Applying Boltzmann Equation to Starch Enzymatic Hydrolysis Modeling

A.PATSIOURA, E.REPPAS, G.MANIATIS and V.GEKAS
Department of Environmental Engineering,
Technical University of Crete,
Polytechnioupolis, Chania 73100, Crete,
GREECE

vassilis.gekas@enveng.tuc.gr annapatsioura@yahoo.gr
gmaniatis88@gmail.com lastexit@windowslive.com

Abstract: - In certain enzymatic actions, such as the case of α-amylase hydrolyzing the starch, the mode of splitting the molecule is random. Various models, mostly non-deterministic, have been applied to account for this type of action, as for example the Monte-Carlo procedure. In this paper a new approach is applied based on the Boltzmann entropy equation where the multiplicity, W, has the meaning of various possible products after a given number of hydrolysis stages. Through this analysis the probability of obtaining the desirable product, which is maltose, is estimated and compared to experimental results.

Key-words: - Enzymes, enzymatic hydrolysis, starch, α-amylase, maltose, Boltzmann equation, enzymatic modeling

1 Introduction

In Herakleion, July 2008, we have presented a paper on the application of the Boltzmann equation in various branches of Physics [1] the S being the entropy, information, sensation or even time with a proper definition of the multiplicity, W, in each case. In Cambridge, Feb. 2008 we have reported the experimental results of starch hydrolysis in an effort to study the kinetics of the hydrolysis [2]. In our laboratory we have faced the challenge of modeling an enzymatic system acting in a random mode due to the application of α-amylase among the enzymes. Instead of following a Monte Carlo approach, as it was the case in previous studies [3-6], and after we solved the problem by applying simple combinational analysis, we have applied the Boltzmann type of equation where W, is the multiplicity, the number of possible product mixtures after a given stage of hydrolysis, and S, is the size of the substrate molecule. This relationship is shown to be a logarithmic one as it is expected from Thermodynamics. Then, from the multiplicity we estimate the probability, P_{malt}, of having the desired product, maltose, in the final product. At last we compare this probability to experimental results obtained in our laboratory.

2 Starch Hydrolysis

One way of healthy non-dairy milk production is by enzymatic hydrolysis of oat starch and this product as well as other similar products is commercially now available in the international market. The system of enzymes used consists mainly of α-amylase and β-amylase, so that the oat milk contains maltose instead of lactose, which is contained in dairy milk. It is known that lactose causes intolerance effects to some individuals and even in populations. Many people around the world do not have the ability to produce the enzyme lactase and therefore cannot properly digest the milk sugar lactose found in dairy milk.

Another health aspect of the oat milk is the low lipid concentration and the ability of decreasing the low density cholesterol in the organism. The latter happens because oats are rich in water-soluble fibers such as β-glucans. Foods that contain soluble dietary fibers have been shown to decrease the plasma cholesterol concentration, probably by reducing the absorption of cholesterol and bile acids.

Both α- and β-amylase during the oat starch hydrolysis are of special importance as they are responsible for the solubilization of starch. α-amylase is an endo-amylase and catalyzes the hydrolysis in a random manner in the interior of the starch molecule. It catalyzes the hydrolysis of internal α-1,4-glucosidic linkages and is thus ideally suited to cutting a starch polymer into smaller fragments, linear and branched oligosaccharide of various chain lengths. Both bacterial and fungal α-amylases have found widespread use in industrial processes because of their high thermo-stability, but their efficient expression systems have also made them attractive for such purposes.

A literature review [7] has shown that the studies carried on until now have led either to empirical equations directly from multi-factorial experimental designs or to fundamental even kinetic models but without coupling terms in the equations that should correspond to the observed interaction effects. This paper covers the case of the viscosity thinning enzyme α-amylase, being the first step in starch hydrolysis.

