PEPTIDE AND PROTEIN ASSEMBLY USING THE STAUDINGER LIGATION

by

Bradley L. Nilsson

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2003

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Committee's Page. This page is not to be hand-written except for the signatures

A dissertation entitled

PEPTIDE AND PROTEIN ASSEMBLY USING

THE STAUDINGER LIGATION

submitted to the Graduate School of the
University of Wisconsin-Madison
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

by

Bradley Linden Nilsson

Date of Final Oral Examination: December 11, 2003

Month & Year Degree to be awarded: **December** 2003 **May** **August**

****

Approval Signatures of Dissertation Committee

Signature, Dean of Graduate School

Martin Cadwallader/EH

Committee's Page. This page is not to be hand-written except for the signatures

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PEPTIDE AND PROTEIN ASSEMBLY USING THE STAUDINGER LIGATION

Bradley L. Nilsson

Under the supervision of Professor Ronald T. Raines

At the University of Wisconsin–Madison

The technique of native chemical ligation enables the total chemical synthesis of proteins. In native chemical ligation, the thiolate of an *N*-terminal cysteine residue of one peptide attacks the *C*-terminal thioester of a second peptide. An amide linkage forms after rapid S→N acyl transfer. This method is limited, however, by an absolute requirement for a cysteine residue at the ligation juncture. Here, this restriction is overcome with a new chemical ligation method, the Staudinger ligation.

The Staudinger ligation can be used to couple a peptide with a *C*-terminal phosphinothioester to another with an *N*-terminal α-azido group to form a single peptide that contains no residual atoms. A phosphinobenzenethiol is used to link a thioester and azide in moderate yields. The product is an amide with no residual atoms. Diphenylphosphinomethanethiol thioesters are shown to give high isolated yields for this transformation. These findings provide precedent for a powerful and versatile new method for the total synthesis of proteins.

The Staudinger ligation for amide bond formation is orthogonal and complementary to other ligation methods. Herein, is described the first use of the Staudinger ligation to couple peptides on a solid support. The fragment thus produced is used to assemble

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

functional ribonuclease A via native chemical ligation. The synthesis of a protein by this route expands the versatility of chemical approaches to protein production.

A phosphinothiol reagent incorporating carboxyl-functionality in one of the side chains could expand the utility of the Staudinger ligation. A phosphinomethanethiol reagent that incorporates a carboxyl group as one of the phosphine substituents is synthesized. This functionalization imparts water solubility and provides a handle for the immobilization of the phosphinothiol to a solid support.

Existing phosphinothiols do not mediate efficient Staudinger ligations at Xaa–Yaa sites in which neither Xaa or Yaa is glycine. Herein, a phosphinocresol reagent is synthesized and shown to give improved yields for couplings at an Ala–Ala ligation site. Its design is based on the promotion of steric crowding in the iminophosphorane intermediate, as revealed by molecular modeling.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgements

Many individuals have contributed to the production of this thesis. I consider myself extremely fortunate in the associations I have enjoyed during my time here at the University of Wisconsin. On both a professional and personal level I have worked with people who have made my graduate school experience extremely enjoyable. I am profoundly grateful for the impact my associates have had on me.

I wish to especially acknowledge Professor Ron Raines. He has been a tremendous advisor. He gives his students a great deal of freedom to develop new ideas. His insight and creativity have taught me to be more imaginative in my own approach to science. I am honestly more excited about the challenges of scientific research now than I was at the beginning of my graduate work. Ron deserves a great deal of the credit for this. I thank Ron for his scientific and career counsel as well as for his financial support during my graduate studies.

Professor Laura Kiessling also deserves special acknowledgement. The original idea for a traceless form of the Staudinger ligation was hers. At the time I started working on this project we thought it would be a quick paper and a chapter in my thesis. It turns out her idea has formed the entire theme of my graduate work presented herein.

My interactions with all the faculty in the department of chemistry have been very positive. I particularly need to acknowledge Professor Hans Reich for sharing his expertise in dealing with phosphines and thiols with me. I also acknowledge Professor Dan Rich for his helpful comments with our early manuscripts on this project.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The students and postdocs I've worked with in the Raines lab have given me stories to tell for a long time to come. I'm grateful to know them and for the impact they've had on my work, and for making my experience here as fun as it has been. Dr. Rob Hondal was my early collaborator on ligation chemistry in the Raines lab. He deserves thanks for introducing me to this exciting area of research. Matt Soellner has been a great collaborator on team Staudinger. The Raines lab will be a much quieter place once the Soellner-Nilsson combination is disbanded. Dr. Ulrich Arnold and Matt Hinderacker are also deserving of thanks for their contributions to my experience. I also thank Ken, Cara, Josh, Lynn, Jonathan, Annie, Eric, and Frank (the chemists) for tolerating my noise and for making the lab a great place to work. The Raines lab biochemists (too numerous to mention individually) also deserve recognition for their helpful scientific advice and for keeping me entertained during the occasional walkthroughs.

The scientific staff at the University of Wisconsin is also tremendous. I thank Dr. Gary Case for his help and instruction in peptide synthesis. I thank Dr. Milo Westler of the NMR facility for his assistance with carrying out many of our NMR studies.

I thank my parents for their profound impact on my life. Their support in all forms can never be repaid. Their love and encouragement have been a constant to me.

Finally I need to thank Trista for her neverending love, support, encouragement and patience. The importance of her understanding and companionship cannot be overstated. We came to Wisconsin as a couple and we leave Wisconsin with Emma, Liam, and Sadie. I thank them for the perspective and humor they bring into my life. The inherent discouragement of scientific research was so much easier to handle knowing that I had them to go home to.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Contents

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Figures

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Schemes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Tables

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Abbreviations

Ala	alanine
Ac	acetyl
Asn	asparagine
Bn	benzyl
Boc	*tert*-butoxycarbonyl
CBD	chitin binding domain
Cys	cysteine
DABCO	1,4-diazabicyclo[2.2.2]octane
DCC	dicyclohexylcarbodiimide
DCU	dicyclohexylurea
DIEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
DTT	dithiothreitol
ESI	electrospray ionization
EtOAc	ethyl acetate
Fmoc	9-fluorenylmethoxycarbonyl
GFP	green fluorescent protein
Glu	glutamic acid

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Gly	glycine
HATU	(*N*-[dimethylamino)-1*H*-1,2,3-triazolo[4,5-b]pyridin-1-ylmethylene-*N*-methylmethanaminium hexafluorophosphate *N*-oxide
HOBt	*N*-hydroxybenzotriazole
HPLC	high pressure liquid chromatography
Leu	leucine
Lys	lysine
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
NMR	nuclear magnetic resonance
Phe	phenylalanine
Pro	proline
PyBOP	benzotrizole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
rDNA	recombinant DNA
RNase A	ribonuclease A
RNase S	ribonuclease S
*t*Bu	*tert*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
Trt	trityl
UV	ultraviolet
Val	valine

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 1

CHEMICAL SYNTHESIS OF PROTEINS

1.1 Introduction

The human genome contains 30,000 or so genes (Lander et al., 2001; Venter et al., 2001). Scientists from a broad range of disciplines are now working to reveal the structure and function of the proteins encoded by these genes. Their findings could lead to the solution of a multitude of problems in biology and medicine. In addition to structure—function analyses of extant proteins, chemical biologists are working to create new proteins with desirable properties, either by *de novo* design or by altering natural frameworks.

The study of natural proteins and the creation of nonnatural ones requires the ability to access and manipulate proteins. The isolation of proteins from their natural source is often tedious, idiosyncratic, and impractical. In contrast, the production of proteins by using recombinant DNA (rDNA) technology, either in a heterologous host or *in vitro*, can provide access to large quantities of protein and allows for the exchange of one of the 20 natural amino acid residues for another. Yet, aggregation often limits the yield of properly folded proteins produced by recombinant proteins. Moreover, the restrictions of the genetic code severely limit the possible modifications.

The chemical synthesis and semisynthesis of proteins harbors the potential to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

overcome many of the disadvantages of current protein production methods (Casi & Hilvert, 2003; Kent, 2003). Chemical synthesis using established solid-phase techniques could be rapid and easily automated, and facilitate purification. The development of methodology for the chemical synthesis of proteins is entering a phase of rapid growth. The application of existing and emerging chemical protein synthesis technologies is facilitating the study of important questions pertaining to the biological function of proteins (Borgia & Fields, 2000). In addition, the power of chemical synthesis enables the facile incorporation of nonnatural amino acids or modules into proteins. This nonnatural functionality can be tailored to answer specific questions that mere mutation of a given residue to another of the 20 common amino acids cannot begin to address.

The chemical synthesis of proteins has long been identified as an important objective. On December 12, 1902, Emil Fischer delivered his Nobel Prize lecture in Stockholm, Sweden, saying in part (Fischer, 1902):

> Of the chemical aids in the living organism the ferments—mostly referred to nowadays as enzymes—are so pre-eminent that they may justifiably be claimed to be involved in most of the chemical transformations in the living cell. The examination of the synthetic glucosides has shown that the action of the enzymes depends to a large extent on the geometrical structure of the molecule to be attacked, that the two must match like lock and key. Consequently, with their aid, the organism is capable of performing highly specific chemical transformations which can never be accomplished with the customary agents. To equal Nature here, the same means have to be applied, and I therefore foresee the day when physiological chemistry will not only make extensive use of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

natural enzymes as agents, but when it will also prepare synthetic ferments for its purposes.

A century later, Fischer's vision is becoming reality. Enzymes and other proteins are not only accessible targets for synthetic chemistry, but are poised to become dominant targets of the 21st century. Herein, we discuss current efforts toward preparing proteins synthetically, focusing on the development of powerful new methodologies for splicing together peptide fragments in a convergent strategy for the total chemical synthesis of proteins.

1.2 Advances in Peptide Synthesis

A discussion of the chemical synthesis of proteins must include mention of the prodigious advances in peptide synthesis over the last century. Emil Fischer's synthesis of glycyl glycine is the first reported synthesis of a dipeptide and is also the first instance of the use of the term "peptide" to refer to amino acid polymers (Fischer & Fourneau, 1901). An important synthetic advance for efficient peptide synthesis was the use of reversible protection for the amine functionality (Bergmann & Zervas, 1932). With the emergence of protecting group strategies, it became possible to synthetically access small peptide hormones. For example, in 1953 Vincent du Vigneaud and coworkers reported a solution-phase synthesis of the octapeptide hormone oxytocin (Du Vigneaud et al., 1953). Even though fifty years had passed since the synthesis of the first peptide by Fischer, these types of syntheses were still only accomplished with considerable effort (Merrifield, 1996).

The advent of solid phase methods heralded a revolution for peptide synthesis

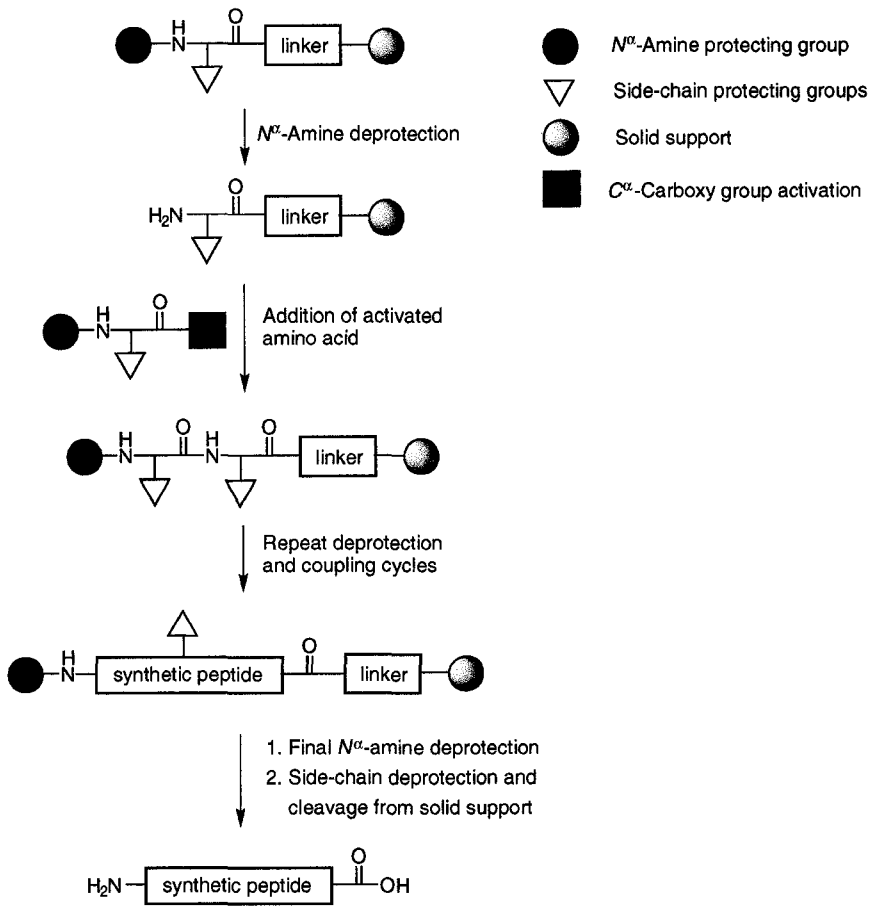
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Merrifield, 1986). In 1963, Bruce Merrifield published the first solid-phase synthesis of a peptide, a tetrapeptide (Merrifield, 1963). He attached an amino acid to an insoluble support via its carboxyl group and attached the next amino acid having a protected amine and an activated carboxyl group. The amine protection was removed and the next residue was coupled in like manner. Within a few years, Merrifield reported the development of an instrument for the automated synthesis of peptides (Merrifield et al., 1966). In short order, he and Bernd Gutte applied this new method to achieve the total chemical synthesis of a protein, ribonuclease A (RNase A) (Gutte & Merrifield, 1969; Gutte & Merrifield, 1971). Concurrently, a group led by Ralph Hirschmann reported the chemical synthesis of RNase S by solution-phase segment condensations (Hirschmann et al., 1969). The Merrifield work was largely performed by a single individual, while the work by the Hirschmann team required eight scientists to complete. This distinction highlighted a major advantage of solid-phase peptide synthesis.

Automated solid-phase peptide synthesis is commonplace today. Proteins and peptides of short to moderate length are routinely synthesized by the strategy shown in Scheme 1.1. An amino acid with both amine and side-chain protection is immobilized to a resin. Several strategies for amine protection are used, namely the use of the acid sensitive *tert*-butoxycarbonyl (Boc) protection or the base sensitive 9-fluorenylmethyloxycarbonyl (Fmoc) protection (Wellings & Atherton, 1997). These N^{α}-amine protecting groups can be removed quickly and a protected amino acid utilizing enthalpic activation of the carboxyl group can then be coupled to the free resin-bound amine. The coupling reactions are forced to completion by the use of an excess of the activated soluble amino acid. The cycle of deprotection and coupling is repeated to complete the sequence, and the peptide

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

can be liberated from the resin and have its side-chain protection removed to yield the desired peptide (Guy & Fields, 1997; Stewart, 1997).



Scheme 1.1 *Solid-phase peptide synthesis.*

The efficiency of solid-phase peptide synthesis is consistently improving. The development of new solid supports has increased the length of accessible peptides (Meldal, 1997). New linkers between the support and the peptide have diversified the types of chemistry that can be used to liberate a synthetic peptide (Songster & Barany, 1997). New side-chain protection strategies have minimized deleterious side reactions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Improved coupling reagents have increased the speed and efficiency of amino acid couplings while reducing the risk of amino acid epimerization (Albericio & Carpino, 1997; Miranda & Alewood, 1999).

Solid-phase peptide synthesis has enabled the total chemical synthesis of proteins. Since the pioneering work of Merrifield, proteins that contain as many as 166 amino acid residues have been synthesized in this fashion (Merrifield, 1996). These syntheses have, in some cases, been critical to structure–function analyses of the target proteins. Notably, the chemical synthesis of HIV-1 protease enabled the determination of important crystal structures of the protease in complex with inhibitors (Miller et al., 1989; Wlodawer et al., 1989; Swain et al., 1990). In addition, an enantiomeric, all-D amino acid form of this protease was synthesized as a demonstration of its chiral specificity for an all-D substrate (deL. Milton et al., 1992). Another notable example is a recent synthesis of a modified B1 domain of streptococcal G protein (Odaert et al., 1999). This synthesis incorporated a completely nonnatural dibenzofuran-based β-turn mimic as a conformational probe. These studies highlight the true power and potential of the total chemical synthesis of proteins in enhancing our understanding of protein structure and function.

Despite the significant gains made in advancing the technology of solid-phase peptide synthesis, there remain limitations. Modern peptide synthesis is typically limited to peptides of $\leq$40 residues (Bray, 2003). Peptides and proteins of greater length can be prepared, although not routinely. Hence, most proteins cannot by synthesized by the stepwise assembly of amino acid monomers.

The convergent assembly of protected or partially protected peptide segments, both in solution and on solid-phase, is one approach used to access proteins greater than 40

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

amino acids in length (Lloyd-Williams et al., 1992; Albericio et al., 1997; Aimoto, 1999; Sato & Aimoto, 2003). In this approach, peptides are synthesized in a protected form and are then assembled. The 238 amino acid residue precursor molecule of the green fluorescent protein (GFP) was synthesized using this approach (Nishiuchi et al., 1998). A total of 26 peptide fragments corresponding to parts of the GFP were synthesized and then assembled in solution followed by deprotection. The resulting protein exhibited a fluorescence profile identical to the native protein upon standing in solution. In addition to demonstrating the ability to synthesize this protein, this study also demonstrated that the formation of the GFP chromophore is not dependent on any external cofactors.

1.3 Total Synthesis of Proteins by Chemoselective Ligation of Peptide Fragments

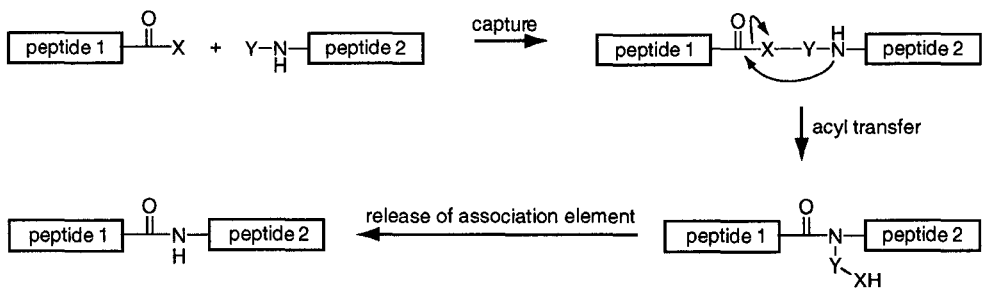
The formation of an amide bond between protected peptide fragments by nucleophilic attack of a free amine on an enthalpically activated carboxyl group can be problematic when dealing with large peptides (Coltart, 2000). With large peptides coupling efficiency is reduced due to an entropic barrier. Namely, the reactivity between the free amine and the activated carboxyl group is not inherently high enough for acyl transfer to occur rapidly enough at the relatively dilute concentrations necessary to solubilize large peptides.

In nature, amide bond formation occurs by transfer of a *C*-terminal acyl group of a nascent peptide chain to the N^{α}-amino group of an aminoacyl *t*RNA in the ribosome. The *C*-terminal acyl group is activated only as en ester by a *t*RNA and the transfer is aided, in part, by the two coupling partners being held in close proximity to one another in the proper conformation (Nissen et al., 2000; Yin & Steitz, 2002). New peptide bond forming

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

strategies have emerged that take advantage of these principles for the intermolecular coupling of large peptide fragments in solution.

These emerging methods are now referred to as *peptide ligation* methods (Scheme 1.2). These ligation-based techniques have become practical and powerful methods for the assembly of synthetic peptides, often as unprotected peptide fragments in aqueous solution. These strategies utilize an initial capture step followed by an intramolecular acyl transfer. The capture step involves the preassociation of the coupling partners by a chemical reaction that is more efficient and rapid than intermolecular acyl transfer to an amine. The acyl transfer step is then rendered intramolecular and can thus occur far more readily (Page & Jencks, 1971; Menger, 1985). The final step of these strategies is the release of the association element following acyl transfer. In some cases, this occurs spontaneously, while in other instances it requires a separate chemical step.



Scheme 1.2 *A general strategy for peptide ligation.*

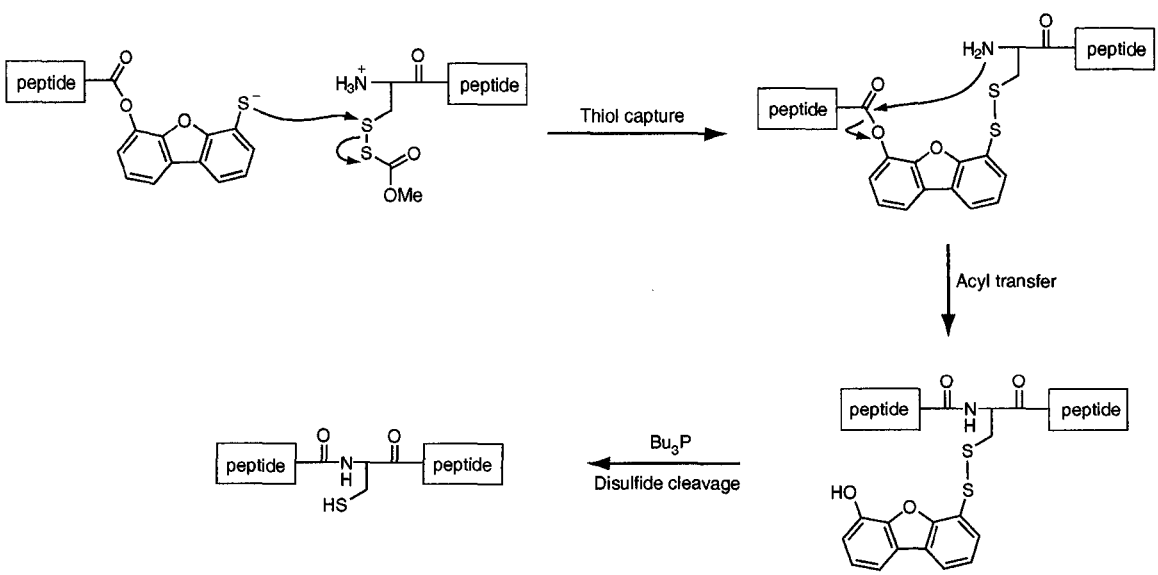
Herein, we discuss the current peptide ligation strategies along with their relative advantages and disadvantages. Some important examples of the application of peptide ligation strategies will also be highlighted. Our focus is on ligation strategies that yield

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

native amide bonds. There are existing ligation approaches that yield non-native bonds between peptide fragments (Tam et al., 1999; Tam et al., 2001; Eom et al., 2003). Although these approaches contribute to the synthesis of non-native proteins, they will not be discussed herein.

1.4 Prior Thiol Capture

The modern application of peptide ligation approaches for assembling synthetic peptide fragments traces its beginnings to Kemp's *prior thiol capture* strategy (Scheme 1.3). In this approach, a 4,6-disubstituted benzofuran (4-hydroxy-6-mercaptodibenzofuran) is used as the critical association element to bring the peptide coupling partners together (Kemp & Galakatos, 1986; Kemp et al., 1986a). This strategy uses highly efficient thiol–disulfide exchange in a prior capture step preparatory to the acyl transfer step.



Scheme 1.3 *Prior thiol capture approach to peptide assembly.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the *prior thiol capture* approach the disubstituted furan is derivatized as a *C*-terminal peptide 4-(acyloxy)-6-mercaptodibenzofuran. A second peptide possessing an *N*-terminal cysteine derivatized as the mixed methoxycarbonylsulfenyl (Scm) disulfide (Kemp & Carey, 1989) can undergo disulfide exchange with these benzofuran peptides in an initial thiol capture step that covalently brings the ligation partners into proximity (Kemp & Fotouhi, 1987; Fotouhi et al., 1989). An intramolecular *O*→*N*-acyl transfer then occurs to produce the desired amide bond. The *N*-terminal cysteine of peptide 2 exists as a disulfide with the dibenzofuran association element. This acyl transfer step occurs with a halflife ranging from 0.1 to 50 h depending on the bulk of the side chain of the *C*-terminal acyloxy-derived amino acid (Kemp et al., 1986b). Upon completion of the acyl transfer step, phosphine reduction of the mixed disulfide yields the native peptide. The amide-bond forming acyl transfer occurs without racemization of the coupled amino acids (McBride & Kemp, 1987).

The utility of this approach was demonstrated by the synthesis of several peptide fragments. In one study, the *C*-terminal 29-residue fragment of BPTI was synthesized from four segments (amino acids 30–37, 38–50, 51–54, 55–56) (Fotouhi et al., 1989). Each segment possessed an *N*-terminal cysteine and was derivatized as a *C*-terminal dibenzofuran. The synthesis commenced from the two fragments at the extreme *C*-terminus and proceeded towards the *N*-terminus sequentially, one fragment at a time. The *N*-terminal cysteines were protected orthogonally until they were ready for coupling, at which time they were deprotected and derivatized as the mixed Scm-disulfides. In a separate study, a 39-residue peptide and a 25-residue peptide were synthesized in high

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

yield utilizing side-chain protection only at cysteine residues (Kemp & Carey, 1993). In this study it was shown that the ligation of fully protected peptides was 50-fold slower than ligation of peptides protected only at cysteine.

Kemp's *prior thiol capture* strategy is an important contribution to the development of peptide splicing methodology. Although it has not seen widespread use for the synthesis of proteins via segment condensation, it represents the first significant effort that demonstrates the chemoselective ligation of unprotected peptide fragments. In addition, it represents the first effort toward the systematic application of the use of proximity effects to reduce acyl transfer between peptides to an efficient intramolecular reaction.

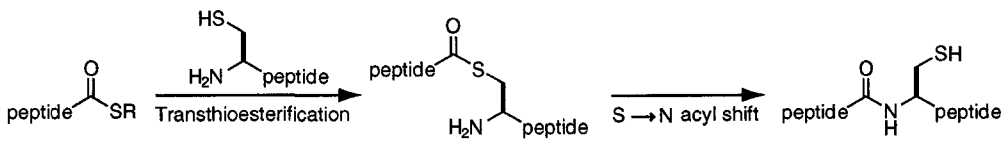
1.5 Native Chemical Ligation

1.5.1 Synthesis of Proteins by Native Chemical Ligation

Currently, the most common ligation method is *native chemical ligation* (Scheme 1.4). This ligation method was initially reported by Wieland in 1953, when the reaction of ValSPh and CysOH in aqueous buffer was shown to yield the dipeptide: ValCysOH (Wieland et al., 1953). This reaction was shown to proceed through the intermediacy of a thioester containing the sulfur of the cysteine residue. In the 1990's, this seminal discovery was developed into a practical method to ligate large peptide fragments (Dawson et al., 1994; Dawson & Kent, 2000). In native chemical ligation, the thiolate of an *N*-terminal cysteine residue of one peptide attacks the *C*-terminal thioester of a second peptide to effect a transthioesterification. An amide linkage forms after rapid $S \rightarrow N$ acyl transfer. *Expressed protein ligation* is an extension of native chemical ligation in which the *C*-terminal thioester is produced by recombinant DNA (rDNA) technology rather than

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

chemical synthesis (Muir et al., 1998; Muir, 2003).



Scheme 1.4 *Native chemical ligation.*

Native chemical ligation is a powerful approach for the total chemical synthesis of proteins. Ligations typically occur entirely in aqueous solution with unprotected peptides (Muir et al., 1997). The rates of these ligations can be modulated through the addition of thiol additives that form more reactive thioester intermediates that can undergo transthioesterification more rapidly (Dawson et al., 1997). Ligation can also occur between cysteine and any of the 20 amino acids, although ligation to hindered amino acids such as valine, isoleucine, and proline occur more slowly (Hackeng et al., 1999). While native chemical ligation is normally carried out in aqueous solution, this method has also been applied to the ligation of unprotected peptide fragments on a solid support (Camarero et al., 1998; Canne et al., 1999; Brik et al., 2000).

Native chemical ligation has seen wide application to the assembly of an ever increasing number of totally synthetic proteins. Some recent examples of the preparation of native proteins include human matrix Gla protein (84 residues) (Hackeng et al., 2001), the anticoagulant microprotein S (116 residues) (Hackeng et al., 2000), and human neutrophil pro α-defensin-1 (75 residues) (Wu et al., 2003b). Proteins with post-translational modifications such as glycosylation (MacMillan & Bertozzi, 2000) that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cannot be accessed in their native form using rDNA techniques have been successfully synthesized using native chemical ligation. These include the glycoproteins diptericin (58 residues) (Shin et al., 1999) and lymphotactin (93 residues) (Marcaurelle et al., 2001). The prion protein has also been synthesized with appropriate mimics of the glycosylphosphatidyl inositol anchor (Ball et al., 2001).

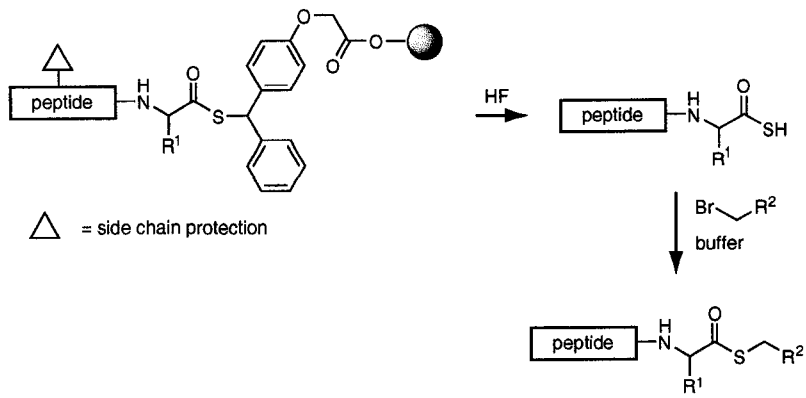
The most powerful aspect of the total chemical synthesis of proteins is the ability to incorporate non-natural residues as probes of mechanism and function, or for therapeutic purposes. Native chemical ligation has yielded many important examples of modified synthetic proteins. The protooncogene H-Ras (166 residues) and its binding partner, the Ras-binding domain (RBD, 81 residues), have been synthesized with the specific incorporation of fluorescent tags to better understand the interactions between these two proteins (Becker et al., 2001; Becker et al., 2003). In a particularly noteworthy example, a synthetic polyethylene glycol-modified erythropoiesis protein (166 residues) has been designed and synthesized (Kochendoerfer et al., 2003). This protein displays monodisperse polymers at specific points on its surface, which serve to prolong its potent biological activity. This work reveals the significant potential therapeutic applications of the total chemical synthesis of proteins.

1.5.2 Solid-Phase Peptide Synthesis of C-Terminal Thioesters

The synthesis of *C*-terminal peptide thioesters is an important consideration in the application of native chemical ligation to the total synthesis of proteins (Alsina & Albericio, 2003). Initial reports of native chemical ligation used Boc-solid phase peptide synthesis on specialized resins that yielded thioacids upon cleavage (Dawson et al., 1994;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Canne et al., 1995; Dawson et al., 1997). These thioacid peptides were then treated with an electrophile to give thioesters (Scheme 1.5).

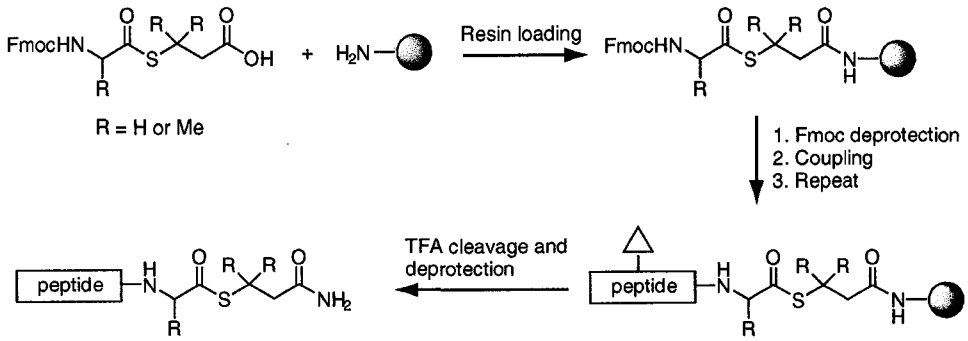


Scheme 1.5 *Thioester synthesis using Boc-based solid phase methods.*

There has been much recent effort focused on the development of Fmoc-based strategies for the synthesis of thioesters. The use of Fmoc strategies avoids the use of harsh HF cleavage conditions and is thus more easily accessible in a laboratory context. One strategy has been to load a presynthesized Fmoc-amino acid thioester onto a solid support and proceed with deprotection and coupling cycles (Scheme 1.6). A caveat of this technique is that nontraditional Fmoc deprotection cocktails must be used, since thioesters are particularly prone to aminolysis by piperidine (Li et al., 1998). A number of deprotection cocktails have been reported, but the best is a mixture of DBU and HOBt in DMF according to thioester stability studies (Bu et al., 2002b). Upon completion of the solid-phase synthesis, TFA cleavage and peptide deprotection can be affected normally, as thioesters are stable to these conditions. These conditions have been used successfully to assemble thioester peptides for subsequent enzymatic cyclization (Bu et al., 2002a; Wu

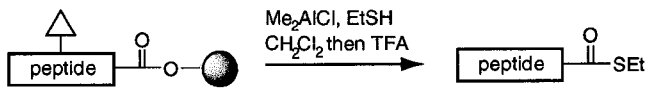
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

et al., 2003a) as well as for the preparation of thioesters of side chain *N*-glycosylated peptides (Hojo et al., 2003).



Scheme 1.6 *Fmoc-based strategy of peptide thioester synthesis.*

A second Fmoc strategy makes use of novel cleavage conditions from the solid support (Swinnen & Hilvert, 2000). An advantage of this method is that standard resins and peptide synthesis conditions are employed. Treatment the final peptide on solid support with an alkylaluminum reagent and an alkyl thiol in methylene chloride cleaves the peptide as a thioester. A separate TFA treatment is required for deprotection of the side chain protecting groups.

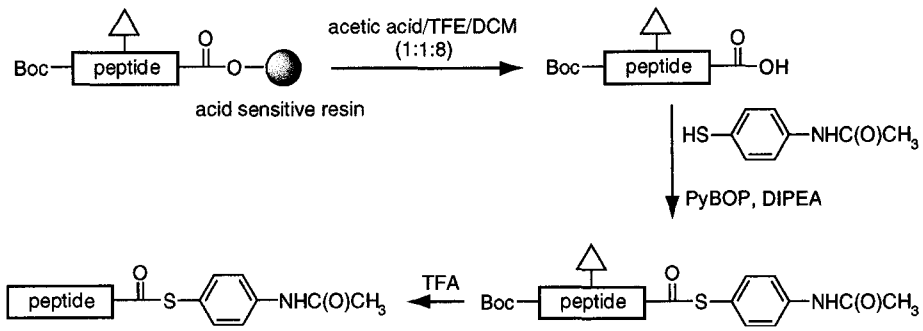


Scheme 1.7 *Fmoc thioester cleavage conditions.*

Solution-phase synthesis of thioesters from protected peptides has also been reported (Scheme 1.8) (von Egglekraut-Gottanka et al., 2003). Solid-phase peptide synthesis is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

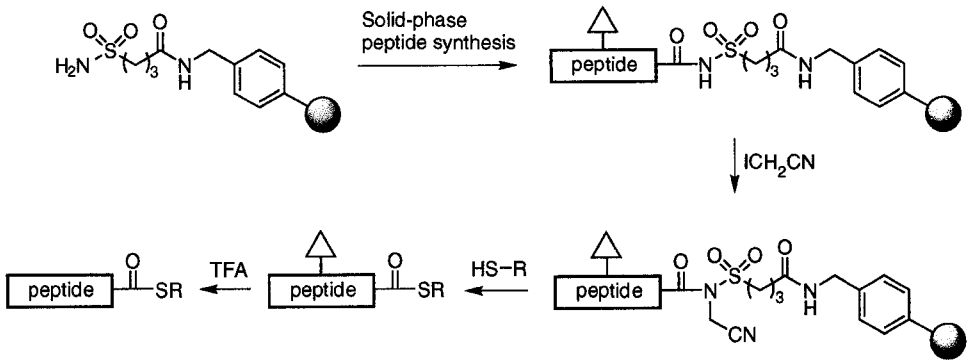
carried out using a chlorotrityl resin that is sensitive to acid in very low concentrations. The peptide can be cleaved from the resin in its protected form, after which thioester formation occurs using standard coupling procedures. The thioester thus formed can be deprotected efficiently using TFA conditions and purified prior to use.



Scheme 1.8 *Solution-phase peptide thioester synthesis.*

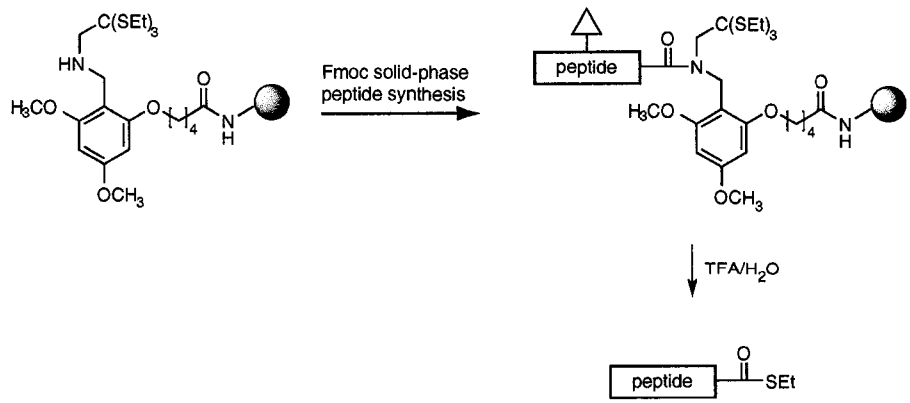
Specialized resins show particular promise as general methods for the synthesis of peptide thioesters. Several groups have used Kenner's sulfonamide safety-catch linker for the synthesis of these types of thioesters (Scheme 1.9) (Backes & Ellman, 1999; Ingenito et al., 1999). This strategy involves loading of the nascent peptide on a sulfonamide linker that is stable to acid and base. Solid-phase peptide synthesis is carried out using standard protocols. The linker is then activated by alkylating the sulfonamide nitrogen. Upon activation, the carbonyl of the sulfonamide becomes extremely electrophilic. The peptide can be liberated from the resin as a thioester by introducing a thiol, which serves as a nucleophile against the sulfonamide electrophile. The liberated peptide is then deprotected using standard TFA conditions. The use of chaotropic salts has been shown to improve thioester yields (Quaderer & Hilvert, 2001).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 1.9 *Synthesis of peptide thioesters using sulfonamide safety-catch resins.*

A second type of specialized linker utilizes trithioortho esters as masked thioesters (Scheme 1.10) (Brask et al., 2003). This strategy is useful only for the synthesis of peptide thioesters in which glycine is the *C*-terminal amino acid. Peptide synthesis occurs using standard Fmoc conditions on a novel backbone amide linker in which the thioester is masked as a trithioorthoester. Upon the completion of solid-phase peptide synthesis, the resin is treated with TFA, thus deprotecting the peptide and unmasking the thioester.

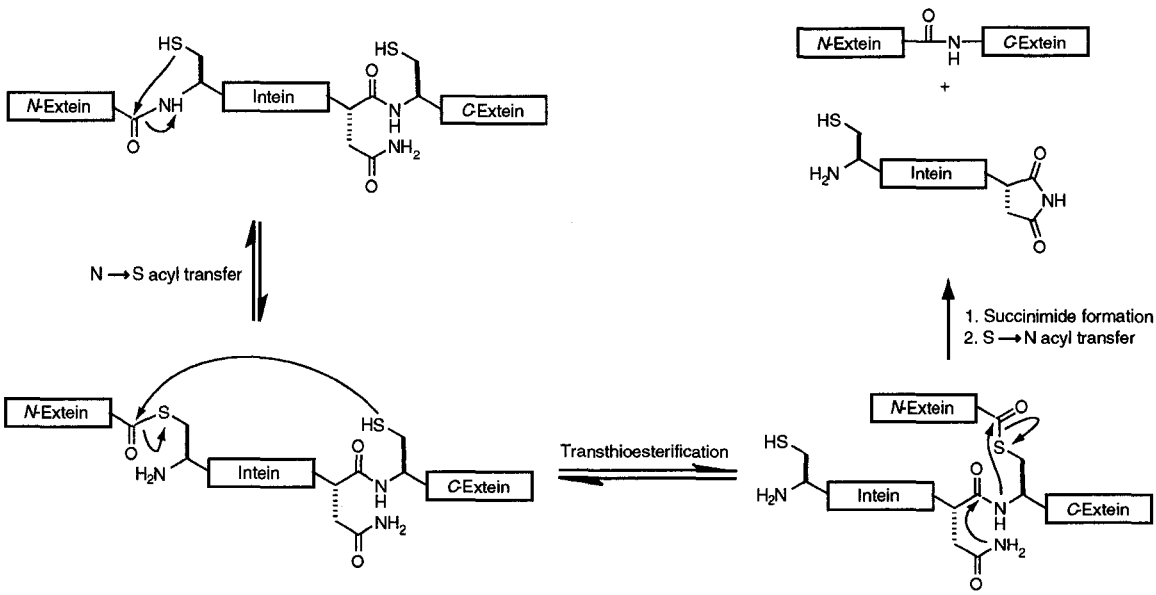


Scheme 1.10 *Synthesis of peptide thioester using masked trithioorthoesters.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.5.3 Expressed Protein Ligation

Expressed protein ligation is a powerful protein semisynthesis technique that combines the tools of chemistry and biology (Blaschke et al., 2000; Goody et al., 2002; Hofmann & Muir, 2002; Muir, 2003). In essence, expressed protein ligation uses the tools of recombinant DNA (rDNA) technology to access the thioester fragment for native chemical ligation. The inspiration for this method is intein-mediated, post-transcriptional protein splicing (Wallace, 1993). In Nature, an intein is a protein splicing element analogous to the intron in nucleic acids. It mediates its own excision from a peptide sequence through a series of acyl transfer reactions that ultimately result in the splicing of the flanking peptides (Scheme 1.11).

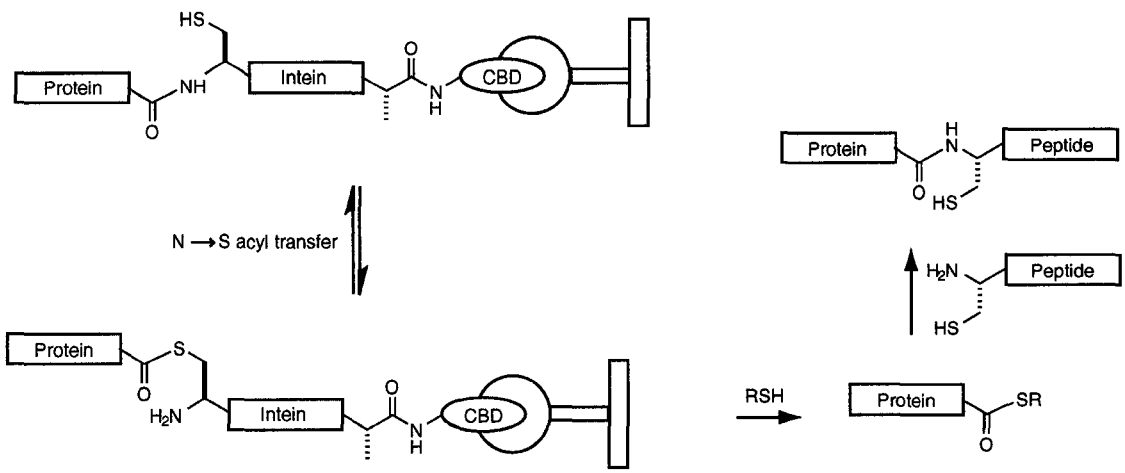


Scheme 1.11 *Mechanism of protein splicing.*

Expressed protein ligation utilizes a modified intein as part of a fusion protein in order

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to access a target protein fragment as a *C*-terminal thioester. The fusion protein consists of the protein target fragment, a modified intein (the *C*-terminal asparagines is mutated to alanine), and a chitin-binding domain (CBD) (Scheme 1.12). The intein is modified by mutating the *C*-terminal asparagine to an alanine, thus making the intein unable to perform the final splicing step of succinimide formation followed by S→N acyl transfer. The fusion protein can be applied to a chitin column where the resultant thioester can be trapped using a solution of an exogenous thiol, thus eluting the target protein as a *C*-terminal thioester and leaving the remainder of the fusion protein bound to the chitin column.



Scheme 1.12 *Expressed protein ligation.*

The semisynthesis of proteins using expressed protein ligation represents a powerful union of chemistry and biology. The use of rDNA to produce protein fragments increases the size of proteins that are accessible to native chemical ligation. The bulk of a protein can be accessed by biosynthesis, while a small peptide fragment can be prepared by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

chemical synthesis and ligated to complete the protein. This technique is gaining widespread use as a potent tool for protein engineering. The power of this technique is illustrated in two recent examples of protein prosthesis. In these examples, RNase A was prepared via expressed protein ligation incorporating two nonnatural turn mimics, dimethylproline (Arnold et al., 2003) and dinipecotic acid (Arnold et al., 2002), at a critical β-turn (Figure 1.1). The incorporation of these turn mimics resulted in the stabilization of these engineered RNase A variants to thermal denaturation compared to the wild-type enzyme.

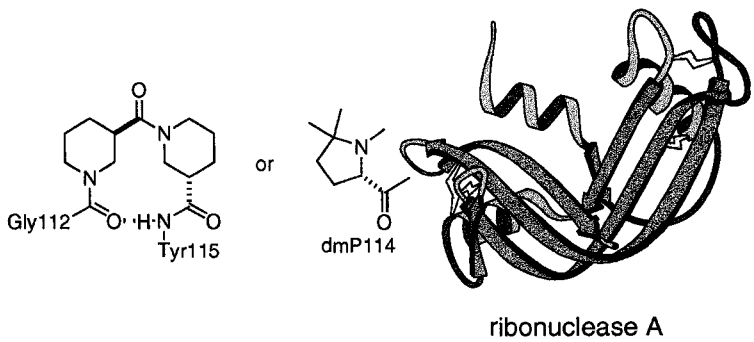


Figure 1.1 *Protein prosthetics for engineered RNase A.*

Intein-mediated protein splicing has recently been reported to occur within living cells. Conditional protein trans-splicing is an elegant example of this work (Mootz & Muir, 2002; Mootz et al., 2003). Trans-splicing strategies involve splitting an intein in two. When these fragments are noncovalently reconstituted, the activity of the intein is regained. Two fusion proteins are produced, each containing one half of a desired protein, one half of a functional intein, and a dimerization domain. The dimerization event can be initiated by the addition of a small molecule to the cells. Upon dimerization, association

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the split intein occurs, followed by splicing of the protein fragments. This approach has also been used to tag synthetic molecular probes to proteins within cells (Giriat & Muir, 2003).

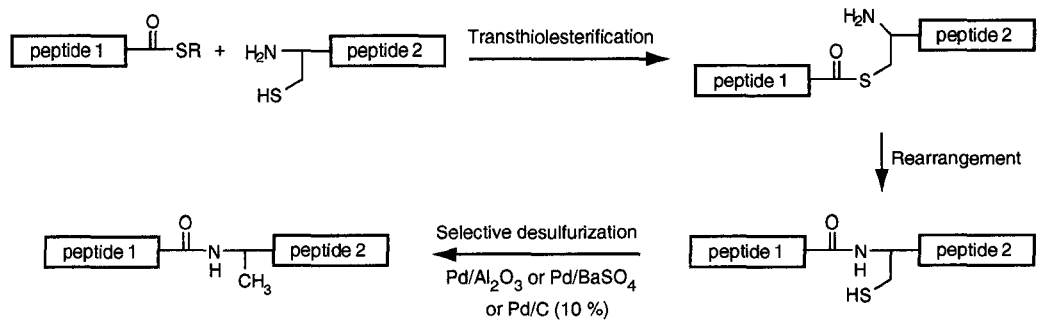
1.5.4 Native Chemical Ligation and Expressed Protein Ligation Followed by Chemical Modification

The utility of the native chemical ligation approach has been expanded to include chemical modification. In these strategies an additional chemical modification step is utilized to allow ligation using cysteine or cysteine derivatives, yet giving a final protein product that does not contain cysteine at the ligation juncture. To date, these approaches are quite limited and have been used only in specific instances.

One such example is native chemical ligation at cysteine followed by desulfurization (Yan & Dawson, 2001). This strategy employs standard native chemical ligation using a cysteine residue. Following the ligation step, a chemical desulfurization of the cysteine residue gives an alanine at the ligation site (Scheme 1.13). The desulfurization has been effected using a variety of metal reagents, including Pd/Al_2O_3, Pd/carbon (10%) and $Pd/BaSO_4$ in the presence of hydrogen gas. Three different proteins were synthesized using this technique: streptococcal protein G B1 domain (56 residues), a variant of the ribonuclease, barnase (110 residues), and the cyclic antibiotic microcin J25 (21 residues). This approach has seen limited use due to the fact that all cysteines in a protein target will be reduced by the desulfurization conditions. Disulfides, however, are stable to those conditions. One could envision a system in which a protein could be oxidatively refolded leaving only one free cysteine residue, which could then be reduced selectively. No

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

examples of such a synthesis have been reported.



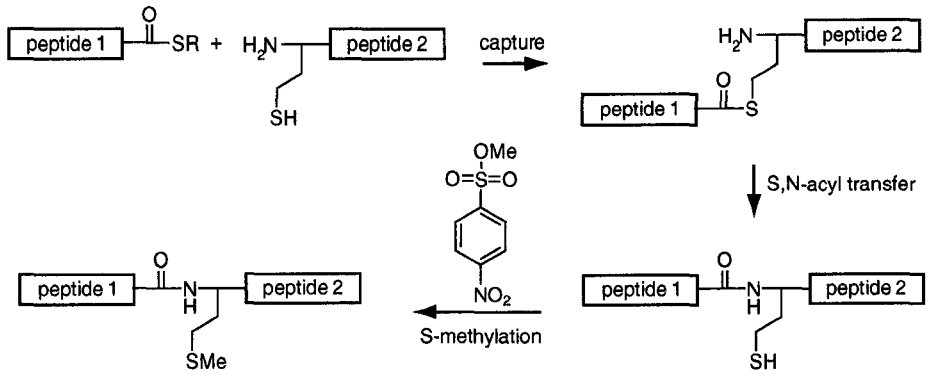
Scheme 1.13 *Native chemical ligation followed by reductive desulfurization.*

A second example yields methionine at the splice site by ligation with homocysteine followed by chemical methylation (Tam & Yu, 1998). This ligation proceeds initially in a nearly identical fashion to native chemical ligation. Homocysteine effects *S*→*N* acyl transfer efficiently despite having an additional carbon in its side chain. Once acyl transfer is complete, sulfur methylation is carried out using methyl *p*-nitrobenzenesulfonate. An excess of the methylating reagent is used, but by limiting reaction times the methylation of lysine amines is avoided. This method has been used in the synthesis of parathyroid hormones, but its generality is limited, as only targets with one sulfur-containing residue can be synthesized.

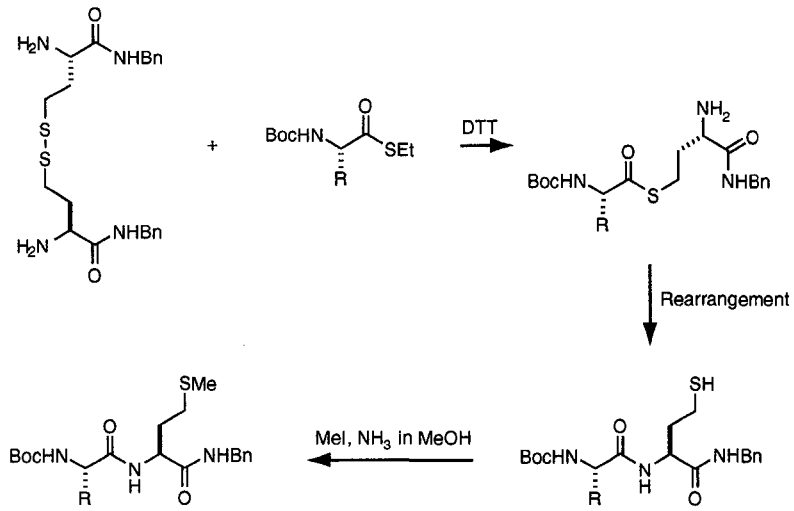
Alternative conditions for the application of this system have been explored in model systems (Scheme 1.15) (Pachamuthu & Schmidt, 2003). These conditions start from the disulfide homocystine that is reduced *in situ* with dithiothreitol (DTT) in the presence of a thioester amino acid. Native chemical ligation occurs efficiently as expected. Modified conditions are used to effect *S*-methylation. The homocysteine amide is treated with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

methyl iodide and ammonia in methanol for several hours at 0 °C, followed by 30 min at room temperature. Although good yields are reported, these conditions could be rather impractical and unselective for the treatment of large unprotected proteins.



Scheme 1.14 *A methionine ligation strategy.*



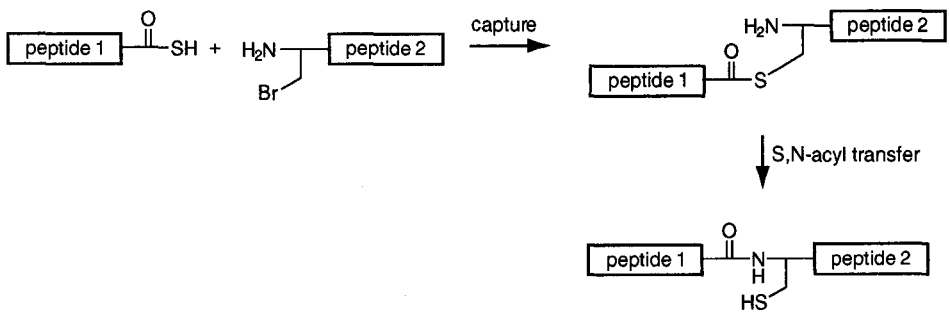
Scheme 1.15 *Native chemical ligation resulting in methionine at the ligation site.*

1.5.5 Native Chemical Ligation Through Acyl-Initiated Capture

Another original approach to native chemical ligation is the *in situ* generation of a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cysteine thioester using an acyl-intiated capture strategy (Tam et al., 1995). In this method (Scheme 1.16) a peptide containing a *C*-terminal thioacid is coupled to a second peptide possessing β-bromoalanine as its *N*-terminal amino acid. An intermediate thioester is generated by attack of the thioacid on the bromoalanine to give the intermediate thioester that then undergoes $S \rightarrow N$ acyl transfer resulting in a native amide bond with cysteine at the ligation junction. An advantage of this method is that the ligation can proceed at more acidic pH levels (pH 5) than native chemical ligation (pH 7.6). A disadvantage of this method is that bromoalanine is prone to undesired side reactions and can thus be impractical to work with.



Scheme 1.16 *Native chemical ligation by acyl-initiated capture.*

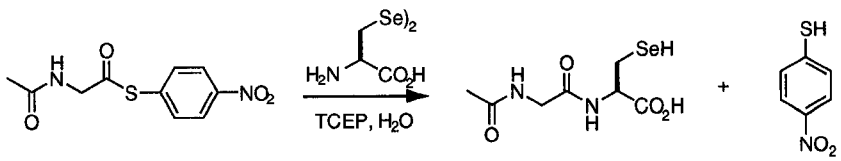
1.5.6 Native Chemical Ligation with Selenocysteine

Native chemical ligation and expressed protein ligation have recently been expanded to include ligation at selenocysteine (Gieselman et al., 2002). Selenocysteine (Sec, U) is incorporated translationally and has its own $tRNA^{Sec}$ and codon, UGA, which is also the *opal* stop codon. Decoding a UGA codon for selenocysteine requires a unique structure in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the 3′ untranslated region of the mRNA called a *s*elenoc*y*steine *i*nsertion *s*equence (SECIS) element. The production of eukaryotic selenocysteine-containing proteins in prokaryotes is problematic because eukaryotic and prokaryotic cells use a different SECIS element (Arner et al., 1999).

The feasibility of using selenocysteine in native chemical ligation has been demonstrated through the use of a model thioester (Scheme 1.17) (Hondal et al., 2001). In addition, pH–rate profiles have demonstrated the increased rate of selenocysteine ligations compared to cysteine ligations at lower pH, as expected from the lower pK_a of selenols and the somewhat higher reactivity of selenolates. For example, the reaction with selenocysteine is 10^3-fold faster than that with cysteine at pH 5.0.



Scheme 1.17 *Native chemical ligation using selenocysteine.*

Selenocysteine has also been used in expressed protein ligation using ribonuclease A (RNase A; EC 3.1.27.5) as a model protein (Hondal et al., 2001; Hondal & Raines, 2002). Recombinant DNA technology was used to prepare a fragment corresponding to residues 1–109 of RNase A with a *C*-terminal thioester. Standard solid-phase methods were used to synthesize a peptide corresponding to residues 110–124, but with selenocysteine as residue 110. The thioester fragment and the peptide fragment were then ligated and the product was refolded and purified. The desired C110U RNase A protein was verified by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mass spectrometry and found to have wild-type enzymatic activity. These data indicate that C110U RNase A is not only an intact protein, but a correctly folded enzyme with an Se–S bond between selenocysteine 110 and cysteine 56. The use of selenocysteine in the native chemical ligation of peptides has also been reported (Gieselman et al., 2001; Quaderer et al., 2001).

A synthesis of a selenocysteine variant of the 128 amino acid protein azurin has also been carried out using expressed protein ligation (Berry et al., 2002). Cysteine 112 was replaced with selenocysteine and fragment 112–128 was ligated to fragment 1–111. The variant azurin protein displayed interesting spectral properties and was used to assess the nature of the S–Cu interaction in the native protein.

Isomorphous replacement of sulfur with selenium can be effected with expressed protein ligation. This could be advantageous in terms of stabilizing a protein. For example, proteins containing selenosulfide (Se–S) bonds should have greater conformational stability in a reducing environment than proteins with disulfide (S–S) bonds as the reduction potential of a selenosulfide bond is less than that of the corresponding disulfide bond (Besse et al., 1997). This use of selenocysteine to stabilize an enzyme represents another form of protein prosthesis.

Selenomethionine incorporation has also been reported using a strategy similar to that used for the incorporation of methionine (Roelfes & Hilvert, 2003). Native chemical ligation of peptide fragments occurred using selenohomocysteine to yield the diselenide product. After reduction, the diselenide was methylated to give the selenomethionine analog. This method could be applied generally as the methylation of selenium should be chemoselective at acidic pH.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.6 Chemical Ligation at Non-Cysteine Residues

1.6.1 Limitation of Native Chemical Ligation

A limitation of native chemical ligation is its intrinsic reliance on having a cysteine residue at the ligation juncture. Cysteine is uncommon, comprising only 1.7% of all residues in proteins (McCaldon & Argos, 1988a). Modern peptide synthesis is typically limited to peptides of ≤40 residues (Bray, 2003), whereas most proteins contain hundreds of residues. Hence, most proteins cannot be prepared by any method that allows for peptides to be coupled only at cysteine residues. The removal of the cysteine limitation by applying a more general ligation technology would greatly expand the utility of total protein synthesis (Cotton & Muir, 1999). An ideal technique should efficiently and rapidly ligate between any two amino acid residues without detectable levels of racemization.

1.6.2 Conformationally Assisted Ligation

It is possible to ligate a *C*-terminal thioester fragment with a second peptide without cysteine at the *N*-terminus when there is a strong conformational preference for the association of the two fragments (Beligere & Dawson, 1999). Chymotrypsin inhibitor 2 (CI2) can be subdivided into 2 fragments corresponding to residues 1–40 and 41–64. These fragments self-associate to form a stable protein. When fragment 1–40 was synthesized as a *C*-terminal thioester, complete ligation to fragment 41–64 (residue 41 is methionine) was complete in under 2 h. This synthesis of CI2 has recently been applied to the preparation of a fluorophore-linked analog of CI2 for studies of protein folding

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Deniz et al., 2000).

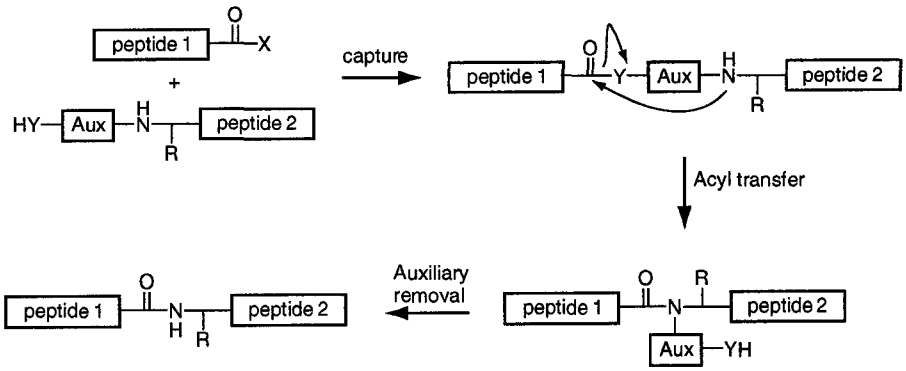
RNase S and the S-peptide is another set of protein fragments that self-associate strongly in solution, and has also been examined for conformationally assisted ligation (Beligere & Dawson, 1999). RNase A 1–20 was synthesized as a *C*-terminal thioester and incubated with RNase A 21–124 (residue 21 is serine). The ligation was complete in 10 h.

One limitation of conformationally assisted ligation is that Xaa–Yaa coupling is slow when Xaa is a β-branched residue (Beligere & Dawson, 1999). Another limitation of the technique is that it is necessarily limited to proteins·peptide fragments that have an inherent propensity for self association.

1.6.3 The Use of Removable Auxiliaries for Peptide Ligation

An emerging strategy has been the use of auxiliaries that act as cysteine surrogates to mediate the chemical ligation of peptide fragments (Scheme 1.18). These auxiliaries are synthetically appended to the *N*-terminus of a peptide. A second peptide is activated at the *C*-terminus in a manner that is specifically complementary to capture by the auxiliary. Upon capture, an acyl transfer event occurs. The final step is removal of the auxiliary to yield the native amide bond. Often, removal of the auxiliary involves chemical treatment of the assembled peptides.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

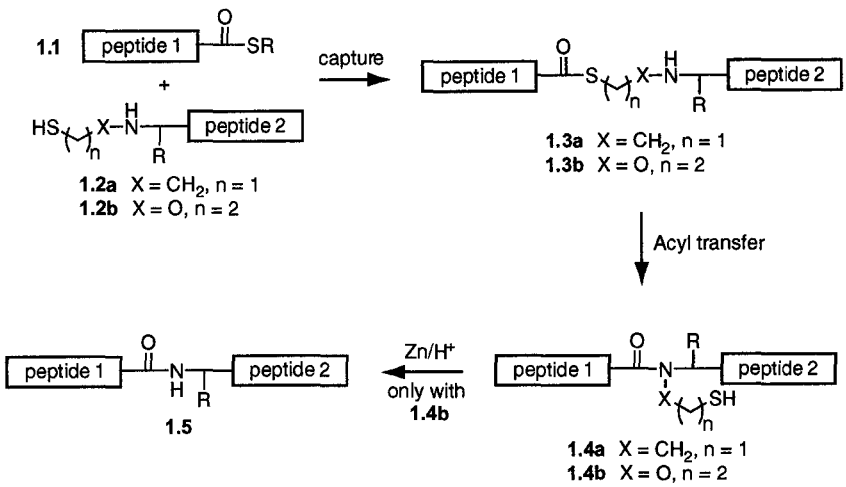


Scheme 1.18 *Auxiliary-mediated peptide ligation.*

The first reported use of auxiliaries for peptide ligation at non-cysteine residues was the use of N^{α}-ethanethiol and N^{α}-oxyethanethiol (Scheme 1.19) (Canne et al., 1996). Peptides containing these auxiliaries at the *N*-terminus (**1.2a** and **1.2b**) were synthesized and coupled to peptides with *C*-terminal thioester under conditions similar to native chemical ligation (6 M guanidine, 0.1 M phosphate buffer, pH 7.5). Reaction temperatures ranged from 25 to 37 °C, and reaction times ranged from 4 to 22 h. The most efficient couplings occurred when both the *N*-terminal and *C*-terminal amino acids were glycine (75-90% rearranged products **1.3a** and **1.3b** respectively). No unrearranged products were observed. Rearrangement yields were between 52% and 64% when one of the amino acids at the ligation juncture was phenylalanine and the other was glycine. No rearranged products could be detected between two nonglycyl residues. Thus, even though coupling could be carried out at a noncysteine residue, some sequence limitations were observed. After rearrangement, the auxiliary was removed by zinc reduction under acidic conditions to yield a native amide bond. Only the N^{α}-oxyethanethiol auxiliary could be removed efficiently from the rearranged product. The N^{α}-oxyethanethiol

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

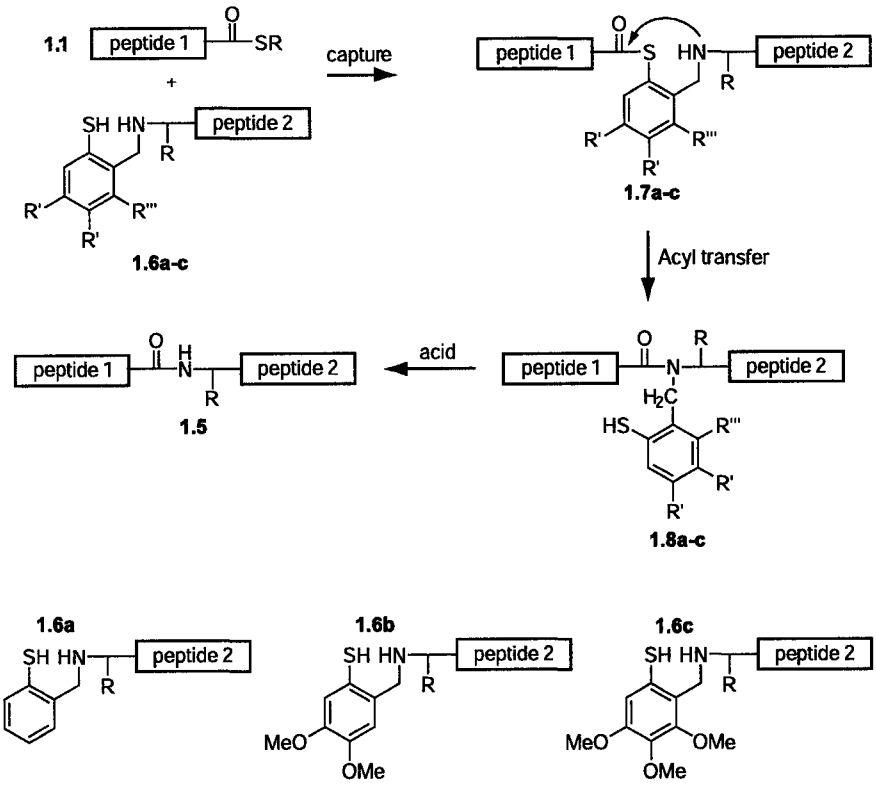
auxiliary has also been used in the cyclization of synthetic peptides (Shao et al., 1998).



Scheme 1.19 *Native chemical ligation using ethanethiol and oxyethanethiol auxiliaries.*

N^{α}-2-Mercaptobenzylamine-derived auxiliaries represent a second class of auxiliary-assisted chemical ligation (Scheme 1.20). Three related forms of this class of auxiliary have been used in peptide ligations (compounds **1.6a-c**). Peptides incorporating these auxiliaries at the *N*-terminus react with *C*-terminal peptide thioesters in a manner similar to cysteine-containing peptides. Acyl transfer occurs through a six-membered transition state. The auxiliaries, with appropriate phenyl ring substitution, can be removed from the acyl transfer products under acidic conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 1.20 *N^{α}-2-Mercaptobenzylamine-assisted chemical ligations.*

Three auxiliaries of this class have been reported. Auxiliary **1.6a** was the first auxiliary of this class to be used for chemical ligation (Offer & Dawson, 2000). This compound gives rearranged peptides of type **1.8a** in the case of sequences in which glycine is at either the *C* or *N* position of the ligation junction. An exception is if a β-branched amino acid is at the ligation site of one of the coupling partners. Auxiliary **1.6a** could not be removed by acid treatment of the rearranged peptides. Auxiliary **1.6b** was subsequently reported, although its use was demonstrated only in a single peptide ligation where both the *C* and *N* ligation positions were glycine (Kawakami et al., 2001). An advantage of this auxiliary over **1.6a** is that treatment with a trifluoromethanesulfonic acid/thioanisole/trifluoroacetic acid cocktail removed the auxiliary completely from the

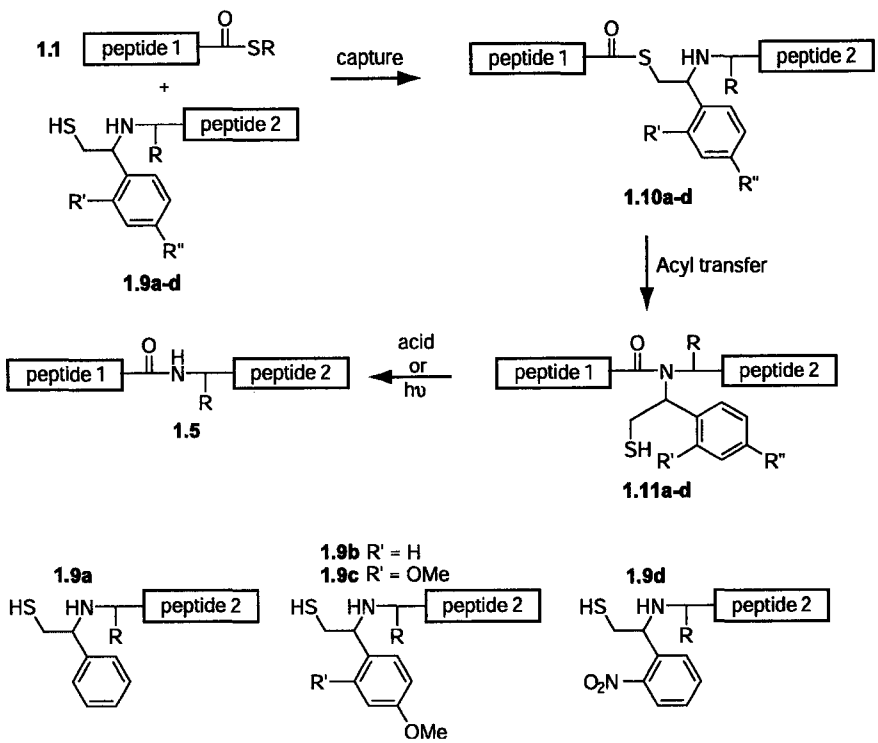
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

rearranged acyl-transfer product to give a native peptide bond. Auxiliary **1.6c** is the most recent ligation agent of this class to be reported (Offer et al., 2002). Several ligation attempts have been reported with various amino acids at the *N* and *C* splice sites. Half-lives for acyl transfer to the rearranged product were 0.2 h at Gly–Gly, 2.0 hours at Lys–Gly, 5.0 hours at Gly–Ala. No rearrangement was observed between Ala–Ala. By increasing the electron density of the phenyl ring with tri-methoxy substitution, the auxiliary can be removed from the rearranged products using either TFA or HF. The 63 residue SH3 domain of α-spectrin was prepared by chemical ligation between Lys27 and Gly28 using this auxiliary. This auxiliary is an efficient agent for coupling at glycyl residues. Although this method is not sequence independent, it does relax significantly the cysteine limitation. Glycine is among the most common amino acids in proteins (McCaldon & Argos, 1988a).

N^{α}-(1-Phenyl-2-mercaptoethyl) auxiliaries represent a second class of auxiliary-assisted chemical ligation (Scheme 1.21). These auxiliaries represent a similar strategy to the N^{α}-2-mercaptobenzylamine agents described previously. They incorporate a thiol in their design in order to effect efficient transthioesterification as the capture step. With the aforementioned N^{α}-2-mercaptobenzylamine auxiliaries, the acyl transfer rearrangement occurred through a six-membered transition state. With N^{α}-(1-phenyl-2-mercaptoethyl) auxiliaries, this rearrangement occurs through a smaller, five-membered transition state. Substitution on the phenyl ring affects the acid lability of the auxiliary in the rearranged product. In addition, the nature of the substituent can allow for removal of the auxiliary under more mild conditions. For example, auxiliary **1.9d** can be removed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

photochemically when *o*-nitro substitution of the phenyl ring is employed.



Scheme 1.21 *N^{α}-(1-Phenyl-2-mercaptoethyl)-assisted chemical ligation.*

Initial studies of the N^{α}-(1-phenyl-2-mercaptoethyl) class of auxiliaries were undertaken using compounds with no substitution on the phenyl ring (compound **1.9a**) (Marinzi et al., 2001). Ligations between several amino acids were explored in order to examine the sequence requirements for rearrangement to occur. It was found that the half-lives for model peptide rearrangements were 30-40 minutes at Gly–Gly, 120-140 minutes for Ala–Gly, with no rearrangement observed at Gly–Ala or Ala–Ala. The transthioesterification capture step occurred efficiently regardless of sequence, thus it was ascertained that the acyl transfer rearrangement step was rate-limiting. No removal of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

N^{α}-(1-phenyl-2-mercaptoethyl) auxiliary from rearranged products has been reported to date.

Two ring-substituted forms of the same auxiliary were reported concurrently (compounds **1.9b** and **1.9c**). The N^{α}-(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary (**1.9b**) and the N^{α}-(1-(2,4-dimethoxyphenyl)-2-mercaptoethyl) auxiliary (**1.9c**) were both shown to ligate efficiently at Xaa–Gly sites where Xaa was Gly, His, Ala, or Lys. Rearrangement yields ranged from 76 to 98% with the typical rearrangement complete in 16 hours. Ligations using **1.9c** at Xaa–Gly sites where Xaa was not Gly required longer reaction times (40 h). The auxiliaries could be removed from rearranged products using HF (**1.9b**) or TFA (**1.9c**). These auxiliaries have been applied to the total chemical synthesis of cytochrome b562 (106 residues) (Low et al., 2001) and for the cyclization of 11 residue cyclic peptides (Cardona et al., 2003).

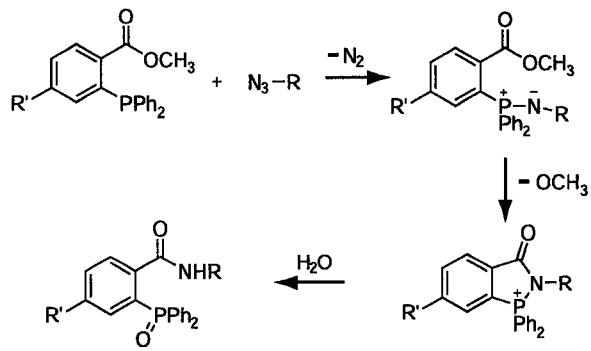
The N^{α}-(1-(2-nitrophenyl)-2-mercaptoethyl) (**1.9d**) auxiliary represents a photolabile derivative of this class of ligation reagents. Capture and rearrangement was demonstrated by the ligation of two 10 residue peptides at a Gly–Gly site. Rearrangement was complete after 24 h. The auxiliary was removed efficiently by photochemical irradiation at 365 nm for 6 h in phosphate buffer, pH 7.4. Although the use of the N^{α}-(1-(2-nitrophenyl)-2-mercaptoethyl) auxiliary is still subject to the glycine limitation, it can be removed under far milder conditions than many of its analogs.

1.6.4 Staudinger Ligation as a Peptide Ligation Strategy

The Staudinger reaction provides another alternative to native chemical ligation with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the promise of removing the cysteine limitation. In the Staudinger reaction, a phosphine is used to reduce an azide to an amine: $PR_3 + N_3R' + H_2O \rightarrow O{=}PR_3 + H_2NR' + N_2$ (g) (Staudinger & Meyer, 1919; Gololobov & Kasukhin, 1992). This reaction occurs via a stable intermediate, an iminophosphorane ($R_3P^+{-}^-NR'$). (This iminophosphorane is sometimes referred to by the less precise term, "aza-ylid".) Vilarrassa and others have shown that the nitrogen of the iminophosphorane can be acylated, both in intermolecular and intramolecular reactions (Garcia et al., 1984; Garcia et al., 1986; Bosch et al., 1995; Mizuno et al., 1999b; Velasco et al., 2000). Hydrolysis of the resulting amidophosphonium salt gives an amide and phosphine oxide. Bertozzi has shown that the phosphine can itself serve as the acyl donor in a strategy used to tag azido sugars (Scheme 1.22) (Saxon & Bertozzi, 2000b; Saxon et al., 2002; Vocadlo et al., 2003). A drawback of the initial Bertozzi method for "Staudinger ligation" is that the phosphine oxide becomes incorporated into the final amide product.

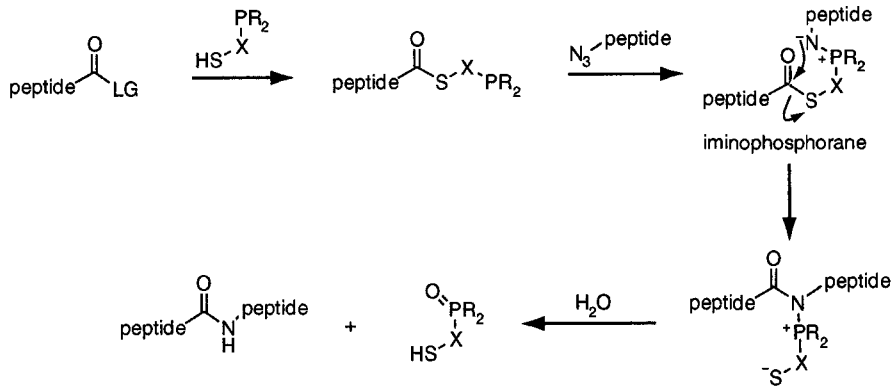


Scheme 1.22 *Staudinger ligation for sugar labeling.*

To apply the Staudinger reaction to peptide synthesis, a phosphinothiol has been used

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to unite a thioester and azide in a traceless fashion. A putative mechanism for this version of the Staudinger ligation (Nilsson et al., 2000; Nilsson et al., 2001) is shown in Scheme 1.23. The ligation begins by transthiolesterification with the phosphinothiol. Coupling of the resulting phosphinothioester with a peptide azide leads to the formation of the reactive iminophosphorane in the capture step. Iminophosphorane nitrogen attack on the thioester constitutes the acyl transfer step and leads to an amidophosphonium salt. Hydrolysis of the amidophosphonium salt produces the desired amide and a phosphine oxide. Significantly, no atoms from the phosphinothiol remain in the amide product.

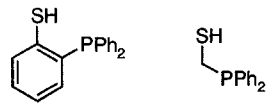


Scheme 1.23 *Traceless Staudinger ligation employing phosphinothiols.*

Several phosphinothiol compounds have been analyzed for their ability to mediate the Staudinger ligation (Figure 1.2). Compound **1.12** was the first phosphinothiol screened (Nilsson et al., 2000). It was coupled to several *N*-acetyl amino acids (glycine and phenylalanine) as a thioester and incubated with N_3GlyNHBn to determine whether dipeptide was formed. The desired dipeptides were formed, albeit in low yields (≤35%). Phosphinothiol **1.12** thus provided proof of principle for the efficacy of a traceless

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Staudinger ligation. The development of **1.12** as a reagent for effecting the Staudinger ligation is discussed in detail in Chapter 2 of this thesis and the synthesis of a water soluble derivative of this molecule is described in Chapter 5.



1.12 **1.13**

Figure 1.2 *Phosphinothiols used in the traceless Staudinger ligation.*

Phosphinothiol **1.13** was a second generation phosphinothiol (Nilsson et al., 2001; Soellner et al., 2003). The development of this compound is discussed in Chapter 3. Staudinger ligation with phosphinothiol **1.12** occurs through an intermediate with a six-membered ring. It was reasoned that reducing the size of this ring would bring the nucleophilic imide nitrogen more proximal to the electrophilic thioester carbon and improve the yields for the Staudinger ligation products. To access a transition state with a smaller ring, the *o*-phenyl group of phosphinothiol **1.12** was replaced with a single methylene group. Thioesters of **1.13** derived from AcOH, AcGlyOH, and AcPheOH were prepared either by transthioesterification or by coupling with dicyclohexylcarbodiimide. These amino acid thioesters incorporating compound **1.13** formed dipeptides with N_3GlyNHBn in >90% yields. Additionally, AcGlySCH_2PPh_2 was coupled with both the D- and L-forms of N_3Ser(Bn)NHBn, N_3PheNHBn, and N_3Asp(OBn)NHBn (Soellner et al., 2002). Each of these dipeptide ligations proceeded in >90% yield. These ligations also confirmed that the Staudinger ligation proceeds with retention of stereochemistry at the α-carbon of the azido acid. As with the auxiliary-assisted ligations discussed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

previously, the Staudinger ligation with compounds **1.12** and **1.13** is less efficient when glycine does not comprise at least one of the coupling partners (Nilsson et al.).

Having demonstrated the efficient use of $HSCH_2PPh_2$ (**1.12**) in mediating the Staudinger ligation, this new synthetic methodology was exploited by using it as one of four amide bond-forming methods in the orthogonal assembly of a protein (Nilsson et al., 2003). These efforts are described in Chapter 4. RNase A was used as a model system for semisynthesis to demonstrate the potential advantage of the orthogonality of the Staudinger ligation to other amide-bond forming techniques. The amide bond-forming techniques used included mRNA translation by the ribosome, expressed protein ligation, solid-phase peptide synthesis, and the Staudinger ligation. A judicious combination of these tools holds great promise for the expansion of our ability to access proteins by semisynthesis.

Table 1.1 summarizes how the various amide bonds in this semisynthetic RNase A were formed. Fragment 1–109 was prepared as a *C*-terminal thioester by rDNA techniques. The amide bonds in this segment were formed by mRNA translation by the ribosome. Fragment 110–124 was prepared by a several methods. The amide bonds between residues 110–111 and 113–124 were formed by standard solid-phase peptide synthesis utilizing HATU activation. The protected fragment 110–111 was elaborated as a phosphinothioester of $HSCH_2PPh_2$ (**1.12**). The protected fragment 113–124 was elaborated as an *N*-terminal azide while immobilized to resin. The phosphinothioester of 110–111 was coupled with this 113–124 fragment on the solid phase using the Staudinger ligation. Finally, fragment 110–124 was liberated from resin and deprotected. Residue 110 is cysteine. This enabled coupling of fragment 110–124 to 1–109 via expressed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

protein ligation to give semisynthetic RNase A. The k_{cat}/K_m of this semisynthetic RNase A was 9.4×10^6 $M^{-1}s^{-1}$, compared to 1.1×10^7 $M^{-1}s^{-1}$ for biosynthetic RNase A. The identity of the semisynthetic enzyme was further characterized by mass spectrometry. It is also noteworthy that this protein incorporated a single carbon-13 NMR label at position 114.

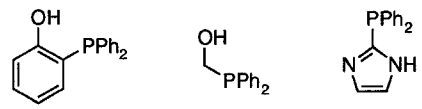
Bond	Coupling Method
1-109	mRNA Translation by Ribosome (sequential)
109-110	Expressed Protein Ligation (convergent)
110-111	Solid-Phase Peptide Synthesis (sequential)
111-112	Staudinger Ligation (convergent)
112-124	Solid-Phase Peptide Synthesis (sequential)

Table 1.1 *Amide bond-forming methods used in an orthogonal assembly of RNase A.*

Other phosphines have also been explored as reagents for carrying out traceless forms of the Staudinger ligation (Figure 1.3) (Saxon et al., 2000). In the initial report for these compounds, an acyl transfer between an acetyl ester (or amide in the case of compound **1.16**) derivative of the respective phosphine and a 5'-azido nucleoside was used to test the efficacy of these compound for use in the Staudinger ligation. Only compounds **1.14** and **1.16** were shown to promote efficient ligation, while no ligated product was observed when compound **1.15** was used. No attempts were initially made to ligate two amino acids together. A second study has shown that phosphine **1.14** can promote Staudinger ligation between two non-glycyl amino acids (Merkx et al., 2003). Under specialized

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reaction conditions (anhydrous THF, 47 °C, followed by addition of water) Gly–Phe and Ala–Phe couplings occurred in 32–36% yields. These results represent the first successful nonglycyl couplings via the Staudinger ligation. The reaction conditions are not, however, readily adapted to unprotected peptides in aqueous solution. These results do represent an important step in outlining some of the important chemical factors in improving the Staudinger ligation as a sequence independent coupling method. Chapter 6 of this thesis describes the use of molecular modeling calculations in the development of an analog of compound **1.14** that gives improved yields in a model Staudinger ligation.



