Structural and phylogenetic characterization of human SLURP-1, the first secreted mammalian member of the Ly-6/uPAR protein superfamily

KNUT ADERMANN,¹ FRANK WATTLER,² SIGRID WATTLER,^{1,2} GABRIELE HEINE,¹
MARKUS MEYER,¹ WOLF-GEORG FORSSMANN,¹ AND MICHAEL NEHLS^{1,2}

¹Lower Saxony Institute for Peptide Research (IPF), Feodor-Lynen-Strasse 31, D-30625 Hannover, Germany

²Lexicon Genetics Inc., 4000 Research Forest Drive, The Woodlands, Texas 77381

(RECEIVED August 5, 1998; ACCEPTED November 19, 1998)

Abstract

Members of the Ly-6/uPAR protein family share one or several repeat units of the Ly-6/uPAR domain that is defined by a distinct disulfide bonding pattern between 8 or 10 cysteine residues. The Ly-6/uPAR protein family can be divided into two subfamilies. One comprises GPI-anchored glycoprotein receptors with 10 cysteine residues. The other subfamily includes the secreted single-domain snake and frog cytotoxins, and differs significantly in that its members generally possess only eight cysteines and no GPI-anchoring signal sequence. We report the purification and structural characterization of human SLURP-1 (secreted mammalian Ly-6/uPAR related protein 1) from blood and urine peptide libraries. SLURP-1 is encoded by the ARS (component B)-81/s locus, and appears to be the first mammalian member of the Ly-6/uPAR family lacking a GPI-anchoring signal sequence. A phylogenetic analysis based on the SLURP-1 primary protein structure revealed a closer relationship to the subfamily of cytotoxins. Since the SLURP-1 gene maps to the same chromosomal region as several members of the Ly-6/uPAR subfamily of glycoprotein receptors, it is suggested that both biologically distinct subfamilies might have co-evolved from local chromosomal duplication events.

Keywords: CD59; disulfide bonds; frog toxin; gene locus; Ly-6 antigen; secretion; snake toxin; urokinase-type plasminogen activator receptor (uPAR)

The receptor of the urokinase-type plasminogen activator (uPAR) is one of the founding members of a superfamily of membrane-bound glycoproteins (Behrendt et al., 1991). The uPAR protein lacks an obvious transmembrane segment and is anchored into the membrane by the post-translational addition of a carboxy-terminal glycosylphosphatidylinositol (GPI) moiety (Ploug et al., 1991). The mature uPAR is composed of three repeat units of approximately 90 amino acids (Roldan et al., 1990). The consensus of these repeats is defined by a conserved pattern of 10 cysteine residues, which is shared by a diverse group of single-repeat glycoproteins (Behrendt et al., 1991). This group was termed the

Ly-6/uPAR superfamily, based on the detection of the same cysteine bonding pattern in certain mammalian membrane-bound leukocyte antigens, known as the Ly-6 family (LeClair et al., 1986; Kamiura et al., 1992).

A common ancestry between the diverse Ly-6/uPAR family members is also inferred by a shared carboxy-terminal GPI-anchoring signal sequence and, more importantly, by the similar genomic organization of the exons that encode the Ly-6/uPAR domains (Bothwell et al., 1988; Petranka et al., 1992). Furthermore, most members of the human Ly-6/uPAR superfamily, including RIG-E (Mao et al., 1996), the E48 antigen (Brakenhoff et al., 1995), and the prostate stem cell antigen (PSCA) (Reiter et al., 1998), map to chromosome 8q24 that is compatible with ancient local gene duplication events. An exception is the membrane inhibitor of reactive lysis, known as CD59, MAC1F, or protectin, located on chromosome 11p13 (Forsberg et al., 1989; Bickmore et al., 1993). The human 8q24 locus is syntenic to a region of mouse chromosome 15 to which most murine members of the Ly-6/uPAR family have been mapped (Hogarth et al., 1987; Gumley et al., 1995).

Chemical analysis of the cysteine-rich domain structure from several members of the Ly-6/uPAR superfamily has revealed a conserved pattern of disulfide bridging (Ploug et al., 1993; Fletcher

Reprint requests to: Michael Nehls, Lexicon Genetics Inc., 4000 Research Forest Drive, The Woodlands, Texas 77381; e-mail: mnehls@lexgen.com.

Abbreviations: ANUP, anti-neoplastic urinary protein; CZE, capillary zone electrophoresis; DTT, dithiothreitol; ESMS, electrospray mass spectrometry; GPI, glycosylphosphatidylinositol; HF, hemofiltrate; LC-ESMS, liquid chromatography electrospray mass spectrometry; MALDI-MS, matrix assisted laser desorption/ionization-mass spectrometry; MeCN, acetonitrile; PSCA, prostate stem cell antigen; RP-HPLC, reverse-phase high performance liquid chromatography; SLURP-1, secreted mammalian Ly-6/uPAR related protein 1; TFA, trifluoroacetic acid; uPAR, urokinase-type plasminogen activator receptor; UV, ultraviolet.

et al., 1994). This pattern is similar to the well-studied single-domain snake venom α -neurotoxins and cardiotoxins (Fleming et al., 1993; Ploug & Ellis, 1994). These secreted proteins are approximately the same size as the Ly-6/uPAR domains, and most possess at least 8 of the 10 cysteine residues. The cysteine residues are comparably spaced and, most importantly, identically paired (Ploug & Ellis, 1994). All Ly-6/uPAR family members including the snake and frog cytotoxins share the consensus sequence motif CCXXXXCN at the carboxy-terminal end. Another important link between the GPI-anchored receptors of the Ly-6/uPAR family and the secreted cytotoxins was established by the characterization of the chromosomal structure of the erabutoxin-c gene (Fuse et al., 1990). An intron central to the Ly-6/uPAR domain is found in both this snake gene and the mammalian genes (Ploug & Ellis, 1994). We regard the vertebrate Ly-6/uPAR domain containing proteins as members of two distinct subfamilies consisting of either the secreted cytotoxins or the membrane-bound glycoprotein receptors.

