Structure of plasma and tissue kallikreins

Monika Pathak; Szu Shen Wong; Ingrid Dreveny; Jonas Emsley
Centre for Biomolecular Sciences, School of Pharmacy, University of Nottingham, Nottingham, UK

Summary

The kallikrein kinin system (KKS) consists of serine proteases involved in the production of peptides called kinins, principally bradykinin and Lys-bradykinin (kallidin). The KKS contributes to a variety of physiological processes including inflammation, blood pressure control and coagulation. Here we review the protein structural data available for these serine proteases and examine the molecular mechanisms of zymogen activation and substrate recognition focusing on plasma kallikrein (PK) and tissue kallikrein (KLK1) cleavage of kininogens. PK circulates as a zymogen bound to high-molecular-weight kininogen (HK). PK is activated by coagulation factor XIIa and then cleaves HK to generate bradykinin and factor XI to generate further XIIa. A structure has been described for the activated PK protease domain in complex with the inhibitor benzamidine. Kallikrein-related peptidases (KLKs) have a distinct domain structure and exist as a family of 15 genes which are differentially expressed in many tissues and the central nervous system. They cleave a wide variety of substrates including low-molecular-weight kininogen (LK) and matrix proteins. Crystal structures are available for KLK1, 3, 4, 5, 6 and 7 activated protease domains typically in complex with S1 pocket inhibitors. A substrate mimetic complex is described for KLK3 which provides insight into substrate recognition. A zymogen crystal structure determined for KLK6 reveals a closed S1 pocket and a novel mechanism of zymogen activation. Overall these structures have proved highly informative in understanding the molecular mechanisms of the KKS and provide templates to design inhibitors for treatment of a variety of diseases.

Keywords
Plasma kallikrein, tissue kallikrein, kininogen, zymogen activation, substrate recognition

Introduction

The kallikrein enzymes are serine proteases divided into the two major categories of plasma kallikrein (PK) and kallikrein-related peptidases (KLKs) (1, 2). These two categories differ significantly in their location in plasma or the tissues, respectively. The molecular weight, substrate specificity, gene structure and kinin produced are also different. A central component of the kallikrein-kinin system (KKS) is the cleavage of kininogen to generate short vasoactive peptides termed kinins. Illustrated schematically in Figure 1 are the two pathways involving plasma and tissue kallikrein KLK1 leading to the generation of the principle peptide bradykinin (BK) (3, 4). The participation of KKS in inflammation, cancer, and in pathologies related to cardiovascular, renal and central nervous systems have been discussed in a recent review (5).

The two forms of kininogen which act as substrate for kallikreins are termed high-molecular-weight kininogen (HK) and low-molecular-weight kininogen (LK), respectively (4). These result from differential splicing of the kininogen gene (symbol KNG1) to generate proteins which differ at the C-terminus. In the mature polypeptide HK is 626 amino acids in length with six domains (D1-D6). Notably D2 interacts with endothelial cell receptor C1q, D3 interacts with platelets, D4 harbours the BK peptide and D5 contains a histidine-rich region that binds to negatively charged surfaces (6). The C-terminal D6 domain interacts with PK and coagulation factor XI (FXI) (6). The splice variant LK is 409 amino acids in length and the principal difference is the loss of D6. LK is a substrate for tissue kallikrein KLK1. Other members of the KLKs family cleave a variety of protein and peptide substrates which has been reviewed recently (7, 8). Notably, KLK3 or prostate specific antigen (PSA) cleaves semenogelin and fibronectin. This proteolytic action is involved in the liberation of sperm (7).

Plasma kallikrein (gene symbol KLKB1) becomes activated by coagulation factor XIIa (FXIIa) and cleaves HK to generate BK (4). Human tissue kallikrein (KLK1), unlike PK, cleaves LK to generate Lys-bradykinin (kallidin) that is subsequently converted to BK by a second aminopeptidase cleavage (4).

In addition to cleaving HK, PK is a central component of the contact system (9). Once activated, PK is capable of activating FXII and thereby amplifies the generation of FXIIa that reciprocally activates PK through a feedback system. Although PK and kininogen deficiency in humans does not manifest a clinical abnor-}

© Schattauer 2013

Thrombosis and Haemostasis 110.3/2013

For personal or educational use only. No other uses without permission. All rights reserved.
The activated form of PK is involved in blood pressure regulation through cleavage of HK and the release of the vasodilator BK (13). A gain of function mutation in FXII has been characterised in humans as hereditary angioedema (HAE type III) which manifests clinically as recurrent acute skin swelling, abdominal pain, and laryngeal oedema (14). PK is also activated on the surface of endothelial cells by a less well understood FXII independent process involving the enzyme prolylcarboxypeptidase (PRCP) (15).

