

Forschungsbeiträge/Research Papers

Computer simulation of proteolysis. Peptic hydrolysis of partially demasked β -Lactoglobulin

M. M. Vorob'ev and I. A. Goncharova

Computer assistance was used for simulation of proteolysis of partially demasked β -Lactoglobulin by pepsin. Total proteolysis kinetics and kinetics for intermediate peptides were computed by PROTEOLYSIS program for different rates of demasking. The increase of demasking rate by a factor of 100 gives twofold increase in the total hydrolysis rate. The identification of cleavage sites on the basis of limited

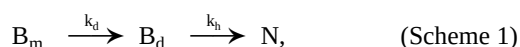
number of peptides was found to depend on the depth of hydrolysis. When the number of identified peptides was only 10, the computer simulation has allowed to predict a half of experimentally determined cleavage sites. The influence of the dimension of enzyme active site and of the secondary structure of substrate on the efficiency of prediction of cleavage sites was analysed.

1 Introduction

The computer assistance seems to be very useful for quantitative studies of proteolysis kinetics because one can analyse in details the complicated patterns of fragment evolution during the process. For the random degradation of homopolymers, the exact mathematical description of the kinetics for each of the fragments was published [1]. But for enzymatic hydrolysis of protein, which is heteropolymer, the kinetics is complicated by the different splitting potentialities of different bonds in polypeptide chain as well as by the process of enzyme saturation. To overcome these difficulties, the PC software PROTEOLYSIS has been elaborated to simulate enzymatic hydrolysis of any protein with arbitrary amino acid sequence [2].

In spite of the importance of proteolysis in food science and biology, no general model has been created up to date to describe the proteolysis quantitatively for different enzymes, substrates, for various reagent concentrations, at arbitrary degrees of proteolysis. Meanwhile, three major steps to approach the quantitative description of proteolysis were made. Firstly, the increments of the rate constants were introduced which evaluated the contributions of various amino acid residues [3]. This allows to calculate the splitting potentialities (rate constants) for any peptide bond in any peptide chain on the basis of amino acid sequence of this chain. Secondly, the role of the proteolysis degree was established as a single argument for the functions which describe the changes of different properties for hydrolysates during proteolysis [4]. Thirdly, the functions, instead of rate constants, were used for description of kinetic parameters which can be compared for different substrates at an equivalent degree of peptide bond hydrolysis [5].

Proteolysis is regarded in the present work as two-stage process, including the demasking of masked peptide bonds B_m on the first stage and the hydrolysis of demasked bonds B_d on the second stage [6]:



where k_d is the rate constant of demasking, k_h is the rate constant of hydrolysis of demasked bonds, and N is the amino nitrogen, the chemical outcome of the reaction. Demasking parameter k_d/k_h plays an important role in the kinetic regulation of proteolysis, especially at the initial stage of the process [7, 8].

Pepsin is a well known example of a proteolytic enzyme with extended active site of the dimension of approximately 5–10 amino acid residues [3]. In this case, the rate constant for the hydrolysis of a bond is determined not only by amino acid residue containing the CO group of this bond (primary specificity), but also by other amino acid residues (secondary specificity).

Specificity of the proteolytic enzymes is quantitatively evaluated by the set of hydrolysis rate constants for different bonds in various polypeptide chains. To evaluate specificity, some approaches were suggested including statistical analysis of a multitude of proteolysis products [3] and kinetic studies of hydrolysis of low-molecular-weight substrates [5].

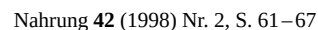
β -Lactoglobulin (BLG) is the major milk whey protein. At acidic pH, BLG is resistant to peptic action [9], however, it can be hydrolysed after some breaking of its tertiary structure by high pressure [10], chemical modification [11], or interaction with the ethanolic media [12]. In the terms of Scheme 1, the treatment of native BLG gives some new substrate with higher extent of demasking, i.e. the partially demasked BLG with higher k_d/k_h ratio.

The production of valuable individual peptides with high level of functional properties [13] and physiologically active peptides [14] is of interest for scientists now and will be of interest for practice in the nearest future. It requires the elaboration of new experimental methods of determination and identification of individual peptides in the complex mixtures as well as the elaboration of new computer algorithms and programs for prediction of peptide fragmentation [15].

