ACYLATION OF MEMBRANE PROTEINS TO ENHANCE THEIR WATER SOLUBILITY

By
ROBERT CHRISTOPHER MORTON, B.Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
1991
WATER SOLUBILIZATION OF MEMBRANE PROTEINS
DOCTOR OF PHILOSOPHY (1991)  McMaster UNIVERSITY  
(Biochemistry)  Hamilton, Ontario

TITLE:  Acylation of Membrane Proteins to Enhance Their Water Solubility

AUTHOR:  Robert Christopher Morton, B.Sc.  (McMaster University)

SUPERVISOR:  Professor G.E. Gerber

NUMBER OF PAGES:  xvi, 220
ABSTRACT

Membrane proteins are an integral part of the structure of membranes providing the functional aspect of the membrane with respect to the transport of water soluble molecules, signal recognition and reaction catalysis. Primary, secondary and tertiary structure of these types of proteins warrant investigation in order to better understand their individual function. Both DNA sequencing and standard protein sequencing can provide information on the protein's primary sequence. Protein sequencing can identify sites of post-translational peptide bond hydrolysis and translocation and aid in the elucidation of the secondary and tertiary structure. However, the hydrophobic nature of these membrane proteins necessitates the use of detergents or chaotropic agents to achieve the solubilization of the protein and even its peptide fragments under aqueous conditions; in the absence and sometimes even in the presence of these agents, membrane proteins aggregate and precipitate out of solution. Standard protein sequencing techniques are, therefore, hampered by this protein aggregation and insolvability, resulting in incomplete enzyme and chemical cleavage of the protein, general losses of both the protein and the peptide fragments and the overall poor isolation of hydrophobic peptide fragments from one another.

The focus of this research was on the development of a method by which membrane proteins could be chemically derivatized resulting in the formation of a stable product that was soluble in aqueous solutions, in the absence of detergents and chaotropic agents. The methodology consists of the following: 1. Solubilization of the membrane protein in anhydrous pyridine
through the utilization of tetrabutylammonium ions which provide a hydrophobic counter ion to the carboxylate groups on the protein. 2. Complete derivatization of the amino acid side chain amine and hydroxyl groups with di-trimethylsilylethyl trimesic anhydride, in the presence of the acylation catalyst, 4-pyrrolidinopyridine. 3. Regeneration of the carboxyl groups on the trimesic acid moiety through the selective removal of the trimethylsilylethyl ester groups by tetrabutylammonium fluoride. 4. Isolation of the trimesylated protein by column chromatography or dialysis. Apobacteriorhodopsin was used as the model membrane protein and after this procedure, the trimesylated protein was found to be completely soluble in simple aqueous solutions. This water soluble derivative was shown to chromatograph on Sephadex type columns that were eluted with buffers containing low salt concentrations. It was further shown to be a suitable substrate for the enzyme pyroglutamate aminopeptidase resulting in the quantitative removal of the N-terminal pyroglutamate residue; the native protein is a poor substrate due to its insolubility, under these aqueous conditions.

This method of trimesylation is applicable to all proteins, both membranous and water soluble, generating a trimesylated protein derivative that is completely water soluble. This chemical modification incorporates two carboxyl groups at each site of modification, Lys, Tyr, Ser and Thr side chains; due to the frequency of the hydroxyl amino acid residues even in hydrophobic domains, this modification would also enable peptide fragments to be water soluble.
ACKNOWLEDGEMENTS

I would like to thank Dr. Gerber for his supervision and discussions that enabled this research to materialize from a conceptual stage to a workable methodology; the training in problem solving was invaluable. I would also thank my family and friends, especially P. Leblanc and D. Mangroo, for their support and encouragement. A special thanks goes to my wife, Marion, for her love and understanding.
# TABLE OF CONTENTS

Page

- i  Title Page
- ii Descriptive Note
- iii Abstract
- v Acknowledgements
- vi Table of Contents
- x Table of Abbreviations
- xiii List of Figures and Tables

1  I. Introduction:

3  1. Amino Acid Sequencing of Membrane Proteins

4    A. Enzymatic Protein Fragmentation

4    B. Organic Solvent Solubilization and Chemical Fragmentation

5    C. Protein Fragment Isolation

7    D. Automated Protein Sequencing

8    E. Alternative Sequencing Methods

9  2. Terminally Blocked Proteins

13  3. Deduced Amino Acid Sequence from DNA

14  4. Topological Models of Membrane Proteins

21  5. Putative Transmembrane Segments

25  6. Chemical Derivatization of Membrane Proteins

30  7. Chemical Reagent Requirements

34  8. Model Systems
A. Model Peptides

B. Model Protein

9. Rationale and Objective

II. Materials and Methods:

1. Chemicals and Solvents

2. High Pressure Liquid Chromatography (HPLC)

3. Alkaline Hydroxylamine Ester Assay

4. Tritium Exchange on Trimesic Acid

5. Synthesis of DT-TM Acid

6. Formation of DT-TM Anhydride

7. DNP-Proline Synthesis

8. DNP-Peptide Synthesis

    A. DNP-prolylserinamide and DNP-prolylthreoninamide

    B. DNP-prolylglycinamide and DNP-prolylglycyltyrosinamide

9. Formation of Tetrabutylammonium Salt

10. Standard Trimesylation

11. Preparation of Apobacteriorhodopsin

12. Solubilization of Apobacteriorhodopsin in Anhydrous Pyridine

13. Polyacrylamide Gel Electrophoresis

    A. Laemmeli

    B. Acidic Lithium Dodecylsulphate

14. Amino Acid Analysis

    A. Laboratory Procedure

    B. Protein Services Procedure

15. Protein Sequencing

III. Results and Discussion:
1. Apobacteriorhodopsin Solubilization

2. Succinylation
   A. Reaction Characterization
   B. Characterization of Succinyl-Apobacteriorhodopsin
   C. Succinyl Ester Stability
   D. Conclusion

3. Model System; DNP-Peptides

4. Trimesylation
   A. DT-TM Imidazolide
      i. DT-TM Acid Formation
      ii. DT-TM Imidazolide Stability
      iii. Characterization of the Trimesylation Reaction
      iv. Stability of the Trimesyl Ester
      v. Apobacteriorhodopsin Trimesylation
      vi. Trimesyl-Apobacteriorhodopsin Characterization
      vii. Conclusion
   B. DT-TM Anhydride
      i. DT-TM Acid Anhydride Formation
      ii. DT-TM Anhydride Reactivity and Stability
      iii. Characterization of the Trimesylation Reaction
      iv. Apobacteriorhodopsin Trimesylation and Isolation
      v. Characterization of Trimesyl-Apobacteriorhodopsin
         a. Column Chromatography
         b. Enzyme Digestion
         c. Amino Acid Sequencing
      vi. Conclusion
      vii. Future Investigations
Complete nucleotide and derived amino acid sequence of cDNA encoding the mitochondrial uncoupling protein of rat brown adipose tissue: lack of a mitochondrial targeting presequence

Cloning of cDNA encoding the uncoupling protein of brown adipose tissue mitochondria: lack of a mitochondrial targeting presequence

Water solubilization of membrane proteins. Extensive derivatization with a novel polar derivatizing reagent

Amino acid analysis by dinitrophenylation and reverse phase high pressure liquid chromatography.

A novel method of complete activation by carbonyldiimidazole: application to ester synthesis.
# Table of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, Ala</td>
<td>alanine</td>
</tr>
<tr>
<td>C, Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>D, Asp</td>
<td>aspartic acid</td>
</tr>
<tr>
<td>E, Glu</td>
<td>glutamic acid</td>
</tr>
<tr>
<td>F, Phe</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>G, Gly</td>
<td>glycine</td>
</tr>
<tr>
<td>H, His</td>
<td>histidine</td>
</tr>
<tr>
<td>I, Ile</td>
<td>isoleucine</td>
</tr>
<tr>
<td>K, Lys</td>
<td>lysine</td>
</tr>
<tr>
<td>L, Leu</td>
<td>leucine</td>
</tr>
<tr>
<td>M, Met</td>
<td>methionine</td>
</tr>
<tr>
<td>N, Asn</td>
<td>asparagine</td>
</tr>
<tr>
<td>P, Pro</td>
<td>proline</td>
</tr>
<tr>
<td>Q, Gln</td>
<td>glutamine</td>
</tr>
<tr>
<td>R, Arg</td>
<td>arginine</td>
</tr>
<tr>
<td>S, Ser</td>
<td>serine</td>
</tr>
<tr>
<td>T, Thr</td>
<td>threonine</td>
</tr>
<tr>
<td>V, Val</td>
<td>valine</td>
</tr>
<tr>
<td>W, Trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Y, Tyr</td>
<td>tyrosine</td>
</tr>
<tr>
<td>apoBR</td>
<td>apobacteriorhodopsin</td>
</tr>
<tr>
<td>BR</td>
<td>bacteriorhodopsin</td>
</tr>
</tbody>
</table>
Bu$_4$N$^+$  tetrabutylammonium
CCl$_4$  carbon tetrachloride
CDI  carbonyldiimidazole
CHCl$_3$  chloroform
CH$_3$CN  acetonitrile
DCC  dicyclohexylcarbodiimide
DMAP  4-dimethylaminopyridine
DMF  dimethylformamide
DNP  dinitrophenyl
DT-TM  di-trimethylsilyl ethyl trimesyl
EDTA  ethylenediaminetetraacetic acid
Et$_3$N; Et$_3$NH$^+$  triethylamine; triethylammonium
EtOH  ethanol
FDNB  1-fluoro-2,4-dinitrobenzene
HFBA  heptafluorobutyric acid
Im  imidazolyl
LDS  lithium dodecylsulphate
MeOH  methanol
Me$_3$SiCH$_2$CH$_2$OH  2-(trimethylsilyl)ethanol
Me$_2$SO  dimethylsulfoxide
β-MSH  β-mercaptoethanol
MT-TM  mono-trimethylsilyl ethyl trimesyl
ODS  octadecylsilica
PAGE  polyacrylamide gel electrophoresis
PGAPase  pyroglutamate aminopeptidase
PITC  phenylisothiocyanate
PPY  4-pyrrolidinopyridine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyr</td>
<td>pyridine</td>
</tr>
<tr>
<td>Quadrol</td>
<td>N,N,N,N-tetrakis(2 hydroxypropyl)ethylenediamine-</td>
</tr>
<tr>
<td></td>
<td>TFA buffer in n-propanol/ water (3:4 v/v), pH 9.0</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecylsulphate</td>
</tr>
<tr>
<td>TFA; TFA⁻</td>
<td>trifluoroacetic acid; trifluoroacetate</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TM</td>
<td>trimesyl</td>
</tr>
<tr>
<td>Trimesic acid</td>
<td>1,3,5-benzenetricarboxylic acid</td>
</tr>
<tr>
<td>TT-TM</td>
<td>tri-trimethylsilylethyl trimesyl</td>
</tr>
</tbody>
</table>
LIST OF FIGURES AND TABLES

Page

22 Table 1: Amino Acid Content and Charge Contribution in Putative Transmembrane Segments of Various Membrane Proteins
55 Figure 1: Succinylation of DNP-Amino Acids
56 Figure 2: Time Course of DNP-Amino Acid Succinylation
59 Table 2: Succinyl-Ester Quantitation
61 Figure 3: Identification of the DNP-Threonine Ester Products
63 Table 3: Reaction Conditions for the Assessment of the Effect of the Molar Ratio of Trifluoroacetic Acid to Triethylamine on the Rate of Succinylation
64 Figure 4: The Effect of the Molar Ratio of Trifluoroacetic Acid to Triethylamine on the Rate of Succinylation
66 Table 4: Reaction Conditions for the Assessment of the Effect of the Molar Ratio of Succinic Anhydride to Trifluoroacetic Acid on the Rate of Succinylation
67 Figure 5: The Effect of the Molar Ratio of Succinic Anhydride to Trifluoroacetic Acid on the Rate of Succinylation
68 Figure 6: Proposed Equilibria Involved in the Succinylation of DNP-Threonine in the Presence of Trifluoroacetic Acid
70 Table 5: The Effect of Different Salts on the Rate of Succinylation
72 Figure 7: Amino Acid Sequence of Apobacteriorhodopsin: Sites of Succinylation
Table 6: Net Negative Charge of Apobacteriorhodopsin: Before and After Succinylation

Figure 8: Elution Profile of Succinyl-Apobacteriorhodopsin from a Sephadex G-50 Column

Figure 9: Enzymatic Removal of the N-Terminal Pyroglutamyl Residue

Table 7: Stability of Succinylated DNP-Serine

Table 8: Stability of [14C]Succinyl-Apobacteriorhodopsin

Table 9: Stability of Succinylated DNP-Serine Methyl Ester

Figure 10: Model Peptide Synthesis

Figure 11: Synthesis of Diprotected Trimesic Acid

Table 10: Formation of Diprotected Trimesic Imidazolide

Figure 12: Reactivity of CDI and DT-TM Imidazolide

Figure 13: Reactivity of CDI and DT-TM Imidazolide with Sephadex LH-20

Figure 14: Time Course of DNP-Prolylserinamide Trimesylation with Sephadex LH-20 Treated or Untreated DT-TM Imidazolide

Figure 15: HPLC Analysis of the Reaction Products

Figure 16: Dowex 50x2 Removal of Imidazole

Figure 17: Stability of Diprotected Trimesic Imidazolide

Figure 18: Stability of Diprotected Trimesic Imidazolide in Me2SO

Figure 19: Effect of Et3N on the Rate of Trimesylation with DT-TM Imidazolide in DMF

Figure 20: Time Course of Model Peptide Trimesylation with DT-TM Imidazolide in DMF

Figure 21: Trimethylsilyl-ethyl Ester Deprotection
Figure 22: Tritiated Trimesylation of DNP-Prolylserinamide with DT-TM Imidazolide and Subsequent Bu₄N⁺F⁻ Deprotection

Figure 23: Proposed Equilibria Involved with DNP-Proline During Trimesylation with DT-TM Imidazolide

Figure 24: Time Course of Trimesylation of Model Peptides with DT-TM Imidazolide in Anhydrous Me₂SO

Figure 25: Stability of the Trimesyl Derivative of DNP-Prolylglycyltyrosinamide

Figure 26: Time Course of Apobacteriorhodopsin Trimesylation with DT-TM Imidazolide in DMF

Figure 27: Time Course of Apobacteriorhodopsin Trimesylation with DT-TM Imidazolide in Anhydrous Me₂SO

Figure 28: Elution Profile of Trimesylated Apobacteriorhodopsin from a Sephadex G-150 Column

Figure 29: Determination of Apparent Molecular Weight of Trimesylated Apobacteriorhodopsin by Gel Permeation Chromatography on Sephadex G-150

Figure 30: Elution Profile of Trimesylated Apobacteriorhodopsin from a Sephadex G-150 Column

Table 11: Adsorption of Trimesyl-Apobacteriorhodopsin to Sephadex G-150

Table 12: Summary of the Electrophoretic Results for Trimesylated Apobacteriorhodopsin on Acidic LDS Polyacrylamide Gels

Table 13: Charge Density of Proteins During S(L)DS-PAGE

Figure 31: Formation of Diprotected Trimesic Anhydride

Table 14: Generation of Dicyclohexyl-N-(DT-TM)-urea (DCAU) During DT-TM Anhydride Formation by DCC
Table 15: DT-TM Anhydride Reactivity in Various Solvents

Figure 32: Stability of DT-TM Anhydride

Figure 33: The Effect of Amphipathic salt on the Rate of Trimesylation

Figure 34: The Effect of Catalyst on the Rate of Trimesylation

Figure 35: Time Course of Model Peptide Trimesylation

Figure 36: Time Course of Trimesylation of Apobacteriorhodopsin

Figure 37: Isolation of Diprotected Trimesyl-Apobacteriorhodopsin from a Sephadex LH-60 Column

Figure 38: Isolation of Deprotected Trimesyl-Apobacteriorhodopsin from a Sephadex LH-60 Column

Figure 39: Elution Profile of Trimesyl-Apobacteriorhodopsin from a Sephadex G-150 Column

Figure 40: $K_{av}$ versus the Molecular Weight of Nine Standard Proteins as observed by Gel Permeation Chromatography on a HPLC TSK-250 Column

Figure 41: Elution Profile of Trimesyl-Apobacteriorhodopsin from a HPLC TSK-250 Column

Figure 42: Removal of the N-Terminal Pyroglutamyl Residue from Trimesyl-Apobacteriorhodopsin by Pyroglutamate Aminopeptidase

Table 16: Determination of Pyroglutamic Acid Release by Pyroglutamate Aminopeptidase Digestion

Table 17: Amino Acid Sequencing Results from Deblocked Trimesyl-Apobacteriorhodopsin
I. Introduction:

Biological membranes are primarily composed of a semi-permeable phospholipid bilayer in which a variety of proteins are embedded to varying extents. These proteins can move laterally within the plane of the lipid bilayer, as schematically illustrated by the "fluid mosaic model" of Singer and Nicolson (1972). These membrane associated proteins are of great importance because it is their individual functions that give rise to the properties of the biological membrane in which they are found. In general, these functions are attributed to the transport of water soluble molecules across the phospholipid bilayer, the catalysis of chemical reactions or the recognition of specific extracellular or intracellular chemical signals.

Membrane proteins are divided into two groups on the basis of their strength of association with the membrane. Peripheral membrane proteins are weakly associated and can be removed from the membrane by high concentrations of salt or urea. Integral membrane proteins have strong hydrophobic interactions with the phospholipids of the membrane; these proteins usually transverse the lipid bilayer, at least once. Detergents are required for the solubilization of integral membrane proteins in aqueous buffers; in the absence of detergents they aggregate and precipitate out of solution. This behaviour is due to the hydrophobic amino acid residues in the proteins' transmembrane segments, which strongly interact in the absence of detergents or chaotropic agents, resulting in large molecular weight protein aggregates (Khorana, 1980, Ozols et al., 1985 and Chen et al., 1986).

Even in the presence of detergents and chaotropic agents, some
intrinsic membrane proteins, such as bacteriophage M13 coat protein, cytochrome P450 and lac permease and highly amphiphilic proteins, such as apolipoprotein B are difficult to solubilize under aqueous conditions. The coat protein of the bacteriophage M13 was found to remain as a dimer even in the presence of high concentrations of the detergent sodium dodecylsulphate (SDS) (Knippers and Hoffmann-Berling 1966). The native form of cytochrome P450 was observed to exist as a dimer, even under the reducing and denaturing conditions utilized for SDS-polyacrylamide gel electrophoresis (PAGE) (Ozols et al., 1985). The lac permease aggregates even in the presence of detergents and phospholipids (Roepe and Kaback, 1989). Apolipoprotein B forms large protein aggregates in the presence of 7.8 M guanidine hydrochloride (Steele and Reynolds, 1979a and Steele and Reynolds, 1979b) and in the presence of SDS during PAGE (Meredith, 1984).

The ultimate goal is to better understand and characterize membrane proteins' structure function relationship, specifically the topology of the protein within the membrane and the biochemical processes by which they transport, recognize or catalyze reactions on specific molecules. In order to achieve this, the primary structure of the membrane protein must be determined. The standard procedures that are utilized for determining the amino acid sequence for water soluble proteins involves the fragmentation of the protein, either enzymatically or chemically, followed by the isolation and sequencing of the peptides. The entire sequence of the protein can be reconstructed from the peptide sequences through overlapping segments (Needleman, 1975 and Needleman, 1977). The capability of sequencing picomolar amounts of water soluble proteins has been realized with the advent of tandem gel and continuous elution electrophoresis, micro-bore HPLC, fast atom bombardment mass spectrometry and electrotransference of
both proteins and peptides onto polyvinylidene difluoride (PVDF) membranes for direct amino acid analysis, cleavage and sequencing (Matsudaira, 1987, Shively et al., 1989 and Wilson, 1989).

I.1. Amino Acid Sequencing of Membrane Proteins

Progress in the determination of the amino acid sequence of integral membrane proteins by actual protein sequencing has been much slower than for water soluble proteins. It is the membrane proteins' inherent lack of water solubility and poor manageability that hinders the standard procedures of protein fragmentation, peptide fragment purification and sequencing (Khorana et al., 1979, Gerber and Khorana, 1982, Ozols, 1986 and Shively et al., 1989).

The polytopic transmembrane protein, lac permease from *Escherichia coli*, was recently characterized in an unusual soluble monomeric form after the overexpression of the lacY gene. It appears to be peripherally associated with the membrane, being extractable from the membrane with 5 M urea and it remains in aqueous solution after the urea is removed. This protein appears to have adopted a nondenatured conformation in that it can bind ligand, it will undergo reconstitution with *E. Coli* phospholipids and then is no longer extractable with urea. However, this soluble form of the membrane protein, lac permease, was observed to aggregate in a time- and concentration-dependent manner (Roepe and Kaback, 1989). This unique form of a membrane protein would probably be difficult to manage during standard protein fragmentation and sequencing procedures. Furthermore, the formation of this type of water soluble membrane protein may not be universally applicable and the extent of water solubilization would definitely be dependent on the protein itself.
I.1.A. Enzymatic Protein Fragmentation

Fragmentation of membrane proteins by the utilization of enzymes is generally incomplete due to the inability of the enzymes to effectively function on a protein suspension or under the conditions required to achieve protein solubility. Furthermore, the cleavage site on the protein may not be exposed to the enzyme due to protein aggregation or due to the detergent micelle (Hunkapiller et al., 1984, and Chen et al., 1986). Some enzymes may be tolerant to the addition of organics; pyroglutamate aminopeptidase is functional in 10% methanol (Szewczuk and Mulczyk, 1969), 4% dimethylsulfoxide (Me₂SO; Browne and O’Cuinn, 1983), 3% dimethylformamide (DMF; Friedman et al., 1985) or 5% glycerol (Capecchi and Loudon, 1984). However, these small amounts of organic solvents are not in themselves sufficient to improve the solubility of membrane proteins under aqueous conditions. Furthermore, only a few enzymes are capable of functioning in the presence of detergents or chaotropic agents. Endoprotease V8 and endoprotease lys-C which cleave at glutamyl and lysyl residues, respectively, are both functional in the presence of 0.5% SDS (Hunkapiller et al., 1984 and Cleveland et al., 1977) as are proteinase K and carboxypeptidase A (Henry et al., 1986). Endoprotease V8 is also functional in the presence of 6 M urea or 5.5 M guanidinium hydrochloride (Drapeau, 1980). Limited protein digestion was achieved in the presence of 0.5% SDS and 10% glycerol for chymotrypsin, endoproteinase V8 and papain (Cleveland et al., 1977). However, Henry et al. (1986) found trypsin and chymotrypsin to be rapidly denatured in the presence of SDS.

I.1.B. Organic Solvent Solubilization and Chemical Fragmentation

Membrane proteins and their fragments are significantly more
soluble in organic solvents, such as trifluoracetic acid (TFA), Me₂SO, DMF and formic acid (Khorana et al., 1979, Llinas et al., 1980, Huang et al., 1981, Gerber and Khorana, 1982 and Hennessey and Scarborough, 1989). The novel solvent, sulfolane (thiophene, tetrahydro-1,1-dioxide), was proposed for solubilizing hydrophobic proteins, though it was only tested on water insoluble zein and globin chains (Vecchio et al., 1984). Its usefulness for very hydrophobic integral membrane proteins has not yet been assessed.

There has also been some limited investigation into the utilization of enzymes in organic solvents; lipases were studied and were found to be functional as a suspension in a variety of organic solvents containing only 0.02 - 1.0% water. However, upon complete solubilization of the enzyme in DMF or Me₂SO, no activity was observed, presumably due to the denaturation of the enzyme (Zaks and Klibanov, 1985). The usefulness of this type of procedure for protein fragmentation is uncertain due to the inability of the lipases, and presumably other enzymes, to function when dissolved in the solvents that are known to solubilize membrane proteins.

The major problem for the chemical cleavage of membrane proteins is the insolubility of the protein and hence the poor rate of reaction due to the two phase system. Some chemical cleavages have been improved by the use of organic solvents that solubilize membrane proteins, at least partially: cyanogen bromide (CNBr) cleavage (Gross and Witkop, 1962) can be performed in 70% formic acid (Craven et al., 1965), o-iodosobenzoic acid fragmentations is done in 80% acetic acid and 7 M guanidinium hydrochloride (Mahoney and Hermodson, 1979) and mild acid cleavage of Asp-Pro peptide bonds is performed in 70% formic acid (Hunkapiller et al., 1984).

I.I.C. Protein Fragment Isolation
Isolation of the protein fragments is complicated by the insolubility and aggregation of hydrophobic fragments even in the presence of detergents. Organic solvent mixtures have aided the isolation of peptides from Sephadex LH-20 and LH-60 gel permeation chromatography and from high pressure liquid chromatography (HPLC). Some of the useful solvent mixtures that have been employed were 88% formic acid, ethanolic formic acid, propanolic formic acid, 0.5% TFA and acetonitrile (CH$_3$CN) and 0.5% TFA or heptafluorobutyric acid (HFBA) and propanol (Gerber et al., 1979, Khorana et al., 1979, Takagaki et al., 1980, Walsh et al., 1981, Gerber and Khorana, 1982, Regnier, 1983 and Hermodson and Mahoney, 1983). However, fragment aggregation and poor recovery of very hydrophobic fragments still occurs (Khorana et al., 1979, Gerber and Khorana, 1982, Ozols, 1986 and Reddy et al., 1986).

Cytochrome P450 peptides, that were generated by CNBr cleavage or trypsin cleavage (after succinylation of lysine side chains) were separated by reversed phase HPLC using 0.1% TFA or HFBA and 75% propanol. Even though these solvent systems are regarded as being the best for protein and peptide solubilization and elution (Hermodson and Mahoney, 1983), some hydrophobic peptides were found to be irreversibly adsorbed to the column (Ozols et al., 1985). The basis of protein adsorption is presumably through conformational changes upon exposure to the reversed phase column (C$_8$ or C$_{18}$). 1-propanol was able to elute these proteins by allowing the proteins' conformation to revert back (Katzenstein et al., 1986). Furthermore, the recovery of peptides is generally observed to be quite poor from reversed phase HPLC; only 20% of the applied tryptic peptides from a large N-terminal CNBr peptide of succinylated adenine nucleotide translocater (AN$_1$) was recovered from the HPLC column (Boulay and Vignais (1984). Current usage
of micro-bore HPLC systems has improved the overall recoveries, resolution and detection sensitivity (15-20 fold). However, the major problem for certain peptides is the losses due to surface adsorption (Wilson, 1989).

I.I.D. Automated Protein Sequencing

When sequencing a hydrophobic protein fragment using the automated Edman degradation sequencer (Edman and Begg, 1967), the fragment can be lost from the spinning cup during the standard organic solvent washes (benzene, ethylacetate, 1-chlorobutane) within each sequencing cycle. The addition of detergents has been utilized to improve the sequencing of these types of hydrophobic fragments (Bailey et al., 1977 and Hunkapiller and Hood, 1978). However, the addition of Polybrene, a polymeric quaternary ammonium salt, to the spinning cup provides the best results for assisting the non-covalent adhesion of the peptide to the cup, thereby reducing peptide losses during the organic solvent washes and extraction of the 2-anilino-5-thiazolinone derivatives of the amino acids (Tarr et al., 1978). Even with these methods, sequencing membrane proteins is difficult (Gerber and Khorana, 1982) and errors can arise in the assignment of the amino acid sequence (Ovchinnikov et al., 1979).

Improvements to the system such as mechanical alterations and improved purity and storage of reagents, have resulted in the ability to sequence "micro" amounts of protein or peptide (Walsh et al., 1981 and Begg et al., 1984)). This was further improved with the development of the gas liquid phase sequenator in which the reagents are added in a gaseous phase to a chamber containing the protein or peptide on a porous glass support (Hewick et al., 1981, Hunkapiller et al., 1983 and Esch, 1984). This miniaturized system enables the volumes of organic solvents (heptane,
ethylacetate and 1-chlorobutane) to be minimized in the attempt to further reduce peptide washout. A technique that can be utilized is the electroblotting of proteins from SDS-PAGE onto activated glass paper (negatively or positively charged) to permit direct gas-liquid phase sequencing (Aebersold et al., 1986) or onto a PDVF membrane (Shively et al., 1989). Even with these methods, peptide washout may still occur; any small loss of protein becomes much more significant as the amount of protein is reduced to the "microsequencing" scale (Hunkapiller et al., 1984).

Sequence studies on integral membrane proteins and their fragments still remain generally difficult due to the inability to perform classical protein sequencing techniques on long stretches of hydrophobic amino acid residues, particularly those that are found in transmembrane segments. This was observed in the Band 3 protein of erythrocytes (Jay and Cantley, 1986) as well as segments found in the interior of water soluble proteins, β-subunit of luciferase from Vibrio harveyi (Johnston et al., 1986).

I.1.E. Alternative Sequencing Methods

Protein sequencing by tandem mass spectrometry requires the protein to undergo fragmentation via the standard enzymatic or chemical procedures and the peptides subsequently isolated on a reversed phase HPLC system. The peptides are then dissolved in glycerol and subjected to ionization by a liquid secondary-ion mass spectrometry on a triple quadrupole mass spectrometer. This technique is limited to sequencing peptides that are under 1,800 daltons (Hunt et al., 1986). Fast atom bombardment mass spectrometry (FAB/MS) has also gained considerable use for the actual amino acid sequence and for the confirmation of post-translational modifications of specific residues (Shively et al., 1989). These sequencing approaches when
used for membrane proteins will obviously have the inherent problems of complete protein fragmentation and peptide isolation.

A novel procedure for the positioning of amino acid residues in the sequence of a protein involves the labelling of the N-terminus of the protein with an amino specific reagent (radioiodinated Bolton-Hunter reagent) after all lysine side chains are derivatized with phenylisothiocyanate (PITC) and the first residue is removed by Edman degradation (Jay, 1984). Subsequently the protein is cleaved selectively and the resulting labelled peptides sized, thereby, positioning the site of cleavage. This procedure is limited to proteins that are amenable to both the cleavage and the sizing of the fragments.

I.2. Terminally Blocked Proteins

Proteins that are blocked at their N-terminus present a fairly major problem in protein sequencing due to their inability to react with PITC at their amino terminus; the Edman degradation, whether manual or automated, can not proceed. Several types of N-terminal blocking groups have been detected and include acetyl (which is the most common), pyroglutamyl (pyrrolidone carboxyl), fatty acyl, α-ketoacyl, glucuronyl and methyl (Wold, 1981). The N-terminal formyl group, though present on the initiator methionine residue for all prokaryotic protein syntheses, has rarely been found in isolated proteins (Wold, 1984 and Vaaler et al., 1986); this is likely due to the action of a deformylase (Thibodeau et al., 1978) or possibly due to a N-α-acylpeptide hydrolase (Gade and Brown, 1978). A notable exception is gramicidin A which has a formylated valine as its first amino acid residue (Urry et al., 1971).

Evidence is accumulating which suggests that blocking of the N-terminus of proteins is a general phenomenon for eukaryotic proteins
especially those that are intracellular (Tsunasawa and Sakiyama 1984); it has been estimated that 90% of the soluble proteins of mouse L cells and 80% of those in Erlich ascites cells are N-α-acetylated (Brown and Roberts 1976). It is further suggested that N-α-acetylation is a cotranslational event that is catalyzed by a ribosomal bound N-α-acyltransferase using acetyl-CoA as the donor molecule (Tsunasawa and Sakiyama 1984). It was discovered that fructose P2 aldolase when isolated from skeletal muscle in the presence of phenylmethanesulfonyl fluoride, a serine protease inhibitor, had an acetylated methionine residue as its N-terminus (Lebherz et. al., 1984); the previously determined N-terminal proline residue from the protein isolated in the absence of any protease inhibitor (Tolan et. al., 1984) is actually the penultimate amino acid residue. The N-terminal sequence had previously been incorrectly identified presumably due to the exposure of the protein to endogenous acetylamino acyl hydrolases which removed the N-terminal acetylated amino acid.

F. Wold (1984) postulated that proteins are N-α-acetylated during translation and that cellular N-α-acetylamino acyl hydrolases (also known as N-α-acylpeptide hydrolases) may normally be involved in processing the N-α-acetylated protein, by removing the acetylated amino acid residue prior to further N-α-acetylation of the protein or prior to some other type of posttranslational modification. N-α-acetylation maybe a general feature of eukaryotic proteins, however, no biological role has been correlated with it. The involvement of N-α-acylpeptide hydrolases with protein biosynthesis is also not known (Jones and Manning, 1985).

A group of enzymes that are capable of removing the blocked amino acid from the rest of the protein are known as N-α-acylpeptide hydrolases. Some of them have been isolated and characterized from bovine
liver (Gade and Brown, 1978) and human erythrocytes (Jones and Manning, 1985). The highest catalytic activity was observed for substrates whose N-terminal residue was acetyl alanine, however, a variety of acetylated peptides were found to be suitable substrates. Even formyl-Met-Val could be hydrolyzed by this enzyme when at its pH optimum, though at a relatively slow rate. (The enzyme has different pH profiles depending on which type of acyl group is present on the substrate.) (Jones and Manning, 1985).

In spite of the biological uncertainties of this enzyme, it has been utilized to aid in the sequencing of N-α-acetylated proteins (Tsunasawa and Sakiyama, 1984). The N-α-acylpeptide hydrolase was not found to be functional on either native or denatured ovalbumin (Gagnon et al., 1978). However, upon CNBr fragmentation of the ovalbumin, and subsequent isolation of the N-terminal octapeptide the enzyme readily hydrolyzed the acetylated amino acid residue permitting the N-terminus to be sequenced. The N-terminal residue was identified by comparing the amino acid composition of the entire octapeptide to the established amino acid sequence from residue 2 through 8 (Tsunasawa and Sakiyama, 1984).

The other common blocking group at the N-terminus is pyroglutamate and like the acetyl group its biological significance is also unclear except that some of the blocked proteins are biologically active peptides, such as neurotensin and thyrotropin-releasing hormone (Friedman et al., 1985). An enzyme that specifically hydrolyzes the pyroglutamate residue from peptides has been purified (Armentrout and Doolittle, 1969 and Doolittle, 1972) and is now commercially available as pyroglutamate aminopeptidase (EC. 3.4.11.8).

This enzyme is ubiquitous and appears to be a thiol protease which is sensitive to both sulphydryl group blocking reagents and heavy metal ions.
(Doolittle, 1972 and Fujiwara et al., 1981). Furthermore, the specificity of pyroglutamate aminopeptidase appears to be quite broad with respect to the penultimate amino acid; of 11 different N-pyroglutamyl peptides, only the peptide bond between pyroglutamate and proline failed to be hydrolyzed (Doolittle, 1972 and Browne and O'Cuinn, 1983). The enzyme could also catalyze the hydrolysis of pyroglutamate when the penultimate amino acid was replaced with β-naphthylamine (Szewczuk and Mulczyk, 1969) or by 7-amino-4-methylcoumarin (Browne and O'Cuinn, 1983). However, pyroglutamate aminopeptidase is quite specific for the 5 membered ring of the pyroglutamate residue and will not recognize a 6 membered ring (Capecci and Loudon, 1985). The enzyme can be inactivated with L-pyroglutamyl chloromethylketone (Fujiwara et al., 1981) and 5-oxoproline which appears to be a transition state aldehyde inhibitor (Friedman et al., 1985).

Pyroglutamate aminopeptidase has a molecular weight of approximately 24,000 (Browne and O'Cuinn, 1983) and appears to have its N-terminal residue blocked with an undetermined group (Armentrout and Doolittle, 1969 and Doolittle, 1972). This enzyme readily hydrolyzes biological peptides (Browne and O'Cuinn, 1983) and has also been used to deblock proteins such as bovine fibrinogen (Armentrout and Doolittle, 1969) and human serum mucoid (Szewczuk and Mulczyk, 1969) enabling Edman degradation to proceed. These proteins obviously must have their N-terminus exposed sufficiently to allow the enzyme to hydrolyze the blocking residue. Other proteins may have their N-terminus "buried", thereby being inaccessible to the enzyme or the protein of interest may not be water soluble, such as a membrane protein (ie. bacteriorhodopsin). In these cases, even though the N-terminal residue maybe pyroglutamate, the protein would remain blocked. This inaccessibility of the cleavage site would be a common
problem for all deblocking enzymes, such as the N-α-acylpeptide hydrolases.

Mass spectrometry provides an alternative procedure for determining the N-terminal sequence of blocked proteins. Proteins of interest are digested with a proteinase, permethylated and the blocked N-terminal peptides (mono, di, tri and tetrapeptides) form volatile derivatives that are detected using a g.l.c.-mass spectrometer (Rose et al., 1984). In this manner, human adenine phosphoribosyltransferase was found to be N-α-acetylated (Wilson et al., 1986). The proteinase digestion could be a potential problem for membrane proteins due to their water insolubility which would result in incomplete cleavage and poor recovery of N-terminal peptides.