1.14 **1.15** **1.16**

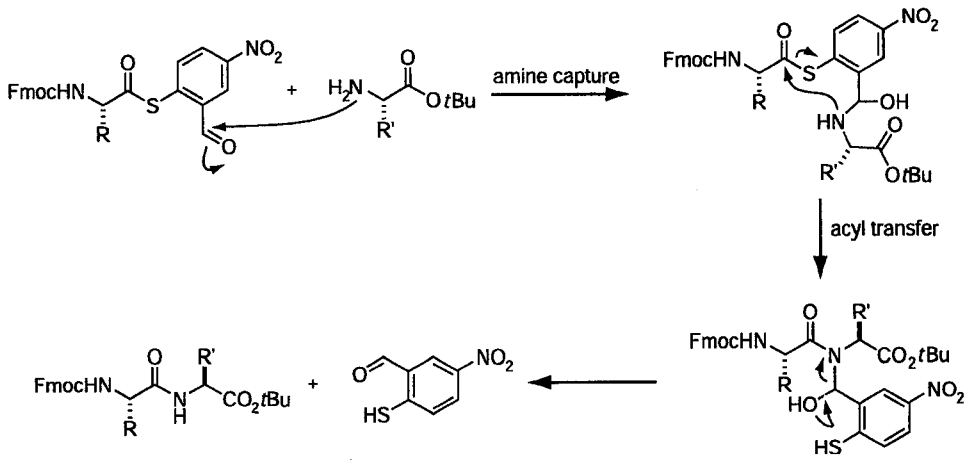
Figure 1.3 *Non-thiol-containing phosphines used for the traceless Staudinger ligation.*

1.6.5 Peptide Bond Formation using a Formyl-substituted Nitrophenylthioester

There has been a recent report detailing the use of formyl-substituted nitrophenylthioesters as peptide ligation agents (Scheme 1.24) (Ishiwata et al., 2003). The capture step is attack of an amine on the aldehyde group to give the intermediate hemiaminal. This hemiaminal intermediate can rearrange in an acyl transfer step. The rearranged product then decomposes to give the native peptide. This method has been demonstrated with model dipeptides. Couplings at AcGly-Val*O*t*Bu* and AcVal-Val*O*t*Bu* proceed in >90% yield in 2.5 h. Although this is an efficient coupling between two β-branched amino acids, the highly activated *p*-nitrophenylthioester is not generalizable to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

unprotected peptides because the formyl group has no inherent selectivity to a particular type of amine.



Scheme 1.24 *Peptide formation using formyl-nitrophenylthioesters.*

1.7 Conclusion

The total chemical synthesis of proteins represents a powerful method to probe the structure and function of proteins in ways that are difficult or impossible with standard expression using recombinant DNA methods. Indeed, it is becoming possible not only to synthesize a protein, but to endow natural protein scaffolds with desirable properties. The key advances in peptide synthesis of the last century are a strong foundation for forays into the total synthesis of proteins. Emerging technology for the chemoselective ligation of synthetic peptide fragments is building upon this foundation and making the total chemical synthesis of proteins an increasingly accessible endeavor.

The development and refinement of ligation methods for protein assembly has entered an explosive phase of growth. There are currently multiple research groups striving to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

achieve the goal of a ligation technology that will practically and efficiently ligate between any two amino acids in a protein sequence. Technology such as auxiliary-assisted chemical ligation and the Staudinger ligation represent a promising start. There are without question other imaginative and innovative approaches that will soon be developed that will make the total chemical synthesis of proteins, not only possible and practical, but routine. The marriage of chemical synthesis and biology in the realm of protein structure and function will be among the most dynamic fields of discovery for the next century.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 2*

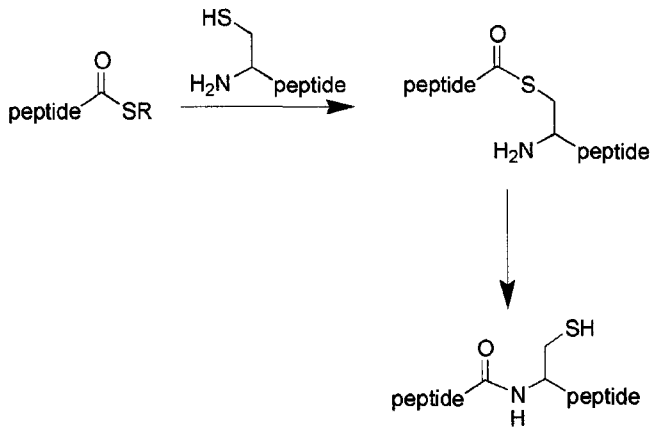
STAUDINGER LIGATION: A PEPTIDE FROM A THIOESTER AND AN AZIDE

New methods are facilitating the total chemical synthesis of proteins (Merrifield, 1986; Kent, 1988; Kaiser, 1989). In particular, Kent and others have developed an elegant means to stitch together two unprotected peptides in aqueous solution (Wieland et al., 1953; Dawson et al., 1994; Muir et al., 1997; Wilken & Kent, 1998; Kochendoerfer & Kent, 1999). In this method, which is called 'native chemical ligation', the thiolate of an *N*-terminal cysteine residue in one peptide attacks the carbon of a *C*-terminal thioester in another peptide to produce, ultimately, an amide bond between the two peptides (Scheme 2.1). Recently, Muir and others have expanded the utility of native chemical ligation by demonstrating that the thioester fragment can be produced readily with recombinant DNA (rDNA) techniques (Evans et al., 1998; Holford & Muir, 1998; Muir et al., 1998; Cotton & Muir, 1999; Ayers et al., 2000).

---

* This chapter has been published previously under the same title. Reference: Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939-1941.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



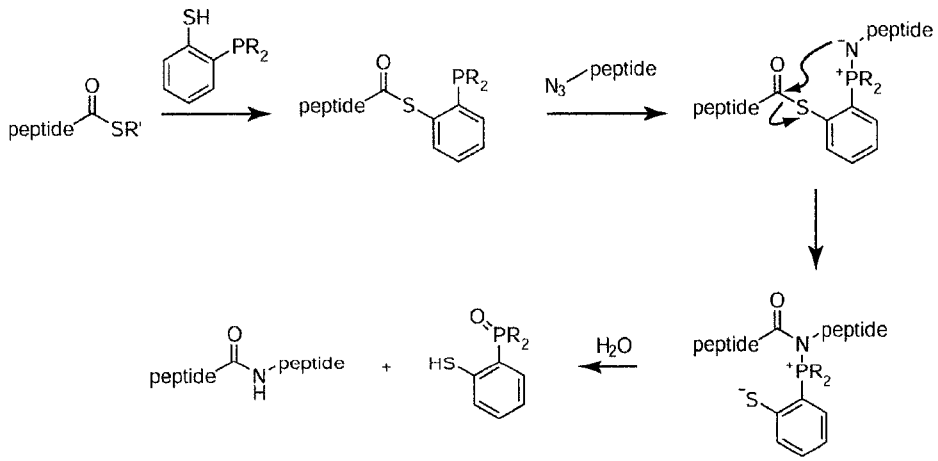
Scheme 2.1 *Native chemical ligation.*

Though powerful, native chemical ligation has a serious limitation. The method has an absolute reliance on the formation of a peptide bond to a cysteine residue (Homandberg & Laskowski, 1979; Wuttke et al., 1993; Vogel & Chmielewski, 1994; Beligere & Dawson, 1999). Creating a linkage at a natural Xaa–Cys bond is not always possible, as cysteine comprises only 1.7% of the residues in globular proteins (McCaldon & Argos, 1988b). Installing an extra cysteine residue is often undesirable. Cysteine is by far the most reactive residue towards disulfide bonds, $O_2(g)$, and other electrophiles (Raines, 1997). In addition, the sulfhydryl group of cysteine can suffer β-elimination to form dehydroalanine, which can undergo further reaction (Friedman, 1999). Thus, the impact of native chemical ligation would be even greater were it not limited to creating an Xaa–Cys bond.

Offer and Dawson have described a means to remove the limitation inherent in native chemical ligation (Offer & Dawson, 2000). In their method, a peptide bond is formed from a thioester and an *o*-mercaptobenzylamine. Though effective, this method is engrammic, leaving *o*-mercaptobenzylamine in the ligation product.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Here, we describe a method for peptide ligation that eliminates the need for a cysteine residue and leaves no residual atoms in the peptide product. Our method is inspired by the Staudinger reaction (Staudinger & Meyer, 1919; Gololobov et al., 1981; Gololobov & Kasukhin, 1992). In the Staudinger reaction, a phosphine is used to reduce an azide to an amine: $PR_3 + N_3R' + H_2O \rightarrow O{=}PR_3 + H_2NR' + N_2(g)$. The intermediate in the reaction is an iminophosphorane ($R_3P^+{-}{}^-NR'$), which has a nucleophilic nitrogen. Vilarrasa and others have shown that this nitrogen can attack an acyl donor in an intermolecular or intramolecular reaction (Bosch et al., 1993; Bosch et al., 1995; Bosch et al., 1996; Shalev et al., 1996; Ariza et al., 1998; Tang & Pelletier, 1998). The final product, after hydrolysis of the amidophosphonium salt, is an amide. Saxon and Bertozzi have shown that the acyl group can originate from the phosphine itself and be transferred to the iminophosphorane nitrogen in an intramolecular reaction in water (Saxon & Bertozzi, 2000a). Their product is an amide containing a phosphine oxide.



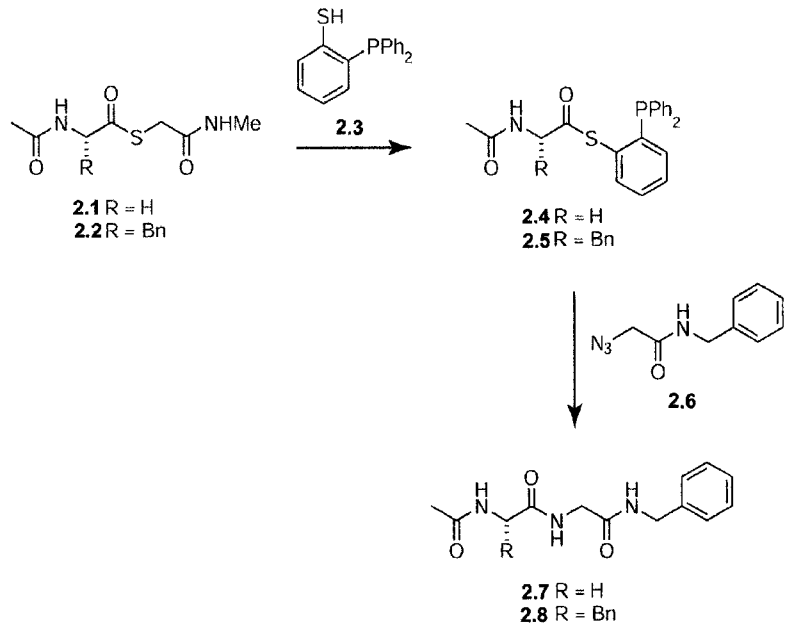
Scheme 2.2 *A traceless Staudinger ligation.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

To apply the Staudinger reaction to peptide synthesis, we use a phosphinothiol to unite a thioester and azide. A putative mechanism for this version of the 'Staudinger ligation' (Saxon & Bertozzi, 2000a) is shown in Scheme 2.2. The ligation begins by transthioesterification with the phosphinothiol. Coupling of the resulting phosphinothioester with a peptide azide leads to the formation of the reactive iminophosphorane. Attack of the iminophosphorane nitrogen on the thioester leads to an amidophosphonium salt. Hydrolysis of the amidophosphonium salt produces the desired amide and a phosphine oxide. Significantly, no atoms from the phosphinothiol remain in the amide product .

A critical aspect in effecting the Staudinger ligation of a thioester and azide is selecting an appropriate phosphinothiol. We chose an *o*-phosphinobenzenethiol ($R_2PC_6H_4$-*o*-SH) because it allows a six-membered ring to form in the transition state for acyl transfer. Moreover, $R_2PC_6H_4$-*o*-SH does not allow for the formation of an episulfide and a stable amidophosphine ($R_2PNR'C(O)R''$) by C–P bond cleavage in the amidophosphonium salt, as would simple alkyl thiols such as $R_2PCH_2CH_2SH$. Further, thiophenol itself is known to effect the transthioesterification of thioesters during native chemical ligation (Dawson et al., 1997).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 2.3 *Demonstration of traceless Staudinger ligation.*

We demonstrated the efficacy of the Staudinger ligation by effecting the transformations shown in Scheme 2.3 (R = H, Bn). In these transformations, the peptides AcGlyGlyNHBn (**2.7**) and AcPheGlyNHBn (**2.8**) were synthesized from a glycyl (**2.1**) and phenylalanyl (**2.2**) thioesters and a glycyl azide (**6**) by the action of *o*-(diphenylphosphino)benzenethiol (**2.3**) (Block et al., 1989). Thioesters **2.4** and **2.5** were prepared in quantitative yield by the transthioesterification of thioesters **2.1** and **2.2** with an excess of phosphinobenzenethiol **2.3** in DMF containing diisopropylethylamine (DIEA) (Vedejs & Diver, 1993). Excess thiol was removed by covalent immobilization to a Merrifield resin (chloromethylpolystyrene-divinylbenzene). Azide **2.6** (1 equivalent) was added to a solution of thioester **2.4** or **2.5** in unbuffered THF:H_2O (3:1), and the resulting solution was stirred at room temperature for 12 h. The reaction was then acidified by the addition of 2 N HCl, and solvents were removed under reduced pressure.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chromatography on silica gel gives purified amide **2.7** in 15% yield and amide **2.8** in 35% yield. The other major product was GlyNHBn, which can derive from the Staudinger reaction. The effect of alternative solvent conditions on the coupling efficiency of thioester **3** and azide **4** was explored. The reaction was performed in THF:H_2O (3:1) buffered at pH 2, 4, 8, and 13.5. The reaction was also performed in methylene chloride or dimethyl formamide, followed by an acidic aqueous workup. None of these conditions improved the yield of product compared to that obtained in unbuffered THF:H_2O (3:1).

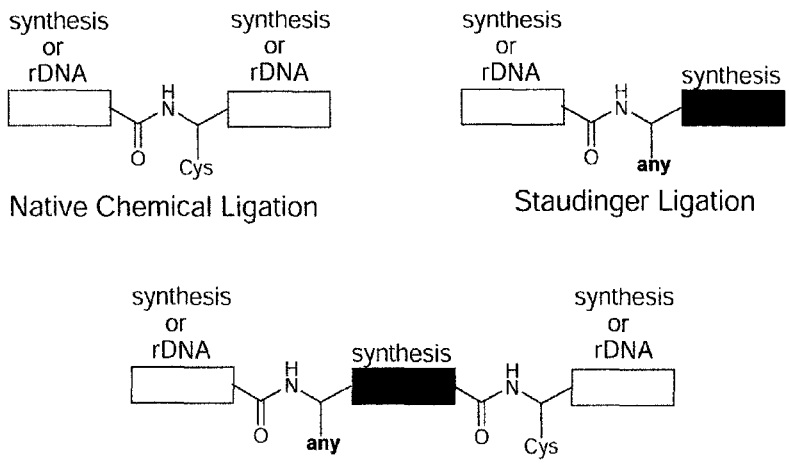
Amides **2.7** and **2.8** could be formed by a mechanism other than that in Scheme 2.2. Specifically, the amide product of the Staudinger ligation could, in theory, arise from the reduction of the azide followed by acyl transfer to the resulting amine. To test for this occurrence, we mixed thioester **2.5** and authentic GlyNHBn under conditions (reactant concentrations, solvent, temperature, and time) identical to those used to effect the ligation of thioester **2.5** and glycyl azide **2.6**. We saw no evidence for the formation of amide **2.8**. This result argues against the alternative mechanism.

Phosphinothiol **2.3** has the attributes necessary to effect the Staudinger ligation. Still, phosphinothiol **2.3** has low aqueous solubility and bestows a yield that may be too low for some applications. We anticipate, however, that these limitations can be overcome by structural optimization.

We note that an optimized version of the Staudinger ligation would expand the scope of protein synthesis. Staudinger ligation of a thioester fragment (Ingenito et al., 1999) and an azide fragment (Zaloom & Roberts, 1981) would be orthogonal to native chemical

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ligation as well as other strategies (Tam et al., 1995; Tam et al., 1999) for the coupling of unprotected peptides. Scheme 2.4 depicts the simplest proteins that would be accessible by native chemical ligation alone, the Staudinger ligation alone, and a sequential combination of the two methods. The use of thiol- or azide-protecting groups would extend the versatility of these methods even further.



Scheme 2.4 *Expanded scope of protein synthesis.*

Finally, we suggest that the Staudinger ligation could also be used with natural thioesters. For example, both the biosynthesis of polyketides and the nonribosomal biosynthesis of peptides proceed via the elaboration of thioester intermediates (Katz, 1997; Marahiel et al., 1997; vonDohren et al., 1997; Cane et al., 1998; Keating & Walsh, 1999). Interception of these intermediates with a phosphinothiol would allow for Staudinger ligation to an azide. Most significantly, ligation of a biosynthetic library of thioesters with a chemical library of azides could be a facile means to increase molecular

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

diversity.

Experimental Section

General Experimental

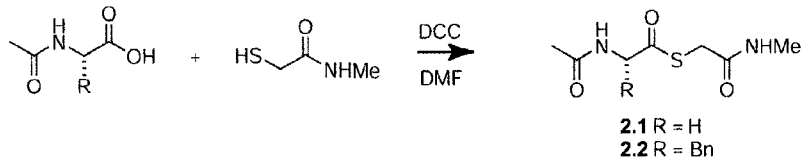
Chemicals and solvents were from Aldrich®, with the exception of *N*-methylmercaptoacetamide (Fluka®), bromoacetyl bromide (Acros®), and Merrifield resins (Novabiochem®). Merrifield resins (chloromethylpolystyrene-divinylbenzene) were 200–400 mesh (substitution 0.63 mmol/g) and 70–90 mesh (1.26 mmol/g). The progress of reactions was monitored by thin-layer chromatography using Whatman® TLC plates (AL SIL G/UV) with visualization by illumination with ultraviolet light or staining with I_2. NMR spectra were obtained with Bruker AC-300 and Varian UNITY-500 spectrometers. Phosphorus-31 NMR spectra were proton-decoupled and referenced against an external standard of deuterated phosphoric acid. Mass spectra were obtained using electrospray ionization (ESI) techniques at the University of Wisconsin Biotechnology Center.

Thioesters 2.1 and 2.2

An *N*-acetyl amino acid (*N*-acetyl glycine or *N*-acetyl phenylalanine) and one equivalent of *N*-methylmercaptoacetamide (NMA) were charged to a flame-dried reaction vessel under an argon atmosphere and dissolved in dry DMF to a final concentration of 0.5–0.7 M. 1,3-Dicyclohexylcarbodiimide (DCC; 1.1 equivalents) was added, and the resulting mixture was stirred at room temperature for 10–12 h. The 1,3-dicyclohexylurea (DCU) by-product was removed by filtration, and solvent was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

removed under reduced pressure. Products were recrystallized from CH_2Cl_2 and hexanes. Thioester **2.1** was obtained in a 90% yield, and thioester **2.2** was obtained in a 92% yield.



Scheme 2.5 *Synthesis of NMA thioesters.*

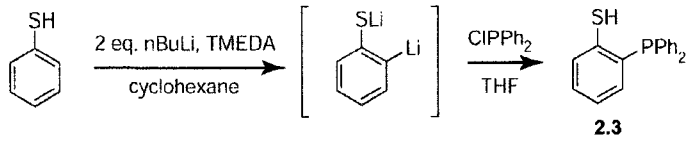
Thioester 2.1. ^{1}H NMR (DMSO-d_6, 300 MHz) δ8.62 (t, J = 6 Hz, 1 H), 8.05 (bs, 1 H), 4.00 (d, J = 6 Hz, 2 H), 3.56 (s, 2 H), 2.59 (d, J = 4.5 Hz, 3 H), 1.93 (s, 3 H) ppm; ^{13}C NMR (DMSO-d_6, 75 MHz) δ 197.95, 170.07, 167.13, 48.54, 31.81, 25.84, 22.26 ppm; MS (ESI) *m/z* 204.25 (M^+ = 204.9, fragments at 105.9, 100.0, 72.0).

Thioester 2.2. ^{1}H NMR ($CDCl_3$:CD_3OD, 1:1, 500 MHz) δ7.30–7.27 (m, 2 H), 7.24–7.19 (m, 3 H), 4.79 (dd, J = 10, 5 Hz, 1 H), 3.57 (apparent q, J = 15 Hz, 2 H), 3.24 (ABX, J = 14, 5 Hz, 1 H), 2.91 (ABX, J = 14, 10 Hz, 1 H), 2.76 (s, 3 H), 1.95 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$:CD_3OD, 1:1, 125 MHz) δ200.26, 172.83, 169.92, 136.93, 129.47, 128.98, 127.40, 61.24, 37.73, 32.72, 26.72, 22.40 ppm; MS (ESI) *m/z* 294.37 (MH^+ = 295.0 fragments at 190.0, 162.2, 120.2).

Phosphinothiol 2.3

Phosphinothiol **2.3** was prepared by the method of Block and coworkers, and NMR data (^{1}H and ^{31}P) correlated with their published data (Block et al., 1989).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



2.3

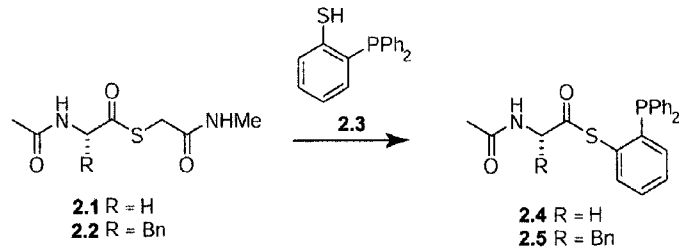
Scheme 2.6 *Synthesis of phosphinothiol* **2.3**.

Additional spectral data. ^{13}C NMR ($CDCl_3$, 125 MHz) δ137.71 (d, J = 30 Hz), 135.93 (d, J = 8.75 Hz), 135.35 (d, J = 9.75 Hz), 133.98, 133.83, 130.45, 129.25, 129.00, 128.67 (d, J = 6.75 Hz), 125.92 ppm; MS (ESI) *m/z* 294.35 (MH^+ = 295.0).

Thioesters 2.4 and 2.5

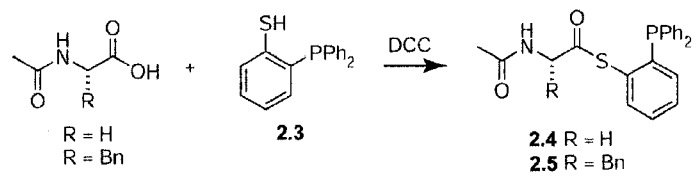
Method A (transthioesterification). Thioester **2.1** or **2.2** (1 equivalent) and phosphinothiol **2.3** (10 equivalents) were charged to a flame-dried reaction vessel under an argon atmosphere and dissolved in dry DMF (0.25 M). Dry argon was bubbled through the mixture, and diisopropylethylamine (DIEA, 5 equivalents) was added. The resulting mixture was stirred for 12 h, after which additional DIEA (5 equivalents) was added. Merrifield resin (either high or low loading capacity) having a loading capacity at least equivalent to the molar amount of phosphinothiol **2.3** was added to the mixture to remove excess phosphinothiol **2.3** and NMA. The resulting slurry was stirred for an additional 12 h under argon, and the resin was removed by filtration. Solvent was removed under reduced pressure, the residue was taken up in CH_2Cl_2, and the insoluble DIEA salts were removed by filtration. Solvent was again removed, and the residue was used in the subsequent coupling reaction without further purification. The reaction appeared to proceed in quantitative yield, as judged by TLC.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 2.7 *Transthioesterification using phosphinothiol* **2.3**.

Method B (DCC coupling). Thioester **2.1** or **2.2** (1 equivalent) and phosphinothiol **2.3** (1 equivalent) were added to a flame-dried reaction vessel under an argon atmosphere. DCC (1.1 equivalents) was added, and the mixture was stirred for 12 h. The DCU by-product was removed by filtration, and solvent was removed under reduced pressure. Thioesters **2.4** and **2.5** were purified by flash chromatography (silica gel, ethyl acetate:hexanes 1:1 followed by 100% ethyl acetate). Thioester **2.4** was obtained in 61% yield, and thioester **2.5** was obtained in 52% yield.



Scheme 2.8 *DCC coupling of phosphinothiol* **2.3** *to amino acids.*

Thioester 2.4. ^{1}H NMR ($CDCl_3$, 500 MHz) δ7.48 (ddd, J = 5.5, 4, 1.5 Hz, 1 H), 7.41 (td, J = 7.5, 1.5 Hz, 1 H), 7.37–7.32 (m, 7 H), 7.28–7.24 (m, 4 H), 6.92 (ddd, J = 7.5, 3, 1.5 Hz, 1 H), 5.86 (bs, 1 H), 4.07 (d, J = 6 Hz, 2 H), 2.02 (s, 3 H) ppm; ^{13}C NMR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

($CDCl_3$, 125 MHz) δ 194.83, 170.63, 144.04, 138.00, 136.68 (d, J = 10.75 Hz), 134.84, 134.69, 130.95, 130.30, 129.68, 129.34 (d, J = 6.88 Hz), 49.78, 23.69 ppm; ^{31}P NMR ($CDCl_3$, 202 MHz) δ–9.91 ppm; MS (ESI) *m/z* 393.44 (MH^+ = 394.2, fragments at 295.2, 225.2).

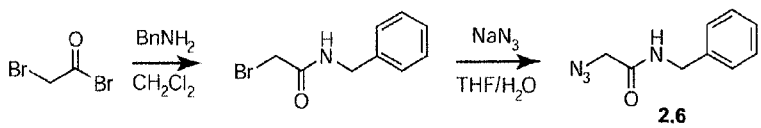
Thioester 2.5. ^1H NMR ($CDCl_3$, 500 MHz) δ 7.44 (ddd, J = 7.5, 4, 1.5 Hz, 1 H), 7.40 (td, J = 7.5, 1.5 Hz, 1 H), 7.36–7.31 (m, 7 H), 7.28–7.21 (m, 7 H), 7.12–7.10 (m, 2 H), 6.89 (ddd, J = 8, 3, 1 Hz, 1 H), 5.63 (d, J = 13.5 Hz, 1 H), 4.92 (m, 1 H), 2.95 (ABX, J = 14.5, 5.5 Hz, 1 H), 2.64 (ABX, J = 14, 8 Hz, 1 H), 1.91 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 197.47, 170.43, 137.94, 136.40, 134.83, 134.75, 134.67, 134.58, 130.79, 130.30, 129.94, 129.64 (d, J = 5.9 Hz), 129.33, 127.77, 60.29, 38.24, 23.79 ppm; ^{31}P NMR ($CDCl_3$, 202 MHz) δ–10.33 ppm; MS (ESI) *m/z* 483.56 (MH^+ = 484.2, fragment at 295.2).

Azide 2.6

Benzyl amine (20.4 mL, 186 mmol) and methylene chloride (186 mL) were added to a flame-dried reaction vessel under an argon atmosphere and the solution was cooled to 0 °C in an ice bath. Bromoacetyl bromide (8.1 mL, 93 mmol) was added dropwise to the solution. A precipitate, presumably the HBr salt of benzyl amine, formed almost immediately. The reaction mixture was warmed to room temperature and stirred for 1 h. The precipitate was removed by filtration, and the organic phase was washed with 2 N HCl (2 × 75 mL). The organic layer was dried over anhydrous magnesium sulfate and filtered, and solvent was removed under reduced pressure. The resulting white solid was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

dissolved in THF (200 mL) and water (50 mL). Sodium azide (30.3 g, 466 mmol) was added, and the resulting mixture was stirred vigorously at reflux for 17 h. The organic layer was then separated from the aqueous layer, washed with a solution of saturated brine (2 × 75 mL), dried over anhydrous magnesium sulfate, and filtered. Solvent was removed under reduced pressure. Azide **2.6** was isolated in 98% yield, and was used without further purification.



Scheme 2.9 *Synthesis of azide* **2.6**.

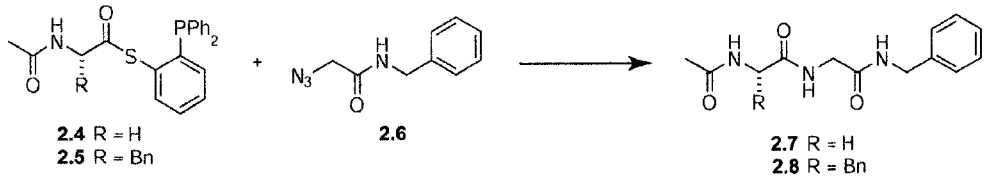
Spectral data. ^{1}H NMR ($CDCl_3$, 300 MHz) δ 7.39–7.27 (m, 5 H), 6.71 (bs, 1 H), 4.47 (d, J = 5.7 Hz), 4.00 (s, 2 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 166.66, 137.39, 128.43, 127.45, 127.35, 52.06, 43.08 ppm; MS (ESI) *m/z* 190.20 (MH^+ = 191.0, fragment at 91.2).

Amides 2.7 and 2.8

Thioester **2.4** or **2.5** (1 equivalent) and azide **2.6** (1 equivalent) were dissolved in THF:H_2O (3:1) to a concentration of 0.2 M. A yellow color, presumably from liberated thiolate, formed quickly. The mixture was stirred at room temperature for 12–16 h, and then acidified with 2 N HCl until the yellow color became clear. Solvent was removed under reduced pressure, and the amide products were separated from the phosphine oxide

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

by-products (which was characterized spectrally; data not shown) by flash chromatography (silica gel, 2.5–10% methanol in methylene chloride). Yields for amide **2.7** and amide **2.8** ranged from 15 to 35%.

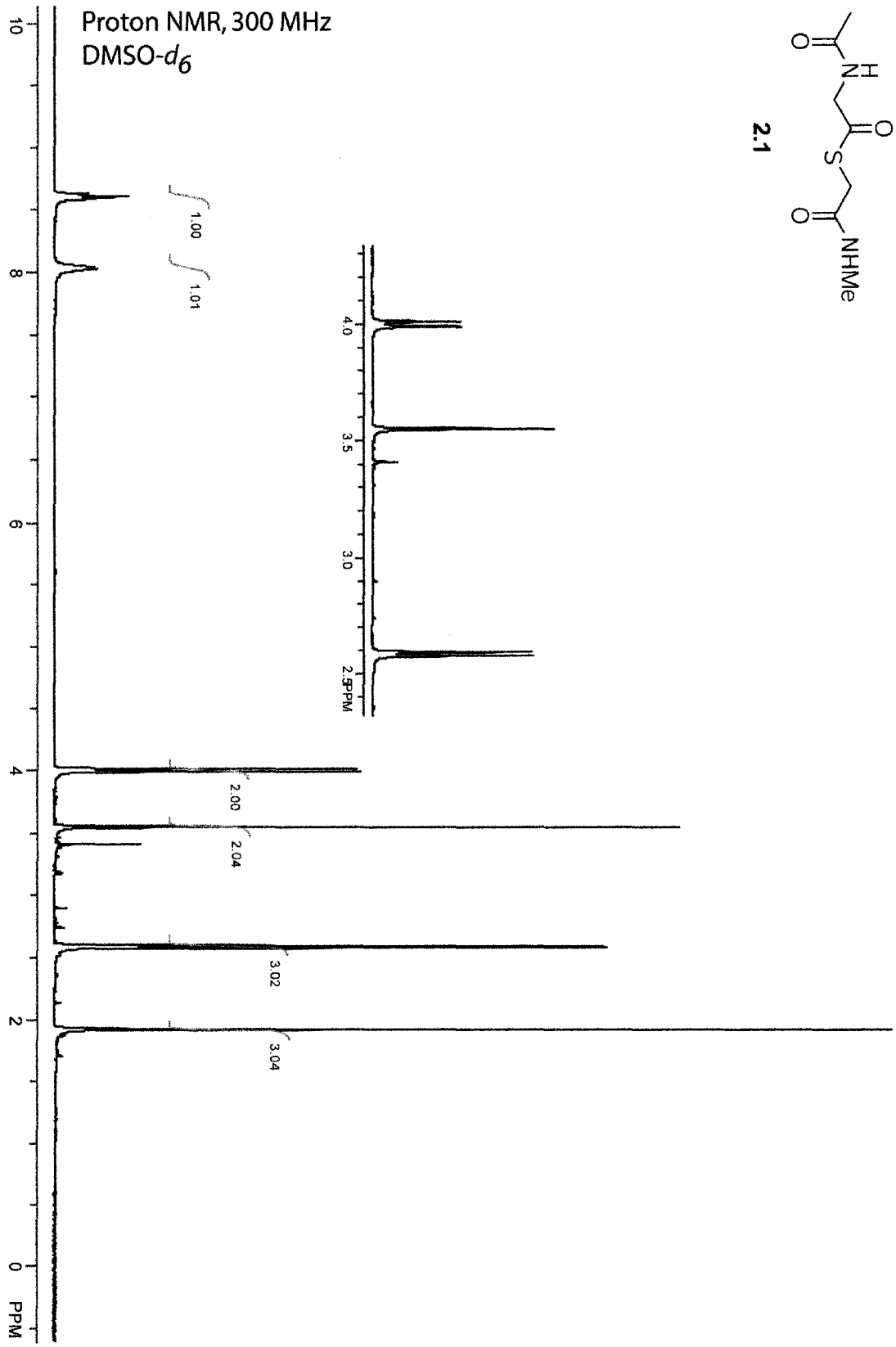


Scheme 2.10 *Staudinger ligation of phosphinothioesters and azides.*

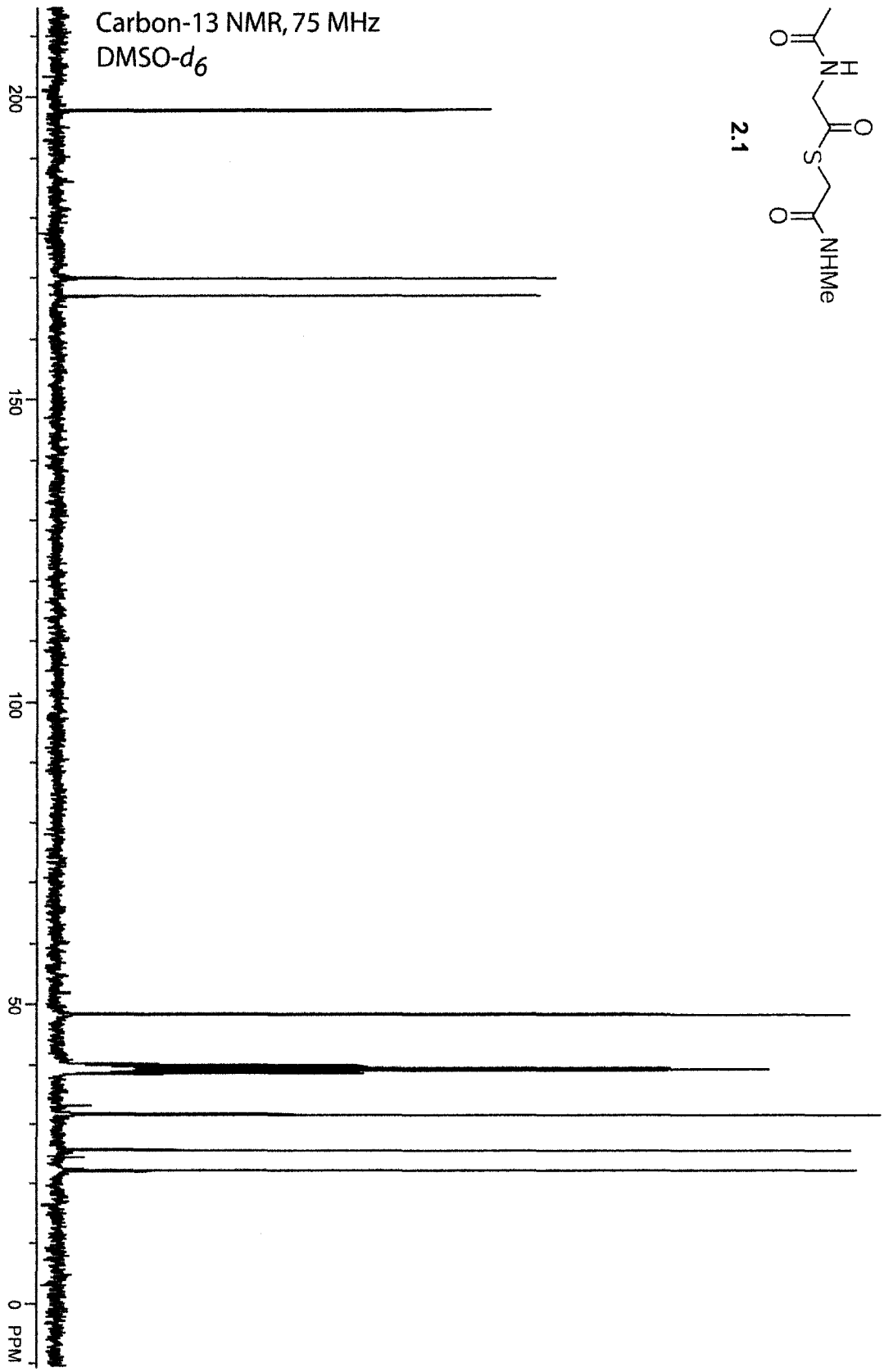
Amide 2.7. ^{1}H NMR ($CDCl_3$:CD_3OD, 1:1, 500 MHz) δ 7.33–7.22 (m, 5 H), 4.41 (s, 2 H), 3.92 (s, 2 H), 3.86 (s, 2 H), 2.01 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$:CD_3OD, 1:1, 125 MHz) δ 173.56, 171.52, 170.67, 138.83, 129.04, 218.02, 127.78, 43.70, 43.62, 43.16, 22.45 ppm; MS (ESI) *m/z* 263.29 (MH^+ = 264.0).

Amide 2.8. ^{1}H NMR ($CDCl_3$:CD_3OD, 1:1, 500 MHz) δ 7.32–7.19 (m, 10 H), 4.48 (apparent t, J = 7.5 Hz, 1 H), 4.44 (d, J = 15 Hz, 1 H), 4.34 (d, J = 14.5 Hz, 1 H), 3.95 (d, J = 16.5 Hz, 1 H), 3.71 (d, J = 16.5 Hz, 1 H), 3.11 (dd, J = 13.5, 7 Hz, 1 H), 2.94 (dd, J = 14, 8 Hz, 1 H), 1.88 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$:CD_3OD, 1:1, 125 MHz) δ 173.42, 172.81, 170.34, 138.61, 137.18, 129.55, 129.01, 128.93, 127.86, 127.68, 127.40, 56.12, 43.53, 43.18, 37.76, 22.36 ppm; MS (ESI) *m/z* 353.42 (MNa^+ = 376.2, MH^+ = 354.2 fragments at 165.2, 120.2, 91.2).

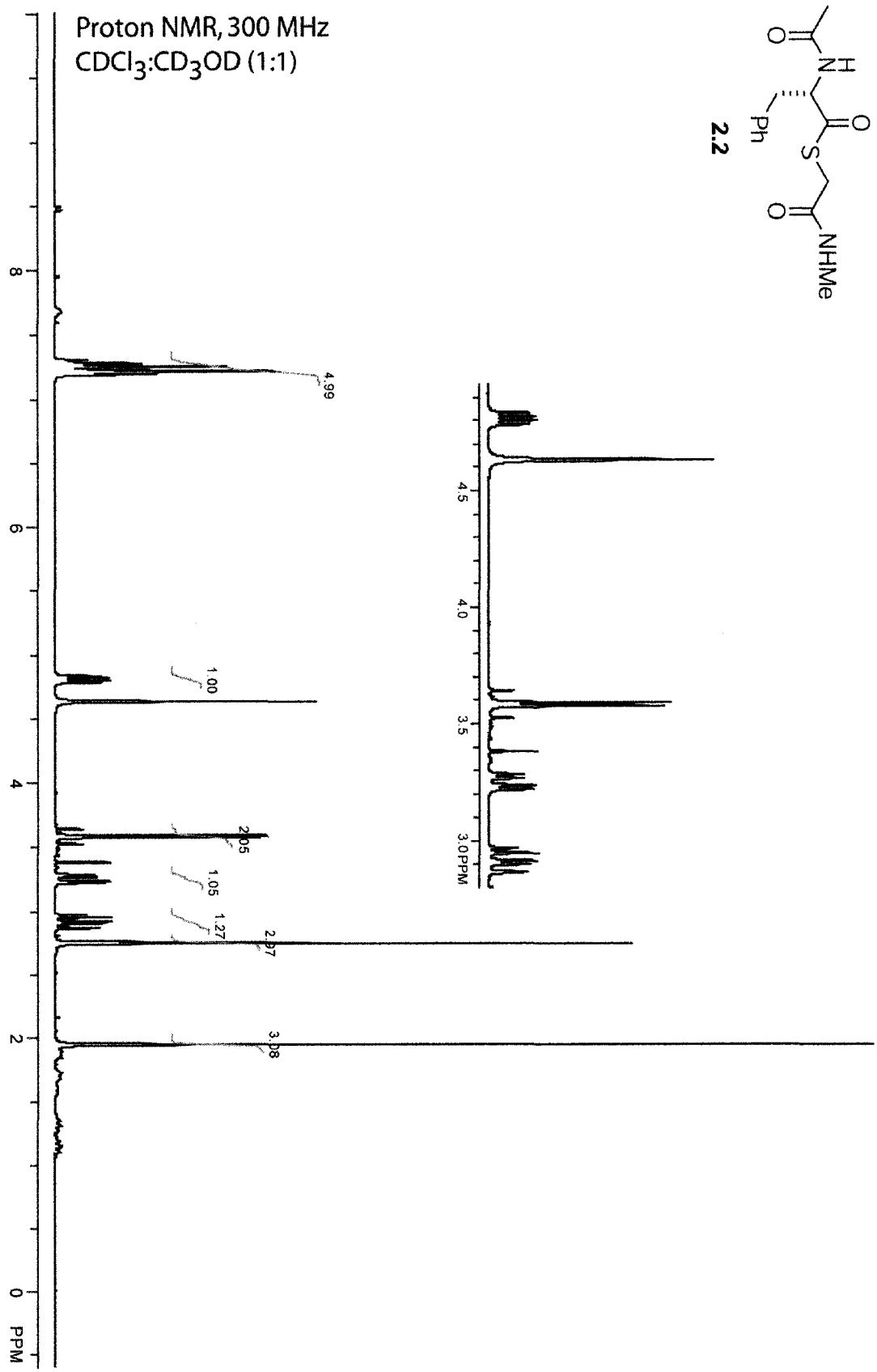
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



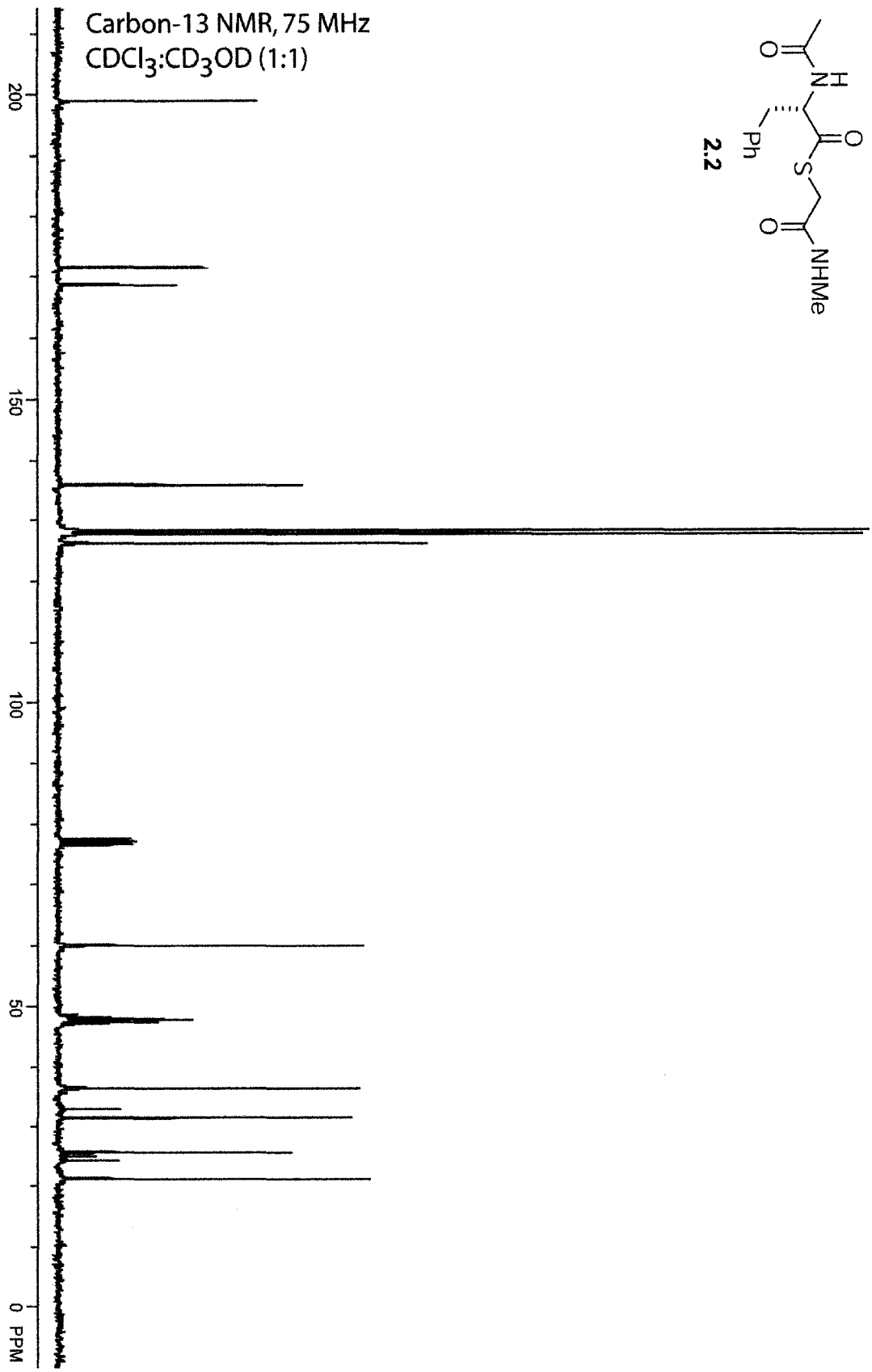
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



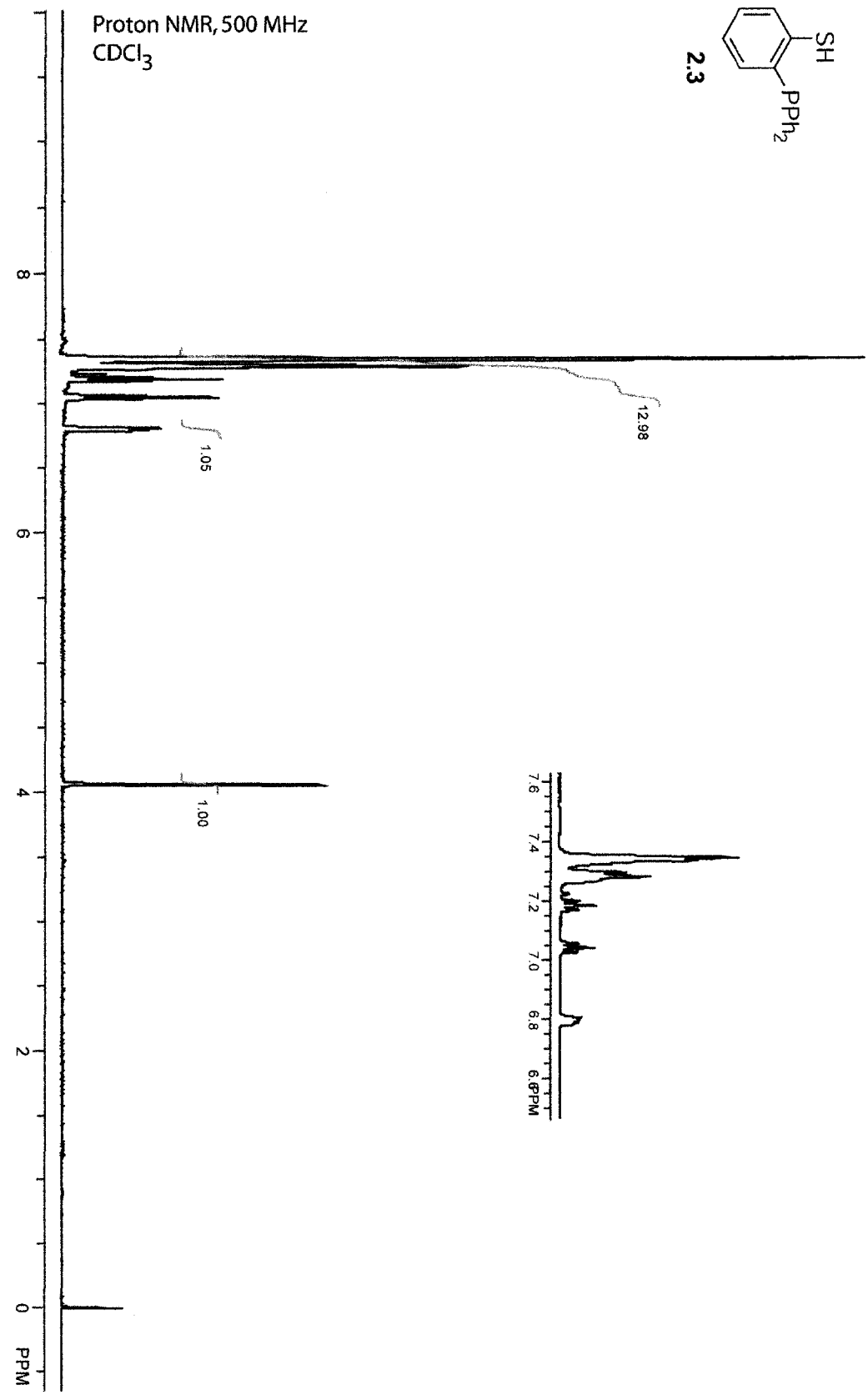
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



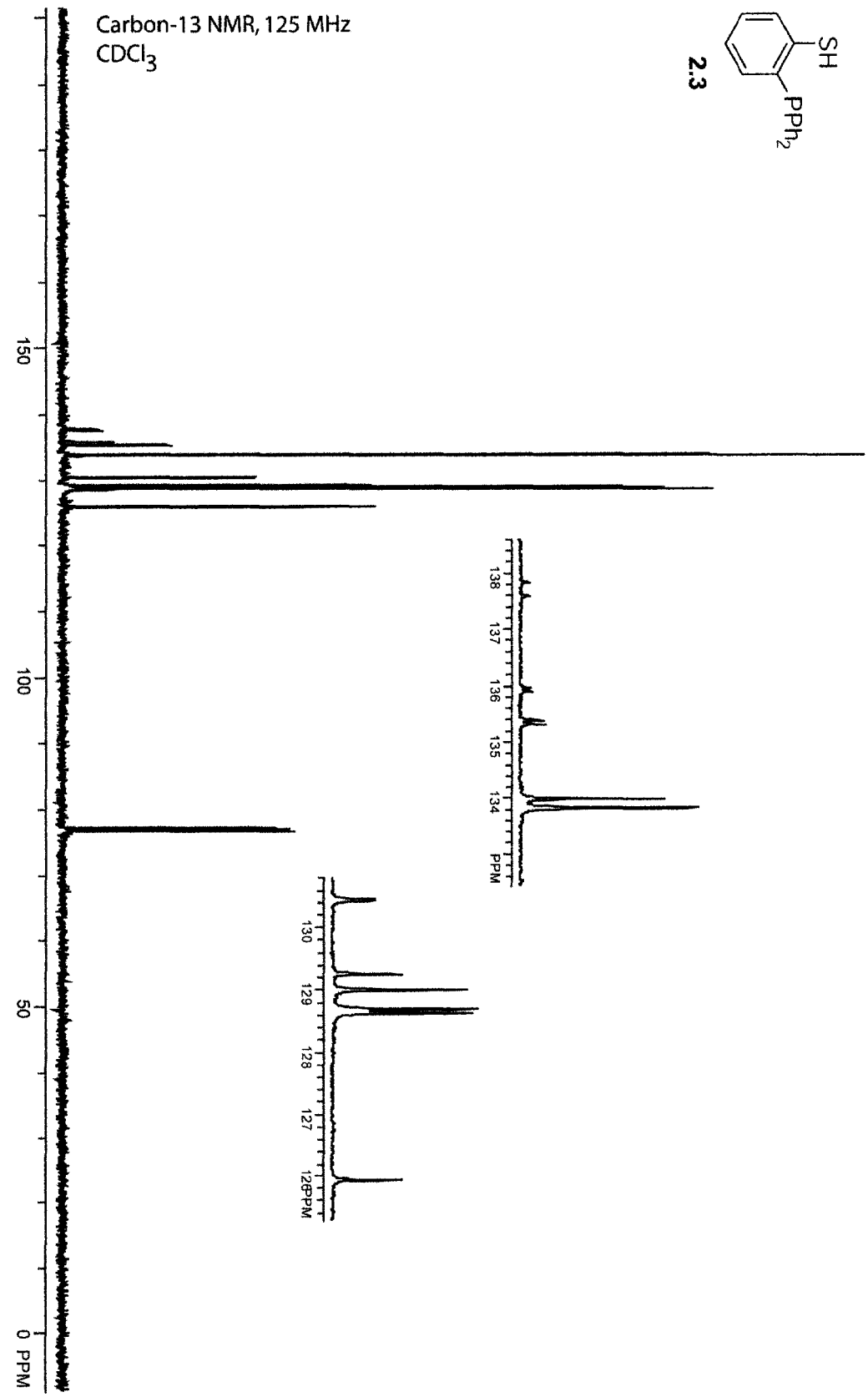
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



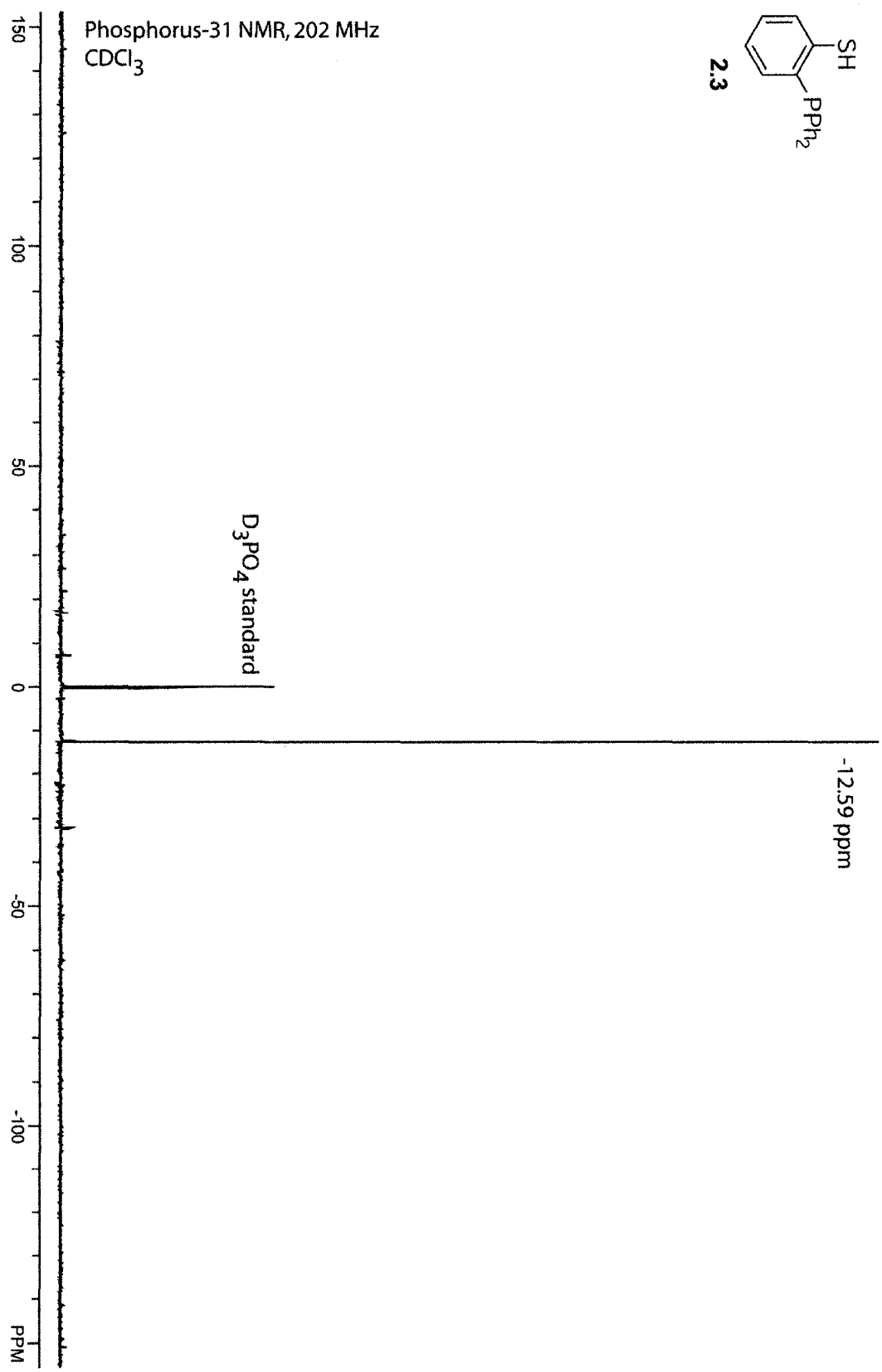
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



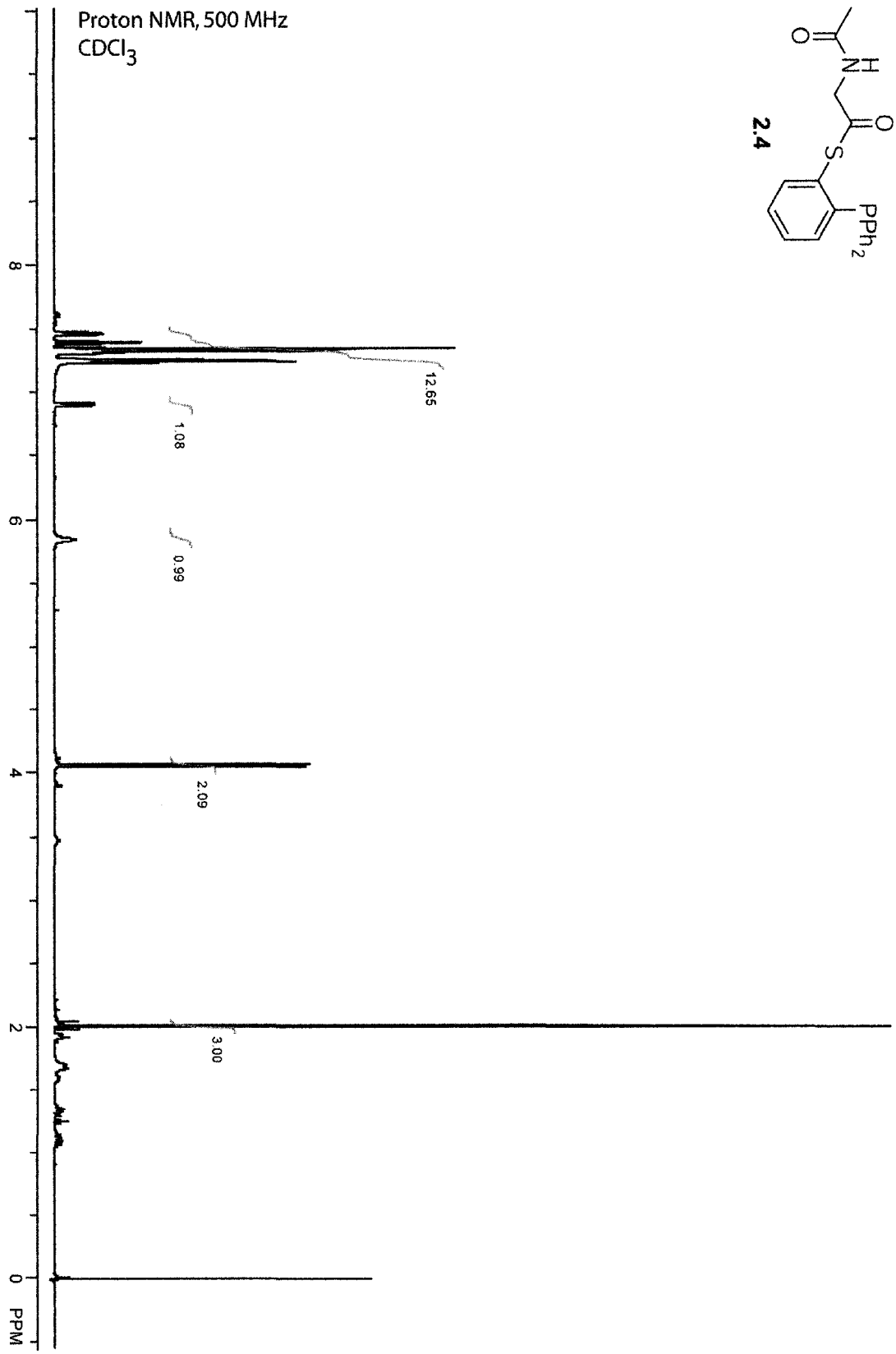
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

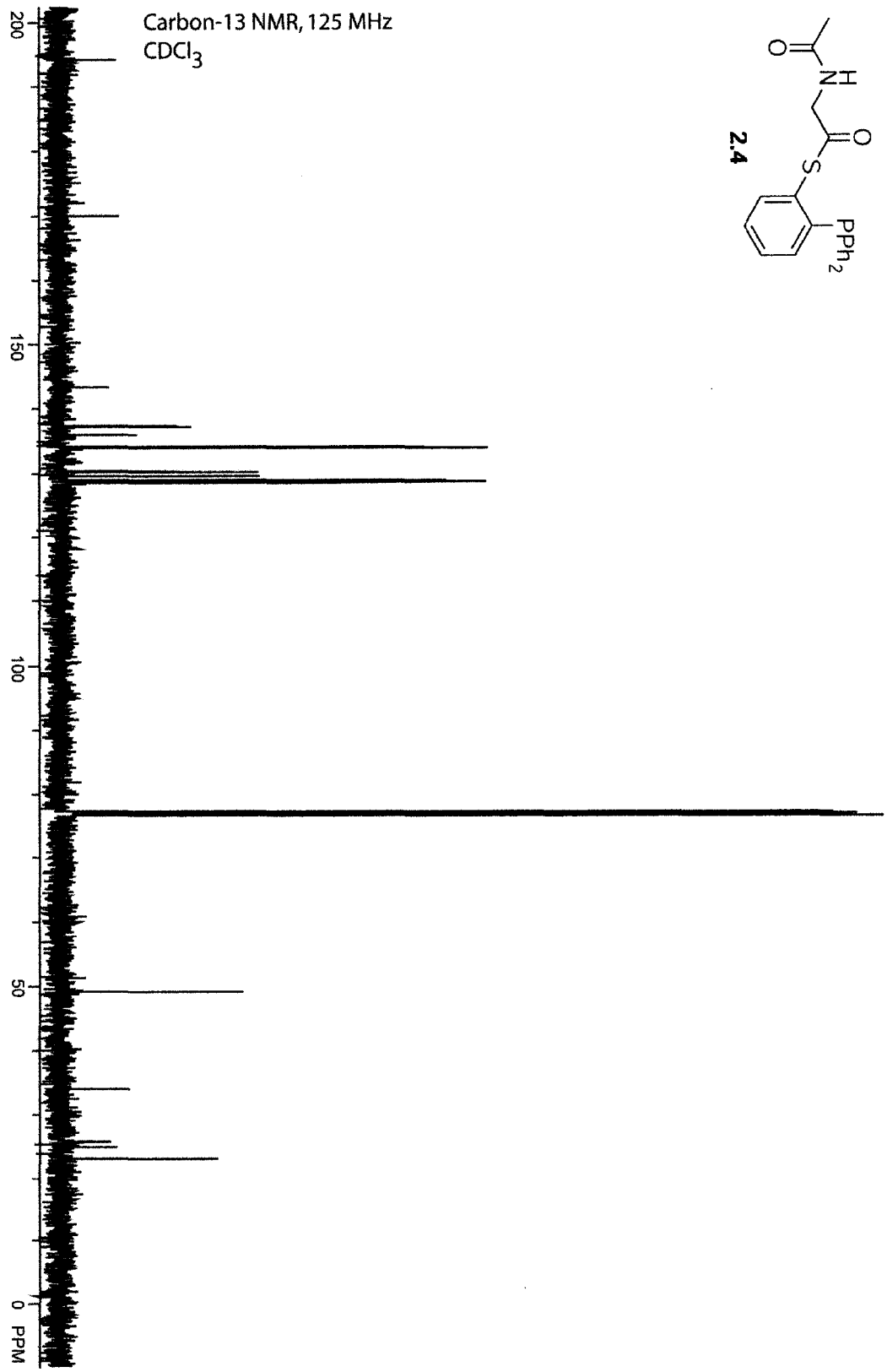


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

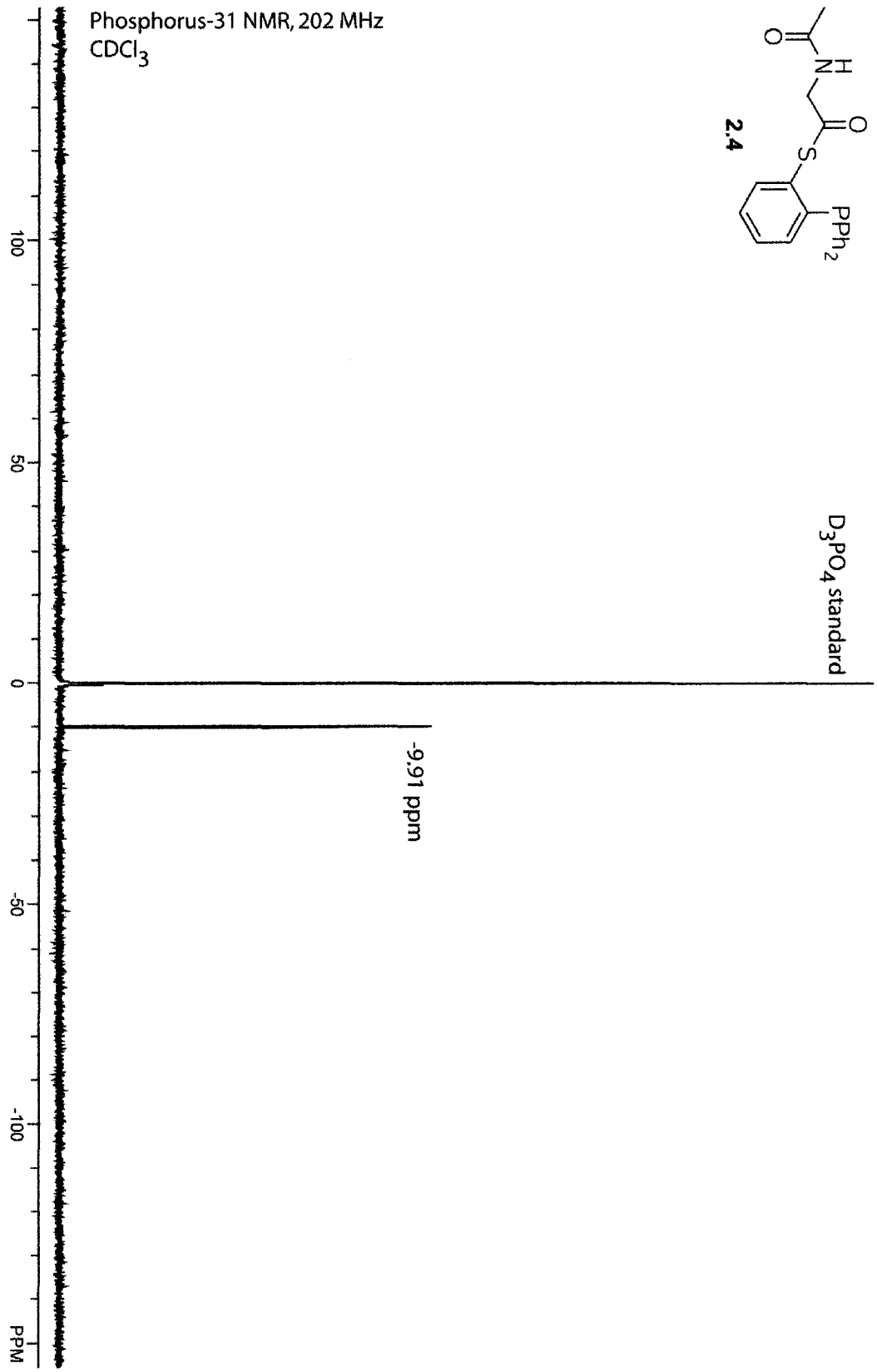


2.4

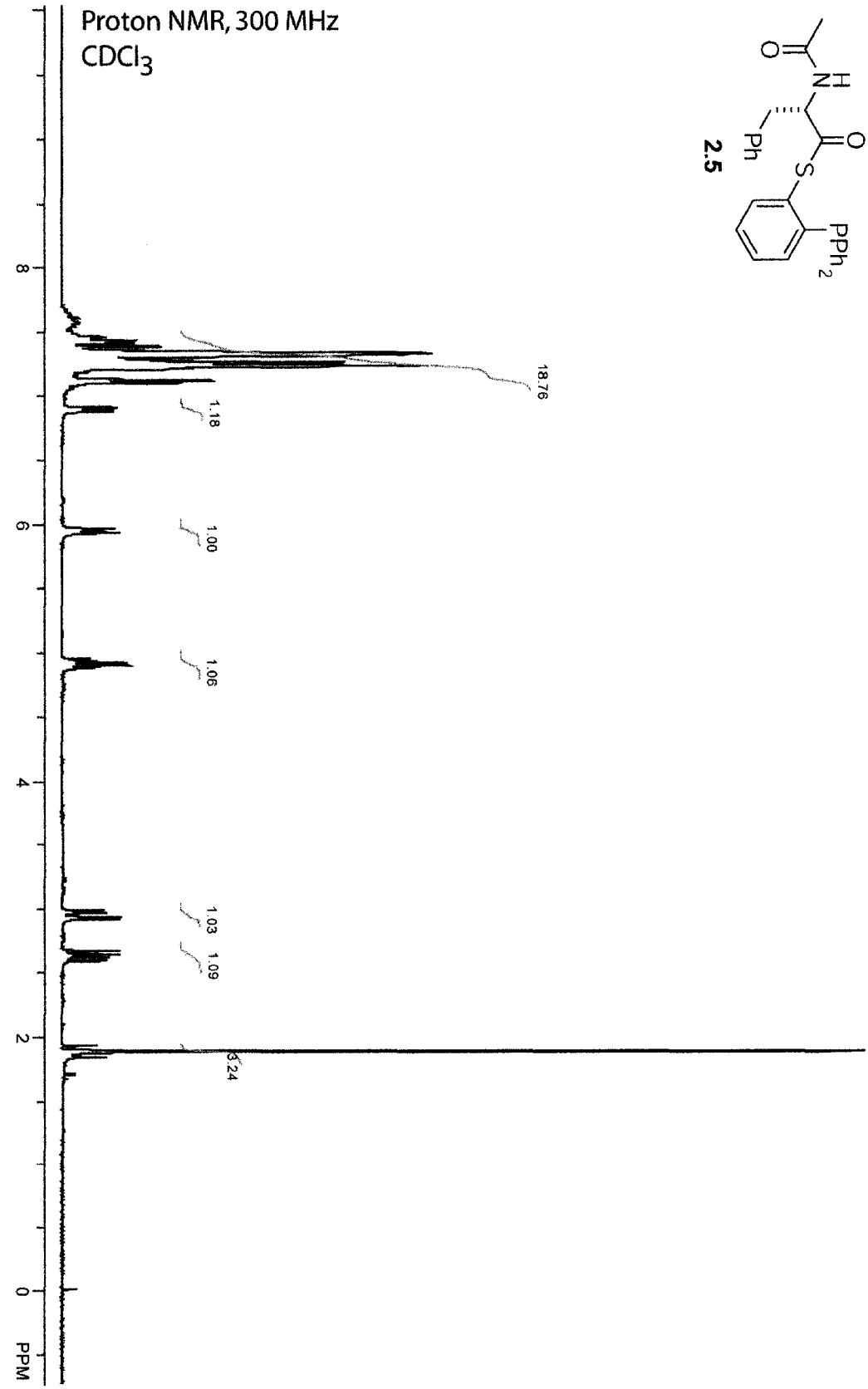
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

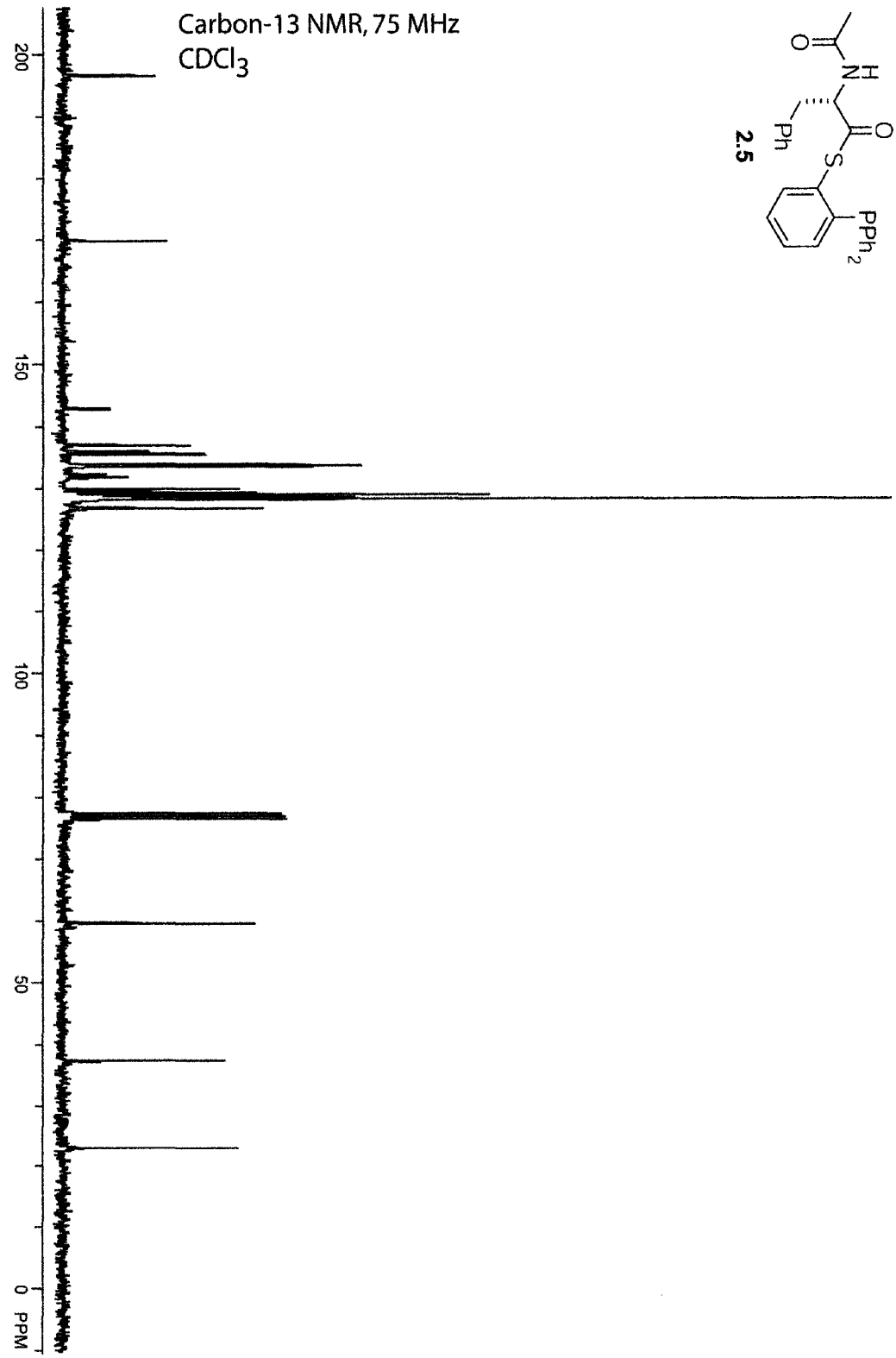


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

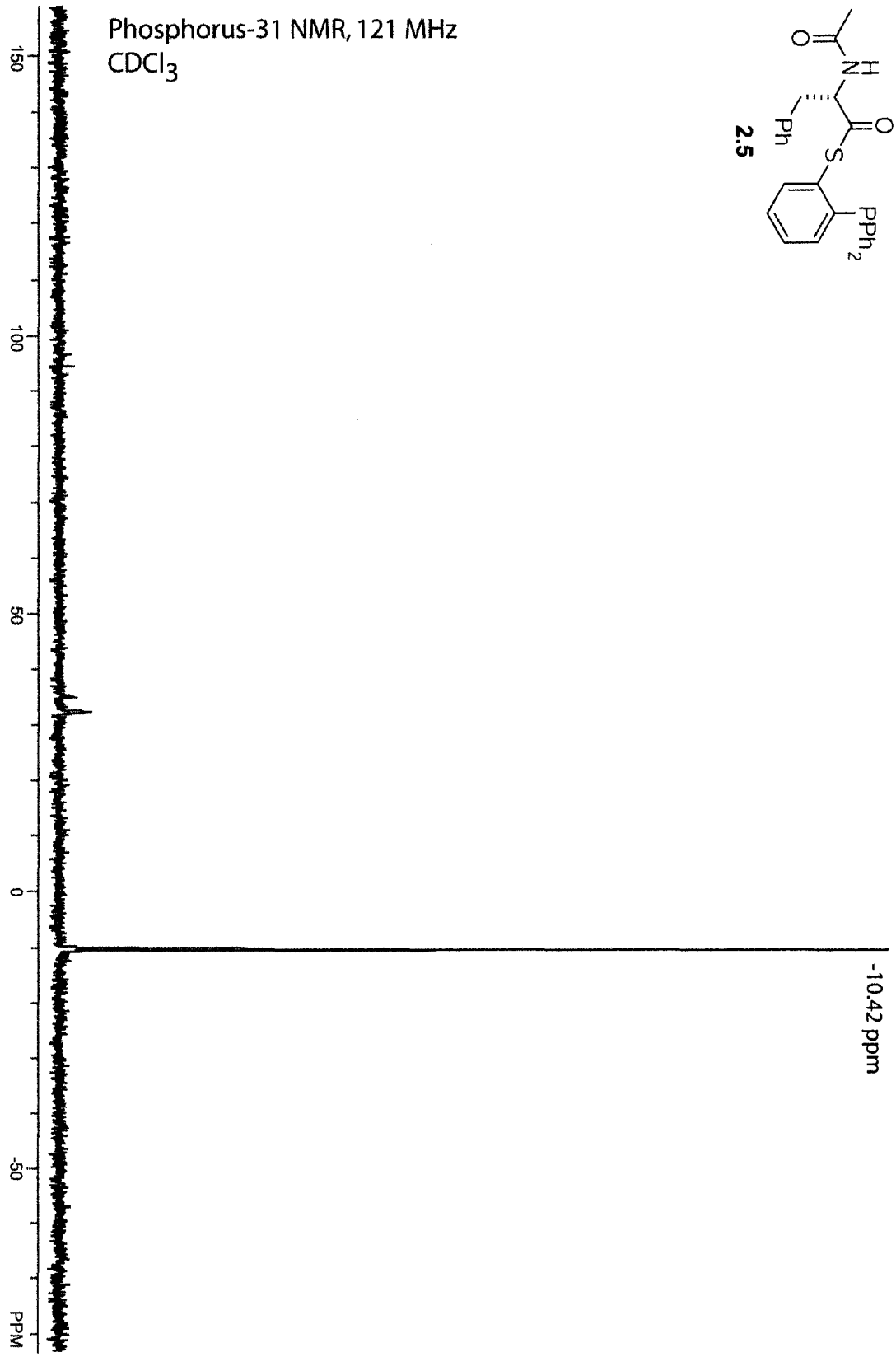


2.5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

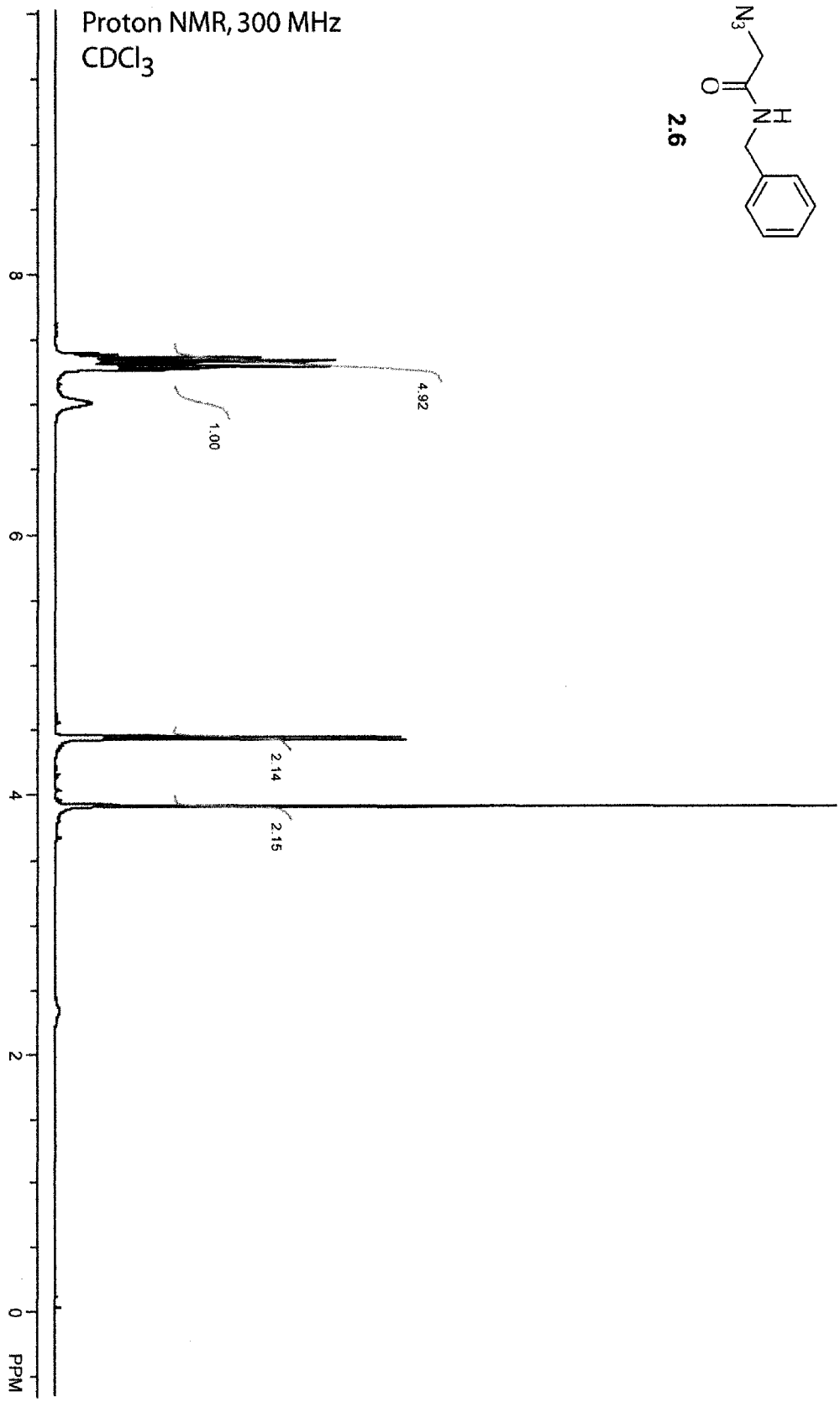


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

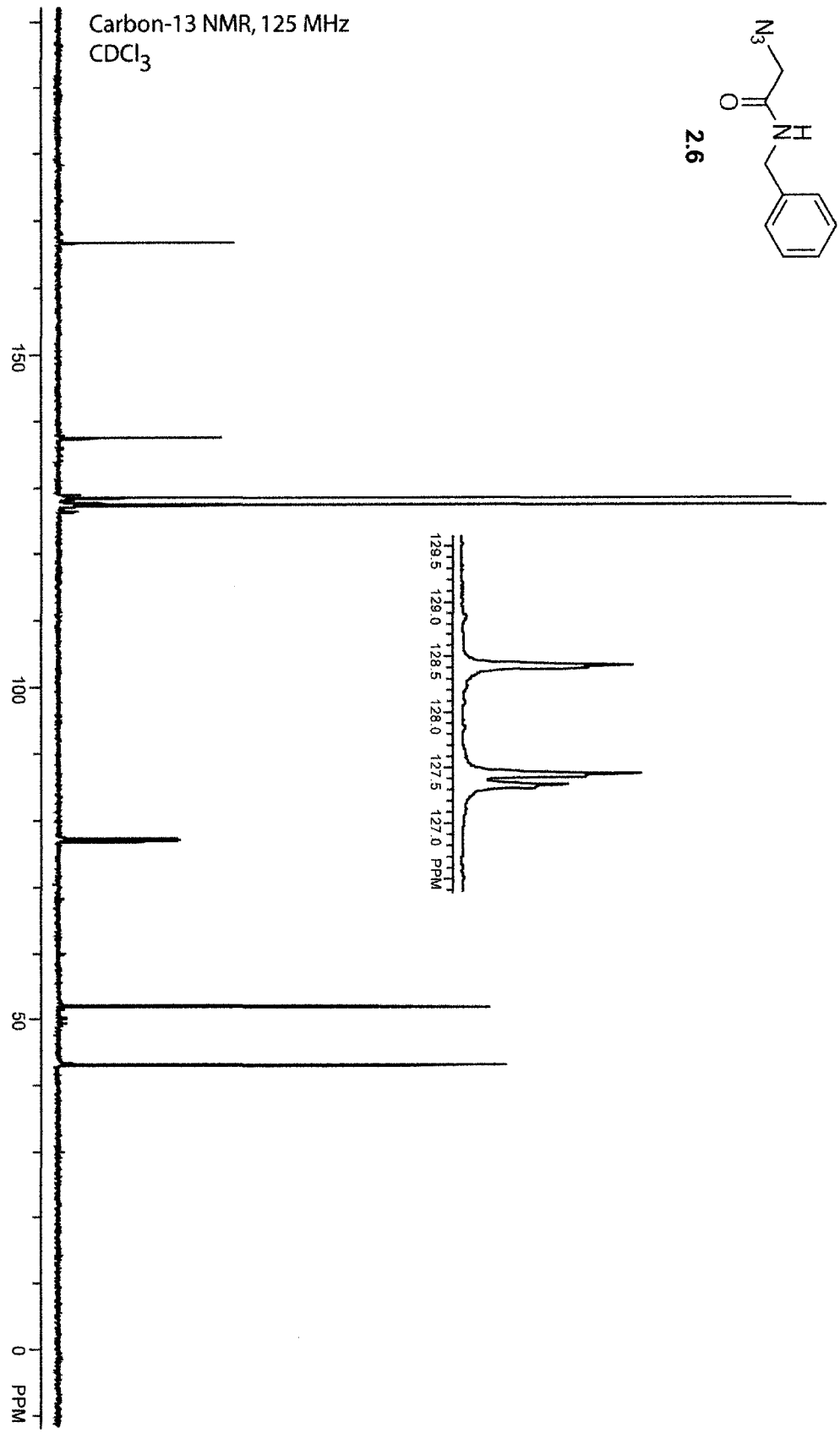


2.5

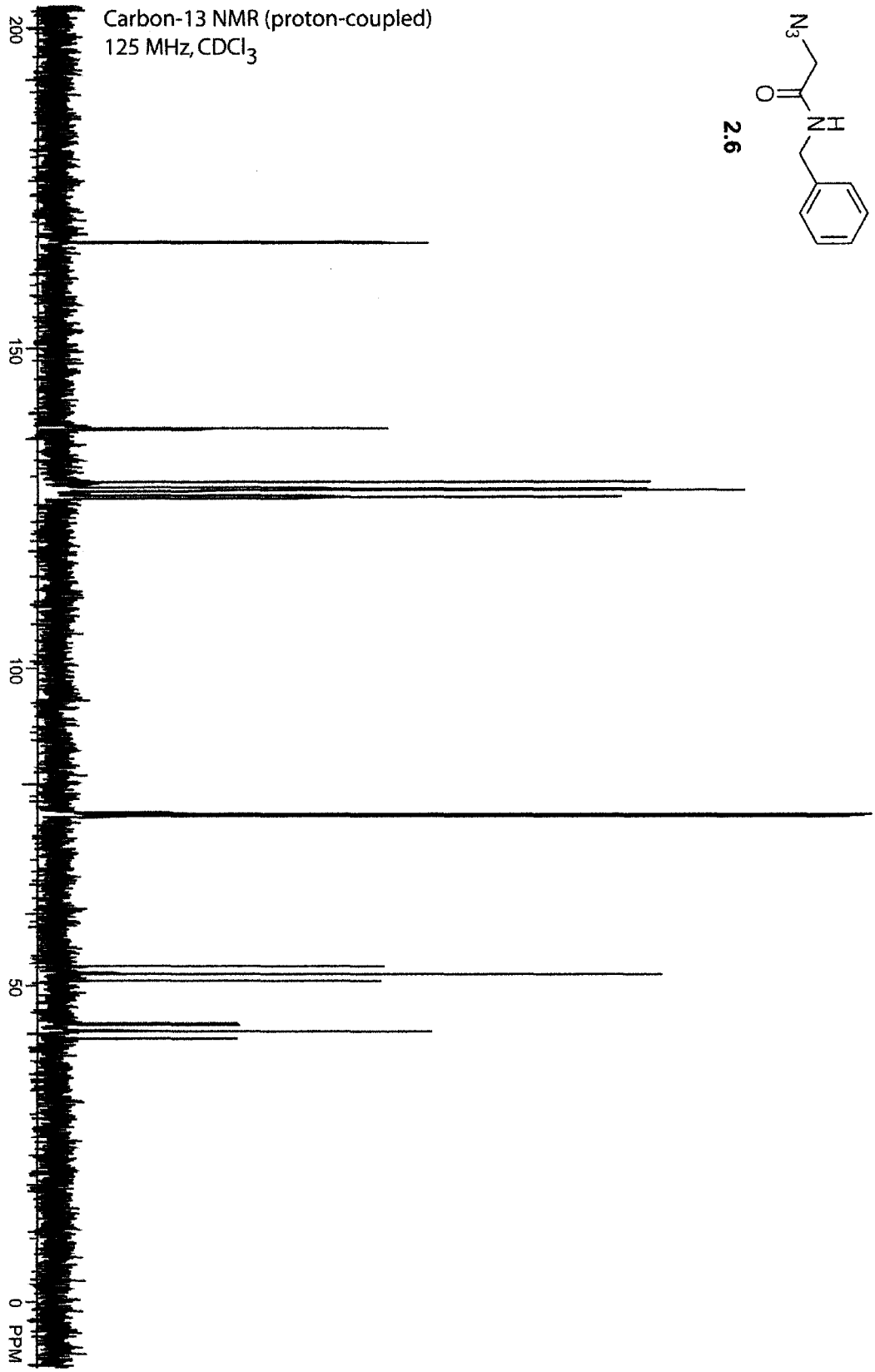
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



