EXPRESSION OF HUMAN METHIONINE AMINOPEPTIDASE IN *ESCHERICHIA COLI*

by

DAVID MADDEN

A THESIS

Presented to the Department of Biology
and the Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

June 1997

APPROVED: [redacted]

Todd Lowther, Ph.D.

An Abstract of the Thesis of

David Thomas Madden for the degree of Bachelor of Arts

in the Department of Biology to be taken June 1997

Title: EXPRESSION OF HUMAN METHIONINE AMINOPEPTIDASE IN *ESCHERICHIA COLI*

Approved: [redacted]

Todd Lowther, Ph.D.

Overexpressing protein in *Escherichia coli* is often used as a means of obtaining large quantities of protein. However, the overexpression of eukaryotic genes in *Escherichia coli* can become problematic because of the inherent differences in protein synthesis between eukaryotes and prokaryotes. While there are currently a number of strategies employed to overcome these problems, none can guarantee success. In order to study the human methionine aminopeptidase, which also acts to protect the α subunit of eukaryotic initiation factor 2 (eIF-2) from phosphorylation, I have performed a series of experiments to develop an overexpression system in *Escherichia coli*. Although every attempt to express the full length protein failed, overexpression of the last 323 amino acids, or catalytic domain, produced large quantities of insoluble protein.

This thesis is dedicated to Hannah Celeste, and the
long scientific career ahead of her.

ACKNOWLEDGMENTS

I would like to pay special thanks to Todd Lowther, who guided me through every obstacle, and taught me everything I presently understand about molecular biology and biochemistry. His patience and good humor were a blessing for my countless mistakes. Thanks also to Joan Wozniak, who skillfully introduced me to x-ray crystallography, and took the brunt of my ignorance. I thank Larry Weaver, a kind man and amazing teacher. His careful explanations gave me hope for tackling this bear of a science. Thanks especially to Brian Matthews, for always having good faith in my intentions and reminding me to follow my dreams. And to everyone in the Matthews lab, who understand that playing hard makes a nice complement to working hard. Thanks to my parents and family for continued love and support. And great thanks to Karna Dawson, my most cherished friend and confidant.

TABLE OF CONTENTS

Introduction

Today, research in biology is rapidly transforming as every discipline in the sciences discovers the common threads that unite all living things. As new studies continue to pour into the literature, researchers today understand that problems regarding genetic defects, infectious disease, or cancer may best be answered by studying fruit flies, yeast, or bacteria. Modern technology permits scientists to explore biological systems to the extent that experiments are only limited by imagination and creativity.

As this revolution in science continues, the direction of most studies is focused at the molecular scale. Scientists no longer take for granted that single mutations in a gene's DNA sequence can spell catastrophe for the cell in addition to seriously compromising an organism's ability to survive (sickle-cell anemia is the result of a singe base pair mutation in a hemoglobin gene). In order to prevent 16.4 million people from dying each year worldwide from infectious diseases, or help the million Americans currently infected with HIV, or give hope to the forty thousand women in the United States who are diagnosed with breast cancer each year, these diseases must be attacked at the molecular level.

One of the most important classes of molecules that demands our attention in order to understand disease pathology is proteins. Proteins are the most abundant molecules in a cell next to water, and their importance to the cell justifies this. These molecules not only provide a structural scaffolding for cells but also facilitate the thousands of chemical reactions that must take place in order for an organism to extract energy from nutrients, respond to changes in environment, grow and differentiate, and ultimately reproduce.

An important cellular function carried out by proteins involves the processing of newly made proteins. Methionine aminopeptidase is an enzyme that removes the amino-terminal methionine from these nascent polypeptides. The most common first codon in a nucleic acid sequence encoding protein is AUG, the codon that specifies the inclusion of the amino acid methionine. However, many proteins in the cell do not have methionine at the amino-terminus because during the maturation process, methionine aminopeptidase

cleaves this initiator methionine from the rest of the chain [1]. Methionine aminopeptidase activity has been shown to be essential in *Escherichia coli* and *Salmonella typhimurium*; removal of the gene encoding methionine aminopeptidase from these cells kills them [2, 3]. Knockout experiments like these demonstrate that correct processing of the amino-terminal methionine on nascent polypeptides is critical to cell health.

