Identification of the Novel Plasminogen Receptor, Plg-R_{KT}

Lindsey A. Miles\textsuperscript{1}, Nicholas M. Andronicos\textsuperscript{2}, Emily I. Chen\textsuperscript{3}, Nagyung Baik\textsuperscript{1}, Hongdong Bai\textsuperscript{4}, Caitlin M. Parmer\textsuperscript{5}, Shahrzad Lighvani\textsuperscript{1}, Samir Nangia\textsuperscript{4,6}, William B. Kiosses\textsuperscript{1}, Mark P. Kamps\textsuperscript{6}, John R. Yates III\textsuperscript{1} and Robert J. Parmer\textsuperscript{4,6}

1. Introduction

1.1 The plasminogen activation system

Initiation of the plasminogen activation system results in generation of the broad spectrum serine protease, plasmin, from the circulating zymogen, plasminogen. Plasminogen is activated to plasmin by plasminogen activators (PA’s), either urokinase type-plasminogen activator (u-PA) or tissue-type plasminogen activator (t-PA), via specific proteolytic cleavage (Castellino & Ploplis, 2005). Plasmin is the major enzyme responsible for degradation of fibrin clots (fibrinolysis) to maintain normal blood homeostasis (Bugge et al., 1995; Ploplis et al., 1995). Dysregulation of the plasminogen activation system can result in hemorrhage (excess fibrinolysis) or thrombosis (insufficient fibrinolysis). The plasminogen activation system is regulated by direct inhibition of plasmin (by the circulating serpin, α\textsubscript{2}-antiplasmin) and by synthesis and secretion of plasminogen activators and the serpin, plasminogen activator inhibitor 1 (PAI-1)] (Collen, 1999). In a key regulatory step, plasminogen activation is promoted when plasminogen and its activator, t-PA, bind concomitantly to lysine residues on the surface of fibrin clots, resulting in a marked reduction in the Km for activation of plasminogen compared with the reaction in solution (Hoylaerts et al., 1982).

1.2 Functions of plasminogen receptors

In the past 25 years an additional mechanism for positive regulation of plasminogen activation has been recognized: co-localization of plasminogen and PA’s on cell surfaces markedly decreases the Km for plasminogen activation in a mechanism analogous to assembly of components of the plasminogen activation system on fibrin (Miles et al., 2005). The plasmin produced is retained on the cell surface where it is protected from its inhibitor, α\textsubscript{2}-antiplasmin (Figure 1) (Hall et al., 1991; Plow et al., 1986). Thus, the cell surface becomes

\textsuperscript{1}The Scripps Research Institute, USA
\textsuperscript{2}CSIRO Livestock Industries, Armidale, NSW, Australia
\textsuperscript{3}Stony Brook University, USA
\textsuperscript{4}Veterans Administration San Diego Healthcare System, USA
\textsuperscript{5}Yale University, USA
\textsuperscript{6}University of California, San Diego, USA
armed with the broad spectrum proteolytic activity of plasmin. Cell surface plasmin plays a key role in processes in which cells must degrade an extracellular matrix in order to migrate, including inflammation (Busuttil et al., 2004; Ploplis et al., 1998; Plow et al., 1999), wound healing (Creemers et al., 2000; Romer et al., 1996), metastasis (Palumbo et al., 2003; Ranson et al., 1998) and neurite outgrowth (Gutierrez-Fernandez et al., 2009; Jacovina et al., 2001). Cell surface plasmin also plays a key role in myogenesis (Lopez-Alemany et al., 2003) and prohormone processing (Jiang et al., 2001, 2002).

![Diagram of Plasmin Activation](image)

Activation of cell-associated plasminogen (Plg) to plasmin (Pm) by cell-associated plasminogen activators (PA) is markedly enhanced compared to the reaction in solution. The Pm formed remains on the cell surface where it is relatively protected from its inhibitor, $\alpha_2$-antiplasmin ($\alpha_2$-AP).

Fig. 1. Enhancement of plasminogen activation on the cell surface.

1.3 Characteristics of plasminogen receptors
Cellular plasminogen binding sites are very broadly distributed on both prokaryotic and eukaryotic cells. Of the many cell types examined to date, only red cells do not exhibit detectable plasminogen binding ability (Miles et al., 2005). The interactions of plasminogen with cells are of very high capacity, reaching $3 \times 10^7$ molecules/cell on lung fibroblasts (Plow et al., 1986), for example. Thus, the plasminogen binding capacity of a cell is made up of contributions from a set of distinct cell surface proteins. An important aspect of the mechanism of promotion of plasminogen activation on cell surfaces is that a subset of carboxypeptidase-B-sensitive plasminogen binding proteins is responsible for enhancement of plasminogen activation on eukaryotic cells. When cells are treated with carboxypeptidase B, the ability to stimulate plasminogen activation is lost (Félez et al., 1996). Since carboxypeptidase B removes C-terminal basic residues, these results imply that proteins exposing C-terminal basic residues on cell surfaces are responsible for stimulation of plasminogen activation (Figure 2,A). Known carboxypeptidase-B-sensitive cell surface plasminogen receptors could previously be divided into two classes: 1) proteins synthesized with C-terminal lysines and having
additional known intracellular functions, including α-enolase (Miles et al., 1991; Redlitz et al., 1995), cytokeratin 8 (Hembrough et al., 1995, 1996), S100A10 (Choi et al., 2003; Kassam et al., 1998), TIP49a (Hawley et al., 2001) and histone H2B (Herren et al., 2006) and; 2) proteins requiring proteolytic processing in order to reveal a C-terminal basic residue (lysine), including actin (Dudani & Ganz, 1996; Miles et al., 2006) and annexin 2 (Hajjar et al., 1994). However, until recently, no integral membrane plasminogen binding proteins synthesized with a C-terminal basic residue had been identified. The existence of a plasminogen receptor with the latter characteristics would reveal a novel mechanism for stimulation of plasminogen activation because release and rebinding of intracellular proteins or proteolytic cleavage of membrane proteins to expose C-terminal basic residues would not be required.

