Food Research International 165 (2023) 112485

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

The path of proteolysis by bovine chymotrypsin

Gijs J.C. Vreeke, Jean-Paul Vincken, Peter A. Wierenga *

Laboratory of Food Chemistry, Wageningen University & Research, P.O. 17, 6708 AA Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Label-free quantification
Biochemistry
Peptide release kinetics
Protease selectivity
Subsite model
Substrate inhibition
Digestive enzyme
Food digestion

ABSTRACT

Chymotrypsin is one of the major proteases in intestinal protein digestion. Observations about the type of bonds that are hydrolysed (specificity and preference) were in the past derived from the peptide composition after digestion or hydrolysis rates of synthetic peptides. In this study, the path of hydrolysis by bovine chymotrypsin, *i.e* formation and degradation of peptides, were described for α-lactalbumin, β-lactoglobulin and β-casein. The peptide compositions, determined with UPLC-PDA-MS at different time points were used to determine the digestion kinetics for individual cleavage sites. It was evaluated how statements on (secondary) specificity from literature were reflected in the release kinetics of peptides. β-Lactoglobulin reached the highest degree of hydrolysis (10.9 ± 0.1 %) and was hydrolysed fastest (28 ± 1 $mM_{peptide\ bonds}/s/mM_{enzyme}$), regardless of its globular (tertiary) structure. Chymotrypsin showed a preference towards aromatic amino acids, methionine and leucine, but was also tolerant to other amino acids. For the cleavage sites within this preference,~73% of the cleavage sites were hydrolysed with high or intermediate selectivity. For the missed cleavages within the preference, 45 % was explained by hindrance of proline, which affected hydrolysis only when in positions *P*3, *P*1′ or *P*2′. No clear indication (based on primary structure) was found to explain the other missed cleavages. A few cleavage sites were hydrolysed extremely efficient in α-lactalbumin (F9, F31, W104) and β-casein (W143, L163, F190). This study gave unique and quantitative insight in peptide formation and degradation by chymotrypsin in the digestion of proteins. The approach used showed potential to explore the path of hydrolysis for other proteases with less defined specificity.

1. Introduction

Chymotrypsin is commonly known for the involvement in intestinal protein digestion in many animals species (Gorrill & Thomas, 1967; Gorrill & Friend, 1970; Guyonnet, Tłuscik, Long, Polanowski, & Travis, 1999) and humans (Feinstein, Hofstein, Koifmann, & Sokolovsky, 1974). Nowadays, the protease is also used in the production of bioactive peptides (Kamath, Niketh, Chandrashekar, & Rajini, 2007; Srinivas & Prakash, 2010) and as alternative to trypsin for proteomics (Giansanti, Tsiatsiani, Low, & Heck, 2016). In the 20th century, many studies in biochemistry have been performed to unravel the hydrolysis mechanism of chymotrypsin (Blow, 1976; Steitz, Hendekson, & Blow, 1969). In these studies, parameters for hydrolysis kinetics as k-cat were determined for synthetic (amide) substrates to find the relationship between amino acids around the cleavage site in the substrate and protease activity. These findings about the mechanism were (later) confirmed with X-ray crystallography. These efforts led to several hypotheses about the effect of amino acids in the different binding positions of the protease. Recent advances in mass spectrometry and peptide quantification enabled us to actually measure the hydrolysis rates of cleavage sites during proteolysis instead of analysing synthetic model substrates. The aim of this study is to describe the path of hydrolysis for three milk proteins with varying structures. Current statements on specificity from synthetic substrates will be evaluated during actual hydrolysis of intact proteins.

1.1. Prior knowledge about chymotrypsin's mechanism, specificity and preference

Chymotrypsin is a serine protease, in which three amino acid residues (Ser195, His57, Asp102) form a catalytic triad that facilitates the hydrolysis of the peptide bond via a two-step mechanism. First, a temporary bond (acylation) is made between the peptide bond in the substrate and the serine residue in the catalytic site. Second, the serine residue is de-acylated by releasing the amide part and subsequently the carboxylic acid product. The type of peptide bonds that can be

* Corresponding author.
E-mail address: peter.wierenga@wur.nl (P.A. Wierenga).

https://doi.org/10.1016/j.foodres.2023.112485
Received 13 October 2022; Received in revised form 20 December 2022; Accepted 12 January 2023
Available online 21 January 2023

Table 1
Reported specificity and preference of chymotrypsin.

Assumed specificity / statement about specificity	Preference	Source
Aromatic and hydrophobic residues		(Adler-Nissen, 1986)
F, H, L, M, W, Y	F, W, Y > H, L, M	(Keil, 1992)
A, F, G, H, I, M, S, T, V, W, Y	F, W, Y > M, I, S, V, H, G, A, T	(Burrell, 1993)
F, L, M, W, Y	W > Y > F > M > L	(Lin et al., 1999)
F, K, L, M, R, W	F, W, Y > L, M > K, R	(Hudáky, Kaslik, Venekei, & Gráf, 1999)
Correlates with hydrophobicity		(Hedstrom, 2002)
F, L, M, N, Q, W, Y	W > Y > F > M > L > Q, N	(Vorob'ev, 2013)
F, Y, L		MEROPS database (Rawlings, Waller, Barrett, & Bateman, 2014)
F, H, L, M, W, Y	F, W, Y > H, L, M	ExPASy peptide cutter

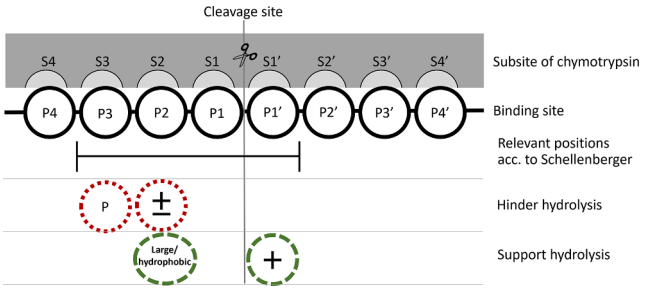

Fig. 1. Illustration of the subsite model of Schechter and Berger with influence of binding site positions on chymotrypsin hydrolysis. + and - indicate positively charged and negatively charged residues, P indicates proline.

hydrolysed (protease specificity), depends on whether the amino acid in the *P*1 position of the substrate fits the S1 pocket of the enzyme. Chymotrypsin consists of a relatively deep S1 pocket, in which the amino acid residue in the *P*1 position can make hydrophobic interactions with for instance Ser190, Cys191, Cys220, Val213, Trp215 and Tyr228 (Blow, 1976; Steitz et al., 1969). Considering these interactions, it is generally believed that chymotrypsin has a specificity for hydrophobic and aromatic residues in the *P*1 position. But, the exact reported specificity is broader (Table 1). For instance, occurrences of hydrolysis by chymotrypsin were described after methionine, glycine or arginine residues. These observations seem to imply that the S1 pocket of chymotrypsin could accommodate a wider range of amino acid residues. Besides protease specificity, studies report that not all amino acids are hydrolysed with the same likelihood, often described with the term "preference". In literature, the preferred amino acids are mostly phenylalanine (F), tryptophan (W) and tyrosine (Y) (Table 1).

1.2. Observations of secondary specificity

Several studies indicated that not all theoretical cleavage sites within the assumed preference were hydrolysed by chymotrypsin. For example, Galvão *et al.* described that only 56 % of the theoretical cleavage sites were hydrolysed in whey proteins (Galvão, Souza Silva, Custódio, Monti, & Giordano, 2001). Deng *et al.* observed that 54 % of the theoretical maximum degree of hydrolysis ($DH_{max,\ theo}$) was reached for chymotrypsin hydrolysis of apo alpha-lactalbumin (Deng, Butré, & Wierenga, 2018). In a proteomics study by Giansanti *et al.* a similar observation was done. They reported that 80 % of the peptides identified in chymotrypsin digests of *E. coli* lysate still contained 1–4 intact cleavage sites (Giansanti et al., 2016). Possibly, these observations could be explained by neighbouring amino acids that hinder the hydrolysis of the cleavage sites, also called "secondary specificity".

