A new approach for modelling and optimizing batch enzymatic proteolysis

Sophie Beaubier, Xavier Framboisier, Frantz Fournier, Olivier Galet, Romain Kapel

Manuscript

A new approach for modelling and optimizing batch enzymatic proteolysis

Sophie Beaubier[1, 2], Xavier Framboisier[1], Frantz Fournier[1], Olivier Galet[2], Romain Kapel[1]*

1 *Laboratoire Réaction et Génie des Procédés, UMR-7274, plateforme SVS, 13 rue du bois de la Champelle, F-54500 Vandœuvre-lès-Nancy, France*

2 *Avril SCA, 11 rue de Monceau, F-75008 Paris, France*

Email address for corresponding author:
**** romain.kapel@univ-lorraine.fr***

[1]Abbreviations

aa: amino acids ; AU: Anson Units ; DH: degree of hydrolysis ; DoE: design of experiments ; E/S: enzyme/substrate ratio ; k: kinetic constant ; Naa: mean number of amino acids by peptide ; pI: isolectric point ; RA: rapeseed albumins ; RE: relative error ; RSD: relative standard deviation ; SE-HPLC: size-exclusion high-performance liquid chromatography ; T: temperature ; Xp: protein conversion rate

Abstract:

In this manuscript, an original approach is described for modelling and optimizing batch enzymatic proteolysis. It was the first time that a multicriteria optimization methodology of the enzymatic proteolysis that integrates enzymatic cost and reaction duration was proposed. First, a simulation procedure was developed to predict the kinetics of the protein conversion rate and the degree of hydrolysis at any set of pH, temperature, and Enzyme/Substrate ratio conditions. This was achieved by a hybrid approach based on second order kinetic models and design of experiments methodology**.** The applicability of the methodology was positively validated with the hydrolysis of rapeseed albumins by Alcalase 2.4L. The ANOVA analysis showed that both models were reliable (R^2 for k_{Xp}= 0.95; for k_{DH}= 0.85) and no significant lack of fit was observed (p-value < 0.05). The approach was also validated with good and significant (p-value < 0.05) correlations between experimental and predicted values for the rapeseed albumins hydrolysis with two other proteases in both existing proteolysis mechanisms. A generic multicriteria optimization tool was then applied to search for the best reaction duration/ enzymatic cost trade-offs. The kinetic equations were implemented in a genetic-evolutionary algorithm to generate the Pareto's front (duration/ enzymatic cost trade-offs) and domain (corresponding sets of operating conditions). The approach offers an attractive tool for industrial to reduce important reaction costs to produce hydrolysate of interest. This easy-to-use methodology would present important advances in proteolysis process implementation.

Keywords*:* proteolysis; kinetics; modelling; operating conditions; production optimization; enzymatic cost

1. Introduction

Enzymatic proteolysis has been widely used in food industry for decades. The process is commonly implemented to improve protein functional properties [1], digestibility, to reduce their allergenicity [2], or to produce bioactive peptides [3-4]. In this process, proteases catalyze peptide bonds hydrolysis at more or less specific protein sequence portions (protease specificity); [5]. It results in the production of complex mixtures of peptides called hydrolysates. Hydrolysates properties are related to their peptide or protein composition which depends on protease specificity, reaction advancement and proteolysis mechanism [1].

There are two proteolysis advancement parameters: the degree of hydrolysis (DH) and the protein conversion rate (Xp) [6]. The DH represents the ratio of peptide bonds cleaved on protein peptide bonds expressed in percentage [7]. Xp represents the ratio of protein hydrolyzed on the initial protein [8]. According to Linderstrom-Lang theory, there are two main reaction mechanisms: the one-by-one type and the zipper type mechanisms. In one-by-one mechanism, proteins are progressively hydrolyzed into the same peptide mixture. Xp kinetics are rather slow and peptide compositions remain the same. Only, their concentration increase with the DH. In zipper mechanism, proteins are quickly hydrolyzed in a large peptide population, which are further hydrolyzed into smaller peptides. Xp kinetics are fast and the peptide compositions change with increasing DH. These mechanisms can be efficiently identified by monitoring proteolysis by size exclusion chromatography [8]. These mechanisms are relative to the enzyme specificity and the cleavage site accessibility which depend on the protein structure [9-10].

The main operating conditions that govern batch enzymatic proteolysis are reaction pH and temperature (T). Enzyme and substrate concentration also impact the process. These are most often gathered into a single Enzyme/Substrate ratio (E/S) parameter in proteolysis studies [11-13]. These three operating conditions act on protease catalytic properties and protein structure. Hence, they both modulate reaction kinetics and proteolysis mechanism (and consequently the peptide composition); [11;13-15]. Enzymatic proteolysis reactions are associated to three main performance criteria. The first one is the level of targeted functionality or property. The two others are techno-economic criteria: the enzymatic cost and the reaction duration.

To date, proteolysis optimization approaches only consist in identifying the appropriate set of operating conditions that maximize hydrolysate functionality. To do so, empirical “one factor at the time” or “black box” design of experiments (DoE) approaches [12-13; 16] are classically applied. DoE approaches can also be applied to identify the DH related to a maximized functionality [17-18]. However, to our knowledge, no methodology took the two other criteria (reaction duration and enzyme cost) into consideration. This constitutes a strong bottleneck in proteolysis development since the enzyme cost is considered as the main killer for industrial applications. Since hydrolysate functionality is either related to DH or Xp in a given protease and proteolysis mechanism, an original approach to do so would consist in implementing DH or Xp kinetics model in numerical tools for multicriteria optimization for searching the best cost/ duration trade-off.

Kinetics model based on Michaelis-Menten rate equation [19-22], order 1 [23-25] or 2 models were used with success to fit proteolysis kinetics [26-28]. Recently, Deng et al.

[28] demonstrated that a 2nd order model better described DH kinetics of tryptic hydrolysis of α-lactalbumin and β-casein. The main limitation for kinetics simulation lies on calculate kinetic terms whatever proteolysis conditions (pH, T and E/S). Recently, Valencia et al. [23] proposed a modeling methodology that predicted peptide bond hydrolysis kinetics as a function of operating conditions. To do so, they correlated T and E/S to kinetic terms (obtained by data fitting under different conditions) using a DoE methodology. This approach was validated with salmon muscle and whey protein proteolysis by Alcalase 2.4L. Such an approach could be easily adapted to get DH and Xp kinetic equations as a function of each proteolysis conditions.

The aim of present work was to propose an original methodology of multicriteria optimization of the enzymatic proteolysis that integrates enzymatic cost and reaction duration. To do so, DH and Xp kinetics (DH or Xp) simulation procedure based on the works of Valencia et al. [23] was proposed. In this procedure, 2nd order kinetic model parameters were obtained by data regression in different set of conditions. Then, correlations between experimental conditions and kinetic parameters were obtained by DoE methodology. DH and Xp kinetic models were then used in a generic multicriteria optimization tool developed and used with success in our laboratory for various applications related to protein transformations [29-30]. This tool used genetic-evolutionary algorithms exploiting the Pareto's domination concept to generate the Pareto's front and domain [31]. In this case, Pareto's front represented all the acceptable trade-offs between enzymatic cost and reaction duration for a target DH or Xp value while Pareto's domain represented the corresponding sets of operating conditions. The proposed methodology was implemented and validated with the rapeseed albumins

(RA) proteolysis by Alcalase 2.4L. This substrate was chosen because of its increasing interest for food industry applications as an alternative to animal proteins. Alcalase 2.4L is a well-known non-specific serine protease from *Bacillus licheniformis*, which was implemented in most of the studies of enzymatic proteolysis. Validation of the methodology was also achieved with hydrolysis of RA by two other proteases (Prolyve and Flavourzyme) operating in different reaction condition ranges and following different enzymatic mechanisms.

2. Methodology

The overall multicriteria optimization approach requires to be able to simulate DH and Xp kinetics whatever the process operating conditions (pH, T and E/S). Kinetics modelling were based on three main assumptions.

1) DH (and so Xp) follow second order kinetic equations (Eq. 1 and Eq. 2). This hypothesis was supported by Deng et al. [28].

$$DH\ (\%) = DH_{max} - \frac{DH_{max}}{(1 + k_{DH} \times DH_{max} \times t)} \quad (1)$$

$$Xp\ (\%) = Xp_{max} - \frac{Xp_{max}}{(1 + k_{Xp} \times Xp_{max} \times t)} \quad (2)$$

with k_{DH} and k_{Xp}, kinetic constant of DH and Xp kinetics respectively, and $DH_{max,}$ and Xp_{max}, the maximum degree of hydrolysis and the maximum protein conversion rate, respectively.

2) DH_{max} and Xp_{max} parameters depend on protein cleavage site accessibility and protease specificity. Rapeseed albumins structure are highly thermostable [32] but may

be impacted by the pH [33-34]. Hence, in this case, DH_{max} and Xp_{max} only depend on the protease used and the pH.

3) k_{Xp} and k_{DH} parameters depend on both cleavage site accessibility and operating conditions effect on protease catalysis. Hence, these kinetic parameters depend on pH, T and E/S.

These assumptions are valid for a given protein/protease couple in a given enzymatic mechanism. The zipper type and the one-by-one type mechanisms were described according to Linderstrom-Lang theory [9-10]. These mechanisms are related to the protein structure which can be modified according to the reaction conditions (pH, T and E/S). Hence, a preliminary step is necessary for optimization to ensure the reaction mechanism in the range of operating conditions chosen.

According to these assumptions, Xp and DH kinetic models as a function of proteolysis conditions were obtained in three steps (Figure 1):

The first step consisted in choosing an operating conditions area (pH, T and E/S) that covers the activity range of the protease (given by the supplier). From this area, full factorial design is then defined, i.e. the 3 factors (pH, T and E/S) at 3 levels (codified as -1; 0; +1).

The second step consisted in experimental determination of DH_{max} and Xp_{max} parameters for the three pH levels (after 24 hours of reaction time). Polynomial regressions of the data were used to get correlations between pH and DH_{max} and Xp_{max}.

