Supporting Information

Distributed Drug Discovery, Part 3: Using D³ Methodology to Synthesize Analogs of an Anti-Melanoma Compound

William L. Scott,^{*,†} Christopher O. Audu,[†] Jeffery L. Dage,[#] Lawrence A. Goodwin,[#] Jacek G. Martynow,[†] Laura K. Platt,[†] Judith G. Smith,[†] Andrew T. Strong,[†] Kirk Wickizer,[†] Eric M. Woerly,[†] and Martin J. O'Donnell^{*,†}

[†]Department of Chemistry and Chemical Biology, Indiana University Purdue University Indianapolis, Indianapolis, Indiana 46202-3274. [#]Medicinal Analytical Chemistry, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana, 46285.Analytical Chemistry, Lilly Research Laboratories, Indianapolis, IN 46285.

Table of Contents

1. Tutorial for using the Distributed Drug Discovery (D^3) database	
2. Proton NMR Spectra of Products	
Compound 8{9}a	17
Compound 8{10}a	18
Compound 8{11}a	19
Compound 8{12}a	20
Compound 8{13}a	21
Compound 8{14}a	22
Compound 8{15}a	23
Compound 8{16}a	24
Compound 8{17}a	25
Compound 8{18}a	26
Compound 8{19}a	27
Compound 8{20}a	28
Compound 8{21}a	29
Compound 8{22}a	30
Compound 8{23}a	31
Compound 8{24}a	32
Compound 8{25}a	33
Compound 8{26}a	34
Compound 8{27}a	35
Compound 8{28}a	36
Compound 8{9}b	37
Compound 8{10}b	38
Compound 8{11}b	39
Compound 8{12}b	40
Compound 8{13}b	41
Compound 8{14}b	42
Compound 8{15}b	43
Compound 8{16}b	44
Compound 8{17}b	45
Compound 8{18}b	46
Compound 8{19}b	47
Compound 8{20}b	48
Compound 8{21}b	49
Compound 8{22}b	50
Compound 8{23}b	51
Compound 8{24}b	52
Compound 8{25}b	53
Compound 8{26}b	54
Compound 8{27}b	55
Compound 8{28}b	56

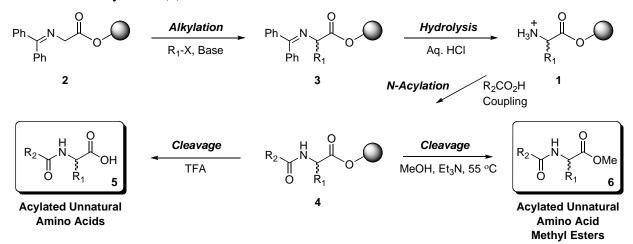
Tutorial for Using the Distributed Drug Discovery (D³) Database on the Collaborative Drug Discovery (CDD) Website

The Distributed Drug Discovery (D^3) database is a virtual catalog of 48,608 unique acylated unnatural amino acid derivatives obtained from a combinatorial enumeration. For access to this database, register for a free read-download account with Collaborative Drug Discovery (CDD) at http://www.collaborativedrug.com/register/iupui-d3 (accessed November 2, 2008).

The following discussion is d	ivided into three parts:	
1) Synthesis	2) Enumeration	3) Tutorial

Compound identifiers are those described in the accompanying *Journal of Combinatorial Chemistry* papers [Scott, W. L.; O'Donnell, M. J. et al. *J. Comb. Chem.* DOI: 10.1021/cc800184v (http://dx.doi.org/10.1021/cc800184v) and DOI: 10.1021/cc800185z (http://dx.doi.org/10.1021/cc800185z) and DOI: 10.1021/cc800183m (http://dx.doi.org/10.1021/cc800183m)]. The user is referred to these papers for an in-depth discussion of the D³ project.

Synthesis. Scheme 1 describes the chemistry involved in the combinatorial enumeration to prepare a virtual catalog of 48,608 unique acylated unnatural amino acids **5** or their methyl esters **6**. The starting material is the Wang resin-bound benzophenone imine of glycine (**2**). The first diversity element (R_1) is attached to the glycine scaffold by reaction of an alkyl halide or Michael acceptor (both abbreviated as " R_1 -X") in the presence of base to give **3** in an alkylation or conjugate addition reaction. The imine activating group is removed by hydrolysis and then the second diversity element is added by N-acylation using a standard coupling reaction. Cleavage from the resin is accomplished with either trifluoroacetic acid to give the acylated unnatural amino acid (**5**) or by transesterification with methanol to yield the acylated unnatural amino acid methyl ester (**6**).



Scheme 1. Synthetic Routes to Acylated Unnatural Amino Acids 5 and Methyl Esters 6.

A total of 100 commercially available electrophiles ($R_1X = alkyl$ halide or Michael acceptor) and 100 commercially available carboxylic acids (R_2CO_2H) were used in an enumeration creating all the possible combinations of R_1 and R_2 for **5** and **6**.

Enumeration. Enumeration with the 100 electrophiles (alkyl halides and Michael acceptors) and 100 carboxylic acids gave 24,416 acylated amino acids **5** and 24,192 acylated amino acid methyl esters **6**. The reason these numbers exceed 10,000 in each case is discussed below.

Issues of stereochemistry were addressed in this enumeration. Examples illustrating the possibilities for the amino acid products **5** are shown (Table 1). The products are identified in three ways: row/column (e.g. **A1**), number in D^3 database (e.g. DDD-000002150), or number from *J. Comb. Chem.* paper [e.g. (**R**)-5{9,119}]. The diversity elements, reactants R₁X and R₂CO₂H, include achiral (**1** and **A**), optically pure (**B**), racemic (**2** and **C**) and prochiral (**3**) examples. Only one of the possible stereoisomers for each product is shown in this table, designed to illustrate three main points.

(1) When achiral or optically pure reactants (R_1X and R_2CO_2H) were used, racemic products (2 stereoisomers) resulted, e.g. A1 and its enantiomer or B1 and its diastereomer at the α carbon. The database contains 15,308 unique molecules 5 (7,654 pairs) of this class.

(2) When one of the reactants (R_1X and R_2CO_2H) was racemic or, in the case of Michael acceptors, prochiral, 4 stereoisomers were obtained, e.g. A2 or A3 and their stereoisomers. The database contains 8,068 unique molecules 5 (2,017 sets of four stereoisomers) of this class.

(3) When both reactants (R_1X and R_2CO_2H) were racemic or prochiral and racemic, 8 stereoisomers were formed. Examples: C2 or C3 and their stereoisomers. The database contains 1,040 such unique molecules 5 (130 sets of eight stereoisomers) of this class.