Plasminogen binding to carboxyl terminal lysines on the cell surface. Panel A) The binding of plasminogen to cell surface proteins occurs via receptors exposing carboxyl terminal lysines to the extracellular environment. Cell surface proteins with carboxyl terminal lysines that are masked or in other inaccessible orientations on the cell surface, or membrane-associated proteins with carboxyl terminal lysines that are located on the inner face of the membrane, cannot serve as plasminogen receptors. Panel B) CpB treatment of intact cells removes carboxyl terminal lysines from plasminogen receptors, and plasminogen binding to the cell surface is reduced. Reprinted with permission from (Hawley, Green, and Miles 2000, 84:882-890).

Fig. 2. Plasminogen binding to carboxyl terminal lysines on the cell surface.

1.4 Need for a proteomics approach to identify integral membrane plasminogen receptor(s) with C-terminal basic residues
Previous methodologies and characteristics of plasminogen binding proteins may have precluded identification of an integral membrane plasminogen binding protein with a C-terminal basic residue. The identification of plasminogen receptors has relied previously on cell surface labeling followed by affinity chromatography on plasminogen-Sepharose columns and N-terminal sequencing of fractions eluted from SDS gels. Thus, many
intracellular proteins that are also present on the cell surface were readily identified because protein fractions that bound to plasminogen-Sepharose included the labeled, surface-associated protein, as well as nonlabeled intracellular protein. Using these methods, a lower abundance integral membrane plasminogen binding protein might not have been detectable.

Previously, we used a proteomics approach to examine monocytoid cell membranes for the presence of proteins exposing carboxyl terminal lysines on the extracellular face of the cell membrane (Hawley et al., 2000). We compared plasminogen ligand blots of 2-D gels of membrane fractions of intact cells treated with carboxypeptidase B with untreated membranes (e.g. Figure 2,B). We eluted a prominent carboxypeptidase B-sensitive protein from the 2-D gels and obtained two peptide sequences using tandem mass spectrometry. These peptide sequences corresponded to TATA-binding protein-interacting protein (TIP49a) (Hawley et al., 2001). However, TIP49a is a member of the class of cell surface plasminogen receptors synthesized with a C-terminal lysine and also having intracellular functions and is not an integral membrane protein.

The methodology used to identify TIP49a and other plasminogen receptors has required elution of candidate proteins from 2-D SDS polyacrylamide gels. However, many membrane proteins are not well resolved on SDS polyacrylamide gels. Therefore, we used an isolation method that used column chromatography instead of SDS polyacrylamide gel analysis: We took advantage of the exquisite sensitivity of multidimensional protein identification technology (MudPIT) to search for integral membrane plasminogen receptor(s) exposing a C-terminal basic residue on the cell surface and present on viable cells.

2. Methods

2.1 Plasminogen receptor isolation
Plasminogen receptor isolation was performed as described (Andronicos et al., 2010). Briefly, progenitor and M-CSF-differentiated Hoxa9-ER4 cells were separately biotinylated, using EZ-Link Amine-PEO3-Biotin. The cells were then subjected to dead cell removal on annexin V-coated magnetic microspheres that resulted in a 99% enrichment of viable cells. Membrane fractions were prepared from the viable cells by dounce homogenization in the presence of Complete Protease Inhibitor Cocktail in Invitrosol, followed by centrifugation steps as used in our laboratory (Hawley et al., 2000, 2001) and were applied to a plasminogen-Sepharose affinity column as described (Miles et al., 1991). The column was washed in phosphate buffered saline containing 1 X Invitrosol until no protein was detected at 280 nm followed by elution with the washing buffer containing 0.2 M ε-aminocaproic acid (EACA). The eluant from the plasminogen-Sepharose column was incubated with 50μl of immobilized avidin for 30 minutes at 4°C. The proteins bound to the immobilized avidin were resuspended in Invitrosol and heated at 60°C. Then, 80% acetonitrile was added and the samples were digested with trypsin. After 24 h, the solvent was evaporated in a speedvac, and peptides were dissolved in buffer A (95% H2O, 5% acetonitrile, and 0.1% formic acid).

2.2 Multidimensional chromatography and tandem mass spectrometry
Multidimensional chromatography and tandem mass spectrometry were performed as described (Andronicos et al., 2010). Briefly, the protein digest was subjected to MudPIT [reviewed in (Eng et al., 1994)]. Peptide mixtures were resolved by strong cation exchange
liquid chromatography upstream of reversed phase liquid chromatography (Larmann, Jr. et al., 1993; Link et al., 1999; Opitcek & Jorgenson, 1997; Wolters et al., 2001). Eluting peptides were electrosprayed onto an LTQ ion trap mass spectrometer equipped with a nano-LC electrospray ionization source. Full MS spectra were recorded over a 400–1600 m/z range, followed by three tandem mass (MS/MS) events sequentially generated in a data-dependent manner on the first, second, and third most intense ions selected from the full MS spectrum (at 35% collision energy). Mass spectrometer scan functions and HPLC solvent gradients were controlled by the Xcalibur data system.

