

# INSTRUCTORS' NOTES

## Complete Analysis of a Biologically Active Tetrapeptide: A Project Utilizing Thin-Layer Chromatography and Mass Spectrometry

### Joseph W. LeFevre

Chemistry Department, State University of New York at Oswego, Oswego, NY 13126;  
*lefevre@oswego.edu*

### David W. Dodsworth

Bristol-Myers Squibb, P.O. Box 4755, Syracuse, NY 13221-4755; *dodsword@bms.com*

## I. Comments on the Experiment

### A. *N*-Terminal Amino Acid Analysis

The identification of DNS-alanine as the *N*-terminal amino acid was straightforward using the 15 standard DNS-amino acids provided. All but three amino acids were quickly eliminated. Dansyl-alanine, DNS-methionine, and di-DNS-lysine have similar  $R_f$ s in solvent A (see Table 1). These three DNS-amino acids were distinguished either by careful double spotting as described in the STUDENT HANDOUT #1, Section B.4. or by using solvents B and C, which easily resolved these three compounds.

### B. Identification of the Remaining Three Amino Acids

By comparing the  $R_f$ s of all four DNS-amino acids in the mixture with those of the 15 standards using solvent A, arginine, aspartic acid, glutamic acid, serine and threonine were quickly eliminated as possibilities. The purification of the four DNS-amino acids on 5 x 20 cm normal-phase TLC plates using solvent A, took approximately 2 hours for full development. A typical separation appears in Figure 2 in the JOURNAL COPY. A trace of dansyl-amide (DNS-NH<sub>2</sub>), a reaction by-product, appears near the solvent

front. Also, a trace of an unknown compound appears below DNS-NH<sub>2</sub>. Dansyl-glycine was easily identified since it was cleanly resolved from the other standards using solvent A. Dansyl-phenylalanine was more difficult to identify since its *R<sub>f</sub>* in solvent A was very close to that of DNS-methionine. These two compounds could not be separated in solvent B either. However, solvent C was successful in resolving them.

Dansyl-leucine could not be distinguished from DNS-valine in solvent A. However, solvents B and C were successful in separating these two compounds. Solvent C is used in conjunction with more expensive reversed-phase plates. It is only necessary to use this system to distinguish between DNS-phenylalanine and DNS-methionine if the double spotting technique using solvent A fails. The identification process can be simplified by eliminating DNS-methionine and DNS-valine as possibilities. The pertinent *R<sub>f</sub>*s for solvents A, B, and C are shown in Table 1. The TLC separations are significantly improved if the plates are predeveloped for a few seconds in methanol (MeOH), which concentrates the spots as narrow bands just above the origin. The average times for 6 cm development of solvents A, B, and C are included at the bottom of Table 1. Solvent A should always be used in a fume hood since pyridine has an obnoxious odor.

DNS-Amino Acid	Normal-phase		Reversed-phase
	Solvent A	Solvent B	Solvent C
Arg	0.00		
Asp	0.01		
Ser	0.07		
Glu	0.09		
Thr	0.10		
<b>Gly</b>	0.18		
Trp	0.25		
di-Tyr	0.28		
di-Lys	0.33	0.46	0.45
<b>Ala</b>	0.36	0.28	0.71
Met	0.38	0.38	0.64

<b>Phe</b>	0.41	0.36	0.54
Pro	0.47		
Val	0.52	0.36	0.67
<b>Leu</b>	0.56	0.30	0.60

Table 1.  $R_f$ s of DNS-Amino Acids in Various Solvents

Development time (min)	Solvents:
20	A = toluene:pyridine:AcOH/30:10:1 (v/v/v)
15	B = CH <sub>2</sub> Cl <sub>2</sub> :MeOH:EtOAc/12:4:1 (v/v/v)
20	C = MeOH:2% AcOH/75:25 (v/v/v)

### C. Determination of the Stereochemistry of the Amino Acids

Two 1 x 3" reversed-phase TLC plates per pair of students are needed for this part of the experiment. Each purified DNS-amino acid from the peptide was spotted next to its corresponding DL-racemate. Glycine is achiral. Dansyl-DL-phenylalanine, DNS-DL-leucine, DNS-glycine, and DNS-L-alanine are commercially available, but racemic DNS-alanine is not. Therefore, this compound must be synthesized as described below.

Racemic DNS-DL-alanine is prepared by reacting 500  $\mu$ L of 1 mM DL-alanine in 0.2 M NaHCO<sub>3</sub> (pH = 9.5) with 500  $\mu$ L of 10 mM DNS-Cl in acetone. After 1 hour the solution is acidified to pH = 2 – 4 by adding 1 M HCl. One mL of saturated NaCl is then added. The solution is extracted three times (1 mL each) with ethyl acetate, EtOAc. The EtOAc layers are combined, washed once with 2 mL of H<sub>2</sub>O, dried with anhydrous NaSO<sub>4</sub>, and filtered through a Pasteur pipet. The sample is sufficiently pure at this point for use in reversed-phase TLC.

The plates were developed in a solvent composed of acetonitrile (CH<sub>3</sub>CN):0.2 M  $\beta$ -CD/25:75 (v/v). The aqueous portion of the solvent was saturated first with urea to increase the solubility of the  $\beta$ -CD. The solution is also 0.6 M in NaCl to prevent the solvent from dissolving the binder which holds the reversed-phase stationary phase to the glass plate. The plates take about one hour to develop. Typical separations appear in Figure 3 in the JOURNAL COPY. The  $R_f$ s and  $\alpha$  values (separation factors) appear in Table 2.

Standard DNS-DL-amino acids	$R_{fD}$	$R_{fL}$	$\alpha^*$	$R_f$ of amino acid from peptide	stereochemistry
alanine	0.52	0.44	1.38	0.52	D
glycine**	0.60	—	—	0.60	—
phenylalanine	0.33	0.27	1.33	0.27	L
leucine	0.37	0.28	1.51	0.37	D

\*  $\alpha = [(1 - R_{fL\text{-isomer}})/R_{fL\text{-isomer}}]/[(1 - R_{fD\text{-isomer}})/R_{fD\text{-isomer}}]$

\*\* There is no D- or L-isomer.

Table 2. Reversed-Phase TLC Separation Data of the DNS-Amino Acids

Rather than sending students to the literature to determine that the L-isomers have smaller  $R_f$ s than the corresponding D-isomers, it is possible to obtain this information by spotting commercial samples of the L-isomers of Ala, Phe, and Leu next to the corresponding DL-racemates of these compounds.

If enough of the DNS-amino acids (except for glycine) from the peptide is spotted, a small amount of the opposite isomer may be visible on the TLC plates. This is due to partial racemization during the 18 hour hydrolysis of the tetrapeptide. This is especially true for Leu.

Several students questioned why the standard DNS-glycine sample gave only one spot. After close examination of the structure, they were able to determine that DNS-glycine has no enantiomer.

#### D. Determination of the Sequence of the Peptide

**It is important to wait until all of the wet chemistry is completed and the amino acids are identified before handing out the mass spectrum.** Otherwise the students will be able to figure out the identity of the amino acids from the mass spectrum alone. The sole purpose of the mass spectrum in this experiment is to determine the sequence after the individual amino acids have been identified.

The fragmentation pattern of the tetrapeptide appears in Figure 3, showing the  $Y^+$  and  $B^+$  ion series. See reference 7 in the JOURNAL COPY. The data is summarized in Table 3.

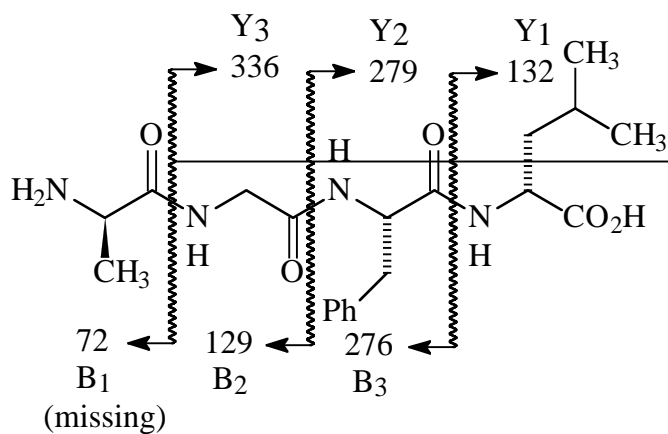


Figure 3. Fragmentation pattern of the tetrapeptide

High Mass	Low Mass	$\Delta$ Mass	Ion Type	Amino Acid Lost	Sequence
389*	276	113	B <sub>3</sub>	Leu	Leu
276	129	147	B <sub>2</sub>	Phe	Leu-Phe
129	72**	57	B <sub>1</sub> **	Gly	
407	336	71	Y <sub>3</sub>	Ala	Ala
336	279	57	Y <sub>2</sub>	Gly	Ala-Gly
279	132	147	Y <sub>1</sub>	Phe	Ala-Gly-Phe
132					

\* 389 = 407 (parent ion) – 18 (H<sub>2</sub>O)

\*\* The B<sub>1</sub> ion at 72 is not present.

