Modelling the Kinetics of the Enzymatic Hydrolysis of Sodium Carboxymethyl Cellulose

by

Louis E. Daigneault

A thesis
presented to the University of Ottawa
in partial fulfillment of the
requirements for the degree of
Master of Applied Science
in
Chemical Engineering

Department of Chemical Engineering
University of Ottawa
Ottawa, Ontario, 1986

ISBN 0-315-33356-1

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

As the stock of non-renewable resources dwindles, the attention of researchers in biotechnology turns to biomass. The most abundant renewable resource, cellulose, has great promise as a feedstock for processes to produce food, liquid fuel and chemicals.

A fundamental step in each process is the breakdown of the large cellulosic polymers to smaller molecules. Enzymatic hydrolysis is one of the most viable routes to achieving this but the mechanism of action of cellulases on cellulose are not clearly understood. One of the problems has been the complexity inherent in a two-phase system (solid/liquid). In this thesis, the enzymatic hydrolysis of the soluble cellulose derivative sodium carboxymethyl cellulose (CMC) was examined to overcome this problem.

Three different types of CMC were studied. The degree of polymerization, DP, had a range of 400 to 3200. The degree of substitution, DS, had a range of 0.511 to 0.947.

Here, a mechanistic model is proposed to represent the overall production rates of glucose and cellobiose from the hydrolysis of substituted cellulose derivatives. This rate is a function of the concentration of enzyme in the reactor, and DS, DP and concentration of the substrate. The main feature of the model is the realization that bonds are substrates to the enzymes.

The major findings were:

- Three enzymes contribute to the kinetics of the reaction.
- The equation describing the rate of direct action by endoglucanase, E_x, on its substrate bonds is:

$$\text{Rate of cleavage of bonds by } E_x = -\frac{k_{1E_x} \cdot E_x \cdot s}{\beta} \qquad (1)$$

where s is the concentration of bonds between two unsubstituted anhydroglucosidic units except for cellobiose and β is the total concentration of β-glucosidic linkages in the reactor.

- The equation describing the rate of direct action by cellobiase, E_1, on its substrate bonds is:

$$\text{Rate of cleavage of bonds by } E_1 = -k_{1E_1} \cdot E_1 \cdot s \tag{2}$$

where s in this case is the concentration of bonds at the non-reducing end of a polymer or oligomer of CMC provided the (non-reducing) end unit is unsubstituted.

- As has been suggested in the literature, it appears that exoglucanase releases cellobiose into solution by attacking eligible non-reducing ends of the polymer.

Furthermore, the problems associated with quantifying the concentration of any given cellulase component and the difficulties in modelling the data obtained in this enzymatic depolymerization are discussed.

ACKNOWLEDGEMENTS

This thesis was written making use of software available on the University of Ottawa Amdahl 470/V8 computer. The software processing the text of this work is TEX. In reading through this work, the effectiveness of TEX in formatting mathematical equations, tables and user-designed sketches should become evident. TEX represents the state-of-the-art in computer typesetting. It is easy to see why Donald Knuth, the system's creator, makes that claim in the TEXbook [1].

The plots found in this thesis were also made by computer. The TELLAGRAF software package is a copyright of ISSCO [2].

Both the text and plots were output on a QMS lasergrafix printer. Although both packages produce device-independent files which are then sent to the QMS printer for printing, the text and plots are output separately for each system. It is to the advantage of all involved in the production of scientific work that an interface to the two systems be designed so that text and graphics may be output together. In this way, the high quality output of TEX and TELLAGRAF would be combined into one package producing beautiful technical papers.

Several people have, directly and indirectly, contributed to the ideas leading to the development of the kinetic model describing, in part, the enzymatic hydrolysis reaction by cellulases. In the first place, this work would not have possible without the support and ideas of my thesis co-supervisors, Drs. David McLean and Gérald André. It was their determination that did get the project oriented in the right direction.

Thanks are in order to Dr. Ralph Overend of the National Research Council. During the summer of 1983, Dr. Overend impressed on me the importance of biotechnology and how it relates to chemical engineering.

The help of the technical staff of the Department of Chemical Engineering at

the University of Ottawa was very appreciated. Mr. A. Bonaldo, Mr. D. Lefebvre and Mr. J. Gasperetti deserve much credit for their fine work.

Most of all, I want to thank my family. My mother Colombe and my sister Marie, through their encouragement, made possible the production of this thesis.

Finally, I gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PROGRAMS

NOMENCLATURE

A_{ij} • event of the next substituent group binding to hydroxyl j of an anhydroglucosidic unit which is unsubstituted ($i = 0$) or substituted ($i = 1$)

bond • refers to β-1,4 glucosidic bonds

CBA • Cellobiase activity of an enzyme solution as a function of enzyme DF ..... (μmol glucose/min/mL)

$\widehat{\text{CBA}}$ • Cellobiase activity of an enzyme solution at standard conditions ..... (I.U./mL)

CMC • Sodium carboxymethyl cellulose

CMC_0 • Initial concentration of CMC in solution ..... (g/L)

c.v. • Coefficient of variation

c_0 • Fraction of anhydroglucosidic units which are unsubstituted

c_1 • Fraction of anhydroglucosidic units which are substituted once

c_2 • Fraction of anhydroglucosidic units which are substituted twice

c_3 • Fraction of anhydroglucosidic units which are substituted thrice

C_2 • Carbon 2 in the anhydroglucosidic unit

C_3 • Carbon 3 in the anhydroglucosidic unit

C_6 • Carbon 6 in the anhydroglucosidic unit

DF • Dilution factor

$\widehat{\text{DF}}$ • Dilution factor yielding standard conditions in an enzyme test

DP • Number-average degree of polymerization of a solution of glucosidic polymers. This is an intensive variable

DP_{CMC} • Number-average degree of polymerization of CMC. This is an intensive variable.

DS • Degree of substitution of CMC. DS is an intensive variable

E_1 • Cellobiase or concentration thereof ..... (I.U./L)

E_2 • Exoglucanase or concentration thereof ..... (I.U./L)

E_x • Endoglucanase or concentration thereof ..... (I.U./L)

e • Concentration of an arbitrary enzyme in a mixture

e_0 • Initial concentration of an arbitrary enzyme in a mixture

FPA • Filter paper activity of an enzyme mixture as a function of DF

..... (μmol glucose/min/mL)

$\widehat{\text{FPA}}$ • Filter paper activity of an enzyme mixture at standard conditions

..... (I.U./mL)

G_1 • Glucose or concentration thereof ..... (M)

$\hat{G}_1$ • Predicted concentration of glucose ..... (M)

$G1$ • Glucose concentration ..... (g/L)

G_2 • Cellobiose or concentration thereof ..... (M)

$\hat{G}_2$ • Predicted concentration of cellobiose ..... (M)

$G2$ • Cellobiose concentration ..... (g/L)

G_x • Cellulose or concentration thereof

HEC • Hydroxyethyl cellulose

I.U. • International Units; quantity of enzyme that produces a set amount of product under documented conditions (see experimental procedure)

k_i • Rate constant of substrate i depletion in an arbitrary enzymatic reaction

k_{1E_x} • Rate constant of endoglucanase, assuming Michaelis-Menten kinetics

..... (M/(I.U.·s))

k_{1E_1} • Rate constant of cellobiase, assuming Michaelis-Menten kinetics

..... (M/(I.U.·s))

K_i • Michaelis constant of substrate i depletion in an arbitrary enzymatic reaction ..... (M)

K_{iG_1} • Dissociation constant of a G_1–E_1 complex ..... (M)

K_{mE_2} • Michaelis constant of endoglucanase ..... (M)

K_{mE_1} • Michaelis constant of cellobiase ..... (M)

LSE • Least squares' estimate

M • Concentration units ..... (moles/litre)

n • Number of experimental measurements

N • Number of anhydroglucose units

p • Number of parameters in a model

p • Product concentration ..... (M)

s_i • Concentration of substrate i in an arbitrary enzymatic reaction (M)

S^2 • Variance, the random variable

Se_2 • Selectivity of the substituent group toward carbon#2

Se_3 • Selectivity of the substituent group toward carbon#3

Se_6 • Selectivity of the substituent group toward carbon#6

SSR • Sum of squared residuals

t • "Student's t" probability distribution function

t • Time .....(hours)

T • Temperature ..... (°Celsius)

TOL • Lower limit of component concentration before assuming nil concentration in numerical solution of model ..... (M)

Δu • Incremental number of substituent groups added to a chain of units

UF • Uranium fraction.

unit • Refers to an anhydroglucose unit whether substituted or not

v_i • Rate of reaction with respect to arbitrary substrate i ... (M/hours)

x_2 • Fraction of the anhydroglucose units that have a substituent in hydroxyl position 2

x_3 • Fraction of the anhydroglucose units that have a substituent in hydroxyl position 3

x_6 • Fraction of the anhydroglucose units that have a substituent in hydroxyl position 6

GREEK LETTERS

α • Probability level for a statistical test

β • Total concentration of β-1,4 bonds ..... (M)

β_{CMC} • Concentration of β-1,4 bonds from the CMC polymer ..... (M)

β_{nrc} • Bond between a non-reducing end unsubstituted unit and a non-end, unsubstituted unit or concentration thereof ..... (M)

β_{nrcs} • Bond between a non-reducing end unsubstituted unit and a substituted unit or concentration thereof ..... (M)

β_{rc} • Bond between a reducing end unsubstituted unit and a non-end, unsubstituted unit or concentration thereof ..... (M)

β_{rcs} • Bond between a reducing end unsubstituted unit and a substituted unit or concentration thereof ..... (M)

β_{snr} • Bond between a non-reducing side, non-end unsubstituted unit and a substituted unit or concentration thereof ..... (M)

β_{sr} • Bond between a reducing side, non-end unsubstituted unit and a substituted unit or concentration thereof ..... (M)

β_{ss} • Bond between two substituted units or concentration thereof ... (M)

β_u • Bond between two non-end unsubstituted units or concentration thereof ..... (M)

ϵ • Experimental error associated with a variable

θ • Model parameter

ν • Degrees of freedom

σ^2 • True variance of a random variable

$\hat{\sigma}^2$ • Estimate of σ^2 obtained from a sample of data

$\hat{\sigma}_p^2$ • Pooled estimate of the sample variance

ω_{a1} • Weighing factor for the glucose response based on chromatograph peak area

ω_{h1} • Weighing factor for the glucose response based on chromatograph peak height

ω_{h2} • Weighing factor for the cellobiose response

Ω • Shaker speed ..... (rpm)

Chapter 1

INTRODUCTION

1.1 Enzymatic Hydrolysis of Cellulosics: Research Objectives

The motivation for this work comes from the prospect that man's future lies in the renewable resources of the earth. Cellulose is the most abundant of these resources [3][4]. The prospects for using biomass to produce sugars as intermediates that lead to more useful chemicals may well depend on getting to understand the mechanism of the enzymatic hydrolysis of cellulose.

This thesis addresses one related problem in the vast field of cellulase research; what is the mechanism of the enzymatic hydrolysis of a substituted derivative of cellulose? The first objective of this work is to develop the theory that applies to enzymatic hydrolyses. In the course of the theory development, a mechanistic model should be obtained which, although initially applied to a specific ether of cellulose, could be expanded to mathematically describe the hydrolysis of all cellulosics i.e., esters of cellulose cellodextrins and cellulose. With the collection of data and through the use of modelling techniques, kinetic parameters inherent to a proper mechanistic model will be evaluated. Furthermore, the model should be able to predict the concentration of glucose and cellobiose as well as the degree of polymerization of the mixture as the reaction proceeds, once the kinetic parameters have been evaluated.

Our specific reaction can be looked at as depolymerization of the soluble polymer sodium carboxymethyl cellulose by a catalyst. As elsewhere, sodium carboxymethyl cellulose is henceforth called CMC. In Figure 1, some possible compositions of the repeating units of CMC are shown. At some points in this work, CMC may be referred to as 'cellulose gum' [5] with regard to its commercial name.

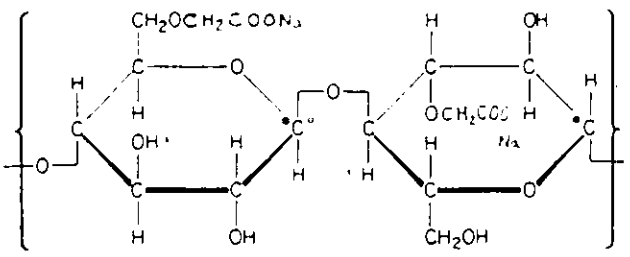


Figure 1: Typical Anhydroglucosidic Units of Carboxymethyl Cellulose

The catalyst of CMC or cellulose hydrolysis may be either acid or enzymes. In terms of practical applications, acid hydrolysis is more common. A two-stage acid process is employed in over 40 Soviet wood hydrolysis plants [4]. It has the problems of requiring acid-resistant equipment and yielding a poor grade of sugar. Cellulase enzymes obtained from fungi or bacteria do not have these problems. On a mass basis, it takes 100,000 times as much acid as enzyme to obtain a similar degree of hydrolysis [3].

The view of most researchers is that the cellulase enzyme system is composed of three enzymes: 1,4-β-D-glucan glucanohydrolase, 1,4-β-D-glucan cellobiohydrolase and β-glucosidase. These names are rather cumbersome. Each of the enzymes has a common name in widespread use in the literature amongst its several names. The list of names is tabulated elsewhere [3]. The names are descriptive of the action that each of these enzymes is thought to carry out during the hydrolysis of cellulose. In this work, the three enzymes are called endoglucanase, exoglucanase and cellobiase respectively. The text will also refer to these as E_x, E_2 and E_1 respectively. Endoglucanase is thought to act randomly on the cellulolytic polymers. Exoglucanase is thought to act on the non-reducing end of the polymer, removing 2 units hence the "exo" action. Cellobiase is not a cellulase. But it enhances cellulase action [6]. Although in dispute, cellobiase is thought to act at the non-reducing end of the polymer, removing one unit at a time. As opposed to exoglucanase, cellobiase does hydrolyze cellobiose, hence its name. In this work, references to the 'cellulase

enzyme system' addresses mixtures containing all three enzymes. In our research, these enzymes are free in solution i.e., the reaction mixture is homogeneous except for some activity tests on enzyme solutions where the substrate is insoluble filter paper.

1.2 Hydrolysis of Substituted Cellulose Derivatives

The experimental data that apply to the hydrolysis of partially substituted celluloses reflects the fact that the substituent group affects the extent of hydrolysis [7]. Under otherwise identical conditions, soluble derivatives are initially hydrolyzed more rapidly than insoluble cellulose. As the hydrolysis time increases however, the maximum reducing sugars concentration that can be reached is lower for hydrolysis of substituted derivatives of cellulose than for cellulose hydrolysis.

It will become evident when an overall hydrolysis model is developed, that the distribution of the substituent groups on the CMC must be treated in a mathematical way. Two intensive variables describe the CMC: the degree of polymerization and the degree of substitution henceforth referred to as DP and DS respectively. The DP is the average chain length of the CMC, calculated as the average number of anhydroglucose units per CMC molecule. More important to the distribution is the DS. The carboxymethyl groups substitute on any of three free hydroxyl groups on each anhydroglucose unit (with the loss of one hydrogen atom per hydroxyl group). The average number of substituted hydroxyl groups per anhydroglucose unit is known as DS. Thus the values of DS range from 0 to 3. The anhydroglucose units at either end of the polymer chain have four hydroxyl groups that may be substituted. In our discussion, this will be assumed to be three. Any difference with the physical reality will be essentially diluted by the length of the chain. Hence it is assumed that this "end effect" is negligible.

Three types of CMC were used in the hydrolysis experiments: 4H, 7L and

9M31F. All are produced by Hercules Inc. The first digit in the product designation is approximately 10 times the DS.