3 Development of the new model, the symmetric case

The symmetric case means that given a molecule or an amylose chain of a given size, the α-amylase does not make distinction between the two ends, i.e. the reducing end and the non reducing end. The main presupposition is that starch is mainly composed of amylose because the amylopectin has already been debranched by a suitable enzyme, i.e. the pullulanase. Starting from a given initial size n, the α-amylase cuts the molecule in the first stage of hydrolysis at the positions 3 to n-2 (which means 3 non-productive attacks; 2 bonds adjacent to the non-reducing end and 1 bond adjacent to the reducing end). As an example a molecule with an initial size of ten glucose units will be cleaved in a way shown by Fig.1. From a mathematical viewpoint, the combination of the products after each step of hydrolysis is presented as a sum, thus helping us considering it better as a matter of combinational analysis.

The model development implies the study of the relationship connecting the multiplicity, W, with molecule size, S. Also, the relationship between the multiplicity, W, and the maltose probability, P_{malt}, is given.

It has to be referred that the model is based on the following assumptions:

1. The substrate is a linear molecule.
2. Through a linear molecule, each adjacent bond to the reducing and non-reducing end, respectively, cannot be cleaved. For decamer, the bonds at the positions 1-2 and 9-10 are non-productive attacks.
3. The enzyme cannot cleave a molecule smaller than a pentamer.
4. In decamer, the bond at the position 5-5 is counted only once.
5. For the calculations in the model there is no limitation for the number of glucose units (S) in the molecule; nevertheless, according to the MW of amylose or starch, the physical limitation emerges.
6. In longer molecules, after each step of hydrolysis, the products may be further hydrolyzable, more than one with the same probability. In that case all hydrolyzable molecules have to account for the model. For the 13-mer for example, the enzyme may cleave at the position 7-6, which means that both products can be hydrolyzed with the same probability.
7. During the third step of hydrolysis, or even later, when the same product is appeared twice, it is count only once. For example for 12-mer (12→7+5→5+2+5), in the third step of hydrolysis, the pentamer will be further hydrolyzed only once.

The first relationship, after the first stage of hydrolysis is shown to be:

$$W_1 = S-3 \qquad (1)$$

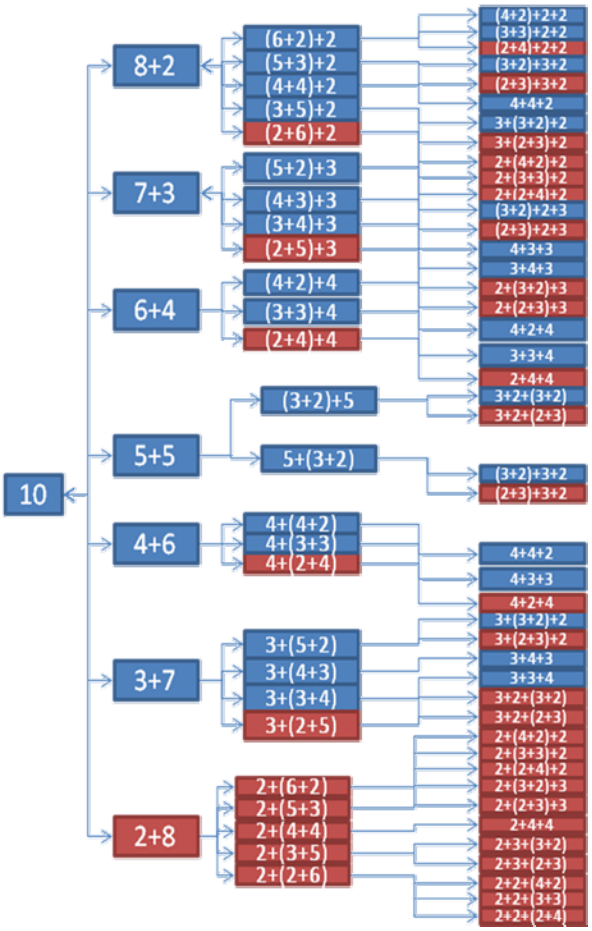


Figure 1. A schematic diagram of a three-stage hydrolysis of a decamer and the possible combinations of products

In the figure below, the symmetric model is represented in both red and blue boxes, while the real assymetric model is shown only with the blue ones.