2.3 Database search and interpretation of MS/MS datasets
Database searching and interpretation of MS/MS Datasets were performed as described (Andronicos et al., 2010). Briefly, tandem mass spectra were extracted from raw files, and a binary classifier (Bern et al., 2004), previously trained on a manually validated data set, was used to remove low quality MS/MS spectra. Remaining spectra were searched against a mouse protein database containing 50,370 protein sequences downloaded as FASTA-formatted sequences from EBI-IPI and 124 common contaminant proteins, for a total of 66,743 target database sequences (Peng et al., 2003). To calculate confidence levels and false positive rates, a decoy database containing the reverse sequences of the 66,743 proteins appended to the target database and the SEQUEST algorithm (Yates, III, 1998) were used to find the best matching sequences from the combined database.

SEQUEST searches were done on an Intel Xeon 80-processor cluster running under the Linux operating system. The peptide mass search tolerance was set to 3 Da. No differential modifications were considered. No enzymatic cleavage conditions were imposed on the database search, so the search space included all candidate peptides whose theoretical mass fell within the 3 Da mass tolerance window, despite their tryptic status. The validity of peptide/spectrum matches was assessed in DTASelect2 (Tabb et al., 2002) using SEQUEST-defined parameters, the cross-correlation score (XCorr) and normalized difference in cross-correlation scores (DeltaCN). The search results were grouped by charge state (+1, +2, and +3) and tryptic status (fully tryptic, half-tryptic, and non-tryptic), resulting in 9 distinct sub-groups. In each one of the sub-groups, the distribution of XCorr and DeltaCN values for direct and decoy database hits was obtained, and the two subsets were separated by quadratic discriminant analysis. Outlier points in the two distributions (for example, matches with very low Xcorr but very high DeltaCN) were discarded. Full separation of the direct and decoy subsets is not generally possible; therefore, the discriminant score was set such that a false positive rate of 5% was determined based on the number of accepted decoy database peptides. This procedure was independently performed on each data subset, resulting in a false positive rate independent of tryptic status or charge state.

3. Results
3.1 Isolation of an integral membrane plasminogen receptor exposing a C-terminal lysine on the cell surface
We used specific proteolysis followed by MudPIT to probe the membrane proteome of differentiated mouse monocyte progenitor cells (Hoxa9-ER4) for the presence of integral membrane plasminogen receptor(s) exposing a C-terminal basic residue on the cell surface, as outlined in Figure 3. [The Hoxa9-ER4 cell line was derived from primary murine bone
marrow myeloid precursors immortalized with an estrogen regulated conditional oncoprotein, HoxA9-ER4 (Wang et al., 2006). The Hoxa9-ER4 line is factor-dependent and differentiates to monocytes when estrogen is removed from the medium, thereby inactivating the Hoxa9-ER protein. The mature monocytes respond to M-CSF (Odegaard et al., 2007). First, the Hoxa9-ER4 monocyte progenitor cells were differentiated with macrophage colony stimulating factor (M-CSF), which induces plasminogen receptors on these cells (Andronicos et al., 2010). Then intact cells were biotinylated. Because early apoptotic and non-viable/necrotic cells exhibit markedly enhanced plasminogen binding ability (Mitchell et al., 2006; O'Mullane & Baker, 1998, 1999) we passed the biotinylated cells over a dead cell removal column to enrich for live cells. Cells were then lysed and membrane fractions prepared and passed over a plasminogen-Sepharose affinity column and specifically eluted with ε-aminocaproic acid (EACA), a lysine analog that blocks the binding of plasminogen to cells (Miles & Plow, 1985). Biotinylated proteins bound to the avidin column and were digested with trypsin while still on the column. The peptide digest was then subjected to MudPIT.

Monocyte (Hoxa9-ER4) progenitor cells were differentiated with macrophage colony stimulating factor (M-CSF), which induces plasminogen receptors (▲) on these cells. Then intact cells were biotinylated (●) and passed over a dead cell removal column. Live cells were then lysed and membrane fractions prepared and passed over a plasminogen-Sepharose affinity column and specifically eluted. Biotinylated plasminogen receptors (▲●) were then bound to an avidin column and digested with trypsin.

Fig. 3. Isolation of plasminogen receptors.
In MudPIT, the peptide mixtures were first resolved by strong cation exchange liquid chromatography followed by reversed phase liquid chromatography. Eluting peptides were electrosprayed onto an LTQ ion trap mass spectrometer and full MS spectra were recorded over a 400-1600 m/z range, followed by three tandem mass events. The spectra obtained were searched against a mouse protein database. Using this method, only one protein with a predicted transmembrane sequence and a C-terminal basic residue was identified: the hypothetical protein, C9orf46 homolog (IPI 00136293), homologous to the protein predicted to be encoded by human chromosome 9, open reading frame 46. The peptides corresponding to C9orf46 homolog that were obtained in the MudPIT analysis are shown in Table 1. We have designated the protein, Plg-R\textsubscript{KT}, to indicate a plasminogen receptor with a C-terminal lysine and having a transmembrane domain (see below).