The KLKs family of serine proteases is a large family of 15 genes located on the same chromosomal locus each with diverse functions and substrates. They possess at least 30% sequence identity and the mature proteins are approximately 230 amino acid residues in length. KLKs are synthesised as zymogens, secreted then proteolytically processed to generate the active forms by liberation of a hydrophobic residue (Ile,Leu) which inserts into the core of the protease domain. In the human KLKs family, KLK1 liberates kinin from kininogen efficiently and the other KLKs possess greatly reduced or no kininogen cleavage activity (7, 16). KLK1 has trypsin-like and chymotrypsin-like specificity (17).

The domain structure of PK and KLKs is shown with a typical serine protease catalytic domain at the C-terminus. The unusual feature of four repeats of the 90 residue apple domain which occurs at the N-terminus in PK has not been described in terms of a crystal structure. A model of full length PK based on the homologous factor XI zymogen structure reveals an assembled disc arrangement of the apple domains shown in Figure 2B (18). The available crystal structures for PK and KLKs are listed in Table 1. Here we describe these structures focusing on the understanding of zymogen activation, substrate recognition and the cleavage of kininogen.

### The KLK1 protease structure

The catalytic triad residues His57, Asp102, and Ser195 are conserved in all of the KLKs (classic chymotrypsin residue numbering will be used throughout) and structures are available for KLK1,3,4,5,6 and 7 (Table 1). Of the 24 crystal structures described, 18 are human proteins and the remaining six are from thrombus formation (10-12). The activated form of PK is involved in blood pressure regulation through cleavage of HK and the release of the vasodilator BK (13). A gain of function mutation in FXII has been characterised in humans as hereditary angioedema (HAE type III) which manifests clinically as recurrent acute skin swelling, abdominal pain, and laryngeal oedema (14). PK is also activated on the surface of endothelial cells by a less well understood FXII independent process involving the enzyme prolylcarboxypeptidase (PRCP) (15).

The KLKs family of serine proteases is a large family of 15 genes located on the same chromosomal locus each with diverse functions and substrates. They possess at least 30% sequence identity and the mature proteins are approximately 230 amino acid residues in length. KLKs are synthesised as zymogens, secreted then proteolytically processed to generate the active forms by liberation of a hydrophobic residue (Ile,Leu) which inserts into the core of the protease domain. In the human KLKs family, KLK1 liberates kinin from kininogen efficiently and the other KLKs possess greatly reduced or no kininogen cleavage activity (7, 16). KLK1 has trypsin-like and chymotrypsin-like specificity (17).

The domain structure of PK and KLKs is shown with a typical serine protease catalytic domain at the C-terminus. The unusual feature of four repeats of the 90 residue apple domain which occurs at the N-terminus in PK has not been described in terms of a crystal structure. A model of full length PK based on the homologous factor XI zymogen structure reveals an assembled disc arrangement of the apple domains shown in Figure 2B (18). The available crystal structures for PK and KLKs are listed in Table 1. Here we describe these structures focusing on the understanding of zymogen activation, substrate recognition and the cleavage of kininogen.