After the 2nd stage of hydrolysis the relationship becomes:

for S=even

w2

$$w2 = 2\sum_{k=5}^{k=s-2} w1(k) - w1\left(\frac{s}{2}\right) \quad (2)$$

for S=odd

$$w2 = 2\sum_{k=5}^{k=s-2} w1(k) \quad (3)$$

Then it is shown that for higher stages of hydrolysis the multiplicity is a function of the multiplicities of the lower stages. As an example for a molecule of S=13:

$$W_{3(13)}=2[W_{2(11)}+W_{2(10)}+W_{2(9)}+2W_{1(5)}W_{1(8)}+2W_{1(6)}W_{1(7)} +2W_{1(6)}+2W_{1(5)}] \quad (4)$$

Based on the sixth assumption it should be noted that the resulting products could be interpreted also analytically; during the hydrolysis of the 13-mer e.g., when 8+5 come as products at the first step of hydrolysis, the resulting combination of products during the third step of hydrolysis is going to be: $W_{1(8)}$ combinations of products for the 8-mer and $W_{1(5)}$ combinations of products for the 5-mer. Each of the $W_{1(8)}$ combinations comes with a pentamer which has not been hydrolyzed (respectively the $W_{1(5)}$ combination comes with unhydrolyzed octamer). Consequently, during the third step of hydrolysis, $W_{1(8)}*W_{1(5)}$ combinations of products will come from the octamer, while $W_{1(5)}*W_{1(8)}$ combinations from the pentamer. In total, $2*W_{1(5)}*W_{1(8)}$ will come from the hydrolysis of 8+5.

To show it better we demonstrate this principle in Fig.2.

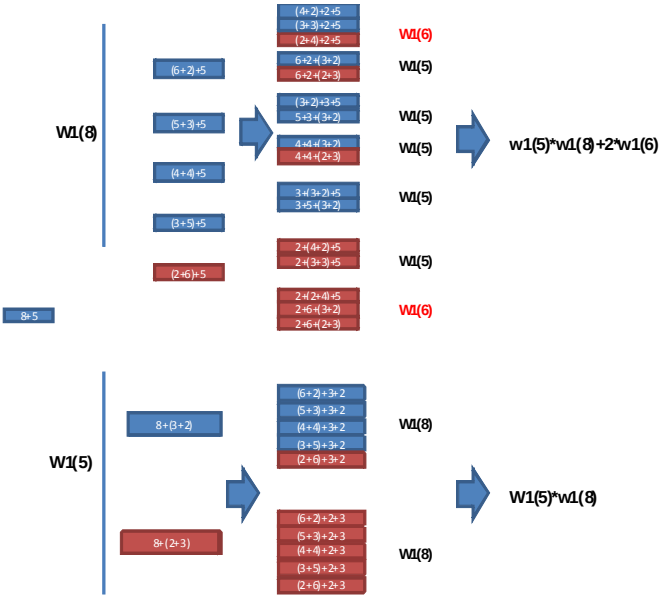


Figure 2. Schematic diagram of the hydrolysis of a 13-mer. The first stage gives 8+5, the second stage gives $W_{1(8)}$ +$W_{1(5)}$ and the third stage gives the expression in the right hand of equation 4.

From the aforementioned equations that have been developed, it can be concluded that the enzyme, α-amylase represents a non-random number generator. In other words, the molecule represents an array of non-random numbers that edits and re-associates them logically, in a non-random manner. In the case of the decamer, the enzyme's action could be described by the following steps:

i. The molecule includes an array of numbers, from 2 until 8.
ii. "Gets" one by one the terms of the array subtracting them from the size of the molecule, so that a new array (descending array this time) will arise, from 8 until 2.
iii. Re-associates the terms of these two arrays one by one, in a way that, the combinations of the products of the first step of hydrolysis will come up, e.g. (2+8), (3+7), (4+6), etc.

