 0.492473 -0.006897
6 -2.755967 0.756424 1.241958
6 -1.313967 0.554137 0.745422
1 -0.635522 0.258772 1.547483
1 -0.936339 1.491401 0.321598
1 -2.916662 1.743994 1.679207
1 -2.994871 0.006739 2.003788
1 -3.515329 1.339038 -0.696817
1 -4.659567 0.334647 0.210325
1 -3.138622 -0.906847 -1.663603
1 -3.236422 -1.630321 -0.052236
1 -0.788172 -1.795589 -1.839387
1 1.440962 -0.843762 -1.940618
```

GP(–H)=P *A*15b

$C_{12}H_{19}NO$, MP2/6-31+G(d), +72.1º

```
7 -1.474202 0.152976 -0.126678
6 -0.736954 -0.730493 0.803533
6 0.543762 -1.261670 0.199241
6 1.657392 -0.537504 -0.022227
6 2.946735 -1.089287 -0.598966
6 4.006981 -0.098858 -0.101529
6 3.278355 1.246599 -0.201821
6 1.857360 0.932844 0.300390
1 1.089101 1.567342 -0.152440
1 1.785995 1.089626 1.386307
1 3.763043 2.046500 0.368400
1 3.243634 1.563055 -1.252621
1 4.257681 -0.312839 0.946106
1 4.934904 -0.133459 -0.682986
1 3.127298 -2.129074 -0.304075
1 2.911176 -1.058308 -1.698376
1 0.559612 -2.325072 -0.051096
6 -1.686620 1.466313 0.148798
6 -2.608597 2.213700 -0.790172
1 -2.468339 1.919509 -1.833439
1 -3.654662 2.027526 -0.521045
1 -2.409891 3.281102 -0.682347
8 -1.147056 2.050442 1.117767
6 -2.355692 -0.647155 -1.007766
6 -2.302721 -2.063768 -0.406253
```

```
6 -1.759997 -1.849051 1.011857
1 -1.316401 -2.751619 1.445139
1 -2.554507 -1.496353 1.680556
1 -1.621765 -2.697373 -0.983061
1 -3.287257 -2.540733 -0.418187
1 -3.369886 -0.240637 -0.979519
1 -2.000993 -0.621467 -2.043764
1 -0.532912 -0.152111 1.710538
```

GP(–H)=P *A*15c

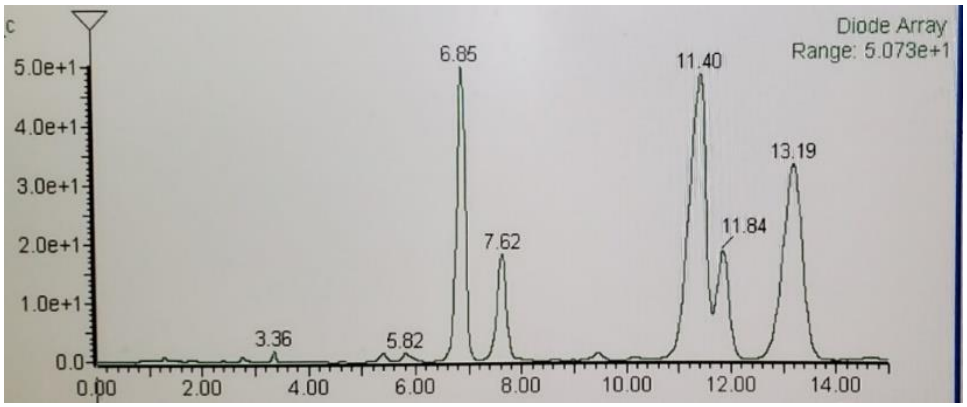
$C_{12}H_{19}NO$, MP2/6-311+G(2d,p), +119.6º

```
6 0.664769 -0.647449 -0.581773
7 1.788655 0.143153 -0.041092
6 1.745777 1.490593 -0.001696
6 2.974685 2.179476 0.545295
1 3.114568 1.932006 1.600429
1 3.874028 1.868550 0.009137
1 2.846116 3.255123 0.444201
8 0.750163 2.128179 -0.388218
6 2.954295 -0.683621 0.305195
6 2.407350 -2.106479 0.201206
6 1.349621 -1.984392 -0.898487
1 0.638705 -2.811774 -0.913588
1 1.836237 -1.922959 -1.876553
1 1.942575 -2.398636 1.145826
1 3.190469 -2.828952 -0.031030
1 3.761356 -0.520810 -0.416954
1 3.325730 -0.443800 1.302595
6 -0.437607 -0.780391 0.425077
6 -1.689811 -0.337949 0.244477
6 -2.814845 -0.477829 1.253257
6 -4.080083 -0.227632 0.422275
6 -3.621426 0.849734 -0.566537
6 -2.224191 0.368386 -0.989256
1 -1.563028 1.174407 -1.315043
1 -2.314773 -0.345708 -1.816502
1 -4.300734 0.983019 -1.410928
1 -3.539838 1.808740 -0.043861
1 -4.350149 -1.136937 -0.125496
1 -4.938918 0.073439 1.025732
1 -2.801388 -1.441263 1.767965
1 -2.722537 0.307490 2.013623
1 -0.181983 -1.280445 1.360057
1 0.290485 -0.153675 -1.480420
```

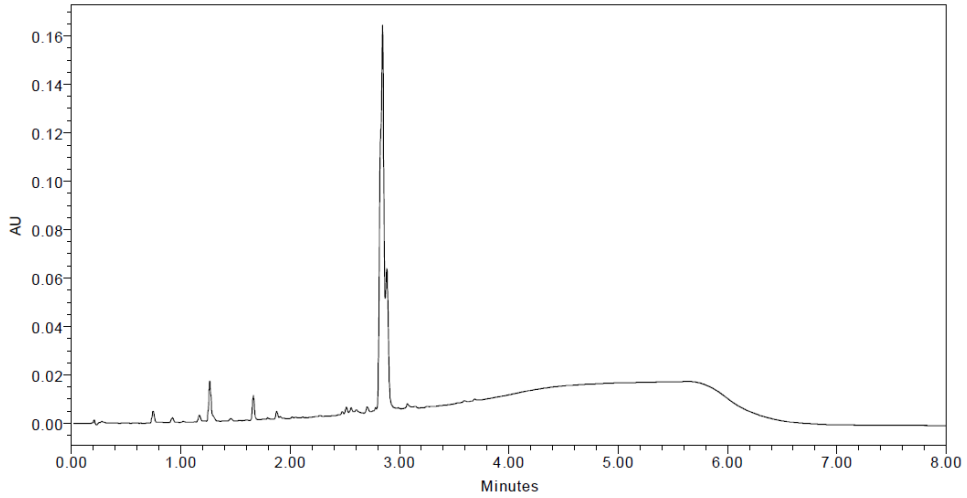
Appendix B

Supplementary Information for Chapter 3

A)



B)



C)

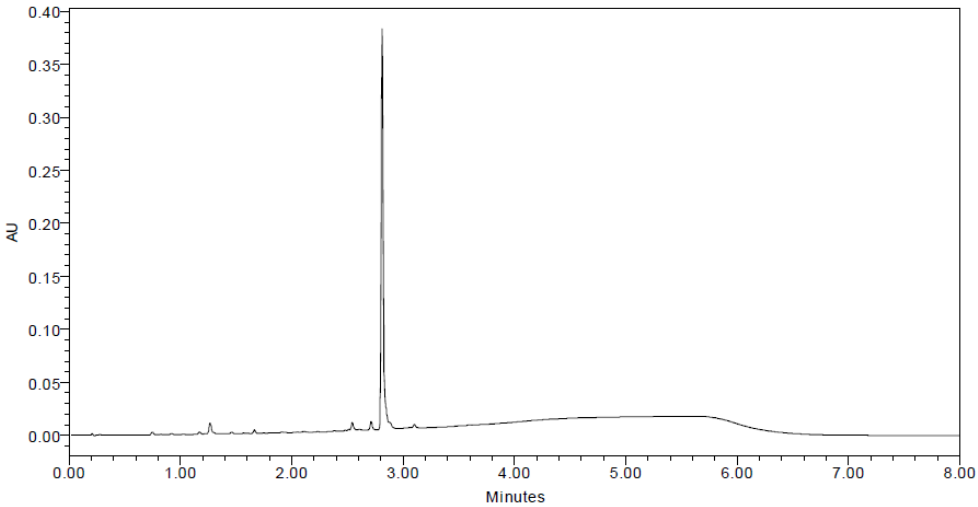


Figure B.1. HPLC chromatograph of Benzyl Ester **3.13-(*Z*)** for separation of diastereomers. **A)** The mixture of diastereomers. The peaks at 11.40 min and 11.84 min correspond to the two rotamers of the major diastereomer **3.13-(*S*,*Z*)**. The peak at 13.19 min corresponds to the minor diastereomer **3.13-(*R*,*Z*)** (rotamer peak not observed). **B)** UPLC chromatograph of the separated major diastereomer **3.13-(*S*,*Z*)**. **C)** UPLC chromatograph of the separated minor diastereomer **3.13-(*R*,*Z*)**.