### The KLK1 protease structure

The catalytic triad residues His57, Asp102, and Ser195 are conserved in all of the KLKs (classic chymotrypsin residue numbering will be used throughout) and structures are available for KLK1,3,4,5,6 and 7 (Table 1). Of the 24 crystal structures described, 18 are human proteins and the remaining six are from...
Figure 2: PK and KLKs domain organisation and sequence. A) Plasma kallikrein illustrated at the top with four apple domains (A1-4) and three amino acid residues of the catalytic triad His57 (H), Asp102 (D), and Ser195 (S) indicated. Disulfide bridge (C-C bond, Cys364 and Cys484) connecting heavy chain and light chain of PK is also shown. In contrast, to PK, kallikrein-related peptidases are composed of only a protease domain with a pro-peptide (P). B) Cartoon diagram illustrating the three-dimensional domain architecture of PK as determined by homology modelling using the FXI structure. The protease domain is positioned on top of a disc shaped arrangement of the apple domains shown on the right rotated by 90 degrees. Each apple domain comprises of seven $\beta$-sheets (cyan) and one $\alpha$-helix (red). Active site pocket loops are shown in black and are likely to be disordered in the zymogen. C) Amino acid sequence alignment of plasma and KLK1,3,6. Catalytic site residues are indicated by an arrow. Residues from the kallikrein loop and Asn residues from N-linked glycosylation sites are boxed and shown in pink and yellow respectively. In PK, the S2 pocket residue Gly99 is boxed and indicated in red. In KLK3 the S1 pocket has a chymotryptic nature with Ser189 present in place of Asp189, as boxed and shown in green.
other species (▶Table 1). The majority have an inhibitor such as benzamidine bound in the S1 pocket (▶Figure 3). KLKs are both trypsin and chymotrypsin like in their substrate specificity range (7). Most, like KLK1, have a trypsin like Asp189 residue in the S1 pocket with the exception being KLK3 which has Ser189 and is chymotrypsin like (▶Figure 4A). KLK1 nevertheless exhibits both trypsin- and chymotrypsin-like specificities (19). The KLK1 cleavage of the physiological substrate LK releases the peptide Lys-bradykinin (kallidin), by processing at two pairs of amino acid residues (Met-Lys and Arg-Ser) (20). KLK1 displays trypsin-like specificity hydrolysing the Arg-Ser peptide bond at the C-terminal cleavage site in LK (17). However the N-terminal cleavage is uncharacteristic and has a Met sidechain at P1 (▶Figure 4A). The chymotrypsin-like activity is evident in the cleavage of kallistatin and somatostatin, after a Phe–Phe dipeptide (21). KLK3 has a chymotryptic specificity cleaving after a Tyr sidechain for substrate and free crystal structures of KLKs (19). There is no substrate mimetic complex structure for KLK1 showing how it would recognise the LK cleavage sites. However, a KLK3 structure is described in complex with a substrate mimetic peptide derived from the sequence of matrix protein Semenogelin I (▶Figure 3C) (22). This latter structure is highly informative illustrating the P1 Tyrosine sidechain inserted into the chymotryptic S1 pocket and shows how the kallikrein loop closes over the top of the N-terminal regions of the peptide substrate to bury residues at P2, P3, P4 which are in a β-strand conformation.

Table 1: Plasma and tissue kallikreins: Crystal structure PDB codes, gene and protein nomenclature.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Protein name</th>
<th>Other protein names/symbols</th>
<th>PDB code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Human plasma kallikrein</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLKB1</td>
<td>PK</td>
<td>Human plasma kallikrein</td>
<td>2ANW, 2ANY</td>
<td>Activated protease in complex with benzamidine (glycosylated and deglycosylated forms)</td>
</tr>
<tr>
<td><em>Human Tissue/Glandular Kallikrein (Kallikrein-Related Peptidases)</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLK1</td>
<td>hk1</td>
<td>Tissue kallikrein, pancreatic/renal kallikrein, hPRK</td>
<td>1SPJ</td>
<td>Structure of activated human tissue kallikrein protease domain with vacant active site</td>
</tr>
<tr>
<td>KLK3</td>
<td>hk3</td>
<td>Prostate specific antigen, PSA</td>
<td>22CL, 2ZCK, 2ZCH</td>
<td>Crystal structure of human prostate specific antigen (PSA) complexed with an antibody and a peptide substrate mimetic.</td>
</tr>
<tr>
<td>KLK4</td>
<td>hk4</td>
<td>Prostase, KLK-L1 protein, EMSP1</td>
<td>2BDI, 2BDH, 2BDG</td>
<td>Human kallikrein 4 complex with zinc, cobalt ions and benzamidine</td>
</tr>
<tr>
<td>KLK5</td>
<td>hk5</td>
<td>KLK-L2 protein; HSCTE</td>
<td>2PSY, 2PSX</td>
<td>Crystal structure of human kallikrein 5 in complex with zinc ions and leupeptin</td>
</tr>
<tr>
<td>KLK6</td>
<td>hk6</td>
<td>Zyme, Protease M, Neurosin</td>
<td>1L06, 1L2E, 1GVL</td>
<td>Human kallikrein 6 (hk6) active Ffrm in complex with benzamidine inhibitor, Human prokallikrein 6 (hk6)/ prozyme/ proprotease m/ prostatein, zymogen</td>
</tr>
<tr>
<td>KLK7</td>
<td>hk7</td>
<td>HSCCE</td>
<td>2QXG, 2QXJ, 2QXI, 2QXH, 3BSQ</td>
<td>Crystal structures of human kallikrein 7 in the presence inhibitors; The 3BSQ structure is an apo form of KLK7.</td>
</tr>
</tbody>
</table>