Other experiments have also demonstrated the importance of the integrity of the amino-terminus of the protein. Most notably, Gonda *et al.* demonstrated that the identity of the amino acid at the amino-terminus of a protein strongly influences the half-life of the protein; larger or charged amino acids at the amino-terminus signal rapid degradation, while smaller or uncharged amino acids at the amino-terminus confer slow degradation [4]. They called this pattern in protein degradation the "N-end Rule". As it turns out, the same amino acids that the N-end rule predicts will stabilize a protein when positioned at the amino-terminus are the same amino acids that also permit methionine removal via methionine aminopeptidase when positioned at the penultimate amino acid on the amino-terminus [5, 6]. As a result of these similar activities, it can be deduced that the function of methionine aminopeptidase is to remove the amino-terminal methionine only if the resultant protein will be stable to the degradation pathway inferred by the N-end rule. Since many regulatory proteins need to be degraded to preserve cell vitality, the fact that methionine aminopeptidase activity is essential to some organisms is explained.

The three dimensional crystal structure of the methionine aminopeptidase from *Escherichia coli* was determined in 1992 [7]. The structure revealed an internal pseudo two-fold symmetry and coordination to two cobalt ions (Co^{2+}). Using this unique internal symmetry and peptide sequence from the *E. coli* methionine aminopeptidase, Bazan *et al.* searched protein and DNA data bases for sequence similarities in other genes [8]. Besides finding other aminopeptidases like aminopeptidase P, prolidase, and creatinase, they also found a protein called p67. p67 is a eukaryotic initiation factor 2 (eIF-2)-associated 67-kDa protein, and at the time of Bazan's study no protease activity had been reported. Sometime

later, Li *et al.* demonstrated that the gene encoding the human p67 was in fact human methionine aminopeptidase [9].

It was the discovery of other distinct forms of methionine aminopeptidases that prompted the division of these enzymes into two classes. Class one methionine aminopeptidases (including those from *E. coli*, *S. typhimurium*, *B. subtilis*, and *S. cerevisiae*) have a smaller catalytic domain while class two methionine aminopeptidases (including those from *S. cerevisiae*, *M. fervidus*, pig, and human) have a catalytic domain that includes a sixty-five amino acid insertion [10]. Several of the methionine aminopeptidases have long amino-terminal extensions that may include zinc fingers (as in the yeast type I enzyme), or polybasic and polyacidic tracks of up to ten amino acids each (as in the human enzyme). The function of these extensions has been postulated to play a role in protein/nucleic acid or protein/protein interactions but the precise function remains unknown [11].

In addition to being a methionine aminopeptidase, human methionine aminopeptidase is also a translational regulator. The process in which genes transcribed into mRNA are read as instructions for protein synthesis is called translation, and in eukaryotes translation is heavily regulated. The most powerful regulation of translation in eukaryotes takes place during translation initiation, the same place where p67 functions [12]. One of the major steps required for translation initiation to occur is the binding of a GTP bound eukaryotic initiation factor called eIF-2 to methionyl-tRNA, which in turn binds to the 40S ribosomal subunit as a ternary complex (for reviews see [13-15]). After this step, the mRNA to be translated may also be loaded onto the 40S ribosomal subunit. Next, eIF-2 cleaves GTP, dissociating all other factors bound to the 40S with the exception of the methionyl-tRNA and the bound mRNA. As a result, the 60S ribosomal subunit may bind, and translation commences. After having been released, eIF-2 requires the nucleotide exchange factor eIF-2B in order to discharge the bound GDP and replace it with GTP.

Once this is accomplished, the eIF-2•GTP binary complex can bind to another methionyl-tRNA and initiate another round of protein synthesis.

However, when a cell senses that protein synthesis should be shut down, that is when confronted with high temperature, high hydrostatic pressure, or double-stranded RNA to name a few cases, the translation initiation machinery is slowed by the phosphorylation of the α subunit of eIF-2 [14]. The effect of eIF-2α phosphorylation is not immediate, as phosphorylated eIF-2 can initiate another single round of translation. However, the binding of phosphorylated eIF-2 to the nucleotide exchange factor eIF-2B has been estimated to be 400 times stronger than the same interaction with phosphate free eIF-2 [16]. As a result, any eIF-2B that comes into contact with phosphorylated eIF-2 is sequestered. Since eIF-2B is limiting, eIF-2B sequestration acts as a powerful translation repressor. The other activity of human methionine aminopeptidase, or p67, is to protect the eIF-2α subunit from being phosphorylated by the kinases that respond to environments that suppress translation. But eIF-2α kinases do not upregulate their phosphorylation activity when the cell is confronted with the conditions mentioned. Instead p67, or human methionine aminopeptidase, is degraded, and thereby permits the kinases access to eIF-2α [15]. As a result, eIF-2α is phosphorylated and translation is repressed.

To summarize, human methionine aminopeptidase acts not only to remove the methionine from nascent polypeptides but also acts as a key player in the regulation of protein synthesis. The fact that a single protein could have two such distinct activities is intriguing, making work on this protein all the more attractive. The human methionine aminopeptidase has such dramatic differences in its peptide sequence from the *Escherichia coli* enzyme that structural analysis of the human enzyme should prove fascinating. Specifically, the function of the sixty-five amino acid insertion in the catalytic domain of the human methionine aminopeptidase might be revealed in a structural analysis. In addition, information about how the polybasic and polyacidic tracks of amino acids on the amino-

terminal extension of the human enzyme give functionality to the protein could be deduced by the structural elements that contain them.

Currently there are a variety of techniques available for studying a protein's function or structure or both. However, most all of these techniques require that the protein of interest be very pure, and some techniques, such as x-ray crystallography, also require a large quantity of protein. Obtaining large quantities of protein that is also highly purified is difficult as the purification process always pins purity against quantity, obtaining one at the expense of the other. Traditional methods of protein purification, called raw extraction, rely on the relative abundance of a protein in the tissue of an organism from which the protein is to be extracted. For example, to isolate a protein involved in nerve cell signaling in mammals, several kilograms of cow brains are used from which only a few milligrams of pure protein can be isolated. Should some naive person attempt to conduct crystallographic studies of the same protein using raw extraction, the quantity of cow brains required could fill the better half of a large laboratory. But the structural information learned from crystallographic studies of proteins is so important to biology and chemistry that strategies for protein purification which are less resource intensive than raw extraction are aggressively sought.

One of the most popular of these alternate strategies is to cause bacteria, like *Escherichia coli*, to make a relatively large quantity of the protein of interest. *E. coli* serve as an ideal host for most overexpression purposes because it is easy to handle in the lab, having doubling times of just twenty minutes and well developed genetics. Using bacterial overexpression systems, it is common to see yields of hundreds of milligrams of protein from just four liters of *E. coli* culture. As a result of the high quantity of target protein in the bacteria, it is possible to obtain large quantities of very pure protein.

Overexpression in bacteria requires that the gene of interest has been cloned from the original organism and also that the gene can be positioned into a circular DNA molecule called a plasmid. On this plasmid are other genes coding for repressors, promoters,

terminators, and enzymes that permit resistance to antibiotics; all of these genes are required for overexpression to work properly. This plasmid can be transformed into an *E. coli* cell, after which the plasmid is replicated by the cells machinery and becomes a stable factor in the cell line. The advantage of this system is that once the gene of interest is turned on by the addition of the chemical inducer isopropylthiogalactoside (IPTG), the *E. coli* cells will transcribe the target gene into mRNA and subsequently translate the mRNA into protein, just as if the gene was one of the cell's own.

There are problems with overexpressing genes from eukaryotes (e.g. yeast, animals, insects, and us) in prokaryotic cells (*E. coli* and other bacteria) because as one might imagine, the way in which a cow makes a protein in its brain is far different than the way *Escherichia coli* would make the same protein it not only has no use for, but could also be toxic to the cell. These problems are largely due to differences in the way proteins are synthesized in eukaryotes versus prokaryotes, namely differences in transcription, translation, codon preference, cellular redox environment, inclusion body formation, protein glycosylation and other cotranslational modifications to name a few. Despite all of these differences, it is still possible to overexpress eukaryotic proteins in prokaryotic cells because many technologies have been developed to address these problems. Therefore, developing an overexpression system in which an *E. coli* cell (a prokaryote) is to make a protein normally found in a eukaryote requires a systematic survey of protocols that promise to alleviate problems arising from the inherent differences between eukaryotes and prokaryotes.

I have developed an overexpression system in *Escherichia coli* for the human enzyme methionine aminopeptidase so that future studies on the structural, biochemical, kinetic, and physiological importance of this protein can be better addressed *in vitro*. Currently there is some literature on the function of this protein in mammals, but there are many details about human methionine aminopeptidase that have yet to be determined. Developing an overexpression system for the human methionine aminopeptidase is

important because it will permit other researchers to have access to a large quantity of protein required for all of the experiments necessary to answer questions about the role this protein plays in cells.