Computer simulation allows to utilize available experimental data for the calculation of the parameters entering into the model. Then these parameters can be used for prediction of proteolysis kinetics for new protein substrates with desirable amino acid sequences and demasking parameters. The most interesting is a question of the correspondence of predicted results with the experiment.

Foodinform Ltd., ul. Vavilov 28, RUS-117813 Moscow, Russia.

Correspondence to:
Dr. M. M. Vorob'ev.



Help	Protein	Enzyme	Kinetic conditions	Simulation	Results	Quit
				Specificity Parameters Run		
Protein :		beta-lactoglobulin B				
Concentration (mg/ml)		S = 5				
Enzymes :		pepsin				
Concentration (mg/ml) :		E1 = 0.1				
Proteolysis time (min) :		t1 = 12				
pH :		pH1 = 3				
Temperature (°C) :		T1 = 37				
Step = 00.30 Calculation Secondary						
t = 3.3 N = 162 d = 0.0064						
Wait, please ... ESC-Abort						

Figure 3. Printscreen pattern for the PROTEOLYSIS program. Definition of the kinetic conditions and the procedure of calculations. Parameters S, E, t, pH and temperature are defined by user. The progress of calculations is demonstrated by presentation of current t and D values.

The numerical input data consist of E_0 , S_0 , T, pH and proteolysis time t (Fig. 2, block 1). After integration of differential equation set consisting of eqs. (1)–(2) (Fig. 2, block 4) with *Runge-Kutta* algorithm, the program represents kinetic output data in the graphic form (Fig. 2, block 5). In the case of known structure-activity relationship, the program can be used for the prediction of the changes of hydrolysis properties during proteolysis (Fig. 2, block 6). This program was written in C, and it does not require the WINDOWS operation system.

The additional information about PROTEOLYSIS program is available on the request by E-mail: mmvor@ineos.ac.ru.

3 Results and discussion

3.1 Total proteolysis kinetics

Amino nitrogen N, i.e. the sum of the molar concentrations over all amino compounds (intact protein, all peptides and amino acids), increases in the course of proteolysis, owing to continuous splitting of peptide bonds. Amino nitrogen $N(t)$ and proteolysis degree $D(t) = N(t)/(B_m(t) + B_d(t) + N(t))$ are calculated by PROTEOLYSIS program for any moment of proteolysis. These functions determine the total proteolysis kinetics describing the process as a whole, and are usually experimentally determined with TNBS [16], pH-stat technique [17] and other methods, which give proportional response to the molar concentration.

The exact values of k_d/k_h parameter for BLG partially demasked by different ways mentioned before are unknown. However, for the estimation of this parameter, the data on the initial proteolysis kinetics of whey milk proteins and caseins can be used [7]. At neutral pH, whey proteins (BLG, α -lactalbumin and bovine serum albumin) are easily accessible for action of chymotrypsin, trypsin and pancreatin [7, 8, 17, 18]. The initial rates of hydrolysis by chymotrypsin of α - and β -casein (unmasked polypeptides chains) were shown to be more than two orders of magnitude higher than the rates of the degradation of milk whey proteins limited by the protein globule demasking [7]. Therefore, for k_d/k_h parameter, the range of 4–0.04 was assumed to correspond to the interval between fully demasked and partially demasked substrate.

$N(t)$ is shown in Fig. 4 for peptic proteolysis of BLG, when the k_d/k_h parameter was under modeling. The increase of k_d/k_h

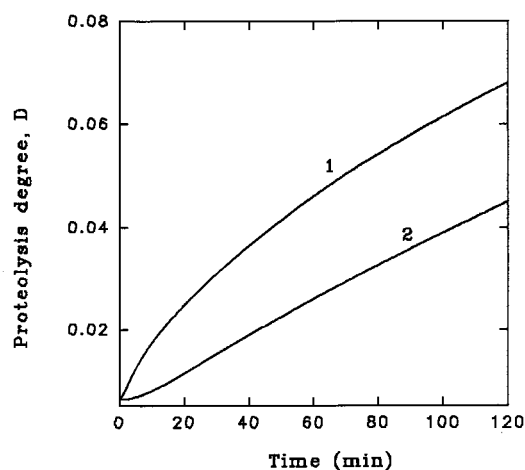


Figure 4. Dependence of the proteolysis degree on the proteolysis time. Proteolysis conditions: $S_0 = 5$ mg/ml, $E_0 = 0.1$ mg/ml, $T = 37^\circ\text{C}$, pH = 3.0; 1 – $k_d/k_h = 4$; 2 – $k_d/k_h = 0.04$.

parameter by a factor of hundred leads only to an approximately twofold increase in the total proteolysis rate which is registered after some initial period of proteolysis. Such close total kinetics for substrates with different level of peptide bond masking were not surprising. For example, the rate of product accumulation upon the proteolysis of milk whey proteins by trypsin was about two times lower than that of caseins [18].