**Tissue kallikrein structures from different species**

- **Toni**n
  Rat submaxillary gland serine protease, tonin
  1TON
  Activated protease domain in complex with zinc ions

- **Kallikrein A**
  Porcine Pancreatic Kallikrein A
  2PKA
  Porcine pancreatic kallikrein A activated protease structure Complex with BPTI

- **Kallikrein**
  Porcine kallikrein
  1HIA
  Porcine kallikrein complexed with inhibitor hirustasin

- **Kallikrein-13**
  Mouse kallikrein-13 (prorenin converting enzyme)
  1A05
  Activated protease domain

- **Horse PSA**
  Prostate specific antigen
  1GVZ
  Activated protease from stallion seminal plasma
Pathak et al. Structure of kallikreins

KLK1 and PK substrate specificity and inhibition

Enzyme kinetic data using peptide mimics of the cleavage sites in kininogen have been utilised to characterise the determinants of KLK1 substrate specificity. Li et al. analysed KLK1 substrate specificity using phage display and a large random octapeptide library (17). This screen identified peptides from both the chymotryptic and trypsin-like specificity with Tyr and Arg preferred at P1. This is illustrated as a consensus sequence in ▶ Figure 4B calculated using across 12 chymotryptic peptides and three trypsin-like peptides where the taller characters represent more abundant amino acids (23). Consistent with the C-terminal cleavage site in kininogen the tryptic KLK1 preference for Arg at P1 and Ser at P1’ was observed in all three peptides identified. By contrast, the 12 chymotryptic peptides identified did not resemble the N-terminal Met-Lys cleavage site but instead were typical of other KLKs with Tyr the most common amino acid at P1 (compare ▶ Figure 4B with the cleavage site for KLK3 in substrate semenogelin I shown in ▶ Figure 4A). This data also characterised that Ser/Arg are pre-
Kallikreins: Pathobiology and association to disease

A

Kininogen

Gly-Met-Ile-Ser-Leu-Met-Lys-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg-Ser-Ser-Arg-Ile-Gly

PK

Factor XII

Ser-Leu-Ser-Ser-Met-Thr-Arg-Val-Val-Gly-Gly

Semenogelin I

Lys-Gly-Ile-Ser-Ser-Gln-Tyr-Ser-Asn-Thr-Gln

B

C

KLK1

PK

Abz-Gly-Phe-Ser-Pro-Phe-Arg-Ser-Ser-Arg-Ile-QEDnpp

0.2

17.7

88500

2.2

0.45

205

(24)

Abz-Gly-Phe-Ser-Pro-Phe-Arg-Ser-Leu-Arg-Ile-QEDnpp

0.36

4.6

12778

1.5

0.6

333

(24)

Abz-Phe-Arg-Ser-Ser-Arg-EDCnpp

1.5

1.6

1067

(25)

Abz-Met-Ile-Ser-Leu-Met-Lys-Arg-Pro-EDCnpp

0.7

0.33

471

(25)

Abz-Met-Ile-Ser-Leu-Met-Lys-Leu-Pro-EDCnpp

9.0

0.02

2

(25)
ferred at P1’ which has been observed for KLK1, KLK3 and KLK6 (19). Further away from the P1 site preferences are more consistent with the N-terminal cleavage site and Arg is identified in peptides at P2’ (19) and hydrophobic residue at P2 (Phe, Leu) as illustrated for LK cleavage sites (Leu-Met-Lys and Phe-Arg-Ser, Figure 4A). Further preference for small sidechains such as Ala/Thr/Ser at position P3 and P4 of the chymotryptic cleavage sites is consistent with this region of the substrate being buried by interactions with the kallikrein loop in a fashion similar to that observed in the KLK3-peptide complex structure (Figure 3C). This data shows cleavage of the Arg-Ser bond by KLK1 is defined by the P1-P1’ residues; however, the atypical Met-Lys bond cleavage is driven largely by subsites remote from P1-P1’.