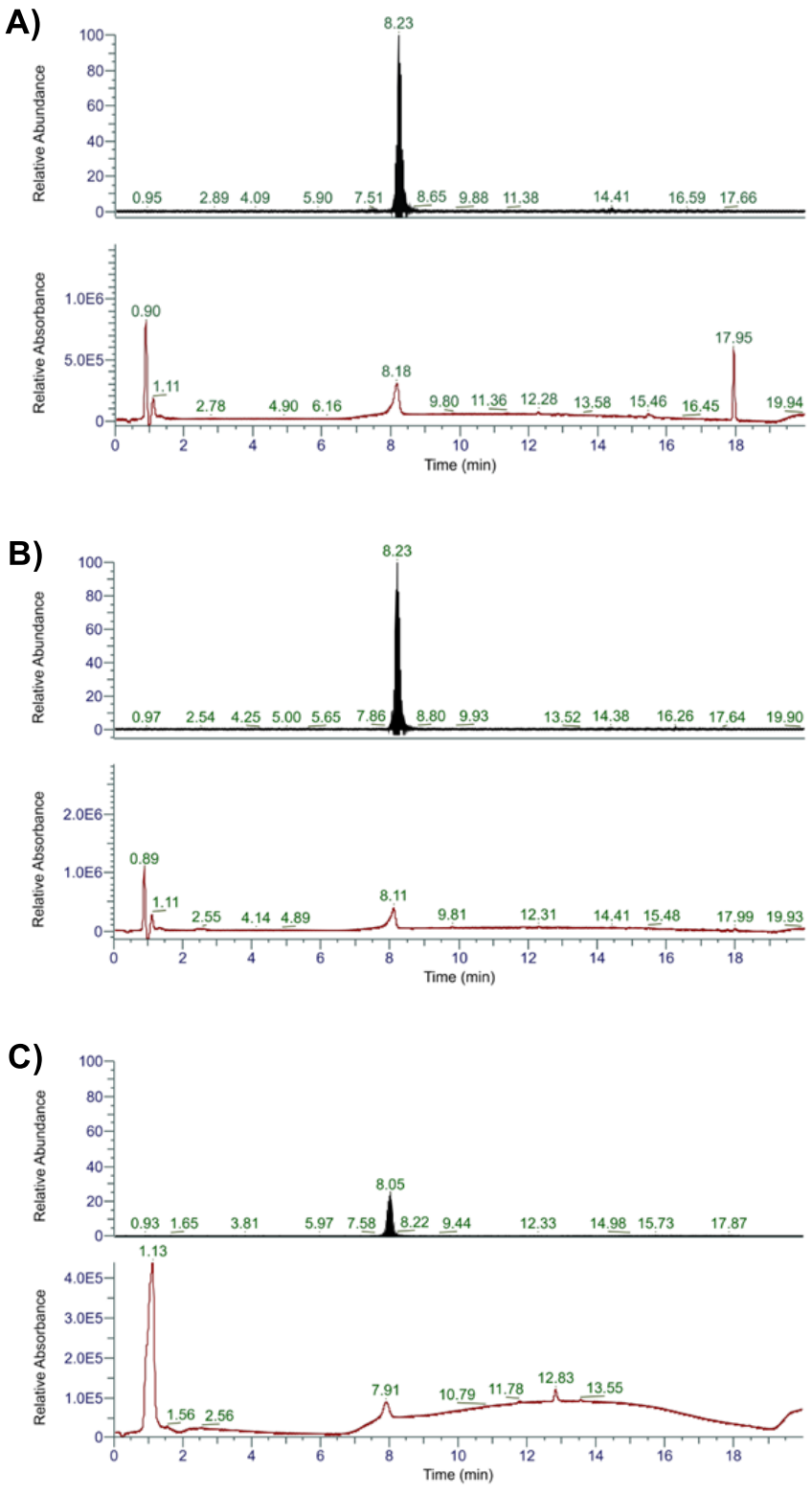
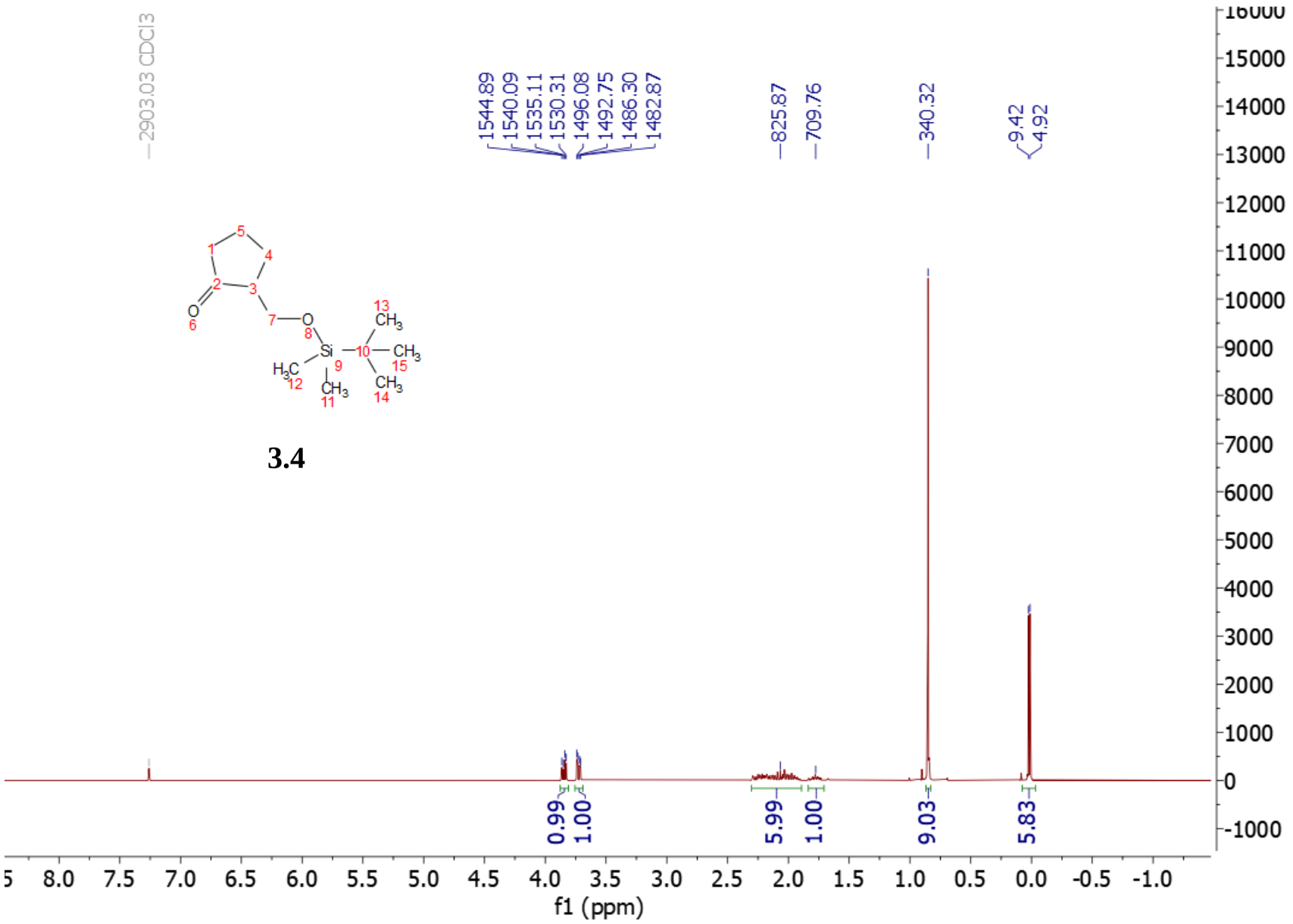
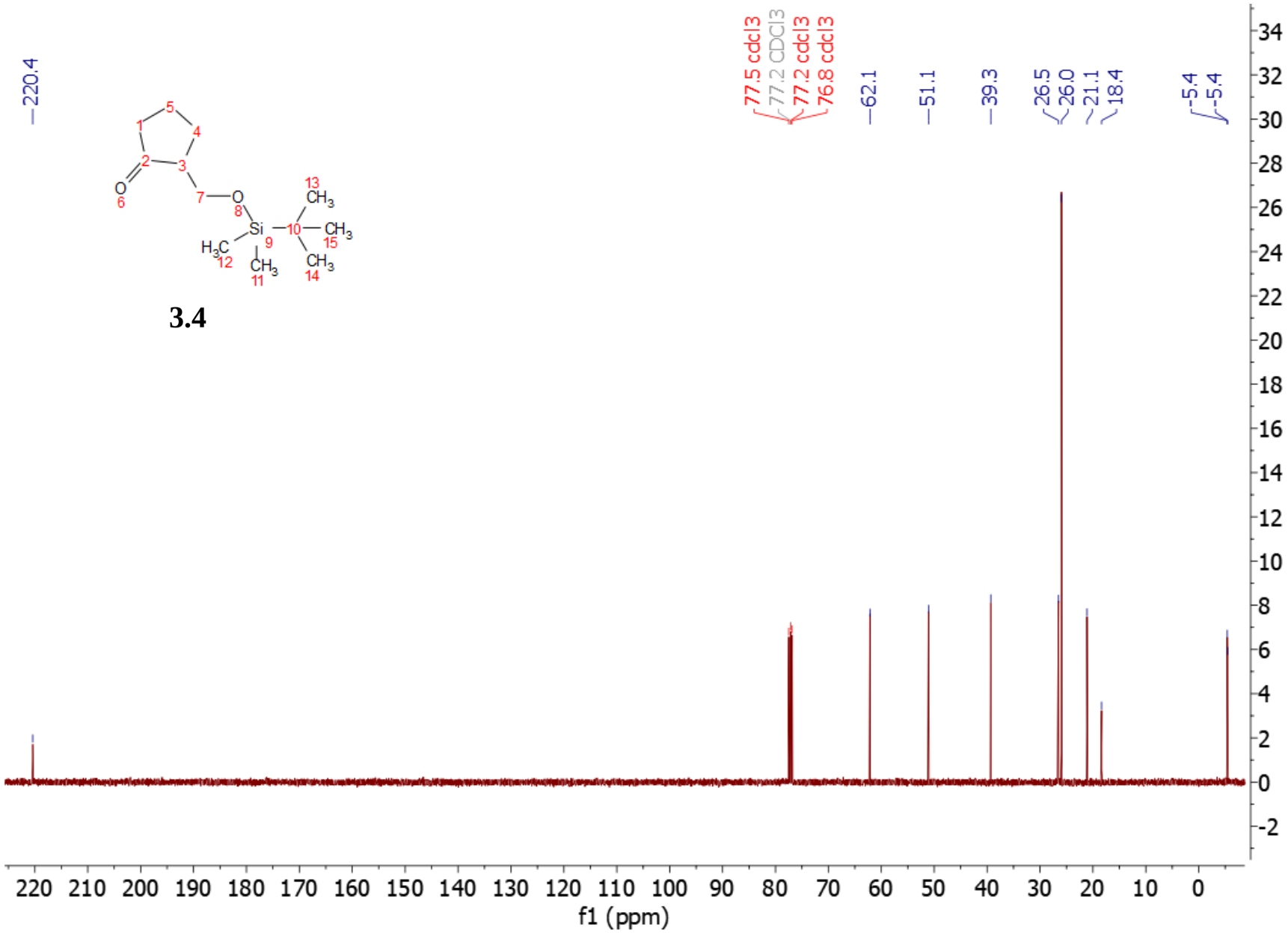


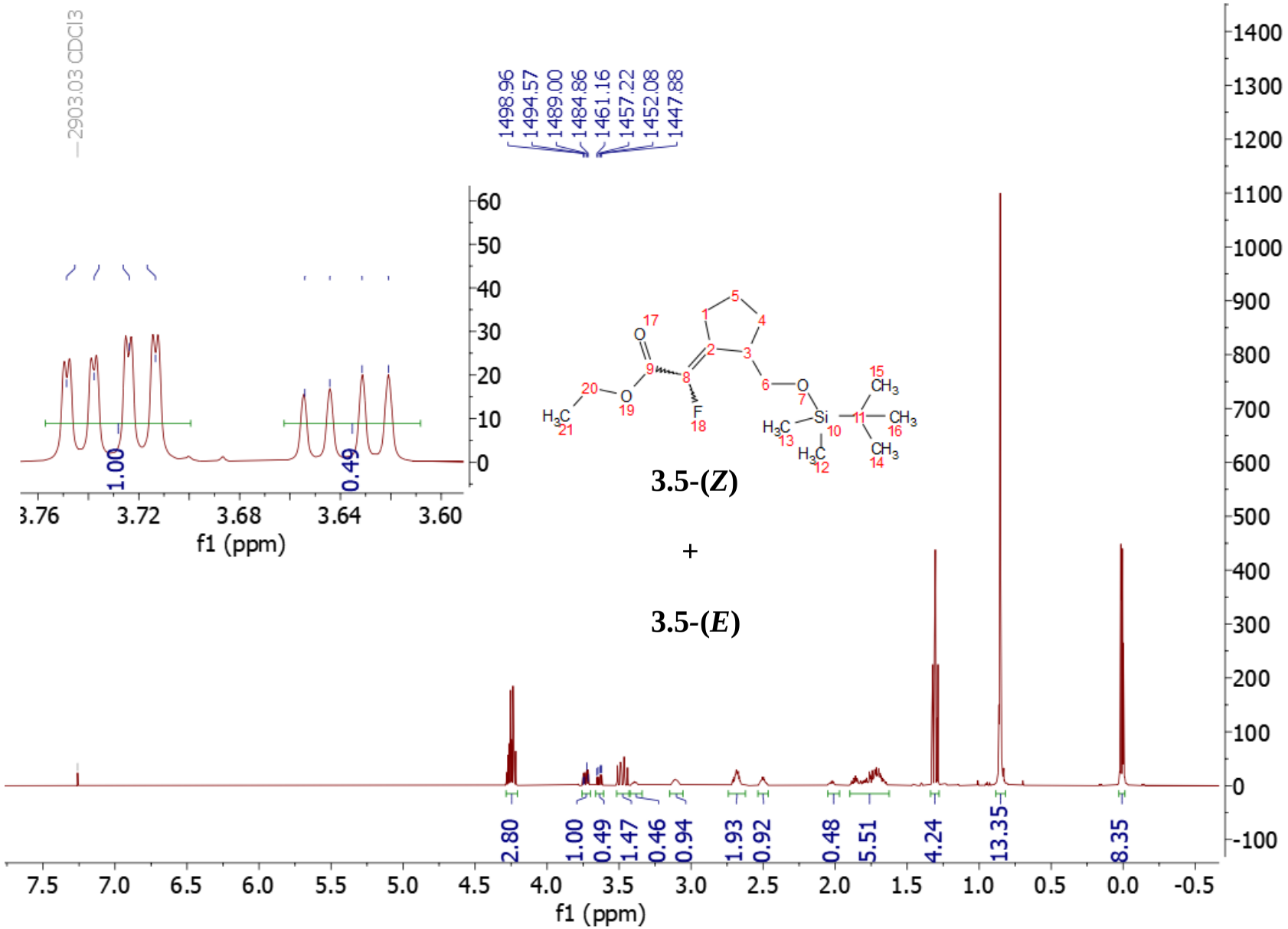
Figure B.2. LC-MS chromatograms showing total ion count and UV absorption at 280 nm of peptides. **A) 3.2-(*R,Z*)**, **B) 3.2-(*S,Z*)** and **C)** control **3.3**.

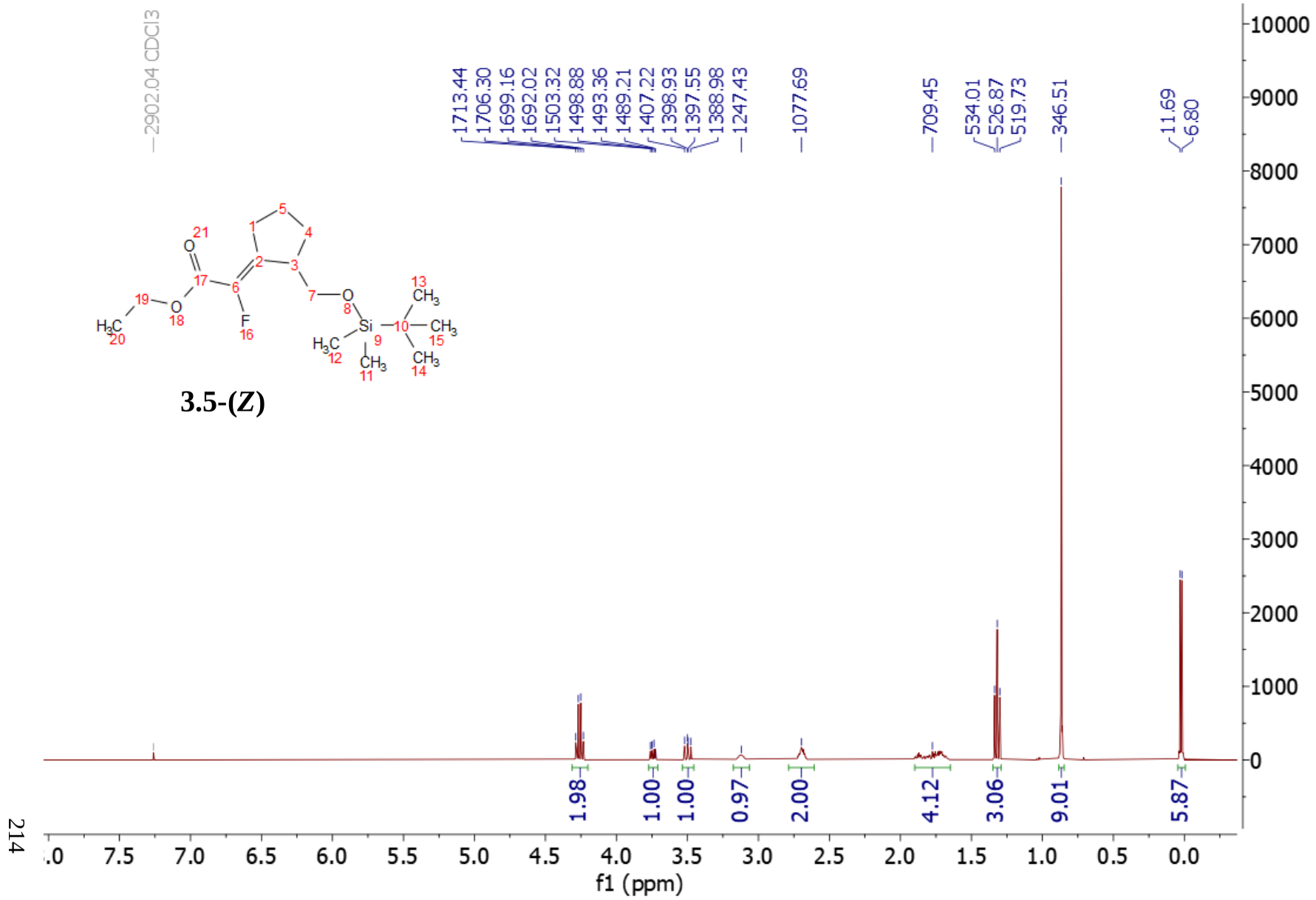

2903.03 CDCl3
1544.89
1540.09
1535.11
1530.31
1496.08
1492.75
1486.30
1482.87
825.87
709.76
340.32
9.42
4.92
0.99
1.00
5.99
1.00
9.03
5.83
f1 (ppm)

3.4

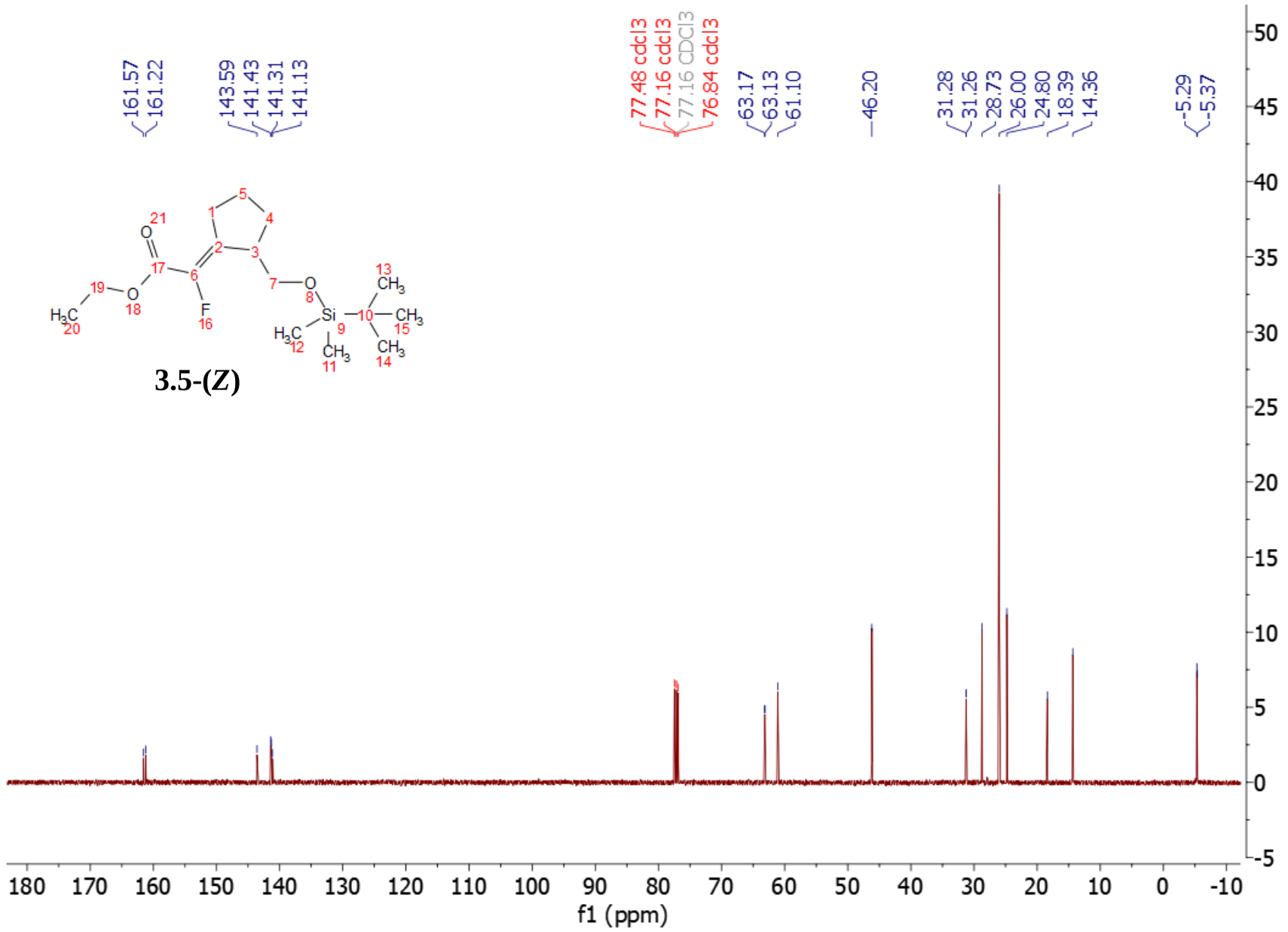

220.4
77.5 cdcl3
77.2 CDCl3
77.2 cdcl3
76.8 cdcl3
62.1
51.1
39.3
26.5
26.0
21.1
18.4
-5.4
-5.4
f1 (ppm)

3.4

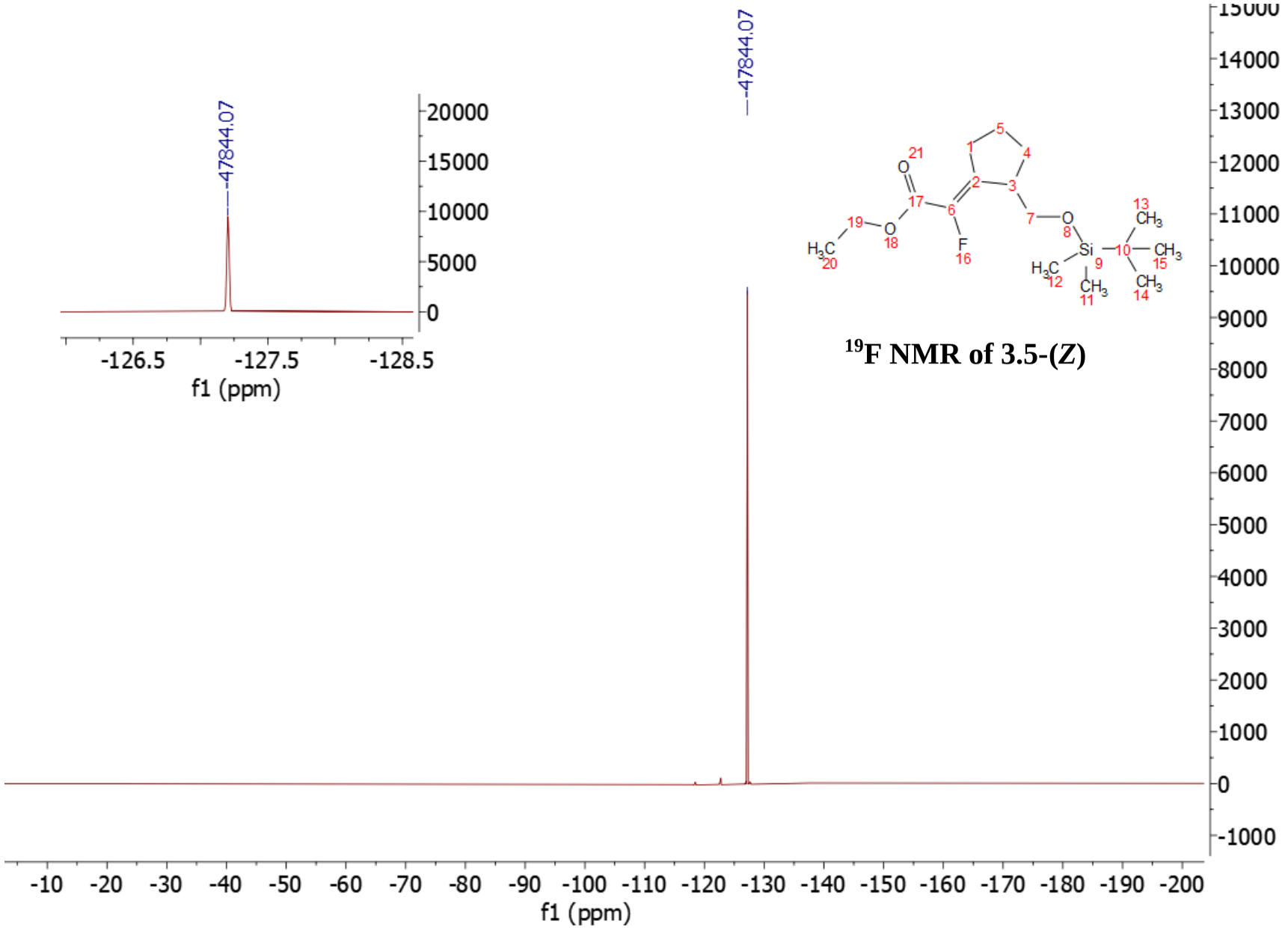

2903.03 CDCl3
1498.96
1494.57
1489.00
1484.86
1461.16
1457.22
1452.08
1447.88
1.00
0.49
f1 (ppm)
3.5-(Z)
+
3.5-(E)
2.80
1.00
0.49
1.47
0.46
0.94
1.93
0.92
0.48
5.51
4.24
13.35
8.35
f1 (ppm)


2902.04 CDCl3
1713.44
1706.30
1699.16
1692.02
1503.32
1498.88
1493.36
1489.21
1407.22
1398.93
1397.55
1388.98
1247.43
1077.69
709.45
534.01
526.87
519.73
346.51
11.69
6.80
1.98
1.00
1.00
0.97
2.00
4.12
3.06
9.01
5.87
f1 (ppm)

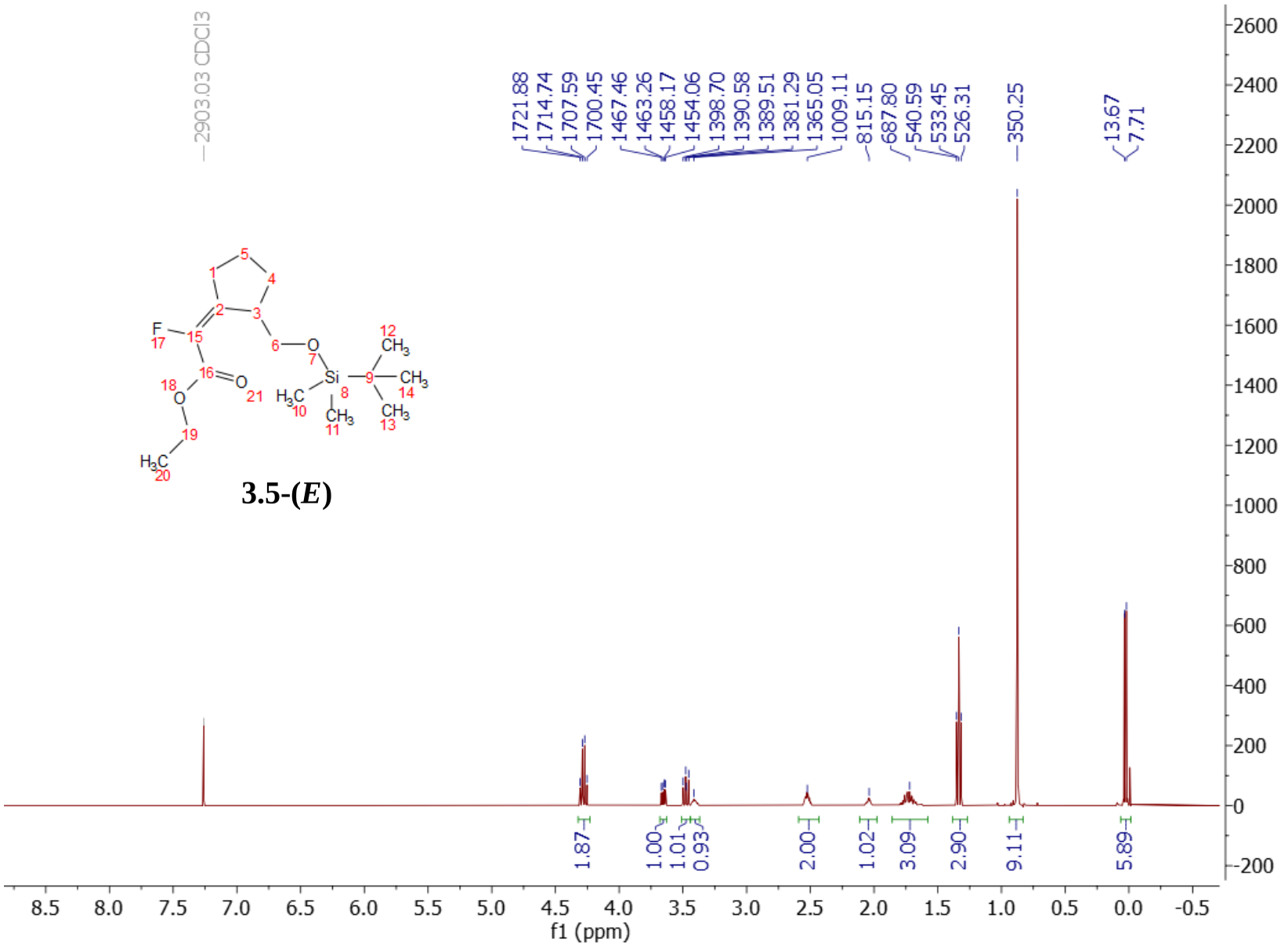
3.5-(Z)


161.57
161.22
143.59
141.43
141.31
141.13
77.48 cdcl3
77.16 cdcl3
77.16 CDCl3
76.84 cdcl3
63.17
63.13
61.10
46.20
31.28
31.26
28.73
26.00
24.80
18.39
14.36
-5.29
-5.37
f1 (ppm)

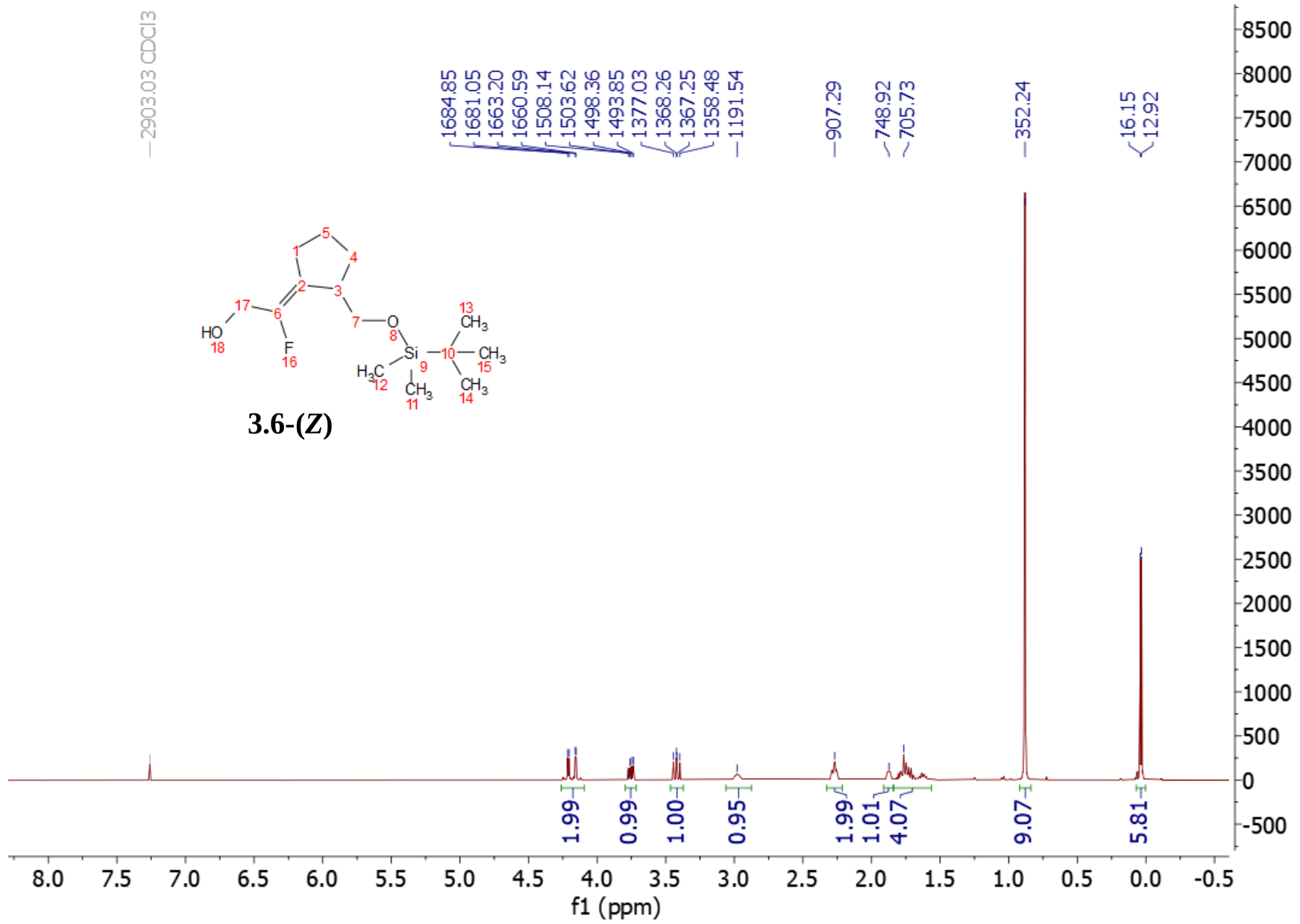
3.5-(Z)



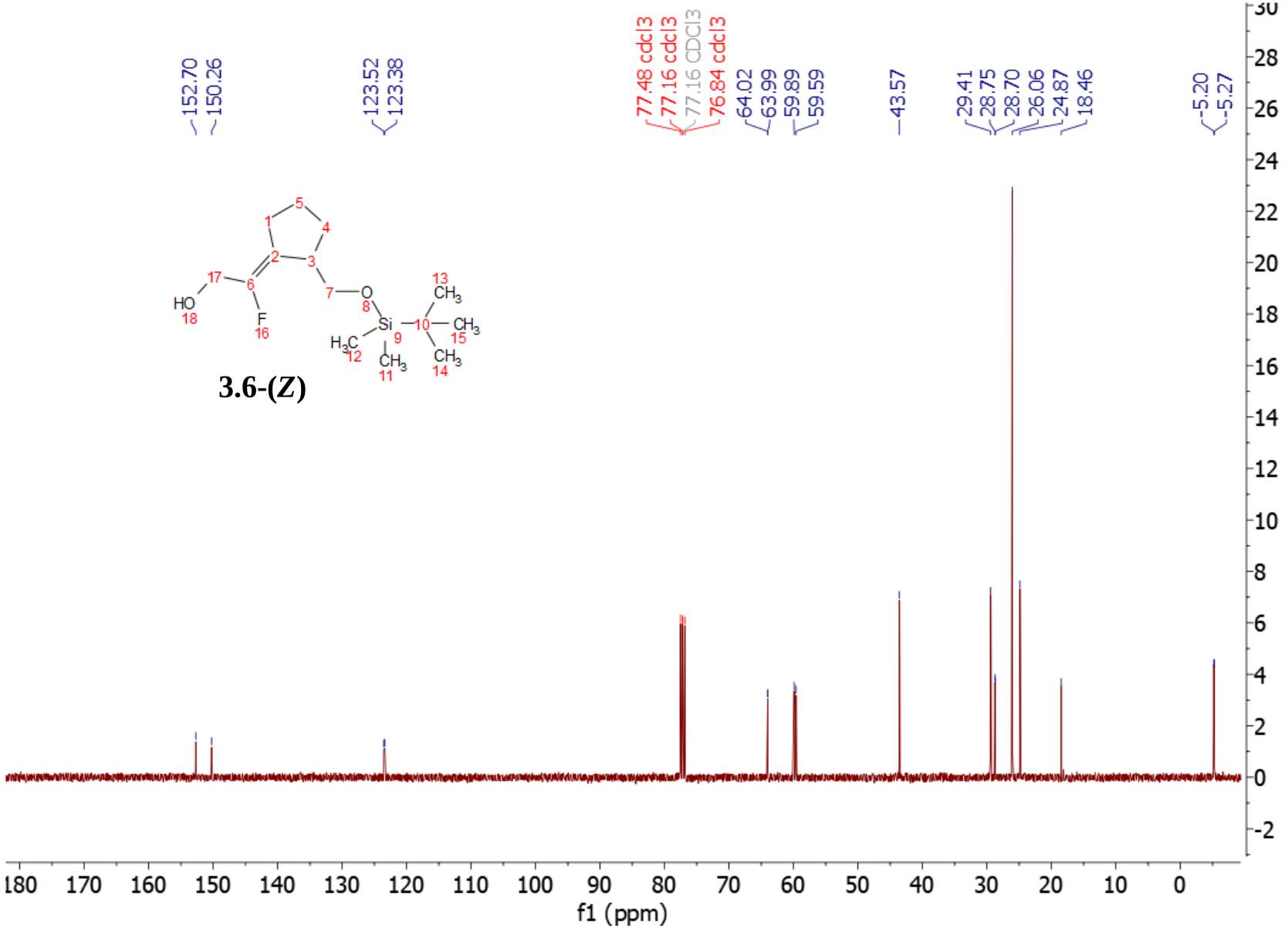
^{19}F NMR of 3.5-(*Z*)


2903.03 CDCl3
1721.88
1714.74
1707.59
1700.45
1467.46
1463.26
1458.17
1454.06
1398.70
1390.58
1389.51
1381.29
1365.05
1009.11
815.15
687.80
540.59
533.45
526.31
350.25
13.67
7.71
1.87
1.00
1.01
0.93
2.00
1.02
3.09
2.90
9.11
5.89
f1 (ppm)

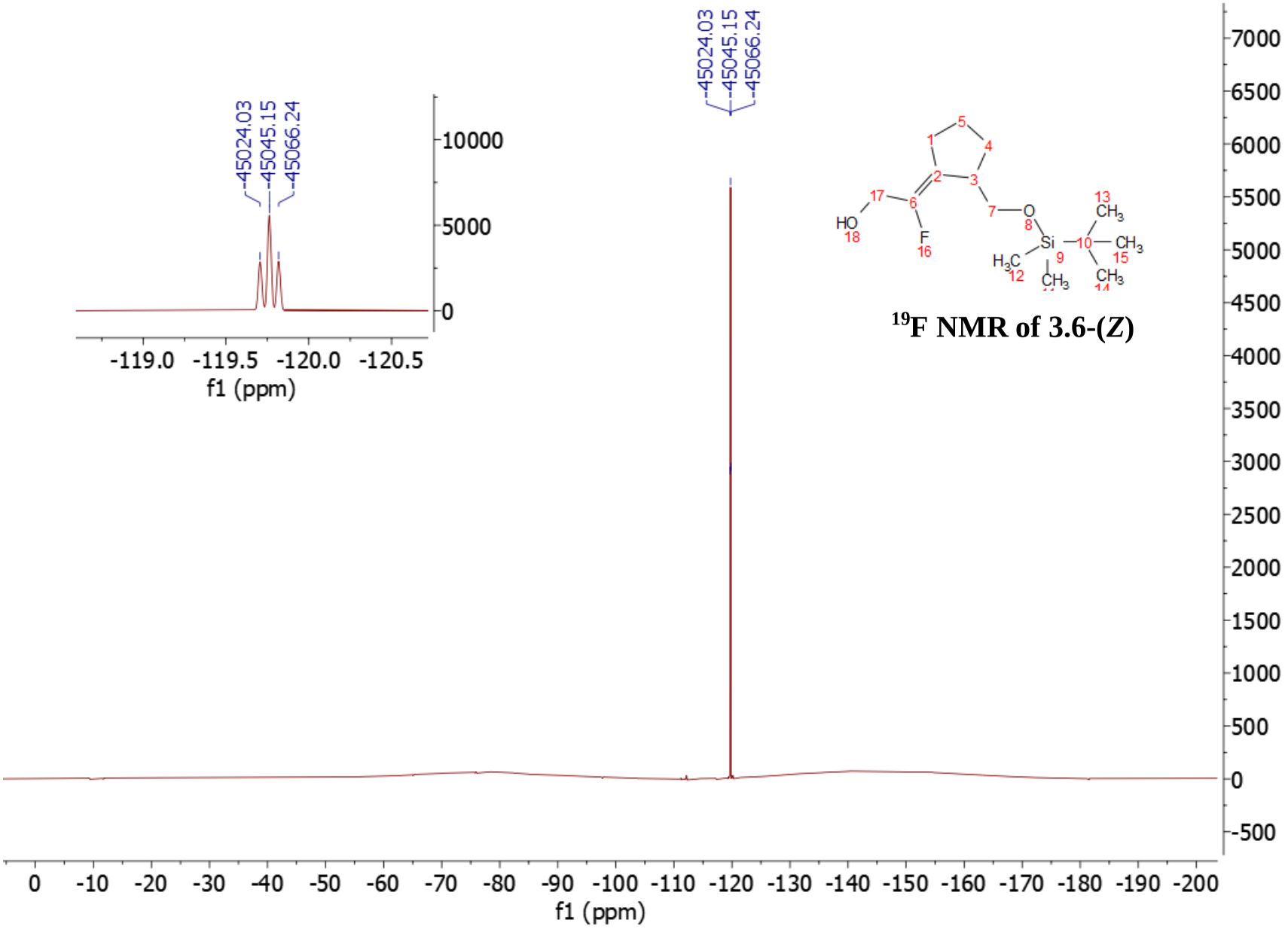
3.5-(*E*)


2903.03 CDCl3
1684.85
1681.05
1663.20
1660.59
1508.14
1503.62
1498.36
1493.85
1377.03
1368.26
1367.25
1358.48
1191.54
907.29
748.92
705.73
352.24
16.15
12.92
HO
F
Si
CH3
f1 (ppm)

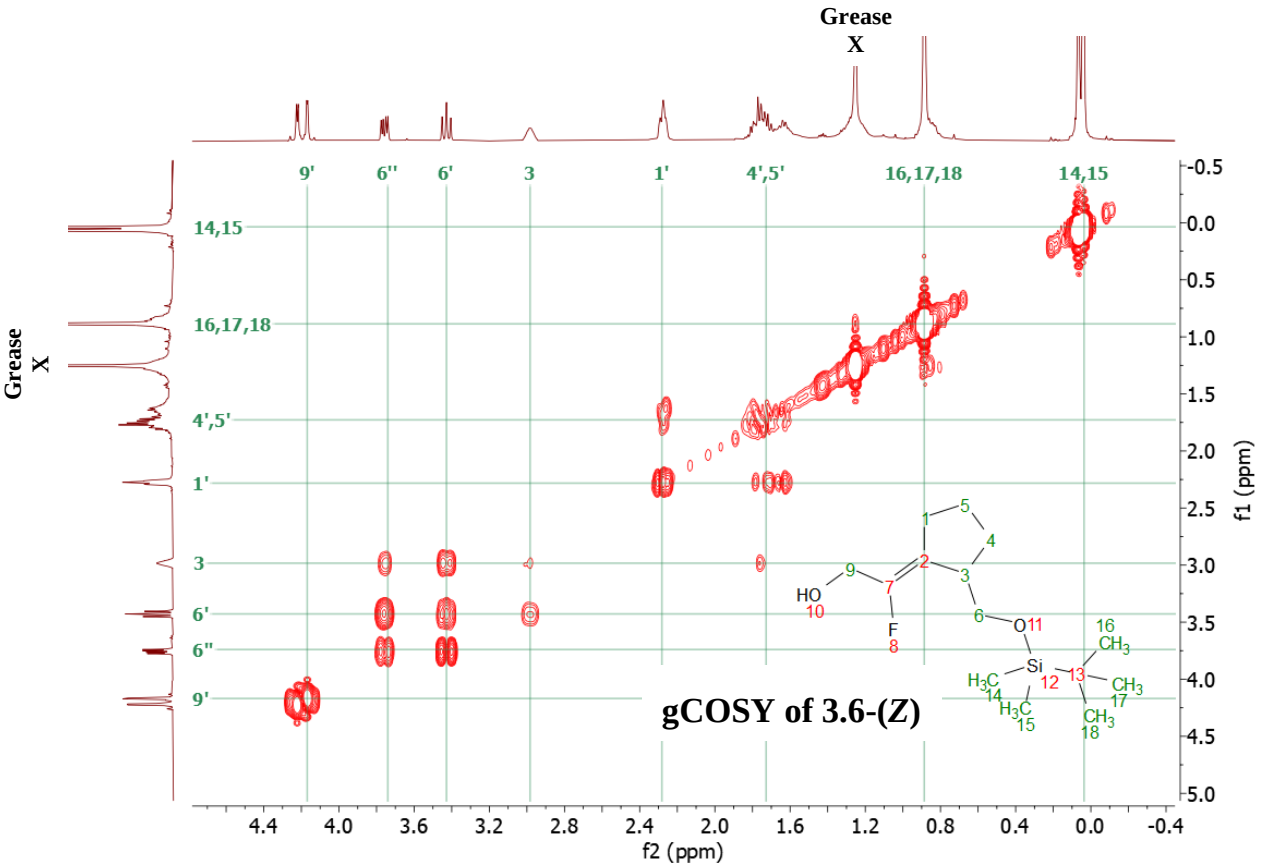
3.6-(Z)

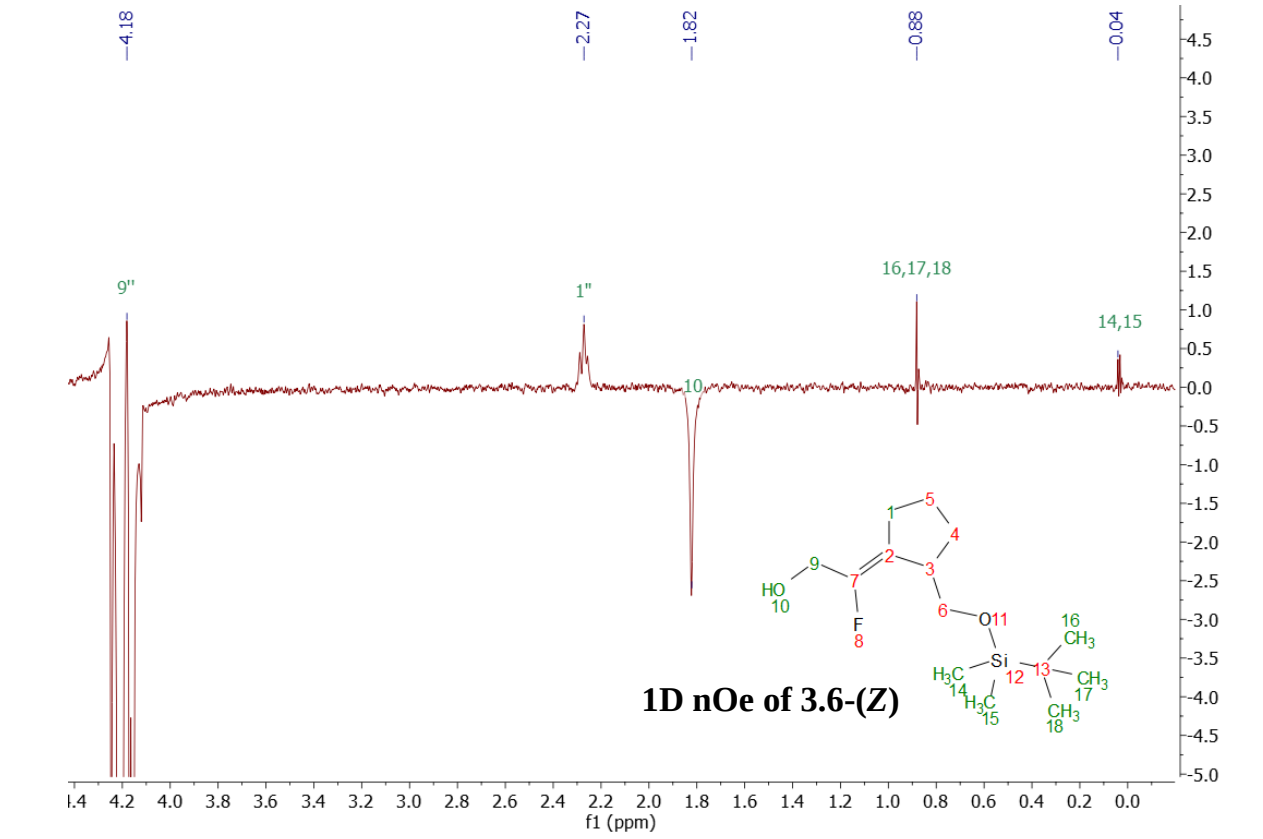

152.70
150.26
123.52
123.38
77.48 cdcl3
77.16 cdcl3
77.16 CDCl3
76.84 cdcl3
64.02
63.99
59.89
59.59
43.57
29.41
28.75
28.70
26.06
24.87
18.46
-5.20
-5.27
HO
18
17
6
F
16
1
2
3
4
5
7
O
8
Si
9
10
H3C
12
CH3
11
CH3
13
CH3
14
CH3
15
f1 (ppm)

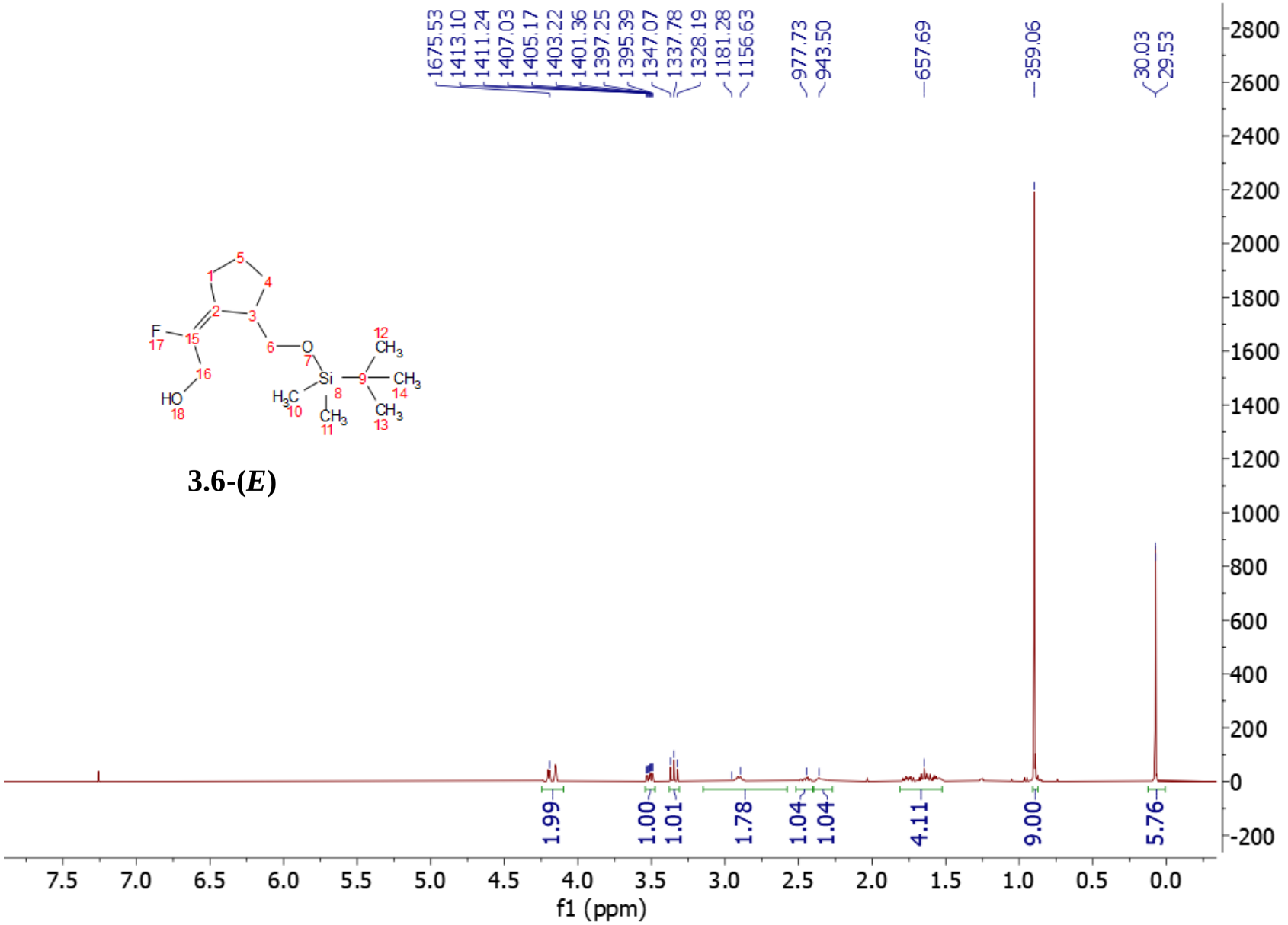
3.6-(Z)



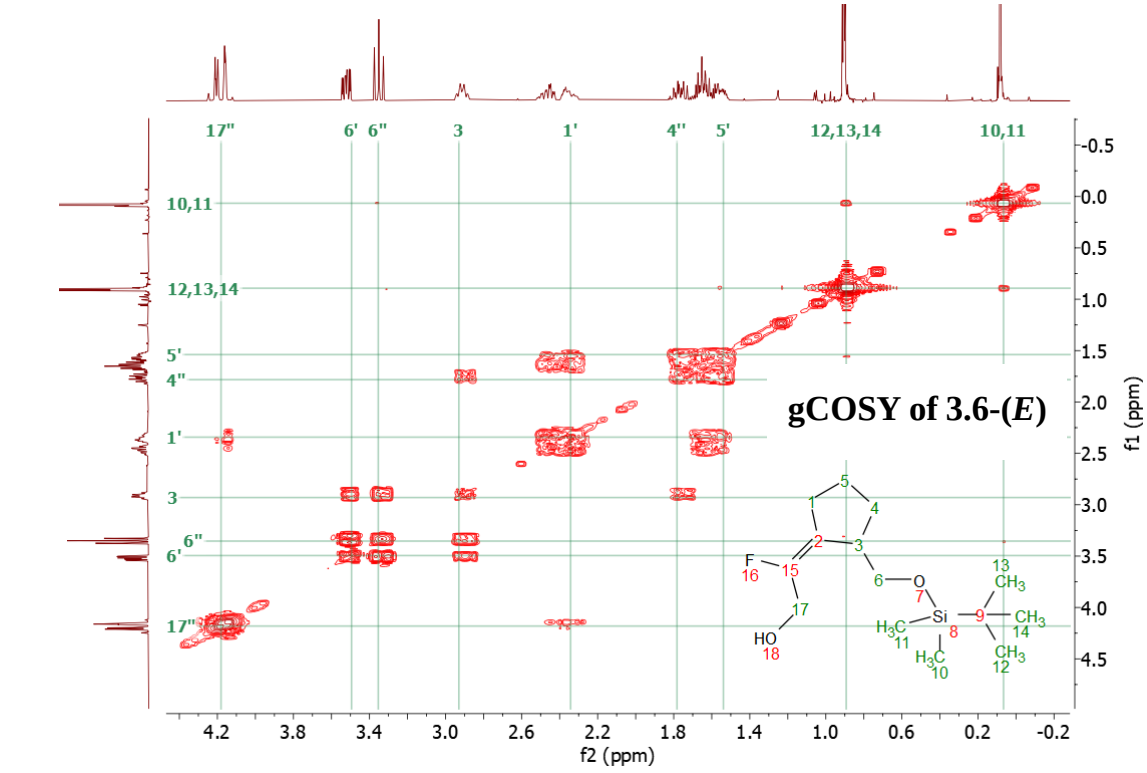
^{19}F NMR of 3.6-(*Z*)

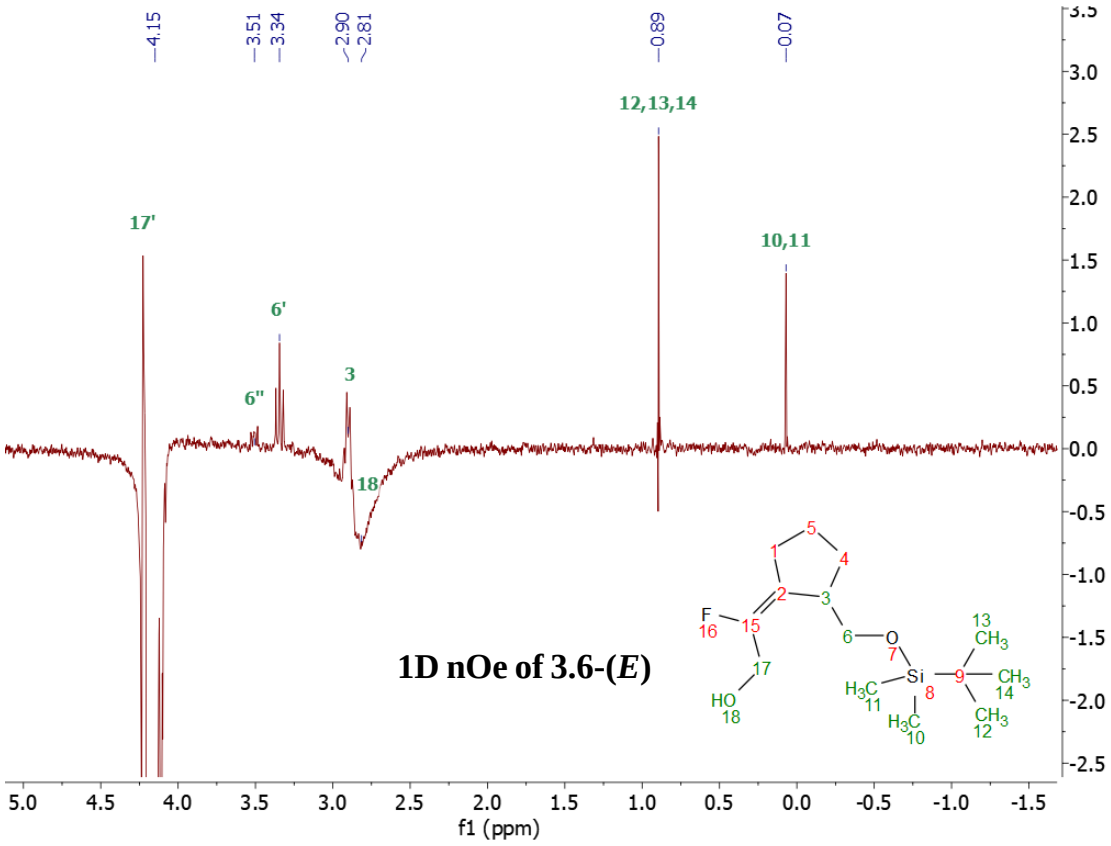

Grease
X
Grease
X
9'
6''
6'
3
1'
4',5'
16,17,18
14,15
f2 (ppm)
f1 (ppm)
gCOSY of 3.6-(Z)

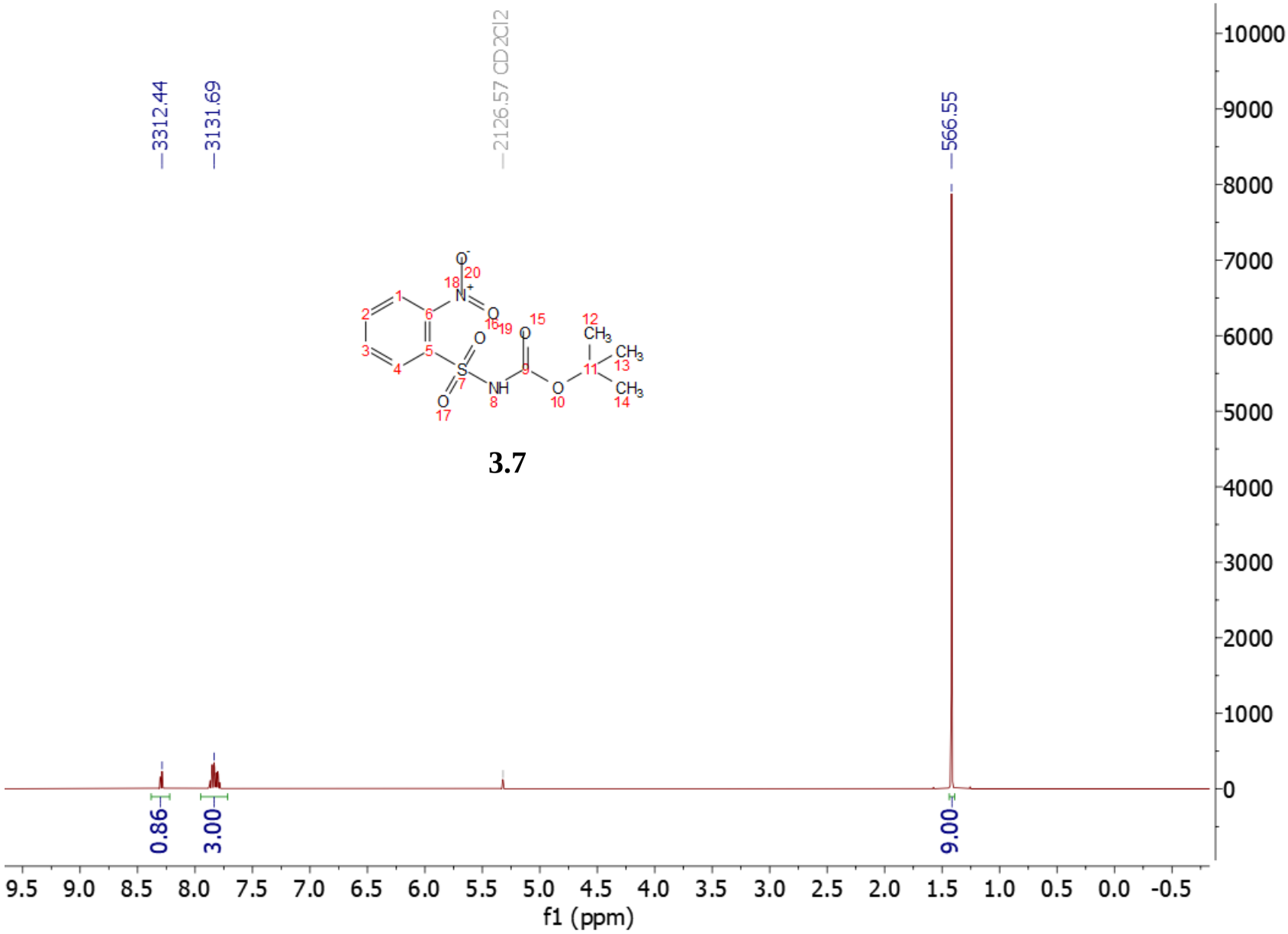

4.18
2.27
1.82
0.88
0.04
9''
1''
16,17,18
14,15
10
f1 (ppm)
1D nOe of 3.6-(Z)


1675.53
1413.10
1411.24
1407.03
1405.17
1403.22
1401.36
1397.25
1395.39
1347.07
1337.78
1328.19
1181.28
1156.63
977.73
943.50
657.69
359.06
30.03
29.53
1.99
1.00
1.01
1.78
1.04
1.04
4.11
9.00
5.76
f1 (ppm)

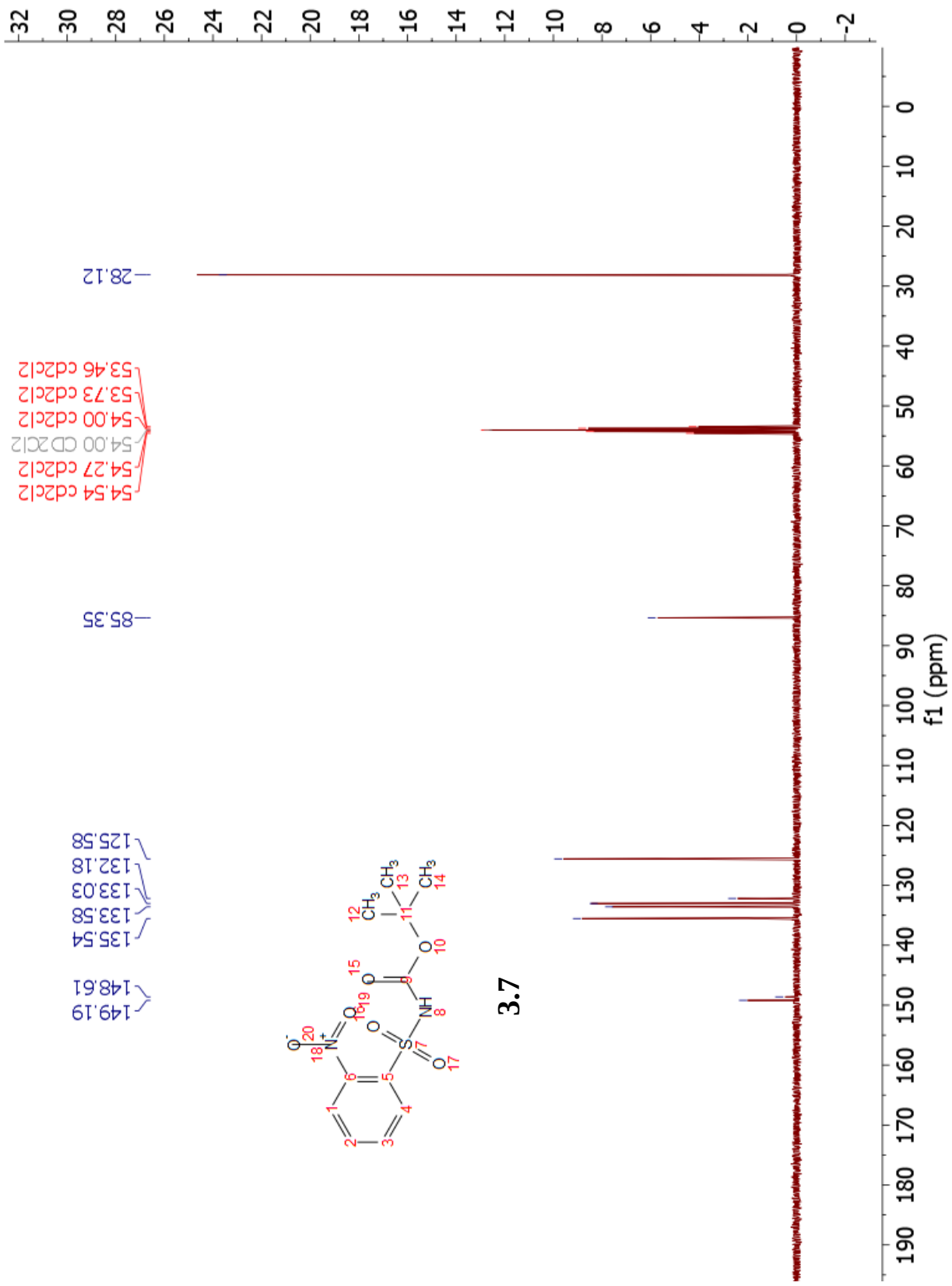
3.6-(*E*)