3.2 Kinetic simulations for individual peptides

There are three different types of individual kinetic curves. The first one is the monotonously dropping curve, which corresponds to the intact protein (Fig. 5 a, b). At a low rate of demasking ($k_d/k_h = 0.04$), the degradation of intact BLG was computed to be finished in 68 min, while at a higher rate of demasking ($k_d/k_h = 4$) that was computed to be done in 25 min. Kinetics of the hydrolysis of the intact protein is mostly determined by k_d parameter and can be used for the estimation of the demasking stage (Scheme 1), as it was shown by us previously [7].

The second one corresponds to the intermediate peptide products, which are accumulated and then consumed as the

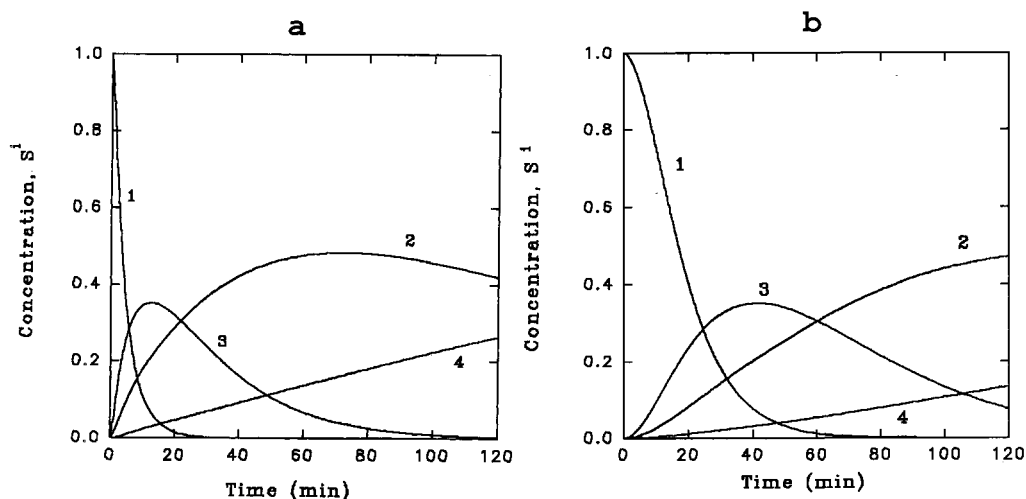


Figure 5. Dependence of the concentration of individual components on the proteolysis time. 1 – intact BLG, 2 – Leu¹³³-Ile¹⁶², 3 – Lys⁷⁵-Ile¹⁶², 4 – Glu¹⁵⁷-Ile¹⁶².
a) Demasking parameter $k_d/k_h = 4$.
b) Demasking parameter $k_d/k_h = 0.04$.

result of hydrolysis of intrinsic peptide bonds. These peptides being the intermediate products are accumulated and then consumed in the course of proteolysis. Intermediate products represent the relatively long peptides, which can be surfactants or possess physiological activities. The PROTEOLYSIS program allows to predict the time intervals providing the accumulation of desirable peptides. An illustration is shown in Fig. 5a, b on the example of two peptides Lys⁷⁵-Ile¹⁶² and Leu¹³³-Ile¹⁶².

Third type of curves is related to the residual proteolysis products consisting of amino acids and short peptides with unspecific peptide bonds (Fig. 5a, b; peptide Glu¹⁵⁷-Ile¹⁶²).