I.3. Deduced Amino Acid Sequence from DNA

Sequencing of cellular proteins, especially those that exist in relatively small amounts, can be accomplished through recombinant DNA technology. There are two methods that are generally employed: 1. Generation of a cDNA library from isolated mRNA and then the identification of the genomic DNA fragment. Sequencing of even one protein fragment helps to confirm the right selection of the reading frame from the six possible choices. 2. Extracting the protein coding region from a genomic DNA or cDNA library by utilizing a synthesized DNA probe based on the partial N-terminal sequence of the isolated protein. In either case, the obtained amino acid sequence is only deduced; confirmation of the sequence must come from the protein itself. Furthermore, DNA sequencing alone can only predict possible sites of post-translational modifications (ie glycosylation at -Asn-Xaa-Ser(Thr)- sequences (Wold, 1985)) but not the final structure of the mature, biologically active protein (Hunkapiller et al., 1984). Based on the DNA sequence of apolipoprotein B, there are 20 potential sites of
glycosylation but analysis of the mature protein revealed that only 13
asparagine residues were modified (Chen et al., 1986).

Post-translational events are believed to be the most important
aspect of proteins with respect to their biological activity and targetting within
the cell (Wickner and Lodish, 1985, Yan et al., 1989 and Hartmann et al., 1989).
The extent and types of post-translational events are constantly growing: A
unique post-translational event was identified in concanavalin A
biosynthesis which involves the intrapolypeptide cleavage, transposition and
ligation of the glycosylated pro-concanavalin A (Carrington et al., 1985 and
Chrispeels et al., 1986). There appears to be evidence implicating the existence
of a peptidyl transferase enzyme in HL60 cells (a human cell line) and rabbit
peritoneal granulocytes which covalently attaches the tetrapeptide, tufts in
onto a 100,000 dalton membrane acceptor protein (Najjar and Bump, 1989).

Mutant studies and site directed mutagenesis can provide
additional information on the importance of specific residues in the protein.
Work of this type has been performed on bacteriorhodopsin (aMogi et al.,
1989, bMogi et al., 1989 and Stern and Khorana, 1989) and on ATP synthetase
as summarized by Futai et al (1989). There is a need for both DNA and
protein sequencing in order to unravel protein biosynthesis and its structure
function relationship.

I.4. Topological Models of Membrane Proteins

X-Ray diffraction and electron microscopy can be employed to assess
the three dimensional structure of proteins. Neither of these methods
require prior knowledge of the amino acid sequence but they do require three
dimensional or two dimensional crystals, respectively. These procedures are
not easily applied to membrane proteins due to the difficulty in obtaining
adequate crystals. Few membrane proteins have been crystallized but new techniques are being developed that may improve crystal formation and analysis (Dodson, 1986). A crystal of the bacterial photosynthetic reaction center was generated by Michel (1982) from which its high resolution structure was determined (Deisenhofer et al., 1985 and Deisenhofer and Michel, 1991). Recent improvements in electron diffraction microscopy are yielding high resolution protein structures (Henderson et al., 1990 and Kuhlbrandt and Wang, 1991). Furthermore, new techniques in nmr, such as heteronuclear 3° and 4° nmr (Clore and Gronenborn, 1991) and rotational resonance nmr (Creuzet et al., 1991) are also providing unique tools by which to study protein topology.

Alternatively, circular dichroism spectra can assess the percentage of the protein that is in one secondary structure or another. This procedure is generally used for assessing membrane proteins. However, when other methods are used, such as the laser Raman spectra, the proportions of the same protein found in one secondary structure or another are different. This was observed for the bacteriophage M13 coat protein (Henry et al., 1986). Another limitation to these methods is that neither can identify the actual sequences that are involved in the proposed secondary structures.

The primary sequence of many membrane proteins are currently being predicted from the DNA sequence of the corresponding gene or from the cDNA sequence generated from the mRNA. This deduced amino acid sequence is then being used to predict the topology of the protein within the membrane by employing one of the standard methods for predicting secondary and tertiary protein structure of water soluble proteins. These methods include:

1. Analysis of the primary sequence via the Chou and Fasman
algorithm which assesses segments of the protein that may exist in an \( \alpha \)-helix, \( \beta \)-sheet or \( \beta \)-turn conformation (ie. Leu occurs frequently in the middle of \( \alpha \)-helices and \( \alpha \)-helices normally begin with a negatively charged residue and end with a positively charged amino acid (Chou and Fasman, 1978)). This is an empirical method which is based on the frequency that a particular amino acid is present in one secondary structure or another, in different proteins, all of which were water soluble proteins (Chou and Fasman 1978). The best results for accuracy of predicting secondary structures was obtained when five empirical algorithms were jointly used in a computer program; \( \sim 72\% \) for \( \alpha \)-helix, 68\% for \( \beta \)-sheets and 66\% for \( \beta \)-turns (Chou and Fasman 1978 and Garnier et al., 1978). However, the Chou and Fasman algorithm is generally used due to their reasonable accuracy for \( \alpha \)-helix predictions. The Chou and Fasman algorithm is generally used for analyzing the amino acid sequence of membrane proteins, such as the acetylcholine receptor (Criado et al., 1985), even though there is little evidence to indicate that it is strictly applicable to these types of proteins.

2. Sequence homologies are also used for structural predictions and there are computer programs that contain thousands of protein sequences in their data base (Toh et al., 1983). The cooperativity of 6 or more amino acid residues must be identical in order for them to have similar secondary structures (Kabsch and Sander, 1984). These types of secondary structures have been well studied for soluble proteins; bond angles and lengths for both conformation and packing of such structures has been assessed (Chothia, 1984 and Chothia and Finkelstein, 1990). The direct application to membrane proteins has not been assessed; the local enviroments of the amino acids would be different even though the sequences may be identical.

3. A method that is generally used on membrane proteins, in
conjunction with the Chou and Fasman predictions, is the hydropathy index or hydrophobicity profile (Kyte and Doolittle, 1982). This is a measurement of the amino acid's relative affinity for a hydrophobic phase. The best assessment for the amino acid residue was found when a "window" of 7 - 11 amino acids are considered and over which the values are averaged. This method is used to identify a hydrophobic amino acid segment within the protein that contains no residues bearing a charge on their side chain; such a protein segment which is α-helical in structure and long enough to span the lipid bilayer, is considered to be a transmembrane segment (Eisenberg, 1984).

A good procedure, though one that is generally not used, is a computer algorithm which uses a 21 amino acid "window" to find membrane spanning helices. This 21 amino acid window was based on each residue in an α-helix spanning 1.5 Å, thus totalling 32 Å which is the approximate thickness of the lipid bilayer. If this "window" was found to be highly hydrophobic then it was considered that this α-helix was an 'initiator' of membrane association. A limitation of this procedure is that it can only select α-helices; no other secondary structures are considered (Eisenberg, 1984).

Though Kyte and Doolittle (1982) are often cited, no one method is generally accepted for the measurement and calculation of the relative affinity of the amino acid residue for the hydrophobic phase. This is further complicated by the use of different scales which makes direct comparison difficult (Eisenberg, 1984 and Fauchère, 1985).

4. The hydrophobic moment is a measure of the amphilicility (hydrophobicity asymmetry) of a segment of the protein, which identifies segments of amino acids that when placed in an α-helix arrangement result in one side being polar while the other side is nonpolar, with respect to the side chains of the amino acids. This procedure, however, is rarely utilized
and is prone to misinterpretation (Eisenberg, 1984).

5. A superior method but one that is not generally used is the hydrophobic moment plot in which the entire protein is assumed to be α-helical in structure and an 11 amino acid "window" is assessed with respect to its hydrophobic moment and its hydrophobicity index. The plot of the hydrophobic moment versus the hydrophobicity index positions the hydrophobic α-helix into areas on the graph distinguished by 'globular', 'surface' and 'transmembrane' α-helices. This method is capable of distinguishing between helices buried in globular proteins to those spanning the lipid bilayer (Eisenberg, 1984).

A potential flaw in all these methods is that only α-helical secondary structures in conjunction with their hydrophobicity are considered for transmembrane segments even though other secondary structures are equally plausible within the lipid bilayer: 1. β-sheets which could form a β-barrel (Kennedy, 1978); evidence for the ability of β-sheets to interact with lipid was observed with the bacteriochlorophyll protein which is composed of 3 subunits each having 15 strands of twisted β-sheets which provides a hydrophobic core for the seven bacteriochlorophyll A molecules while presenting a polar surface to the solvent (Mathews et al., 1979). 2. The novel β-helix proposed by Urry et al. (1971) for the gramicidin A transmembrane protein in which the side chains of the amino acid residues are in the middle of the helix. 3. Evidence indicates the existence of an 11 amino acid straight chain (in an non-α-helical structure) that spans the membrane bilayer in the acetylcholine receptor (Criado et al., 1985). α-helices appear to be the most common type of secondary structure in proteins as well as the α-helical coiled coil structural motif (through short regions of helix-helix packing). They also appear to occur for some membrane proteins, such as the bacterial
photosynthetic reaction center (Rees et al., 1985), which tends to confirm the utility of the hydropathy plots. However, assumptions that only α-helices can form transmembrane segments could result in erroneous assignments and inaccurate topological models.

Some topological models of membrane proteins were based on assumptions that transmembrane segments could not contain charged amino acid residues; this limitation is being modified with the realization that amphipathic helices could provide hydrophilic pores through which ions and molecules could be transported (Cohen and Parry, 1986). Furthermore, ion pairs between basic and acidic side chains ("salt bridges") which are known to occur in soluble proteins, even when buried and devoid of solvent, are likely to occur in membrane proteins within the lipid bilayer, providing stability and/or function (Honig and Hubbell, 1984).

Even with this improved interpretation, the methods used for the identification of the putative transmembrane segments of membrane proteins have not altered (Cohen and Parry, 1986). A notable example of the inadequacies of the models purported is the acetylcholine receptor. In 1983, the deduced amino acid sequence of the α-subunit was obtained; 4 α-helices, composed mostly of hydrophobic amino acids and containing no residues bearing a charged side chain, were proposed to be transmembrane segments (Noda et al., 1983). In 1984, a additional fifth transmembrane α-helix was proposed on the basis of different theoretical analyses (Finer-Moore and Stroud, 1984). The addition of a sixth and seventh transmembrane segment occurred in 1985, based on evidence of monoclonal antibody crossreactivity to the native protein. One transmembrane segment was an amphipathic α-helix while the other segment was only 11 amino acids in length and thus was purported to be a linear peptide chain, in order for it to span the
membrane (Criado et al., 1985). Finally in 1986, two of the previously proposed transmembrane α-helices were deleted (the fourth and fifth) due to immunological evidence with monoclonal antibodies that indicated that they were in the cytoplasm and not spanning the lipid bilayer (Ratnam et al., 1986).

Topological models for membrane proteins provide a hypothesis to which experiments can be directed in the attempt to prove or disprove the model. However, from the example of the acetylcholine receptor, it is clear that the basis by which these models are developed is inadequate and by the current methodologies, extensive work is required to collect evidence to validate or correct the topological model.

More direct evidence for topological data involves the identification of the sites on the native membrane protein that are sensitive to chemical modification, enzymatic cleavage or immunological detection (Jennings, 1989). The more points of reference on the topology of the protein within the membrane, the more accurate the proposed model will be. In order to accumulate points of reference on intrinsic membrane proteins, a variety of 'probes' can be utilized to investigate the protein's secondary and tertiary structure, its topology in the membrane, the site(s) of post-translational modification and the amino acid residues that are involved in the catalytic or binding site. The types of probes consist of membrane permeant labels (eg. lipid permeable photolabelling probes (Brunner et al., 1980, Brunner and Richards, 1980 and Hoppe et al., 1984) such as photoreactive fatty acid derivatives (Leblanc et al., 1982)) and nonpenetrating labels (eg. isethionyl acetimidate (Whiteley and Berg, 1974) and diazobenzene [35S]sulfonate (Zhang et al., 1984)), sensitivity to enzymatic fragmentation (eg. trypsin (Clarke, 1976), proteinase K (Brdiczka and Krebs, 1973)), radioactive substrates for protein modification (eg. fatty acids (Capone et al., 1983), phosphate
(Rozengurt et al., 1983)) and selective chemical derivatization of amino acid residues (eg. N-ethylmaleimide (Boulay and Vignais, 1984), suicide inhibitors (Abeles, 1978)). The identification of the site of modification or cleavage position on the membrane protein necessitates that either the entire molecule or selected fragments must be sequenced. However, the major limitation to the topological studies is the inherent hydrophobic nature of these fragments, which results in their insolubility, aggregation and their poor sequencing behaviour, as previously stated.

I.5. Putative Transmembrane Segments

Table 1 consists of a list of several putative transmembrane segments from various membrane proteins, as assigned by the authors based on their interpretations of the accumulated data on the appropriate membrane protein. Most of the topological models are based solely on the hydropathy profile of the deduced amino acid sequence in conjunction with \( \alpha \)-helix predictions, while some proteins (BR, ACHR) have actually undergone more rigorous testing for their orientation in the membrane; the proteins' sensitivity to chemical modification or proteolytic digestion (Gerber et al., 1977 and Engelman et al., 1982) or to its interaction with monoclonal antibodies (Ratnam et al., 1986). There are some transmembrane segments that are only tentatively assigned on the basis of very few points of reference. (The seventh and eighth transmembrane segments of cytochrome P450 may or may not traverse the membrane (Ozols et al. 1985).) Also the amino acid sequence of maltase-glucoamylase was tentative (Norén et al., 1986).

The major observation from this table is that relatively few residues within these putative transmembrane segments contain charged side chains (a low frequency of Lys, Arg, Asp and Glu). Even though these are tentative
Table 1: Amino acid content and charge contribution in putative transmembrane segments of various membrane proteins. The explanation of the table titles are as follows: "DNA Seq" denotes that the amino acid sequence was deduced solely from the DNA sequence of the protein (+) or that other topological data was also utilized in the assignment of the transmembrane segment (-). "AA Residues 'start' and 'total'" identifies the starting amino acid residue (numbered from the N-terminus) in the transmembrane segment and the total number of residues involved in spanning the membrane. The "Total Number of Charges" at pH 7 is subdivided into those found in the "native" protein prior to chemical modification of amine and hydroxyl groups, while "succinyl" and "trimesyl" imply that the corresponding protein has been completely derivatized with succinic anhydride or with activated trimesic acid. The proteins that are included in the table have been designated by an abbreviation and are identified as follows:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M6 prot</td>
<td>M6 protein (Hollingshead et al., 1986)</td>
</tr>
<tr>
<td>prot A</td>
<td>protein A (Hollingshead et al., 1986)</td>
</tr>
<tr>
<td>AGPR</td>
<td>rat asialoglycoprotein receptor (hepatic lectin) (Drickamer et al., 1984)</td>
</tr>
<tr>
<td>MGA</td>
<td>maltase-glucoamylase (Norén et al., 1986)</td>
</tr>
<tr>
<td>GP A</td>
<td>glycophorin A (Marchesi et al., 1976)</td>
</tr>
<tr>
<td>GP C</td>
<td>glycophorin C (Colin et al., 1986)</td>
</tr>
<tr>
<td>UDP-GT</td>
<td>UDP-glucuronosyltransferase (Mackenzie, 1986)</td>
</tr>
<tr>
<td>FNR (α)</td>
<td>fibronectin receptor (α-subunit) (Argraves et al., 1986)</td>
</tr>
<tr>
<td>M13 Coat</td>
<td>bacteriophage M13 coat protein (Nakashima and Konigsberg, 1974)</td>
</tr>
<tr>
<td>VSV G</td>
<td>glycoprotein G of VSV (Gallione and Rose, 1985)</td>
</tr>
<tr>
<td>ANT</td>
<td>adenine nucleotide translocater (Saraste and Walker, 1982)</td>
</tr>
<tr>
<td>BR</td>
<td>apobacteriorhodopsin (Abercrombie and Khorana, 1986)</td>
</tr>
<tr>
<td>cyt P450</td>
<td>liver cytochrome P450; isozyme 3b (Ozols et al., 1985)</td>
</tr>
<tr>
<td>ACHR(α)</td>
<td>nicotinic acetylcholine receptor (α-subunit) (Ratnam et al., 1986)</td>
</tr>
<tr>
<td>βAdrenR</td>
<td>avian β-adrenergic receptor (Yarden et al., 1986)</td>
</tr>
<tr>
<td>Lead Pep</td>
<td>E. Coli leader peptidase (Dalbey and Wickner, 1986)</td>
</tr>
<tr>
<td>GalHTrns</td>
<td>β-galactoside: H⁺ transporter (lac permease) (Foster et al., 1983)</td>
</tr>
<tr>
<td>CaMg ATP</td>
<td>Ca²⁺Mg²⁺ ATPase of sarcoplasmic reticulum (Brandl et al., 1986)</td>
</tr>
<tr>
<td>NaK ATP</td>
<td>Na⁺K⁺ ATPase (Shull et al., 1985)</td>
</tr>
<tr>
<td>GluTrns</td>
<td>glucose transporter (Mueckler et al., 1985)</td>
</tr>
<tr>
<td>Band 3</td>
<td>erythrocyte anion exchange protein (Kopito and Lodish, 1985)</td>
</tr>
<tr>
<td>Membrane Protein</td>
<td>DNA Seq</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>M6 prot</td>
<td>+ 24820</td>
</tr>
<tr>
<td>prot A</td>
<td>+ 48921</td>
</tr>
<tr>
<td>AGPR</td>
<td>+ 3927</td>
</tr>
<tr>
<td>MGA</td>
<td>+ 818</td>
</tr>
<tr>
<td>GP A</td>
<td>+ 7319</td>
</tr>
<tr>
<td>GP C</td>
<td>+ 5823</td>
</tr>
<tr>
<td>UDP-GT</td>
<td>+ 49517</td>
</tr>
<tr>
<td>FNR (α)</td>
<td>+ 17334</td>
</tr>
<tr>
<td>M13 Coat</td>
<td>- 2119</td>
</tr>
<tr>
<td>VSVG</td>
<td>+ 46320</td>
</tr>
<tr>
<td>ANT</td>
<td>- 1126</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>- 823</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>cyt P450</td>
<td>- 319</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ACHR(α)</td>
<td>+ 14211</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>β-AdrenR</td>
<td>+ 4424</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Membrane Protein</td>
<td>DNA Seq</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Lead Pep</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>GalHTns</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>CaMg ATP</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>NaK ATP</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>GluTrns</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Band 3</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
assignments and the basis of their selection is for stretches of amino acid residues that are hydrophobic and uncharged, it is these types of segments that self aggregate in aqueous solutions, some aggregate even in the presence of detergents and chaotropic agents (Ozols et al., 1985). The hydroxyl amino acids (Tyr, Ser and Thr) are present in 95 of the 98 putative transmembrane segments and generally at a much higher frequency than the charged residues. These hydroxyl residues alone do not improve the solubility of these transmembrane segments, as is observed by the general behaviour of proteolytic fragments of membrane proteins (Jay and Cantley, 1986). However, if these hydroxyl amino acid residues could be chemically modified with a reagent that incorporates a charge at each site (such as one negative charge per site upon succinylation or two negative charges per site upon trimesylation), this would greatly increase the overall charge density of the hydrophobic transmembrane segments thereby enhancing their water solubility.

I6. Chemical Derivatization of Membrane Proteins

The objective of this research is to develop methods of extensive chemical derivatization for membrane proteins as a means to render them water soluble in the absence of detergents and chaotropic agents. This conversion of membrane proteins into water soluble ones should facilitate their enzymatic fragmentation and greatly improve the behaviour of the resulting fragments during Edman degradation in automated amino acid sequencers.

During the 1950's and 1960's, it was established that proteins were capable of undergoing chemical derivatization which resulted in an alteration of their conformation and/or enzyme activity (Hass, 1964, Gounaris and
Perlmann, 1967, Freedman et al., 1968 and Hartley, 1970). The possible reactive sites on a protein include the amino and carboxyl termini as well as the functional groups present on the side chains of the amino acid residues. In general, the types of modifications performed in these studies were relatively specific and few in number and the studies were not concerned with the solubility of the proteins and their fragments.

Denaturation of water soluble proteins to ensure complete chemical or enzymatic fragmentation can result in an insolubility problem due to the association of the exposed hydrophobic domains, causing aggregation and precipitation. Conversion of the positive charge of the lysine residues to a negative one via acylation of the ε-amino group with succinic anhydride (Klotz, 1967), substantially changed the overall charge density of the denatured protein and its fragments to render them water soluble (Klapper and Klotz, 1972). Furthermore, the derivatized lysine residues, now negatively charged, would electrostatically repel each other as well as aspartate and glutamate residues, thereby denaturing the protein and permitting more efficient cleavage by enzymes (Klotz, 1967).

The poor solubility and aggregation of membrane proteins even in the presence of detergents and chaotropic agents has made difficult the assessment of the monomeric molecular weight of some proteins. As an example of extensive chemical derivatization of membrane proteins, a novel method was developed to more accurately assess the molecular weight of these types of proteins through the dinitrophenylation of the N-terminus followed by permethylation of the hydrophilic groups in the side chain of the appropriate amino acid residues thereby rendering them more hydrophobic. This modification enables the protein to be soluble in CHCl₃/ MeOH (4:1) and then be accurately sized on Sepharose CL-6B (Meredith, S.C., 1984).
In the attempt to improve membrane protein fragment behaviour the amino group at the N-terminus and/or the ε-amino groups of the lysine residues were covalently modified with a variety of derivatizing reagents. This was generally performed in order to decrease the number of tryptic peptides; due to the specificity of trypsin, cleavage only occurs at the unmodified arginine residues (Klotz, 1967). The derivatization was also attempted to increase the overall negative charge density of the protein and its fragments, in the hope to enhance their water solubility.

Succinylation of membrane proteins has been attempted under aqueous conditions using a slightly modified procedure of Klotz (1967); a 200 or a 1000 fold excess of succinic anhydride was used over the free amino groups of the protein, in the presence of 6 or 7 M guanidinium hydrochloride (Boulay and Vignais, 1984 and Ozols et al., 1985, respectively). The large molar excess was required solely to modify the lysine residues. In general, succinylation was found to enhance the solubility of the denatured membrane protein, however, succinylation of the lysine residues was insufficient to improve the water solubility of the hydrophobic fragments due to the low frequency or absence of this residue from the transmembrane regions; the hydrophobic fragments still aggregated (Khorana et al., 1979, Ozols et al., 1985, Reddy et al., 1986 and Ozols, 1986). Furthermore, not all succinylated membrane proteins are rendered water soluble; succinylated cytochrome P450 was found to precipitate when dialyzed against 0.1 M ammonium acetate, pH 8.1 (Ozols et al., 1985).

Citraconylation of the lysine residues has also been utilized as a means to reduce the number of tryptic fragments and partially improve protein fragment behaviour (Aquila et al., 1982, Kawabata et al., 1986 and Ozols, 1986). However, this reagent is also limited to reacting with the ε-
amino groups of lysine residues when performed under aqueous conditions and thus, is not likely to substantially improve the water solubility of transmembrane segments.

To improve the solubility of the very hydrophobic N-terminal fragment of bacteriorhodopsin, 4-sulfophenylisothiocyanate groups were attached to the C-terminal carboxyl group (Gerber et al., 1979) prior to its derivatization with succinic anhydride. This method, though successful in the generation of water-soluble products, may not be generally useful due to the paucity of the sites of succinylation.

Therefore, the objective was to modify a sufficient number of amino acid residues on the protein so as to afford the protein a substantially increased charge density that would render the membrane protein and its fragments water soluble. The hydroxyl groups of the residues tyrosine, serine and threonine, due to their relatively high frequency within the putative transmembrane segments (Table 1), are the desired sites of modification. The resulting incorporation of one or two charges at each site of derivatization would substantially increase the charge density of these hydrophobic segments. The protein after extensive derivatization would probably be in a denatured conformation and thus be more amenable to enzymatic and chemical cleavage. The resulting fragments would be water soluble and more manageable with respect to their isolation and automated sequencing.

Developing such a method for improving the hydrophilicity of these transmembrane segments would permit more rapid topological studies. This in turn would improve the rate at which evidence is gathered to prove or disprove topological models and assess the involvement of structural features that are purported to be associated with the proteins' function. There was a hypothesis that suggested that the structure function relationship of
proline residues within the transmembrane segments of membrane proteins was for the transport of ions or molecules (Brandl and Deber, 1986). However, amino acid substitutions of the membrane-embedded proline residues of bacteriorhodopsin did not support this proposal (Mogi et al., 1989). This proposed method of enhancing the water solubility of hydrophobic proteins in conjunction with standard topological and improved sequencing techniques would greatly aid the study of membrane proteins.

Furthermore, the proposed extensive chemical derivatization of membraneous or water soluble proteins, with a reagent that would incorporate a negative charge at each site of modification, as stated, would result in the denaturation of the protein due to charge repulsion. The increased charge density would enable the protein to be water soluble, even in this denatured state. In this form the N-terminus of the protein would be readily accessible by such N-terminal deblocking enzymes as the N-acylpeptide hydrolase or pyroglutamate aminopeptidase; the identification of the type of blocking group present on the protein could be assessed by mass spectrometry. Inaccessibility of a protein's termini to enzymes was observed for pepsinogen; carboxypeptidase A is not able to cleave its terminal amino acid residue in its native conformation but upon succinylation the C-terminus was accessible to the enzyme (Gounaris and Perlmann, 1967). Therefore, the water solubilization of a protein in a denatured conformation would improve the efficiency of deblocking the protein, enabling it to be directly sequenced from its N-terminus, starting at the second residue. This procedure would eliminate the need to either generate a N-terminal fragment of a protein or to sufficiently denature the protein in order to ensure that the N-terminal residue is accessible to the enzyme. In light of the fairly broad specificity of the deblocking enzymes, it is anticipated that they
would still hydrolyze the blocked N-terminal residue on a modified protein
even if one of the sites of derivatization was on the first residue, in the case of
N-acylpeptide hydrolases or on the penultimate amino acid residue, in the
case of pyroglutamyl aminopeptidase.

I.7. Chemical Reagent Requirements

The objective of chemically converting membrane proteins and
their possible fragments into water soluble derivatives can be achieved if the
derivatizing reagent meets certain requirements. The first consideration was
for a reagent that is capable of converting the positive charge of the ε-amino
group of lysine into a negative one. There are several chemical reagents that
can achieve this type of conversion and they include β-sulfopropionyl
chloride, 3,3-tetramethylene glutaric anhydride, maleic anhydride, citraconic
anhydride, tetrafluorosuccinimide anhydride, succinic anhydride, 4-fluoro-3-
nitrobenzenesulfonate, and 4-sulfophenylisocyanate (Vallee and Riordan,
1969 and Stark, 1970). However, as previously stated, derivatization of only
lysine residues is insufficient to maintain the solubility of membrane
proteins due to their relatively low frequency, especially in the
transmembrane segments. Therefore, the chemical reagent must also be
capable of reacting with other functional sites on the protein, namely the
hydroxyl groups of the tyrosine, serine and threonine residues.

The possible list of reagents is further reduced by the necessity of
using one that employs a carboxylate group for the incorporation of the
negative charge. This requirement is due to the chemistry involved during
the Edman degradation in automated sequencers; the 2-anilino-5-
thiazolinone amino acid derivatives are extracted from the protein film with
1-chlorobutane. This organic solvent extraction occurs fairly efficiently for all
the amino acids, even for Asp and Glu which bear an additional carboxyl group on their side chain because the carboxyl group is sufficiently nonpolar when protonated. However, when the anilino-thiazolinones contain a phosphate or sulfate group, less than 1% of the product is recovered during the extraction cycle due to their substantial polarity (Tremblay et al., 1989). Therefore, the additional charges incorporated on a protein must be in the form of a carboxyl group which will not adversely affect the sequencing of the protein.

Of the aforementioned reagents, the most suitable appears to be succinic anhydride due to its strong preference for $\alpha$ and $\varepsilon$ amino groups and its capability of succinylating sulfhydryl, imidazole and aliphatic or aromatic hydroxyl groups. The types of chemical bonds that would be formed during succinylation would include: 1. succinyl amide bonds, though stable, are reversible upon acid hydrolysis (6 N HCl, 100°C). 2. succinyl thioesters and 3. succinyl imidazolide derivative of histidines both of which are expected to hydrolyze over time. 4. succinyl ester bonds which are relatively stable except they are acid labile and can be cleaved with alkaline hydroxylamine. An exception is the succinyl ester on tyrosine, it apparently spontaneously decomposes, even at neutral pH (Gounaris and Perlmann, 1967).

An additional feature of succinic anhydride is that it is already reactive and requires no form of activation; it readily undergoes nucleophilic attack on one of its carbonyl groups. Therefore, under the appropriate conditions, the derivatization of a protein with succinic anhydride would be extensive, resulting in the incorporation of a large number of negative charges which would increase the overall negative charge density and thus render any membrane protein water soluble.

Succinylation is a well characterized reaction and has been used for
protein modification for both water soluble and membranous proteins. Derivatization of water soluble proteins, such as pepsinogen (Gounaris and Perlmann, 1967), aldolase (Hass, 1964 and Meighan and Schachman, 1970) and bovine serum albumin (Meighan and Schachman, 1970 and Klapper and Klotz, 1972) was performed in basic aqueous solvent systems (pH 7 - 10). These proteins were observed to undergo alterations in their conformation and a decrease in their enzyme activity upon succinylation (Hass, 1964, Gounaris and Perlmann, 1967, Freedman et al., 1968 and Hartley, 1970). Furthermore, enzymatic cleavage of the modified protein by trypsin was now restricted to the underivatized arginine residues (Klotz, 1967).

The derivatization of membrane proteins, such as cytochrome P450 (Ozols et al., 1985) and adenine nucleotide translocater (Boulay and Vignais, 1984) was also performed under aqueous conditions but in the presence of guanidinium hydrochloride to aid the solubility of the protein. The derivatization was to result in an improvement in the solubility of the protein and its fragments, as well as reduce the number of tryptic fragments for sequencing considerations. However, there was limited success with respect to the solubilization (Ozols et al., 1985), probably due to the paucity of lysine residues, as previously discussed, and to incomplete derivatization as a result of protein aggregation and initial insolubility in the aqueous solvent system.

The utilization of an aqueous solvent system for extensive derivatization of proteins by succinic anhydride is clearly inappropriate from the following observations: 1. A fifteen fold molar excess of succinic anhydride per free amino group was required in order to obtain adequate modification of even the free amino groups in a water soluble protein (Gounaris and Perlmann, 1967). The primary reaction was probably reagent
hydrolysis. 2. Succinylation of the less nucleophilic hydroxyl groups would require greater amounts of the succinic anhydride. Under these aqueous reaction conditions the complete acylation of the hydroxyl groups would be difficult to achieve, especially for the secondary hydroxyl group of threonine residues. 3. Membrane proteins are generally insoluble in aqueous solvents; even in the presence of detergents and chaotropic agents these proteins have the potential for aggregation resulting in site inaccessibility and therefore, incomplete derivatization. Therefore, if the succinylation reaction was performed in a suitable anhydrous solvent then reagent hydrolysis would be minimized while the desired modification reaction of both amines and hydroxyl groups would be much more efficient. Furthermore, membrane proteins would have an enhanced solubility thereby providing a better substrate for acylation.

A possible problem with succinic anhydride, though as of yet unsubstantiated, is the stability of the succinyl esters. The spontaneous loss of the succinyl group from tyrosine residues (Gounaris and Perlmann, 1967) and the general stability of the 5 membered ring of the cyclic anhydride, implicates the participation of the free carboxylate group in a nucleophilic attack on the carbonyl group of the succinyl ester resulting in the recyclization of the succinic anhydride and the concommitant regeneration of the tyrosine hydroxyl group. If this is true then the same mechanism could occur on succinyl derivatives of the aliphatic hydroxylamino acids, though presumably at a slower rate. The succinyl ester of the aliphatic hydroxylamino acids was shown to be stable by succinylating free serine or threonine with succininc anhydride and quantitating the amount of succinyl ester present (alkaline hydroxylamine ester assay), before and after a 24 hour incubation in sodium phosphate buffer, pH 8 (Gounaris and Perlmann, 1967). In a similar manner,
the stability of the succinyl esters on the derivatized pepsinogen were assessed and shown to be stable in the phosphate buffer at pH 8, after 24 hours. However, only 15 - 20% of the serine and threonine residues had been derivatized (Gounaris and Perlmann, 1967). Whether this incompletely derivatized protein presents a true picture of the stability of O-succinylated amino acid residues is uncertain. Further investigation is needed to more accurately assess the stability of these succinyl esters.

1.8. Model Systems
1.8.A. Model Peptides

An appropriate model system was required for the assessment of the rate of acylation by the designated reagent. This model system must include hydroxylamino acids, those residues that would be targeted for derivatization in proteins. Therefore, each of the hydroxylamino acids tyrosine, serine and threonine were separately incorporated. It was reasoned that if the rate of reaction with hydroxyl groups is optimized then the reaction with the ε-amino group of Lys residues should also be maximal. Furthermore, the attachment of a chromophore to these amino acids was included in the design of the model system to simplify the characterization of the acylation reaction; reversed phase HPLC could be utilized for the analysis of the reaction mixture, at a wavelength specific to the chromophore so that only the peptide and its derivatives would be detected. The chromophore that was chosen was the dinitrophenyl group (Sanger's derivative; DNP) which has a extinction coefficient of approximately 10,000 at 365 nm (Sanger, 1949 and Sanger and Tuppy, 1951). However, the proximity of DNP group and of any other charged or bulky groups on the model peptide must be considered with respect to the position of the hydroxyl group on the side
chain of the amino acid in order for the model peptide to accurately represent
the hydroxylamino acid residues within a protein.

The derivatization reaction should be optimized with the peptide
containing the Thr residue because it contains a secondary hydroxyl group
which will be the slowest to react. However, the stability of the ester product
must be assessed by the aromatic ester generated on the Tyr residue as it will
be the most sensitive to nucleophilic attack.

I.8.B. Model Protein

Bacteriorhodopsin was selected as the model protein for study for
the following reasons: 1. It is a very hydrophobic intrinsic membrane protein
whose polypeptide backbone spans the lipid bilayer at least seven times. 2.
The purple membrane of Halobacteria halobium is commercially available
(Sigma Chemical Company) providing a ready source of the protein. 3. The
apoprotein can be generated, delipidated and isolated in a fairly simple
procedure (Gerber et al., 1977). 4. The amino acid sequence of the
apobacteriorhodopsin is known (Khorana et al., 1979; Figure 8). 5. It is one of
very few membrane proteins that have actually been sequenced using
standard protein sequencing techniques in conjunction with novel
procedures that were developed to manage this protein, especially for dealing
with the insolubility and the aggregation of the protein's hydrophobic
fragments (Gerber and Khorana, 1982). Therefore, apobacteriorhodopsin will
be a good model to study the efficiency of extensive chemical derivatization
and the behaviour of the acylated product with respect to its water solubility,
its suitability as a substrate for enzymes and ultimately its behaviour in
automated protein sequencers.
I.9. Rationale and Objective:

Integral membrane proteins mediate the various functions of biological membranes. It is of great interest, therefore, to study the structure function relationship of these proteins to better understand their mechanism of action. The primary sequence of many membrane proteins are being readily deduced from DNA sequencing procedures. However, in order to confirm the amino acid sequence, to assess protein biosynthesis, with respect to post-translational modifications and to accurately determine their topology within the membrane, the protein and its fragments must be sequenced.

The standard procedures of protein fragmentation, fragment purification and the actual sequencing of the isolated fragments are all plagued by the membrane protein's inherent lack of water solubility and the problems associated therewith. Acylation of the lysine residues with succinic anhydride, to increase the overall negative charge density, renders denatured water soluble proteins and their fragments water soluble. However, this extent of derivatization is generally insufficient to solubilize intrinsic membrane proteins or their hydrophobic fragments. Transmembrane segments consist mainly of hydrophobic amino acid residues with a low frequency of lysine or any other charged residue; hydrophobic interactions between such segments accounts for the aggregation of membrane proteins or protein fragments. Hydroxylamino acids are present in transmembrane segments at a much higher frequency and incorporation of a negative charge at these sites would substantially increase the charge density which would render membrane proteins and their fragments water soluble.

Therefore, the objective of this research is to develop methods of extensive chemical derivatization for membrane proteins as a means to render them water soluble in the absence of detergents and chaotrophic agents.
Depending on the selected reagent, there would be an attachment of one or more carboxylate groups at each site of derivatization. The attached carboxylate groups would enhance the overall charge density of the protein and its possible fragments. To achieve efficient acylation of even the secondary hydroxyl groups of threonine residues, the reaction must be performed in a solvent that is chemically inert, anhydrous and one that supports membrane protein solubility. This conversion of hydrophobic membrane proteins into water soluble ones should facilitate their enzymatic fragmentation, fragment isolation and greatly improve their behaviour during automated sequencing.