Lima et al. compared KLK1 and PK enzyme kinetics of cleavage for a 10-mer synthetic peptide series and carried out a time course measuring the hydrolysis of a 19-mer peptide derived from kinogen which spans both cleavage sites (24). These data demonstrated the requirement of Ser at P1’ as critical for both PK and KLK1 cleavage with only very low hydrolysis measured for substitutions into this position for PK and reduced KLK1 hydrolysis for all peptides with substitutions at P1’. Cleavage of the native 10-mer peptide substrates revealed a kcat/Km which was 500 times lower for PK than KLK1 (Figure 4C) whereas the longer peptide spanning both cleavage sites was equally efficient showing PK likely requires interactions distant from the active site pockets (24). A substitution at P2’ affects a marked reduction of the KLK1 kcat/Km whereas values for PK are increased (Figure 4C).

Chagas et al. characterised KLK1 cleavage of a series of peptides examining the Met-Lys and Arg-Ser proteolysis in parallel and derived kinetic parameters of kcat, kcat/Km and km/kcat (Figure 4C) (25). This evaluated that substitutions at P2’ in most cases resulted in no hydrolysis measured or a large reduction in the rate of cleavage for the substitution of Arg for Leu as illustrated in Figure 4C. This study also revealed a role for Pro at P3’. This data also illustrated that the kinetics of cleavage were most optimal for peptides with Phe at P1 (17).

The structures of inhibitor complexes of KLKs have been summarised in a recent review (26). Interesting features of the porcine KLKA complexes with BPTI (Kunitz-type) and hirustasin (antistasin-type inhibitor) are that both insert an Arg sidechain into the S2’ pocket (27, 28). This occupancy of S2’ by the Arg sidechain is almost identical for both inhibitors. Kallistatin is a serpin with a unique P1 Phe, which has inhibitory specificity towards KLK1 (29). An investigation of the P1-P3 residues of human kallistatin by site-directed mutagenesis revealed the P1 Phe displays a better selectivity for KLK1 than P1 Arg. For the P2 and P3 variants, the mutants with hydrophobic and bulky amino acids at P2 and basic amino acids at P3 displayed better binding activity with KLK1. Molecular modelling of the kallistatin P2 Phe sidechain predicted that this residue can occupy the KLK1 S34 pocket (29).

A model of the KLK1-kininogen interaction

The crystal structures of KLKs complexes and the functional data described above define three distinct regions in the protease domain important to recognition of substrate and inhibitors; (i) the S1 pocket (ii) kallikrein loop and S2-S4 pockets (iii) S1-2’ pockets. There are no substrate mimetic complex crystal structures for KLK1 to describe precisely how it utilises these three regions however the structure of the ligand free KLK1, KLK3-peptide complex and the KLKA inhibitor complexes with hirustatin and BPTI provide templates to model the KLK1 interactions with kinogen.

Molecular modelling was carried out with COOT in three stages (30). As the kinogen sequence and the data from the KLK1 phage screening fits with a similar pattern of interactions described for KLK3 and semenogelin I this was used as a template to model the Met-Lys cleavage site peptide interaction. Thus in stage 1 the kinogen MISLMKR peptide sequence was constructed using the COOT mutate function based on the coordinates of the semenogelin I peptide from the KLK3 complex (2ZCL). In stage 2 this was then superimposed onto the KLK1 ligand-free structure (1SPJ) utilising the close similarity with the KLK3 protease structure as a template (COOT structure alignment function). The kallikrein loop of KLK1 has a number of non-conservative substitutions in the sequence compared to KLK3 and this precluded accurate modelling of this structure (Figure 2B). The loop is partially disordered in the ligand free KLK1 structure and this region is hence illustrated as a dashed line coloured pink in Figure 4D. There is evidence that Arg at P2’ is important for the Met-Lys cleavage as kinetic studies reveal a substition at this position reduces cleavage (Figure 4C). Two crystal structures of KLKA inhibitor complexes reveal a fit of the Arg sidechain into the S2’ pocket which is formed by the Phe40 and Phe151 residues which are conserved in human KLK1. In stage 3 these structures (1HIA, 2KAI) were used as a template to model the KNG1 P1’-P2 residues Lys-Arg into the KLK1 S1-2’ pockets. The S2’ pocket residues are illustrated in red in Figure 4D. Overall this complex is a good fit and the only steric clash is with the P2 Leu sidechain and Tyr99 in the S2 pocket. The Tyr99 sidechain was thus adjusted to match the sidechain rotamer observed in the ligand bound KLKA inhibitor complex structures. The hirustatin structure complexed with KLKA also has a similar conformation of the P2-P4
residues with the P4 residue Val sidechain occupying the S4 pocket as shown for the kininogen P4 residue Ile in Figure 4D.

