Biol. Chem., Vol. 388, pp. 1215–1225, November 2007 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2007.124

Defining the extended substrate specificity of kallikrein 1-related peptidases

Carla A. Borgoño[1], Julie-Ann Gavigan[2], Juliano Alves[3], Ben Bowles[1], Jennifer L. Harris[2,3], Georgia Sotiropoulou[4] and Eleftherios P. Diamandis[1,*]

[1] Department of Laboratory Medicine and Pathobiology, University of Toronto and Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto M5G 1X5, ON, Canada
[2] Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121, USA
[3] Scripps Research Institute, Department of Molecular Biology, La Jolla, CA 92037, USA
[4] Department of Pharmacy, School of Health Sciences, University of Patras, Rion, GR-26500 Patras, Greece

* Corresponding author
e-mail: ediamandis@mtsinai.on.ca

Abstract

Human kallikrein 1-related peptidases (KLKs) form a subfamily of 15 extracellular (chymo)tryptic-like serine proteases. KLKs 4, 5, 13 and 14 display altered expression/activity in diverse pathological conditions, including cancer. However, their distinct (patho)physiological roles remain largely uncharacterized. As a step toward distinguishing their proteolytic functions, we attempt to define their primary and extended substrate specificities and identify candidate biological targets. Heterologously expressed KLKs 4, 5, 13 and 14 were screened against fluorogenic 7-amino-4-carbamoylmethylcoumarin positional scanning-synthetic combinatorial libraries with amino acid diversity at the P1–P4 positions. Our results indicate that these KLKs share a P1 preference for Arg. However, each KLK exhibited distinct P2–P4 specificities, attributable to structural variations in their surface loops. The preferred P4–P1 substrate recognition motifs based on optimal subsite occupancy were as follows: VIQSAV-QL-R for KLK4; YFWGPV-RK-NSFAM-R for KLK5; VY-R-LFM-R for KLK13; and YW-KRSAM-HNSPA-R for KLK14. Protein database queries using these motifs yielded many extracellular targets, some of which represent plausible KLK substrates. For instance, cathelicidin, urokinase-type plasminogen activator, laminin and transmembrane protease serine 3 were retrieved as novel putative substrates for KLK4, 5, 13 and 14, respectively. Our findings may facilitate studies on the role of KLKs in (patho)physiology and can be used in the development of selective KLK inhibitors.

Keywords: combinatorial libraries; enzyme kinetics; protease substrates; proteolytic enzyme; specificity profiling.

Introduction

Proteolytic enzymes, encoded by 2% of all human genes, play pivotal roles in the control of most physiological processes, ranging from cell-cycle regulation and tissue remodeling to inflammation (Barrett et al., 2004). Alterations in protease expression, structure and/or activity drive and sustain a number of common pathological conditions, including cancer, cardiovascular disease and arthritis (Barrett et al., 2004). Consequently, the therapeutic targeting of many proteases is under intense investigation (Southan, 2001). Serine proteases (SPs) represent 30% of all proteases within the human degradome, the second largest mechanistic class after metalloproteases (Puente et al., 2003). According to the classification system of Rawlings et al. (2006), SPs are categorized into 12 clans, each comprising numerous families based on tertiary structure and amino acid sequence similarity. The S1 family within clan PA(S) is the most densely populated and includes the human kallikrein 1-related peptidases (KLKs) (Lundwall et al., 2006).

KLKs (EC 3.4.21) constitute a subgroup of 15 extracellular (chymo)tryptic-like SPs encoded by the largest contiguous cluster of protease genes in the genome on chromosome 19q13.4 (Yousef et al., 2000b; Puente et al., 2003; Borgono et al., 2004; Lundwall et al., 2006). Until a decade ago, the KLK family was thought to comprise the three 'classical' KLKs (i.e., KLK1, tissue/pancreatic/renal/urinary kallikrein; KLK2, human glandular kallikrein 1/kallikrein 1-related peptidase 2; and KLK3, prostate-specific antigen/PSA/kallikrein 1-related peptiase 3) and has only recently been expanded to 15 members. Although the roles of KLK1 in kinin generation (Bhoola et al., 1992) and of KLK2 and KLK3 in fertilization via seminal clot liquefaction (Lilja, 1985; Deperthes et al., 1996) are well established, the biological significance of most newly identified KLKs remains to be elucidated.

KLK4 (prostase/EMSP1/PRSS17/ ARM1/KLK-L1/kallikrein 1-related peptidase 4) (Nelson et al., 1999; Stephenson et al., 1999; Yousef et al., 1999; Hu et al., 2000), KLK5 (human stratum corneum tryptic enzyme/KLK-L2/kallikrein 1-related peptidase 5) (Brattsand and Egelrud, 1999; Yousef and Diamandis, 1999), KLK13 (KLK-L4/kallikrein 1-related peptidase 13) (Yousef et al., 2000a) and KLK14 (KLK-L6/kallikrein 1-related peptidase 14) (Hooper et al., 2001; Yousef et al., 2001) are members of the expanded KLK family and share 40–50% amino acid similarity (Yousef and Diamandis, 2001). Similar to other KLKs, they are predominantly expressed, and often co-expressed, in the glandular epithelia of a wide range of organs, including the prostate and breast, and are presumed to act in cascade pathways with other KLKs and proteases (Borgono et al., 2004). Inferred from their expression profiles and/or the known functions of rodent orthologs, these KLKs likely participate in several normal

physiological processes, including the maturation of dental enamel (KLK4) (Hu et al., 2000), skin desquamation (KLK5, KLK14) (Caubet et al., 2004; Stefansson et al., 2006) and seminal plasma liquefaction (KLK4, KLK5) (Takayama et al., 2001; Michael et al., 2006).

Dysregulated KLK4, 5, 13 and 14 expression and/or activity is associated with multiple diseases, including cancer (e.g., prostate, breast, ovarian, testicular, lung; KLK4, 5, 13, 14) (Borgono and Diamandis, 2004), skin conditions (e.g., psoriasis, atopic dermatitis, Netherton and peeling-skin syndromes; KLK5, 13, 14) (Komatsu et al., 2005; Descargues et al., 2006; Hachem et al., 2006; Komatsu et al., 2006) and dental defects (e.g., amelogenesis imperfecta; KLK4) (Hart et al., 2004; Wright et al., 2006). Functional studies also indicate that these KLKs can digest substrates directly or indirectly implicated in the development or progression of the above disorders *in vitro*, including extracellular matrix (ECM) molecules (Bernett et al., 2002; Magklara et al., 2003; Kapadia et al., 2004a; Ghosh et al., 2004; Felber et al., 2005; Obiezu et al., 2006), urokinase plasminogen activator (uPA) (Takayama et al., 2001), (corneo)desmosomal adhesion molecules (Caubet et al., 2004) and proteinase-activated receptors (PARs) (Oikonomopoulou et al., 2006). In effect, these KLKs may represent potentially useful biomarkers and drug targets.

Therefore, delineation of biological pathways involving KLK4, 5, 13 and 14 action is of particular interest given the preliminary findings implicating these KLKs in disease pathology. Although these KLKs may possess common functions based on their overlapping tissue expression, cellular localization and *in vitro* substrate repertoire, the association of particular KLKs with specific (patho)physiological processes implies that they likely exert unique proteolytic activities as well. Their distinctive actions may be reflected in their substrate specificity preferences, that is, their ability to discriminate among substrates, which can be attributed to structural differences in surface loop topology that comprise the substrate binding sites (Perona and Craik, 1995; Hedstrom, 2002). Although the S1 subsite is recognized as the primary determinant of substrate specificity in SPs, additional subsites (e.g., S4–S4′) also contribute to substrate discrimination (Hedstrom, 2002) [the Schechter and Berger (1967) nomenclature is used to describe the interaction between protease subsites (Sn–S1;S1′–Sn′) and corresponding substrate residues (Pn–P1;P1′–Pn′), where P1–P1′ denotes the scissile bond]. Based on the presence of an acidic Asp residue at the S1 subsite at position 189 (based on the chymotrypsinogen numbering scheme; Hartley, 1964), this KLK subgroup is predicted to cleave C-terminal to basic amino acids (Arg, Lys), consistent with a trypsin-like specificity (Borgono et al., 2004). Indeed, biochemical studies on recombinant KLK4 (Takayama et al., 2001; Matsumura et al., 2005; Debela et al., 2006b; Obiezu et al., 2006), KLK5 (Brattsand et al., 2005; Michael et al., 2005; Debela et al., 2006b), KLK13 (Sotiropoulou et al., 2003; Kapadia et al., 2004b) and KLK14 (Brattsand et al., 2005; Felber et al., 2005) to date, employing both conventional (e.g., fluorogenic/chromogenic peptides) (Zimmerman et al., 1977) and novel technologies (e.g., phage display substrates, positional-scanning synthetic combinatorial libraries, PS-SCL) (Harris et al., 2000; Deperthes, 2002), have confirmed their predicted trypsin-like activity and also indicate a P1 preference for Arg over Lys. Uniquely, KLK14 manifests a dual trypsin- and chymotrypsin-like specificity, as it can accommodate both basic and aromatic residues at its S1 subsite (Felber et al., 2005).