The third step established correlations between k_{Xp} and k_{DH} and the 3 operating conditions (pH, T and E/S). To do so, k_{Xp} and k_{DH} are determined for each condition of

the DoE by fitting the models (Eq. (1) and Eq. (2)) with experimental kinetics (for 3 hours). Polynomial correlations between kinetic parameters and operating conditions are established by DoE methodology.

Then, the model obtained were used to search for the best duration/ enzymatic cost trade-offs. To do this, a genetic-evolutionary algorithm exploiting the Pareto's domination concept is used to identify the Pareto's front (set of all acceptable duration/ enzymatic cost trade-offs) and Pareto's domain (the corresponding operating conditions area).

The proposed methodology is summarized in Figure 1.

3. Material and methods

3.1. Materials

Alcalase 2.4L was purchased from Novozymes (Bagsvaerd, Denmark). This enzyme is a non-specific serine endopeptidase from *Bacillus licheniformis* with a specific activity of 2.4 Anson Units (AU-A). g^{-1} and a density of 1.18 g. mL^{-1}. Flavourzyme (1000L) and Prolyve (PAC 30L) were purchased from Novozymes (Bagsvaerd, Denmark) and Soufflet Biotechnologies (France), respectively. The industrial rapeseed meal was supplied by the AVRIL Group (France) and the albumin fraction was isolated according to Nioi et al. [35]. HPLC grade eluents (water and acetonitrile) were both purchased from Fisher Scientific (Geel, Belgium). Synthetic peptides used for the column calibration were purchased (GeneCust, Dudelange, Luxembourg). All other chemicals and reagents used were of analytical grade.

3.2. Proteolysis reactions

Rapeseed albumins were hydrolyzed by Alcalase 2.4L, Flavourzyme and Prolyve, individually. The recommended domains of operating conditions (pH, T and E/S) are displayed in Table 1. These conditions were given by suppliers. Proteolysis were carried out in a stirred and thermally controlled batch reactor of 50 mL under magnetic agitation. The protein solution was prepared by suspended the freeze-dried isolate in distilled water. Initial protein substrate concentration was 1% (w/v; in terms of protein content, N × 6.25). pH and temperature of the substrate solution were then adjusted at the desired values. Eventually, proteases were added according to the chosen E/S ratio. Enzymatic activities were checked prior use. pH was maintained constant by addition of NaOH 0.5N using an automatic titration system (902 Titrando, Metrohm Ltd., Herisau, Switzerland). Samples were taken for further analysis, after 15 and 30 min, 1 and 3 hours, i.e. 4 points by kinetic experiment. The reaction was stopped by pH shift assuring the inactivation of the enzyme prior storage at -20 °C.

The maximum protein conversion rate (Xp_{max}) and the maximum degree of hydrolysis (DH_{max}) were determined after 24 hours of hydrolysis reaction with Alcalase 2.4L at 60 °C, E/S of 1/15 ($1.73E^{-01}$ AU-A. g^{-1} protein) in the 3 pH conditions of the matrix, i.e pH 7.0; 8.5 and 10.0.

3.3. Estimation of Xp and DH parameters

Xp and DH values were simultaneously determined from SE-HPLC chromatograms according to Beaubier et al. [8]. SE-HPLC analysis was made using a Superdex peptide 10/300 GL column (10 × 300 mm, GE Healthcare, USA) as reported previously [36]. 10 µL were injected onto the column kept at 35 °C, connected to a Shimadzu model

LC20 system (Shimadzu Corporation, Japan). The isocratic elution was performed at 0.5 mL.min^{-1} with a water/acetonitrile/TFA (69.9/30/0.1; v/v) solvent. UV signal was recorded at 214 nm using a cell with an optical path of 0.5 cm. Column was calibrated with synthesized standard peptides eluted in the same conditions. Chromatograms at 214 nm were exported in Excel spreadsheets to determine the protein conversion rate and the DH values.

Eq. 3 was applied to quantify the protein conversion rate of each sample.

$$Xp_t\,(\%) = \left(1-\frac{A_t}{A_0}\right)\times 100 \qquad (3)$$

with A_t and A_0 the integrated protein signal at the proteolysis duration “t” and before enzyme addition.

Eq. 4 was applied to calculate the degree of hydrolysis:

$$DH_t\,(\%) = \frac{n_p}{n_{aa}}\times Xp_t\,(\%) \qquad (4)$$

with, n_p and n_{aa} are respectively the amount of peptides and amino acids (in mole) in the hydrolysate calculated according to Bodin et al. [36] as follows:

$$n_p = Qv \int \left(\frac{A_x}{\varepsilon_x l}\right) dt \qquad (5)$$

$$n_{aa} = Qv \int \left(\frac{A_x}{\varepsilon_x l} \times Naa_x\right) dt \qquad (6)$$

with Q_v, the elution flow rate and dt, a fraction of the elution time. A_x and ε_x are respectively the absorbance and the molar extinction coefficient for point ‘x’ and l the path length of the light beam. Naa_x is the mean amino acid number of the peptide

mixture for each point ‘x’ and it corresponds to the ratio between the molar mass of the point ‘x’ and the mean amino acid molar mass of the hydrolysate.

3.4. Design of experiments

Correlation models between pH, T and E/S and kinetic parameters (k_{Xp} and k_{DH}) were obtained by DoE. To do so, a full factorial design was used. The matrix was composed of three factors (pH, T and E/S) at 3 coded levels (Table 2). Hence, the matrix consisted of 30 experiments (27 combinations and 3 replicates at the center point).

The response considered were k_{Xp} (Y_1) and k_{DH} (Y_2) obtained by regression of experimental kinetics with second order kinetic models (Eq. 1 and 2). Correlation models were polynomial equations with the three operating conditions as variables (X_1, X_2 and X_3) and intercept (b_0), linear (b_i), quadratic (b_{ii}) and interaction (b_{ij}) coefficients:

$$Y_x = b_0 + \sum_{i=1}^{3} b_i X_i + \sum_{i=1}^{3} b_{ii} X_i^{2} + \sum_{i<j=1}^{3} b_{ij} X_i X_j \quad (7)$$

MODDE 12.0.1 software (Umetrics, Sartorius Stedim Biotech, Sweden) was used to fit the data to the polynomial equation with adapted transformation if necessary. The regression coefficients were obtained by multivariable regression. An ANOVA was used to evaluate the statistical significance of these coefficients. Final correlation models were obtained by suppressing the non-significant terms (p-value > 0.05). The coefficients R^2 and Q^2, the relative standard deviation (RSD), the reproducibility and the lack of fit were analyzed using the MODDE software to characterize the goodness of fit for each model (Y_1 and Y_2). Response surface plots were also drawn with this software.

3.5. Pareto’s front and domain generation

To identify the best duration/ enzymatic cost trade-offs, a routine based on genetic-evolutionary algorithms was applied [31]. The objective function corresponded to the minimization of 2 antagonist criteria: the enzymatic cost and the reaction duration, for a given DH or Xp value. The kinetics models (DH and Xp) were exploited to describe those 2 criteria. The reaction time (t) was directly isolated from the model equations 1 or 2. The enzymatic cost was calculated from E/S value and the enzyme price, in $€_{enzyme}$. kg^{-1} of substrate. In this study, the cost of Alcalase 2.4L provided by Novozymes (Bagsvaerd, Denmark) was 48 €. kg^{-1} of protease.

The Pareto's front represented the set of all acceptable trade-offs of the performances (cost and duration) while the Pareto's domain represented the corresponding operating conditions area. A large population (2000 individuals) was randomly generated by the program. Each individual of the population was represented with a set of the 3 operating conditions (pH, T and E/S) and was characterized by its phenotype, depicted by two chromosomes (genotype) and a recessive/dominance index. Performance criteria of each individual were evaluated for the given Xp or DH value. Cross-over and mutation operators on homozygote and heterozygote individuals were applied to generate new populations. These operators were also coupled with elitist selection to lead the population, along generation, towards a non-dominated population.

3.6. Model validation and statistical analysis

To validate the predictive model of each response variable (k_{Xp} and k_{DH}), new sets of operating conditions (pH, T and E/S) were implemented in the domain (Table 2). These were different from those included in the full factorial design. For the validation,

proteolysis kinetics were monitored for 24 hours. Validation were also carried out at different initial substrate concentration (up to 5% w/v; in terms of protein content, N × 6.25).

Statistical analysis was performed using the freeware R (3.4.1.) and the Analysis ToolPak of Excel (Microsoft Office, Redmond, Washington, USA) to compare experimental and predicted values. Data were analyzed by using the one-way analysis of variance (ANOVA), after testing normality and homogeneity of variance. Then, the "Tukey test" ($p < 0.05$) was performed to find parameters that are significantly different from each other. The relative error (RE) was determined by comparing the observed and predicted values of each model as follows:

$$RE\ (\%) = \frac{\text{Observed value } (k_x) - \text{Predicted value } (k_x)}{|\text{Predicted value } (k_x)|} \quad (8)$$

RE less than 15% was considered acceptable.

4. Results and discussions

4.1. Modelling rapeseed albumin proteolysis kinetics by Alcalase 2.4L

4.1.1. Experimental kinetics

Figure 2 shows Xp and mean peptide size (expressed in number of aminoacids, N_{aa}) as a function of DH in various operating conditions (pH from 7.0 to 10.0, T from 45 °C to 75 °C at E/S ranging from 1/15 (g/g) to 1/150 (g/g)). This conditions domain was chosen according to supplier data for this enzyme (Table 1). E/S range was selected to fit classical industrial implementation conditions and corresponded to the reported ranges in many papers on proteolysis optimization using DoE methodology [12-13]. For DH values comprised from 0 to 18%, Xp and N_{aa} values were observed between 0 to 100%

and 7 aa to 5 aa. The curves of Xp and N_{aa} versus DH were very close whatever the selected reaction conditions. Values of N_{aa} decreased very slightly (from 7 to 5 aa) and Xp variation was almost linear.