Here we report the isolation of a protein from human hemofiltrate and urine, which shares the carboxy-terminal consensus sequence, and the similarly spaced 10 cysteine residues with members of the Ly-6/uPAR family. Our chemical analysis is in agreement with the predicted cysteine bridging pattern, and defines the protein as a new Ly-6/uPAR family member. In contrast to other mammalian members of the Ly-6/uPAR family, a comparison of the protein's primary structure with the predicted sequence from chromosomal and cDNA sequences reveals the lack of a typical GPI-anchoring signal sequence. We named this new gene product secreted mammalian Ly-6/uPAR related protein 1, or SLURP-1, based on its primary structural characteristics. SLURP-1 provides an important evolutionary link between the GPI-anchored mammalian receptors and the secreted cytotoxins. This conclusion is supported by a phylogenetic analysis, which places the SLURP-1 protein closer to certain snake neurotoxins. We anticipate that this evolutionary relationship will help to elucidate the function of the SLURP-1 protein.

Results

Identification of a new secreted Ly-6/uPAR-related protein in blood and urine

The Ly-6/uPAR superfamily is divided into GPI-anchored cell surface receptors and nonmammalian secreted cytotoxins. Several lines of evidence suggest that both subfamilies co-evolved but no mammalian secreted member of this family has been described so far. We carried out an extensive GenBank and SwissProt database search with the carboxy-terminal Ly-6/uPAR consensus sequence XXXXXXXCCXXDLCNXXX, whereby D and L appeared to be the most common amino acids at the indicated position, and searched all matching sequences for those that did not contain a carboxy-terminal GPI-anchoring signal sequence. The ARS (component B)-81/s locus (accession number X99977 and P55000 in GenBank and SwissProt, respectively), encodes a putative protein with these characteristics. Based on assumed splicing and conceptual translation of the genomic sequences, we termed this putative protein SLURP-1 for secreted mammalian Ly-6/uPAR related protein 1, since it shares additional characteristics with the mammalian Ly-6/uPAR family besides the carboxy-terminal Ly-6/uPAR consensus sequence. These include the size of the prospective mature protein with a calculated molecular weight of 8,843 Da and the existence of 10 appropriately spaced cysteine residues. Using oli-

gonucleotide primers derived from the putative 5' and 3' untranslated transcribed sequences, we obtained PCR amplicons from reverse transcribed (RT) human RNA that was obtained from vagina and tonsil tissue. Sequences of the amplified cDNA products were in agreement with the predicted exonic structure of the genomic database entry, accession number X99977. However, unlike all other mammalian members of the Ly-6/uPAR family, the SLURP-1 transcript does not encode a GPI-anchoring signal sequence as predicted by the genomic sequence. This observation suggested that SLURP-1 is a secreted protein and could be isolated and characterized from human blood.

To identify SLURP-1, aliquots of all subfractions from our peptide library that was recently created from human blood ultrafiltrate (Forssmann et al., 1993; Schulz-Knappe et al., 1996b, 1997; Jürgens et al., 1997) were subjected to analysis that allowed the determination of protein size and the disulfide number. We performed a dithiothreitol (DTT) reduction of hemofiltrate subfractions, followed by S-carboxamidomethylation by treatment with iodoacetic amide. MALDI-MS analysis of these samples compared with nonderivatized samples of blood ultrafiltrate subfractions resulted in the detection of a protein in subfraction 13 of pH-pool 2, with characteristics of SLURP-1. Figure 1 shows the MALDI-MS analysis of this fraction before and after DTT/iodoacetic amide treatment. The original sample contained a component of 8,851 Da, which is within an acceptable experimental deviation to the calculated mass of 8,843 Da for the new Ly-6/

Fig. 1. MALDI-MS analysis of human blood ultrafiltrate subfraction 13, pH-pool 2. (A) Original sample and (B) after DTT reduction and subsequent carboxamidomethylation. In B, signals corresponding to peptides without DTT-reducible cysteine residues, which did not result in a shift of m/z , are marked with an asterisk; disulfide-bond containing components are marked with triangles.

uPAR protein family member, based on the translated cDNA sequence after subtraction of the putative amino-terminal signal peptide. A protein with this mass was not detected in the reduced and carboxamidomethylated sample, but instead a component of 9,434 Da was identified that was not present in the untreated sample. The difference of 583 Da is consistent with a 10-fold carboxamidomethylation of the corresponding protein, revealing that the 8,851 Da-protein contains the expected 10 cysteine residues. It was evident that each of the 10 cysteine residues is part of a disulfide bond since no alkylation was observed in the sample that was not treated with DTT prior to alkylation. Based on molecular weight and cysteine content, the newly identified protein corresponds closely to the expected characteristics of SLURP-1.

To exactly determine the molecular weight of the identified protein and to verify that it is indeed SLURP-1, a LC-ESMS analysis of blood filtrate subfraction 13/pH-pool 2 was carried out (Fig. 2). By analysis of extracted ions with m/z values from the LC-ESMS run according to fourfold, fivefold, and sixfold charged ions, the protein could be unambiguously identified. Calculations based on ions with charges +4 to +7 determined a molecular weight of 8,842 Da, corresponding accurately with the amino acid sequence of the new putative Ly-6/uPAR protein family member encoded by the ARS (component B)-81/s locus. This result also confirmed our initial results obtained by the MALDI-MS screening. Likewise, SLURP-1 was detected by LC-ESMS analysis in a chromatographic fraction obtained from human urine coeluting with an α 1-microglobulin fragment (Fig. 2).