One can evaluate the lifetime τ of the intermediate peptide, as it is shown in Fig. 6. For example, for the peptide Lys⁷⁵-Ile¹⁶² τ is 35 min ($k_d/k_h = 4$) and 73 min ($k_d/k_h = 0.04$), that corresponds to $D_\tau = 2.6\%$, where D_τ is the lifetime evaluated in the terms of degree of hydrolysis, as it is shown in Fig. 6. If one assumes the maximum proteolysis degree for peptic hydrolysates as 15% [19], the ratio $D_\tau/15 = 0.17$ presents the dimensionless lifetime. For comparison, this parameter is 0.027 for the intact BLG. For the preparation of some peptide by proteolysis from an appropriate protein, the lifetime of this peptide in the reaction mixture is practically important, indicating, for example, the requirement on the narrowness of time interval for the collection of some fraction containing this peptide.

The main difficulty of the theoretical analysis of proteolysis kinetics consists in a huge number of proteolysis products, while their concentrations continuously change during the process. Even in the precise analytical experiments the number of the identified peptides in the hydrolysates is limited [20]. This makes difficult the prediction of major components, which could be found in hydrolysates by an appropriate analytical method. In this study, the number of identified compounds is assumed to be 10.

The PROTEOLYSIS program made selection over all proteolytic components and represented a list of 10 components, which have maximum concentrations during the defined time interval. Using this information one can simulate the fragmentation of the BLG hydrolysed by pepsin, as it can be deter-

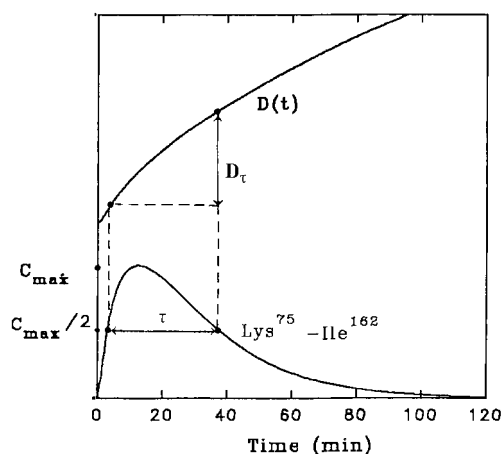


Figure 6. Definition of the lifetime of the intermediate peptides. τ corresponds to the time interval, when the concentration of a peptide does not drop lower than a half of maximum concentration.

mined by some analytical methods with a finite number of analysed compounds (Fig. 7, $k_d/k_h = 0.04$ and Fig. 8, $k_d/k_h = 4$). In this series of computer experiments, the dimension of the active site for pepsin was assumed to be 10 amino acid residues. The increase of time interval of proteolysis allows to reveal new peptide fragments, which are shown in Fig. 7 and 8. These new peptide fragments allow to identify new cleavage sites, which were not predicted in the previous periods of proteolysis. Thus, in the case of a limited number of analysed compounds, the result of identification of cleavage sites may depend on the depth of hydrolysis. It means, that the average length of peptides (degree of proteolysis) should be taken into consideration, when one makes conclusion about possible cleavage sites on the basis of protein fragmentation.

The demasking parameter is less important, especially at the second stage of proteolysis, when the demasking stage does not limit the hydrolysis stage (Scheme 1). One can expect the similar distributions of peptides in the hydrolysates obtained at different kinetic conditions but with the close degrees of pep-