A limitation of shotgun proteomics, such as MudPIT, is that they typically under sample a proteome because they use data dependent data acquisition (a computer-driven data acquisition approach). This can lead to variations in the proteins identified, particularly amongst the lower abundant proteins. Thus, we cannot exclude the possibility that other membrane proteins exposing C-terminal lysines were present in the membrane proteome.

### Table 1. Peptides obtained corresponding to C9orf46 homolog

<table>
<thead>
<tr>
<th>Xcorr</th>
<th>DeltCN</th>
<th>Conf%</th>
<th>ObsM+H+</th>
<th>CalcM+H+</th>
<th>Peptide Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8378</td>
<td>0.2884</td>
<td>99.8%</td>
<td>2195.8743</td>
<td>2196.4788</td>
<td>K.SMENMKNQEQFMVTHAR.L (3+)</td>
</tr>
<tr>
<td>2.6749</td>
<td>0.1167</td>
<td>95.2%</td>
<td>1359.5521</td>
<td>1361.5181</td>
<td>K.NQEQFMVTHAR.L (2+)</td>
</tr>
<tr>
<td>2.6771</td>
<td>0.2534</td>
<td>99.8%</td>
<td>1160.4321</td>
<td>1160.3514</td>
<td>R.HTMQNEMR.E (2+)</td>
</tr>
<tr>
<td>4.7468</td>
<td>0.3052</td>
<td>100%</td>
<td>1523.5322</td>
<td>1523.6954</td>
<td>R.MKSEAEDILETEK.T (2+)</td>
</tr>
<tr>
<td>5.1774</td>
<td>0.3788</td>
<td>100%</td>
<td>2335.5544</td>
<td>2333.6997</td>
<td>R.MKSEAEDILETEKTNLPLK.G (3+)</td>
</tr>
<tr>
<td>3.775</td>
<td>0.3164</td>
<td>100%</td>
<td>1264.0922</td>
<td>1264.3287</td>
<td>K.SEADELETKE (2+)</td>
</tr>
<tr>
<td>2.995</td>
<td>0.0655</td>
<td>96.8%</td>
<td>1137.2722</td>
<td>1137.3184</td>
<td>K.GLITFESLEK.R (2+)</td>
</tr>
<tr>
<td>2.893</td>
<td>0.2591</td>
<td>99.7%</td>
<td>1364.4922</td>
<td>1364.5848</td>
<td>K.GLITFESLEKAR.R (2+)</td>
</tr>
</tbody>
</table>

SEQUEST-defined parameters (Xcorr, DeltCN, and Conf%) are shown for each peptide. (core: cross-correlation score; DeltCN: normalized difference in cross-correlation scores; Conf%: confidence level of the peptide; ObsM+H+: observed peptide mass; CalcM+H+: theoretical peptide mass). Observed peptide mass, theoretical peptide mass, and charges of the peptide identified (3+ or 2+) are also shown to demonstrate accurate peptide identification. This research was originally published in Blood, Andronicos, N.M., Chen, E.I., Bai, N., Bai, H., Parmer, C.M., Kiosses, W.B., Kamps, M.P., Yates, J.R., III, Parmer, R.J., Miles, L.A., Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-R\textsubscript{KT}, a major regulator of cell surface plasminogen activation, Blood. 2010, 115: 1319-30.

A key advantage of MudPIT is that proteins in a given proteome can be identified simultaneously. As proof of principle of our isolation method, peptides corresponding to other proteins previously identified as plasminogen binding proteins on monocytes were also detected in the membrane preparations: α-enolase, gamma actin, S100A10, histone H2B, annexin 2, and β\textsubscript{2} integrin.
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>VTARALQER</td>
<td>HLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Human</td>
<td>MGIF1KGMN</td>
<td>EINMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQSKME</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Rat</td>
<td>MGIF1KGMN</td>
<td>EINMKQKRFPM</td>
<td>VMHARLQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Dog</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LVNARLQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Cow (arabidopsis)</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Alpaca</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>VMDWLQER</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Dolphin</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Gibbon</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Guinea Pig</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Horse</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Lemur</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Opossum</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Orangutan</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Panda (Giant)</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Pig</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Rabbit</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Rhesus Monkey</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Tarsier</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Tree Shrew</td>
<td>MGIF1KGMN</td>
<td>ENMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Lizard (arboval)</td>
<td>MGIF1KGMN</td>
<td>EINMKQKRFPM</td>
<td>LNNRALQER</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
<tr>
<td>Frog (xenopus)</td>
<td>MGSLIKGLIT</td>
<td>TQMKQKQIT</td>
<td>LMNQAQIQLE</td>
<td>QLTMQNMKE</td>
<td>RGMQIAWS</td>
<td>REFLKFGTF</td>
</tr>
</tbody>
</table>

Table 2. Alignment of Orthologs of Plg-R_{KT}
3.2 Conservation of Plg-R_{K_T} across species

The C9orf46 homolog/Plg-R_{K_T} murine DNA sequence encodes a protein of 147 amino acids with a molecular mass of 17,261 Da and a C-terminal lysine (Table 2, first line). We blasted the C9orf46 homolog/Plg-R_{K_T} sequence against all species using NCBI Blast and obtained unique human, rat, dog, cow, giant panda, gibbon, horse, pig, rabbit, and rhesus monkey predicted orthologs, with high identity and homology (e.g. human versus rhesus monkey = 99% similarity), high identity (e.g. human vs rhesus monkey = 98% identity) and no gaps in the sequence (Table 2). Of key importance, a C-terminal lysine was predicted for all of the mammalian orthologs obtained in the blast search. In a query of the Ensembl Gene Report, DNA sequences of all 10 other sequenced mammalian orthologs encoded C-terminal lysines (Table 2).

