# Aggregation of peptides in soy protein isolate hydrolysates

***The individual contributions of glycinin- and β-conglycinin-derived peptides***

Bas J.H. Kuipers

| | |
|---|---|
| Promotor: | **Prof. dr. ir. H. Gruppen**<br>Persoonlijk hoogleraar bij de leerstoelgroep Levensmiddelenchemie |
| Promotiecommissie: | **Prof. dr. R.J. Bino**<br>Wageningen Universiteit/PRI<br>**Prof. dr. ir. J. Delcour**<br>Katholieke Universiteit Leuven, België<br>**Prof. dr. C.G. de Kruif**<br>Universiteit Utrecht, NIZO food research<br>**Dr. A.L. de Roos**<br>DSM Food Specialties |

*Dit onderzoek is uitgevoerd binnen de onderzoekschool VLAG (Voeding, Levensmiddelentechnologie, Agrobiotechnologie en Gezondheid).*

# Aggregation of peptides in soy protein isolate hydrolysates

## *The individual contributions of glycinin- and β-conglycinin-derived peptides*

Bas J.H. Kuipers

**Proefschrift**
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M.J. Kropff
in het openbaar te verdedigen
op maandag 21 mei 2007
des namiddags te half twee in de Aula

Kuipers, Bas J.H.

Aggregation of peptides in soy protein isolate hydrolysates;
*The individual contributions of glycinin- and β-conglycinin-derived peptides*

Ph.D. Thesis Wageningen University, The Netherlands, 2007

ISBN: 90-8504-609-2

**ABSTRACT**

B.J.H Kuipers Aggregation of peptides in soy protein isolate hydrolysates

*The individual contributions of glycinin- and β-conglycinin-derived peptides*

Ph.D. Thesis Wageningen University, Wageningen, The Netherlands, 2007

Keywords: Soy proteins, glycinin, β-conglycinin, enzymatic hydrolysis, peptides, aggregation, gelation, identification, mass-spectrometry, mapping

The aim of the work presented in this thesis is to understand how limited enzymatic hydrolysis can influence the aggregation behavior of soy protein material. This is performed by uncoupling the enzymatic degradation and the aggregation process. Subsequently, the regions in soy proteins from which the aggregating peptides originate were identified.
Hydrolysates of soy protein isolates (SPI) were prepared using the protease subtilisin Carlsberg. The enzyme was inhibited when the desired degree of hydrolysis (DH) was reached. It was shown that with increasing DH, the pH at which a gel could be formed was increasing as well. The aggregation behavior of SPI, glycinin- and β-conglycinin-derived hydrolysates showed that glycinin-derived peptides are responsible for the aggregation of SPI-derived peptides. This was also found when glycinin was hydrolyzed with other enzymes (bromelain, papain, and chymotrypsin) using SPI, β-conglycinin, and bovine whey proteins as reference materials. Only hydrolysis of glycinin by trypsin did not result in strong aggregating peptides.
Subsequently, a new method was developed, denoted accumulative-quantitative-peptide-mapping. This method reveals those regions in the parental protein from which the aggregating peptides originate. This method comprises a second hydrolysis of the aggregating peptides, followed by separation, identification, and quantification of individual peptides obtained. Quantification was performed based on absorbance at 214 nm, corrected for the calculated molar extinction coefficient based on the amino acid composition. This novel method revealed that mainly the basic polypeptide and that part of the acidic polypeptide close to the location of the disulfide bridge connecting the basic to the acidic polypeptide, are the predominant regions of glycinin yielding aggregating peptides. These regions have a relative high hydrophobicity compared to that of the total protein. Upon hydrolysis the net hydrophobicity of the remaining glycinin is increasing, eventually resulting in aggregation and subsequent gelation.

**TABLE OF CONTENTS**

# CHAPTER 1

# General introduction

## INTRODUCTION

Soybeans and soy products, such as soymilk, tofu, soy sauce and tempeh, are common in Asia since the ancient times. In Europe and the USA, the use of soybeans is common since a few decennia, due to the use of soybean oil for frying oils and margarines. The side stream of the oil production, the defatted soybean meal, was primarily used as protein source for animal feeds. The low price of the side products and good nutritional quality of the constituents triggered the food industry to focus on the application of soy proteins in food formulations. Due to advances in soy processing technology this resulted in new food products (e.g. soy based meat and dairy analogs) and new ingredients (soy isolates and concentrates) that can perform several functional properties in food systems *(1, 2)*.

One of the processing treatments of soy proteins to improve its use as a functional ingredient in food systems is enzymatic hydrolysis. This may improve the solubility, emulsifying and foaming properties of soy protein preparations *(3, 4)*. Another functional property of soy proteins are their gelling properties, which contributes to the texture of the food system *(5, 6)*. The present study was initiated with the aim to understand how limited hydrolysis can strengthen the aggregation behavior of soy proteins. More particular, it is focused on the contribution of the individual storage proteins and the regions of these proteins from which peptides originate that have strong aggregating properties.

## SOY PROTEINS

Proteins are the main constituents of soybeans, constituting on average 40 % of the dry matter. Based on their biological functions in plants, seed proteins can be classified into two groups: metabolic proteins and storage proteins. The metabolic proteins are active in the normal cellular activities, including the synthesis of the storage proteins. The storage proteins provide a source of nitrogen and carbon for the development of the seedling during germination *(2)*.

Based on their sedimentation coefficients, when dissolved in a pH 7.6, 0.4 M sodium-phosphate buffer containing 0.01 M mercaptoethanol, soybean proteins can be classified into 2S (13-18 %), 7S (30-46 %), 11S (36-53 %), and 15 S (0-4 %) proteins *(7)*. The 11S and 15S fractions consist of glycinin and glycinin polymers, respectively *(8)*. The 7S fraction is more heterogeneous, but the majority is β-conglycinin, which is next to glycinin the major storage protein present in soybeans *(9)*. The 2S fraction consists of Bowman-Birk and Kunitz trypsin inhibitors, cytochrome C, and α-conglycinin *(8, 10)*. The above nomenclature based on sedimentation constants has been used a lot in literature. However, these constants largely depend on the conditions of solubilization (buffer, ionic strength, and pH). Nowadays, based on the known DNA sequences of the soybean proteins, soybean proteins can be classified based on their genes encoding for glycinin, β-conglycinin, and the other proteins present in soybeans. As stated above, glycinin and β-conglycinin are the two main storage proteins of

soybeans, in which glycinin is in general present in a higher proportion. The ratio glycinin/β-conglycinin was observed to vary from 1.31 to 2.45 based on densitometric analysis of SDS-PAGE gels of various soybean cultivars *(11-13)*.

**Soy glycinin**

In total, 6 different glycinin genes have been identified from soybeans. These are Gy1 to Gy5 and Gy7, representing the glycinin subunits G1 to G5 and G7, respectively. The G7 subunit has been recently identified *(14)* and is related to the other five soybean subunits. It is poorly expressed and, therefore, not further taken into account in this study. With the exception of G4 each glycinin subunit consists of one acidic (A) and one basic (B) polypeptide, which are connected with a disulfide bridge (Figure 1). For the 5 major glycinin subunits 6 different acidic (A1a, A2, A1b, A5, A4 and A3) and 5 basic (B2, B1a, B1b, B3 and B4) polypeptides have been identified *(15)*. The A4 polypeptide of G4 is the only acidic polypeptide not covalently linked to a basic polypeptide. The molecular masses of the basic polypeptide are around 20 kDa, while those of the acidic polypeptides vary between 10.5 and 36.4 kDa. In Figure 1 the details of the different subunits are given.

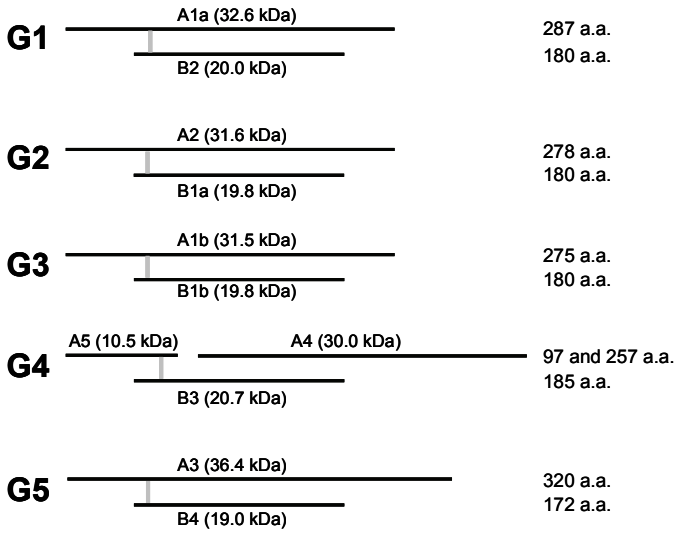


**Figure 1**: Schematic overview of the five major glycinin subunits and their corresponding polypeptides with molecular masses and number of amino acids (a.a.) per chain. The grey lines represent the disulfide bridge connecting the acidic and basic polypeptide.

The 5 subunits can be divided into two subfamilies that are designated as Group-I (G1, G2, and G3) and Group-II (G4 and G5) *(16, 17)*. Homologies between members of the same group range from 80 to 90 %, and between the groups the homology is less than 50 % *(15)*. Figure 2 shows the hydrophobicity, according to Eisenberg *(18)*, of Group-I and Group-II glycinins as a function of their amino acid sequence. The amino acid sequences as present in Swiss-Prot (www.expasy.org) with the following primary accession numbers: P04776 (G1),

P04405 (G2), P11828 (G3), P02858 (G4), and P04347 (G5), were used for the plots. The calculations were performed using the ProtScale software from the Swiss Institute of Bioinformatics (www.expasy.org), using a "window size" of 15 (= peptide range over which the average hydrophobic score is calculated). In Figure 2, Glycinin 2 and 3 are aligned to Glycinin 1, and Glycinin 5 is aligned to Glycinin 4. In the aligned graphs the cysteines participating in the disulfide bridge are at the same position. Regions in the amino acid sequence in which the hydrophobicity score is above 0 are regarded as being hydrophobic regions of the protein. The dotted line represents the location of the disulfide bridge that connects the acidic and the basic polypeptide. It can be seen that the hydrophobicity scores of the basic polypeptides are higher than that of the acidic polypeptides, indicating that the basic polypeptides are more hydrophobic than the acidic polypeptides. The high similarity in hydrophobicity distribution along the peptide chains between subunits within one group (80-90%) reflects the high sequence homology within each group. Between the two groups there is also some homology, but it is lower than 50 % *(15)*.

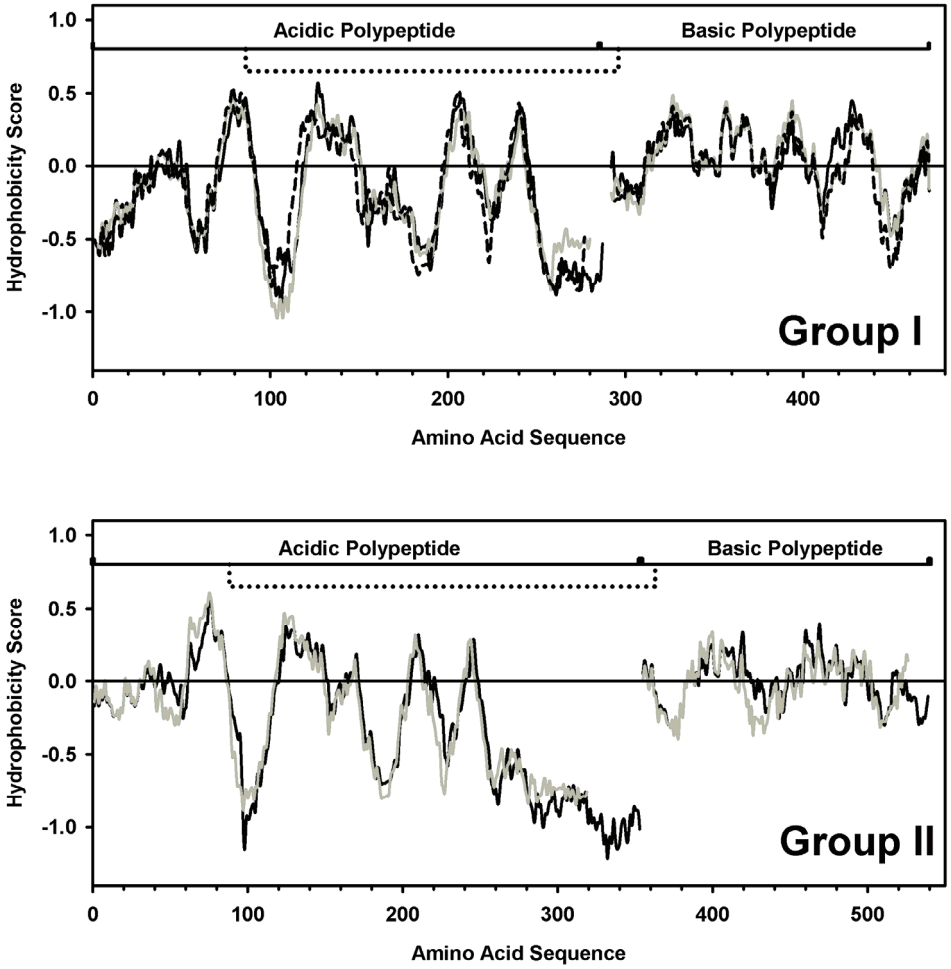


**Figure 2**: Hydrophobicity score of glycinin according to Eisenberg *(18)* for Group I (G1 (——) G2 (——) and G3 (— —), and Group II (G4 (——) and G5 (——)). The dotted lines represent the location of the disulfide bridges between the acidic and basic polypeptides.

At ambient temperatures and pH 7.6 ($I$ = 0.5 M), glycinin is present as hexamers (11S) with molecular masses of about 300-360 kDa, while at pH 3.8 glycinin is present as trimers (7S). Lowering the ionic strength at neutral pH from 0.5 to below 0.1 M, also induces dissociation of 11S glycinin into 7S glycinin *(19)*. Figure 3 shows the crystal structure of a glycinin trimer *(20)*. The core of the subunit consists of two jelly-roll β-barrels and two extended helix domains. These helix domains stack on the helix domains of the adjacent glycinin subunit.

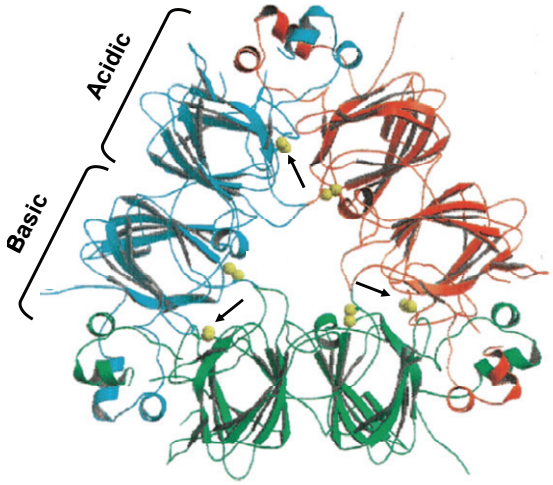


**Figure 3**: Crystal structure of three subunits in the Glycinin G1 trimer. The yellow dots represent disulfide bridges. The arrows indicate the disulfide bridge that connects the acidic and basic polypeptide. Reprinted from Adachi and co-workers 2001 *(20)*, with permission from Elsevier.

### Soy β-conglycinin

β-Conglycinin is a glycoprotein consisting of three different subunits, α, α', and β *(21)* as presented in Figure 4. The α and α' subunit have masses of 63.5 and 67.2 kDa, respectively, whereas the β-subunit has a lower mass of 47.8 kDa. The α and α' subunits each contain two, and the β subunit contains one carbohydrate chain attached to asparagine. The carbohydrate chain can vary in length. The difference is in the number of mannose units. In total 9 to 11 monosaccharides constitute the carbohydrate chain ($(GlcNAc)_2$-$(Man)_{7\text{-}9}$) *(22, 23)*. As a trimer (7S), β-conglycinin can exist in seven different combinations (βββ, ββα', ββα, βαα', βαα, ααα', and ααα) *(24, 25)*.

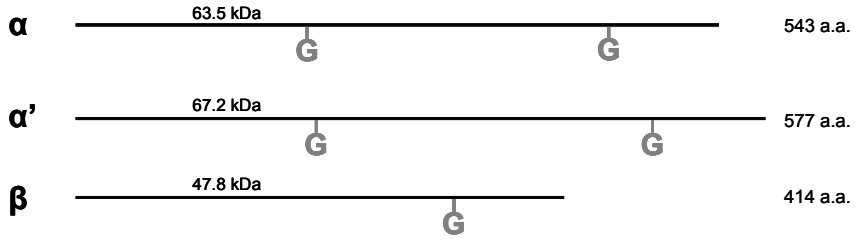


**Figure 4**: Schematic overview of β-conglycinin subunits and the number of amino acids (a.a.) per subunit. The molecular masses are only based on the polypeptide chain. The glycosylation sites are indicated with a G.

Figure 5 shows the hydrophobicity plots of the α-, α'-, and β-subunits of β-conglycinin, similar to the plots of glycinin in Figure 2. The amino acid sequences, as present in Swiss-Prot (www.expasy.org), with the following primary accession numbers were used: P13916 (α-subunit), P11827 (α'-subunit), and P25974 (β-subunit). The α'- and β-subunit are aligned to the α-subunit. In the aligned graph the glycosylation sites that are the closest to the C-terminus are at the same position as indicated with the arrow. It can be seen that there is a high homology between the different β-conglycinin subunits, with the main difference that the β-subunit is shorter, and does not contain the highly hydrophilic N-terminal extension.

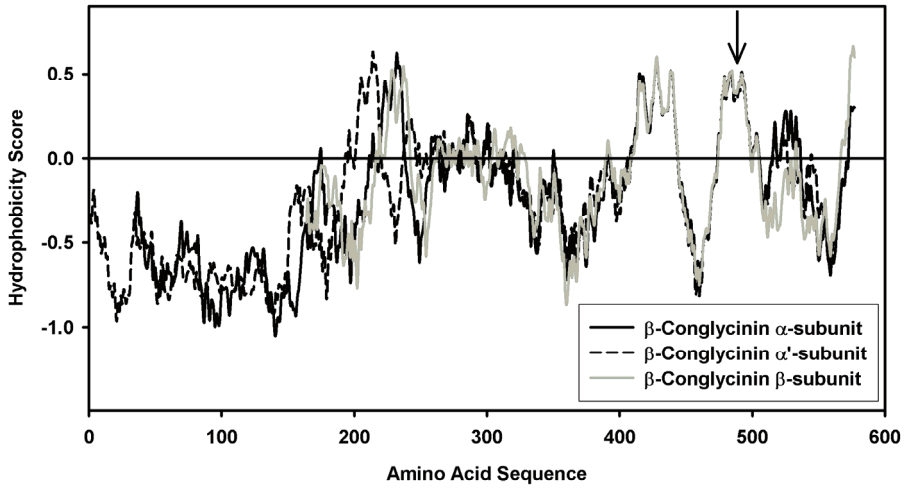


**Figure 5**: Hydrophobicity score of β-conglycinin according to Eisenberg *(18)* for the α-subunit (——), α'-subunit (— —), and the β-subunit (——). The arrow points out the position of the glycosylation site to which the chains are aligned.

## ENZYMATIC HYDROLYSIS OF SOY PROTEINS

To use soy proteins in food applications, depending on the application, the proteins must possess appropriate functional properties such as solubility, emulsifying, foaming or gelling properties *(26, 27)*. Hydrolysis of proteins is known to be a means to improve these functional properties *(3, 4, 28)*.

The extent to which the functional properties of a protein preparation may be altered by hydrolysis is dependent on the degree to which the protein has been hydrolyzed (degree of hydrolysis; DH). Even more important are the positions at which the protease can cleave (specificity) and will cleave (selectivity). For a protein or a protein preparation the DH is defined as the number of peptide bonds cleaved ($h$) divided by total number of peptide bonds present ($h_{tot}$):

$$\mathrm{DH} = \frac{h}{h_{tot}} \times 100\% \qquad [1]$$

The three most common methods to determine the DH, are the TNBS *(29)*, OPA *(30, 31)*, and pH-stat method *(28)*. The OPA and TNBS methods are both based on a reaction with the primary amino groups, of which the number is increasing with increasing DH. At slightly alkaline pH the pH-stat method is based on the net release of protons during the hydrolysis into the surrounding medium, leading to a decrease in pH. The number of peptide bonds cleaved can be determined from the amount of base required to maintain a constant pH during the enzymatic reaction using Formula 2:

$$DH = B \times N_b \times \frac{1}{\alpha} \times \frac{1}{MP} \times \frac{1}{h_{tot}} \times 100\ \% \qquad [2]$$

Where B is the base consumption (mL), $N_b$ is the normality of the base (M), $\alpha$ the average degree of dissociation of the $\alpha$-$NH_2$ groups, MP the mass of the protein being hydrolyzed (g) and $h_{tot}$ the total number of peptide bonds in the protein substrate (meqv $g^{-1}$ protein). The advantage of the pH-stat method over the OPA and TNBS methods is that with the pH-stat method the DH can be followed in time during the hydrolysis. To stop the reaction the enzyme has to be inactivated. This can be performed by changing the pH, heating the hydrolysate to denature the enzyme, or by the addition of an enzyme inhibitor *(32)*.

The position in a protein at which a protease can cleave depends on the specificity of the enzyme. Whether a cleavable peptide bond will actually be cleaved depends on the susceptibility of the peptide bond and the selectivity of the enzyme. With susceptibility, the possibility at which the peptide bond can be reached by the enzyme physically in meant. With selectivity of the enzyme the preference is meant for peptide bonds that are equally susceptible.

Specificity is a characteristic of the enzyme since the reaction takes place in a particular region of the enzyme, the active site. The active site is able to recognize and bind to the substrate due to structural complementarities between amino acid residues in the substrate and amino acid residues in the vicinity of the active site. Proteases with a narrow specificity have high interaction requirements and, therefore, only recognize a narrow range of peptide bonds. Proteases with a broad specificity have flexible interaction requirements. As a consequence those proteases recognize a wider range of peptide bonds *(33)*. In the protein substrate, structural complementarities with the enzyme are ensured by the amino acid residues that are at the carboxyl side of the cleaved peptide bond, denoted $P_n$, and those at the amino side of the cleaved peptide bond $P'_n$ *(32)*.

In literature quite some information can be found about the influence of enzymatic hydrolysis on for example the solubility or aggregation of soy proteins *(34-39)*. Due to different hydrolysis parameters in these studies (e.g. type of enzyme, pH, and substrate pretreatment) differences in specificity and selectivity can be expected.

**Table 1**: Literature overview of research performed in the field of enzyme induced aggregation of soy proteins

| Substrate | Reference | Enzyme | Focus | Main conclusions |
|---|---|---|---|---|
| Soymilk | Lee, 1997 *(40)* | *Bacillus licheniformis* protease | Texture of soy milk curds obtained by different coagulation methods. | Enzyme-induced curd result in a more smooth texture in comparison with curd obtained by acidification or addition of $CaSO_4$. |
| Soymilk | Murata, 1987 *(41, 42)* | Various commercial proteases | Which enzymes can induce aggregation of soymilk, and what are optimum aggregation conditions. | High proteolytic activity coincides with high aggregation activity. Microbial alkaline and neutral proteases have better aggregation properties than microbial acidic and animal proteases and plant proteases other than bromelain. |
| Soymilk | Murata, 1988 *(43)* | Subtilisin Carlsberg | Comparison of functional properties of enzyme, calcium salt and acid-induced soymilk curds. | Enzyme-induced curds show good emulsifying and foaming stability, and contribute to a smooth texture. |
| Soymilk | Park, 1985/87 *(44, 45)* Kobayashi, 1988 *(46)* | *Bacillus sp.* K-295G-7 protease | Characterization of the enzyme in comparison to other proteases with known soymilk aggregation capacity. | Two serine proteases were identified, Enzyme I having similar specificity as subtilisin Carlsberg, Enzyme II had lower similarity compared to other tested proteases. |
| Soymilk | Yasuda, 1999 *(47)* Aoyama, 2000 *(48)* | *Bacillus pumilus*-derived protease | Characterization of the enzyme, (optimum pH and temperature, and specificity). Degradation mechanism of soy proteins. | Serine protease with a high similarity to subtilisin Carlsberg based on amino acid composition and specificity. Basic polypeptide of glycinin was resistant towards degradation. |
| SPI, glycinin, β-conglycinin, | Mohri, 1984 *(39)* | Bromelain | Study the enzyme-induced aggregation mechanism, with respect to the contribution of glycinin and β-conglycinin. | Glycinin is responsible for the enzyme-induced aggregation of SPI, mainly due to hydrophobic interactions and disulfide bridges, basic polypeptide was degraded less rapid than the acidic polypeptide. |
| Water extractable soy proteins. | Fuke, 1985 *(49)* | Bromelain | Study the enzyme-induced aggregation and gelation of SPI. | Degradation of soy proteins leads to the formation of a gel with fine network structure composed of low molecular weight fragments. |
| SPI | Inouye, 2002 *(50)* Nagai, 2004 *(51)* | Subtilisin Carlsberg | Investigate the mechanism/kinetics of enzyme-induced SPI aggregation. | Glycinin and β-conglycinin are completely degraded. The degradation of the hydrophobic regions in the inner areas of the protein results in aggregation due to hydrophobic interactions. |
| Glycinin | Zhong, 2006 *(37)* | Papain | Effect of incubation temperature and enzyme concentration on the gelation of glycinin peptides. Study the molecular interactions involved in the gelation. | Gel strength was independent of enzyme concentration and decreased with increasing temperature. Hydrophobic interactions are the main forces involved in the gelation mechanism. |
| SPI | Zhong, 2007 *(38)* | Various commercial proteases | Study the enzyme-induced aggregation of SPI, using various commercial proteases. | Each protease resulted in enzyme-induced gelation, with papain resulting in the strongest gels, followed by Alcalase and bromelain. |

## AGGREGATION OF SOY PROTEIN HYDROLYSATES

In general, hydrolysis is assumed to be unfavorable for the gelling properties of proteins. This is because it increases the number of charged groups and reduces the molecular weight, which hamper gelation *(52, 53)*. On the other hand, the structure of the protein is altered during hydrolysis and buried hydrophobic groups become exposed, which might result in aggregate formation. Two well known enzyme-induced aggregation systems, with enzymes having narrow specificity, are the aggregation of casein upon hydrolysis with chymosin in the cheese processing *(54)*, and the aggregation of fibrin that is formed upon hydrolysis of fibrinogen by trombin in the process of blood platelet aggregation *(55)*. Also in bovine whey and soy systems, hydrolysis has been observed to enhance the gelling properties of the protein, which receives quite some attention in literature and is referred to as enzyme-induced aggregation *(33, 37, 48, 49, 56-58)*.

### Enzyme-induced aggregation of soy proteins

Table 1 gives an overview of research performed in the field of enzyme-induced aggregation of soy proteins. It can be observed that a lot of research is performed with soymilk as a substrate for enzyme-induced aggregation as a means to produce curd for tofu. In general, these studies aimed at improving the texture of tofu or finding new enzyme preparations that can induce aggregation. Enzyme-induced aggregation results in a more smooth texture when compared to other aggregation methods, such as acid or calcium-induced aggregation *(40, 43)*. It was shown that many different commercial available proteases have the property to induce aggregation of soy proteins *(38, 41, 42)*.

There are three enzymes, originating from micro organisms that have been studied more intensively with respect to their ability to induce aggregation of soy proteins. These are *Bacillus sp*. K-295 G-7 protease *(44-46)*, *Bacillus pumilus* TYO-67 protease *(47, 48)* and subtilisin Carlsberg from *Bacillus licheniformis (50, 51)*. These *Bacillus sp*. enzymes are all serine proteases, having a quite similar, but not identical, specificity towards the Insulin B Chain *(48)* as can be observed in Figure 6. Enzyme II from *Bacillus sp*. K-295 G-7 shows the strongest deviation from the other enzymes.

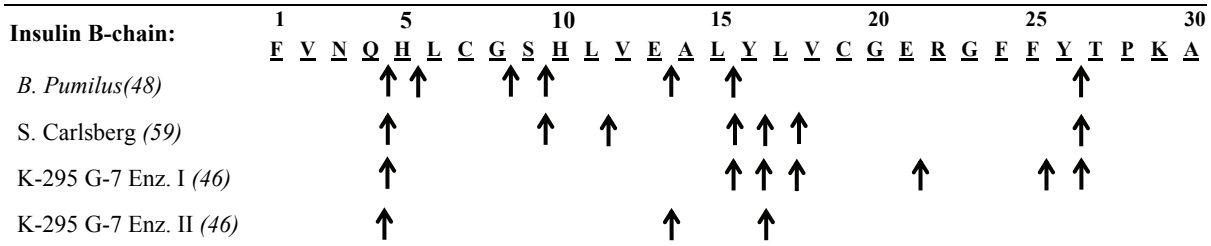


**Figure 6**: Comparison of the cleavage sites (indicated by the arrows) of oxidized insulin B-chain treated with different *Bacillus sp*.-derived proteases having the property to induce aggregation of soy proteins (adapted from Aoyama *(48)*).

The observed differences in specificity in Figure 6 have also been observed in the degradation of the different polypeptides of glycinin. The *Bacillus pumilus* TY-67 protease was not able to degrade the basic polypeptide of glycinin, whereas this polypeptide was degraded upon hydrolysis with subtilisin Carlsberg *(48)*.
Nagai and Inouye *(51)* proposed a mechanism for the enzyme-induced aggregation of soy proteins. During the first stage of hydrolysis, the hydrophilic surface areas of the proteins are degraded, not resulting in aggregation. In the second stage of the hydrolysis the hydrophobic core of the protein is cleaved. This results in a strong decomposition of the tertiary structure of the proteins followed by aggregation due to hydrophobic interactions.
Next to proteases from micro-organisms, also proteases from plants have the ability to induce aggregation of soy proteins. Bromelain, from pineapple, can induce aggregation of soy proteins *(39, 49)*. Papain, from papaya, was also shown to be able to induce aggregation of glycinin *(37)*. In the bromelain-induced aggregation of soy protein isolate (SPI), glycinin appeared to be responsible for the enzyme-induced aggregation in SPI. In the same study it was also shown that the acidic polypeptide from glycinin was degraded faster compared to the basic polypeptide*(39)*. The driving force in the aggregation of soy peptides after hydrolysis with bromelain or papain appeared to be mainly hydrophobic interactions *(37, 39)*.

**Acid-induced aggregation of hydrolysates**

In all the studies presented in Table 1, the aggregation coincides with the enzymatic hydrolysis. If no pH-stat setup is used this hydrolysis results in a decrease in pH *(38)*. So, in these studies two aggregation promoting effects play a role at the same time. On the one hand the formation of peptides that have a stronger tendency to aggregate compared to the parental protein, and on the other hand a decrease in pH towards the p*I* of the protein, which may also promote aggregate formation.
This standard approach, as presented in Figure 7, makes it difficult to study the aggregation of a hydrolysate with a fixed DH as a function of pH. Up till now, a relation between enzyme-induced aggregation and DH is only found indirectly. This is done by varying the enzyme/protein ratio or by varying the incubation time. To study the aggregation of a protein hydrolysate as a function of pH with a controlled DH, an alternative approach has to be used in which the aggregation is uncoupled from a change in pH. Such an approach is presented in Figure 7. This approach will be used in this thesis. When using a pH-stat set-up at pH 8.0 degradation of the protein occurs upon the addition of a protease. At this pH no (extensive) aggregation takes place. At each desired DH the hydrolysis can be stopped by using a chemical inhibitor. When the pH of this hydrolysate is subsequently decreased, aggregation occurs. Consequently, the change in aggregation behavior as a function of pH can be studied and compared to the aggregation behavior of the parental protein. In this alternative approach not the enzyme-induced aggregation is studied, as in the standard approach, but the acid-induced aggregation of peptides.

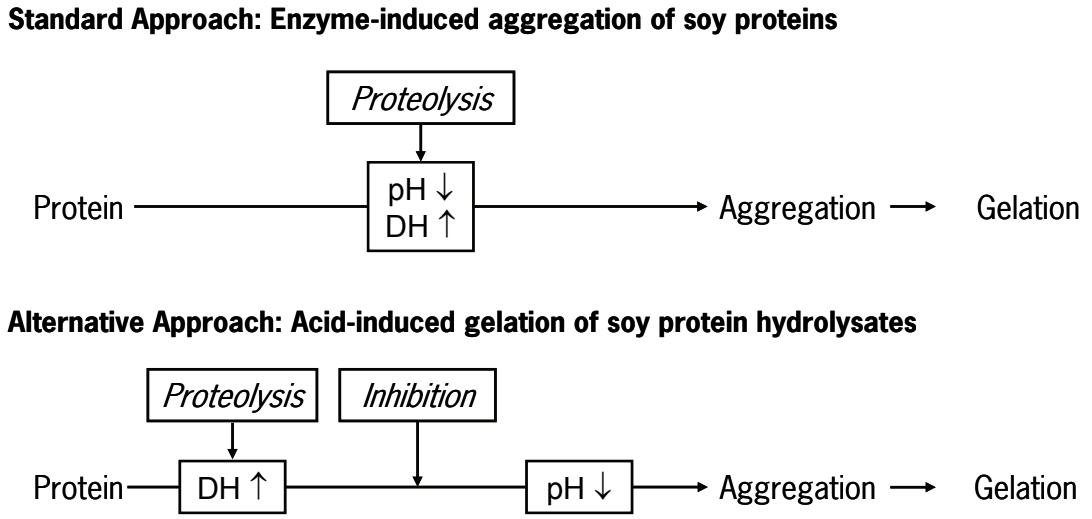


**Figure 7**: Schematic representation of two approaches to study the influence of hydrolysis on aggregation: Enzyme-induced aggregation of proteins (Standard Approach) and acid-induced aggregation of protein hydrolysates (Alternative Approach).

## AIM AND SETUP OF THIS THESIS

The aim of the work presented in this thesis is to understand how limited hydrolysis can strengthen the aggregation behavior of soy protein material, with a focus on the individual contribution of the storage proteins glycinin and β-conglycinin. This is done by uncoupling the enzymatic degradation and aggregation. Subsequently, it is focused on the identification of those regions in soy proteins from which peptides originate that have strong aggregating properties.

In **Chapter 2** it is described how the onset-pH of gelation of soy protein isolates, hydrolyzed with subtilisin Carlsberg, changes as a function of pH and ionic strength. **Chapter 3** elucidates the contribution of glycinin and β-conglycinin-derived peptides to the gelation of soy protein hydrolysates. **Chapter 4** describes how other enzymes than subtilisin Carlsberg, namely papain, bromelain, chymotrypsin, and trypsin change the aggregation behavior of SPI, glycinin, and β-conglycinin.

**Chapter 5, 6, and 7** together present a new approach to characterize those regions in soy proteins from which peptides originate that have strong aggregating properties. In **Chapter 5** data are presented for the prediction of the absorbance of peptides at 214 nm based on their individual amino acid sequences. **Chapter 6** describes a novel method, using quantitative RP-HPLC-MS for determination of those regions in proteins from which peptides, having similar functional properties, are released. In **Chapter 7**, the method presented in **Chapter 6** is applied to soy glycinin with the objective to elucidate which fragments from glycinin have the strongest tendency to aggregate after enzymatic hydrolysis. Finally, a general discussion on the work presented in this thesis is given in **Chapter 8**.

## REFFERECES


1. Endres, J.G. *Soy Protein Products; Characteristics, Nutritional Aspects and Utilization*. AOCS Press: Champaign, IL, 2001.
2. Liu, K. *Soybeans; Chemistry, Technology and Utilization*. 1st ed.; Chapman & Hall: New York, 1997.
3. Kilara, A.; Panyam, D. Peptides from milk proteins and their properties. *Crit. Rev. Food Sci. Nutr.* **2003,** *43*, (6), 607-633.
4. Panyam, D.; Kilara, A. Enhancing the functionality of food proteins by enzymatic modification. *Trends Food Sci. Technol.* **1996,** *7*, (4), 120-125.
5. Utsumi, S.; Kinsella, J.E. Forces involved in soy protein gelation - Effects of various reagents on the formation, hardness and solubility of heat-induced gels made from 7S, 11S, and soy isolate. *J. Food Sci.* **1985,** *50*, (5), 1278-1282.
6. Kohyama, K.; Nishinari, K. The effect of glucono-δ-lactone on the gelation time of soybean 11S protein: Concentration dependence. *Food Hydrocolloids* **1992,** *6*, (3), 263-274.
7. Saio, K.; Kamiya, M.; Watanabe, T. Food processing characteristics of soybean-11S and soybean-7S proteins. 1. Effect of difference of protein components among soybean varieties on formation of tofu-gel. *Agric. Biol. Chem.* **1969,** *33*, (9), 1301-1308.
8. Wolf, W.J.; Nelsen, T.C. Partial purification and characterization of the 15S globulin of soybeans, a dimer of glycinin. *J. Agric. Food Chem.* **1996,** *44*, (3), 785-791.
9. Utsumi, S.; Matsumura, Y.; Mori, T. Structure-function relationships of soy proteins. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Praf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 257-291.
10. Catsimpoolas, N.; Ekenstam, C. Isolation of α-, β- and γ-conglycinins. *Arch. Biochem. Biophys.* **1969,** *129*, (2), 490-497.
11. Fehr, W.R.; Hoeck, J.A.; Johnson, S.L.; Murphy, P.A.; Nott, J.D.; Padilla, G.I.; Welke, G.A. Genotype and environment influence on protein components of soybean. *Crop Science* **2003,** *43*, (2), 511-514.
12. Ji, M.P.; Cai, T.D.; Chang, K.C. Tofu yield and textural properties from three soybean cultivars as affected by ratios of 7S and 11S proteins. *J. Food Sci.* **1999,** *64*, (5), 763-767.
13. Cai, T.D.; Chang, K.C. Processing effect on soybean storage proteins and their relationship with tofu quality. *J. Agric. Food Chem.* **1999,** *47*, (2), 720-727.
14. Beilinson, V.; Chen, Z.; Shoemaker, R.C.; Fischer, R.L.; Goldberg, R.B.; Nielsen, N.C. Genomic organization of glycinin genes in soybean. *Theor. Appl. Genet.* **2002,** *104*, (6-7), 1132-1140.
15. Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallon, B.J.; Fischer, R.L.; Sims, T.L.; Drews, G.N.; Goldberg, R.B. Characterization of the glycinin gene family in soybean. *Plant Cell* **1989,** *1*, (3), 313-328.
16. Scallon, B.; Thanh, V.H.; Floener, L.A.; Nielsen, N.C. Identification and characterization of DNA clones encoding Group-II glycinin subunits. *Theor. Appl. Genet.* **1985,** *70*, (5), 510-519.
17. Fischer, R.L.; Goldberg, R.B. Structure and flanking regions of soybean seed protein genes. *Cell* **1982,** *29*, (2), 651-660.
18. Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984,** *179*, 125-142.
19. Lakemond, C.M.M.; de Jongh, H.H.J.; Hessing, M.; Gruppen, H.; Voragen, A.G.J. Soy glycinin: Influence of pH and ionic strength on solubility and molecular structure at ambient temperatures. *J. Agric. Food Chem.* **2000,** *48*, (6), 1985-1990.

20. Adachi, M.; Takenaka, Y.; Gidamis, A.B.; Mikami, B.; Utsumi, S. Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **2001,** *305*, (2), 291-305.
21. Brooks, J.R.; Morr, C.V. Current aspects of soy protein fractionation and nomenclature. *J. Am. Oil Chem. Soc.* **1985,** *62*, (9), 1347-1354.
22. Thanh, V.H.; Shibasaki, K. β-Conglycinin from soybean proteins - Isolation and immunological and physicochemical properties of monomeric forms. *Biochim. Biophys. Acta* **1977,** *490*, (2), 370-384.
23. Yamauchi, F.; Yamagishi, T. Carbohydrate sequence of a soybean-7S protein. *Agric. Biol. Chem.* **1979,** *43*, (3), 505-510.
24. Yamauchi, F.; Sato, M.; Sato, W.; Kamata, Y.; Shibasaki, K. Isolation and identification of a new type of β-conglycinin in soybean globulins. *Agric. Biol. Chem.* **1981,** *45*, (12), 2863-2868.
25. Sykes, G.E.; Gayler, K.R. Detection and characterization of a new β-conglycinin from soybean seeds. *Arch. Biochem. Biophys.* **1981,** *210*, (2), 525-530.
26. Kinsella, J.E. Functional-properties of soy proteins. *J. Am. Oil Chem. Soc.* **1979,** *56*, (3), 242-258.
27. Wolf, W.J. Soybean proteins - Their functional, chemical, and physical properties. *J. Agric. Food Chem.* **1970,** *18*, (6), 969-976.
28. Adler-Nissen, J. *Enzymic Hydrolysis of Food Proteins*. 1st ed.; Elsevier Applied Science Publishers: London, England, 1986.
29. Adler-Nissen, J. Determination of the degree of hydrolysis of food protein hydrolysates by trinitrobenzenesulfonic acid. *J. Agric. Food Chem.* **1979,** *27*, 1256-1262.
30. Hernandez, M.J.M.; Domingo, E.B.; Camanas, R.M.V.; Alvarezcoque, M.C.G. Use of the ortho-phthalaldehyde and N-acetyl-L-cysteine reagent in the evaluation of milk-proteins. *J. Dairy Sci.* **1991,** *74*, (6), 1779-1785.
31. Hernandez, M.J.M.; Camanas, R.M.V.; Cuenca, E.M.; Alvarezcoque, M.C.G. Determination of the protein and free amino-acid content in a sample using O-phthalaldehyde and N-acetyl-L-cysteine. *Analyst* **1990,** *115*, (8), 1125-1128.
32. Beynon, R.; Bond, J.S. *Proteolytic Enzymes*. 2nd ed.; Oxford University Press: Oxford, England, 2001.
33. Creusot, N. Enzyme-induced aggregation of whey proteins with *Bacillus licheniformis* protease. *Ph.D. Thesis*, Wageningen University, Wageningen, The Netherlands, **2006**.
34. Jung, S.; Murphy, P.A.; Johnson, L.A. Physicochemical and functional properties of soy protein substrates modified by low levels of protease hydrolysis. *J. Food Sci.* **2005,** *70*, (2), C180-C187.
35. Jung, S.; Roussel-Philippe, C.; Briggs, J.L.; Murphy, P.A.; Johnson, L.A. Limited hydrolysis of soy proteins with endo- and exoproteases. *J. Am. Oil Chem. Soc.* **2004,** *81*, (10), 953-960.
36. Kim, S.Y.; Park, S.W.; Rhee, K.C. Functional properties of proteolytic enzyme modified soy protein isolate. *J. Agric. Food Chem.* **1990,** *38*, 651-656.
37. Zhong, F.; Xin, Y.; Yue, L.; Shoemaker, C.F. Papain-induced gelation of soy glycinin (11S). *J. Food Sci.* **2006,** *71*, (5), E232-E237.
38. Zhong, F.; Wang, Z.; Xu, S.Y.; Shoemaker, C.F. The evaluation of proteases as coagulants for soy protein dispersions. *Food Chem.* **2007,** *100*, (4), 1371-1376.
39. Mohri, M.; Matsushita, S. Improvement of water absorption of soybean protein by treatment with bromelain. *J. Agric. Food Chem.* **1984,** *32*, 486-490.
40. Lee, C.W., and Jo, G.Y. The microstructures of soybean milk curds prepared by different coagulation methods. *J. Food Sci. Nutr.* **1997,** *2*, 259-262.

41. Murata, K.; Kusakabe, I.; Kobayashi, H.; Akaike, M.; Park, Y.W.; Murakami, K. Studies on the coagulation of soymilk-protein by commercial proteinases. *Agric. Biol. Chem.* **1987,** *51*, (2), 385-389.

42. Murata, K.; Kusakabe, I.; Kobayashi, H.; Kiuchi, H.; Murakami, K. Selection of commercial enzymes suitable for making soymilk-curd. *Agric. Biol. Chem.* **1987,** *51*, (11), 2929-2933.

43. Murata, K.; Kusakabe, I.; Kobayashi, H.; Kiuchi, H.; Murakami, K. Functional-properties of 3 soymilk curds prepared with an enzyme, calcium salt and acid. *Agric. Biol. Chem.* **1988,** *52*, (5), 1135-1139.

44. Park, Y.W., Kusakabe, I., Kobayashi, H., and Murakami, K. Production and properties of a soymilk-clotting enzyme system from a microorganism. *Agric. Biol. Chem.* **1985,** *49*, (11), 3215-3219.

45. Park, Y.W.K., H, Kusakabe, I. and Murakami, K. Purification and characterization of soymilk-clotting Enzymes from Bacillus sp. K-295G-7. *Agric. Biol. Chem.* **1987,** *51*, (9), 2343-2349.

46. Kobayashi, H.; Park, Y.W.; Kaneda, M.; Kusakabe, I.; Yoshida, S.; Murakami, K. Substrate-specificity of soymilk-clotting enzyme-I and enzyme-II from *Bacillus Sp.* K-295G-7 toward peptide-substrates. *Agric. Biol. Chem.* **1988,** *52*, (9), 2351-2353.

47. Yasuda, M., Aoyama, M., Sakaguchi, M., Nakachi, K. and Kobamoto, N. Purification and characterization of a soybean-milk-coagulating enzyme from Bacillus pumilus TYO-67. *Appl. Microbiol. Biotechnol.* **1999,** *51*, (4), 474-479.

48. Aoyama, M.; Yasuda, M.; Nakachi, K.; Kobamoto, N.; Oku, H.; Kato, F. Soybean-milk-coagulating activity of *Bacillus pumilus* derives from a serine proteinase. *Appl. Microbiol. Biotechnol.* **2000,** *53*, (4), 390-395.

49. Fuke, Y.; Sekiguchi, M.; Msuoka, H. Nature of stem bromelain treatments on the aggregation and gelation of soybean proteins. *J. Food Sci.* **1985,** *50*, 1283-1288.

50. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.

51. Nagai, K.; Inouye, K. Insights into the reaction mechanism of the coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2004,** *52*, (15), 4921-4927.

52. Nielsen, P.M. Functionality of protein hydrolysates. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Paraf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 443-472.

53. Mahmoud, M.I. Physicochemical and functional-properties of protein hydrolysates in nutritional products. *Food Technol.* **1994,** *48*, (10), 89-95.

54. Walstra, P.; Wouters, J.T.M.; Geurts, T.J. *Dairy Science and Technology*. 2nd ed.; CRC Press, Taylor & Francis Group: Boca Raton, 2006.

55. Berg, J.M.; Tymoczko, J.L.; Stryer, L. *Biochemistry*. 5th ed.; W.H. Freeman and Company: New York, 2002.

56. Ju, Z.Y.; Otte, J.; Madsen, J.S.; Qvist, B. Effects of limited proteolysis on gelation and gel properties of whey protein isolate. *J. Dairy Sci.* **1995,** *78*, 2119-2128.

57. Doucet, D.; Gauthier, S.F.; Foegeding, E.A. Rheological characterization of a gel formed during extensive enzymatic hydrolysis. *J. Food Sci.* **2001,** *66*, 711-715.

58. Otte, J.; Ju, Z.Y.; Faergemand, M.; Lomholt, S.B.; Qvist, K.B. Protease-induced aggregation and gelation of whey proteins. *J. Food Sci.* **1996,** *61*, 911-923.

59. Johansen, J.T.; Ottesen, M.; Svendsen, I.; Wybrandt, G. Degradation of B-chain of oxidized insulin by 2 subtilisins and their succinylated and N-carbamylated derivatives. *C. R. Trav. Lab. Carlsberg* **1968,** *36*, (20), 365-384.

# CHAPTER 2

# Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins

**ABSTRACT**

Acid-induced cold gelation of soy protein isolate hydrolysates was studied. Hydrolysates with degrees of hydrolysis (DH) up to 9.0 % were prepared by using subtilisin Carlsberg. The enzyme was inhibited to uncouple the hydrolysis from the subsequent gelation; the latter was induced by addition of glucono-δ-lactone. Visual observations, confocal scanning laser microscopy images and the elasticity modulus showed that hydrolysates gelled at higher pH values with increasing DH. The non-hydrolyzed soy protein isolate gelled around pH 6.0, whereas a DH = 4.5 % hydrolysate gelled around pH 7.6. Gels made from hydrolysates had a softer texture when manually disrupted and showed syneresis below a pH of 5-5.5. Monitoring of gelation by measuring the development of the storage modulus could be replaced by measuring the pH-onset of aggregate formation ($pH_{Aggr\text{-}onset}$) using turbidity measurements. The rate of acidification was observed to also influence this $pH_{Aggr\text{-}onset}$. Changes in ionic strength (0.03, 0.2 and 0.5 M) had only a minor influence on the $pH_{Aggr\text{-}onset}$, indicating that the aggregation is not simply a balance between repulsive electrostatic and attractive hydrophobic interactions, but is much more complex.

Based on: Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J., *J. Agric. Food Chem.* **2005,** *53*, 1031-1038.

## INTRODUCTION

Soy proteins are used in the food industry for their ability to form heat-induced gels *(1)*. This gelation involves the formation of aggregates, which subsequently form a more or less continuous network. The properties of heat-induced gels depend on the conditions during aggregate formation, such as protein concentration, temperature and duration of heat treatment, ionic strength and pH *(2, 3)*. Gels can also be prepared at ambient temperature, as reviewed by Bryant and McClements *(4)*. In this so-called cold gelation process two stages can be distinguished: (i) the preparation of a heat-denatured globular protein solution; (ii) the induction of gelation at ambient temperature. Using heat-denatured soy proteins, gelation can be induced by gradual acidification *(5)*. When the pH is lowered towards the iso-electric point the electrostatic repulsion between proteins decreases *(6)*, which results in a turbid gel due to random aggregation of the proteins *(7)*. Upon lactic acid fermentation, during which sugars are metabolized into lactic acid, the pH decreases gradually, eventually resulting in formation of a gel when approaching the iso-electric pH of the protein. This gel formation can be studied in the absence of acid-producing micro-organisms, by the use of glucono-δ-lactone (GDL) *(8)*. Since GDL is water-soluble and slowly hydrolyses into gluconic acid, its addition results in a gradual and homogeneous decrease in pH without the need for stirring. This is in contrast to acidification by addition of organic or inorganic acids, which requires mixing during acidification, resulting in inhomogeneous gel formation.

In addition to heating and acidification, enzymatic hydrolysis can be used to improve the functional properties of proteins *(9)*. In general hydrolysis is assumed to be unfavorable for the gelling properties of proteins, since it increases the number of charged groups and reduces the molecular weight, which hamper gelation *(10)*. On the other hand, the structure of the protein is altered during hydrolysis and buried hydrophobic groups become exposed, which might result in aggregate formation. Under certain conditions these aggregates can form gel networks, and in some cases, hydrolysis has been observed to enhance the gelling properties of the protein *(11-14)*.

Limited hydrolysis of whey proteins by *Bacillus licheniformis* protease (BLP) was shown to result in an increased gel strength after heating of hydrolysates at neutral pH *(11)*. Doucet and co-workers *(13)* and Otte and co-workers *(14)* observed that during extensive hydrolysis of bovine whey proteins with Alcalase and limited hydrolysis with BLP, respectively, gels were formed. For soy proteins it was found that by the action of bromelain gelation could be induced *(12)*. It has also been shown that different, mainly alkaline and neutral, proteinases of microbial origin, can induce coagulation of proteins in soy milk *(15)*. In accordance with this, Inouye *(16)* recently reported that subtilisin Carlsberg is able to induce aggregation during hydrolysis of soy protein isolates, but gelling properties were not studied. It might be expected that under certain conditions these aggregates would form a gel network upon hydrolysis. Since the hydrolysis of soy proteins can take place at ambient temperature, hydrolysis may be a way to modulate the structure of soy proteins to improve their

performance in the cold gelation process. In order to be able to modulate the texture of soy protein products, the aggregation behavior of the protein hydrolysates and the parameters that influence this behavior, such as the degree of hydrolysis (DH), pH and ionic strength, should be understood. To our knowledge no detailed information is available on the influence of these parameters on the gelation of soy protein hydrolysates.

The purpose of this study was, therefore, to investigate the influence of DH, pH and ionic strength on the aggregation and gelation behavior of soy protein hydrolysates, made with subtilisin Carlsberg at ambient temperature. In particular, subtilisin Carlsberg was chosen because it is the main enzyme in Alcalase, which is a commercially available, food-grade, industrial enzyme mixture *(17)*.

## MATERIALS AND METHODS

### Materials and chemicals

Hyland soybeans (non-GMO) were supplied by Fa. L.I. Frank (Twello, The Netherlands). Subtilisin Carlsberg from *Bacillus licheniformis* (10 units/mg solid: one unit releases color, equivalent to 1.0 μmole (181 μg) of tyrosine per min from casein at pH 7.5 at 37°C as measured using the Folin-Ciocalteu reagent); glucono-δ-lactone (GDL) and phenylmethylsulphonyl fluoride (PMSF) were obtained from Sigma Chemical Co. (St. Louis, MO, USA; article no. P-5380, P-7626 and G-4750, respectively). All other chemicals were of analytical grade and were purchased from Merck (Darmstadt, Germany).

### Analysis of the protein content

The nitrogen content of various samples was estimated by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milan, Italy) according to the manufacturer's instructions. In this work soy protein isolate (SPI) is regarded to only containing glycinin and β-conglycinin in a 64/36 ratio as determined by size-exclusion chromatography. Based on this ratio, the nitrogen conversion factor for SPI is 5.58, based on the average amino acid composition of glycinin and β-conglycinin as found in Swiss-Prot and only based on the polypeptide chain (www.expasy.org; Primary accession numbers used were: P04776 (Glycinin 1; A1a and B2 polypeptides), P04405 (Glycinin 2; A2 and B1a polypeptides), P11828 (Glycinin 3; A1b and B1b polypeptides), P02858 (Glycinin 4: A5, A4 and B3 polypeptides), P04347 (Glycinin 5; A3 and B4 polypeptides)), P11827(β-Conglycinin, α'-chain), P13916 (β-Conglycinin, α-chain) and P25974 (β-Conglycinin, β-chain)).

### Preparation of the soy protein isolate (SPI)

Soybean meal (SBM) was prepared by crushing soybeans with a Condux-Werk LV 15M (Condux-Werk, Wolfgang bei Hanau, Germany), followed by milling the crushed beans in a Fritsch Pulverisette 14702 using a 0.5 mm sieve (Fritsch Gmb, Albisheim, Germany). Milling

was performed in the presence of solid $CO_2$ to prevent excessive heating (volume ratio soybean : $CO_2$ was approximately 4 : 1). The SBM was defatted three times with hexane at room temperature (w/v ratio SBM : hexane = 1 : 10) followed by drying to the air.
To obtain the SPI, the defatted SBM was suspended in a 30 mM Tris-HCl buffer (pH 8.0), containing 10 mM 2-mercaptoethanol, and stirred at ambient temperature for 1.5 hours (w/v ratio SBM : buffer = 1 : 10). After removal of the insoluble parts by centrifugation (30 min, 12,000 x *g*, 10ºC) the supernatant was brought to pH 4.8 with 2 M HCl to induce precipitation of the proteins. After 2 hours at room temperature the dispersion was centrifuged (20 min., 12,000 x *g*, 10ºC). The precipitate was washed twice with Millipore water (approximate v/v ratio precipitate : water = 1 : 9) followed by suspending it in water and adjusting its pH to 8.0 before it was freeze-dried. The protein content (N x 5.58) of the freeze-dried SPI was 83.3 (± 1.6) % (w/w).