To achieve the objective of this thesis, the following specific goals had to be met: 1. The solvent had to be selected and evaluated on the basis of protein solubility and reagent reactivity. 2. A model system had to be designed and synthesized for characterizing the acylation reaction on the hydroxylamino acids of interest. 3. The acylation reagent had to be considered and designed to meet the specifications of reacting with hydroxyl groups as well as amino groups, of forming a stable derivative and resulting in the incorporation of at least one carboxylate group per site of modification. 4. The reagent then had to be activated and evaluated on the basis of its stability and reactivity towards the model substrates. 5. The stability of the resulting ester for all of the hydroxylamino acid model substrates were assessed. 6. The derivatization of a membrane protein had to be assessed on the basis of complete acylation of the desired functional groups. The acylated membrane protein could then be assessed with respect to its water solubility by gel permeation, its electrophoretic mobility on PAGE systems, its suitability as a substrate for enzymes, and its behaviour under sequencing conditions.
II. Materials and Methods

II.1. Chemicals and Solvents

Chemicals that were purchased from Sigma included trimesic acid (1,3,5-benzenetricarboxylic acid), 1,1'-carbonyldiimidazole (CDI), dicyclohexylcarbodiimide (DCC), 4-dimethylaminopyridine (DMAP), 4-pyrrolidinopyridine (PPY), 1-fluoro-2,4-dinitrobenzene (FDNB), L-proline and the hydrochloride salts of L-serinamide, L-threoninamide, L-glycyltyrosinamide and L-glycinamide. Tetrabutylammonium hydroxide (Bu₄N⁺OH⁻, 1.54 M aqueous solution) and trimethylsilylethanol (Me₃SiCH₂CH₂-OH) were obtained from Aldrich. Triethylamine (Et₃N) and trifluoracetic acid (TFA) were purchased from Pierce. The Et₃N was distilled from FDNB under nitrogen to remove any primary or secondary amine contaminants; storage of the distillate was at 4°C under nitrogen, over Fisher 4 Å molecular sieve and protected from light. Sephadex G-150, LH-20 and LH-60 were obtained from Pharmacia. All other chemicals were reagent grade.

HPLC grade methanol, tetrahydrofuran (THF), dimethylformamide (DMF), dimethylsulfoxide (Me₂SO), benzene, toluene, petroleum ether and diethylether were purchased from Caledon. Anhydrous DMF was made by a modified procedure of Thomas and Rochow (1957); DMF was initially dried by the addition and subsequent removal under reduced pressure of 0.1 volume of benzene followed by storage over 4 Å molecular sieve and under nitrogen. Dry benzene was obtained by distillation from calcium hydride and was stored under nitrogen and over molecular sieve (4 Å). Pyridine, obtained from Baker, was incubated overnight with KOH pellets, decanted and then
distilled from barium oxide under nitrogen. The anhydrous pyridine was then stored under nitrogen and over 4 Å molecular sieve in amber bottles. Carbontetrachloride (CCl₄), purchased from Fisher, was distilled from phosphorous pentoxide under nitrogen. All other solvents were reagent grade. The petroleum ether fraction used was the one boiling between 30-60°C.

The purple membrane of *Halobacterium halobium* cells was purchased from Sigma.

II.2. High Pressure Liquid Chromatography (HPLC)

The HPLC system (Millipore-Waters) consisted of two M-60 pumps, a 720 system controller, a M 730 data module, a WISP 710B automatic injector and a 440 absorbance detector with the appropriate filter for 365 nm, 280 nm or 254 nm detection. The columns that were routinely used for the analysis of DNP-peptides were a µBondapak ODS (4.6 x 250 mm) and an Altex Ultrasphere (4.6 x 150 mm) both of which were comprised of 10 µm C₁₈ bonded particles that had been end capped after the column was packed. The elution buffers were Caledon methanol (HPLC grade) and BDH sodium acetate dissolved in glass distilled water; the pH was adjusted with BDH glacial acetic acid.

II.3. Alkaline Hydroxylamine Ester Assay

This assay that was developed by Hestrin (1949) quantitatively determines the amount of ester present in a sample. Two volumes of the alkaline hydroxylamine reagent, which consists of equal parts of a 2 M NH₂OH·HCl and a 3.5 M NaOH solution, are mixed with one volume of the sample and are allowed to react for 5 minute at room temperature. To this
reaction, one volume of a HCl solution was immediately added, that was prepared by diluting concentrated HCl with two equal parts of water. After an additional 5 minutes incubation at room temperature, the reaction was centrifuged in a table top centrifuge (3,000 rpm) in order to pellet any protein precipitate. One volume of a 0.37 M FeCl₃ solution in 0.1 N HCl was thoroughly mixed with the supernatant and the absorbance of the solution at 540 nm was measured after one minute. The modification introduced in this procedure was to decrease the incubation time with the alkaline hydroxylamine reagent from 5 to 1 minute; this was found to reduce the background absorbance. When ethylacetate was treated in this manner one micro.mole of ester was found to yield an absorbance of 0.465.

II.4. Tritium Exchange on Trimesic Acid

The trimesic acid (660 mg) was placed into the bottom of an 18 x 150 mm Pyrex test tube and the tube was then drawn out. The tritiated water (0.50 mL containing 2.5 Ci) was then added and rinsed in with 50 mL of 12 N HCl. The contents of the tube were then frozen in liquid nitrogen, the tube was evacuated, purged with nitrogen and then sealed. The sealed tube was placed in a pressure device containing water to equalize the pressure and heated at 275°C for 31 days. After cooling, the tube was opened, the [³H]H₂O removed by evaporation and the dry residue recrystallized twice from 15 mL of boiling water. The specific activity of the product was 5 mCi/mmol and it was found to co-elute with authentic trimesic acid on HPLC analysis.

II.5. Synthesis of DT-TM Acid (Figure 11; IV)

All steps in the synthesis of the diprotected trimesic acid were performed in a closed vessel under nitrogen and at room temperature. The
reactive triimidazolide derivative of trimesic acid (II) was generated by the addition of 4.8 mmol of trimesic acid (I) in 6 mL of anhydrous DMF, in approximately 1 mL aliquots, to a screw cap tube containing 2.33 g CDI (3 equiv.) and 4 mL of dry DMF. During the addition, the mixture was continuously vortexed. After 1 hour, two equivalents of Me₃SiCH₂CH₂-OH were added to the activated trimesic acid (II) and the subsequent esterification reaction was allowed to proceed overnight. Following a dilution with an equal volume of dry DMF, the remaining imidazolides were hydrolyzed by the addition of 2 mL of 2.5 M aqueous NaOH. Under these basic conditions, three petroleum ether extractions afforded the complete removal of the triester (VI) contaminant. The aqueous solution was then acidified with HCl to pH 1 allowing the desired product, DT-TM acid, to be selectively extracted into diethyl ether/ petroleum ether (1:9). The monoester side product remained in the aqueous phase and could only be extracted with 100% diethyl ether. The extracts containing DT-TM acid were pooled, washed with 0.1 N HCl and then once with 4 M NaCl. The solvent was evaporated under a stream of nitrogen and the product was repeatedly (3 times) dissolved in dry benzene and the solvent removed under reduced pressure. The dried DT-TM acid was stored as a 0.2 M solution in dry benzene, over molecular sieve (4 Å) and under nitrogen.

II.6. Formation of DT-TM Anhydride

The reactive anhydride derivative of DT-TM acid was generated as a 0.2 M solution at room temperature by the addition of 1.0 equivalent of DCC to 2.25 equivalents of DT-TM acid in dry CCl₄. After 1 hour, this mixture was diluted to 50 mM with anhydrous benzene and the dicyclohexylurea precipitate was removed by filtration through glass wool. This afforded a
stock solution of the activated reagent from which appropriate aliquots could be used for trimesylation, after the removal of the benzene under reduced pressure.

II.7. DNP-Proline Synthesis

During the synthesis and subsequent storage of all DNP containing products, care was taken to protect these compounds from prolonged exposure to light. L-proline (150 μmol) was dissolved in ethanol/water (3:1, 2 mL) in the presence of 2 equivalents of Et₃N; the dinitrophenylation reaction was initiated by the addition of FDNB (1.2 equiv.). After three hours, at room temperature, the reaction was diluted with two volumes of aqueous 10 mM sodium phosphate (dibasic), pH 7. Following the removal of the ethanol under reduced pressure, the excess FDNB was selectively extracted into diethyl ether. Subsequent acidification of the aqueous phase to pH 1, by titration with 6 N HCl, permitted the extraction of DNP-proline into ethyl acetate. The extracts were washed with water and then evaporated under reduced pressure. Residual water was removed by drying the DNP-proline several times from a mixture of anhydrous benzene and DMF. DNP-proline was stored as a 0.25 M solution in anhydrous DMF over a molecular sieve (4 Å) and under nitrogen.

II.8. DNP-Peptide Synthesis

II.8.A. DNP-prolylserinamid and DNP-prolylthreoninamid

DNP-proline was activated by the addition of 1.1 equivalents of CDI in dry THF as a 60 mM solution and incubated at room temperature for 1 hour under nitrogen. To the reactive imidazolidone derivative, 3 equivalents of the hydrochloride salt of serinamide or threoninamide was added as a
freshly prepared 1.54 M aqueous solution in the presence of 1 equivalent of 
Bu$_4$N$^+$OH$^-$.

II.8.B. DNP-prolylglycinamide and DNP-prolylglyclytyrosinamide

The formation of DNP-prolyl-imidazolide was carried out under the same conditions as outlined above except that the solvent was dry DMF and the concentration was 80 mM. At the appropriate time, 3 equivalents of the hydrochloride salt of glycinamide or glyclytyrosinamide was added as an anhydrous 80 mM solution in DMF, in the presence of 2 equivalents of Et$_3$N.

After 12 hours, each peptide reaction was diluted with a 0.5 volume of water and the desired DNP-peptide extracted into ethyl acetate. Isolation and purification of the peptide products was accomplished by preparative HPLC on a reversed phase μBondapak C$_{18}$ column which was equilibrated and run isocratically with 1% acetic acid and the following specified percentages of methanol: DNP-prolylglycinamide and DNP-prolylserinamide required 10% methanol while DNP-prolylthreoninamide and DNP-prolylglyclytyrosinamide required 15% and 20% methanol, respectively. The DNP-peptides were dried several times from anhydrous pyridine and then stored as a 25 mM solution in dry pyridine over molecular sieve and under nitrogen.

II.9. Formation of Tetrabutylammonium Salt

Tetrabutylammonium trifluoroacetate (Bu$_4$N$^+$TFA$^-$) was made by the addition of 1.1 equivalents of TFA to 1.0 equivalent of tetrabutylammonium hydroxide. The excess acid was removed under reduced pressure and the resulting amphipathic salt was repeatedly (3 times) dissolved in benzene and the solvent was removed under reduced pressure.
The resulting residue was then stored as a 1 M solution in absolute ethanol under nitrogen.

II.10. Standard Trimesylation

The DNP-peptides were trimesylated under nitrogen as a 0.5 mM solution in dry pyridine in the presence of 50 mM activated reagent (DT-TM anhydride), 100 mM PPY and 10 mM Bu₄N⁺TFA⁻. The addition of the DNP-peptide of choice initiated the reaction and at specific times an aliquot of the reaction was quenched in 10% H₂O and 0.1% Et₃N in DMF. The extent of peptide acylation was determined by HPLC analysis of a portion of the quenched reaction which was applied to a µBondapak C₁₈ reversed phase column. Employing the HPLC conditions specified in Figure 22, the following retention times for the DNP-peptides and their corresponding derivatives were obtained; 5.0 and 15.9 minutes for the serine dipeptide, 6.4 and 16.2 minutes for the threonine dipeptide and 7.8 and 17.5 minutes for the tyrosine tripeptide.

Following complete DNP-peptide derivatization, the excess reagent was deactivated by the addition of Me₃SiCH₂CH₂OH to a final concentration of 0.1 M while remaining under nitrogen and at room temperature. The alcoholic reagent quench was allowed to proceed for 0.5 hour. The carboxyl groups were then deblocked by the addition of at least 2 equivalents of Bu₄N⁺F⁻ for each 2-(trimethylsilyl)ethyl group present (Lipshutz and Pegram, 1980). This deprotection reaction was essentially instantaneous.

II.11. Preparation of Apobacteriorhodopsin

The purple membrane of *Halobacterium halobium* cells was suspended (2 mg/mL) in 0.5 mL Me₂SO and 0.5 mL 4.0 M NaCl containing 1
M NH₂OH·HCl (pH 7). The suspension was stirred for 1 hour at 37°C to ensure complete disappearance of the purple colour (Oesterheld et al. 1973 and Gerber et al., 1977). The apomembrane was diluted 8 fold with H₂O, pelleted by a 100,000 x g centrifugation for one hour, resuspended in 1 mL distilled H₂O and lyophilized. The residue was redissolved in 88% formic acid and applied to a Sephadex LH-20 column (1 x 40 cm) that was equilibrated and run in formic acid/ethanol (30/70) to achieve the delipidation of the apobacteriorhodopsin. Fractions (0.5 mL) were collected under nitrogen and the void volume peak was pooled and stored at -20°C.

II.12. Solubilization of Apobacteriorhodopsin in Anhydrous Pyridine

The delipidated protein was dried under reduced pressure in the presence of Bu₄N⁺TFA⁻ (100 μmoles/mg of protein) which was added as a 1 M solution in absolute ethanol and after the addition of two volumes of dry toluene and absolute ethanol (2:1, v/v). The protein/salt residue was then twice redissolved and redried in dry toluene and absolute ethanol (2:1, v/v). The resulting residue was dissolved in 200 μL dry pyridine which was subsequently removed under reduced pressure. The resulting residue was redissolved in 50 μL pyridine.

II.13. Polyacrylamide Gel Electrophoresis

II.13.A. Sodium Dodecylsulfate

The standard system of Laemmli (1970) was employed for the running of a SDS polyacrylamide gel at pH 6.8. The SDS-PAGE system containing 8 M urea at pH 6.8 of Swank and Munkres (1971) was also utilized.
II.13.B. Acidic Lithium Dodecylsulfate

The electrophoretic systems of Jones et al. (1981) and Lichtner and Wolf (1979) were utilized for the development of acidic polyacrylamide gels at pH 4 and 2, respectively. One modification of the acidic (pH 2) gel was the inclusion of 8 M urea, according to Swank and Munkres (1971).

II.14. Amino Acid Analysis

II.14.A. Laboratory Procedure

Protein quantitation and recovery was assessed by hydrolyzing a portion of the sample in constant boiling HCl for 12 hours at 115°C. The hydrolysate was then dried and then redried from aqueous Et$_3$N (EtOH/H$_2$O/Et$_3$N (2:2:1)). The amino acid residue was then derivatized with the PITC reagent (EtOH/H$_2$O/Et$_3$N/PITC (7:1:1:1)) for 30 min at room temperature, according to the "PicoTag" method of Millipore-Waters (Heinrikson and Meredith, 1984). The resulting phenylthiocarbamyl derivatives were quantitated by HPLC at 254 nm by utilizing their specified solvent system: A solution was prepared containing 19 g sodium acetate trihydrate dissolved in 1 L of water to which 0.5 mL Et$_3$N was added and then the pH was adjusted to 6.4 with glacial acetic acid. Solvent A then consisted of 940 mL of this solution mixed with 60 mL CH$_3$CN and then filtered through a 0.45 μm Millipore-Waters filter. Solvent B consisted of 600 mL CH$_3$CN and 400 mL H$_2$O that was mixed and degassed by filtering through a 0.45 μm Millipore-Waters filter. The following modified gradient program gave the best separation of the phenylthiocarbamoyl derivatives of the amino acids on a Novopak C$_{18}$ column that was operated at 38°C: The column was initially equilibrated in solvent A at 1 mL/min. 1.5 min after sample injection the
proportion of solvent B was increased linearly such that at time 11.5 minutes the HPLC was pumping 54% solvent A and 46% solvent B. Within another 0.5 min the percentage of solvent B was then 100%.

II.14.B. Protein Services Procedure

The service at the University of Toronto generally uses the HCl gas phase procedure for 24 hours to achieve the hydrolysis of protein samples. The trimesylated protein samples required HCl/propionic acid (50:50, v/v) for complete hydrolysis. Ion exchange chromatography was then employed for the separation of the amino acids followed by their detection with ninhydrin.

II.15. Protein Sequencing

Protein samples were sent to the Amino Acid and Protein Services Department at the University of Toronto and they were sequenced on a gas-liquid phase sequenator that was operated by M. Blum. The samples were applied to glass fibre filter discs that had been previously treated with Polybrene. The sample was dried and then sequenced automatically. The 2-anilino-5-thiazolinone amino acid derivatives were directly converted to the PTH derivative and injected in-line onto a HPLC reversed phase column for identification of the PTH amino acid.
III. Results and Discussion:

III.1. Apobacteriorhodopsin Solubilization

In the absence of detergents and chaotropic agents, membrane proteins are generally insoluble in aqueous systems. This property makes their manipulation and processing through the standard procedures for protein fragmentation and amino acid sequence determination difficult. Extensive chemical derivatization of membrane proteins may afford an enhancement in their water solubility. However, complete acylation of hydroxylamino acid residues is not readily achieved under aqueous conditions; this is especially true for the secondary hydroxyl groups of threonine residues. Even in the presence of 100 molar excess of reagent, the threonine residues of the water soluble protein pepsinogen were poorly derivatized (Gounaris and Perlman, 1967). The utilization of detergents and chaotropic agents are not always sufficient to achieve complete solubility of some membrane proteins or highly amphiphilic proteins in aqueous systems. Protein aggregation and detergent micelles would result in the protection of the desired sites of derivatization. Furthermore, the activated reagent would be consumed by hydrolysis. An appropriate anhydrous solvent system is therefore required.

Membrane proteins can be solubilized in anhydrous TFA and ethanolic formic acid (30%, v/v) (Khorana et al., 1979); therefore, under the appropriate conditions, these types of proteins should be soluble in an anhydrous aprotic organic solvent such as DMF, Me₂SO or pyridine. Such a solvent system would enhance acylation reactions, thereby ensuring complete
derivatization of all possible reactive sites on the side chains of the amino acid residues in the protein.

The first approach that was studied employed the solubilization capabilities of anhydrous TFA (Gerber and Khorana, 1982). Delipidated apobacteriorhodopsin, that had been prepared and stored at -20°C in 88% formic acid and ethanol (3:7, v/v), as outlined in Materials and Methods, was dried under reduced pressure and the resulting protein residue was found to be readily soluble in anhydrous TFA at concentrations of 60 - 100 mg/mL. This protein solution was then diluted with dry DMF, to a concentration of approximately 5 mg/mL. The protein remained in solution provided the DMF was added in small aliquots while continually vortexing the protein solution. This acidic solution was then neutralized and then made basic by the addition of 1.5 equivalents of Et$_3$N (pK$_a$ 11) with respect to the TFA. (Under anhydrous conditions, the concept of pH, though not strictly applicable, is used with reference to the degree of ionization of particular functional groups of interest.) The addition of the anhydrous base in small aliquots with continual mixing generally ensured complete solubility of the protein. Occasionally at this stage, a portion of the apobacteriorhodopsin would come out of solution; however, it was always found to redissolve after all of the Et$_3$N was added and the solution was subjected to brief sonication in a water bath type sonicator. This protein solubilization procedure was not found to be applicable with Me$_2$SO or pyridine as the solvent; upon dilution of the anhydrous TFA protein solution, with either Me$_2$SO or pyridine, virtually all of the apobacteriorhodopsin precipitated out of solution and could not be solubilized even in the presence of Et$_3$N and after extensive sonication.
The suitability of this solvent system which achieved membrane protein solubilization was assessed with respect to the efficiency of acylation of hydroxyl groups. However, the presence of Et$_3$NH$^+$TFA$^-$ at a concentration of 1.35 M was found to severely reduce the rate of acylation of the hydroxyl groups of DNP-Ser and DNP-Thr by succinic anhydride (section III.3.A.). Therefore, other methods were investigated in an attempt to reduce or eliminate trifluoroacetate from the solvent system, while maintaining the solubility of the membrane protein.

Initial studies involved the dialysis of the protein salt mixture against 0.2 to 1.0 M Et$_3$N in DMF in an attempt to remove the trifluoroacetate while leaving the solubilized apobacteriorhodopsin in DMF and under basic conditions. However, under strictly anhydrous conditions, apobacteriorhodopsin was found to precipitate out of solution; its solubility was only maintained when 100 mM H$_2$O was included in the dialysate. This amount of water was not permissible in the solvent system because not only would it readily react with the acylating reagent but upon its reaction the loss of water may result in the membrane protein precipitating out of solution.

In the presence of Et$_3$N, the carboxyl groups of apobacteriorhodopsin would be ionized and Et$_3$NH$^+$ would be the counterion to the carboxylate groups. The fact that water was required to maintain the protein in solution suggested that Et$_3$NH$^+$ ion may not be sufficiently hydrophobic to permit the solubilization of the apobacteriorhodopsin in anhydrous DMF. The more hydrophobic counterion Bu$_4$N$^+$ was therefore considered for the carboxylate groups on the protein. Apobacteriorhodopsin was solubilized by the procedure outlined above and dialyzed against anhydrous DMF containing 100 mM Et$_3$N and 50 mM Bu$_4$N$^+$TFA$^-$ or Bu$_4$N$^+$Cl$^-$. After three exchanges of the dialysate, the apobacteriorhodopsin remained in solution. It
was further noted, that if the amphipathic salt was added to the TFA solubilized protein in DMF, prior to the addition of Et₃N, the apobacteriorhodopsin was more apt to remain in solution as the mixture was titrated with base. This enhanced solubility was thought to be due to the Bu₄N⁺ providing a more hydrophobic counterion to the carboxylate groups on the protein as they are generated by the addition of the Et₃N.

The major problem with this solubilization procedure is that during the dialysis in anhydrous DMF, the dialysis bag dehydrates and becomes brittle. Without extreme care, mere handling of the bag could cause it to break open, spilling the contents. To avoid this, gel permeation chromatography was investigated on the basis that it should readily separate the solubilized protein from the Et₃NH⁺TFA⁻ salt by the difference in their molecular weight. Sephadex LH-20 columns were chosen because their exclusion limits are 1 x 10³ to 2.0 x 10⁴ daltons and the support is stable in organic solvents. The solubilized protein, having a molecular weight of 26,000 (excluding the weight contributed by the hydrophobic Bu₄N⁺ counterions), would be totally excluded from the gel, while the Et₃NH⁺TFA⁻ salt, having a molecular weight of 158, would elute near the bed volume of the column. The overall result would be the complete elimination of trifluoroacetate from the solubilized apobacteriorhodopsin.

The mixture containing the apobacteriorhodopsin and the Et₃NH⁺TFA⁻ salt in DMF was applied to a Sephadex LH-20 column (1 x 28 cm) that had been equilibrated and eluted with anhydrous DMF containing 50 mM Bu₄N⁺Cl⁻ and 100 mM Et₃N. The elution of the protein was monitored by the Isco Absorbance Monitor at 280 nm. As anticipated, the solubilized apobacteriorhodopsin eluted near the void volume of the column based on the 280 nm absorption. Analysis of the rate of derivatization of the model
peptides in the presence of this effluent, established that the trifluoracetate had been entirely removed (data not shown).

This method was also found to be applicable for the solubilization of apobacteriorhodopsin in anhydrous Me₂SO or pyridine; apobacteriorhodopsin was initially solubilized in DMF using the TFA procedure and then applied to Sephadex LH-20 columns that were equilibrated and eluted with either Me₂SO or pyridine containing Bu₄N⁺Cl⁻ and Et₃N. Irrespective of the solvent used, the apobacteriorhodopsin still eluted near the void volume of the column and in the absence of the triethylammonium trifluoroacetate salt. Only after extensive analysis were DMF and Me₂SO eliminated as possible solvents due to the instability of the activated reagents under acylation conditions over time.

Further work was performed in an attempt to simplify the protein solubilization process; this resulted in the following improved procedure: The lipid free protein, 1 mg/mL in 88% formic acid ethanol (3:7, v/v), was first dried under reduced pressure, in the presence of Bu₄N⁺TFA⁻ (100 μmoles/ mg of protein). The protein residue was then twice redissolved and dried from 6 volumes of dry toluene and absolute ethanol (2:1, v/v) after which the protein residue was readily soluble in dry pyridine. This procedure ensures that both formic acid and water are completely removed from the protein, as assessed by ditrimethylsilyl trimesic anhydride reactivity and [³H]H₂O removal, respectively (data not shown). Membrane proteins, as well as water soluble proteins, such as apomyoglobin (data not shown), are thus rendered soluble in anhydrous pyridine as the tetrabutylammonium salt, in a form suitable for the efficient acylation of the amine and hydroxyl groups on the side chains of their corresponding amino acid residues.
III.2. Succinylation

By achieving the solubilization of the very hydrophobic membrane protein, apobacteriorhodopsin, in anhydrous organic solvent systems, the objective of extensive chemical derivatization of the protein should now be feasible. The generation of a negative charge at the site of modification can be achieved by any one of several reagents. Succinic anhydride is a suitable reagent due to its ability to react with and incorporate a carboxylate group at α and ε amine groups, sulfhydryl, imidazole and aromatic and aliphatic hydroxyl groups. Furthermore, succinic anhydride is already reactive, requiring no additional procedure for activation. Therefore, performing the succinylation reaction in the anhydrous solvent system developed for the solubilization of membrane proteins should be efficient and result in the complete modification of all amine and hydroxyl groups.

III.2.A. Reaction Characterization

In order to fully characterize the succinylation reaction, the hydroxylamino acids, serine and threonine were chosen because they were usually the least succinylated under aqueous conditions. If the conditions for derivatization were developed such that the rates of succinylation of these amino acids are optimized, then the rates for the other functional groups (Tyr and Lys) should also be maximal. The 2,4-dinitrophenyl derivatives of these hydroxylamino acids were utilized as model systems. The DNP group provides a convenient chromophore having a large extinction coefficient at 365 nm ($10^4 \text{M}^{-1}\text{cm}^{-1}$, Sanger and Tuppy, 1951) and thus enables all the amino acid related products generated during the reaction, to be observed.

Initial studies on succinylation were done in anhydrous DMF in the presence of $1.35 \text{ M Et}_3\text{NH}^+\text{TFA}^-$ and $1.35 \text{ M Et}_3\text{N}$. The DNP-amino acids
were acylated in the presence of 0.5 M succinic anhydride. The succinylation reaction was assessed by injecting an aliquot of the reaction, at the indicated times, onto a reversed phase HPLC column that was equilibrated and developed according to the legend of Figure 1. The effluent was monitored at 254 nm which was the fixed wavelength on the Beckman HPLC, and by integrating the area under the absorbance peaks, the DNP-amino acid related compounds could be quantitated. This procedure provided a rapid and reliable method for assessing the extent of succinylation of the model DNP-amino acid residues.

Representative chromatograms in Figure 1, illustrate the conversion of the starting DNP-amino acid into the corresponding succinyl-ester derivative. As observed, there was a complete loss in the starting DNP-amino acid, with time and the subsequent unambiguous formation of a new DNP-containing product having a distinct retention time. The succinyl ester products of both DNP-Ser and DNP-Thr elute later than the respective starting material, under these isocratic HPLC conditions at pH 4.1. This behaviour was due to the combined effects of the extra methylene groups of the succinyl group on the DNP-amino acids, the additional carboxyl group which would be partially protonated at pH 4.1 and to the low ionic strength of the eluent (10 mM sodium acetate), thereby enabling the modified DNP-amino acid to partition more readily into the ODS (octadecylsilica), the stationary phase of the reversed phase column.

Figure 2 summarizes the time course of succinylation of the DNP-amino acids, for the various time points analyzed by this HPLC method. As observed, DNP-Ser was fully derivatized within 2 hours while greater than 90% of DNP-Thr was succinylated within 6 hours. The rate of succinylation of the primary hydroxyl group of Ser was approximately 10 times faster than
Figure 1: Succinylation of DNP-amino acids. A 10 nmol aliquot of the succinylation reaction (Figure 2) was injected directly on to an Altex Ultrasphere ODS reversed phase column (15 x 250 mm) that was equilibrated and run isocratically in 40% methanol and 60% 10 mM aqueous sodium acetate, pH 4.1; the effluent was monitored at 254 nm. The acylation of DNP-serine was assessed at 0 (panel A), 30 (panel B) and 70 minutes (panel C). The DNP-threonine reaction was injected at 0 (panel D), 80 (panel E) and 280 minutes (panel F). The extent of acylation was determined by integrating the area under the peptide related peaks.
Figure 2: Time course of DNP-amino acid succinylation. DNP-amino acids, serine (▲), threonine (●), and glutamic acid (■) were separately reacted at 34 mM in DMF with 0.5 M succinic anhydride in the presence of 1.35 M Et₃NH⁺TFA⁻ salt and 1.35 M Et₃N. The extent of amino acid acylation was determined on a Beckman HPLC, as described in Figure 1.
the rate for the secondary hydroxyl group of Thr. These acylation rates reflect the both the differences in the pK\textsubscript{a} of the corresponding hydroxyl groups, which are 15 for a primary hydroxyl group and 16 for a secondary one (CRC Handbook, 1974) and the steric effects that are involved. These values indicate that Ser would be ten times better as a nucleophile than Thr. In this anhydrous solvent system, the succinylation reaction was efficient and complete derivatization of even a secondary hydroxyl group was achieved.

DNP-glutamic acid was reacted under identical conditions to establish that there was no site on the molecule that could undergo succinylation. The inability of the DNP-Glu to become acylated would therefore indicate that it must be due to the presence of the hydroxyl group on DNP-Ser and DNP-Thr that was solely responsible for the generation of the new DNP-amino acid derivatives observed with these substrates. Furthermore, the modification of the hydroxyl group was indeed monitored by this method. The formation of a mixed anhydride on the carboxylate group, under the reaction conditions, was not detected.

The O-succinylation of hydroxyl amino acids can also be quantitated by the alkaline hydroxylamine reaction of Hestrin (1949) (Moore and Stein, 1954, Balls and Wood, 1956 and Gounaris and Perlmann, 1967). This ester assay quantitatively determines the amount of ester present in a given sample by reacting the ester with hydroxylamine, at pH 10.5. The resulting hydroxymate derivative of the ester, upon acidification and subsequent addition of FeCl\textsubscript{3}, forms a soluble purple brown complex which absorbs strongly at 540 nm. The ester concentration can be determined from the absorbance based on the 0.465 absorbance that was obtained when one \textmu mole of ethylacetate was reacted under these conditions.
A comparison between the two methods of succinyl product detection, reversed phase HPLC and the alkaline hydroxylamine ester assay, was made to ensure that the percentage of product obtained and observed on the HPLC chromatograms (Figure 1) corresponded to the expected percent total ester calculated colorimetrically (Table 2). Within experimental error, a good correlation exists between HPLC quantitation of the product and the amount of ester present, as determined by the hydroxylamine assay. Furthermore, this experiment confirms that the new DNP containing product that was generated in the succinylation reaction, was indeed an ester derivative of the DNP-amino acids. This ester derivative could either be the succinyl or the trifluoracetetyl ester. This uncertainty was due to the presence of TFA in the succinylation reaction which would likely generate a small portion of the mixed anhydride, $\text{F}_3\text{C-CO-O-CO-CH}_2\text{CH}_2\text{CO-O}^-$, even though the succinic anhydride is a relatively stable cyclic structure. TFA anhydride is well known as a condensing reagent resulting in the esterification of a weaker carboxylic acid; its presence results in the generation of the mixed anhydride. The mixed anhydride can also be generated by the addition of TFA to a symmetric anhydride of a weaker carboxylic acid. The mixed anhydride is generally more reactive than the symmetric anhydride. However, the formation and reactivity of the mixed anhydride is subject to many factors, such as the type of solvent, the stability of the symmetric anhydride and the difference between the acid strengths of the two acids. Such a mixed anhydride, between TFA and succinic anhydride, has been shown to only give rise to the succinyl ester, though in relatively low yield, primarily due to the difference in the acid strengths and to the stability of the cyclic anhydride (Tedder, 1955). However, the reactivity of TFA mixed anhydrides are generally found to be optimal under acidic conditions and are
<table>
<thead>
<tr>
<th>Model Substrate</th>
<th>Reaction Time (h)</th>
<th>Product Yield (% theoretical)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HPLC</td>
</tr>
<tr>
<td>DNP-Ser</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>100</td>
</tr>
<tr>
<td>DNP-Thr</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>0.67</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>1.33</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>4.67</td>
<td>91</td>
</tr>
</tbody>
</table>

Table 2: Succinyl-ester quantitation. DNP-Ser and DNP-Thr were succinylated under the conditions outlined in Figure 2. At the indicated times two aliquots of the reaction were removed to assess the extent of acylation by two different methods, HPLC analysis (Figure 1) and by the hydroxylamine ester assay, as described in Materials and Methods.
reduced under basic conditions, in some instants giving rise to the trifluoroacetyl ester (Tedder, 1955). Therefore, the major concern of the presence of the TFA under the basic conditions employed for the succinylation is the type of DNP-amino acid product formed from this reactive mixed anhydride molecule, whether it would be the succinyl or the trifluoroacetyl ester. The hydroxylamine ester assay is limited to assessing the amount of ester and is not capable of distinguishing between different types of esters. However, the reversed phase HPLC would be able to address the structural differences between the two esters.

To confirm that the single product of succinylation observed by HPLC was indeed the succinyl ester, DNP-Thr was separately acylated with succinic anhydride and trifluoroacetic anhydride, under the same reaction conditions specified in Figure 2. An aliquot of each reaction was then injected onto the HPLC column that was developed according to the legend in Figure 3. As observed in the chromatograms (Figure 3), each ester product eluted in a unique fashion, having a specific retention time for the isocratic gradient; succinyl-DNP-threonine eluted at 17.4 minutes while the trifluoroacetyl-DNP-threonine eluted at 21.9 minutes. The trifluoroacetyl ester of DNP-Thr was more hydrophobic than the corresponding succinyl ester, even though the carboxyl group of the succinyl moiety was protonated under these conditions. This behaviour is presumably due to the hydrophobic nature of the trifluoromethyl group and the absence of a polar carboxyl group on the trifluoroacetyl ester.

The crucial point illustrated here, was that there was no absorbance peak, in the standard succinylation reaction, that eluted in the corresponding position of the trifluoroacetyl ester. Furthermore, the rate constant for trifluoroacetylation of DNP-Thr in the presence of 0.67 M Et₃N
Figure 3: Identification of the DNP-threonine ester products. DNP-threonine was derivatized under the reaction conditions specified in Figure 2 using succinic anhydride (panel A) in one reaction and trifluoroacetic anhydride (panel B) in another. After allowing the reactions to go to completion, a 10 nmol aliquot was injected on an Altex Ultrasphere ODS reversed phase column that was equilibrated and eluted isocratically with 45% aqueous methanol containing 0.5% formic acid.
was only $2.8 \times 10^{-5} \text{M}^{-1}\text{s}^{-1}$, which was 20 times slower than the rate constant for succinylation under standard reaction conditions. Therefore, the only ester that was generated under these conditions of succinylation was the succinyl ester, even in the presence of TFA.

A comparison of the rate constants of succinylation was used to assess the different reaction conditions, in the attempt to select one that achieves the optimal rate of succinylation for the hydroxylamino acids. Succinylation is a second order reaction and its rate constant can be calculated from the equation (Barrow, 1966):

$$k_t = \frac{1}{A-B} \ln \frac{B(A-X)}{A(B-X)}; \text{ where } A = \text{[succinic anhydride] initial}
B = \text{[DNP-threonine] initial}
X = \text{[DNP-threonine] initial} - \text{[DNP-threonine] at time } t$$

This equation was employed to calculate the apparent succinylation rate constants for DNP-threonine in a series of experiments in which the concentrations of TFA and Et$_3$N were varied (Table 3). These rate constants were then plotted against the corresponding molar ratio of TFA to Et$_3$N (Figure 4).

In the absence of TFA, 90% of DNP-threonine was succinylated within 1 hour, corresponding to a rate constant of $4.9 \times 10^{-3} \text{M}^{-1}\text{s}^{-1}$. In the presence of TFA, the rate constant was reduced by a factor of 6 while the molar ratio of TFA to Et$_3$N was only increased from 0 to 0.35, maintaining the pH of the solvent system at 11. These results imply that the TFA must affect the succinylation reaction by a more direct route than through solvent pH; it is likely that the TFA reacts with the succinic anhydride to form the mixed anhydride, which is less reactive under these basic conditions.

Further assessment of the reaction conditions determined the
<table>
<thead>
<tr>
<th>Reagent Concentration (mM)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DNP-Thr</td>
<td>Succinic Anhydride</td>
<td>TFA</td>
<td>Et&lt;sub&gt;3&lt;/sub&gt;N</td>
<td>Molar Ratio TFA:Et&lt;sub&gt;3&lt;/sub&gt;N</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0</td>
<td>135</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>135</td>
<td>386</td>
<td>0.35</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>135</td>
<td>270</td>
<td>0.50</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>135</td>
<td>196</td>
<td>0.69</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>135</td>
<td>136</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 3: Reaction conditions for the assessment of the effect of the molar ratio of trifluoroacetic acid to triethylamine on the rate of succinylation. DNP-Thr was derivatized under the reaction conditions specified in the table, and the extent of acylation after one hour incubation was assessed by reversed phase HPLC, as outlined in Figure 1. The apparent rate constant was then calculated from this value.
Figure 4: The effect of the molar ratio of trifluoroacetic acid to triethylamine on the rate of succinylation. The apparent rate constants were assessed for the succinylation of DNP-threonine under the reaction conditions specified in Table 3. These are plotted against the corresponding molar ratio of trifluoroacetic acid to triethylamine.
effect on the succinylation rate constant of varying the concentration of the TFA with respect to the concentration of the succinic anhydride. The apparent rate constant was calculated for a series of experiments in which the molar ratio of succinic anhydride to TFA was varied while the molar ratio of TFA to Et₃N was kept constant at 0.5 (Table 4). The apparent rate constants of succinylation were then plotted against the corresponding molar ratio of succinic anhydride to TFA (Figure 5).