The concept of secondary specificity was introduced with the subsite model of Schechter and Berger (Schechter & Berger, 1967) (Fig. 1). Schellenberger *et al.* described in 1989 that chymotrypsin hydrolysis was influenced by the amino acids in the *P*3-*P*1′ positions (Schellenberger, Schellenberger, Mitin, & Jakubke, 1990). A study from 1972 using synthetic substrates, observed that a proline residue in the *P*3 position hindered hydrolysis (Segal, 1972). In 1991, Schellenberger *et al.* performed a Quantitative Structure-Activity Relationship (QSAR) study on chymotrypsin hydrolysis of synthetic substrates tested in literature (Schellenberger, Braune, Hofmann, & Jakubke, 1991). The model showed preference for large residues in the *P*2 position, no effect of proline in the *P*2 position and confirmed the earlier observed hindrance by proline in the *P*3 position. Hedstrom reported a slight preference for hydrophobic amino acid residues in the *P*2 position and a preference for positively charged residues in the *P*1′ position (Hedstrom, 2002). Another study, focussing on the influence of the amino acid in the *P*1′ position in relation to that in the *P*1 position, also described the positive contribution of lysine and arginine in the *P*1′ position (Keil, 1992). The presence of a charged residue in the *P*2 position would negatively affect the probability of hydrolysis (Keil, 1992).

Besides the primary substrate structure, a study of Wright suggested a conformational specificity, in which the tertiary structure of the amino acids in the *P*3 and *P*4 position would influence the hydrolysis (Wright, 1977). Similarly, research of Vorob'ev *et al.* suggested that some peptide bonds in proteolysis have to be demasked before other bonds become accessible to chymotrypsin (Vorob'ev, 2013).

1.3. Peptide quantification and enzyme selectivity

The previous studies on chymotrypsin determined hydrolysis kinetics for synthetic substrates or analysed the peptides formed after proteolysis. For the latter, the identified peptides were used to determine the preference by analysing the amino acids in all cleavage sites at which hydrolysis was observed, sometimes corrected for frequency of occurrence. Typically, the concentrations of the peptides formed were not measured and considered. In this manuscript, all individual peptides (at each timepoint) will be absolutely and label-free quantified, based on UV absorbance and predicted molar extinction coefficient according to Kuipers *et al.* (Kuipers & Gruppen, 2007). The peptide composition at

Table 2
Protein materials used in this study with DH_{max} for chymotrypsin hydrolysis.

Protein	Uniprot code	N-factor[1] [g of protein / g N]	Protein content [w/w]	Protein purity	h_{tot}[2] [mmol/g protein]	F, Y, W in sequence (%)	DH_{max} experimental at E:S ratio 1:25
α-LA	P00711	6.25	93 %	90 %	8.6	9.8 %	8.4 ± 0.3 %
β-LG	P02752	6.29	96 %	100 %	8.8	6.2 %	10.9 ± 0.1 %
β-cas	P02666	6.39	90 %	90 %	8.7	6.7 %	8.2 ± 0.3 %
CT	P00766	5.99	86 %	100 %			

[1] From Uniprot (https://www.uniprot.org).
[2] h_{tot} is the amount of peptide bonds per gram of protein calculated with the amino acid sequence and molecular weight of the protein reported in Uniprot.

different time points allows us to determine the hydrolysis rate of each individual cleavage site in relation to the total hydrolysis rate, previously introduced as protease selectivity (Butré, Sforza, Gruppen, & Wierenga, 2014). The determination of the selectivity makes it possible to see whether previous observations about secondary specificity based on synthetic substrates are also valid for real protein substrates. Ultimately, the aim is to evaluate whether the path of proteolysis of any protein by chymotrypsin, and the respective peptides released, can be predicted.

2. Materials and methods

2.1. Materials

α-Lactalbumin was obtained from Davisco Foods International Inc. (Le Sueur, MN, USA). The protein isolate was further treated into apo α-Lactalbumin (α-LA) by removal of the calcium ions with EDTA, similarly done as described by Deng *et al.* (Deng, Butré, et al., 2018). β-Lactoglobulin (β-LG, L0130), β-casein (β-cas, C6905), bovine α-chymotrypsin (C-3142) and aprotinin from bovine lung (A6279) were purchased from Sigma-Aldrich (St. Louis, MO, USA). The protein contents of the protein samples were 93 % (w/w) for α-LA, 96 % (w/w) for β-LG and 90 % for β-cas and 86 % (w/w) for chymotrypsin (based on Dumas) (Table 2). The protein purities were 90 % for α-LA, 100 % for β-LG and 90 % for β-cas. The other proteins in α-LA were β-lactoglobulin and bovine serum albumin. The other proteins in β-cas were other type of caseins. The chymotrypsin did not contain any trypsin, according to SDS-PAGE and UPLC-PDA-MS measurements. According to the supplier, the chymotrypsin had an activity of ≥ 40 *N*-benzoyl-L-tyrosine-ethyl-ester units/mg protein and was treated with Nα-Tosyl-L-lysine chloromethyl ketone hydrochloride (TLCK) to inactivate residual trypsin activity. Aprotinin was present at 2.3 mg/mL, based on previously determinations with RP-UPLC (Deng, van der Veer, Sforza, Gruppen, & Wierenga, 2018).

2.2. Methods

2.2.1. Protein hydrolysis by chymotrypsin

Hydrolysis was performed for 3 h with a pH-stat device (Metrohm, Herisau, Switzerland) equipped with 0.2 M NaOH. α-LA, β-LG and β-cas were dissolved at 1 % (w/v) in Millipore water. Afterwards, the pH was adjusted to 8.0 and the solutions were equilibrated for 30 min at 37 °C. Hydrolysis was performed at enzyme to substrate ratios of 1:100 and 1:25, both in duplicate. Samples (200 μL) were taken before and during the hydrolysis at various time points, after which aprotinin was immediately added to inactivate the chymotrypsin. To the samples with an E:S ratio of 1:100, 6 μL of aprotinin was added and to the samples at E:S ratio of 1:25, 24 μL of aprotinin was added. The inactivation of chymotrypsin activity was confirmed by addition of aprotinin to a (ongoing) hydrolysis in the pH-stat after 1 min, in which the increase in DH stopped at moment of addition. Hydrolyses without enzyme were performed to exclude potential consumption of NaOH due to CO_2 dissolution. The degree of hydrolysis was calculated using equation (1).

$$DH_{stat}[\%] = V_b \times N_b \times \frac{1}{\alpha} \times \frac{1}{m_p} \times \frac{1}{h_{tot}} \times 100\% \quad (1)$$

where V_b [mL] is the volume of added NaOH; N_b [mol/L] is the normality of NaOH; α is the average degree of dissociation of the α-NH group ($1/\alpha = 1.3$ at 37 °C and pH 8 (Butré, Wierenga, & Gruppen, 2014)); m_p [g] is the amount of protein in solution; h_{tot} [mmol/g] is the number of peptide bonds per gram of protein. The DH in time was fitted with second order reaction kinetics to determine the overall hydrolysis rate constant (k_h) (equation (2)). The k_h was used to calculate the turnover number [$mM_{peptidebonds}$ /s/mM_{EZ}], which was the amount of peptide bonds in solution hydrolysed by an equal molar amount of enzyme each second.

$$DH_{fit}[\%] = DH_{max,fit} - DH_{max,fit}/(1 + k_h \times DH_{max,fit} \times t) \quad (2)$$

2.2.2. Sample preparation

The protein hydrolysates were mixed (1:1) [v/v] with 20 mM dithiothreitol (DTT) and 100 mM Tris-HCl buffer at pH 8.0 and incubated for 2 h to reduce disulphide bonds. Afterwards, the intact protein samples were diluted 10x in eluent A and the hydrolysates were diluted 3x [v/v] in eluent A. The samples were centrifuged (10 min, 14,000 × g, 20 °C) and the supernatant was injected on the UPLC-MS. The endpoint hydrolysate of β-LG was injected three times to estimate reproducibility.

2.2.3. Reverse phase ultra-high performance liquid chromatography (RP-UPLC)

The peptides were analysed on an Acquity Premier UPLC coupled to a Photodiode Array (PDA) detector, both from Waters (Milford, MA, USA). A gradient was used of 1 % ACN + 0.1 % TFA in UPLC-grade water (eluent A) and 1 % UPLC-grade water + 0.1 % TFA in ACN (eluent B). The gradient was 0–2 min isocratic on 3 % B; 2–10 min linear gradient from 3 to 22 % B; 10–16 min linear gradient 22–30 % B; 16–21 min linear gradient 30–100 % B; 21–26 min isocratic on 100 % B; 26–28 min linear gradient 100–3 % B and 28–32 min isocratic on 3 % B. The column was the Acquity Premier peptide column BEH C18 2.1 mm *150 mm 300 Å 1.7 μm, thermostated at 30 °C. The applied flow was 350 μL/min and the volume injected was 4 μL.