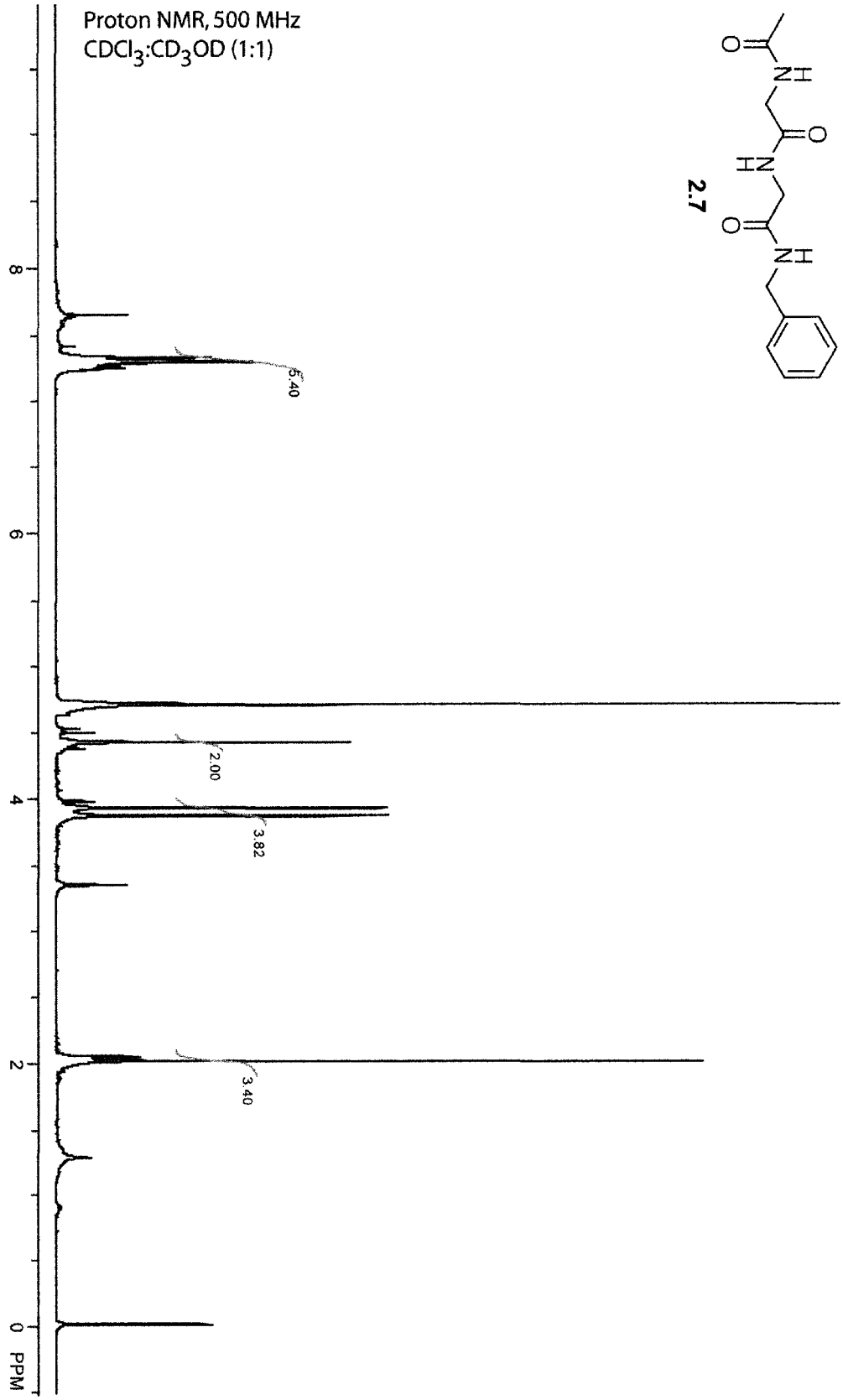
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



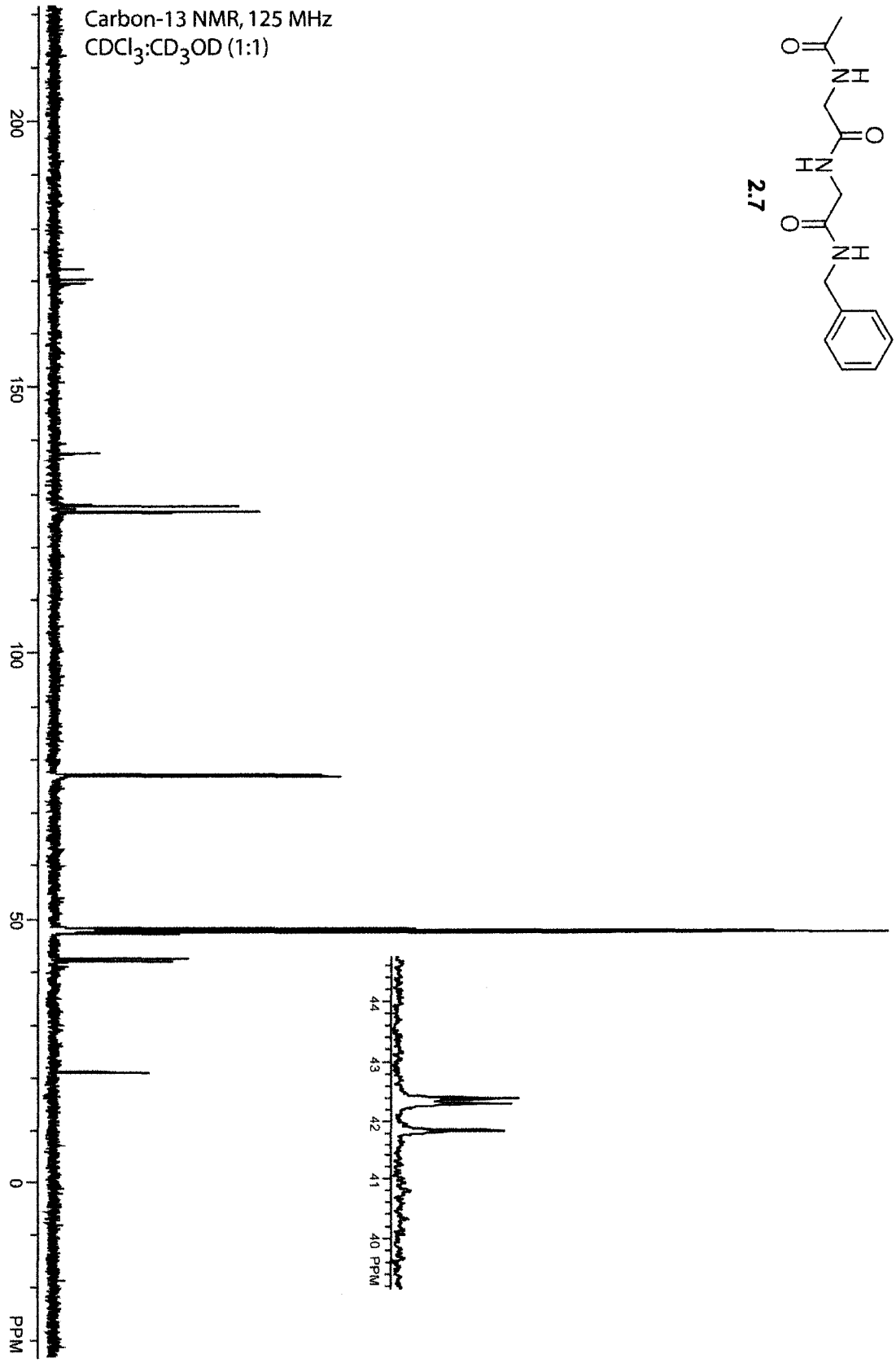
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

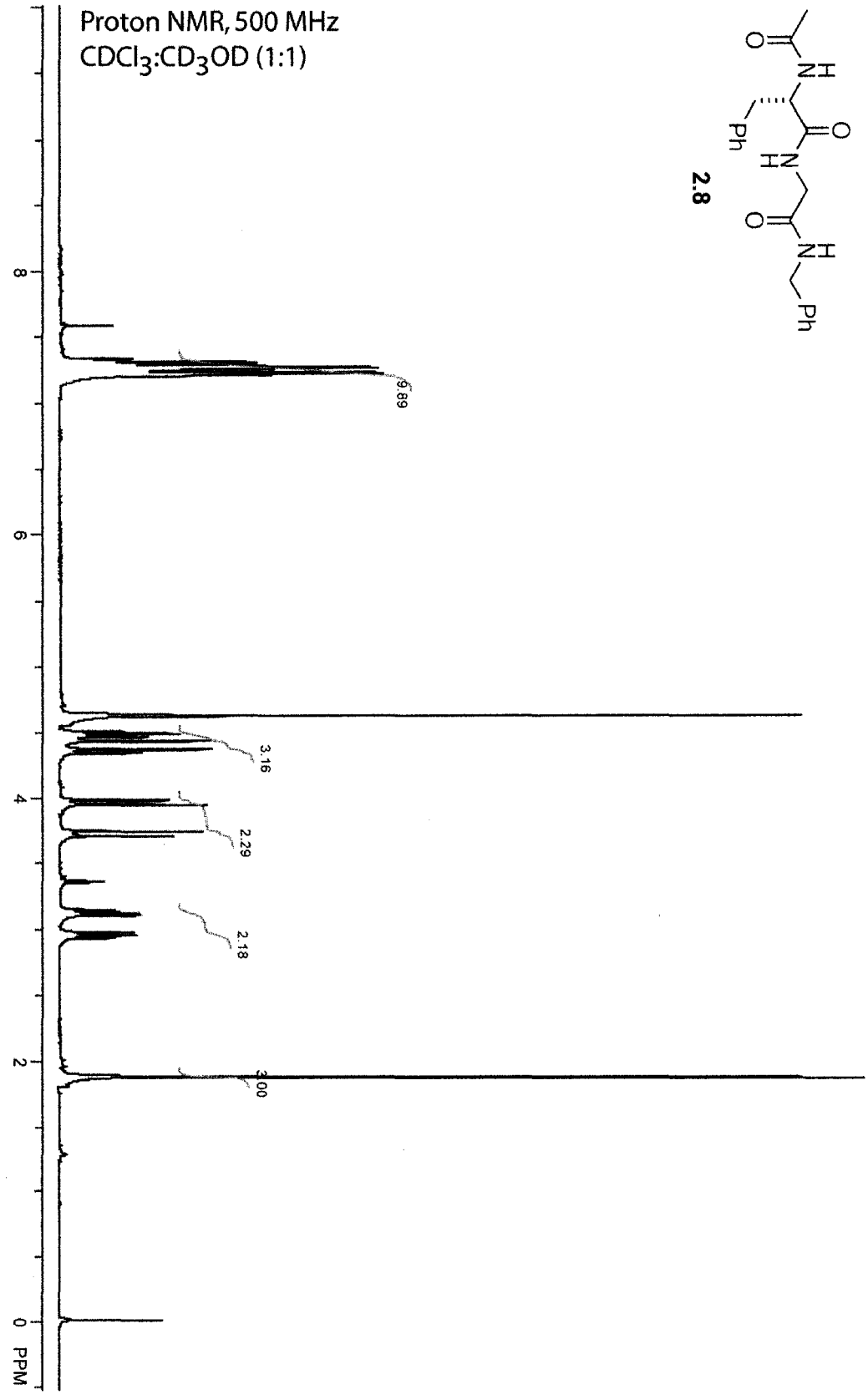


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



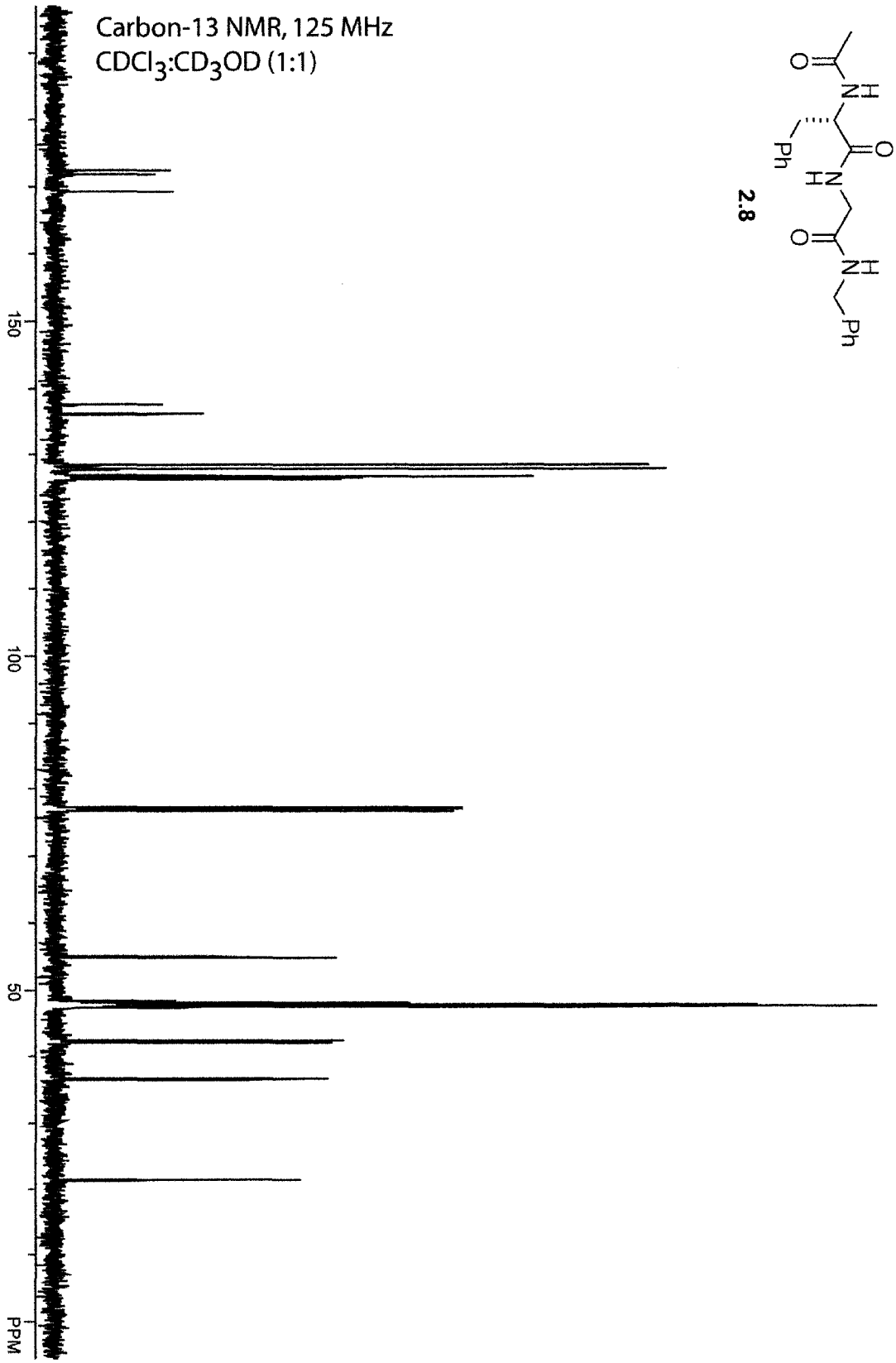
2.7

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



2.8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



2.8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 3[†]

HIGH-YIELDING STAUDINGER LIGATION OF A PHOPHINOTHIOESTER AND AZIDE TO FORM A PEPTIDE

The ligation of synthetic peptides provides a convergent route for the total chemical synthesis of proteins. Currently, the most common ligation method is "native chemical ligation". This ligation method was discovered in 1953, when the reaction of ValSPh and CysOH in aqueous buffer was shown to yield the dipeptide: ValCysOH (Wieland et al., 1953). In the 1990's, this seminal discovery was developed into a practical method to ligate large peptide fragments (Dawson et al., 1994; Tam et al., 1999; Borgia & Fields, 2000; Dawson & Kent, 2000). In native chemical ligation, the thiolate of an *N*-terminal cysteine residue of one peptide attacks the *C*-terminal thioester of a second peptide. An amide linkage forms after rapid S→N acyl transfer (Scheme 3.1). "Expressed protein ligation" is an extension of native chemical ligation in which the *C*-terminal thioester is produced by recombinant DNA (rDNA) technology rather than chemical synthesis (Holford & Muir, 1998; Muir et al., 1998; Cotton & Muir, 1999).

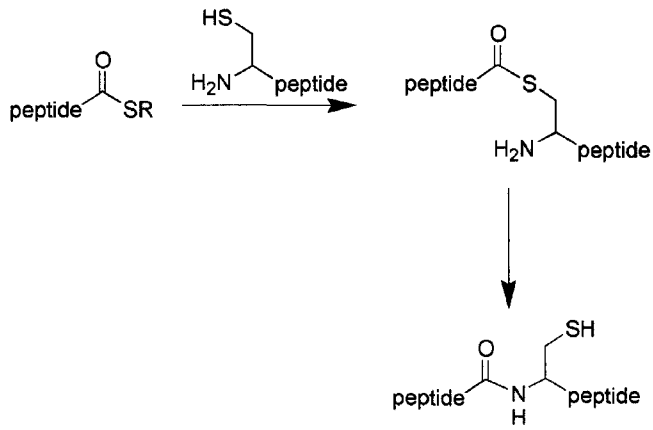
A limitation of native chemical ligation is its intrinsic reliance on having a cysteine

---

[†] This chapter has been published previously under the same title. Reference: Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2001**, *3*, 9-12. The information contained in Scheme 3.4 has been published previously in: Soellner, M. B.; Nilsson, B. L.; Raines, R. T. *J. Org. Chem.* **2002**, *67*, 4993-4996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

residue at the ligation juncture. Cysteine is uncommon, comprising only 1.7% of all residues in proteins (McCaldon & Argos, 1988b). Modern peptide synthesis is typically limited to peptides of ≤50 residues (Tam et al., 1999; Borgia & Fields, 2000; Dawson & Kent, 2000). Hence, most proteins cannot be prepared by any method that allows for peptides to be coupled only at cysteine residues.



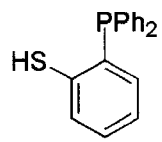
Scheme 3.1 *Native chemical ligation.*

The removal of the cysteine limitation by applying a more general ligation technology would greatly expand the utility of total protein synthesis. The Staudinger reaction provides such an alternative. In the Staudinger reaction, a phosphine is used to reduce an azide to an amine: $PR_3 + N_3R' + H_2O \rightarrow O{=}PR_3 + H_2NR' + N_2(g)$ (Staudinger & Meyer, 1919; Gololobov & Kasukhin, 1992). This reaction occurs via a stable intermediate, an iminophosphorane ($R_3P^+{-}^-NR'$), which has a nucleophilic nitrogen. Vilarrassa and others have shown that this nitrogen can be acylated, both in intermolecular and intramolecular reactions (Garcia et al., 1984; Garcia et al., 1986; Urpi & Vilarrasa, 1986; Inazu &

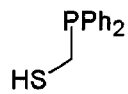
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Kobayashi, 1993; Molina & Vilaplana, 1994; Bosch et al., 1996; Shalev et al., 1996; Maunier et al., 1997; Afonso, 1998; Ariza et al., 1998; Mizuno et al., 1999a; Malkinson et al., 2000; Velasco et al., 2000). Hydrolysis of the resulting amidophosphonium salt gives an amide and phosphine oxide. Saxon and Bertozzi have shown that the phosphine can itself serve as the acyl donor (Saxon & Bertozzi, 2000a).

Recently, we reported the use of the "Staudinger ligation" to form a peptide (Nilsson et al., 2000). In this initial work, an amide bond was formed between a thioester of phosphinothiol **3.1** and an azide. The reaction likely proceeds by the intramolecular rearrangement of an iminophosphorane intermediate to give an amidophosphonium salt, as shown in Scheme 3.2. This salt hydrolyzes to leave an amide and *o*-(diphenylphosphinyl)benzenethiol. It is noteworthy that the Staudinger ligation has no reliance upon a cysteine or any other specific residue at the *N*-terminus of the peptide fragment. In addition, the method is traceless—no residual atoms from the phosphinothiol remain in the peptide product.



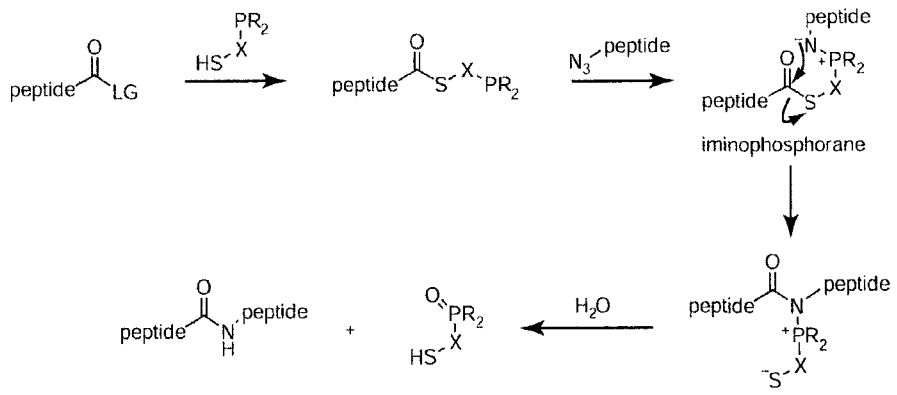
3.1



3.2

The model amides synthesized with this original phosphinothiol are listed in Table 1 (Nilsson et al., 2000). Although phosphinothiol **3.2** does enable amide bond formation, the isolated yields for these Staudinger ligations are too low for some applications.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



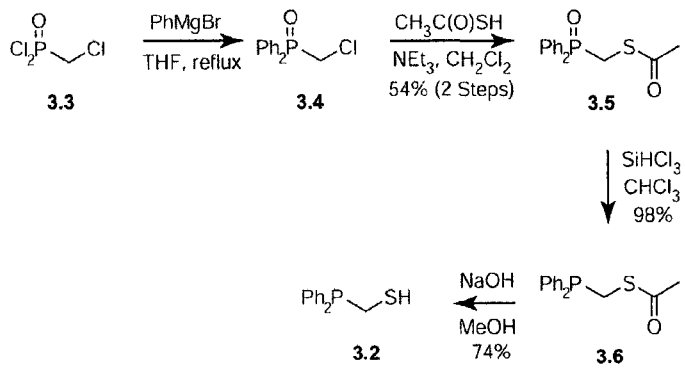
Scheme 3.2 *Traceless Staudinger ligation to form a native amide bond.*

Staudinger ligation with phosphinothiol **3.1** occurs through a transition state with a six-membered ring. We reasoned that reducing the size of this ring would bring the nucleophilic imide nitrogen more proximal to the electrophilic thioester carbon and improve the yields for the Staudinger ligation products. To access a transition state with a smaller ring, we replaced the *o*-phenyl group of phosphinothiol **3.1** with a single methylene group. We retained the other two electron-withdrawing phenyl groups, which make the phosphorus less nucleophilic and thereby minimize the susceptibility of the phosphine to deleterious oxidation by O_2(g).

The synthesis of the previously unknown phosphinothiol, **3.2**, is shown in Scheme 3.3. Phenylmagnesium bromide was added to chloromethylphosphonic dichloride (**3.3**), and the resulting Grignard reaction refluxed for 12 h to give phosphine oxide **3.4**. A mixture of **3.4** with thioacetic acid and triethylamine in dry THF was heated at reflux for 12 h (Lamoureux & Whitesides, 1993; Woycechowsky et al., 1999). After purification by flash chromatography and treatment with decolorizing charcoal, thiophosphine oxide **3.5** was isolated in a 54% combined yield for the two steps. An excess of trichlorosilane in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

chloroform for 72 h was used to reduce **3.5** to phosphinothioester **3.6** (Charrier & Mathey, 1978), which was isolated by flash chromatography in nearly quantitative yield. Hydrolysis of the phosphinothioester **3.6** with sodium hydroxide in methanol for 2 h gave phosphinothiol **3.2**. During this reaction, Ar(g) was bubbled through the reaction mixture to prevent oxidation of the resultant thiol. Hydrolysis under acidic conditions was unsuccessful. Phosphinothiol **3.2** was purified by chromatography over alumina and isolated in 74% yield. The overall yield for the process in Scheme 3.3 was 39%.

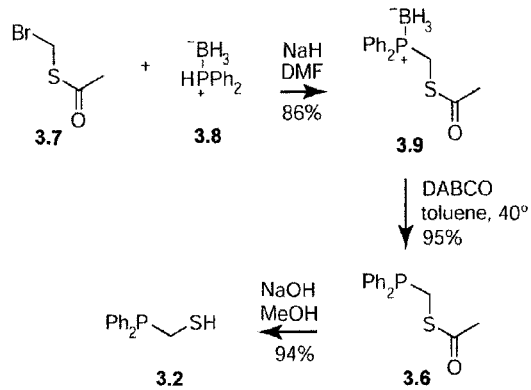


Scheme 3.3 *Synthesis of phosphinothiol* **3.2**.

The previously described synthesis of phosphinothiol **3.1** required four steps, two of which were problematic, with an overall yield of only 39%. We have developed an improved synthesis that uses air-stable borane protection of the phosphine (Scheme 3.4) (Brunel et al., 1998; Carboni & Monnier, 1999). The synthesis is based on the easily prepared alkylating agent **3.7** (Farrington et al., 1989) and the commercially available borane–diphenylphosphine complex **3.8**. Compound **3.8** is deprotonated by sodium hydride in DMF followed by addition of **3.7** to give borane-complex **3.9** (86% yield)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Imamoto et al., 1990). Complex **3.9** is stable to air and moisture, and can be stored on the shelf at room temperature for months without any sign of oxidation or decomposition. The borane complex is disrupted by mild heating with DABCO in toluene for 4 h (95% yield) (Brisset et al., 1993). The protecting group of the resulting acyl phosphinothiol **3.6** is removed as described previously except we used only one equivalent of sodium hydroxide. Phosphinothiol **3.2** was obtained in an improved 94% yield. The overall yield for this three-step synthesis is 74%.



Scheme 3.4 *Improved synthesis of phosphinothiol* **3.2**.

Next, we determined the efficacy of phosphinothiol **3.2** in effecting the Staudinger ligation (Scheme 3.4). Thioesters of **3.2** derived from AcOH, AcGlyOH, and AcPheOH were prepared by either transthioesterification or coupling with dicyclohexylcarbodiimide. After these reactions were complete by TLC analysis, Merrifield resin was used to immobilize unreacted **3.2**. After workup and chromatography, the purified thioesters were isolated in >90% yields. To effect the Staudinger ligation, each thioester was stirred with $N_3CH_2C(O)NHBn$ (Nilsson et al., 2000) (1 equiv) in THF/H_2O (3:1) at room temperature for 12 h. Solvents were removed

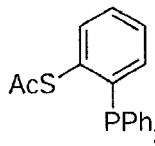
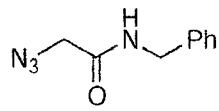
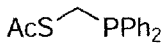
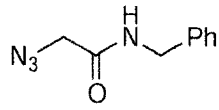
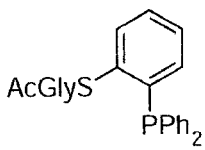
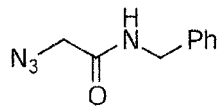
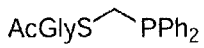
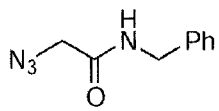
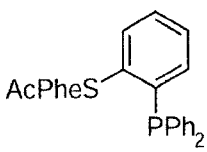
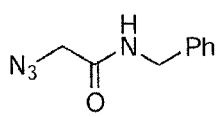
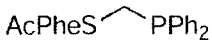
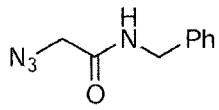
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

under reduced pressure, and the product amides were purified by chromatography. After the ligation reactions, phosphinothiol **3.2** was regenerated from its phosphine oxide by reduction with an excess of trichlorosilane in chloroform.

The yields of amide product using phosphinothiol **3.2** are far greater than those using phosphinothiol **3.1** (Table 1). AcGlyNHBn was obtained in 91% isolated yield with **3.2**, compared to a trace yield with **3.1**. AcGlyGlyNHBn was obtained in 80% yield using **2**, compared to 15% using **3.1**. AcPheGlyNHBn was obtained in 92% yield with **3.2**, compared to 35% with **3.1**. These dramatic improvements indicate that phosphinothiol **3.2** is a superior reagent for effecting the Staudinger ligation of a thioester and azide to form an amide. Bertozzi and coworkers have assessed the ability of the oxo analogs of phosphinothiols **1** and **2** to effect a Staudinger ligation (Saxon et al., 2000). Surprisingly, they found that $Ph_2PC_6H_4$-*o*-OH gives a higher yield than does Ph_2PCH_2OH. The basis for the antipodal reactivity of thioesters and esters is unclear.

What is the basis for the high yields obtained with phosphinothiol **3.2**? One contributing factor could be the proximity of the nucleophile and electrophile. The key intermediate in the Staudinger ligation is the iminophosphorane (Scheme 3.2). The transition state leading from the iminophosphorane of **3.2** to the amidophosphonium salt contains a 5-membered ring. Both the C–S and P–N bonds in this ring have significant double-bond character. Thus, the iminophosphorane can adopt relatively few conformations. In contrast, reaction of $N_3(CH_2)_{10}C(O)SPy$ and PBu_3 to form a lactam proceeds via a transition state with a 12-membered ring. The yield of this reaction is only 28% (Bosch et al., 1993).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Phosphinothioester	Azide	Peptide	Isolated Yield (%)
		AcGlyNHBn	<10
		AcGlyNHBn	91
		AcGlyGlyNHBn	15
		AcGlyGlyNHBn	80[a]
		AcPheGlyNHBn	35
		AcPheGlyNHBn	92

[a] Isolated by recrystallization. Other ligation products were isolated by silica-gel chromatography.

Table 3.1 *Yields for Staudinger ligation with phosphinothiols* **3.1** *and* **3.2**.

Another factor that could contribute to the high yields obtained with phosphinothiol **3.2** is a stable conformation that facilitates amide formation. Molecular mechanics calculations indicate that the iminophosphorane intermediate can adopt a β-turn–like conformation (Figure 3.1). A β-turn is stabilized by an O···HN hydrogen bond that defines a 10-membered ring. The thioester, imide, and amide groups of the iminophosphorane are situated in positions that correspond to the three amide groups of a β-turn. In this conformation, the nucleophilic imide nitrogen is within 3.0 Å of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

electrophilic thioester carbon. Moreover, the O···HN hydrogen bond would polarize the thioester, making its carbon even more electrophilic. Finally, the bulk of the two phenyl groups could accelerate acyl transfer by increasing the fraction of iminophosphorane in the β-turn–like conformation (Bruice & Pandit, 1960; Jung & Gervay, 1991). The favorable conformation depicted in Figure 3.1 would be inaccessible during a Staudinger ligation with **1**, as well as in the ligation of a thioester with a non-peptidyl azide.

Phosphinothiol **3.2** has an additional intrinsic advantage over phosphinothiol **3.1**. In general, aliphatic thiols (such as **3.2**) have higher pK_a values than do aromatic thiols (such as **3.1**). Because thioester hydrolysis rates correlate inversely with their thiol pK_a values (Janssen, 1969), aliphatic thioesters have a longer half-life in aqueous solution. This long half-life is important, as the hydrolysis of the thioester, either before or after iminophosphorane formation, is likely to be a competing side reaction for the Staudinger ligation.

Staudinger ligation with phosphinothiol **3.2** could enable facile protein synthesis. Peptides with a *C*-terminal thioester can be produced by Fmoc-based solid-phase synthesis (Ingenito et al., 1999; Shin et al., 1999; Swinnen & Hilvert, 2000) or rDNA technology (Holford & Muir, 1998; Muir et al., 1998; Cotton & Muir, 1999). Azido acids are readily accessible (Zaloom & Roberts, 1981) and can be used in solid-phase synthesis (Meldal et al., 1997; Tang & Pelletier, 1998). A process based on the Staudinger ligation of thioesters and azides could be a viable source of proteins for both basic research and drug discovery.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

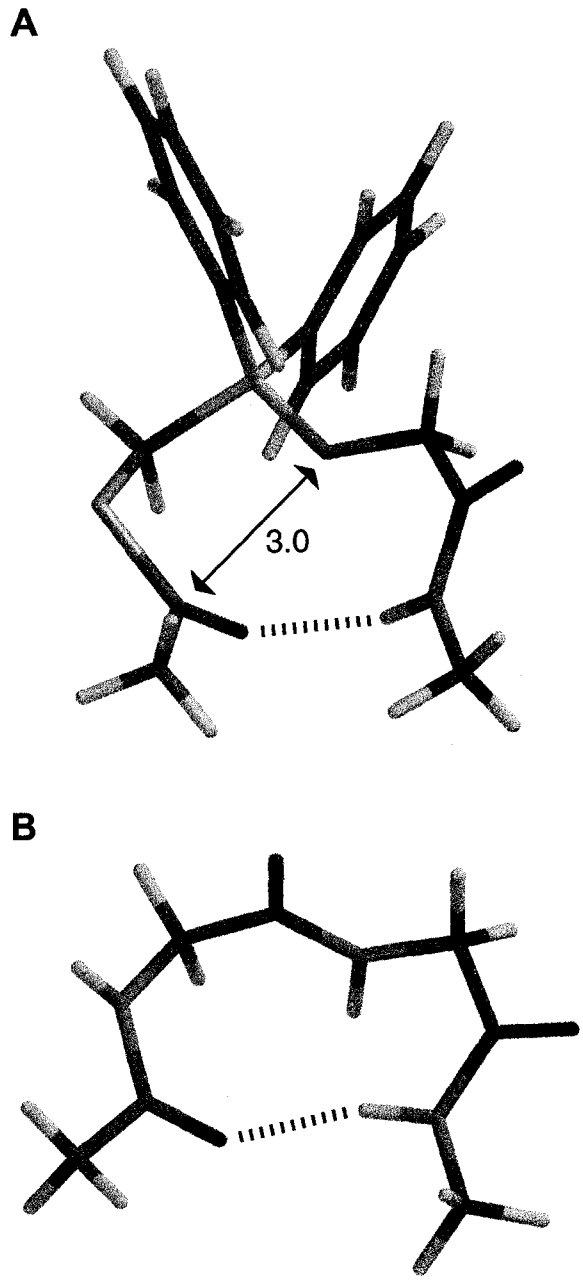


Figure 3.1 (A) Structure of the iminophosphorane formed during a Staudinger ligation ($AcSCH_2PPh_2$ + $N_3CH_2C(O)Me$) with phosphinothiol **3.2** ($r_{N\cdots C}$: 3.0 Å; ∠N···C=O: 65°). (B) Structure of an analogous β-turn (AcGlyGlyNMe). Each molecule is depicted in a conformation of minimal energy according to the MMFF94 force field in MacSpartan Pro v1.3.5 (Wavefunction, Irvine, CA)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Experimental Section

General Experimental

Chemicals and solvents were purchased from Aldrich® with the exception of Merrifield resins (Novabiochem®). Merrifield resins used were 200-400 mesh (substitution 0.63 mmol/g) and 70-90 mesh (1.26 mmol/g). Reactions were monitored by thin layer chromatography using Whatman® TLC plates (AL SIL G/UV) and visualized by UV or I_2. NMR spectra were obtained using Bruker AC-300 or Varian UNITY-500 spectrometers. Phosphorus-31 NMR spectra were proton-decoupled and referenced against an external standard of deuterated phosphoric acid. Mass spectra were obtained using electrospray ionization (ESI) techniques by the University of Wisconsin Biotechnology Center at the University of Wisconsin-Madison.

Diphenyl-phosphinoylmethyl chloride (3.4)

Chloromethylphosphonic dichloride (20 g, 120 mmol) was dissolved in freshly distilled THF (240 mL). A 1 M solution of phenylmagnesium bromide in THF (240 mL, 240 mmol) was added dropwise over 1 h. The resultant mixture was stirred at reflux for 24 h. The reaction was then quenched by the addition of 20 mL of water and solvent was removed under reduced pressure. The residue was taken up in CH_2Cl_2 and washed once with water (50 mL) and once with brine (50 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered, and solvent removed under reduced pressure. The residue was purified by chromatography (silica gel, 3% methanol in methylene chloride). Compound **3.4** was isolated as a white solid in 63% yield. **Spectral data.** ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.84-7.79 (m, 4 H), 7.62-7.58 (m, 2 H), 7.54-7.50 (m, 4 H), 4.05 (d,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

J = 7 Hz, 2 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 132.60, 131.51 (d, J = 9.6 Hz), 129.64 (d, J = 103.9 Hz), 128.72 (d, J = 11.6 Hz), 37.64 (d, J = 71.9 Hz) ppm; ^{31}P NMR ($CDCl_3$, 202 Hz) δ28.46 ppm; MS (ESI) *m/z* 250.03 (MH^+ = 251.0, M_2H^+ = 501.2 fragments at 173.0, 143.0, 91.0).

Thioacetic acid S-(diphenyl-phosphinoylmethyl) ester (3.5)

Compound **3.4** (18.94 g, 75.6 mmol) was dissolved in 450 mL THF. Thiolacetic acid (34.3 mL, 480 mmol) was added and the solution was cooled in an ice bath and argon was bubbled through the reaction mixture for 1 h. Diisopropylethyl amine (83.6 mL, 480 mmol) was added dropwise and the mixture was refluxed for 24 h. Another aliquot of thiolacetic acid (35.2 mL, 492 mmol) was then added followed by triethyl amine (69.0 mL, 492 mmol). The reaction was refluxed for another 24 h after which solvent was removed under reduced pressure in a well-ventilated hood (stench!). The resultant black oil was taken up in methylene chloride and washed once each with 2 N HCl, saturated sodium bicarbonate solution, and brine. The organic layer was dried over anhydrous $MgSO_4$ and filtered. Activated charcoal was added to this solution and it was heated at reflux for 30 minutes and filtered. The solvent was evaporated and the residue was purified by chromatography (silica gel, 70% ethyl acetate in hexanes). The pooled fractions were taken up in methylene chloride and treatment with activated charcoal was repeated. Upon evaporation of solvent compound **3.5** was obtained as an orange oil that solidified upon standing at room temperature. The yield for this reaction was 85%. **Spectral data.** ^1H NMR ($CDCl_3$, 500 MHz) δ 7.80-7.75 (m, 4 H), 7.56-7.52 (m, 2 H),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.49-7.46 (m, 4 H), 3.77 (d, J = 8 Hz, 2 H), 2.25 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 192.82, 132.11, 131.05 (d, J = 102 Hz), 130.86 (d, J = 9.75 Hz), 128.46 (d, J = 12.63 Hz), 29.83, 27.12 (d, J = 69.88 Hz) ppm; ^{31}P NMR ($CDCl_3$, 202 MHz) δ 29.14 ppm; MS (ESI) *m/z* 290.05 (MH^+ = 291.0, M_2H^+ = 581.2, fragments at 249.2, 171.0, 125.0).

Thioacetic acid S-[(diphenylphosphanyl)-methyl] ester (3.6).

Compound **3.5** (18.65 g, 64.2 mmol) was dissolved in anhydrous chloroform (160 mL). To this solution was added trichlorosilane (97 mL, 963 mmol) and the mixture was stirred under argon for 72 h. Solvent was removed under reduced pressure (NOTE: excess trichlorosilane in removed solvent was quenched by slow addition of saturated sodium bicarbonate solution in a well ventilated hood) and the residue was purified by chromatography (silica gel, 3% methanol in methylene chloride). Compound **3.6** was obtained in 98% yield. **Spectral data.** ^1H NMR ($CDCl_3$, 500 MHz) δ 7.43-7.40 (m, 4 H), 7.33-7.30 (m, 6 H), 3.50 (d, J = 4 Hz, 2 H), 2.23 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 194.01, 136.42 (d, J = 13.6 Hz), 132.28 (d, J = 19.4 Hz), 128.69, 128.11 (d, J = 6.8 Hz), 29.83, 25.41 (d, J = 23.4 Hz) ppm; ^{31}P NMR ($CDCl_3$, 202 MHz) δ –15.11 ppm; MS (ESI) *m/z* 274.06 (MH^+ = 275.0, fragments at 233.0, 199.2, 121.2).

(Diphenylphosphino)methanethiol (3.2)

Compound **3.6** (17.27 g, 63.0 mmol) was dissolved in anhydrous methanol and argon was bubbled through the solution for 1 h. Sodium hydroxide (5.04 g, 126 mmol) was then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

added and the reaction was stirred under argon for 2 h. Solvent was then removed under reduced pressure and the residue was taken up in methylene chloride and washed twice with 2 N HCl and once with brine. The organic layer was dried over $MgSO_4$, filtered, and solvent was evaporated. The residue was purified by chromatography (alumina, 25% ethyl acetate in hexanes) to afford **3.2** as a clear oil in 74% yield. **Spectral data.** ^{1}H NMR ($CDCl_3$, 300 MHz) δ7.41-7.38 (m, 4 H), 7.33-7.26 (m, 6 H), 3.02 (d, J = 7.8 Hz, 2 H), 1.38 (t, J = 7.5 Hz, 1 H) ppm; ^{13}C NMR ($CDCl_3$, 75 MHz) δ 132.54 (d, J = 17.1 Hz), 128.86, 128.36, 128.14, 20.60 (d, J = 21.7 Hz) ppm; ^{31}P NMR ($CDCl_3$, 121 MHz) δ-7.94 ppm; MS (ESI) *m/z* 232.05 (MH^+ = 233.0, fragments at 183.0, 155.0, 139.0, 91.2).

Borane-thioacetic acid S-[(diphenylphosphanyl)-methyl] ester complex (3.9).

Borane–diphenylphosphine complex **3.8** (10.33 g, 51.6 mmol) was dissolved in dry DMF under Ar(g) and cooled to 0 °C. NaH (1.24 g, 51.6 mmol) was added slowly, and the mixture was stirred at 0 °C until bubbling ceased. Alkylating agent **7** (8.73 g, 51.6 mmol) was then added, and the mixture was allowed to warm to room temperature and stirred for 12 h. Solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 10% v/v EtOAc in hexanes). Compound **3.9** was isolated as a colorless oil in 86% yield. **Spectral data.** ^{1}H NMR (300 MHz, $CDCl_3$) δ7.74-7.67 (m, 4 H), 7.54-7.41 (m, 6 H), 3.72 (d, J = 6 Hz, 2 H), 2.23 (s, 3 H), 1.51-0.53 (broad m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 192.94, 132.26 (d, J = 9.2 Hz), 131.61 (d, J = 2.3 Hz), 128.71 (d, J = 10.2 Hz), 127.43 (d, J = 55.4 Hz), 29.87, 23.59 (d, J = 35.5 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ19.40 (d, J = 59.3 Hz) ppm; MS (ESI)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

m/z 311.0806 (MNa^+ [$C_{15}H_{18}BOPSNa$] = 311.0807).

Thioacetic acid S-[(diphenylphosphanyl)-methyl] ester (3.6).

Compound **3.9** (4.00 g, 13.9 mmol) was dissolved in toluene (0.14 L) under Ar(g). DABCO (1.56 g, 13.9 mmol) was added, and the mixture was heated at 40 °C for 4 h. Solvent was removed under reduced pressure, and the residue was dissolved in CH_2Cl_2 and washed with both 1 N HCl and saturated brine. The organic layer was dried over $MgSO_4$(s), and the solvent was removed under reduced pressure. Compound **3.6** was isolated in 95% yield, and was used without further purification. **Spectral Data.** As reported previously.

Thioester 3.7 (entry 6 of Table 1)

Method A. Compound **3.2** (500 mg, 2.15 mmol) and N-acetylphenylalanine (446 mg, 2.15 mmol) were dissolved in 15 mL DMF under argon. DCC (489 mg, 2.37 mmol) was then added and the reaction was stirred for 12 h at room temperature. The DCU by-product was filtered off, the solvent was evaporated, and the residue was purified by chromatography (50% ethyl acetate in hexanes). Thioester **3.7** was isolated in 84% yield. **Method B.** Phosphinothiol **3.2** (500 mg, 2.2 mmol) was dissolved in 5 mL of dry THF. The solution was deoxygenated by bubbling argon for 0.5 h. To this solution was added NaH (51.6 mg, 2.2 mmol). The mixture forms a slurry to which is added 2 mL of DMF to dissolve any precipitate. The N-methylmercapoacetamide (NMA) thioester of N-acetylphenylalanine (63 mg, 0.22 mmol) was added and the reaction stirred for 8 h.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Excess thiolate of **3.2** was removed by addition of Merrifield resin (1.5 g, 1.26 mmol/g) and stirring for 6 h. Resin was filtered off and the residue was purified by chromatography as in method A. Product **3.7** was isolated in 92% yield. **Spectral data.** ^{1}H NMR ($CDCl_3$, 300 MHz) δ 7.44-7.39 (m, 4 H), 7.35-7.33 (m, 6 H), 7.26-7.21 (m, 3 H), 7.11-7.09 (m, 2 H), 6.29 (d, J = 8.4 Hz, 1 H), 4.98-4.91 (m, 1 H), 3.57-3.44 (m, 2 H), 3.09 (dd, J = 14.1, 5.4 Hz, 1 H), 2.93 (dd, J = 14.1, 7.5 Hz, 1 H) 1.88 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 198.91, 169.86, 135.50, 132.62 (d, J = 19.4 Hz), 129.11 (d, J = 9.8 Hz) 128.79 (d, J = 35.9 Hz), 128.50, 128.45, 126.99, 59.56, 37.99, 25.61 (d, J = 24.4 Hz), 22.88 ppm; ^{31}P NMR ($CDCl_3$, 121 MHz) δ –14.66 ppm.

Thioester 3.8 (entry 4 of Table 1)

Phosphinothiol **3.2** (100 mg, 0.43 mmol) and N-acetylglycine (55 mg, 0.47 mmol) were dissolved in 3 mL of DMF under argon. DCC (98 mg, 0.47 mmol) was added and the reaction was stirred for 12 h at room temperature. The DCU by-product was filtered off, the solvent was evaporated, and the residue was purified by chromatography (50% ethyl acetate in hexanes). Thioester **3.8** was isolated in 67% yield. **Method B.** Phosphinothiol **3.2** (500 mg, 2.2 mmol) was dissolved in 5 mL of dry THF. The solution was deoxygenated by bubbling argon for 0.5 h. To this solution was added NaH (51.6 mg, 2.2 mmol). The mixture forms a slurry to which is added 2 mL of DMF to dissolve any precipitate. The N-methylmercapoacetamide (NMA) thioester of N-acetylglycine (44 mg, 0.22 mmol) was added and the reaction stirred for 8 h. Excess thiolate of **3.2** was removed by addition of Merrifield resin (1.5 g, 1.26 mmol/g) and stirring for 6 h. Resin

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was filtered off and the residue was purified by chromatography as in method A. Product **3.8** was isolated in 91% yield. **Spectral data.** ^{1}H NMR ($CDCl_3$, 300 MHz) δ 7.46-7.39 (m, 4 H), 7.38-7.36 (m, 6 H), 6.44 (bs, 1 H), 4.15 (d, J = 5.7 Hz, 2 H), 3.53 (d, J = 3.6 Hz, 2 H), 2.02 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$, 75 MHz) δ 196.13, 170.29, 136.45 (d, J = 13.6 Hz), 132.62 (d, J = 19.1 Hz), 129.17, 128.54 (d, J = 6.7 Hz), 48.98, 25.29 (d, J = 24.2 Hz), 22.84 ppm; ^{31}P NMR ($CDCl_3$, 121 MHz) δ -15.20 ppm; MS (ESI) *m/z* 331.08 (MH^+ = 332.2, MK^+ = 370.0).

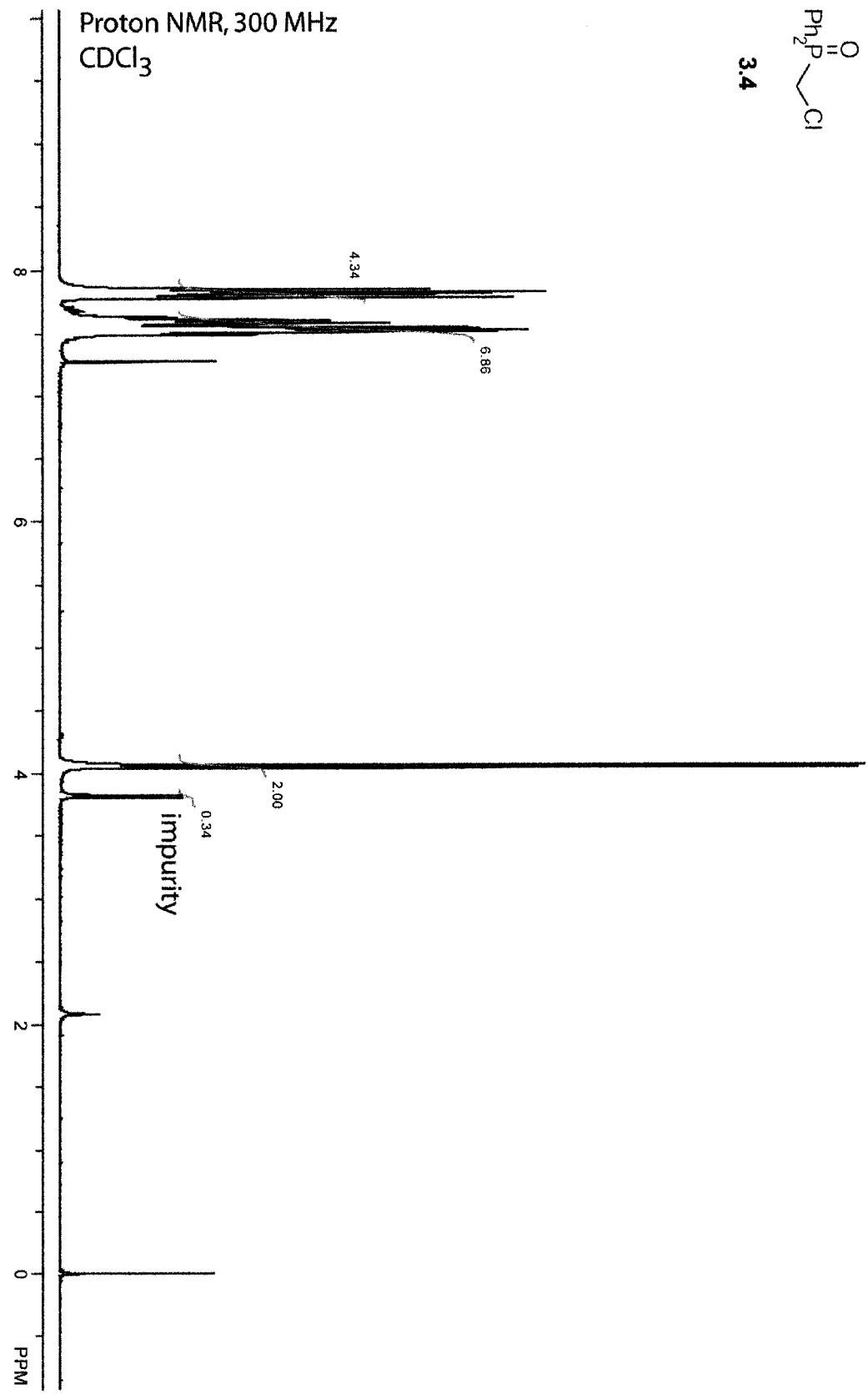
Amide 3.10 (AcGlyNHBn)

Compound **3.5** (271 mg, 0.99 mmol) and azide N_3GlyNHBn (187 mg, 0.99 mmol) were dissolved in 9.4 mL of THF:H_2O (3:1) and stirred at room temperature for 12 h. The solvent was evaporated and the residue was purified by chromatography (5% methanol in methylene chloride). Amide **3.10** was obtained in 91% yield. **Spectral data.** ^{1}H NMR ($CDCl_3$:CD_3OD, 1:1, 300 MHz) δ 7.32-7.24 (m, 5 H), 4.40 (s, 2 H), 3.88 (s, 2 H), 2.65 (s, 3 H) ppm; ^{13}C NMR ($CDCl_3$:CD_3OD, 1:1, 125 MHz) δ 171.76, 169.37, 137.49, 127.83, 126.77, 126.59, 42.50, 42.09, 21.32 ppm; MS (ESI) *m/z* 206.11 (MH^+ = 207.0).

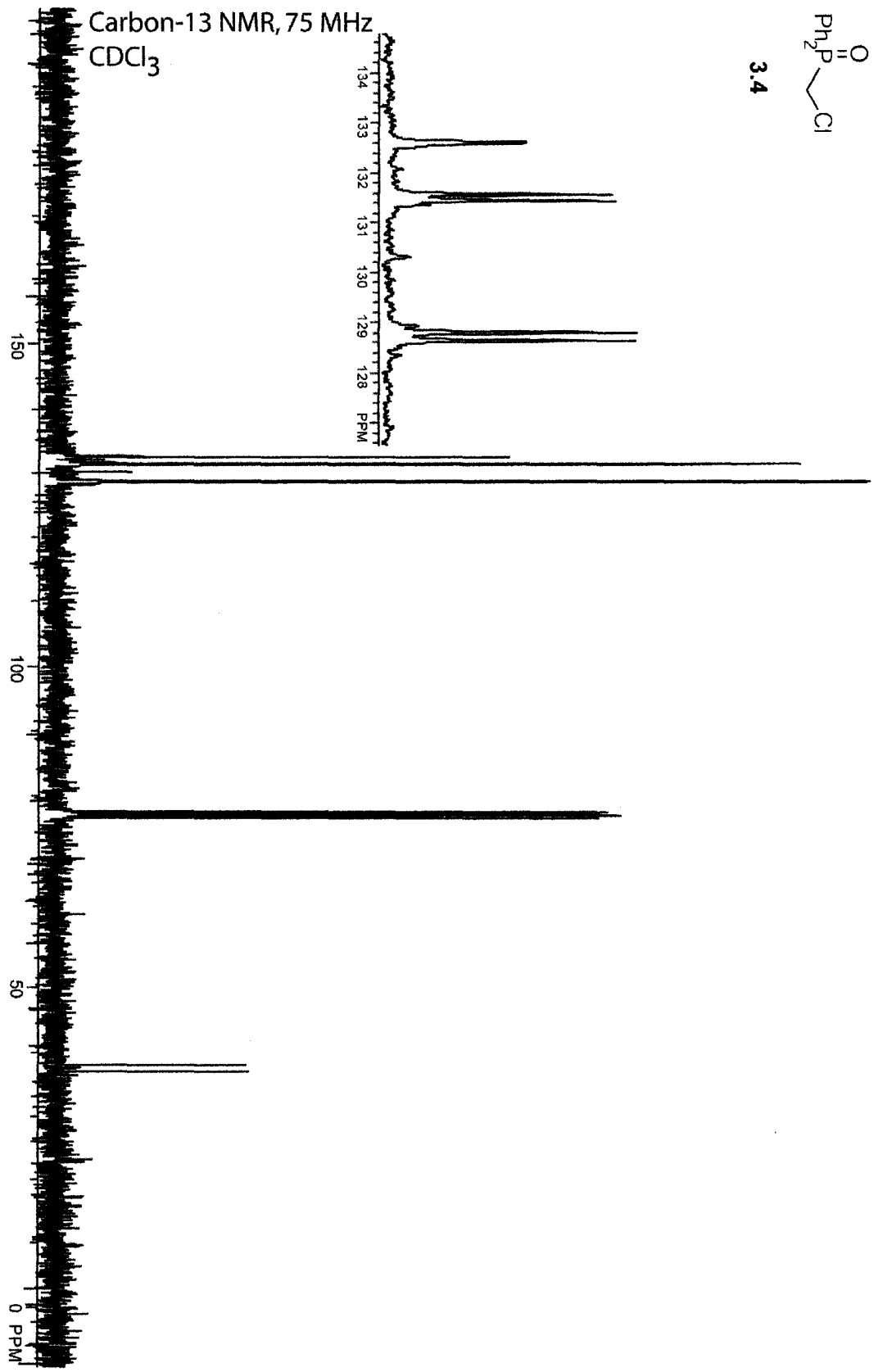
Others

Experimental and spectral information for the other amide products and for the NMA thioesters mentioned above can be found in the Supporting Information of Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939-1941.

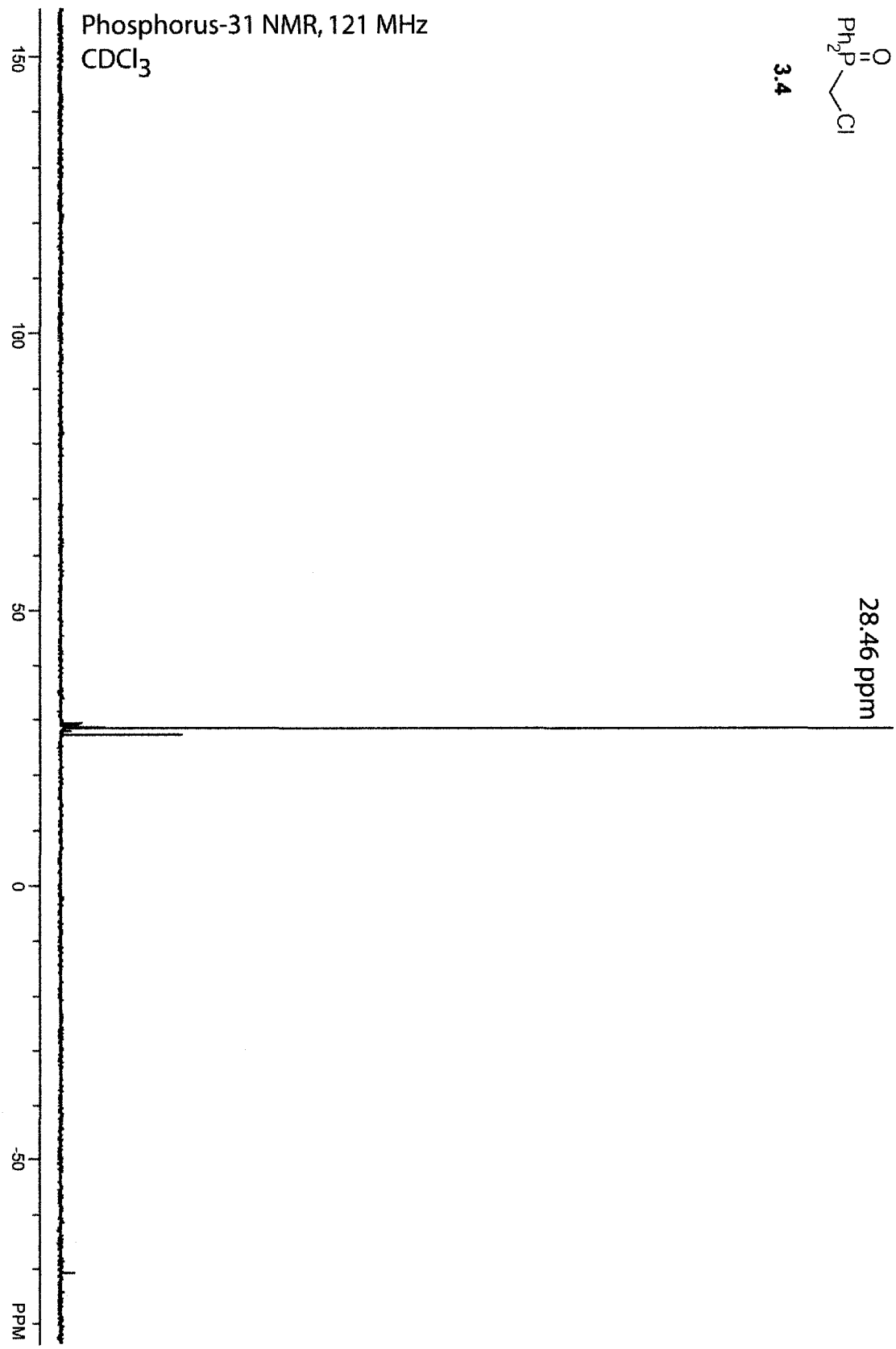
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



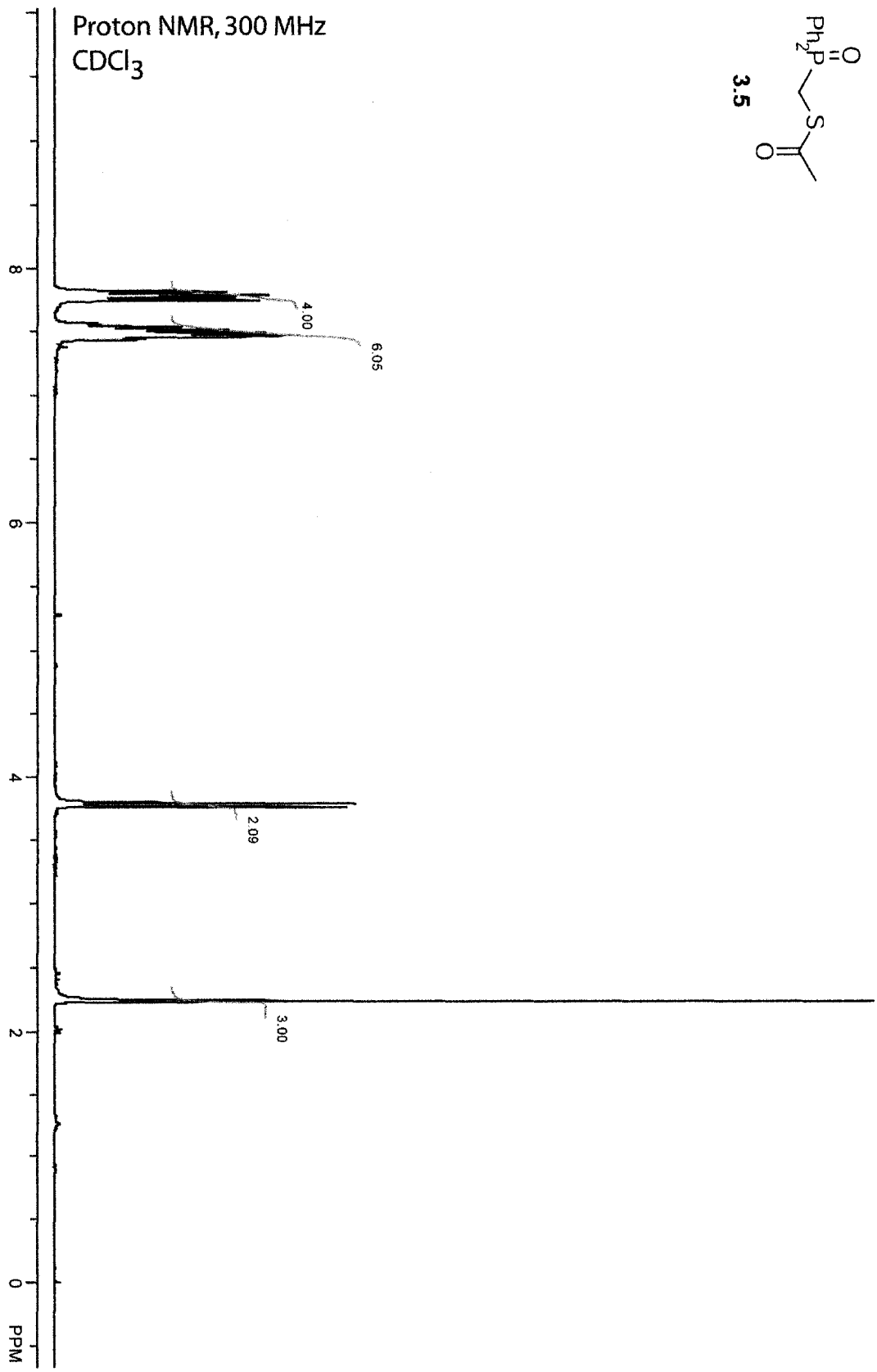
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



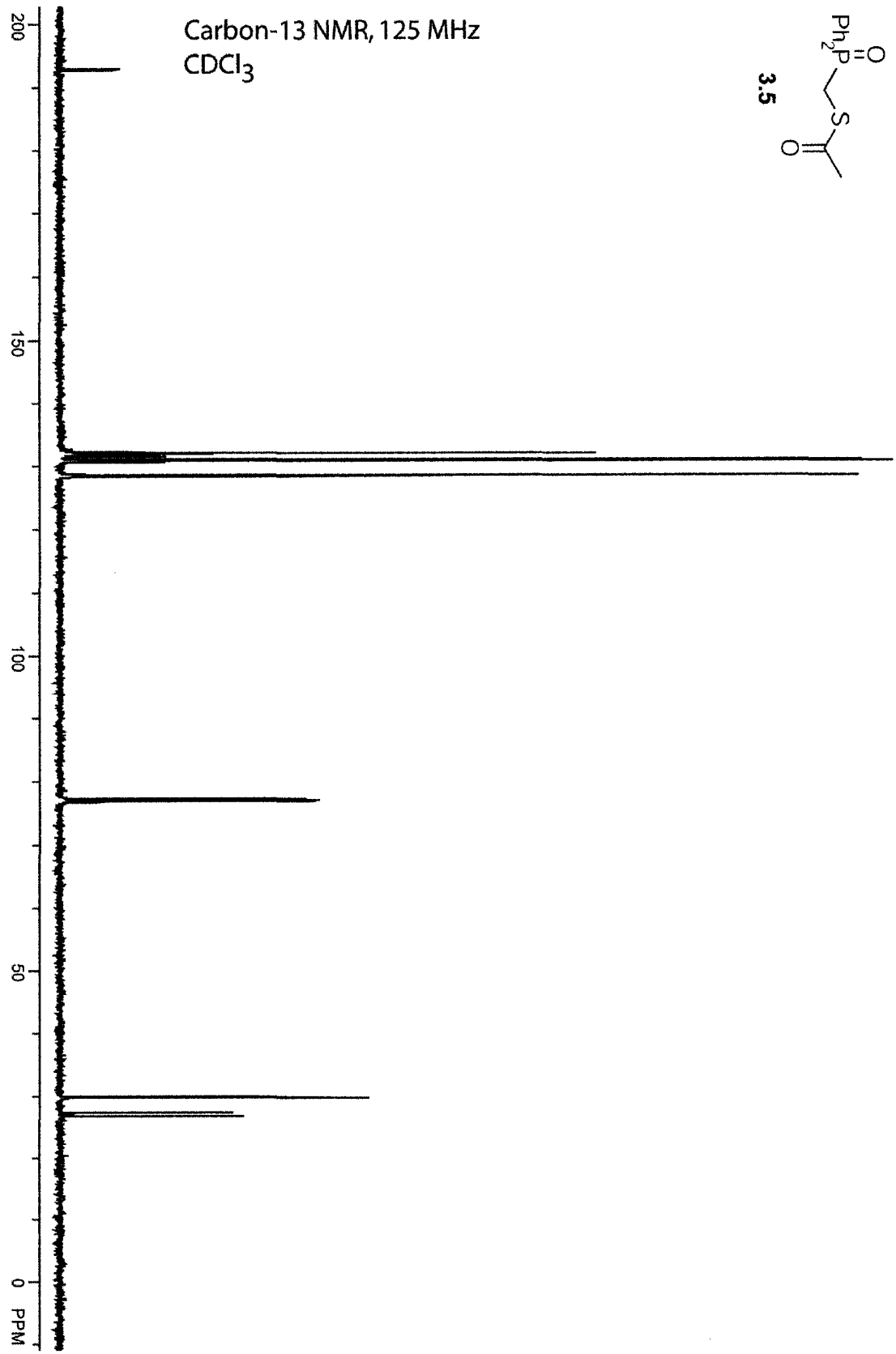
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



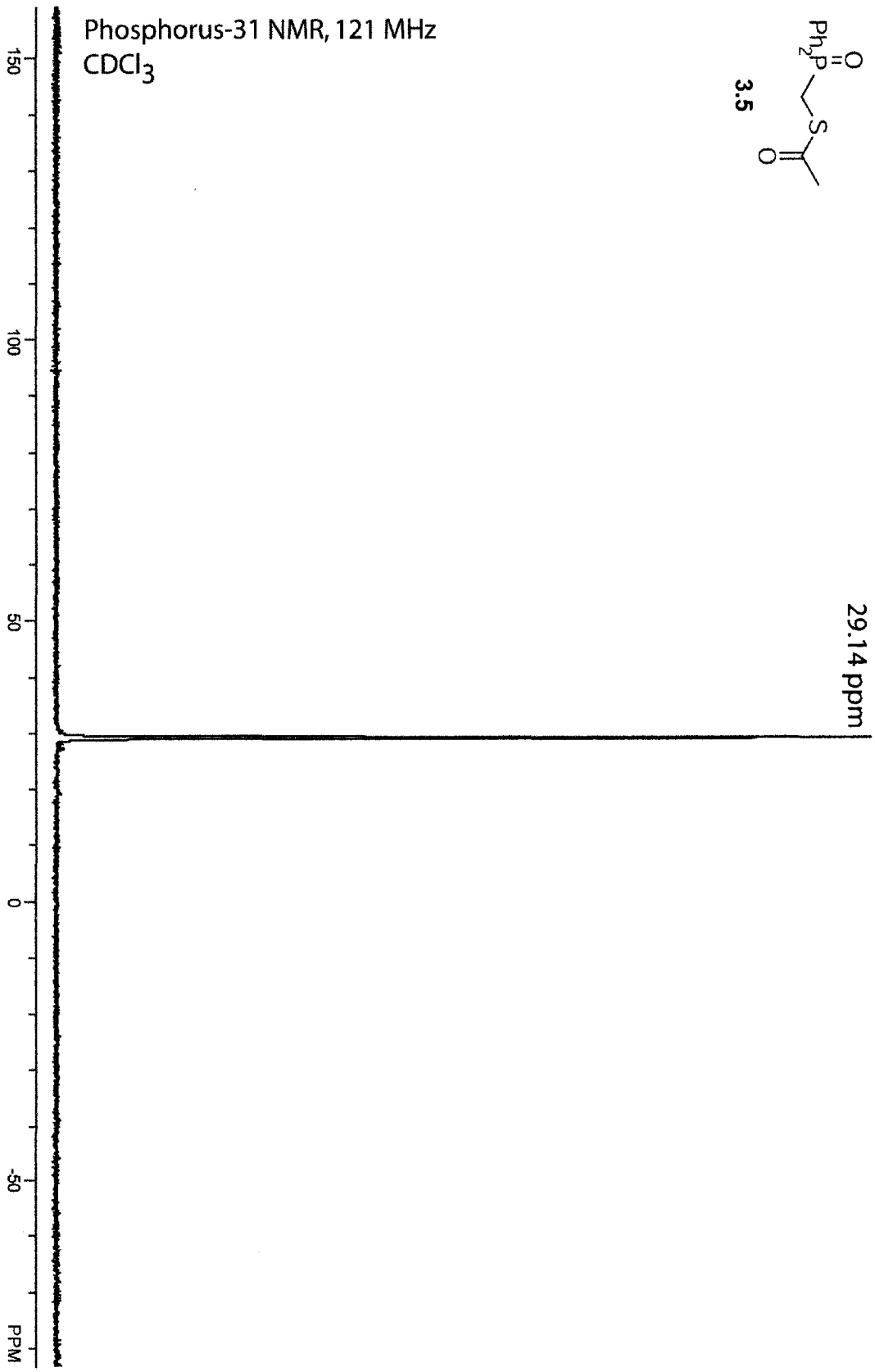
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



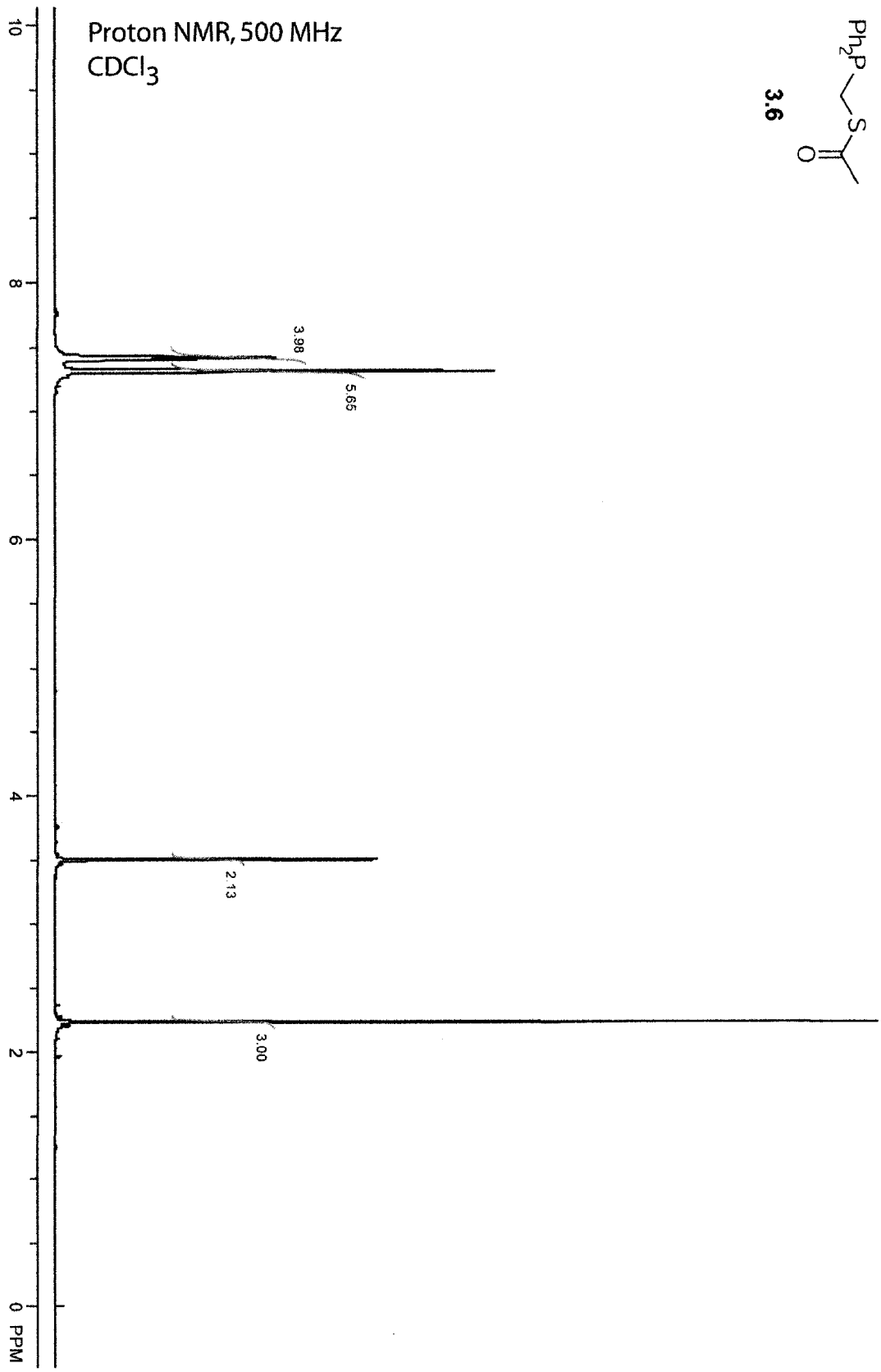
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



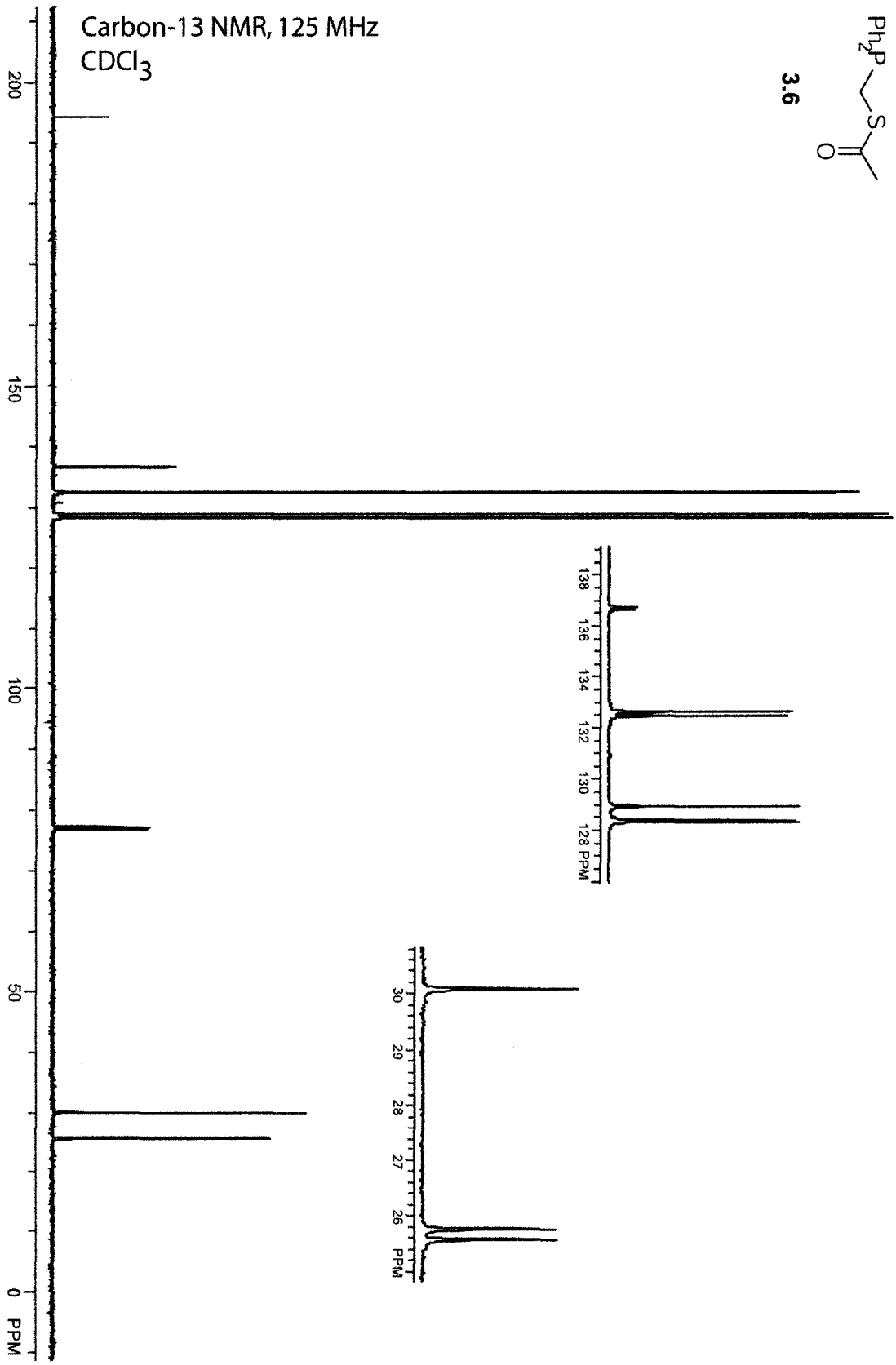
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