When the molar ratio of succinic anhydride to TFA was greater than one, the apparent succinylation rate constant approaches \(6 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}\). At low molar ratios, the succinic anhydride was primarily in the form of the mixed anhydride due to the concentration of TFA. The mixed anhydride had a very low rate constant of succinylation in the presence of base (\(2.8 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}\)), approximately 20 fold less than the maximal rate constant obtained in the absence of TFA.

Figure 6 illustrates the proposed equilibria that were believed to be involved in the succinylation of DNP-Thr when in the presence of TFA. \(k_1\) and \(k_{-1}\) are the rate constants for the formation and the breakdown of the mixed trifluoroacetic succinic anhydride. \(k_2\) is the rate constant for the succinylation of DNP-Thr by succinic anhydride. \(k_3\) is the rate constant for the succinylation of DNP-Thr by the mixed anhydride. \(k_{\text{app}}\) is the observed rate constant of succinylation and is dependent on all the other rate constants. As stated, \(k_{\text{app}} = k_2\) when the molar ratio of succinic anhydride to TFA exceeds ~1, while \(k_{\text{app}} = k_3\) when the molar ratio of succinic anhydride to TFA approaches 0. Experimentally, \(k_2\) was found to be substantially greater than \(k_3\); \(5.9 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}\) as compared to \(2.8 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}\). As previously ascertained, there was no formation of any trifluoroacetyl-DNP-Thr under these reaction conditions (Figure 3).
<table>
<thead>
<tr>
<th>DNP-Thr (mM)</th>
<th>Suc Anh (mM)</th>
<th>TFA (M)</th>
<th>Et&lt;sub&gt;3&lt;/sub&gt;N (M)</th>
<th>Molar Ratio</th>
<th>Molar Ratio</th>
<th>Apparent Rate Constant (x 10&lt;sup&gt;4&lt;/sup&gt; M&lt;sup&gt;-1&lt;/sup&gt;s&lt;sup&gt;-1&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>0.14</td>
<td>0.28</td>
<td>0.5</td>
<td>3.70</td>
<td>5.9</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>0.67</td>
<td>1.35</td>
<td>0.5</td>
<td>0.74</td>
<td>5.4</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>1.35</td>
<td>2.70</td>
<td>0.5</td>
<td>0.37</td>
<td>3.8</td>
</tr>
<tr>
<td>34</td>
<td>500</td>
<td>1.35</td>
<td>2.70</td>
<td>0.5</td>
<td>0.37</td>
<td>3.9</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>1.35</td>
<td>2.70</td>
<td>0.5</td>
<td>0.04</td>
<td>1.4</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1.35</td>
<td>2.70</td>
<td>0.5</td>
<td>0.01</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 4: Reaction conditions for the assessment of the effect of the molar ratio of succinic anhydride to trifluoroacetic acid on the rate of succinylation. DNP-Thr was derivatized under the conditions specified in the table and the extent of derivatization after a one hour reaction was assessed by HPLC, as outlined in Figure 1. The apparent rate constant was calculated from this value.
Figure 5: The effect of the molar ratio of succinic anhydride to trifluoroacetic acid on the rate of succinylation. The apparent rate constant was assessed for the succinylation of DNP-threonine under the reaction conditions specified in Table 4. These are plotted against the corresponding molar ratio of succinic anhydride to trifluoroacetic acid.
Figure 6: Proposed equilibria involved in the succinylation of DNP-threonine in the presence of trifluoroacetic acid
Therefore, in order to achieve a maximal rate of succinylation while in the presence of TFA, the molar ratio of succinic anhydride to TFA must be kept above 1 and the amount of Et₃N must be at least twice the concentration of the TFA that is present in the reaction mixture. Alternatively, the rate could be enhanced by a factor of 6, to approximately 4.9 x 10⁻³ M⁻¹s⁻¹, if the TFA could be entirely removed from the succinylation reaction.

In order to ensure rapid and complete derivatization of all possible amine and hydroxyl groups on the protein, the rate of succinylation should be optimal. Due to the volume of TFA needed to initially solubilize the membrane protein residue, the amount of succinic anhydride that would be required for a rapid derivatization would exceed the number of functional groups on the protein by several hundred fold. A simpler procedure of removing the TFA from the protein solution prior to the addition of the succinic anhydride would permit the use of more quantitative amounts of the reagent. As previously presented (section III.1.), the removal of the TFA was accomplished by dialysis or gel permeation of the protein solution in the presence of a hydrophobic counterion (Bu₄N⁺), which acts as a phase transfer reagent by maintaining the membrane protein's solubility in the anhydrous organic solvent system. The rate constant for succinylation was therefore reassessed in the presence of different amphipathic salts, using DNP-Ser as the model substrate (Table 5). In the presence of Et₃NH⁺TFA⁻, which is the salt that is normally generated during the solubilization of apobacteriorhodopsin, the kₓ for the succinylation of DNP-Ser was found to be approximately 6.7 x 10⁻³ M⁻¹s⁻¹. This rate constant was 10 times that observed for DNP-Thr, which reflects both the differences in the pKₐ of the hydroxyl groups and steric hindrance. When the aprotic
<table>
<thead>
<tr>
<th>Amphipathic Salt (50 mM)</th>
<th>Apparent Rate Constant (x $10^3$ M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Et$_3$NH$^+$TFA$^-$</td>
<td>6.7</td>
</tr>
<tr>
<td>Bu$_4$N$^+$TFA$^-$</td>
<td>6.7</td>
</tr>
<tr>
<td>Bu$_4$N$^+$Cl$^-$</td>
<td>12.4</td>
</tr>
</tbody>
</table>

Table 5: The effect of different salts on the rate of succinylation. The apparent rate constants were determined for the succinylation of 1 mM DNP-Ser in DMF with 50 mM succininc anhydride in the presence of 10 mM Et$_3$N and 50 mM of the specified amphipathic salt.
amphipathic salt, Bu₄N⁺TFA⁻ was used there was no significant change in the rate constant. This indicated that the TFA ion reduced the rate constant possibly by forming the mixed anhydride and that even a strong counterion like Bu₄N⁺, cannot significantly shift the equilibrium back to the symmetric anhydride. However, when TFA was entirely removed from the reaction by using Bu₄N⁺Cl⁻ as the phase transfer reagent, the k_app was increased to approximately 1.24 x 10⁻² M⁻¹s⁻¹. Furthermore, this Bu₄N⁺Cl⁻ salt was found to maintain the solubility of the apobacteriorhodopsin in anhydrous organic solvents (section III.1.).

The increase in the rate constant when the trifluoroacetate ion was replaced with the chloride ion was presumably due to the inability of the chloride ion to react with the succinic anhydride. However, the increase in the rate constant was only by a factor of 2 and not by a factor of 6, as was expected on the basis of the increase in the rate constant observed for the succinylation of DNP-Thr when TFA was eliminated from the reaction conditions (from 8.3 x 10⁻⁴ to 4.9 x 10⁻³ M⁻¹s⁻¹; Figure 4). However, in the DNP-Thr situation no other amphipathic salt was added to the reaction mixture. Therefore, the polarity of the solvent system must be the contributing factor that was involved in the succinylation reaction. The increased polarity (dielectric constant) of the solvent due to the amphipathic salt, may decrease the nucleophilic character of the oxygen of the hydroxyl group or it may stabilize intermediates in the reaction, thereby reducing the rate of succinylation.

III.2.B. Characterization of Succinyl-Apobacteriorhodopsin

The amino acid sequence of the apobacteriorhodopsin (Figure 7) provides the basis from which to theoretically determine the number of sites
Figure 7: Amino acid sequence of apobacteriorhodopsin: sites of derivatization. From the published primary sequence of apobacteriorhodopsin (Khorana et. al., 1979) the theoretical sites of acylation (*) and cleavage sites of trypsin (ARG) can be identified. In this schematic representation of the protein <GLU represents the pyroglutamic acid residue on the N-terminus.
of derivatization on this membrane protein. There are 49 possible sites of succinylation including the Tyr residues. However, due to the instability of aromatic succinyl esters, the theoretical total is 38 sites: 7 Lys, 13 Ser, and 18 Thr residues. Upon derivatization and incorporation of a negative charge at each of these sites there would be a large increase in the net negative charge density on the protein; specifically, an increase from one charge per 3714 molecular weight to one charge per 500 molecular weight, a 7.43 fold increase (Table 6). This increase in the net negative charge density on the apobacteriorhodopsin should greatly increase its water solubility in the absence of detergents and chaotropic agents, thereby making it generally more manageable under aqueous conditions and amenable to enzymatic or chemical fragmentation and to protein sequencing.

Apobacteriorhodopsin was solubilized in DMF, using the TFA procedure (section III.1) in the presence of Et$_3$N and without the addition of the Bu$_4$N$^+$TFA$^-$ salt. The membrane protein was succinylated in the presence of 0.5 M succinic anhydride for 5.5 hours. In order to observe the behaviour of the modified membrane protein on a molecular sizing column, the reaction mixture was diluted with one volume of 50 mM NH$_4$HCO$_3$, pH 8.8 and applied to a 40 mL Sephadex G-50 column which was equilibrated and run with the same ammonium bicarbonate buffer. Sephadex G-50 is capable of fractionating globular proteins between $1.5 \times 10^3$ to $3.0 \times 10^4$ daltons.

The elution profile in Figure 8 shows three absorbance peaks: 1. A peak eluting near the void volume of the column that has a mass of $3.0 \times 10^4$ daltons or larger. This probably represents partially succinylated protein aggregates that were fully excluded from the gel. 2. A partially retained peak of about $2.8 \times 10^4$ daltons which likely corresponds to the succinylated apobacteriorhodopsin monomer, which has an expected molecular weight of
<table>
<thead>
<tr>
<th>Location of Charges</th>
<th>Apobacteriorhodopsin native</th>
<th>succinyl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp residue</td>
<td>-9</td>
<td>-9</td>
</tr>
<tr>
<td>Glu residue</td>
<td>-11</td>
<td>-11</td>
</tr>
<tr>
<td>C' terminal</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Arg residue</td>
<td>+7</td>
<td>+7</td>
</tr>
<tr>
<td>Lys residue</td>
<td>+7</td>
<td>-7</td>
</tr>
<tr>
<td>succinyl-Ser</td>
<td>0</td>
<td>-13</td>
</tr>
<tr>
<td>succinyl-Thr</td>
<td>0</td>
<td>-18</td>
</tr>
<tr>
<td>Net Negative Charge</td>
<td>7</td>
<td>52</td>
</tr>
<tr>
<td>Mol Wt per Charge</td>
<td>3714</td>
<td>500</td>
</tr>
</tbody>
</table>

Table 6: Net negative charge of apobacteriorhodopsin: before and after succinylation. A comparison of the theoretical charges that would exist at pH 7 on apobacteriorhodopsin in its native state and subsequent to succinylation and the resulting net negative charge density of the protein.
Figure 8: Elution profile of succinyl-apobacteriorhodopsin from a Sephadex G-50 column (1.5 x 30 cm). Apobacteriorhodopsin (0.5 mg) was solubilized in 0.5 mL of DMF in the presence of 1.35 M Et₃NH⁺TFA⁻ salt and 1.35 M Et₃N. Succinic anhydride (100 mg) was added to the solution and then left at room temperature for 5.5 h. The reaction mixture was then diluted with an equal volume of 50 mM NH₄HCO₃, pH 8.8 and applied to a Sephadex G-50 column that was equilibrated and eluted with 50 mM NH₄HCO₃, pH 8.8. The effluent was monitored at 280 nm by an Isco Absorbance Monitor (model UA-5).
2.98 x 10^4. The succinylated protein appears to have been retarded possibly due to adsorption to the Sephadex backbone through a hydrophobic effect. An absorbance peak that elutes near the bed volume of the column, corresponding to a molecular mass of less than 1.5 x 10^3 daltons, was thought to be due to the DMF component of the reaction mixture. The apparent trailing of the succinyl-apobacteriorhodopsin absorbance peak was believed to be due to the DMF with which it was applied.

The succinylated apobacteriorhodopsin appeared to be fully water soluble and eluted as a fairly sharp peak from Sephadex G-50. The rationale for enhancing membrane proteins' water solubility through extensive chemical modification was sound and the objective was achieved via succinylation.

Having rendered apobacteriorhodopsin water soluble through the modification by succinic anhydride, it was desired to assess the suitability of this succinylated membrane protein as a substrate for the enzyme, pyroglutamate aminopeptidase. This enzyme is specific for the hydrolysis of the amide bond between N-terminal pyroglutamate residue and any other amino acid. Even though its first residue is pyroglutamate (Figure 7), apobacteriorhodopsin is not a good substrate for this enzyme, it will not undergo cleavage. However, the N-terminal CNBr fragment of apobacteriorhodopsin was found to be a suitable substrate but only after it was rendered water soluble by initially reacting its terminal carboxyl group with ethylenediamine and then derivatizing it with 4-sulfophenylisothiocyanate (Gerber et. al., 1979). Therefore, the first peptide bond in the amino acid sequence of apobacteriorhodopsin is cleavable by pyroglutamate aminopeptidase, under the right conditions. The insolubility of the membrane protein under the conditions of the enzyme digestion accounts for
it being a poor substrate. Furthermore, the enzyme is not functional in the presence of detergents and chaotropic agents, and though it can tolerate low percentages (less than 10%) of organic solvents these are not sufficient to solubilize membrane proteins, such as apobacteriorhodopsin. Therefore, if succinylation has truly rendered apobacteriorhodopsin water soluble, then this derivative of the membrane protein should be a reasonably good substrate on which pyroglutamate aminopeptidase will act.

The elaborate procedure for the detection of pyroglutamate, as outlined in Figure 9, was required to ensure the quantitative recovery of the pyroglutamate, its efficient conversion and derivatization to DNP-glutamate and the accurate integration of DNP-Glu on the HPLC. This procedure removed the reactive contaminants that contributed to the background on the HPLC chromatogram. The best results were obtained when the enzyme was dialyzed against the potassium phosphate buffer prior to the digestion of the succinylated protein, in order to remove the stabilizing agents with which the enzyme was packaged. After incubation with the enzyme, the β-mercaptoethanol in the digest was reacted with iodoacetic acid in order to derivatize the thiol groups. Excess cysteine was then used to react with the remaining iodoacetic acid. Passage through the Dowex-50 × 2 column ensured the removal of all molecules bearing an amine group. In this manner, all the contaminants that interfered with DNP-Glu identification, by reversed phase HPLC (Figure 9), were removed.

There was a 96% recovery of the DNP-Glu after removal of the pyroglutamic acid residue from succinyl-apobacteriorhodopsin and its subsequent hydrolysis and dinitrophenylation. This recovery from 5 nmoles of protein was corrected for the actual yields that were obtained for a sample containing 5 nmoles of pyroglutamate standard, having undergone this
Figure 9: Enzymatic removal of the N-terminal pyroglutamic acid residue. Succinyl-apobacteriorhodopsin (5 nmoles) was dialyzed against 100 mM K_2HPO_4, pH 8, 10 mM EDTA and 10 mM β-mercaptoethanol for 12 hours and then incubated for 24 hours at 37°C with 1.5 Units of freshly dialyzed pyroglutamate aminopeptidase. This reaction mixture was then processed according to the figure to yield DNP-glutamic acid. The detection and quantitation of the modified amino acid was achieved by HPLC at 254 nm on an Altex Ultrasphere ODS reversed phase column that was equilibrated with 18% methanol and 82% 10 mM aqueous sodium acetate, pH 5.0. A gradient from 18% to 70% methanol was employed to optimize its detection.
Step 1: PGAPase; predialyzed
37°C for 20 h
100 mM K₂HPO₄, pH 8
10 mM β-MSH, 10 mM EDTA

Step 2: 35 mM iodoacetic acid, 15 min
Step 3: 45 mM cysteine, 15 min
Step 4: Dowex 50×2 (H⁺)
Step 5: 2 N HCl, 100°C for 1 h

Step 6: 1 M Na HCO₃, pH 9.5
160 mM FDNB (EtOH), 1 h

Dinitrophenyl-glutamate
procedure. This established that for the membrane protein, apobacteriorhodopsin, extensive succinylation not only rendered it water soluble but allowed it to become a suitable substrate for pyroglutamate aminopeptidase.

Protein sequencing by the Edman degradation procedure cannot be done on apobacteriorhodopsin due to the N-terminal pyroglutamate residue which 'blocks' the protein; there is no N-terminal amine group with which the PITC can react. Succinylation and subsequent digestion with the N-terminal deblocking enzyme, pyroglutamate aminopeptidase, would render apobacteriorhodopsin amenable to N-terminus sequencing, starting from its second amino acid residue.

This procedure could be utilized for other membrane proteins that are similarly blocked by pyroglutamate residues, and possibly for those proteins that are effectively blocked by the acetylation of their first amino acid residue. Rendering the proteins water soluble via succinylation would substantially improve their suitability as a substrate for the enzyme N-α-acylpeptide hydrolase. This would deblock the protein and allow protein sequencing to occur from the N-terminus, starting from the second amino acid residue.

III.2.C. Succinyl Ester Stability

The stability of the succinyl-ester on Ser and Thr residues was required in order for this derivatization procedure of rendering membrane proteins water soluble to be useful. To investigate the ester stability, the succinylated model substrates were reacted in various buffers, at room temperature. After 18 hours an aliquot was injected onto the HPLC reversed phase column to quantitate the loss of the succinyl ester and the concomitant
formation of the DNP-amino acids. The aqueous systems that were tested included: H$_2$O, 100 mM NaCl, 0.1 M NH$_3$OH; pH 10.5, 50 mM NH$_4$HCO$_3$; pH 8.8, 50 mM Tris buffer; pH 8.1, 10 mM β-mercaptoethanol, 50 mM cysteine and 50 mM sodium glycinate. Under all of these solvent conditions, it was found that the succinylated product was stable except for its reaction with alkaline hydroxylamine which resulted in 100% ester cleavage, as anticipated (Table 7).

The succinylated model substrates were then analyzed for their stability under protein sequencing (spinning cup) and CNBr fragmentation conditions. They were shown to be stable in anhydrous heptafluorobutyric acid (HFBA) at 5°C (cleavage conditions of the phenylthiocarbamoyl amino acid residue during sequencing), in Quadrol buffer at 57°C (0.1 M Quadrol; (N,N,N,N-tetrakis(2-hydroxypropyl) ethylenediamine)-TFA- buffer in n-propanol/water (3:4 v/v), pH 9.0) and in 70% formic acid at room temperature, overnight (CNBr cleavage conditions).

These results established that the succinyl ester on the modified DNP-Ser and DNP-Thr model substrates was quite stable. On the basis of these findings and under the assumption that the model system accurately represents the protein situation, it was anticipated that the succinyl esters will be as stable on the amino acid residues of proteins.

If $[^{14}$C]succinic anhydride was used as the derivatizing reagent the behaviour of the radioactive membrane protein product could be better monitored and characterized. This radioactive modification of apoBR would help to assess the Sephadex elution profile, elucidate the enzymatic fragmentation pattern that would be generated by trypsin or staphlococcal aureus V8 on a PAGE system, accurately assess the stability of the succinyl
<table>
<thead>
<tr>
<th>Buffer System</th>
<th>Incubation Time (h)</th>
<th>Remaining Succinyl-DNP-Ser (% theoretical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>0.1 M NaCl</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>0.1 M NH₂OH, pH 10.5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>50 mM NH₄HCO₃, pH 8.8</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>50 mM Tris, pH 8.1</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>10 mM β-MSH</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>50 mM cysteine</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>50 mM sodium glycinate</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>HFBA (anhydrous), 57°C</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>0.1 M Quadrol, 57°C</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>70% formic acid</td>
<td>12</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7: Stability of succinylated DNP-serine. DNP-Ser was succinylated at a concentration of 40 mM in DMF with 2.0 M succinic anhydride and in the presence of 100 mM Et₃N. The resulting product was diluted to 1 mM in the various buffers and incubated for the specified period of time. A 10 nmol aliquot was then analyzed using the same conditions as those outlined in Figure 1.
ester on the protein and ultimately aid in the analysis of the automated amino acid sequencing of succinyl-apoBR, with respect to the repetitive yields of Ser and Thr residues.

The optimum anhydrous solvent system was required for an efficient derivatization with $[^{14}$C]succinic anhydride; the solubilized apoBR (2.5 mg) in 1 mL of the TFA/ Et$_3$N/ DMF mixture was dialyzed for 6 hours at room temperature against 3 times 20 mL of a 50 mM Bu$_4$N$^+$Cl$^-$ and 50 mM Et$_3$N in dry DMF. The succinic anhydride concentration could now be decreased to 100 mM and still result in complete derivatization. It was further shown that an aliquot of the protein derivatizing reaction mixture was capable of modifying DNP-serine even after an overnight reaction (data not shown). (The amount of reactive contaminants, such as H$_2$O and primary and secondary amines was found to be approximately 30 mM by reagent titration of the optimal solvent system in the presence of the model substrates.) The specific activity of the succinic anhydride was 1 mCi/mmole which would result in 1 nmole of the fully derivatized apoBR having 8.36 x 10$^4$ dpm (38 sites of succinylation).

Upon derivatization, a portion of the $[^{14}$C]succinyl-apoBR (40 μg) was subjected to dialysis to determine the amount of non-dialyzable counts in the retentate. The modified apoBR was diluted with 0.5 mL 1 N NaCl and then dialyzed against 500 mL 1 N NaCl. The solution was changed 3 times, once every 2 hours with the last dialysis proceeding overnight (16 hours). The retentate contained only 1.8 x 10$^4$ dpm of the expected 1.28 x 10$^5$ dpm. To ensure that the protein had not dialyzed out, a dialysis bag with a molecular weight limit of 10$^3$ was used; the retentate after overnight only contained 2.0 x 10$^4$ (16%) of the expected 1.28 x 10$^5$ dpm. A Lowry protein determination further verified that all the protein was still present in the dialysis bag. The
conclusion from this work was that either the succinylation was incomplete or the succinyl esters were too labile and were hydrolyzing off the protein.

To systematically analyze this problem, the succinylation reaction was performed with two solvent systems of apoBR solubilization, in the presence of TFA and in its absence through dialysis against Bu₄N⁺Cl⁻. In both systems, 200 µg of apoBR was succinylated for 40 hours with [¹⁴C]succinic anhydride having a specific activity of 1 mCi/mmole. An appropriate aliquot of each reaction, containing approximately 85 µg, was then diluted with H₂O to a volume of 0.5 mL and then dialyzed three times against 500 mL H₂O for 3 hours. The recovered nondialyzable counts from both aliquots were approximately 2.86 x 10⁵ which was within experimental error to the expected 2.73 x 10⁵ dpm, indicating that the succinylation reaction was complete.

To test the stability of the succinyl ester on the protein the nondialyzable counts were pooled and divided into four and each portion was diluted with 0.5 mL of a specified buffer and then dialyzed against 25 mL of the same buffer at room temperature for 12 h. The buffers that were used included 0.1 M NH₂OH; pH 10.5, 50 mM NH₄HCO₃; pH 8.8, 375 mM Tris; pH 8.8 and 126 mM Tris; pH 6.8. The last two buffers correspond to the amount of base in the Laemmli stacking gel and separating gel, respectively. The radioactivity was assessed for the retentate and the dialysate (Table 8).

The results indicated that the hydroxylamine treatment effectively removes the O-succinyl groups leaving the succinyl amide linkages intact; lysine residues make up 18.4% of the total sites of succinylation on the protein and after reaction with hydroxylamine, 16% of the initial radioactivity was left in the retentate. In all the other buffers tested, greater than 50% of the radioactive succinyl groups were lost over the 12 hour time period. The succinyl esters on the modified apoBR hydrolyze off, independent of the
<table>
<thead>
<tr>
<th>Dialysis Buffer</th>
<th>Dialysis Component</th>
<th>Radioactivity (x 10^{-4} dpm)</th>
<th>Percentage of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 M NH₂OH pH 10.5</td>
<td>dialysate retentate</td>
<td>10.4 2.0</td>
<td>84 16</td>
</tr>
<tr>
<td>50 mM NH₄HCO₃ pH 8.8</td>
<td>dialysate retentate</td>
<td>8.7 6.6</td>
<td>57 43</td>
</tr>
<tr>
<td>375 mM Tris pH 8.8</td>
<td>dialysate retentate</td>
<td>8.0 7.0</td>
<td>53 47</td>
</tr>
<tr>
<td>126 mM Tris pH 6.8</td>
<td>dialysate retentate</td>
<td>8.4 6.1</td>
<td>58 42</td>
</tr>
</tbody>
</table>

Table 8: Stability of [¹⁴C]succinyl-apobacteriorhodopsin. Apobacteriorhodopsin that had been succinylated with [¹⁴C]succinic anhydride, under two different solvent conditions, were dialyzed exhaustively against water and assessed by radioactivity as being completely acylated. Aliquots of these protein solutions were pooled, divided equally into four, diluted with an equal volume with one of the four buffers indicated and then was subsequently dialyzed against 25 mL of the corresponding buffer for 12 hours at room temperature. The total amount of radioactivity was then determined for both the retentate and the dialysate; an aliquot of each was diluted into Amersham ACS and the radioactivity determined using a Beckman LS-7800 scintillation counter.
concentration of the amine groups in the buffer. This result may be explained by the ability of the succinyl group to cyclize back to the 'stable' 5 membered cyclic anhydride, as was observed for the succinyl-Tyr residues (Gounaris and Perlmann, 1967).

These results contradict the stability studies performed with the derivatized model substrates, succinyl-DNP-Ser and succinyl-DNP-Thr. These model substrates must not be an accurate representation of the amino acid residues in the protein due to either the proximity of the DNP chromophore or the carboxylate group at the C-terminus of the amino acid itself. It is likely due to the electrostatic repulsion of the negative charge of the carboxylate group that prevents the free carboxylic acid end of the succinyl group from cyclizing to result in the cleavage the ester.

To prove this, the methyl ester of DNP-serine was formed by reacting DNP-serine with freshly generated diazomethane (\( \text{CH}_2^-\text{N}^+\equiv\text{N} \)). The resulting product was then isolated, succinylated in the absence of TFA and subsequently diluted into each of the following solutions and reacted at room temperature: \( \text{H}_2\text{O}, 100 \text{ mM NaCl, 0.1 M NH}_2\text{OH; p}\text{H 10.5, 50 mM NH}_4\text{HCO}_3; \text{p}\text{H 8.8, 50 mM Tris; p}\text{H 8.1, 50 mM cysteine and 50 mM sodium glycinate.} \) The conversion of the succinylated product back to its starting material, the DNP-Ser methyl ester, was monitored with time by HPLC, using the conditions specified in Table 9. These results appear to show some of the instability that was observed for the succinyl esters on the modified apobacteriorhodopsin; other than in \( \text{H}_2\text{O}, \) the succinyl ester on the DNP-Ser methyl ester appeared to hydrolyze off at different rates depending on the reaction buffer (Table 9).
Table 9: Stability of succinylated DNP-Ser methyl ester. The methyl ester of DNP-Ser was succinylated at a concentration of 10 mM in DMF with 1 M succinic anhydride in the presence of 20 mM Et₃N. After complete derivatization, 0.5 μmol aliquots were diluted into the various buffers and reacted at room temperature for the specified time. An aliquot of 20 nmol was then analyzed by HPLC as in Figure 1, except that the column was equilibrated in 35% methanol and run using a gradient of 35% to 70% over 10 minutes.
III.2.D. Conclusion:

Succinylation of the membrane protein, apobacteriorhodopsin rendered it completely water soluble in the absence of detergents and chaotropic agents as atested by its ability to be eluted from Sephadex G-50 and to be a suitable substrate for pyroglutamate aminopeptidase. Even though the O-succinyl groups are not stable on the derivatized protein, the loss of some of these succinyl groups did not appear to adversely affect the solubility of succinyl-apobacteriorhodopsin during the time periods involved. Further studies, such as N-terminal sequencing of the deblocked succinyl-apobacteriorhodopsin, were not performed for the following reasons: 1. The instability of the succinyl groups on the membrane protein would result in protein heterogeneity, which could result in incomplete reactions, such as protein fragmentation, and ultimately permit protein aggregation to occur possibly culminating in protein precipitation. 2. An entirely different chemical reagent is required to overcome the instability of the succinyl group; the resulting modified apobacteriorhodopsin will behave differently than the succinyl-apobacteriorhodopsin.

The succinylation experiments, served to demonstrate the feasibility of the approach: The succinylated membrane protein was sufficiently water soluble to be able to elute from Sephadex columns in aqueous buffers in the absence of detergents. Succinyl-apobacteriorhodopsin was sufficiently soluble to permit removal of the blocking pyroglutamic acid residues from the amino terminus by pyroglutamate aminopeptidase. In addition, the development of the solvent system utilizing hydrophobic ion pairing (Bu₄N⁺Cl⁻) permitted the solubilization of proteins under strictly anhydrous conditions.
III.3. Model System; DNP-Peptides

The utilization of the HPLC and DNP-hydroxylamino acids as model substrates proved to be a rapid procedure for monitoring the acylation reaction. However, these model compounds do not provide an accurate representation of the stability of the ester derivatives of the hydroxylamino acid residues in the protein (section III.2.C.). The proximity of the negative charge of the C-terminal carboxylate group appears to enhance the stability of the ester. The choice and design of a model system is very critical and obviously must take into account the proximity of any unusual group such as the DNP chromophore and the "terminal" carboxyl group. The reactivity of the functional group on the side chain of amino acids in the model system should be representative of its counterpart in a protein; this would necessitate that the model peptide containing a DNP chromophore should be synthesized in such a way as to keep the chromophore at a reasonable distance from the site of interest. The terminal carboxylate group could be dealt with in a similar manner or possibly converted to the unreactive amide derivative. The amide derivative would be superior to an ester derivative because it could not readily participate in any transesterification or hydrolysis reaction, as might an ester.

The amide derivatives of DNP-peptides were synthesized in which glycine was used to lengthen the distance between the chromophore and the amino acid residue of interest. Thus model peptides had a structural formula of DNP-Gly-X-NH₂, where X was either Ser, Thr or Gly-Tyr. However, further experiments identified a possible flaw in this model peptide; the amino group of glycine was sensitive to acylation by trifluoroacetic anhydride. Therefore, this secondary amine group was changed to an unreactive tertiary amine by replacing the glycine residue with
a proline. These final alterations resulted in making the model DNP-peptides more representative of hydroxylamino acid residues in a protein.

The syntheses of these DNP-peptides are schematically outlined in Figure 10 and detailed in the Materials and Methods. The syntheses of DNP-prolylserinamide or DNP-prolylthreoninamide model peptides were not performed under anhydrous conditions so as to minimize the undesired dehydration side-reactions of these hydroxylamino acids. The yields for DNP-prolylglycyltyrosinamide and DNP-prolylglycinamide were optimal by performing these syntheses under anhydrous conditions. All DNP-peptides were isolated and purified by reversed phase HPLC and were judged as being better than 99% pure, based on their chromatogram, monitored at 365 nm. These DNP-peptides were found to be excellent model systems for the characterization of the acylation reactions and the accurate assessment of the stability of the resulting esters.

III.4. Trimesylation

A variety of chemicals were considered as possible protein derivatization reagents which would avoid the drawbacks of succinylation. The chemical that was finally chosen was trimesic acid (TM acid; 1,3,5-benzenetricarboxylic acid, Figure 11; I) based on the following considerations: 1. It is sterically impossible to have the intramolecular cyclization which may be responsible for the proposed succinyl ester hydrolysis similar to that observed for succinyl-tyrosine residues (Gounaris and Perlmann, 1967). 2. Aromatic esters have greater stability over aliphatic esters due to the electronic effect of the aromatic ring. 3. The two extra carboxylate groups on the trimesic acid molecule further increase the electron density of the aromatic ring, which would enhance the stability of the ester linkage. 4. The
Figure 10: Model peptide synthesis. This is a schematic representation of the reactions employed for the synthesis of the model DNP-peptides in which Im denotes an imidazolyl group and HCl and NH₂ represent the hydrochloride salt and the amide derivative of the corresponding amino acid, respectively.
Figure 11: Synthesis of diprotected trimesic acid. In this schematic representation of the synthetic reactions, Im represents an imidazoly1 group and \( \text{Me}_3\text{SiCH}_2\text{CH}_2 \) denotes a 2-(trimethylsilyl)ethyl group.
bulky benzene ring would sterically hinder the hydrolysis of the ester. Therefore, it was expected that upon protein derivatization with appropriately activated trimesic acid, the resulting trimesyl esters, both aliphatic and aromatic, were expected to be extremely stable. It was expected that the incorporation of two negative charges by this reagent at each site of modification (lysine, serine, threonine, and tyrosine residues) would render even the most hydrophobic membrane protein completely water soluble.

In order to develop a useful reagent for derivatization, several methods of activating the trimesic acid were studied and the following complications had to be considered: 1. The multifunctional nature of the reagent posed some problems. Activation of only one or two carboxyl groups resulted in extensive intermolecular crosslinking; formation of polymeric anhydrides were observed with the reagent itself as well as with the protein and model peptide substrates. This indicated the need to protect the extra carboxyl groups. 2. The same electronic factors responsible for the increased stability of aromatic esters also result in decreasing the reactivity of the activated carboxyl group; the reduction of the partial positive charge on its carbonyl carbon would mean that it is less reactive towards nucleophiles. Therefore, the extra carboxylate groups on the ring must have their negative charge temporarily eliminated, while the acylating carboxyl group must be activated with a strong electron withdrawing group. When these criteria are met, a substantial partial positive charge will be generated on the carbonyl carbon of the activated carboxyl group, thereby enhancing the acylating potential of the reagent (increased reactivity towards nucleophiles). After derivatization the extra carboxylate groups should be regenerated in order to stabilize the desired trimesyl product.

Activating TM acid to its tri-imidazolide derivative (Figure 11; II)
(Staab, 1962) was initially investigated and even though it was found to have the requisite activity with respect to the rate of acylation of the model peptides (data not shown), it was not considered to be a useful reagent. During protein derivatization, the multifunctional nature of the reagent could result in both intramolecular and intermolecular crosslinking. Furthermore, the presence of any trace amine contaminants could also result in reagent polymerization and possibly protein crosslinking.

The trimesic acid would be a potentially effective reagent if it was converted from a multifunctional to a monofunctional reagent by chemically blocking two of the carboxyl groups, in a reversible manner. This could be achieved by converting two of the three carboxylate groups of TM acid into esters. The trimethylsilyl ethyl ester was chosen as the protecting group for the following reasons: 1. The formation of the ester linkage is easily achieved by the addition of 2-(trimethylsilyl)ethanol to the activated imidazolide derivative of the carboxyl group (Figure 11). 2. The resulting ester derivative would be very hydrophobic and would enhance the solubility of the reagent in the anhydrous organic solvent system. 3. This protecting group can be quantitatively removed by naked (non-hydrated) fluoride ions, preferably as a quaternary ammonium fluoride, such as tetrabutylammonium fluoride (Bu₄N⁺F⁻), in anhydrous DMF or tetrahydrofuran (THF) (Sieber, 1977 and Gerlach, 1977). This deprotection reaction is essentially instantaneous in the presence of at least two equivalents of Bu₄N⁺F⁻, resulting in the regeneration of the carboxylate groups (Lipshutz and Pegram, 1980).

Figure 11 illustrates the reaction scheme that was used for the synthesis of the diprotected trimesic acid (DT-TM acid). All of the carboxyl groups of the trimesic acid were initially activated by the addition of 3 equivalents of 1,1'-carbonyldiimidazole (CDI) in dry DMF (Staab, 1962). To
the resulting tri-imidazolide derivative of trimesic acid, 2.2 equivalents of 2-(trimethylsilyl)ethanol was added to form mostly the di-protected reagent, although, some mono- and tri-protected reagent is also generated. In order to purify DT-TM acid, the activating group is hydrolyzed off by a basic aqueous quench and the required diprotected trimesic acid (IV) is then easily obtained by a selective extraction procedure, as outlined in Materials and Methods. The selective ether extraction affords pure diprotected trimesic acid, as determined by reversed phase HPLC or straight phase thin layer chromatography. The purified DT-TM acid was subjected to nmr analysis and was found to have the requisite ratio of protons; nmr δ (TMS): 0.5 (s, 18H, (CH$_3$)$_3$-Si-), 1.0 (t, 4H, -Si-CH$_2$-), 3.2 (t, 4H, -O-CH$_2$-), 7.0 (s, 3H, aromatic). This result confirms the presence of two trimethylsilyl ethyl groups for every trimesic acid molecule. The isolated product yield for the DT-TM acid was found to be 55 - 65%. The DT-TM acid was repeatedly dissolved and dried under reduced pressure from anhydrous benzene and then stored under N$_2$ as a 0.5 M solution in dry benzene.