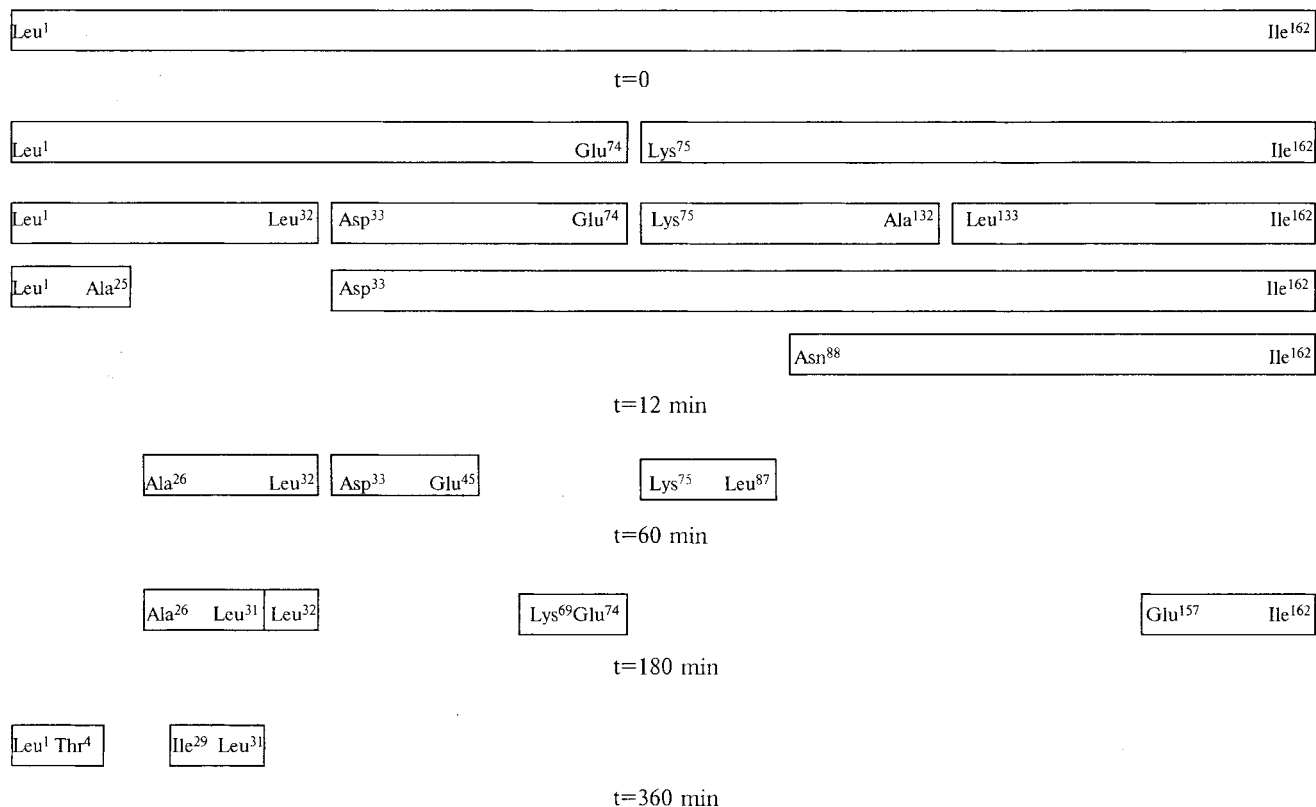


Figure 7. Major components of the proteolytic reaction, as it was simulated by the PROTEOLYSIS program with $k_d/k_h = 0.04$. Dimension of the active site of pepsin was accepted to be 10 amino acid residues.