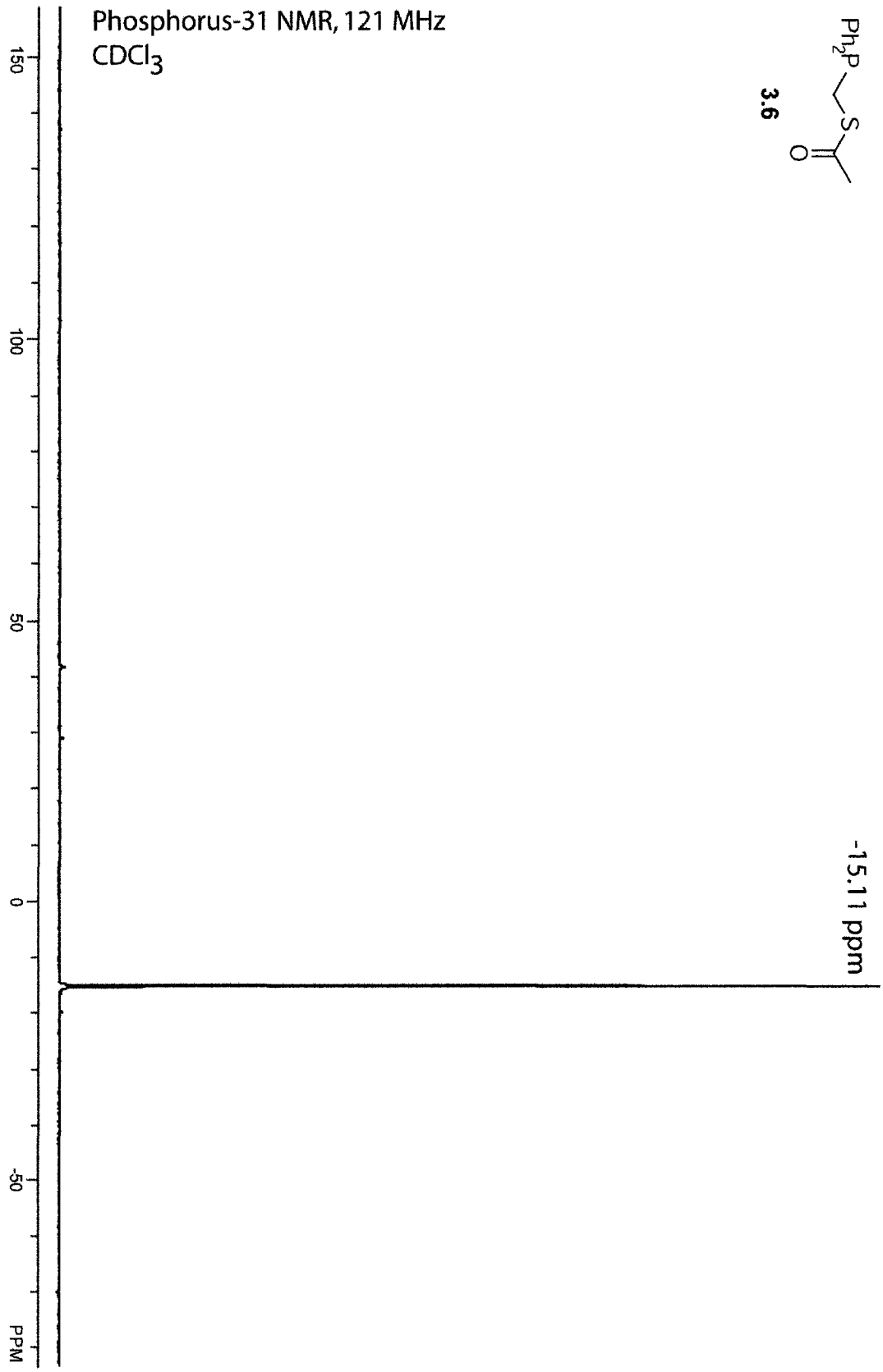


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

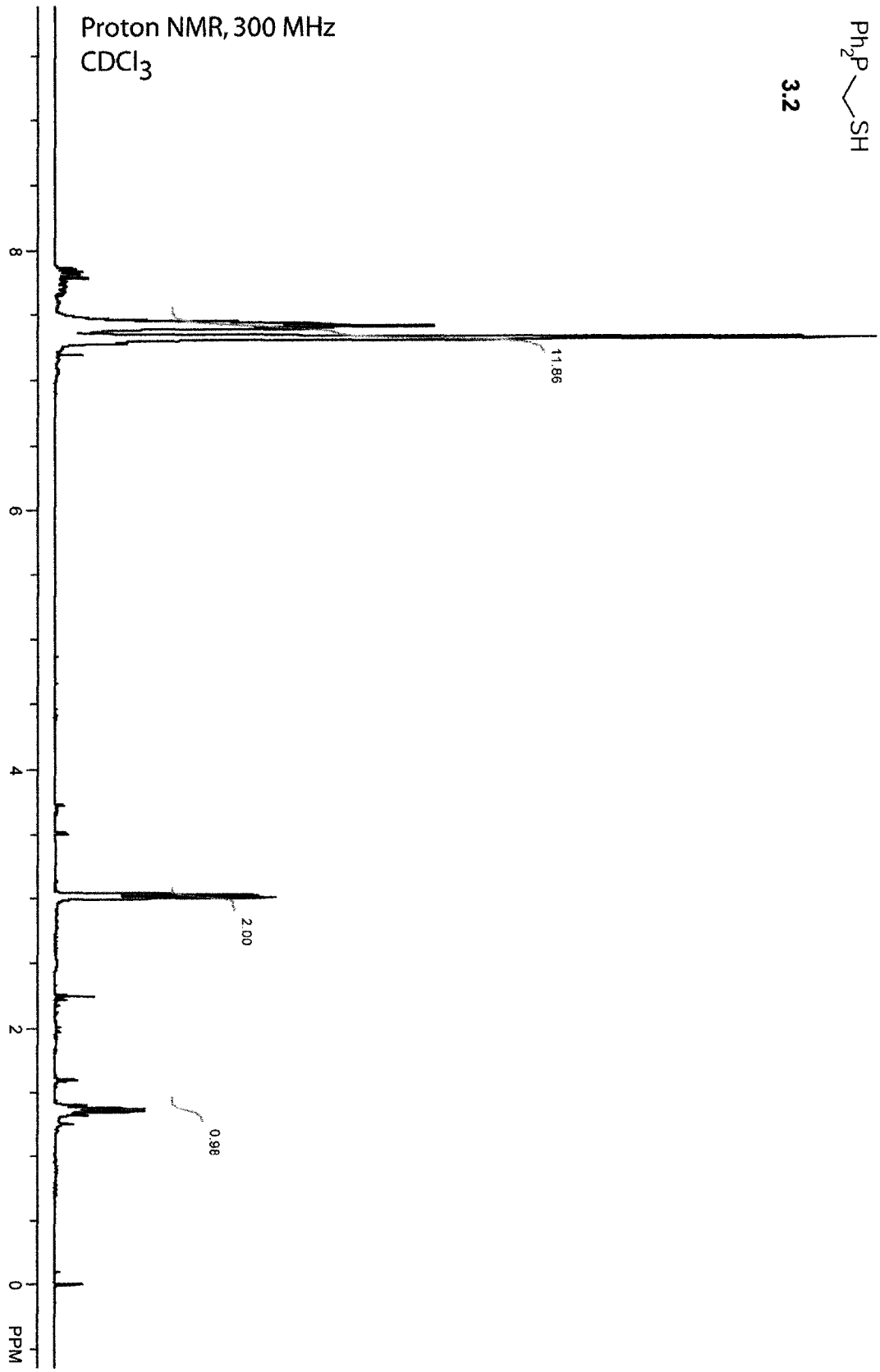


3.6

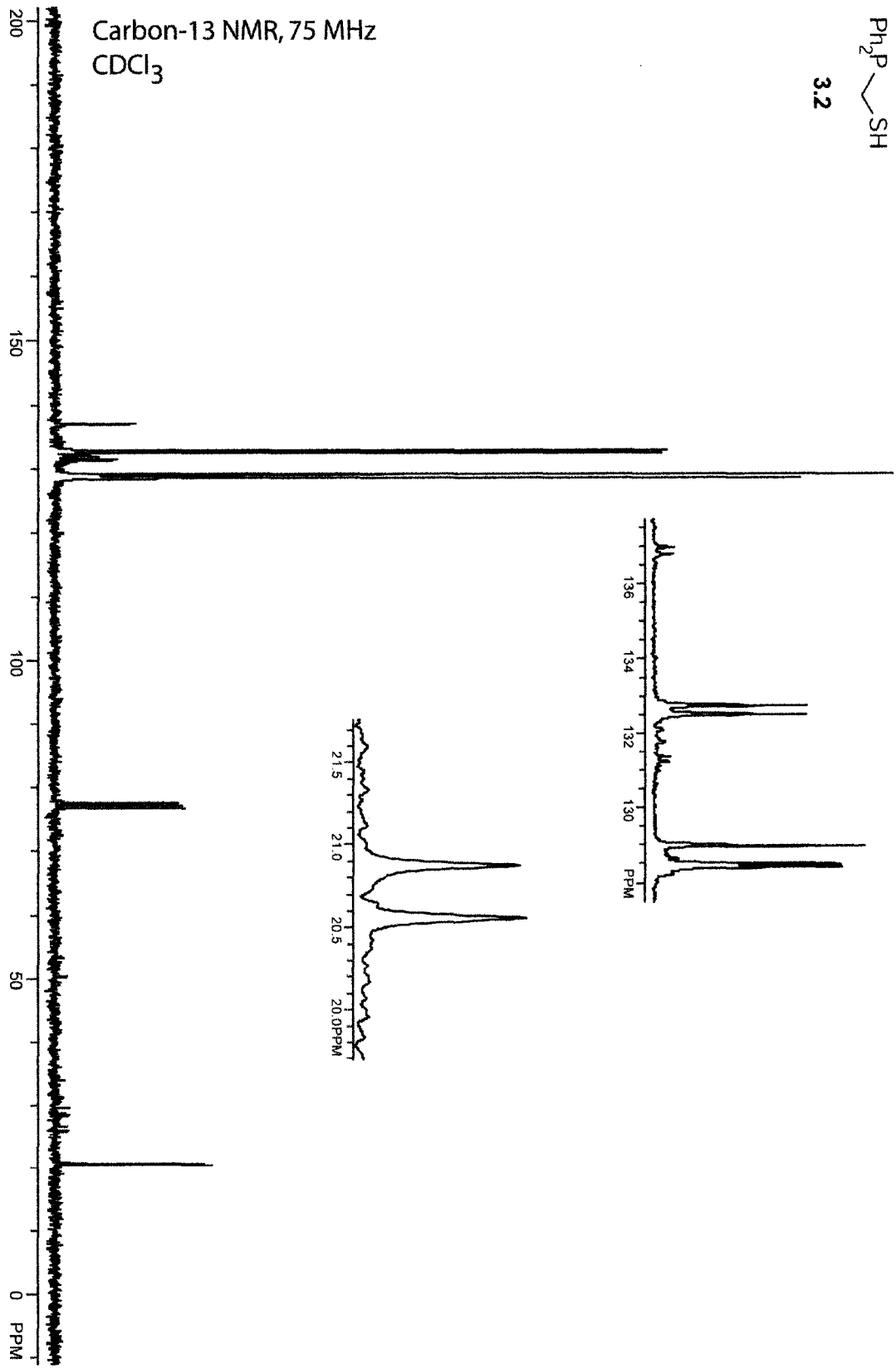
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



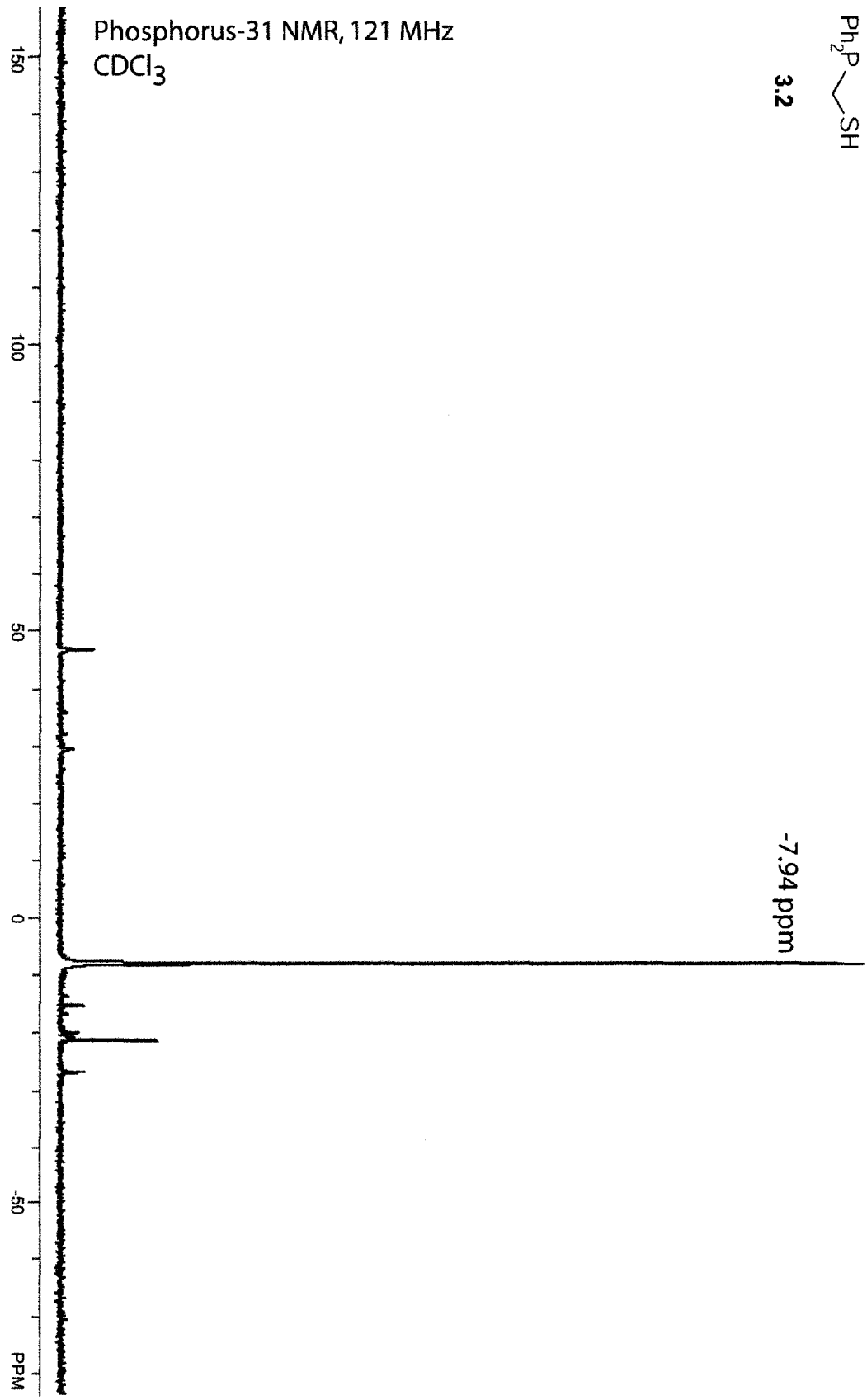
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



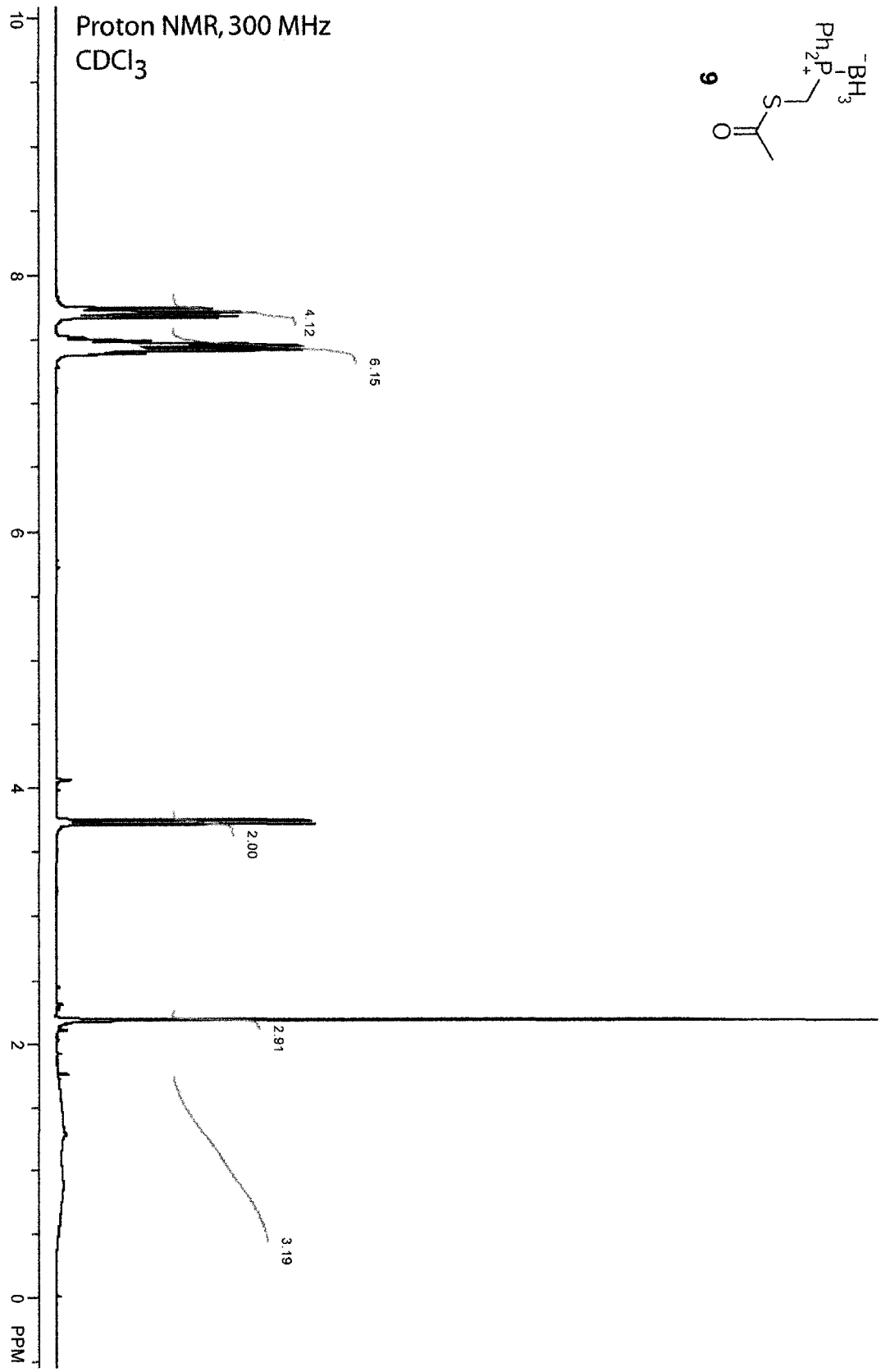
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



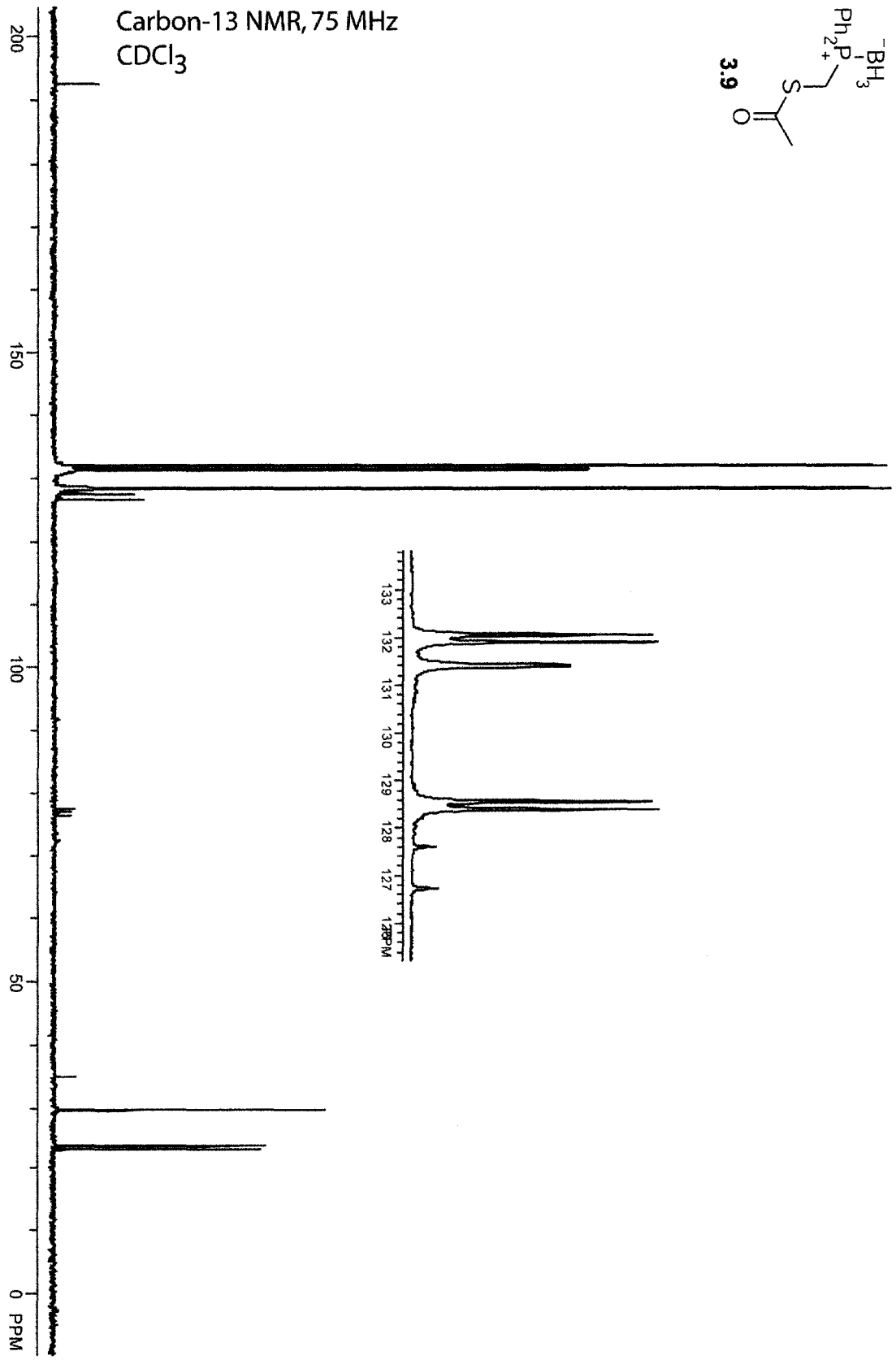
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



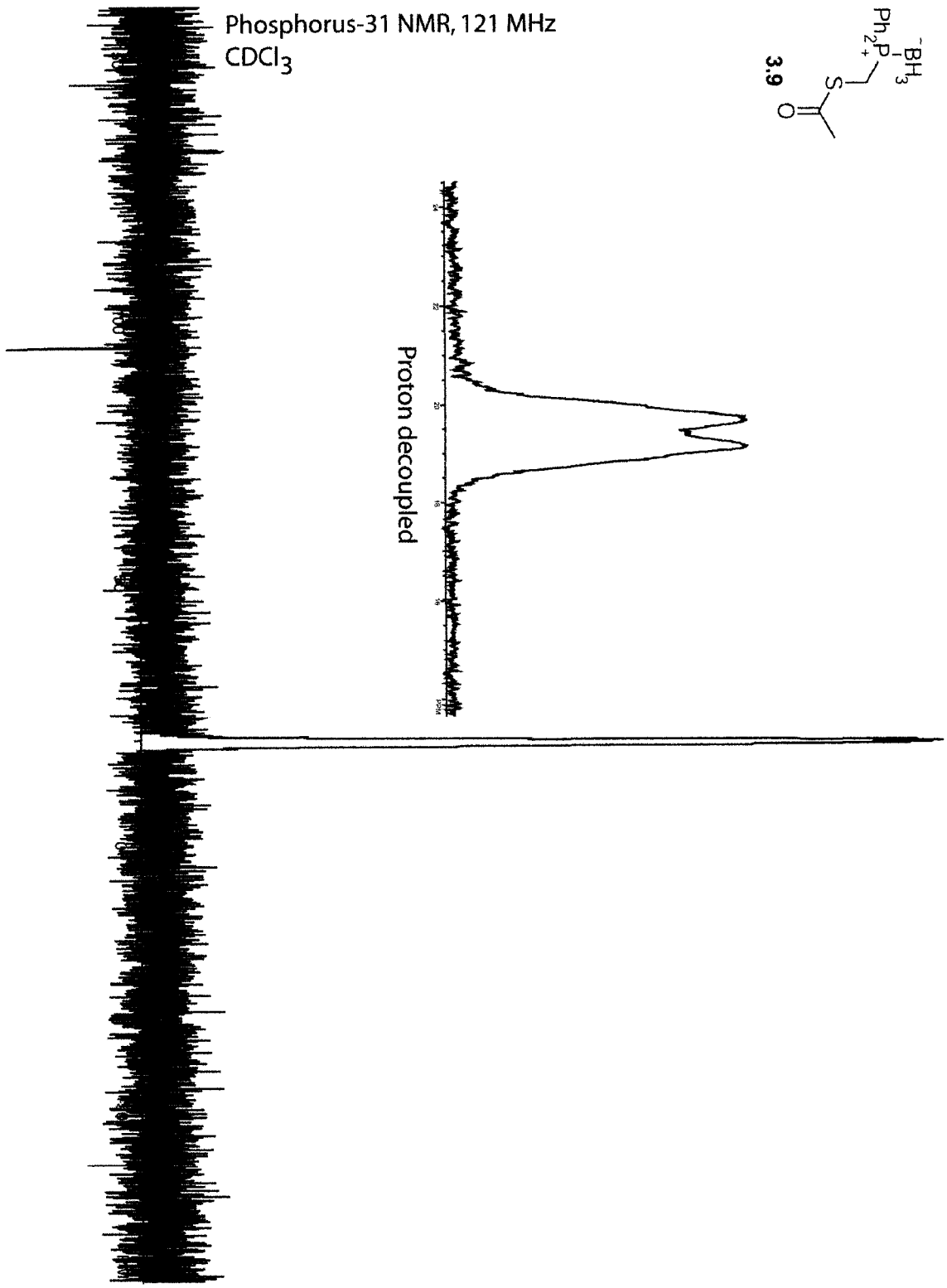
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



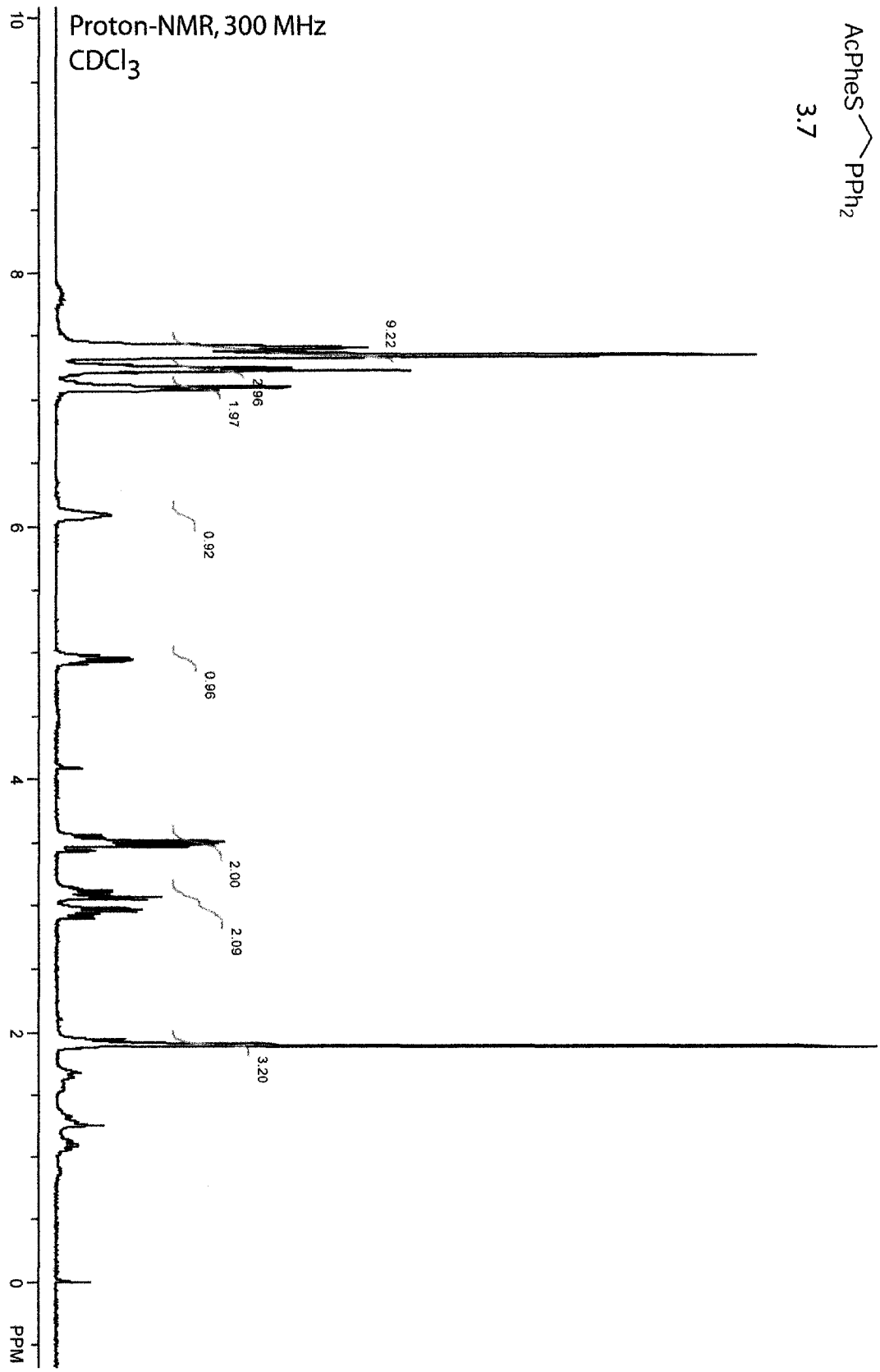
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



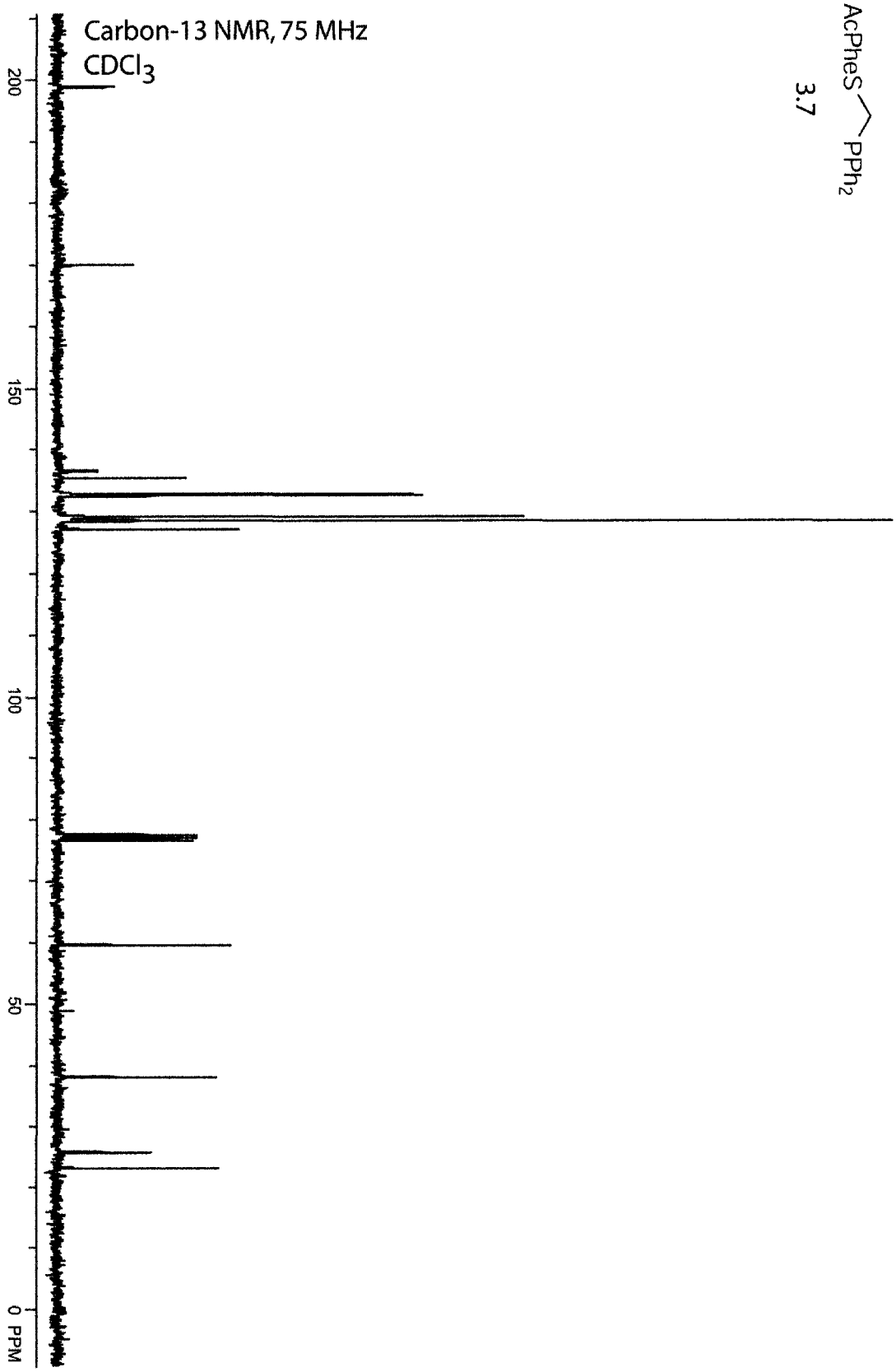
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



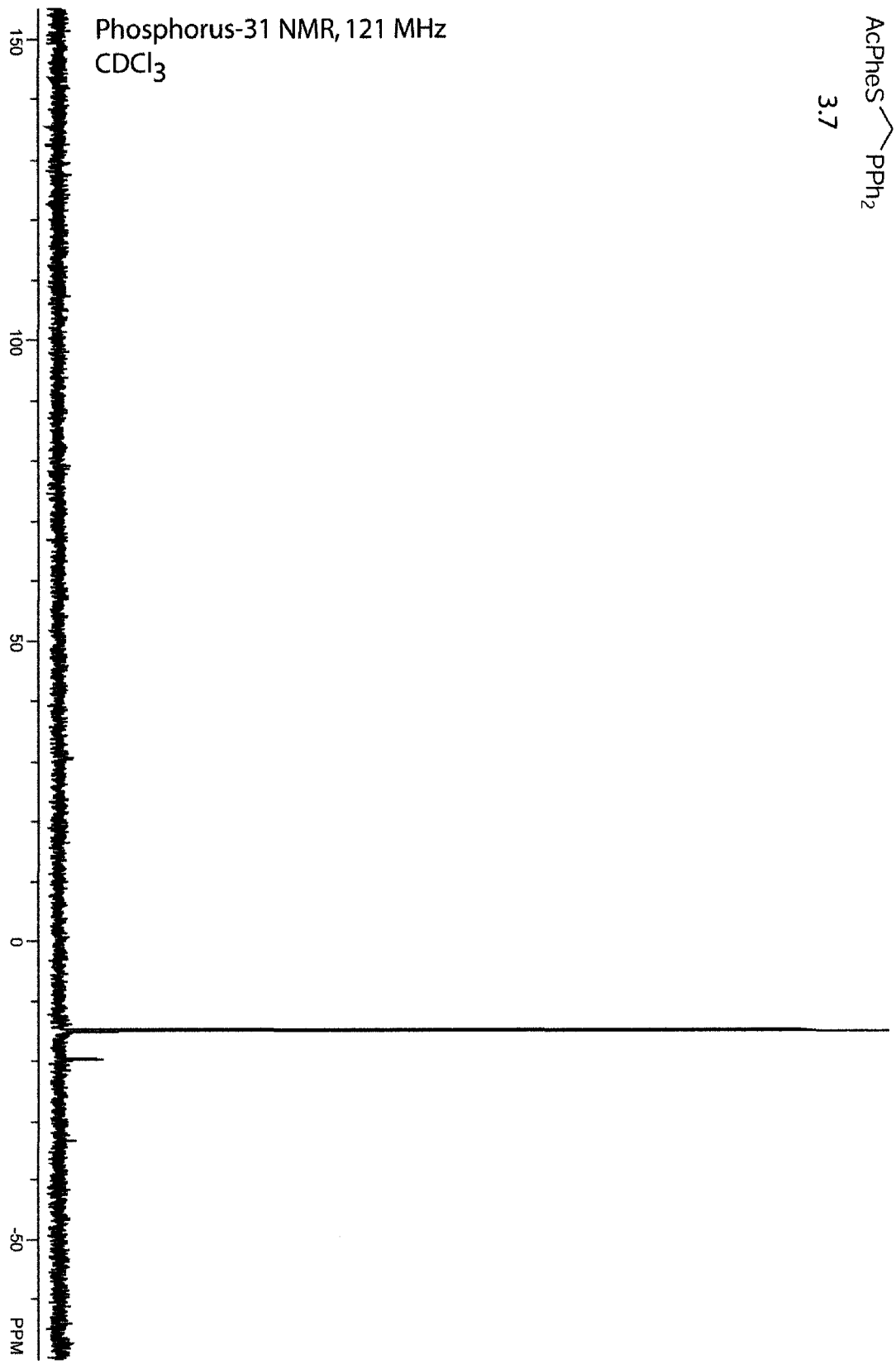
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



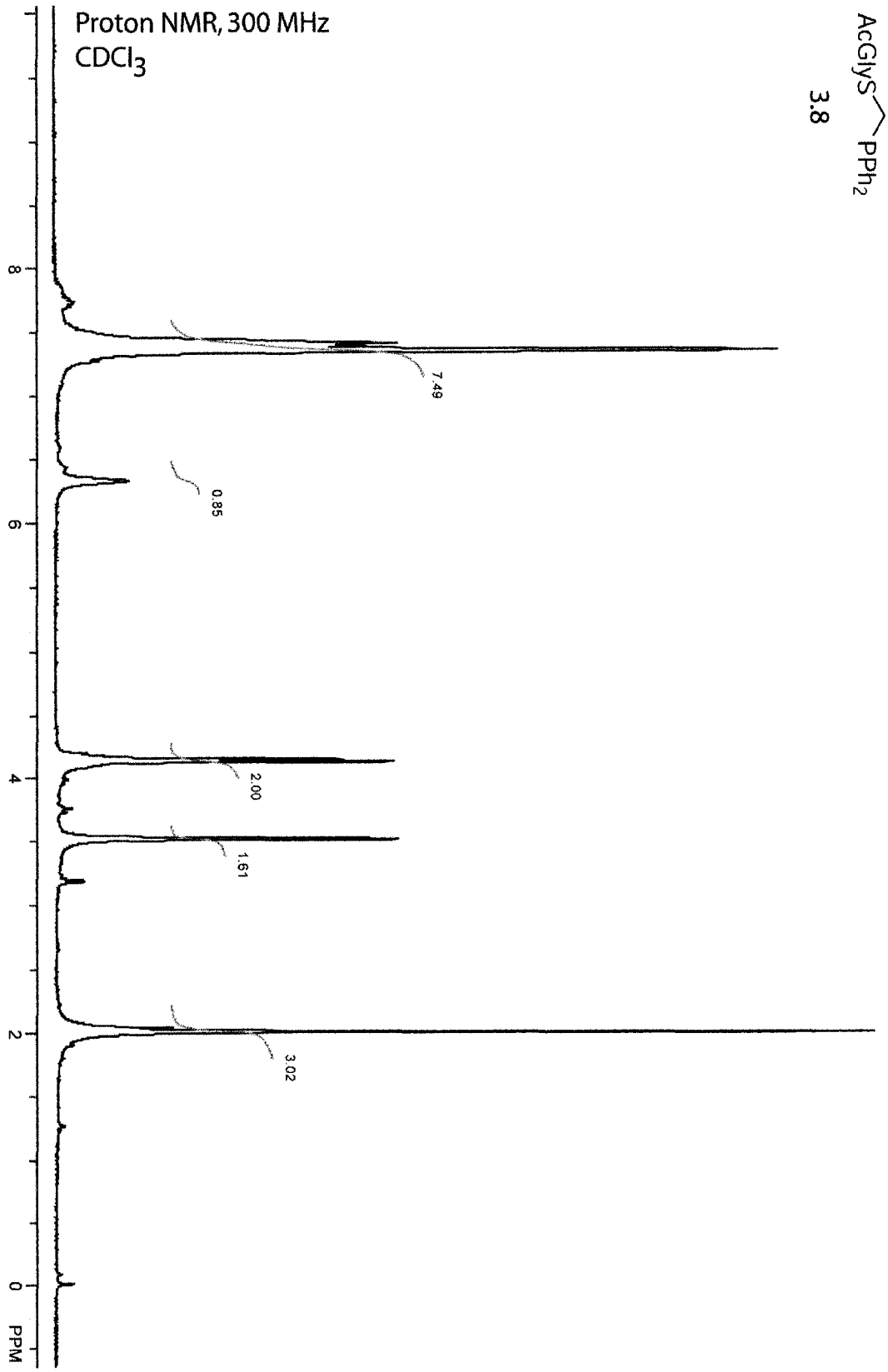
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



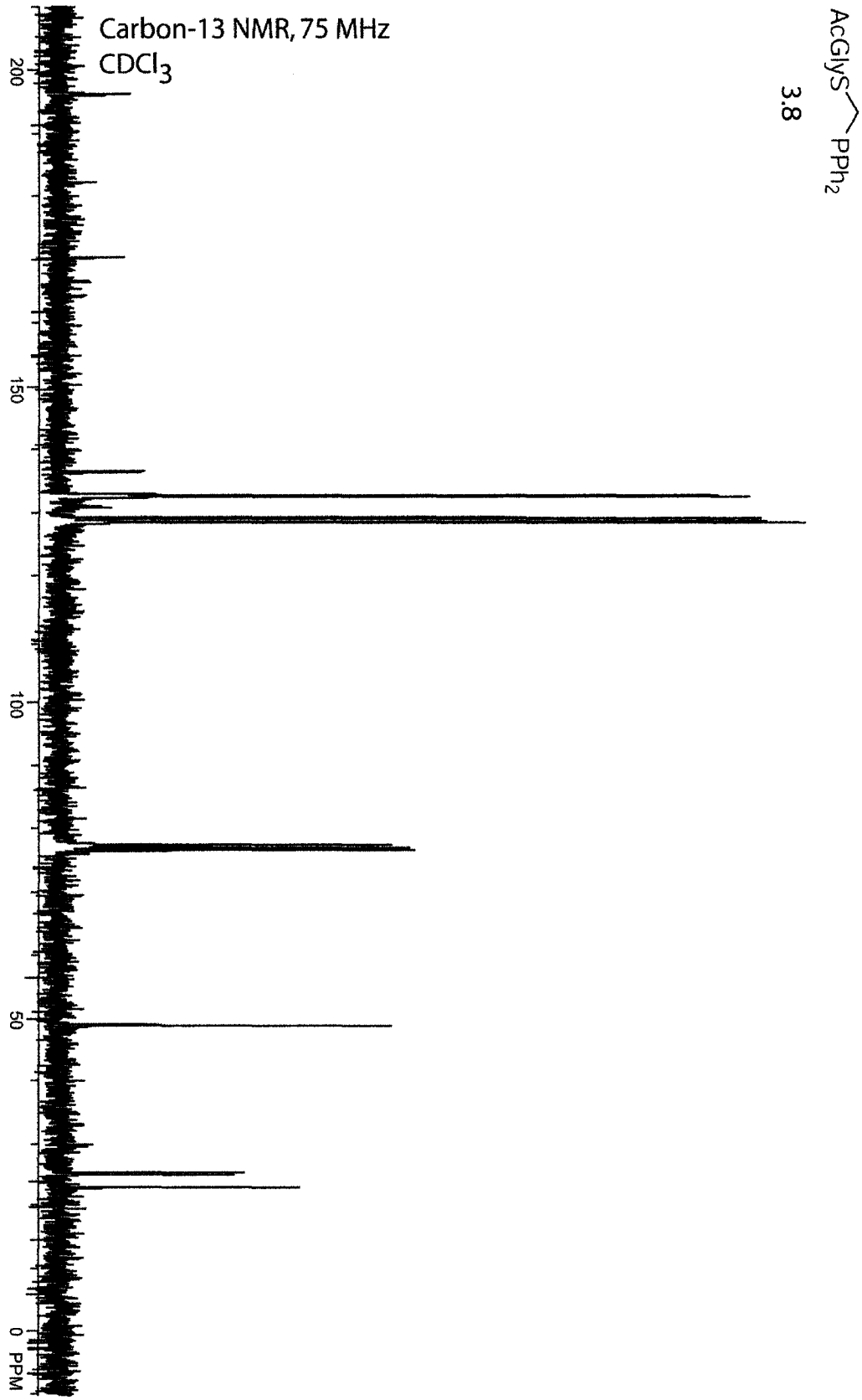
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



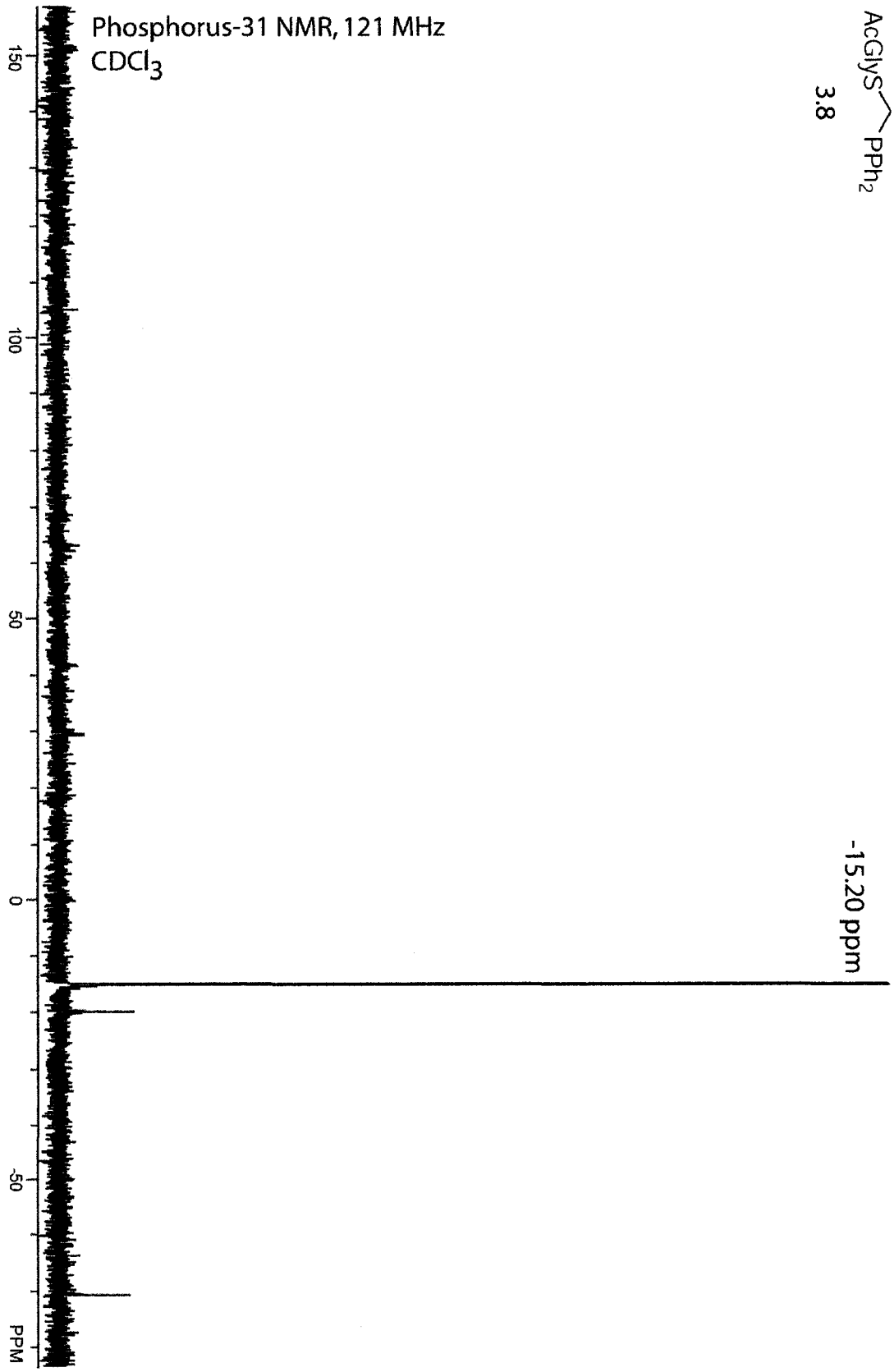
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



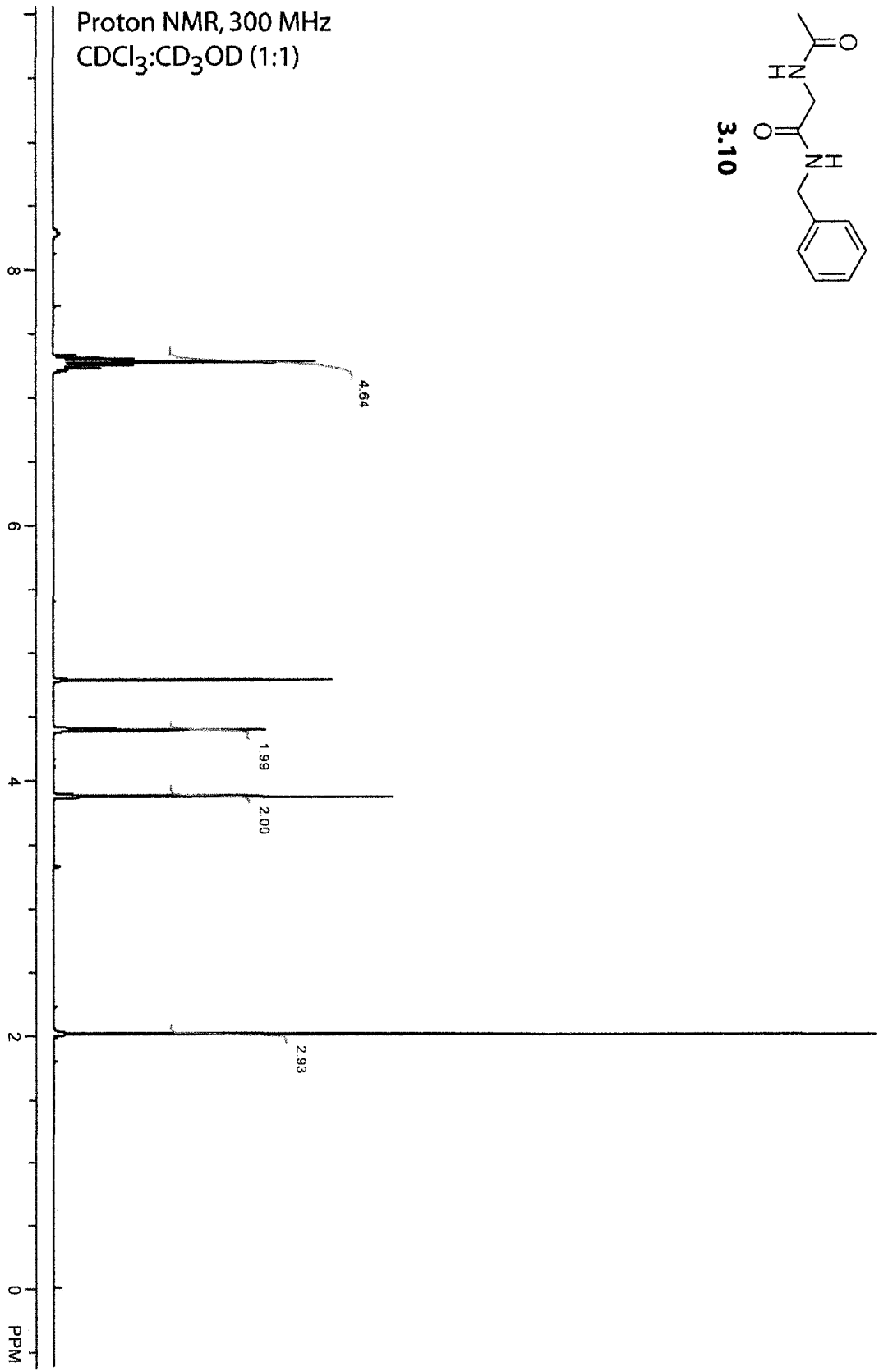
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



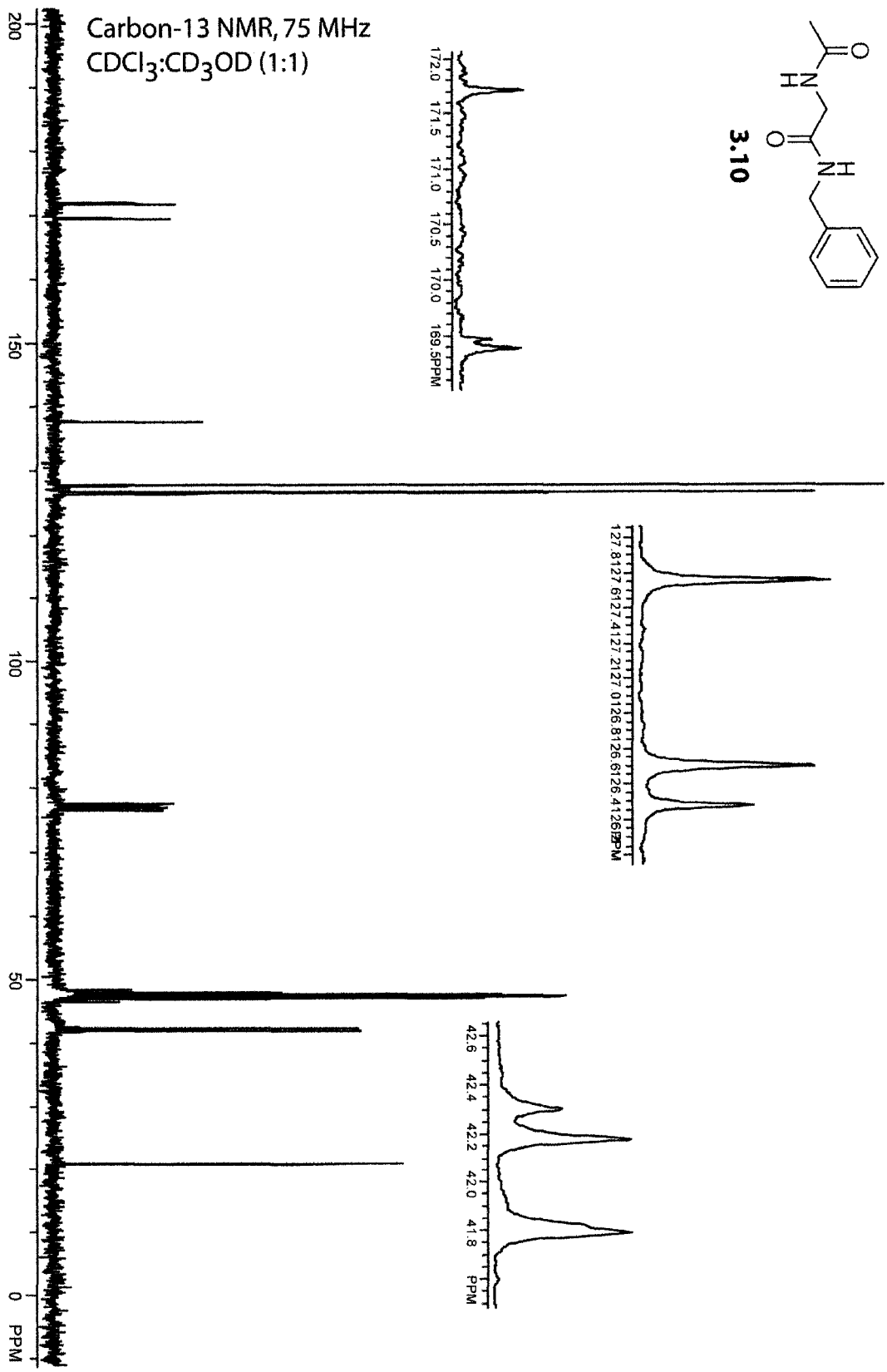
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

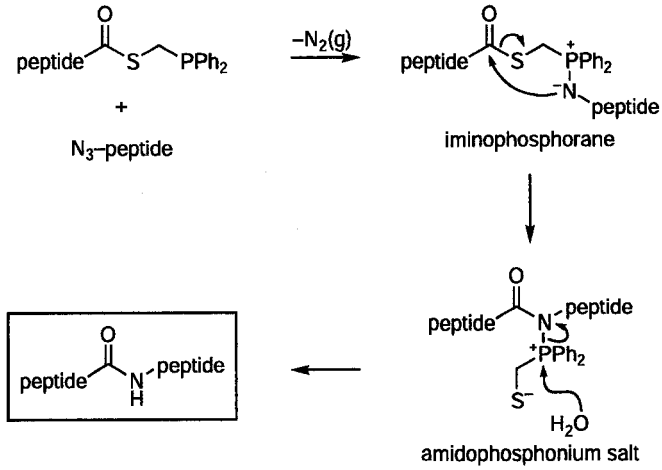
Chapter 4[‡]

PROTEIN ASSEMBLY BY ORTHOGONAL CHEMICAL LIGATION METHODS

Chemical synthesis harbors the potential to provide ready access to natural proteins as well as to create nonnatural ones. Indeed, numerous proteins have already been assembled from synthetic peptides (Dawson & Kent, 2000; Kent, 2003). "Native chemical ligation"—the coupling of a peptide (Wieland et al., 1953; Dawson et al., 1994) (or protein (Muir et al., 1998; Severinov & Muir, 1998; Muir, 2003)) containing a C-terminal thioester with another peptide containing an *N*-terminal cysteine residue—has been especially efficacious (Dawson & Kent, 2000). Emerging strategies for protein assembly avoid the need for a cysteine residue at the ligation junction. The Staudinger ligation is one such strategy (Bosch et al., 1993; Nilsson et al., 2000; Saxon et al., 2000; Saxon & Bertozzi, 2000a; Nilsson et al., 2001; Soellner et al., 2002).

[‡] This chapter has been published previously under the same title. Reference: Nilsson, B. L.; Hondal, R. H.; Soellner, M. B.; Raines, R. T. *J. Am. Chem. Soc.* **2003**, *125*, 5268-5269.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 4.1 *Staudinger ligation.*

In our version of the Staudinger ligation, a peptide containing a C-terminal phosphinothioester reacts with a peptide containing an N-terminal azide to give an amide with no residual atoms, as in Scheme 4.1 (Nilsson et al., 2001; Soellner et al., 2002). The initial intermediate is an iminophosphorane, which rearranges to an amidophosphonium salt. Hydrolysis yields the amide and a phosphine oxide. Previously, we showed that this reaction can be used to form dipeptides in high isolated yield (>90%) and with the retention of α-carbon stereochemistry (Nilsson et al., 2001; Soellner et al., 2002). Herein, we describe the first use of the Staudinger ligation to couple peptides on a solid support. We use the fragment thus produced to assemble a protein via native chemical ligation. The synthesis of a protein by this route expands the versatility of chemical approaches to protein production.

As a model protein, we chose ribonuclease A (RNase A (Raines, 1998); 124 amino acid residues), which was the first protein to succumb to total synthesis (Gutte & Merrifield, 1969; Hirschmann et al., 1969). Herein, RNase A was successfully assembled

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

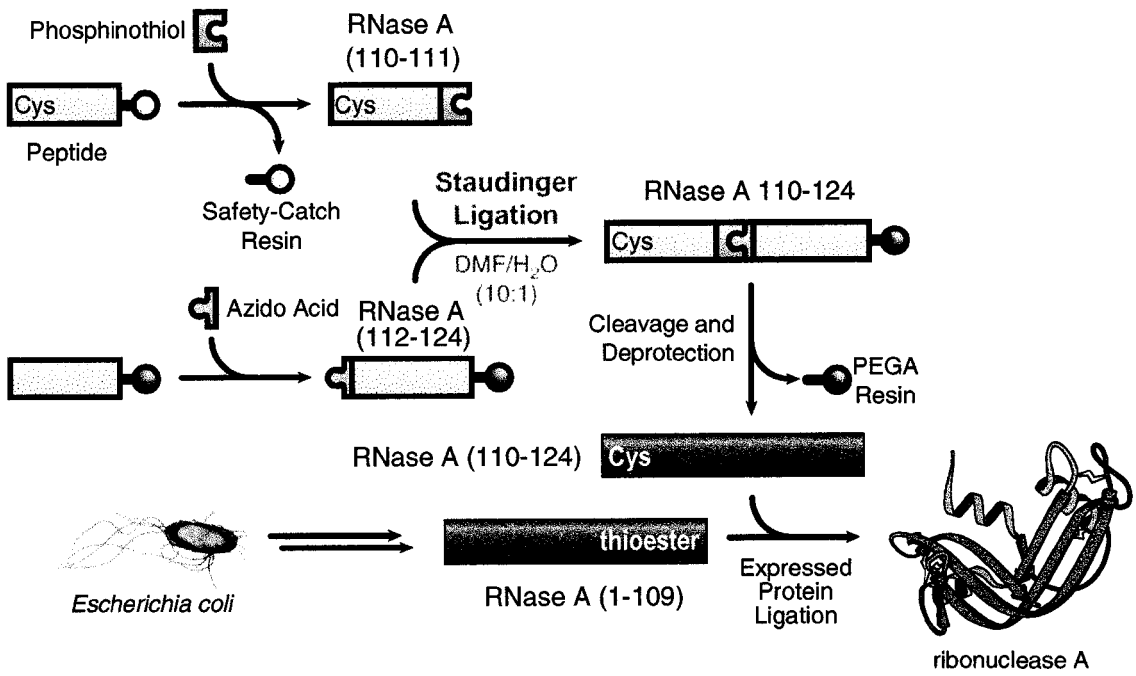
from three fragments, which comprised residues 1–109, 110–111, and 112–124. Four distinct amide-bond forming methods were used (Table 4.1). The route is shown in Scheme 4.2.

Four Distinct Methods to Form an Amide Bond

Bond	Method
1–109	*m*RNA Translation by Ribosome
109–110	Expressed Protein Ligation
110–111	Solid-Phase Peptide Synthesis
111–112	Staudinger Ligation
112–124	Solid-Phase Peptide Synthesis

Table 4.1 *Methods of amide bond formation in RNase A assembly.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 4.2 *Strategy for the orthogonal assembly of RNase A.*

The initial effort to effect the Staudinger ligation on the solid phase was attempted with fragment 110–112 (FmocCys(Trt)Glu(O*t*Bu)GlySCH$_2$PPh$_2$) to fragment N$_3$-113–124 on resin. Four equivalents of 110-112 were used in THF:H$_2$O (10:1) or DMF:H$_2$O (10:1) over 12 h. In both cases only traces of the ligated product (110-124) were isolated upon cleavage from resin and side group deprotection of the product peptide. The major isolated peptide as characterized by MALDI mass spectrometry following HPLC purification was fragment 113–124 with the *N*-terminus reduced to the amine. A subsequent attempt to couple at this ligation site utilized AcGlySCH$_2$PPh$_2$ as a coupling partner for the resin bound N$_3$-113–124 under the same conditions described above (THF:H$_2$O solvent). This ligation was successful, yielding the desired ligation product (Ac-112–124) exclusively. A final ligation experiment was attempted in solution.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fragment 110–112 (FmocCys(Trt)Glu(O*t*Bu)GlySCH$_2$PPh$_2$) and the benzamide of azido asparagine (N_3-Asn-Bn) were mixed in equimolar amounts in THF:H_2O (3:1). After silica gel chromatography the desired ligation product was isolated in 25% yield with the major products being the reduced Asn-Bn and the oxidized phosphinothioester of fragment 110–112. These results suggested to us some potential sequence specificity to the ligation of bulkier phosphinothioesters.

We then explored a frameshift for the ligation to fragment 110–111 (FmocCys(Trt)Glu(O*t*Bu)SCH$_2$PPh$_2$) and fragment N_3-112-124. We reasoned that ligation of a bulky phosphinothioester to a glycyl azide may be less encumbered. We initially attempted a model coupling in solution using fragment 110-111 and N_3-Gly-Bn. The desired ligation product was isolated in 80% yield after chromatographic purification. With this result in hand we incubated four equivalents of fragment 110–111 with resin bound N_3-112–124 in either THF:H_2O (10:1) or DMF:H_2O (10:1). In THF the ligation was unsuccessful. We were surprised to discover that the major product was N_3-112–124. Perhaps the azide remains intact due to resin collapse under the reaction conditions. We were gratified to find, however, that in DMF the ligation proceeds efficiently with the desired ligation product (RNase A 110–124) isolated in 61% yield upon cleavage from resin, side group deprotection, and HPLC purification. RNase A fragment 110-124 was then synthesized once again under these conditions incorporating a proline containing two carbon-13s (α and carbonyl) and one nitrogen-15 at position 117 and a proline with a single carbon-13 (carbonyl) at position 114 as specific NMR tags.

RNase A(110–111) was synthesized as a C-terminal phosphinothioester using a sulfonamide-linker ("safety-catch") resin (Kenner et al., 1971; Backes & Ellman, 1999).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The fully loaded resin was activated with iodoacetonitrile (Backes & Ellman, 1999). Treatment with an excess of diphenylphosphinomethanethiol (Nilsson et al., 2001; Soellner et al., 2002) in DMF for 18 h liberated FmocCys(Trt)Glu(O*t*Bu)SCH_2PPh_2, which was isolated in 64% yield.

RNase A(112–124) was synthesized as an *N*-terminal azide. The $n-1$ peptide was synthesized by using standard Fmoc-protection and HATU activation on a hydroxyethylpolystyrene (PEGA) resin, which has diverse solvent compatibility. α-Azido glycine (Nilsson et al., 2000; Nilsson et al., 2001) (residue 112) was used to cap the $n-1$ peptide by its activation with PyBOP, HOBt, and DIPEA in DMF. The identity of N_3CH_2C(O)Asn(Trt)ProTyr(*t*Bu)ValProValHis(Trt)PheAsp(O*t*Bu)AlaSer(*t*Bu)Val was confirmed by cleaving a small amount of the α-azido peptide from the resin with TFA and analyzing by MALDI mass spectrometry.

RNase A(110–111) and RNase A(112–124) were coupled by Staudinger ligation directly on the PEGA resin. As described above, four equivalents of RNase A(110–111) in DMF/H_2O (10:1) were incubated with the resin over 12 h. After cleavage from the resin, side-chain deprotection, and HPLC purification, RNase A(110–124) was isolated in 61% yield.

RNase A(110–124) was synthesized again by this route, now incorporating an NMR probe of protein structure. Specifically, [$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]proline (a gift from the National Stable Isotope Resource at Los Alamos) was inserted at position 114. The Asn113–Pro114 peptide bond resides in the *cis* (*E*) conformation in the properly folded protein, but a mixture of *cis* and *trans* conformations in peptide fragments (An et al., 1999).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RNase A(1–109) was produced by biosynthesis as a *C*-terminal thioester with N-methylmercaptoacetamide, as described previously (Evans et al., 1998; Hondal et al., 2001; Arnold et al., 2002). Both unlabeled and labeled RNase A(110–124) contained an *N*-terminal cysteine residue. Ligation of RNase A(1–109) and RNase A(110–124) in aqueous buffer, folding, and purification yielded intact RNase A. Its molecular mass was verified by MALDI mass spectrometry.

The ability to incorporate labeled amino acids at specific sites is a distinct advantage of producing proteins by chemical synthesis (Borgia & Fields, 2000; Dawson & Kent, 2000; Miranda & Alewood, 2000; Tam et al., 2001). The route in Scheme 2 was used to incorporate [$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]Pro114 into RNase A, and a 1D HSQC NMR experiment was used to probe for proper folding of the resulting protein. In Fmoc[$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]ProOH, the carbamyl C–N bond was a mixture of *cis* and *trans* isomers (Figure 1). In contrast, the Asn113–[$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]Pro114 peptide bond in labeled RNase A was a single species, consistent with this C–N bond being only in the *cis* conformation. Moreover, the chemical shift of the α-proton of [$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]Pro114 in the synthetic RNase A was identical to that of unlabeled Pro114 in natural RNase A (Rico et al., 1989).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

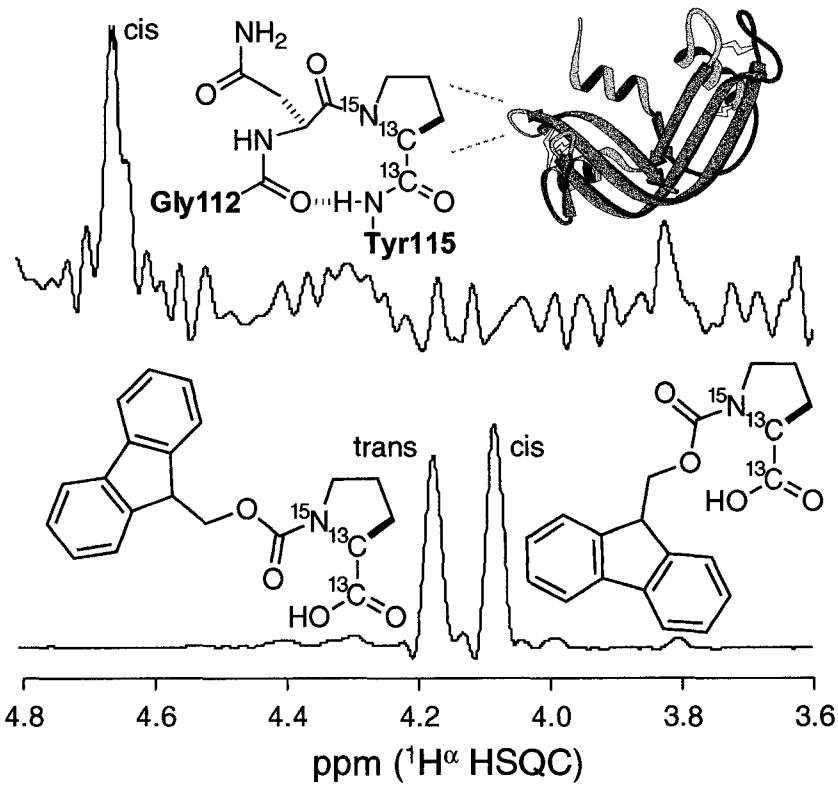


Figure 4.1 Observation of the α-proton of [$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]proline by 1D ^{15}N-CT-HSQC NMR spectroscopy. Top: [$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]Pro114 RNase A in D_2O (δ 4.66 ppm). Bottom: Fmoc[$^{13}C'$,$^{13}C^{\alpha}$,^{15}N]ProOH in D_2O/CD_3OD (1:1), assigned by 2D HMBC NMR spectroscopy.

Enzymatic activity provides an extremely sensitive measure of protein structure (Knowles, 1987). The enzymatic activity of the RNase A synthesized as in Scheme 2 ($k_{cat}/K_M = 0.94 \times 10^7\ M^{-1}s^{-1}$) was nearly equal to that of the wild-type enzyme produced by recombinant DNA technology ($k_{cat}/K_M = 1.1 \times 10^7\ M^{-1}s^{-1}$) (Kelemen et al., 1999; Hondal et al., 2001; Arnold et al., 2002).

Thus, the solid-phase assembly of peptides with the Staudinger ligation has not only been realized, but used to assemble a functional enzyme. This method for amide bond formation is orthogonal and complementary to other ligation methods (Staudinger &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Meyer, 1919; Bosch et al., 1993; Nilsson et al., 2000; Saxon et al., 2000; Saxon & Bertozzi, 2000a; Nilsson et al., 2001; Soellner et al., 2003). The enzyme created herein is remarkable in that its peptide bonds were synthesized in four distinct processes: *m*RNA translation by a ribosome, solid-phase peptide synthesis, native chemical ligation, and solid-phase Staudinger ligation (Scheme 2). We anticipate that the two solid-phase processes alone could be the basis for an automated means to assemble proteins, making enzymes and other proteins more accessible targets for synthetic chemistry.

Acknowledgment. We thank W. M. Westler for help and advice with NMR experiments. This work was supported by grant GM44783 (NIH). B.L.N. was supported by the Abbott Laboratories Fellowship in Synthetic Organic Chemistry (University of Wisconsin–Madison). R.J.H. was supported by Postdoctoral Fellowship GM20190 (NIH). HSQC and HMBC NMR spectroscopy was performed at the National Magnetic Resonance Facility at Madison, which was supported by grant P41 RR02301 (NIH). The SIR^9 was supported by grant RR02231 (NIH).

Experimental Section

General Experimental. Chemicals and solvents were from Aldrich® with the exception of Fmoc-protected amino acids and alkanesulfonamide safety-catch resins (Novabiochem®). Reactions were monitored by thin-layer chromatography and visualized by UV light or staining with I_2. Peptide synthesis was performed by using standard Fmoc-protection strategies on an Applied Biosystems Pioneer automated synthesizer with HATU activation. HPLC was performed on a C18 reverse-phase column. NMR spectra were obtained with a Bruker AC-300, Bruker DMX-400, Varian UNITY-500, or Varian

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

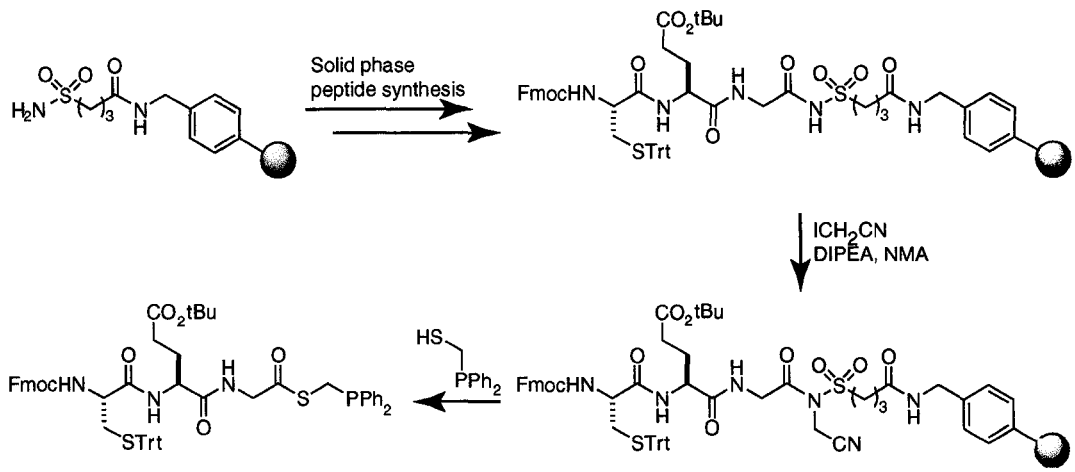
Inova-600 spectrometer. Phosphorus-31 NMR spectra were proton-decoupled and referenced against an external standard of deuterated phosphoric acid. Mass spectra were obtained using electrospray ionization (ESI) or matrix assisted laser desorption ionization (MALDI) techniques.

NMR spectral assignments of FmocProOH *cis/trans* isomers. The *cis/trans* assignments for the $^{1}H^{\alpha}$ resonances of FmocProOH (Figure 1, lower spectrum) were made by correlating the resonances of the α-proton to those of the β- and γ-carbons using a 2D HMBC experiment. The β- and γ-carbons of proline amides each have two distinct ^{13}C resonances due to *cis/trans* amide bond isomerization. The ^{13}C chemical shifts for the *cis* and *trans* isomers are constant among different proline derivatives (Dorman & Bovey, 1973; Wuthrich, 1976). The ^{13}C shifts in FmocProOH were correlated to the two $^{1}H^{\alpha}$ resonances of FmocProOH, and the peak at 4.18 ppm was thereby assigned to arise from the *trans* isomer and the peak at 4.08 ppm to arise from the *cis* isomer.

Synthesis of FmocCys(Trt)Glu(O*t*Bu)GlySCH$_2$PPh$_2$. Fmoc-Gly-OH was loaded onto 4-sulfamylbutyryl resin according to the method of Backes and Ellman. The protocol was repeated twice to give optimal loading. Fmoc-deprotection was performed using 30% piperidine in DMF. Couplings were performed by premixing amino acid (3 eq.), PyBOP (3 eq.), HOBt (3 eq.), and DIPEA (6 eq.) in DMF and adding the mixture to the preswollen resin for 90 minutes. The resin containing the synthetic peptide was activated with iodoacetonitrile by the method of Backes and Ellman. The final

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phosphinothioester was liberated from the activated resin by incubating the resin (1.15 g, 0.48 mmol peptide loading) with Ph_2PCH_2SH (882 mg, 3.8 mmol) in DMF (20 mL) for 18 h. The resin was then filtered and washed (5X 10 mL DMF, 5X 10 mL CH_2Cl_2) and the organic solvents were removed under reduced pressure. The residue was purified by chromatography (silica gel, 30% EtOAc in hexanes). The desired peptide phosphinothioester was isolated in 57% yield.



Scheme 4.3 *Safety–catch synthesis of RNase A 110–112 phosphinothioester.*

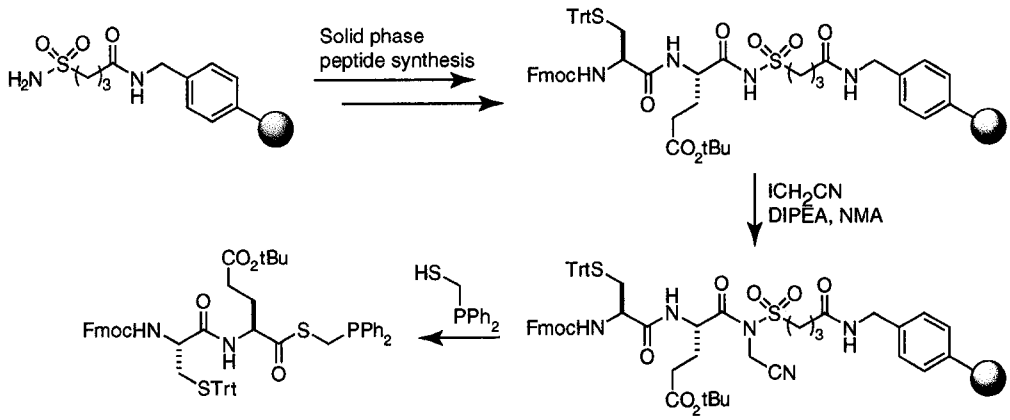
Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.77-7.73 (m, 2 H), 7.60-7.56 (m, 2 H), 7.41-7.18 (m, 29 H), 5.16 (d, 1 H, J = 6.0 Hz), 4.46-4.31 (m, 3 H), 4.19 (t, 1 H, J = 6.0 Hz), 4.13-4.05 (m, 1 H), 3.83-3.76 (m, 2 H), 3.49-3.39 (m, 1 H), 2.77-2.61 (m, 2 H), 2.49-2.26 (m, 2 H), 2.13-1.90 (m, 2 H), 1.41 (s, 9 H) ppm; ^{13}C NMR (63 MHz, $CDCl_3$) δ 195.45, 173.02, 170.88, 170.06, 155.78, 143.91, 143.44 (d, J = 12.5 Hz), 140.99, 136.46 (d, J = 14.0 Hz), 132.44 (d, J = 17.6 Hz), 129.21, 128.89, 128.31 (d, J = 6.6 Hz), 127.94, 127.52, 126.83 (d, J = 2.3 Hz), 124.84, 119.74, 80.72, 67.06, 66.92, 53.90, 52.65, 48.48,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

46.80, 33.69 (d, J = 7.2 Hz), 31.41, 27.80, 26.51, 25.33, 24.95, 24.73 ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ -15.58 ppm; MS (ESI) *m/z* 1064.3486 (MNa^+ [$C_{61}H_{60}N_3O_7PS_2Na$] = 1064.3508).

Synthesis of FmocCys(Trt)Glu(O*t*Bu)SCH$_2$PPh$_2$. Method A. FmocGlu(O*t*Bu)OH was loaded onto 4-sulfamylbutyryl resin according to the method of Backes and Ellman (Kenner et al., 1971; Backes & Ellman, 1994; Backes et al., 1996; Backes & Ellman, 1999; Ingenito et al., 1999; Shin et al., 1999). The protocol was repeated twice to give optimal loading. Fmoc-deprotection was performed using 30% (v/v) piperidine in DMF. Coupling was performed by premixing the protected amino acid (3 eq), PyBOP (3 eq), HOBt (3 eq), and DIPEA (6 eq) in DMF, and adding the mixture to the pre-swollen resin for 90 min. The resin containing the synthetic peptide was activated with iodoacetonitrile according to the method of Backes and Ellman. The final phosphinothioester was liberated by incubating the resin (1.0 g, 1.12 mmol peptide loading) with Ph_2PCH_2SH (2.1 g, 9.0 mmol) in DMF (15 mL) for 18 h. The resin was then filtered and washed (5 × 10 mL DMF, 5 × 10 mL CH_2Cl_2), and solvents were removed under reduced pressure. The residue was purified by chromatography (silica gel, 30% EtOAc in hexanes) to yield 706 mg (0.72 mmol, 64% yield based on a 1.12-mmol resin loading) of FmocCys(Trt)Glu(O*t*Bu)SCH$_2$PPh$_2$.

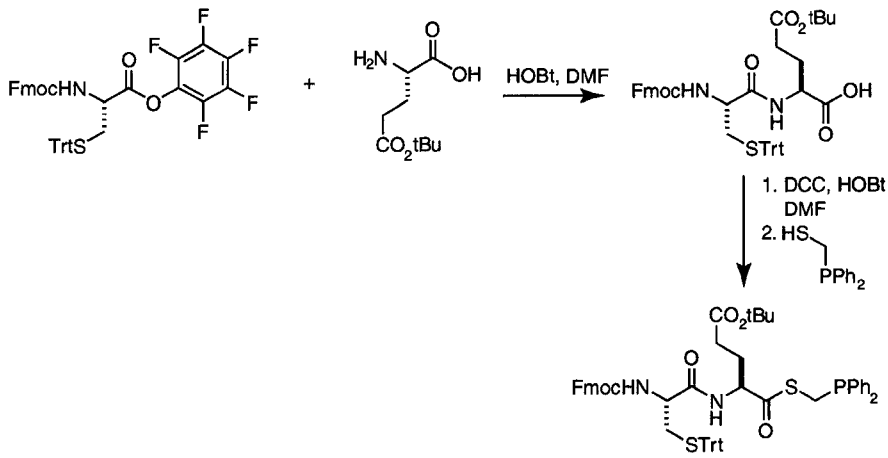
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 4.4 *Safety-catch synthesis of RNase 110–111 phosphinothioester.*

Method B. The pentafluorophenolic ester of Fmoc-Cys(Trt)-OH (4.68 g, 6.2 mmol) and HOBt (0.84 g, 6.2 mmol) were dissolved in dry DMF (25 mL) in a flame-dried reaction flask under argon. To this mixture was added H-Glu(O*t*Bu)-OH (1.26 g, 6.2 mmol). The reagents were stirred under argon for 12 h and the solvent was removed under reduced pressure. The residue was purified by chromatography (silica gel, 4% MeOH in CH_2Cl_2). FmocCys(Trt)Glu(O*t*Bu)OH was isolated in 98% yield. FmocCys(Trt)Glu(O*t*Bu)OH (2.00 g, 2.6 mmol) was charged to a flame-dried reaction flask under argon and dissolved in dry CH_2Cl_2 (25 mL). To this solution was added HOBt (0.35 g, 2.6 mmol) followed by DCC (0.54 g, 2.6 mmol). Upon formation of DCU by-product Ph_2PCH_2SH (0.66 g, 2.9 mmol) was added and the mixture was stirred for 4 h, DCU was filtered, and solvent was removed under reduced pressure. The residue was purified by chromatography (silica gel, 35% EtOAc in hexanes). FmocCys(Trt)Glu(O*t*Bu)SCH_2PPh_2 was isolated in 85% yield.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 4.5 *Solution-phase synthesis of RNase A 110–111 phosphinothioester.*

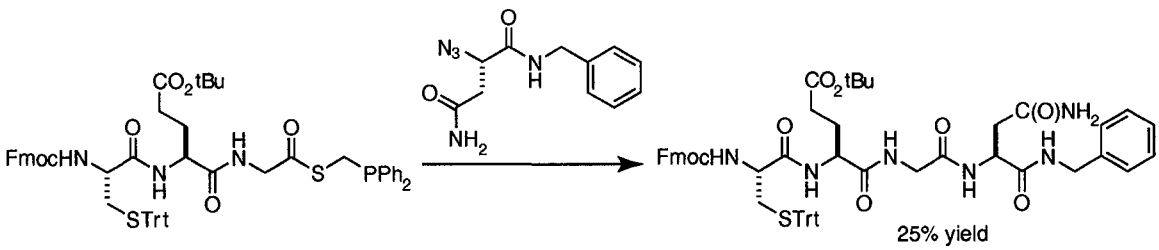
Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75–7.70 (m, 2H), 7.57–7.55 (m, 2H), 7.42–7.14 (m, 29H), 6.68 (d, J = 6.6 Hz, 1H), 5.13 (d, J = 8.1 Hz, 1H), 4.56–4.50 (m, 1H), 4.36–4.34 (m, 2H), 4.19–4.17 (m, 1H), 3.81–3.80 (m, 1H), 3.44–3.38 (m, 2H), 2.78–2.68 (m, 1H), 2.61–2.57 (m, 1H), 2.27–2.23 (m, 2H), 2.11–1.95 (m, 1H), 1.83–1.70 (m, 1H), 1.37 (s, 9H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 198.14, 171.89, 170.19, 155.81, 144.17, 143.59, 143.46, 141.10, 136.49 (d, J = 14.0 Hz), 132.69 (d, J = 4.2 Hz), 132.44 (d, J = 4.1 Hz), 129.41, 128.98, 128.38 (d, J = 6.6 Hz), 127.93, 127.59, 126.94, 126.74, 124.93, 119.80, 80.74, 67.21, 67.00, 58.50, 53.83, 46.89, 31.00, 27.85, 27.23, 25.45 (d, J = 24.8 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ –14.51 ppm; MS (ESI) *m/z* 1007.3340 (MNa^+ [$C_{59}H_{57}N_2O_6PS_2Na$] = 1007.3371).