III.4.A. DT-TM Imidazolide

The procedure that was initially investigated for the activation of the carboxylate group of the DT-TM acid, involved the utilization of 1,1'-carbonyldiimidazole (CDI) to generate the diprotected trimesic acid imidazolide (DT-TM imidazolide, Figure 11; III). This reactive species was reported to be an excellent acylation reagent that was capable of reacting with amine groups within one to two hours at room temperature in inert solvents, such THF or DMF (Staab, 1962). The formation of esters was much slower but the rate can be substantially improved by the addition of catalytic amounts of sodium ethoxide, sodamide or any reasonably strong inert base.
such as triethylamine (Et$_3$N, pK$_a$ = 11.01, CRC Handbook) which can result in the ionization of the hydroxyl group of the corresponding alcohol, either directly or indirectly, thereby increasing the nucleophilicity of the alcohol. In this manner, the esterification of the secondary hydroxyl group of a threonine residue was accomplished (Staab, 1962).

The proposed reaction scheme involves the reaction of one equivalent of CDI with DT-TM acid resulting in the liberation of CO$_2$ and the formation of the derivatization reagent, DT-TM imidazolide as well as one equivalent of imidazole. This activation system was selected for the following reasons: 1. The purity of the CDI can be quantitatively assessed. 2. The activation process is rapid and familiar. 3. The resulting DT-TM imidazolide should not require any purification prior to its use as an acylating reagent.

III.4.A.i. DT-TM Imidazolide Formation

Initial studies indicated that the addition of only one equivalent of reactive CDI, as determined by titration with trimesic acid, was insufficient to achieve complete activation of DT-TM acid to the corresponding imidazolide. The suspected side reaction was the formation of DT-TM anhydride due to the limiting nature of the CDI. Therefore, in order to optimize the conditions of activation, the procedure was assessed with respect to the solvent, the presence of Et$_3$N, the molar ratio of CDI to DT-TM acid, and the number of equal portions in which the DT-TM acid was divided and added sequentially to the CDI solution.

This analysis involved an indirect assay in which DT-TM acid was activated according to the conditions specified in Table 10, followed by the addition of 10 equivalents of cyclohexylamine with respect to the amount of
Table 10: Formation of diprotected trimesic imidazolidine. Diprotected trimesic acid was activated by its addition to a solution of CDI under the conditions specified in the table. The activation was allowed to proceed at room temperature for a period of 1 h after the last addition of reagent. Cyclohexylamine (10 molar excess to CDI) was added and left for an additional 0.5 h. Aliquots of 100 nmol were analyzed by HPLC at 280 nm on a Waters μBondapak C_{18} reversed phase column (0.46 x 25 cm) that was equilibrated at 20% methanol and 80% 10 mM aqueous sodium acetate, pH 4.4 and was developed with a methanol gradient from 20% to 100% over 10 minutes. The amount of DT-TM anhydride generated was assessed by normalizing the integrated area corresponding to DT-TM acid in the chromatogram, to the total area associated with DT-TM acid (i.e., the cyclohexylamide derivative of DT-TM acid).
CDI utilized in the activation. The cyclohexylamine will react with DT-TM imidazolide to form the corresponding amide derivative. This base will also react with any DT-TM anhydride that is present, resulting in the formation of one equivalent of the amide derivative and one equivalent of an underivatized molecule of DT-TM acid. These products were isolated and quantitated by applying an aliquot of the reaction mixture to a reversed phase HPLC column that was developed according to the specifications in the legend to Table 10 while monitoring the effluent at 280 nm. The proportion of underivatized DT-TM acid in the sample would correlate to the amount of DT-TM anhydride that was formed under the activation conditions. The percentage of DT-TM imidazolide formed could then be quantitated by multiplying the percentage of DT-TM anhydride formed by two (to account for the two molecules of DT-TM acid) and then subtracting it from 100.

These results (Table 10) clearly show that DMF is a relatively poor solvent for CDI activation of DT-TM acid because no matter how the procedure was varied, the formation of DT-TM anhydride could not be prevented. The presence of Et$_3$N was found to double the amount of DT-TM anhydride formed as compared to that obtained in its absence; apparently, the DT-TM carboxylate ion undergoes a nucleophilic attack on the DT-TM imidazolide thereby forming the anhydride. Furthermore, increasing the CDI or the number of additions of DT-TM acid, in equal portions, was found to make little difference in reducing the anhydride formation. However, in the benzene/DMF solvent mixture (3:1, v/v) and in the absence of Et$_3$N, the formation of DT-TM imidazolide was 100%, irrespective of how the DT-TM acid was added, all at once or in 4 equal aliquots. Furthermore, the activation reaction was found to be complete within 0.5 h. Thus the best procedure for the complete activation of DT-TM acid by CDI to its imidazolide derivative
requires a fairly non-polar solvent system and a slight excess of CDI. Complete reagent activation could still be achieved in 0.5 h when the amount of CDI was reduced to 1.2 equivalents (data not shown); further reductions in CDI resulted in reagent incompletely activated due to the limiting nature of the CDI.

However, it was observed that residual CDI after DT-TM acid activation, reacted with the site of acylation (hydroxyl group of Ser) thereby effectively blocking the desired trimesylation reaction (see Figure 14). Removal of the excess CDI would ensure complete reagent activation and permit acylation of all the desired functional groups. If a reasonable difference in the reactivities of CDI and DT-TM imidazolide exists, the excess CDI could be inactivated selectively while the majority of the DT-TM imidazolide would still be present.

In order to assess the reactivities, the reaction of CDI and DT-TM imidazolide towards methanol was analyzed in different anhydrous solvents. The reactions were monitored spectrally using two wavelengths, one for each of the activated molecules of interest, as outlined in the legend to Figure 12. The following rate constants were calculated from the results in Figure 12: CDI had a rate constant of $9.7 \times 10^{-2} \text{ M}^{-1}\text{s}^{-1}$ when the reaction was performed in benzene but only $2.4 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1}$ in DMF. The rate constants for DT-TM imidazolide were $5.3 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ in benzene and $9.7 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ in DMF. It appears that a less polar solvent increases the reactivity of CDI towards a primary alcohol by 40 fold while the reactivity of DT-TM imidazolide decreases by a factor of 1.8. Therefore, CDI has a rate constant that is 175 times that of DT-TM imidazolide towards MeOH when in benzene; this large difference may permit the selective removal of the CDI from the desired DT-TM imidazolide.
Figure 12: Reactivity of CDI and DT-TM imidazolide. HPLC grade methanol was added to a 6.7 mM solution of CDI (triangles) or DT-TM imidazolide (squares) in DMF (open symbols) or benzene (closed symbols) to give a final concentration of 100 mM. The amount of the activated species remaining over time was monitored spectrally on a Varian Cary 210 Spectrophotometer, 290 nm for CDI ($\epsilon_{290}= 280 \text{ M}^{-1}\text{cm}^{-1}$) and 310 nm for DT-TM imidazolide ($\epsilon_{310}= 210 \text{ M}^{-1}\text{cm}^{-1}$).
The reaction of an alcohol (R-OH) with the excess CDI, after the activation of the DT-TM acid would result in a product (R-O-CO-Im) which would still be reactive and soluble in benzene. Sephadex LH-20 (Pharmacia) was considered because it is composed of a glucose matrix that has undergone hydroxypropylation and is able to swell in anhydrous solvents. Therefore, Sephadex LH-20 would provide a source of hydroxyl groups with which the excess CDI could react, even though it would be a two phase reaction. The resulting product, though still reactive, would be covalently attached to the gel matrix which could be pelleted by centrifugation. This would afford a solution of the DT-TM imidazolide which was entirely devoid of the CDI.

The rate at which CDI and DT-TM imidazolide react with swollen Sephadex LH-20 in benzene was determined by reacting them individually in the presence of the gel matrix and assessing the absorbance of an aliquot of the supernatant at the appropriate wavelengths, as specified in Figure 12. The rate constants that were obtained were $5.3 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1}$ for CDI and $0 \text{ M}^{-1}\text{s}^{-1}$ for DT-TM imidazolide (Figure 13). This procedure provides an excellent method for the complete removal of excess CDI within one hour, while maintaining 100% of the DT-TM imidazolide in the solution.

In order to illustrate the problems associated with CDI activation of DT-TM acid and the effectiveness of the Sephadex LH-20 procedure to remove contaminating CDI, the time course of DNP-prolylserinamide acylation was assessed when Sephadex LH-20 treated or untreated DT-TM imidazolide was utilized (Figure 14). The activation of the DT-TM acid with 1.0 or 1.5 equivalents of CDI results in acylation which rapidly levels off at a low product yield. Analysis of the reaction products by HPLC (Figure 15) indicated the formation of an adduct between CDI and DNP-prolylserinamide that occurs in the presence of DT-TM acid that was activated with 1.5
Figure 13: Reactivity of CDI and DT-TM imidazolide with Sephadex LH-20. A 20 mM solution of CDI (▲) and a 40 mM solution of DT-TM imidazolide (■) in anhydrous benzene was reacted with 4 equivalents of dry Sephadex LH-20 (1.15 x 10^-5 hydroxyl groups per dry mg based on CDI titration). At the specified times the LH-20 was pelleted and an aliquot of the solution was removed, diluted to 4 mM with benzene and the remaining activated species determined spectrally, as described in Figure 12.
Figure 14: Time course of DNP-prolylserinamide trimesylation with Sephadex LH-20 treated or untreated DT-TM imidazolide. DNP-prolylserinamide was trimesylated with reagent imidazolide that had been generated with 1.0 equivalent (triangles) or 1.5 equivalents (squares) of CDI and then reacted for 1 hour in dry benzene with (closed symbols) or without (open symbols) Sephadex LH-20 (4 equivalents per equivalent of CDI). The extent of DNP-peptide acylation at specific times was determined by HPLC analysis of an aqueous DMF quench of an aliquot of the reaction, as described in Figure 22.
Figure 15: HPLC analysis of the reaction products. DT-TM imidazolide was generated in the presence of 1.5 equivalents of CDI and subsequently treated with or without 4 equivalents of Sephadex LH-20 in dry benzene, as described in Figure 13. DNP-prolylserinamide was reacted under standard reaction conditions without reagent (A) or with untreated 200 mM DT-TM imidazolide (B), 50 mM CDI (C) or 200 mM Sephadex treated DT-TM imidazolide. After one hour, aliquots of the reactions were diluted with aqueous DMF and assessed by HPLC according to Figure 22.
equivalents of CDI (panel B) and when the model peptide is reacted with CDI alone, under the reaction conditions (panel C). The excess CDI is thus preventing peptide acylation by DT-TM imidazolide. It was found that when the excess CDI was first removed by reacting the activation mixture with Sephadex LH-20 in dry benzene (as described in Figure 13), the CDI-peptide adduct is absent from the HPLC chromatogram and the acylation reaction goes to completion (Figure 15; panel D).

Furthermore, increasing the CDI from 1.0 to 1.5 equivalents results in lower reaction rates and product yields, however, if the excess CDI is first removed by the LH-20 procedure, a higher reaction rate is obtained. The absence of CDI-peptide adducts and the fact that the reaction proceeds to 100%, in both of these cases indicates that the slightly lower rate observed in the case of 1.0 equivalent of CDI, is likely due to incomplete formation of the DT-TM imidazolide and hence a lower reagent concentration. Therefore, this procedure ensures the complete formation of DT-TM imidazolide by using excess CDI in a benzene/DMF solvent mixture, followed by the reaction of the reaction mixture with Sephadex LH-20 to selectively remove the residual CDI.

III.4.A.ii. DT-TM Imidazolide Stability

The DT-TM imidazolide was anticipated to be a stable reagent under the acylation conditions because both the solvent, DMF and the base, Et$_3$N were considered to be chemically inert. Reagent stability would be important for complete protein derivatization based on the observations for the time course of reaction observed for the model peptides (see Figure 20). However, the DT-TM imidazolide was not found to be stable by itself in DMF and Et$_3$N and therefore, the cause of the instability was investigated.

During the activation reaction of DT-TM acid to its imidazolide
derivative, an equivalent of imidazole is generated and remains with the acylating reagent. This side product is the only other component present in the reaction mixture and it was reasoned that this weak base may have some effect on the stability of the reagent over time, under the acylation conditions. In the presence of Et₃N, the imidazole would become deprotonated resulting in the possibility of increasing the nucleophilicity of one of its carbon atoms which upon reacting with DT-TM imidazolide would result in a loss of the activated reagent. Therefore, a means to remove the imidazole was considered.

The use of Dowex 50x2 (μparticle) was considered as a viable procedure for the following reasons: The Dowex 50x2 could be added to the benzene solution containing DT-TM imidazolide after the Sephadex LH-20 had been removed. It would react with the imidazole, in an acid base type reaction, resulting in a strong ionic interaction. The Dowex 50x2 could then be pelleted by centrifugation, in a similar manner as the Sephadex LH-20, thus affording the removal of the imidazole from the solution. This procedure would result in the effective purification of DT-TM imidazolide and would be beneficial as long as the DT-TM imidazolide was essentially unreactive towards the Dowex 50x2.

The Dowex 50x2 was initially titrated with imidazole and was found to be capable of removing 4.74 μequivalents of imidazole from a benzene solution per dry mg. In order to determine the usefulness of this procedure, imidazole, CDI and DT-TM imidazolide were reacted in benzene with the Dowex 50x2 (Figure 16). As observed by residue weight (detailed in Figure 16 legend), the removal of imidazole was effectively achieved within 30 minutes, when two equivalents of Dowex 50x2 were
Figure 16: Dowex-50x2 removal of imidazole. Imidazole (▲), CDI (●) or DT-TM imidazolide (■) were reacted with 2 equivalents of μparticle Dowex 50x2 (4.74 μequivalents per dry mg based on imidazole titration) in benzene. At the specified times, the Dowex 50x2 was pelleted and an aliquot of the imidazole solution was dried under reduced pressure and the weight of the dry residue was determined. For CDI and DT-TM imidazolide, an aliquot of the solutions were diluted to 4 mM and assessed spectrally, as described in Figure 12.
added to the solution of imidazole. Spectral assessment indicated that there was only a slight loss of CDI during this time period; CDI would be more efficiently removed via the Sephadex LH-20 procedure. There was essentially no loss of the DT-TM imidazolide over the time period studied, based on its 310 nm absorbance. Therefore, the particle Dowex 50x2 proved to be an effective way in which imidazole could be removed with no appreciable loss of the desired DT-TM imidazolide. This procedure for imidazole removal enables the stability of the DT-TM imidazolide to be assessed in the presence or absence of imidazole. This will ensure that the reagent remains reactive over the required time period to achieve complete trimesylation of the secondary hydroxyl groups and hence for all of the desired functional groups.

The DT-TM imidazolide was added to benzene and to 1 M Et$_3$N in DMF, which were the conditions in which it was stored as a stock solution and in which it was used during the acylation reaction, respectively. As indicated by Figure 17, DT-TM imidazolide was completely stable for over 60 hours in anhydrous benzene, in the presence or absence of imidazole. However, its stability under the acylation conditions was relatively poor, with a loss of 60% in the initial rate constant within the first 12 hours, with or without imidazole present in the solution. The initial rate constant for trimesylation was found to continually decrease to less than 10% of the optimal within 48 hours.

The stability of DT-TM imidazolide was also assessed over time, under the acylation conditions established in anhydrous Me$_2$SO (Figure 18); the concentration of 0.8 M Et$_3$N was found to be optimal as well as being the maximum concentration obtainable in anhydrous Me$_2$SO. The DT-TM imidazolide was also found to be unstable under these conditions; regardless whether imidazole was present or absent, the rate of loss in the initial rate
Figure 17: Stability of diprotected trimesic imidazolide. Diprotected trimesic imidazolide was dissolved in benzene (squares) or in 1 M Et₃N in DMF (triangles) before (closed symbols) and after (open symbols) the activated reagent was treated with 4 equivalents of µparticle Dowex 50x2. At the specified time, an aliquot of the activated reagent was used to trimesylate DNP-prolylserinamide. The initial rate constant was then calculated from the HPLC analysis of the acylation reaction (as outlined in Figure 22) and plotted against the time of reaction as a percentage of the rate constant observed prior to the reaction.
Figure 18: Stability of diprotected trimesic imidazolide in Me₂SO. DT-TM imidazolide was treated with (closed squares) or without (open squares) 4 equivalents of μparticle Dowex 50×2 in benzene prior to its addition in Me₂SO at a concentration of 0.2 M, in the presence of 0.8 M Et₃N. At the indicated times an aliquot of the DT-TM imidazolide was added to DNP-prolylserinamide and the subsequent trimesylation reaction was monitored by HPLC as outlined in Figure 22. The initial rate constants were calculated and plotted against the corresponding time of reaction as a percentage of the initial rate constant at zero reaction time.
constant of trimesylation in Me₂SO with 0.8 M Et₃N, was found to be approximately the same as that found for 1 M Et₃N in DMF.

The instability of DT-TM imidazolide was found to be concentration independent in either DMF or Me₂SO. The reagent appears to react with both of these solvents; the DT-TM imidazolide may react with DMF and Me₂SO in a manner analogous to the Vilsmeier reaction and the Moffat oxidation (Pfitzner and Moffat, 1965), respectively.

Further investigation into the DT-TM imidazolide instability was not done on the basis that complete trimesylation of the model substrate, DNP-prolylthreoninamide could be achieved regardless of whether the acylation reaction was performed in DMF (see Figure 20) or Me₂SO (see Figure 24). However, the acylation of the secondary hydroxyl groups of a protein may require a substantially longer time than the model peptides, due to possible steric hindrance caused by the folding of the protein under these anhydrous solvent conditions. The length of time required for complete protein trimesylation may present problems with respect to DT-TM imidazolide instability and incomplete derivatization. In order to circumvent this postulated problem, higher concentrations of DT-TM imidazolide were utilized for the acylation of proteins; this was done on the basis of increasing the initial rate of reaction and to ensure reagent reactivity over the required time of reaction. It was found that a concentration of 0.8 M DT-TM imidazolide could be achieved in anhydrous Me₂SO if the imidazole was first removed by the μparticle Dowex 50x2 procedure and the temperature was increased to 45°C.

III.4.A.iii. Characterization of the Trimesylation Reaction

The reaction mechanism by which DT-TM imidazolide acylates a
hydroxyl group involves the nucleophilic attack by the hydroxyl oxygen on the activated carbonyl of the imidazolide derivative. Therefore, it was anticipated, that as the Et₃N concentration was increased, the rate of trimesylation would also increase due to the shift in the equilibrium in favour of the polarization of the hydroxyl groups resulting in an increase in their nucleophilicity. A maximum rate of acylation would be reached with Et₃N because its pKₐ is only 11 while the pKₐ of the Ser and Thr hydroxyl groups are approximately 15 and 16, respectively.

Therefore, the optimal concentration of Et₃N with respect to the rate of trimesylation of DNP-prolylserinamide with DT-TM imidazolide was assessed by calculating the initial rate constant for the acylation reaction for each concentration of base analyzed (Figure 19). The initial rate constant for trimesylation was found to increase as the Et₃N concentration increased and as anticipated, it appeared to level off at a maximal value of 1.7 x 10⁻² M⁻¹s⁻¹ as the Et₃N concentration approaches 2 M; at this concentration, 28% of the volume of the solvent is Et₃N which may significantly alter the solvent properties of DMF. At 1 M Et₃N, the rate constant is 1.4 x 10⁻² M⁻¹s⁻¹ which is 84% of the maximum while the volume of the Et₃N is half of that required to achieve the maximal rate constant. Therefore, 1 M Et₃N was routinely used as the standard amount of base for the trimesylation reaction.

Under the reaction conditions of 1 M Et₃N in anhydrous DMF, the time course of trimesylation of each of the model peptides with DT-TM imidazolide was assessed, in the presence and absence of the amphipathic salt Bu₄N⁺Cl⁻. Bu₄N⁺Cl⁻ was utilized as the phase transfer reagent that improves the solubilization of apobacteriorhodopsin in anhydrous DMF (section III.1.). This experiment would establish the length of time required for the complete derivatization of aromatic, primary and secondary hydroxyl groups in the
Figure 19: Effect of Et$_3$N on the rate of trimesylation with DT-TM imidazolide in DMF. DNP-prolylserinamide (1 mM) was trimesylated in DMF in the presence of 0.2 M DT-TM imidazolide and varying concentrations of Et$_3$N. The acylation reaction was monitored with time by HPLC (as described in Figure 22). The initial rate constant was calculated and plotted against the corresponding Et$_3$N concentration.
model peptides.

As observed in Figure 20, complete trimesylation of the aromatic hydroxyl group of DNP-prolylglycyltyrosinamide was accomplished within 10 minutes, while the primary and secondary hydroxyl groups of DNP-prolylserinamide and DNP-prolylthreoninamide were achieved within 1 hour and 10 hours, respectively. The ten fold difference in the time required for complete derivatization between the primary and secondary hydroxyl groups was expected on the basis of the difference in their pKₐ values and steric hinderance. The model peptide, DNP-prolylglycinamide was not derivatized under these conditions, indicating that there is no other nucleophilic site that exists on the model peptides except for the hydroxyl functional group. Under these conditions, complete and fairly rapid acylation of even the secondary hydroxyl groups on the model peptides was achieved.

The presence of the Bu₄N⁺Cl⁻ salt was found to increase the rate constant for DNP-prolylserinamide from 6.9 x 10⁻³ to 1.4 x 10⁻² M⁻¹s⁻¹. Furthermore, the rate constant was observed to remain at ~1.4 x 10⁻² M⁻¹s⁻¹ even when the Bu₄N⁺Cl⁻ concentration was increased to 100 mM or to 200 mM. This increase in the rate constant may reflect the increase in the polarity of the solvent system which results in the destabilization of the reaction intermediate.

On the basis of these model peptide results, it was anticipated that for protein trimesylation longer than 10 hours will be required to achieve complete derivatization due to possible steric hinderance that may arise from the folding of the protein; there may be acylation sites that are buried within the protein that are less accessible to the DT-TM imidazolide. Furthermore, during this long reaction time there will likely be a substantial loss in the concentration of the DT-TM imidazolide and hence the rate of reaction will
Figure 20: Time course of model peptide trimesylation with DT-TM imidazolide in DMF. DNP-prolylserinamide (squares), DNP-prolylthreoninamide (triangles), DNP-prolylglycinamide (circles) and DNP-prolylglycyltyrosinamide (inverted triangles, inset) were trimesylated at 1 mM with 0.2 M DT-TM imidazolide in 1 M Et₃N in anhydrous DMF in the presence (closed symbols) and absence (open symbols) of 50 mM Bu₄N+Cl⁻. At the indicated times, an aliquot of the reaction was diluted into aqueous DMF and analyzed for the extent of acylation by HPLC, as described in Figure 22.
also diminish over time (Figure 16). However, it was thought that as the derivatization proceeds, the protein would tend to become more unfolded due to the incorporation of the hydrophobic protected trimesyl group thereby aiding the derivatization of all possible sites of acylation. Therefore, the DT-TM imidazolide concentration was increased to 0.4 M for the derivatization of the protein to ensure a large initial rate of trimesylation and to maintain a higher level of reactivity over the same period of time.

As stated previously, the selection of the trimethylsilylethyl ester as the protecting group of two of the carboxylates of trimesic acid was on the basis of the selective deprotection reaction. Bu₄N⁺F⁻ reacts specifically with 2-(trimethylsilyl)ethyl groups resulting, in the case of ester deprotection, the formation of Me₃SiF, ethylene and the regeneration of the carboxylate group, as the Bu₄N⁺ salt (Lipshutz and Pegram, 1980). In order to determine the number of equivalents of Bu₄N⁺F⁻ that were required for the complete deprotection of the Me₃SiCH₂CH₂ ester, an acylation reaction of DNP-prolylthreoninamide was deactivated by dilution into Me₃SiCH₂CH₂OH, followed by a reaction with varying equivalents of Bu₄N⁺F⁻, based on the total Me₃SiCH₂CH₂ groups present (Figure 21). This permitted the deprotection reaction to be monitored by following the simultaneous deprotection of the trimesylated DNP-prolylthreoninamide at 365 nm.

The deprotection of the carboxylate groups, was observed indirectly through the conversion of the hydrophobic trimesylated DNP-peptide to a substantially more polar product (Figure 22). Three equivalents of Bu₄N⁺F⁻ per Me₃SiCH₂CH₂ ester were found to be the minimum requirement for complete deprotection (Figure 21). Due to the relatively small volumes that were used for this assay and the larger volumes generally needed for protein trimesylation, the possibility of adversely affecting the
Figure 21: Trimethylsilylethyl ester deprotection. DNP-prolylthreoninamide (1 mM) was trimesylated under the standard reaction conditions, as described in Figure 20. After acylation was complete, DT-TM imidazolide was reacted with Me₃SiCH₂CH₂OH (0.4 M) for 1 hour. Aliquots of this reaction were then added to varying amounts of Bu₄N⁺F⁻, left at room temperature for 15 minutes and then analyzed by HPLC for the extent of deprotection; the HPLC conditions used were those described in Figure 22.
Figure 22: Tritiated trimesylation of DNP-prolylserinamide with DT-TM imidazolide and subsequent Bu₄N+F⁻ deprotection. Panels A, B, and C are representative HPLC chromatograms of DNP-prolylserinamide at different stages of the trimesylation reaction that were obtained by applying samples containing 10 nmol of the DNP-peptide onto a µBondapak C₁₈ reversed phase column (0.46 x 25 cm) that was equilibrated with 35% methanol and 65% 10 mM aqueous sodium acetate, pH 4.4 and developed as indicated in panel A. The behaviour of the pure DNP-prolylserinamide prior to trimesylation is shown in panel A. The chromatogram, in panel B, was obtained after complete acylation of the DNP-peptide under the conditions used in Figure 20 and following the inactivation of the DT-TM imidazolide with 0.4 M Me₃SiCH₂CH₂OH. Fractions at 0.8 minute intervals were collected, diluted with Amersham aqueous counting scintillant and the radioactivity determined using a Beckman LS-7800 scintillation counter; the results shown were obtained when the acylation reaction was performed in the presence (○—○) and absence (○—○) of DNP-prolylserinamide. Panel C was obtained after deprotection of the trimesylated DNP-peptide with Bu₄N+F⁻. As in panel B, the extent of radioactivity was determined in fractions that were collected at 0.8 minute intervals. The symbols i, ii, iii, and iv in panels B and C indicate the relative positions at which TM acid, MT-TM acid, DT-TM acid and TT-TM acid respectively are eluted under these HPLC conditions.
Diagram showing absorbance and percent methanol over time.
deprotection reaction by the presence of trace amounts of contaminants (H₂O) was minimized by the general utilization of 5 equivalents of Bu₄N⁺F⁻ per Me₂SiCH₂CH₂ ester for the standard deprotection reaction.

The synthesis of [³H]trimesic acid was performed in order to provide radioactive reagent for the direct evaluation of the model peptide and the protein derivatization. It had been shown that the aromatic protons of benzoic acid can be exchanged with water at 275°C under acid catalysis (Werstiuk and Kadai, 1973 and Werstiuk and Kadai, 1974); this method was therefore used to tritiate the trimesic acid, as described in Materials and Methods. Recrystallization of the product from a minimum volume of boiling water afforded the tritiated trimesic acid. Subsequent analysis by reversed phase HPLC indicated a single peak of radioactivity co-eluting with authentic trimesic acid (data not shown).

The synthesis and acylation of model peptides was routinely monitored by reversed phase HPLC. The use of DNP-peptide derivatives permitted detection of all peptide-containing products formed and the integration of chromatograms monitored at 365 nm was used as the basis of all quantitation. The purity of each of the synthetic peptides was assessed by HPLC, as shown in Figure 22A for the DNP-prolylserinamide; the peptides had all been purified by preparative HPLC and were thus better than 99% pure as judged by their elution profile monitored at 365 nm.

The characterization of the acylated model peptide, DNP-prolylserinamide, and its subsequent deprotection was performed with the tritiated DT-TM imidazolide reagent in DMF in the presence of 1 M Et₃N. The reactions were monitored by analyzing an aliquot of the appropriate reaction on reversed phase HPLC, detecting the absorbance of the effluent at 365 nm and assessing the radioactivity of fractions of the collected effluent, as
outlined in the legend of Figure 22. The trimesylation reaction resulted in the complete disappearance of the starting material and the concomitant formation of a single very hydrophobic product (Figure 22B). Although the results are shown only for the serine peptide, comparable results were observed for all three model peptides using identical procedures and the same analytical system.

The fact that the new hydrophobic product was trimesylated was demonstrated in two ways. First, the hydrophobic peptide product was very sensitive to Bu₄N⁺F⁻, as shown in Figure 22C; since this is a highly selective reagent for the removal of the trimethylsilyl ethyl blocking group and a brief reaction with this reagent converted the hydrophobic product to a polar one, it was concluded that the blocking groups were in fact attached to the peptide. The results show that under the standard reaction conditions, the conversion was essentially complete with no protected starting material remaining (Figure 22C). The use of only 0.5 equivalent of the Bu₄N⁺F⁻ in this reaction resulted in partial deprotection, generating an intermediate product having a retention time of 8.6 minutes; the addition of more Bu₄N⁺F⁻ resulted in the complete conversion of this product to one that eluted at 5.3 minutes. The conclusion was that the intermediate had one of the trimethylsilyl ethyl groups still remaining and that the hydrophobic product eluted at 15.9 minutes was the desired product, the DNP-prolylserinamide acylated by the diprotected trimesic acid.

Secondly, in order to confirm the trimesylation of the serine peptide, the effluent from the HPLC reversed phase column was assessed for radioactivity (Figure 22B). In this case, after complete acylation, the excess reagent, having a specific activity of 0.65 mCi/mmol, was quenched by the addition of 2-(trimethylsilyl)ethanol. The figure indicates that in the case of
the reagent blank (open circles) this resulted in the complete conversion of
the reagent to the triprotected derivative of TM acid (TT-TM acid). In the
presence of peptide, a single new peak of radioactivity was observed (closed
circles) to coincide with the new peptide product formed. The amount of
radioactivity in this peak was approximately $9 \times 10^3$ cpm which corresponds to
$1.38 \times 10^4$ dpm for 10 nmoles of peptide or 0.63 mCi per mmole of peptide.
This confirms the incorporation of one mole of trimesic acid per mole of
DNP chromophore into this product. These results thus demonstrate that the
hydrophobic product obtained contains one molecule of trimesic acid and two
molecules of the trimethylsilylethyl substituents per molecule of peptide and
was the desired trimesylated model peptide.

The deprotection of the radioactive trimesylated peptide is shown
in Figure 22C. The entire quenched reaction mix depicted in panel B was
subjected to Bu$_4$N$^+$F$^-$ treatment in order to remove the protecting groups. In
the case of the reagent blank (open circles), the TT-TM acid disappeared
completely and a single product was eluted in the position of trimesic acid. In
the DNP-prolylserinamide case (closed circles), the hydrophobic peptide
product completely disappeared and a new polar product was observed, in
addition to the expected trimesic acid. The amount of radioactivity found in
this peak (approximately $8 \times 10^3$ cpm) corresponded exactly to the attachment
of a single trimesyl group. Comparable results were observed for DNP-
prolylthreoninamide and DNP-prolylglycylytyrosinamide.

These results indicate that the DT-TM imidazolide readily acylates
peptide hydroxyl groups and subsequent treatment with Bu$_4$N$^+$F$^-$ results in
the complete removal of the trimethylsilylethyl groups. This method
produced the desired peptide product that was acylated by trimesic acid; the
peptide now contained two carboxylate ions where there was only one
hydroxyl group previously. These results would imply that upon trimesylation of a protein with DT-TM imidazolide and subsequent deprotection with Bu₄N⁺F⁻, the protein would be extensively derivatized at all of its amine, aromatic hydroxyl, and aliphatic hydroxyl groups; the presence of these carboxylate groups would greatly increase the protein's overall negative charge density which should render even the most hydrophobic membrane protein water soluble.

There was some uncertainty as to the behaviour of the side chain carboxylate groups of Asp and Glu during the trimesylation reaction with respect to their possible activation through the formation of a mixed anhydride. The potential problem of protein crosslinking, either intramolecularly or intermolecularly, is dependent on which of the two carbonyl groups of the mixed anhydride is most electrophilic. Therefore, the formation and the behaviour of the mixed anhydride between the carboxyl group of DNP-proline and DT-TM imidazolide was assessed and monitored by reversed phase HPLC at 365 nm (according to the legend of Figure 23).

DNP-Pro was reacted with DT-TM imidazolide in the presence or absence of 1 M Et₃N in DMF and subsequently diluted by the addition of an equal volume of 1 M Me₃SiCH₂CH₂OH in DMF. Aliquots of the reactions were quenched in H₂O/DMF (1:9, v/v) and assessed by HPLC. Only DNP-Pro was observed in the chromatogram of the reaction containing Et₃N and appeared to remain unaltered even after several hours of reaction. However, in the absence of base, a new hydrophobic DNP containing product was formed and was found to be sensitive to Bu₄N⁺F⁻, resulting in the regeneration of DNP-Pro. This product was concluded to be the trimethylsilyl)ethyl ester derivative of DNP-Pro, indicating that under certain conditions the carboxyl group could be activated to the mixed anhydride and
subsequently esterified. Analysis of the reaction without Et$_3$N before the addition of Me$_3$SiCH$_2$CH$_2$OH yielded only DNP-Pro, implying that the mixed anhydride was hydrolyzed during the aqueous quench prior to it being applied to the HPLC. Furthermore, the addition of Et$_3$N to the DNP-Pro-DT-TM mixed anhydride just before the addition of Me$_3$SiCH$_2$CH$_2$OH prevented the formation of the trimethylsilyl ethyl ester of DNP-Pro. Presumably, the presence of Et$_3$N alters the reactivity of the carbonyl groups in the mixed anhydride such that the DT-TM carbonyl group is the one that becomes esterified resulting in the regeneration of the carboxyl group of DNP-Pro. Figure 23 schematically illustrates the reactions that appear to be involved. The overall conclusion was that under the standard conditions for trimesylation with DT-TM imidazolide (1 M Et$_3$N), protein crosslinking would not be a possible side reaction; if mixed anhydrides on the Asp or Glu side chains formed and underwent nucleophilic attack by an amine or hydroxyl group on the protein, the end result would be the regeneration of the Asp or Glu carboxyl group. Furthermore, these putative mixed anhydrides would be at a relatively low concentration with respect to the 0.4 M DT-TM imidazolide present in the reaction mixture, for competition of any of the nucleophilic sites on the protein.

The investigation into the utilization of Me$_2$SO as the solvent for protein solubilization prompted the assessment of the rate of trimesylation of the model peptides with DT-TM imidazolide in anhydrous Me$_2$SO. The optimal concentration of 0.8 M was used for both Et$_3$N and DT-TM imidazolide, which necessitated the removal of the imidazole (Figure 16) and the reaction performed at 45°C. Under these acylation conditions and in the presence of 50 mM Bu$_4$N$^+$Cl$^-$, complete trimesylation of DNP-prolylglycyl-tyrosinamide was obtained within 1 minute while DNP-prolylserinamide
Figure 23: Proposed equilibria involved with DNP-proline during trimesylation with DT-TM imidazolidine. DNP-proline (4 mM) was reacted with 0.2 M DT-TM imidazolidine in the presence and absence of 1 M Et$_3$N. Aliquots of each reaction were either diluted into 1 M Me$_3$SiCH$_2$CH$_2$OH in DMF or had Et$_3$N added to achieve 1 M, left for 10 minutes and then diluted into the alcoholic DMF solution. Aliquots of these reactions were then analyzed using the HPLC conditions described in Figure 22.
and DNP-prolylthreoninamide required 15 minutes and 3 hours, respectively (Figure 24). The presence of Bu₄N⁺Cl⁻ was found to increase the rate constant of trimesylation for all the DNP-peptides; the rate constant for DNP-prolylserinamide increased from 4.3 x 10⁻³ to 7.9 x 10⁻³ M⁻¹s⁻¹, correlating to a 1.84 fold increase when in the presence of Bu₄N⁺Cl⁻.

Under these conditions, a fairly rapid acylation was obtained of even the secondary hydroxyl group of DNP-prolylthreoninamide. The complete acylation of apobacteriorhodopsin would likely require more than 3 hours due to possible steric hindrance as a result of protein folding and due to the loss in reagent reactivity over time. However, the overall rate may be better than that obtained in DMF because proteins that are dissolved in Me₂SO, appear to be more fully denatured; in DMF, proteins appear to retain a substantial portion of their secondary structure (Llinas et al., 1980).

III.4.A.iv Stability of the Trimesyl Ester

The most critical aspect of trimesylation was the stability of the trimesyl ester product in simple aqueous buffers. It was expected to be much more stable than the succinyl ester for all the reasons presented in the introduction to trimesylation. The most sensitive of the trimesyl esters generated would be the aromatic ester on the side chain of tyrosine, thus, the trimesylated DNP-prolylglycyltyrosinamide was utilized to assess the stability of the trimesyl ester under various solvent conditions, as outlined in the legend to Figure 25.