2.2.4. Electro spray ionisation time of flight mass spectrometry

The mass spectra were acquired from 50 to 3000 *m/z* with the Select Series Cyclic IMS (Waters). The machine was operated in Time-of-flight and V-mode without using the ion mobility dimension. Ionisation was done with the electrospray ionisation source at capillary voltage of 2.5 kV and a temperature of 150 °C. The sample cone was operated at 40 V and nitrogen was used as desolvation gas (500 °C, 800 L/h) and cone gas (200 L/h). Lock mass data was acquired by injection of 50 pg/μL Leucine-Enkephalin at 10 μL/min, at a capillary voltage of 2.7 kV. An automatic quadrupole profile was applied. Fragmentation was done in the trap-cell by changing the voltage for MS (6 V) into a MSe ramp from 28 V to 56 V.

2.2.5. Peptide identification

Peptides were identified using the recently developed guidelines and methodology of Vreeke *et al.* with UNIFI software version 1.8 (Vreeke, Lubbers, Vincken, & Wierenga, 2022). Processing was done semi-specific with a specificity of F, Y, W, M, H, L and (fully) a-specific. The semi-specific processing option identified peptides that matched chymotrypsin's specificity on at least the N- or C-terminal side, whereas the a-specific digest considered potential hydrolysis of all peptide bonds. The a-specific method was used in the end for all analyses. The threshold set in peak detection was 1000 counts for MS peaks and 250 counts for MS/MS peaks. There was no maximum set to the number of included MS and MS/MS signals. The threshold on mass error was 10 ppm for the MS and 20 ppm for MS/MS fragments. The requirements on number of MS/MS fragments to include a peptide were similar as described before (Vreeke et al., 2022). The limit of detection (LOD) and limit of annotation (LOA) were determined by injection of a concentration series (2.5 mg/L to 5 g/L) of a tryptic α-LA hydrolysate. For this mass spectrometer under the applied conditions, the average LOD for a peptide was $1.6 \cdot 10^4$ Counts and the average LOA was $2.1 \cdot 10^5$ Counts. Peptides consisting of >41 amino acids and intact proteins were not detected in the a-specific analysis, and therefore obtained from the semi-specific processing result. Adducts and in-source fragments that were recognised in UNIFI were excluded from the analysis. In-source fragments that were not recognised as such were removed using PeptQuant, an in-house written script in Matlab v2018b. Annotations were considered as in-source fragment if the peptide and potential in-source fragment eluted at the same retention time, the in-source fragment included the same

sequence as the peptide and the in-source fragment had a lower MS intensity than the peptide.

2.2.6. Peptide quantification

Quantification of the peptides was done using the UV absorbance at 214 nm. The UV peak areas were integrated in Masslynx software version 4.2 using the parameters as described in Vreeke *et al.* (Vreeke et al., 2022). The areas were converted into absolute peptide concentrations, $C_{peptide}$ [μM], using the molar extinction coefficients, predicted based on Kuipers *et al.* (Kuipers & Gruppen, 2007) and the calculation of Butré *et al.* (equation (3)).

$$C_{peptide}\ [\mu M] = \frac{A_{214} \bullet Q}{\varepsilon_{214} \bullet l \bullet V_{inj} \bullet k_{cell}} \tag{3}$$

where A_{214} [μAU min] is the UV peak area at 214 nm, V_{inj} [μL] is the volume of sample injected, Q [μL min^{-1}] is the flow rate, ε_{214} [L Mol^{-1} cm^{-1}] the predicted molar extinction coefficient of the peptide and l [cm] is the path length of the UV cell, which is 1 cm according to the manufacturer. The cell constant, k_{cell} for the UV detector was 0.78 (Vreeke et al., 2022).

2.2.7. Completeness of analysis

The completeness of the peptide identification was evaluated with the amino acid sequence coverage and the peptide sequence coverage (equation (4) & (5)), the completeness of quantification was evaluated with the protein recovery and the molar sequence coverage (equation (6) & (7)). All these parameters were introduced by Butré *et al.* (Butré et al., 2014).

$$AA\ seq.\ cov.\ [\%] = \frac{\#unique\ annotated\ AA}{\#AA\ in\ protein\ sequence} \bullet 100\% \tag{4}$$

$$Peptide\ seq.\ cov.\ [\%] = \frac{\#annotated\ AA}{\#annotated\ AA + \#missing\ AA} \bullet 100\% \tag{5}$$

$$Protein\ recovery\ [\%] = \left(\frac{\left(\frac{\sum C_n}{\#AA_{protein}} \right)}{C_0} \right) \bullet 100\% \tag{6}$$

where C_n [μM] is the concentration of each individual AA (n) in the protein sequence, and $\#AA_{protein}$ is the number of amino acids in the initial protein and C_0 [μM] is the initially injected protein concentration.

$$Molar\ seq.\ cov.\ [\%] = \left(1 - \frac{\sqrt{\frac{\sum (C_n - C_0)^2}{(\#AA_{protein} - 1}}}{C_0} \right) \bullet 100\% \tag{7}$$

where C_n [μM] is the concentration of each individual AA (n) in the protein sequence, C_0 [μM] is the initially injected protein concentration and $\#AA_{protein}$ is the number of amino acids in the initial protein.

2.2.8. Description of the endpoint hydrolysate with protease specificity and preference

Protease specificity is defined as the type of amino acid on the carboxylic side next to which the enzyme can hydrolyse peptide bonds. In literature, this is typically determined by analysis of the start and end amino acid residues of the peptides formed after hydrolysis. For comparison, this was done similarly in this study. However, the specificity does not give any information on the probability that cleavage occurred or the extent of hydrolysis. Therefore, protease preference was calculated. This was defined as the proportion to which hydrolysis occurs compared to the theoretical probability based on the frequency of occurrence of the amino acids (equation (8)). When a type of amino acid residue has a preference of 1, it is hydrolysed to the same extent as expected with the frequency of occurrence.

$$Preference[-] = \frac{\left(\frac{\sum C_{endpoint,i}}{\sum C_{endpoint,total}} \right)}{\left(\frac{\#AA_i}{\#AA_{tot} - 1} \right)} \tag{8}$$

where $C_{endpoint,i}$ [μM] is the concentration cleavage site products formed after amino acids of type i in the endpoint hydrolysate, $C_{endpoint,total}$ [μM] is the total amount of cleavage site products formed in the endpoint hydrolysate, $\#AA_i$ is the frequency of occurrence of amino acid residues of type i in the sequence of the substrate, except the C-terminal amino acid. The $\#AA_{tot}$ is the total number of amino acid residues in the sequence.

The concentration of cleavage site products $C_{t,j}$ for cleavage site j at timepoint t was derived from the peptide concentrations (equation (9)).

$$C_{t,j}[\mu M] = \sum \{ C_{peptide}[x-y] | j = x-1 \cup j = y \} \tag{9}$$

where $C_{t,j}$ [μM] equals the sum of all peptide concentrations ($C_{peptide}$) with sequence x-y, which are released after amino acid j or which end by amino acid j.

2.2.9. Determination of the hydrolysis rate constants

For each cleavage site, the formation of cleavages site products in time was fitted using second order reaction kinetics (equation (10)) and the sequential (second order) reaction kinetics (equation (11)).

$$C_{j,t}[\mu M] = 2(C_0 - C_0/(1 + k_{j,app} \times t \times C_0)) \tag{10}$$

$$C_{j,t}[\mu M] = 2(C_0 - \frac{C_0 \times \left(\frac{k_2}{k_2 - k_1} \right)}{1 + C_0 \times k_1 \times t} + \frac{C_0 \times \left(\frac{k_1}{k_1 - k_2} \right)}{1 + C_0 \times k_2 \times t}) \tag{11}$$

where $C_{j,t}$ [μM] is the concentration cleavage site products for cleavage site j at timepoint t, C_0 [μM] is the (expected) protein concentration and $k_{j,app}$ [s^{-1}] is the apparent hydrolysis rate constant for cleavage site j. Both the C_0 value and the apparent hydrolysis rate constant were fitted. The C_0 was fitted because for some cleavage sites the experimental plateau value was lower than C_0. If in these cases the theoretical C_0 would be used, the $k_{j,app}$ would be underestimated. For some cleavage sites, product formation started in a later stage of the hydrolysis and the data could better be described by sequential kinetics. The k_2 of the sequential fit was used when the R^2 improved by 0.10 in comparison with the regular second order fit and the R^2 of the sequential fit was > 0.50.