**Heat denaturation of SPI**

Prior to each hydrolysis experiment an SPI solution was freshly prepared by suspending the freeze-dried SPI in Millipore water containing 0.032 % (w/v) sodium azide to a concentration of 8.0 % (w/w) (6.7 % (w/w) protein). If necessary, the pH was adjusted to 8.0 with 2 M NaOH. The dispersion was stirred overnight at 4ºC. After equilibration to ambient temperature the pH was readjusted to 8.0, if necessary. The slightly turbid solution was subsequently centrifuged (30 min, 22,000 x *g*, 20ºC) and the supernatant was heated in a water bath at 95ºC for 30 min in a closed bottle, followed by immediate cooling in a water bath at room temperature. To determine whether the heating step of the SPI solution was sufficient to irreversibly denature the proteins, DSC experiments were performed on a VP-DSC Micro Calorimeter (MicroCal Inc., Northampton, MA, USA). DSC experiments were conducted with the SPI solution before and after the heat treatment (30 min, 95ºC). The SPI solutions were diluted to a protein concentration of approximately 5 mg/mL with 10 mM potassium phosphate buffer (pH 7.6) containing 0.5 M NaCl. Thermograms were recorded from 20 to 105ºC with a heating rate of 1 K/min. The non-heated SPI showed two endothermic transitions at about 79ºC and 95ºC, representing the denaturation of β-conglycinin and glycinin respectively *(18)*. The heated SPI showed no endothermic effect. The heated SPI solution, which will further be denoted as SPI solution (DH = 0 %), had a clear, slightly opaque appearance and had a protein concentration of 6.2 ± 0.1 (w/w) %.

**Proteolysis of SPI**

Hydrolysates with various DH values (proportion of peptide bonds cleaved; 1.1, 2.2, 3.4, 4.5, 5.9, 6.7 and 9.0 %) were prepared from the SPI solution by hydrolysis at pH 8.0 and at 40ºC using subtilisin Carlsberg. Hydrolysates were freshly prepared prior to each experiment. The pH and DH were controlled using the pH-stat method by using a 719S Titrino (Metrohm ion analysis, Herisau Switzerland) *(19)*. The $h_{tot}$ used was 8.77 meqv/gram and was calculated,

based on the amino acid composition of SPI, as also used for calculation of the nitrogen conversion factor for the Dumas method The enzyme concentration varied from 3 to 6 units/gram substrate, while the molarity of the NaOH solution used to maintain the pH varied from 0.1 to 0.5 M. The higher the desired DH was, the higher the amount of enzyme and the higher the molarity of NaOH that was used. The enzyme was dissolved in Millipore water and directly added to the SPI solution. When the desired DH was reached, the enzymatic hydrolysis was stopped *(20)* by addition of a 100 mM PMSF stock solution in 2-propanol to a final concentration of 1 mM. Instability of the PMSF in aqueous solutions *(21)* resulted in a slight decrease in pH directly after addition. Approximately 10 min after PMSF addition, the pH was stable, and the hydrolysate was cooled down to room temperature. Although PMSF did not inhibit subtilisin Carlsberg completely, the effect of the remaining activity was negligible for the timescales used in all experiments, as tested by size exclusion chromatography. The average hydrolysis time was approximately 30 min. At DH values below 9.0 %, the hydrolysates had a similar appearance to the heated SPI solution. During more extensive hydrolysis (DH ≥ 9.0 %), however, the solution became turbid.

**Preparation of cold-set gels**

Gels were prepared from the SPI solution, and the hydrolysates thereof, by addition of GDL, which was freshly dissolved in ice water before each experiment. Two different gels were prepared from each sample, one with a final pH of 5.2 (±0.2) and the other with a final pH of 3.1 (±0.1), by adding 0.45-0.6 % (w/w) and 5-6 % (w/w) GDL, respectively. The buffering capacity of the hydrolysates increased with increasing DH and, therefore, more GDL was added at the higher DH values.

To correct for the different amounts of NaOH added during hydrolysis, the ionic strength and the protein concentration in the gels were standardized by the addition of NaCl solution (2 or 5 M) and/or Millipore water before all further experiments. The final protein content of the gels was approximately 4.0 % (w/w) with ionic strengths of 0.03, 0.2 or 0.5 M.

**Confocal Scanning Laser Microscopy (CSLM)**

Imaging was carried out of gels made from the SPI solution and hydrolysates with a DH of 2.2 and 4.5 %, all at an ionic strength of 0.03 M. A Confocal Scanning Laser Microscope type TCS-SP (Leica Microsystems AG, Wetzlar, Germany), configured with an ArKr laser for single-photon excitation (JDS Uniphase, San Jose, CA) was used. By adding increasing amounts of GDL, gels with various final pH values (pH 5-8) were obtained. The protein gels were stained by applying 10 µl of a 0.2 % (w/v) Rhodamine B solution to 1 mL of gel. The 568 nm laser line was used for excitation to induce a fluorescent emission of Rhodamine B, which was monitored between 600 and 700 nm. The pore-size and area-% covered by pores on the picture were estimated using Leica Qwin software (Leica Microsystems AG, Wetzlar, Germany). Directly after the picture was made, the gel was disrupted and the pH was

measured using a Metrohm Biotrode micro pH glass electrode (Metrohm ion analysis, Herisau, Switzerland).

**Turbidity experiments**

Aggregate formation in the samples during the GDL acidification was followed as described by Alting and co-workers *(22)*, by measuring the turbidity as the absorbance at 500 nm at 25ºC on a Cary 1E UV-vis spectrophotometer (Varian Nederland BV, The Netherlands) equipped with a temperature controller. The pH was monitored simultaneously in samples placed in a water bath at 25ºC. Samples were measured in glass cuvettes with a path length of 1 mm to prevent sedimentation. The parallel measurement of the pH as function of time allowed us to monitor the turbidity as a function of pH at various DH values (0, 1.1, 2.2, 3.4, 4.5, and 5.9 %) and ionic strengths (0.03, 0.2 and 0.5 M).

**Rheological characterization of the gels**

The formation of a gel network in SPI solutions with DH values 0, 2.2 and 4.5 % ($I$ = 0.03 M) was followed by simultaneously monitoring the storage modulus (G'), the turbidity, and the pH as function of time at 25ºC. Small-amplitude oscillatory measurements were made with a Carri-Med $CLS^2$ 500 rheometer (TA Instruments, New Castle, DE, USA) using a double-walled-cylinder measuring unit ($R_1$, $R_2$, $R_3$, $R_4$ and h are 20, 20.38, 21.96, 22.38 and 20.5 mm, respectively). Immediately after the addition of GDL, samples were brought into the measuring unit and covered with a thin layer of paraffin oil to prevent evaporation. All experiments were conducted in oscillation at a strain of 1 % and a frequency of 1 Hz. The strain was within the linear region.

**Light scattering**

Dynamic light-scattering experiments were done using a Malvern Autosizer IIC Submicron particle size distribution analyzer (Malvern Instruments Ltd., Worcestershire, United Kingdom). The system consisted of a Malvern PCS41 optics unit with a 5 mW He-Ne laser and a Malvern K7032-ES correlator used in serial configuration. The Autosizer IIC worked at a fixed scattering angle of 90º, and the wavelength of the laser beam was 632.8 nm, resulting in a wave vector of 0.0187 $nm^{-1}$. Hydrolysates were diluted with 20 mM Tris-HCl buffer (pH 8.15), resulting in a protein concentration of approximately 20 mg/mL, before being transferred to 10 mm quartz cuvettes. The average apparent diameter of the protein aggregates was monitored at 20 (±0.1) ºC. The Malvern software was used in "Easy Mode", which means that the interval time was automatically adjusted to the "optimal" value. The apparent diameter of the particles in solution was calculated from a cumulant fit of the intensity autocorrelation function *(23)*. Aggregate sizes were estimated at least in triplicate.

### High-Performance Size-Exclusion Chromatography (HP-SEC)

Aliquots (300 µL) of SPI solution and hydrolysate solutions (approximately 10 mg protein/mL) were mixed with 510 µL 0.15 M Tris-HCl buffer (pH 8.0) containing 100 mM 1,4-dithiothreitol (DTT) and 8 M guanidinium chloride. After 1 h of stirring at ambient temperature, 310 µL acetonitrile and 200 µL 8 M guanidinium chloride, containing 2 % (v/v) trifluoroacetic acid (TFA), were added and the mixture was stirred for another 1 h. After centrifugation (15 min, 22,000 x *g*, 20ºC) the supernatants were applied to a Shodex Protein KW-803 column (300 x 8 mm; Showa Denko K.K., Japan). The column was equilibrated with 40 % (v/v) aqueous acetonitrile containing 0.1 % (v/v) TFA. The flow rate was 0.6 mL/min and the absorbance of the eluate was monitored at 220 nm. According to the manufacturer's specification, the void volume of the column was approximately 5.7 mL. The included volume was estimated using GlyTyr (238 Da) and was observed to be approximately 12.4 mL. For the SPI solution separate fractions were collected, and dried in an ALPHA-RVC CMC-1 rotating vacuum concentrator (CHRIST, Osterode am Harz, Germany) followed by analysis by gel electrophoresis.

### Gel electrophoresis

The protein composition of SPI, hydrolysates of various DH values and the various fractions of SPI collected during HP-SEC was determined using SDS-PAGE under reducing conditions (10 mM 2-mercaptoethanol) on a Mini-protean II System (Bio-Rad Laboratories, USA) according to the instructions of the manufacturer. Gels (12 %) were calibrated with marker proteins with molecular masses ranging from 14.4 to 94 kDa (Amersham Biosciences, Roosendaal, The Netherlands) and were stained with Coomassie Brilliant Blue.

## RESULTS

### Cold gelation of hydrolysates

Gels of the SPI solution and hydrolysates thereof with various DH values were made by gradual acidification. Upon lowering the pH from 8.0, one of the most striking effects was the observation that the higher the DH of the hydrolysates, the higher the pH at which gelation started. Just below this gelation pH the gels had an opaque appearance, which became more turbid upon further pH decrease. Around pH 5 syneresis could be observed in gels made from the hydrolysates, while in gels from SPI no syneresis was observed. Compared to gels made from SPI, gels made from hydrolysates could be disrupted more easily.

In Figure 1 the storage modulus as function of pH is presented for DH = 0, 2.2 and 4.5 % samples at $I$ = 0.03 M. Because of syneresis, this experiment was not carried out at pH values lower than 5. Upon lowering the pH of the DH 0, 2.2 and 4.5 % solutions, G' starts to increase at pH 6.3, 6.8 and 7.2, respectively. This confirms the visual observations described above. The maximum G'-values were in the range of 1,000 to 3,000 Pa.

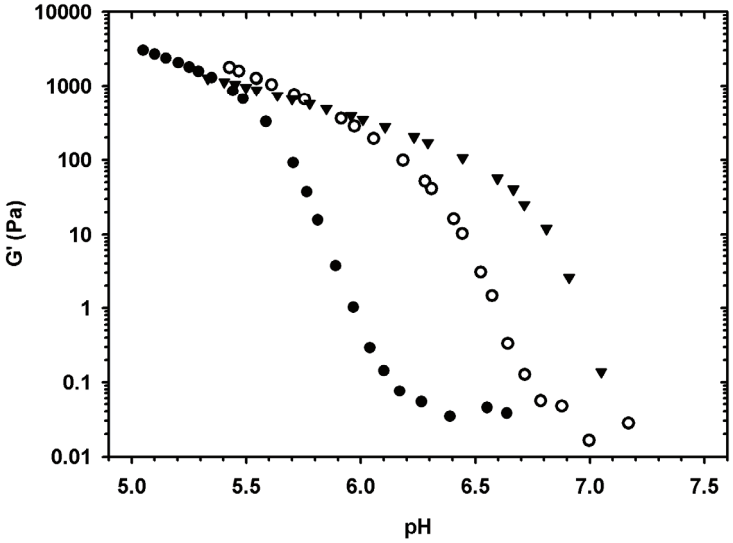


**Figure 1**: Development of the storage modulus (G') as a function of pH for SPI (●) and hydrolysates with DH 2.2 % (○) and 4.5 % (▼) at $I$ = 0.03M.

**CSLM of cold-gelled hydrolysates**

The microstructures of the cold-set gels were subsequently characterized by CSLM. Figure 2 shows the CSLM images of the gels prepared from SPI solutions with a DH of 0, 2.2 and 4.5 %, made at 16 h after the addition of GDL. The orange colored parts represent protein-bound Rhodamine, while the dark spots represent places were little or no protein is present, which can be assumed to be pores in the protein gel network. Visual observations, the area-% covered by the pores and the average pore diameter of the gels corresponding to the images shown in Figure 2 are presented in Table 1. When Figure 2 is compared with Table 1, the visual observations show that conditions at which no gel was formed coincided with a homogeneous distribution of the proteins. When the formation of gels was observed this coincided with the formation of a network, observed as the appearance of pores. These pores increased in average size with decreasing pH. The development of the total area-% covered by pores showed the same trend (Table 1).

When comparing samples at one DH value it can be observed that a coarser network is formed with decreasing pH, for all three DH values (Figure 2). At DH = 4.5 % *e.g.* no pores were observed at pH 8.1, whereas at pH 7.6 very small pores, with an average pore diameter of 0.22 μm were observed. At pH 6.9 the pores became much more abundant and had an average diameter of 0.56 μm. The largest pores, with an average diameter of 0.69 μm, were observed at pH 5.4 at DH = 4.5 %. In this sample syneresis was also observed.

When comparing samples at a fixed pH (Figure 2), the network seems to become coarser with increasing DH. This is confirmed by the data in Table 1, which show that at pH 5.4/5.5 the average pore diameter increases from 0.44 μm to 0.69 μm, when the DH is increasing from 0 to 4.5 %.

**Turbidity of cold-gelled hydrolysates**

Aggregate formation often precedes gelation and can be monitored in acid-induced gels as an increase in turbidity, because gels that are formed around the p*I* of the protein usually consist of large random aggregates, and often result in a coarse protein network *(24, 25)*. Turbidity measurements were therefore used as a tool to study gel formation indirectly, with the advantage that different measurements could be performed simultaneously. Experiments were performed at various ionic strengths to obtain information about the contribution of electrostatic interactions in the aggregation process. The turbidity at 500 nm as a function of pH for DH 0, 2.2 and 4.5 % at $I$ = 0.03, 0.2 and 0.5 M is shown in Figure 3.

At all the conditions used it was observed that with decreasing pH the turbidity started to increase at a specific pH, which indicates that at these conditions aggregates were formed. The development of the turbidity as function of pH followed a kind of S-curve until a maximum was reached. The pH at which aggregates start to form is denoted *pH of aggregation onset* ($pH_{Aggr\text{-}onset}$), which is defined as the intersection of the tangent at the point of inflection of the S-curve with the $x$-axis.

If turbidity is studied over a pH range from 3 to 8 the amount of GDL used to reach pH 3 results in too high a rate of acidification in the pH 5 to 8 range. Therefore, also a gel was prepared of the same SPI solution or hydrolysate, using a lower amount of GDL. The curves obtained of one hydrolysate using two different levels of GDL did not follow the same course (Figure 3), indicating that the rate of acidification affects the aggregation process. At a particular pH, the turbidity of the slowly acidified sample was higher compared to the fast-acidified sample; a low rate of acidification gives the proteins more time to aggregate.

The curves at $I$ = 0.03 M reached a maximum in turbidity around pH 4.5 which was followed by a decrease in turbidity below pH 4.5. The turbidity maximum decreased with increasing DH (Figure 3), while the pH of maximum turbidity remained constant. With increasing DH, the $pH_{Aggr\text{-}onset}$ increased to 5.9, 6.9 and 7.1 for DH 0, 2.2 and 4.5 %, respectively. This shift in aggregation pH to higher pH values with increasing DH is similar to that found in the rheological experiments and during the visual observations described above.

At higher ionic strengths similar turbidity patterns were observed as for $I$ = 0.03 M but some differences became apparent. At $I$ = 0.2 and 0.5 M the maximum turbidities reached were around pH 4.2 and 3.7, respectively (Figure 3). Thereafter, at $I$ = 0.2 M a slight decrease in turbidity was observed (Figure 3) below the pH of maximum turbidity, whereas at $I$ = 0.5 M this decrease was absent (Figure 3).

Table 2 shows the $pH_{Aggr\text{-}onset}$ obtained for different DH values at all the ionic strengths studied. These results show that the shift in $pH_{Aggr\text{-}onset}$ with increasing DH is on average the largest at $I$ = 0.2, while at $I$ = 0.03 M and 0.5 M, the $pH_{Aggr\text{-}onset}$ is slightly lower.

The observed shift in $pH_{Aggr\text{-}onset}$ must be the result of the proteolytic activity of the subtilisin Carlsberg, because simultaneous addition of subtilisin Carlsberg and its inhibitor PMSF to the SPI solution at 40ºC resulted in similar turbidity patterns as for the non-hydrolyzed SPI (no further data shown).

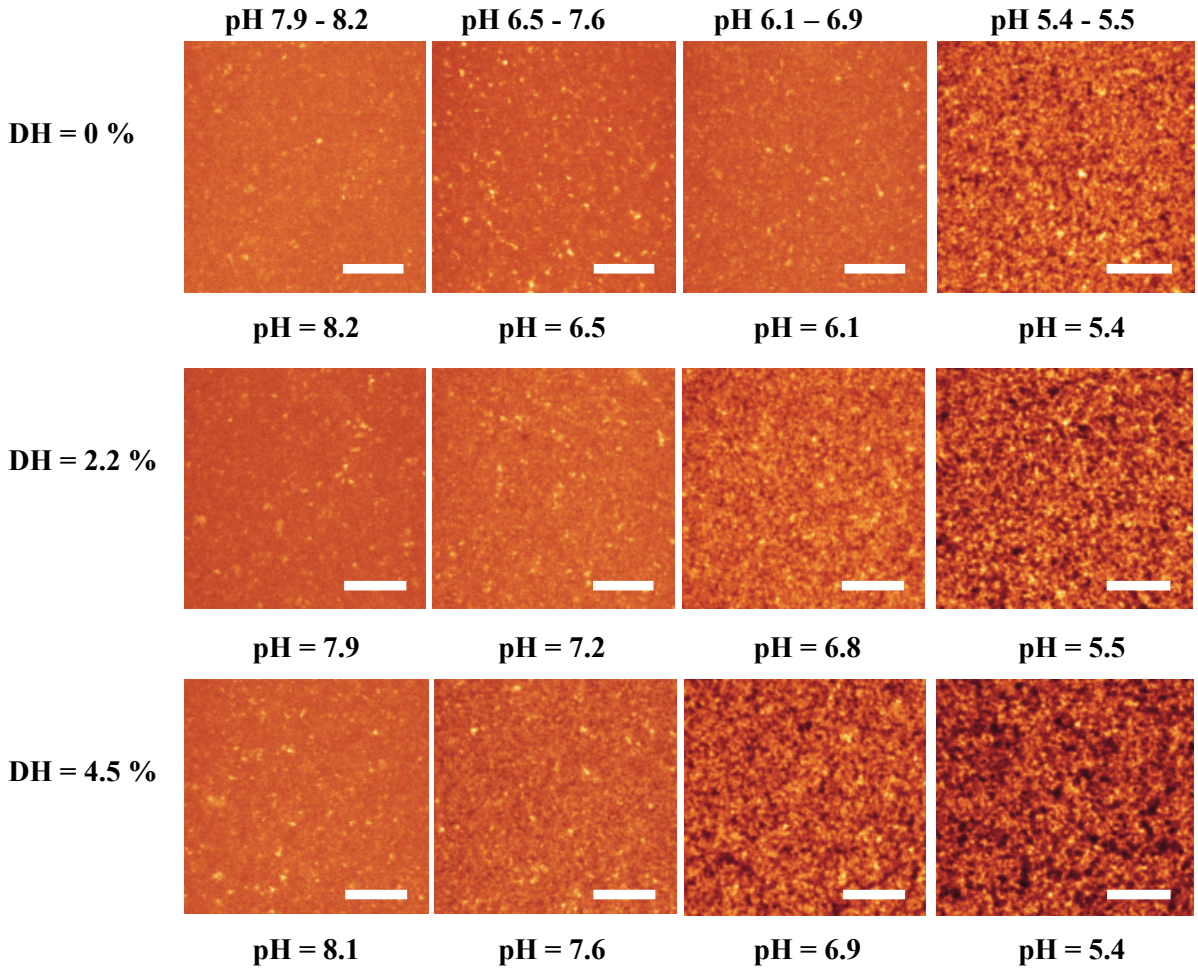


**Figure 2**: CSLM images of 4.0 % (w/w) protein gels prepared from SPI solutions with a DH of 0, 2.2 and 4.5 % ($I$ = 0.03 M). The bars represent a length of 10 µm. The pH-values of the gels are presented at the bottom of each image.

**Table 1**: Characteristics of the CSLM images as presented in Figure 2 of 4.0 % (w/w) protein gels prepared from SPI solutions with a DH of 0, 2.2 and 4.5 % ($I$ = 0.03 M).

| | | **pH 7.9-8.2** | **pH 6.5-7.6** | **pH 6.1–6.9** | **pH 5.4-5.5** |
|---|---|---|---|---|---|
| **DH = 0 %** | *Visual observation* | *Liquid* | *Liquid* | *Liquid* | *Turbid gel* |
| | Area % covered by pores | 0.1 % | 0.1 % | 0.1 % | 7.2 % |
| | *Average pore diameter (µm)* | *0.20 (±0.01)*[a] | *0.20 (±0.01)*[a] | *0.20 (±0.01)*[a] | *0.44 (±0.23)* |
| **DH = 2.2 %** | *Visual observation* | *Liquid* | *Liquid* | *Opaque gel* | *Turbid gel* |
| | Area % covered by pores | 0.1 % | 0.1 % | 2.4 % | 12.5 % |
| | *Average pore diameter(µm)* | *0.20 (±0.01)*[a] | *0.20 (±0.04)*[a] | *0.33 (±0.16)* | *0.57 (±0.33)* |
| **DH = 4.5 %** | *Visual observation* | *Liquid* | *Opaque gel* | *Turbid gel* | *Turbid gel*[b] |
| | Area % covered by pores | 0.1 % | 0.5 % | 11.1 % | 13.1 % |
| | Average pore diameter*(µm)* | *0.20 (±0.01)*[a] | *0.22 (±0.05)* | *0.56 (±0.35)* | *0.69 (±0.42)* |

[a] The minimum pore size that can be measured is 0.2 µm

[b] Syneresis was observed in this sample

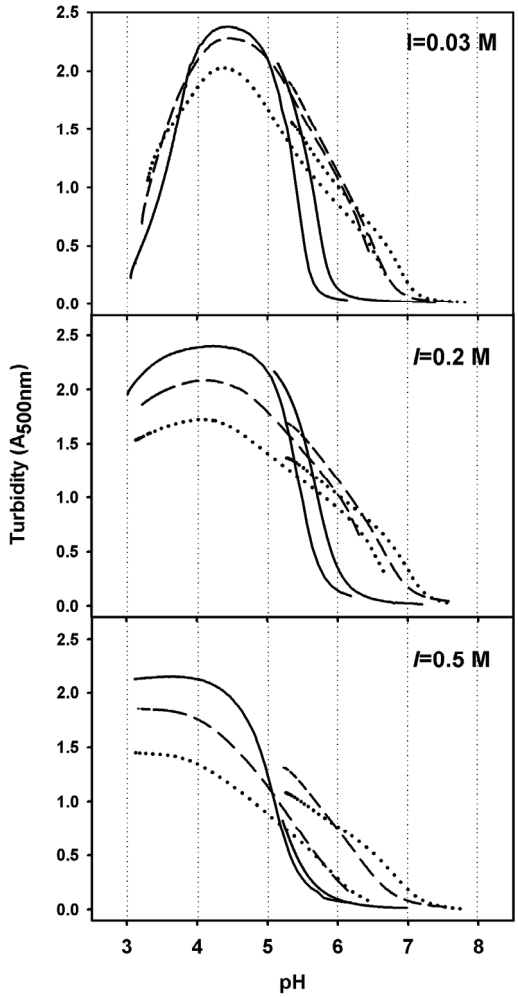


**Figure 3**: Turbidity at 500 nm for SPI (——) and SPI hydrolysates with DH 2.2 % (– – –) and 4.5 % (•••••) as a function of pH at $I$ = 0.03, 0.2 and 0.5 M.

**Table 2**: $pH_{Aggr\text{-}onset}$ of SPI and hydrolysates thereof with different DH values at $I$ = 0.03, 0.2 and 0.5 M.

| DH value (%) | $pH_{Aggr\text{-}onset}$ $I$ = 0.03 | $pH_{Aggr\text{-}onset}$ $I$ = 0.2 | $pH_{Aggr\text{-}onset}$ $I$ = 0.5 |
|---|---|---|---|
| 0 | 5.9 | 6.1 | 5.7 |
| 1.1 | 6.6 | 6.8 | 6.4 |
| 2.2 | 6.9 | 7.1 | 6.8 |
| 3.4 | 7.0 | 7.2 | 7.1 |
| 4.5 | 7.1 | 7.3 | 7.2 |
| 5.9 | 7.3 | 7.6 | 7.4[a] |

[a] Already resulted in a clear gel after ionic strength was adjusted to 0.5M.

### Aggregate size

Heat and hydrolysis induces the formation of aggregates. This aggregate formation was studied by light scattering. Table 3 shows the estimated average aggregate sizes obtained for the non-heated and heated SPI solution, and hydrolysates at various DH values.

**Table 3**: Effect of heating and hydrolysis on the hydrodynamic diameter of SPI and its hydrolysates.

| Sample | Diameter (nm)[a] |
|---|---|
| Non-heated SPI solution | 131 ± 2 |
| DH = 0 % | 197 ± 3 |
| DH = 1.1 % | 137 ± 2 |
| DH = 2.2 % | 137 ± 1 |
| DH = 4.5 % | 176 ± 2 |
| DH = 6.7 % | 413 ± 11 |
| DH = 9.0 % | 1584 ± 353 |

[a] Errors represent the standard error of the cumulant fits within one measurement.

Heating the SPI solution resulted in the formation of aggregates, which was observed as a slightly opaque SPI solution. Upon hydrolysis a decrease in aggregate size was observed at DH = 1.1 and 2.2 %, followed by a strong increase in aggregate size with a further increase in DH, as also observed by Ju and Kilara *(26)*. As the estimated average size of the aggregates at DH = 9.0 % was found to be close to the limit of the apparatus, care should be taken with the interpretation of these results.

### Protein and peptide size distribution

HP-SEC was carried out to obtain an indication of the size of the peptides present in the SPI solution and the hydrolysates.

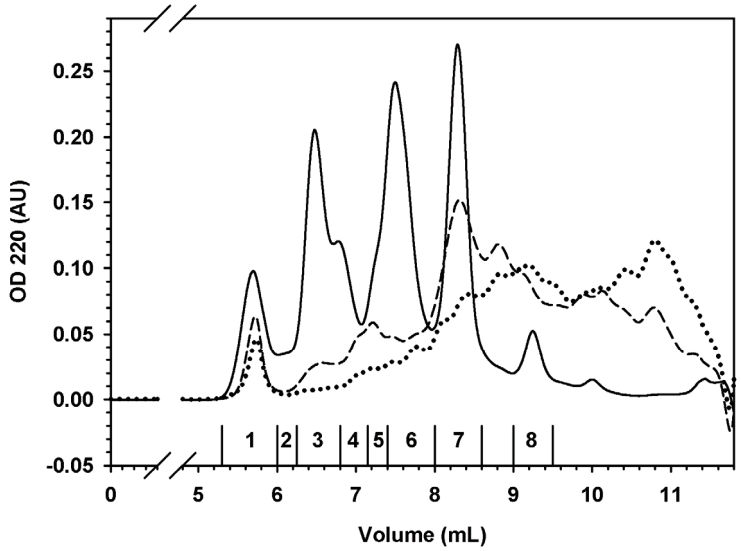


**Figure 4**: HP-SEC chromatogram of SPI (——), and its hydrolysates at 2.2 % (– – –) and 4.5 % (•••••). The numbers 1 to 8 refer to fractions collected of the DH = 0 % sample.

Figure 4 shows the HP-SEC chromatogram of the DH = 0, 2.2 and 4.5 % samples. The chromatogram of DH = 0% contains five main peaks (fractions 1, 3, 6, 7, and 8) and two shoulders (fractions 4 and 5). The fractions collected were analyzed by SDS-PAGE as shown in Figure 5, which also shows the protein composition of SPI. Fractions 1 and 2 elute around 5.7 mL, the void volume of the column, and therefore probably contain aggregates not dissociated by DTT and guanidinium. As can be observed in Figure 4, subtilisin Carlsberg was able to degrade the aggregates present in the first peak since the area of the peak decreased with increasing DH.

Figure 5 shows that fraction 3 contains several protein bands with molecular masses in the range of 55 to 70 kDa, probably representing the α-fraction of β-conglycinin. Fraction 4 shows a band around 50 kDa, which could be the β-subunit of β-conglycinin. A protein band around 36 kDa is present in fraction 5, representing the acid polypeptide A3 of glycinin. Fraction 6 has protein bands around 32 kDa, probably representing the acidic polypeptide of glycinin A1a, A2, A1b and A4. Fraction 7 contains a band around 20 kDa, which probably consists of all the basic polypeptides of glycinin *(27)*. A protein band around 10 kDa can be observed in fraction 8, probably representing the acidic polypeptide A5 of glycinin.

The chromatograms in Figure 4 clearly show that smaller fragments are formed with increasing DH and that already almost all the intact protein has been degraded at DH = 2.2 %. The SDS-PAGE gel in Figure 6, which shows the peptide patterns observed at various DH values, confirms this. Some new bands become apparent in the hydrolysates, which are not present in the non-hydrolyzed SPI solution, indicating that some more or less stable peptides are formed. In particular, the peptide with a molecular mass of approximately 25 kDa seems quite resistant to enzymatic hydrolysis, as also reported by Inouye *(16)*.
When the HP-SEC experiment was performed with samples prepared under non-reducing conditions, a minor shift to higher peptide sizes could be observed (data not shown). This was confirmed in an SDS-PAGE experiment under non-reducing conditions (data not shown). The minor shift to higher peptide size became less apparent with increasing DH.

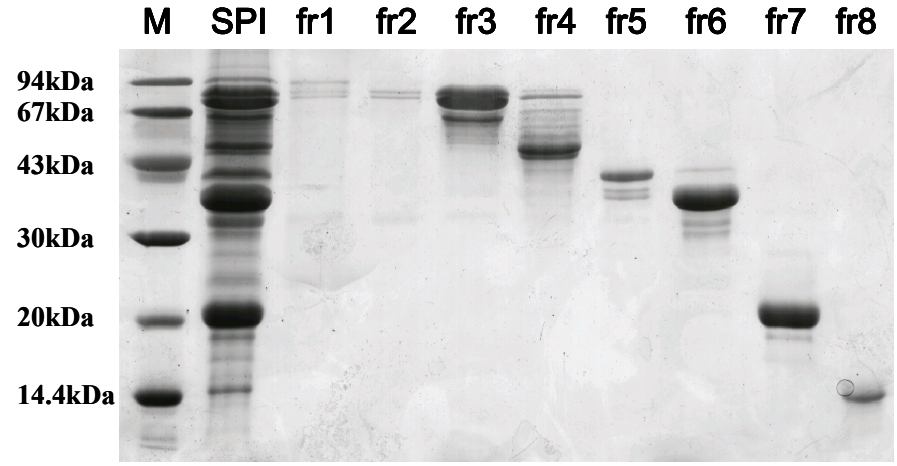


**Figure 5**: SDS-PAGE gel with the low molecular mass marker (M), the intact protein (SPI), and collected fractions 1 to 8 as indicated in the chromatogram of Figure 4.

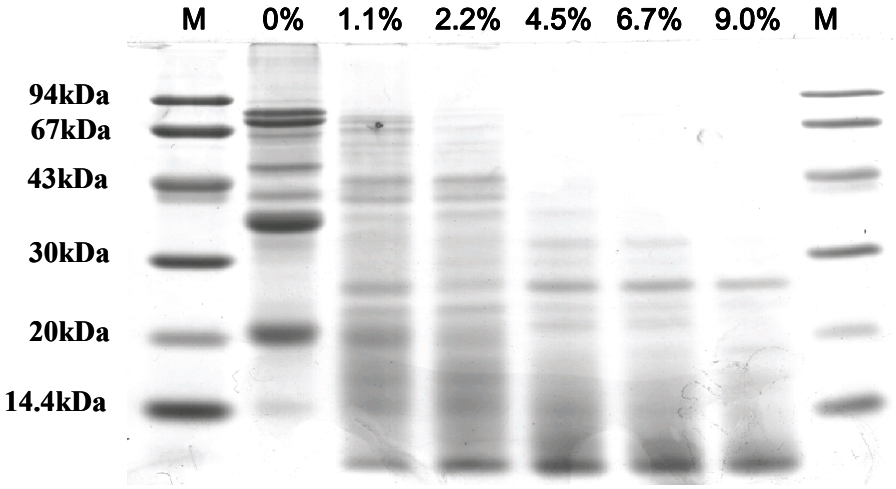


**Figure 6**: SDS-PAGE gel with the protein composition of different SPI samples with increasing DH value (0, 1.1, 2.2, 4.5, 6.7 and 9.0 %). The two outer lanes represent the marker (M).

## DISCUSSION

### Uncoupling of hydrolysis and gelation

In our experimental set-up, the preparation of soy protein hydrolysate gels consisted of heating an SPI solution, followed by hydrolysis and finally acid-induced cold gelation in separate steps. A prerequisite for this approach is that gelation occurred only upon acidification. These conditions were met by, on the one hand, choosing a protein concentration not resulting in extensive aggregation or gelation upon heating, and on the other hand by stopping the hydrolysis before visible aggregates were formed. This allowed hydrolysis and gelation to be studied separately. This approach is different from studies performed by *e.g.* Doucet and co-workers *(13)* and Otte and co-workers *(14)*, where gelation occurred already during the hydrolysis. The latter could therefore be defined as enzyme-induced gelation. Our approach is defined as acid-induced gelation of hydrolysates.

In our preparations enzyme-induced gelation also occurred, around an estimated DH of 9.0 %, when hydrolysis was performed without stirring and pH control. This is what could be expected based on the increase in aggregate size at DH = 9.0 % (Table 3), as also reported by Inouye and co-workers *(16)*. Because our aim was to uncouple the hydrolysis from the gelation, such high DH values were not studied in detail.

**Shift in aggregation and gelling pH**

The uncoupling of hydrolysis and gelation made it possible to study the effect of the pH on the gel formation. The rheological experiments showed that the higher the DH, the higher the pH at which gels could be formed. With the hydrolysates, gels could be made at pH values at which the non-hydrolyzed SPI did not yet form a gel. This makes the use of hydrolyzed SPI interesting for foods at neutral and slightly acidic pH. In addition, lower amounts of protein are needed when compared to heat-induced gels. The gels formed may also have a new texture that could be of interest of the food industry. Our results are in agreement with the observations made by Ipsen and co-workers *(28)*, who studied enzyme-induced gelation of whey protein systems at various pH values, and found that at pH values further away from the p*I* of the parental protein, a higher DH was necessary to form a gel.

When the shift in aggregation pH during acidification is studied, timescale effects have to be taken into account. These timescale effects can be observed clearly in Figure 3 as the difference between the slow and fast acidification. The curves of gels prepared from the same hydrolysate at the same ionic strength do not overlap, due to the difference in acidification rate. The slower the acidification rate, the more time the proteins will spend at one particular pH to form aggregates, resulting in a higher turbidity at that pH. These timescale effects become even more apparent when comparing data from measurements performed during acidification (rheological and turbidity measurements) with data from experiments performed after the acidification had finished (visual observations and CSLM imaging). The DH = 4.5 % ($I$ = 0.03 M) hydrolysate had formed a gel at pH 7.6 after 16 h of incubation (Table 1), while no increase in turbidity was observed at pH 7.6 for this hydrolysate during acidification (Figure 3).

Rheological measurements are a useful tool to study the shift in aggregation pH, but they have the disadvantage that they are time consuming, and cannot be used in systems that show syneresis, as occurred at pH values below 5-5.5. Measurements that are less time consuming and able to provide information about a broader pH range would therefore be more favorable. Turbidity experiments have the advantage that various samples can be measured in parallel, and by choosing a cuvet with a short light path, visible syneresis could in this case be prevented. These advantages enable measurements in a broader pH range. The turbidity measurements relate well with rheological experiments when G' is plotted on a logarithmic scale (Figure 7A), and can therefore be used to provide reliable indications about the gel formation of the preparations studied. When G' is plotted on a linear scale it can be seen that

aggregation, measured as increase in turbidity, precedes gelation, measured as the increase of G' (Figure 7B), as also observed by Alting and co-workers *(25)* .

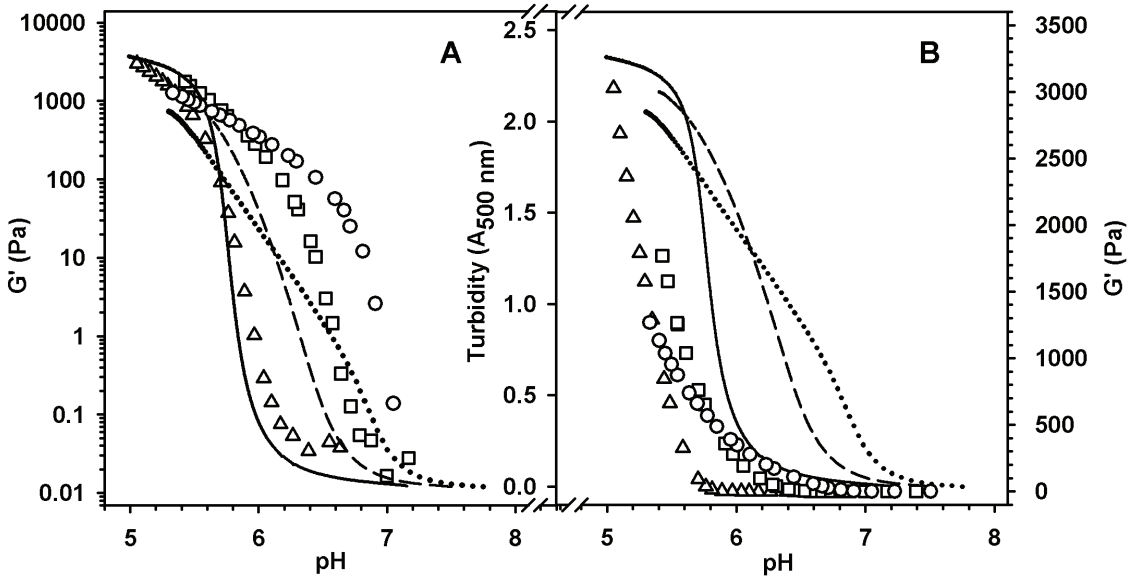


**Figure 7**: Development of the storage modulus (G') for SPI (Δ), and its hydrolysates at DH 2.2 % (□) and 4.5 % (○) and turbidity ($A_{500nm}$) of the SPI (——), and its hydrolysates at DH 2.2 % (– – –) and 4.5 % (•••••) as a function of pH at $I$ = 0.03 M, with the storage modulus on a logarithmic (**A**) and linear (**B**) scale.

**Proposed mechanism of peptide aggregation**

The gels formed from hydrolysates were not built up from intact soy proteins, since at DH values of 2.2 % and higher, almost no intact soy proteins were present (Figure 4). This is what might be expected when 2.2 % of the peptide bonds are cleaved. In addition, subtilisin Carlsberg is a relatively nonspecific protease known for its high specificity for Leu, Tyr, Phe and Trp and also cleaves next to Gln, Ser, and Tyr *(17, 29)*, so a broad spectrum of peptides can be expected. Some of these peptides are held together by disulfide bridges, but this effect seems to be minor. Part of the peptides form soluble aggregates during hydrolysis below a DH of 9.0 %. These aggregates increase in diameter with increasing DH. Around DH = 9.0 % insoluble aggregates are already formed during hydrolysis (Table 3). Upon acidification the aggregates formed during hydrolysis might aggregate further with peptides or other aggregates. Since the shift in $pH_{Aggr\text{-}onset}$ increases with increasing DH this shift is probably not the result of the release of some very specific peptides that accumulate with increasing DH.

Possible explanations for the shift in aggregation pH upon hydrolysis have been previously suggested in studies concerning the hydrolysis of whey proteins with Alcalase *(30)* or BLP *(31)*. Part of the peptides formed were observed to have a higher p*I* *(31, 32)*, when compared to the parental protein. A change in p*I* changes the balance between electrostatic and hydrophobic forces, and can thus result in aggregate formation around the p*I* of the peptides.

The similarity between these observations with whey proteins and the observations made for soy proteins as presented here, suggests a similar mechanism.

In general, the balance between electrostatic repulsion and hydrophobic attraction determines whether a protein forms aggregates or not. If this is the dominant mechanism in our gels, then at higher pH values aggregates are expected to be formed with increasing ionic strength, due to screening of the repulsive electrostatic forces, enabling the hydrophobic forces to dominate. In Figure 3, however, no large influence of the ionic strength can be observed on the $pH_{Aggr\text{-}onset}$. This indicates that the observed aggregation is not simply a balance of repulsive electrostatic and attractive hydrophobic interactions, but is more complex. Possibly also specific electrostatic attractions play a role, making the aggregation mechanism much more delicate and less predictable. Moreover, effects on aggregation kinetics may not become apparent within the timescale of the experiment. When a mixture of peptides is present with a wide variety of p*I*'s, then at a certain pH, the peptides present will have a net charge that is either positive, negative or zero. This might lead to aggregation due to hydrophobic attraction of peptides with a net charge of zero, while electrostatic attraction takes place between oppositely charged peptides. In addition to these interactions, also peptides with a high polarity or hydrophobicity may be formed that will or will not form aggregates, independent of the pH.

## ACKNOWLEDGEMENT

We thank Saskia de Jong and Jan Klok from NIZO food research for assisting with the rheology and CSLM experiments, respectively. The soybeans were kindly provided by Fa. L.I. Frank (Twello, The Netherlands).

## REFERENCES

1. Utsumi, S.; Kinsella, J.E. Forces involved in soy protein gelation - Effects of various reagents on the formation, hardness and solubility of heat-induced gels made from 7S, 11S, and soy isolate. *J. Food Sci.* **1985,** *50*, (5), 1278-1282.
2. Renkema, J.M.S.; Gruppen, H.; van Vliet, T. Influence of pH and ionic strength on heat-induced formation and rheological properties of soy protein gels in relation to denaturation and their protein compositions. *J. Agric. Food Chem.* **2002,** *50*, (21), 6064-6071.
3. Renkema, J.M.S.; van Vliet, T. Heat-induced gel formation by soy proteins at neutral pH. *J. Agric. Food Chem.* **2002,** *50*, (6), 1569-1573.
4. Bryant, C.M.; McClements, D.J. Molecular basis of protein functionality with special consideration of cold-set gels derived from heat-denatured whey. *Trends Food Sci. Technol.* **1998,** *9*, (4), 143-151.
5. Kohyama, K.; Nishinari, K. The effect of glucono-δ-lactone on the gelation time of soybean 11S protein: Concentration dependence. *Food Hydrocolloids* **1992,** *6*, (3), 263-274.
6. Damodaran, S. Amino acids, peptides, and proteins. In *Food Chem.*, 3rd ed.; Fennema, O. R., Ed. Marcel Dekker, Inc.: New York, 1996; pp 322-429.
7. Doi, E. Gels and gelling of globular proteins. *Trends Food Sci. Technol.* **1993,** *4*, 1-5.

8. Kruif de, C.G. Skim milk acidification. *J. Colloid Interface Sci.* **1997,** *185*, (1), 19-25.
9. Panyam, D.; Kilara, A. Enhancing the functionality of food proteins by enzymatic modification. *Trends Food Sci. Technol.* **1996,** *7*, (4), 120-125.
10. Nielsen, P.M. Functionality of protein hydrolysates. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Paraf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 443-472.
11. Ju, Z.Y.; Otte, J.; Madsen, J.S.; Qvist, B. Effects of limited proteolysis on gelation and gel properties of whey protein isolate. *J. Dairy Sci.* **1995,** *78*, 2119-2128.
12. Fuke, Y.; Sekiguchi, M.; Msuoka, H. Nature of stem bromelain treatments on the aggregation and gelation of soybean proteins. *J. Food Sci.* **1985,** *50*, 1283-1288.
13. Doucet, D.; Gauthier, S.F.; Foegeding, E.A. Rheological characterization of a gel formed during extensive enzymatic hydrolysis. *J. Food Sci.* **2001,** *66*, 711-715.
14. Otte, J.; Ju, Z.Y.; Faergemand, M.; Lomholt, S.B.; Qvist, K.B. Protease-induced aggregation and gelation of whey proteins. *J. Food Sci.* **1996,** *61*, 911-923.
15. Murata, K.; Kusakabe, I.; Kobayashi, H.; Kiuchi, H.; Murakami, K. Selection of commercial enzymes suitable for making soymilk-curd. *Agric. Biol. Chem.* **1987,** *51*, (11), 2929-2933.
16. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.
17. Ottesen, M.; Svendsen, A. The subtilisins. *Methods Enzymol.* **1970,** *19*, 199-215.
18. Hermansson, A.-M. Physico-chemical aspects of soy proteins structure formation. *J. Texture Stud.* **1978,** *9*, 33-58.
19. Adler-Nissen, J. *Enzymic Hydrolysis of Food Proteins*. 1st ed.; Elsevier Applied Science Publishers: London, England, 1986.
20. Beynon, R.; Bond, J.S. *Proteolytic Enzymes*. 2nd ed.; Oxford University Press: Oxford, England, 2001.
21. James, G.T. Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Anal. Biochem.* **1978,** *86*, 574-579.
22. Alting, A.C.; de Jongh, H.H.J.; Visschers, R.W.; Simons, J. Physical and chemical interactions in cold gelation of food proteins. *J. Agric. Food Chem.* **2002,** *50*, (16), 4682-4689.
23. Hoffmann, M.A.M.; Roefs, S.; Verheul, M.; VanMil, P.; DeKruif, K.G. Aggregation of β-lactoglobulin studied by *in situ* light scattering. *J. Dairy Res.* **1996,** *63*, (3), 423-440.
24. Walstra, P. Soft Solids. In *Physical Chemistry of Foods*, 1st ed.; Walstra, P., Ed. Marcel Dekker, Inc.: New York, 2003; pp 683-771.
25. Alting, A.C.; Weijers, M.; De Hoog, E.H.A.; van de Pijpekamp, A.M.; Stuart, M.A.C.; Hamer, R.J.; De Kruif, C.G.; Visschers, R.W. Acid-induced cold gelation of globular proteins: Effects of protein aggregate characteristics and disulfide bonding on rheological properties. *J. Agric. Food Chem.* **2004,** *52*, (3), 623-631.
26. Ju, Z.Y.; Kilara, A. Gelation of hydrolysates of a whey protein isolate induced by heat, protease, salts and acid. *Int. Dairy J.* **1998,** *8*, (4), 303-309.
27. Utsumi, S.; Matsumura, Y.; Mori, T. Structure-function relationships of soy proteins. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Praf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 257-291.
28. Ipsen, R.; Otte, J.; Lomholt, S.B.; Qvist, K.B. Standardized reaction times used to describe the mechanism of enzyme-induced gelation in whey protein systems. *J. Dairy Res.* **2000,** *67*, (3), 403-413.

29. Johansen, J.T.; Ottesen, M.; Svendsen, I.; Wybrandt, G. Degradation of B-chain of oxidized insulin by 2 subtilisins and their succinylated and N-carbamylated derivatives. *C. R. Trav. Lab. Carlsberg* **1968,** *36*, (20), 365-384.

30. Doucet, D.; Otter, D.E.; Gauthier, S.F.; Foegeding, E.A. Enzyme-induced gelation of extensively hydrolyzed whey proteins by Alcalase: Peptide identification and determination of enzyme specificity. *J. Agric. Food Chem.* **2003,** *51*, (21), 6300-6308.

31. Otte, J.; Lomholt, S.B.; Ipsen, R.; Stapelfeldt, H.; Bukrinsky, J.T.; Qvist, K.B. Aggregate formation during hydrolysis of β-lactoglobulin with a Glu and Asp specific protease from *Bacillus licheniformis*. *J. Agric. Food Chem.* **1997,** *45*, 4889-4896.

32. Doucet, D.; Gauthier, S.F.; Otter, D.E.; Foegeding, E.A. Enzyme-induced gelation of extensively hydrolyzed whey proteins by Alcalase: Comparison with the plastein reaction and characterization of interactions. *J. Agric. Food Chem.* **2003,** *51*, (20), 6036-6042.

# CHAPTER 3

# Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates

**ABSTRACT**

The aggregation behavior as a function of pH was studied for hydrolysates obtained by hydrolysis of soy protein isolate (SPI) and glycinin- and β-conglycinin-rich protein fractions with subtilisin Carlsberg. The substrates were hydrolyzed up to degrees of hydrolysis (DH) of 2.2 and 6.5 %. Compared to non-hydrolyzed SPI, a decrease in solubility was observed for the hydrolysates of SPI (0.8 (w/w) % protein, $I$ = 0.03 M) around neutral pH. At pH 8.0, glycinin hydrolysates had a much lower solubility (~ 43 and 60 %, respectively for DH = 2.2 and 6.5 %) than SPI and β-conglycinin-derived hydrolysates, which were almost completely soluble. Peptides that aggregated were all larger than 5 kDa, and as estimated by size-exclusion chromatography their composition was almost independent of the aggregation pH. The solubility of hydrolysates of SPIs with a varying glycinin and β-conglycinin composition showed that glycinin-derived peptides are the driving force for the lower solubility of SPI hydrolysates. The solubility of SPI hydrolysates at pH 8.0 was shown not to be the sum of that of glycinin and β-conglycinin hydrolysates. Assuming that the separate hydrolysis of glycinin and β-conglycinin did not differ from that in the mixture (SPI), this indicates that β-conglycinin-derived peptides have the ability to inhibit glycinin-derived peptide aggregation.

Based on: Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. *Food Biophys*. **2006**, 1, 178-188.

## INTRODUCTION

Soy protein aggregation followed by the formation of a space-filling gel network plays a major role in the structure formation of various soy-based foods, such as meat analogues, tofu or fermented soy milk *(1)*. In addition to these existing food structures, new structures would open possibilities to expand or to improve the use of soy proteins in foods. A possible way to obtain such new structures may be via the modification of the aggregation properties of proteins. Such modifications may be induced physically, e.g. via heat treatment, or chemically, for example by using proteolytic enzymes *(2)*.

The improvement of soy protein aggregation and the subsequent change in soy protein gelation by pre-treatment with proteases, has not yet received much attention in the literature, although hydrolysis can be a means of expanding the gelling properties of whey proteins *(3, 4)* and soy proteins *(5, 6)*.

The probable reason for this lack of attention is the fact that hydrolysis is assumed to be unfavorable for the gelling properties of proteins, since hydrolysis increases the number of charged groups and reduces the molecular mass, both of which generally hamper gelation *(7)*. On the other hand, during hydrolysis the folded structure of the protein is altered, resulting in the exposure of buried hydrophobic groups, which might induce aggregate formation.

Interestingly, upon hydrolysis of proteins, aggregates may be formed *(8, 9)*, indicating that peptides are formed with an aggregation behavior different from that of the non-hydrolyzed protein. Under specific conditions these aggregates can form gel networks, which have improved properties compared to the non-hydrolyzed protein *(4, 5, 10, 11)*. In contrast to the formation of aggregates, peptides can also prevent aggregation. Peptides derived from milk or soy proteins were shown to have the ability to inhibit blood platelet aggregation *(12, 13)*.

In a previous study *(6)* it was shown that hydrolysis of soy protein isolates results in the formation of peptides that aggregate at a pH value higher than the pH at which the non-hydrolyzed proteins aggregate. Understanding the characteristics and origin of these aggregating peptides will enable optimization of the hydrolysis process with respect to the yield in peptides having aggregating properties, which can be used as functional ingredients in foods.

The two main proteins present in soy are glycinin and β-conglycinin, which represent ~42 and ~34 % of all the proteins present in soybeans, respectively *(14)*. Glycinin is composed of an acidic (acidic p*I*; ~35 kDa) and a basic polypeptide (basic p*I*; ~20 kDa) connected by one disulfide bridge *(15)* and mainly exists as a hexamer (11S; pH 7.6, $I$ = 0.5 M). Glycinin can dissociate into its 7S and 3S form upon decreasing the ionic strength or changing the pH *(16)*. β-Conglycinin exists as a trimer with a molecular mass of 150-200 kDa, and contains three major subunits: α' (72 kDa), α (68 kDa), and β (52 kDa) *(15)*. Glycinin and β-conglycinin have different gelling properties: Upon heating, glycinin is able to form stronger gels than β-conglycinin *(17)*. In heat-induced gels made of mixtures of glycinin and β-conglycinin, the gel hardness of the mixed preparations was intermediate between gels of glycinin and β-conglycinin *(18, 19)*. Analogous to this, it can be hypothesized that peptides originating from

glycinin or β-conglycinin may also have a different contribution to the aggregation behavior of hydrolysates made from soy protein isolate (SPI). Therefore, the main purpose of the present study was to elucidate the individual contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of SPI hydrolysates.

## MATERIALS AND METHODS

### Materials and chemicals

Defatted soybean meal (SBM) was used, after being prepared from Hyland soybeans (non-GMO) supplied by Fa. L.I. Frank (Twello, The Netherlands) as described previously *(6)*. Subtilisin Carlsberg from *Bacillus licheniformis* (10 units/mg solid: one unit releases color, equivalent to 1.0 μmole (181 μg) of tyrosine per min from casein at pH 7.5 at 37°C as measured using the Folin-Ciocalteu reagent) was obtained from Sigma Chemical Co. (St. Louis, MO, USA; article no. P-5380). All other chemicals were of analytical grade and were purchased from Merck (Darmstadt, Germany) or Sigma (St. Louis, MO, USA).

### Analysis of the protein content

The nitrogen content of various samples was determined in duplicate by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milano, Italy) according to the manufacturer's instructions. Methionine was used as a standard. The nitrogen conversion factors for SPIs, glycinin and β-conglycinin were 5.58, 5.57 and 5.60, respectively. The values were based on their average amino acid compositions as found in Swiss-Prot (www.expasy.org). SPI is regarded to only containing glycinin and β-conglycinin in a 64/36 (w/w) ratio as determined by size-exclusion chromatography. Primary accession numbers used were: P04776 (Glycinin 1; A1a and B2 subunits), P04405 (Glycinin 2; A2 and B1a subunits), P11828 (Glycinin 3; A1b and B1b subunits), P02858 (Glycinin 4: A5, A4 and B3 subunits), P04347 (Glycinin 5; A3 and B4 subunits)), P11827(β-Conglycinin, α'-chain), P13916 (β-Conglycinin, α-chain) and P25974 (β-Conglycinin, β-chain)).
Since β-conglycinin is a glycoprotein and is reported to contain ~4.5% carbohydrates *(14, 20)*, the nitrogen conversion factor can be calculated to be ~5.87. However, this value would be appropriate when the β-conglycinin content is determined. During proteolysis only the polypeptide chain is cleaved and all the sample preparations are standardized at the same polypeptide content. As a consequence, for β-conglycinin a nitrogen conversion factor of 5.60 (based on the amino acid sequence) was used.