The aromatic trimesyl ester on the tyrosine side chain was found to be very stable in acidic conditions with no appreciable loss over 10 hours in 60% formic acid or after 2 hours in 25% aqueous TFA at 57°C. It was quite stable in neutral pH conditions and under sequencer conditions;
Figure 24: Time course of trimesylation of model peptides with DT-TM imidazolide in anhydrous Me$_2$SO. DNP-prolylglycylytyrosinamide (squares), DNP-prolylserinamide (triangles), DNP-prolylglycinamide (circles) and DNP-prolylthreoninamide (inverted triangles, inset) were derivatized at a concentration of 2 mM with 0.8 M DT-TM imidazolide and 0.8 M Et$_3$N at 45°C in the presence (closed symbols) and absence (open symbols) of 50 mM Bu$_4$N$^+$Cl$^-$. At the indicated times, aliquots of the reactions were diluted with aqueous DMF (10% H$_2$O) and analyzed for the extent of trimesylation by HPLC using similar conditions as those specified in Figure 22.
Figure 25: Stability of the trimesyl derivative of DNP-prolylglycyltyrosinamide. DNP-prolylglycyltyrosinamide was trimesylated at a concentration of 0.18 M in pyridine with 0.24 M DT-TM anhydride in the presence of one equivalent of PPY. After 6 hours, Me₃SiCH₂CH₂OH was added to achieve 0.48 M and after an additional 6 hours an aliquot was removed and diluted into 20 volumes of 1 M Bu₄N⁺F⁻ in THF. Aliquots of the deprotected mixture was then diluted into 25% aqueous TFA at 57°C (■), 60% formic acid (□), 20 mM Na₂HPO₄, pH 6.8 (▲), Quadrol buffer, pH 9 (▲), 50 mM Tris buffer, pH 8.0 (●), 20 mM NH₄HCO₃, pH 8.0 (○), 50 mM NH₄HCO₃, pH 9.5 (×), 100 mM K₂HPO₄, pH 7.8, 10 mM EDTA and 50 mM β-mercaptoethanol (□), and NH₃OH, pH 10.5 (□). At the indicated times of reaction, the trimesylated DNP-prolylglycyltyrosinamide was quantitated by applying an aliquot to an Altex Ultrasphere ODS reversed phase column (4.6 x 15 cm) that was equilibrated and eluted with 35% methanol and 65% aqueous Na₂HPO₄, pH 6.8.
approximately a 5% loss over 12 hours in 20 mM Na₂HPO₄, pH 6.8 and in Quadrol (pH 9). Under basic conditions, such as pH 8 or pH 9.5 the ester was found to slowly hydrolyze at a rate of 13% and 21%, respectively, over 12 hours. The trimesyl ester was found to be acutely reactive with β-mercaptoproethanol resulting in a 100% loss within 12 hours. Alkaline hydroxylamine at pH 10.5 was found to efficiently cleave the ester with 12 minutes.

The trimesyl ester of DNP-prolylserinamide and DNP-prolylthreoninamide were also assessed and were found to be entirely stable under all conditions studied, except in NH₂OH, pH 10.5, as expected (data not shown). Furthermore, the diprotected trimesyl ester of all the model peptides are stable under the conditions of the Me₃SiCH₂CH₂OH quench of the DT-TM imidazolidone reagent and under the conditions of the trimethylsilyl ethyl ester deprotection.

As anticipated, the trimesyl ester was superior to the succinyl ester, being stable in a variety of different buffers and solvents, especially in acidic and neutral buffers. The trimesyl esters would therefore be stable under the conditions for protein fragmentation by CNBr (70% formic acid), under automated spinning cup sequencer conditions (Quadrol at 57°C and in anhydrous HFBA) and finally under the conversion conditions for the generation of the PTH amino acid derivatives from the anilinothiazolinone derivatives (25% aqueous TFA at 57°C). The ester derivatives of the hydroxyl groups of Ser and Thr were perfectly stable in all the buffers tested while the more sensitive aromatic ester of Tyr was hydrolyzed under basic conditions over several hours and it appeared to be selectively removed by β-mercaptoproethanol. This would imply that under the pyroglutaminate aminopeptidase conditions for the removal of the N-terminal pyroglutamic
acid residue, the Tyr residues would be regenerated. The loss of 11 trimesyl
groups from the modified apobacteriorhodopsin would not be expected to
result in any solubility problem of the membrane protein because the
remaining 38 trimesyl groups would be stable and would still provide a
substantial negative charge density.

III.4.A.v. Apobacteriorhodopsin Trimesylation

After completing the characterization and the optimization of the
conditions required for the acylation reaction with the aid of the model DNP-
peptides, the trimesylation was then performed on the membrane protein,
apobacteriorhodopsin (apoBR). On the basis of its amino acid sequence
(Figure 8), there are 49 possible sites of trimesylation. The tritiated DT-TM
imidazolide was utilized for the acylation of the membrane protein to
provide a means by which the extent of trimesylation could be quantitated by
the amount of radioactivity incorporated into a known amount of protein.

Solubilization of the apoBR was achieved in anhydrous DMF by
using anhydrous TFA to initially dissolve the protein, followed by dilution
with DMF and Et₃N and the subsequent isolation of the protein from a
Sephadex LH-60 column (1 x 20 cm) that was equilibrated and developed with
DMF containing 10 mM Bu₄N⁺Cl⁻ and 10 mM Et₃N (section III.1.). The
trimesylation of the protein was achieved with 0.4 M tritiated DT-TM
imidazolide (specific activity of 0.7 mCi/ mmol) in the presence of 1 M Et₃N
and 0.2 M Bu₄N⁺Cl⁻ (as detailed in the legend of Figure 26). The extent of
trimesylation was determined by quantitating the radioactivity and
normalizing it to the amount of protein recovered from the column (Figure
26). The protein quantity was assessed by amino acid analysis using the
"PicoTag" method of Millipore-Waters, as described in Materials and
Figure 26: Time course of apobacteriorhodopsin trimesylation with DT-TM imidazolide in DMF. Apobacteriorhodopsin, that had been solubilized in anhydrous DMF in the presence of Bu₄N+Cl⁻, was reacted at a concentration of 1 mg/mL with 0.4 M tritiated DT-TM imidazolide, having a specific activity of 0.7 mCi/mmol, in the presence of 1 M Et₃N and 0.2 M Bu₄N+Cl⁻. An aliquot of the reaction was removed at the indicated times and diluted into DMF containing 1 M Me₃SiCH₂CH₂OH and 1 M Et₃N. After 2 hours, the trimethylsilylethyl groups were deprotected by the addition of 5 equivalents of Bu₄N+F⁻ for every Me₃SiCH₂CH₂ group and then the sample was applied to a Sephadex LH-60 column (1 x 28 cm) that was equilibrated and eluted in DMF (0.1% Et₃N). Fractions of 0.5 mL were collected and the radioactivity assessed by diluting an aliquot into Amersham aqueous counting scintillant and analyzing it on a Beckman LS-7800 scintillation counter. Radioactivity eluting near the void volume of the column was totalled as % theoretical yield normalized to the protein amount established by amino acid analysis.
Methods.

As observed in Figure 26, there was a fairly rapid initial acylation in which 50% of the sites were derivatized within the first 0.5 hour. This would correlate roughly to the complete trimesylation of all of the more reactive e-amino groups of Lys and the aromatic hydroxyl groups of Tyr residues; these residues comprise 36.7% of the total 49 possible sites of derivatization. The acylation rate then appeared to slow down, achieving 70% derivatization within 4 hours; this would likely indicate the trimesylation of the less reactive primary hydroxyl groups of the Ser residues (Lys, Tyr and Ser residues comprise 63.2%). Finally, there was a slow acylation rate corresponding to the secondary hydroxyl groups of Thr residues, resulting in 100% trimesylation within 24 hours. The overall decrease in the rate of trimesylation with time was also due, in part, to the loss in reagent reactivity with time, as was observed in Figure 17.

This assessment of 100% trimesylation of all possible sites on apoBR within 24 hours was based on the radioactivity incorporated onto the protein and normalized to the amount of protein recovered from the Sephadex LH-60 column; the actual protein recovery was only 40 - 50% based on amino acid analysis. The total length of time required for complete trimesylation of the protein was about 5 times longer than that required for complete acylation of the secondary hydroxyl groups on the model peptide DNP-prolylthreoninamide; this approximation incorporates the increase in the concentration of DT-TM imidazolide to 0.4 M for the protein reaction. The increase in time was anticipated to be due to steric hindrance within the protein and due to the loss of DT-TM imidazolide reactivity over time (Figure 17).

The time course of apobacteriorhodopsin trimesylation with DT-
TM imidazolide was also assessed in anhydrous Me₂SO (Figure 27). Even though the rate of trimesylation in Me₂SO was 25% of that in DMF (Figure 24) and the DT-TM imidazolide was not any more stable under the reaction conditions in Me₂SO (Figure 18), it was anticipated that the protein would be completely derivatized within a shorter period of time, due to the protein being more fully denatured in Me₂SO (Llinas et al. 1980).

The solubilization of apoBR in Me₂SO was achieved by initially dissolving it in TFA, followed by dilution with DMF containing Bu₄N⁺TFA⁻ and then neutralized with 1.5 equivalents of Et₃N, with respect to TFA. The protein was then isolated from a Sephadex LH-20 column that was equilibrated and eluted with Me₂SO containing 50 mM Bu₄N⁺Cl⁻ and 50 mM Et₃N. The apobacteriorhodopsin was derivatized according to the conditions detailed in Figure 27 and the extent of acylation of the membrane protein with tritiated DT-TM imidazolide was assessed with time by the amount of radioactivity incorporated onto the protein and normalized to the amount of protein recovered. It was found that the trimesyl-apoBR recovery was fairly low, approximately 20%, after its isolation from two Sephadex LH-60 columns, as outlined in the legend of Figure 27.

The overall pattern of the extent of apobacteriorhodopsin trimesylation with time (Figure 27) was similar to the pattern observed with DMF (Figure 26) which again illustrated the different types of functional groups undergoing acylation and their corresponding differences in reactivities. Underlining this change in the rate of acylation was the general loss in reagent reactivity over the same time period (Figure 18). However, the acylation of the apoBR was complete within 18 hours, which was approximately 6 times longer than the time required to completely derivatize DNP-prolylthreoninamide under the same conditions. This rate of protein
Figure 27: Time course of apobacteriorhodopsin trimesylation with DT-TM imidazolide in anhydrous Me$_2$SO. Apobacteriorhodopsin that had been solubilized in anhydrous Me$_2$SO in the presence of Bu$_4$N$^+$Cl$^-$, was reacted at a concentration of 2 mg/mL with 0.8 M tritiated DT-TM imidazolide (specific activity of 0.7 mCi/mm mol) at 45°C in the presence of 0.8 M Et$_3$N and 0.4 M Bu$_4$N$^+$Cl$^-$. At the indicated times, an aliquot of the reaction was removed, diluted with an equal volume of 0.8 M Et$_3$N in Me$_2$SO and applied to a Sephadex LH-60 column (1 x 28 cm) that was equilibrated and eluted with Me$_2$SO. The radioactivity associated near the void volume of the column was collected, concentrated and deprotected with 5 equivalents of Bu$_4$N$^+$F$^-$. After 15 minutes, the sample was then diluted with 2 volumes of Me$_2$SO and applied and eluted from another Sephadex LH-60 column (1 x 28 cm) with Me$_2$SO. Fractions at 0.5 minute intervals were collected and diluted into Amersham aqueous counting scintillant and the radioactivity assessed by a Beckman LS-7800 scintillation counter. The radioactivity eluting near the void volume was totalled as percent theoretical yield based on the 24 hour sample whose radioactivity was normalized to its protein amount, as determined by amino acid analysis.
derivatization in Me₂SO was better than that observed in DMF; when 0.4 M DT-TM imidazolide was used in DMF, complete apoBR trimesylation required 24 hours (Figure 26). As a comparison, 0.8 M DT-TM imidazolide in Me₂SO would be equivalent to 0.2 M in DMF based on the reactivity observed with the model peptides and as such, would require at least 48 hours if the protein acylation rates were identical. Therefore, the protein acylation must be faster in Me₂SO than in DMF, by a factor of approximately 2.7. This must solely be due to the different conformational states of the protein in the different solvents; apoBR is more fully denatured in Me₂SO and hence, its possible acylation sites more accessible to DT-TM imidazolide, while in DMF, the secondary stuctures of the protein must hinder the trimesylation reaction (Llinas et al., 1980).

III.4.A.vi. Trimesyl-Apobacteriorhopsin Characterization

Having achieved the complete trimesylation of apobacteriorhodopsin, the behaviour of this modified protein, TM-apoBR, was assessed on a Sephadex G-150 molecular sieving column that was equilibrated and developed with a simple aqueous buffer consisting of 50 mM NH₄HCO₃, at pH 8. The molecular weight exclusion limits for the Sephadex G-150 matrix are 5.0 x 10³ to 3.0 x 10⁵ for globular proteins and 1.0 x 10³ to 1.5 x 10⁵ for dextrans. The trimesylated apobacteriorhodopsin has a calculated molecular weight of 3.58 x 10⁴ and was expected to behave more like a dextran than a globular protein, due to charge repulsion of the trimesyl groups. Therefore, Sephadex G-150 should be capable of resolving any covalently crosslinked proteins, if they were present.

The effluent from the column was monitored at 280 nm, creating
the chromatograms observed in Figure 28. As described in the legend, panels A and B represent different solubilization procedures and reaction conditions that were utilized for apoBR trimesylation in DMF. For panel A, the apoBR was initially solubilized in TFA, diluted with DMF and Et3N and then applied and collected from a Sephadex LH-60 column that was developed with Bu4N+Cl− in DMF. The recovered apoBR was concentrated prior to its acylation with DT-TM imidazolide. The protein for panel B, was TFA solubilized, followed by the addition of Bu4N+TFA−, DMF and Et3N and then reacted with DT-TM imidazolide. Panel C contained no apoBR, but was processed in a similar manner as the protein sample for panel A, thereby providing a reaction blank for comparison.

The protein recoveries from the various columns that were utilized for the isolation of the TM-apoBR sample of panel A, as assessed by amino acid analysis, were as follows: DT-TM-apoBR yield from LH-60 (Me2SO) was 59%. The TM-apoBR yield from LH-60 (Me2SO) was 67%. Therefore, from the initial 200 μg of apobacteriorhodopsin, approximately 38 μg (19% of the original) was applied to the Sephadex G-150 column and the resulting recovery of TM-apoBR was 57%. These yields were based on the radioactivity recovered.

As observed in panel A, a broad absorbance peak was obtained that appeared to be slightly retained by the column and with which the radioactivity essentially coeluted. These results indicated the water soluble nature of the derivatized apobacteriorhodopsin and the covalent attachment of the desired trimesyl group to the protein. There was little radioactivity associated with the absorbance peak that eluted near the void volume of the column, which would indicate the absence of large solubilized protein aggregates being generated during the trimesylation reaction through some
Figure 28: Elution profile of trimesylated apobacteriorhodopsin from a Sephadex G-150 column. Apobacteriorhodopsin (220 μg) that had been solubilized in anhydrous DMF in the absence of TFA was trimesylated at a concentration of 2 mg/mL with 0.4 M DT-TM imidazole in the presence of 1 M Et₃N and 0.13 M Bu₄N+Cl-. After 18 hours, at 45°C, the reaction was diluted with an equal volume of 0.8 M Et₃N in Me₂SO and was applied to a Sephadex LH-60 column (1 x 28 cm) that was equilibrated and run with Me₂SO. The radioactivity observed near the void volume of the column was collected, concentrated and deprotected with 5 equivalents of Bu₄N+F- per Me₃SiCH₂CH₂ group. The sample was then applied to another Sephadex LH-60 column (1 x 28 cm) that was equilibrated and eluted with Me₂SO. The radioactivity associated near the void volume was collected, concentrated, diluted with 50 mM NH₄HCO₃, pH 8.5 and applied to a Sephadex G-150 column (1 x 120 cm) that was equilibrated and eluted with 50 mM NH₄HCO₃, pH 8.5 and 0.02% NaN₃ (panel A). Fractions of 2.7 mL were collected and diluted into Amersham aqueous counting scintillant and the radioactivity quantitated by a Beckman LS-7800 scintillation counter. Apobacteriorhodopsin that had been solubilized in anhydrous DMF in the presence of TFA, was trimesylated at a concentration of 0.31 mg/mL with 0.31 M DT-TM imidazole in the presence of 0.385 M Bu₄N+TFA-, 0.31 M Et₃NH+TFA-, 60 mM Et₃N and 0.62 M 4-pyrrolidinopyridine. After 16 hours at room temperature, the DT-TM imidazole was deactivated by the addition of Me₃SiCH₂CH₂OH (0.8 M) for 0.5 hour. The sample was then processed in a similar manner as in panel A, except that after deprotection the sample was diluted with 50 mM NH₄HCO₃, pH 8.5 and directly applied to the Sephadex G-150 column (panel B). When the same reaction process as that for panel A was performed in the absence of apobacteriorhodopsin and applied to the Sephadex G-150 column, the absorbance profile in panel C was obtained.
form of protein crosslinking. Furthermore, there was no radioactivity associated with the absorbance peak eluting near the total volume of the column; as anticipated, the trimesyl esters on the protein were not readily undergoing hydrolysis under these aqueous conditions. The presence of both the absorbance peaks eluting near the void volume and the total volume of the column in panel C, indicate that they were derived from the acylation reaction or the deprotection reaction but were not related to the apobacteriorhodopsin.

The molecular weight of TM-apoBR was assessed on the basis of its $K_{av}$, with respect to globular proteins (Figure 29). The $K_{av}$ of the eluted protein was determined by the following equation:

$$K_{av} = V_e - V_o$$

where $V_o$ is the void volume,

$$V_t - V_o$$

$V_t$ is the total or bed volume and

$V_e$ is the volume required to elute the protein

The TM-apoBR in panel A, eluted in a position that indicated its molecular weight was about 90,000, based on its $K_{av}$ of 0.18 (Figure 29). The fully derivatized apobacteriorhodopsin has a calculated molecular weight of only 35,800 and within experimental error, the TM-apoBR in panel A was assessed as being 100% derivatized, based on the radioactivity normalized to the amount of protein. This molecular weight though determined with respect to standard globular proteins, was also estimated to be too large when extrapolated to the exclusion limits for dextrans; the TM-apoBR molecular weight would be approximately 45,000. This nominal molecular weight was arrived at through the assumption that the difference in the size exclusion of dextrans and globular proteins is a factor of 0.5 ($1.5 \times 10^5$ as compared to $3.0 \times 10^5$).

The chromatogram of Figure 28, panel B illustrated the broad
Figure 29: Determination of apparent molecular weight of trimesylated apobacteriorhodopsin by gel permeation chromatography on Sephadex G-150. The column (1 x 120 cm) was equilibrated and eluted with 50 mM NH$_4$HCO$_3$, pH 8.5 and 0.02% sodium azide. The standard proteins used were: glucose-6-phosphate dehydrogenase, $M_r = 105,000$ (A); bovine serum albumin, $M_r = 68,000$ (B); chicken ovalbumin, $M_r = 45,000$ (C); carbonic anhydrase, $M_r = 29,000$ (D); whale myoglobin, $M_r = 17,000$ (E); and horse cytochrome C, $M_r = 12,500$ (F). The observed behaviour for trimesyl-apobacteriorhodopsin is indicated by squares for the reactions performed in DMF (closed symbol for panel A and open symbol for panel B of Figure 28), a circle for its acylation in Me$_2$SO (Figure 30) and a cross for the reaction performed in pyridine (Figure 39).
absorbance peak that was observed when apobacteriorhodopsin was derivatized with cold DT-TM imidazolide, in the presence of TFA, under the conditions specified in the legend. The absorbance peaks that eluted near the void volume and the total volume of the column were again derived from the acylation reaction or the deprotection reaction; the increase in their absorbance was likely due to the elimination of the Sephadex LH-60 column after the deprotection reaction of the DT-TM-apoBR, prior to the Sephadex G-150 column. The $K_{av}$ of the retained broad peak was found to be 0.42 which correlated to a molecular weight of 36,000 (Figure 29). The middle third of this absorbance peak was collected, concentrated and reapplied to the Sephadex G-150 column and the resulting chromatogram indicated the regeneration of a broad absorbance peak having a $K_{av}$ of 0.42 and there was no evidence of a void volume peak (data not shown). This indicated that the behaviour of the protein with respect to its elution position and the shape of the absorbance peak were inherent properties of the modified protein and were not dependent on how the protein was applied to the column or on the other components of the sample mixture that were coapplied.

Assessment of the different conditions utilized for this reaction indicated that in the presence of 1 M Et$_3$N, the same rate of trimesylation of the model peptides by DT-TM imidazolide was observed, even if the amphipathic salt, Bu$_4$N$^+$TFA$^-$ was present. The acylation rate could be enhanced by a factor of 4, if 2 equivalents of the acylation catalyst, 4-pyrrolidinopyridine (PPY) were also added when Bu$_4$N$^+$TFA$^-$ was present (data not shown). Both preparations of TM-apoBR were entirely derivatized and yet their behaviour on the Sephadex G-150 column were substantially different. The explanation of this may be due to the differences in the processing of the apobacteriorhodopsin prior to its modification; whether this
gave rise to the larger molecular weight species (possibly protein dimers) in the case of the reaction of Figure 28, panel A was not known and was not investigated further.

The molecular weight of the TM-apoBR generated in Figure 28; panel B, was very close to the calculated molecular weight of 35,800 after complete derivatization. This would appear to indicate that the monomeric TM-apoBR behaves as a globular protein. This behaviour was not anticipated due to the large number of negative charges that have been incorporated over the entire protein molecule, the result of which would be the inhibition of protein folding due to charge repulsion. The modified protein was expected to have become linearized upon derivatization and hence behave more like a dextran than a globular protein. The molecular weight exclusion limits of the Sephadex G-150 for dextrans is $1.0 \times 10^3$ to $1.5 \times 10^5$ and therefore, it was anticipated that the modified protein would elute slightly closer to the void volume of the column and have a higher apparent molecular weight when compared to globular proteins. The observation must indicate that other factors must be involved in the elution of the trimesylated protein from Sephadex columns.

Similar results were obtained for the trimesylation reactions that were performed in Me$_2$SO. The apobacteriorhodopsin that had undergone solubilization and trimesylation in Me$_2$SO was isolated, as described in the legend to Figure 30 and the yield of the modified protein from each of the columns were as follows: There was a 95% recovery of the DT-TM-apoBR from the Sephadex LH-60 that was eluted with Me$_2$SO. After deprotection and isolation of the TM-apoBR on another Sephadex LH-60 column in Me$_2$SO, a 63% yield of the applied radioactivity was found. Therefore, from the original 130 μg of apobacteriorhodopsin, only about 27 μg (21% of the
initial amount of protein) was applied to the Sephadex G-150 column. The TM-apoBR yield from the G-150 column was 67%. In all cases, the yields were based on the radioactivity that was applied to the columns.

The chromatogram in Figure 30, depicts the broad absorbance peak at 280 nm that was obtained for the TM-apoBR applied to the Sephadex G-150 column. On the basis of the $K_a$ of 0.41 of the retained peak, the apparent molecular weight of the modified protein was determined to be about 38,000 (Figure 29), which was very close to the calculated molecular weight of TM-apoBR, 35,800 and to the apparent molecular weight of the TM-apoBR when derivatized in DMF and in the presence of TFA (36,000; Figure 28, panel B). The radioactivity was found to essentially coelute with the absorbance peak indicating that there was a covalent attachment of the trimesyl group onto the protein and that all of the desired sites of modification on the protein were completely trimesylated. The bulk of the radioactivity was associated with the protein except for a slight amount near the total volume of the column, the origin of which was not known but may be due to free trimesic acid that was not separated from the TM-apoBR during its isolation. The absence of both an absorbance peak and a peak of radioactivity eluting near the void volume of the column signified that protein aggregation had not occurred during the trimesylation reaction which would have resulted in the formation of a partially solubilized protein mass having a relatively large molecular weight.

On the basis of the relatively low recovery (50 to 60%) of the trimesylated protein from all of the different types of Sephadex columns that were utilized in its isolation, there appears to be some type of adsorption or affinity towards the glucose matrix. Any form of affinity for the gel would be expected to cause some general retention and hence peak broadening, which
Figure 30: Elution profile of trimesylated apobacteriorhodopsin from a Sephadex G-150 column. Apobacteriorhodopsin (130 μg) that had been solubilized in anhydrous Me₂SO in the presence of Bu₄N⁺Cl⁻, was reacted at a concentration of 1.3 mg/mL with 0.8 M tritiated DT-TM imidazolidine (specific activity of 0.7 mCi/mmol), 0.8 M Et₃N and 0.4 M Bu₄N⁺Cl⁻ for 18 hours at 45°C. The reaction was then diluted with an equal volume of 0.8 M Et₃N in Me₂SO and was applied and eluted from a Sephadex LH-60 column (1 x 28 cm) with Me₂SO. The radioactivity associated near the void volume of the column was collected, concentrated and deprotected with 5 equivalents of Bu₄N⁺F⁻. After 0.5 hour, this reaction was applied and eluted from another Sephadex LH-60 column (1 x 28 cm) with Me₂SO. The radioactivity near the void volume was collected, concentrated and diluted with 2 volumes of 50 mM NH₄HCO₃, pH 8.5. This was then applied to a Sephadex G-150 column (1 x 120 cm) that was equilibrated and developed with 50 mM NH₄HCO₃, pH 8.5. Fractions of 2.4 mL were collected and diluted into Amersham aqueous counting scintillant and the amount of radioactivity quantitated by a Beckman LS-7800 scintillation counter.
was exactly what was observed. Therefore, the elution profile of the TM-apoBR from the Sephadex G-150 column must be the result of several factors affecting the elution of the modified protein and not just one of molecular sieving and protein shape.

In order to accurately assess the adsorption phenomenon of the trimesylated apobacteriorhodopsin to Sephadex, apobacteriorhodopsin that had been derivatized with tritiated DT-TM imidazolide and isolated from the Sephadex G-150 column (1 x 120 cm), was divided into small portions and applied to small Sephadex G-150 columns that were equilibrated and eluted with the buffers specified in Table 11. The purpose was to assess the effect of varying the salt concentration or the organic modifiers in the elution solvent on the adsorption of the acylated protein to Sephadex G-150. As indicated in Table 11, increasing the Na$_2$HPO$_4$ concentration in the elution buffer, appears to result in the TM-apoBR adsorbing more strongly to the Sephadex; at 0 M, there was 96% recovery of the acylated protein as compared to 0% recovery at 1 M. The presence of 8 M urea was also found to be effective in eluting the TM-apoBR (94%), even in the presence of 50 mM Na$_2$HPO$_4$, pH 7.5, which in the absence of urea was found to yield only 52% of the protein. The addition of organic modifiers, such as 10% ethanol or 10% isopropanol were also found to reduce the recovery of the TM-apoBR by 30 and 24%, respectively.

The adsorption behaviour of the trimesylated protein appeared to be due to the hydrophobic effect as opposed to an ionic effect; as the salt concentration increased the adsorption of the TM-apoBR to the Sephadex also increased. This would account for the low protein recoveries that were observed from all the different types of Sephadex columns utilized for the modified protein's isolation. This type of adsorption would cause general retention of the trimesylated protein on the column and therefore would not
<table>
<thead>
<tr>
<th>Eluent</th>
<th>Organic Modifier</th>
<th>TM-apoBR Recovery (% theoretical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄, pH7.5 (mM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>10% ethanol</td>
<td>70</td>
</tr>
<tr>
<td>0</td>
<td>10% propanol</td>
<td>76</td>
</tr>
<tr>
<td>50</td>
<td>8 M urea</td>
<td>94</td>
</tr>
</tbody>
</table>

Table 11: Adsorption of trimesyl-apoBR to Sephadex G-150. Tritiated trimesyl-apobacteriorhodopsin (0.7 mCi/mmol trimesyl group) that had been isolated from a Sephadex G-150 column was reapplied to small Sephadex G-150 columns (0.75 x 3 cm) that were separately equilibrated and eluted with one of the buffers specified in the table. Fractions of 1.0 mL were collected, diluted into Amersham aqueous counting scintillant and the radioactivity quantitated with a Beckman LS-7800 scintillation counter. The eluted radioactivity was totalled and expressed as a percentage of the amount applied.
give a true assessment of the molecular weight whether based on the $K_{av}$ of globular proteins or linear dextrans. Furthermore, the trimesylated protein would likely adsorb more strongly if the hydrophobic counterion, $Bu_4N^+$, was associated with the trimesyl-apoBR when the modified protein was applied to the Sephadex G-150 column; this would necessitate the removal of this counterion prior to elution through an aqueous column.

Therefore, the elution profiles of TM-apoBR that were observed in Figure 28 and in Figure 30, illustrate that the mobility of trimesylated proteins through the Sephadex matrix was the result of several factors exerting their individual effects. These factors include the molecular sieving properties of the Sephadex, the affinity of the modified protein towards the matrix presumably through a hydrophobic effect, the actual number of sites on the protein that were trimesylated and the resulting extent of linearization of the protein.

Polyacrylamide gel electrophoresis was considered as a possible method to assess the molecular weight of the trimesylated apobacteriorhodopsin and to determine the extent of any protein crosslinking that may have occurred during the acylation reaction. The standard SDS-PAGE system of Laemmli (1970) and the 8 M urea SDS gel of Swank and Munkres (1971) were both tried at pH 6.8, however, the mobility of the TM-apoBR was such that it migrated faster than the bromophenol blue dye that was used to monitor the electrical front (data not shown).

These PAGE systems were then modified such that they were buffered at pH 4.0, as outlined in Materials and Methods. However, the TM-apoBR was still found to migrate faster than the dye front (data not shown). This increased mobility was believed to be primarily due to the enhanced negative charge density of the modified protein as a result of the 98 additional
carboxylate groups that were incorporated with the attached trimesyl groups. In order to further reduce the mobility of the TM-apoBR, the PAGE system was run at pH 2.0; under these acidic conditions, approximately 99% of one of the trimesyl carboxylate groups and 90% of the other would be protonated (pKa1 = 3.1 and pKa2 = 3.9; Morrison and Boyd, 1966), thereby reducing the modified protein's negative charge density and hence, its mobility. At low pH, the dodecylsulphate must be used as the lithium salt to maintain its solubility (Lichtner and Wolf, 1979), hydrogen peroxide, ascorbic acid and FeSO4 are used as the catalyst system for acrylamide polymerization (Jones et al., 1981) and the electrical front must be marked with an appropriate dye, one that is negatively charged at pH 2, such as Safranin C.

The tritiated TM-apoBR samples that were generated in DMF and Me2SO and then isolated from the Sephadex G-150 column (Figure 28; panel A and Figure 30, respectively) were run on an acidic (pH 2) LDS-PAGE system containing 8 M urea (Table 12). Urea was included in the acrylamide gel in order to reduce the pore size of the gel (Swank and Munkres, 1971) and thus, further reduce the mobility of the modified protein. The detection of the proteins in this gel was accomplished by fluorography; the acrylamide gel was soaked in Me2SO in the presence of the fluor, PPO, and then placed next to an X-ray film at -70°C, for two weeks (Bonner and Laskey, 1974). The mobility of all of the proteins were so severely restricted that they were only just able to migrate into the stacking gel after 20 hours of electrophoresis at 180 mV; this was indicated by the banding pattern of the standard molecular weight protein markers (data not shown). The poor resolution and separation of the standard molecular weight proteins made the estimation of the relative molecular weight of the trimesylated apobacteriorhodopsin inaccurate. However, the protein banding pattern indicated that the DMF and the Me2SO acylation reactions had generated a
<table>
<thead>
<tr>
<th>Reaction Solvent</th>
<th>Figure</th>
<th>8 M Urea in Gel</th>
<th>Protein Detection</th>
<th>Apparent Mol Wt (x10^-3)</th>
<th>PAGE (R_f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMF</td>
<td>28; A</td>
<td>+ (2 bands)</td>
<td>fluor</td>
<td>90</td>
<td>n.d.</td>
</tr>
<tr>
<td>CI_2SO</td>
<td>30</td>
<td>+ (1 band)</td>
<td>fluor</td>
<td>38</td>
<td>n.d.</td>
</tr>
<tr>
<td>DMF</td>
<td>28; B</td>
<td>-</td>
<td>silver</td>
<td>150</td>
<td>*</td>
</tr>
</tbody>
</table>

n.d. = not determined
* = no protein band observed

Table 12: Summary of the electrophoretic results for trimesylated apobacteriorhodopsin on acidic LDS polyacrylamide gels. Aliquots of the pooled fractions that were collected for the trimesylated apobacteriorhodopsin isolated from the Sephadex G-150 column, were processed, applied and run on an acidic (pH 2) lithium dodecylsulphate polyacrylamide gel in the presence or absence of 8 M urea, as indicated in the table, according to the modified procedure of Swank and Munkres (Materials and Methods). After electrophoresis, the proteins were detected by processing the gels appropriately for fluorography (fluor) or for silver staining (silver).
protein that had the same mobility under these conditions. In the case of the acylation reaction performed in Me₂SO, this was the only protein band observed and it was a clear distinct band, indicative of the monomeric TM-apoBR (Table 12). An additional protein product was observed in the DMF reaction, a slower migrating band having an apparently larger molecular weight and whose intensity appeared to be three times that of the TM-apoBR band that comigrated with the Me₂SO protein product. The higher molecular weight protein product can only originate from apobacteriorhodopsin and as indicated by the associated radioactivity, had trimesyl groups covalently attached to it; the presence of this product lends support to the concept that under the acylation conditions in DMF, the apobacteriorhodopsin may not be fully denatured and as such may associate as an aggregate which could result in intermolecular crosslinking.

The same acidic (pH 2) LDS polyacrylamide gel system was again employed but now in the absence of 8 M urea and it generated a much superior standard molecular weight protein banding pattern due to the increased mobility through the larger pores of the gel. These proteins were detected by the silver stain technique due to its reported high sensitivity and the requirement of carboxylate groups for the staining of the proteins (Merril et al., 1984); the presence of trimesyl groups on apobacteriorhodopsin should enhance its detection by this method. (Staining of TM-apoBR was found to be very poor with the standard Coomassie Blue method (Fairbanks et al., 1971); the dye apparently becomes associated with proteins through the ε-amino groups on Lys residues, which in the case of TM-apoBR are covalently modified.) On this type of acidic gel, the absorbance peaks isolated from the Sephadex G-150 column for the trimesylation reaction of apobacteriorhodopsin performed in DMF, in the presence of TFA (Figure 28;
panel B), were analyzed (Table 12). The peak eluting near the void volume of the column was found to have no detectable protein bands associated with it and therefore, was believed to be reagent derived from the deprotection reaction. The major peak correlating to a 36,000 dalton protein appears to have a relative molecular weight of only 11,000. Furthermore, the slight forward shoulder of the major absorbance peak, eluted from the Sephadex column having a $K_{av}$ that correlated to a relative molecular weight of 65,000 and yet it was found to migrate on the polyacrylamide gel as a fairly distinct band, corresponding to a 16,000 dalton protein.

These discrepancies in the molecular weight of the modified apobacteriorhodopsin can be resolved by the assessment of the charge densities of the proteins involved, at the pH of electrophoresis, as indicated in Table 13. In this table, the detergent, dodecylsulfate (DS$^-$), was assumed to associate with proteins, both native and trimesylated, in a ratio of 1.4 : 1 by weight which correlates to one molecule of detergent for every 190 molecular weight of protein. It was not expected that the trimesyl groups whether protonated or ionized would alter the packing of the DS$^-$, though it was anticipated that the trimesyl carboxylate groups would tend to be oriented away from the protein backbone. Therefore, upon trimesylation of apobacteriorhodopsin, the amount of bound DS$^-$ would increase from 137 to 188 due to the additional 9,800 molecular weight.