Determination of the primary structure and disulfide bonds of SLURP-1

As a next step, larger quantities of the protein were isolated to verify the amino acid sequence and to elucidate the pattern of disulfide bonds of SLURP-1. Starting with 50 mg of lyophilized subfraction 13 from blood ultrafiltrate, pH-pool 2, which corresponds to approximately 2,000 L hemofiltrate, SLURP-1 was purified by applying two subsequent reverse-phase HPLC steps (Fig. 3). After a C4 reverse-phase HPLC, the SLURP-1-containing fractions were subjected to a final separation step by C18 HPLC, yielding approximately 2 mg of highly purified SLURP-1 protein, as confirmed by capillary zone electrophoresis and ESMS (Fig. 3). An equivalent purification of SLURP-1 was carried out using 270 mL of human urine. The preparative purification was achieved by subsequent cation exchange and RP-HPLC (Heine et al., 1997). Figure 3 shows that SLURP-1 isolated from urine was again contaminated with α 1-microglobulin, known from the initial analytical experiments.

SLURP-1 purified from human blood was cleaved using specific and nonspecific endoproteases with and without prior reduction of disulfide bonds, and the resulting peptide fragments were separated by RP-HPLC, and subjected to ESMS analysis and Edman degradation. Protease treatment by trypsin, chymotrypsin, Glu-C, and Arg-C revealed a high stability of SLURP-1 toward these enzymes. After 4–6 h, only one or two cleaved fragments were detected by ESMS. Since the disulfide-maintained structure of SLURP-1 may protect against proteolytic cleavage, the complete amino acid sequence was unambiguously assembled with peptide maps derived from both specific (after disulfide reduction), and nonspecific endoprotease cleavages (thermolysin, pepsin, and subtilisin, without reduction). The sequences of the complete protein and of selected peptide fragments are shown in Figure 4A.

The calculated molecular weight of 8,843 Da of the determined primary structure of isolated SLURP-1, including five disulfide bridges, confirms exactly the sequence of the translated mature protein sequence derived from the SLURP-1 cDNA amplicon and our previous analytical results.

As shown in Figure 4B, the cleavages of SLURP-1 by thermolysin, subtilisin, and pepsin produced several peptide fragments, which were used to determine the pattern of disulfide bridges that maintain the tertiary structure of the protein. ESMS analysis and Edman degradation of these SLURP-1 fragments allowed an unambiguous assignment of disulfide bonds Cys3–Cys28 and Cys6–Cys15. To determine the connectivity of Cys77, the subtilisin-derived fragment in which the segments SC, CCF, and RDLNCS were linked by two disulfide bonds (Fig. 4B) was subjected to one single step of manual Edman degradation and subsequently analyzed by nanospray ESMS. The obtained product had a molecular weight of 817 Da that is in exact accordance with a segment CF disulfide-linked to DLNCS. Thus, Cys77 is unambiguously linked to Cys72. In contrast, the connection of Cys21 remained unclear due to the internal segment SCSSC that made an unambiguous disulfide assignment difficult. However, the disulfide bond pattern of SLURP-1, as far as it could be determined by the applied techniques (Fig. 4C), is in agreement with the known pattern of all other members of the Ly-6/uPAR superfamily.

SLURP-1 is a secreted Ly-6/uPAR family member related to snake and frog cytotoxins

As outlined above, all previously known members of the Ly-6/uPAR family except the secreted cytotoxins from snake venom and frog skin are GPI-anchored cell membrane glycoprotein receptors. Since the SLURP-1 gene does not apparently encode a GPI-anchoring signal sequence and hence could be purified as a secreted protein from human blood and urine (Fig. 5A), we tried to determine the relationship of SLURP-1 to the two biologically distinct Ly-6/uPAR subfamilies, the receptor proteins and the secreted cytotoxins. We carried out a phylogenetic analysis based on sequences common to both subfamilies, thereby excluding disulfide bond B since most secreted toxins lack both this and the GPI-anchoring signal sequences that are typical for the receptor subfamily. Figure 5A shows the result after applying the Clustal algorithm for phylogenetic alignment with a PAM250 residue weight table. In this figure we also show the GPI-anchoring signal sequences of the protein precursors that were excluded during the Clustal alignment. As shown in Figure 5B, the Clustal phylogenetic tree analysis separated the secreted and GPI-anchored proteins into four subgroups, two for each Ly-6/uPAR subfamily. Surprisingly, SLURP-1 is phylogenetically more closely related to two snake cytotoxins than it is to the mammalian GPI-anchored receptors. One of the two groups of the GPI-anchored proteins contains only the human CD59 protein.

We mapped SLURP-1 using the PCR-based G3 human/hamster radiation hybrid panel from Research Genetics (Huntsville, Alabama). With oligonucleotide primers derived from the genomic sequences encoding the amino- and carboxy-terminal end, human SLURP-1 could be amplified specifically from hybrid genomic DNA (data not shown). Based on the specific retention of human chromosomal DNA containing SLURP-1, we could map the gene to chromosome 8 at 33.23 cR. The most closely linked marker with a LOD score of 7.22 was WI-7711, which maps close to the locus encoding the proline-rich tyrosine kinase 2 gene at 8q24–8qter.

