Comparison of Protein Hydrolysis Catalyzed by Bovine, Porcine, and Human Trypsins

Yuxi Deng, Harry Gruppen, and Peter A. Wierenga*

Laboratory of Food Chemistry, Wageningen University, P.O. Bus 17, 6708 AA Wageningen, The Netherlands

Supporting Information

ABSTRACT: Based on trypsin specificity (for lysines and arginines), trypsins from different sources are expected to hydrolyze a given protein to the same theoretical maximum degree of hydrolysis (DH\textsubscript{max,threo}). This is in contrast with experiments. Using \(\alpha\)-lactalbumin and \(\beta\)-casein, this study aims to reveal if the differences among experimental DH\textsubscript{max} (DH\textsubscript{max,exp}) by bovine, porcine, and human trypsins are due to their secondary specificity. Peptide analysis showed that \(\sim78\%\) of all the cleavage sites were efficiently hydrolyzed by porcine trypsin, and \(\sim47\%\) and \(\sim53\%\) were efficiently hydrolyzed by bovine and human trypsins, respectively. These differences were explained by the enzyme secondary specificity, that is, their sensitivities to the amino acids around the cleavage sites. The DH\textsubscript{max} predictions based on the secondary specificity were 4 times closer to the DH\textsubscript{max,exp} than the predictions based on trypsin specificity alone (DH\textsubscript{max,threo}). Proposed preliminary relations between binding sites and trypsin secondary specificity allow DH\textsubscript{max,exp} estimations of tryptic hydrolysis of other proteins.

KEYWORDS: LC-MS, peptide release kinetics, tryptic hydrolysis, protein digestibility, secondary specificity

INTRODUCTION

In vivo protein digestion in the intestines is often simulated by \textit{in vitro} hydrolysis with pancreatin, in which the main proteases are trypsin and chymotrypsin. Trypsins specifically hydrolyze peptide bonds on the carboxylic sides of lysines and arginines. In principle, when the specificity of the enzyme and the amino acid (AA) sequence of the protein are known, the theoretical maximum degree of hydrolysis (DH\textsubscript{max,threo}) can be calculated. In the case of the tryptic hydrolysis of \(\alpha\)-lactalbumin, the DH\textsubscript{max,threo} is 10.7\% (see Table 2). When this protein was hydrolyzed by porcine trypsin, the experimental maximum degree of hydrolysis (DH\textsubscript{max,exp}, \(\sim10.7\%\)) was indeed similar to the DH\textsubscript{max,threo}. Using bovine trypsin, surprisingly, the DH\textsubscript{max,exp} was only \(\sim3.6\%\). Peptide analysis showed that the lower DH\textsubscript{max,exp} reached by bovine trypsin was due to 8 of the 14 cleavage sites (CSs) on the protein being hydrolyzed by the enzyme only at low rates. It was proposed that these low hydrolysis rate constants were due to the hindrance of bovine trypsin by the charged AAs surrounding the CSs (Table 1). The differences in the hydrolysis rate constants of CSs in a protein are quantitatively described by enzyme selectivity. The current study aims to investigate hydrolysis by bovine, porcine, and human trypsins and to determine if the differences in DH\textsubscript{max,exp} could be explained based on the enzyme selectivity. A detailed understanding of the differences in the enzyme selectivity of the trypsins will allow a better understanding of the reported \textit{in vitro} digestibility of a protein.

In commercial trypsins extracted from bovine and porcine pancreases, two variants of trypsin, cationic (trypsin-1) and anionic (trypsin-2), are present. Trypsin-1 is the dominant variant. Human trypsin consists of three variants of trypsin; the contents of trypsin-1 (cationic), trypsin-2 (anionic), and trypsin-3 (mesotrypsin) were found to be 60, 30, and 10\%, respectively. Unless stated otherwise, in this study, bovine and porcine trypsins refer to trypsin-1, and human trypsin refers to trypsin-2. For bovine trypsin, the experimental maximum degrees of hydrolysis (DH\textsubscript{max,exp}) of \(\alpha\)-lactalbumin, \(\beta\)-casein, and \(\beta\)-lactoglobulin were reported to be 5.1\%, 4.3\%, and 8.0\%, respectively. These values were only 48, 60, and 75\% of the theoretical maximum degrees of hydrolysis (DH\textsubscript{max,threo}) of the proteins. The peptide analysis of the \(\alpha\)-lactalbumin hydrolysates showed that 2 out of 13 CSs were not hydrolyzed. For the hydrolyzed CSs, the hydrolysis rate constants of bovine trypsin toward individual CSs varied from \(4.7 \times 10^{-3}\) to \(3.0 \times 10^{-3}\) s\(^{-1}\). mg\textsuperscript{-1}\ enzyme which were equivalent to selectivity values (relative hydrolysis rate constants) ranging from 0.4 to 27.6\%. The two CSs that were not hydrolyzed were referred to as missed cleavages. The occurrence of missed cleavages was also shown in bovine tryptic hydrolysates of human serum albumin hydrolysates, in which \(\sim100\) peptides with missed CSs were identified. The occurrence of missed cleavages in bovine tryptic hydrolysis has been attributed to the presence of charged AAs surrounding the CSs. When a protease binds to a substrate, part of the enzyme (the S4′→S4′ positions) interacts with part of the substrate AA sequence (the P4→P4′ binding site of the CS). P4→P4′ positions are numbered from the N- to C-terminal sides of the CS, and P1 identifies the AA for which the enzyme is specific. Recently, this concept was applied to describe bovine trypsin selectivity toward CSs in \(\alpha\)-lactalbumin, \(\beta\)-casein, and \(\beta\)-lactoglobulin. To obtain a better understanding of bovine tryptic hydrolysis, the relations between the experimental selectivity and molecular properties of the binding site positions were investigated. Based on the relations, CSs could be differentiated into three categories by the charges of the AAs in the P2 and P2′ positions.
Table 1. Selectivity Rules for Bovine Trypsin$^2$ and Numbers of Cleavage Sites (CSs) Correctly Predicted by the Rules out of the Total Numbers of CSs in $\alpha$-Lactalbumin and $\beta$-Casein for Bovine, Porcine, and Human Trypsins$^6$

| Category | Binding site | Number of correctly predicted CSs/total CSs |
|----------|--------------|------------------------------------------|
| High selectivity | All others | 11/12 | 14/21 | 9/14 |
| Intermediate selectivity | KRDE | 5/9 | 0/2 | 2/3 |
| Low/zero selectivity | KR | 1/7 | 2/5 | 3/11 |
| C | 1/7 | | | |

$^2$Red indicates charged amino acids (KRDE) and green indicates neutral amino acids (all others). The letters in the columns represent the abbreviations of the amino acids. A dash (−) means that there are no amino acids present. To avoid repetition, the right part of the table, showing the numbers of correctly predicted CSs out of the total numbers of CSs, was from experimental data that will be discussed in the Links between Binding Site Sequences and the Selectivity of Porcine and Human Trypsins section.

Table 2. Nitrogen to Protein Conversion Factors (N-factors), Protein Contents and Purities, Molar Weights, Numbers of Peptide Bonds Per Gram of Protein ($h_{\text{max}}$), Numbers of Amino Acids (AAs), Molar Extinction Coefficients at 214 nm ($\epsilon_{214}$), Numbers of Maximum Degrees of Hydrolysis (DH$_{\text{max,theo}}$) of the Three Trypsins (Bovine, Porcine, and Human) and Two Proteins ($\alpha$-Lactalbumin and $\beta$-Casein) and Numbers of Unique Peptides in Each Hydrolysate

| protein | Uniprot code | N-factor | protein content [gram of N] | purity [%] | molar weight [Da] | $h_{\text{max}}$ [mmol/g] | CS | AA | DH$_{\text{max,theo}}$ [%] | $\epsilon_{214}$ [M$^{-1}$ cm$^{-1}$] | $\alpha$-LA$^a$ peptides | $\beta$-cas$^b$ peptides |
|---------|--------------|----------|-----------------------------|-----------|------------------|------------------------|----|----|-----------------------|-------------------------|------------------|------------------|
| $\alpha$-LA | P00711 | 6.25 | 93% | 90 | 14 186 | 8.6 | 13 | 209 | 7.2 | 424 047 |
| $\beta$-cas | P02666 | 6.39 | 90% | 90 | 23 983 | 8.7 | 15 | 209 | 7.2 | 30 461 |
| BT | P00760 | 5.97 | 80% | 100 | 23 305 | 223 | 223 | 223 | 223 | 30 36 |
| PT | P00761 | 5.84 | 86% | 100 | 23 475 | 223 | 223 | 223 | 223 | 30 36 |
| HT | P07478 | 5.98 | 77% | 100 | 24 041 | 224 | 224 | 224 | 224 | 30 36 |

$^a$From Uniprot (http://www.uniprot.org). $^b$The protein purity is defined as the percentage of the total UV$_{280}$ peak area that is the UV$_{280}$ peak area of the target protein, determined by RP-UPLC.