However, with the exception of the recent study by Debela et al. (2006b), definitions and comparisons of extended KLK substrate specificities have not been intensely investigated. Thus, the present study explores the primary (P1) and extended (P2–P4) specificity profiles of heterologously expressed recombinant KLK4, 5, 13 and 14 using PS-SCL technology in an attempt to further distinguish the specificities of these closely related SPs and to identify biologically relevant substrates.

Results

Substrate specificities of KLK4, KLK5, KLK13 and KLK14

Primary (P1) subsite preferences Two-position fixed PS-SCLs (P1/P2, P1/P3, P1/P4) were used to determine the P1 specificity profiles of KLK4, 5, 13 and 14 (Figure 1A; Table 1). All KLKs displayed common and strong selectivity for basic residues at the P1 position, with a stringent preference for Arg, as predicted by the presence of Asp^{189} at their S1 subsites (Yousef and Diamandis, 2001). P1-Lys-containing substrates were also hydrolyzed by KLK4, KLK5 and KLK14, but at considerably lower efficiencies.

These results are in general accordance with prior studies on the primary substrate specificity of KLK4 (Takayama et al., 2001; Matsumura et al., 2005; Obiezu et al., 2006; Debela et al., 2006b), KLK5 (Brattsand et al., 2005; Michael et al., 2005; Debela et al., 2006b), KLK13 (Sotiropoulou et al., 2003; Kapadia et al., 2004b) and KLK14 (Brattsand et al., 2005; Felber et al., 2005). However, for KLK4, subtle differences exist among the three PS-SCL-generated P1 profiles reported to date (Matsumura et al., 2005; Debela et al., 2006b). Although all three studies demonstrate that KLK4 has an appreciable P1 preference for Arg and Lys, the work of Debela et al. (2006b) indicates that KLK4 can also accommodate Gly, Glu and Asn at its S1 subsite, whereas our findings suggest that KLK4 may also tolerate Phe at P1. These discrepancies may be attributed to the nature of the recombinant KLK4 protein used in each study, a result of the different expression systems [e.g., *Pichia pastoris*, *Drosophila melanogaster* S2 cells (Matsumura et al., 2005) and *Escherichia coli* (Debela et al., 2006b)] and distinct methods (e.g., refolding and/or enterokinase activation) used to obtain the enzymatically active enzyme. Further studies are required to clarify these inconsistencies. In addition, our P1 profile of KLK13 implies that Asn and Ala are also tolerated at this position, which has not been reported elsewhere (Sotiropoulou et al., 2003; Kapadia et al., 2004b). Although we have previously ascribed a dual trypsin- and chymotrypsin-like specificity to KLK14 by phage display substrate technology (Felber et al., 2005), we did not detect its chymotrypsin-like P1-

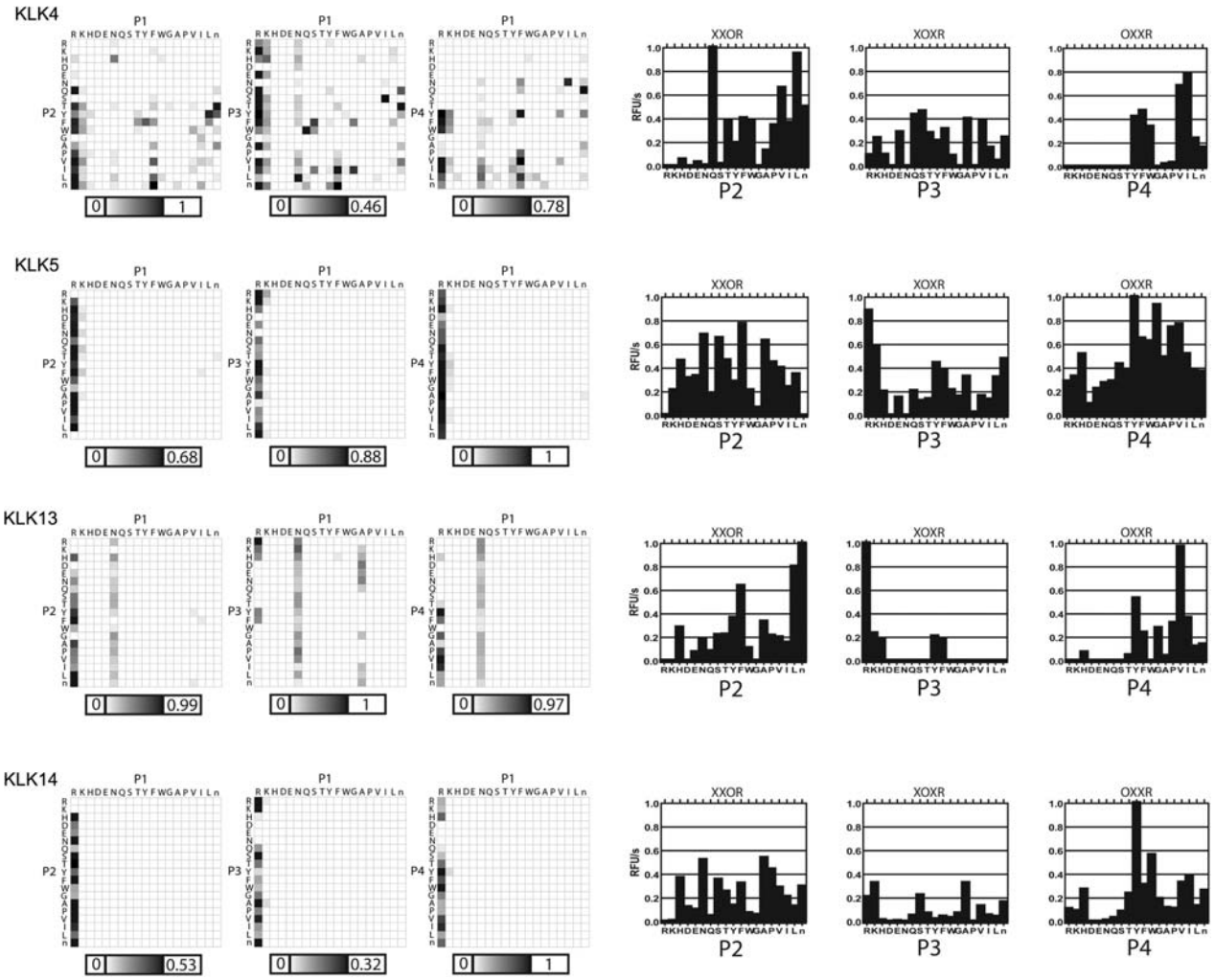

Figure 1 Primary and extended substrate specificities of KLK4, 5, 13 and 14 as determined by PS-SCL screening.
(A) 2D array results illustrating KLK activity in the two-position fixed (P1–P4) ACC PS-SCL libraries. Each array shows the rate of substrate cleavage [in relative fluorescence units (RFU) over time per nanomol of KLK (RFU/s/nM)], normalized on a scale of 0–1 and represented by the relative degree of shading. The shade of each square indicates the rate of substrate cleavage, ranging from white for no activity to black for highest activity. The relative cleavage rates were normalized within each library, with the highest rate observed in that library represented in the grayscale box below each matrix (scale of 0–1, 0 being the lowest activity and 1 being the highest activity observed across the libraries for a given enzyme). The x- and y-axes represent amino acid residues at the P1, P2, P3 and P4 positions. The two positions in each substrate that are not held constant contain an equimolar mixture of 19 amino acids (n represents norleucine, a synthetic amino acid structurally similar to Met; Met and Cys are excluded), for a total of 361 substrates/well. (B) Histograms depicting the extended N-terminal substrate specificities of KLKs determined by two-position fixed libraries in which the P1 position is held constant as Arg. The identity of the amino acid residue in the fixed position is indicated on the x-axis. The y-axis is reaction velocity in RFU/s.

Tyr preference by PS-SCL. This may suggest that residues C-terminal to the scissile bond (i.e., P1′–P4′), which are not included in our PS-SCLs, may determine the chymotrypsin-like specificity of KLK14. The latter is true for KLK1, which manifests a dual enzymatic activity similar to KLK14, since its cleavage after P1-Met and P1-Phe requires specific residues from P2 to P2′ and Pro at P3′ (Chagas et al., 1995).