LC-ESMS detection of SLURP-1

Fig. 2. Detection of SLURP-1 in blood ultrafiltrate and urine. Left, urine: (A) LC-ESMS analysis of subfraction 7, (B) total ion counts (TIC) for fourfold, fivefold, and sixfold charged ions of the indicated scans of A, and (C) ESMS of the indicated scans of A. Asterisks mark an α 1-microglobulin fragment, which co-eluted with SLURP-1. Right, hemofiltrate: (A) LC-ESMS analysis of subfraction 13, pH pool 2, (B) total ion counts (TIC) for fourfold, fivefold, and sixfold charged ions of the indicated scans of A, and (C) ESMS of the indicated scans of A.

Discussion

Use of blood peptide library for the database-guided protein identification

We report the identification and characterization of the secreted SLURP-1 protein, a non-GPI-anchored member of the Ly-6/uPAR superfamily from human blood ultrafiltrate and urine peptide li-

braries. The isolation process was guided by the information from conceptionally translated human genomic and cDNA sequences. The existence of SLURP-1 was suspected after the identification of a Ly-6/uPAR carboxy-terminal consensus sequence putatively encoded by the human ARS (component B)-81/s locus.

In the past few years, several new circulating bioactive compounds from our large-scale peptide library from human blood ultrafiltrate and urine have been isolated. New peptide hormones

Purification of SLURP-I

Fig. 3. Isolation and purity analysis of human SLURP-1 from blood ultrafiltrate and urine. Left, urine: (A) C4 and C18 RP-HPLC of subfraction 7, (B) capillary zone electrophoresis of purified SLURP-1, and (C) ESMS analysis of purified SLURP-1. Asterisks mark an α 1-microglobulin fragment, which co-eluted with SLURP-1. Right, hemofiltrate: (A) C4 and C18 RP-HPLC of subfraction 13/pH-pool 2, (B) CZE of purified SLURP-1, and (C) ESMS of purified SLURP-1.

such as guanylin (Kuhn *et al.*, 1993) and uroguanylin (Hess *et al.*, 1995) were found by screening these peptide libraries using specific bioassays. A more systematic approach (Jürgens *et al.*, 1997; Schulz-Knappe *et al.*, 1997) resulted in the discovery of new chemokines (Schulz-Knappe *et al.*, 1996a) and the first human β -defensin (Bensch *et al.*, 1995). The isolation of SLURP-1 presented in this paper documents a very practical approach for isolating predicted polypeptides based on the corresponding cDNA and/or genomic sequence from our human peptide libraries. This might become a very useful application since a large number of new nucleic acid sequences encoding novel peptides and proteins, which are functionally interesting, are constantly generated by the Human Ge-

nome Project, in particular via the sequence analysis of cDNA library-derived expressed sequence tags (EST). For example, several novel chemokines have been identified using EST databases (Wells & Peitsch, 1997). However, certain genes might not be represented in EST databases and their existence can only be assumed by the analysis of genomic sequences. The reported simplicity of the identification of SLURP-1 using standard MALDI-MS and LC-ESMS techniques is an additional advantage of this new application for the discovery of secreted peptides and proteins.

As the human hemofiltrate was obtained from patients suffering chronic renal disease, SLURP-1 concentrations in urine from healthy people were quantified in an independent study. Since similar con-

Fig. 4. Sequence assignment and determination of disulfide bonds of purified SLURP-1. **A:** Complete primary structure by selected overlapping peptide fragments obtained from proteolytic cleavage. The amino acid sequence of SLURP-1 can be accessed at the Protein Information Resource, accession number A58842. **B:** Disulfide-linked fragments from unspecific proteolysis with pepsin, thermolysin, and subtilisin. For fragment 3 obtained by thermolysin cleavage, it could not be differentiated between the possible segments SCSS, CSSS, and SSSC. **C:** Scheme of the identified disulfide bond pattern.

centrations of at least 10 nM in both sources were found, we do not consider SLURP-1 as a disease-related protein. To determine exact physiological concentrations of SLURP-1 in blood and other body liquids, an ELISA utilizing a specific anti-SLURP-1 antibody is necessary and is currently being developed.

SLURP-1 is a secreted member of the Ly-6/uPAR superfamily

As previously outlined, SLURP-1 is, to the best of our knowledge, the first described mammalian member of the Ly-6/uPAR protein family that is secreted, i.e., does not encode a typical carboxy-terminal GPI-anchoring signaling sequence common to all other mammalian members currently described. Interestingly, we could map the human gene to chromosome 8q24-8qter, a region that contains most GPI-anchored receptors of the Ly-6/uPAR superfamily with the exception of CD59. The co-localization with the majority of the membrane-bound proteins suggests a common evolutionary origin of the two biologically diverse subgroups. Phylogenetic studies can commence to investigate the evolutionary relationship between receptor-type membrane receptors and hormone-like secreted proteins. As a first attempt toward this goal, we analyzed the relationship of SLURP-1 with members of both biologically distinct subfamilies, i.e., the secreted cytotoxins found in snake venom and in skin of the African clawed frog and GPI-anchored receptors. Our phylogenetic analysis places SLURP-1 into one of the two subgroups of the subfamily of secreted cytotoxins (see Fig. 5B). This indicates that SLURP-1, albeit chromosomally linked to the GPI-anchored receptors, has an amino acid composition that relates more closely to the cytotoxins. Since it is a secreted molecule itself, these data suggest a co-evolution of both subfamilies (see below). The GPI-anchored subfamily was also divided by this analysis into two subgroups, one with CD59 as a single member. This was not unexpected since CD59 was already shown to be located at a distinct chromosomal locus, suggesting an

early separation from the other members of this subfamily during evolution.