This meant that Alcalase 2.4L can more easily and quickly hydrolyze the intermediate peptides of RA than the intact proteins in the applied conditions [7]. Hence, the weak decrease of peptide size and the linear disappearance of proteins suggested a single enzymatic mechanism of one-by-one type, according to Linderstrom-Lang theory [9-10]. This mechanism seemed to apply in the whole domain of conditions. This was suggested by the fact that a DH value was associated to the same N_{aa} and Xp whatever the operating conditions (as examples, Xp of around 53% and N_{aa} around 6.5 aa at DH 8% and Xp of around 92% and N_{aa} around 5.2 aa at DH 16%). Thus, the operating conditions (pH, T and E/S) only influence the hydrolysis kinetics (DH and Xp) but not the hydrolysis mechanism. This can be explained by the high structural stability of RA in these operating ranges. It has indeed been reported that its structure is highly thermostable and that their far-UV circular dichroism spectra were almost similar in the pH range of the applied hydrolysis [32-33]. Besides, the high Xp value reached (100% at DH around 18%) indicated that Alcalase 2.4L in the selected operating conditions efficiently hydrolyzed RA. This was already observed with various proteins including RA [8]. Hence, the choice of protease and operating conditions domain was appropriate for the methodology development.

4.1.2. Determination of Xp_{max} and DH_{max} parameters

Figure 3A shows Xp and DH kinetics at pH 7.0, 8.5 and 10.0. Hydrolysis curves were classically characterized by two steps. The first one corresponded to a high initial

reaction rate. In the second one, a decrease of the hydrolysis rate over time was observed. Figure 3B shows maximal values of Xp and DH (obtained after 24 h) at these three pH. Correlations models (pH *vs* limit terms) obtained by regressions with polynomial models were displayed in the figure. Model form and terms are displayed in Table 3.

Experimental kinetics and Xp_{max} and DH_{max} parameters were very close for conditions at pH 7.0 and 8.5. These differed more at pH 10.0 with an obvious decrease in limit Xp and DH (from 100% to 89% and from 23% to 15% respectively). Thus, Xp_{max} and DH_{max} depended on pH conditions as expected. This can be explained either by a loss of protease activity or to a change in proteolysis site accessibility on the protein (by protein denaturation or aggregation). Alcalase 2.4L is known to have high stability over a wide pH and temperature range (supplier data). So, changes in Xp_{max} and DH_{max} were probably explained by a decrease of the cleavage site accessibility. Proteins are known to aggregate around their isoelectric point (pI) and rapeseed albumins have a high pI (pI > 10; [37]).

4.1.3. Modelling of the effect of pH, T and E/S on second order kinetic parameters (k_{Xp} and k_{DH})

Correlations between kinetic constants (k_{Xp} and k_{DH}) and the proteolysis conditions (pH, E/S, and T) were obtained by DoE. Table 2 shows the matrix of experiments used. For each set of operating conditions in the matrix, Xp and DH kinetics were monitored during 3 h. Kinetics were regressed using second order model in order to get the corresponding k_{Xp} and k_{DH}. Figure 4 shows the comparison between experimental and calculated DH and Xp at the different kinetic points (at 0.25, 0.5, 1 and 3 h) in each experimental

condition of the matrix. Good linearity between the Xp values (R^2= 0.98) and between the DH values (R^2= 0.98) were found. The “Tukey test” ($p < 0.05$) was also performed and no significant differences were observed between the experimental and calculated values. This validated the application of the modified second order model implemented by Deng et al. [28] to reliably describe the enzymatic kinetic system.

Table 4 shows regression coefficients of predicted quadratic polynomial models for k_{Xp} and k_{DH} yielded by DoE. The intercept, linear, quadratic and interaction coefficients are presented for both models. Among all operating conditions, the responses k_{Xp} (Y_1) and k_{DH} (Y_2) were significantly influenced ($p < 0.05$) by linear and quadratic terms of E/S. The pH had a negative effect on both responses. Interaction between pH and E/S was found for k_{Xp} and between T and E/S for k_{DH}.

Table 4 also displays the statistical results of the analysis of variance (ANOVA test). It appeared that the probabilities for the regression of both responses were significant at 95%. This indicated that the models were statically good. Moreover, both models for k_{Xp} prediction (R^2 = 0.946) and for k_{DH} prediction (R^2= 0.854) fitted well the experimental data. Q^2 values (Table 4) expressed useful models to well predict new data. The probabilities for the lack of fit were not significant at 95%. Thus, these models allowed good prediction and could be applied to simulate the hydrolysis kinetics whatever the operating conditions (pH, T and E/S) in the considered domain.

Figure 5 A and B shows the impact of pH and T on k_{Xp} and k_{DH} for different E/S values (1/150; 1/27 and 1/15; g/g). Obviously, both kinetic constants increased with E/S values. This may be explained by an increase in the number of active sites at higher E/S values,

resulting in an increase in the number of cleaved peptide bonds. Hence, the higher the E/S value, the faster the kinetics whatever the values of pH and T.

The effect of the pH appeared slightly different on k_{Xp} and k_{DH}. k_{Xp} values decreased from pH 7.0 to 8.5 and increased beyond that pH while k_{DH} values kept increasing throughout the pH range. Whatever E/S value, this figure showed that optimum of enzymatic activity was in a pH range comprising between 9.5 and 10.0. As specified above, the pH did not modify the structure of RA in this range [32-33]. Hence, this can be explained by the maximum range activity of Alcalase 2.4L between pH 9.0 and 10.0 (> 80% according to supplier) rather than a change in proteolysis site exposure.

The analysis of the temperature influence on both kinetic constants showed a clear trend. Kinetic constants increased with increase in temperature up to 60 °C and slightly decreased beyond that temperature. This optimum reflects the classical trade-off between positive effect of T on chemical reaction and protease denaturation. The same behavior had already been reported for the hydrolysis of food proteins (chicken meat) with Alcalase 2.4L [38].

The fastest kinetics of Xp and DH could thus be found in the optimum operating area of pH 10.0, E/S = 1/15 ($1.73E^{-01}$ AU-A. g^{-1} protein) and around 60 °C. These observations were consistent with supplier recommendations concerning Alcalase activity (optimal activity at pH from 8.0 to 10.0 and T from 60 to 75 °C, based on enzymatic activity assays with synthetic protein substrates). The data given by our methodology gives specific insight on rapeseed albumins proteolysis by Alcalase 2.4L and might differ with another protein.

4.2. Validation of the simulation of hydrolysis kinetics

Two new sets of conditions (A and B) in the DoE matrix were implemented to validate the simulation of hydrolysis kinetics**.** Table 5 shows the operating condition sets and the value ranges of kinetic constants evaluated by the correlation models. The prediction interval for kinetic constants were calculated using the relative standard deviation (RSD, Table 4) of each model. Condition B was achieved at a higher substrate concentration (5% w/v; in terms of protein content, N × 6.25) to get closer of industrial implementations. This would also validate that the observed E/S effect is independent of the initial substrate concentration and protease concentration values. Experimental k_{Xp} and k_{DH} in the two validation conditions are also shown in Table 5.

For condition A, both experimental k_{Xp} and k_{DH} were found in the prediction intervals of the models. This validated the simulation methodology for a given substrate concentration. For condition B, the experimental value of k_{Xp} was found in the prediction interval. The experimental value of k_{DH} was slightly higher than the upper limit of this interval. This would suggest that the methodology would apply in a particular substrate concentration range. Nevertheless, the experimental values of DH during the reaction were compared to the predicted values of DH and a relative error below 15% was found for each point of the enzymatic kinetic.

Figure 6 shows the prediction intervals of the kinetic curves of Xp (A) and DH (B) and the experimental kinetics for both points of validation. The prediction intervals were really close due to the very low value of RSD (Table 4). This can be explained by the fact that the methodology applied to monitor the experimental kinetic curves used SE-HPLC which is a sensitive and precise tool. From the beginning of the reaction up to 3

hours of reaction, experimental Xp and DH kinetics were in the prediction intervals in both condition A and B. Beyond 8 hours, some experimental points revealed slightly over the prediction interval. A slight underestimation of the models was thus observed at high reaction time, but it remained acceptable since the relative error was below 15%. This can be explained by the assumption 2 concerning DH_{max} and Xp_{max}. The estimation of these parameters could be slightly more complex with influence of other parameters, like reaction temperature.

Figure 6 C and D expose the experimental versus predicted values of Xp and DH, respectively. Correlations between experimental and predicted Xp and DH were analyzed. Significant and good correlations were found for Xp values ($R^2 = 0.995$, $p < 0.05$) and DH values ($R^2 = 0.997$, $p < 0.05$). The “Tukey test” was also performed to compare experimental and predicted values. No significant differences were observed between the experimental and calculated values of Xp (p-value for point A = 0.817; for point B = 0.947) and DH (p-value for point A = 0.922; for point B = 0.891). Hence, both models of prediction of k_{Xp} and k_{DH} were validated for both validation points. Moreover, the observed effect of E/S was independent of the initial substrate concentration and the proposed modelling approach may therefore be applied on a larger scale (close to the industrial scale).

The described methodology for establishing models was also applied to the RA hydrolysis with two other proteases (Flavourzyme and Prolyve). These proteases were selected because they were active in different operating pH and T ranges compared to Alcalase 2.4L (Table 1).

Proteolysis operating conditions applied in this study are shown in Table 6. Figure 7 A and B show the Xp and the N_{aa} variations versus DH for these hydrolysis cases. A significant decrease of N_{aa} was observed (8 to 3 aa with Flavourzyme and 11 to 6 aa with Prolyve) with a fast and non-linear increase of Xp values (at DH 5%, 40% with Flavourzyme and 60% with Prolyve). Interestingly, these observations indicated that proteolysis followed the second type of proteolysis mechanism (zipper mechanism according to Linderstrom-Lang theory [9-10]). This meant that Flavourzyme and Prolyve enzymes can more easily and quickly hydrolyze the intact proteins than Alcalase 2.4L in the applied conditions. At pH 4.0 and 5.0, RA are positively charged (pI > 10; [37]) and strong electrostatic repulsions avoiding protein – protein aggregations that often take place around their pI (i.e. under Alcalase conditions). This may explain the difference in initial protein hydrolyze rate with both Flavourzyme and Prolyve compared to Alcalase.