Table 1. *Degree of Substitution of CMC: Hercules Inc. Specifications* [5]

Type	Substitution Range	% Sodium Content
4	0.38-0.55	4.5-6.1
7	0.65-0.90	7.0-8.9
9	0.85-0.95	8.5-9.2
12	> 1.20	> 10.7

The second digit is the viscosity type. Tables 1 and 2 indicate the specifications of the manufacturer relating to our CMC substrates.

Table 2. *Degree of Polymerization of CMC: Hercules Inc. Specifications* [5]

Viscosity Type	Degree of Polymerization	Molecular Weight
H=High	3200	700000
M=Medium	1100	250000
L=Low	400	90000

The DP does not affect the distribution of the substituent groups in theory. As will be explained later, some experimental evidence indicates that an imperfect substitution reaction may alter the distribution. Ignoring this fact for a moment, a relationship must be found between the DS and the fraction of the CMC anhydroglucose units which are unsubstituted.

A DS of 0.7 means that there are an average of 7 substituent groups per ten anhydroglucose units. But this does not mean that 30% of the units are unsubstituted. Some of the anhydroglucose units may be substituted more than once. An equation relating the DS and the "fraction of unsubstituted units" is developed in the Theory chapter.

1.3 Schematic Diagrams of Anhydroglucosidic Units

Before proceeding with a survey of the literature, standard diagrams are now introduced. Diagrams will be useful in illustrating some mechanisms that apply to specific sequences of some fragments of the polymer chain.

Unsubstituted glucose molecule

Substituted glucose molecule

Unsubstituted reducing-end unit

Unsubstituted non-reducing-end unit

Unsubstituted mid-molecule unit

Substituted reducing-end unit

Substituted non-reducing-end unit

Substituted mid-molecule unit

The units illustrated above are distinguished by their state of substitution, their orientation and their position on the polymer chain. A substituted anhydroglucose unit is drawn larger than the substituted unit to reflect the significant increase in size due to the carboxymethyl group. No distinction is made between units which are substituted once, twice or thrice.

Each unit is drawn asymmetrically. One of the ends in the polymer is deemed reducing while the other end is non-reducing. In our diagrams, the half of a unit nearer to the reducing end is to the right and marked with an '**R**'.

Chapter 2

LITERATURE SURVEY

Here are some aspects of the literature to review for their importance to modelling hydrolysis of cellulosics:

- the multiplicity and mechanisms of the cellulase enzyme system
- the optimal reaction conditions
- the enzyme activity assay procedures
- the choice of a substrate
- previous attempts at modelling the hydrolysis of cellulose or its derivatives

2.1 Hydrolysis Reaction Kinetics: Enzyme Multiplicity

Wood and McCrae [8] enumerate the principle enzyme components involved in the hydrolysis of cellulose.

> "As currently understood, cell-free enzyme preparations that can solubilize native cellulose contain at least two types of enzymes, so-called E_2 and E_x. A third enzyme, a β-glucosidase or cellobiase, is normally, but not always, present as an extracellular enzyme." [8]

This quotation was singled out because, in the derivation of an enzymatic hydrolysis model, it is necessary to clearly identify all cellulase components that act on cellulose and its derivatives. And our starting point is that there are basically three important enzymes.

Yet Wood and McCrae found two separate exoglucanases for *T. koningii* and four exoglucanases from *F. Solani.* Five endoglucanases were isolated from *T. koningii.* Those five CMCases are reported to have different randomness in attacking cellulose. To explain this multiplicity, Wood and McCrae suggest, from Rantela's work, that the cellulase producing organism can alter the active site of E_x depending on its carbon source.

Endoglucanase preparations usually contain from 3 to 5 enzyme components differing in molecular weight and isoelectric point according to two review papers [3][6]. Gritzali and Brown [9] report that 4 different endoglucanases and 4 different exoglucanases have been isolated in the cellulase system of *Trichoderma viride*.

Some papers emphasize the large number of components but the limited basic types of reaction. Lee et al [10] report that the total cellulase enzyme system has been found to contain as many as 14 components but they reiterate that there are only a few—4 for their mechanism—basic types of reaction. The multiplicity of the enzyme system is also recognized by Ladisch et al. [11] but only 3 major types of components are recognized: endoglucanase, E_x, exoglucanase, E_2 and cellobiase, E_1. Again, these are the ones we shall use in our modelling.

There is another opinion of the cause of the apparent multiplicity. In Gong's opinion [12], and with experimental evidence, there are thought to be three basic enzymes, E_x, E_2 and E_1 for β-1,4 bond hydrolysis. Multiple enzymes are derived from the same parent enzyme and arise from partial proteolysis of the enzyme in question. Gong's three purified cellobiases are chromatographically distinct yet kinetically similar. This has been reported elsewhere [13]. In fact 9 other references listed in Mandels's review paper support the proteolytic action on cellulases. Yet, the latter reports that some components have been well characterized as having different modes of action [6].

Some protease enzyme in a broth might lead to post-translational modification of a parent enzyme. The increase in multiplicity with age of the culture seems to support the proteolysis theory [3].

This work is based in part on support of the proteolysis theory and of the three-enzyme theory. The proposal of Wood and McCrae [8] where the enzyme active site is altered as a function of substrate is not conducive to research looking for the mode of action of cellulases. In that case, the number of enzyme components would

prohibit the use of a simple mechanistic model to describe the hydrolysis.

2.2 Separation of Cellulases

Two commercial enzyme solutions are at our disposition. The first, Celluclast, is a highly concentrated cellulase mixture whereas the second, Novozym, is a highly concentrated cellobiase. Each solution contains different concentrations of the three enzyme components, E_x, E_2 and E_1. An ideal kinetic study would involve separating these components and investigating their separate effect on hydrolysis rates and their effect when combined in different ratios.

Darbyshire [14] discusses the whole spectrum of enzyme separation methods. The favorite method for separating cellulases is adsorption chromatography. Its high cost makes this method suitable for research or analytical purposes, not large-scale recovery.

The cellulase fractionation involves chromatography on Sephadex gels, ion exchange resins, electrophosphoresis on polyacrilamide gels or electrofocusing in a pH gradient. Best separations combine these methods. Poulsen and Petersen [13] use gel filtration and ion-exchange chromatography to isolate two endoglucanases from *Cellulomonas*. Gong et al. [12], and others [3][6] also enumerate the above techniques for separating cellulases and add avicel columns or immobilized concanavalin A as more recent techniques.

The complete separation of the components in a mixture is not easy. Because the reported loss of activity or loss of synergism of the enzymes on cotton varies widely due to their separation, there is a suggestion that even researchers are not capable of separating the enzymes completely [8]. We are here reminded of Gong's warning; the methods of purifying enzymes are always tedious and the recovery of enzymes is very low.

Because the separation of cellulases is a project in itself, and because we can

vary the ratio of the enzyme components by changing the ratio of the commercial enzyme mixtures, it was decided not to purify the cellulase components.

2.3 Hydrolysis Reaction Kinetics: Enzyme Mechanisms

A lot of information has accumulated since Reese first proposed a two-enzyme cellulase system [15]. And in spite of extensive research, the debate on the mode of action of cellulases continues [9][11][12][16]. Despite all the uncertainty, enough information must be extracted from the present state of cellulase kinetic research to build a mechanistic mathematical model. The mechanisms of the individual enzymes, E_x, E_2 and E_1 as well as their aggregate action is reviewed in this section.

2.3.1 Endoglucanase

In an early paper, when only endoglucanase, E_x, and cellobiase, E_1, were thought to act together to hydrolyze cellulose, G_x, Reese [17] proposed that E_x cleaves bonds between unsubstituted anhydroglucose units. Virtually all papers since then have reported this random action [6][10].

Furthermore, some modelling [10] assumes that E_x has a uniform affinity for the β-1,4 bonds regardless of location or degree of polymerization, DP. This assumption is necessary in the model development to keep the model form simple. Otherwise the substrate becomes too complex to describe mathematically. We have not found evidence in the literature that shows that the location of the β-1,4 bonds affects the rate of cleavage by the enzyme.

The proposed substrates of E_x are listed in review papers [3][6][18]. As previously discussed, all substrates contain β-1,4 bonds between two unsubstituted anhydroglucose units. Soluble cellulose derivatives (including CMC), provided they have a low degree of substitution, amorphous (non-crystalline) cellulose and cellodextrins are hydrolyzed by E_x. The activity on cellodextrins decreases as the chain length decreases and when it's cellobiose, chain length 2, there is almost no

activity. There is almost no activity on crystalline cellulose by the purified E_x. But on amorphous (reprecipitated) cellulose, there is a rapid decrease in the DP of the solution [3]. This list is repeated by Lee et al. [19] who goes further in stating that there is no G_2 hydrolysis by E_x. But Ladisch reports that both E_x and E_1 hydrolyze G_2 [11]. Conflicting reports are not rare in the treatment of cellulase kinetics. This shows the difficulty in analyzing data when no proper mathematical approach has been developed.

Note that all molecules containing β-1,4 linkages are substrates of E_x. The action on cellobiose is low to nil which explains the apparent reduced rate of reaction on low DP cellodextrins. Upon scission of the latter, G_1 and G_2 are immediately obtained, preventing further hydrolysis by E_x of the β-1,4 bonds of G_2.

After identifying the substrates of E_x, let us identify the products. Oligosaccharides and glucose are reported to be the hydrolysis products of cellulose attacked by E_x and E_2 [6][9].

Endoglucanase is said to act randomly by Ladisch et al. [11] but they list only glucose, cellobiose and cellotriose as the scission products. Fan and Lee [18] agree that E_x acts randomly. But they say cellobiose is the major hydrolysis product. It is difficult to reconcile these two statements because a totally random scission of the cellulose chain by an enzyme that cannot hydrolyze G_2 will yield equimolal amounts of G_1 and G_2 at completion of hydrolysis. One difficulty in identifying products is the inability to measure concentrations of cellulose oligomers of DP>6. The concentration of those species can only be lumped with the cellulose concentration because, like cellulose, they are insoluble. Yet the non-specificity of endoglucanase in releasing products. is pointed out in the literature.

Of interest to the model development also are the inhibitors of E_x. In Gong's review paper [12], E_x is said to act randomly to hydrolyze G_x. No inhibitors of E_x are reported. But some papers [20][21] report the inhibition of endoglucanases by

the accumulating cellobiose as a fact. So does Mandels in her review paper [6]. The latter adds methylcellulose—also known as methocel—as an inhibitor of E_x.

Ohmine et al. [22] found cellobiose to competitively inhibit E_x and E_2. However they also found that glucose was not an inhibitor of either E_x or E_2.

What is of most interest to us is the fact that the inhibitors of E_x all contain β-1,4 bonds that are not cleaved by E_x. In the case of cellobiose, it is a single bond but when high DS methylcellulose is involved it is those bonds which include at least one substituted unit. This inhibitory effect of uncleavable bonds on E_x will be verified by the modelling.

2.3.2 Exoglucanase

Exoglucanase is perhaps the least studied cellulase. It is the most difficult to isolate. And if its mode of action on cellulose is not well known, researchers are even more in the dark about its mechanism on substituted derivatives of cellulose.

A cellulase review [3] reports that enzyme fractionation studies first led to the conclusion that E_2 acts in an endwise manner to split off cellobiose. Further research by another group narrowed the mechanism to the non-reducing end of the polymer chain. This is supported by most cellulase research groups [11][18][23].

For example, only G_2 is obtained when G_4 and G_6 are hydrolyzed by E_2 but G_1 and G_2 are obtained when G_3 and G_5 are hydrolyzed by E_2 [24].

Action that some attribute to exoglucanase has been attributed by others to two separate enzymes, exoglucanase and cellobiosylhydrolase [10][19]. The former would split off glucose units from the non-reducing end of the polymer whereas the latter would split off cellobiose units from the non-reducing end. With this theory, cellobiase is assumed to hydrolyze cellobiose to two glucose molecules. Furthermore the cellobiosylhydrolase has no action on CMC. One group [19] admits that the first kind of exoglucanase—also known as glucohydrolase—has rarely been reported. This "double exoenzyme" theory is not used here initially because of its isolated

reporting. It will be seen that most researchers support the theory that E_1 has the dual purpose of hydrolyzing cellobiose and splitting glucose from the non-reducing end of a polymer.

Quantitative information is available; exoglucanase acts by removing the dimer from the non-reducing end of the chain with inversion of the β-glucosidic configuration [6] and an average Michaelis constant for E_2 has been reported [25] as $K_{mE_2} = 5 \cdot 10^{-3}$ M.

Polymers with two consecutive unsubstituted units at the non-reducing end are eligible substrates of E_2. They include amorphous cellulose, avicel (microcrystalline cellulose) and cellodextrins, with increasing activity as chain length increases [6]. Action of E_2 on amorphous cellulose decreases the DP in a slow, linear fashion [3]. There is no action on cotton or cellobiose by the purified enzyme. But Mandels claims limited action by E_2 on CMC. Of course, the presence of substituent groups reduces the probability of finding two consecutive unsubstituted anhydroglucose units at the CMC extremities.

Thus cellobiose is the specific product of cellulosic hydrolysis by E_2. However in keeping with the uncertainty of the mechanism of cellulase kinetics, one paper refers to the production of mono-saccharides by E_2 [26]. We repeat that this glucohydrolase action will be attributed to E_1 in this work.

As we did for endoglucanase, we now identify the inhibitors of exoglucanase. In their summary of the literature Gong et al. [12] report that E_2 produces G_2 through and endwise cleavage of the cellulose chain and that G_2 is an inhibitor of E_2. Others report inhibition of exoglucanase by cellobiose [3][6][11][20][22]. The latter group [22] found that G_1 is not an inhibitor of E_2. Methylcellulose is also a strong inhibitor according to Mandels [6]. One report of glucose being an inhibitor of E_2 is questionable [11] as they name virtually all compounds to inhibit all enzymes. Reports that G_2, the product, is an inhibitor of E_2 agree well with enzyme theory

stating that a product of an enzymatic reaction is a likely inhibitor. This is feedback control in nature.

Finally, one should be aware of warnings about assuming that E_2 attacks cellulose in an endwise manner just because a high concentration of G_2 is noted in hydrolysis mixtures. A study of the oligosaccharide chain of cellulose shows that only alternate β-1,4 bonds are exposed to attack by cellulases. The other bonds are protected because they face into the body of the solid. Alternate bonds would be cleaved leaving fragment of DP 2, 4, 6... held to the crystalline cellulose by hydrogen bonding. Freeing a tetrasaccharide into solution requires the breakage of more hydrogen bonds than freeing a cellobiosyl. The less probable splitting off of four consecutive units makes those units candidates to be further hydrolyzed to two G_2 molecules. Hence, we must be careful to interpret the presence of cellobiose as evidence of endwise attack [3].

With all of the evidence presented concerning the mechanism of E_2 action, our initial attempts at modelling the enzymatic hydrolysis of CMC will assume that E_2 effect is negligible because of the presence of substituent groups. However if a significant concentration of cellobiose is found in the hydrolyzates, we can modify the proposed model to take into account the significant hydrolysis rate of E_2.

2.3.3 Cellobiase

In one of the earliest papers on cellulase enzyme mechanisms, Reese [17] proposed that E_1 cleaves units one at a time at the non-reducing end of the polymer chain provided said unit is unsubstituted. This proposition has been challenged in part by researchers claiming a separate enzyme, exoglucanase, separate from cellobiosylhydrolase, cleaves polymer ends. In this case, cellobiase is assumed to act only on cellobiose [10].

In this work, we assume that the former theory holds. If we were to assume the cellobiase action is by two separate enzymes, the model would take a more

complex form since four instead of three enzymes would act on CMC. Besides, the four enzyme theory is only a conjecture of some researchers. It is unsupported by experimental evidence.