Which came first, the secreted or the GPI-anchored versions of this group of cysteine-rich proteins? This question cannot be answered with currently available data. Nevertheless, it is a speculative observation that a deletion of a single base pair in the stop codon of SLURP-1 elongates the reading frame by another 23 amino acids, with 26% leucine residues having some resemblance to a functional GPI-anchoring signal sequence (Yan et al., 1998). The identification of orthologs from other species, as well as of paralogous genes, might support our current hypothesis that a secreted molecule has been generated by gene duplication and mutagenesis of the sequence encoding the GPI-anchoring signal peptide. Recently, proteolysis of the carboxy-terminal GPI-anchoring signal sequence, independent of GPI modification, was reported to generate secreted forms of receptors (Wang et al., 1997). This alternative may result in a more dynamic way of utilizing the gene product than the proposed model of germline mutations of the GPI-anchoring signal, since it allows the cell to express the same protein both in a secreted and in a membrane-bound form.

Structure comparison between SLURP-1 and related proteins

As a member of the Ly-6/uPAR family, SLURP-1 is expected to grossly fold similarly to the snake neurotoxins. Such a structural similarity constrained by the cysteine bridges was already shown when one of the uPAR subdomains was predicted by computer simulation to have a three-dimensional structure similar to an α -neurotoxin (Ploug & Ellis, 1994). We hope that resolving the three-dimensional structure of SLURP-1 will aid in the understanding of the protein's function.

An exhaustive database search using the detected primary protein structure of SLURP-1 revealed no significant sequence homology to other proteins except for a partial homology with the

Fig. 5. Sequence comparison and phylogenesis of SLURP-1. (A) Sequence homology between SLURP-1 and related proteins of the Ly-6/uPAR protein superfamily and (B) phylogenetic tree of the Ly-6/uPAR protein superfamily. The SwissProt database accession numbers of the analyzed sequences are Q09022 for frog Xenoxin-1, P80494 for black mamba muscarinic toxin alpha, 229625 for cobra toxin CM2a, and Y12267 and P14534, for krait neurotoxin and neurotoxin-2, respectively. The GPI-anchored proteins have the accession numbers L34554 (GenBank) for stem cell antigen 2 (Sca-2) from chicken, P13987 (SwissProt) for human CD59, Q14210 (SwissProt) for human E48 antigen, AF043498 (GenBank) for human prostate stem cell antigen, U42376 (GenBank) for human RIG-E precursor, P09568 (SwissProt) for murine lymphocyte antigen Ly-6c.2/Ly-6c.1 precursor, and P35459 (SwissProt) for murine thymocyte B cell antigen precursor (MmTHB).

anti-neoplastic urinary protein (ANUP). This protein was originally described to have an anti-tumor activity on human tumor cell lines, but not on benign human diploid cells (Sloane et al., 1986; Pottathil et al., 1990; Ridge & Sloane, 1996). There are several lines of evidence that show that ANUP and SLURP-1 are not the same protein. For ANUP, only the amino-terminal segment starting with a pyroglutamic residue followed by LKCYTCKEPMTSAA has yet been identified (Ridge & Sloane, 1996). SLURP-1 has no glutamine as the biochemical precursor of pyroglutamate but alanine at position -1. Another difference between SLURP-1 and ANUP is a serine instead of alanine at position 14. Most importantly, ANUP was recently characterized as a 16 kDa monomeric protein with three cysteine residues while SLURP-1 is a 9 kDa protein containing 10 cysteines (Sloane & Davis, 1996). Hence, there is no obvious structural relation between SLURP-1 and ANUP. A recent SwissProt database annotation (accession number P55000) that linked these two proteins is probably incorrect.

In conclusion, we provide evidence of a powerful new application for isolating novel secreted molecules from blood and urine peptide libraries. As a practical example, we described the identification of human SLURP-1, a new protein that appears to be a phylogenetic link between two biologically distinct subfamilies of the Ly-6/uPAR superfamily. We are currently isolating and characterizing the gene encoding the murine ortholog of human SLURP-1 that will allow us to study in more detail the function of the SLURP-1 protein in an experimentally more accessible system. Ablation of gene function by targeted recombination in embryonic stem cells, as well as a more detailed comparative expression analysis, will be carried out. Both approaches will provide information regarding the source and physiological target tissues of circulating SLURP-1.

Materials and methods

Purification of SLURP-1 from human blood ultrafiltrate

Human hemofiltrate (HF) was obtained from patients with chronic renal failure (provided by the Nephrologisches Zentrum Niedersachsen, Hannoversch-Münden, Germany). The sterile filtrate was acidified to pH 3 at 4 °C and the peptides were extracted by batch elution from a cation exchanger (Fractogel TSK 650(M), Merck, Darmstadt, Germany) with 0.5 M NH₄Ac. Five combined batches equivalent to about 4,000 L hemofiltrate were loaded and eluted stepwise from the cation exchanger using buffers with increasing pH (3.6, 4.5, 5.0, 5.6, 6.6, 7.4, and 9.0; pH-pool numbers 1–7). Each of the seven pH-pools was then subfractionated by reverse-phase separation using a Source RPC 15 column (Pharmacia, Freiburg, Germany, 125 × 10 mm, 15 µm). Elution was performed with a linear gradient of 100% solvent A (0.01 M HCl) to 60% solvent B (MeCN/0.01 M HCl 4:1). From each of the seven pH-pools, 2 mL fractions were collected, subsequently lyophilized and stored at -20 °C. The total of 300 subfractions represents our two-dimensional peptide library from human blood ultrafiltrate (based on pK and hydrophobicity). A detailed description of the experimental procedures has been published (Jürgens et al., 1997; Schulz-Knappe et al., 1997).