### Preparation of soy protein isolates (SPIs)

From the defatted SBM, an SPI was prepared by precipitation of the proteins at pH 4.8 using 1 M HCl as described previously *(6)*. Densitometric analysis of the SDS-PAGE gel showed that at least 90 % of the protein bands present in the SPI originate from glycinin or β-

conglycinin (Figure 4 (1)). This SPI is further denoted as SPI 64/36, in which 64/36 (w/w) is the ratio of glycinin/β-conglycinin in the preparation (Table 1), as determined by high-performance size-exclusion chromatography (HP-SEC). In the preparation of the SPI, the clear solution at pH 8.0, obtained just after centrifugation and just before the pH decrease to pH 4.8, is denoted as pH 8.0 protein extract.

In addition to the SPI 64/36, SPIs were also prepared by precipitation at other pH intervals, which resulted in SPIs with different glycinin/β-conglycinin ratios by making use of the difference in the pH-dependent solubility of the two proteins *(21)*. Two sequential precipitations were performed, both starting with the pH 8.0 protein extract from SBM.

The first sequential precipitation resulted in three SPIs precipitated in the pH intervals 8.0-6.2, 6.2-5.0, and 5.0-4.8. First, a precipitate was prepared by decreasing the pH to 6.2. The pH of the supernatant obtained at pH 6.2 was subsequently decreased to pH 5.0 to precipitate another part of the proteins. Next, the pH of the supernatant obtained at pH 5.0 was set to 4.8 followed by centrifugation. This first sequential precipitation resulted in three SPIs: SPI 92/8, SPI 49/51, and SPI 22/78.

The second sequential precipitation resulted in two SPIs precipitated in the pH intervals 8.0-5.6 and 5.6-4.8. First, a precipitate was prepared at pH 5.6. The pH of the supernatant obtained at pH 5.6 was decreased to pH 4.8 followed by centrifugation. This second sequential precipitation resulted in two SPIs: SPI 74/21 and SPI 35/65. The compositions of the five SPIs obtained can be found in Table 1.

In the two sequential precipitations, the pH was decreased by addition of 1 M HCl. All suspensions were stirred for at least 1 h before centrifugation (15 min, 12,000 x *g*, 20ºC). All the pellets, obtained at different pH values (6.2, 5.6, 5.0 and 4.8), were washed once with Millipore water ((v/v) ratio precipitate : water ~ 1 : 9) followed by suspension in water and adjustment of the pH to 8.0 before they were freeze-dried.

**Purification of glycinin**

Glycinin was purified in a similar way as the SPI 92/8 preparation with the difference that the precipitate obtained at pH 6.2 was directly suspended in water at pH 8.0 overnight at 4°C. After centrifugation (30 min, 12,000 x *g*, 20ºC) the supernatant was dialyzed against distilled water. The pH of the dialyzed protein dispersion was adjusted to pH 8.0 and the clear solution obtained was freeze-dried. The freeze-dried product was denoted “Glycinin” (Table 1).

**Purification of β-conglycinin**

To obtain purified β-conglycinin, the supernatant from the glycinin purification (pH 6.2) was adjusted to pH 5.0, stirred for 1 h at ambient temperature, and centrifuged (30 min, 18,000 x *g*, 20ºC). The pellet was discarded and the pH of the supernatant was adjusted to 4.8 to induce precipitation of β-conglycinin. The suspension was stirred for 1 h, followed by centrifugation (30 min, 12,000 x *g*, 20ºC). The precipitate was suspended overnight at 4°C in 35 mM sodium

phosphate buffer (pH 8.0). After centrifugation (30 min, 18,000 x *g*, 20ºC) the supernatant was applied onto a Superdex 200 column (60 x 10 cm, Amersham, Uppsala, Sweden) and eluted with a 35 mM sodium phosphate buffer (pH 8.0) at a flow rate of 40 mL/min at ambient temperature. The eluate was monitored by UV detection at 220 nm and 200 mL fractions were collected. Fractions containing β-conglycinin, as analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; molecular mass bands above 43 kDa), were pooled and dialyzed against distilled water. The pH of the dialyzed protein dispersion was adjusted to pH 8.0 and the clear solution obtained was freeze-dried. The freeze-dried product was denoted "β-Conglycinin" (Table 1).

**Hydrolysis of SPIs, Glycinin and β-Conglycinin**

The SPIs, Glycinin, and β-Conglycinin were suspended in Millipore water at a concentration of 1.5 - 2 % (w/w) protein and stirred overnight at 4°C. The pH was adjusted to 8.0, if necessary. SPI solutions were centrifuged (18,000 x *g*, 20 min, 20°C), while the Glycinin and β-Conglycinin solutions were filtered using a 0.45 μm filter (Schleicher & Schuell GmbH, Dassel, Germany). The protein content of the supernatant and filtrates was determined using the Dumas method. Solutions were diluted to 1 % (w/w) protein using Millipore water followed by heating at 95°C for 30 minutes. The heated solutions all had a clear to slightly opaque appearance. Hydrolysates with degrees of hydrolysis (DH) of 2.2 and 6.5 % were prepared from the heated SPI, Glycinin, and β-Conglycinin solutions by hydrolysis at pH 8.0 at 40ºC, using subtilisin Carlsberg. The pH and DH were controlled using the pH-stat method by using a 719S Titrino (Metrohm ion analysis, Herisau, Switzerland) *(22)*. The $h_{tot}$ used for the calculation of the DH was calculated based on the amino acid composition of the various preparations, assuming that only glycinin ($h_{tot}$ = 8.87 meqv/g) and β-conglycinin ($h_{tot}$ = 8.60 meqv/g) were present (Table 1). The enzyme concentrations used were ~3 or 13 units/gram protein for the DH values of 2.2 and 6.5 %, respectively. The enzyme was dissolved in Millipore water and directly added to the solutions. The molarity of the NaOH solution used to maintain the pH at pH 8.0 varied from 0.1 to 0.3 M. When the desired DH was reached, the enzymatic hydrolysis was stopped by addition of a 100 mM phenylmethylsulphonyl fluoride (PMSF) stock solution in 2-propanol to a final concentration of 1 mM. The pH-stat experiment was stopped after the pH remained stable at pH 8.0. Hydrolysates were freshly prepared prior to each experiment.

**Gel electrophoresis**

The protein composition of SPIs, Glycinin and β-Conglycinin and their hydrolysates was examined using SDS-PAGE under reducing conditions (10 mM 2-mercaptoethanol). A 10-20 % Tris/HCl Ready Gel (Bio-Rad Laboratories Hercules, CA, USA) was used on a Mini-protean II System (Bio-Rad Laboratories, USA) according to the instructions of the manufacturer. The gel was calibrated with marker proteins with molecular masses ranging

from 14.4 to 94 kDa (Amersham Biosciences, Roosendaal, The Netherlands) and was stained with Coomassie Brilliant Blue.

**Solubility of the SPI, Glycinin and β-Conglycinin hydrolysates**

The aggregation behavior of the hydrolysates and non-hydrolyzed solutions was studied as the (in-)solubility at a protein concentration of 0.8 % (w/w) ($I$ = 0.03 M). For the solubility of the non-hydrolyzed samples, the heated solutions, as also used for the hydrolysis experiments, were used.

The ionic strength of the heated protein solution was arbitrary regarded as 0 M. The ionic strength of 0.03 M was defined by the total amount of NaOH added during the hydrolysis and NaCl added after hydrolysis. During hydrolysis the ionic strength of the solution increases, due to the liberation of charged groups. The increase in ionic strength during hydrolysis is assumed to be equal to the amount of NaOH added. This results in an increase of 0.003 and 0.009 M for hydrolysates with a DH of 2.2 and 6.5 %, respectively. To end up with a final protein concentration of 0.8 % (w/w) and an ionic strength of 0.03 M Millipore water and 0.5 M NaCl were added. The solutions were acidified stepwise from pH 8.0 to pH 3.0 using 0.2 M HCl. A sample was taken every 0.5 pH units. The samples, set at various pH values, were mixed in a head-over-tail rotator for 1 h followed by centrifugation (20 min, 22,000 x g, 20°C). The proportion of the total amount of peptides present that remained in the supernatant after centrifugation was defined as solubility. The solubility was determined by diluting the supernatant 20 times with 1 % (w/v) SDS in Millipore water, followed by determination of the absorbance at 280 nm ($A_{280}$), using a spectrophotometer (UV 1601, Shimadzu, Japan). A 0.95 % (w/v) SDS solution, served as a blank. In addition, all the (20 times) diluted samples were diluted 10 times with 1% SDS, followed by analysis of their absorbances at 220 nm ($A_{220}$). The measured absorbances were corrected for dilutions during acidification with HCl.

The solubility was calculated by expressing the $A_{220}$ or $A_{280}$ of the supernatant as proportion of the $A_{220}$ or $A_{280}$ of the total hydrolysate at pH 8.0. For SPI 64/36 and Glycinin the solubility curves were prepared in duplicate and were observed to have an average standard deviation < 5 % of the mean for each data point when obtained at pH 4.0 and higher, while at pH 3.0 and 3.5 standard deviations up to 20 % were obtained.

The pellets, containing the aggregating peptides, were washed once with Millipore water and again centrifuged (20 min, 22,000 x $g$, 20°C). The pellets were re-suspended in 0.4 % (v/v) trifluoroacetic acid (TFA) in Millipore water followed by freeze-drying. TFA was added to cause a pH decrease sufficient to irreversibly inhibit the remaining minor activity of subtilisin Carlsberg. In this study we define aggregating peptides as those peptides present in the pellets after centrifugation, although we are fully aware of the possibility that soluble aggregates, composed of peptides, remain present in the supernatant.

### High-performance size-exclusion chromatography (HP-SEC)

SPI 64/36, Glycinin, β-Conglycinin, and their hydrolysates and pellets were dissolved in 500 µL of 0.15 M Tris-HCl buffer (pH 8.0) containing 100 mM 1,4-dithiothreitol (DTT) and 8 M guanidinium chloride. The final protein concentration was ~3 mg/mL. After mixing in a head-over-tail rotator for 1 h at ambient temperature, 215 µL of acetonitrile, containing 2 % (v/v) TFA, was added and the mixture was mixed in a head-over-tail rotator for another 1 h. The solutions had a clear appearance, but were centrifuged as a precaution. After centrifugation (15 min, 22,000 x *g*, 20ºC) 20 µL of the supernatants were applied to a Shodex Protein KW-803 column (300 x 8 mm; Showa Denko K.K., Tokyo, Japan). The column was equilibrated and run with 6 M urea containing 30 % (v/v) acetonitrile and 0.1 % (v/v) TFA. The flow rate was 0.5 mL/min and the absorbance of the eluate was monitored at 220 nm. The column was calibrated using various proteins in a molecular mass range of 0.308 to 669 kDa (glutathione, bradykinin, angiotensin, insulin chain B, aprotinin, ribonuclease A, β-lactoglobulin, chymotrypsin, aldolase, ovalbumin, albumin, catalase, ferritin, and thyroglobulin). The void volume of the column was ~ 5.8 mL, as determined from the volume at which thyroglobulin (669 kDa) eluted. The included volume was estimated using glutathione (308 Da) and was observed to be 12.7 mL.

To enable comparison between the size distribution of the non-hydrolyzed proteins with those of hydrolysates and the aggregating peptides, the total area under the chromatograms was normalized. In order to facilitate comparison, the total area under the chromatograms of the aggregated peptides was corrected for the proportion of peptides that was aggregated.

To determine the glycinin/β-conglycinin ratios in the SPIs, Glycinin and β-Conglycinin samples the area under the peaks (corrected for the difference in absorbance between glycinin and β-conglycinin at 220 nm) in the SEC chromatograms at 220 nm was used. On a weight basis, the absorbance of glycinin appeared to be a factor 1.25 higher than the absorbance of β-conglycinin. Proteins eluting between 6.1-7.0 ml and 7.0-10 ml were considered to be β-conglycinin and glycinin, respectively, as described previously *(6)*. Due to the high solubility of trypsin inhibitors over a broad pH range *(23)*, their presence in the SPIs, Glycinin and β-Conglycinin samples is unlikely and is, therefore, neglected.

## RESULTS

### Differences in solubility of the Glycinin, β-Conglycinin and SPI 64/36 hydrolysates

To elucidate the individual contributions of glycinin- and β-conglycinin-derived peptides in the SPI hydrolysates, first the aggregation behavior of hydrolysates from Glycinin and β-Conglycinin was compared with the aggregation behavior of hydrolysates from SPI 64/36.

The most remarkable observation in this study was that the Glycinin hydrolysates had much lower solubilities in the pH range from 7.0 to 8.0 than non-hydrolyzed Glycinin, as indicated by the arrow in Figure 1A. Before ionic strength adjustment at pH 8.0, the Glycinin hydrolysate at DH 2.2 % already showed a white haze. It became even more turbid upon ionic

strength adjustment to 0.03 M. The Glycinin hydrolysate at DH 6.5 % was already turbid when the desired DH was reached, and no extra aggregation could be visually observed upon adjustment of the ionic strength. After stepwise acidification of the Glycinin hydrolysate, the solubility of the DH = 2.2 % sample slightly decreased until pH 4.5, while it increased upon further acidification. With decreasing pH, the Glycinin hydrolysate at DH = 6.5 % showed almost no further decrease in solubility until pH 4.5. Below this pH the solubility increased slightly during acidification to pH 3.0. In the pH range from pH 4.0 to 8.0, the solubility at DH = 6.5 % is higher that at DH = 2.2 %.

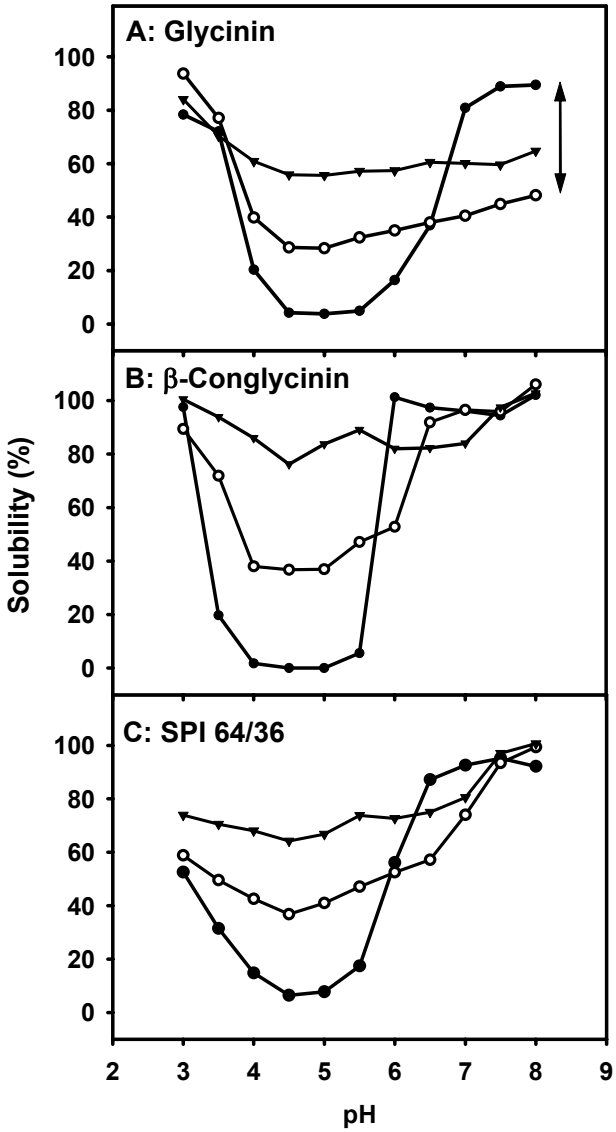


**Figure 1:** Solubility curves for DH = 0 (●), 2.2 (○), and 6.5 (▼) % for Glycinin (A), β-Conglycinin (B), and SPI 64/36 (C) at 0.8 % (w/w) protein concentration ($I$ = 0.03 M). The solubility was determined from the absorbance at 220 nm. The arrow indicates the decrease in solubility of the Glycinin hydrolysates compared to the non-hydrolyzed Glycinin.

In contrast to the Glycinin hydrolysates, the β-Conglycinin hydrolysates had a clear appearance, before and after ionic strength adjustment at pH 8.0, which is reflected in the ~100 % solubility at this pH (Figure 1B). Upon decreasing the pH of the hydrolyzed β-Conglycinin the pH at which the solubility started to decrease, increased with increasing DH.
Interestingly, Figure 1C shows that around pH 8.0 the solubility of the non-hydrolyzed SPI 64/36 and its hydrolysates was almost 100%. The overall solubility profile of SPI 64/36 and their hydrolysates as a function of pH was, therefore, similar to that obtained for β-Conglycinin and its hydrolysates, with the main difference that the pH at which the solubility started to decrease was now the same for both DHs studied.
All the hydrolysates had a minimum solubility in the pH 4 – 5 range. The solubility in this range increased with increasing DH, which agrees with the general observation that hydrolysis increases protein solubility *(24, 25)*. Upon further decrease of the pH to 3, the solubility increased with decreasing pH. Glycinin and β-Conglycinin and their hydrolysates reached a solubility of ~80 % or higher at pH 3.0, whereas the solubility of SPI 64/36 and its hydrolysates at pH 3.0 remained below 80%.

**Contribution of tryptophan and tyrosine to the composition of aggregating peptides**

The solubility as a function of pH for SPI 64/36, Glycinin and β-Conglycinin and their hydrolysates as calculated from the $A_{220}$ and $A_{280}$ absorbances are shown in Figure 2.

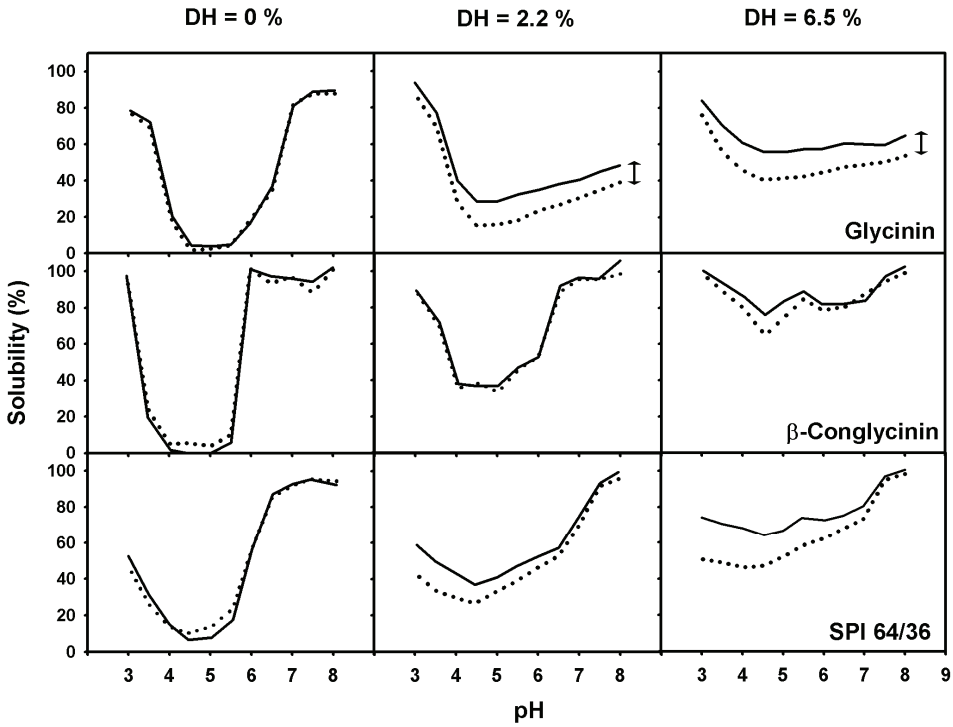


**Figure 2**: Solubility curves as a function of pH for Glycinin, β-Conglycinin, and SPI 64/36 at 0.8 % (w/w) protein (I = 0.03 M), with DH values of 0, 2.2 and 6.5 %. The solubility was determined from the absorbance at 220 nm ($A_{220}$; ——) and 280 nm ($A_{280}$; ·····). The arrows indicate the difference in solubility for Glycinin when measured at different wavelengths.

For both Glycinin hydrolysates the solubility measured as $A_{280}$ was lower than the solubility measured as $A_{220}$ over the whole pH range for the two hydrolysates, whereas for β-Conglycinin hydrolysates, as well as for all three non-hydrolyzed preparations, the solubility measured as $A_{280}$ was similar to the solubility measured as $A_{220}$. This indicates that, compared to the aggregating peptides from β-conglycinin, the aggregating peptides from glycinin seem to be enriched in the aromatic amino acids tryptophan and tyrosine that absorb light at 280 nm.

For the hydrolysates of SPI 64/36 the solubility measured as $A_{220}$ is the same as the solubility measured as $A_{280}$ at neutral pH, but the solubility measured as $A_{280}$ decreases further with decreasing pH than the solubility measured as $A_{220}$.

**Molecular size distribution of aggregating peptides**

The HP-SEC chromatograms of the non-hydrolyzed proteins, of their hydrolysates at DH 2.2 and 6.5 %, and of the peptides aggregating at pH 5.0 are shown in Figure 3. For all hydrolysates it can be observed that the aggregating peptides represent all the larger peptides present in the hydrolysates. The aggregating peptides at DH = 6.5 % are smaller compared to the size of aggregating peptides at DH = 2.2 %. Peptides with a molecular mass below 3 kDa do not precipitate at pH 5.0 (black arrows).

In Figure 3 only the chromatograms of the aggregating peptides at pH 5.0 are shown, since all pellets originating from one hydrolysate showed similar patterns at the various precipitation pH values (pH 3, 4, 5, 6, 7, and 8), and differed only in amount. The only exception to this were the peptide or peptides present in the peak eluting around 8.1 mL (~24 kDa; grey arrows), originating from β-conglycinin, which precipitated to a lower extent at the pH values tested above pH 5. Although we do not know whether this concerns one peptide, or various peptides with a similar mass, we denote peptides with an estimated mass of 24 kDa, as the 24 kDa peptide.

When comparing the chromatograms of the DH = 2.2 % hydrolysates with those of the DH = 6.5 % hydrolysates (Figure 3), it is remarkable to see that a higher proportion of the larger peptides seem to aggregate at a lower DH than at a higher DH. The 24 kDa peptide present in the β-Conglycinin hydrolyzate is the best example of this phenomenon, since it remained almost completely soluble at DH 6.5 %, whereas at DH = 2.2 %, it precipitated almost completely.

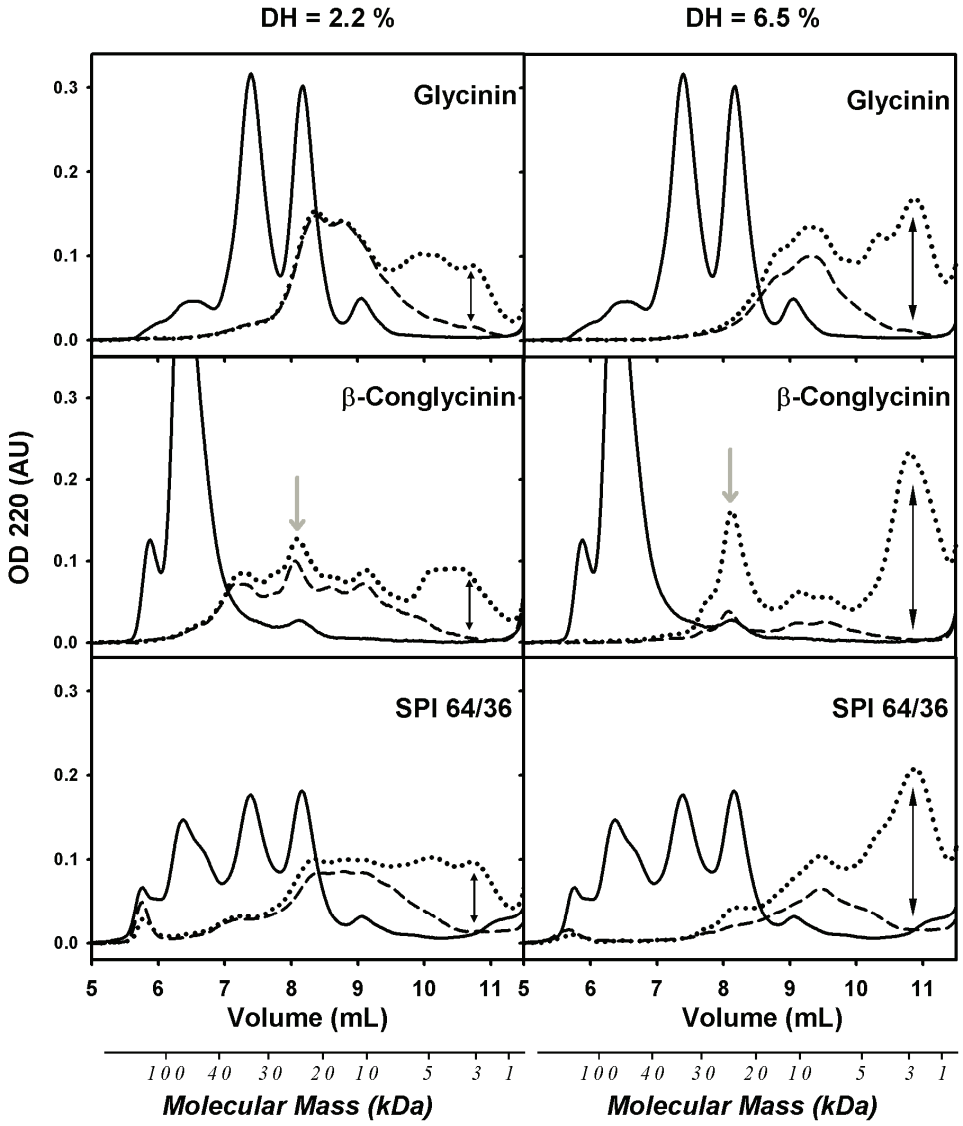


**Figure 3**: HP-SEC chromatograms under denaturing conditions of Glycinin, β-Conglycinin, and non-hydrolyzed SPI 64/36 (——), hydrolysates (• • • •) thereof at DH = 2.2 and 6.5 %, and aggregating peptides (------) of the hydrolysates at pH 5.0. The samples were reduced prior to analysis. The secondary *x*-axis shows the apparent molecular mass eluting at that retention time. The black and grey arrows indicate the high solubility of small peptides and the presence of the hydrolysis-resistant ~24 kDa peptide, respectively.

Figure 4 shows the SDS-PAGE gel of non-hydrolyzed SPI 64/36, Glycinin and β-Conglycinin and their hydrolysates. The absence of β-conglycinin bands in Glycinin (lane 4), and the absence of glycinin bands in β-Conglycinin (lane 7) indicate that Glycinin and β-Conglycinin had a high purity. It should be noted that in the β-Conglycinin the β-subunit (~48 kDa) *(15)* was present in a relatively low amount compared to SPI 64/36.

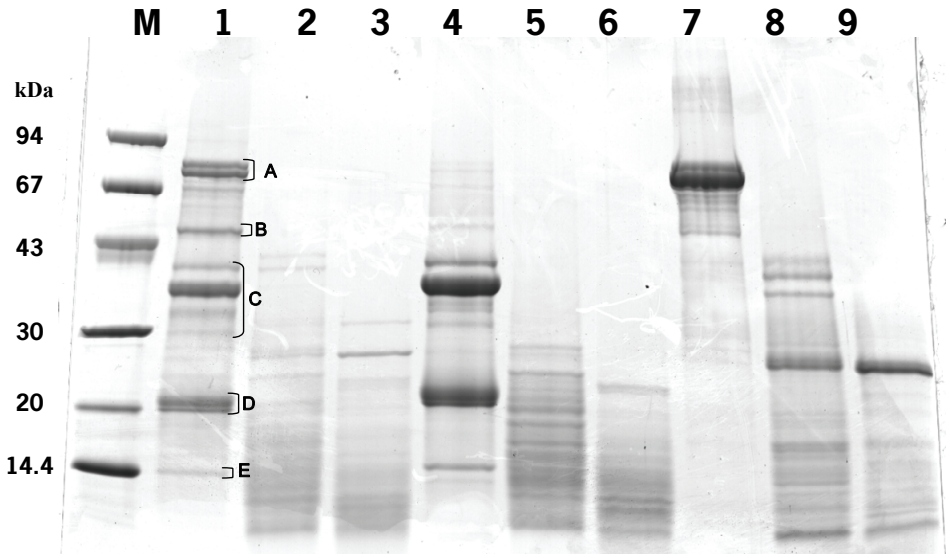


**Figure 4**: SDS-PAGE gel under reducing conditions showing the marker (M), SPI 64/36 at DH 0, 2.2 and 6.5 % (1-3), Glycinin at DH 0, 2.2 and 6.5 % (4-6), and β-Conglycinin at DH 0, 2.2 and 6.5 % (7-9), respectively. In the SPI lane (1), the α/α' subunits of β-Conglycinin (A), β subunit of β-Conglycinin (B), the acidic polypeptides of Glycinin (except A5) (C), the basic polypeptide of Glycinin (D) and the acidic polypeptide A 5 of Glycinin (E) are indicated.

The DH = 2.2 % samples show that at this DH non-hydrolyzed protein was no longer present. In the hydrolysates separate peptide bands were observed, corresponding to specific peptides that represent intermediate peptides in the degradation of the protein towards smaller peptides. The 24 kDa peptide that seems to be resistant to enzymatic hydrolysis, as shown in Figure 3, can also be observed in the SDS-PAGE gel and is present in both hydrolysates of SPI 64/36 and β-Conglycinin, although its intensity in SPI 64/36 is lower due to the lower amount of β-Conglycinin present in the SPI 64/36, when compared to the β-Conglycinin preparation itself.

## Aggregation behavior of peptides in hydrolysates of SPIs having different glycinin/β-conglycinin ratios

To determine how the individual aggregation behavior of glycinin and β-conglycinin peptides can be translated to "normal" SPI, SPIs with different glycinin/β-conglycinin ratios were prepared, as shown in Table 1. This table shows that an almost linear distribution of glycinin and β-conglycinin ratios was obtained over the six SPIs, which can also be observed from the SDS-PAGE gel as presented in Figure 5. This setup enabled us to control the protein composition, and use glycinin/β-conglycinin ratios that vary more than the variation present in cultivated soybeans *(26-28)*.

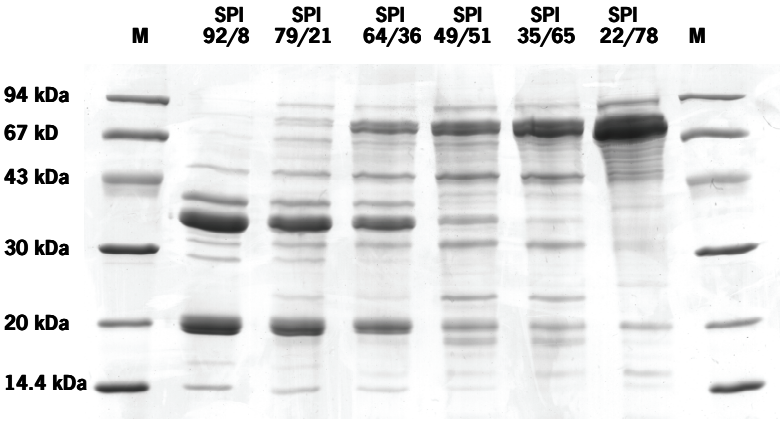


**Figure 5**: SDS-PAGE gel under reducing conditions showing the marker (M), SPI 92/8, SPI 79/21, SPI 64/36, SPI 49/51, SPI 35/65, and SPI 22/78.

**Table 1:** Glycinin and β-conglycinin content of Glycinin, SPIs and β-Conglycinin, the pH interval at which they were obtained, $h_{tot}$, and the protein content of the freeze-dried preparations.

| Sample | Precipitation pH interval | Glycinin (%) [a] | β-Conglycinin (%) [a] | $h_{tot}$ [b] (meqv/g) | Protein content (%)[c] |
|---|---|---|---|---|---|
| Glycinin | 6.2-8.0 | 92 | 8 | 8.85 | 90.4 (± 3.7) |
| SPI 92/8 | 6.2-8.0 | 92 | 8 | 8.85 | 85.9 (± 0.6) |
| SPI 79/21 | 5.6-8.0 | 79 | 21 | 8.81 | 81.3 (± 2.9) |
| SPI 64/36[d] | 4.8-8.0 | 64 | 36 | 8.77 | 83.3 (± 1.6) |
| SPI 49/51 | 5.0-6.2 | 49 | 51 | 8.73 | 76.2 (± 1.7) |
| SPI 35/65 | 4.8-5.6 | 35 | 65 | 8.70 | 74.2 (± 1.0) |
| SPI 22/78 | 4.8-5.0 | 22 | 78 | 8.66 | 78.1 (± 3.9) |
| β-Conglycinin | 4.8-5.0 | 12 | 88 | 8.63 | 80.3 (± 4.4) |

[a] As determined by HP-SEC (Weight %)
[b] $h_{tot}$: Total number of peptide bonds in the protein substrate (meqv/g protein) *(22)*.
[c] The protein content is expressed as (w/w) % and measured in duplicate.
[d] SPI 64/36 is same as used previously and can be considered as a reference SPI *(6)*.

The solubility curves as a function of pH obtained for hydrolysates of the various SPIs at DH = 2.2 and 6.5 % are shown in Figure 6. It can be clearly seen that the higher the glycinin content of the SPI, the higher the pH at which the solubility started to decrease (SPI 64/36, 49/51, 35/65 and 22/78) or the lower the solubility at pH 8.0 (SPI 92/8 and 79/21) as indicated with the arrow (Figure 6A). This phenomenon can also be observed at DH = 6.5 % (Figure 6B), but it is less clear than at DH = 2.2 %. All preparations displayed the same pH dependency around pH 4.5, followed by an increase in solubility upon further decrease of the pH.

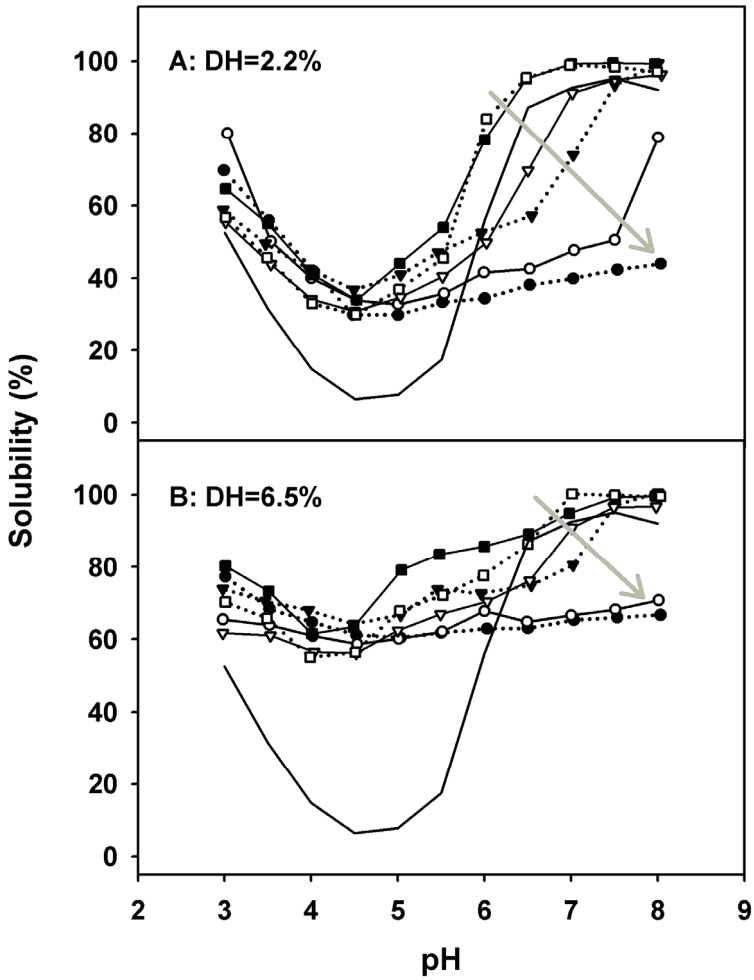


**Figure 6**: Protein solubility as a function of pH for non-hydrolyzed SPI 64/36 (——) , DH= 2.2 (A) and 6.5 % (B) for SPI 92/8(●•••●), SPI 79/21 (○——○), SPI 64/36 (▼•••▼), SPI 49/51 (∇——∇) SPI 35/65 (□•••□) and SPI 22/78 (■——■) at 0.8 % (w/w) protein (I = 0.03 M). The shift in aggregation pH is indicated with the arrows.

## DISCUSSION

### Differences in aggregation behavior between glycinin and β-conglycinin hydrolysates

The most remarkable result in this study was the difference in aggregation behavior of glycinin- and β-conglycinin-derived peptides. Around neutral pH, glycinin-derived peptides have a much stronger tendency to aggregate than β-conglycinin-derived peptides (Figure 1). This tendency of glycinin-derived peptides to form aggregates seems, therefore, to be the driving force in the previously observed shift in aggregation pH (i.e. the pH at which visible aggregation occurs) for hydrolysates of SPI when compared to non-hydrolyzed SPI *(6)*. This was confirmed by the solubility behavior of hydrolysates of the various SPIs (Figure 6), showing that the higher the glycinin content, the lower the solubility around neutral pH. When translating the behavior of hydrolysates of Glycinin and β-Conglycinin to SPI, it was

observed that, when hydrolyzed separately, the hydrolysis of β-Conglycinin is up to 15 % slower than the hydrolysis of Glycinin (results not shown); a difference that is not very large. When the SDS-PAGE patterns of the hydrolysates of SPI are compared with those of the hydrolysates of the Glycinin and β-Conglycinin (Figure 4), it can be observed that several peptides present in the SPI hydrolysate also seem to be present in the hydrolysates of Glycinin and β-Conglycinin. This may indicate that the two purified proteins were indeed hydrolyzed into the same fragments when hydrolyzed individually or in a mixture, such as in the SPI.
This difference in contribution to the total aggregation behavior between peptides derived from different proteins was also reported for the enzyme-induced gelation of a mixture of α-lactalbumin and β-lactoglobulin *(29)*. α-Lactalbumin-derived peptides were able to form much stronger gels than β-lactoglobulin-derived peptides. Moreover, the presence of β-lactoglobulin-derived peptides even inhibited gelation of α-lactalbumin.

**Characteristics of peptides responsible for aggregation**

Subtilisin Carlsberg is known to have a broad specificity. It strongly prefers to cleave next to Leu, Tyr, Phe and Trp and also cleaves next to Gln, Ser, and Tyr *(30, 31)*. Nevertheless, using SDS-PAGE clearly distinctive bands above 10 kDa can be seen in the hydrolysates. This indicates that the degradation of the proteins develops according to a specific pathway, resulting in a limited set of intermediate peptides that are degraded upon further hydrolysis. The only exception to this seems to be a 24 kDa peptide derived from β-conglycinin. This peptide is likely to be a stable end-product instead of an intermediate peptide and seems to accumulate with increasing DH (Figure 3 and 4), since the intensity increases with increasing DH, whereas the intensity of bands at higher molecular weight decreases. This stepwise degradation, via intermediate peptides, was also observed during hydrolysis of α-lactalbumin with a *Bacillus licheniformis* protease (BLP), in which α-lactalbumin was degraded into various intermediates of which one specific peptide proved to be resistant to further hydrolysis *(32)*.
The broad specificity of subtilisin Carlsberg would theoretically result in a huge number of peptides in the hydrolysates. Even though, as described above, there seems to be some degree of specificity, it remains difficult to speculate on the aggregation mechanism of glycinin peptides based on their amino acid composition. Nevertheless, amino acids containing aromatic rings that strongly absorb at 280 nm (tryptophan and tyrosine) seem to accumulate in the precipitate. This accumulation was mainly occurring for the aggregating peptides from glycinin, and almost absent in the precipitate of β-Conglycinin (Figure 2). Since tryptophan and tyrosine are relatively hydrophobic *(33)*, this may suggest that hydrophobic interactions play a dominant role in the aggregation of glycinin-derived peptides. This hypothesis can be confirmed by comparing the amino acid sequences of glycinin and β-conglycinin. If hydrophobic interactions play a dominant role in the aggregation of glycinin-derived peptides, then in the hydrophobic patches of glycinin a relatively high proportion of tryptophan and tyrosine should be present. Therefore, hydrophobic patches were defined in all sequences of

glycinin and β-conglycinin subunits, by calculating the hydrophobicity according to Eisenberg *(33)*. The calculation was performed using the ProtScale software from the Swiss Institute of Bioinformatics (www.expasy.org), using a "window size" of 15 (= peptide range over which the average hydrophobic score is calculated). Those regions in the amino acid sequence for which the hydrophobic score was above 0, for at least 10 amino acids in a row, were regarded as hydrophobic patches. In these patches, as well as in the total protein, the proportion of tryptophan and tyrosine (mole %) was calculated and is presented in Table 2.

**Table 2:** Proportion of tryptophan (Trp) and tyrosine (Tyr) in the proteins (mole %), average hydrophobicity and the p*I* of each polypeptide of glycinin and subunit of β-conglycinin, as well as the proportion of tryptophan and tyrosine present in the defined hydrophobic patches (mole %).

| | **Subunit** | **Poly-peptide** | **Average Hydrophobicity** [b] | **p*I*** | **Proportion of Trp & Tyr in Protein**[a] | **Proportion of Trp & Tyr in Patches**[a] |
|---|---|---|---|---|---|---|
| Glycinin | Gly 1 | *A1a* | -0.20 | 5.2 | 3.5 | 7.1 |
| | | *B2* | **0.03** | **6.3** | 2.8 | 3.3 |
| | Gly 2 | *A2* | -0.20 | 4.9 | 3.3 | 5.2 |
| | | *B1a* | **0.04** | **6.1** | 3.3 | 3.3 |
| | Gly 3 | *A1b* | -0.19 | 5.1 | 3.2 | 4.8 |
| | | *B1b* | **0** | **8.0** | 2.2 | 2.9 |
| | Gly 4 | *A5* | -0.03 | 5.5 | 3.1 | 6.6 |
| | | *A4* | -0.42 | 4.7 | 2.7 | 4.8 |
| | | *B3* | **0** | **9.3** | 5.9 | 8.9 |
| | Gly 5 | *A3* | -0.24 | 5.0 | 3.1 | 6.6 |
| | | *B4* | **-0.02** | **9.5** | 5.2 | 6.5 |
| β-Conglycinin | α | - | -0.24 | 4.9 | 2.6 | 2.8 |
| | α' | - | -0.24 | 5.6 | 2.6 | 3.0 |
| | β | - | -0.11 | 5.2 | 2.9 | 3.1 |

[a] mole %
[b] Average hydrophobicity of the polypeptides according to Eisenberg *(33)*

The proportion of tryptophan and tyrosine in the hydrophobic patches of all the acidic polypeptides of glycinin and the basic polypeptide of Glycinin 4 appeared to be more than 1.5 times higher than the average proportion of tryptophan and tyrosine present in the complete protein (as indicated in grey). This indicates that the differences between the solubilities measured as the $A_{220}$ and $A_{280}$ could be due to the accumulation of the hydrophobic patches of the acidic polypeptides of glycinin and the basic polypeptide of Glycinin 4 in the precipitates. For β-conglycinin and the basic polypeptides of Glycinin 1, 2, 3, and 5, the tryptophan and tyrosine are less accumulated in the hydrophobic patches, i.e. more evenly distributed over the sequence, resulting in a similar development of the solubility measured as $A_{220}$ and $A_{280}$.
Besides aggregation of the hydrophobic patches, also peptides originating from the basic polypeptides of Glycinin 1, 2, 3 and 5 were likely to be present in the precipitated glycinin-

derived peptide fractions, since the basic polypeptides have a higher average hydrophobicity and p*I* than the acidic polypeptides as indicated in bold in Table 2, which might promote aggregation of peptides originating form these basic polypeptides.
Although for β-conglycinin the proportion of tryptophan and tyrosine, as well as the average hydrophobicity and the p*I* are similar to those of the acidic polypeptide of glycinin, there are no indications that β-conglycinin-derived peptides show the same aggregation behavior at neutral pH.

**Inhibition of aggregation**

Glycinin hydrolysates have a solubility at pH 8.0 of ~40 and 60 % for DH = 2.2 and 6.5 %, respectively, while the SPI and β-Conglycinin hydrolysates are ~100 % soluble (Figure 1). This is remarkable because glycinin accounts for ~64 % of the total protein in SPI 64/36. Therefore, if it is assumed that glycinin, when present in SPI, is degraded into the same peptides as when hydrolyzed individually, a lower solubility of the SPI hydrolysate would be expected at pH 8.0. A possible explanation for this difference could be that in SPI peptides derived from β-conglycinin are present that inhibit aggregation. This can also be observed in Figure 3, in which it can be seen that with increasing DH, the total amount of precipitating peptides in the SPI hydrolysate decreases relatively more than the total amount of larger peptides in the hydrolysate. This is illustrated by the peak representing the 24 kDa peptide in β-Conglycinin. At DH 2.2 % the 24 kDa peptide is almost completely present in the pellet, whereas at DH = 6.5 % most of this peptide is present in the supernatant.
Therefore, it can be hypothesized that peptides are present that can inhibit or change aggregate formation in soy protein hydrolysates. This is in agreement with observations by Sodini and co-workers *(34)*, who observed that dairy yoghurts, supplemented with hydrolysates, showed a decreased viscosity with increasing supplementation. That peptides can change the properties of non-hydrolyzed proteins was also observed by Barbeau and co-workers *(35)*, who found that specific peptide fractions stabilized the structure of β-lactoglobulin, and hypothesized that these peptides might have prevented interactions between denatured β-lactoglobulin and casein. The phenomenon that peptides can inhibit aggregation was also noticed in studies of blood platelet aggregation. It was found that milk- and soy-derived “bioactive” peptides could inhibit blood platelet aggregation *(12, 13)*.

Our results indicate that in the case of soy proteins the peptides derived from β-conglycinin are predominantly responsible for the inhibition of peptide aggregation, whereas the peptides derived from glycinin have the opposite effect. The presence of peptides that have the ability to prevent aggregation may be interesting for the food industry, for use in products where aggregation of proteins is unwanted.

## ACKNOWLEGEMENTS

The soybeans were kindly provided by Fa. L.I. Frank (Twello, The Netherlands).

## REFERENCES

1. Liu, K. *Soybeans; Chemistry, Technology and Utilization*. 1st ed.; Chapman & Hall: New York, 1997.
2. Foegeding, E.A.; Davis, J.P.; Doucet, D.; McGuffey, M.K. Advances in modifying and understanding whey protein functionality. *Trends Food Sci. Technol.* **2002,** *13*, 151-159.
3. Ju, Z.Y.; Otte, J.; Zakora, M.; Qvist, K.B. Enzyme-induced gelation of whey proteins: Effect of protein denaturation. *Int. Dairy J.* **1996,** *7*, 71-78.
4. Doucet, D.; Gauthier, S.F.; Foegeding, E.A. Rheological characterization of a gel formed during extensive enzymatic hydrolysis. *J. Food Sci.* **2001,** *66*, 711-715.
5. Fuke, Y.; Sekiguchi, M.; Msuoka, H. Nature of stem bromelain treatments on the aggregation and gelation of soybean proteins. *J. Food Sci.* **1985,** *50*, 1283-1288.
6. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. *J. Agric. Food Chem.* **2005,** *53*, 1031-1038. *Chapter 2 in this thesis*.
7. Nielsen, P.M. Functionality of protein hydrolysates. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Paraf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 443-472.
8. Caessens, P.W.J.R.; Gruppen, H.; Visser, S.; van Aken, G.A.; Voragen, A.G.J. Plasmin hydrolysis of β-casein: Foaming and emulsifying properties of the fractionated hydrolysate. *J. Agric. Food Chem.* **1997,** *45*, 2935-2941.
9. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.
10. Ju, Z.Y.; Otte, J.; Madsen, J.S.; Qvist, B. Effects of limited proteolysis on gelation and gel properties of whey protein isolate. *J. Dairy Sci.* **1995,** *78*, 2119-2128.
11. Otte, J.; Ju, Z.Y.; Faergemand, M.; Lomholt, S.B.; Qvist, K.B. Protease-induced aggregation and gelation of whey proteins. *J. Food Sci.* **1996,** *61*, 911-923.
12. Lee, K.A.; Kim, S.H. SSGE and DEE, new peptides isolated from a soy protein hydrolysate that inhibit platelet aggregation. *Food Chem.* **2005,** *90*, 389-393.
13. Rutherfurd, K.J.; Gill, H.S. Peptides affecting coagulation. *Br. J. Nutr.* **2000,** *84*, S99-S102.
14. Koshiyama, I. Chemical and physical properties of a 7S protein in soybean globulins. *Cereal Chem.* **1968,** *45*, 394-404.
15. Utsumi, S.; Matsumura, Y.; Mori, T. Structure-function relationships of soy proteins. In *Food Proteins and their Applications*, 1st ed.; Damodaran, S.; Praf, A., Eds.; Marcel Dekker, Inc.: New York, 1997; pp 257-291.
16. Wolf, W.J.; Briggs, D.R. Studies on the cold-insoluble fraction of the water-extractable soybean proteins. 2. Factors influencing conformation changes in the 11S component. *Arch. Biochem. Biophys.* **1958,** *76*, 377-393.
17. Nakamura, T.; Utsumi, S.; Mori, T. Mechanism of heat-induced gelation and gel properties of soybean 7S-globulin. *Agric. Biol. Chem.* **1986,** *50*, 1287-1293.
18. Nakamura, T.; Utsumi, S.; Mori, T. Interactions during heat-induced gelation in a mixed system of soybean-7S and 11S-globulins. *Agric. Biol. Chem.* **1986,** *50*, 2429-2435.

19. Renkema, J.M.S.; Knabben, J.H.M.; van Vliet, T. Gel formation by β-conglycinin and glycinin and their mixtures. *Food Hydrocolloids* **2001,** *15*, 407-414.
20. Hou, H.J.; Chang, K.C. Structural characteristics of purified β-conglycinin from soybeans stored under four conditions. *J. Agric. Food Chem.* **2004,** *52*, (26), 7931-7937.
21. Thanh, V.H.; Shibasaki, K. Major proteins of soybean seeds. A straightforward fractionation and their characterization. *J. Agric. Food Chem.* **1976,** *24*, 1117-1121.
22. Adler-Nissen, J. *Enzymic Hydrolysis of Food Proteins*. 1st ed.; Elsevier Applied Science Publishers: London, England, 1986.
23. Sorgentini, D.A.; Wagner, J.R. Comparative study of structural characteristics and thermal behavior of whey and isolate soybean proteins. *J. Food Biochem.* **1999,** *23*, 489-507.
24. Wu, W.U.; Hettiarachchy, N.S.; Qi, M. Hydrophobicity, solubility, and emulsifying properties of soy protein peptides prepared by papain modification and ultrafiltration. *J. Am. Oil Chem. Soc.* **1998,** *75*, 845-850.
25. Adler-Nissen, J. Enzymatic hydrolysis of proteins for increased solubility. *J. Agric. Food Chem.* **1976,** *24*, (6), 1090-1093.
26. Fukuda, T.; Maruyama, N.; Kanazawa, A.; Abe, J.; Shimamoto, Y.; Hiemori, M.; Tsuji, H.; Tanisaka, T.; Utsumi, S. Molecular analysis and physicochemical properties of electrophoretic variants of wild soybean Glycine soja storage proteins. *J. Agric. Food Chem.* **2005,** *53*, 3658-3665.
27. Hughes, S.A.; Murphy, P.A. Varietal influence on the quantity of glycinin in soybeans. *J. Agric. Food Chem.* **1983,** *31*, 376-379.
28. Mori, T.; Utsumi, S.; Inaba, H.; Kitamura, K.; Harada, K. Differences in subunit composition of glycinin among soybean cultivars. *J. Agric. Food Chem.* **1981,** *29*, 20-23.
29. Ipsen, R.; Bulow-Olsen, K.; Otte, J.; Qvist, K.B. Protease induced gelation of mixtures of α-lactalbumin and β-lactoglobulin. *Milchwissenschaft* **2001,** *56*, 492-495.
30. Johansen, J.T.; Ottesen, M.; Svendsen, I.; Wybrandt, G. Degradation of B-chain of oxidized insulin by 2 subtilisins and their succinylated and N-carbamylated derivatives. *C. R. Trav. Lab. Carlsberg* **1968,** *36*, (20), 365-384.
31. Ottesen, M.; Svendsen, A. The subtilisins. *Methods Enzymol.* **1970,** *19*, 199-215.
32. Otte, J.; Ipsen, R.; Bauer, R.; Bjerrum, M.J.; Waninge, R. Formation of amyloid-like fibrils upon limited proteolysis of bovine alpha-lactalbumin. *Int. Dairy J.* **2005,** *15*, 219-229.
33. Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984,** *179*, 125-142.
34. Sodini, I.; Lucas, A.; Tissier, J.P.; Corrieu, G. Physical properties and microstructure of yoghurts supplemented with milk protein hydrolysates. *Int. Dairy J.* **2005,** *15*, 29-35.
35. Barbeau, J.; Gauthier, S.F.; Pouliot, Y. Thermal stabilization of beta-lactoglobulin by whey peptide fractions. *J. Agric. Food Chem.* **1996,** *44*, 3939-3945.

# CHAPTER 4

# Soy glycinin as a dominant source of aggregating peptides using different proteases

**ABSTRACT**

This study was carried out to investigate whether peptides obtained after limited proteolytic hydrolysis of soy glycinin have the generic property to aggregate at neutral pH, and to what extent this depends on the protease used. Soy-derived proteins (soy protein isolate, glycinin, and β-conglycinin) and, as a reference, bovine whey-derived proteins (whey protein isolate, α-lactalbumin, β-lactoglobulin) were heated and hydrolyzed to a degree of hydrolysis of 2.2 %. Subtilisin Carlsberg, chymotrypsin, trypsin, bromelain, and papain were used as enzymes. The solubility of the hydrolysates obtained was studied as a function of pH. As well as glycinin peptides obtained after hydrolysis with subtilisin Carlsberg, at neutral pH, glycinin hydrolysates obtained by hydrolysis with chymotrypsin, bromelain, and papain also showed a lower solubility compared to the non-hydrolyzed glycinin. This decrease in solubility was not observed for glycinin hydrolyzed by trypsin, and also not for hydrolysates from β-conglycinin, soy protein isolate, and whey-derived protein preparations. These results show that glycinin has a strong tendency to aggregate upon limited hydrolysis in comparison to the other substrates. Since chymotrypsin prefers to cleave next to hydrophobic amino acids, which are relatively abundant in the hydrophobic regions of glycinin, a stronger exposure of hydrophobic groups as a consequence of hydrolysis is probably the driving force in the aggregate formation of glycinin hydrolysates.