Table 3. MS/MS data from the tetrapeptide

Putting all of the data together, the tetrapeptide is (*N*-term) D-Ala-Gly-L-Phe-D-Leu (*C*-term).

The MS/MS spectrum was acquired using the following conditions:

Spectrometer Manufacturer – Micromass

Model – VG Quattro

Configuration – Tandem quadrupole mass spectrometer with a QhQ (quadrupole mass analyzer, hexapole collision, quadrupole mass analyzer) configuration

Ion source – Electrospray Ionization (ESI)

Ionization – Positive ion

Collision gas – Argon

Calibration – Performed using the ammoniated molecular ion series of polyethylene glycol (PEG)

Sample – A portion of the tetrapeptide (96 µg) was dissolved in 1.0 mL of H<sub>2</sub>O:CH<sub>3</sub>CN/80:20 (v/v) containing 0.1% formic acid. A 10 µL injection was made. A Harvard 20 Syringe Pump provided continuous infusion of the dissolved analyte.

#### E. Answers to Questions

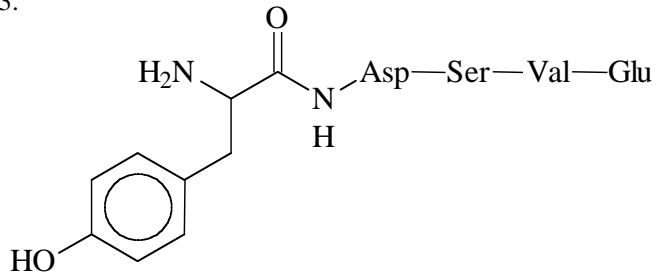
1. At lower pH the amino group is protonated so no reaction can occur with DNS-Cl. At higher pH OH<sup>-</sup> hydrolyzes DNS-Cl to DNS-OH.

2. Section A:  $(0.001 \text{ moles}/1000 \text{ mL}) \times 0.1 \text{ mL} = 1 \times 10^{-7} \text{ mol}$

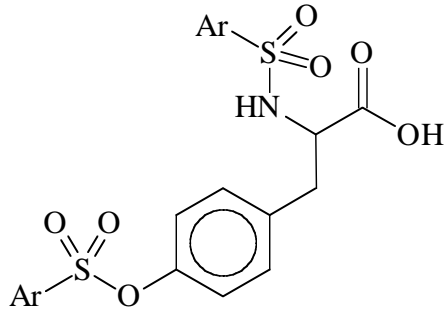
$$1 \times 10^{-7} \text{ mol} \times 407 \text{ g/mol} \times 10^6 \text{ µg/g} = 40.7 \text{ µg}$$

Section B: 200 µL was used which corresponds to  $2 \times 10^{-7} \text{ mol}$  or 81.4 µg

3.



1) DNS-Cl  
 2) H<sup>+</sup>  
 3) 6 M HCl, 110°C,  
 18 hr



+ Asp HCl + Ser HCl  
 + Val HCl + Glu HCl

di-DNS-tyrosine

4.

Least polar  
 (highest  $R_f$ )

Most polar  
 (lowest  $R_f$ )

val < gly < thr < asp

Side chain: -CH(CH<sub>3</sub>)<sub>2</sub>      -H      -CH(CH<sub>3</sub>)OH      -CH<sub>2</sub>CO<sub>2</sub>H

## II. Equipment

### A. General

Eppendorf pipets (10 – 100  $\mu$ L and 100 – 1000  $\mu$ L)

Infrared heat lamps

Long-wave (366 nm) UV lamp

Oven

Rotary evaporator

Scissors

### B. Individual

Two 10-mL beakers

Three 5-mL conical vials (for extractions)

Cotton

Five 1-dram vials

10-mL graduated cylinder

Microspatula

Pasteur pipets and bulbs

pH Paper

10-mL Round-bottom flask

Tweezers

Watchglass (6 cm)

### C. TLC

One 5 x 20 cm developing chamber with lid

Three jars, 4 oz. screw cap

Microburner

Micropipets

Reversed-phase TLC plates (KC18F, 1 x 3", Alltech Associates)

Silica gel TLC plates (plastic backed, #13181, Eastman Kodak). Cut 20 x 20 cm sheets into 3.5 x 7.5 cm and 5 x 20 cm strips.

### III. Reagents

The materials listed here are sufficient for at least 10 pairs of students.

Acetonitrile (150 mL)

0.2 M  $\beta$ -CD (100 mL)

Dansyl chloride (15 mL of a 10 mM solution in acetone)

Ethyl acetate (100 mL)

1 M HCl (5 mL)

6 M HCl (5 mL)

12 M HCl (5 mL)

Methanol (200 mL)

NaCl (10 mL of a saturated solution)

0.2 M NaHCO<sub>3</sub>, pH = 9.5 (15 mL)

Anhydrous Na<sub>2</sub>SO<sub>4</sub> (1 g)

1 mM tetrapeptide solution (3 mL)

#### IV. TLC Solvents

Toluene:pyridine:acetic acid (500 mL)

CH<sub>2</sub>Cl<sub>2</sub>:MeOH:EtOAc/12:4:1 (100 mL)

MeOH:2% acetic acid/75:25 (100 mL)

#### V. Preparations

- A. 1 mM tetrapeptide: Dissolve 5 mg of [des-Tyr<sup>1</sup>-D-Ala<sup>2</sup>-D-Leu<sup>5</sup>] enkephalin (Sigma #E 5390, FW = 406.5) in 12.3 mL of deionized water. Three mL will serve 10 pairs of students.
- B. 10 mM dansyl chloride: Dissolve 40 mg of dansyl chloride (Aldrich #D14,335-9, FW = 269.75) in 15 mL of reagent grade acetone. Filter the solution to remove traces of residual dansyl sulfonic acid, which is insoluble in acetone.
- C. 0.2 M NaHCO<sub>3</sub>: Dissolve 1.68 g of reagent grade NaHCO<sub>3</sub> (FW = 84.01) in 80 mL of deionized water in a 100-mL volumetric flask. Add 2 N NaOH dropwise until the pH = 9.5. Add water to the 100-mL line. Cap and shake well.
- D. 0.2 M β-CD: Saturate 60 mL of deionized water with reagent grade urea. Filter 80 mL of the solution through a fluted filter paper and glass funnel into a 100-mL volumetric flask. Add 22.7 g of β-cyclodextrin hydrate (CERESTAR, USA, FW = 1135) and 3.5 g of NaCl. Add more saturated urea solution to the 100-mL line, cap and mix well. The cloudiness will dissipate after several hours.
- E. Toluene:pyridine:acetic acid/30:10:1 (v/v/v): For 820 mL, combine 600 mL toluene, 200 mL of pyridine and 20 mL of glacial acetic acid. Mix well and store in a tightly stoppered bottle in a fume hood.
- F. CH<sub>2</sub>Cl<sub>2</sub>:MeOH:EtOAc/12:4:1 (v/v/v): For 170 mL, combine 120 mL of CH<sub>2</sub>Cl<sub>2</sub>, 40 mL of MeOH and 10 mL of EtOAc. Mix well and store in a tightly stoppered bottle.

- G. MeOH:2% acetic acid/75:25 (v/v): For 100 mL combine 75 mL of MeOH and 25 mL of 2% acetic acid in water. Mix well and store in a tightly stoppered bottle.

