INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

ENHANCING THE RUMEN STABILITY OF MB-1, A *DE NOVO* PROTEIN WITH POTENTIAL APPLICATION IN RUMINANT NUTRITION

A Thesis

Submitted to the Graduate Faculty

in Partial Fulfilment of the Requirements

for the Degree of

Master of Science

in the Department of Pathology and Microbiology

Faculty of Veterinary Medicine

University of Prince Edward Island

Jennifer J. Morrison

Hampton, P.E.I.

January, 1999

CONDITIONS OF USE

SIGNATURE PAGES

iii-iv

REMOVED

ABSTRACT

MB-1, or Milk Bundle-1, is a *de novo* protein with potential application in ruminant nutrition. Its artificially high content of the essential amino acids methionine, threonine, lysine, and leucine renders it a high quality source of limiting nutrients for the animal. However, whether expressed in rumen microbes or grown in transgenic crops, MB-1 is not stable enough for its intended use. In fact, it is labile at rumen temperature and would be rapidly degraded if exposed to rumen enzymes. With the goal of enhancing the rumen stability of MB-1, three mutants were designed and analyzed.

The first mutant, MB-1-Cys, was designed to have an exposed cysteine residue and be capable of forming an intermolecular disulfide bond. Titration of the mutant protein with *p*-hydroxymercuribenzoate demonstrated that the single cysteine had been introduced and that it was capable of binding heavy metals. Disulfide bond formation was spontaneous upon removal of reducing agents, as indicated by SDS-PAGE. The resulting MB-1-Cys Dimer was approximately twice the molecular weight of the monomer, as expected. Exposure to Pronase E revealed that the presence of the disulfide bond rendered the new protein twice as resistant to proteolytic degradation as MB-1. The two proteins had the same melting temperatures, thus the enhanced stability was not due to a change in conformational stability. Rather, it resulted from the reduction in exposed surface area associated with the close apposition of the MB-1-Cys monomers.

Construction of the second mutant involved the introduction of two cysteine mutations in such a way that they would be capable of forming an intramolecular disulfide bond and potentially increasing the thermal stability of the protein. Two different mutants were designed to serve as a mode of determining the connectivity of MB-1. For the purpose of purification, the mutations were to be introduced into the MB-1-His plasmid as well. Since the mutagenesis kit used was known to be efficient for introducing multiple mutations, attempts were made to introduce the two mutations into the plasmids simultaneously. Whether due to experimental conditions or to inefficiency of the kit used, only a few mutations were successful. However, the results obtained have suggested ways of improving the techniques for future trials.

For the third mutant, MB-1-Trp, the lone tyrosine in MB-1 was designed to be replaced by a tryptophan in order to improve the spectral properties of the protein and possibly increase its thermal stability. A mutant was obtained with the correct Trp codon, but an unexpected frame-shift in the 5' region of the tryptophan oligonucleotide used predicted a defective gene. Although the mutant gene is not of use for protein production, the results showed that the kit is efficient for introducing single mutations, and that completion of the mutants would be quick provided the cysteines for the intramolecular disulfide bond are introduced one at a time and the new tryptophan oligonucleotide is properly incorporated.

ACKNOWLEDGMENTS

I simply could not have completed this thesis project without the support of several individuals. First of all, I would like to thank my supervisor, Dr. Marc Beauregard. With his guidance, I have grown not only as a researcher, but also as a person. Thank-you, Marc. Dr. Fred Kibenge, my co-supervisor, and Mrs. Patty McKenna were instrumental in preparing me for the use of molecular biology techniques. Thank-you for your time and patience, especially during this past year. Dr. Liz Wirtanen and Dr. Jean Grundy were the two post-doctoral fellows that I had the opportunity to work with at the beginning of my program. Their advice was very helpful when trying to master new protocols, but it was with my colleagues Jillian MacCallum, Dayre McNally, Satnam Nijjar, Martin Williams, Nancy Haché, and Mylene Gagnon that I spent so many late nights in the lab. They celebrated with me when experiments worked and they were supportive when experiments failed. Their friendship made going to the lab everyday an enjoyable experience.

The members of my supervisory committee, Dr. Marc Beauregard, Dr. Fred Kibenge, Dr. Fred Markham, and Dr. M.S. Nijjar, had the difficult task of keeping my project focussed for the duration of the two years, and ensuring that my thesis was in a presentable format. Thank-you again for your suggestions and for taking such an interest in my progress. I am also grateful for the willingness of faculty and staff in the Chemistry, and Pathology and Microbiology Departments to assist me anytime I needed help. So often overlooked are the people in Graphics and Photography. I would like to say a special thank-you to Michelle, Shelley, and Tom for all the times they have come to my rescue with last minute slides and figures. Last, but certainly not least, I thank my family and friends for their ongoing love, patience, and moral support. I know you will always be with me wherever I am.

Over the past two years, this research was supported by Dairy Farmers of Canada, Canadian Space Agency, a PGS-A Natural Sciences and Engineering Research Council Grant, and most recently, The G. Murray and Hazel Hagerman Scholarship. This financial assistance was greatly appreciated. I would like to thank these agencies and families for their contribution to my education.

TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

TERM	ABBREVIATION
Commonly Used Terms	
Essential amino acid	EAA
Milk Bundle - 1	MB-1
maltose binding protein	MBP
Methionine	Met, M
Threonine	T
Lysine	K
Leucine	Leu, L
Tyrosine	Tyr, Y
Tryptophan	Trp, W
Cysteine	Cys, C
Histidine	His, H
three dimensional	3D
volatile fatty acid	VFA
Escherichia coli	*E. coli*
relative mobility	R_f
optical density	OD
Standard Units of Measurements	
degree Celsius	°C
centrifugal force	g
Dalton	Da
kiloDalton	kDa
litre	L
millilitre	mL
microlitre	μL
molar	M
millimolar	mM
micromolar	μM
molecular weight	MW, M_r
molecular weight cut off	MWCO
microgram	μg
nanogram	ng
nanometre	nm
maximum wavelength	λ_{max}

hydrogen ion activity	pH
isoelectric point	pI
percent	%
hours	h
minutes	min
seconds	sec
photons per second	photons/sec
melting temperature	T_m
absorbance units	a.u.

Chemicals

dithiothreitol	DTT
ethylenediaminetetraacetic acid	EDTA
ethylene glycol-bis(β-aminoethylether)-N,N,N',N'-tetraacetic acid	EGTA
calcium chloride	$CaCl_2$
sodium azide	NaN_3
sodium chloride	NaCl
sodium hydroxide	NaOH
isopropyl β-D-thiogalactopyranoside	IPTG
phenylmethylsulfonyl fluoride	PMSF
ribonuclease A	RNAse A
Cytochrome C	Cyt C
sodium dodecyl sulphate	SDS
polyacrylamide gel electrophoresis	PAGE
diethylaminoethyl	DEAE
p-hydroxymercuribenzoate	*p*-HMB

CHAPTER 1 - INTRODUCTION

1.1 PROTEINS

1.1.1 General Structure

Proteins are the most complex and important macromolecules found in the cells of living organisms. As polymers of amino acids, they perform a variety of functions which are necessary for an organism to exist. Some occupy a structural role, giving the cells shape as they either associate with lipids to form cell membranes or aggregate together as part of the cytoskeleton. Others serve as the major constituents of muscle or connective tissue. Enzymes comprise yet another class of proteins which function metabolically as catalysts, directing and accelerating biochemical reactions. In order to perform these tasks, a protein must fold into a conformation that is specific for the interactions involved (Zubay, 1998, Cheftel *et al.*, 1985).

Twenty amino acids serve as the building blocks for all naturally occurring proteins. Each amino acid has an amino group and a carboxylic acid group bonded to the same carbon, the α-carbon. Also bonded to the α-carbon is a side chain group termed an R-group (**Figure 1.1a**). It is this R-group which confers unique properties to each of the twenty amino acids. On the basis of their side chain properties, the amino acids can be divided into three main groups: polar, non-polar and charged (Campbell, 1991). **Figure 1.2** shows representative amino acids from each group. Methionine and leucine, with their long hydrocarbon side chains, are considered non-polar. Threonine has a free hydroxyl group which makes it polar in nature. Lysine is listed with the charged amino acids since it has a free amino group at the end of its side chain. At physiological pH, this amino group is charged.

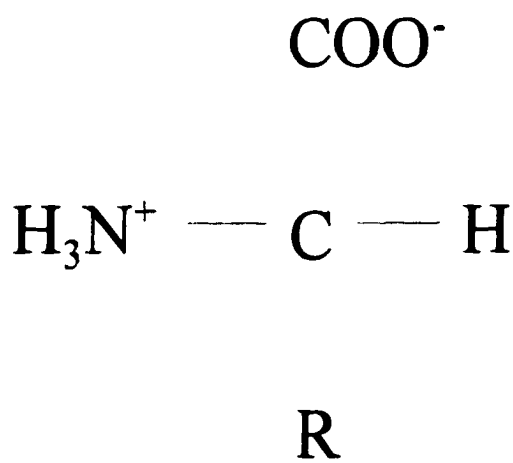


Figure 1.1a The Basic Structure of an Amino Acid. An amine group and a carboxylic group are separated by a single carbon (the α-carbon). The composition of the R group distinguishes between the 20 naturally occurring amino acids.

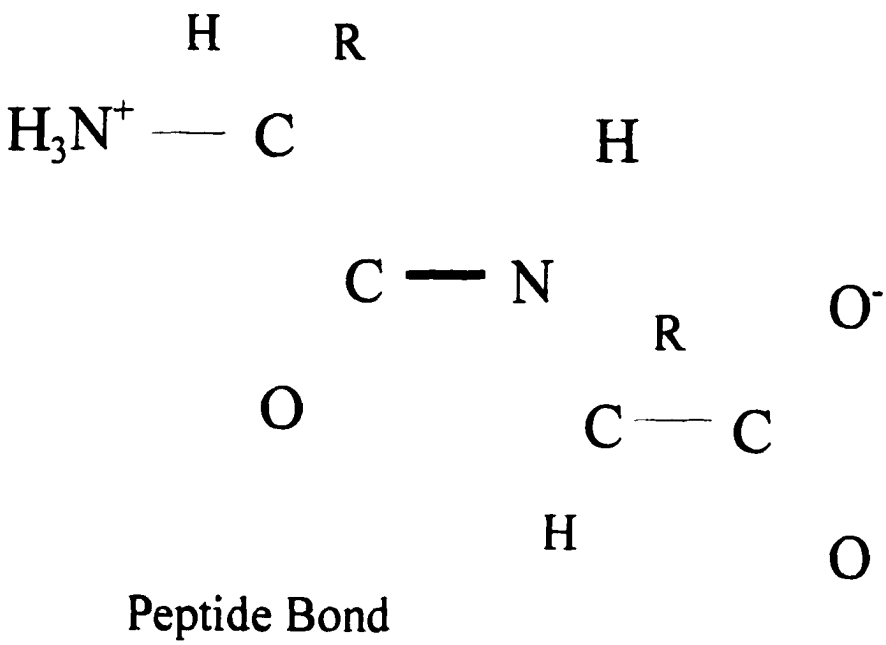
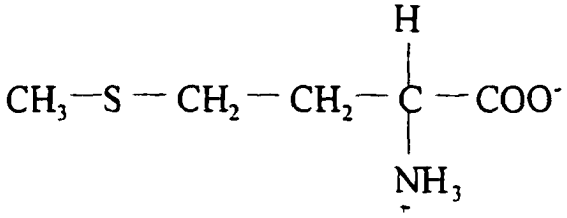
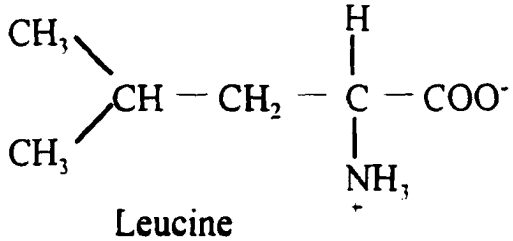


Figure 1.1b The Peptide Bond. In a protein, the amino acids are joined together by peptide linkage. A molecule of water is lost in the reaction as a peptide bond forms between the carboxyl group of one amino acid and the amino group of the other.

NON-POLAR

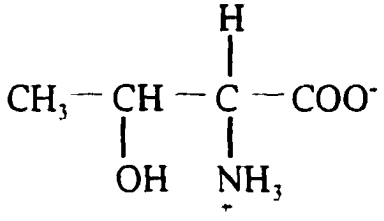


Methionine



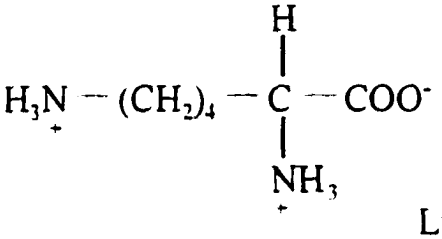
Leucine

POLAR



Threonine

CHARGED



Lysine

Figure 1.2 Amino Acids Representing the Three Main Groups. Methionine and leucine are non-polar on account of their long hydrocarbon chains. Threonine has a free OH group on its side chain, making it polar. Lysine has a large hydrocarbon portion, but its side chain amino group makes it charged at physiological pH.

In a given protein, the amino acids are linked via peptide bonds to form a linear sequence (**Figure 1.1b**). The difference in polarities of the amino acids, and their distribution within this sequence, determines the fold of the protein relative to the solvent it is dissolved in. Three levels of organization can be seen in the fold of a protein. As shown in **figure 1.3**, the linear sequence of amino acids represents the primary structure. Based on the amino acid arrangement in this sequence, certain regions will form local regular secondary structures, like α-helices or β-strands. Such structural elements are characterized by extensive hydrogen bond networks and repeating dihedral angles along the peptide bonds. These features specify the 3-D structure of the backbone, or polypeptide chain without R groups. Further packing of secondary structural elements to minimize the hydrophobic surface area results in the formation of one or several compact globular units called domains. At this level, the overall spatial arrangement of the peptide chain in three dimensions is referred to as the protein's tertiary structure (Zubay, 1998). The importance of the protein's primary structure should not be overlooked, however. It is this simple linear sequence of amino acids that dictates the ultimate structure and function of the protein.

1.1.2 Sources of Amino Acids

Amino acids required by most animals can be obtained either from dietary or bodily sources. Ingested protein is generally broken down into amino acids in the true stomach and intestine, and absorbed from the intestine into the bloodstream, providing an amino acid pool for bodily usage. In times of shortage, muscle protein stores can be mobilized or secretory cells can be stimulated to synthesize amino acids and proteins *de novo* (Mepham, 1986; Holmes and Wilson, 1984). Amino acids are classified as essential or non-essential. Histidine (H), Isoleucine (I), Lysine (K), Leucine (L),

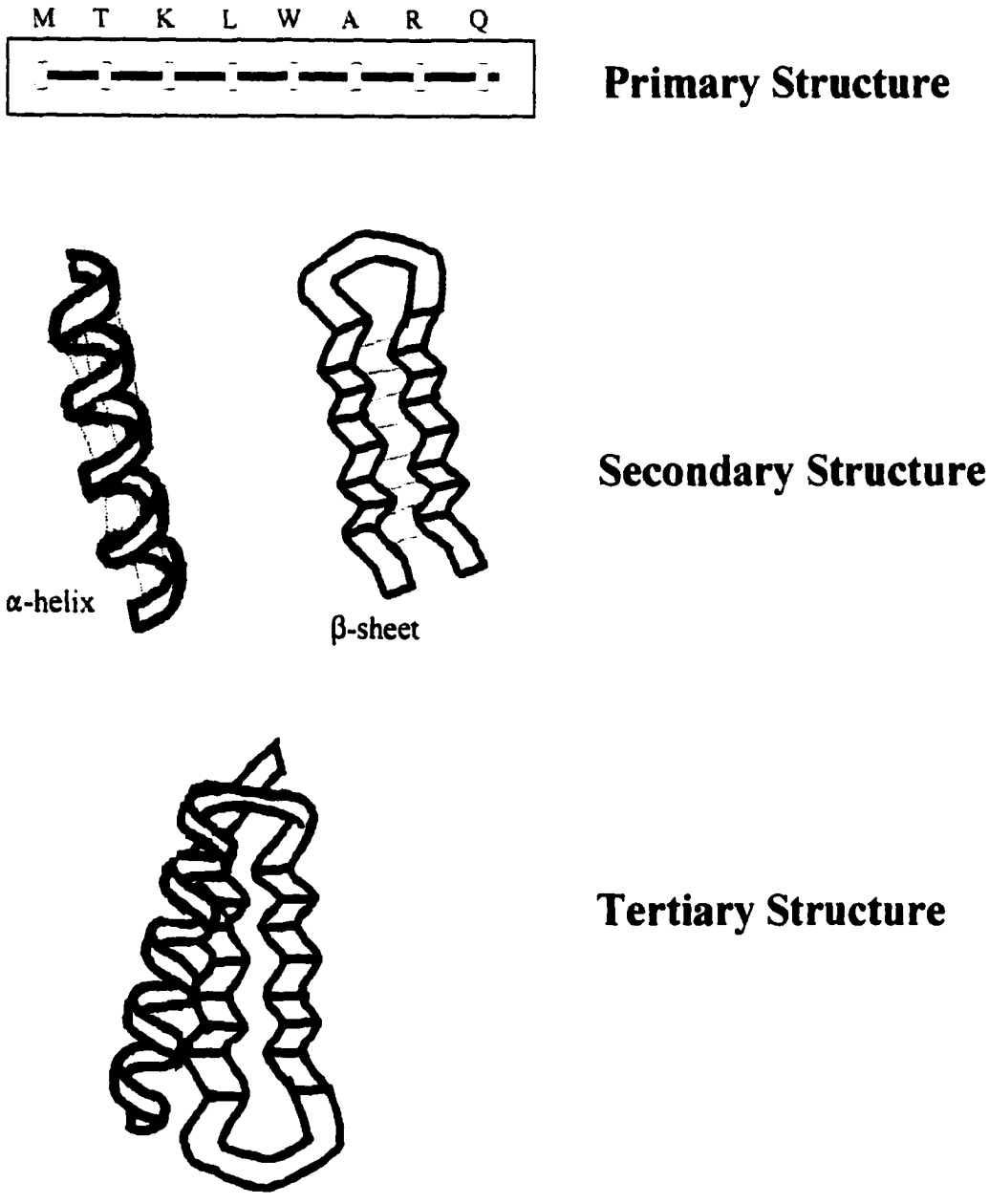


Figure 1.3 The Three Levels of Organization in a Folded Protein. Distribution of the amino acids within the linear sequence (or primary structure) dictates the fold and function of the protein. Hydrogen bond networks (---) and repeating backbone angles are seen in local secondary structures like α-helices and β-sheets. The tertiary structure of the protein is the 3D arrangement of these structural elements so as to minimize hydrophobic surface.

Methionine (M), Threonine (T), Phenylalanine (F), Tryptophan (W), and Valine (V) are considered essential amino acids (EAAs). Unlike the others, these nine cannot be synthesized by the body in times of need, therefore they must be taken in through the diet (Purser, 1970). The absorption of these EAAs from digested protein is therefore vital for the maintenance, reproduction, and growth of all animals, but it is of particular importance for high-producing dairy cattle.

1.2 PROTEIN IN HIGH-PRODUCING DAIRY CATTLE

1.2.1 Bodily Demand for Protein

The lactating mammary epithelial cell has been deemed one of the most metabolically active cells found in nature. During lactation, the mammary tissues achieve metabolic priority over all other tissues of the body. At peak times, approximately 80% of the energy available to the body is utilized by the mammary cells, powering various enzymes and transporters that are involved in substrate uptake, milk synthesis, and secretion (Larson, 1985). Except for energy, fiber, and water, lactating animals require more protein than all other nutrients combined (Miller, 1979). To accommodate the increased requirements of the priority tissue, the animal alters its pattern of nutrient partitioning among tissues. Hormone changes associated with lactation appear to regulate the metabolism of the tissues in ways which enhance the rate of substrate entry into the metabolic pool, diminish their utilization by other tissues non-essential to lactation, augment their partitioning to mammary glands by cardiovascular control, and facilitate that mammary uptake by stimulating appropriate transport systems located in cells (Mepham, 1986; Larson, 1985). Essentially, the onset of lactogenesis signals a nutrient drain towards the mammary gland to meet its metabolic demands. Central to metabolism is the synthesis of proteins, because all enzymes and most

membrane transporters which regulate cellular metabolism are proteins. The onset of parturition stimulates gene transcription and synthesis of these proteins as well as milk proteins (Eckert *et al.*, 1988). For this synthesis to take place, the amino acids must be in adequate supply.

1.2.2 Path of Digestion and Absorption

In ruminants, such as dairy cattle, approximately 60 to 70% of the protein that is taken into the body is degraded by rumen microorganisms when it reaches the rumen. Pathways illustrating rumen microbial activity and ruminant digestion are well defined (**Figure 1.4**). The microbes housed there have primarily membrane-bound proteases and peptidases that break solubilized protein down into peptides, amino acids, and ammonia (Hutjens *et al.*, 1996; Church, 1988). This hydrolysis at the cell surface permits direct access to the products of degradation and rapid absorption into the cell. The end products serve not only as nitrogen sources for the microorganisms, but also as sources of building blocks for other molecules or sources of carbon and energy when the peptides/amino acids are in excess (Church, 1988). Some of the deaminated amino acids are converted to important branched-chain fatty acids, for instance (n-valerate, isovalerate, isobutyrate, and 2-methylbutyrate). These particular fatty acids are essential growth factors and serve as precursors for bacterial cell wall components (Church, 1988). The rumen microbes use the ammonia, some sulfur and non-protein nitrogen, and a large portion of the liberated amino acids in order to synthesize their own protein. It is this microbial protein which serves as the primary source of amino acids for the ruminant animal (Hutjens *et al.*, 1996; Church, 1988; Chalupa, 1975; Hungate, 1966).

Rumen microbes generally contain about 65% of their dry matter as crude protein (Miller, 1979). They are harvested by the animal, and together with feed protein which

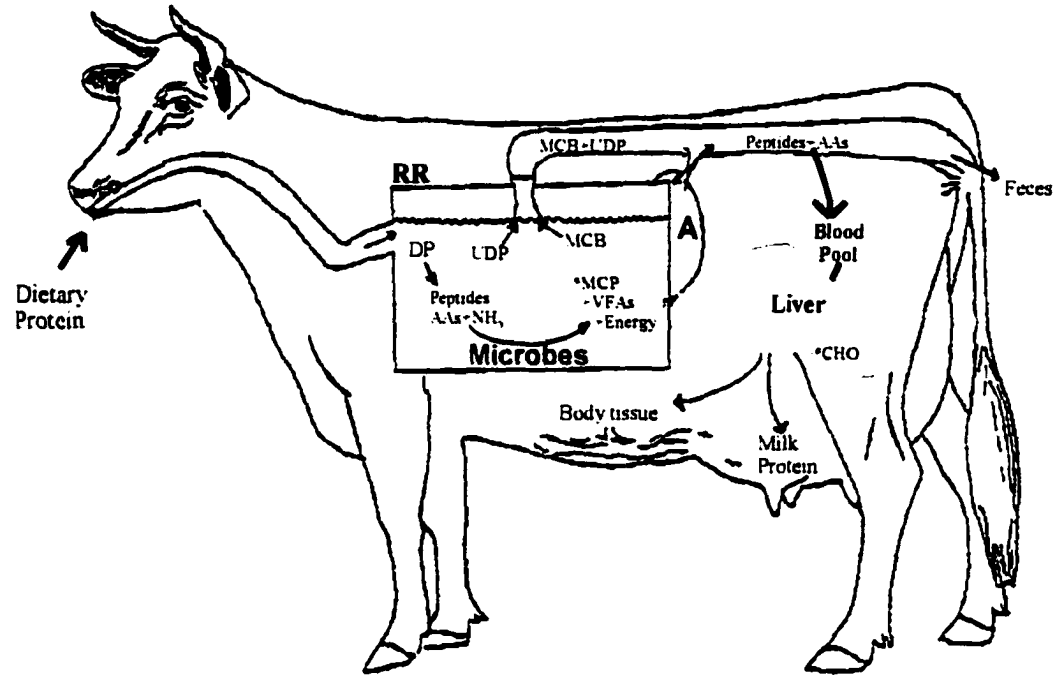


Figure 1.4 Path of Digestion and Absorption in the Ruminant. 60-70% of the protein that is taken into the body is degraded (DP) when it reaches the reticulo-rumen (RR). Microbes that are present there use the protein as a source of nitrogen, peptides, and amino acids. Although they use most of these endproducts to produce their own protein (MCP), they depend on some of the amino acids as a source of building blocks for other molecules like volatile fatty acids (VFA). The microbes (MCB) and undigested dietary protein (UDP) travel into the abomasum (A) where the low pH and activity of pepsin act to lyse the bacterial cells and hydrolyze the proteins. Further digestion of the proteins is achieved by the digestive enzymes (pancreatic and hepatic secretions) that are met in the small intestine. Amino acids and some small peptides are absorbed there and then transported by the bloodstream to the liver. To meet the high demands of the mammary glands, a large portion of the amino acids are carried to the mammary epithelial cells to be used for bodily and milk protein synthesis. Other amino acids are used for maintenance and growth of body tissues. In times of need, the animal draws upon amino acids as a source of glucose and energy (CHO). This gluconeogenesis takes place in the liver. Protein that is not digestible by the animal is excreted in the feces. (Modified from Church, 1988.)

escapes or bypasses degradation in the rumen, travel further down the digestive tract into the omasum and then the abomasum or true stomach (Church, 1988; Hungate, 1966). There, the low pH (2) and to some extent the change in osmotic pressure induce lysis of the microbial cell walls and unfolding of the proteins. Since the cell walls are high in lipoprotein, pepsin in the gastric juice augments the lysis process and initiates protein breakdown through cleavage of peptide bonds (Asplund, 1994). As the bolus passes into the small intestine, alkaline secretions are added from the liver, pancreas, and intestinal mucosa. When the pH approaches neutrality (7-7.5), the trypsin, chymotrypsin, and carboxypeptidase A in the secretions begin to degrade the proteins and peptides more extensively (Church, 1988). Liberated amino acids and short peptides are then transported across the intestinal wall and carried via the portal vein to the liver. There, they may pass directly into the systemic circulation and travel to other tissues of the body, or they may join the body's general amino acid pool to be used for protein synthesis. In addition to supplying building blocks for milk and bodily proteins in the mammary gland and bodily proteins in other tissues, the amino acids may be drawn upon to produce glucose in the liver during times of energy shortage. This is a metabolically expensive pathway, however, because the nitrogen released during deamination must be captured in the form of urea and excreted (Hutjens *et al.*, 1996). Any excess amino acids pass from the tissues back to the liver where they are either used for one of the above synthetic pathways, or converted to ammonia and keto acids.

1.2.3 Essential Amino Acid Requirements

About 60% of the amino acids of the animal body consist of essential ones compared to 40% that are non-essential (Miller, 1979). The tissues of cattle, as in other animals, cannot synthesize the carbon chain of these essential amino acids. So, in order

to maintain such a high level of essential amino acids, the ruminant depends upon the dietary protein and microbes that reach the intestine. In many ruminants, the EAA content of microbial protein and undigested feed protein is sufficient to meet their dietary needs. However, the amino acid synthesis by microbes is not sufficient to meet the requirements for rapid growth and high production in domestic ruminants (Church, 1988). In fact, microbial degradation of feed proteins during lactation may exceed microbial protein synthesis and result in a net loss of dietary protein. The impact of this rumen microbial degradation is not limited to protein quantity; it affects protein quality as well. Microbial protein synthesized in the rumen does not possess an ideal balance of essential amino acids. The result is a fairly large deficit in the amino acids methionine (M) and lysine (K), and a lesser deficit in other amino acids including threonine (T) and leucine (L) (Orskov *et al.*, 1987; Chalupa, 1975; Stern *et al.*, 1985; Ashes *et al.*, 1995). Farmers try to meet these requirements by using expensive supplements like roasted soybean and/or protected EAAs or other by-products such as fish meal.

1.3 OPTIMIZED FEEDING TECHNOLOGY

1.3.1 Benefit of EAA Enriched Proteins

Various diets have been evaluated in terms of digestion and nutritional value for dairy cattle. These diets included forages, concentrates, and protein supplements - the usual source of protein. As of yet, no diet has been found to provide an essential amino acid profile that meets the nutritional requirements of high-producing dairy cattle (Rulquin *et al.*, 1995; Orskov *et al.*, 1987). The amount and quality of protein produced by the animal is therefore limited by the availability of essential amino acids. Since enzymes, transport proteins, and milk proteins are necessary for milk production, the quantity and quality of milk produced could be decreased when EAA levels are not

adequate. The health of the lactating cow may also be compromised. Insufficient protein can have multiple effects on the animal, depending on the extent of the shortage. For instance, when deficient in protein, cows will lose weight as microbial fermentation is decreased and less feed is available for ingestion. The animals also become less resistant to diseases and infections (Miller, 1979). Consequently, farmers have to feed their animals a larger quantity of bulk protein to provide the limiting ingredients - a situation which is neither efficient nor cost effective. A search for a more efficient feeding technology has long been the focus of many researchers.

Studies have shown that milk and milk protein production are significantly increased when livestock are fed rumen protected essential amino acids or protein that is high in essential amino acids and resistant to microbial breakdown (Rulquin *et al.*, 1995; Armentano *et al.*, 1997; Orskov *et al.*, 1987; Faldet and Satter, 1991; Broderick and Clayton, 1992; Papas *et al.*, 1984; Robinson *et al.*, 1995). Methods used to protect these proteins and amino acids from ruminal degradation include roasting, chemical treatment, and encapsulation (SmartAmineTM by RhonePoulenc) (Chalupa, 1975; Stern *et al.*, 1985; Ashes *et al.*, 1995; McNiven *et al.*, 1994). These techniques may render the protein resistant to degradation in the rumen, but they also increase the cost of feeding, and often make the proteins indigestible in the intestine (Tamminga *et al.*, 1979; Tanner *et al.*, 1994; Stern *et al.*, 1985; Chalupa, 1975). In some cases, the availability of certain EAAs have been selectively decreased by such treatments (Clark *et al.*, 1992; Robinson *et al.*, 1995).

Postruminal injection of high quality proteins or essential amino acids alone has also been found to increase the quantity and quality of milk (Rulquin *et al.*, 1995; Orskov *et al.*, 1987; King *et al.*, 1990; Rogers *et al.*, 1979; Broderick *et al.*, 1971; Schwab *et al.*, 1976). This technique would certainly not be useful for feeding animals on a daily basis.

Furthermore, naturally occurring proteins rich in essential amino acids like M, T, K, and L are rare (Beauregard *et al.*, 1995). With the advances in protein design that have been made over the past decade, a possible alternative for delivering the EAAs post-ruminally would be to create a protein that is high in essential amino acids, and naturally resistant to rumen degradation. Such a protein would be beneficial not only to the dairy industry, but also to beef and sheep wool industries which depend on high levels of growth and production in ruminants.

1.3.2 Earlier Attempts

Previous attempts to create an EAA enriched proteins involved altering the sequence of natural proteins to the extent that they would serve as a candidate for protein supplement. The changes in amino acid sequences merely destabilized the polypeptides, however, thus efforts of researchers have leaned towards designing and producing a protein *de novo* (Dyer *et al.*, 1993). The first synthetic genes coding for amino acid enriched proteins took the form of simple, repeating DNA segments that, when expressed in *Escherichia coli (E. coli)* cells, translated into essential amino acids (Doel *et al.*, 1980; Jaynes *et al.*, 1985; Beauregard *et al.*, 1994). This approach presented several problems. Highly repetitive DNA segments are quite unstable unless appropriate conditions are met (bacterial strains should be recombination deficient [rec-]). Their expression and lifetime in the bacterial cytosol were so limited that they required radio-labelling or *in vitro* translation for detection despite the use of optimal bacterial strains and vectors. Moreover, the translated polypeptides with their repeating sequences deviate from the natural protein format. In terms of nutrition, the stability of such proteins is inadequate for effective use as a feed additive, and in fact, they can have toxic effects (Hagen and Warren, 1982; Boebel and Baker, 1982; Beauregard *et al.*, 1995; Maldague *et al.*, 1991).

Beauregard *et al.* (1995) postulated that by finding a suitable structure for the selected essential amino acids and encoding the required folding elements into the primary sequence, a more stable, non-toxic protein could be produced. In short, they attempted to design a protein that would adopt a stable tertiary fold in spite of an artificially high content of the essential amino acids M, T, K, and L. The end product would represent the first *de novo* protein designed to have a potential application in ruminant nutrition (Beauregard *et al.*, 1995).

1.3.3 MB-1: First *de novo* Attempt

MB-1, or Milk Bundle - 1, is the *de novo* protein that was designed by Beauregard *et al* (1995). All previous *de novo* designs had concentrated on choosing the amino acids that would best fit a desired structure. In this case, the researchers were credited with being the first to design a protein to fit the desired sequence (Beauregard *et al.*, 1995). Information on the 3D structures of natural proteins had revealed many rules and restrictions that apply to the folding of proteins. For instance, the earliest recognized, best understood, and most fundamental structural motif of natural proteins is the α-helix. It was known that certain amino acids, because of their characteristics, would increase a peptide chain's propensity towards assuming a certain conformation like the α-helix. Since three of the four desired amino acids were considered good helix formers (M, K, and L), the α-helical bundle was chosen as the tertiary fold to model in the creation of MB-1 (Beauregard *et al.*, 1995). Furthermore, of the different types of protein folds that exist, the rules governing α-helical bundle folding were the best understood (Bryson *et al.*, 1995). It was hoped that incorporation of the essential amino acids into this type of fold would result in the production of a stable protein which could be used for nutritional purposes.

1.4 MB-1: A DESIGNED PROTEIN WITH A TAILORED CONTENT OF EAAs

1.4.1 Design and Structure

As mentioned, MB-1 was designed to adopt an α-helical bundle fold. In order to understand the design process, it is imperative to describe the proposed structure of MB-1. **Figure 1.5** shows a ribbon model of a right-handed α-helical bundle. It is comprised of four α-helices joined together by loop segments with the hydrophobic residues arranged in such a way that they collapse together hiding themselves from the polar solvent. This collapse of the hydrophobic core is the driving force behind bundle formation (Betz and DeGrado, 1996). A top view of the bundle is also seen in **figure 1.5**. From this angle, the recurring pattern of amino acids is evident every two turns of the helix. The heptad pattern is denoted by the letters of the alphabet, with residues b,c, and f occupying exposed positions and a and d occupying buried positions (Cohen and Parry, 1990).

The design of MB-1 helices involved looking at four naturally occurring α-helical bundles and finding a consensus for the positioning of amino acids within the heptads. In addition to their statistical occurrence, the amino acids found in each position were evaluated in terms of their hydrophobicity, volume, and charge. This consensus residue or residue profile approach reflects the assumptions that the micro-environment of each position in a protein is specific and is reflected in the properties of the amino acids that fit that particular position (Beauregard *et al.*, 1995). Once a consensus was obtained for each of the heptad positions, trial helices were constructed so as to maximize essential amino acid content and yet maintain proper folding elements (e.g. charge distribution, residue volume, pI, weight). For some helices, many cycles of optimization and secondary structure prediction were required before the sequence satisfied the considerations. Loops connecting the helices were also designed so as to maximize EAA

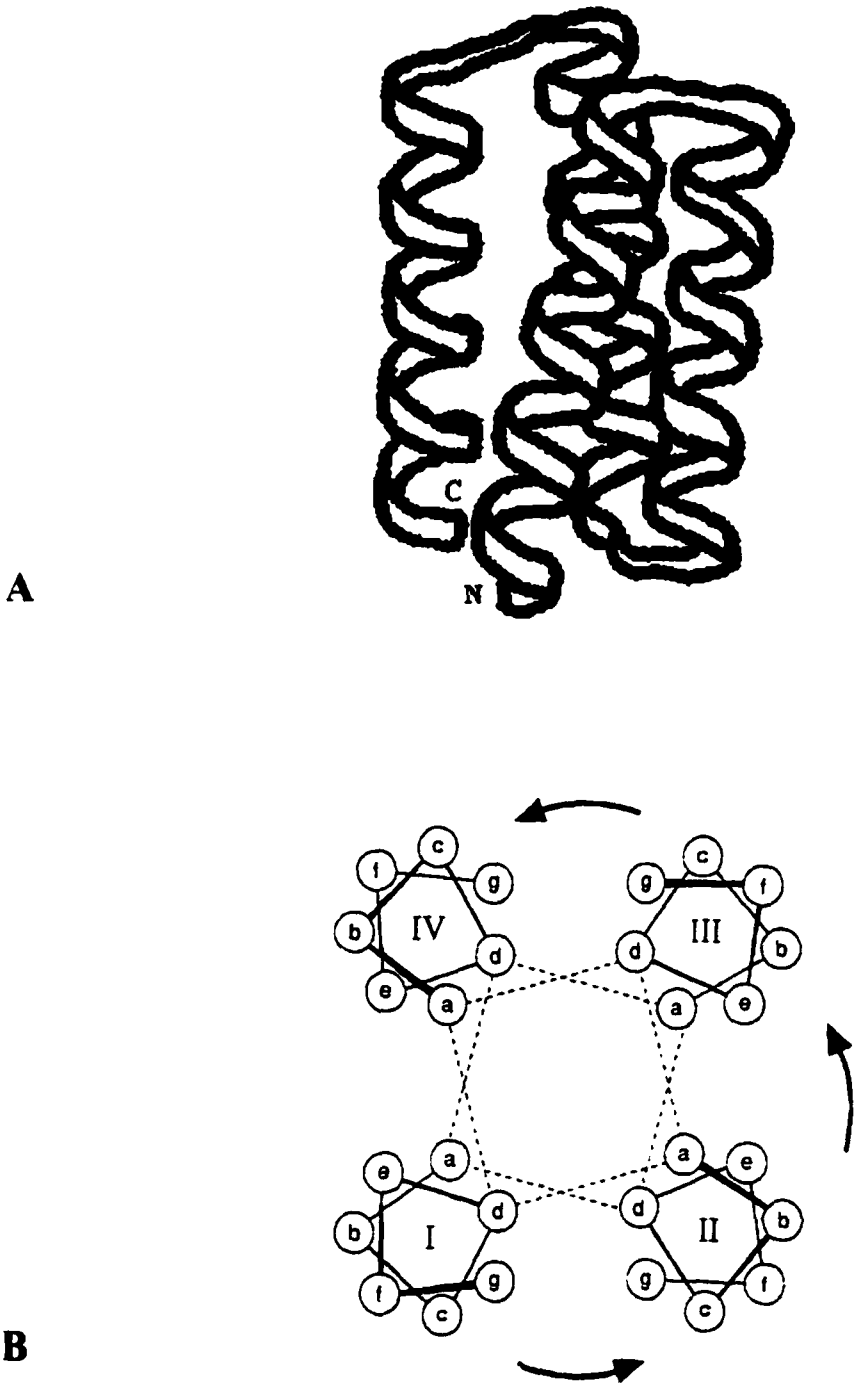


Figure 1.5 The Designed Fold of MB-1. A ribbon diagram (A) represents the right-handed α-helical bundle. Four α-helices are joined together by short loop segments with the hydrophobic amino acids arranged in such a way that they collapse together to form the bundle core. Looking down at the top of the bundle (B), the hydrophobic amino acids are seen primarily in positions 'a' and 'd' of the helix heptads. Residues in positions b, c, and f are more polar as they are exposed to the environment. 'e' and 'g' positions are borderline.

content, but their lengths varied considerably in naturally occurring proteins, so no consensus was obtained. Rather, short segments of small and hydrophilic residues were used to refine properties of the whole protein (e.g. pI and charge). Amino acids that would have a high turn forming propensity, but have minimal effect on adjacent helices were favored. In order to improve the protein lifetime in *E. coli*, aromatic and paired basic residues were avoided in loops, and the size and pI of the intact protein were optimized (less than 35kDa and pI between 7 and 8; Beauregard *et al.*, 1995).