In their summaries of the literature, Gong et al. [12] and Fan and Lee [18] report that E_1 hydrolyzes G_2 and perhaps other cellodextrins to give G_1. G_1 is said to be an inhibitor of E_1. In another review of the literature [6], cellobiose, β linked glucose dimers, cellodextrins, aryl β-glucosides, (salicin, p-nitrophenyl β-glucoside) are reported as E_1 substrates. Furthermore, the activity on cellodextrins increases as the chain length decreases which is opposite to endoglucanase action. This chain-length dependency is also noted by Lee et al. [19]. No action of E_1 on G_x was reported. The transfer of a glucose unit to other sugars and alcohols is suggested.

The chain length dependency of the activity is, in our opinion, only apparent. The actual substrate of E_1 is really the non-reducing end bond. At a given mass concentration of a dextrin, there are more non-reducing end bonds for a lower DP dextrin, hence more substrate for E_1, hence the rate of bond cleavage by E_1 is higher.

Ladisch et al. [11] point out the high activity of E_1 toward cellobiose and only speculate on possible activity with respect to cellodextrins.

The data of Matteau and Saddler [20] show that E_1 cleaves the bonds of salicin at the same rate as that of G_2. After 12 hours of hydrolysis, glucose yield from G_2 hydrolysis is approximately 1.25 mmol G_1/mmol initial G_2. The yield from salicin hydrolysis is 0.62 mmol G_1/mmol initial salicin. Only one of the salicin's two units is glucose, accounting for the 2-factor.

It appears from these observations that E_1 cleaves β-glucoside bonds at a constant rate, regardless of the reducing side unit on the non-reducing end bond of the molecule. This is to be an important assumption of our model. Applied to all three

enzymes in the solution, it means that all substrate bonds—to one enzyme— are cleaved at a rate independent of the units which are linked by the said bond.

The high activity of cellobiose on its substrates is evident in the literature. An example of the effect of cellobiase on the hydrolysis of cellulose can be found in a paper by Gritzali and Brown [9]. Little cellobiose is found when 0.001% E_1 is added to a cellulose hydrolysis reactor that contained the whole cellulase enzyme system. Without the extra cellobiase, cellobiose and glucose concentrations are comparable for the first 24 hours of the reaction. With the addition of cellobiase, the cellobiose concentration is virtually nil in the reactor as the reaction proceeds to completion.

Humphrey [4] also shows a typical enzymatic digestion of cellulose. On his curve, the maximum cellobiose concentration occurs at 50 hours. Where this maximum occurs is a function of the concentrations of E_x, E_2 and E_1. Cellobiose always exhibits a maximum between its initial and final concentration ($G_2 = 0$ at $t = \infty$). The reason for cellobiose presence is the poor β-glucosidase activity of *Trichoderma* says Humphrey. The poor cellobiase activity of *Trichoderma* is echoed by Gong et al. [12] and Mandels [6], the latter showing a 20% increase in avicel and milled pulp hydrolysis reaction conversion by the addition of E_1 to equal 1 unit/filter paper unit in the hydrolysis reactor.

The product of the hydrolysis of any cellulosic by E_1 is acknowledged by all to be G_1.

There is some quantitative information available on E_1 action. An average Michaelis constant for E_1 has been reported [25] as $K_{mE_1} = 10^{-4}$ M. Cellobiase is reported inhibited by glucose ($K_i = 10^{-3}$ M) and by substrates cellotriose and cellotetraose when these are highly concentrated. Amazingly, the same authors report that the hydrolysis of cellobiose by E_1 yields G_1, G_2 and... G_3! For example, 30 g/L of cellobiose yielded approximately 21 g/L cellobiose, 5 g/L glucose and 4 g/L cellotriose after 2 hours of reaction time. This supports their glucosyl transfer theory

[9].

We have already mentioned the product glucose as an inhibitor of E_1. Ohmine et al. [22] have inserted a term to account for glucose as a competitive inhibitor of E_1 in their initial cellulose hydrolysis model. Glucose was found to be an inhibitor of only E_1 because its addition to a mixture reduced the glucose yield while not affecting the reducing sugar concentration.

Montenecourt and Eveleigh [21] report other papers showing glucose to be a non-competitive inhibitor of cellobiase. From their work, Fan and Lee [18] find glucose to inhibit E_1 competitively. Matteau and Saddler [20] accept it as an inhibitor, without specifying whether it is competitive or non-competitive. One review paper [11] reports that both cellobiose and glucose have been found to inhibit E_1.

With enzyme mixtures containing three enzyme components, it is unlikely that we can show cellobiose to be an inhibitor of all of these. Thus we will only verify for G_1 inhibition of E_1 in the model. The model form can be changed to take into account both competitive and non-competitive inhibition.

2.3.4 Synergism

We have discussed the individual modes of action of cellulases. A recurring theme in the literature is synergism. The action of the sum of the enzyme mixture on cellulose is greater than the sum of the action on cellulose of the three components individually. That is the phenomenon that cellulase researchers agree on: synergism [6][18], especially between E_x and E_2 [11].

The initial hydrolysis rate is high as the accessible cellulose is hydrolyzed but if one of the cellulase components is missing, then the rate of cellulose hydrolysis falls off rapidly [6][19].

At least E_x and E_2 are required to hydrolyze crystalline cellulose, but either of the two will hydrolyze amorphous cellulose [16]. Purified E_1, E_2 or E_x cannot degrade crystalline cellulose (cotton) separately but upon recombining these com-

ponents, the initial activity is recovered [8]. Individually, E_x and E_2 hydrolysis of crystalline cellulose is about 40% of their combined or recombined rate of hydrolysis [19].

There are however limitations on elevating the hydrolysis rate by adding one component of the enzyme system. For example, the stimulatory effect of adding cellobiase to a reaction mixture is limited to 1.5 cellobiase units per filter paper unit. Any more cellobiase does not show any improvement in hydrolysis rate. The same phenomena limits the addition of E_x or E_2. But if all enzyme concentrations are increased concurrently, then the hydrolysis rate of cellulose will increase [6]. Matteau and Saddler [20] have also reported the usefulness of adding cellobiase to the reaction mixture to increase the glucose yield. In their case, the cellobiase was immobilized.

Synergism between enzymes of different fungi is present only if the fungal source produces E_x and E_2 capable of degrading cellulose. The synergism of cellulases is low on less well-ordered cellulose and absent on soluble cellulose derivatives. However since the reported loss of activity varies widely, Wood and McCrae [8] suggest that many of the enzyme separations are incomplete.

The mechanism on cellulose of E_2 and E_x has not been established and the order of action of those two cellulases in the hydrolysis is in dispute. A good discussion on arguments for and against one of the enzymes acting first can be found in [3], [6] and [8]. Does E_x initiate the attack generating new chain ends to be hydrolyzed by E_2 or does E_2 produce some disorder in the cellulose chain to be further hydrolyzed by E_x at the weak point? The authors of the discussion papers support the initial action by E_x to produce free chain ends for the exoglucanase and—not mentioned by the authors—cellobiase.

Other mechanisms have been proposed [3]; E_x and E_2 may combine to form the complete hydrolysis enzyme. One of the enzymes, the "hydrolytic" factor can

hydrolyze soluble derivatives but the hydrolytic factor must be held in position on insoluble cellulose by the "affinity" factor, E_2.

An interesting result of Gong et al. [12] is that the combination of E_x and E_2 when incubated together with substrate Avicel gives double the conversion of E_2 used alone. E_x used alone has little activity. Of course E_x may have been equally active in the presence of, or without E_2 but since Gong was measuring sugars released, the concentration of smaller fragments would not be recorded. But those smaller fragments may provide more substrate for E_2.

The way in which synergism has been handled by most groups attempting to model cellulosic hydrolysis is shabby. One group has shown that by assuming each enzyme has its own particular substrate, the synergism is explained. This will be discussed in the section Hydrolysis Modelling.

For example, assuming that E_x is a slow, random acting enzyme and E_1 is a fast acting enzyme cleaving bonds at the non-reducing end of a polymer, initially the concentration of available ends for E_1 is low. However as the reaction proceeds, E_x increases the substrate of E_1. The solubilization of G_x is quick. If either of the two enzymes is missing from a mixture, the hydrolysis rate is reduced considerably.

In the model that we shall develop, synergism is taken into account by the normal mode of action of the three enzymes.

While on the subject of the overall enzyme system, we now discuss what components researchers have found to generally inhibit the overall system.

Products (reducing sugars) are reported to be inhibitors of the cellulase enzyme system [3]. Cellobiose is reported to be an inhibitor of the overall enzyme system [9]. Most agree that it is a stronger inhibitor of cellulase action than glucose [16]. Cellobiose, in contrast to the above reports, has been described as an activator when CMC is the substrate [6]. Glucose is reported to be an inhibitor of the overall enzyme system [7][16][27], with [7] reporting because the yield of glucose decreased

with increasing initial substrate but the overall conversion was not affected.

Fan and Lee [18] report that the products of hydrolysis are probable non-competitive inhibitors of the surface reaction on cellulose, slowing down the action of E_x and E_1.

Perhaps most revealing about the state of understanding of cellulase inhibition are Howell's paper [23] and Ladisch's review paper [11]. The former reports that cellulases have been found to be inhibited by cellobiose, the lower oligomers and by excess cellulose. The latter reports that the soluble products G_1 and G_2 are inhibitors of the cellulase complex and of all individual enzyme components, E_x, E_2 and E_1. It is possible that researchers always assume inhibition is responsible for the lowering of an apparent hydrolysis rate upon addition of an extra component. However it has been shown that the hydrolysis rate of the pre-addition component will go down when more of a substrate is added to a reaction mixture [28].

One of the important aspects of cellulose hydrolysis research has concentrated on the phenomena of adsorption of the cellulase components onto the cellulose. Not all three components adsorb at the same rate. In fact, β-glucosidase shows little adsorption but E_x and E_2 show a lot of initial adsorption followed by a levelling off of the quantity of enzyme adsorbed. A good review of cellulase adsorption can be found elsewhere [3]. Since our substrate is to be soluble, the enzyme adsorption onto the substrate is non-existent and it will not be further discussed here.

2.4 Optimal Hydrolysis Reaction Conditions

The hydrolysis reaction rate is a function of both temperature and pH, as are most enzymatic reactions. In this section one can find a quick survey of the literature to determine what are the optimal temperature and pH of the reaction. There are maxima in both cases because of tolerance levels of the enzymes. It will be seen that the optimal temperature for *Trichoderma reesei* (our enzyme source) is 50°C and the optimal pH is 4.8.

The reaction kinetic parameters in this work are evaluated at the optimum conditions. The analysis of the kinetic behavior as a function of the temperature and pH for cellulases is a project in itself. We make use of the presently known optima.

2.4.1 Optimal pH

Kanda et al. [24] report a pH range of 3.5 to 6.0 where the activity of E_2 from *Irpex lacteus* shows little variation. The optimal was 5.0. Heptinstall et al. [29] reports an optimal pH of 5.2 for E_2 from *Aspergillus fumigatus*. The optimal pH for the *Thermomonospora sp.* cellulase system minus β-glucosidase is 6.5 [6].

Zabriskie [7] used 5 different sources of enzymes, including *Trichoderma reesei*, and conducted his cellulose acetate hydrolysis experiments at a pH of 5.0 in all cases. The pH was controlled with a 0.05 M acetate buffer. For *Trichoderma* he was at the reported optimal pH of 4.8 for the endo and exoglucanases [6][11][20]. The pH of 4.8 is reported used for the hydrolysis of ball milled solka floc [3], acid swelled solka floc [16] and avicel [22] by *Trichoderma viride* cellulase. From the earliest times [15], a pH of around 4.8–5.0 has been used for the enzymatic hydrolyses by *Trichoderma*. We will be no different, conducting our experiments at the optimum pH of 4.8.

2.4.2 Optimal Temperature

Kanda et al. [24] report an optimal temperature of 50°C for hydrolysis by *Irpex lacteus* cellulases.

The optimal temperature for the cellulase system of *Thermomonospora sp.* minus β-glucosidase is 50°C [6].

Zabriskie [7] used 5 different sources of enzymes, including *Trichoderma reesei*, and conducted his cellulose acetate hydrolysis experiments at a temperature of 50°C in all cases. For *Trichoderma* he was at the reported optimal temperature of 50°C for the endo and exoglucanases [6]. This temperature has also been used to assess the kinetics of the hydrolysis of G_2 and salicin by cellobiase [20]. This temperature of 50°C is reported used for the hydrolysis of ball milled solka floc [3], acid swelled solka floc [16] and avicel [22] by *Trichoderma viride* cellulase. Ladisch et al. [11] are more conservative, stating that the optimal hydrolysis temperature for *Trichoderma reesei* cellulase is in the range 40–50°C.

We will conduct our hydrolysis experiments at the optimum temperature of 50°C.

2.5 Enzyme Activity Assay Procedures

The enzyme concentration cannot be conveniently measured in units of moles of enzyme per volume of solution. Instead standard tests are run on how much substrate uptake is realized by a known quantity of enzyme in a given period of time. The substrate uptake rate is of course proportional to the concentration of enzyme in the assayed solution. The respective enzyme dilution curves must be linear for the test to be valid [30]. The concentration of substrate should be high enough that substrate depletion does not limit the reaction. The extent of conversion should be slight, hence the test time short, before depletion of the substrate or product inhibition affects the reaction rate.

Canevascini and Gattlen [26] describe the results of many different cellulase assay procedures tested on enzyme mixtures of different origin.

There is no total agreement on the best choice of substrate for measuring the activities of the three cellulase components. This is partly due to a lack of knowledge about the mechanism of each enzyme.

And if the substrate cannot be agreed upon, then certain branches of science certainly disagree on the overall test procedure.

> "It seems unlikely that any single assay procedure for enzyme mixtures acting on native cellulosics can satisfy biochemists trying to understand synergism, microbiologists evaluating strains and chemical engineers involved in process development" [6]

Demeester [31] goes on to say that determination of the activities of the individual components after purification is very time-consuming and not suitable for practical work. However Canevascini [26] states that as long as the different components have not been separated, the determination of their activity is a difficult task, well reflected in the diversity of analytical methods.

One paper discussing the methods of estimating the concentration of cellulases in a mixture contains a table of substrates used for the measurement of individual components and components acting in concert. For example, four substrates are listed to assay E_1, three to assay E_x and three to assay E_2. These will be discussed later. The same paper [32] blames confusion in the cellulase assay procedures on the bewildering array of substrates, enzyme actions, units, and activities that have been used.

In cellulase enzyme assays, the test substrate will usually contain β-glucosidic units. These are hydrolyzed to reducing sugars, the concentrations of which are measured as the indicator of the extent of reaction.

The dinitrosalicylic acid (DNS) procedure can be used to measure the concentration of reducing sugar in a mixture. It is the procedure used in this work. It is

described as part of the filter paper activity test description in the Experimental Procedure section of this thesis. Several cellulase researchers use DNS to determine the amount of reducing sugars in a mixture [7][16][22][23][32]. However it has been reported that the citrate buffer—see the Experimental Procedure—causes a decrease in the reducing power of sugars [33]. This may reduce the sensitivity of the test. The partial destruction of glucose by alkaline oxidation results in biased results especially at low glucose concentrations.

An alternative method to DNS is the Somogyi and Nelson procedure [33][34][35]. Several cellulase researchers use this method to evaluate the reducing power of their hydrolysis mixtures [13][26][36]. It is also a colorimetric method but used as an iodometric method, it gives more precise results.

Lindner et al. used both the DNS and the Nelson and Somogyi assays for monitoring the reducing sugars produced by the hydrolysis of CMC. Their results and subsequent discussion show support for Nelson and Somogyi's procedure. It is more accurate and yields linear dilution curves at conditions where DNS is non-linear.

The DNS procedure was nonetheless used in this work for historical reasons; it had been used in our biochemical engineering group from its beginning in 1980.

There is a warning issued in the same paper by Lindner et al. [30]. The product estimation by most colorimetric reducing sugar assays is non-stoichiometric since the aldehyde group reacts differently depending on the residue to which it is attached. This is especially true for CMCase activity due to the wide product variety. Thus the reducing sugar test cannot be used to measure the concentration of reducing ends in a mixture where short cellulose derivative oligomers are present. This was a further motivation for us to measure the specific concentrations of glucose and cellobiose in the reaction mixtures using HPLC.