Synthesis of FmocCys(Trt)Glu(O*t*Bu)GlyAsn(Trt)Bn by coupling FmocCys(Trt)Glu(O*t*Bu)GlySCH_2PPh_2 and N_3-Asn-Bn

FmocCys(Trt)Glu(O*t*Bu)GlySCH_2PPh_2 (0.25 g, 0.23 mmol) and N_3-Asn-Bn (0.56 g,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

0.23 mmol) were dissolved in THF:H_2O (3:1, 3.8 mL) and stirred for 12 h. The solvent was removed under reduced pressure and the products were purified by chromatography (silica gel, 5% MeOH in CH_2Cl_2). FmocCys(Trt)Glu(O*t*Bu)GlyAsn(Trt)Bn was isolated in 25% yield.



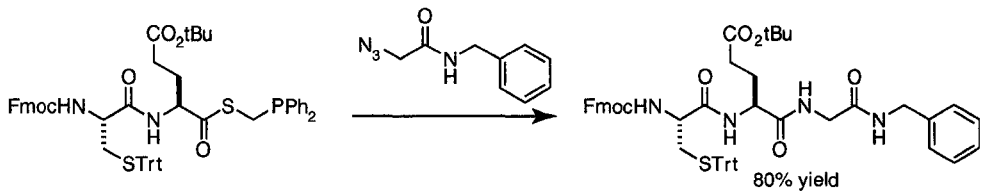
Scheme 4.6 *Staudinger ligation of RNase A 110–112 thioester to N_3-Asn-Bn.*

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$:CD_3OD (1:1)) δ 7.80-7.76 (m, 2 H), 7.65-7.61 (m, 2 H), 7.46-7.26 (m, 24 H), 4.78 (t, 1 H, J = 5.4 Hz), 4.28-4.22 (m, 2 H), 3.99-3.96 (m, J = 1 Hz), 3.84-3.54 (m, 2 H), 2.81-2.74 (m, 2 H), 2.68-2.58 (m, 2 H), 2.30 (t, J = 6.9 Hz), 2.16-2.08 (m, 1 H), 1.88-1.81 (m, 1 H), 1.41 (s, 9 H), 1.28 (s, 2 H), 0.92-0.88 (m, 1 H) ppm; ^{13}C NMR (125 MHz, $CDCl_3$:CD_3OD (1:1)) δ 173.31, 172.51, 172.19, 171.12, 170.69, 169.35, 156.37, 143.84, 143.39, 143.21, 140.92, 137.56, 129.07, 128.06, 127.73, 127.41, 126.92, 126.74, 124.61, 119.59, 80.71, 66.85, 53.81, 52.80, 49.92, 46.65, 42.93, 42.62, 36.14, 33.35, 31.27, 29.25, 27.45, 25.70 ppm; MS (ESI) *m/z* 1053.4166 (MNa^+ [$C_{59}H_{62}N_6O_9SNa$] = 1053.4197).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Synthesis of FmocCys(Trt)Glu(O*t*Bu)GlyBn by coupling FmocCys(Trt)Glu(O*t*Bu)SCH_2PPh_2 and N_3-Gly-Bn

FmocCys(Trt)Glu(O*t*Bu)SCH_2PPh_2 and N_3-Gly-Bn were treated as described above. The residue was purified by chromatography (silica gel, 3% MeOH in CH_2Cl_2) and FmocCys(Trt)Glu(O*t*Bu)GlyBn was isolated in 80% yield.



Scheme 4.7 *Staudinger ligation of RNase A 110–111 to N_3-Gly-Bn.*

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$:CD_3OD (1:1)) δ 7.81-7.76 (m, 2 H), 7.63-7.61 (m, 2 H), 7.43-7.19 (m, 24 H), 4.45-4.32 (m, 4 H), 4.29-4.18 (m, 2 H), 3.92 (t, 1 H, J = 6 Hz), 3.89-3.71 (m, 2 H), 2.66-2.54 (m, 2 H), 2.35-2.30 (m, 2 H), 2.20-2.06 (m, 1 H), 1.98-1.85 (m, 1 H), 1.41 (s, 9 H) ppm; ^{13}C NMR (300 MHz, $CDCl_3$:CD_3OD (1:1)) δ 143.80, 143.33, 143.15, 140.87, 137.53, 128.99, 127.98, 127.62, 127.33, 127.10, 126.78, 126.66, 126.64, 126.54, 124.51, 119.51, 80.67, 66.75, 66.63, 53.67, 53.01, 46.61, 42.73, 42.01, 33.21, 31.18, 33.21, 31.18, 27.33, 25.74 ppm; MS (ESI) *m/z* 939.3733 (MNa^+ [$C_{55}H_{56}N_4O_7SNa$] = 939.3767).

Solid Phase Staudinger Ligation. Asn(Trt)ProTyr(tBu)ValProValHis(Trt)Phe-Asp(OtBu)AlaSer(tBu)Val was synthesized on a polystyrene resin with PEG linkers by using standard methods. α-Azido glycine (Nilsson et al., 2000; Nilsson et al., 2001) was coupled using PyBOP/HOBt/DIPEA preactivation of the azido acid in DMF prior to its

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

addition to the resin. The coupling reaction was allowed to proceed for 2 h at room temperature. Four equivalents of FmocCys(Trt)Glu(OtBu)SCH_2PPh_2 were added to the resin-linked peptide in DMF/H_2O (10:1). Approximately 1.5 mL of solvent were used per 180 mg of resin (25-μmol loading). The mixture was agitated gently for 12 h, after which the solvent was removed by filtration and the resin was rinsed with DMF (5 × 10 mL) and CH_2Cl_2 (5 × 10 mL). The resin was dried under vacuum and then treated with a cleavage cocktail (95% TFA, 2.5% H_2O, 2.5% ethanedithiol) for 2 h. The cleavage cocktail was separated from the resin by filtration and added to ice-cold diethyl ether to precipitate the deprotected peptide product. The peptide product was purified by reverse-phase HPLC and analyzed by MALDI mass spectrometry. Other ligations were performed under identical conditions. The results are listed in Table 4.2.

Phosphinothioester peptide	Azido peptide	Peptide product	MALDI-MS	
			expected	found
Fmoc-110-112-R	N_3-113-124-R'	NH_2-113-124	1344.9	1344.7
Ac-Gly-R	N_3-113-124-R'	Ac-112-124	1443.7	1443.9
Fmoc-110-111-R	N_3-112-124-R'	NH_2-110-124	1633.7	1633.4
Fmoc-110-111-R	N_3-112-124-R' ($3X^{13}C$, $1X^{15}N$)	NH_2-110-124 ($3X^{13}C$, $1X^{15}N$)	1637.7	1637.6

R = SCH_2PPh_2, R' = resin, all peptide starting materials are side-chain protected

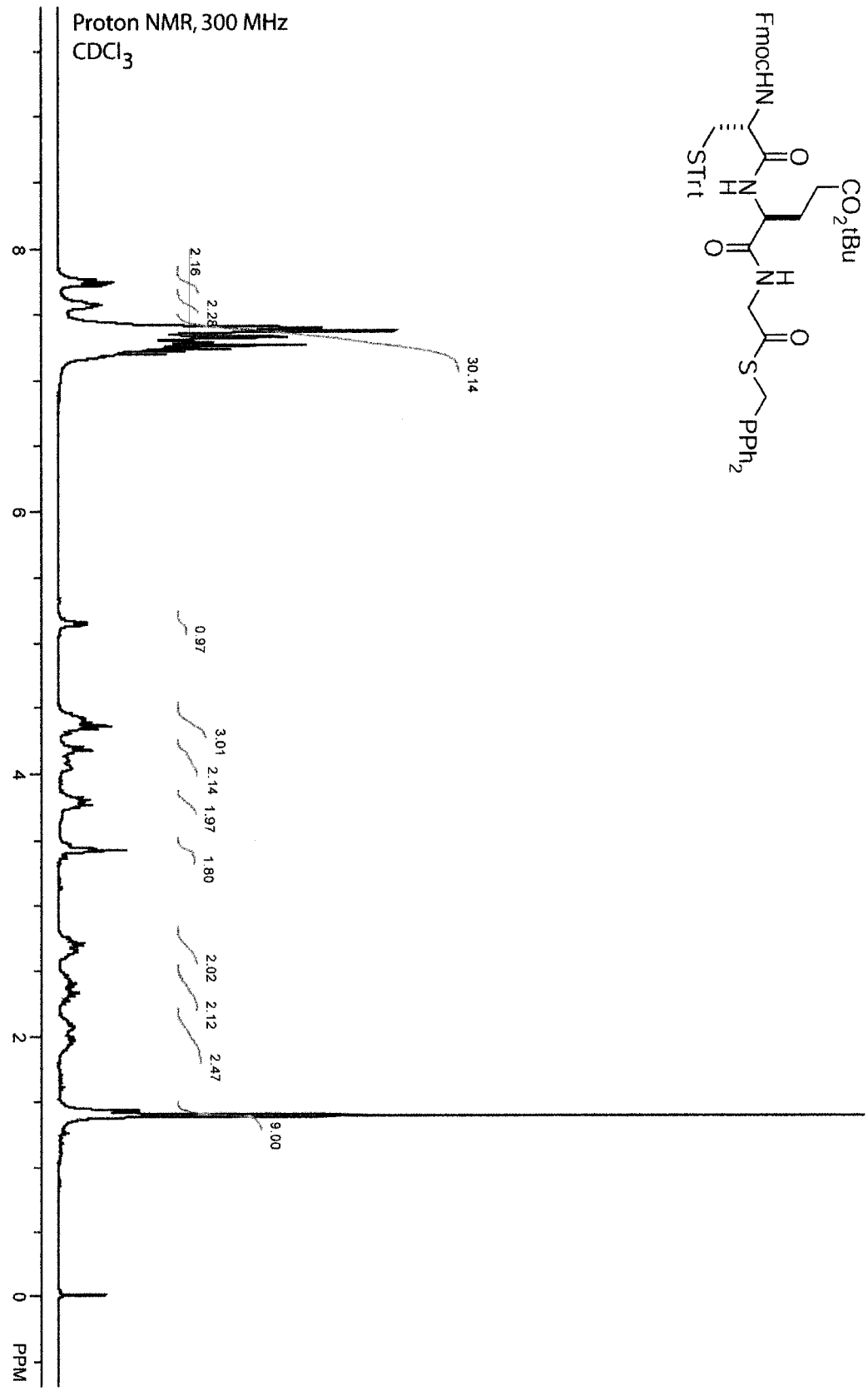
Table 4.2 *Selected mass spectrometry data for solid-phase peptide ligation products.*

Expressed protein ligation. Expressed protein ligation of RNase A(1–109) thioester, which also has an N-terminal methionine residue [Met(–1)], to RNase A(110–124) was performed as described previously (Hondal et al., 2001; Arnold et al., 2002; Hondal &

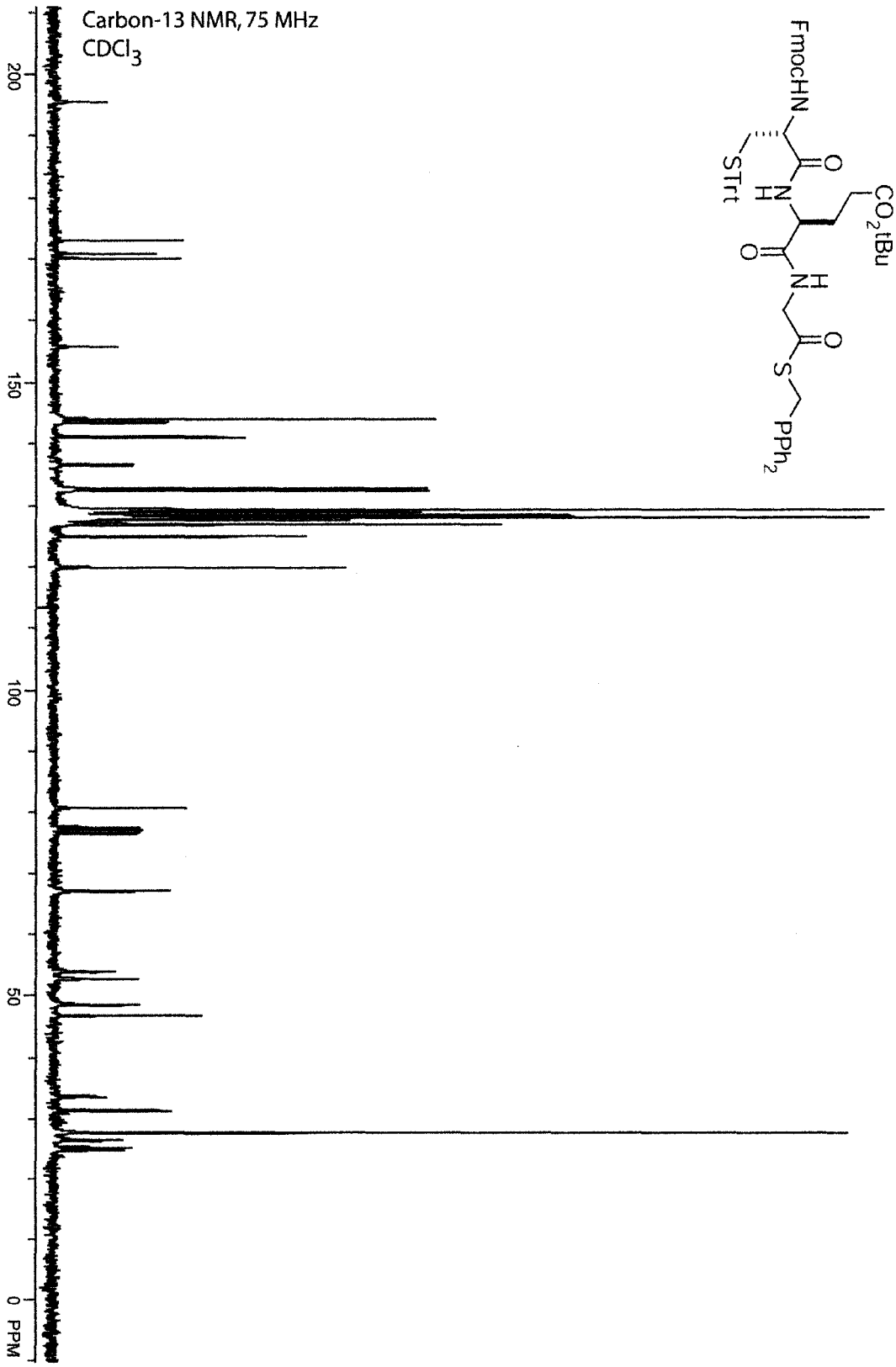
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Raines, 2002). MALDI *m/z* 13,832; expected, 13,819.

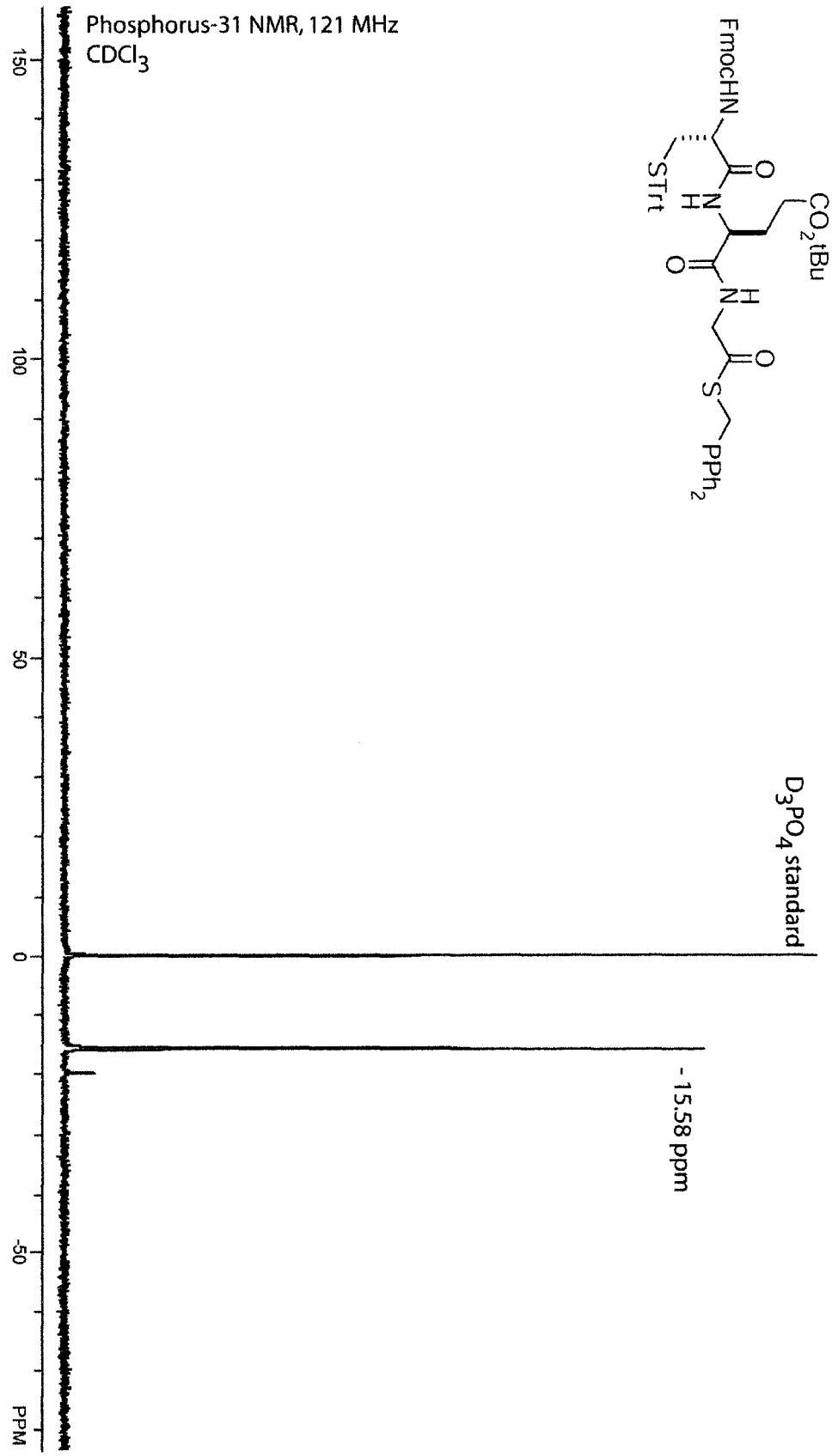
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



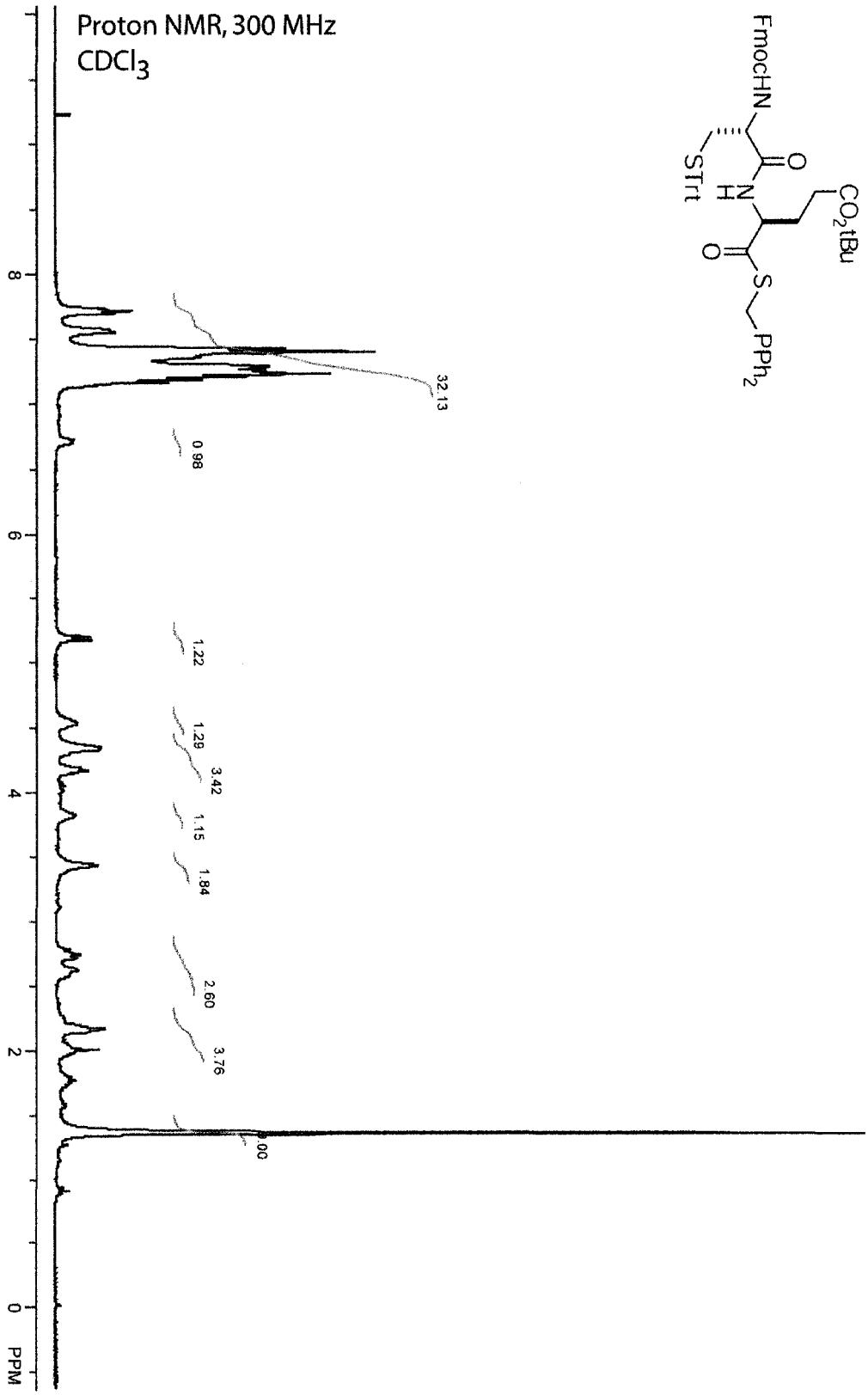
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



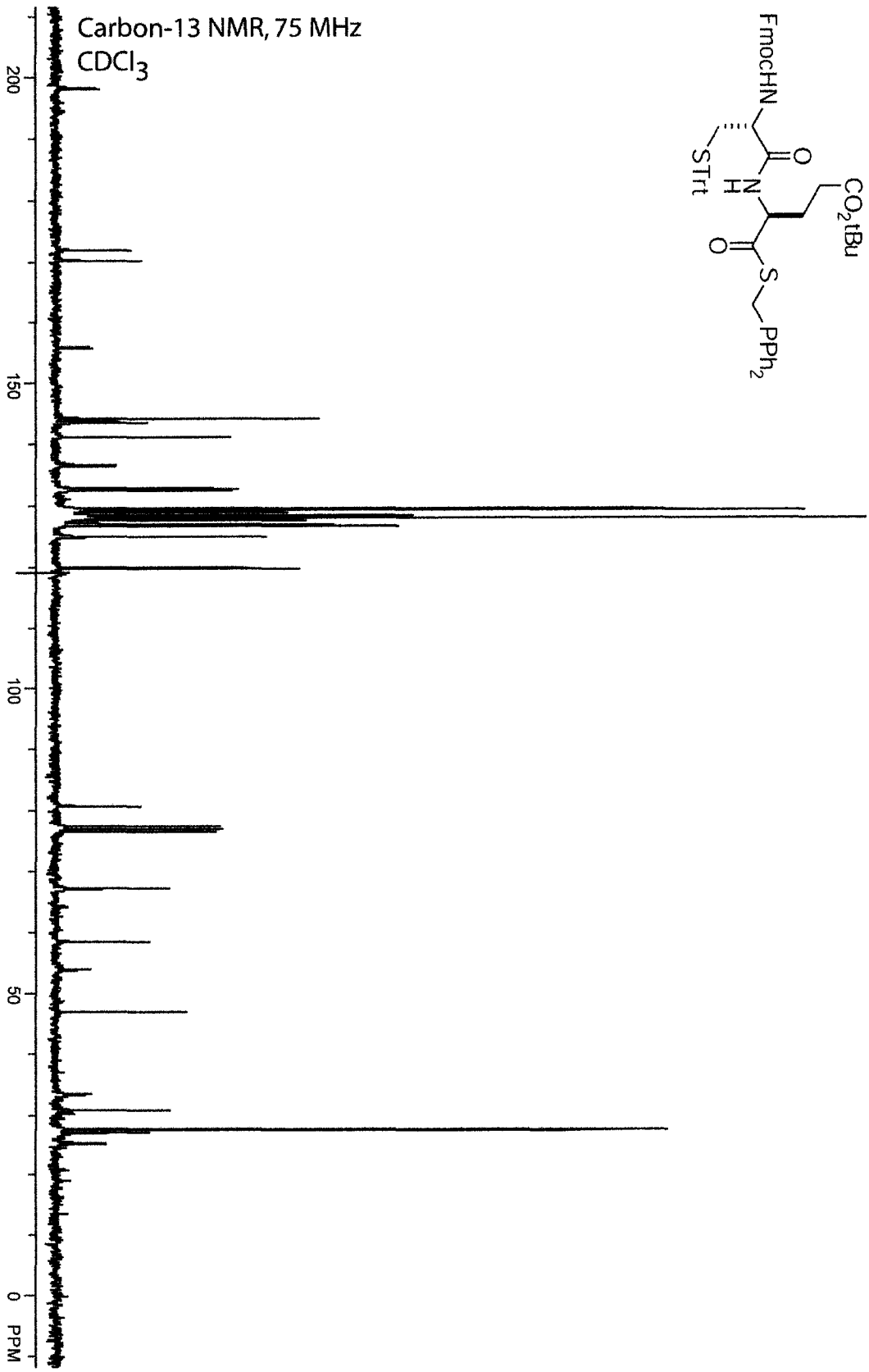
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



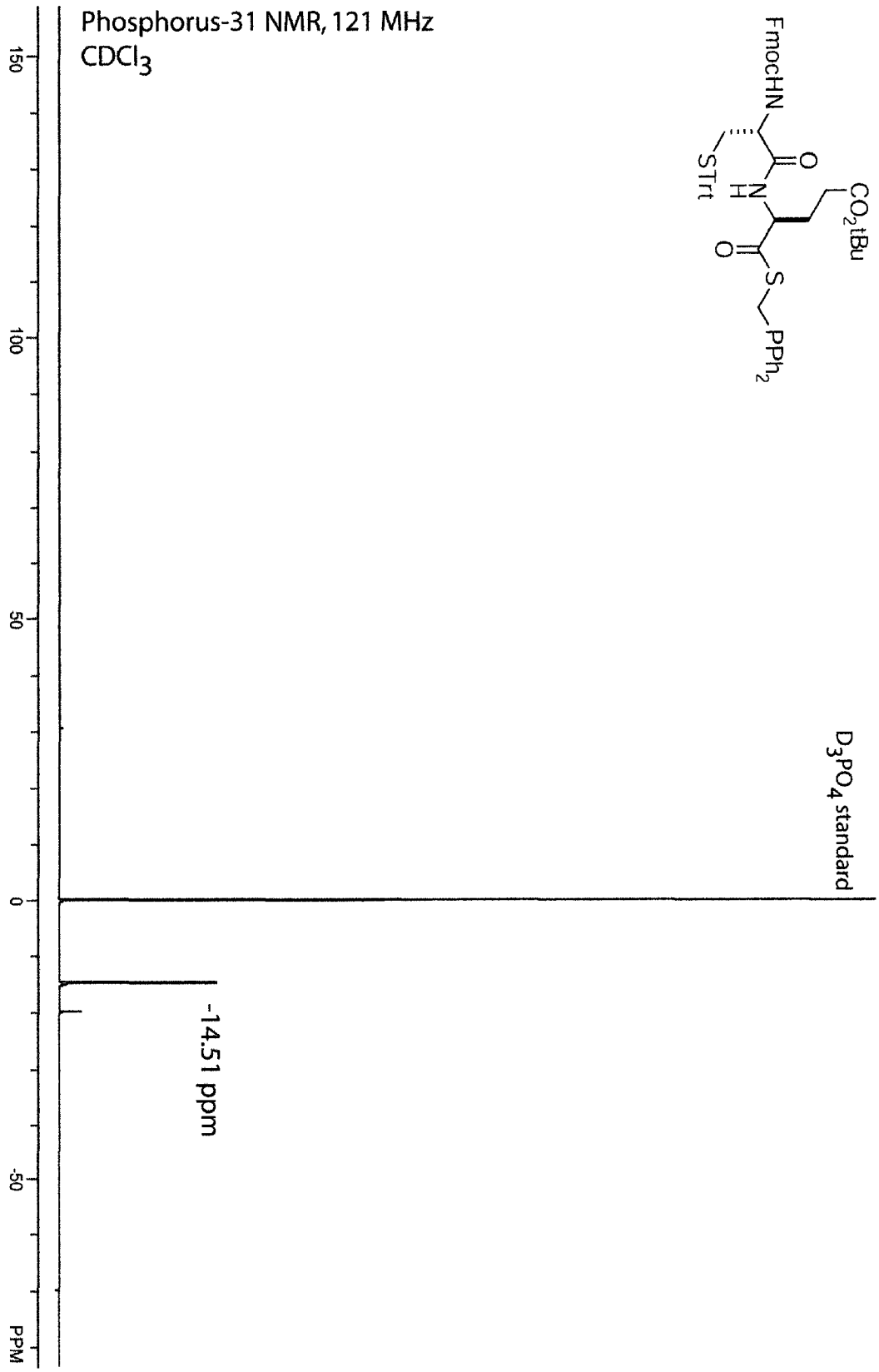
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



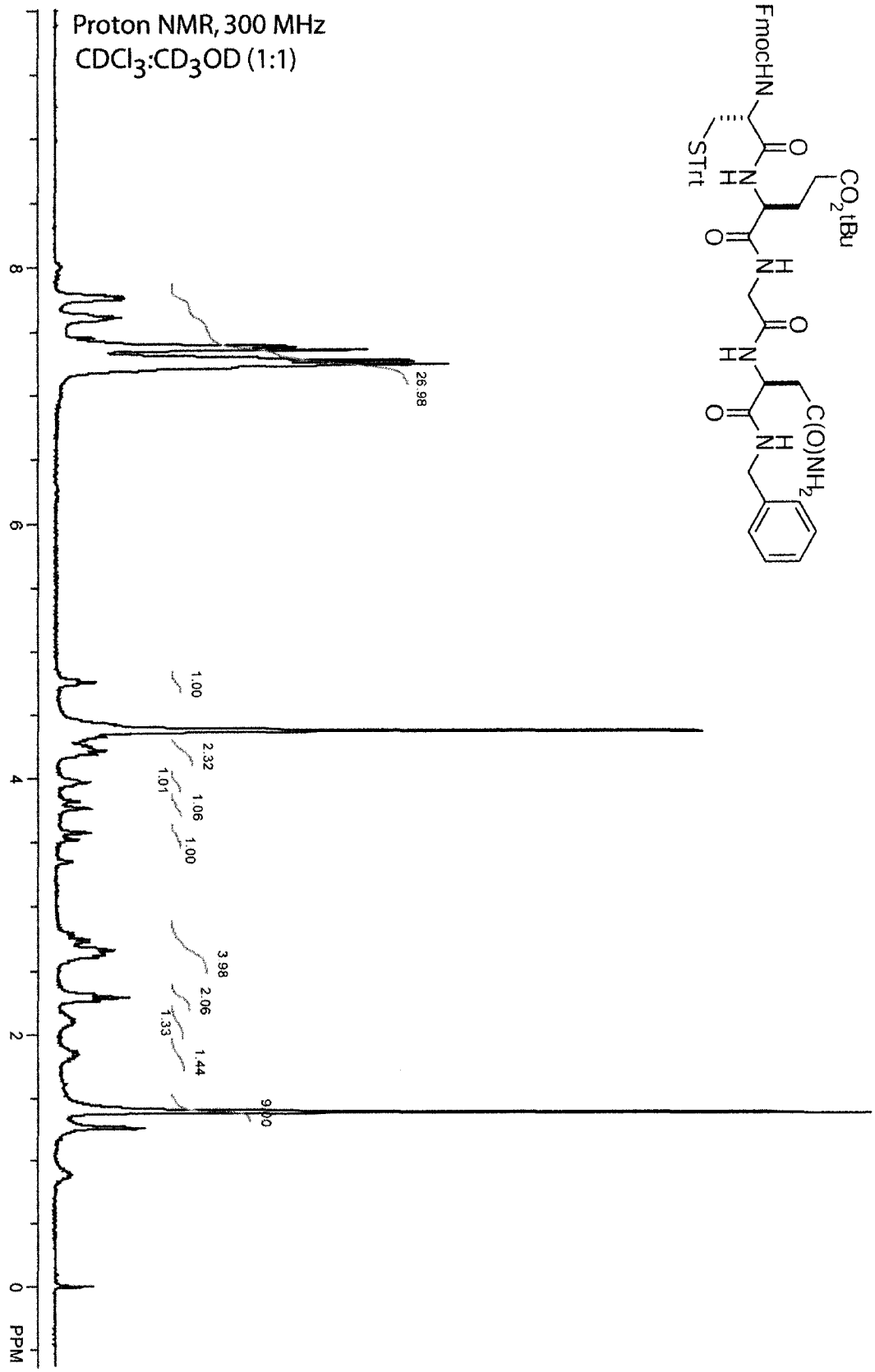
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



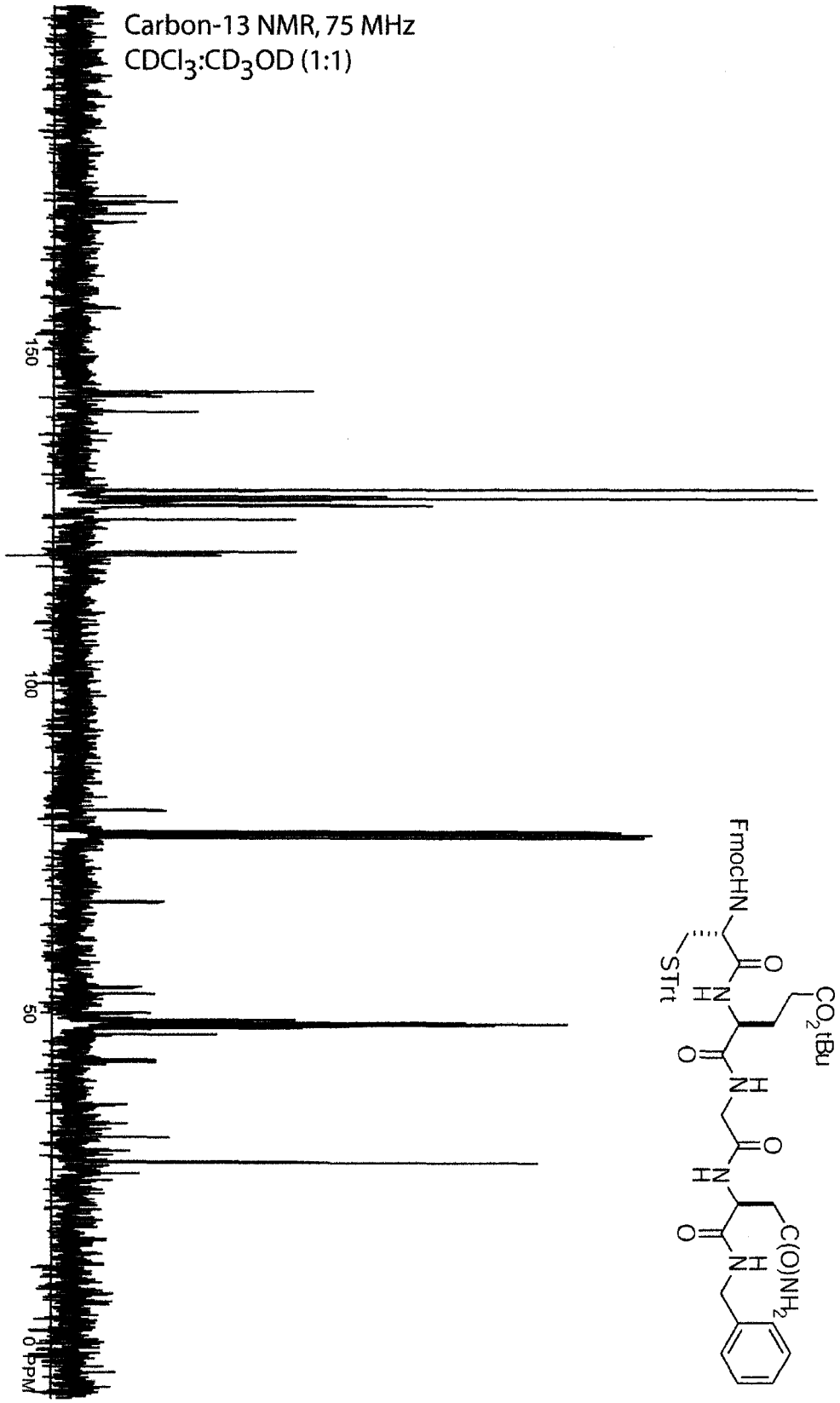
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



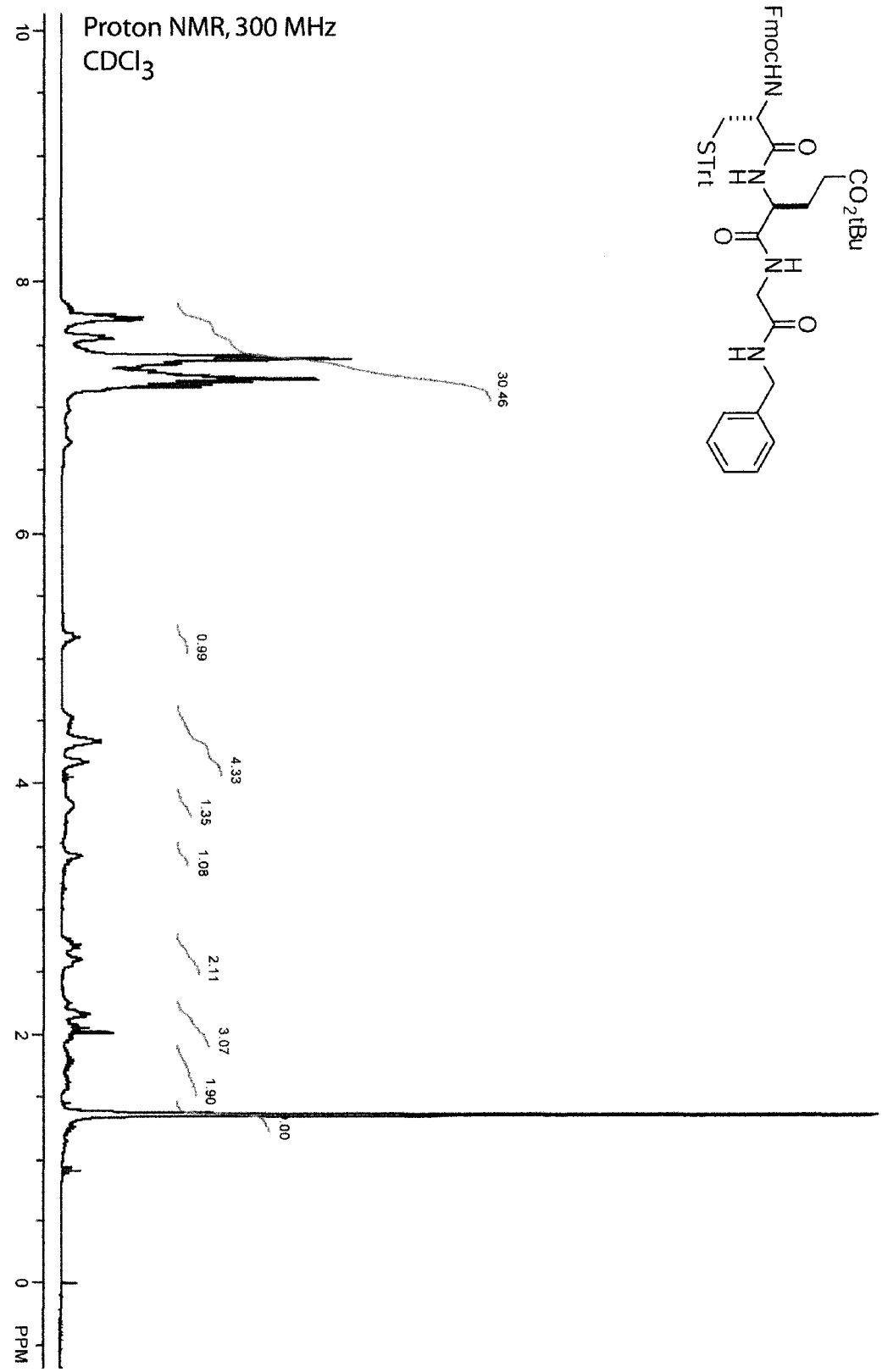
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



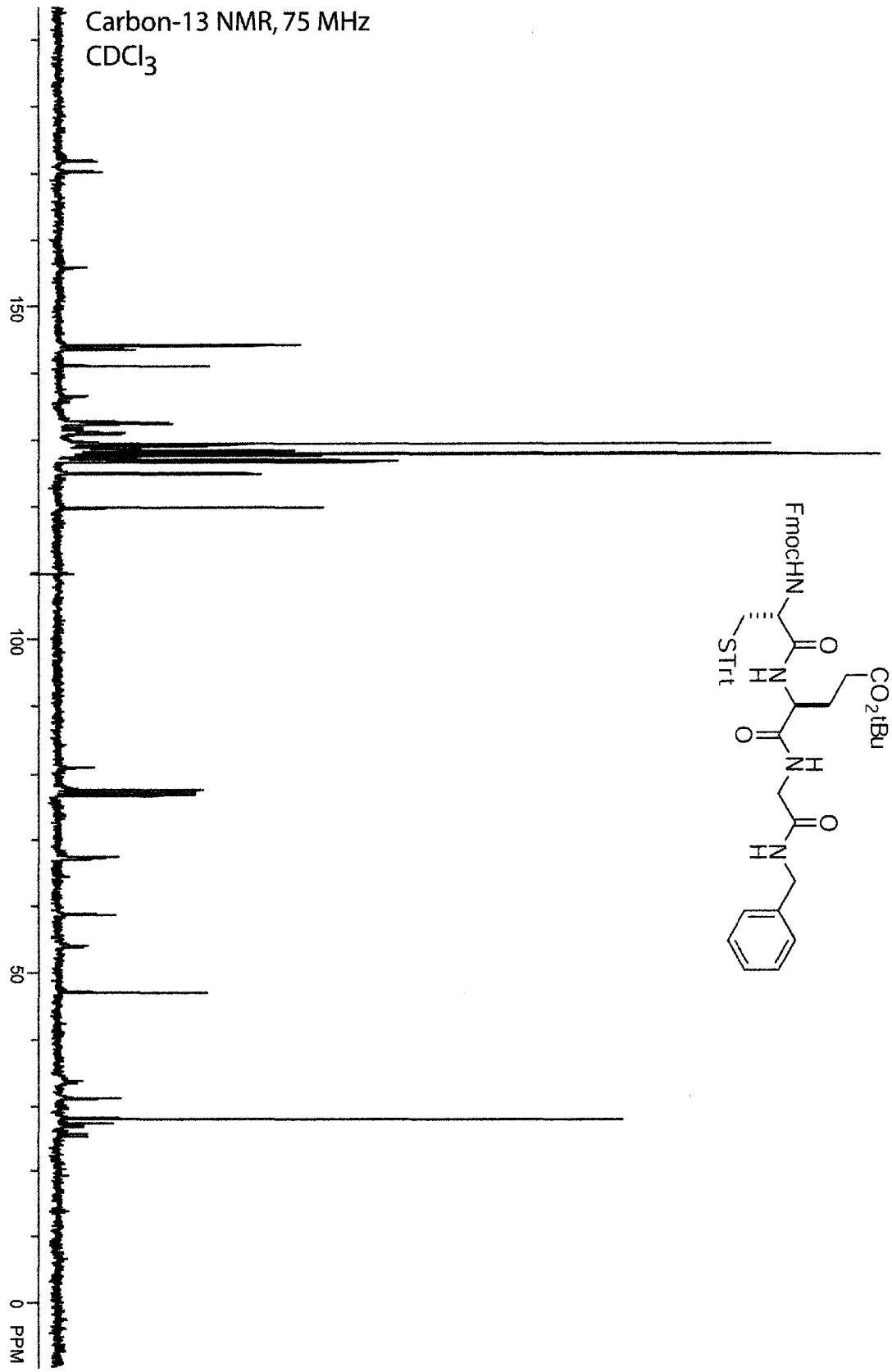
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 5

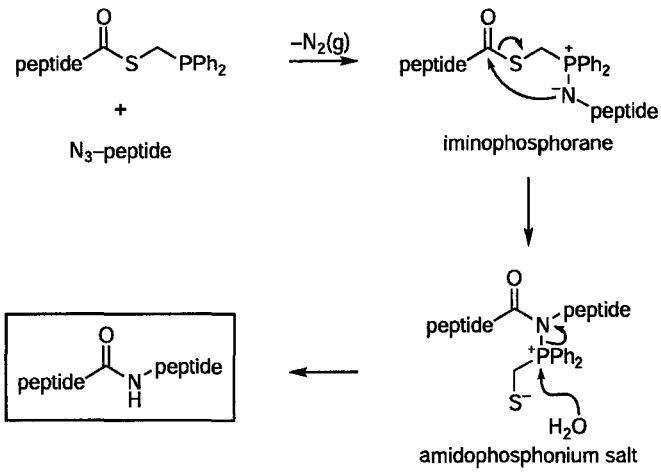
PREPARATION OF A CARBOXYL-FUNTIONALIZED STAUDINGER LIGATION REAGENT

Total chemical synthesis is beginning to provide ready access to natural proteins as well as enable the creation of nonnatural ones. Many proteins have already been assembled from synthetic peptides. "Native chemical ligation"—the coupling of a peptide (or protein) containing a *C*-terminal thioester with another peptide containing an *N*-terminal cysteine residue—has been especially powerful (Dawson & Kent, 2000; Kent, 2003). "Expressed protein ligation" is a method by which *C*-terminal thioesters for native chemical ligation can be accessed using recombinant DNA techniques (Muir, 2003). The thioester fragment is expressed in a heterologous host as part of a fusion protein consisting of the target fragment, a modified intein splicing element, and a chitin binding domain. The target fragment can then be cleaved as a thioester from a chitin stationary phase by elution with a thiol buffer. This innovative method for protein thioester generation has greatly expanded the utility of native chemical ligation for protein semisynthesis.

Emerging strategies for protein assembly avoid the need for a cysteine residue at the ligation junction. The Staudinger ligation is one such strategy (Scheme 5.1). In one form of the Staudinger ligation, a peptide with a *C*-terminal phosphinothioester is coupled with a second peptide having an *N*-terminal azido acid through the intermediacy of an

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

iminophosphosphorane. The iminophosphorane can rearrange to give an amidophosphonium salt as an acyl transfer product that is hydrolyzed in the presence of water. The final ligation product contains a native amide bond without any residual atoms (Nilsson et al., 2000; Nilsson et al., 2001; Nilsson et al., 2003). This technique occurs without racemization (Soellner et al., 2002). The Staudinger ligation has been used in the orthogonal assembly of a protein (RNase A) (Nilsson et al., 2003) and for the site-specific immobilization of peptides to a surface (Soellner et al., 2003). Different reagents and methods have also been explored in applying the Staudinger ligation to peptide synthesis and *in vivo* labeling experiments (Saxon et al., 2000; Saxon & Bertozzi, 2000a; Saxon et al., 2002; Vocadlo et al., 2003)



Scheme 5.1 *Staudinger ligation.*

We have envisioned that a phosphinothiol reagent incorporating carboxyl-functionality in at least one of the side chains would expand the utility of the Staudinger ligation. Such a funtionalization of a phosphinothiol should impart water solubility and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

allow the use of the resulting phosphinothiol for the generation of protein fragment phosphinothioesters in expressed protein ligation systems. In addition, a carboxyl group could be used as a handle for immobilization of phosphinothiols to solid support for the solid-phase synthesis of peptide thioesters. Such a phosphinothiol-functionalized solid support could represent a new form a "safety-catch" resin in which cleavage from the resin occurs upon addition of azide via a Staudinger ligation.

We have considered several different carboxyl-functionalized phosphinothiols as targets (Figure 5.1). Compound **5.1** represents a monocarboxyl reagent that could be used ideally for immobilization to a solid support without orthogonal protection of a separate carboxyl group. This compound is the only member of the series possessing chirality at phosphorus. When coupled with a chiral amino acid, it should be possible to separate the resulting diastereomers and it may be interesting to evaluate how chirality at phosphorus affects the outcome of the Staudinger ligation, particularly at difficult couplings. Compounds **5.2-4** are symmetrical reagents that vary in substitution about the phosphorus with groups that affect the electron density of the phosphorus atom.

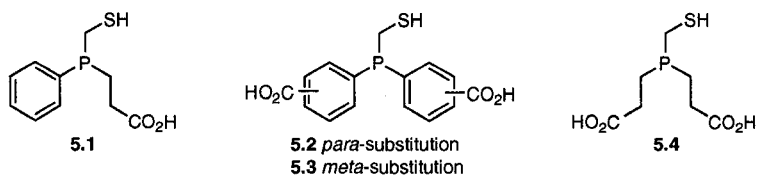
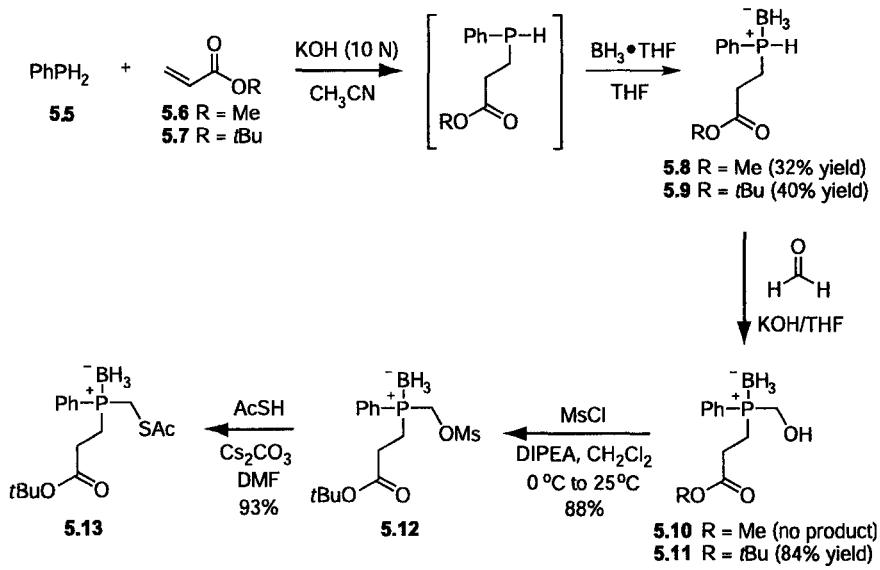


Figure 5.1 *Target carboxyl-functionalized phosphinothiol reagents.*

Compound **5.1** was chosen as the first target for synthesis (Scheme 5.2). The borane protected bisphosphines **5.8** and **5.9** were synthesized by the addition of phenyl

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phosphine to methyl acrylate and *tert*-butyl acrylate repectively (Rampal et al., 1981). The resultant bisphosphines were not isolated due to their reactive nature, but were diluted in THF and treated with borane to give the desired stable borane-bisphosphine adducts. Compounds **5.8** and **5.9** were isolated in 32% and 40% yields following purification by silica gel chromatography. Our original form of this phosphinothiol (Nilsson et al., 2001) was synthesized by addition of diphenylphosphine-borane to the alkylating agent $BrCH_2SAc$ to directly give a protected form of the desired reagent. Treatment of phosphines **5.8** or **5.9** with the aforementioned alkylating agent failed to give desired product (compound **5.13** or the methyl ester form). Instead, a more indirect route to the protected intermediate was taken.



Scheme 5.2 *Synthesis of a protected form of phosphinothiol* **5.13**.

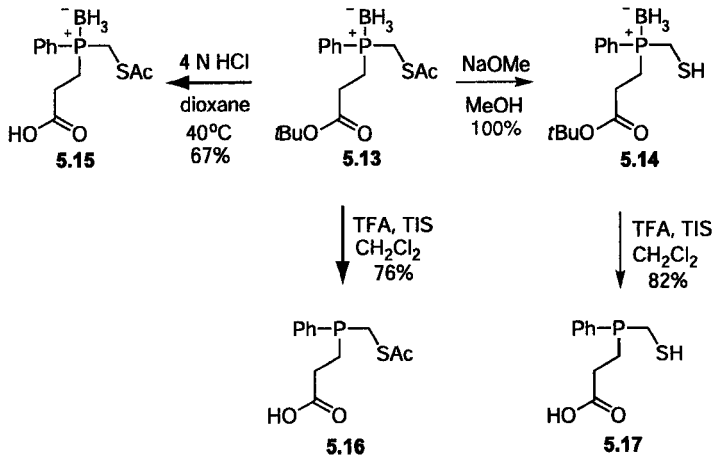
The protected intermediate **5.13** was synthesized through the intermediacy of an alcohol, which was converted to a leaving group and displaced with thioacetic acid (He et

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

al.). Compounds **5.8** and **5.9** were added to a solution of aqueous formaldehyde, potassium hydroxide, and THF. Phosphine **5.9** successfully added to formaldehyde to give the phosphine alcohol **5.11** in 84% yield. The methyl ester of phosphine **5.8** did not survive the basic conditions of this reaction, and no product **5.10** was obtained. The alcohol of compound **5.11** was converted to the mesylate in 88% using standard conditions. Addition of the cesium salt of thioacetic acid to mesylate **5.12** gave the desired protected phosphine **5.13** in 93% yield. This protected form of the desired phosphine is stable to storage for several weeks without decomposition.

We next examined conditions for the deprotection of phosphine **5.13** (Scheme 5.3). Ideally, we hoped to be able to find conditions to orthogonally remove the protecting groups individually. Treatment of phosphine **5.13** with 4 N HCl in dioxane at 50 °C for 2 h gave a mixture of compounds **5.15** and **5.16**. The two compounds could not be separated chromatographically. Treatment of this mixture of compounds with mCPBA oxidized the unprotected phosphine **5.16** and allowed for clean chromatographic separation from compound **5.15**. Using these conditions, compound **5.15** was obtained in 67% yield. Treatment of **5.13** with TFA in methylene chloride using triisopropylsilane as a cation scavenger resulted in the clean deprotection of both the *tert*-butyl ester and the borane complex of the phosphine to give compound **5.16** in 76% yield. The acetate could be removed quantitatively using sodium methoxide in methanol to give compound **5.14**. Treatment of **5.14** with TFA and triisopropylsilane in methylene chloride again removed the *tert*-butyl ester and the borane to give the completely deprotected carboxyl-functionalized phosphinothiol **5.17**. The yield of the final deprotection step was 82%.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



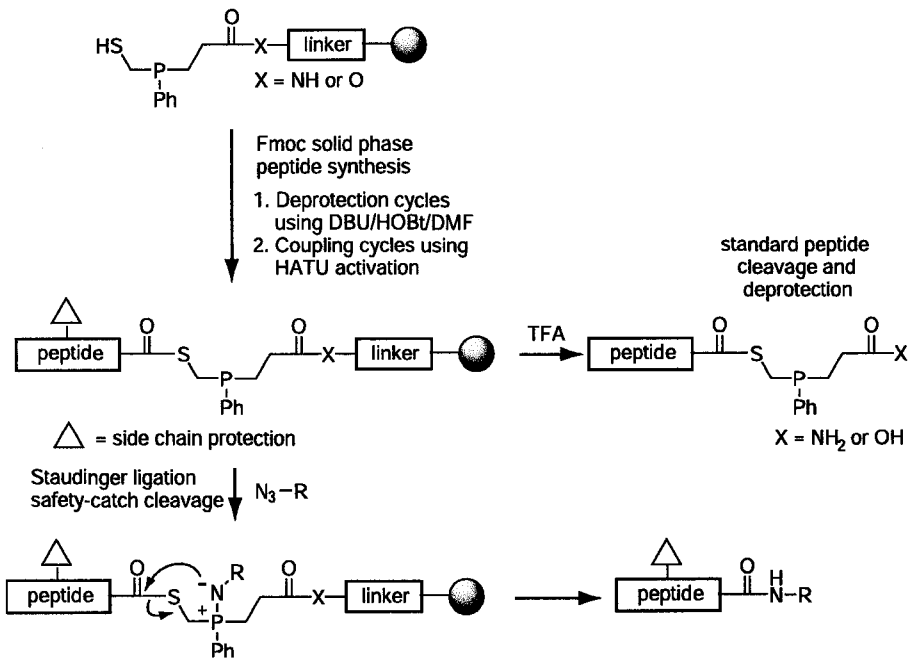
Scheme 5.3 *Deprotection of phosphinothiol* **5.13**.

We next determined the water solubility of the completely deprotected phosphinothiol **5.17**. Compound **5.17** was added to 100 mM potassium phosphate buffer, pH 7.6. These are the buffer conditions for the liberation of *C*-terminal thioesters from fusion proteins on a chitin stationary phase during a standard expressed protein ligation protocol (Muir et al., 1998). A saturated solution of **5.17** was made in buffer as evidenced by the formation of a small amount of white precipitate. This precipitate was removed by centrifugation and the concentration of **5.17** in solution was determined by assaying spectrophotometrically with Ellman's reagent (5,5'-dithiobis(2-nitrobenzoic acid) or DTNB), which reacts with both the phophine and thiol groups. The concentration of phosphinothiol **5.17** was determined to be 40 mM in these buffer conditions. During expressed protein ligation, thiol concentrations of 30–50 mM are usually used. Thus, compound **5.17** should be applicable to the creation of *C*-terminal phosphinothioesters using expressed protein ligation protocols.

A second application that we intend to explore for compound **5.17** is its use in solid-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phase peptide synthesis (Scheme 5.4). The carboxyl group should be an efficient handle for immobilization to resins or other solid supports. It should be possible to use existing conditions to synthesize phosphinothioesters on the solid phase with resin bearing this compound (Bu et al., 2002a; Bu et al., 2002b). In addition, it should be possible to cleave peptides synthesized on these types of resins by the addition of organic azides. These types of resins would effectively be "safety-catch" resins (Kenner et al., 1971).



Scheme 5.4 *Solid-phase peptide synthesis applications for carboxyl-functionalized phosphinothiol.*

We have also explored the synthesis of compounds **5.2–4**. A number of different strategies have been employed, but none have been successful to date. Efforts continue toward the synthesis of these additional bis-carboxyl-functionalized phosphinothiols.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Experimental Section

General Experimental

Chemicals and solvents were purchased from Aldrich® with the exception of dry THF, CH_2Cl_2, and DMF which were purchased from VWR®. Reactions were monitored by thin layer chromatography using Whatman® TLC plates (AL SIL G/UV) and visualized by UV or I_2. NMR spectra were obtained using Bruker AC-300 or Varian UNITY-500 spectrometers. Phosphorus-31 NMR spectra were proton-coupled and referenced against an external standard of deuterated phosphoric acid unless otherwise stated. Mass spectra were obtained using electrospray ionization (ESI) techniques.

3-Phenylphosphanyl-propionic acid methyl ester-borane complex (5.8). Phenylphosphine (**5.5**, 10 g, 0.09 mol) was charged to a flame-dried round bottom flask under an argon atmosphere and dissolved in acetonitrile (10 mL) that had been degassed under vacuum and then bubbled with argon. KOH (10N, 1 mL) was added, and the resulting solution was cooled to 0 °C in an ice bath. Methyl acrylate (**5.6**, 0.09 mol, 8.1 mL) was then added at a rate that kept the internal reaction temperature below 35 °C. Upon complete addition of the methyl acrylate, the solution was stirred at 0 °C for 30 min and then heated at 50 °C for 8 h. After 8 h, the solution was diluted with dry THF (150 mL), and the resulting solution was cooled to 0 °C. Borane-THF (1.0 M in THF, 91 mmol, 91 mL) was added dropwise over 10 min. The reaction was allowed to warm to room temperature and then stirred for 10 h. The solvent was then removed under reduced pressure, and the residue was purified by chromatography (silica gel, 35% v/v ethyl acetate in hexanes) to give compound **5.8** as a colorless oil in 32% yield (6.00 g, 28.6

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ7.74-7.67 (m, 2 H), 7.55-7.44 (m, 3 H), 5.6 (dm, J_{HP} = 376.9 Hz, 1 H), 3.66 (s, 3 H), 2.72-2.45 (m, 2 H), 2.40-2.10 (m, 2 H), 1.25-0.33 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ173.84 (d, J = 12.4 Hz), 132.57 (d, J = 8.4 Hz), 131.68 (d, J = 2.2 Hz), 128.86 (d, J = 10.8 Hz), 124.44 (d, J = 56.2 Hz), 51.78, 28.09, 18.58 (d, J = 36.6 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ-4.08 (d, J = 383.7 Hz) ppm; MS (ESI) *m/z* 233.0885 (MNa^+ [$C_{10}H_{16}BO_2PNa$] = 233.0879).

3-Phenylphosphanyl-propionic acid *tert*-butyl ester-borane complex (5.9). Phenylphosphine (**5.5**, 10 g, 0.09 mol) was charged to a flame-dried round bottom flask under an argon atmosphere and dissolved in acetonitrile (10 mL) that had been degassed under vacuum and then bubbled with argon. KOH (10 N, 1 mL) was then added, and the resulting solution was cooled to 0 °C in an ice bath. *tert*-Butylacrylate (**5.7**, 0.09 mol, 13.3 mL) was added at a rate that kept the reaction temperature below 35 °C. Upon complete addition of the *tert*-butylacrylate, the solution was stirred at 0 °C for 30 min and then heated at 50 °C for 8 h. The solution was then diluted with dry THF (150 mL), and the resulting solution was cooled to 0 °C. Borane-THF (1.0 M in THF, 91 mmol, 91 mL) was added dropwise over 10 min. The reaction was allowed to warm to room temperature, and was then stirred for 10 h. The organic solvent was removed under reduced pressure, and the residue was purified by chromatography (silica gel, linear gradient of 0–20% v/v ethyl acetate in hexanes) to give compound **5.9** as a colorless oil in 40% yield (9.09 g, 36 mmol).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74-7.67 (m, 2 H), 7.54-7.44 (m, 3 H), 5.59 (dm, J_{HP} = 375.3 Hz), 2.64-2.40 (m, 2 H), 2.38-2.04 (m, 2 H), 1.44 (s, 9 H), 1.21-0.35 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ170.70 (d, J = 11.4 Hz), 132.61 (d, J = 9 Hz), 131.63 (d, J = 2.8 Hz), 128.86 (d, J = 9.6 Hz), 124.72 (d, J = 55.6 Hz), 81.00, 29.41, 27.76, 18.67 (d, J = 37.9 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ -5.47 (d, J = 365.3 Hz) ppm; MS (ESI) *m/z* 275.1354 (MNa^+ [$C_{13}H_{22}BO_2PNa$] = 275.1348).

3-(Hydroxymethyl-phenyl-phosphanyl)-propionic acid tert-butyl ester-borane complex (5.11). Bisphosphine borane **5.9** (7.28 g, 29 mmol) was charged to a flame-dried round bottom flask equipped with a stir bar under an argon environment, and dissolved in THF (23 mL). Formaldahyde (37% in water, 16.4 mL, 202 mmol) was then added followed by potassium hydroxide (1.62 g, 29 mmol). The reaction mixture was stirred at room temperature under argon for 4 h. The organic solvent was then concentrated under reduced pressure. The residual water was diluted with H_2O (75 mL) and washed 3 times with 80 mL of ethyl acetate. The combined organic layers were washed with 2 N HCl (2 $\times$ 25 mL) and once with brine (25 mL). The organic layer was dried over anhydrous magnesium sulfate, which was then removed by filtration. The ethyl acetate was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 20% v/v ethyl acetate in hexanes). Compound **5.11** (6.85 g, 24.3 mmol) was isolated as a colorless oil in 84% yield (6.85g, 24.3 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.85-7.79 (m, 2 H), 7.56-7.44 (m, 3 H), 4.10 (s, 2 H), 3.41 (bs, 1 H), 2.76-2.61 (m, 1 H), 2.54-2.35 (m, 2 H), 2.17-2.02 (m, 1 H),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.45 (s, 9 H), 0.99-0.13 (bm, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ172.45 (d, J = 11.4 Hz), 132.21 (d, J = 8.3 Hz), 131.67 (d, J = 2.6 Hz), 126.49 (d, J = 54 Hz), 81.58, 59.52 (d, J = 39.5 Hz), 28.57, 27.81, 16.48 (d, J = 37.3 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ16.25 ppm; MS (ESI) *m/z* 305.1451 (MNa^+ [$C_{14}H_{24}BO_3PNa$] = 305.1454).

3-(Methanesulfonyloxymethyl-phenyl-phosphanyl)-propionic acid tert-butyl ester-borane complex (5.12). Compound **5.11** (6.85 g, 24.3 mmol) was charged to a flame dried round bottom flask under argon and is dissolved in dry methylene chloride (65 mL). This solution was cooled to 0 °C. Diisopropylethyl amine (6.34 mL, 36.4 mmol) was then added, followed slowly by methanesulfonyl chloride (2.64 mL, 34.0 mmol). The reaction was allowed to warm slowly to room temperature after which it was stirred for 12 hours at ambient temperature. The reaction mixture was then washed twice with water (2 × 10 mL), 2 N HCl (2 × 10 mL), and again water (10 mL). The organic layer was dried over anhydrous magnesium sulfate, filtered, and concentrated *en vacuo*. The crude residue was purified by flash chromatography (silica gel, 35% v/v ethyl acetate in hexanes) to give the desired compound **5.12** (7.72 g, 21.4 mmol) as a yellow oil in 88% yield (7.72 g, 21.4 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.85-7.79 (m, 2 H), 7.64-7.51 (m, 3 H), 4.71-4.55 (m, 2 H), 3.01 (s, 3 H), 2.66-2.55 (m, 1 H), 2.46-2.31 (m, 3 H), 1.44 (s, 9 H), 1.04-0.25 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ170.86 (d, J = 15.9 Hz), 132.61, 132.54 (d, J = 8.5 Hz), 129.09 (d, J = 9.8 Hz), 123.97 (d, J = 53.5 Hz), 81.36, 64.52 (d, J = 36.2 Hz), 37.19, 28.27, 27.90, 16.7 (d, J = 38.7 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

δ16.63 ppm; MS (ESI) *m/z* 383.1244 (MNa^+ [$C_{15}H_{26}BO_5PSNa$] = 383.1229).

3-(Acetylsulfanylmethyl-phenyl-phosphanyl)-propionic acid tert-butyl ester – borane complex (5.13). Compound **5.12** (7.72 g, 21.4 mmol) was charged to a flame dried round bottom flask equipped with a stir bar under an argon environment and dissolved in dry DMF (50 mL). This solution was cooled to 0 °C. Thiolacetic acid (1.74 mL, 24.4 mmol) was then added. Cesium carbonate (6.98 g, 21.4 mmol) was added, and the reaction mixture, a slurry, was stirred at 0 °C for 15 min and then allowed to warm slowly to room temperature. The reaction mixture was stirred for 12 h at room temperature and then filtered to remove the solid, and the organic layer was concentrated under reduced pressure. The crude residue was purified by flash chromatography (silica gel, 20% v/v ethyl acetate in hexanes). Compound **5.13** was isolated as a pale yellow oil in 93% yield (6.76 g, 19.9 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.81-7.75 (m, 2 H), 7.59-7.47 (m, 3 H), 3.41 (d, J = 6.1 Hz, 2 H), 2.61-2.46 (m, 1 H), 2.33 (s, 3 H), 2.36-2.22 (m, 3 H), 1.43 (s, 9 H), 1.14-0.30 (m, 3 H)192.86, 170.94 (d, J = 17.3 Hz), 132.06 (d, J = 11.1 Hz), 128.72 (d, J = 10.1 Hz), 125.88 (d, J = 52.5 Hz), 80.98, 29.91, 28.46 (d, J = 2.2 Hz), 27.81, 23.65 (d, J = 32.2 Hz), 18.89 (d, J = 36.7 Hz) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 192.84, 170.94 (d, J = 16.6 Hz), 132.08 (d, J = 8.7 Hz), 132.01, 128.73 (d, J = 9.9 Hz), 125.87 (d, J = 53.4 Hz), 80.97, 29.92, 28.46 (d, J = 2.3 Hz), 27.81, 23.66 (d, J = 32.3 Hz), 18.87 (d, J = 37.4 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ 18.11 ppm; MS (ESI) *m/z* 363.1344 (MNa^+ [$C_{16}H_{26}BO_3PSNa$] = 363.1331).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3-(Mercaptomethyl-phenyl-phosphanyl)-propionic acid tert-butyl ester – borane complex (5.14). Compound **5.13** (500 mg, 1.47 mmol) was dissolved in methanol (13 mL) that had been deoxygenated (first by vacuum, then by bubbling with argon) under an argon atmosphere. Sodium methoxide (80 mg, 1.47 mmol) was added, and the reaction mixture was stirred under argon at room temperature and the reaction progress was monitored by TLC. After 30 min, approximately half of the starting material was consumed. Another equivalent of sodium methoxide was then added (80 mg, 1.47 mmol), and the reaction was complete after and 30 min. The reaction mixture was quenched by the addition of potassium phosphate buffer (0.25 M, pH 6.2, 10 mL). The reaction solution was then extracted with ethyl acetate (3 × 10 mL). The combined organic layers were dried over anhydrous magnesium sulfate, filtered, and concentrated under reduced pressure. The desired free thiol compound **5.14** was isolated in quantitative yield and good purity without any additional purification.

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.81-7.74 (m, 2 H), 7.57-7.47 (m, 3 H), 2.92-2.87 (m, 2 H), 2.64-2.52 (m, 1 H), 2.41-2.24 (m, 3 H), 1.82-1.74 (m, 1 H), 1.43 (s, 9 H), 1.14-0.27 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ171.07 (d, J = 15.4 Hz), 132.09 (d, J = 8.8 Hz), 131.90 (d, J = 4.2 Hz), 128.82 (d, J = 10.2 Hz), 125.95 (d, J = 53.2 Hz), 81.00, 28.57, 19.55 (d, J = 31.1 Hz), 18.34 (d, J = 38.1 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ 20.83 ppm; MS (ESI) *m/z* 321.1240 (MNa^+ [$C_{14}H_{24}BO_2PSNa$] = 321.1225).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3-(Acetylsulfanylmethyl-phenyl-phosphanyl)-propionic acid – borane complex (5.15). Compound **5.13** (500 mg, 1.47 mmol) was charged to a flame-dried round bottom flask in an argon atmosphere. A solution of 4 N HCl in dioxane (3 mL) was added, and the reaction mixture was stirred at room temperature for 3 h. After 3 h, TLC analysis indicated a significant amount of starting material remained. Additional 4 N HCl in dioxane (6 mL) was added, and the resulting solution was heated at 60 °C for an additional 1 h. At this point all the starting material was consumed by TLC. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 5% v/v methanol in methylene chloride). NMR analysis of the isolated material indicated a mixture of the desired compound **5.15** and the borane deprotected compound **5.16**. No solvent conditions were found that could efficiently separate the compounds by chromatography. The mixture was dissolved in methylene chloride (5 mL) and mCPBA (77%, 67 mg, 0.3 mmol) to oxidize the free phosphine, resulting in a mixture that could be purified by chromatography (silica gel, 5% v/v methanol in methylene chloride). Compound **5.15** was isolated as a colorless oil in 67% yield (280 mg, 0.98 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.94 (bs, 1 H), 7.79-7.23 (m, 2 H), 7.56-7.46 (m, 3 H), 3.39 (d, J = 6 Hz, 2 H), 2.74-2.61 (m, 1 H), 2.49-2.37 (m, 1 H), 2.35-2.23 (m, 2 H), 2.32 (s, 3 H), 1.06-0.27 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 193.27, 177.90 (d, J = 16.5 Hz), 132.19, 132.13 (d, J = 9.9 Hz), 128.92 (d, J = 9.9 Hz), 125.67 (d, J = 53.8 Hz), 30.03, 27.49, 23.74 (d, J = 33.3 Hz), 18.75 (d, J = 38.8 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ 18.22 ppm; MS (ESI) *m/z* 283.0733 (M - H^- [$C_{12}H_{17}BO_3PS^-$] = 283.0729).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3-(Acetylsulfanylmethyl-phenyl-phosphanyl)-propionic acid (5.16). Compound **5.13** (500 mg, 1.47 mmol) was charged to a flame-dried round bottom flask under argon and dissolved in dry methylene chloride (3 mL). Triisopropyl silane (0.753 mL, 3.67 mmol) was added, followed by trifluoroacetic acid (1.47 mL, 19.1 mmol). This reaction mixture was stirred at room temperature for 1 h. The solvent was then removed under reduced pressure. The residue was dissolved in ethyl acetate and purified by flash chromatography (silica gel, 40% v/v ethyl acetate in hexanes). Compound **5.16** was isolated as a colorless oil in 76% yield (300 mg, 1.11 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.79 (bs, 1 H), 7.54-7.48 (m, 2 H), 7.39-7.36 (m, 3 H), 3.24 (d, J = 4.2 Hz), 2.53-2.39 (m, 2 H), 2.30 (s, 3 H), 2.14-2.09 (m, 2 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 194.95, 178.57 (d, J = 13.3 Hz), 132.12 (d, J = 10 Hz), 129.61, 128.53 (d, J = 7.4 Hz), 30.17 (d, J = 18.9 Hz), 30.14, 25.39 (d, J = 26.1 Hz), 21.13 (d, J = 11.9 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ -23.46 ppm; MS (ESI) *m/z* 269.0399 (M - H$^-$ [$C_{12}H_{14}O_3PS^-$] = 269.0401).

3-(Mercaptomethyl-phenyl-phosphanyl)-propionic acid (5.17). Compound **5.14** (886 mg, 2.97 mmol) was charged to a flame-dried round bottom flask under argon and dissolved in dry methylene chloride (6 mL). Triisopropyl silane (1.21 mL, 5.94 mmol) was added, followed by the addition of trifluoracetic acid (2.98 mL, 38.6 mmol). The reaction was stirred at room temperature for 30 min after which solvent was removed under reduced pressure. The residue was purified by flash chromatography. Two columns were necessary to get pure product (silica gel, once in 35% v/v ethyl acetate in hexanes,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

once in 5% v/v methanol in methylene chloride). Compound **5.17** was isolated as a clear oil in 82% yield (556 mg, 2.4 mmol).

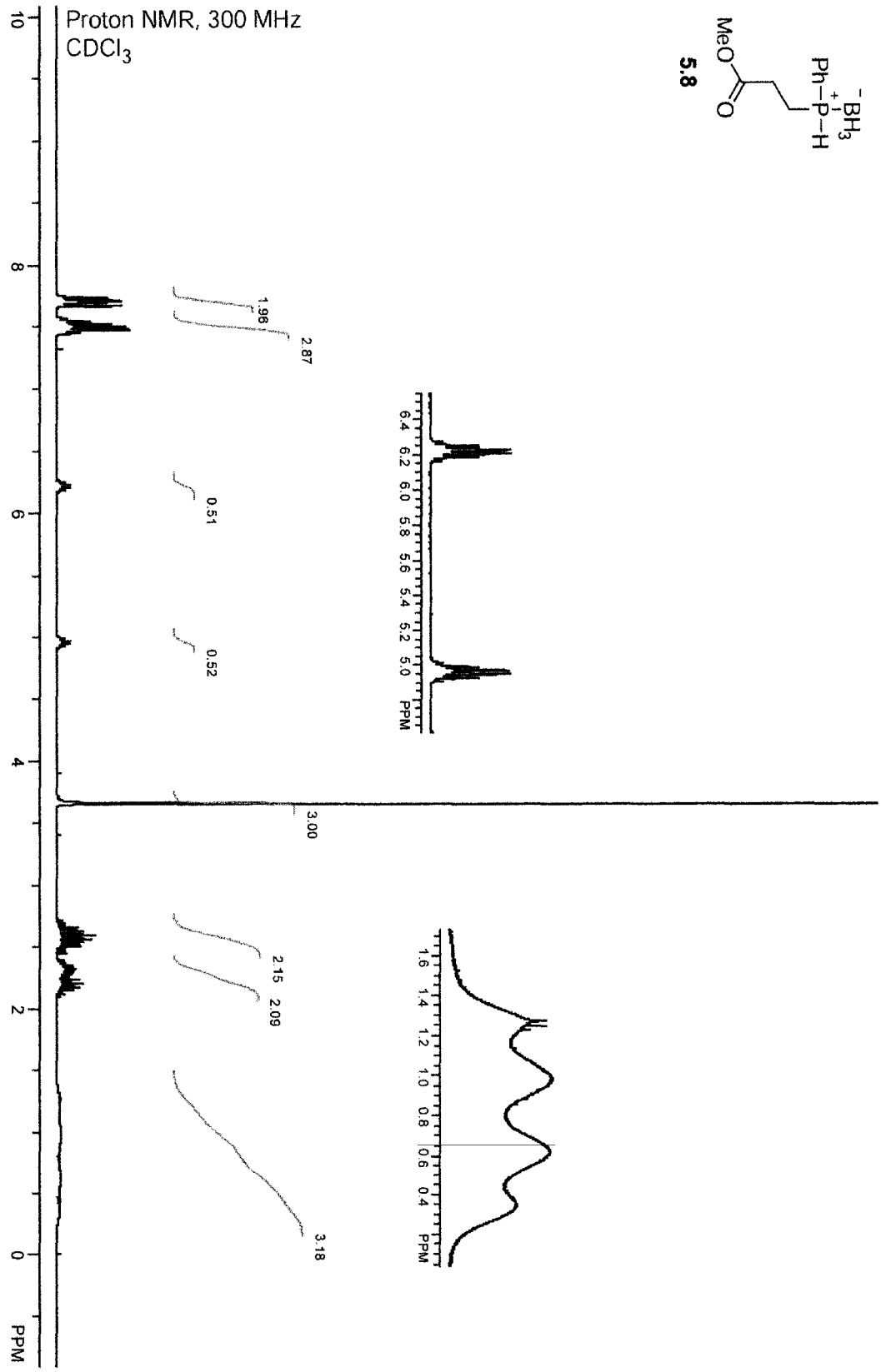
Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 11.42 (bs, 1 H), 7.53-7.46 (m, 2 H), 7.39-7.36 (m, 3 H), 2.75 (s, 2 H), 2.51-2.36 (m, 2 H), 2.22-2.07 (m, 2 H), 1.30 (s, 1 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 179.27 (d, J = 12.9 Hz), 135.23 (d, J = 14.6 Hz), 132.16 (d, J = 19.5 Hz), 129.46, 128.52 (d, J = 6.1 Hz), 30.15 (d, J = 19.7 Hz), 20.82, 20.73 (d, J = 39.4 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ -16.91 ppm; MS (ESI) *m/z* 227.0305 (M – H$^-$ [$C_{10}H_{12}O_2PS^-$] = 227.0296).

Determination of the water solubility of compound 5.17. A saturated solution of **5.17** in water (100 mM potassium phosphate buffer, pH 7.6) was prepared. Saturation was determined by the appearance of white precipitate, which was pelleted by centrifugation. The aqueous layer was then decanted and assumed to be saturated. The concentration of compound **5.17** in this solution was determined by using a DTNB (5,5'-dithiobis(2-nitrobenzoic acid) assay (Woycechowsky et al., 2003; Woycechowsky & Raines, 2003). DTNB is a disulfide that can be reduced by thiols and phosphines in aqueous solution. The formation of thionitrobenzoic acid (thiolate) can be monitored by UV spectroscopy at a wavelength of 412 nm (ϵ = 14150 $M^{-1}cm^{-1}$ at pH 7.6 (Han & Han, 1994). One equivalent of thiolate is formed for every equivalent of thiol in solution and two equivalents of thiolate are formed for every equivalent of phosphine in solution (Han & Han, 1994). The assay was performed by mixing 589 μL of 100 mM potassium phosphate, pH 7.6 with 10 μL of DTNB (5 mg/mL in the same buffer) in a cuvette. This solution was used to blank the spectrometer at 412 nm. Then, 1 μL of a 1:10 dilution of

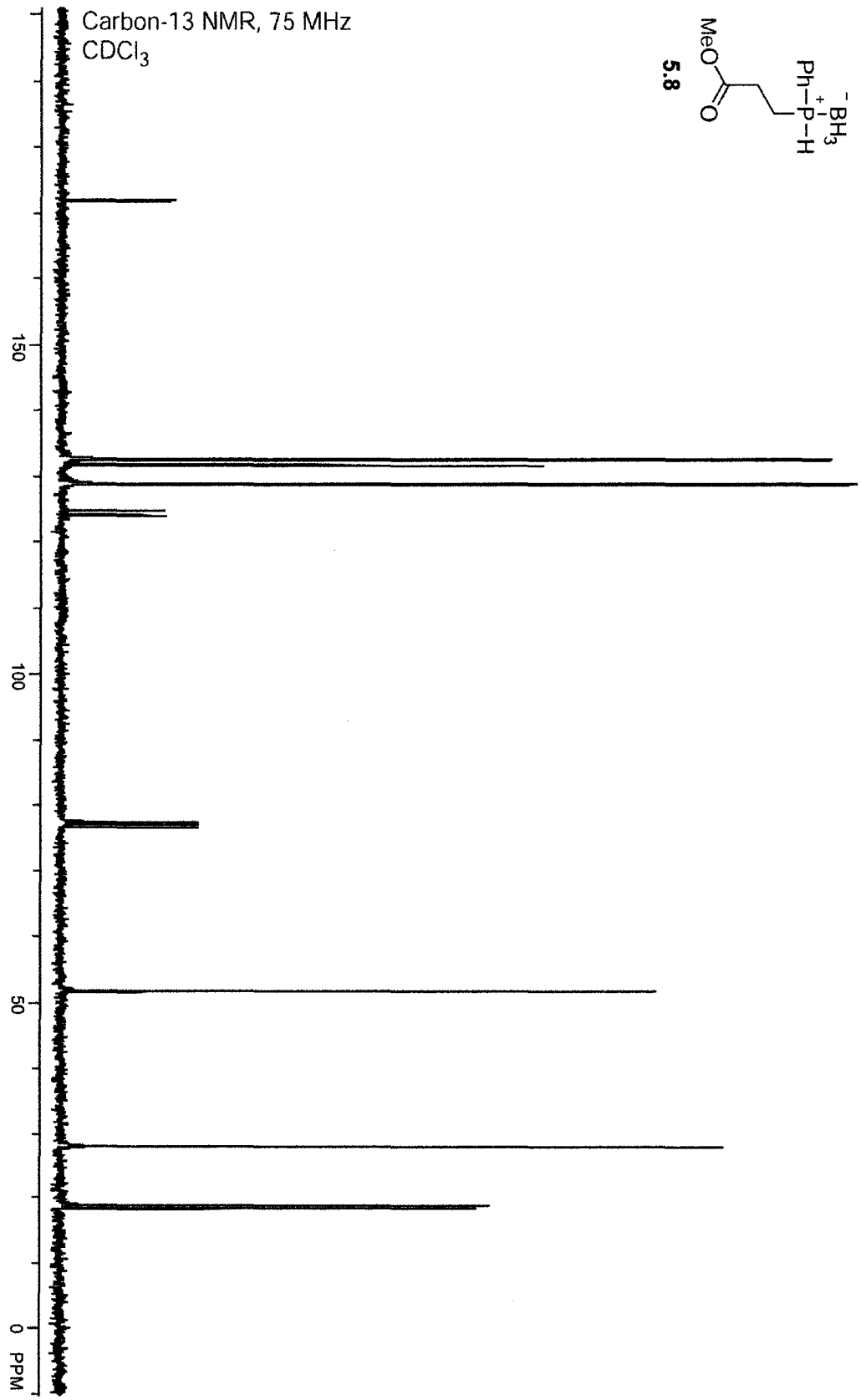
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the buffer solution of **5.17** was added to the cuvette, and the resulting solution was mixed by inversion. The absorbance at 412 nm (A_{412}) was observed until reaction was complete (approximately 5 minutes). The concentration of **5.17** was determined by the following equation [**5.17**] (mM) = $[(A_{412}/14150\ M^{-1}cm^{-1}) \times 600 \times 1000 \times 10]/3$. The average of three determinations gave a concentration of 39 mM.