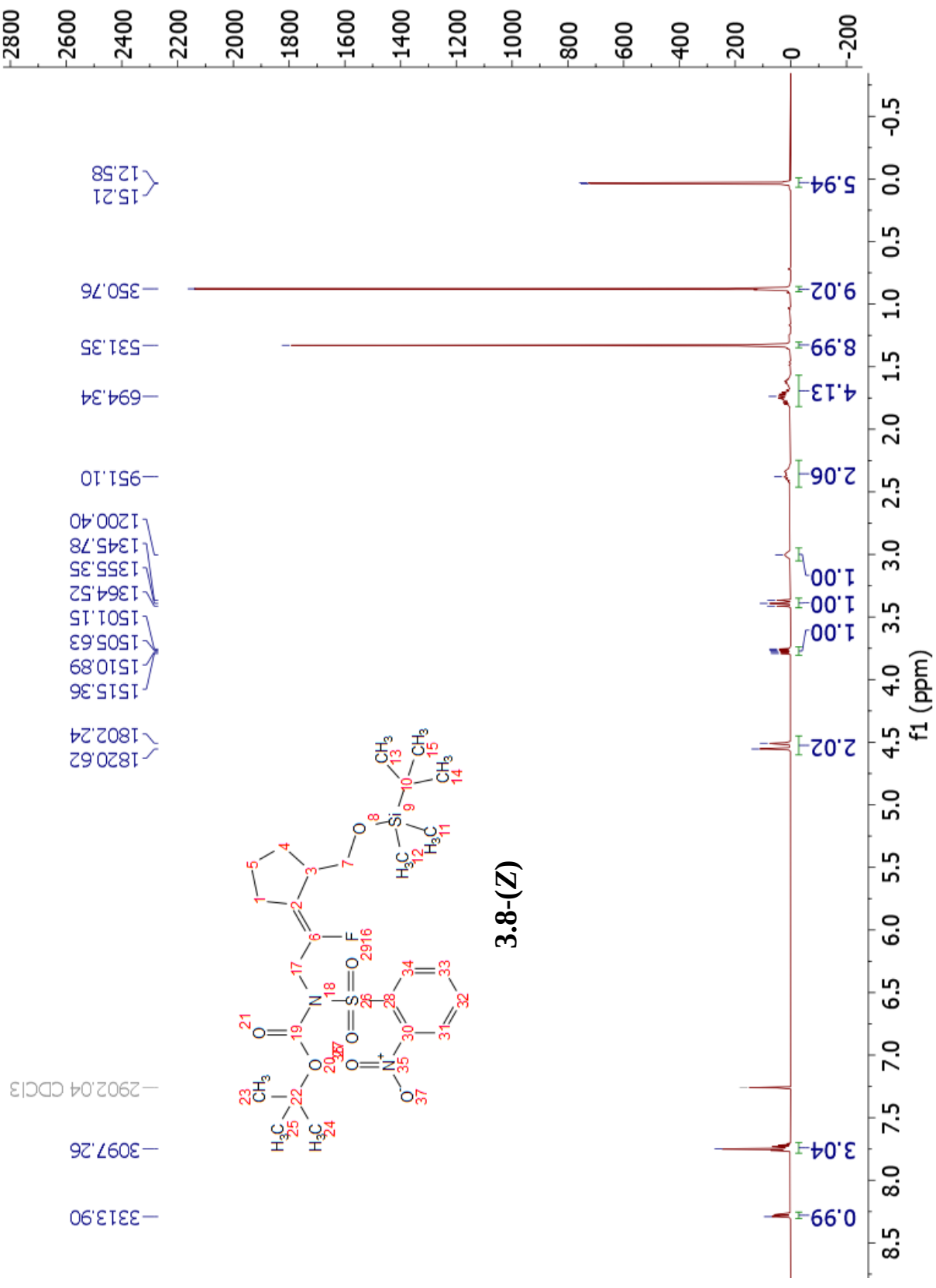

gCOSY of 3.6-(E)
17"
6'
6"
3
1'
4"
5'
12,13,14
10,11
f2 (ppm)
f1 (ppm)

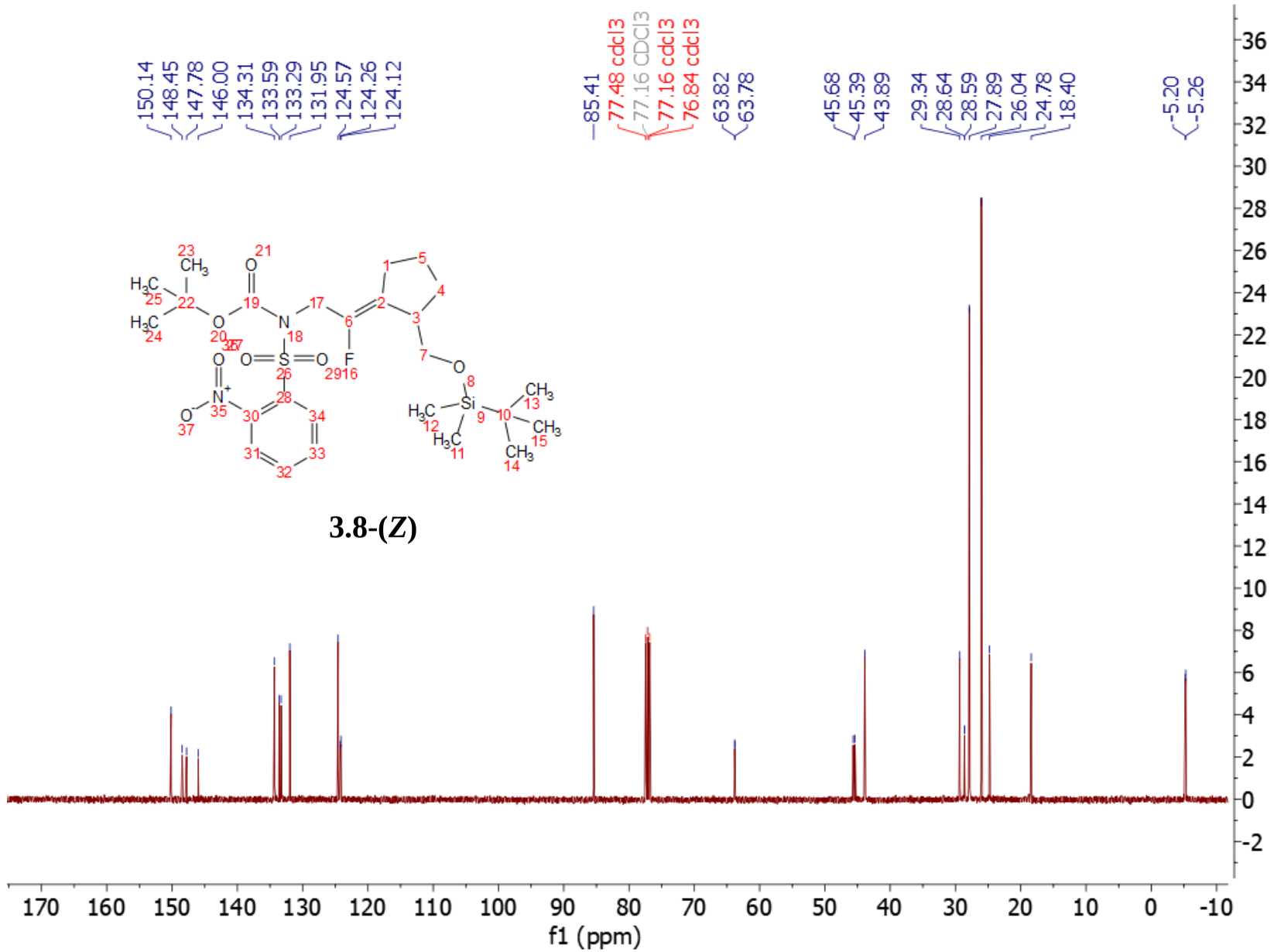

1D nOe of 3.6-(E)
4.15
3.51
3.34
2.90
2.81
0.89
0.07
17'
6'
6"
3
18'
12,13,14
10,11
f1 (ppm)



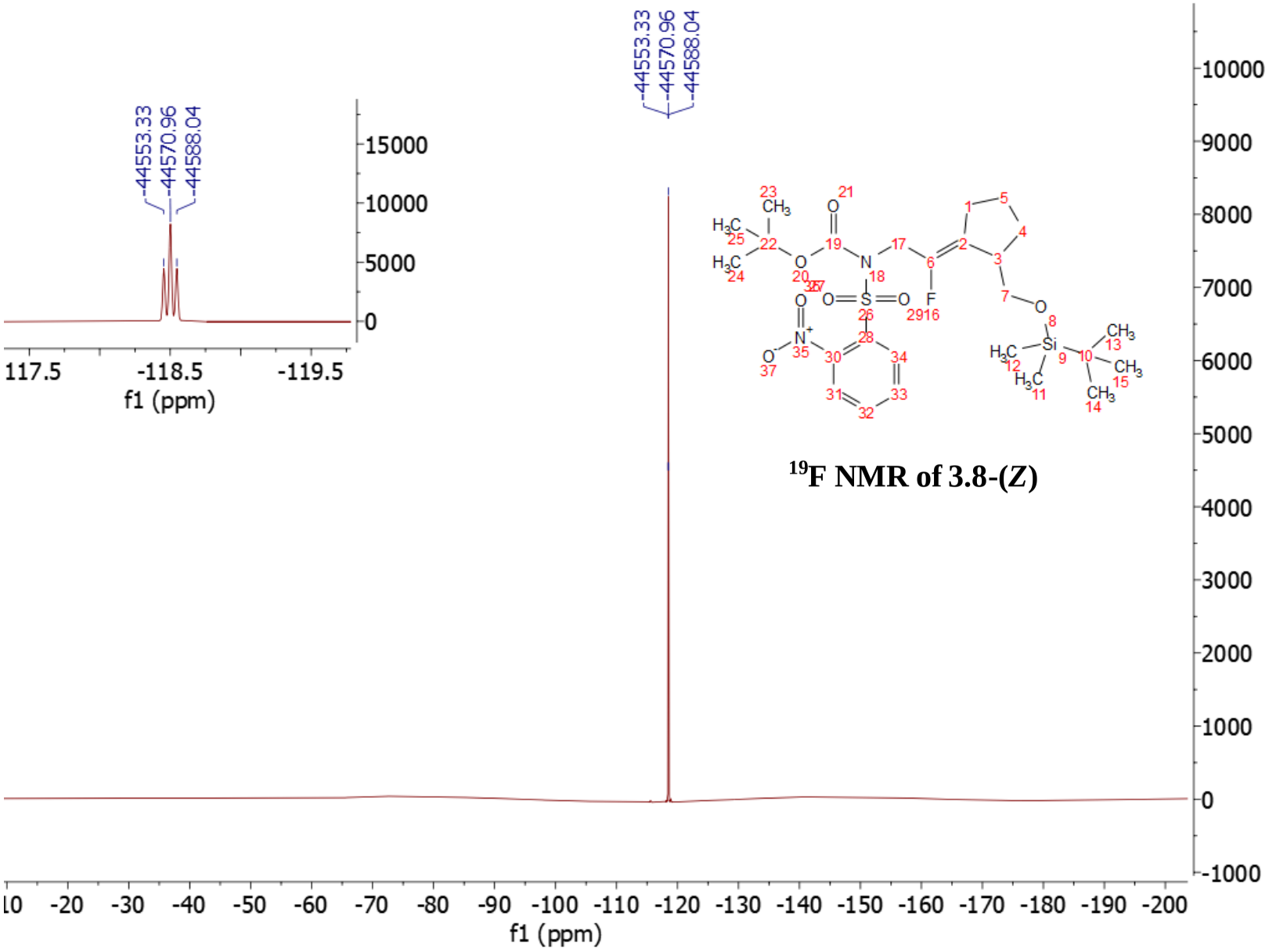
3.7


28.12
53.46 cd2cl2
53.73 cd2cl2
54.00 cd2cl2
54.00 CD2Cl2
54.27 cd2cl2
54.54 cd2cl2
85.35
125.58
132.18
133.03
133.58
135.54
148.61
149.19
3.7
f1 (ppm)

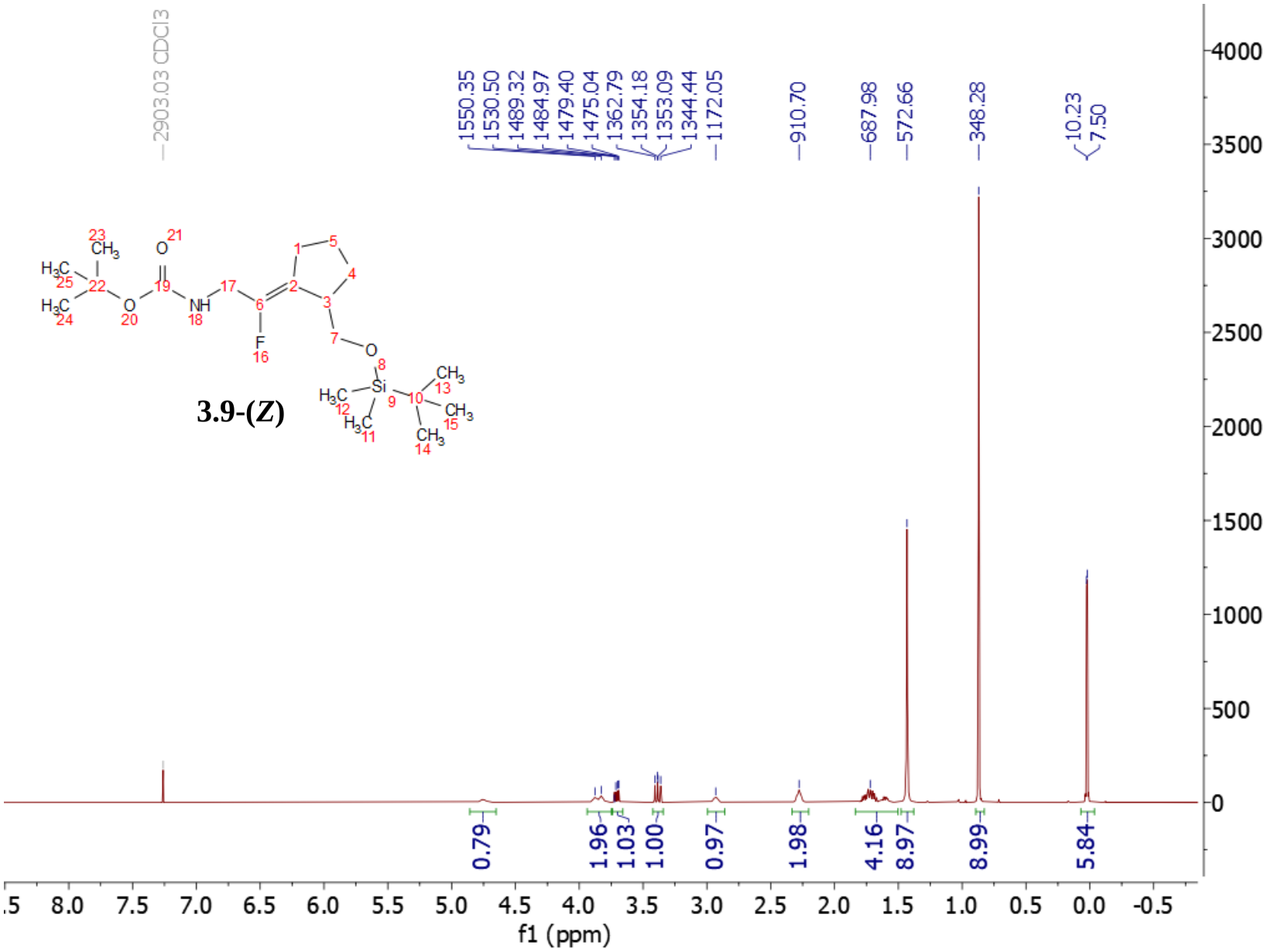

f1 (ppm)
5.94
9.02
8.99
4.13
2.06
1.00
1.00
1.00
2.02
3.04
0.99
15.21
12.58
350.76
531.35
694.34
951.10
1200.40
1345.78
1355.35
1364.52
1501.15
1505.63
1510.89
1515.36
1802.24
1820.62
2902.04 CDCl3
3313.90
3097.26
3.8-(Z)

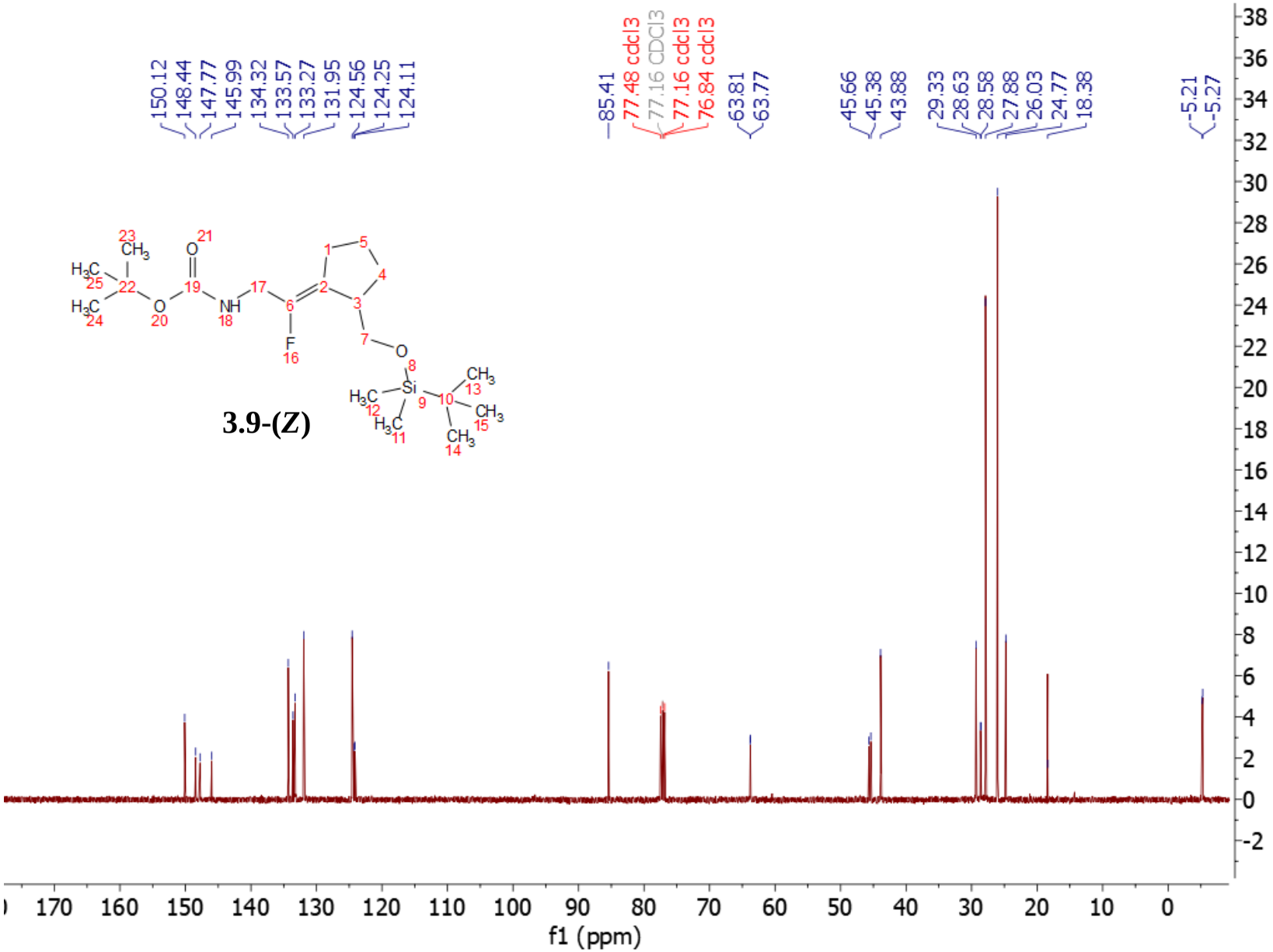

150.14
148.45
147.78
146.00
134.31
133.59
133.29
131.95
124.57
124.26
124.12
85.41
77.48 cdcl3
77.16 CDCl3
77.16 cdcl3
76.84 cdcl3
63.82
63.78
45.68
45.39
43.89
29.34
28.64
28.59
27.89
26.04
24.78
18.40
-5.20
-5.26
f1 (ppm)

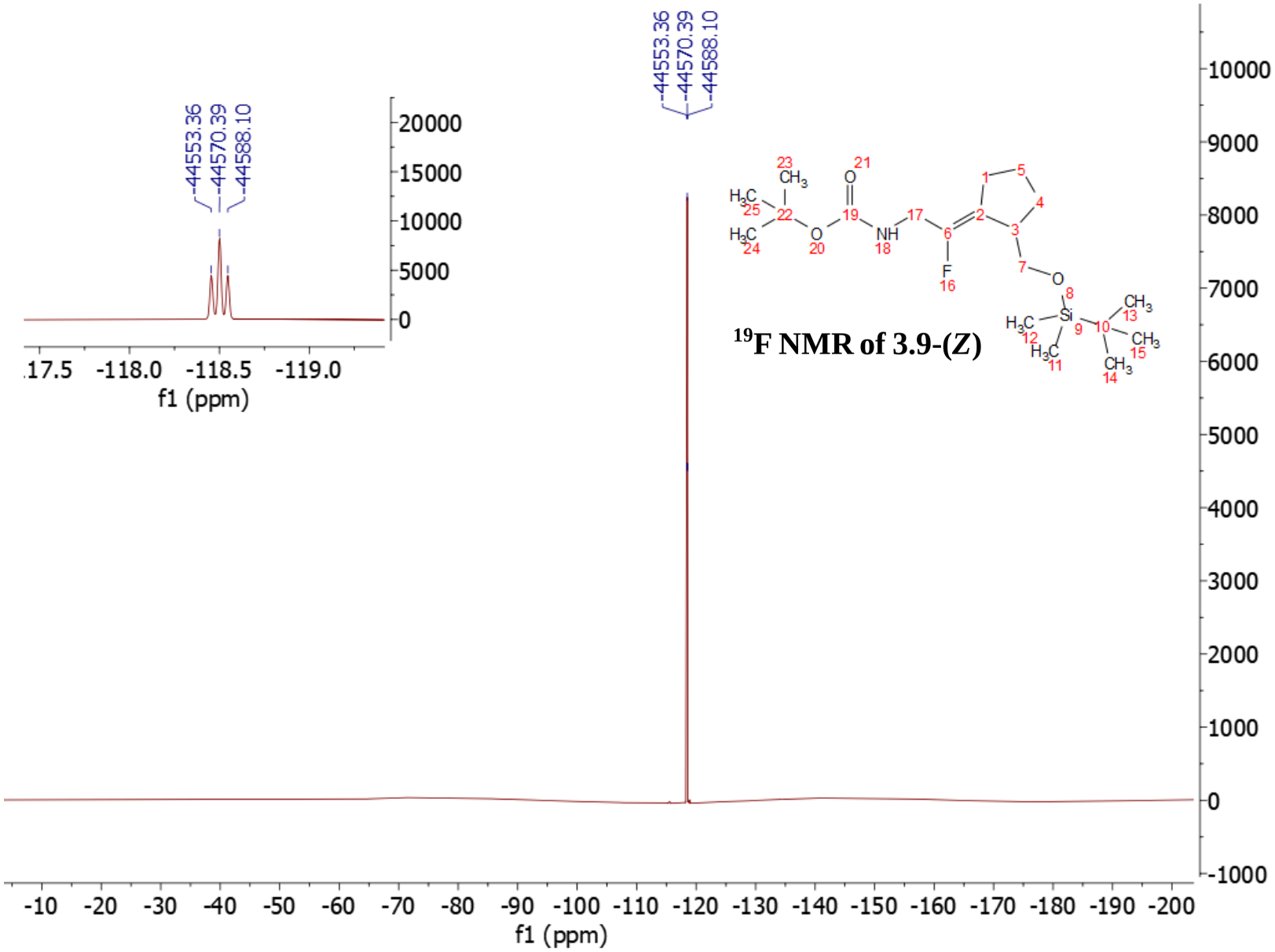
3.8-(Z)



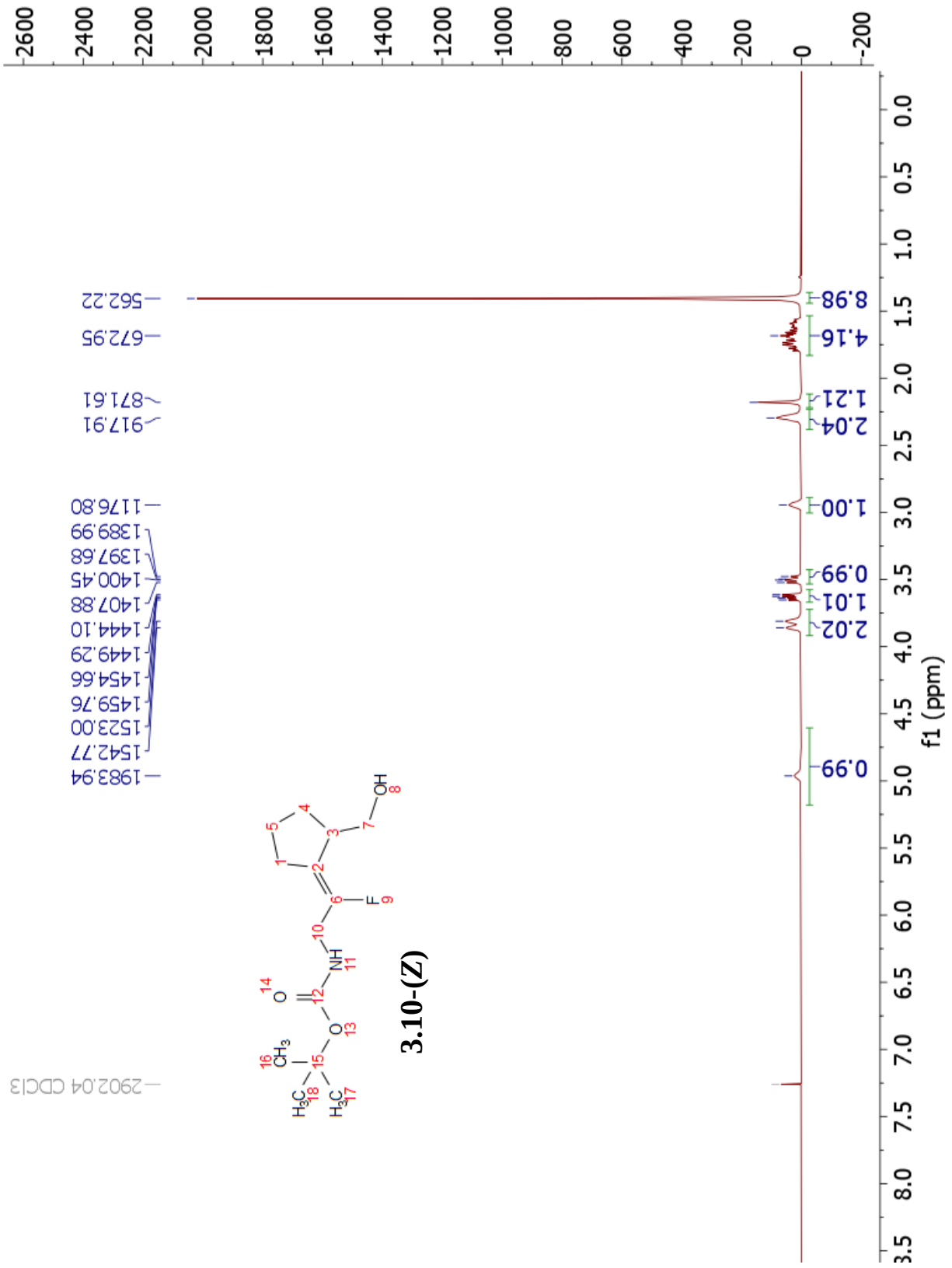
^{19}F NMR of 3.8-(*Z*)

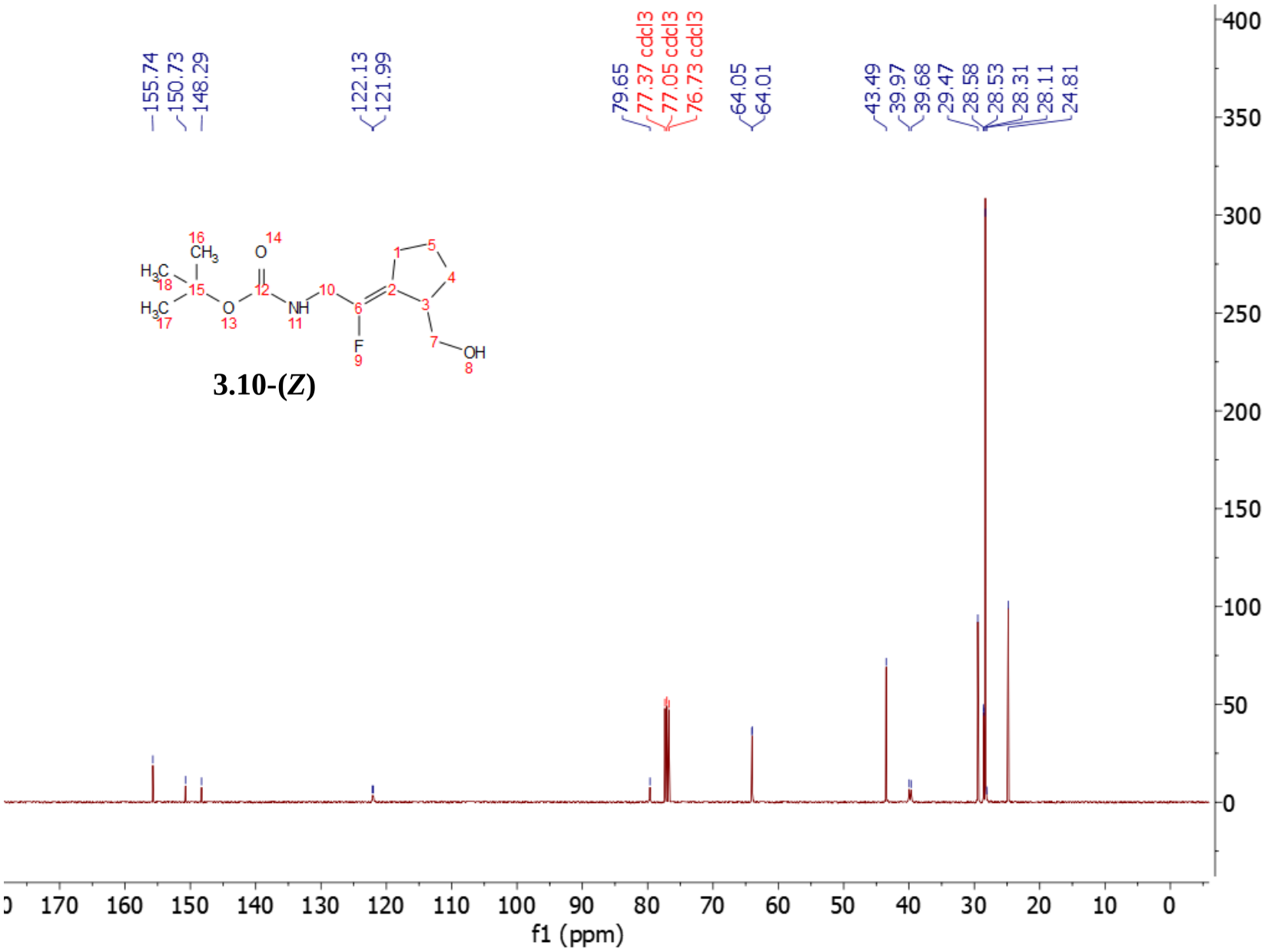

2903.03 CDCl3
1550.35
1530.50
1489.32
1484.97
1479.40
1475.04
1362.79
1354.18
1353.09
1344.44
1172.05
910.70
687.98
572.66
348.28
10.23
7.50
3.9-(Z)
0.79
1.96
1.03
1.00
0.97
1.98
4.16
8.97
8.99
5.84
f1 (ppm)


150.12
148.44
147.77
145.99
134.32
133.57
133.27
131.95
124.56
124.25
124.11
85.41
77.48 cdcl3
77.16 CDCl3
77.16 cdcl3
76.84 cdcl3
63.81
63.77
45.66
45.38
43.88
29.33
28.63
28.58
27.88
26.03
24.77
18.38
-5.21
-5.27
3.9-(Z)
f1 (ppm)

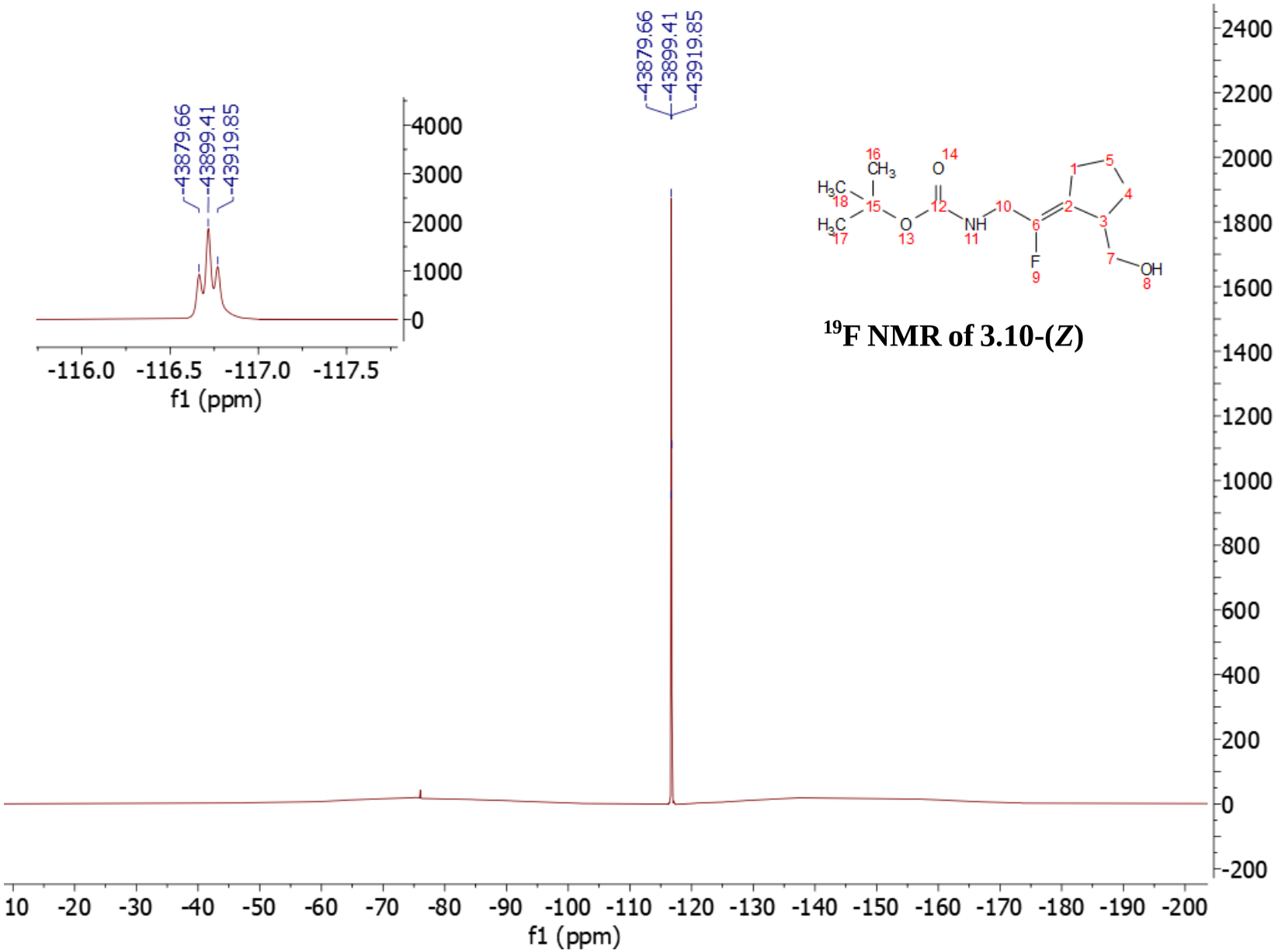


^{19}F NMR of 3.9-(*Z*)

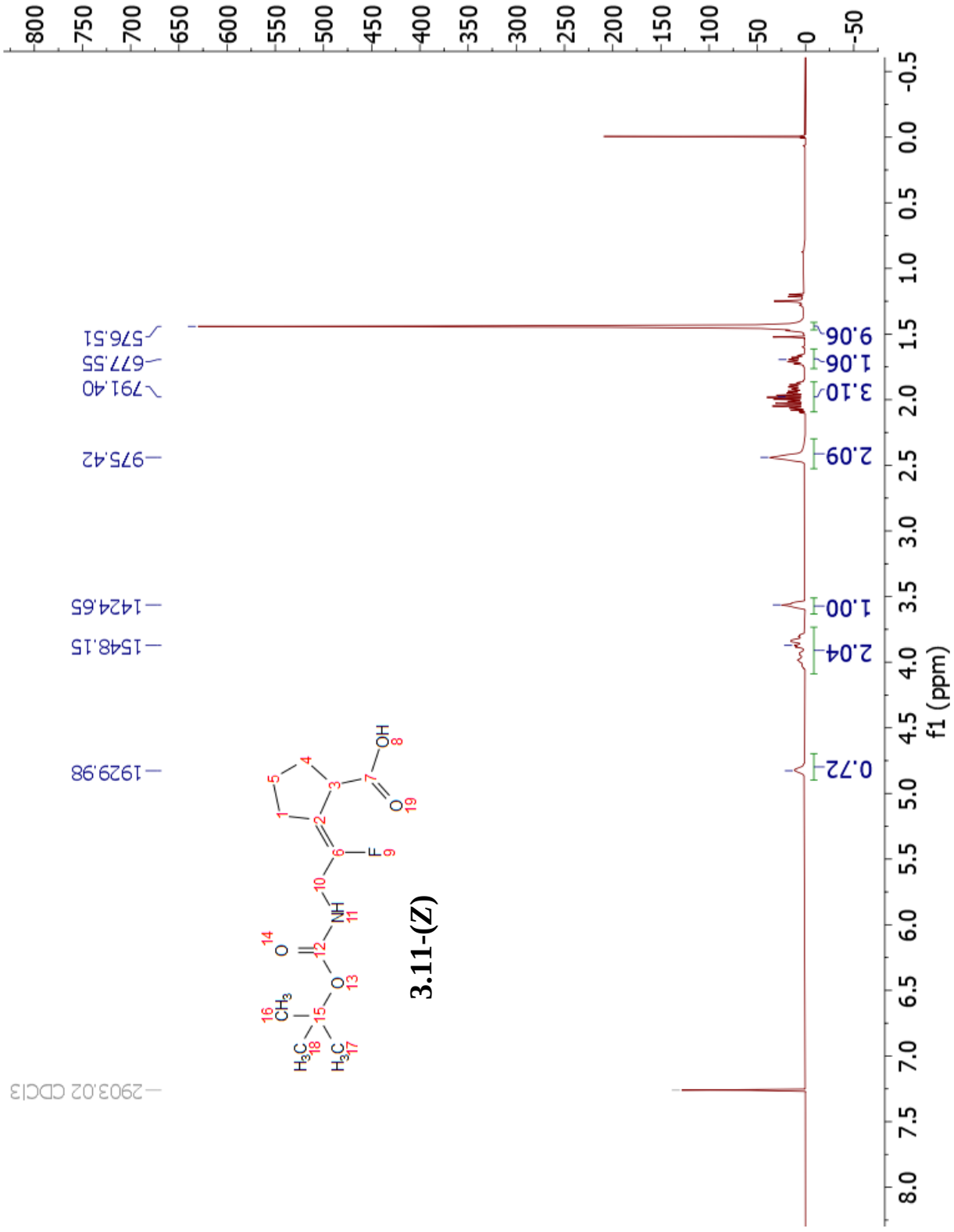

8.98
4.16
1.21
2.04
1.00
0.99
1.01
2.02
0.99
f1 (ppm)
562.22
672.95
871.61
917.91
1176.80
1389.99
1397.68
1400.45
1407.88
1444.10
1449.29
1454.66
1459.76
1523.00
1542.77
1983.94
2902.04 CDCl3
3.10-(Z)

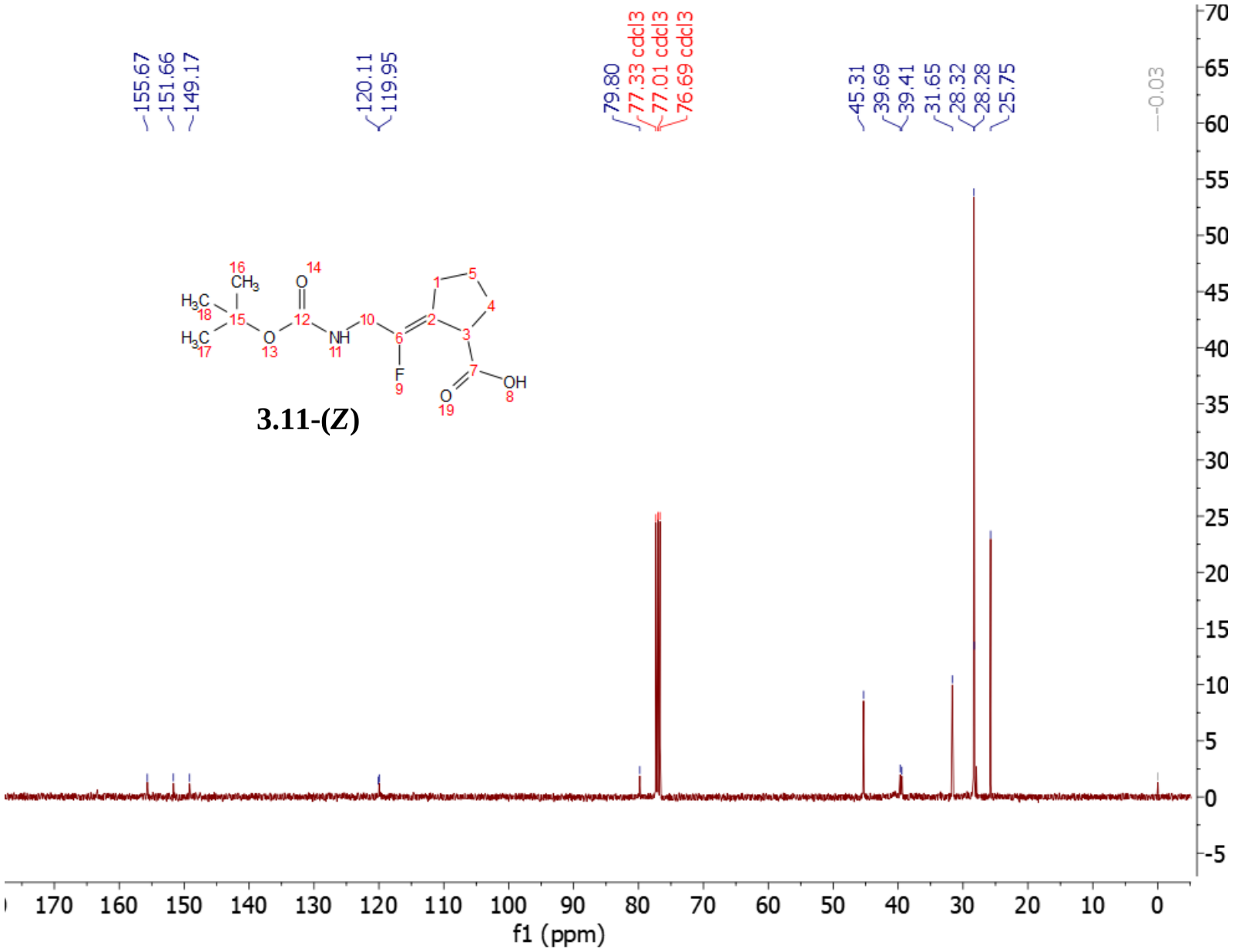

155.74
150.73
148.29
122.13
121.99
79.65
77.37 cdcl3
77.05 cdcl3
76.73 cdcl3
64.05
64.01
43.49
39.97
39.68
29.47
28.58
28.53
28.31
28.11
24.81
f1 (ppm)

3.10-(Z)

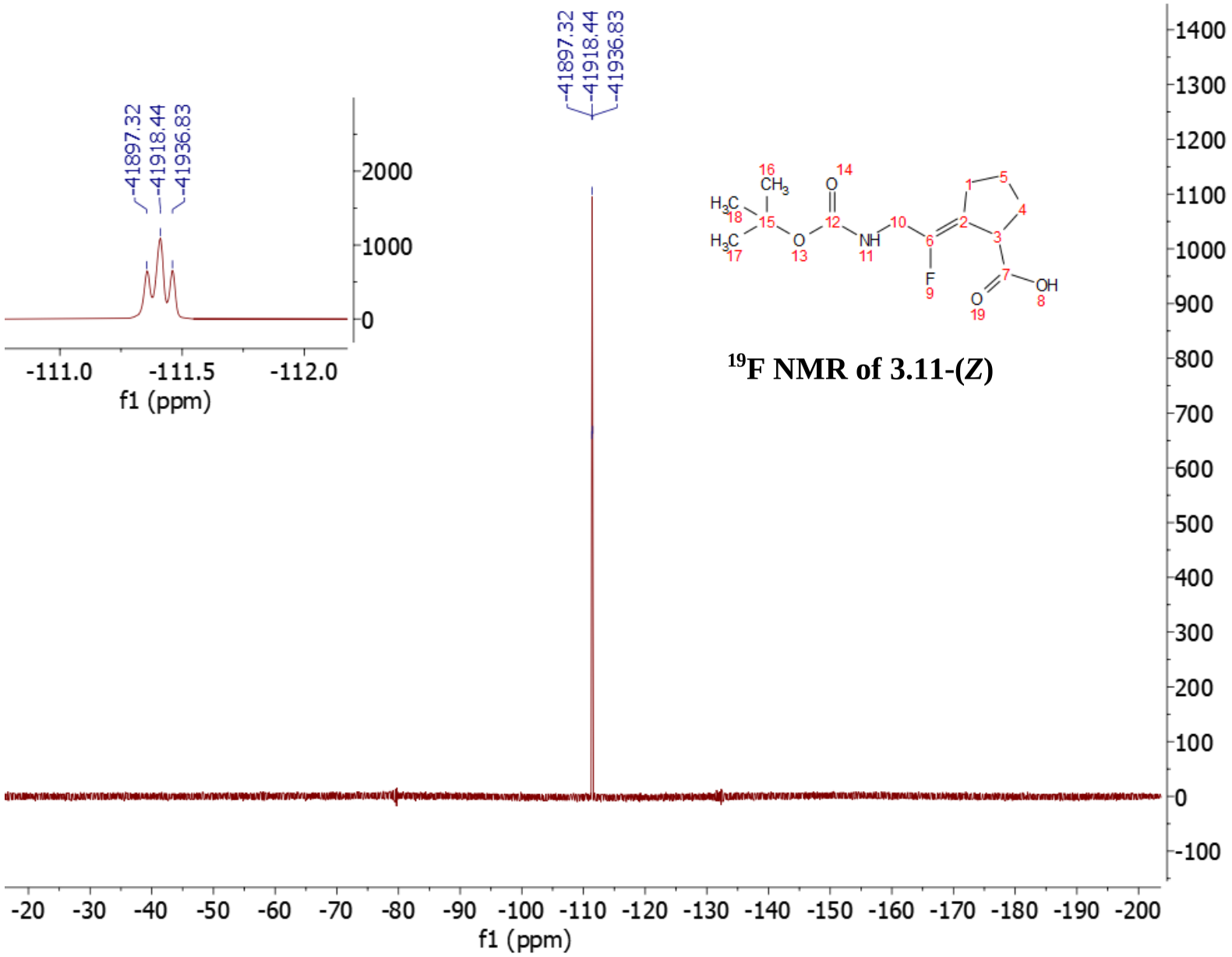


^{19}F NMR of 3.10-(*Z*)

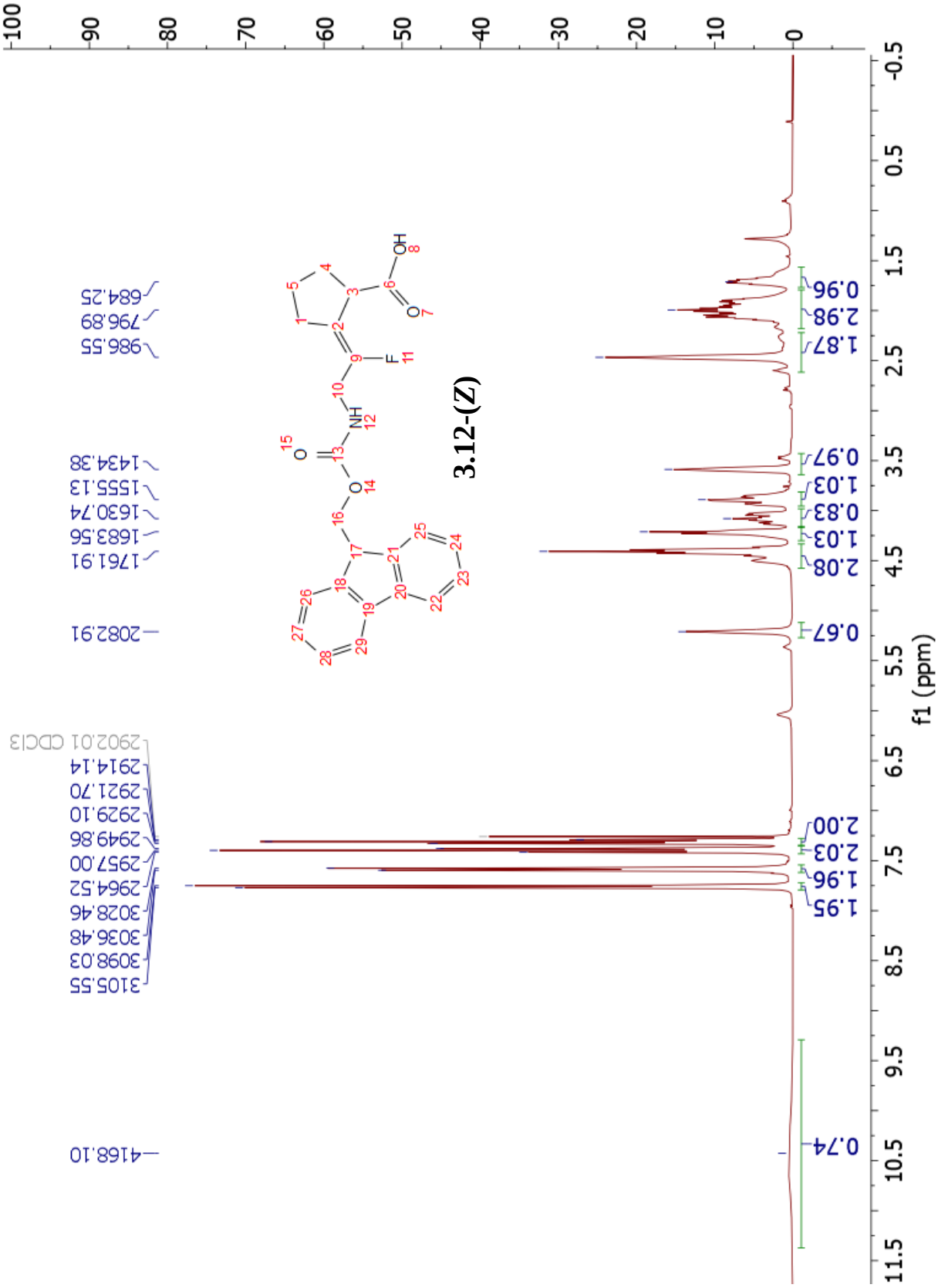

2903.02 CDCl3
1929.98
1548.15
1424.65
975.42
791.40
677.55
576.51
0.72
2.04
1.00
2.09
3.10
1.06
9.06
f1 (ppm)
3.11-(Z)

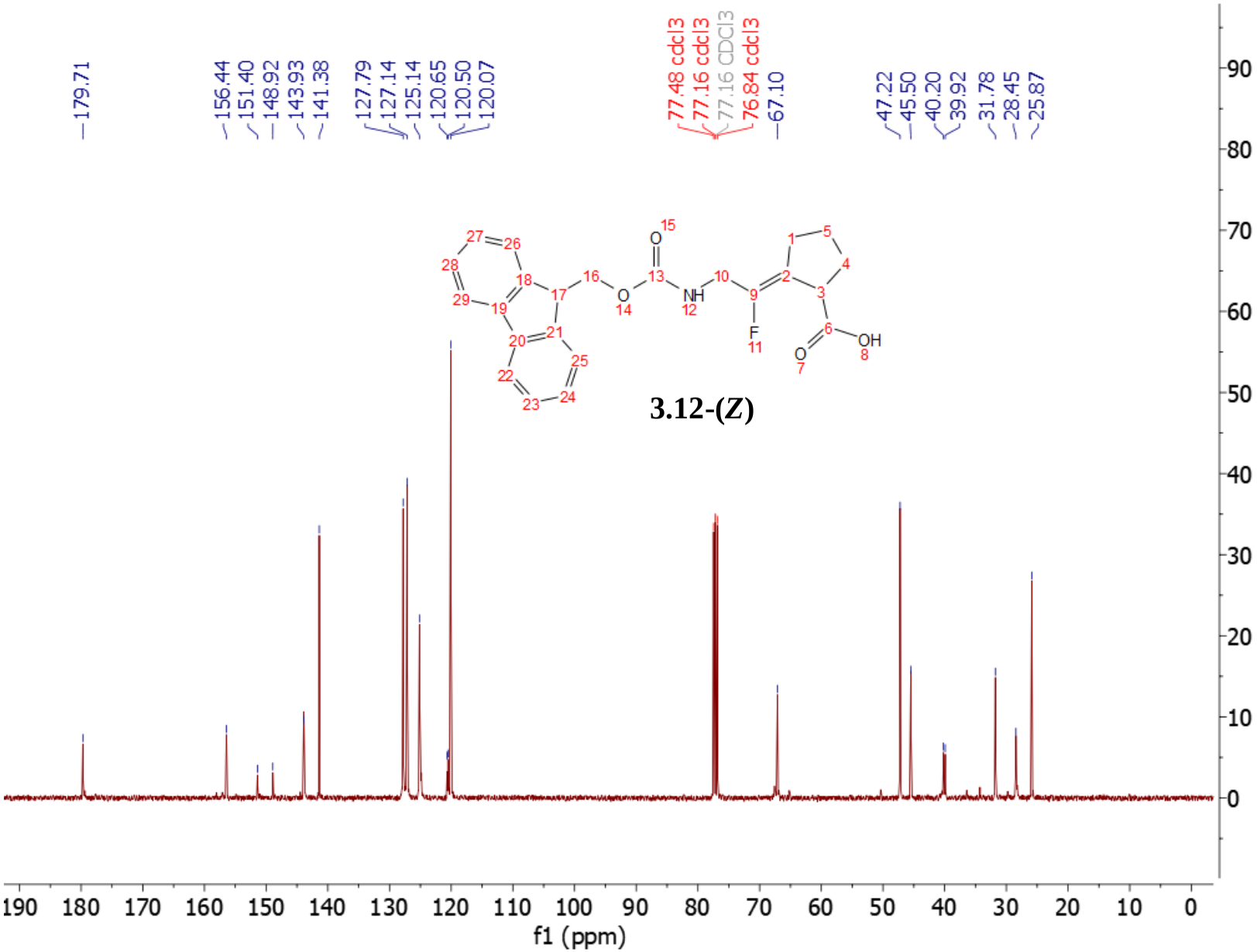

155.67
151.66
149.17
120.11
119.95
79.80
77.33 cdcl3
77.01 cdcl3
76.69 cdcl3
45.31
39.69
39.41
31.65
28.32
28.28
25.75
0.03
f1 (ppm)

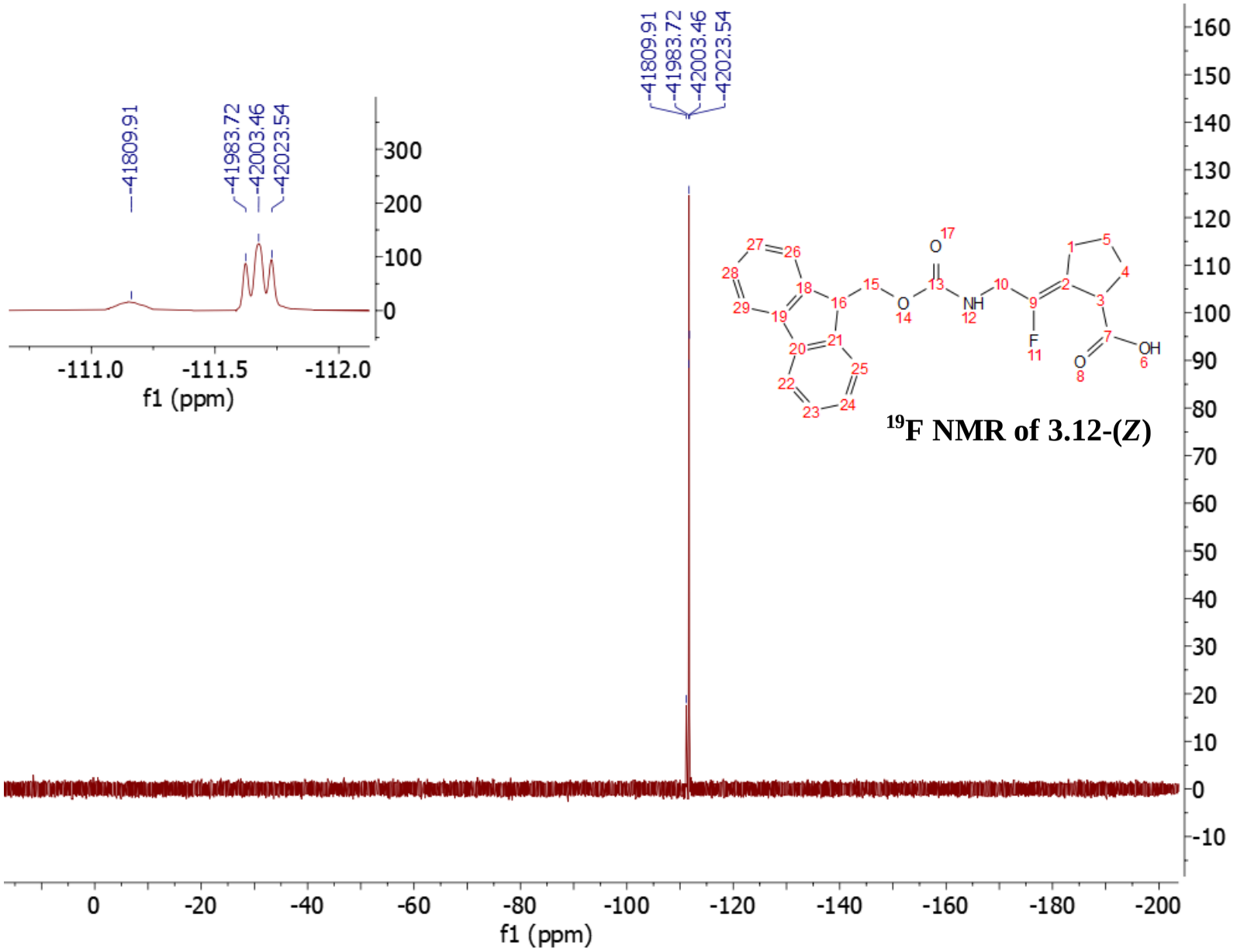
3.11-(Z)



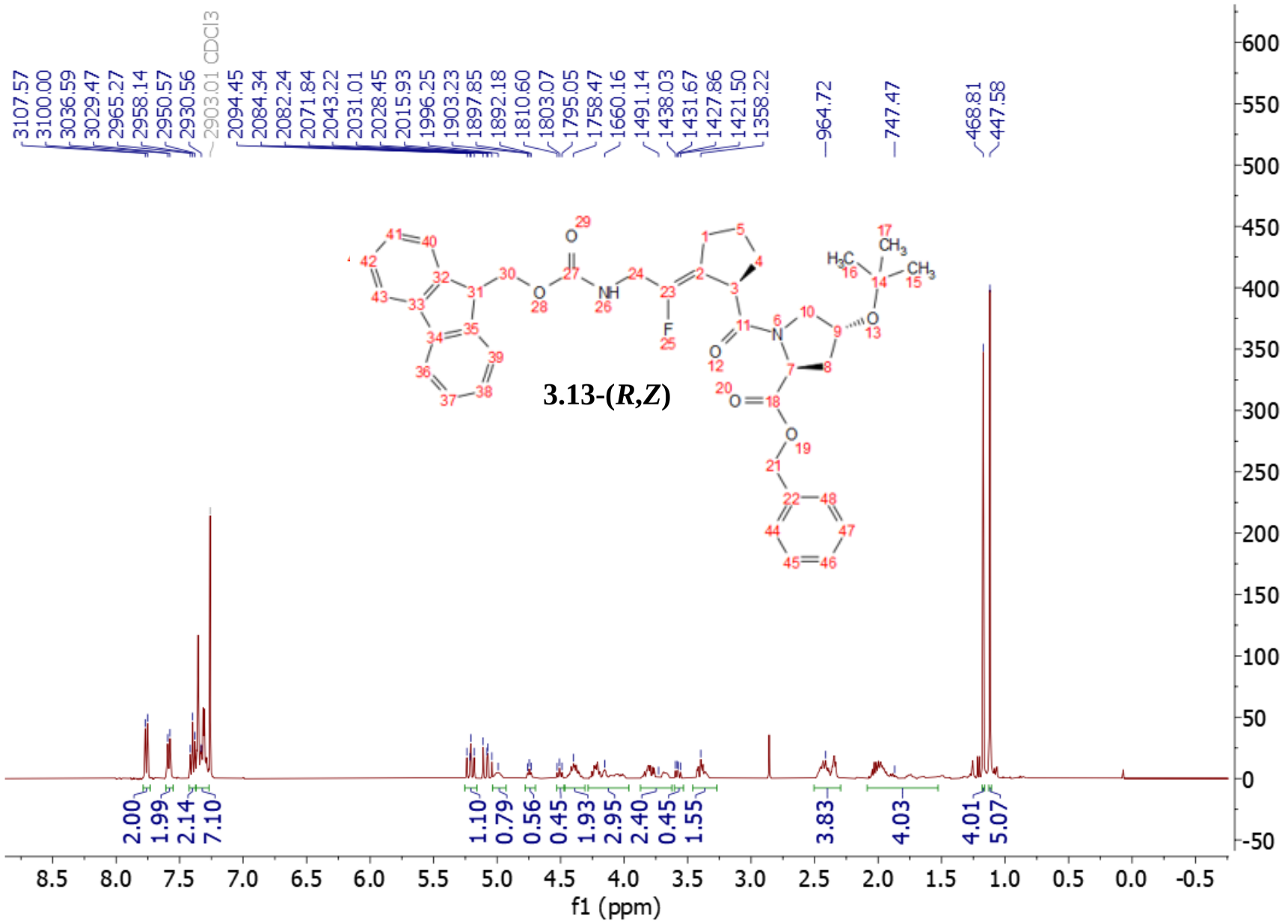
^{19}F NMR of 3.11-(*Z*)

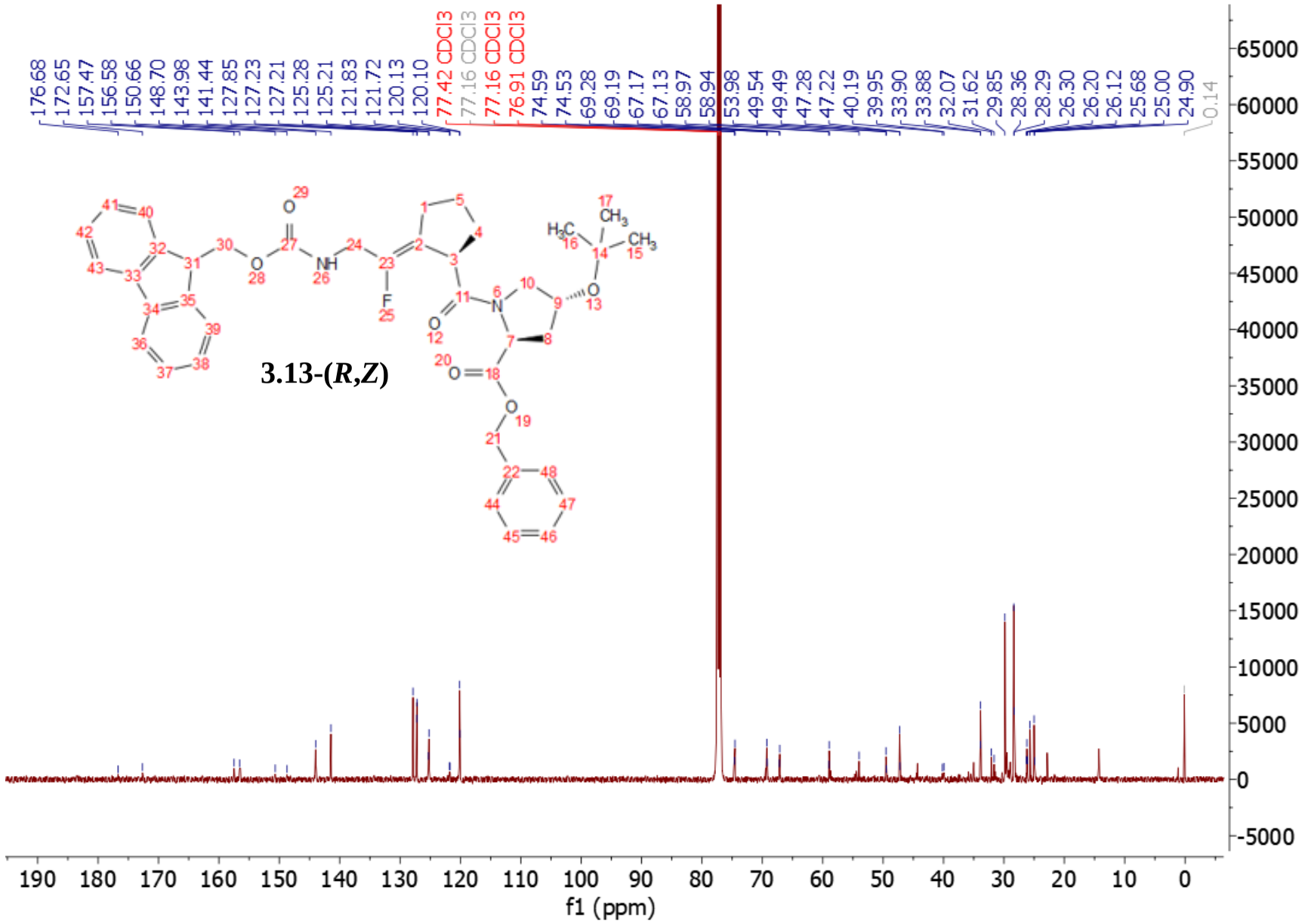

3.12-(Z)
f1 (ppm)
0.74
1.95
1.96
2.03
2.00
0.67
2.08
1.03
0.83
1.03
0.97
1.87
2.98
0.96
4168.10
3105.55
3098.03
3036.48
3028.46
2964.52
2957.00
2949.86
2929.10
2921.70
2914.14
2902.01 CDCl3
2082.91
1761.91
1683.56
1630.74
1555.13
1434.38
986.55
796.89
684.25

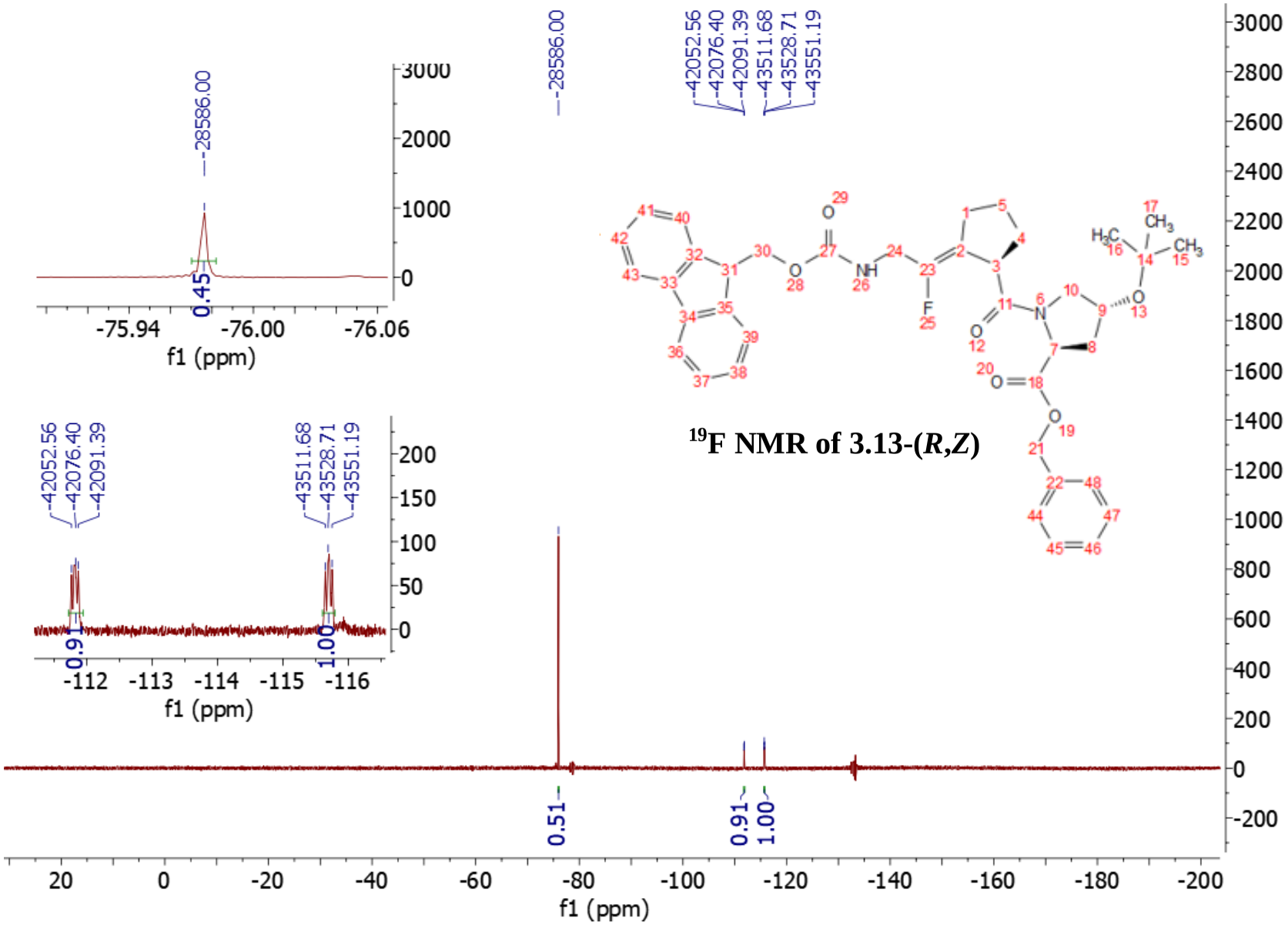

179.71
156.44
151.40
148.92
143.93
141.38
127.79
127.14
125.14
120.65
120.50
120.07
77.48 cdcl3
77.16 cdcl3
77.16 CDCl3
76.84 cdcl3
67.10