DH_{max}, and Xp_{max} parameters were then experimentally determined [Xp_{max}= 94.8% ± 0.1% and DH_{max}= 36.1% ± 1% for RA hydrolysis with Flavourzyme; Xp_{max}= 98.4% ± 0.9% and DH_{max}= 19.8% ± 0.5% for RA hydrolysis with Prolyve]. Models coefficients for kinetic constants were identified by DoE as described in Fig.1. Each model was statistically analyzed and can be applied reliably [R^2 for k_{Xp} model = 0.939 for Flavourzyme case and 0.985 for Prolyve case; R^2 for k_{DH} model = 0.98 for Flavourzyme case and 0.98 for Prolyve case]. Hence, new sets of operating conditions were implemented for both hydrolysis cases to validate the models, presented in Table 6.

Figure 8 shows the prediction intervals of the kinetic curves of Xp (A) and DH (B) and the experimental kinetics as well as the experimental versus predicted values of Xp (C) and DH (D) for both hydrolysis cases. Statistical comparisons were also made between

experimental and predicted values. No significant differences were observed between the experimental and calculated values of Xp (p-value for Prolyve case= 0.816; for Flavourzyme case= 0.944) and DH (p-value for Prolyve case= 0.887; for Flavourzyme case= 0.919). Both models of prediction of k_{Xp} and k_{DH} were also validated for RA hydrolysis with Prolyve and with Flavourzyme. In this way, the proposed modelling methodology can be easily implemented for the simulation of the whole hydrolysis kinetics of Xp and DH with a high goodness of fit whatever the enzyme and the type of enzymatic mechanism.

4.3. Search for the best duration/ enzymatic cost trade-offs

The obtained kinetic models of DH and Xp for the hydrolysis of RA with Alcalase 2.4L were then used in a multicriteria optimization tool [31]. Figure 9 shows the Pareto's fronts of 25; 50 and 75% of Xp (A) and of 3; 5; 8 and 10% of DH (B). These large Xp and DH ranges cover all potential applications such as improving digestibility or functional properties and obtaining bioactive peptides [1]. As expected, a decrease of the enzymatic cost implied an increase in the reaction time for both Xp and DH Pareto's fronts. This was explained by the E/S effect on kinetic constants, whatever pH and T conditions (Fig. 5A and B). This trend was particularly marked at high Xp or DH. At Xp = 75%, a reduction of 50% of the enzymatic cost (from 3.2 €. kg^{-1} protein to 1.6 €. kg^{-1} protein) was reached at the expense of an increase of the reaction duration of approximatively 40% (from 1.4 to 2.4 hours of reaction).

Enzymatic reactions are industrially implemented between 1 and 5 hours due to the technical constraints. Indeed, there may be a risk of microbial development beyond this duration range. Below 1 h of reaction, the reaction stops (by T or pH shift) could be too

close to the reaction implementation duration. At low Xp (25 and 50%) and DH values (3 and 5%), steeper Pareto's fronts were found. The reaction time ranges were thus limited, and implementation would be preferentially made around 1 hour. In this way, a factor 4 could be saved in terms of enzymatic cost (from 3.2 €. kg^{-1} protein to 0.8 €. kg^{-1} protein) at 5% of DH. At high Xp and DH values, the reaction times covered a broader range (from 1 h and 24 min to 6 h and 15 min at Xp 75% and from 54 min to 5 h and 36 min at DH 10%). Hence, at DH 10%, the reaction implementation of 5 hours instead of 1 hour would reduce the enzymatic cost by a factor 6 (from 2.6 €. kg^{-1} protein to 0.4 €. kg^{-1} protein).

The Pareto's fronts can be translated to the Pareto's domains by obtaining the optimal combinations of operating conditions (pH, T and E/S) for each acceptable trade-off, as shown in Figure 10. It exposes, in a three-dimension space, the optimal area for Xp of 50% (A) and for DH of 10% (B). As expected, the figure indicates that optimal conditions (or acceptable trade-offs) can be obtained in a wide range of E/S ratio (from 1/15 to 1/150 g/g). This is due to the fact that one of the criteria (the enzyme cost) is directly related to this factor and the two criteria (enzyme cost and reaction duration) are antagonist. More surprisingly, all acceptable trade-offs for Xp of 50% and for DH of 10% were found for very similar pH and T duets (pH 10.0 and 62°C for 50% Xp, and pH 9.4 and 59°C for 10% DH). For those cases, the optimal areas were close to the conditions that maximize kinetic constants (Fig. 5A and B). It meant that, in the spite of interactions observed between E/S, pH and T (Table 4) on their impact on proteolysis kinetics, the best reaction duration/ enzymatic cost trade-off has to be searched by tuning E/S ratio in the optimal pH / T conditions. To our knowledge, this is the first time that this interesting

effect was observed. This observation is valid for the Alcalase 2.4L / RA system but this effect may vary for other enzyme/substrate system.

In any cases, once the Pareto's front and domain applied according to the proposed methodology, it is then for each industrial decision maker to choose the best solution by defining its own preferences between criteria for the process implementation. A decision support tool [31] could also help to find the most appropriate duration/ enzymatic cost trade-off.

5. Conclusion

The present study proposed an original approach to rationally identify the optimal conditions to implement for producing hydrolysate in terms of reaction duration/ enzymatic cost trade-off. To our knowledge, this has never been reported before and constitutes a significant improvement since the enzymatic cost is often considered as a killer for implementation of batch enzymatic proteolysis. This required the development of a reliable kinetic modelling approach as a function of the operating conditions. The novelty of the proposed modelling approach is the simulation of both the Xp and DH hydrolysis kinetics, whatever the pH, T and E/S conditions as well as the enzyme/substrate couple. The kinetic model establishment has been validated in various cases of RA hydrolysis with different proteases and in both existing enzymatic mechanism types (one-by-one and zipper). Goodness of fit and predictability of the methodology was highlighted for each enzyme/substrate couple. A generic multicriteria optimization tool was then used to search for the best duration/ enzymatic cost trade-offs and identify the optimal reaction conditions for a given Xp or DH value (associated to an

optimum performance criterion). This methodology can be used to compare the catalytic and cost efficiency of different proteases, and to aptly describe the hydrolysis reaction of different enzyme/substrate couples and enzymatic mechanisms. This is a significant advance since the optimization approaches previously used in proteolysis applied DoE only to maximize a hydrolysate functionality.

This study represents a reliable tool that can be applied to any batch proteolysis process, with a minimal number of experimentations. The Pareto's front and domains obtained by the multicriteria optimization approach could not be obtained by sole experimental approach. Interestingly, this innovative modelling approach can be easily associated to the prediction of hydrolysate fractionation by ultrafiltration [39] to simulate the production of target peptides in ultrafiltration compartments of an enzymatic membrane reactor. Finally, it can arouse interest for industrials to reduce the production costs of hydrolysates while meeting technical implementation constraints. By this way, it can improve the processing of protein isolates to develop new bioactive, functional or nutritive peptides with lower costs. Hence, it could help the sustainable valorization of by-products (like plant meals, slaughterhouse or fish by-products) in food, cosmetic or pharmaceutical industry.

Acknowledgements:

Conflicts of interest: none

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

References:

[1] Panyam, D., & Kilara, A. (1996). Enhancing the functionality of food proteins by enzymatic modification. *Trends in food science & technology*, *7*(4), 120-125.

[2] Cabanillas, B., Pedrosa, M. M., Rodriguez, J., Muzquiz, M., Maleki, S. J., Cuadrado, C., ... & Crespo, J. F. (2012). Influence of enzymatic hydrolysis on the allergenicity of roasted peanut protein extract. *International archives of allergy and immunology*, *157*(1), 41-50.

[3] Gauthier, S. F., Pouliot, Y., & Saint-Sauveur, D. (2006). Immunomodulatory peptides obtained by the enzymatic hydrolysis of whey proteins. *International dairy journal*, *16*(11), 1315-1323.

[4] Harnedy, P. A., & FitzGerald, R. J. (2012). Bioactive peptides from marine processing waste and shellfish: A review. *Journal of functional foods*, *4*(1), 6-24.

[5] Tavano, O. L. (2013). Protein hydrolysis using proteases: An important tool for food biotechnology. *Journal of Molecular Catalysis B: Enzymatic*, *90*, 1-11.

[6] Adler-Nissen, J. (1976). Enzymic hydrolysis of proteins for increased solubility. *Journal of Agricultural and Food Chemistry*, *24*(6), 1090-1093.

[7] Adler-Nissen, J. (1986). *Enzymic hydrolysis of food proteins*. Elsevier applied science publishers.

[8] Beaubier, S., Framboisier, X., Ioannou, I., Galet, O., & Kapel, R. (2019). Simultaneous quantification of the degree of hydrolysis, protein conversion rate and mean molar weight of peptides released in the course of enzymatic proteolysis. *Journal of Chromatography B*, *1105*, 1-9.

[9] Vorobèv, M. M., Paskonova, E. A., Vitt, S. V., & Belikov, V. M. (1986). Kinetic description of proteolysis Part 2. Substrate regulation of peptide bond demasking and hydrolysis. Liquid chromatography of hydrolyzates. *Food/Nahrung*, *30*(10), 995-1001.

[10] Linderstrom-Lang, K. (1953). The initial phases of the enzymatic degradation of proteins. *Bulletin de la Société de chimie biologique*, *35*(1-2), 100-116.

[11] Noman, A., Xu, Y., AL-Bukhaiti, W. Q., Abed, S. M., Ali, A. H., Ramadhan, A. H., & Xia, W. (2018). Influence of enzymatic hydrolysis conditions on the degree of hydrolysis and functional properties of protein hydrolysate obtained from Chinese sturgeon (Acipenser sinensis) by using papain enzyme. *Process Biochemistry*, *67*, 19-28.

[12] Chabeaud, A., Dutournié, P., Guérard, F., Vandanjon, L., & Bourseau, P. (2009). Application of response surface methodology to optimise the antioxidant activity of a saithe (Pollachius virens) hydrolysate. *Marine Biotechnology*, *11*(4), 445-455.

[13] Guo, Y., Pan, D., & Tanokura, M. (2009). Optimisation of hydrolysis conditions for the production of the angiotensin-I converting enzyme (ACE) inhibitory peptides from whey protein using response surface methodology. *Food Chemistry*, *114*(1), 328-333.