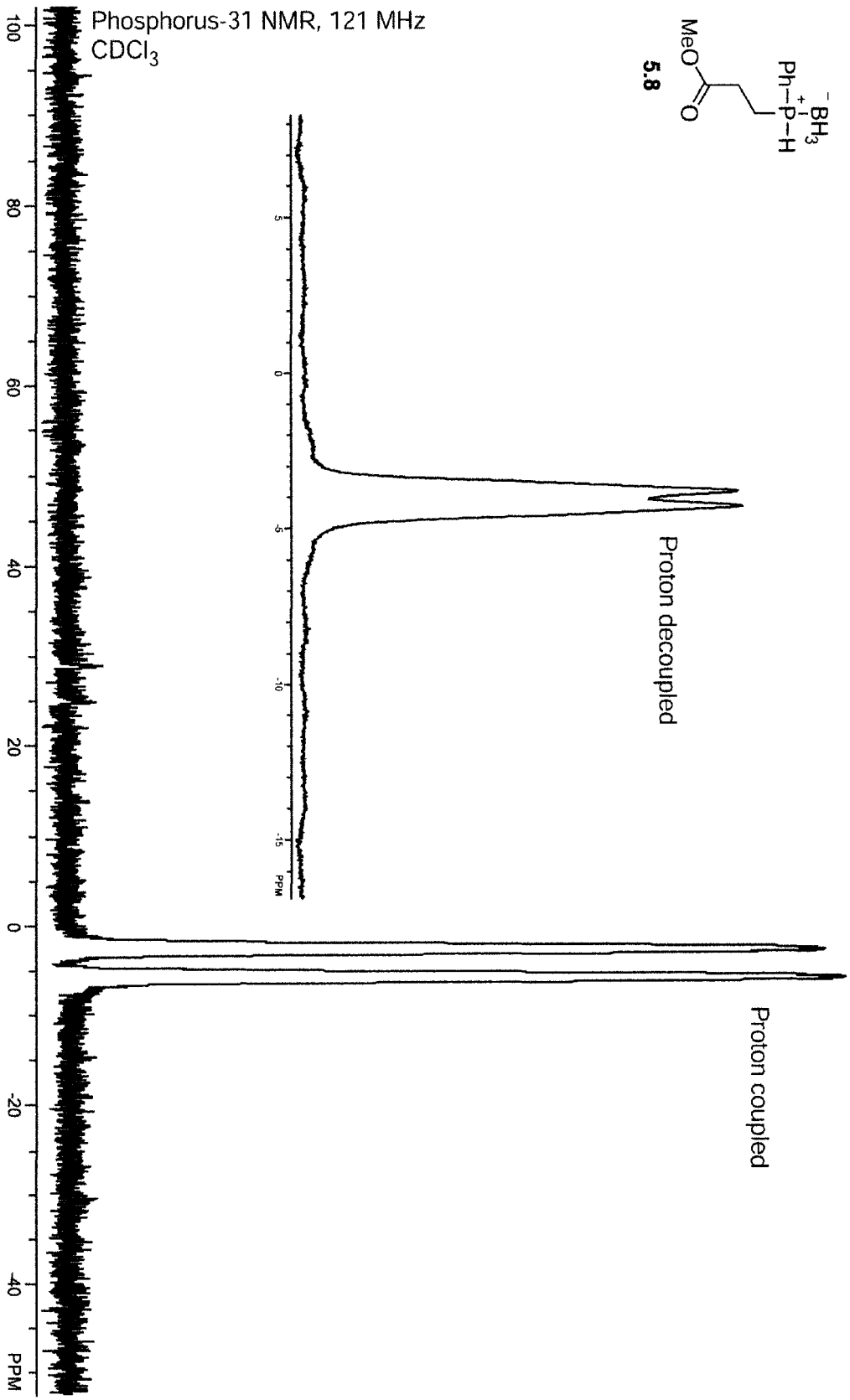
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



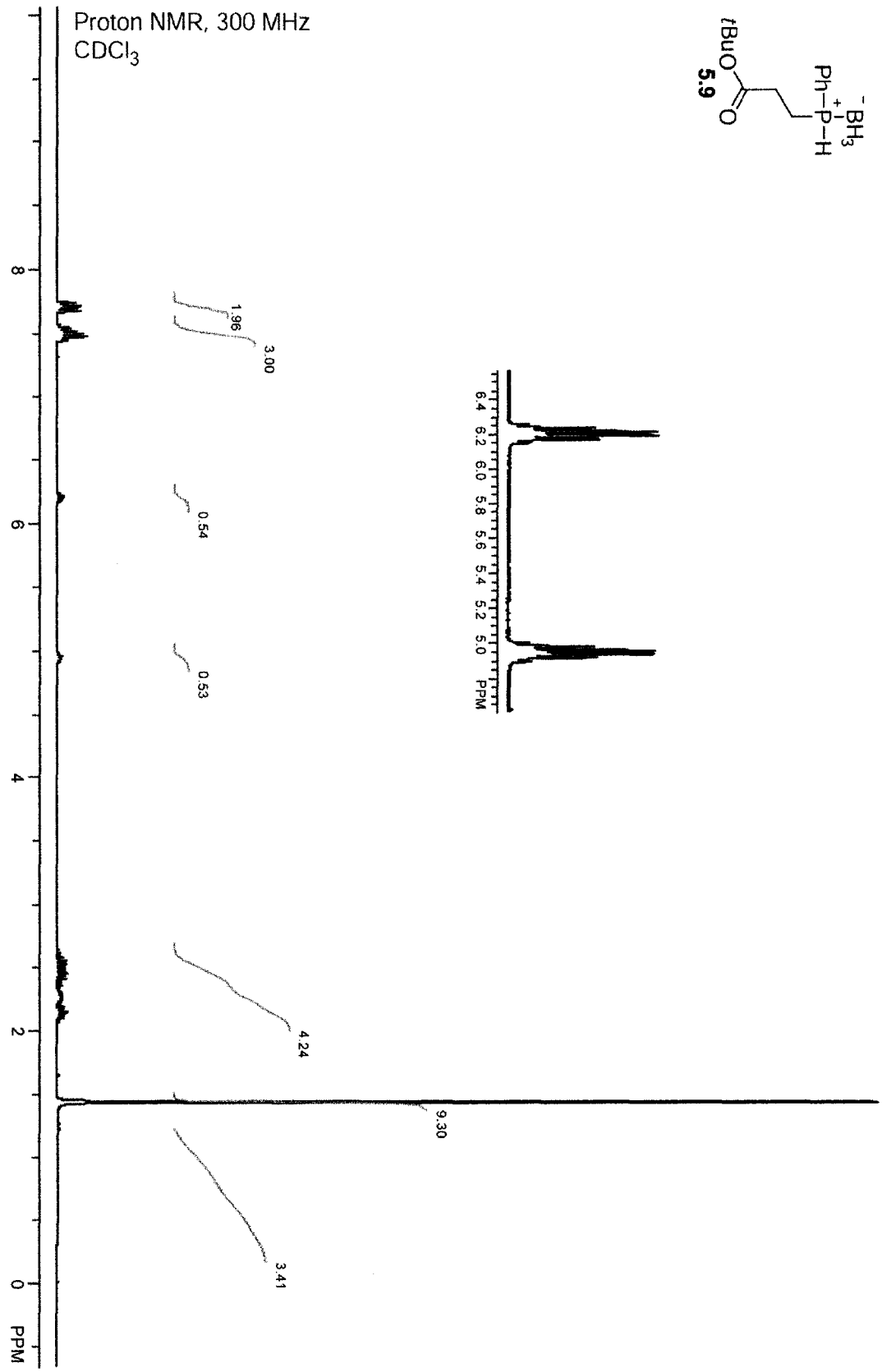
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



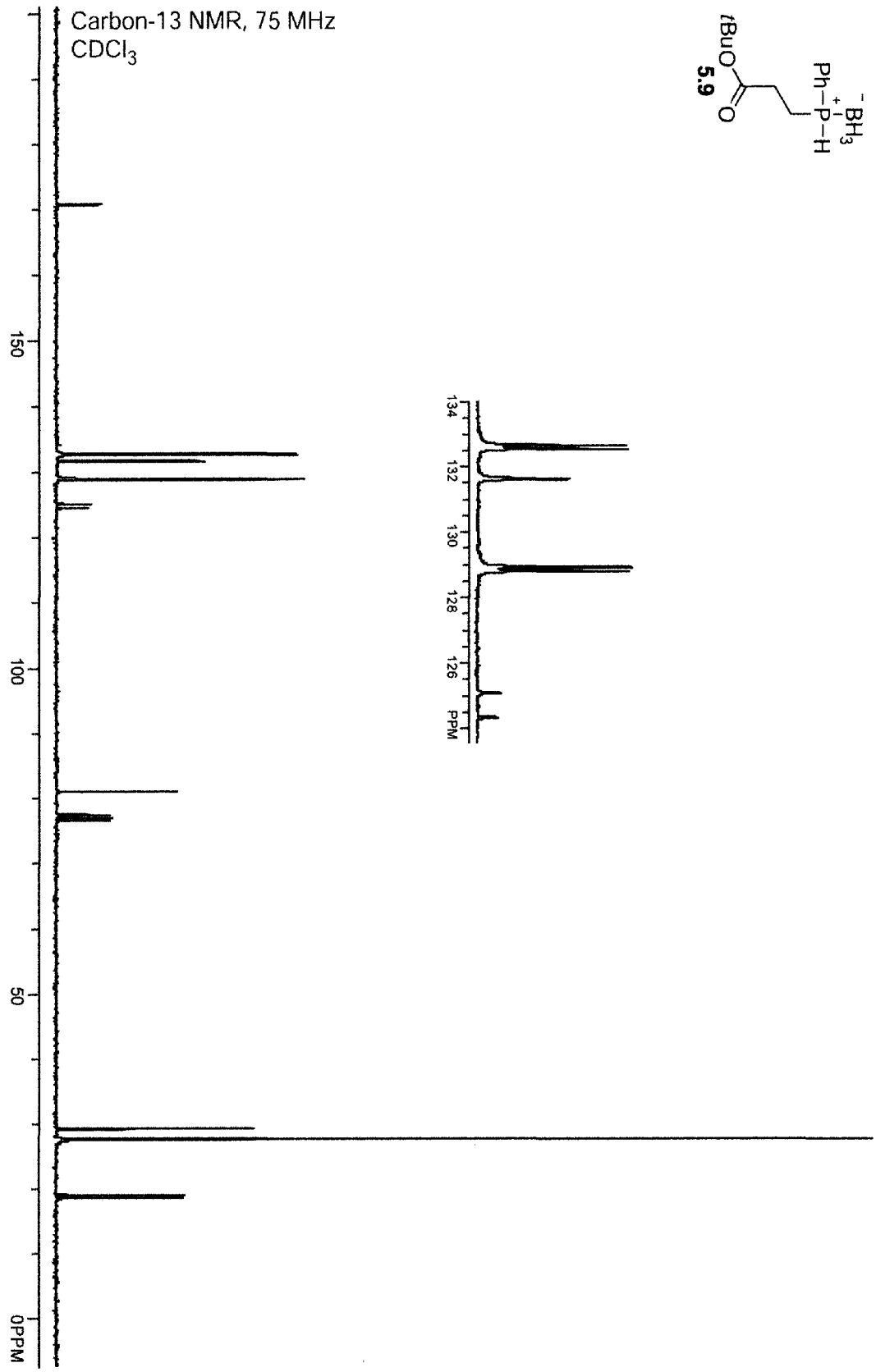
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



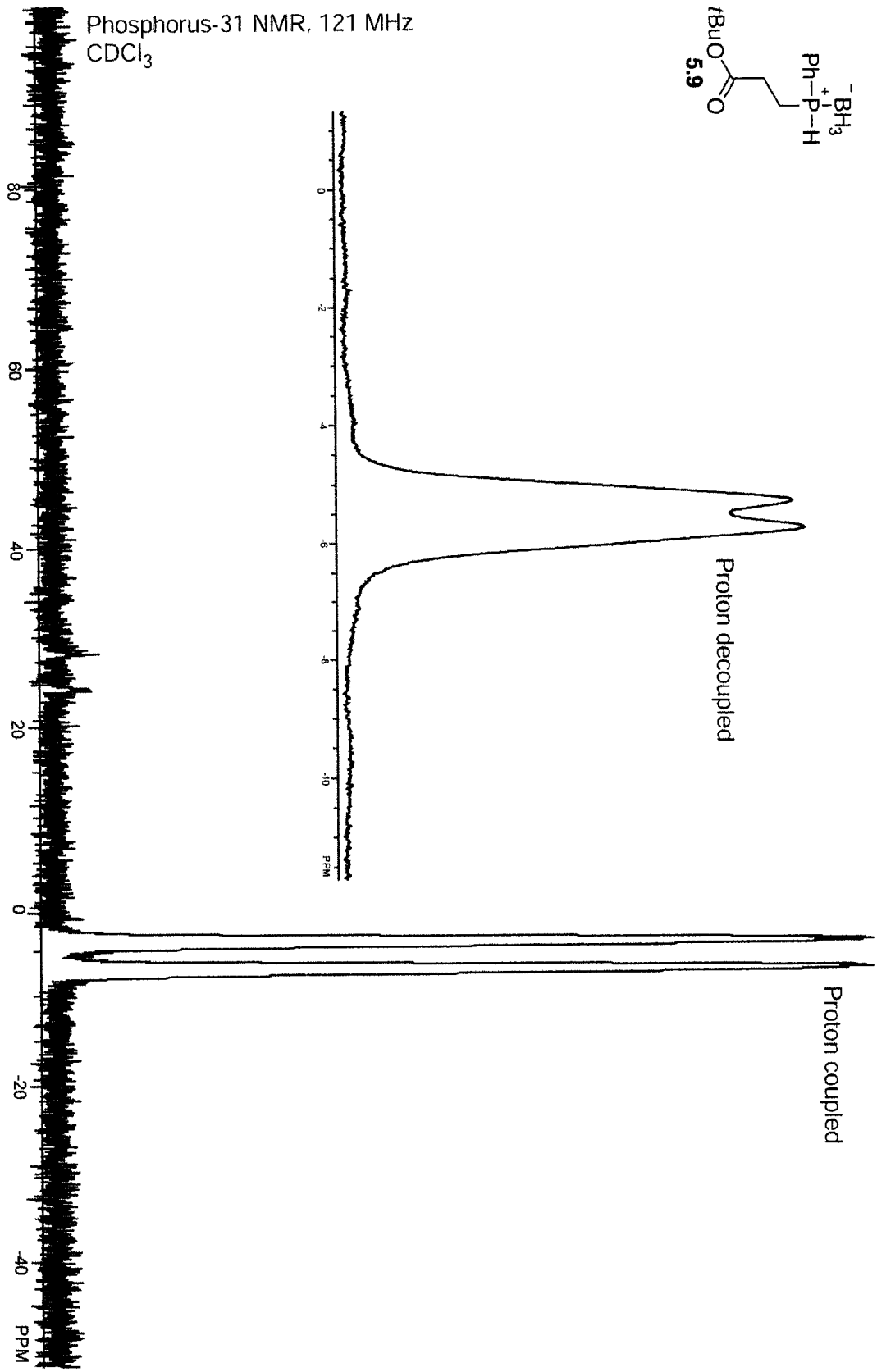
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



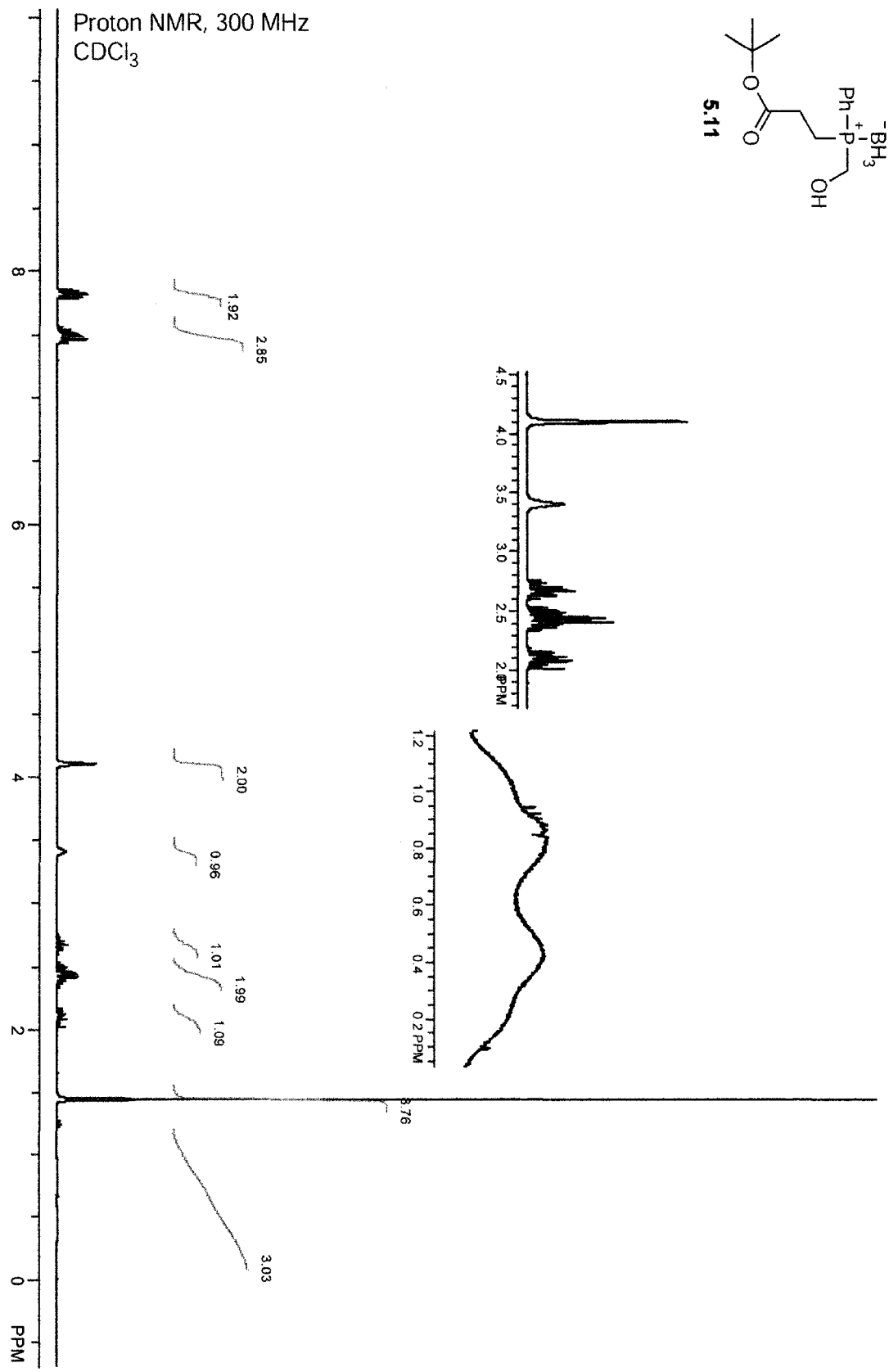
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



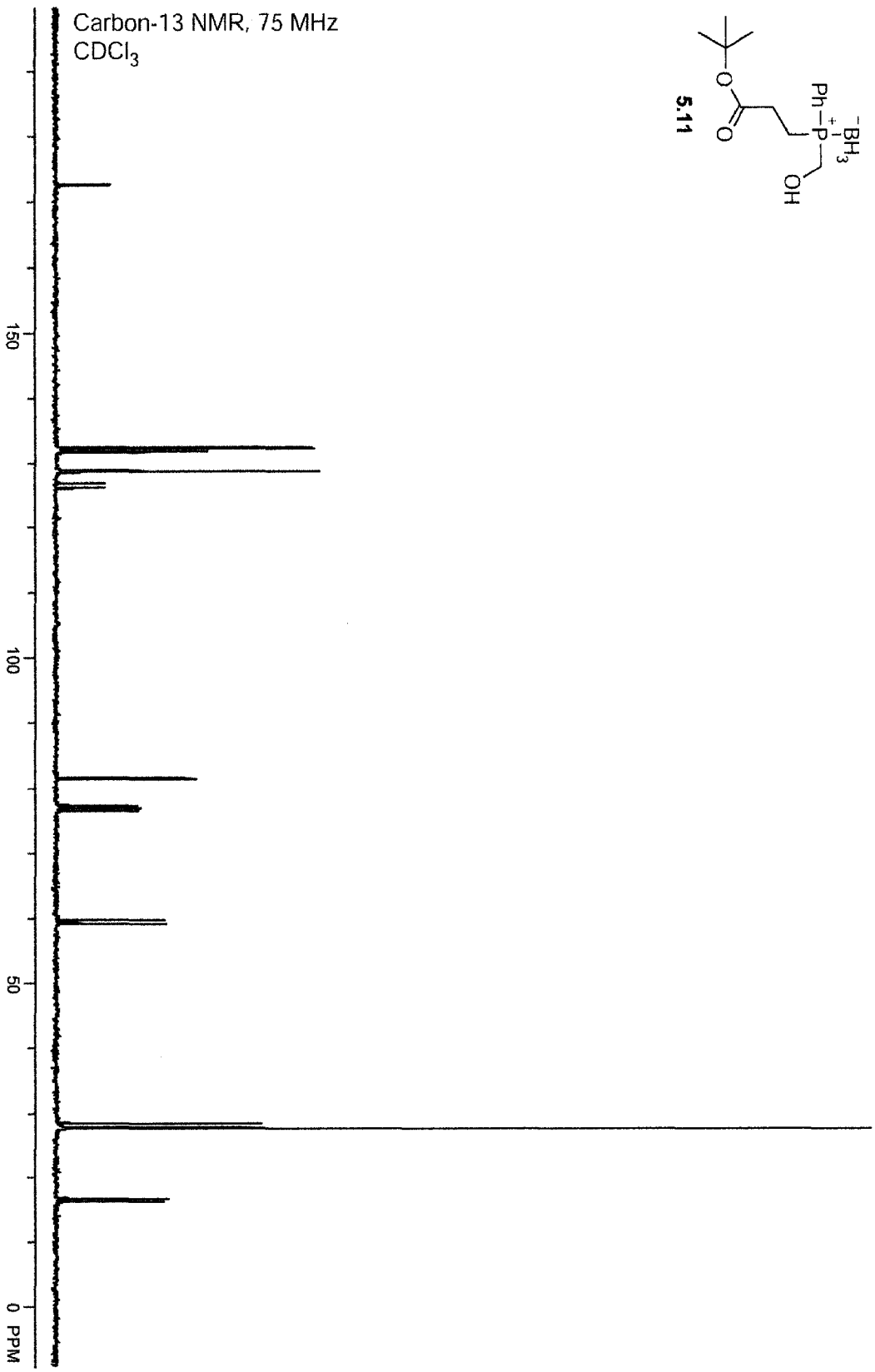
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



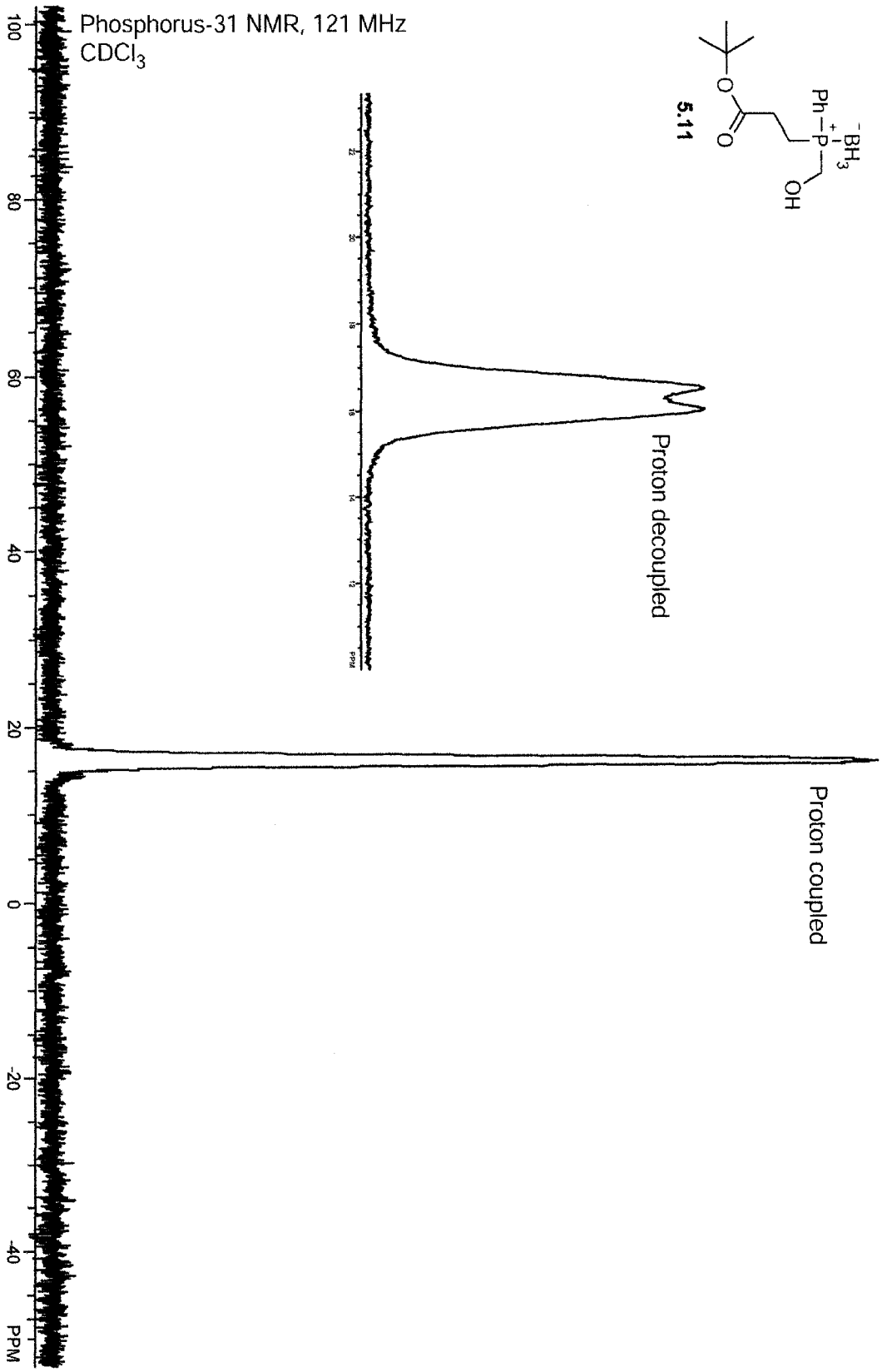
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



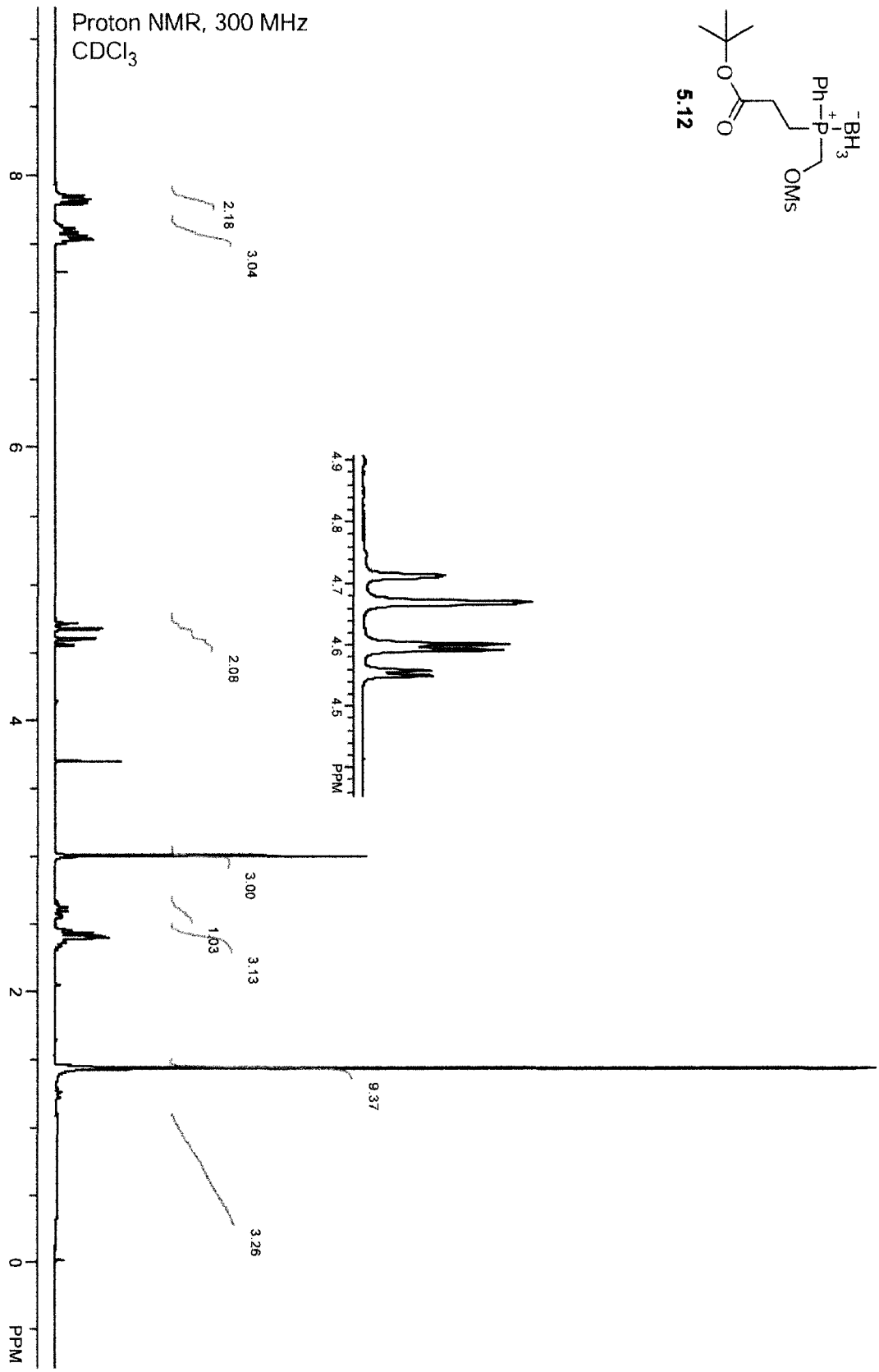
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



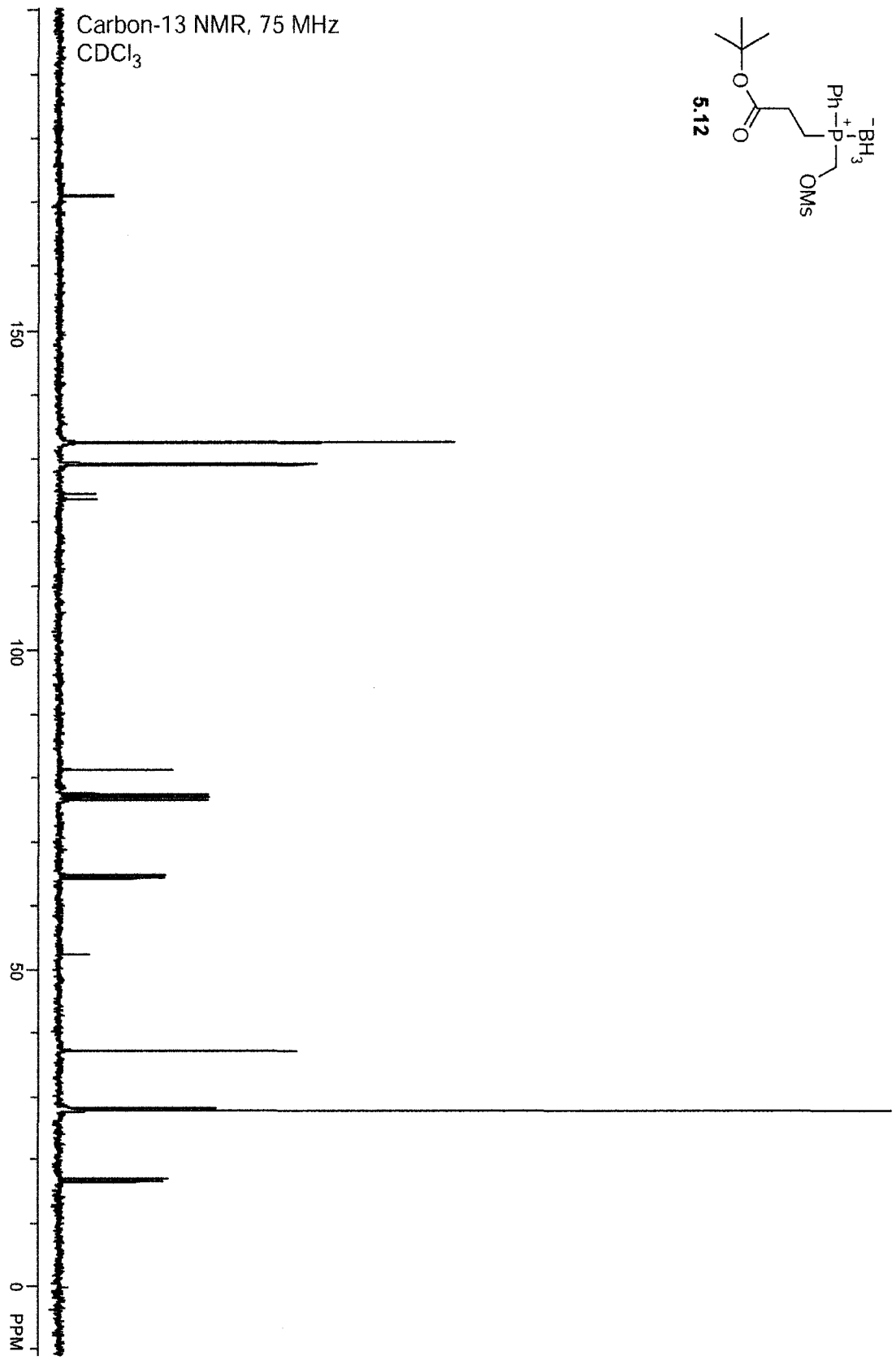
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



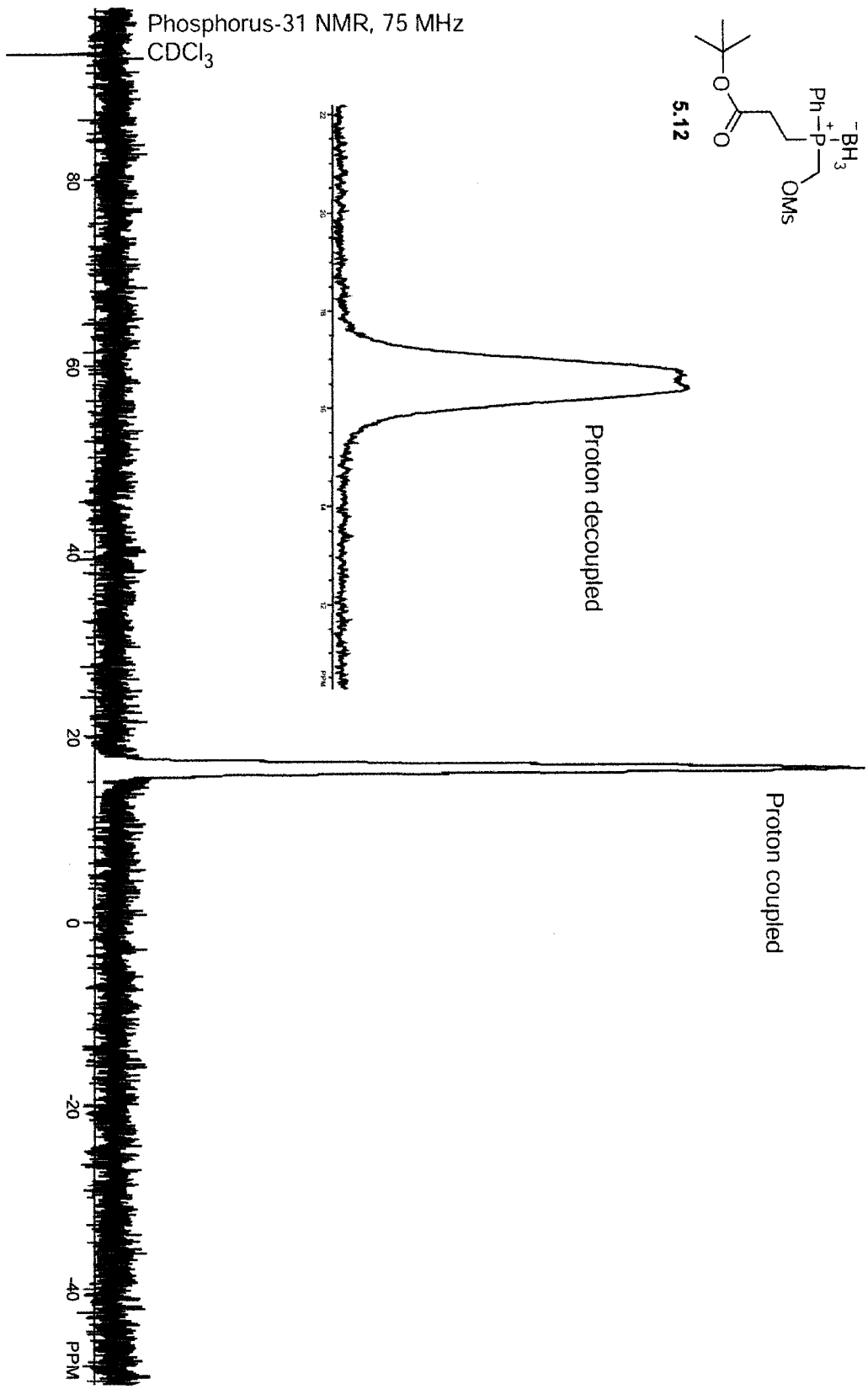
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



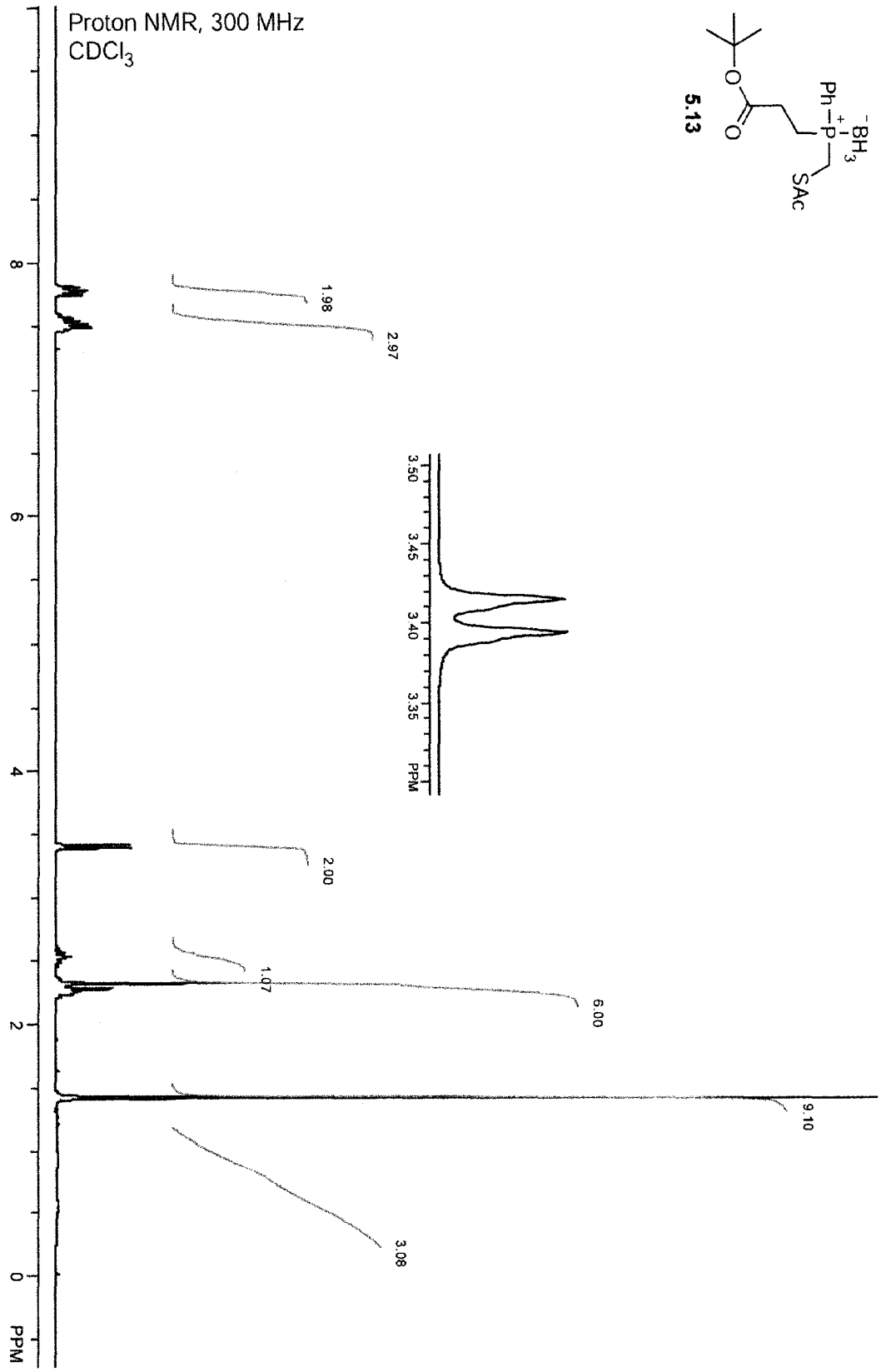
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



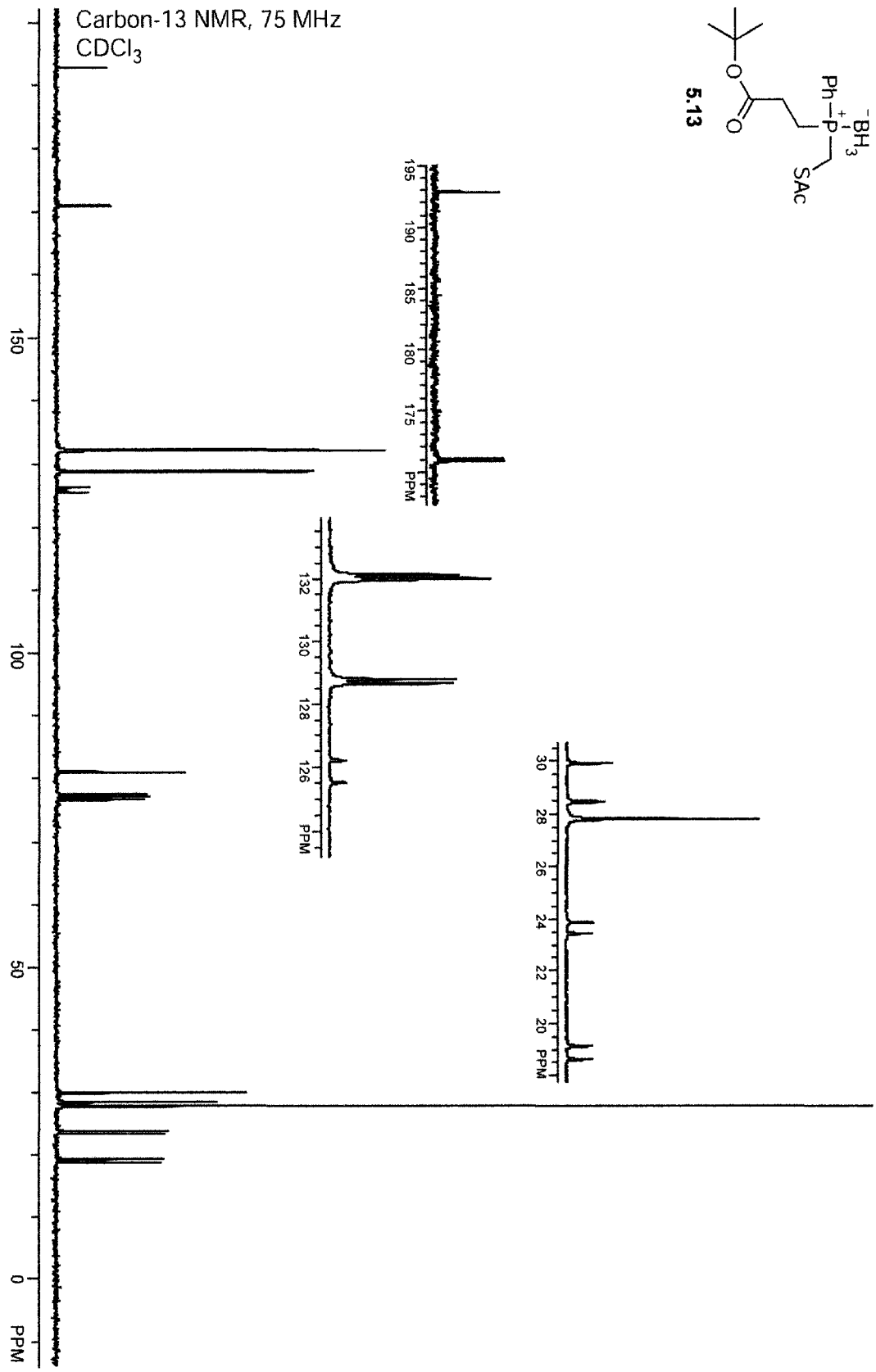
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



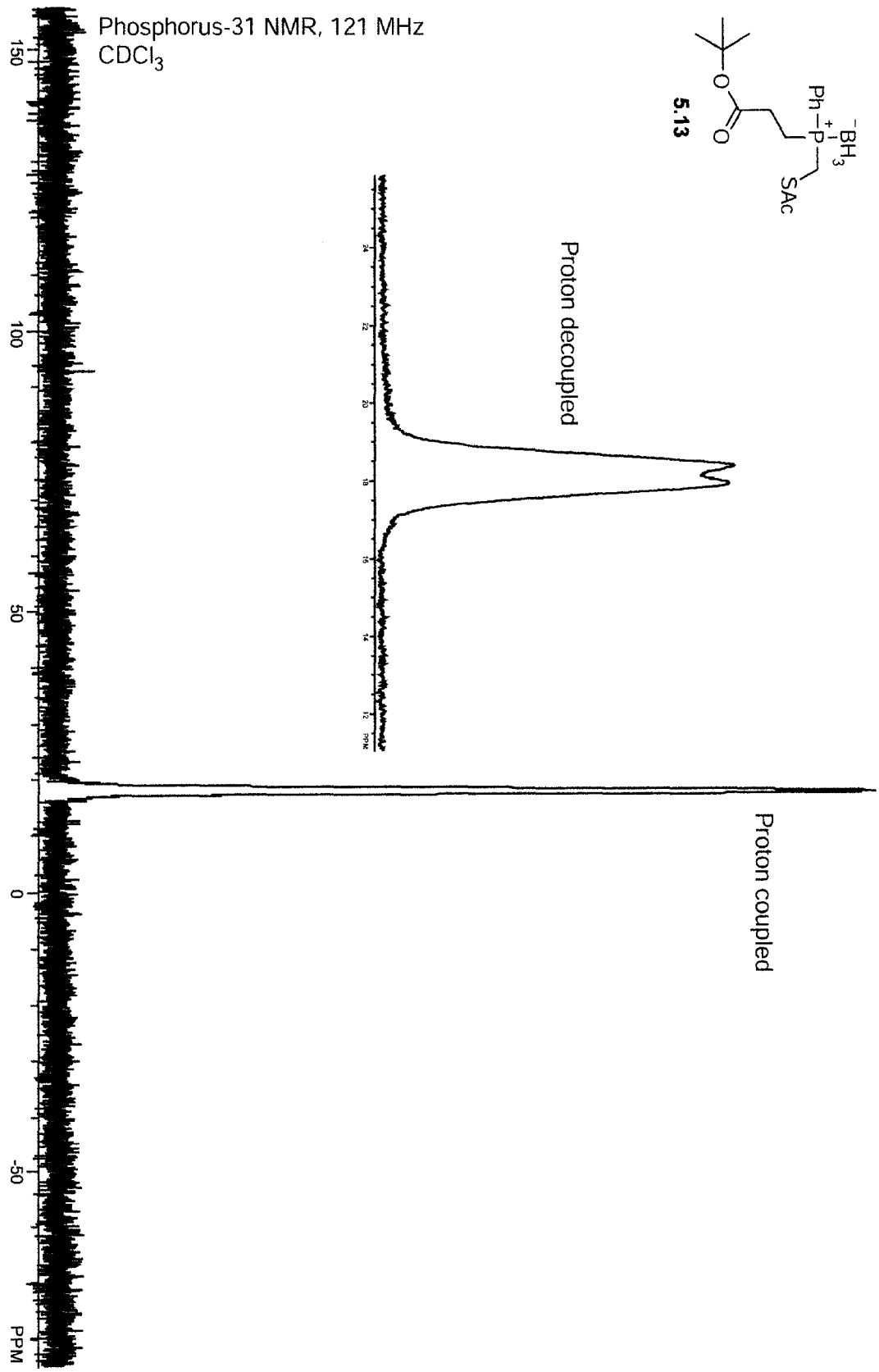
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



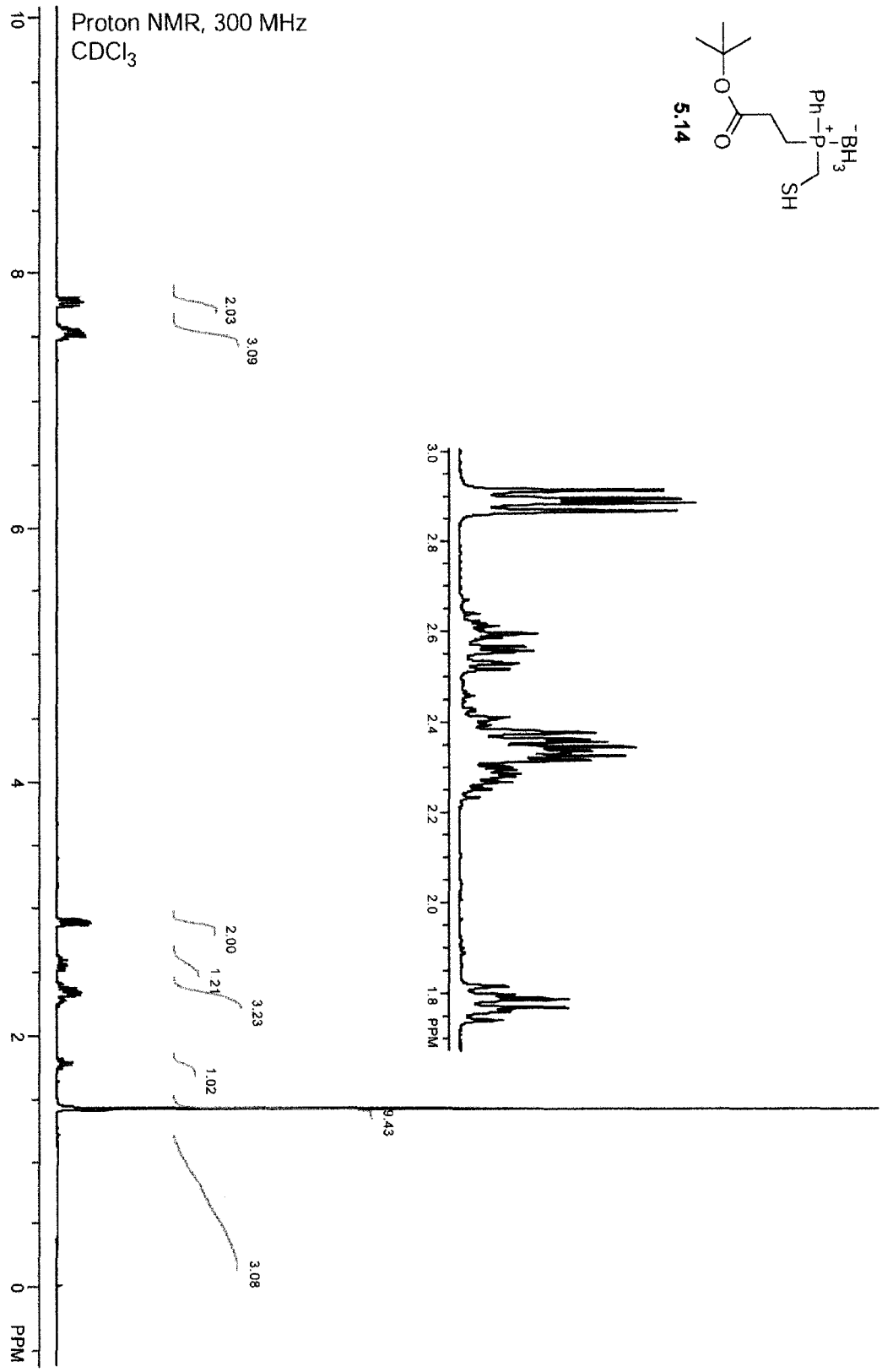
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



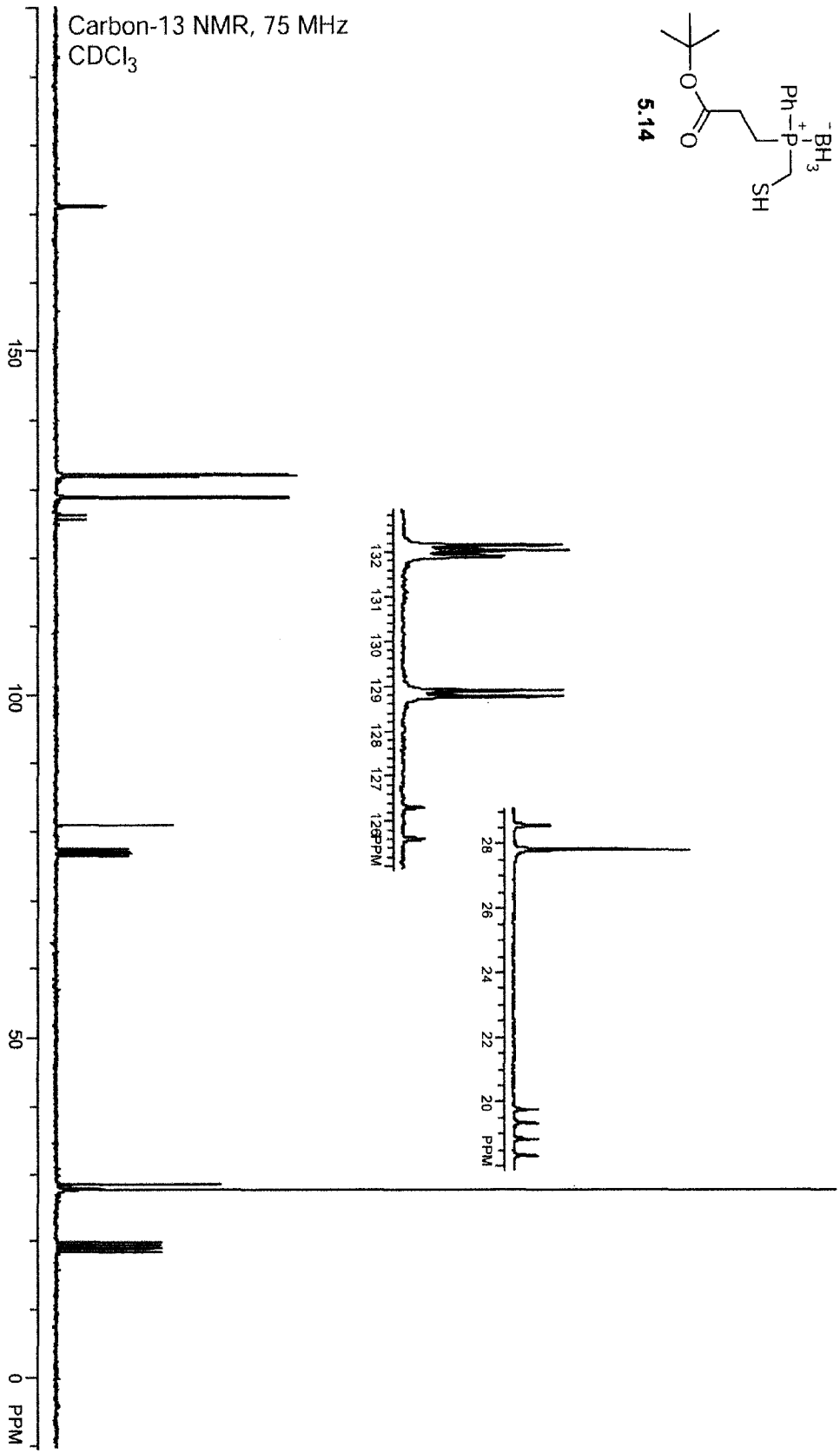
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



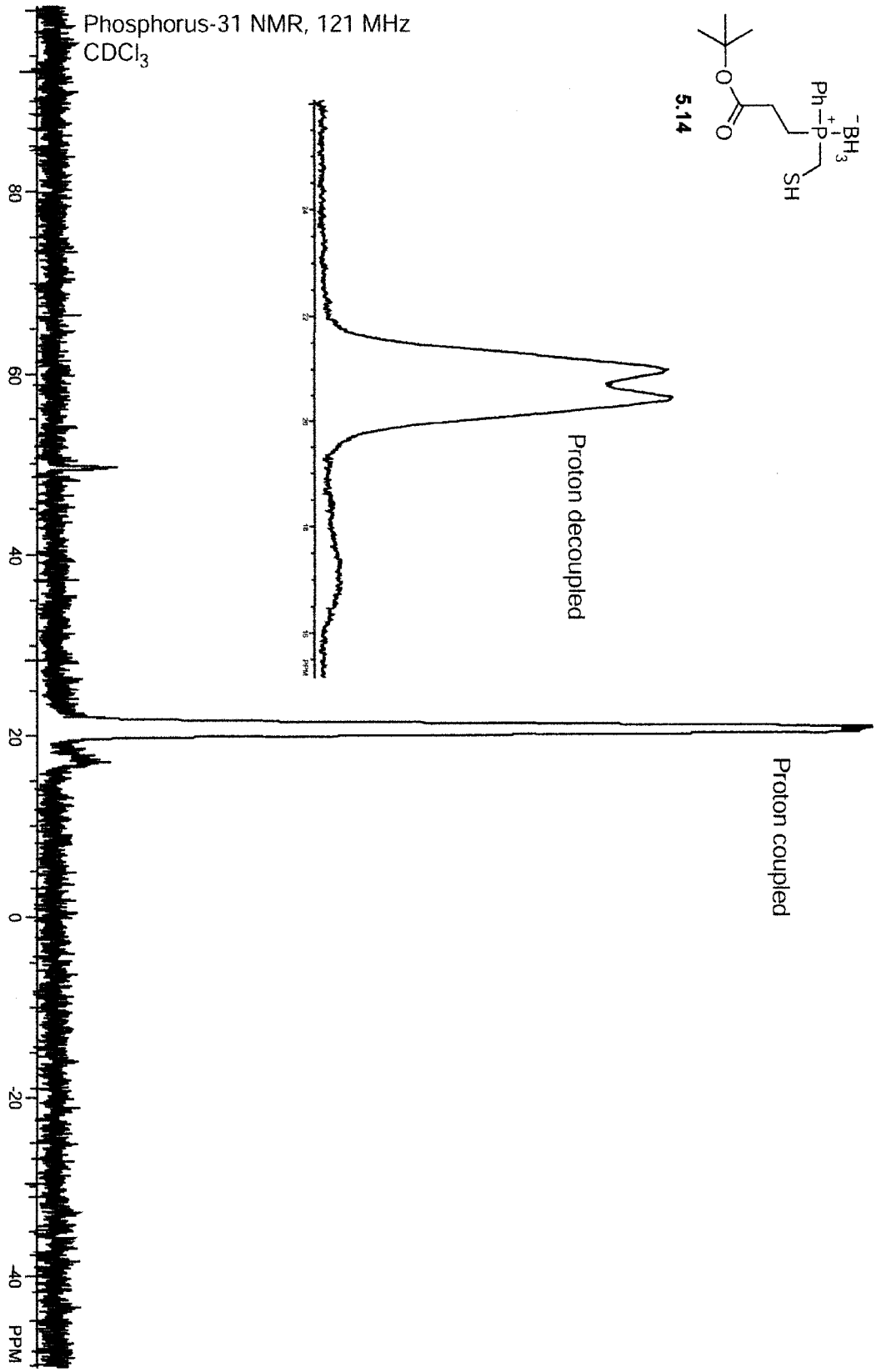
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



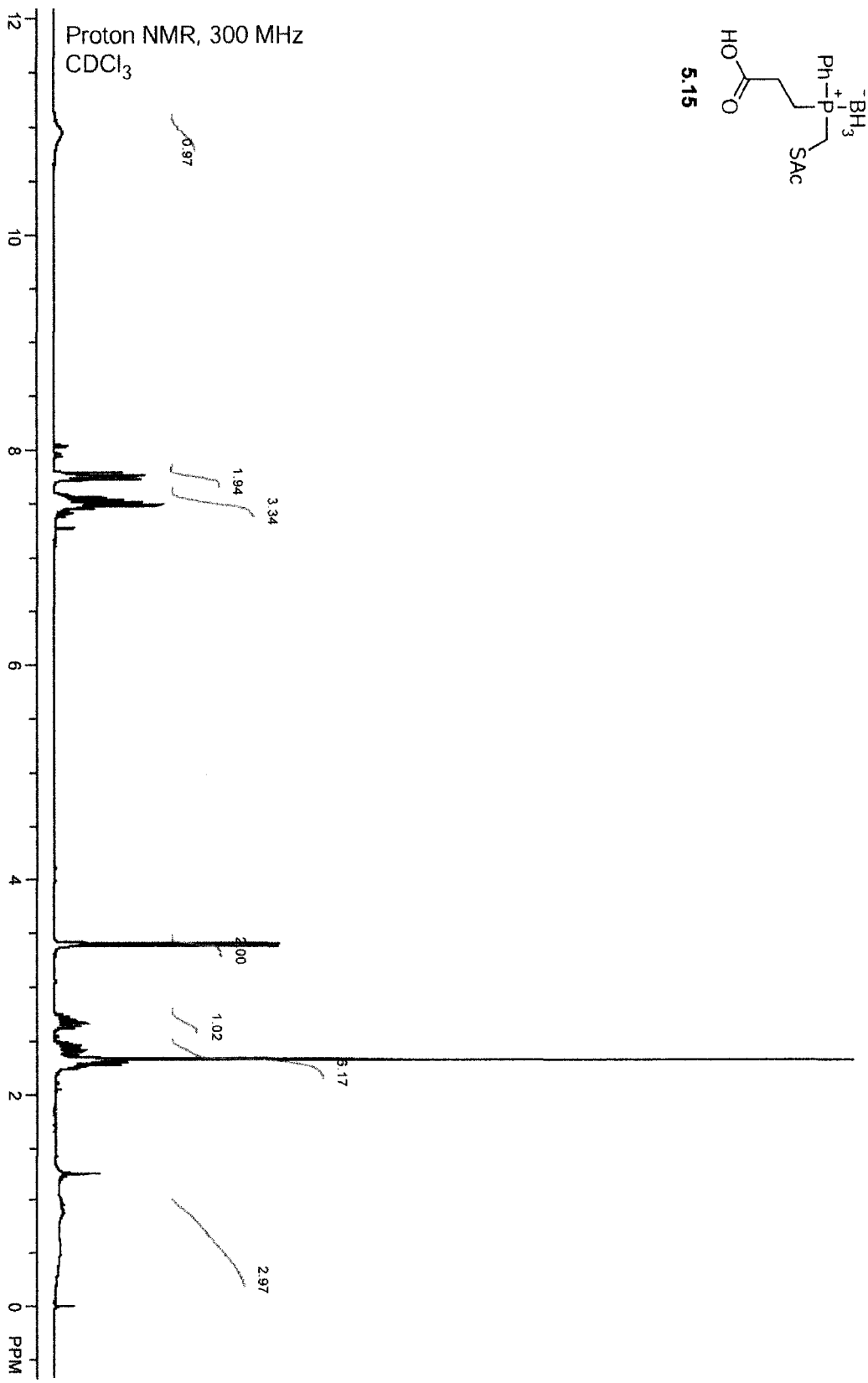
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

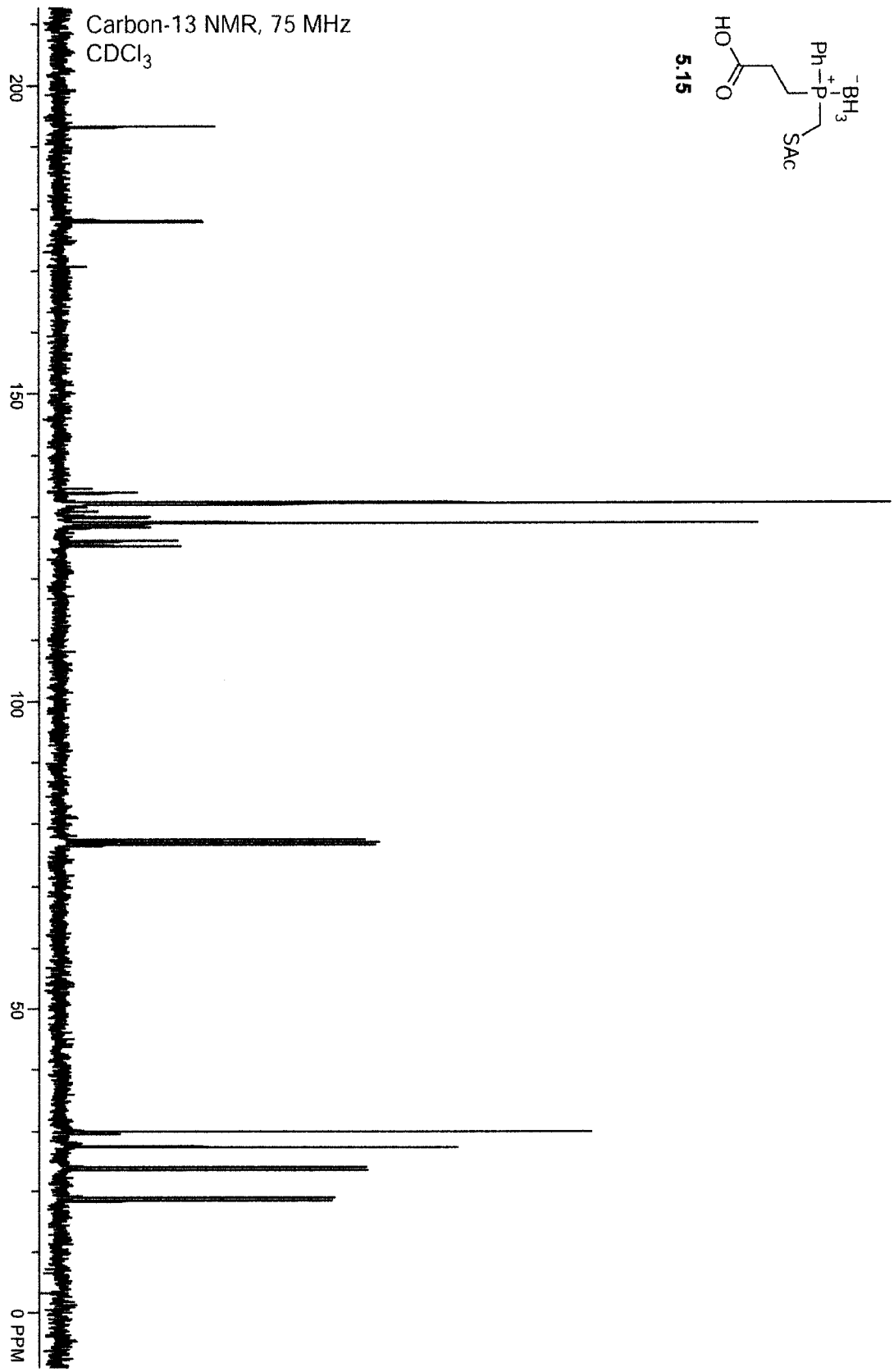


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



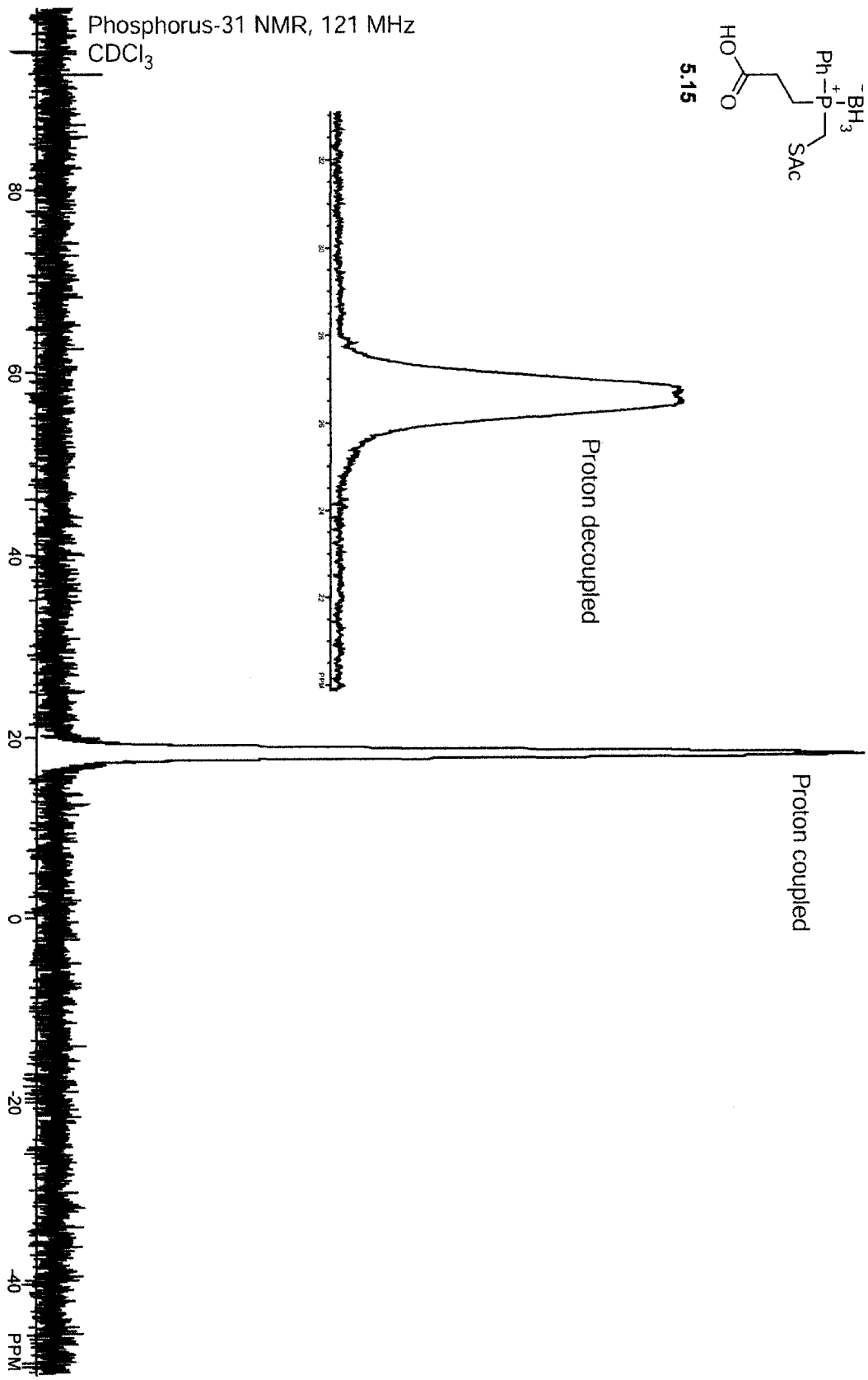
5.15

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

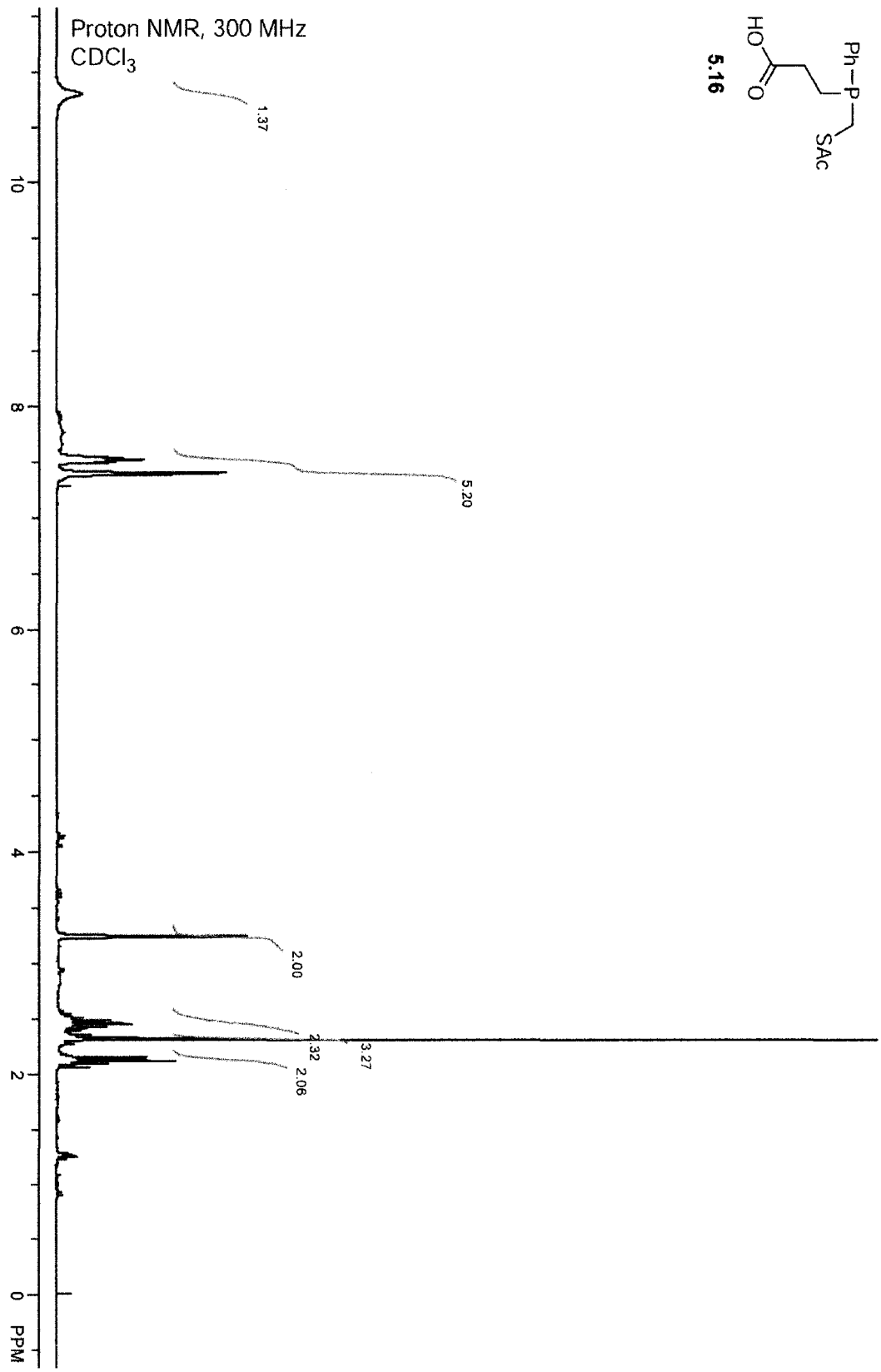


5.15

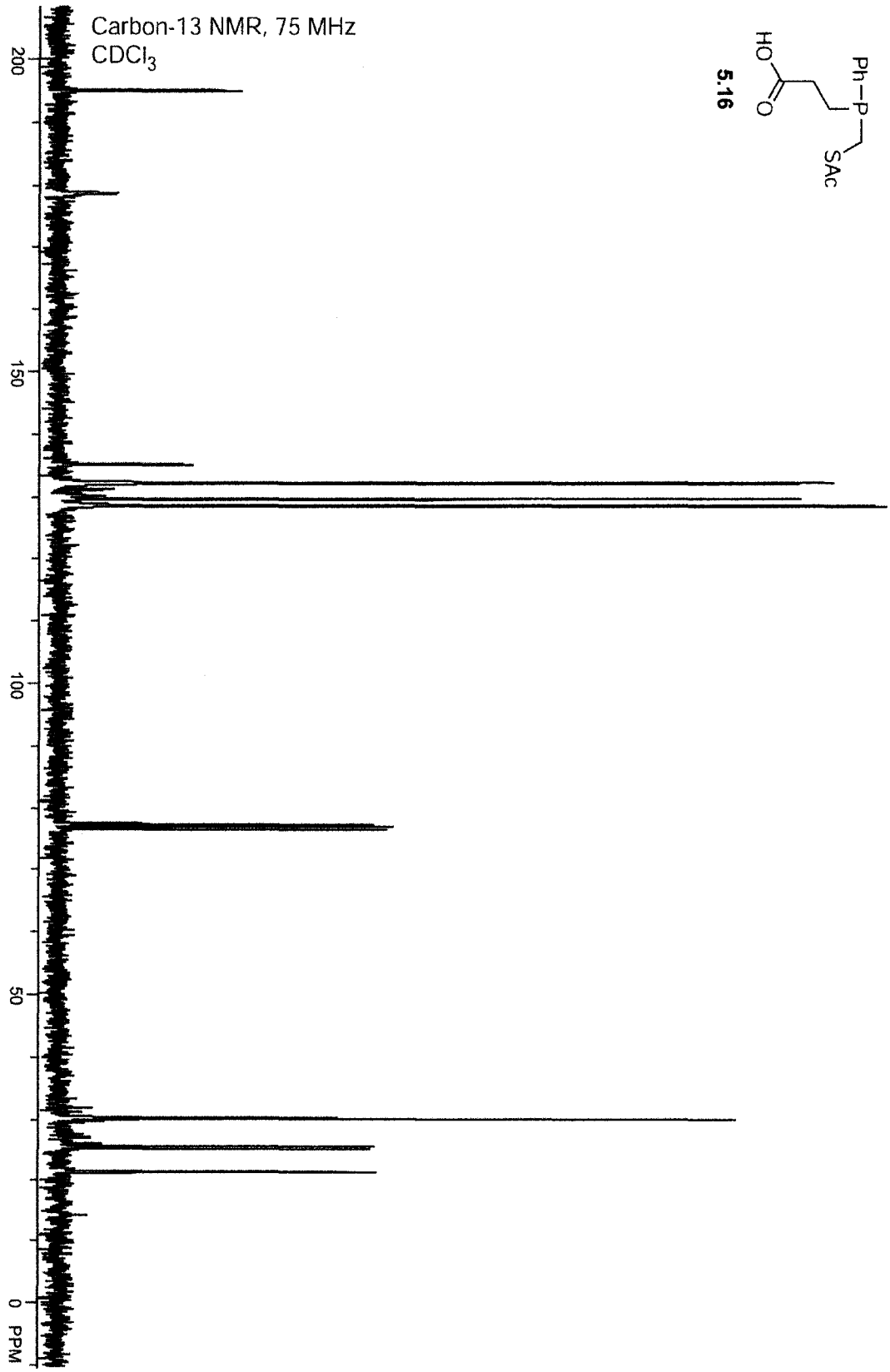
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

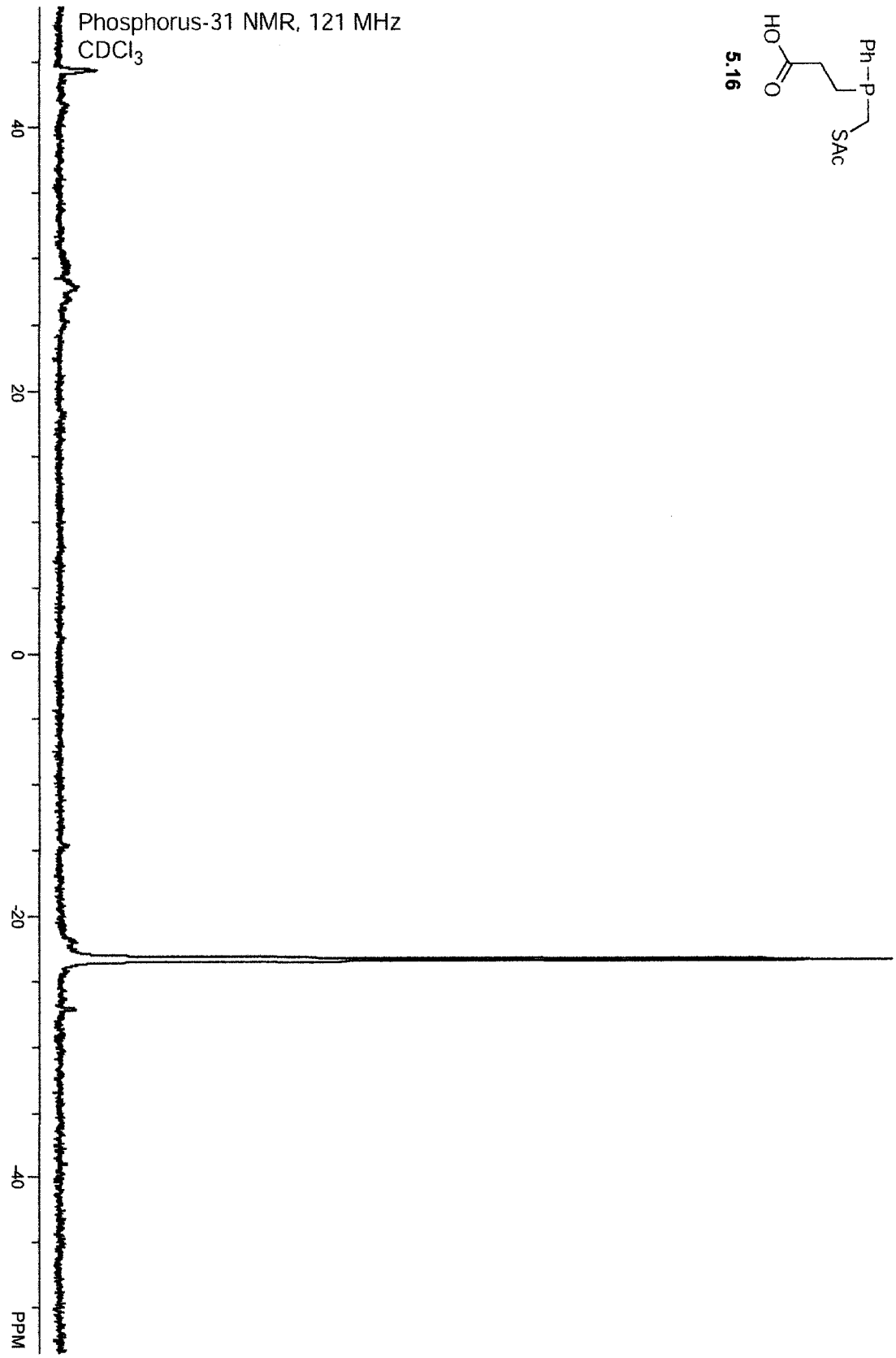


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



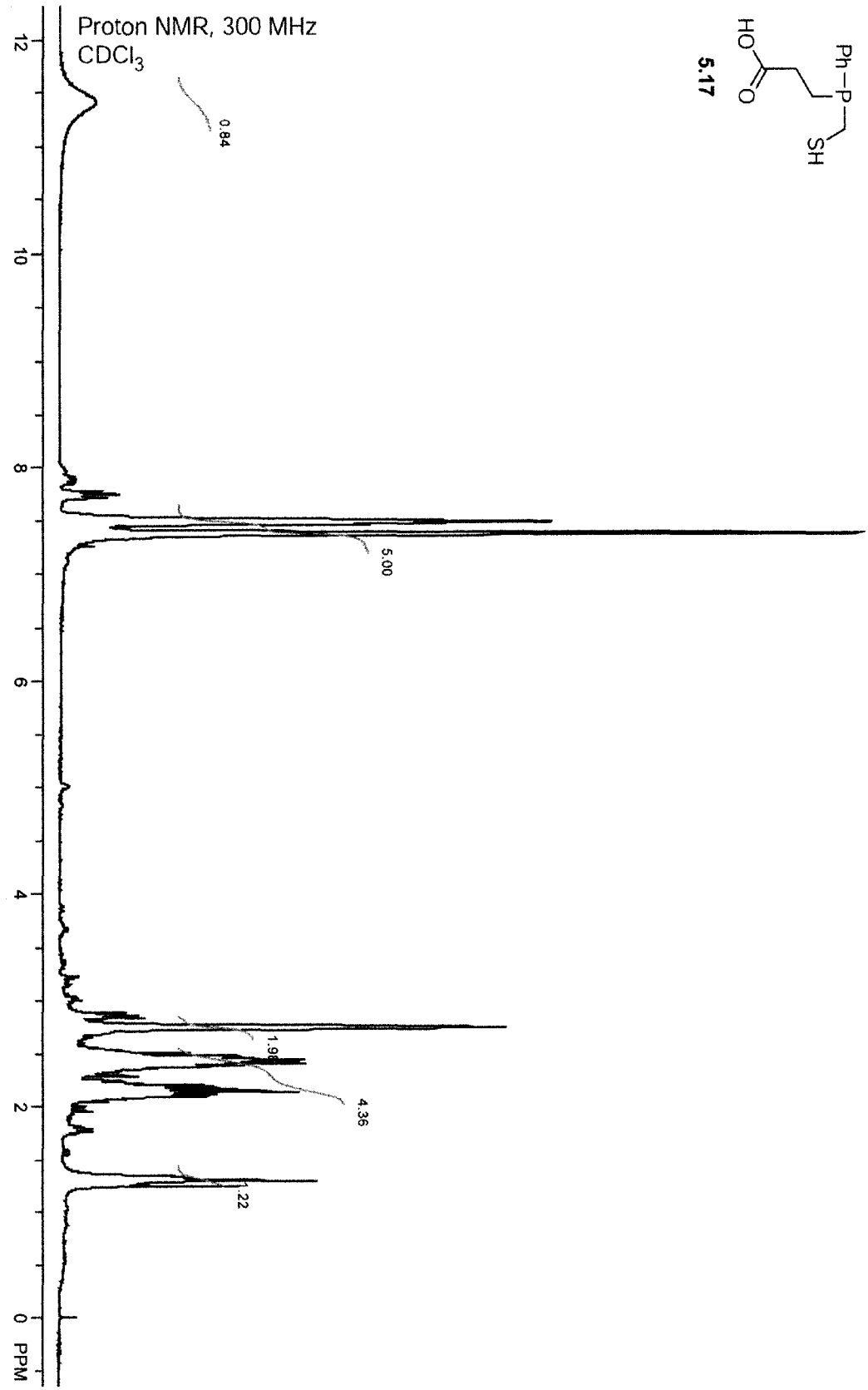
5.16

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

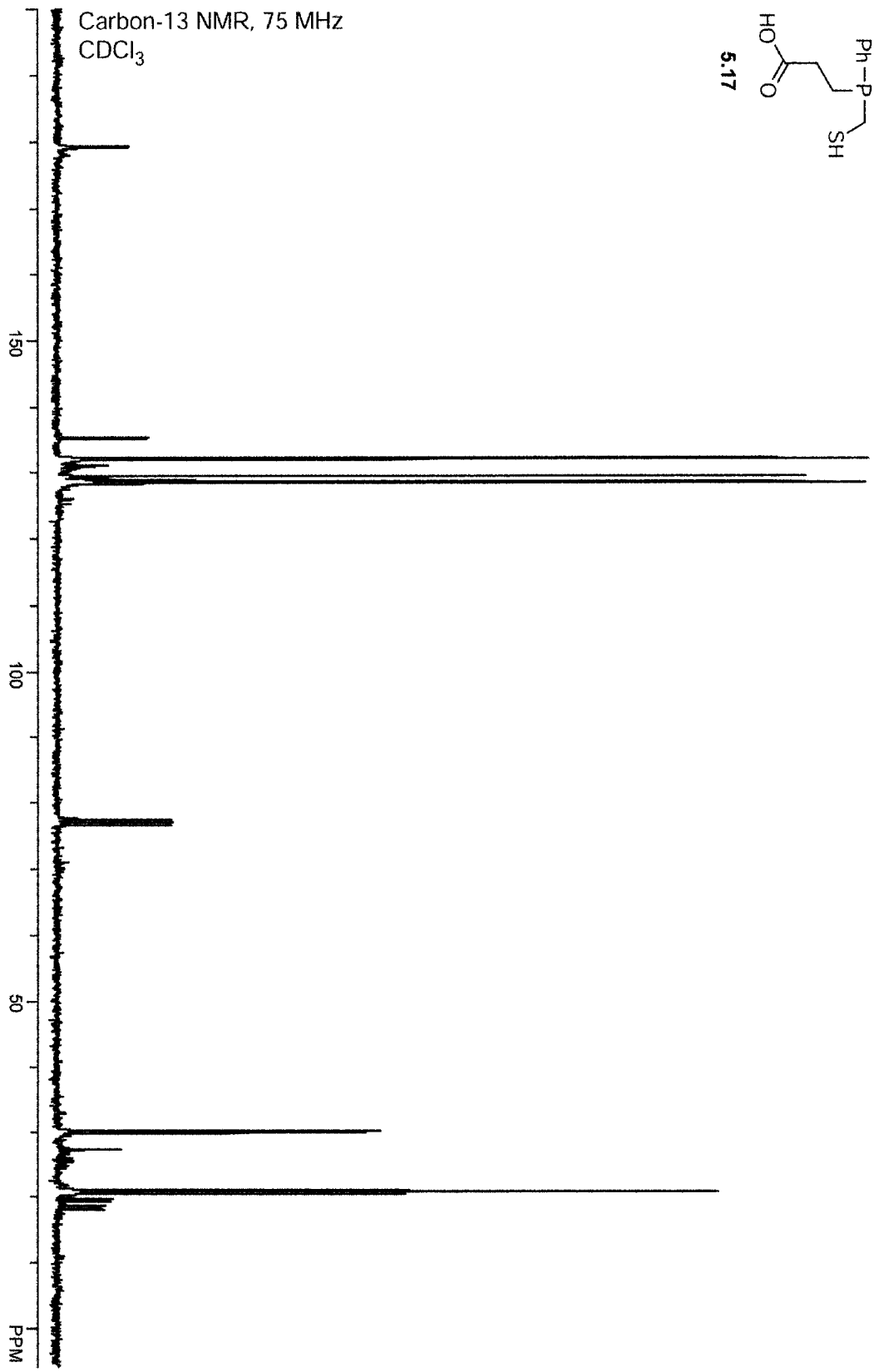


5.16

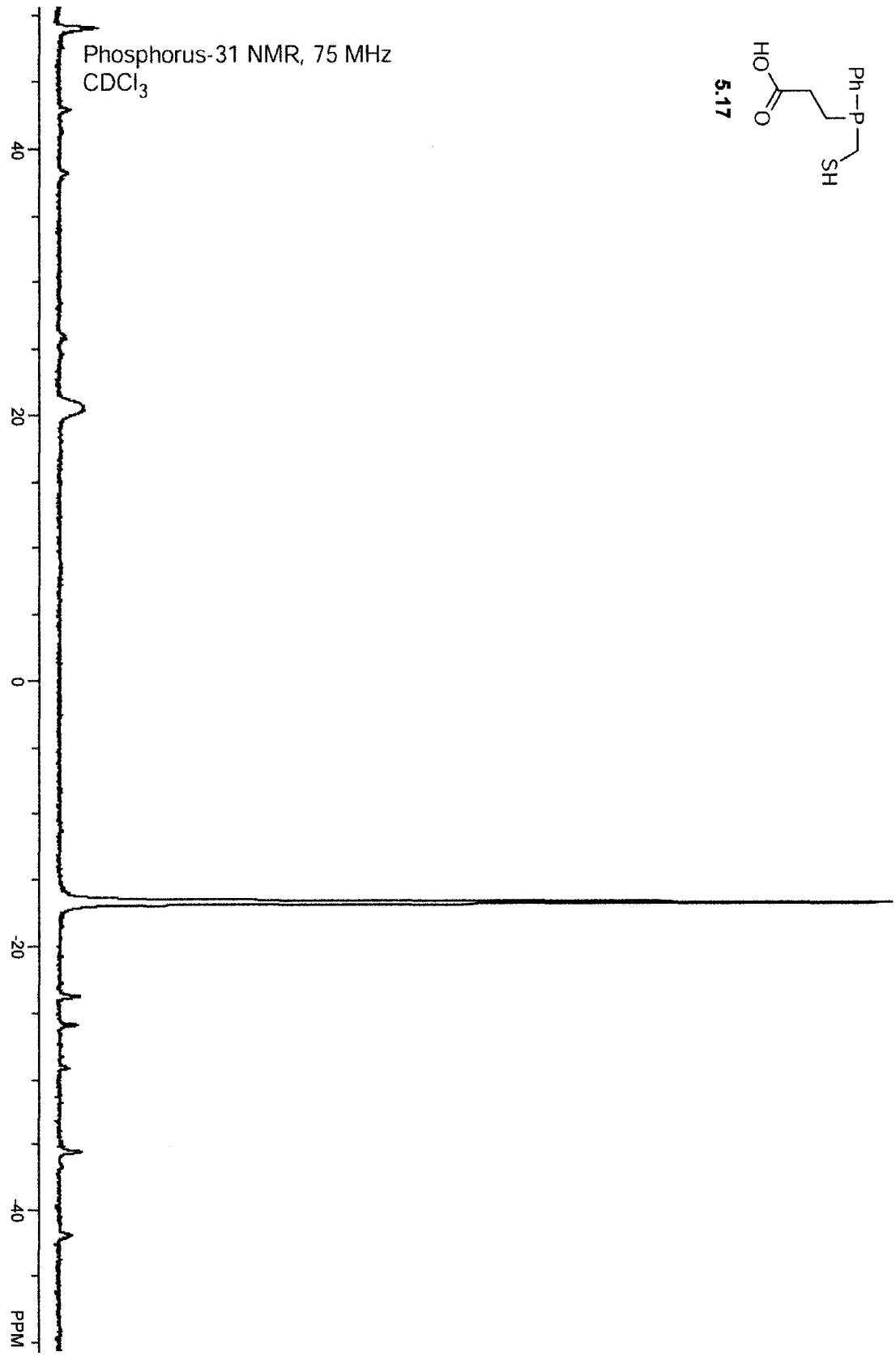
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



5.17

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 6

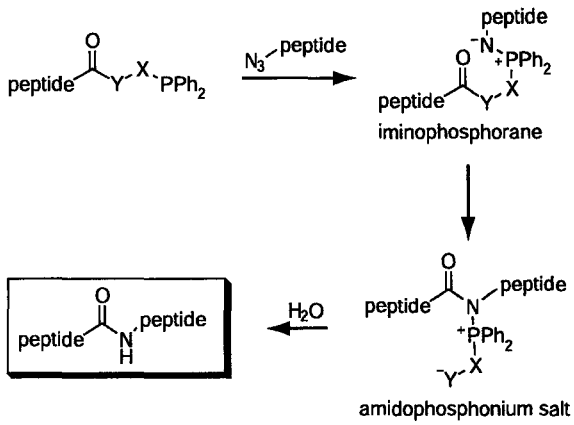
IMPROVED PHOSPHINES FOR STAUDINGER LIGATION AT NON-GLYCINE RESIDUES PREDICTED USING MOLECULAR MODELING

The advent of methodology for the chemoselective ligation of peptide fragments has made proteins accessible targets for total chemical synthesis. Many proteins have already been assembled from synthetic peptides using prior capture strategies. "Native chemical ligation"—the coupling of a peptide (or protein) containing a *C*-terminal thioester with another peptide containing an *N*-terminal cysteine residue—has become the most widely applied capture strategy for the total chemical synthesis of proteins (Dawson & Kent, 2000; Kent, 2003). "Expressed protein ligation" is an application of native chemical ligation in which *C*-terminal thioesters for native chemical ligation are accessed using recombinant DNA techniques (Muir, 2003). These powerful methods are limited by the requirement for a cysteine residue at the ligation junction.

Emerging strategies for protein assembly avoid the need for a cysteine residue at the ligation junction. The Staudinger ligation is one such strategy (Scheme 6.1). In one form of the Staudinger ligation, a peptide with a *C*-terminal phosphinothioester is coupled with a second peptide having an *N*-terminal azido acid through the intermediacy of an iminophosphosphorane. The iminophosphorane can rearrange to give an amidophosphonium salt as an acyl transfer product that is hydrolyzed in the presence of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

water. The final ligation product contains a native amide bond without any residual atoms (Nilsson et al., 2000; Nilsson et al., 2001; Nilsson et al., 2003). The overall reaction occurs without racemization (Soellner et al., 2002). The Staudinger ligation has been used in the orthogonal assembly of a protein (RNase A) (Nilsson et al., 2003) and for the site-specific immobilization of peptides and small molecules to a surface (Köhn et al., 2003; Soellner et al., 2003). Different reagents and methods have also been explored in applying the Staudinger ligation to other problems, including *in vivo* labeling experiments (Saxon et al., 2000; Saxon & Bertozzi, 2000a; Saxon et al., 2002; Vocadlo et al., 2003)

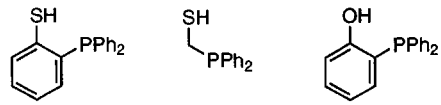


Scheme 6.1 *Staudinger ligation.*

Three separate reagents have been studied specifically for applications to peptide synthesis (Figure 6.1). Compound **6.1** was able to provide dipeptides from phosphinothioesters and azido acids, albeit in moderate to poor yields of ≤35% (Nilsson et al., 2000). Compound **6.2** represents a second generation reagent that successfully affects the Staudinger ligation in high yield for model couplings (>90%) (Nilsson et al., 2001). These successful couplings are performed in mixed THF/water or DMF/water

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

solvents with a stoichiometric ratio of reagents. We have found, however, that for couplings at Xaa–Yaa sites, either Xaa or Yaa must be glycine for efficient ligation to occur. Reagent **6.2** ligates inefficiently when both Xaa and Yaa are non-glycine residues. In a recent paper, the Liskamp group has reported that amino acid phophinooxyesters of reagent **6.3**, previously reported for use in non-amino acid Staudinger ligations (Saxon et al., 2000), can mediate the Staudinger ligation between non-glycyl residues in moderate yields (36% for Ala–Phe) under certain conditions (Merkx et al., 2003). The coupling partners were heated at 47 °C in anhydrous THF, followed by the addition of water at a later stage. Presumably, hydrolysis of the iminophosphorane intermediate does not compete with the acyl transfer step under these conditions.



6.1 6.2 6.3

Figure 6.1 *Reagents used to form peptides via the Staudinger ligation.*

These results prompted us to analyze the iminophosphorane intermediates with molecular modeling. Specifically, we performed conformational searches on the iminophosphorane intermediates formed by phophinothio- or oxyesters of reagents **6.1-3** and non-glycyl azido acids to reveal potential structures of these intermediates. We were curious about the distance between the nucleophilic nitrogen of the iminophosphorane and the carbon of the electrophilic carbon of the thio- or oxyester. Although the results are not entirely conclusive in explaining why some ligations succeed while others fail,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

they did show some interesting trends, particularly when comparing thioesters of **6.1** to oxyesters of **6.3**.

Conformational searching of the iminophosphorane intermediates of phosphinothioesters of compound **6.2** gave surprisingly uniform results, even when comparing ligations that are known to succeed experimentally and ligations that are known to fail. The interatomic distance between the iminophosphorane nitrogen and the thioester carbonyl carbon was consistently on the order of 5.60 Å (Figure 6.2). This distance is not conducive to reaction between these activated atoms. We then performed conformational searches on examples where this distance is constrained to a reactive distance of 3.0 Å. Our assumption upon viewing these results is that steric interactions between amino acid functionality in Xaa–Yaa couplings in which neither residue is a glycine could preclude the proper approach of the activated atoms to each other.

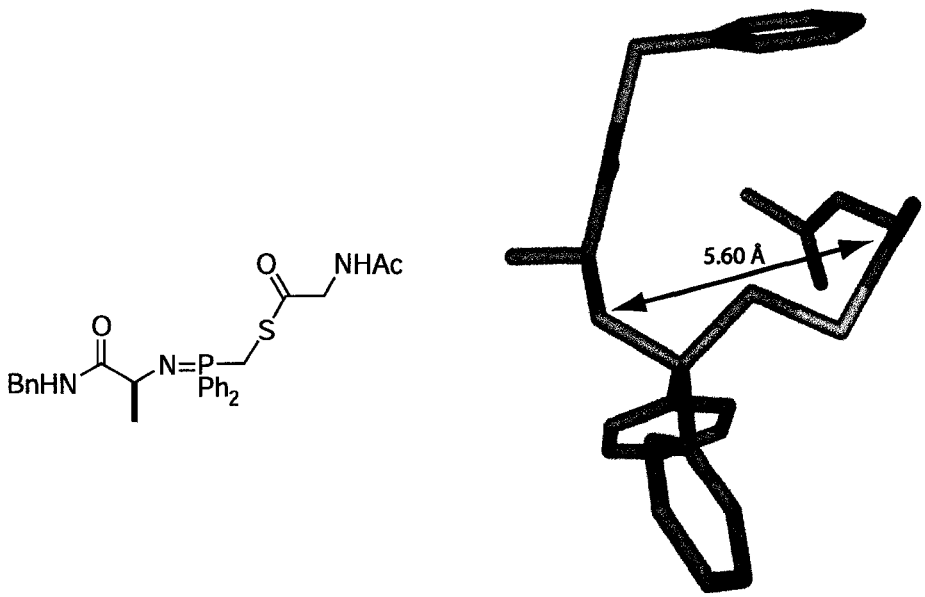


Figure 6.2 *Nitrogen–carbon distance of AcGly–AlaBn iminophosphorane with compound* ***6.2****.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Next, we modeled iminophosphoranes of the phosphinothio- or oxy-esters of compounds **6.1** or **6.3** respectively. We chose leucine as the azido acid component since leucine is the most common of the amino acids in proteins (McCaldon & Argos, 1988a). We arbitrarily chose phenylalanine as the ester component, which would represent a reasonably challenging Xaa–Yaa target ligation sequence with a leucine azide. The interatomic distance for the iminophosphorane of phosphinothiol **6.1** is >5.10 Å (Figure 6.3). For phosphinophenol **6.3** the interatomic distance is 4.15 Å (Figure 6.4). This distance is closer by nearly 1 Å, due largely to the smaller atomic radius of oxygen versus sulfur. This shorter N–C distance could explain, in part, why Staudinger ligation to form Xaa–Yaa peptides occurs more efficiently using compound **6.3**. The reactive atoms in the iminophosphorane intermediate approach each other more closely, and this closer approach increases the chance for a successful acyl transfer event. In addition, the decreased ring size is likely to allow more favorable geometry for the approach of the nucleophile to the electrophile.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

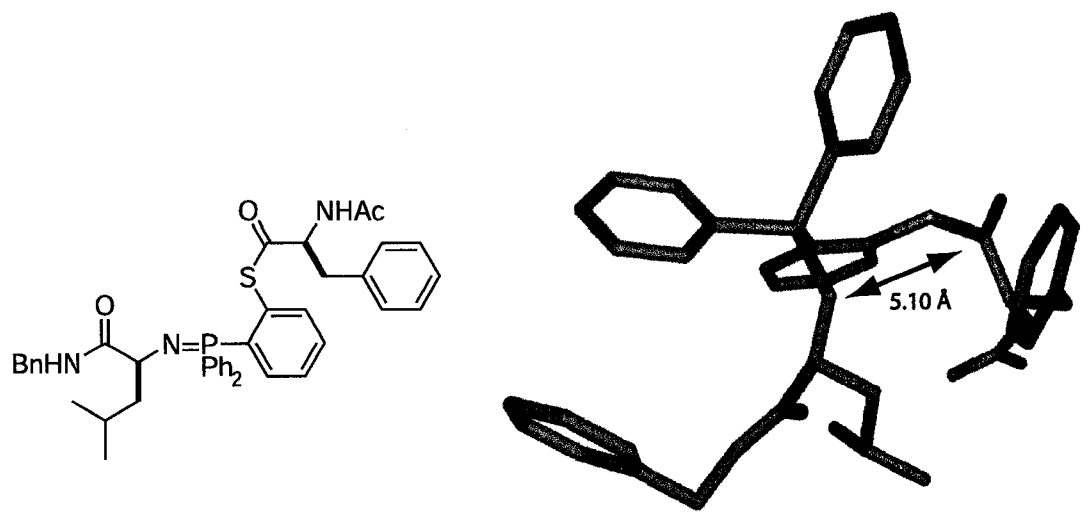


Figure 6.3 *Nitrogen–carbon distance of AcPhe–LeuBn iminophosphorane with compound* ***6.1****.*

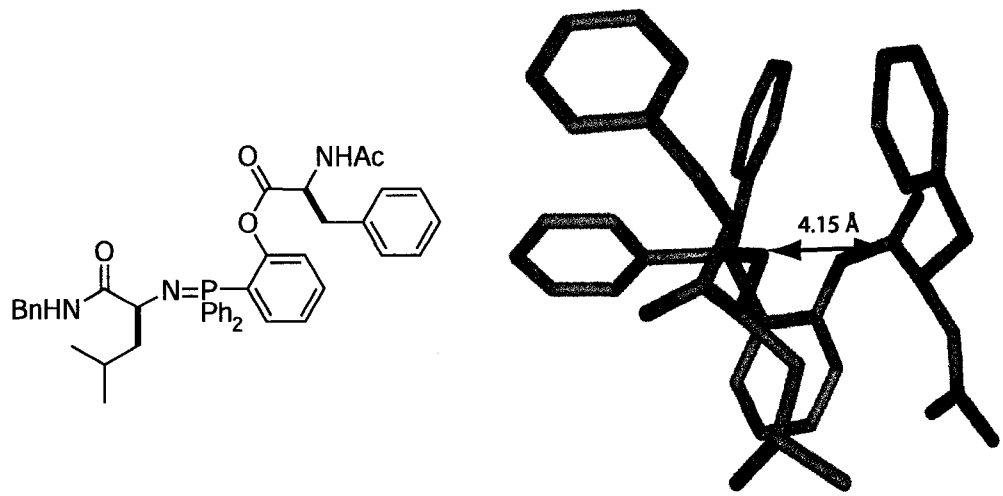
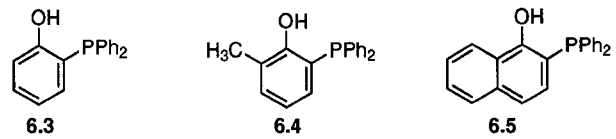


Figure 6.4 *Nitrogen–carbon distance of AcPhe–LeuBn iminophosphorane with compound* ***6.3****.*

In considering these results, we wondered if it might be possible to design elements into a phosphine reagent that could make the distance between the iminophosphorane nitrogen and the ester carbon even less. We hypothesized that placing a second substituent *ortho* to the phenol oxygen (or thiophenol sulfur) in the aromatic ring could

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

force the ester carbon of the iminophosphorane intermediate to be even closer to the iminophosphorane nitrogen. Such steric crowding in the iminophosphorane intermediate could result in improved yields of Staudinger ligation product.



6.3 6.4 6.5

Figure 6.5 *Designed phosphines incorporating crowding elements.*

Phosphines **6.4** and **6.5** represent possibilities for such improved reagents (Figure 6.5). We applied the tools of molecular modeling to each of these phosphines in addition to modeling an *o*-methyl derivative of our phosphinothiophenol **6.1**, which is analogous to phosphinophenol **6.4**. The results of these conformational searches were encouraging. By placing a methyl-substituent *ortho* to the sulfur in compound **6.1**, the reactive atoms in the iminophosphorane intermediate came to within 3.92 Å of one another (Figure 6.6). This is distance is nearly 1.2 Å closer than that without the methyl group (Figure 6.3). The modeling for compound **6.4** yielded an N–C distance of 3.26 Å (Figure 6.7). This distance is near or at a distance where acyl transfer should occur readily. The modeling for the naphthol-phosphine **6.5** yielded an N–C distance of 3.24 Å, which is nearly identical to that for compound **6.4** (Figure 6.8).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

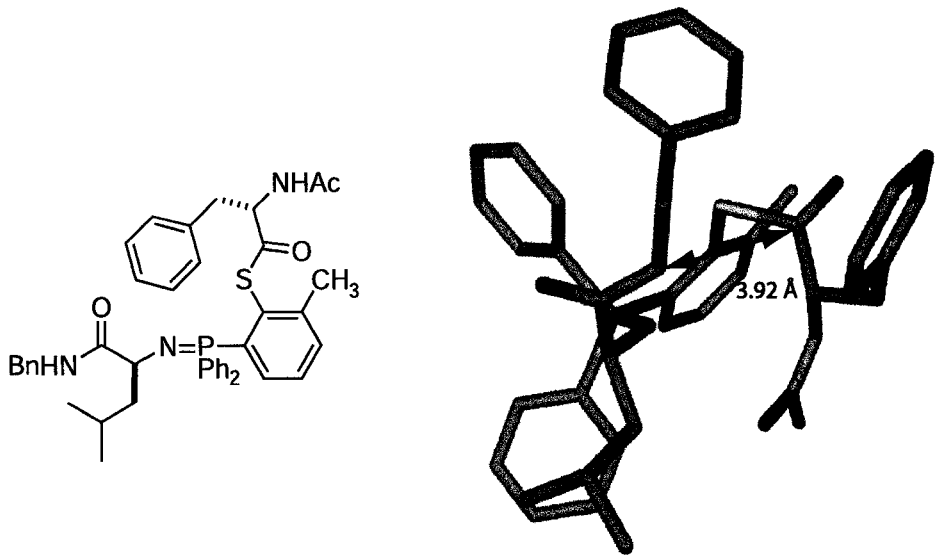


Figure 6.6 *Nitrogen–carbon distance of AcPhe–LeuBn iminophosphorane with the ortho-methyl derivative of compound* ***6.1****.*

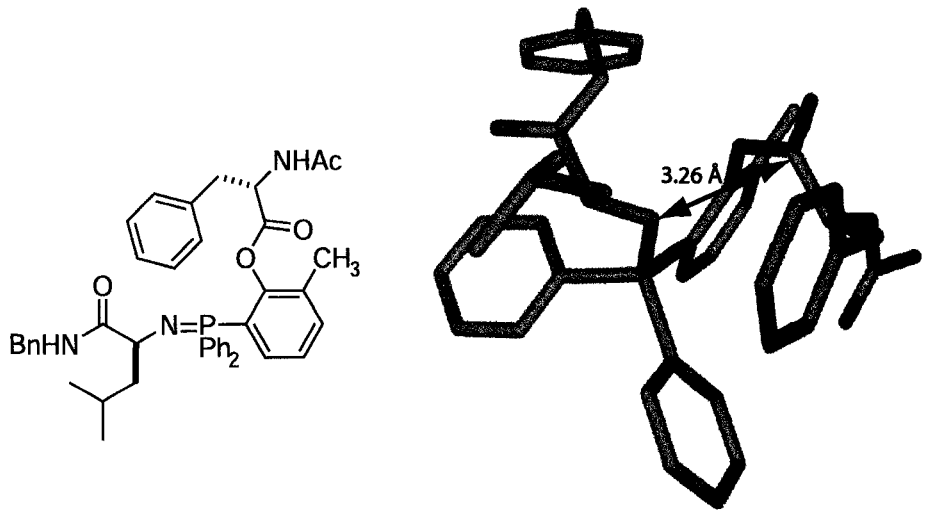


Figure 6.7 *Nitrogen–carbon distance of AcPhe–LeuBn iminophosphorane with compound* ***6.4****.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

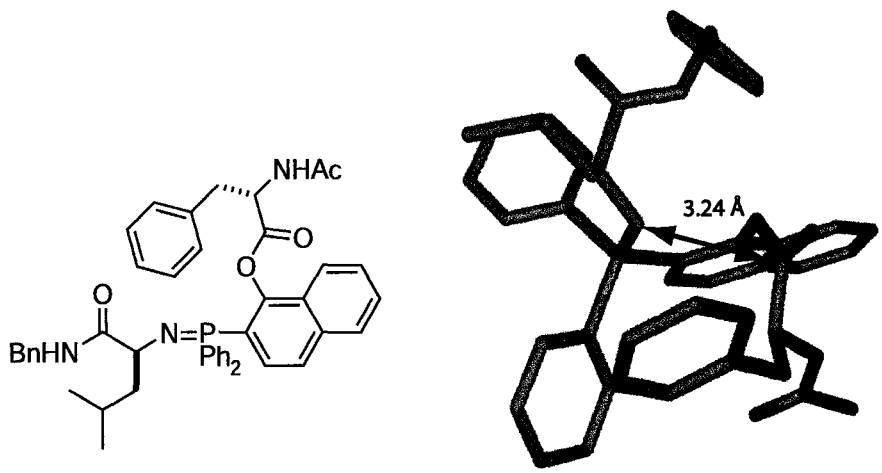
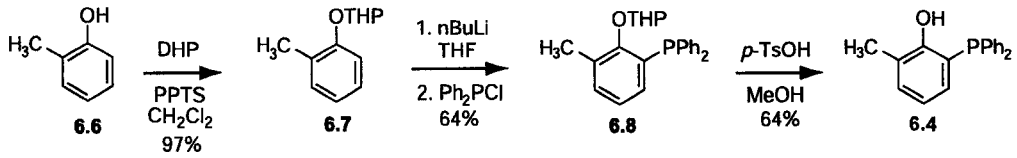


Figure 6.8 *Nitrogen–carbon distance of AcPhe–LeuBn iminophosphorane with compound* ***6.5****.*

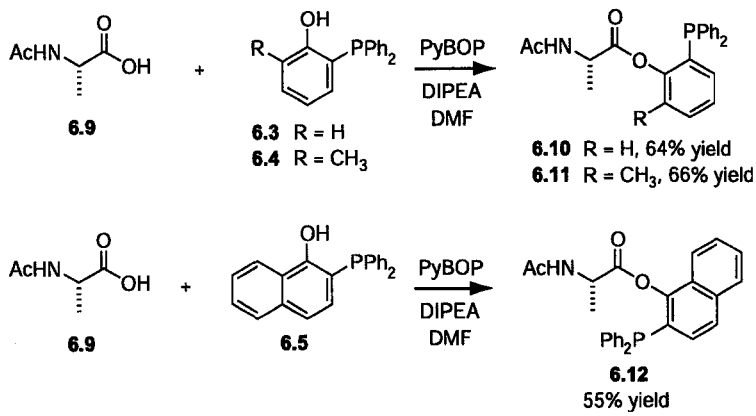
We decided to synthesize phosphines **6.3**, **6.4**, and **6.5** in order to compare the efficiency with which each mediates the Staudinger ligation. In addition, we were interested to test the efficacy of our molecular modeling in predicting improved reagents for amide bond formation. Compounds **6.3** (Herd et al., 1998; Saxon et al., 2000) and **6.5** (Blume et al., 2002) have been reported previously, and we prepared them by the published protocols. Phosphine **6.4** was synthesized as shown in Scheme 6.2. The phenolic oxygen of *o*-cresol was protected as the tetrahydropyran derivative (**6.7**) in nearly quantitative yield. The oxygens of the THP-protected phenol served as *ortho*-directors for the lithiation of the phenyl ring. Chlorodiphenyl phosphine was added to the lithiated species to give protected phosphine **6.8** in 64% yield. Deprotection of the phenolic oxygen under acidic conditions gave phosphine **6.4**, again in 64% yield.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Scheme 6.2 *Synthesis of phosphine* ***6.4****.*

The three phosphinophenols were then coupled to AcAlaOH prior to carrying out test ligations. The oxyesters were formed by PyBOP-mediated coupling to the phenolic phosphines (Scheme 6.3). The desired phosphinophinooxyesters, compounds **6.10**, **6.11**, and **6.12**, were obtained in 64%, 66%, and 55% yields, respectively. These reactions were complete in 6 h when PyBOP is used as the condensation agent. When dicyclohexylcarbodiimide (DCC) couplings were attempted, the isolated yields were similar, but the reactions took from 12–16 hours to go to completion.



Scheme 6.3 *Synthesis of alanyl phosphinooxyesters for Staudinger ligation.*

With these compounds in hand, model Staudinger ligation reactions were performed using N_3AlaNHBn as the azide. We used the Liskamp conditions to carry out these

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

couplings (Merkx et al., 2003). The reactions were carried out using 1:1 stoichiometry of phosphinooxyester to azide. The phosphinooxyester was dissolved in dry THF and heated to 47 °C. The azide was then added and the reaction mixtures were stirred for 12 h, followed by the addition of water and heating for another 1 h. Solvent was then removed under reduced pressure and the residue was purified by chromatography in order to isolate the desired dipeptide, AcAlaAlaNHBn and ascertain its yield. The results of these experiments are summarized in Table 6.1.

The unsubstituted phosphinophenol gave lower yields of ligated product than did the *ortho*-substituted derivative. The desired AcAlaAlaNHBn dipeptide was isolated in 22% yield using phosphinophenol **6.3**. The dipeptide yield was better when the cresol-derived phosphine **6.4** was used. In this case AcAlaAlaNHBn was isolated in 37% yield. When this coupling was carried out in DMF as a solvent, the isolated yield was 38%. These two reagents were also used in an NMR experiment using alanyl azide incorporating carbon-13 label at the α-carbon. The yields could be determined through integration of the α-carbon peaks. The NMR yields were 45% dipeptide using reagent **6.11** versus 35% dipeptide using compound **6.10**. Couplings between naphthol-derived compound **6.12** and the alanyl azide (Table 1, entry 3) afforded dipeptide in a disappointing 21% yield. This low yield may be due to problems with purification of the dipeptide since the dipeptide and one of the by-products were difficult to separate using chromatography. The low yield may also be due to some electronic effect brought about by the aromatic substitution para to the phosphine in these naphthol-derived phosphines.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

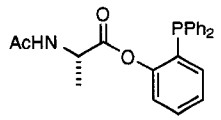
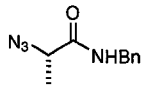
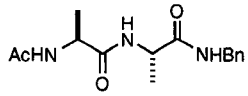
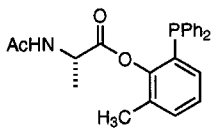
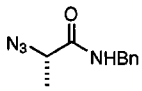
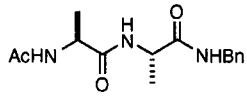
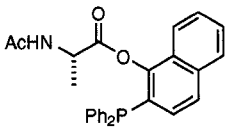
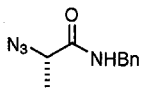
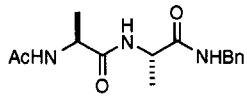
Phosphinoester	Azido acid	Product	Yield
 6.10	 6.13	 6.14	22%
 6.11	 6.13	 6.14	37%
 6.12	 6.13	 6.14	21%

Table 6.1 *Yields for Staudinger ligations mediated by phosphines* ***6.3-5****.*

As predicted by molecular modeling, the yields of the Staudinger ligation using phenol-derived phosphines improved when *ortho*-substitution of the phenolic ring was employed. Presumably, the reactive atoms are being sterically forced together in such a manner as to make rearrangement more favorable. This work shows the utility of computational chemistry in directing design of improved chemical systems. In addition, it also begins to elucidate some important factors for designing efficient reagents for carrying out the Staudinger ligation at non-glycyl ligation sites. Although the phosphinophenols analyzed in this work do not likely represent the end-game for phosphine design for this reaction, they do highlight the importance of bringing the reactive groups near one another for efficient rearrangement to occur. The Staudinger ligation competes with simple hydrolysis as a side reaction, and steric effects like those described herein may represent a favorable method for tipping this balance away from

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

hydrolysis and towards ligation. Future efforts should focus on other potential factors, including how the electronic nature of the phosphine may affect the activation of the iminophosphorane nitrogen, and how varying the activation of the (thio)ester may improve ligation rates.

Experimental Section

General Experimental

Chemicals and solvents were from Aldrich® with the exception of dry THF, CH_2Cl_2, and DMF, which were from VWR®. Reactions were monitored by thin layer-chromatography using Whatman® TLC plates (AL SIL G/UV) and visualized by UV illumination or I_2 staining. NMR spectra were obtained using Bruker AC-300 or Varian UNITY-500 spectrometers. Phosphorus-31 NMR spectra were proton-coupled and referenced against an external standard of deuterated phosphoric acid unless stated otherwise. Mass spectra were obtained using electrospray ionization (ESI) techniques.

Molecular Modeling

Molecular modeling was carried out using the Monte-Carlo Multi-Minimum (MCMM) algorithm (Chang et al., 1989; Guida et al., 1992) in the program Macromodel (Mohamadi et al., 1990). Calculations were performed using the OPLS-AA force field (Jorgensen et al., 1996) on molecules en vacuo utilizing a dielectric constant of 11.4, which approximates the dielelectric constant of THF/water (3:1) (Kinart & Kinart, 1994).

2-*o*-Tolyloxy-tetrahydro-pyran (6.7). *o*-Cresol (**6.6**, 5.0 g, 46.2 mmol) was charged

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to a flame-dried round bottom flask under argon. 3,4-2*H*-Dihydropyran (12.6 mL, 139 mmol) and pyridinium *p*-toluenesulfonate (1.16 g, 4.62 mmol) were also added, and the resulting mixture was dissolved in dry methylene chloride (100 mL) and stirred for 18 h at room temperature. The mixture was then washed with half-saturated brine, and the aqueous phase was extracted with ethyl acetate (3 × 50 mL) and the combined organic phase was dried over anhydrous magnesium sulfate, filtered, and concentrated under reduced pressure. The residue was purified by flash chromatography (silica gel, 10% v/v ethyl acetate in hexanes). The desired THP-protected cresol was obtained as a clear liquid in 97% yield (8.57 g, 44.6 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.13-7.04 (m, 3 H), 6.86 (td, J = 7.2 Hz, J = 1.6 Hz, 1 H), 5.39 (t, J = 3.3 Hz, 1 H), 3.91-3.83 (m, 1 H), 3.59-3.53 (m, 1 H), 2.26 (s, 2 H), 2.06-1.95 (m, 1 H), 1.87-1.81 (m, 2 H), 1.68-1.53 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 154.95, 130.44, 126.97, 126.58, 121.01, 113.87, 95.86, 61.70, 30.43, 25.15, 18.78, 16.07 ppm; MS (EI) *m/z* 192.1150 ($M^{\cdot+}$ [$C_{12}H_{16}O_2$] = 192.1150).

[3-Methyl-2-(tetrahydro-pyran-2-yloxy)-phenyl]-diphenyl-phosphane (6.8). Compound **6.7** (1.0 g, 5.2 mmol) was charged to a flame-dried round bottom flask under an argon atmosphere and dissolved in dry THF (20 mL). This solution was cooled to -78 °C and *n*-butyllithium (1.6 M in hexanes, 3.9 mL, 6.2 mmol) was added dropwise over 10 min. The solution was stirred at -78 °C for 15 min, and then stirred at room temperature for 1 h. The solution was again cooled to -78 °C and chlorodiphenylphosphine (1.2 mL,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6.2 mmol) was added dropwise. The reaction was allowed to warm slowly to room temperature and stirred for 4 h. After 4 h the reaction was quenched by the addition of a solution of saturated sodium bicarbonate (15 mL). The organic layer was separated, and the water phase was extracted with ethyl acetate (3 × 15 mL). The combined organic phase was dried over anhydrous magnesium sulfate, filtered, and concentrated. The residue was purified by silica gel chromatography (10% v/v ethyl acetate in hexanes) to give phosphine **6.8** as a white solid in 64% yield (1.26 g, 3.35 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ7.32-7.28 (m, 10 H), 7.18 (d, J = 7.5 Hz, 1 H), 6.92 (t, J = 7.9 Hz), 6.67-6.63 (m, 1 H), 5.12-5.08 (m, 1 H), 4.04-3.97 (m, 1 H), 3.37-3.30 (m, 1 H), 2.38 (s, 3 H), 1.95-1.81 (m, 3 H), 1.51-1.42 (m, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ159.12 (d, J = 18.2 Hz), 137.24 (d, J = 11.5 Hz), 137.03 (d, J = 11.5 Hz), 134.47 (d, J = 20 Hz), 133.06 (d, J = 20 Hz), 132.31 (d, J = 12 Hz), 130.64 (d, J = 14.3 Hz), 128.53, 128.44, 128.34, 128.31, 128.26, 128.17, 124.24, 64.37, 30.83, 25.04, 20.31, 17.83 ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ –16.44 ppm; MS (EI) *m/z* 376.1592 ($M^{\cdot+}$ $[C_{24}H_{25}O_2P]$ = 376.1577).

2-Diphenylphosphanyl-6-methyl-phenol (6.4). Phosphine **6.8** (1.2 g, 3.2 mmol) was dissolved in deoxygenated methanol (23 mL) under argon. *p*-Toluenesulfonic acid (910 mg, 4.8 mmol) was then added. The reaction mixture was stirred under argon at room temperature for 2 h and then quenched by the addition of brine (23 mL). The organic layer was washed with 2 N HCl (15 mL) and brine (15 mL), dried over anhydrous magnesium sulfate, filtered, and concentrated under reduced pressure. The residue was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

purified by flash chromatography (silica gel, 10% v/v ethyl acetate in hexanes) to give the substituted phosphinophenol **6.4** as a white solid in 64% yield (596 mg, 2.04 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.33-7.28 (m, 10 H), 7.14-7.11 (m, 1 H), 6.86-6.81 (m, 1 H), 6.77 (t, J = 7.1 Hz, 1 H), 6.43 (d, J = 7.9 Hz, 1 H), 2.42 (s, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 157.50 (d, J = 16.8 Hz), 135.07 (d, J = 4.3 Hz), 133.30 (d, J = 18.8 Hz), 132.79, 132.22 (d, J = 3.6 Hz), 128.85, 128.58 (d, J = 7.9 Hz), 124.42, 120.62 (d, J = 2.5 Hz), 120.01 (d, J = 3.5 Hz), 16.17 (d, J = 2.1 Hz) ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ –29.65 ppm; MS (ESI) *m/z* 291.0950 (M – H$^-$ [$C_{19}H_{17}OP$] = 291.0939).

***N*-Acetylalanine 2-diphenylphosphanyl-phenyl ester (6.10)**. *N*-Acetylalanine (500 mg, 3.8 mmol), phosphine **6.3** (1.06 g, 3.8 mmol), and PyBOP (1.98 g, 3.8 mmol) were charged to a flame-dried round bottom flask under argon. These reagents were dissolved in dry DMF (16 mL), after which diisopropylethyl amine was added (1.33 mL, 7.6 mmol). This solution was stirred at room temperature for 10 h. The solvent was removed under reduced pressure, and the residue was taken up in ethyl acetate. This organic solution was washed with 2 N HCl (2 × 25 mL), saturated sodium bicarbonate (2 × 25 mL), and brine (25 mL). The organic layer was dried over magnesium sulfate, filtered, and concentrated *en vacuo*. The residue was then purified by flash chromatography (silica gel, 70% v/v ethyl acetate in hexanes) to give the alanyl ester **6.10** as a white solid in 66% yield (983 mg, 2.5 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.58 (d, J = 7.1 Hz), 7.40-7.25 (m, 10 H), 7.21-7.17 (m, 1 H), 7.11 (t, J = 7.6 Hz, 1 H), 6.81-6.77 (m, 1 H), 4.57-4.48 (m, 1 H), 1.94

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(s, 3 H), 1.29 (d, J = 8.7 Hz) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ170.95, 170.69, 152.29 (d, J = 17 Hz), 134.91 (d, J = 5.6 Hz), 134.79 (d, J = 5.6 Hz), 133.35 (d, J = 19.7 Hz), 133.32, 129.69, 129.35 (d, J = 15 Hz), 128.63 (d, J = 2.8 Hz), 128.17 (d, J = 6.5 Hz), 126.02, 121.80, 48.13, 21.61, 16.27 ppm; ^{31}P NMR (121 MHz, $CDCl_3$) δ –17.09 ppm; MS (ESI) *m/z* 414.1215 (MNa^+ [$C_{23}H_{22}NO_3PNa$] = 414.1235).

***N*-Acetylalanine 2-diphenylphosphanyl-6-methyl-phenyl ester (6.11).** *N*-Acetylalanine (267 mg, 2.04 mmol), phosphine **6.4** (596 mg, 2.04 mmol), and PyBOP (1.06 g, 2.04 mmol) were charged to a flame-dried round bottom flask under argon. The reagents were dissolved in dry DMF (8.5 mL), after which diisopropylethyl amine was added (0.71 mL, 4.08 mmol). This solution was stirred at room temperature for 12 h. The solvent was removed under reduced pressure, and the residue was dissolved in ethyl acetate. This organic solution was washed with 2 N HCl (2 × 25 mL), saturated sodium bicarbonate (2 × 25 mL), and brine (25 mL). The organic layer was dried over magnesium sulfate, filtered, and concentrated *en vacuo*. The residue was then purified by flash chromatography (silica gel, 70% v/v ethyl acetate in hexanes) to give the alanyl ester **6.11** as a white solid in 64% yield (528 mg, 1.30 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$:CD_3OD (1:1)) δ7.55 (bs, 1 H), 7.33-7.27 (m, 6 H), 7.27-7.22 (m, 5 H), 7.05 (t, J = 8.3 Hz, 1 H), 6.65-6.61 (m, 1 H), 4.69-4.60 (m, 1 H), 2.18 (s, 3 H), 1.94 (s, 3 H), 1.42 (d, J = 7.4 Hz, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$:CD_3OD (1:1)) δ170.86, 170.28, 151.18 (d, J = 19.1 Hz), 135.20 (d, J = 9.7 Hz), 135.09 (d, J = 9.7 Hz), 133.49 (d, J = 9.7 Hz), 133.20 (d, J = 10.7 Hz), 131.80, 131.26,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

130.49, 129.64 (d, J = 14 Hz), 128.52, 128.10 (d, J = 7.5 Hz), 126.13, 47.86, 21.66, 16.61, 15.71 ppm; ^{31}P NMR (121 MHz, $CDCl_3$:CD_3OD (1:1)) δ –17.71 ppm; MS (ESI) *m/z* 428.1404 (MNa^+ [$C_{24}H_{24}NO_3PNa$] = 428.1392).

***N*-Acetylalanine 2-diphenylphosphanyl-naphthalen-1-yl ester (6.12).** *N*-Acetylalanine (315 mg, 2.40 mmol), phosphine **6.4** (788 mg, 2.40 mmol), and PyBOP (1.25 g, 2.40 mmol) were charged to a flame-dried round bottom flask under argon. The reagents were dissolved in dry DMF (10 mL), after which diisopropylethyl amine was added (0.84 mL, 4.80 mmol). This solution was stirred at room temperature for 12 h. The solvent was removed under reduced pressure and the residue was dissolved in ethyl acetate. This organic solution was washed with 2 N HCl (2 × 25 mL), saturated sodium bicarbonate (2 × 25 mL), and brine (25 mL). The organic layer was dried over magnesium sulfate, filtered, and concentrated *en vacuo*. The residue was then purified by flash chromatography (silica gel, 75% v/v ethyl acetate in hexanes) to give the alanyl ester **6.12** as a white solid in 55% yield (591 mg, 1.34 mmol).

Spectral data. ^{1}H NMR (300 MHz, $CDCl_3$:CD_3OD) δ 8.05 (d, J = 6.3 Hz, 1 H), 7.70-7.67 (m, 1 H), 7.51 (d, J = 8.3 Hz, 2 H), 7.39-7.35 (m, 3 H), 7.20-7.14 (m, 9 H), 6.86-6.82 (m, 1 H), 4.65-4.56 (m, 1 H), 1.76 (s, 3 H), 1.40 (d, J = 7.4 Hz) ppm; ^{13}C NMR (75 MHz, THF-d_6) δ 172.25, 170.32, 152.36 (d, J = 21 Hz), 137.68 (d, J = 4.6 Hz), 137.53 (d, J = 4.6 Hz), 136.26, 134.72 (d, J = 9.1 Hz), 134.45 (d, J = 9.1 Hz), 130.26, 129.72, 129.62, 129.58, 129.51, 129.49, 129.43, 128.38 (d, J = 3.6 Hz), 128.35 (d, J = 21.1 Hz), 127.89, 127.20, 126.75 (d, J = 14.8 Hz), 124.07, 49.53, 22.52, 17.72 ppm; ^{31}P NMR (121

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MHz, $CDCl_3$:CD_3OD) δ –18.78 ppm; MS (ESI) *m/z* 442.1568 (MH^+ [$C_{27}H_{25}NO_3P$] = 442.1572).

***N*-Azidoalanine benzamide (6.13)**. Azido alanine (Lundquist & Pelletier, 2001) and the benzamide (Nilsson et al., 2001; Soellner et al., 2002) thereof were synthesized using the protocols described previously.

Spectral data. ^1H NMR (300 MHz, $CDCl_3$) δ 7.34-7.21 (m, 5 H), 6.99 (bs, 1 H), 4.38 (d, J = 6 Hz, 2 H), 3.99 (q, J = 6.9 Hz, 1 H), 1.49 (d, J = 6.9 Hz, 3 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.72, 137.52, 128.45, 127.37, 127.32, 58.76, 43.17, 16.84 ppm; MS (ESI) *m/z* 205.1079 (MH^+ [$C_{10}H_{13}N_4O$] = 205.1089).

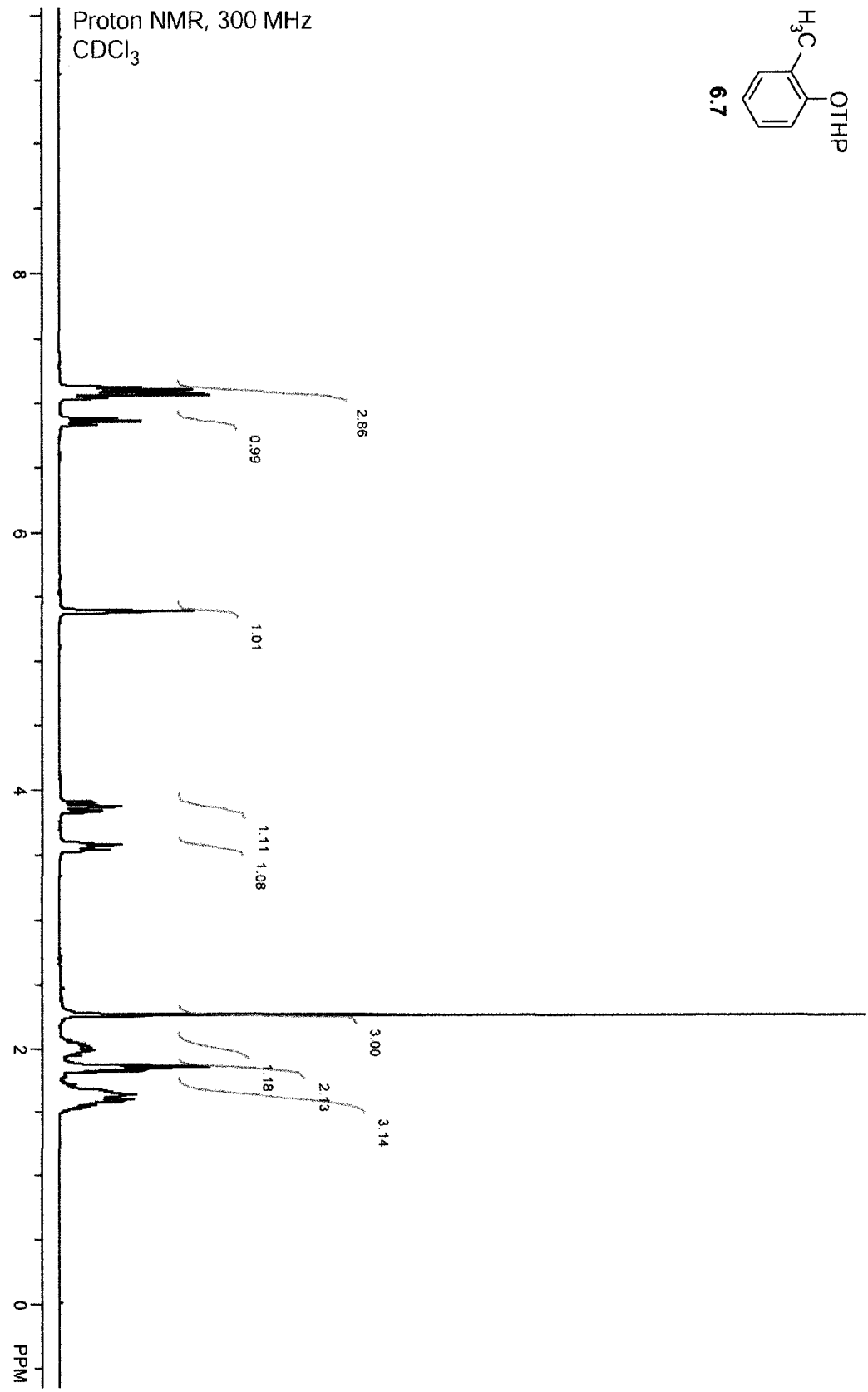
Staudinger ligations. The Staudinger ligation trials were carried out by dissolving the respective phosphine-ester (**6.10-12**, 0.5 mmol) in dry THF or DMF (5 mL) and heating to 47 °C. One equivalent of N_3AlaNHBn (100 mg, 0.5 mmol) was then added and the resulting solution was heated for 10 h. Water (1 mL) was then added and heating was allowed to proceed for an additional 1 h. A small amount of 2 N HCl (0.1 mL) was then added, and the solvents were removed under reduced pressure. The residue was dissolved in methylene chloride and purified by flash chromatography (silica gel, 5–10% v/v methanol in methylene chloride) to give AcAlaAlaNHBn in various yields.

Spectral data. ^1H NMR (300 MHz, $CDCl_3$:CD_3OD (1:1)) δ 7.32-7.35 (m, 5 H), 4.51-4.22 (m, 4 H), 1.98 (1.87) (s, 3 H), 1.41 (d, J = 7.5 Hz, 3 H), 1.34 (d, J = 6.9 Hz, 3 H)

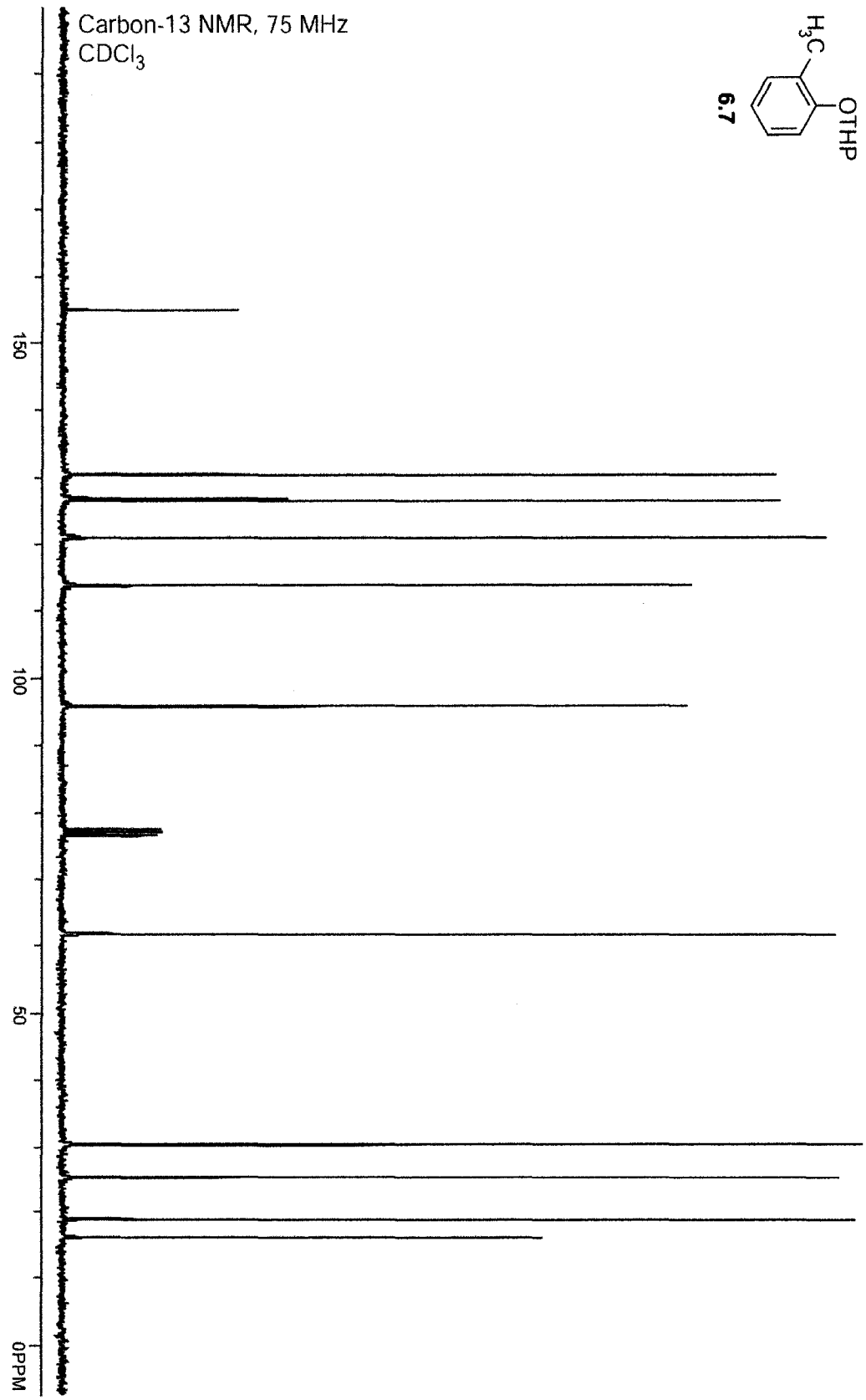
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ppm; ^{13}C NMR (75 MHz, $CDCl_3$:CD_3OD (1:1), numbers in parentheses indicate doubling due to rotational isomerism) δ173.27 (172.78), 172.67 (172.54), 171.44, 137.59 (137.51), 127.39 (127.49), 126.28 (126.36), 126.02 (126.17), 48.92 (48.70), 48.59 (48.34), 41.99 (42.06), 20.50 (20.76), 15.99 (16.23), 15.45 ppm; MS (ESI) *m/z* 314.1481 (MNa^+ [$C_{15}H_{21}N_3O_3Na$] = 314.1482).

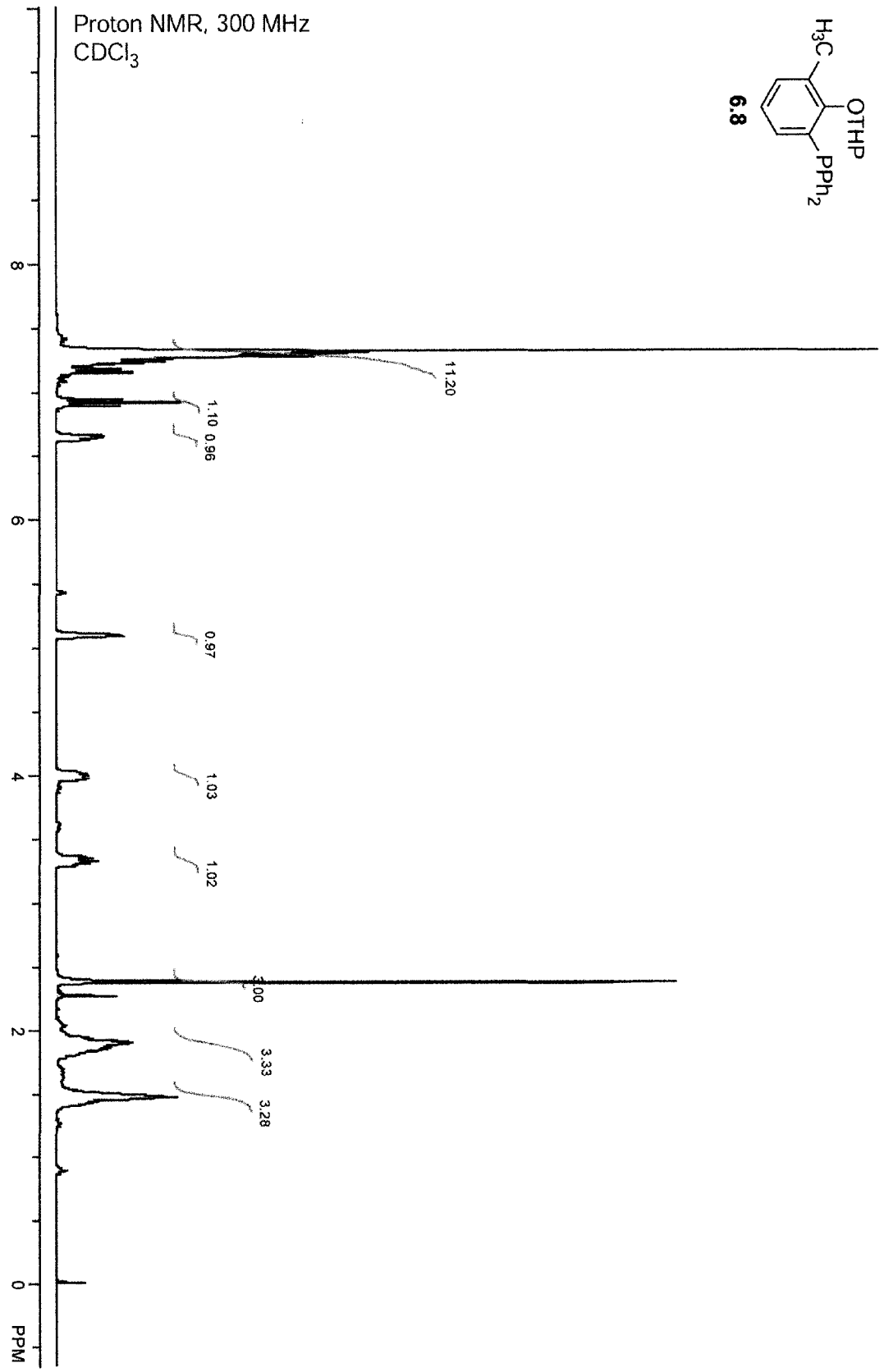
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



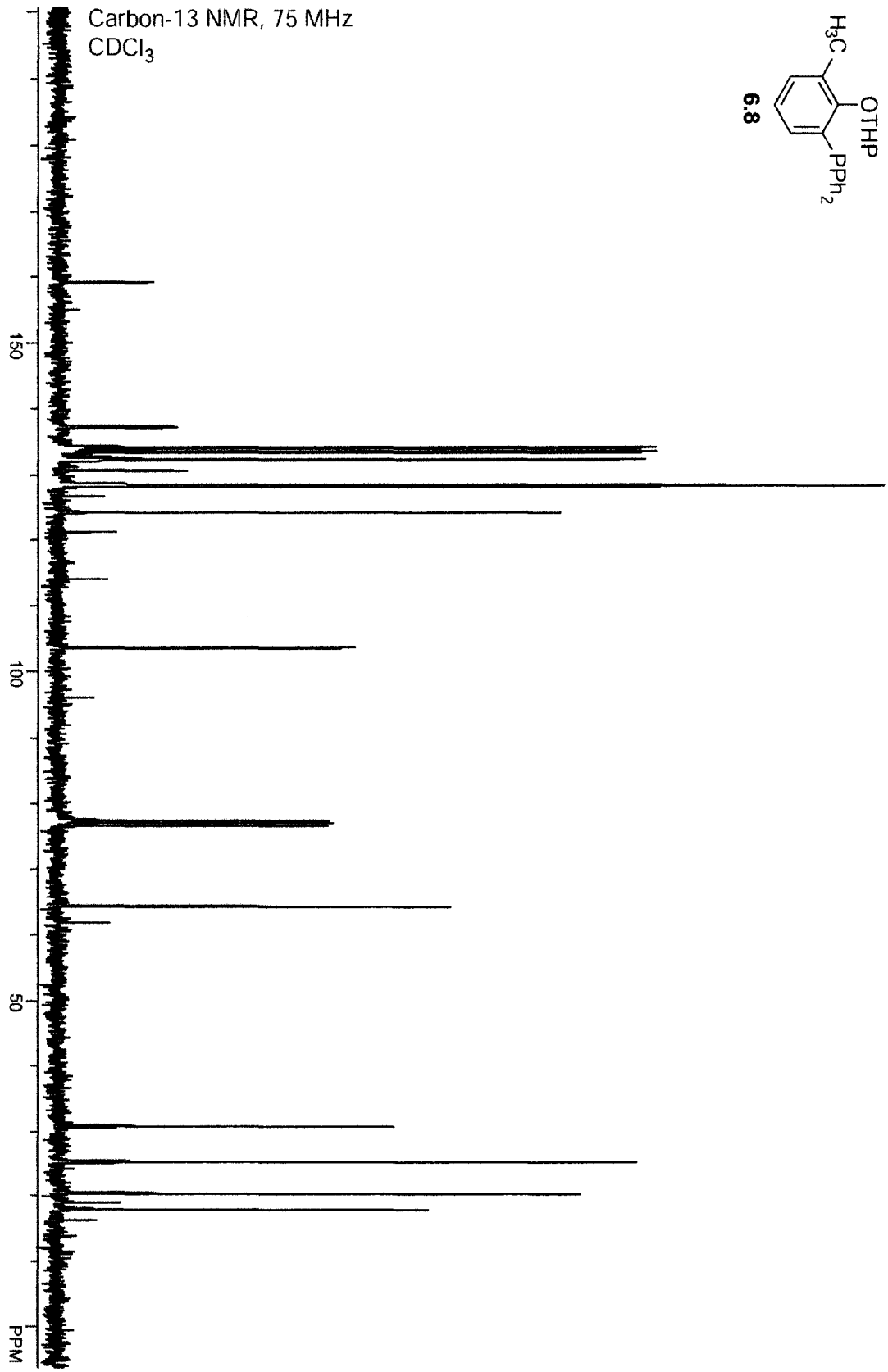
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



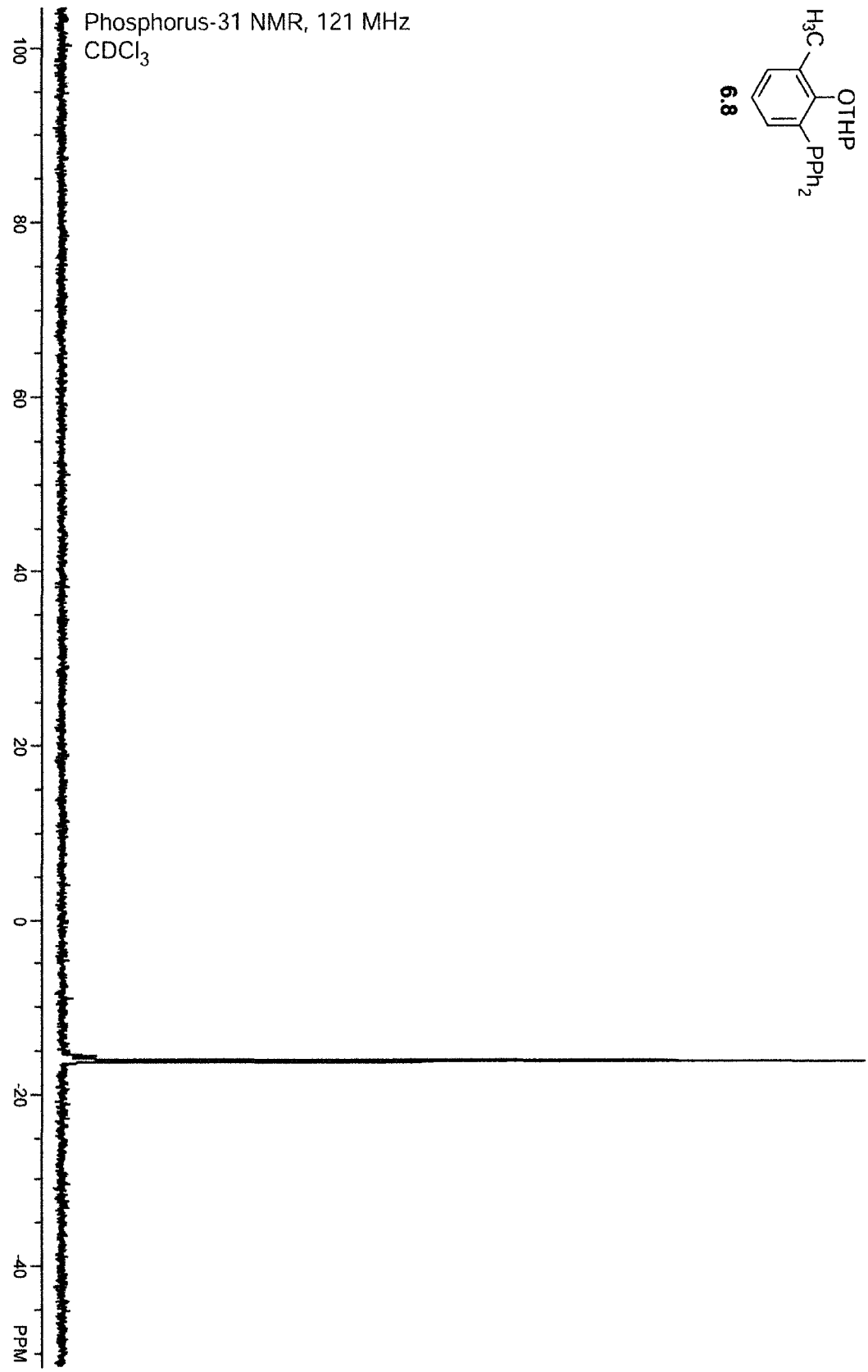
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



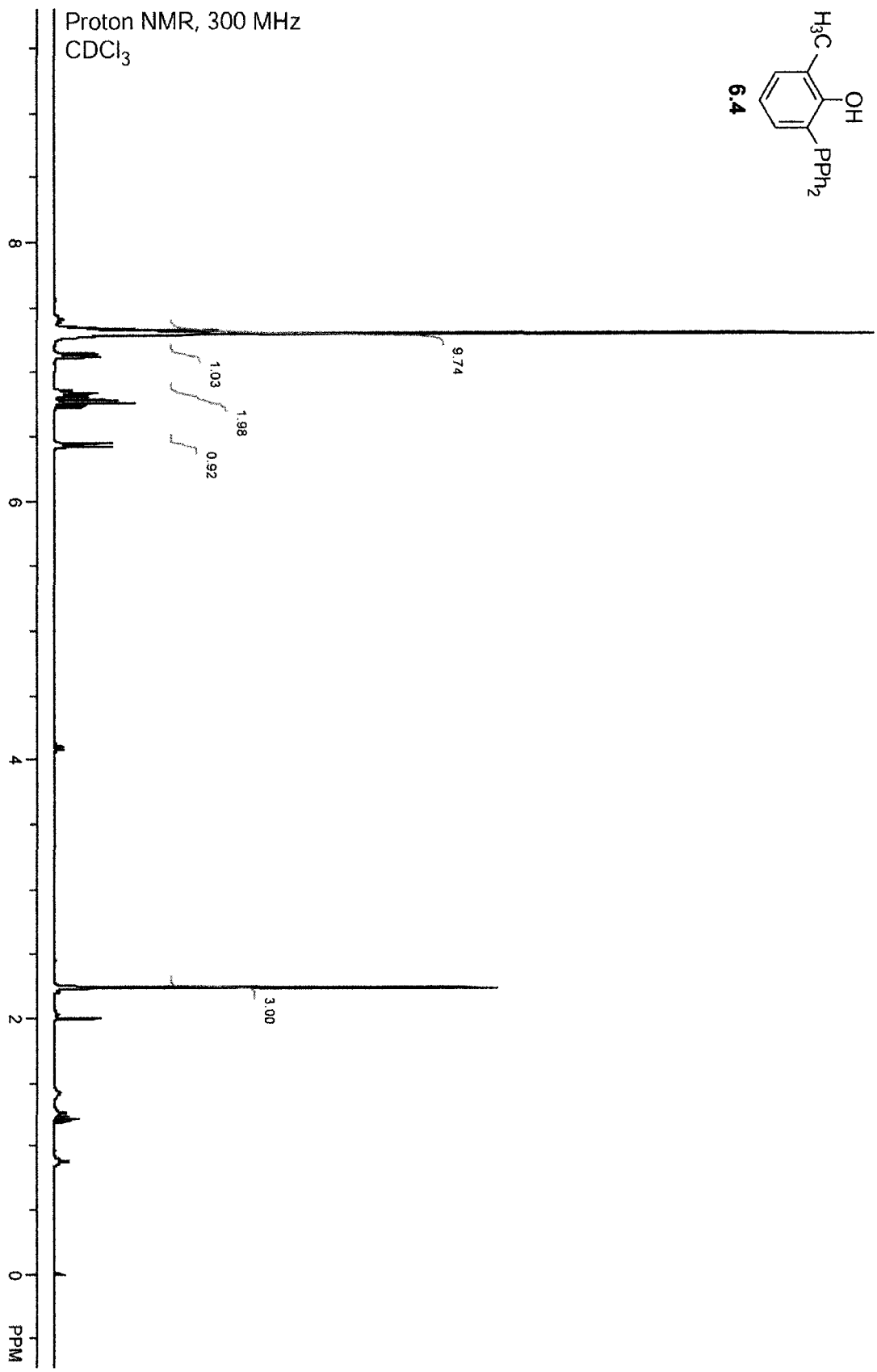
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



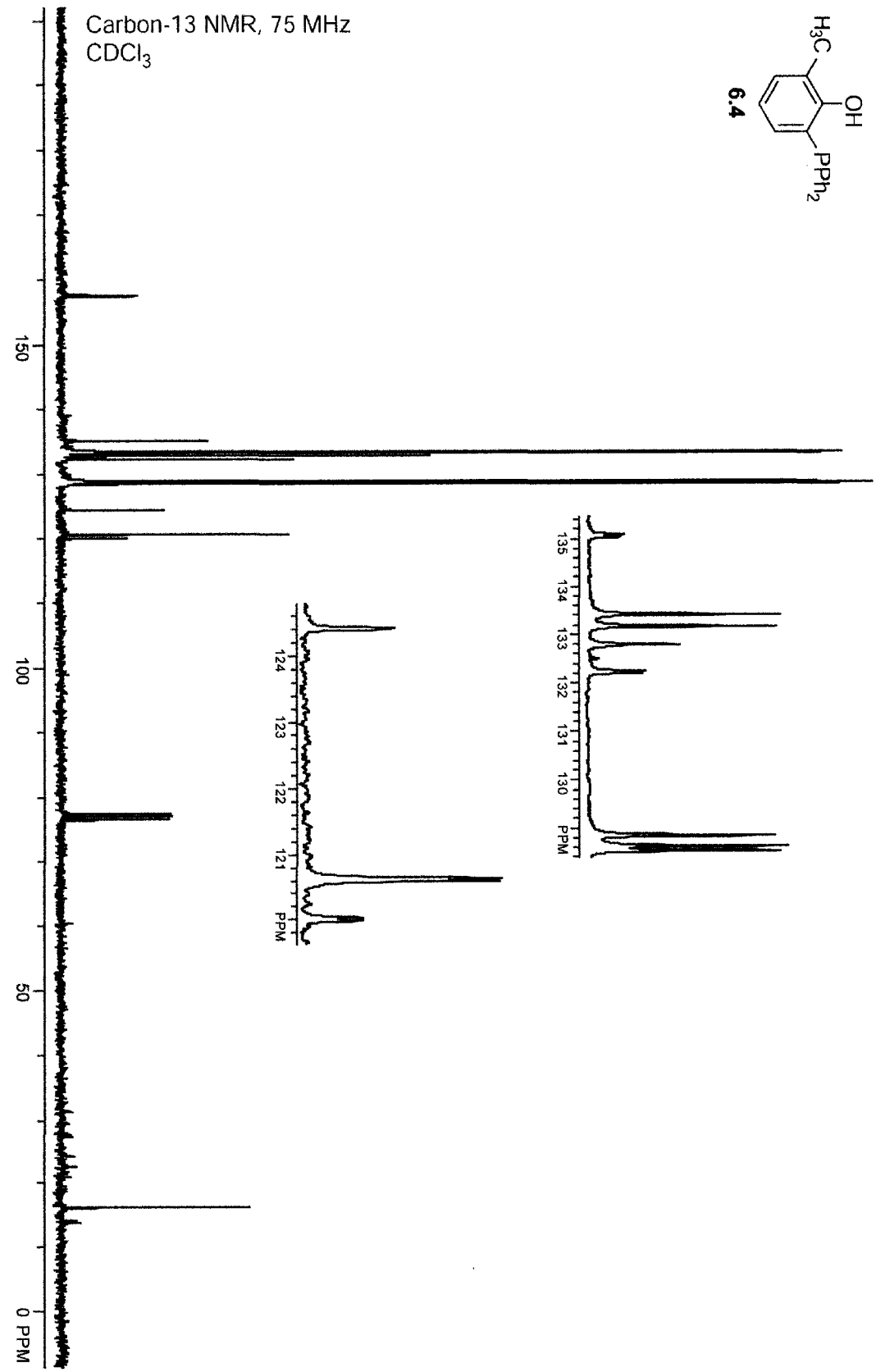
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



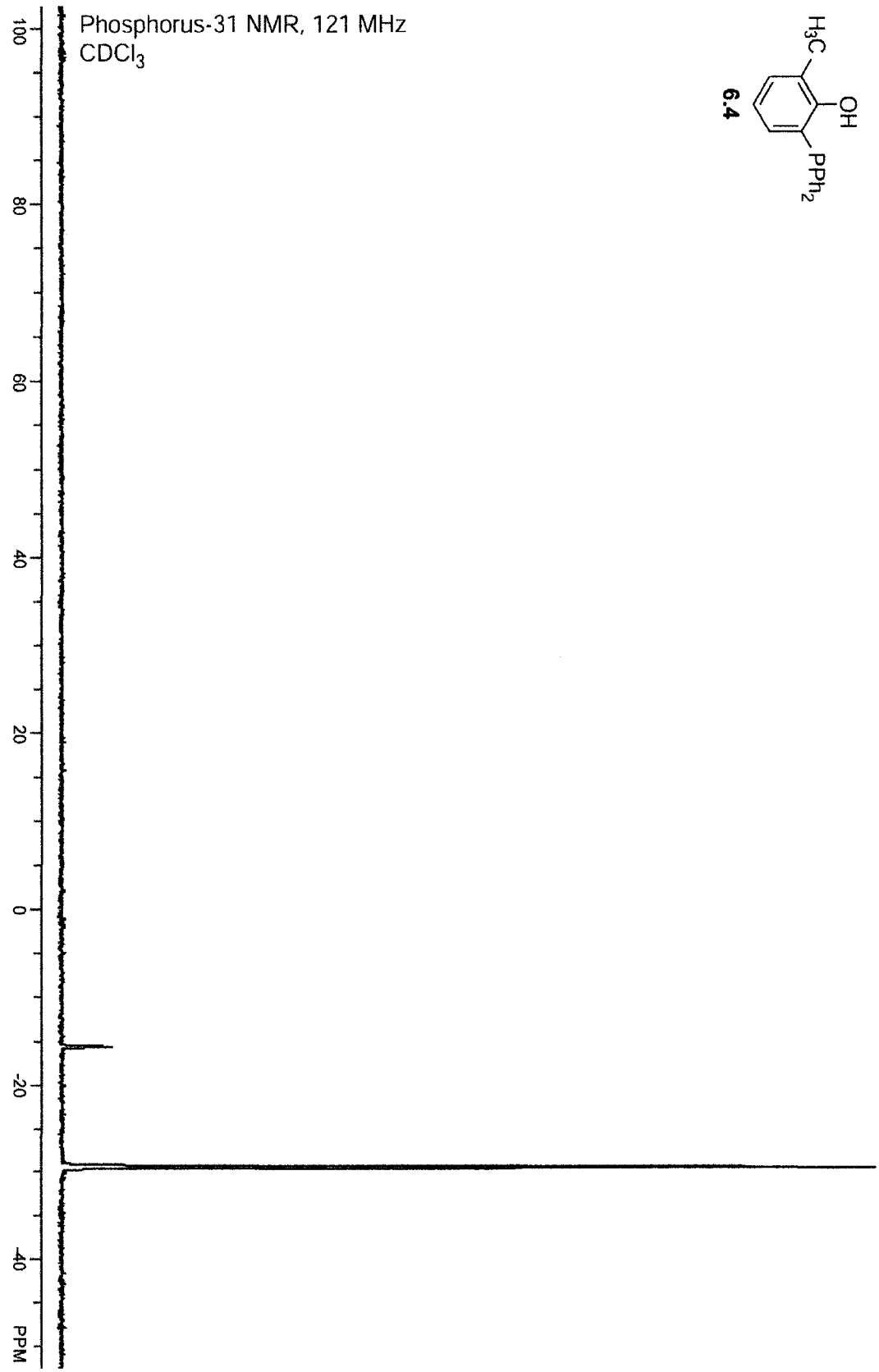
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



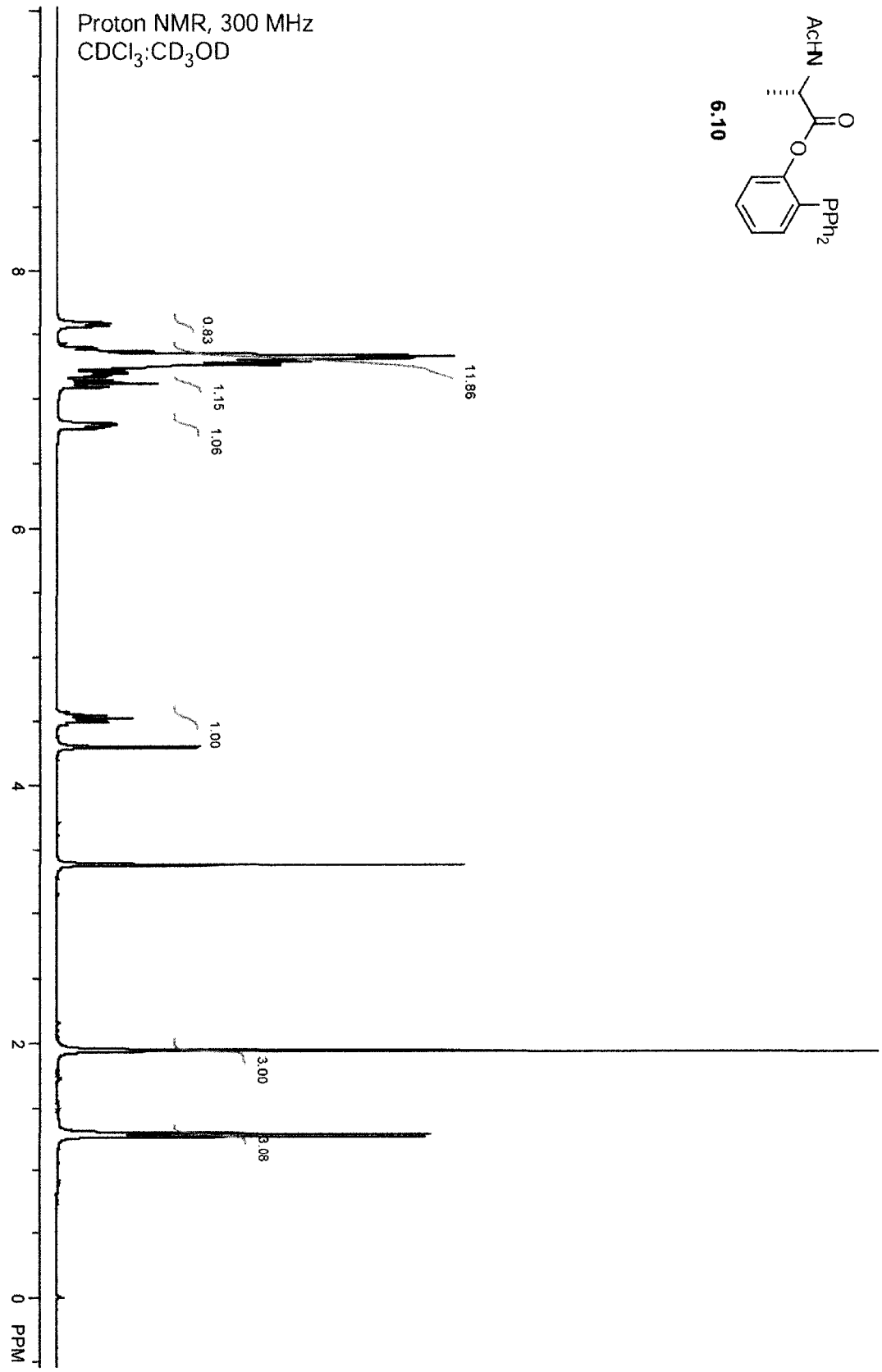
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