#### IV. Hazard Alert

- A. Acetic Acid [*Registry of Toxic Effects of Chemical Substances (RTECS)*; 8 vols.; U. S. Department of Health and Human Services, National Institute for Occupational Safety and Health, U. S. Government Printing Office: Washington, D. C., 1987 #AF1225000]: Corrosive. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- B. Acetone [*RTECS* #AL3150000]: Flammable and irritant. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- C. Acetonitrile [*RTECS* #AL7700000]: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- D. Beta-cyclodextrin hydrate [*RTECS* #GU22993000]. Avoid inhaling dust and ingesting the compound.
- E. Dansyl chloride [*RTECS* #QK3688000]: Corrosive. Moisture sensitive. Irritant. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
- F. Dichloromethane [*RTECS* #PA8050000]: Toxic and irritant. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- G. Ethyl acetate [*RTECS* #AH5425000]: Flammable and irritant. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- H. Hydrochloric acid [*RTECS* #MW4025000]: Highly toxic and corrosive. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- I. Methanol [*RTECS* #PC1400000]: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- J. Pyridine [*RTECS* #UR8400000]: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- K. Sodium chloride [*RTECS* #VZ4725000]: Irritant and hygroscopic. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.

- L. Sodium hydrogen carbonate [RTECS #VZ0950000]: Moisture sensitive. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
- M. Sodium hydroxide [RTECS #WB4900000]: Corrosive and toxic. Prevent eye, skin and clothing contact. Avoid ingesting the compound.
- N. Sodium sulfate [RTECS #WE1650000]: Irritant and hygroscopic. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
- O. Silica gel [RTECS #VV8850000]: Hygroscopic. Avoid inhaling dust or ingesting the compound.
- P. Toluene [RTECS #XS5250000]: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
- Q. Urea [RTECS #YR6250000]: Irritant. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.

# STUDENT HANDOUT #1

## Complete Analysis of a Biologically Active Tetrapeptide: A Project Utilizing Thin-Layer Chromatography and Mass Spectrometry

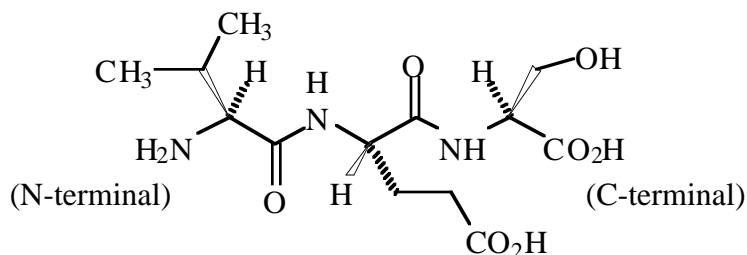
Joseph W. LeFevre\*

Chemistry Department, State University of New York at Oswego, Oswego, NY 13126;

\**lefevre@oswego.edu*

### I. INTRODUCTION

In peptides, amino acids are linked together one after another via amide bonds. Shown below is a tripeptide composed of L-valine-L-glutamic acid-L-serine. Valine is the *N*-terminal amino acid and serine is the *C*-terminal one.



The goals of the experiment will be four-fold. First, you will determine the *N*-terminal amino acid residue by reacting the tetrapeptide with 5-dimethylamino-1-naphthalenesulfonyl (dansyl, DNS-) chloride. After hydrolysis of the DNS-tetrapeptide, the *N*-terminal DNS-amino acid will be identified by normal-phase thin-layer chromatography (TLC). Second, following hydrolysis of a fresh portion of the peptide and conversion to the corresponding DNS-derivatives, you will determine the identities of the remaining three amino acids by TLC. Third, once you have identified the four amino acids present, you will determine the stereochemistry of each using reversed-phase TLC. Finally, you will determine the sequence of the amino acids in the peptide by mass spectrometry. In order to complete the work in four weeks, you will work in pairs.

## II. DISCUSSION

Dansyl chloride (DNS-Cl) has been used in the determination of the *N*-terminal amino acid residue of a peptide or protein since 1963 (1,2). It has found wide application because of its intense fluorescence following derivatization with an amino group enabling one to study very small quantities of peptides, proteins and amino acids. In *N*-terminal analysis, DNS-Cl reacts under basic conditions (pH 8.5 - 10.5) with the free amino group to form an *N*-DNS-peptide containing a primary sulfonamide. Proline is an exception in that it forms a secondary sulfonamide. To complete the process, the amide bonds of the peptide are hydrolyzed using 6 M HCl and heat to liberate the *N*-terminal DNS-amino acid and the free amino acids. Primary and secondary sulfonamides are stable to these hydrolysis conditions. These reactions are illustrated in Figure 1 for a hypothetical tetrapeptide where R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> and R<sub>4</sub> may vary depending upon the particular amino acids present.

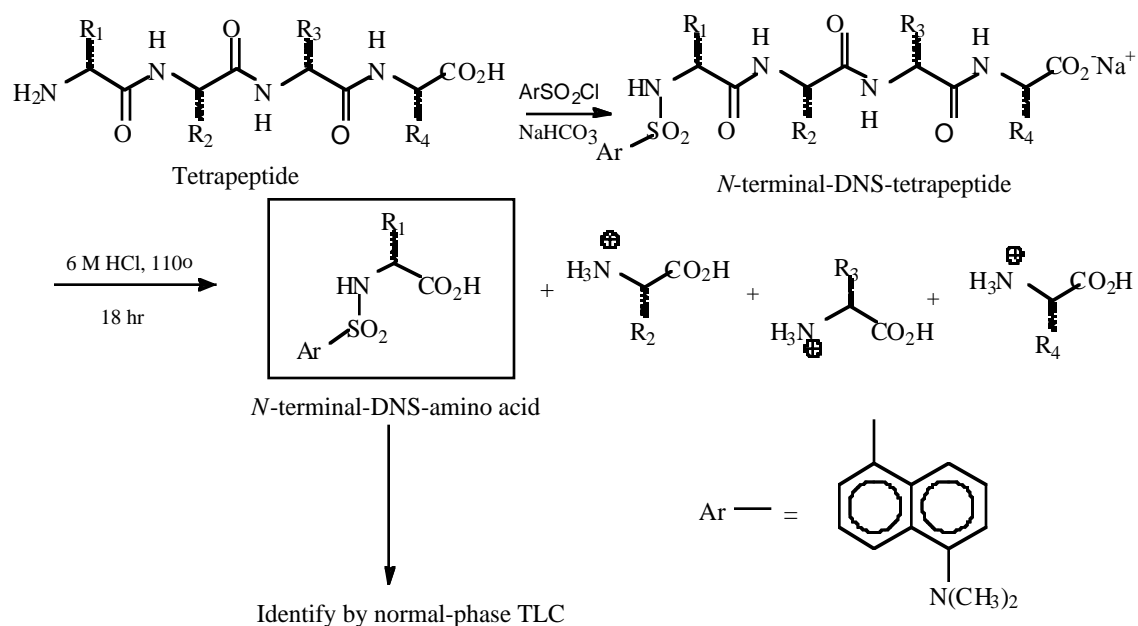


Figure 1. Formation and Hydrolysis of an *N*-DNS-tetrapeptide

For *N*-terminal amino acids with reactive side chains such as an amino group (lysine), an imidazole group (histidine), or a phenolic hydroxyl group (tyrosine), di-DNS-derivatives are formed as shown in Figure 2.

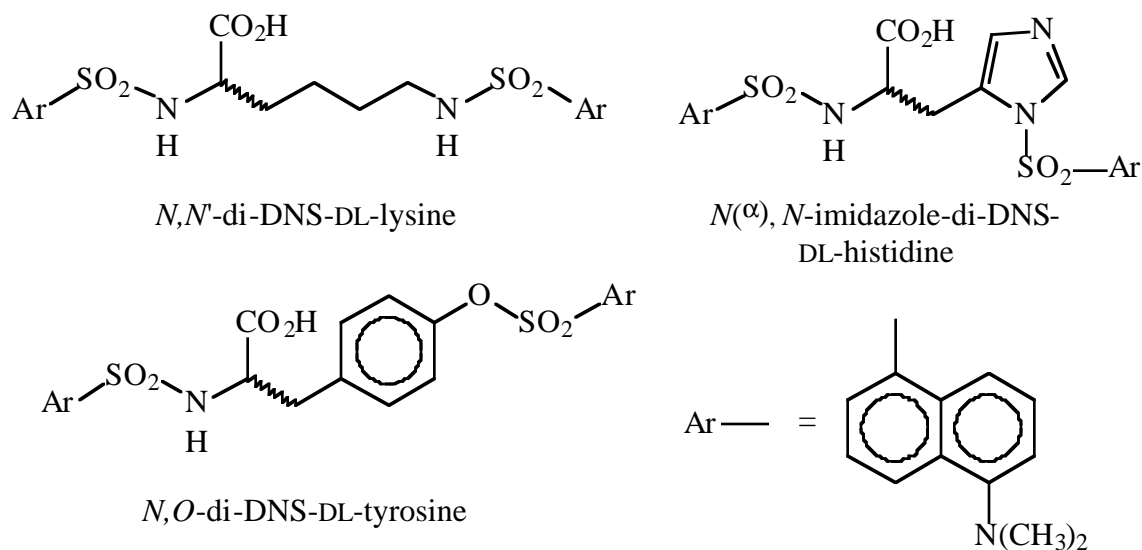


Figure 2. Structures of the di-DNS-amino acids

The DNS-amino acid can be extracted into a water-immiscible organic solvent such as ethyl acetate (EtOAc), leaving the polar free amino acids behind in the aqueous layer. By using normal-phase TLC and comparing the retention factor ( $R_f$ ) of the *N*-terminal DNS-amino acid with those of standard samples, the *N*-terminal residue can be identified.