In addition, the DNA sequences of xenopus and the green lizard also encoded C-terminal lysines (Table 2). Furthermore, Plg-R_{K_T} orthologs with 149 amino acids with a C-terminal lysine were encoded in bony fish (salmon and zebrafish) and the high similarity with a mammalian ortholog is illustrated in the alignment of the mouse and zebrafish proteins in Table 3.

The Plg-R_{K_T} sequence also encodes a putative conserved DUF2368 domain (encompassing amino acids 1-135), an uncharacterized protein with unknown function conserved from nematodes to humans. Notably, Plg-R_{K_T} orthologs of lower organisms were of different predicted lengths and did not consistently predict C-terminal lysines. It is interesting to note that the evolutionary origin of plasminogen is currently believed to originate with protochordates (Liu & Zhang, 2009), so that lower organisms without plasminogen would not utilize the C-terminal lysine of Plg-R_{K_T} to bind plasminogen.

| Mouse   | 1 | GFIFISKMNENMKQNQFMEVTHARLQLERHLMQMNERQMQAMQIAMWSRFKLYFGTF | 60 |
| Zebrfish| 1 | GFIFISKMNENMKQNQFMEVTHARLQLERHLMQMNERQMQAMQIAMWSRFKLYFGTF | 60 |
| Mouse   | 61 | FGATISLATGALKRRKPAFLVPIVPLSFITYQYDLGYGTLLIQRMKSEAEDILETE | 120 |
| Zebrfish| 61 | FGATISLATGALKRRKPAFLVPIVPLSFITYQYDLGYGTLLIQRMKSEAEDILETE | 120 |
| Mouse   | 121 | LELPGLITTFESLEKARERQSKL--FSOK | 147 |
| Zebrfish| 121 | LELPGLITTFESLEKARERQSKL--FSOK | 147 |

Table 3. Alignment of Mouse and Zebrafish Plg-R_{K_T} Sequences

It is also noteworthy that the primary sequence of C9orf46 homolog/Plg-R_{K_T} is apparently tightly conserved in humans, with no validated polymorphisms (eSNPs) within the 6 exons encoded by the gene (on chromosome 9p24.1) in the NCBI human genome sequence variation database (dbSNP, http://www.ncbi.nlm.nih.gov/SNP).

3.3 Topology of Plg-R_{K_T}

The C9orf46 homolog/Plg-R_{K_T} sequence was analyzed in the TMpred site (www.ch.embnet.org/cgi-bin/TMPRED). The model predicted two transmembrane helices extending from F_{53}-L_{73} (secondary helix, oriented from outside the cell to inside the cell) and P_{78}-Y_{99} (primary helix, oriented from inside the cell to outside the cell) (Figure 4). Hence a 52 amino acid N-terminal region and a 48 amino acid C-terminal tail with a C-terminal lysine were predicted to be exposed on the cell surface.

Fig. 4. Structural model of Plg-R KT.

We experimentally tested predictions of the model. First, we raised a monoclonal antibody against the synthetic peptide, CEQSKLFSDK (corresponding to the nine C-terminal amino acids of murine Plg-R KT with an amino terminal cysteine added for coupling). To examine subcellular localization, membrane and cytoplasmic fractions from progenitor and differentiated Hoxa9-ER4 monocyte progenitor cells were electrophoresed and western blotted with anti-Plg-R KT mAb or isotype control mAb. A specific immunoreactive band migrating with an Mr app of ~17,000, was detected in membrane fractions of differentiated monocyte progenitor cells, clearly demonstrating the existence of this new protein (Figure 5,A). The protein was not detected in undifferentiated cells or in the cytoplasmic fraction of the differentiated cells.

To test the prediction that Plg-R KT is an integral membrane protein, membranes from differentiated monocyte progenitor cells were subjected to phase separation in Triton X-114 as described (Bordier, 1981; Estreicher et al., 1989). In this method, integral membrane proteins form mixed micelles with the nonionic detergent and are recovered in the Triton X-114 detergent phase, whereas hydrophilic proteins remain in the aqueous phase. An immunoreactive band migrating with an Mr app of ~17,000 was detected in the detergent phase in western blotting with anti-Plg-R KT mAb, but was not detected in the aqueous phase (Figure 5,B). These data support the prediction that Plg-R KT is an integral membrane protein.