[14] Yu, L., Sun, J., Liu, S., Bi, J., Zhang, C., & Yang, Q. (2012). Ultrasonic-assisted enzymolysis to improve the antioxidant activities of peanut (Arachin conarachin L.) antioxidant hydrolysate. *International journal of molecular sciences*, *13*(7), 9051-9068.

[15] Cheison, S. C., Leeb, E., Toro-Sierra, J., & Kulozik, U. (2011). Influence of hydrolysis temperature and pH on the selective hydrolysis of whey proteins by trypsin and potential recovery of native alpha-lactalbumin. *International Dairy Journal*, *21*(3), 166-171.

[16] Gbogouri, G. A., Linder, M., Fanni, J., & Parmentier, M. (2004). Influence of hydrolysis degree on the functional properties of salmon byproducts hydrolysates. *Journal of food science*, *69*(8), C615-C622.

[17] Seo, H. W., Jung, E. Y., Go, G. W., Kim, G. D., Joo, S. T., & Yang, H. S. (2015). Optimization of hydrolysis conditions for bovine plasma protein using response surface methodology. *Food chemistry*, *185*, 106-111.

[18] Singh, T. P., Siddiqi, R. A., & Sogi, D. S. (2019). Statistical optimization of enzymatic hydrolysis of rice bran protein concentrate for enhanced hydrolysate production by papain. *LWT*, *99*, 77-83.

[19] Barros, R. M., & Malcata, F. X. (2004). A kinetic model for hydrolysis of whey proteins by cardosin A extracted from Cynara cardunculus. *Food chemistry*, *88*(3), 351-359.

[20] Valencia, P., Pinto, M., & Almonacid, S. (2014). Identification of the key mechanisms involved in the hydrolysis of fish protein by Alcalase. *Process Biochemistry*, *49*(2), 258-264.

[21] Marquez, M. C., & Vázquez, M. A. (1999). Modeling of enzymatic protein hydrolysis. *Process biochemistry*, *35*(1-2), 111-117.

[22] Demirhan, E., Apar, D. K., & Özbek, B. (2011). A kinetic study on sesame cake protein hydrolysis by Alcalase. *Journal of food science*, *76*(1), C64-C67.

[23] Valencia, P., Espinoza, K., Ceballos, A., Pinto, M., & Almonacid, S. (2015). Novel modeling methodology for the characterization of enzymatic hydrolysis of proteins. *Process Biochemistry*, *50*(4), 589-597.

[24] Margot, A., Flaschel, E., & Renken, A. (1997). Empirical kinetic models for tryptic whey-protein hydrolysis. *Process Biochemistry*, *32*(3), 217-223.

[25] Zhou, C., Yu, X., Qin, X., Ma, H., Yagoub, A. E. A., & Hu, J. (2016). Hydrolysis of rapeseed meal protein under simulated duodenum digestion: Kinetic modeling and antioxidant activity. *LWT-food science and technology*, *68*, 523-531.

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

[27] Deng, Y., Butré, C. I., & Wierenga, P. A. (2018). Influence of substrate concentration on the extent of protein enzymatic hydrolysis. *International Dairy Journal*, *86*, 39-48.

[28] Deng, Y., van der Veer, F., Sforza, S., Gruppen, H., & Wierenga, P. A. (2018). Towards predicting protein hydrolysis by bovine trypsin. *Process Biochemistry*, *65*, 81-92.

[29] Slabi, S. A., Mathe, C., Basselin, M., Framboisier, X., Ndiaye, M., Galet, O., & Kapel, R. (2020). Multi-objective optimization of solid/liquid extraction of total sunflower proteins from cold press meal. *Food Chemistry*, *317*, 126423.

[30] Sara, A. S., Mathé, C., Basselin, M., Fournier, F., Aymes, A., Bianeis, M., ... & Kapel, R. (2020). Optimization of sunflower albumin extraction from oleaginous meal and characterization of their structure and properties. *Food Hydrocolloids*, *99*, 105335.

[31] Muniglia, L., Kiss, L. N., Fonteix, C., & Marc, I. (2004). Multicriteria optimization of a single-cell oil production. *European journal of operational research*, *153*(2), 360-369.

[32] Schwenke, K. D., Drescher, B., Zirwer, D., & Raab, B. (1988). Structural Studies on the Native and Chemically Modified Low-Molecular Mass Basic Storage Protein (Napin) from Rapeseed (*Brassica napus L.*). *Biochemie und Physiologie der Pflanzen*, *183*(2-3), 219-224.

[33] Perera, S., McIntosh, T., & Wanasundara, J. (2016). Structural properties of cruciferin and napin of Brassica napus (canola) show distinct responses to changes in pH and temperature. *Plants*, *5*(3), 36.

[34] Krzyzaniak, A., Burova, T., Haertlé, T., & Barciszewski, J. (1998). The structure and properties of Napin-seed storage protein from rape (*Brassica napus L.*). *Food/Nahrung*, *42*(03-04), 201-204.

[35] Nioi, C., Kapel, R., Rondags, E., & Marc, I. (2012). Selective extraction, structural characterisation and antifungal activity assessment of napins from an industrial rapeseed meal. *Food chemistry*, *134*(4), 2149-2155.

[36] Bodin, A., Framboisier, X., Alonso, D., Marc, I., & Kapel, R. (2015). Size-exclusion HPLC as a sensitive and calibrationless method for complex peptide mixtures quantification. *Journal of Chromatography B*, *1006*, 71-79.

[37] Rodrigues, I. M., Coelho, J. F., & Carvalho, M. G. V. (2012). Isolation and valorisation of vegetable proteins from oilseed plants: Methods, limitations and potential. *Journal of Food Engineering*, *109*(3), 337-346.

[38] Kurozawa, L. E., Park, K. J., & Hubinger, M. D. (2008). Optimization of the enzymatic hydrolysis of chicken meat using response surface methodology. *Journal of Food Science*, *73*(5), C405-C412.

[39] Kapel, R., Klingenberg, F., Framboisier, X., Dhulster, P., & Marc, I. (2011). An original use of size exclusion-HPLC for predicting the performances of batch ultrafiltration implemented to enrich a complex protein hydrolysate in a targeted bioactive peptide. *Journal of membrane science*, *383*(1-2), 26-34.

Table captions:

Table 1: Domain of the operating conditions for the hydrolysis of rapeseed albumins with Alcalase 2.4L, Flavourzyme and Prolyve

Table 2: Experimental values and level distribution of the variables used in the hydrolysis of rapeseed albumins with Alcalase 2.4L

Table 3: Regression coefficients of polynomial models for Xp_{max} and DH_{max} parameters

Table 4: Regression coefficients of predicted quadratic polynomial models for k_{Xp} and k_{DH} and ANOVA analysis results

Table 5: Predicted and experimental values of the kinetic parameters of Xp (k_{Xp}) and DH (k_{DH}) for the validation points of the models

Table 6: Domains of the operating conditions values applied for the hydrolysis of rapeseed albumins with Flavourzyme and Prolyve

Figures captions:

Figure 1: Scheme of the proposed methodology for models establishment (A) to simulate the kinetics of the protein conversion rate and the degree of hydrolysis as a function of the 3 operating conditions (pH, T and E/S) and to search for the best duration/ enzymatic cost trade-offs (B) for the enzymatic hydrolysis of proteins

Figure 2: Identification of enzymatic mechanism for the hydrolysis of rapeseed albumins with Alcalase 2.4L in the chosen operating conditions area by plotting Xp values (A) and the mean number of amino acids by peptide, Naa (B) versus DH values

Figure 3: Enzymatic hydrolysis of rapeseed albumins with Alcalase 2.4L over 24 h at 60 °C, E/S: 1/15 (g/g) and pH 7.0 (blue curves); 8.5 (red curves) and 10.0 (green curves). Kinetics of Xp (dotted lines) and DH (solid lines) are exposed (A) and correlations between the maximum Xp values (blue), the maximum DH values (red) and the pH condition (B)

Figure 4: Experimental values of Xp (green points) and DH (blue points) versus calculated values of Xp and DH, respectively

Figure 5: Response surface plots showing effects of temperature, pH, and E/S ratio (fixed values; g/g) on the enzymatic hydrolysis constants of Xp (A) and DH (B) for hydrolysis of rapeseed albumins with Alcalase 2.4L

Figure 6: Experimental kinetics (points) and predicted intervals of kinetics (dotted lines) of Xp (A) and DH (B) for the validation points (A in blue and B in green) and correlations between experimental and predicted values of Xp (C) and DH (D) for point A (circles) and B (cross)

Figure 7: Identification of enzymatic mechanism for the hydrolysis of rapeseed albumins with Prolyve (blue) and Flavourzyme (green) in the applied operating conditions area by plotting Xp values (A) and the mean number of amino acids by peptide, Naa (B) versus DH values

Figure 8: Experimental kinetics (points) and predicted intervals of kinetics (dotted lines) of Xp (A) and DH (B) for the validation points (hydrolysis of rapeseed albumins with Prolyve in blue and with Flavourzyme in green) and correlations between experimental and predicted values of Xp (C) and DH (D) for hydrolysis of rapeseed albumins with Prolyve (circles) and with Flavourzyme (cross)

Figure 9: Identification of Pareto's front for Xp values (A) of 25% (blue); 50% (red) and 75% (green) and for DH values (B) of 3% (blue); 5% (red); 8% (green) and 10% (purple) in the case of the applied study

Figure 10: Identification of Pareto's domain for Xp value of 50% (A) and for DH value of 10% (B) in the case of the applied study

Table 1: Domain of the operating conditions for the hydrolysis of rapeseed albumins with Alcalase 2.4L, Flavourzyme and Prolyve

Enzyme	**pH**		**T** *(°C)*		**E/S ratio** (*g enzyme /g substrate*)	
	Minimum	*Maximum*	*Minimum*	*Maximum*	*Minimum*	*Maximum*
Alcalase	7.0	10.0	45	75	1/150	1/15
Flavourzyme	5.0	8.5	40	60	1/150	1/15
Prolyve	2.5	5.5	45	65	1/150	1/15

Table 2: Experimental values and level distribution of the variables used in the hydrolysis of rapeseed albumins with Alcalase 2.4L