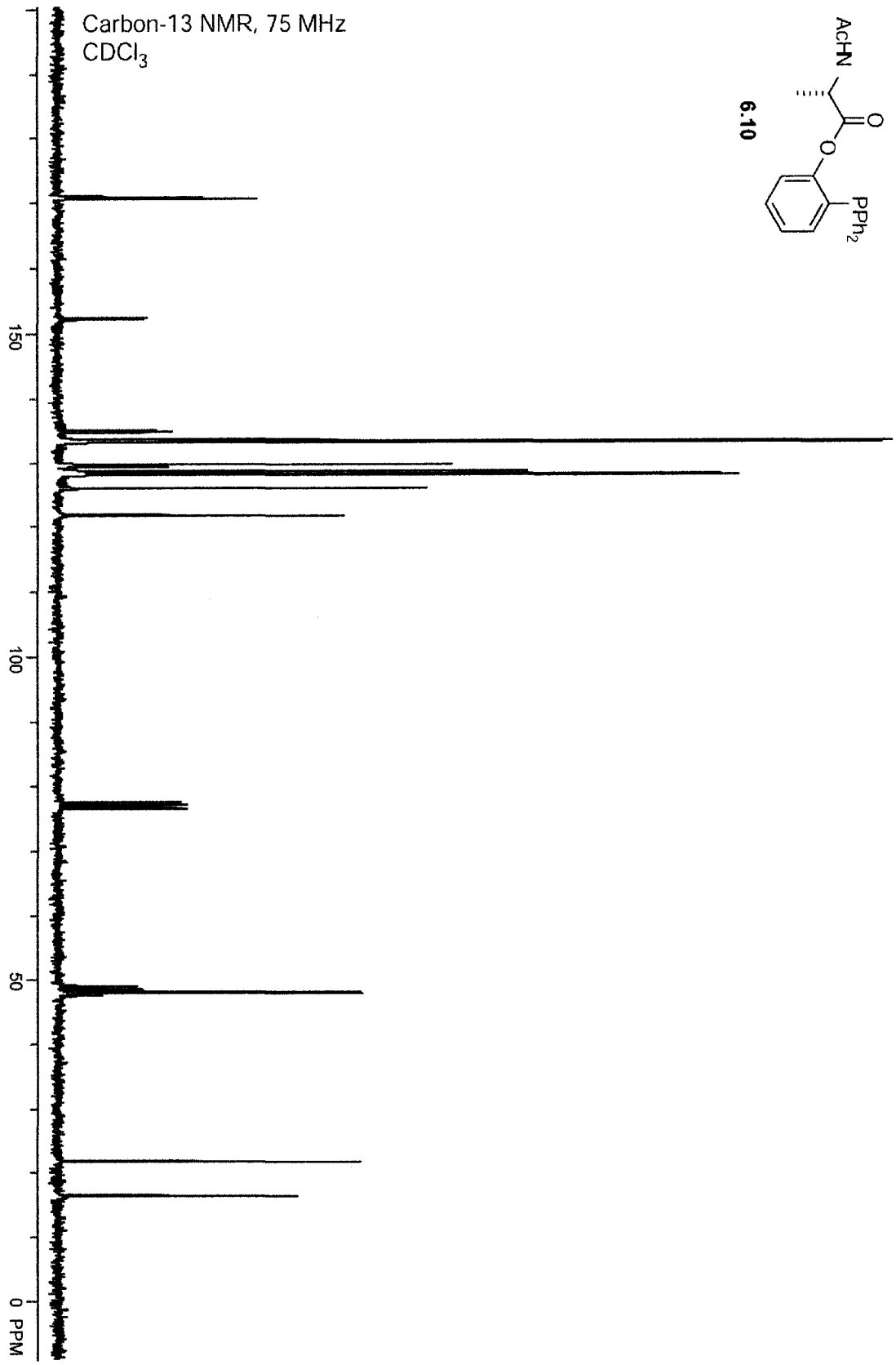
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



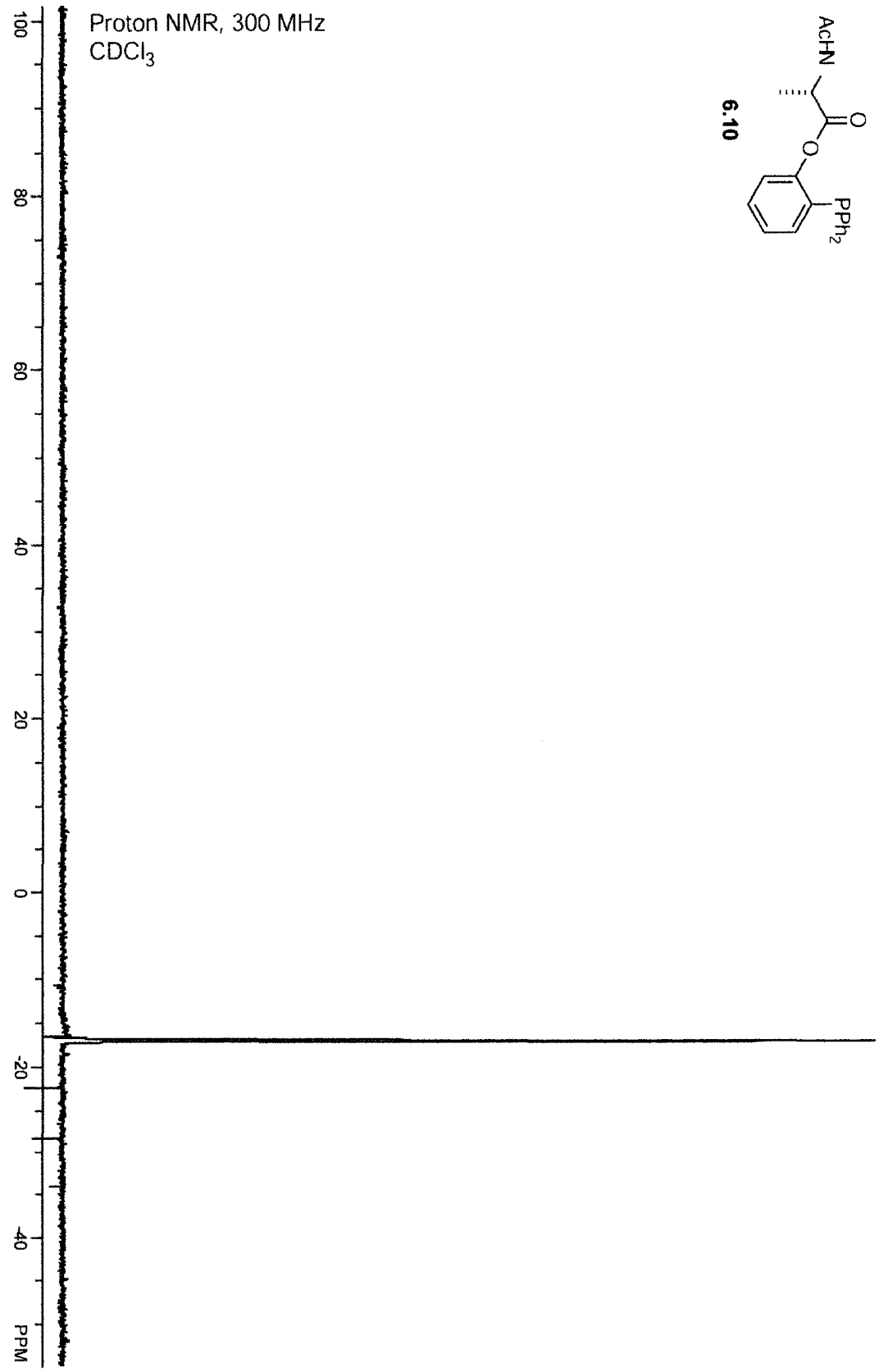
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



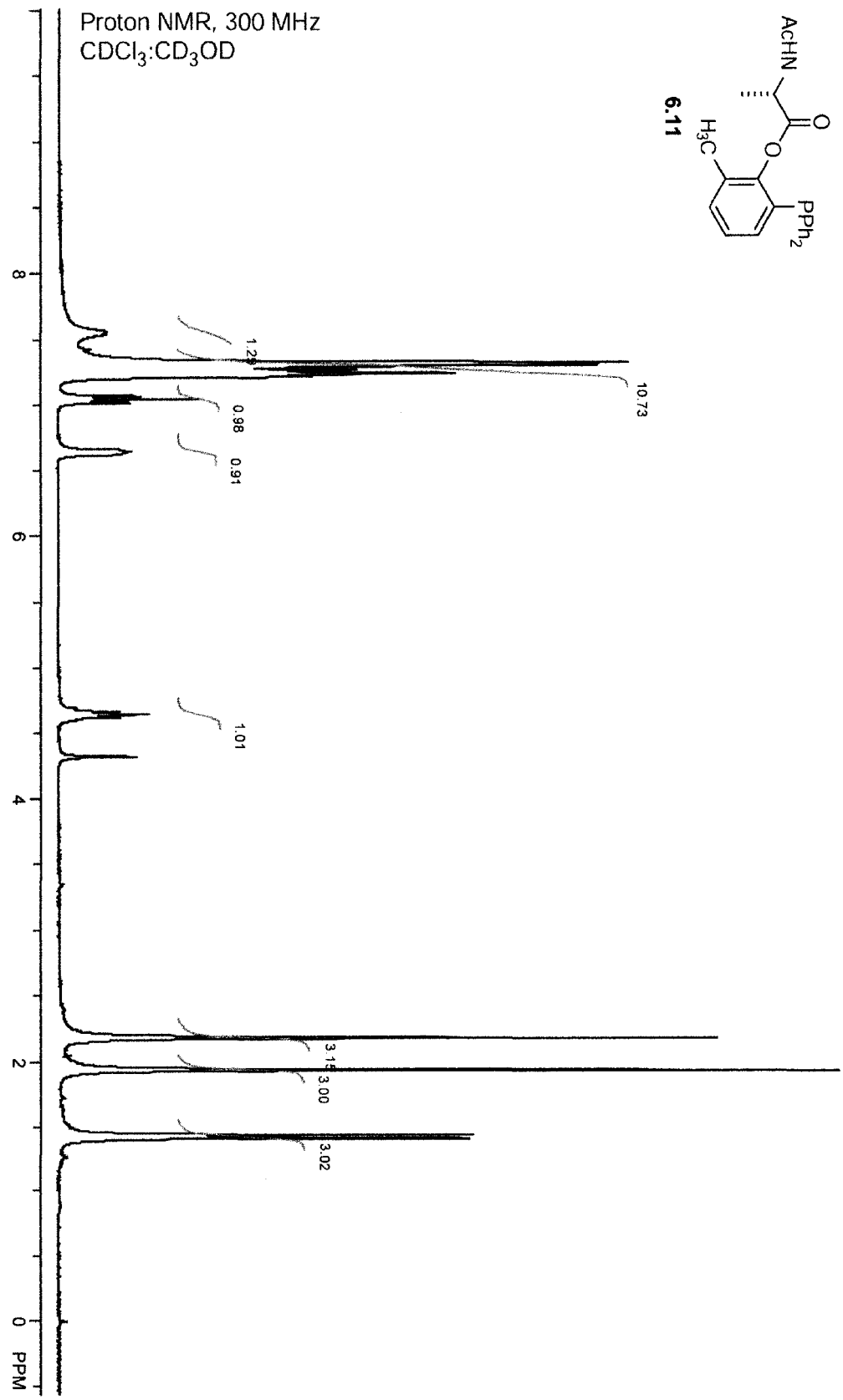
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



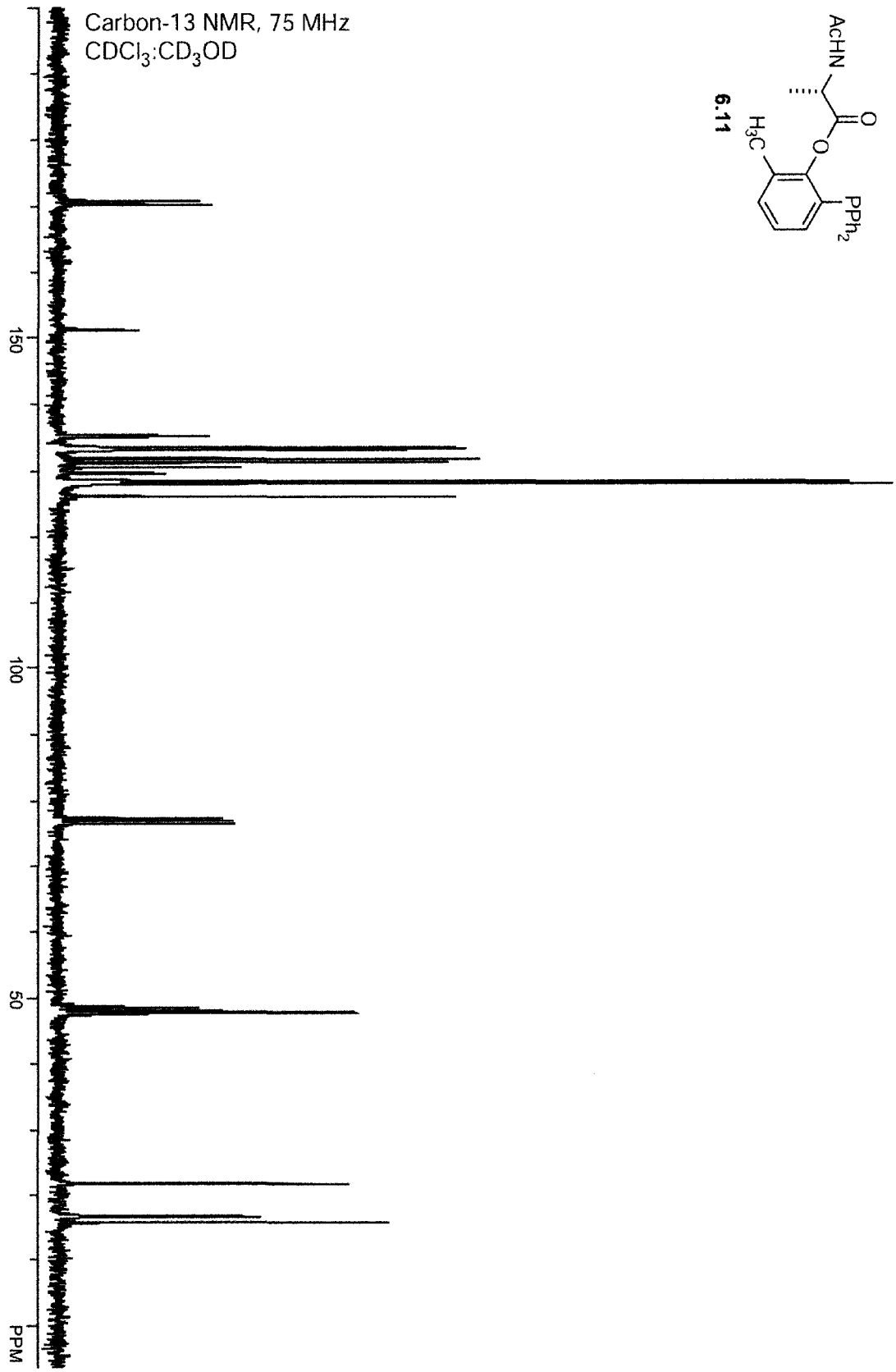
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



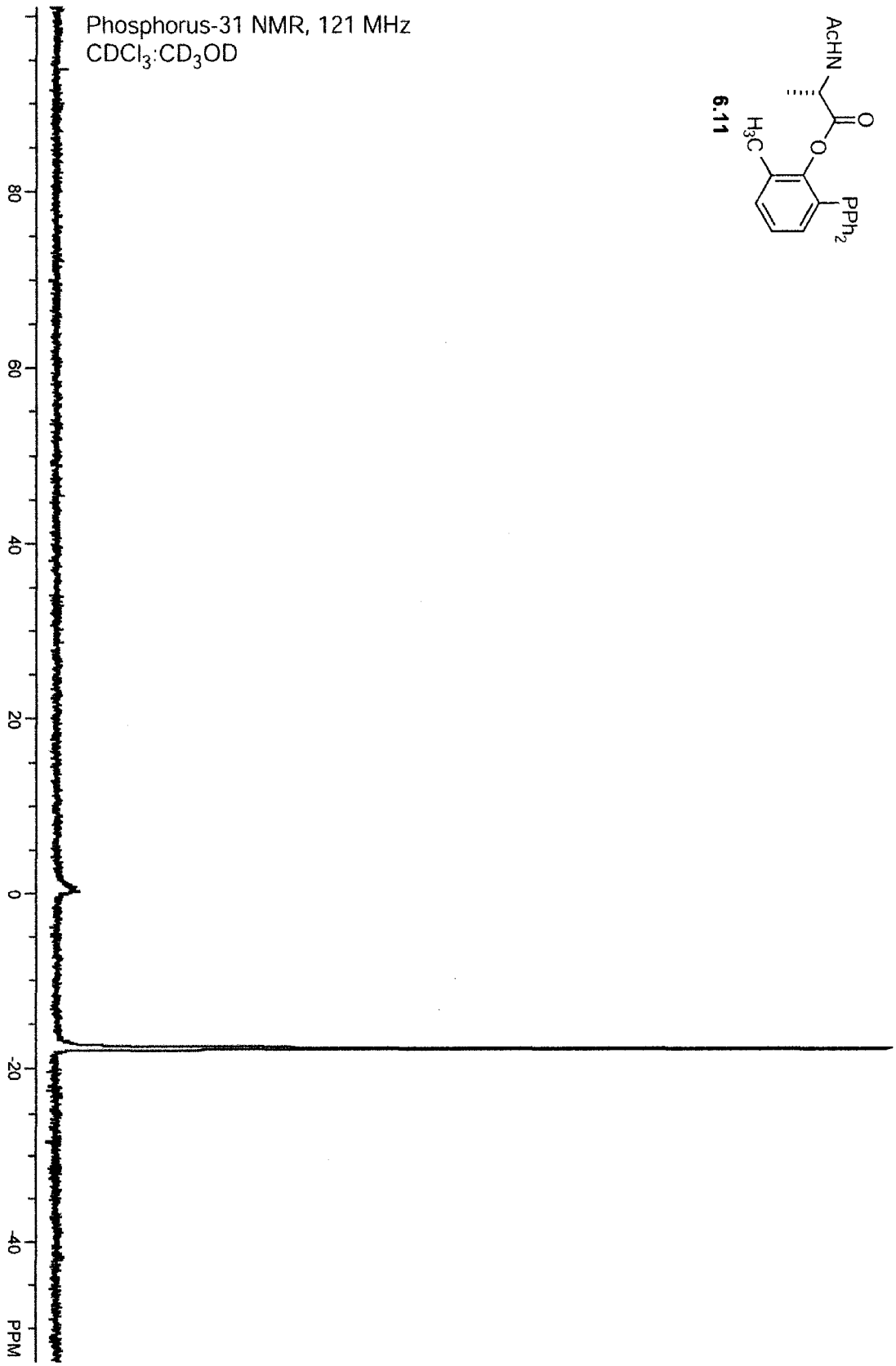
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

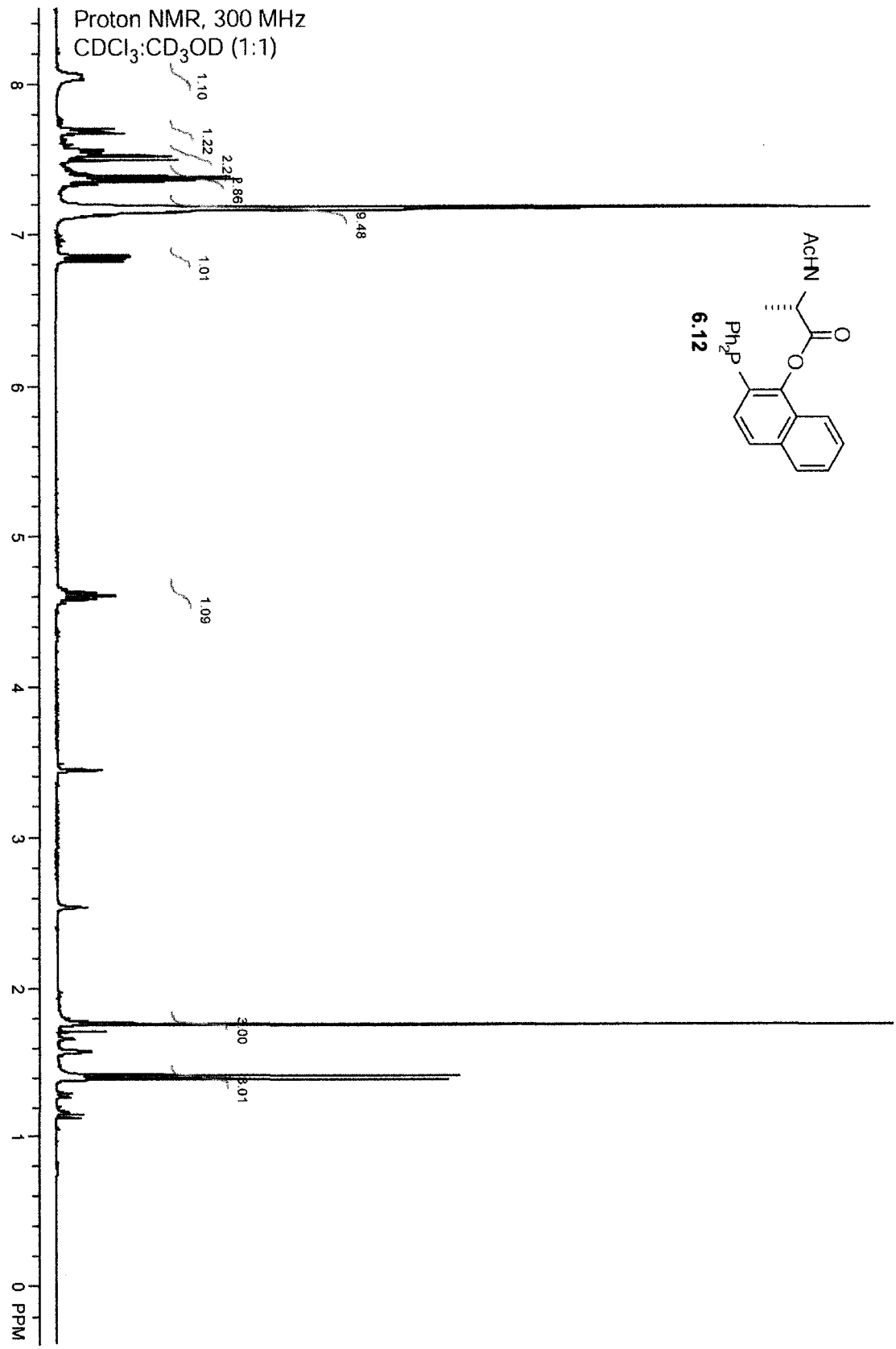


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

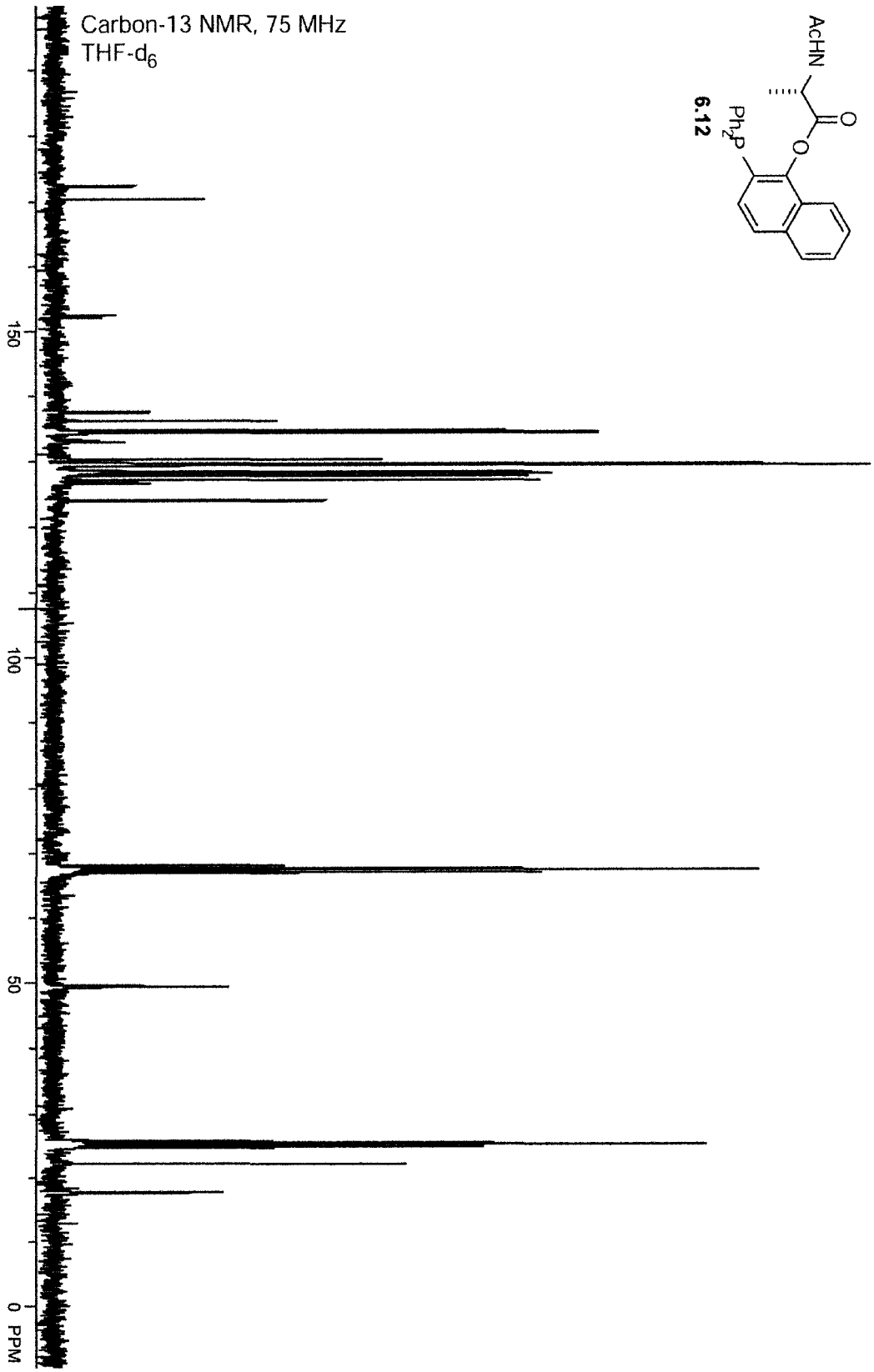


6.11

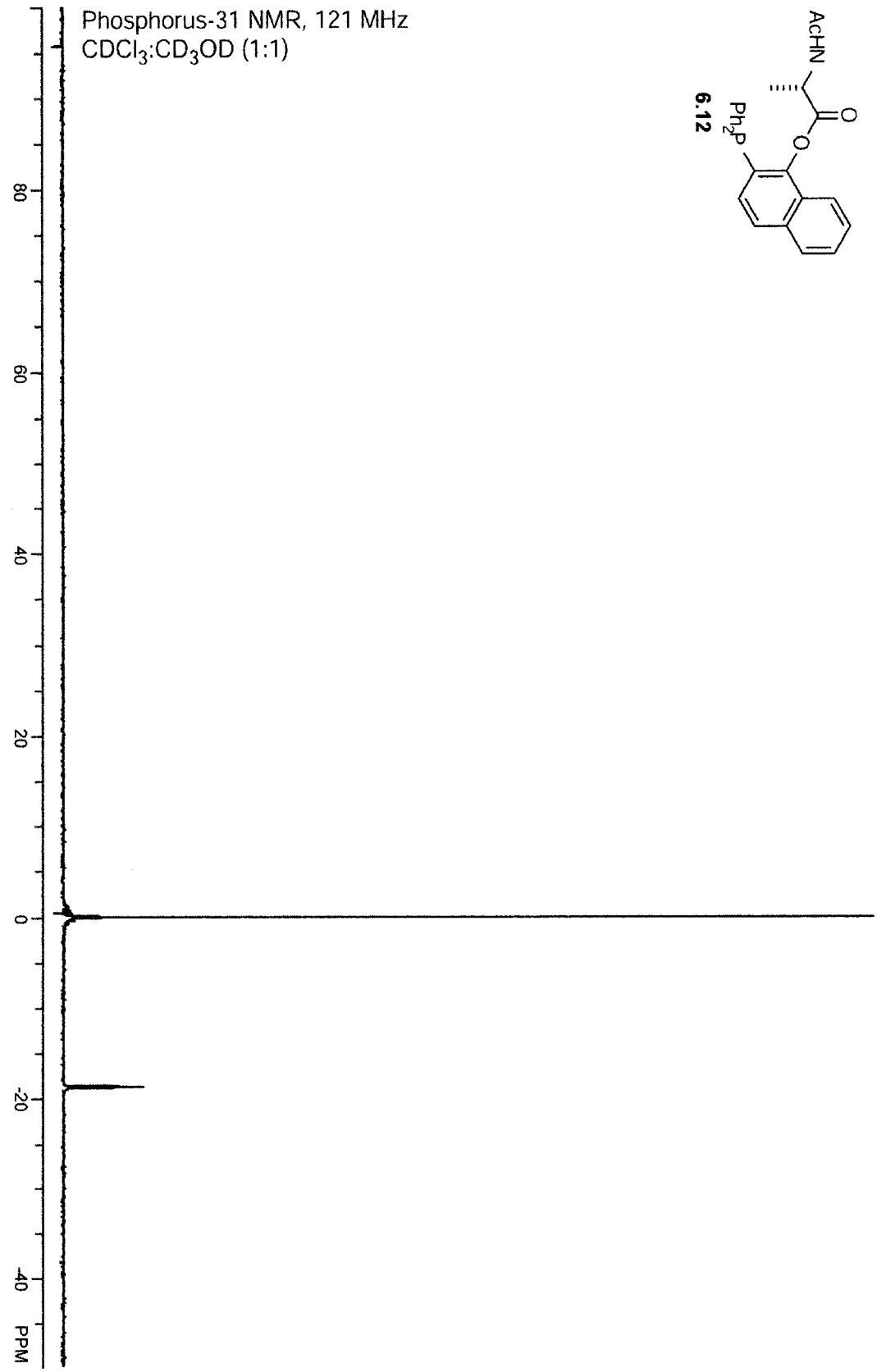
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



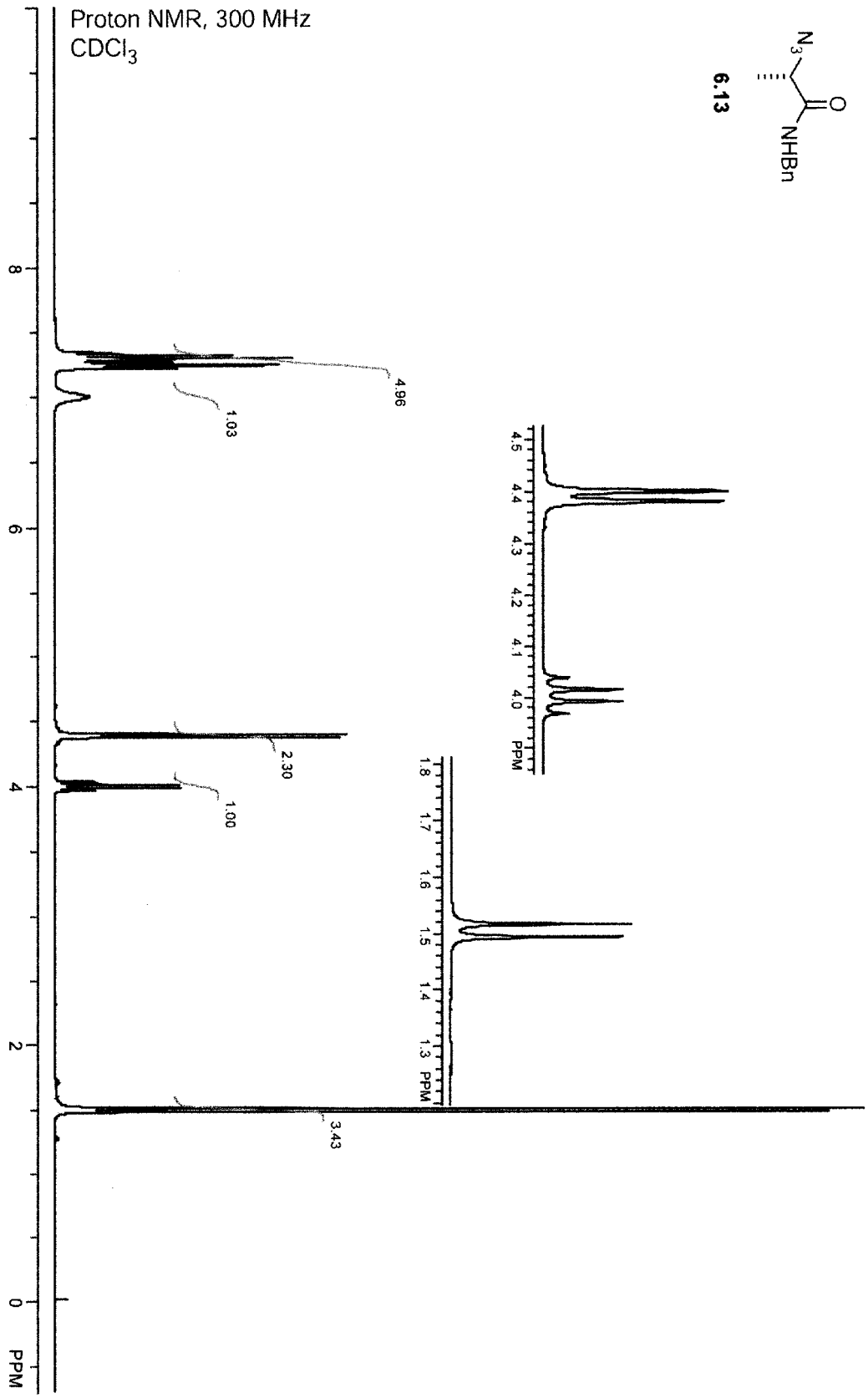
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



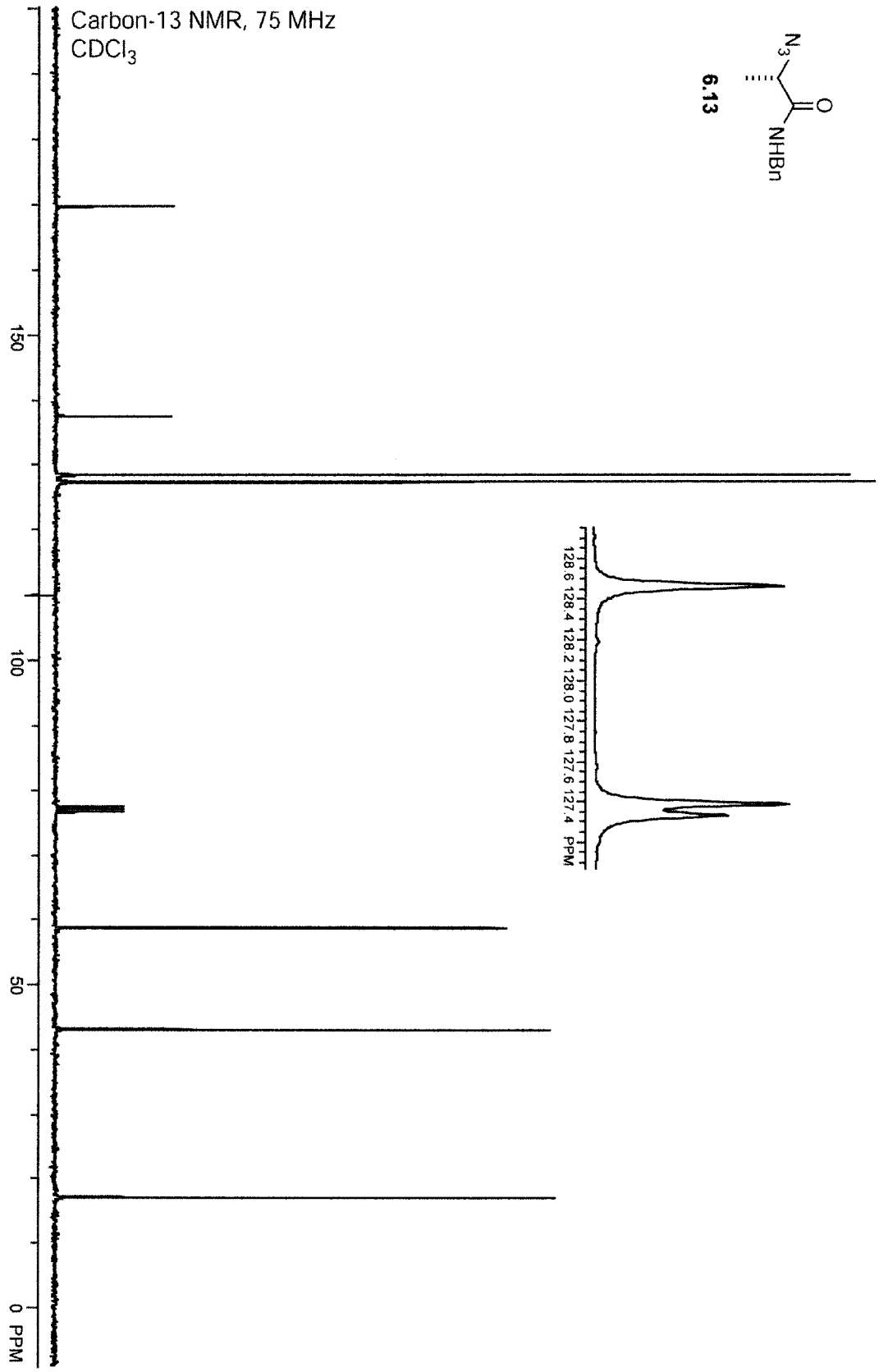
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



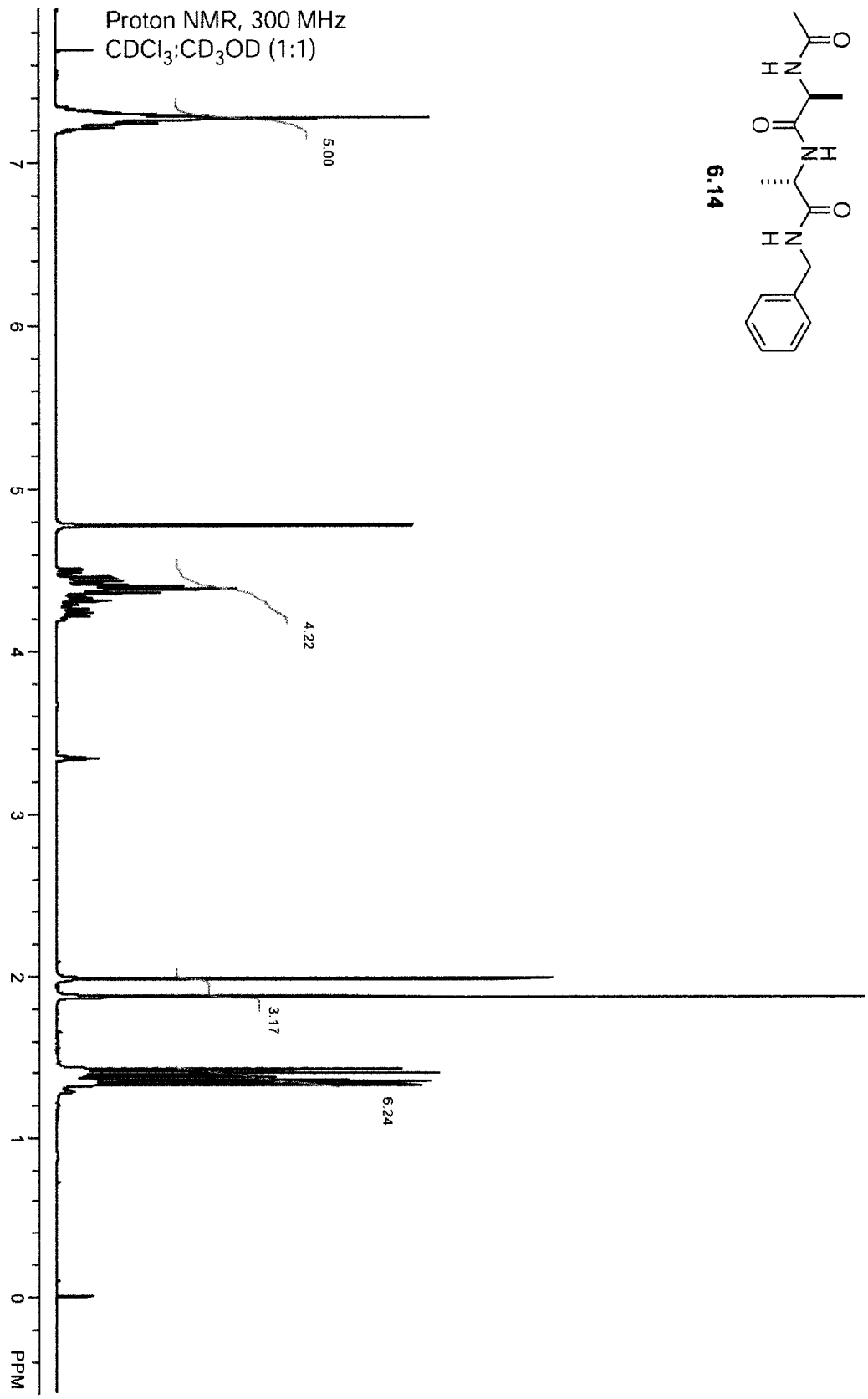
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



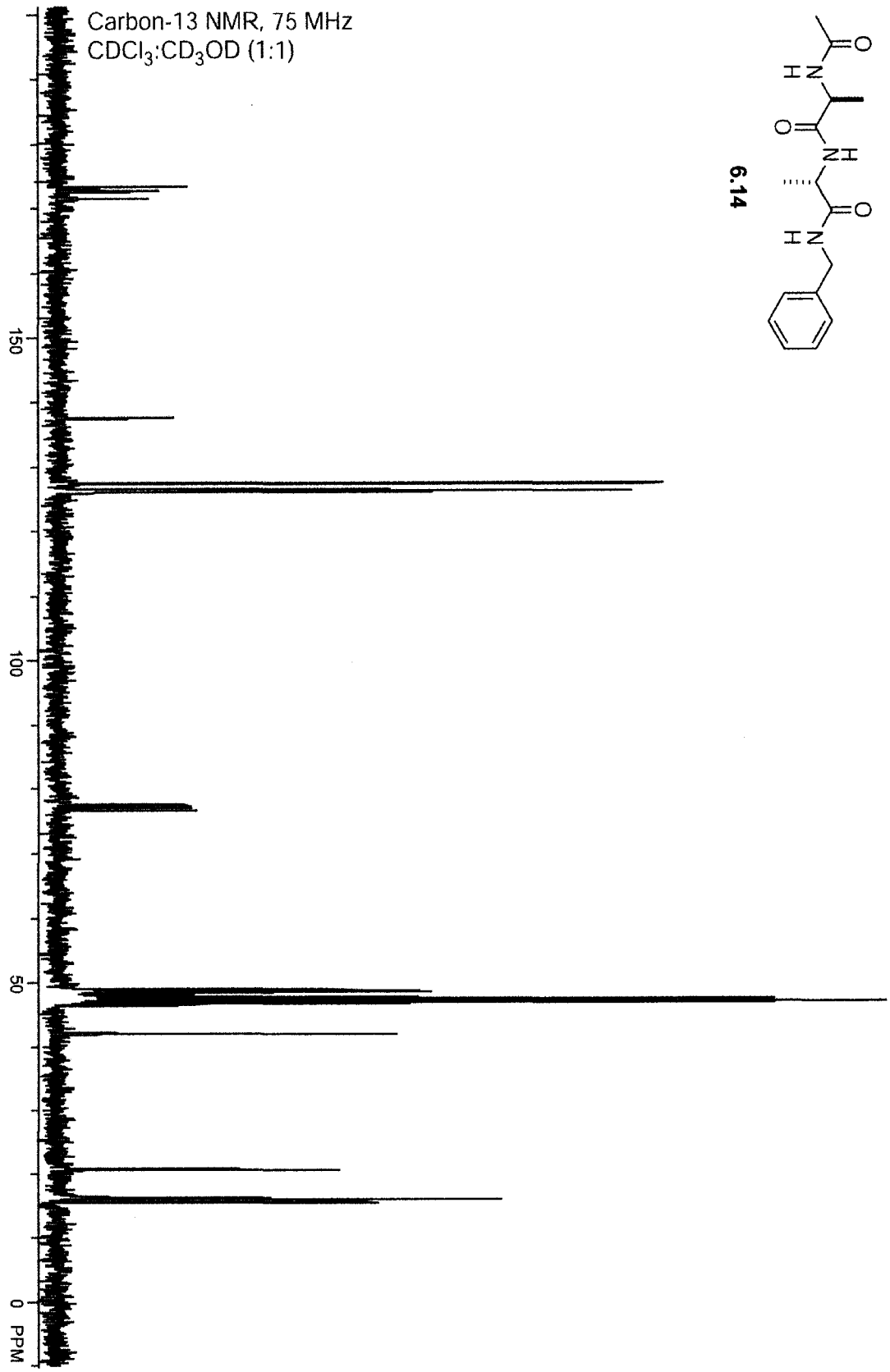
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



6.14

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References

Afonso, C. A. M. (1998). Studies on the transformation of azido-group to N-(t-butoxycarbonyl)amino group via Staudinger reaction. *Synth. Commun.* **28**, 261-276.

Aimoto, S. (1999). Polypeptide synthesis by the thioester method. *Biopolymers* **51**, 247-265.

Albericio, F. & Carpino, L. A. (1997). Coupling reagents and activation. *Methods Enzymol.* **289**, 104-126.

Albericio, F., Lloyd-Williams, P. & Giralt, E. (1997). Convergent solid-phase peptide synthesis. *Methods Enzymol.* **289**, 313-336.

Alsina, J. & Albericio, F. (2003). Solid-phase synthesis of C-terminal modified peptides. *Biopolymers* **71**, 454-477.

An, S. S. A., Lester, C. C., Peng, J.-L., Li, Y.-J., Rothwarf, D. M., Welker, E., Thannhauser, T. W., Zhang, L. S., Tam, J. P. & Scheraga, H. A. (1999). Retention of the cis proline conformation in tripeptide fragments of bovine pancreatic ribonuclease A containing a non-natural proline analogue, 5,5-dimethylproline. *J. Am. Chem. Soc.* **121**, 11558-11566.

Ariza, X., Urpi, F., Viladomat, C. & Vilarrasa, J. (1998). One-pot conversion of azides to Boc-protected amines with trimethylphosphine and Boc-ON. *Tetrahedron Lett.* **39**, 9101-9102.

Arner, E. S., Sarioglu, H., Lottspeich, F., Holmgren, A. & Bock, A. (1999). High-level expression in Escherichia coli of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the selA, selB and selC genes. *J. Mol. Biol.* **292**, 1003-1016.

Arnold, U., Hinderaker, M. P., Köditz, J., Golbik, R., Ulbrich-Hoffmann, R. & Raines, R.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

T. (2003). Protein prosthesis: A nonnatural residue accelerates folding and increases stability. *J. Am. Chem. Soc.* **125**, 7500-7501.

Arnold, U., Hinderaker, M. P., Nilsson, B. L., Huck, B. R., Gellman, S. H. & Raines, R. T. (2002). Protein prosthesis: A semisynthetic enzyme with a β-peptide reverse turn. *J. Am. Chem. Soc.* **124**, 368-369.

Ayers, B., Blaschke, U. K., Camarero, J. A., Cotton, G. J., Holford, M. & Muir, T. W. (2000). Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers* **51**, 343-354.

Backes, B. J. & Ellman, J. A. (1994). Carbon-carbon bond-forming methods on solid support. Utilization of Kenner's "safety-catch" linker. *J. Am. Chem. Soc.* **116**, 11171-11172.

Backes, B. J. & Ellman, J. A. (1999). An alkanesulfonamide "safety-catch" linker for solid-phase synthesis. *J. Org. Chem.* **64**, 2322-2330.

Backes, B. J., Virgilio, A. A. & Ellman, J. A. (1996). Activation method to prepare a highly reactive acylsulfonamide "safety-catch" linker for solid-phase synthesis. *J. Am. Chem. Soc.* **118**, 3055-3056.

Ball, H. L., King, D. S., Cohen, F. E., Prusiner, S. B. & Baldwin, M. A. (2001). Engineering the prion protein using chemical synthesis. *J. Peptide Res.* **58**, 357-374.

Becker, C. F. W., Hunter, C. L., Seidel, R., Kent, S. B. H., Goody, R. S. & Engelhard, M. (2003). Total chemical synthesis of a functional interacting protein pair: The protooncogene H-Ras and the Ras-binding domain of its effector c-Raf1. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5075-5080.

Becker, C. F. W., Hunter, C. L., Seidel, R. P., Kent, S. B. H., Goody, R. S. & Engelhard, M. (2001). A sensitive fluorescence monitor for the detection of activated Ras: Total chemical synthesis of site-specifically labeled Ras binding domain of c-Raf1 immobilized on a surface. *Chem. Biol.* **8**, 243-252.

Beligere, G. S. & Dawson, P. E. (1999). Conformationally assisted protein ligation using C-terminal thioester peptides. *J. Am. Chem. Soc.* **121**, 6332-6333.

Bergmann, M. & Zervas, L. (1932). Über ein allgemeines verfahren der peptid-synthese. *Chem. Ber.* **65**, 1192-1201.

Berry, S. M., Gieselman, M. D., Nilges, M. J., Van der Donk, W. A. & Lu, Y. (2002). An engineered azurin variant containing a selenocysteine copper ligand. *J. Am. Chem. Soc.* **124**, 2084-2085.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Besse, D., Siedler, F., Diercks, T., Kessler, H. & Moroder, L. (1997). The redox potential of selenocystine in unconstrained cyclic peptides. *Angew. Chem. Int. Ed. Engl.* **36**, 883-885.

Blaschke, U. K., Silberstein, J. & Muir, T. W. (2000). Protein engineering by expressed protein ligation. *Methods Enzymol.* **328**, 478-496.

Block, E., Oforiokai, G. & Zubieta, J. (1989). 2-Phosphinobenzenethiols and 2-phosphinylbenzenethiols - new ligand types. *J. Am. Chem. Soc.* **111**, 2327-2329.

Blume, F., Zemolka, S., Fey, T., Kranich, R. & Schmalz, H.-G. (2002). Identification of suitable ligands for a transition metal-catalyzed reaction: screening of a modular ligand library in the enantioselective hydroboration of styrene. *Adv. Synth. Catal.* **344**, 868-883.

Borgia, J. A. & Fields, G. B. (2000). Chemical synthesis of proteins. *Trends Biotechnol.* **18**, 243-251.

Bosch, I., Gonzalez, A., Urpi, F. & Vilarrasa, J. (1996). On the reaction of acyl chlorides and carboxylic anhydrides with phosphazenes. *J. Org. Chem.* **61**, 5638-5643.

Bosch, I., Romea, P., Urpi, F. & Vilarrasa, J. (1993). Alternative procedures for the macrolactamization of omega- azido acids. *Tetrahedron Lett.* **34**, 4671-4674.

Bosch, I., Urpi, F. & Vilarrasa, J. (1995). Epimerization-free peptide formation from carboxylic-acid anhydrides and azido derivatives. *J. Chem. Soc.-Chem. Commun.*, 91-92.

Brask, J., Albericio, F. & Jensen, K. J. (2003). Fmoc solid-phase synthesis of peptide thioesters by masking as trithioortho esters. *Org. Lett.* **5**, 2951-2953.

Bray, B. L. (2003). Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat. Rev. Drug. Discov.* **2**, 587-593.

Brik, A., Keinan, E. & Dawson, P. E. (2000). Protein synthesis by solid-phase chemical ligation using a safety catch linker. *J. Org. Chem.* **65**, 3829-3835.

Brisset, H., Gourdel, Y., Pellon, P. & Lecorre, M. (1993). Phosphine-borane complexes - direct use in asymmetric catalysis. *Tetrahedron Lett.* **34**, 4523-4526.

Bruice, T. C. & Pandit, U. K. (1960). The effect of geminal substitution, ring size, and rotamer distribution on the intramolecular nucleophilic catalysis of the hydrolysis of monophenyl esters of dibasic acids and the solvolysis of the intermediate anhydrides. *J. Am. Chem. Soc.* **82**, 5858-5865.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Brunel, J. M., Faure, B. & Maffei, M. (1998). Phosphine-boranes: synthesis, characterization and synthetic applications. *Coord. Chem. Rev.* **178-180**, 665-698.

Bu, X., Wu, X., Xie, G. & Guo, Z. (2002a). Synthesis of tyrocidine A and its analogues by spontaneous cyclization in aqueous solution. *Org. Lett.* **4**, 2893-2895.

Bu, X., Xie, G., Law, C. W. & Guo, Z. (2002b). An improved deblocking agent for direct Fmoc solid-phase synthesis of peptide thioester. *Tetrahedron Lett.* **43**, 2419-2422.

Camarero, J. A., Cotton, G. J., Adeva, A. & Muir, T. W. (1998). Chemical ligation of unprotected peptides directly from a solid support. *J. Peptide Res.* **51**, 303-316.

Cane, D. E., Walsh, C. T. & Khosla, C. (1998). Biochemistry - harnessing the biosynthetic code: Combinations, permutations, and mutations. *Science* **282**, 63-68.

Canne, L. E., Bark, S. J. & Kent, S. B. H. (1996). Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* **118**, 5891-5896.

Canne, L. E., Botti, P., Simon, R. J., Chen, Y., Dennis, E. A. & Kent, S. B. H. (1999). Chemical protein synthesis by solid phase ligation of unprotected peptide segments. *J. Am. Chem. Soc.* **121**, 8720-8727.

Canne, L. E., Walker, S. M. & Kent, S. B. H. (1995). A general method for the synthesis of thioester resin linkers for use in the solid phase synthesis of peptide-alpha-acids. *Tetrahedron Lett.* **36**, 1217-1220.

Carboni, B. & Monnier, L. (1999). Recent developments in the chemistry of amine- and phosphine- boranes. *Tetrahedron* **55**, 1197-1248.

Cardona, V. M. F., Hartley, O. & Botti, P. (2003). Synthesis of cyclic peptides from unprotected precursors using removable N-alpha-(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary. *J. Peptide Res.* **61**, 152-157.

Casi, G. & Hilvert, D. (2003). Convergent protein synthesis. *Curr. Opin. Struct. Biol.* **13**, 589-594.

Chang, G., Guida, W. C. & Still, W. C. (1989). An internal coordinate method for searching conformational space. *J. Am. Chem. Soc.* **111**, 4379-4386.

Charrier, C. & Mathey, F. (1978). Diphenyl(cyclopentadienylmethyl)phosphine and its ferrocene and cymantrene complexes. *Tetrahedron Lett.* **27**, 2407-2410.

Coltart, D. M. (2000). Peptide segment coupling by prior ligation and proximity-induced

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

intramolecular acyl transfer. *Tetrahedron* **56**, 3449-3491.

Cotton, G. J. & Muir, T. W. (1999). Peptide ligation and its application to protein engineering. *Chem. Biol.* **6**, R247-R256.

Dawson, P. E., Churchill, M. J., Ghadiri, M. R. & Kent, S. B. H. (1997). Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* **119**, 4325-4329.

Dawson, P. E. & Kent, S. B. H. (2000). Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923-960.

Dawson, P. E., Muir, T. W., Clarklewis, I. & Kent, S. B. H. (1994). Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779.

deL. Milton, R. C., Milton, S. C. F. & Kent, S. B. H. (1992). Total chemical synthesis of a D-enzyme: The enantiomers of HIV-1 protease show demonstration of reciprocal chiral substrate specificity. *Science* **256**, 1445-1448.

Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S., Dawson, P. E., Schultz, P. G. & Weiss, S. (2000). Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5179-5184.

Dorman, D. E. & Bovey, F. A. (1973). Carbon-13 magnetic resonance spectroscopy. Spectrum of proline in oligopeptides. *J. Org. Chem.* **38**, 2379-2383.

Du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W., Katsoyannis, P. G. & Gordon, S. (1953). The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* **75**, 4879-4880.

Eom, K. D., Miao, Z., Yang, Z.-l. & Tam, J. P. (2003). Tandem ligation of multipartite peptides with cell-permeable activity. *J. Am. Chem. Soc.* **125**, 73-82.

Evans, T. C., Benner, J. & Xu, M. Q. (1998). Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Science* **7**, 2256-2264.

Farrington, G. K., Kumar, A. & Wedler, F. C. (1989). A convenient synthesis of diethyl (mercaptomethyl)phosphonate. *Org. Prep. Proced. Int.* **21**, 390-392.

Fischer, E. (1902). Syntheses in the purine and sugar group. http://www.nobel.se/chemistry/laureates/1902/fischer-lecture.html.

Fischer, E. & Fourneau, E. (1901). Ueber einige derivate des glykocolls. *Chem. Ber.* **34**,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2868-2877.

Fotouhi, N., Galakatos, N. G. & Kemp, D. S. (1989). Peptide synthesis by prior thiol capture. 6. Rates of the disulfide bond forming capture reaction and demonstration of the overall strategy by synthesis of the C-terminal 29-peptide sequence of BPTI. *J. Org. Chem.* **54**, 2803-2817.

Friedman, M. (1999). Lysinoalanine in food and in antimicrobial proteins. *Adv. Exp. Med. Biol.* **459**, 145-159.

Garcia, J., Urpi, F. & Vilarrasa, J. (1984). New synthetic tricks - triphenylphosphine-mediated amide formation from carboxylic-acids and azides. *Tetrahedron Lett.* **25**, 4841-4844.

Garcia, J., Vilarrasa, J., Bordas, X. & Banaszek, A. (1986). New synthetic tricks - one-pot preparation of N-substituted phthalimides from azides and phthalic-anhydride. *Tetrahedron Lett.* **27**, 639-640.

Gieselman, M. D., Xie, L. & van der Donk, W. A. (2001). Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.* **3**, 1331-1334.

Gieselman, M. D., Zhu, Y., Zhou, H., Galonic, D. & van der Donk, W. A. (2002). Selenocysteine derivatives for chemoselective ligations. *ChemBioChem* **3**, 709-716.

Giriat, I. & Muir, T. W. (2003). Protein semi-synthesis in living cells. *J. Am. Chem. Soc.* **125**, 7180-7181.

Gololobov, Y. G. & Kasukhin, L. F. (1992). Recent advances in the Staudinger reaction. *Tetrahedron* **48**, 1353-1406.

Gololobov, Y. G., Zhmurova, I. N. & Kasukhin, L. F. (1981). Sixty years of Staudinger reaction. *Tetrahedron* **37**, 437-472.

Goody, R. S., Alexandrov, K. & Engelhard, M. (2002). Combining chemical and biological techniques to produce modified proteins. *ChemBioChem* **3**, 399-403.

Guida, W. C., Bohacek, R. S. & Erion, M. D. (1992). Probing the conformational space available to inhibitors in the thermolysin active site using Monte Carlo energy minimization techniques. *J. Comput. Chem.* **13**, 214-228.

Gutte, B. & Merrifield, R. B. (1969). The total synthesis of an enzyme with ribonuclease A activity. *J. Am. Chem. Soc.* **91**, 501-502.

Gutte, B. & Merrifield, R. B. (1971). The synthesis of ribonuclease A. *J. Biol. Chem.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

246, 1922-1941.

Guy, C. A. & Fields, G. B. (1997). Trifluoroacetic acid cleavage and deprotection of resin-bound peptides following synthesis by Fmoc chemistry. *Methods Enzymol.* **289**, 67-83.

Hackeng, T. M., Fernandez, J. A., Dawson, P. E., Kent, S. B. & Griffin, J. H. (2000). Chemical synthesis and spontaneous folding of a multidomain protein: Anticoagulant microprotein S. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14074-14078.

Hackeng, T. M., Griffin, J. H. & Dawson, P. E. (1999). Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10068-10073.

Hackeng, T. M., Rosing, J., Spronk, H. M. H. & Vermeer, C. (2001). Total chemical synthesis of human matrix Gla protein. *Protein Sci.* **10**, 864-870.

Han, J. C. & Han, G. Y. (1994). A procedure for quantitative determination of tris(2-carboxyethyl)phosphine, an odorless reducing agent more stable and effective than dithiothreitol. *Anal. Biochem.* **220**, 5-10.

He, Y., Hinklin, R. J. & Kiessling, L. L. unpublished results.

Herd, O., Hessler, A., Hingst, M., Machnitzki, P., Tepper, M. & Stelzer, O. (1998). Palladium catalyzed P-C coupling - a powerful tool for the syntheses of hydrophilic phosphines. *Catal. Today* **42**, 413-420.

Hirschmann, R., Nutt, R. F., Veber, D. F., Vitali, R. A., Varga, S. L., Jacob, T. A., Holly, F. W. & Denkewalter, R. G. (1969). Studies on the total synthesis of an enzyme. V. The preparation of enzymatically active material. *J. Am. Chem. Soc.* **91**, 507-508.

Hofmann, R. M. & Muir, T. W. (2002). Recent advances in the application of expressed protein ligation to protein engineering. *Curr. Opin. Biotechnol.* **13**, 1-7.

Hojo, H., Haginoya, E., Matsumoto, Y., Nakahara, Y., Nabeshima, K., Toole, B. P. & Watanabe, Y. (2003). The first synthesis of peptide thioester carrying N-linked core pentasaccharide through modified Fmoc thioester preparation: synthesis of an N-glycolsylated Ig domain of emmprin. *Tetrahedron Lett.* **44**, 2961-2964.

Holford, M. & Muir, T. W. (1998). Adding 'splice' to protein engineering. *Structure* **6**, 951-956.

Homandberg, G. A. & Laskowski, M. (1979). Enzymatic resynthesis of the hydrolyzed peptide bond(s) in ribonuclease S. *Biochemistry* **18**, 586-592.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Hondal, R. J., Nilsson, B. L. & Raines, R. T. (2001). Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **123**, 5140-5141.

Hondal, R. J. & Raines, R. T. (2002). Semisynthesis of proteins containing selenocysteine. *Methods Enzymol.* **347**, 70-83.

Imamoto, T., Oshiki, T., Onozawa, T., Kusumoto, T. & Sato, K. (1990). Synthesis and reactions of phosphine boranes - synthesis of new bidentate ligands with homochiral phosphine centers via optically pure phosphine boranes. *J. Am. Chem. Soc.* **112**, 5244-5252.

Inazu, T. & Kobayashi, K. (1993). A new simple method for the synthesis of N(alpha)-Fmoc-N(beta)- glycosylated-L-asparagine derivatives. *Synlett*, 869-870.

Ingenito, R., Bianchi, E., Fattori, D. & Pessi, A. (1999). Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry. *J. Am. Chem. Soc.* **121**, 11369-11374.

Ishiwata, A., Ichiyanagi, T., Takatani, M. & Ito, Y. (2003). Chemoselective peptide bond formation using formyl-substituted nitrophenylthio ester. *Tetrahedron Lett.* **44**, 3187-3190.

Janssen, M. J. (1969). In *The Chemistry of Carboxylic Acids and Esters* (Patai, S., ed.), pp. 730-736. Interscience Publishers, New York.

Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225-11236.

Jung, M. E. & Gervay, J. (1991). Gem-dialkyl effect in the intramolecular Diels-Alder reaction of 2-furfuryl methyl fumarates - the reactive rotamer effect, enthalpic basis for acceleration, and evidence for a polar transition-state. *J. Am. Chem. Soc.* **113**, 224-232.

Kaiser, E. T. (1989). Synthetic approaches to biologically-active peptides and proteins including enzymes. *Acc. Chem. Res.* **22**, 47-54.

Katz, L. (1997). Manipulation of modular polyketide syntheses. *Chem. Rev.* **97**, 2557-2575.

Kawakami, T., Akaji, K. & Aimoto, S. (2001). Peptide bond formation mediated by 4,5-dimethoxy-2-mercaptobenzylamine after periodate oxidation of the N-terminal serine residue. *Org. Lett.* **3**, 1403-1405.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Keating, T. A. & Walsh, C. T. (1999). Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **3**, 598-606.

Kelemen, B. R., Klink, T. A., Behlke, M. A., Eubanks, S. R., Leland, P. A. & Raines, R. T. (1999). Hypersensitive substrate for ribonucleases. *Nucleic Acids Res.* **27**, 3696-3701.

Kemp, D. S. & Carey, R. I. (1989). Boc-L-Dmt-OH as a fully N,S-blocked cysteine derivative for peptide synthesis by prior thiol capture. Facile conversion of N-terminal Boc-L-Dmt-peptides to H-Cys(Scm)-peptides. *J. Org. Chem.* **54**, 3640-3646.

Kemp, D. S. & Carey, R. I. (1993). Synthesis of a 39-peptide and a 25-peptide by thiol capture ligations: Observation of a 40-fold rate acceleration of the intramolecular O,N-acyl-transfer reaction between peptide fragments bearing only cysteine protective groups. *J. Org. Chem.* **58**, 2216-2222.

Kemp, D. S. & Fotouhi, N. (1987). Peptide synthesis by prior thiol capture-V. The scope and control of disulfide interchange during the acyl transfer step. *Tetrahedron Lett.* **28**, 4637-4640.

Kemp, D. S. & Galakatos, N. G. (1986). Peptide-synthesis by prior thiol capture. 1. A convenient synthesis of 4-hydroxy-6-mercaptodibenzofuran and novel solid-phase synthesis of peptide-derived "4-(acyloxy)-6- mercaptodibenzofurans. *J. Org. Chem.* **51**, 1821-1829.

Kemp, D. S., Galakatos, N. G., Bowen, B. & Tan, K. (1986a). Peptide synthesis by prior thiol capture. 2. Design of templates for intramolecular O,N-acyl transfer. 4,6-Disubstituted dibenzofurans as optimal spacing elements. *J. Org. Chem.* **51**, 1829-1839.

Kemp, D. S., Galakatos, N. G., Dranginis, S., Ashton, C., Fotouhi, N. & Curran, T. P. (1986b). Peptide synthesis by prior thiol capture. 4. Amide bond formation: The effect of a side-chain substituent on the rates of intramolecular O,N-acyl transfer. *J. Org. Chem.* **51**, 3320-3324.

Kenner, G. W., McDermott, J. R. & Sheppard, R. C. (1971). Safety catch principle in solid phase peptide synthesis. *J. Chem. Soc., Chem. Commun.*, 636-637.

Kent, S. (2003). Total chemical synthesis of enzymes. *J. Peptide Sci.* **9**, 574-593.

Kent, S. B. H. (1988). Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **57**, 957-989.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Kinart, C. M. & Kinart, W. J. (1994). Studies on intermolecular interactions in liquid tetrahydrofuran-water binary mixtures. *Pol. J. Chem.* **68**, 339-342.

Knowles, J. R. (1987). Tinkering with enzymes: what are we learning? *Science* **236**, 1252-1258.

Kochendoerfer, G. G., Chen, S.-Y., Mao, F., Cressman, S., Traviglia, S., Shao, H., Hunter, C. L., Low, D. W., Cagle, E. N., Carnevali, M., Gueriguian, V., Keogh, P. J., Porter, H., Stratton, S. M., Wiedeke, M. C., Wilken, J., Tang, J., Levy, J. J., Miranda, L. P., Crnogorac, M. M., Kalbag, S., Botti, P., Schindler-Horvat, J., Savatski, L., Adamson, J. W., Kung, A., Kent, S. B. H. & Bradburne, J. A. (2003). Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* **299**, 884-887.

Kochendoerfer, G. G. & Kent, S. B. H. (1999). Chemical protein synthesis. *Curr. Opin. Chem. Biol.* **3**, 665-671.

Köhn, M., Wacker, R., Peters, C., Schröder, H., Soulère, L., Breinbauer, R., Niemeyer, C. M. & Waldmann, H. (2003). Staudinger ligation: A new immobilization strategy for the preparation of small-molecule arrays. *Angew. Chem. Int. Ed. Engl.*, In press.

Lamoureux, G. V. & Whitesides, G. M. (1993). Synthesis of dithiols as reducing agents for disulfides in neutral aqueous-solution and comparison of reduction potentials. *J. Org. Chem.* **58**, 633-641.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Harkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.-F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Li, X., Kawakami, T. & Aimoto, S. (1998). Direct preparation of peptide thioesters using an Fmoc solid-phase method. *Tetrahedron Lett.* **39**, 8669-8672.

Lloyd-Williams, P., Albericio, F. & Giralt, E. (1992). Convergent solid-phase peptide synthesis. *Tetrahedron* **49**, 11065-11133.

Low, D. W., Hill, M. G., Carrasco, M. R., Kent, S. B. H. & Botti, P. (2001). Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6554-6559.

Lundquist, J. T. & Pelletier, J. C. (2001). Improved solid-phase peptide synthesis method utilizing alpha- azide-protected amino acids. *Org. Lett.* **3**, 781-783.

MacMillan, D. & Bertozzi, C. R. (2000). New directions in glycoprotein engineering. *Tetrahedron* **56**, 9515-9525.

Malkinson, J. P., Falconer, R. A. & Toth, I. (2000). Synthesis of C-terminal glycopeptides from resin-bound glycosyl azides via a modified Staudinger reaction. *J. Org. Chem.* **65**, 5249-5252.

Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.

Marcaurelle, L. A., Mizoue, L. S., Wilken, J., Oldham, L., Kent, S. B. H., Handel, T. M. & Bertozzi, C. R. (2001). Chemical synthesis of lymphotactin: A glycosylated chemokine with a C-terminal mucin-like domain. *Chem. Eur. J.* **7**, 1129-1132.

Marinzi, C., Bark, S. J., Offer, J. & Dawson, P. E. (2001). A new scaffold for amide ligation. *Bioorg. Med. Chem.* **9**, 2323-2328.

Maunier, V., Boullanger, P. & Lafont, D. (1997). One-pot synthesis of glycosyl amides from glycosyl azides using a modified Staudinger reaction. *J. Carbohydr. Chem.* **16**, 231-235.

McBride, B. J. & Kemp, D. S. (1987). Peptide synthesis by prior thiol capture. . .III. Assessment of levels of racemization during two typical thiol capture coupling reactions. *Tetrahedron Lett.* **28**, 3435-3438.

McCaldon, P. & Argos, P. (1988a). Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **4**, 99-122.

McCaldon, P. & Argos, P. (1988b). Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide-sequences. *Proteins* **4**, 99-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

122.

Meldal, M. (1997). Properties of solid supports. *Methods Enzymol.* **289**, 83-104.

Meldal, M., Juliano, M. A. & Jansson, A. M. (1997). Azido acids in a novel method of solid-phase peptide synthesis. *Tetrahedron Lett.* **38**, 2531-2534.

Menger, F. M. (1985). On the source of intramolecular and enzymatic reactivity. *Acc. Chem. Res.* **18**, 128-134.

Merkx, R., Rijkers, D. T. S., Demink, J. & Liskamp, R. M. J. (2003). Chemoselective coupling of peptide fragments using the Staudinger ligation. *Tetrahedron Lett.* **44**, 4515-4518.

Merrifield, B. (1986). Solid phase synthesis. *Science* **232**, 341-347.

Merrifield, B. (1996). The chemical synthesis of proteins. *Protein Sci.* **5**, 1947-1951.

Merrifield, R. B. (1963). Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154.

Merrifield, R. B., Stewart, J. M. & Jernberg, N. (1966). Instrument for automated synthesis of peptides. *Anal. Chem.* **38**, 1905-1914.

Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H. & Wlodawer, A. (1989). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**, 1149-1152.

Miranda, L. P. & Alewood, P. F. (1999). Accelerated chemical synthesis of peptides and small proteins. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1181-1186.

Miranda, L. P. & Alewood, P. F. (2000). Challenges for protein chemical synthesis in the 21st century: Bridging genomics and proteomics. *Biopolymers* **55**, 217-226.

Mizuno, M., Haneda, K., Iguchi, R., Muramoto, I., Kawakami, T., Aimoto, S., Yamamoto, K. & Inazu, T. (1999a). Synthesis of a glycopeptide containing oligosaccharides: Chemoenzymatic synthesis of eel calcitonin analogues having natural N-linked oligosaccharides. *J. Am. Chem. Soc.* **121**, 284-290.

Mizuno, M., Muramoto, I., Kobayashi, K., Yaginuma, H. & Inazu, T. (1999b). A simple method for the synthesis of N-beta-glycosylated-asparagine and -glutamine derivatives. *Synthesis*, 162-165.

Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C.,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chang, G., Hendrickson, T. & Still, W. C. (1990). MacroModel - an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **11**, 440-467.

Molina, P. & Vilaplana, M. J. (1994). Iminophosphoranes - useful building-blocks for the preparation of nitrogen-containing heterocycles. *Synthesis*, 1197-1218.

Mootz, H. D., Blum, E. S., Tyszkiewicz, A. B. & Muir, T. W. (2003). Conditional protein splicing: A new tool to control protein structure and function in vitro and in vivo. *J. Am. Chem. Soc.* **125**, 10561-10569.

Mootz, H. D. & Muir, T. W. (2002). Protein splicing triggered by a small molecule. *J. Am. Chem. Soc.* **124**, 9044-9045.

Muir, T. W. (2003). Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249-289.

Muir, T. W., Dawson, P. E. & Kent, S. B. H. (1997). Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. In *Solid-Phase Peptide Synthesis*, Vol. 289, pp. 266-298.

Muir, T. W., Sondhi, D. & Cole, P. A. (1998). Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6705-6710.

Nilsson, B. L., Hondal, R. J., Soellner, M. B. & Raines, R. T. (2003). Protein assembly by orthogonal chemical ligation methods. *J. Am. Chem. Soc.* **125**, 5268-5269.

Nilsson, B. L., Kiessling, L. L. & Raines, R. T. (2000). Staudinger ligation: A peptide from a thioester and azide. *Org. Lett.* **2**, 1939-1941.

Nilsson, B. L., Kiessling, L. L. & Raines, R. T. (2001). High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide. *Org. Lett.* **3**, 9-12.

Nilsson, B. L., Soellner, M. B. & Raines, R. T. unpublished results.

Nishiuchi, Y., Inui, T., Nishio, H., Bódi, J., Kimura, T., Tsuji, F. I. & Sakakibara, S. (1998). Chemical synthesis of the precursor molecule of the Aequorea green fluorescent protein, subsequent folding, and development of fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13549-13554.

Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920-930.

Odaert, B., Jean, F., Boutillon, C., Buisine, E., Melnyk, O., Tartar, A. & Lippens, G. (1999). Synthesis, folding, and structure of the β-turn mimic modified B1 domain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of streptococcal protein G. *Protein Sci.* **8**, 2773-2783.

Offer, J., Boddy, C. N. C. & Dawson, P. E. (2002). Extending synthetic access to proteins with a removable acyl transfer auxiliary. *J. Am. Chem. Soc.* **124**, 4642-4646.

Offer, J. & Dawson, P. E. (2000). N-Alpha-2-mercaptobenzylamine-assisted chemical ligation. *Org. Lett.* **2**, 23-26.

Pachamuthu, K. & Schmidt, R. R. (2003). Synthesis of methionine containing peptides related to native chemical ligation. *Synlett*, 659-662.

Page, M. I. & Jencks, W. P. (1971). Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1678-1683.

Quaderer, R. & Hilvert, D. (2001). Improved synthesis of C-terminal peptide thioesters on "safety-catch" resins using LiBr/THF. *Org. Lett.* **3**, 3181-3184.

Quaderer, R., Sewing, A. & Hivert, D. (2001). Selenocysteine-mediated native chemical ligation. *Helv. Chim. Acta* **84**, 1197-1206.

Raines, R. T. (1997). Nature's transitory covalent bond. *Nat. Struct. Biol.* **4**, 424-427.

Raines, R. T. (1998). Ribonuclease A. *Chem. Rev.* **98**, 1045-1066.

Rampal, J. B., Berlin, K. D., Pantaleo, N. S., McGuffy, A. & Helm, D. V. d. (1981). Carbon-phosphorus heterocycles. Synthesis of phosphorus-containing cannabinoid precursors and a single-crystal analysis of 1,2,3,4-tetrahydro-10-hydroxy-8-n-pentyl-5H-phosphorino[3,4-c][1]benzopyran-5-one 3-oxide. *J. Am. Chem. Soc.* **103**, 2032 - 2036.

Rico, M., Bruix, M., Santoro, J., Gonzalez, C., Neira, J. L., Nieto, J. L. & Herranz, J. (1989). Sequential H-1-NMR assignment and solution structure of bovine pancreatic ribonuclease A. *Eur. J. Biochem.* **183**, 623-638.

Roelfes, G. & Hilvert, D. (2003). Incorporation of selenomethionine into proteins through selenohomocysteine-mediated ligation. *Angew. Chem. Int. Ed. Engl.* **42**, 2275-2277.

Sato, T. & Aimoto, S. (2003). Use of thiosulfonate for the protection of thiol groups in peptide ligation by the thioester method. *Tetrahedron Lett.* **44**, 8085-8087.

Saxon, E., Armstrong, J. I. & Bertozzi, C. R. (2000). A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds. *Org. Lett.* **2**, 2141-2143.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Saxon, E. & Bertozzi, C. R. (2000a). Cell surface engineering by a modified Staudinger reaction. *Science* **287**, 2007-2010.

Saxon, E. & Bertozzi, C. R. (2000b). Designing chemical reactions for a cellular environment: Cell surface engineering via a modified Staudinger reaction. *Science* **287**, 2007-2010.

Saxon, E., Luchansky, S. J., Hang, H. C., Yu, C., Lee, S. C. & Bertozzi, C. R. (2002). Investigating cellular metabolism of synthetic azidosugars with the Staudinger ligation. *J. Am. Chem. Soc.* **124**, 14893-14902.

Severinov, K. & Muir, T. W. (1998). Expressed protein ligation, a novel method for studying protein–protein interactions in transcription. *J. Biol. Chem.* **273**, 16205-16209.

Shalev, D. E., Chiacchiera, S. M., Radkowsky, A. E. & Kosower, E. M. (1996). Sequence of reactant combination alters the course of the Staudinger reaction of azides with acyl derivatives. Bimanes .30. *J. Org. Chem.* **61**, 1689-1701.

Shao, Y., Lu, W. & Kent, S. B. H. (1998). A novel method to synthesize cyclic peptides. *Tetrahedron Lett.* **39**, 3911-3914.

Shin, Y., Winans, K. A., Backes, B. J., Kent, S. B. H., Ellman, J. A. & Bertozzi, C. R. (1999). Fmoc-based synthesis of peptide-(alpha)thioesters: Application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.* **121**, 11684-11689.

Soellner, M. B., Dickson, K. A., Nilsson, B. L. & Raines, R. T. (2003). Site-specific protein immobilization by Staudinger ligation. *J. Am. Chem. Soc.* **125**, 11790-11791.

Soellner, M. B., Nilsson, B. L. & Raines, R. T. (2002). Staudinger ligation of α-azido acids retains stereochemistry. *J. Org. Chem.* **67**, 4993-4996.

Songster, M. F. & Barany, G. (1997). Handles for solid-phase peptide synthesis. *Methods Enzymol.* **289**, 126-174.

Staudinger, H. & Meyer, J. (1919). New organic compounds of phosphorus. III. Phosphinemethylene derivatives and phosphinimines. *Helvetica Chim. Acta* **2**, 635-46.

Stewart, J. M. (1997). Cleavage methods following Boc-based solid-phase peptide synthesis. *Methods Enzymol.* **289**, 29-44.

Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H. &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Wlodawer, A. (1990). X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8805-8809.

Swinnen, D. & Hilvert, D. (2000). Facile, Fmoc-compatible solid-phase synthesis of peptide C- terminal thioesters. *Org. Lett.* **2**, 2439-2442.

Tam, J. P., Lu, Y. A., Liu, C. F. & Shao, J. (1995). Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12485-12489.

Tam, J. P., Xu, J. X. & Eom, K. D. (2001). Methods and strategies of peptide ligation. *Biopolymers* **60**, 194-205.

Tam, J. P. & Yu, Q. (1998). Methionine ligation strategy in the biomimetic synthesis of parathyroid hormone. *Biopolymers* **46**, 319-327.

Tam, J. P., Yu, Q. T. & Miao, Z. W. (1999). Orthogonal ligation strategies for peptide and protein. *Biopolymers* **51**, 311-332.

Tang, Z. L. & Pelletier, J. C. (1998). Preparation of amides from acids and resin bound azides: Suppression of intramolecular lactam formation. *Tetrahedron Lett.* **39**, 4773-4776.

Urpi, F. & Vilarrasa, J. (1986). New synthetic tricks advantages of using triethylphosphine in some phosphorus-based reactions. *Tetrahedron Lett.* **27**, 4623-4624.

Vedejs, E. & Diver, S. T. (1993). Tributylphosphine - a remarkable acylation catalyst. *J. Am. Chem. Soc.* **115**, 3358-3359.

Velasco, M. D., Molina, P., Fresneda, P. M. & Sanz, M. A. (2000). Isolation, reactivity and intramolecular trapping of phosphazide intermediates in the Staudinger reaction of tertiary phosphines with azides. *Tetrahedron* **56**, 4079-4084.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Miklos, G. L. G., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reiner, K., Remington, K., Abu-Threide, J., Beasle, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturved, K., Deng, Z., Di Francesco,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gon, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R.-R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z. Y., Wang, A., Wang, X., Wang, J., Wei, M.-H., Wides, R., Xiao, C., Yan, C., et al. (2001). The sequence of the human genome. *Science* **291**, 1304-1351.

Vocadlo, D. J., Hang, H. C., Kim, E.-J., Hanover, J. A. & Bertozzi, C. R. (2003). A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9116-9121.

Vogel, K. & Chmielewski, J. (1994). Rapid and efficient resynthesis of proteolyzed triose phosphate isomerase. *J. Am. Chem. Soc.* **116**, 11163-11164.

von Egglekraut-Gottanka, R., Klose, A., Beck-Sickinger, A. G. & Beyermann, M. (2003). Peptide alpha-thioester formation using standard Fmoc-chemistry. *Tetrahedron Lett.* **44**, 3551-3554.

vonDohren, H., Keller, U., Vater, J. & Zocher, R. (1997). Multifunctional peptide synthetases. *Chem. Rev.* **97**, 2675-2705.

Wallace, C. J. A. (1993). The curious case of protein splicing: Mechanistic insights suggested by protein semisynthesis. *Protein Sci.* **2**, 697-705.

Wellings, D. A. & Atherton, E. (1997). Standard Fmoc protocols. *Methods Enzymol.* **289**, 44-67.

Wieland, T., Bokelmann, E., Bauer, L., Lang, H. U., Lau, H. & Schafer, W. (1953). Polypeptide syntheses. VIII. Formation of sulfur containing peptides by the intramolecular migration of aminoacyl groups. *Liebigs Ann. Chem.* **583**, 129-49.

Wilken, J. & Kent, S. B. H. (1998). Chemical protein synthesis. *Curr. Opin. Biotechnol.* **9**, 412-426.

Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J. & Kent, S. B. H. (1989). Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease. *Science* **245**, 616-621.

Woycechowsky, K. J., Hook, B. A. & Raines, R. T. (2003). Catalysis of protein folding by an immobilized small-molecule dithiol. *Biotechnology Progress* **19**, 1307-1314.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Woycechowsky, K. J. & Raines, R. T. (2003). The CXC motif: A functional mimic of protein disulfide isomerase. *Biochemistry* **42**, 5387-5394.

Woycechowsky, K. J., Wittrup, K. D. & Raines, R. T. (1999). A small-molecule catalyst of protein folding in vitro and in vivo. *Chem. Biol.* **6**, 871-879.

Wu, X., Bu, X., Wong, K. M., Yan, W. & Guo, Z. (2003a). Biomimetic synthesis of gramicidin S and analogues by enzymatic cyclization of linear precursors on solid support. *Org. Lett.* **5**, 1749-1752.

Wu, Z., Prahl, A., Powell, R., Ericksen, B., Lubkowski, J. & Lu, W. (2003b). From pro defensins to defensins: Synthesis and characterization of human neutrophil pro alpha-defensin-1 and its mature domain. *J. Peptide Res.* **62**, 53-62.

Wuthrich, K. (1976). *NMR in Biological Research: Peptides and Proteins*, American Elsevier Publishing Company, New York.

Wuttke, D. S., Gray, H. B., Fisher, S. L. & Imperiali, B. (1993). Semisynthesis of bipyridyl-alanine cytochrome c mutants: novel proteins with enhanced electron-transfer properties. *J. Am. Chem. Soc.* **115**, 8455-8456.

Yan, L. Z. & Dawson, P. E. (2001). Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **123**, 526-533.

Yin, Y. W. & Steitz, T. A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science* **298**, 1387-1395.

Zaloom, J. & Roberts, D. C. (1981). Preparation of azido derivatives from amino acids and peptides by diazo transfer. *J. Org. Chem.* **46**, 5173-6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.