Materials. $\alpha$-Lactalbumin ($\alpha$-LA) was obtained from Davisco Foods International Inc. (Le Sueur, MN). According to the results from a circular dichroism study, $\sim$72% of the $\alpha$-LA was in the apo form, and the rest was in holo form.$^1$ $\beta$-Casein ($\beta$-cas, C6905), bovine trypsin (BT, T1426), porcine trypsin (PT, T0303), and aprotinin (A6279) were all purchased from Sigma-Aldrich (St. Louis, MO). Recombinant human trypsin-2 (henceforth referred to as human trypsin and abbreviated as HT) was purchased from Shanghai Yaxin Biotechnology Company, Ltd. (Shanghai, China). The molecular properties, including the protein contents and purities, of the two proteins (i.e. $\alpha$-LA and $\beta$-cas) and the three trypsins (i.e. BT, PT, and HT) were listed in Table 2.
enzymes, all three trypsins were analyzed using ultrahigh-performance liquid chromatography with mass spectrometry (UPLC-MS). There were no chymotrypsin peaks found in the chromatograms of any of the trypsins. The activities of the bovine, porcine, and human trypsins were specified to be \( \geq 10,000, 13,000 - 20,000 \), and \( \geq 7500 \) benzoyl-L-arginine-ethyl-ester units per milligram of protein by the manufacturer, respectively. Aprotinin was used to inhibit trypsin hydrolysis, as described below. Aprotinin, in a 0.9% sodium chloride and 0.9% benzylalcohol solution, was determined to have a concentration of 2.3 mg/mL based on the peak area at 214 nm (\( A_{214} \)) determined by RP-UPLC-UV. All the other chemicals were of analytical grade and purchased from Sigma or Merck (Darmstadt, Germany).

**Methods. Enzymatic Hydrolysis of the Proteins.** The enzymatic hydrolysis of the proteins was performed as described previously. The \( \alpha \)-LA and \( \beta \)-cas solutions (1%, w/v; 10 mL in Millipore water; Millipore, Burlington, MA) were adjusted to pH 8.0 and equilibrated at 37 °C (Switzerland). A trypsin solution (100 \( \mu \)L, 10 mg/mL in Millipore water) was added to each equilibrated protein solution to reach enzyme to substrate ratios of 1:100 (w/w). Hydrolysis was performed at 37 °C for 3 h, and 0.2 M NaOH was used to keep the pH constant. The degree of hydrolysis (\( DH_{\text{stat}} \)) was calculated using eq 1.11

\[
DH_{\text{stat}} (%) = \frac{V_0 \times N_0 \times \frac{1}{\alpha} \times \frac{1}{m_p} \times \frac{1}{h_{\text{tot}}} \times 100}{1}
\]

(1)

where \( V_0 \) (mL) is the volume of NaOH added, \( N_0 \) (mol/L) is the normality of the NaOH, \( \alpha \) is the average degree of dissociation of the \( \alpha \)-NH groups (1/\( \alpha \) = 1.3 at 37 °C and pH 8.5) \( m_p \) (g) is the mass of the protein in the solution (taking into account the protein content in the powder based on the Dumas results), and \( h_{\text{tot}} \) (mmol/g) is the total number of millimoles of peptide bonds per gram of protein substrate (Table 2). For each hydrolysis, 200 \( \mu \)L samples were taken at set DH values. The DH values at which the samples were taken depended on the \( DH_{\text{max}} \). For all the hydrolysates were analyzed on an H-class Acquity Reversed-Phase Ultrahigh-Performance Liquid Chromatography (RP-UPLC). The hydrolysates were analyzed on an H-class Acquity UPLC system (Waters, Milford, MA) equipped with a BEH C18 column (1.7 \( \mu \)m, 2.1 \( \times \) 100 mm, Waters), connected to an Acquity UPLC PDA detector (Waters). The disulfide bridges were reduced by incubating the samples at protein concentrations of 0.5% for 2 h with 100 mM dithiothreitol (DTT) in 50 mM Tris-HCl buffer at pH 8.0. The reduced samples were further diluted to protein concentrations of 0.1% (w/v) and centrifuged (10 min, 14,000 rpm) before injection (4, s/1).

**Electron-Spray-Ionization Time-of-Flight Mass Spectrometry (ESI-Q-TOF-MS).** The MS and MS/MS (MS/MS) data of the peptides were collected with an online SYNAPT G2-Si high-definition mass spectrometer (Waters) coupled to the RP-UPLC system. The MS system was calibrated with sodium iodide. Online lock-mass data (angiotensin II, [M + 2H] \( ^2+ \): 523.7751) was acquired. Based on the differences between the measured and theoretical lock masses, corrections were applied to the mass to charge ratios (m/z) of the peptides during data collection. The data were analyzed manually using MassLynx software version 4.1 (Waters) and UNIFI software version 1.8 (Waters).

**Estimation of the Number of Possible Peptides.** For a protein with \( N \) cleavage sites, the theoretical number of possible peptides (NPP \( \text{max,} \text{theo} \)) was calculated based on enzyme specificity using eq 4.15

\[
\text{NPP}_{\text{max,} \text{theo}} = \frac{(N + 1)(N + 2)}{2}
\]

where \( N \) is the number of CSs in the protein sequence. Experimentally, the enzymes only efficiently hydrolyzed a few high selectivity CSs (HSSs). Therefore, the number of possible peptides corrected based on the enzyme selectivity (NPP \( \text{max,HSS} \)) was calculated by replacing \( N \) in eq 4 with the number of HSSs (HSS) estimated using eq 3.

**Peptide Identification and Quantification.** The mass tolerance between the theoretical masses and measured masses for the accepted annotations was set at 100 ppm. Two AA modifications, methionine oxidation (+16 Da) in \( \alpha \)-LA and phosphorylation (+80 Da per phosphoserine) in \( \beta \)-cas, were taken into account. The peptides were annotated based on the MS spectra and were confirmed by the identification of the b and y fragments in the MS/MS spectra when possible. In cases of coelution, the concentration of each of the coeluting peptides was calculated by using the intensity of the total ion count to divide the UV area at 214 nm. The concentration of each peptide was calculated using eq 5.

\[
C_{\text{peptide}} (\mu \text{M}) = \frac{A_{214} \cdot Q}{\varepsilon_{214} \cdot 1 \cdot V_{\text{expt}}} \cdot k_{\text{cell}}
\]

where \( C_{\text{peptide}} (\mu \text{M}) \) is the peptide concentration, \( A_{214} (\mu \text{AU-min}) \) is the UV area at 214 nm, \( V_{\text{expt}} (\mu \text{L}) \) is the injected volume of the sample, Q (\( \mu \text{L-min}^{-1} \)) is the flow rate, and \( 1 \) (cm) is the path length of the UV cell, which is 1 cm according to the manufacturer. \( \varepsilon_{214} (1/\mu \text{M/cm}) \) is the molar extinction coefficient at 214 nm, which was calculated as described previously. Because of the multiple reflections caused by the coating of the UV cell, the effective path length of the light through the cell is not the same as the length specified by the producer. This
issue was also reported in earlier studies. To correct for this, the cell constant of the UV detector \(k_{da}\) was determined using a series of standard solutions made with \(\beta\)-lactoglobulin, \(\beta\)-cas, and angiotensin II, and the approach was as described elsewhere.