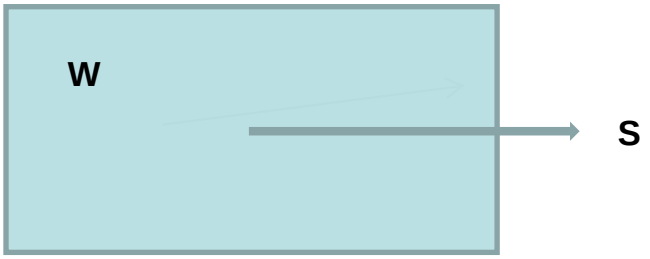
Because the analytical expressions become more and more complicated the higher the stage number of hydrolysis is reached, we have proceeded algorithmically having developed a suitable computer software. This is a sample of a program that calculates multiplicity for molecules with odd size.

```
  w1=0
  w3=0
  w2=0
  s=13
  for i=2:(s-2)
  w1=w1+1
  a=[i (s-i)]
  if (i>4)
  for j= 2:(i-2)
  b=[j (i-j) (s-i)]
  w2=w2+1
  if (j>4)
  for k=2:(j-2)
  s=[k (i-j) (s-i) (j-k)]
  w3=w3+1
  end
  if ((j-k)>4)
for k=2:(j-k-2)
z=(i j (s-i-j) (s-i-j-k)
w3=w3+1
end
end
if ((s-i)>4)
for j=2:(s-i-2)
x=[i j (s-i-j)]
w2=w2+1
if (s-i-j)>4
for o=2:(s-i-j-2)
w3=w3+1
end
if (j>4)
w3=w3+1
end
end
end
```

4 A Generalized Boltzmann Approach

4.1 The origin

We have already shown that the Boltzmann equation, although of a strict thermodynamic origin can be extended in a generalized manner in all those situations where a logarithmic relationship exists between an extensive property, whose the value could be estimated by microstates or by combination theory and an intensive property, S.



In the above schematic representation, the meaning of W is:

- Multiplicity
- Microscopic
- Accumulative
- Internal
- Structural details
- An input
- «Γίγνεσθαι»

whereas the meaning of S is:

- A trend, a tendency
- Macroscopic
- Appearing
- Qualitative
- Subjective
- Output
- «Φαίνεσθαι»

Time could be either in the place of W or in the case of S as we have shown in Herakleion last year

- Time as W, the objective time
- Time as S, the feeling of the duration of the time, the subjective time, the "temps vaicu"

This means that if applying the Boltzmann equation with the two notions of time, as an objective or astronomical concept and as feeling of its duration we shall understand the need of a modification of the simple S=klogW equation. Time as Chronus, or as Kronos, the God of Chronus or Time (ATH TE AMA EAP) Decay but also Growth: Time is of an ambivalent nature.

4.2. S as Size

- Products of Hydrolysis in the case of a random enzymatic action

- Subunit sequences in biopolymer products and connection with their mechanical properties

The kind of logarithm could be

The decadic log10 ,

The natural or neperian loge, or ln

The binary log2

The natural or neperian logarithm expresses better the ideas of this paper and our efforts to understanding the Boltzmann equation in connection of its various applications. However as Plato puts it, the Rational coexists with the Irrational everywhere in the Universe or Cosmos.

As far as the constant k is concerned:

- Its meaning depends on the pair of S-W used
- Its dimensions is consequently the matter of a dimensional analysis depending also on the kind of logarithm used
- The Boltzmann constant in the S as Entropy
- More alternatives for S as the Information