The 100 amino acid sequence chosen for MB-1 is shown in **figure 1.6**. Here the bundle is represented in an opened out and flattened manner with dashed line circles depicting the buried positions in the bundle core. As indicated in the picture, 57% of the protein is made up of the essential amino acids methionine (M), threonine (T), lysine (K), and leucine (L). This high content of essential amino acids makes MB-1 a possible candidate for ruminant nutrition.

1.4.2 Production and Structural Characterization of MB-1

Since MB-1 is a *de novo* protein, it initially had to be produced via recombinant DNA technology. However, unlike the short repetitive peptides discussed earlier, the sequence for MB-1 resembles natural proteins, containing 16 different amino acids. As a result, highly repetitive and unstable DNA sequences could be avoided when creating the gene for recombinant expression (**Figure 1.7**; Beauregard *et al.*, 1995). The 11 kDa protein was therefore produced in *E. coli*. Amino acid analysis of MB-1 confirmed its expected content of essential amino acids (57%), while circular dichroism (CD) spectra indicated that it is indeed of considerable helical content (approximately 60%) (Beauregard *et al.*, 1995). Quenching studies revealed the position of the lone tyrosine residue (Y62) in MB-1 as being buried as per design (MacCallum *et al.*, 1996). The fact

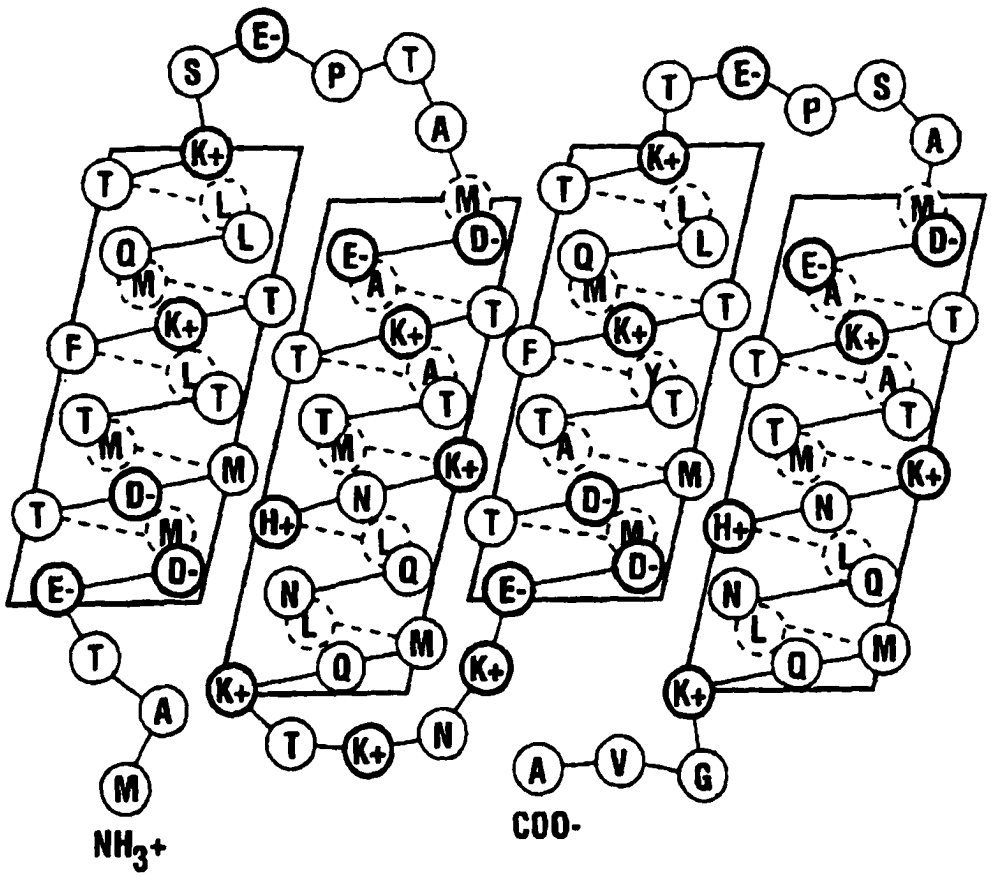


Figure 1.6 Amino Acid Sequence of MB-1. This opened out, flattened view of the MB-1 sequence shows the hydrophilic face of the protein in solid lines and the hydrophobic face in dotted lines. Fifty-seven percent of the 100 amino acids are selected EAAs methionine (M), threonine (T), lysine (K), and leucine (L).

```
5' - ATG GCT ACG GAA GAC ATG ACC GAC ATG ATG
      M   A   T   E   D   M   T   D   M   M

     ACC ACC CTG  TTT AAA ACT ATG CAG CTG TTG
      T   T   L   F   K   T   M   Q   L   L

     ACC AAG TCG GAA CCC ACG GCT ATG GAC GAG
      T   K   S   E   P   T   A   M   D   E

     GCC ACT AAA ACG GCT ACT ACA ATG AAG AAT
      A   T   K   T   A   T   T   M   K   N

     CAT CTT CAA AAC CTG ATG CAG AAG ACT AAG
      H   L   Q   N   L   M   Q   K   T   K

     AAC AAA GAA GAC ATG ACG GAC ATG GCC ACT
      N   K   E   D   M   T   D   M   A   T

     ACG TAC TTC AAA ACG ATG CAG TTG TTA ACG
      T   Y   F   K   T   M   Q   L   L   T

     AAG ACC GAC CCC TCG GCC ATG GAC GAG GCC
      K   T   E   P   S   A   M   D   E   A

     ACG AAG ACG GCT ACA ACC ATG AAA AAT CAT
      T   K   T   A   T   T   M   K   N   H

     CTG CAG AAC TTG ATG CAA AAA GGC GTA GCT
      L   Q   N   L   M   Q   K   G   V   A

     TAA - 3'
```

Figure 1.7 Sequence of MB-1 and Corresponding Gene. Each amino acid in MB-1 is denoted by a single letter below its corresponding codon. Since MB-1 contains 16 of the 20 amino acids, Beauregard *et al.* (1995) were able to create this stable, non-repetitive gene for expression in bacteria.

that the protein is highly helical and that the tyrosine is buried within the core suggests that MB-1 is folded as predicted.

1.5 PROTEOLYTIC STABILITY

1.5.1 Proteolytic Stability in the Rumen

For MB-1 to be used as an essential amino acid supplement for dairy cattle, it must be easily delivered to the animal. Two methods of protein delivery are currently under investigation. In the first case, rumen bacteria are being genetically engineered to produce the protein themselves (Teather *et al.*, Ruminal production system for MB-1 derivatives, patent pending). Introduced into the rumen environment, they would produce the protein intracellularly and serve as a major source of protein for the animal, provided the protein is resistant to intracellular proteases. This approach would minimize cost while providing protection from the rumen environment. The second approach involves production of transgenic feed crops by introducing the MB-1 gene into the plant DNA. Since the sowing, growing, and harvesting of such crops is no different than for the wild-types, this alternative would be easy to integrate into farming practices. In any case, both methods have several advantages over chemically and physically modified "by-pass" proteins: they do not involve industrial processing, they are not likely to provide overprotection in the intestine, and they simply provide selected amino acids known to be limiting for animal performance.

Since MB-1 is folded and relatively stable in *E.coli*, it is believed that it will be stable in selected rumen microbes. Nevertheless, there are three reasons why the stability of MB-1 in the rumen environment outside the microbes also deserves attention. First, the production of MB-1 within the rumen microbe must be controlled and limited so as to minimize the metabolic load on the microbe. This could be accomplished by tailoring

the MB-1 gene to secrete the protein into the rumen fluid. Although this would allow for increased production of the protein and avoid detrimental effects due to intracellular accumulation, it would entail exposure of the protein to rumen proteases. Second, if MB-1 is maintained within the microbe, there is a chance that the cells will be lysed before reaching the abomasum, therefore exposing the protein to the rumen environment (MacCallum *et al.*, 1997). Finally, the other alternative for farming practices (feeding cattle plants that contain MB-1) would imply exposure to and solubilization in the rumen. Under any of these circumstances, the protein must resist degradation in the rumen in order for the essential amino acids therein to be assimilated by the cow.

Several techniques are used to evaluate the degradability of proteins under rumen conditions. These include *in vivo*, *in situ*, and *in vitro* studies. Measurements *in vivo* are usually performed using animals equipped with cannulae in their abomasum or small intestine. The flow of total protein into the abomasum is monitored and the proportion of dietary protein escaping microbial degradation is determined by subtracting the amount of total protein that is of microbial origin (estimate). Such trials are laborious. They require surgically prepared animals, and the results differ from laboratory to laboratory as they depend on the method used and the flow rate of the digesta into the abomasum. Although the dynamics of the rumen are considered using this technique, the estimates of degradation are subject to a high degree of error, therefore other options were sought to evaluate MB-1 stability (Tamminga, 1979).

The *in situ* technique involves placing non-digestible dacron bags containing feedstuffs inside the rumen of a fistulated animal and subsequently assessing the amount of non-degraded protein left in the bag at specific time intervals (Mehrez and Orskov, 1977; Nocek *et al.*, 1983). This method also exposes the proteins to rumen conditions, but only those found in the rumen fluid. A large portion of the microorganisms in the

rumen are associated with the feed particles themselves, and some are even attached to the rumen wall. These sources of degradability are not accounted for, nor is the effect of the solid material on protein stability (Church, 1988). In addition, the technique requires a large amount of feedstuff, and MB-1 is currently being produced only in milligram quantities. This technique has been shown to have poor repeatability between labs and requires correction for the microbial contamination of feed residues in the bag (Luchini *et al.*, 1996). Finally, the pores of the bag have to be large enough to allow large protease molecules to pass through. MB-1, with its small molecular weight, would simply diffuse through the pores.

The alternative to *in vivo* and *in situ* techniques is to study protein degradation *in vitro*. Techniques used for these in-laboratory studies involve exposure of the proteins to either rumen extracts (fluid or crude protease) or commercially available protease extracts. Incubation of the proteins with freshly collected rumen fluid is still widely used today. Using this method, the rate of ammonia accumulation is determined as an estimation of degradation. In this setting, the ammonia which is reutilized for microbial synthesis is not taken into account. As a result, the technique gives low rates of degradation (Broderick, 1978). To overcome these difficulties, inhibitors of amino acid and ammonia utilization (deaminase inhibitors) have been employed. Yet, the specificity of their action is not well established (Wallace and Kopencny, 1983).

Extracting proteases from the rumen and using them to assess the degradation rate is another *in vitro* option. The method used by Mahadevan *et al.* (1987) involves the extraction of intracellular proteases from the rumen, and the subsequent exposure of these enzymes to the proteins being studied. Developing this protocol for the extraction of extracellular rumen proteases was considered too demanding at the time. Since the immediate goal was to compare the proteolytic stability of MB-1 to other known

proteins, a standard procedure used in feed evaluation laboratories was employed instead. In such studies, the protein is degraded with non-rumen proteases of microbial origin to estimate the rate of feed degradation (Poos-Floyd *et al.*, 1985; Pichard and Van Soest, 1977; Roe *et al.*, 1991; Krishnamoorthy *et al.*, 1983). The use of commercially available proteases for studying pure and soluble proteins like MB-1 is suitable, and their rate of disappearance can easily be monitored using SDS-PAGE (Spencer *et al.*, 1988).

1.5.2 MB-1 Degradability

The enzymes chosen to evaluate MB-1 degradability in the rumen were Pronase E (bacterial protease type XIV from *Streptomyces griseus*) and Neutrase (bacterial protease from *Bacillus subtilis*). Commercially available Pronase E is a broad spectrum protease extract that has both exopeptidase and endopeptidase activity as do proteases found in the rumen (Roe *et al.*, 1991). Neutrase is also a broad spectrum protease extract, but it only has endopeptidic activity. Its use in estimating ruminal degradation has been documented. however, giving a highly significant correlation ($R^2 = 0.92$) with the *in situ* dacron bag method (Assoumani *et al.*, 1990; Roe *et al.*, 1991).

Using these two proteases under rumen-like conditions (pH 6-7, temperature 38-42 °C), degradability studies performed on MB-1 indicated that it would be rapidly and totally degraded if exposed to the rumen environment (MacCallum *et al*, 1997). For comparison, two natural proteins of similar size to MB-1 were also evaluated. These proteins, RNase A and Cytochrome C (Cyt C), are known for their folding stability and their resistance to proteolytic degradation (Privalov, 1979; Imoto *et al.*, 1974). As predicted, they were not degraded as quickly as MB-1. No difference was observed between Pronase E and Neutrase degradation. However, the apparently low resistance of MB-1 to rumen degradation is not unlike that found for a number of proteins common to

ruminant feedstuffs. Pea proteins convicilin and vicilin, for instance, show similar resistance to rumen proteases (Spencer *et al.*, 1988; McNabb *et al.*, 1994). The degradability of MB-1 may compare to other dietary proteins, but the lack of stability would limit expression yields in crops or rumen microbes, and promote digestion in the rumen. These results have prompted studies to determine the cause for such low proteolytic stability.

1.5.3 Factors Affecting Proteolytic Stability

Numerous factors affect the rate at which feed protein is digested in the rumen. Among the most important are those related to the nature of the protein. Such factors include the structural stability of the protein, its solubility, and properties of the solvent in which it is dissolved (Roe *et al.*, 1991; Church, 1988). The structural or conformational stability of a protein has a major impact on its proteolytic stability (Goldberg *et al.*, 1978). Unfolding of a protein due to low conformational stability exposes regions which are targeted by proteases. For natural and synthetic proteins then, folding stability is important for resistance to proteolytic degradation (Goldberg *et al.* 1986; Huang *et al.* 1994; Liao, 1993; Schein, 1989; Yang and Tsou, 1995). Solubility of proteins is also an important factor in ruminal degradation of protein (Shirley, 1986; Church, 1988). Studies with rumen fluid of comparable ionic strength and pH has confirmed that the degradation of proteins by rumen microorganisms is directly related to solubility of the protein. Soluble proteins are more susceptible to proteolytic attack (Church, 1988; Crooker *et al.*, 1978; Wohlt *et al.*, 1973). Among the solvent properties that affect the stability of proteins are temperature, pH, ionic strength and the presence of amphiphilic molecules (Privalov, 1979). Amphiphilic molecules are compounds having both polar and nonpolar regions. The main amphiphilic molecules in rumen fluid that

would affect protein stability are the volatile fatty acids. Like detergents, amphiphilic compounds are capable of destabilizing the fold of a protein and promoting degradation in the rumen (Privalov, 1979; Creighton, 1989).

1.6 CONFORMATIONAL STABILITY

1.6.1 The Unfolding of Proteins

The conformational or structural stability of a protein is defined as the Gibb's free energy difference between the folded and the unfolded conformations of that protein under physiological conditions (Bradshaw, 1990). For many small globular proteins, unfolding has been found to approach a two state unfolding mechanism (Creighton, 1989). Assessing the conformational stability therefore requires determining the equilibrium constant and free energy change of the following reaction:

Folded (F) <----------> Unfolded (U)

Although it is difficult to define the specific process whereby early folding is initiated, the main pattern involves ordering of the backbone into specific structural elements and packing together of sidechains to form the native structure. Which of these happens first requires detailed experimentation to resolve, but whatever the case, the main driving force in protein folding and stability is believed to be the hydrophobic effect (Engelhard and Evans, 1996). Attraction of non-polar groups plays only a small role in this effect; rather, it is the properties of the surrounding water molecules that cause the hydrophobic residues to collapse together. A molecule of water has a significant dipole due to the greater electronegativity of the oxygen relative to the hydrogen. This leads to the formation of strong hydrogen bonding networks between

neighboring molecules. Regions of proteins that cannot integrate and contribute to this network are usually excluded from it. As a result, hydrophobic or non-polar side chains collapse together to hide themselves from surrounding water molecules (Zubay, 1998; Privalov, 1990).

When heat is applied to a protein solution, the network of hydrogen bonds between the water molecules becomes progressively less organized. As a result, the solvent begins to lose its polarity, and the protein gradually unfolds (Pfeil and Privalov, 1976). The greater the stability of the protein, the less polar the solvent must become before unfolding occurs. Unfolding of a protein with addition of heat is a well documented phenomenon. Privalov (1979), for example, reported the change in Gibb's free energy of unfolding with change in temperature for a series of proteins. **Figure 1.8** shows that, in general, as temperature is increased, the Gibbs free energy for unfolding of a protein reaches a maximum and then decreases until the free energy difference of unfolding reaches zero and the protein unfolds. A parameter referred to as the T_m, or melting temperature can be obtained from an unfolding curve. This is the temperature at which 50 percent of the protein population is unfolded. When a protein sample is placed at a temperature near or above its T_m then, at least half of the molecules will be unfolded and more susceptible to proteolytic degradation (Privalov, 1979).

1.6.2 Measuring Conformational Stability

Fluorescence spectroscopy is a powerful technique used to detect the unfolding of proteins and calculate thermodynamic parameters (Jaenicke, 1996; Kwon *et al.*, 1996; Permyakov, 1992; Stellwagen and Wilgus, 1978; Pfeil and Privalov, 1976). There are three naturally occurring amino acid residues, phenylalanine, tyrosine, and tryptophan, that serve as intrinsic fluorophores and contribute to the UV fluorescence of a protein.

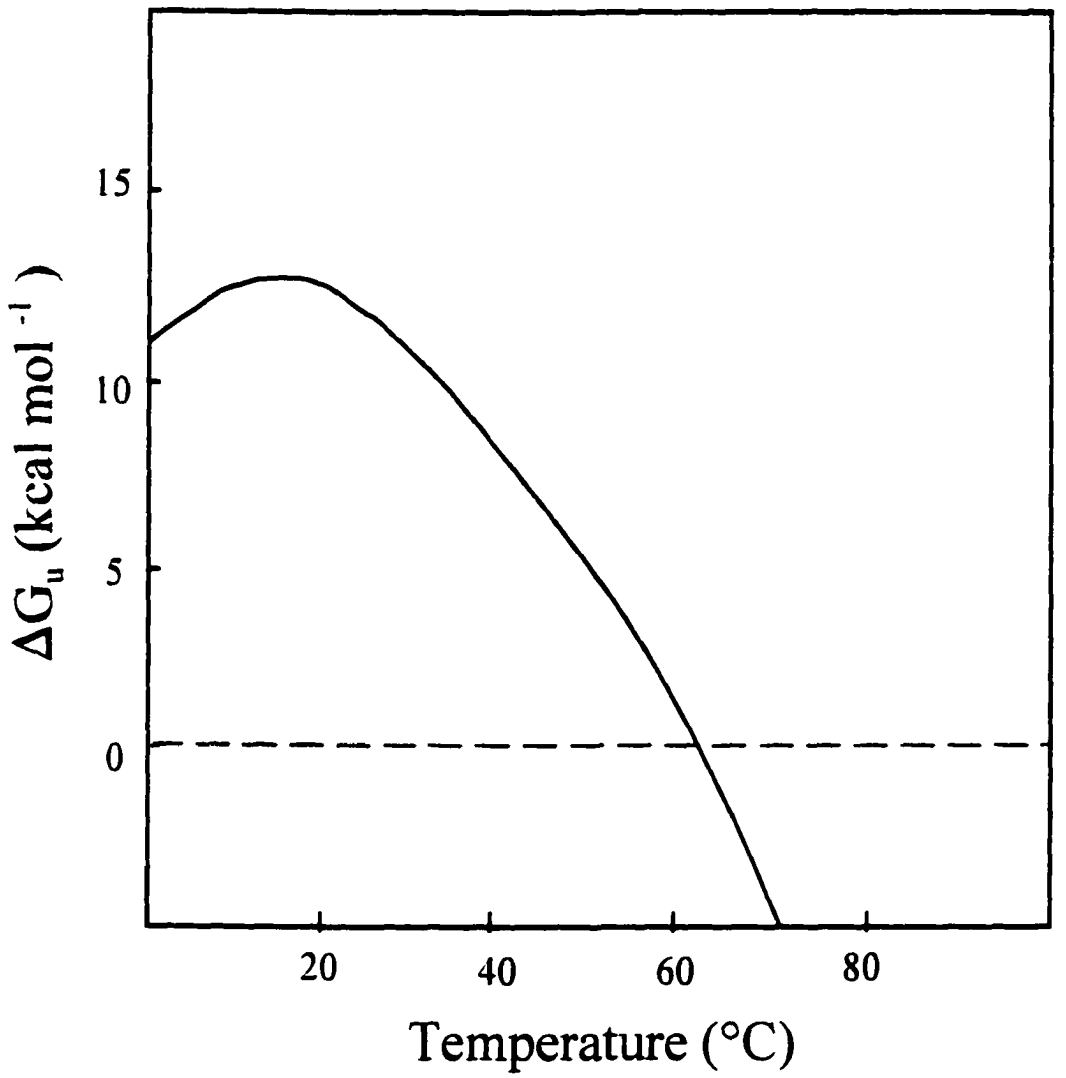


Figure 1.8 General Trend of Free Energy Change (ΔG_u) as a Function of Temperature. The peak of the curve shows the temperature at which protein stability is maximal. Unfolding occurs gradually with temperature, and becomes spontaneous when the free energy is less than zero.

When such a fluorophore absorbs photons, it may do several things. It may produce heat, giving way to a rise in temperature; it may change its chemistry, putting energy into altered bonding structures; or it may become luminescent (Tinoco, Jr. *et al.*, 1995). Luminescence refers to the emission of photons from electronically excited states.

The absorption and emission of light by a fluorophore are best illustrated by the Jablonski diagram **(Figure 1.9)**. Absorption of a photon can excite the fluorophore from the ground singlet state (S_0) to any vibrational level of a higher electronic state (S_1 or S_2). Vibrational relaxation and internal conversion immediately return the photon to the first vibrational level of the S_1 state. Fluorescence occurs only from this level. As the fluorophore returns to the ground state, a photon is given off which can be detected by the photomultiplier in a fluorometer. The emitted light is of lower energy than that absorbed, therefore the emission spectrum of the fluorophore is shifted to a longer wavelength (**Figure 1.10**; Lakowicz, 1983; Tinoco, Jr. *et al.*,1995). This loss in energy is not only a result of vibrational relaxation and internal conversion (IC), but also it results from other competing processes. These processes include intersystem crossing (ISC) of S_1 molecules to the triplet state, further internal conversion (IC) of S_1 molecules to the ground state, quenching (q) of the fluorescence by other amino acids or by components of the solvent, and competing photochemical reactions (D). The following equation defines the quantum yield of fluorescence, which is the ratio of photons emitted during fluorescence to those that get absorbed:

$$\Phi_F = \frac{k_F}{k_F + k_{ISC} + k_{IC} + k_D + k_q [Q]}$$

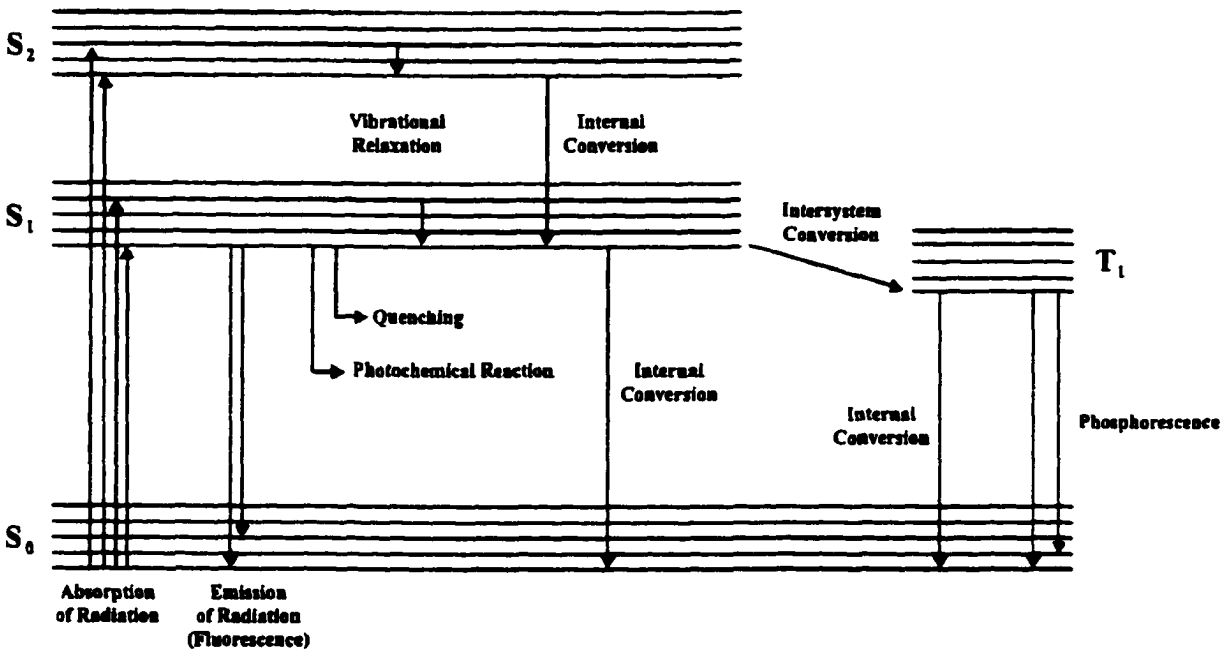


Figure 1.9 A Jablonski Diagram Illustrating Fluorescence. Shown are the ground singlet state -S_0, excited states - S_1 and S_2, and a triplet state - T_1. A fluorophore absorbs light, exciting it to any vibrational level in a higher electronic state (S_1 or S_2). Fluorescence emission occurs only from the first vibrational level in the S_1 state, so vibrational relaxation (and internal conversion from S_2) returns the excited fluorophore to this level. From here, the fluorophore returns back to the ground state, and emits a photon. The observed intensity of this fluorescence depends upon competing processes. Such processes include intersystem crossing to the triplet state, internal conversion to the ground state, quenching, and competing photochemical reactions.

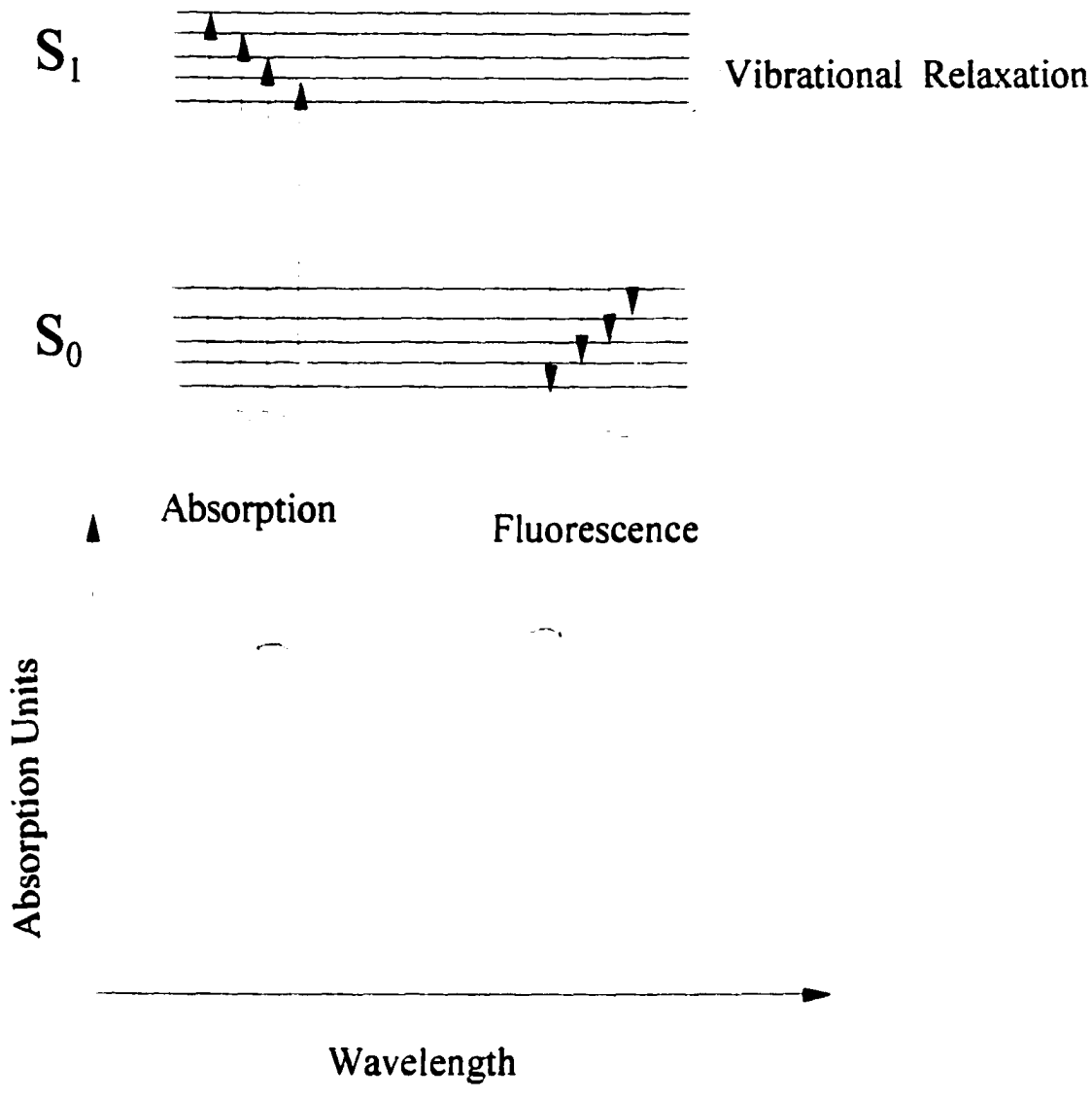


Figure 1.10 Stokes' Shift in Fluorescence. Vibrational relaxation of photons within the higher energy state results in a net loss of energy. Since the emitted light is of lower energy than that absorbed, the emission spectrum of the fluorophore is shifted to a longer wavelength.

As the equation states, the intensity of fluorescence seen from any molecule depends upon the magnitude of the rate constant for photon emission (or radiation; k_F) relative to the sum of the rate constants for the radiative processes (k_F) and the non-radiative processes (k_{ISC}, k_{IC}, k_D, $k_q[Q]$). In other words, the observed intensity increases with the rate of photon emission, but decreases overall as the rates of competing processes increase (Lakowicz, 1983; Tinoco, Jr. *et al.*, 1995).

When a fluorophore is moved between environments, fluorescence quenching can increase or decrease depending on the constituents in the different environments and their concentrations. Such changes in the levels of competing processes (in this case, $k_q[Q]$) are reflected by changes in the fluorescence intensity (Stryer, 1968).

1.6.3 Relationship Between Conformational Stability, Proteolysis Verified

As mentioned earlier, several studies have indicated that unfolding promotes proteolytic degradation (Goldberg *et al.*, 1978; Liao, 1993; Schein, 1989; Yang and Tsou, 1995; Huang *et al.*, 1994). Exploring MB-1's folding stability under rumen conditions was therefore an important step in trying to explain its proteolytic instability. MB-1 was designed to contain a single tyrosine residue (Y) in position 62 of the bundle. This corresponds to position 'd' in the heptad of helix III. In this region, the tyrosine was predicted to be hidden inside of the protein core. Quenching studies indicated that its side chain is indeed buried as per design (MacCallum *et al.*, 1997). Since the denaturation of a protein containing tyrosine will expose it to aqueous solution, the degree of fluorescence quenching is altered and the fluorescence intensity changes. Unfolding of MB-1 and the associated changes in conformation can therefore be inferred based on changes in the spectral emissions of its tyrosine residue (Stryer, 1968).

Preliminary studies on MB-1 were conducted using rumen fluid that had been obtained from fistulated cows. Despite the removal of proteins and microorganisms, several substances remained that would interfere with fluorescence measurements. Chromophores such as chlorophyll and carotenoids were among these substances, absorbing light between 300 and 400nm (Van Soest, 1983; Katz, 1994). To overcome this problem, thermal denaturation studies were carried out in a pigment-free solution with pH, salt, and volatile fatty acid (VFA) content similar to rumen levels. The intent was to find out whether MB-1 would be folded at rumen temperature, and if the presence of amphiphilic VFAs would destabilize the protein. From these studies, MB-1 was found to have a T_m of approximately 39°C (MacCallum *et al.*, 1997). This means that half the population of MB-1 molecules would be unfolded at rumen temperature (38-42°C) and susceptible to rumen proteases (Church, 1988). A parallel study using the same buffer in the absence of VFAs showed no change in the T_m or unfolding curve, indicating that the amphiphilic VFAs were not the destabilizing factors in MB-1's low conformational stability. Clearly, conformational stability is a limiting factor, and is at least in part responsible for the high degradation rate seen under rumen-like conditions.

1.7 THESIS OBJECTIVES

Milk Bundle-1 protein was designed to be used as a feed supplement for ruminant nutrition. Its artificially high content of essential amino acids renders it a high quality source of limiting nutrients for the animal. For it to be used efficiently, however, it must be capable of passing through the rumen virtually unaltered in sequence. Whether expressed in rumen microbes or grown in transgenic crops, MB-1 is not yet stable enough for its intended use. Proteolytic stability studies under rumen-like conditions indicated

that it would be rapidly degraded when exposed to rumen enzymes. Subsequent thermal denaturation studies revealed that the reason for this lack of proteolytic stability is at least in part the fact that it is 50% unfolded at rumen temperature. Logically, the next step is to enhance the rumen stability of MB-1.

The main objective of this work is to introduce mutations into MB-1 that will render it more stable under rumen conditions. To accomplish this, several short term goals were set. They were:

- to determine what stabilizing features would be the best to engineer
- to decide upon specific positioning of the mutations
- to create the mutations in the MB-1 gene using site-directed mutagenesis
- to optimize production and purification of the mutant proteins
- to assess any change in proteolytic or conformational stability relative to that of MB-1

CHAPTER 2 - GENERAL MATERIALS AND METHODS

2.1 MB-1 PRODUCTION AND PURIFICATION

2.1.1 Recombinant Protein Expression

The MB-1 gene was expressed in *E. coli* (TB-1 strain) using the pMALc2 vector shown in **figure 2.1** (New England Biolabs, Mississauga, ON). Overnight cultures of the recombinant plasmid were diluted 1/100 in 1 litre of LB Miller medium (Difco) and selected for resistance to ampicillin (100 μg amp/mL broth). The subculture was grown for about 1.5 h at 37°C with agitation (300 rpm) until the optical density (OD) at 600 nm was 0.4 to 0.5 absorbance units (a.u.). At this point, when bacterial growth was approaching the logarithmic phase, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM in order to induce transcription and translation of the protein. After addition of the IPTG, incubation was continued for approximately 2 h until the A_{600} was between 1.2 and 1.4 a.u.

2.1.2 Cell Harvest and Protein Fractionation

Extraction of recombinant proteins from bacteria must be performed quickly once expression has reached its optimal level. Otherwise, the cellular, membrane-bound and secreted proteases of the bacteria will be given more time to degrade them (Cull and McHenry, 1990). Consequently, the cells were harvested by centrifugation (3200 x g, 1 h) at 4°C so as to slow protease activity while the cells were being pelleted from the growth medium. After centrifugation, the cells were resuspended in Buffer A [10 mM Tris, 200 mM NaCl, 10mM ethylenediaminetetraacetic acid (EDTA) and 1 mM NaN_3, pH 7.4] to which protease inhibitors were added [10 mM EDTA and ethylene glycol-

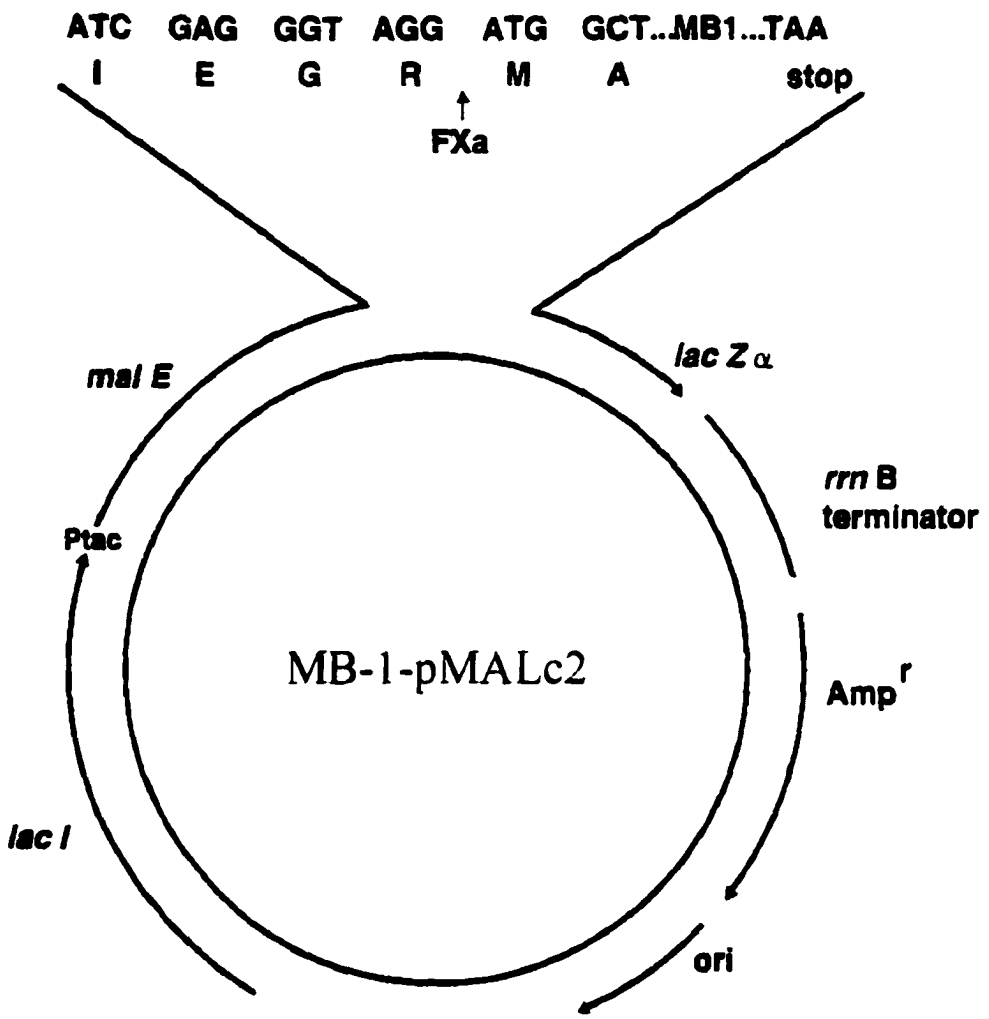


Figure 2.1 Recombinant Expression Construct. In order for it to be expressed in *E. coli*, the gene for MB-1 was inserted into the pMALc2 vector (New England BioLabs). Using this system, the *de novo* protein is produced in fusion with the maltose binding protein (malE or MBP). The MBP is used as a tag for affinity chromatography. A factor Xa cleavage site is located between the two proteins to facilitate separation of MB-1 from MBP.

bis(β-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA), 2 mM benzamide and benzamidine HCl, and 0.1 mM phenylmethylsulfonyl fluoride (PMSF)].