To experimentally test whether the C-terminal lysine of Plg-R KT was exposed on the cell surface, we treated intact biotinylated cells with carboxypeptidase B prior to performing our isolation procedure. Under this condition, C-terminal lysines exposed on the cell surface are...
Identification of the Novel Plasminogen Receptor, Plg-RKT

A. Membrane fractions or cytoplasmic fractions from either undifferentiated or M-CSF-treated Hoxa9-ER4 cells were electrophoresed on 12% sodium dodecyl sulfate polyacrylamide gels under reducing conditions and western blotted with either anti-Plg-RKT mAb, anti-α-enolase mAb as a loading control, or isotype control mAb. B. M-CSF-treated Hoxa9-ER4 cell membranes were solubilized in 3% Triton X-114. After heating at 37°C and separation of the phases by centrifugation, an aliquot of both phases was electrophoresed and western blotted with anti-Plg-RKT mAb. This research was originally published in *Blood*, Andronicos, N.M., Chen, E.I., Baik, N., Bai, H., Parmer, C.M., Kiosses, W.B., Kamps, M.P., Yates, J.R., III, Parmer, R.J., Miles, L.A., Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-RKT, a major regulator of cell surface plasminogen activation, *Blood*. 2010, 115: 1319-30.

Fig. 5. Plg-RKT behaves as a regulated integral membrane protein.

removed but intracellular C-terminal lysines are protected (see Figure 2,B). Under this condition, no peptides corresponding to Plg-RKT were obtained in the MudPIT analysis, consistent with cell surface exposure of the C-terminal lysine of Plg-RKT.

In order to experimentally evaluate whether the N-terminus of Plg-RKT was exposed on the cell surface, PC12 (rat pheochromocytoma) cells were stably transfected with V5-pClneo-Plg-RKT that expressed a V5 tag at the N-terminus of Plg-RKT. (The V5 sequence was added in front of the mammalian expression vector, pClneo using PCR and then the full-length 443 bp Plg-RKT cDNA was subcloned into the V5-pClneo vector using the XhoI and SmaI cloning sites. Constructs were transfected into cells using Lipofectamine 2000 and stable transfectants were selected with G418.)

A specific band migrating with a Mrapp of 17,000 was detected in cell membranes of the stably transfected cells with both anti-V5 mAb and anti-Plg-RKT mAb (Figure 6, lane 1). The band was not detected by either mAb after trypsin digestion of the isolated membrane fraction (lane 2). When intact cells were incubated with trypsin and the trypsin neutralized with SBTI prior to preparation of the membrane fraction, the majority of the band detectable with either anti-V5 or anti-Plg-RKT was lost (lane 4). In controls, treatment with soybean trypsin inhibitor (SBTI) fully neutralized the ability of trypsin to degrade the V5-tagged Plg-RKT in purified membrane fractions (lane 3), demonstrating that the trypsin had been neutralized prior to membrane fractionation of the treated cells. These results suggest that the N-terminus of Plg-RKT is accessible to trypsin proteolysis of intact cells and is, therefore,
Membrane fractions of PC12 cells stably transfected with V5-pCIneo-Plg-RKT were incubated with either buffer (lane 1), trypsin (1mg/ml) (lane 2) or trypsin 1mg/ml + soybean trypsin inhibitor (SBTI) (2mg/ml) (lane 3) for 30 minutes at 37°C or intact PC12 cells were incubated with 1mg/ml trypsin for 2 hr at 37°C, followed by 2mg/ml SBTI for 15 min. Following neutralization of trypsin with SBTI, the membrane fraction was prepared from the treated, intact cells (lane 4). 30µg/lane of membrane fractions were electrophoresed on 18% SDS PAGE under reducing conditions and western blotted with either anti-V5, anti-Plg-RKT mAb or isotype control.

Fig. 6. The N-termini and C-termini of Plg-RKT are exposed on the cell surface.

exposed on the extracellular face. Furthermore, because the anti-Plg-RKT mAb reacts with the C-terminus of Plg-RKT, these data also confirm the exposure of the C-terminus on the extracellular face of the cell membrane.

3.4 Role of the C-terminal lysine of Plg-RKT in plasminogen binding

We further addressed the exposure of the C-terminus of Plg-RKT on the cell surface using confocal microscopy with a mAb raised against the Plg-RKT C-terminal peptide. (The mAb reacted with the C-terminal peptide of murine Plg-RKT and blocked plasminogen binding to CEQSKLFSDK). When cells were incubated with anti-Plg-RKT mAb and a polyclonal anti-plasminogen antibody, Plg-RKT and plasminogen were both immunodetected in small aggregates dispersed over the cell surface (Figure 7,A), in a similar distribution to that published for confocal analyses of monocyte-associated plasminogen (Das et al., 2007). Most importantly, after preincubation of monocytes with plasminogen, immunodetection of Plg-RKT was reduced by half (Figure 7,A,B). These results demonstrate that the C-terminus of Plg-RKT is exposed on the cell surface. Furthermore, these results show that plasminogen binds to the C-terminal domain of Plg-RKT on the cell surface.
A. M-CSF-differentiated (Hoxa9-ER4) cells were grown on coverslips and preincubated with either phosphate buffered saline (- plasminogen) or 2 μM plasminogen (+ plasminogen), then fixed in 1% formaldehyde, washed and stained with polyclonal anti-plasminogen IgG or anti-Plg-R\textsubscript{KT} mAb and stained with a combination of Alexa 488- F(ab')\textsubscript{2} of goat anti-rabbit IgG and Alexa 568- F(ab')\textsubscript{2} fragment of goat anti-mouse IgG.

B. The number and size of each labeled aggregate was determined. The results reflect counts from over 40 cells in 2 independent experiments. Data represent mean ± SEM. *p < 0.001. This research was originally published in Blood, Andronicos, N.M., Chen, E.I., Baik, N., Bai, H., Parmer, C.M., Kiosses, W.B., Kamps, M.P., Yates, J.R., III, Parmer, R.J., Miles, L.A., Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-R\textsubscript{KT}, a major regulator of cell surface plasminogen activation, Blood. 2010, 115: 1319-30.