2.5.1 Filter Paper Activity

According to Mandels et al. [32], for a *practical* measurement of saccharifying cellulase, measurement of any single enzyme concentration is unsatisfactory. For this reason, the filter paper assay was introduced. We suggest that for research purposes, as long as the mechanism of the cellulase system has not been identified, all enzymes should be assayed separately. Having one assay for a three enzyme system is questionable especially if the ratio of these enzymes in a mixture affects the concentration of any given product. The difficulty of assaying enzymes separately is the doubt surrounding the particular substrate that will be acted on by only one enzyme of the cellulase system.

Filter paper activity of an enzyme solution measures the overall activity of an enzyme mixture [26]. The description of the test can be found in the Experimental Procedure section of this thesis. It has been modified in places [7] to use an acetate buffer instead of a citrate buffer to minimize interference on the DNS method. We use the FPA test as the assay of endoglucanase concentration.

2.5.2 Endoglucanase Activity

Gong et al. [12] used $5\,g/\mathrm{L}$ CMC as the substrate for measuring E_x activity. The reducing sugars are monitored with DNS coloring for detection on the spectrophotometer. Heptinstall [29] reports that CMC is the substrate to use in monitoring E_x activity. Reese et al. [15] recommended using low DS CMC for the purpose of assaying E_x but that was before the theory of existence of E_2. With enzyme multiplicity, it is now thought that the hydrolysis rate of a low DS CMC will be a higher function of the enzyme solution composition.

There are two methods of keeping track of the activity during the test: measure viscosity or the increase in reducing sugar [26][29]. There is a stronger case for the former [26] because the viscometric method is specific, sensitive and free of interference. When measuring reducing power, E_2 is thought to contribute to the

measurement [26]. Lindner and co-workers make a strong case against the reducing sugar test using CMC as substrate. They, as most, conduct their CMCase activity test at the optimum reaction conditions for E_x, 50°C and pH 4.9. They note that non-linear dilution curves are obtained when monitoring reducing sugars as a function of enzyme at the constant assay reaction time. The non-linearity problem can be solved by diluting the amount of enzyme further but the reducing sugar concentration is so low that the test loses accuracy. And they further warn that the hydrolysis of CMC monitored by reducing sugar concentration will be a function of enzyme concentration and composition. This must be assumed until the reaction mechanism is eventually established.

Again according to some [26][32], there is variation in the DS, the time of incubation and the reaction temperature between researchers measuring enzyme activities. This is why tests of CMCase activity are in need of standardization. Because the CMC hydrolysis lasts one hour and the increase in concentration of reducing sugar is very non-linear, the determination of the rate of hydrolysis over that period of time may be much lower than the initial rate of hydrolysis [13]. Then a dilution series of the enzyme becomes necessary.

Mandels et al. [32] in their review of the literature list three substrates that may be used to assay E_x activity: amorphous cellulose, CMC, cellodextrins. However the use of cellodextrins and amorphous cellulose to assay E_x must be questionned; amorphous cellulose is also included as assay substrate of E_2. Cellodextrins are also included as assay substrate of E_1 and exoglucanase, E_2.

Demeester's proposed test [31] may be the best measurement of endoglucanase activity yet. A high molecular weight hydroxyethyl cellulose (HEC) with a DS of 1.3 is the substrate. This is ideal since only E_x would be able to break any bonds if it is true that E_1 and E_2 can only cleave unsubstituted units from a non-reducing end of the polymer chain. The test instrument is a capillary viscometer. The response

variable is viscosity. To try to measure reducing sugar in the test is useless because little is produced at high DS. And those reducing sugars are also a concentration of E_2 and E_1. Against the test is the necessity to calibrate the viscometer and to characterize the substrate.

2.5.3 Exoglucanase Activity

Gong et al. [12] used avicel as the substrate for measuring E_2 activity. The reducing sugars are monitored with DNS coloring for detection on the spectrophotometer. Use of avicel or filter paper is suitable for the conjoint determination of E_x and E_2 claims Canevascini [26] in rejecting that test procedure.

Mandels et al. [32] in their review of the literature list three substrates that may be used to assay E_2 activity: amorphous cellulose, crystalline cellulose and cellodextrins. However the use of cellodextrins and amorphous cellulose to assay E_2 must be questionned; amorphous cellulose is also included as substrate to assay E_x. Cellodextrins are also included as substrate to assay E_1 and E_x. In fact, a 1982 review paper [6] claims that no measurement procedure existed at that time for E_2 because of synergism.

Heptinstall et al. [29] recently claim to have developed a more reliable assay procedure for exocellulase to eliminate the influence of the E_x concentration on the test results when avicel is used as the substrate. The substrate used by Heptinstall is MU-cellobiose (β-D-Glcp-(1$\rightarrow$4)-β-D-Glcp-7-O-4-methylumbelliferone). However, to overcome interference by E_1, D-glucono-1,5-lactone must be included in the assay. The reaction progress is followed fluorimetrically.

In our work, the exoglucanase concentration is not measured. Our first assumption is that its effect on CMC is negligible. But following initial model fitting, it is introduced to explain the high cellobiose concentration. The variance in the G_2 data is so high that an assumption of virtually infinite concentration of E_2 in the mixture is satisfactory. More detailed work by cellulase researchers will eventually

need an exact measurement of E_2.

2.5.4 Cellobiase Activity

Zabriskie et al. [7] measured the release of glucose from a 30 minute hydrolysis of a $5\,g/\mathrm{L}$ solution of cellobiose at 50°C and pH 5.0. Gong et al. [12] used cellobiose as the substrate for measuring cellobiase activity of purified E_1 but *p*-nitrophenyl-β-glucoside was their substrate of choice for measuring cellobiase activity during purification procedures. The product glucose was measured with a Beckman glucose analyzer in the first assay and the release of nitrophenol was monitored colorimetricaly in the second assay. It is presumed that *p*-nitrophenyl-β-D-glucoside does not react with E_x or E_2 and thus makes it ideal for monitoring E_1 activity in a crude cellulase mixture. However Mandels [6] states that cellobiose is the substrate of choice for cellulase work since there are some aryl β-glucosidases which do not hydrolyze cellobiose.

Heptinstall [29] and Mandels [6] also mention the two substrates as the most popular for monitoring β-glucosidase activity. The latter suggests the aglucone saligenin can be followed. This third possible substrate is the one used by Matteau [20], salicin, to follow E_1 activity. The test can be followed colorimetrically since glucose and saligenin do not interfere with each other on the spectrophotometer. Their results comparing salicin and cellobiose hydrolysis by E_1 are consistent, indicating the same activity of E_1 on both substrates.

Hägerdal et al. [37] have also used the two substrates to measure E_1 activity. They give different names to the activities as measured by the release of glucose from cellobiose and as measured by the release of nitrophenol from *p*-nitrophenyl-β-D-glucoside. The former is called cellobiase activity and the latter is called β-glucosidase activity. The two measurements are not in the same ratio when measured in different mixtures. Thus they are not a measure of the same enzyme(s) concentration. This is also reported by Canevascini [26], however the precision as-

sociated with the results is not reported by either author. From the latter, one might conclude that the cellobiase activity is about twice the aryl-β-glucosidase activity. The aryl-β-glucosidase activity is lower than or equal to cellobiase activity in virtually all cases.

Mandels et al. [32], in their review of E_1 activity tests lists the three substrates previously discussed, G_2, salicin and p-nitrophenyl β-glucoside plus another, cellodextrin. However the use of cellodextrin must be questioned since it is also listed as an assay substrate for E_2 and E_x.

2.5.5 Protein Content

Note that the protein content of a mixture can be taken as indicator of the total concentration of enzyme in the solution if enzymes are the only protein in the solution. This is not discussed here because the procedure is even less specific as an indicator of cellulase concentration than the FPA test.

2.6 Enzyme Deactivation

As higher and higher conversions are desired in the hydrolysis reaction, the time of reaction must increase. The hydrolysis rate will slow however due to product inhibition and enzyme deactivation. If the hydrolysis substrate is cellulose, the rate also slows due to increasing resistance of the residual substrate [6]. This section is concerned with enzyme deactivation.

Howell and Stuck [23] note that cellulases are not very sensitive to deactivation by high temperatures or in the presence of poisons. They remain active from pH 2.0–8.5, at 50°C for at least 24 hours. There is also no loss of activity in a 4 M urea solution.

Enzyme stability of cellulases are reported by Hägerdal et al. [37]. For *Thermoactinomyces sp.* only 15% of the CMCase activity is lost over 24 hours at 60°C. At 60°C, 50% of the avicelase activity is left after 24 hours but at 55°C the stability

is only marginally affected. Cellobiase activity is most affected by temperature; the half-life of E_1 is 8 hours at 55°C and 1 hour at 60°C. Note that these data apply to *Thermoactinomyces* and that the optimum pH range for hydrolysis by *Thermoactinomyces* is 6.0–7.3.

Hägerdal et al. also cite a study by Andren where the stability of the cellulolytic enzyme system from *Trichoderma viride* QM 9414 was reported as 40% loss of activity when maintaining the enzyme system at 50°C for 24 hours. But at 30°C no loss of activity was found after 24 hours.

The half-life of E_1 from *Trichoderma E-58* is 1 day at 50°C according to Matteau and Saddler [20].

Ohmine et al. [22] determined that E_2 and E_1 are relatively stable at 50°C. E_x was deactivated to a half-life of 280 hours at that same temperature. The components were in a mixture produced by *Trichoderma viride* and the incubation period was 30 minutes.

Mandels in her review paper reports the thermal stability of *Trichoderma reesei* E_1. At 40°C, there is 95% loss after 1 hour incubation. But at 60°C, there is 100% loss. However this is for the intracellular enzyme. For the extracellular E_1 used in kinetic studies most of the time, there was no reported loss for a 1 hour incubation at 60°C.

In Mandels' review paper [6], there are references to papers showing the β-glucosidases and endoglucanases to be more resistant to changes in pH, temperature and chemical inhibitors than exoglucanases.

All of the papers enumerated above show that the enzyme deactivation rate of cellulases is dependent on the component—E_x, E_2 or E_1—, its source and the temperature amongst other things. For cellulases, this rate may vary from 0 to 80% after one day. A significant drawback of any work, including this one, is the failure to measure and account for enzyme deactivation in the hydrolysis reaction.

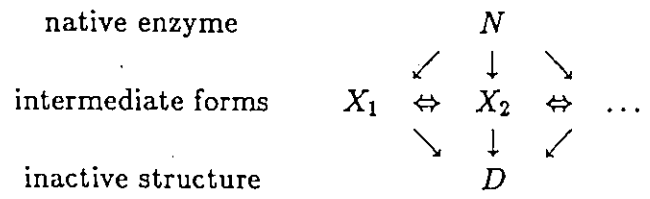
The alternative to measuring the deactivation of enzymes is to modify the model and insert parameters that describe deactivation as it is currently understood in the literature.

The deactivation of enzymes often follows a first order decay law [28]:

$$\frac{de}{dt} = -k_d \cdot e \qquad e(t) = e_0 \cdot \exp(-k_d t) \tag{3}$$

The decay constant, k_d often has an Arrhenius type dependency on temperature. Since our experiments are to be done at a constant T, it would be simple to add some parameters to the model, one parameter for each enzyme to estimate its rate of decay during hydrolysis. However, we will discuss later why a parsimoniously parameterized model is preferred.

The enzyme deactivation may not follow such a simple pattern. Gianfreda et al. [38] provide a general discussion—not specifically related to cellulases—on complex enzyme deactivation. Non-exponential decays may be two-sloped on a logarithmic plot. Another phenomena is the time-progressive increase in the rate of deactivation. The latter is explained by assuming that the native enzyme reaches its final, totally inactive form through intermediate, partially inactive proteic structures [38] as shown here:



Therein may be a link with the proposed proteolytic action on cellulases. To verify this link is beyond the scope of this work. But if the cellulase deactivation rate is important at 50°C, and it is not simple in form, then the overall hydrolysis model would be complicated by the addition of several parameters to take this mechanism into account.

Most models pertaining to enzymatic hydrolysis by cellulases choose to ignore the deactivation of the enzymes [3]. Thirteen out of fifteen reported models did not take deactivation into account. In order to maintain parsimony in the parameters, we also do not include deactivation in the model. But this section is included as a *caveat* to any user of a fitted model which disregards deactivation.

2.7 Choice of a Substrate

Several substrates may be used to investigate the mechanism of enzymatic hydrolysis by cellulases: cellulose, cellodextrins, cellulose esters or cellulose ethers. By a process of elimination, an ether of cellulose was chosen. The justification is presented here.

2.7.1 Cellulose

The hydrolysis reaction can be surface limited because the substrate is solid [4]. The rate of reaction should be a function of surface area because direct physical contact must occur between cellulose and cellulase for the reaction to occur [3][18]. This is probably why soluble cellulose derivatives are hydrolyzed more rapidly than cellulose [7]. The initial hydrolysis rate for CMC is 4.1 times faster than for solka floc under otherwise identical conditions. This difference is greater if the cellulose substrate has even higher crystallinity [7]. This is due to inter-chain hydrogen bonding.

The structure of cellulose is not perfectly established. Three postulated representations of elementary fibrils in native cellulose are presented in a recent review paper [3]. Yet it is known that cellulase produces extensive changes in cellulose prior to producing a measurable quantity of reducing sugar: fragmentation, swelling, cracking, lowering of the DP.

Here, we repeat the theory that cellulases adsorb onto cellulose. The protein content in the hydrolysis reactor has been reported to drop from its initial level

when insoluble cellulose is the substrate [16]. Refer to Gilbert and Tsao [3] for a complete discussion on adsorption.

A great deal of enzyme is required to saccharify cellulose at high conversions. A filter paper activity test lasts only one hour and the conversion is low. If one is to model cellulose hydrolysis, data are needed at high conversion. At the high conversions, the enzyme efficiency—ratio of achieved conversion to that predicted by the enzyme unit value—is one-tenth of predicted by the FPA test [6]. Thus shake-flask experiments will require a great deal of enzyme when a high conversion of solid cellulose is desired.

Perhaps the most important reason for rejecting cellulose as substrate lies with the initial objective of this work. The hydrolysis reaction would be heterogeneous. Intrinsic kinetic data on endo and exoglucanase are lacking, so that it is necessary to separate the kinetic and diffusional effects [11][39].

2.7.2 Cellodextrins

Cellodextrins are short cellulose oligomer chains. They are soluble in water at a DP$\leq$ 6 [3]. They have no attached substituent groups.

Cellodextrin hydrolyses have been modelled, though not always successfully. One paper [11] reports an increasing rate of reaction with chain length when exposure is to endoglucanase and a reduced rate when exposure is to cellobiase. One kinetic model is presented. However its range of validity is limited to cellodextrins of DP less than 5. Each new dextrin of DP greater than 4 would increase the number of parameters in the model. This is one example that shows the poor nomenclature is causing problems in the modelling of cellulase kinetics. Ladisch's paper does however encourage the use of cellodextrins for kinetic studies of the cellulases' hydrolysis mechanism because of their solubility.

Mandels et al. [32] in their efforts to find the 'one' substrate that can be used to measure the overall cellulase activity rejected cellotetraose because it is not available

commercially and preparing it is a major research effort in its own right.

The preparation of cellodextrins is described by Huebner et al. [40]. An acid hydrolysis of cellulose is the key step of six steps in the preparation of the cellodextrin mixture. The cellulose solids at the end of the reaction are filtered off. The resulting cellodextrin solution can be concentrated 25 fold by evaporation. During the acid hydrolysis, there is partial degradation of the sugars. This is the first difficulty. The second one is more important. The first six steps of the cellodextrin production leave a mixture of G_1–G_6. The last step in Huebner et al.'s method is the separation of the dextrins by preparative chromatography. This is time-consuming yet the alternatives are worse according to the authors.