To extract the protein, bacteria were lysed by sonication in an ice/methanol bath, using ten 30 sec pulses with a 30 sec pause for cooling between each. Sodium chloride (NaCl) was then added to the lysate to a final concentration of 700 mM. This allowed for precipitation of cell debris during centrifugation (11,100 x g, 30 min). The resulting bacterial extract contained bacterial proteins as well as a fusion protein of maltose binding protein (MBP) and milk bundle-1 (MBP-MB1). To separate the fusion from these other proteins, the supernatant was loaded onto an amylose column for affinity chromatography.

2.1.3 Amylose affinity column chromatography

MB-1 was expressed as a fusion with the maltose binding protein (MBP) in order to provide an affinity tag for purification. A 20 mL column of amylose resin (New England Biolabs, Mississauga, ON) was first washed with 10 volumes of Buffer A. A solvent delivery system was then attached to provide a constant flow rate of 1 mL/min. After an equilibration period of 20 min, the bacterial extract was loaded onto the column through a sample loop. Over a 2 h period, all other components of the bacterial supernatant were washed through the column while the MBP-MB1 fusion protein was retained. Elution of the fusion protein was accomplished by washing with maltose buffer (Buffer A + 10 mM maltose, pH 7.4). The higher affinity of the MBP for maltose allowed for complete recovery of the fusion protein. Absorbance of the eluted sample at 280 nm was monitored and fractions corresponding to the eluted peak were pooled for cleavage.

2.1.4 Cleavage of Fusion

A factor Xa recognition site is located at the exact point where MBP and MB1 are joined (**see figure 2.1**). Cleavage of the fusion protein was therefore achieved by the addition of this enzyme. The pooled fractions were placed in a dialysis bag (Spectra/Por, Houston, TX; MWCO 3,500 Da) and 50 µL of factor Xa (1 mg/mL, New England Biolabs, Mississauga, ON) was added. Dialysis against 2 L of cleavage buffer (20 mM Tris, 100 mM NaCl, 2 mM $CaCl_2$, pH 8.0) continued at 4°C overnight.

2.1.5 Ion-Exchange Chromatography

After overnight cleavage of the fusion protein, the dialysis bag was equilibrated in 2 L of TE buffer for 2 h. In order to separate MB-1 from the other proteins, a 20 mL column of fast flow diethylaminoethyl sepharose (DEAE pH 8.0; Pharmacia, Baie D'Urfé, Québec) was equilibrated with 10 volumes of TE buffer (Buffer B; 10 mM Tris, 1 mM EDTA, pH 8.0). The diethylaminoethyl, when ionized, has a net positive charge (Rossomando, 1990). When the solution containing MB-1, MBP, and factor Xa was loaded onto the DEAE column and washed with 100 volumes of TE, the two larger, more negatively charged proteins were retained. Elution of MB-1 was immediate and did not require a salt gradient. Fractions (4 mL) were therefore collected immediately after sample addition and assayed for protein content using SDS-PAGE. A schematic diagram of the production procedure is shown in **figure 2.2.**

2.1.6 Electrophoresis (SDS-PAGE)

Denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used in this thesis work for monitoring the purity, quantity, and molecular

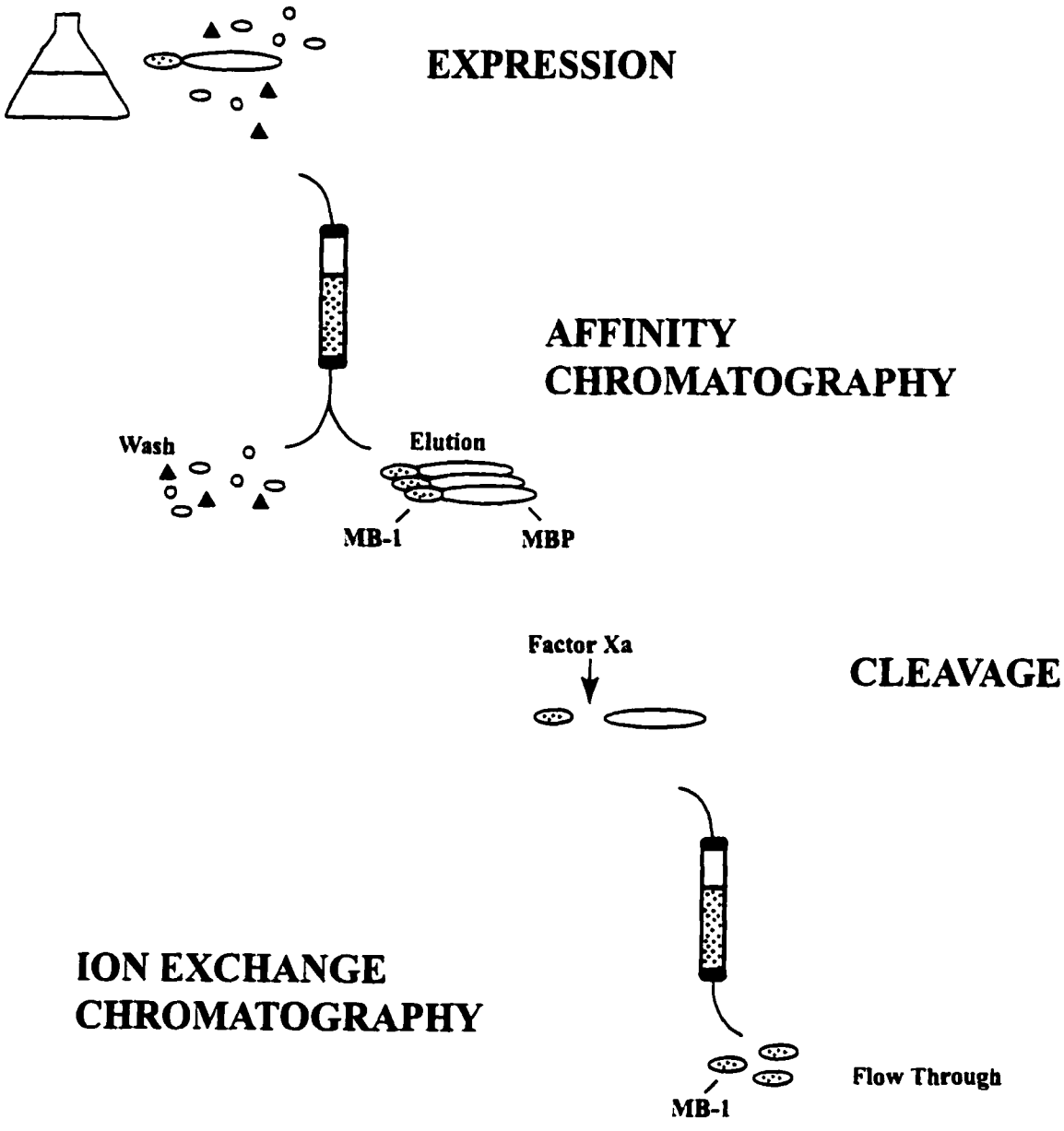


Figure 2.2 General Scheme of MB-1 Production. The MB-1-pMALc2 plasmid is expressed in *E.coli* to yield a fusion of maltose binding protein (MBP) and MB-1. Once harvested, the bacterial cells are lysed and cell debris is pelleted. The supernatant containing bacterial proteins and the fusion protein is loaded onto an amylose column for affinity chromatography. Bacterial proteins are washed through while the fusion remains bound to the amylose. Elution of the fusion is accomplished with maltose buffer. Cleavage with factor Xa and purification on a DEAE sepharose column results in extraction of MB-1 alone.

weight (M_r) of protein in a given sample (Ornstein, 1964; Davis, 1964). The basic procedure involves passing the protein samples through a microscopically porous gel matrix with the application of an electric current. Since MB-1 is a small molecular weight protein (M_r = 11 kDa), a tris-tricine method was employed. This method is capable of resolving proteins in the range of 5 to 20 kDa (Gallagher, 1995; Schägger and von Jagow, 1987). Based on the Laemmli system, the tris-tricine method uses a discontinuous gel and a discontinuous buffer system to separate the proteins. Each polyacrylamide gel was made to have a 12% resolving gel and a 3% stacking gel. Electrophoresis of the proteins was conducted in a Miniprotean Bio-Rad apparatus (Beverly, MA). Tris-tricine buffer (0.1 M Tris, 0.1 M Tricine, 0.1% SDS, pH 8.25) was placed in the upper reservoir of the gel unit to serve as the cathode buffer, while 0.2 M Tris, pH 8.9 was placed in the bottom chamber as the anode buffer. Protein samples were boiled in tricine loading buffer (200 mM Tris-HCl, 40% glycerol, 0.04% Coomassie Blue G-250, 2% SDS) for 3 min, before being loaded into the wells with a pipet or Hamilton syringe. Once all of the samples had been loaded, the gel was run for 1 h at 100 V and 30 mA.

2.1.7 Gel Staining

When protein samples are run on SDS-PAGE, they cannot be visualized until they have been stained. SDS-PAGE by the method of Laemmli followed by Coomassie Brilliant Blue R-250 staining was standard procedure for protein researchers. Detection of proteins with Coomassie Blue requires a large amount of sample, however (0.1-1 μg / band). With MB-1 being produced in only milligram quantities, the use of this procedure was deemed wasteful, and often times, the presence of pure protein or contaminants

would go undetected. Since silver staining is faster and 10-50 fold more sensitive than Coomassie staining (detects 1-10 ng / band), MB-1 gels were stained using a silver staining kit from Bio-Rad (Bio-Rad Bulletin 1089; Bio-Rad 1998/9 Catalogue; Gallagher, 1995; Merril, 1990).

Figure 2.3 shows a typical SDS-PAGE analysis at various stages of MB-1 purification. Samples were taken both before and after cleavage of the fusion protein and checked on gel to ensure that factor Xa activity was optimal. Aliquots from the final purified fractions were also loaded onto the gel in order to determine which ones contained MB-1. Purity of the samples after ion-exchange chromatography was based on the absence of bands having a different molecular weight than MB-1.

2.1.8 Concentration

Those fractions identified as containing pure protein were pooled together and concentrated under nitrogen using an Amicon ultrafiltration unit (Amicon, Oakville, ON). Further concentration was achieved using Centricon 10 concentrating tubes (Amicon, Oakville, ON).

2.2 SAMPLE PREPARATION

2.2.1 Protein Quantitation

Before a protein sample can be used in an experiment, its concentration must be determined. The concentration of protein in a given solution was detected using the Bicinchoninic Acid Assay Kit (BCA; Sigma, Oakville, ON). This offers a colorimetric assay in which the copper (II) sulfate of the standard working reagent (SWR) is reduced by the proteins. The Cu (I) that is formed reacts with bicinchoninic acid in the SWR to

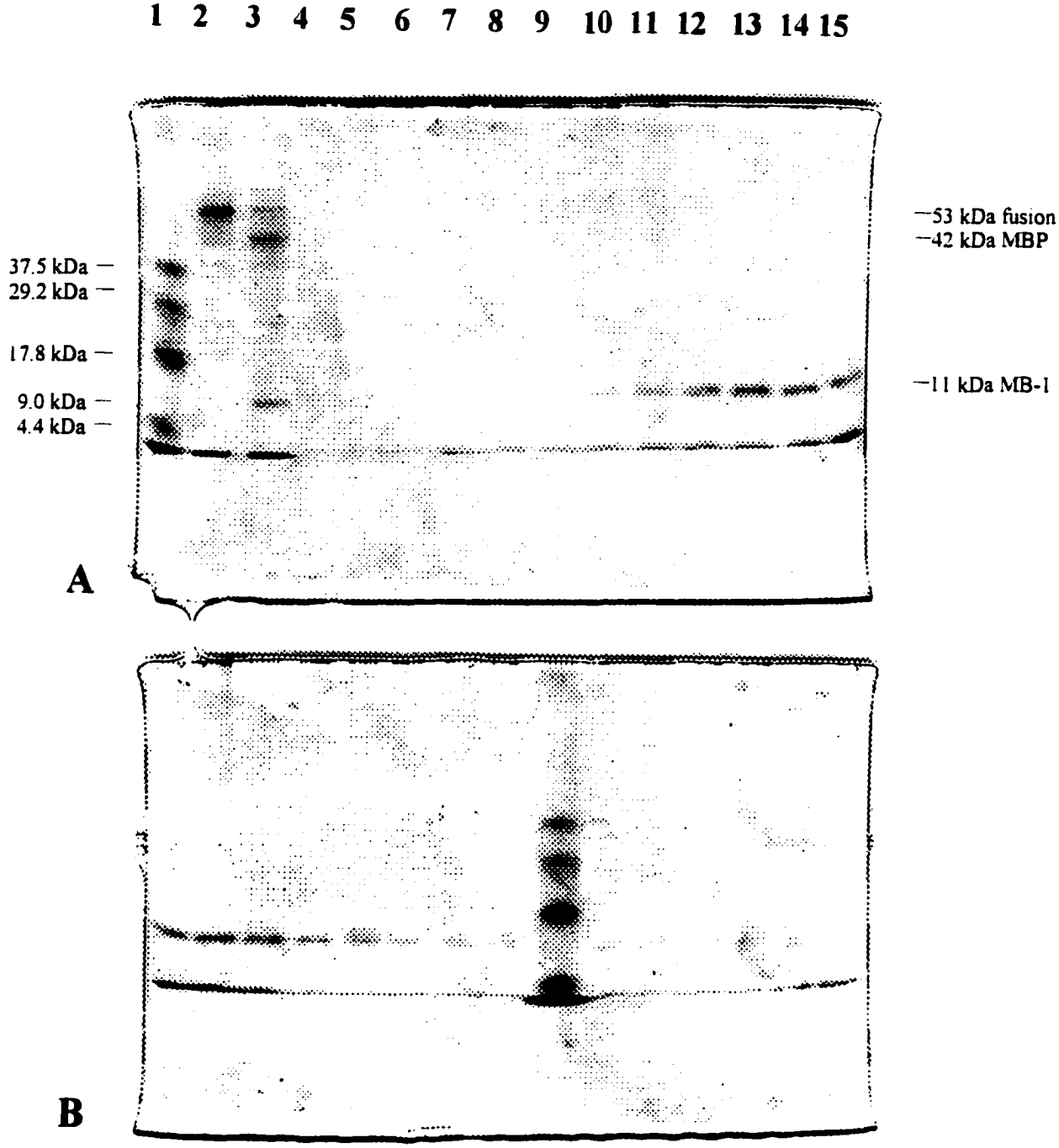


Figure 2.3 SDS-PAGE Analysis During MB-1 Purification. In gel A, Lane 1 contains molecular weight markers, lane 2 shows the fusion protein before cleavage, lane 3 shows the same sample after cleavage, and the rest of the lanes show fractions collected after ion-exchange chromatography. Gel B continues from the first gel, showing the subsequent DEAE fractions and molecular markers in lane 9. MB-1 of >95% purity can be seen in lanes 10-15 on gel A and 1-8 on gel B. Faint bands are also present in lanes 10-14.

yield a purple complex. Since the color change is proportional to the amount of protein in solution, the concentration of an unknown protein sample can be calculated by comparison with standards of known quantity (Bovine Serum Albumin; 1 mg/mL supplied with the BCA kit, Sigma). The absorbance of the sample at 562 nm was therefore measured using a spectrophotometer and compared with the absorbance of the standards.

2.2.2 Assessment of Purity

For all of the experiments performed, it was important to assess the purity of the protein sample. Under certain circumstances, preparations of MB-1 show the presence of a degradation product, having a molecular weight of 8 kDa (**Figure 2.4**). Based on its molecular weight and its ability to co-purify with MB-1, the contaminant is believed to be a truncated unit of MB-1. If the purity of a sample is below 90%, it cannot be used for certain types of experiments as the lower molecular weight contaminants would interfere with the analysis. SDS-PAGE was therefore carried out to ensure that samples were greater than 90 percent pure for most studies. A lower purity was acceptable for the degradation experiments, however, since they involved the use of a densitometer which could be directed to measure only the desired band.

2.2.3 Buffer Exchange

The experiments presented in this thesis required that the proteins be dissolved in solutions other than the final purification buffer (TE buffer, pH 8.0). Two techniques were utilized to transfer the protein between buffers. In the first technique, protein samples were dialysed by placing them in porous bags of 3.5 kDa molecular weight cut

off (MWCO) and immersing them in 1000 volumes of the appropriate buffer. Agitation at 4°C for a minimum of twelve hours with two buffer changes resulted in adequate buffer exchange. In the second technique, buffer exchange was accomplished using size exclusion chromatography. A 5 mL Sephadex G-10 column was equilibrated with 60 mL of desired buffer before loading the protein sample. When the 1-1.5 mL sample was loaded, the components of the original buffer were retained at the top. Because the protein was much larger than these particles, it flowed through quickly and was eluted with the new buffer. One mL fractions were collected, with the protein being eluted in fractions 4 through 8.

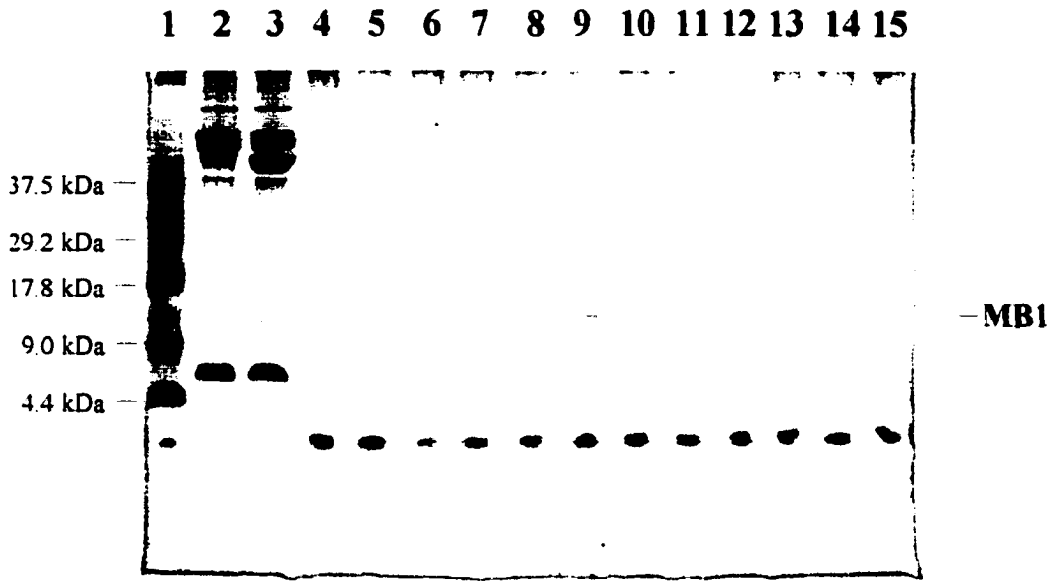


Figure 2.4 SDS-PAGE Showing the 8 kDa Degradation Product. Lane 1 contains molecular weight markers, lane 2 shows a sample before cleavage, lane 3 shows the same sample after cleavage, and the rest show fractions collected from the DEAE Sepharose column. MB-1 is more abundant, migrating at 11 kDa, but the truncated unit is still too concentrated for this sample to be used in certain types of studies.

CHAPTER 3 - MUTANT 1 : MB-1-CYS DIMER

3.1 INTRODUCTION

3.1.1 Design of the Mutant

Previous studies have shown that MB-1 is not stable enough for its desired application in ruminant nutrition. In its present form, MB-1 would be rapidly degraded by rumen proteases (MacCallum *et al.*, 1997). An alternative form with increased rumen stability thus needs to be created. This protein would be capable of passing through the rumen virtually unaltered in sequence, and delivering a high quantity of essential amino acids further down the digestive tract for digestion and absorption.

There are two ways in which the stability of MB-1 could be enhanced. Since it is 50% unfolded at rumen temperature, the protein could be rendered more resistant to degradation by increasing its thermal stability, causing it to be more stable and folded at rumen temperature. Many researchers have shown that a greater folding stability results in greater proteolytic resistance (Goldberg *et al.*, 1986; Huang *et al.*, 1994; Liao, 1993; Schein, 1989; Yang and Tsou, 1995). For MB-1, this could be achieved either by engineering fold specifiers into its tertiary fold, or by enhancing the packing and hydrophobicity of the bundle core.

Another way in which the rumen stability of MB-1 could be increased is by decreasing the amount of surface area that is available to rumen enzymes. This can be accomplished by linking monomeric units together, either through a covalent bond or through expression as a polymer. With the units so closely apposed, one would mask a certain portion of protease targets on the other and thereby confer greater proteolytic resistance to the linked protein.

It was this latter approach that was taken to design the first MB-1 mutant. In order to reduce the surface area exposed by each MB-1 monomer, a mutant capable of disulfide linkage was engineered. As seen in **figure 3.1**, however, a disulfide bond can only form when the sulfhydryl groups of two cysteine residues are near one another and exposed to oxidizing conditions (Friedman, 1973; Creighton, 1989). Since MB-1 has no cysteine residues, the first step in the creation of the disulfide mutant was to determine the best position for the cysteine residue. With its free sulfhydryl group, this cysteine mutation would not only serve as a model for studying the effect of oligomerization on stability, but also it would provide a means of binding heavy metal ions to the protein to facilitate analysis by x-ray diffraction (Friedman, 1973; Rees *et al.*, 1992).

3.1.2 Protein Engineering

For the sulfhydryl group of the cysteine residue to be capable of forming a disulfide bond with a neighboring MB-1 unit, it had to be engineered in a position that would be exposed to the aqueous environment. Referring to the top down view of the bundle **(figure 1.5, p. 15)**, it is easy to see that positions 'b', 'c' and 'f' of the heptad repeat are located on the hydrophilic, exposed surface of the four helices. Any of these three positions were deemed suitable for the cysteine mutation. When selecting the appropriate amino acid to be replaced, care had to be taken not only to avoid altering the essential amino acid content, but also to ensure that the cysteine would offer similar properties to the position it would be filling. It is a well-established fact that the micro-environment of the heptad positions allow for incorporation of only certain amino acids (Paliakasis and Kokkinidis, 1992; Beauregard *et al.*, 1995). Otherwise, the structure and stability of the bundle could be distorted in order to accommodate the "foreign residue"

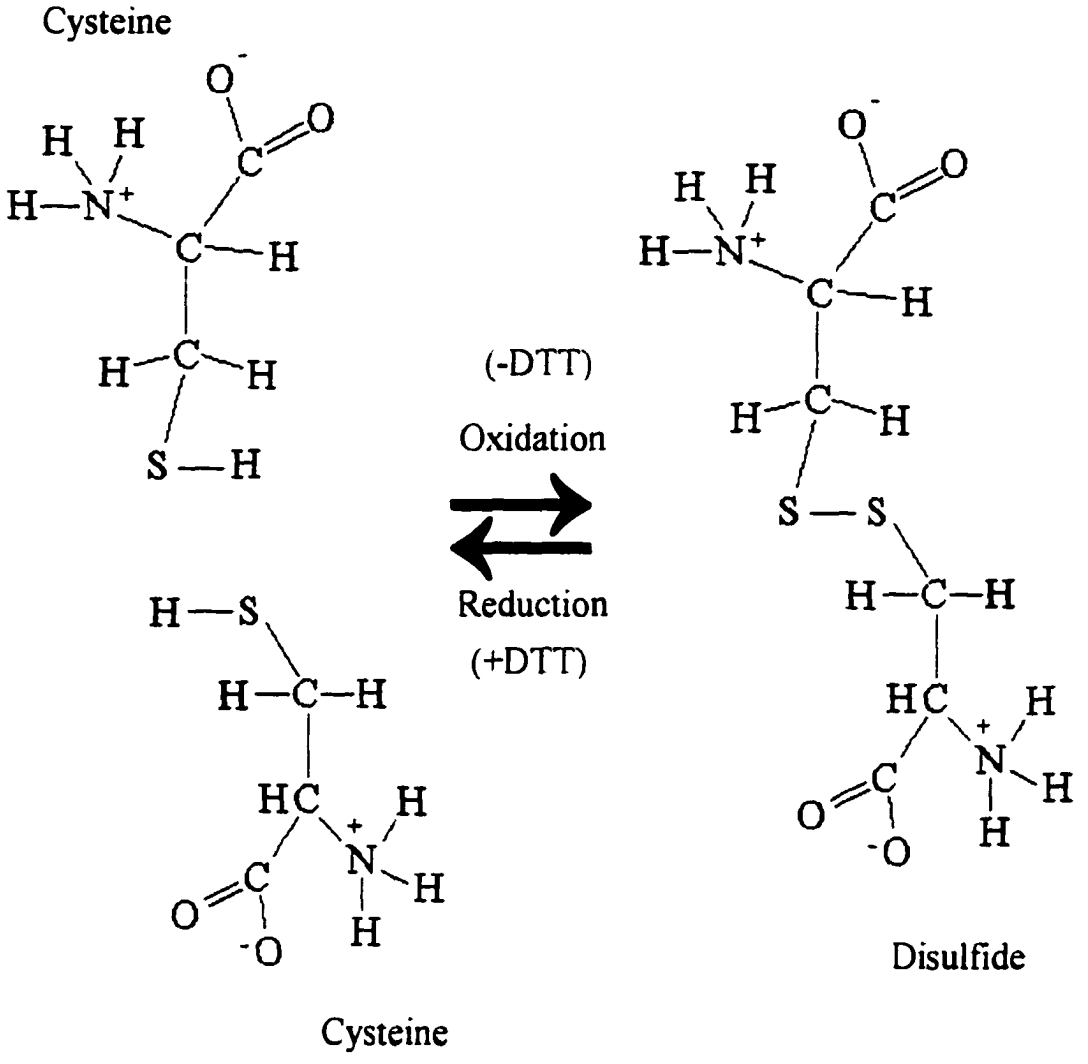


Figure 3.1 Formation of a Disulfide Bond. A disulfide bond forms when the sulfhydryl groups of two closely apposed cysteine residues undergo oxidation. The bond formed is a strong covalent bond which can only be broken by reducing agents like dithiothreitol (DTT) or β-mercaptoethanol.

(Bordo and Argos, 1991). As a result, an asparagine residue (N) in position 93 of the bundle was chosen for replacement. Being in position 'c', it was predicted to be exposed to the environment and its properties are similar to those of cysteine, i.e. polar and uncharged. As shown in **figure 3.2**, the cysteine would be located near the end of the fourth helix and capable of forming an intermolecular disulfide with a neighboring monomer.

3.1.3 Genetic Engineering

Once the position of the cysteine mutation was determined, the next step was to decide on the best and easiest mode of mutating the MB-1-pMALc2 plasmid. TGC was the codon chosen for incorporation into the mutant plasmid, for its preferred expression in *E.coli*. By looking at the restriction sites on the plasmid, it was found that the restriction enzymes *Pst*I and *Hin*dIII each cut MB-1 at a unique site in the region where the mutation would be introduced **(Figure 3.3)**. It was decided that replacement of this region with a mutated segment (i.e. "cassette") would be the simplest and quickest method of mutating the gene. Cassette mutagenesis with directional sub-cloning was therefore the mode chosen to replace the native MB-1 sequence with the mutation. Oligonucleotides Cys1 and Cys2 were designed such that, when annealed, they would form a dsDNA cassette that contained the TGC codon instead of the AAC codon and corresponded to the region between *Pst*I and *Hin*dIII sites.

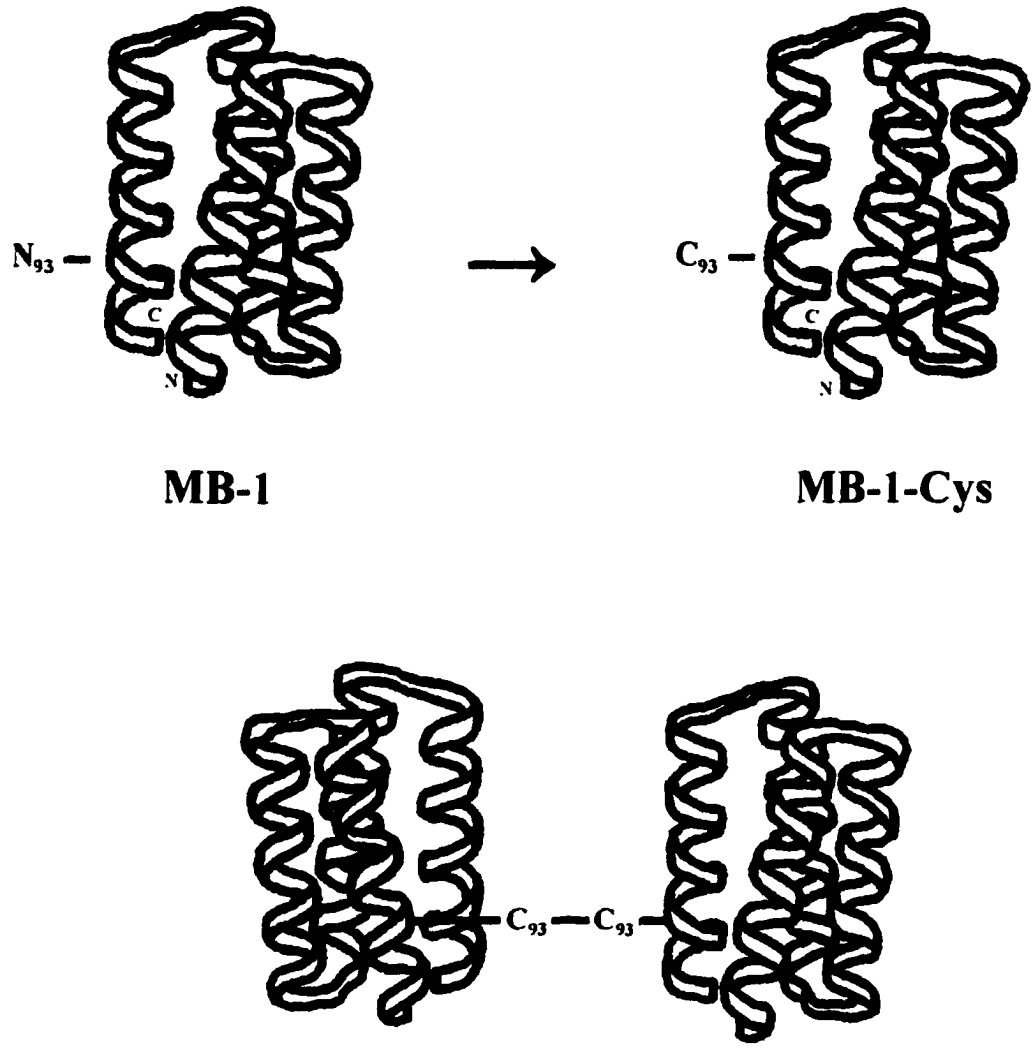


Figure 3.2 Creation of Mutant 1: MB-1-Cys Dimer. An asparagine (N) in position 93 of MB-1 will be replaced by a cysteine (C) to form the MB-1-Cys monomer. In this position, the sulfhydryl group of the cysteine residue is predicted to be exposed and available to form a disulfide bond with a neighboring MB-1-Cys monomer. Each unit in MB-1-Cys Dimer would protect a certain portion of the other unit, thereby decreasing the surface area exposed to proteases.

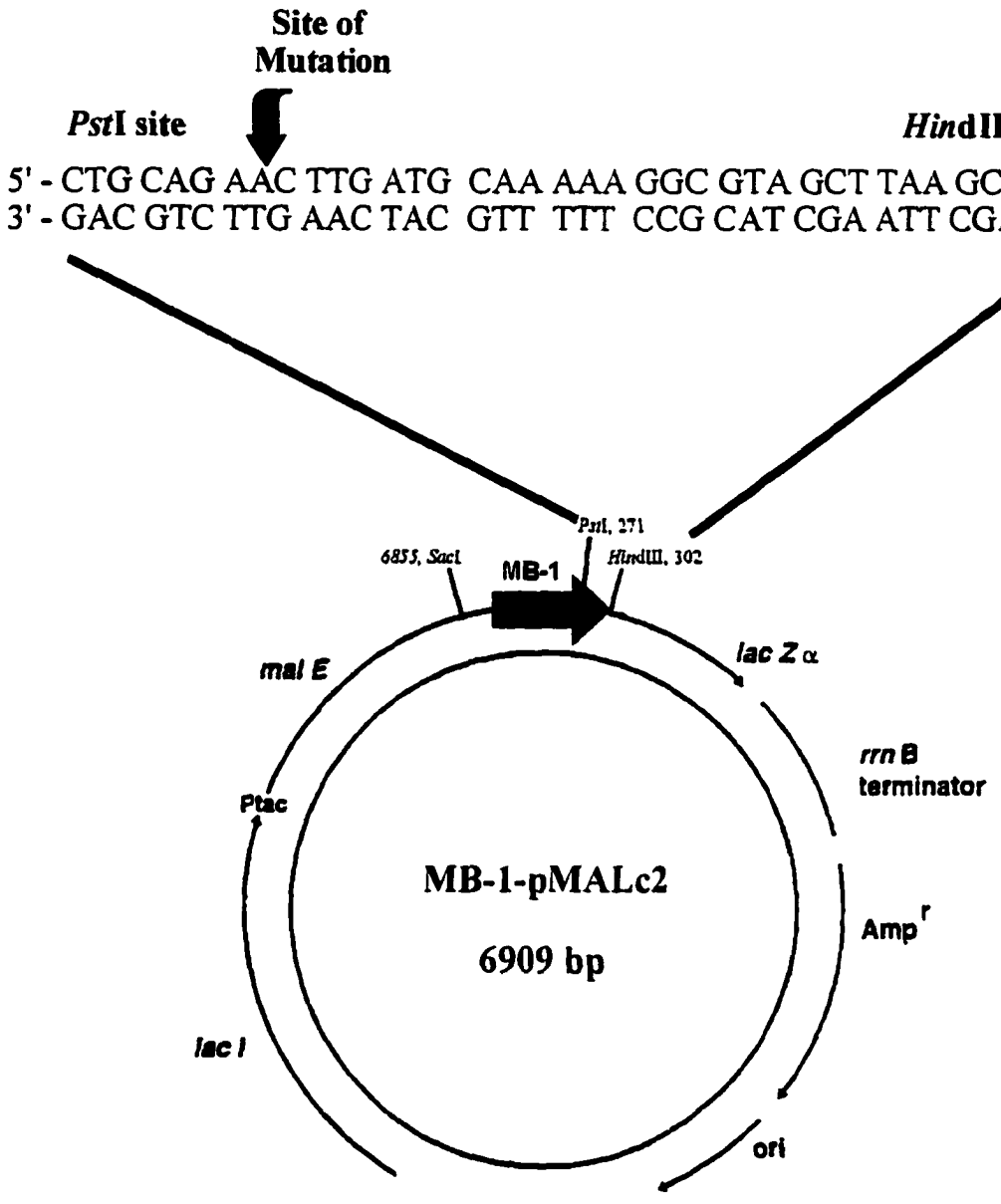


Figure 3.3 Restriction Sites in the Region to be Mutated. *Pst*I cuts the MB-1-pMALc2 plasmid just upstream of the mutation site, while *Hind*III cuts downstream. Using these two sites, the wild-type "cassette" containing the AAC of asparagine will be removed from the plasmid and replaced by a cassette containing the TGC of cysteine.

3.1.4 Chapter Objectives

The main objectives for this chapter were:

- ❑ To introduce the mutation into the MB-1-pMALc2 plasmid
- ❑ To produce and purify the mutant protein, optimizing the original protocol
- ❑ To confirm the presence of the single cysteine residue
- ❑ To determine whether or not the disulfide bond was formed
- ❑ To compare the proteolytic stability of MB-1-Cys Dimer to that of MB-1
- ❑ To compare the conformational stability of MB-1-Cys Dimer to MB-1

3.2 MATERIALS AND METHODS

3.2.1 Cassette Mutagenesis

As mentioned, cassette mutagenesis was used in order to introduce the cysteine mutation onto the MB-1-pMALc2 plasmid. A schematic diagram of the process is shown in **figure 3.4**. Synthetic oligonucleotides Cys1 and Cys2 were obtained as a gift from Dr R.M. Teather, Agriculture Canada, Ottawa. The oligos were purified by first running the samples on a 15% polyacrylamide, 44% urea gel using Tris Borate EDTA (TBE) running buffer (30 min prerun, 1 h migration @ 100 V). DNA bands were visualized by ethidium bromide (EtBr) staining, and then cut from the gel. After the excised bands were placed in labelled test tubes and crushed with a pipet tip, the DNA was eluted from the polyacrylamide by incubation in 400 μL water overnight at 37°C. Most of the polyacrylamide remnants were pelleted by centrifugation, and the resulting DNA-containing supernatants were desalted on sephadex G-25 spun columns (350 x g, 2 min) to remove any residual gel or EtBr particles. Once purified, the oligos were quantitated by measuring the OD at 260 nm in Tris EDTA (TE) buffer pH 8.0. The sample

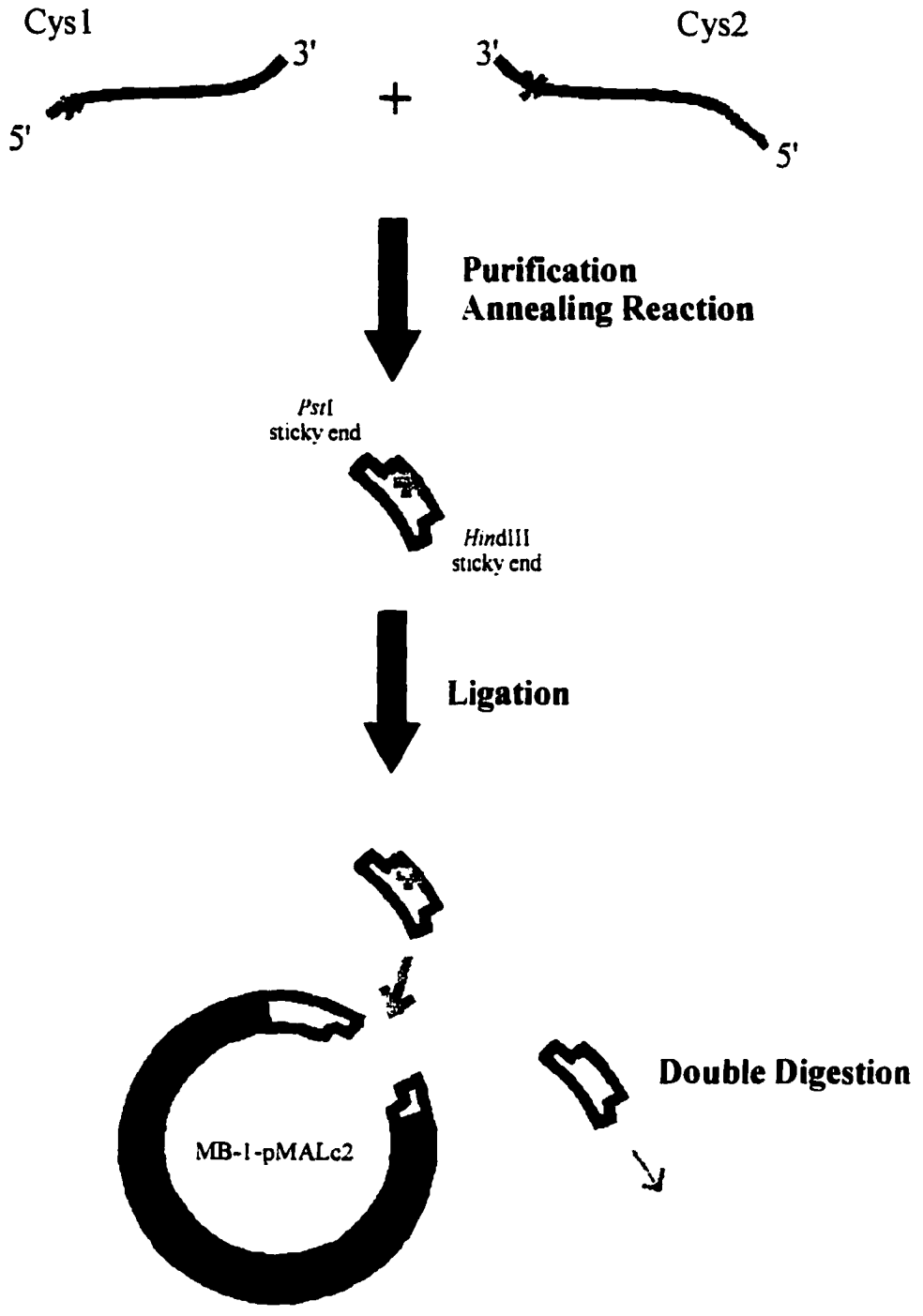


Figure 3.4 Schematic of Cassette Mutagenesis. The Cys1 and Cys2 oligonucleotides were purified and annealed, yielding the mutant cassette with *Pst*I and *Hin*dIII sticky ends. Digestion of MB-1-pMALc2 with these two enzymes resulted in removal of the wild-type cassette and linearization of the plasmid. Ligation of the mutant cassette to the vector was achieved by incubation with T4 DNA ligase.

concentrations were calculated using the conversion factor 1.0 A_{260} unit = 33 μg/mL ssDNA (New England BioLabs Catalogue 96/97).