The second goal of the experiment is to identify the remaining three amino acid residues. You will not determine the correct sequence of these amino acids at this point, as this is quite time consuming to do by chemical means. The remaining three amino acids will be identified by hydrolyzing a sample of the tetrapeptide into its individual amino acids and converting each to its corresponding DNS-derivative. The resulting four DNS-amino acids will be purified by normal-phase TLC. Then, by comparing  $R_f$ s with those of standard DNS-amino acids, the three unknown amino acids will be identified. This process is illustrated in Figure 3.

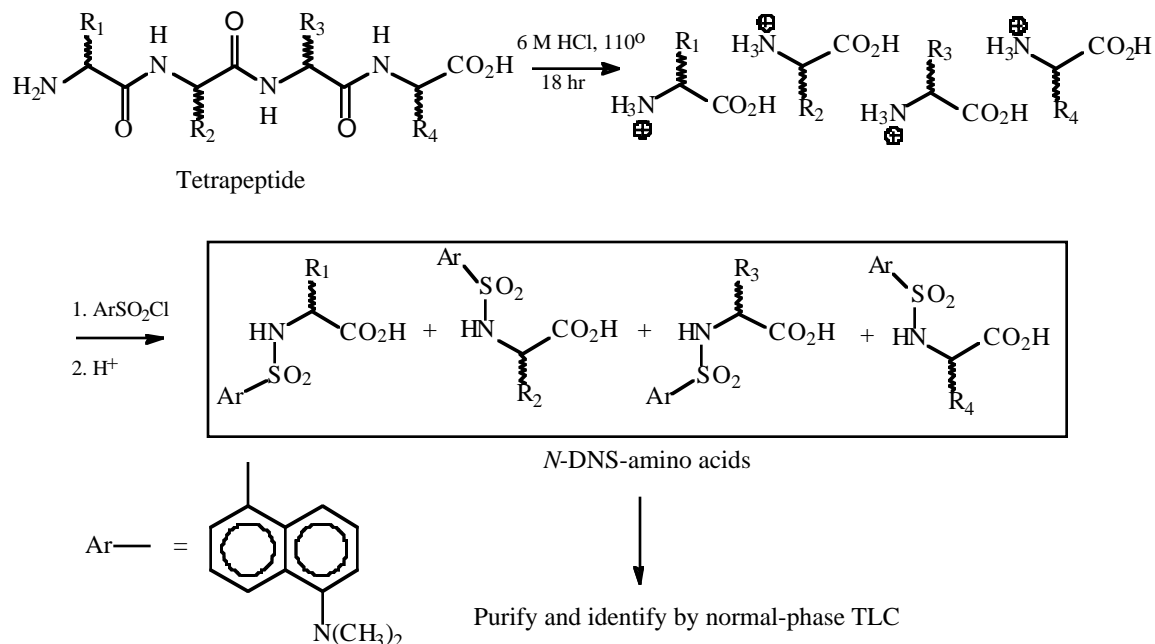
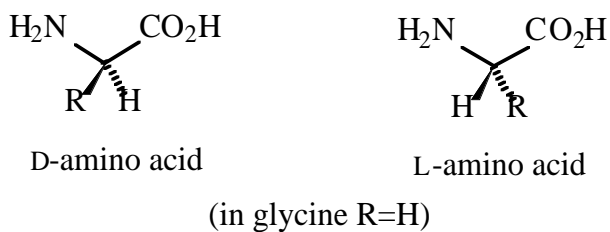


Figure 3. Formation of DNS-Amino Acids

Once you know the identities of the four amino acids in the tetrapeptide, you will then, as the third goal, proceed to identify them as either D- or L- in stereochemistry. Examination of the structure of an  $\alpha$ -amino acid shows that a chiral center exists. Because of this fact, there are two possible enantiomers for each common protein amino acid. An exception is glycine, which has no chiral center and, therefore, no enantiomer. Threonine and isoleucine, each of which has two chiral centers, can exist as four possible stereoisomers. Mirror image stereoisomers are designated D- or L-, as shown below.



The amino acids found in proteins are of the L-configuration. However, increasing numbers of small, biologically active peptides are being discovered which have D-amino acids in addition to the normal L-amino acids. For example, the peptide antibiotic actinomycin D contains a D-valine unit in addition to L-threonine, L-proline and L-N

methylvaline (3). Therefore, you cannot assume that the amino acids in the tetrapeptide are all of the L-configuration.

Recall that enantiomers have identical physical properties with two exceptions. They rotate the plane of polarized light in equal but opposite directions, and they react at different rates with chiral reagents. Therefore, they cannot be separated by conventional chromatography. In order to separate enantiomers, some sort of diastereomeric relationship must be introduced using additional chirality. This can be accomplished in a number of ways. We will use the chiral molecule beta-cyclodextrin ( $\beta$ -CD) to introduce this extra chirality. This molecule consists of seven  $\alpha$ -D-glucose units in a cyclic arrangement as shown in Figure 4.

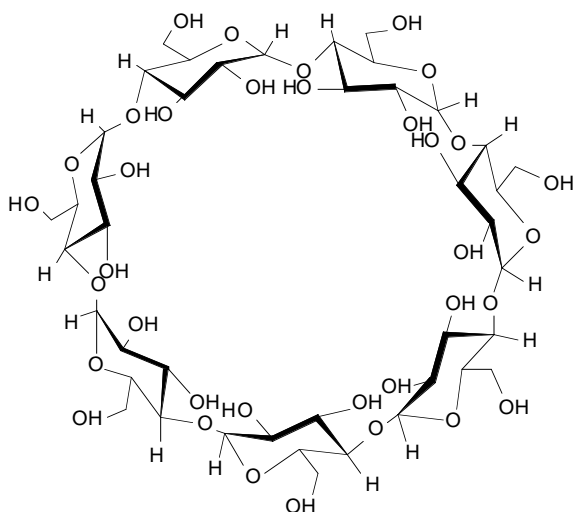


Figure 4. Structure of  $\beta$ -cyclodextrin

The  $\beta$ -cyclodextrin forms diastereomeric inclusion complexes with the DNS-amino acids. These complexes permit selective hydrogen-bonding interactions to occur between DNS- D- and L-amino acids, resulting in the separation of the enantiomers. All common protein amino acids except proline and tryptophan have been completely resolved by using  $\beta$ -CD in the mobile phase of reversed-phase TLC systems (4,5).

By comparing the  $R_f$ s of the DNS-amino acids in the tetrapeptide with DL-standards and by knowing whether the DNS- D- or L-amino acid has the higher  $R_f$ , you can identify the amino acids in the sample as D- or L- in configuration. For example, assume you have identified one of your amino acids by normal-phase TLC (via its DNS-derivative). To determine whether it is of the D- or L-configuration, spot your sample on a reversed-phase TLC plate next to a standard sample of the same DNS-amino acid which is a mixture of both the D- and L-isomers. After development in a solvent system containing  $\beta$ -CD, compare the  $R_f$  of your sample with those of the standard.

By knowing which spot is D- and which is L- in the reference sample (see references 4 and 5 to get this information on all of the DNS-DL-amino acid pairs), you can immediately determine the stereochemistry of your DNS-amino acid sample. This is illustrated in Figure 5 below.