The Stoke's radius of the trimesylated apobacteriorhodopsin would be independent of pH when in the presence of SDS or LDS; both modified and unmodified proteins would be linear rods in shape, when in the presence of this detergent. Therefore, the mobility of proteins during electrophoresis would primarily be due to its negative charge density at the prescribed pH of the PAGE system. Furthermore, it was thought that the
<table>
<thead>
<tr>
<th>Protein Sample</th>
<th>Total Asp</th>
<th>Total Glu</th>
<th>Total Lys</th>
<th>Total Arg</th>
<th>Total His NH₃⁺</th>
<th>Total COO⁻</th>
<th>Assoc Grps</th>
<th>Net Neg Charge at pH</th>
<th>Net Neg Charge 10[3] Mol Wt; pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>apoBR</td>
<td>2.6</td>
<td>0</td>
<td>9</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>137.2</td>
<td>142.3</td>
</tr>
<tr>
<td>ANT</td>
<td>2.9</td>
<td>0</td>
<td>15</td>
<td>6</td>
<td>22</td>
<td>17.3</td>
<td>0</td>
<td>153.3</td>
<td>133.2</td>
</tr>
<tr>
<td>BSA</td>
<td>6.6</td>
<td>0</td>
<td>54</td>
<td>78</td>
<td>62</td>
<td>23.17</td>
<td>1</td>
<td>347.2</td>
<td>379.24</td>
</tr>
<tr>
<td>Oval</td>
<td>4.5</td>
<td>0</td>
<td>(33)</td>
<td>20</td>
<td>15</td>
<td>7.1</td>
<td>1</td>
<td>237.2</td>
<td>229.2</td>
</tr>
<tr>
<td>C.Anh</td>
<td>3.0</td>
<td>0</td>
<td>14</td>
<td>13</td>
<td>18</td>
<td>7.11</td>
<td>1</td>
<td>158.2</td>
<td>150.12</td>
</tr>
<tr>
<td>Chymo</td>
<td>2.4</td>
<td>0</td>
<td>9</td>
<td>5</td>
<td>14</td>
<td>4.2</td>
<td>1</td>
<td>126.2</td>
<td>120.11</td>
</tr>
<tr>
<td>Mb</td>
<td>1.7</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>19</td>
<td>4.12</td>
<td>1</td>
<td>89.7</td>
<td>75.6</td>
</tr>
<tr>
<td>Lyso</td>
<td>1.43</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>75.6</td>
<td>67.6</td>
</tr>
<tr>
<td>Cyto C</td>
<td>1.25</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>19</td>
<td>2.3</td>
<td>1</td>
<td>66.5</td>
<td>54.4</td>
</tr>
<tr>
<td>TM-apoBR</td>
<td>3.58</td>
<td>49</td>
<td>9</td>
<td>10</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>188.2</td>
<td>299.260</td>
</tr>
<tr>
<td>TM-ANT</td>
<td>4.12</td>
<td>61</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>17.3</td>
<td>0</td>
<td>217.339</td>
<td>293.204</td>
</tr>
<tr>
<td>TM-BSA</td>
<td>9.46</td>
<td>143</td>
<td>54</td>
<td>78</td>
<td>0</td>
<td>23.17</td>
<td>0</td>
<td>498.879</td>
<td>725.476</td>
</tr>
<tr>
<td>TM-Oval</td>
<td>6.18</td>
<td>84</td>
<td>(33)</td>
<td>0</td>
<td>15</td>
<td>7</td>
<td>0</td>
<td>325.506</td>
<td>438.313</td>
</tr>
<tr>
<td>TM-C.Anh</td>
<td>4.42</td>
<td>71</td>
<td>14</td>
<td>13</td>
<td>0</td>
<td>7.11</td>
<td>0</td>
<td>233.386</td>
<td>329.224</td>
</tr>
<tr>
<td>TM-Chymo</td>
<td>3.80</td>
<td>70</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>4.2</td>
<td>0</td>
<td>200.349</td>
<td>300.202</td>
</tr>
<tr>
<td>TM-Mb</td>
<td>2.38</td>
<td>34</td>
<td>6</td>
<td>14</td>
<td>0</td>
<td>4.12</td>
<td>0</td>
<td>125.199</td>
<td>168.113</td>
</tr>
<tr>
<td>TM-Lyso</td>
<td>1.97</td>
<td>27</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>104.157</td>
<td>136.96</td>
</tr>
<tr>
<td>TM-Cyto C</td>
<td>1.93</td>
<td>34</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>102.178</td>
<td>152.101</td>
</tr>
</tbody>
</table>

**Average Net Negative Charge per 10³ Molecular Weight in the presence of DS⁻ at pH**

<table>
<thead>
<tr>
<th></th>
<th>6.5</th>
<th>4</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane Proteins</td>
<td>5.0</td>
<td>4.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Water Soluble Proteins</td>
<td>4.9</td>
<td>4.4</td>
<td>3.8</td>
</tr>
<tr>
<td>TM-Membrane Proteins</td>
<td>8.3</td>
<td>7.2</td>
<td>5.1</td>
</tr>
<tr>
<td>TM-Water Soluble Proteins</td>
<td>8.7</td>
<td>7.6</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Table 13: Charge Density of Proteins During S(L)DS-PAGE. In this table, the following abbreviations were used to designate proteins: apoBR, apobacteriorhodopsin; ANT, adenine nucleotide translocater; BSA, bovine serum albumin; Oval, ovalbumin; C.Anh, carbonic anhydrase; Chymo, chymotrypsinogen; Mb, sperm whale myoglobin; Lyso, egg white lysozyme; Cyto C, cytochrome C.
trimesylated protein would tend to be in a rod-like shape in the absence of
detergents when the pH was above the pKₐ of the carboxylate groups of the
attached trimesyl group.

The net negative charge per 10³ molecular weight would directly
correlate with the relative mobility of a protein during SDS (or LDS) PAGE;
the greater the charge density of the protein, the faster the rate of migration of
the protein through the polyacrylamide gel. As observed from the averages
of the charge densities (Table 13), the unmodified membrane proteins (i.e.
apoBR and ANT) have a charge density that are comparable to that of the
water-soluble proteins, under all of the pHs assessed. This illustrated the
ability of SDS-PAGE to assess the molecular weight of all types of protein
relative to those proteins used as standard molecular weight markers.

Upon trimesylation of the membrane proteins, their average
charge density increased from 5.0 to 8.3 net negative charges per 10³ molecular
weight, at pH 6.8. This would correlate to an increase in their rate of
migration by approximately 1.7 times, with respect to an unmodified protein
of similar molecular weight. Decreasing the pH of the system to 4.0, results in
a small reduction in the average charge density of the trimesylated membrane
proteins and only a minor alteration in its rate of migration; approximately
1.5 times the migration rate of an unmodified protein of similar molecular
weight. This large negative charge density at both of these pHs reflects the
state of ionization of the trimesyl carboxylate groups on the modified protein
and that the ε-amino groups of the Lys residues, now trimesylated, do not
carry a positive charge. By decreasing the pH to 2.0, the carboxyl groups on the
attached trimesyl groups would be almost entirely protonated thereby
reducing the average negative charge density to 5.1 charges per 10³ molecular
weight. Therefore, at pH 2.0, the TM-apoBR would still have a slightly larger
charge density, by a factor of approximately 1.2 than that calculated for standard proteins at this pH and hence the modified protein would migrate faster and thus behave as if it had a lower molecular weight, as was observed.

The front shoulder of the apparent 36,000 dalton protein (Figure 28, panel B), which elutes with a $K_{av}$ of a 65,000 dalton protein, was observed to electrophorese as a 16,000 dalton protein. Slight differences in the extent of trimesylation would not have a significant effect on the mobility of the TM-apoBR at pH 2.0, therefore, it may represent a small portion of the protein that has undergone crosslinking.

In order to get a more accurate assessment of the true molecular weight of the TM-apoBR, it would be necessary to derivatize the standard water soluble proteins to create trimesylated molecular weight markers. As indicated in Table 13, trimesylated water soluble proteins would have comparable negative charge densities at all of the pHs, to those of the modified membrane proteins. Lowering the pH of the LDS-PAGE system to 1 would not significantly decrease the negative charge density of the TM-apoBR and therefore would have little effect on the mobility of the trimesylated protein; the trimesyl carboxyl groups are almost completely protonated at pH 2. The reason that the mobility of the TM-apoBR will not approximate that of standard water soluble proteins is due to the elimination of the positive charge of the e-amino groups on the Lys residues in derivatized proteins.

The major observation from this electrophoretic work was the generation of a reproducible protein product upon trimesylation, one that eluted from Sephadex G-150 with a $K_{av}$ corresponding to a 36,000 dalton protein and that electrophoresed on an acidic PAGE system as a discrete band of apparently lower relative molecular weight (11,000) due to its negative charge density.
III.4.B.vii. Conclusion

In conclusion, the following could be stated from the work performed in DMF: 1. DT-TM imidazolide was a good acylating reagent but was not stable under the reaction conditions (1 M Et$_3$N in DMF), for undetermined reasons. 2. ApoBR was solubilized in the presence of Bu$_4$N+Cl$^-$ in DMF but it may not be sufficiently denatured to render all acylation sites accessible to DT-TM imidazolide. Proteins are known to maintain their secondary structure in DMF, (Llinas et al., 1980) which may lead to protein association and possible aggregation. 3. Even though the extent of apoBR trimesylation was assessed as being 100%, it could be conceivable that one or two sites may not be derivatized or have participated in protein crosslinking. This would account for the observed behaviour of the TM-apoBR on the Sephadex (ie. high apparent molecular weight of 55,000; Figure 28, panel A) and for the observation of 2 protein bands on acidic LDS-PAGE (Table 12). Altering the solubilization procedure to include TFA, appears to ensure complete trimesylation, as assessed by the elution behaviour from Sephadex G-150 (apparent molecular weight of 36,000; Figure 28, panel B) and the presence of one discrete band on acidic LDS-PAGE (Table 12).

Me$_2$SO was tried as an alternate solvent system due to its ability to completely denature proteins (Llinas et al., 1980). DT-TM imidazolide was found not to be stable under the reaction conditions (0.8 M Et$_3$N in Me$_2$SO), again for undetermined reasons and that its reactivity was only 25% of that found in DMF for the same concentration of DT-TM imidazolide and temperature. Apobacteriorhodopsin could be solubilized in Me$_2$SO but only by initially using the technique employed to solubilize the protein in DMF.
and then removing the DMF and the TFA by passing the protein through a Sephadex LH-20 column equilibrated and eluted with Me₂SO containing Bu₄N⁺Cl⁻ and Et₃N. By this procedure the apoBR appeared to be solubilized as judged by PAGE (Laemmli) analysis of the 280 nm absorbance peak eluting near the void volume of the column. This solubilized protein was found to be efficiently trimesylated within 18 hours in the presence of 0.8 M DT-TM imidazolide, 0.8 M Et₃N and 0.4 M Bu₄N⁺Cl⁻ at 45°C. The trimesylated protein behaved as if it were a 38,000 dalton globular protein on a Sephadex G-150 column and it electrophoresed as a single distinct band on a LDS-PAGE system at pH 2, containing 8 M urea (Table 12).

Even with the limitations of DT-TM imidazolide instability and poor reactivity in Me₂SO, the objective of enhancing the water solubility of apoBR through extensive chemical modification was realized. However, there was concern with some aspects of the procedure, such as (a) the laborious protein solubilization process, (b) the large DT-TM imidazolide concentration required for an adequate rate of trimesylation, (c) the high Et₃N concentration, (d) the relatively high temperature (45°C) and (e) the long reaction time. These factors when combined may enhance undesired side reactions (i.e. oxidation reactions, which are known to occur in Me₂SO (Pfitzner and Moffat, 1965). Therefore, having established that the trimesylated apobacteriorhodopsin was water soluble, an improved reagent and a simpler procedure was investigated.

III.4.B. Diprotected Trimesic Anhydride (DT-TM Anhydride)

The selection of the imidazolide derivative as the method of generating a reactive acylating reagent was based on its reported high reactivity and excellent product yield (Staab, 1962). After extensive
investigation, the acylation of the apoBR with the diprotected trimesic acid imidazolide was achieved, generating a water soluble derivative of apoBR. However, there were two main concerns with the imidazolide reagent: 1. The instability of the reagent in both DMF and Me₂SO. 2. The extreme reaction conditions (high reagent and base concentrations and elevated temperature) that were necessary to obtain an adequate rate of derivatization. In an attempt to overcome these problems, pyridine was assessed as an alternate solvent for derivatization with the diprotected trimesic acid imidazolide reagent.

Pyridine is an excellent solvent for acylation reactions (Guibe-Jampel et al., 1979) and in contrast to DMF and Me₂SO, it was found that DT-TM imidazolide was stable in pyridine for over 12 hours, even in the presence of 1 M Et₃N (data not shown). However, the major drawback of this solvent was the poor reactivity of the DT-TM imidazolide; it was significantly lower than that found in DMF or Me₂SO (approximately 10% of the initial rate constant found in DMF). Furthermore, no improvement in the reactivity was achieved even in the presence of equimolar amounts of the acylation catalysts, 4-dimethylaminopyridine (DMAP) and 4-pyrrolidinopyridine (PPY) (Hofle et al., 1978).

These results necessitated the investigation of a better acylation reagent. Preliminary experiments indicated that the reactivity of the diprotected trimesic anhydride was five times better than that of the DT-TM imidazolide. These findings were consistent with the observations noted for the activation of fatty acids (Mangroo and Gerber, 1988). Anhydrides readily react with amines and aromatic hydroxyl groups but have a much slower rate of reaction for aliphatic hydroxyl groups. The rate of reaction is substantially improved when pyridine is utilized as the solvent because it itself acts as a catalyst. The rate of acylation can be further improved by a factor of 10⁴ over
that of pyridine when the reaction is performed in the presence of the acylation catalysts DMAP (Steglich and Hofle, 1969 and Guibe-Jampel et al., 1979) or PPY (Neises and Steglich, 1978 and Hofle et al., 1978). The resulting catalyzed anhydride acylation reagent would be more reactive than the corresponding acid chloride (Mangroo and Gerber, 1988). In the presence of such catalysts, DT-TM anhydride should rapidly and efficiently derivatize amine and hydroxyl groups on the side chains of amino acid residues in membrane proteins. The DT-TM anhydride was, therefore, systematically investigated as a possible protein derivatizing reagent.

III.4.B.i. DT-TM Anhydride Formation

The synthesis of the DT-TM anhydride can be achieved by the reaction of 2 equivalents of the DT-TM acid with 1 equivalent of dicyclohexylcarbodiimide (DCC) and the concomitant formation of N,N'-dicyclohexylurea (DCU) which precipitates out of the CCl₄ solution. However, generation of the undesired side product, dicyclohexyl-N-acyl urea (DCAU), can occur depending on the acid and the solvent that are used (Khorana, 1953). The best solvent for the synthesis of long chain fatty acid anhydrides, with a minimum of DCAU formation, was found to be CCl₄ (over that for benzene, CHCl₃ or pyridine) and the reaction was complete within 0.5 hours (Selinger and Lapidot, 1966).

Formation of the diprotected trimesic anhydride by dicyclohexylcarbodiimide and the subsequent adduct formation with the acylation catalyst are schematically represented in Figure 31. The diprotected trimesic acid (I) is converted to the reactive symmetric DT-TM anhydride (II) by DCC and the resulting dicyclohexylurea (DCU) precipitates out of solution. Filtration of the reaction mixture through glass wool efficiently removes the
Figure 31: Diprotected trimesic anhydride: formation, catalyst activation and protein acylation. In this schematic representation of the reaction, DCC denotes dicyclohexylcarbodiimide, while PPY and DMAP represent 4-pyrrolidinopyridine and 4-dimethylaminopyridine, respectively.
DCU, thereby, affording the anhydride. The acylation catalyst (PPY or DMAP) forms a highly reactive adduct with the DT-TM anhydride (III), which is susceptible to nucleophilic attack by amine or hydroxyl groups, resulting in the trimesylation of these sites.

It was therefore anticipated that under the appropriate reaction conditions, the DT-TM anhydride in the presence of DMAP or PPY, would prove to be a very reactive reagent that would rapidly trimesylate the amine and hydroxyl groups on the side chains of the amino acid residues of a protein. Furthermore, this reactivity would be achieved without the need for harsh reaction conditions, such as high base concentrations or elevated temperatures, as were required with DT-TM imidazolide, which could result in undesired side reactions such as dehydration and oxidation.

During the formation of the DT-TM anhydride, there is the possibility of generating the undesired side product, dicyclohexyl-N-acylurea (DCAU) which would effectively reduce the amount of anhydride. This can occur through the intramolecular rearrangement of the dicyclohexyl-O-(DT-TM)urea prior to its attack by a second molecule of DT-TM acid (Khorana, 1953). The extent of this side reaction would become significant if the DT-TM acid concentration became limiting. It was, therefore, investigated in the attempt to optimize the formation of the symmetric anhydride while minimizing the generation of the DCAU side product.

The trimesyl group absorbs light at 280 nm which enables all of its derivatives to be detected at this wavelength. The trimethylsilyl ethyl ester derivatives of trimesic acid can be resolved and quantitated by reversed phase HPLC (see Figure 22). During the formation of the DT-TM anhydride with DCC, an aliquot was removed and diluted into pyridine containing 1 M Me₃SiCH₂CH₂OH for one hour. This should result in the generation of one
molecule of DT-TM acid and one molecule of TT-TM acid which can be resolved by HPLC. However, the HPLC chromatogram indicated the presence of a third peak that had a distinctive retention time, eluting between DT-TM acid and TT-TM acid when a gradient of methanol (50 to 100% over 15 minutes) was used to develop the reversed phase C18 column (data not shown). This peak did not change even when longer reaction times with the trimethylsilylethanol were carried out indicating that the peak was not the DT-TM anhydride. This peak was presumed to be the dicyclohexyl-N-(DT-TM)-urea and its presence indicated that the formation of the desired anhydride was not complete. By quantitating the extent of this side product during different activation procedures, one could, therefore, determine the conditions that are required for optimal DT-TM anhydride formation by DCC.

DCC was initially purified by being dissolved in dry benzene, pelleting any insoluble material by centrifugation, drying an aliquot of the supernatant under reduced pressure in a tared tube, and finally adding an appropriate volume of CCl4 to make a 0.4 M solution. The activity of this DCC solution was assessed as 98%, based on the addition of excess DT-TM acid. The amount of dicyclohexyl-N-(DT-TM)-urea formed was monitored by HPLC at 280 nm and quantitated from the area integrated under the absorbance peak having the appropriate retention time for the presumed dicyclohexyl-N-(DT-TM)-urea divided by the total area associated with the peaks corresponding to TM acid and its derivatives. The calculated dicyclohexyl-N-(DT-TM)-urea (DCAU) was expressed in Table 14, as a percentage of the reactive DCC by employing the following equation:

\[
\text{DCAU} (\%) = \left( \frac{\text{area of DCAU}}{\text{total area}} \right) \times 2.0 \text{ equiv. DT-TM per DCC} \times 100 \\
\text{(# equiv. DT-TM acid added)}
\]
Table 14: Generation of dicyclohexyl-N-(DT-TM)-urea (DCAU) during DT-TM anhydride formation by DCC. Diprotected trimesic acid was activated by the addition of DCC under the specified conditions, to achieve a final concentration of 0.2 M in CCl₄. After 1 hour at room temperature, an aliquot was removed and diluted into 1 M Me₃SiCH₂CH₂OH in pyridine. An aliquot of this pyridine solution was applied to a µBondapak C₁₈ reversed phase column previously equilibrated with 50% methanol and 50% 10 mM aqueous sodium acetate, pH 4.4 and developed with a gradient of methanol which reached 100% in 15 minutes. The amount of dicyclohexylacetylurea formed in the reaction was quantitated by integration of the DCAU (dicyclohexyl-N-(DT-TM)-urea) peak normalized to 2.0 equivalents of diprotected trimesic acid per DCC.
The results in Table 14 indicate that the formation of the DT-TM anhydride occurs poorly when the DCC solution was added to solid DT-TM acid; under these conditions, the requisite 2 equivalents of DT-TM acid are not entirely in solution, resulting in the formation of dicyclohexyl-N-(DT-TM)-urea (DCAU) from 45% of the DCC. Even when the DCC solution was added to a solution containing the 2 equivalents of the DT-TM acid, only 85% of the desired anhydride was formed. The best synthesis of the DT-TM anhydride was achieved when at least 2.4 equivalents of DT-TM acid was reacted with 1 equivalent of the DCC; under these conditions 96% of the DCC was utilized in the synthesis of the desired anhydride. This procedure was used routinely to ensure that no reactive DCC remained and that the formation of the DT-TM anhydride was optimal.

III.4.B.ii. DT-TM Anhydride Reactivity and Stability

In order to identify the best solvent system with respect to the reactivity of the DT-TM anhydride, DNP-prolylthreoninamide was derivatized in the presence of an acylation catalyst and the initial rate constant of trimesylation was assessed. The details of the components of each reaction mixture are given in Table 15. On the basis of the calculated initial rate constants, the order of the solvents with respect to DT-TM anhydride reactivity was as follows: The utilization of CHCl₃ or DMF as the solvent resulted in comparable initial rate constants. In pyridine, the rate of reaction was found to be 0.3 M⁻¹s⁻¹ which was 10 times slower than when the same reaction was performed in CHCl₃ (3 M⁻¹s⁻¹). In Me₂SO, the initial rate constant of the anhydride was at least a factor of 200 times slower than when pyridine was utilized. Furthermore, a difference in the rate constants was consistently observed between the two acylation catalysts in each of the
<table>
<thead>
<tr>
<th>Solvent</th>
<th>Catalyst</th>
<th>Molar Ratio Of Catalyst to Anhydride</th>
<th>Initial Rate Constant (M⁻¹s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>chloroform</td>
<td>DMAP</td>
<td>1.0</td>
<td>3</td>
</tr>
<tr>
<td>pyridine</td>
<td>DMAP</td>
<td>0.2</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>PPy</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>DMF</td>
<td>DMAP</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>Me₂SO</td>
<td>PPy</td>
<td>0.2</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

Table 15: DT-TM anhydride reactivity in various solvents. DNP-prolyl-threoninamide was trimesylated at a concentration of 1 mM with 100 mM DT-TM anhydride in different solvents in the presence of either DMAP or PPy, as specified in the table. The initial rate constants for the individual reactions were calculated from the HPLC analyses of quenched aliquots of the acylation reaction, as described in Figure 22.
solvents; PPY resulted in an initial rate constant that was twice that observed with DMAP, at the same concentration.

The choice of solvent for the acylation reaction must be made on the basis of its polarity and hence the reactivity of the reagent but also on its ability to support protein solubilization. Although CHCl₃ was one of the better solvents with respect to the reactivity of the DT-TM anhydride, it was not found to be an appropriate solvent for the solubilization of proteins, even in the presence of a hydrophobic counterion (data not shown). In DMF, the reagent reactivity was very good and furthermore, apobacteriorhodopsin was able to be solubilized in it when tetrabutylammonium ions were present (section III.1.). Pyridine, on the other hand, is routinely used in peptide chemistry and was found to be a very good solvent for protein solubilization (section III.1.). However, the initial rate constants of trimesylation in pyridine was 8 times slower than that found in DMF.

In order to select either DMF or pyridine as the appropriate solvent system, an additional criterion must be met; the DT-TM anhydride be stable under the protein derivatization conditions. The protein must be quantitatively and completely derivatized. To ensure complete modification, even at those sites that may be sterically protected, a long reaction time may be required. In order to assess the stability of the anhydride under the reaction conditions, the initial rate constants were calculated for the trimesylation of the model peptide, DNP-pcylthreoninamide, after different periods of preaddition of the anhydride in either DMF or pyridine, as detailed in the legend of Figure 32. The initial rate constant of trimesylation was found to be essentially identical over a period of 20 hours in pyridine when in the presence of equimolar amounts of PPY, clearly indicating that the DT-TM anhydride was very stable under these conditions. In DMF, the reactivity
Figure 32: Stability of DT-TM anhydride. Diprotected trimesic acid was added in pyridine (■) and in DMF (□) in the presence of 1 equivalent of PPY. At the specified time, an aliquot of the activated reagent was used to trimesylate DNP-prolylthreoninamide. The initial rate constant was calculated from the HPLC analyses of the acylation reaction (as outlined in Figure 22) and plotted against the time of addition as a percentage of the rate constant observed at time zero.
of the DT-TM anhydride was observed to rapidly decrease to 37% of its original rate within one hour and to less than 3% over 20 hours of preaddition. It was further noted that in the absence of the acylation catalysts, the initial rate constant was still found to decrease in DMF but at a slower rate (about 79% reactive after 5 hours when in the presence of 0.1 M Et₃N). It appears that as the reactivity of the acylating reagent increases in DMF, its stability decreases.

The overall conclusion, based on DT-TM anhydride reactivity and stability, was that pyridine was the superior solvent. The most appropriate trimesylation conditions for peptides and proteins would utilize the DT-TM anhydride in pyridine and in the presence of the acylation catalyst, PPy.

III.4.B.iii. Characterization of the Trimesylation Reaction

The solubilization of the membrane protein, apoBR, in anhydrous solvents requires the presence of a amphipathic salt, such as Bu₄N+Cl⁻ or Bu₄N+TFA⁻, which provides a hydrophobic counterion to the carboxyl groups of the protein, as described previously (section III.1.). These amphipathic salts may reduce the reactivity of DT-TM anhydride either through increased solvent polarity or in the case of Bu₄N+TFA⁻ through the formation of a mixed anhydride between DT-TM acid and TFA (as was observed with the DT-TM imidazolide). Therefore, in order to assess the effect of the presence of the phase transfer reagent, the initial rate of trimesylation was quantitated in the presence of increasing amounts of both of the amphipathic salts, Bu₄N+Cl⁻ and Bu₄N+TFA⁻.

As observed in Figure 33, the presence of either amphipathic salt, Bu₄N+Cl⁻ or Bu₄N+TFA⁻, at a concentration of only 5 mM decreased the initial rate constant to about 35% of its initial value. Further increases in the
Figure 33: The effect of amphipathic salt on the rate of trimesylation. DNP-prolylthreoninamide (25 μM) was trimesylated in the presence of 50 mM DT-TM anhydride, 50 mM PPY and with varying amounts of the amphipathic salt, Bu₄N⁺Cl⁻ (■) or Bu₄N⁺TFA⁻ (□). The initial rate constant was calculated from the HPLC analyses of quenched aliquots of the acylation reaction and was plotted as a percentage of the initial rate constant in the absence of salt and against the corresponding tetrabutylammonium salt concentration.
salt concentration did not substantially alter the initial rate constant; at 10, 15 and 20 mM salt concentration, the initial rate constant of trimesylation remained about 30-35% of its maximum value. These identical observations for the two amphipathic salts, led to the conclusion that the detrimental effect on the rate of reaction of the DT-TM anhydride was primarily due to the increase in the solvent polarity which stabilizes and therefore, reduces the reactivity of the anhydride/PPY adduct (Neises and Steglich, 1978). It would appear that upon stabilization of this acylium ion, any additional increase in the salt concentration does not result in additional stabilization therefore, its reactivity remains unaltered.

In general, catalysts of chemical reactions are employed in trace amounts due to their regeneration upon completion of the desired reaction. The catalysts DMAP and PPY are not consumed during the acylation reaction. Figure 34 illustrates the effect of increasing the concentration of each of the catalysts on the acylation of DNP-prolylthreoninamide. This was performed in the attempt to assess the optimal catalyst concentration required to achieve the maximum rate of acylation of a secondary hydroxyl group, under the conditions of apobacteriorhodopsin solubilization.

In the absence of the amphipathic salt, Bu₄N⁺TFA⁻, it was observed that as the molar ratio of catalyst to the DT-TM anhydride increases, the initial rate constant for trimesylation also increases, in an almost linear fashion. Independent of the catalyst employed, it appears that a maximal rate of trimesylation exists, approximately 8.0 x 10⁻¹ M⁻¹s⁻¹ in the absence of Bu₄N⁺TFA⁻. However, the amount of catalyst required to achieve this maximum rate was 3.0 equivalents of DMAP per DT-TM anhydride while only 1.5 equivalents of PPY were found to be necessary. These results appear to indicate that the large amount of catalyst ensures that the anhydride is
Figure 34: The effect of catalyst on the rate of trimesylation. DNP-prolylthreoninamide was reacted with 50 mM DT-TM anhydride at a concentration of 0.15 mM in pyridine in the presence (closed symbols) or absence (open symbols) of 20 mM Bu₄N⁺TFA⁻ and with varying concentrations of PPY (squares) or DMAP (triangles). Aliquots of the reaction were removed at specific times and the activated reagent quenched in aqueous DMF. These samples were then analyzed by HPLC under the conditions specified in Figure 22. The initial rate constant was determined from the HPLC chromatograms and plotted against the corresponding molar ratio of catalyst to anhydride.
primarily in the form of the reactive PPY adduct, thereby maximizing its concentration and hence the rate of acylation.

In the presence of 20 mM Bu$_4$N$^+$TFA$^-$, the optimum initial rate constant decreases to 2.2 x 10$^{-1}$ M$^{-1}$s$^{-1}$ (30% of initial) while the molar equivalents of PPY required to achieve this was only about 0.5. The presence of the amphipathic salt increases the polarity of the solvent which decreases the reactivity of the adduct through its stabilization. The reduction in the number of equivalents of catalyst required to achieve the maximum rate constant would indicate that in the presence of the salt, the formation of the acylium ion is not rate limiting. The optimal rate of trimesylation with DT-TM anhydride can therefore be accomplished in the presence of the amphipathic salt when 0.5 to 1.0 molar equivalents of PPY are used, with respect to the anhydride.

Having established the optimal proportions of catalyst to DT-TM anhydride, it was necessary to assess the rate of trimesylation of the model peptides under the conditions of apobacteriorhodopsin solubilization. All three of the model peptides were derivatized under the conditions outlined in the legend of Figure 35, however, only the results of the DNP-prolylthreoninamide were plotted; in the presence of PPY, DNP-prolylserinamide and DNP-prolylglycyltyrosinamide were found to be completely trimesylated in less than one minute, irrespective of whether Bu$_4$N$^+$TFA$^-$ was present or not. As indicated in the Figure, complete acylation of DNP-prolylthreoninamide was achieved within 3 minutes in the absence of Bu$_4$N$^+$TFA$^-$ and within 7.5 minutes in the presence of the salt. Furthermore, the rate of trimesylation was very poor in the absence of PPY, as anticipated.

These results indicate the ability of the DT-TM anhydride, under
Figure 35: Time course of model peptide trimesylation. DNP-prolyl-threoninamide was trimesylated at a concentration of 0.5 mM with 100 mM DT-TM anhydride in the presence (squares) and absence (triangles) of 100 mM PPY and in the presence (closed symbols) or absence (open symbols) of 20 mM Bu₄N⁺TFA⁻. An aliquot of the reaction was diluted into aqueous DMF and analyzed for the extent of acylation by HPLC, using the conditions specified in Figure 22.
these reaction conditions, to rapidly modify even the secondary hydroxyl
group of threonine. The complete derivatization of a protein, under similar
conditions, could require several times longer than the time required for
DNP-threoninamide derivatization, assuming that some of the acylation sites
within the protein may be sterically protected. Even if a long time period is
found to be necessary for complete protein acylation, the DT-TM anhydride is
stable under the reaction conditions for at least 20 hours (Figure 32).

III.4.B.iv. Apobacteriorhodopsin Trimesylation and Isolation

Prior to assessing apobacteriorhodopsin as a substrate for
trimesylation, the membrane protein had to be processed in order for it to be
soluble in the anhydrous pyridine, as was previously discussed in section III.1.
and detailed in Materials and Methods. To reiterate, the delipidated protein,
present in a solution of formic acid ethanol, was dried under reduced
pressure, in the presence of Bu$_4$N$^+$TFA$^-$ (100 μmoles/ mg of protein) and
from a mixture containing 4 volumes of toluene and 2 volumes of absolute
ethanol. The protein residue was twice redissolved in the toluene/ ethanol
mixture and redried under reduced pressure. The residue was then dissolved
in a small volume of pyridine, redried and then redissolved in an appropriate
volume of pyridine. The solutions of DT-TM anhydride and PPY were then
added to this anhydrous protein solution.

In order to adequately assess the time course of
apobacteriorhodopsin trimesylation, tritiated DT-TM anhydride was utilized
to aid the quantitation of the extent of acylation of the protein, as indicated in
the legend to Figure 36. Under the specified reaction conditions, complete
and fairly rapid trimesylation of apoBR was achieved; 100% of all Lys, Tyr, Ser
and Thr residues on the protein were derivatized within 1.5 hours. The
Figure 36: Time course of trimesylation of apobacteriorhodopsin. Apobacteriorhodopsin (150 μg) that had been solubilized in anhydrous pyridine in the presence of 15 μmoles of Bu₄N⁺TFA⁻, was reacted with 100 mM DT-TM anhydride (specific activity of 0.7 mCi/mmol) and 100 mM PPY. An aliquot of the reaction, containing 10 μg of protein, was removed at the specified times and was diluted into 0.5 M Me₃SiCH₂CH₂OH in pyridine and left for at least 0.5 hour. The sample was then applied to a Sephadex LH-20 column that was equilibrated and eluted with 10 mM Bu₄N⁺Cl⁻ in pyridine. Fractions of 0.5 mL were collected and the radioactivity assessed by diluting an aliquot of the fraction into Amersham aqueous counting scintillant and analyzed on a Beckman LS-7800 scintillation counter. The radioactivity eluting near the void volume of the column was expressed as the percent theoretical yield, normalized to the amount of protein recovered as established by amino acid analysis.
extent of trimesylation was determined by normalizing the radioactivity recovered on the protein to the actual amount of protein recovered from the Sephadex LH-20 column, which was assessed by amino acid analysis. (Amino acid analysis indicated that the protein recovery from the column was only 50-60% of that applied.) The time required for complete apoBR derivatization was about 10 times longer than the time required for the Thr residue on DNP-prolylthreoninamide to be completely acylated (90 minutes as compared to 7.5 minutes). This difference in time was likely due to steric hindrance within the protein thereby reducing the accessibility of particular sites for modification.

The trimesylation of apobacteriorhodopsin was found to be rapid and efficient under these anhydrous reaction conditions. In comparison to the DT-TM imidazolide derivatization of apoBR, the time required for complete trimesylation correlates to an approximate 80 fold decrease over the reaction in DMF and a 30 fold decrease in Me2SO.

In order to isolate and purify the trimesylated apoBR, the Me3SiCH2CH2OH quenched acylation reaction was applied to a Sephadex LH-60 column that was equilibrated and eluted with DMF, thereby, isolating the modified protein near the void volume of the column from the rest of the reaction mixture which would elute near the bed volume (Figure 37). As expected, this column provides good separation of the protein from the bulk of the quenched reaction mixture due to the molecular weight exclusion limits of the column which are 1.5 x 10^3 to 3.0 x 10^4. The recovery of the DT-trimesylated protein was 75% based on amino acid analysis.

The absorbance at 280 nm of a solution of this protein (337 µg in 1 mL) was found to give a reading of 3.0. (This absorbance was defined as an optical density unit (OD) due to the uncertainty of the extinction coefficient
Figure 37: Isolation of diprotected trimesyl-apobacteriorhodopsin from a Sephadex LH-60 column. Apobacteriorhodopsin (450 μg) that had been solubilized in anhydrous pyridine in the presence of 45 mmoles of Bu₄N⁺TFA⁻ was trimesylated with 100 mM DT-TM anhydride and 100 mM PPY. The reaction was left for 15 hours, at which time Me₃SiCH₂CH₂OH was added to a final concentration of 0.2 M and left to quench the activated reagent for an additional one hour. The entire sample was then applied to a Sephadex LH-60 column (1 x 28 cm) that was equilibrated and eluted with anhydrous DMF. Fractions of 1.15 mL were collected and the absorbance at 280 nm was assessed for each fraction using a Varian Cary 210 Spectrophotometer.
for this protein product.) This would correlate to a value of 8.9 OD/mg of DT-trimesyl apoBR, while underivatized apoBR has only 2.0 OD/mg. This increase in absorbance is presumed to be due to the 49 trimesyl groups that are now covalently attached to the protein.

This isolated modified protein still has the trimesyl carboxylate groups protected as the trimethyilsilylethyl ester. The DT-TM apoBR that was associated with the void volume was deprotected by the addition of Bu₄N+F⁻ and then applied to another Sephadex LH-60 column in order to isolate the deprotected trimesylated apobacteriorhodopsin from the hydrophobic counterion, Bu₄N⁺. With the use of an acidic eluent, such as 5% formic acid in ethanol, the carboxylate groups of the attached trimesyl groups would become protonated resulting in the formation of tetrabutylammonium formate which would elute near the bed volume of the column.

As observed in the chromatogram, (Figure 38) the TM-apoBR eluted as a sharp peak near the void volume but then trailed asymmetrically, as detected by absorbance at 280 nm. Amino acid analysis of the fractions of the effluent that were collected, confirmed the presence of protein at 6.5 through to 13 mL and in a relative proportion that mimicked the optical density peak shape. The total protein recovery from this column was only 67%. The observed trailing of the protein was likely due to the large amount of Bu₄N+F⁻ with which the trimesylated protein was applied; Bu₄N+F⁻ would enhance the hydrophobic interaction of the modified apobacteriorhodopsin with the matrix of the column. This retention would also account for the low protein recovery from the column.