Results and Discussion

For many of the problems encountered when overexpressing proteins from eukaryotes in prokaryotic cells, plasmids are available that address each problem by incorporating different control elements or accessory proteins. To this same extent, various *E. coli* strains have also been developed for overexpression purposes. In order to provoke these *E. coli* cells into synthesizing the human methionine aminopeptidase protein, I have cloned the gene for the enzyme into a number of different plasmids and transformed them into several *E. coli* overexpression strains. In some cases it seemed advantageous to try to express just the catalytic domain of the enzyme, in which case truncated versions of the full length gene were used for overexpression. To summarize the results, all of the overexpression constructs that contained the full length human methionine aminopeptidase failed to express any detectable protein by three hours post induction. However, the two overexpression constructs that contained only the protease domain from human methionine aminopeptidase both succeeded to produce the expected size protein (see Table 1).

The first problem in overexpressing eukaryotic proteins in prokaryotic cells arises from differences in the way genes are transcribed. Transcription in eukaryotes is under considerable regulation, while transcription in prokaryotes is much more simple. Furthermore, prokaryotic RNA polymerases do not even recognize the promoters from eukaryotes. To overcome this problem, only the coding region of human methionine aminopeptidase was positioned into an expression vector, leaving all of the eukaryotic regulatory DNA behind. Instead, the vector contains a promoter from a bacteriophage called T7. The gene encoding T7 RNA polymerase resides on the *E. coli* chromosome, and has been manipulated to be induced by IPTG. Since T7 RNA polymerase will only recognize its own promoter sequence, and no other T7 promoters reside on any *E. coli* DNA, only the human methionine aminopeptidase gene is transcribed upon induction by IPTG [17, 18].

Table 1

Constructs tested for expression of soluble human methionine aminopeptidase.

Plasmid Name	N-term fusion	Insert[1]	Expected Size (kDa)	Vector	Rationale	Expression?[2]	Soluble?
hupet28b		HuMetAP	52.9	pET28b	T7 promoter driven vector.	No	
hishupet28b	poly-his	HuMetAP	55.1	pET28b	Reduces GC content of initiating sequence.	No	
hupsbet		HuMetAP	52.9	pSBET	Coexpresses the rare arg tRNA.	No	
hishupsbet	poly-his	HuMetAP	55.1	pSBET	Reduces GC content of initiating sequence and coexpresses the rare arg tRNA.	No	
gsthupgex	GST	HuMetAP	79.7	pGEX	More significant initiating sequence than poly-his tag.	No	
hisprothupet28b	poly-his	ProtHuMetAP	38.3	pET28b	Truncates HuMetAP to the core catalytic domain.	Yes[3]	No
gstprothupgex	GST	ProtHuMetAP	63.1	pGEX	GST can increase the solubility of insoluble proteins.	Yes	No[4]

1. Abbreviations: HuMetAP, full length human methionine aminopeptidase; ProtHuMetAP, human methionine aminopeptidase truncated to core catalytic domain.
2. Expression is defined by the appearance of a band three hours after induction visualized by coomassie-blue stained 12.5% SDS-PAGE.
3. While all these plasmids were also transformed into BL21(DE3)pTrx and BL21(DE3)pGroEL, this construct was the only plasmid to express protein in BL21(DE3)pGroEL. No constructs expressed protein in BL21(DE3)pTrx cells.
4. This construct may have expressed trace soluble protein (see text).

Once human methionine aminopeptidase mRNA is made, other problems are encountered. The beginning of the full length human methionine aminopeptidase gene is very GC rich and Watson-Crick base pairing predicts G-C pairs to be more stable than A-T pairs. This is problematic because if an intramolecular or intermolecular complementary strand of RNA comes into contact with the initiating sequence of the human methionine aminopeptidase, it could bind strongly to this region and possibly mask the Shine-Dalgarno sequence, a binding site for prokaryotic ribosomes. If this happens, ribosomes will be prevented from accessing the mRNA and no protein will be made.

There are two simple strategies to alleviate this problem. Since complementary mRNA sequence on the same mRNA molecule will be the most potent translation suppresser, scanning the inserted gene of interest for stretches of complementary sequence can eliminate this problem. However, complementary RNA sequences can come from other transcribed genes as well, so an alternative strategy is to fuse the 5' end of the gene of interest with a stretch of protein encoding sequence that has lower GC content. One construct tested for expression was a fusion protein consisting of the full length human methionine aminopeptidase gene and a poly-histidine tag. The other fusion protein tested consisted of the full length human methionine aminopeptidase gene and another gene encoding glutathione-S-transferase (GST). Neither of these constructs expressed protein.