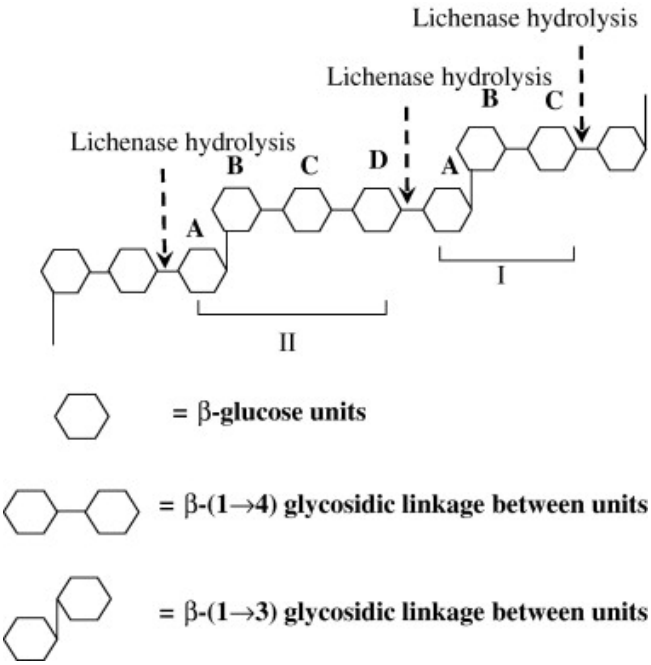
Before entering to our technological case studies it is needed to express the one part of the Boltzmann equation as p, the probability of one event instead of the multiplicity, It follows the brief description of the two technological case studies as well as a third application concerning chess.

- W is the multiplicity, p is the probability of the multiple alternatives
- W = 1/p
- Therefore if p is used instead of W the Boltzmann equation becomes:

$$S=- k \log\{p\}$$

4.3 Introduction in the beta-glucan application

The beta glucans from cereals contain mainly cellotrioses (DP3) cellotetraoses (DP4) and a small percentage of DPn, where n higher than four. Sequences of cellotrioses, for example three or higher trioses in following one another mean good mechanical properties of the beta glucans, for example bettwer elasticity. The problem is"given a composition and the size (the MW) of a certain beta glucan in DP's which is the probability of the appearance of a desired triose sequence, for example, DP33, that is three trioses in series?



Again we prefer not to apply the Monte Carlo or any other non-deterministic method, but we intend to apply our direct estimation of probabilities through the application of the universal Boltzmann equation. This problem is easier than the previous because the case is symmetric

The steps are the following

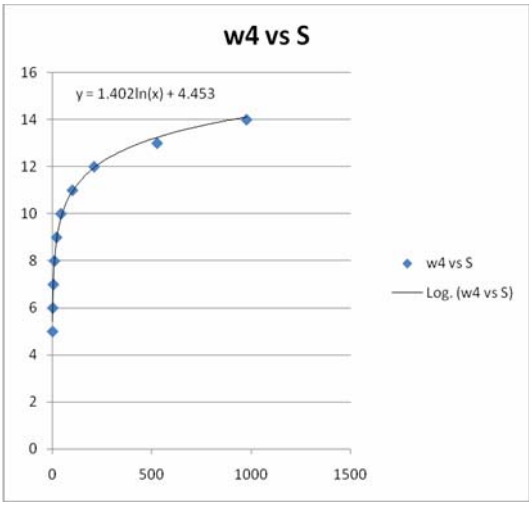
Calculate the W's for small sizes

Study the tendencies (expected to follow the Boltzmann-like relationship)

Extending for higher sizes

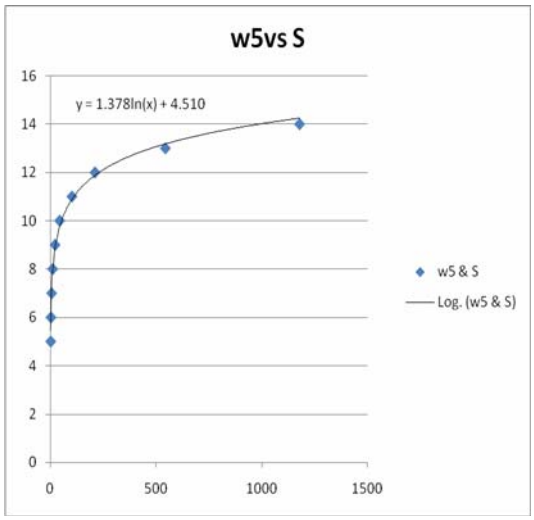
Estimate the probability of the desired sequence(s)

4.4. The application in Hydrolysis



In the case of hydrolysis the microstates are the probable states of the molecule after its enzymatic cleavage. We expect this relationship to be logarithmic as it is exhibited in Fig 3.