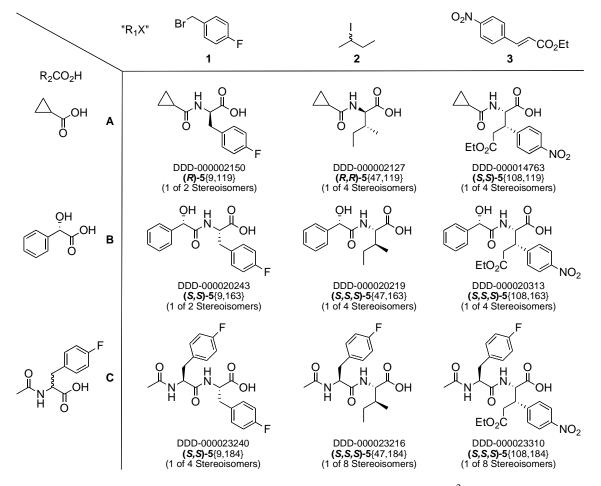
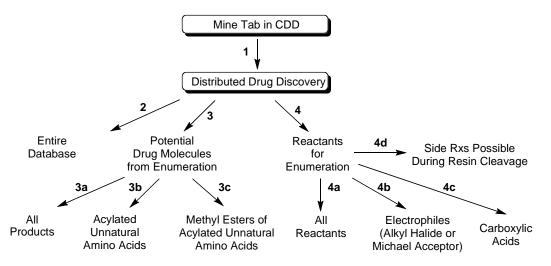


Table 1. Sample 3 x 3 Enumeration Representative of 48,608 Member D³ Catalog.

Tutorial. This tutorial is a guide to using the IUPUI Distributed Drug Discovery (D^3) database residing on the Collaborative Drug Discovery (CDD) website. The tutorial is accessible from: a) the Supporting Information in Scott, W. L.; O'Donnell, M. J. et al. J. Comb. Chem. DOI: 10.1021/cc800184v (http://dx.doi.org/10.1021/cc800184v) and DOI: 10.1021/cc800185z (http://dx.doi.org/10.1021/cc800185z) b) the CDD website; or c) the IUPUI Department of Chemistry and Chemical Biology website [http://chem.iupui.edu/ (accessed November 2, 2008] under faculty & staff directory for O'Donnell or Scott. The tutorial available at (ii) and (iii) will be updated periodically.

- 1) To Obtain Username and Password for the IUPUI Distributed Drug Discovery (D^3) database. For access to the IUPUI - Distributed Drug Discovery (D^3) database, register for a free Collaborative read-download with Drug Discovery (CDD) account https://www.collaborativedrug.com/register/iupui-d3 (accessed November 2, 2008) by completing the "Sign up for IUPUI – Distributed Drug Discovery (D3)" information.
- 2) Login to the IUPUI Distributed Drug Discovery (D^3) database. URL: http://www.collaborativedrug.com/ (accessed November 2, 2008) Username: Your Username Password: Your Password Select: "Mine" on the Dashboard page (upper right-hand corner)
- 3) Mining.

Allows quick searching of the entire D^3 database (or a subset thereof) using "protocols". Searches can be further refined with Structures, Chemical Properties and/or Molecule Keywords. Scheme 2 outlines the D^3 database organization and protocol structure:



Select: Distributed Drug Discovery from pull-down "protocol" -> Distributed Drug Discovery Work Site
 Select: (any Distributed Drug Discovery protocol) -> Entire Distributed Drug Discovery Database

- Select: Potential Drug Molecules from Enumeration
- **3a**. Option: (any readout) -> All Enumerated Products

3b. Option: Acylated Unnatural Amino Acids -> All Enumerated Acylated Unnatural Amino Acids

- 3c. Option: Methyl Esters of Acylated Unnatural Amino Acids -> All Enumerated Methyl Esters
- 4. Select: Reactants for Enumeration
 - 4a. Option: (any readout) -> All Reactants
 - 4b. Option: Electrophiles (Alkyl Halide or Michael Acceptor) -> All Electrophile Reactants
 - 4c. Option: Carboxylic Acids -> All Carboxylic Acid Reactants

4d. Option: Side Rxs Possible During Resin Cleavage -> Reactant Functionality Transformed During Cleavage

Scheme 2. D³ Database Organization and Protocol Structure.

From the Mine tab, "any protocol type", "any protocol", and "any readout" pull-down menus can be specified. One or more selections lead to all possible search results within the chosen category. Table 2 lists the results obtained from searches in the indicated categories (numbers in parentheses refer to steps in Scheme 2):

I /	
Protocols	Search Results ^a
(any protocol type)	
(any protocol)	Default
Distributed Drug Discovery (1)	
(any Distributed Drug Discovery protocol) (2)	48,818
Potential Drug Molecules from Enumeration (3)	
(any readout) (3a)	48,608
Acylated Unnatural Amino Acids (3b)	24,416
Methyl Esters of Acylated Unnatural Amino Acids	(3c) 24,192
Reactants for Enumeration (4)	
(any readout) (4a)	212
Electrophiles (Alkyl Halide or Michael Acceptor) (4	4b) 112
Carboxylic Acids (4c)	100
Side Rx Possible During Resin Cleavage (4d)	22
a	

^a "Add a term" to modify a search by: Structures, Chemical Properties and/or Molecule Keywords

 Table 2. Results Obtained from Different Protocol Searches.

Example 1: All Products Search

Select: Mine tab Select: New Query Select the following from the pull-down protocol menus: *Distributed Drug Discovery*

Potential Drug Molecules from Enumeration

(any readout)

Note: "(any readout)" means the default has been automatically chosen.

- Select: Search Molecules
- Result: Displays first 100 of 48,608 acylated unnatural amino acids and methyl esters from enumeration.
- Select: Third molecule (DDD-000001000) in list of products. Note that the first two products in the list are both reactant carboxylic acids (DDD-0000000267 and DDD-000000274) and enumeration products (DDD-000001010 and DDD-000013558, respectively).
- Result: Left column information about the product and the reactants used for its preparation in the enumeration; right column CDD-generated product information (see Appendices 1 and 2 for a listing of product and reactant information).