Kuipers, B.J.H.; Alting, A.C.; Gruppen, H. *Submitted*

## INTRODUCTION

To use soy proteins in foods, depending on the application, these proteins must possess appropriate functional properties such as solubility, emulsifying, foaming, or gelling properties *(1)*. A good solubility of proteinaceous material is a prerequisite for the manifestation of many functional properties. To obtain these functional properties the protein can be hydrolyzed, since this is known to be a means to improve the solubility, and emulsifying and foaming properties *(2)*. However, enzymatic hydrolysis may also lead to an increase in hydrophobic exposure, which can prevent the solubility improvement.

Upon hydrolysis of soy and whey proteins *(3-6)* the solubility is indeed not always increased. In these cases, hydrolysis resulted in the formation of insoluble aggregates, indicating that peptides are formed with a stronger tendency to aggregate in comparison to the parental proteins. Enzymatic hydrolysis that promotes aggregation can be of industrial importance since it can for example be used in the production of soy milk curds (tofu) with an improved microstructure *(7)*. Understanding the molecular basis for the different aggregation behavior of soy proteinaceous material after enzymatic hydrolysis can, therefore, be of help to improve the texture of protein rich foods.

In the literature a lot of information can be found about the influence of enzymatic hydrolysis of proteins on the solubility or aggregation behavior of soy *(8-13)* and whey *(6, 14, 15)* proteins. However, it is difficult to compare these results. This is due to the fact that among these studies parameters influencing the aggregation behavior (e.g. pretreatment of the substrate, enzyme used, methods to determine the degree of hydrolysis, and the ionic strength and temperature) vary greatly. Therefore, accurate comparison of the effects of different proteases on soy proteins is lacking. Moreover, no information on the capability of the different soy proteins to yield aggregating peptides is available.

Soy protein preparations mainly consist of the two storage proteins glycinin and β-conglycinin, which in the literature are often referred to as 11S and 7S protein, respectively *(16)*. In soy protein isolate (SPI) the proportion of glycinin is approximately 2 times higher than the proportion of β-conglycinin *(17, 18)*.

Recently, Zhong and co-workers *(12)* showed that various commercial protease preparations have the ability to induce gelation of SPI. However, the individual contributions of glycinin and β-conglycinin in these enzyme-induced gels from SPI were not studied. In a previous study *(17)* we have shown that limited hydrolysis of soy glycinin by the protease subtilisin Carlsberg results in a decrease in solubility at neutral pH. In contrast, β-conglycinin-derived peptides remained completely soluble at this pH. To investigate whether this phenomena is unique for the enzyme/substrate combination used, in the present study also other enzymes (bromelain, papain, trypsin, and chymotrypsin) were used. The effects of hydrolysis were studied with respect to the solubility as a function of pH, using the same conditions for each protease to enable a valid comparison. In addition, since whey proteins are known for their ability to aggregate upon hydrolysis *(5, 6, 19)*, in this study also whey protein hydrolysates,

obtained via hydrolysis using subtilisin Carlsberg, chymotrypsin, and trypsin, were studied in order to determine the substrate dependency.

## MATERIALS AND METHODS

### Materials and chemicals

SPI, glycinin and β-conglycinin were prepared from defatted soybean meal, prepared from Hyland soybeans (non-GMO) supplied by Fa. L.I. Frank (Twello, The Netherlands) as described previously *(20)*. Densitometric analysis of an SDS-PAGE gel of SPI (10-20 % Tris-HCl Ready Gel; Bio-Rad, Hercules, CA), showed that, by ascribing protein bands of 10 to 40 kDa to glycinin and 50 to 90 kDa to β-conglycinin, at least 90 % of the protein bands present in the SPI originate from glycinin or β-conglycinin. The glycinin/β-conglycinin ratio was 64/36 (w/w) as determined by size-exclusion chromatography. The glycinin and β-conglycinin preparations obtained have a protein content of 87 % (N x 5.57) and 81 % (N x 5.60), respectively. Both preparations have a purity > 95 % as determined by densitometric analysis of an SDS-PAGE gel containing both preparations (no data shown). BiPro™, BioPURE-Alphalactalbumin™, and BioPURE-Betalactoglobulin™ were obtained from Davisco Foods International (Eden Prairie, MO). These protein preparations are further denoted as WPI, α-lactalbumin, and β-lactoglobulin, respectively. As specified by the supplier, all three whey protein preparations have a protein content of at least 95 % (w/w). α-Lactalbumin and β-lactoglobulin both have a purity of at least 90 %. WPI was regarded to contain β-lactoglobulin, α-lactalbumin and BSA in a 0.80 : 0.14 : 0.06 weight ratio, as given by the supplier.
Subtilisin Carlsberg from *Bacillus licheniformis*, chymotrypsin (TLCK treated), trypsin (TPCK treated), bromelain from pineapple stem, and papain from *Papaya latex*, were obtained from Sigma Chemical Co. (St. Louis, MO, USA, article numbers P-5380, C-3142, T-1426, B-4882, and P-4762, respectively).
Phenylmethylsulphonylfluoride (PMSF), Bowman-Birk Inhibitor (BBI), iodoacetamide (IAA) and E-64 were obtained from Sigma. All other chemicals were of analytical grade and were purchased from Merck (Darmstadt, Germany) or Sigma.

### Hydrolysis of protein preparations

The SPI, glycinin, β-conglycinin, WPI, α-lactalbumin, and β-lactoglobulin preparations were suspended in Millipore water at concentrations of 1.5 - 2 % (w/w) protein and stirred overnight at 4°C. The pH was adjusted to 8.0, if necessary. The SPI suspension was centrifuged (20 min, 18,000 x *g*, 20°C), whereas the other preparations were filtered using a 0.45 μm filter (Schleicher & Schuell GmbH, Dassel, Germany). The protein concentrations of the supernatant and filtrates were determined using the Dumas method. Solutions were subsequently diluted to 1 % (w/w) protein using Millipore water followed by heating at 95°C

for 30 minutes. The heated solutions all had a clear opaque appearance. Hydrolysates with degrees of hydrolysis (DH) of 2.2 % were prepared at pH 8.0 and 40ºC using subtilisin Carlsberg, trypsin, chymotrypsin, bromelain, and papain. The pH and DH were controlled using the pH-stat method by using a 719S Titrino (Metrohm ion analysis, Herisau, Switzerland). The $h_{tot}$ values used for the determination of the DH were calculated on the basis of the amino acid compositions of the proteins used as found in Swiss-Prot (www.expasy.org; Primary accession numbers used were: P04776 (Glycinin 1; A1a and B2 polypeptides), P04405 (Glycinin 2; A2 and B1a polypeptides), P11828 (Glycinin 3; A1b and B1b polypeptides), P02858 (Glycinin 4: A5, A4 and B3 polypeptides), P04347 (Glycinin 5; A3 and B4 polypeptide)), P11827(β-Conglycinin, α'-chain), P13916 (β-Conglycinin, α-chain), and P25974 (β-Conglycinin, β-chain)).

Since β-conglycinin is a glycoprotein and is reported to contain ~4.5% carbohydrates *(21, 22)*, the nitrogen conversion factor can be calculated to be ~5.87. However, this value would be appropriate when the β-conglycinin content is determined. During proteolysis only the polypeptide chain is cleaved and all the sample preparations are standardized at the same polypeptide content. As a consequence, for β-conglycinin a nitrogen conversion factor of 5.60 (based on the amino acid sequence) was used. For glycinin a nitrogen conversion factor of 5.57 is calculated. In this work we regard SPI as containing only glycinin and β-conglycinin in a ~64/36 (w/w) ratio *(17)*. Therefore, a nitrogen conversion factor of 5.58 was used for SPI.

For WPI, α-lactalbumin, and β-lactoglobulin, the nitrogen conversion factors used were 6.32, 6.25 and 6.34, respectively. They were based on their amino acid compositions as found in Swiss-Prot (www.expasy.org; Primary accession numbers used were: P02769 (BSA), P00711 (α-lactalbumin), and P02754 (β-lactoglobulin).

The $h_{tot}$ values calculated were 8.77, 8.85, and 8.63 for SPI, glycinin, and β-conglycinin *(17)*, respectively, and 8.78, 8.60, 8.80 for WPI, α-lactalbumin, and β-lactoglobulin, respectively.

The enzyme concentrations used, as presented in Table 1, were experimentally determined to be sufficient to reach a DH of 2.2 % within ~20 minutes. The enzymes were dissolved in Millipore water and directly added to the solutions. The molarity of the NaOH solution used to maintain the pH at pH 8.0 was 0.2 M, and accurately determined by titration prior to each hydrolysis experiment. When the desired DH was reached, the enzymes were inhibited using the inhibitors at the concentrations presented in Table 1. Subtilisin Carlsberg and chymotrypsin were both inhibited using PMSF from a stock solution in 2-propanol. Trypsin was inhibited using BBI. Bromelain was inhibited by a mixture of IAA and E-64, whereas papain was inhibited solely by E-64. The pH-stat experiment was stopped after the enzyme was inhibited completely (the pH was constant at pH 8.0). Hydrolysates were freshly prepared prior to each experiment.

**Table 1:** Enzymes and corresponding inhibitors used during the hydrolysis.

| **Enzyme** | | **Inhibitor** | |
|---|---|---|---|
| *Type* | *Concentration (% (w/w))* | *Type* | *Concentration (mM)* |
| Subtilisin Carlsberg | 0.001 | PMSF | 1 |
| Chymotrypsin | 0.002 | PMSF | 1 |
| Trypsin | 0.002 | BBI | 0.005 |
| Bromelain | 0.01 | IAA + E-64 | 0.2 + 0.02 |
| Papain | 0.02 | E-64 | 0.02 |

**Solubility of hydrolysates**

The aggregation behavior of the hydrolysates and non-hydrolyzed solutions was studied via the solubility behavior at different pH values at a protein concentration of 0.8 % (w/w) ($I$ = 0.03 M). For determining the solubility of the non-hydrolyzed samples, the heated protein prior to hydrolysis was used.

The ionic strengths of the heated, non-hydrolyzed, protein solutions were arbitrarily taken as 0 M. During hydrolysis the ionic strength of the solutions increases due to the release of charged groups. The increase in ionic strength during hydrolysis is assumed to be equal to the amount of NaOH added to keep the pH at 8.0. This resulted in an increase in ionic strength of 0.003 M for hydrolysates with a DH of 2.2 %. To end up with a final protein concentration of 0.8 % and an ionic strength of 0.03 M, both Millipore water and a 0.5 M NaCl solution were added. The solutions were acidified stepwise from pH 8.0 to pH 3.0 using 1 M HCl. Samples were taken every 0.5 pH units. The samples were incubated in a head-over-tail rotator for 1 h. Part of the sample was kept apart and denoted as "total" sample. The remainder was centrifuged (20 min, 22,000 x *g*, 20°C). The concentration of proteinaceous material in the supernatant was determined by diluting the supernatant 200 times with 1 % (w/v) SDS in Millipore water. The same was done for the total sample. Subsequently, their absorbances at 220 nm ($A_{220}$) were determined using a Shimadzu UV-1601 spectrophotometer (Shimadzu, Kyoto, Japan). The absorbances measured were corrected for the minor dilutions occurring during the acidification with HCl. As a consequence, the solubility was calculated by expressing the $A_{220}$ of the supernatant as proportion of the $A_{220}$ of the total hydrolysate at pH 8.0.

**Analysis of protein content**

The nitrogen content of the various samples was determined in duplicate by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milano, Italy) according to the manufacturer's instructions. Methionine was used as a standard. The nitrogen conversion factors for SPI, glycinin, and β-conglycinin are 5.58, 5.57, and 5.60, respectively. They were based on the average amino acid compositions of the proteins as found in Swiss-Prot, as shown above.

## RESULTS

### Soy-derived proteins and their hydrolysates

The intact, heated soy-derived proteins have solubilities between 90 and 100 % around pH 8.0 and have a clear appearance in solution. Upon decreasing the pH, the solubility decreases and reaches its minimum around the p*I* of the protein and subsequently increases upon further acidification. Upon lowering the pH from 8.0, the solubility of glycinin starts to decrease at a higher pH value than β-conglycinin as can be observed in Figure 1. At a pH below the p*I*, β-conglycinin starts to increase in solubility at a lower pH than glycinin. This difference is in line with the difference in solubility of native glycinin and β-conglycinin *(23)*. In fact, this difference is also used in the purification of the glycinin and β-conglycinin preparations by acid precipitation over different pH intervals *(17)*. Figure 1 shows the solubility curves of soy-derived protein hydrolysates (DH = 2.2 %) obtained with subtilisin Carlsberg, chymotrypsin, trypsin, bromelain, and papain, together with the parental proteins. For all hydrolysates it can be observed that, independent of the enzyme used, the solubility increases around the p*I* (~pH 5). This agrees with the general observations that hydrolysis increases protein solubility.

For the SPI hydrolysates obtained using subtilisin Carlsberg, chymotrypsin, bromelain, and papain, it can be observed that the hydrolysates are almost completely soluble above pH 7. They start decreasing in solubility below pH 7 and reach their lowest solubility around pH 4.5. At pH 6.5 and 7.0, the solubilities of the hydrolyzed SPIs are lower than the solubilities of the non-hydrolyzed SPI (arrows A, D, G, and J). For the glycinin hydrolysates it can be observed that already at pH 8.0 hydrolysis results in a strong decrease in solubility when subtilisin Carlsberg, chymotrypsin, and papain are used (arrows B, E, and K). When using bromelain this decrease starts at pH 7.5 (arrow H).

Upon hydrolysis of β-conglycinin with subtilisin Carlsberg and chymotrypsin, the hydrolysates are only slightly less soluble at pH 6 when compared to the solubility of non-hydrolyzed β-conglycinin (arrows C and F). However, the bromelain and papain hydrolysates of β-conglycinin result in a lower solubility around neutral pH (arrows I and L).

In general, it can be concluded that chymotrypsin, subtilisin Carlsberg, bromelain, and papain exhibit similar effects and yield hydrolysates with lower solubilities than that of the non-hydrolyzed samples (at pH 7-8). This is in contrast to the solubilities of the hydrolysates obtained with trypsin, which are at each pH (with pH 3 as the only exception) similar or higher than those of the parental proteins.

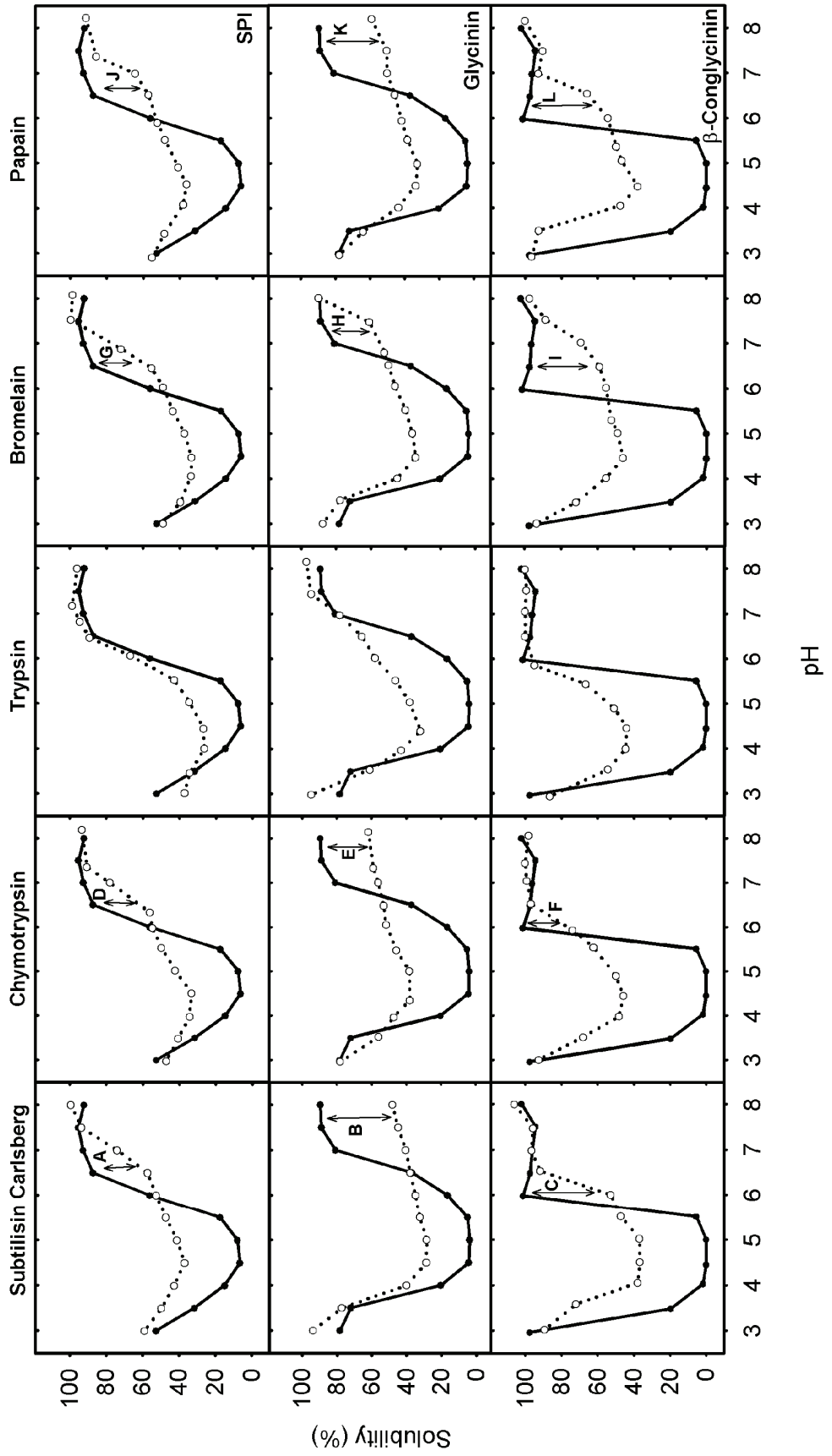


**Figure 1**: Solubility curves (0.8% (w/w) protein, $I = 0.03$ M) as a function of pH for intact (●), heated SPI, heated glycinin and heated β-conglycinin, and their DH 2.2 % hydrolysates (○) obtained with different proteases.

**Whey-derived proteins and their hydrolysates**

As can be observed in Figure 2, the intact whey-derived proteins used in this study, after heating for 30 minutes at 90°C at neutral pH, exhibit solubilities of ~100% at pH 6 and higher. The solubilities decrease to values between 20 and 40 % around the p*I* of the proteins. These are 4.80 and 4.83 for α-lactalbumin and β-lactoglobulin, respectively, based on their amino acid compositions as presented in Swiss-Prot. Upon further acidification, the solubility increases towards 100 % again.

In general, native whey protein isolates have a good solubility at their p*I*. However, upon heating, the solubility around the p*I* is decreased *(24)*. This can also be observed in Figure 2. The non-heated WPI, α-lactalbumin, and β-lactoglobulin have a solubility of ~ 85, 40, and 90 %, respectively, around their p*I* (0.8% (w/w) protein, $I$ = 0.03 M; results not shown). This shows that upon heating the solubility around the p*I* of all the whey-derived protein preparations decreased.

As well as the solubility curves for the intact proteins, Figure 2 also shows the solubility curves for hydrolyzed whey-derived proteins (DH = 2.2 %) using subtilisin Carlsberg, chymotrypsin, and trypsin. The most remarkable observation was that, in contrast to soy glycinin (Figure 1), for all hydrolysates the solubility follows a similar trend to that of the corresponding non-hydrolyzed preparation. The only difference was that upon hydrolysis around the p*I* an increase in solubility can be observed.

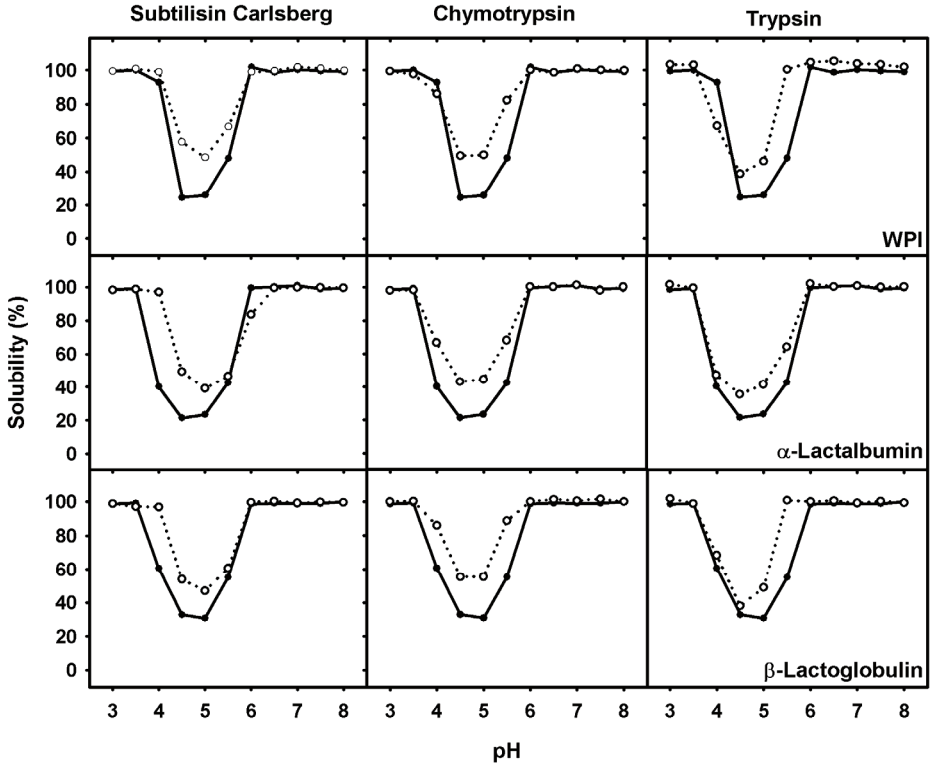


**Figure 2**: Solubility curves (0.8% (w/w) protein, $I$ = 0.03 M) as a function of pH for intact (●), heated WPI, heated α-lactalbumin, and heated β-lactoglobulin and their DH 2.2 % hydrolysates (○) obtained with different proteases.

## DISCUSSION

### Substrate and enzyme dependency of the solubility of hydrolysates

In our previous work *(17)* we showed that upon hydrolysis of glycinin, subtilisin Carlsberg was able to decrease the solubility at neutral pH. In contrast, β-conglycinin hydrolysates remained soluble at this pH. To find out whether the decrease in solubility of subtilisin Carlsberg-hydrolyzed glycinin was unique for this substrate-enzyme combination the present study was performed in order to investigate whether also other enzymes, resulting in hydrolysates with different peptide compositions, were able to decrease the solubility of glycinin. Other substrates were also tested using the same enzymes.

The most remarkable result of this study was that as well as subtilisin Carlsberg, when hydrolyzed to the same DH, also the actions of chymotrypsin, bromelain, and papain were able to strongly decrease the solubility of glycinin at neutral pH. Incubation with subtilisin Carlsberg also slightly decreases the solubility of SPI, whereas for trypsin no decrease in solubility was observed for any of the hydrolyzed substrates (Figure 1). Chymotrypsin and subtilisin Carlsberg both seem to result in hydrolysates with the same solubility behavior around neutral pH. They show a decrease in solubility for glycinin, and no change in solubility for β-conglycinin. Papain and bromelain were also able to slightly decrease the solubility of β-conglycinin.

For whey protein isolates it is known that upon hydrolysis with *Bacillus licheniformis* protease (BLP) the whey proteins are hydrolyzed into peptides that have stronger aggregating properties compared to the parental protein *(15, 19)*. For this reason, WPI, α-lactalbumin, and β-lactoglobulin were hydrolyzed by subtilisin Carlsberg, chymotrypsin, and trypsin to investigate whether these enzymes can also induce aggregation of whey proteins. The hydrolysates were studied under exactly the same conditions (0.8 % (w/w) protein, DH = 2.2 %, $I$ = 0.03 M) as used for the soy-derived proteins. From the results (Figure 2) it is clear that for each substrate-enzyme combination used there was no pH in the range of 3 to 8 at which the solubility was strongly decreased upon hydrolysis, although enzymes were used that have the capability to decrease soy protein solubility.

Overall, it can be concluded that glycinin, compared to β-conglycinin and whey-derived proteins, has a strong tendency to aggregate upon hydrolysis by the proteases used in this study. This shows that there is a strong substrate dependency for the aggregation of proteins upon enzymatic hydrolysis. As well as the substrate dependency, there is also an enzyme dependency. Glycinin does not aggregate upon hydrolysis with trypsin, whereas it does aggregate upon hydrolysis with the other enzymes.

### Mode of action of proteases used

As discussed above, glycinin has the strongest capacity to yield peptides that aggregate at neutral pHs. Trypsin was the only enzyme used that was not able to form strong aggregating peptides from glycinin.

Trypsin has a well-defined specificity, cleaving next to the hydrophilic amino acids Lys and Arg at the $P_1$ position *(25)*. Chymotrypsin has a broader specificity, strongly preferring cleavage next to aromatic amino acids (Tyr, Trp, Phe) and Leu in the $P_1$ position, and to a much lower extent next to a few other amino acids *(26)*. This difference between trypsin, cleaving next to hydrophilic amino acids, and chymotrypsin, cleaving next to hydrophobic amino acids, might be the key to the differences observed between the solubility profiles of glycinin hydrolysates obtained with these two enzymes.
In contrast to trypsin, subtilisin Carlsberg is a relatively nonspecific protease, strongly preferring to cleave next to Leu, Tyr, Phe, and Trp and also cleaving next to Gln, Ser, and Tyr *(27, 28)*. When only hydrolyzed to a DH of 2.2 % it is likely that cleavage first occurs at peptide bonds for which the enzyme possesses the highest selectivity. For subtilisin Carlsberg this seems to be similar to chymotrypsin. Only upon hydrolysis to higher DH values, significant cleavage next to amino acids to which the enzyme possesses a lower selectivity is expected. This may explain why SPI, glycinin, and β-conglycinin hydrolysates obtained after hydrolysis with subtilisin Carlsberg and chymotrypsin up to DH 2.2 % show the same solubility behavior.
Papain is known to have a rather broad specificity. Nevertheless, there is a slight preference for Lys and Arg at the $P_1$ position, and bulky hydrophobic (eg. Phe) amino acids at the $P_2$ position of the active site *(29)*, which seems to be an intermediate between the activities of trypsin and chymotrypsin. Bromelain is known to have a broad specificity, but prefers hydrophilic amino acids in the $P_1$ and $P_1$' positions, with the highest specificity for Glu, Asn, Lys, or Arg ($P_1$) and Ala, Thr, or Ser ($P_1$') *(30)*. Due to the broad specificity of papain and bromelain, it is difficult to speculate which peptide bonds will be split initially.
Hydrophobic interactions are known to be the dominant driving force in peptide aggregation *(11, 31)*. To induce aggregation upon hydrolysis it therefore might be of importance that the protease used is able to cut in the hydrophobic regions. This results in a disruption of these hydrophobic regions with an increased exposure of hydrophobic groups as a result.

**Distribution of cleavage sites**

Since trypsin and chymotrypsin have the best-defined specificities, their cleavage positions were studied in relation to the hydrophobicity score regions present in the protein. Figure 3 shows the hydrophobicity of Glycinin 1 according to Eisenberg *(32)* as a function of its amino acid sequence. The dotted line between the acidic and basic polypeptide represents the disulfide bridge. Those regions in the amino acid sequence for which the hydrophobic score was above 0 for at least 10 amino acids in a row were regarded as hydrophobic regions, as indicated in grey in Figure 3. The dashes on top of the graph represent the position at which trypsin ("Try": Arg and Lys) and chymotrypsin ("Chy"; Phe, Tyr, Trp, and Leu) prefer to cleave. These graphs were also prepared for Glycinin 2 to 5, β-conglycinin α, α', and β, as well as for α-lactalbumin and β-lactoglobulin, but are not shown here. For Glycinin 1 it can be observed that chymotrypsin has more theoretical cleavage sites in the hydrophobic (grey)

regions than trypsin. Table 2 shows the proportions of amino acids present in the defined hydrophobic regions, the numbers of theoretical cleavage sites for trypsin and chymotrypsin for each protein, as well as the proportion of these cleavage sites present in the hydrophobic regions defined.

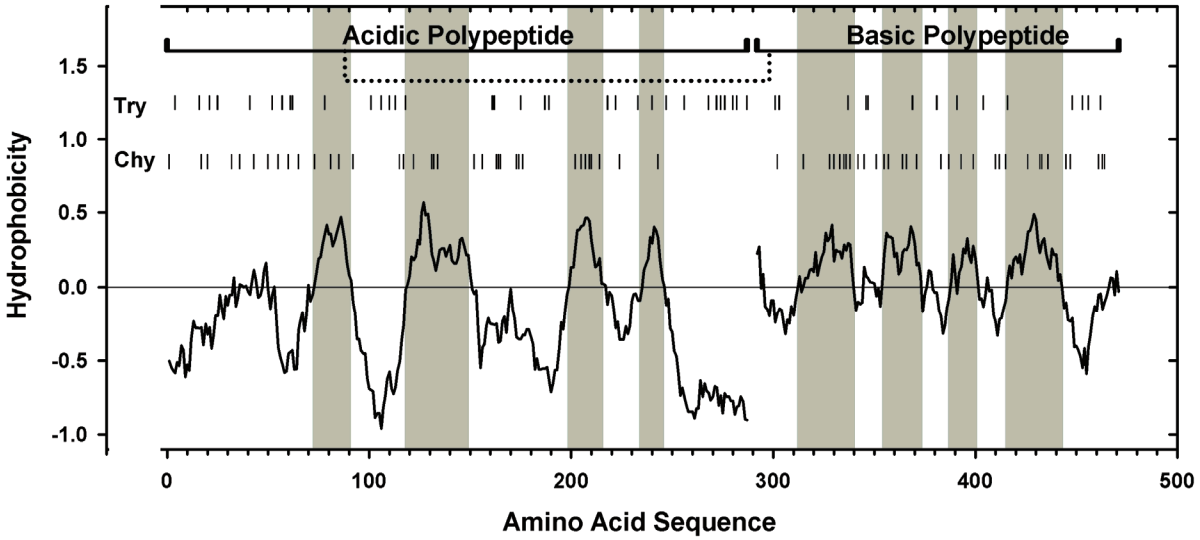


**Figure 3**: Hydrophobicity chart of Glycinin 1, according to Eisenberg *(32)*. The dotted line between the acidic and basic polypeptide represents the disulfide bridge. The grey regions represent the defined hydrophobic regions and the dashes represent the places where trypsin (Try) and chymotrypsin (Chy) can cleave.

Table 2 shows that the proportions of amino acids present in the hydrophobic regions are rather similar for all proteins studied (~ 30 %), with the α`-chain of β-conglycinin (~22 %) as an exception. For glycinin and for β-conglycinin 47 - 54 % and 32 - 42 % of the chymotrypsin cleavage sites are present in the hydrophobic regions of the protein, respectively. For trypsin the cleavage sites for glycinin and β-conglycinin are 7 - 17 % and 7 - 8 % located in the hydrophobic regions, respectively. If selectivity is not taken into account, the probability that chymotrypsin cleaves in the hydrophobic regions of the protein is larger than that of trypsin. Cleavage in a hydrophobic region might yield a stronger exposure of hydrophobic groups. This might explain the stronger aggregation when hydrolyzed with chymotrypsin than with trypsin.

**Table 2**: Total number of trypsin and chymotrypsin cleavage sites in glycinin, β-conglycinin, α-lactalbumin, and β-lactoglobulin, and their proportion as present in the hydrophobic regions.

| | | Trypsin | | Chymotrypsin | |
|---|---|---|---|---|---|
| | % of AA in hydrophobic regions | Total number of cleavage sites | % in hydrophobic regions | Total number of cleavage sites | % in hydrophobic regions |
| Glycinin 1 | 34.7 | 47 | 17 | 68 | 40 |
| Glycinin 2 | 34.3 | 43 | 12 | 67 | 49 |
| Glycinin 3 | 34.3 | 45 | 7 | 69 | 49 |
| Glycinin 4 | 28.9 | 63 | 13 | 72 | 47 |
| Glycinin 5 | 32.3 | 51 | 14 | 68 | 54 |
| β-Conglycinin α | 28.7 | 74 | 7 | 86 | 35 |
| β-Conglycinin α' | 22.2 | 76 | 7 | 85 | 32 |
| β-Conglycinin β | 30.0 | 49 | 8 | 81 | 42 |
| α-Lactalbumin | 30.1 | 13 | 31 | 25 | 36 |
| β-Lactoglobulin | 33.7 | 18 | 33 | 32 | 31 |

For α-lactalbumin and β-lactoglobulin, the proportions of chymotrypsin cleavage sites present in the hydrophobic regions are quite similar (between 31 and 36 %) and lower than those for glycinin. The proportion of trypsin cleavage sites in the hydrophobic regions of α-lactalbumin and β-lactoglobulin is relatively high compared to those for glycinin and β-conglycinin. Nevertheless, they are substantially lower than the proportions of chymotrypsin cleavage sites present in the hydrophobic regions of glycinin. The latter might partly explain why α-lactalbumin and β-lactoglobulin do not form aggregates upon hydrolysis with chymotrypsin.

It can be concluded that glycinin is rather susceptible to aggregation upon limited hydrolysis using various non-specific proteases. The results further suggest that cleavage of peptide bonds in the hydrophobic regions might be important for the capacity to yield aggregating peptides.

## ACKNOWLEDGEMENTS

The soybeans were kindly provided by Fa. L.I. Frank (Twello, The Netherlands). Leonie Hermans is acknowledged for performing some of the experiments.

## REFERENCE

1. Kinsella, J.E. Functional-properties of soy proteins. *J. Am. Oil Chem. Soc.* **1979,** *56*, (3), 242-258.
2. Kilara, A.; Panyam, D. Peptides from milk proteins and their properties. *Crit. Rev. Food Sci. Nutr.* **2003,** *43*, (6), 607-633.
3. Fuke, Y.; Sekiguchi, M.; Msuoka, H. Nature of stem bromelain treatments on the aggregation and gelation of soybean proteins. *J. Food Sci.* **1985,** *50*, 1283-1288.

4. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.
5. Ju, Z.Y.; Otte, J.; Zakora, M.; Qvist, K.B. Enzyme-induced gelation of whey proteins: Effect of protein denaturation. *Int. Dairy J.* **1996,** *7*, 71-78.
6. Doucet, D.; Gauthier, S.F.; Foegeding, E.A. Rheological characterization of a gel formed during extensive enzymatic hydrolysis. *J. Food Sci.* **2001,** *66*, 711-715.
7. Lee, C.W., and Jo, G.Y. The microstructures of soybean milk curds prepared by different coagulation methods. *J. Food Sci. Nutr.* **1997,** *2*, 259-262.
8. Jung, S.; Murphy, P.A.; Johnson, L.A. Physicochemical and functional properties of soy protein substrates modified by low levels of protease hydrolysis. *J. Food Sci.* **2005,** *70*, (2), C180-C187.
9. Jung, S.; Roussel-Philippe, C.; Briggs, J.L.; Murphy, P.A.; Johnson, L.A. Limited hydrolysis of soy proteins with endo- and exoproteases. *J. Am. Oil Chem. Soc.* **2004,** *81*, (10), 953-960.
10. Kim, S.Y.; Park, S.W.; Rhee, K.C. Functional properties of proteolytic enzyme modified soy protein isolate. *J. Agric. Food Chem.* **1990,** *38*, 651-656.
11. Zhong, F.; Xin, Y.; Yue, L.; Shoemaker, C.F. Papain-induced gelation of soy glycinin (11S). *J. Food Sci.* **2006,** *71*, (5), E232-E237.
12. Zhong, F.; Wang, Z.; Xu, S.Y.; Shoemaker, C.F. The evaluation of proteases as coagulants for soy protein dispersions. *Food Chem.* **2007,** *100*, (4), 1371-1376.
13. Mohri, M.; Matsushita, S. Improvement of water absorption of soybean protein by treatment with bromelain. *J. Agric. Food Chem.* **1984,** *32*, 486-490.
14. Mutilangi, W.A.M.; Panyam, D.; Kilara, A. Functional properties of hydrolysates from proteolysis of heat-denatured whey protein isolate. *J. Food Sci.* **1996,** *61*, (2), 270-274.
15. Ju, Z.Y.; Kilara, A. Gelation of hydrolysates of a whey protein isolate induced by heat, protease, salts and acid. *Int. Dairy J.* **1998,** *8*, (4), 303-309.
16. Liu, K. *Soybeans: Chemistry, Technology and Utilization*. Chapman & Hall: New York, 1997.
17. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophys.* **2006,** *1*, 178-188. *Chapter 3 in this thesis*.
18. Fehr, W.R.; Hoeck, J.A.; Johnson, S.L.; Murphy, P.A.; Nott, J.D.; Padilla, G.I.; Welke, G.A. Genotype and environment influence on protein components of soybean. *Crop Science* **2003,** *43*, (2), 511-514.
19. Creusot, N.; Gruppen, H.; van Koningsveld, G.A.; de Kruif, C.G.; Voragen, A.G.J. Peptide-peptide and protein-peptide interactions in mixtures of whey protein isolate and whey protein isolate hydrolysates. *Int. Dairy J.* **2006,** *16*, (8), 840-849.
20. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. *J. Agric. Food Chem.* **2005,** *53*, 1031-1038. *Chapter 2 in this thesis*.
21. Hou, H.J.; Chang, K.C. Structural characteristics of purified β-conglycinin from soybeans stored under four conditions. *J. Agric. Food Chem.* **2004,** *52*, (26), 7931-7937.
22. Koshiyama, I. Chemical and physical properties of a 7S protein in soybean globulins. *Cereal Chem.* **1968,** *45*, 394-404.
23. Thanh, V.H.; Shibasaki, K. Major proteins of soybean seeds. A straightforward fractionation and their characterization. *J. Agric. Food Chem.* **1976,** *24*, 1117-1121.

24. Bertrand-Harb, C.; Baday, A.; Dalgalarrondo, M.; Chobert, J.M.; Haertle, T. Thermal modifications of structure and co-denaturation of α-lactalbumin and β-lactoglobulin induce changes of solubility and susceptibility to proteases. *Nahrung* **2002,** *46*, (4), 283-289.
25. Chen, J.M.; Ferec, C. Trypsin. In *Handbook of Proteolytic Enzymes*, 2nd ed.; Barrett, A. J.; Rawlings, N. D.; Woessner, J. F., Eds.; Elsevier: London, 2004; pp 1483-1493.
26. Graf, L.; Szilagyi, L.; Venekei, I. Chymotrypsin. In *Handbook of Proteolytic Enzymes*, 2nd ed.; Barrett, A. J.; Rawlings, N. D.; Woessner, J. F., Eds.; Elsevier: London, 2004; pp 1495-1501.
27. Johansen, J.T.; Ottesen, M.; Svendsen, I.; Wybrandt, G. Degradation of B-chain of oxidized insulin by 2 subtilisins and their succinylated and N-carbamylated derivatives. *C. R. Trav. Lab. Carlsberg* **1968,** *36*, (20), 365-384.
28. Ottesen, M.; Svendsen, A. The subtilisins. *Methods Enzymol.* **1970,** *19*, 199-215.
29. Menard, R.; Storer, A.C. Papain. In *Handbook of Proteolytic Enzymes*, 2nd ed.; Barrett, A. J.; Rawlings, N. D.; Woessner, J. F., Eds.; Elsevier: London, 2004; pp 1125-1128.
30. Napper, A.D.; Bennett, S.P.; Borowski, W.; Holdridge, M.B.; Leonard, M.J.C.; Rogers, E.E.; Duan, Y.J.; Laursen, R.A.; Reinhold, B.; Shames, S.L. Purification and characterization of multiple forms of the pineapple-stem-derived cysteine proteinases Ananain and Comosain. *Biochem. J.* **1994,** *301*, 727-735.
31. Nagai, K.; Inouye, K. Insights into the reaction mechanism of the coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2004,** *52*, (15), 4921-4927.
32. Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984,** *179*, 125-142.

# CHAPTER 5

# Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative RP-HPLC-MS analysis

**ABSTRACT**

The molar extinction coefficients of 20 amino acids and the peptide bond were measured at 214 nm in the presence of acetonitrile and formic acid to enable quantitative comparison of peptides eluting from reversed-phase high-performance-liquid-chromatography, once identified with mass-spectrometry (RP-HPLC-MS). The peptide bond has a molar extinction coefficient of 923 $M^{-1}cm^{-1}$. Tryptophan has a molar extinction coefficient that is ~30 times higher than that of the peptide bond, whereas the molar extinction coefficients of phenylalanine, tyrosine and histidine are ~6 times higher than that of the peptide bond. Proline, as an individual amino acid has a negligible molar extinction coefficient. However, when present in the peptide chain (except at the N-terminus) it absorbs ~3 times more than a peptide bond. Methionine has a similar molar extinction coefficient as the peptide bond while all other amino acids have much lower molar extinction coefficients. The predictability of the molar extinction coefficients of proteins and peptides, calculated by the amino acid composition and the number of peptide bonds present, was validated using several proteins and peptides. The measured and calculated molar extinction coefficients were in good agreement, which shows that it is possible to compare peptides analyzed by RP-HPLC-MS in a quantitative way. This method enables a quantitative analysis of all peptides present in hydrolysates once identified with RP-HPLC-MS.

Kuipers, B.J.H.; Gruppen, H. *J. Agric. Food Chem., accepted for publication.*

## INTRODUCTION

Food-derived peptides may have a positive impact on body functions or condition and ultimately influence human health *(1-4)*. In addition, peptides may also exert technofunctional properties (e.g. solubility and gelation) in foods. Using reversed-phase high-performance-liquid-chromatography mass-spectrometry (RP-HPLC-MS) the sequence of peptides present in protein hydrolysates can be easily obtained when the sequence of the parental proteins is known *(5, 6)*. Besides knowing the sequences of peptides, it is also important to know the quantities of the peptides of interest present in a peptide mixture.

Quantification of all peptides present in a hydrolysate can be carried out via the determination of the peptide bonds as these are present in all peptides. The π electrons in the peptide bond are to some extent delocalized over three atoms: the peptide nitrogen, carbon and oxygen. At low wavelengths (180-220 nm), the $\pi \rightarrow \pi^*$ transition in the peptide bond absorbs light *(7)*. This absorbance can be used to estimate the relative amounts of peptides present in a hydrolysate. However, quantification based on the absorbance at 214 nm is not a common method, since in this wavelength region the absorption of the peptide bond is also conformation dependent *(8)*. Furthermore, besides the peptide bond, several amino acid residues are reported to contribute significantly to the absorption *(9, 10)*. Moreover, the absorbance of peptides is also pH dependent due to the absorption of the carboxylic acid group, which depends on whether the carboxyl group is protonated or not. This results in a variation of the absorbance around the p$K_a$ (~pH 3) of the free carboxylic acid in the peptides *(11)*. This effect plays the most dominant role in dipeptides, and is decreasing with increasing peptide length *(12, 13)*.

Due to the complexity of absorption at 214 nm researchers are, in general, skeptic towards quantification of proteins at this wavelength. But for peptides devoid of tryptophan or tyrosine, there is no good alternative *(14)*. As explained above, the complexity of the absorbance is much higher compared to the absorbance at 280 nm, resulting in a higher expected error at 214 nm. Nevertheless, this has to be accepted due to absence of good alternatives. The high complexity of the absorbance at 214 nm is probably the reason why, to our knowledge, no research is performed to validate whether the absorbance of peptides at 214 nm can be predicted based on its amino acid composition. However, the rapid development of RP-HPLC-MS in the last years enables such an approach and the increase in research on (bio-active) peptides requires such an approach. To quantify the amounts of all peptides present in protein hydrolysates, the molar extinction coefficients of the individual peptides have to be known. In literature several, somewhat conflicting, data have been reported. Buck and co-workers *(15)* and GPMAW software *(16)* both calculate the molar extinction coefficient of a protein/peptide at 214 nm as the sum of the number of peptide bonds and the sum of the number of each type of amino acid multiplied by their individual molar extinction coefficients (ε). The molar extinction coefficients used by Buck and co-workers *(15)* and GPMAW *(16)* can be found in Table 1. In the GPMAW software *(16)* no

reference or experimental data are provided for the molar extinction coefficient of the amino acids and the peptide bond. In addition in Table 1, the molar extinction coefficients of individual amino acids, manually deduced from the 200-230 nm wavelength scans given by Saidel and co-workers *(9)* and Wetlaufer *(10)* are shown as well. The values of Saidel and co-workers *(9)* and Wetlaufer *(10)* were not aimed at representing the absorbances of individual amino acids in a peptide chain. Therefore, in these studies no value is presented for the absorbance of the peptide bond. In general, it can be seen that tryptophan, histidine, phenylalanine, and tyrosine are strong contributors to the absorbance at 214 nm. The values reported for these amino acids agree rather well with each other. However, there are large differences between the four studies with respect to the contributions of the peptide bond, asparagine, glutamine, cysteine, and methionine.

**Table 1**: Molar extinction coefficients at 214 nm ($M^{-1}cm^{-1}$) of amino acids and peptide bond as found in literature and software.

| | **GPMAW *(16)*** | **Buck *(15)*** | **Saidel *(9)*[a]** | **Wetlaufer *(10)*[a]** |
|---|---|---|---|---|
| Peptide bond | 1,000 | 2,846 [c] | *-*[e] | *-*[e] |
| Tyrosine | 5,000 | 5,755 | *5,450* | *6,100* |
| Histidine | 5,000 | 6,309 | *5,700* | *5,550* |
| Phenylalanine | 5,000 | 7,208 | *5,050* | *5,800* |
| Tryptophan | 33,000 | 22,735 | *32,250* | *31,950* |
| Asparagine | ~0[b] | 2,846[d] | *140* | *-*[e] |
| Glutamine | ~0[b] | 2,846[d] | *140* | *-*[e] |
| Cysteine | 1,000 | ~0[b] | *275* | *210* |
| Methionine | 1,000 | ~0[b] | *1,050* | *1,100* |
| Proline | ~0[b] | ~0[b] | *25* | *-*[e] |

[a] Manually deduced from wavelength scans given in the article.
[b] Regarded not to have a significant influence on the absorption at 214 nm.
[c] Determined as the difference in extinction coefficient at 214 nm between N-acetylphenylalanine and phenylalanine.
[d] Regarded to be the same as the peptide bond.
[e] Not determined.

Characterization of protein hydrolysates using RP-HPLC-MS would be improved if a quick and relatively easy method to quantify the amounts of individual peptides present based on their UV absorbances is available. A method as presented by Buck and co-workers *(15)* can be applied. However, one might doubt whether molar extinction coefficients and in particular those of asparagine and glutamine used in that work are correct. In this respect the values, although manually deduced, from the work of Saidel and co-workers *(9)* and from Wetlaufer *(10)* might be more correct. Another issue is the presence of proline. Proline, as a free amino acid, is reported to have a low molar extinction coefficient of ~25 $M^{-1}cm^{-1}$ *(9)*, or not regarded as being a major absorbing group in proteins and peptides *(15, 16)*. However, due to the cyclic nature of the three-carbon side chain to the nitrogen of the peptide backbone, proline

might show different absorption properties when present in a peptide chain (except when present at the N-terminus).
In the present study we, therefore, measured the molar extinction coefficients of all 20 amino acids in the presence of 20 % (v/v) acetonitrile, and 0.1 % (v/v) formic acid (~pH 3). This condition is similar to conditions during RP-HPLC-MS analysis. First, the individual contributions of the peptide building blocks were measured: the individual amino acids and the peptide bond. The next step is the validation of the hypothesis that the molar extinction coefficients can be rather well predicted based on the individual contribution of the building blocks by measuring the molar extinction coefficients of several proteins and peptides.

## MATERIALS AND METHODS

### Materials and chemicals

Amino acids were all purchased from Fluka (Buchs, Switzerland); alanine (A; 5130), arginine (R; 11040), asparagine (N; 11150), aspartic acid (D; 1190), cysteine (C; 30090), glutamine (Q; 49420), glutamic acid (E; 49450), glycine (G; 50050), histidine (H; 53370), isoleucine (I; 58880), leucine (L; 61820), lysine (K; 62930), methionine (M; 64320), phenylalanine (F; 78020), proline (P; 81710), serine (S; 84960), threonine (T; 89180), tryptophan (W; 93660), tyrosine (Y; 93830), and valine (V; 94620).
Gly-Gly-Gly (H-3355), Gly-Gly-Pro (H-3470), Gly-Pro-Gly (H-9745), Pro-Gly-Gly (M-1730), Gly-Tyr-Gly (GYG; H-3670), RGDS (H-1155), YGGFLRR (Dynorphyn A (1-7); (H-2660), GPRP (H-2935), and RPPGFSP (Bradikinin (1-7); H-1955) were obtained from Bachem (Heidelberg, Germany). RPPGFSPFR (Bradikinin; B-3259) was obtained from Sigma (St. Louis, MO, USA). RINKKIEK, EQLSTSEENSK, and YIPIQYVLSR, with a purity of ≥ 95 % as determined by RP-HPLC, were obtained from Ansynth (Ansynth Service BV, Roosendaal, The Netherlands).
Polylysine (P2658) was obtained from Sigma (St. Louis, MO, USA) with a DP of ~190 as indicated by the supplier. Soy glycinin and β-conglycinin were prepared as described previously *(17)*. α-Lactalbumin (Bovine; L5385), β-lactoglobulin (Bovine; L0130), BSA (Bovine; A4503), α-casein (Bovine; C6780), and lysozyme (Egg White; L6876) were obtained from Sigma (St. Louis, MO, USA). All other chemicals used were of analytical grade and purchased from Sigma or Merck.

### Determination of the protein/peptide concentration

The nitrogen content of various proteins, peptides or solutions thereof was determined in duplicate by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milano, Italy) according to the manufacturer's instructions. Methionine was used as a standard. The nitrogen conversion factor for the various proteins was based on the average amino acid compositions as found in the Swiss-Prot database (www.expasy.org). The primary accession numbers used are: Glycinin (P04776, P04405, P11828, P02858, and

P04347), β-Conglycinin (P13916, P25974, and P11827), α-Lactalbumin (P00711), β-Lactoglobulin (P02754), BSA (P02769), α-Casein (P02663, and P02662), and Lysozyme (P0698). In case more than one variety was present they were all regarded to be present in a 1 : 1 ratio.

**Determination of the molar extinction coefficients of peptide building blocks**

All 20 amino acids, polylysine and Gly-Gly-Gly, Gly-Gly-Pro, Gly-Pro-Gly, and Pro-Gly-Gly were dissolved in 20 % (v/v) acetonitrile containing 0.1 % (v/v) formic acid, up to an accurately known concentration of ~5 mM. Aspartic acid, glutamic acid and tyrosine did not fully dissolve in the solvent used. Therefore for these amino acids, HCl was added up to 20 mM, as well as to their corresponding sample blanks. The absorbance was measured at 214 nm using a UV Shimadzu UV-1601 spectrophotometer (Shimadzu, Kyoto, Japan). Several dilutions were prepared to obtain a calibration curve in the linear region of the spectrophotometer.

**Determination of the molar extinction coefficients of proteins and peptides.**

To validate whether the molar extinction coefficient of a protein or peptide can be predicted based on the amino acid composition as found in the Swiss-Prot database (www.expasy.org), the molar extinction coefficients of various proteins and peptides were analyzed. Of all proteins and peptides the protein content was determined using the Dumas method. Peptides that were hygroscopic or present in a too low quantity to weigh accurately (GPRP, YGGFLRR, RINKKIEK, EQLSTSEENSK, YIPIQYVLSR, RPPGFSP, and RPPGFSPFR) were pre-dissolved in water followed by determining the protein concentration in the solution using the Dumas method. Of each protein and peptide a stock solution was prepared containing an accurately known concentration between 0.5 and 1.0 mg/mL in 20 % (v/v) acetonitrile and 0.1 % (v/v) formic acid (v/v). These stock solutions were diluted 25 and 50 times with the following 4 solutions: (1) 5 % (v/v) acetonitrile/0.09925 % (v/v) formic acid, (2) 20 % (v/v) acetonitrile/0.097 % (v/v) formic acid, (3) 35 % (v/v) acetonitrile/0.09475 % (v/v) formic acid, and (4) 50 % (v/v) acetonitrile/0.0925 % (v/v) formic acid. The acetonitrile/formic acid ratio was chosen to mimic RP-HPLC-MS analysis in which 100 % water/0.1% (v/v) formic acid is frequently *(18)* used as buffer A and 100 % acetonitrile/0.085 % (v/v) formic acid is used as buffer B. No higher concentrations of acetonitrile were chosen, since in general peptides elute from an RP-HPLC column before 50 % (v/v) acetonitrile concentration is reached *(5, 19, 20)*. To determine the molar extinction coefficient at 214 nm, the absorbance was measured using a UV Shimadzu UV-1601 spectrophotometer (Shimadzu, Kyoto, Japan).

### Alkylation of α-lactalbumin

α-Lactalbumin, dissolved in 8 M Urea, 50 mM Tris-HCl pH 8.0, was reduced using 50 mM DTT at 40°C for 2 hrs, followed by alkylation for 1 hour in the dark using iodoacetamide (IAA) up to 150 mM. The pH was adjusted to 8.0 after IAA addition. After dialysis against Millipore water and freeze-drying, the alkylated, as well as the untreated α-lactalbumin, were dissolved in water containing 0.1 % (v/v) TFA up to ~ 1 mg/mL. Subsequently, the samples were mixed in a 1 : 1 ratio with matrix solution (10 mg/mL sinapinic acid in 50% (v/v) acetonitrile and 0.3 % (v/v) TFA) and applied onto a MALDI plate. MALDI-TOF MS analysis was performed to verify whether alkylation was sufficient, using an Ultraflex workstation (Bruker Daltonics, Hamburg, Germany) equipped with a nitrogen laser of 337 nm. The mass spectrometer was used in the positive mode. After a delayed extraction time of 340 ns, the ions were accelerated to a kinetic energy of 25 kV. The ions were detected in the linear mode. External calibration of the mass spectrometer was performed using protein calibration standard I (mass range 5000-20000 Da; Part-No. # 206355, Bruker Daltonics, Hamburg, Germany).

The MALDI-TOF mass spectrum of untreated and alkylated α-lactalbumin showed a main peak at 14,177.8 and 14,642.1 Da, respectively (mass spectra not shown). Due to the presence of 8 cysteines in α-lactalbumin, upon alkylation an increase in mass of 464.4 Da (8 x 58.05) is expected, which is the exact difference in observed masses. Of both the untreated and alkylated α-lactalbumin the molar extinction coefficient was determined in duplicate using 20 % (v/v) acetonitrile and 0.1 % (v/v) formic acid as described above.

## RESULTS

### Molar extinction coefficients of individual amino acids and the peptide bond

Table 2 shows the determined molar extinction coefficients of all individual amino acids. Standard deviations were all below 4 % and, therefore, not individually presented in the table. It is clear that tryptophan has the highest molar extinction coefficient (29,050 $M^{-1}cm^{-1}$), followed by tyrosine, phenylalanine and histidine. These amino acids all three have similar molar extinction coefficients just above ~5,000 $M^{-1}cm^{-1}$. Of the sulfur containing amino acids methionine (980 $M^{-1}cm^{-1}$) has a higher absorbance than cysteine (225 $M^{-1}cm^{-1}$). All other amino acids have a molar extinction coefficient below 200 $M^{-1}cm^{-1}$.