The apparent hydrolysis rate constants k, or in case of sequential kinetics k_2, were divided by the mass of enzyme used in mg (equation (12)) and used to calculate the selectivity (equation (13)). Cleavage sites were excluded in the selectivity analysis when cleavage site products were observed in (only) one timepoint, in only one of the duplicates hydrolyses or at concentrations below 1 μM.

$$k_i\ \left[s^{-1} mg^{-1} \right] = \frac{k_{i,app}}{m_E} \tag{12}$$

$$Selectivity[\%] = \frac{k_i \times C_0}{\sum (k_i \times C_0)} \times 100\% \tag{13}$$

where $k_{i,app}$ [s^{-1}] is the apparent hydrolysis rate constant calculated with the second order, or sequential order fit, m_E is the mass of enzyme in mg. The C_0 [μM] was derived from the fit.

2.2.10. Clustering of cleavage sites and subsite analysis

The cleavages sites were manually clustered into high selectivity sites (HSS), intermediate selectivity sites (ISS), low selectivity sites (LSS) and no hydrolysis (NH) by analysis of the cleavage site product formation in time. HSS cleavage sites showed formation of $>$ 50 μM of cleavage site products within 3 min of hydrolysis. ISS cleavage sites formed $>$ 50 μM of cleavage site products, but not within 3 min of hydrolysis. LSS showed product formation in multiple timepoints but at concentrations $<$ 50 μM. Cleavages sites were categorised as no

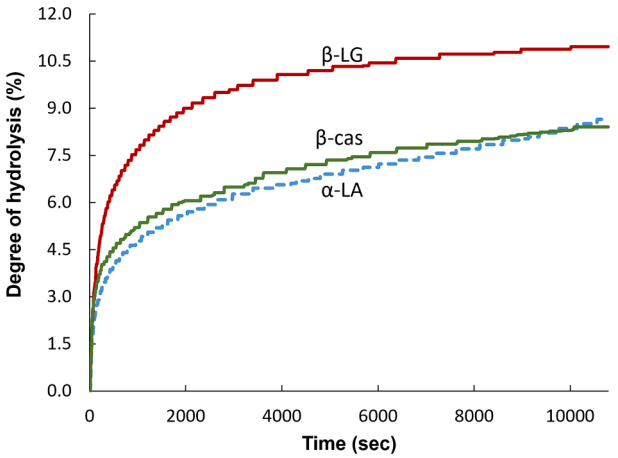

Fig. 2. Degree of hydrolysis in time for 1% α-LA (▬ ▬), β-LG (▬) and β-cas (▬) with bovine chymotrypsin at an E:S ratio of 1:25.

hydrolysis when product formation was < 1 μM, or only observed in one of the duplicates or at only one of the timepoints. After clustering, the amino acids in the *P*4-*P*4′ position of the cleavage sites were analysed in relation to their relative occurrence to the four selectivity clusters. It was taken into consideration that the number of cleavage sites in each cluster was different and whether the amino acid in the *P*1 position was preferred or not. For each rule, it was also tested how often the inverse was correct. To have as many observations as possible the separate analyses of α-LA, β-LG and β-cas were combined.

3. Results & discussion

3.1. Description of the hydrolysis

Based on the frequency of occurrence of the aromatic residues (F, Y, W) a DH_{max} was expected of 9.8 % for α-LA, 6.2 % for β-LG and 6.7 % for β-cas, respectively. Hydrolysis of 1 % α-LA, β-LG and β-cas with chymotrypsin reached degree of hydrolysis values of 6.5 ± 0.1 %, 8.7 ± 0 % and 5.7 ± 0.4 % at an E:S ratio of 1:100. These values did neither match with the DH_{max} values expected based on aromatic residues, nor with the order of the DH_{max} values of the substrates. The assumed preference and amino acid composition did not solely explain the observed degree of hydrolyses. To confirm that the degree of hydrolysis reached was the absolute maximum, the experiment was repeated at higher E:S ratio (1:25). For all substrates a significant increase in DH_{max} was observed, to 8.4 ± 0.3 %, 10.9 ± 0.1 % and 8.2 ± 0.3 %, respectively (Fig. 2). For chymotrypsin, the increase in $DH_{max,\ exp}$ with increasing E:S ratio was previously also reported at hydrolysis of α-LA at low (0.1 % w/v) substrate concentration (Deng, Butré, et al., 2018). Similar shifts in DH_{max} were also observed for trypsin (Deng, Butré, et al., 2018), and pepsin (Ruan, Chi, & Zhang, 2010) and suggested to be the result of the formation of inhibitory peptides (Deng, Butré, et al., 2018; Loveday, Peram, Singh, Ye, & Jameson, 2014). Further experiments on peptide release were all done using the samples at an E:S ratio of 1:25, to minimise the effect of inhibition. The turnover rate of chymotrypsin was slightly higher for β-LG (28 ± 1 mM/s/mM) than for β-cas and α-LA (17 ± 0 and 20 ± 4 mM/s/mM, respectively). This gave a first impression that the activity of chymotrypsin is not affected by protein structure. The turnover rate was in line with the activity provided by the supplier (≥17 mM/s/mM). During the hydrolysis, samples were taken at different time points to determine the peptides present and their concentrations.

3.2. Peptide identification

Peptide identification was performed with semi-specific and a-specific data-processing in UNIFI, to check for possible differences between both processing options. Almost all peptides in the semi-specific processing were also identified with a-specific processing. One peptide sequence was identified with semi-specific processing, which exceeded the maximum length for a-specific processing (α-LA 61–104). The a-specific processing yielded a few fully a-specific peptides as for instance β-LG 71–76 (K)IIAEKT (confirmed with 6 MS/MS fragments). The corresponding *m/z* value was assigned to β-LG 1–6 LIVTQT (3 MS/MS fragments) with semi-specific processing. For this example, the a-specific annotation was clearly a better match to the MS/MS spectrum. The a-specific analysis yielded ~10 % of in-source fragments that were not correctly recognised as such by the software, but were annotated as unique peptide. All these annotations were removed automatically in further analyses, based on sequence and retention time. In the endpoint hydrolysate of β-LG, the 61 unique peptides were identified in three replicate injections. Each replicate contained on average 54 ± 2 peptides. Of the 61 unique entries, 82 % had a MS intensity above the average limit of annotation and 80 % was identified in all three replicates. In total, the hydrolysates of α-LA, β-LG and β-cas consisted of 78, 169 and 141 peptides, respectively, in the various time points (**Supplementary data A**).

3.3. Peptide quantification

All peptides were quantified based on UV absorbance and their predicted molar extinction coefficient. On average, ~94 ± 1 % of the UV peak area present in the α-LA chromatograms was assigned to annotated peptides. For β-LG and β-cas, these values were lower, 84 ± 3 % and 73 ± 7 %, respectively. The peptide concentrations during hydrolysis varied between 0.06 μM and 304 μM, with an average relative standard deviation of 5.7 % on individual peptide concentrations.

3.4. Evaluation of completeness of analysis

The identified peptides of α-LA covered the amino acid sequence for 100 %, for all time points. The amino acid sequence coverages for β-LG and β-cas were on average 97 ± 5 % and 90 ± 8 %, respectively (Table 3). Parts of the protein sequence that were not covered by peptides were for example, [106–122] of β-LG and [7–42] for β-cas. The chromatograms were investigated manually with extracted ion chromatograms for the expected masses, but this did not lead to the identification of these sequences. The average peptide sequence coverages obtained were 96 ± 2 % for α-LA, 92 ± 5 % for β-LG and 86 ± 5 % for β-cas hydrolysates. These coverages were similar to those reported

Table 3
Amino acid sequence coverages, peptide sequence coverages, protein recoveries and molar sequence coverages for the chymotrypsin hydrolysates of α-LA, β-LG and β-cas.