The ratio between the measured and expected \(A_{1314}\) was taken as the \(k_{da}\) value. For the UV cell used in this study, \(k_{da}\) was reported to be 0.78. The linear region of the peak area \(A_{1314}\) in the UV detector ranged from \(5 \times 10^{-10} \text{ to } 6 \times 10^{-10} \mu\text{AU}\text{min}^{-1}.\) Therefore, peptide quantification was performed for peptides with \(A_{1314} \geq 5 \times 10^{-10} \mu\text{AU}\text{min}^{-1}.\) In all the hydrolysates, the total peak areas in the chromatograms were 95 ± 6% of the expected values, and 97 ± 3% of the total \(A_{1314}\) was assigned to annotated peptides. The standard errors for the peptide concentrations for hydrolysates injected in triplicate were reported to be 6% using the same quantification method as described above. The selectivity determined from duplicate hydrolysates was reported to have standard errors of 15%. In this work, which had duplicate experiments, the average standard errors for the peptide concentrations were ~15%.

### Enzyme Affinities toward Intact Proteins

The percentages of remaining intact proteins in the hydrolysates were plotted against hydrolysis times. The initial hydrolysis rate constants of the intact proteins, \(k_{enzyme,protein} \text{ (s}^{-1})\), were calculated from the initial slopes of the curves (the linear slopes of the first three points as functions of time). According to the Linderstrøm-Lang theory, the amount of intact protein as a function of DH illustrates the enzyme affinity toward intact proteins relative to the enzyme affinity toward intermediate peptides.\(^{19}\) Based on this theory, the percentages of intact proteins in the hydrolysates were also plotted against the DHmax (%). Because the hydrolysates by the three trypsins did not reach the same DHmax for the DHmax were normalized by dividing them by the DHmax.

#### Quality Check of the Peptide Analysis Annotation and Quantification

The quality of the peptide analysis was analyzed using three parameters, the amino acid, peptide, and molar sequence coverages, which were calculated with eqs (6), (7), and (8), respectively. These checks were performed for all the hydrolysates. For each enzyme–substrate combination, the average value for each parameter was calculated as the average of that parameter for the duplicate experiments at all DH values.

\[
\text{amino acid sequence coverage (\%)} = \frac{\text{unique annotated AA}}{\text{AAs}_{\text{protein}}} \times 100\% 
\]

\[
\text{peptide sequence coverage (\%)} = \frac{\text{AA (annotated peptides)}}{\text{AA (annotated peptides) + AA (missing peptides)}} \times 100\% 
\]

\[
\text{molar sequence coverage (\%)} = \left[ 1 - \frac{\sum (C_i - C_0)^2}{C_0} \right] \times 100\% 
\]

\(C_\alpha, (\mu\text{M})\) represents the concentration of each individual AA (n) in the protein sequence, \(C_0 (\mu\text{M})\) is the initial injected protein concentration, and \(\text{AAs}_{\text{protein}}\) is the number of AAs in the sequence of the parental protein. The injected initial concentrations \(C_\alpha\) were 59 and 34 \(\mu\text{M}\) for \(\alpha\)-LA and \(\beta\)-cas, respectively (calculated based on the protein contents and purities listed in Table 2). In the \(\alpha\)-LA and \(\beta\)-cas samples at DH values of 0%, the experimental \(C_\alpha\) were measured to be 58 ± 2 and 33 ± 1 \(\mu\text{M}\), respectively, using eq 5. The missing peptides are the peptides that were not found but should have been present given the annotated peptides.\(^{3}\) To further check the quality of the peptide analysis for each hydrolysate, the degree of hydrolysis based on the peptide analysis (DH\(_{\text{max,app}}\)) was calculated using eq 9.

\[
\text{DH}_{\text{max,app}} (\%) = \frac{\sum C_{i,t}}{2 \times \left( \text{peptide bonds} \right) \times C_0} \times 100\% 
\]

where peptide bonds refers to the number of peptide bonds, and \(C_\alpha\) is the concentration of the cleavage site products, calculated using eq 10.

#### Determination of the Hydrolysis Rate Constants

For each CS, the concentration of cleavage site products \(C_{i,t}\) that originated from the hydrolysis of that peptide bond was calculated using eq 10.

\[
C_{i,t} (\mu\text{M}) = \sum \left\{ C_{\text{peptide},[x-y]} \right\} \text{for } i = x - 1 \text{ to } i = y 
\]

where \(C_{i,t} (\mu\text{M})\) refers to the concentration of CS products formed at each time point, \(t\), and \(i\) equals the sum of the concentrations of all the peptides of the sequence \([x-y]\) for which \(i = (x - 1)\) or \(i = y\). The apparent hydrolysis rate constant was obtained by fitting the second-order reaction equation (eq 11) to the experimental data.

\[
k_{\text{enzyme,protein}} (\mu\text{M}) = 2 \times \left( C_0 \text{enzyme,protein} - C_{i,t} \text{enzyme,protein} \right) \left( 1 + k_{\text{app}} \text{enzyme,protein} \times C_0 \text{enzyme,protein} \times t \right) 
\]

where both \(C_0 \text{enzyme,protein} (\mu\text{M})\) and \(k_{\text{app}} \text{enzyme,protein} (\text{sink-1})\) are the fitting parameters, and the superscripts refer to the enzyme–substrate combinations. The apparent hydrolysis rate constant, \(k_{\text{enzyme,protein}} \text{ (sink-1)}\), is referred to as \(k_{\text{enzyme,protein}} \text{ (sink-1)}\). The hydrolysis rate constant, \(k_{\text{enzyme,protein}} \text{ (sink-1)}\), of each hydrolysis was calculated using eq 12.

\[
k_{\text{enzyme,protein}} (\text{sink-1}) = \frac{k_{\text{enzyme,protein}} (\text{sink-1})}{m_0} 
\]

where \(m_0 (\text{mg})\) is the mass of the enzyme added for the hydrolysis. \(C_{i,t}\) was also calculated based on the concentration of remaining nonhydrolyzed CSs \(C_{\text{remaining}}\) in the hydrolysates \(C_{i,t} = C_0 - C_{\text{remaining}}\). The hydrolysis rate constants calculated using the two methods were compared.

#### Enzyme Selectivity

The selectivity (%) of the enzymes toward the CSs present in the two substrate proteins were calculated by dividing the hydrolysis rate constant, \(k_{\text{enzyme,protein}} \text{ (sink-1)}\), of each CS by the sum of the hydrolysis rate constants for all the CSs for that enzyme–substrate combination (eq 13).

\[
\text{selectivity (enzyme,protein) (%) } = \frac{k_{\text{enzyme,protein}} (\text{sink-1})}{\sum (C_0 \text{enzyme,protein} \times C_0 \text{enzyme,protein})} \times 100\% 
\]

In a recent work on bovine trypsin, the high (HSS), intermediate (ISS), and low or zero (LSS) selectivity sites in \(\alpha\)-LA and \(\beta\)-cas were clustered using the MATLAB method k-means clustering. As a partitioning method, k-means clustering uses the squared Euclidean metric to determine the distances and the k-means++ algorithm for the cluster center initialization. It is used to find clusters that divide data points with minimized total sums of distances. To compare the hydrolysis rate constants of porcine and human trypsins toward CSs with those of bovine trypsin for each substrate, the lowest hydrolysis rate constants of bovine trypsin toward the CSs in each category (HSS, ISS, and LSS) were set as the thresholds. Using the thresholds, the numbers of HSSs, ISSs, and LSSs of porcine and human trypsins toward CSs in \(\alpha\)-LA and \(\beta\)-cas were determined.

#### Prediction of the Maximum Degrees of Hydrolysis

Previously, it was found that for bovine trypsin, the average \(k_{\text{HSS,protein}}\) of the ISSs and LSSs were ~24 and ~3% of the average \(k_{\text{HSS,protein}}\) of the HSSs, respectively. This means that within the time of the hydrolysis experiment when most HSSs were hydrolyzed, ~24 and ~3% of the ISSs and LSSs were hydrolyzed, respectively. Hence, the maximum degrees of hydrolysis (DH\(_{\text{max,app}}\)) predicted based on the enzyme selectivity can be calculated using eq 14.

\[
\text{DH}_{\text{max,app}} (\%) = \frac{\text{HSS } + 0.24 \times \text{ISS } + 0.03 \times \text{LSS}}{\text{peptide bonds}} \times 100\% 
\]
where HSS, ISS, and LSS denote the numbers of high, intermediate, and low or zero selectivity sites, respectively. Because the numbers of HSSs, ISSs, and LSSs of porcine and human tryptic hydrolyses of α-LA and β-cas were obtained based on the thresholds set from bovine trypsin selectivity, their DH\text{max,pre} were also predicted using eq 14.