4) File Downloads

The three files described below are available for download at several locations in the database. The most direct route to these downloads follows:

Select: Archive tab Select: Protocols Select: "Potential Drug Molecules from Enumeration" Select (under Associated Files): Filename

- a) "IUPUI D3 Enumerated Products.sdf": The SD File of the 48,608 enumerated products. This file contains the fields described in Appendix 3.
- b) "IUPUI D3 Tutorial.pdf": The tutorial will be updated periodically and be available through CDD at this site and also on the IUPUI Department of Chemistry and Chemical Biology website [http://chem.iupui.edu/ (accessed November 2, 2008)] under the faculty & staff directory for O'Donnell or Scott.
- c) "IUPUI D3 Refs Rxs Electrophiles.pdf": This file contains selected lead references, cited in Scott, W. L.; O'Donnell, M. J. et al. J. Comb. Chem. DOI: 10.1021/cc800184v (<u>http://dx.doi.org/10.1021/cc800184v</u>) for reaction of each electrophile (alkyl halide or Michael acceptor) with Schiff bases of amino acid derivatives under a variety of basic conditions.
- 5) Substructure or Full Structure Searches

Select: Mine tab

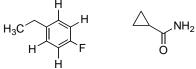
Select: New Query

Select: Structures – by substructure

Select: "Launch the editor" to build a structure for this search (opens MarvinSketch applet) Note: A dialog box may appear with information about a certificate or digital signature. If this happens, select "Trust" or "Run."

a) Example 2: Substructure Search

Draw the following search query in MarvinSketch (see Appendix 4 for detailed drawing instructions).



Select: Use This Structure

Result: Structures in Mine Protocols and Molecules structure box.

Note: Sometimes a "?" appears in the resulting structure box, indicating the image did not load but the structure is still stored and one can continue.

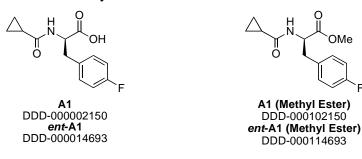
Select the following from the pull-down protocol menus:

Distributed Drug Discovery

Potential Drug Molecules from Enumeration (any readout)

Select: Search Molecules

Result: 4 matching compounds from the 48,606 enumerated products; the two enantiomeric carboxylic acids and their two enantiomeric methyl esters:



 b) Example 3: Substructure Search Select: Mine tab Select: New Query Select the following from the pull-down protocol menus: Distributed Drug Discovery Potential Drug Molecules from Enumeration Acylated Unnatural Amino Acids
 Draw the following search query in MaryinSketch (see Appendix 4 for

Draw the following search query in MarvinSketch (see Appendix 4 for detailed drawing instructions):

Select: Use This Structure

Select: Search Molecules

Result: 1518 hits from the database of 24,416 amino acid products (Table 3). The 24 hits from Table 1 are summarized below:

Structure in Table 1	Number of Stereoisomers
A1	2
B1	2
C1	4
C2	8
C3	8

Table 3. Stereochemical Results from Example 3.

 c) <u>Example 4</u>: Substructure Search Demonstrating Stereochemistry and Side Reactions Select: Mine tab Select: New Query

Select: New Query

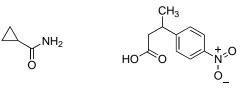
Select the following from the pull-down protocol menus:

Distributed Drug Discovery

Potential Drug Molecules from Enumeration

(any readout)

Draw the following search query in MarvinSketch (see Appendix 4 for detailed drawing instructions):



Select: Use This Structure

Select: Search Molecules

Result: 8 matching results from the database of 48,608 amino acid and methyl ester products (Table 4). The 4 stereoisomers of the acylated unnatural amino acids (5) are represented in the top row while the 4 stereoisomers of the methyl esters (6)

are in the bottom row. For the methyl ester products a possible side reaction during resin cleavage (transesterification of the ethyl ester in the Michael acceptor) has been carried out. Notation of this side reaction plus a lead reference are obtained by searching for the electrophile used in this preparation (Michael acceptor DDD-000000121).

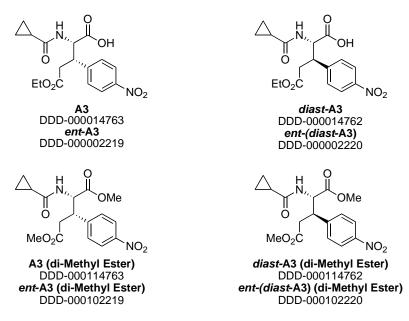
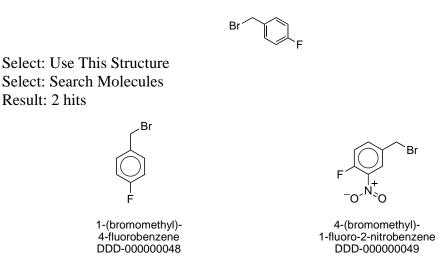


Table 4. Results from Example 4.

- 6) Obtaining Information About a Particular Molecule
 - a) <u>Example 5</u>: Product Information Search
 - From a Search: Under Search Results, select structure or DDD number
 Using a D³ Number: From the original Mine page:
 - From the original Mine page:
 Select: New Query
 In Molecule Keywords type: DDD-000002150 (shortcut: DDD 2150)
 Select: Search Molecules
 Result: 11 hits. The CDD search strategy returns molecules that begin with "2150": DDD-000002150 and the 10 molecules DDD-000021500 through DDD-000021509. Lower DDD registry numbers appear first in the search results, so DDD-000002150 is the first entry.
 b) Example 6: Reactant Information Search (see Appendix 2 for a listing of reactant

information) Select: Mine tab Select: New Query Select the following from the pull-down protocol menus: Distributed Drug Discovery Reactants for Enumeration (any readout) Draw the following search query in MarvinSketch (see Appendix 4 for detailed drawing instructions). To insert the Br: Position the cursor over the H_3C and type Br or select "More", choose Br, Close, Click on H_3C .



Select: Structure or name [1-(bromomethyl)-4-fluorobenzene], to obtain information about this reactant (see Appendix 2 for a listing of reactant information)

7) Selecting Other Public Data Sets

Other public data sets besides the Distributed Drug Discovery (D3) database are available as follows:

Select: Collaborate tab Select: Data Sets Select: Select all or make individual selection(s) Select: Save selections Select: Mine tab and proceed with search

Information about Individual Enumeration Products

- 1) Information about the enumeration product and the reactants used for its generation (left column).
 - Definition (category content): Name (DDD Number) Synonyms Description (Blank) Structure (SMILES or InChI or InChIKey) Molecular Formula Product Type Reactant: Electrophile (Chemical Name) Reactant: Electrophile DDD No. Reactant: Carboxylic Acid (Chemical Name) Reactant: Carboxylic Acid DDD No.

2) <u>CDD-generated information about the enumeration product (right column)</u>.