(a)



(b)

Figure 3. (a) W_4 vs S & (b) W_5 vs S. It is obvious that the relationship is logarithmic.

In Fig. 4 below, the linearization of the S vs W_5, i.e. the size of the amylose molecule and the 5^{th} –stage-hydrolysis multiplicity, is shown.

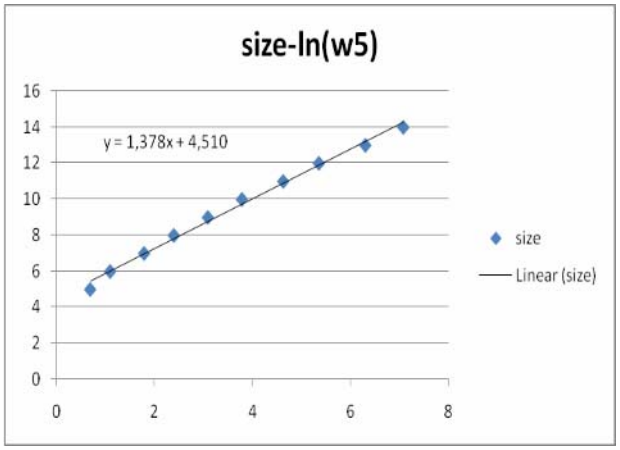


Figure 4. The linearization of the symmetric model. In this diagram the size is in a linear scale while the multiplicity is in a Neperian logarithmic scale.

5 Relationship between the ideal symmetric and the real asymmetric case

Next we calculate the multiplicity for the "real" asymmetric model, where the minimal distance from the reducing end is two glucose units and the minimal distance, respectively, three glucose units from the non-reducing end of the substrate molecule, assumed to be linear, as explained in chapter 3 above. This remark differentiates this model from the symmetric one; the assumptions are the same as for the symmetric model, apart from assumption 2 that doesn't apply to this case. In Tables 1 & 2 as well as in Figs 5-6 are given the results of the symmetric, respectively the "real" asymmetric model.

Table 1. W vs S for the symmetric model

Size, S	w1	w2	w3	w4	w5
5	2	**2**	**2**	**2**	**2**
6	3	**3**	**3**	**3**	**3**
7	4	6	6	6	6
8	5	11	11	11	11
9	6	18	22	22	22
10	7	26	44	44	44
11	8	40	94	102	102
12	9	51	151	211	211
13	10	70	276	528	544
14	11	84	398	978	1178

Table 2. W vs S for the "real" asymmetric model

Size, S	w1	w2	w3	w4	W5
5	1	**1**	**1**	**1**	**1**
6	2	**2**	**2**	**2**	**2**
7	3	3	3	3	3
8	4	5	5	5	5
9	5	9	9	9	9
10	6	16	17	17	17
11	7	25	33	33	33
12	8	34	62	63	63
13	9	49	125	141	141
14	10	61	200	279	280

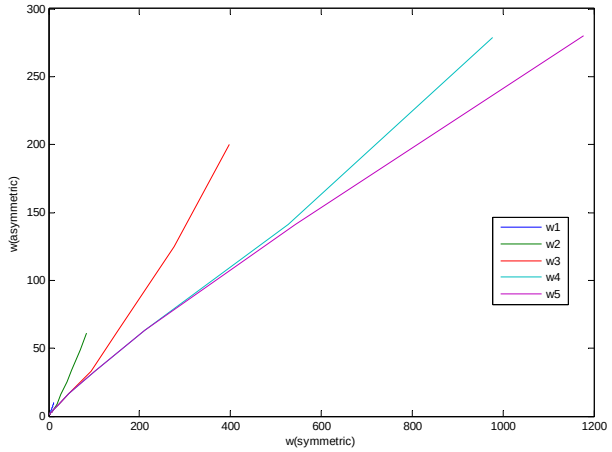


Figure 5. Comparison between the multiplicities of the two models, the symmetric and asymmetric one.