Prior to MS of peptides, usually the cysteines are reduced and alkylated to prevent the formation of new disulfide bridges *(21, 22)*. In order to determine the effect of alkylation on the molar extinction coefficient of cysteine, α-lactalbumin was reduced and alkylated. After verifying that the alkylation was complete (material and method section), the molar extinction coefficient of the reduced and both reduced and alkylated α-lactalbumin was determined. There appeared to be no significant difference between the molar extinction coefficients of untreated and alkylated α-lactalbumin (no further data shown). Therefore, it can be concluded that alkylation does not change the molar extinction coefficient cysteine.

**Table 2**: Molar extinction coefficients of free amino acids ($M^{-1}cm^{-1}$) at 214 nm in 20 % (v/v) acetonitrile and 0.1 % (v/v) formic acid.

| Amino Acid | ε ($M^{-1}cm^{-1}$)[a] | Amino Acid | ε ($M^{-1}cm^{-1}$) [a] |
|---|---|---|---|
| Alanine (A) | 32 | Leucine (L) | 45 |
| Arginine (R) | 102 | Lysine (K) | 41 |
| Asparagine (N) | 136 | Methionine (M) | 980 |
| Aspartic acid (D) | 58 | Phenylalanine (F) | 5,200 |
| Cysteine (C) | 225 | Proline (P) | 30 |
| Glutamine (Q) | 142 | Serine (S) | 34 |
| Glutamic acid (E) | 78 | Threonine (T) | 41 |
| Glycine (G) | 21 | Tryptophan (W) | 29,050 |
| Histidine (H) | 5,125 | Tyrosine (Y) | 5,375 |
| Isoleucine (I) | 45 | Valine (V) | 43 |

[a] Standard deviations were all below 4 %.

In order to estimate the molar extinction coefficient of the peptide bond, the molar extinction coefficient of polylysine was determined using Formula 1.

$$\varepsilon_{\text{peptide bond}} = \frac{\varepsilon_{\text{polypeptide}} - (\varepsilon_{\text{lysine}} * n_{\text{lysine}})}{n_{\text{peptide bonds}}} \qquad [1]$$

The molar extinction coefficient for polylysine was determined to be 182,200 ± 450 $M^{-1}cm^{-1}$. The molar extinction coefficient of lysine is 41 $M^{-1}cm^{-1}$ as presented in Table 2. Given the fact that polylysine has a DP of 190, this results in a molar extinction coefficient for the peptide bond of 923 $M^{-1}cm^{-1}$.

### Absorption of proline in a peptide chain

Table 3 shows the molar extinction coefficients of the tri-peptides tri-glycine and the tripeptides with two glycines and one proline, in which proline is present at the N-terminus, C-terminus and in the middle of the peptide chain. It can be observed that Gly-Pro-Gly and Gly-Gly-Pro have extinction coefficients much higher that of Pro-Gly-Gly. The latter value is in the same range as that of the tripeptide Gly-Gly-Gly.

**Table 3**: Measured molar extinction coefficients of the tri-glycine and tri-peptides with 2 glycines and 1 proline in 20 % (v/v) acetonitrile and 0.1 % (v/v) formic acid.

| | ε ($M^{-1}cm^{-1}$)[a] | | ε ($M^{-1}cm^{-1}$)[a] |
|---|---|---|---|
| Gly-Gly-Gly | 1,080 ± 35 | Gly-Pro-Gly | 3,620 ± 5 |
| Pro-Gly-Gly | 950 ± 5 | Gly-Gly-Pro | 3,880 ± 65 |

[a] Measured in duplicate.

**Measured molar extinction coefficient of proteins and peptides.**

Table 4 shows the measured molar extinction coefficients of the several proteins and peptides analyzed. It can be clearly seen that, as expected, the protein exhibit higher molecular extinction coefficient than peptides. The values are the average of the observed molar extinction coefficients measured in duplicate at the 4 different acetonitrile concentrations (5, 20, 35, and 50 % (v/v)) acetonitrile and ~0.1 % (v/v) formic acid. The low standard deviations presented in Table 4 show that there is not a significant influence of the acetonitrile concentration on the absorbance. In the observed variation between different acetonitrile concentrations no trend could be observed (results not shown). These results show that when comparing peaks in one RP-HPLC chromatogram, peaks can be compared without taking into account the acetonitrile concentration at the moment of elution.

**Table 4**: Measured molar extinction coefficients (214 nm) of protein and peptides.

| **Proteins** | $\varepsilon_{measured}$ **($M^{-1}cm^{-1}$)** | **Stdev (%)** | **Peptides** | $\varepsilon_{measured}$ **($M^{-1}cm^{-1}$)** | **Stdev (%)** |
|---|---|---|---|---|---|
| Glycinin | 972,700 | 0.8 | RGDS | 2,960 | 4.4 |
| β-Conglycinin | 841,500 | 1.8 | RINKKIEK | 7,900 | 2.4 |
| α-Lactalbumin | 292,500 | 2.0 | EQLSTSEENSK | 8,410 | 5.6 |
| β-Lactoglobulin | 302,900 | 0.5 | GYG | 5,835 | 1.3 |
| BSA | 1,155,700 | 1.0 | YGGFLRR | 18,000 | 4.5 |
| α-Casein | 495,000 | 3.0 | YIPIQYVLSR | 22,000 | 5.2 |
| Lysozyme | 258,400 | 1.2 | GPRP | 8,980 | 1.4 |
| | | | RPPGFSP | 23,100 | 3.0 |
| | | | RPPGFSPFR | 31,600 | 3.6 |

## DISCUSSION

**Molar extinction coefficients of protein and peptide building blocks**

The molar extinction coefficient of the *peptide bond* is determined based on the absorbance of polylysine at pH 3. This polypeptide was deliberately chosen since at this pH the polypeptide is present in a random coil conformation *(23)*. The molar extinction coefficient was calculated to be 923 $M^{-1}cm^{-1}$. This value strongly deviates from the one reported by Buck and co-workers as presented in Table 1 *(15)*. Based on data of Goldfarb *(11)*, the molar extinction coefficient of a peptide bond, also based on the absorbance of polylysine at 214 nm should be ~ 1000 $M^{-1}cm^{-1}$, which is close to our findings and similar to the value used in GPMAW *(16)*. Therefore, for all further calculations in this report, for the peptide bond a molar extinction coefficient of 923 $M^{-1}cm^{-1}$ is used.

In Figure 1 the measured values of protein and peptide building blocks, as presented in Table 2, are compared with values deduced from Saidel and co-workers *(9)*, Wetlaufer *(10)*, and those reported by GPMAW*(16)*, and Buck and co-workers *(15)*. The amino acids can be divided into four groups, based on their contributions relative to the molar extinction

coefficient of the peptide bond itself. The first group has a high contribution to the absorption at 214 nm, being larger than the absorption of the peptide bond ($\varepsilon_{amino\ acid} > \varepsilon_{peptide\ bond}$ : Pro (not at N-terminus), Trp, Tyr, His, and Phe). The second group has a rather similar molar extinction coefficient as the peptide bond ($\varepsilon_{amino\ acid} \approx \varepsilon_{peptide\ bond}$ : Met). The third and fourth groups have molar extinction coefficients being lower than that of the peptide bond. The molar extinction coefficient of the third group is between 10 and 100 % of that of the peptide bond (10 % < $\varepsilon_{amino\ acid}$ < 100 % $\varepsilon_{peptide\ bond}$ : Arg, Asn, Gln, and Cys), whereas the fourth group has an almost negligible molar extinction coefficient, being less than 10 % of the extinction coefficient of the peptide bond ($\varepsilon_{amino\ acid}$ < 10% $\varepsilon_{peptide\ bond}$ : Pro (at N-terminus), Gly, Ala, Ser, Lys, Thr, Val, Ile, Leu, Asp, and Glu). Below, each group will be discussed individually.

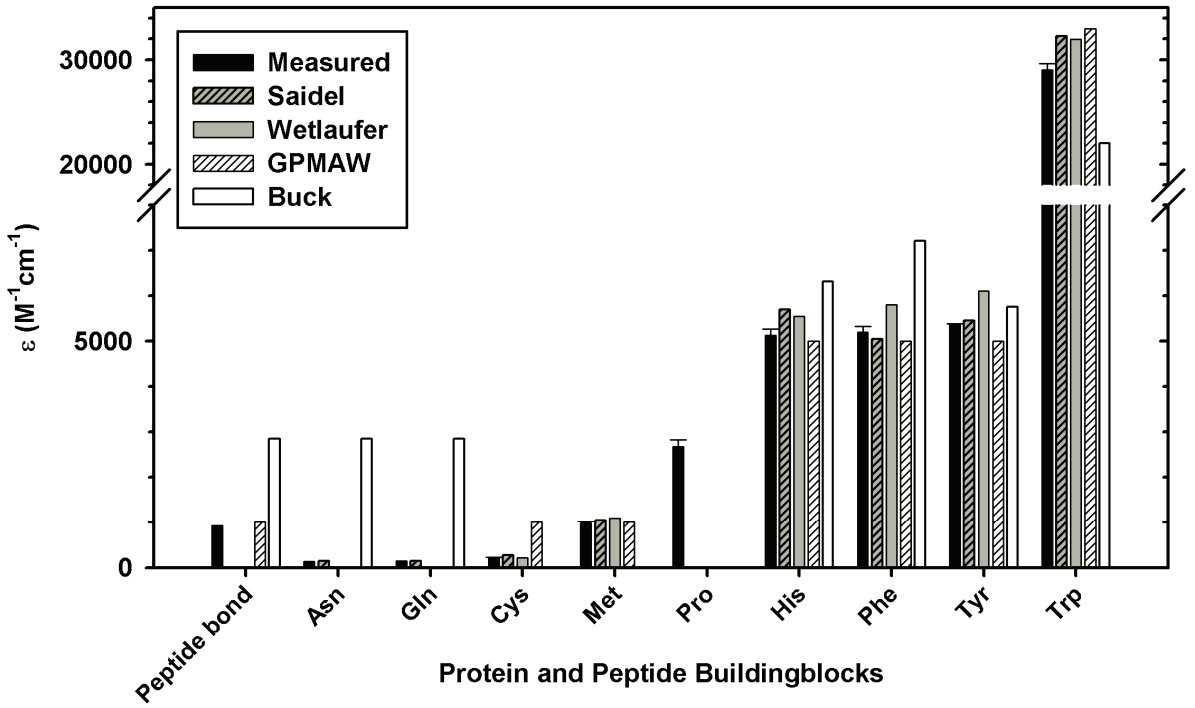


**Figure 1**: Molar extinction coefficients at 214 nm of measured and literature-values for the protein/peptide building blocks (peptide bond and individual amino acids) as presented by Saidel and co-workers *(9)*, Wetlaufer *(10)*, GPMAW*(16)* and Buck and co-workers *(15)*.

*Group I* $\varepsilon_{amino\ acid} > \varepsilon_{peptide\ bond}$

Just as presented in all references, tryptophan is the most dominant contributor to the absorbance at 214 nm. Histidine, phenylalanine, and tyrosine have measured values around 5,000 - 6,000 $M^{-1}cm^{-1}$, which is in the same range as the previously reported values.

As an individual amino acid, proline does not show a significant absorbance (30 $M^{-1}cm^{-1}$) as also shown by Saidel and co-workers *(9)*. This low absorbance is probably the reason why Buck and co-workers *(15)* and GPMAW*(16)* do not regard proline to be a significant

contributor to the absorbance of peptides and proteins. However, when the (secondary) amino group of proline participates in the peptide bond (Gly-Pro-Gly and Gly-Gly-Pro), the molar extinction coefficient strongly increases (Table 3). It should be noted that this increase is not observed when proline is present at the N-terminus of the peptide. The contribution of proline to the molar extinction coefficient of a protein or a peptide, when proline is present in other positions than the N-terminus, can be calculated by the difference in absorbance between Gly-Gly-Gly and the average molar extinction coefficient of Gly-Pro-Gly and Gly-Gly-Pro. This results in a molar extinction coefficient for proline of 2,675 (± 155) $M^{-1}cm^{-1}$. Only when present at the N-terminus of a protein or peptide, we regard the contribution of proline to the absorbance to be 30 $M^{-1}cm^{-1}$, which is close to the value deduced from the work of Saidel and co-workers *(9)*.

*Group II:* $\varepsilon_{amino\ acid} \sim \varepsilon_{peptide\ bond}$

Methionine is the only amino acid having a rather similar molar extinction coefficient as the peptide bond. The result found in this study is found to be in the same range as the ones deduced from Saidel and co-workers *(9)* and Wetlaufer *(10)*. It is also quite similar to values used by GPMAW. In contrast, Buck and co-workers *(15)* assumed that methionine would not contribute significantly in the absorption at 214 nm.

*Group III:* $10\ \% < \varepsilon_{amino\ acid} < 100\ \% \varepsilon_{peptide\ bond}$

Group III comprises cysteine, asparagine, glutamine and arginine. The measured molar extinction coefficient of cysteine was similar to the absorbances deduced from Saidel and co-workers *(9)* and Wetlaufer *(10)*. GPMAW*(16)*, expects a higher contribution for cysteine, which is not in line with other reported values. It might be that GPMAW *(16)* deduced its values from Saidel and co-workers *(9)* but mixed up the value for cysteine with cystine. Alkylation was shown not to have a significant effect on the molar extinction coefficient of α-lactalbumin. This indicates that upon alkylation as often used in peptide analysis, the molar extinction coefficient of alkylated cysteine will remain ~230 $M^{-1}cm^{-1}$ as presented for cysteine in Table 2.

Buck and co-workers *(15)* assumed that glutamine and asparagine both have the same molar extinction coefficient as the peptide bond due to the presence of the amide group. This assumption is to our opinion not valid. Our measured values and the values presented by Saidel and co-workers *(9)* show a much lower molar extinction coefficient of glutamine and asparagine, indicating that the assumption of Buck and co-workers *(15)* is indeed invalid.

*Group IV:* $\varepsilon_{amino\ acid} < 10\% \varepsilon_{peptide\ bond}$

This group contains amino acids that hardly contribute to the absorbance of proteins and peptides when compared to the peptide bond. Proline, when present at the N-terminus, is also present in this group.

**Validation of the molar extinction coefficients of proteins and peptides**

To validate whether the molar extinction coefficient of a protein or peptide can be calculated based on the absorbance of the individual building blocks, the molar extinction coefficients of the various proteins and peptides analyzed were calculated using Formula 2. This calculation is based on the assumption that the molar extinction coefficient of proteins and peptides at 214 nm is only defined by the contribution of the peptide bonds present together with the sum of the contribution of the individual amino acids present.

$$\varepsilon_{\text{protein / peptide}}\ (M^{-1}cm^{-1}) = \varepsilon_{\text{peptide bond}}\ \text{x}\ n_{\text{peptide bonds}} + \sum_{i=1}^{20} \varepsilon_{\text{amino acid(i)}}\ \text{x}\ n_{\text{amino acid(i)}} \qquad [2]$$

In Formula 2 the $\varepsilon_{\text{peptide bond}}$ is 923 $M^{-1}cm^{-1}$ and for $\varepsilon_{\text{amino acid}}$ the values from Table 2 are used, with the only exception that for proline 2,675 $M^{-1}cm^{-1}$ (when not present at the N-terminus) is used. In Figure 2, the calculated molar extinction coefficients are compared with the values measured as presented in Table 4. In addition to this, also the molar extinction coefficients solely based on the absorption of the peptide bonds ($\varepsilon_{\text{peptide bond}}$ x $n_{\text{peptide bonds}}$) are presented. The difference between the measured value and the calculated value only based on the peptide bond illustrates the significant contribution of the amino acids itself to the overall absorbance.

*Validation of the molar extinction coefficients of proteins*

It can be observed that the measured values for glycinin, β-conglycinin, α-lactalbumin and β-lactoglobulin and BSA are quite close to the calculated values (error < 11 %), whereas the errors for α-casein and lysozyme are higher (20 and 25 %, respectively). For the calculation of the molar extinction coefficient of β-conglycinin (containing 2 or 4 GlcNAc per protein molecule *(24, 25)*), the molar extinction coefficient of GlcNAc (experimentally determined to be ~500 $M^{-1}cm^{-1}$) was not taken into account. The results show that the concentration of proteins can be rather well estimated by their absorbance at 214 nm, although the error is larger than when measuring the absorbance at 280 nm. Using this wavelength the error is reported to be below 2 % *(26)*. The differences between calculated and measured values might have several reasons. The presence of impurities might be one of the reasons for the large errors. In this paper all protein and peptide preparations were regarded to be 100 % pure. Another reason for a high error when measuring at 214 nm might be the influence of the protein conformation *(8)*. This is difficult to take into account in the calculation. However, it should be stated that the conformation has a larger effect on the absorbance at lower wavelengths (190-200 nm) than at 214 nm *(8)*.

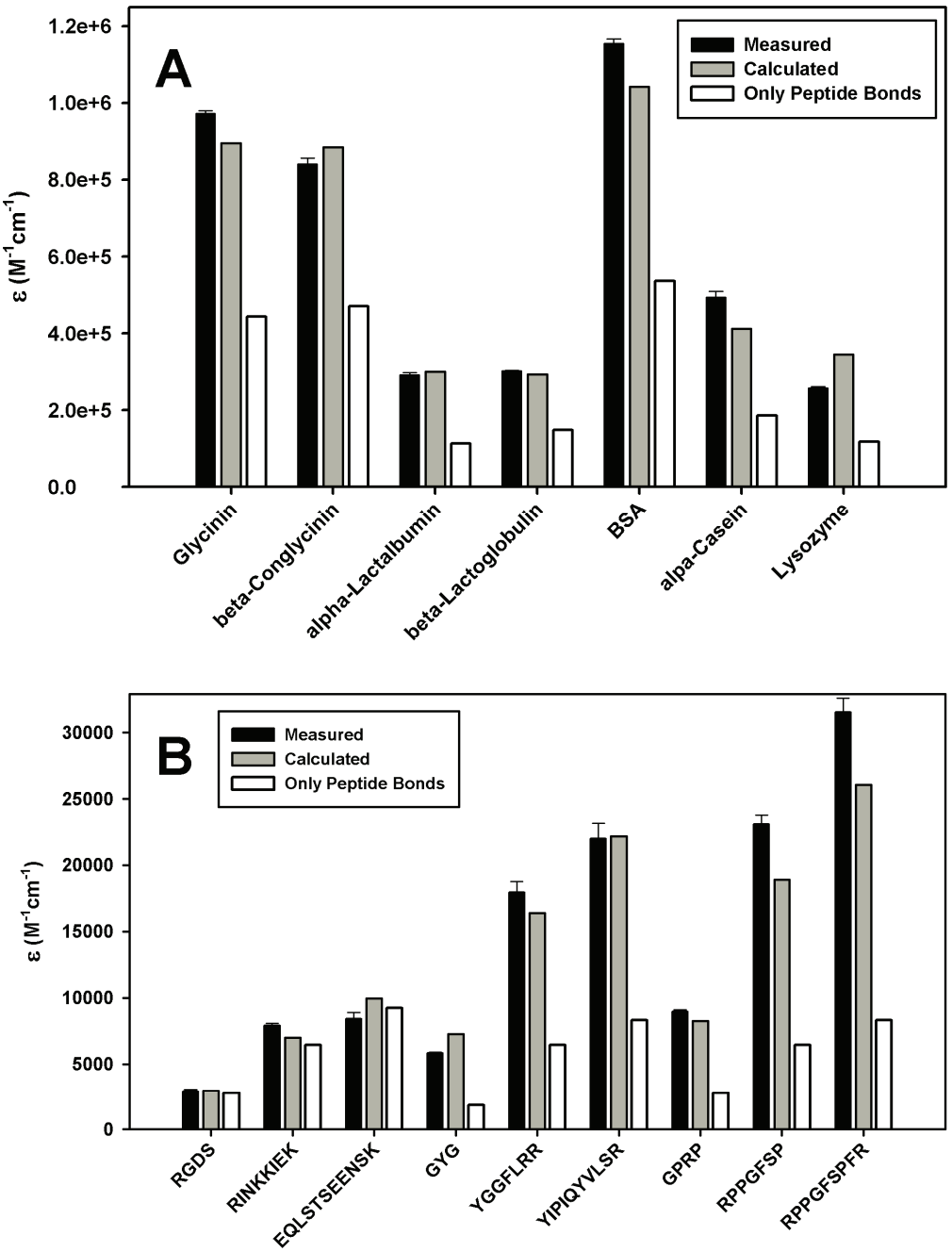


**Figure 2**: Measured and calculated molar extinction coefficients at 214 nm for various proteins (2A) and peptides (2B). Measured values are determined in the average of duplicate measurements at 4 different acetonitrile concentrations.

*Validation of the molar extinction coefficients of peptides*

From Figure 2B it can be concluded that the molar extinction coefficient can be predicted rather well for the different peptides as can be observed by the small difference between the

measured and calculated values. Only the difference between the measured and calculated value of RPPGFSP and RPPGFSPFR is relatively large.
When peptides do not contain amino acids that play a significant role in the absorbance at 214 nm (Group III and IV; RGDS, RINKKIEK, and EQLSTSEENSK) the molar extinction coefficient is only defined by the absorbance of the peptide bond. Therefore, for some peptides it has a linear relation with the length of the peptide *(27)*.
When tyrosine is present in the peptides (GYG, YIPIQYVLSR and YGGFLRR), this clearly results in an increase in molar extinction coefficients and is rather well predictable based on the amino acid composition.
The peptides that contain proline (GPRP) and proline with phenylalanine (RPPGFSP and RPPGFSPFR) clearly show a higher absorbance than only based on their peptide bonds. It shows, taking into account the relative high molar extinction coefficient of phenylalanine, the strong influence of proline on the absorption. However, for RPPGFSP and RPPGFSPFR, there is still a relative large error between the calculated and measured molar extinction coefficient. This might indicate that the contribution of proline to the molar extinction coefficient is not yet fully understood. It might be that when prolines are present next to each other, the absorbance is strengthened as in RPPGFSP(FR), but this aspect was not studied.
The large differences observed in the molar extinction coefficients between peptides with a similar number of peptide bonds (e.g. RINKKIEK and RPPGFSP), illustrates the already mentioned necessity to find a method to estimate peptide concentrations based on their molar extinction coefficients during RP-HPLC analysis.

Table 3 shows that Gly-Gly-Gly has a molar extinction coefficient of 1080 $M^{-1}cm^{-1}$, whereas a value of 1909 $M^{-1}cm^{-1}$ was expected according to Formula 2 and Table 2. A similar observation can be made for the absorbance of GYG, for which a value of 7,263 $M^{-1}cm^{-1}$ was expected and a value of 5,835 $M^{-1}cm^{-1}$ was measured. These results indicate that for these peptides probably the peptide bond does not have a molar extinction coefficient of 923 $M^{-1}cm^{-1}$, but a lower value. However, the molar extinction coefficients for RGDS, RINKKIEK, and EQLSTSEENSK show that a molar extinction coefficient of 923 $M^{-1}cm^{-1}$ for the peptide bond is a good estimation. The reason for the lower molar extinction coefficient for Gly-Gly-Gly can be due to influence of the terminal carboxyl group. If the terminal carboxyl group is charged, this influences the absorption of the peptide bond. The absorbance is lower below the p$K_a$ of the carboxyl group (~ pH 3) than above the p*Ka (11)*. Since this only counts for the C-terminal carboxyl group, its influence will decrease with increasing chain length *(12, 13)*. Also, there is no difference between the measured and calculated molar extinction coefficient of RGDS. This shows that not for all small peptides a lower value is measured when compared to the calculated value.

## Conclusion

It can be concluded that based on the absorbance at 214 nm, and the known amino acid sequence of a peptide, the molar extinction coefficient can be predicted rather well. Therefore, Table 5 presents molar extinction coefficients of protein and peptide building blocks to be used in RP-HPLC-MS quantification. We are aware that due to the complexity of the peptide and protein absorbance at 214 nm there is still an error in the prediction of the molar extinction coefficients. Therefore, more research with the aim to better understand the absorbance of peptides at 214 nm will be needed. Nevertheless, when using the values presented in Table 5 a good estimation can be made of the peptide concentration, preventing large over- or underestimations of peptide amounts present.

**Table 5**: Molar extinction coefficients of protein and peptide building blocks for calculation of the molar extinction coefficients of peptides and proteins.

| Building Block | ε ($M^{-1}cm^{-1}$) | Building Block | ε ($M^{-1}cm^{-1}$) |
|---|---|---|---|
| ***Group I:*** $\varepsilon_{amino\ acid} > \varepsilon_{peptide\ bond}$ | | | |
| Proline (not at N-terminus) (P) | 2,675 | Tyrosine (Y) | 5,375 |
| Histidine (H) | 5,125 | Tryptophan (W) | 29,050 |
| Phenylalanine (F) | 5,200 | | |
| ***Group II:*** $\varepsilon_{amino\ acid} \sim \varepsilon_{peptide\ bond}$ | | | |
| Peptide bond | 923 | Methionine (M) | 980 |
| ***Group III:*** $10\ \% < \varepsilon_{amino\ acid} < 100\ \% \ \varepsilon_{peptide\ bond}$ | | | |
| Arginine (R) | 102 | Glutamine (Q) | 142 |
| Asparagine (N) | 136 | Cysteine (C) | 225 |
| ***Group IV:*** $\varepsilon_{amino\ acid} < 10\% \ \varepsilon_{peptide\ bond}$ | | | |
| Glycine (G) | 21 | Valine (V) | 43 |
| Proline (at N-terminus) (P) | 30 | Isoleucine (I) | 45 |
| Alanine (A) | 32 | Leucine (L) | 45 |
| Serine (S) | 34 | Aspartic acid (D) | 58 |
| Lysine (K) | 41 | Glutamic acid (E) | 78 |
| Threonine (T) | 41 | | |

## ACKNOWLEDGEMENT

We thank Dr. Ben van den Broek of AFSG, Wageningen, for kindly supplying the peptides RINKKIEK, EQLSTSEENSK, and YIPIQYVLSR.

## REFERENCES

1. Kitts, D.D.; Weiler, K. Bioactive proteins and peptides from food sources. Applications of bioprocesses used in isolation and recovery. *Curr. Pharm. Des.* **2003,** *9*, (16), 1309-1323.
2. Korhonen, H.; Pihlanto, A. Bioactive peptides: Production and functionality. *Int. Dairy J.* **2006,** *16*, (9), 945-960.

3. Korhonen, H.; Pihlanto, A. Food-derived bioactive peptides - Opportunities for designing future foods. *Curr. Pharm. Des.* **2003,** *9*, (16), 1297-1308.
4. Li, G.H.; Le, G.W.; Shi, Y.H.; Shrestha, S. Angiotensin I-converting enzyme inhibitory peptides derived from food proteins and their physiological and pharmacological effects. *Nutr. Res.* **2004,** *24*, (7), 469-486.
5. Kodera, T.; Nio, N. Identification of an angiotensin I-converting enzyme inhibitory peptides from protein hydrolysates by a soybean protease and the antihypertensive effects of hydrolysates in spontaneously hypertensive model rats. *J. Food Sci.* **2006,** *71*, (3), C164-C173.
6. Caessens, P.; Daamen, W.F.; Gruppen, H.; Visser, S.; Voragen, A.G.J. β-Lactoglobulin hydrolysis. 2. Peptide identification, SH/SS exchange, and functional properties of hydrolysate fractions formed by the action of plasmin. *J. Agric. Food Chem.* **1999,** *47*, (8), 2980-2990.
7. Cantor, C.R.; Schimmel, P.R. *Biophysical Chemistry Part II: Techniques for the study of biological structure and function*. W.H. Freeman and company: New York, 1980.
8. Rosenheck, K.; Doty, P. Far ultraviolet absorption spectra of polypeptide and protein solutions and their dependence on conformation. *Proc. Natl. Acad. Sci. U. S. A.* **1961,** *47*, (11), 1775-1785.
9. Saidel, L.J.; Goldfarb, A.R.; Waldman, S. The absorption spectra of amino acids in the region 200 to 230 millimicrons. *J. Biol. Chem.* **1952,** *197*, (1), 285-291.
10. Wetlaufer, D.B. Ultraviolet spectra of proteins and amino acids. *Adv. Protein Chem.* **1962,** *17*, 303-390.
11. Goldfarb, A.R. Absorption spectrum of the peptide bond. 2. Influence of chain length. *J. Biol. Chem.* **1953,** *201*, (1), 317-320.
12. Saidel, L.J. Ultraviolet absorption spectra of peptides. 2. Glycine peptides containing more than one substituted amide link. *Arch. Biochem. Biophys.* **1955,** *56*, (1), 45-58.
13. Saidel, L.J.; Lieberman, H. Ultraviolet absorption spectra of peptides. 4. Alanine residue. *Arch. Biochem. Biophys.* **1958,** *76*, (2), 401-409.
14. Moffatt, F.; Senkans, P.; Ricketts, D. Approaches towards the quantitative analysis of peptides and proteins by reversed-phase high-performance liquid chromatography in the absence of a pure reference sample. *J. Chromatogr., A* **2000,** *891*, (2), 235-242.
15. Buck, M.A.; Olah, T.A.; Weitzmann, C.J.; Cooperman, B.S. Protein estimation by the product of integrated peak area and flow-rate. *Anal. Biochem.* **1989,** *182*, (2), 295-299.
16. Lighthouse-Data *GPMAW 7.01;* www.gpmaw.com, 2005.
17. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophys.* **2006,** *1*, 178-188. *Chapter 3 in this thesis*.
18. Garcia, M.C. The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry. *J. Chromatogr., B* **2005,** *825*, (2), 111-123.
19. Hernandez-Ledesma, B.; Ramos, M.; Recio, I.; Amigo, L. Effect of β-lactoglobulin hydrolysis with thermolysin under denaturing temperatures on the release of bioactive peptides. *J. Chromatogr., A* **2006,** *1116*, (1-2), 31-37.
20. Van der Ven, C.; Gruppen, H.; de Bont, D.B.A.; Voragen, A.G.J. Reversed phase and size exclusion chromatography of milk protein hydrolysates: relation between elution from reversed phase column and apparent molecular weight distribution. *Int. Dairy J.* **2001,** *11*, (1-2), 83-92.
21. Biemann, K. Mass-spectrometry of peptides and proteins. *Annu. Rev. Biochem.* **1992,** *61*, 977-1010.

22. Sechi, S.; Chait, B.T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **1998,** *70*, (24), 5150-5158.
23. Doty, P.; Imahori, K.; Klemperer, E. The solution properties and configurations of a polyampholytic polypeptide - copoly-L-lysine-L-glutamic acid. *Proc. Natl. Acad. Sci. U. S. A.* **1958,** *44*, (5), 424-431.
24. Thanh, V.H.; Shibasaki, K. β-Conglycinin from soybean proteins - Isolation and immunological and physicochemical properties of monomeric forms. *Biochim. Biophys. Acta* **1977,** *490*, (2), 370-384.
25. Yamauchi, F.; Yamagishi, T. Carbohydrate sequence of a soybean-7S protein. *Agric. Biol. Chem.* **1979,** *43*, (3), 505-510.
26. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption-coefficient of a protein. *Protein Sci.* **1995,** *4*, (11), 2411-2423.
27. Becklin, R.R.; Desiderio, D.M. The amount of ultraviolet absorbency in a synthetic peptide is directly proportional to its number of peptide-bonds. *Anal. Lett.* **1995,** *28*, (12), 2175-2190.

# CHAPTER 6

# Functional region identification in proteins by accumulative-quantitative-peptide-mapping using RP-HPLC-MS

**ABSTRACT**

A new method was developed to identify regions in proteins from which peptides are derived with specific functional properties. This method is applicable for systems in which peptides of a hydrolyzed protein posses specific functional properties, but are too large to be sequenced directly and/or that the peptide mixture is too complex to purify and characterize each peptide individually. In the present work, aggregating peptides obtained by proteolytic hydrolysis of soy glycinin were used as a case study. The aggregating peptides are isolated and subsequently further degraded with trypsin to result in peptides with a mass < 5,000 Da in order to enable sequence identification using RP-HPLC-MS in combination with MS/MS. Prior to RP-HPLC the peptides are fractionated using anion and cation exchange chromatography. The fractions obtained are analyzed with RP-HPLC-MS. The peptides, with identified sequences, were quantified using the peak areas of the RP-HPLC chromatograms measured at 214 nm. Next, the peak areas were corrected for the molar extinction coefficient of the individual peptides, followed by accumulative-quantitative-peptide-mapping. The results show that in complex systems, based on the method described, the regions in the parental protein from which the functional peptides originate can be properly identified.

Kuipers, B.J.H.; Bakx, E.J.; Gruppen, H. *Submitted.*

## INTRODUCTION

In literature several studies can be found on the ability of proteases to yield peptides that have strong aggregating properties *(1-3)*. Upon limited hydrolysis of soy glycinin we have shown previously *(4, 5)* that various proteases, having a broad specificity, are able to form a wide range of peptides with strong aggregating properties. These peptides have molecular masses ranging from ~5,000 to 25,000 Da. Due to their large molecular masses it can be assumed that many different peptides are present, which largely overlap in sequence.

Knowing the sequences of the aggregating peptides in a protein hydrolysate is of prime interest to understand the aggregation mechanism. The easiest method to analyze the sequences of peptides in a protein hydrolysate is by direct analysis on RP-HPLC-MS. However, identification of peptide sequences using ESI Iontrap MS in combination with Tandem Mass Spectrometry (MS/MS) *(6)* requires peptide masses below ~ 4,000 Da for positive sequence identification. If the molecular masses of the peptides of interest are higher, as is the case for the aggregating peptides from glycinin, a different method has to be used. The most obvious method is first to isolate all aggregating peptides individually, followed by quantification of the individual peptide concentration. Methods for separation might be RP-HPLC or 2-D electrophoresis. For determination of the peptide sequences, each individual peptide fraction can be studied by N-terminal sequencing followed by determination of the total mass by using, e.g., Maldi-TOF *(7-10)*. However, if the peptides of interest comprise a wide range of peptides which overlap in sequence, accurate separation of the individual peptides is not feasible.

Therefore, if functional peptides have masses above ~5,000 Da and largely overlap in sequence, a new method has to be used. Such a method is described in this paper. This method, denoted *accumulative-quantitative-peptide-mapping*, does not have the aim to identify the sequence of the peptides of interest, but aims to identify those regions in the parental protein from which the functional peptides originate. This method is schematically presented in Figure 1.

The functional peptides (e.g. peptides with strong aggregating properties) are first isolated from peptides not contributing to the functionality of interest. To decrease the peptide mass, enabling positive sequence identification with RP-HPLC-MS, subsequently tryptic digestion is performed. Because of the expected overlap in sequence, tryptic digestion of the overlapping sequences from different functional peptides mainly results in the accumulation of identical peptides from the overlapping regions. The tryptic digest is fractionated (e.g. RP-HPLC) followed by identification (MS/MS), quantification (absorbance at 214 nm *(11)*) and peptide mapping projected towards the sequence of the parental protein. The amino acids in the sequence of the parental protein, for which the strongest accumulation takes place, reveal those regions in the parental protein from which the functional peptides mainly originate. These regions can be defined as the major functional peptide region of the protein (in Figure 1: around the amino acids 150 to 200).

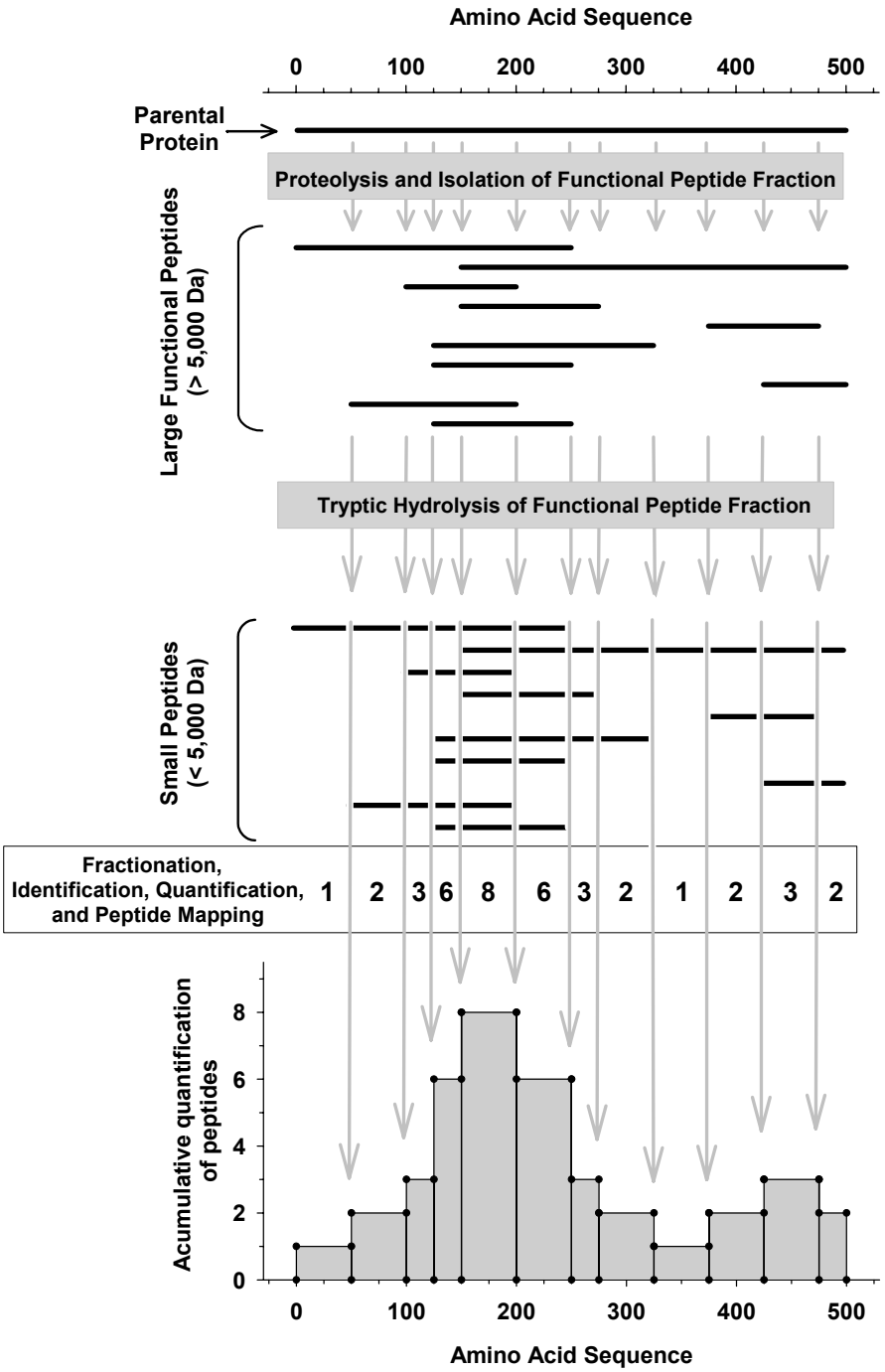


**Figure 1**: Identification of the functional regions of proteins by *accumulative-quantitative-peptide-mapping*. The parental protein is selectively degraded into "large functional peptides" that are isolated in a functional peptide fraction based on a specific functional property, followed by a second, complete degradation of the peptides in the functional peptide fraction by trypsin. Each peptide is identified and quantified. Next, all peptides are mapped onto the amino acid sequence of the parental protein, taking into account their abundance. Those regions in the protein sequence with the highest accumulation per amino acid reveal the functional regions in the parental protein from which the peptides of interest originate. The grey arrows (↓) indicate the position at which the enzyme can cleave. To simplify the scheme, the specificity of the first proteolytic digestion is taken to be the same as trypsin.

The aim of the present study was to develop a quantitative method to identify those regions in a protein from which the peptides predominantly originate that have specific functional properties. In this study aggregating peptides from soy glycinin are used as an example. It is envisaged that this method can also be applied to other function peptide mixtures, containing peptides being too large to be analyzed directly by RP-HPLC-MS, and too complex to be separated into its individual peptides.

## MATERIALS AND METHODS

### Materials and chemicals

Glycinin was prepared from defatted SBM (from Hyland soybeans) by precipitation of the pH 8.0 soluble proteins at pH 6.2 using 1 M HCl, as described previously *(5)*. Glycinin has a protein content of 87 % (w/w) (N x 5.57) as determined by the Dumas method and a purity > 95 % as determined by densitometric analysis of the SDS-PAGE gel (no data shown). Chymotrypsin (TLCK treated), trypsin (TPCK treated), and bradykinin were obtained from Sigma Chemical Co. (St. Louis, MO, USA; article numbers C-3142, T-1426, and B-3259, respectively). The eluents used for RP-HPLC-MS were all of HPLC grade. All other chemicals were of analytical grade and were purchased from Merck (Darmstadt, Germany) or Sigma (St. Louis, MO, USA).

### Analysis of the protein content

The nitrogen contents of the various samples were determined in duplicate by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milano, Italy) according to the manufacturer's instructions. Methionine was used as a standard. The nitrogen conversion factor for purified glycinin is 5.57, based on the average amino acid compositions of the 5 different subunits of glycinin as found in the Swiss-Prot database (www.expasy.org; Primary accession numbers used were: P04776 (Glycinin 1; A1a and B2 polypeptides), P04405 (Glycinin 2; A2 and B1a polypeptides), P11828 (Glycinin 3; A1b and B1b polypeptides), P02858 (Glycinin 4: A5, A4 and B3 polypeptides), and P04347 (Glycinin 5; A3 and B4 polypeptides)). Equal molar abundance of the subunits was assumed.

### Hydrolysis of glycinin

Glycinin was suspended in Millipore water at a concentration of 1.5 - 2 % (w/w) protein and stirred overnight at 4°C. The pH was adjusted to 8.0, if necessary. The glycinin suspension was filtered using a 0.45 μm filter (Schleicher & Schuell GmbH, Dassel, Germany). The protein concentration of the filtrate was determined using the Dumas method. The solution was diluted to 1 % (w/w) protein using Millipore water followed by heating at 95°C for 30 minutes. The heated solution had a clear appearance. The heated glycinin solution (150 mL) was hydrolyzed by chymotrypsin up to a degree of hydrolysis (DH) of 2.2 % at pH 8.0 at

40°C. The pH and DH were controlled using the pH-stat method by using a 719S Titrino (Metrohm ion analysis, Herisau, Switzerland). The $h_{tot}$ used for the calculation was 8.85 meqv/g, calculated based on the average amino acid composition of the glycinin subunits. The enzyme : protein ratio was 1 : 350 (w/w). The chymotrypsin was dissolved in Millipore water and directly added to the glycinin solutions. The molarity of the NaOH solution used to maintain the pH at 8.0 was 0.1 M. When the desired DH was reached, the enzymatic hydrolysis was stopped by addition of a 100 mM phenylmethylsulphonyl fluoride (PMSF) solution in 2-propanol to a final concentration of 1 mM. The pH-stat experiment was stopped when the pH remained stable at pH 8.0. After cooling down to room temperature, Millipore water and NaCl (from a 2 M stock solution) were added to end up with a protein concentration of 0.8 % (w/w) and an ionic strength of $I$ = 0.03 M, as described previously *(5)*. The pH was adjusted to 8.0 if necessary. The addition of NaCl induces aggregation. After 1 hr stirring, a sample was taken, whereas the rest was centrifuged (5 min, 5,000 x *g, 20°C*). Analysis of the protein content of the sample before centrifugation and the supernatant after centrifugation showed that ~43 % of the proteinacious material was aggregated. The pellet was washed by re-suspending it in Millipore water containing 0.03 M NaCl (~ 1 : 7 ratio of pellet : water) and centrifuged again (5 min, 5,000 x *g, 20°C*). This washing step was repeated twice. During the washing the pH of the suspension remained 8.0. After washing, the pH of the pellet was decreased to ~2.5 using trifluoroacetic acid to resolubilize the peptides, followed by freeze drying. These freeze dried peptides are further denoted aggregating peptides.

**Tryptic digestion of aggregating peptides**

Aggregating peptides (20 mg) were dissolved in 600 µL of 20 mM Tris-HCl buffer (pH 8.0) containing 8 M urea and 15 mM DTT. The solution was kept for 30 minutes at 60°C. Iodoacetamide (IAA) was added up to 30 mM and incubation took place in the dark for 30 minutes to alkylate the peptides. Every 10 minutes the pH was verified and adjusted to 8.0 if necessary using 1 M NaOH. The samples were diluted to 1.6 M urea by the addition of 20 mM Tris-HCl buffer (pH 8.0). This resulted in partial aggregation of the peptides as observed by an increase in turbidity. The alkylated aggregating peptides were degraded by trypsin in 3 cycles. First, trypsin was added in an enzyme protein ratio (w/w) 1 : 50 followed by incubation at 37°C for 4 hrs. The hydrolyzed sample was centrifuged (15 min, 24,000 x *g*, 20°C). The pellet was re-suspended in 20 mM Tris-HCl buffer (pH 8.0) containing 8 M urea. The solution obtained was diluted to 1.6 M urea by the addition of 20 mM Tris-HCl buffer (pH 8.0). Trypsin (0.1 mg) was added to the re-suspended pellet and incubated overnight at 37°C in the presence of 0.02 % (w/v) sodium azide. After incubation, the sample was treated the same way as described above, by re-suspending in 8 M urea, dilution to 1.6 M urea and incubation for 2 hrs with 0.1 mg trypsin at 37°C in the presence of 0.02 % (w/v) sodium azide. The hydrolysate was mixed with the supernatants of the first and second tryptic

hydrolysis cycle, followed by centrifugation (15 min, 24,000 x *g*, 20°C) prior to fractionation using ion exchange chromatography.
In addition to the aggregating peptide fraction, also the total hydrolysate after chymotryptic hydrolysis, containing the aggregating as well as the non-aggregating peptides, was hydrolyzed in 3 cycles with trypsin as described above.

**Ion exchange chromatography (IEX)**

An aliquot (1.4 mL) of the supernatant of the tryptic digest of the aggregating peptides was diluted with 0.8 mL Millipore water and filtered through a 0.45 µm filter (Schleicher & Schuell, Dassel, Germany). Next, 2 mL of the filtered peptide solution was applied onto a 1 mL Mono Q 5/50GL anion exchange column (Amersham Biosciences, Uppsala, Sweden). Solvent A (20 mM Tris-HCl buffer, pH 8.0) and solvent B (20 mM Tris-HCl buffer, pH 8.0 containing 2 M NaCl) formed the eluent in the following consecutive steps: 3 mL isocratic elution at 100 % A; 10 mL sample injection; linearly to 50 % B in 60 mL; linearly to 100 % B in 15 mL; 5 mL isocratic elution at 100 % B; linearly to 100 % A in 3 mL; 12 mL isocratic at 100 % A. The flow rate was 2 mL/min, except during the injection stage at which the flow rate was 1 mL/min. Fractions of 1 mL were collected. The eluent was monitored at 214 nm. The non-bound fractions (fractions 1-12) were pooled and applied on a cation exchange column as described below.
The pH of the non-bound fractions on the anion exchanger was decreased to pH 3.5 using formic acid. Millipore water was added up to 23 mL and filtered through a 0.45 µm filter (Schleicher & Schuell, Dassel, Germany). Only 20 mL (4 times 5 mL) of the 23 mL unbound peptide fraction was applied on a 1 mL cation exchanger, Mono S 5/50GL (Amersham Biosciences, Uppsala, Sweden). Solvent A (20 mM formic acid) and solvent B (20 mM formic acid containing 2 M NaCl) formed the eluent in the following subsequent steps: 3 mL isocratic elution at 100 % A; 38 mL sample injection (every ~7 mL, a 5 mL sample was injected); linearly to 30 % B in 15 mL; linearly to 100 % B in 17 mL; 10 mL isocratic elution at 100 % B; linearly to 100 % A in 3 mL; 12 mL isocratic at 100 % A. The flow rate was 2 mL/min, except during the injection stage where the flow rate was 1 mL/min. Fractions of 1 mL were collected. The eluate was monitored at 214 nm.

**RP-HPLC-MS**

Of all bound fractions collected from anion (25 fractions) and cation (27 fractions) exchange chromatography, 300 µL was taken. Next, to all fractions 16.5µL acetonitrile was added, followed by the addition of 14 µL of 5 % (v/v) and 3.2 % (v/v) formic acid for the anion and cation exchange chromatography fractions, respectively. All unbound fractions (38 mL) from the cation exchange column were pooled and 6 mL was freeze dried and subsequently dissolved in 300 µL Millipore water followed by the addition of 16.5µL acetonitrile and 14 µL 3.2 % (v/v) formic acid. Next to the bound and unbound samples, a sample of the tryptic

digest of the total chymotryptic hydrolysate was diluted 8 times with a solution of 5 % (v/v) acetonitrile in water containing 0.1 % (v/v) formic acid.
After 1 hr mixing (head-over-tail) all samples were centrifuged (15 min, 24.000 x *g*, 20°C) and 100 µL was injected onto a reversed phase C18 column (218MS52; 250 x 2.1 mm 5µm) (Grace Vydac, Hesperia, CA, USA), installed on a Spectra System HPLC (Thermo Separation Products, Fremont, CA, USA). A flow rate of 0.2 mL/min was used. The solvents used were Millipore water containing 0.1 % (v/v) formic acid (Solvent A) and acetonitrile containing 0.085 % (v/v) formic acid (Solvent B). The elution used was as follows: from 0 to 10 min, 5 % B (isocratic); linear gradient up to 45 % B till 80 min; linear gradient up to 95 % B till 90 min; isocratic elution with 95 % B for 5 minutes; linear gradient to 5 % B till 96 min followed by isocratic elution with 5 % B till 110 minutes. All HPLC operations and data processing were controlled by Chromeleon version 6.7 software (Dionex Corporation, Sunnyvale, USA).
After UV detection at 214 nm, a flow splitter (type Acurate, LC Packings, Amsterdam, The Netherlands) decreased the flow to 0.05 mL/min that was applied on an LCQ Deca XP Max (Thermo Finnigan, San Jose, CA, USA) with the use of electrospray ionization and detection in the positive ion mode. The capillary spray voltage was 4 kV, and the capillary temperature was 200°C. Before use the instrument was tuned with a 0.01 mg/mL bradykinin solution. The instrument was controlled by Xcalibur software v1.3 (Thermo Finnigan, San Jose, CA, USA). The scan range was set from *m*/*z* 400 to 2000. The MS/MS functions were performed in data dependent mode. The collision-energy value was 35 %. BioWorks software, version 3.1 (Thermo Electron, San Jose, CA, USA) was used for automatic sequencing and database search for the sequences in a database only containing sequences of glycinin G1 to G5 and trypsin and chymotrypsin, as present in the Swiss-Prot database. In the database search the possible modifications of the proteins, oxidation of methionine (+15.99 Da) and alkylation of cysteine (+57.05 Da), were included. To discriminate between correct and incorrect peptide sequence assignment the cross correlation value (Xcorr) for each identified peptide was used as a criterion *(12, 13)*. For positive identification of the peptides a Xcorr threshold of 1.5 for single charged peptides, 2.0 for double charged peptides and 2.5 for triple charged peptides was used.

**Quantification of peptides based on the peak area at 214 nm**

The molar extinction coefficient for each of the identified peptides was calculated using Formula 1 as described previously *(11)*:

$$\varepsilon_{\text{peptide}}\ (\text{M}^{-1}\text{cm}^{-1}) = \varepsilon_{\text{peptide bond}}\ \text{x}\ \text{n}_{\text{peptide bonds}} + \sum_{\text{i=1}}^{20} \varepsilon_{\text{amino acid(i)}}\ \text{x}\ \text{n}_{\text{amino acid(i)}} \qquad [1]$$

The molar extinction coefficients used for the peptide building blocks (individual amino acid contribution and the peptide bond) are presented in Table 1. Since the peak area is used as a

measure for the absorbance and the light path is not defined, but similar in all analysis, the concentration is not expressed as molarity, but as molar equivalents (Meqv).

**Table 1**: Molar extinction coefficients ($M^{-1}cm^{-1}$) of protein and peptide building blocks for calculation of the molar extinction coefficient in the peptide chain at 214 nm *(11)*.

| Building block | ε ($M^{-1}cm^{-1}$) | Building block | ε ($M^{-1}cm^{-1}$) |
|---|---|---|---|
| Peptide bond | 923 | Leucine | 45 |
| Alanine | 32 | Lysine | 41 |
| Arginine | 102 | Methionine | 980 |
| Asparagine | 136 | Phenylalanine | 5,200 |
| Aspartic acid | 58 | Proline (at N-terminus) | 30 |
| Cysteine | 225 | Proline (not at N-terminus) | 2,675 |
| Glutamine | 142 | Serine | 34 |
| Glutamic acid | 78 | Threonine | 41 |
| Glycine | 21 | Tryptophan | 29,050 |
| Histidine | 5,125 | Tyrosine | 5,375 |
| Isoleucine | 45 | Valine | 43 |

To enable comparison of all RP-HPLC chromatograms in a quantitative way, the area of the peaks from the cation IEX were proportionally corrected for the amount of unbound fraction that was not applied onto the cation exchange column. In addition to that, the area of the peaks in the unbound fraction of cation IEX was multiplied by 1.9 to correct for the dilution step of this fraction. The analysis of the 53 RP-HPLC analysis revealed that, based on the peak areas, less than 3 % of all the peptides were present in the unbound fraction after cation IEX.

### High-performance size-exclusion-chromatography (HP-SEC)

Glycinin, its DH 2.2 hydrolysate, the aggregating peptide fraction thereof and the tryptic hydrolysate of the aggregating peptides were analyzed with size exclusion chromatography. Samples were dissolved under denaturing (8 M guanidiniumchloride) and reducing conditions (100 mM DTT), and separated with a Shodex Protein KW-803 column (300 x 8 mm; Showa Denko K.K., Tokyo, Japan) equilibrated and eluted with 6 M urea containing 30 % (v/v) acetonitrile and 0.1 % (v/v) TFA. The flow rate was 0.5 mL/min and the eluate was monitored at 214 nm. The final protein concentration was ~3 mg/mL. In order to facilitate comparison of the different samples, the total area under the chromatograms was corrected for the proportion of peptides that was aggregated. The column was calibrated using proteins with known molecular masses as described previously *(5)*.

## RESULTS AND DISCUSSION

### Monitoring the hydrolysis

To monitor the protein degradation of the sequential hydrolysis of glycinin with chymotrypsin, followed by the hydrolysis of the aggregating peptide fraction with trypsin, HP-SEC was performed. It can be observed (Figure 2) that at a DH of 2.2 % there is probably no more intact protein present. Table 2 shows the molecular mass range of the proteins and peptides present in the various analyzed fractions. The chromatogram of the intact glycinin polypeptides shows some aggregates and/or β-conglycinin eluting before 7 mL (> 40,000 Da). The polypeptides eluting around 7.4 mL (25,000 to 40,000 Da) represent the acidic polypeptides A1 to A4, polypeptides eluting around 8.2 mL (15,000 to 25,000 Da) represent all the basic polypeptides, and the peak eluting around 9 mL (8,000 to 15,000 Da) represents the A5 acidic polypeptide of glycinin *(14)*. The aggregating peptide fraction mainly contains peptides with masses in the range of ~5,000 to 25,000 Da (Table 2). These are reduced to peptides with masses < 5,000 Da upon tryptic hydrolysis. Prior to tryptic digestion, the aggregating peptides were (partly) present as aggregates in the hydrolysis buffer. In Figure 2 it can be observed that the aggregating peptide fraction was degraded into smaller fragments, indicating that the aggregation did not prevent tryptic degradation of the peptides. After the repetitive digestion with trypsin there were still some aggregates present in the hydrolysate. Nevertheless, HP-SEC of the total tryptic hydrolysate and its supernatant, revealed that > 90 % of aggregating peptides was soluble after tryptic degradation (results not shown).