The unique dual specificity of KLK14 may be due to its distinct amino acid constituents within two surface loops (loops G and H; Figure 2) comprising residues 185–188 and 221–225, which, in addition to residue 189,

Table 1 Summary of P4–P1 residues most preferred by KLKs.

KLK	Amino acid position			
	P4	P3	P2	P1
KLK4	VI (aliphatic)	QSAR (polar, aliphatic)	QL (polar, aliphatic)	R
KLK5	YFWGPV (aliphatic, aromatic)	RK (basic)	NSFAM (polar, aliphatic)	R
KLK13	VY (aliphatic, aromatic)	R (basic)	LFM (aromatic, aliphatic)	R
KLK14	YW (aromatic)	RKSAM (basic, aliphatic)	HNSPA (polar, small)	R

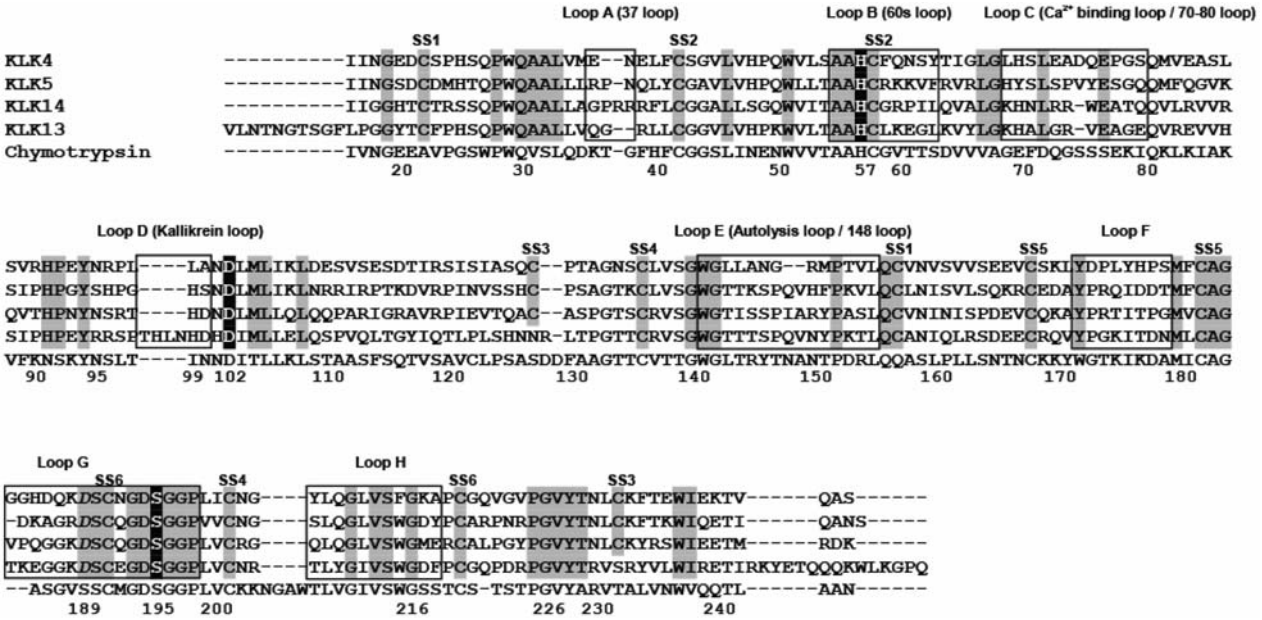

Figure 2 Multiple sequence alignment of KLK4, 5, 13 and 14 and bovine chymotrypsin.
The chymotrypsinogen numbering (Hartley, 1964) used throughout is shown on the bottom. Dashes represent gaps that bring the sequences to better alignment. Fully conserved amino acids are highlighted in gray. The amino acids of the catalytic triad (H57, D102, S195) are shown in white font on a black background. Residue 189 at the bottom of the S1 binding pocket is shown in italics. Disulfide bond-forming cysteine residues are labeled SS1–SS6 above each residue. Surface loop regions are boxed and labeled above.

are also known to bind to P1 residues (Cohen et al., 1981) and contribute to P1 specificity (Hedstrom et al., 1992). Examination of these loop regions reveals that KLK14 contains residues that are not present in other human KLKs at similar positions (Yousef and Diamandis, 2001), including the residues Gln and Tyr at positions 185 and 224, respectively, which may facilitate the binding of P1-Tyr residues and confer its chymotryptic-like specificity.

Extended (P2–P4) subsite preferences Compared to the primary subsite, KLK subsite preferences on the N-terminal side of the scissile bond are generally less stringent (Figure 1B, Table 1). The P2–P4 specificity profiles for all KLKs described below were obtained using P1-Arg fixed sublibraries.

P2 preferences In general, liberal substitutions were tolerated at P2 for all KLKs studied. For instance, a variety of polar (Asn, Asp), aliphatic (Ala, Pro, Val, Leu, Ile) and aromatic (Phe, Tyr, Trp) residues were accepted at the S2 subsite to varying degrees. Among all the KLKs, KLK5 and KLK14 displayed the most analogous P2 specificity profiles; P2 substituents that resulted in greater than 50% activity were generally Asn, Ala, Phe, Ser and His, whereas those giving rise to 15% activity or less were commonly Gln and Gly. KLK13 exerted a relatively more stringent P2 specificity, with a P2 preference for Leu and Phe (<50% activity) and, akin to KLK5, KLK14 did not favor Gln and Gly (>10% activity). In stark contrast to the intolerance for P2-Gln of the latter KLKs, Gln was the most favored P2 residue for KLK4, followed by Leu and Val, consistent with previous findings (Matsumura et al., 2005; Debela et al., 2006b). Furthermore, with the exception of KLK5, which can accept Lys at P2, all KLKs demonstrate a common intolerance for the basic amino acids Arg and Lys at their S2 subsites, although considerable activity was displayed against P2-His-containing substrates. Uniquely, KLK4 could not accommodate any charged P2 residue.

Variations in P2 specificity among KLK4, 5, 13 and 14 can be mainly attributed to the evolved diversity in a surface loop denoted the 'kallikrein loop' that forms the S2 subsite (Katz et al., 1998; Kishi et al., 1999) (Figure 2). The kallikrein loop, unique to the KLK family, is a 3–11-aa insertion between the sixth and seventh β-sheets and is only present in its entirety in the 'classical' KLKs, KLK1–3 (Yousef and Diamandis, 2001). Among the KLKs studied, KLK13 has the longest kallikrein loop consisting of a 5-aa insertion (Figure 2) and was also found to exhibit the most stringent P2 preferences. Thus, the comparatively long kallikrein loop of KLK13 may be responsible for restricting the size, shape and/or physicochemical properties of its S2 subsite, as reported for other KLKs (Katz et al., 1998; Kishi et al., 1999; Carvalho et al., 2002), which is reflected in the narrower range of P2 residues that it can accommodate. KLK4 demonstrated a unique P2 preference for hydrophobic residues and intolerance for charged residues, whereas the exact opposite was observed for the remaining KLKs. The latter is likely a function of the residue at position 99 within the kallikrein loop (Figure 2). With the exception of KLK4, which harbors Leu at position 99, all other KLKs possess a His residue (Figure 2). The latter could increase the polarity of the S2 subsite and favor interactions with oppositely charged P2 residues (i.e., Arg, Lys), as evidenced in this study. However, Leu99 in KLK4 would increase the hydrophobicity of S2 and favor hydrophobic but not

charged P2 residues, as observed. The significance of residue 99 in determining P2 specificity has also been shown for KLK1, which harbors a Tyr at position 99 that mainly confers its specificity for P2-Phe (Chen and Bode, 1983; Deshpande and Burton, 1992).

P3 preferences Compared to their low P2 specificities, the KLKs exhibited relatively higher selectively for substituents at P3. Among the KLKs, KLK13 displayed the most stringent specificity at P3, followed by KLK14, KLK5 and KLK4, which could tolerate a broader range of amino acids at this position. KLK5, 13 and 14 all demonstrated a strong preference for basic, positively charged amino acids (Arg and/or Lys) at P3. KLK14 could also hydrolyze P3-Ala- and Ser-containing substrates to a similar extent. On the contrary, KLK4 did not favor basic P3 residues, but instead preferred substrates with aliphatic P3 residues, namely Ser, Gln, Ala, and Val. The negatively charged residues Asp and Glu were generally excluded from all S3 subsites, with the exception of P3-Glu, which was moderately tolerated by KLK4 and KLK5 (<15% activity).

Our data on the P3 specificity of KLK4 and KLK5 are in general agreement with prior reports (Matsumura et al., 2005; Michael et al., 2005; Debela et al., 2006b; Obiezu et al., 2006). For example, previous studies on KLK4 also indicate a P3-preference for aliphatic residues (i.e., Gln, Ser, Ala, Val) (Matsumura et al., 2005; Obiezu et al., 2006; Debela et al., 2006b). However, in contrast to our work, Debela et al. (2006b) found a lower specificity at P3 compared with P2 for KLK4 and KLK5, whereas we report the opposite trend. These discrepancies may be due to differences in enzyme and/or PS-SCL preparation.

With respect to the crystal structure of KLK4, Debela et al. (2006a) described a wide S3 pocket bordered by Ala218 and Asn192 within loops G and H, respectively, which are likely responsible for the relatively low P3 specificity observed. KLK5 and 14 harbor Gln at position 192, but possess distinct residues at position 218 (Figure 2), which may account for their partially overlapping yet distinct P3 preferences. KLK5 possesses an aromatic Tyr at 218, which may reflect its relatively broader specificity at P3 and ability to interact with polar residues. Furthermore, KLK14 has an acidic Glu residue at 218, which may account for its preference for basic P3 amino acids. Interestingly, KLK13 possesses Glu at position 218 and a Phe residue at 192 (Figure 2), which may explain its ability to uniquely accommodate only basic (Arg, Lys, His) and aromatic (Tyr, Phe) P3 substituents.

P4 preferences As for the S2 subsite, S4 also exhibited low specificity among KLKs, yet general preferences were evident. KLK5 was the least restrictive in its P4 preferences, followed by KLK14, KLK4 and KLK13, which tolerated fewer resides at this position. KLK4, 5, 13 and 14 efficiently digested (>50% activity) substrates containing hydrophobic aliphatic and/or aromatic residues, yet their activities were lowest (<10%) when S4 was occupied by acidic, negatively charged amino acids. KLK4 preferred aliphatic residues (e.g., Val, Ile), while KLK14 favored aromatic amino acids (e.g., Tyr, Trp, Phe).

Analogous results for the P4 specificity profile of KLK4 and KLK5 were obtained by previous PS-SCL screening studies (Matsumura et al., 2005; Debela et al., 2006b). All data on the P4 specificity of KLK4 indicate that hydrophobic residues, particularly Val and Ile, are most favored at this position (Matsumura et al., 2005; Debela et al., 2006b). Our work and that of Debela et al. (2006b) demonstrates that KLK5 prefers bulky (e.g., Tyr) and small (e.g., Gly) P4 substituents. Furthermore, our study is the first to report the P4 specificity profile of KLK13 and KLK14.

Surface loop F, and in particular residue 175, represents an important determinant of P4 specificity (Figure 2). For instance, the S4 subsite of KLK4 comprises Leu175, Leu99 and Phe215 and forms a hydrophobic pocket that ideally accommodates aliphatic residues such as Val and Ile (Debela et al., 2006a). KLK5 possesses Gln175 and exhibits a low tolerance for acidic residues. KLK13 and KLK14 have polar residues at position 175, namely Lys and Thr, respectively, which may account for their more stringent P4 preferences.

Activity of KLK14 towards single AMC substrates

The PS-SCL specificity profiles of all KLKs indicate that, apart from their S1 requirement for basic residues (i.e., Arg≫Lys), the S2–S4 subsites may represent critical determinants of specificity. To validate the PS-SCL screening results, the dependence of KLK14 on extended interactions was determined by quantifying kinetic parameters for several single tetrapeptide AMC substrates (Table 2). The preference of KLK14 for Arg over Lys at the P1 position was verified with the substrates Ac-YAAR-AMC and Ac-YAAK-AMC, with the latter hydrolyzed with 2.5-fold lower catalytic efficiency than the former. The dependence of KLK14 specificity on P2–P4 residues was assessed using a series of suboptimal P4–P2 substrates, Ac-WVHR-AMC, Ac-FIGR-AMC and Ac-DDKR-AMC, which were designed based on the PS-SCL screen using P1-Arg fixed libraries (Figure 1B). The P2–P4 substituents of Ac-WVHR-AMC individually yielded approximately 50% maximal activity in the PS-SCL screen; those in Ac-FIGR-AMC resulted in approximately

Table 2 Steady-state kinetic parameters for KLK14 hydrolysis of single AMC substrates.

Substrate	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})	Activity (%)
Ac-YAAR-AMC	0.1242±0.007223	567.57	4569.81	100
Ac-YAAK-AMC	0.2558±0.005942	449.09	1755.63	38.4
Ac-WVHR-AMC	0.2260±0.01485	169.030	747.92	16.37
Ac-FIGR-AMC	0.2146±0.01185	53.770	250.557	5.48
Ac-DDKR-AMC	NR	–	–	–

NR, no reaction.

25% activity and the residues in Ac-DDKR-AMC yielded no or minimal activity (Figure 1B). As expected, comparable results were obtained with the single AMC substrates, that is, the activity of KLK14 against Ac-WVHR-AMC, Ac-FIGR-AMC and Ac-DDKR-AMC was 16%, 5% and 0%, respectively.

Identification of putative KLK substrates

The unique combination of individual S1–S4 subsite preferences of each KLK likely translates into distinctive *in vivo* substrate repertoires and (patho)physiological functions. The PS-SCL screening results imply that the most effective KLK substrates may contain the P4–P1 residues summarized in Table 1. A search of NCBI non-redundant protein databases revealed a number of candidate KLK substrates that contain the optimal P4–P1 recognition motifs predicted by PS-SCL (Table 3). Among the substrates identified, several have previously been shown to be cleaved by KLKs *in vitro*. For example, KLK4 can digest the α chain of fibrinogen (Obiezu et al., 2006) and KLK5 cleaves fibronectin (Michael et al., 2005). Although most of these putative substrates have yet to be experimentally verified, many represent logical KLK targets based on: (i) KLK specificity profiles; (ii) co-localization/co-expression with KLKs in tissues/fluids; and/or (iii) their relevance to (patho)physiological processes in which KLK action is implicated. For instance, KLK4 is co-expressed with cathelicidin in the epidermis (Yamasaki et al., 2006) and may play a role in innate immunity by processing the precursor form at a site known to generate active antimicrobial peptides, a function that has recently been ascribed to KLK5 and KLK7 (Yamasaki et al., 2006). KLK4 is also found at high levels in seminal plasma (Obiezu et al., 2005), where it may contribute to seminal clot liquefaction, along with other KLKs (e.g., KLK3, 5) (Malm et al., 2000; Michael et al., 2006), through digestion of semenogelin-1. KLK13 may affect the activity of a disintegrin and metalloproteinase with thrombospondin motifs 3 (ADAMTS-3), which may in turn alter its ability to process collagen II within the ECM of cartilage (Fernandes et al., 2001). KLK14 represents an activator of the transmembrane SP TMPRSS3, which is implicated in ovarian (Sawasaki et al., 2004) and pancreatic (Iacobuzio-Donahue et al., 2003) cancers and in epithelial sodium channel activation (Guipponi et al., 2002).

Discussion

In recent years, combinatorial approaches such as PS-SCL screening have emerged as useful quantitative methods not only to confirm and elucidate the specificity of a number of proteases (e.g., trypsin, chymotrypsin, thrombin, plasmin, tryptases, caspases, granzymes, hepsin, prostasin, chymase) (Thornberry et al., 1997; Backes et al., 2000; Harris et al., 2000, 2001; Raymond et al., 2003; Shipway et al., 2004; Herter et al., 2005), but also for the identification of novel substrates and the production of selective inhibitors.

Using two-position fixed PS-SCLs with exhaustive diversity at the P1–P4 positions, we were able to define primary and extended substrate specificities of homologous extracellular SPs of the tissue kallikrein family (Yousef and Diamandis, 2001; Borgono et al., 2004), namely KLKs 4, 5, 13 and 14, and identify distinctive subsite preferences (Table 1). All four KLKs display a common, almost exclusive preference for basic P1 residues, primarily Arg. However, each KLK exhibited a unique extended (P4–P2) specificity. Indeed, such extended interactions are known to critically influence the binding and hydrolysis of substrates by SPs (Hedstrom, 2002). In the case of KLK4, the majority of its extended N-terminal specificity resides within its S2 and S4 subsites, whereas liberal substitutions are tolerated at S3. For KLK5, P3 substituents primarily influence specificity, whereas for KLK13 and KLK14, both P3 and P4 may be important determinants of extended specificity. Notably, the optimal subsite occupancy across S1–S4 for KLK4, 5, 13 and 14 is unique not only among KLKs, but also compared to those of other SPs (Thornberry et al., 1997; Backes et al., 2000; Harris et al., 2000, 2001; Raymond et al., 2003; Shipway et al., 2004; Herter et al., 2005). As described above, variations in substrate discrimination among KLK4, 5, 13 and 14 can be attributed to the diversity in surface loop structures that surround the substrate-binding site/active site and comprise the extended subsites. Additional studies are needed to delineate the probable contribution of residues C-terminal to the scissile bond to KLK substrate specificity. Such residues are known to be important for substrate discrimination in other KLKs (e.g., S1′–S3′ for KLK1 and S2′ for KLK3) (Chagas et al., 1995; Coombs et al., 1998) and proteases (e.g., S4′ for prostasin) (Shipway et al., 2004).

Verification of PS-SCL specificity profiling for KLK14 was kinetically assessed using single AMC tetrapeptide substrates. Moreover, our PS-SCL results are also consistent with previous studies in which KLK specificity was kinetically determined directly using AMC tripeptide substrates and indirectly via serpin inhibitors. For instance, KLK4 was able to hydrolyze VPR-AMC at a five-fold higher efficiency than GPR-AMC (Obiezu et al., 2006), indicative of a preference for Val over Gly at P3, in agreement with our data. With respect to KLK14, minimal catalytic activity was observed against AMC substrates containing basic P2 residues (i.e., Arg, Lys) and acidic P3–P4 residues (i.e., Asp), whereas the highest efficiencies were observed for P2-Ala, which is also consistent with our present data (Borgono et al., 2007). Inhibitory serpins behave as irreversible suicide inhibitors and are cleaved by their protease targets at a scissile bond (P1–P1′) within their exposed reactive site loops (Gettins, 2002). The P4–P1′ sequences for the serpins α_2-antiplasmin (AP) and antithrombin III (AT) are AMSR$^{364}\downarrow$M and IAGR$^{393}\downarrow$S, respectively, where the arrow denotes the cleavage site. Michael et al. (2005) have shown that the inhibitory reaction between KLK5 and AP was 24-fold faster than that with AT, which may be partially attributed to the highly favored P2 substituent in AP (Met) vs. AT (Gly). The residues present in the P3 and P4 positions in each serpin are similarly favored by KLK5 according to our screening results.

Since the consequence of KLK proteolytic action may be largely determined by its extended substrate specificity, knowledge of optimal KLK P4–P1 preferences may

Table 3 Subset of putative KLK protein targets containing PS-SCL-based motifs.

KLK	Putative substrate (GenBank accession no.)	Motif				Motif location	Function
		P4	P3	P2	P1		
KLK4	Cathelicidin antimicrobial peptide precursor (NP_004336)	I	V	Q	R	At a known proteolytic cleavage site	Microbial defense and innate immunity
	Hepsin (NP_002142)	V	S	L	R	Within catalytic domain	Transmembrane serine protease involved in blood coagulation and prostate cancer progression
	KLK13 (NP_056411)	I	Q	L	R	Within catalytic domain	Kallikrein serine protease implicated in tumor metastasis
	Semenogelin-1 (NP_002998)	V	S	Q	R	At a known proteolytic cleavage site	Gel-forming protein of seminal plasma
	Fibrinogen α chain (P02671)	I	Q	Q	R	Within C-terminal globular domain	Plasma protein involved in coagulation; degradation fragments regulate cell adhesion and spreading and have vasoconstrictor, chemotactic and mitogenic activities
KLK5	Urokinase-type plasminogen activator (NP_002649)	G	R	S	R	Within catalytic domain	Serine protease involved in conversion of plasminogen to plasmin; ECM degradation; tumor cell migration and proliferation
	Thyroglobulin (NP_003226)	G	R	F	R	Between second and third thyroglobulin type I repeat domains	Precursor of thyroid hormones (T3 and T4)
	Fibronectin (NP_997647)	P	R	A	R	FN3 domain	Plasma and ECM protein involved in cell adhesion and migration processes including embryogenesis, wound healing, blood coagulation, host defense and metastasis
KLK13	ADAMTS-3 (NP_055058)	V	R	F	R	Between ADAM cysteine-rich domain and first thrombospondin type 1 repeat domain	Metalloprotease domain cleaves the propeptides of type II collagen prior to fibril assembly within cartilage and skin
	Neurotrypsin (NP_003610)	V	R	L	R	Between Sushi and first SR domain	Serine protease implicated in neuronal plasticity, learning and memory
	Laminin α5 chain (NP_005551)	V	R	L	R	LAMnt domain	Component of the BL
	Collagen α1 (VII) chain (NP_000085)	V	R	L	R	FN3 domain	Component of the BL between the epidermal/dermal junction
	Cadherin-16 (NP_004053)	V	R	L	R	CAD domain	Adhesion molecule implicated in tumor suppression
KLK14	Matrilin-4 (NP_003824)	W	A	A	R	vWA domain	Component of cartilage ECM
	Collagen α2 (IV) (NP_001837)	Y	A	S	R	Canstatin region	Component of the BL
	Fibrinogen β chain (NP_005132)	Y	R	A	R	Before fibrinogen domain	Plasma protein involved in coagulation; degradation fragments regulate cell adhesion and spreading, and have vasoconstrictor, chemotactic and mitogenic activities
	Transmembrane protease, serine 3 (NP_076927)	Y	S	S	R	Activation site	Transmembrane serine protease associated with deafness and malignancy
	Desmoglein 3 (NP_001935)	Y	S	A	R	Between CAD domains	Calcium-binding desmosomal cadherin that mediates cell-cell adhesion within the epidermis

ADAMTS-3, a disintegrin and metalloproteinase with thrombospondin motifs 3; BL, basal lamina; CAD, cadherin; ECM, extracellular matrix; FN3, fibronectin type 3 domain; KLK, kallikrein 1-related peptidase; LAMnt, laminin N-terminal domain; SR, scavenger receptor Cys-rich domain; vWA, von Willebrand factor type A domain.

allow the identification of novel KLK substrates. In fact, a multitude of macromolecular targets containing the preferred PS-SCL-based P4–P1 recognition motifs for each KLK were retrieved from protein databases. The subset of putative KLK substrates listed in Table 3 is particularly relevant, as these further support the notion that KLKs contribute to tumor progression via digestion of extracellular matrix components (e.g., collagens, laminins, fibronectin; Borgono and Diamandis, 2004), tumor growth regulation by regulating IGF activity (e.g., IGFBP-6 degradation; Borgono and Diamandis, 2004), the generation of antimicrobial peptides in the skin (e.g., cathelicidin processing; Yamasaki et al., 2006), the liquefaction of seminal clots (e.g., semenogelin degradation;

Michael et al., 2006) and proteolytic cascade pathways (e.g., possible regulation/activation of hepsin, neurotrypsin, TMPRSS3; Borgono et al., 2004; Michael et al., 2006). Interestingly, matrilin-4, a component of cartilage ECM, was identified as a KLK14 substrate in the present study and by phage display substrate (Felber et al., 2005). Further studies are warranted to validate these findings.

Indeed, PS-SCL technology has facilitated the discovery of novel substrates for several proteases. Takeuchi et al. (2000) utilized the collective results from two combinatorial peptide-based specificity profiling technologies, PS-SCL and phage display substrate, to identify the preferred cleavage sequence and putative substrates (e.g., single chain uPA and proteinase-activated receptor-2) of membrane-type SP 1. Using PS-SCL screening alone, other groups have identified and verified albumin as a novel chymase substrate (Raymond et al., 2003) and hepatocyte growth factor as a target for hepsin (Herter et al., 2005). PS-SCL has also guided the development of efficient and specific inhibitors for several proteases, including β-tryptases (Harris et al., 2001) and chymase (Raymond et al., 2003).

In conclusion, our data provide novel information on specificity differences among closely related members of the KLK family and may guide future research into the design of specific small-molecule KLK inhibitors, identification of biological KLK targets and the role of KLKs in normal and pathological processes.

Materials and methods

Materials

Single fluorogenic tetrapeptides with the general structure Ac-P4-P3-P2-P1-AMC (where Ac represents an acetyl group and AMC denotes 7-amino-4-methylcoumarin), Ac-Tyr-Ala-Ala-Arg-AMC (Ac-YAAR-AMC), Ac-Tyr-Ala-Ala-Lys-AMC (Ac-YAAK-AMC), Ac-Trp-Val-His-Arg-AMC (Ac-WVHR-AMC), Ac-Phe-Ile-Gly-Arg-AMC (Ac-FIGR-AMC), Ac-Asp-Asp-Lys-Arg-AMC (Ac-DDKR-AMC), were synthesized by Enzyme Systems Products, Inc. (Livermore, CA, USA), diluted to a final concentration of 20 mM in dimethyl sulfoxide and stored at -20°C. Peptides showed >98% purity by HPLC analysis and their molecular weights were confirmed by mass spectrometry. Recombinant KLK4, KLK5, KLK13 and KLK14 were produced in-house using the Easyselect™ *P. pastoris* expression system (Invitrogen, Carlsbad, CA, USA) as previously described in detail (Sotiropoulou et al., 2003; Felber et al., 2005; Michael et al., 2005; Obiezu et al., 2006). Briefly, the sequences encoding pro-KLK5 and pro-KLK13 were cloned into the *P. pastoris* expression vector pPIC9 and transformed into yeast strain KM71. Although both enzymes were cloned in zymogen form, enzymatically active KLK5 and KLK13 were obtained after purification, indicating that autoactivation had occurred. The sequences encoding mature KLK4 and KLK14 were cloned into the *P. pastoris* expression vector pPICZαA and transformed into yeast strains KM71H and X-33, respectively. KLKs were purified to homogeneity (>95% purity on Coomassie Blue-stained polyacrylamide gels) via a one- or two-step chromatographic procedure (KLK4, anion exchange and benzamidine affinity chromatography; KLK5, cation exchange and reversed-phase chromatography; KLK13, hydrophobic interaction chromatography; KLK14, cation exchange and soybean trypsin inhibitor affinity chromatography) and their identities were confirmed by tandem mass spectrometry and N-terminal sequencing analyses.

PS-SCL screening

PS-SCLs were synthesized and screened as previously described by Harris et al. (2000). PS-SCLs comprised three separate two-position fixed sublibraries in which the P1 position and a second position (P2, P3 or P4) was held constant, while the remaining positions consisted of an equimolar mixture of 19 amino acids (Cys was omitted and Nle was substituted for Met). PS-SCLs were added to 384-well optical black plates (Nunc, Rochester, NY, USA), with each well containing 361 individual substrates at a concentration of approximately 0.25 μM/substrate/well to a final volume of 30 μl in buffer. Each sublibrary consisted of 361 substrates/well for a total of 130 321 individual substrates. Hydrolysis of PS-SCL substrates was initiated by the addition of KLK4, KLK5, KLK13 or KLK14 to each well at a final concentration of 100 nM. Initial rates of KLK-mediated peptide hydrolysis were monitored by measuring the release of free ACC on a Molecular Devices Gemini EM instrument (Molecular Devices, Sunnyvale, CA, USA), with excitation at 380 nm and emission at 450 nm, at 37°C for 30 min. Kinetic values are reported as relative fluorescence units per second and normalized across libraries for each individual enzyme.

Single-substrate kinetic assays

Based on the resultant PS-SCL-based KLK14 specificity profile, the optimal peptide substrate (Ac-YAAR-AMC), as well as six suboptimal substrates (Ac-YAAK-AMC, Ac-YKNR-AMC, Ac-YKNK-AMC, Ac-WVHR-AMC, Ac-FIGR-AMC, Ac-DDKR-AMC), was synthesized and hydrolyzed by KLK14. Individual peptides at various final concentrations (0.004–2 mM) were incubated with recombinant KLK14 (final concentration 12 nM) at a final volume of 100 μl under optimal buffer conditions (100 mM Na_2HPO_4, pH 8, 0.01% Tween-20) in 96-well white polystyrene microtiter plates. Reaction mixtures contained less than 5% (v/v) dimethyl sulfoxide. The initial reaction rate of KLK14-mediated peptide hydrolysis was monitored by measuring free AMC fluorescence on the Wallac 1420 Victor^{2}™ fluorimeter (Perkin-Elmer, Wellesley, MA, USA) with excitation and emission filters of 380 nm and 480 nm, respectively, at 1-min intervals for 20 min at 37°C. KLK14-free reactions were used as negative controls and the background counts obtained were subtracted from each value. A standard curve was constructed using known concentrations of AMC to convert the rates of reaction from AMC fluorescence counts/min to free AMC produced/min. The slope of the resultant AMC standard curve was 19.2 AMC fluorescence counts/nM free AMC. The steady-state (Michaelis-Menten) kinetic constants (k_{cat}/K_m) were then calculated by non-linear regression analysis using Enzyme Kinetics Module 1.1 (Sigma Plot, SSPS, Chicago, IL, USA). All experiments were performed in triplicate and repeated at least twice.

Database searches for potential KLK substrates

Tetrapeptide substrate motifs were designed for each KLK based on the most preferred amino acid(s) (i.e., amino acids at P1–P4 positions that yielded approximately >50% activity, where % activity was determined by dividing the fluorescence intensity of an individual amino acid at each position by the amino acid that resulted in the highest fluorescence at each position, for each KLK individually) at S1–S4 subsites as deduced from PS-SCL profiles (Figure 2). The substrate motifs, written in PROSITE syntax, were as follows: [VI]-[QSAV]-[QL]-R for KLK4; [YFWGPV]-[RK]-[NSFAM]-R, for KLK5; X-[AP]-[HNMA]-R, for KLK13; and [YW]-[KRSAM]-[HNSPA]-R, for KLK14. The non-

redundant protein databases of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) were then searched for matches to these motifs, using the Pattinprot algorithm (http://npsa-pbil.ibcp.fr) at the Network Protein Sequence Analysis web server, which scans protein databases for pattern(s) (Combet et al., 2000). The matching proteins retrieved by this process were filtered by selecting only physiologically relevant macromolecular protein substrates for KLKs, namely, proteins whose expression profiles and cellular localization (i.e., proteins targeted to the secretory pathway, including secreted proteins and membrane proteins in which the substrate motif was located in an extracellular region) coincided with those of the KLKs. Proteins from non-human species, as well as human proteins that were classified as intracellular, 'hypothetical' or had an unknown function, were excluded from further analysis. The remaining proteins were characterized in terms of structure, function, expression and association with disease pathology using several bioinformatic databases, including the Human Protein Reference Database (http://www.hprd.org), SMART (http://smart.embl-heidelberg.de), Entrez Gene (http://www.ncbi.nih.gov/entrez), OMIM (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) and the Gene Ontology (GO) database (http://www.godatabase.org/cgi-bin/amigo/go.cgi).

References

Backes, B.J., Harris, J.L., Leonetti, F., Craik, C.S., and Ellman, J.A. (2000). Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. Nat. Biotechnol. *18*, 187–193.

Barrett, A.D., Rawlings, N.D., and Woessner, J.F. (2004). Handbook of Proteolytic Enzymes, 2nd edition (London, UK: Academic Press).

Bernett, M.J., Blaber, S.I., Scarisbrick, I.A., Dhanarajan, P., Thompson, S.M., and Blaber, M. (2002). Crystal structure and biochemical characterization of human kallikrein 6 reveals that a trypsin-like kallikrein is expressed in the central nervous system. J. Biol. Chem. *277*, 24562–24570.

Bhoola, K.D., Figueroa, C.D., and Worthy, K. (1992). Bioregulation of kinins: kallikreins, kininogens, and kininases. Pharmacol. Rev. *44*, 1–80.

Borgono, C.A. and Diamandis, E.P. (2004). The emerging roles of human tissue kallikreins in cancer. Nat. Rev. Cancer. *4*, 876–890.

Borgono, C.A., Michael, I.P., and Diamandis, E.P. (2004). Human tissue kallikreins: physiologic roles and applications in cancer. Mol. Cancer Res. *2*, 257–280.

Borgono, C.A., Michael, I.P., Shaw, J.L., Luo, L.Y., Ghosh, M.C., Soosaipillai, A., Grass, L., Katsaros, D., and Diamandis, E.P. (2007). Expression and functional characterization of the cancer-related serine protease, human tissue kallikrein 14. J. Biol. Chem. *26*, 2405–2422.

Brattsand, M. and Egelrud, T. (1999). Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation. J. Biol. Chem. *274*, 30033–30040.

Brattsand, M., Stefansson, K., Lundh, C., Haasum, Y., and Egelrud, T. (2005). A proteolytic cascade of kallikreins in the stratum corneum. J. Invest. Dermatol. *124*, 198–203.

Carvalho, A.L., Sanz, L., Barettino, D., Romero, A., Calvete, J.J., and Romao, M.J. (2002). Crystal structure of a prostate kallikrein isolated from stallion seminal plasma: a homologue of human PSA. J. Mol. Biol. *322*, 325–337.

Caubet, C., Jonca, N., Brattsand, M., Guerrin, M., Bernard, D., Schmidt, R., Egelrud, T., Simon, M., and Serre, G. (2004). Degradation of corneodesmosome proteins by two serine proteases of the kallikrein family, SCTE/KLK5/hK5 and SCCE/KLK7/hK7. J. Invest. Dermatol. *122*, 1235–1244.

Chagas, J.R., Portaro, F.C., Hirata, I.Y., Almeida, P.C., Juliano, M.A., Juliano, L., and Prado, E.S. (1995). Determinants of the unusual cleavage specificity of lysyl-bradykinin-releasing kallikreins. Biochem. J. *306*, 63–69.

Chen, Z. and Bode, W. (1983). Refined 2.5 Å X-ray crystal structure of the complex formed by porcine kallikrein A and the bovine pancreatic trypsin inhibitor. Crystallization, Patterson search, structure determination, refinement, structure and comparison with its components and with the bovine trypsin-pancreatic trypsin inhibitor complex. J. Mol. Biol. *164*, 283–311.

Cohen, G.H., Silverton, E.W., and Davies, D.R. (1981). Refined crystal structure of γ-chymotrypsin at 1.9 Å resolution. Comparison with other pancreatic serine proteases. J. Mol. Biol. *148*, 449–479.

Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. (2000). NPS@: network protein sequence analysis. Trends Biochem. Sci. *25*, 147–150.

Coombs, G.S., Bergstrom, R.C., Pellequer, J.L., Baker, S.I., Navre, M., Smith, M.M., Tainer, J.A., Madison, E.L., and Corey, D.R. (1998). Substrate specificity of prostate-specific antigen (PSA). Chem. Biol. *5*, 475–488.

Debela, M., Magdolen, V., Grimminger, V., Sommerhoff, C., Messerschmidt, A., Huber, R., Friedrich, R., Bode, W., and Goettig, P. (2006a). Crystal structures of human tissue kallikrein 4: activity modulation by a specific zinc binding site. J. Mol. Biol. *362*, 1094–1107.

Debela, M., Magdolen, V., Schechter, N., Valachova, M., Lottspeich, F., Craik, C.S., Choe, Y., Bode, W., and Goettig, P. (2006b). Specificity profiling of seven human tissue kallikreins reveals individual subsite preferences. J. Biol. Chem. *281*, 25678–25688.

Deperthes, D. (2002). Phage display substrate: a blind method for determining protease specificity. Biol. Chem. *383*, 1107–1112.

Deperthes, D., Frenette, G., Brillard-Bourdet, M., Bourgeois, L., Gauthier, F., Tremblay, R.R., and Dube, J.Y. (1996). Potential involvement of kallikrein hK2 in the hydrolysis of the human seminal vesicle proteins after ejaculation. J. Androl. *17*, 659–665.

Descargues, P., Deraison, C., Prost, C., Fraitag, S., Mazereeuw-Hautier, J., D'Alessio, M., Ishida-Yamamoto, A., Bodemer, C., Zambruno, G., and Hovnanian, A. (2006). Corneodesmosomal cadherins are preferential targets of stratum corneum trypsin- and chymotrypsin-like hyperactivity in Netherton syndrome. J. Invest. Dermatol. *126*, 1622–1632.

Deshpande, M.S. and Burton, J. (1992). Mapping the binding site of tissue kallikrein: preparation and testing of all possible substrate analog inhibitors homologous with the sequence of kininogen between Ser386 and Gln392. J. Med. Chem. *35*, 3094–3102.

Felber, L.M., Borgono, C.A., Cloutier, S.M., Kundig, C., Kishi, T., Ribeiro, C.J., Jichlinski, P., Gygi, C.M., Leisinger, H.J., Diamandis, E.P., et al. (2005). Enzymatic profiling of human kallikrein 14 using phage-display substrate technology. Biol. Chem. *386*, 291–298.

Fernandes, R.J., Hirohata, S., Engle, J.M., Colige, A., Cohn, D.H., Eyre, D.R., and Apte, S.S. (2001). Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. J. Biol. Chem. *276*, 31502–31509.

Gettins, P.G. (2002). Serpin structure, mechanism, and function. Chem. Rev. *102*, 4751–4804.

Ghosh, M.C., Grass, L., Soosaipillai, A., Sotiropoulou, G., and Diamandis, E.P. (2004). Human kallikrein 6 degrades extracellular matrix proteins and may enhance the metastatic potential of tumour cells. Tumour Biol. *25*, 193–199.

Guipponi, M., Vuagniaux, G., Wattenhofer, M., Shibuya, K., Vazquez, M., Dougherty, L., Scamuffa, N., Guida, E., Okui, M., Rossier, C., et al. (2002). The transmembrane serine protease (TMPRSS3) mutated in deafness DFNB8/10 activates the epithelial sodium channel (ENaC) *in vitro*. Hum. Mol. Genet. *11*, 2829–2836.

Hachem, J.P., Wagberg, F., Schmuth, M., Crumrine, D., Lissens, W., Jayakumar, A., Houben, E., Mauro, T.M., Leonardsson, G., Brattsand, M., et al. (2006). Serine protease activity and residual LEKTI expression determine phenotype in Netherton syndrome. J. Invest. Dermatol. *126*, 1609–1621.

Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A., and Craik, C.S. (2000). Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. USA *97*, 7754–7759.

Harris, J.L., Niles, A., Burdick, K., Maffitt, M., Backes, B.J., Ellman, J.A., Kuntz, I., Haak-Frendscho, M., and Craik, C.S. (2001). Definition of the extended substrate specificity determinants for β-tryptases I and II. J. Biol. Chem. *276*, 34941–34947.

Hart, P.S., Hart, T.C., Michalec, M.D., Ryu, O.H., Simmons, D., Hong, S., and Wright, J.T. (2004). Mutation in kallikrein 4 causes autosomal recessive hypomaturation amelogenesis imperfecta. J. Med. Genet. *41*, 545–549.

Hartley, B.S. (1964). Amino acid sequence of bovine chymotrypsinogen-A. Nature *201*, 1284-1287.

Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem. Rev. *102*, 4501–4524.

Hedstrom, L., Szilagyi, L., and Rutter, W.J. (1992). Converting trypsin to chymotrypsin: the role of surface loops. Science *255*, 1249–1253.

Herter, S., Piper, D.E., Aaron, W., Gabriele, T., Cutler, G., Cao, P., Bhatt, A.S., Choe, Y., Craik, C.S., Walker, N., et al. (2005). Hepatocyte growth factor is a preferred *in vitro* substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. Biochem. J. *390*, 125–136.

Hooper, J.D., Bui, L.T., Rae, F.K., Harvey, T.J., Myers, S.A., Ashworth, L.K., and Clements, J.A. (2001). Identification and characterization of klk14, a novel kallikrein serine protease gene located on human chromosome 19q13.4 and expressed in prostate and skeletal muscle. Genomics *73*, 117–122.

Hu, J.C., Zhang, C., Sun, X., Yang, Y., Cao, X., Ryu, O., and Simmer, J.P. (2000). Characterization of the mouse and human PRSS17 genes, their relationship to other serine proteases, and the expression of PRSS17 in developing mouse incisors. Gene *251*, 1–8.

Iacobuzio-Donahue, C.A., Ashfaq, R., Maitra, A., Adsay, N.V., Shen-Ong, G.L., Berg, K., Hollingsworth, M.A., Cameron, J.L., Yeo, C.J., Kern, S.E., et al. (2003). Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies. Cancer Res. *63*, 8614–8622.

Kapadia, C., Ghosh, M.C., Grass, L., and Diamandis, E.P. (2004a). Human kallikrein 13 involvement in extracellular matrix degradation. Biochem. Biophys. Res. Commun. *323*, 1084–1090.

Kapadia, C., Yousef, G.M., Mellati, A.A., Magklara, A., Wasney, G.A., and Diamandis, E.P. (2004b). Complex formation between human kallikrein 13 and serum protease inhibitors. Clin. Chim. Acta *339*, 157–167.

Katz, B.A., Liu, B., Barnes, M., and Springman, E.B. (1998). Crystal structure of recombinant human tissue kallikrein at 2.0 Å resolution. Protein Sci. *7*, 875–885.

Kishi, T., Kato, M., Shimizu, T., Kato, K., Matsumoto, K., Yoshida, S., Shiosaka, S., and Hakoshima, T. (1999). Crystal structure of neuropsin, a hippocampal protease involved in kindling epileptogenesis. J. Biol. Chem. *274*, 4220–4224.

Komatsu, N., Saijoh, K., Toyama, T., Ohka, R., Otsuki, N., Hussack, G., Takehara, K., and Diamandis, E.P. (2005). Multiple tissue kallikrein mRNA and protein expression in normal skin and skin diseases. Br. J. Dermatol. *153*, 274–281.

Komatsu, N., Suga, Y., Saijoh, K., Liu, A.C., Khan, S., Mizuno, Y., Ikeda, S., Wu, H.K., Jayakumar, A., Clayman, G.L., et al. (2006). Elevated human tissue kallikrein levels in the stratum corneum and serum of peeling skin syndrome-type B patients suggests an over-desquamation of corneocytes. J. Invest. Dermatol. *126*, 2338–2342.

Lilja, H. (1985). A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. J. Clin. Invest. *76*, 1899–1903.

Lundwall, A., Band, V., Blaber, M., Clements, J.A., Courty, Y., Diamandis, E.P., Fritz, H., Lilja, H., Malm, J., Maltais, L.J., et al. (2006). A comprehensive nomenclature for serine proteases with homology to tissue kallikreins. Biol. Chem. *387*, 637–641.

Magklara, A., Mellati, A.A., Wasney, G.A., Little, S.P., Sotiropoulou, G., Becker, G.W., and Diamandis, E.P. (2003). Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors. Biochem. Biophys. Res. Commun. *307*, 948–955.

Malm, J., Hellman, J., Hogg, P., and Lilja, H. (2000). Enzymatic action of prostate-specific antigen (PSA or hK3): substrate specificity and regulation by Zn^{2+}, a tight-binding inhibitor. Prostate *45*, 132–139.

Matsumura, M., Bhatt, A.S., Andress, D., Clegg, N., Takayama, T.K., Craik, C.S., and Nelson, P.S. (2005). Substrates of the prostate-specific serine protease prostase/KLK4 defined by positional-scanning peptide libraries. Prostate *62*, 1–13.

Michael, I.P., Pampalakis, G., Mikolajczyk, S.D., Malm, J., Sotiropoulou, G., and Diamandis, E.P. (2006) Human tissue kallikrein 5 is a member of a proteolytic cascade pathway involved in seminal clot liquefaction and potentially in prostate cancer progression. J. Biol. Chem. *281*, 12743–12750.

Michael, I.P., Sotiropoulou, G., Pampalakis, G., Magklara, A., Ghosh, M., Wasney, G., and Diamandis, E.P. (2005). Biochemical and enzymatic characterization of human kallikrein 5 (hK5), a novel serine protease potentially involved in cancer progression. J. Biol. Chem. *280*, 14628–14635.

Nelson, P.S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999). Molecular cloning and characterization of prostase, an androgen-regulated serine protease with prostate-restricted expression. Proc. Natl. Acad. Sci. USA *96*, 3114–3119.

Obiezu, C.V., Shan, S.J., Soosaipillai, A., Luo, L.Y., Grass, L., Sotiropoulou, G., Petraki, C.D., Papanastasiou, P.A., Levesque, M.A., and Diamandis, E.P. (2005). Human kallikrein 4: quantitative study in tissues and evidence for its secretion into biological fluids. Clin. Chem. *51*, 1432–1442.

Obiezu, C.V., Michael, I.P., Levesque, M.A., and Diamandis, E.P. (2006). Human kallikrein 4: enzymatic activity, inhibition, and degradation of extracellular matrix proteins. Biol. Chem. *387*, 749–759.

Oikonomopoulou, K., Hansen, K.K., Saifeddine, M., Tea, I., Blaber, M., Blaber, S.I., Scarisbrick, I., Andrade-Gordon, P., Cottrell, G.S., Bunnett, N.W., et al. (2006). Proteinase-activated receptors, targets for kallikrein signaling. J. Biol. Chem. *281*, 32095–32112.

Perona, J.J. and Craik, C.S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Sci. *4*, 337–360.

Puente, X.S., Sanchez, L.M., Overall, C.M., and Lopez-Otin, C. (2003). Human and mouse proteases: a comparative genomic approach. Nat. Rev. Genet. *4*, 544–558.

Rawlings, N.D., Morton, F.R., and Barrett, A.J. (2006). MEROPS: the peptidase database. Nucleic Acids Res. *34*, D270–D272.

Raymond, W.W., Ruggles, S.W., Craik, C.S., and Caughey, G.H. (2003). Albumin is a substrate of human chymase. Prediction by combinatorial peptide screening and development of a selective inhibitor based on the albumin cleavage site. J. Biol. Chem. *278*, 34517–34524.

Sawasaki, T., Shigemasa, K., Gu, L., Beard, J.B., and O'Brien, T.J. (2004). The transmembrane protease serine (TMPRSS3/TADG-12) D variant: a potential candidate for diagnosis and therapeutic intervention in ovarian cancer. Tumour Biol. *25*, 141–148.

Schechter, I. and Berger, A. (1967). On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. *27*, 157–162.

Shipway, A., Danahay, H., Williams, J.A., Tully, D.C., Backes, B.J., and Harris, J.L. (2004). Biochemical characterization of prostasin, a channel activating protease. Biochem. Biophys. Res. Commun. *324*, 953–963.

Sotiropoulou, G., Rogakos, V., Tsetsenis, T., Pampalakis, G., Zafiropoulos, N., Simillides, G., Yiotakis, A., and Diamandis, E.P. (2003). Emerging interest in the kallikrein gene family for understanding and diagnosing cancer. Oncol. Res. *13*, 381–391.

Southan, C. (2001) A genomic perspective on human proteases as drug targets. Drug Discov. Today *6*, 681–688.

Stefansson, K., Brattsand, M., Ny, A., Glas, B., and Egelrud, T. (2006). Kallikrein-related peptidase 14 may be a major contributor to trypsin-like proteolytic activity in human stratum corneum. Biol. Chem. *387*, 761–768.

Stephenson, S.A., Verity, K., Ashworth, L.K., and Clements, J.A. (1999). Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3–13.4. J. Biol. Chem. *274*, 23210–23214.

Takayama, T.K., McMullen, B.A., Nelson, P.S., Matsumura, M., and Fujikawa, K. (2001). Characterization of hK4 (Prostase), a prostate-specific serine protease: activation of the precursor of prostate specific antigen (pro-PSA) and single-chain urokinase-type plasminogen activator and degradation of prostatic acid phosphatase. Biochemistry *40*, 15341–15348.

Takeuchi, T., Harris, J.L., Huang, W., Yan, K.W., Coughlin, S.R., and Craik, C.S. (2000). Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. J. Biol. Chem. *275*, 26333–26342.

Thornberry, N.A., Rano, T.A., Peterson, E.P., Rasper, D.M., Timkey, T., Garcia-Calvo, M., Houtzager, V.M., Nordstrom, P.A., Roy, S., Vaillancourt, J.P., et al. (1997). A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. J. Biol. Chem. *272*, 17907–17911.

Wright, J.T., Daly, B., Simmons, D., Hong, S., Hart, S.P., Hart, T.C., Atsawasuwan, P., and Yamauchi, M. (2006). Human enamel phenotype associated with amelogenesis imperfecta and a kallikrein-4 (g.2142G>A) proteinase mutation. Eur. J. Oral Sci. *114* (Suppl. 1), 13–17.

Yamasaki, K., Schauber, J., Coda, A., Lin, H., Dorschner, R.A., Schechter, N.M., Bonnart, C., Descargues, P., Hovnanian, A., and Gallo, R.L. (2006). Kallikrein-mediated proteolysis regulates the antimicrobial effects of cathelicidins in skin. FASEB J. *20*, 2068–2080.

Yousef, G.M. and Diamandis, E.P. (1999). The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation. J. Biol. Chem. *274*, 37511–37516.

Yousef, G.M. and Diamandis, E.P. (2001). The new human tissue kallikrein gene family: structure, function, and association to disease. Endocr. Rev. *22*, 184–204.

Yousef, G.M., Obiezu, C.V., Luo, L.Y., Black, M.H., and Diamandis, E.P. (1999). Prostase/KLK-L1 is a new member of the human kallikrein gene family, is expressed in prostate and breast tissues, and is hormonally regulated. Cancer Res. *59*, 4252–4256.

Yousef, G.M., Chang, A., and Diamandis, E.P. (2000a). Identification and characterization of KLK-L4, a new kallikrein-like gene that appears to be down-regulated in breast cancer tissues. J. Biol. Chem. *275*, 11891–11898.

Yousef, G.M., Chang, A., Scorilas, A., and Diamandis, E.P. (2000b). Genomic organization of the human kallikrein gene family on chromosome 19q13.3–q13.4. Biochem. Biophys. Res. Commun. *276*, 125–133.

Yousef, G.M., Magklara, A., Chang, A., Jung, K., Katsaros, D., and Diamandis, E.P. (2001). Cloning of a new member of the human kallikrein gene family, KLK14, which is down-regulated in different malignancies. Cancer Res. *61*, 3425–3431.

Zimmerman, M., Ashe, B., Yurewicz, E.C., and Patel, G. (1977). Sensitive assays for trypsin, elastase, and chymotrypsin using new fluorogenic substrates. Anal. Biochem. *78*, 47–51.

Received March 15, 2007; accepted June 15, 2007