Once translation begins, the prokaryotic ribosomes may have problems when they encounter codons that are only infrequently translated in prokaryotes. The genetic code is conserved between all species, but over the course of evolution, species have developed preferences for which codons they use to translate mRNA. In particular, the codons AGG and AGA, which code for arginine, are used frequently in eukaryotes like humans, but are rarely used in *Escherichia coli*. As a result, *E. coli* ribosomes may misincorporate amino acids, stall, skip, or terminate the protein when they encounter these rare codons. While tandem AGG or AGA codons have been shown to be particularly problematic, lone AGA or AGG codons may also disrupt proper translation [19].

In order to overcome this problem, the tRNA responsible for translating these rare arginine codons can be coexpressed with human methionine aminopeptidase, which has 11 AGA codons and 3 AGG codons. To do this, the gene for human methionine aminopeptidase was cloned into an expression vector called pSBET [19]. Besides having all of the necessary components required to overexpress protein, pSBET also contains the arg U gene encoding the rare arginine tRNA so that the intracellular concentrations of this factor will be higher during protein synthesis. Neither of the constructs made to take advantage of coexpressing rare tRNAs produced any protein after three hours of induction. Furthermore, the constructs containing just the protease domain of human methionine aminopeptidase expressed protein well, even though this domain still contained 6 AGA and 2 AGG codons. These results suggest that prokaryotic ribosomes, as in *Escherichia coli*, do not have difficulty translating these codons when they are not adjacent to other rare codons.

The plasmid constructs that did succeed in making protein were truncated versions of the full length human methionine aminopeptidase in which sequence encoding the first 155 amino acids was removed, leaving only the coding region for the last 323 amino acids of the protein. This truncation was fused to both a poly-histidine sequence and the gene encoding GST. At three hours post induction, significant bands appeared on the SDS-PAGE gels indicating expression. However, these truncated protein fusions were not soluble. Strategies to express soluble protein often include inducing at lowered temperatures, inducing with less IPTG, and growing the cells in shaker flasks. I expressed the poly-histidine fusion (*hisprothupet28b*) under all of these conditions and none produced soluble protein. However, I have yet to do these same experiments for the GST fusion (*gstprothupgex*).

The full length methionine aminopeptidase also has 13 cysteine residues which is problematic because the cytoplasmic environment in *Escherichia coli* is more reducing than that in mammalian cells. As a result, eukaryotic proteins that normally form disulfide

bonds when synthesized in mammalian cells are unable to do so in *E. coli* [20]. This can be particularly problematic to solubility, since correct disulfide bond formation may be required to prevent aggregation.

There have been a few types of *E. coli* overexpression cells developed to overcome the formation of inclusion bodies. AD494 cells are a thioredoxin reductase deficient strain that permit the formation of disulfide bonds. In addition, BL21(DE3)pLysS cells have been modified to overexpress thioredoxin in the place of T7 lysozyme (BL21(DE3)pTrx) in order to catalyze proper disulfide bond formation in *Escherichia coli* cells [21]. None of the plasmids produced protein in this cell line (see Table 1).

Incorrect disulfide bond formation may explain some proteins' tendency to be expressed as insoluble aggregates in *Escherichia coli*, but many proteins fail to remain soluble for other reasons. In order to accommodate these proteins, other BL21(DE3)pLysS modified cells in which the T7 lysozyme gene has been replaced by the *E. coli* protein folding chaperones from the GroESL operon may be used to help catalyze proper protein folding [21]. These cells are called BL21(DE3)pGroEL. In addition, amino-terminal fusions with proteins that are known to be soluble (like GST) have helped promote the solubility of target proteins [22]. I expected that plasmids which failed to express any protein in BL21(DE3) cells would also fail in BL21(DE3)pTrx or BL21(DE3)pGroEL because these cells do not correct problems associated with transcription or translation. The fact that *hisprothupet28b* (poly-his fused with the protease domain of human methionine aminopeptidase) only expressed protein in BL21(DE3)pGroEL and not BL21(DE3)pTrx was surprising because these cell strains promote solubility and should not inhibit protein synthesis. I suspect that *gstprothupgex* (GST fused with the protease domain of human methionine aminopeptidase) failed to express protein in either of these strains because the pGEX vector is Taq promoter driven, and overexpression of T7 RNA polymerase (as in λDE3 lysogens) might have disrupted the *E. coli* polymerases from transcribing this gene.