Structure Lipinski properties Molecular weight log P H-bond donors H-bond acceptors Lipinski Rule of 5 Additional properties Formula pKa Exact mass Atom count Composition Topological polar surface area (PSA)

Information about Individual Reactants

1) Information about the reactant (left column).

Definition (category content): Name (DDD Number) Synonyms (Chemical Name(s)) Description (Blank) Structure (SMILES or InChI or InChIKey) Reactant Type Molecular Formula *Reference No. for Rxs with Schiff Bases of Amino Acid Derivatives *Side Rxs Possible During Resin Cleavage *These categories are only shown when an entry is present.

2) CDD-generated information about the reactant (right column).

Structure Lipinski properties Molecular weight log P H-bond donors H-bond acceptors Lipinski Rule of 5 Additional properties Formula pKa Exact mass Atom count Composition Topological polar surface area (PSA)

Information Available for Enumeration Products in SD File Download

Structure **SMILES** Name (DDD Number) Product Type (Acylated Unnatural Amino Acid or Methyl Esters of Unnatural Amino Acids) Molecular weight log P H-bond donors H-bond acceptors Lipinski Rule of 5 Formula рКа Exact mass Atom count Composition Topological polar surface area (PSA) Reactant: Electrophile (Chemical Name) Reactant: Electrophile DDD No. Reactant: Carboxylic Acid (Chemical Name) Reactant: Carboxylic Acid DDD No.

Detailed Instructions for Drawing Structures for Example 2 (page 5)

Repeat Login Procedure

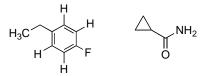
Select the following from the pull-down protocol menus:

Distributed Drug Discovery Potential Drug Molecules from Enumeration

(any readout)

Note: "(*any readout*)" means the default has been automatically chosen.

Target Substructures to Draw in Marvin Sketch:



From within the CDD Database:

Select: Mine tab

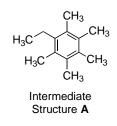
Select: New Query

Select: Structures – by substructure

Select: "Launch the editor" to build a structure for this search (opens MarvinSketch applet) Note: A dialog box may appear with information about a certificate or digital signature. If this happens, select "Trust" or "Run."

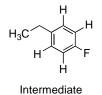
From within the Structure Editor of MarvinSketch:

- (1) Select the benzene ring from the structure palate.
- (2) Move the cursor to position the benzene ring on the left side of the drawing surface then click.
- (3) Select the bond tool (below "Help").
- (4) Position the cursor above one of the carbons so that a blue circle appears then click. Alternatively, you can click/drag/release to position the bond.
- (5) Repeat step 4 for all benzene ring carbons.
- (6) Position the cursor above one of the CH_3 groups then click.



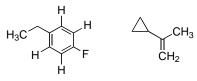
- (7) The hydrogens and fluorine on the benzene ring can be added in one of two ways:
 - (a) Position the cursor above a CH₃ group so that a blue circle appears then type the atom symbol or
 - (b) Select the atom ("H" or "F") in the palate then click above each CH_3 to be changed. If the atom symbol is not in the palate, select "More" for a complete list.

(8) Since not explicitly designated otherwise, the chain H₃C remaining is assumed to be a carbon with 0-3 hydrogens attached.



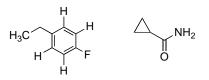
Structure **B**

- (9) Select the bond tool and click in the open space to the right of the previously drawn structure.
- (10) Build the carbon framework of the right-hand structure as before. Click/drag/release to complete the final bond of the cyclopropane ring. To create the second bond of the double bond, position the cursor over the center of the single bond (parentheses will appear) then click.



Intermediate Structures C

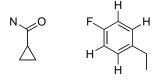
- (11) The heteroatoms can be added to the right-hand structure in one of two ways:
 - (a) Select the atom ("O" or "N") in the palate then click above the appropriate atom to be changed.
 - (b) Position the cursor above the CH_2 or CH_3 group (a circle will appear) then type either an O or an N, respectively.
- (12) As before, since not explicitly designated otherwise, the NH₂ is assumed to be a nitrogen with 0-2 hydrogens attached.



Final Search Structures **D** (in Structure Editor of MarvinSketch)

(13) Select "Use This Structure".

From within the CDD Database: Result: Structures in Mine Protocols and Molecules structure box.



Final Search Structures **E** (in CDD Search Window)

Note: Sometimes a "?" appears in the resulting structure box, indicating the image did not load but the structure is still stored and one can continue.

Select: Search Molecules

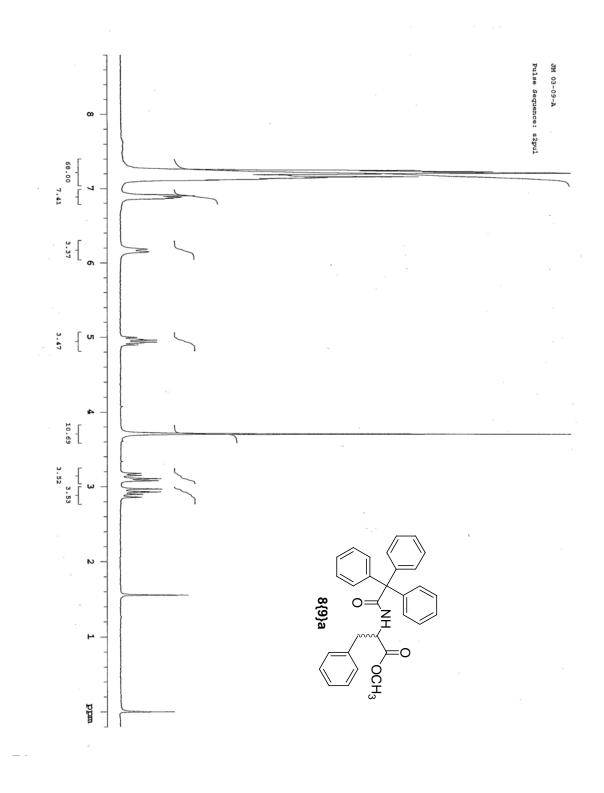
Result: 4 matching results from the 48,606 enumerated products; the two enantiomeric carboxylic acids and their two enantiomeric methyl esters:

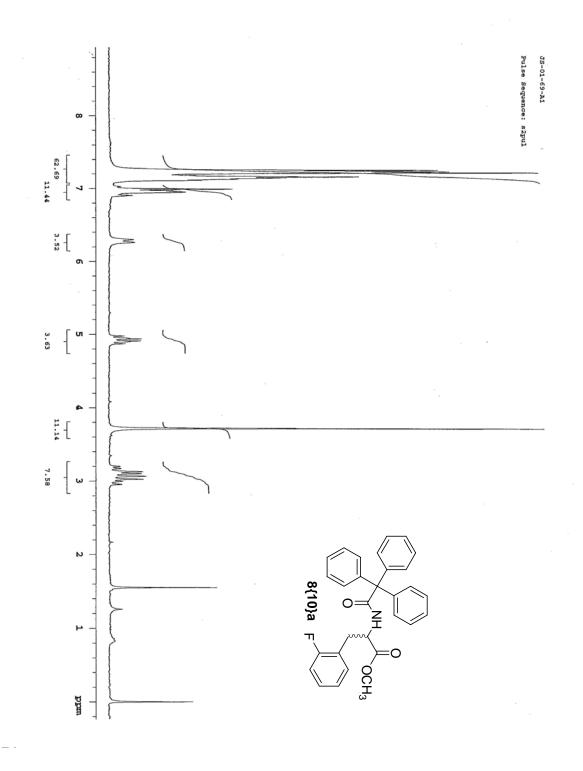
OH ö

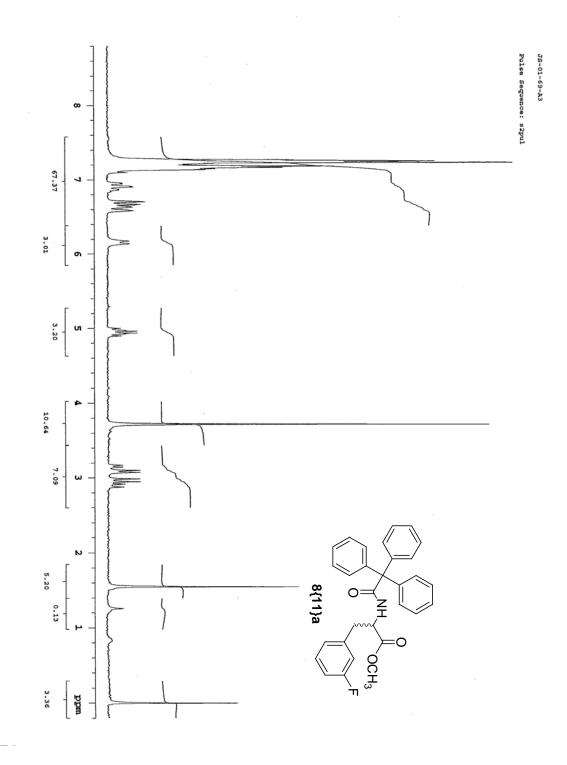
A1 DDD-000002150 *ent-*A1 DDD-000014693

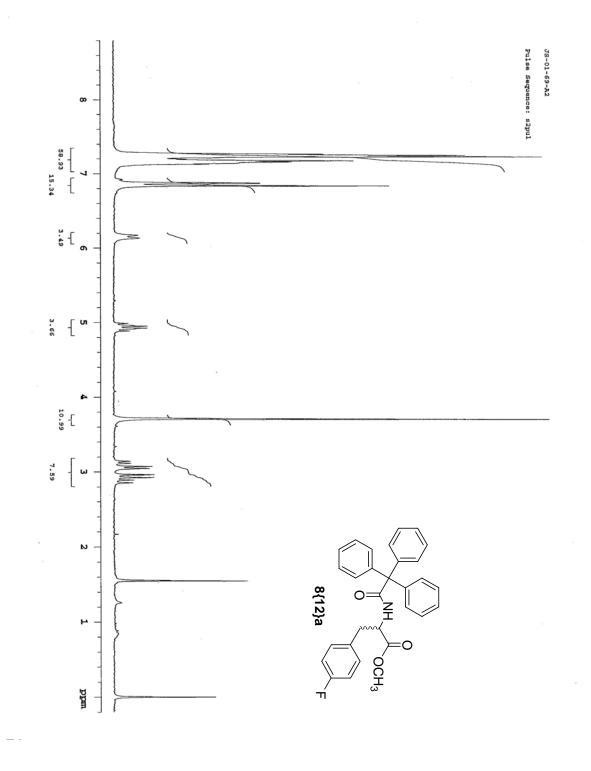
OMe Ô

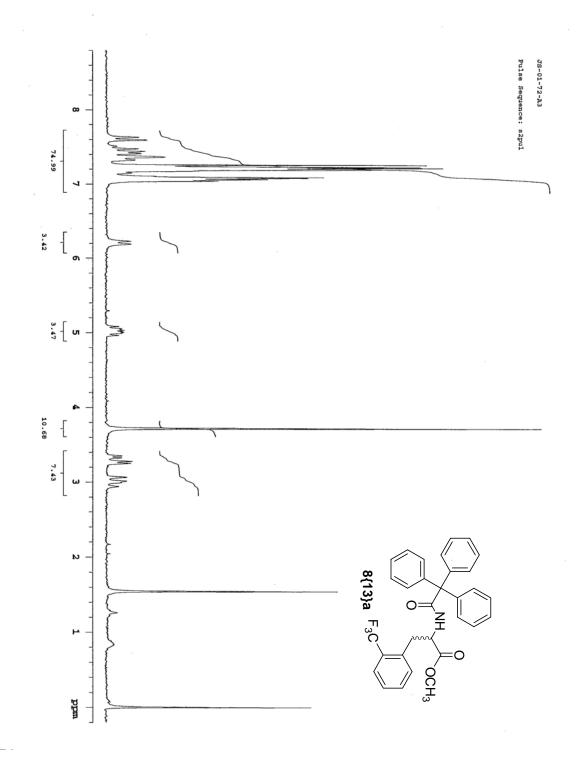
A1 (Methyl Ester) DDD-000102150 ent-A1 (Methyl Ester) DDD-000114693

