Arginyl aminopeptidase-like 1 (RNPEPL1): from hypothetical reading frame

toward functional characterization

By:

Karen Beasley Maynard

Doctorate of Philosophy in Molecular Bioscience

Middle Tennessee State University

Murfreesboro, TN

December 2014

Dissertation Committee:

Dr. Rebecca Seipelt-Thiemann, Chair

Dr. Frank Bailey

Dr. A. Bruce Cahoon

Dr. Paul Kline

Dr. J. Brian Robertson

It is with great excitement that I dedicate this dissertation to my family:

My parents that have always loved and supported me, without which I

would not have been prepared or able to take on this great opportunity,

My loving husband, that helped me survive the long hours and frustrations,

and share the joys that accompanied this process,

And, lastly, my wonderful daughter, with the hope that she takes this as

encouragement to make her own life what she wants it to be.

It is my hope to make you all proud.

ACKNOWLEDGEMENTS

ABSTRACT

Arginyl aminopeptidase-like 1 (RNPEPL1) was first reported as a hypothetical reading frame during various genomic studies to complete the human genome. This study aimed to move the gene from hypothetical status to verified, as well as move toward establishing a biological role or function in inflammation based on gene family similarity.

First, this investigation confirmed the gene was transcribed and the transcript encoded a functional enzyme with M1 aminopeptidase characteristics. RNA was expressed in 20 human tissues, with three splice variants being identified in the process. Protein was expressed, purified, and characterized. RNPEPL1 protein did have enzymatic activity with a high specificity for cleaving methionine, citrulline, and glutamine N-terminal residues. The enzyme also demonstrated a broad pH range and similar inhibitor sensitivities to other members of the M1 family.

The M1 aminopeptidase family was then studied in depth to identify conservations within the gene structures that may indicate potential functional role(s) of RNPEPL1. This study followed the twelve human members, eleven mouse members, and eleven rat members of the M1 aminopeptidase family through multiple alignments, phylogenetic trees,

synteny studies, and exon structure analyses. All members shared the conserved $HEXXHX_{18}E$ and G/AMEN domains crucial for inclusion into the family. Phylogenetic analyses at the amino acid level indicated the presence of three groups within the M1 family. The first group contained three genes (LTA_4H, RNPEP, and RNPEPL1), while the second group contained five highly related genes (ENPEP, ANPEP, APQ, TRHDE, and NPEPPS), and the final group contained three related genes (LNPEP, ERAP1, and ERAP2). ERAP2, which was only found in humans, resulted in a synteny study among the three species. Results suggested that ERAP2 may have been destroyed during a chromosomal rearrangement in the rat/mouse lineage, among others. Finally, exon structure analyses showed high conservation of most of the second and third groups, with the LTA_4H group appearing to have diverged earlier.

Due to the clustering of RNPEPL1 with the LTA_4H family in the above studies, it was possible that RNPEPL1, like RNPEP and LTA_4H, might have a role in inflammation. To test this, RNPEPL1 expression levels were analyzed to determine what changes, if any, were present in cells, blood plasma, and tissue sections undergoing an inflammatory response. Results indicated that RNPEPL1 expression does decline between 25% – 30% after 24 hours of treatment.

In conclusion, RNPEPL1 is a true gene that transcribes a functional M1 aminopeptidase. This family is one of great divergence, indicating that it is an old gene family. RNPEPL1 is best placed within the LTA_4H group, in which the other members play various roles in inflammation. RNPEPL1 does decrease significantly during inflammatory response, however the change is slight.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

LIST OF ABBREVIATIONS

ANPEP   alanyl (membrane) aminopeptidase

APO   aminopeptidase O

APQ   aminopeptidase Q / laeverin

BCA   bicinchoninic acid

BMDMO   bone marrow derived macrophage

BP   base pair

ENPEP   glutamyl aminopeptidase

ERAP1   endoplasmic reticulum aminopeptidase 1

ERAP2   endoplasmic reticulum aminopeptidase 2

HRP   horseradish peroxidase

IHC   Immunohistochemistry

IMAC   immobilized metal affinity chromatography

JTT   Jones-Taylor-Thronton

LNPEP   placental leucine aminopeptidase

LPS   Lipopolysaccharide

LTA_4 leukotriene A_4

LTA_4H leukotriene A_4 hydrolase

LTB_4 leukotriene B_4

MEGA   Molecular Evolutionary Genetics Analysis

MHC   Major Histocompatibility Complex

NCBI   National Center Biotechnology Information

NIDDM   non-insulin dependent diabetes mellitus

NNI   Nearest Neighbor Interchange

NPEPPS   puromycin-sensitive aminopeptidase

NT   Nucleotide

ORF   open reading frame

PBS   phosphate buffered saline

PCR   polymerase chain reaction

RNPEP   arginyl aminopeptidase

RNPEPL1   arginyl aminopeptidase-like 1

RPM   rotations per minute

RT-PCR   reverse transcriptase polymerase chain reaction

SDS-PAGE   sodium dodecyl sulfate polyacrylamide gel electrophoresis

TBS   tris buffered saline

TRHDE   thyrotropin-releasing hormone-degrading ectoenzyme

UCSC   University of California Santa Cruz

CHAPTER ONE: INTRODUCTION

The M1 family of aminopeptidases consists of 12 members in humans to date that are all known to bind a single zinc ion. These members share a common $HEXXH(X)_{18}E$ motif, the zinc binding domain that is crucial for inclusion into the family, and GAMEN motif, thought to be the substrate binding domain for the majority of the enzymes [1]. Additionally, all members are believed to have aminopeptidase activity, yet some of the members are known to be bifunctional. Few members of this family have been fully characterized; however, what is known has allowed them to be grouped into sub-groups by amino acid similarity, activity, and/or genomic structure.

The first group, the oxytocinases, contains three structurally similar members. Placental leucine aminopeptidase (LNPEP) is reported to be the main regulator of oxytocin during pregnancy, degrades vasopressin, cleaves angiotensin III [2], and serves as the angiotensin IV receptor in the brain [3]. All of these roles are functions of the cleaved, soluble form of the protein. Additionally, a membrane bound form of the protein associates with GLUT4 post stimulation with insulin and is involved with various areas

of the GLUT4 response pathways. All three members share functional involvement in preparing antigens for presentation to MHC Class I molecules [4], however, endoplasmic reticulum aminopeptidase1 (ERAP1) and endoplasmic reticulum aminopeptidase 2 (ERAP2) vary in other functions and location [2]. ERAP1 and ERAP2 are localized in the endoplasmic reticulum, where both are most well-known for cleaving antigens for presentation to MHC Class I molecules [2]. In addition, ERAP1 inactivates angiotensin II and III, whereas the ERAP2 role in the pathway is specified to only cleaving angiotensin III [5]. ERAP1 also appears to play a role in hypertension and ventricular hypertrophy through the angiotensin function [2] and suppresses angiogenesis through VEGF interaction [6].

The second functionally similar set of five M1 proteases shares some angiotensin pathway cleavage roles, similar to the oxytocinases, as well as roles in the brain regulating blood pressure and apoptosis [7]. Laeverin (aminopeptidase Q, APQ) cleaves angiotensin III [7]. It is reported to be expressed as a cell-surface protein on human extravillous trophoblasts [8], yet RNA is expressed in many tissues (data not published). Like LNPEP, it is believed to play important roles in regulating pregnancy [7]. Puromycin-sensitive aminopeptidase (NPEPPS) inhibition has been shown to be involved with apoptosis inhibition during cell cycle [9], protein catabolism,

reproduction, and pain and anxiety regulation in the brain [10].

Thyrotropin-Releasing Hormone-Degrading Ectoenzyme (THRDE) is

involved in regulation of the hypothalamic regions of the brain through

both hormone and neurotransmitter actions [11]. THRDE shares great

structural similarity with the remaining two members of this group,

aminopeptidase A and aminopeptidase N. Aminopeptidase A (ENPEP,

glutamyl aminopeptidase) hydrolyzes angiotensin II to angiotensin III,

which is implicated as an important step in blood pressure regulation [12].

Aminopeptidase N (ANPEP, alanyl (membrane) aminopeptidase) is also

known as CD13. Enzymatically, ANPEP plays a major role in the digestion

of proteins. In addition to this, it acts as the receptor for corona virus [13].

Aminopeptidase O (APO) is the only one of the four members of the

final group, the ARM group, which shows cleavage in the angiotensin

pathway, converting angiotensin III to angiotensin IV [10]. APO is classified

by structure and other activity to the ARM subfamily of M1 peptidases,

grouped by the presence of an armidillo repeat protein-protein

interaction domain that is unique to the remaining four members of the

family [10]. Leukotriene A_4 Hydrolase (LTA$_4$H) is known for proinflammatory

hydrolase activity, converting LTA$_4$ to LTB$_4$ [14] and is critical to the study of

M1 peptidases, as it has been structurally studied in depth, unlike the other

members. Arginyl aminopeptidase (RNPEP) plays a role in processing glucagon to mini glucagon [15], somatostatin-28 to somatostatin-14, and other species tissue specific cleavage roles [1]. Previous studies are unclear on the ability of RNPEP to perform the LTA_4H hydrolysis, with some studies reporting that RNPEP is able to perform the hydrolysis [1,16,17], yet one reporting that hydrolysis activity is absent [18]. Arginyl aminopeptidase like-1 (RNPEPL1, APB-like) is most structurally related to RNPEP, yet shows more enzymatic similarity to LTA_4H [19]. To date, only basic enzymatic characterization has been reported. This enzyme will be the focus of study.

The first area of this study confirmed and characterized RNPEPL1. RNPEPL1 was initially placed into the last sub-family by computational predictions upon gene discovery. This placement was confirmed in 2008 when initial characterization reported the transcription, translation, and aminopeptidase activity necessary for inclusion in the M1 aminopeptidase family. Enzymatic characterization of RNPEPL1 shows a strong preference for citrulline and methionine. Another interesting kinetic characterization of RNPEPL1 lies in the broad pH preference. This is slightly varied from the other members of the M1 family. These data shows that RNPEPL1 prefers a pH range between 6.5 and 8.2 with a maximum specificity at 7.2.

Inhibitory kinetics also support the placement into the M1 family, showing

that RNPEPL1 is inhibited by many of the same protease inhibitors as other

members. Along with enzymatic characterization, the preliminary report

showed 3 normally occurring RNA splice variations, leading to the

possibility of multiple forms and/or lengths of the protein. Finally, the initial

study of the genome indicated that RNPEPL1 encoded a 494 amino acid

protein. This information was the basis of the characterization. However,

recent comparison with other genomes has led to an updated

computational prediction that additional upstream coding region in exon

1 exist, leading to the possibility of a larger RNPEPL1 protein 725 amino

acids in length [19]. Variation in the protein primary structure, also known

as splice variation, can influence the function of the protein. Splice

variations occur naturally to allow for a large amount of protein variation

from a smaller gene pool. Variations include the addition and exclusion

of exons and introns, yielding insertions and deletions of amino acids [20].

The second area of this study completes an in-depth analysis of the

M1 family to develop better insight into RNPEPL1 and possible activities.

As previously described, the M1 family appears to subgroup by common

functions, yet the family as a whole acts in many functional roles. To

determine how the family came to be and understand the divergence

seen, this study will analyze phylogenetic trees, gene and exon structure, and synteny of the members to better understand the relationships with the family subgroups and as a family.

Results from the previous two sections suggested that RNPEPL1 might play roles in biological processes similar to its most related family members. Therefore, the final goal of this research plan was to gather cellular and animal-level evidence as to the expression of RNPEPL1 changes in inflammation and whether expression changes to a significant degree to warrant further investigation. Microarray data showed preliminary evidence that RNPEPL1 expression is altered in several different inflammatory models. These preliminary studies showed that expression is reduced in inflamed heart and increased in the inflamed liver, but to small degrees.

Together, these research goals provide a more thorough and complete biochemical description of RNPEPL1, expand the current understanding of its family and divergence, and suggest that RNPEPL1's most significant functional role is unlikely to be in inflammation.

CHAPTER TWO: ARGINYL AMINOPEPTIDASE-LIKE 1(RNPEPL1) IS AN ALTERNATIVELY PROCESSED AMINOPEPTIDASE WITH SPECIFICITY FOR METHIONINE, GLUTAMINE, AND CITRULLINE RESIDUES

Michael W. Thompson, Karen A. Beasley, Matthew D. Schmidt, and Rebecca L. Seipelt

Department of Biology, Middle Tennessee State University, Murfreesboro, TN 37132

ABSTRACT

A previously uncharacterized member of the M1 family of zinc metallopeptidases, arginyl aminopeptidase-like 1 (RNPEPL1), was cloned and expressed, and the recombinant enzyme characterized. RNPEPL1 was a broad specificity aminopeptidase with preference for a P1 methionine, glutamine, or citrulline residue, and exhibited a broad pH preference, with maximal activity observed between pH 6.6 and 8.0. The enzyme was inhibited by calcium ions but unaffected by chloride ions, and was insensitive to specific inhibitors of the closely-related arginyl aminopeptidase, indicating similarity to leukotriene A_4 hydrolase. RT-PCR analysis of RNPEPL1 expression revealed a ubiquitous tissue distribution, consistent with a general housekeeping function, but also revealed alternative splicing of the mRNA in all tissues examined. The inclusion of intron 5 was predicted to result in a truncated protein product, while an alternative 3' splice site of exon 9 of the reference sequence was predicted to result in the omission of a conserved eleven amino acid stretch from the C-terminal domain.

Keywords: aminopeptidase; metallopeptidase; protease; zinc; calcium; alternative splicing

INTRODUCTION

The M1 family of zinc metallopeptidases is a family of enzymes that utilize Zn^{2+} as reactive site centers. Most are aminopeptidases, exopeptidases which sequentially hydrolyze one or more amino acid residues from the N-terminus of a peptide or protein substrate. Members of this family have been implicated in many diverse functions, from the breakdown of peptides and recycling of amino acids from general cellular proteolysis [1] to the regulation of physiological signaling cascades through limited proteolysis of peptide hormones [2], and in the processing of peptides displayed on class I MHC molecules [3,4]. Accordingly, members of this family have garnered attention as potential drug targets. Inhibitors of these enzymes have shown promise in treating many conditions, such as hypertension [5], inflammation [6], and even some cancers [7,8]. In humans, the M1 family is comprised of at least twelve enzymes, and may be further subdivided into three groups, according to differences in subcellular localization, catalytic activity, sensitivity to various inhibitors, and genomic structure: the membrane-bound aminopeptidases, the oxytocinase-like enzymes, and the soluble leukotriene A_4 hydrolase-like enzymes.

The membrane-bound members of the M1 family include laeverin (LAEV, APQ), thyrotropin releasing hormone-degrading ectoenzyme (TRHDE), glutamyl aminopeptidase (APA, ENPEP), and aminopeptidase N/alanyl aminopeptidase (APN, ANPEP). These enzymes are type I integral membrane proteins, and three of them (TRHDE, ANPEP, and ENPEP) have been hypothesized to participate in the breakdown of peptide hormones at the cell surface [2,9,10]. The enzymes of this group also exhibit genomic structures that are more similar to each other than to the other members of the M1 family, especially in the gene regions encoding the transmembrane domain.

The oxytocinase-like members of the M1 family share many genomic structural characteristics with the membrane-bound enzymes, but most are cytoplasmic and lack transmembrane domains. This group includes oxytocinase/insulin-regulated aminopeptidase (LNPEP), puromycin-insensitive leucine-specific aminopeptidase (ERAP1), puromycin-sensitive aminopeptidase (NPEPPS), and endoplasmic reticulum aminopeptidase-2 (ERAP2). While these enzymes are not thought to regulate the degradation of peptide hormones, much evidence suggests the involvement of at least ERAP1 and ERAP2 in the processing of peptide antigens for display on class I MHC [3,11-13]. LNPEP

has been implicated in memory and learning via its role as the angiotensin IV receptor [14]. It also has the ability to cleave neuropeptides, such as vasopressin, oxytocin, and met-enkephalin [15], although its involvement in this process remains to be determined.

Although included with the oxytocinase-like group of the M1 family, NPEPPS could theoretically be placed into a group of its own because it has a unique genomic structure [1] and an unusual sensitivity to the protein synthesis inhibitor puromycin. Like ERAP1 and 2, NPEPPS has been implicated in class I MHC antigen presentation [3]; thus, it is possible that peptides resulting from cellular proteolysis are processed in the cytoplasm by NPEPPS before trimming in the endoplasmic reticulum by ERAP1 and 2, in order to generate the proper peptide length and N-terminus for binding to class I MHC. Together, it is likely that these enzymes contribute to the huge repertoire of antigens that are displayed on class I MHC at the cell surface. NPEPPS has also been implicated in sexual development, reproduction, and brain function [3,16-18].

The enzymes of the third group, the leukotriene A_4 hydrolase-like enzymes, share a common genomic structural pattern that differs from other members of the M1 family. They also show more homology to each other at the amino acid level than to the other members of the M1 family,

and are thought to have arisen from a single ancestral gene. This group includes leukotriene A$_4$ hydrolase (LTA$_4$H), arginyl aminopeptidase/aminopeptidase B (RNPEP), aminopeptidase O (APO), and aminopeptidase B-like 1 (RNPEPL1), a previously uncharacterized member of the M1 family identified through numerous genome-wide surveys [19-22]. The enzymes of this subgroup are also catalytically different than other M1 family members, as most show a preference for basic amino acid residues, and may also possess an unusual epoxide hydrolase activity that shares an overlapping active site with the peptidase activity. LTA$_4$H, the most well characterized enzyme of this subgroup, has been implicated in inflammation, converting leukotriene A$_4$ into leukotriene B$_4$ [23], and in normal cardiac function [24]. Likewise, RNPEP has been implicated in numerous cellular processes, most notably the processing of peptide hormones [25-28], but reports differ as to whether or not RNPEP possesses an epoxide hydrolase activity analogous to that of LTA$_4$H [29,30].

RNPEPL1 initially attracted interest as a candidate gene for genetic predisposition to non-insulin-dependent diabetes mellitus (NIDDM) because of its proximity to the calpain-10 gene, but no links were found [31]. Thus, despite its close homology to RNPEP and LTA$_4$H and proximity to

a disease-associated chromosomal locus, RNPEPL1 has yet to be characterized at either the genetic, biochemical, or physiological level. In order to characterize RNPEPL1, its cDNA was amplified from a library and subcloned into an expression vector, and the recombinant enzyme generated. Kinetic studies of recombinant RNPEPL1 indicate that it is a broad-specificity aminopeptidase with preference for methionine, glutamine, and citrulline residues. Alternative mRNA splicing studies of the RNPEPL1 mRNA indicate that the gene is expressed in all tissues examined, and that the pre-mRNA is processed into three isoforms in normal tissues: one that encodes the full-length protein, one that that generates a truncated protein, and one that generates an isoform lacking eleven conserved amino acids from the C-terminal domain.

MATERIALS AND METHODS

Reagents

Amino acid 7-amino-4-methylcoumarin substrates were purchased from Sigma Chemical Company or Bachem. Imidazole, agarose, glucose, galactose, puromycin, and bestatin were purchased from Sigma

Chemical Company. Pfx, TEV protease, and nickel-nitrilotriacetic acid (Ni-NTA) agarose were purchased from Invitrogen. Total human mRNA from normal tissues was purchased from Ambion. Reverse-transcriptase polymerase chain reaction (RT-PCR) kits and enzymes were purchased from Qiagen. Oligonucleotides were purchased from Midland Certified Reagent Company and Integrated DNA Technologies.

Molecular Cloning of RNPEPL1

The wild-type RNPEPL1 cDNA representing the reference sequence available from GenBank (accession number NM_018226) was amplified by PCR from the MegaMan human transcriptome cDNA library (Stratagene) using Pfx polymerase (Invitrogen) to yield a 1.5 kb PCR product consisting of the RNPEPL1 open reading frame (ORF) lacking its natural stop codon. The primers used for the PCR reaction were ATGAGTGCCACCCGGACTGCATAC (forward) and GGCAGACACATTGACGTCCCTGAG (reverse). This PCR product was purified by precipitation and used as a template for a second PCR reaction utilizing the same forward primer as the first reaction and a second reverse primer,

GGCCTGAAAATACAGGTTTTCGGCAGACACATTGACGTCCCTGAG, to yield an amplimer consisting of the RNPEPL1 ORF fused to a C-terminal TEV-protease recognition site. This PCR product was directly ligated into pYES2.1-Topo (Invitrogen) to facilitate expression of the encoded protein fused to a C-terminal his$_6$ affinity tag in yeast. All clones were sequenced to ensure the correct orientation, to verify that the cDNA matched the NCBI reference sequence, and to ensure that the construct was in frame with the his$_6$ affinity tag and stop codon of the vector.