	α-LA		β-LG		β-cas	
	Average + stdev [%]	Min-Max [%-%]	Average + stdev [%]	Min-Max [%-%]	Average + stdev [%]	Min-Max [%-%]
Amino acid sequence coverage (%)	100 ± 0	100	97 ± 5	86–100	90 ± 8	83–100
Peptide sequence coverage (%)	96 ± 2	92–99	92 ± 5	80–98	86 ± 5	79–95
Protein recovery (%)	94 ± 10	70–110	82 ± 14	58–111	71 ± 13	55–105
Molar sequence coverages (%)	80 ± 7	65–89	63 ± 14	36–81	34 ± 14	9–62

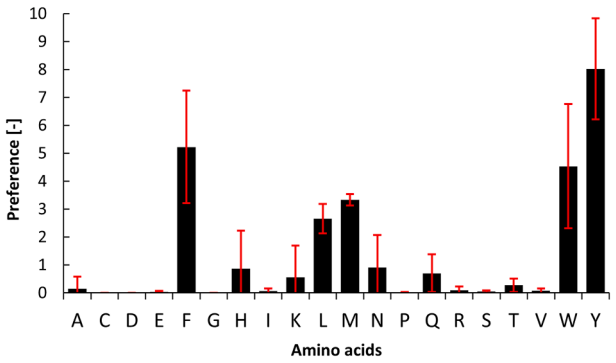

Fig. 3. Preference of chymotrypsin for the amino acid residue in the *P*1 position. The average and standard deviation are shown of the preference values determined against α-LA, β-LG and β-cas.

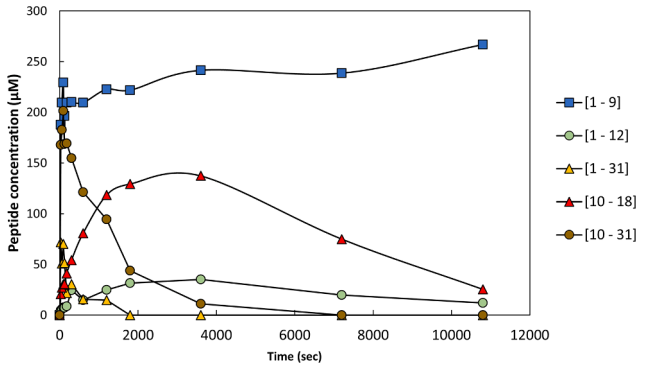

Fig. 4. The absolute concentration in time of a few α-LA peptides to illustrate peptide formation and degradation by chymotrypsin.

previously for tryptic hydrolysates of the same substrates (Vreeke et al., 2022). For α-LA, the average amino acid concentration in a certain position on the sequence was 94 ± 10 % of the injected protein concentration. This resulted in a molar sequence coverage of 80 ± 7 %. For β-LG and β-cas, the average retrieved amino acid concentrations were lower than the injected concentrations, resulting in (average) protein recoveries of 82 ± 14 % for β-LG and 71 ± 13 % for β-cas. As a result, the molar sequence coverages were also lower than observed in the past. For β-LG, this was mainly due to low coverage of sequence region 106–122. For β-cas, sequence region 194–209 was quantified~2x the expected concentration, while other part of the sequence were quantified below the injected (protein) concentration.

3.5. *Chymotrypsin's specificity and preference*

The specificity of chymotrypsin was determined based on the peptides present in the endpoint hydrolysates. Peptides were released by hydrolysis of α-LA peptide bonds after the aromatic amino acids (phenylalanine, tyrosine and tryptophan) but also after asparagine, glutamic acid, leucine, lysine, histidine, methionine, serine, threonine, and valine residues. Besides these amino acids, hydrolysis of β-LG showed peptides released after alanine, aspartic acid, cysteine, glutamine, isoleucine and proline. Based on these observations, chymotrypsin could better be classified as a-specific protease than as specific protease. Within the broad specificity, chymotrypsin showed a preference to hydrolyse cleavage sites with phenylalanine, tryptophan and tyrosine in the *P*1 position. To describe this preference quantitatively, the amount of product released after a certain AA was compared to the amount that one would expect for completely random enzymatic hydrolysis (all AA similarly preferred). For phenylalanine, 5.2 (±2.0) times more hydrolysis products were released than one would expect based on the presence of phenylalanine in the sequences of α-LA, β-LG and β-cas (Fig. 3). Tryptophan and tyrosine had preference values of 4.5 ± 2.2 and 8.0 ± 1.8. Other amino acids that were preferred by chymotrypsin in the *P*1 position were methionine and leucine with preferences of 3.3 ± 0.2 and 2.7 ± 0.5. The residues histidine, lysine, glutamine and asparagine had a

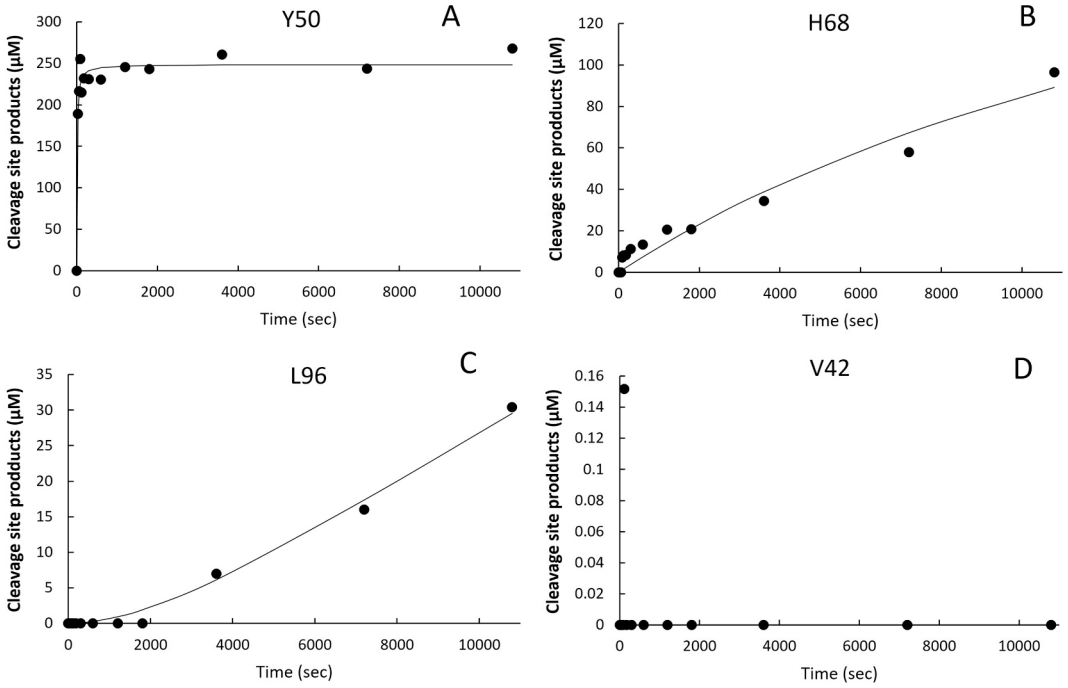

Fig. 5. Release of cleavage site products in time for Y50 (A), H68 (B), L96 (C) and V42 (D) upon chymotrypsin hydrolysis of α-LA. The examples illustrating cleavage sites in the HSS cluster (A), ISS cluster (B), LSS cluster (C) and NH cluster (D). Markers are experimental data, lines are from second-order fit (A,B) and sequential second-order fit (C).

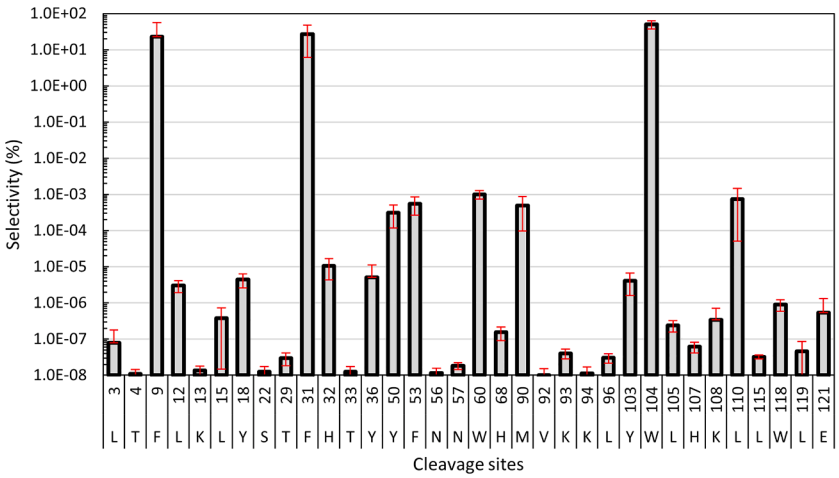

Fig. 6. Selectivity (%) of chymotrypsin for hydrolysed cleavage sites of α-LA, plotted on logarithmic scale.