While there are techniques available to alleviate most of the problems encountered when expressing a eukaryotic protein in a prokaryotic cell, no technique can compensate for the differences in cotranslational and post-translational modifications between these cells. Eukaryotes frequently attach glycosyl moieties to different sites on the protein while prokaryotes do not. As a result, any protein that is synthesized with glycosyl groups in eukaryotes will lack these modifications when it is expressed in *E. coli* cells.

The inability of *E. coli* to glycosylate proteins is significant because there is evidence that the homologue to human methionine aminopeptidase, or p67, in rabbits is heavily glycosylated [23]. Furthermore, it has been shown that the first step in the degradation of p67 is the removal of these glycosyl groups; once the protein is free of these moieties the protein is rapidly degraded [15]. Because *Escherichia coli* cells do not glycosylate proteins, the human methionine aminopeptidase protein that is produced might be targeted for degradation in a similar way as what was found in rabbits. Since the glycosyl groups are all O-linked, they require serine or threonine residues to attach to. While the full length human methionine aminopeptidase contains 50 serine and threonine residues, 17 of these residues are located on the 155 amino acid amino-terminal extension, more than enough to accommodate for the 12 observed glycosyl groups. It could be that this region of the protein contains a degradation signal, and that when the glycosyl groups are present, the signal is masked. Once the glycosyl groups are removed, the signal is exposed and the protein is rapidly degraded. This could explain why the only plasmids that expressed protein were those that coded for just the catalytic domain of the enzyme and not the full length protein. However, more experiments would be required to justify this claim.

Conclusion

Currently, one of the most successful ways to obtain significant quantities of protein is to overexpress the protein in *Escherichia coli* cells. But the fundamental differences between eukaryotes and prokaryotes can stand in the way of overexpressing a eukaryotic gene, like human methionine aminopeptidase, in a prokaryote like *Escherichia coli*. Although there have been many techniques developed that attempt to compensate for these differences, none can guarantee success.

While most of the plasmids employed to produce human methionine aminopeptidase failed to express any protein, the two constructs that encode only the protease domain of this enzyme show promise. Although the protein made by these two constructs is not soluble, they both express well. Future experiments to express soluble protein may include expressing at different temperatures and in shaker flasks for the *gstprothupgex* construct. Another approach may be to refold the protein from the insoluble aggregates in the inclusion bodies.

The functions of many features of human methionine aminopeptidase are currently unknown, and no structural information exists for the long inserts that distinguish this enzyme from the *Escherichia coli* methionine aminopeptidase. However, once a protocol has been developed to produce large quantities of the human methionine aminopeptidase, structural analyses and other studies can get under way to answer questions about how this enzyme carries out its function in cells.

Materials and Methods

Construction of Expression Vectors

HUPET28b. Using 100 pmol of the oligonucleotide 5'-GAATCCATGG(NcoI)CGGGC-GTGGAGGAGGTAGCGGCCTCC (forward) and 100 pmol of the oligonucleotide 5'-ATC-TTACTCGAG(XhoI)TTAATAGTCATCTCCTCTGCTGACA-ACTTCT (reverse), polymerase chain reaction was carried out using a plasmid W1 (gift from Dr. Ralph Bradshaw, UC Irvine) containing the cDNA for human methionine aminopeptidase as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 55°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 55°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 1455 bp PCR product was digested with NcoI and XhoI and ligated into NcoI/XhoI cut pET28b (Novagen) to generate the plasmid *hupet28b.*

HISHUPET28b. Using 100 pmol of the oligonucleotide 5'-GGAATTCCATATG(NdeI) GCGGGCGTGGAGGAGGTAGC (forward) and 100 pmol of the oligonucleotide 5'-AT-CTTACTCGAG(XhoI)TTAATAGTCATCTCCTCTGCTGACA-ACTTCT (reverse), polymerase chain reaction was carried out using the plasmid *hupet28b* (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 55°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 55°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 1459 bp PCR product was blunt-end ligated into the EcoRV sites of pSBET and the resultant vector was digested with NdeI and XhoI. The 1439 bp fragment was isolated and ligated into NdeI/XhoI cut pET28b to generate the plasmid *hishupet28b.*