An RNPEPL1 cDNA for expression in Sf9 cells was generated by PCR using the yeast expression construct as a template. The primers used were CACCATGGCGAGCGGCGAGCATTCCCCCG (forward) and TCAACTGCCCTTGGGTGCCACGATCTG (reverse). A 1.5 kb PCR amplimer containing the RNPEPL1 ORF was obtained and ligated into the shuttle vector pENTR (Invitrogen). Enzyme-mediated homologous recombination was used to generate recombinant baculovirus DNA, which was transfected into Sf9 cells using the Baculodirect expression system (Invitrogen Corporation).

Sf9 cells were maintained at 28°C in TNM-FH medium supplemented with 10% fetal bovine serum, 0.25 µg/ml amphotericin B, 100 µg/ml penicillin, 100 µg/ml streptomycin, and 50 µg/ml gentamycin or 50%

supplemented TNM-FH/50% SF 900 II SFM medium (Invitrogen) with antibiotics and antifungal agents at the same concentrations. Cells were seeded at 20-30% confluence in T25 flasks for stationary cultures, or to approximately 3×10^5 cells/ml in plastic Erlenmeyer flasks for shaking cultures, which were shaken at 50 RPM. Transfected Sf9 cells were grown in TNM-FH supplemented with ganciclovir. Recombinant virus was amplified three times in cell culture prior to use in protein expression.

Expression Analysis of RNPEPL1

The tissue distribution and splicing pattern of the RNPEPL1 transcript were determined by RT-PCR. Sequences of mRNAs available at AceView (NCBI) allowed identification of potentially alternative regions in the pre-mRNA. Four alternatively spliced regions were identified along with five different polyadenylation sites, including two alternative terminal exons. Total mRNAs from normal human adipose, brain, cervix, esophagus, colon, kidney, ovary, placenta, prostate, skeletal muscle, small intestine, spleen, testes, thymus, thyroid, heart, liver, trachea, and lung tissues (Ambion) were used in one-step RT-PCR assays (Qiagen) utilizing three primer sets designed to amplify a specific region of the cDNA surrounding the

putatively alternatively spliced regions. Alternative intron and exon

numbers are noted as in AceView with the corresponding reference

sequence exon and intron number given in brackets (NM_018226).

Primers used to assay intron 5 [intron 5] inclusion were:

GTGGAGCAGTGGCTGAGTG (I-F) and CAGTCTCCAGGCAGGTGAAG (I-

R); primers used to assay exon 11 exclusion [an alternative exon found in

some cDNAs in intron 6 between exons 6 and 7] and intron 8 exclusion

[intron 7] simultaneously were: GTCAGCAAACTGCAGGTCAA (II-F) and

CTCAGGAAGGAGTCCAGCAG (II-R); and the primers used to assay intron

8 exclusion and alternative 3' splice site use in exon 17/18 [exon 9] were:

GCGCTTTGATGACTTTCTCC (III-F) and GCTGGATCCCTGAGACAGGT (III-

R). Products were fractionated on 1.5% agarose gels for size determination

prior to being ligated into a general cloning vector (pCR4, Invitrogen

Corporation) to confirm the identity of each PCR product.

Expression and Purification of Recombinant RNPEPL1

The RNPEPL1-pYES2.1 construct was transformed into

Saccharomyces cerevisiae 1971 (*MATα leu2-3 ura3-52 pep4-3*) for

expression of recombinant protein. Expression and purification methods

were similar to those published previously for LTA$_4$H [32], with a few modifications. Briefly, an overnight 4 ml yeast culture was diluted into 50 ml of yeast SC medium lacking uracil and grown for 4-6 hours at 30°C. RNPEPL1 expression was induced by dilution of this culture into 1 l of yeast SC medium lacking uracil and containing 1.1% galactose (w/v). After 60-72 hours, the cultures were harvested and the cell pellets stored at -80°C until ready for use.

Yeast cultures were lysed by agitation with glass beads at 4°C for 10 minutes in lysis buffer (50 mM HEPES, 0.125M NaCl, pH 7.4) on a vortex shaker, followed by centrifugation at 2,000 x g for 10 minutes at 4°C to remove the glass beads. The supernatant was further clarified by centrifugation at 30,000 x g for 30 minutes at 4°C in a Sorvall RC-6 centrifuge to remove cell debris.

Amplified RNPEPL1 baculovirus was used to infect a large Sf9 cell culture for expression of recombinant protein. Cells were harvested after 72 hours by centrifugation, and stored at -80°C until ready for use. The cells were resuspended in 5 ml lysis/storage buffer (50 mM HEPES, 0.125M NaCl, pH 7.4) and lysed by sonication on ice at 30% power for 5 minutes with repeating 30 second pulse/30 second rest cycles. The supernatant

was clarified by centrifugation at 30,000 x g for 30 minutes at 4°C in a Sorvall RC-6 centrifuge to remove cell debris.

Recombinant RNPEPL1 was purified from the yeast or Sf9 cell supernatant by chromatography on a 1 ml Ni-NTA column equilibrated with lysis/storage buffer at 4°C. RNPEPL1 was bound to the column by passing the lysate over the column three times. The column was subsequently washed with 0.5 l wash buffer (50 mM HEPES, 1.0 M NaCl, pH 7.4), and bound proteins eluted with 50 mM HEPES, 0.3 M NaCl, 0.2 M imidazole, pH 7.4. Fractions containing protein were identified by absorbance at 280 nm, combined, and dialyzed against 8 L TEV cleavage buffer (50 mM Tris-Cl, 0.15 M NaCl, pH 8.0) overnight at 4°C with one buffer change.

For affinity tag removal, DTT was added to a concentration of 1 mM and 5000 units of his_6-tagged TEV protease (Invitrogen) added. Digestion was conducted for 48-72 hours at 4°C, followed by dialysis against 2 l lysis/storage buffer to remove DTT. Recombinant processed RNPEPL1 (typically 30-40% of the reaction) was subsequently purified by passing the reaction mixture over a 2 ml Ni-NTA column and collecting the unbound fraction. Repurified enzyme was concentrated to a volume of less than 5 ml, dialyzed against 2 l lysis/storage buffer, and stored in small aliquots at -

80°C. Protein concentrations were determined simultaneously by the BCA method and by absorbance at 280 nm using an extinction coefficient of 1.501 ml/mg. Molar concentrations were calculated using a molecular weight of 55300 g/mol.

Kinetic Analysis of Recombinant RNPEPL1

Kinetic analyses of recombinant RNPEPL1 were performed at 37°C using the 7-amido-4-methylcoumarin (AMC) derivatives of all twenty naturally occurring amino acids, citrulline, and ornithine. A Spectramax M2e spectrophotometer (Molecular Devices) was used to measure the fluorescence of the free AMC product liberated by enzymatic hydrolysis of the substrate in assay buffer (50 mM HEPES, 0.125 M NaCl, pH 7.4). The instrument was set to an excitation wavelength of 370 nm and an emission wavelength of 460 nm, with both excitation and emission slits set to 9 nm and the PMT sensitivity set to medium. The extinction coefficient of the product was determined by measuring the fluorescence of multiple concentrations of AMC under the same conditions. The kinetic parameters (K_M and k_{cat}) were determined by fitting of the Michaelis-Menten equation to the obtained data sets using Scientist software

(Micromath). The pH dependence of enzymatic activity was determined by multiple enzyme assays containing 82 nM RNPEPL1, 0.5 mM Met-AMC (K_M=0.38 mM), and a buffer containing 50 mM HEPES, 125 mM NaCl at various pH levels ranging from 3.6 to 8.8.

The sensitivity of the enzyme to inhibitors was determined by using the compound as an alternate substrate inhibitor in an assay containing 100 nM RNPEPL1 and 0.15 mM methionine-AMC (K_M= 0.38 mM) in assay buffer. Enzyme and inhibitor were preincubated for 5 min at 37°C prior to addition of substrate. The K_i for each inhibitor was determined by fitting of the equation for competitive inhibition to the data using Scientist software (Micromath). Similar results for both substrates and inhibitors were obtained from at least two experiments utilizing multiple independent enzyme preparations.

The effects of chloride and calcium ions on RNPEPL1 catalytic activity were determined by incubating 82-280 nM enzyme in the presence of increasing amounts of NaCl or $CaCl_2$, respectively, in an assay mixture containing 100 nM RNPEPL1 and 50 mM HEPES, pH 7.4. The mixture was preincubated for 5 minutes, the assay initiated by the addition of 0.15 mM Met-AMC or 0.1 mM Cit-AMC, and the initial rates

determined. The quadratic binding equation was fit to the resulting data and a dissociation constant calculated.

RESULTS

Molecular Cloning and Expression of RNPEPL1

The RNPEPL1 cDNA was amplified by PCR from a human cDNA pool and ligated into the vector pYes2.1/V5/His/Topo for expression in *Saccharomyces cerevisiae*. Several clones containing the RNPEPL1 cDNA in the forward orientation were subjected to sequencing to verify their identities; of these, one clone was identical to the reference sequence in the NCBI database. This plasmid was subsequently used as a PCR template to generate recombinant plasmid DNA for construction of recombinant baculovirus and subsequent expression of the protein in Sf9 cells.
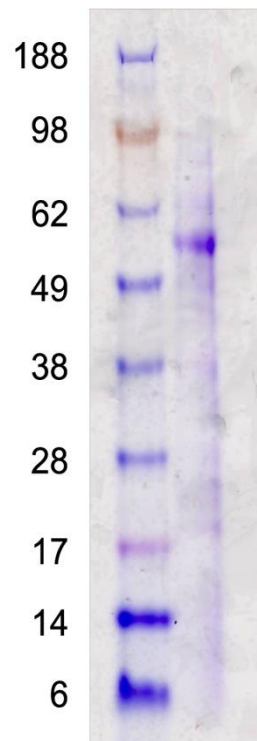
Figure 1. Purification of recombinant human RNPEPL1.

Histidine-tagged human RNPEPL1 was purified from Sf9 insect cells by

immobilized metal affinity chromatography (IMAC). Following affinity tag

removal, the protein was subjected to SDS-PAGE analysis. Identical results

were obtained from RNPEPL1 purified from both yeast and Sf9 cells.

Recombinant RNPEPL1 was expressed and purified from both yeast

and Sf9 cells as described in Materials and Methods. Typical yields were

approximately 0.5 mg/liter culture in yeast and 4 mg per 0.5 l culture in Sf9 cells. The his_6 affinity tag was subsequently removed by the addition of TEV protease, and the enzyme further purified over a second Ni-NTA column to yield recombinant RNPEPL1 that was >90% homogenous. RNPEPL1 exhibited an approximate molecular weight of 55 kDa by visual inspection on an SDS-PAGE gel (Fig. **1**).

Kinetic Analysis of RNPEPL1

The general kinetic properties of recombinant RNPEPL1 were determined using amino acid AMC substrates. Recombinant RNPEPL1 exhibited a broad substrate specificity with a preference for P1 methionine and glutamine residues; however, citrulline was also a favorable substrate (Fig. **2**, Table **1**). Even though RNPEPL1 is most similar in amino acid sequence to arginyl aminopeptidase (RNPEP), which has specificity for arginine and lysine amino termini (Pierotti et al 1995), hydrolysis of the AMC derivatives of these two amino acids by RNPEPL1 was relatively low (Fig. **2**). By k_{cat} values alone, RNPEPL1 substrate specificity was
Met>Gln>Cit>Ala>Leu>Phe>His>Arg>Ser>Tyr>Trp>Asn>Ile>Lys, with only

minor differences between the seven least favorable substrates (Table **1**).

By k_{cat}/K_M values, the preference was

Met>Cit>Gln>Leu>His>Tyr>Trp>Phe>Ala>Arg>Asn>Ile>Lys (Figure **2**).
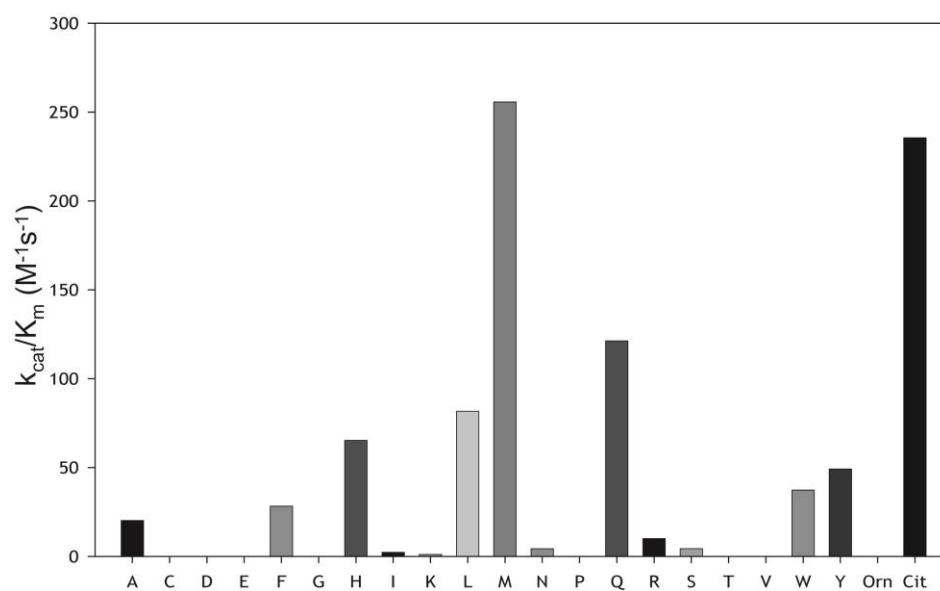

Figure 2. Substrate specificity of recombinant RNPEPL1.
The substrate specificity of RNPEPL1 was determined by determining the

k_{cat}/K_M for the aminoacyl AMC derivatives of all twenty naturally-occurring

amino acids, citrulline, and ornithine. Individual K_M and k_{cat} values for

each substrate are given in Table **2**.

Table 1. Kinetic parameters for RNPEPL1 for amino acid AMC substrates

K_M, V_{max}, and k_{cat} values were determined by fitting of the Michaelis-Menten equation to experimental data and are reported ± standard error.

	K_M (mM)	k_{cat} (min^{-1})		K_M (mM)	k_{cat} (min^{-1})
Ala-AMC	1.72 ± 0.09	2.09 ± 0.06	Leu-AMC	0.34 ± 0.09	1.69 ± 0.20
Arg-AMC	0.91 ± 0.14	0.55 ± 0.04	Lys-AMC	1.72 ± 0.09	0.23 ± 0.02
Asn-AMC	0.89 ± 0.19	0.23 ± 0.02	Met-AMC	0.37 ± 0.03	5.70 ± 0.20
Cit-AMC	0.17 ± 0.06	2.45 ± 0.30	Phe-AMC	0.65 ± 0.12	1.11 ± 0.10
Gln-AMC	0.41 ± 0.06	3.01 ± 0.20	Ser-AMC	1.76 ± 0.22	0.46 ± 0.04
His-AMC	0.18 ± 0.05	0.69 ± 0.07	Trp-AMC	0.12 ± 0.03	0.27 ± 0.02
Ile-AMC	1.05 ± 0.21	0.14 ± 0.02	Tyr-AMC	0.12 ± 0.03	0.35 ± 0.04

RNPEPL1 did not hydrolyze the AMC derivatives of glycine, tyrosine, aspartate, glutamate, proline, threonine, or valine. The pH dependence of the enzyme was relatively broad and slightly alkaline relative to that of other M1 metallopeptidases, with maximal activity observed between pH 6.6 and 8.0 (Fig. **3**).

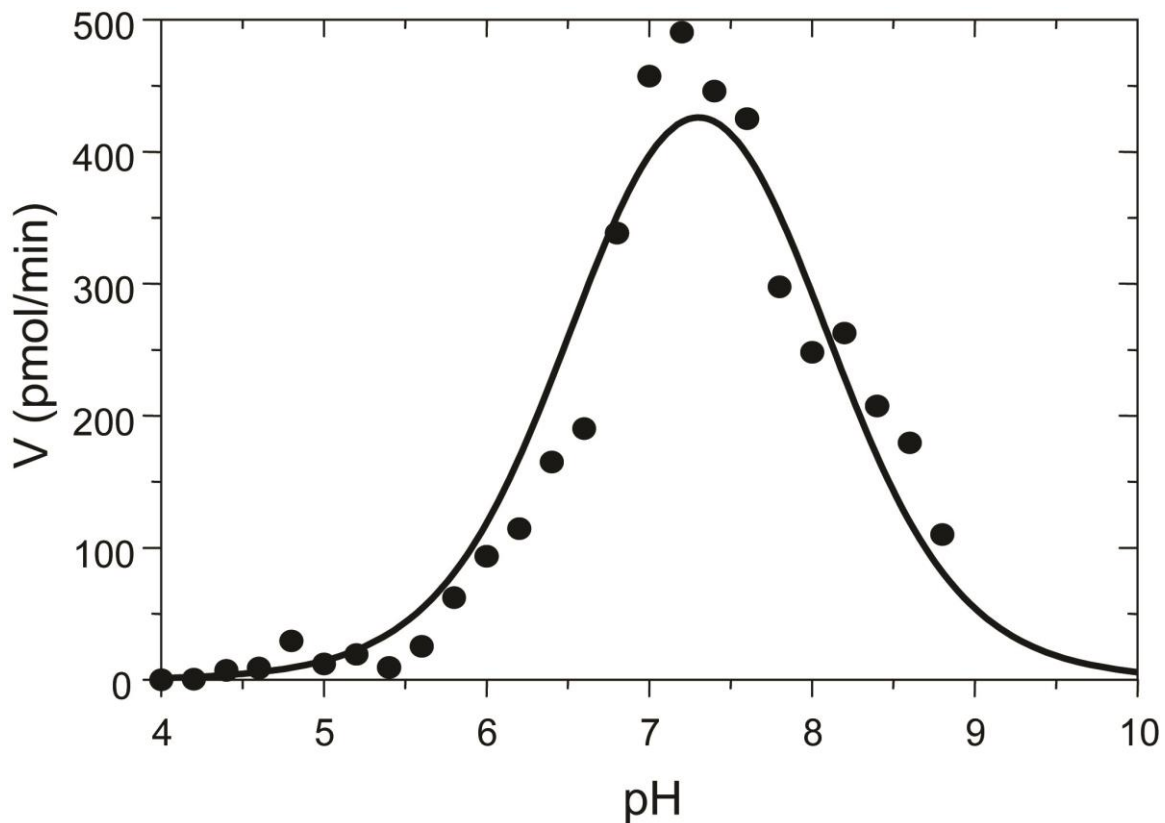Figure 3. pH optimum of RNPEPL1. The pH optimum for hydrolysis of 0.5 mM Met-AMC by 82 nM RNPEPL1 was determined by measurement of the catalytic rate in 50 mM HEPES, 0.125 M NaCl with a pH ranging from 3.6 to 8.8.

The sensitivity of RNPEPL1 to various inhibitors of metalloproteases and other proteolytic enzymes was also evaluated. RNPEPL1 exhibited a high sensitivity to amastatin and a moderate sensitivity to bestatin, similar

to most other members of the M1 family (Table **2**). RNPEPL1 showed low sensitivity to puromycin, and was totally unaffected by the RNPEP-specific inhibitors captopril and arphamenine A. Arphamenine B was an ineffective inhibitor as well, inhibiting catalysis by 22% at the highest concentration tested (3 mM). Inhibitors of serine proteases (PPACK I and II) and cysteine proteases (E-64) had no discernable effect on RNPEPL1 activity (data not shown).