For the preparative scale isolation of SLURP-1, 50 mg of fraction 13 of pool 2 (eluting at pH 4.5) derived from processing of 4,000 L of hemofiltrate were separated by RP-HPLC (Parcosil C4, Biotek, Östringen, Germany, 125 × 10 mm, 5 µm, 300 Å; solvent A, 0.1% TFA; solvent B, 0.1% TFA in MeCN/water 4:1; gradient,

5–50% B in 50 min; flow rate, 1.8 mL/min; UV detection at 214 nm). The fraction containing SLURP-1 as detected by MALDI-MS or ESMS was subjected to a final RP-HPLC purification step (Vydac C18, The Separations Group, Hesperia, California, 250 × 10 mm, 300 Å, 5 µm; solvent A, 0.1% TFA; solvent B, 0.1% TFA in MeCN/water 4:1; gradient, 25–50% B in 50 min; flow rate, 1.8 mL/min; UV detection at 214 nm). The purified protein was subsequently characterized by capillary zone electrophoresis (P/ACE 2000, Beckman, München, Germany; buffer, 0.1 M NaH₂PO₄ containing 0.2 mg/mL hydroxypropylmethylcellulose, pH 2.5; uncoated capillary, 50 cm × 50 µm, pressure injection; constant current, 50 µA; UV detection at 200 nm, 25 °C), and by ESMS, sequencing and endoproteinase treatment. For the microanalytical identification of SLURP-1, fraction 13/pH-pool 2 was subjected directly to LC-ESMS.

Purification of SLURP-1 from urine

Human urine was collected from healthy volunteers and treated by tangential flow filtration and C₁₈ solid-phase extraction (Heine et al., 1997). Equivalents of 800 mL were subjected to a cation exchange HPLC (PolySulfoethyl A, PolyLC, Columbia, Maryland, 100 × 4.6 mm, 5 µm, 300 Å; solvent A, 20 mM HOAc, 10% MeCN, pH 3.2; solvent B, 20 mM HOAc, 10% MeCN, 1 M NH₄Ac, pH 6.6; gradient, 10–80% B in 40 min; flow rate, 0.4 mL/min; UV detection at 280 nm). Fractions were collected every 2 min, lyophilized, stored at -20 °C, and represent our human urine peptide library. Detection of SLURP-1 was carried out by direct LC-ESMS analysis of these fractions. For isolation of SLURP-1, the fraction containing the protein (fraction 7, a volume equivalent to 270 mL urine was applied per run) was further purified by RP-HPLC (ReproSil-Pur, Dr. A. Maisch, Ammerbuch, 250 × 2 mm, 120 Å, 5 µm; solvent A, 0.1% TFA; solvent B, 0.07% TFA in MeCN/water 4:1; gradient, 5–100% B in 50 min; flow rate, 0.2 mL/min; UV detection at 220 nm). The fractions obtained were characterized as described above using capillary zone electrophoresis, ESMS, and sequencing.

Proteolytic cleavage of SLURP-1

For the determination of primary structure and disulfide bonds, samples of purified SLURP-1 isolated from blood ultrafiltrate were treated with specific (trypsin, chymotrypsin, Arg-C, and Glu-C) and nonspecific endoproteinases (subtilisin, thermolysin, and pepsin). Digests were carried out for 10 h using standard procedures as recommended by the manufacturer. Proteolysis with specific endoproteinases was performed after reduction of disulfide bonds by DTT, while nonspecific proteinases were used without disulfide reduction to detect disulfide-linked fragments. All derived peptide mixtures were separated using RP-HPLC (Vydac C18, 250 × 4.6 mm, 300 Å, 5 µm; solvent A, 0.1% TFA; solvent B, 0.1% TFA in MeCN/water 4:1; gradient, 5–70% B in 65 min; flow rate, 0.7 mL/min; UV detection at 214 and 280 nm). Peaks of the peptide maps were collected and subjected to a further characterization by Edman degradation, ESMS, or nanospray collision induced dissociation.

Carboxamidomethylation of blood ultrafiltrate subfractions

About 10 µg of lyophilized hemofiltrate subfractions including fraction 13, pH-pool 2, were dissolved in 10 µL 0.1% TFA and

90 μL 0.2 M N-methylmorpholine (pH 8.3). After addition of 5 μL 0.4 M DTT under nitrogen atmosphere, the mixture was kept at 40°C for 60 min. Subsequently, 20 μL 0.5 M iodoacetic amide were added and reacted for 30 min. The carboxamidomethylation was stopped by addition of 2 μL 50% HOAc. The samples were then desalted by filtration over POROS 50 R2 (PerSeptive Biosystems, Wiesbaden, Germany) and eluted with 80 μL 0.1% TFA in MeCN/water 4:1, diluted with 30 μL 0.1% TFA, and directly used for MALDI-MS.