Relation between Binding Site Sequences and Enzyme Selectivity toward CSs. The correlations between the binding site sequences and the hydrolysis rate constants were analyzed using the information for all the CSs. For this, at each binding site position, the standard errors of the molecular weight, hydrophobic moment, and charge of the AA were analyzed using the same approach as reported previously. The scale of hydrophobic moments ranged from that of the most hydrophilic AA (arginine, −1.8) to that of the most hydrophobic AA (isoleucine, 0.73). For the analysis at pH 8, arginine and lysine, the CSs, were positively charged, and the glutamic and aspartic acids were negatively charged. Histidine was not included as a charged AA because only 0.9% of the histidines were charged at pH 8. In addition, cysteine was considered a charged AA because the sulfhydryl group was charged at pH 8. The sulfhydryl groups of the cysteines could be free in the intact proteins or could be released as the result of disulfide bridge shuffling. If the sulfhydryl group of a cysteine was involved in a disulfide bridge, it would not carry a charge, but it could still affect hydrolysis as a result of steric hindrance for enzyme–substrate binding.

RESULTS AND DISCUSSION

General Description of α-LA and β-Cas Hydrolysis by the Three Trypsins. The curves (DH\text{stat versus} time) of the hydrolyses of α-LA by the bovine, porcine, and human trypsins were fitted using eq 2. The overall hydrolysis rate constants (k\text{DH,HT/PT/HT,α-LA}) of the hydrolyses of α-LA by the bovine (1.6 ×
10^{-3} \pm 0.2 \times 10^{-3} \text{ s}^{-1})^2$ and porcine ($1.7 \times 10^{-3} \pm 0.0 \times 10^{-3} \text{ s}^{-1}$) trypsins were slightly higher than those by human trypsin ($1.1 \times 10^{-3} \pm 0.3 \times 10^{-3} \text{ s}^{-1}$). Larger differences were found in the $\text{DH}_{\text{max,exp}}$ for the hydrolyses by the three trypsins. The $\text{DH}_{\text{max,exp}}$ values derived from eq 2 for the hydrolyses by the three trypsins were compared. For porcine trypsin, the $\text{DH}_{\text{max,fit}}$ of $\alpha$-LA hydrolysis (10.0 $\pm$ 0.3%, Figure 1A) was similar to $\text{DH}_{\text{max,fit}}$ (10.7%, Table 2). This suggested that porcine trypsin efficiently cleaved $\sim$94% of the CSs in $\alpha$-LA, suggesting that around 12 out of the 13 CSs were high selectivity sites (HSSs, eq 3). The $\text{DH}_{\text{max,fit}}$ of $\alpha$-LA hydrolyzed by human and bovine trypsins were 6.8 $\pm$ 0.1 and 5.1 $\pm$ 0.2%, respectively, which were much lower than the $\text{DH}_{\text{max,fit}}$ for bovine and human trypsin hydrolysates, only 48% (6 out of 13) and 63% (8 out of 13) of the CSs in $\alpha$-LA, respectively, were estimated to be HSSs. That porcine trypsin hydrolysis reached the $\text{DH}_{\text{max,fit}}$, but bovine trypsin hydrolysis did not was in line with the data reported in literature. The $\text{DH}_{\text{max,exp}}$ of porcine trypsin hydrolysis of whey protein concentrate at 50°C was performed at 37°C with a pH of 8.0 was reported to be 14%, which was slightly higher than the $\text{DH}_{\text{max,exp}}$ of intact $\alpha$-LA hydrolyzed by human and bovine trypsin (6.8 $\pm$ 0.1 and 5.1 $\pm$ 0.2%, respectively). Because $\alpha$-lactalbumin denatures at 37°C (the hydrolysis temperature) and $\beta$-casein is a random coil protein, the differences in hydrolysis between the two proteins were not caused by the compactness of the protein tertiary structure. It is possible that the $\beta$-casein molecules formed micellar structures in solution, which may have caused the differences between the two proteins in the overall hydrolysis rates and the hydrolysis rates of the intact proteins.

Because no differences in enzyme affinities toward the intact proteins were found in the hydrolysates by the three trypsins, the differences in the $\text{DH}_{\text{max,exp}}$ were most likely due to variations in the amounts of HSSs for each hydrolysis. This was analyzed in detail using a quantitative peptide analysis.

### Peptide Analysis of $\alpha$-LA Hydrolysis by the Three Trypsins

#### Peptide Annotation of $\alpha$-LA Hydrolysis by the Three Trypsins

Based on the estimated number of HSSs, the predicted maximum numbers of peptides (NPP$_{\text{max,HSS}}$) were 28, 91, and 45 for bovine, porcine, and human trypsin hydrolysates, respectively (eq 4). Experimentally, 23, 30, and 31 unique peptides were annotated in bovine, porcine, and human tryptic hydrolysates of $\alpha$-LA, respectively (Table 2 and Appendix A). The numbers of annotated peptides were lower than the NPP$_{\text{max,HSS}}$ particularly for porcine trypsin. This could be because a large number of the intermediate peptides that should be formed during hydrolysis only accumulated at very low concentrations (only 5% of the expected $A_{110}$ was not annotated) that were below the detection or quantification limit. The quality of peptide annotation and quantification was analyzed to identify how much of the total hydrolysate was represented by the quantified peptides.

### Quality Check of the Peptide Analysis of $\alpha$-LA Hydrolysis by the Three Trypsins

The quality of peptide analysis was considered to be sufficient based on an analysis by four parameters: the average AA, peptide, and molar sequence coverages and the DH calculated from the peptide analysis (DH$_{\text{AA}}$). The AA sequence coverages of all the samples were 100% (Table 3). The average peptide sequence coverages for the bovine, porcine, and human trypsin hydrolysates were comparable: 96 $\pm$ 3, 90 $\pm$ 2, and 95 $\pm$ 5%, respectively. The average molar sequence coverages for the bovine, porcine, and human trypsin hydrolysates were 80 $\pm$ 9, 70 $\pm$ 11, and 72 $\pm$ 7%, respectively (Table 3 and Appendix C). The sequence coverages in this work were higher than the thresholds set based on the data obtained from previous studies (5,25,26) (i.e. 80% with the average peptide and molar sequence coverages, respectively). Further confirmation was obtained by compar-

| enzyme | $\alpha$-LA | $\beta$-cas | $\alpha$-LA | $\beta$-cas | $\alpha$-LA | $\beta$-cas |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|
| bovine | 100 $\pm$ 0 | 100 $\pm$ 0 | 96 $\pm$ 3  | 83 $\pm$ 10 | 80 $\pm$ 9  | 74 $\pm$ 8  |
| porcine| 100 $\pm$ 0 | 99.5 $\pm$ 1| 90 $\pm$ 2  | 87 $\pm$ 8  | 70 $\pm$ 11 | 72 $\pm$ 7  |
| human  | 100 $\pm$ 0 | 100 $\pm$ 0 | 95 $\pm$ 5  | 91 $\pm$ 4  | 78 $\pm$ 8  | 81 $\pm$ 8  |
ing the DHMS with the DH values taken from the pH-stat titration (DHstat). Averaged over all the hydrolysates, the standard error between DHstat and DHMS was ~7.3% (Figure 3A). For the three porcine tryptic hydrolysates, toward the end of the hydrolysates, DHMS was smaller than DHstat, which was also reflected by the low molar sequence coverages (≤60%, Appendix C). This only had a minor effect on the fitting parameters, because for the fitting of the data with the equation of the second-order reaction, the results were mostly determined by the initial five data points. Those five hydrolysates had high molar sequence coverages (≥70%).

In the bovine tryptic hydrolysates, only specific cleavages were found, respectively. Such a-specific cleavages were previously found in the later stage of whey protein isolate hydrolysis by the Bacillus licheniformis protease (BLP). These sites were observed to be hydrolyzed only in certain specific peptides. Later on, these peptides were synthesized and incubated in the absence and presence of BLP. Even without the presence of the enzyme, the a-specific cleavages still occurred on those unstable peptide bonds. This autolytic degradation of intrinsically unstable peptide bonds was referred to as spontaneous cleavage. In this study, the peptides that formed after a specific cleavages were found only at the end of hydrolysis, and these sites were cleaved to low extents (≤6 μM). Therefore, they were not included as CSs.

Figure 3. Comparison of DH values calculated based on peptide analysis (DHMS) and pH-stat titration (DHstat) for (A) α-LA and (B) β-cas hydrolysates by (◇) bovine, (□) porcine, and (△) human trypsins. The dotted line (---) indicates the function y = x. The error bars are typically smaller than the marker sizes.

Figure 4. Concentrations of peptides (A) α-LA[14–16] and (B) α-LA[1–10] and cleavage site products (C) α-LA K13–14 and (D) α-LA R10–11 formed during the hydrolysates by (◇) bovine, (□) porcine, and (△) human trypsins. The dotted lines (---) represent the initial injected concentrations of α-LA. In panels (C) and (D), the dashed lines (—) represent the fits using the second-order reaction equation. The markers represent the averaged values of duplicate hydrolysates.
Peptide Release Kinetics of α-LA Hydrolysis by the Three Trypsins. In all the hydrolysates, the peptide concentrations were followed throughout the hydrolysates (Appendix E). The differences among the hydrolysates by the three enzymes were illustrated for the peptides α-LA[14–16] and α-LA[1–10] (Figure 4A,B). During the bovine and porcine trypsin hydrolysates, α-LA[14–16] was released rapidly. Because α-LA[14–16] was a final peptide (i.e., a peptide that did not contain CSs), its concentration only increased. In the porcine trypsin hydrolysis, the concentration of α-LA[14–16] in the final hydrolysate reached the injected concentration of α-LA ($C_0 = 59 \mu M$, Figure 4A). This illustrated that CSs K13–14 and K16–17 were efficiently hydrolyzed by porcine trypsin. During the bovine trypsin hydrolysis, the concentration of peptide α-LA[14–16] reached $\sim 10 \mu M$ at the end of hydrolysis. During the human trypsin hydrolysis, peptide α-LA[14–16] was released slowly, and the concentration reached $\sim 34 \mu M$ at the end. Peptide α-LA[1–10] is an intermediate peptide whose concentration can decrease after the hydrolysis of the CS K5–6 (Figure 4B). For the hydrolysates by all the trypsins, α-LA[1–10] formed rapidly in the beginning. However, for the hydrolysates by the different trypsins, the concentrations reached different values before they started to decrease. In the final porcine and human trypsin hydrolysates, α-LA[1–10] was not found, which means that K5–6 was fully hydrolyzed. Unlike the hydrolysates by these two trypsins, the concentration of α-LA[1–10] in the final bovine trypsin hydrolysate was $\sim 10 \mu M$. This means that K5–6 was not fully hydrolyzed by bovine trypsin. This was also seen in the final cleavage site products of K5–6, which will be discussed in the next section.

Hydrolysis Rate Constants of the CSs in α-LA Hydrolysis by the Three Trypsins. Based on the peptide concentrations at each DH, the concentration of the cleavage site products for each CS during α-LA hydrolysis was determined. There were clear differences in the hydrolysis rates of each trypsin for the individual CSs, as shown, for instance, for CS K13–14 (Figure 4C). This CS was rapidly hydrolyzed by bovine and porcine trypsins but not by human trypsin. CS K13–14 was fully hydrolyzed by porcine trypsin, as indicated by the final CS product concentrations as well as by the peptide release kinetics of α-LA[14–16]. For CS K5–6, the concentrations of the CS products increased rapidly in the hydrolysates by all the trypsins (Figure 4D), which was expected based on the peptide release kinetics of peptide α-LA[1–10], discussed above. Still, although the concentrations of the CS products at the end of the porcine and human trypsin hydrolysates reached 59 and 55 μM, respectively, for the bovine trypsin hydrolysate, it only reached 34 μM. During the porcine trypsin hydrolysates, most CSs reached $C_0$, meaning that they were fully hydrolyzed. This was not the case for the bovine and human trypsins. For some CSs, the hydrolysis rate constants of the bovine and human trypsins were low, meaning that they were not yet hydrolyzed within the experiment time. In other cases, the molar sequence coverages of some parts of the sequences were not complete, resulting in low concentrations of CS products. In addition, for the bovine and human trypsin hydrolysates, $\sim 12\%$ of the intact α-LA was not hydrolyzed at the end of the hydrolysates. This means that the remaining nonhydrolyzed CSs were in the remaining intact α-LA.

From the CS product formation kinetics, the hydrolysis rate constants ($k_{\text{PT},\text{HT},\text{AT},\text{f},\text{c},\alpha-LA}$) of the three enzymes toward the CSs in α-LA were determined (Table 4). Differences between the enzymes were seen, for instance, in the hydrolysis kinetics of CS K62–63. This CS was efficiently hydrolyzed by porcine trypsin ($1.7 \times 10^{-3} \text{ s}^{-1} \text{m}^{-1}$). However, it was not seen hydrolyzed by bovine trypsin at all and was only slowly hydrolyzed by human trypsin ($0.07 \times 10^{-3} \text{ s}^{-1} \text{m}^{-1}$).

Overall, the total hydrolysis rate constants of porcine trypsin for the CSs in α-LA were $\sim 2$ times higher than those of bovine and human trypsins. The $k_{\text{PT},\text{H},\text{A},\text{CS}}$ for bovine trypsin toward the CSs in α-LA was equally distributed in the range of $10^{-5}$ to $10^{-4}$, and $10^{-3} \text{ s}^{-1} \text{m}^{-1}$. A range similar to that of the $k_{\text{HT},\text{H},\text{A},\text{CS}}$ was found for the $k_{\text{HT},\text{H},\text{A},\text{CS}}$ of human trypsin. Unlike these two trypsins, for porcine trypsin, 10 out of the 13 CSs had similar hydrolysis rate constants, which were all $\geq 10^{-3} \text{ s}^{-1}$. mg{\text{enzyme}}^{-1}$. For porcine and human trypsins, the hydrolysis rate constants of the CSs in α-LA were also calculated based on the concentrations of the remaining nonhydrolyzed CSs ($C_{\text{remain}}$) in the hydrolysates. For the CSs in the parts of the AA sequences where the molecular sequence coverages were high, the hydrolysis rate constants calculated based on the two methods have an average standard error of 26%. This was 10% higher than the average standard error of hydrolysates in the duplicate experiments (16%). The hydrolysis rate constants of the CSs calculated using both the methods were used to further calculate the enzyme selectivity.

Using the same approach as that used in the peptide analysis of the α-LA hydrolysates, the hydrolysis kinetics, peptide release kinetics, and hydrolysis rate constants of the three trypsins toward CSs in β-cas were analyzed.

Peptide Analysis of β-Cas Hydrolysis by the Three Trypsins. Peptide Annotation of the Hydrolysis of β-Cas by the Three Trypsins. Based on the estimated numbers of HSSs, the predicted maximum numbers of peptides (NPP_{\text{max}}) were 55, 120, and 78 for the bovine, porcine, and human trypsin hydrolysates of β-cas, respectively. Experimentally, 34, 35, and 40 unique peptides were annotated in the bovine, porcine, and human trypsin hydrolysates, respectively (Table 2 and Appendix B). Similar to those for α-LA hydrolysis, the annotated numbers of peptides were lower than predicted.
Quality Check of the Peptide Analysis of β-Cas Hydrolysis by the Three Trypsins. Overall, for all the β-cas hydrolysates, the average AA, peptide, and molar sequence coverages were in the same ranges as those for α-LA (Table 3 and Appendix D). Only the porcine tryptic hydrolysates at a DH of 6.7% had an AA sequence coverage less than 100%. This was because this hydrolysate mainly contained final peptides, of which the dipeptides β-cas[98–99] and β-cas[106–107] were not found in the chromatograms. Because dipeptides have low molar extinction coefficients, the loss of UV signal was not visible from the total UV signal. The quality of peptide quantification was further confirmed by the agreement between the calculated DHstat of each β-cas hydrolysate and DHstat (Figure 3B). Averaged over all the samples, the standard error between the DHstat and DHMS was ~7.7%. In the bovine, porcine, and human tryptic hydrolysates of β-cas, peptides with four, five, and four α-specific cleavages were found, respectively. Because of the low concentrations of CS products for these sites (≤8 μM), they were not included as CSs.