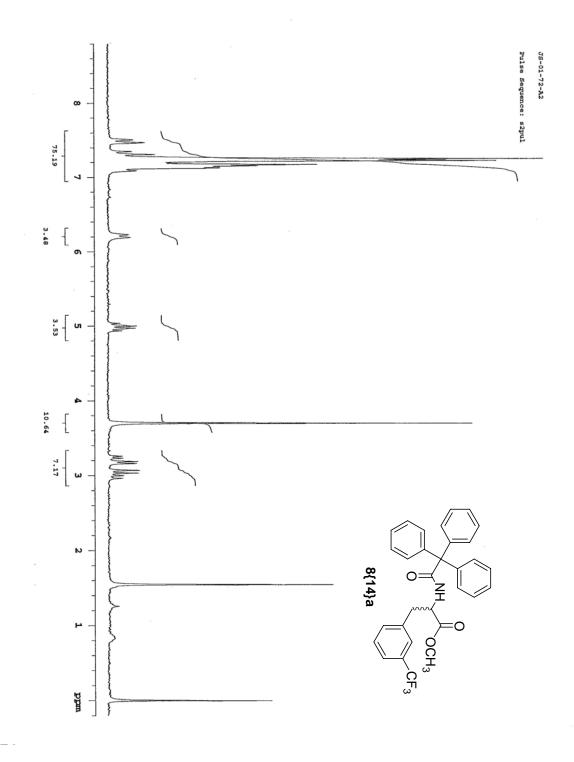


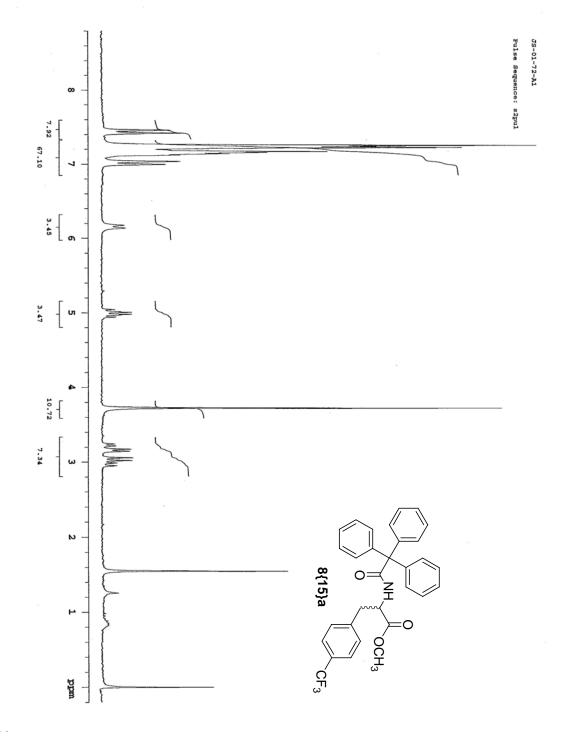


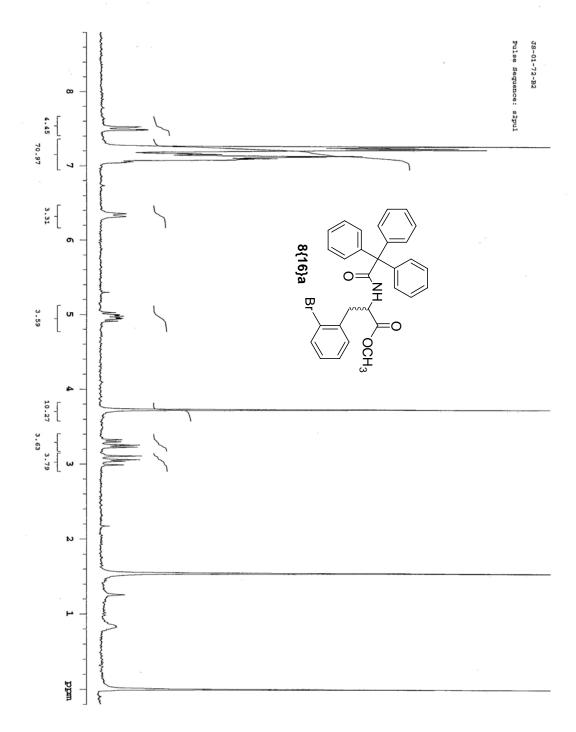




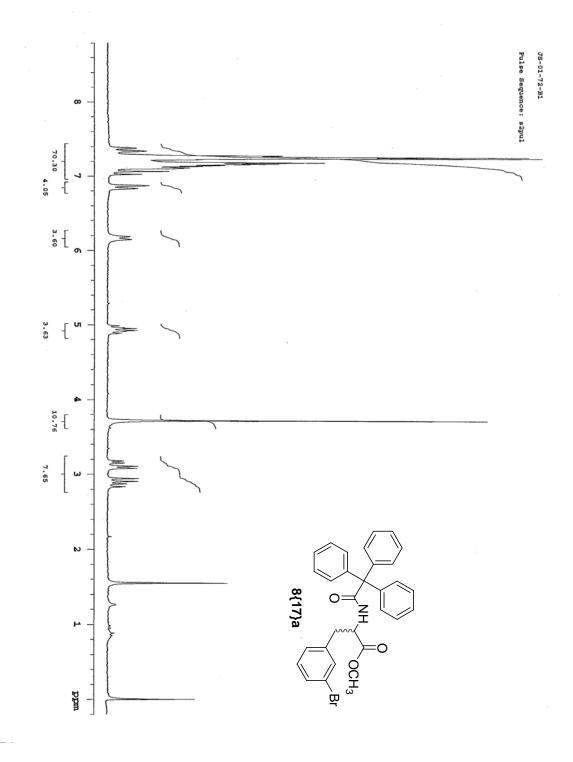


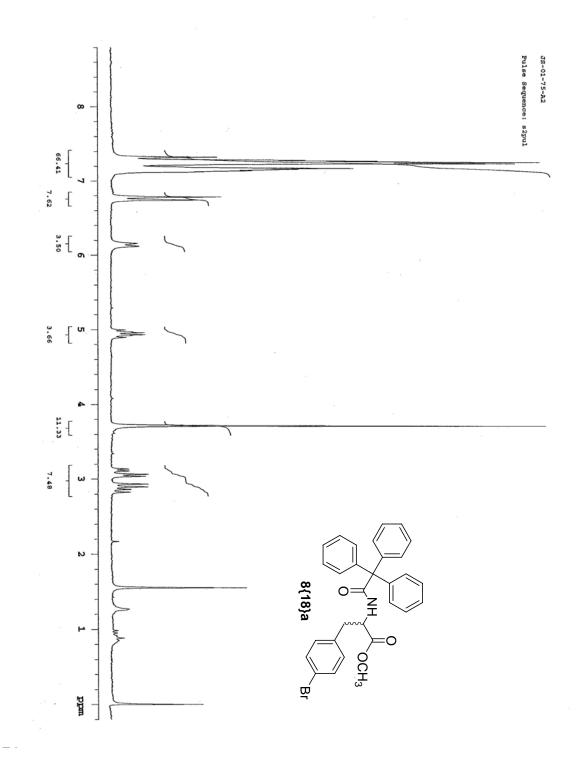


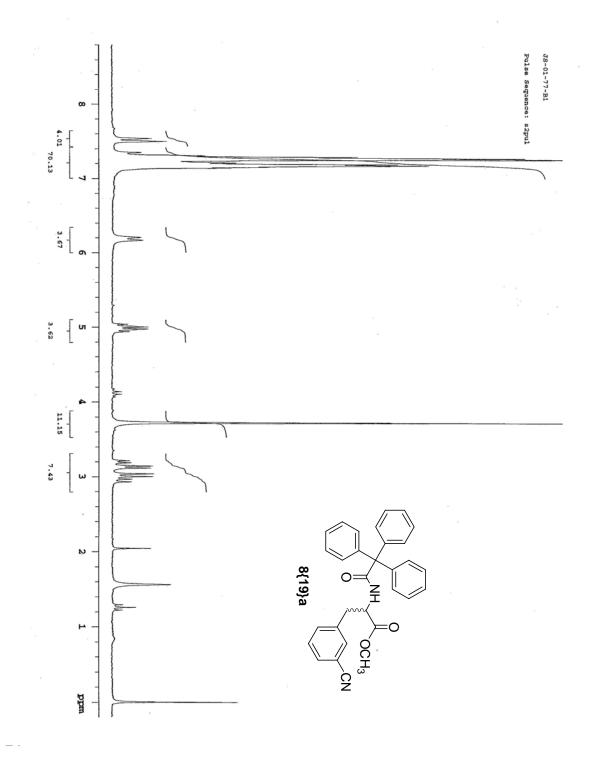


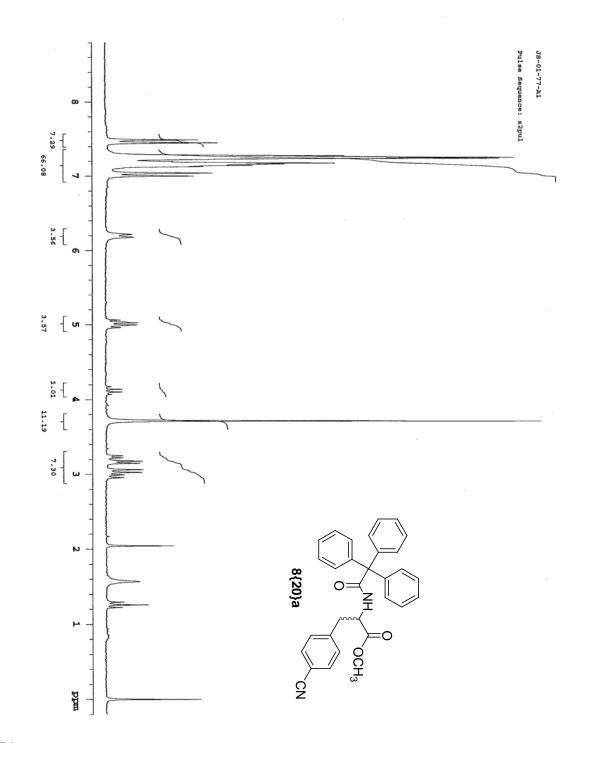


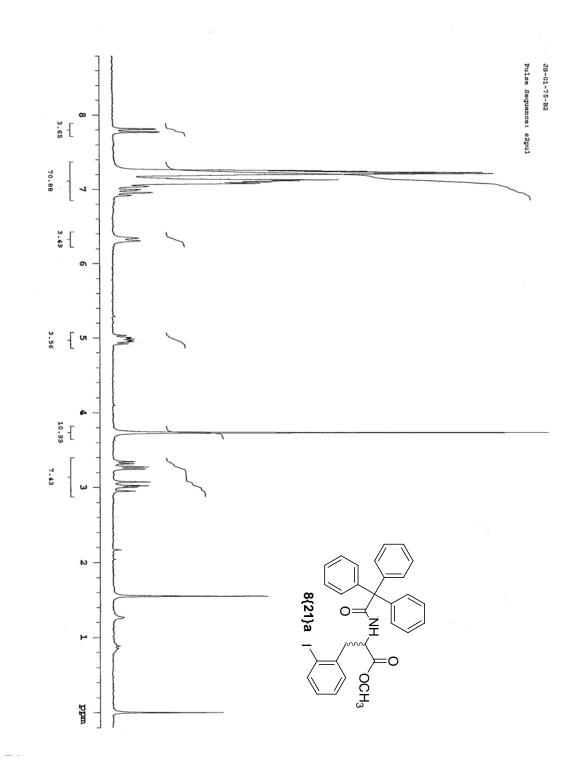
- -

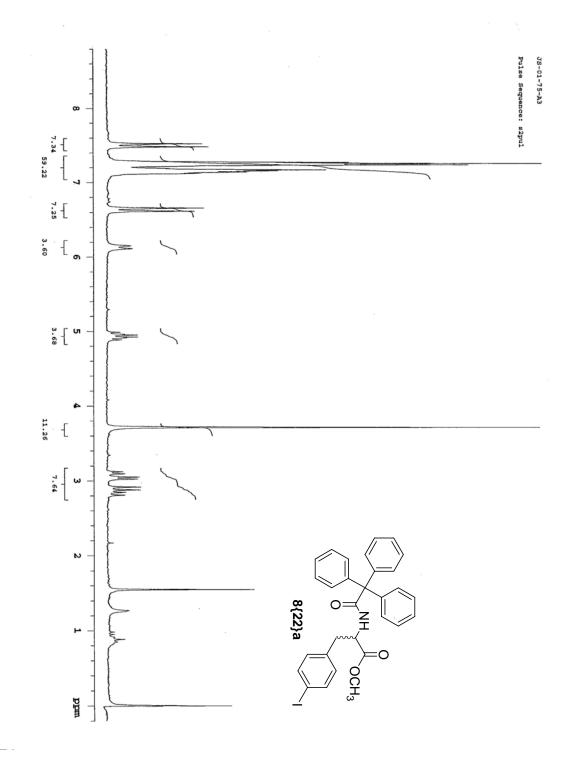


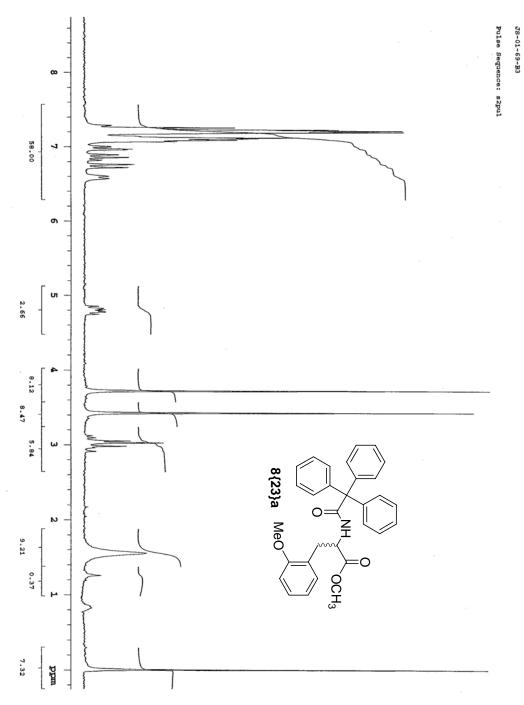