47.22
45.50
40.20
39.92
31.78
28.45
25.87
3.12-(Z)
f1 (ppm)



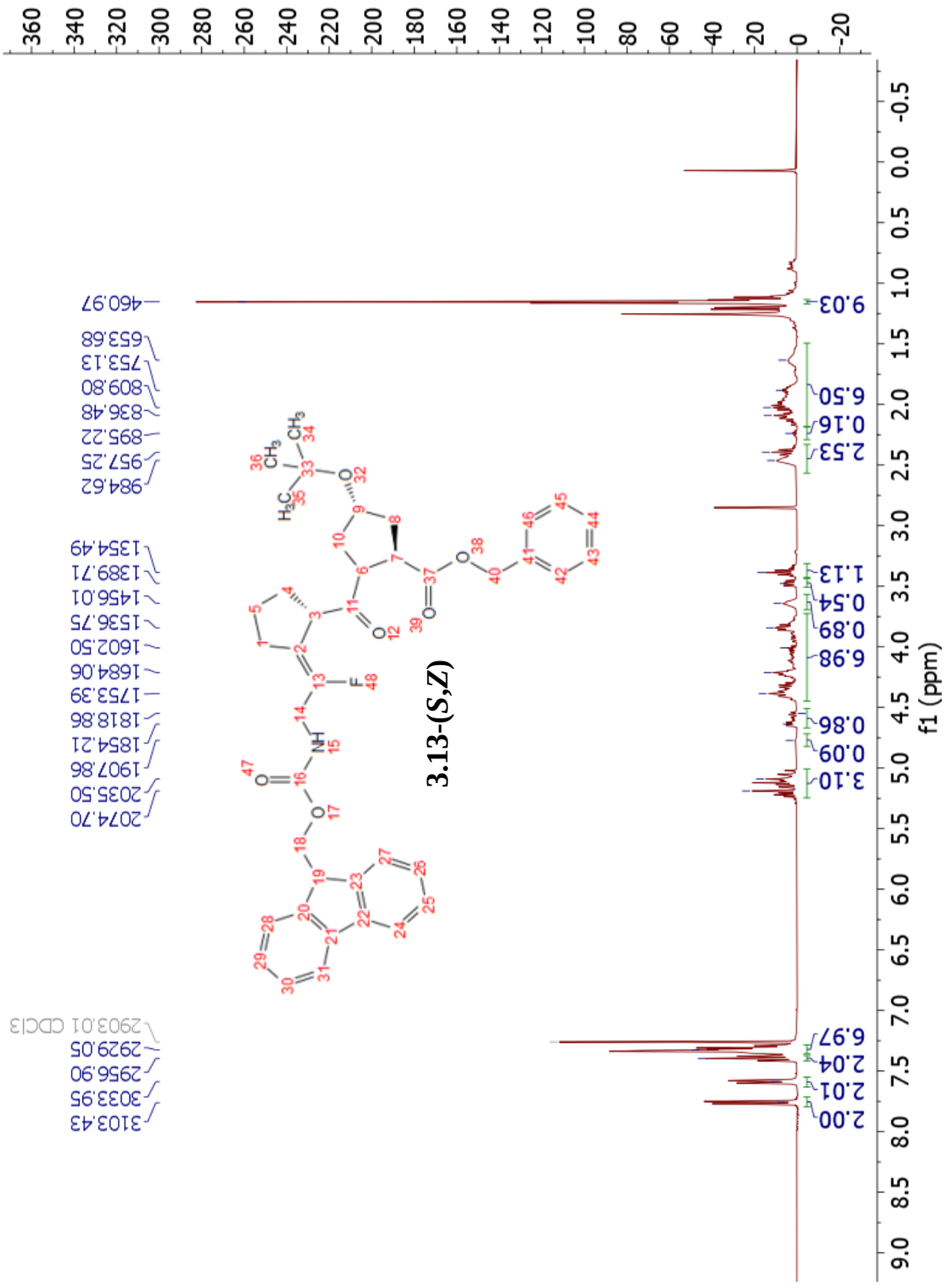
^{19}F NMR of 3.12-(*Z*)

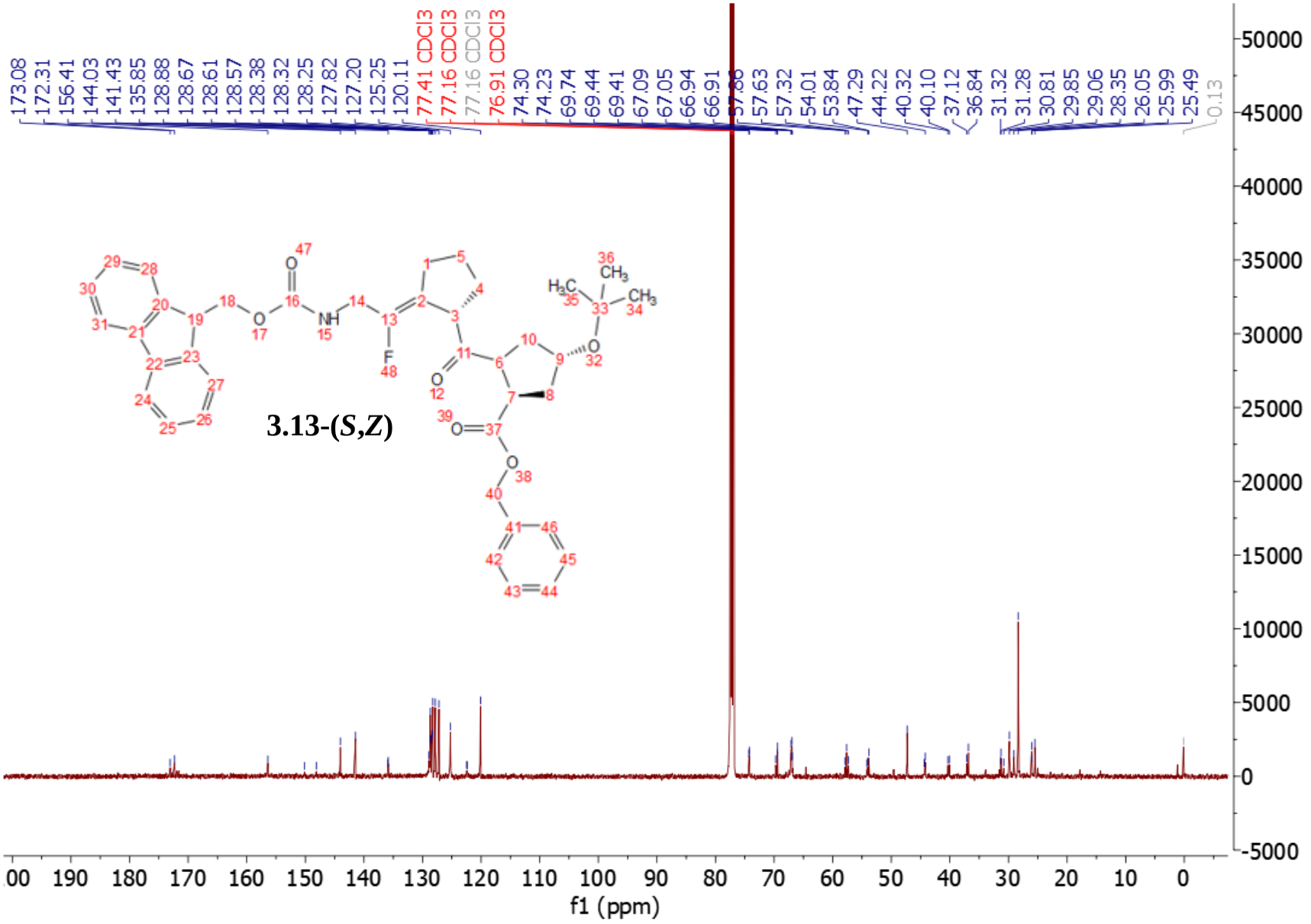

3107.57
3100.00
3036.59
3029.47
2965.27
2958.14
2950.57
2930.56
2903.01 CDCl3
2094.45
2084.34
2082.24
2071.84
2043.22
2031.01
2028.45
2015.93
1996.25
1903.23
1897.85
1892.18
1810.60
1803.07
1795.05
1758.47
1660.16
1491.14
1438.03
1431.67
1427.86
1421.50
1358.22
964.72
747.47
468.81
447.58
3.13-(R,Z)
2.00
1.99
2.14
7.10
1.10
0.79
0.56
0.45
1.93
2.95
2.40
0.45
1.55
3.83
4.03
4.01
5.07
f1 (ppm)


176.68
172.65
157.47
156.58
150.66
148.70
143.98
141.44
127.85
127.23
127.21
125.28
125.21
121.83
121.72
120.13
120.10
77.42 CDCl3
77.16 CDCl3
77.16 CDCl3
76.91 CDCl3
74.59
74.53
69.28
69.19
67.17
67.13
58.97
58.94
53.98
49.54
49.49
47.28
47.22
40.19
39.95
33.90
33.88
32.07
31.62
29.85
28.36
28.29
26.30
26.20
26.12
25.68
25.00
24.90
0.14
3.13-(R,Z)
65000
60000
55000
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

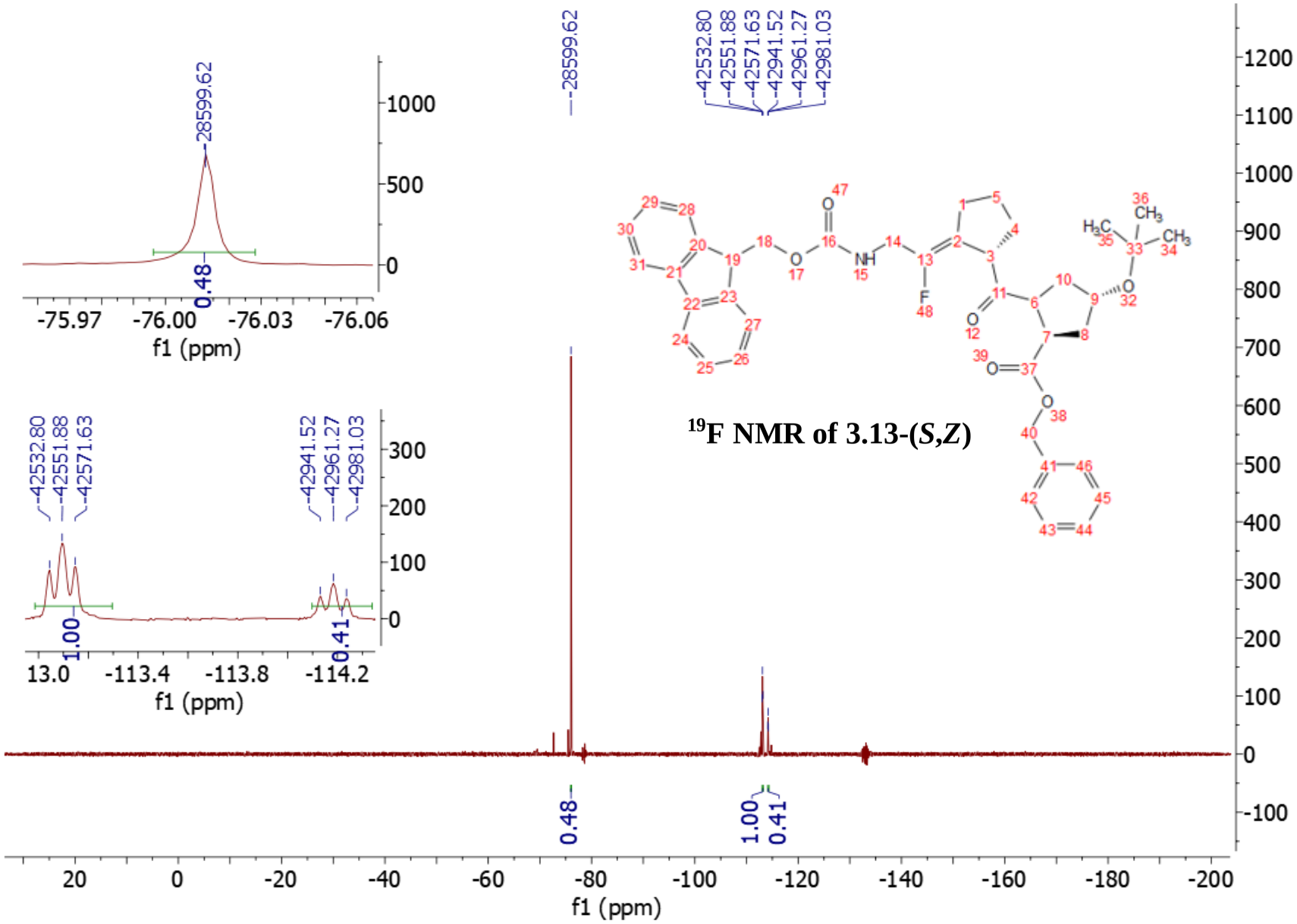


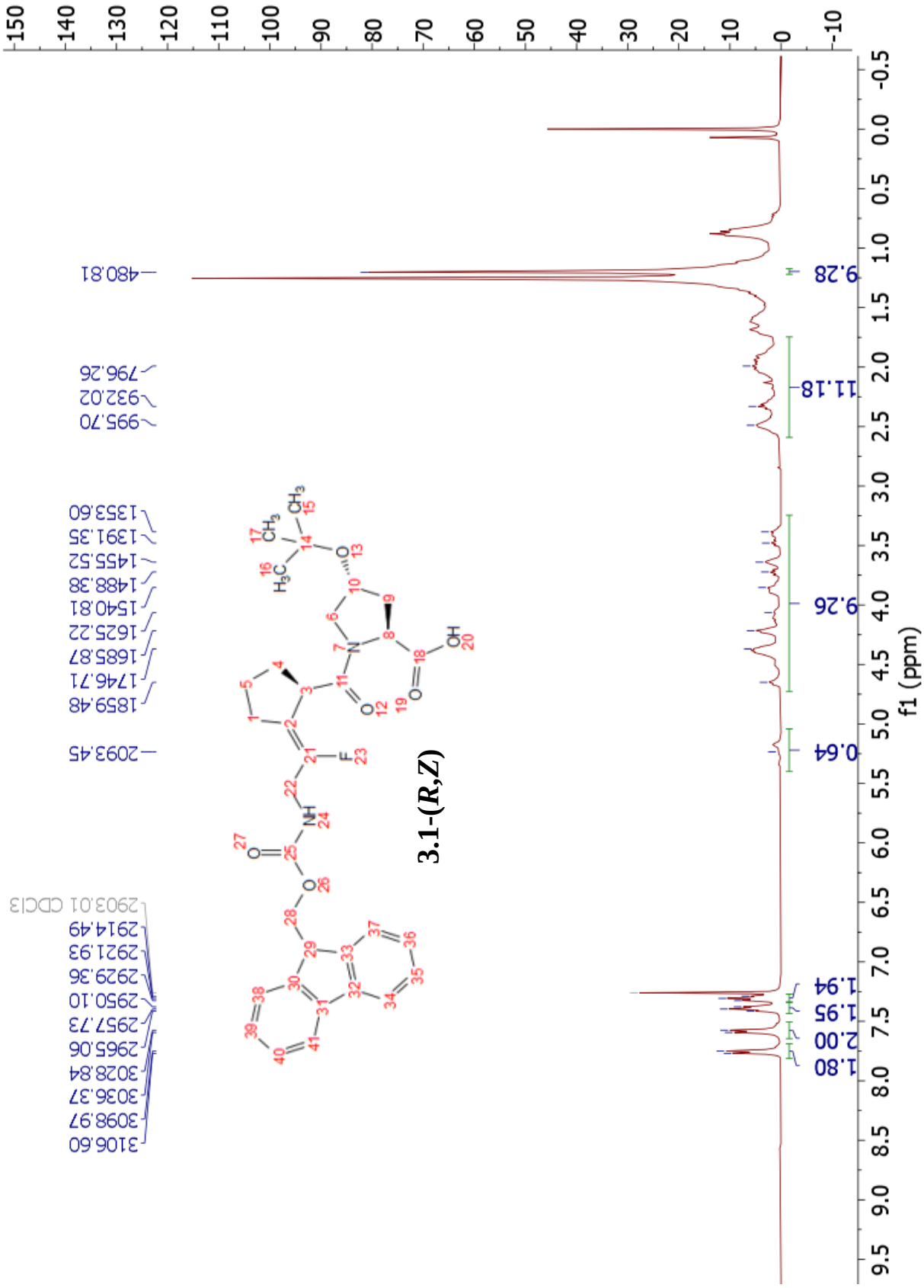
^{19}F NMR of 3.13-(*R*,*Z*)

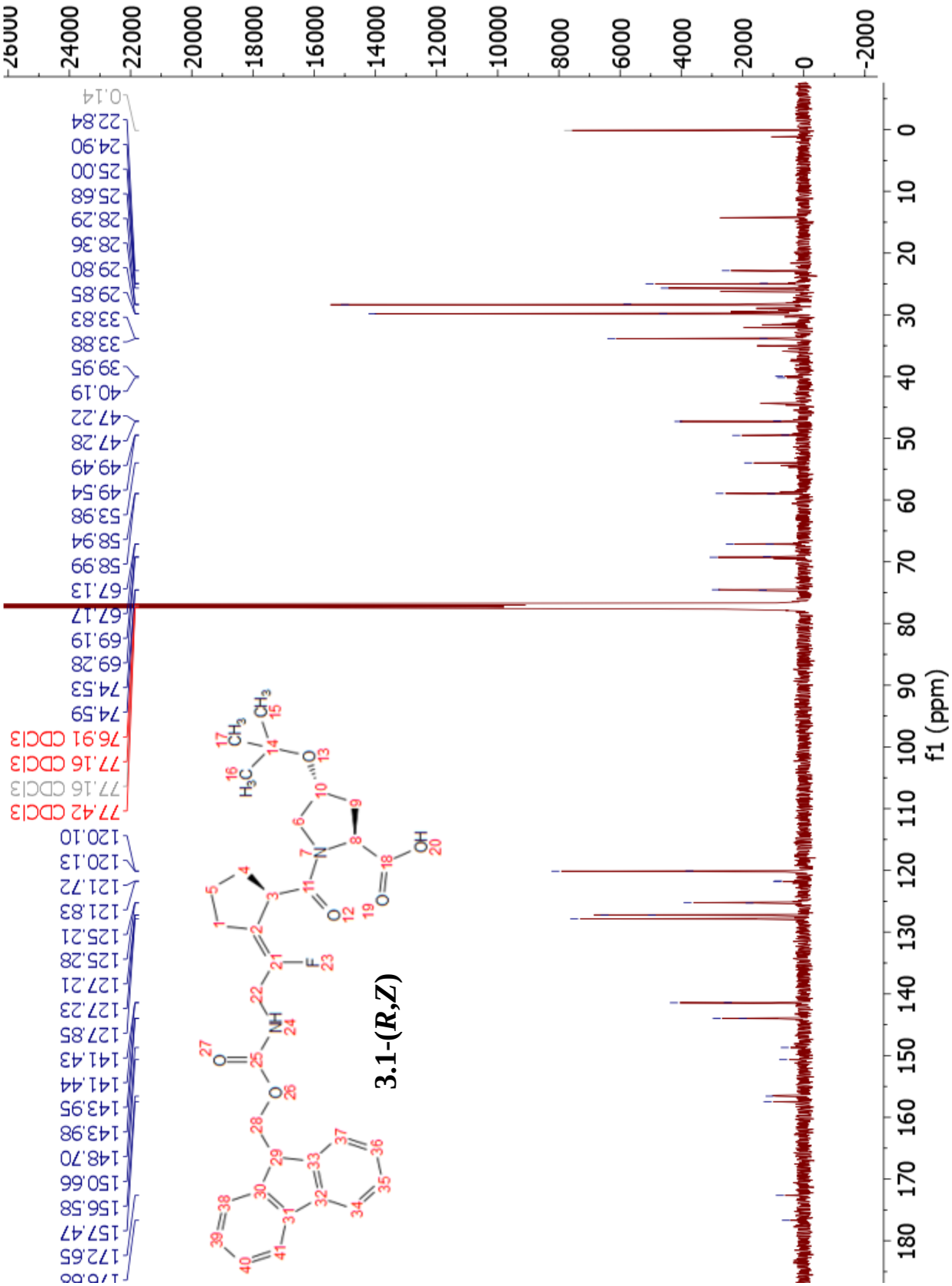

3.13-(S,Z)
f1 (ppm)
9.03
6.50
0.16
2.53
1.13
0.54
0.89
6.98
0.86
0.09
3.10
2.00
2.01
2.04
6.97
460.97
653.68
753.13
809.80
836.48
895.22
957.25
984.62
1354.49
1389.71
1456.01
1536.75
1602.50
1684.06
1753.39
1818.86
1854.21
1907.86
2035.50
2074.70
2903.01 CDCl3
2929.05
2956.90
3033.95
3103.43

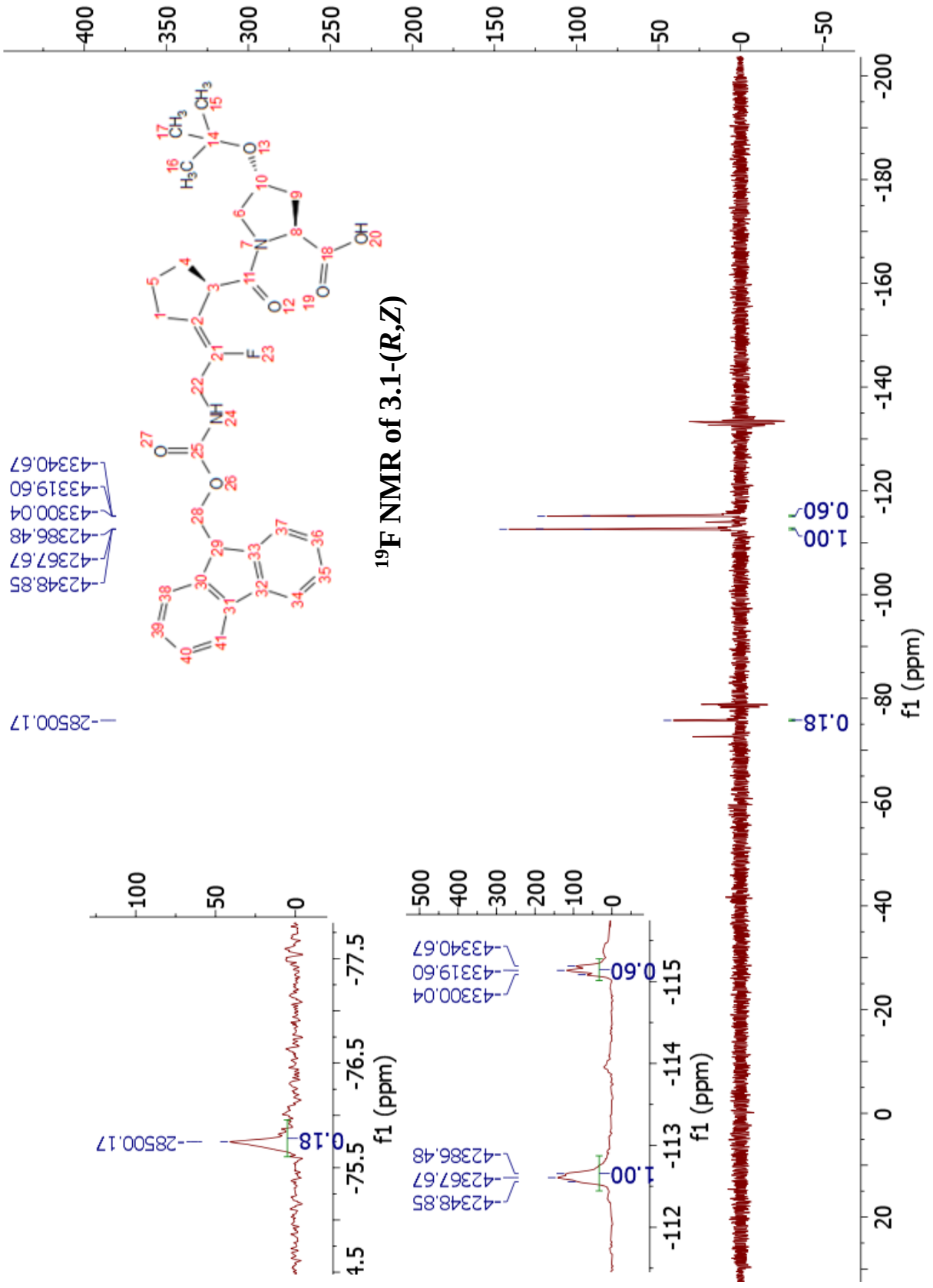

173.08
172.31
156.41
144.03
141.43
135.85
128.88
128.67
128.61
128.57
128.38
128.32
128.25
127.82
127.20
125.25
120.11
77.41 CDCl3
77.16 CDCl3
77.16 CDCl3
76.91 CDCl3
74.30
74.23
69.74
69.44
69.41
67.09
67.05
66.94
66.91
57.86
57.63
57.32
54.01
53.84
47.29
44.22
40.32
40.10
37.12
36.84
31.32
31.28
30.81
29.85
29.06
28.35
26.05
25.99
25.49
0.13
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
00 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

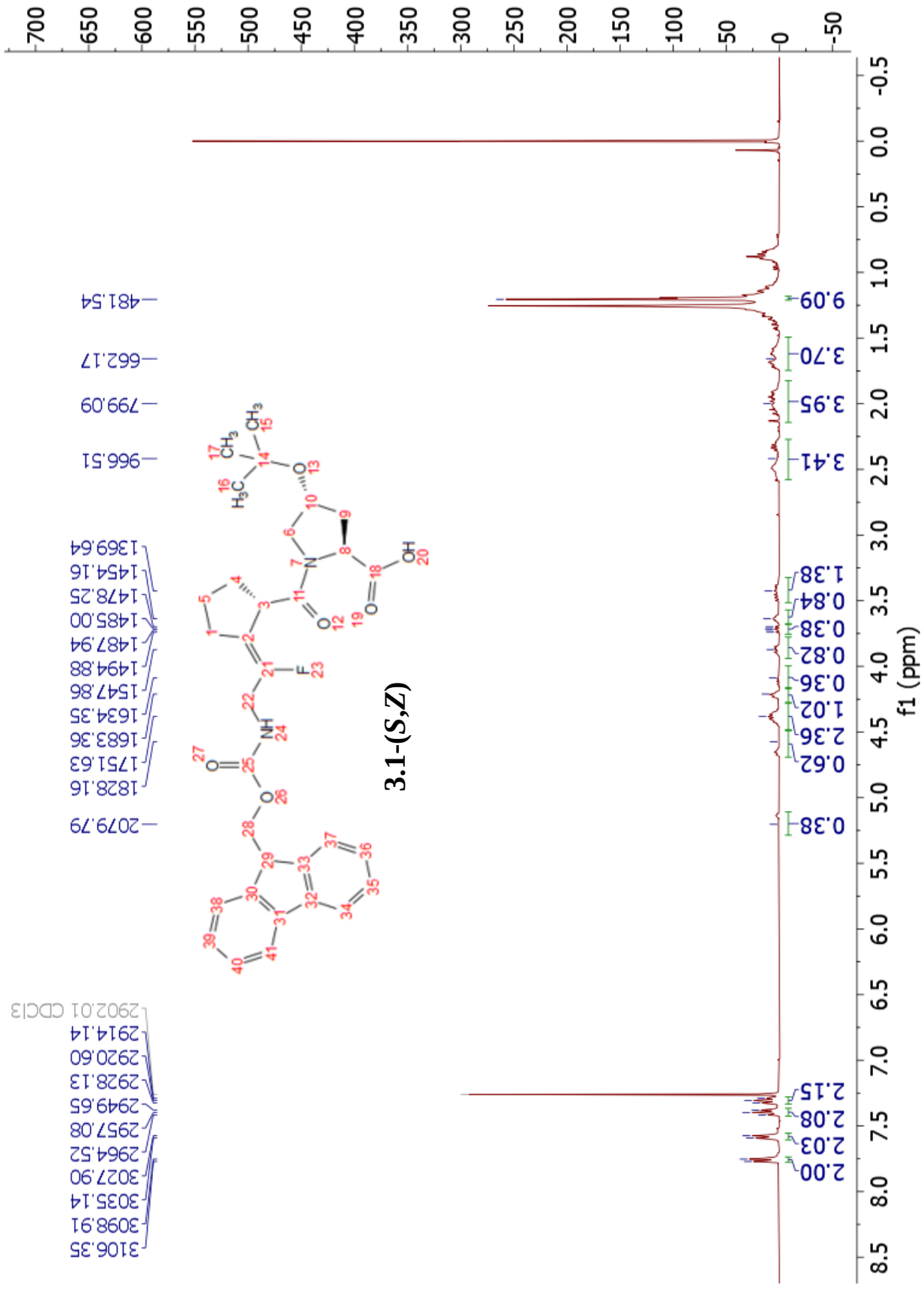
3.13-(*S*,*Z*)

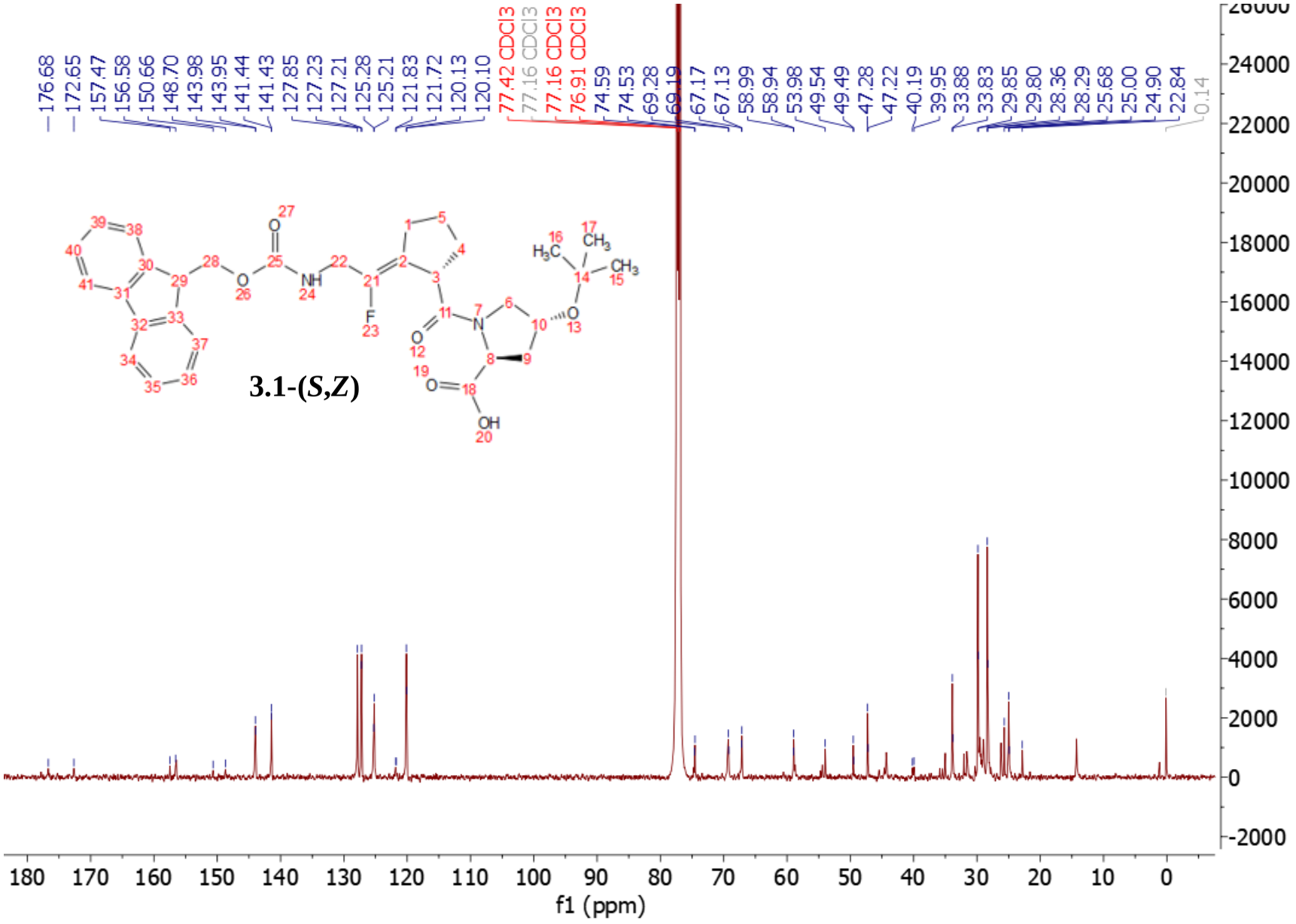

-28599.62
-42532.80
-42551.88
-42571.63
-42941.52
-42961.27
-42981.03
0.48
1.00
0.41
f1 (ppm)
19F NMR of 3.13-(S,Z)

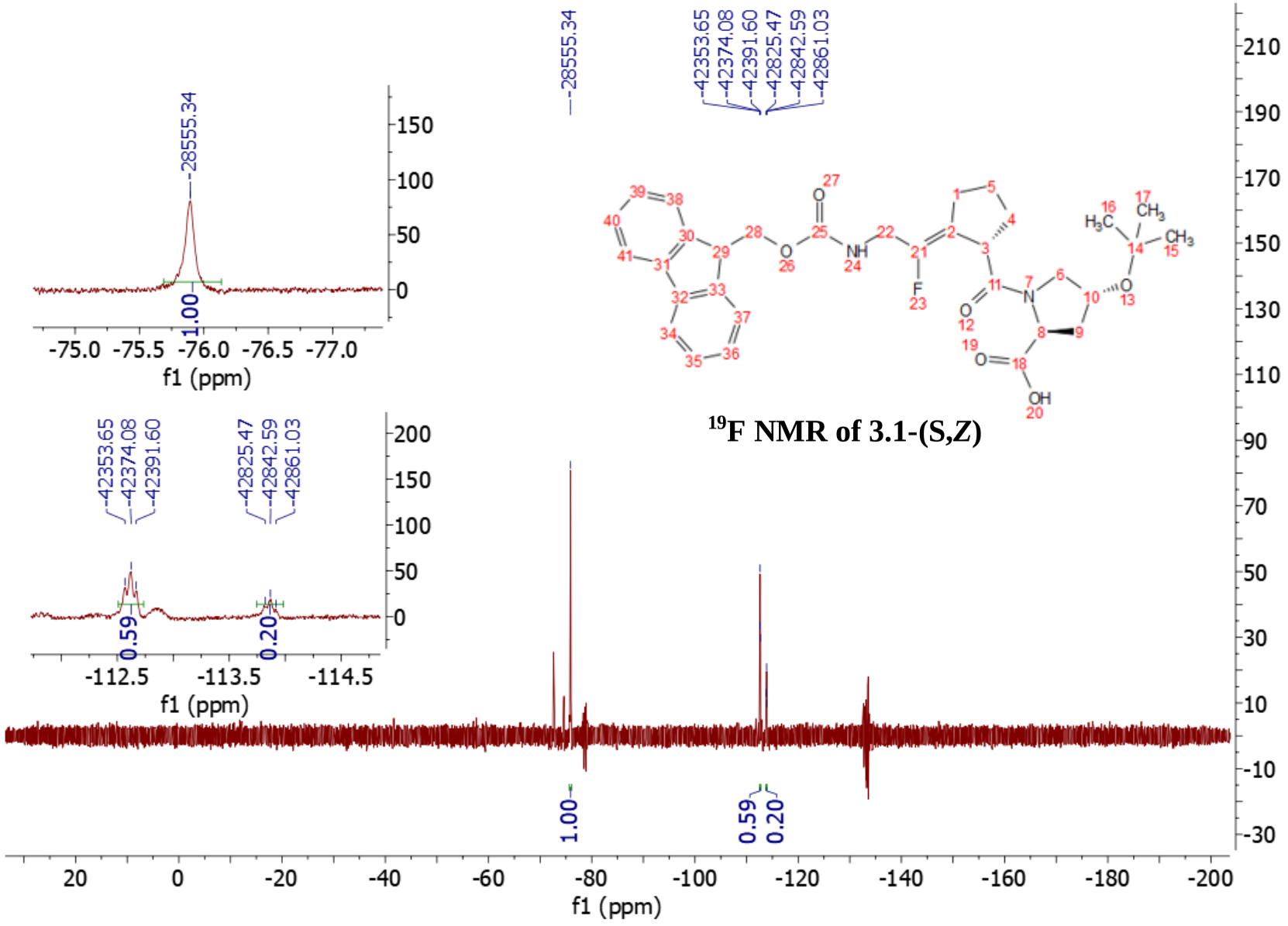

3.1-(R,Z)
f1 (ppm)
9.28
11.18
9.26
0.64
1.94
1.95
2.00
1.80
480.81
995.70
932.02
796.26
1553.60
1391.35
1455.52
1488.38
1540.81
1625.22
1685.87
1746.71
1859.48
2093.45
3106.60
3098.97
3036.37
3028.84
2965.06
2957.73
2950.10
2929.36
2921.93
2914.49
2903.01 CDCl3


f1 (ppm)
3.1-(R,Z)

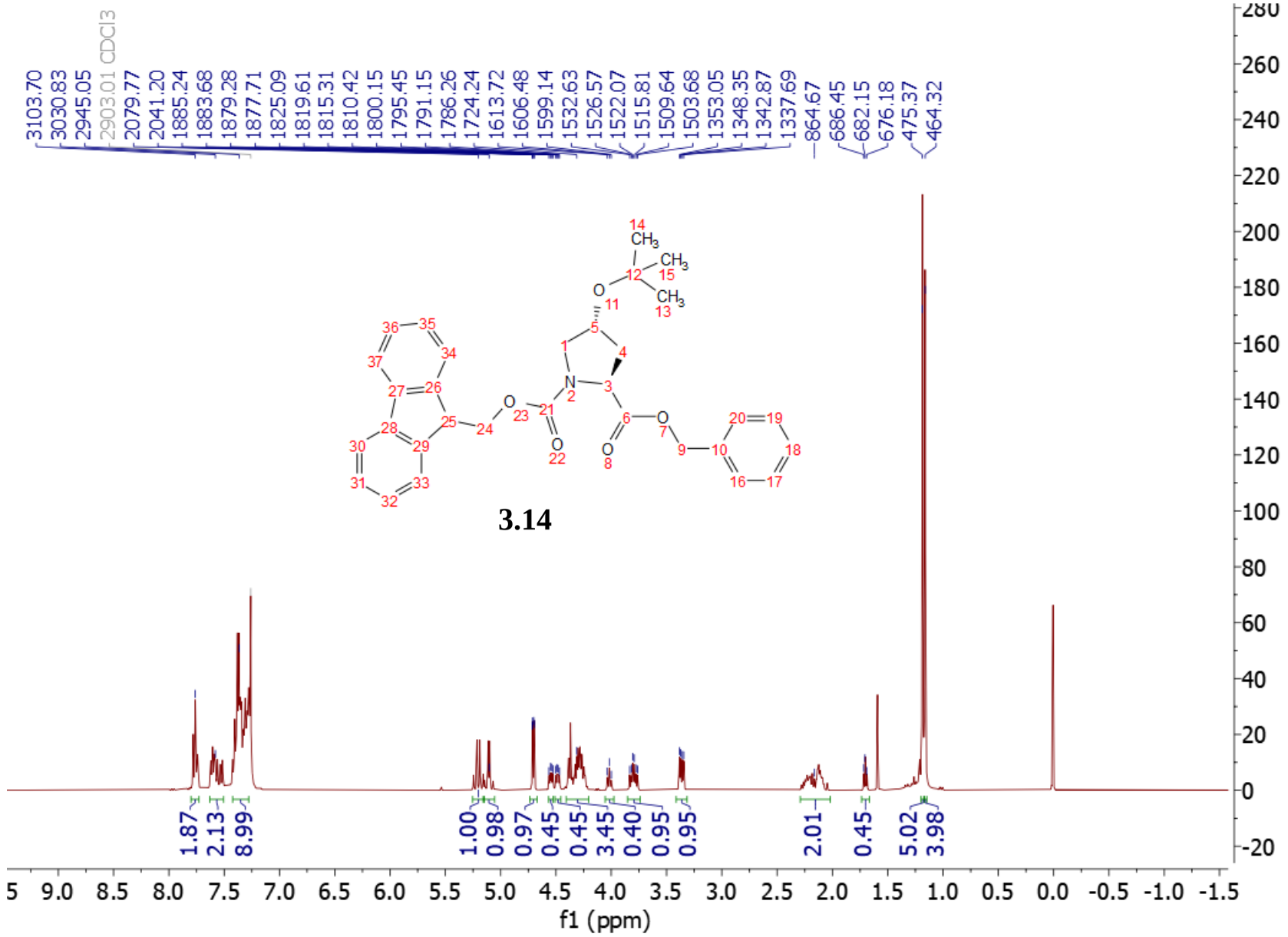

19F NMR of 3.1-(R,Z)
f1 (ppm)
0.18
0.60
1.00
-28500.17
-42348.85
-42367.67
-42386.48
-43300.04
-43319.60
-43340.67

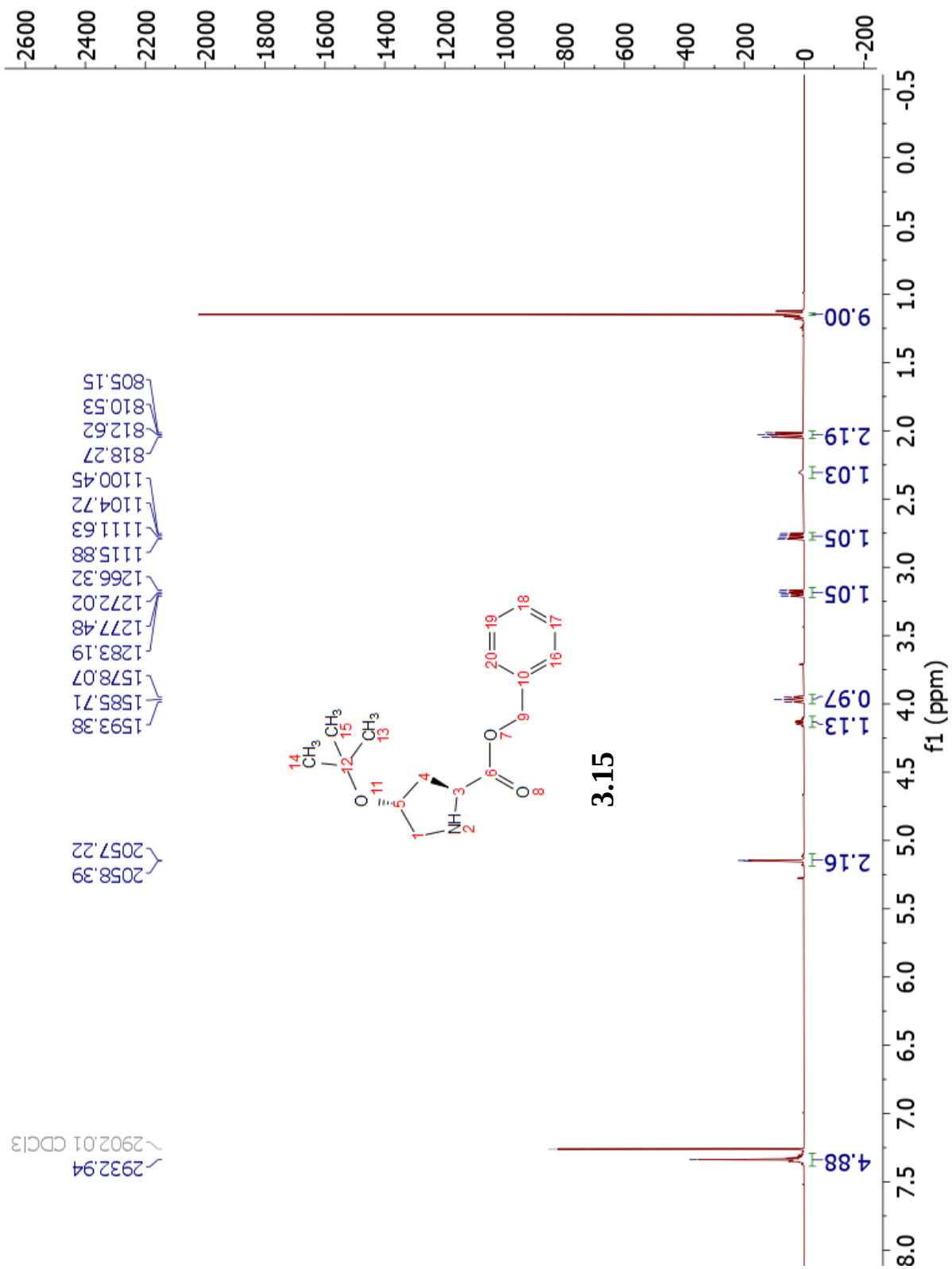

3.1-(S,Z)
f1 (ppm)
9.09
3.70
3.95
3.41
1.38
0.84
0.38
0.82
0.36
1.02
2.36
0.62
0.38
2.15
2.08
2.03
2.00
481.54
662.17
799.09
966.51
1369.64
1454.16
1478.25
1485.00
1487.94
1494.88
1547.86
1634.35
1683.36
1751.63
1828.16
2079.79
2902.01 CDCl3
2914.14
2920.60
2928.13
2949.65
2957.08
2964.52
3027.90
3035.14
3098.91
3106.35

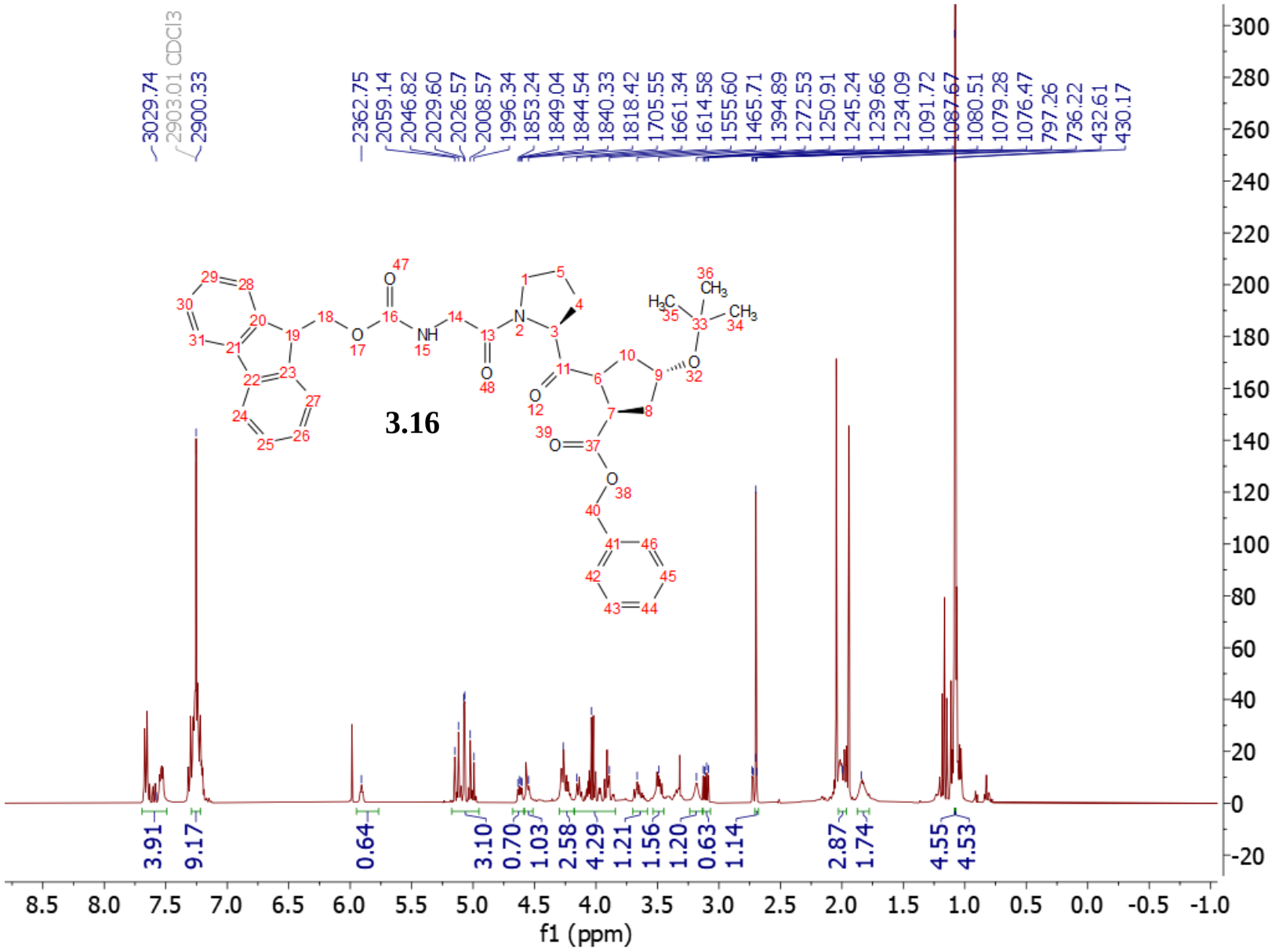

176.68
172.65
157.47
156.58
150.66
148.70
143.98
143.95