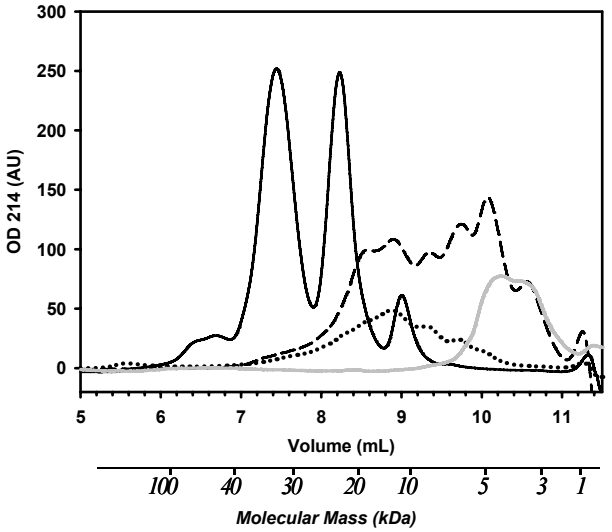


**Figure 2**: HP-SEC chromatograms under denaturing conditions of glycinin (——), glycinin hydrolysate at DH 2.2 % (— —), the aggregating peptides (• • • •) and the tryptic digest of aggregating peptides (——). Samples were denatured and reduced prior to analysis.

**Table 2**: Molecular mass ranges of glycinin and glycinin-derived peptides.

| **Sample** | **Molecular mass range (Da)** |
|---|---|
| Intact glycinin polypeptides | 8,000 - 40,000 |
| Glycinin hydrolysate at DH = 2.2 % | < 25,000 |
| Aggregating peptides of glycinin at DH = 2.2 % | 5,000 - 25,000 |
| Tryptic digest of aggregating peptides | < 5,000 |

**Peptide fractionation**

Figure 3 shows the ion exchange chromatograms of the tryptic digest of the aggregating peptides. Anion exchange chromatography yielded 25 fractions, which eluted after applying the NaCl gradient (%B). The unbound fraction of the anion exchange fractionation (3A) was applied on the cation exchange column (3B).

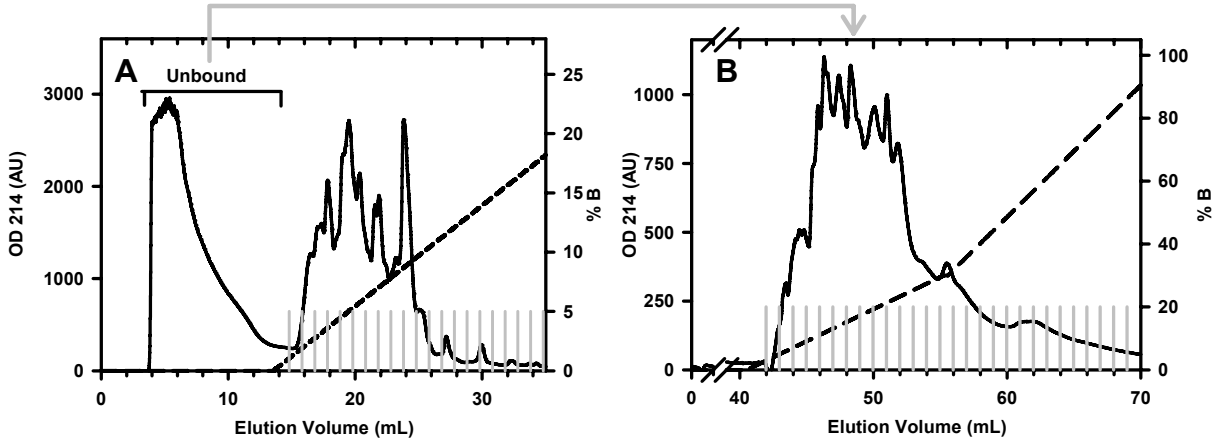


**Figure 3:** Anion exchange chromatogram (**A**) of tryptic digest and cation exchange chromatogram (**B**) of the unbound fraction from anion exchange chromatography. Grey lines represent the collected fractions. The dashed line represents the 2 M NaCl gradient (%B).

The 52 bound fractions and the unbound fraction of cation exchange chromatography and the tryptic hydrolysate of the total chymotryptic hydrolysate were all analyzed on RP-HPLC-MS. The NaCl gradient in the anion and cation chromatography were chosen in such a way to result in fractions that, when separated on RP-HPLC, give a separation more or less identical to the RP-HPLC chromatogram shown in Figure 4. All 53 chromatograms from the RP-HPLC analysis were integrated. All peaks with a peak area smaller than 0.4 mAU*min were rejected since this peak area was too close to the background noise. In total 682 peaks were integrated with a combined peak area of 3,861mAU*min.

**Identification and quantification of a peptide**

Figure 4 shows a typical example of how the RP-HPLC analysis was combined with MS to assign peptide sequences to the peaks in RP-HPLC chromatograms. The RP-HPLC chromatogram in Figure 4A shows the separation of the peptides in a particular ion exchange chromatography column. The number of peaks observed is average compared to other chromatograms. This means that also chromatograms were present with more peaks, but also chromatograms with a lower number of peaks. The presented base peak MS spectrum (Figure 4B) shows a high similarity to the UV214 pattern in the RP-HPLC spectrum, taking into account the elution delay between the RP-HPLC and the MS of approximately 0.25 minutes. In Figure 4C the mass scan at 41.77 minutes is presented.

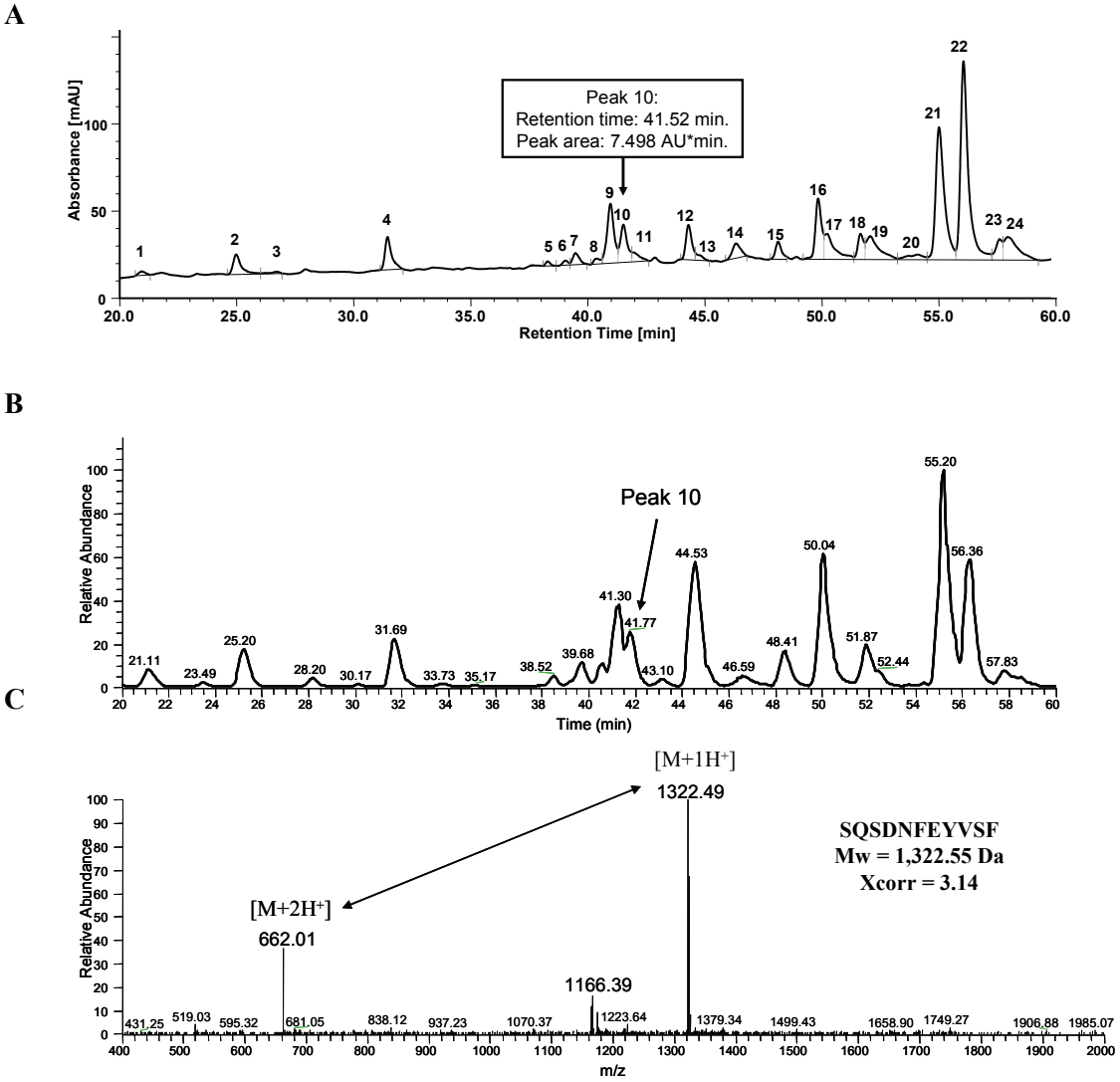


**Figure 4**: RP-HPLC-UV graph at 214 nm of an anion exchange fraction (**4A**), its base-peak MS spectrum (**4B**), and the mass scan at 41.77 minutes (**4C**).

Based on one example of a RP-HPLC peak, in 5 steps it will be explained how the data that belong to this peak are used for the peptide mapping, which is described below.

*Step 1: Peak/Peptide selection:* For this particular example peak 10 (as indicted with the arrow) in the RP-HPLC chromatogram (Figure 4A) with its maximum at 41.52 minutes is taken. This peak corresponds to the top of the peak at 41.77 minutes in the base-peak MS spectrum (Figure 4B). Figure 4C shows the full MS scan at 41.77 minutes. It can be seen that the spectrum of a peptide with an m/z of 1322.49 is the most dominant one. The MS/MS spectra of the minor peak with an m/z of 662.01 (+2) was similar to the MS/MS spectrum of the m/z 1322.49 (+1) peak. This confirms that both peaks belong to the same peptide, with a difference in charge. The other peaks present in Figure 4C, around an m/z of 1166.39 belong

to the peptide that is responsible for peak 9 in Figure 4A, but is still slightly present in peak 10 due to overlap of both peaks (no further data shown).

*Step 2: Peptide identification:* Using BioWorks software and database searching (glycinin sequences as presented in Swiss-Prot), the sequence of the peptide in peak 10 was determined to be SQSDNFEYVSF, based on the MS/MS data from a double charged ion (m/z = 662.01), with an Xcorr of 3.14. As this Xcorr is well above the threshold value of 2.0, the sequence was accepted.

*Step 3: Peptide location*: Next, it is determined in which subunit of glycinin this sequence is present. Glycinin can be present in 5 subunits, Glycinin 1 to 5 *(15)*. For peptide mapping, the contribution of one peptide to the peptide map has to be corrected for the molar proportion in which the different glycinin subunits, containing the identified peptide, are present. The 5 subunits can be divided into two subfamilies that are designated as Group-I (Glycinin 1, 2, and 3) and Group-II (Glycinin 4, and G5) *(16, 17)*. Homologies between members of the same group range from 80 to 90 %, and between the groups the homology is less than 50 % *(15)*. The database search towards the protein sequences of glycinin revealed that the sequence SQSDNFEYVSF is present in the basic polypeptides of glycinin G1 [f405-415], glycinin G2 [f396-406], and glycinin G3 [f389-399].

*Step 4: Peptide quantification per glycinin subunit*: Based on the amino acid sequence determined in step 2, the molar extinction coefficient was calculated using Formula 1 and the data presented in Table 1, resulting in a $\varepsilon_{SQSDNFEYVSF}$ = 25,564 $M^{-1}cm^{-1}$. The area of peak 10 was 7.498 AU*min indicating that the total concentration of the peptide with sequence SQSDNFEYVSF is 0.293 mMeqv using Formula 2.

$$c_{peptide}\,(\text{Meqv}) \sim \frac{A_{peak}}{\varepsilon_{peptide}} \qquad [2]$$

If a peptide was present in more than one subunit, the subunits in which this peptide was present were all a member of the same group. To calculate the concentration of one peptide in the peptide map of one subunit, its concentration has to be corrected for the molar proportion in which the subunits containing the sequenced peptide are present ($F_{subunit}$). $F_{subunit}$ can be calculated by dividing the molarity of the subunit in which a peptide is present by the sum of the molarities of all the subunits in which the peptide occur. In this study, for simplicity reasons, due to the high sequence homology within Group I and Group II, it was assumed that within each group the subunits were present in similar molar ratios.

In addition to the correction for the presence of a peptide in more than one subunit, it might occur that different peptides constitute one peak in the RP-HPLC chromatogram. This was the case in 30 of the 185 peaks. In these cases it was assumed that each peptide contributes to the absorbance in an equal amount. The correction was done by dividing the peak area by the number of peptides present in one RP-HPLC peak ($n_{peptides\ in\ peak}$).

Formula 3 shows the calculation of the relative concentration of a peptide in one subunit, which is similar to Formula 2 but now extended with the correction factors $F_{subunit}$ and $n_{peptides\ in\ peak}$.

$$c_{peptide\ in\ 1\ subunit}\,(\text{Meqv}) \sim \frac{A_{peak}}{\varepsilon_{peptide}} \times \frac{1}{n_{peptides\ in\ peak}} \times F_{subunit} \qquad [3]$$

Since SQSDNFEYVSF was the only peptide responsible for the absorbance of peak 10 and is present in 3 glycinin subunits, for $F_{subunit}$ and $n_{peptides\ in\ peak}$, 0.333 and 1 are used, respectively. This results in a relative concentration of SQSDNFEYVSF in one subunit of 0.098 mMeqv (0.293/3). Consequently, in the peptide mapping of an individual subunit, for each amino acid in this sequence, a value of 0.098 mMeqv is assigned.

*Step 5: Peptide mapping:* Figure 5 shows the peptide map for Glycinin 1, only containing the relative concentration of SQSDNFEYVSF of 0.098 mMeqv as calculated above. The grey regions in the peptide map represent the part of the Glycinin 1 amino acid sequence that can be identified with MS. These data were obtained from the RP-HPLC-MS of the tryptic hydrolysate of the total chymotryptic hydrolysate, containing the aggregating and non-aggregating peptides. For Glycinin 1 a coverage of ~90 % was obtained.
When analyzing a digest of a protein with RP-HPLC-MS, in general not 100 % of the protein sequence is covered with the peptides identified *(18, 19)*. Reasons for a coverage lower than 100 % could be that peptides are too hydrophilic to be separated on RP-HPLC, post translational modifications, not all peptides can be ionized, or that the amino acid sequence of the glycinin used does not fully match with the amino acid sequence used as present in Swiss-Prot.

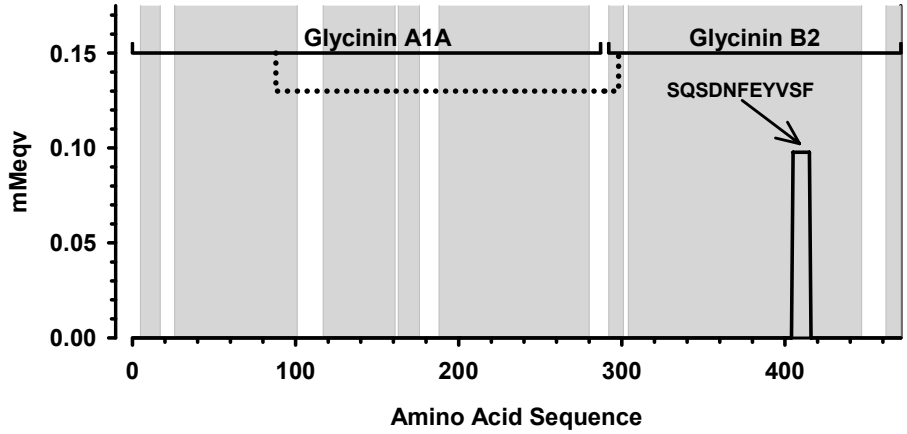


**Figure 5**: Illustration how the calculated relative concentration of SQSDNFEYVSF is used in the peptide mapping to identify the region of Glycinin 1 from which this peptide originates. The dotted line shows the location in the amino acid sequence where the disulfide bridge connects the acidic and the basic polypeptide. The grey regions indicate the coverage of the amino acid sequence that could be analyzed with MS.

**Accumulative-quantitative-peptide-mapping**

Of the 682 peaks present in the 53 chromatograms, all large peaks (185 peaks), responsible for 75 % (2,896 mAU*min) of the total peak area of all the integrated chromatograms, were selected to identify their constituent peptide sequence bases on MS and MS/MS data. For all these 185 RP-HPLC peaks the 5 steps as presented above were followed.

In total 218 peptides were present in the 185 RP-HPLC peaks. The number of 218 is due to the presence of 2 or more peptides responsible for one peak. For 132 of the 218 peptides a sequence was positively identified and originated from glycinin. The reasons for not finding 218 sequences might be diverse. One reason could be the coverage of the total amino acid sequence of the protein, which was 90 % for Glycinin 1, but equal or lower values were obtained for the other glycinin subunits (90, 87, 68 and 55 % for Glycinin 2, 3, 4 and 5, respectively). If the sequence of a peptide is part of the protein sequence that does not correspond for 100 % with the sequence used in the database, the peptide will not be positively identified. Another reason for not identifying all peptide sequences might be that peptides originate from sources other than glycinin. They might originate from the enzymes (trypsin and chymotrypsin) or from β-conglycinin, which was present in low concentration in the glycinin preparation *(5)*.

The 132 peptides contained 79 different unique glycinin peptide sequences. Due to the aspecific fractionation in the ion exchange chromatography experiment, peptides having the same sequence were present in more than one fraction. In addition to this, also methionine can be oxidized, which resulted in an additional peak for the oxidized peptide in the RP-HPLC chromatograms. The 79 peptides with a known sequence, represent a peak area of 2,025 mAU*min. This means that 70 % of the 75 % RP-HPLC peak area is positively linked to a part of the glycinin peptides, indicating that 53 % of the total peak area is assigned to peptide sequences. This value can be increased by determining more than 75 % of the peak area. However, this will not change the general picture of the regions in the protein from which the functional peptides originate. We have found that when the total peak area studied is extended from 66 to 75 % the additionally identified peptide sequences mainly originated from those regions of the parental protein that already showed a significant quantitative accumulation when only 66 % of the total peak area was analyzed (no further data shown).

An overview of the results of the interpretation of the 53 RP-HPLC chromatograms is given in Table 3.

**Table 3**: Overview of different parameters related to the interpretation of the 53 RP-HPLC chromatograms for the total peak area and the total analyzed peak area.

| | Parameter | Value |
|---|---|---|
| *Total peak area* | Total number of RP-HPLC Chromatograms | 53 |
| | Total number of peaks in 53 RP-HPLC Chromatograms | 682 |
| | Peak area (mAU*min) | 3,861 |
| *Total analyzed peak area* | Peak area (mAU*min) | 2,896 |
| | Number of peaks | 185 |
| | Number of peptides | 218 |
| | Number of positively identified peptide sequences | 132 |
| | Peak area of the 132 identified peptide peaks (mAU*min) | 2,025 |
| | Number of unique peptides | 79 |

Figure 6 shows the quantitative peptide map of functional regions of Glycinin 1. In this figure all the peptides from the 79 unique peptide sequences that are present in Glycinin 1 are accumulatively mapped. The accumulative peptide map gives an overview of those regions of Glycinin 1 that are likely to be present in the aggregating region. Mainly peptides from the basic polypeptide (B2) seem to be present in the aggregating regions and a part of the acidic polypeptide (A1A), around the amino acid 80.

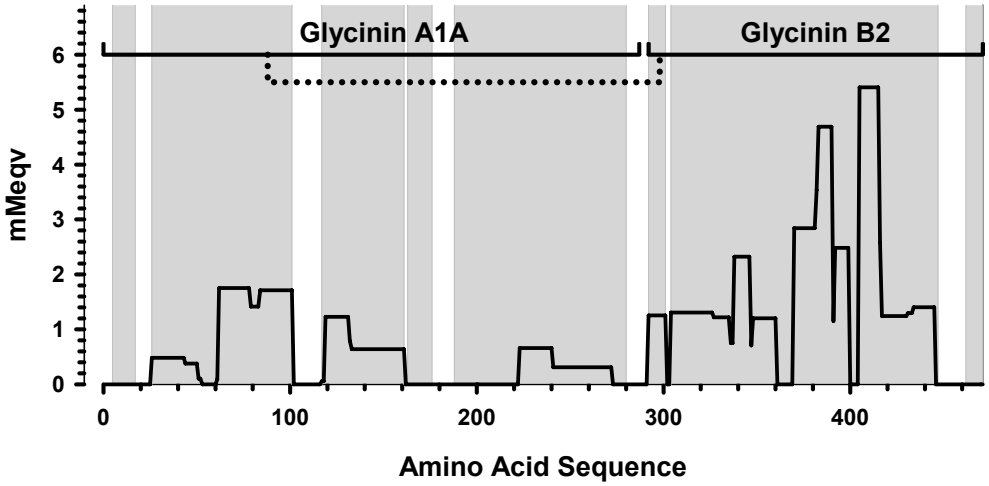


**Figure 6**: Quantitative peptide map of functional regions of glycinin 1 from which the aggregating peptides originate. The dotted line shows the location in the amino acid sequence where the disulfide bridge connects the acidic and the basic polypeptide. The grey regions indicate the coverage of the amino acid sequence that could be analyzed with MS.

It can be concluded that the method described in this paper for quantitative peptide mapping using RP-HPLC-MS, is sufficiently discriminating to identify which regions in a protein are predominantly responsible for the functional peptides derived from these proteins. This method is useful for protein hydrolysates, where functional peptides can be separated from peptides not contributing to the functionality of interest. It can e.g. also be applied to proteins that show resistance towards enzymatic degradation *(20)*. This is of interest to allergy

research as well as in other digestion studies, in particular when proteins interact with other components, which hinder the digestion of proteins *(21, 22)*.

**ACKNOWLEGEMENTS**

The soybeans were kindly provided by Fa. L.I. Frank (Twello, The Netherlands).

**REFERENCES**

1. Zhong, F.; Xin, Y.; Yue, L.; Shoemaker, C.F. Papain-induced gelation of soy glycinin (11S). *J. Food Sci.* **2006,** *71*, (5), E232-E237.
2. Mohri, M.; Matsushita, S. Improvement of water absorption of soybean protein by treatment with bromelain. *J. Agric. Food Chem.* **1984,** *32*, 486-490.
3. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.
4. Kuipers, B.J.H.; Alting, A.C.; Gruppen, H. Soy glycinin as a dominant source of aggregating peptides using different proteases. *Submitted; Chapter 4 in this thesis*. **2007**.
5. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophys.* **2006,** *1*, 178-188. *Chapter 3 in this thesis*.
6. Covey, T. Liquid chromatography/mass spectrometry for the analysis of protein digest. In *Protein and Peptide Analysis by Mass Spectrometery*, Chapman, J. R., Ed. Humana Press: Totowa, New Jersey, 1996; p 350.
7. Je, J.Y.; Park, P.J.; Byun, H.G.; Jung, W.K.; Kim, S.K. Angiotensin I converting enzyme (ACE) inhibitory peptide derived from the sauce of fermented blue mussel, *Mytilus edulis*. *Bioresour. Technol.* **2005,** *96*, (14), 1624-1629.
8. Saiga, A.; Okumura, T.; Makihara, T.; Katsuta, S.; Shimizu, T.; Yamada, R.; Nishimura, T. Angiotensin I-converting enzyme inhibitory peptides in a hydrolyzed chicken breast muscle extract. *J. Agric. Food Chem.* **2003,** *51*, (6), 1741-1745.
9. Chatterton, D.E.W.; Sorensen, E.S.; Petersen, T.E.; Lonnerdal, B. Heparin binding peptides from the 23 N-terminal residues of α(s1)-casein present in commercial preparations of whey protein concentrate, glycomacropeptide and α-lactalbumin. *Milchwissenschaft* **2001,** *56*, (3), 143-146.
10. Forster-Waldl, E.; Marchetti, M.; Scholl, I.; Focke, M.; Radauer, C.; Kinaciyan, T.; Nentwich, I.; Jager, S.; Schmid, E.R.; Boltz-Nitulescu, G.; Scheiner, O.; Jensen-Jarolim, E. Type I allergy to elderberry (*Sambucus nigra*) is elicited by a 33.2 kDa allergen with significant homology to ribosomal inactivating proteins. *Clin. Exp. Allergy* **2003,** *33*, (12), 1703-1710.
11. 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 RP-HPLC-MS analysis. *J. Agric. Food Chem., accepted for publication.* **2007**. *Chapter 5 in this thesis*.
12. Link, A.J.; Eng, J.; Schieltz, D.M.; Carmack, E.; Mize, G.J.; Morris, D.R.; Garvik, B.M.; Yates, J.R. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999,** *17*, (7), 676-682.
13. Storms, H.F.; van der Heijden, R.; Tjaden, U.R.; van der Greef, J. Optimization of protein identification from digests as analyzed by capillary isoelectric focusing-mass spectrometry. *J. Chromatogr., B* **2005,** *824*, (1-2), 189-200.

14. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. *J. Agric. Food Chem.* **2005,** *53*, 1031-1038. *Chapter 2 in this thesis*.
15. Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallon, B.J.; Fischer, R.L.; Sims, T.L.; Drews, G.N.; Goldberg, R.B. Characterization of the glycinin gene family in soybean. *Plant Cell* **1989,** *1*, (3), 313-328.
16. Scallon, B.; Thanh, V.H.; Floener, L.A.; Nielsen, N.C. Identification and characterization of DNA clones encoding Group-II glycinin subunits. *Theor. Appl. Genet.* **1985,** *70*, (5), 510-519.
17. Fischer, R.L.; Goldberg, R.B. Structure and flanking regions of soybean seed protein genes. *Cell* **1982,** *29*, (2), 651-660.
18. Lim, H.; Eng, J.; Yates, J.R.; Tollaksen, S.L.; Giometti, C.S.; Holden, J.F.; Adams, M.W.W.; Reich, C.I.; Olsen, G.J.; Hays, L.G. Identification of 2D-gel proteins: A comparison of MALDI/TOF peptide mass mapping to micro LC-ESI tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2003,** *14*, (9), 957-970.
19. Mann, M.; Jensen, O.N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **2003,** *21*, (3), 255-261.
20. Maleki, S.J.; Kopper, R.A.; Shin, D.S.; Park, C.W.; Compadre, C.M.; Sampson, H.; Burks, A.W.; Bannon, G.A. Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *J. Immunol.* **2000,** *164*, (11), 5844-5849.
21. Rawel, H.M.; Czajka, D.; Rohn, S.; Kroll, J. Interactions of different phenolic acids and flavonoids with soy proteins. *Int. J. Biol. Macromol.* **2002,** *30*, (3-4), 137-150.
22. Rawel, H.M.; Kroll, J.; Rohn, S. Reactions of phenolic substances with lysozyme - physicochemical characterisation and proteolytic digestion of the derivatives. *Food Chem.* **2001,** *72*, (1), 59-71.

# CHAPTER 7

# Identification of strong aggregating regions in soy glycinin upon enzymatic hydrolysis

**ABSTRACT**

Upon hydrolysis with chymotrypsin, soy glycinin has a strong tendency to aggregate. The regions of glycinin from which the aggregating peptides originate were identified by accumulative-quantitative-peptide-mapping. To this end, the aggregating peptides were further hydrolyzed with trypsin to obtain peptides of which the sequence can be identified using RP-HPLC-MS/MS. This resulted in a hydrolysate in which 90 % of the proteinaceous material was dissolved. The soluble fraction was analyzed using the method of accumulative-quantitative-peptide-mapping: Fractionation using ion exchange chromatography, followed by identification of peptides by RP-HPLC-MS/MS, quantification based on the absorbance at 214 nm, and finally peptide mapping. For the peptide mapping the proportions in which each of the 5 glycinin subunits are present, as determined by Edman degradation, were taken into account. The results showed that mainly the basic polypeptide and a part of the acidic polypeptide, close to the location of the disulfide bridge between the basic and acidic polypeptide, are present in the aggregating peptide fraction. Based on the results obtained an aggregation mechanism was proposed. The hydrophilic acidic polypeptides shield the hydrophobic basic polypeptides and the former are preferentially degraded upon hydrolysis. This results in a net increase in hydrophobicity of the remaining material, which mainly consists of the basic polypeptide fragments. This increase in hydrophobicity is proposed to be the driving force in the aggregation of chymotrypsin derived peptides of glycinin.

## INTRODUCTION

Various proteases derived from micro-organisms, plants and animals are reported to have the ability to induce aggregation of soy proteins. This aggregation of soy proteins might be a good approach to improve the texture of soy protein foods such as soy protein based analogs of dairy products *(1, 2)*.

Three microbial serine proteases have been studied in more detail with respect to their ability to induce soy protein aggregation. These proteases are subtilisin Carlsberg from *Bacillus Licheniformis (3-5)* and proteases that originate from *Bacillus sp*. K-295G-7 *(6-8)*, and *Bacillus pumilus* TYO-67 *(9, 10)*. The plant proteases bromelain *(11, 12)* and papain *(13)* and the animal-derived protease chymotrypsin *(14)* were also shown to induce aggregation of soy proteins. Also several industrial available protease preparations have the ability to induce soy protein aggregation *(15, 16)*.

Various studies have been performed to understand the aggregation behavior of hydrolyzed soy proteins. Mechanistic studies to understand enzyme-induced aggregation of hydrolyzed soy proteins comprised rheological *(16)* and turbidity *(3, 4, 17)* measurements. Forces involved in the aggregation were shown to be mainly hydrophobic interactions, whereas the presence of disulfide bridges seemed to play a minor role *(12, 13)*.

In the aggregation of soy protein isolate (SPI), it was shown that especially peptides derived from glycinin are the driving force for the aggregation upon hydrolysis with bacterial, plant, and animal-derived proteases *(12, 14)*. It was concluded *(14)* that the ability of several proteases to induce aggregation of soy glycinin indicates that not the enzyme specificity is of major importance to induce aggregation, but the ability of the substrate, glycinin, to yield strong aggregating peptides. This is in contrast to enzyme-induced aggregation of whey proteins in which the specificity of the protease is of major importance *(18, 19)*.

Due to the complex composition of glycinin, understanding the aggregation behavior of soy glycinin hydrolysates is rather difficult. In total 6 different glycinin genes have been identified from soybeans. These are Gy1 to Gy5 and Gy7, representing the glycinin subunits G1 to G5 and G7, respectively. Gy7 has been recently *(20)* identified and is related to the other five soybean subunits. Since Gy 7 is poorly expressed it is not further taken into account in this study. With the exception of G4, each glycinin subunit consists of one acidic (A) and one basic (B) polypeptide, which are connected with via a disulfide bridge. For the 5 major glycinin subunits six different acidic polypeptides (A1a, A2, A1b, A5, A4 and A3) and 5 basic (B2, B1a, B1b, B3 and B4) polypeptides have been identified *(21)*. The A4 polypeptide of G4 is the only acidic polypeptide that is not covalently linked to a basic polypeptide. Based on their amino acid sequence as present in the Swiss-Prot database (www.expasy.org), the molecular masses of the basic polypeptide can be calculated to be around 20 kDa, while those of the acidic polypeptides vary between 10.5 and 36.4 kDa. The 5 subunits can be divided into two subfamilies that are designated as Group-I (G1, G2, and G3) and Group-II (G4 and

G5) *(22, 23)*. Homologies between members of the same group range from 80 to 90 %, while between the groups the homology is less than 50 % *(21)*. At pH 7.6 and $I$ = 0.03 M, glycinin can be present as a trimer and as a hexamer *(24)* having molecular masses of 150-180 and 300-360 kDa, respectively.

The aim of the present study is to understand the aggregation behavior of soy glycinin-derived peptides using the recently described method of accumulative-quantitative-peptide-mapping *(25)*. This method identifies those regions in proteins from which peptides originate with specific functional properties. The method is applicable in systems in which the peptides of interest are too large to be identified directly with mass spectrometry and comprise a wide range of peptides with overlapping sequences that can not be isolated individually.

## MATERIALS AND METHODS

### Materials and chemicals

Soy glycinin was prepared from defatted soybean meal, prepared from Hyland soybeans (non-GMO) supplied by Fa. L.I. Frank (Twello, The Netherlands) as described previously *(4)*. Chymotrypsin (TLCK treated) and was obtained from Sigma Chemical Co. (St. Louis, MO, USA; C-3142). Eluents used for RP-HPLC-MS were all of HPLC grade. All other chemicals were of analytical grade and were purchased from Merck (Darmstadt, Germany) or Sigma.

### Analysis of the protein content

The nitrogen content of various samples was determined in duplicate by the Dumas method using an NA2100 Nitrogen and Protein Analyzer (CE INSTRUMENTS, Milano, Italy) according to the manufacturer's instructions. Methionine was used as a standard. The nitrogen conversion factor for purified glycinin is 5.57, based on the amino acid compositions of the 5 subunits as calculated from the Swiss-Prot database (www.expasy.org; Primary accession numbers used were: P04776 (Glycinin 1; A1a and B2 polypeptides), P04405 (Glycinin 2; A2 and B1a polypeptides), P11828 (Glycinin 3; A1b and B1b polypeptides), P02858 (Glycinin 4: A5, A4 and B3 polypeptides), and P04347 (Glycinin 5; A3 and B4 polypeptides)).

### Determination of the proportions of different glycinin subunits

Glycinin was dissolved in 700 µL 8 M guanidiniumchloride, 50 mM Tris-HCl buffer (pH 8.0) containing 15 mM DTT. The protein was reduced for 30 minutes at 60°C. Iodoacetamide (IAA) was added up to 50 mM and the mixture was incubated in the dark for 30 minutes to alkylate the peptides. Every 10 minutes the pH was checked and, if necessary, adjusted to 8.0 using 1 M NaOH. The alkylated glycinin was applied onto a Shodex Protein KW-803 size exclusion chromatography column (300 x 8 mm; Showa Denko K.K., Tokyo, Japan) equilibrated and run with 30 % (v/v) aqueous acetonitrile containing 0.1 % (v/v)

trifluoroacetic acid (TFA). The flow rate was 0.5 mL/min and the absorbance of the eluate was monitored at 214 nm *(4)*. Fractions of 0.2 mL were collected. To separate the glycinin from salts and other proteinaceous impurities, fractions containing glycinin were pooled and freeze dried.

The proportion of the various subunits of glycinin in the freeze-dried sample was estimated from the Initial Yields of the sequences of the individual acidic (and basic) polypeptides as reported in Swiss-Prot. These Initial Yields were obtained by performing a 6 cycle N-terminal sequence analysis on a 23 µg sample of glycinin loaded onto a pre-washed and Biobrene coated glass filter. N-terminal sequence analysis was performed by Edman degradation *(26, 27)* with a Procise 494 (Applied Biosystems, Foster City, CA, U.S.A) automated sequencing system *(28)*, equipped with a 140 C Microgradient System and a 758A Absorbance detector, and using protocols, reagents, chemicals and materials from Applied Biosystems (Foster City, CA, USA). The proportion (mole %) of the various subunits was calculated based on the known N-termini of the all polypeptides in glycinin. The 6 acidic polypeptides all have a unique N-terminus, whereas the basic polypeptides within each Group (Group I: G1, G2, and G3; Group II, G4 and G5) are identical in this respect. The results of 3 N-terminal sequence analysis cycles were essentially sufficient to calculate the ratio. Nevertheless, 6 cycles were used in the calculation to verify the correctness of the ratio found after the first 3 cycles.

**Accumulative-quantitative-peptide-mapping**

To identify those regions in glycinin from which aggregating peptides originate the method of accumulative-quantitative-peptide-mapping was used as previously described in detail *(25)*. The method is schematically summarized in Figure 1 and will be briefly described here.

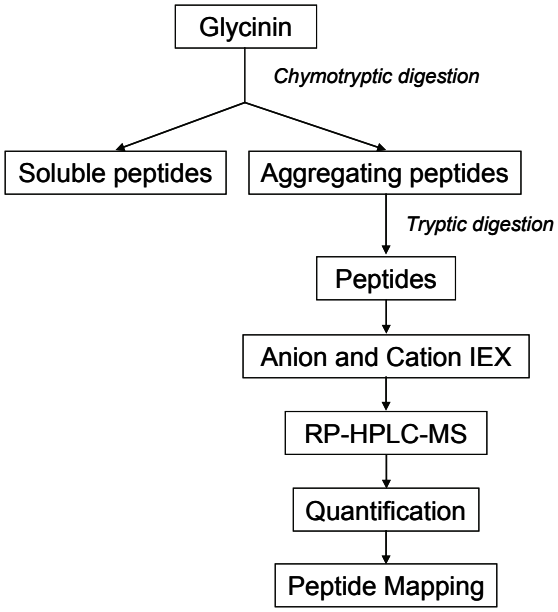


**Figure 1**: Schematic overview of accumulative-quantitative-peptide-mapping to identify regions in glycinin from which the aggregating peptides originate *(25)*.

A 1 % (w/w) solution of heated glycinin was hydrolyzed with chymotrypsin up to a DH of 2.2 % using the pH-stat method at pH 8.0 and 40°C. When the desired DH was reached, the enzymatic hydrolysis was stopped by addition of phenylmethylsulphonyl fluoride (PMSF). Water and NaCl were added to the hydrolysate to result in a final protein concentration of 0.8 % (w/w) and an ionic strength of 0.03 M, as described previously *(5)*. After centrifugation (5 min, 5,000 x *g*, 20°C), analysis of the protein concentration (Dumas method) in the supernatant showed that ~43 % of the proteinaceous material was aggregated. The pellet was washed three times by re-suspending it in Millipore water containing 0.03 M NaCl. After washing, the pH of the pellet was decreased to ~2.5 using TFA to resolubilize the peptides, followed by freeze-drying. These freeze-dried peptides were further denoted aggregating peptides.

The aggregating peptides were dissolved in 20 mM Tris-HCl buffer (pH 8.0) containing 8 M urea and subsequently reduced and alkylated. After dilution to 1.6 M urea, the aggregating peptides were further hydrolyzed using trypsin at 37°C. This tryptic digestion was performed in three cycles.

The hydrolyzed sample was centrifuged (15 min, 24,000 x *g*, 20°C). The pellet was re-suspended in 20 mM Tris-HCl buffer (pH 8.0) containing 8M urea. The solution obtained was diluted to 1.6 M urea by the addition of 20 mM Tris-HCl buffer (pH 8.0). Trypsin was added to the re-solubilized pellet and incubated at 37°C again. After incubation, the sample was treated in the same way as described above, by re-suspending the pellet in 20 mM Tris-HCl buffer (pH 8.0) containing 8 M urea, dilution to 1.6 M urea with 20 mM Tris-HCl buffer (pH 8.0) and subsequent incubation with trypsin followed by centrifugation. The supernatants of the three tryptic digestion cycles were mixed prior to fractionation using ion exchange chromatography.

The combined supernatants of the tryptic digests was applied onto a 1 mL Mono Q 5/50GL anion exchange column (Amersham Biosciences, Uppsala, Sweden) equilibrated with a 20 mM Tris-HCl buffer (pH 8.0). Fractions of 1 mL were collected. The unbound fractions were decreased in pH to pH 3.5 using formic acid and applied onto a 1 mL Mono S 5/50GL cation exchanger (Amersham Biosciences, Uppsala, Sweden) equilibrated with a 20 mM formic acid buffer (pH 3.5). Fractions of 1 mL were collected. Elution of the bound proteins on the anion and cation exchanger was performed using their corresponding equilibration buffers containing 2 M NaCl. The eluate was monitored at 214 nm. The unbound fraction of cation exchange chromatography was concentrated by freeze drying.

All bound fractions collected from anion (25 fractions), and cation (27 fractions) exchange, and the unbound fraction of the cation exchange were analyzed quantitatively using RP-HPLC-MS using a C18 column (218MS52; 250 x 2.1 mm 5µm) (Grace Vydac, Hesperia, CA, USA), installed on a Spectra System HPLC (Thermo Separation Products, Fremont, CA,

USA). A flow rate of 0.2 mL/min was used. The solvents used were Millipore water containing 0.1 % (v/v) formic acid (Solvent A) and acetonitrile containing 0.085 % (v/v) formic acid (Solvent B). All operations and data processing were controlled by Chromeleon version 6.7 software (Dionex Corporation, Sunnyvale, USA). The eluate was applied onto an LCQ Deca XP Max (Thermo Finnigan, San Jose, CA, USA) with the use of electrospray ionization and detection in the positive ion mode. The capillary spray voltage was 4 kV, and the capillary temperature was 200°C. The instrument was controlled by Xcalibur software v1.3 (Thermo Finnigan, San Jose, CA, USA). The MS/MS functions were performed in data dependent mode. BioWorks software, version 3.1 (Thermo Electron, San Jose, CA, USA) was used for automatic sequencing and database search for the sequences in a database only containing sequences of glycinin G1 to G5 and trypsin and chymotrypsin. To discriminate between correct and incorrect peptide sequence assignment the cross correlation value (Xcorr) for each identified peptide was used as a criterion *(29, 30)*. For positive identification of the peptides a Xcorr threshold of 1.5 for single charged peptides, 2.0 for double, and 2.5 for triple charged peptides was used.

Of the 682 peaks present in the 53 chromatograms, all the largest peaks (185 peaks), responsible for 75 % (2,896 mAU*min) of the total peak area of all 53 integrated chromatograms, were selected to identify their constituent peptide sequences based on MS and MS/MS data. In total, 218 peptides were present in the 185 RP-HPLC peaks. From these 218 peptides the sequences were positively identified of 132 peptides that originate from glycinin. The 132 peptides contained 79 different unique peptide sequences. The 79 peptides with a known sequence represent an area of 2,025 mAU*min, which means that 70 % of the 75 % RP-HPLC area is positively linked to a part of the glycinin sequence.
Quantitative analysis of the peptide concentration was performed using Formula 1:

$$c_{\text{peptide in 1 subunit}}(\text{Meqv}) \sim \frac{A_{\text{peak}}}{\varepsilon_{\text{peptide}}} \text{ x } \frac{1}{n_{\text{peptides in peak}}} \text{ x } F_{\text{subunit}} \qquad [1]$$

In Formula 1 the RP-HPLC peak area ($A_{peak}$) of each peak was divided by the theoretically calculated molar extinction coefficient ($\varepsilon_{peptide}$) *(31)*, corrected for the total number of different peptides present in the RP-HPLC peak ($n_{peptides\ in\ peak}$), and multiplied by the molar proportion ($F_{subunit}$) in which the subunits, containing the sequenced peptide, are present *(25)*. In total 79 sequences were identified that were used for accumulative-quantitative-peptide mapping using all 5 glycinin subunits.
All 79 peptides identified, together with their relative concentrations were used to build the quantitative peptide map as a projection on the amino acid sequence of each individual glycinin subunit (G1 to G5).

The maximum sequence coverage that can be obtained was analyzed using a tryptic digestion of the total hydrolysate after chymotryptic hydrolysis of heated glycinin. This hydrolysate contains the aggregating, as well as the non-aggregating peptides, and was analyzed by RP-HPLC-MS in a similar way as the 53 fractions described above. For all glycinin subunits a total coverage was obtained of 90, 90, 87, 68 and 55 % for G1, G2, G3, G4, and G5, respectively.
In addition to the peptides present in the supernatant, the final pellet after tryptic digestion was dissolved in 5 % (v/v) acetonitrile in Millipore water, containing 0.1 % (v/v) formic acid and also analyzed on RP-HPLC-MS similar to all other analyzed samples.

### High-performance size-exclusion chromatography (HP-SEC)

Aliquots (100 µL) of the 1 % (w/w) glycinin solution used for the chymotryptic hydrolysis were taken before and after the heat treatment (30 min, 95°C) and mixed with 900 µL 10 mM sodium phosphate buffer (pH 8.0).
After 1 h of stirring at ambient temperature the samples were applied onto a Superdex 200 HR 10/30 column (300 x 10 mm, Amersham Biosciences, Uppsala, Sweden). The column was equilibrated with 10 mM sodium phosphate buffer (pH 8.0). The flow rate was 0.5 mL/min and the absorbance of the eluate was monitored at 214 nm. The void volume of the column was approximately 7.8 mL as determined by injection of blue dextran. The column was calibrated using various proteins in a molecular mass range of 43 to 669 kDa (ovalbumin, catalase, ferritin and thyroglobulin).
In addition to the above described HP-SEC analysis, glycinin, its DH 2.2 % hydrolysate, the aggregating peptide fraction, the total tryptic hydrolysate and its supernatant and pellet were all analyzed using size exclusion chromatography under denaturing conditions *(5)*. To allow quantitative comparison of the chromatograms obtained, the pellet after tryptic digestion was resuspended in the same volume as present prior to the centrifugation. Samples were dissolved under denaturing (8 M Guanidiniumchloride) and reducing conditions (100 mM DTT), and run on a Shodex Protein KW-803 column (300 x 8 mm; Showa Denko K.K., Tokyo, Japan) equilibrated and run with 6 M urea containing 30 % (v/v) acetonitrile and 0.1 % (v/v) TFA. The aggregating peptide fraction was also applied on the Shodex KW-803 column in a reduced form. The flow rate was 0.5 mL/min and the absorbance of the eluate was monitored at 214 nm. The column was calibrated using proteins with known molecular masses as described previously *(5)*.

## RESULTS AND DISCUSSION

### Influence of heat treatment on the quaternary structure of glycinin

To study the influence of the heat treatment (prior to hydrolysis) on the quaternary structure of glycinin, HP-SEC was performed of glycinin before and after heating. Figure 2 shows the molecular size distribution patterns of glycinin before and after heating. In the non-heated

glycinin, 4 different fractions can be observed (A to D). Based on the calibration of the column, fractions B and C most likely represent the glycinin hexamer (300-360 kDa) and trimer (150-180 kDa), respectively. These hexamers and trimers are known to be present next to each other at neutral pH and low ionic strength *(24)*. Fraction A has a higher apparent molecular mass than the hexamer and is probably a polymer of glycinin, also denoted as 15S *(32)*. Fraction D probably represents free glycinin subunits (50-62 kDa) or even acidic and basic polypeptides (10-36 kDa).

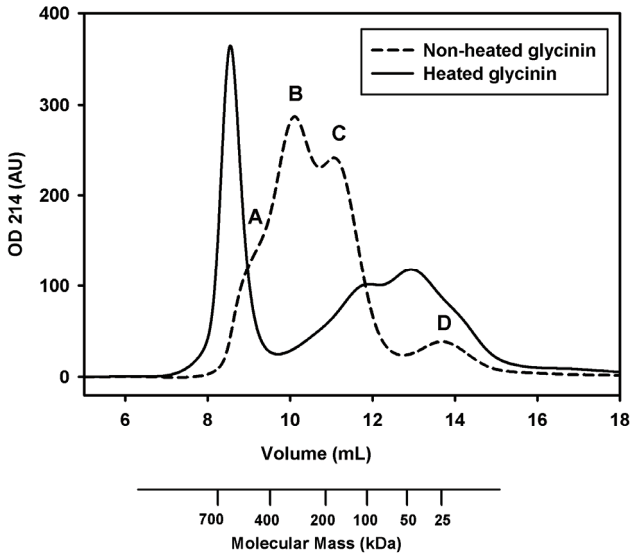


**Figure 2**: HP-SEC chromatograms of non-heated glycinin (-----) and heated glycinin (——) in 10 mM sodium phosphate buffer (pH 8.0). The letters A, B, C, and D represent 4 fractions of the non-heated glycinin representing the 15S, 11S, 7S, and dissociation products, respectively.

Upon heating it can be observed that on the one hand the hexamers and trimers form aggregates (peak around 8.2 mL) and on the other hand dissociate. The dissociated fraction contains proteins with sizes similar, but mainly smaller that the glycinin trimer. These hexamers and trimers are probably dissociated into trimers, dimmers, monomers and probably also individual polypeptides that are dissociated by the reduction of the disulfide bridge. When comparing the peak areas of the chromatograms before and after heating, it can be approximated that ~ 50 % of all the proteinaceous material is present in the aggregates, while the other ~ 50 % is present as dissociation products. Since the heated glycinin has a clear appearance, the formed aggregates can be denoted as soluble aggregates. The observations are in agreement with the work of Mori and co-workers *(33)* and Yamauchi and co-workers *(34)* who described the formation of soluble aggregates upon heating of glycinin (pH 7.6, $I$ = 0.5 M), as well as a dissociation of the individual subunits in their constituent polypeptides. However, the insolubilization of the basic polypeptides, as described by Mori and co-workers *(33)*, was not observed, which might be due to the low ionic strength (proteins were dissolved only in Millipore water) during heating in the present study.

### Effect of disulfide bridges on the size of aggregating peptides

Due to the heat-induced aggregation (Figure 2) prior to hydrolysis, it might be that new intermolecular disulfide bridges are formed *(35)* that can be of influence on the observed aggregation of peptides. Figure 3A shows the size distribution obtained by HP-SEC under denaturing conditions of the reduced glycinin and reduced and non-reduced aggregating peptide fraction. When comparing the reduced hydrolyzed sample with the non-reduced hydrolyzed sample it can be seen that there is a slight shift to larger sizes from the non-reduced sample. This indicates that in the aggregating peptides disulfide bridges are present that slightly increase the size of the fragments. Nevertheless, this shift is not large, indicating that no large number of intermolecular disulfide bridges are formed upon heating.

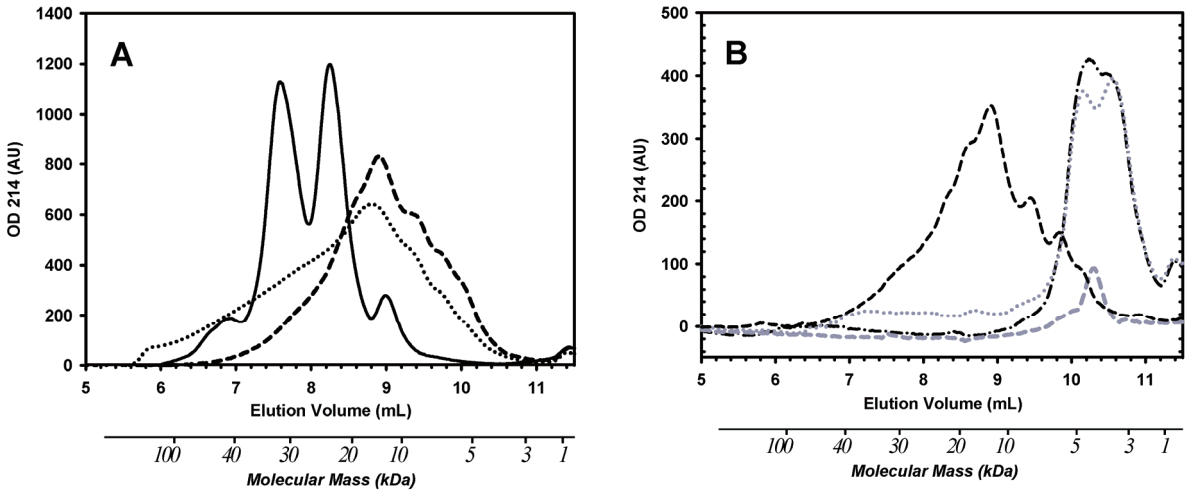


**Figure 3**: HP-SEC chromatograms under denaturing conditions (6M urea, 30 % (v/v) acetonitrile, 0.1 % (v/v) TFA). Figure **3A** shows the intact reduced glycinin (——), and its reduced (— —), and non-reduced (• • • •) aggregating peptides. Figure **3B** shows the aggregating peptides (— —), and its corresponding total tryptic hydrolysate of aggregating peptides (— • —), its supernatant (• • • •) and pellet (— —).

### Monitoring the tryptic hydrolysis

Since the aggregating peptides are too large to be identified directly with MS, these peptides were further degraded with trypsin prior to analysis. To monitor the hydrolysis of the isolated aggregating peptides with trypsin, HP-SEC under denaturing conditions was performed for the aggregating peptides and the total tryptic hydrolysate as well as its supernatant and pellet. It can be observed in Figure 3B that the aggregating peptide fraction mainly contains peptides with apparent molecular masses in the range of ~5,000 to 25,000 Da. Upon tryptic hydrolysis the aggregating peptides are reduced to peptides with apparent molecular masses below 5,000 Da. After the repeated hydrolysis with trypsin, there were still some aggregates present in the hydrolysate. These aggregates could be dissolved and analyzed by HP-SEC under denaturing conditions. The molecular masses of these peptides range between 3,000 and 5,000 Da. The peak area of the pellet represents ~10 % of the peak area of the total tryptic hydrolysate. It is

assumed that this 10 % of the peak area also represents 10 % of the proteinaceous material. This means that 90 % of the proteinaceous material of the aggregating peptides was present in the supernatant that was further fractionated with ion exchange chromatography and analyzed with RP-HPLC-MS followed by quantification and peptide mapping.

**Relative proportions of the 5 glycinin subunits**

In *accumulative-quantitative-peptide-mapping* it may occur that a peptide that is identified is present in more than one subunit of the parental protein *(25)*. In order to present a correct peptide mapping, the distribution of the quantity of such a peptide over the different subunits in which the peptide is present should be according to the molar proportions in which these subunits are present in the total parental protein. In order to be able to apply this correction, the molar ratios between all 5 glycinin subunits was determined using N-terminal sequence analysis of the glycinin polypeptides.

Table 2 shows the absolute concentration measured (pmol) and the relative proportion (mole %) of the individual polypeptides based on the sequence yield of the various amino acids. These values are determined using the known N-terminal sequences of all acidic and basic polypeptides as present in the Swiss-Prot database. While the N-termini of the various acidic polypeptides differ from each other, the basic polypeptides of G1, G2, and G3 have the same N-terminus, which also counts for the N-termini of G4 and G5. Therefore, 8 different N-terminal sequences should in principle be distinguishable during the Edman degradation (results not shown). It can, however be seen that not all amino acid residues are unique for one acidic polypeptide (i.e. one glycinin subunit). The Phe-signal in the first Edman cycle, for example, is a combination of the contributions of A1a and A1b. Similar to that, the Ile-signal in the first cycle is due to both A5 and A3. On the other hand, the Leu-signal and the Arg-signal in the first cycle are unique and thus give a direct indication of the relative contributions of subunits G2(A2) and G4(A5), respectively. In a similar way the data in the other Edman cycles can be used to estimate the relative contributions of the different subunits. In addition, these cycles are used to verify the validity of these estimations. The signals belonging to the basic polypeptides are used as a final verification. The sum of the contributions of the acidic polypeptides of G1, G2, and G3 (72 %), and the sum of the contributions of G4 and G5 (27/28 %), are equal to the relative contributions of the basic polypeptides of G1-3 (73 %) and G 4-5 (27 %), respectively.