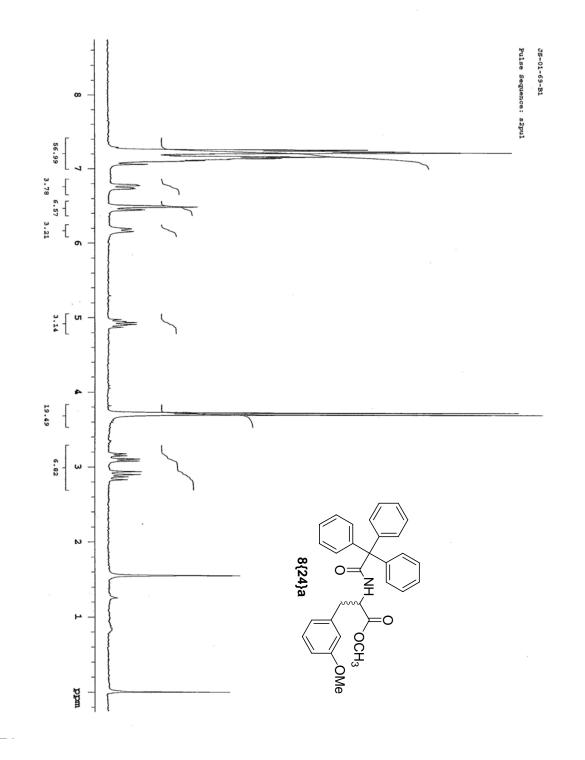


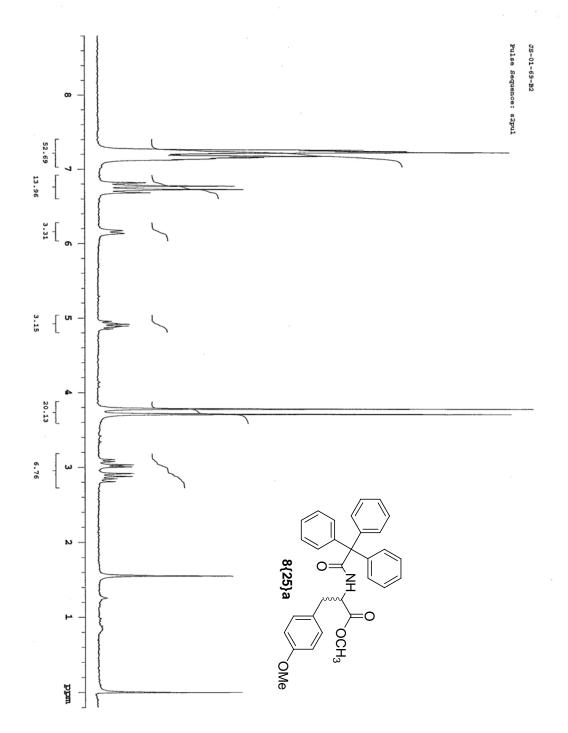




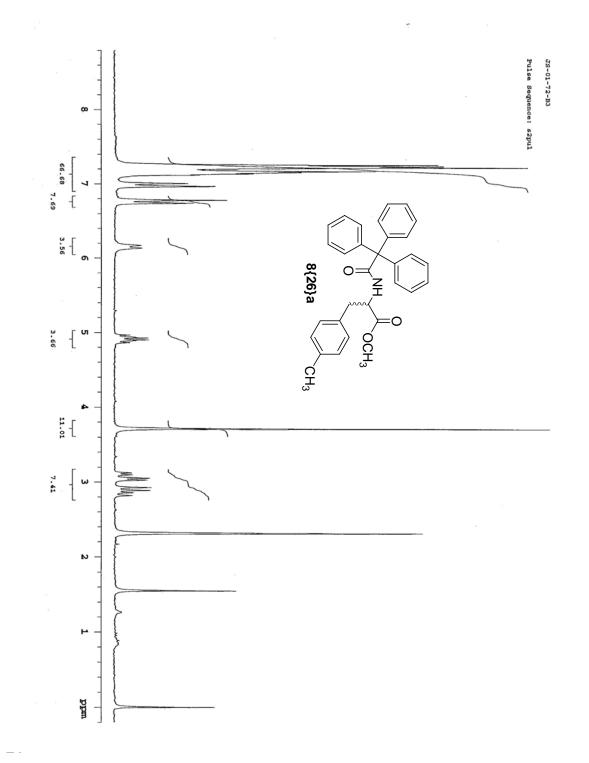


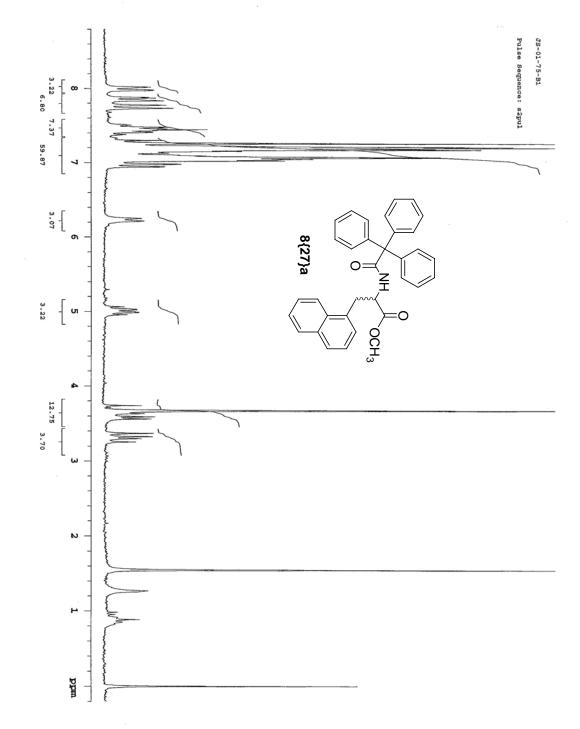






- -





- -