HUPSBET. Using 100 pmol of the oligonucleotide 5'-GGAATTCCATATG(NdeI)GCG-GGCGTGGAGGAGGTAGC (forward) and 100 pmol of the oligonucleotide 5'-ATCTT-AGATATC(EcoRV)TTAATAGTCATCTCCTCTGCTGACAACTTCT (reverse), polymerase chain reaction was carried out using the plasmid *hupet28b* (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 55°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 55°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 1459 bp PCR product was blunt-end ligated into the EcoRV sites of pSBET and the resultant vector was digested with NdeI and EcoRV. The 1441 bp fragment was isolated and ligated into NdeI/EcoRV cut pSBET to generate plasmid *hupsbet.*

HISHUPSBET. Using 100 pmol of the oligonucleotide 5'-TAATACGACTCACT-ATAGGG (forward) and 100 pmol of the oligonucleotide 5'-ATCTTAGATATC (EcoRV)TTAATAGTCATCTCCTCTGCTGACAACTTCT (reverse), polymerase chain reaction was carried out using the plasmid *hishupet28b* (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 53.5°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 53.5°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 1596 bp PCR product was digested with XbaI and EcoRV and ligated into XbaI/EcoRV cut pSBET to generate plasmid *hishupsbet.*

GSTHUPGEX. Using 100 pmol of the oligonucleotide 5'-GGGCCATGG(NcoI) CGGGCGTGGAGGAGGTAGCGGCCTCC (forward) and 100 pmol of the oligonucleotide 5'-ATCTTACTCGAG(XhoI)TTAATAGTCATCTCCTCTGCTGACA-ACTTCT (reverse), polymerase chain reaction was carried out using the plasmid

hishupet28b (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at a higher melting temperatures (40 seconds at 96°C, 40 seconds at 59.4°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 59.4°C, 2 minutes at 72°). The reaction was terminated after 7 minutes of incubation at 72°C. This 1454 bp PCR product was digested with XhoI and ligated into SmaI/XhoI cut pGEX-4T-3 (Pharmacia) to generate the plasmid *gsthupgex.*

HISPROTHUPET28b. Using 100 pmol of the oligonucleotide 5'-GGGCTCATA-TG(NdeI)GCATTAGATCAGGC (forward) and 100 pmol of the oligonucleotide 5'-ATCTTACTCGAG(XhoI)TTAATAGTCATCTCCTCTGCTGACAACTTCT (reverse), polymerase chain reaction was carried out using the plasmid *hishupet28b* (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 58.3°C, 2 minutes at 72°C); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at 58.3°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 992 bp PCR product was blunt-end ligated into the EcoRV sites of pSBET and the resultant vector was digested with NdeI and XhoI. The 974 bp fragment was isolated and ligated into NdeI/XhoI cut pET28b to generate the plasmid *hisprothupet28b.*

GSTPROTHUPGEX. Using 100 pmol of the oligonucleotide 5'-GGGCTCATATG (NdeI)GCATTAGATCAGGC (forward) and 100 pmol of the oligonucleotide 5'-ATCTTACTCGA-G(XhoI)TTAATAGTCATCTCCTCTGCTGACAACTTCT (reverse), polymerase chain reaction was carried out using the plasmid *hisprothupet* (see above) as template and 2 units of vent polymerase. The first three cycles of the reaction were carried out at higher melting temperatures (40 seconds at 96°C, 40 seconds at 58.4°C, 2 minutes at 72°); the last 22 cycles at lower melting temperatures (40 seconds at 94°C, 40 seconds at

58.4°C, 2 minutes at 72°C). The reaction was terminated after 7 minutes of incubation at 72°C. This 992 bp PCR product was digested with XhoI and ligated into SmaI/XhoI cut *gsthupgex* (see above) to generate the plasmid *gstprothupgex*.

Mini-expression Experiments

To test each construct's ability to overexpress human methionine aminopeptidase, the Taq promoter driven vectors (*gstprothupgex* and *gsthupgex*) were transformed into BL21 cells while T7 promoter driven vectors (*hupet28b*, *hishupet28b*, *hupsbet*, *hishupsbet*, and *hisprothupet28b*) were transformed into DE3 lysogens of BL21 cells. 500 µls of overnight culture was used to inoculate 25 mls of L-broth and antibiotic (kanomycin, ampicillin, or chloramphenicol), and the cells were grown at 37°C while being shaken 300 rpm. When the OD_{600} reached 0.5 OD, IPTG was added to bring the final concentration to 1 mM. Samples were taken at zero and three hours, and prepared for 12.5% SDS-PAGE. The gels were subsequently stained using coomassie dye, and the zero and three hour time points were compared; appearance of a band in the three hour sample indicated expression.