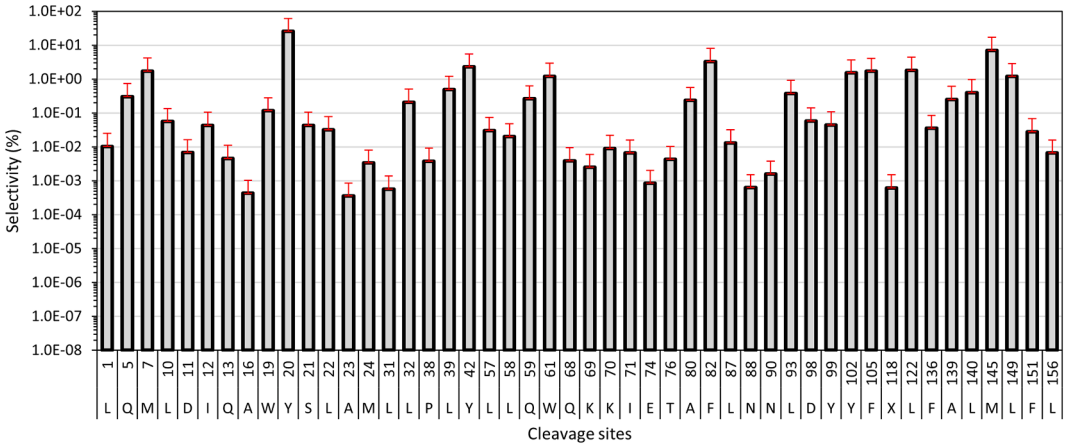

Fig. 7. Selectivity (%) of chymotrypsin for hydrolysed cleavage sites of β-LG, plotted on logarithmic scale.

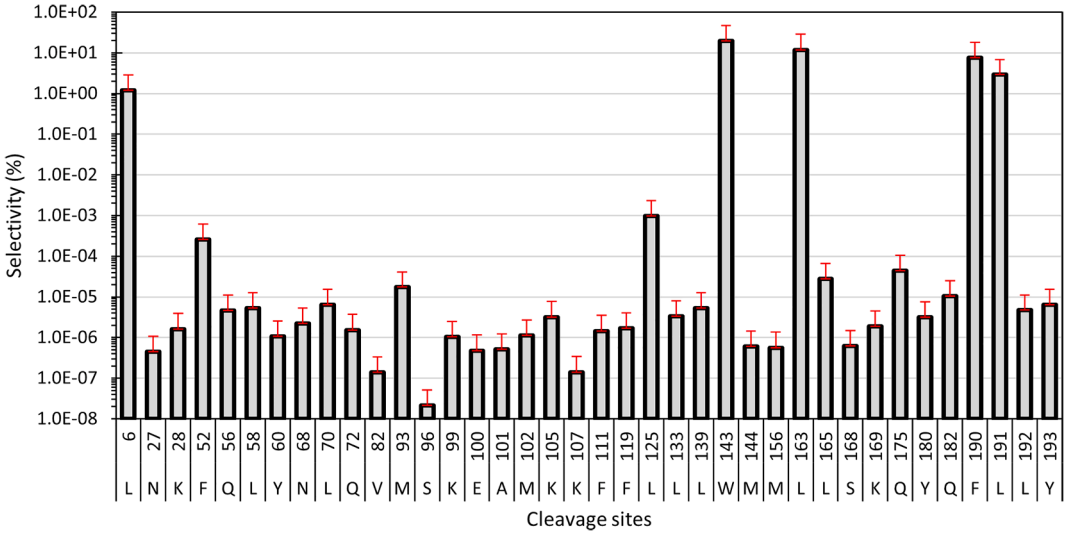

Fig. 8. Selectivity (%) of chymotrypsin for hydrolysed cleavage sites of β-cas, plotted on logarithmic scale.

preference around 1, but the individual preference values for each substrate varied substantially (e.g. hydrolysis of these residues was not observed for all substrates.) The order of the preference values matched reported preferences in literature (Table 1). In order to see whether cleavage sites with similar amino acid in the *P*1 position, were hydrolysed at similar rate, the kinetics of each individual cleavage site had to be determined.

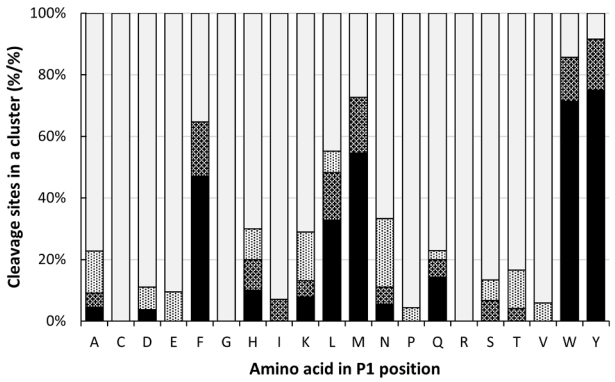

Fig. 9. Occurrence (%/%) of cleavage sites of α-LA, β-LG and β-cas within the clusters containing the high selectivity sites (HSS) (■), intermediate selectivity sites (ISS) (▩), low selectivity sites (LSS) (▨) or not hydrolysed (NH) (□), categorised based on the amino acid in the *P*1 position.

3.6. Hydrolysis rate kinetics of individual cleavage sites

By analysis of the peptides and their concentrations at various timepoints (Fig. 4), the hydrolysis rate constants were calculated for each individual cleavage site. A few cleavage sites in α-LA (F9, F31, and W104) and β-cas (W143, L163 and F190) were hydrolysed extremely fast. For these, >90 % of the cleavage site was hydrolysed in the 30 s^{-1} hydrolysate sample. As a result, all intact proteins were hydrolysed at 90 s^{-1} for α-LA and at 30 s^{-1} for β-cas. The degree of hydrolysis calculated from the peptide composition was for the first time points higher than the degree of hydrolysis based on pH-stat. Theoretically, both should be similar. A possible explanation is that after inactivation, hydrolysis of these cleavage sites continued. Possibly, chymotrypsin had a higher affinity towards these few peptide bonds in α-LA and β-cas than towards the inhibitor (aprotinin). Another hypothesis could be that the enzyme-aprotinin complex dissociated (partly) during sample preparation for LC-MS, leading to a continuation of the hydrolysis of these sites.

The fast depletion of intact protein was also observed in a previous study for α-LA hydrolysed with chymotrypsin (Deng, Butré, et al., 2018). Another study also observed that F31 and W104 in α-LA were hydrolysed relatively early by chymotrypsin compared to other cleavage sites (Catiau, Delval-Dubois, Guillochon, & Nedjar-Arroume, 2011). For β-cas, it has been reported that certain peptide bonds have to be hydrolysed first (de-masked) by chymotrypsin before remaining cleavage sites become accessible (Vorob'ev, 2013). For the cleavage sites, the data were best described by fitting the hydrolysis rate constant and C_0 as variables, since in many cases the maximum concentration (μM) reached was lower than the initial protein concentration. Several cleavage sites were hydrolysed from the start of the incubation and were fully hydrolysed within a few minutes (Fig. 5**A**). For these high selectivity sites (HSS), hydrolysis is efficient and seems independent of the hydrolysis of other peptide bonds. For intermediate selectivity sites (ISS), cleavage site products were released from the start of the hydrolysis and no clear plateau concentration was reached within 3 h (Fig. 5**B**), indicating a much lower rate of hydrolysis by chymotrypsin. For some cleavage sites, the hydrolysis products were not observed in the first few timepoints, but started to release at a later stage of the hydrolysis (Fig. 5**C**). After the lag time the release of cleavage site products increased to substantial concentrations. It seems that the hydrolysis of these cleavage sites is efficient, but requires other peptide bonds to be hydrolysed in advance (de-masking). For these cleavage sites, the sequential kinetics did better describe the data than the normal second order kinetics. For all hydrolysed and non-hydrolysed cleavage sites, the amino acids in the *P*4-*P*4′ positions were analysed to evaluate whether the primary structure can be used to predict the hydrolysis efficiency of a peptide bond during proteolysis.

3.7. Chymotrypsin selectivity and the effect of the P1 position

Analysis of the hydrolysis rate constant of each cleavage site showed differences in the rate of hydrolysis of bonds with similar amino acid in the *P*1 position (Figs. 6-8). For example, some cleavage sites with phenylalanine (F) in the *P*1 position in β-cas were not hydrolysed (F33,

Table 4
Clustering of cleavage sites with proline at different positions of the cleavage site.