The PFRSS peptide of kininogen was modelled using a similar procedure. This peptide is typical for a tryptic cleavage whereby the Arg sidechain inserts into the S1 pocket and forms a salt bridge to Asp189. Thus the Arg sidechain from the KLKA-hirustatin complex can be used as a template to model the kininogen P1 Arg. The interesting feature of the P2-P3 sequence of kininogen is the presence of Pro at P3. This residue will not allow the polypeptide to adopt the \( \beta \)-strand extended conformation which is observed in the semenogelin I peptide (bound to KLK3) and in the hirustatin structure complexed with KLKA. Hence BPTI was utilised as a template to model these residues as it also has Pro at P3 and the Cys sidechain at P2 allowing modelling of the kininogen Phe into the S2 pocket. There are no clear templates available to model the KLK1 interactions with the kininogen Ser-Ser residues at P1-2'. The requirement for Ser at P1' implies a hydrogen bond is formed with the sidechain OH group and this is modelled interacting with Gln41.

The two models presented are similar to the docking calculations previously reported for the KLK1 complexes with tetrapeptides ARSA and FYSQ (17) and the P2 Phe and Leu residue interactions with S2 are the same as those modelled in the study on kalilistatin interactions with KLK1 (29). Overall these models provide a useful synthesis of information from previous structural and functional studies and provide a framework for further experimentation.

The PFRSS peptide of kininogen was modelled using a similar procedure. This peptide is typical for a tryptic cleavage whereby the Arg sidechain inserts into the S1 pocket and forms a salt bridge to Asp189. Thus the Arg sidechain from the KLKA-hirustatin complex can be used as a template to model the kininogen P1 Arg. The interesting feature of the P2-P3 sequence of kininogen is the presence of Pro at P3. This residue will not allow the polypeptide to adopt the \( \beta \)-strand extended conformation which is observed in the semenogelin I peptide (bound to KLK3) and in the hirustatin structure complexed with KLKA. Hence BPTI was utilised as a template to model these residues as it also has Pro at P3 and the Cys sidechain at P2 allowing modelling of the kininogen Phe into the S2 pocket. There are no clear templates available to model the KLK1 interactions with the kininogen Ser-Ser residues at P1-2'. The requirement for Ser at P1' implies a hydrogen bond is formed with the sidechain OH group and this is modelled interacting with Gln41.

The two models presented are similar to the docking calculations previously reported for the KLK1 complexes with tetrapeptides ARSA and FYSQ (17) and the P2 Phe and Leu residue interactions with S2 are the same as those modelled in the study on kalilistatin interactions with KLK1 (29). Overall these models provide a useful synthesis of information from previous structural and functional studies and provide a framework for further experimentation.

Figure 5: Crystal structure and model of PK. A) Cartoon diagram of the activated PK crystal structure (PDB:2ANY). Shown as sticks are the side chains of active site residues His57, Asp102, Ser195 and S1 pocket residue Asp189 interacting with benzamidine (green). B) Close up view of the active site with Gly99 shown in the S2 pocket which is postulated to accommodate a bulky hydrophobic residue from the (Met, Phe) P2 position of kininogen. C) Molecular surface of activated PK depicted with the inhibitor benzamidine (shown as stick in green) bound to sub-site S1 (cyan). Different sub-sites S2-S2' are also marked. D) Cartoon diagram of the homology model of the PK zymogen protease domain. Loops in black are flexible. E) Overlay of PK zymogen model (blue) and activated PK (lince green) structure with the activation loop connecting to the apple domains is shown. Inset is the activated protease domain structure interaction between the positive charge of the N-terminus from residue Ile16 and the side chain of Asp194. The side chain of Asp194 forms an internal salt bridge with the N-terminus of the protease.
KLK1 zymogen activation

KLKs are secreted as zymogens which are subsequently activated by proteolysis. There is no structure for the zymogen of KLK1 but a study of zymogen activation has been described for KLK6 where structures are available for the zymogen and the activated form (Table 1). The KLK6 structure revealed a highly unusual zymogen structure when compared to well studies serine protease zymogens such as trypsin, chymotrypsin (31) and thrombin (32). This well studied mechanism has been described as “molecular sexuality” whereby the activating cleavage results in the terminal hydrophobic residue Ile16 from the activation loop inserting into the hydrophobic core (33) (shown as Ile16 IN and OUT in Figure 4D, E). In the example of trypsin the catalytic triad residues retain an identical conformation in both, the zymogen and the activated protease and zymogenicity is maintained by the substrate binding pocket loops being disordered and the absence of an oxyanion hole (33). The N-terminal residue Ile16 forms a salt bridge with Asp194 inducing conformational changes in this loop to form the oxyanion hole required for stabilising the proteolytic transition state intermediate. In addition, these loops form the S1 pocket and present Asp189 into a position favourable for interacting with the P1 Arg sidechain of the substrate. Although overall this mechanism is also observed in KLK6 two clear differences are described; (i) the KLK6 zymogen loops are well defined in a conformation that occludes the S1 pocket, (ii) the active site Ser195 is out of position in the zymogen and the sidechain points away from the Asp102 and His57 side chains of the catalytic triad. Upon activation Ser195 undergoes a conformational change and re-orients to form the typical triad. The Asp189 residue forms the base of the S1 pocket in a manner similar to trypsin although the conformational change is again distinct in KLK6 (31).

PK active site structure and substrate specificity

Two crystal structures are available for the activated protease domain of PK (34). Both describe a complex with benzamidine bound in the S1 pocket for de-glycosylated and glycosylated forms of the protein respectively. Figure 5A illustrates the topology of this structure showing the characteristic two β-barrels with catalytic triad indicated. PK maintains its trypsin like activity in the cleavage of HK (Lys-Arg, Arg-Ser) and also cleaves FXII (Arg-Val), in each case with a basic residue at the P1 position. The structure defines distinctive features around the active site and the S2 pocket is deeper when compared to trypsin and thrombin (Figure 5B, C) (34). In trypsin a preference exists for Pro at P2 whereas the bulky hydrophobic side chain of Phe or Met at P2 in kininogen can be modeled in the deeper S2 pocket of PK (34). The extent of the substrate binding pockets is shown in the PK protease domain surface representation in Figure 5C. Peptides with Pro at P2 are poor substrates for PK although the FXII cleavage site has Thr at P2. Figure 4 shows the presence of Ser at the P4 position in all three cleavage sites of PK which may indicate further interactions at S4 and a β-strand conformation for P3-P5 as observed in the semenogelin I peptide KLK3 complex structure (Figure 4C) (22).

PK has a well characterised binding site for the HK D6 domain within the apple 2 domain which likely provides selectivity for binding HK over LK (18, 35-37). The interaction with HK is thought to localise PK to the correct membrane and serve as an exosite interaction to facilitate efficient HK cleavage. The interaction of PK with HK has been studied in detail and separate methods have localised the interaction to an interface formed between the C-terminal region of the D6 domain of HK and the apple 2 domain of PK. The HK peptide containing the key determinants of PK binding spans ~30 amino acids (residues 565-595) close to the C-terminus (38, 39).

PK zymogen activation

PK is secreted as a zymogen which only becomes proteolytically active when cleaved at the Arg371-Ile372 bond (40). FXIIa is the best characterised physiological activator of PK. Small quantities of FXIIa are generated by a slow autoactivation of FXII in the presence of negatively charged surfaces (41). The tandem repeat of four apple domains is unique to PK and the homologous protein FXI which shares 58% amino acid sequence identity (40). There is no crystal structure for full-length PK; however, a crystal structure is available for FXI (42). This allowed homology modelling of the PK structure showing the overall inter-domain architecture is preserved with a disc-like arrangement of the apple domains positioned underneath the protease domain by specific interfaces as illustrated in Figure 2B and C (18). This model predicts a globular architecture unlike the elongated structures defined for the vitronectin-K-dependent coagulation factors which is confirmed in images of the PK zymogen obtained by electron microscopy (43). The model of PK shows a short activation loop is held close to the body of the protease domain by the domain interfaces formed with the apple domains (18). In the FXI zymogen crystal structure the loops responsible for forming the S1 pocket (including Asp189) are not observed and are assumed to be flexible (42). PK may have a more typical zymogen activation mechanism whereby residue Ile16 on the outside of the protease domain inserts into the core to form a salt bridge with Asp194 which precipitates further conformational changes resulting in the formation of the S1 pocket and oxyanion hole (Figure 5D, E).

Prolyl carboxypeptidase (PRCP) was also shown to facilitate PK activation on the surface of endothelial cells in a process that is independent of FXII (44). Whereas FXIIa activates both PK and FXI, PRCP is a selective PK activator (44). An inhibitor of PRCP was shown to block BK generation (45). Examination of the PK model revealed that a key difference between the FXI and PK structures is located at the C-terminus with the addition of an extra 10 residues in PK where there is a proline residue at the -1 position (18). This sequence is a potential substrate for PRCP within PK and then a second cleavage of Arg371-Ile372 could take place by PK autoactivation.
Summarisation

Structural information on kallikreins has provided the molecular detail to understand protease substrate recognition,zymogen activation and provided templates for structure based drug design and development (26). Structural studies are more advanced for KLKs than PK largely as they have been seen as important targets for developing new cancer treatments. In total, 22 crystal structures exist for KLKs whereas only two have been determined for PK. A uniquezymogen activation mechanism has been described for KLK6 which may also occur in other KLKs including KLK1. The KLK1 cleavage of the kininogen Met-Lys sequence is atypical and a variety of data indicated this cleavage is driven by interactions from amino acids at P2-P5 and P2’ rather than the P1 residue interaction with the S1 pocket. The phage screen against KLK1 did not detect Met at the S1 position or Lys at the S2 position (►Figure 4B).

To date only the activated protease domain structure of PK has been reported. A kininogen like substrate mimetic complex with the PK protease would help the understanding of how the active site utilises the pockets surrounding the catalytic triad. Both the kininogen cleavage sites (Leu-Lys-Arg, Phe-Arg-Ser) have a large hydrophobic residue at P2 which has been modelled into the PK S2 pocket (34). It is interesting to note that the only common residue present across all three PK cleavage sequences is the Ser residue at P4 (►Figure 4A).

Electron microscopy images of the full-length complex formed between PK and HK revealed a model for how the PK apple domains and protease domain interact with D6 and D4 domains of HK, respectively (43). A full-length zymogen structure for PK is needed to confirm that the domain arrangement is the same as FXI and this may help understand the basis of zymogen activation and recognition of HK (46). FXI is the homologue of PK and also binds HK through an interaction with the apple domains. It is interesting to note the FXIa cleavage of substrate FIX is also a double cleavage which releases a peptide from the FIX activation loop. The kinetics of this cleavage are that no intermediate is released and thus FXIa remains bound to FIX during the cleavage (47).

The FXI zymogen does not bind its substrate coagulation FIX until activated to FXIa which a clear contrast to the PK zymogen which circulates in complex with substrate HK. Electron microscopy of FXI and FIXa has described a large conformational change upon FXI zymogen activation involving re-arrangements of the apple and protease domain (48). Thus cleavage of the PK activation loop may precipitate conformational changes that extend beyond the protease domain to include larger scale domain motions including the apple 2 domain which may precipitate a more favourable presentation of the HK D4 domain cleavage sites towards the PK active site.

PK and HK have recently been explored as emerging targets for development of anti-thrombotics (10, 49) and genome wide studies in humans have found polymorphisms in PK and HK linked to thrombosis (50, 51). The KLKs have been implicated in the pathogenesis of cancer,neurodegeneration,inflammation, and asthma (52). Orally bioavailable small molecule inhibitors targeting coagulation proteases thrombin and FXa have recently come into use for the treatment of cardiovascular disease, which opens up the possibility that other serine protease inhibitors could be developed targeting kallikreins (53). Crystal structures will continue to provide important templates not only to understand the molecular basis of the KKS but also as a tool for medicinal chemists to design and develop new inhibitors.

Acknowledgements

We would like to thank the British Heart Foundation project grant (PG/09/025/27136) and program grant (RG/07/002/23) (MPIDJE) and the Danish Agency for Science Technology and Innovation together with Novonordisk for funding (SW).

Conflicts of interest

None declared.

References