Factors	**Levels**	**Codified values**		
		-1	0	1
pH (X_1)	3	7.0	8.5	10.0
T (X_2, °C)	3	45	60	75
E/S ratio (X_3, g/g)	3	1/150	1/27	1/15

Table 3: Regression coefficients of polynomial models for Xp_{max} and DH_{max} parameters

Variables	**Xp_{max}**	**DH_{max}**
Transformation	None	None
Coefficients		
Intercept (b_0)	-0.181	-0.103
Linear (b_i)		
pH	0.310	0.0986
Quadratic (b_{ii})		
pH pH*	-0.0203	-0.00731

Table 4: Regression coefficients of predicted quadratic polynomial models for k_{Xp} and k_{DH} and ANOVA analysis results

Variables	**Kinetic constant k_{Xp}**	**Kinetic constant k_{DH}**
Transformation	Logarithmic	Logarithmic
Coefficients		
Intercept (b_0)	-2.410	-1.297
Linear (b_i)		
pH	-1.541	-1.683
T	0.2239	0.2480
E/S	56.20	13.65
Quadratic (b_{ii})		
pH pH*	0.1060	0.1125
T T*	$-1.782E^{-03}$	$-2.159E^{-03}$
E/S E/S*	-217.7	-207.4
Interaction (b_{ij})		
pH E/S*	-3.175	-
T E/S*	-	0.2465
Analysis of variance		
R^2	0.946	0.854
R^2 adj	0.927	0.801
Q^2	0.888	0.719
RSD	$9.96E^{-02}$	0.173
Regression (p value)	$2.77E^{-11}$	$10.0E^{-07}$
Lack of fit (p value)	0.119	0.448

Table 5: Predicted and experimental values of the kinetic parameters of Xp (k_{Xp}) and DH (k_{DH}) for the validation points of the models

	Variables				Model prediction intervals		Experimental values	
Validation point	pH	T	E/S	S_0 (%, w/v)	$\mathbf{k_{Xp}}$	$\mathbf{k_{DH}}$	$\mathbf{k_{Xp}}$	$\mathbf{k_{DH}}$
A	7.5	55	1/50	1	[0.207 – 0.406]	[0.780 – 1.13]	0.391	1.06
B	8	50	1/20	5	[0.680 – 0.879]	[1.91 – 2.26]	0.832	2.41

Table 6: Domains of the operating conditions values applied for the hydrolysis of rapeseed albumins with Flavourzyme and Prolyve

Enzyme	**Flavourzyme**		**Prolyve**	
	Experimental domain	*Validation point*	*Experimental domain*	*Validation point*
pH	[5.0]	5.0	[4.0]	4.0
T *(°C)*	[40 °C – 60 °C]	45	[45 °C – 65 °C]	60
E/S ratio (*g enzyme/g substrate*)	[1/15 - 1/150]	1/20	[1/15 - 1/150]	1/20

Figure 1: Scheme of the proposed methodology for models establishment (A) to simulate the kinetics of the protein conversion rate and the degree of hydrolysis as a function of the 3 operating conditions (pH, T and E/S) and to search for the best duration/ enzymatic cost trade-offs (B) for the enzymatic hydrolysis of proteins

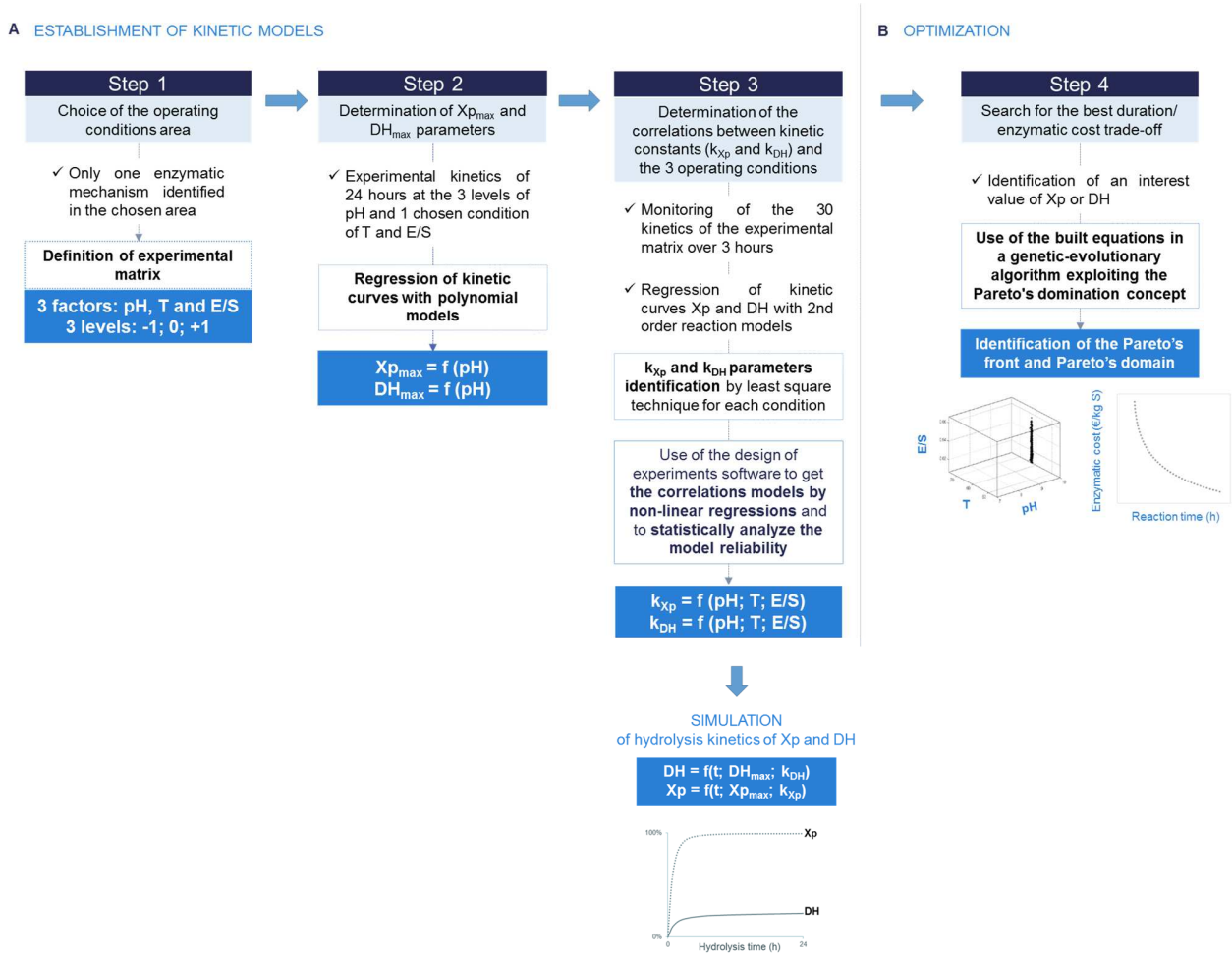


Figure 2: Identification of enzymatic mechanism for the hydrolysis of rapeseed albumins with Alcalase 2.4L in the chosen operating conditions area by plotting Xp values (A) and the mean number of amino acids by peptide, Naa (B) versus DH values

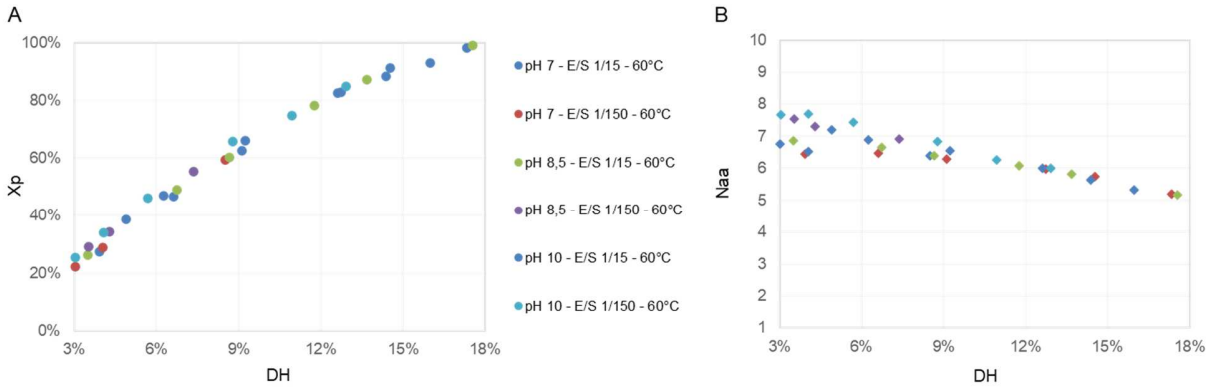


Figure 3: Enzymatic hydrolysis of rapeseed albumins with Alcalase 2.4L over 24 h at 60 °C, E/S: 1/15 (g/g) and pH 7.0 (blue curves); 8.5 (red curves) and 10.0 (green curves). Kinetics of Xp (dotted lines) and DH (solid lines) are exposed (A) and correlations between the maximum Xp values (blue), the maximum DH values (red) and the pH condition (B)

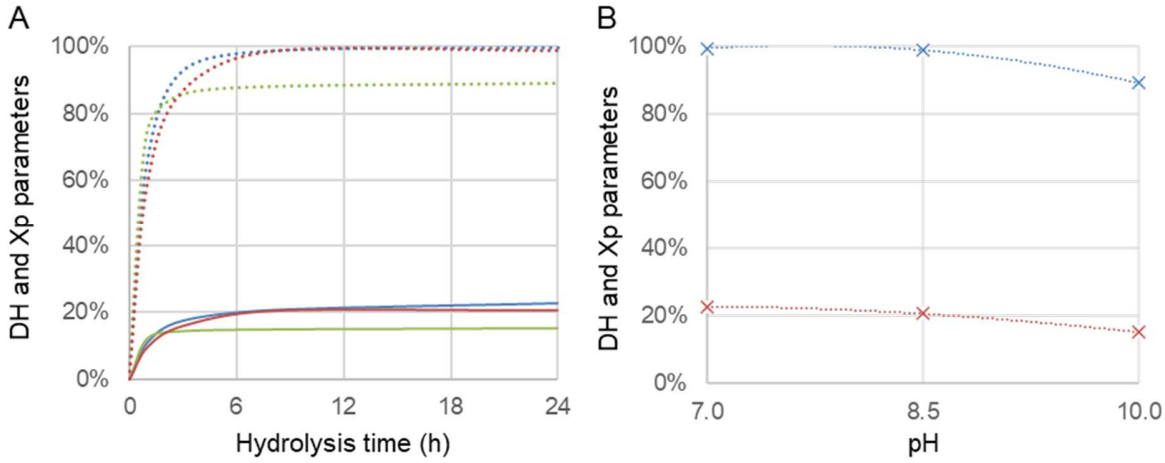


Figure 4: Experimental values of Xp (green points) and DH (blue points) versus calculated values of Xp and DH, respectively