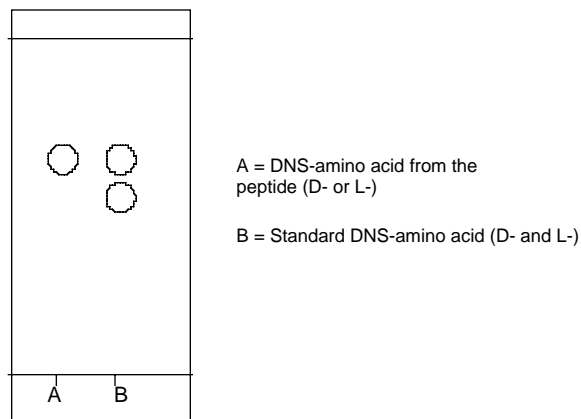


Figure 5. Reversed-Phase TLC Plate Showing the Separation of a DNS-DL-Amino Acid

The fourth goal in the experiment involves determination of the sequence of the amino acids in the tetrapeptide by mass spectrometry. A handout that discusses this technique will be made available to you later. The schedule of the lab work appears below. You will work in pairs.

### Week 1

1. Begin hydrolysis of the tetrapeptide the afternoon before the first lab period (student 2, Section B.1.)
2. Synthesize the four DNS-amino acids from the hydrolyzed peptide (student 2, Section B.2.)
3. Synthesize the DNS-tetrapeptide (student 1, Section A.1.)

### Week 2

1. Begin hydrolysis of the DNS-tetrapeptide the afternoon before the second lab period (student 1, Section A.2.)
2. Identify the *N*-terminal DNS-amino acid (student 1, Section A.3)
3. Purify the four DNS-amino acids by preparative TLC (student 2, Section B.3.)

### Week 3

1. Identify the remaining three DNS-amino acids (students 1 and 2, Section B.4)

#### Week 4

1. Determine the stereochemistry of each DNS-amino acid (students 1 and 2, Section C)
2. Receive the mass spectrum for determination of the amino acid sequence (students 1 and 2)

### III. EXPERIMENTAL

**Caution: Wear departmentally approved safety goggles at all times in the laboratory. Many chemicals are potentially harmful. Keep chemicals away from flames or other heat sources. Prevent eye, skin, and clothing contact. Avoid inhaling the vapors and ingesting the reagents. If you spill any reagent, notify your lab instructor immediately. Hazards for specific chemicals are listed below.**

1. Acetic Acid: Corrosive. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
2. Acetone: Flammable and irritant. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
3. Acetonitrile: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
4. Beta-cyclodextrin hydrate: Avoid inhaling dust and ingesting the compound.
5. Dansyl chloride: Corrosive. Moisture sensitive. Irritant. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
6. Dichloromethane: Toxic and irritant. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
7. Ethyl acetate: Flammable and irritant. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
8. Hydrochloric acid: Highly toxic and corrosive. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
9. Methanol: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.

10. Pyridine: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
11. Sodium chloride: Irritant and hygroscopic. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
12. Sodium hydrogen carbonate: Moisture sensitive. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
13. Sodium hydroxide: Corrosive and toxic. Prevent eye, skin and clothing contact. Avoid ingesting the compound.
14. Sodium sulfate: Irritant and hygroscopic. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.
15. Silica gel: Hygroscopic. Avoid inhaling dust or ingesting the compound.
16. Toluene: Flammable and toxic. Fire hazard. Prevent eye, skin and clothing contact. Avoid inhaling the vapors and ingesting the compound.
17. Urea: Irritant. Prevent eye, skin and clothing contact. Avoid inhaling dust and ingesting the compound.

A. *N*-Terminal Amino Acid Analysis

1. Formation of the DNS-Tetrapeptide (Student 1, week 1)

A 1 millimolar (mM) solution of the tetrapeptide in H<sub>2</sub>O will be provided for you. Take 100 microliters (μL) of this solution and add to it 100 μL of 0.2 M NaHCO<sub>3</sub> (pH = 9.5) in a 1 dram vial. Use an 10-100 μL Eppendorf pipet for all transfers of solvents. Add 200 μL of 10 mM DNS-Cl in acetone, cap tightly, and shake for 1 minute. Let the vial sit for 1.5 hours at room temperature. The yellow color of the DNS-Cl will fade over time. When the reaction is complete, adjust the pH of the solution to between 2 and 4 using 50 μL of 1 M HCl. Use a micropipet and pH paper to measure the pH. If the pH is still above 5 at this point, add more 1 M HCl in 10 μL portions until the pH is between 2 and 4. Label the vial with your name, and place the uncapped vial in a 10 mL beaker in the hood until the next lab period to allow the solvent to evaporate.

2. Hydrolysis of the DNS-Tetrapeptide (Begin the day before week 2; student 1)

At a later time (to be determined by your instructor) add 500  $\mu\text{L}$  of 6 M HCl to the DNS-tetrapeptide using a 100-1000  $\mu\text{L}$  Eppendorf pipet and tightly cap the vial. Label it with your name, put it in a 10 mL beaker, and place it in an oven at 105-110°C for 18 hours.

After hydrolysis of the DNS-tetrapeptide, carefully remove the cap and let the sample cool to room temperature. In a fume hood, transfer the sample to a 6 cm diameter watchglass using a Pasteur pipet. Use three small portions of water (2-3 drops each) to rinse the 1 dram vial to ensure complete transfer of the sample to the watchglass. Place the watchglass 2-3 inches from an infrared heat lamp until all of the aqueous HCl has evaporated.

**Caution: The heat lamp and watchglass will get very hot.**

3. Identification of the *N*-Terminal DNS-Amino Acid (Student 1, week 2)

Let the crude residue cool to room temperature. Dissolve it in 5 drops of wet EtOAc using a Pasteur pipet, and mix the solution well with a clean, dry microspatula. The wet EtOAc is prepared by shaking together 1 mL of  $\text{H}_2\text{O}$  and 2 mL of EtOAc. Use the **top layer** for dissolving your sample. *Make sure that you do not get traces of the bottom water layer in your pipet.* Using a ruler and pencil (not a pen!), lightly draw a line 7-8 mm from the bottom of a 3.5 x 7.5 mm plastic-backed normal-phase TLC plate. Spot your sample on the pencil line at the center of the plate. If the wet EtOAc evaporates from the watchglass simply add a few drops more as needed to keep the compound in solution. At 5 mm intervals on either side of your sample, place five additional spots corresponding to five known DNS-amino acid standards. Since there are 15 different standards, you will have to run three plates to complete the analysis. These plates can be run simultaneously to save time. Remember to spot your sample on each plate.

**Caution: Ultraviolet (UV) radiation can cause severe eye damage. Wear goggles. Do not look directly into the UV lamp.**

Before developing your plates, view them under a long-wave UV lamp to make sure that there is a fluorescent spot visible for each sample at the origin. If there is not, spot more sample. Pre-develop the

plates in methanol (MeOH) for a second or two, removing the plate from the MeOH when it reaches the origin. This polar solvent will concentrate the spots and result in better separations. Dry the plate for 1 minute in an oven at 110°C to evaporate the MeOH before developing it in the next solvent.

**Caution: Pyridine has an obnoxious odor. Always develop and view the TLC plates in a fume hood.**

Develop the plates in 5 mL of solvent composed of toluene:pyridine:acetic acid (AcOH)/30:10:1 (by volume) in a 4 oz screw cap jar. Use a Pasteur pipet to dispense the solvent **in the hood**. After development, immediately mark the solvent front with a pencil. Leave all of your plates in a safe place in the hood. View the plates under long-wave UV light (366 nm) in the hood. Identify the *N*-terminal DNS amino acid by comparing its migration distance with those of the standards. Ignore any blue fluorescent spot at the origin in your sample. If it is visible it is due to dansyl sulfonic acid, a reaction byproduct. Circle the spots carefully with a pencil. Calculate all  $R_f$ s. Note that you may have to use the double-spotting technique and/or a different solvent system in order to definitively identify the *N*-terminal DNS-amino acid. These things are described in Section B.4.

The standard DNS-amino acids are:

alanine	leucine	serine
aspartic acid	lysine ( <i>N,N'</i> -di-DNS)	threonine
arginine	methionine	tryptophan
glycine	phenylalanine	tyrosine ( <i>N,O</i> -di-DNS)
glutamic acid	proline	valine

Notice that asparagine and glutamine have not been included in the list. This is because these amino acids would be hydrolyzed to aspartic acid and glutamic acid, respectively, by 6 M HCl.

**Clean-up: Dispose of the used TLC solvent in a container labeled “Used normal-phase TLC solvent” in a fume hood.**

#### B. Identification of the Remaining Three Amino Acids

1. Hydrolysis of the Tetrapeptide (Begin the day before week 1; student 2)

Using an Eppendorf pipet, transfer 200  $\mu\text{L}$  of a 1 mM solution of the tetrapeptide in  $\text{H}_2\text{O}$  to a clean 1 dram vial. Add 200  $\mu\text{L}$  of concentrated HCl, cap tightly, and place it in the oven at 105–110°C for 18 hours.

After the hydrolysis is complete, let the sample cool to room temperature. Transfer it to a clean watchglass using a Pasteur pipet. Rinse the vial three times with a small amount of water (2-3 drops each time) to ensure complete transfer of the sample. In a fume hood, place the watchglass 2-3 inches from an infrared heat lamp until the aqueous HCl has evaporated.

**Caution: The heat lamp and watchglass will be very hot.**

2. Formation of the DNS-Amino Acids (Student 2, week 1)

Let the residue cool to room temperature. Add 500  $\mu\text{L}$  of 0.2 M  $\text{NaHCO}_3$  using an Eppendorf pipet. Transfer the solution back to the 1 dram vial using a Pasteur pipet. Rinse the watchglass with another 500  $\mu\text{L}$  portion of 0.2 M  $\text{NaHCO}_3$  and add it to the same vial. Using a micropipet for spotting TLC plates and pH paper, make sure that the pH of this solution is 9 - 10. **This is critical to the success of the dansylation reaction!** Notify your instructor if the pH is less than 9. Add 1000  $\mu\text{L}$  of 10 mM DNS-Cl in acetone. Cap the vial, shake it for 1 minute, and let it sit in a 10 mL beaker for 1.5 hours at room temperature.

Adjust the pH to between 2 and 4 by adding 400  $\mu\text{L}$  of 1 M HCl. If the pH is above 5, continue to add 1 M HCl in 25  $\mu\text{L}$  portions until the pH is between 2 and 4. Transfer the solution to a 5 mL conical vial using a Pasteur pipet. Add 1 mL of saturated NaCl solution. Extract the aqueous solution three times with EtOAc (1.0 mL each). Save the aqueous layer in case more extractions are necessary. Combine the EtOAc layers and dry with 100 mg of anhydrous  $\text{Na}_2\text{SO}_4$ . Transfer the dried EtOAc in several small portions to a clean, dry 10 mL round bottom flask via a Pasteur filter pipet. Use three small portions of fresh EtOAc (4-5 drops each) to rinse the vial and ensure complete transfer of the sample to the round bottom flask. Evaporate the EtOAc to dryness using a rotary evaporator. Stopper and label the flask and turn it in to your instructor so that it can be refrigerated until the next lab period.

3. Purification of the DNS-Amino Acids (Student 2, week 2)

Dissolve the crude mixture of the four DNS-amino acids synthesized above in part B.2 in 50  $\mu\text{L}$  of wet EtOAc using an Eppendorf pipet. Obtain a 5 x 20 cm plastic-backed, normal-phase TLC plate (0.1 mm thickness). Using a ruler and pencil (not a pen!), very lightly draw a pencil line 1 cm from the bottom of the plate. *Be careful not to disrupt the silica gel.* Spot the crude *N*-DNS-amino acid 5 mm from the left-hand edge of the plate. This will be used as a reference so that you will know which of the four amino acids is the *N*-terminal one. Spot the crude DNS-amino acid mixture on this line across the plate using a micropipet, beginning 5 mm from the *N*-DNS-amino acid spot and ending 5 mm from the right-hand edge. Make the spots as small as possible (ca. 1-2 mm in diameter) and as close to each other as possible. It will probably be necessary to make more than one pass across the plate to deliver all of the sample. **Save the remaining sample in the 10-mL round bottom flask.** Pre-develop the plate in MeOH for a few seconds up to the origin. Dry in the oven for 1 minute at 110°C.

**Caution: Pyridine has an obnoxious odor. Always develop and view the TLC plates in a fume hood.**

Develop the plate in 10 mL of the toluene:pyridine:AcOH/30:10:1 (by volume) solvent system. The plate takes approximately 2 hours to fully develop.

While this plate is developing, add 4-5 drops of wet EtOAc to dissolve what remains of the crude DNS-amino acid mixture in the 10-mL round bottom flask. If the solvent evaporates, add more. Spot the sample on 3.5 x 7.5 mm plastic-backed normal-phase TLC plates along with the standard DNS-amino acids as in Section A.3. Pre-develop the plates in MeOH, oven dry them, and then develop them in the same solvent system used above in a fume hood.

**Caution: Ultraviolet (UV) radiation can cause severe eye damage. Wear goggles. Do not look directly into the UV lamp.**

View the plates in a under a long-wave UV lamp. By comparing migration distances, four or five of the possibilities can be eliminated immediately. Unambiguous identifications will be made using the purified DNS-amino acids from the peptide as described in the next section.

After development of the 5x20 cm plate, remove it from the tank and immediately mark the solvent front before it evaporates. Visualize the bands under the UV-light and calculate all  $R_f$ s. Carefully cut out the 4 major UV-absorbing bands *with the lowest  $R_f$ s* that are greater than zero. Label the vials A-D, with A corresponding to the DNS-amino acid with the lowest  $R_f$  and D the highest. Ignore any other bands -- these represent reaction by-products. Place the four strips in *separate* labeled 1 dram vials. Add 1 mL of MeOH to each of the four vials, cap them, and shake them for a couple of minutes. The MeOH will elute the purified DNS-amino acids from the silica gel. When all of the fluorescence has been removed from the silica gel, remove the strips with tweezers and discard them. Place the **uncapped** vials in a small beaker labeled with your name and put them in a fume hood for a day or two to evaporate the MeOH. Cap the vials and place them in a refrigerator until the next laboratory period.

4. Normal-phase TLC Identification of the Purified DNS-Amino Acids (Students 1 and 2, week 3)

From your partner's work on the *N*-terminal residue in Section A, you may already know the identity of one of the four DNS-amino acids. If so, the *N*-terminal DNS-amino that was just purified can be set aside until later use in determining its stereochemistry. Analyze the remaining three DNS-amino acids one at a time. Dissolve each in 100  $\mu$ L of wet EtOAc and proceed as described in Section A.3. You should be able to eliminate most of the possible DNS-amino acids quickly. If you are unable to distinguish between two choices (that is, if the  $R_f$  of your compound is very close to two other compounds), try the following double spotting technique. Spot your sample twice on the plate (ca. 5 mm apart) and then spot one of the other choices directly on top of one of the spots of your compound. Repeat with another set of spots for the other compound in question. Develop the plate. If the compounds are the same, only one spot will be seen where you double spotted. If the compounds are different, you should be able to distinguish two spots (although the spots may overlap). If this double-spotting technique does not give you a positive identification, try a different solvent system. Another good one to try is  $\text{CH}_2\text{Cl}_2$ :MeOH:EtOAc/12:4:1 (by volume). If these two solvents fail you can try MeOH:2% acetic acid/75:25 (by volume). This solvent is used with a 1 x 3" glass-backed reversed-phase TLC plate, and will not work with normal-phase TLC plates. **Please use these reversed-phase plates sparingly** as they are rather expensive. When you have successfully identified one of the three

DNS-amino acids, go on to the next one until all have been correctly identified. **Check with your instructor to make sure you have identified all four DNS-amino acids correctly before you proceed to the next section.**

C. Determination of the Stereochemistry of the Amino Acids (Students 1 and 2, week 4)

For this part of the experiment, you will use two 1 x 3" glass-backed, reversed-phase plates. For the stereochemical analysis of the *N*-terminal amino acid, use the one that was purified in Section B.3. Dissolve each of the four purified DNS amino acids in 3-4 drops of wet EtOAc. Using a ruler and pencil, very lightly draw a pencil line 7 mm from the bottom of each plate. Starting 5 mm from the left hand edge, spot the sample labeled A and its corresponding standard next to it, 5 mm to the right. **Make the spots as small as possible.** Then spot the sample labeled B and next to it the corresponding standard, again using 5 mm intervals between spots. In a similar manner spot samples C and D next to their corresponding standards on the second plate. Predevelop the plates in MeOH for a second or two up to the origin. Dry the plates in the oven at 110°C for 1 minute. Let them cool to room temperature, and then develop each in 4 mL of a solvent composed of acetonitrile (CH<sub>3</sub>CN):aqueous 0.2 M β-CD/25:75 (by volume). The aqueous portion is also 0.6 M in NaCl and is saturated with urea. The NaCl prevents the highly aqueous solvent from dissolving the binder which holds the C-18 derivatized silica gel to the glass plate. The urea increases the solubility of the β-CD in water approximately 10-fold. After developing the plates (approximately 1 hr), carefully compare the migration distances of the DNS-D- or L-amino acid spots from the peptide with those of the corresponding DNS-DL-standards under the long-wave UV light. Let the plate dry for 15 minutes before marking the spots lightly with a sharp pencil. See references 4 and 5 to determine which spot corresponds to the D-isomer and which corresponds to the L-isomer.

**Clean-up: Dispose of the used TLC solvent in a container labeled "Used reversed-phase TLC solvent" in a fume hood.**

#### IV. REFERENCES

1. W. R. Gray and D. S. Hartley, *Biochem. J.*, **89** (1963) 59.

2. W. R. Gray, *Methods in Enzymology*, **25B** (1972) 121.
3. R. Walter and J. Meienhofer (Eds.) Peptides. *Chemistry, Structures and Biology*, Ann Arbor Science Publishers, Ann Arbor, 1975.
4. D. W. Armstrong, F.-Y. He and S. M. Han, *J. Chromatogr.*, **448** (1988) 345.
5. J. W. LeFevre, *J. Chromatogr. A*, **653** (1993) 293.

Thanks to David Clement for generating Figure 4.

## V. QUESTIONS

1. In the formation of the DNS-amino acids, why must the pH of the solution be between pH 8.5 and 10.5?  
(See reference 2 for help.)
2. Calculate the amount of tetrapeptide in moles and micrograms that was used in sections A.1 and B.1 (MW = 406).
3. The pentapeptide Tyr-Asp-Ser-Val-Glu was subjected to the reaction sequence shown below. Draw the structures of the products that result.
  - A. DNS-Cl, NaHCO<sub>3</sub>
  - B. H<sup>+</sup> (to pH=4)
  - C. 6 M HCl, 110°C, 18 hr.
4. Based upon your normal phase TLC results, list the following DNS-amino acids in order of increasing polarity (least polar to most polar). Briefly explain your answer.
  - A. Asp
  - B. Val
  - C. Thr
  - D. Gly

# STUDENT HANDOUT #2

## Peptide Sequencing by Tandem Quadrupole Mass Spectrometry

**David W. Dodsworth\***

Bristol-Myers Squibb, P.O. Box 4755, Syracuse, NY 13221-4755; \**dodsworth@bms.com*

Classical peptide sequencing is usually performed by the Edman degradation. This method is performed by attaching a derivatizing agent on the *N*-terminus of the peptide, and then hydrolyzing the adjacent amide bond, thus releasing the derivatized *N*-terminal amino acid as a phenylthiohydantoin (PTH) derivative. This process generates a new *N*-terminus on the peptide. The PTH amino acid is then identified by a chromatographic method such as TLC. Because only the *N*-terminal amino acid is removed, the process can be serially repeated (automated) until the entire peptide is sequenced. The practical limit of this method is 25-30 amino acid residues due to incomplete derivatization and/or incomplete hydrolysis of the derivatized amino acid. Also, if the *N*-terminus is blocked due to a modification of the protein/peptide, as in a cyclic peptide, this method will not work at all. A third concern related to this method is that it is time consuming, as each automated cycle typically takes two hours. As a result, the analyst can easily take more than a day to sequence an eight residue peptide.

The same information can be obtained in one experiment by tandem quadrupole mass spectrometry (MS/MS). It is an instrumental technique that allows the operator to obtain the mass spectrum of a single component in the presence of others. The MS/MS instrument is, in effect, two quadrupole mass spectrometers connected in series by a third quadrupole collision cell. A sample (the peptide being investigated) is vaporized and ionized in the ion source of the mass spectrometer. Once the sample is ionized, a normal full scan mass spectrum is acquired and the molecular ion is determined. The first quadrupole region is then set to pass only the molecular ion of the material to be analyzed. This ion is then passed into the collision cell where it undergoes multiple collisions with an inert collision gas, such as argon. The collisions cause the molecular ion to randomly fragment at the weaker (more labile) bonds. Under the low collision energies of a quadrupole instrument, fragmentation typically occurs at the amide

bonds in the backbone of a peptide. The resulting fragment ions are passed into the third quadrupole region and a product ion mass spectrum is acquired.

Once the product ion spectrum of the peptide has been obtained, its amino acid sequence may be determined by interpreting the fragment ions. Roepstorff and Fohlman developed a nomenclature system that is now universally used for identifying the peptide fragment ions (*I*). Typically, the two main fragment ions that form in the quadrupole MS/MS experiment arise by random cleavage of a single amide bond in the peptide backbone. The single positive charge will remain on one fragment or the other. This is partially determined by the amino acid composition of the two fragments. If the charge remains on the *N*-terminal containing fragment, it is referred to as a B ion. If charge retention is on the carboxylic acid terminal of the peptide, it is referred to as a Y ion. The way to distinguish between the two ions will be addressed in the example.

The main advantage of sequencing peptides by mass spectrometry is the speed with which it is accomplished. An eight residue peptide can typically be sequenced (within certain constraints) in an hour or less. The two major disadvantages to sequencing by MS/MS are as follows:

1. The quadrupole experiment cannot distinguish between leucine and isoleucine or between glutamine and lysine.
2. One cannot distinguish between D- and L-configurations of the amino acids. We, of course, have already solved this problem using reversed-phase TLC.

Clearly, MS/MS does not replace the Edman degradation method. However, it can be used in conjunction with the Edman method to ease and/or speed the sequencing process.

Another important use of sequencing by MS/MS is in peptide synthesis. In this case the analyst has prior knowledge of the suspected sequence of the peptide and can quickly determine if the peptide was prepared correctly.

The MS/MS spectrum of the tetrapeptide you are investigating is presented in Figure 1. It contains all the ions required to fully sequence the peptide.

#### Unknown Lab Peptide

OSW EGO 2 1 (0.823)

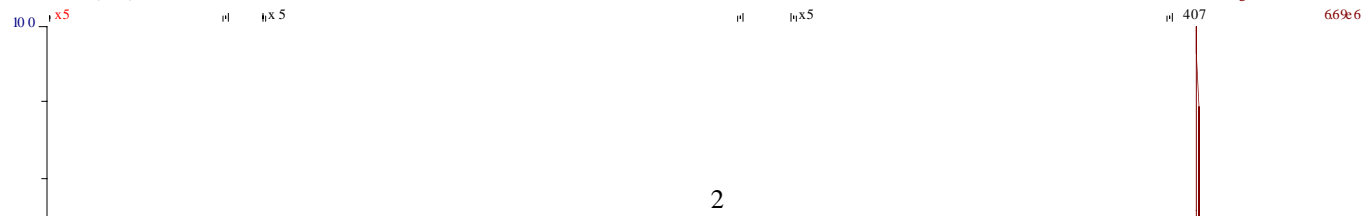


Figure 1. MS/MS spectrum of the tetrapeptide

To provide assistance in the interpretation of the mass spectrum, the following example will be worked for you.

A peptide was submitted for determination of its sequence. The following information was obtained prior to MS/MS analysis:

1. Treatment with dansyl chloride followed by hydrolysis in 6 M HCl and TLC analysis resulted in the identification of dansyl-methionine.
2. Total amino acid analysis yielded two glycine units, three valines, and one methionine (Gly<sub>2</sub>, Val<sub>3</sub>, and Met).

The full scan mass spectrum of the peptide was then acquired on a tandem quadrupole mass spectrometer using standard operating parameters. The peptide's molecular weight was found to be 560.30. A product ion MS/MS spectrum was then acquired using the protonated molecular ion (MH<sup>+</sup>, m/z 561) as the precursor ion. This spectrum is presented in Figure 2. To interpret the MS/MS spectrum, one needs to proceed as described in the paragraphs that follow.

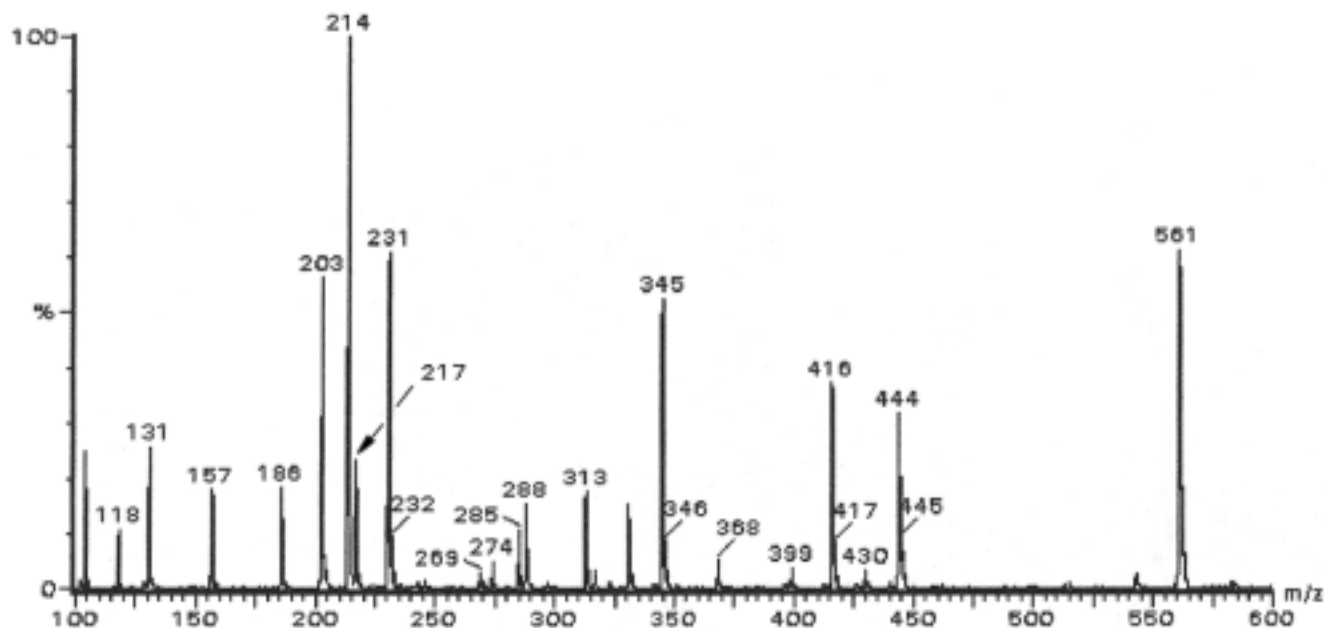


Figure 2. MS/MS Spectrum of the example peptide

First, all mass losses are the residual masses (RMs) of the amino acids. The RM is the molecular weight of the amino acid minus water (18). The names and symbols of the common amino acids along with the RM compositions and values and appear in Table 1 below.

Name	Symbol	Residual Mass and Composition
Alanine	Ala A	71.03711 (C <sub>3</sub> H <sub>5</sub> NO)
Arginine	Arg R	156.10111 (C <sub>6</sub> H <sub>12</sub> N <sub>4</sub> O)
Asparagine	Asn N	114.04293 (C <sub>4</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub> )
Aspartic acid	Asp D	115.02694 (C <sub>4</sub> H <sub>5</sub> NO <sub>3</sub> )
Cysteine	Cys C	103.00919 (C <sub>3</sub> H <sub>5</sub> NOS)
Glutamic Acid	Glu E	129.04259 (C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub> )
Glutamine	Gln Q	128.05858 (C <sub>5</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub> )
Glycine	Gly G	57.02146 (C <sub>2</sub> H <sub>3</sub> NO)
Histidine	His H	137.05891 (C <sub>6</sub> H <sub>7</sub> N <sub>3</sub> O)
Isoleucine	Ile I	113.08406 (C <sub>6</sub> H <sub>11</sub> NO)

Leucine	Leu L	113.08406 (C <sub>6</sub> H <sub>11</sub> NO)
Lysine	Lys K	128.09496 (C <sub>6</sub> H <sub>12</sub> N <sub>2</sub> O)
Methionine	Met M	131.04049 (C <sub>5</sub> H <sub>9</sub> NOS)
Phenylalanine	Phe F	147.06841 (C <sub>9</sub> H <sub>9</sub> NO)
Proline	Pro P	97.05276 (C <sub>5</sub> H <sub>7</sub> NO)
Serine	Ser S	87.03203 (C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> )
Threonine	Thr T	101.04768 (C <sub>4</sub> H <sub>7</sub> NO <sub>2</sub> )
Tryptophan	Trp W	186.07931 (C <sub>11</sub> H <sub>10</sub> N <sub>2</sub> O)
Tyrosine	Tyr Y	163.06333 (C <sub>9</sub> H <sub>9</sub> NO <sub>2</sub> )
Valine	Val V	99.06841 (C <sub>5</sub> H <sub>9</sub> NO)

Table 1. Residual masses and compositions of the commonly occurring amino acid residues.

Second, always start the interpretation by looking for the loss of the *C*-terminal amino acid. If it is present in the spectrum, it will give you the B ion series. This is done by subtracting 18 from the protonated molecular ion ( $m/z$  561-18 = 543 ion), and then subtracting the residual masses of the amino acids, one at a time, until a mass agrees with an observed fragment ion. **Be Careful!** Start with the smallest amino acid (Gly, RM = 57) and work to the largest or the wrong *C*-terminal amino acid may be chosen. The ion at  $m/z$  444 confirms that Val (RM = 99) is the *C*-terminus, and it is the B<sub>5</sub> ion. There are six amino acids in this peptide and the ion at  $m/z$  444 arises from the loss of Val. Therefore, this ion is due to the remaining five amino acids, hence B<sub>5</sub> designation.

Third, now that the *C*-terminal ion is identified, look for the *N*-terminal ion. If it is present in the spectrum, it will give you the Y ion series. This is done by subtracting the RM of each amino acid until a match is found. **Caution!** Frequently, only one series is present and it could be either the Y or B series. From the preliminary data, Met is the *N*-terminal amino acid. The ion at  $m/z$  430 is due to the loss of Met (RM = 131), and it is the Y<sub>5</sub> ion.

Fourth, work each series down to the low masses using the amino acid composition data for assistance. There is no need to look for the RM loss of an amino acid when it is not present in the sample. For instance,  $m/z$  444 to  $m/z$  345 is a change in mass of 99, which is the RM of Val. The

known sequence from the C-terminus is now Val-Val. Continue until the entire peptide is correctly sequenced. The data needed to sequence the peptide is found in Table 2.

High Mass	Low Mass	$\Delta$ Mass	Ion Type	Amino Acid Lost	Sequence
543	444	99	B <sub>5</sub>	Val (V)	V
444	345	99	B <sub>4</sub>	Val (V)	VV
345	288	57	B <sub>3</sub>	Gly (G)	GVV
288	231	57	B <sub>2</sub>	Gly (G)	GGVV
			B <sub>1</sub> *		
561	430	131	Y <sub>5</sub>	Met (M)	M
430	331	99	Y <sub>4</sub>	Val (V)	MV
331	274	57	Y <sub>3</sub>	Gly (G)	MVG
274	217	57	Y <sub>2</sub>	Gly (G)	MVGG
217	118	99	Y <sub>1</sub>	Val (V)	MVGGV
118				Val (V)	MVGGVV

\* Note that the B<sub>1</sub> ion at 132 [which arises from B<sub>2</sub> at 231-99 (Val)] is not present.

Table 2. Molecular ion data.

The sequences being built in Table 2 are from opposite ends of the peptide. When the data is put together the resulting sequence is:

(N-terminal) Met-Val-Gly-Gly-Val-Val (C-terminal)

Apply the same methodology to determine the amino acid sequence of your tetrapeptide.

(Remember that you already know the N-terminal residue and the amino acid content of the rest of the peptide from your earlier work). Place your data in a table similar to Table 2.

#### Reference:

1. Roepstorff, P.; Folhman, J.; *Biomed. Mass Spectrom.*, 1984, **11**, 601