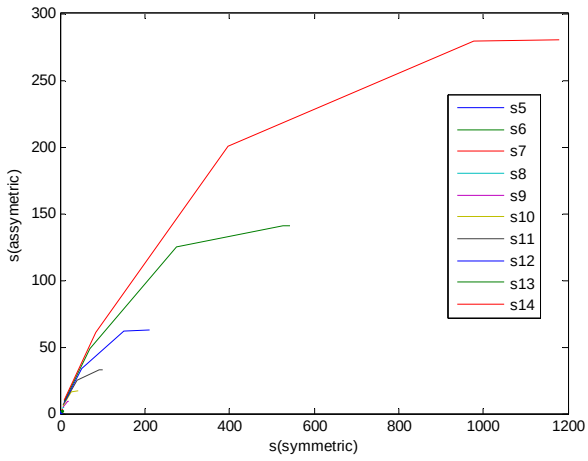


Figure 6. Comparison between the substrate sizes of the two models, the symmetric and asymmetric one.

The results above show that there is a clear converging tendency between the two models for higher degrees of hydrolysis. This enables us to apply the more handy symmetric model. In the final step of our model we have estimated the maltose probability **Table 2. W vs S for the "real" asymmetric model**

6 Applying the method directly in the asymmetric model

Using the results from the symmetric approach we applied the same theory directly on the fifth hydrolysis of the asymmetric case, in order to discover a model behavior independent from symmetry. After linearization we had this result:

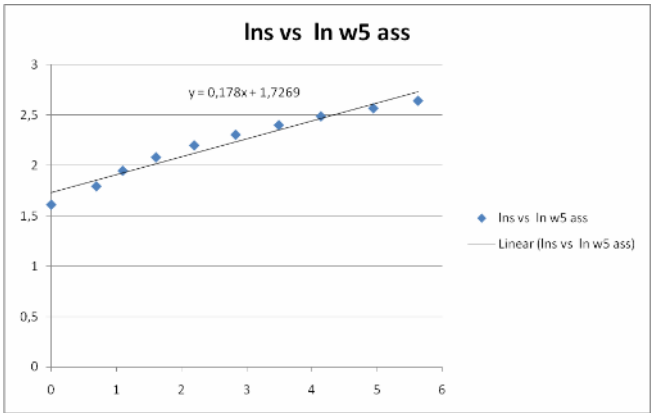


Figure 7. Linearization of the asymmetric model. In this diagram the size is in a linear scale while the multiplicity is in a Neperian logarithmic scale.

The equation of this line (Fig. 7, Y=0.178X+1.72) could be used as a general type of predicting the multiplicity after the last hydrolysis of any size, for engineering purposes, because the model is developed enough for this generalization. The behavior of the asymmetric model is exactly the same as the symmetric one, but, as we predicted, with a standard deviation. The only difference appears at the slope of the line which is about 0.16 bigger than the symmetric one.

7 Calculating the probability of maltose

Symmetric model		
Size, S	**Maltose**	**Maltose percentage**
5	2/4	50%
6	2/6	33%
7	8/16	50%
8	14/32	47%
9	36/74	49%
10	77/164	47%
11	210/458	46%
12	469/954	49%
13	1453/2654	55%
14	3111/6423	49%

Table 3a. Above, the maltose probability of the symmetric model

Table 3b Below, the maltose probability in the real case

"Real" model		
Size, S	**Maltose**	**Maltose percentage**
5	1/2	50%
6	1/2	50%
7	6/9	67%
8	9/18	50%
9	27/46	52%
10	57/111	51%
11	156/291	54%
12	371/683	54%
13	1098/2097	52%
14	2615/4990	52%

As shown in the Tables 3a-3b, maltose represents usually the 47%-55% of the final products, after the end hydrolysis. Outliers of this range of values are observed for small size starch molecules (dextrins) and do not interfere with our practical conclusions.