		P4	P3	P2	P1	P1'	P2'	P3'	P4'	Total for all cleavages sites
Occurrence P in position		45	45	45	45	45	45	45	45	
Occurrence with FYWML in the P1 position		7	7	13	0	16	8	12	9	
Occurence of a cleavage site in a cluster, when proline occurs in that respective position.	HSS	6	0	9	0	2	1	8	7	59
	ISS	2	0	4	0	2	2	4	3	29
	LSS	2	1	2	2	0	1	2	3	35
	NH	35	44	30	43	41	41	31	32	375
Proline in position, observed / expected[1]	HSS	1.1	0.0	1.7	0.0	0.4	0.2	1.5	1.3	
	ISS	0.8	0.0	1.5	0.0	0.8	0.8	1.5	1.1	
	LSS	0.6	0.3	0.6	0.6	0.0	0.3	0.6	0.9	
	NH	1.0	1.3	0.9	1.3	1.2	1.2	0.9	0.9	

[1]The ratio observed/expected was calculated by dividing the number of cleavage sites in a specific cluster with proline in a certain position by the expected number of cleavage sites for that category based on the distribution of all cleavage sites.

Table 5
Clustering of cleavage sites with charged residues at different positions of the cleavage site.

		P4	P3	P2	P1	P1'	P2'	P3'	P4'
Occurrence K or R in position		45	45	46	46	46	45	45	45
Occurrence with FYWML in P1 position		6	18	9	0	9	9	8	12
Occurrence K, R, E or D in position		113	114	115	115	115	113	112	112
Occurrence with FYWML in P1		21	39	22	0	23	25	23	28
K of R in position, cleavage site in	HSS	2	8	5	3	10	7	6	9
	ISS	1	6	3	2	6	5	3	3
	LSS	5	3	4	6	11	5	5	1
	NH	37	28	35	35	19	28	31	32
K of R in position, expected/observed	HSS	0.4	1.5	0.9	0.6	1.8	1.3	1.1	1.7
	ISS	0.4	2.3	1.1	0.7	2.2	1.9	1.1	1.1
	LSS	1.6	0.9	1.2	1.9	3.4	1.6	1.6	0.3
	NH	1.1	0.8	1.0	1.0	0.5	0.8	0.9	0.9
K,R,D,E in position, cleavage site in	HSS	5	17	9	4	15	11	16	15
	ISS	4	13	6	2	8	8	5	7
	LSS	9	5	7	12	18	8	8	6
	NH	95	80	94	97	75	86	83	84
K, R, D, E in position, expected/observed	HSS	0.4	1.2	0.7	0.3	1.1	0.8	1.2	1.1
	ISS	0.6	1.9	0.9	0.3	1.2	1.2	0.8	1.1
	LSS	1.1	0.6	0.9	1.5	2.2	1.0	1.0	0.8
	NH	1.1	0.9	1.1	1.1	0.9	1.0	1.0	1.0

F87, F157 and F205), and some at slow rate, ($9 \pm 1 \cdot 10^{-4}$ $s^{-1} \cdot mg^{-1}_{enzyme}$; F52) up to a hydrolysis rate of 8.3 $s^{-1} \cdot mg^{-1}_{enzyme}$ (F190). Furthermore, it was observed that cleavage sites with non-aromatic amino acids could be HSS. This was for example the case for cleavage sites H32 and M90 in α-LA and D98 and A139 in β-LG (**Supplementary data B**). In order to predict the extent of hydrolysis, it should be considered how often cleavages sites with a certain AA in the *P*1 position are hydrolysed. For the cleavage sites in α-LA, β-LG and β-cas with phenylalanine, tyrosine or tryptophan in the *P*1 position, hydrolysis was observed for 28 out of 36 occurrences (Fig. 9). For methionine, chymotrypsin hydrolysed 8 out of 11 occurrences, despite the absence of an aromatic ring. Leucine was hydrolysed less frequently (49 %) and at relatively slower hydrolysis rates than the other preferred residues. This suggests that leucine is less attractive for chymotrypsin to be positioned in the S1 binding pocket, or, that the effect of surrounding amino acids seem more dominant than for phenylalanine, tyrosine, tryptophan and methionine. Cleavage sites with non-preferred amino acids were hydrolysed, but only in<30 % of the occurrences and at hydrolysis rates $< 1 \cdot 10^{3}$ s^{-1}. Possibly, the low hydrolysis rates of cleavage sites with preferred amino acids in the *P*1 position could be explained by amino acids neighbouring the *P*1 position (secondary specificity).

3.8. The effect of neighbouring amino acids: Proline

Literature suggested that proline could hinder chymotrypsin hydrolysis when present around the cleavage site. In the sequences of α-LA, β-LG and β-cas, proline was present in each binding site position 45 times. The cleavage sites with proline in the position *P*4, *P*2, *P*3′ and *P*4′ were similarly distributed in the HSS/ISS/LSS/NH clusters as one would expect based on the relative number of cleavage sites in each cluster (Table 4). Therefore, a proline in these positions does not seem to influence the hydrolysis rate. On the contrary, a proline in the *P*3 position did never yield a cleavage site in HSS or ISS, despite that for 7 of these cleavage sites, a preferred amino acid was in the *P*1 position. For cleavage sites with a proline in the *P*1′ and *P*2′ positions, 1 and 2 cleavage sites were respectively clustered as HSS. This was also 2.5 times and 5 times less than expected. Cleavage sites that had a phenylalanine, tyrosine, tryptophan or methionine in the *P*1 position and that were classified as LSS or NH (11 occurrences), 8 had a proline in the *P*3, *P*1′ or *P*2′ position. For leucine, 10 out of 29 cleavage sites in LSS or NH had a proline in the *P*3, *P*1′ or *P*2′ position. The proline rule explained 73 % of the missed cleavages for phenylalanine, tyrosine, tryptophan and methionine, but only 34 % of the missed cleavages for leucine.

3.9. The effect of neighbouring amino acids: Charged amino acids

A study from Hedstrom suggested that a lysine or arginine residue in the *P*1′ position would enhance hydrolysis. A lysine or arginine in the *P*1′ occurred 46 times, of which 9 in combination with a preferred AA residue in the *P*1 position. Of the 46 cleavage sites with a positive residue in the *P*1′ position, 10 were HSS, 6 were ISS, 11 were LSS and 19 were not hydrolysed (Table 5). The number of cleavage sites in the respective clusters were 1.8x (HSS), 2.2x (ISS), 3.4x (LSS) more than expected. The trend as described by Hedstrom was confirmed, but was not as dominant as the observations with proline. A study from Keil suggested that a (positively or negatively) charged residue in the *P*2 position would hinder hydrolysis. Out of the 115 occurrences, of which 22 with a preferred residue in the *P*1 position, 9 cleavage sites were classified as HSS and 6 as ISS. The occurrences were respectively 0.7 and 0.9 times the expected number of cleavage site in HSS and ISS and thereby within the expectation. The effect of charged amino acid residues on the *P*1′ position and *P*2 position were not the main reason behind low hydrolysis rates.

3.10. The effect of neighbouring amino acids: Size and hydrophobicity

For the *P*3 position, the large residues alanine and arginine would fit best according to the QSAR model of Schellenberger. Cleavage sites with these residues in the *P*3 position were 1.7 and 1.1 times the expectation in HSS and ISS, respectively. These values were not significantly different from the values in other positions, considering the number of observations. Previous studies described for the *P*2 position a good fit when either the residues leucine and valine, or, (other) hydrophobic residues were in the *P*2 position. From our dataset, both did not show considerable effects. Although these observations were done with synthetic substrates, their effect in proteolysis seems minor.

4. Conclusion

In this study, the aim was to describe the path of hydrolysis with chymotrypsin. Digestion kinetics for individual cleavage sites allowed us to relate subsite composition with chymotrypsin's activity. Although chymotrypsin is generally considered to have a specificity for aromatic and hydrophobic residues, this study showed that it is able to hydrolyse peptide bonds after almost all amino acid residues. A preference was observed for phenylalanine, tyrosine, tryptophan and methionine (hydrolysed in 78 % of the cleavage sites) and to a lower extent leucine (hydrolysed in 49 % of the cleavage sites). For cleavage sites with these preferred amino acids in the *P*1 position, still a substantial number of cleavage sites were not hydrolysed. The negative effect of proline around the cleavage site was found to be position dependent. A proline residue in the *P*3, *P*1′ or *P*2′ position hindered chymotrypsin hydrolysis, whereas a proline in other positions had no effect. Hindrance by proline explained 45 % of the missed-cleavages for cleavage sites within the preference. Charge or hydrophobic amino acids as neighbouring amino acids in the cleavage sites did not show major effects. The approach taken gave fundamental insight in chymotrypsin and is promising to apply to other digestive or novel (commercial) proteases.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Gijs J.C. Vreeke: Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Writing – original draft. **Jean-Paul Vincken:** Conceptualization, Supervision, Writing – review & editing. **Peter A. Wierenga:** Conceptualization, Software, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2023.112485.