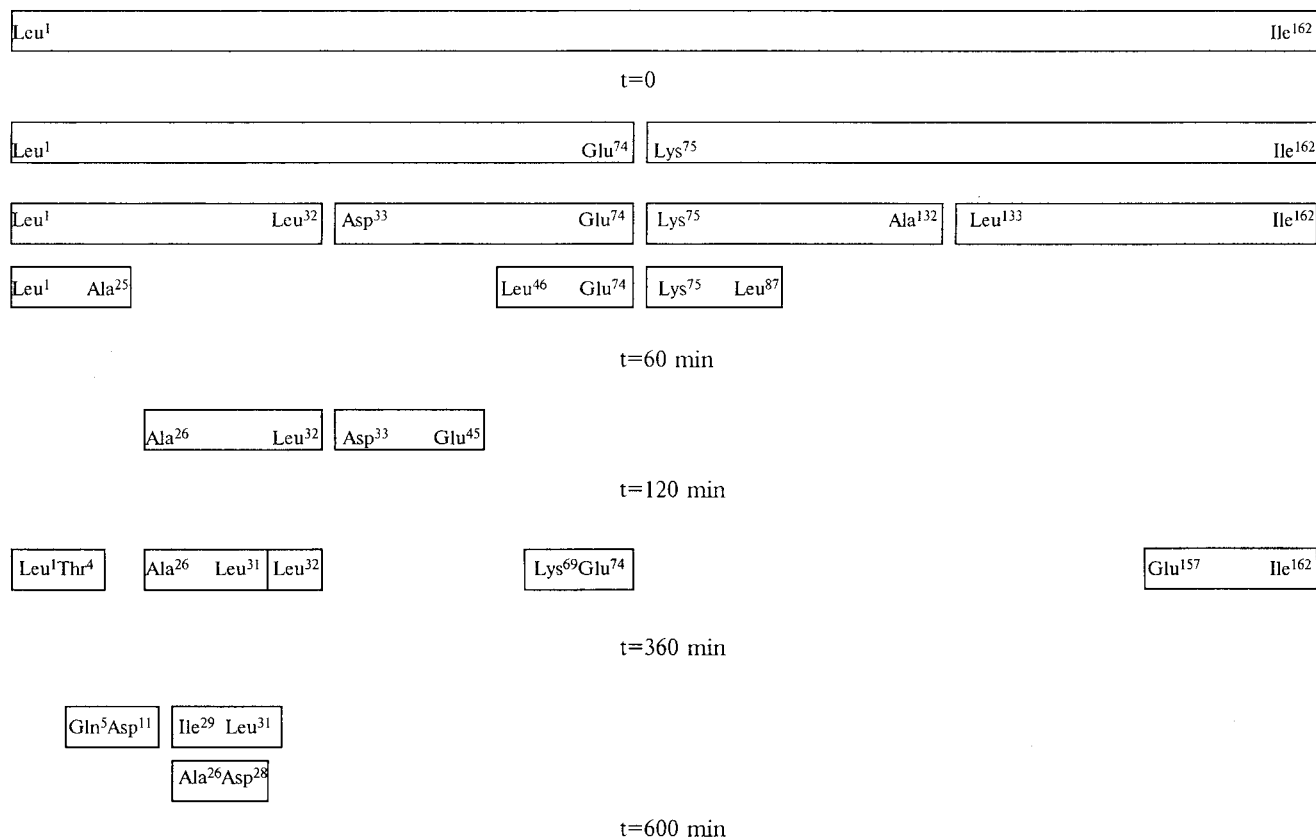


Figure 8. Major components of the proteolytic reaction, as it was simulated by the PROTEOLYSIS program with $k_d/k_h = 4$. Dimension of the active site of pepsin was accepted to be 10 amino acid residues.