Cys1 and Cys2 oligos were then annealed in a 1:1 ratio in Tris NaCl EDTA buffer (TE+50 mM NaCl) by incubating them at 100°C for 2 min and then turning the block heater off to let them cool slowly to room temperature. At this point, the oligos formed a cassette housing the cysteine-coding mutation, flanking MB-1 sequences, and sticky ends that corresponded to *Pst*I and *Hin*dIII cleavage sites on the plasmid. MB-1-pMALc2 was then simultaneously digested with 2U *Pst*I and 3U *Hin*dIII per μg DNA in NEBuffer 2 (New England BioLabs, Mississauga, Ontario). Samples were incubated at 37°C for 2 h to optimize enzymatic activity. The digested DNA was mixed with an equal volume of 2X loading buffer (10% glucose, 1% bromophenol blue) and loaded into the single well of a 0.9% low melting point agarose gel for purification [Tris Acetate EDTA (TAE) running buffer; 3 h @ 30 V]. The band corresponding to the linear plasmid DNA was cut from the gel and purified by phenol:chloroform extraction and precipitation with ammonium acetate and 100% ethanol. By combining the linearized plasmid and the annealed mutant cassette in a 1:2 ratio (by concentration) and incubating them overnight at 16°C with 10U T4 DNA ligase, the mutant cassette was ligated into the plasmid. After transformation of the ligation mixture into competent *E. coli* B834 cells, the bacteria were plated with ampicillin selection. Plasmid DNA was extracted from the resulting colonies according to Sambrook *et al.* (1989) and prepared for dideoxy sequencing.

3.2.2 DNA Sequencing

Dideoxy-sequencing, first described by Sanger *et al.* (1977), employs chain-terminating dideoxynucleotides to produce 5' co-terminal fragments of varying lengths.

A primer annealed to the DNA is elongated by DNA polymerase in the presence of deoxynucleotides (dNTPs). [α^{-35}S]-labelled dATP is also incoporated into the growing DNA chain to facilitate visualization of the fragments during subsequent autoradiography. These labelling reactions are then divided into four tubes, each containing one type of dideoxynucleotide (ddATP, ddCTP, ddGTP, or ddTTP). Because these dideoxynucleotides lack a 3'-OH group, their incorporation results in termination of the DNA chain. However, since the proportions of ddNTPs to dNTPs have been optimized, all possible fragments with the same 5' end are obtained during the reaction. For instance, the tube that contains ddATP will contain each possible fragment starting from the primer and ending at a 'ddATP'. When these different sized oligonucleotides are separated from the DNA template by denaturation, they can be fractionated by electrophoresis on the gel. The same holds true for each ddNTP reaction. Therefore, when the contents of the 'A', 'C', 'G', 'T' tubes are loaded side by side and detected by autoradiography, the DNA sequence can be read directly, starting with the fastest-moving or smallest bands at the bottom of the gel and moving up (Sambrook *et al.*, 1989; Zubay, 1998).

In order to confirm whether or not the mutation had been incorporated, dideoxy sequencing was performed on the plasmid DNA using a *Bst* Sequencing Kit (Bio-Rad, Mississauga, Ontario). This kit was employed since *Bst* polymerase (Polymerase I, large fragment from *Bacillus stearothermophilus*) was said to exhibit several properties superior to the other polymerases used in DNA sequencing. For instance, with an optimal reaction temperature of 65°C, the *Bst* DNA polymerase can be used at higher temperatures than most polymerases, thereby reducing secondary structure in the DNA template and allowing for more precise sequencing ladders. Also, with its high synthesis

rate (120 nucleotides/second), only a small amount of DNA is required (1.0 μg of double-stranded plasmid DNA is usually sufficient). The even band intensity, lessened artifacts, and low background that can be accomplished using this kit made it the one of choice at the time (Bio-Rad Bulletin 1474; *Bst* DNA Sequencing Kit Instruction Manual).

DNA samples (approx. 2 μg) were prepared for sequencing by alkaline denaturation (2 M NaOH, 2 mM EDTA pH 8.0) and precipitation (2 M sodium acetate pH 4.6, 3 volumes 100% ethanol). After washing each of the DNA pellets with 70% ethanol, they were dried in a vacuum concentrator, and dissolved in 7 μL of TE buffer pH 8.0 or sterile water. In reaction buffer (100 mM Tris pH 8.5, 100 mM $MgCl_2$), with a 5 min incubation at 75°C and subsequent cooling, the M13/pUC (-47) (24mer) sequencing primer was then annealed to the denatured DNA template just downstream of the *Hin*dIII site. Approximately 15 μCi of [α^{-35}S] dATP and 1U *Bst* Polymerase were then added to the annealed template and mixed thoroughly. A 2.5 μL aliquot of this final mix was added to each of the four prewarmed (65°C) reaction tubes labelled A,C,G, and T and containing 2 μL of dNTPs and corresponding ddNTP. All four tubes were then incubated at 65°C for 2 min to allow for primer extension, radiolabelling, and termination of DNA synthesis. The reactions were then chased with a mixture of dNTPs (10 mM each) and incubated for another 2 min. This chase reaction completed the synthesis of any unreacted primer. Addition of stop solution (95% deionized formamide, 10 mM EDTA, 0.05% xylene cyanole FF, 0.05% bromophenol blue) terminated the final reactions. Care was taken to spin the samples down when required in between steps (microcentrifuge @ 14,000 x g). Reaction tubes were stored at -20°C until the gel was ready to load. While the gel was being prerun (30 min, 55 W), the reaction samples were thawed and denatured for 2 min at 75°C. Five μL of each reaction were loaded in A, C,

G, T order onto a 10% polyacrylamide, 42% urea sequencing gel (Hoefer SQ3 Sequencer, Hoefer Scientific, San Francisco, CA). The samples were resolved at 55 W for 90 min in TBE buffer. Following the run, the gel was fixed in a 10% methanol / 10% acetic acid solution for 1 h. The gel was then briefly washed with water, and transferred to 3 Whatman chromatography paper and dried. A temperature of 80°C under vacuum was used in the drying unit (Dry Gel Sr. Model SE1160, Hoefer Scientific, San Francisco, CA) for 3 h. The gel was then exposed to x-ray film (Hyperfilm MP, Amersham) in an autoradiography cassette for 48 h at room temperature. The film was developed in an automated developer at the Atlantic Veterinary College.

3.2.3 Mutant Protein Production

The next step after confirming the sequence of the mutant plasmid was to produce the mutant protein. Initially, MB-1-Cys was purified using the protocol for original MB-1 (described in **section 2.1, p. 35**). During these purification trials, it was noted that several contaminating bands were present, and it was proposed that these bands were due to the formation of intermolecular disulfide bonds between MB-1-Cys and other sulfhydryl-containing proteins. Upon inclusion of 2 mM dithiothreitol (DTT) in all solutions pertaining to purification (except for cleavage buffer), the contaminating bands disappeared.

3.2.4 *p*-Hydroxymercuribenzoate Assay

Once pure MB-1-Cys monomer was obtained, the presence of the cysteine residue had to be confirmed. Spectrophotometric measurement of mercaptide formation by *p*-hydroxymercuribenzoate and available cysteine groups on MB-1-Cys was therefore

performed using a modified version of the procedure described by Boyer (1954). One half of the 1-2 mL of purified MB-1-Cys was dialyzed against 0.33 M sodium acetate, 2 mM DTT buffer, pH 4.6. The other half was allowed to dimerize as described in **section 3.2.5**. This latter step was necessary as the DTT present in the purified monomer sample interfered with BCA quantitation of the protein (described in **section 2.2, p. 39**). The amount of MB-1-Cys in the stock solution was subsequently determined by back-calculating from the concentration of the dimer.

A sample containing approximately 0.66 mg/mL of MB-1-Cys in 0.33 M sodium acetate, 2 mM DTT buffer (pH 4.6) was prepared and its absorbance at 255 nm was measured. An aliquot (18 μL) of *p*-HMB was then added to a final concentration of 60 μM, mixed quickly, and measured immediately for change in A_{255}. Blanks corresponding to the absorbance of the buffer and the absorbance of the buffer and protein were subtracted from the absorbances observed to yield the change in absorbance of the *p*-HMB upon binding.

The amount of protein-*p*-HMB complex formed was calculated using the Lambert-Beer Law, $A=\epsilon \times c \times \ell$, which relates the absorbance observed (A) to the molar extinction coefficient ($\epsilon_M = 6.2 \times 10^3\ M^{-1}cm^{-1}$) and molar concentration (c) of the complex, and to the pathlength (ℓ) of the cuvette (Creighton, 1989). All assays were conducted at room temperature in a Hellma Quartz cuvette (ℓ=1 cm) using a Hewlett-Packard 8453 UV-Vis Diode Array spectrophotometer.

3.2.5 Disulfide Bond Formation

Given that the cysteine residue was present in MB-1-Cys, attempts were made to introduce the disulfide bridge between monomer units (**Figure 3.1, p.46**). Oxidation

trials using glutathione, air, and oxygen sources were carried out only to find that complete conversion to disulfide product could be accomplished by simply removing the DTT from the final purification buffer (TE + 2 mM DTT, pH 8.0). Dialysis against Tris buffer, pH 8.2 containing 0.4 μM $CuCl_2$, facilitated the oxidation process to convert MB-1-Cys monomer to MB-1-Cys Dimer. Four buffer changes over an 18 h period at room temperature resulted in rapid and complete oxidation. The conversion of MB-1-Cys to MB-1-Cys Dimer was monitored using SDS-PAGE.

The 2% SDS in the loading buffer used for electrophoresis renders the proteins denatured and of the same negative charge. As a result, when an electrical current is applied, their migration towards the anode depends not on charge or shape, but rather on their molecular weight. Smaller proteins pass through the matrix more quickly than larger proteins, thus they migrate further down on the gel. It follows that the relative mobility of a protein on the gel (R_f) is actually inversely proportional to the logarithm (log) of its molecular weight. The molecular weights of MB-1-Cys and MB-1-Cys Dimer were therefore calculated by comparing the sample's migration to that of known molecular weight standards (Hames, 1981; Garfin, 1990).

3.2.6 Proteolytic Stability Studies

With MB-1-Cys Dimer formed, the next step was to determine the impact of dimerization on proteolytic stability. When MB-1 was first tested, it was exposed to two types of rumen-like protease extracts, Neutrase (bacterial protease from *Bacillus subtilis*) and Pronase E (protease type XIV from *Streptomyces griseus*). These enzymes gave similar results, therefore only Pronase E was used in the proteolytic stability studies.

The experimental set-up used rumen-like pH (6.8) and temperature (39°C) to approximate the rumen environment. The 1-2 mL samples (containing 3-5 mg) of MB-1, and MB-1-Cys Dimer obtained after production, purification, concentration and estimation of degradation (described in **sections 2.1 and 2.2**) were exchanged into the appropriate buffer by dialysis (described in **section 2.2.3**). Assays were conducted in borate phosphate buffer (BP buffer: 7.6 g/L of NaH_2PO_4 H_2O + 13.17 g/L of $Na_2B_4O_7$ $10H_2O$, pH 6.8). Protein stock solutions were equilibrated at room temperature overnight in order to ensure that the proteins were fully folded prior to the experiment. The concentration of the stocks was then determined as described in **section 2.2, p. 39**, and for each, a sample was prepared such that the final concentration after enzyme addition was 0.2 mg/mL. The samples were incubated at rumen temperature (39°C) for 5 min before addition of the enzyme (6.6 U per g sample). A 10 μL aliquot was taken from each reaction tube immediately after enzyme addition. This "Time Zero (T_0)" sample was placed in a tube containing 20 μL tricine loading buffer and boiled for 1 min to stop the reaction. Samples were taken at subsequent time intervals and treated in the same manner before being placed at -20°C for storage. By staggering the sampling times, the two proteins could be assessed under approximately the same conditions. An initial trial was run to determine whether or not the temperature resulted in added degradation of the sample. Since no change was observed when the samples were incubated in the absence of Pronase E, it was apparent that the temperature used for the experiments (39°C) did not contribute to additional degradation.

When the assay was complete, samples were thawed, denatured in a boiling water bath, and then loaded onto a 12% SDS-polyacrylamide gel. The gel was run and stained as described in **sections 2.1.6 and 2.1.7**. Care was taken not to stain two gels in the same

container since this would result in uneven staining. After the gels had been silver stained, they were subjected to laser densitometry (LKB 2222-020 UltroScan LX, Pharmacia). The densitometer measures the absorbency through a gel, reporting a peak wherever a silver-stained protein band is detected. The size of the peak is proportional to the intensity of the band. Thus, by using a densitometer, it was possible to detect the slight changes in band intensity observed during a degradation study. A typical scan is shown in **Appendix A**. By setting the area under the curve value for the T_0 samples equivalent to 100% protein, the amount of protein remaining at subsequent time intervals was determined.

3.2.7 Thermal Stability Studies

In order to determine whether the cysteine insertion and disulfide bond formation had conferred greater folding stability to MB-1, both the disulfide mutant and MB-1 were exposed to thermal denaturation studies. Despite the removal of proteins and microorganisms, rumen fluid from a fistulated cow could not be used in previous trials for MB-1, since contaminating chromophores still remained and interfered with fluorescence measurements (Van Soest, 1983; Katz, 1994). Rather, a buffered volatile fatty acid (VFA) mixture was employed to approximate the ionic strength, pH and amphiphilic content of the rumen liquor (Stewart, 1975; Hungate, 1966). The inclusion of these properties in stability studies is important since they are believed to have the greatest impact on protein folding in the rumen. The same VFA buffer was therefore used to assess the folding stability of MB-1-Cys Dimer.

VFA buffer was made to contain 50 mM sodium acetate, 20 mM sodium propionate, 20 mM sodium butyrate, and 17 mM potassium phosphate at a pH of 6.5.

Addition of sodium azide (NaN_3, 1 mM) prevented growth of airborne microorganisms. Samples of MB-1, and MB-1-Cys Dimer (1-2 mL) obtained after production, purification, concentration and estimation of degradation (described in **sections 2.1 and 2.2**) were exchanged into this buffer as described in **section 2.2.3, p. 41**. Once in the appropriate buffer, the quantity of protein in the stock solutions had to be determined as described in **section 2.2.1**. For each protein, a sample was prepared such that it would have a final concentration of 0.4 mg/mL. The samples were then allowed to equilibrate for 12 h at 15°C before starting the denaturation study. Denaturation experiments were conducted by monitoring the change in tyrosine fluorescence as the samples were heated from 15°C to 65°C. An equilibration period was allotted between each temperature to ensure that the corresponding protein conformation had been reached. Any change in protein fold was reflected by the spectral emission of the lone tyrosine residue.

To monitor the fluorescence of tyrosine, the protein samples were placed in the sample chamber of a fluorescence spectrophotometer (PTI Model RF-M2004; K.C. Irving Chemistry Centre, UPEI). Sample temperature was initially set to 15°C with a circulating bath (Lauda R6S). Then computer software was used to set the motorized excitation and emission monochromators to the desired wavelengths. Light entered the sample chamber from a 75 watt xenon lamp as seen in **figure 3.5** and was absorbed by the tyrosine at 284 nm, exciting the electrons to a higher energy level. As the excited fluorophore relaxed back down to the ground singlet state, photons were emitted in all directions. The photons emitted at a 90° angle were detected by a photomultiplier tube (PMT) while the monochromator scanned from 297 to 350 nm. A signal fed from the tube to the attached computer was recorded as a graphical representation of the photons emitted per second at each wavelength. **Figure 3.6** shows a typical emission spectrum

PTI Model RF-M2004
Fluorescence Spectrophotometer

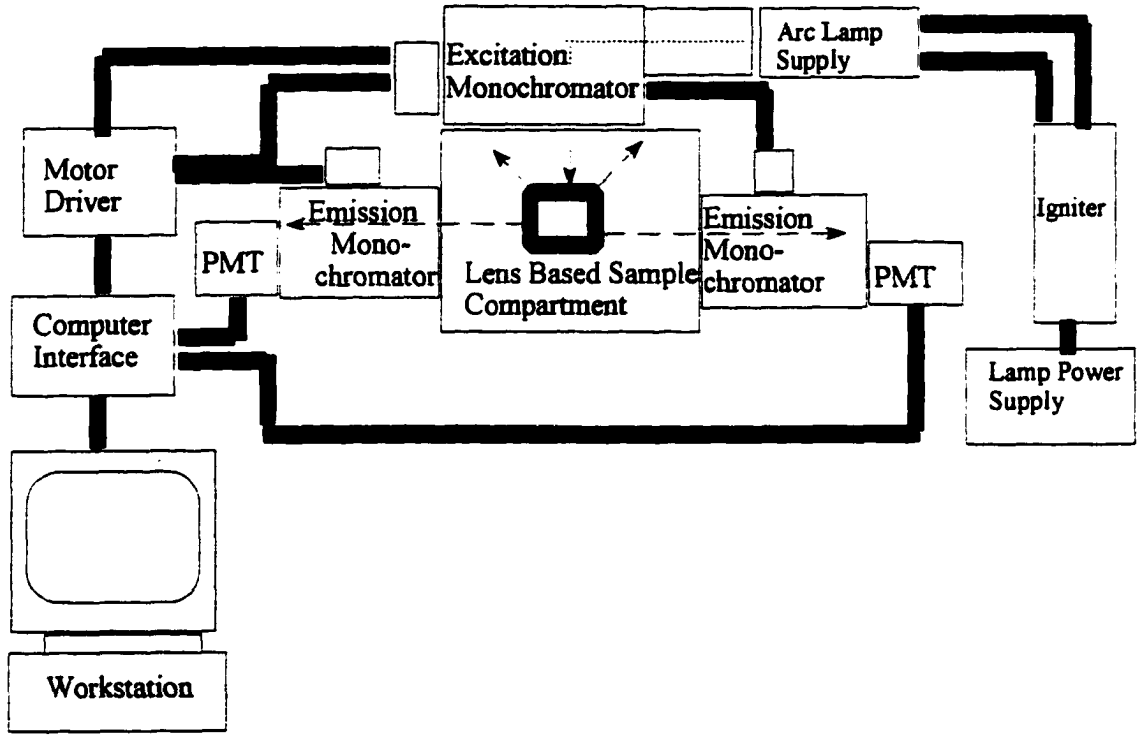


Figure 3.5 General Lightpath through the Fluorometer. As shown by the dashed lines, light enters the excitation monochromator from the lamp supply. The excitation wavelength is set to 284 nm using the computer interface. Light of this wavelength enters the sample compartment and is absorbed by the fluorophores in the sample. The fluorophores emit photons in all directions as they relax back down to the ground state. Those photons emitted at right angles are detected by the photomultiplier tube (PMT) as the emission monochromator scans from 297 to 350 nm. A signal from the PMT is fed to the attached computer and is recorded as a graphical representation of photons emitted per second at each wavelength.

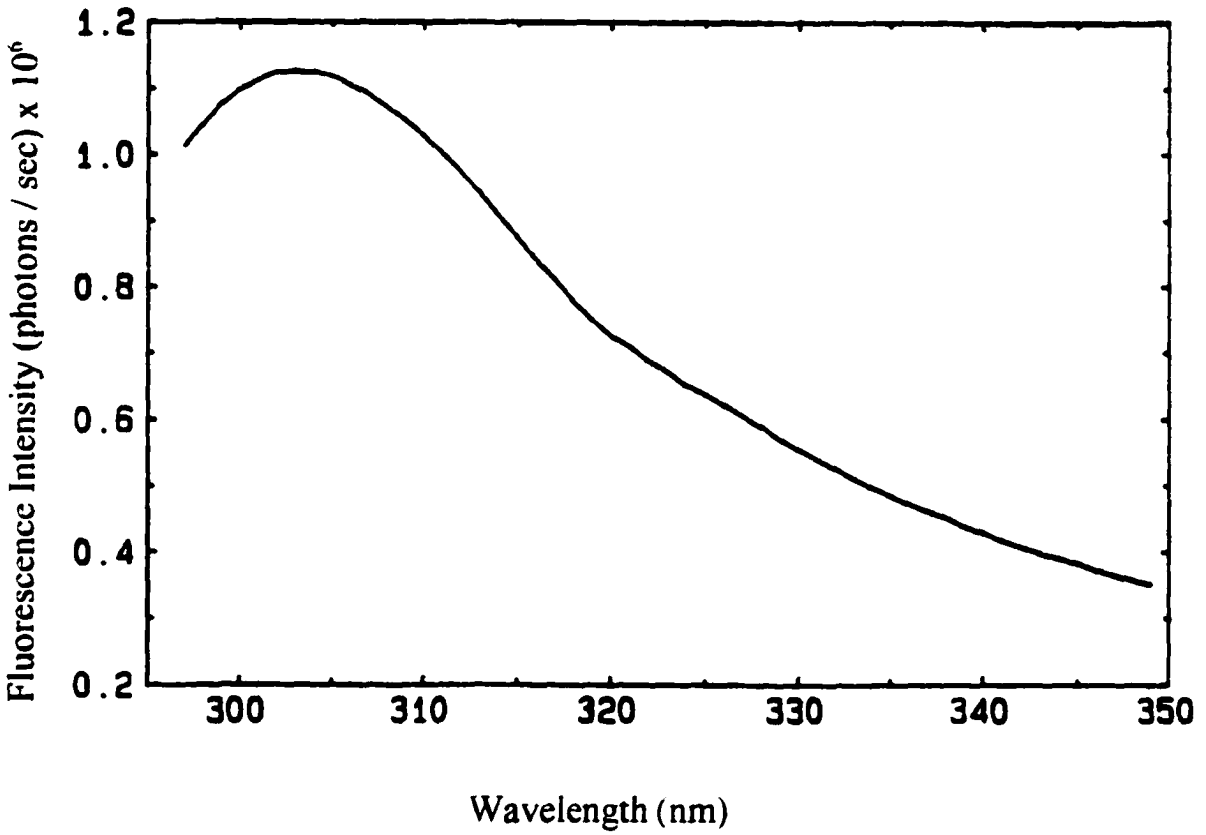


Figure 3.6 Typical Tyrosine Emission. The signal from the photomultiplier tube of the fluorometer is recorded by the attached computer as a spectrum of photons emitted per second at each wavelength. This spectrum shows peak tyrosine emission at approximately 302 nm when the fluorophore is excited at 284 nm.

for tyrosine. The spectra obtained at each temperature were corrected for buffer and water emission, and for loss of efficiency in the detector and gratings. They represent the photon flux emitted at each wavelength, over a pre-determined wavelength interval, as determined by slit widths and dispersion of the emission monochromator.

Inner filter effect was also taken into account during the fluorescence studies. It refers to the non-linear relationship between sample concentration (fluorophore concentration) and fluorescence emission. The corrected fluorescence is approximated by the equation:

$$F_{corr} = F_{obs} \text{ antilog } [(OD_{ex}+OD_{em})/2]$$

where $OD = \text{Log } (I_o/I)$ and OD_{ex} and OD_{em} represent the optical densities of the sample at the excitation and emission wavelengths, respectively (Lakowicz, 1983). However, since the same concentration was maintained throughout an experiment, the OD values did not change. Each data point was therefore multiplied by the same correction factor, changing only the magnitude of the fluorescence while the shape of the curve remained the same.

Measuring the conformational stability of a protein requires determining the equilibrium constant (K) and the free energy change (ΔG) for the reaction, Folded (F) ⇌Unfolded (U). Since the unfolding of many small globular proteins has been found to closely approach a two-state folding mechanism, the thermal stability was calculated assuming a unimolecular process as described by MacCallum *et al.* (1997). The fluorescence intensity, or physical parameter Y, observed at any point on the unfolding curve is defined by the equation, $Y = F_f Y_f + F_u Y_u$, where Y_f and Y_u represent the

theoretical values of Y characteristic of the folded and unfolded states, respectively, under the conditions where Y is being measured. F_f and F_u refer to the fraction of protein present in the folded (f) and unfolded conformations (u) and their sum is set to equal 1 (i.e. $F_f + F_u = 1$). By combining these two equations, the fraction of protein that is unfolded (F_u) at a given temperature is equal to $(Y_f - Y)/(Y_f - Y_u)$. In the transition region, values for Y_f and Y_u must be calculated by extrapolating the pre- and post-transition regions and using a least squares analysis to determine representative linear equations (Y_f and Y_u).

Provided the unfolding of the protein is reversible, the equilibrium constant, K, and the free energy change, ΔG, can be calculated using the following equations:

$$K_u = \frac{F_u}{(1-F_u)} = \frac{F_u}{F_f} = \frac{(Y_f - Y)}{(Y - Y_u)}$$

and, $$\Delta G_u = -RT \ln K = -RT \ln \frac{(Y_f - Y)}{(Y - Y_u)}$$ where R=gas constant, T=absolute temperature

It was therefore imperative to test the reversibility of unfolding for each protein. In order to do so, the protein samples were allowed to equilibrate at the starting temperature after the experiment was complete. If the reaction was reversible, the protein would refold to its original pre-transition state and give the same fluorescence emission spectrum. One of the causes of irreversibility in protein unfolding is the formation of aggregates at temperatures above the melting temperature. Aggregation was also monitored, by visibly checking the samples after each measurement. A cloudy appearance or the presence of small particles would indicate protein aggregation.

3.3 RESULTS

3.3.1 Mutant Sequence

Plasmid DNA was isolated from the two ampicillin resistant colonies obtained during the mutagenesis procedure. Using the M13/pUC (-47) (24mer) sequencing primer, the *Bst* polymerase synthesized the non-coding strand of DNA 5' to 3', from just downstream of the *Hin*dIII site across the gene. **Figures 3.7 and 3.8** show the sequencing results. Reading 5' to 3', the GTT corresponding to the AAC codon for asparagine (N) in MB-1 was not replaced in clone 1. Being the wild-type sequence for MB-1, it worked well as a control. In the second clone, the GTT was replaced by GCA which corresponds to the TGC codon for the cysteine residue in the translated protein.

3.3.2 Production of Pure MB-1-Cys Monomer

The introduction of a single mutation into a protein can alter its structure and properties. As a result, the production of the mutant often requires modifying the procedures used for the wild-type protein. This was the case for MB-1-Cys. The inclusion of 2 mM DTT in the purification buffers was required in order to prevent formation of mixed disulfide bonds. Such a simple modification resulted in the production of pure MB-1-Cys monomer as seen in **figure 3.9**.

3.3.3 *p*-HMB Binding to MB-1-Cys

p-Hydroxymercuribenzoate, shown in **figure 3.10**, has a specific binding affinity for free thiol groups. When added to a solution containing a cysteinyl protein, its mercury atom forms a covalent bond with the sulfur atom on the protein. Formation of this complex results in an increase in absorbance at 255 nm. Since this change in

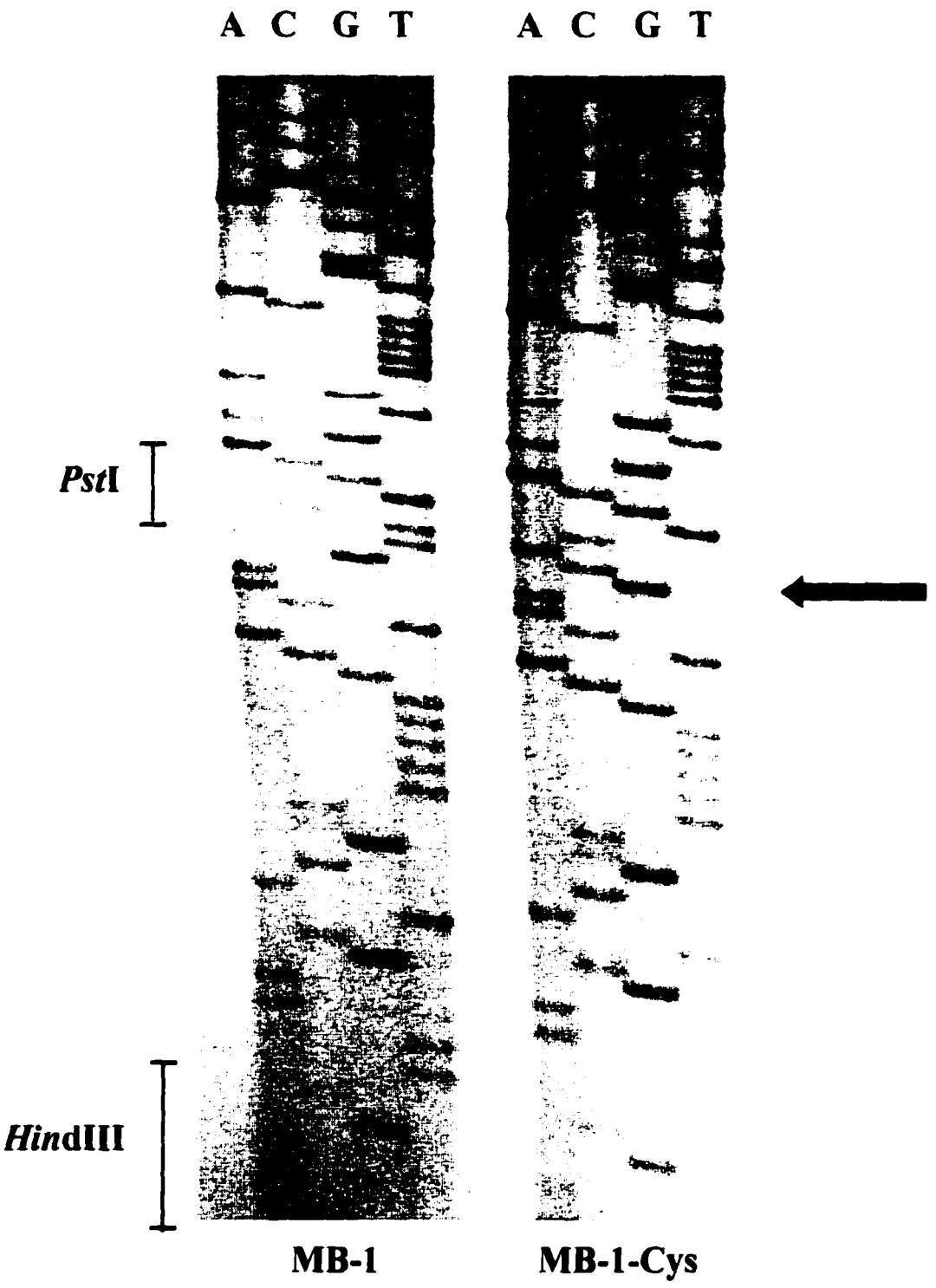


Figure 3.7 Sequencing Gels for MB-1 and MB-1-Cys Plasmids. Non-coding strands were sequenced 5' to 3' in the C-terminal region. The mutation is indicated by an arrow as the sequence is read from the bottom to the top of the gel. The GTT corresponding to asparagine (N93) in MB-1 was changed to a GCA. This corresponds to the TGC codon for cysteine in the mutant. *Pst*I and *Hin*dIII sites are shown for reference points, and the read sequence in the region of the cassette is shown in **figure 3.8** for clarification

*Pst*I
sticky end

*Hin*dIII
sticky end

*C L M Q K G V A

5' - G **TGC** TTG ATG CAA AAA GGC GTA GCT TA -3'
3' - AC GTC **ACG** AAC TAC GTT TTT CCG CAT CGA ATT CGA -5'

Figure 3.8 Sequence of Mutant Cassette. Cys1 (top) and Cys 2 (bottom) oligonucleotides were annealed to form the cassette shown. This segment of dsDNA houses the cysteine mutation (TGC) and flanking MB-1 sequences. The amino acid sequence is indicated by the single letters above the cassette. When the M13/pUC (-47) (24mer) sequencing primer was used to sequence the two clones, it annealed just downstream of the *Hin*dIII site and sequenced the non-coding or bottom strand in the 5' to 3' direction (**Figure 3.7**). As indicated by the underlined region, the mutation observed on the gel was GCA which replaced the GTT corresponding to asparagine in MB-1.

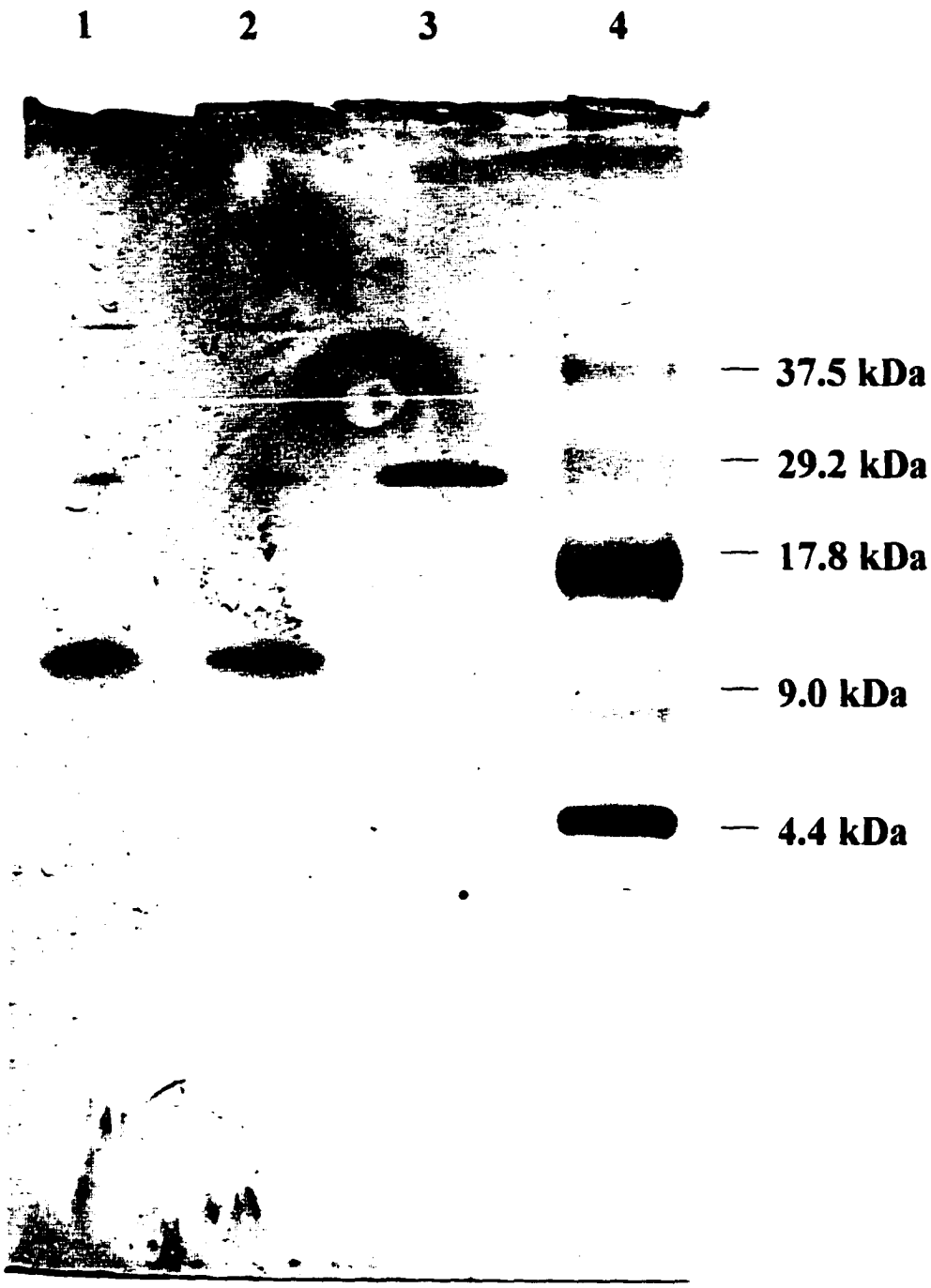


Figure 3.9 SDS-PAGE Analysis of Disulfide Formation. Lanes 1 and 2 show MB-1-Cys monomer after purification using dithiothreitol (DTT) to prevent formation of mixed disulfides. After removal of the DTT, dimer formation was spontaneous as seen in lane 3. Polypeptide markers were loaded in lane 4 for comparison (37.5 - 4.4 kDa).

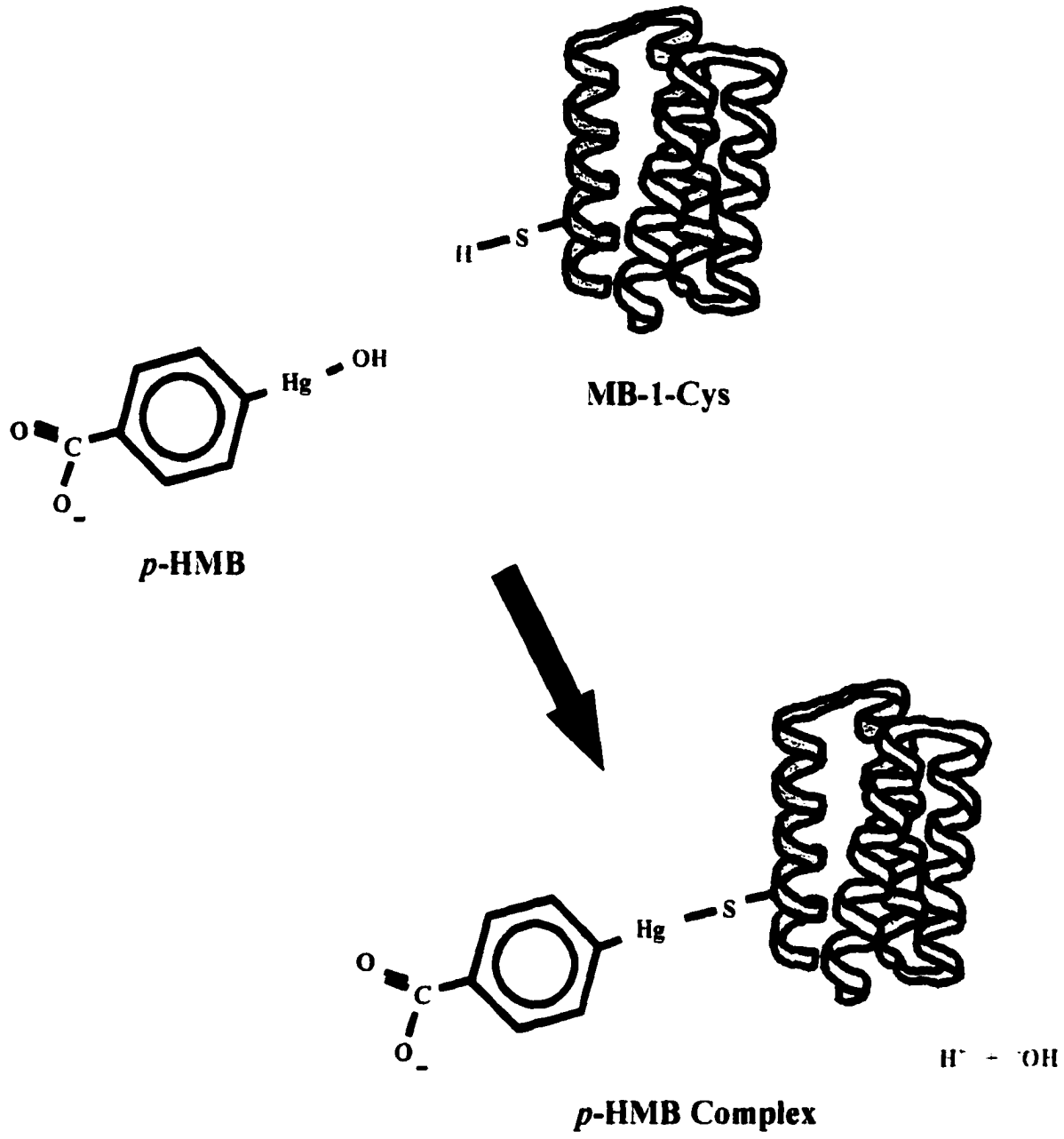


Figure 3.10 *p*-Hydroxymercuribenzoate Assay. When *p*-HMB is added to a solution containing cysteinyl protein, its mercury (Hg) atom forms a covalent bond with the sulfur of the protein. The formation of *p*-HMB complex causes an increase in the absorbance of *p*-HMB at 255 nm. By measuring the ΔA_{255}, the number of cysteine residues in the protein can be determined.

absorbance is proportional to the number of complexes formed, it is indicative of and can be used to detect the number of cysteines in the protein (Boyer, 1954; Friedman, 1973).

Titration of MB-1-Cys with *p*-HMB showed an absorbance change of 0.327 a.u. at 255 nm when *p*-HMB was added to the MB-1-Cys. An absorbance of 0.327 corresponds to a *p*-HMB-Cys derivative concentration of 54.6 μM (Boyer, 1954). The concentration of MB-1-Cys present in the assay was estimated to be 50.8 μM, thus all of the Cys residues available on MB-1-Cys reacted with *p*-HMB. Not only did this assay identify the number of sulfhydryl groups on the protein (~1), but also it showed that the MB-1-Cys monomer is capable of binding mercuric salts. A comparison experiment done on MB-1 showed no *p*-HMB binding.

3.3.4 Demonstration of Disulfide Bond Formation

Initial characterization trials have indicated that MB-1-Cys readily dimerizes in the absence of reducing agent and that this capacity is not increased by the use of glutathione-catalysed oxidation or exposure to air or oxygen. The spontaneous reaction of Cys residues on adjacent proteins shows that as per design, the residue is not buried in the hydrophobic core but is accessible to the solvent. SDS-PAGE showed the conversion of MB-1-Cys monomer to disulfide product upon removal of DTT by dialysis into 0.1 M Tris, 0.4 μM $CuCl_2$ buffer, pH 8.2 **(figure 3.9, p. 68).** A plot of log M_r versus relative mobility was used in order to determine the molecular weights of MB-1-Cys monomer and MB-1-Cys Dimer. **Figure 3.11** shows one of the three standard curves created using the log molecular weight values for each protein standard and their relative mobilities. A linear regression analysis was performed on each set of data, and the slope and constant values obtained were used to calculate the molecular weight of each protein. Comparison

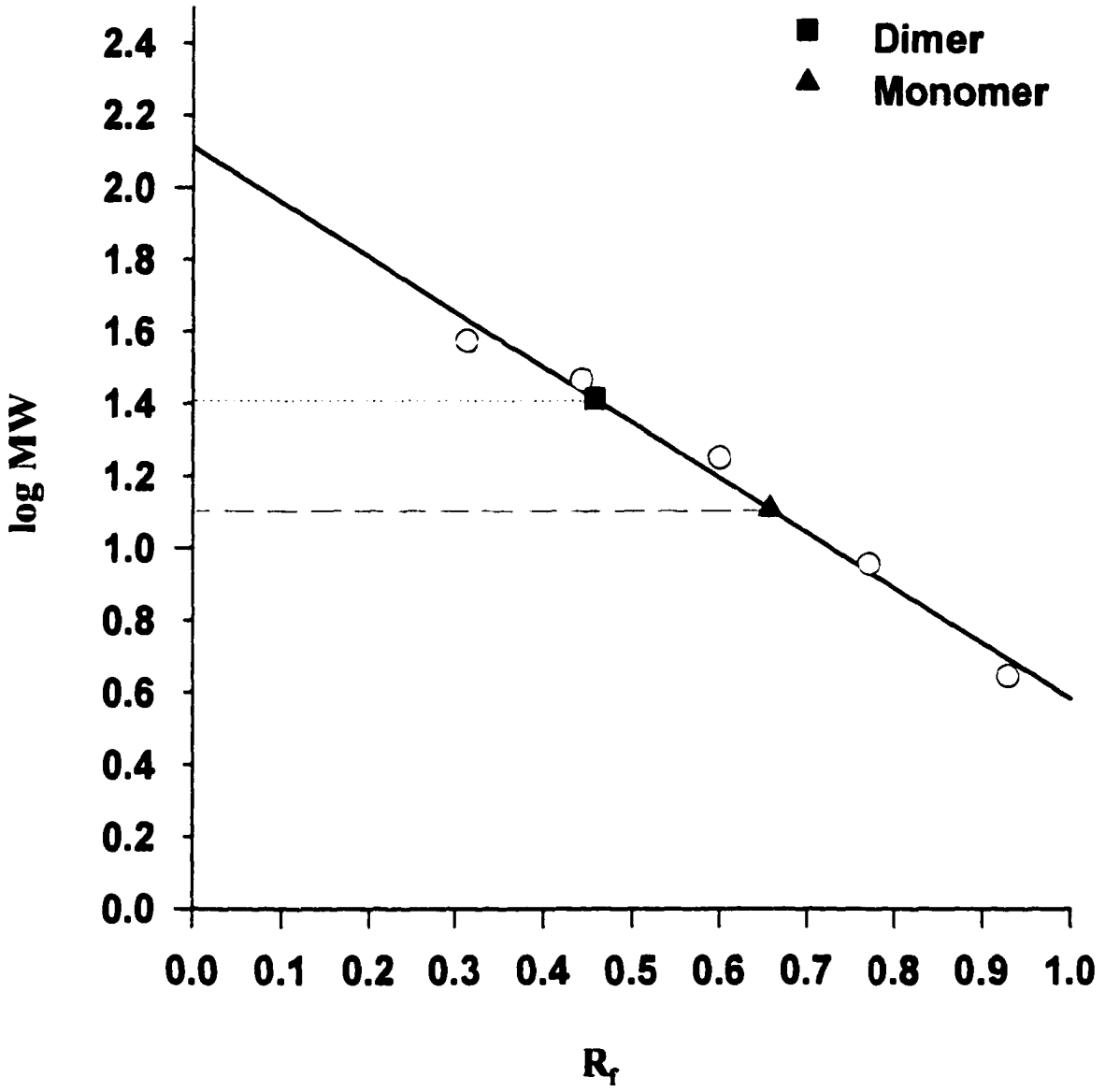


Figure 3.11 Typical Standard Curve for Determination of Molecular Weight. For each of three trials, a linear regression was performed on the log MW and R_f data for the molecular weight standards. This typical result shows molecular weight standards represented by open circles and a standard curve with slope -1.532 and y-intercept 2.1139. The relative mobility of the Dimer was calculated to be 0.4571 for this trial, while that for the monomer was 0.6571. Using the slope and y-intercept values, the corresponding log molecular weight values were 1.4136 and 1.1072, respectively. These values are indicated by the dotted line for the Dimer (square), and correspond to the position of the triangle for the monomer. By taking the antilog of these values, the molecular weights were 25.9 for the Dimer and 12.8 for the monomer. The average values over the three trials were approximately 26 kDa and 13 kDa, respectively.

to the molecular weight markers revealed that the monomer is approximately 13 kDa while the Dimer is approximately twice that (26 kDa).

3.3.5 Degradability of MB-1-Cys Dimer vs MB-1

Exposure of the proteins to Pronase E resulted in gels similar to the one in **figure 3.12**. As the protein was degraded, the band intensity decreased. This was reflected by the area under the curve values from the densitometer.

Degradability results could not be obtained for MB-1-Cys monomer. The DTT required to keep the protein in its reduced form was shown to cause problems during the experiments. Studies comparing Cyt C degradation in the presence and absence of DTT revealed that the activity of the proteases was inhibited by the reducing agent. Since it was the difference between MB-1 and MB-1-Cys Dimer stabilities that was of interest, no further studies were attempted on MB-1-Cys monomer.

Studies exposing MB-1 and MB-1-Cys Dimer to Pronase E were performed in triplicate. **Figure 3.13** shows the average amount of protein remaining at each time interval for both proteins. As the graph illustrates, MB-1 disappeared after only 50 min of exposure to the enzymes, while MB-1-Cys Dimer was still slightly visible after 120 min.

3.3.6 Conformational Stability of MB-1-Cys Dimer Compared to MB-1

The folding stabilities of MB-1 and MB-1-Cys Dimer were determined by monitoring the changes in tyrosine fluorescence as temperature was increased. **Figure 3.14** shows the difference in fluorescence emission of tyrosine in the folded and unfolded

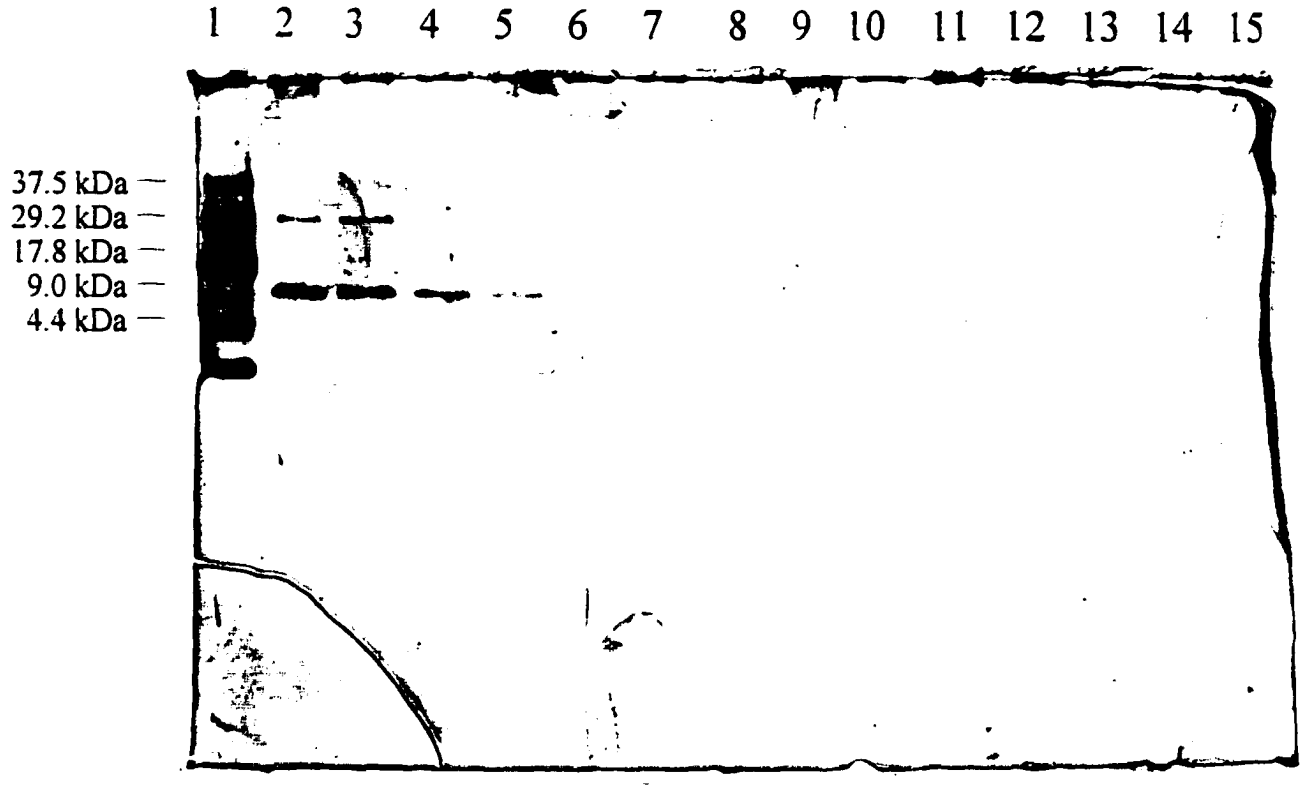


Figure 3.12 Typical Gel of a Degradation Trial Using Pronase E. Lane 1 shows molecular weight markers, lane 2 contains the time zero (T_0) sample. and lanes 3-15 show samples taken at subsequent time intervals (10 min, 20 min. 30 min, 40 min, and so forth). The intensity of the bands tend to decrease with time as the protein is degraded.

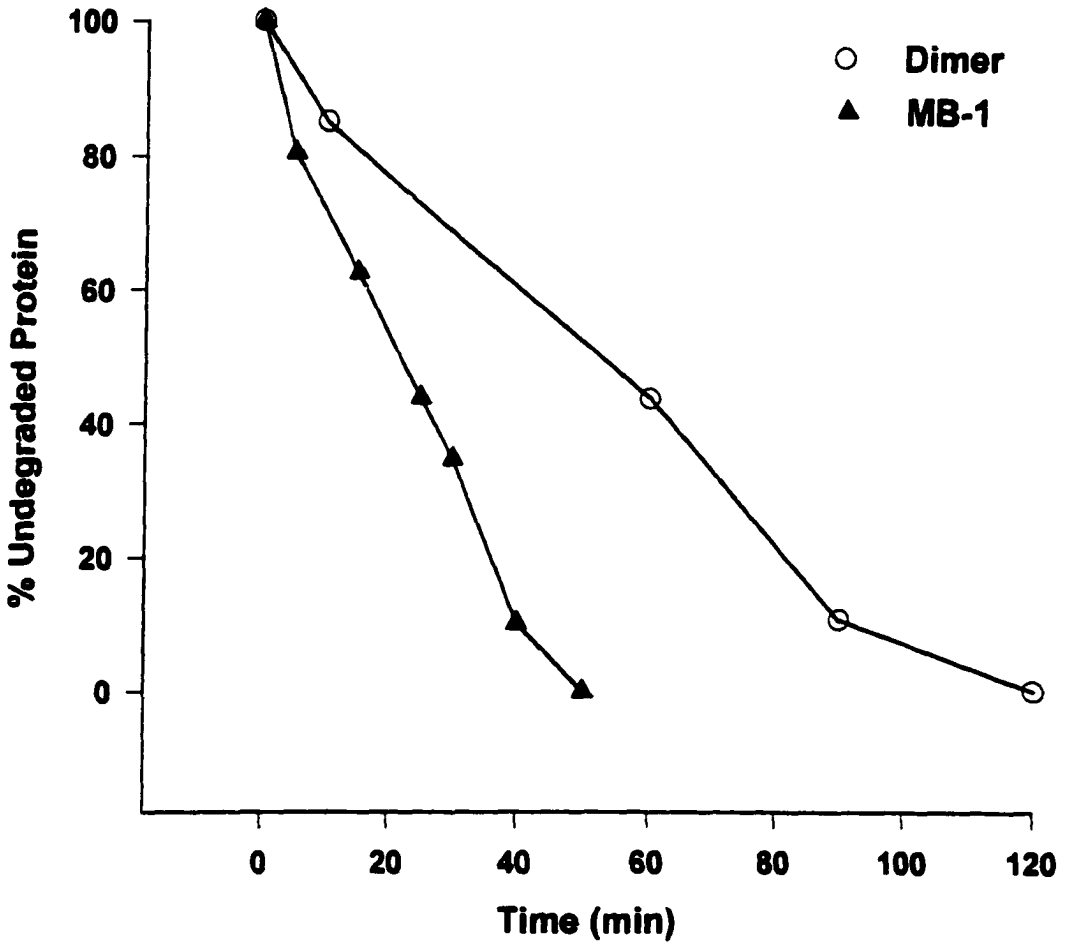


Figure 3.13 Digestion of MB-1 and MB-1-Cys Dimer using Pronase E. The studies were run in triplicate. This graph shows the average values at each time interval. Original MB-1 was fully digested after a 50 minute incubation period while the dimer was still faintly visible after 120 minutes. Average SD = 10%.

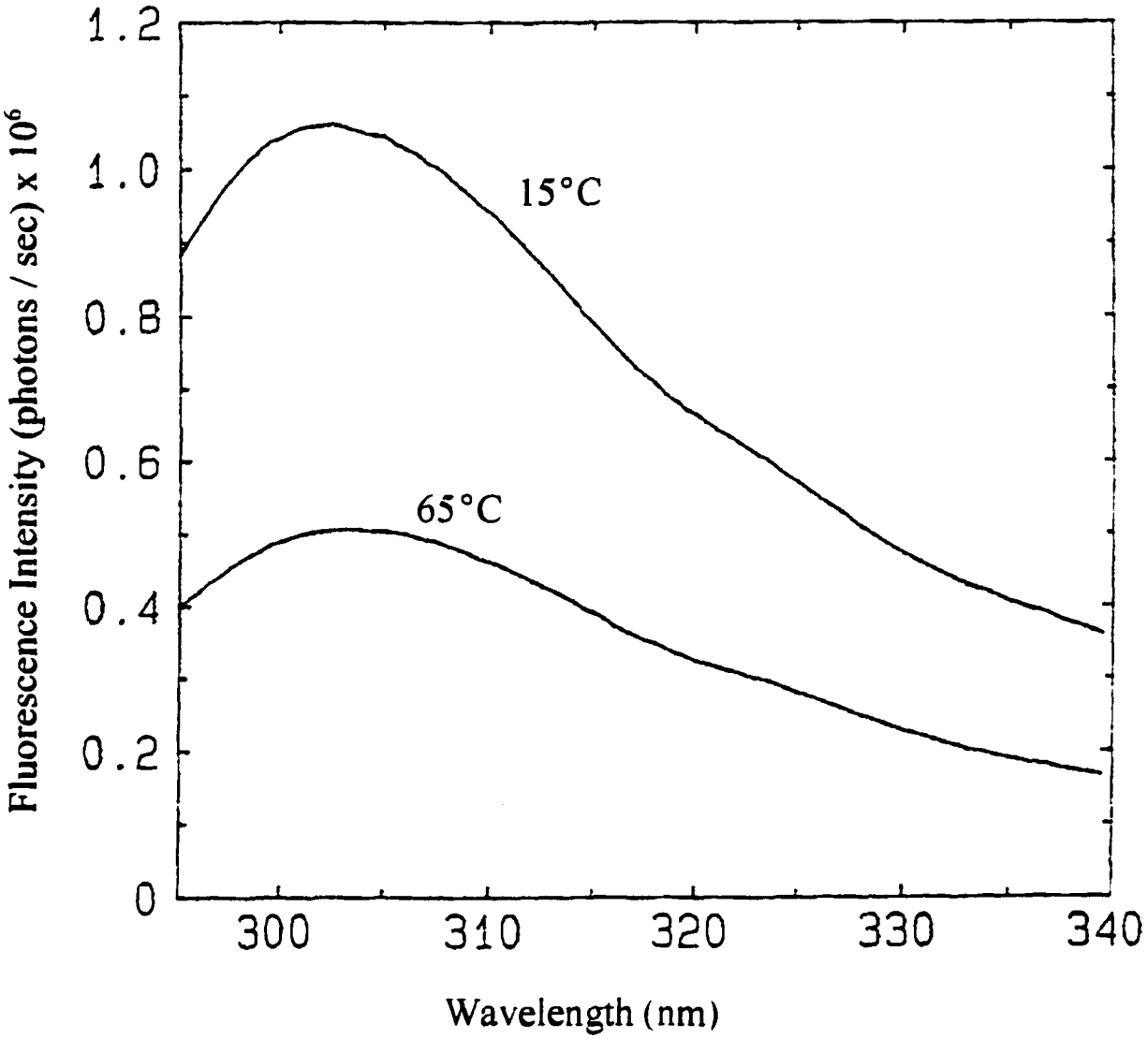


Figure 3.14 Fluorescence Emission Spectra for MB-1. The top curve shows a fluorescence spectrum of MB-1 in VFA buffer at 15°C. The bottom curve shows a spectrum of the same sample at 65°C. The decrease in fluorescence intensity is a result of more efficient quenching by the buffer in the unfolded state.

states. As seen in the graph, the maximal emission for tyrosine is at 302 nm. When the intensity at 302 nm (I_{302}) was plotted as a function of temperature, the unfolding curve shown in **figure 3.15** was obtained. As the temperature was increased, the protein gradually unfolded, exposing the tyrosine to buffer components capable of quenching its fluorescence. This increased quenching is evident on the curve as a small, but distinct decrease in I_{302}.

The unfolding curves obtained for MB-1 and MB-1-Cys are very similar as seen in **figures 3.15 and 3.16**. Each has three regions: a pre-transition region, a transition region, and a post-transition region. Points in the pre-transition region were used to define curve Y_f. This curve reflects how the fluorescence of the folded protein would change with temperature if it never unfolded. Similarly, points in the post-transition region were used to define curve Y_u, illustrating the trend for the unfolded protein if it never folded. Over short temperature ranges, these curves are linear and can be defined using linear regression analysis.

The equation for Y_f **(see figures 3.15, 3.16)** was used to calculate the fluorescence intensities that would have been observed at each data point if all of the proteins in the sample had remained folded. Each of these values (Y_f values) therefore represented a theoretical value for Y at a given temperature. Equation Y_u was employed in a similar fashion to determine the theoretical fluorescence intensities, or Y_u values, for a completely unfolded sample. Since the observed fluorescence intensities (experimental Y values) fell between these theoretical values, it was possible to use the equation $F_u = (Y_f-Y)/(Y_f-Y_u)$ to calculate the fraction of unfolded protein (F_u) at a given

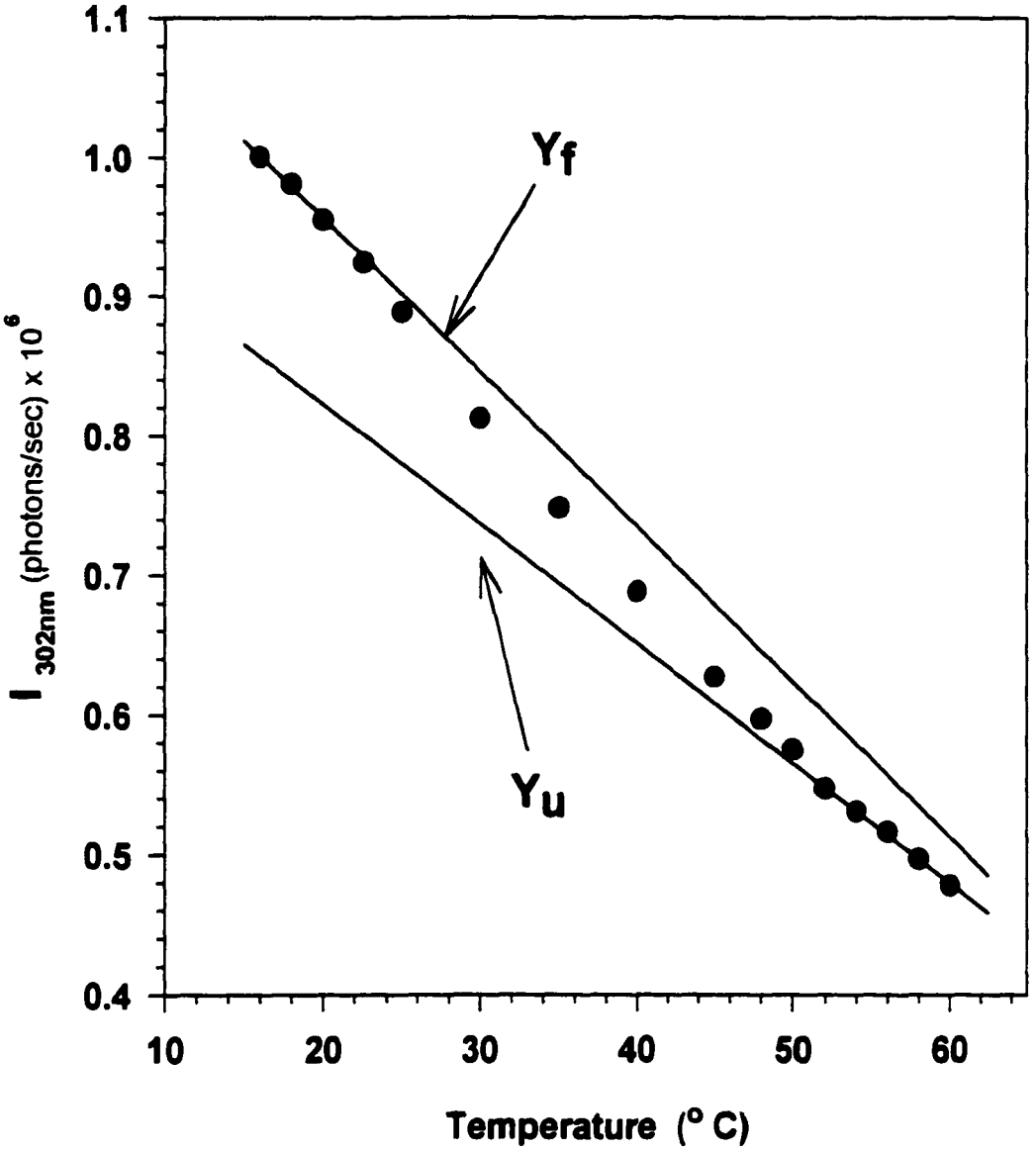


Figure 3.15 Unfolding Curve for MB-1. Effect of temperature on the fluorescence of tyrosine in MB-1. As the temperature is increased, the protein unfolds and the fluorescence intensity decreases. Y_f defines the change in fluorescence of a 100% folded sample with change in temperature. Y_u defines the change in fluorescence of an unfolded sample with change in temperature. Equations for Y_f (-0.0111x + 1.17) and Y_u (-0.0086x + 0.99) were used to calculate theorectical values for Y at a given temperature. These theoretical values were used, in turn, to determine the fraction of protein unfolded.

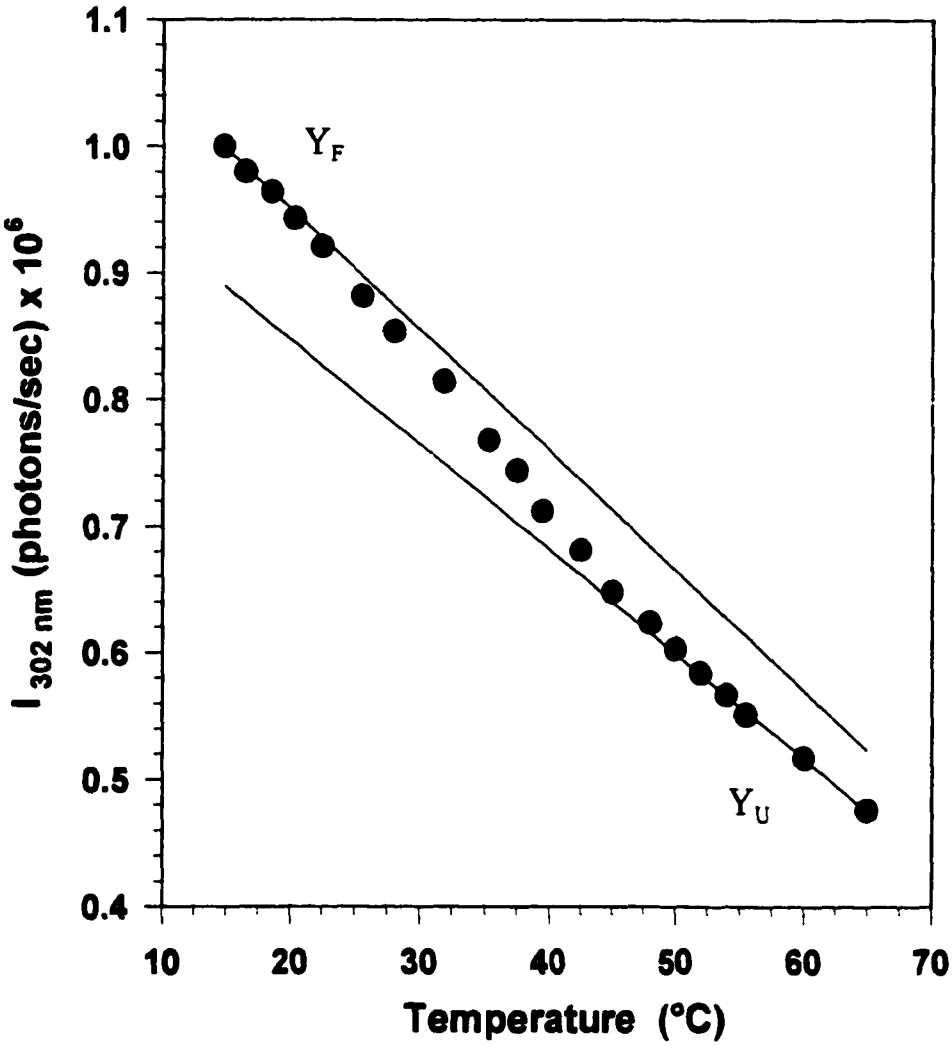


Figure 3.16 Unfolding Curve for MB-1-Cys Dimer. Effect of temperature on the fluorescence of MB-1-Cys Dimer. The lines indicate the equations defined for Y_f (-0.0095x + 1.14) and Y_u (-0.0083x + 1.01).

temperature. The fraction folded (F_f) was then determined for each data point using the formula $F_f = 1 - F_u$. It is from a fraction folded plot like that in **figure 3.17** that the melting temperature (T_m) of the protein can be determined. The T_m is the point at which 50% of the protein population is unfolded (or when $F_f = 0.5$). For the MB-1 trial shown, the T_m was found to be 38°C.

Since the fraction folded curve for the MB-1 study revealed approximately the same T_m as previously measured (39±1.7°C; MacCallum *et al.*, 1997), denaturation of MB-1 was not repeated. Fraction folded curves from the two Dimer studies were averaged and can be seen compared to MB-1 in **figure 3.18**. At a F_f value of 0.5, the T_m for both MB-1 and MB-1-Cys Dimer was found to be approximately 38°C. This means that the introduction of the cysteine mutation and disulfide linkage has not altered the conformational stability of the protein.

In order to calculate equilibrium thermodynamic parameters such as the equilibrium constant (K_u) and change in free energy (ΔG_u), the unfolding of MB-1 and MB-1-Cys must be reversible. Both proteins were found to unfold irreversibly, however, as they aggregated at temperatures above the T_m and did not return to the native fold when allowed time to renature. The reaction that took place was therefore Folded (F) → Unfolded (U) → Aggregated (A) rather than Folded (F) ⇌ Unfolded (U). Consequently, a detailed equilibrium thermodynamic analysis was not performed.

3.4 DISCUSSION

3.4.1 The Construction of MB-1-Cys Dimer

The first MB-1 mutant was successfully constructed. Dideoxy-sequencing confirmed that the cysteine mutation had been incorporated into the MB-1 gene, and

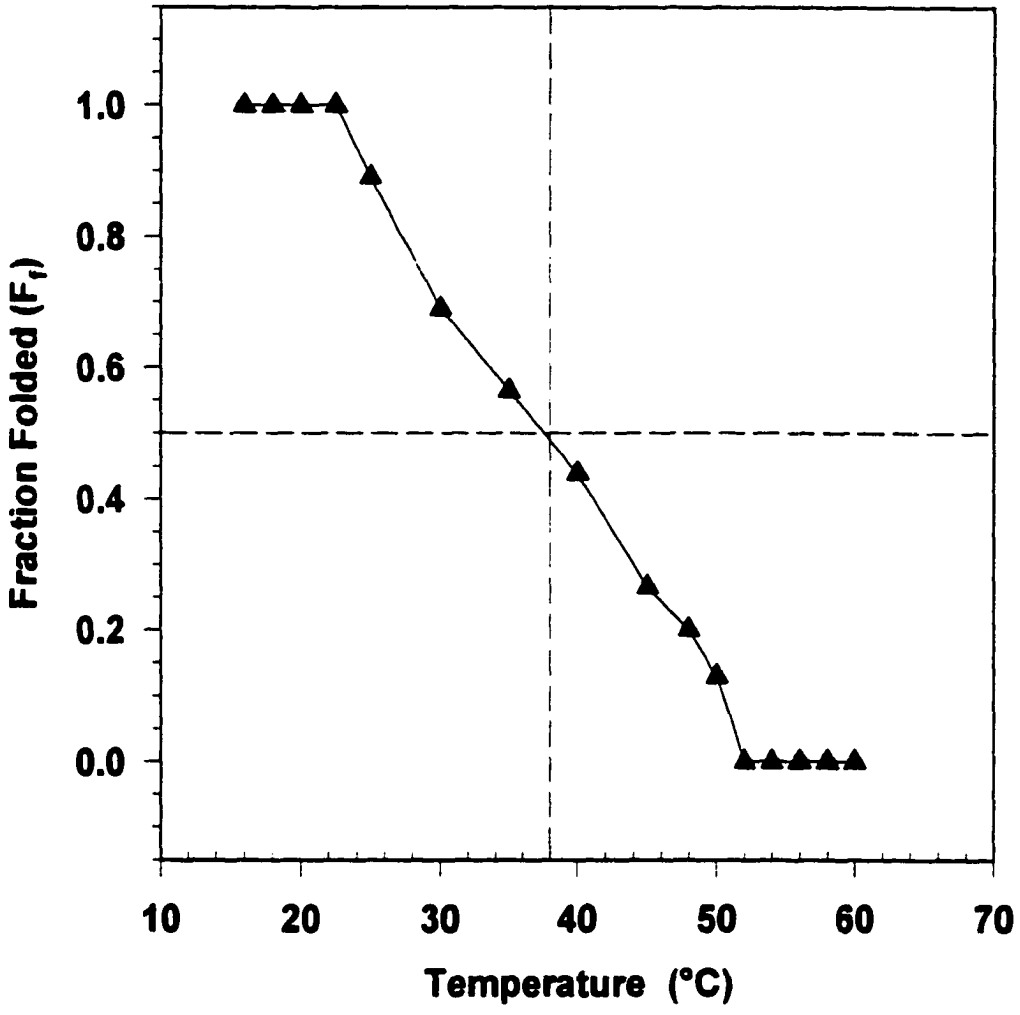


Figure 3.17 Fraction Folded Curve for MB-1. The effect of temperature on the fraction of the MB-1 population that is folded. F_f starts off at a value of 1, where all of the population is folded. Then, it gradually decreases until it levels off at 0 where all of the MB-1 population is unfolded. The melting temperature (T_m) of 38 °C was determined from the point where $F_f = 0.5$

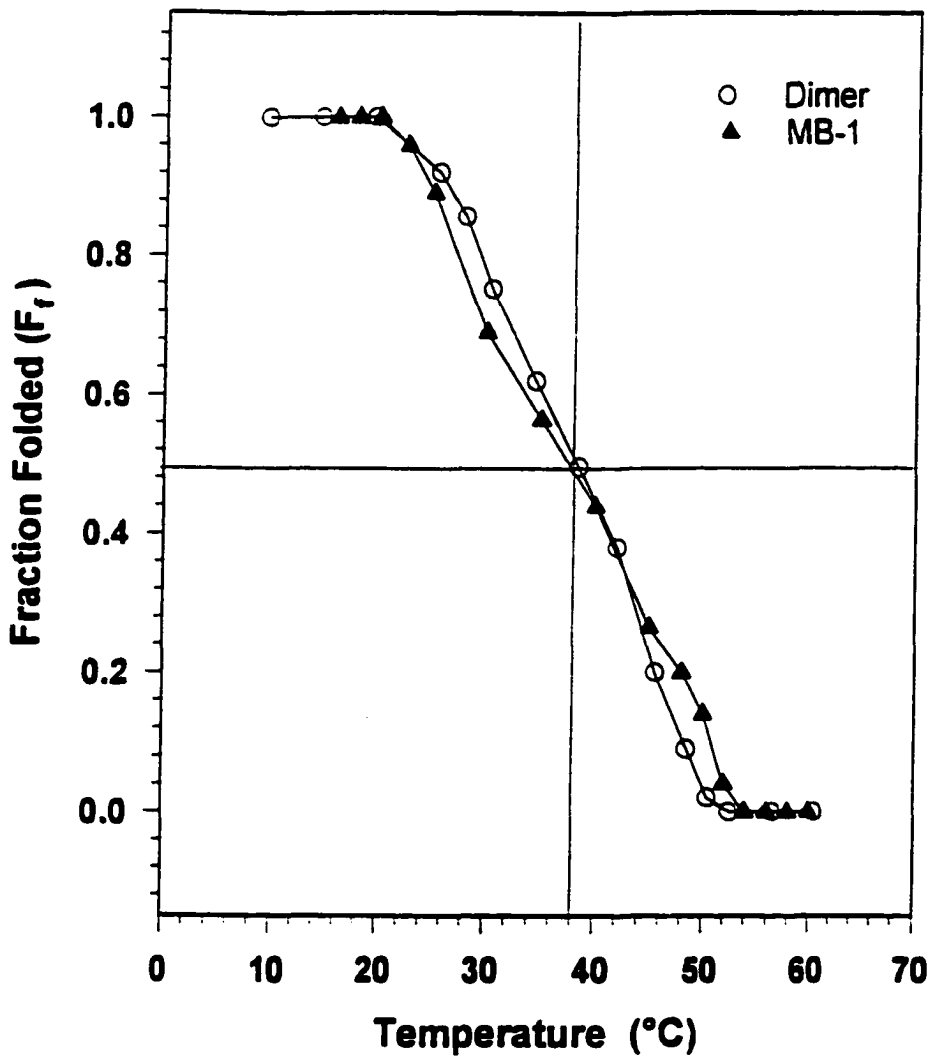


Figure 3.18 Fraction Folded Curves for MB-1 and MB-1-Cys Dimer. The fraction folded (F_f) starts at a value of 1 when the protein is folded, then decreases with temperature as the protein is unfolded. When the protein is completely unfolded, the F_f shows a value of 0. The T_m is the point at which half of the protein population is unfolded ($F_f = 0.5$). As shown in the crosshairs, both MB-1 and MB-1-Cys Dimer have a T_m of approximately 38°C.

subsequent expression in *E.coli* resulted in production of the mutant protein, MB-1-Cys. Extraction of pure MB-1-Cys from the bacterial lysate required that a strong reducing agent (DTT) be included in all purification buffers (except cleavage buffer). Upon addition of the DTT, contaminating bands that had been observed during initial production trials disappeared. These bands were obviously mixed disulfides, as suggested by SDS-PAGE (**figure 3.9**). Their formation served as the first indication that the sulfhydryl group of the cysteine was exposed.