Based on the proportions of acidic polypeptides it can be observed that G1 and G2 are the most dominant subunits present in glycinin. The subunit G3 is present in a rather low proportion. G4 and G5 are present in a 1 to 1 ratio. The relative proportions (mole %) as shown in Table 2 are used to calculate $F_{subunit}$ (Formula 1) for each peptide that was identified to be present in more than one subunit.

**Table 2**: Initial yield (absolute amount (pmol)) and the relative proportion (mole %) of individual polypeptides in glycinin, as determined by N-terminal sequencing.

| Polypeptide | N-terminus | Initial Yield (pmol) | Relative Prop. (mole %) |
|---|---|---|---|
| ***Acidic polypeptides (100 %)*** | | | |
| G1 (A1a) | Phe-Ser-Ser-Arg-Glu-Gln | 30 | 36 |
| G2 (A2) | Leu-Arg-Gly-Gln-Ala-Gln | 25 | 30 |
| G3 (A1b) | Phe-Arg-Glu-Gln-Pro-Gln | 5 | 6 |
| G4 (A5) | Ile-Ser-Ser-Ser-Lys-Leu | 11 | 13 |
| G4 (A4) | Arg-Arg-Gly-Ser-Arg-Ser | 12 | 14 |
| G5 (A3) | Ile-Thr-Ser-Ser-Lys-Phe | 12 | 14 |
| ***Basic polypeptides (100 %)*** | | | |
| G1 - G3 (B2, B1a, B1b) | Gly-Ile-Asp-Glu-Thr-Ile | 55 | 73 |
| G4 - 5 (B3, B4) | Gly-Val-Glu-Gly-Asn-Ile | 20 | 27 |

## Peptide mapping for all glycinin subunits

Peptide mapping for the 5 glycinin subunits is performed on the basis of the results obtained from the 90 % soluble proteinaceous material after tryptic digestion of the aggregating peptides. Figures 4 and 5 show the accumulative peptide maps for Glycinin 1 to 3 and Glycinin 4 and 5, respectively. The accumulative peptide map gives an overview of those regions in each subunit that most likely yield aggregating peptides. The higher the bar at a specific region in the sequence, the stronger the contribution of the amino acid sequence of this region to the aggregating properties of the peptides derived from the corresponding subunit.

On top of each graph it is indicated which part of the sequence represents which polypeptide. The dotted lines between the different polypeptides represent the disulfide bridge that connects the acidic and basic polypeptide. The grey regions represent the coverage of the amino acid sequence that could be identified with MS. These coverage data were obtained using RP-HPLC-MS of the tryptic hydrolysate of the total chymotryptic hydrolysate, containing both the aggregating and the non-aggregating peptides as described previously *(25)*.

Glycinin 1 and 2 show (Figure 4) the highest quantitative response in the peptide maps compared to Glycinin 3, 4, and 5. When the acidic and basic polypeptides are compared, it can be seen that a relative high proportion of the basic polypeptide is identified to be present in the aggregating peptides. In addition, the concentration determined (height of the bars) is higher for the basic polypeptides than for the acidic polypeptide. For the acidic polypeptide of Glycinin 1 it seems that the highest proportion of positively identified peptides can be observed around the location of the disulfide bridge. This is less apparent for Glycinin 2.

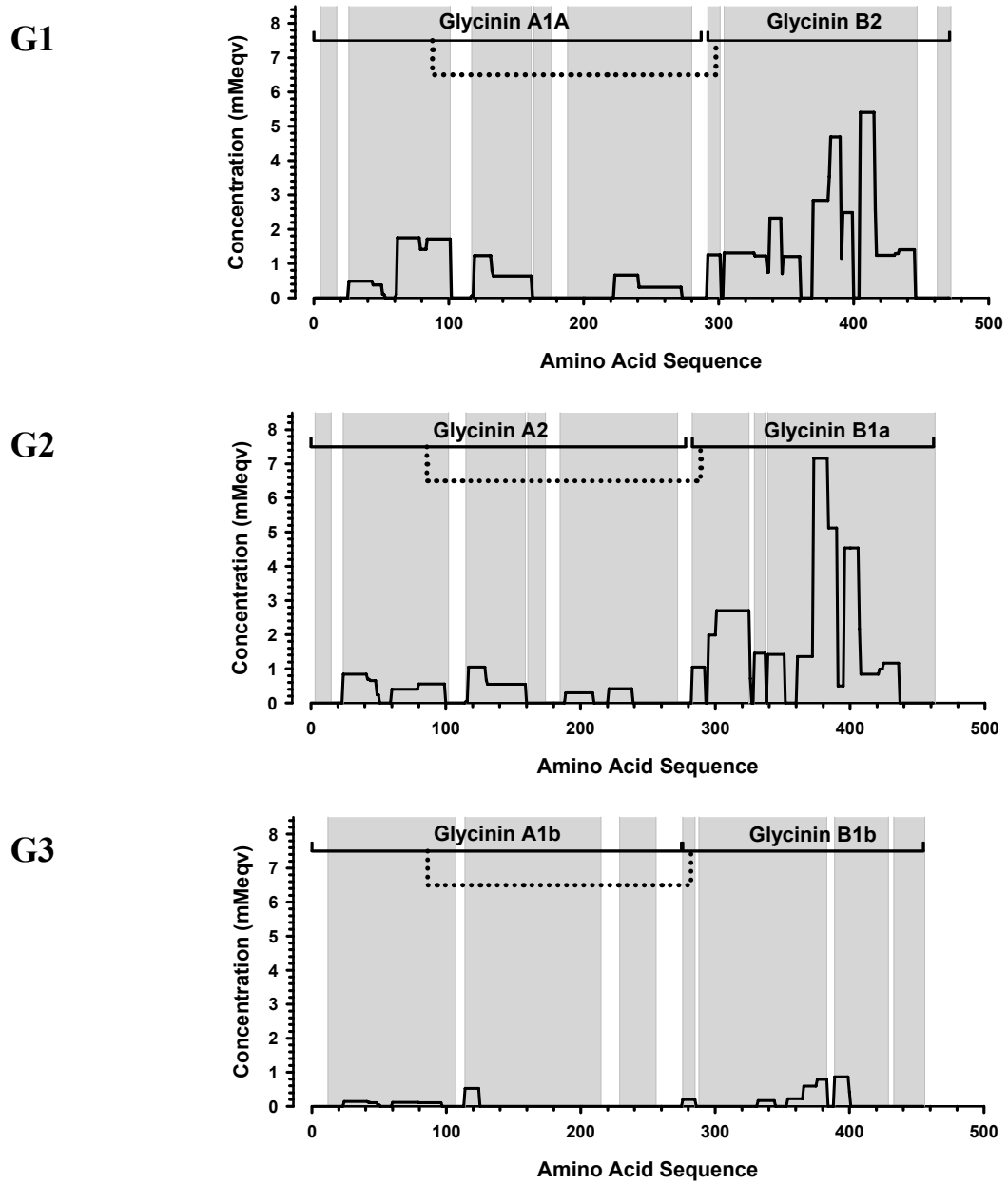


**Figure 4**: Accumulative quantitative peptide maps of Glycinin 1, 2, and 3 from which the aggregating peptides originate. The dotted line shows the location in the amino acid sequence at which the disulfide bridge connects the acidic and the basic polypeptide. The grey regions indicate the coverage of the amino acid sequence that could be analyzed with MS.

For Glycinin 3 the response is much lower compared to the responses of Glycinin 1 and 2. This is probably due to the much lower proportion of Glycinin 3 in the glycinin preparation (Table 2). Nevertheless, for Glycinin 3 as well it can be observed that the response and coverage in the basic polypeptide are higher than those of the acidic polypeptide.

For both Glycinin 4 and 5 a lower coverage of the amino acid sequence is obtained (grey regions), which may contribute to their lower overall responses. Figure 5 shows that also for Glycinin 4 the response for the basic polypeptide is higher than in the acidic polypeptide.

Also the acidic polypeptide gives the highest response around the disulfide bridge. No peptides were detected from the A4 polypeptide of Glycinin 4. For Glycinin 5 a low response for the basic polypeptide could be expected, due to its very low overall sequence coverage from the MS data. As a consequence, Glycinin 5 shows the highest response in the acidic polypeptide around the disulfide bridge.

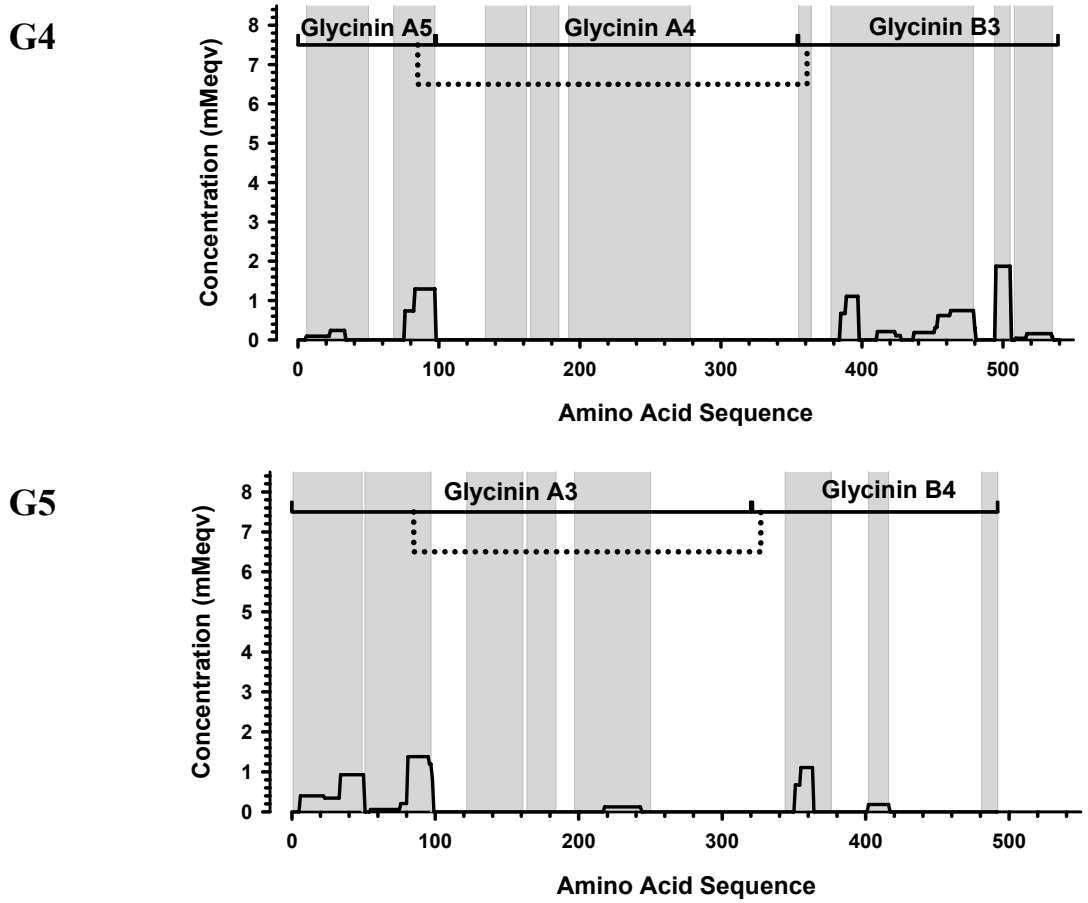


**Figure 5**: Accumulative quantitative peptide maps of regions of Glycinin 4 and 5 from which the aggregating peptides originate. The dotted line shows the location in the amino acid sequence at which the disulfide bridge connects the acidic and the basic polypeptide. The grey regions indicate the coverage of the amino acid sequence that could be analyzed with MS.

Overall it can be observed that next to the basic polypeptide, the region of the acidic polypeptide, which contains the disulfide bridge connecting the acidic with the basic polypeptide is dominantly present in the aggregating peptides in all 5 glycinin subunits. It is shown that the presence of disulfide bridges slightly increases the size of the aggregating peptides (Figure 3A). Therefore, it is likely that this disulfide bridge is still present in the aggregating peptides between the peptides originating from the basic and acidic polypeptide. As a result, this may facilitate the (co-)precipitation of peptides of the acidic polypeptide that are covalently connected to the aggregating peptide of the basic polypeptide. However, the disulfide bridge between the acidic and basic polypeptide is located in a hydrophobic region

of the acidic polypeptide. If this disulfide bridge would be reduced upon heating (Figure 2) it seems likely that peptides originating from this region of the acidic polypeptide have the ability to aggregate when not covalently connected to the basic polypeptide.
It can be concluded that most aggregating peptides originate from the basic polypeptides. Since the molecular masses of the aggregating peptides are between 5,000 and 25,000 Da, and the basic polypeptides are not intact anymore, it can be concluded that the basic polypeptide is only slightly degraded, and thus seems to show a rather high resistance towards enzymatic hydrolysis. This is in agreement with previous observations that the basic polypeptides are degraded more slowly than the acidic polypeptide, when hydrolyzed with bromelain *(12)* and is not degraded at all upon hydrolysis with a serine protease of *Bacillus pumilus (9)*.

**Aggregating peptides in the pellet after tryptic digestion**

After repetitive tryptic digestion, peptides with a strong tendency to aggregate were still present, yielding a precipitate. They comprise ~10 % of the total amount of aggregating peptides. These peptides were analyzed with RP-HPLC-MS (chromatogram not shown) to identify their amino acid sequences followed by determination of their locations in the sequence of the various glycinin subunits. Since 10 % of the peptides was present in the pellet after tryptic digestion, the total peak area of the integrated peaks was adjusted to result in a peak area ratio of 9 : 1 between the RP-HPLC chromatograms of the soluble and insoluble peptide fractions after tryptic digestion. This correction enables a quantitative comparison with the peptide maps in Figures 4 and 5. The 8 largest peaks, responsible for 75 % of the total peak area in the pellet fraction, were studied. From 6 out of these 8 peaks the amino acid sequences of the constituting peptides could be positively identified. These peptides can be observed in Table 3 together with their corresponding hydrophobicity. One of the 6 peptides is present in 2 glycinin subunits (Glycinin 1 and 3). These 6 peaks account for 91 % of the total peak area of the 8 largest peaks and are mapped quantitatively in Figure 6. The location of the polypeptides and the disulfide bridges are shown for all 5 glycinin subunits.

**Table 3**: Peptides identified in the final pellet obtained after repetitive tryptic digestion, their calculated hydrophobicities, and their location in the different subunits.

| Subunit | Sequence | Hydrophobicity [a] | Location |
|---|---|---|---|
| Glycinin 1 | NAMFVPHYNLNANSIIYALNGR | 0.171 | B2 f[348-369] |
| Glycinin 2 | RPSYTNGPQEIYIQQGNGIFGMIFPGCPSTY | 0.198 | A2 f[60-90] |
| | NAMFVPHYTLNANSIIYALNGR | 0.204 | B1a f[339-359] |
| Glycinin 3 | NAMFVPHYNLNANSIIYALNGR | 0.171 | B1b f[332-353] |
| Glycinin 4 | NGIYSPHWNLNANSVIY | 0.210 | B4 f[411-427] |
| | NGIYSPHWNLNANSVIYVTR | 0.104 | B4 f[411-430] |
| Glycinin 5 | NGSHLPSYLPYPQMIIVVQGK | 0.227 | A3 f[55-75] |

[a] Calculated according to Eisenberg *(36)*

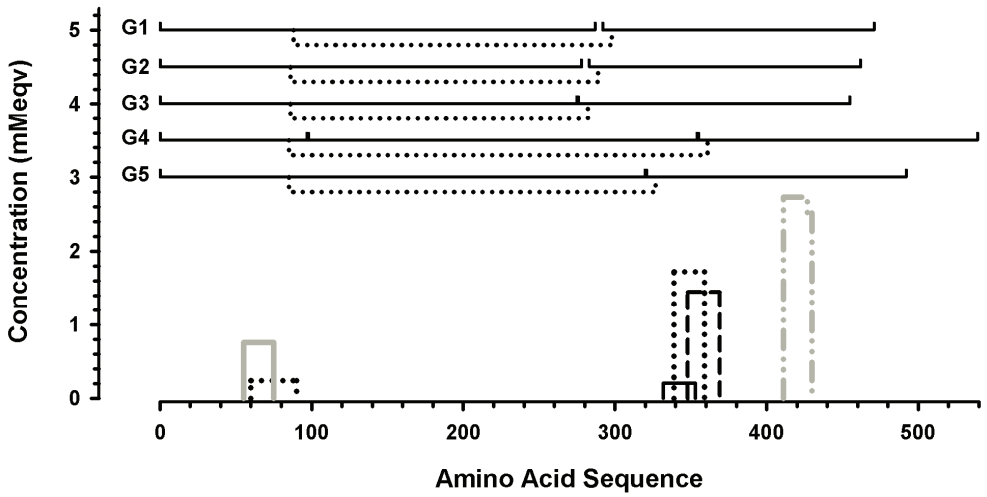


**Figure 6**: Accumulative quantitative peptide map of peptides in the pellet fraction after repetitive tryptic digestion of the aggregating peptides originating from G1 (— —), G2 (• • • •), G3 (——), G4 (— • • —), and G5(——).The dotted line shows the location in the amino acid sequence at which the disulfide bridge connects the acidic and the basic polypeptide.

It can be seen that 4 of the 6 identified peptides originate from basic polypeptides, while 2 of them originate from the acidic polypeptide, covering the region around the location of the disulfide bridge. All identified peptides originate from a region in glycinin that was already defined in the peptide maps (Figures 4 and 5) as being a region contributing to the aggregation of glycinin upon chymotryptic hydrolysis. The average hydrophobicity score of all intact acidic and basic polypeptides is -0.228 (± 0.048) and 0.018 (± 0.034), respectively. The higher the score, the higher the hydrophobicity. When comparing these values with values of the peptides in Table 3, it can be concluded that the latter peptides are relatively hydrophobic. This relative high hydrophobicity is, therefore, probably the reason why these peptides are not soluble.

It can be concluded that the sequences of peptides that aggregate after repetitive tryptic digestion do not reveal new regions of glycinin that strongly contribute tot the aggregating regions. They were shown to originate from the same regions as already defined as aggregating regions based on the analysis of 90 % of the aggregating peptides.

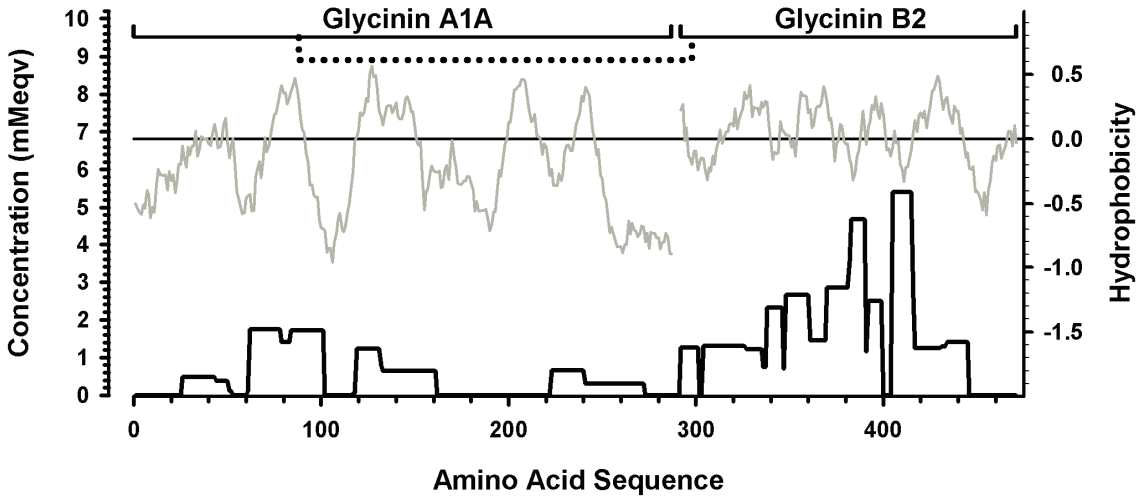


**Figure 7**: Hydrophobicity plot according to Eisenberg *(36)* (——) for Glycinin 1 combined with its corresponding accumulative quantitative peptide map (——), including the identified peptide in the pellet fraction as shown in Table 3.

As already mentioned above, the hydrophobicity is likely to play an important role in the aggregation of glycinin-derived peptides. Therefore, in Figure 7 the hydrophobicity according to Eisenberg *(36)* (grey line) as a function of the amino acid sequence for Glycinin 1 is shown. In addition, the total accumulative quantitative peptide map, being a combination of Figures 4 and 6, is shown. The calculation of the hydrophobicity was performed using the ProtScale software from the Swiss Institute of Bioinformatics (www.expasy.org), using a "window size" of 15 (= peptide range over which the average hydrophobic score is calculated). If the hydrophobicity score is above 0, the region in the sequence can be regarded to be hydrophobic. It shows that the hydrophobic regions are strongly present in the aggregating peptide, revealing the large contribution of hydrophobic interactions in the aggregation mechanism of the glycinin-derived peptides. This graph is only shown for Glycinin 1, since all subunits showed the same pattern.

**Mechanism of enzyme-induced aggregation**

Recently, Nagai and Inouye *(17)* proposed a mechanism for the enzyme-induced aggregation of soy protein isolate upon hydrolysis with subtilisin Carlsberg. First the hydrophilic surface areas of the proteins are degraded, subsequently followed by the degradation of the hydrophobic core of the protein. This results in a strong decomposition of the tertiary structure of the proteins followed by aggregation due to hydrophobic interactions. However, no attention was paid to the quaternary structure of the proteins prior to hydrolysis and the aggregating fractions were not identified with respect to the regions from which they originate in the SPI. In the present study the quaternary structure of the proteins prior to hydrolysis was studied together with the regions from which the peptides originate.

Figure 8 visualizes the probable aggregation mechanism of soy glycinin-derived peptides. It schematically shows the quaternary structure of heated glycinin molecules that comprises ~50 % aggregated material, whereas the other ~50 % comprises the glycinin dissociation products (trimers, dimers, monomers and individual acidic and basic polypeptides). The possible different contributions of soluble aggregates and glycinin dissociation products on the aggregation of the glycinin peptides was not studied. Therefore, in the proposed model it is assumed that each form in which glycinin is present prior to hydrolysis participates equally in the subsequent hydrolysis and aggregation.

In the hexameric and trimeric structure of non-heated glycinin the basic polypeptides are more or less buried in the tri- or hexamer, shielded by the hydrophilic acidic polypeptides that face the hydrophilic solution *(24)*, which promotes a good solubility. As the aggregates and monomers formed upon heating are soluble, they still might posses a more or less organized structure with a hydrophilic outside (acidic polypeptides) and a hydrophobic core (basic polypeptide) as shown in Figure 8 and also described in the model of Nagai and Inouye *(17)*.

Upon enzymatic hydrolysis the hydrophilic outside of the soluble glycinin aggregates and glycinin dissociation products (acidic polypeptides) might be preferentially degraded. In contrast, the hydrophobic core of the soluble aggregates and monomers (basic polypeptide) is only degraded to a low extent. The remaining proteinaceous material, being mainly peptides originating from the basic polypeptide and a fraction of the acidic polypeptide, increases in hydrophobicity with increasing DH and at a certain point starts to form aggregates.

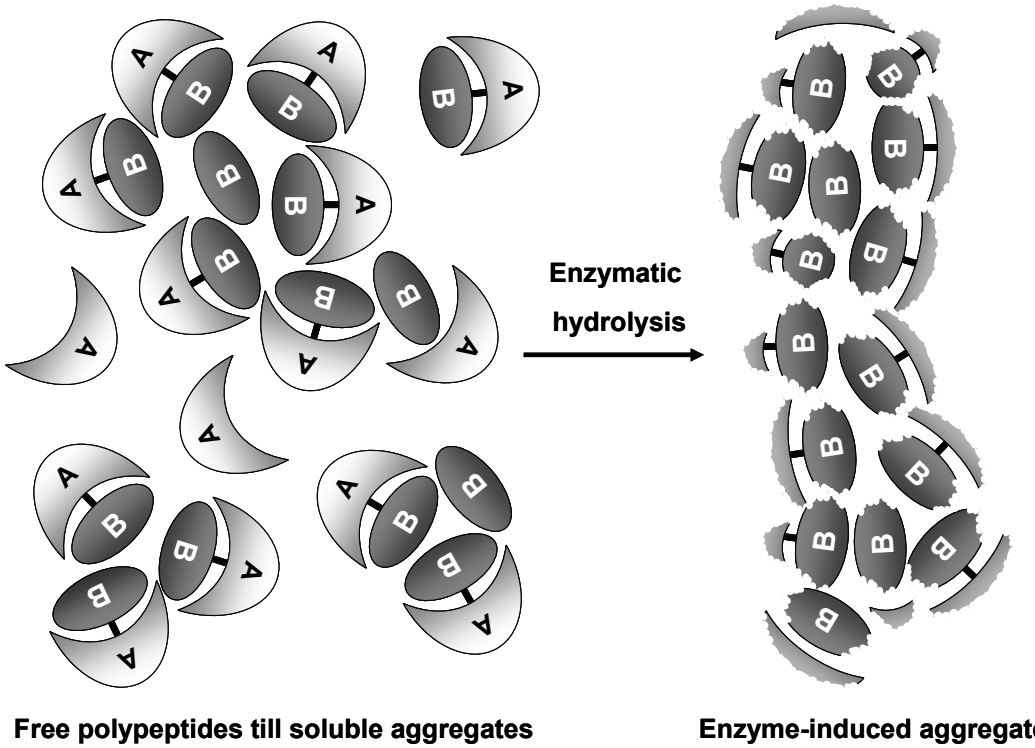


**Figure 8**: Proposed model for aggregate formation upon hydrolysis of soluble glycinin aggregates and its dissociation products.

In conclusion, the results show that aggregation of heated glycinin upon chymotryptic hydrolysis is mainly due to the aggregation of peptides originating from the basic polypeptide. In addition, those parts of the acidic polypeptide in the region where the disulfide bridge connects the basic polypeptides to the acidic polypeptides are also present in the aggregating peptides. Since the aggregation of glycinin is mainly substrate, and to a lesser extent, enzyme-dependent *(14)*, it is likely that enzyme-induced aggregation of glycinin follows a similar mechanism upon hydrolysis with other enzymes.

## ACKNOWLEGEMENTS

The soybeans were kindly provided by Fa. L.I. Frank (Twello, The Netherlands).

## REFERENCES

1. Lee, C.W., and Jo, G.Y. The microstructures of soybean milk curds prepared by different coagulation methods. *J. Food Sci. Nutr.* **1997,** *2*, 259-262.
2. Murata, K.; Kusakabe, I.; Kobayashi, H.; Kiuchi, H.; Murakami, K. Functional-properties of 3 soymilk curds prepared with an enzyme, calcium salt and acid. *Agric. Biol. Chem.* **1988,** *52*, (5), 1135-1139.
3. Inouye, K.; Nagai, K.; Takita, T. Coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2002,** *50*, 1237-1242.
4. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. *J. Agric. Food Chem.* **2005,** *53*, 1031-1038. *Chapter 2 in this thesis*.
5. Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophys.* **2006,** *1*, 178-188. *Chapter 3 in this thesis*.
6. Park, Y.W., Kusakabe, I., Kobayashi, H., and Murakami, K. Production and properties of a soymilk-clotting enzyme system from a microorganism. *Agric. Biol. Chem.* **1985,** *49*, (11), 3215-3219.
7. Park, Y.W.K., H, Kusakabe, I. and Murakami, K. Purification and characterization of soymilk-clotting Enzymes from Bacillus sp. K-295G-7. *Agric. Biol. Chem.* **1987,** *51*, (9), 2343-2349.
8. Kobayashi, H.; Park, Y.W.; Kaneda, M.; Kusakabe, I.; Yoshida, S.; Murakami, K. Substrate-specificity of soymilk-clotting enzyme-I and enzyme-II from *Bacillus Sp.* K-295G-7 toward peptide-substrates. *Agric. Biol. Chem.* **1988,** *52*, (9), 2351-2353.
9. Aoyama, M.; Yasuda, M.; Nakachi, K.; Kobamoto, N.; Oku, H.; Kato, F. Soybean-milk-coagulating activity of *Bacillus pumilus* derives from a serine proteinase. *Appl. Microbiol. Biotechnol.* **2000,** *53*, (4), 390-395.
10. Yasuda, M., Aoyama, M., Sakaguchi, M., Nakachi, K. and Kobamoto, N. Purification and characterization of a soybean-milk-coagulating enzyme from Bacillus pumilus TYO-67. *Appl. Microbiol. Biotechnol.* **1999,** *51*, (4), 474-479.
11. Fuke, Y.; Sekiguchi, M.; Msuoka, H. Nature of stem bromelain treatments on the aggregation and gelation of soybean proteins. *J. Food Sci.* **1985,** *50*, 1283-1288.

12. Mohri, M.; Matsushita, S. Improvement of water absorption of soybean protein by treatment with bromelain. *J. Agric. Food Chem.* **1984,** *32*, 486-490.
13. Zhong, F.; Xin, Y.; Yue, L.; Shoemaker, C.F. Papain-induced gelation of soy glycinin (11S). *J. Food Sci.* **2006,** *71*, (5), E232-E237.
14. Kuipers, B.J.H.; Alting, A.C.; Gruppen, H. Soy glycinin as a dominant source of aggregating peptides using different proteases. *Submitted; Chapter 4 in this thesis*. **2007**.
15. Murata, K.; Kusakabe, I.; Kobayashi, H.; Akaike, M.; Park, Y.W.; Murakami, K. Studies on the coagulation of soymilk-protein by commercial proteinases. *Agric. Biol. Chem.* **1987,** *51*, (2), 385-389.
16. Zhong, F.; Wang, Z.; Xu, S.Y.; Shoemaker, C.F. The evaluation of proteases as coagulants for soy protein dispersions. *Food Chem.* **2007,** *100*, (4), 1371-1376.
17. Nagai, K.; Inouye, K. Insights into the reaction mechanism of the coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2004,** *52*, (15), 4921-4927.
18. Ju, Z.Y.; Otte, J.; Madsen, J.S.; Qvist, B. Effects of limited proteolysis on gelation and gel properties of whey protein isolate. *J. Dairy Sci.* **1995,** *78*, 2119-2128.
19. Creusot, N.; Gruppen, H.; van Koningsveld, G.A.; de Kruif, C.G.; Voragen, A.G.J. Peptide-peptide and protein-peptide interactions in mixtures of whey protein isolate and whey protein isolate hydrolysates. *Int. Dairy J.* **2006,** *16*, (8), 840-849.
20. Beilinson, V.; Chen, Z.; Shoemaker, R.C.; Fischer, R.L.; Goldberg, R.B.; Nielsen, N.C. Genomic organization of glycinin genes in soybean. *Theor. Appl. Genet.* **2002,** *104*, (6-7), 1132-1140.
21. Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallon, B.J.; Fischer, R.L.; Sims, T.L.; Drews, G.N.; Goldberg, R.B. Characterization of the glycinin gene family in soybean. *Plant Cell* **1989,** *1*, (3), 313-328.
22. Fischer, R.L.; Goldberg, R.B. Structure and flanking regions of soybean seed protein genes. *Cell* **1982,** *29*, (2), 651-660.
23. Scallon, B.; Thanh, V.H.; Floener, L.A.; Nielsen, N.C. Identification and characterization of DNA clones encoding Group-II glycinin subunits. *Theor. Appl. Genet.* **1985,** *70*, (5), 510-519.
24. Lakemond, C.M.M.; de Jongh, H.H.J.; Hessing, M.; Gruppen, H.; Voragen, A.G.J. Soy glycinin: Influence of pH and ionic strength on solubility and molecular structure at ambient temperatures. *J. Agric. Food Chem.* **2000,** *48*, (6), 1985-1990.
25. Kuipers, B.J.H.; Bakx, E.J.; Gruppen, H. Functional region identification in proteins by accumulative-quantitative-peptide-mapping using RP-HPLC-MS. **2007**. *Submitted; Chapter 6 in this thesis*.
26. Ilse, D.; Edman, P. Formation of 3-phenyl-2-thiohydantoins from phenylthiocarbamyl amino acids. *Aust. J. Chem.* **1963,** *16*, (3), 411-416.
27. Edman, P. On the mechanism of the phenyl isothiocyanate degradation of peptides. *Acta Chem. Scand.* **1956,** *10*, (5), 761-768.
28. Hewick, R.M.; Hunkapiller, M.W.; Hood, L.E.; Dreyer, W.J. A gas-liquid solid-phase peptide and protein sequenator. *J. Biol. Chem.* **1981,** *256*, (15), 7990-7997.
29. Link, A.J.; Eng, J.; Schieltz, D.M.; Carmack, E.; Mize, G.J.; Morris, D.R.; Garvik, B.M.; Yates, J.R. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999,** *17*, (7), 676-682.
30. Storms, H.F.; van der Heijden, R.; Tjaden, U.R.; van der Greef, J. Optimization of protein identification from digests as analyzed by capillary isoelectric focusing-mass spectrometry. *J. Chromatogr., B* **2005,** *824*, (1-2), 189-200.

31. 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 RP-HPLC-MS analysis. *J. Agric. Food Chem., accepted for publication.* **2007**. *Chapter 5 in this thesis*.
32. Saio, K.; Kamiya, M.; Watanabe, T. Food processing characteristics of soybean-11S and soybean-7S proteins. 1. Effect of difference of protein components among soybean varieties on formation of tofu-gel. *Agric. Biol. Chem.* **1969,** *33*, (9), 1301-1308.
33. Mori, T.; Nakamura, T.; Utsumi, S. Gelation mechanism of soybean 11S-globulin - formation of soluble aggregates as transient intermediates. *J. Food Sci.* **1981,** *47*, (1), 26-30.
34. Yamauchi, F.; Yamagischi, T.; Iwabuchi, S. Molecular understanding of heat-induced phenomena of soybean protein. *Food Rev. Int.* **1991,** *7*, (3), 283-322.
35. Visschers, R.W.; de Jongh, H.H.J. Disulphide bond formation in food protein aggregation and gelation. *Biotechnol. Adv.* **2005,** *23*, (1), 75-80.
36. Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984,** *179*, 125-142.

# CHAPTER 8

# General discussion

## AGGREGATION OF SOY PROTEIN DERIVED PEPTIDES

In the present study the aggregation behavior of soy peptides has been investigated. The main prerequisite for all experiments was that there was no enzyme activity at the moment that the (visible) aggregation of the peptides was induced. This was obtained by uncoupling the hydrolysis from the subsequent aggregation, which is in contrast to all the studies presented in literature (Table 1, Chapter 1) in which *enzyme-induced aggregation* was studied under conditions at which the aggregation coincides with hydrolysis. Only Nagai and Inouye *(1)* stop the hydrolysis at a certain point to study the subsequent aggregation of the hydrolysate. The uncoupling in the present study is done by keeping the pH stable during the hydrolysis (pH-stat method), followed by inhibition of the enzyme at the moment that the DH of interest is reached. Subsequently, the hydrolysate obtained (with a fixed DH) was used to study its aggregation behavior as a function of pH. Using this approach enables a control of pH and DH during the aggregation process.

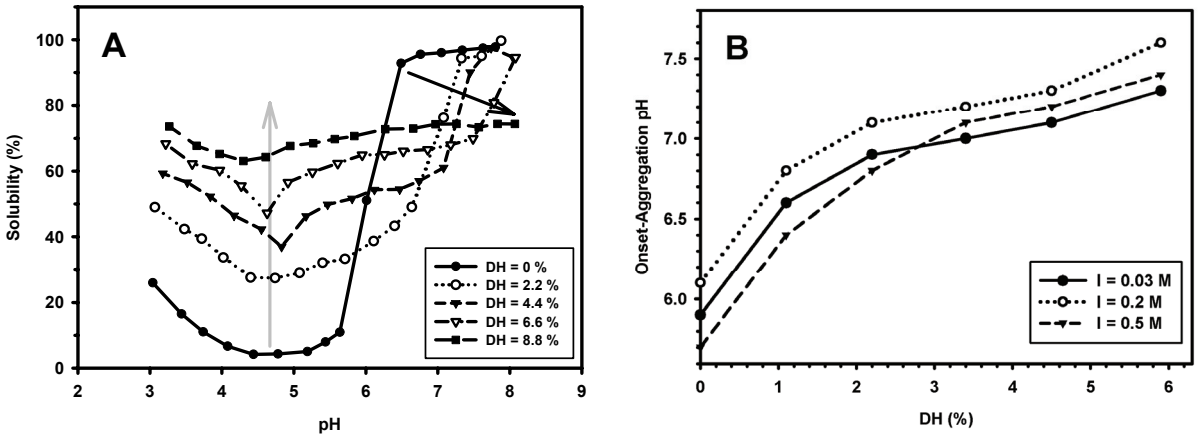


**Figure 1**: Solubility (4 % (w/w) protein, $I$ = 0.03 M) of SPI hydrolysates as a function of pH, determined using the Dumas method (**1A**) and the aggregation pH as a function of DH for different ionic strengths (**1B**; data deduced from Chapter 2, Table 2). Hydrolysates were obtained by hydrolysis of heated SPI solutions using subtilisin Carlsberg.

The first results obtained by uncoupling the hydrolysis from subsequent aggregation, clearly showed the relation between the DH and the pH at which aggregates are formed (Chapter 2). It was shown that with increasing DH, the solubility decreased at higher pH values, as indicated by the black arrow in Figure 1A. In Figure 1B the onset-aggregation pH as a function of DH is shown for different ionic strengths. The ionic strength hardly influenced this change in onset-aggregation pH. This can be observed in Figure 1B, which is obtained from turbidity experiments at various ionic strengths and DHs (Chapter 2). In the pH range of 6.5-8.0, the solubility of the hydrolyzates is lower than the solubility of the parental proteins. However, around the *pI* of the parental protein (pH 5.34 and 5.23 for glycinin and β-

conglycinin, based on their average amino acid composition as found in the Swiss-Prot database), an increase in solubility can be observed (indicated with the grey arrow). This is in agreement with the general observation that enzymatic hydrolysis improves protein solubility *(2)*.

Glycinin-derived peptides appeared to be the driving force in the strong aggregation of SPI at neutral pH (Chapter 3). This was extensively investigated for subtilisin Carlsberg (Chapter 3). Although not proven, the results obtained in Chapter 4 indicate that this statement is also valid for the use of a number of other proteases, not including trypsin. In Chapter 7 a mechanism is proposed for the aggregation of glycinin derived peptides upon hydrolysis. The mechanism is based on the observations that mainly peptides that originate from the basic polypeptides of glycinin are present in the aggregating peptides. In addition to this, also peptides originating from the acidic polypeptide around the location of the disulfide bridge that connects the acidic to the basic polypeptide have the tendency to aggregate.

Upon heating, prior to hydrolysis, the glycinin was partly aggregated as well as partly dissociated into probably its trimers or monomers or even into its individual polypeptides, as observed by size exclusion chromatography (Chapter 7). Although it was not part of the present study, it would be interesting to study to what extent the quaternary structure of the proteins, present prior to hydrolysis, are of influence for the subsequent hydrolysis. In other words, to study whether the soluble aggregates are more susceptible to aggregation upon hydrolysis than the hexamers, trimers, monomers or the individual acidic and basic polypeptides. Understanding the preferred quaternary structure required for strong aggregation upon hydrolysis enables optimizing the pretreatment of soy protein isolates, ultimately resulting in optimum aggregation.

In this thesis, there was no focus on the nature of interactions that play a dominant role in the aggregation of peptides, although we showed that electrostatic interactions play a minor role on the pH at which the peptides start to aggregate, as can be observed from Figure 1B. Previous studies have shown that hydrophobic interactions are the driving force in the aggregation mechanism of enzyme-induced aggregation, whereas the formation of disulfide bridges played a minor role *(1, 3, 4)*. Our results (Chapter 7) point into the same direction. The regions from which the aggregating peptides mainly originate are predominantly hydrophobic. Although studies on the nature of the interactions between the proteins were not performed in this study, the approach of uncoupling the hydrolysis from the aggregation is very suitable for these studies. This is because one hydrolysate batch can be used for several aggregation experiments (e.g. different temperatures or ionic strengths).

## RESISTANCE OF BASIC POLYPEPTIDE TOWARDS ENZYMATIC HYDROLYSIS

As already mentioned above, a second result of this study was that peptides from the basic polypeptides of glycinin are the main constituents of the aggregating proteinaceous material (Chapter 7). Since the aggregating peptides have molecular masses of 5,000 to 25,000 Da, it is likely that the basic polypeptide is only degraded to a low extent at a DH of 2.2 %.

Usually soy proteins are heated prior to consumption. This heating aims at destroying the Bowman Birk and Kunitz protease inhibitors to enable the use of the full nutritional potential of soybean proteins after consumption *(5)*. For this reason, in the study presented in this thesis the soy protein solutions were also heated prior to proteolytic degradation. As an extension, next to the heated glycinin, also the hydrolysis of the non-heated glycinin by chymotrypsin was studied with respect to the degradation pattern of the proteins. The results of these studies have not yet been discussed in the previous chapters.

For the heated and non-heated glycinin the size distributions of their aggregating peptides at pH 4.8 were studied for several DH values using size exclusion chromatography as shown in Figure 2. The acidic polypeptides elute around 7.6 and 9.4 mL, the basic polypeptides elute around 8.4 mL, and a minor amount of β-conglycinin present in the glycinin preparation elutes around 6.8 mL (Chapter 2). These results show that for the non-heated glycinin, up to a DH of 2.2 % the basic polypeptides are resistant towards chymotryptic hydrolysis. In contrast, for the heated glycinin the basic and the acidic polypeptide are both degraded. The observation that the basic polypeptides show a higher resistance towards enzymatic hydrolysis is in agreement with several other studies *(3, 6-9)*.

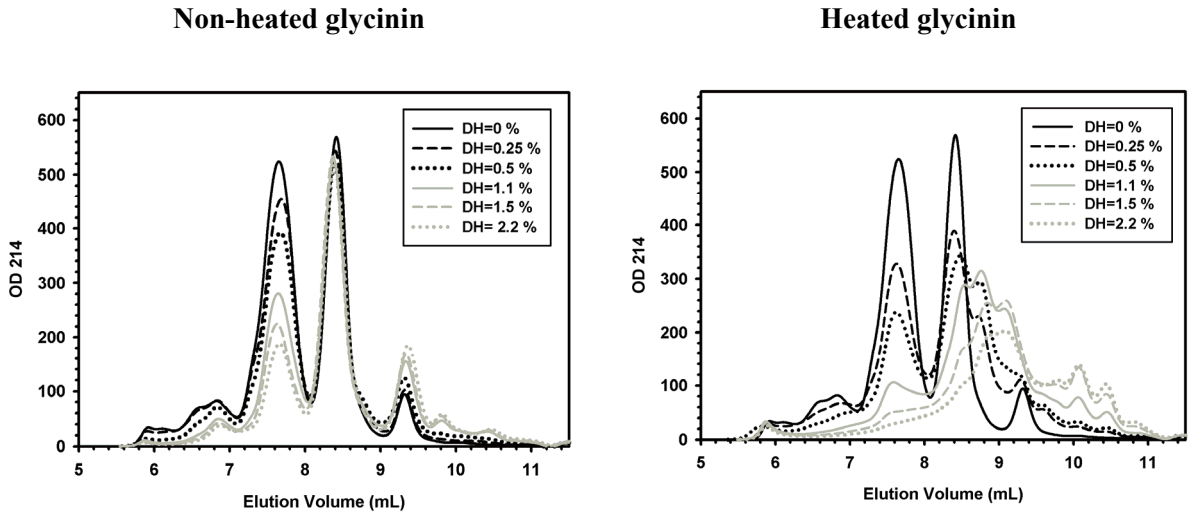


**Figure 2**: Size exclusion chromatograms under denaturing conditions (6M urea, 30 % (v/v) acetonitrile, 0.1 % (v/v) TFA) of aggregating peptides after chymotryptic hydrolysis at pH 4.8 of non-heated and heated glycinin at various DH values. The samples were denatured and reduced prior to analysis. The peaks around 7.6 and 9.4 mL represent the acidic polypeptide. The peak around 8.4 mL represents the basic polypeptide (Chapter 2).

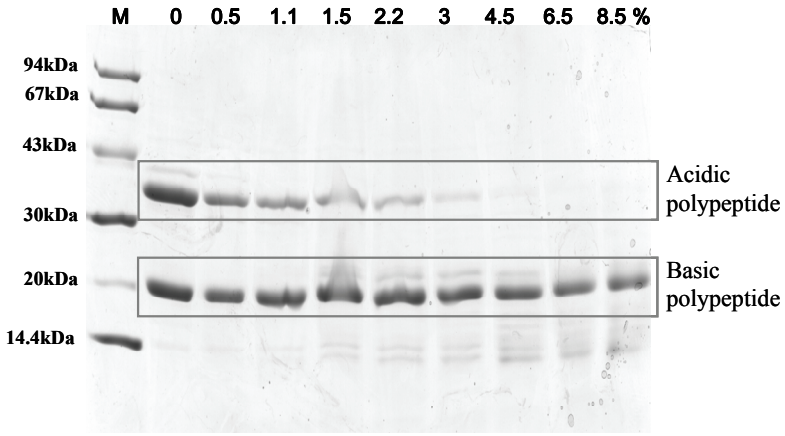


**Figure 3**: SDS-PAGE of marker (M) aggregating peptides at pH 4.8 for various DH values (0-8.5 %) of non-heated glycinin after hydrolysis with chymotrypsin.

At DH values higher than 2.2 % the basic polypeptide of the non-heated glycinin was also degraded as observed by size exclusion chromatography (no further results shown) and by the weak decrease in density of the ~ 20 kDa protein band as present at the SDS-PAGE gel shown in Figure 3. Nevertheless, at a DH of 8.5 % still a relative high proportion of the intact basic polypeptide was present.

In the non-heated glycinin, probably the acidic polypeptide is shielding the basic polypeptide rather well, thereby preventing its degradation. This occurs till the point at which the basic polypeptide becomes exposed to the solvent and therewith becomes better susceptible to the enzyme. Upon heating the quaternary structure of glycinin is changed in such a way that the acidic polypeptide does not protect the basic polypeptide sufficiently anymore, resulting in a better susceptibility of the basic polypeptide towards chymotryptic hydrolysis.

## USE OF SPIs WITH DIFFERENT GLYCININ AND β-CONGLYCINN RATIOS

A third result of the present study was the ease of preparation of SPIs having different ratios of glycinin and β-conglycinin (Chapter 3). These SPIs can be used to find the optimum ratio for a corresponding food product. For tofu processing it is known that the glycinin/β-conglycinin ratio is an important parameter for the final product attributes such as yield and firmness *(10, 11)*. However, it should be noted that not only the glycinin/β-conglycinin ratio defines the quality of the final product, but also the method of processing plays a significant role *(12)*.

The strong aggregating properties of glycinin peptides can be used in food products where strong aggregation is required, such as in the preparation of curd for tofu. For optimum use of this property it is important to work with soybeans or soy protein isolates that have a high glycinin content. This high glycinin content can be obtained by selecting soybean varieties with a high level of glycinin or by choosing an appropriate procedure for SPI preparation as shown in Chapter 3. Although purification of glycinin and β-conglycinin based on differential

acidic precipitation has been used extensively, to our knowledge this acidification has not been applied before with the aim to obtain SPIs with different glycinin/β-conglycinin ratios.
In cultivated soybeans there is not much variation in the glycinin/β-conglycinin ratios, as visually observed from SDS-PAGE analysis (13, 14). Densitometric analysis of various soybean cultivars showed a glycinin/β-conglycinin ratio varying from 1.31 to 2.45 (10-12, 15). This variation in ratio is much lower than observed in related protein sources such as pea, in which the ratio of legumin (analogue to soy glycinin) / vicilin (analogue to soy β-conglycinin) varied from 0.13 to 0.83 (16). To obtain soy varieties with higher glycinin/β-conglycinin ratio breading trials should be considered, as these breeding trials have been reported to result in soybeans with a glycinin/β-conglycinin ratio of 6.03 and 3.82 (17). In addition, the use of wild type varieties can be considered, since these are known for their higher genetic variation than cultivated soybeans (18).
As stated above by making use of appropriate conditions SPIs with different glycinin/β-conglycinin ratios can be obtained. This was done by making use of the difference in solubility of glycinin and β-conglycinin as a function of pH *(19)*. The difference in solubility as a function of pH for glycinin and β-conglycinin is depicted in Figure 4. At various pH values, the ratios between insoluble glycinin and β-conglycinin differ. Based on this difference, soy protein isolates can be prepared having different glycinin/β-conglycinin ratios. Consequently, this can be done without the need of reconstituting purified glycinin and β-conglycinin preparations to SPIs with appropriate glycinin/β-conglycinin ratios. Three acidification procedures were used all starting with the pH 8.0 soluble soy protein fraction (A, B, and C). After precipitating the proteins at a given pH, the pH of its supernatant was further decreased to the next precipitation pH. The glycinin/β-conglycinin ratios found in Chapter 3 varied from 11.50 to 0.28, which is a much higher variation than the ones observed in soybean cultivars as described above. This approach is of interest due to the simplicity to obtain isolates with various glycinin/β-conglycinin ratios from the same genetic soybean variety. It enables studying the contribution of glycinin or β-conglycinin to a functional property of interest. For commercial applications the advantage of preparation of specific SPIs is that different SPIs can be prepared from one batch of soybeans, thereby avoiding the need for using several soybean cultivars in the process facilities, which keeps logistics more simple.
When the approach of interval precipitation is used to obtain glycinin-rich isolates, the fraction that is left is enriched in β-conglycinin. This fraction can as well be of interest, since β-conglycinin has different aggregation properties compared to glycinin *(20, 21)*. In addition to this, β-conglycinin also receives increased attention with respect to its physiological functions *(22-24)*, making β-conglycinin-rich SPIs interesting as health improving ingredients.

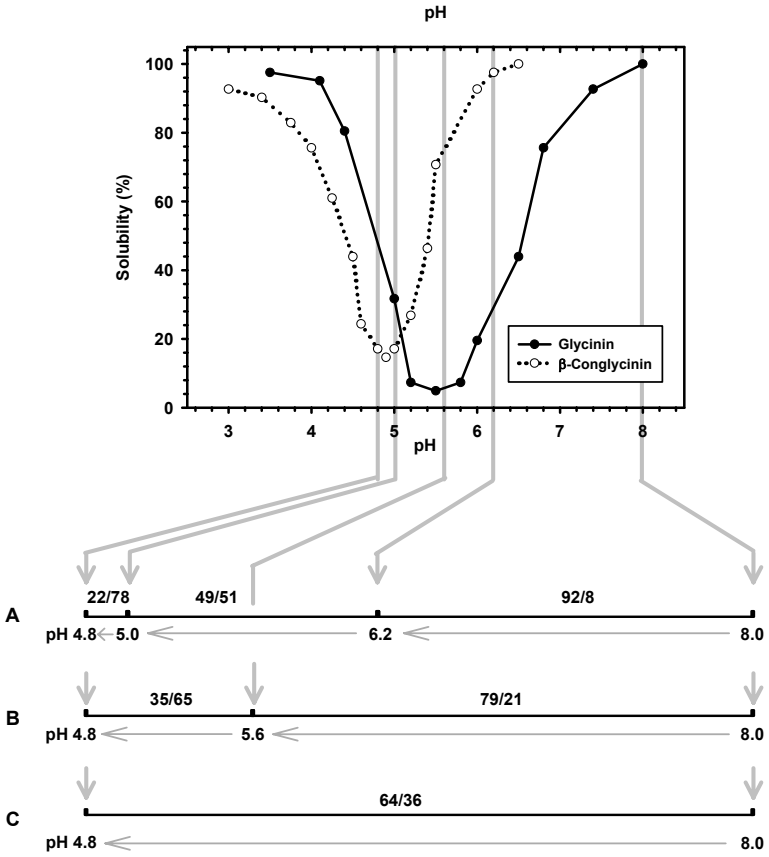


**Figure 4**: Solubility profile of soy glycinin and β-conglycinin as a function of pH combined with its 3 different sequential precipitation intervals used (A, B, and C) to obtain SPIs having different glycinin/β-conglycinin ratios (e.g. 64/36 for the precipitation interval C).

## PEPTIDE AND PROTEIN QUANTIFICATION AT 214 NM

A fourth result of this study is the establishment of a method to quantify peptides based on their absorbance at 214 nm, once identified in a protein hydrolysate. With the increased interest in bio-active peptides *(25-27)*, there is a demand for rapid methods to quantify a range of individual peptides. For proteins, quantification based on their absorbances at 280 nm is a commonly used method. The molar extinction coefficient of a protein at 280 nm is only defined by the proportion of tyrosine, tryptophan, and cystine *(28)*. This method can only be applied for peptides containing tryptophan or tyrosine *(29)*, which is certainly not always the case as tryptophan and tyrosine are usually present in low proportions in proteins; 1.3 and 3.2 %, respectively *(30)*. As a consequence, at 280 nm not all peptides can be quantified. Quantification of all peptides present in a hydrolysate, therefore, requires a method at a wavelength at which each peptide absorbs light. The best approach for this will be by measuring the absorbance around 214 nm, as the peptide bond absorbs light around this wavelength.

As no complete database was available for the contribution of all amino acids and the peptide bond to the absorbance of a protein or peptide at 214 nm, in this study the contribution of

each amino acid and the peptide bond were measured (Chapter 5). We have shown that, based on the known amino acid sequence of peptides and proteins, the molar extinction coefficient can be predicted rather well. Especially when one considers that no easy-to-use alternative is at hand.

Protein quantification at 214 nm has been described more often in literature *(31-34)*. Protein quantification at 214 nm appears to be suitable for determining the protein concentration when the amino acid composition of the protein is not known. This is due to the fact that the amino acid composition of most proteins is rather similar. At 214 nm, therefore, the absorbance correlates well with the concentration (w/w) for most proteins. This correlation can be observed in Table 1. In this table, for a given set of proteins and peptides, the ratio between the measured molar extinction coefficient and the calculated molar extinction coefficients only based on the number of peptide bonds of the proteins and peptides are given. For the calculation of the ratios the molar extinction coefficients as determined in chapter 5 are used. The higher the ratio, the higher the relative contribution of the amino acid side chains to the contribution of the peptide bond. For the limited set of proteins used in this study this ratio is on average 2.23 (±0.30). For the peptides analyzed the average ratio is 2.49 (± 1.13).

The relative low standard deviation for the proteins shows that the absorbance at 214 nm will give a rather good estimation of the protein concentration. The higher standard deviation for peptides illustrates well that the absorbance of peptides much more requires a correction for the amino acid composition of the peptides due to the strong variation of amino acid composition between peptides. This illustrates the importance of the establishment of the method described in Chapter 5.