Peptide Annotation of β-Cas Hydrolysis by the Three Trypsins. The differences in the peptide release kinetics of the hydrolysates of β-cas by the three trypsins were compared (Appendix F). This was illustrated for the peptides β-cas[108–113] and β-cas[183–209] (Figure 5A,B). As expected, for porcine tryptic hydrolysis, the final peptide β-cas[108–113] formed rapidly in the beginning and reached ~35 μM after 1000 s, which was close to the initial injected concentration (34 μM, Figure 5A). This indicated that the CSs K108–109 and K114–115 were fully hydrolyzed by porcine trypsin. During bovine tryptic hydrolysis, β-cas[108–113] was released gradually, and the concentration reached ~23 μM. For the hydrolysis by human trypsin, β-cas[108–113] was not formed in the first ~3500 s but so rapidly accumulated after this point that it still reached the initial injected concentration (34 μM) before the end of the experiment. Previously, such a delay in the release kinetics of a peptide was attributed to the “demasking” phenomenon. Demasking is when during protein enzymatic hydrolysis, some CSs only become accessible after the cleavage of other peptide bonds in the protein.28 The concentration of the intermediate peptide β-cas[183–209] increased rapidly from t = 0, started to decrease at a DH of 2% in porcine tryptic hydrolysis, and was completely degraded at ~1000 s (Figure 5B). For both bovine and human tryptic hydrolysates, β-cas[183–209] formed gradually during the hydrolysates and started to decrease at ~400 and ~800 s, respectively. There was ~5 μM β-cas[183–209] remaining in the final bovine tryptic hydrolysate, whereas for the other two trypsins, it was completely degraded.

Hydrolysis Rate Constants of the CSs in β-Cas Hydrolysis by the Three Trypsins. During the porcine tryptic hydrolysis, the concentrations of the CS products of K113–114 increased rapidly at the start, reaching the initial injected concentration (34 μM) within ~1500 s (Figure 5C). The hydrolysis of K113–114 by bovine trypsin was slower than that by porcine trypsin. For bovine trypsin, the concentrations of the CS products only reached ~27 μM at the end. For human trypsin, as indicated by the peptide release kinetics of β-cas[108–113], the cleavage of K113–114 was masked in intact β-cas. Indeed, no CS products of K113–114 were formed in the beginning of the hydrolysis. However, after a DH of 4% was reached, the
concentration increased rapidly and reached ∼33 μM (C₀ = 34 μM). This indicated that human trypsin could hydrolyze K113–114 once it was demasked. This also explained why the fit using the second-order reaction equation was not good for this CS. Clear evidence of demasking effects were observed in only four cases (R202–203 for the three enzymes and K113–114 for human trypsin) in all the hydrolyses. The equation for consecutive reactions, as describe previously, was used to calculate the hydrolysis rate constants of the CSs. For CS K183–184, all the trypsins hydrolyzed the CS efficiently, illustrated by the fact that all the CS products at the end of the hydrolyses reached C₀ (Figure 5D). However, porcine trypsin hydrolyzed K183–184 faster than the other two trypsins.

The hydrolysis rate constants (kβ-cas) of the hydrolyses of all the CSs in β-cas by the three enzymes were determined (Table 5). Similar to those of α-LA hydrolysis, the human trypsins, the hydrolysis rate constants of the CSs were also calculated based on the concentrations of the remaining nonhydrolyzed CSs in the hydrolysates. For the CSs that were in the parts of the AA sequences with high molar sequence coverages, the hydrolysis rate constants of the CSs determined with both methods had average standard errors of 28%.

It should be noted that even for the HSSs, the hydrolysis rate constants of the three trypsins toward the CSs in β-cas were higher than those for the CSs in α-LA. This might be due to the differences in the hydrolysis rate constants of the intact proteins (kintact) during the hydrolyses. The kintact values of the three trypsins toward β-cas were 8–14 times higher than their kintact values toward α-LA. This was also reflected in the total hydrolysis rate constants of the two proteins. This then suggests that the measured hydrolysis rate constant of the CSs are the products of the kintact and hydrolysis rate constants of the CSs in the peptides.

### Enzyme Selectivity and Predictions of the Maximum Degrees of Hydrolysis

Based on the hydrolysis rate constants calculated using the concentrations of the CS products and the remaining nonhydrolyzed CSs (Cremaining) in the hydrolysates, the selectivity of the bovine, porcine, and human trypsins toward the CSs in α-LA and β-cas was calculated using eq 13. According to the two methods, for the CSs in the parts of AA sequences where the molecular sequence coverages were high, the selectivity values had an average standard error of 26%. The selectivity values calculated based on the concentrations of CS products were used for further comparisons of the selectivity of each trypsin (Figure 6A,C). The selectivity values of porcine trypsin toward most of the CSs in α-LA were similar, whereas the ranges of selectivity values for bovine and human trypsin selectivity were much larger. This means that the bovine and human trypsins were more selective than porcine trypsin toward CSs in α-LA (Figure 6A). For β-cas hydrolysis, the total hydrolysis rate constant of porcine trypsin hydrolysis was 3 times higher than those for bovine and human trypsin hydrolysates (Figure 6D). Although the ranges of the selectivity values were relatively similar for the hydrolysates by the three trypsins, porcine trypsin efficiently hydrolyzed more CSs than the other two trypsins. This finding was in line with another study, which reported a human serum albumin hydrolysate by bovine trypsin contained ∼60% more peptides with missed CSs in the AA sequence than that by porcine trypsin.

Based on the threshold Lβ-cas/PT,HT,α-LA values from the bovine trypsin selectivity results, the high, intermediate, and low or zero selectivity sites (HSSs, ISSs, and LSSs, respectively) in α-LA and β-cas for the porcine and human trypsins were categorized (Table 6). For the CSs in α-LA, the amounts of HSSs for each trypsin were close to but slightly lower than the predictions based on DHmax,pt which were 6, 12, and 8 for the bovine, porcine, and human trypsins, respectively. This was also observed for β-cas hydrolysis. That lower numbers of HSSs obtained experimentally were due to the fact that the ISSs also contributed to the DHmax,pt. Based on the numbers of HSSs, ISSs, and LSSs, the maximum degrees of hydrolysis (DHmax,pt) predicted based on the enzyme selectivity were calculated using eq 14 and compared with the DHmax,pt (Figure 7). The average standard error (|DHmax,pt/theo − DHmax,exp|/DHmax,exp × 100%) between the DHmax,pt predicted based on the selectivity and the experimental DHmax,exp was ∼14%, which was ∼4 times more accurate than the prediction based on the enzyme specificity (DHmax,thru).

### Table 5. Individual and Total Hydrolysis Rate Constants (±15% Standard Error) of All Cleave Sites (CSs) in β-cas during Hydrolyses by Bovine, Porcine, and Human Trypsins Based on CS Product Concentrations

| enzyme | bovine trypsin | porcine trypsin | human trypsin |
|--------|----------------|----------------|---------------|
| R1     | 0.0 (±0.0)     | 0.0 (±0.0)     | 0.0 (±0.0)    |
| R25    | 1.4 (±0.4)     | 11.1 (±1.3)    | 64.4 (±7.6)   |
| K28    | 16.7 (±0.9)    | 114.8 (±21.9)  | 56.2 (±2.6)   |
| K29    | 1.8 (±0.1)     | 4.4 (±0.3)     | 0.1 (±0.0)    |
| K32    | 1.6 (±0.7)     | 8.0 (±0.14)    | 0.9 (±0.1)    |
| K48    | 0.002 (±0.000) | 0.2 (±0.0)     | 0.7 (±0.1)    |
| K97    | 1.3 (±0.3)     | 53.1 (±3.7)    | 53.7 (±2.6)   |
| K99    | 31.4 (±1.9)    | 17.5 (±0.8)    | 0.5 (±0.0)    |
| K105   | 47.8 (±3.5)    | 92.6 (±21.9)   | 12.1 (±0.3)   |
| K107   | 5.7 (±2.0)     | 5.9 (±1.8)     | 0.7 (±0.2)    |
| K113   | 2.3 (±0.2)     | 10.1 (±0.7)    | 0.1 (±0.0)    |
| K169   | 66.6 (±8.2)    | 191.5 (±13.1)  | 19.4 (±0.7)   |
| K176   | 23.9 (±7.1)    | 98.7 (±23.9)   | 13.9 (±1.1)   |
| R100   | 0.0 (±1.9)     | 17.8 (±2.9)    | 0.2 (±0.1)    |
| R202   | 0.4 (±0.2)     | 2.5 (±0.3)     | 0.04 (±0.0)   |
| total  | 207.6          | 651.2          | 232.9         |

*CSs classified as high, intermediate, and low selectivity sites were highlighted with light gray, gray, and dark gray, respectively.*
Based on the coherent outcomes of $\alpha$-LA and $\beta$-cas hydrolyses by bovine, porcine, and human trypsins, it was concluded that the differences in the DH$_{\text{max,exp}}$ of the hydrolysates by the three trypsins could be explained by the enzyme selectivity differences. Based on the percentages of HSSs out of the total CSs for the two substrates, the bovine, porcine, and human trypsins efficiently hydrolyzed 43, 75, and 50%, respectively.

Links between Binding Site Sequences and the Selectivity of Porcine and Human Trypsins. Using the rules of bovine trypsin selectivity, 57% (16 out of 28) and 50% (14 out of 28) of the CSs for porcine and human trypsin selectivity were correctly predicted, respectively (Table 1). To improve the predictions, the link between the binding site sequences and the selectivity of porcine and human trypsins was studied. The selectivity of porcine and human trypsins toward CSs was not related to the hydrophobicities and molecular weights of the AAs in the binding site sequences. For both trypsins, the standard errors of the molecular weights of the AAs in all the positions in the three categories (HSS, ISS, and LSS) were similar (10–22%) and close to the standard error over all 20 protein AAs (22%). For the hydrophobicities, the absolute values of the standard errors for the hydrophobic moments of all the binding site positions in the three categories were between 95 and 8220%, and the absolute standard error over all 20 protein AA was $\sim$420%. There was a relation between the selectivity of porcine trypsin and the charges of the AAs in the P2 positions. All five LSSs had negatively charged AAs in their P2 positions (Table 7). Out of the 21 HSSs, only 2 CSs ($\alpha$-LA K62 and K98) had negatively charged AAs in their P2 positions (if cysteine was taken as a charged AA). For all the other binding site positions,
no clear differences were found between the HSSs and LSSs. This rule that porcine trypsin efficiently hydrolyses CSs with no negatively charged AAs in the P2 positions (Table 8) explained the porcine trypsin selectivity toward 24 out of the 28 (∼86%) CSs in α-LA and β-cas. Positively charged AAs did not seem to hinder porcine trypsin selectivity because they were not present in the binding site sequences of the LSSs but were mostly present in the binding site sequences of the HSSs.

For human trypsin, it was found that 11 of the 15 ISSs and LSSs had negatively charged AAs in either the P2 or P1′ positions of the binding site sequences (Table 9). For the HSSs, only 3 of the 13 CSs had negatively charged AAs in either the P2 or P1′ positions of the binding site sequences (if cysteine was taken as a charged AA). This means that the selectivity of human trypsin toward 21 of the 28 CSs (∼82%) can be explained by the hindrance of negatively charged AAs in either the P2 or P1′ positions (Table 8). No clear observations were found on the effects of positively charged AAs on human trypsin selectivity. That porcine trypsin was only hindered by a charged AA in one position of the binding site sequence explained why it reached higher DH_{max,exp} than the other two trypsins.

In this article, large differences were found for two proteins (α-lactalbumin and β-casein) in their DH_{max,exp} as well as their hydrolysate compositions after they were hydrolyzed by bovine, porcine, and human trypsins. Based on the full quantitative peptide analysis, porcine trypsin efficiently hydrolyzed ∼78% of the cleavage sites, whereas the values for the bovine and human trypsins were only ∼47 and ∼53%, respectively. These data explained why the porcine tryptic hydrolysis reached ∼90% of the theoretical maximum, whereas the bovine and human tryptic hydrolysates reached ∼54 and 68%, respectively. This outcome means that in vitro protein digestibility determined by porcine trypsin hydrolysis should be ∼2 times higher than that determined by bovine or human trypsin hydrolysis. Knowing the enzyme selectivity, the average standard error of the DH_{max} predictions decreased from 52% (based solely on enzyme specificity) to 13%. Preliminary relations between the amino acid sequences of the binding site positions and the selectivity of each trypsin were proposed, which would allow estimations of the DH_{max,exp} of hydrolysis of other substrate proteins by porcine trypsin.
Table 9. Human Tryptic Hydrolysis Rate Constants and Amino Acid Sequences of Binding Site Positions (P4 to P4′) of Cleavage Sites in α-LA and β-Cas

| cleavage site | $k_{\text{H}}$ (s⁻¹.mg⁻¹.protein⁻¹) | binding site |
|---------------|-----------------------------------|--------------|
| High Selectivity |                                   |              |
| α-LA K10      | 2.7 × 10⁻³                       | E VF R E L K D |
| α-LA K58      | 2.2 × 10⁻³                       | I N K I W C K |
| α-LA K94      | 1.6 × 10⁻³                       | G V K L L D K |
| α-LA K5       | 1.3 × 10⁻³                       | Q L T G E V F |
| α-LA K16      | 1.3 × 10⁻³                       | K D K Y Y G G |
| α-LA K108     | 1.2 × 10⁻³                       | L A H K A L C S |
| β-cas R25     | 6.4 × 10⁻²                       | S I T R I N K K |
| β-cas K28     | 5.6 × 10⁻²                       | R I N K K I E K |
| β-cas K97     | 5.3 × 10⁻²                       | G V S K V K E A |
| β-cas K169    | 1.9 × 10⁻²                       | S Q S K V L P V |
| β-cas K176    | 1.4 × 10⁻²                       | V P Q K A V P Y |
| β-cas K105    | 1.2 × 10⁻²                       | M A P K H K E M |
| β-cas R183    | 1.0 × 10⁻²                       | P Y Q R O D M P I |
| Intermediate Selectivity |                      |              |
| α-LA K98      | 8.5 × 10⁻⁴                       | I L D K V G I N |
| α-LA K79      | 3.2 × 10⁻⁵                       | S C D K F L D D |
| β-cas K93     | 9.3 × 10⁻⁵                       | M C V K K I L D |
| Low/zero Selectivity |                      |              |
| α-LA K13      | 2.3 × 10⁻⁴                       | R E L K D L K G |
| α-LA K62      | 6.9 × 10⁻⁵                       | I W G K D D Q N |
| α-LA K114     | 3.6 × 10⁻⁶                       | C S E K L D Q W |
| α-LA K122     | 2.2 × 10⁻⁶                       | L C E K L L － － |
| β-cas K107    | 7.4 × 10⁻⁵                       | P K H K E M P F |
| β-cas K48     | 7.0 × 10⁻⁴                       | L Q D K H I H P F |
| β-cas K113    | 8.0 × 10⁻⁵                       | P F P P K Y P V E |
| β-cas K29     | 7.1 × 10⁻⁵                       | I N K K I E K F |
| β-cas K99     | 5.1 × 10⁻⁵                       | S K V K E A M A |
| β-cas R202    | 3.8 × 10⁻⁵                       | G P V R G P F P |
| β-cas R1      | 0                                  | － － － － － － |

“α” Amino acids with negatively charged side chains (DE) are marked with red. Amino acids with positively charged side chains (KR) are marked with orange except when they are in P1. Cysteines are marked with blue. A dash (—) means that there is no amino acid present. The green boxes emphasize the P2 and P1’ positions of the binding sites.

“Prophosphoserine.” Masked cleavage site in the intact protein.

These trypsins. This work can help to explain the differences in the reported in vitro digestibilities of proteins and define expectations of the outcome of future experiments using these trypsins. It is a step forward in obtaining a better understanding of the hydrolysis process.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.8b00679.

List of peptides derived from α-LA hydrolyzed by the three trypsins; list of peptides derived from β-cas hydrolyzed by the three trypsins, molar sequence coverages of α-LA hydrolysates catalyzed by bovine, porcine, and human trypsins at all DH values, molar sequence coverages of β-cas hydrolysates catalyzed by bovine, porcine, and human trypsins at all DH values, list of peptide concentrations at all DH values in α-LA hydrolyzed by porcine and human trypsins, list of peptide concentrations at all DH values in β-cas hydrolyzed by porcine and human trypsins (PDF)

**AUTHOR INFORMATION**

Corresponding Author

*Tel.: +31 317483786. E-mail: peter.wierenga@wur.nl.

ORCID

Yuxi Deng: 0000-0002-9772-0340

Peter A. Wierenga: 0000-0001-7799-6676

Funding

This work is supported by NanoNextNL, a micro- and nanotechnology consortium of the government of The Netherlands and 130 partners. This work is financially supported by FrieslandCampina, The Netherlands.

Notes

The authors declare no competing financial interest.

**ABBREVIATION USED**

DH, degree of hydrolysis; DHmax, theoretical maximum degree of hydrolysis; DHmax exp, experimental maximum degree of hydrolysis; DHmax fit, fitted maximum degree of hydrolysis; kH, overall hydrolysis rate constant; kH int, hydrolysis rate constant of intact proteins; kH cleavage, hydrolysis rate constant of cleavage sites; C0, initial injected protein concentration; HSS, high selectivity site; ISS, intermediate selectivity site; LSS, low or zero selectivity site; CS, cleavage site; AA, amino acid; BT, bovine trypsin; PT, porcine trypsin; HT, human trypsin; A134, UV peak area at 214 nm; RP-UPCL, reversed-phase ultrahigh-performance liquid chromatography; ESI-Q-TOF, MS, electron-spray-ionization time-of-flight mass spectrometry; α-LA, α-lactalbumin; β-cas, β-casein

**REFERENCES**

(1) Deng, Y.; Wierenga, P. A.; Schols, H. A.; Sforza, S.; Gruppen, H. Effect of Maillard induced glycation on protein hydrolysis by lysine/arginine and non-lysine/arginine specific proteases. *Food Hydrocolloids* 2017, 69, 210–219.

(2) Deng, Y.; van der Veer, F.; Sforza, S.; Gruppen, H.; Wierenga, P. A. Towards predicting protein hydrolysis by bovine trypsin. *Process Biochem.* 2018, 65, 81–92.

(3) Butré, C. J.; Sforza, S.; Gruppen, H.; Wierenga, P. A. Introducing enzyme selectivity: A quantitative parameter to describe enzymatic protein hydrolysis. *Anal. Bioanal. Chem.* 2014, 406 (24), 5827–5841.

(4) Rinderknecht, H.; Renner, I. G.; Carmack, C. Trypsinogen variants in pancreatic juice of healthy volunteers, chronic alcoholics, and patients with pancreatitis and cancer of the pancreas. *Gut* 1979, 20 (10), 886–891.

(5) Fernández, A.; Riera, F. β-Lactoglobulin tryptic digestion: A model approach for peptide release. *Biochem. Eng. J.* 2013, 70, 88–96.

(6) Walmsley, S. J.; Rudnick, P. A.; Liang, Y.; Dong, Q.; Stein, S. E.; Nesvizhskii, A. I. Comprehensive analysis of protein digestion using six trypsins reveals the origin of trypsin as a significant source of variability in proteomics. *J. Proteome Res.* 2013, 12 (12), 5666–5680.

(7) Keil, B. Specifity of proteolysis; Springer-Verlag: Berlin, 1992.

(8) Vandermauliere, E.; Mueller, M.; Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrom. Rev.* 2013, 32 (6), 453–465.

(9) Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 1967, 27 (2), 157–162.

(10) Dallas Johnson, K.; Clark, A.; Marshall, S. A functional comparison of ovine and porcine trypsins. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* 2002, 131 (3), 423–431.

(11) Adler-Nissen, J. Enzymic hydrolysis of food proteins; Elsevier Applied Science Publishers: London, UK, 1986.

(12) Butré, C. J.; Wierenga, P. A.; Gruppen, H. Influence of water availability on the enzymatic hydrolysis of proteins. *Process Biochem.* 2014, 49, 1903–1912.
(13) Butré, C. I. Introducing enzyme selectivity as a quantitative parameter to describe the effects of substrate concentration on protein hydrolysis. Ph.D. Thesis, Wageningen University & Research, Wageningen, The Netherlands, 2014.

(14) Kuipers, B. J. H.; Gruppen, H. Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. J. Agric. Food Chem. 2007, 55 (14), 5445–5451.

(15) Bruin, G. J. M.; Stegeman, G.; Van Asten, A. C.; Xu, X.; Kraak, J. C.; Poppe, H. Optimization and evaluation of the performance of arrangements for UV detection in high-resolution separations using fused-silica capillaries. Journal of Chromatography A 1991, 559 (1–2), 163–181.

(16) Chervet, J. P.; Van Soest, R. E. J.; Ursem, M. Z-shaped flow cell for UV detection in capillary electrophoresis. Journal of Chromatography A 1991, 543 (C), 439–449.

(17) Moring, S. E.; Reel, R. T.; van Soest, R. E. J. Optical improvements of a Z-shaped cell for high-sensitivity UV absorbance detection in capillary electrophoresis. Anal. Chem. 1993, 65 (23), 3454–3459.

(18) Kosters, H. A.; Wierenga, P. A.; de Vries, R.; Gruppen, H. Characteristics and effects of specific peptides on heat-induced aggregation of β-lactoglobulin. Biomacromolecules 2011, 12 (6), 2159–2170.

(19) Linderstrøm-Lang, K. The initial phases of the enzymatic degradation of proteins. Bull. Soc. Chim. Biol. 1953, 35 (1–2), 100–116.

(20) Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C.; Wilcox, W. Hydrophobic moments and protein structure. Faraday Symp. Chem. Soc. 1982, 17 (0), 109–120.

(21) Caessens, P. W. J. R.; Daamen, W. F.; Gruppen, H.; Visser, S.; Voragen, A. G. J. J. Agric. Food Chem. 1999, 47 (8), 2980–2990.

(22) Kamau, S. M.; Lu, R.-R. The effect of enzymes and hydrolysis conditions on degree of hydrolysis and DPPH radical scavenging activity of whey protein hydrolysates. Current Research in Dairy Sciences 2011, 3, 25–35.

(23) Cheison, S. C.; Leeb, E.; Toro-Sierra, J.; Kulozik, U. Influence of hydrolysis temperature and pH on the selective hydrolysis of whey proteins by trypsin and potential recovery of native alpha-lactalbumin. Int. Dairy J. 2011, 21 (3), 166–171.

(24) Panyam, D.; Kilara, A. Emulsifying peptides from the tryptic hydrolysis of casein. J. Food Sci. 2004, 69 (3), FCT154–FCT163.

(25) Butré, C. I.; Sforza, S.; Wierenga, P. A.; Gruppen, H. Determination of the influence of the pH of hydrolysis on enzyme selectivity of Bacillus licheniformis protease towards whey protein isolate. Int. Dairy J. 2015, 44, 44–53.

(26) Butré, C. I.; Sforza, S.; Gruppen, H.; Wierenga, P. A. Determination of the influence of substrate concentration on enzyme selectivity using whey protein isolate and Bacillus licheniformis protease. J. Agric. Food Chem. 2014, 62 (42), 10230–10239.

(27) Butré, C. I.; Buhler, S.; Sforza, S.; Gruppen, H.; Wierenga, P. A. Spontaneous, non-enzymatic breakdown of peptides during enzymatic protein hydrolysis. Biochim. Biophys. Acta, Proteins Proteomics 2015, 1854 (8), 987–994.

(28) Vorobiev, M. M.; Paskonova, E. A.; Vitt, S. V.; Belikov, V. M. Kinetic description of proteolysis. Part 2. Substrate regulation of peptide bond demasking and hydrolysis. Liquid chromatography of hydrolysates. Nahrung 1986, 30 (10), 995–1001.

(29) Vorobiev, M. M.; Butré, C. I.; Sforza, S.; Wierenga, P. A.; Gruppen, H. Demasking kinetics of peptide bond cleavage for whey protein isolate hydrolysed by Bacillus licheniformis protease. J. Mol. Catal. B: Enzym. 2016, 133, S426–S431.