Titration of the protein with *p*-hydroxymercuribenzoate not only verified the presence of the single cysteine residue, but also proved that MB-1-Cys monomer is capable of binding mercuric salts. The fact that the amount of complex formed during the titration (54.6 μM) was greater than the amount of MB-1-Cys available (50.8 μM) can be easily explained by the mode of protein quantitation used. Since the presence of DTT interfered with quantitation of MB-1-Cys, half of the sample had to be dialysed into buffer with no DTT, and resulted in dimer formation. Some of the sample could have been lost during the dialysis, either through leakage or through adhesion to the dialysis bag when the sample was retrieved. The amount of protein in the dialysed sample would therefore be less than that in the original buffer. As a result, the concentration used to determine the molar ratio of *p*-HMB complex formed to MB-1-Cys available would have been underestimated, making the ratio higher than expected.

Nevertheless, formation of the MB-1-Cys Dimer was spontaneous upon removal of the DTT from purification buffers. This demonstrates that the cysteine was exposed and available for disulfide formation, as per design. The molecular weight of MB-1-Cys Dimer was 26 kDa, approximately twice that of the monomer. However, the exact

weight may be slightly less since the presence of the disulfide inhibits optimal binding of SDS and mobility through the gel (Creighton, 1989).

3.4.2 Assessment of Proteolytic and Conformational Stability

The ultimate goal of linking the two monomers together by a disulfide bond was to decrease the surface area exposed by each and limit the amount of degradation by rumen enzymes. Studies exposing MB-1 and MB-1-Cys Dimer to Pronase E revealed that indeed, this was the case. The MB-1-Cys Dimer was twice as resistant to proteolytic degradation. A portion of this enhanced stability was undoubtedly due to the decrease in exposed protease targets; yet, contributions from another potentially stabilizing source had to be investigated.

Many research groups have been successful in increasing the conformational stability of a protein - often with only a single amino acid mutation. For instance, a Gln16 to Leu mutation of λ-repressor was shown to increase the T_m by 14°C (Pakula and Sauer, 1989). As well, the Asn57 to Ile mutation of yeast iso-1-cytochrome c increased the T_m by 17°C (Das *et al.*, 1989). With the introduction of a cysteine residue and subsequent disulfide linkage between MB-1 units, the possibility existed that its conformational stability had been enhanced, rendering it more stable at rumen temperature and resistant to proteolytic degradation. In order to determine whether such a change had taken place in MB-1-Cys Dimer, thermal unfolding studies were performed on both the Dimer and MB-1. Results indicated that each protein had a T_m of approximately 38°C. Since the melting temperature of MB-1 had not been altered as a result of the mutation, the enhanced proteolytic stability of the dimer could not be attributed to an increased thermal stability.

The use of Pronase E extract to estimate degradation rates in the rumen has been criticized by some researchers. They found that it did not correlate well with *in vivo* techniques (Nocek *et al.*, 1983; Sniffen *et al.*, 1979; Krishnamoorthy *et al.*, 1983; Poos-Floyd *et al.*, 1985; Roe *et al.*, 1991; Luchini *et al.*, 1996). However, even the validity and reproducibility of *in vivo* and *in situ* methods have been questioned. The rumen environment is extremely complex, and impossible to mimic with accuracy. It not only varies from animal to animal, but also it changes according to dietary intake and microorganism diversity (Asplund, 1994; Krause and Russell, 1996). Protein degradation rates depend on the stability of the protein, the conditions of the rumen (temperature, pH, types of microorganisms), the retention time in the rumen, and also on the type and amount of food in the rumen (Church, 1988; Taminga, 1979; King *et al.*, 1990). Since so many factors influence the rate of rumen degradation, it would not be logical to assume that the degradation rate experienced under rumen conditions could be approximated by a single protease extract.

3.4.3 Conclusion

While it may not reflect exactly the degradation in the rumen, digestion with Pronase E is reliable for studying the relative proteolytic stability of two proteins. In this thesis work, the MB-1-Cys Dimer was found to be twice as resistant to proteolytic degradation compared to MB-1. These results demonstrate that the decrease in exposed surface area has conferred stability to MB-1. The three experiments were reproducible to within approximately 10% standard deviation for each protein. However, the relative degradability of MB-1-Cys Dimer compared to MB-1 was the same for each trial. Furthermore, studies recently performed on the two proteins using a rumen protease

extract (RPE) lend support to the Pronase E results. The RPE studies showed a faster degradation rate for both, but maintained that the MB-1-Cys Dimer is twice as resistant to proteolysis (Morrison *et al.*, in preparation; Navidzadeh, 1998).

CHAPTER 4 - MUTANT 2: INTRAMOLECULAR DISULFIDE

4.1 INTRODUCTION

4.1.1 Design of the Mutant

Chapter 3 described the construction of a mutant that was twice as resistant to proteolytic degradation as MB-1 due to its decrease in exposed surface area. As seen in **section 3.1.1**, this was only one of the approaches which could be taken to enhance the rumen stability of the protein. The other would be to increase its conformational stability, either through the introduction of fold specifiers, or through the manipulation of the bundle core. In this chapter, an attempt is made to enhance the proteolytic stability of MB-1 with the introduction of fold specifiers.

When MB-1 was designed, no fold specifiers were included to stabilize its tertiary structure (Beauregard *et al.*, 1995). Such features include metal binding sites, salt bridges, and disulfide bonds. Each of these is known to stabilize the fold of a protein, but at the same time presents a challenge to engineer. The introduction of a metal binding site would require the correct geometric arrangement of at least two or three metal binding amino acids like histidines or cysteines (Regan and Clarke, 1990; Handel and DeGrado, 1990; Wrede and Schneider, 1994). Furthermore, engineering a metal binding site into MB-1 would create problems for its intended use. To be an effective feed additive for ruminant nutrition, the protein must be inert. With the ability to bind metals in the body, such a mutant could induce biochemical stress in whatever organism it is found. In view of this, it was decided not to engineer a metal binding site into MB-1.

The second type of fold specifier which could be introduced is the salt bridge. A salt bridge is formed between oppositely charged amino acids like glutamic acid (-) and lysine (+). The side chains of the amino acids must be in close enough proximity to allow for charge interaction (within 3-5 Å; Dao-pin *et al.*, 1991). When engineered into a protein, they can be placed so as to favor one type of topology through charge attraction or to disfavor another through repulsive forces (Graddis *et al.*, 1993; Bryson *et al.*, 1995; Betz and DeGrado, 1996). Like the metal-binding site, a salt-bridge could be introduced to favor a particular conformation or fold of MB-1, but its effect on the thermal stability of the protein can not be predicted. In most cases, however, since the interactions are not covalent, their contribution to the thermal stability of the protein is negligible (Dao-pin *et al.*, 1991; Hill *et al.*, 1990; Stitger and Dill, 1990; Matthews, 1993).

The other option for increasing MB-1's conformational stability was to introduce a disulfide bond between two of its helices. This covalent linkage is found in many naturally occurring proteins, and has been introduced by mutation into others (Zhou *et al.*, 1993; Wetzel *et al.*, 1988; Kanaya *et al.*, 1991; Luckey *et al.*, 1991). The ability of an intramolecular disulfide to enhance the conformational stability of a protein has been well established (Wrede and Schneider, 1994; Creighton, 1989). It does so by destabilizing the unfolded state of the protein. Since the unfolded conformation is restricted by the bond and not allowed as much freedom (or entropy), it is not as favored as usual (**figure 4.1**; Zhou *et al.*, 1993; Creighton *et al.*, 1995). The reaction is therefore pushed more towards the folded conformation. As a result, a greater amount of heat is required to unfold the protein. The insertion of intramolecular disulfides into T4 lysozyme, for instance, resulted in a significant increase in its conformational stability (Matsumura *et al.*, 1989 a,b). Single disulfide bond insertions increased the T_m by as

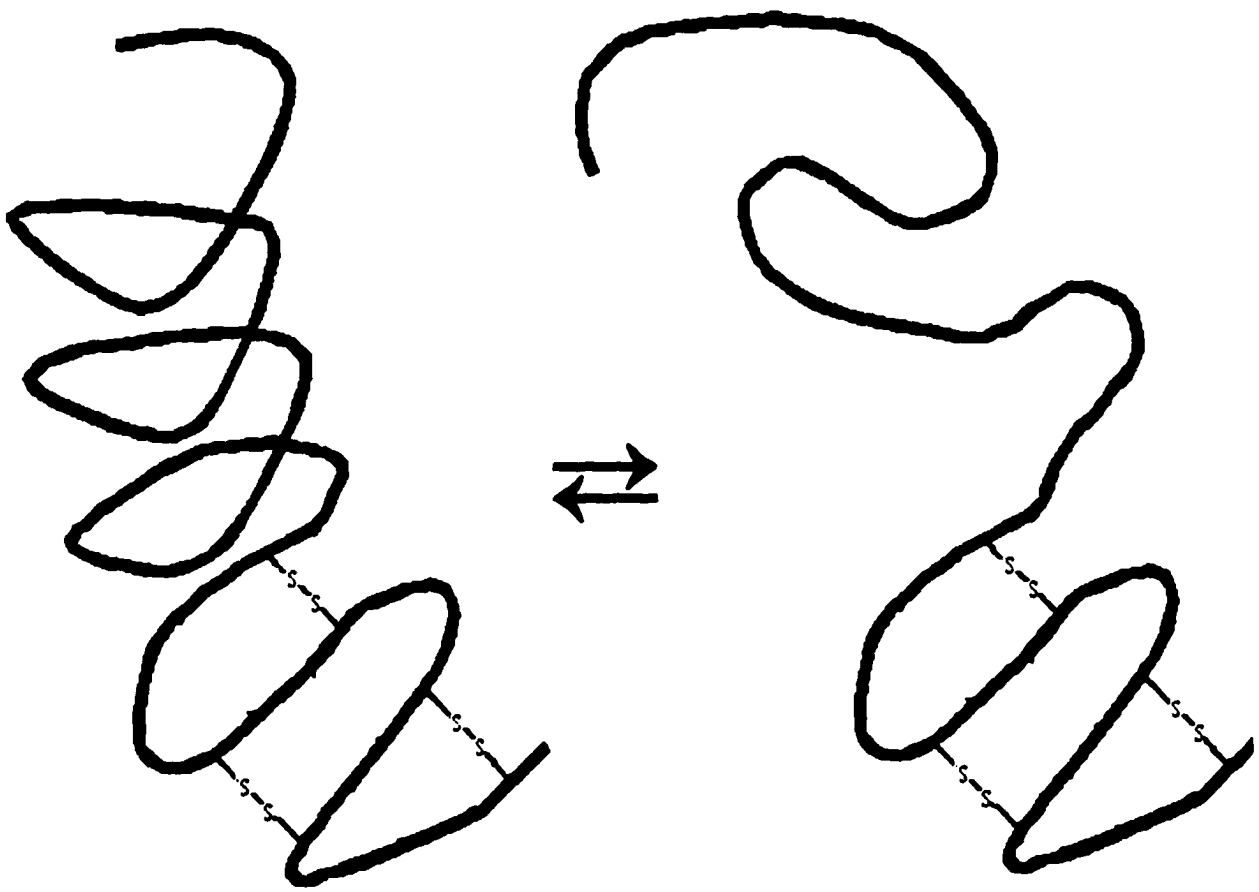


Figure 4.1 Schematic of Unfolding. The presence of intramolecular disulfide bonds destabilizes the unfolded state by restricting the freedom of the protein chain. As a result, the free energy of unfolding is increased and it takes more heat to unfold the protein.

much as 11 °C while double and triple disulfide insertions did so to an even larger extent (up to 23.4 °C).

What is more important, in terms of MB-1, is that the presence of disulfide bonds is known to confer resistance against degradation in the rumen (Hancock *et al.*, 1994; McNabb *et al.*, 1994; Spencer *et al.*, 1988; Mahadevan *et al.*, 1980). The reasons for introducing an intramolecular disulfide into MB-1 were therefore two-fold. First, the disulfide bond should increase the conformational stability of the protein, rendering it folded at rumen temperature. And second, this enhanced conformational stability should be reflected by an increase in resistance to rumen proteases.

4.1.2 Protein Engineering

Since MB-1 contains no cysteine residues, the first step in construction of the disulfide bond was to determine where the two cysteines should be positioned. Not only would they have to be closely apposed in the folded protein, but also they should be placed in a region that would offer the most stability. The disulfide bonds that provide the greatest stability in proteins are those found in the most closely packed region - the hydrophobic core (Creighton, 1988). This is indeed the case for most naturally occurring proteins, but the difficulty in introducing them by mutation lies in choosing the best position. The stabilizing effect of the disulfide is only attained if the strict stereochemical requirements of the bundle interior are met (Zhou *et al.*, 1993; Creighton *et al.*, 1995). If the introduction of the disulfide creates strain on the protein, its presence would be counter-productive and destabilizing. Therefore, great care must be taken to ensure that the positions of the cysteine residues are optimal.

MB-1 was designed to fold into a square α-helical bundle (Beauregard *et al.*, 1995). However, as of yet, the exact tertiary structure of the protein is unknown. Without crystallographic data on the structure, it is impossible to use modelling software to help visualize the positioning of the amino acids. The task of engineering a specific folding interaction into the protein is therefore even more difficult. In order to do so, it must be assumed that the protein is folded as per design. As described in **section 1.4.2, p.16**, the results of several studies suggest that it is. Unfortunately, the "handed-ness" of the folded protein has not been determined. This refers to the connectivity between the helices of the bundle. When the loop connecting helices I and II goes to the right, the bundle is said to be right-handed. If it turns to the left, the bundle has left-handed connectivity. The difference between the two is illustrated in **figure 4.2**. Since one of the natural proteins used in the design of MB-1 is a left-handed bundle (ROP), a possibility exists that the connection between helices goes to the left rather than to the right. For an intramolecular disulfide to form in a left-handed bundle, the cysteines must be positioned differently than in a right-handed bundle. Consequently, the introduction of a disulfide bond between MB-1 helices offers a way of determining the connectivity of the protein.

To pursue this angle of study, two mutants have been designed with the potential to form a disulfide bond. The first of the mutants will house cysteine residues in position 'a' of the heptad repeat **(figure 4.3)**. Placed near one another in the core of the protein, the cysteines should form a disulfide bond. That is, provided the protein is a right-handed bundle. The second mutant will house cysteines in position 'd' of the bundle. These residues would also be located in the core of the protein, but can only form a disulfide if the helices show left-handed connectivity. Since only one mutant will

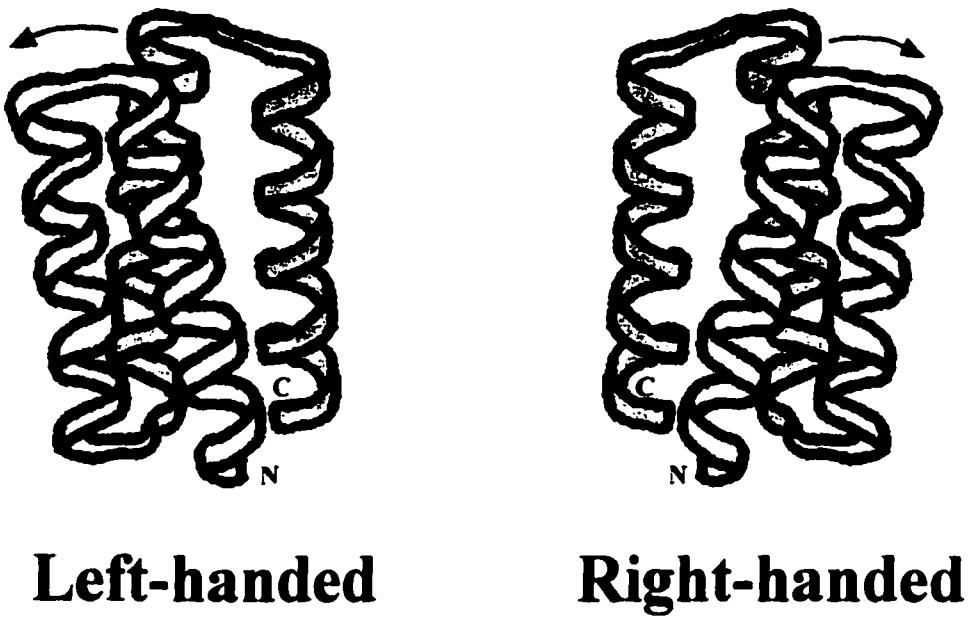


Figure 4.2 Left-handed vs Right-handed Connectivity. In a left-handed bundle, the loop connecting the first and second helices turns to the left. MB-1 was designed to be a right-handed bundle. If it folds as predicted, the connection between the first and second helices should go to the right as depicted above.

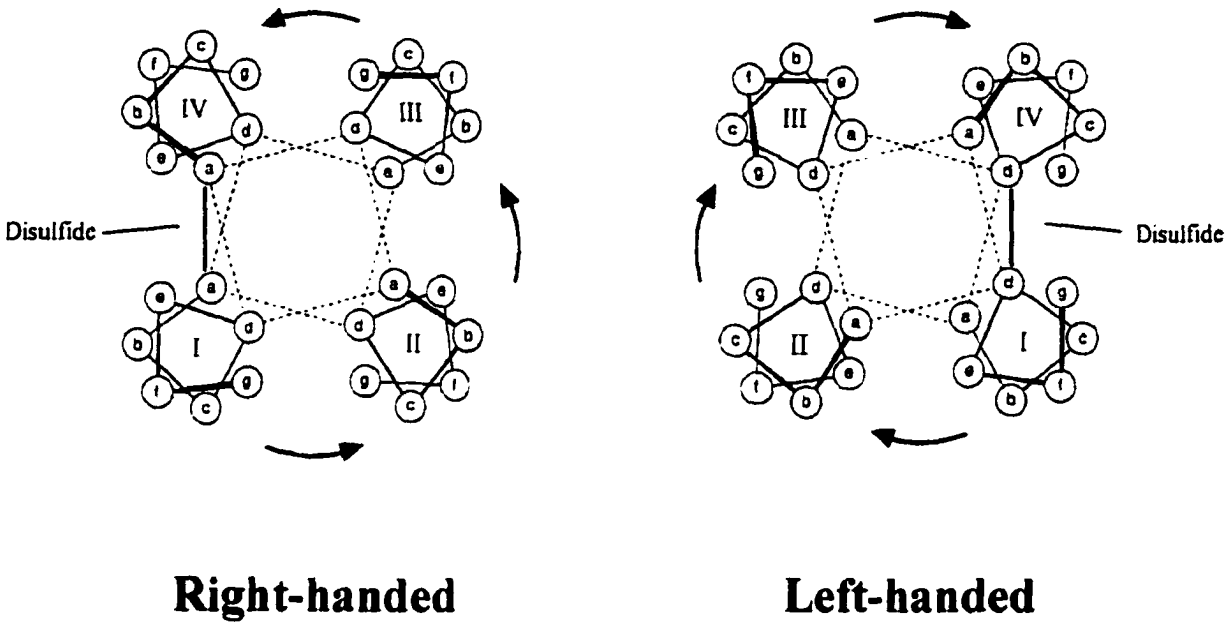


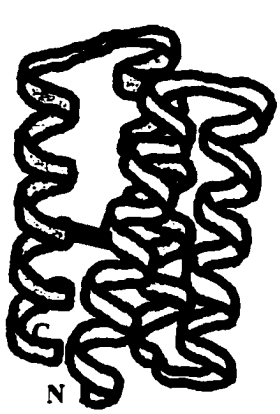
Figure 4.3 Disulfide Position for Right-handed and Left-handed Bundles. For a disulfide to form between helices I and IV in a right-handed bundle, the two cysteine residues must be placed in position 'a' of the heptad. If located near one another in the folded protein, the sulfhydryl groups of the cysteines should be capable of forming the disulfide. That is, provided the protein is right-handed. If the protein is left-handed, the cysteines would have to be placed in position 'd'.

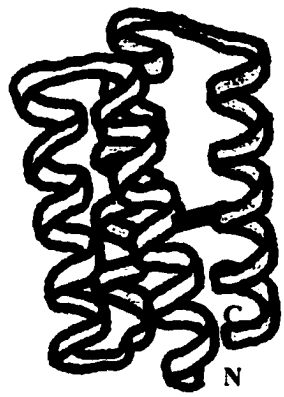
actually be capable of forming a disulfide, the connectivity of the protein can be deduced accordingly.

The disulfide will be engineered between core residues in helices I and IV. In this location the bridge would hold the end strands of the protein in a fixed conformation, preventing them from opening up and exposing core residues. Such a fold specifier would offer added protection from proteases by protecting target regions found in the core. To create the mutations in this region, the cysteines will be inserted as shown in **figure 4.4**. For the "right-handed" disulfide, a methionine in position 10 and leucine in position 91 will be replaced (M10C, L91C). These residues are predicted to be in 'a' positions of the bundle, and located in helix I and IV, respectively. The "left-handed" mutant will have cysteine residues in place of leucine 13 in helix I and methionine 87 in helix IV (L13C, M87C). Predicted to be in position 'd' of the heptad, these residues would be closely apposed and capable of forming a disulfide if MB-1 is left-handed.

4.1.3 Genetic Engineering

With the two cysteine mutations located so far apart in the sequence of the protein, it was decided that the easiest way to create the mutant plasmids was to perform oligo-directed mutagenesis on the protein-coding gene. An Altered Sites® II *in vitro* Mutagenesis System (Promega Corporation, Madison, WI) was therefore employed to make the mutagenesis procedure more efficient. **Figure 4.5** shows a schematic representation of how the Promega Kit works. Essentially, the gene to be mutated is first cloned into the pALTER-1 mutagenesis vector. This vector is resistant to tetracycline and sensitive to ampicillin. Once the insertion of the gene has been verified by DNA sequencing, the recombinant plasmid is alkaline-denatured. Purified oligonucleotides





Right-handed

Helix I: M10C
Helix IV: L91C

Left-handed

Helix I: LI3C
Helix IV: M87C

Figure 4.4 Positioning of the Cysteine Mutations. To create the right-handed disulfide mutant, methionine in position 10 and leucine in position 91 will be replaced by cysteine residues (M10C, L91C). If MB-1 is folded as per design, this mutant should be capable of forming a disulfide bond. However, the possibility exists that MB-1 has left-handed connectivity. In order to investigate this possibiltiy, a second mutant will be created. It will house cysteines in place of the leucine in position 13 and the methionine in position 87 (L13C, M87C). These cysteines will only be capable of forming a disulfide if the bundle is left-handed. The connectivity of MB-1 can therefore be deduced by determining which mutant forms the disulfide.

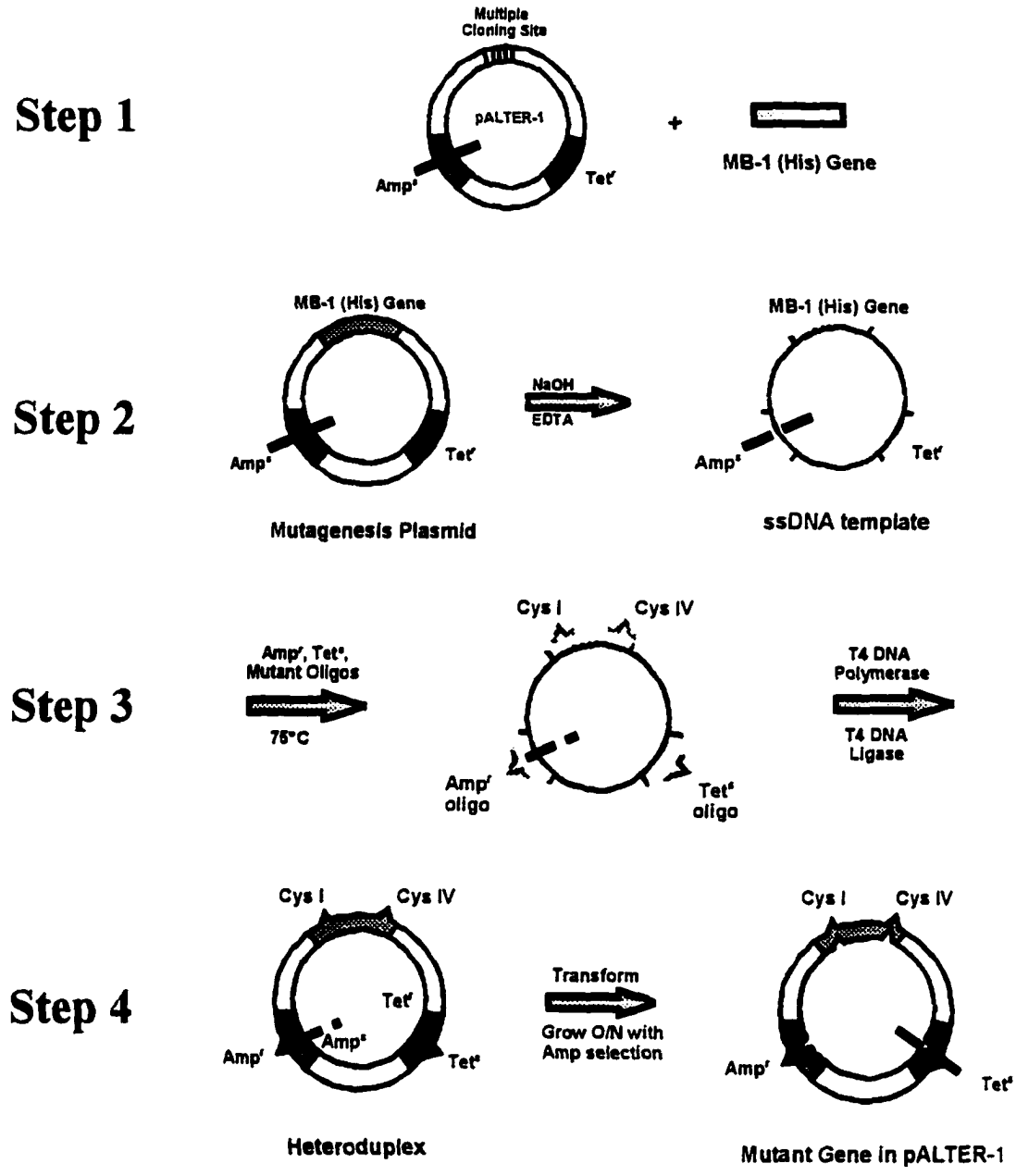


Figure 4.5 Schematic of the Promega *in vitro* Mutagenesis Kit. First the MB-1 or MB-1-His gene must be cloned into the pALTER-1 vector. The resulting mutagenesis plasmid is then alkaline-denatured. In step 3, the mutant, repair, and knock-out oligos are annealed to the DNA template and the mutant strand is produced by addition of T4 DNA polymerase and T4 DNA ligase. In the final step, the heteroduplex of DNA is transformed into a repair minus strain of bacteria and selected for ampicillin resistance. Only bacteria containing the mutant plasmid should grow.

housing the cysteine mutations are then annealed to the denatured template along with a tetracycline knock-out oligo and an ampicillin repair oligo. Addition of T4 DNA polymerase and T4 DNA ligase results in synthesis and ligation of the mutant strand. Finally, the heteroduplex is allowed to replicate in a repair minus strain of bacteria while selecting for ampicillin resistance. Since the mutant plasmid would be rendered ampicillin resistant by the repair oligo, any bacteria containing mutant plasmid will grow. Bacteria containing wild-type plasmid should not grow as they would still be sensitive to the antibiotic. In this way, the Promega Kit offers a way to ensure that the majority of the colonies obtained contain the mutant plasmid. Screening for the cysteine mutations should in turn be relatively simple.

Oligonucleotides for the mutagenesis procedure were designed such that they would anneal to the same template strand as the repair and knock-out oligos. The 5'-3' DNA strands corresponding to the cysteine codon (UGC) and flanking MB-1 sequences were therefore used to create the oligos (**Figure 4.6**). With 12 base pairs on either side of the mutation, these oligos should anneal strongly to the DNA template and permit incorporation of the mutation into the synthesized strand.

Another MB-1 mutant was prepared and characterized when these experiments began. Designed to have a tag of 6 histidines at the C-terminal of the protein, the MB-1-His mutant can be purified using a nickel column rather than the amylose and DEAE sepharose columns described in **section 2.1, p. 35**. The ability to use this recyclable column greatly reduces the time and money spent making protein (Grundy *et al.*, 1998). Since future work in the lab is shifting towards the use of this protocol for all MB-1 proteins, it was deemed beneficial to introduce the disulfide mutations into MB-1-His at the same time.

Left-handed

wt: 5' - ATG ATG ACC ACC CTG TTT AAA ACT ATG - 3'
C13: 5' - ATG ATG ACC ACC TGC TTT AAA ACT ATG - 3'

wt: 5' - ACG GCT ACA ACC ATG AAA AAT CAT CTG -3'
C87: 5' - ACG GCT ACA ACC TGC AAA AAT CAT CTG - 3'

Right-handed

wt: 5' - ATG ACC GAC ATG ATG ACC ACC CTG TTT -3'
C10: 5' - ATG ACC GAC ATG TGC ACC ACC CTG TTT - 3'

wt: 5' - ATG AAA AAT CAT CTG CAG AAC TTG ATG -3'
C91: 5' - ATG AAA AAT CAT TGC CAG AAC TTG ATG - 3'

Figure 4.6 The Mutagenic Oligonucleotides. Both the wild-type and the mutant sequences are depicted for the "left-handed "and "right-handed " disulfide mutants. In order to create the mutant oligos, the underlined wild-type triplets were replaced by TGC. As seen for the left-handed mutant, this cysteine coding sequence was used to replace a CTG coding for leucine 13 and an ATG coding for methionine 87. For the right-handed mutant, the ATG coding for methionine 10 and the CTG coding for leucine 91 were replaced. Since the oligos were designed to have 12 wild-type nucleotides on either side of the mutation, they should anneal strongly to the template DNA and allow for efficient incorporation of the mutation.

4.1.4 Chapter Objectives

Objectives of this chapter were:

- ❑ To introduce two cysteine mutations at specific positions into the MB-1 gene in order to allow for formation of an intramolecular disulfide bond
- ❑ To produce and purify the mutant protein by optimizing the original protocol
- ❑ To verify the presence of the two cysteine residues
- ❑ To determine whether or not the disulfide bond was formed
- ❑ To compare the proteolytic stability of the mutant protein to that of MB-1
- ❑ To assess the conformational stability of the mutant protein relative to MB-1

4.2 MATERIALS AND METHODS

4.2.1 Construction of the Mutagenesis Vectors

The genes to be mutated had to be removed from their pMALc2 expression vectors and cloned into the pALTER-1 mutagenesis vector. Using *Sac*I and *Hin*dIII, a double digest in NEBuffer 1 (+100 μg/mL BSA) was performed on both expression vectors (MB-1 and MB-1-His in pMALc2) as well as on pALTER-1. After adding 4 U of each enzyme for every μg of plasmid DNA, the samples were incubated at 37°C for 2.5 h to optimize the enzymatic activity. The digested DNA was then mixed with an appropriate volume of 6X loading buffer (40% sucrose, 0.25% bromophenol blue) and loaded into the wells of a 1.5% agarose gel. An aliquot of ethidium bromide (EtBr: final concentration 0.5 μg/mL) had been added to the gel in order to visualize the DNA bands. The samples were migrated in TAE running buffer at 80 V for 1 h. Bands corresponding to the linearized plasmids and the genes of interest were cut out with a razor blade and placed in labelled test tubes for further purification.

Extraction of the DNA from the agarose was accomplished using a Bio-Rad Prep-a-Gene® Kit (Bio-Rad Laboratories, Mississauga, ON). This kit employs a silica matrix to selectively bind the plasmid DNA and allow it to be separated from contaminating particles like agarose and EtBr. Once bound, the DNA/matrix pellet was washed with 70% ethanol. Then the DNA was eluted from the matrix with a 5 min incubation in Tris EDTA buffer (TE, pH 8.0) at 37°C. Pure expression vector (pMALc2) was stored at –20°C while the purified inserts, MB-1 and MB-1-His, were subcloned into pALTER-1. By combining the linearized pALTER-1 plasmid with either the MB-1 or MB-1-His insert in a 1:4 ratio (by concentration), and incubating them with 20 U T4 DNA ligase (overnight at 16°C), the given genes were incorporated into mutagenesis plasmids. After transformation of the ligation mixtures into competent JM109 cells, the bacteria were plated with tetracycline selection (12.5 μg/mL). Plasmid DNA was then extracted from the resulting colonies for dideoxy sequencing. Confirmation of the plasmid constructs was achieved using the *Bst* Sequencing Kit as described in **section 3.2.2, p. 52**.

4.2.2 Oligo-Directed Mutagenesis

When the subcloning had been confirmed, the templates and oligos were prepared for mutagenesis. Approximately 2 μg of each DNA sample was alkaline-denatured using 2 M NaOH, 2 mM EDTA, and precipitated with 2 M ammonium acetate pH 4.6 and 100% ethanol. The denatured templates were either stored at –20°C or used immediately for mutagenesis.

Synthetic oligonucleotides with the desired mutations were purchased from Medicorp, Inc. (Montreal, Quebec). The lyophilized oligos were resuspended in 1 mL TE buffer and checked by migration on a 15% polyacrylamide, 44% urea gel using Tris

Borate EDTA (TBE) running buffer (30 min prerun, 1 h migration @ 100 V). Bands were visualized by ethidium bromide staining. Half (500 μL) of each oligo solution was desalted on a Sephadex G-25 spun column (350 x g, 2 min) to remove contaminants. Once purified, the oligos were quantitated by measuring the OD at 260 nm in TE buffer. The concentration of each sample was determined using the conversion factor, 1 A_{260} unit = 33 μg / mL ssDNA (New England BioLabs 96/97 Catalogue). Then, just prior to the mutagenesis reactions, the oligos were 5'-phosphorylated with T4 polynucleotide kinase (5 U enzyme : 0.9 μg oligo).

In one reaction, the two mutant oligos (C10 and C91 for right-handed; C13 and C87 for left-handed) and the ampicillin repair and tetracycline knock-out oligos were annealed to the denatured template in a 25:5:1 ratio (mutant oligo : repair/knock-out oligo : template). The components were combined and incubated for 5 min in a block heater set at 75°C. Then the block heater was turned off and the samples were allowed to cool to about 45°C (30 min). After the tubes had been placed on ice for 10-15 min, 10 U of T4 DNA polymerase and 3 U of T4 DNA ligase were added to synthesize and ligate the mutant DNA strand, respectively. An incubation at 37°C for 90 min permitted production of the DNA heteroduplex (one mutant strand, one wild-type strand). Initial transformation of the DNA into repair minus ES1301 cells allowed for production of homoduplex DNA in an environment where mismatch repair was suppressed. In this way, the chance that the mutagenic or antibiotic mismatches would be repaired was decreased. Cells were plated on LB agar with ampicillin selection (100 μg/mL). Ampicillin-resistant plasmids were isolated using the Bio-Rad Prep-a-Gene® Kit and transformed into the final host, JM109. Dideoxy sequencing was then performed on the DNA to screen for mutations.

4.2.3 DNA Sequencing

In order to confirm both the cloning of the gene into the mutagenesis vector and whether or not the mutations were present, dideoxy sequencing was performed on the plasmid DNA. Either the *Bst* Sequencing Kit as described in **section 3.2.2, p. 52**, or the Sequenase Version 2 Sequencing Kit from Amersham was used. For the Sequenase Kit, DNA samples (approx. 5 μg) were prepared as in **section 3.2.2**. They underwent alkaline denaturation with 2 M NaOH, 2 mM EDTA pH 8.0 and precipitation with 2 M sodium acetate pH 4.6, 3 volumes 100% ethanol. After washing each of the DNA pellets with 70% ethanol, they were dried in a vacuum concentrator, and dissolved in 7 μL TE buffer. With a 2 min incubation at 65°C and subsequent cooling, the appropriate primer was then annealed to the denatured DNA template [SP6 sequencing primer for reverse, universal M13 (-20) (17mer) for forward]. Dithiothreitol, diluted labelling mix (containing dCTP, GTP, dTTP), 10 μCi of [α^{-35}S]dATP, and diluted Sequenase were then added to the annealed template and mixed thoroughly. A 5 min incubation at room temperature allowed for extension of the primer and incorporation of radiolabelled ATP.

A 3.5 μL aliquot of this total mixture was then added to each of the four pre-warmed (37°C) reaction tubes labelled A, C, G, and T. As in the *Bst* Sequencing Kit, these tubes had been filled with premixed dNTPs and corresponding ddNTP (10:1 ratio; 2.5 μL). After the primed template/enzyme/radiolabel mix was added to all four tubes, the tubes were incubated at 37°C for 5 min. This allowed for rapid DNA chain elongation and termination of the chain as ddNTPs were incorporated. The reactions were then stopped with the addition of 2X stop solution (95% formamide, 20mM EDTA, 0.05% Bromophenol Blue, 0.05% Xylene Cyanol FF) and placed at -20°C until the gel was ready to load. Care was taken to spin the samples down (microcentrifuge @ 14,000

x g) between steps. Prior to loading, the samples were thawed and then denatured for 2 min at 75°C. Five μL of each was loaded in A, C, G, T order onto a 10% polyacrylamide, 42% urea sequencing gel. The samples were migrated at 55 W for the appropriate time in TBE running buffer (30 min prerun, 90 min for reverse reaction, 3 h 15 min for forward). The gel was then fixed in 10% methanol / 10% acetic acid and then dried as described in **section 3.2.2, p. 52**. Autoradiography of the dried gel was accomplished in the same manner as described earlier for the *Bst* Sequencing Kit **(Section 3.2.2)**.

4.3 RESULTS

4.3.1 Construction of the Mutagenesis Plasmid

Plasmid DNA was extracted from three colonies on each plate (MB-1 and MB-1-His plates). The purified DNA was then sequenced using the SP6 sequencing primer with the *Bst* Sequencing Kit. This primer anneals just downstream of the *Hin*dIII site and allows sequencing across the gene on the non-coding strand. Results are shown in **figure 4.7** for wild-type and mutagenesis plasmids. The ends of the MB-1 and MB-1-His genes are underlined. With the mutagenesis vectors produced, the next step was to introduce the cysteine mutations.