After having isolated the trimesylated apobacteriorhodopsin from the Sephadex LH-20 column that was equilibrated and run in 5% ethanolic formic acid, the fractions containing the modified protein were pooled and
Figure 38: Isolation of deprotected trimesyl-apobacteriorhodopsin from a Sephadex LH-60 column. Isolated diprotected trimesyl-apobacteriorhodopsin (Figure 37) that had been concentrated to 300 μL, was deprotected by the addition of Bu₄N+F⁻ (10 equivalents per Me₃SiCH₂CH₂ group). After 0.5 hour, this reaction was diluted with an equal volume of 5% ethanolic formic acid and applied to a Sephadex LH-60 column (1 x 28 cm) that was equilibrated and eluted with 5% ethanolic formic acid. Fractions of 0.8 mL were collected and their absorbance measured at 280 nm on a Varian Cary 210 Spectrophotometer.
the pH slowly adjusted to 7-8 with Na₂CO₃ while continually vortexing the protein solution. The ethanol was then removed under reduced pressure leaving an aqueous solution of the trimesyl-apobacteriorhodopsin. This indicated that the membrane protein upon modification had now become completely soluble in an aqueous system. Subsequent work found that the derivatized protein, upon deprotection with Bu₄N+F⁻, could be directly transferred into water without the need of this ethanolic formic acid desalting column.

The total absorbance at 280 nm of the trimesylated protein upon isolation was found to be 2.35 OD/ 226 μg of recovered protein, which correlates to 10.5 OD/mg of trimesyl-apoBR. This increase in absorbance per mg weight of protein from 8.9 to 10.5 OD/mg was presumed to be due to the deprotection of the carboxyl groups.

III.4.B.v. Characterization of Trimesyl-Apobacteriorhodopsin

Having achieved the goal of converting the membrane protein, apobacteriorhodopsin, into a fully water soluble form, the behaviour of the modified protein under some of the standard techniques that are used to manipulate proteins was investigated. Specifically, its behaviour on molecular sieving columns for molecular weight assessment, its suitability as a substrate for enzymes and its behaviour on the amino acid sequencer.

III.4.B.v.a. Column Chromatography

The water soluble trimesyl-apobacteriorhodopsin was found to elute quantitatively in a symmetrical peak near the void volume of a Sephadex G-50 column that was equilibrated and eluted with only water (data not shown). This result confirmed the previous chromatographic
observations that were summarized in Table 11. However, use of this type of column matrix only showed that the protein was completely excluded, as would be expected.

A Sephadex G-150 column was then utilized in order to observe the behaviour of the trimesyl-apobacteriorhodopsin on a molecular sieving column, in a simple aqueous buffer. A 200 µg portion of the freshly deprotected trimesyl-apobacteriorhodopsin, that had been directly transferred into water and not passed through the ethanolic formic acid column (Figure 38), was applied to the Sephadex G-150 column. This column was equilibrated and eluted with 50 mM NH₄HCO₃, pH 8.5 and the resulting optical density profile (280 nm) of the trimesyl-apobacteriorhodopsin is illustrated in Figure 39. The protein yield from the G-150 column was not assessed by amino acid analysis, however, on the basis of the 280 nm absorbance, it appears that the protein recovery was only 25% of that applied.

The trimesyl-apoBR eluted as a fairly sharp symmetrical peak having a $K_{av}$ that corresponds to a 36,000 dalton globular protein (as determined from Figure 29) which is very close to its calculated molecular weight of 35,800. The previously discussed factors that influenced the elution behaviour of the trimesylated protein, after modification in DMF or Me₂SO, from Sephadex G-150 (Figure 28 and 30) would also be applicable here. These factors were in reference to the linearization of the trimesylated protein as the extent of modification increased and the retention of the modified protein due to hydrophobic interactions with the matrix of the column (Table 11). This particular sample of trimesylated apoBR was not isolated from a Sephadex LH-20 column after the deprotection reaction and the presence of the hydrophobic ion, Bu₄N⁺ may be the cause of the poor recovery of the protein from the column; it would be the counterion to the carboxylate
Figure 39: Elution profile of trimesyl-apobacteriorhodopsin from a Sephadex G-150 column. Apobacteriorhodopsin (200 μg) that had undergone trimesylation and subsequent isolation from a Sephadex LH-60 column, using the conditions specified in Figure 37, was deprotected with Bu₄N⁺F⁻ (10 equivalents per Me₃SiCH₂CH₂ group), diluted with 2 volumes of 50 mM NH₄HCO₃, pH 8.5 and applied to a Sephadex G-150 column (1 x 120 cm) that was equilibrated and eluted with 50 mM NH₄HCO₃, pH 8.5 and 0.02% NaN₃. The absorbance of the effluent was monitored at 280 nm by an Isco Absorbance Monitor.
groups on the protein thereby increasing the hydrophobic interaction with the matrix.

The absence of any absorbance eluting near the void volume of the column indicated that no large solubilized protein aggregate had been generated through some form of crosslinking of the protein with itself. Furthermore, the absence of a void volume peak also indicated that under these reaction and deprotection conditions no solvent related side-reaction occurred, as was observed in both DMF and Me₂SO when DT-TM imidazolide was used as the acylating reagent (Figure 28 and 30).

In the attempt to better assess the molecular weight of the trimesylated apobacteriorhodopsin, a Bio-Sil TSK 250 HPLC gel filtration column was utilized. The exclusion limits of the TSK column are $1.5 \times 10^3$ to $3.0 \times 10^5$ molecular weight for globular proteins and $1.0 \times 10^3$ to $1.0 \times 10^5$ for dextrans. The matrix of the TSK column is comprised of crosslinked silica which is then "end capped" which means that the hydroxyl groups of the silica are trimethylated.

Several water soluble proteins were applied and eluted with 20 mM Na₂HPO₄, pH 6.8 containing 0.1 M NaCl, in order to develop a standard curve for the log of the molecular weight versus the $K_{av}$ of the protein (Figure 40). For this particular column, bovine thyroglobulin ($M_r = 6.7 \times 10^5$) was used to determine the void volume ($V_0$) of the column while cyanocobalamin ($M_r = 1.35 \times 10^3$) was used to assess the total volume ($V_t$). The method of calculating the $K_{av}$ for the eluted protein was the same as that used for Figure 29. This curve was generated such that the apparent molecular weight of the trimesyl-apobacteriorhodopsin could be determined through its $K_{av}$ from the TSK column.

Figure 41 illustrates the elution profiles of trimesyl-apoBR from
Figure 40: $K_{av}$ versus the molecular weight of nine standard proteins as observed by gel permeation chromatography on a HPLC TSK-250 column. The HPLC TSK-250 column (0.75 x 30 cm) was equilibrated and eluted with 0.1 M NaCl and 20 mM Na$_2$HPO$_4$, pH 6.8 at 40°C. The standard proteins used were: human fibrinogen, $M_r$ = 340,000 (A); bovine intestinal alkaline phophatase, $M_r$ = 84,500 (B); bovine serum albumin, $M_r$ = 66,000 (C); chicken ovalbumin, $M_r$ = 45,000 (D); carboxypeptidase B-DBF, $M_r$ = 34,000 (E); carbonic anhydrase, $M_r$ = 30,000 (F); chymotrypsinogen, $M_r$ = 24,000 (G); sperm whale myoglobin, $M_r$ = 17,000 (H); and aprotinin, $M_r$ = 6,500 (I).
Figure 41: Elution profile of trimesyl-apobacteriorhodopsin from a HPLC TSK-250 column. An aliquot of the BioRad standard protein mixture that contained 40 µg of each of thyroglobulin, bovine gamma globulin and chicken ovalbumin, 20 µg of myoglobin and 4 µg of cyanocobalamin, was applied to a BioRad TSK-250 column (0.75 x 30 cm) that was equilibrated and eluted with 0.1 M NaCl and 20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 6.8 containing 20% Me<sub>2</sub>SO at 40°C (Panel A). In panel B, the same eluent was used to elute 2 µg of trimesyl-apobacteriorhodopsin that had been isolated from a Sephadex G-50 column (1 x 25 cm) that was equilibrated and eluted with H<sub>2</sub>O. For panels C, D and E, the TSK column was equilibrated and developed with 60% formic acid and 30% isopropanol at 40°C. An aliquot of the trimesylation reaction containing 16 µg of apobacteriorhodopsin was applied in panel C while for panel D, 2 µg of the derivatized protein was applied after it had been isolated from the Sephadex LH-60 column (Figure 37). This isolated diprotected trimesylated protein was then diluted into 5 equivalents of Bu<sub>4</sub>N<sup>+</sup>F<sup>-</sup> and 2 µg was reapplied to the TSK column (panel E).
the HPLC TSK column under various eluent solvents. Panel A shows the
typical profile observed when the standard protein mixture supplied by
BioRad was applied to the column; the inclusion of 20% Me$_2$SO was not
found to adversely effect their elution pattern. However, the trimesylated
apobacteriorhodopsin that had been isolated from the Sephadex G-150
column, did not elute from the TSK column in the presence of 0.1 M NaCl
and 20% Me$_2$SO (panel B).

The inability to elute the water soluble derivative of
apobacteriorhodopsin under standard elution conditions for typical water
soluble proteins indicated that a strong interaction must be involved between
the protein and the column matrix. This interaction appeared to be
hydrophobic in nature. In order to elute the trimesyl-apoBR, a mixture of
formic acid and isopropanol (60%: 30%) was utilized for the eluent. However,
under these conditions, all of the proteins in the BioRad standard mixture
eluted at the void volume of the column except for cyanocobalamin which
eluted near the bed volume of the TSK column. The modified
apobacteriorhodopsin was found to elute with the formic acid/ isopropanol
mixture (60%: 30%), resulting in the generation of an absorbance peak after
applying the neat trimesylation reaction, the isolated diprotected trimesyl-
apoBR and the deprotection reaction, as illustrated in panels C, D and E,
respectively (Figure 41). (The void volume contaminant in panels B and E
was shown to be derived from the deprotection reaction and occurs even in
the absence of protein.) When radioactive DT-TM anhydride was utilized in
the derivatization reaction, the protein peak having a 9.2 minute retention
time, correlated to being 100% derivatized as determined by the recovered
radioactivity after normalization to the amount of protein recovered, based
on amino acid analysis.
Trimesyl-apobacteriorhodopsin recovery from the TSK column was found to be only 32% based on the applied radioactivity. The poor protein recovery even under the isopropanol formic acid conditions was presumably due to the hydrophobic interaction between the trimesylated protein and the matrix of the column; the trimethylated silica of the column imparts a reversed phase nature to the column while the carboxylate groups of the trimesyl groups on the protein would be almost fully protonated under these elution conditions. Therefore, the modified protein tended to adsorb strongly to the matrix even in the presence of 30% isopropanol.

The fact that the standard water soluble proteins of various molecular weights were not retained within the matrix when the formic acid isopropanol mixture was used as the eluent indicates that the primary basis of retention observed for the trimesylated apobacteriorhodopsin must be due to its affinity for the matrix of the column and not due to molecular sieving. Hence, the TSK column was not very useful for the determination of the molecular weight of the modified protein nor in its isolation.

III.4.B.v.b. Enzyme Digestion

The suitability of the isolated water soluble trimesyl-apobacteriorhodopsin as a substrate for the deblocking enzyme, pyroglutamate aminopeptidase was assessed. This enzyme should remove the N-terminal pyroglutamyl residue from trimesyl-apobacteriorhodopsin provided that the protein is soluble and the N-terminus sufficiently exposed to be correctly positioned in the active site of the enzyme.

Prior to the addition of pyroglutamate aminopeptidase to the trimesyl-apobacteriorhodopsin, the enzyme was initially dissolved in 100 mM K₂HPO₄, pH 8 containing 10 mM EDTA and 30 mM dithiothreitol to achieve
a concentration of 100 Units/ mL. This enzyme solution was then dialyzed against 2 L of 10 mM K$_2$HPO$_4$, pH 8 containing 1 mM EDTA and 20 mM β-MSH at 4°C. The dialysate was changed twice with a minimum of 3 hours equilibration for each buffer. This was performed in order to remove the stabilizing reagents that were packaged with the enzyme, especially the sugars. (During the acid hydrolysis of the pyroglutamyl residue, the resulting amino group could undergo an acid catalyzed condensation with the aldehyde group of the sugar which would result in the formation of a N-substituted imine.) The activity of the enzyme was assessed before and after dialysis by monitoring the rate at which it hydrolyzes pyroglutamic acid β-naphthylamide. The rate of reaction was consistently found to be 0.96 - 1.0 nmoles/ min/ Unit enzyme before dialysis and 0.7 - 0.8 after dialysis.

In order to quantitate the enzymatic removal of the pyroglutamyl residue from trimesyl-apobacteriorhodopsin, the blocked amino acid residue was isolated by passing the enzyme digest reaction through a small Dowex 50x2 column that was equilibrated in the H$^+$ form. This column should bind the protein components of the reaction mixture and any compounds bearing an amine group. The pyroglutamic acid should be fully eluted with three bed volumes of water. The water was then removed under reduced pressure and the pyroglutamyl residue converted to glutamic acid by its reaction in 2 N HCl at 100°C for 1 hour. Prior to this reaction, 10 nmoles of either lysine or valine were added to each sample in order to ensure that any amine reactive contaminant will preferentially react with the amine groups of the added amino acid. Identification of the glutamic acid was achieved by derivatizing the freshly generated amine group with PITC, according to the reaction scheme in Figure 42. The derivatization procedure was done according to the PicoTag method of Millipore-Waters for amino acid analysis except that twice
Figure 42: Removal of the N-terminal pyroglutamyl residue from trimesyl-apobacteriorhodopsin by pyroglutamate aminopeptidase. Trimesyl-apobacteriorhodopsin (8.7 nmol) that had been isolated from Sephadex LH-60 in 5% ethanolic formic acid (Figure 38) was dried in the presence of 15 μmoles of K₂HPO₄ and 1.5 μmoles of EDTA. The protein residue was dissolved in 150 μL water, to which 10 μL of MeOH, 2 μL 1.4 M β-MSH and 24 units of pyroglutamate aminopeptidase were added. The digestion was performed at 37°C for 20 hours, with an addition of 2 μL of 1.4 M β-MSH after 5 hours. The reaction was then applied to a Dowex 50x2 column (in the H⁺ form) and the effluent was then processed according to the figure to yield the phenylthiocarbamoyl derivative of glutamic acid. The detection and quantitation of the modified amino acid was achieved by HPLC at 254 nm on an Altex Ultrasphere ODS column (0.46 x 15 cm) that was equilibrated with 3% of a 60% aqueous CH₃CN and 97% of an aqueous solution containing 140 mM sodium acetate and 3.6 mM triethylammonium acetate, pH 6.8. The column was developed according to the following gradient: in 1 minute the aqueous CH₃CN increased to 15% and then to 20% over the next 4 minutes.
Trimesyl-apobacteriorhodopsin

PGAPase: predialyzed

37°C for 18 h
100 mM K$_2$HPO$_4$, pH 8
50 mM 8-MSH, 10 mM EDTA

Pyroglutamate

Dowex 50x2 (H$^+$)
2 N HCl, 100°C for 1 h

Glutamic Acid

EtOH/ H$_2$O/ Et$_3$N/ PITC
(7 : 1 : 1 : 1)

Phenylthiocarbamoyl-glutamate
the reagent and Et$_3$N were used to ensure complete derivatization.

The resulting phenylthiocarbamoyl-glutamate derivative could be quantitated by applying it to a reversed phase HPLC column that was eluted with a shallow gradient of acetonitrile, as indicated in the legend. By this procedure, the phenylthiocarbamoyl-glutamate could be separated from the background contaminants. This was in contrast to the elaborate procedure that was required for the detection of the DNP-glutamate derivative (Figure 10).

In order to assess this procedure of quantitating the N-terminal residue, the following control samples were processed as per the reaction scheme of Figure 42. 10 nmoles of standard pyroglutamate was sequentially isolated, hydrolyzed and derivatized and the final recovery of the phenylthiocarbamoyl-glutamate was observed to be 90%. Secondly, 10 nmoles of the standard pyroglutamic acid β-naphthylamide was utilized to assess the extent of enzymatic cleavage that occurs under the specified conditions; based on the previous control, 87% was recovered and quantitated. When 8.7 nmoles of trimesyl-apoBR was used as the substrate, 6.8 nmoles of the expected phenylthiocarbamoyl-glutamate was obtained which correlates to a 98% recovery, after taking into account the losses observed with the standard samples.

The pyroglutamate aminopeptidase digestion experiment was then repeated using approximately 1 n mole of trimesyl-apoBR with different enzyme to substrate ratios in order to try to reduce the amount of enzyme required for complete removal of the pyroglutamyl residue. Several additional control samples were included: 1 n mole of pyroglutamic acid β-naphthylamide in the presence and absence of the enzyme and trimesyl-apobacteriorhodopsin without the enzyme present. These samples under
went incubation and passage through the Dowex 50x2 column as previously described and were assessed at the Protein and Amino Acid Services at the University of Toronto. The quantitation of the amino acids in the samples were performed by ion exchange chromatography followed by ninhydrin detection and the results are listed in Table 16.

As observed, the recovery of glutamic acid from the substrate pyroglutamic acid β-naphthylamide was 0.3 nmoles and 0.8 nmoles over the background affiliated with the enzyme when the enzyme to substrate ratio was 2:1 and 10:1, respectively. With these same two enzyme to substrate ratios, the net amount of glutamic acid from the trimesyl-apoBR were found to be 0.62 and 0.63 nmoles. Based on the overall recovery of glutamic acid from the pyroglutamic acid β-naphthylamide, these values would correlate to approximately 0.8 nmoles from the initial one nmole of protein. However, the control sample that contained trimesyl-apoBR but no enzyme also yielded 0.92 nmoles. Furthermore, repeating these experiments resulted in generating identical values for glutamic acid for each sample and as well as obtaining approximately 1 nmole from the trimesyl-apoBR that had not been exposed to the enzyme.

The conclusion from these results was that the trimesyl-apoBR must be eluting through the Dowex 50x2 column and that the subsequent acid hydrolysis is sufficient to cleave the pyroglutamyl residue off the trimesyl-apoBR as well as to open the ring structure. This was confirmed by determining the amino acid composition of the Dowex 50x2 effluent in the presence of HCl/propionic acid (50:50, v/v). The right amino acid composition for bacteriorhodopsin was obtained from the sample containing trimesyl-apoBR and pyroglutamate aminopeptidase while essentially no protein was obtained from a sample containing just the enzyme. The ability
<table>
<thead>
<tr>
<th>Sample</th>
<th>Reaction Components (nmol)</th>
<th>Glutamic Acid (nmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PGβNA</td>
<td>PGAPase</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 16: Determination of pyroglutamic acid release by pyroglutamate aminopeptidase digestion. Each sample was processed according to the legend of Figure 42 with respect to the digestion conditions, the passage through the Dowex 50×2 column and the conditions for acid hydrolysis. The samples were processed at the Protein and Amino Acid Services at the University of Toronto for amino acid analysis by ion exchange chromatography and ninhydrin detection.
of the deblocked trimesylated protein to pass through Dowex 50x2, even though it contains 7 arginine residues and has an amino group at its N-terminus, must be due to its overall large negative charge density. This unique behaviour may provide a novel approach for the removal of the enzyme, PGAPase, prior to sequencing the deblocked trimesyl-apoBR.

The detection of the pyroglutamyl residue after enzymatic cleavage of 1.4 nmol of trimesyl-apoBR was accomplished by concentrating the Dowex 50x2 effluent, adding 10 nmol of Lys and applying it to a Sephadex G-25 column that was equilibrated and eluted in water. Under these conditions the trimesyl-apoBR is excluded from the matrix and exhibits no peak trailing due to hydrophobic interactions (Table 11). The bed volume was pooled and concentrated under reduced pressure and then subjected to the acid hydrolysis conditions for conversion of the pyroglutamic acid to glutamic acid. These samples were also analyzed at the University of Toronto. After normalizing the values with respect to the Lys recovery, the PGAPase alone (Table 16; sample 9) yielded a value of 0.72 nmol while the enzyme digest of trimesyl-apoBR (Table 16; sample 10) yielded 2.44 nmol or 1.72 nmol after subtraction of the contribution made by the presence of the enzyme. This corrected value was approximately 20% higher than the amount of trimesyl-apoBR that was originally added to the reaction mixture, however, it was considered to be within experimental error.

The results indicate that the trimesylated protein was being utilized as a water soluble substrate for the enzyme, pyroglutamate aminopeptidase. This deblocked modified protein would be suitable for N-terminal sequencing through Edman degradation on an automated spinning cup sequenator or a gas-liquid phase sequenator.
III.4.B.v.c. Amino Acid Sequencing

The sequencing of the trimesylated apobacteriorhodopsin from the N-terminus was undertaken in order to assess its behaviour on the automated sequencer. This would be of interest because hydrophobic proteins are normally difficult to sequence; they have a tendency to be extracted by the organic solvents that are used to remove the sequencing reagents and the 2-anilino-5-thiazolinone amino acid derivatives upon cleavage from the protein. (Benzene, heptane, and butyl chloride are used for the spinning cup sequencer while heptane, ethylacetate and butyl chloride are used in the gas-liquid phase sequencer.) It was anticipated that this water soluble derivative of the membrane protein, apobacteriorhodopsin, would readily adhere to the spinning cup or the glass fibre filter when in the presence of the polymeric quarternary ammonium salt, Polybrene, through an ionic interaction with its negative carboxyl groups. Therefore, it was not expected to readily dissolve into any of the organic solvents.

Sequencing of the trimesyl-apoBR would also confirm that the modified membrane protein had been deblocked, that its N-terminal residue, pyroglutamate, has been removed through action of the pyroglutamate aminopeptidase. Sequencing will not occur if the pyroglutamate is still present on the protein because this particular residue has its amino group protected through an internal amide linkage with its side chain carboxyl group; its amine group is not available to undergo derivatization with PITC.

To date, all of the attempts to sequence the trimesylated apoBR have not yielded any amino acid sequence and there is no definitive explanation as to why this has occurred. In the experiments outlined in Table 17, the molar ratio of the PGAPase to the trimesyl-apoBR was 10:1 and was found to fully deblock the modified protein (previous section). Even though
<table>
<thead>
<tr>
<th>Experiment</th>
<th>TM-apoBR (nmoles)</th>
<th>PTH-Amino Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In Digest</td>
<td>After Dowex</td>
</tr>
<tr>
<td>A</td>
<td>1.27</td>
<td>1.27</td>
</tr>
<tr>
<td>B</td>
<td>0.80</td>
<td>n.d.</td>
</tr>
<tr>
<td>C</td>
<td>1.27</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.70</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17: Amino acid sequencing results from deblocked trimesyl-apobacteriorhodopsin. For each experiment, the amount of TM-apoBR was assessed in the deblocking digest, in the Dowex 50x2 and the Sephadex G-25 effluent and in the amino acid sequencer application. The PTH-amino acid residues that were observed in the HPLC chromatograms were listed for the first three cycles. "n.d." means that the amount of TM-apoBR was not determined while "---" indicates that the sample was not applied to that particular column. "*" identifies the PTH-amino acids that were obtained after subtracting the residues obtained for the PGAPase alone.
the reaction conditions were identical, the protein recovery from the Dowex 50x2 column appeared to vary from experiment to experiment; the protein recovery in Experiment A was 100% while only 66% and 5.5% was obtained in Experiment B and C, respectively. In all cases, these proteins were sequenced on the gas-liquid phase sequenator at the University of Toronto.

In Experiment A, the HPLC chromatograms for each cycle generally appeared to have polar contaminants as indicated by the absorbance peaks which occurred at the beginning of the gradient and which decreased with each subsequent cycle. Furthermore, a peak eluting at the position of PTH-Phe was present in all of the chromatograms; its identity was unknown. In the first cycle, a distinct peak corresponding to PTH-Ala was detected and which correlated to approximately 20 pmoles (2.2% yield). This was the correct amino acid residue that was expected based on the sequence of the deblocked apobacteriorhodopsin, even though its yield was low. The next cycle was expected to contain PTH-Gln. Due to the polar contaminants at the beginning of each HPLC chromatogram, the peak corresponding to this residue was not resolved. The third cycle was expected to contain PTH-Ile, however, no PTH-amino acid peaks were observed in the chromatogram. The glass filter to which the protein was applied was treated with HCl/propionic acid (50:50, v/v) and the protein content assessed. After the 15 sequencing cycles, it was found to contain 2 μg of the 23 μg that was originally applied. This would correlate to a 15% loss in the total protein after each sequencing cycle which would indicate that the protein was not being readily "washed" off the glass fibre filter.

In Experiment B, the effluent of the Dowex 50x2 column was concentrated and applied to a Sephadex G-25 column (9 mL total volume). The fractions eluting near the bed volume of the column were pooled and
the amount of glutamic acid was determined to be 1.72 nmoles (Table 16; sample 10) indicating that the pyroglutamyl residue was removed from the trimesylated apoBR. The fractions eluting near the void volume of the column which contained the trimesyl-apobacteriorhodopsin (as determined by amino acid composition) were pooled, concentrated and then applied to the sequencer. No PTH amino acid residues were observed in the HPLC chromatograms.

Experiment C was similar to Experiment A, except the recovery of the trimesyl-apobacteriorhodopsin from the Dowex 50x2 column was found to be quite low (5.5%). Sequencing of this protein yielded three PTH amino acids (Ala, Met and Glu) in the first cycle, all of which were less than 10% in yield. No PTH-amino acid residues were detected in subsequent cycles.

In Experiment D, the Dowex 50x2 column was eliminated in the attempt to sequence the trimesyl-apoBR in the presence of the PGAPase and then subtract from it any PTH-amino acid residues obtained from the enzyme alone. The PGAPase is known to have its N-terminus blocked and was, therefore, not expected to yield any PTH-amino acid residues. However, ten PTH-amino acid residues were detected in the first cycle for the PGAPase alone. A large quantity of enzyme (5 nmol) was utilized for the deblocking of trimesyl-apobacteriorhodopsin and it must have contained some protein impurities and/or fragmented enzyme. After subtracting the PTH-amino acid residues obtained for each cycle of the enzyme alone from those obtained for the trimesyl-apoBR plus the PGAPase, the resulting PTH-amino acids did not correlate to the expected sequence for the deblocked apobacteriorhodopsin. Furthermore, the expected PTH-amino acids were not even present in the chromatogram at the appropriate sequencing cycle.

All of the PGAPase digestion reactions were identical with respect
to their enzyme to substrate ratio, component concentrations, incubation
time and temperature. Under these conditions, it was shown that the
pyroglutamyl residue was removed and isolated near the bed volume of the
Sephadex G-25 column in Experiment B. Therefore, the trimesyl-apoBR in
each of the experiments would be expected to be deblocked. Even if the
digestion was incomplete due to minor experimental variations, the portion
of the protein that had undergone deblocking should be sequencable and yield
the correct PTH-amino acid residue for each cycle.

The absence of even the first amino acid residue which should be
Ala may indicate that the sequencing problem exists at either the PITC
coupling stage of the sequencing cycle or at the TFA cleavage step or at both of
these crucial steps. At both of these stages in the sequencing cycle, base or acid
are passed over the glass fibre filter as gases carried by a stream of argon: Prior
to the addition of the PITC reagent as a 5% solution in heptane, the protein
film is exposed to a stream of argon that has been bubbled through a 12%
solution of methylamine. This should neutralize the protein film and allow
the PITC to react with the amine group under basic conditions. After the
coupling stage, the cleavage of the derivatized amino acid residue is
accomplished by exposing the protein film to a stream of argon that has
passed through neat TFA. It may be possible that as a dry film on the glass
fibre filter, the trimesylated apoBR does not readily pick up the gaseous base
or acid thereby resulting in poor coupling and/or poor cleavage. This
conclusion is in part supported by the amino acid composition observations:
No amino acid residues were detected when the dried residue of trimesyl-
apoBR was subjected to HCl gas for 48 hours. The requisite amino acid
composition of BR was obtained when HCl/propionic acid (50:50, v/v) was
utilized for the protein hydrolysis.
III.4.B. vi. Conclusion

The activation of the diprotected trimesic acid to the symmetric anhydride provided a highly reactive derivatization reagent when it was used in pyridine, in the presence of the acylation catalyst, PPY. Under these conditions, the anhydride was stable for over 20 hours and resulted in the complete modification of all the amino acid side chain ε-amine and hydroxyl groups on the membrane protein, apobacteriorhodopsin. The deprotected trimesyl-protein derivative was found to have the following characteristics:

1. It was water soluble in the absence of detergents and chaotrophic agents. 2. Complete protein recovery, after its application to a Sephadex type column was achieved if the salt concentration of the eluent was kept below 20 mM. Higher salt concentrations resulted in significant losses in the protein, presumably due to adsorption to the column matrix through hydrophobic interactions.

3. The amino acid composition of the trimesyl-apoBR could not be obtained when the HCl gas phase hydrolysis procedure was employed; the correct composition was achieved when HCl/propionic acid (50:50, v/v) was used. 4. The trimesyl-apoBR was found to be a suitable substrate for the deblocking enzyme, pyroglutamate aminopeptidase; the pyroglutamyl residue was cleaved after an overnight digest at 37°C. 5. The deblocked trimesyl-apoBR was found to elute through a Dowex 50x2 column (H⁺ form) that had been equilibrated and eluted with water. This unique behaviour provided a novel procedure by which the enzyme was eliminated from the digestion mixture prior to the modified protein's application to the amino acid sequencer.

6. No PTH-amino acid residue was obtained when the deblocked trimesyl-apoBR was sequenced on the gas liquid phase sequencer. A possible explanation is that the modified protein film will not readily wet with the base and/or the acid which are both delivered in the gas phase in this type of
sequencer; without proper interaction of the protein with these reagents, then the PITC coupling stage and the cleavage stage will be compromised.

III.4.B.vii. Future Investigation

Research should be initially focussed on the behaviour of the trimesylated proteins on amino acid sequencers, both the gas-liquid phase and the pulsed liquid phase sequencer. This latter model delivers both the base and the acid in the liquid phase; this would eliminate the suspected problem of the trimesylated protein film being unable to react with the acid and base when in the gas phase. Unblocked trimesylated water soluble proteins could be used for this study in order to eliminate the complication of the PGAPase digest that is required for the trimesylated apobacteriorhodopsin. Prior to trimesylating unblocked proteins, their N-terminus must first be reversibly blocked with citraconic anhydride (Klapper and Klotz, 1972) or with 9-fluorenylmethyl chloroformate (Fmoc-chloride; Meienhofer et al., 1979). During the protection of the N-terminal amine group, the ε-amine group on the Lys side chain would also be protected from trimesylation. Deblocking of the protected N-terminus (and Lys side chains) after trimesylation, could be achieved by using either acidic conditions for the citraconic anhydride or 20% piperidine in an aprotic solvent for the Fmoc-chloride (Chang et al., 1980). The behaviour of these proteins and their ability to be sequenced could then be assessed on both types of sequencers. Furthermore, their application on to the glass fibre filter may require some modification (ie TFA activation of the filter) or the sequencing program may require some alteration with respect to reagent and solvent delivery times.

Additional work on apobacteriorhodopsin (or any other protein) would be substantially aided if the protein was initially labelled by reductive
methylation using $[^{14}\text{C}]$formaldehyde. This would enable the protein recovery to be easily and accurately quantitated after each reaction step and each column isolation. The extent of trimesylation could be monitored by the incorporation of $[^3\text{H}]$trimesic acid. With this dual labelled trimesylated apoBR, one could accurately monitor its recovery from the Dowex 50x2 column and assess its behaviour on both types of sequencers.

Another protein that could provide sequencer information is the N-Ac-β-endorphin, which is a small water soluble protein (31 amino acid residues) which has its N-terminus acetylated and it contains 4 Lys, 2 Ser, 3 Thr and 2 Tyr residues. This protein could be trimesylated under the standard derivatization conditions, treated with CNBr to cleave it into two fragments at its single Met residue (position 5) and then the digest could be directly applied to the sequencer. Only the C-terminal fragment would be amenable for Edman degradation. This simple procedure would generate information on the behaviour of trimesylated fragments on the sequencer without any additional chemical reactions or the inclusion of extraneous protein if an enzyme digest was performed.
IV. Conclusions:

The objective of this research was achieved: the membrane protein, apobacteriorhodopsin was extensively derivatized and rendered water soluble in the absence of detergents and chaotropic agents. This accomplishment was realized by establishing and achieving several goals which were as follows:

1. The solubilization of the apoBR in an aprotic anhydrous solvent which was required to ensure complete chemical derivatization. The solvents that were found to be suitable were DMF, Me₂SO and pyridine. However, in order for even the hydrophobic membrane protein to be soluble, tetrabutylammonium ions were required to provide hydrophobic counterions to the carboxylate groups on the protein.

2. The model peptides, DNP-prolylserinamide, DNP-prolylthreoninamide and DNP-prolylglyclytyrosinamide, were synthesized and used to assess and optimize the reactivity of the derivatizing reagent with primary, secondary and aromatic hydroxyl groups, respectively.

3. The derivatization reagent was selected on the basis of its ability to react with both amine and hydroxyl groups resulting in their conversion to carboxylate groups. Initially, succinic anhydride was chosen and was found to fully derivatize the apoBR resulting in a water soluble derivative. However, the succinyl esters were found to be unstable. To overcome this product instability, trimesic acid (1,3,5-benzenetricarboxylic acid) was selected as the derivatization reagent. Trimesic acid would also provide two carboxyl groups at each site of derivatization (Lys, Tyr, Ser and Thr residues). This chemical
was first converted to a monofunctional reagent by protecting two of its carboxyl groups as trimethylsilylethyl esters. Regeneration of the carboxyl groups can be quantitatively and selectively achieved by the addition of tetrabutylammonium fluoride, under anhydrous conditions.

4. The activation of the diprotected trimesic acid reagent required extensive investigation. Using the imidazolide derivative, complete derivatization of the apoBR in DMF or Me₂SO was achieved but only under rigorous conditions and lengthy reaction periods. Furthermore, the imidazolide reagent was not stable in either of these solvents. However, diprotected trimesic acid anhydride was found to have the highest reactivity towards the model peptides in pyridine, in the presence of the acylation catalyst, 4-pyrrolidinopyridine. Furthermore, under these reaction conditions the reagent was entirely stable for over 20 hours.

5. The deprotected trimesyl derivative at the sites of acylation on the protein should be stable: specifically the amide and ester bonds that are generated between one of the carboxyl groups on the trimesic acid and the ε-amine group or the hydroxyl group on the side chain of the amino acid residues, Lys, Ser, Thr and Tyr. The stability of the derivative was evaluated through the use of the model peptides and it was shown to be very stable in several aqueous solutions; only the Tyr peptide derivative exhibited a slow rate of hydrolysis in basic buffers and in the presence of β-mercaptoethanol.

6. Apobacteriorhodopsin was used as the model membrane protein and it was found to be extensively and completely derivatized under these reaction conditions, as determined by the incorporation of radioactive trimesic acid. Furthermore, the deprotected trimesylated apobacteriorhodopsin (TM-apoBR) was found to be completely water soluble in simple aqueous buffers (Morton and Gerber, 1988).
The trimesylated protein product was found to be unique with respect to its behaviour under standard protein manipulating techniques: 1. The TM-apoBR would elute through a Sephadex column provided the salt concentration was kept below 20 mM. Greater salt concentrations resulted in substantial protein loss, presumably due to hydrophobic interactions between the trimesyl groups and the sugar moieties in matrix of the column. Due to this type of hydrophobic interaction, the TM-apoBR would not elute from the HPLC TSK-250 protein sizing column. The novel behaviour of the TM-apoBR was further illustrated by the fact that it passed through a Dowex 50x2 column when the eluent was only water. 2. The electrophoretic mobility of the TM-apoBR was found to be so great that even when the pH of the polyacrylamide gel system was decreased to 2, its rate of migration was not reduced to that of unmodified standard proteins. This mobility was due impart to the large negative charge density imparted by the two carboxyl groups on each trimesyl group but primarily due to the modification of the ε-amino groups of all the Lys residues thereby eliminating their positive charges. 3. The TM-apoBR was shown to be a suitable substrate for the deblocking enzyme, pyroglutamate aminopeptidase; the N-terminal pyroglutamyl residue was quantitatively released over an 18 hour incubation. 4. It was also observed that when the TM-apoBR was dried to a film and subjected to HCl gas for over 48 hours, no amino acid residues were detected in the resulting residue. When HCl/ propionic acid (50:50, v/v) was utilized for 24 hours, the requisite amino acid composition was obtained. 5. Another unique behaviour of the deblocked trimesylated protein was that when it was applied and processed on the gas-liquid phase sequencer, no amino acid sequence was obtained. This may be due to the inability of the modified protein film on the glass fibre filter to wet with the gaseous trimethylamine
thereby preventing PITC coupling to the free amino group on the N-terminus or its inability to pick up the trifluoracetic acid thereby preventing the cleavage of the PTC-amino acid from the rest of the protein. This result was entirely unexpected and requires future investigation to fully assess and remedy the situation.

This method of converting a membrane protein into a water soluble form should greatly aid the general manipulation of these types of proteins for protein fragmentation and fragment isolation. The unique behaviour of these modified proteins will necessitate some minor modifications to the existing procedures, such as keeping the salt concentration in the eluent below 20 mM when performing column chromatography with a Sephadex type matrix. Furthermore, when their behaviour on the amino acid sequencers are fully understood, the chemically modified protein and peptide fragments should be amenable to sequencing.
V. References:


Drapeau, G.R. (1980) Qualityline 3, pp.11, Miles Laboratories, Inc.


Nature 302, 528-532.


VI. Appendix:


