Short Communication

Factor Xa subsite mapping by proteome-derived peptide libraries improved using WebPICS, a resource for proteomic identification of cleavage sites

Oliver Schilling[1],*, Ulrich auf dem Keller[2] and Christopher M. Overall[3,4]

[1] Institute for Molecular Medicine and Cell Research, University of Freiburg, Germany

[2] Institute of Cell Biology, Swiss Federal Institute of Technology, Zürich, Switzerland

[3] Department of Biochemistry and Molecular Biology, Centre for Blood Research, University of British Columbia, Vancouver, Canada

[4] Department of Oral Biological and Medical Sciences, Centre for Blood Research, University of British Columbia, Vancouver, Canada

* Corresponding author
  e-mail: oliver.schilling@mol-med.uni-freiburg.de

Abstract

Proteomic identification of protease cleavage site specificity (PICS) is a recent proteomic approach for the easy mapping of protease subsite preferences that determines both the prime- and non-prime side specificity concurrently. Here we greatly facilitate user access by providing an automated and simple web-based data-analysis resource termed WebPics (http://clipserve.clip.ubc.ca/pics/). We demonstrate the utility of WebPics analysis of PICS data by determining the substrate specificity of factor Xa from P6-P6', an important blood coagulation protease that proteolytically generates thrombin from prothrombin. PICS confirms existing data on non-prime site specificity and refines our knowledge of factor Xa prime-site selectivity.

Keywords: peptide library; prime-site specificity; protease specificity.

The intrinsic and extrinsic pathways of blood coagulation merge at the factor Xa-mediated activation of prothrombin. *In vivo*, prothrombin activation involves formation of a membrane surface complex between factor X and factor Va (Krishnaswamy et al., 1988). Hence, factor Xa, a very selective secreted S1 family serine protease, executes an essential role in hemostasis. As such, factor Xa is directly targeted by anticoagulant drugs such as rivaroxaban, apixaban or edoxaban (Harbrecht, 2011). In addition to prothrombin, factor Xa proteolytically activates factors V, VII, and VIII (Bajaj et al., 1981; Foster et al., 1983; Parker et al., 1997) as well as

protease-activated receptor 2 (Camerer et al., 2000), but little else.

Factor Xa is highly specific and indeed is frequently used for removal of tags from recombinant fusion proteins (Nagai et al., 1985). Across species, the physiological cleavage sites of factor Xa in prothrombin, factor VII and factor V share a P1 arginine and a P2 glycine residue (Brandstetter et al., 1996). This specificity profile has been supported by biochemical and genetic profiling strategies. Thus, at S1 factor Xa is considered to be specific for arginine (Matthews and Wells, 1993; Scholle et al., 2006) although one study employing synthetic peptide libraries also revealed specificity profiles having a P1 lysine (Gosalia et al., 2005). S2 also has a strong contribution to factor Xa specificity. In addition to the canonical specificity for glycine (Cho et al., 1984; Kawabata et al., 1988; Matthews and Wells, 1993; Scholle et al., 2006), there is increasing evidence for a partial preference for aromatic and aliphatic residues in S2, in particular phenylalanine (Harris et al., 2000; Bianchini et al., 2002; Furlong et al., 2002; Gosalia et al., 2005; Hsu et al., 2008).

Further removed from the scissile bond and based on several physiological cleavage sites, factor Xa exhibits a S3 preference for acidic residues. As well, this has been corroborated with synthetic peptide libraries (Cho et al., 1984; Kawabata et al., 1988). However, other studies indicated that negatively charged residues are disfavored in S3 (Bianchini et al., 2002) with P3 arginine being partially preferred in two substrate phage display screens (Scholle et al., 2006; Hsu et al., 2008) as well as in a synthetic peptide library screen (Furlong et al., 2002). The latter study also indicated a second, partial S3 preference for aromatic and aliphatic residues including phenylalanine, tyrosine, and leucine. At S4, a combinatorial peptide library determined a minor specificity for isoleucine, arginine or proline (Furlong et al., 2002); however, this was not confirmed by a second study that did not reveal P4 specificity determinants (Harris et al., 2000).

Profiling of prime-site specificity yields contrasting results. While one study suggests an equal contribution of S2' and S1' to overall factor Xa specificity (Bianchini et al., 2002), the results of another study point towards a limited influence of prime-site selectivity (Ludeman et al., 2003). S1' is described as preferring small or hydrophilic residues (serine, alanine, threonine) (Bianchini et al., 2002). However, a second specificity screen suggests a mixed preference for alanine alongside aromatic or aliphatic residues (Ludeman et al., 2003). In S2', both studies determine partial selectivity for glycine and alanine together with a minor preference for

aromatic or aliphatic residues (Bianchini et al., 2002; Ludeman et al., 2003). For S3', no preference has emerged (Bianchini et al., 2002; Ludeman et al., 2003). It is not clear whether it is the different techniques or inter-laboratory variability that has led to these differences in the profiles for prime and non-prime site specificity of factor Xa. In view of the medical importance of factor Xa and it being such an important drug target, the specificity of this enzyme needs to be completely understood.

A recent novel proteomic approach for determining both prime and nonprime side specificity in the same experiment termed 'proteomic identification of protease cleavage site specificity' (PICS) (Schilling and Overall, 2008) is designed for rapid protease selectivity determination. PICS focuses on biologically relevant sequences by employing proteome-derived peptide libraries that consequently are database searchable. In contrast to substrate phage display and fluorescence resonance energy transfer based synthetic peptide libraries, PICS unambiguously locates the scissile peptide bond. PICS determines individual prime-side cleavage sequences by liquid chromatography (LC)-tandem mass spectrometry (MS/MS) with the corresponding non-prime sequences being derived using bioinformatics. The large cohort of cleavage sequences identified in each PICS experiment allows for the analysis of potential subsite cooperativity as has been demonstrated for HIV protease 1 (Ridky et al., 1996; Schilling and Overall, 2008).

PICS libraries are first prepared from cellular proteomes by digesting first with specific endopeptidases, such as trypsin, chymotrypsin or Glu-C. This protease is then inactivated. Then, thiol and amine groups (N-terminal α-amines and lysine ε-amines) are chemically protected. The peptide mixture is then cleaned up and the proteome-derived peptide library is purified. Aliquots of the library are then used as substrate for the test protease in the PICS assay. Unlike the non-cleaved peptides in the library, which have blocked primary amines, the cleaved prime-side cleavage products display a reactive N-terminal primary amine on their neo-N-termini, allowing for specific biotinylation by NHS chemistry, affinity isolation, and sequence identification by LC-MS/MS. The corresponding non-prime sequences are determined using bioinformatics.

While PICS employs standard reagents without the need for peptide synthesis, its broad usability was hampered by a multi-step protocol for data analysis that requires expert knowledge for command line utilities and spreadsheet applications. This has now been addressed by a web-based integrated series of data handling scripts behind a convenient and easy user interface termed WebPICS (Schilling et al., 2011). The resource is freely accessible as a web-based application (http://clipserve.clip.ubc.ca/pics/) and functions with all major web browsers. WebPICS greatly facilitates data processing and information access by both occasional and expert users. In WebPICS, positional amino acid frequencies ranging from P6–P6' are calculated from the list of reconstructed cleavage sequences and presented as heatmaps with customizable cutoffs for minimal and maximal frequency values. Heatmaps are shown for relative positional frequencies of individual amino acids (in %) and positional frequencies normalized to the natural abundance of each amino acid. Absolute, relative, and normalized positional occurrences are also displayed in tables. For the generation of sequence logos and the highly useful iceLogos (Colaert et al., 2009), cleavage sequences are accessible in a format that allows for direct interaction with the iceLogo website via 'copy and paste'. Here, we report the application of WebPICS. Because factor Xa subsite specificity is well characterized, yet also somewhat controversial, we profiled these important proteases by PICS with data analysis handled by WebPICS as experimental proof of concept.

PICS libraries were derived from cell lysates of cultured human embryonic kidney cells and generated by digestion with either chymotrypsin (cleavage C-terminal to aromatic residues as well as leucine and methionine) or Glu-C (cleavage C-terminal to glutamic acid and, to a lesser extent, aspartic acid). Tryptic PICS libraries are not suited for factor Xa subsite mapping because trypsin cleaves C-terminal to lysine or arginine. Hence, tryptic PICS libraries lack internal basic residues and are incompatible with the primary specificity of factor Xa for P1 arginine.

For specificity screening, we used factor Xa preparations that are supplied for the untagging of recombinant fusion proteins (2 U/μl, Novagen). Based on vendor's recommendation, 2 U factor Xa were incubated with 200 μg PICS library for 16 h at 37°C at pH 8.0. Prime side cleavage product isolation and LC-MS/MS analysis using a QSTAR Pulsar tandem mass spectrometer (Applied Biosystems, Foster City, CA, USA) coupled to an LC Packings capillary LC system (Dionex, Sunnyvale, CA, USA) were performed exactly as reported in our recent step-by-step protocol for PICS (Schilling and Overall, 2008; Schilling et al., 2011). Details on spectrum-to-peptide sequence assignment are provided in the legend to Figure 1. Non-enzyme constraint spectrum-to-sequence assignment by X! Tandem, in conjunction with PeptideProphet (Keller et al., 2002; Craig and Beavis, 2004) for statistical modeling of the false discovery rates identified 90 non-redundant sequences from the chymotryptic PICS library and 34 non-redundant sequences from the Glu-C PICS library (Supplementary Tables 1 and 2). In the chymotryptic assay, a small number of cleavage sites have P1 residues which correspond to a chymotryptic cleavage site. These cases might be related to incomplete amine protection during library preparation. Hence, prime-side cleavage products with chymotryptic amino-termini are omitted from the PICS analysis as well as prime-side cleavage products with amino termini that correspond to a protein amino terminus. This exclusion strategy yields 86 non-redundant factor Xa cleavage sequences for the chymotryptic PICS library.

The factor Xa specificity profiles for the chymotryptic and Glu-C PICS libraries are shown in Figure 1 in heatmap-style representation and as sequence logos (Colaert et al., 2009). PICS employs proteome-derived peptide libraries with natural amino acid abundances. To correct for this bias, positional amino acid occurrences are shown in relation to their natural abundance. In order to discern subsite specificity from random fluctuations of positional frequencies, we focus
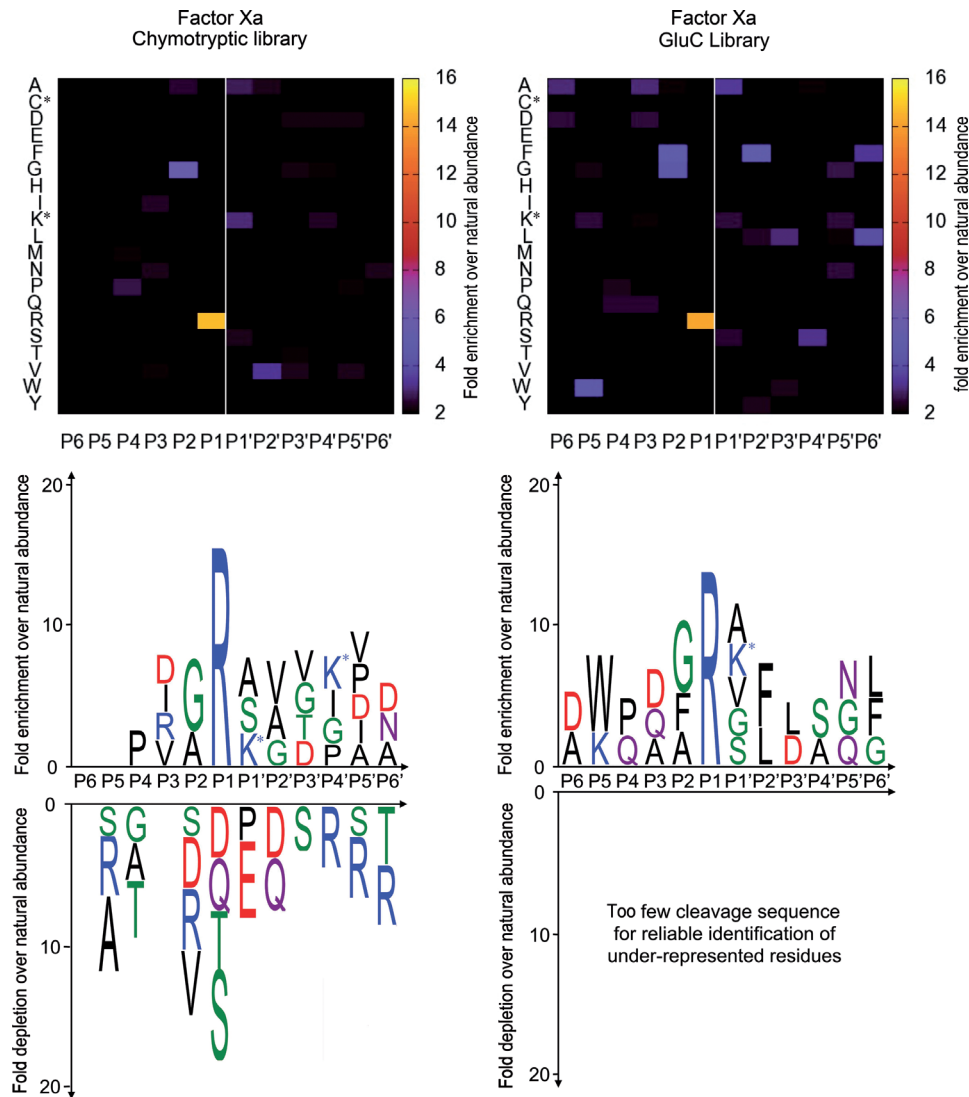
Figure 1 PICS profiling of factor Xa with a chymotryptic and a Glu-C peptide library.
Chymotryptic and Glu-C PICS libraries were prepared as described previously (Schilling and Overall, 2008). 2 U factor Xa (Novagen) were incubated with 200 μg PICS library for 16 h at 37°C in 50 mM HEPES, pH 8.0, 100 mM NaCl, 5 mM CaCl$_2$. Prime-side cleavage products were isolated and identified by LC-MS/MS as described (Schilling and Overall, 2008). Non-enzyme constraint spectrum-to-sequence assignment was applied using X! Tandem (Craig and Beavis, 2004) with a decoy search strategy in conjunction with PeptideProphet at 95% confidence level (Keller et al., 2002). Mass tolerance was 100 ppm for parent ions and 0.1 Da for fragment ions. Static modifications were carboxyamidomethylation of cysteine residues (+57.02 Da), dimethylation of lysine ε-amines (+28.03 Da), and thioacylation of amino-termini (+88.00 Da). The corresponding non-prime sequences were derived bioinformatically (Schilling and Overall, 2008; Schilling et al., 2011). Positional occurrences are shown as enrichment or depletion over natural abundance with human amino acid abundances being derived from the International Protein Index (Kersey et al., 2004). For heatmap-style representation, only enrichments larger than two-fold are displayed. Sequence logos were generated with iceLogo (Colaert et al., 2009). Amino acids are stated in one-letter code, cysteine (C*) is carboxyamidomethylated and lysine (K*) is dimethylated.

on residues exceeding their natural abundance more than two-fold. A corresponding cut-off has been applied for the heatmap-style representation of positional frequencies but not for the sequence logo-type depiction.

Using both libraries, PICS confirmed the canonical specificity for P1 arginine. Cleavage at P1 lysine, as described by Gosalia et al. (Gosalia et al., 2005), was not detected as expected since PICS employs peptide libraries with dimethylated lysine residues. Although dimethylation preserves

the basic character of lysine and matrix metalloproteases show canonical P2' specificity for dimethylated lysine (Schilling and Overall, 2008), serine proteases with trypsin-like specificity fail to cleave C-terminally to dimethylated lysine. Hence PICS is not suited to probe for lysine specificity. In S2, both the chymotryptic and Glu-C PICS assays determined the previously reported preferences for glycine and alanine. A partial preference for phenylalanine is observed in the Glu-C PICS assays. Here, P2 phenylalanine

Table 1   Numerical representation of PICS specificity profiles.

(A) Factor Xa subsite mapping with a chymotryptic PICS library – positional occurrences relative to natural abundance.

	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'	P6'
A	0.2	0.2	0.3	0.5	2.3	0.0	2.8	2.3	1.2	1.3	1.7	1.7
C	0.0	0.5	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.5	0.0
D	1.0	1.2	1.2	2.0	0.3	0.3	0.0	0.3	2.0	2.0	2.2	2.0
E	0.5	0.7	0.8	0.7	1.5	0.0	0.2	0.8	1.0	1.3	0.7	1.5
F	0.0	0.0	0.0	0.0	0.0	0.0	0.3	1.6	1.0	0.3	0.3	1.0
G	0.7	1.4	0.3	0.7	5.6	0.0	1.4	1.7	2.3	2.1	1.2	0.7
H	0.0	0.0	0.5	0.5	0.9	0.0	0.5	0.0	0.0	0.0	0.5	0.5
I	0.5	0.5	1.1	2.4	0.5	0.0	0.5	0.8	1.3	1.9	1.6	1.3
K	1.4	1.4	0.4	0.6	0.4	0.0	3.1	0.0	1.2	2.5	0.4	1.0
L	0.0	0.2	0.1	0.0	0.0	0.0	0.1	1.1	0.6	0.2	0.9	0.9
M	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.6	0.0	0.5	0.5	1.0
N	0.3	0.7	0.3	2.3	1.0	0.0	0.0	0.0	0.3	0.0	1.0	2.3
P	0.4	0.7	2.7	0.7	1.1	0.0	0.4	0.5	1.2	1.4	2.2	0.7
Q	0.5	0.5	1.0	0.7	0.7	0.2	0.7	0.5	0.0	0.5	0.0	1.0
R	0.2	0.2	1.0	1.8	0.2	15.0	0.8	0.4	0.0	0.2	0.2	0.2
S	0.4	0.4	0.7	1.1	0.4	0.1	2.4	0.8	0.4	0.7	0.4	0.6
T	1.1	0.4	0.2	0.6	0.4	0.2	1.3	0.9	2.1	1.1	0.9	0.2
V	0.2	1.3	1.3	1.9	0.2	0.0	0.8	3.3	2.3	0.8	2.3	1.9
W	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0
Y	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.9	0.0	0.5	0.0	0.0

(B) Factor Xa subsite mapping with a GluC PICS library – positional occurrences relative to natural abundance.

	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'	P6'
A	2.9	0.4	1.7	2.9	2.1	0.0	3.4	0.8	0.4	2.1	0.8	0.8
C	1.3	0.0	0.0	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0
D	2.5	0.0	0.0	2.5	0.0	0.0	0.0	0.0	1.9	0.6	0.6	0.6
E	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.4
F	0.0	0.0	0.0	0.0	4.9	0.0	0.0	4.9	0.8	0.8	0.8	3.3
G	1.3	2.2	0.4	0.0	4.4	0.0	1.8	0.4	1.3	0.4	2.6	1.8
H	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0
I	0.7	0.7	0.7	0.7	0.0	0.0	0.7	1.4	0.7	1.4	0.7	1.4
K	0.5	2.6	0.0	2.1	0.5	0.5	2.6	0.5	0.5	1.5	2.6	0.5
L	0.3	0.6	1.2	0.9	0.6	0.3	0.3	2.4	3.0	0.3	2.1	3.9
M	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N	0.8	0.0	0.0	1.7	1.7	0.0	0.0	0.0	0.0	0.8	2.5	1.7
P	0.0	0.4	2.3	0.4	0.9	0.0	0.0	0.0	0.9	1.8	0.9	0.4
Q	0.0	0.0	2.5	2.5	0.6	0.0	0.0	0.6	0.0	1.2	1.8	0.6
R	0.0	0.0	1.0	0.5	0.0	14.2	0.0	0.5	1.0	0.0	0.0	0.0
S	0.0	0.7	0.0	1.0	0.3	0.0	2.5	1.7	1.4	3.2	0.0	0.0
T	0.0	0.0	1.1	0.5	0.0	0.0	1.6	1.1	0.5	0.0	0.0	1.1
V	0.0	0.5	0.0	0.5	0.0	0.0	2.0	1.5	1.5	1.0	0.5	0.0
W	0.0	4.5	0.0	0.0	0.0	0.0	0.0	0.0	2.2	0.0	0.0	0.0
Y	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3	1.1	0.0	1.1	0.0

Positional amino acid occurrences are shown as enrichment over natural abundance with human amino acid abundances being derived from the International Protein Index (Kersey et al., 2004).

exceeds its natural abundance 4.9-fold (Table 1). As previously discussed, the dual preference for small residues such as glycine or alanine alongside phenylalanine is in agreement with a number of studies on factor Xa specificity.

S3 contributes less to factor Xa specificity than S2. Both PICS assays detect a minor preference for aspartic acid, which exceeds its natural abundance two-fold and 2.5-fold, respectively. Further S3 specificity determinants fail to emerge clearly. Slight preferences for isoleucine and valine as well as arginine are observed in the chymotryptic PICS assay. Although these S3 preferences are corroborated by previously discussed factor Xa profiling experiments (Furlong et al., 2002; Scholle et al., 2006; Hsu et al., 2008) they are not confirmed by the Glu-C PICS assay. In S4, both PICS assays determine a minor preference for proline. This is corroborated by a specificity study employing synthetic peptide libraries (Furlong et al., 2002). However, factor Xa PICS profiling does not substantiate a preference for P4 isoleucine which has been observed with the aforementioned synthetic peptide library screen and which is frequently used to direct factor Xa cleavage in tag removal of recombinant fusion proteins (Nagai et al., 1985).

Figure 2 The web-based platform for PICS data analysis.
This online resource is publicly accessible at http://clipserve.clip.ubc.ca/pics/. It primarily serves to derive non-prime cleavage sequences for prime-site PICS cleavage sequences and generates heatmap-style, graphical representation of protease specificity (Schilling et al., 2011).

In S1', both PICS assays determined slight selectivity for small amino acids, such as serine and alanine. In addition, both PICS assays indicate an S1' preference for dimethylated lysine. In S2', the Glu-C PICS assay indicates a preference for phenylalanine, leucine, and, to a lesser extent, tyrosine. In comparison to S1', S2' appears to contribute more strongly to factor Xa specificity. With the exception of missed chymotryptic cleavage sites, chymotryptic PICS libraries lack internal phenylalanine and leucine and so are not well suited for specificity profiling for these residues. Neverthe-

less, the chymotryptic PICS assay showed valine as the most preferred residue in S2' and a total of 16 factor Xa cleavage sites from the chymotryptic PICS library have an S2' phenylalanine, leucine or tyrosine, possibly stemming from PICS peptides with missed chymotryptic cleavages that could then be recognized by factor Xa. In conclusion, both the Glu-C and chymotryptic PICS assays indicate an S2' preference for aromatic and aliphatic residues. For S3', both PICS assays detect a minor preference for aspartic acid. Overall however, S3' appears to have a very limited contribution to factor Xa

selectivity. Altogether, subsite mapping of factor Xa substantiates previous findings about its extended active site specificity and refines our knowledge about prime-site specificity determinants.

The present PICS profiling of factor Xa represents the first reported application of WebPICS, our new web-based platform for automated PICS data analysis, http://clipserve.clip.ubc.ca/pics/ (Figure 2). At the same site, instructions for use are also found online at http://clipserve.clip.ubc.ca/pics/manual.pdf. Notably, the new WebPICS platform also processes data from *in vivo* cleavage experiments that employ native protein substrates (Kleifeld et al., 2010) to enable comparison of specificity determinants for peptidic and native protein substrates. In addition, prime-side cleavage sequences from any N-terminal proteomics experiment can be used, not only PICS (Schilling and Overall, 2008), but also TAILS (Kleifeld et al., 2010), COFRADIC (Van Damme et al., 2005), and N-terminal biotinylation schemes (Timmer et al., 2007; Mahrus et al., 2008). Such data can be easily entered via 'copy and paste' and rendered non-redundant. Input sequences can originate from any MS/MS analysis software. For PeptideProphet, non-sequence information such as preceding residues or modified masses are automatically removed. Prime-side sequences are then extended N-terminally until the first potential cleavage site of the digestion protease used in peptide library generation (Schilling and Overall, 2008). Trypsin, chymotrypsin, and Glu-C are supported as digestion proteases and human, mouse, and *E. coli* are supported as proteome sources. *Arabidopsis* will be added soon. In summary, the web-based PICS platform makes PICS specificity profiling widely accessible and removes the need for expert knowledge in data processing.

Accession codes

The LC-MS/MS data of this publication have been made publicly available at the PRIDEPRoteomics IDEntifications database (Vizcaíno et al., 2010), accession numbers 17155-56.

Acknowledgments

References

Bajaj, S.P., Rapaport, S.I., and Brown, S.F. (1981). Isolation and characterization of human factor VII. Activation of factor VII by factor Xa. J. Biol. Chem. *256*, 253–259.

Bianchini, E.P., Louvain, V.B., Marque, P.-E., Juliano, M.A., Juliano, L., and Le Bonniec, B.F. (2002). Mapping of the catalytic groove preferences of factor Xa reveals an inadequate selectivity for its macromolecule substrates. J. Biol. Chem. *277*, 20527–20534.

Brandstetter, H., Kühne, A., Bode, W., Huber, R., von der Saal, W., Wirthensohn, K., and Engh, R.A. (1996). X-ray structure of active site-inhibited clotting factor Xa. Implications for drug design and substrate recognition. J. Biol. Chem. *271*, 29988–29992.

Camerer, E., Huang, W., and Coughlin, S.R. (2000). Tissue factor- and factor X-dependent activation of protease-activated receptor 2 by factor VIIa. Proc. Natl. Acad. Sci. USA *97*, 5255–5260.

Cho, K., Tanaka, T., Cook, R.R., Kisiel, W., Fujikawa, K., Kurachi, K., and Powers, J.C. (1984). Active-site mapping of bovine and human blood coagulation serine proteases using synthetic peptide 4-nitroanilide and thio ester substrates. Biochemistry *23*, 644–650.

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. Nat. Methods *6*, 786–787.

Craig, R. and Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. Bioinformatics *20*, 1466–1467.

Foster, W.B., Nesheim, M.E., and Mann, K.G. (1983). The factor Xa-catalyzed activation of factor V. J. Biol. Chem. *258*, 13970–13977.

Furlong, S.T., Mauger, R.C., Strimpler, A.M., Liu, Y.-P., Morris, F.X., and Edwards, P.D. (2002). Synthesis and physical characterization of a P1 arginine combinatorial library, and its application to the determination of the substrate specificity of serine peptidases. Bioorg. Med. Chem. *10*, 3637–3647.

Gosalia, D.N., Salisbury, C.M., Maly, D.J., Ellman, J.A., and Diamond, S.L. (2005). Profiling serine protease substrate specificity with solution phase fluorogenic peptide microarrays. Proteomics *5*, 1292–1298.

Harbrecht, U. (2011). Old and new anticoagulants. Hamostaseologie *31*, 21–27.

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.

Hsu, H.-J., Tsai, K.-C., Sun, Y.-K., Chang, H.-J., Huang, Y.-J., Yu, H.-M., Lin, C.-H., Mao, S.-S., and Yang, A.-S. (2008). Factor Xa active site substrate specificity with substrate phage display and computational molecular modeling. J. Biol. Chem. *283*, 12343–12353.

Kawabata, S., Miura, T., Morita, T., Kato, H., Fujikawa, K., Iwanaga, S., Takada, K., Kimura, T., and Sakakibara, S. (1988). Highly sensitive peptide-4-methylcoumaryl-7-amide substrates for blood-clotting proteases and trypsin. Eur. J. Biochem. *172*, 17–25.

Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. *74*, 5383–5392.

Kersey, P.J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004). The International Protein Index: an integrated database for proteomics experiments. Proteomics *4*, 1985–1988.

Kleifeld, O., Doucet, A., Auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat. Biotechnol. *28*, 281–288.

Krishnaswamy, S., Jones, K.C., and Mann, K.G. (1988). Prothrom-

binase complex assembly. Kinetic mechanism of enzyme assembly on phospholipid vesicles. J. Biol. Chem. *263*, 3823–3834.

Ludeman, J.P., Pike, R.N., Bromfield, K.M., Duggan, P.J., Cianci, J., Le Bonniec, B., Whisstock, J.C., and Bottomley, S.P. (2003). Determination of the P1', P2' and P3' subsite-specificity of factor Xa. Int. J. Biochem. Cell Biol. *35*, 221–225.

Mahrus, S., Trinidad, J.C., Barkan, D.T., Sali, A., Burlingame, A.L., and Wells, J.A. (2008). Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell *134*, 866–876.

Matthews, D.J. and Wells, J.A. (1993). Substrate phage: selection of protease substrates by monovalent phage display. Science *260*, 1113–1117.

Nagai, K., Perutz, M.F., and Poyart, C. (1985). Oxygen binding properties of human mutant hemoglobins synthesized in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *82*, 7252–7255.

Parker, E.T., Pohl, J., Blackburn, M.N., and Lollar, P. (1997). Subunit structure and function of porcine factor Xa-activated factor VIII. Biochemistry *36*, 9365–9373.

Ridky, T.W., Cameron, C.E., Cameron, J., Leis, J., Copeland, T., Wlodawer, A., Weber, I.T., and Harrison, R.W. (1996). Human immunodeficiency virus, type 1 protease substrate specificity is limited by interactions between substrate amino acids bound in adjacent enzyme subsites. J. Biol. Chem. *271*, 4709–4717.

Schilling, O., Huesgen, P.F., Barré, O., Auf dem Keller, U., and Overall, C.M. (2011). Characterization of the prime and non-prime active site specificities of proteases by proteome-derived peptide libraries and tandem mass spectrometry. Nat. Protoc. *6*, 111–120.

Schilling, O. and Overall, C.M. (2008). Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. Nat. Biotechnol. *26*, 685–694.

Scholle, M.D., Kriplani, U., Pabon, A., Sishtla, K., Glucksman, M.J., and Kay, B.K. (2006). Mapping protease substrates by using a biotinylated phage substrate library. Chembiochem. *7*, 834–838.

Timmer, J.C., Enoksson, M., Wildfang, E., Zhu, W., Igarashi, Y., Denault, J.-B., Ma, Y., Dummitt, B., Chang, Y.-H., Mast, A.E., et al. (2007). Profiling constitutive proteolytic events in vivo. Biochem. J. *407*, 41–48.

Van Damme, P., Martens, L., Van Damme, J., Hugelier, K., Staes, A., Vandekerckhove, J., and Gevaert, K. (2005). Caspase-specific and nonspecific in vivo protein processing during Fas-induced apoptosis. Nat. Methods *2*, 771–777.

Vizcaíno, J.A., Côté, R., Reisinger, F., Barsnes, H., Foster, J.M., Rameseder, J., Hermjakob, H., and Martens, L. (2010). The Proteomics Identifications database: 2010 update. Nucleic Acids Res. *38*, D736–D742.