Table 2. Affinity of RNPEPL1 for aminopeptidase inhibitors K_i values were determined by fitting of the competitive inhibition equation to the data and are reported ± standard error. ND=inhibition not detectable

	K_i (Met)
Bestatin	5.7 ± 1.2
Amastatin	0.19 ± 0.04
Puromycin	800 ± 140
Captopril	ND
Arphamenine A	ND[1]
Arphamenine B	ND[2]

[1] less than 5% inhibition observed at 0.6 mM

[2] 22% inhibition observed at 3 mM

Regulation of RNPEPL1 By Chloride and Calcium Ions

Many members of the M1 family of aminopeptidases are activated by halide salts [33,34] and/or regulated by Ca^{2+} ions [35]. The effects of chloride and calcium ions on RNPEPL1 activity were determined by examining the effect of increasing concentrations of NaCl or $CaCl_2$ on the hydrolysis of Met-AMC. RNPEPL1 activity was unaffected by chloride ions at concentrations up to 400 mM (Fig. **4A**), but was inhibited by calcium in a dose-dependent manner (Fig. **4B**).
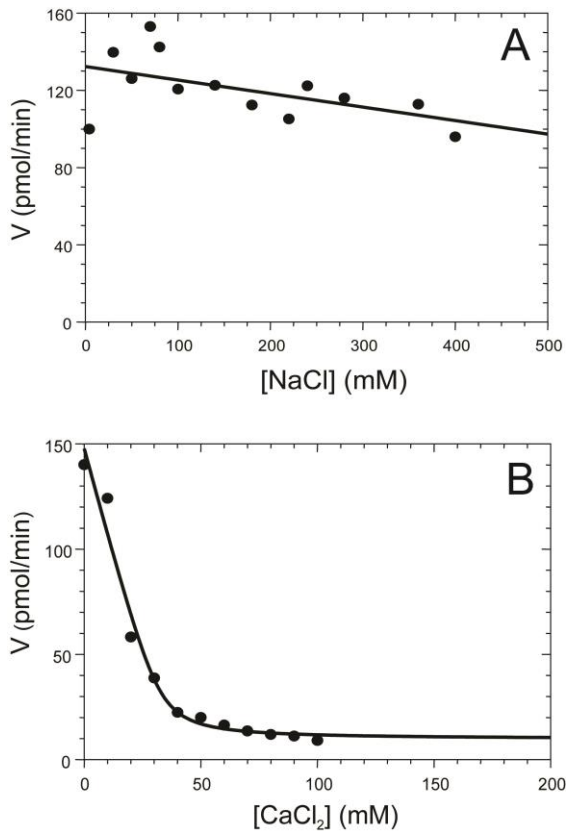
Figure 4. RNPEPL1 is inhibited by calcium ions, but is unaffected by chloride ions. The effects of Cl- and Ca^{2+} ions on the hydrolysis of Met-AMC were determined by addition of NaCl (A) or CaCl$_2$ (B) to an assay containing 0.125 mM Met-AMC and 82 nM (A) or 280 nM (B) RNPEPL1 at 37°C. Similar results were obtained using 0.1 mM Cit-AMC as substrate (data not shown).

Similar results were obtained using Cit-AMC as a substrate (data not shown). Calcium bound to the enzyme with an approximate K_D of 50 mM, as determined by fitting of the quadratic binding equation to the data.

Expression Analysis and Alternative mRNA Splicing of RNPEPL1

To determine the extent of alternative splicing of the RNPEPL1 mRNA, sequence information available from AceView at the National Center for Biotechnology Information (NCBI) [36] and from the University of California Santa Cruz (UCSC) Genome Browser [37] were examined to identify potentially alternative splicing events. Alternative intron and exon numbers are given relative to the AceView numbering system with the corresponding intron or exon for the reference sequence NM_018226 given in brackets. Four possible alternative splice choices were identified: region a, inclusion of intron 5 [intron 5]; region b, exclusion of exon 11 [an alternative exon within intron 6]; region c, inclusion of intron 8 [intron 7]; and region d, an alternative 3' splice site choice at exon 17/18 [exon 9] (Fig. **5A**). Three reverse-transcriptase polymerase chain reaction (RT-PCR) assays were performed on RNA from nineteen normal human tissues (Fig. **5B**) to examine these possibilities.
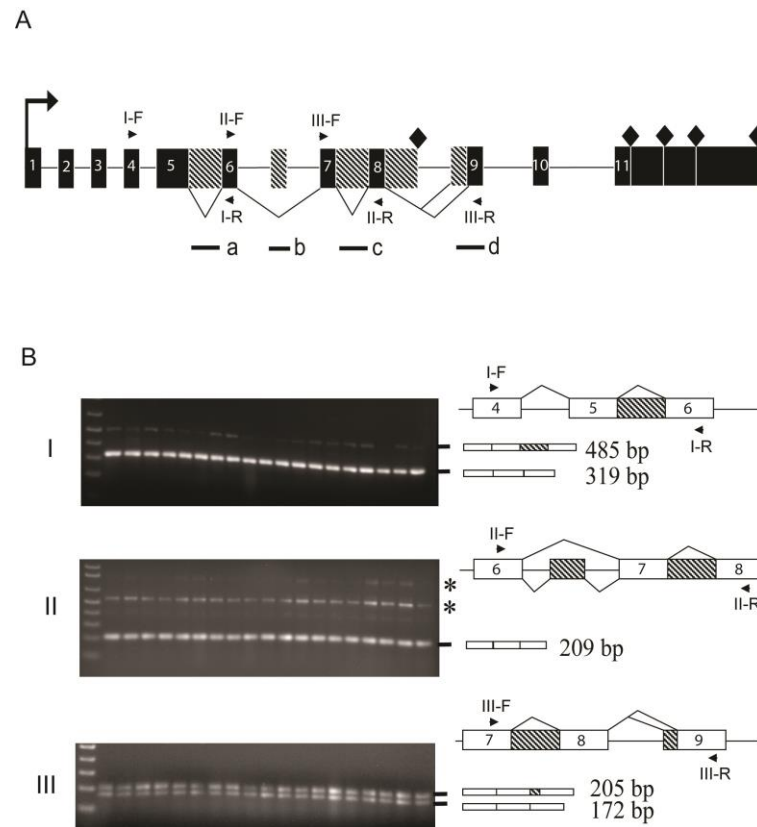
Figure 5. Human RNPEPL1 is alternatively spliced in normal tissues.

Alternative splicing events (a-d) within four regions of the RNPEPL1 pre-

mRNA (A) were investigated using three RT-PCR assays. Exons are shown

as solid boxes and numbered relative to the reference sequence

NM_018226; alternative exons are shown as hatched boxes. PCR products

(B) resulting from reaction I (region a using I-F and I-R primers), reaction II

(regions b and c using II-F and II-R primers), and reaction III (regions c and

d using III-F and III-R primers) were fractionated on 1.5% agarose gels for

nineteen normal human tissues (left to right: adipose, brain, cervix, colon,

esophagus, heart, kidney, liver, lung, ovary, placenta, prostate, skeletal muscle, small intestine, spleen, testes, thymus, thyroid, and trachea). PCR products confirmed by sequencing are diagrammed on the right. *=nonspecific product due to mispriming; ♦ =potential polyadenylation site. Drawings are not to scale.

PCR products were observed in all nineteen tissues in all three reactions, indicating a ubiquitous tissue distribution of RNPEPL1 mRNA. These reactions also revealed that two of the potential alternative splicing choices were utilized, and that these choices were made in all nineteen tissues examined. The first reaction revealed PCR products indicating both the inclusion (485 bp) and exclusion (319 bp) of intron 5 [intron 5] (Fig **5A** region a, Fig. **5B** reaction I) in all tissues tested. The inclusion of intron 5 [intron 5] was predicted to produce a truncated protein that lacks substantial portions of the catalytic pocket (Fig. **5B** reaction I, Fig. **6** I5_inc).

```
RNPEPL1     1  AQRRITTETYGAAFTCLETAFRLDALHRQMKLLGEDSPVSKLQVKLEPGVNPSHLMNLFT
alt_3'ss    1  AQRRITTETYGAAFTCLETAFRLDALHRQMKLLGEDSPVSKLQVKLEPGVNPSHLMNLFT
I5_inc      1  AQRRITTETYGAARAGEGSHGGPWALVQGQRA............................

RNPEPL1    61  YEKGYCFVYYLSQLCGDPQRFDDFLRAYVEKYKFTSVVAQDLLDSFLSFFPELKEQSVDC
alt_3'ss   61  YEKGYCFVYYLSQLCGDPQRFDDFLRAYVEKYKFTSVVAQDLLDSFLSFFPELKEQSVDC
I5_inc         ............................................................

RNPEPL1   121  RAGLEFERWLNATGPPLAEPDLSQGSSLTRPVEALFQLWTAEPLDQAAASASAIDISKWR
alt_3'ss  121  RA..........GPPLAEPDLSQGSSLTRPVEALFQLWTAEPLDQAAASASAIDISKWR
I5_inc         ............................................................

RNPEPL1   181  TFQTALFLDRLLDGSPLPQEVVMSLSKCYSSLLDSMNAEIRIRWLQIVVRNDYYPDLHRV
alt_3'ss  170  TFQTALFLDRLLDGSPLPQEVVMSLSKCYSSLLDSMNAEIRIRWLQIVVRNDYYPDLHRV
I5_inc         ............................................................

RNPEPL1   241  RRFLESQMSRMYTIPLYEDLCTGALKSFALEVFYQTQGRLHPNLRRAIQQILSQGLGSST
alt_3'ss  230  RRFLESQMSRMYTIPLYEDLCTGALKSFALEVFYQTQGRLHPNLRRAIQQILSQGLGSST
I5_inc         ............................................................

RNPEPL1   301  EPASEPSTELGKAEADTDSDAQALLLGDEAPSSAISLRDVNVSA
alt_3'ss  290  EPASEPSTELGKAEADTDSDAQALLLGDEAPSSAISLRDVNVSA
I5_inc         ...........................................
```

Figure 6. Predicted protein isoforms of RNPEPL1 generated by alternatively spliced mRNAs. The predicted mRNAs resulting from alternative splicing events observed in normal human tissues were translated *in silico*. Inclusion of intron 5 (I5_inc) is predicted to result in a truncated protein, while usage of the downstream 3' splice site at exon 17/18 (alt_3'ss) was predicted to omit 11 amino acids (Gly[273]-Thr[283]) from the protein.

Although neither exclusion of exon 11 [alternative exon within intron 6] (Fig. **5A** region b, Fig. **5B** reaction II) nor inclusion of intron 8 [intron 7] (Fig. **5A** region c, Fig. **5B** reaction III) occurred in any of the tissues tested, the third reaction indicated that a 3'splice site choice at exon 17/18 [exon 9] was utilized (Fig. **5A** region d, Fig **5B** reaction III). PCR products

indicating the usage of both the upstream 3' splice site (205 bp) and the downstream 3' splice site (172 bp) were observed in all tissues examined. When the downstream splice site is utilized, an 11 amino acid deletion is predicted to result (Fig. **6** alt_3'ss).

DISCUSSION

Large-scale sequencing of the human genome revealed a novel gene, RNPEPL1, encoding a protein with extensive sequence homology to arginyl aminopeptidase. Although expression of an mRNA from this gene has been reported [19-22], no attempt to characterize the gene or encoded protein has been made until now. To this end, primers were designed to amplify the RNPEPL1 cDNA and subclone it into expression vectors. Expression and purification of the recombinant protein from both yeast and Sf9 cells, followed by removal of the his$_6$ affinity tag, yielded a protein of approximately 55 kDa as predicted from the cDNA sequence (Fig. **1**). Affinity tag removal was deemed necessary because of previous studies with human NPEPPS and yeast LTA$_4$H that demonstrated interference of the his$_6$ epitope tag with catalytic activity, raising the K_M and lowering the k_{cat} for amino acid β-naphthylamide substrates [32,38].

This step further purified the recombinant protein to near homogeneity by removing impurities that interacted nonspecifically with the affinity resin.

Like most members of the M1 metallopeptidase family, RNPEPL1 exhibits a broad substrate specificity, preferring a P1 methionine, glutamine, or citrulline residue, although alanine and leucine were also suitable substrates (Fig. **2**, Table **1**). While the closely-related RNPEP shows a preference for an N-terminal arginine or lysine residue [39], these amino acids were poor substrates for RNPEPL1, indicating that it is catalytically distinct despite its close sequence homology. RNPEPL1 activity is more similar to that of LTA_4H, which preferentially hydrolyzes a P1 alanine or arginine residue, but may also cleave proline, leucine, methionine, and lysine residues at this position. Unlike LTA_4H, however, RNPEPL1 did not cleave proline, and arginine and lysine were relatively poor substrates (Fig. **2**, Table **1**). Together, these data indicate that RNPEPL1 is a distinct cytoplasmic aminopeptidase with an activity that overlaps significantly with that of other peptidases. Redundant peptidase activities likely ensure the complete breakdown of peptides in the cytoplasm and indicate a possible housekeeping function for RNPEPL1, consistent with its ubiquitous tissue distribution (Fig. **5B**).

The effects of inhibitors on RNPEPL1 activity were also determined to further compare RNPEPL1 to other members of the M1 family. Predictably, RNPEPL1 was moderately sensitive to bestatin, exhibiting a K_i value in the low micromolar range (Table **2**). In this regard, RNPEPL1 demonstrated similarity to LTA_4H, which also exhibits a K_i in the low micromolar range for this inhibitor [40], whereas this inhibitor binds much more tightly to RNPEP [41]. RNPEPL1 also exhibited similarity to LTA_4H in that it was very sensitive to amastatin (K_i=0.19 μM), unlike RNPEP, which is only weakly inhibited by this compound [42]. This is likely due to differences in substrate specificity. The first two side chains of the tripeptide analog amastatin are valine, which are not well recognized at the P1 or P1' position by RNPEP [43]. Although a P1 valine was not recognized by RNPEPL1, it is likely that a P1' valine may be recognized.

Reports in the literature indicate that the activities of some members of the M1 family are modulated by calcium ions. For example, glutamyl aminopeptidase (ENPEP) exhibits a stronger preference for substrates containing acidic amino acids in the presence of calcium [35]. This effect is due to coordination of the metal between the side chain of Asp^{221} and the P1 side chain of the substrate [44]. Similarly, the catalytic activity of the puromycin-sensitive aminopeptidase is accelerated by calcium

[45,46], but most likely by a different mechanism since Asp[221] is replaced

by glutamine in this enzyme. To investigate the possibility that RNPEPL1

catalytic activity is affected by Ca^{2+}, the activity of the enzyme towards

Met-AMC (Fig. **4B**) or Cit-AMC (data not shown) was determined in the

presence of increasing levels of Ca^{2+}. Surprisingly, RNPEPL1 was inhibited

by Ca^{2+} with an apparent K_i of 51.0 mM, indicating that in some instances,

changes in cytoplasmic calcium levels could possibly modulate the

activity of this enzyme. RNPEPL1 activity would be expected to remain

high under normal circumstances, since the intracellular concentration of

calcium within most cells is generally around 0.1 μM [47], with

concentrations rising about an order of magnitude as a result of

intracellular signaling. At these concentrations, it is unlikely that Ca^{2+}

concentrations would increase enough to temporarily inhibit RNPEPL1

activity. However, the occurrence of more intense releases of Ca^{2+} potent

enough to inhibit RNPEPL1 in some tissues cannot be ruled out. For

example, loss of membrane integrity, as observed in synoviocytes of

rheumatoid arthritis patients [48], may allow Ca^{2+} levels to rise high

enough to cause RNPEPL1 inhibition. Furthermore, observations that

citrullinated peptides are found on class I and II MHC in the cells of

patients with rheumatoid arthritis, coupled with the finding that RNPEPL1

exhibits high activity towards peptides containing an N-terminal citrulline residue, suggest that RNPEPL1 and other similar peptidases may trim such peptides in the cytoplasm of cells. It is possible that inhibition of this activity could contribute to the pathogenesis of the disorder.

Recent studies of alternative splicing suggest that more than 60% of human genes produce multiple mRNAs, which sometimes encode alternative protein isoforms. These isoforms may perform regulatory functions in the cell [49] by encoding truncated proteins, as in the case of telomerase [50], by encoding non-functional proteins, as in the case of hyaluronidase [51], by altering the function of the protein, as in the case of bcl-2 gene family members [52-55], or by altering the subcellular localization of the protein, as in the case with immunoglobulins [56] and aminopeptidase O [57]. Furthermore, gene expression can be regulated by producing mRNAs by alternative splicing that are subject to nonsense mediated decay, as occurs with the nucleolar protein NOL5 [58]. Within the M1 family, alternative mRNA splicing resulting in isoforms of aminopeptidase O with altered properties have been reported [57]; thus, it is likely that other members of this family are similarly regulated.

In order to investigate possible alternative splicing of the RNPEPL1 mRNA, RT-PCR analysis of nineteen normal human tissues was performed.

Transcripts were detected in all tissues examined, and two alternative

splicing events were observed. The first alternative splicing event, inclusion

of intron 5 [intron 5] (Fig. **5B** reaction I), was predicted to produce an

mRNA isoform containing a premature stop codon at codon 181 (Fig. **6**

I5_inc), which would produce a truncated protein product lacking

catalytic activity. Although some tissues may appear to have different

ratios of the two products, this type of analysis is qualitative, not

quantitative. Still, conditions that favor the inclusion of intron 5 could thus

reduce the amount of catalytically functional RNPEPL1 activity in the cell

or tissue, or might limit the amount of functional enzyme available to bind

an interacting protein.

The second alternative splicing event observed was the utilization of

an alternative downstream 3' splice site (Fig. **5B** reaction III) that is

predicted to produce an mRNA isoform encoding a protein lacking

Gly^{273}-Thr^{283} (Fig. **6** alt_3'ss). This stretch of amino acids is conserved among

the enzymes of the leukotriene A_4 hydrolase-like group (data not shown)

and occurs immediately upstream of a region of RNPEPL that is

homologous to a putative armadillo repeat in aminopeptidase O [59,60].

Although the effect of omitting these conserved eleven amino acids on

catalytic activity is still unknown, it is possible that their deletion may alter

protein-protein interactions, change the subcellular localization of the protein, or even modify the enzyme's catalytic activity. Recent reports that the C-terminal domains of glutamyl aminopeptidase [61] and of aminopeptidase N [62] may act as chaperones and regulate enzymatic activity favor the latter two conclusions.

ACKNOWLEDGEMENTS

REFERENCES

[1] Thompson, M.W., Tobler, A., Fontana, A., and Hersh, L.B. Cloning and analysis of the gene for the human puromycin-sensitive aminopeptidase. *Biochemical and Biophysical Research Communications*, **1999**, *258*, 234-40.

[2] Zini, S., Fournie-Zaluski, M.C., Chauvel, E., Roques, B.P., Corvol, P., and Llorens-Cortes, C. Identification of metabolic pathways of brain angiotensin II and III using specific aminopeptidase inhibitors: predominant role of angiotensin III in the control of vasopressin release. *Proceedings of the National Academy of Science U S A*, **1996**, *93*, 11968-73.

[3] Stoltze, L., Schirle, M., Schwarz, G., Schroter, C., Thompson, M.W., Hersh, L.B., Kalbacher, H., Stevanovic, S., Rammensee, H.G., and Schild, H. Two new proteases in the MHC class I processing pathway. *Nature Immunology*, **2000**, *1*, 413-8.

[4] Yan, J., Parekh, V.V., Mendez-Fernandez, Y., Olivares-Villagomez, D., Dragovic, S., Hill, T., Roopenian, D.C., Joyce, S., and Van Kaer, L. In vivo role of ER-associated peptidase activity in tailoring peptides for presentation by MHC class Ia and class Ib molecules. *Journal of Experimental Medicine*, **2006**, *203*, 647-59.

[5] Reaux, A., Fournie-Zaluski, M.C., David, C., Zini, S., Roques, B.P., Corvol, P., and Llorens-Cortes, C. Aminopeptidase A inhibitors as potential central antihypertensive agents. *Proceedings of the National Academy of Science U S A*, **1999,** *96*, 13415-20.

[6] Whittle, B.J., Varga, C., Berko, A., Horvath, K., Posa, A., Riley, J.P., Lundeen, K.A., Fourie, A.M., and Dunford, P.J. Attenuation of inflammation and cytokine production in rat colitis by a novel selective inhibitor of leukotriene A4 hydrolase. *British Journal of Pharmacology*, **2008,** *153*, 983-91.

[7] Tsukamoto, H., Shibata, K., Kajiyama, H., Terauchi, M., Nawa, A., and Kikkawa, F. Aminopeptidase N (APN)/CD13 inhibitor, Ubenimex, enhances radiation sensitivity in human cervical cancer. *BMC Cancer*, **2008**, *8*, 74.

[8] Urabe, A., Mutoh, Y., Mizoguchi, H., Takaku, F., and Ogawa, N. Ubenimex in the treatment of acute nonlymphocytic leukemia in adults. *Annals of Hematology*, **1993**, *67*, 63-6.

[9] Miller, B.C., Ackroyd, A., Hersh, L.B., and Cottam, G.L. Methionine enkephalin metabolism by murine macrophage ectopeptidase(s). *Regulatory Peptides*, **1994**, *50*, 87-98.

[10] Papadopoulos, T., Heuer, H., and Bauer, K. Analysis of the thyrotropin-releasing hormone-degrading ectoenzyme by site-directed mutagenesis of cysteine residues. Cys68 is involved in disulfide-linked dimerization. *European Journal of Biochemistry*, **2000**, *267*, 2617-23.

[11] Evnouchidou, I., Momburg, F., Papakyriakou, A., Chroni, A., Leondiadis, L., Chang, S.C., Goldberg, A.L., and Stratikos, E. The internal sequence of the peptide-substrate determines its N-terminus trimming by ERAP1. *PLoS ONE*, **2008**, *3*, e3658.

[12] Serwold, T., Gonzalez, F., Kim, J., Jacob, R., and Shastri, N. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, **2002**, *419*, 480-3.

[13] Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G., and van Endert, P.M. Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nature Immunology*, **2005**, *6*, 689-97.

[14] Albiston, A.L., McDowall, S.G., Matsacos, D., Sim, P., Clune, E., Mustafa, T., Lee, J., Mendelsohn, F.A., Simpson, R.J., Connolly, L.M., and Chai, S.Y. Evidence that the angiotensin IV (AT(4)) receptor is the enzyme insulin-regulated aminopeptidase. *Journal of Biological Chemistry*, **2001**, *276*, 48623-6.

[15] Lew, R.A., Mustafa, T., Ye, S., McDowall, S.G., Chai, S.Y., and Albiston, A.L. Angiotensin AT4 ligands are potent, competitive inhibitors of insulin regulated aminopeptidase (IRAP). *Journal of Neurochemistry*, **2003**, *86*, 344-50.

[16] Hui, M., and Hui, K.S. Neuron-specific aminopeptidase and puromycin-sensitive aminopeptidase in rat brain development. *Neurochemical Research*, **2003**, *28*, 855-60.

[17] Osada, T., Watanabe, G., Kondo, S., Toyoda, M., Sakaki, Y., and Takeuchi, T. Male reproductive defects caused by puromycin-sensitive aminopeptidase deficiency in mice. *Molecular Endocrinology*, **2001**, *15*, 960-71.

[18] Osada, T., Watanabe, G., Sakaki, Y., and Takeuchi, T. Puromycin-sensitive aminopeptidase is essential for the maternal recognition of pregnancy in mice. *Molecular Endocrinology*, **2001**, *15*, 882-93.

[19] Gerhard, D.S., Wagner, L., Feingold, E.A., Shenmen, C.M., Grouse, L.H., Schuler, G., Klein, S.L., Old, S., Rasooly, R., Good, P., Guyer, M., Peck, A.M., Derge, J.G., Lipman, D., Collins, F.S., Jang, W., Sherry, S., Feolo, M., Misquitta, L., Lee, E., Rotmistrovsky, K., Greenhut, S.F., Schaefer, C.F., Buetow, K., Bonner, T.I., Haussler, D., Kent, J., Kiekhaus, M., Furey, T., Brent, M., Prange, C., Schreiber, K., Shapiro, N., Bhat, N.K., Hopkins, R.F., Hsie, F., Driscoll, T., Soares, M.B., Casavant, T.L., Scheetz, T.E., Brown-stein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Piao, Y., Dudekula, D.B., Ko, M.S., Kawakami, K., Suzuki, Y., Sugano, S., Gruber, C.E., Smith, M.R., Simmons, B., Moore, T., Waterman, R., Johnson, S.L., Ruan, Y., Wei, C.L., Mathavan, S., Gunaratne, P.H., Wu, J., Garcia, A.M., Hulyk, S.W., Fuh, E., Yuan, Y., Sneed, A., Kowis, C., Hodgson, A., Muzny, D.M., McPherson, J., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madari, A., Young, A.C., Wetherby, K.D., Granite, S.J., Kwong, P.N., Brinkley, C.P., Pearson, R.L., Bouffard, G.G., Blakesly, R.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Griffith, M., Griffith, O.L., Krzywinski, M.I., Liao, N., Morin, R., Palmquist, D., et al. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Research*, **2004**, *14*, 2121-7.

[20] Hillier, L.W., Graves, T.A., Fulton, R.S., Fulton, L.A., Pepin, K.H., Minx, P., Wagner-McPherson, C., Layman, D., Wylie, K., Sekhon, M., Becker, M.C., Fewell, G.A., Delehaunty, K.D., Miner, T.L., Nash, W.E., Kremitzki, C., Oddy, L., Du, H., Sun, H., Bradshaw-Cordum, H., Ali, J., Carter, J., Cordes, M., Harris, A., Isak, A., van Brunt, A., Nguyen, C., Du, F., Courtney, L., Kalicki, J., Ozersky, P., Abbott, S., Armstrong, J., Belter, E.A., Caruso, L., Cedroni, M., Cotton, M., Davidson, T., Desai, A., Elliott, G., Erb, T., Fronick, C., Gaige, T., Haakenson, W., Haglund, K., Holmes, A., Harkins, R., Kim, K., Kruchowski, S.S., Strong, C.M., Grewal, N., Goyea, E., Hou, S., Levy, A., Martinka, S., Mead, K., McLellan, M.D., Meyer, R., Randall-Maher, J., Tomlinson, C., Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Shah, N., Swearengen-Shahid, S., Snider, J., Strong, J.T., Thompson, J., Yoakum, M., Leonard, S., Pearman, C., Trani, L., Radionenko, M., Waligorski, J.E., Wang, C., Rock, S.M., Tin-Wollam, A.M., Maupin, R., Latreille, P., Wendl, M.C., Yang, S.P., Pohl, C., Wallis, J.W., Spieth, J., Bieri, T.A., Berkowicz, N., Nelson, J.O., Osborne, J., Ding, L., Meyer, R., Sabo, A., Shotland, Y., Sinha, P., Wohldmann, P.E., Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Jones, T.A., She, X., Ciccarelli, F.D., et al. Generation and annotation of the DNA sequences of human chromosomes 2 and 4. *Nature*, **2005**, *434*, 724-31.

[21] Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T.O., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., et al. Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nature Genetics*, **2004**, *36*, 40-5.

[22] Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proceedings of the National Academy of Science U S A*, **2002**, *99*, 16899-903.

[23] Haeggstrom, J.Z. Leukotriene A4 hydrolase/aminopeptidase, the gatekeeper of chemotactic leukotriene B4 biosynthesis. *Journal of Biological Chemistry*, **2004**, *279*, 50639-42.

[24] Helgadottir, A., Manolescu, A., Helgason, A., Thorleifsson, G., Thorsteinsdottir, U., Gudbjartsson, D.F., Gretarsdottir, S., Magnusson, K.P., Gudmundsson, G., Hicks, A., Jonsson, T., Grant, S.F., Sainz, J., O'Brien, S.J., Sveinbjornsdottir, S., Valdimarsson, E.M., Matthiasson, S.E., Levey, A.I., Abramson, J.L., Reilly, M.P., Vaccarino, V., Wolfe, M.L., Gudnason, V., Quyyumi, A.A., Topol, E.J., Rader, D.J., Thorgeirsson, G., Gulcher, J.R., Hakonarson, H., Kong, A., and Stefansson, K. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nature Genetics*, **2006**, *38*, 68-74.

[25] Cadel, S., Pierotti, A.R., Foulon, T., Creminon, C., Barre, N., Segretain, D., and Cohen, P. Aminopeptidase-B in the rat testes: isolation, functional properties and cellular localization in the seminiferous tubules. *Molecular and Cellular Endocrinology*, **1995**, *110*, 149-60.

[26] Cohen, P. Proteolytic events in the post-translational processing of polypeptide hormone precursors. *Biochimie*, **1987**, *69*, 87-9.

[27] Gainer, H., Russell, J.T., and Loh, Y.P. An aminopeptidase activity in bovine pituitary secretory vesicles that cleaves the N-terminal arginine from beta-lipotropin60-65. *FEBS Letters*, **1984**, *175*, 135-9.

[28] Fontes, G., Lajoix, A.D., Bergeron, F., Cadel, S., Prat, A., Foulon, T., Gross, R., Dalle, S., Le-Nguyen, D., Tribillac, F., and Bataille, D. Miniglucagon (MG)-generating endopeptidase, which processes glucagon into MG, is composed of N-arginine dibasic convertase and aminopeptidase B. *Endocrinology*, **2005**, *146*, 702-12.

[29] Cadel, S., Foulon, T., Viron, A., Balogh, A., Midol-Monnet, S., Noel, N., and Cohen, P. Aminopeptidase B from the rat testis is a bifunctional enzyme structurally related to leukotriene-A4 hydrolase. *Proceedings of the National Academy of Science U S A*, **1997**, *94*, 2963-8.

[30] Fukasawa, K.M., Fukasawa, K., Harada, M., Hirose, J., Izumi, T., and Shimizu, T. Aminopeptidase B is structurally related to leukotriene-A4 hydrolase but is not a bifunctional enzyme with epoxide hydrolase activity. *Biochemical Journal*, **1999**, *339 (Pt 3)*, 497-502.

[31] Horikawa, Y., Oda, N., Cox, N.J., Li, X., Orho-Melander, M., Hara, M., Hinokio, Y., Lindner, T.H., Mashima, H., Schwarz, P.E., del Bosque-Plata, L., Horikawa, Y., Oda, Y., Yoshiuchi, I., Colilla, S., Polonsky, K.S., Wei, S., Concannon, P., Iwasaki, N., Schulze, J., Baier, L.J., Bogardus, C., Groop, L., Boerwinkle, E., Hanis, C.L., and Bell, G.I. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nature Genetics*, **2000**, *26*, 163-75.

[32] Thompson, M.W., Archer, E.D., Romer, C.E., and Seipelt, R.L. A conserved tyrosine residue of Saccharomyces cerevisiae leukotriene A4 hydrolase stabilizes the transition state of the peptidase activity. *Peptides*, **2006**, *27*, 1701-9.

[33] Fukasawa, K.M., Fukasawa, K., Kanai, M., Fujii, S., and Harada, M. Molecular cloning and expression of rat liver aminopeptidase B. *Journal of Biological Chemistry*, **1996**, *271*, 30731-5.

[34] Wetterholm, A., Medina, J.F., Radmark, O., Shapiro, R., Haeggstrom, J.Z., Vallee, B.L., and Samuelsson, B. Leukotriene A4 hydrolase: abrogation of the peptidase activity by mutation of glutamic acid-296. *Proceedings of the National Academy of Science U S A*, **1992**, *89*, 9141-5.

[35] Goto, Y., Hattori, A., Ishii, Y., Mizutani, S., and Tsujimoto, M. Enzymatic properties of human aminopeptidase A. Regulation of its enzymatic activity by calcium and angiotensin IV. *Journal of Biological Chemistry*, **2006**, *281*, 23503-13.

[36] Thierry-Mieg, D., Thierry-Mieg, J., Potdevin, M., Sienkiewicz, M., and V, S. AceView: Identification and functional annotation of cDNA-supported genes in eukaryotes. *Aceview* [Accessed 2007-2009]. www.ncbi.nlm.nih.gov/IEB/Research/Acembly

[37] Karolchik D, Kuhn, R., Baertsch R, Barber GP, Clawson H, Diekhans M, Giardine B, Harte RA, Hinrichs AS, Hsu F, Miller W, Pedersen JS, Pohl A, Raney BJ, Rhead B, Rosenbloom KR, Smith KE, Stanke M, Thakkapallayil A, Trumbower H, Wang T, Zweig AS, Haussler D, and WJ., K. The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Research*, **2008**, *36*, D773-9.

[38] Thompson, M.W., Govindaswami, M., and Hersh, L.B. Mutation of active site residues of the puromycin-sensitive aminopeptidase: conversion of the enzyme into a catalytically inactive binding protein. *Archives of Biochemistry and Biophysics*, **2003**, *413*, 236-42.

[39] Cadel, S., Gouzy-Darmon, C., Petres, S., Piesse, C., Pham, V.L., Beinfeld, M.C., Cohen, P., and Foulon, T. Expression and purification of rat recombinant aminopeptidase B secreted from baculovirus-infected insect cells. *Protein Expression and Purification*, **2004**, *36*, 19-30.

[40] Muskardin, D.T., Voelkel, N.F., and Fitzpatrick, F.A. Modulation of pulmonary leukotriene formation and perfusion pressure by bestatin, an inhibitor of leukotriene A4 hydrolase. *Biochemical Pharmacology*, **1994**, *48*, 131-7.

[41] Ocain, T.D., and Rich, D.H. Synthesis of sulfur-containing analogues of bestatin. Inhibition of aminopeptidases by alpha-thiolbestatin analogues. *Journal of Medicinal Chemistry*, **1988**, *31*, 2193-9.

[42] Nagata, Y., Mizutani, S., Nomura, S., Kurauchi, O., Kasugai, M., and Tomoda, Y. Purification and properties of human placental aminopeptidase B. *Enzyme*, **1991**, *45*, 165-73.

[43] Foulon, T., Cadel, S., Piesse,C. Cohen,P., Aminopeptidase B. in: Barrett, A.J., Rawlings,N.D. Woessner,J.F., (Ed.), Handbook of Proteolytic Enzymes, Elsevier, London, 2004, pp. 328-332.

[44] Goto, Y., Hattori, A., Mizutani, S., and Tsujimoto, M. Asparatic acid 221 is critical in the calcium-induced modulation of the enzymatic activity of human aminopeptidase A. *Journal of Biological Chemistry*, **2007**, *282*, 37074-81.

[45] Dando, P.M., Young, N.E., Barrett A.J., Aminopeptidase PS: A Widely Distributed Cytosolic Peptidase. in: Hopsu-Havn V.K., J.N., Kirschke H., (Ed.), Proteolysis in Cell Functions, IOS Press, Amsterdam, **1997**, pp. 88-95.

[46] Backstrom, J.R., Miller, C.A., and Tokes, Z.A. Characterization of neutral proteinases from Alzheimer-affected and control brain specimens: identification of calcium-dependent metalloproteinases from the hippocampus. *Journal of Neurochemistry*, **1992**, *58*, 983-92.

[47] Nussey, S., and Whitehead, SA, Endocrinology: An Integrated Approach, Taylor & Francis, London, 2001.

[48] Kinloch, A., Lundberg, K., Wait, R., Wegner, N., Lim, N.H., Zendman, A.J., Saxne, T., Malmstrom, V., and Venables, P.J. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. *Arthritis & Rheumatology*, **2008**, *58*, 2287-95.

[49] Kim, E., Magen, A., and Ast, G. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Research*, **2007**, *35*, 125-31.

[50] Lincz, L.F., Mudge, L.M., Scorgie, F.E., Sakoff, J.A., Hamilton, C.S., and Seldon, M. Quantification of hTERT splice variants in melanoma by SYBR green real-time polymerase chain reaction indicates a negative regulatory role for the beta deletion variant. *Neoplasia*, **2008**, *10*, 1131-7.

[51] Lokeshwar, V.B., Schroeder, G.L., Carey, R.I., Soloway, M.S., and Iida, N. Regulation of hyaluronidase activity by alternative mRNA splicing. *Journal of Biological Chemistry*, **2002**, *277*, 33654-63.

[52] Bae, J., Leo, C.P., Hsu, S.Y., and Hsueh, A.J. MCL-1S, a splicing variant of the antiapoptotic BCL-2 family member MCL-1, encodes a proapoptotic protein possessing only the BH3 domain. *Journal of Biological Chemistry*, **2000**, *275*, 25255-61.

[53] Hossini, A.M., Geilen, C.C., Fecker, L.F., Daniel, P.T., and Eberle, J. A novel Bcl-x splice product, Bcl-xAK, triggers apoptosis in human melanoma cells without BH3 domain. *Oncogene*, **2006**, *25*, 2160-9.

[54] Renshaw, S.A., Dempsey, C.E., Barnes, F.A., Bagstaff, S.M., Dower, S.K., Bingle, C.D., and Whyte, M.K. Three novel Bid proteins generated by alternative splicing of the human Bid gene. *Journal of Biological Chemistry*, **2004**, *279*, 2846-55.

[55] U, M., Miyashita, T., Shikama, Y., Tadokoro, K., and Yamada, M. Molecular cloning and characterization of six novel isoforms of human Bim, a member of the proapoptotic Bcl-2 family. *FEBS Letters*, **2001**, *509*, 135-41.

[56] Seipelt, R.L., Spear, B.T., Snow, E.C., and Peterson, M.L. A nonimmunoglobulin transgene and the endogenous immunoglobulin mu gene are coordinately regulated by alternative RNA processing during B-cell maturation. *Molecular and Cellular Biology*, **1998**, *18*, 1042-8.

[57] Axton, R., Wallis, J.A., Taylor, H., Hanks, M., and Forrester, L.M. Aminopeptidase O contains a functional nucleolar localization signal and is implicated in vascular biology. *Journal of Cellular Biochemistry*, **2008**, *103*, 1171-82.

[58] Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., and Ares, M., Jr. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & Development*, **2007**, *21*, 708-18.

[59] Diaz-Perales, A., Quesada, V., Sanchez, L.M., Ugalde, A.P., Suarez, M.F., Fueyo, A., and Lopez-Otin, C. Identification of human aminopeptidase O, a novel metalloprotease with structural similarity to aminopeptidase B and leukotriene A4 hydrolase. *Journal of Biological Chemistry*, **2005**, *280*, 14310-7.

[60] Zdobnov, E.M., and R, A.A. InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*, **2001**, *17*, 847-8.

[61] Rozenfeld, R., Muller, L., El Messari, S., and Llorens-Cortes, C. The C-terminal domain of aminopeptidase A is an intramolecular chaperone required for the correct folding, cell surface expression, and activity of this monozinc aminopeptidase. *Journal of Biological Chemistry*, **2004**, *279*, 43285-95.

[62] Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T., and Yoshimoto, T. Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from Escherichia coli and conformational change of methionine 260 involved in substrate recognition. *Journal of Biological Chemistry*, **2006**, *281*, 33664-76.

CHAPTER THREE: EVOLUTIONARY ANALYSIS OF THE MAMMALIAN M1 AMINOPEPTIDASES REVEALS CONSERVED EXON STRUCTURE AND GENE DEATH

Karen B. Maynard, Shannon Smith, Anthony Davis, Andrew Trivette, and Rebecca L. Seipelt-Thiemann

Program for Molecular Biosciences, Biology Department, Middle Tennessee State University, Murfreesboro, TN, USA

ABSTRACT

The M1 aminopeptidase family share conserved domains, yet show functional divergence within the family as a whole. In order to better understand this family, this study analyzed the mammalian members in depth at exon, gene, and protein levels. The twelve human members, eleven rat members, and eleven mouse members were first analyzed in multiple alignments to visualize both reported and unreported conserved domains. Phylogenetic trees were then generated for human, rat, mouse, and all mammals to determine how closely related the homologs were and to gain insight to the divergence in the family members. This produced three groups with similarity within the family. Next, a synteny study was completed to determine the present locations of the genes and changes that had occurred. It became apparent that gene death likely resulted in the lack of one member in mouse and rat. Finally, an in-depth analysis of the exon structure revealed that nine members of the human family and eight in mouse, are highly conserved within the exon structure. Taken together, these results indicate that the M1 aminopeptidase family is a divergent family with three subgroups and that

genetic evidence mirrors categorization of the family by enzymatic function.

INTRODUCTION

The rapid development of sequencing technologies and the subsequent ease and speed of genomic sequence data discovery has given researchers the ability to detect large scale differences in the genomes of organisms to illuminate both answers and new questions [1]. The availability of this data has led to an increasing interest in and the ability to study genome and gene family evolution. The evolutionary history of a number of gene families have been clarified due to the availability of data, such as the calpains [2] and the insulin growth factor binding proteins [3]. A gene family is a group of genes that were generated by gene and/or genome duplication of an ancestral gene. Over time, these duplicate genes underwent additional duplications, mutations, or gene rearrangements, creating a family of genes with multiple members that retain sequence, structural, and ultimately functional similarities. Genes can also degenerate to pseudogenes due to accumulation of mutations or rearrangements. More recent duplicates

will have fewer changes than more distant duplicates, while still retaining some features of the ancestral gene, either by amino acid similarity, chromosomal context, gene structure, and/or function. The level of divergence among members of the gene family can be utilized to predict the relatedness or evolutionary history of its members. For a review of genome architecture and the forces that alter it, see Koonin 2009 [4].

One gene family for which sequencing data has become available to study gene family evolutionary history is the M1 metalloproteinase gene family. This gene family has twelve members in the human genome and eleven in both mouse and rat genomes.   Three conserved protein motifs identify aminopeptidases as part of the M1 aminopeptidase family. All members share a HEXXHX18E motif that participates in zinc-binding, which is a hallmark of the metalloproteinase. Additionally, there is a conserved G/AMEN motif that is believed to be the substrate binding domain. Finally, all members contain two conserved active site residues. The first is the glutamate (E) residue of the HEXXH zinc binding domain. The second is a distal tyrosine (Y) residue. The human family members that fulfill these requirements are: placental leucine aminopeptidase (LNPEP), endoplasmic reticulum aminopeptidase 1 (ERAP1), endoplasmic reticulum aminopeptidase 2 (ERAP2), laeverin/aminopeptidase Q (APQ),

puromycin-sensitive aminopeptidase (NPEPPS), thyrotropin-releasing hormone-degrading ectoenzyme (TRHDE), aminopeptidase A (ENPEP), aminopeptidase N (ANPEP), aminopeptidase O (APO), leukotriene A4 hydrolase (LTA$_4$H), arginyl aminopeptidase B (RNPEP), and arginyl aminopeptidase B like-1 (RNPEPL1). Each member is known or predicted to have enzymatic activity to remove single amino acids from the amino terminal ends of proteins and peptides. In addition, several are known to be bifunctional. Members of this gene family are known to have important roles in the angiotensin system, the immune system, and in inflammation. For example, LNPEP, ERAP1, ERAP2, APQ, ENPEP, and APO all have roles in activation or inactivation of various components of the angiotensin system [5 - 10]. M1 aminopeptidases play roles in important biological processes from pregnancy (LNPEP and APQ) [5, 10], pain and apoptosis (NPEPPS) [9, 11], and blood pressure regulation (ENPEP) [8] to inflammation (LTA$_4$H) [12] and antigen presentation (LNPEP, ERAP1 and ERAP2) [13].

In order to examine the evolutionary history of this gene family, a detailed comparison of the phylogenetics, chromosomal context, and gene structure of the family members in human, mouse, and rat was undertaken. These comparative analyses provide a picture of an old

gene family with three distinct groups with multiple members that share various aspects of gene structure, amino acid similarity, and genome organization. In addition the absence of one family member, ERAP2, in both mouse and rat, is likely explained by a chromosomal rearrangement event.

METHODS

Multiple Alignment and Phylogenetic Tree Analysis

Human, mouse, and rat protein and nucleotide reference sequences were obtained from NCBI (Table 1). In cases where two or more isoforms were found, the most complete and/or extensive isoform was chosen to include all possible sequence traits. Alignments were created using ClustalW [14] and visually enhanced with Boxshade to allow for ease of analysis and identification of conserved domains. ClustalW was additionally used to align sequences within Molecular Evolutionary Genetics Analysis, MEGA 6 for production of phylogenetic trees [15]. MEGA 6 was then used to create and test multiple phylogenetic trees. The phylogenetic reconstructions were completed using maximum likelihood with bootstrap values over 500 simulations, with partial gap

deletions. For the protein sequence trees, the Jones-Taylor-Thronton (JTT)

model was used. All trees utilized Nearest Neighbor Interchange for tree

interference with very strong branch swap filters.

Table 1: Accession numbers and chromosome locations (NCBI) of the

twelve human (Hs), eleven mouse (Mm), and eleven rat (Rn) members of

the M1 aminopeptidase family.

Gene Name	Location	Nucleotide #	Amino Acid #	GI #
ANPEP_Hs	15q25-26	NM_001150.2	NP_001141.2	GI: 157266299
ANPEP_Mm	7	NM_008486.2	NP_032512.2	GI: 225637486
ANPEP_Rn	1q31	NM_031012.1	NP_112274.1	GI: 13591913
APO_Hs	9q22.32	NM_001193331.2	NP_001180260.1	GI: 308387394
APO_Mm	13	AJ810420.1	CAH17902.1	GI: 58737214
APO_Rn	17q14	NM_001012346.1	NP_001012346.1	GI: 59933289
APQ_Hs	5q23.1	NM_173800.4	NP_776161.3	GI: 194239712
APQ_Mm	18	NM_029008.1	NP_083284.1	GI: 254588009
APQ_Rn	18q11	XM_577617.4	XP_577617.3	GI: 564393383
ENPEP_Hs	4q25	NM_001977.3	NP_001968.3	GI: 132814466
ENPEP_Mm	3	NM_007934.3	NP_031960.1	GI: 12486800
ENPEP_Rn	2q42	NM_022251.2	NP_071587.2	GI: 140970580
ERAP1_Hs	5q15	NM_016442.3	NP_057526.3	GI: 94818900
ERAP1_Mm	13	NM_030711.4	NP_109636.1	GI: 146149238
ERAP1_Rn	2q11	NM_030836.1	NP_110463.1	GI: 13540637
ERAP2_Hs	5q15	NM_022350.3	NP_071745.1	GI: 194306627
LNPEP_Hs	5q15	NM_005575.2	NP_005566.2	GI: 61742776
LNPEP_Mm	17	NM_172827.3	NP_766415.1	GI: 146198621
LNPEP_Rn	1q12	NM_001113403.1	NP_001106874.1	GI: 164663800
LTA4H_Hs	12q22	NM_000895.2	NP_000886.1	GI: 375331875
LTA4H_Mm	10	NM_008517.2	NP_032543.2	GI: 116734869
LTA4H_Rn	7q13	NM_001030031.1	NP_001025202.1	GI: 71795620
NPEPPS_Hs	17q21	NM_006310.3	NP_006301.3	GI: 158937235
NPEPPS_Mm	11	NM_008942.2	NP_032968.2	GI: 68226730
NPEPPS_Rn	10q31	NM_080395.1	NP_536320.1	GI: 158749539

Table 1 continued

Gene Name	Location	Nucleotide #	Amino Acid #	GI #
RNPEP_Hs	1q32	NM_020216.3	NP_064601.3	GI: 40316914
RNPEP_Mm	1	NM_145417.3	NP_663392.2	GI: 227499102
RNPEP_Rn	13q13	NM_031097.1	NP_112359.1	GI: 13592046
RNPEPL1_Hs	2q37.3	NM_018226.4	NP_060696.4	GI: 318065088
RNPEPL1_Mm	1	NM_181405.4	NP_852070.3	GI: 318065084
RNPEPL1_Rn	9	NM_001271177.1	NP_001258106.1	GI: 403224984
TRHDE_Hs	12q15-q21	NM_013381.2	NP_037513.1	GI: 218751879
TRHDE_Mm	10	NM_146241.2	NP_666353.1	GI: 256542286
TRHDE_Rn	7q22	NM_001108991.1	NP_001102461.1	GI: 157823372

Chromosome and Gene Conservation

Gene order/chromosome conservation for the M1 metalloproteases was first investigated by determining the chromosome locations of these genes in *H. sapiens*, *M. musculus* and *R. norvegicus*. The GENE tool was used to identify genes that flank M1 metalloproteases in human, mouse, and rat. Additionally, this data was used to determine the relative positions and direction of transcription of these flanking genes [16]. Gene order was visually inspected and resulting summary figure drawn. Further, the NCBI synteny tool was utilized with respect to *LNPEP*, *ERAP1*, and *ERAP2* to gather data in support of a chromosomal rearrangement in mouse and rat.

Gene and Exon Structure Analysis

Exon sequences were obtained from UCSC Genome Browser [17] (Table 1). For genes with more than one isoform, the most extensive reference sequence was chosen. Sequences were obtained with 5' and 3' UTRs and divided by exons, with introns excluded. The location of functional motifs was obtained from the MEROPS database [18]. Reference sequences were kept consistent throughout the various stages of data collection. Exon size, exon reading frame, and alignment of conserved amino acids was determined using PERL scripts and manual inspection. The exon that encoded sequences of the active site amino acids was selected as an anchor exon. For seven of twelve human genes and six of eleven mouse genes, the anchor exon was 155 bp exon (*LNPEP*, *ENPEP*, *ANPEP*, *APQ*, *ERAP1*, *ERAP2*, and *TRHDE*). Using this exon as an immobile base, upstream and downstream exon lengths and reading frames were manually compared among the genes. For two genes in human and mouse, the anchor exon was a 236 bp exon (*RNPEP*, *RNPEPL1*). Gene structure diagrams were constructed based on conservation of reading frame. PERL scripts are available on request.

RESULTS

Phylogenetic Analyses

To gain a clearer picture of the evolutionary history of the M1 gene family,

several lines of evolutionary evidence were gathered including multiple

alignment, phylogenetic analyses, gene order/synteny among three

species, and gene structure conservation. Amino acid alignments were

completed for *H. sapiens*, *M. musculus*, and *R. norvegicus* M1 family

aminopeptidases to determine the level of identity and similarity among

the proteins as well as critical motifs necessary for placement into the M1

family (Figure 1). As expected from how the family is defined, the active

site and zinc binding motifs were conserved within nearly all sequences

aligned. The G/AMEN motif is a critical part of the active site and is highly

conserved. Aminopeptidase O (APO), while still defined as a member of

the family by function does differ by having a GMAS motif in place of the

G/AMEN motif. This difference was found consistently in all APO genes

that were examined for this study (data not shown). All aminopeptidases

also showed conservation of a downstream tyrosine (Y) residue that is

believed to be involved with the active site [18].   The zinc binding motif

(HEXXHX$_{18}$E) was conserved in all proteins in all species examined, including APO.

Figure 1: Aligned amino acids for M1 aminopeptidases in A. human, B. rat, and C. mouse. The boxed areas highlight the amino acids involved in the active site and zinc binding motifs, which characterize this family. Amino

acid sequences were aligned with CLUSTAL W and shaded with

Boxshade. Red indicates identity and blue indicates similarity.

To gain a better view of the relatedness and relationships among the

individual members of the M1 aminopeptidase group, phylogenetic trees

were then constructed individually using human, rat, and mouse amino

acid sequences (Figure 2, A-C). Additionally, a more comprehensive tree

was constructed which included all M1 amino acids sequences found in

mammals, fish, reptiles, and birds (Figure 2, D). Within all the trees

completed, three groups were consistently observed. The largest group

contained seven human aminopeptidases and six rat/mouse

aminopeptidases: APQ, TRHDE, ANPEP, ENPEP, NPEPPS, LNPEP, ERAP1, and

ERAP2 when present. The second group contained LTA$_4$H, RNPEP, and

RNPEPL1. The third group consistently contained a single aminopeptidase,

APO, which appeared to be independent. This is consistent with the

degree of amino acid dissimilarity from the other aminopeptidases. The

large group could be further subdivided into two more closely related

groups, the first containing ENPEP, NPEPPS, TRHDE, ANPEP, and APQ and

the second subgroup containing ERAP1, ERAP2, LNPEP.
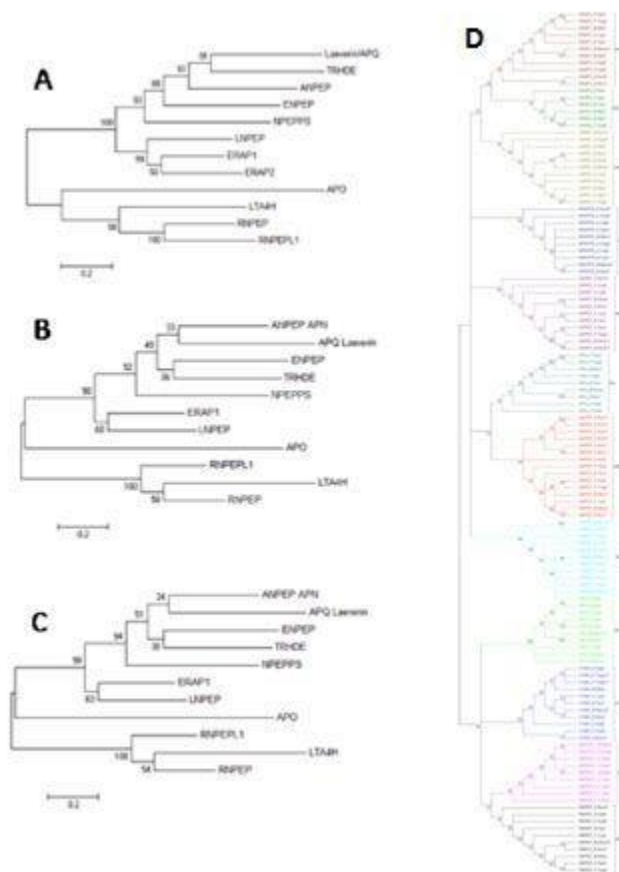
Figure 2: Phylogenetic analyses of A. human, B. rat, C. mouse amino acid sequences for M1 aminopeptidases. Analyses showing these species and sequences found for all mammals fish, birds, and reptiles are shown in D. Neighbor joining trees were constructed within MEGA6 with bootstrap resampling at 500. Bootstrap values are shown as percentages. Homologous proteins are color-matched in D to highlight similarity to the pattern observed in the individual species trees.

Gene Order/Chromosomal Context Analyses

Since gene families are known to result from gene and/or genome duplication, it was of interest to examine the chromosomal context and conservation of each M1 aminopeptidase gene relative to the others in human, mouse, and rat. M1 aminopeptidase family genes and also flanking genes within each genome, chromosome locations were retrieved from NCBI for all human, rat, and mouse family members. Analyses of these maps showed several features. First, for all individual genes studied, most were not found on the same chromosome, indicating that if tandem gene duplication occurred, then it occurred in the distant past. Second, at least one conserved gene neighbor was consistently present for all three species (Figure 3) indicating the gene duplication event included flanking genes as well as the aminopeptidase genes or they results from a whole genome duplication event. Four genes (*ANPEP*, *AQPEP*, *APO*, *RNPEPL1*) showed identical gene order and direction of transcription in human, mouse, and rat (Fig. 3 A, B, H, I). Seven of the eleven genes showed a conserved chromosomal context between human and rat/mouse, but were directly inverted with respect to both gene order and direction of transcription (Figure 3C-G, J, K; *ENPEP*, *NPEPPS*, *LNPEP*, *TRHDE*, *ERAP1*, *LTA$_4$H*, RNPEP). Lastly, in humans, *ERAP1*,

ERAP2, and *LNPEP* were found together on chromosome 5. However, *ERAP2* is absent in rat and mouse and *LNPEP* (Figure 4C) is located on another chromosome. Analysis of syntenic regions of the human, mouse, and rat genome suggests that *ERAP2* may have been destroyed in a chromosomal rearrangement that occurred between the *ERAP1* and *CAST* genes (Figure 4A). This may explain why *ERAP2* is absent in mouse and rat and present in other species such as monkeys, dogs, cows, and zebrafish (Figure 4B).
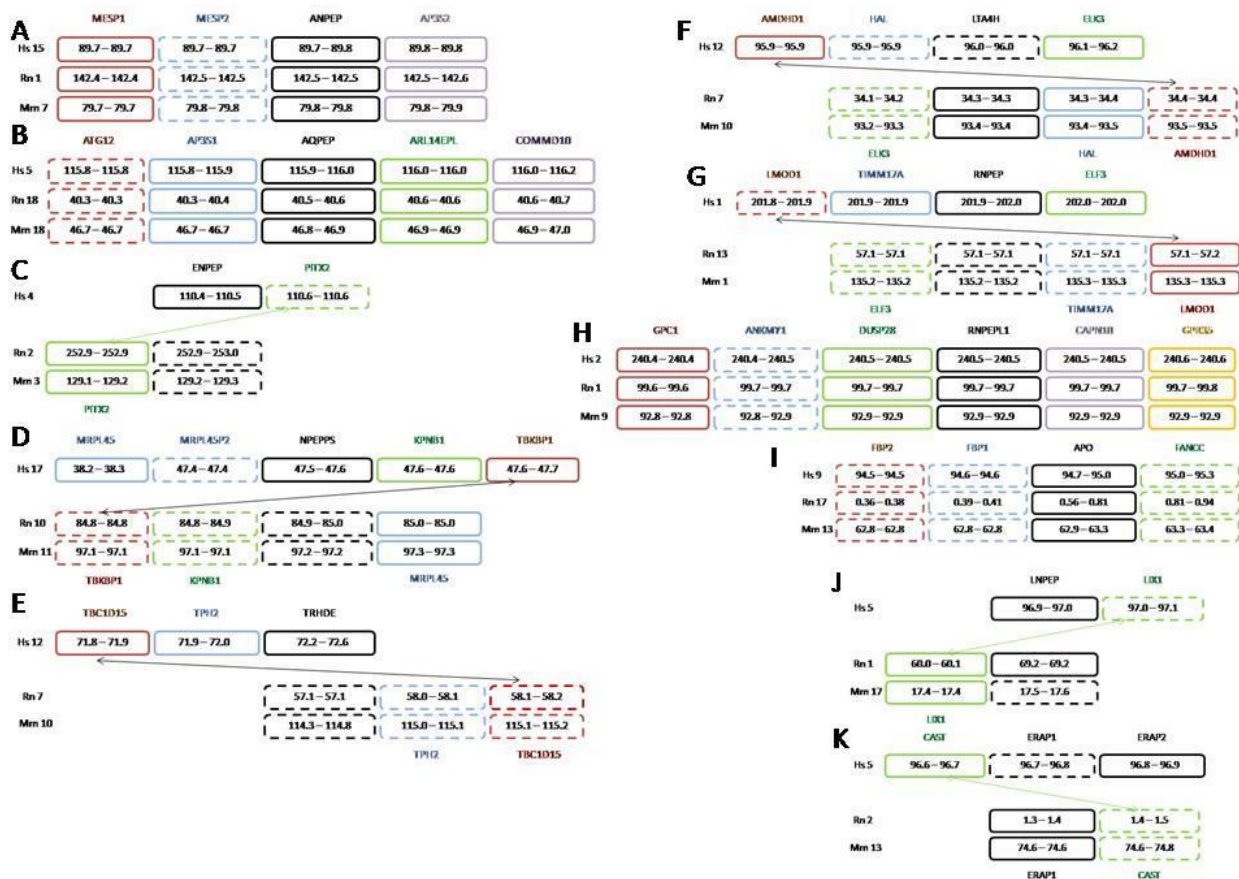
Figure 3: Comparison of the chromosomal context of each M1 aminopeptidase in human, mouse, and rat. Aminopeptidase genes are noted in black boxes, flanking genes are colored, dashed boxes denote genes in reversed orientation and the numbers in the boxes denote chromosomal location.

To facilitate a clearer understanding of the presence/absence of *ERAP2* in a variety of species, the M1 aminopeptidases present in species used in this study were compared with an evolutionary species tree [19, 20]. Genomes from human, primates, dog, cow, a fish, and a nematode contained *ERAP2* (Fig. 4D, solid lines), while other genomes did not (rodents, chicken, frog). The presence of *ERAP2* in a variety of divergent species suggests it is not a recent addition to these genomes, but that *ERAP2* absence is due to gene death in a distinct lineage.
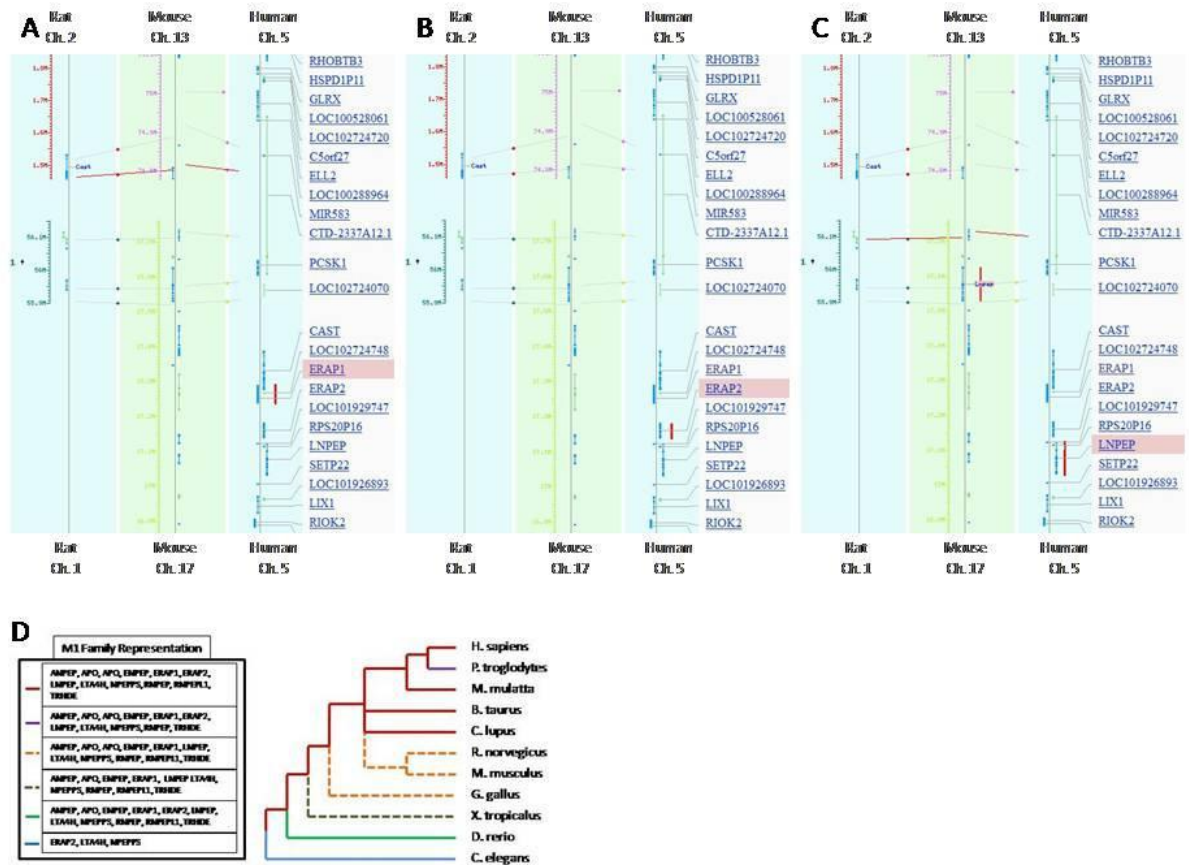
Figure 4: Syntenic regions of human, mouse and rat chromosomes containing *ERAP1* (A), *ERAP2* (B), and *LNPEP* (C). Chromosome numbers and species are noted at the top and bottom of each chromosomal representation as visualized in the synteny tool at NCBI. D: Tree of species highlighting the presence or lack of *ERAP2*. Solid lines denote species with *ERAP2* and dashed lines represent species that are missing *ERAP2*. The line coloring highlights which members of the family are present, as noted in the key.

Gene Structure Analysis

Due to the origin of gene family members through gene and genome duplication followed by divergence, members of a gene family should also show features of relatedness at the gene structure level. The more similar the gene structure, the more closely related the genes are to each other. A comprehensive analysis of the gene structure of the human M1 aminopeptidase family was completed to study the relatedness of the M1 aminopeptidase family. For the human genes, exon number ranged from 11 exons (*RNPEP* and *RNPEPL1*) to 23 exons (*NPEPPS*) with a total mRNA length ranging from 2060 nucleotides (nt) (*LTA₄H*) to 5682 nt (*ERAP2*). M1 family members are found on human chromosomes 1, 2, 4, 5, 9, 12, 15, and 17. Mouse exon number ranged from 11 exons (*RNPEP* and *RNPEPL1*) to 23 exons (*NPEPPS*) with the smallest total mRNA at 2,066 nt (*LTA₄H*) and largest total mRNA at 5,722 nt (*TRHDE*). Mouse M1 genes were found on chromosomes 1, 3, 7, 10, 11, 13, 17, and 18.

To examine relatedness among human genes with differing exon numbers, the exon containing the functional motif HEXXH was used as an anchor exon and exons upstream and downstream from this exon were

compared for reading frame conservation. Exon lengths that maintained downstream reading frames were observed through the entire 3' end and most of the 5' end of two groups of genes. The largest group consisted of seven genes (*ANPEP*, *APQ*, *TRHDE*, *ENPEP*, *LNPEP*, *ERAP1*, and *ERAP2*) (Figure 5A). A reference number for each exon is noted at the bottom of the diagram (1-22) to reduce confusion. Some important features were observed which allow these seven genes to be further subgrouped. First, three instances of exon fusion and/or exon splitting (*ERAP1*, *ERAP2*, *LNPEP*) from the main group were observed (reference exons 9/10, 14/15, and 17/18). These fused exons maintain the same reading frame and were found aligned in a multiple alignment with the individual exons 9 and 10, 14 and 15, and 17 and 18 respectively. Of these three genes, two genes (*ERAP1* and ERAP2) showed insertion of an in frame exon (reference exon 11) that was not found in any other gene.   The exons in the 5' ends of all seven genes showed more variation, but still maintained the exon reading frame conservation in many cases. This analysis showed that *ANPEP*, *APQ*, *TRHDE*, *ENPEP*, *LNPEP*, *ERAP1*, and *ERAP2* are structurally related with closer relationships among *ANPEP*, *APQ*, *TRHDE*, and *ENPEP* as a group and *LNPEP*, *ERAP1*, and *ERAP2* as a group.

The smaller group of genes with conserved exon structures consisted of *RNPEP* and *RNPEPL1* (Figure 5B). These genes were nearly identical in reading frame, but also exon size. The remainder of genes, *APO*, *LTA$_4$H*, and *NPEPPS* appeared to be structurally dissimilar to these prior groups, despite amino acid similarity to particular groups of genes, and to each other suggesting divergence from each other and/or the prior groups. Nearly identical results were observed with mouse exon structure analyses except that mouse lacks *ERAP2*, so only one gene (*ERAP1*) showed insertion of an in frame exon (reference exon 11) not found in any other aminopeptidase in that group (Figures 6).

Figure 5: Conserved exon structure/reading frame analyses in human M1 aminopeptidases. A. Group 1 and B. Group 2. Each circle represents an exon and the number(s) inside note the exon size. Circles with solid lines represent exons that are exactly the same size in each gene. Circles with dashed lines represent exons that maintain the reading frame, but are not exactly the same size. Each line of a distinct color and solidity represents a different gene's structure (see key for line color/solidity). A shaded circle represents an entirely non-coding exon.
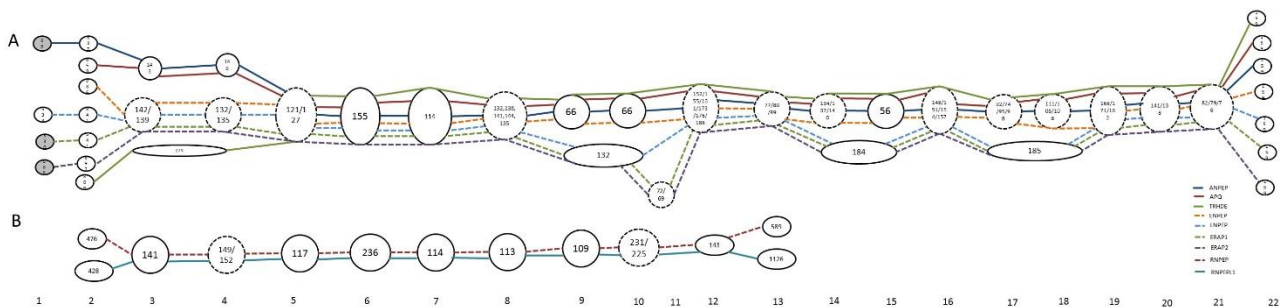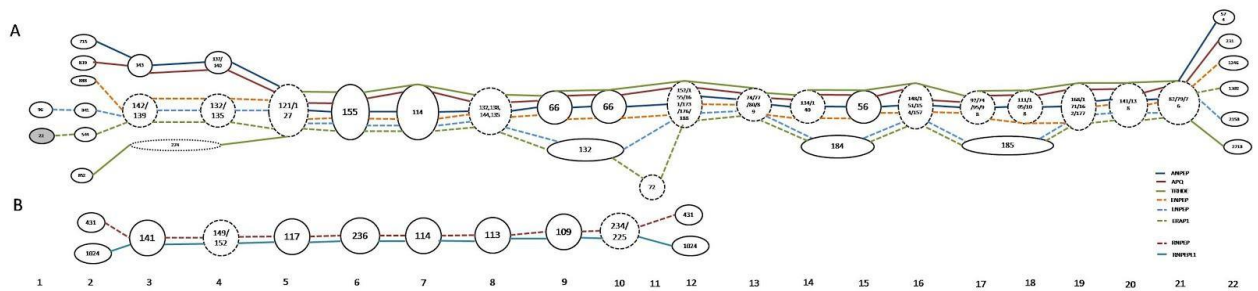
Figure 6. Conserved exon structure/reading frame analyses in mouse M1 aminopeptidases. A. Group 1 and B. Group 2. Each circle represents an exon and the number(s) inside note the exon size. Circles with solid lines represent exons that are exactly the same size in each gene. Circles with dashed lines represent exons that maintain the reading frame, but are not exactly the same size. Each line of a distinct color and solidity represents a different gene's structure (see key for line color/solidity). A shaded circle represents an entirely non-coding exon.

DISCUSSION

This study aimed to critically examine the evolutionary history and relatedness of the M1 aminopeptidase family, a diverse group of conserved metalloproteases, using multiple lines of evidence, including multiple alignment, phylogenetic analyses, gene order/synteny, and gene

structure. The family shares critical traits such as a HEXXHX$_{18}$E zinc binding

domain that is required for enzymatic activity, a G/AMEN domain that is

believed to be the active site pocket, and a 3' tyrosine residue that

completes the enzymatic site. Multiple alignment analysis revealed

conservation of the critical known domains of interest across three chosen

species; human, rat, and mouse, insure that all proteins being studied

were in fact true members of the M1 aminopeptidase family.

The most different and divergent member of the M1 family appears to be

aminopeptidase O. Aminopeptidase O contained a GMAS domain

instead of a G/AMEN domain in all species studied. This change to the

most critical domain for an aminopeptidase, along with its overall amino

acid dissimilarity to other M1 family members, suggests that

aminopeptidase O is an early member of the family and is long diverged.

This was supported by phylogenetic analysis, which showed that APO was

separate from all other members of the family and not within a group for

all species studied. Gene structure analysis showed that aminopeptidase

O was not similar to any other aminopeptidase in exon structure for any

part of the gene (data not shown).

Both phylogenetic and gene structure analyses suggest moderately

recent additions to the M1 gene family are represented by the large

group including *LNPEP, ERAP1, ERAP2, APQ, ENPEP, ANPEP, TRHDE*, and *NPEPPS* with *NPEPPS* being the most diverged in amino acid sequence and exon structure (data not shown). This large group could be further split into two smaller groups with the more diverged group being *LNPEP, ERAP1*, and *ERAP2* (when present), due to the exon structure differences observed in this group (exon joining and exon insertion). In addition, these three genes encode proteins that are known to share some functional characteristics as oxytocinases [5]. The remaining five genes in this large group are most similar to each other in amino acid sequence and exon structure, suggesting either that they are more recent additions or have critical non-redundant roles in cellular processes.

Phylogenetic analyses indicate the last group should consist of *LTA$_4$H, RNPEP*, and *RNPEPL1*. The nearly identical exon structure and high amino acid similarity of *RNPEP* and *RNPEPL1* suggests these two genes are most recent gene duplication. Biochemical studies indicate that *RNPEPL1* is more similar in substrate preference to LTA$_4$H [21] than any other M1 aminopeptidase, supporting the grouping of these three aminopeptidases.

Lastly, a study of the synteny/gene order for these genes in human, mouse, and rat has provided some further evolutionary history for the *ERAP1*, *ERAP2*, and *LNPEP* group. In humans, *ERAP2* resides near *ERAP1* and *LNPEP* on the same chromosome (5q15), and are therefore likely a result of gene duplication. Despite being present in humans, monkeys, dogs, cows, zebrafish, and even *C. elegans*, *ERAP2* is absent in rat and mouse. Prior studies suggested that *ERAP2* was a very recent gene duplication of *ERAP1* in humans, however, that explanation is unable to account for the seemingly "random" observation within certain species. Furthermore, the conservation of *ERAP2* within a diverse set of species is far too high for this duplication to have happened in each species independently without a common precursor gene. The gene order/synteny suggests that a chromosomal rearrangement between the *ERAP1* and *CAST* genes occurred somewhere in the mouse/rat lineage that effectively bisected *ERAP2* sequences and destroyed the gene. Taken together, these lines of evidence indicate that the M1 family of aminopeptidases is a complex and old gene family with three major groups that can be further subdivided. Some members are more recently duplicated as shown by high levels exon structure conservation, while other members are long diverged with minimal amino acid conservation.

Furthermore, one family member may have been eliminated by a chromosomal rearrangement in distinct species lineages.

CONCLUSIONS/HIGHLIGHTS

Phylogenetic analysis of M1 aminopeptidase amino acids shows three subgroups that are consistent with subgroups assigned based on biological process.

Exon structure is conserved in nine human and eight mouse M1 aminopeptidases.

Aminopeptidase genes grouped by exon structure conservation mirror aminopeptidase functional grouping and phylogenetic analysis.

The absence of *ERAP2* in mouse and rat is likely due to gene death due to a chromosomal rearrangement in the rat/mouse lineage.

ACKNOWLEDGEMENTS

REFERENCES

[1] Hahn, M.W.; Bie, T.D.; Stajich, J.E., Nguyen, C.; Cristianini, N. Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genomic Research*, **2005**, *15*, 1153-1160.

[2] Macqueen, D.J.;Wilcox,A.H. Characterization of the definitive classical calpain family of vertebrates using phylogenetic, evolutionary and expression analyses. *Open Biology*, **2014**, *4,130219*; doi:10.1098/rsob.130219.

[3] Daza, D.O.; Sundstrom, G.; Bergqvist, C.A.; Duan, C,; Larhammar, D. Evolution of the insulin-like growth factor binding protein (IGFBP) family. *Endocrinology,* **2011,** *152(6),* 2278 - 2289.

[4] Koonin, E.V. Evolution of genome architecture. *International Journal of Biochemistry and Cell Biology,* **2009,** *41(2),* 298 - 306.

[5] Tsujimoto, M.; Hattori, A. The oxytocinase subfamily of M1 aminopeptidases. *Biochimica et Biophysica Acta - Proteins and Proteomics,* **2005,** *1751(1),* 9-18.

[6] Albiston, A.L.; Fernando, R.; Ye, S.; Peck, G.R.; Chai, S.Y. Alzheimer's, angiotensin IV and an aminopeptidase. *Biological & pharmaceutical bulletin.* **2004,** *27(6),* 765 - 767

[7] Tanioka, T.; Hattori, A.; Masuda, S.; Nomura, Y.; Nakayama, H.; Mizutani, S.; Tsujimoto, M. Human Leukocyte-derived Arginine Aminopeptidase. The third member of the oxytocinase subfamily of aminopeptidases. *Journal of Biological Chemistry,* **2003,** *278(34),* 32275-32283.

[8] Albiston, A.L.; Ye, S.; Chai, S.Y. Membrane Bound Members of the M1 Family: More Than Aminopeptidases. *Protein & Peptide Letters,* **2004,** *11(5),* 491-500.

[9] Diaz-Perales, A.; Quesada, V.; Sanchez, L.M.; Ugalde, A.P.; Suarez, M.F.; Fueyo, A.; Lopez-Otin, C. Identification of Human Aminopeptidase O, a Novel Metalloprotease with Structural Similarity to Aminopeptidase B and Leukotriene A4 Hydrolase. *Journal of Biological Chemistry,* **2005,** *280(14),* 14310 - 14317.

[10] Maruyama, M.; Hattori, A.; Goto, Y.; Ueda, M.; Maeda, M.; Fujiwara, H.; Tsujimoto, M. Laeverin / Aminopeptidase Q, a Novel Bestatin-sensitive Leucine Aminopeptidase Belonging to the M1 Family of Aminopeptidases. *Journal of Biological Chemistry,* **2007,** *282(2),* 20088 - 20096.

[11] Constam, B.D.; Tobler, A.R.; Rensing-Ehl, A.; Kemler, I.; Hersh, L.B.; Fontana, A. Puromycin-sensitive aminopeptidase. Sequence analysis, expression, and functional characterization. *Journal of Biological Chemistry,* **1995,** *270,* 26931 - 26939.

[12] Helgstrand, C.; Hasan, M.; Uysal, H.; Haeggstrom, J.Z.; Thunnissen, M. A Leukotriene A4 Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding. *Journal of Molecular Biology,* **2011,** *406(1),* 120 - 134.

[13] Georgiadou, D.; Heam, A.; Evnouchidou, I.; Chroni, A.; Leondiadis, L.; York, I.A.; Rock, K.L.; Stratikos, E. Placental leucine aminopeptidase efficiently generates mature antigenic peptides in vitro but in patterns distinct from endoplasmic reticulum aminopeptidase 1. *Journal of Immunology,* **2010,** *185(3),* 1584 - 1592.

[14] Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. ClustalW and ClustalX version 2. *Bioinformatics,* **2007,** *23(21)*, 2947-2948.

[15] Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution,* **2013,** *30*, 2725-2729.

[16] Sayers, E.W.; Barrett, T.; Benson, D.A.; Bryant, S.H.; Canese, K.; Chetvernin, V.; Church, D.M.; DiCuccio, M.; Edgar, R.; Federhen, S.; Feolo, M.; Geer, L.Y.; Helmberg, W.; Kapustin, Y.; Landsman, D.; Lipman, D.J.; Madden, T.L.; Maglott, D.R.; Miller, V.; Mizrachi, I.; Ostell, J.; Pruitt, K.D.; Schuler, G.D.; Sequeira, E.; Sherry, S.T.; Shumway, M.; Sirotkin, K.; Souvorov, A.; Starchenko, G.; Tatusova, T.A.; Wagner, L.; Yaschenko, E.; Ye, J. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, **2009,** *37(9),* 3124.

[17]Kent, W.J.; Sugnet, C.W.; Furey, T.S.; Roskin, K.M.; Pringle, T.H.; Zahler, A.M.; Haussler, D. The human genome browser at UCSC. *Genomic Research,* **2002,** *12(6),* 996 - 1006.

[18] Rawlings, N.D.; Waller, M.; Barrett, A.J.; Bateman, A. *MEROPS*: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research,* **2014,** *42*, D503-D509.

[19] Letunic, I.; Bork, P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acid Research,* **2011,** *39,* W475-W478.

[20] Maddison, D.R.; Maddison, W.P. The Tree of Life Web Project. Internet address: http://tolweb.org (Accessed July 10, 2014),

[21] Thompson, M.W.; Beasley, K.A.; Schmidt, M.D.; Seipelt, R.L. Arginyl Aminopeptidase-Like 1 (RNPEPL1) Is an Alternatively Processed Aminopeptidase with Specificity for Methionine, Glutamine, and Citrulline Residues. *Protein & Peptide Letters,* **2009,** *16,* 1256 - 1266.

CHAPTER FOUR: RNPEPL1 IN INFLAMMATION

INTRODUCTION

Inflammation is a very well characterized physiological response to infection or trauma. The key signs of inflammation include redness, heat, swelling, and pain to the injury or infection sight. Inflammatory response can be mediated through several different pathways depending on the triggering agent. Bacterial lipopolysaccharide (LPS) exposure is well known to induce inflammatory response [1]. Numerous animal models report the use of LPS both locally and systemically to induce inflammation for study of the physiological responses that follow. Namely, intraperitoneal injection with LPS is known to induce systemic inflammatory response through the NFKB pathway [2]. These pathways are very complex and affect numerous cellular components.

The M1 aminopeptidase family members are known to regulate proteins and peptides by exopeptidase activities. They cleave the N-terminal residues from small proteins and peptides to activate or inactivate the targets. Many members that have been characterized to date are reported to play various roles in inflammatory response. LNPEP

[3,4], ERAP1 [3,5], ERAP2 [5], APQ [6], ENPEP [7], and APO [8] all have activity in the angiotensin pathway, which is intertwined with inflammation [9]. APQ was reported as one of the three most-highly overexpressed genes in rheumatoid arthritis [10], an inflammatory disease. LTA_4H, the most similar member to RNPEPL1 (as previously shown) is a major proinflammatory enzyme in the leukotriene pathway [11]. Furthermore, similarity to LTA_4H and RNPEP based on computational predictions suggests that it may play a role in inflammation.

This study intended to further current, limited knowledge of one specific M1 aminopeptidase whose function is yet unknown in an inflammatory model. Arginyl aminopeptidase-like 1 (RNPEPL1) was initially characterized in 2008 with the cloning of a transcript, expression and purification of the protein, as well as characterization of splice variants and aminopeptidase activities. Since that characterization, no further information has been published on RNPEPL1 function. Due to the lack of information on RNPEPL1's biological role, this initial study aimed to determine whether RNPEPL1 expression changed in a model of the inflammatory response, and if so, to what extent at both the cellular and systemic levels. A significant change in RNPEPL1 RNA or protein expression level would strengthen support for RNPEPL1's possible role in inflammation

and provide direction for more thorough and mechanistic studies. However, the changes in RNPEPL1 expression at the RNA and protein level in all experiments were consistently small, but statistically significant. This suggests that RNPEPL1's role may not be dramatic in inflammation, its activity may be regulated at the protein level, it may not have significant a role in this inflammatory model, or it may function in a different biological process altogether. This limited success suggests that other avenues should be explored such as testing for LTA_4H hydrolase activity.

MATERIALS AND METHODS

Cell Culture

Bone marrow derived macrophages from FVBN/J mice were isolated and cultured per standard technique by the James West Lab (Vanderbilt University) [12]. Growth conditions were as follows: Standard DMEM growth medium with 10% FBS added, temperature was held constant at 37°C, and carbon dioxide levels were monitored and held at 5%. Inflammatory response was induced with 1ug/ml lipopolysaccharide (LPS) (E. coli B, Sigma) added to standard media for the desired

incubation time. A non-inflammatory negative control was achieved by retaining one flask of cells that had received only standard media without any LPS addition for each replicate. Cells were then harvested for RNA and/or total protein isolation as described below.

RNA Analysis

Microarray data was obtained through collaboration with the James West Lab (Vanderbilt University) [12]. Briefly, inflammatory response was induced as noted above with 4 and 24 hour time points. A non-inflammatory negative control was prepared as described above. RNA was isolated and processed for microarray analysis via Qiagen RNEasy Kit. RNA samples were then prepared for CORE processing per instructions. Microarray results were analyzed using one –way ANOVA.

Protein Analysis

Total protein was isolated by adding TRIzol Reagent (Ambion, Ref # 15596026) directly to the cell culture and processed as directed by the manufacturer instructions. The resulting protein solution was then analyzed by Mirco BCA Protein Assay Kit (Pierce, Product # 23235) to allow for standardization of total protein concentration. Western blots

were achieved by first separating samples by size by gel electrophoresis on a SDS-PAGE gel (Genscript Express PAGE Gels (4-20%) Cat. #MG420W12) with a prestained molecular weight marker. Next, the separated protein was transferred to a PVDF membrane (Immobilon, Cat # IPFL00010)) via a semi-dry transfer machine with cutoff values of 30V, 90mA, and 300W for 4 hours. The membranes were then processed for visualization. This began with the membranes being washed once in TBS - T (1x TBS with .2% Tween 20) to remove any gel residue and blocked in 10-15mls of 5% dry milk in TBS for 1 - 2 hours at room temperature or overnight at 4C to prevent any non-specific binding. Following removal of the blocking buffer, primary antibody incubation was completed by adding anti-RNPEPL1 primary antibody (GeneTex, 1:2500) and GAPDH primary antibody, a control standard, (GeneTex, 1:2500) for 1 - 2 hours at room temperature or overnight at 4C. All antibodies used were diluted in blocking buffer described above. Following the incubation, the primary antibodies were decanted and the membrane was washed 3 times for 5 - 10 minutes each in 10 - 15 mls of TBS – T to remove any antibody residue. Alkaline phosphatase conjugated secondary antibodies (Goat pAb to Rabbit IgG AP – abcam, ab6722-1) for RNPEPL1 @ 1:2500, (AKP anti-mouse IgG, BD Biosciences Pharmingen, Cat # 553394) for GAPDH @

1:2500 were utilized to allow for color development and visualization. Secondary incubation was processed as described above. Once secondary incubation was completed, the secondary antibodies were decanted and the membrane was washed 2 times for 5 – 10 minutes in TBS – T and 2 times for 5 – 10 minutes each in 10mls alkaline phosphatase substrate buffer. Finally, 33ul of NBT and 66ul of BCIP was added to 15mls alkaline phosphatase substrate buffer and added to the membrane immediately following the second alkaline phosphatase substrate buffer wash. The solution was allowed to incubate on the membrane until color development was achieved to desired darkness. The membrane was then rinsed with diH2O to stop the development process. All blotting washes and incubations were completed at room temperature, unless otherwise noted, and shaking at 30 – 50 rotations per minute.

Blood Samples and Tissues for IHC

The mouse treatment and sacrifice described below were not done for the purpose of this current study. Unused tissues were generously donated post-sacrifice from a previous, unrelated study in the West Lab.

Blood plasma samples and tissue slides for IHC from mice with and without inflammatory response due to LPS insult were received from the James West Lab. Briefly, FVBN/J mice were randomly divided into two groups, control and LPS. The control animal group received standard care with free access to food and water. The LPS group received standard care with the addition of one intraperitoneal injection of 1mg/kg of LPS. The mouse groups were then sacrificed 24 hours post LPS injection.

Mouse Tissues IHC

Mouse tissues were harvested from sacrificed animals, fixed in 10% formalin, and paraffin embedded. Tissue blocks were then sectioned in 5 micron slices and baked onto glass slides. Slides were then prepared for staining by deparaffinizing the tissue in two 5 minute washes in xylene. Sections were then rehydrated in a series of alcohol washes. The slides were then processed through antigen retrieval by steaming for 20 minutes in citrate buffer (pH 6.0). Slides were then allowed to cool and rinsed 2 times in diH$_2$O. Next, the slides were incubated in 3% H$_2$O$_2$ in methanol to eliminate any non-specific peroxidase activity. Slides were withdrawn from the peroxide solution and rinsed three times with diH2O and twice with PBS. Slides were then blocked in 10% Universal Blocking

Reagent (Biogenex # HK085-5K) in PBS and incubated in a humid

chamber for 15 minutes. Primary antibody (anti-RNPEPL1, GENETEX) was

added at 1:200 dilution in PBS and left overnight at 4C. Slides were then

rinsed three times for 10 minutes each in 0.4% Triton X-100 in PBS to remove

any unbound antibody. The secondary antibody was applied and the

slides were incubated for 30 minutes at room temperature. After the

secondary incubation, the slides were rinsed in 0.4% Triton X-100 in PBS

three times for 10 minutes each. Next, the Biogenex Stretavidin HRP

(Innogenex Streptavidin CJ-1005-25) conjugate was applied and left for

20 minutes at room temperature. The slides were rinsed in 0.4% Triton X-100

in PBS three times for 10 minutes each. To visualize the protein staining,

Vector Red was applied and the first slide was immediately placed under

the microscope. A timer was started. The Vector Red was allowed to

develop to desired color intensity. The time was noted at three

minutes. All remaining slides were processed with the Vector Red for the

same three minutes to ensure that the level of color development was the

same for each slide. Finally, the slides were counterstained with Methyl

Green, which stains the nucleus to aid in identification of the cellular

components and allow for a positive staining control for the negative

antibody control slide. Lastly, the slides were cover slipped with mounting

media. The slides were then analyzed and three pictures, of similar physiologically nature and identical microscope and camera settings, were taken of each slide, for both the test section and the negative control section that was not processed with primary anti-RNPEPL1 antibody. The pictures were then uploaded and processed for red color intensity through Image J, a standard program for IHC [13]. The results were analyzed by averaging the percentage of red staining for all slides in each group.

Blood Samples

Blood samples from sacrificed mice were allowed to clot at room temperature for 2-4 hours. The samples were then centrifuged at 500 rpms for 5 minutes to separate the cells from the plasma. The plasma was then snap frozen in liquid nitrogen and stored at -80C for further processing. Upon thawing, TRIzol (Ambion) was added and the samples were processed as directed by the manufacturer. Western blot analysis was performed consistently as described above.

RESULTS

RNA Expression

RNA extracted from bone-marrow derived macrophages was processed and run on microarray chips for three different conditions. Condition one was control macrophages that were cultured under standard protocols. Condition two contained the same macrophages under the same culture conditions with the addition of 1ug/ml of LPS added to the media 4 hours before harvest. This group will be referred to as the LPS 4 hour group. The third treatment condition consisted of the same cells cultured under identical conditions with LPS treatment occurring 24 hours before harvest. This group will be referred to as LPS 24 hour group. The addition of LPS is to induce inflammatory response as described above. The purpose of testing time points at 4 and 24 hours is to distinguish between expression levels of developing inflammatory response (early in the insult) and fully developed inflammatory response (once inflammatory state has been achieved). Figure 1 shows the results from the microarray comparison for the three different conditions. RNPEPL1 expression is compared to LTA_4H expression for a reference of a well-characterized and closely related member of the M1

family. RNPEPL1 RNA expression change is not found to be significant at 4 hours, reporting at a decrease of 2.2%. RNPEPL1 RNA expression change is found to decrease by 25% at 24 hours, which is significant by one-way ANOVA with a p-value 0.0017.
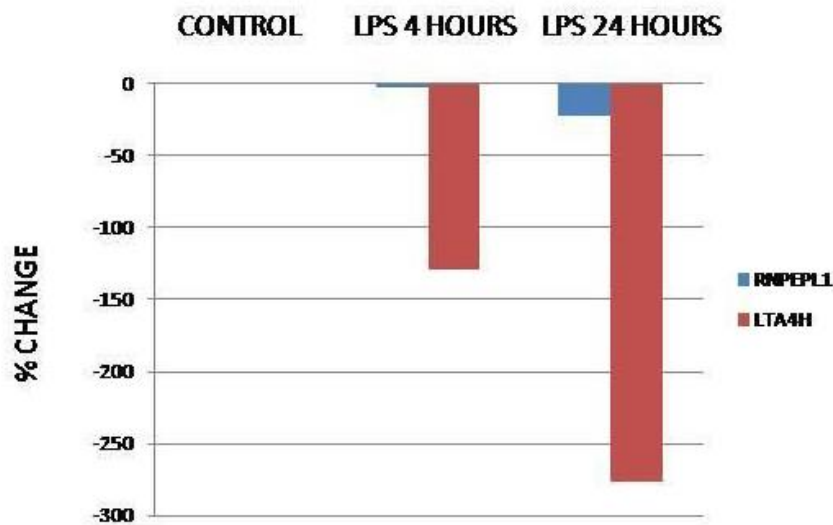

Figure 1: Percent change of RNPEPL1 RNA expression levels, quantified by microarray, for three conditions: control macrophages, macrophages with 4 hours of LPS treatment, and macrophages with 24 hours of LPS treatment. Note that the percent change seen for these two specific genes are negative, showing a decrease in RNA expression.

Protein Analysis – Macrophages

To determine if protein expression decreased similar to RNA expression,

Western blots on total protein were run. Bone-marrow derived

macrophages from FVBN/J mice were isolated and cultured as described

above. The cells were then divided into 7 groups. The control group

underwent standard treatment. The remaining six groups underwent LPS

treatment at a dosage of 1ug/ml with the following time points: 1 hour, 2

hours, 4 hours, 8 hours, 12 hours, and 24 hours. The various time points

allowed for tracking of RNPEPL1 protein throughout the inflammatory

response. Each group was run in triplicate with protein samples for each

time point being combined. Figure 2A represents a Western blot run with

the protein harvested for each time point. Lane 1 shows the molecular

weight marker used for a size standard. Lane 2 represents the control cell

group cultured without any LPS treatment. Lanes 3-6 show the LPS

treatment groups with time of treatment noted below each lane. Lane 7

represents a RNPEPL1 control from purified protein from the human

transcript. The larger band seen at the 67 kDa area is RNPEPL1

expression. The smaller band seen at 43 kDa is the GAPDH loading

standard. Note the size of the RNPEPL1 band when compared to the

human standard. This is due to the mouse gene containing a larger exon

one than the human transcript used. Figure 2B represent the quantitation

of the intensity of each band as measured by BIO-RAD imaging

software. Each band was measured for intensity with a blank from directly

above or below each sample. The intensities were then standardized for

total loaded protein based on the known loading control GAPDH

standard. Finally, the change in intensity was determined and graphed,

represented in Figure 2B. There is a marked increase in protein

concentration during the first four hours of inflammatory response,

followed by a drop off at 8 and 12 hours. Finally, the 24 hour time point

shows that RNPEPL1 expression was 25% less than control or non-
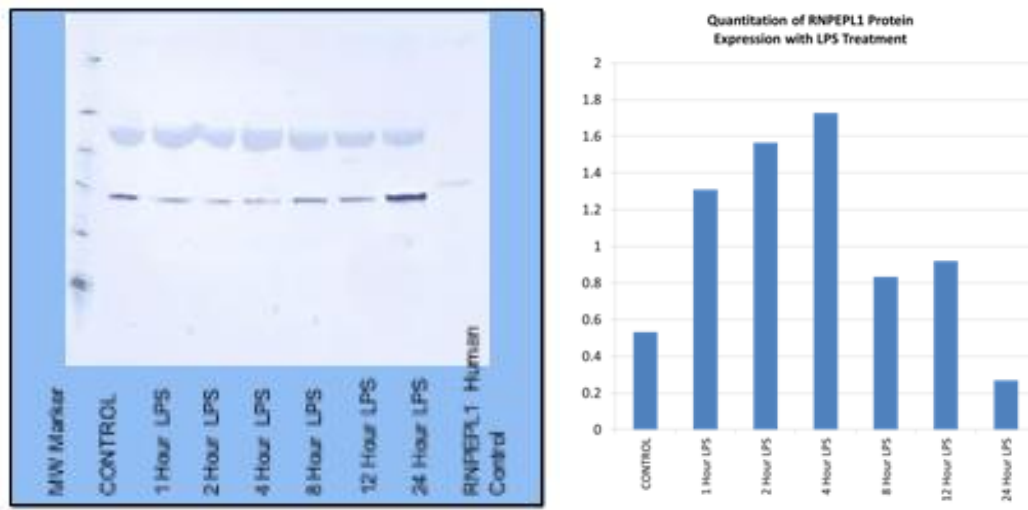
inflammatory levels.

Figure 2: RNPEPL1 protein expression in macrophages. (A) represents a characteristic Western blot from the time point LPS treated macrophages. (B) shows the quantification of the ratios observed between RNPEPL1 and the GAPDH standard control. Measurements are in relative intensity units.

Immunohistochemistry

With *in vitro* studies reporting consistent decreases of RNPEPL1 expression of 25%, an *in vivo* Immunohistochemistry approach was taken. Tissues for IHC were obtained from the West Lab. The tissues were from two groups. Group 1 contained tissues from mice that were given standard care. Group 2 was the LPS group. These mice were treated with one IP injection of bacterial LPS (1mg/kg) and sacrificed 24 hours post

injection. Slides were prepared and processed as described

above. Pictures were taken on an Olympus BX-60 microscope with a 100x

oil immersion objective. Pictures were then analyzed using Image

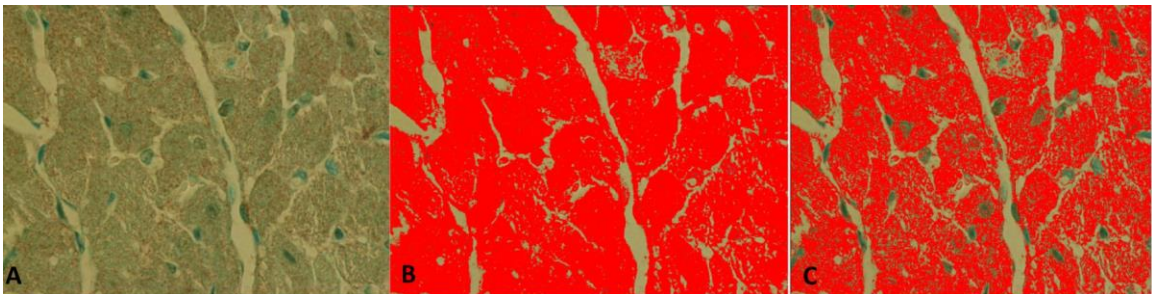J. Tissues studied include heart, kidney, liver, lung, and testes.

Figure 3: Measurement of RNPEPL1 in IHC. (A) represents a tissue sample as

observed. (B) shows the same sample with the cell coverage area

highlighted. (C) Is the same sample with only the red staining highlighted.

Figure 3 shows the calculation process that was performed. (A) shows the

picture as taken. (B) has the total cell area highlighted in red. (C)

represents the red staining area highlighted in red. To calculate the

percentage of red staining, the red stained area was divided by the total

cell area. This was repeated for all pictures in all groups. This method of

calculation was chosen to avoid areas of black space and to ensure that

each slide had comparable cellular coverage for a true comparison.

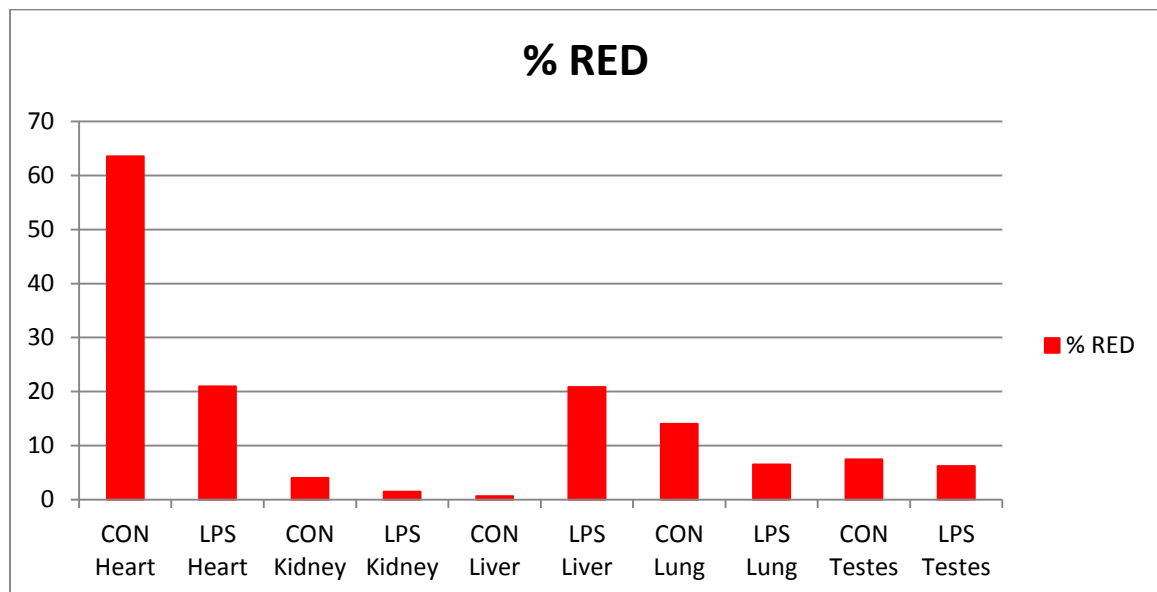Figure 4 shows the percentages of red staining for all groups (labelled on

the graph).

Figure 4: Comparison of RNPEPL1 staining in control and LPS tissues.

The most notable tissue studied was the heart tissue. Figure 5 shows the control heart tissue (A), the LPS treated heart tissue (B), and the no-primary negative control (C). You can see a large difference in staining, with the control tissue containing 63% red (A) and the LPS treated tissue dropping to 20% red (B). The negative control reported 1.5% red (C), which is considered background color.

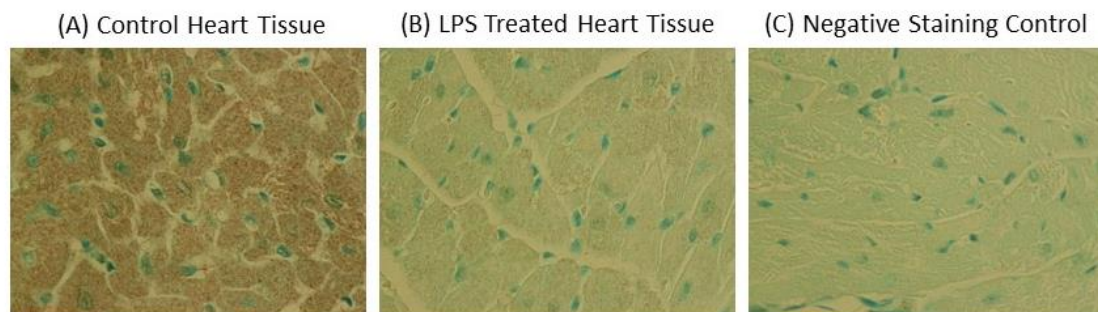Figure 5: Comparison of control and LPS heart tissue. This image shows the variation found between control and LPS treated heart tissue. RNPEPL1 is stained red. The LPS treated tissue (B) shows 43% less red staining than control tissue (A). (C) is an example of the negative staining control results. This was done for each slide to give a measure of any background or non-specific staining present.

Blood Sample Western

Since tissue specific changes in RNPEPL1 expression were found during the IHC study, it was of interest to test the systemic level changes through blood samples. Blood plasma protein was isolated per TRIzol as described above. The samples were analyzed by Western blot analysis. Images were analyzed via Bio-Rad software. Results showed that RNPEPL1 protein expression decrease from 21% - 33%, with the percentage of decrease being significant with all p-values being less than 0.05.

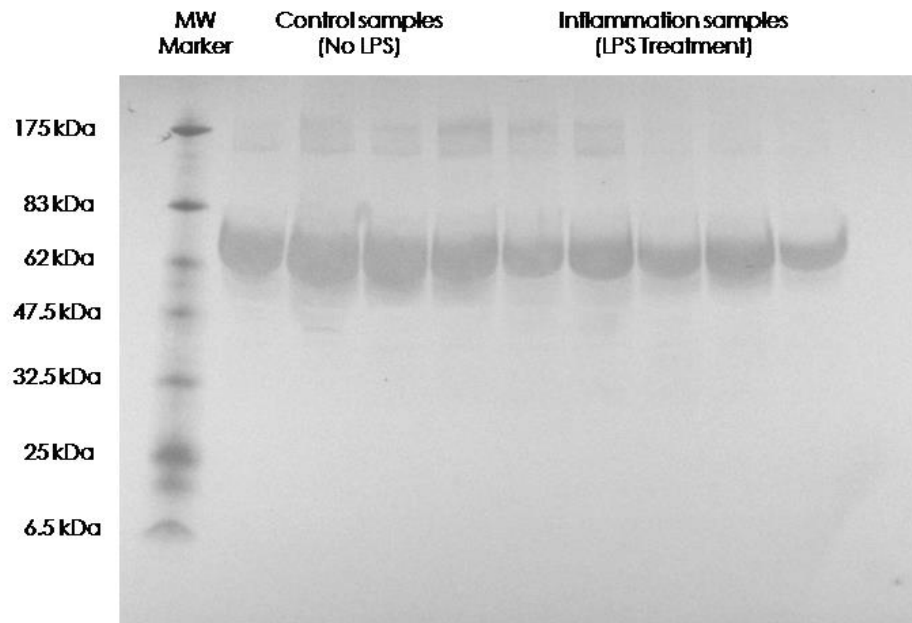Figure 6: RNPEPL1 protein in plasma. RNPEPL1 protein was analyzed in plasma samples from control and LPS treated mice. RNPEPL1 concentration was found to decrease on average by 27%. Note the presence on RNPEPL1 in the plasma, indicating that it is secreted in the blood.

DISCUSSION

Many members of the M1 aminopeptidase family play various roles in inflammation. LTA$_4$H is one member that is a known pro-inflammatory

enzyme.   Given the similarity of RNPEPL1 and LTA$_4$H, it was of interest to determine if RNPEPI1 reacted similarly to LTA$_4$H in inflammatory response. This study intended to determine RNPEPL1 RNA and protein expression at both the cellular and systemic levels. Microarray data determined that RNPEPL1 RNA expression decreased 25% after 24 hours of inflammatory response. Similarly, Western blot analyses showed that RNPEPL1 protein levels decreased on average 25% in the same cellular model. These results led to the continuation of the study at the systemic level. Blood plasma samples and tissue slides were donated from the James West lab for control and LPS groups. The tissue slides showed that RNPEPL1 protein expression change was tissue specific. The greatest change was seen in the heart, with an average of 43% decrease. Lung, kidney, and testes reported small decreases less than 10%, which are not considered significant at this time. Interestingly, liver tissue showed a reverse effect. The control tissue reported less than 2% red area, while the LPS tissue was measured at 21%. These results should be repeated with more samples. If this is in fact a true representation, this may be indicative of a tissue specific role that RNPEPL1 plays in the liver under inflammatory conditions. Lastly, blood plasma samples were analyzed though Western blots. Western were run four times to accommodate for the lack of a loading

standard. Samples were standardized for total protein through a BCA

assay. Protein decreases were significant, ranging from 21% - 33%.

In conclusion, RNPEPL1 expression consistently decreases around 25% after

24 hours of inflammatory response. However, the expression appears to

be tissue specific from the preliminary IHC work conducted. This, along

with the initial increase in RNPEPL1 expression seen in the cellular protein

Western, and the content of RNPEPL1 in the plasma, are areas of future

study outside of the scope of this project.

REFERENCES

[1] Hsieh, C.H.; Rau, C.S.; Jeng, J.C.; Chen, Y.C.; Lu, T.H.; Wu, C.J.; Wu, Y.C.; Tzeng, S.L.; Yang, J.C.S. Whole blood-derived microRNA signatures in mice exposed to lipopolysaccharides. *Journal of Biomedical Science*, **2012,** *19(1), 69.*

[2] Teeling, J.l.; Cunningham, C.; Newman, T.A.; Perry, V.H. The effect of non-steroidal anti-inflammatory agents on behavioural changes and cytokine production following systemic inflammation: Implications for a role of COX-1. *Brain, Behavior, and Immunology,* **2010,** *24(3-4), 409-419.*

[3] Tsujimoto, M.; Hattori, A. The oxytocinase subfamily of M1 aminopeptidases. *Biochimica et Biophysica Acta - Proteins and Proteomics,* **2005,** *1751(1), 9-18.*

[4] Albiston, A.L.; Fernando, R.; Ye, S.; Peck, G.R.; Chai, S.Y. Alzheimer's, angiotensin IV and an aminopeptidase. *Biological & Pharmaceutical Bulletin,* **2004,** *27(6), 765 – 767.*

[5] Tanioka, T.; Hattori, A.; Masuda, S.; Nomura, Y.; Nakayama, H.; Mizutani, S.; Tsujimoto, M. Human Leukocyte-derived Arginine Aminopeptidase. The third member of the oxytocinase subfamily of aminopeptidases. *Journal of Biological Chemistry,* **2003,** *278(34), 32275-32283.*

[6] Maruyama, M.; Hattori, A.; Goto, Y.; Ueda, M.; Maeda, M.; Fujiwara, H.; Tsujimoto, M. Laeverin / Aminopeptidase Q, a Novel Bestatin-sensitive Leucine Aminopeptidase Belonging to the M1 Family of Aminopeptidases. *Journal of Biological Chemistry,* **2007,** *282(2), 20088 - 20096.*

[7] Albiston, A.L.; Ye, S.; Chai, S.Y. Membrane Bound Members of the M1 Family: More Than Aminopeptidases. *Protein & Peptide Letters,* **2004,** *11(5), 491-500.*

[8] Diaz-Perales, A.; Quesada, V.; Sanchez, L.M.; Ugalde, A.P.; Suarez, M.F.; Fueyo, A.; Lopez-Otin, C. Identification of Human Aminopeptidase O, a Novel Metalloprotease with Structural Similarity to Aminopeptidase B and Leukotriene A4 Hydrolase. *Journal of Biological Chemistry,* **2005,** *280(14),* 14310 - 14317.

[9] Benigni, A.; Cassis, P.; Remuzzi, G. Angiotensin II revisited: new roles in inflammation, immunology and aging. *EMBO Molecular Medicine,* **2010,** *2(7), 247-257.*

[10] Haas, C.S.; Creighton, C.J.; Pi, X.; Maine, I.; Koch, A.E.; Haines, G.K.; Ling, S.; Chinnaiyan, A.M.; Holoshitz, J. Identification of genes modulated in rheumatoid arthritis using complementary DNA microarray analysis of lymphoblastoid B cell lines from disease-discordant monozygotic twins. *Arthritis & Rheumatology,* **2006,** *54(7), 2047- 2060.*

[11] Helgstrand, C.; Hasan, M.; Uysal, H.; Haeggstrom, J.Z.; Thunnissen, M. A Leukotriene A4 Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding. *Journal of Molecular Biology,* **2011,** *406(1),* 120 - 134.

[12] Talati, M.; West, J.; Zaynagetdinov, R.; Hong, C.C.; Han, W.; Blackwell, T.; Robinson, L.; Blackwell, T.S.; Lane, K. BMP Pathway Regulation of and by Macrophages. *PLoS One,* **2014,** *9(4),* doi:10.1371/journal.pone.0094119.

[13] Abramoff, M.D.; Magalhaes, P.J.; Ram, S.J. Image Processing with ImageJ. *Biophotonics International*, **2004,** *11(7),* 36-42.

CONCLUSIONS

Arginyl aminopeptidase-like 1 was once a hypothetical reading frame predicted by the genomics boom. Through RNA studies, it was confirmed as a transcribed gene with three splice variants found ubiquitously in humans. The transcript was cloned into yeast, which produced a functional protein. This protein was purified and characterized for enzymatic activity. Based on conserved functional domains, similar genomic structure, and basic enzymatic activity, RNPEPL1 was placed in the M1 family of aminopeptidases.

The M1 family of aminopeptidases is made up of 12 human members that are all believed to be enzymatically active based on conserved active sites and a shared critical zinc binding domain. Many members are functionally characterized, although some are lacking in enzymatic role reports. In order to study this family in depth, various genetic studies were completed.

Multiple alignments allowed visualization of the critically conserved domains. The first was the HEXXHX18E zinc binding domain that contains two of the three active site residues, with the third being a downstream tyrosine. The second was the G/AMEN domain believed to be required for enzymatic activity. It was here that one member of the family,

aminopeptidase O (APO) began to diverge from the family. APO contains a GMAS domain in place of the G/AMEN domain.

The second genetic study completed was an analysis of phylogenetic trees for human, mouse, and rat. At this point, it was discovered that mouse and rat do not contain endoplasmic reticulum aminopeptidase 2 (ERAP2). The trees were created using MEGA 6 with bootstrap values (500 trials) and JTT model substitutions. The trees resulted in nearly identical results. The trees were divided into 2 distinct groups. The first contained LTA_4H, RNPEP, and RNPEP. The second group contained all the remaining members, except APO, falling into 2 subgroups. Subgroup A, the MHC Presentation Group, contained LNPEP, ERAP1, and ERAP2 (in humans). Subgroup B contained ANPEP, ENPEP, NPEPPS, APQ, and TRHDE. APO initially separated with the first group, but then separated itself with no bootstrap value available.

To dig even deeper into the genetic conservation, an analysis of the exon structure for each family member was complete. This was a new approach to better understand the order of divergence and duplication that created the gene family found currently. This study looked at each gene's exons including the length, reading frame, and location of critical domains. Seven of the members aligned very well, with three

participating in exon slicing or joining events. These three are also the three found in subgroup A of the family trees. Subgroup B is also very well represented in the exon conservation study, with 4 of the 5 members aligning. The LTA$_4$H group did not fit into the same reading frame or contain the functional groups in the same exon. This is further evidence that the LTA$_4$H group was an early divergence of the family.

While analyzing the genes, it was found that ERAP2 was missing from mouse and rat. An explanation for this is found in the synteny study performed. The results indicate that ERAP2 may have been destroyed during a chromosomal rearrangement in a group of species in the mouse and rat lineage.

With the family study completed, all results pointed to RNPEPL1 being most closely related to LTA$_4$H and RNPEP, both of which are involved in inflammation. The effects of inflammatory response on RNPEPL1 expression in both RNA and protein were analyzed. At the cellular level, RNPEPL1 RNA expression decreases 25% and protein expression appears to increase early in inflammation (4-8hours) and then decrease by about 22% after 24 hours. Systemically, RNPEPL1 protein expression in the blood plasma was found to decreases from 21% – 33%. It is notable that RNPEPL1 is found in the plasma. This indicates a secretion

of RNPEPL1 in the blood, which is a new discovery. Interestingly, the response is now believed to be tissue specific. Heart tissue shows a 43% decrease in RNPEPL1 protein staining, while lung, kidney, and testes showed slight decreases. Liver increased by 20%, contrary to any other evidence. Although these changes are significant, they are not enough to continue this line of research. Other areas of RNPEPL1 research are, however, still open to exploration, such as the ability of RNPEPL1 to hydrolyze LTA_4 to LTB_4 similar to LTA_4H.

REFERENCES FOR INTRODUCTION AND CONCLUSIONS

[1] Pham; Gouzy-Darmon; Pernier; Hanquez; Hook; Beinfeld; Nicolas; Etchebest; Foulon; Cadel. Mutation in the substrate-binding site of aminopeptidase B confers new enzymatic properties. *Biochimie,* **2011,** *93,* 730-741.

[2] Tsujimoto; Hattori. The oxytocinase subfamily of M1 aminopeptidases. *Biochimica et Biophysica Acta,* **2005,** *1751,* 9-18.

[3] Albiston; Fernando; Ye; Peck; Chai. Alzheimer's, Angiotensin IV, and an Aminopeptidase. *Biological Pharmacology Bulletin,* **2004,** *27(6),* 765-767.

[4] Georgiadou; Hearn; Evnouchidou; Chroni; Leodiadia; York; Rock; Stratikos. PLAP efficiently generates mature antigenic peptides in vitro but in patterns distinct from ERAP1. *Journal of Immunology,* **2010,** *185(3),* 1584-1592.

[5] Tanioka; Hattori; Masuda; Nomura; Nakayama; Mizutani; Tsujimoto. Human Leukocyte-derived Arginine Aminopeptidase. The third member of the oxytocinase subfamily of aminopeptidases. *The Journal of Biological Chemistry,* **2003,** *278(34),* 32275-32283.

[6] Watanabe, Shibata, Kikkawa, et al. Adipocyte-Derived Leucine Aminopeptidase Suppresses Angiogenesis in Human Endometrial Carcinoma via Renin-Angiotensin System. *Clinical Cancer Research,* **2003,** *9,* 6497-6503.

[7] Maruyama; Hattori; Goto; Ueda; Maeda; Fujiwara; Tsujimoto. Laeverin/Aminopeptidase Q, a Novel Bestatin-sensitive Leucine Aminopeptidase Belonging to the M1 Family of Aminopeptidases. *The Journal of Biological Chemistry,* **2007,** *282(28),* 20088-20096.

[8] Fujiwara; Higuchi; Yamada; Hirano; Sato; Nishioka; Yoshioka; Tatsumi; Ueda; Maeda; Fujii. Human extravillous trophoblasts express laeverin, a novel protein that belongs to membrane-bound gluzincin metallopeptidases. *Biochemical and Biophysical Research Communications,* **2004,** *313,* 962-968.

[9] Constam; Tobler; Resing-Ehl; Kemler; Hersh; Fontana. Puromycin-sensitive aminopeptidase. Sequence analysis, expression, and functional characterization. *The Journal of Biological Chemistry,* **1995,** *270,* 26931-26939.

[10] Diaz-Perales; Quesada; Sanchez; Ugalde; Suarez; Fueyo; Lopez-Otin. Identification of Human Aminopeptidase O, a Novel Metalloprotease with Structural Similarity to Aminopeptidase B and Leukotriene A$_4$ Hydrolase. *Journal of Biological Chemistry,* **2005,** *280(14),* 14310-14317.

[11] Papadopoulos; Heuer; Bauer. Analysis of the thyrotropin-releasing hormone-degrading ectoenzyme by site-directed mutagenesis of cysteine residues. *European Journal of Biochemistry,* **2000,** *267,* 2617-2623.

[12] Albiston; Ye; Chai. Membrane Bound Members of the M1 Family: More Than Aminopeptidases. *Protein & Peptide Letters,* **2004,** *11(5),* 491-500.

[13] Ito; Nakajima; Onohara; Takeo; Nakashima; Matsubara; Ito; Yoshimoto. Crystal Structure of Aminopeptidase N (Proteobacteria Alanyl Aminopeptidase) from Escherichia coli and Conformational Change of Methionine 260 Involved in Substrate Recognition. *Journal of Biological Chemistry,* **2006,** *281(44),* 33664-33676.

[14] Helgstrand; Hasan; Uysal; Haeggstrom; Thunnissen. A Leukotriene A$_4$ Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding. *Journal of Molecular Biology,* **2011,** *406,* 120-134.

[15] Pham; Cadel; Gouzy-Darmon; Hanquez; Beinfeld; Nicolas; Etchebest; Foulon. Aminopeptidase B, a glucagon-processing enzyme: site directed mutagenesis of the Zn^{2+}-binding motif and molecular modelling. *BMC Biochemistry,* **2007,** *8(21).* doi:10.1186/1471-2091-8-21.

[16] Piesse; Tymms; Garrafa; Gouzy; Lacasa; Cadel; Cohan; Foulon. Human aminopeptidase B (rnpep) on chromosome 1q32.2: complementary DNA, genomic structure and expression. *GENE,* **2002,** *292,* 129-140.

[17] Cadel; Foulon; Viron; Balogh; Midol-Monnet; Noel; Cohen. Aminopeptidase B from the rat testis is a bifunctional enzyme structurally related to leukotriene-A$_4$ hydrolase. *Proceedings of the National Academy of Science,* **1997,** *94,* 2963-2968.

[18] Fukasawa; Fukasawa; Harada; Hirose; Izumi; Shimizu. Aminopeptidase B is structurally related to leukotriene-A$_4$ hydrolase but is not a bifunctional enzyme with epoxide hydrolase activity. *Biochem. J.,* **1999,** *339, 497-502.*

[19] Thompson; Beasley; Schmidt; Seipelt. Arginyl Aminopeptidase-Like 1 (RNPEPL1) Is an Alternatively Processed Aminopeptidase with Specificity for Methionine, Glutamine, and Citrulline Residues. *Protein & Peptide Letters,* **2009,** *16,* 1256-1266.

[20] Florea; Liliana. Bioinformatics of alternative splicing and its regulation. *Briefings in Bioinformatics,* **2006,** *7(1),* 55-69.