Mass spectrometry

Matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS) was carried out in a linear mode using a LaserTec RBT (PerSeptive/Vestec, Houston, Texas) equipped with a 337 nm laser. Positive ions were accelerated at 30 kV. Samples of 1 μL size were applied on polished stainless steel plates using the dried-drop crystallization technique (Karas & Hillenkamp, 1988) with α -cyano-4-hydroxycinnamic acid as matrix (5 mg/mL 0.1% TFA in 50% MeCN containing 5 mg/mL L-fucose as co-matrix). Electrospray mass spectrometry (ESMS) was performed on a triple-stage quadrupole mass spectrometer (Sciex API III, Perkin Elmer, Langen, Germany). Samples were dissolved in MeCN/water 1:1 containing 0.2% formic acid and infused into the ion spray using a syringe pump at a flow rate of 5 $\mu\text{L}/\text{min}$ with a positive potential of 5 kV at the ion spray interface. Data analysis was performed using MacSpec 3.22 (Perkin Elmer/ABI, Weiterstadt, Germany). Coupling of liquid chromatography to electrospray mass spectrometry (LC-ESMS) was carried out using a model 140B dual syringe solvent delivery system and a model 785 UV detector (Perkin Elmer) equipped with a capillary z-cell (LC-Packings, Amsterdam, The Netherlands) on a ODS-AQ C18 column (YMC, Schermbeck, 100 $\times$ 1 mm, 3 μm , 120 Å; solvent A, 0.06% TFA; solvent B, 0.05% TFA in MeCN/water 4:1; gradient, 0–100% B in 60 min, flow rate, 20 $\mu\text{L}/\text{min}$; UV detection at 220 nm). The eluent was infused into the ion spray without splitting. Collision induced fragmentation (CID) was carried out utilizing the nanospray technique (Wilm & Mann, 1996). Samples were infused as described above. The first quadrupole was set to the m/z value of the single-charged peptide as determined before. Argon was used as collision gas at a collision energy of 20–45 eV. The second mass spectrometer was scanned from 10 m/z to the m/z value determined for the single-charged ion at a scan rate of 3 s. At least 100 spectra were sampled. Experiments were carried out with SLURP-1 segments containing two disulfides obtained from subtilisin and thermolysin cleavage.

Sequence analysis

Sequence analysis of SLURP-1 and its proteolytically generated fragments was carried out by automated Edman degradation on a 494 Procise Protein Sequencer (Perkin Elmer) according to standard procedures recommended by the manufacturer. The 50–1,000 pmol of peptide fragments obtained from HPLC separation after endoproteinase treatment and the reduced SLURP-1 were applied on PolyBrene membranes.

Molecular phylogeny studies of SLURP-1

Alignments and phylogenetic tree analysis of SLURP-1 amino acid sequences were conducted using the Clustal method algorithm

with a PAM250 residue weight table provided by the MegAlign program from the DNASTAR Inc. software package.

Acknowledgments

We thank Peter Schulz-Knappe for help in proteolytic analysis, Manfred Raida for performing nanospray MS, and Susanne Neitz and Ulrike Schrameck for excellent technical assistance. Further thanks to Glenn Friedrich, Henry Skinner, and Brian Zambrowicz for reviewing this manuscript prior to submission.

References

- Behrendt N, Ploug M, Pathy L, Houen G, Blasi F, Dano K. 1991. The ligand-binding domain of the cell surface receptor for urokinase-type plasminogen activator. *J Biol Chem* 266:7842–7847.
- Bensch KW, Raida M, Mägert HJ, Schulz-Knappe P, Forssmann WG. 1995. The hBD-1: A novel beta-defensin from human plasma. *FEBS Lett* 368:331–335.
- Bickmore WA, Longbottom D, Oghene K, Fletcher JM, van Heyningen V. 1993. Colocalization of the human CD59 gene to 11p13 with the MIC11 cell surface antigen. *Genomics* 17:129–135.
- Bothwell A, Pace PE, LeClair KP. 1988. Isolation and expression of an IFN-responsive Ly-6C chromosomal gene. *J Immunol* 140:2815–2820.
- Brakenhoff RH, Gerretsen M, Knippels EM, van Dijk M, van Essen H, Weghuis DO, Sinke RJ, Snow GB, van Dongen GA. 1995. The human E48 antigen, highly homologous to the murine Ly-6 antigen ThB, is a GPI-anchored molecule apparently involved in keratinocyte cell-cell adhesion. *J Cell Biol* 129:1677–1689.
- Fleming TJ, O'Huigin C, Malek TR. 1993. Characterization of two novel Ly-6 genes. Protein sequence and potential structural similarity to alpha-bungarotoxin and other neurotoxins. *J Immunol* 150:5379–5390.
- Fletcher CM, Harrison RA, Lachmann PJ, Neuhaus D. 1994. Structure of a soluble, glycosylated form of the human complement regulatory protein CD59. *Structure* 2:185–199.
- Forsberg UH, Bazil V, Stefanova I, Schroder J. 1989. Gene for human CD59 (likely Ly-6 homologue) is located on the short arm of chromosome 11. *Immunogenetics* 30:188–193.
- Forssmann WG, Schulz-Knappe P, Meyer M, Adermann K, Forssmann K, Hock D, Aoki A. 1993. Characterization of natural post-translationally processed peptides from human blood: A new tool in the systemic investigation of native peptides. In: Yanaihara N, ed. *Peptide chemistry 1992*. Leiden ES-COM. pp 553–557.
- Fuse N, Tsuchiya T, Nonomura Y, Menez A, Tamiya T. 1990. Structure of the snake short-chain neurotoxin, erabutoxin c, precursor gene. *Eur J Biochem* 193:629–633.
- Gumley TP, McKenzie IF, Sandrin MS. 1995. Tissue expression, structure and function of the murine Ly-6 family of molecules. *Immunol Cell Biol* 73:277–296.
- Heine G, Raida M, Forssmann WG. 1997. Mapping of peptides and protein fragments in human urine using liquid chromatography-mass spectrometry. *J Chromatogr A* 776:117–124.
- Hess R, Kuhn M, Schulz-Knappe P, Raida M, Fuchs M, Klodt J, Adermann K, Kaever V, Cetin Y, Forssmann WG. 1995. GCAP-II: Isolation and characterization of the circulating form of human uroguanylin. *FEBS Lett* 374:34–38.
- Hogarth PM, McKenzie IF, Sutton VR, Curnow KM, Lee BK, Eicher EM. 1987. Mapping of the mouse Ly-6, Xp-14, and Gdc-1 loci to chromosome 15. *Immunogenetics* 25:21–27.
- Jürgens M, Schrader M, Raida M, Forssmann WG, Schulz-Knappe P. 1997. Multi-dimensional mapping of human blood peptides by mass spectrometry. *J Biomol Techniques* (<http://www.abrf.org/JBT/Articles/JBT0006/JBT0006.html>).
- Kamiura S, Nolan CM, Meruelo D. 1992. Long range physical map of the Ly-6 complex: Mapping the Ly-6 multigene family by field inversion and two-dimensional gel electrophoresis. *Genomics* 12:89–105.
- Karas M, Hillenkamp F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299–2301.
- Kuhn M, Raida M, Adermann K, Schulz-Knappe P, Gerzer R, Heim JM, Forssmann WG. 1993. The circulating bioactive form of human guanylin is a high molecular weight peptide (10.3 kDa). *FEBS Lett* 318:205–209.
- LeClair KP, Palfree RG, Flood PM, Hammerling U, Bothwell A. 1986. Isolation of a murine Ly-6 cDNA reveals a new multigene family. *Embo J* 5:3227–3234.
- Mao M, Yu M, Tong JH, Ye J, Zhu J, Huang QH, Fu G, Yu L, Zhao SY, Waxman S. 1996. RIG-E, a human homolog of the murine Ly-6 family, is induced by