In order to compare the simulated with the experimental results of oat starch hydrolysis, we had to consider amylopectin as a sum of independent straight chains. This observation is geometrically represented on Fig.8 below:

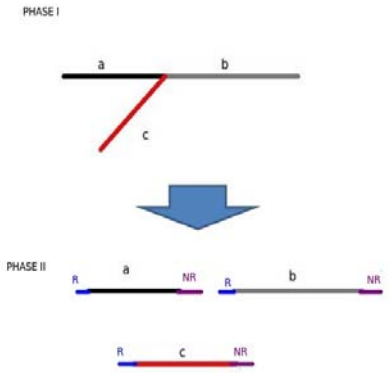


Fig 8. Simplifying amylopectin geometry as a sum of three straight amylose chain molecules, where the reducing and non-reducing ends are shown as R and NR, respectively.

When compared with experimental results [8], the predicted data showed a very good agreement. The kinetics of oat starch hydrolysis has been studied extensively in the laboratory, and the percentage of maltose in the final products was measured to be approximately 50%, which coincides with the predicted values. The agreement is better the purest the starch is from the beginning.

8. Discussion & Concluding Remarks

According to the development described above it was possible to model the aleatory action of the α-amylase enzyme in starch hydrolysis and finally to estimate the probability of maltose existence in the mixture of the final products after a given number of hydrolysis stages, in other words, in a given turnover number of the enzyme. The two models, i.e. the symmetric and the "real" asymmetric one, are found to converge when the number of hydrolysis stages exceeds a certain value. The same occurs for increasing substrate size, beyond a certain number (14). The percentage of maltose converges to the values of a range 47-54% which is proved to be independent of the size, S or the multiplicity value, W. When compared with experimental results [8], the predicted data showed a very good agreement. The kinetics of oat starch

hydrolysis has been studied extensively in the laboratory, and the percentage of maltose in the final products was measured to be approximately 50%, which coincides with the predicted values. The agreement is better the purest the starch is from the beginning.

It is noticed that the experimental results obtained in our Laboratory [8] show values of maltose percentage which fall in the same range, approximately 50%. Therefore an agreement is observed between the model and the experiment.

References:

[1] Maniatis, G., Reppas, E. and Gekas, V. (2008), In depth analogies among entropy, information and sensation. Time in Thermodynamics. *12th IASME/WSEAS Intl. Conf. Engineering Mechanics Structures, Engineering Geology,* Herakleion, 22-25 July, 08

[2] Patsioura, A. and Gekas, V.,(2008) A study of oat-origin enzymatic hydrolysis, *3d IASME WSEAS Intl. Conference on Energy and Environment*, University of Cambridge, UK, Feb 23-25 2008

[3] Wojciechowski, P. M., Koziol, A. and Noworyta, (2001) An Iteration model of starch hydrolysis by amylolytic enzymes, *Biotechnology and Bioengineering*, Vol 75, No 5, December 5, 2001, pp 530-539

[4] Rashid, R; Jamaluddin, H; Amin, (2006) Empirical and feed forward neural networks models of tapioca starch hydrolysis**.** The Lattice Monte Carlo Method For Solving Phenomenological Mass And Heat Transfer, *WSEAS Transaction on Information Science and Applications 3(3)*

[5] Ahmad Faris, Ismail (1996) 2-D And 3-D Cellular Models Using Erosion And Dilation. Vowels Recognition Using Hybrid Evolutionary Feed-forward Neural Network. Effect of Temperature on types of Cyclodextrins produced from Raw Tapioca Starch, *WSEAS Transactions on Information Science and Applications, 5 (6)*

[6] Ratniyomchai,T.,and Kulworawanichpong, (2006) Formation of galacto-oligosaccharides during lactose hydrolysis by a novel beta-galactosidase enzyme. A Monte Carlo method to data stream analysis. *Enformatika Transactions WSEAS Transactions on Circuits and Systems 5 (3): 339-345.*

[7] Patsioura, A. (2009), Enzyme Kinetics modeling as a tool to optimize industrial applications: A pragmatic approach, presented *in the Viterbo Workshop on Hydrolysis, Italym 7-9 August, 2009.*

[8] Patsioura, A., Lazaridou, A., Biliaderis, C. and Gekas, V., (2009) Enzymatic Hydrolysis of starch and Characterization of oat flour hydrolyzates. *To be published. 2009*