141.44
141.43
127.85
127.23
127.21
125.28
125.21
121.83
121.72
120.13
120.10
77.42 CDCl3
77.16 CDCl3
77.16 CDCl3
76.91 CDCl3
74.59
74.53
69.28
69.19
67.17
67.13
58.99
58.94
53.98
49.54
49.49
47.28
47.22
40.19
39.95
33.88
33.83
29.85
29.80
28.36
28.29
25.68
25.00
24.90
22.84
0.14
3.1-(S,Z)
f1 (ppm)

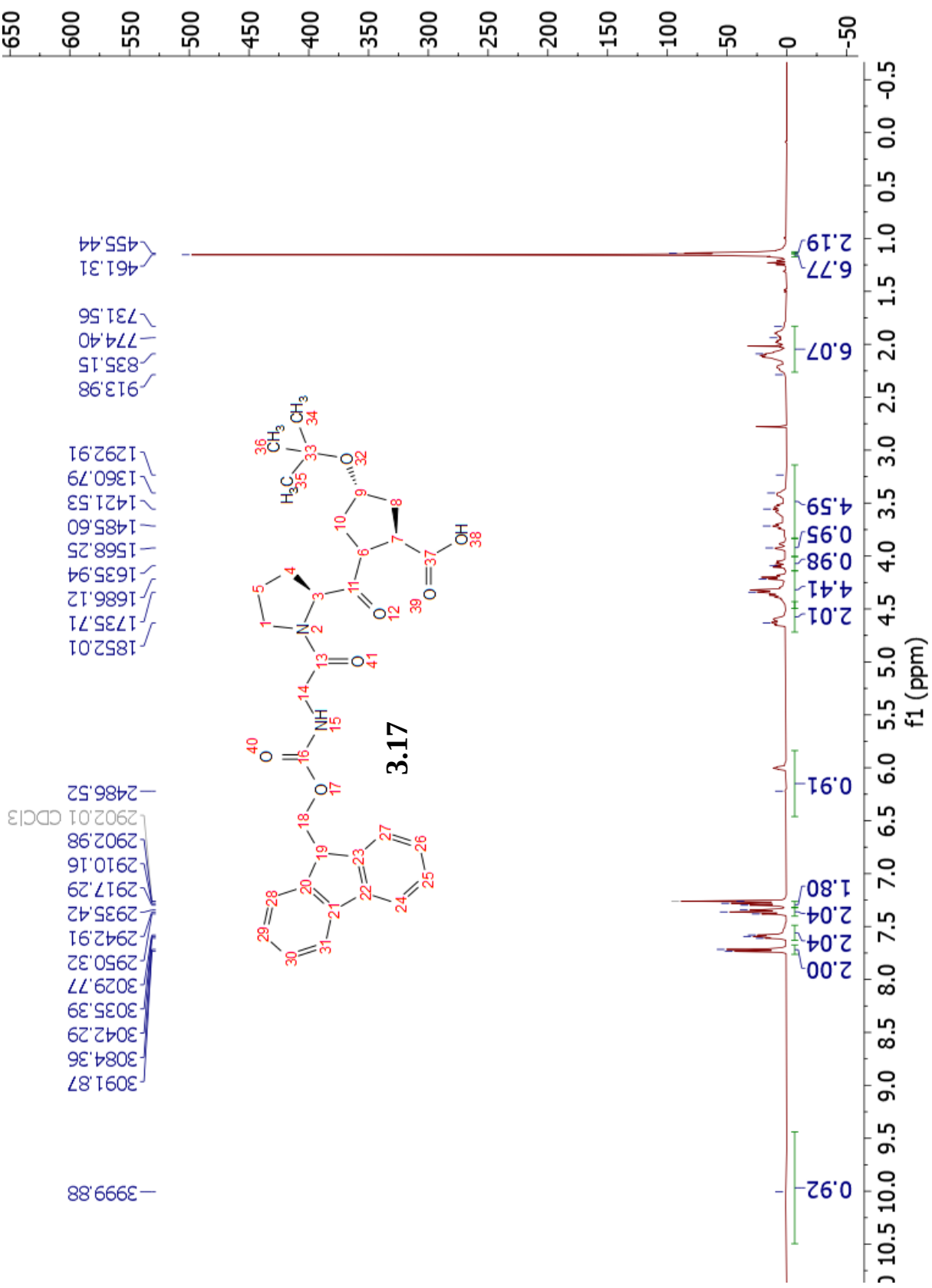


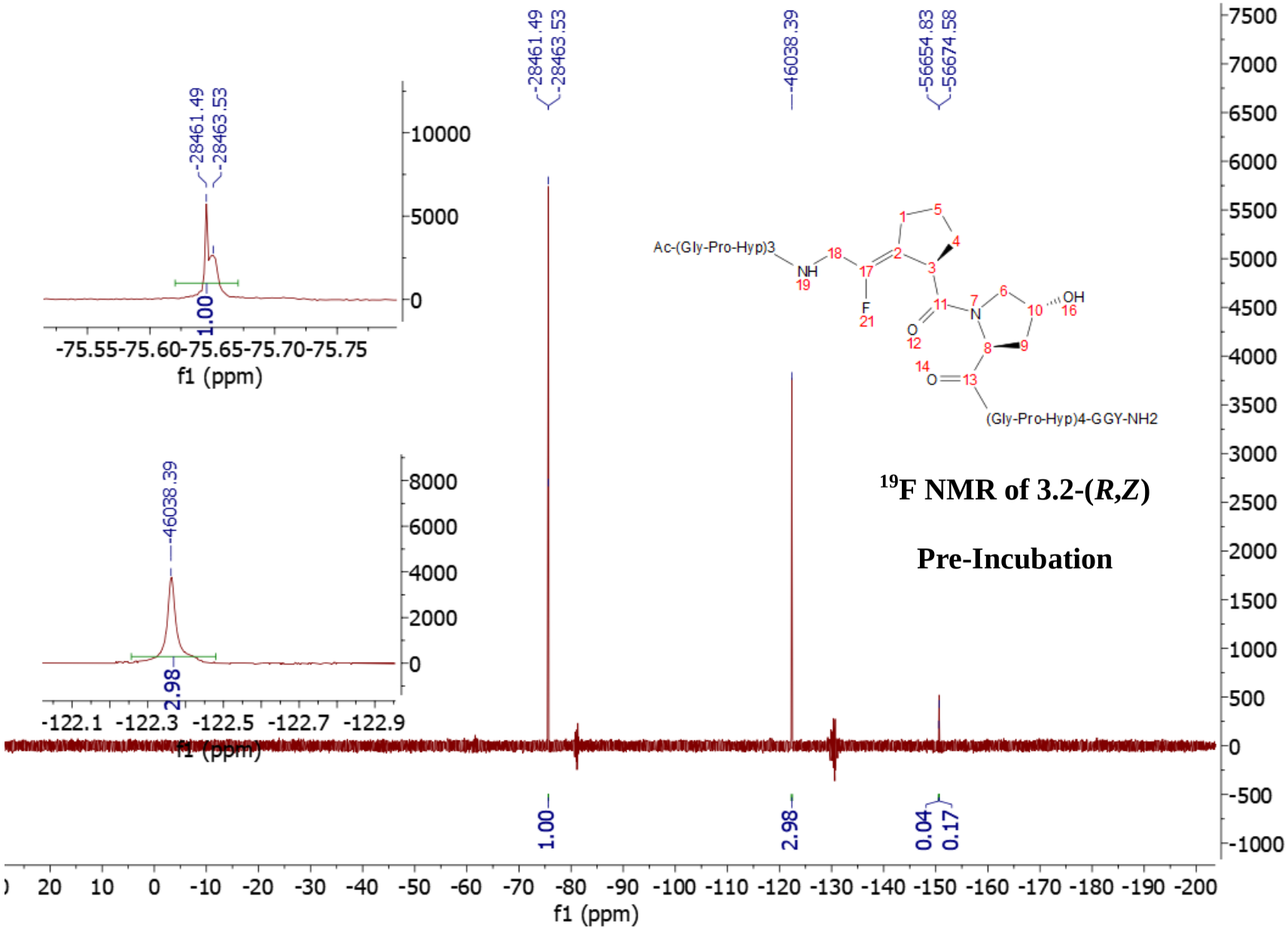
^{19}F NMR of 3.1-(S,*Z*)


3103.70
3030.83
2945.05
2903.01 CDCl3
2079.77
2041.20
1885.24
1883.68
1879.28
1877.71
1825.09
1819.61
1815.31
1810.42
1800.15
1795.45
1791.15
1786.26
1724.24
1613.72
1606.48
1599.14
1532.63
1526.57
1522.07
1515.81
1509.64
1503.68
1353.05
1348.35
1342.87
1337.69
864.67
686.45
682.15
676.18
475.37
464.32
3.14
1.87
2.13
8.99
1.00
0.98
0.97
0.45
0.45
3.45
0.40
0.95
0.95
2.01
0.45
5.02
3.98
f1 (ppm)


3.15
f1 (ppm)
9.00
2.19
1.03
1.05
1.05
0.97
1.13
2.16
4.88
2932.94
2902.01 CDCl3
2058.39
2057.22
1593.38
1585.71
1578.07
1283.19
1277.48
1272.02
1266.32
1115.88
1111.63
1104.72
1100.45
818.27
812.62
810.53
805.15

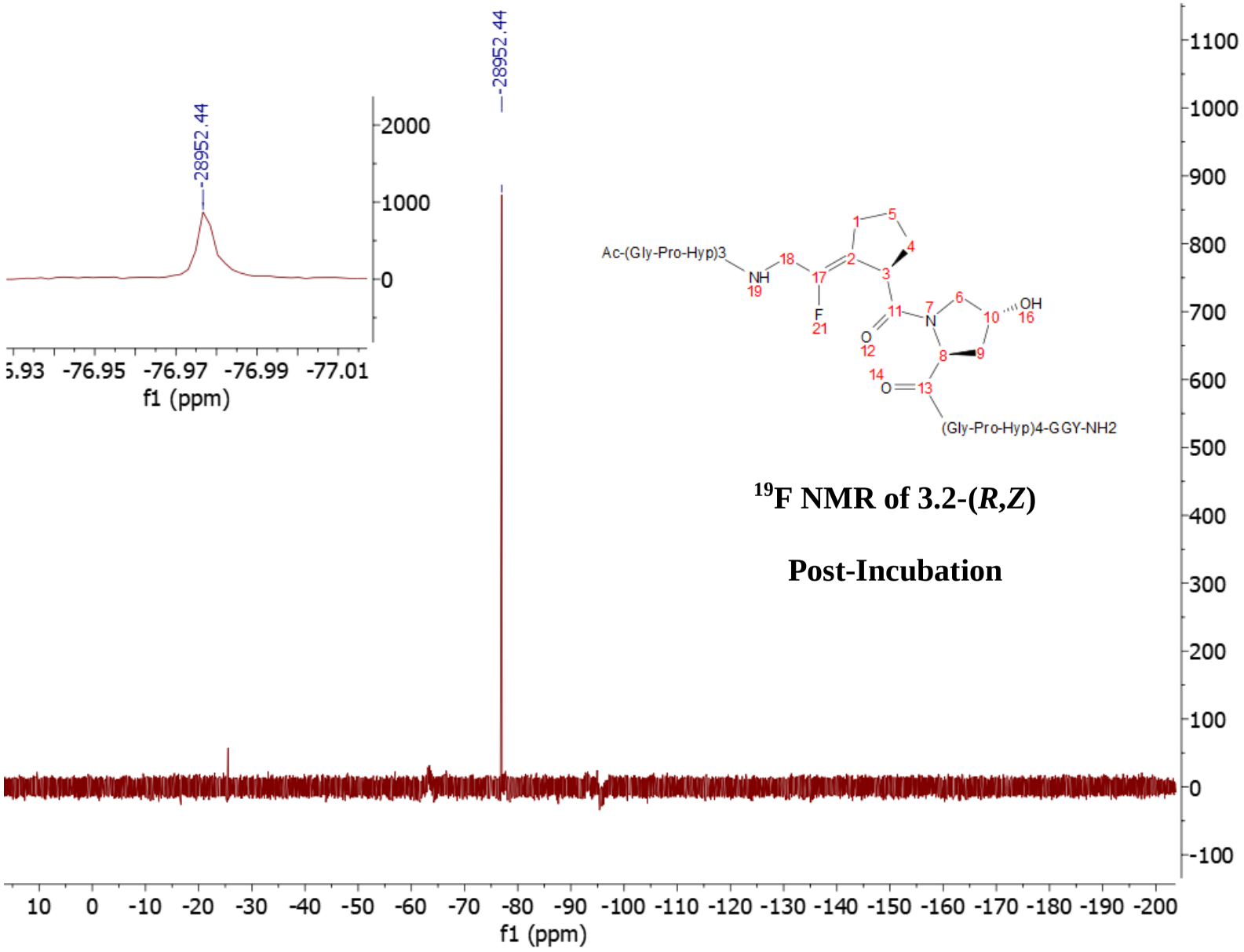

3029.74
2903.01 CDCl3
2900.33
2362.75
2059.14
2046.82
2029.60
2026.57
2008.57
1996.34
1853.24
1849.04
1844.54
1840.33
1818.42
1705.55
1661.34
1614.58
1555.60
1465.71
1394.89
1272.53
1250.91
1245.24
1239.66
1234.09
1091.72
1087.67
1080.51
1079.28
1076.47
797.26
736.22
432.61
430.17
3.16
3.91
9.17
0.64
3.10
0.70
1.03
2.58
4.29
1.21
1.56
1.20
0.63
1.14
2.87
1.74
4.55
4.53
f1 (ppm)


3.17
f1 (ppm)
2.19
6.77
6.07
4.59
0.95
0.98
4.41
2.01
0.91
1.80
2.04
2.04
2.00
0.92
455.44
461.31
731.56
774.40
835.15
913.98
1292.91
1360.79
1421.53
1485.60
1568.25
1635.94
1686.12
1735.71
1852.01
2486.52
2902.01 CDCl3
2902.98
2910.16
2917.29
2935.42
2942.91
2950.32
3029.77
3035.39
3042.29
3084.36
3091.87
3999.88



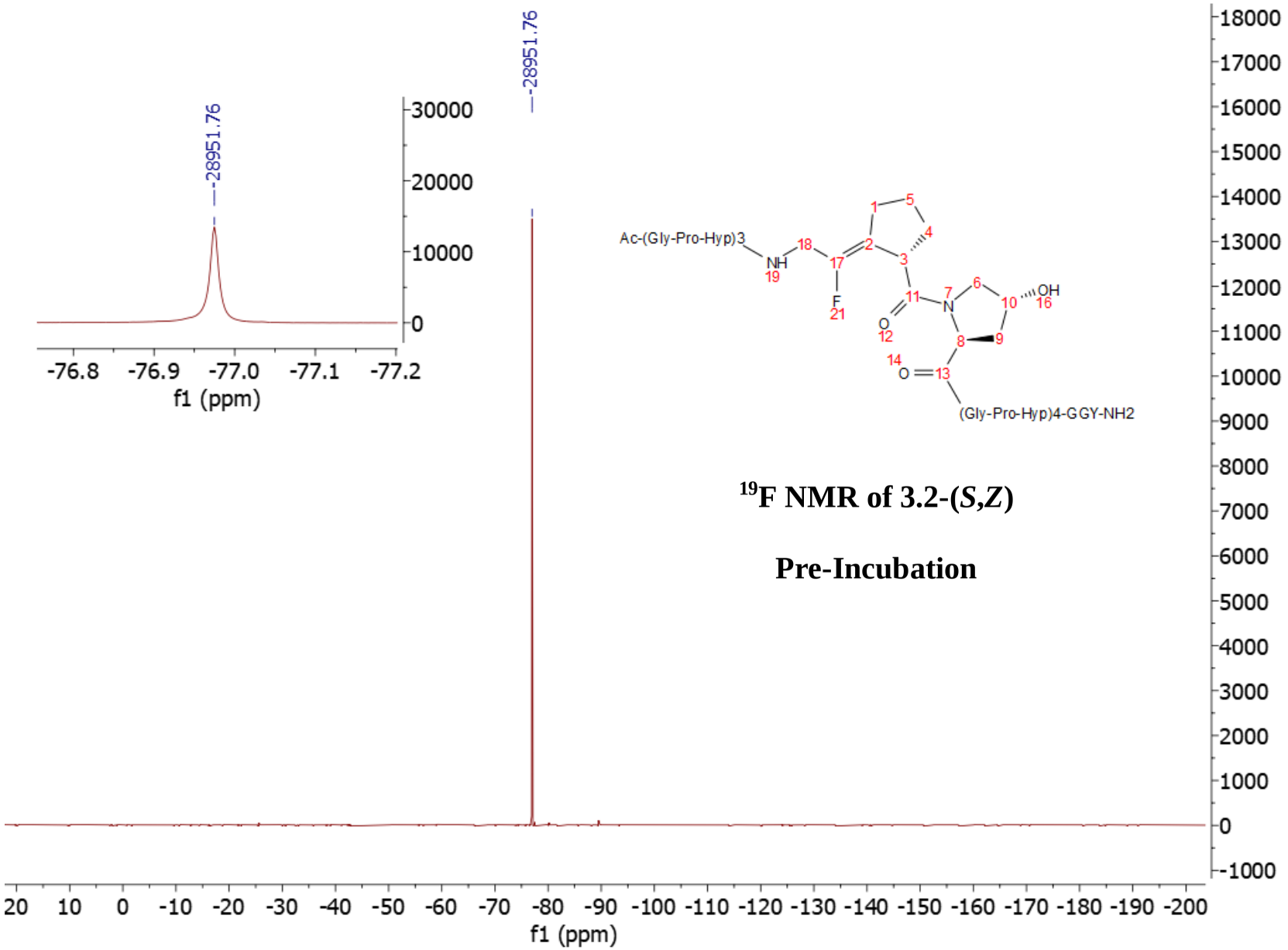
^{19}F NMR of 3.2-(*R,Z*)

Pre-Incubation



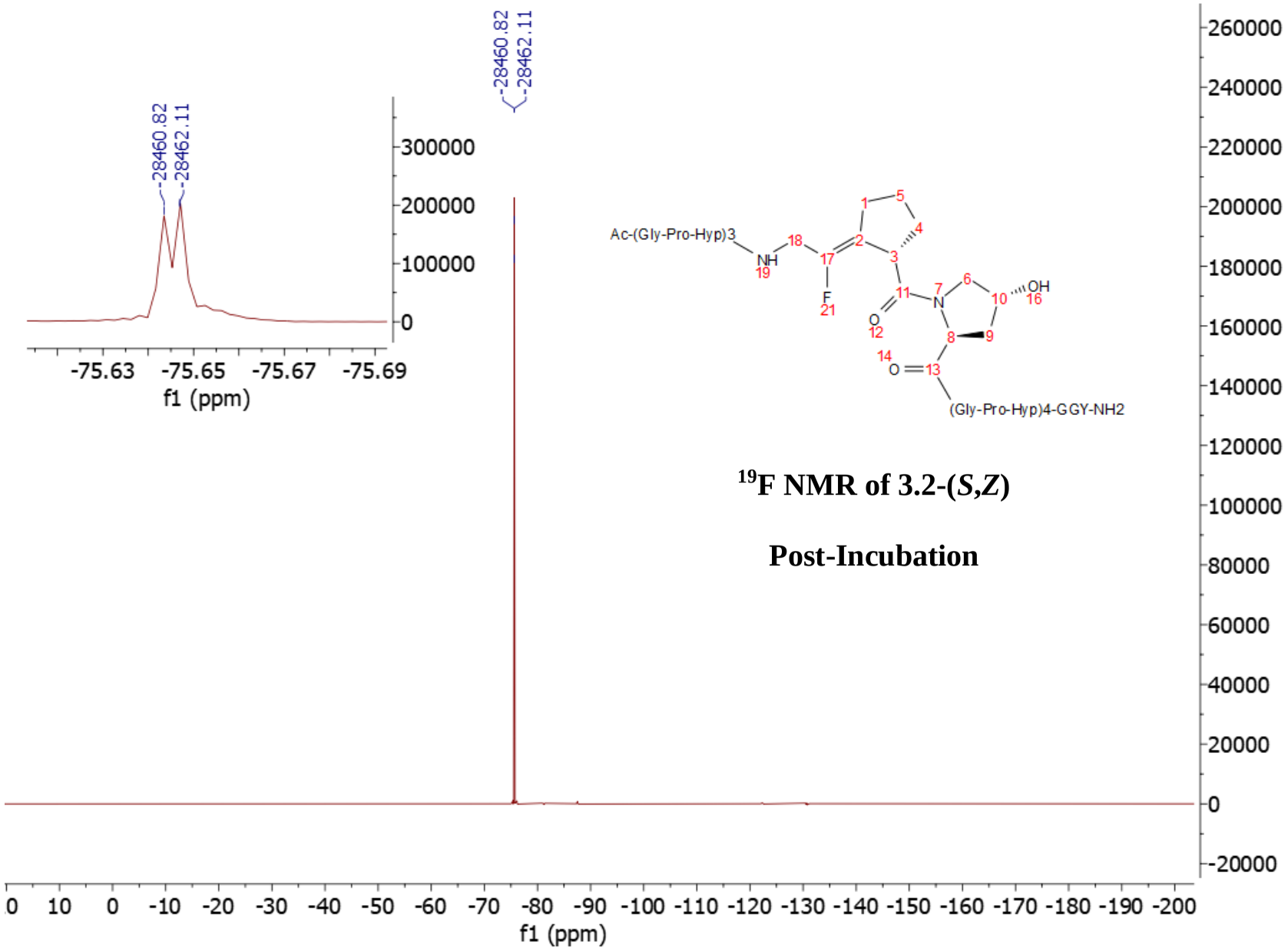
^{19}F NMR of 3.2-(*R*,*Z*)

Post-Incubation



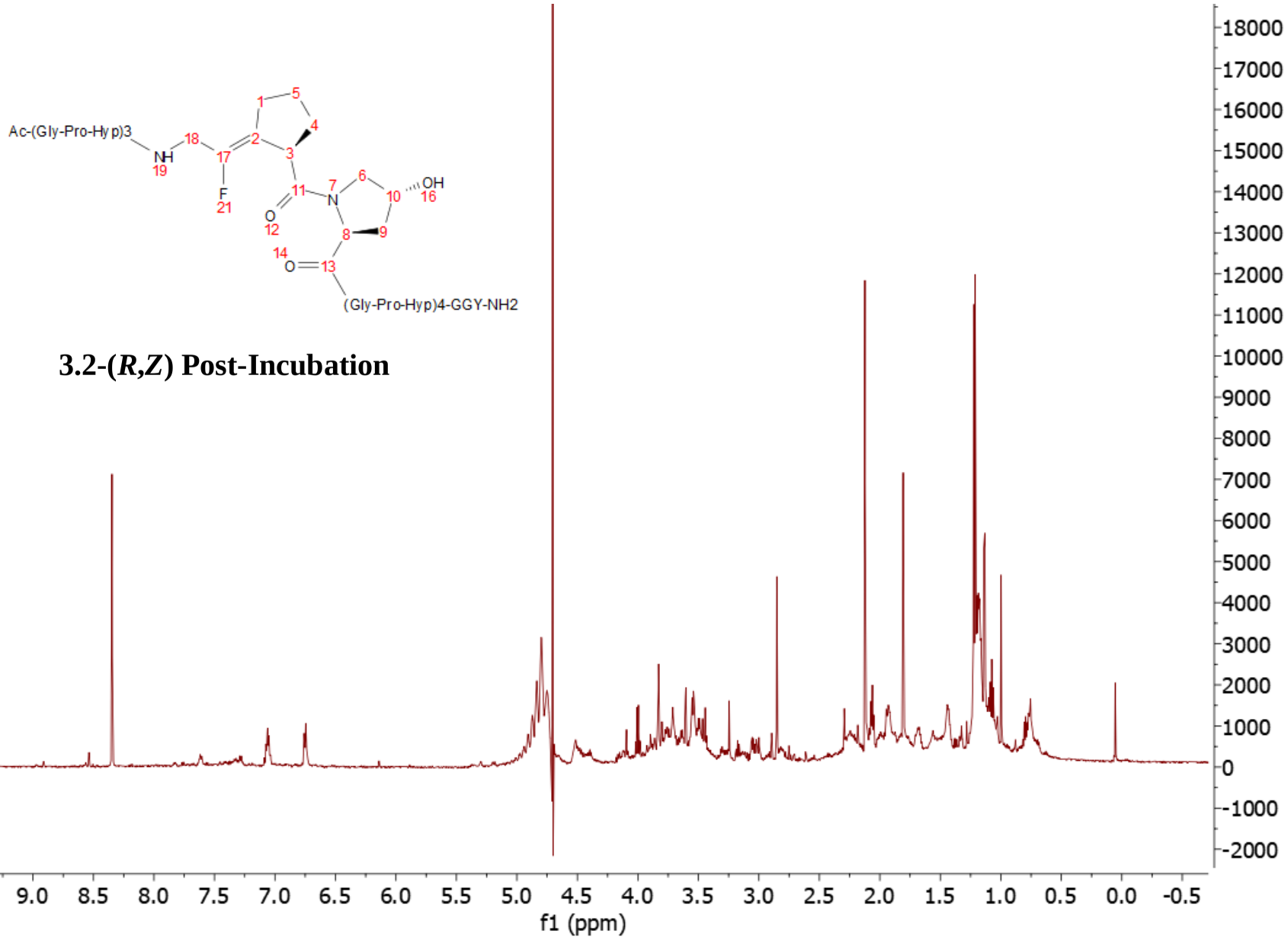
^{19}F NMR of 3.2-(*S*,*Z*)

Pre-Incubation

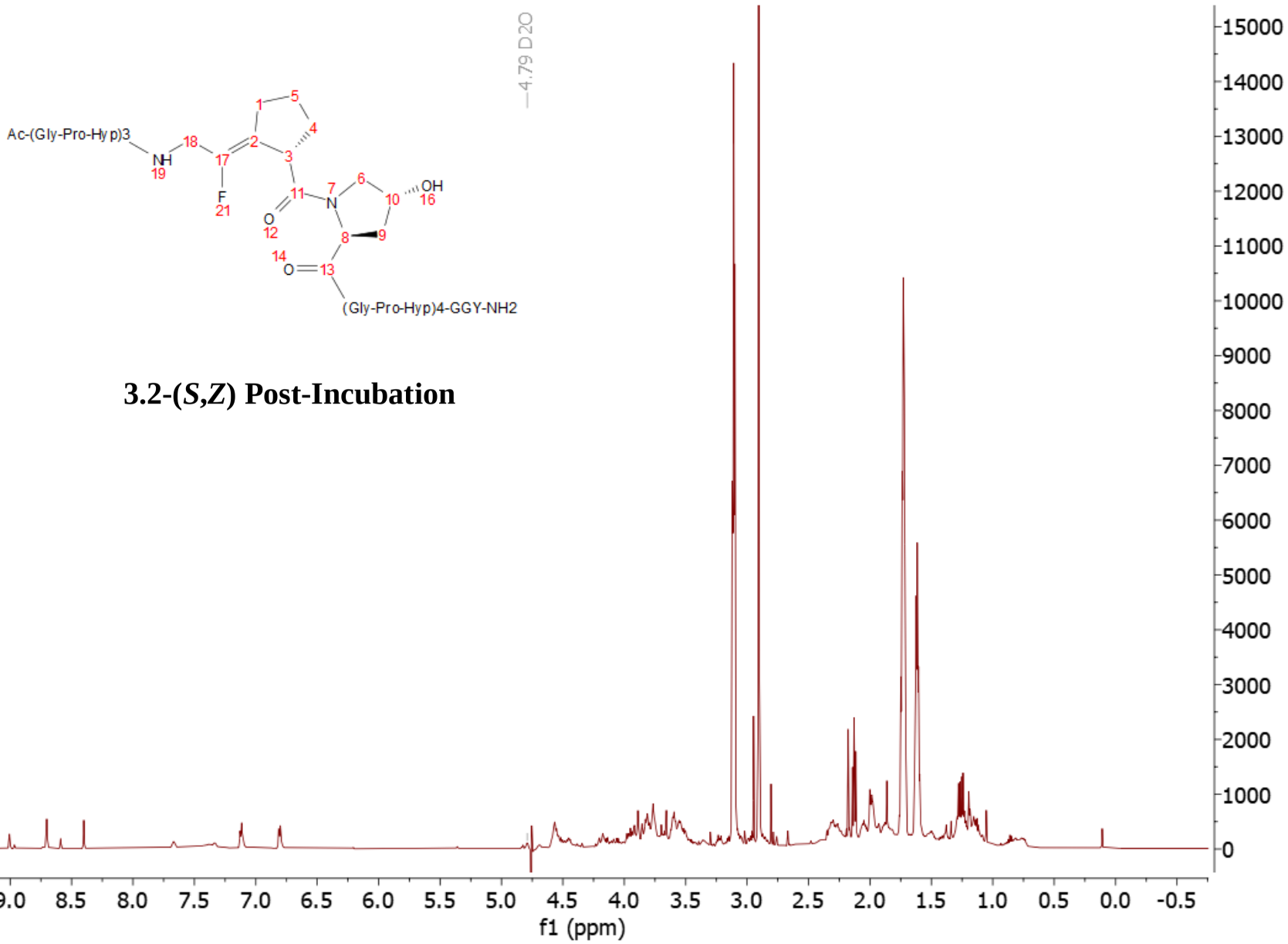


^{19}F NMR of 3.2-(*S*,*Z*)

Post-Incubation



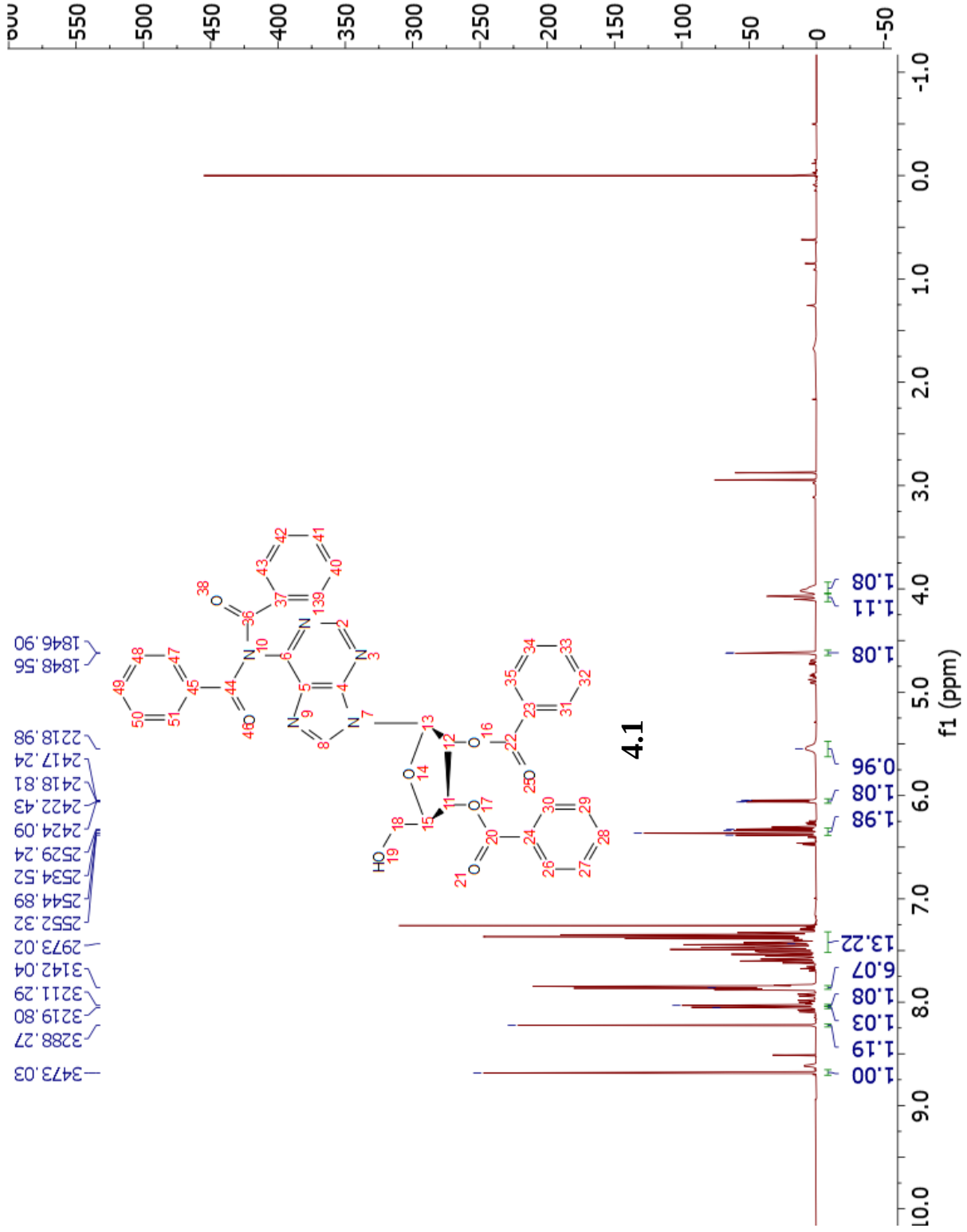
3.2-(*R*,*Z*) Post-Incubation

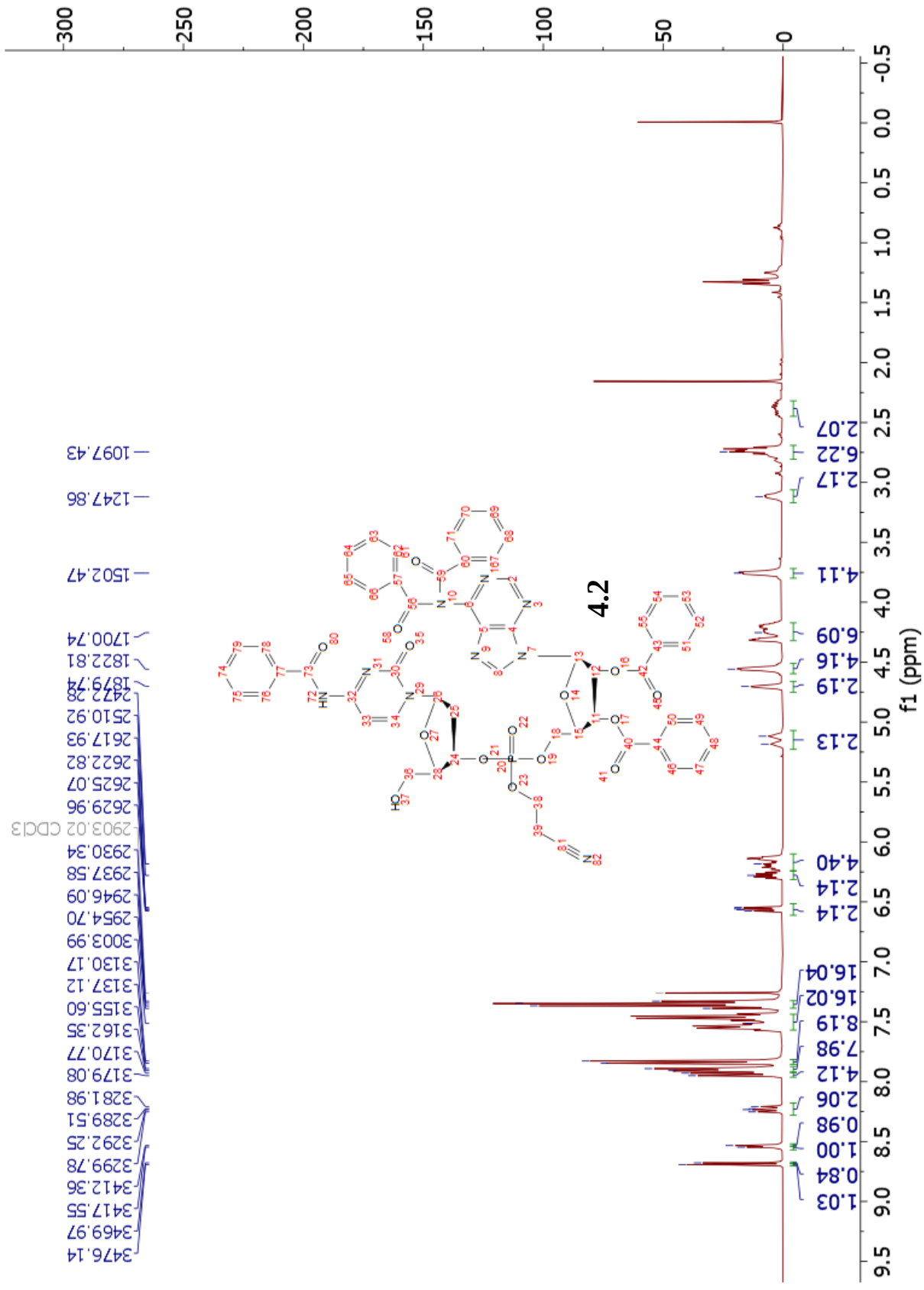


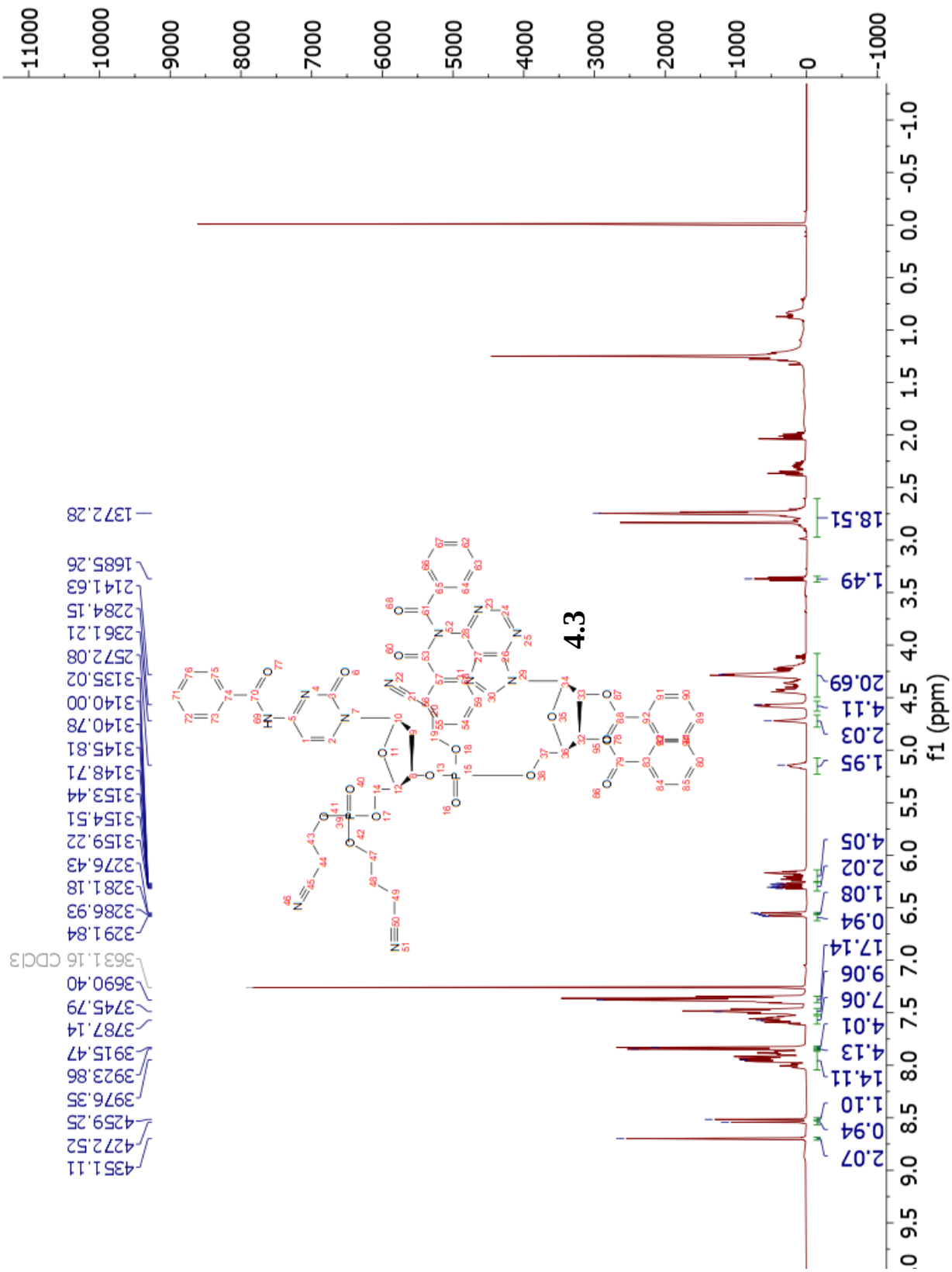
3.2-(*S*,*Z*) Post-Incubation

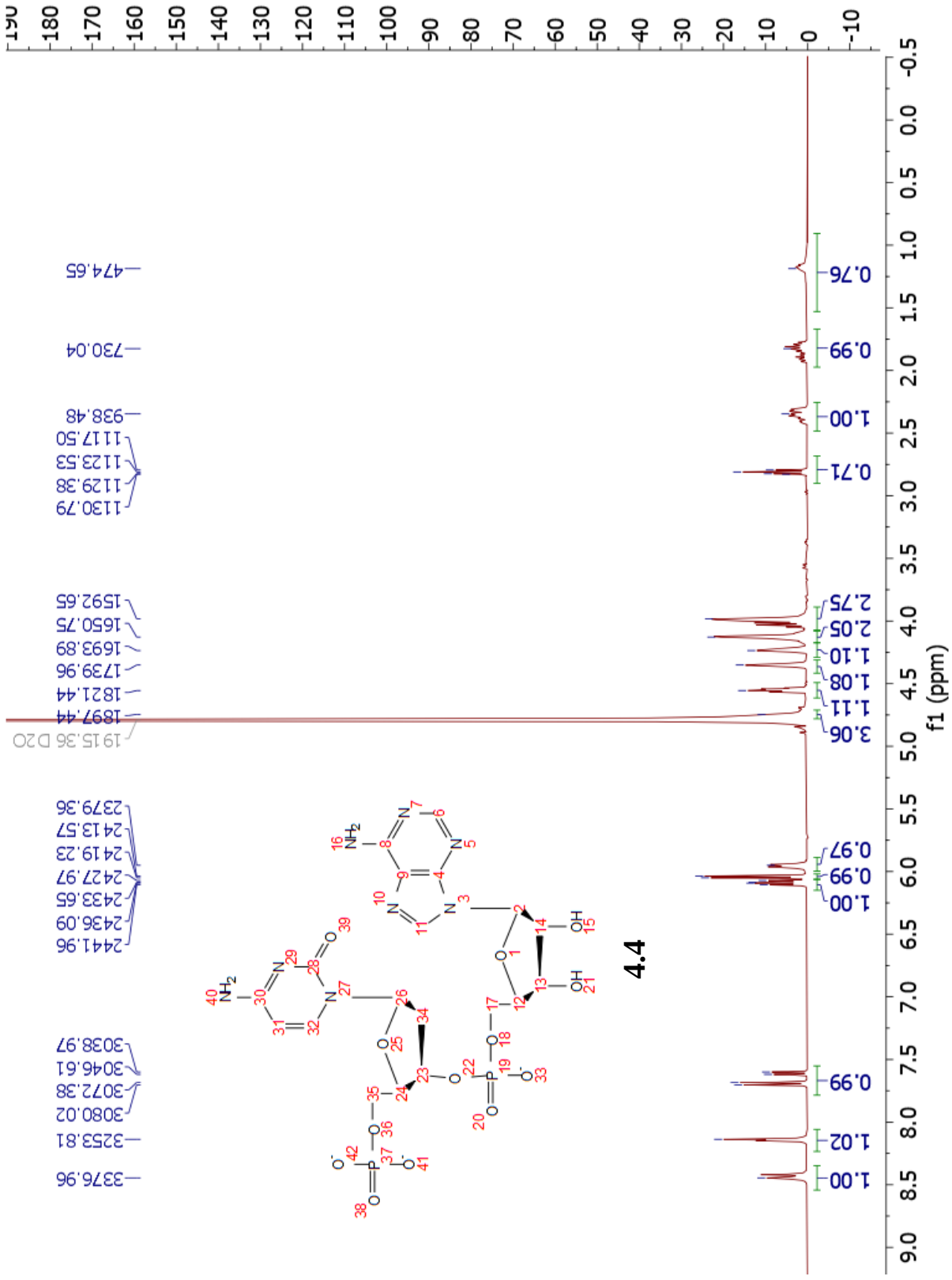
Appendix C

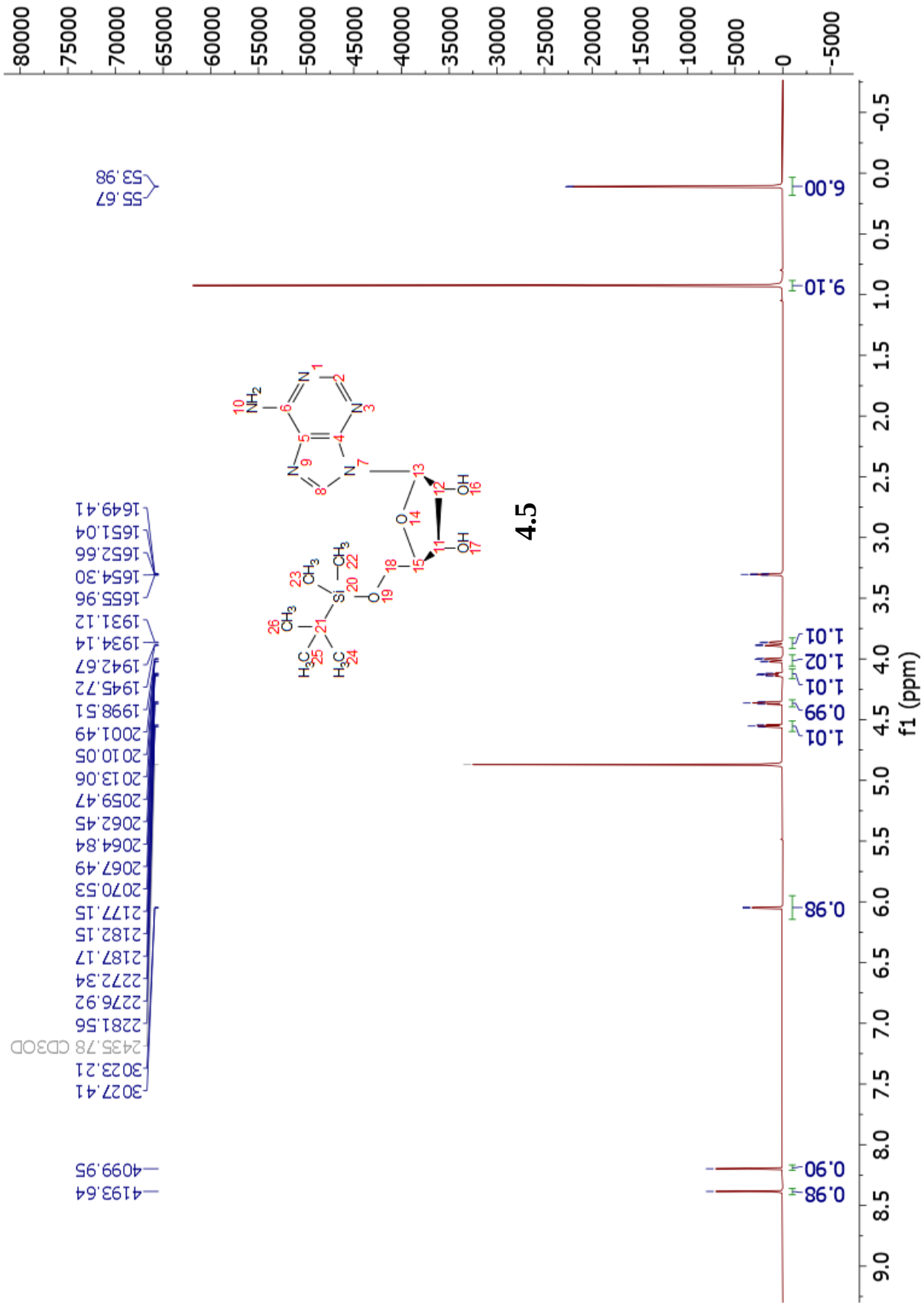
Supplementary Information for Chapter 4

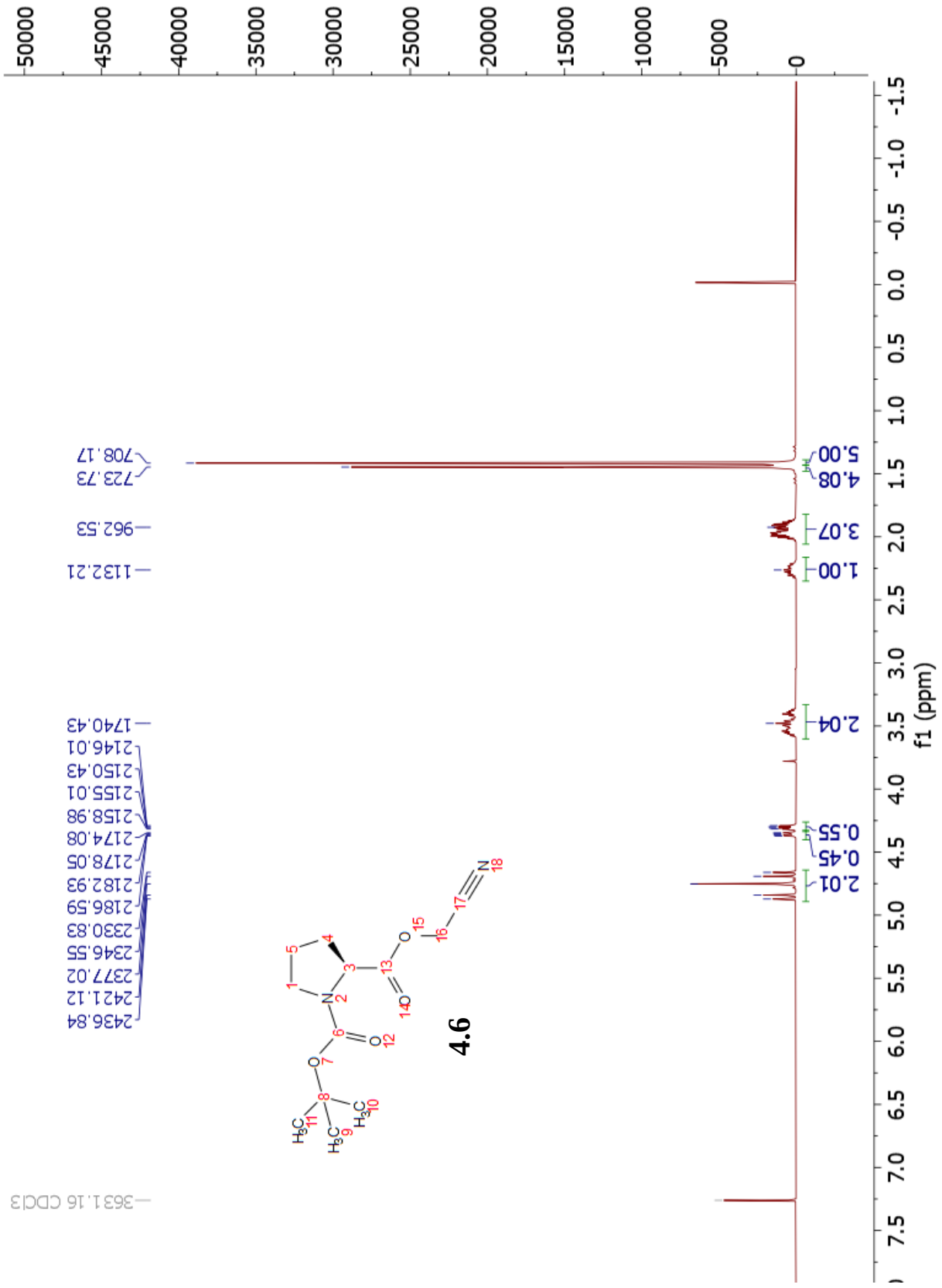


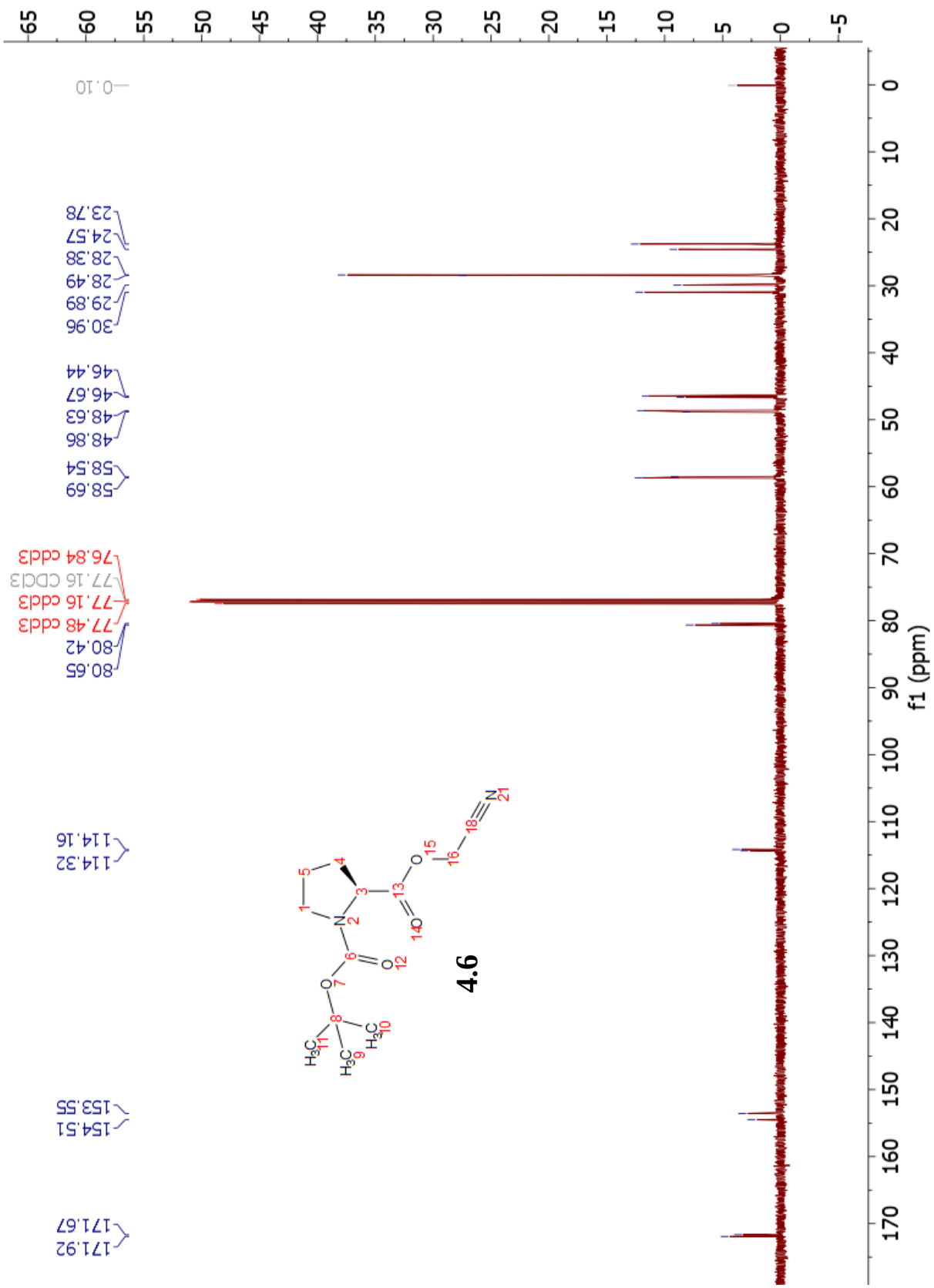

4.2
f1 (ppm)
1.03
0.84
1.00
0.98
2.06
4.12
7.98
8.19
16.02
16.04
7.14
2.14
4.40
2.13
2.19
4.16
6.09
4.11
2.17
6.22
2.07
1097.43
1247.86
1502.47
1700.74
1822.81
1879.74
2472.28
2510.92
2617.93
2622.82
2625.07
2629.96
2903.02 CDCl3
2930.34
2937.58
2946.09
2954.70
3003.99
3130.17
3137.12
3155.60
3162.35
3170.77
3179.08
3281.98
3289.51
3292.25
3299.78
3412.36
3417.55
3469.97
3476.14

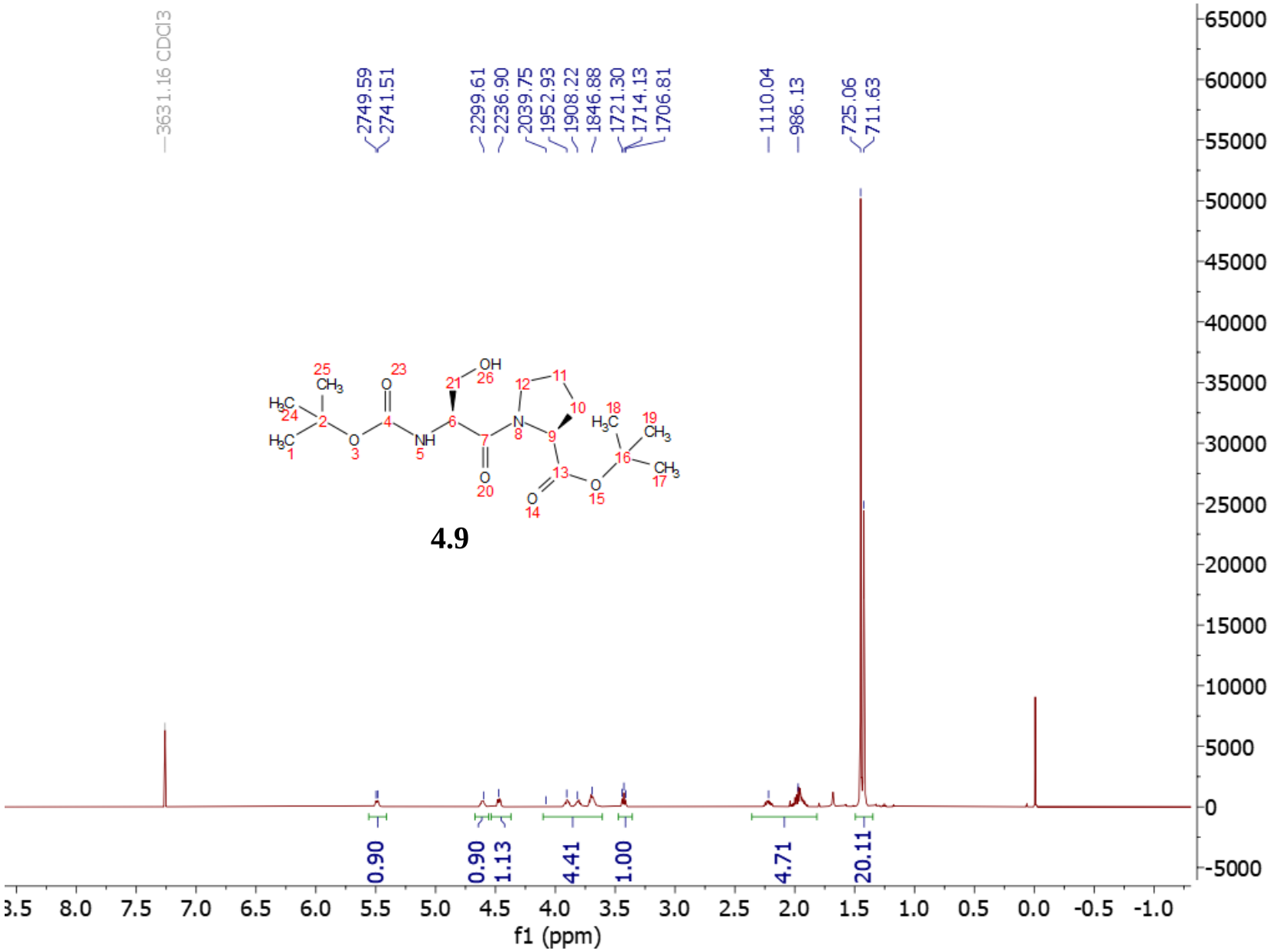

4.3
f1 (ppm)


4.4
f1 (ppm)

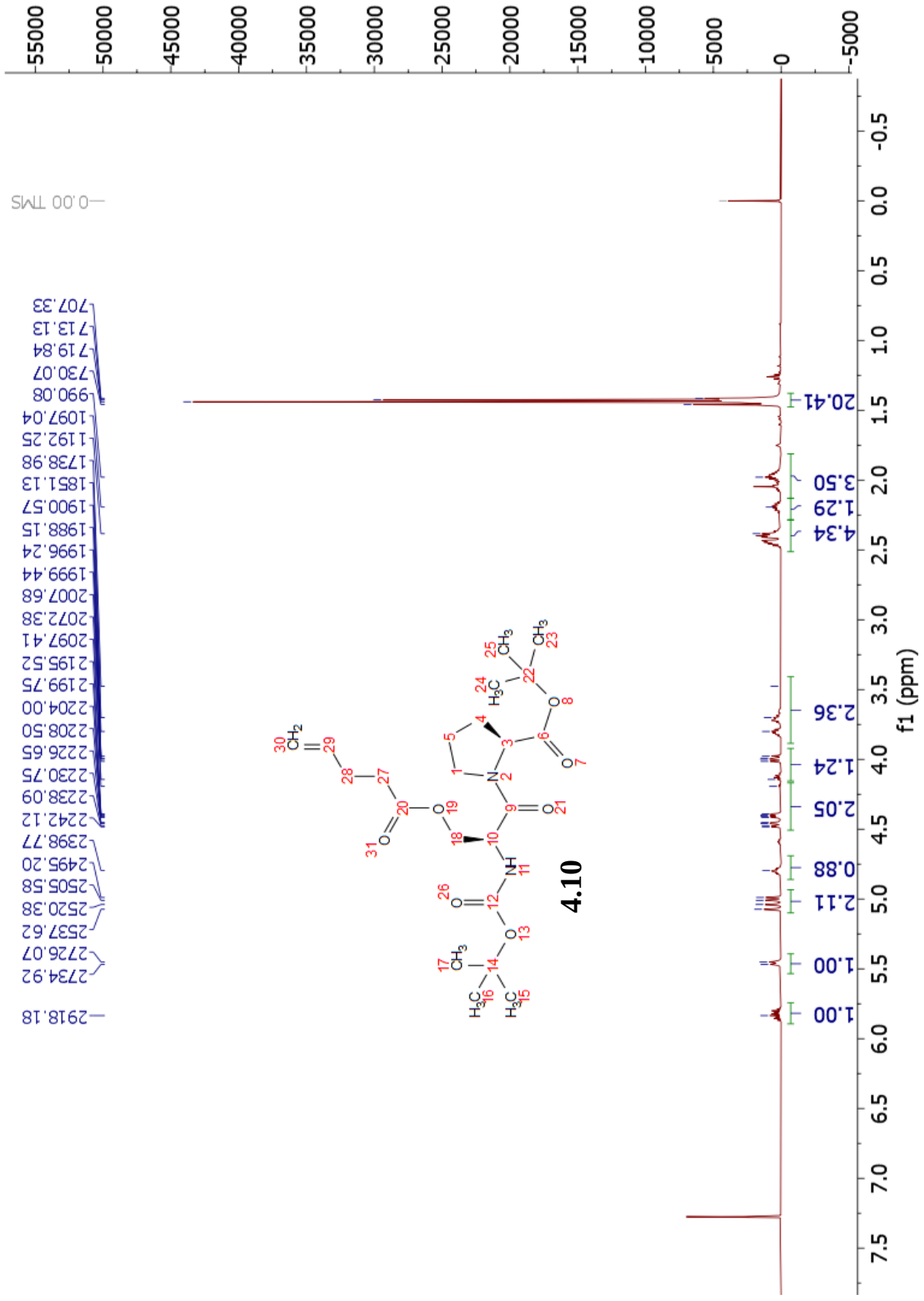

4.5
f1 (ppm)
6.00
9.10
1.01
1.02
1.01
0.99
1.01
0.98
0.90
0.98
CD3OD

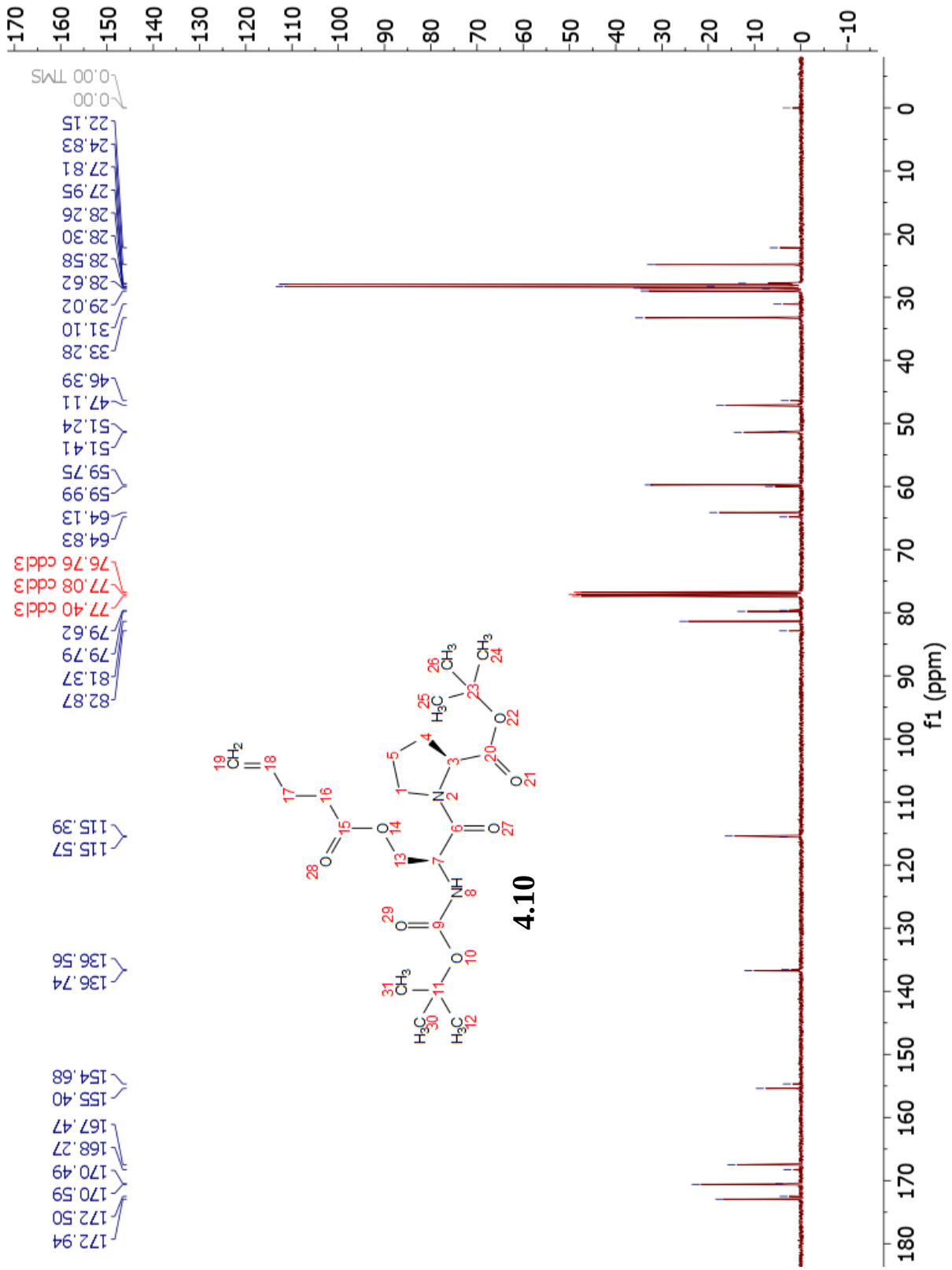

5.00
4.08
3.07
1.00
2.04
2.01
0.45
0.55
708.17
723.73
962.53
1132.21
1740.43
2146.01
2150.43
2155.01
2158.98
2174.08
2178.05
2182.93
2186.59
2330.83
2346.55
2377.02
2421.12
2436.84
363.16 CDCl3
4.6
f1 (ppm)

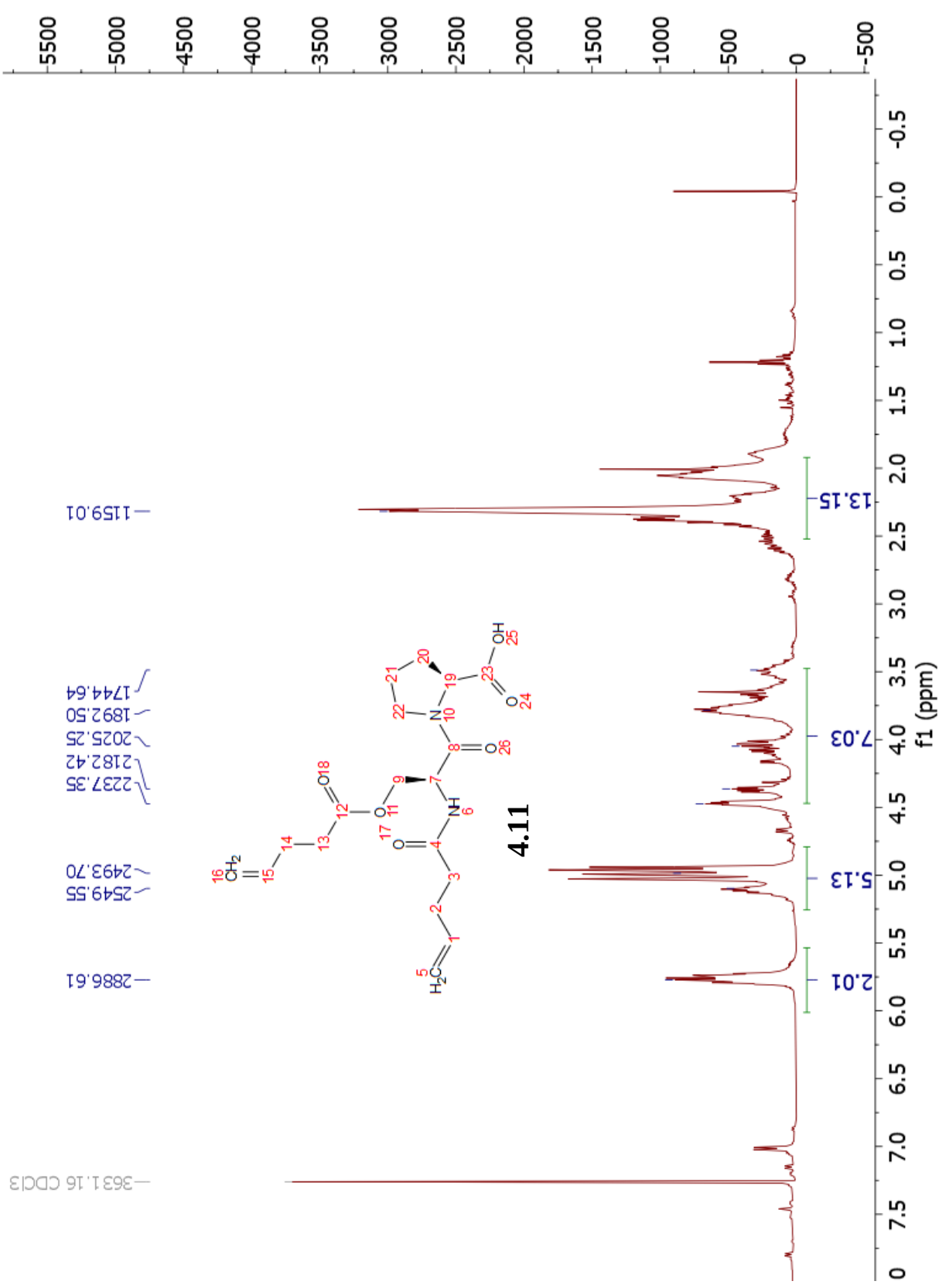

171.92
171.67
154.51
153.55
114.32
114.16
80.65
80.42
77.48 cddl3
77.16 cddl3
77.16 CDCl3
76.84 cddl3
58.69
58.54
48.86
48.63
46.67
46.44
30.96
29.89
28.49
28.38
24.57
23.78
0.10
4.6
f1 (ppm)

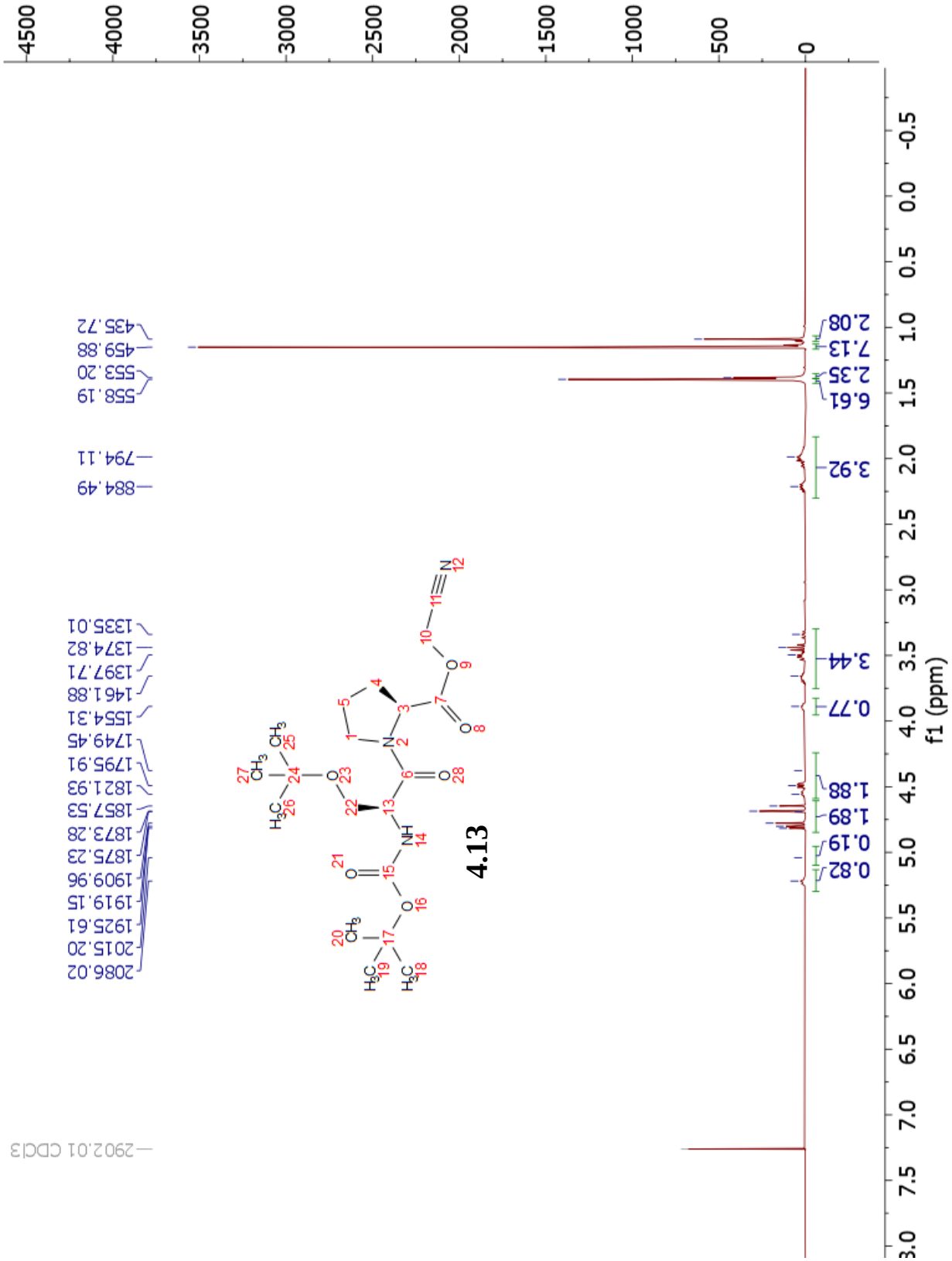

3631.16 CDCl3
2749.59
2741.51
2299.61
2236.90
2039.75
1952.93
1908.22
1846.88
1721.30
1714.13
1706.81
1110.04
986.13
725.06
711.63
0.90
0.90
1.13
4.41
1.00
4.71
20.11
f1 (ppm)

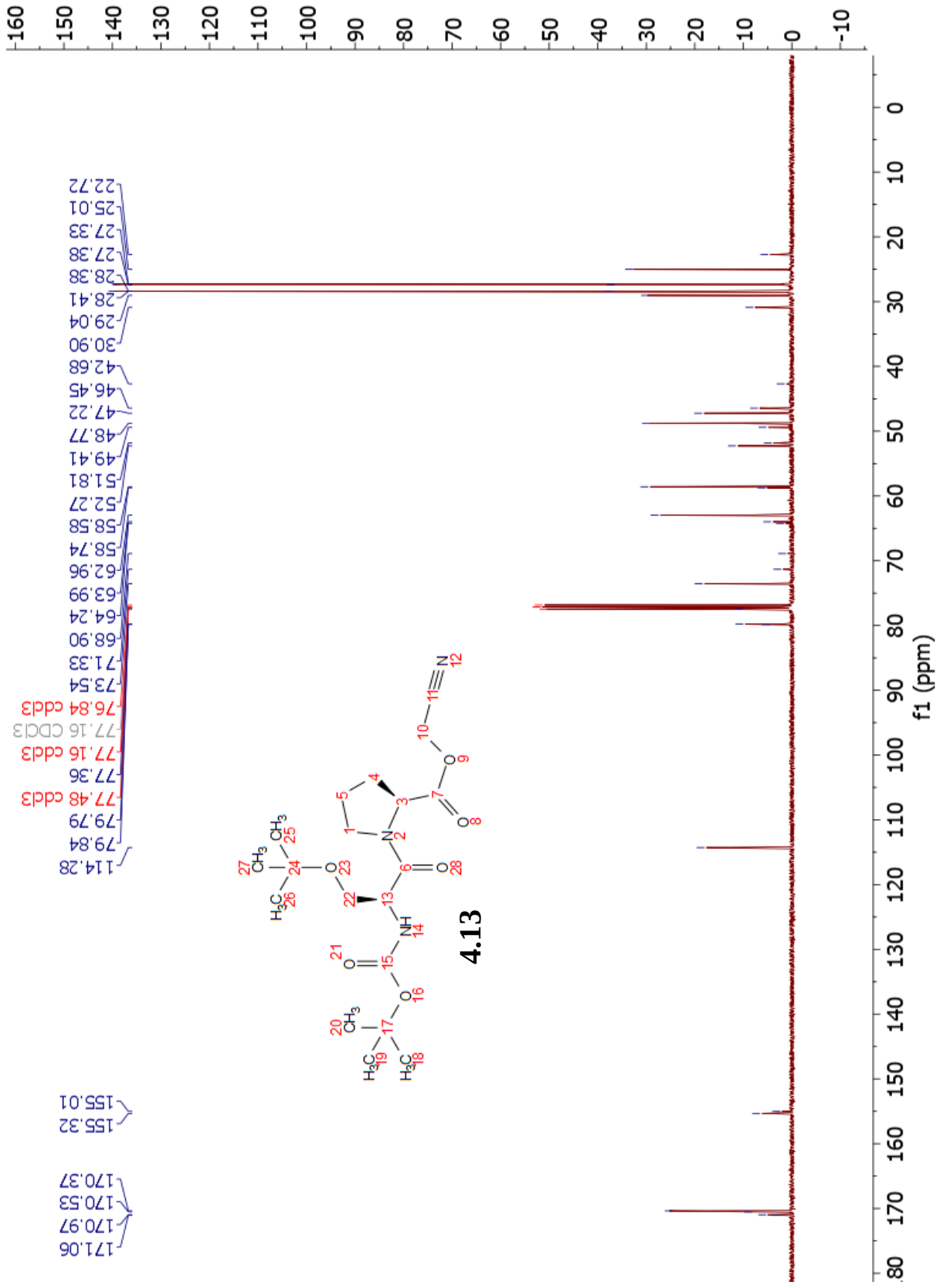
4.9

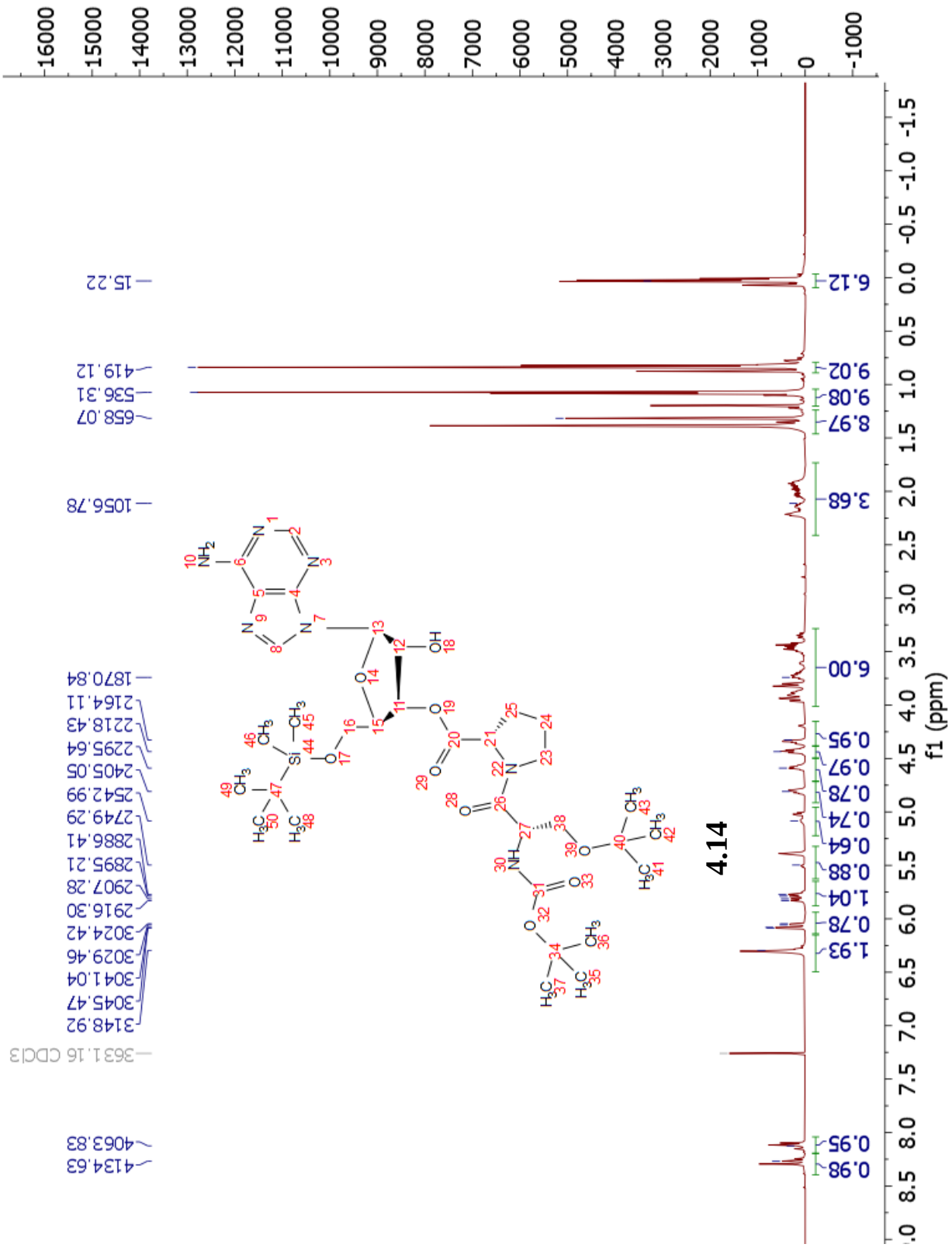

4.10
f1 (ppm)
20.41
3.50
1.29
4.34
2.36
1.24
2.05
0.88
2.11
1.00
1.00
TMS 0.00

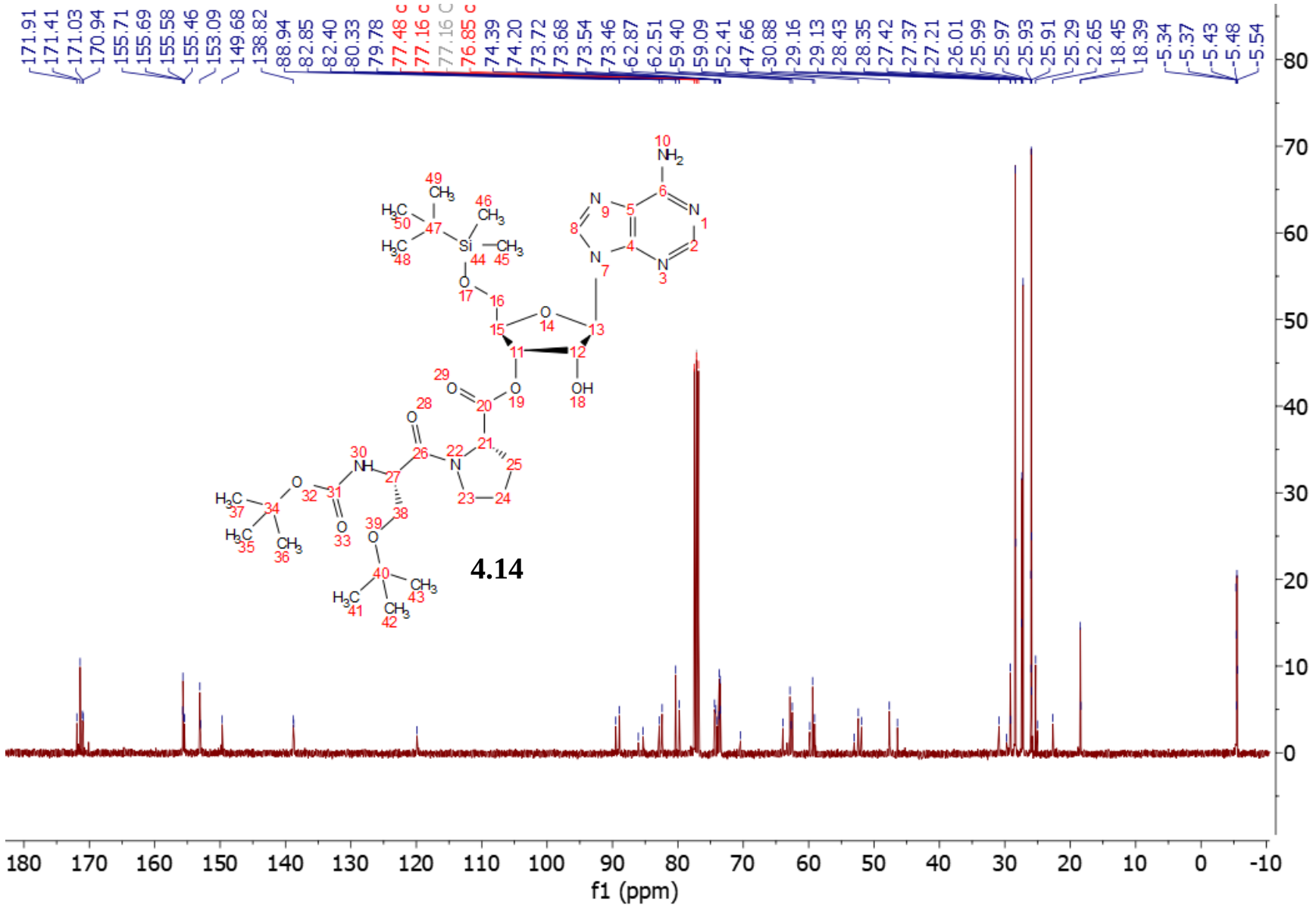

4.10
f1 (ppm)

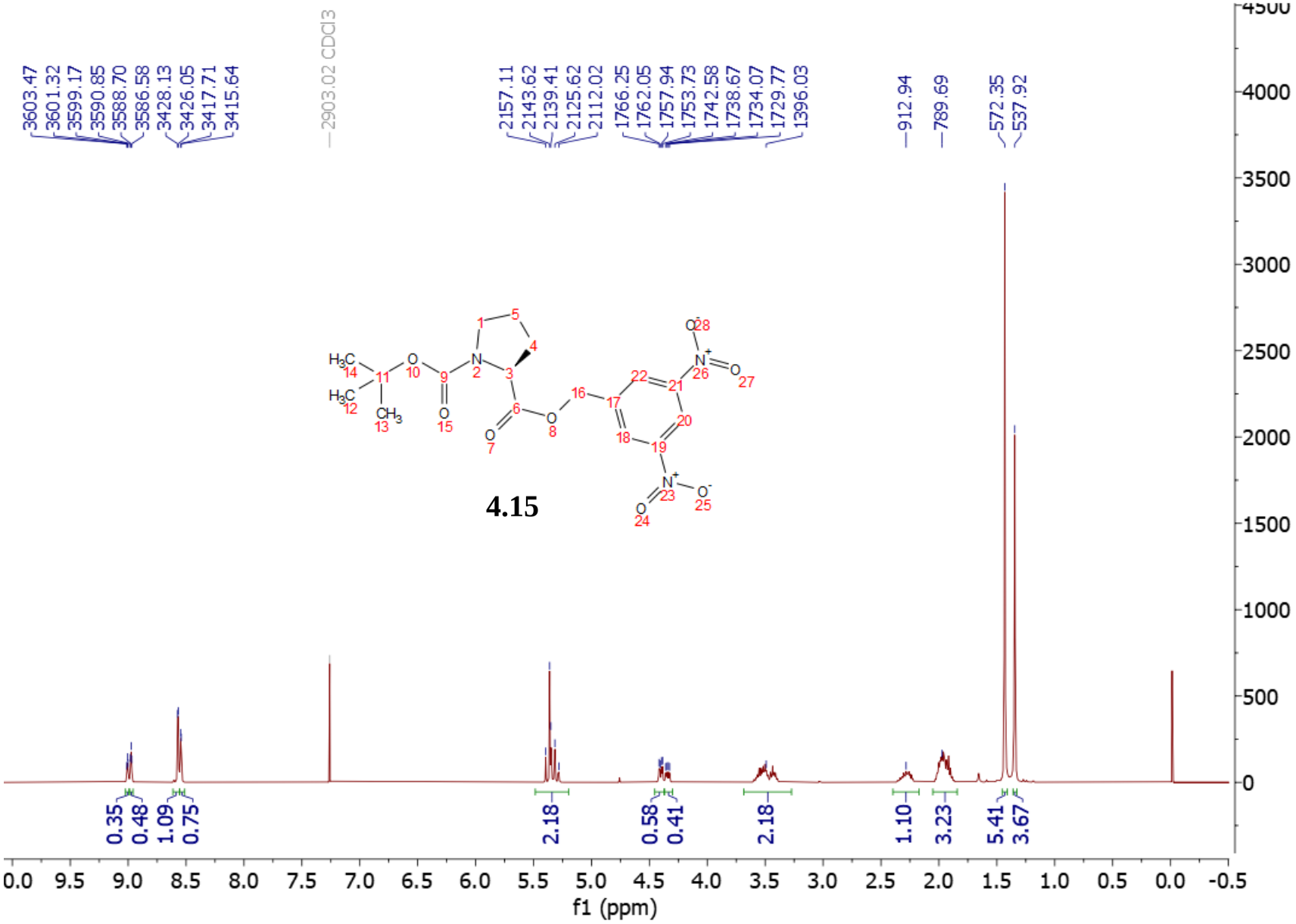

4.11
13.15
7.03
5.13
2.01
f1 (ppm)
3631.16 CDCl3
2886.61
2549.55
2493.70
2237.35
2182.42
2025.25
1892.50
1744.64
1159.01

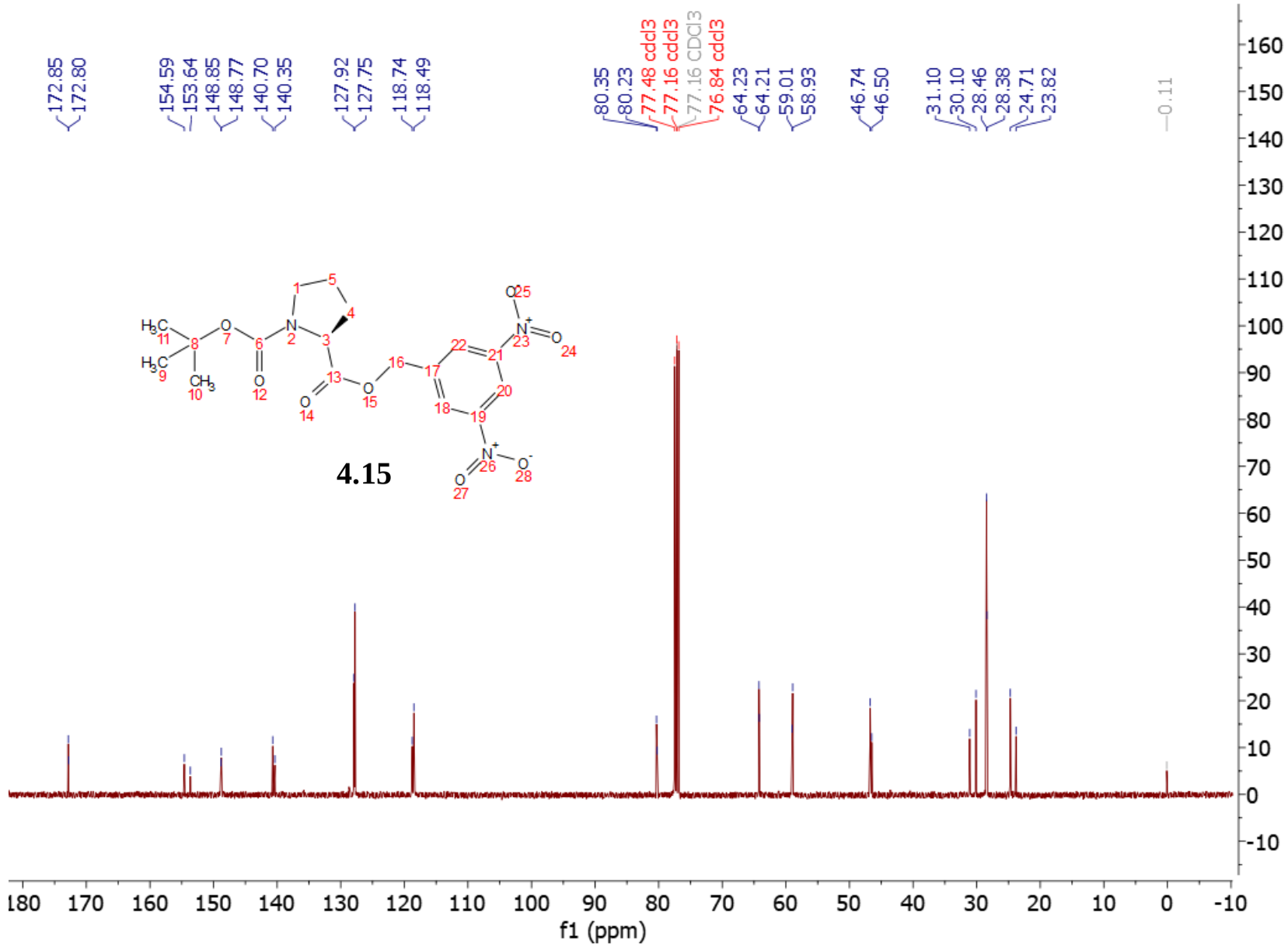

4.13
f1 (ppm)
2902.01 CDCl3

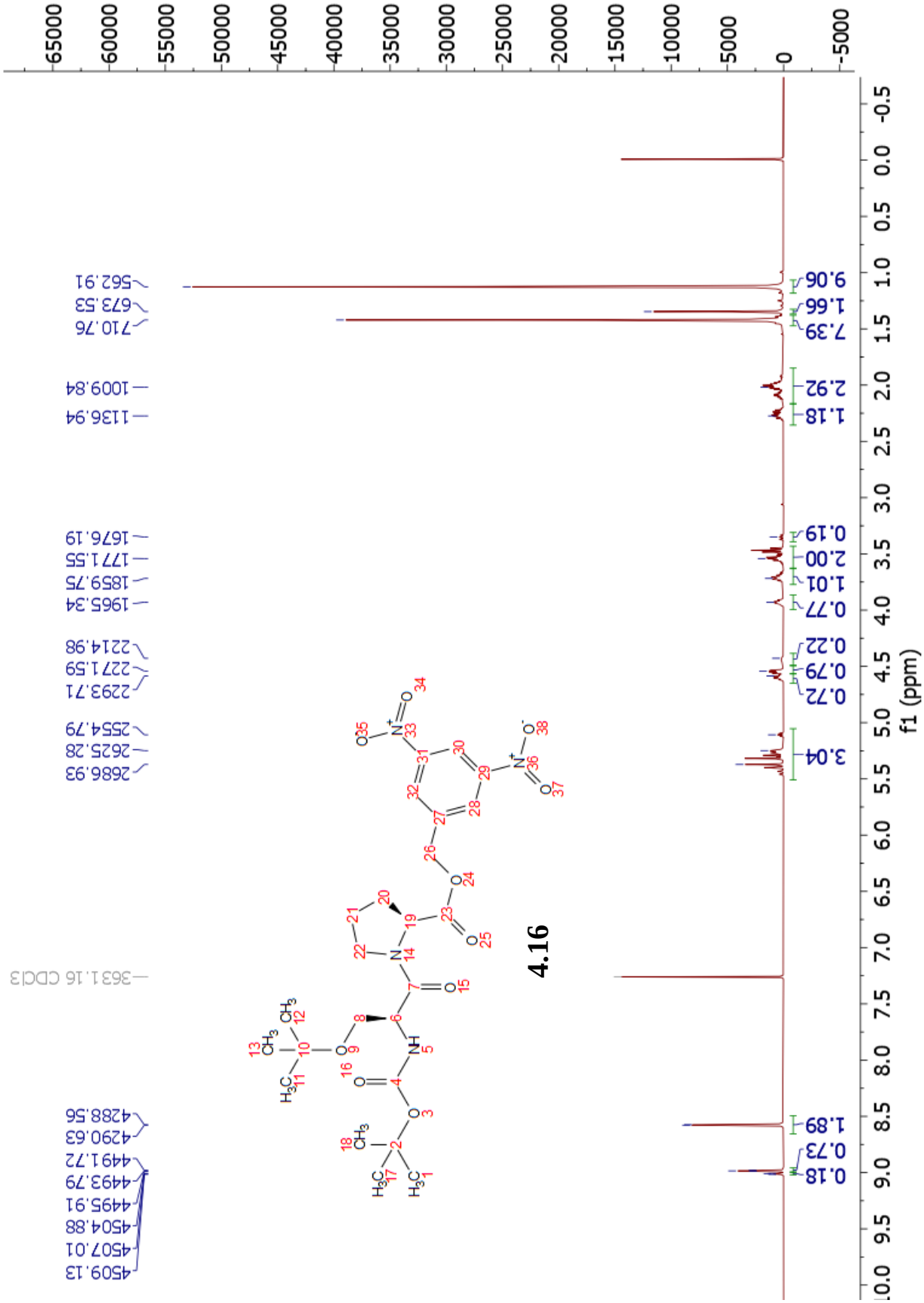

f1 (ppm)
4.13

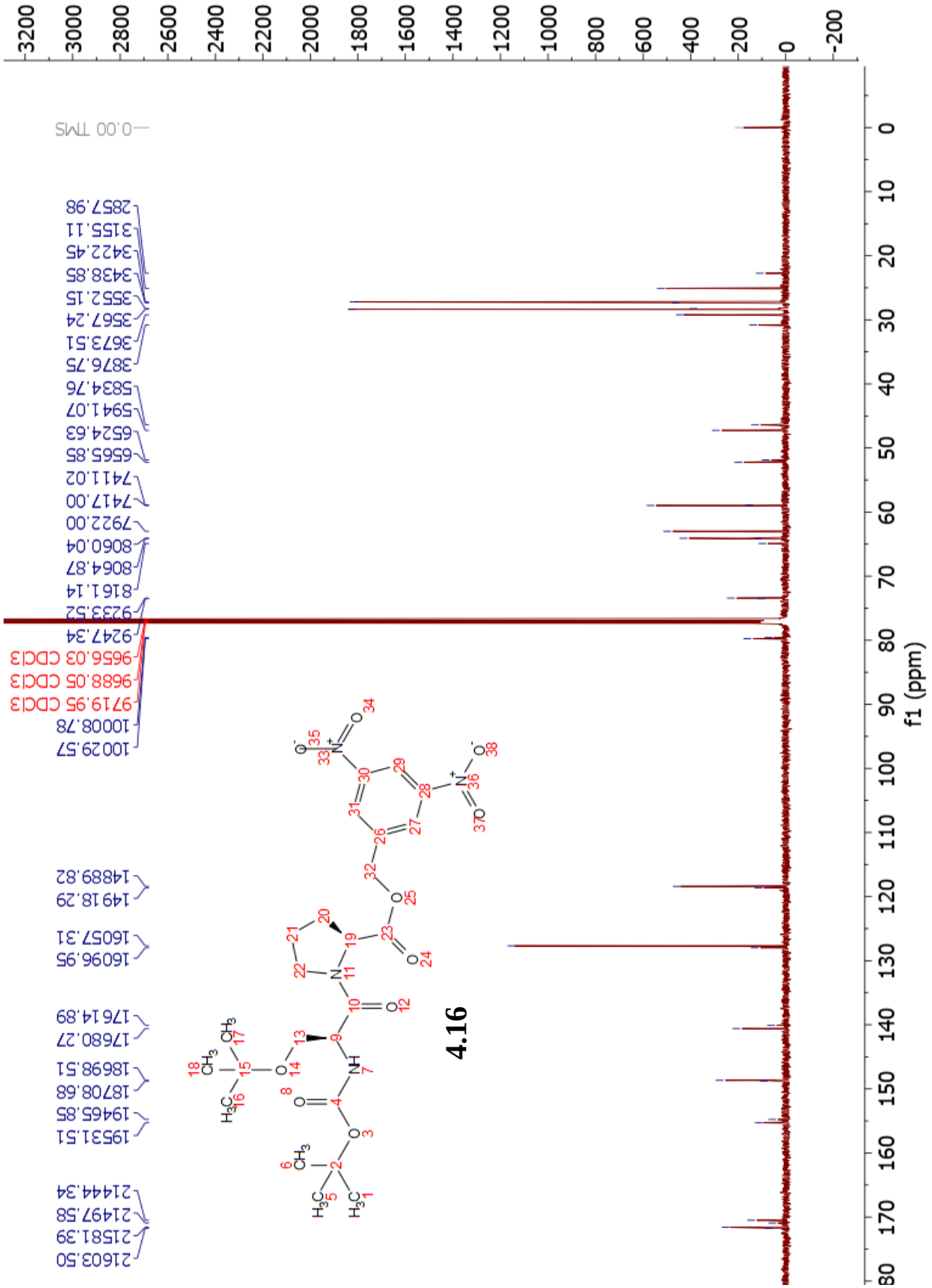

4.14
f1 (ppm)

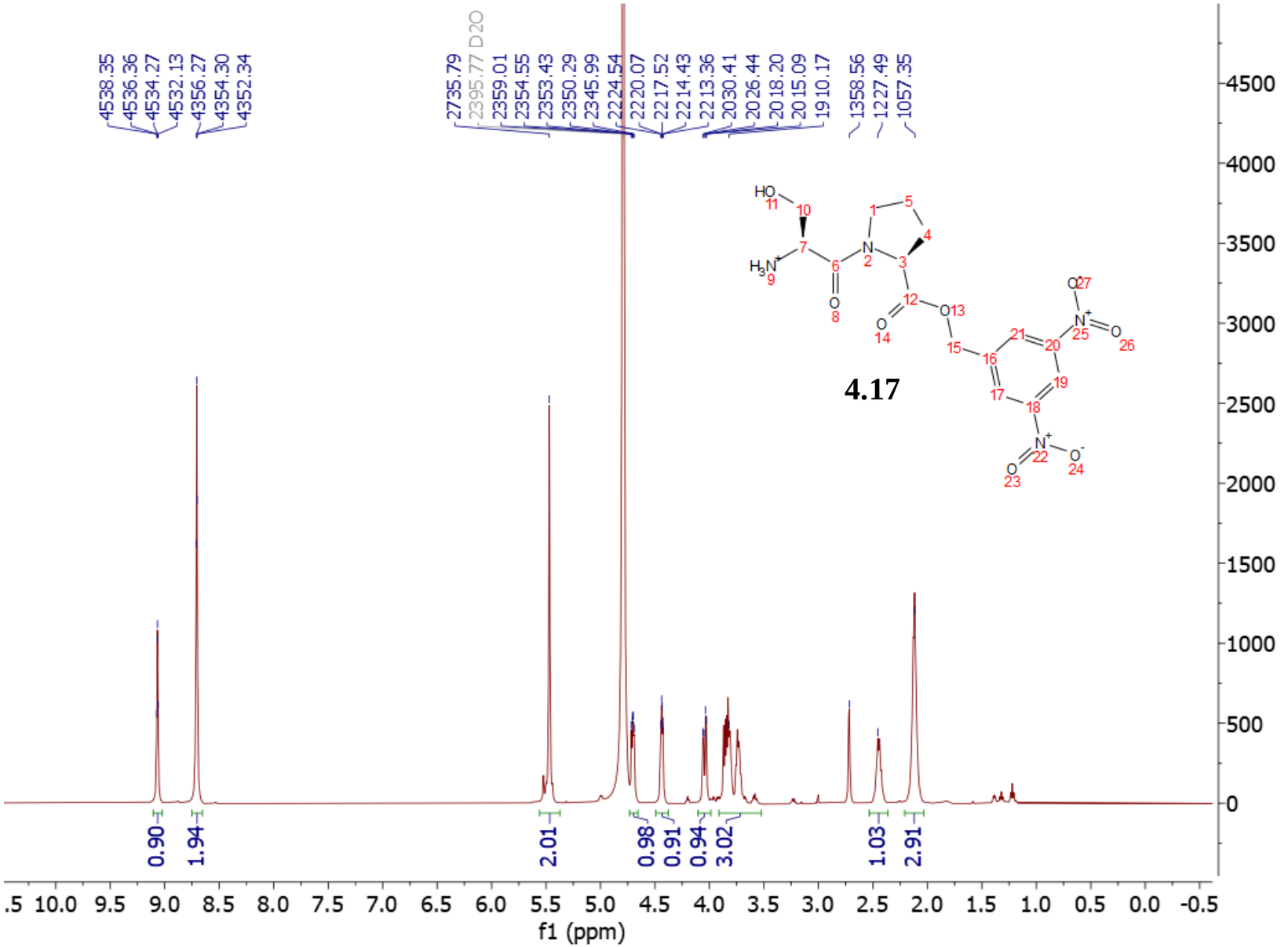

171.91
171.41
171.03
170.94
155.71
155.69
155.58
155.46
153.09
149.68
138.82
88.94
82.85
82.40
80.33
79.78
77.48 c
77.16 c
77.16 C
76.85 c
74.39
74.20
73.72
73.68
73.54
73.46
62.87
62.51
59.40
59.09
52.41
47.66
30.88
29.16
29.13
28.43
28.35
27.42
27.37
27.21
26.01
25.99
25.97
25.93
25.91
25.29
22.65
18.45
18.39
-5.34
-5.37
-5.43
-5.48
-5.54
4.14
f1 (ppm)

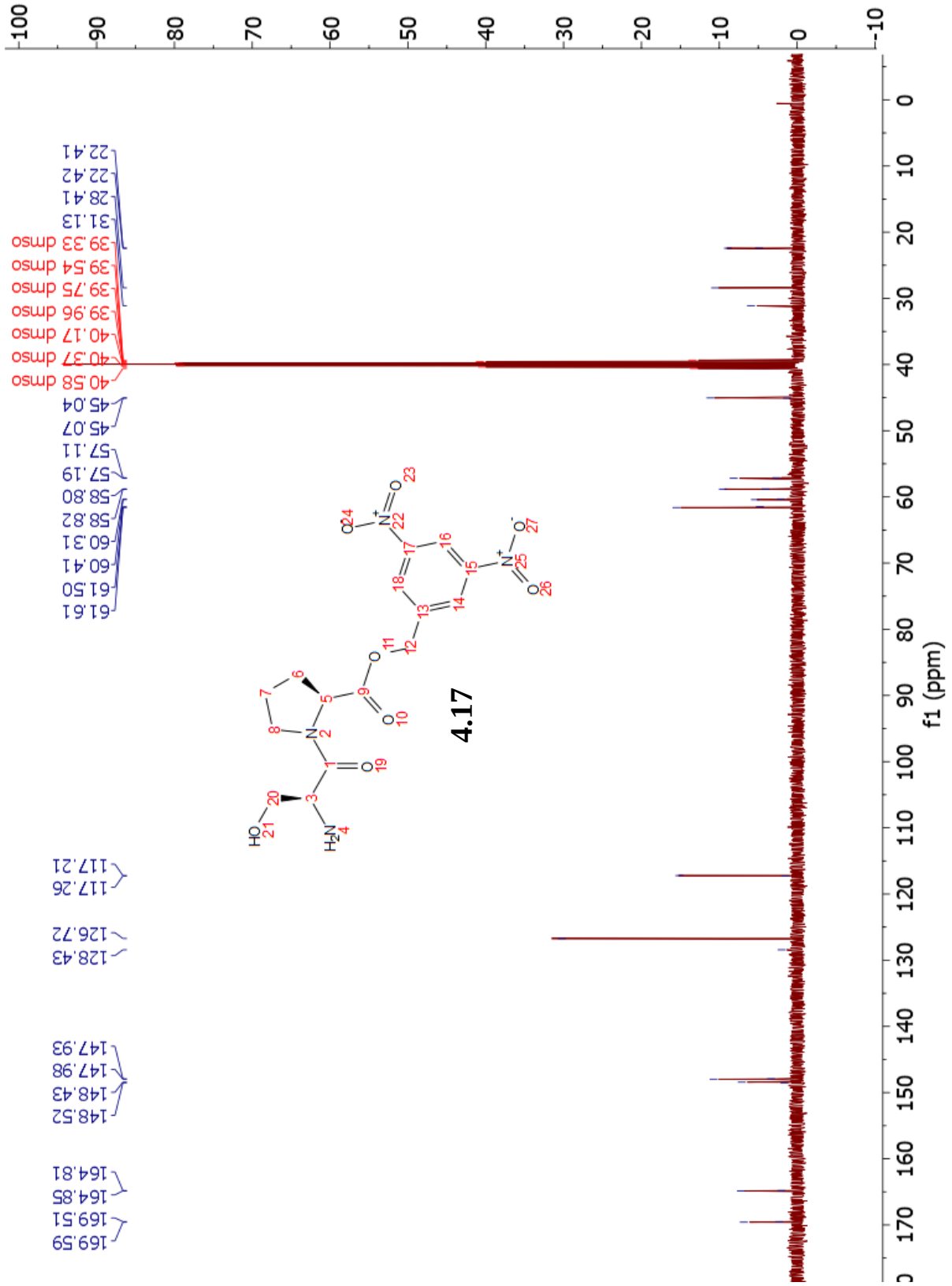

3603.47
3601.32
3599.17
3590.85
3588.70
3586.58
3428.13
3426.05
3417.71
3415.64
2903.02 CDCl3
2157.11
2143.62
2139.41
2125.62
2112.02
1766.25
1762.05
1757.94
1753.73
1742.58
1738.67
1734.07
1729.77
1396.03
912.94
789.69
572.35
537.92
4.15
0.35
0.48
1.09
0.75
2.18
0.58
0.41
2.18
1.10
3.23
5.41
3.67
f1 (ppm)

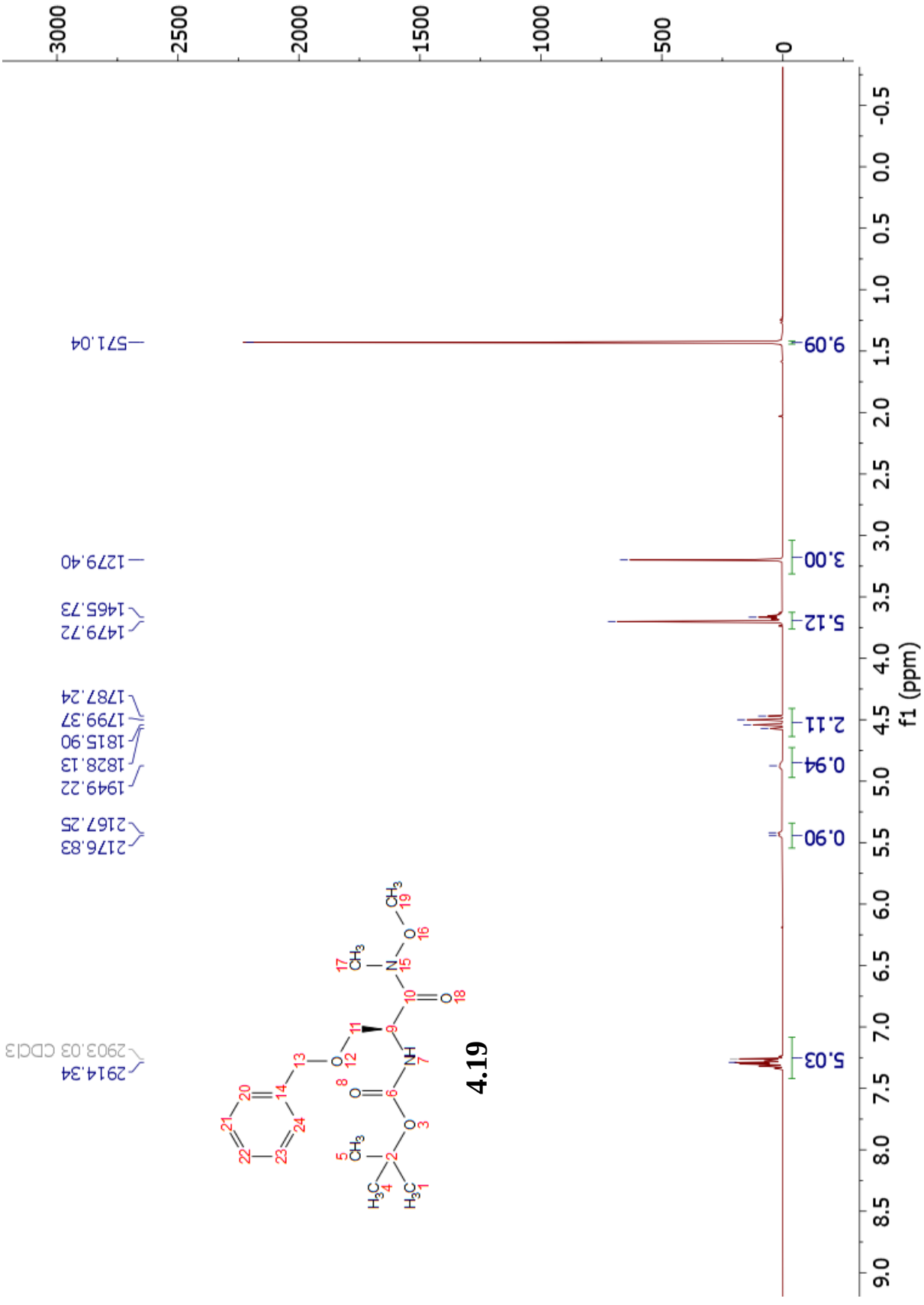

172.85
172.80
154.59
153.64
148.85
148.77
140.70
140.35
127.92
127.75
118.74
118.49
80.35
80.23
77.48 cdcl3
77.16 cdcl3
77.16 CDCl3
76.84 cdcl3
64.23
64.21
59.01
58.93
46.74
46.50
31.10
30.10
28.46
28.38
24.71
23.82
0.11
4.15
f1 (ppm)

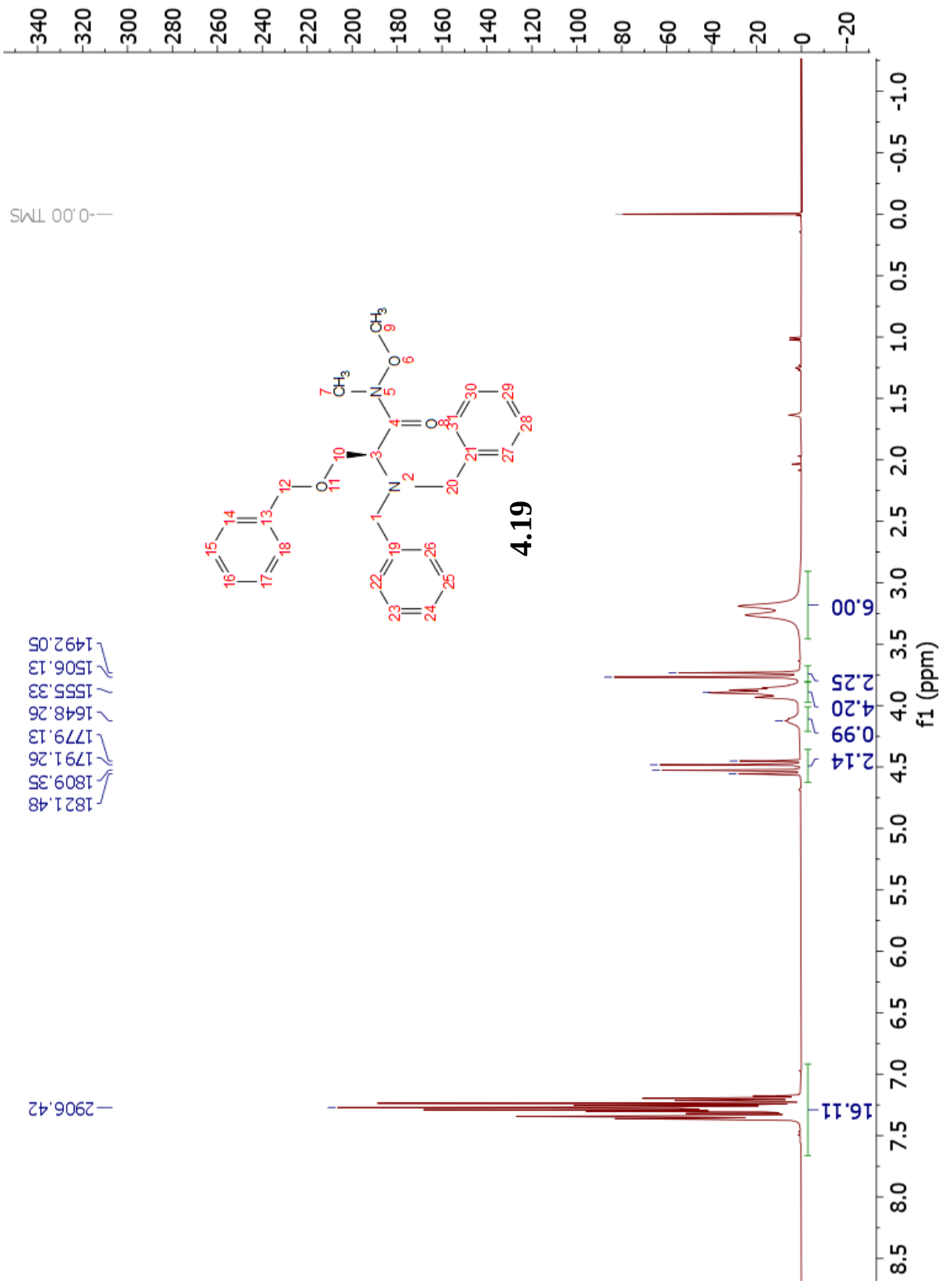

4.16
f1 (ppm)
9.06
1.66
7.39
2.92
1.18
0.19
2.00
1.01
0.77
0.22
0.79
0.72
3.04
1.89
0.73
0.18
562.91
673.53
710.76
1009.84
1136.94
1676.19
1771.55
1859.75
1965.34
2214.98
2271.59
2293.71
2554.79
2625.28
2686.93
4288.56
4290.63
4491.72
4493.79
4495.91
4504.88
4507.01
4509.13
3631.16 CDCl3

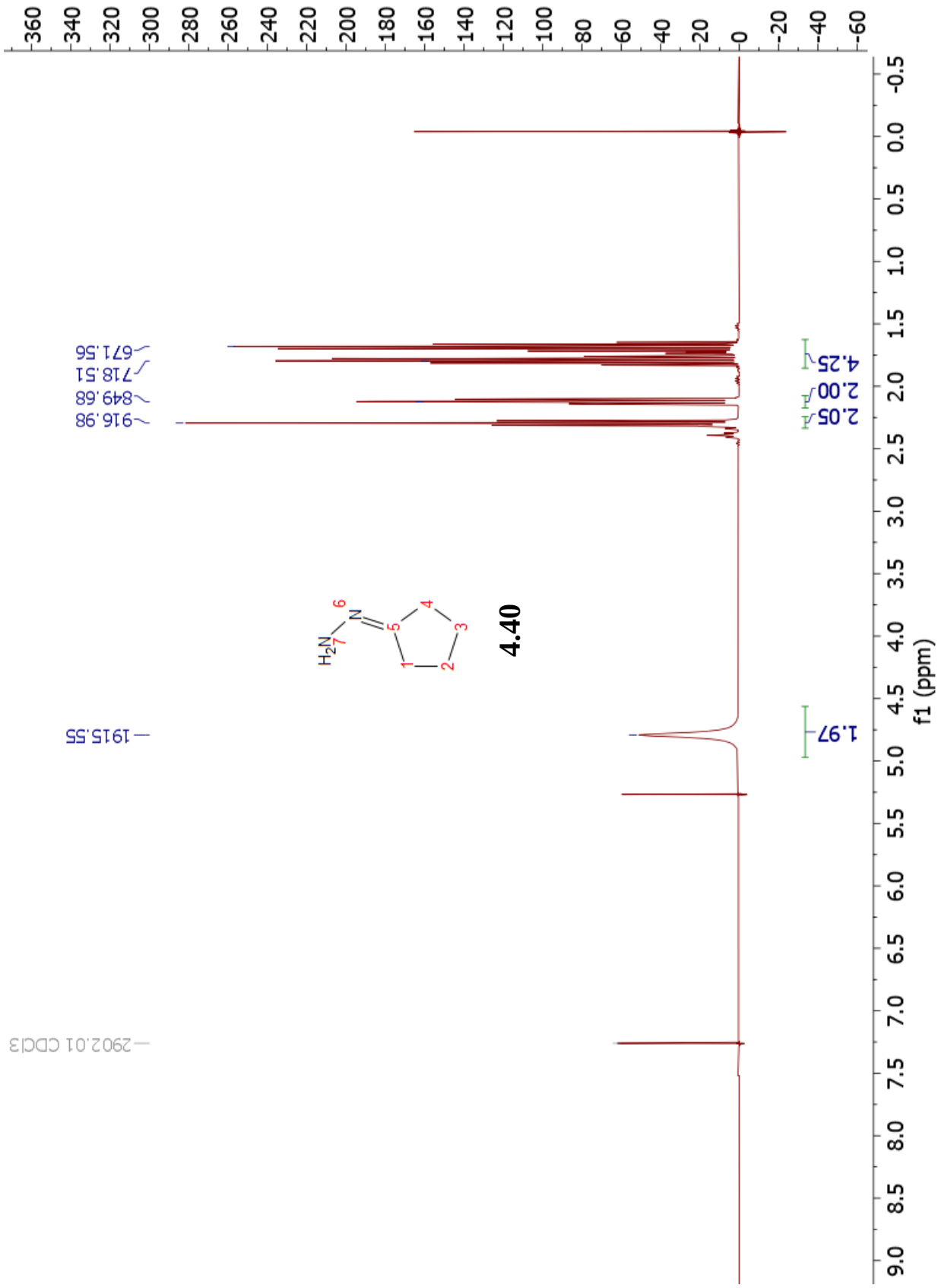

4.16
f1 (ppm)
0.00 TMS
2857.98
3155.11
3422.45
3438.85
3552.15
3567.24
3673.51
3876.75
5834.76
5941.07
6524.63
6565.85
7411.02
7417.00
7922.00
8060.04
8064.87
8161.14
9233.52
9247.34
9656.03 CDCl3
9688.05 CDCl3
9719.95 CDCl3
10008.78
10029.57
14889.82
14918.29
16057.31
16096.95
17614.89
17680.27
18698.51
18708.68
19465.85
19531.51
21444.34
21497.58
21581.39
21603.50

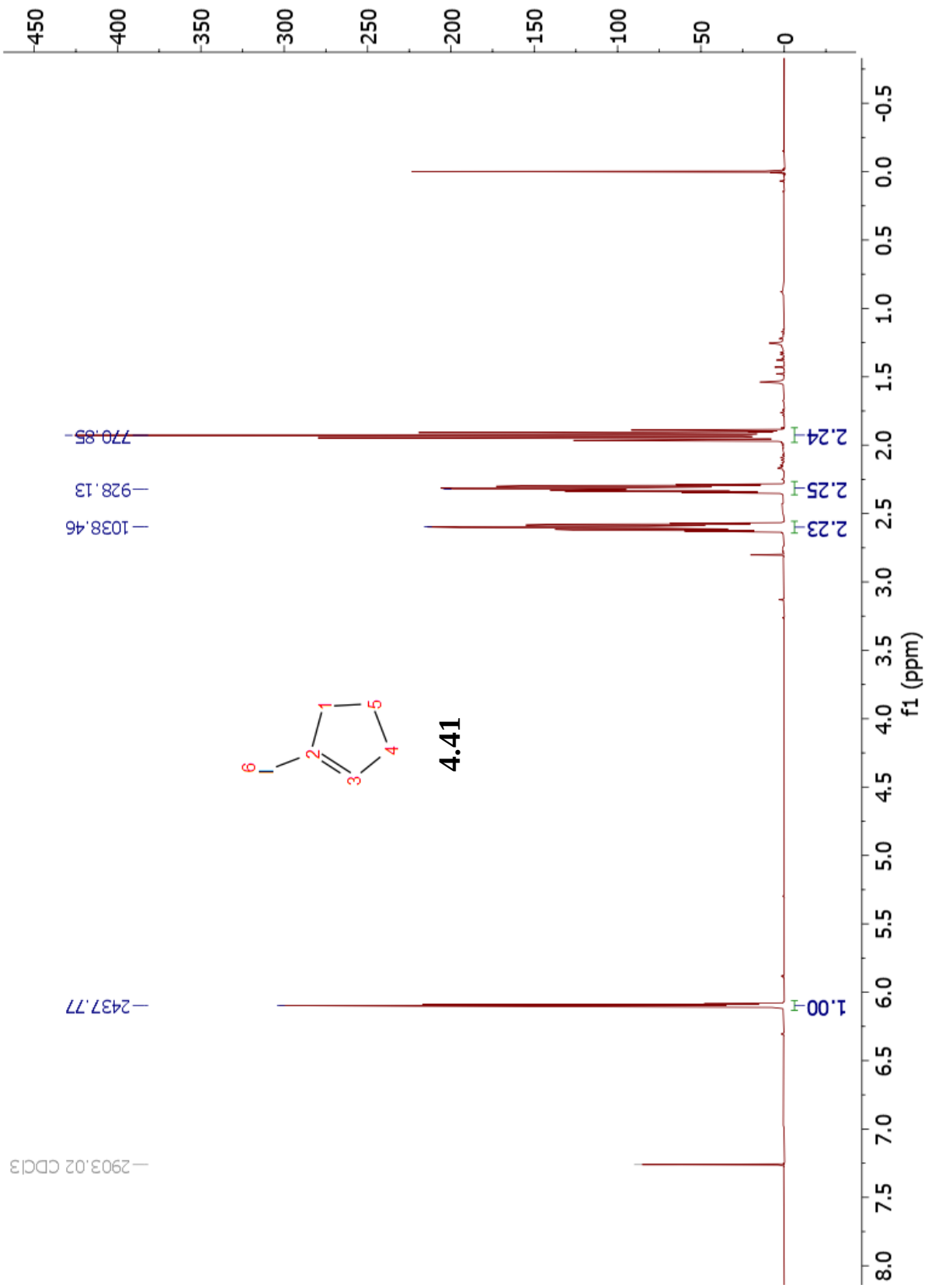

4538.35
4536.36
4534.27
4532.13
4356.27
4354.30
4352.34
2735.79
2395.77 D2O
2359.01
2354.55
2353.43
2350.29
2345.99
2224.54
2220.07
2217.52
2214.43
2213.36
2030.41
2026.44
2018.20
2015.09
1910.17
1358.56
1227.49
1057.35
4.17
0.90
1.94
2.01
0.98
0.91
0.94
3.02
1.03
2.91
f1 (ppm)

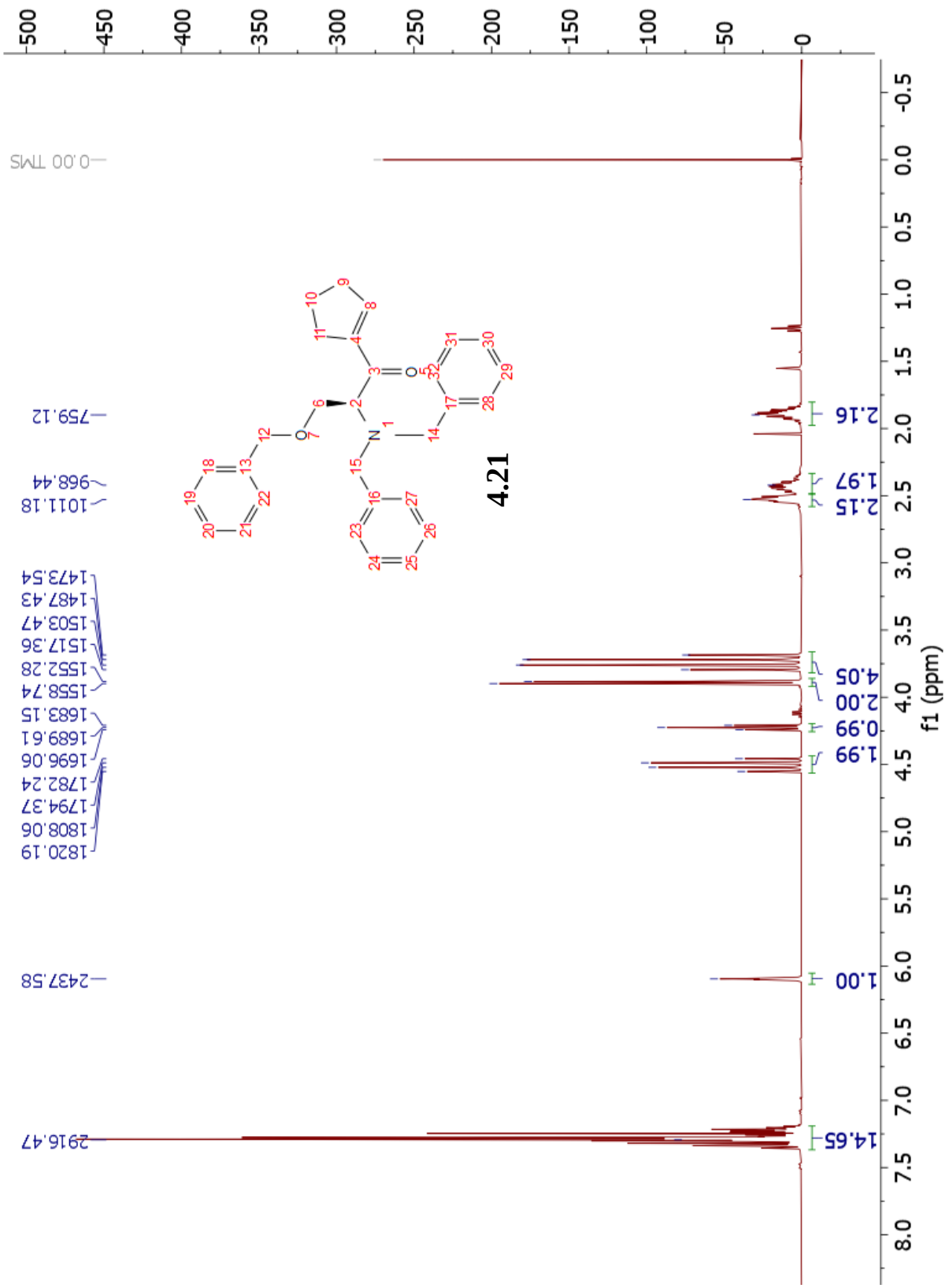

4.17
f1 (ppm)

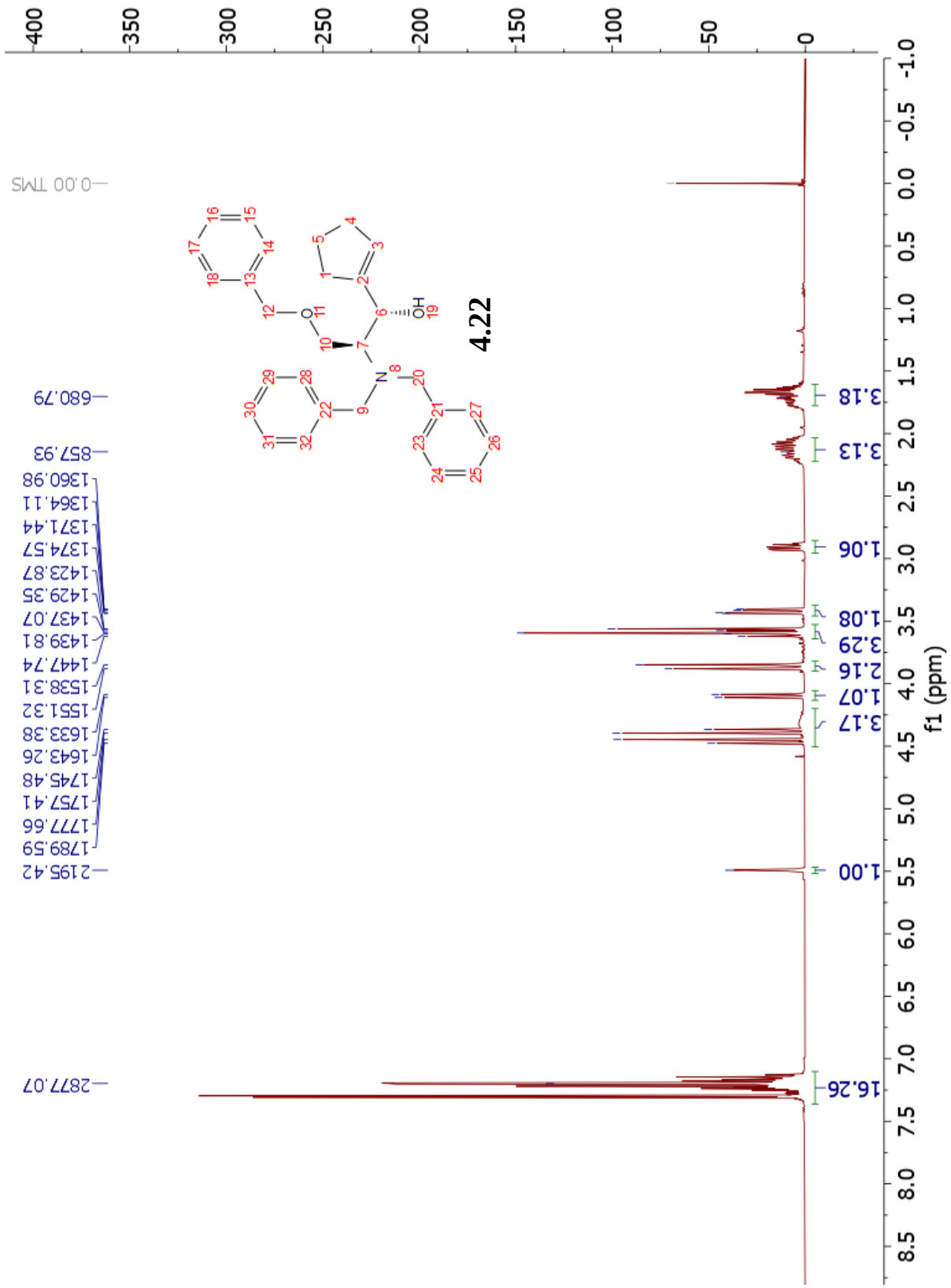

4.19
9.09
3.00
5.12
2.11
0.94
0.90
5.03
f1 (ppm)
571.04
1279.40
1479.72
1465.73
1787.24
1799.37
1815.90
1828.13
1949.22
2167.25
2176.83
2914.34
2903.03 CDCl3

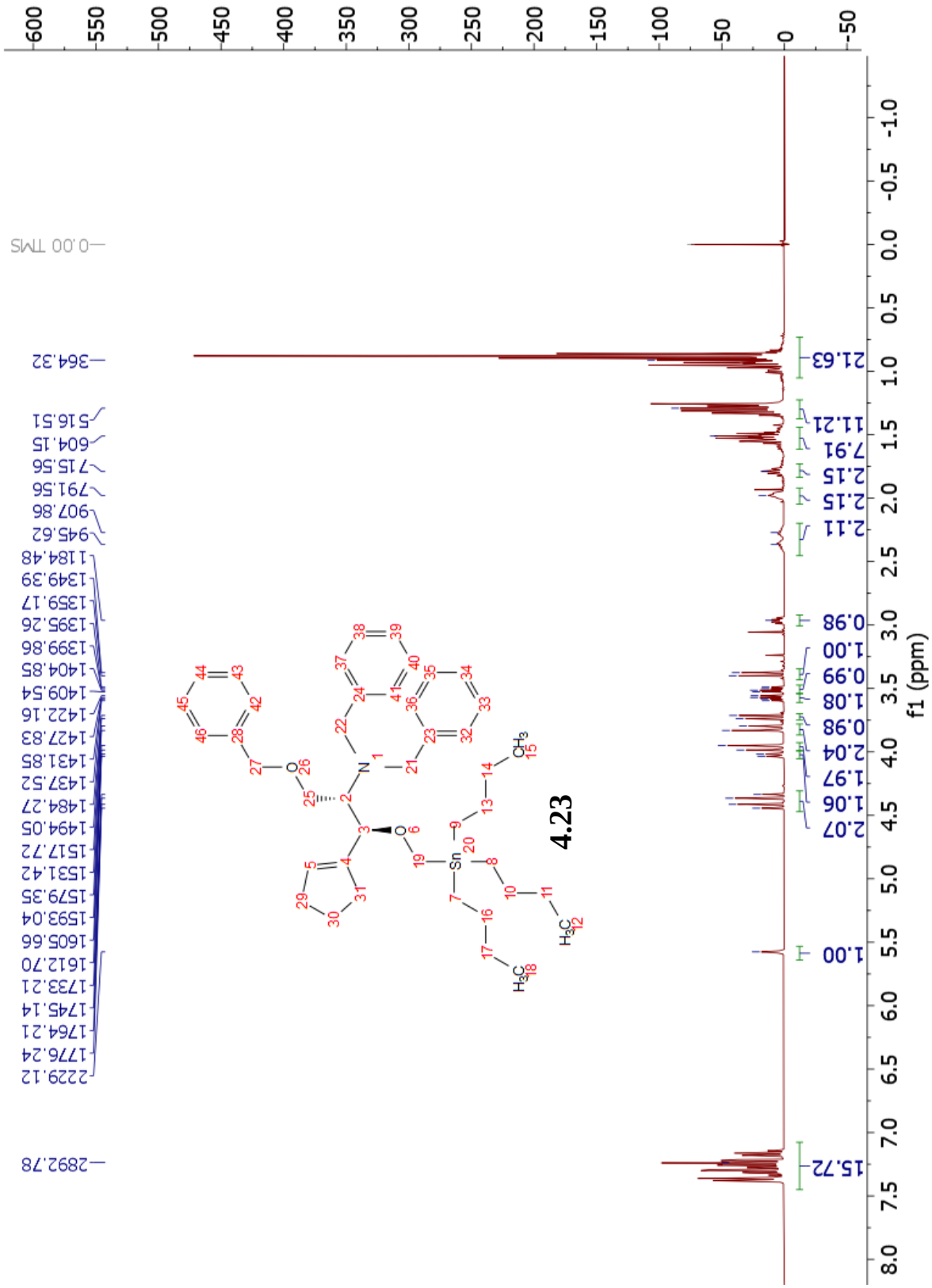

4.19
f1 (ppm)
16.11
2.14
0.99
4.20
2.25
6.00
2906.42
1821.48
1809.35
1791.26
1779.13
1648.26
1555.33
1506.13
1492.05
0.00 TMS

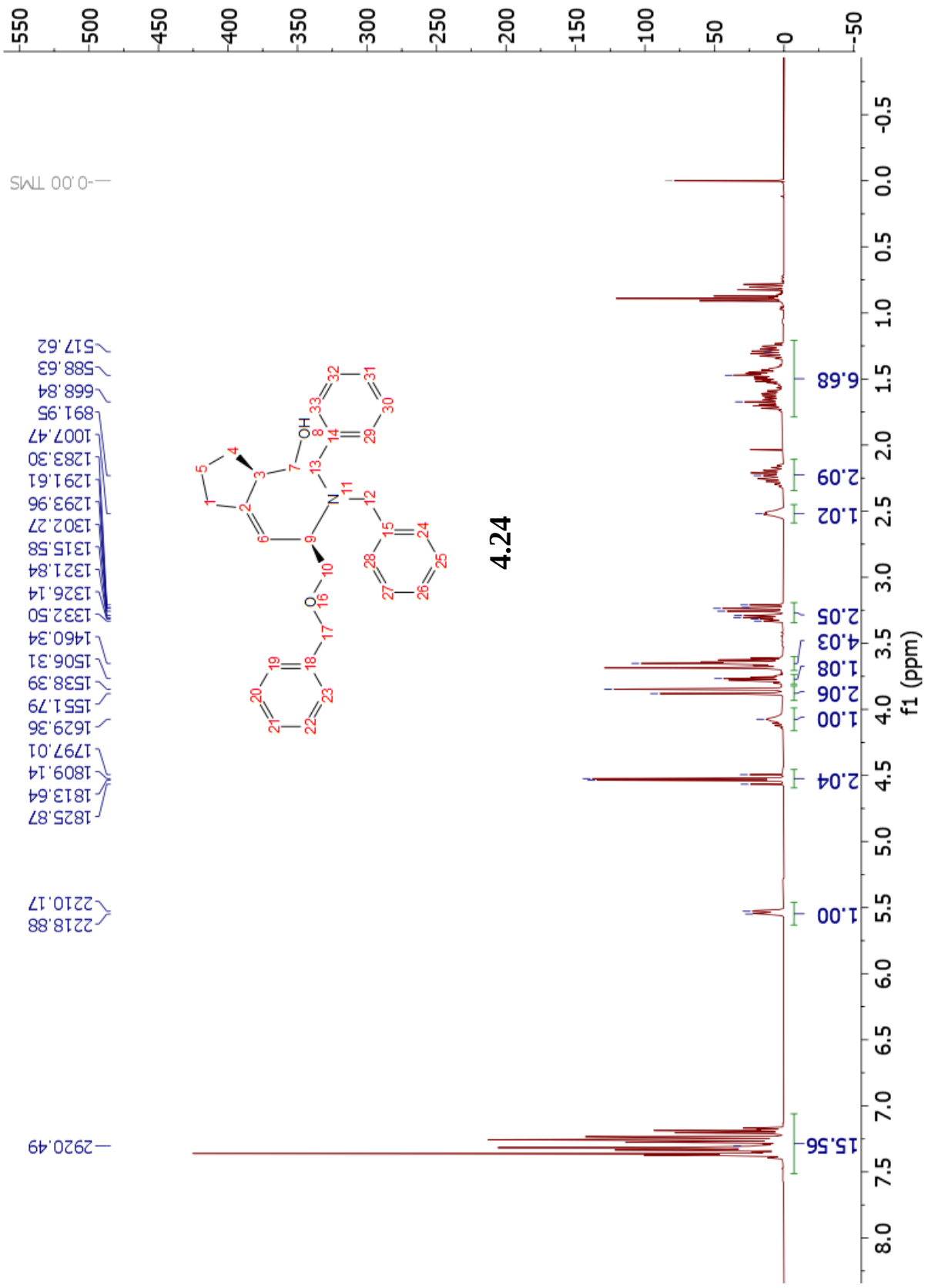

4.40
f1 (ppm)
1.97
2.05
2.00
4.25
916.98
849.68
718.51
671.56
1915.55
2902.01 CDCl3

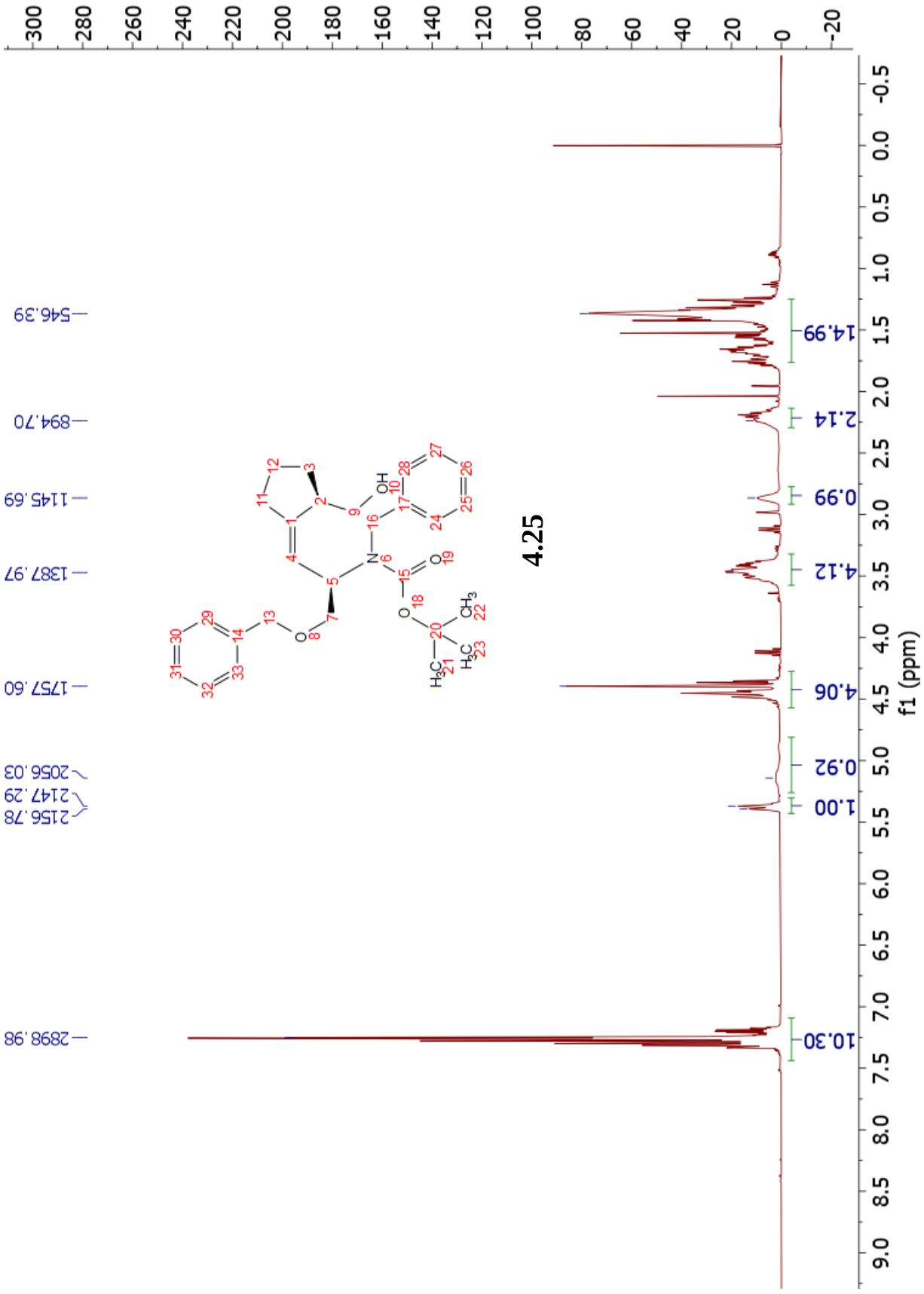

4.41
2903.02 CDCl3
2437.77
1038.46
928.13
770.85
1.00
2.23
2.25
2.24
f1 (ppm)

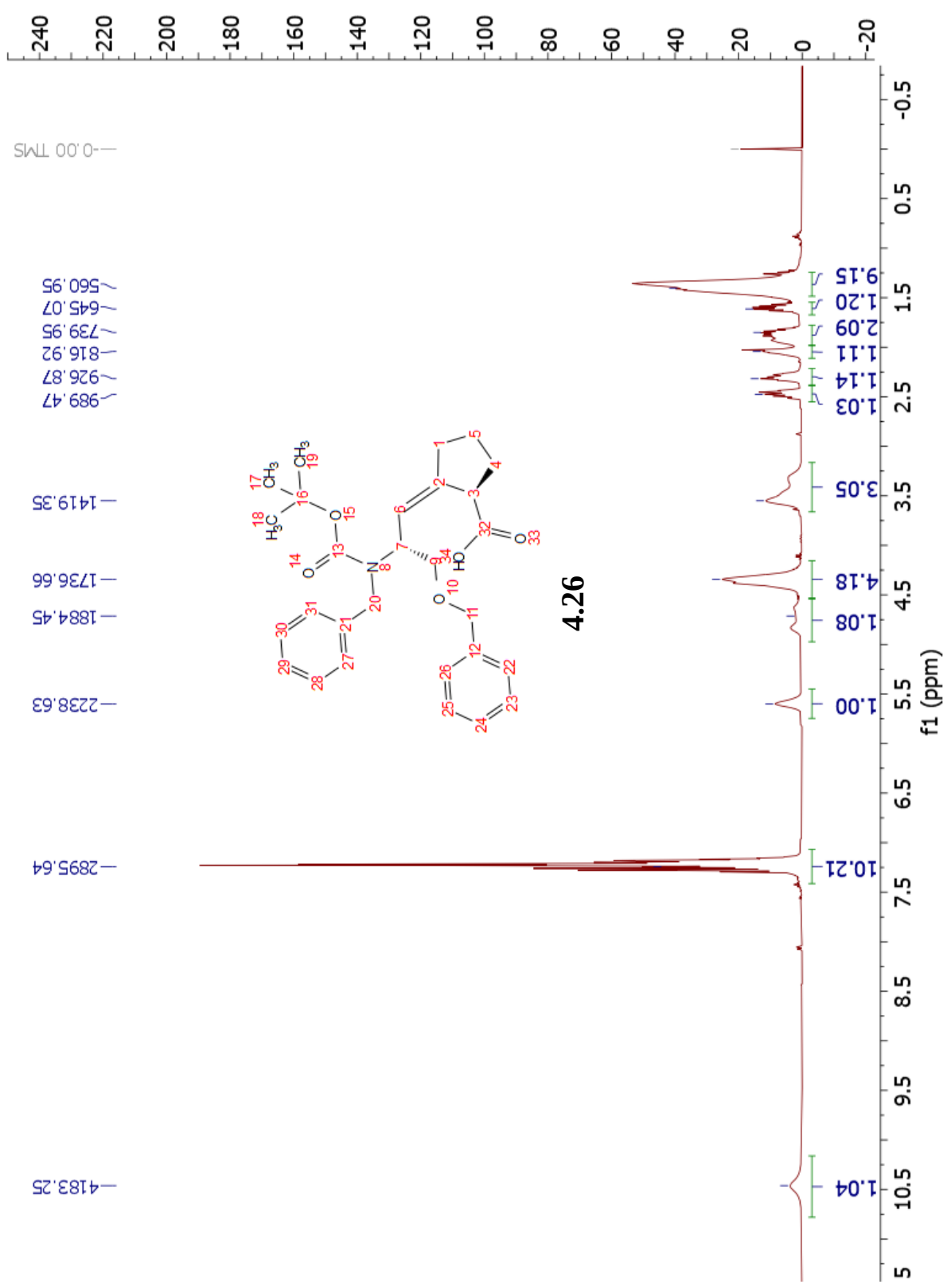

4.21
0.00 TMS
759.12
968.44
1011.18
1473.54
1487.43
1503.47
1517.36
1552.28
1558.74
1683.15
1689.61
1696.06
1782.24
1794.37
1808.06
1820.19
2437.58
2916.47
2.16
1.97
2.15
4.05
2.00
0.99
1.99
1.00
14.65
f1 (ppm)

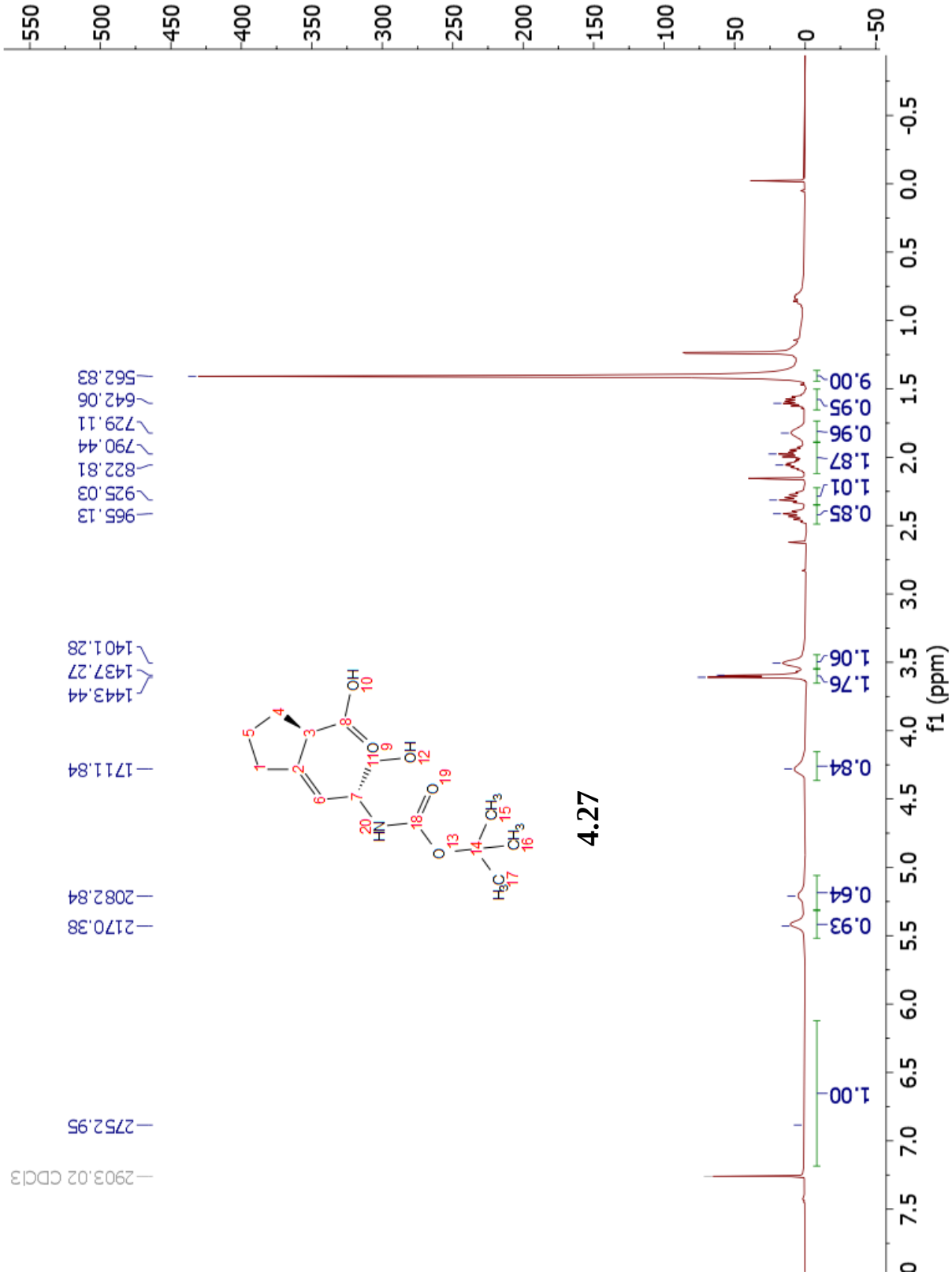

4.22
f1 (ppm)
3.18
3.13
1.06
1.08
3.29
2.16
1.07
3.17
1.00
16.26
0.00 TMS
2877.07
2195.42
1789.59
1777.66
1757.41
1745.48
1643.26
1633.38
1551.32
1538.31
1447.74
1439.81
1437.07
1429.35
1423.87
1374.57
1371.44
1364.11
1360.98
857.93
680.79

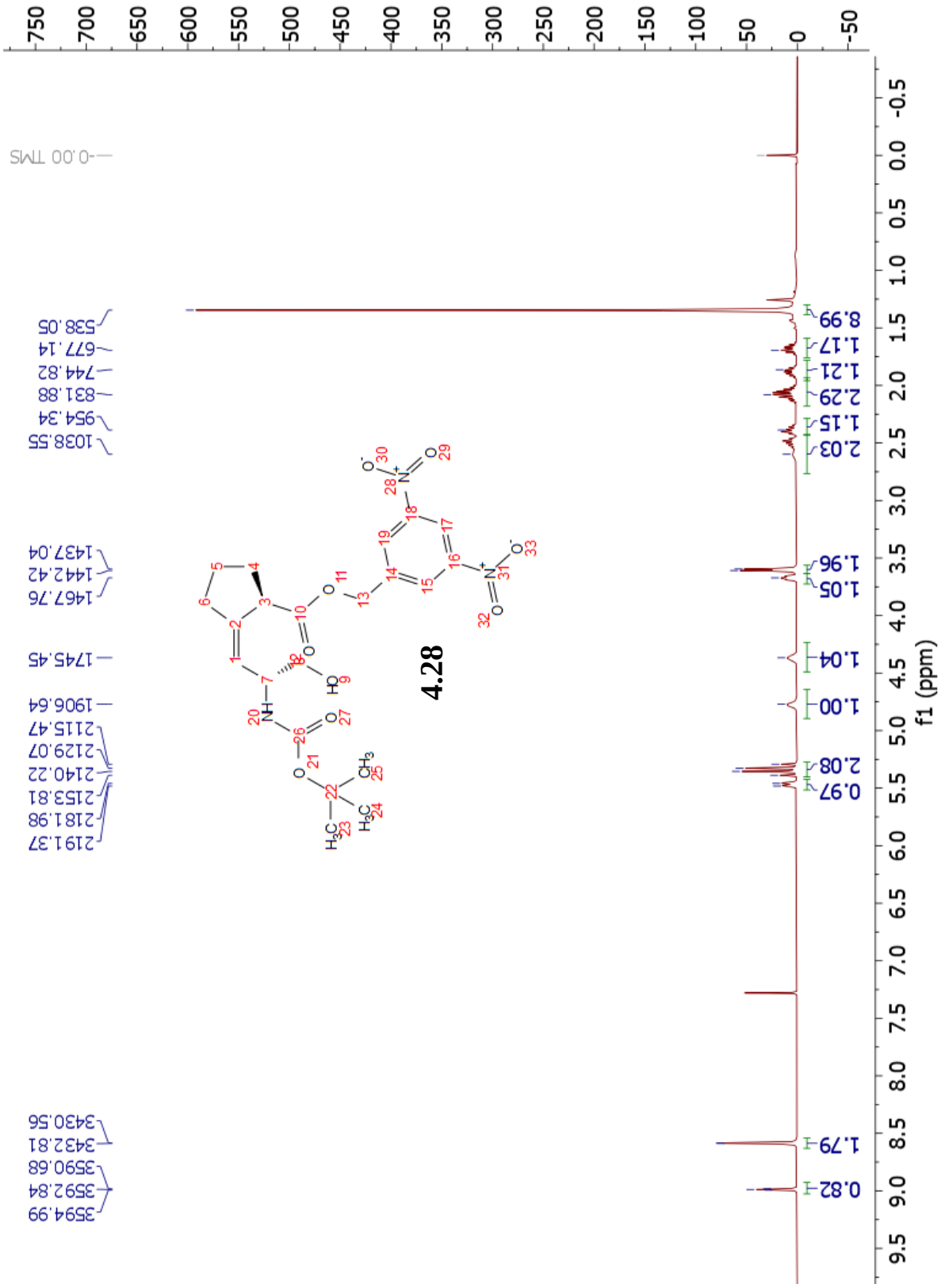

4.23
f1 (ppm)
TMS
0.00
364.32
516.51
604.15
715.56
791.56
907.86
945.62
1184.48
1349.39
1359.17
1395.26
1399.86
1404.85
1409.54
1422.16
1427.83
1431.85
1437.52
1484.27
1494.05
1517.72
1531.42
1579.35
1593.04
1605.66
1612.70
1733.21
1745.14
1764.21
1776.24
2229.12
2892.78
21.63
11.21
7.91
2.15
2.15
2.11
0.98
1.00
0.99
1.08
0.98
2.04
1.97
1.06
2.07
1.00
15.72

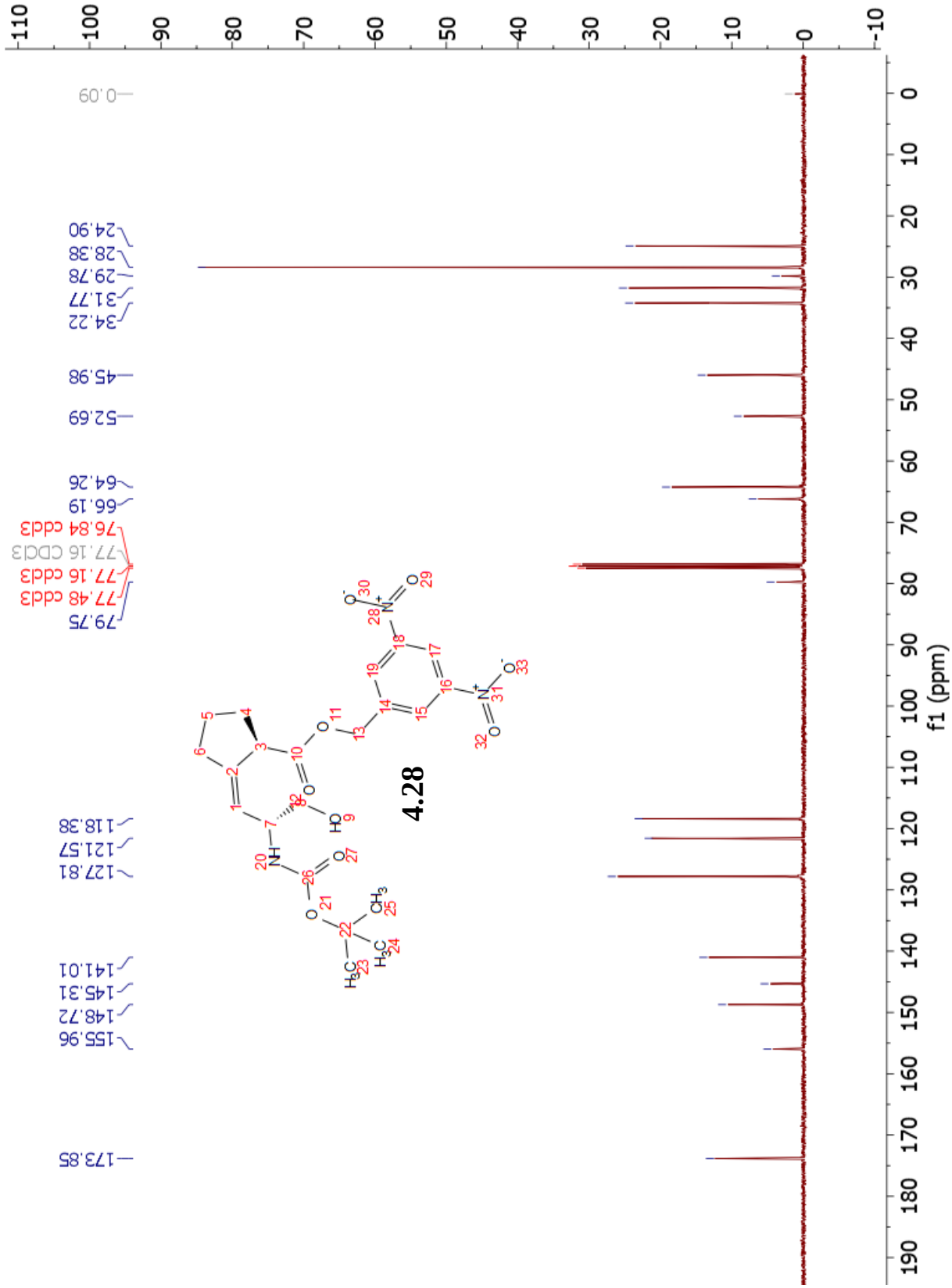

4.24
f1 (ppm)
15.56
1.00
2.04
1.00
2.06
1.08
4.03
2.05
1.02
2.09
6.68
TMS
0.00
2920.49
2218.88
2210.17
1825.87
1813.64
1809.14
1797.01
1629.36
1551.79
1538.39
1506.31
1460.34
1332.50
1326.14
1321.84
1315.58
1302.27
1293.96
1291.61
1283.30
1007.47
891.95
668.84
588.63
517.62

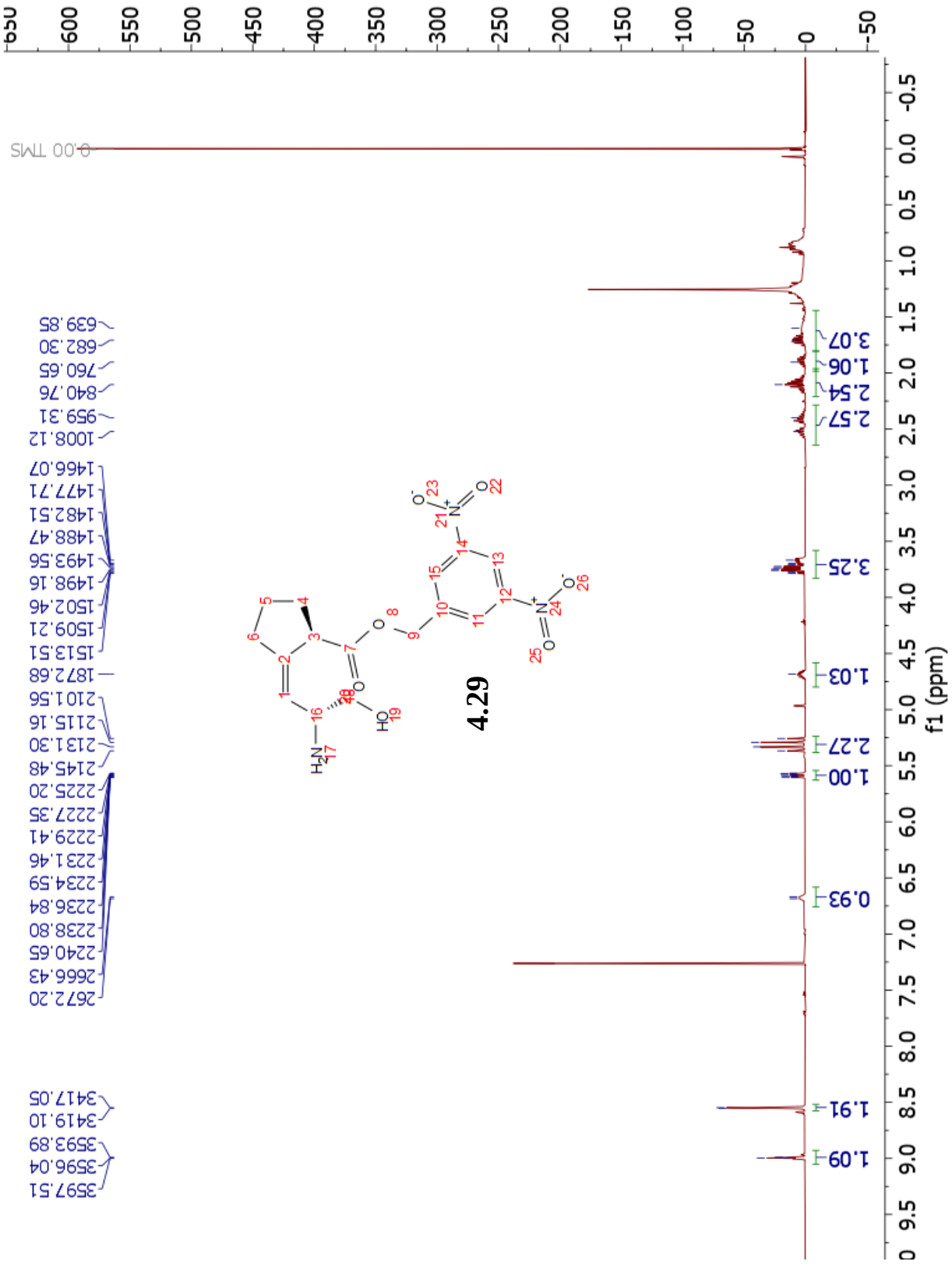

4.25
f1 (ppm)
10.30
1.00
0.92
4.06
4.12
0.99
2.14
14.99
2898.98
2156.78
2147.29
2056.03
1757.60
1387.97
1145.69
894.70
546.39

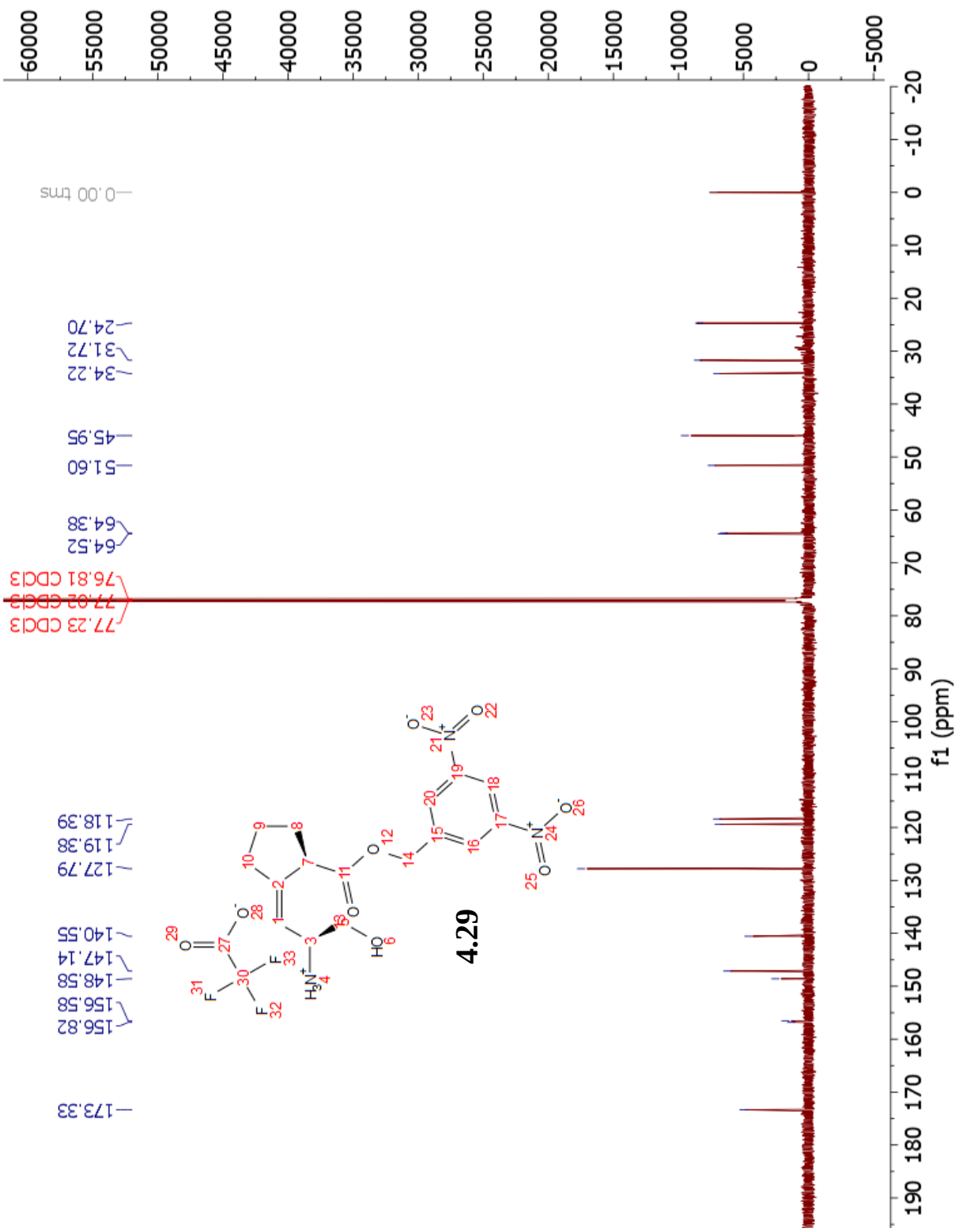

4.26
f1 (ppm)
9.15
1.20
2.09
1.11
1.14
1.03
3.05
4.18
1.08
1.00
10.21
1.04
0.00 TMS
560.95
645.07
739.95
816.92
926.87
989.47
1419.35
1736.66
1884.45
2238.63
2895.64
4183.25

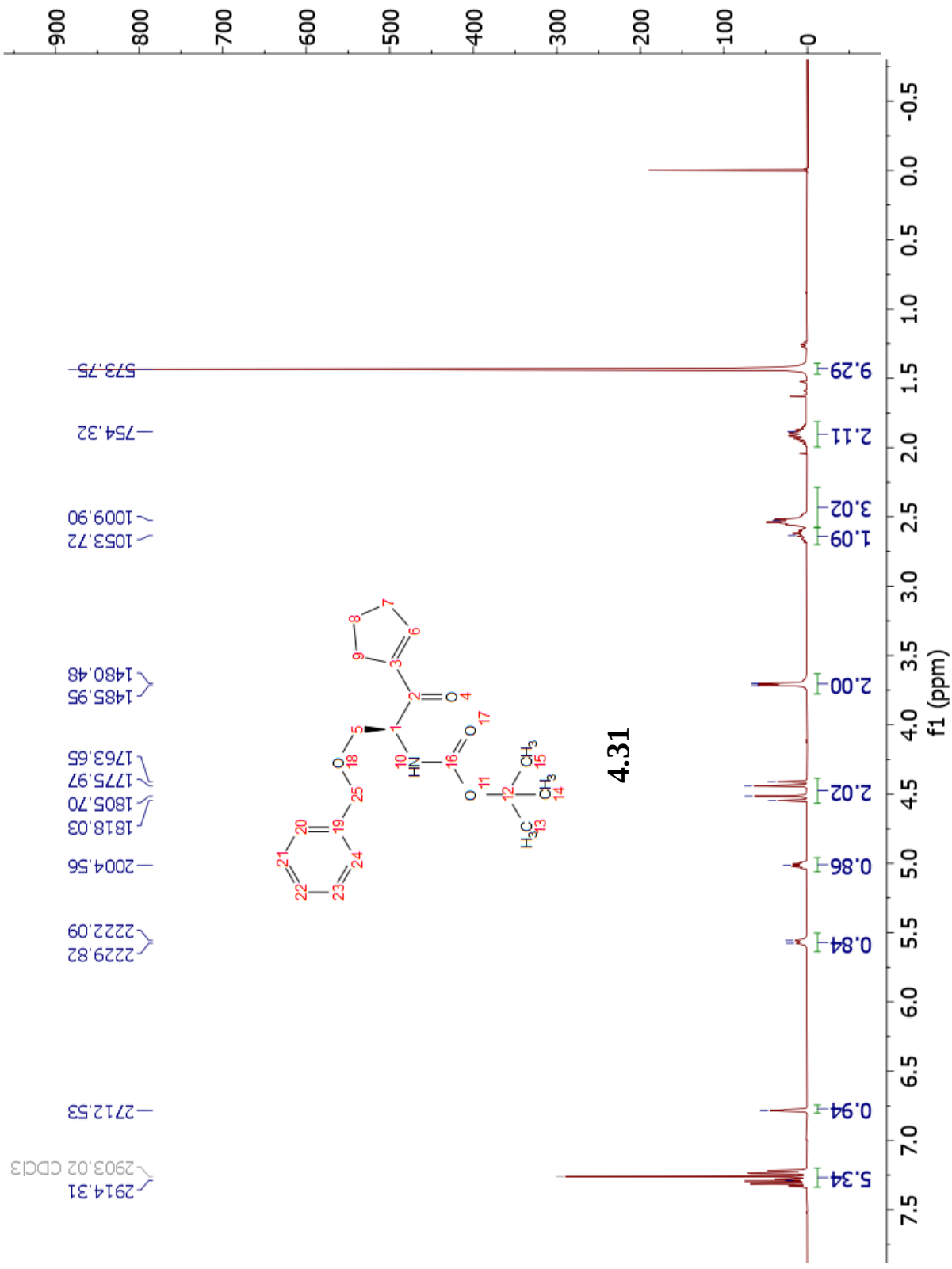

4.27
f1 (ppm)
9.00
0.95
0.96
1.87
1.01
0.85
1.06
1.76
0.84
0.64
0.93
1.00
562.83
642.06
729.11
790.44
822.81
925.03
965.13
1401.28
1437.27
1443.44
1711.84
2082.84
2170.38
2752.95
2903.02 CDCl3

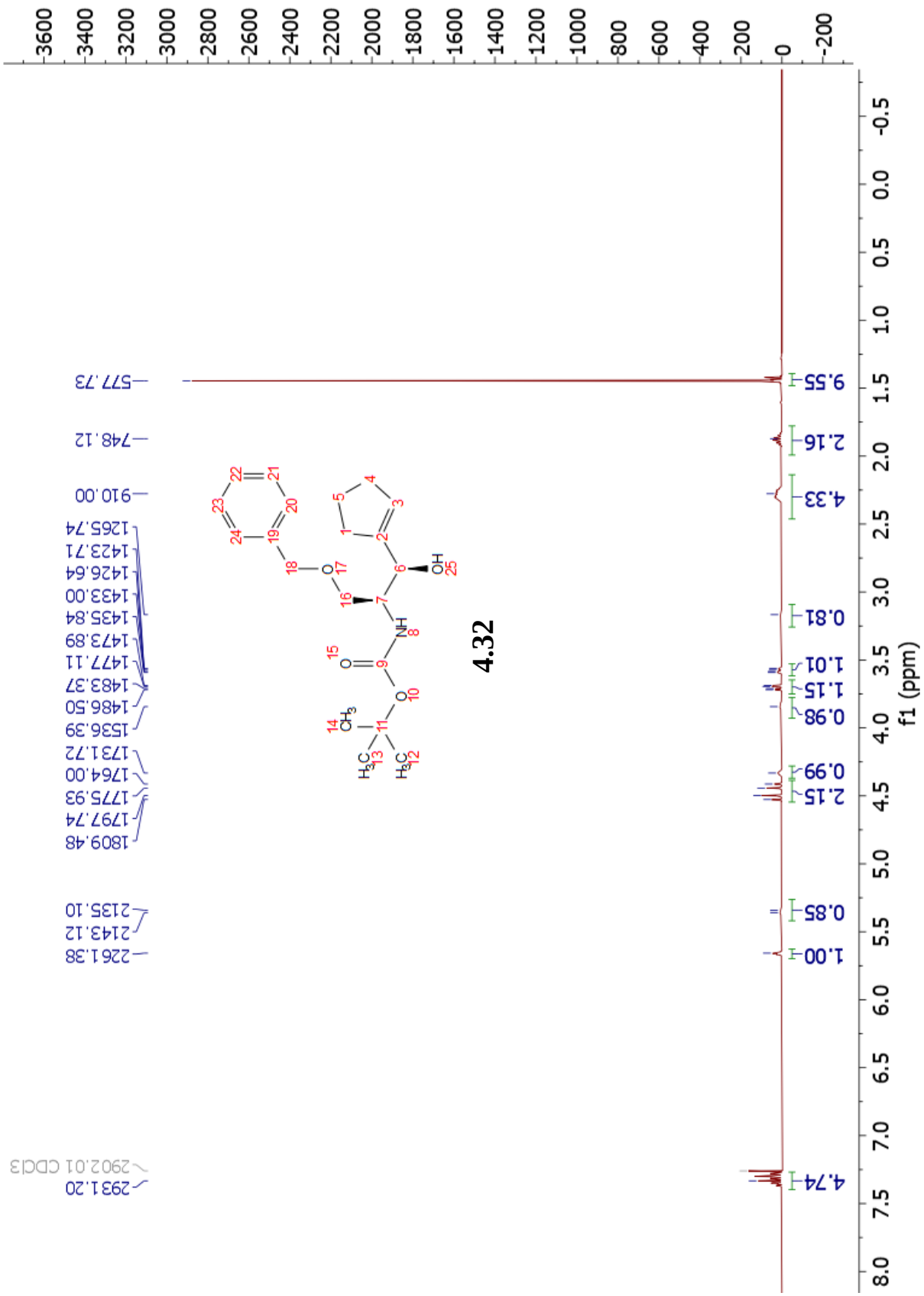

4.28
f1 (ppm)
8.99
1.17
1.21
2.29
1.15
2.03
1.96
1.05
1.04
1.00
2.08
0.97
1.79
0.82
TMS
538.05
677.14
744.82
831.88
954.34
1038.55
1437.04
1442.42
1467.76
1745.45
1906.64
2115.47
2129.07
2140.22
2153.81
2181.98
2191.37
3359.99
3592.84
3590.68
3432.81