- retinoic acid during the differentiation of acute promyelocytic leukemia cell. *Proc Natl Acad Sci USA* 93:5910–5914.
- Petranka JG, Fleenor DE, Sykes K, Kaufman RE, Rosse WF. 1992. Structure of the CD59 encoding gene: Further evidence of a relationship to murine lymphocyte antigen Ly-6 protein. *Proc Natl Acad Sci USA* 89:7876–7879.
- Ploug M, Ellis V. 1994. Structure-function relationships in the receptor for urokinase-type plasminogen activator. Comparison to other members of the Ly-6 family and snake venom alpha-neurotoxins. *FEBS Lett* 349:163–168.
- Ploug M, Kjalke M, Ronne E, Weidle U, Hoyer-Hansen G, Dano K. 1993. Localization of the disulfide bonds in the NH2-terminal domain of the cellular receptor for human urokinase-type plasminogen activator. A domain structure belonging to a novel superfamily of glycolipid-anchored membrane proteins. *J Biol Chem* 268:17539–17546.
- Ploug M, Ronne E, Behrendt N, Jensen AL, Blasi F, Dano K. 1991. Cellular receptor for urokinase plasminogen activator. Carboxyl-terminal processing and membrane anchoring by glycosyl-phosphatidylinositol. *J Biol Chem* 266:1926–1933.
- Pottathil R, Aggarwal S, Davis LH, Bucovac ET, Sloane NH. 1990. Evidence that the anti-neoplastic urinary protein (ANUP) is present in human granulocytes. *Cancer Therapy and Control* 1:193–198.
- Reiter RE, Gu Z, Watabe T, Thomas G, Szigeti K, Davis E, Wahl M, Nisitani S, Yamashiro J, Le Beau MM. 1998. Prostate stem cell antigen: A cell surface marker overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 95:1735–1740.
- Ridge RJ, Sloane NH. 1996. Partial N-terminal amino acid sequence of the anti-neoplastic urinary protein (ANUP) and the anti-tumor effect of the N-terminal nonapeptide of the unique cytokine present in human granulocytes. *Cytokine* 8:1–5.
- Roldan AL, Cubellis MV, Masucci MT, Behrendt N, Lund LR, Dano K, Appella E, Blasi F. 1990. Cloning and expression of the receptor for human urokinase plasminogen activator, a central molecule in cell surface, plasmin dependent proteolysis. *EMBO J* 9:467–474.
- Schulz-Knappe P, Mägert HJ, Dewald B, Meyer M, Cetin Y, Kubbies M, Tomczkowski J, Kirchhoff K, Raida M, Adermann K. 1996a. HCC-1, a novel chemokine from human plasma. *J Exp Med* 183:295–299.
- Schulz-Knappe P, Raida M, Meyer M, Quethorst EA, Forssmann WG. 1996b. Systematic isolation of circulating human peptides: The concept of peptide trapping. *Eur J Med Res* 1:223–236.
- Schulz-Knappe P, Schrader M, Ständker L, Richter R, Hess R, Jürgens M, Forssmann WG. 1997. Peptide bank generated by large-scale preparation of circulating human peptides. *J Chromatogr A* 776:125–132.
- Sloane NH, Davis LH. 1996. In vitro and in vivo anti-tumor activity of a plasma cytokine identical to anti-neoplastic urinary protein: Tumor cell killer protein (TCKP). *Tumor Targ* 2:322–326.
- Sloane NH, Lynn WR, Macleod RM, Hade EP, Pottathil R, Kyriazis AP. 1986. Studies on an anti-neoplastic fraction from human urine. Characterization of the major protein in this fraction. *Biochem J* 234:355–362.
- Wang J, Shen F, Yan W, Wu M, Ratnam M. 1997. Proteolysis of the carboxyl-terminal GPI signal independent of GPI modification as a mechanism for selective protein secretion. *Biochemistry* 36:14583–14592.
- Wells TN, Peitsch MC. 1997. The chemokine information source: Identification and characterization of novel chemokines using the world wide web and expressed sequence tag databases. *J Leukoc Biol* 61:545–550.
- Wilm M, Mann M. 1996. Analytical properties of the nanoelectrospray ion source. *Anal Chem* 68:1–8.
- Yan W, Shen F, Dillon B, Ratnam M. 1998. The hydrophobic domains in the carboxyl-terminal signal for GPI modification, and in the amino-terminal leader peptide, have similar structural requirements. *J Mol Biol* 275:25–33.