References

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

Blow, D. M. (1976). Structure and mechanism of chymotrypsin. *Accounts of Chemical Research, 9*(4), 145–152. https://doi.org/10.1021/ar50100a004

Burrell, M. M. (1993). Methods in Molecular Biology, Vol 16 Enzymes of Molecular Biology. In: Humana.

Butré, C. I., Sforza, S., Gruppen, H., & Wierenga, P. A. (2014). Introducing enzyme selectivity: A quantitative parameter to describe enzymatic protein hydrolysis. *Analytical and Bioanalytical Chemistry, 406*(24), 5827–5841. https://doi.org/10.1007/s00216-014-8006-2

Butré, C. I., Wierenga, P. A., & Gruppen, H. (2014). Influence of water availability on the enzymatic hydrolysis of proteins. *Process Biochemistry, 49*, 1903–1912. https://doi.org/10.1016/j.procbio.2014.08.009

Catiau, L., Delval-Dubois, V., Guillochon, D., & Nedjar-Arroume, N. (2011). Characterization and Identification of a Chymotryptic Hydrolysate of Alpha-Lactalbumin Stimulating Cholecystokinin Release in STC-1 Cells. *Applied Biochemistry and Biotechnology, 165*(5), 1264–1273. https://doi.org/10.1007/s12010-011-9344-6

Deng, Y., Butré, C. I., & Wierenga, P. A. (2018). Influence of substrate concentration on the extent of protein enzymatic hydrolysis. *International Dairy Journal, 86*, 39–48. https://doi.org/10.1016/j.idairyj.2018.06.018

Deng, Y., van der Veer, F., Sforza, S., Gruppen, H., & Wierenga, P. A. (2018). Towards predicting protein hydrolysis by bovine trypsin. *Process Biochemistry, 65*, 81–92. https://doi.org/10.1016/j.procbio.2017.11.006

Feinstein, G., Hofstein, R., Koifmann, J., & Sokolovsky, M. (1974). Human pancreatic proteolytic enzymes and protein inhibitors: Isolation and molecular properties. *European Journal of Biochemistry, 43*(3), 569–581. https://doi.org/10.1111/j.1432-1033.1974.tb03444.x

Galvão, C. M. A., Souza Silva, A. F., Custódio, M. F., Monti, R., & Giordano, R.d. L. C. (2001). Controlled hydrolysis of cheese whey proteins using trypsin and α-chymotrypsin. *Applied Biochemistry and Biotechnology, 91*(1), 761. https://doi.org/10.1385/ABAB:91-93:1-9:761

Giansanti, P., Tsiatsiani, L., Low, T. Y., & Heck, A. J. (2016). Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols, 11*(5), 993–1006. https://doi.org/10.1038/nprot.2016.057

Gorrill, A., & Thomas, J. (1967). Trypsin, chymotrypsin, and total proteolytic activity of pancreas, pancreatic juice, and intestinal contents from the bovine. *Analytical Biochemistry, 19*(2), 211–225. https://doi.org/10.1016/0003-2697(67)90156-X

Gorrill, A. D. L., & Friend, D. W. (1970). Pancreas size and trypsin and chymotrypsin activities in the pancreas and intestinal contents of pigs from birth to 5 weeks of age. *Canadian Journal of Physiology and Pharmacology, 48*(11), 745–750. https://doi.org/10.1139/y70-107

Guyonnet, V., Tłuscik, F., Long, P. L., Polanowski, A., & Travis, J. (1999). Purification and partial characterization of the pancreatic proteolytic enzymes trypsin, chymotrypsin and elastase from the chicken. *Journal of Chromatography A, 852*(1), 217–225. https://doi.org/10.1016/S0021-9673(99)00355-6

Hedstrom, L. (2002). Serine Protease Mechanism and Specificity. *Chemical Reviews, 102*(12), 4501–4524. https://doi.org/10.1021/cr000033x

Hudáky, P., Kaslik, G., Venekei, I., & Gráf, L. (1999). The differential specificity of chymotrypsin A and B is determined by amino acid 226. *European Journal of Biochemistry, 259*(1–2), 528–533. https://doi.org/10.1046/j.1432-1327.1999.00075.x

Kamath, V., Niketh, S., Chandrashekar, A., & Rajini, P. S. (2007). Chymotryptic hydrolysates of α-kafirin, the storage protein of sorghum (Sorghum bicolor) exhibited angiotensin converting enzyme inhibitory activity. *Food Chemistry, 100*(1), 306–311. https://doi.org/10.1016/j.foodchem.2005.10.004

Keil, B. (1992). *Specificity of proteolysis: Springer-Verlag*. New York, USA: Berlin Heidelberg.

Kuipers, B. J. H., & Gruppen, H. (2007). 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. *Journal of Agricultural and Food Chemistry, 55*(14), 5445–5451. https://doi.org/10.1021/jf0703371

Lin, I. C., Sookkheo, B., Phutrakul, S., Chen, S. T., Tseng, M. J., & Wang, K. T. (1999). Combinatorial Peptide Library for Probing the Selectivity of the s-1 Subsite of Proteases. *Journal of the Chinese Chemical Society, 46*(2), 147–152. https://doi.org/10.1002/jccs.199900022

Loveday, S. M., Peram, M. R., Singh, H., Ye, A., & Jameson, G. B. (2014). Digestive diversity and kinetic intrigue among heated and unheated β-lactoglobulin species. *Food & Function, 5*(11), 2783–2791. https://doi.org/10.1039/C4FO00362D

Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research, 42* (D1), D503–D509. https://doi.org/10.1093/nar/gkt953

Ruan, C.-Q., Chi, Y.-J., & Zhang, R.-D. (2010). Kinetics of hydrolysis of egg white protein by pepsin. *Czech journal of food sciences, 28*(5), 355–363. https://doi.org/10.17221/228/2009-cjfs

Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications, 27*(2), 157–162. https://doi.org/10.1016/S0006-291X(67)80055-X

Schellenberger, V., Braune, K., Hofmann, H. J., & Jakubke, H. D. (1991). The specificity of chymotrypsin: A statistical analysis of hydrolysis data. *European Journal of Biochemistry, 199*(3), 623–636. https://doi.org/10.1111/j.1432-1033.1991.tb16163.x

Schellenberger, V., Schellenberger, U., Mitin, Y. V., & Jakubke, H. D. (1990). Characterization of the S′-subsite specificity of bovine pancreatic α-chymotrypsin via acyl transfer to added nucleophiles. *European Journal of Biochemistry, 187*(1), 163–167. https://doi.org/10.1111/j.1432-1033.1990.tb15290.x

Segal, D. M. (1972). Kinetic investigation of the crystallographically deduced binding subsites of bovine chymotrypsin Aγ. *Biochemistry, 11*(3), 349–356. https://doi.org/10.1021/bi00753a007

Srinivas, S., & Prakash, V. (2010). Bioactive Peptides from Bovine Milk [alpha]-Casein: Isolation, Characterization and Multifunctional properties. *International Journal of Peptide Research and Therapeutics, 16*(1), 7–15. https://doi.org/10.1007/s10989-009-9196-x

Steitz, T. A., Hendekson, R., & Blow, D. M. (1969). Structure of crystalline α-chymotrypsin: III. Crystallographic studies of substrates and inhibitors bound to the active site of α-chymotrypsin. *Journal of Molecular Biology, 46*(2), 337–348. https://doi.org/10.1016/0022-2836(69)90426-4

Vorob'ev, M. M. (2013). Quantification of two-step proteolysis model with consecutive demasking and hydrolysis of peptide bonds using casein hydrolysis by chymotrypsin. *Biochemical Engineering Journal, 74*, 60–68. https://doi.org/10.1016/j.bej.2013.02.020

Vreeke, G. J. C., Lubbers, W., Vincken, J.-P., & Wierenga, P. A. (2022). A method to identify and quantify the complete peptide composition in protein hydrolysates. *Analytica Chimica Acta, 1201*, Article 339616. https://doi.org/10.1016/j.aca.2022.339616

Wright, H. T. (1977). Secondary and conformational specificities of trypsin and chymotrypsin. *European Journal of Biochemistry, 73*(2), 567–578. https://doi.org/10.1111/j.1432-1033.1977.tb11352.x