Fig. 7. Plg-R\textsubscript{KT} binds plasminogen on the cell surface.

To further address the plasminogen binding function of the C-terminus of Plg-R\textsubscript{KT}, we tested whether the synthetic peptide, corresponding to the C-terminus of Plg-R\textsubscript{KT} could bind plasminogen. The peptide, CEQSKLFSDK, was coupled to BSA and then coated onto wells of microtiter plates. Biotinylated Glu-plasminogen was incubated with the wells and specific binding was detected with HRP-streptavidin (Figure 8). We tested the ability of the soluble C-terminal peptide to inhibit Glu-plasminogen binding under solution phase equilibrium.
conditions. The soluble peptide competed for Glu-plasminogen binding in a dose-dependent manner with an IC$_{50}$ of 2 μM (Figure 8), similar to the Kd values we have previously determined for Glu-plasminogen binding to cells (Miles et al., 2005). In addition, a mutated peptide with the C-terminal lysine substituted with alanine did not compete for plasminogen binding at concentrations up to 1 mM (Figure 8), further supporting the role of the C-terminal lysine in the interaction of Plg-R$_{KT}$ with plasminogen.

The peptide, CEQSKLFSDK, was coupled to BSA and immobilized on microtiter wells. Biotinylated-Glu-plasminogen (25 nM) was incubated with immobilized CEQSKLFSDK in the presence of increasing concentrations of CEQSKLFSDK (●) or a K147A mutant peptide, CEQSKLFSDA (○). Biotinylated Glu-plasminogen binding was detected with HRP-streptavidin. Data are as mean ± SEM, n=3, for each determination. This research was originally published in Blood, Andronicos, N.M., Chen, E.I., Baik, N., Bai, H., Parmer, C.M., Kiosses, W.B., Kamps, M.P., Yates, J.R., III, Parmer, R.J., Miles, L.A., Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-R$_{KT}$, a major regulator of cell surface plasminogen activation, Blood. 2010, 115: 1319-30.

Fig. 8. Plasminogen binds to the C-terminal peptide of Plg-R$_{KT}$.

3.5 Plg-R$_{KT}$ regulates cell surface plasminogen activation

We verified that plasminogen activation was promoted in the presence of differentiated Hoxa9-ER4 cells. Plasminogen activation was stimulated 12.7-fold in the presence of differentiated monocyte progenitor cells, compared to the reaction in the absence of cells (Figure 9). In order to test the role of Plg-R$_{KT}$ in plasminogen activation, we tested the effect of anti-Plg-R$_{KT}$ mAb raised against the synthetic peptide, CEQSKLFSDK. Anti-Plg-R$_{KT}$ mAb substantially suppressed cell-dependent plasminogen activation (Figure 9). In controls, plasminogen activation in the absence of cells was not affected by anti-Plg-R$_{KT}$ mAb.

3.6 Tissue and cellular distribution and regulation of the Plg-R$_{KT}$ transcript

We searched results of gene expression array analyses for expression of the C9orf46 homolog/Plg-R$_{KT}$ transcript. The transcript is broadly expressed in normal human and mouse tissues [as determined in high-throughput gene expression profiling in which RNA samples from human and murine tissues were hybridized to high-density gene expression arrays (Su et al., 2002; Su et al., 2004)]. The C9orf46 homolog/Plg-R$_{KT}$ transcript has been
Plasminogen activation was determined in either the presence or absence of differentiated Hoxa9-ER4 cells and in the presence of either anti-Plg-R<sub>KT</sub> mAb (filled bars) or isotype control rat IgG2a (open bars). ***p < 0.001, compared to the corresponding isotype control. This research was originally published in Blood, Andronicos, N.M., Chen, E.I., Baik, N., Bai, H., Parmer, C.M., Kiosses, W.B., Kamps, M.P., Yates, J.R., III, Parmer, R.J., Miles, L.A., Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-R<sub>KT</sub>, a major regulator of cell surface plasminogen activation, Blood. 2010, 115: 1319-30.

Fig. 9. Plg-R<sub>KT</sub> regulates cell surface plasminogen activation.

detected in spleen, lymph node, thymus, bone marrow, lung, intestine, adrenal, pituitary, and other endocrine tissues, vascular tissue, kidney, liver, stomach, bladder, and neuronal tissue (hippocampus, hypothalamus, cerebellum, cerebral cortex, olfactory bulb and dorsal root ganglion) (Table 4).

We also searched for C9orf46 homolog/Plg-R<sub>KT</sub> mRNA microarray expression data at http://www.ebi.ac.uk/microarray-as/aew/. C9orf46 homolog/Plg-R<sub>KT</sub> mRNA is present in monocytes, leukocytes, natural killer (NK) cells, T cells, myeloid, dendritic, and plasmacytoid cells, breast cancer, acute lymphoblastic leukemia and Molt-4 acute lymphoblastic leukemia cells (Table 5).

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spleen</td>
<td>Liver</td>
</tr>
<tr>
<td>Thymus</td>
<td>Stomach</td>
</tr>
<tr>
<td>Lymph Node</td>
<td>Bladder</td>
</tr>
<tr>
<td>Lung</td>
<td>Hippocampus</td>
</tr>
<tr>
<td>Intestine</td>
<td>Hypothalamus</td>
</tr>
<tr>
<td>Bone Marrow</td>
<td>Cerebellum</td>
</tr>
<tr>
<td>Adrenal</td>
<td>Cerebral Cortex</td>
</tr>
<tr>
<td>Pituitary</td>
<td>Olfactory Bulb</td>
</tr>
<tr>
<td>Vascular Tissue</td>
<td>Dorsal Root Ganglion</td>
</tr>
<tr>
<td>Kidney</td>
<td></td>
</tr>
</tbody>
</table>

Results of high-throughput gene expression profiling (54).

Table 4. Tissue Distribution of Plg-R<sub>KT</sub> mRNA
These data are consistent with previous reports documenting expression of plasminogen binding sites on peripheral blood leukocytes (Miles & Plow, 1987), breast cancer cells (Correc et al., 1990; Ranson et al., 1998) and other tissues [reviewed in (Miles et al., 2005)]. In addition, results obtained by searching the ArrayExpress Warehouse (http://www.ebi.ac.uk/microarray) indicated that the C9orf46 homolog gene is also regulated in other tissues by lipopolysaccharide, aldosterone, canrenoate, H2O2, and dexamethasone (Table 6).

In a previously published genome-scale quantitative image analysis, overexpression of a cDNA that we now recognize to be the Plg-RKT cDNA, resulted in dramatic increases in cell proliferation whereas knockdown of the corresponding mRNA resulted in apoptosis (Harada et al., 2005). Consistent with an anti-apoptotic role of Plg-RKT, we have shown that cell-bound plasminogen inhibits TNFα-induced apoptosis (Mitchell et al., 2006). In microarray studies, C9orf46 homolog mRNA expression has a high power to predict cervical lymph node metastasis in oral squamous cell carcinoma (Nguyen et al., 2007).

### Table 6. Regulation of C9orf46 homolog/Plg-RKT mRNA in Tissues

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Tissue</th>
<th>Agonist</th>
<th>Effect*</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-MEXP-420</td>
<td>Hippocampal microglial cells</td>
<td>lipopolysaccharide</td>
<td>↓</td>
</tr>
<tr>
<td>E-TABM-229</td>
<td>Kidney</td>
<td>aldosterone</td>
<td>↓</td>
</tr>
<tr>
<td>E-TABM-229</td>
<td>Kidney</td>
<td>canrenoate</td>
<td>↓</td>
</tr>
<tr>
<td>E-MEXP-710</td>
<td>Cholinergic cells</td>
<td>H2O2</td>
<td>↓</td>
</tr>
<tr>
<td>E-MEXP-774</td>
<td>Preadipocytes</td>
<td>Dexamethasone</td>
<td>↓</td>
</tr>
</tbody>
</table>

Data were obtained from ArrayExpress Warehouse (http://www.ebi.ac.uk/microarray).


### 4. Conclusions

In conclusion, MudPIT has allowed us to identify a new protein, Plg-RKT, a novel plasminogen receptor with unique characteristics: integral to the cell membrane and exposing a C-terminal lysine on the cell surface in an orientation to bind plasminogen and promote plasminogen activation. Thus, Plg-RKT is likely to play a key role in plasminogen-dependent functions of cells including inflammation, wound healing, development, metastasis, neurite outgrowth, fibrinolysis, myogenesis and prohormone processing. The
Identification of the Novel Plasminogen Receptor, Plg-RK

broad distribution of the Plg-RK transcript and its regulation in tissues that have been demonstrated to express plasminogen binding sites, suggest that Plg-RK provides plasminogen receptor function that may serve to modulate plasminolytic functions (both physiologic and pathologic) specific to a large number of tissues. Furthermore, the potential function of Plg-RK in the regulation of apoptosis and proliferation may play a key role in cancer and metastasis. Future studies with knockout mice should build on our initial results using MudPIT to elucidate the role of Plg-RK.

5. Acknowledgments

Supported by National Institutes of Health Grants (HL38272, HL45934, and HL081046 to L.A.M., HL50398 to R.J.P., NIH P41 RR011823 to J.R.Y., NIAID sub-contract grant UCSD/MCB0237059 to E.I.C.) and Department of Veterans Affairs to R.J.P. S.L was supported by NIH training grant, T32 HL007195. We thank Dr. Ray Stevens at The Scripps Research Institute and Dr. Nuala Booth and Dr. Ian Booth, University of Aberdeen, for helpful discussions. We thank Ms. Linda Bonafede for manuscript preparation. This is publication #21378 from The Scripps Research Institute.

6. References


Biomedical research has entered a new era of characterizing a disease or a protein on a global scale. In the post-genomic era, Proteomics now plays an increasingly important role in dissecting molecular functions of proteins and discovering biomarkers in human diseases. Mass spectrometry, two-dimensional gel electrophoresis, and high-density antibody and protein arrays are some of the most commonly used methods in the Proteomics field. This book covers four important and diverse areas of current proteomic research: Proteomic Discovery of Disease Biomarkers, Proteomic Analysis of Protein Functions, Proteomic Approaches to Dissecting Disease Processes, and Organelles and Secretome Proteomics. We believe that clinicians, students and laboratory researchers who are interested in Proteomics and its applications in the biomedical field will find this book useful and enlightening. The use of proteomic methods in studying proteins in various human diseases has become an essential part of biomedical research.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