A study of cellodextrin hydrolyses is not practical because of the difficulty in preparing pure solutions of the oligomers. Only a few millilitres per day of cellodextrins could be produced. And the concentration of cellohexaose and cellopentaose is significantly lower than cellotetraose and cellotriose [41]. Since we wanted to work on a larger scale than the test tube, the use of cellodextrins as substrate was rejected.

2.7.3 Cellulose Esters

Reese [17] found that slightly more glucose was obtained from the hydrolysis of cellulose esters—cellulose sulfate, cellulose acetate and cellulose acetate phthalate—than from 4 different cellulose ethers, at equal DS. The reason suspected by Reese is the presence of acetyl esterase in the hydrolysis mixture. This enzyme would deacetylate the substituent group making a subsequent attack by cellulases easier.

The objective of Zabriskie et al. [7] was to accelerate the production of glucose. Cellulose acetate was used as substrate because esterases present with the cellulase are capable of removing the substituent group, thereby improving glucose yield. With the intentionally added esterase, 70% of the cellulose acetate is saccharified before the hydrolysis reaction seems to level out.

Because of interference by esterases, esters were rejected as hydrolysis substrate.

2.7.4 Cellulose Ethers

The ether group on cellulose is very stable [42] but it is possible to regenerate the cellulose by removing the substituent [3].

An ether of celllulose was chosen as our substrate because it did not present some of the problems associated with the other substrates; unlike cellulose, it is soluble; unlike cellodextrins, it is available commercially and thus in large quantities; unlike cellulose esters, its substituent group cannot be eliminated by an enzyme.

CMC is a cellulose ether. Since we had some in large quantities and since it presented the least drawbacks short of hydrolyzing theoretically soluble cellulose, it proved to be our substrate of choice.

2.8 Hydrolysis Modelling

A recent review by Gilbert and Tsao [3] lists models developed to describe cellulose hydrolysis. The important assumptions made in developing the fifteen models are included in the review.

Two models consider the diffusion of the enzyme in the solid substrate, five models consider the enzyme adsorption onto the substrate and of those five, only two consider the decrease in adsorption sites. This shows that many cellulase researchers are not ready to allow for the importance of physical phenomena in cellulose hydrolysis. However, physical phenomena are not significant in the hydrolysis of a soluble cellulose derivative like CMC.

Only six models out of the fifteen listed have considered substrate multiplicity. For cellulose, it should be obvious that the substrate is not uniform; there are some crystalline and amorphous sections. But substrate multiplicity goes deeper than that. Even soluble substrates are not uniform. In substituted cellulose derivatives, the degree of substitution affects the extent of hydrolysis. All cellulosics may be

hydrolyzed in an endwise manner hence the degree of polymerization affects the rate of hydrolysis. We will show in this work that a redefining of the substrate by changing the nomenclature to describe the enzymatic hydrolysis of anything that contains β-1,4 glucosidic bonds leads to an explanation of many phenomena experimentally observed in enzymatic hydrolysis reactions.

The models of the biochemical engineering group of the University of Waterloo are closest to what is to be developed in this work. They are amongst two of the six models placing importance on substrate multiplicity. A third is a modification of the Waterloo model to allow for a different mode of action by the cellulases [10]. Discussion of their approach is found in the following section. The three other models that consider the cellulose substrate as non-uniform separate the cellulose into crystalline and amorphous sections. We now discuss some of the specific approaches to modelling the overall cellulose enzymatic hydrolysis reaction.

Huang [16] looks at the hydrolysis of insoluble amorphous cellulose. He treats it as a one-substrate, one-enzyme reaction. Since all enzymes bind and depolymerize the cellulose, we feel this is an oversimplification. The Langmuir adsorption isotherm is used in the model to describe the (one) enzyme binding to cellulose. Huang's three-parameter equation describes the progress curve of cellulose hydrolysis well. However, we must point out that the ratio of enzymes E_x, E_2 and E_1 in all runs was not changing. One enzyme mixture was simply diluted to change the enzyme concentration. Huang measured the reducing sugars and the glucose produced. The difference between the two is calculated as cellobiose. This assumption may be invalid if the concentration of cellodextrins in the reactor is significant.

Huang's hydrolysis model overestimates the sugar produced when the conversion is above 70%. The Langmuir equilibrium adsorption may not hold and the enzyme may be denaturing. Those are two reasons given by Huang to explain the overestimation.

In any case, we cannot use Huang's hydrolysis model as our own because it doesn't take into account substrate multiplicity, enzyme multiplicity and product multiplicity. We consider the latter to be important because, as discussed in the Introduction, one of our objectives is to predict G_1 and G_2 as the reaction progresses, not only the lumped reducing sugars. In Huang's estimation, the cellobiose/glucose ratio changes from 3:1 after 30 minutes of reaction to 1:1 after 24 hours.

Ohmine et al. [22] first attempted to fit a semi-mechanistic model to cellulose and cellobiose hydrolysis by combining equations from individual pieces of the kinetic model. When they individually accounted for substrate crystallinity, glucose inhibition of E_1 and cellobiose inhibition of E_x and E_2 they predicted each event successfully. However in combining these in an overall model they failed to predict the time course of the hydrolysis. Their eventual solution was an empirical model, not reflecting the true kinetics of the reaction. The empirical model contained a lumped rate retardation parameter.

The Kansas State University group also developed a comprehensive kinetic model by combining models from several aspects of the cellulose hydrolysis reaction [18]. They started by characterizing the substrate by its crystallinity and available surface area. The enzyme-substrate complex is formed by adsorption of the enzyme molecules onto the cellulose. The E_1 reaction does take place concurrently in the aqueous phase.

The principle assumptions made by Fan and Lee [18] in developing their model included:

1. The reaction is composed of two steps:
   - E_x and E_1 synergistically hydrolyze G_x to soluble G_2
   - E_1 hydrolyzes G_2 to G_1.
2. The bicompositional nature of cellulose is ignored i.e., no distinction is made between crystalline and amorphous regions.

3. Mass transfer is not rate-limiting.
4. Enzyme activity is proportional to the amount of adsorbed protein.
5. G_x hydrolysis occurs at the surface of cellulose.
6. G_2 and G_1 are non-competitive inhibitors of this reaction. The G_2 hydrolysis in the aqueous phase is competitively inhibited by G_1.
7. The deactivation of the enzyme and the increase in the fraction of cellulose which is crystalline are responsible for the decrease in the initial rate of hydrolysis.
8. The cellulose concentration is represented by the average bulk concentration.

Item 8 is of particular interest to us because the authors do mention the difficulty of expressing substrate concentration in molar terms unless such a concentration were expressed as the molar concentration of scissile bonds. This last item echoes the concern we have about the Fan and Lee model; its substrate definition is incomplete as far as its accessibility to the enzyme and its heterogeneous nature are concerned.

Otherwise the ten simultaneous equations of their model give a good representation of the cellulose time course hydrolysis. All 11 model parameters—also known as kinetic constants except one—have been evaluated individually. Some residual plots show some trends and of course the lack of parsimony in the parameters (11) and the failure to report confidence in the parameters are of concern to modellers.

Howell and Stuck [23] developed a model where the non-competitive inhibition by cellobiose dominates the kinetics. The main assumptions of their model are:

1. Substrate inhibition by excess cellulose (isolated theory) is negligible at their low operating concentrations.
2. Substrate inhibition by intermediates is negligible.
3. Multiplicity of substrate structure is ignored.
4. Product inhibition by G_1 and G_2 is significant.
5. G_2–G_6 are lumped into a cellobiose concentration term.

6. A single, unidentified reaction is rate-controlling.
7. Equilibrium of enzyme, substrate and products may be expressed by simple equilibrium constants.
8. Substrate concentration is represented by bulk concentration.

Item 3 and 8 have already been criticized as an ineffective characterization of the substrate. Item 5 can be criticized by engineers who want to keep a tab of all species in the reactor.

The model of Howell and Stuck is very limited in its applications. It is valid at relatively low yields or reaction times. At reaction times of 10 hours and at low substrate concentration, $3.4\,g/\mathrm{L}$, the model overpredicts the total carbohydrate yield. Most restrictive of all is its inability to predict the glucose concentration.

A review of many hydrolysis models, including some listed by Gilbert and Tsao [3] can be found in a paper by Lee et al. [43].

In their review, Gilbert and Tsao [3] list some conditions for the use of Michaelis-Menten kinetics in modelling chemical reactions:

1. the amount of substrate available can be expressed in terms of its molar concentration.
2. the substrate concentration greatly exceeds the enzyme concentration, and
3. when integrating the differential form, it is assumed that the concentrations of substrate and product together add up to the initial substrate concentration.

According to Gilbert and Tsao [3], these assumptions do not hold for enzymatic reaction on a solid substrate. The inapplicability of Michaelis-Menten kinetics to the solid substrate is repeated elsewhere [18]. This is not a concern to us as CMC is soluble. But the modelling approach in this work may show that a redefining of the nomenclature used in modelling the kinetics of enzymatic depolymerization can lead to the use of Michaelis-Menten kinetics for even a solid substrate. With

a proper model form and substrate definition, the assumptions above will not be violated.

2.9 Bond Hydrolysis Modelling: Support in the Literature

Many cellulase researchers have, at one time, directly and indirectly, said that the species of interest to us in the hydrolysis of cellulosics are the bonds linking the anhydroglucosidic units. After all, it is a bond that is cleaved by the enzyme. And in a substituted cellulose derivative, the nature of units on both sides of the bonds determines if the bond is eligible to be cleaved.

Our proposed model for CMC hydrolysis is the first that assumes bonds to be the species of interest in the reaction mixture. Let's now review some of the direct and indirect support in the literature for developing a model based on bond hydrolysis by cellulases.

To make the units of activity assays more palatable, Canevascini, in a viscometric endoglucanase test for E_x, transforms his arbitrary viscosity units into 'number of glycosidic linkages broken per unit time'. The idea is correct. The difficulty is that his calculation is not strictly mechanistic. Enzyme kinetics' modelling will need to be mechanistic.

Gilbert and Tsao [3] state that cellulose is a linear polymer of cellobiose. Attala [44] also suggests development of the view that cellobiose, not glucose, should be regarded as the basic repeat unit of cellulose and that we must realize that 2 anhydroglucose units are necessary to define the nature of the linkage between them. Furthermore, steric constraints associated with the β-1,4 linkage determine its reactivity. This is central to our work since a substituent group on an anhydroglucose unit changes the reactivity and the nature of the enzyme-bond complex.

In their review of the enzymatic hydrolysis of biomass, Ladisch et al. [11] state that the process is a function of substrate reactivity and enzyme activity.

"The two must be considered together if an accurate description of

> biomass saccharification is to be developed. To accomplish this, cellulolytic systems must first be modelled using realistic but well-defined substrates so that the optimal cellulolysis conditions can be formulated." [11]

This is indeed our foremost objective. The bond hydrolysis model, should correctly define the substrate to allow the determination of the true kinetics of enzymatic hydrolyses.

Of all the models derived from some mechanistic approach, the cellulose depolymerization kinetic model is our favorite. It has been developed by the biochemical engineering group at the University of Waterloo. In their model, the degree of polymerization of cellulose is most important. A chain splitting mechanism is assumed to follow Michaelis-Menten kinetics with inhibition. Then the model keeps track of the concentrations of the various chain lengths in the reactor. For example, with an initial DP of 1000 for cellulose, there are 1000 different species in the reactor, G_1, $G_2 \ldots G_{999}$, G_{1000}. The analysis by Suga et al. is completely theoretical [27]. In this first analysis, the exoenzyme is assumed to split off monomer units from the end of the chains whereas the endoglucanase is random-acting.

Their simulation results are expected; the exoenzyme alone in the reactor produces more monomer than the endoenzyme alone. Most important is that the rate of release of monomer for two enzymes together is higher than the sum of the rates of release for the two enzymes alone. They are on the correct model development path because their model shows synergism to be a direct consequence of the mechanism of the individual cellulases. One review paper [19] mentions that the Suga model [27] does not consider inhibition by products. The model's importance is in showing that synergism between two enzymes when hydrolyzing a uniform cellulose is an obvious result of even a simple hydrolysis mechanism.

A drawback to their model is that the CMC hydrolysis cannot be predicted on a molecular scale because of the presence of substituents and their effect on the

kinetics.

In a later version of their model [25], the presence of three enzymes—E_x, E_2, E_1—is taken into account. Furthermore the proposed mechanism of exoglucanase has been corrected by assuming that E_2 attacks the non-reducing end of the chain splitting off dimers. E_1 is assumed to attack only G_2, G_3 and G_4. E_x's action is random however its hydrolysis rate on chains of five units or less is assumed to be slower than for longer chains.

The slower hydrolysis of highly crystalline celluloses is blamed on the DP. Okazaki and Moo-Young [25] classify the DP of cellulosics as follows: 200–800 for CMC, 1000–2000 for wood pulp, 3000–5000 for cotton and 10000–14000 for α-cellulose. In our opinion, the DP of the cellulosic will contribute to its crystallinity, however there are other factors that influence crystallinity. The linking of the depolymerization resistance of the substrate solely to its DP is a drawback of their model. It does not take into account the differences between two celluloses of equal DP but different crystallinity. Besides, in our work we use CMC with a DP as high as 3200.

The versatility of the Waterloo approach is evident when one modifies the model form to account for a differing opinion on cellulase mechanism. This has been done by S.E. Lee et al. [10] to account for the separate action of a cellobiosylhydrolase and a cellobiase in the reactor. Of course, more parameters are added to the model.

Okazaki and Moo-Young state that their simulation agrees well with the shape of a CMC hydrolysis curve. However, many mathematical expressions will yield the proper curve shape. The true test of a CMC hydrolysis model lies in its ability to predict the glucose yield, the cellobiose yield and the DP of the solution at any time during the hydrolysis. But in all fairness, the authors admit that comparison to real data as well as taking into account the DS of the CMC is necessary.

Despite its drawbacks, the Waterloo model-development approach is, in our

opinion, the right one. The quantities of all species in the reactor must be accounted for at all times. With the introduction of different nomenclature an "accounting" model can be developed that places importance on both DP and crystallinity. In our work, crystallinity is not a parameter because the substrate, CMC is soluble. However, a method of treating crystallinity will later be discussed regarding the extension of our CMC hydrolysis model to insoluble cellulose hydrolysis.

2.10 Hydrolyses of Cellulose Ethers

Reese [15][17] in early studies of CMC, methocel (methyl cellulose), HEC, carboxymethyl hydroxyethyl cellulose (CMHEC) and sulfoethylcellulose hydrolyses showed the DS to be important in resistance to attack. In the earlier paper [15], Reese correctly concluded that the nominal DS of some HEC was wrong. It had yielded an innordinate amount of reducing sugars for an HEC of DS > 2. In fact its DS was less than 1.

The important points from Reese's papers are that the amount of glucose produced from enzymatic hydrolysis of soluble cellulose derivatives is a high function of the DS and that cellulose esters yield slightly more glucose than cellulose ethers. Reese however suggests that the presence of cellobiase esterase may be the reason for higher glucose from cellulose esters.

Savage, in a separate paper [36], and Reese note that more glucose is obtained from methocel than from the other ethers at equal DSs. We suspect that the distribution of the smaller substituent group on methocel is more random—thus leading to a higher fraction of unsubstituted anhydroglucose units—than for the bigger substituent groups of other ethers. Savage's CMC had a DS of 1.23. After 24 hours of hydrolysis the yield was 5.2 mg reducing sugar per g CMC. By comparison, hydrolysis of methylcellulose, DS 1.24 yielded 22.1 mg reducing sugars per g CMC. Both yields are low, as expected because of the high DS. No precision is given for

the yields.

A study of the CMCase activity test by Lindner et al. [30] may be considered more as a study of CMC hydrolysis. They examine three possible causes for the non-linearity of a curve of reducing sugars as a function of enzyme concentration:

- enzyme deactivation
- cryptic substrate limitation
- product inhibition

No enzyme deactivation was noticed after 30 minutes incubation at 50°C. Then the progress curves of hydrolysis of CMCs DS 0.4 and DS 0.7 were examined. Sure enough, more reducing sugar is obtained at the lower DS. The progress curve of the CMC of DS 0.4 being more linear, the authors conclude that cryptic substrate limitation is the cause of the non-linear dilution curve.

Product inhibition was tested by adding 1.1 g/L G_2. The corrected progress curve shows a 20% reduction in the rate of production of the reducing sugars from CMC. However, because the DS 0.7 curve is still non-linear and the DS 0.4 curve is linear, cryptic substrate limitation is the main cause of non-linearity.

A paper by Banerjee and Laidler [45] reports hydrolysis experiments on CMC. They found a good fit for a Michaelis-Menten model. However they did their experiments with only one enzyme mixture and one type of CMC. The arbitrary Michaelis-Menten model does not take into account the substrate multiplicity or the enzyme multiplicity. They would have noticed this had they tried CMCs of different DS and had they changed ratios of the individual enzyme components.

Poulsen and Petersen [13] verified that DS of CMC is an important influence to the ultimate amount of reducing sugars produced during its enzymatic hydrolysis. The DS varied from 0.4 to 1.2. Although the reducing sugar concentrations are expressed in absorbance units, it can be seen that 27 times as much reducing sugar

is obtained at a CMC DS of 0.4 than 1.2. Their theory, which adds little more information than Reese, suggests that substitution prevents endoglucanases from hydrolyzing bonds a 'certain' distance from the substituted anhydroglucose units.

Poulsen and Petersen [13] do contribute a model to predict the initial rate of CMC hydrolysis from a single time course hydrolysis data point. They use Huang's hydrolysis rate model [16]. At this point in time, one can be very critical of the model applied to CMC hydrolysis. The model is valid only for a DS of 0.7. It only holds true for an enzyme system containing no cellobiase—*Cellulomonas sp.* does not produce E_1. In fact the authors confirm the model fails upon addition of cellobiase. The initial estimate of substrate concentration is obtained from the product concentration at the reaction completion and the model parameters were estimated for one initial substrate concentration. Finally, even though the model was developed by assuming a one-enzyme, one-substrate system, it has four parameters. In all fairness to the authors, they were not in search of a versatile model. The objective of our work is to obtain an overall mechanistic model capable of predicting G_1 and G_2 as well as other fragment concentrations in the mixture under any starting enzymatic hydrolysis conditions.

In his discussion paper, Wood [8] states that there is no synergism between E_x and E_2 in hydrolyzing CMC. He also states that, when highly purified, E_2 shows little or no capacity to hydrolyze CMC. In interpreting the above two statements, one could say that the CMC hydrolysis rate will be the same function of E_x regardless of how much E_2 there is in the solution.

Kanda et al. [24] present some CMC hydrolysis curves. They keep track of the hydrolysis by following absorbance with time. This is unfortunate for us as we cannot find out how much sugar is produced. Nonetheless the curve is the typical CMC time course curve; the initial hydrolysis rate is high but then levels off after several hours—depending on the amount of enzyme. This curve shape was also

reported by Canevascini and Gattlen [26]. There is a problem here. Kanda states that the purified enzyme acting on CMC is exoglucanase. But there is little action of E_2 on CMC according to other sources. Kanda tried CMCs with three different DSs: 0.6, 0.75 and 1.45. The problem is that without the actual glucose or cellobiose concentrations instead of reported absorbance, we cannot conclude much from his data. It makes slightly more sense if, as Kanda claims all product is cellobiose but this must translate into an extremely minute quantity of cellobiose if indeed E_2 is an endwise acting enzyme. Finally there is a warning that he cannot be certain that the enzymes he purified are 100% pure.

In Zabriskie's work, initially the hydrolysis rate for CMC is 4.1 times faster than for solka floc but the hydrolysis reaction of CMC appears to have stopped after saccharifying only 23% of the substrate. [7]. We know that the cellulase enzyme system cannot degrade substituted polymers of cellulose completely.

Demeester [31] has studied hydrolysis of HEC. His first objective is to set up a test for assaying E_x. The DS of HEC is 1.3 thus it was of no use to measure the minute amounts of G_1 and G_2 expected. The most interesting plot in Demeester's paper is the progress of the weight-average molecular weight with time during the hydrolysis reaction. It levels off at a value well above 100,000 which shows that few bonds are cleaved in this HEC hydrolysis due to its high DS.

More important to us is that Demeester's paper shows that another response variable in our model can be followed, DP. We will not measure DP but by measuring it, along with G_1 and G_2, and also the reducing power in the reactor, more precise parameter estimation as well as more checks of the model are possible.

2.11 Summary of the Literature

A new orientation is needed in the analysis of cellulase kinetics. After 20 years of research, the theory is uncertain at best. From our search of the literature, it appears that only general observations can be gathered concerning the cellulase mechanism. The majority of papers accept three major enzymes; E_x acts randomly in cleaving bonds, E_2 produces G_2 from the non-reducing end of a chain and E_1 acts mainly on cellobiose to produce two glucose molecules and on the non-reducing end of a polymer to produce one glucose molecule.

These will be our general guidelines in developing a model. The multiplicity of each enzyme, we feel may be due to proteases attacking the parent enzyme. That is why a three enzyme mechanism should be sufficient to develop a kinetic model. Proteolysis kinetics is a project in itself.

Inhibition effects are not clear. It seems everything in the hydrolysis reaction mixture has been named as a competitive and a non-competitive inhibitor of all three cellulase components. This illustrates the unsatisfactory mathematical approach presently used in analyzing enzymatic hydrolysis of cellulosics.

None of the models in the literature was particularly appealing although the Waterloo model came closest to what we want to develop. Cellulase researchers keep insisting that it is bonds that are cleaved. Yet no mathematical model takes this into account. One of our contributions, we hope, is the theoretical approach to a sound analysis of cellulase kinetics.

Chapter 3

THEORY

In this chapter, a complete development of a structured mathematical model of the kinetics of the enzymatic hydrolysis of soluble cellulose derivatives is presented. The mathematical treatment of the polymer structure is first discussed. Important new nomenclature for the substrate is introduced and then used in a step-by-step building of a first 'basic' model for the hydrolysis kinetics. Finally, a number of alternate models are presented to show the flexibility of our mathematical treatment. Of course these alternate or rival models are necessary since the kinetics of the enzymatic hydrolysis reaction are not known *a priori*.

3.1 Distribution of the Substituent Groups on CMC

As discussed in the previous chapter, the hydrolysis kinetics depend upon the degree of substitution (DS) of the substrate, in particular the number of unsubstituted units. We now develop an equation relating the DS and "fraction of unsubstituted units" in CMC. Some investigators have conducted theoretical work [46] and experimental work [42][47] concerning the distribution of substituent groups on CMC. The nomenclature now introduced is consistent with their papers:

- c_0 is the fraction of units which are unsubstituted in the polymer.
- c_1 is the fraction of units substituted once.
- c_2 is the fraction of units substituted twice.
- c_3 is the fraction of units substituted in all three positions.

Hence $c_0 + c_1 + c_2 + c_3 = 1$, ignoring end effects or disturbances in the cellulose structure.

Let the selectivities Se_2, Se_3 and Se_6 be the probabilities that a carboxymethyl group reacts with a hydroxyl group in position 2, 3 or 6 respectively given that all three hydroxyl groups on an anhydroglucose unit are unsubstituted and that the carboxymethyl group must react i.e., $Se_2 + Se_3 + Se_6 = 1$. Another way to look at selectivity is that, given an unsubstituted cellulose molecule, Se_i is the probability that the first substituent to react with said cellulose, will bind to a hydroxyl group in position i.

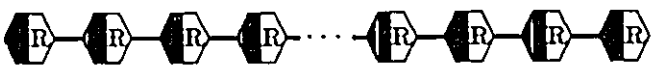
Selectivity nomenclature is introduced to facilitate the development of theory describing the distribution of substituent groups on substituted cellulose derivatives. We consider this approach easier to visualize than the approach in Spurlin's paper [46]. Moreover, the form of the final equation is simpler and more applicable to the purposes of this work.

Finally, x_2, x_3 and x_6 are the fractions of the anhydroglucose units that contain substituents in the 2, 3 and 6 positions respectively. The DS is related to the c's and the x's:

$$\mathrm{DS} \stackrel{\mathrm{def}}{=} c_1 + 2c_2 + 3c_3 \tag{4}$$

$$\mathrm{DS} \stackrel{\mathrm{def}}{=} x_2 + x_3 + x_6 \tag{5}$$

CMC is prepared by the reaction of alkali cellulose with sodium chloroacetate [48]. At the start of the cellulose etherification reaction, all of the anhydroglucose units are free of substituents.



The initial conditions:

$$\text{At DS=0} \begin{cases} c_0 = 1 \\ c_1 = c_2 = c_3 = 0 \\ x_2 = x_3 = x_6 = 0 \end{cases}$$

Consider N anhydroglucose units undergoing etherification. At a given time during

the reaction, the number of free hydroxyl groups at each position is:

$$\begin{cases} N(1-x_2) & \text{\# of free no. 2 hydroxyls} \\ N(1-x_3) & \text{\# of free no. 3 hydroxyls} \\ N(1-x_6) & \text{\# of free no. 6 hydroxyls} \end{cases}$$

If the binding procedure were completely random i.e., $Se_2 = Se_3 = Se_6 = \frac{1}{3}$ then the probability that a substituent attach itself to a given position would be proportional to the fraction of the free hydroxyls in that position.

However the selectivities or preferences of a substituent group for a position change the probabilities. Table 3 is a list of the total number of each type of hydroxyl groups in the N-unit polymer considered as the etherification reaction proceeds.

Table 3. *Hydroxyl Types and their Discrete Amounts in a CMC Solution*

Hydroxyl Type	Total Number
unsub. hydroxyl #2 on unsub. anhyd. unit	$N \cdot c_0$
unsub. hydroxyl #3 on unsub. anhyd. unit	$N \cdot c_0$
unsub. hydroxyl #6 on unsub. anhyd. unit	$N \cdot c_0$
unsub. hydroxyl #2 on sub. anhyd. unit	$N \cdot (1 - x_2 - c_0)$
unsub. hydroxyl #3 on sub. anhyd. unit	$N \cdot (1 - x_3 - c_0)$
unsub. hydroxyl #6 on sub. anhyd. unit	$N \cdot (1 - x_6 - c_0)$
sub. hydroxyl #2 on sub. anhyd. unit	$N \cdot x_2$
sub. hydroxyl #3 on sub. anhyd. unit	$N \cdot x_3$
sub. hydroxyl #6 on sub. anhyd. unit	$N \cdot x_6$
total hydroxyl positions	$3 \cdot N$

Of course there are no substituted hydroxyl groups on unsubstituted units.

To continue the development of the mathematical relationship between DS and c_0 we have to "build" a CMC molecule from a cellulose molecule. We now proceed with this probabilistic buildup by looking at the *events* that occur during this etherification reaction.

A_{02}: next substituent group binds to unsub. hydroxyl #2 on an unsub. unit

A_{03}: next substituent group binds to unsub. hydroxyl #3 on an unsub. unit

A_{06}: next substituent group binds to unsub. hydroxyl #6 on an unsub. unit

A_{12}: next substituent group binds to unsub. hydroxyl #2 on a sub. unit

A_{13}: next substituent group binds to unsub. hydroxyl #3 on a sub. unit

A_{16}: next substituent group binds to unsub. hydroxyl #6 on a sub. unit

Note that the next substituent group cannot bind to a hydroxyl position which is already substituted.

Let A be the event that the next substituent group binds to the CMC we are in the process of building. Since we are forcing the substituent group to attach to the CMC, P(A)=1.

$$\mathrm{P}(A) = \mathrm{P}(A_{02}) + \mathrm{P}(A_{03}) + \mathrm{P}(A_{06}) + \mathrm{P}(A_{12}) + \mathrm{P}(A_{13}) + \mathrm{P}(A_{16}) = 1 \qquad (6)$$

Now assume that for a given hydroxyl type, a substituent group does not discriminate whether that hydroxyl group is on an unsubstituted unit or on a unit with substituents present in the other positions. In other words, assume that the substitution of one hydroxyl in the position 6, for example, will not affect the reactivity of positions 2 or 3.

Hence the ratio of the probabilities of the next substituent group binding to a hydroxyl in position j of an unsubstituted unit as opposed to attaching itself to position j of a substituted unit is equal to the ratio of the amounts of each, i.e.,

$$\frac{\mathrm{P}(A_{0j})}{\mathrm{P}(A_{1j})} = \frac{\mathrm{N} \cdot c_0}{\mathrm{N} \cdot (1 - x_j - c_0)} \qquad (7)$$

$$\text{For the three hydroxyl positions, } \begin{cases} \mathrm{P}(A_{12}) = \mathrm{P}(A_{02}) \cdot (1 - x_2 - c_0)/c_0 \\ \mathrm{P}(A_{13}) = \mathrm{P}(A_{03}) \cdot (1 - x_3 - c_0)/c_0 \\ \mathrm{P}(A_{16}) = \mathrm{P}(A_{06}) \cdot (1 - x_6 - c_0)/c_0 \end{cases} \qquad (8)$$

Eliminating P(A_{12}), P(A_{13}) and P(A_{16}) from equation 6 yields

$$(1 - x_2)\,\mathrm{P}(A_{02}) + (1 - x_3)\,\mathrm{P}(A_{03}) + (1 - x_6)\,\mathrm{P}(A_{06}) = c_0 \qquad (9)$$

Now, we use the definition of selectivity. As listed in Table 3, the number of unsubstituted j hydroxyls on unsubstituted units is the same for all j (and is equal to c_0·N). Assuming that the ratio of reactivities of the hydroxyl groups is independent of the state of substitution of neighbouring hydroxyls or the DS of the chain as a whole, then the ratio of $P(A_{0i})$ to $P(A_{0j})$ is equal to the ratio of the selectivities of hydroxyl positions i and j. Savage [36] does mention some crystalline parts of cellulose are reluctant to undergo etherification. Hopefully, the presence of long unaltered residues of the original cellulose structure is negligible.

$$\frac{P(A_{02})}{P(A_{06})} = \frac{Se_2}{Se_6} \qquad \frac{P(A_{03})}{P(A_{06})} = \frac{Se_3}{Se_6} \tag{10}$$

...with the constraint $Se_2 + Se_3 + Se_6 = 1$.

Combining equations 9 and 10 to eliminate $P(A_{02})$ and $P(A_{03})$ gives

$$P(A_{06}) = \frac{Se_6 \cdot c_0}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

In the same way, the other probabilities can be calculated.

$$P(A_{02}) = \frac{Se_2 \cdot c_0}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

$$P(A_{03}) = \frac{Se_3 \cdot c_0}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

$$P(A_{12}) = \frac{Se_2 \cdot (1-x_2-c_0)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

$$P(A_{13}) = \frac{Se_3 \cdot (1-x_3-c_0)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

$$P(A_{16}) = \frac{Se_6 \cdot (1-x_6-c_0)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}$$

With these equations to calculate the probability that a given hydroxyl type will be the next recipient of a substituent group for all combinations of unsubstituted hydroxyls on substituted or unsubstituted units, we now proceed with the "building" of the CMC molecule.

Again, recall that N is the total number of anhydroglucose units considered in this building process. Let Δu be the incremental number of substituent units to be added. ΔDS is the incremental change in the degree of substitution due to the addition of Δu units to the polymer.

$$\Delta \mathrm{DS} \stackrel{\mathrm{def}}{=} \Delta u / \mathrm{N} \tag{11}$$

The number of units which have a substituent in the j hydroxyl position increases by the increment $\mathrm{P}(A_{0j}) + \mathrm{P}(A_{1j})$ whenever Δu units are substituted on the N anhydroglucose units considered. In other words, the fraction of the Δu substituent groups which will bind to, say position 6 is calculated as

$$\begin{aligned} \mathrm{N}\Delta x_6 &= \Delta u \left[\mathrm{P}(A_{06}) + \mathrm{P}(A_{16})\right] \\ &= \Delta u \left[\frac{Se_6(1-x_6)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}\right] \end{aligned}$$

Replace $\Delta u/\mathrm{N}$ by ΔDS to obtain

$$\Delta x_6 = \Delta \mathrm{DS} \left[\frac{Se_6(1-x_6)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}\right] \tag{12}$$

In the limit as the incremental change in the degree of substitution tends to zero, equation 12 becomes

$$d\mathrm{DS} = \left[\frac{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}{Se_6(1-x_6)}\right] dx_6 \tag{13}$$

Similarly,

$$d\mathrm{DS} = \left[\frac{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}{Se_2(1-x_2)}\right] dx_2 \tag{14}$$

$$d\mathrm{DS} = \left[\frac{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)}{Se_3(1-x_3)}\right] dx_3 \tag{15}$$

The variable c_0 is of considerable interest to us. As before, the change in the number of units which are unsubstituted is calculated from the number of Δu substituents which will bind to an unsubstituted unit.

$$\begin{aligned} \Delta(c_0 \cdot \mathrm{N}) &= -\Delta u \overbrace{\left[\mathrm{P}(A_{02}) + \mathrm{P}(A_{03}) + \mathrm{P}(A_{06})\right]}^{\text{probability of binding to unsub. unit}} \\ &= -\frac{\Delta u \cdot c_0 \cdot (Se_2 + Se_3 + Se_6)}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)} \end{aligned} \tag{16}$$

Recalling that $Se_2 + Se_3 + Se_6 = 1$ and rearranging,

$$\frac{\Delta c_0}{c_0} = -\frac{\Delta u}{\mathrm{N}} \left[\frac{1}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)} \right] \tag{17}$$

Substituting $\Delta \mathrm{DS} = \Delta u/\mathrm{N}$ and taking the limit as as the incremental change in DS tends to zero gives

$$\frac{dc_0}{c_0} = -\frac{d\mathrm{DS}}{Se_2(1-x_2) + Se_3(1-x_3) + Se_6(1-x_6)} \tag{18}$$

Comparing equations 13 and 18 to eliminate DS yields

$$\frac{dc_0}{c_0} = -\frac{dx_6}{Se_6(1-x_6)}$$

This form of equation is simple to integrate. The initial boundary condition is $x_6 = 0$ at $c_0 = 1$.

$$\int_1^{c_0} \frac{dc_0}{c_0} = \int_0^{x_6} \frac{-dx_6}{Se_6(1-x_6)} \tag{19}$$

$$\ln c_0 - \ln 1 = \frac{\ln(1-x_6)}{Se_6} - \frac{\ln(1-0)}{Se_6}$$

Two other similar equations can be developed for the other hydroxyls. The three are listed here after simplifications:

$$\begin{cases} x_2 = 1 - c_0^{Se_2} \\ x_3 = 1 - c_0^{Se_3} \\ x_6 = 1 - c_0^{Se_6} \end{cases} \tag{20}$$

Eliminating x_2, x_3 and x_6 from $\mathrm{DS} = x_2 + x_3 + x_6$.

$$\mathrm{DS} = (1 - c_0^{Se_2}) + (1 - c_0^{Se_3}) + (1 - c_0^{Se_6})$$

$$\mathrm{DS} = 3 - c_0^{Se_2} - c_0^{Se_3} - c_0^{Se_6} \tag{21}$$

One objective of this work has now been satisfied. An equation that allows the calculation of the fraction of anhydroglucose units which are unsubstituted given the DS of a substituted cellulose derivative has been developed. Our equation 21 has been compared with Spurlin's set of equations [46]. The predicted unsubstituted

fractions are the same. We do not reproduce Spurlin's complex equations here. Equation 21 is, in our opinion, easier to understand and certainly easier to work with.

Notwithstanding the two major assumptions concerning etherification, equation 21, as applied to our work, does have its inconveniences. In the first place, it is implicit in c_0. A trial and error routine is necessary to calculate c_0 from DS. In the second place, the selectivities of the carbon positions are not known a *priori*. There are three ways to deal with the latter problem. The selectivities may be simply assumed, estimated as parameters in the hydrolysis modelling or determined by independent chemical analysis.

3.1.1 Selectivities as Model Parameters

One way to determine the values of the selectivities is to include them along with the kinetic constants as parameters to be estimated in the rate equations describing the hydrolysis. This approach has the advantage of ensuring that the selectivities are those specifically related to the substrate being studied, CMC. However, there are several disadvantages. First, the addition of two more parameters increases the complexity of the already difficult estimation problem. (Only two selectivity parameters need be estimated as a result of the constraint $Se_2 + Se_3 + Se_6 = 1$). Second, it is doubtful that typical hydrolysis data contain sufficient information to obtain unambiguous estimates of the selectivities. For these reasons, this approach was not taken in the present study.

3.1.2 Assuming Equal Selectivities

Another approach is to assume, as did Spurlin [46], that the substituent groups attach in a totally random fashion, i.e., $Se_2 = Se_3 = Se_6 = 1/3$. In this case, equation 21 becomes

$$\mathrm{DS} = 3 - c_0^{\frac{1}{3}} - c_0^{\frac{1}{3}} - c_0^{\frac{1}{3}}$$

which may be solved explicitly for c_0 to give

$$c_0 = \left(1 - \frac{1}{3}\mathrm{DS}\right)^3 \tag{22}$$

Although this approach provides a very simple way to calculate c_0 for any DS, its validity is unlikely from both steric considerations and experimental evidence presented in the next section. This method of calculating c_0 will only be used in this work when we examine a hydrolysis model proposed by Reese.

3.1.3 Distribution of the Carboxymethyls Determined Experimentally

There has been some research on the distribution of the carboxymethyl groups on CMC. Timell [42][47] conducted this research. In his studies of CMC with a DS varying from 0.45 to 1.11, he found that the hydroxyl group in the 6^{th} position is the most reactive and that position 3 is least reactive. He also noted that no anhydroglucose units had both positions 2 and 3 substituted. From his experimental data, Timell [47] concluded that a (negatively charged) carboxymethyl group in either of the two secondary positions prevents the adjacent secondary hydroxyl group from being converted into a substituted position. In addition, Timell also concluded that the rate at which substituent groups attach onto position 6 is approximately equal to the sum of the rates of attachment onto positions 2 and 3. Consequently, Timell developed the following equation relating c_0 and DS:

$$c_0 = \left(1 - \frac{1}{2}\mathrm{DS}\right)^2 \tag{23}$$

Note that Timell's experimental work shows that one of the assumptions made in our development of the distribution equation does not hold; the substitution of the

carboxymethyl group onto a hydroxyl—especially the secondary positions—is in fact dependent on the state of substitution of adjacent hydroxyls.

Extrapolating the observations brings up an interesting point. At a DS of 2, all units on the CMC will have two substituent groups attached. Timell discusses the preparation of CMC. Above a DS of 1.5, the preparation of CMC is very difficult, but he does report that carboxymethyl cellulose with a DS as high as 2.8 has been synthesized. However, 23 treatment steps were necessary to achieve such a high etherification. Since high etherification is no problem during methyl cellulose preparation, and since the distributions of CMC and methyl cellulose are different, it is suggested that as more and more negatively charged carboxymethyl groups are introduced on the polymer, interference becomes noticeable between neighbouring anhydroglucose units. This would violate another assumption in our distribution assumption. Fortunately, at the DS of interest to us ($0 \leq \mathrm{DS} \leq 1$), such interference is negligible [47].

Equation 23 is simple and it is based on experimental data. It is the one to be used in our modelling of the kinetics of the enzymatic hydrolysis of CMC.

3.2 Nomenclature and Schematic Diagrams of Relevant Bond Types

To describe the hydrolysis of CMC in terms of the formation of an enzyme-bond complex followed by cleavage of the β-1,4 bond, it becomes necessary to distinguish between the various bond types in the reaction mixture. The molar concentrations of these bond types become the variables in a model representing the mechanism behind the enzymatic hydrolysis of carboxymethyl cellulose by cellulases. The various types of bonds are summarized in Figure 2 in terms of the symbols and nomenclature introduced earlier.

– glucose G_1

substituted glucose

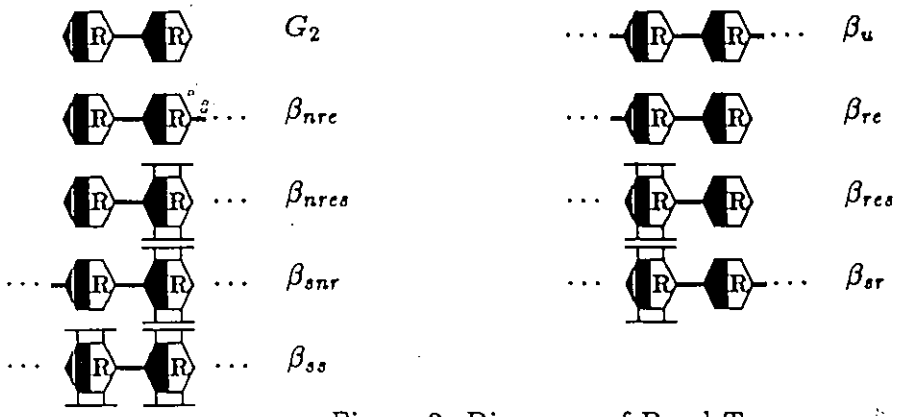


Figure 2: Diagrams of Bond Types

All the bond species illustrated herewith are mutually exclusive. The whole range of possible bond types is covered. For example a cellobiose molecule with substituent groups attached on both glucosidic units would in fact be a β_{ss} type bond.

It is important to observe that the continuation of the polymer chain beyond the next glucosidic unit may or may not be relavant. The diagrams take this into account. For example in case (a) illustrated here, the polymer is the dimer

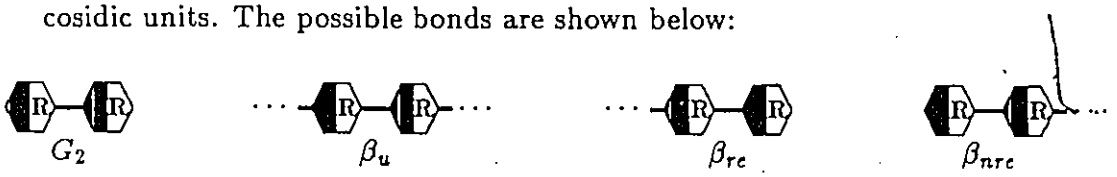
consisting of an unsubstituted non-reducing end unit and a substituted reducing end unit. In case (b), the polymer may or may not continue beyond the reducing side unit. In case (c), the polymer chain must be more than two units long for the discussion to be valid. In both cases (b) and (c), the type of unit in the third position from the non-reducing end is unimportant. But references to a diagram such as (c) in the text imply that there is a third unit in the chain.

3.3 Model 1 Development

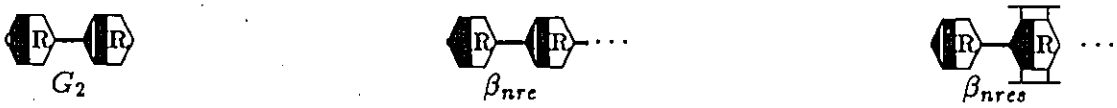
3.3.1 Assumptions and Basic Kinetic Relationships

The form of the enzymatic hydrolysis model for CMC will depend on the assumptions made in its development. A list of the assumptions made in developing the basic model follows.

1. The most important assumption is that, in solution, eligible β-1,4 bonds and the free-enzyme come together to form a complex leading to cleavage of the bond.
2. Two enzymes have an influence on the rate of reaction: endoglucanase and cellobiase. The so-called exoglucanase does not affect the kinetics of the reaction.
3. Endoglucanase cleaves any β-1,4 bond between two unsubstituted anhydroglucosidic units. The possible bonds are shown below:



4. Cellobiase cleaves β-1,4 bonds at the non-reducing end of a molecule where this end unit is unsubstituted. The possible bonds are shown below:



5. The dissociation rate of an enzyme-bond complex is not a function of the type of bond.
6. The system is free of inhibition i.e., the enzyme does not form a complex with products or non-substrate bonds. The literature does not support this but it is nonetheless a convenient simplification in this first-to-be-developed model.
7. Enzyme deactivation is negligible at 50°C and pH 4.8.
8. Michaelis-Menten kinetics apply to the system.

Based on these assumptions, some basic kinetic relationships can be developed.

Keep in mind that the formation of an enzyme-bond complex is the underlying assumption in this development.

Assumptions 3, 5 and 8 lead to an expression for the rate of cleavage of β_u, G_2, β_{re} and β_{nre} bonds by the endoglucanase, E_x:

$$-\frac{d(\beta_u + G_2 + \beta_{re} + \beta_{nre})}{dt} = \frac{E_x \cdot k_{1E_x} \cdot (\beta_u + G_2 + \beta_{re} + \beta_{nre})}{K_{mE_x} + \beta_u + G_2 + \beta_{re} + \beta_{nre}} \tag{24}$$

Equation 24 is a modified Michaelis-Menten expression. To account for the multiple substrates, the usual substrate term has been replaced by the total concentration of the endoglucanase substrates, G_2, β_u, β_{re} and β_{nre}.

Assumption 5—equal dissociation rate of the endoglucanase complex regardless of substrate—implies that the substrate type is transparent to the enzyme. Hence the rate of cleavage of one substrate type by E_x can be calculated if the ratio of this substrate to the total (endoglucanase) substrate is known. For example, the rate of cleavage of cellobiose, G_2, by E_x is

$$-\frac{dG_2}{dt} = \underbrace{\left[\frac{G_2}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]}_{\text{frac of } E_x \text{ substrate which is } G_2} \underbrace{\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]}_{\text{rate at which } E_x \text{ cleaves all substrates}} \tag{25}$$

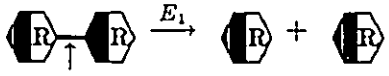
Assumption 4 lists G_2, β_{nre} and β_{nres} as the substrates of cellobiase, E_1. Again the aggregate rate of hydrolysis of these bond types by E_1 can be written as

$$-\frac{d(G_2 + \beta_{nre} + \beta_{nres})}{dt} = \frac{E_1 \cdot k_{1E_1} \cdot (G_2 + \beta_{nre} + \beta_{nres})}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \tag{26}$$

This aggregate rate can be decomposed into a sum of the rates of hydrolysis of the individual substrate types to yield

$$\begin{aligned} -\frac{d(G_2 + \beta_{nre} + \beta_{nres})}{dt} = & \frac{E_1 \cdot k_{1E_1} \cdot G_2}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \\ & + \frac{E_1 \cdot k_{1E_1} \cdot \beta_{nre}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \\ & + \frac{E_1 \cdot k_{1E_1} \cdot \beta_{nres}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \end{aligned} \tag{27}$$

where, for example, the term $E_1 \cdot k_{1E_1} \cdot G_2/(K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres})$ is the rate of cleavage of cellobiose by cellobiase. The two enzymes, E_x and E_1, are acting



concurrently on different G_2 molecules in the reactor by first forming a complex with G_2 and then cleaving its β-1,4 bond. To calculate the overall rate of cellobiose hydrolysis, add the rates for the two enzymes acting on cellobiose.

$$\frac{dG_2}{dt} = -\frac{E_x \cdot k_{1E_x} \cdot G_2}{K_{mE_x} + \beta_u + G_2 + \beta_{re} + \beta_{nre}} - \frac{E_1 \cdot k_{1E_1} \cdot G_2}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \tag{28}$$

Equation 28 is a measure of the cleavage rate of cellobiose by E_x and E_1. It does not necessarily give a measure of the rate of change of cellobiose with time since G_2 can be formed in the hydrolysis of a cellulosic polymer.