For the constructs that did express protein, the entire culture was harvested and resuspended in 35 mls of +T/G buffer (50 mM Hepes Ph 7.9, 10% glycerol, 0.1% Triton X-100, 0.5 M KCl, 5 mM imidazole), 1.4 mgs DNase, and 35 µls 1.0 M $MgCl_2$, and were subsequently french pressed to lyse the cells. After spinning these fractions at 39,000 xg for 45 minutes, the soluble supernatants were collected and the insoluble pellets were resuspended in 35 mls of 10% SDS, and the contents of each fraction were analyzed on 12.5% SDS-PAGE.

References

1. Arfin, S. M. & Bradshaw, R. A. (1988) *Biochemistry* **27,** 7979-7984.
2. Chang, S.-Y. P., McGary, E. C. & Chang, S. (1989) *Journal of Bacteriology* **171,** 4071-4072.
3. Miller, C. G., Kukral, A. M., Miller, J. L. & Movva, N. R. (1989) *Journal of Bacteriology* **171,** 5215-5217.
4. Gonda, D. K., Bachmair, A., Wunning, I., Tobias, J. W., Lane, W. S. & Varshavsky, A. (1989) *Journal of Biological Chemistry* **264,** 16700-16712.
5. Moerschell, R. P., Hosokawa, Y., Tsunasawa, S. & Sherman, F. (1990) *Journal of Biological Chemistry* **265,** 19638-19643.
6. Sherman, F., Stewart, J. W. & Tsunasawa, S. (1985) *Bioessays* **3,** 27-31.
7. Roderick, S. L. & Matthews, B. W. (1993) *Biochemistry* **32,** 3907-3912.
8. Bazan, J. F., Weaver, L. H., Roderick, S. L., Huber, R. & Matthews, B. W. (1994) *Proceedings. National Academy of Sciences (USA)* **91,** 2473-2477.
9. Li, X. & Chang, Y.-H. (1996) *Biochemical and Biophysical Research Communications* **227,** 152-159.
10. Arfin, S. M., Kendall, R. L., Hall, L., Weaver, L. H., Stewart, A. E., Matthews, B. W. & Bradshaw, R. A. (1995) *Proceedings. National Academy of Sciences (USA)* **92,** 7714-7718.
11. Zuo, S., Guo, Q., Ling, C. & Chang, Y.-H. (1995) *Molecular and General Genetics* **246,** 247-253.
12. Hershey, J. W. B. (1989) *Journal of Biological Chemistry* **264,** 20823-20826.
13. Hershey, J. W. B. (1991) *Annual Review of Biochemistry* **60,** 717-755.
14. Jackson, R. J. (1991) in *Translation in Eukaryotes*, ed. Trachsel, H. (CRC Press, Boca Raton, Fla), pp. 193-229.
15. Gupta, N. K., Datta, B., Ray, M. K. & Roy, A. L. (1993) in *Translational Regulation of Gene Expression 2*, ed. Ilan, J. (Plenum Press, New York), pp. 405-431.
16. Panniers, R., Rowlands, A. G. & Henshaw, E. C. (1988) *Journal of Biological Chemistry* **263,** 5519-5525.
17. Studier, F. W. & Moffatt, B. A. (1986) *Journal of Molecular Biology* **189,** 113-130.
18. Studier, F. W., Roesenburg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods in Enzymology* **185,** 60-89.

19. Schenk, P. M., Baumann, S., Mattes, R. & Steinbiß, H.-H. (1995) *BioTechniques* **19,** 196-200.

20. Derman, A. I., Prinz, W. A., Belin, D. & Beckwith, J. (1993) *Science* **262,** 1744-1747.

21. Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T. & Ishii, S. (1995) *Journal of Biological Chemistry* **270,** 25328-25331.

22. Smith, D. B. & Johnson, K. S. (1988) *Gene* **67,** 31-40.

23. Datta, B., Ray, M. K., Chakrabarti, D., Wylie, D. E. & Gupta, N. K. (1989) *Journal of Biological Chemistry* **264,** 20620-20624.