4.3.2 Mutagenesis Procedure

Since Promega claimed to have introduced up to four simultaneous mutations with more than 50% efficiency using their kit, attempts were made to introduce the two cysteine mutations at the same time. For each of the four mutants (MB-1 RH, MB-1 LH, MB-1-His RH, MB-1-His LH), a total of at least ten colonies were screened by dideoxy

pALTER-1: 5' -ACTCAAGCTTGCATGCCTGCAGGTCGACT -3'

MB-1-pALTER-1: 5' -ACTCAAGCTTAAGCTACGCCTTTTTGCAT -3'

MB-1-His-pALTER-1: 5' -ACTCAAGCTTAGTGGTGGTGGTGGTGGTGA -3'

Figure 4.7 Sequences of Mutagenesis Plasmids. Wild-type pALTER-1 and the mutagenesis plasmids were sequenced on the non-coding strand using the SP6 sequencing primer which annealed just downstream of the *Hin*dIII site. The pALTER-1 is the wild-type sequence. The end of the MB-1 gene is underlined in the MB-1-pALTER-1 plasmid. The MB-1-His gene is underlined in the MB-1-His-pALTER-1 plasmid, showing the GTG rich region which corresponds to the 6 His residues at the end of the protein.

sequencing. Those from the first round of mutagenesis were sequenced using only the SP6 primer. This was to check for insertion of the cysteine codon in the helix IV region. Although four colonies were checked for each mutant, none housed the GCA mutation. Consequently, the oligo concentrations were double-checked and the mutagenesis reactions were performed again. At least six colonies for each mutant were then screened using both SP6 and M13 universal (-20) (17mer) primers. This meant that twenty-four plasmid DNA samples were extracted and purified and sequenced at each end of the gene. All in all, the screening process took longer than expected and usually revealed that no mutations had occurred. Even when the reverse reaction showed no mutation, the forward still had to be performed in case a cysteine codon had been introduced at the other end.

4.3.3 Mutation Status

Despite the fact that the efficiency of mutagenesis was low, some mutations were confirmed. These results are summarized in **Table I**. For the right-handed MB-1 mutant, clone 4 had the C-terminal mutation, C91. The sequencing gel is shown in **figure 4.8**. Bands appear across all four lanes for this clone as well as the others that are seen on the same gel. These samples were sequenced at the same time using the Sequenase Kit. Because they were left for too long during both the radiolabelling and the elongation/termination steps, the enzyme likely paused, synthesizing DNA fragments of the same length in all four tubes. Such fragments have the same mobility, and lend to the appearance of bands in all four lanes of the gel. Another possibility is that the alkaline-denaturation procedure did not eliminate the secondary structure of the DNA in that particular region. This would have caused a similar pause as the enzyme reached the

Table I: Summary of Mutant Status. The status of each disulfide mutant is shown. Those clones containing the desired mutations are listed.

Right-Handed	**MB-1**	**MB-1-His**
N-terminal (C10)	None	None
C-terminal (C91)	Clone 4	None
Left-Handed	**MB-1**	**MB-1-His**
N-terminal (C13)	None	Clone 10
C-terminal (C87)	Clone 6 ?	Clone 10 ?

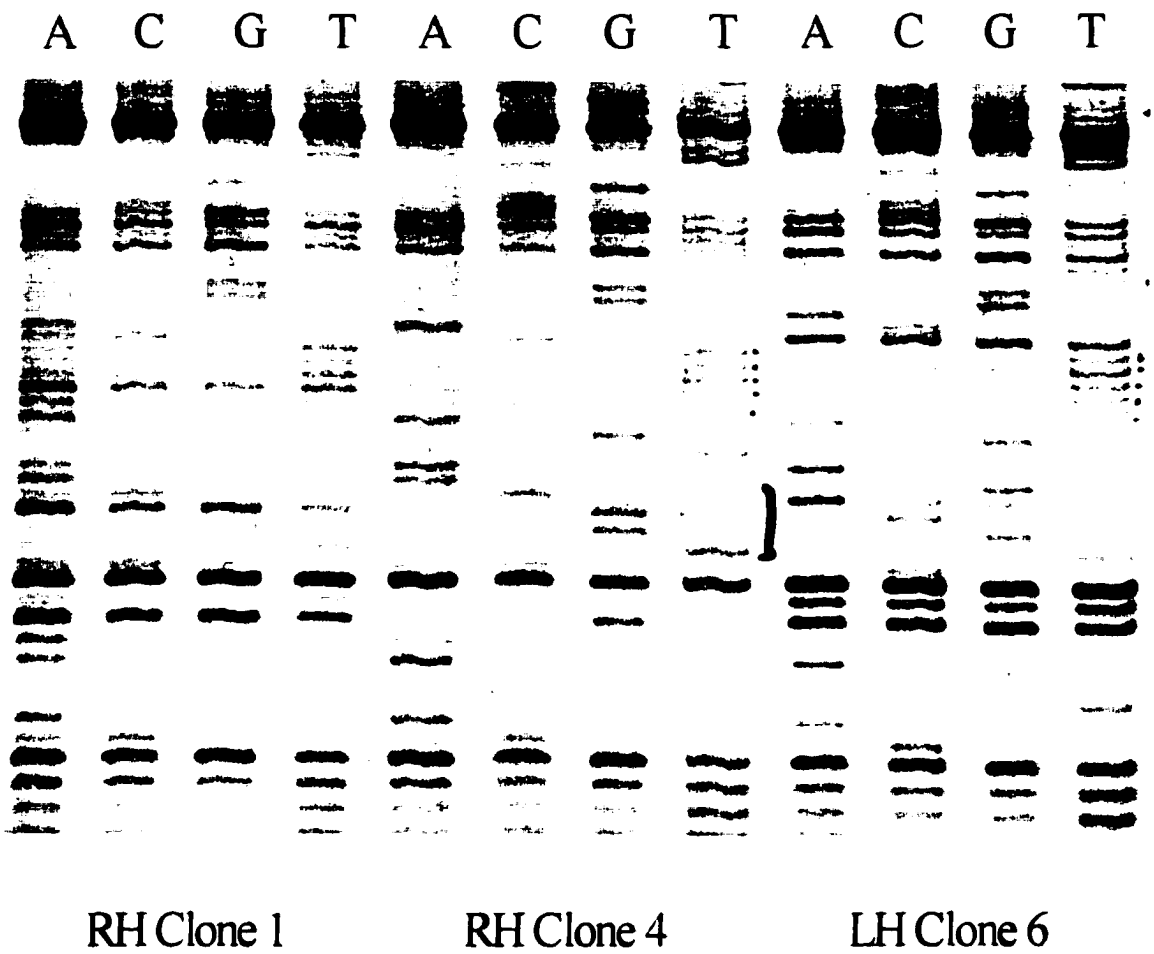


Figure 4.8 Sequencing Gel of MB-1 Clones. These three MB-1 clones were sequenced using the Sequenase Kit and the SP6 sequencing primer. The primer anneals just downstream of the *Hin*dIII site and sequences the non-coding strand 5' to 3'. Frequent artifacts were observed for all three samples. Since they were left for too long during the radiolabelling and elongation/termination reactions, the enzyme paused synthesizing fragments of equal size in all four reaction tubes. Despite this presence of bands in all four lanes, the mutation is obvious for RH Clone 4. Reading in the region marked by the bracket, the sequence is TGGCAA with the CAG coding for leucine replaced by the GCA coding for cysteine. Results for RH Clone 1 are not as clear, but it appears as though LH Clone 6 also has a cysteine mutation. This mutation occurs just above the 5 Ts, yet the pause in this region makes it difficult to confirm the first base of the mutation. The sequence should read TTTTTGCA. Although the pause resulted in a band in all four lanes, this often corresponds to a G in the sequence. However, the clone will have to be sequenced again to ensure that the mutation is correct.

area of strong secondary structure. In any case, the presence of the mutation is obvious. Reading the non-coding strand from 5' to 3', the CAG coding for leucine in MB-1 was replaced by GCA coding for cysteine. Despite the fact that the mutation is present, the sequence in the pause regions should be verified by sequencing on the opposite strand. **Figure 4.8** also shows a left-handed MB-1 clone (#6). It appears as though this clone may have the C-terminal mutation, but a pause artifact is located exactly where the mutation should start. Although the pause resulted in a band in all four lanes, this often corresponds to a G in the sequence. Also, since the 'G' band seems slightly darker than the A, C, and T bands, it is probable that the mutation has been properly incorporated. However, this clone will also have to be sequenced again to ensure that both the mutation and the region around it are as they should be.

As mentioned in **Table I**, there were no right-handed MB-1-His mutations found. One of the left-handed clones did reveal mutations, though. **Figure 4.9** shows the sequencing results for clone 10 in the N-terminal region. Reading from 5' to 3' on the coding strand, the CTG coding for the leucine in the control plasmid was replaced by the TGC coding for cysteine. Results for the reverse sequencing of this clone are shown in **figure 4.10**. Just above the 5 compressed T bands, there should either be a GCA for the mutation or a CAT for the wild-type. A mutation has obviously occurred, and a band may be present in the G lane, but is faint due to compression. Unfortunately, this is not enough to confirm the incorporation of the mutation. Another round of sequencing would be required to confirm this fact.

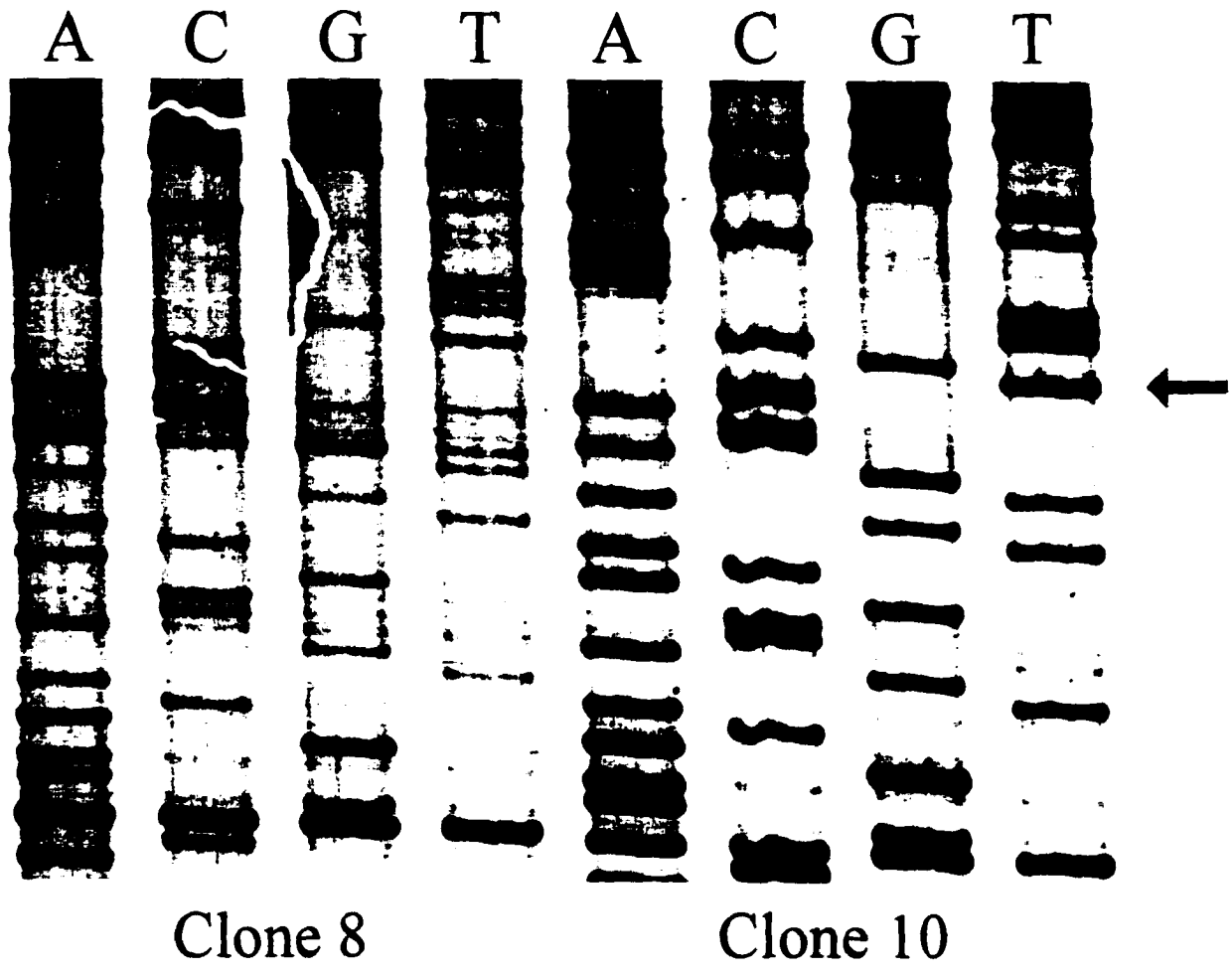


Figure 4.9 Sequencing Gel of Left-handed His Clones. The clones were sequenced in the N-terminal region using the universal (-20) (17mer) sequencing primer. Clone 8 is wild-type, but Clone 10 houses the cysteine mutation. Reading the coding strand from 5' to 3', the mutant sequence is TGC as opposed to the CTG coding for leucine 91.

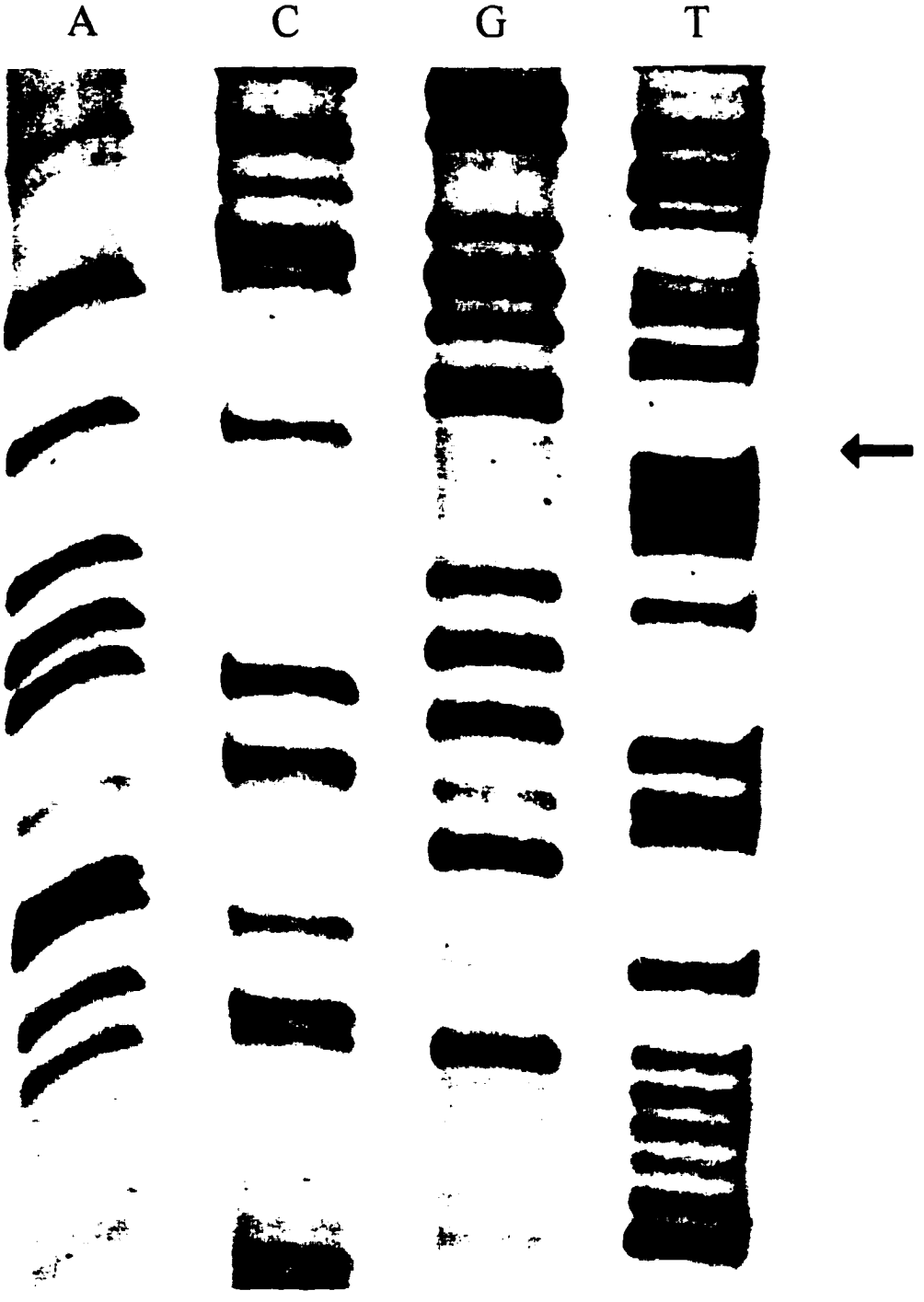


Figure 4.10 Sequencing Gel of His Clone 10 in the C-terminal Region. The C-terminal region was sequenced using the SP6 sequencing primer in order to determine whether or not the cysteine 87 mutation had been incorporated. Just above the 5 T bands, there should be either a GCA coding for the cysteine or a CAT coding for the wild-type methionine. A mutation has obviously occurred, and a G band may be present, but is faint due to compression; another round of sequencing would be required to confirm this fact.

4.4 DISCUSSION

4.4.1 Importance of the Intramolecular Disulfide

As mentioned earlier, the presence of disulfides are known to confer resistance against rumen proteases. Several researchers have established this fact by exposing disulfide-containing proteins to rumen proteases (using either *in situ* or *in vitro* techniques), both in the presence and in the absence of reducing agents. With their disulfide bonds intact, proteins like bovine serum albumin (BSA), ribonuclease A (RNAse A), rice prolamin, maize zein, chicken egg ovalbumin, pea albumin 1, sunflower albumin 8, and pumpkin trypsin inhibitor 1 were highly resistant to rumen degradation (Hancock *et al.*, 1994; McNabb *et al.*, 1994; Spencer *et al.*, 1988; Mahadevan *et al.*, 1980). However, such proteins were rendered up to 30 times as susceptible to degradation upon addition of β-mercaptoethanol. Since this compound reduces disulfide bonds, the proteolytic stability of the native proteins was attributed to disulfide content. Furthermore, for many of these proteins, the degree of resistance was proportional to the number of disulfide bonds present. BSA, for example, contains 16 disulfide bonds and is among the most stable of the proteins listed (Mahadevan *et al.*, 1987). It has been shown to resist rumen degradation for up to 16 hours while proteins like casein and vicilin, which do not contain disulfide bonds, were degraded within an hour (Hancock *et al.*, 1994; McNabb *et al.*, 1994; Spencer *et al.*, 1988).

The ability of intramolecular disulfides to enhance the proteolytic stability of a protein is due, for the most part, to the high conformational stability associated with their presence. By holding the protein tightly in a fixed conformation, less protease targets are available to the enzymes. For MB-1, the conformational stability at rumen temperature is

extremely low. Half of the protein population is unfolded, exposing regions susceptible to degradation. The introduction of an intramolecular disulfide between helices I and IV should, in theory, hold the protein together at this temperature. The bond would therefore be capable of enhancing the proteolytic stability of MB-1 and increasing its lifetime in the rumen. That is, provided it does not disturb the stereochemistry of the residues around it.

Nature is perfectly capable of engineering a disulfide bond so as to maximize the stability of a protein. It can be a very difficult task for researchers, however. As seen in the case of T4 lysozyme discussed earlier, many studies have met with success. Yet, this work dealt with the introduction of disulfide bonds into a natural protein of known tertiary structure. MB-1 was designed *de novo*, meaning that its sequence was created by humans to be unique. It therefore does not resemble that of any known protein. Although MB-1 was designed to fold like natural α-helical bundles, its tertiary structure has not been well-defined. As a result, 3D modelling software could not be used either to optimize the location of the disulfide, or even to determine the connectivity between helices. In their work designing the bundle "Felix", Hecht *et al.* (1990) included two cysteine residues so as to confirm the left-handed connectivity of the protein and confer higher stability to the protein. The fact that the disulfide formed was evidence that the protein had folded as per design. Its presence even promoted a slight increase in the helicity of the protein. However, despite its well-folded helical conformation under native conditions, the bundle protein showed only marginal stability when exposed to denaturants. These results indicate just how hard it can be to engineer a stabilizing disulfide bond within a protein. The positions chosen for MB-1 disulfides may indeed be

successful in producing a more stable mutant, but it must be understood that such a feat may require several attempts.

4.4.2 Mutagenesis Conditions Not Optimal: Suggestions for Future Trials

Construction of the mutagenesis plasmids MB-1-pALTER-1 and MB-1-His-pALTER-1 was rather simple, as confirmed by dideoxy sequencing. The introduction of simultaneous cysteine mutations, on the other hand, was not. Despite the claims made by the company, the efficiency of mutant production was not very high, certainly not greater than 50%. This was likely due to human error, rendering the conditions less than optimal during the reactions. For one thing, the amount of oligonucleotide used at each step in the procedure was not exact. Often, quantitation by absorption spectroscopy leads to under or overestimation of the concentration of DNA in a sample. Unfortunately, this could have led to improper annealing ratios and as a result, the production of more plasmids having ampicillin resistance and no mutation. Another problem that was encountered was the high transformation efficiency of the mutant plasmids into the final host cells. Normally, this would not be considered negative, but when the number of colonies is higher, the chance of finding a given mutation is decreased. Less plasmid was transformed during the second mutagenesis round, yet the percentage of colonies housing the double mutations was still low. In view of this, if the double mutagenesis procedure is to be used for future trials, the screening process will have to be expanded. Often times, the introduction of a mutation results in the insertion or deletion of a recognition site for given restriction enzymes. By first examining the pattern of digestion for each clone, much less time and money would be wasted sequencing. Only those clones that

show the mutant pattern need be sequenced. Changes in restriction sites induced by each mutation have therefore been determined and can be seen in **Appendix B**.

Attempting to introduce the two cysteine mutations simultaneously decreased the probability of finding a plasmid with the desired mutations. Since the Promega Kit comes with the ampicillin knock-out and tetracycline repair oligos as well, successive rounds of mutagenesis could be performed by switching the selectivity of the plasmid back to the original conditions. A more efficient method would therefore be to introduce the mutations one at a time. In this manner, the screening process is simplified. With a mutagenesis rate of 60-90%, the sequencing of 5 colonies should give greater than 95% chance of finding the single, desired mutation. Only one reaction need be performed on each clone. Then once the first mutation has been found, the mutated plasmid can be used in a subsequent reaction to introduce the second mutation. Again, only one round of sequencing need be performed on the resulting clones. Such a technique will be employed to complete the mutants that have already been started. Once the sequences of the MB-1 and MB-1-His clones have been verified in the pause regions, the mutated plasmids can be used to introduce the second mutations. The next round of mutagenesis would therefore render mutant plasmids tetracycline resistant and ampicillin sensitive.

Using this alternative protocol, the creation of the cysteine mutations should be completed relatively soon by the new researchers in the lab. When produced, the left-handed and right-handed mutants will have to be characterized in terms of their ability to form the disulfide bond. The one that is capable of forming the bond will demonstrate the connectivity of MB-1, and then be used to determine the effect of an intramolecular disulfide on its rumen stability.

CHAPTER 5 - MUTANT 3 : TRYPTOPHAN

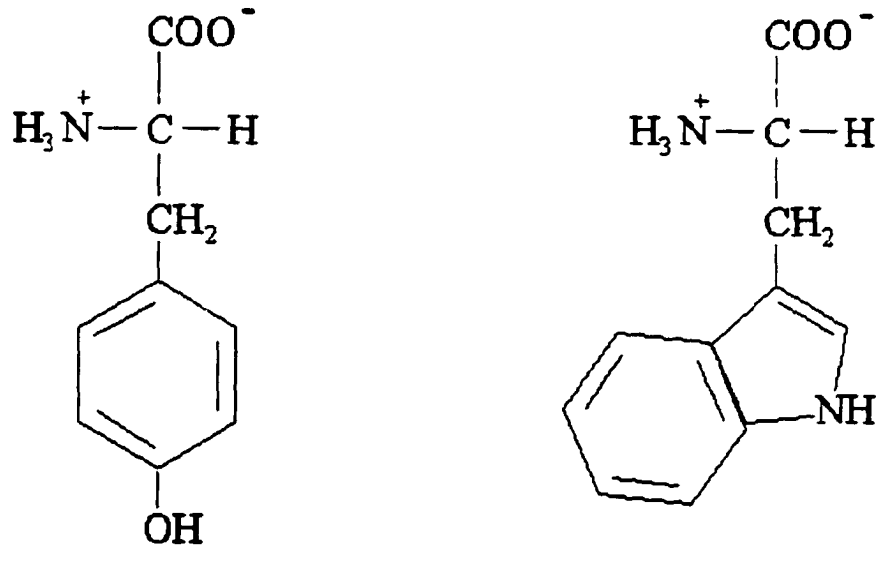
5.1 INTRODUCTION

5.1.1 Design of the Mutant

This chapter deals with the creation of a third mutant, MB-1-Trp. Designed to contain a tryptophan instead of tyrosine, the new mutant is to serve three purposes: 1) to increase the hydrophobicity of the bundle core and possibly enhance the rumen stability of the protein, 2) to provide a better means of monitoring the protein during purification, and 3) to expand the number of experimental options for studying the structure of MB-1.

As previously mentioned, the main driving force in protein folding is the formation of a well-packed hydrophobic core. Since the non-polar amino acids pack tightly together in this region, attempts to bury a polar side chain can be energetically costly. In fact, unless it is capable of forming a hydrogen bond to neutralize its polarity, the presence of a polar amino acid in the core can decrease the protein's conformational stability (Blaber *et al.*, 1993; Kohn and Hodges, 1998). MB-1 was designed to contain a single tyrosine within its core. Housing a free OH group, the tyrosine is polar in nature. By replacing it with a tryptophan, the polar OH group would be exchanged for a less polar side chain (**figure 5.1**; Zubay, 1998). As a result, the overall hydrophobicity of the core would be increased. If sufficient in magnitude, this change could improve the thermal stability of the protein and in turn enhance its resistance to rumen degradation (Handel *et al.*, 1993; Goldberg *et al.*, 1978; Parsell and Sauer, 1989; Liao, 1993).

The second reason for engineering a tryptophan into MB-1 was to confer greater spectral properties to the protein so that it could be easily monitored during purification. MB-1 is produced as a fusion with maltose binding protein (MBP) and purified using

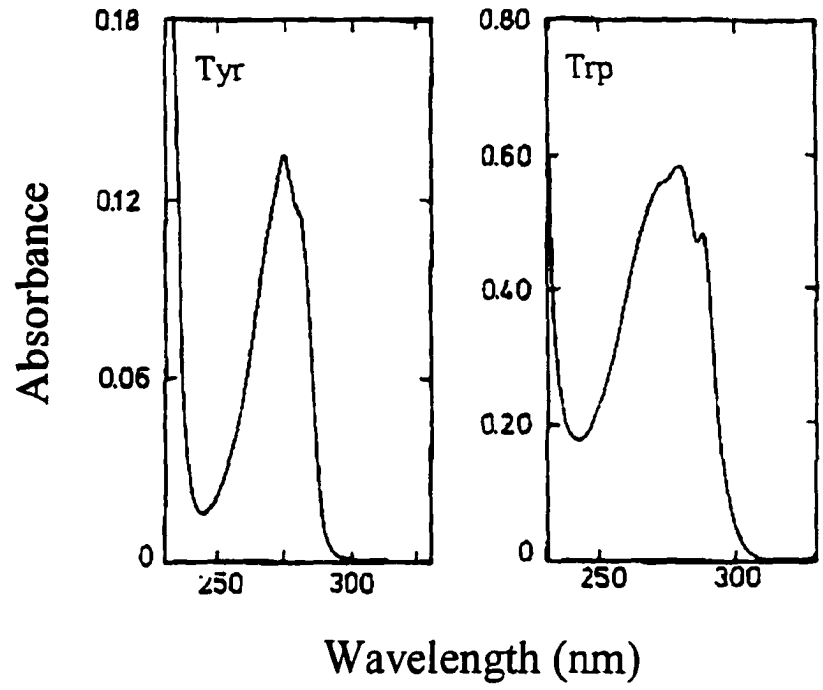


Tyrosine (Y) Tryptophan (W)

Figure 5.1 Tyrosine vs Tryptophan. The free OH group on tyrosine confers polarity to its side chain. By replacing the tyrosine 62 in MB-1 with a tryptophan, the polar OH will be removed from the core of the protein and exchanged for a less polar side chain. The difference in the polarity of the two amino acids is reflected by their hydrophobicity indices, where Y = 13.42 and W = 13.93 (Ragone *et al.*, 1989).

amylose and DEAE Sepharose columns (**section 2.1, p. 35**). As the chromatograms in **Appendix C** illustrate, monitoring the fusion protein as it comes off of the amylose column is simple. This is due to the presence of the MBP. It contains tryptophan which absorbs strongly at 280 nm. Elution of the MB-1-MBP protein is therefore detected as a distinct peak on the chromatogram. The fusion protein is cleaved before being loaded onto the DEAE column, however. Since MBP is retained on the column, MB-1 passes through alone and is diluted in the process. Its tyrosine signal is so weak that the UV monitor cannot detect it. With no peaks appearing on the chromatogram, the fractions collected from the column must be checked for protein using either a colorimetric assay or SDS-PAGE. These latter steps could be omitted if MB-1 contained a tryptophan. Since it has four times the absorption at 280 nm, the tryptophan residue would offer a detectable signal as the protein comes off of the DEAE column (Creighton, 1989).

Figure 5.2 shows the absorption and emission spectra for tyrosine and tryptophan. With its higher absorption and emission levels, the tryptophan offers a stronger signal than tyrosine. Not only is this signal of higher magnitude, but also it is very sensitive to small changes in the environment of the tryptophan (Engelhard and Evans, 1996; Eftink, 1994). As a result, proteins containing tryptophan can be subjected to a wider variety of experiments regarding protein structure. Such studies include time-resolved fluorescence and fluorescence anisotropy decay. These techniques can be used to deduce information about the tertiary and quaternary structure of the protein, e.g. whether or not it has one dominant conformation under physiological conditions (Handel *et al.*, 1993; Willis *et al.*, 1994; Hogue *et al.*, 1992). The introduction of a tryptophan into MB-1 would thus provide a means of gathering important structural information. Since the crystal structure of MB-1 has not yet been determined, this information is



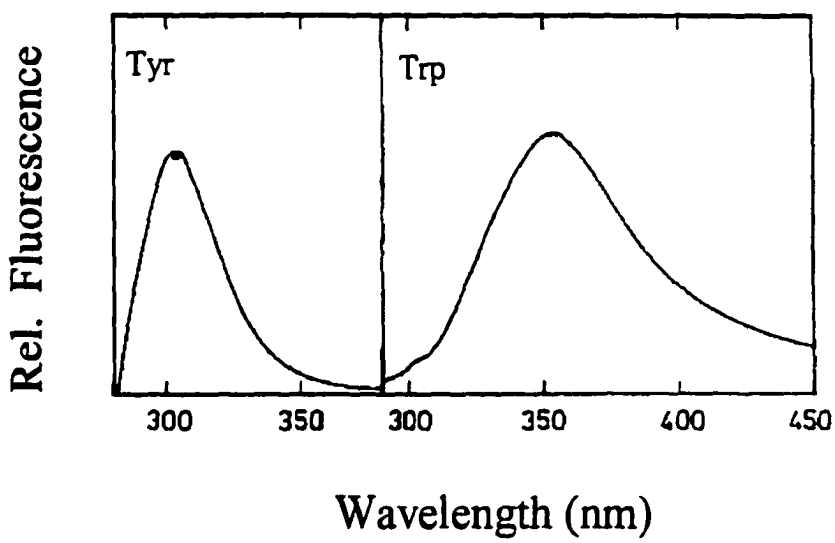


Figure 5.2 Absorption and Emission Spectra. As seen in panel A, tryptophan shows approximately four times the absorption as an equimolar amount of tyrosine. Fluorescence emission spectra are shown in panel B.

imperative for elucidating design problems. Other techniques like size exclusion chromatography and circular dichroism (CD) can be performed on MB-1 as it is. However, the signal of the lone tyrosine is so weak that a large amount of protein is required. Using the MB-1-Trp mutant, these studies could be carried out using less sample. At the same time, the smaller quantity of tryptophan would yield a better signal to noise ratio.

5.1.2 Protein Engineering

As described in **section 1.6.3, p. 30**, MB-1 has a tyrosine (Y62) buried in its core. Found in position 'd' of helix III, the bulky side chain of the tyrosine residue is accommodated by small, neighboring alanine residues (Beauregard *et al.*, 1995). **Figure 5.3** shows a top down view of this "knob in hole" design. The same approach will be used for the tryptophan mutant by simply exchanging the tyrosine residue for a tryptophan. Since a His tag will be used for future purification purposes, the trp mutation will be introduced into MB-1-His as well as MB-1.

5.1.3 Genetic Engineering

In order to create the MB-1-Trp plasmid, the TAC corresponding to tyrosine in MB-1 will be replaced by a TGG coding for tryptophan. The Promega *in vitro* Mutagenesis Kit will be employed as described in **section 4.2.2, p. 99**. However, rather than annealing the two cysteine oligos, only a tryptophan oligo will be annealed to the mutagenesis vectors. The sequence for the mutagenic Trp oligo is shown in **figure 5.4**.

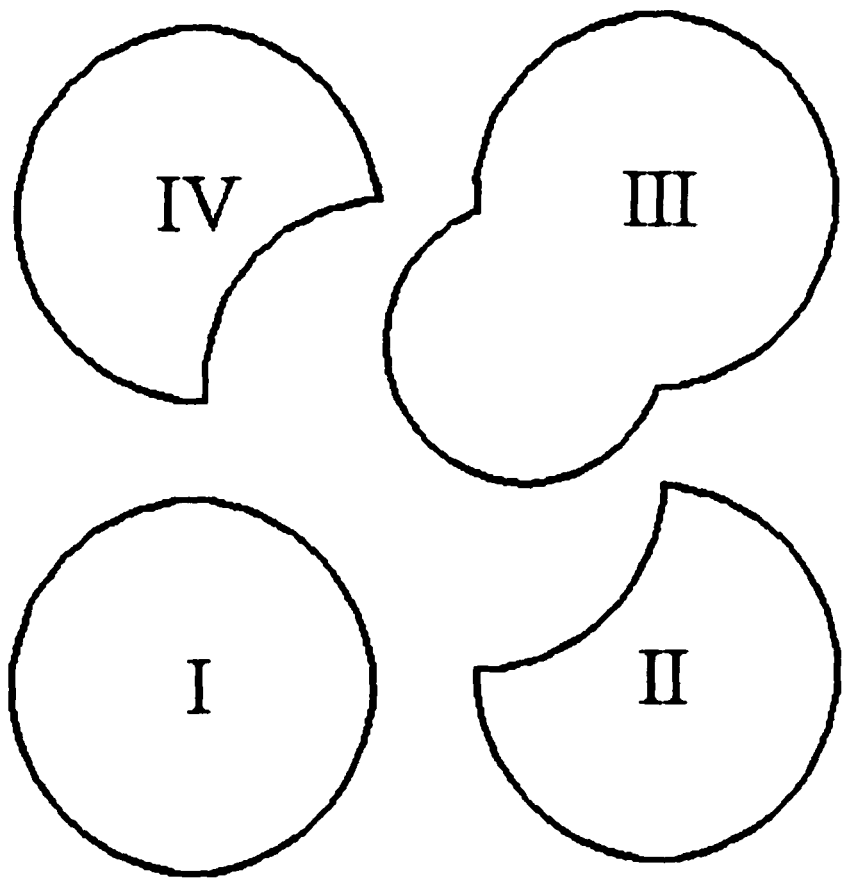


Figure 5.3 "Knob in Hole" Design Approach. The large side chain of tyrosine in the core of MB-1 is accommodated by small alanine residues. Shown as the protruding arc from helix III, the tyrosine is located in position 'd' of the heptad. Alanines in helices II (A35), III (A59, position 'a') and IV (A84) leave room for the bulky aromatic ring. MB-1-Trp will be created by simply replacing the tyrosine with a tryptophan.

M A T T Y F K T M

wt: 5' - ATG GCC ACT ACG TAC TTC AAA ACG ATG -3'

M A T T W F K T M

W62: 5' - ATG GCC ACT ACG TGG TTC AAA ACG ATG -3'

Figure 5.4 The Mutant Oligo. Both the wild-type and mutant sequences are depicted for the MB-1-Trp mutant. In order to create the mutant oligonucleotide W62, the TAC coding for tyrosine (Y) was replaced by a TGG coding for tryptophan (W). Since there are 12 wild-type nucleotides on either side of the mutation. the oligo should anneal strongly to the template DNA and allow for efficient incorporation of the mutation.

5.1.4 Chapter Objectives

Objectives of this chapter were:

- ❑ To introduce the tryptophan mutation into the gene
- ❑ To produce and purify the mutant protein by optimizing the original protocol
- ❑ To confirm the presence of the tryptophan
- ❑ To determine whether or not tryptophan is buried
- ❑ To compare the proteolytic stability of the mutant protein to that of MB-1
- ❑ To assess the conformational stability of the mutant protein relative to MB-1

5.2 MATERIALS AND METHODS

5.2.1 Oligo-Directed Mutagenesis

Introduction of the mutation onto MB-1 and MB-1-His mutagenesis vectors was performed using the same protocol as described in **section 4.2.2, p. 99.** The only difference was that the single Trp oligo was used instead of the two Cys oligos.

5.2.2 DNA Sequencing

Clones resulting from the mutagenesis procedure had to be screened for the Trp mutation. This was accomplished by dideoxy sequencing with the Sequenase Kit described in **section 4.2.3, p. 101**. The samples were sequenced using the SP6 sequencing primer, and run on gel for 2.5 to 3 h. All other steps were performed as described in **section 4.2.3**.

5.3 RESULTS

5.3.1 Mutagenesis Procedure

A total of four clones were sequenced, two from the MB-1 plate and two from the MB-1-His plate. The results are shown in **figure 5.5**. Since the non-coding strand was sequenced 5' to 3', the mutation is read as CCA. This corresponds to the TGG of the coding strand (**figure 5.4, p. 120**), which codes for tryptophan. The CCA should replace the GTA corresponding to the TAC codon for tyrosine. As shown in **figure 5.5**, three of the four clones have incorporated the Trp mutation.

5.3.2 Mutant Status

Figure 5.5 shows that the tryptophan mutations have indeed been incorporated into the plasmids. However, when the sequence of the entire mutant oligo was checked, it was discovered that a G nucleotide in the 3' region of the non-coding strand was missing in all three mutant clones. Comparison to MB-1 Clone 1 in **figure 5.5** reveals that the wild-type sequence should have been readable if it had been present.

5.4 DISCUSSION

5.4.1 Tryptophan: A Valuable Probe

Many research groups have introduced tryptophan or tryptophan analogs into proteins in order to probe their structural dynamics and/or interactions with other molecules (Handel *et al.*, 1993; Hogue *et al.*, 1992; Hogue *et al.*, 1996). This is most beneficial when only one tryptophan is incorporated as the results can be attributed to a specific region of the protein. A problem with signal interpretation arises when other tryptophanyl proteins are present in the solution or when the protein itself contains more

MB-1 Clone 1: 5' - CATCGAAAAGAA**GTA**CGTAGTGGCCAT -3'

MB-1 Clone 2: 5' - CATCGAAAAGAA**CCA**CGTAGTGCCAT - 3'

His Clone 1: 5' - CATCGAAAAGAA**CCA**CGTAGTGCCAT - 3'

His Clone 2: 5' - CATCGAAAAGAA**CCA**CGTAGTGCCAT - 3'

Figure 5.5 Sequences of MB-1 and MB-1-His Clones The plasmid sequences in the region of the tryptophan oligonucleotide are shown. The non-coding or minus strand that corresponds to the coding strand in **figure 5.4** was sequenced. All but MB-1 Clone 1 show the tryptophan mutation. Wild-type and mutant codons are bold and underlined, as the CCA coding for tryptophan in the mutant clones has replaced the GTA coding for tyrosine in wild-type MB-1 Clone 1. Reading further along the sequence, however, a G is missing. From the mutation on, the sequence should read **CCA**CGTAGTGGCC, but instead it reads **CCA**CGTAGTGCC.

than one tryptophan (Engelhard and Evans, 1996). Hogue *et al.* (1992) overcame this type of difficulty by introducing a Trp analog, 5-hydroxytryptophan, into their protein of interest. By replacing the Tyr 57 in oncomodulin with a 5HTrp analog, they were able to study the protein's interactions with antibodies using the unique fluorescent properties of the analog. Such a mutation is usually successful when the interior of the protein is loosely packed and able to accommodate the bulky tryptophan indole ring. If the protein is reasonably well-packed though, the introduction of such a large side chain would perturb the packing and destabilize the protein. This was indeed the case when Handel *et al.* (1993) attempted to replace a buried leucine in α_4 (a four-helix bundle) with a tryptophan. Denaturation with GuHCl revealed that the presence of the L6W mutation destabilized the protein by about 3.6 kcal / mol. These results suggest that the stability of MB-1 could be jeopardized by introducing the tryptophan mutation; however, the potential benefits of introducing the mutation far outweigh the risk of destabilizing the protein.