3430.56

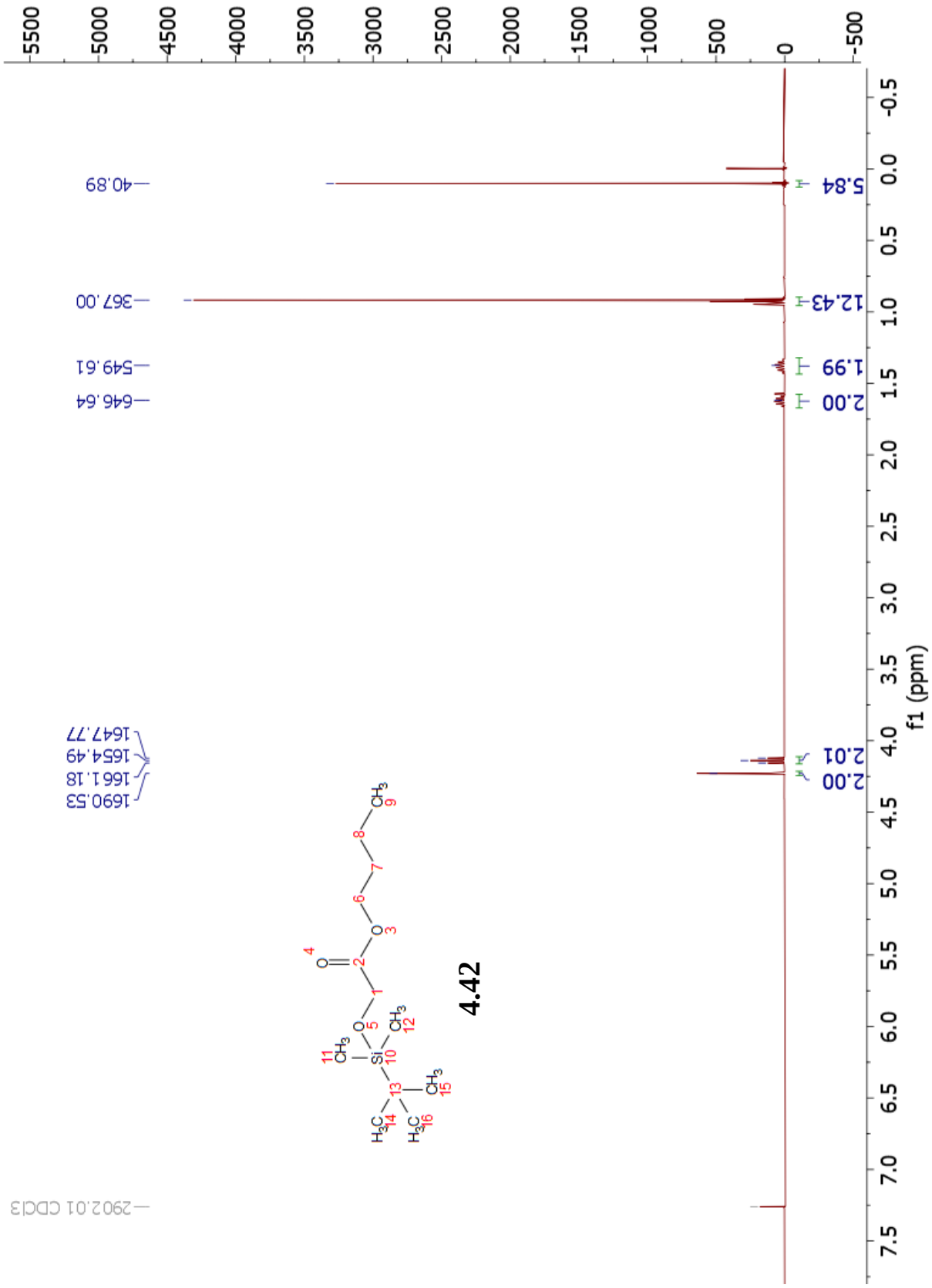

4.28
f1 (ppm)
0.09
24.90
28.38
29.78
31.77
34.22
45.98
52.69
64.26
66.19
76.84 cddi3
77.16 CDCl3
77.16 cddi3
77.48 cddi3
79.75
118.38
121.57
127.81
141.01
145.31
148.72
155.96
173.85

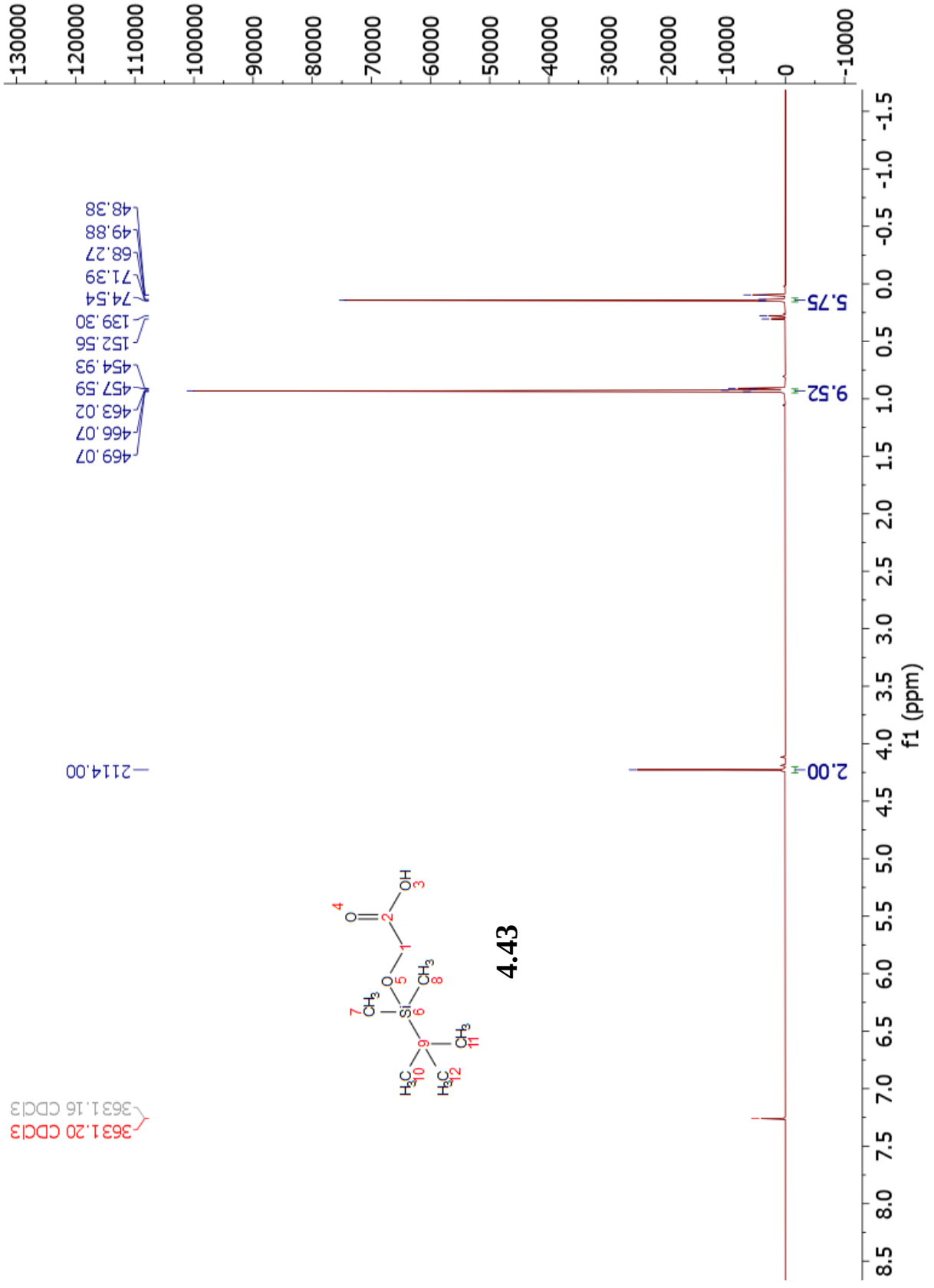

4.29
f1 (ppm)
1.09
1.91
0.93
1.00
2.27
1.03
3.25
2.57
2.54
1.06
3.07
0.00 TMS

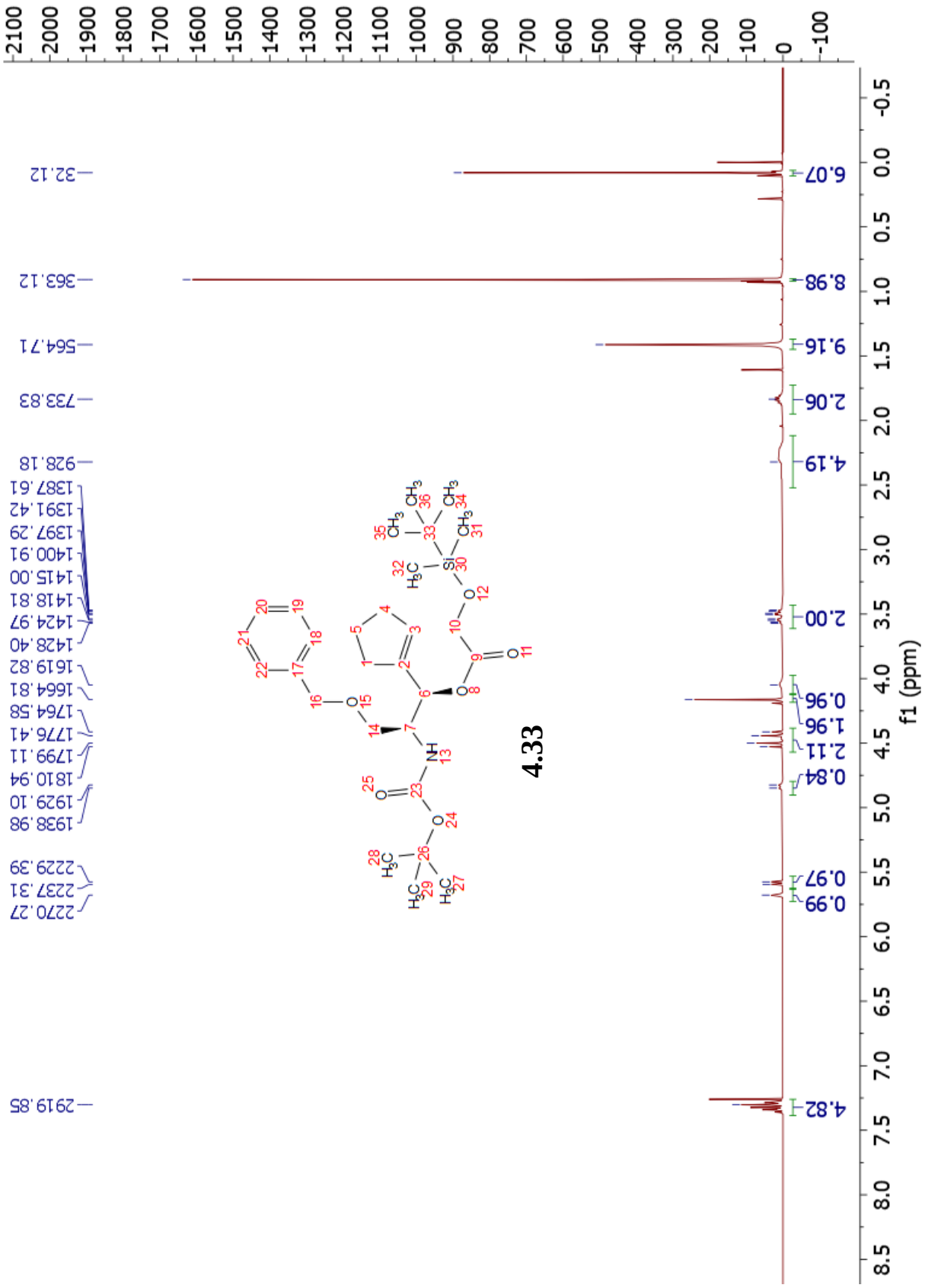

4.29
f1 (ppm)
0.00 tms
24.70
31.72
34.22
45.95
51.60
64.38
64.52
76.81 CDCl3
77.02 CDCl3
77.23 CDCl3
118.39
119.38
127.79
140.55
147.14
148.58
156.58
156.82
173.33

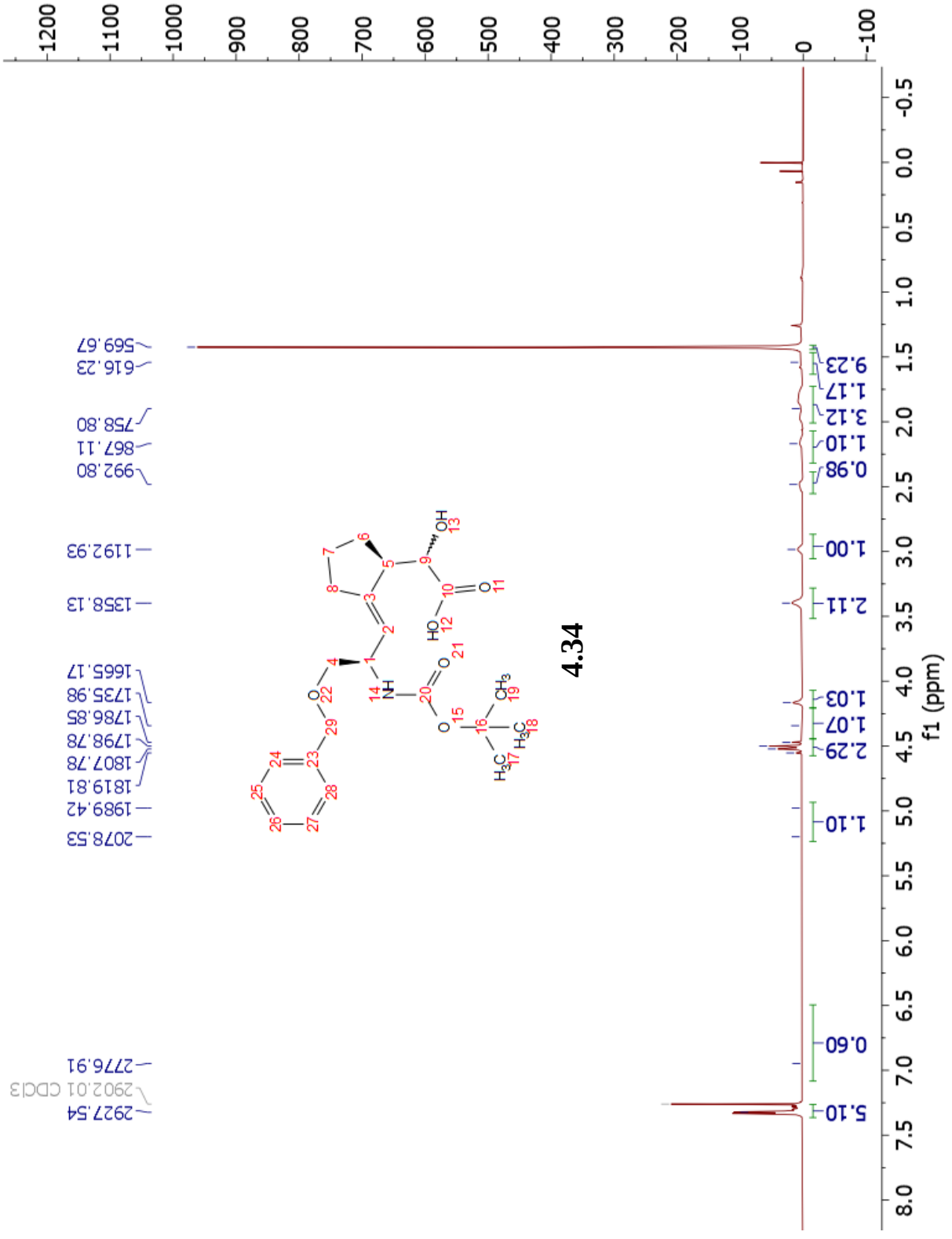

4.31
f1 (ppm)
9.29
2.11
3.02
1.09
2.00
2.02
0.86
0.84
0.94
5.34
573.75
754.32
1053.72
1009.90
1485.95
1480.48
1818.03
1805.70
1775.97
1763.65
2004.56
2229.82
2222.09
2712.53
2914.31
2903.02 CDCl3

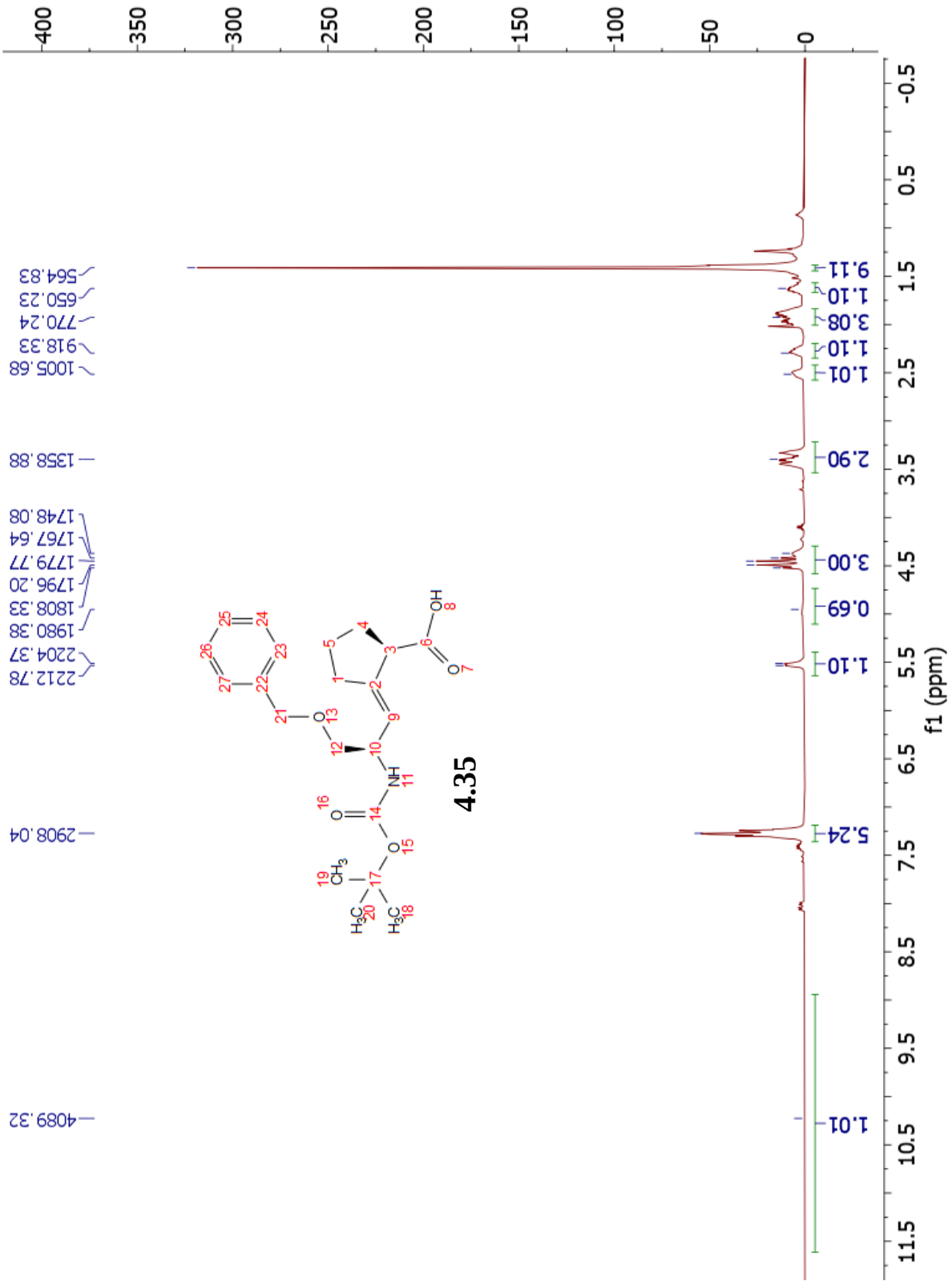

4.32
f1 (ppm)
9.55
2.16
4.33
0.81
1.01
1.15
0.98
0.99
2.15
0.85
1.00
4.74
577.73
748.12
910.00
1265.74
1423.71
1426.64
1433.00
1435.84
1473.89
1477.11
1483.37
1486.50
1536.39
1731.72
1764.00
1775.93
1797.74
1809.48
2135.10
2143.12
2261.38
2931.20
2902.01 CDCl3

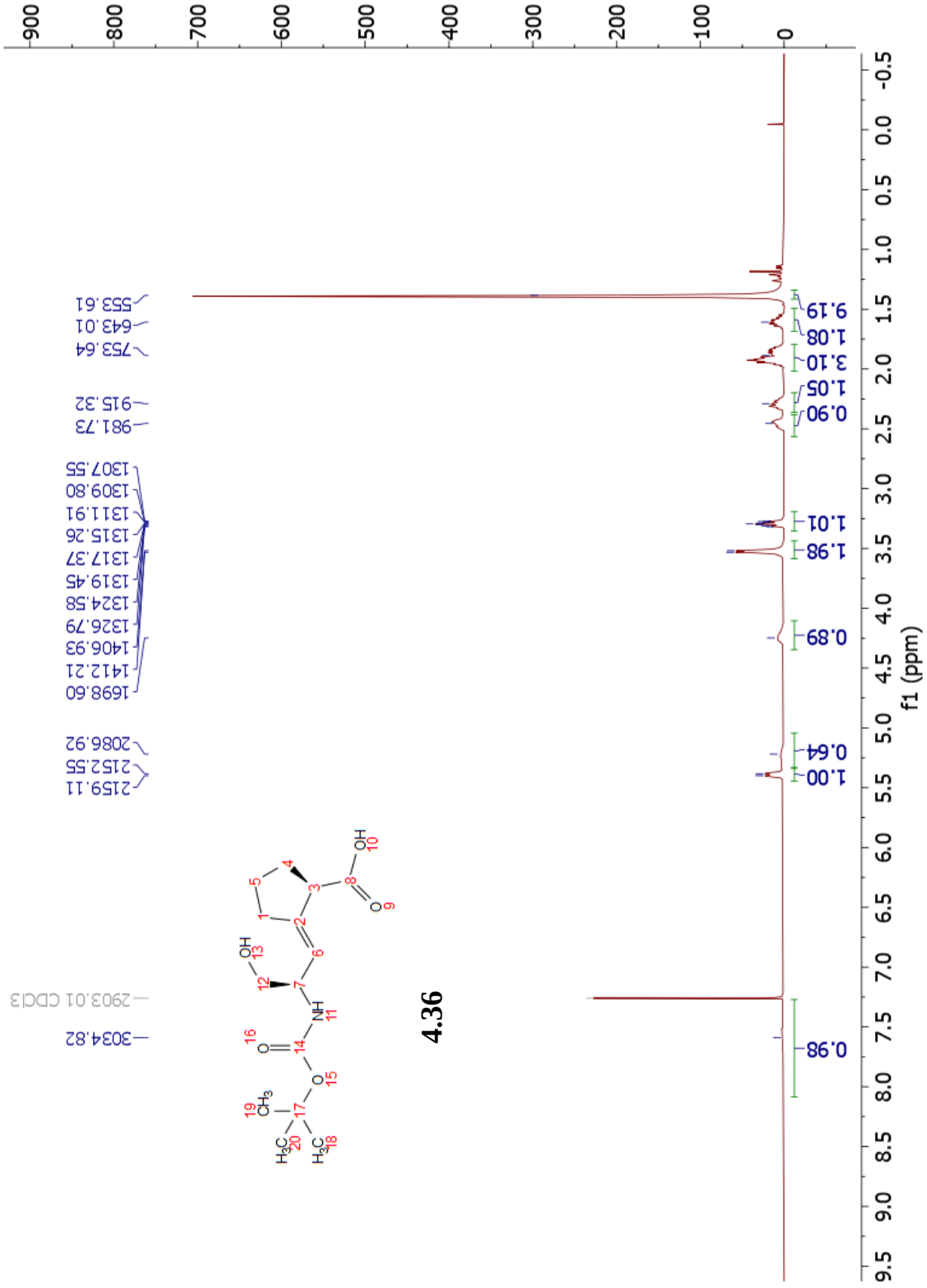

4.42
f1 (ppm)
5.84
12.43
1.99
2.00
2.01
2.00
40.89
367.00
549.61
646.64
1690.53
1661.18
1654.49
1647.77
2902.01 CDCl3

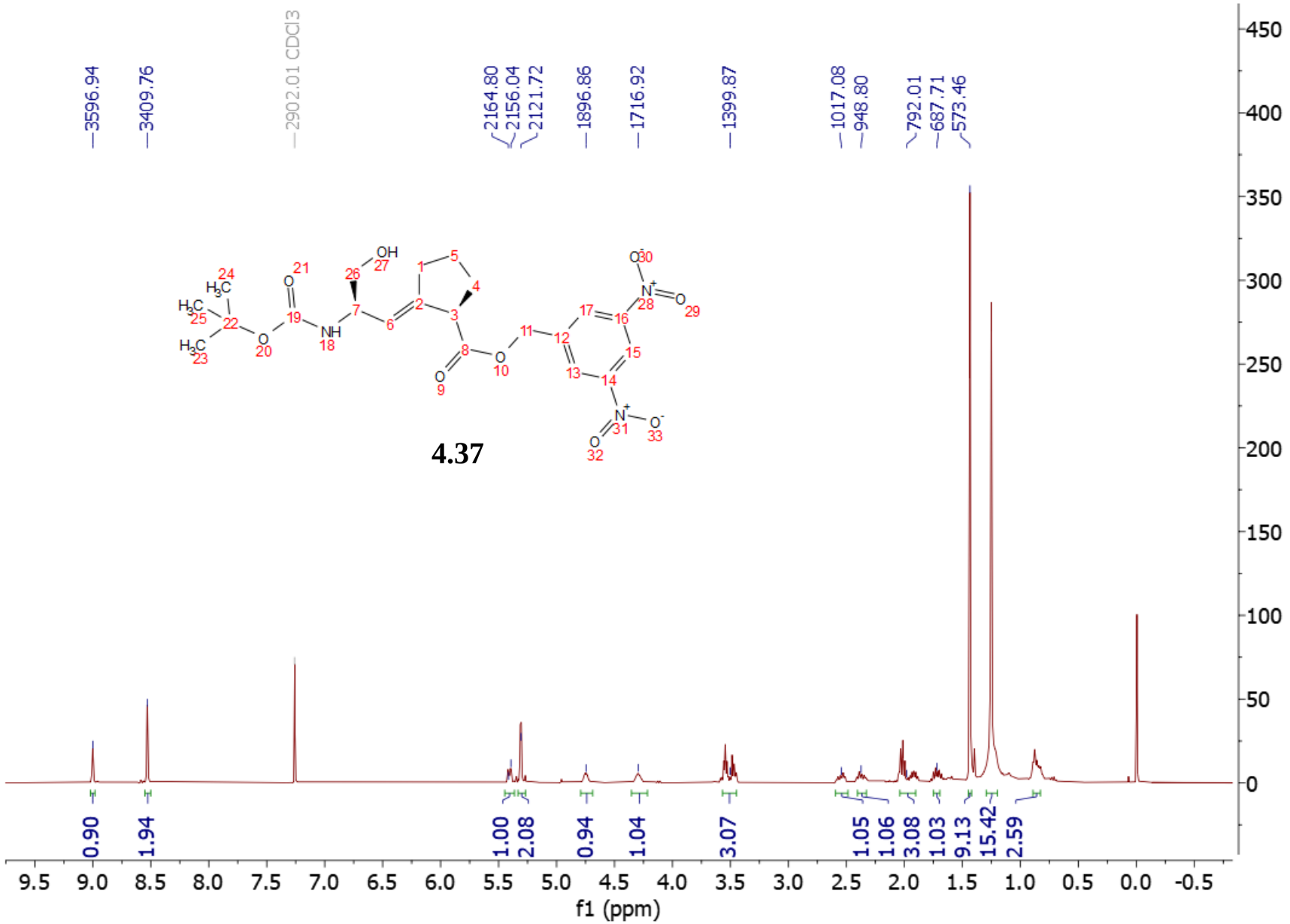

4.43
2.00
9.52
5.75
2114.00
469.07
466.07
463.02
457.59
454.93
152.56
139.30
74.54
71.39
68.27
49.88
48.38
3631.20 CDCl3
3631.16 CDCl3
f1 (ppm)

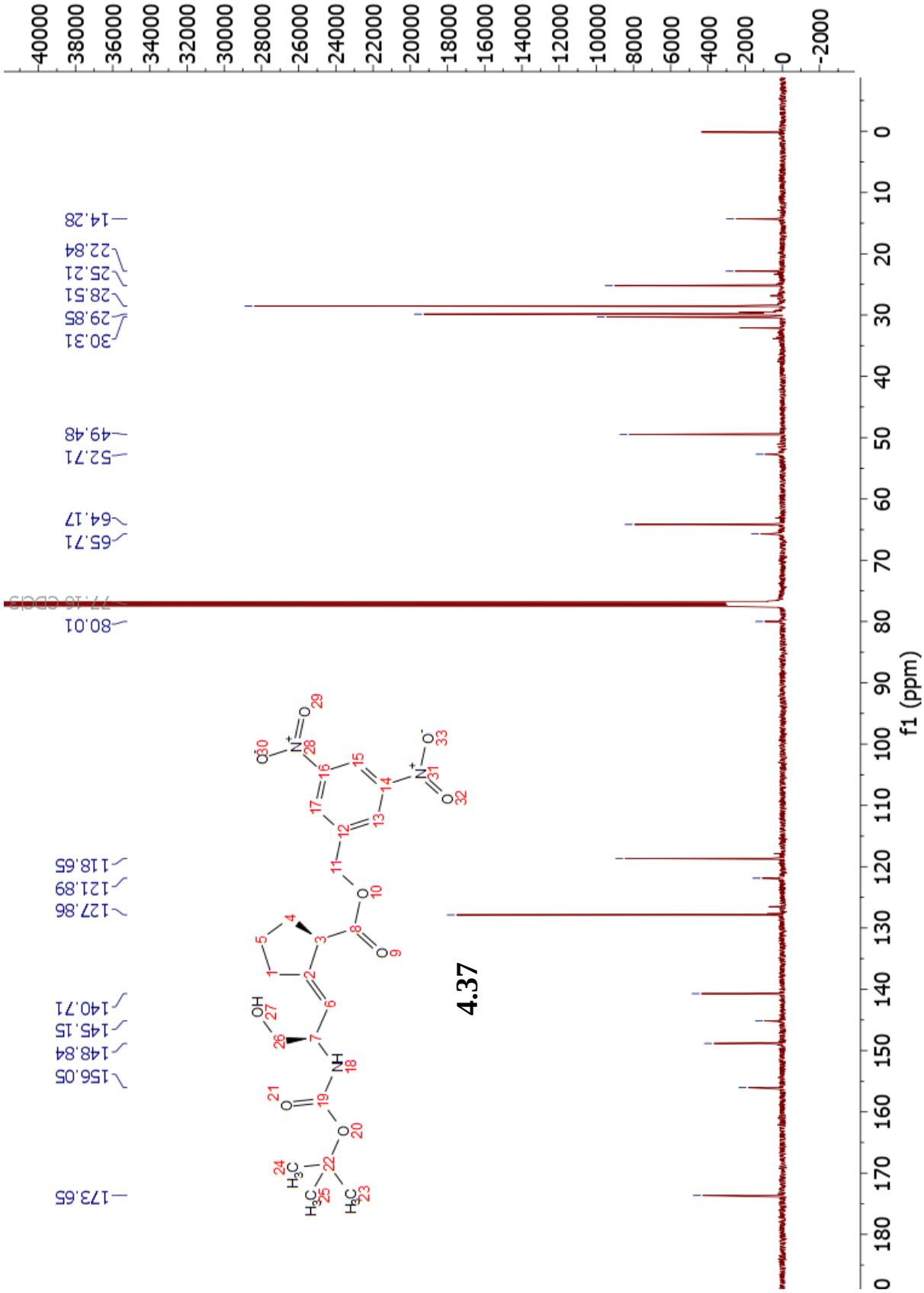

4.33
f1 (ppm)

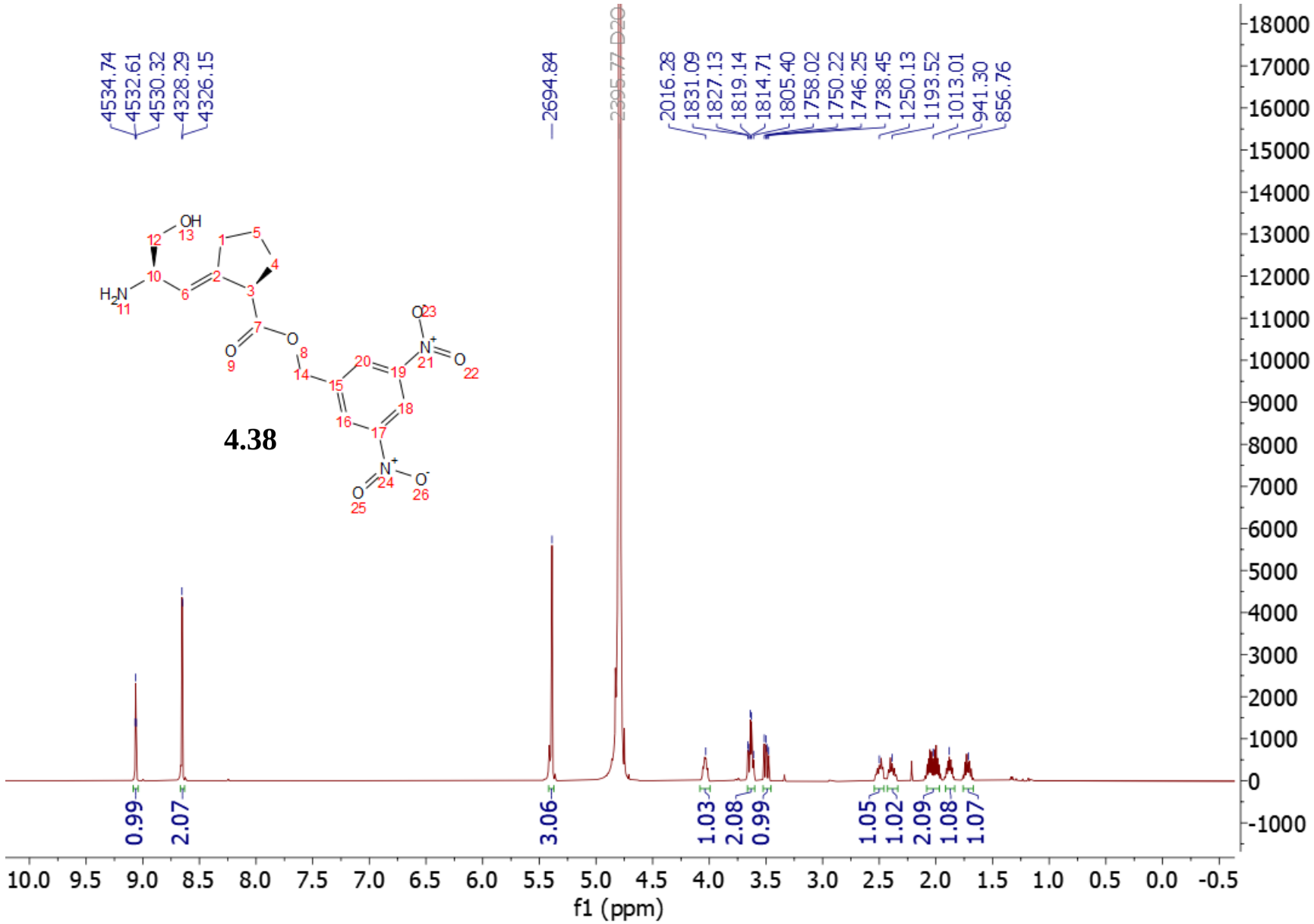

4.34
f1 (ppm)
9.23
1.17
3.12
1.10
0.98
1.00
2.11
1.03
1.07
2.29
1.10
0.60
5.10
569.67
616.23
758.80
867.11
992.80
1192.93
1358.13
1665.17
1735.98
1786.85
1798.78
1807.78
1819.81
1989.42
2078.53
2776.91
2902.01 CDCl3
2927.54

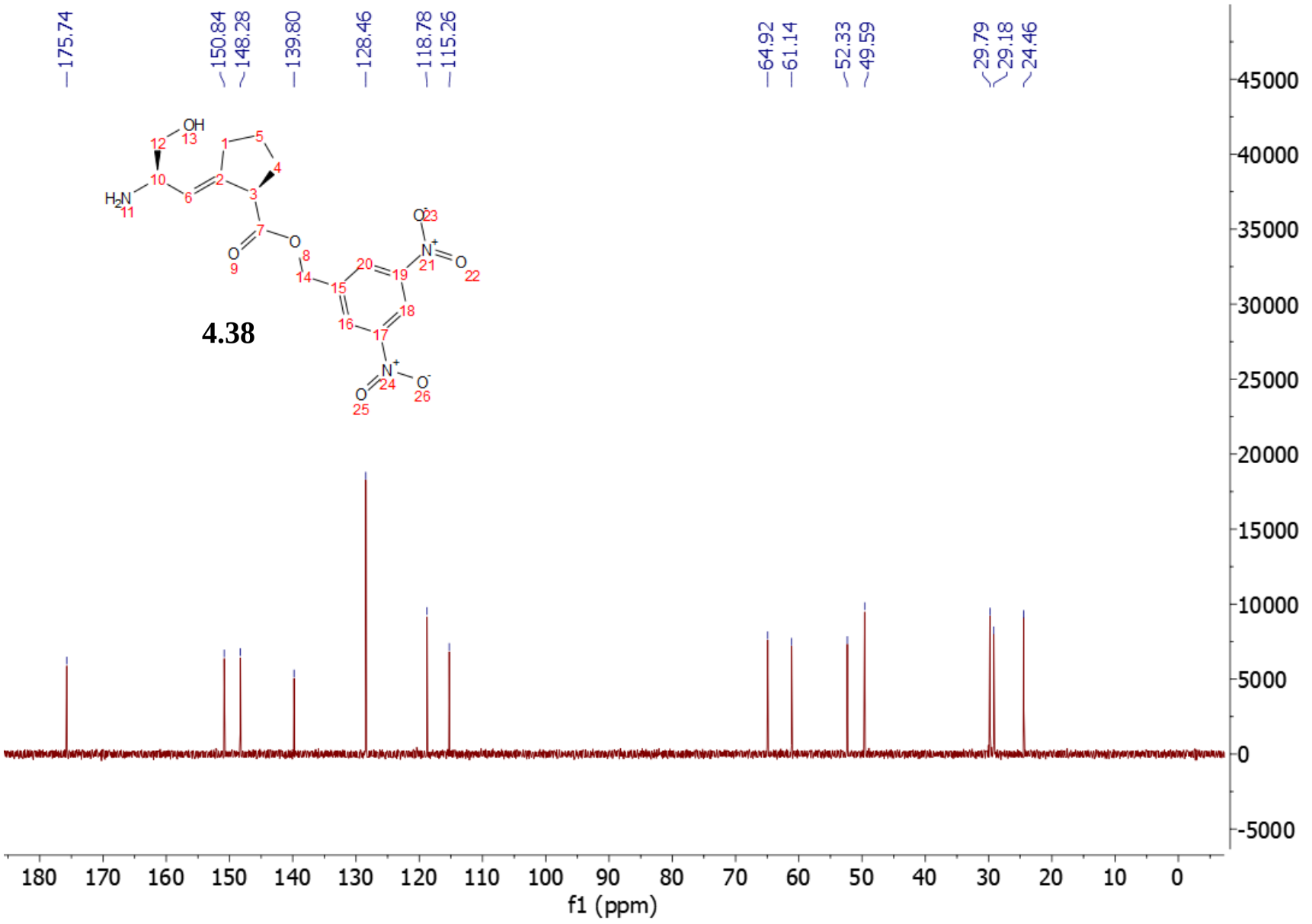

4.35
f1 (ppm)
9.11
1.10
3.08
1.10
1.01
2.90
3.00
0.69
1.10
5.24
1.01
564.83
650.23
770.24
918.33
1005.68
1358.88
1748.08
1767.64
1779.77
1796.20
1808.33
1980.38
2204.37
2212.78
2908.04
4089.32


4.36
f1 (ppm)
0.98
1.00
0.64
0.89
1.98
1.01
0.90
1.05
3.10
1.08
9.19
2903.01 CDCl3
3034.82
2159.11
2152.55
2086.92
1698.60
1412.21
1406.93
1326.79
1324.58
1319.45
1317.37
1315.26
1311.91
1309.80
1307.55
981.73
915.32
753.64
643.01
553.61


3596.94
3409.76
2902.01 CDCl3
2164.80
2156.04
2121.72
1896.86
1716.92
1399.87
1017.08
948.80
792.01
687.71
573.46
4.37
0.90
1.94
1.00
2.08
0.94
1.04
3.07
1.05
1.06
3.08
1.03
9.13
15.42
2.59
f1 (ppm)


14.28
22.84
25.21
28.51
29.85
30.31
49.48
52.71
64.17
65.71
77.16 CDCl3
80.01
118.65
121.89
127.86
140.71
145.15
148.84
156.05
173.65
4.37
f1 (ppm)


4534.74
4532.61
4530.32
4328.29
4326.15
2694.84
2395.77 D2O
2016.28
1831.09
1827.13
1819.14
1814.71
1805.40
1758.02
1750.22
1746.25
1738.45
1250.13
1193.52
1013.01
941.30
856.76
4.38
0.99
2.07
3.06
1.03
2.08
0.99
1.05
1.02
2.09
1.08
1.07
f1 (ppm)


175.74
150.84
148.28
139.80
128.46
118.78
115.26
64.92
61.14
52.33
49.59
29.79
29.18
24.46
OH
12
13
1
5
4
10
2
3
H2N
11
6
7
O
9
O
8
14
15
20
19
21
N+
O
22
O-
23
18
16
17
N+
O
24
O
25
O-
26
4.38
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)