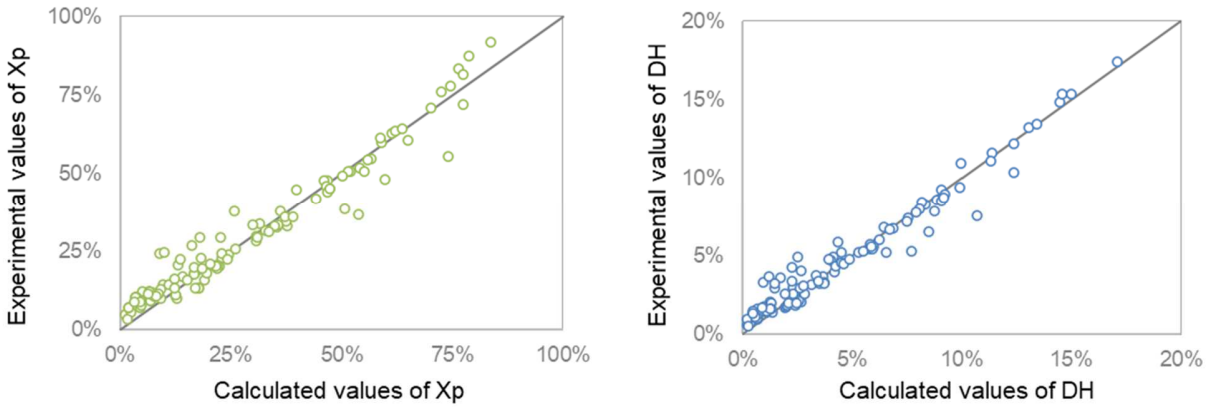


Figure 5: Response surface plots showing effects of temperature, pH, and E/S ratio (fixed values; g/g) on the enzymatic hydrolysis constants of Xp (A) and DH (B) for hydrolysis of rapeseed albumins with Alcalase 2.4L

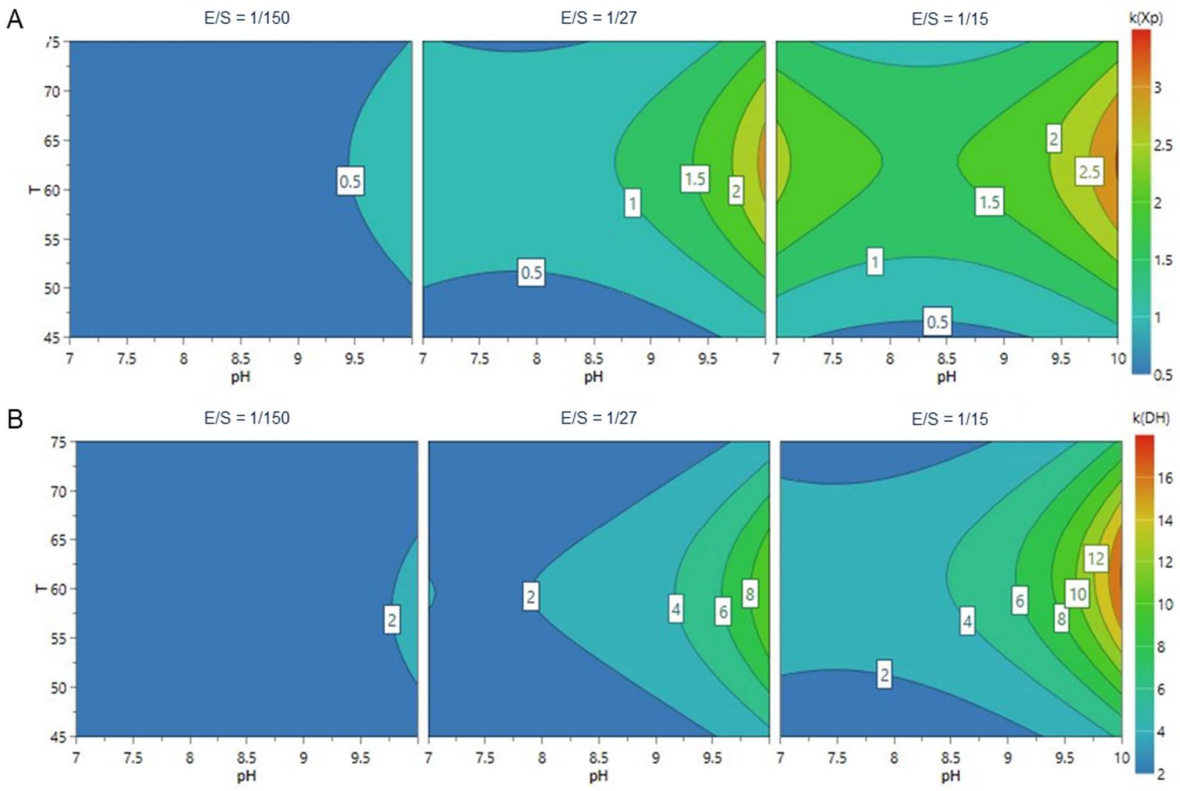


Figure 6: Experimental kinetics (points) and predicted intervals of kinetics (dotted lines) of Xp (A) and DH (B) for the validation points (A in blue and B in green) and correlations between experimental and predicted values of Xp (C) and DH (D) for point A (circles) and B (cross)

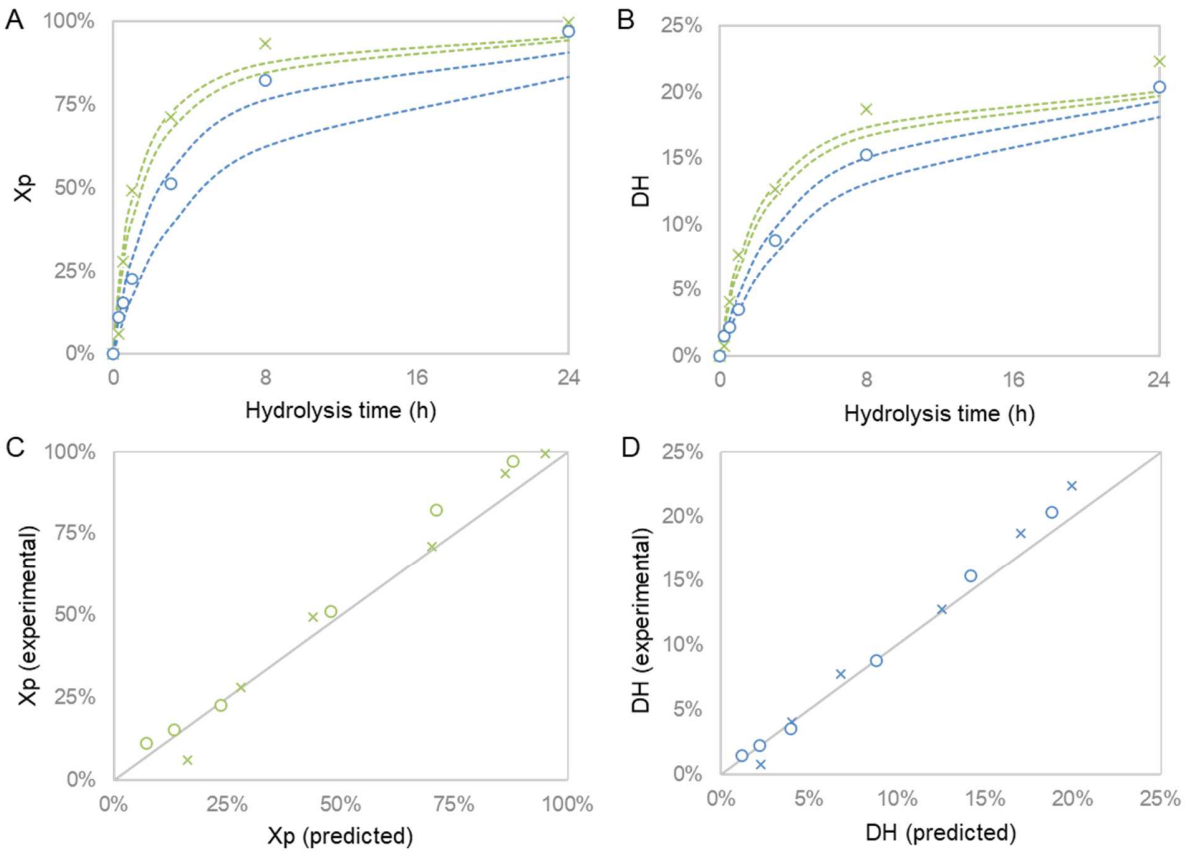


Figure 7: Identification of enzymatic mechanism for the hydrolysis of rapeseed albumins with Prolyve (blue) and Flavourzyme (green) in the applied operating conditions area by plotting Xp values (A) and the mean number of amino acids by peptide, Naa (B) versus DH values