5.4.2 Re-Evaluation of the Mutagenesis Kit

The introduction of the tryptophan mutation into MB-1 and MB-1-His plasmids was very efficient using the Promega Kit. In this case, only one mutagenic oligo was used and three of the four clones that were sequenced housed the desired mutation. These results suggest that the intramolecular disulfide mutants would be created quickly and more efficiently if the cysteine mutations are introduced one at a time. The deletion observed in the Trp mutants was probably not due to the kit. Rather, it was likely the oligonucleotide that rendered the Trp clones of no use for protein production. An apparent single base deletion at the 5' end of the oligo resulted in a frame shift in the

MB-1 reading frame upon incorporation. The expression of such a gene would yield a protein of different sequence if any is produced at all. Consequently, the mutations will have to be introduced again using a new oligonucleotide.

CHAPTER 6 - GENERAL DISCUSSION AND SUGGESTIONS

6.1 GENERAL DISCUSSION

6.1.1 Why Modify Original MB-1

MB-1 is a *de novo* protein designed to contain a large percentage of the essential amino acids methionine, threonine, lysine, and leucine (57%). These amino acids were chosen since they are known to be limiting in high-producing ruminants like dairy cattle. With its lack of antigenicity and ease of production, MB-1 offers a safe, efficient source of these essential amino acids for the animal. Unlike most *de novo* proteins, MB-1 can be expressed *in vivo* to yield a folded protein with several features of native proteins. This ability to fold into a native-like structure makes MB-1 a viable candidate for expression in rumen microbes or transgenic plants. Whichever method of expression is used, the protein will be subjected to proteolytic degradation, either intracellularly or within the rumen environment. Although comparable to some natural plant proteins, MB-1's resistance to rumen-like proteases is rather weak. In order to make the *de novo* protein more appealing to potential industrial partners, its rumen stability must first be improved. If the protein can be made to pass through the rumen virtually unaltered in sequence, then it would provide an efficient means of delivering the EAAs to the body.

6.1.2 MB-1-Cys Dimer: The First of the MB-1 Mutants

With this goal in mind, three mutants of MB-1 have been designed. The first, MB-1-Cys Dimer, was created by linking two MB-1 units together with a disulfide bond. By decreasing the amount of surface area exposed by each monomer, the close apposition of MB-1 units rendered the dimer twice as resistant to rumen-like proteases.

However, despite the fact that it is more stable than MB-1, the new mutant may not be useful in terms of ruminant nutrition. The free sulfhydryl group of the cysteine residue is capable of binding many of the cysteine-containing proteins and some of the metals (Zn, Cu, Fe, etc.) required for normal bodily function. Whether found in the microorganisms, plants, or ruminant animals themselves, the protein may reduce their availability to the organism and induce biochemical stress. Addition of a signal sequence to the N-terminal of the protein may minimize such problems by directing it to other cellular compartments where the stress would be minimal. Such a sequence can be engineered to target the protein either to the periplasmic space in a rumen microbe for disulfide formation and secretion, or to the endoplasmic reticulum in a plant for synthesis, disulfide formation, and secretion (Creighton *et al.*, 1995; Lodish *et al.*, 1995).

6.1.3 The Intramolecular Disulfide Mutant

The second MB-1 mutant was designed to contain an intramolecular disulfide. However, problems were encountered during the mutagenesis procedure and only a few of the desired mutations have been created. Rather than introducing the two cysteine mutations one at a time, an attempt had been made to introduce them simultaneously. Unfortunately, the low efficiency led to difficulties in screening for mutations. A list of restriction site changes has thus been formulated to simplify the screening procedure in future experiments. Both a "left-handed" and a "right-handed" mutant were designed in order to elucidate the connectivity of MB-1. Only one will be capable of forming the disulfide bond, thus the connectivity of the protein will be determined by a matter of deduction. The mutant that forms the disulfide will then be evaluated in terms of conformational and proteolytic stability. If the bond is located in the appropriate

position, it should increase the thermal stability of the protein, rendering it folded at rumen temperature and less susceptible to proteolytic degradation.

6.1.4 MB-1-Trp

MB-1-Trp was the third mutant designed in this thesis. The Promega Mutagenesis Kit worked very efficiently for introducing the single tryptophan mutation into MB-1 and MB-1-His plasmids. This suggests that creation of the intramolecular disulfide mutants would be quicker and more efficient if the cysteine mutations are introduced one at a time. As for the MB-1-Trp and MB-1-His Trp mutants, there was a problem with the 5' end of the oligonucleotide. A new Trp oligo has been received, and mutagenesis is currently being performed. Introduction of the tryptophan into MB-1 should offer greater hydrophobicity to the core of the protein. If the core can accommodate the bulky indole ring, then the thermal and proteolytic stability of the protein may be enhanced. Whether or not the tryptophan enhances the rumen stability, it will offer greater spectral properties to the protein, allowing it to be easily monitored during production and exposed to a wider variety of experiments for structural characterization.

6.2 SUGGESTIONS FOR FUTURE PROJECTS

6.2.1 Exposed Cysteine: A Possible Solution for X-Ray Analysis

As previously mentioned, the exact structure of MB-1 has not yet been elucidated. The power of x-ray crystallography in elucidating structural information about a protein is second to no other technique. For most *de novo* proteins like MB-1, however, it is difficult to obtain protein crystals of suitable quality for x-ray diffraction analysis. A

possible solution to this problem is to bind heavy metals like mercuric salts to the protein prior to the experiment. These metals are known to ease phasing during structure elucidation. Binding metals to the protein may also promote better crystal growth (Wrede and Schneider, 1994; Friedman, 1973). MB-1, as it is, cannot bind heavy metals. As shown in this thesis work, the new cysteine mutant, MB-1-Cys, can. With the wealth of structural information that can be gathered from an x-ray analysis, it would be possible to fully evaluate the design of MB-1 and determine what is causing it to have such a low conformational and rumen stability. The 3D crystallographic data could then be used in modelling software to optimize the position of future mutations in MB-1.

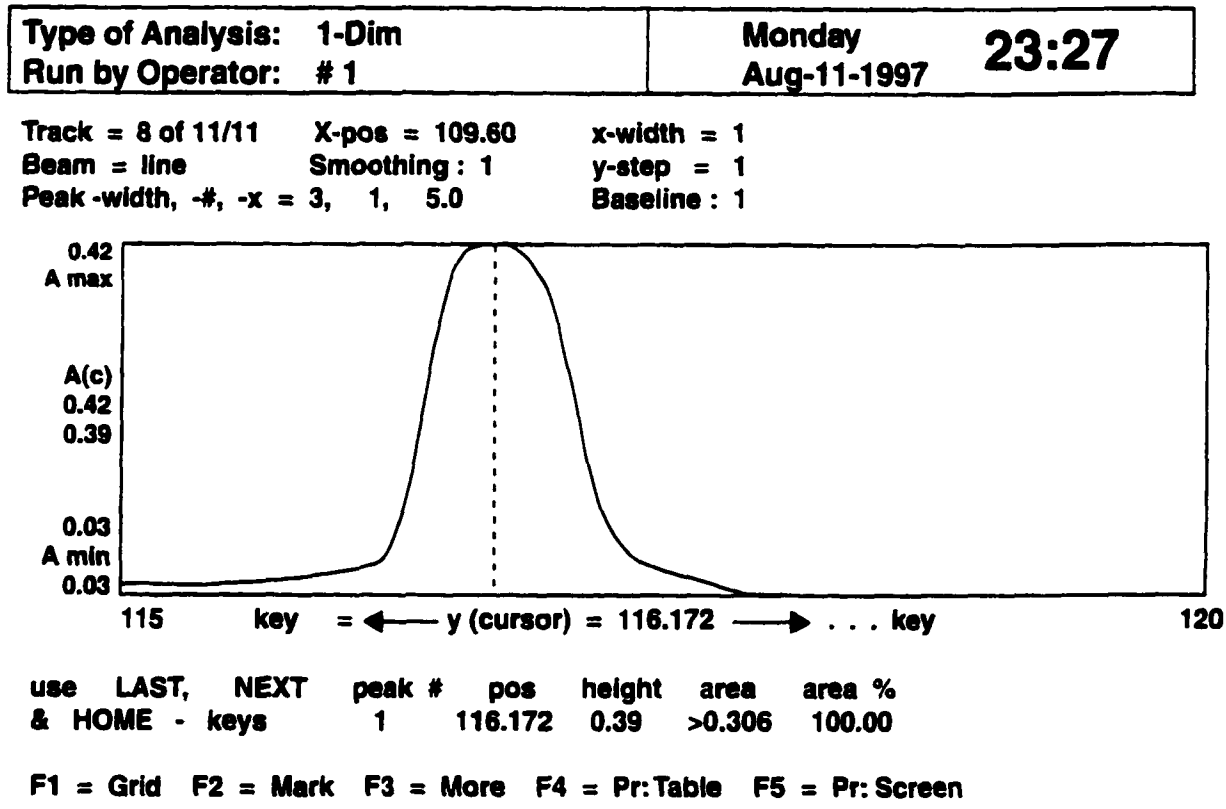
6.2.2 Implications from These Studies

Despite the fact that the MB-1-Cys mutant may not be used for nutritional purposes, the results for the MB-1-Cys Dimer suggest other modes of enhancing the rumen stability of MB-1. By linking MB-1 units together in the same genetic reading frame, and expressing it as a polymer, greater proteolytic resistance could be conferred to the bundle protein. Another possibility would be to introduce the intermolecular and intramolecular disulfide mutations at the same time. In this way, the stabilizing force of oligomerization (inter) would be added to that of enhanced conformational stability (intra). That is, provided the intramolecular disulfide is in the appropriate location.

The packing and hydrophobicity of the bundle core should be optimized in future to promote a more stable conformation at rumen temperature. Preliminary studies suggest that a difference in consensus residues exists between left-handed and right-handed bundles (Morrison, unpublished). Since a left-handed bundle (Rop) was used in the original design of MB-1, the sequence of the bundle core may require substantial

modification in order to enhance the thermal stability of the protein. Two points that certainly need to be addressed are the properties of borderline positions, 'e' and 'g', and the placement of β-branched residues like valine, isoleucine, and threonine. The preliminary results indicate that more hydrophobic residues should be placed in positions 'e' and 'g' rather than the abundance of polar and charged residues seen in MB-1 (T, Y, K^-, H^-). At the same time, the accommodation of a few β-branched residues seems favorable in positions 'a' and 'd'. This suggests that some of the threonines could be moved from borderline to core positions. These were only initial studies, however. Further research and statistical evaluation must be carried out before making any significant changes to the protein core. Perhaps the new MB-1-Cys and MB-1-Trp mutants will someday provide sufficient structural information to aid in this aspect of MB-1 design.

Appendix A: Typical Scan Using the LKB2222-020 UltroScan LX Laser Densitometer. The densitometer measures the absorbancy through the gel, showing peaks like this where a silver-stained band appears. Band intensity decreases as the protein is degraded. This is reflected by the size of the peak and thus the area under the curve. Using the area under the curve values, the amount of protein remaining at each given time interval was determined.

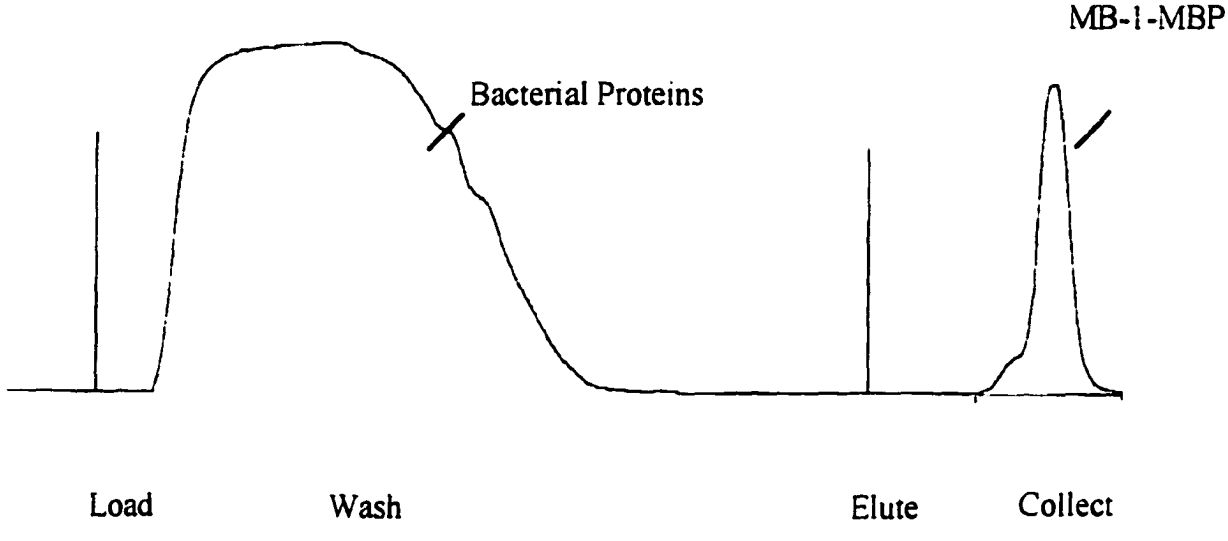


Appendix B: Restriction Site Changes Introduced by Given Mutations. The following table lists the number of restriction sites and sizes of fragments generated for wild-type and mutant MB-1 plasmids. MB-1-His plasmids show the same number of sites as listed for MB-1, but the His tag renders some of the fragments 18 bp larger. These are indicated by the ' * '.

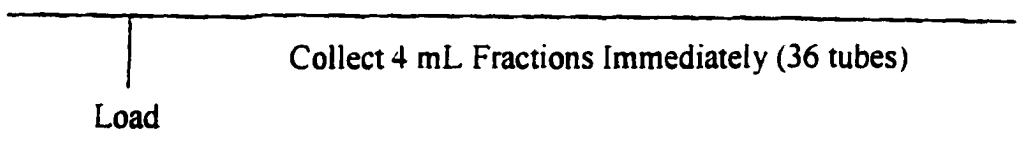
Right-Handed	Enzyme	Wild-Type	MB-1 Mutant
N-terminal (C10)	*Apa*LI	4 recognition sites giving fragments: 497, 498, 1242, *3754 bp	5 recognition sites giving fragments: 497, 498, 789, 1242, *2965 bp
C-terminal (C91)	*Pst*I	1 recognition site giving a fragment of *5991 bp	0 recognition sites thus only circular plasmid
Left-Handed	**Enzyme**	**Wild-Type**	**MB-1 Mutant**
N-terminal (C13)	*Bsp*MI	1 recognition site giving a fragment of *5991 bp	2 recognition sites giving fragments: *1730, 4261 bp
C-terminal (C87)	*Bsp*MI	1 recognition site giving a fragment of *5991 bp	2 recognition sites giving fragments: *1508, 4483 bp

Appendix C: Typical Chromatograms From Amylose and DEAE Sepharose Columns. The top panel shows a chromatogram obtained during affinity chromatography. After loading the sample onto the column, the bacterial proteins are washed through (large peak). Upon addition of maltose buffer, the fusion protein is eluted, showing a small, but distinct peak. Cleavage of the fusion protein results in separation of MB-1 from the tryptophan-containing maltose binding protein (MBP). When loaded onto the DEAE column, MBP is retained while MB-1 passes through. The tyrosine signal of MB-1 is so weak that it is not even detected. No peak is observed on the chromatogram, thus samples must be checked for protein using a colorimetric assay and/or SDS-PAGE.

From Amylose Column



From DEAE Column



REFERENCES

ARMENTANO LE, BERTICS SJ, DUCHARME GA. Response of lactating cows to methionine or methionine plus lysine added to high protein diets based on alfalfa and heated soybeans. J Dairy Sci 1997; 80; 1194-1199.

ASHES JR, GULATI SK, SCOTT TW. The role of rumen protected proteins and energy sources in the diet of ruminants. In: Ivan M, ed. Animal Science Research and Development: Moving Towards a New Century. CFAR Contribution no. 2321. Ottawa, ON: Ministry of Supplies and Services Canada (ISBN 0-662-23589-4), 1995: 177-185.

ASPLUND, JM (ed). Principles of Protein Nutrition of Ruminants. London: CRC Press, 1994.

ASSOUMANI MB, VEDEAU JL, SNIFFEN CJ. Refinement of an enzymatic method for estimating the theoretical degradability of proteins in feedstuffs for ruminants. Proc 41st Annu Mtg Eur Assoc Anim Prod, 1990.

BEAUREGARD M, DUPONT C, TEATHER RM, HEFFORD MA. Design, expression, and initial characterization of MB1, a *de novo* designed protein enriched in essential amino acids. Bio/Technology 1995; 13: 974-981.

BEAUREGARD M, HEFFORD MA, TEATHER RM. Use of monoclonal antibody beta-galactosidase for the detection of alpha peptide fusion. Bio/Techniques 1994; 16: 831-835.

BETZ SF, DEGRADO WF. Controlling topology and native-like behaviour of *de novo* designed peptides: design and characterization of antiparallel four stranded coiled coils. Biochemistry 1996; 35: 6955-6962.

BLABER M, LINDSTROM JD, GASSNER N, XU J, HEINZ DW, MATTHEWS BW. Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala → Ser and Val → Thr substitutions in T4 lysozyme. Biochemistry 1993; 32: 11363-11373.

BOEBEL KP, BAKER DH. Efficacy of methionine peptides as determined by chick bioassay. J Nutr 1982; 112: 1130-1132.

BORDO D, ARGOS P. Suggestions for "safe" residue substitutions in site-directed mutagenesis. J Mol Biol 1991; 217: 721-729.

BOYER PD. Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials. J Am Chem Soc 1954; 76: 4331-4337.

BRADSHAW RA. Proteins: Form and Function. Cambridge: Elsevier Science Publishers, 1990.

BRODERICK GA, CLAYTON MK. Rumen protein degradation rates estimated by non-linear regression analysis of Michaelis-Menten *in vitro* data. Brit J Nutr 1992; 67: 27-42.

BRODERICK GA. *In vitro* procedures for estimation of rates of ruminal protein degradation and proportions of protein escaping the rumen degradation. J Nutr 1978; 108: 181-190.

BRODERICK GA, KOWALCZYK T, SATTER LD. Milk production response to supplementation with encapsulated methionine per os or casein per abomasum. J Dairy Sci 1970; 53: 1714-1720.

BRYSON JW, BETZ SF, LU HS, SUICH DJ, ZHOU HX, O'NEIL KT, DEGRADO WF. Protein design: a hierarchic approach. Science 1995; 270: 935-941.

CAMPBELL MK. Biochemistry. Orlando, Florida: Saunders College Publishing, 1991.

CHALUPA W. Rumen bypass and protection of proteins and amino acids. J Dairy Sci 1975; 58: 1198-1218.

CHEFTEL JC, CUQ JL, LORIENT D. Amino Acids, Peptides and Protein. In: Fennema O, ed. Food Chemistry. New York: Marcel Dekker, Inc, 1985.

CHURCH DC. The Ruminant Animal: Digestive Physiology and Nutrition. Reston: Prentice Hall, 1988.

CLARK JH, KLUSMEYER TH, CAMERON MR. Symposium: Nitrogen Metabolism and Amino Acid Nutrition in Dairy Cattle. Microbial protein synthesis and flows of nitrogen fractions to the duodenum of dairy cows. J Dairy Sci 1992; 75: 2304-2323.

COHEN C, PARRY DAD. α-helical coiled coils and bundles: how to design an α-helical protein. Prot Struct Func Gen 1990; 7: 1-15.

CREIGHTON TE, ZAPUN A, DARBY NJ. Mechanisms and catalysts of disulphide bond formation in proteins. TIBTECH 1995; 13: 18-23.

CREIGHTON TE. Protein Structure: a Practical Approach. Oxford: IRL Press, 1989.

CREIGHTON TE. BioEssays 1988; 8: 57-63.

CROOKER BA, SNIFFEN WH, HOOVER WH, JOHNSON LL. Solvents for soluble nitrogen measurements in feedstuffs. J Dairy Sci 1978; 61: 437-442.

CULL M, MCHENRY CS. Preparation of Extracts from Prokaryotes. Ch 12. In: Deutscher MP, ed. Guide to Protein Purification. Methods in Enzymology Vol 182. San Diego: Academic Press, 1990.

DAO-PIN S, SAUER U, NICHOLSON H, MATTHEWS BW. Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. Biochemistry 1991; 30: 7142-7153.

DAS G, HICKEY DR, MCLENDON G, SHERMAN F. Proc Natl Acad Sci USA 1989; 86: 496-499.

DAVIS BJ. Ann NY Acad Sci 1964; 121: 404-427.

DOEL MT, EATON M, DOOK EA, LEWIS H, PATEL T, CAREY NH. The expression in *E. coli* of synthetic repeating polymeric genes coding for poly (L-aspartyl-L-phenylalanine). Nucl Acids Res 1980; 8: 4575-4592.

DYER JM, NALSON JW, MURAI N. Strategies for selecting mutation sites for methionine enhancement in the bean seed storage phaseolin. J Prot Chem 1993; 12: 545-560.

ECKERT R, RANDALL D, AUGUSTINE G. Animal Physiology: Mechanism and Adaptations. New York: W.H. Freeman and Company, 1988.

EFTINK MR. The use of fluorescence methods to monitor unfolding transitions in proteins. Biophys J; 66: 482-501.

ENGELHARD M, EVANS PA. Experimental investigation of sidechain interactions in early folding intermediates. Folding & Design 1996; 1: R31-R37.

FALDET MA, SATTER LD. Feeding heat-treated full fat soybeans to cows in early lactation. J Dairy Sci 1991; 74: 3047-3054.

FRIEDMAN M. The Chemistry and Biochemistry of the Sulfhydryl Group in Amino Acids, Peptides, and Proteins. Oxford: Permagon Press, 1973.

GALLAGHER SR. One-Dimensional SDS Gel Electrophoresis of Proteins. Ch 10, p.1.1-1.34. In: Coligan JE, Dunn BM, Ploegh HL, Speicher DW, Wingfield PT, eds. Current Protocols in Protein Science. John Wiley & Sons, 1995.

GARFIN DE. One-Dimensional Gel Electrophoresis. Ch 33. In: Deutscher MP, ed. Guide to Protein Purification. Methods in Enzymology Vol 182. San Diego: Academic Press, 1990.

GOLDBERG AL, GOFF SA. The selective degradation of abnormal proteins in bacteria. In: Reznikoff W, Gold L, eds. Maximizing Gene Expression. Stoneham, MA: Butterworths, 1986: 287-314.

GOLDBERG AL, KOWIT JE, ETLINGER J, KLEMES Y. Selective degradation of abnormal protein in animal and bacterial cells. In: Segal HL, Doyle DJ, eds. Protein turnover and lysosome function. New York: Academic Press, 1978.

GRADDIS TJ, MYSZKA DG, CHAIKEN IM. Controlled formation of model homo- and heterodimer coiled coil polypeptides. Biochemistry 1993; 32: 12664-12671.

GRUNDY J, MORRISON JJ, MACCALLUM JD, WIRTANEN L, BEAUREGARD M. Crystallization and stabilization of MB-1, a *de novo* designed protein for optimized feeding technology. J Biotech 1998; 63: 9-15.

HAGEN CE, WARREN GJ. Lethality of palindromic DNA and its use in selection of recombinant plasmids. Gene 1982; 19: 147-151.

HAMES BD. An Introduction to Polyacrylamide Gel Electrophoresis. Ch1. In: Hames BD, Rickwood D, eds. Gel Electrophoresis of Proteins: A Practical Approach. Oxford: IRL Press, 1981.

HANCOCK KR, EALING PM, WHITE DWR. Identification of sulphur-rich proteins which resist rumen degradation and are hydrolysed rapidly by intestinal proteases. Brit J Nutr 1994; 72: 855-863.

HANDEL TM, WILLIAMS SA, MENYHARD D, DEGRADO WF. Introduction of a Trp residue into α_4 as a probe of dynamics. J Am Chem Soc 1993; 115: 4457-4460.

HANDEL TM, DEGRADO WF. *De novo* design of a Zn^{2+}-binding protein. J Am Chem Soc 1990; 112: 6710-6711.

HECHT MH, RICHARDSON JS, RICHARDSON DC, OGDEN RC. *De novo* design, expression, and characterization of Felix: a four-helix bundle protein of native-like sequence. Science 1990; 24: 884-891.

HILL CP, HANDESTON DH, WESSON L, DEGRADO WF, EISENBERG D. Crystal structure of α1: Implications for protein design. Science 1990; 249: 543-546.

HOGUE CWV, DOUBLIÉ S, XUE H, WONG JT, CARTER Jr. CW, SZABO AG. A concerted tryptophanyl-adenylate-dependent conformational change in *Bacillus subtilis* tryptophanyl-tRNA synthetase revealed by the fluorescence of Trp92. J Mol Biol 1996; 260: 446-466.

HOGUE CWV, RASQUINHA I, SZABO AG, MACMANUS JP. A new intrinsic fluorescent probe for proteins: Biosynthetic incorporation of 5-hydroxytryptophan into oncomodulin. Fed Eur Biochem Soc 1992; 310: 269-272.

HOLMES CW, WILSON GF. Milk Production From Pasture. New Zealand: Butterworths, 1984.

HUANG XL, CATIGNANI GL, SWAISGOOD HE. Relative structural stabilities of α-lactoglobulins A and B as determined by proteolytic susceptibility and differential scanning. J Agr Food Chem 1994, 42: 1276-1280.

HUNGATE RE. The Rumen and Its Microbes. New York, NY: Academic Press, 1966.

HUTJENS MF, OVERTON TR, BRAND A. Monitoring Milk Production: Optimizing Rumen Digestion in the Dairy Cow. Ch 4, part 1. In: Brand A, Noordhuizen JPTM, Schukken YH, eds. Herd Health and Production Management in Dairy Practice. The Netherlands: Wageningen Pers, 1996.

IMOTO T, FUKUDA KI, YAGISHITA K. A study of the native-denatured (N=D) transition in lysozyme I. Detection of the transition by product analyses of protease digests. Biochim Biophys Acta 1974, 336: 264-269.

JAENICKE R. Stability and folding of ultrastable proteins: eye lens crystallins and enzymes from thermophiles. FASEB J 1996; 10: 84-92.

JAYNES JM, LANGRIDGE P, ANDERSON K, BOND C, SNADS D, NEWMAN CW, NEWMAN R. Construction and expression of synthetic DNA fragments coding for polypeptides with elevated levels of amino acids. Appl Microbiol Biotechnol 1985; 21: 200-205.

KATZ JJ. Long wavelength chlorophyll. Spectrum 1994; 7: 1-9.

KING KJ, HUBER JT, SADIK M, BERGEN WG, GRANT AL, KING VL. Influence of dietary protein sources on the amino acid profile available for digestion and metabolism in lactating cows. J Dairy Sci 1990; 73: 3208-3216.

KOHN WD, HODGES RS. *De novo* design of α-helical coiled coils and bundles: models for the development of protein-design principles. TIBTECH 1998; 16: 379-389.

KRAUSE DO, RUSSELL JB. Symposium: Ruminal Microbiology. How many ruminal bacteria are there? J Dairy Sci 1996; 79: 1467-1475.

KRISHNAMOORTHY U, SNIFFEN CJ, STERN MD, VAN SOEST PJ. Evaluation of a mathematical model of rumen digestion and an *in vitro* simulation of rumen proteolysis to estimate the rumen-undegraded nitrogen content of feedstuffs. Brit J Nutr 1983; 50: 555-568.

KWON WS, DA SILVA NA, KELLIS JT Jr. Relationship between thermal stability, degradation rate and expression yield of barnase variants in the periplasm of *Escherichia coli*. Prot Eng 1996; 9: 1197-1202.

LAKOWICZ JR. Principles of fluorescence spectroscopy. New York: Plenum Press, 1983: 258-297

LARSON BL (ed). Lactation. Iowa: Iowa State University Press, 1985.

LIAO HH. Thermostable mutants of kanamycin nucleotidyltransferase are also more stable to proteinase K, urea, detergents, and water-miscible organic solvents. Enzyme Microb Technol 1993, 15: 286-292.

LUCHINI ND, BRODERICK GA, COMBS DK. Characterization of the proteolytic activity of commercial proteases and strained ruminal fluid. J Anim Sci 1996; 74: 685-692.

MACCALLUM J, HEFFORD MA, OMAR S, BEAUREGARD M. Prediction of folding stability and degradability of the *de novo* designed protein MB-1 in cow rumen. Appl Biochem Biotech 1997; 66: 83-93.

MACCALLUM J, WIRTANEN L, BEAUREGARD M. Fluorescence quenching studies of the *de novo* designed protein MB-1: Tyrosine 62 is buried in a compact hydrophobic core as per design (Poster). Int Conf Protein Folding and Design, National Institutes of Health, Bethesda, Maryland, 1996: April 23 - 26.

MAHADEVAN S, SAUER FD, ERFLE JD. Preparation of protease from mixed rumen microorganisms and its use for the *in vitro* determination of the degradability of true protein in feedstuffs. Can J Anim Sci 1987; 67: 55-64.

MAHADEVAN S, ERFLE JD, SAUER FD. Degradation of soluble and insoluble proteins by *Bacteroides amylophilus* protease and by rumen microorganisms. J Anim Sci 1980; 50: 723-728.

MALDAGUE P, KISHORE BK, LAMBRICHT P, IBRAHIM S, LAURENT G, TULKENS PM. Nephrotoxicity: Mechanisms, Early Diagnosis, and Therapeutic Management. Bach PJ, ed. New York: Marcel Dekker, Inc., 1991.

MATSUMURA M, BECKTEL WJ, LEVITT M, MATTHEWS BW. Stabilization of phage T4 lysozyme by engineered disulfide bonds. Proc Natl Acad Sci USA 1989a; 86: 6562-6566.

MATSUMURA M, SIGNOR G, MATTHEWS BW. Substantial increase of protein stability by multiple disulphide bonds. Nature 1989b; 342: 291-293.

MATTHEWS BW. Structural and genetic analysis of protein stability. Annu Rev Biochem 1993; 62: 139-160.

MCNABB WC, SPENCER D, HIGGINS TJ, BARRY TN. *In-vitro* rates of rumen proteolysis of ribulose-1,5-bisphophate carboxylase (Rubisco) from lucerne leaves, and of ovalbumin, vicilin and sunflower albumin 8 storage proteins. J Sci Food Agric 1994; 64: 53-61.

MCNIVEN MA, ROBINSON PH, MACLEOD JA. Evaluation of a new high protein variety of soybeans as a source of protein and energy for dairy cows. J Dairy Sci 1994; 77: 2605-2613.

MEHREZ AS, ORSKOV ER. A study of the artificial fibre bag technique for determining the digestibility of feeds in the rumen. J Agric Sci 1977; 88: 645-650.

MEPHAM TB. Physiology of Lactation. Philadelphia: Open University Press, 1986.

MERRIL CR. Gel-Staining Techniques. Ch 36. In: Deutscher MP, ed. Guide to Protein Purification. Methods in Enzymology Vol 182. San Diego: Academic Press, 1990.

MILLER WJ. Dairy Cattle Feeding and Nutrition. Orlando, Florida: Academic Press, 1979.

NOCEK JE, HERBEIN JH, POLAN CE. Total amino acid release rates of soluble and insoluble protein fractions of concentrate feedstuffs by *Streptomyces griseus*. J Dairy Sci 1983; 66: 1663-1667.

ORNSTEIN L. Ann NY Acad Sci 1964; 121: 321-349.

ORSKOV ER, REID GW, TAIR CAG. Effect of fish meal on the mobilization of body energy in dairy cows. Anim Prod 1987; 45: 345-348.

PAKULA AA, SAUER RT. Prot Struc Func Gen 1989; 5: 202-210.

PALIAKASIS CD, KOKKINIDIS M. Relationships between sequence and structure for the four-α-helix bundle tertiary motif in proteins. Prot Eng 1992; 5: 739-748.

PAPAS AM, SNIFFEN CJ, MUSCATO TV. Effectiveness of rumen-protected methionine for delivering methionine postruminally in dairy cows. J Dairy Sci 1984; 67: 545-552.

PARSELL DA, SAUER RT. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. J Biol Chem 1989; 264: 7590-7595.

PERMYAKOV EA. Luminescent spectroscopy of proteins. London: CRC Press, 1992.

PFEIL W, PRIVALOV PL. Thermodynamic investigations of proteins. Phys Chem 1976; 4: 33-40.

PICHARD G, VAN SOEST PJ. Solubility of ruminant feeds. Cornell Nutrition Conference for Feed Manufacturers, 1977: 91-98.

POOS-FLOYD M, KLOPFENSTEIN T, BRITTON RA. Evaluation of laboratory techniques for predicting ruminal protein degradation. J Dairy Sci 1985; 68: 829-839.

PRIVALOV PL. Cold denaturation of proteins. Crit Rev Biochem Mol Biol 1990; 25: 281-306.

PRIVALOV PL. Stability of proteins: small globular proteins. Adv Prot Chem 1979; 33: 167-241.

PURSER DB. Nitrogen metabolism in the rumen: microorganisms as a source of protein for the ruminant animal. J Anim Sci 1970; 30: 988-1001.

RAGONE R, FACCHIANO R, FACCHIANO A, FACCHIANO AM, COLONNA G. Flexibility plot of proteins. Prot Eng 1989; 2: 497-504.

REES AR, STERNBERG MJE, WETZEL R, (eds). Protein Engineering: a Practical Approach. Oxford: IRL Press, 1992.

REGAN L, CLARKE N. A tetrahedral zinc(II)-binding site introduced into a *de novo* designed protein. Biochemistry 1990; 29: 10878-10883.

ROBINSON PH, FREDEEN AH, CHALUPA W, JULIEN WE, SATO H, SUZUKI H. Rumen protected lysine and methionine for lactating dairy cows fed a diet designed to meet microbial and post-ruminal protein requirements. J Dairy Sci 1995; 78: 582-594.

ROE MB, CHASE LE, SNIFFEN CJ. Comparison of *in vitro* techniques to the *in situ* technique for estimation of ruminal degradation of protein. J Dairy Sci 1991; 74: 1632-1640.

ROGERS GL, BRYANT AM, MCLEAY LM. Silage and dairy cow production. NZ J Agric Res 1979; 22: 533-541.

ROSSOMANDO EF. Ion-Exchange Chromatography. Ch 24. In: Deutscher MP, ed. Guide to Protein Purification. Methods in Enzymology Vol 182. San Diego: Academic Press, 1990.

RULQUIN H, VERITE R, GUINARD G, PISULEWSKI PM. In: Ivan M, ed. Animal Science Research and Development: Moving Toward a New Century. CFAR Contribution no. 2321. Ottawa, ON: Ministry of Supplies and Services Canada (ISBN 0-662-23589-4), 1995: 143-160.

SAMBROOK J, FRITSCH EF, MANIATIS T. Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989.

SANGER F, NICKLEN S, COULSON AR. DNA sequencing with chain terminating inhibitors. Proc Nat Acad Sci USA 1977; 74: 5463-5467.

SCHÄGGER H, VON JAGOW G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 1987; 166: 368-379.

SCHEIN CH. Production of soluble recombinant proteins in bacteria. Bio/Technology 1989; 7: 1141-1149.

SCHWAB CG, SATTER LD, CLAY AB. Response of lactating dairy cows to abomasal infusion of amino acids. J Dairy Sci 1976; 59: 1254-1270.

SHIRLEY RL. Nitrogen and Energy Nutrition of Ruminants. Orlando, Florida: Academic Press, 1986.

SNIFFEN CJ, KRISHNAMOORTHY U, VAN SOEST PJ, MUSCATO TV, ROBINSON PH. Report on XV Conference on Rumen Function, Chicago, Illinois, 1979: 30-31.

SPENCER D, HIGGINS TJV, FREER M, DOVE H, COOMBE JB. Monitoring the fate of dietary proteins in rumen fluid using gel electrophoresis. Brit J Nutr 1988; 60: 241-247.

STELLWAGEN E, WILGUS H. Relationship of protein thermostability to accessible surface area. Nature 1978; 275: 342-343.

STERN MD, SANTOS KA, SATTER LD. Protein degradation in rumen and amino acid absorption in small intestine of lactating dairy cattle fed heat-treated whole soybeans. J Dairy Sci 1985; 68: 45-56.

STEWART CS. Some effects of phosphate and volatile fatty acid salts on the growth of rumen bacteria. J General Microbiol 1975; 89: 319-326.

STITGER D, DILL KA. Charge effects on folded and unfolded proteins. Biochemistry 1990; 29: 1262-1271.

STRYER L. Fluorescence spectroscopy of proteins. Science 1968; 162: 526-533.

TAMMINGA S. Protein degradation in the forestomachs of ruminants. J Anim Sci 1979; 49: 1615-1630.

TANNER GJ, MOORE AE, LARKIN PJ. Proanthocyanidins inhibit hydrolysis of leaf proteins by rumen microflora *in vitro*. Brit J Nutr 1994; 71: 947-958.

TINOCO Jr I, SAUER K, WANG JC. Physical Chemistry: Principles and Applications in Biological Sciences. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1995.

VAN SOEST PJ. Nutritional Ecology of the Ruminant. Oregon: O&B Books, 1983.

WILLIS KJ, NEUGEBAUER W, SIKORSKA M, SZABO AG. Probing α-helical secondary structure at a specific site in model peptides via restriction of tryptophan side-chain rotamer conformation. Biophys J 1994; 66: 1623-1630.

WOHLT JE, SNIFFEN CJ, HOOVER WH. Measurement of protein solubility in common feedstuffs. J Dairy Sci 1973; 56: 1052-1055.

WREDE P, SCHNEIDER G (eds). Concepts in Protein Engineering and Design. Berlin: Walter de Gruyter & Co, 1994.

YANG HJ, TSOU CL. Inactivation during denaturation of ribonuclease A by guanidinium chloride is accompanied by unfolding at the active site. Biochem J 1995, 305: 379-384.

ZHOU NE, KAY CM, HODGES RS. Disulfide bond contribution to protein stability: positional effects of substitution in the hydrophobic core of the two-stranded α-helical coiled-coil. Biochemistry 1993; 32: 3178-3187.

ZUBAY G.L. Biochemistry. 4th ed. Dubuque, IA: Wm. C. Brown Publishers, 1998.