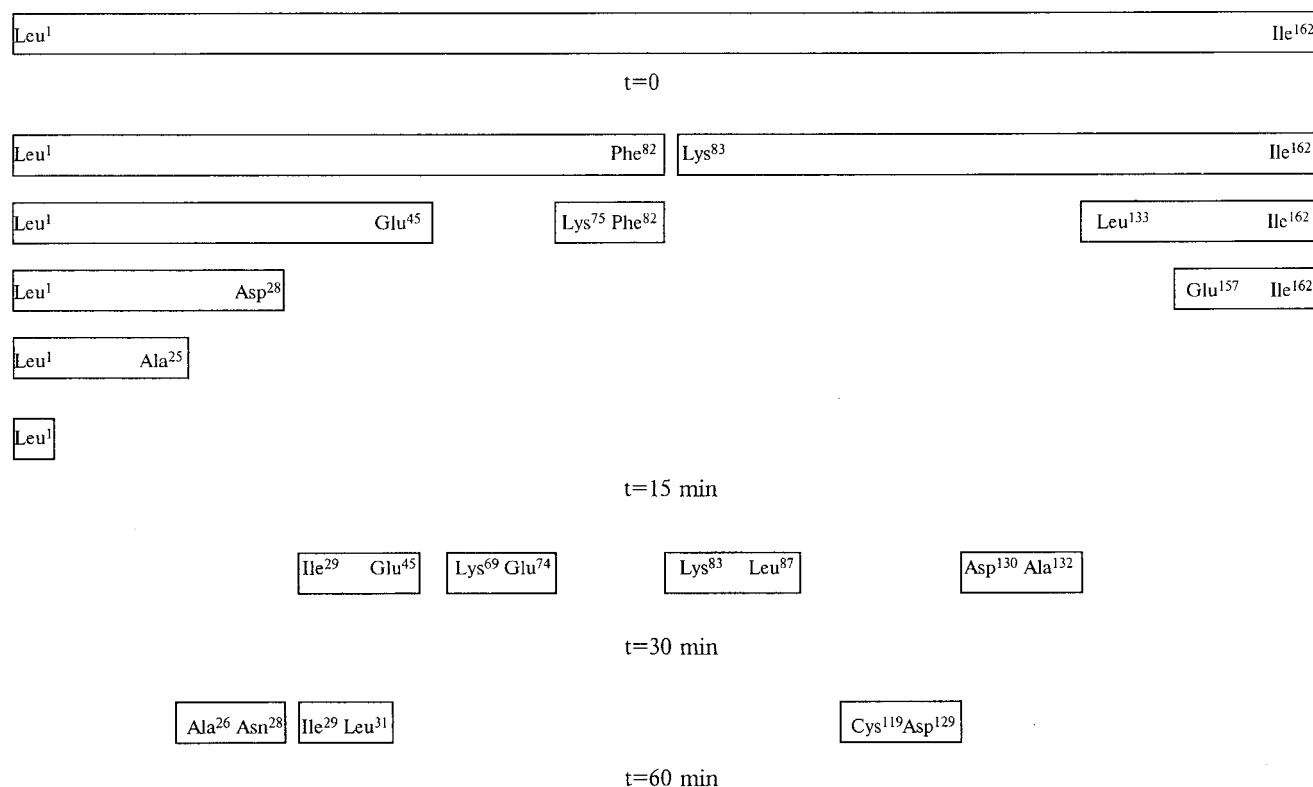


Figure 9. Major components of the proteolytic reaction, as it was simulated by the PROTEOLYSIS program. Dimension of the active site of pepsin was accepted to be 5 amino acid residues ($k_d/k_h = 4$).

tide bond hydrolysis. Very close distributions were found for hydrolysates simulated for 60 min of proteolysis (Fig. 7, $k_d/k_h = 0.04$, $D = 2.5\%$) and for that simulated for 12 min (Fig. 8, $k_d/k_h = 4$, $D = 1.8\%$). In spite of the different demasking parameters, the overall cleavage sites, which were determined on the basis of the fragmentations shown in Fig. 7 and 8, were found to be the same.

3.3 Comparison of the predicted cleavage sites with experiment

The hypothetical decrease of the dimension of the enzyme active site from 10 to 5 amino acid residues leads to a change of the peptide fragment distribution, as it is shown in Fig. 9. But the further decrease of dimension from 5 to 3 gives less realistic results. Table 1 shows 8 coincident cleavage sites, which were identified in the hydrolysates of BLG modified by two methods: with application of high pressure [10] and by the action of ethanol [12]. These coincident peptide bonds were assumed to be hydrolysed independently of the method of globule disturbance, and one can expect, that these bonds can be predicted by the PROTEOLYSIS program. As one can see (Table 1), the computer modelling has allowed to predict a half of experimentally determined cleavage sites, when the dimension of the active site was 10 and 5. At the dimension of the active site of 3, the prediction power of simulations was reduced that can be anticipated because of the loss of information about the pepsin specificity (eq. 5).

This result was obtained in approximation of a limited number of identified peptides (10 peptides), while one should realize that a further increase of the peptide number leads to an

Table 1. The influence of the dimension of the enzyme active site on the prediction of the cleavage sites in the partially demasked BLG.

Cleavage site*	Secondary structure	Result of prediction		
		Dimension of the active site of pepsin		
		3	5	10
Asp ¹¹ ↓Ile ¹²				+
Trp ¹⁹ ↓Tyr ²⁰				
Leu ³¹ ↓Leu ³²		+	+	+
Leu ⁵⁴ ↓Glu ⁵⁵	β-sheet			
Phe ⁸² ↓Lys ⁸³	β-sheet	+	+	
Ala ¹³² ↓Leu ¹³³	α-helix		+	+
Leu ¹⁴⁹ ↓Ser ¹⁴⁸	β-sheet			
Leu ¹⁵⁶ ↓Glu ¹⁵⁷			+	+

* List of the cleavage sites includes coincident cleavage sites, which were determined in both hydrolysates of BLG modified by high pressure [10] and by the action of ethanol [12].

increase in the number of cleavage sites. It is also necessary to stress, that these bonds were computed only on the basis of the primary structure of substrate, without X-ray data.

The set of the predicted cleavage sites (Figs. 7–9) consists of 15 peptide bonds Thr⁴-Gln⁵, Asp¹¹-Ile¹², Ala²⁵-Ala²⁶, Asp²⁸-Ile²⁹, Leu³¹-Leu³², Leu³²-Asp³³, Glu⁴⁵-Leu⁴⁶, Gln⁶⁸-Lys⁶⁹, Glu⁷⁴-Lys⁷⁵, Phe⁸²-Lys⁸³, Leu⁸⁷-Asn⁸⁸, Ala¹¹⁸-Cys¹¹⁹, Asp¹²⁹-Asp¹³⁰, Ala¹³²-Leu¹³³, Leu¹⁵⁶-Glu¹⁵⁷. The native BLG molecule consists of β-sheet, which is formed by 9 strands (residues 16–27, 39–44, 47–58, 62–76, 80–84, 89–97, 102–109,

115–124 and 145–150) and one α -helix (130–140) [21]. Ordered structures include 58% of all residues in the native BLG. Thus, 9 from 15 predicted cleavage sites (60%) are involved in the ordered structures or attached to them.