**Table 1**: Ratios between measured molar extinction coefficients and calculated molar extinction coefficient only based on the absorbance of the peptide bonds.

| **Proteins** | $\varepsilon_{measured}/\varepsilon_{peptide\ bonds}$ | **Peptides** | $\varepsilon_{measured}/\varepsilon_{peptide\ bonds}$ |
|---|---|---|---|
| Glycinin | 2.19 | RGDS | 1.07 |
| β-Conglycinin | 1.79 | RINKKIEK | 1.22 |
| α-Lactalbumin | 2.60 | EQLSTSEENSK | 0.91 |
| β-Lactoglobulin | 2.04 | GYG | 3.16 |
| BSA | 2.15 | YGGFLRR | 2.79 |
| α-Casein | 2.66 | YIPIQYVLSR | 2.65 |
| Lysozyme | 2.19 | GPRP | 3.24 |
| | | RPPGFSP | 3.58 |
| | | RPPGFSPFR | 3.81 |

## DETERMINING THE RATIO OF GLYCININ SUBUNITS.

A fifth result of the present study is the use of an easy and direct method to determine the ratios between the 5 subunits G1 to G5 of glycinin by N-terminal sequencing as described in Chapter 7. The advantage of this method is that prior to analysis, no separation in the different subunits is required.

In literature gel electrophoresis and ion exchange have been reported to be suitable methods to separate individual subunits and quantify their ratios after reduction of their disulfide bonds. Mori and co-workers *(14)* studied the differences in subunit composition of glycinin among different soybean cultivars using gel-electrophoresis. The individual subunits were well separated and their relative proportion was determined densitometrically. Later studies showed that separating all polypeptides on SDS-PAGE is difficult to achieve, due to the overlap in masses of the acidic polypeptides A1a, A2, A1b and A4, which vary from 30 to 32.6 kDa, and the fact that the basic polypeptides, each having molecular masses around 20 kDa. This results in an overlap of protein bands in the SDS-PAGE gel *(10, 35)*. Also, ion exchange chromatography can be used to separate the different acidic polypeptides rather well *(36, 37)*. The peak areas reveal the ratios between the different polypeptides. However, for several peaks some heterogeneity was observed, which complicates accurate determination of the ratios.

A drawback of the determination of the different subunit ratios using gel-electrophoresis or ion exchange chromatography is that next to the separation, a second step is required to determine which polypeptide corresponds to which peak in an ion exchange chromatogram or to which band in a gel. This is in contrast to the method of N-terminal sequencing used in Chapter 7. For this method only the known N-terminal sequences and a relative pure sample of glycinin are the main requirements, as in this method the identification and quantification are integrated.

When the N-termini of all polypeptides are known, N-terminal sequencing showed to be a relatively quick and accurate approach to quantify the ratio between the various polypeptides. Especially, when compared to the other approaches reported in literature. The high accuracy is well demonstrated based on the control that for corresponding N-termini the sum of the acidic polypeptides should be equal to the sum of the amount of basic polypeptides, as shown in Chapter 7.

## GEL STRENGTH MODIFYING PROPERTIES OF SOY PEPTIDES

The sixth result of this study was the elucidation of the opposite effect of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of peptides in a hydrolyzed SPI. The hypothesis was raised (Chapter 4) that peptides of β-conglycinin have aggregation-preventing properties. In SPI they are supposed to prevent the aggregation of the glycinin-derived peptides that have a strong tendency to aggregate. Such aggregation-preventing properties were previously described for the inhibition of blood platelet aggregation by milk- *(38-41)*

and soy protein-derived peptides *(42)*. The peptides that have blood platelet aggregation preventing properties were reported to be hydrophilic and to consist of small peptides up to 11 amino acids.

As peptides have the ability to inhibit aggregation, it is likely that these peptides can also change the aggregation of proteins upon heating, which can be a means to modify the gelling properties of proteins. This is in agreement with observation by Sodini and co-workers *(43)*, who observed that dairy yoghurts, supplemented with hydrolysates, showed a decreased viscosity with increasing supplementation. To study this effect for soy protein-derived peptides we have explored the possibilities of addition of hydrolysates with a DH of 6.5 % derived from SPI, glycinin and β-conglycinin to non-hydrolysed SPI, to tune the properties of heat-induced gels. The hydrolysates were added to a 6.8 % (w/w) SPI solution. The peptide concentrations were 5 and 10 % (w/w) relative to the amount of intact protein. The heat-induced gelation was studied by gradually increasing the temperature up to 90°C. After heating, gels were formed, which shows that the peptides were not able to prevent aggregation. The gel strength after heating was determined using a Texture Analyzer, of which the results can be observed in Figure 5.

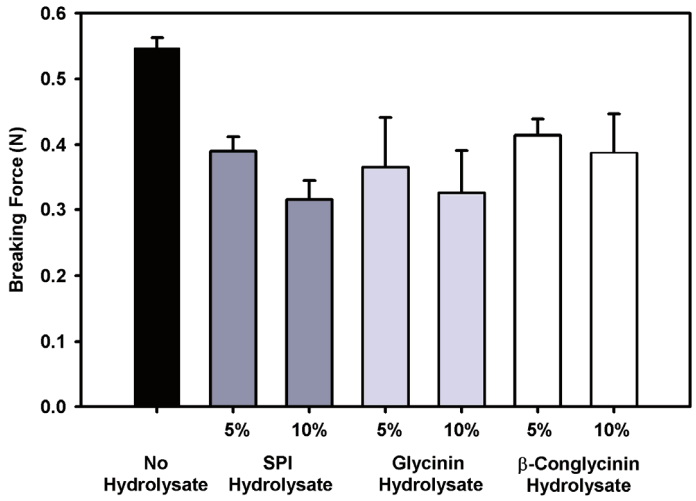


**Figure 5**: Breaking force at the moment of breaking of heat-induced SPI gels (6.8 % protein (w/w) not containing peptides, or supplemented with SPI, glycinin, or β-conglycinin hydrolysates (DH = 6.5 %), at concentrations of 5 and 10 % relative to the SPI concentration. Standard deviations are based on 5 measurements. The diameter of the probe (with a round head) was 5 mm. The moving speed of the probe was 0.2 mm/s.

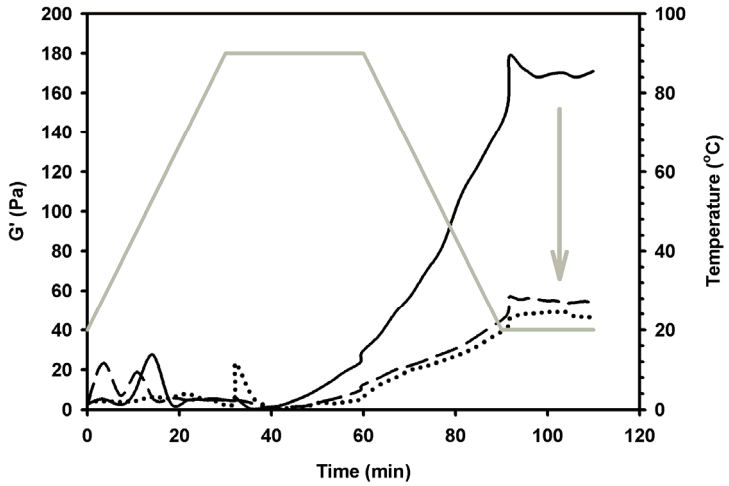


**Figure 6**: Development of the G' as a function of time during heat-induced gelation of SPI (——), SPI + SPI hydrolysate (DH = 6.5 %) up to 5% (w/w) (— — —) and 10 % (w/w) (• • • •) relative to the SPI concentration (6.8 % w/w). The storage modulus (G') was measured at a constant strain of 0.1 and an angular frequency of 0.63 rad/s. The grey line represents the development of the temperature.

Upon the addition of peptides the force needed to break heat-induced soy protein gels containing hydrolysates, was significantly lower when compared to gels not supplemented with hydrolysates (Figure 5). There was no clear difference in the decrease in breaking force between hydrolysates originating from SPI, glycinin or β-conglycinin. In addition to measuring the breaking force, for SPI gels containing SPI-derived hydrolysates, the development of the elastic modulus (G') was followed in time during gradual heating and cooling using a rheometer (Figure 6). These results confirm the results of the breaking force measurement (Figure 5). They show a decrease of in gel strength from ~175 Pa in the absence of hydrolysates to ~50 Pa when hydrolysates are added. Analysis of the protein content of the supernatants of the gels after centrifugation revealed that peptide addition did not decrease the amount of aggregated proteins upon heating. The fact that the amount of aggregating material did not decrease upon addition of peptides but the gel strength did indicates that the peptides change the interactions and therewith change the network structure between intact proteins upon heating.

In the studies on the inhibition of blood platelet aggregation, the peptides responsible for aggregation were reported to be small and hydrophilic. The high hydrophilicity of the peptides reveals that probably electrostatic interactions play a dominant role in the aggregation preventing effect.
Recently, Creusot *(44)* reported that peptides can induce aggregation of proteins, due to the formation of a kind of glue network by the peptides that encloses intact proteins. This results in an additional aggregation of intact proteins upon changing the pH. This is due to the strong aggregating properties of the peptides themselves under the conditions used as the proteins alone do not aggregate at the conditions applied. This phenomenon probably does not play a

role in the heat-induced aggregation of intact SPI in the presence of SPI-derived peptides, since in the present study the soy peptides do not have extra aggregating properties at the conditions used.

Next to the hydrophobic interactions, the formation of disulfide bridges might increase the gel strength *(45)*. Addition of cysteine to proteins prior to heat-induced gelation results in a decrease in molecular weight of heat-induced aggregates due to blocking of free sulfide groups. This blocking prevents the formation of new intermolecular disulfide bridges and so reduces the gel strength *(46)*. This was also observed in the present study upon addition of cysteine to SPI prior to heating (no further results shown). Peptides, containing free sulfhydryl groups (from cysteine), may also have this property to react with a free sulfhydryl group of an intact protein, preventing subsequent disulfide formation with other intact proteins, and as a result decreases the gel strength. Although blocking of free sulfide groups of intact protein by free sulfide groups present in peptides can be in important parameter influencing gel strength, it is not the only reason for the decreased gel strength as observed in Figure 5. β-Conglycinin contains 0 or only 1 cysteine residue and glycinin contains 6 or 8 cysteine residues. Taking into account the low amount of cysteine present in β-conglycinin, and its relative low quantity (5 or 10 % (w/w) relative to the amount of intact protein) added to the SPI, it is not likely that the observed decrease in gel strength as a result of addition of β-conglycinin hydrolysates to SPI is only due to the blocking of free sulfhydryl groups.

A hypothetical mechanism of how peptides may decrease the gel strength of heat-induced gels is the release of amphipathic "gel strength modifying" peptides. The amino acid sequence of glycinin and β-conglycinin both contain distinct hydrophilic and hydrophobic regions as shown in Chapter 1. It can, therefore, be speculated that during hydrolysis of these proteins, amphipathic peptides are formed.

Upon heating, proteins unfold, resulting in an increase in exposure of the hydrophobic interior to the surface of the protein as schematically presented in Figure 7A. Due to hydrophobic interactions between exposed hydrophobic groups, proteins can form aggregates (Figure 7B), which might results in a gel network, further enhanced by disulfide bridge formation. In the presence of amphipathic peptides, these exposed hydrophobic groups of unfolded proteins can be partly "blocked" due to hydrophobic interactions between hydrophobic regions of the protein and peptide (Figure 7C). This results in a protein-peptide complex with an increased net hydrophilic surface, subsequently resulting in a decrease in aggregate size and a decrease in gel strength. The amphipathic peptides might be different from the ones being responsible for the aggregation of the peptides itself (Chapter 7) since these peptides are predominant hydrophobic.

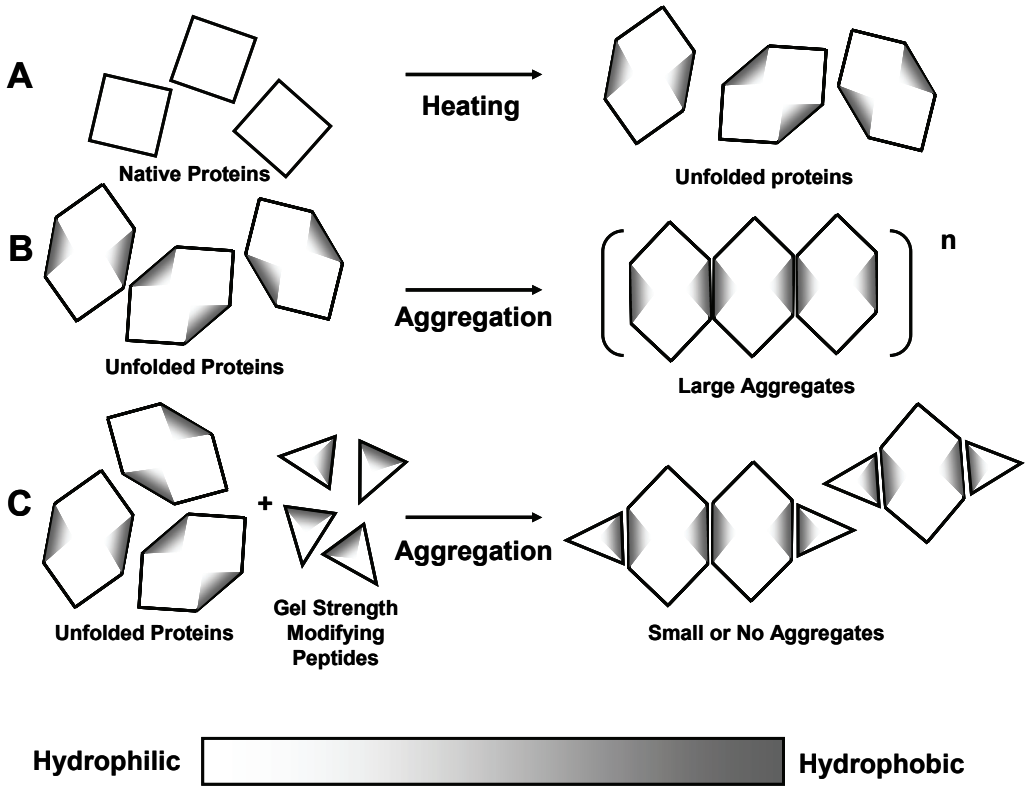


**Figure 7**: Tentative model of the influence of gel strength modifying peptides on the aggregation of intact proteins. Heating of proteins leads to unfolding (**A**), resulting in the exposure of hydrophobic groups, followed by aggregation (**B**). Upon addition of gel strength modifying peptides the aggregation mechanism of proteins is changed (**C**).

As shown above, the ability of soy protein hydrolysates to decrease the gel strength of heat-induced intact soy protein gels can be of interest for food industry as it can be an interesting means to modify the texture of food products.

## REFERENCES

1. Nagai, K.; Inouye, K. Insights into the reaction mechanism of the coagulation of soy protein isolates induced by subtilisin Carlsberg. *J. Agric. Food Chem.* **2004,** *52*, (15), 4921-4927.
2. Adler-Nissen, J. Enzymatic hydrolysis of proteins for increased solubility. *J. Agric. Food Chem.* **1976,** *24*, (6), 1090-1093.
3. Mohri, M.; Matsushita, S. Improvement of water absorption of soybean protein by treatment with bromelain. *J. Agric. Food Chem.* **1984,** *32*, 486-490.
4. Zhong, F.; Xin, Y.; Yue, L.; Shoemaker, C.F. Papain-induced gelation of soy glycinin (11S). *J. Food Sci.* **2006,** *71*, (5), E232-E237.
5. Liener, I.E. Implications of antinutritional components in soybean foods. *Crit. Rev. Food Sci. Nutr.* **1994,** *34*, 31-67.
6. Marsman, G.J.P.; Gruppen, H.; Mul, A.J.; Voragen, A.G.J. In vitro accessibility of untreated, toasted, and extruded soybean meals for proteases and carbohydrases. *J. Agric. Food Chem.* **1997,** *45*, (10), 4088-4095.

7. Dereu, J.C.; Tenwolde, R.M.; Degroot, J.; Nout, M.J.R.; Rombouts, F.M.; Gruppen, H. Protein hydrolysis during soybean tempe fermentation with *Rhizopus oligosporus*. *J. Agric. Food Chem.* **1995,** *43*, (8), 2235-2239.
8. Romagnolo, D.; Polan, C.E.; Barbeau, W.E. Degradability of soybean-meal protein-fractions as determined by sodium dodecyl sulfate-polyacrylamide gel-electrophoresis. *J. Dairy Sci.* **1990,** *73*, (9), 2379-2385.
9. Aoyama, M.; Yasuda, M.; Nakachi, K.; Kobamoto, N.; Oku, H.; Kato, F. Soybean-milk-coagulating activity of *Bacillus pumilus* derives from a serine proteinase. *Appl. Microbiol. Biotechnol.* **2000,** *53*, (4), 390-395.
10. Mujoo, R.; Trinh, D.T.; Ng, P.K.W. Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. *Food Chem.* **2003,** *82*, (2), 265-273.
11. Ji, M.P.; Cai, T.D.; Chang, K.C. Tofu yield and textural properties from three soybean cultivars as affected by ratios of 7S and 11S proteins. *J. Food Sci.* **1999,** *64*, (5), 763-767.
12. Cai, T.D.; Chang, K.C. Processing effect on soybean storage proteins and their relationship with tofu quality. *J. Agric. Food Chem.* **1999,** *47*, (2), 720-727.
13. Hughes, S.A.; Murphy, P.A. Varietal influence on the quantity of glycinin in soybeans. *J. Agric. Food Chem.* **1983,** *31*, 376-379.
14. Mori, T.; Utsumi, S.; Inaba, H.; Kitamura, K.; Harada, K. Differences in subunit composition of glycinin among soybean cultivars. *J. Agric. Food Chem.* **1981,** *29*, 20-23.
15. Fehr, W.R.; Hoeck, J.A.; Johnson, S.L.; Murphy, P.A.; Nott, J.D.; Padilla, G.I.; Welke, G.A. Genotype and environment influence on protein components of soybean. *Crop Science* **2003,** *43*, (2), 511-514.
16. Tzitzikas, E.N.; Vincken, J.P.; De Groot, J.; Gruppen, H.; Visser, R.G.F. Genetic variation in pea seed globulin composition. *J. Agric. Food Chem.* **2006,** *54*, (2), 425-433.
17. Kitamura, K. Breeding trials for improving the food-processing quality of soybeans. *Trends Food Sci. Technol.* **1993,** *4*, (3), 64-67.
18. Fukuda, T.; Maruyama, N.; Kanazawa, A.; Abe, J.; Shimamoto, Y.; Hiemori, M.; Tsuji, H.; Tanisaka, T.; Utsumi, S. Molecular analysis and physicochemical properties of electrophoretic variants of wild soybean Glycine soja storage proteins. *J. Agric. Food Chem.* **2005,** *53*, 3658-3665.
19. Thanh, V.H.; Shibasaki, K. Major proteins of soybean seeds. A straightforward fractionation and their characterization. *J. Agric. Food Chem.* **1976,** *24*, 1117-1121.
20. Nagano, T.; Mori, H.; Nishinari, K. Effect of heating and cooling on the gelation kinetics of 7S globulin from soybeans. *J. Agric. Food Chem.* **1994,** *42*, (7), 1415-1419.
21. Saio, K.; Watanabe, T. Differences in functional properties of 7S and 11S soybean proteins. *J. Texture Stud.* **1978,** *9*, (1-2), 135-157.
22. Adams, M.R.; Golden, D.L.; Franke, A.A.; Potter, S.M.; Smith, H.S.; Anthony, M.S. Dietary soy β-conglycinin (7S globulin) inhibits atherosclerosis in mice. *J. Nutr.* **2004,** *134*, (3), 511-516.
23. Moriyama, T.; Kishimoto, K.; Nagai, K.; Urade, R.; Ogawa, T.; Utsumi, S.; Maruyama, N.; Maebuchi, M. Soybean β-conglycinin diet suppresses serum triglyceride levels in normal and genetically obese mice by induction of β-oxidation, downregulation of fatty acid synthase, and inhibition of triglyceride absorption. *Biosci., Biotechnol., Biochem.* **2004,** *68*, (2), 352-359.
24. Aoyama, T.; Kohno, M.; Saito, T.; Fukui, K.; Takamatsu, K.; Yamamoto, T.; Hashimoto, Y.; Hirotsuka, M.; Kito, M. Reduction by phytate-reduced soybean β-conglycinin of plasma triglyceride level of young and adult rats. *Biosci., Biotechnol., Biochem.* **2001,** *65*, (5), 1071-1075.

25. Korhonen, H.; Pihlanto, A. Bioactive peptides: Production and functionality. *Int. Dairy J.* **2006,** *16*, (9), 945-960.
26. Korhonen, H.; Pihlanto, A. Food-derived bioactive peptides - Opportunities for designing future foods. *Curr. Pharm. Des.* **2003,** *9*, (16), 1297-1308.
27. Li, G.H.; Le, G.W.; Shi, Y.H.; Shrestha, S. Angiotensin I-converting enzyme inhibitory peptides derived from food proteins and their physiological and pharmacological effects. *Nutr. Res.* **2004,** *24*, (7), 469-486.
28. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption-coefficient of a protein. *Protein Sci.* **1995,** *4*, (11), 2411-2423.
29. Moffatt, F.; Senkans, P.; Ricketts, D. Approaches towards the quantitative analysis of peptides and proteins by reversed-phase high-performance liquid chromatography in the absence of a pure reference sample. *J. Chromatogr., A* **2000,** *891*, (2), 235-242.
30. Creighton, T.E. *Proteins. Structures and molecular properties*. Freeman: New York, 1993.
31. Goldfarb, A.R.; Saidel, L.J.; Mosovich, E. The ultraviolet absorption spectra of proteins. *J. Biol. Chem.* **1951,** *193*, (1), 397-404.
32. Mayer, M.M.; Miller, J.A. Photometric analysis of proteins and peptides at 191-194 mmu. *Anal. Biochem.* **1970,** *36*, (1), 91-100.
33. Waddell, W.J. A simple ultraviolet spectrophotometric method for the determination of protein. *J. Lab. Clin. Med.* **1956,** *48*, (2), 311-314.
34. Webster, G.C. Comparison of direct spectrophotometric methods for measurement of protein concentration. *Biochim. Biophys. Acta* **1970,** *207*, (2), 371-373.
35. Zhang, G.Y.; Hayashi, Y.; Matsumoto, S.; Matsumura, Y.; Mori, T. Molecular species of glycinin in some soybean cultivars. *Phytochemistry* **2002,** *60*, (7), 675-681.
36. Kitamura, K.; Takagi, T.; Shibasaki, K. Subunit structure of soybean 11S globulin. *Agric. Biol. Chem.* **1976,** *40*, (9), 1837-1844.
37. Moreira, M.A.; Hermodson, M.A.; Larkins, B.A.; Nielsen, N.C. Partial characterization of the acidic and basic polypeptides of glycinin. *J. Biol. Chem.* **1979,** *254*, (19), 9921-9926.
38. Rutherfurd, K.J.; Gill, H.S. Peptides affecting coagulation. *Br. J. Nutr.* **2000,** *84*, S99-S102.
39. Fiat, A.M.; Miglioresamour, D.; Jolles, P.; Drouet, L.; Sollier, C.B.D.; Caen, J. Biologically-active peptides from milk-proteins with emphasis on 2 examples concerning antithrombotic and immunomodulating activities. *J. Dairy Sci.* **1993,** *76*, (1), 301-310.
40. Qian, Z.Y.; Jolles, P.; Miglioresamour, D.; Schoentgen, F.; Fiat, A.M. Sheep κ-casein peptides inhibit platelet-aggregation. *Biochim. Biophys. Acta, Gen. Subj.* **1995,** *1244*, (2-3), 411-417.
41. Mazoyer, E.; Levytoledano, S.; Rendu, F.; Hermant, L.; Lu, H.; Fiat, A.M.; Jolles, P.; Caen, J. KRDS, a new peptide derived from human lactotransferrin, inhibits platelet-aggregation and release reaction. *Eur. J. Biochem.* **1990,** *194*, (1), 43-49.
42. Lee, K.A.; Kim, S.H. SSGE and DEE, new peptides isolated from a soy protein hydrolysate that inhibit platelet aggregation. *Food Chem.* **2005,** *90*, 389-393.
43. Sodini, I.; Lucas, A.; Tissier, J.P.; Corrieu, G. Physical properties and microstructure of yoghurts supplemented with milk protein hydrolysates. *Int. Dairy J.* **2005,** *15*, 29-35.
44. Creusot, N. Enzyme-induced aggregation of whey proteins with *Bacillus licheniformis* protease. *Ph.D. Thesis*, Wageningen University, Wageningen, The Netherlands, **2006**.
45. Visschers, R.W.; de Jongh, H.H.J. Disulphide bond formation in food protein aggregation and gelation. *Biotechnol. Adv.* **2005,** *23*, (1), 75-80.
46. Wang, C.H.; Damodaran, S. Thermal gelation of globular-proteins - weight-average molecular-weight dependence of gel strength. *J. Agric. Food Chem.* **1990,** *38*, (5), 1157-1164.

## SUMMARY

Enzymatic hydrolysis can be used to improve the functional properties of soy proteins. In a number of studies it is reported that enzymatic hydrolysis of soy proteins can induce aggregation and subsequent gelation. Understanding the characteristics and origin of the aggregating peptides will enable optimization of the hydrolysis process with respect to the yield of peptides having aggregating properties. Hydrolyzates containing high proportions of these peptides can be used as functional ingredients in foods.

The aim of the work presented in this thesis is to understand how limited enzymatic hydrolysis can strengthen the aggregation behavior of soy protein material, with a focus on the individual contributions of the storage proteins glycinin and β-conglycinin. This is done by uncoupling the enzymatic degradation and the aggregation process. Subsequently, it is focused on the identification of those regions in soy proteins from which peptides originate that have strong aggregating properties.

A general introduction to soy proteins, their hydrolysis and their subsequent enzyme-induced aggregation is presented in Chapter 1. A literature overview is presented of the research performed up till now on the enzyme-induced aggregation of soy proteins. It was shown that various proteases have the ability to yield soy peptides that have strong aggregating properties.

In Chapter 2 it is described how the onset-pH of gelation of soy protein isolates (SPIs) hydrolyzed with subtilisin Carlsberg changes as a function of pH and ionic strength. Acid-induced cold gelation of soy protein hydrolysates with a degree of hydrolysis (DH) up to 9 % was studied. The enzyme was inhibited when the desired DH was reached to uncouple the hydrolysis from the subsequent gelation. The latter was induced by addition of glucono-δ-lactone causing a gradual lowering of the pH. Visual observations, confocal scanning laser microscopy images and the elasticity moduli showed that hydrolysates gelled at higher pH values with increasing DH. The non-hydrolyzed SPI gelled around pH 6.0, whereas a DH = 4.5 % hydrolysate gelled around pH 7.6. Gels made from hydrolysates had a softer texture than gels from intact SPI when manually disrupted. They show syneresis below a pH of 5-5.5. Monitoring the gelation by measuring the development of the storage modulus could be replaced by measuring the onset pH of aggregate formation ($pH_{Aggr\text{-}onset}$) using turbidity measurements. Changes in the ionic strength (0.03, 0.2 and 0.5 M) had only a minor influence on the $pH_{Aggr\text{-}onset}$. This indicates that the aggregation is not simply a balance between repulsive electrostatic and attractive hydrophobic interactions, but is more complex.

Chapter 3 elucidates the contributions of glycinin and β-conglycinin-derived peptides to the previously (Chapter 2) described aggregation of soy protein hydrolysates. The aggregation behavior, measured as the solubility of the proteinaceous material, as a function of pH was studied for hydrolysates of SPI, glycinin- and β-conglycinin-rich protein preparations obtained with subtilisin Carlsberg. The substrates were hydrolyzed up to DH 2.2 and 6.5 %. Compared to non-hydrolyzed SPI, a decrease in solubility was observed for the hydrolysates

of SPI (0.8 (w/w) % protein, $I$ = 0.03 M) around neutral pH. At pH 8.0, glycinin hydrolysates had a much lower solubility (~ 43 and 60 % for DH = 2.2 and 6.5 %, respectively) than SPI and β-conglycinin-derived hydrolysates. The latter were almost completely soluble. Peptides that aggregated were all larger than 5 kDa. The solubility behavior of hydrolysates of SPIs with different glycinin and β-conglycinin compositions showed that glycinin-derived peptides are causing the lower solubility of SPI hydrolysates. The solubility of SPI hydrolysates at pH 8.0 was shown not to be simply the sum of those of glycinin and β-conglycinin hydrolysates. Assuming that the separate hydrolysis of glycinin and β-conglycinin did not differ from that in the mixture (SPI), this indicates that β-conglycinin-derived peptides can have the ability to inhibit glycinin-derived peptide aggregation.

In Chapter 4 a number of enzymes (bromelain, papain, trypsin and chymotrypsin) were used to hydrolyze SPI, glycinin, and β-conglycinin. This was done to investigate whether peptides obtained after limited proteolytic hydrolysis of soy glycinin have the generic property to aggregate at neutral pH, and to what extent this depends on the protease used. SPI, glycinin, and β-conglycinin and, as a reference, bovine whey-derived proteins (whey protein isolate, α-lactalbumin, β-lactoglobulin) were heated and hydrolyzed till a DH of 2.2 %. The solubility of the hydrolysates obtained was studied as a function of pH. Next to glycinin-derived peptides obtained after hydrolysis with subtilisin Carlsberg (Chapter 3), at neutral pH glycinin-derived peptide hydrolysates obtained by hydrolysis with chymotrypsin, bromelain and papain also showed a lower solubility compared to the non-hydrolyzed glycinin. This decrease in solubility was not observed for glycinin hydrolyzed by trypsin, and also not for hydrolysates from β-conglycinin, SPI, and whey-derived protein preparations. These results show that glycinin has a strong tendency to aggregate upon limited hydrolysis in comparison to the other substrates. Since chymotrypsin prefers to cleave next to hydrophobic amino acids that are relatively abundant in the hydrophobic regions of glycinin, a stronger exposure of hydrophobic groups as a consequence of hydrolysis is postulated to be the driving force in the aggregate formation of glycinin hydrolysates. Since hydrolysates obtained upon hydrolysis with chymotrypsin result in similar observations as for hydrolysates obtained upon hydrolysis with subtilisin Carlsberg, for further experiments chymotrypsin was used. In this way the number of the peptides that can be expected upon hydrolysis is reduced due to the high specificity of chymotrypsin compared to subtilisin Carlsberg.

To increase the understanding of aggregation of glycinin-derived peptides, Chapter 5, 6, and 7 together present a new method to characterize those regions in soy proteins from which peptides originate that have strong aggregating properties. To determine the concentration of individual peptides, in Chapter 5 data are presented for the prediction of the absorbance of peptides at 214 nm based on their individual amino acid sequences. The molar extinction coefficients of 20 amino acids and the peptide bond were measured at 214 nm in the presence of acetonitrile and formic acid to enable quantitative comparison of peptides eluting from reversed-phase high-performance-liquid-chromatography once identified with mass-spectrometry (RP-HPLC-MS). The peptide bond has a molar extinction coefficient of 923 $M^{-}$

$^{1}cm^{-1}$, as measured via the use of polylysine. Tryptophan has a molar extinction coefficient that is ~30 times higher than that of the peptide bond, whereas the molar extinction coefficients of phenylalanine, tyrosine and histidine are ~6 times higher than that of the peptide bond. Proline, as an individual amino acid has a negligible molar extinction coefficient. However, when present in the peptide chain (except at the N-terminus) it absorbs ~3 times more than a peptide bond. Methionine has a similar molar extinction coefficient as the peptide bond. All other amino acids have a much lower molar extinction coefficient. The predictability of the molar extinction coefficients of proteins and peptides, calculated by the amino acid composition and the number of peptide bonds present, was validated using several proteins and peptides. The measured and calculated molar extinction coefficients were in rather good agreement, which shows that it is possible to compare peptides analyzed by RP-HPLC-MS in a quantitative way. This method enables a quantitative analysis of all peptides present in hydrolysates once identified with RP-HPLC-MS.

The quantification of the peptide concentration as described in Chapter 5 is used in Chapter 6 in which the novel method of accumulative-quantitative-peptide-mapping is presented. This method is applicable for systems in which peptides of a protein hydrolysate posses specific functional properties, but are too large to be sequenced directly and the peptide mixture is too complex to purify and characterize each peptide individually. In the present work, aggregating peptides obtained by proteolytic hydrolysis of soy glycinin were used as a case study. The aggregating peptides are isolated and subsequently further degraded with trypsin to result in peptides with masses < 5,000 Da to enable identification. The isolated peptides are first fractionated using anion and cation exchange chromatography. Subsequently, the fractions obtained are analyzed with RP-HPLC-MS in combination with MS/MS. The peptides, with identified sequences, were quantified using the peak areas of the RP-HPLC chromatograms measured at 214 nm, corrected for their individual molar extinction coefficient (Chapter 5), followed by peptide mapping. The results show that in complex systems, based on the method described, the regions in the parental protein from which the functional peptides originate can be identified.

In Chapter 7 the method of accumulative-quantitative-peptide-mapping is applied to soy glycinin with the objective to elucidate which regions in glycinin have the strongest tendency to aggregate after enzymatic hydrolysis.

Tryptic digestion of the aggregating peptides isolated from a glycinin hydrolysate resulted in a hydrolysate in which 90 % of the proteinaceous material was dissolved. The soluble fraction was used for accumulative-quantitative-peptide-mapping. The molar proportion in which the five glycinin subunits are present in glycinin, as determined by Edman degradation, is taken into account. The results of the accumulative-quantitative-peptide-mapping showed that mainly the basic polypeptide and a part of the acidic polypeptide, close to the location of the disulfide bridge between the basic and acidic polypeptide, are present in the aggregating peptide fraction. Based on the results obtained an aggregation mechanism was proposed. The mechanism shows that the hydrophilic acidic polypeptides, possibly still connected to the

basic polypeptide by a disulfide bridge, shield the hydrophobic basic polypeptides and are degraded upon hydrolysis. This results in a net increase in hydrophobicity of the truncated protein, which mainly consists of the basic polypeptide fragments. This increase in hydrophobicity is the driving force in the aggregation of chymotrypsin-derived peptides of glycinin.

The thesis is concluded with a general discussion (Chapter 8) on the main results obtained in this work with a special attention to possible effect of peptides to intact protein gelation. In addition, attention was paid to the difference between heated and nonheated glycinin as a substrate for hydrolysis.

## SAMENVATTING

Enzymatische hydrolyse kan gebruikt worden om de functionele eigenschappen van soja eiwitten te verbeteren. In verschillende studies is gevonden dat enzymatische hydrolyse van soja eiwit kan resulteren in aggregatie, al dan niet gevolgd door gelering. Het begrijpen van de kenmerken en de herkomst van de aggregerende peptiden kan het mogelijk maken het hydrolyse proces dusdanig te optimaliseren om zo een optimale hoeveelheid aggregerende peptiden te krijgen. Hydrolysaten met een grote hoeveelheid aggregerende peptiden kunnen gebruikt worden als functionele ingrediënten.

Het doel van het werk beschreven in dit proefschrift is het begrijpen hoe gelimiteerde hydrolyse het aggregatiegedrag van soja eiwitten kan versterken, met een focus op de individuele bijdrage van de eiwitten glycinine en β-conglycinine. Dit is gedaan door de enzymatische afbraak van de eiwitten los te koppelen van het aggregatieproces, gevolgd door de identificatie van die gebieden in soja eiwitten van waaruit de peptiden afkomstig zijn die de sterkst aggregerende eigenschappen hebben.

Hoofdstuk 1 van dit proefschrift geeft een algemene introductie van soja eiwitten, hun hydrolyse en de daarop volgende aggregatie van peptiden. Er wordt eveneens een literatuuroverzicht gepresenteerd van het onderzoek dat tot nu toe uitgevoerd is met betrekking tot enzym-geïnduceerde aggregatie van soja eiwitten. Dit overzicht laat zien dan er verschillende proteases zijn die de eigenschap hebben om peptiden te vormen die sterk aggregeren.

In Hoofdstuk 2 wordt beschreven hoe de begin-pH van gelering van de soja eiwit isolaten (SPI), gehydrolyseerd met subtilisine Carlsberg, verandert als een functie van pH en ionsterkte. Deze zuur-geïnduceerde koude gelering van soja eiwit hydrolysaten met een hydrolyse graad (DH) tot 9 % is bestudeerd. Het enzym werd gestopt wanneer de gewenste DH was bereikt om de hydrolyse te ontkoppelen van de daaropvolgende gelering. Deze gelering werd geïnduceerd door de toevoeging van glucono-δ-lacton wat resulteert in een geleidelijke pH daling. Visuele waarnemingen, confocal-scanning-laser microscopie en de elasticiteits modulus lieten zien dat hydrolysaten geleerden bij een hoger pH waarde met toenemende DH. Het niet gehydrolyseerde soja eiwit isolaat vormde een gel rond pH 6,0, terwijl een hydrolysaat met een DH van 4,5 % rond pH 7,6 een gel vormde. Gelen gemaakt van hydrolysaten hadden een zachtere textuur dan gelen van intact SPI wanneer deze handmatig werden vervormd. De gelen gemaakt van hydrolysaten vertoonden synerese bij pH waarden lager dan 5-5,5. Het volgen van de gelering door het meten van de ontwikkeling van de opslag modulus kon worden vervangen door de begin-pH van aggregaat vorming te meten door gebruik te maken van troebelheids metingen. Verandering in ionsterkte (0,03, 0,2 en 0,5 M) had slechts een kleine invloed op de begin-pH van aggregaat vorming. Dit geeft aan dat de aggregatie niet enkel een balans tussen afstotende elektrostatische en aantrekkende hydrofobe interacties is.

Hoofdstuk 3 ontrafelt de bijdrage van peptiden verkregen uit glycinine- en β-conglycinine- aan de voorgaand beschreven (Hoofdstuk 2) aggregatie van de soja eiwit hydrolysaten. Het aggregatie gedrag werd gemeten als de oplosbaarheid van het eiwit materiaal, als functie van de pH. Voor hydrolysaten van SPI, glycinine- en β-conglycinine-rijke eiwit preparaten gehydrolyseerd met subtilisine Carlsberg. De substraten werden gehydrolyseerd tot een DH van 2,2 en 6,5 %. Vergleken met niet-gehydrolyseerd SPI werd rond neutrale pH een daling in oplosbaarheid gemeten voor de hydrolysaten van SPI (0,8 (w/w) % eiwit, $I$ = 0,03 M). Glycinine hydrolysaten hebben bij een pH van 8,0 een veel lagere oplosbaarheid (~ 43 en 60 % voor DH = 2,2 en 6,5 %, respectievelijk) dan hydrolysaten van SPI en β-conglycinine. Deze laatste twee waren bijna geheel oplosbaar. Peptiden die aggregeerden waren allemaal groter dan 5.000 Da. Het oplosbaarheidsgedrag van hydrolysaten van SPIs met verschillende glycinine en β-conglycinine samenstellingen liet zien dat peptiden afkomstig van glycinine verantwoordelijk waren voor de lagere oplosbaarheid van de SPI hydrolysaten. De goede oplosbaarheid van de SPI hydrolysaten bij pH 8,0 liet zien dat dit niet slechts het resultaat was van de som van glycinine en β-conglycinine hydrolysaten. Als aangenomen wordt dat de individuele hydrolyse van glycinine en β-conglycinine niet verschilt van die in de mix (SPI), geeft dit aan dat peptiden van β-conglycinine de eigenschap hebben om aggregatie van glycinine peptiden te remmen.

In Hoofdstuk 4 zijn verschillende enzymen (bromelaïne, papaïne, trypsine en chymotrypsine) gebruik om SPI, glycinine en β-conglycinine te hydrolyseren. Dit werd gedaan om te onderzoeken of peptiden, verkregen na gelimiteerde proteolytische hydrolyse van soja glycinine, de generieke eigenschap hebben om te aggregeren bij neutrale pH en hoe dit afhankelijk is van het gebruikte enzym. SPI, glycinine en β-conglycinine en, als referentie, koemelk wei-eiwitten (wei-eiwit isolaat, α-lactalbumine, β-lactoglobuline) werden verhit en gehydrolyseerd tot een DH van 2,2 %. De oplosbaarheid van de verkregen hydrolysaten werd bestudeerd als een functie van de pH. Naast peptiden van glycinine, verkregen na hydrolyse met subtilisine Carlsberg (Hoofdstuk 3), bij neutrale pH, hadden peptiden van glycinine verkregen door hydrolyse met chymotrypsine, bromelaïne en papaïne ook een lagere oplosbaarheid dan de niet gehydrolyseerde glycinine. Deze daling in oplosbaarheid werd niet waargenomen voor glycinine gehydrolyseerd met trypsine, en ook niet voor hydrolysaten van SPI, β-conglycinine en de wei-eiwitten. Deze resultaten laten zien dat glycinine, in vergelijking met andere eiwitten, een sterke aanleg heeft om te aggregeren na gelimiteerde hydrolyse. Omdat chymotrypsine de voorkeur heeft te knippen naast hydrofobe aminozuren, die in relatief grote hoeveelheden aanwezig zijn in de hydrofobe gebieden van glycinine, heeft dit mogelijk een sterkere blootstelling van hydrofobe groepen tot resultaat. Deze verhoging in hydrofobiciteit wordt verondersteld de drijvende kracht te zijn in de aggregaat vorming in glycinine hydrolysaten.

Omdat hydrolysaten verkregen met chymotrypsine resulteerden in vergelijkbare resultaten als hydrolysaten verkregen door hydrolyse met subtilisine Carlsberg is voor verdere

experimenten chymotrypsine gebruikt. Door de hogere specificiteit van chymotrypsine is het aantal te verwachten peptiden gereduceerd.
Om het begrip van aggregatie van glycinine peptiden te vergroten, wordt in Hoofdstuk 5, 6, en 7 een nieuwe methode gepresenteerd. Deze methode heeft als doel de gebieden in soja eiwitten te karakteriseren van waar de peptiden afkomstig zijn die sterke aggregerende eigenschappen hebben. Om de concentratie van individuele peptiden te bepalen worden in Hoofdstuk 5 data gepresenteerd voor de voorspelling van de absorptie van peptiden bij 214 nm gebaseerd op hun individuele aminozuur sequentie. De molaire extinctiecoëfficiënt van de 20 aminozuren en de peptide binding werd gemeten bij 214 nm in de aanwezigheid van acetonitril en mierenzuur om peptiden kwantitatief te kunnen vergelijken die elueren na reversed-phase high-performance-liquid-chromatography als deze eenmaal geïdentificeerd zijn met behulp van massaspectrometrie (RP-HPLC-MS). De peptide binding heeft een molaire extinctiecoëfficiënt van 923 $M^{-1}cm^{-1}$, gemeten aan de hand van polylysine. Tryptofaan heeft een molaire extinctiecoëfficiënt die ~30 keer hoger is dan die van de peptide binding, terwijl de molaire extinctiecoëfficiënt van fenylalanine, tyrosine en histidine ~6 keer hoger is die van de peptide binding. Proline heeft als een individueel aminozuur een verwaarloosbare molaire extinctiecoëfficiënt. Maar als proline aanwezig is in een peptide (behalve aan de N-terminus) absorbeert deze ~3 keer meer dan een peptide binding. Methionine heeft een vergelijkbare molaire extinctiecoëfficiënt als die van de peptide binding. Alle andere aminozuren hebben een veel lagere molaire extinctiecoëfficiënt. De voorspelbaarheid van de molaire extinctiecoëfficiënt van eiwitten en peptiden, berekend op basis van de aminozuur samenstelling en het aantal peptide bindingen, werd gevalideerd met behulp van verschillende eiwitten en peptiden. De gemeten en berekende molaire extinctiecoëfficiënten kwamen vrij goed overeen. Dit laat zien dat het mogelijk is om peptiden kwantitatief te analyseren met RP-HPLC-MS. Deze methode maakt het mogelijk om alle peptiden die in een hydrolysaat aanwezig zijn te kwantificeren als de sequentie van de peptiden bekend is.
De kwantificering van de peptide concentratie, zoals beschreven in Hoofdstuk 5, is in Hoofdstuk 6 gebuikt, waar een nieuwe methode is beschreven, genaamd: accumulative-quantitative-peptide-mapping. Deze methode is toepasbaar voor systemen waarin peptiden van eiwit hydrolysaten specifieke eigenschappen hebben, maar te groot zijn om direct de sequentie ervan te bepalen en waarvan het hydrolysaat zo complex is dat het niet direct mogelijk is om de peptiden individueel te isoleren en karakteriseren. De aggregerende peptiden van soja glycinine zijn hier gebuikt als een casus. De aggregerende peptiden zijn geïsoleerd en vervolgens verder afgebroken met trypsine wat resulteerde in peptiden met een massa < 5.000 Da hetgeen identificatie mogelijk maakt. De geïsoleerde peptiden zijn eerst gefractioneerd met behulp van ion wisseling chromatografie. Vervolgens zijn de verkregen fracties geanalyseerd met RP-HPLC-MS in combinatie met MS/MS. De peptiden, met bekende sequentie, zijn gekwantificeerd op basis van hun piek oppervlakten van de RP-HPLC chromatogrammen gemeten bij 214 nm, gecorrigeerd voor hun individuele molair

extinctiecoëfficiënt (Hoofdstuk 5), gevolgd door peptide-mapping. De resultaten laten zien dat in complexe systemen, gebaseerd op de beschreven methode, de gebieden in de eiwitten van waar de functionele peptiden afkomstig zijn kunnen worden geanalyseerd.
In Hoofdstuk 7 is de methode van accumulative-quantitative-peptide-mapping toegepast of soja glycinine met als doel te identificeren welke gebieden in glycinine de sterkste tendens hebben om te aggregeren na enzymatische hydrolyse.
Trypsine digestie van de aggregerende peptiden, geïsoleerd uit een glycinine hydrolysaat, resulteerde in een hydrolysaat waarin 90 % van het eiwit materiaal opgelost was. De oplosbare fractie werd gebruikt voor accumulative-quantitative-peptide-mapping. De molaire concentratie waarin de vijf glycinine subunits aanwezig zijn in glycinine, bepaald met behulp van Edman degradatie, is meegenomen in de analyse. De resultaten laten zien dat vooral de basische polypeptide en het deel van de zure polypeptide dicht bij de locatie van de zwavelbrug tussen de zure en basische polypeptide aanwezig zijn in de aggregerende peptide fractie. Gebaseerd op de resultaten is er een aggregatie mechanisme opgesteld. Het mechanisme laat zien dat de hydrofiele zure polypeptide, mogelijk nog steeds verbonden met een zwavelbrug aan de basische polypeptide, de hydrofobe basische polypeptide afschermt en wordt afgebroken gedurende de hydrolyse. Dit resulteert in een netto toename van de hydrofobiciteit van het deels afgebroken, achtergebleven eiwit, welke voornamelijk uit peptiden van de basische polypeptide bestaat. Deze toename in hydrofobiciteit is de drijvende kracht in de aggregatie van glycinine peptiden na hydrolyse met chymotrypsine.
Het proefschrift wordt afgesloten met een algemene discussie (Hoofdstuk 8) van de belangrijkste resultaten verkregen in dit werk met een speciale aandacht voor het mogelijke effect van peptiden om gelering van intact eiwit te veranderen. Tevens wordt ingegaan op de verschillen tussen verhit en onverhit eiwit als substraat voor de hydrolyse.

## DANKWOORD

"*Terug de schoolbanken in*", dat is wat sommige mensen dachten toen ik na twee jaar mijn baan als productontwikkelaar verruilde voor het AIO-schap. Nu ging ik niet echt terug de schoolbanken in, maar ik moet zeggen dat ik veel geleerd heb in mijn tijd als AIO bij Levensmiddelenchemie. Ik heb dit al die jaren met heel veel plezier gedaan. De 5$^{e}$ verdieping van het Biotechnion is een fijne plek om te werken. Hiervoor wil ik alle collega's bedanken, zowel voor jullie behulpzaamheid bij vragen rond mijn onderzoek, als voor de gezelligheid tijdens pauzes en uitjes. Een aantal mensen wil ik graag in het bijzonder noemen.

Allereerst mijn promotor Harry Gruppen. Harry, bedankt dat je van begin af aan het vertrouwen had in mij als je volgende AIO die aan soja-eiwitten ging werken. Ik heb al die jaren prettig met je samengewerkt en vond het fijn dat je me veel vrijheid gaf om zelf te bepalen welke kant het onderzoek op zou gaan. Je kritische houding en je scherpte (vooral in de ochtend) resulteerden vaak in de nodige extra experimenten die wel de puntjes op de i zetten.

Gerrit van Koningsveld, bedankt voor je bijdrage in de eerste jaren van mijn onderzoek zowel als het ging om praktische vaardigheden als het schrijven van artikelen, maar ook als het ging om je bijdrage aan de Top 10 van slechtste grappen.

Arno Alting en Frank Driehuis van NIZO food research, ik wil jullie bedanken voor jullie nuttige bijdrage. Ik heb de samenwerking als zeer prettig ervaren. Arno, door jouw ervaring op het gebied van koude gelering kon ik de experimenten bij het NIZO snel opstarten. Verder wil ik Saskia de Jong en Jan Klok bij NIZO food research bedanken voor de hulp bij de verschillende experimenten die ik daar heb uitgevoerd.

Ik wil Fons Voragen bedanken voor de getoonde interesse en enthousiasme over mijn onderzoek. Jean-Paul Vincken wil ik bedanken voor het kritisch doorlezen van enkele manuscripten en Edwin Bakx voor de hulp en discussies rond massaspectrometrie van peptiden. I also would like to thank Sjef Boeren and Hanem Awad for there support with the mass spectrometry experiments.

Ook wil ik mijn paranimfen Nathalie en Koen bedanken. Leuk dat jullie mijn paranimfen willen zijn. Nathalie I enjoyed working with you for more than 4 years at Lab 503: "the pH-stat expertise centre". Thanks for all the useful discussions about aggregating peptides. Koen, ik mocht bij jou de "moeilijke" begintijd van je promotie onderzoek van dichtbij meemaken, nu kan je bij mij het einde van mijn promotie ook van dichtbij meemaken. Ik kijk met veel plezier terug naar de verschillende activiteiten die we in de tussentijd hebben ondernomen!

Hauke, Karin en Gerd Jan, het gezamenlijk organiseren van de AIO reis naar Japan was een welkome afwisseling tussen alle experimenten door. Ik vond het heel leuk dit met jullie te organiseren, wat resulteerde in een erg mooie reis. Natuurlijk wil ik ook al mijn kamergenoten Francesca, Jolan, Renate, René en Marijke bedanken voor de gezelligheid op kamer 505. Ruurdtje, Corjan, Jing Shao, Leonie, Jing Jiang, Panos, and Joyce thanks a lot for your

contribution to my Ph.D. project during your B.Sc. or M.Sc. thesis projects. I enjoyed working with all of you.
Verder wil ik ook de mannen van Nicolas Maillard (+ spelersvrouwen), de T30@95-vrienden, Hans en Ingeborgh, en Ferdi bedanken voor de belangstelling die jullie toonden voor mijn werk maar vooral voor alle leuke activiteiten die we hebben ondernomen de laatste jaren. Ik hoop dat er nog vele zullen volgen.
Lieve papa en mama ik ben dan nu echt klaar met studeren. Dank jullie wel voor de belangstelling die jullie altijd tonen in mijn werk en ook voor alle steun en liefde die ik van jullie heb gekregen. Lieve Bob en Suzanne jullie zijn een geweldige broer en zus! Joke en Bert bedankt voor de gastvrijheid en gezelligheid als we afreisden tot ver beneden de rivieren of als jullie naar Wageningen kwamen. Ook mijn schoonbroertjes en -zusjes Dagmar, Wouter, Sara en Marcel, Rosa en Fabian bedankt voor alle leuke momenten de afgelopen jaren.
Lieve Julia, jij bent voor mij de allerbelangrijkste! We hebben samen al veel leuke dingen gedaan en er zullen er, samen met Liese, vast nog vele volgen.

## CURRICULUM VITAE

Bas Johan Henri Kuipers is geboren op 7 november 1975 te Winterswijk. In 1992 behaalde hij zijn MAVO diploma aan de RSG Hamaland te Winterswijk, in 1994 gevolgd door het HAVO diploma aan De Driemark te Winterswijk. In 1995 behaalde hij zijn VWO diploma aan de Scholengemeenschap De Waerdenborch in Holten. In datzelfde jaar begon hij aan de opleiding Levensmiddelentechnologie aan de toenmalige Landbouw Universiteit te Wageningen. Hij volgde de specialisatie Levensmiddelenchemie welke werd afgesloten met een afstudeervak bij de leerstoelgroep Levensmiddelenchemie. Hij liep o.a. een stage bij Givaudan Roure in Zwitserland. Zijn tweede afstudeervak heeft hij gedaan bij de leerstoelgroep Levensmiddelennatuurkunde. Hij studeerde af in september 2000 aan de Wageningen Universiteit. Vervolgens trad hij in dienst als productontwikkelaar bij Koninklijke De Ruijter. Per mei 2002 was hij als promovendus in dienst van de Wageningen Universiteit bij de leerstoelgroep Levensmiddelenchemie wat geleid heeft tot dit proefschrift. Sinds maart 2007 is hij werkzaam als productontwikkelaar bij DMV International te Wageningen.

## LIST OF PUBLICATIONS

Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Gruppen, H.; Voragen, A.G.J. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. *J. Agric. Food Chem.* **2005**, 53, 1031-1038.

Kuipers, B.J.H.; van Koningsveld, G.A.; Alting, A.C.; Driehuis, F.; Voragen, A.G.J.; Gruppen, H. Opposite contributions of glycinin- and β-conglycinin-derived peptides to the aggregation behavior of soy protein isolate hydrolysates. *Food Biophys.* **2006**, 1, 178-188.

Kuipers, B.J.H.; Alting, A.C.; Gruppen, H. Soy glycinin as a dominant source of aggregating peptides using different proteases. *Submitted.*

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 RP-HPLC-MS analysis. *J. Agric. Food Chem.*, a*ccepted for publication.*

Kuipers, B.J.H.; Bakx, E.J.; Gruppen, H. Functional region identification in proteins by accumulative-quantitative-peptide-mapping using RP-HPLC-MS. *Submitted.*

Kuipers, B.J.H.; Van Koningsveld, G.A.; Gruppen, H. Identification of strong aggregating regions in soy glycinin upon enzymatic hydrolysis. *To be submitted.*

## OVERVIEW OF COMPLETED TRAINING ACTIVITIES

**Discipline specific activities:**

*Courses:*

VLAG Summer school glycosciences (Wageningen, June 2002)
Rheological Measurements, European School of Rheology (Belgium, 2003)
Industrial Proteins, VLAG and Senter (Wageningen, 2003)
Food Enzymology, VLAG (Wageningen, 2004)

*Meetings:*

Symposium Industrial Proteins, Senter (Ede, 2003)
NIZO Dairy Conference (Arnhem, 2003)
Scientific exchange (Germany, 2004)
Symposium "Transition towards sustainable protein supply chains" (Wageningen, 2004)
Food Colloids (Switzerland, 2006)
Food Chemistry Protein meetings (2002-2006)

**General courses:**

Food Chemistry PhD trip (USA, 2002)
VLAG PhD week (Bilthoven, 2003)
Food Chemistry PhD trip (Japan, 2004)
Scientific Writing (Wageningen, 2004)
Career Perspectives Course (Wageningen, 2005)

**Additional activities:**

Preparation of the PhD research proposal (2002)
Food Chemistry Seminars (2002-2006)
Food Chemistry Colloquia (2002-2006)

*Research presented in this PhD dissertation was financially supported by the Graduate School VLAG*

Printing: Ponsen & Looijen bv, Wageningen

Picture on the cover: With permission from the American Soybean Association