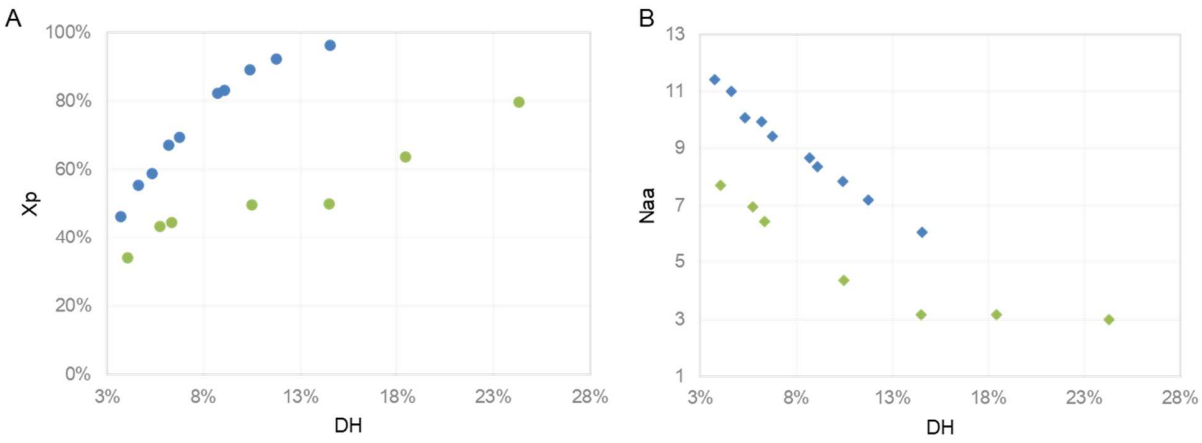


Figure 8: Experimental kinetics (points) and predicted intervals of kinetics (dotted lines) of Xp (A) and DH (B) for the validation points (hydrolysis of rapeseed albumins with Prolyve in blue and with Flavourzyme in green) and correlations between experimental and predicted values of Xp (C) and DH (D) for hydrolysis of rapeseed albumins with Prolyve (circles) and with Flavourzyme (cross)

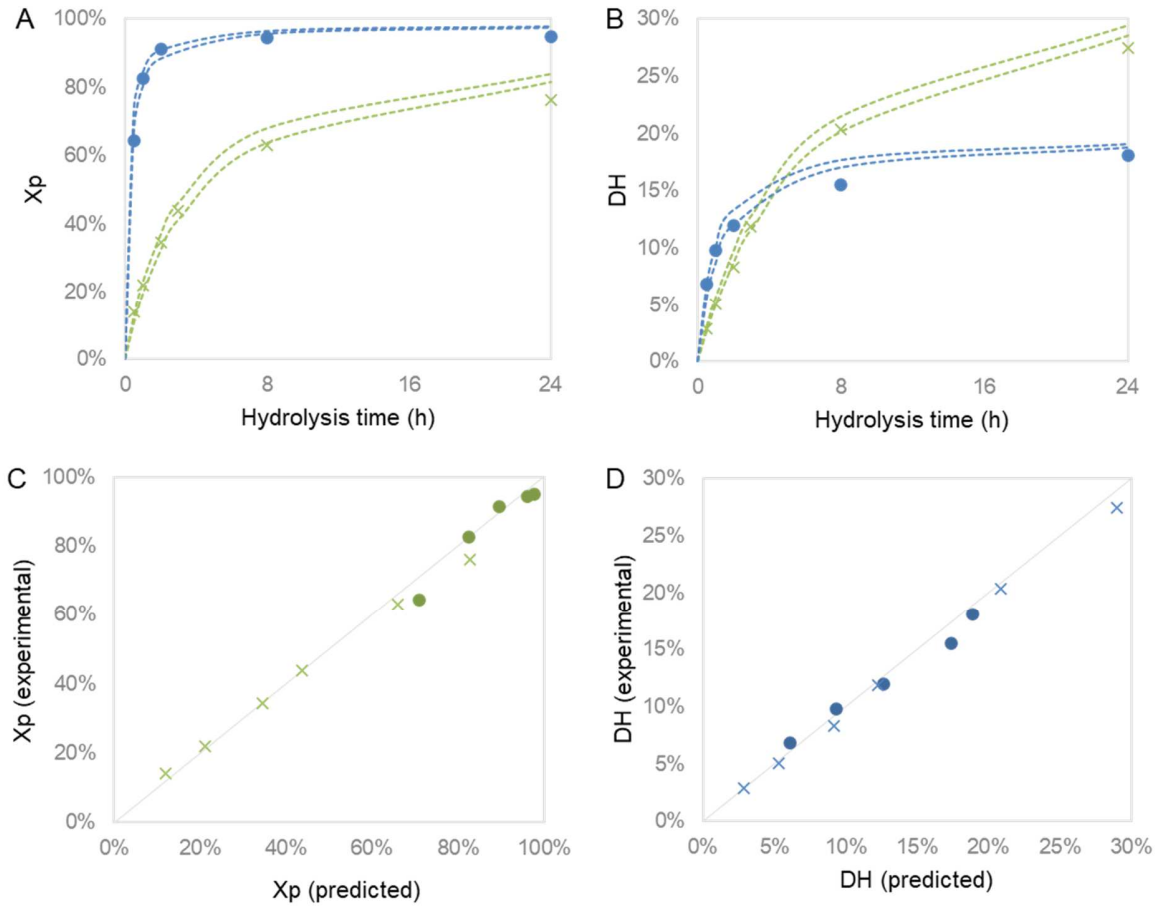


Figure 9: Identification of Pareto's front for Xp values (A) of 25% (blue); 50% (red) and 75% (green) and for DH values (B) of 3% (blue); 5% (red); 8% (green) and 10% (purple) in the case of the applied study

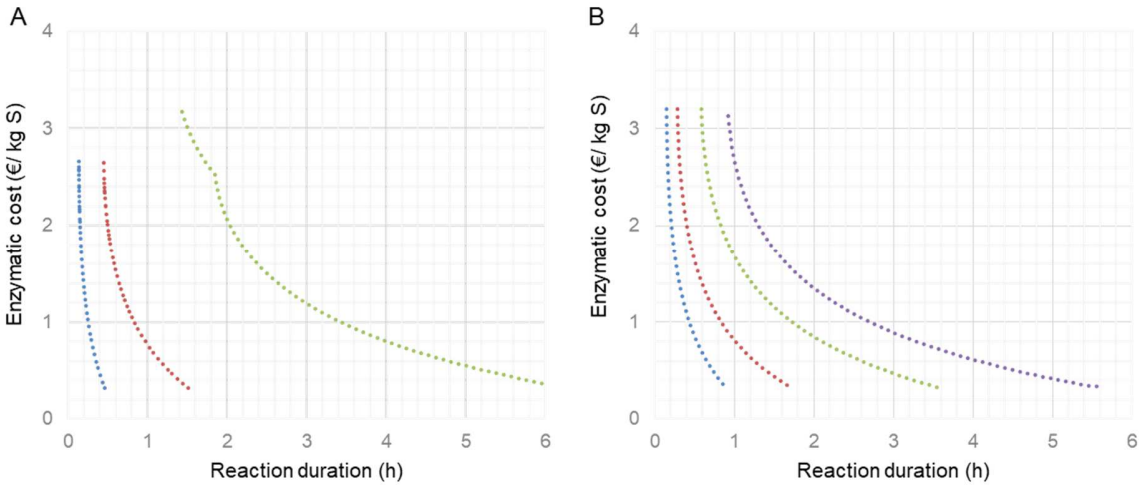
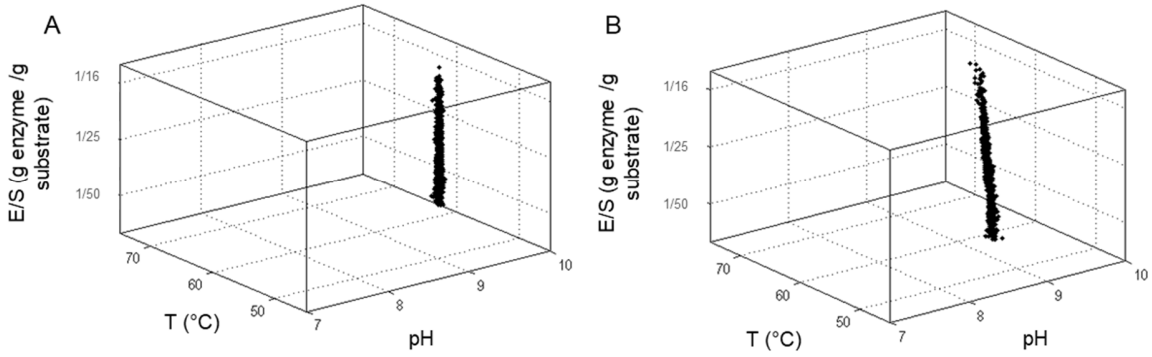


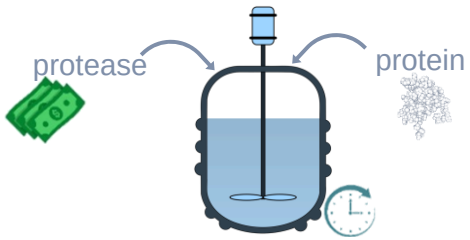
Figure 10: Identification of Pareto's domain for Xp value of 50% (A) and for DH value of 10% (B) in the case of the applied study



Identification of a **hydrolysate of interest** DH_x or Xp_x (linked to one property)

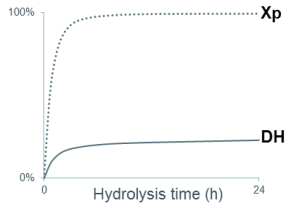
Operating conditions

pH, T, E/S



Kinetic modeling

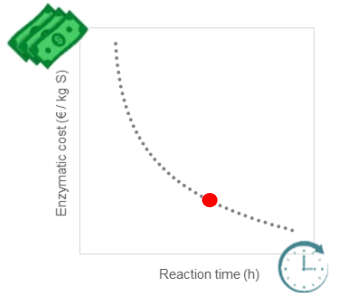
DH = f(t; pH; T; E/S)
Xp = f(t; pH; T; E/S)



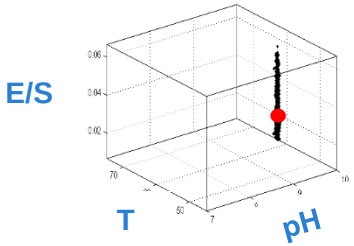
SIMULATION OF HYDROLYSIS KINETICS

Genetic-evolutionary algorithm

Enzymatic cost = f(reaction duration)



Identification of **optimal operating conditions** to produce the target hydrolysate at DH_x or Xp_x



OPTIMIZATION