It is well known that disordered segments of peptide chain possess conformational flexibility and susceptibility to proteolytic attack. Because the PROTEOLYSIS program does not take into consideration the secondary structure of protein, one could anticipate more precise results of the computer prediction for disordered segments, while for β -sheets and α -helix, the favourable conformation for hydrolysis might be unrealizable. In fact, among 5 peptide bonds, which were theoretically predicted and experimentally determined (Table 1), 2 bonds (Phe⁸²-Lys⁸³ and Ala¹³²-Leu¹³³) belong to ordered segments.

References

- [1] Grassie, N., Chemistry of high polymer degradation processes. Butterworths Scientific Publ., London 1956.
- [2] Vorob'ev, M. M., Computer modelling of proteolysis. Proc. First Int. Symposium on: Mathematical modelling and simulation in agriculture and bioindustries, Part 2, ULB, Brussels 1995.
- [3] Antonov, V. K., Chemistry of proteolysis (Rus.). Nauka Publ., Moscow 1983.
- [4] Adler-Nissen, J., Enzymatic hydrolysis of food proteins. Elsevier, London 1986.
- [5] Vorob'ev, M. M., S. V. Vitt and V. M. Belikov, *Nahrung* 31 (1987) 331.
- [6] Vorob'ev, M. M., E. A. Paskonova, S. V. Vitt and V. M. Belikov, *Nahrung* 30 (1986) 995.
- [7] Vorob'ev, M. M., I. Yu. Levicheva and V. M. Belikov, *Prikl. Biochem. Mikrobiol. (Engl. Transl.)* 32 (1996) 237.
- [8] Vorob'ev, M. M., G. Parent and L. Savoie, *Nahrung* 40 (1996) 248.
- [9] Reddy, M. I., N. K. D. Kella and J. E. Kinsella, *J. Agric. Food Chem.* 36 (1988) 737.
- [10] Dufour, E., G. Herve and T. Haertlé, *Biopolymers* 35 (1994) 475.
- [11] Chobert, J.-M., L. Briand, V. Grinberg and T. Haertlé, *Biochim. Biophys. Acta* 1248 (1995) 170.
- [12] Dalgalarondo, M., E. Dufour, J. M. Chobert, C. Bertrand Harb and T. Haertlé, *Int. Dairy J.* 5 (1995) 1.
- [13] Turgeon, S. L., S. F. Gauthier, D. Molle and J. Leonil, *J. Agric. Food Chem.* 40 (1992) 669.
- [14] Schlimme, E., and H. Meisel, *Nahrung* 39 (1995) 1.
- [15] Vorob'ev, M. M., Computer simulation of hydrolysis kinetics of polymer degradation. Proc. Int. Symposium on: Computer assistance to chemical research, Inst. of Organic Chem. RAS, Moscow, 1996.
- [16] Adler-Nissen, J., *J. Chem. Tech. Biotechnol.* 34B (1984) 215.
- [17] Rothenbuhler, E., and J. E. Kinsella, *J. Agric. Food Chem.* 33 (1985) 433.
- [18] Kato, A., K. Komatsu, K. Fujimoto and K. Kobayashi, *J. Agric. Food Chem.* 33 (1985) 931.
- [19] Powers, J. C., D. Harley and D. V. Myers, *Adv. Exp. Biol. Med.* 95 (1977) 141.
- [20] Lemieux, L., and J. Amiot, *J. Chromatogr.* 519 (1990) 299.
- [21] Papiz, M. Z., L. Sawyer, E. E. Eliopoulos, A. C. T. North, J. B. C. Findlay, R. Sivaprasadarao, T. A. Jones, M. E. Newcomer and P. J. Kraulis, *Nature* 324 (1986) 383.

Received: 27 March 1997.

Accepted: 21 October 1997.