3.3.2 Rate Model for G_1

The quantity of glucose in the reactor is strictly increasing as a function of time. Two glucose molecules are produced when a cellobiose molecule is hydrolyzed. The endoglucanase and the cellobiase are both capable of cleaving the β-1,4 glucosidic bond of a cellobiose molecule. Also, with the assumptions made in developing this model, one glucose molecule is produced when a β_{re}, a β_{nre} or a β_{nres} bond is cleaved. Then,

$$\begin{aligned}
\frac{dG_1}{dt} = & \left[\frac{2 \cdot G_2}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (\beta_u + G_2 + \beta_{re} + \beta_{nre})}{K_{mE_x} + \beta_u + G_2 + \beta_{re} + \beta_{nre}}\right] \\
& + \left[\frac{2 \cdot G_2}{G_2 + \beta_{nre} + \beta_{nres}}\right]\left[\frac{E_1 \cdot k_{1E_1} \cdot (G_2 + \beta_{nre} + \beta_{nres})}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right] \\
& + \left[\frac{\beta_{re} + \beta_{nre}}{\beta_u + G_2 + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (\beta_u + G_2 + \beta_{re} + \beta_{nre})}{K_{mE_x} + \beta_u + G_2 + \beta_{re} + \beta_{nre}}\right] \\
& + \left[\frac{\beta_{nre} + \beta_{nres}}{G_2 + \beta_{nre} + \beta_{nres}}\right]\left[\frac{E_1 \cdot k_{1E_1} \cdot (G_2 + \beta_{nre} + \beta_{nres})}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right]
\end{aligned} \tag{29}$$

The first and second lines of equation 29 are the rates of production of glucose from cellobiose hydrolysis by E_x and E_1 respectively. The third line is the rate of

cleavage of β_{re} and β_{nre} bonds by E_x. The fourth line is the rate of cleavage of β_{nre} and β_{nres} by E_1. Equation 29 simplifies to

$$\frac{dG_1}{dt} = \frac{(2G_2 + \beta_{re} + \beta_{nre}) \cdot E_x \cdot k_{1E_x}}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}} + \frac{(2G_2 + \beta_{nre} + \beta_{nres}) \cdot E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}} \qquad (30)$$

The rate of glucose production in a hydrolysis reaction can thus be calculated. The concentration of all the bond types has to be known. Direct experimental determination of these bond concentrations would be one method of calculating the glucose concentration. At present however, analytical techniques are not advanced enough to estimate β_u, β_{re}, β_{nre} and β_{nres} experimentally.

3.3.3 Rate Model for G_2

The expression for the rate of cleavage of cellobiose by E_1 and E_x has already been calculated. Cellobiose concentration in the reactor can also increase. This happens as a result of the cleavage of a bond adjacent to an end bond.

The molecule from which the cellobiose will separate must have a molecular length greater than or equal to 3 anhydroglucose units and the two end units that, after cleavage, form the cellobiose entity have to be unsubstituted:

Both β_{re} and β_{nre} bonds are transformed into a G_2 bond when their adjacent bonds are cleaved. These two cases will now be considered.

Only three types of bonds can be adjacent to a β_{nre} bond: β_u, β_{onr} and β_{re}.

The rate of cleavage of β_u bonds is given by

$$-\left[\frac{\beta_u}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]$$

The probability that β_{nre} is the bond on the non-reducing side of a β_u bond is $\beta_{nre}/(\beta_u + \beta_{nre} + \beta_{sr})$. Thus the rate at which β_{nre} bonds are transformed into G_2 bonds because an adjacent β_u bond was cleaved is

$$\left[\frac{\beta_{nre}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{\beta_u}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]$$

A β_{snr} bond, the other type of bond that can be adjacent to a β_{nre} bond cannot be cleaved by the enzyme system.

The rate of cleavage of β_{re} bonds is given by

$$-\left[\frac{\beta_{re}}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]$$

The probability that β_{nre} is the bond on the non-reducing side of a β_{re} bond is $\beta_{nre}/(\beta_u + \beta_{nre} + \beta_{sr})$. This is actually the fraction of β_{nre} bonds to the total number of possible bonds adjacent to β_{re}. The rate of production of G_2 due to cleavage of β_{re} bonds adjacent to β_{nre} bonds is

$$\left[\frac{\beta_{nre}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]$$

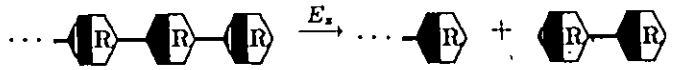
When a β_{nre} and a β_{re} bond are contiguous, they form a cellotriose molecule. We have just considered the action of cleaving β_{re} which releases G_2 into solution. But the β_{nre} bond of the cellotriose can also be cleaved. This also releases a G_2 bond into solution. Furthermore both E_1 and E_x can cleave β_{nre} type bonds. By direct cleavage,

$$\begin{aligned}\frac{d\beta_{nre}}{dt} = &-\left[\frac{\beta_{nre}}{G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{re} + \beta_{nre})}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right] \\ &-\left[\frac{\beta_{nre}}{G_2 + \beta_{nre} + \beta_{nres}}\right]\left[\frac{E_1 \cdot k_{1E_1} \cdot (G_2 + \beta_{nre} + \beta_{nres})}{K_{mE_x} + G_2 + \beta_{nre} + \beta_{nres}}\right]\end{aligned} \tag{31}$$

And the probability that β_{re} is the bond adjacent to a β_{nre} bond is $\beta_{re}/(\beta_u + \beta_{re} + \beta_{snr})$. Multiply equation 31 by the probability of finding a contiguous β_{re} to give the rate at which G_2 is produced by the cleavage of β_{nre} in cellotriose.

$$\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\left(\frac{E_x\cdot k_{1E_x}\cdot\beta_{nre}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right)+\left(\frac{E_1\cdot k_{1E_1}\cdot\beta_{nre}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right)\right]$$

The final way in which the concentration of cellobiose in the reactor can increase



is by the cleavage of a β_u bond adjacent to a β_{re} bond. The rate at which G_2 increases through the transformation of β_{re} into G_2 by the cleavage of an adjacent β_u is

$$\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

Combine the expressions for this rate of production of G_2 through transformations and for the rate of depletion of G_2 by direct cleavage. The resulting equation is the relation for calculating the rate at which cellobiose concentration is changing in the reactor.

$$\begin{aligned}\frac{dG_2}{dt} = &\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\\
&+\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{G_2}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{G_2}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned}$$

Simplifications are possible! It all reduces to:

$$\begin{aligned}\frac{dG_2}{dt} = &\left[\frac{\beta_u\cdot\beta_{re}+\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}+\frac{\beta_u\cdot\beta_{nre}+\beta_{re}\cdot\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}-G_2\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}-G_2\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned}\qquad(32)$$

Now the question may be asked: "How about the quantities that are not experimentally measured..., the β's?". If a complete system is to be defined, a balance must be made on each of those species. We now proceed to develop the pertinent equations.

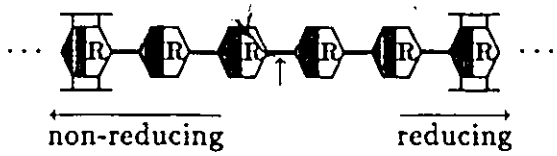
3.3.4 Rate Model for β_u

β_u type bonds are strictly depleted during a reaction. In the first instance, endoglucanases act to cleave these bonds directly. In the second instance, β_u bonds are transformed into one of the end-type bonds—β_{re} or β_{nre}—when an adjacent bond is cleaved. This is represented mathematically as follows:

$$\begin{aligned}
\frac{d\beta_u}{dt} = &- \left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \\
&- \left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \\
&- \left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \\
&- \left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \\
&- \left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \\
&- \left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nrea}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nrea})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nrea}}\right]
\end{aligned} \tag{33}$$

The first line of equation 33 is the direct cleavage rate of β_u bonds by E_x.

The direct cleavage of a β_u bond further reduces the concentration of β_u if any of the adjacent bonds are also of type β_u.



On the reducing side of a β_u bond, we can find either another β_u bond or a β_{re} or a β_{snr} bond. The probability of finding β_u is given by $\beta_u/(\beta_u+\beta_{re}+\beta_{snr})$. Multiply this probability by the rate of cleavage of β_u to obtain the rate at which

reducing-side β_u bonds—adjacent to cleaved β_u bonds—are transformed into β_{nrc} type bonds.

$$\left[\frac{\beta_u}{\beta_u + \beta_{rc} + \beta_{snr}}\right]\left[\frac{\beta_u}{G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{rc} + \beta_{nrc})}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]$$

This is the second line in equation 33.

On the non-reducing side of a cleaved β_u bond, there may also have been a β_u bond. The probability prior to cleavage was $\beta_u/(\beta_u + \beta_{nrc} + \beta_{sr})$. The rate at which non-reducing side β_u bonds—adjacent to cleaved β_u bonds—are transformed into β_{rc} type bonds is

$$\left[\frac{\beta_u}{\beta_u + \beta_{nrc} + \beta_{sr}}\right]\left[\frac{\beta_u}{G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{rc} + \beta_{nrc})}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]$$

This is the third line in equation 33.

If there are three consecutive unsubstituted units at the reducing end of a polymer that is at least 4 units long, then the cleavage of the β_{rc} bond results in the transformation of the β_u bond into a β_{rc} bond. The rate of cleavage of β_{rc} bonds is

$$\left[\frac{\beta_{rc}}{G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{rc} + \beta_{nrc})}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]$$

And $\beta_u/(\beta_u + \beta_{nrc} + \beta_{sr})$ is the probability that it was indeed a β_u besides the cleaved β_{rc}. The fourth line in equation 33 is the rate of transformation of β_u bonds into β_{rc} bonds due to cleavage of adjacent β_{rc} bonds.

To complete the analysis on how β_u changes with time in an enzymatic hydrolysis reaction, we calculate the rate at which β_u bonds are transformed into β_{nrc} bonds through cleavage of β_{nrc} bonds. Three consecutive unsubstituted anhydroglucose units at the non-reducing end of the substituted cellulose are considered. Two enzymes cleave β_{nrc}. The overall rate

of direct cleavage is given by the sum of the two individual rates of E_1 and E_x.

$$\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$+\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

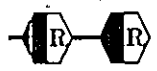
To obtain the rate of depletion of β_u through transformation due to its being adjacent to a cleaved β_{nre} bond, multiply the above overall rate by the probability of finding a β_u adjacent to a β_{nre} type bond, $\beta_u/(\beta_u+\beta_{re}+\beta_{snr})$.

Equation 33 is now complete. It simplifies to

$$\frac{d\beta_u}{dt}=-\left[\frac{\beta_u+\beta_{nre}}{\beta_u+\beta_{re}+\beta_{snr}}+\frac{\beta_u+\beta_{re}}{\beta_u+\beta_{nre}+\beta_{sr}}+1\right][\beta_u]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$-\left[\frac{\beta_u\cdot\beta_{nre}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{34}$$

Equation 34 describes how β_u changes with time in the reactor.

3.3.5 Rate Model for β_{re}



At the reducing end of the cellulose or cellulose derivative, there may be two units which are unsubstituted. According to the nomenclature set up in this work, the β-1,4 glucosidic bond separating these two units is appropriately referred to as β_{re}. Specifically β_{re} is the concentration of these bonds in the reaction mixture, in units of moles/L. Let's set up the differential equation governing the rate of change with time of β_{re} one piece at a time.

The overall rate of reaction of the E_x enzyme cleaving its substrate is calculated as:

$$\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+(G_2+\beta_u+\beta_{re}+\beta_{nre})}$$

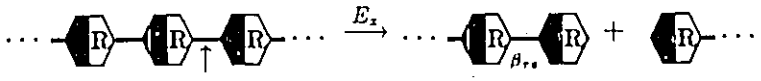
This is a simple Michaelis-Menten equation with rate constant for the E_x enzyme k_{1E_x} and its Michaelis constant K_{mE_x}. For this overall hydrolysis model, one of the listed assumptions concerns the substrates available to the E_x enzyme. They are G_2, β_u, β_{re} and β_{nres}. Those are the quantities which have been used as the

enzyme substrate in the numerator and the denominator of the Michaelis-Menten equation. Also note that there is no product inhibition term in the Michaelis-Menten equation to remain faithful to the assumption of no product inhibition for this particular model.

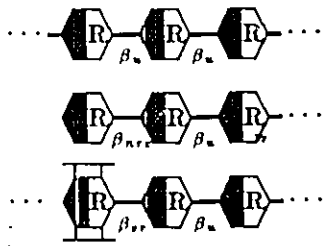
So we have the rate at which E_x cleaves all its available bond substrate. Multiplying the overall E_x rate equation by the ratio of β_u to all substrate available gives the direct rate at which the β_u type bonds are being depleted by the enzyme E_x.

$$\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

β_{re} bonds are created when a β_u bond is cleaved if the bond on the non-reducing side of the β_u bond considered prior to cleavage was also a β_u bond. In other words, if there are two contiguous β_u bonds in a CMC molecule and the one closer to the reducing side is cleaved, then the uncleaved β_u bond has now been transformed into a β_{re} type bond.



To the expression of the rate of cleavage of β_u, add the constraint that the bond next to the β_u bond considered, on the non-reducing side must also be type β_u. Three possible bond types can be adjacent to a β_u bond, on its non-reducing side.



The probability that the adjacent bond considered is indeed β_u type is given as $\beta_u/(\beta_u+\beta_{nre}+\beta_{sr})$. Now combine the likelihood that the adjacent non-reducing side bond is β_u with the rate at which the β_u type bonds are being cleaved to obtain the rate of formation of β_{re} bonds from the cleavage of β_u bonds.

$$\left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

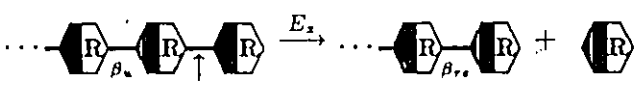
That's only the first part of the overall equation to determine the rate of change of β_{re}.

Another way in which a β_{re} bond is created is when a β_{re} bond is cleaved provided that the adjacent bond was a β_u bond. What types of bond are permitted to be adjacent to β_{re} bonds? β_u (presently examined), β_{nre} (in which case we are looking at a cellotriose molecule) or β_{sr} (if the next anhydroglucose unit considered is substituted). The rate at which the β_{re} bonds are being cleaved equals:

$$\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

And multiply by the probability that the adjacent bond is type β_u to obtain the rate of formation of β_{re} from cleavage of β_{re} contiguous to β_u,

$$\left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$



This last case has been considered despite the fact that the net gain of β_{re} bonds is nil since one must be cleaved to obtain another. Recall the direct cleavage rate of β_{re} bonds:

$$\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

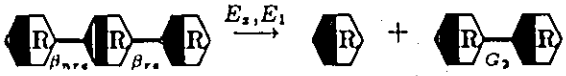
The last two equations can be interpreted in a different way. On cleavage of a β_{nre} bond, there is a net loss of one β_{re} bond if the bond adjacent to it was either a β_{nre} or a β_{sr} bond.

Not all bond reactions that affect the rate of change of β_{re} with time have been covered. One β_{re} bond is lost if a β_u bond is cleaved if the bond next to it was a β_{re} bond. This is the reversed case from above with the same two adjacent bond position considered but a different order of cleavage. The probability that the bond next to the β_u bond on the reducing side is a β_{re} bond is given by $\beta_{re}/(\beta_u+\beta_{re}+\beta_{snr})$.

This has been set at the ratio of β_{re} to the eligible bond types for the position considered, the reducing side of β_u. Thus the rate of change of β_{re} with time due to the cleavage of a β_u bond next to a β_{re} bond is given by

$$-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

There is only one more type of bond reaction which affects the rate of change of β_{re} with time. When a β_{re} bond is adjacent to a β_{nre} bond and that β_{nre} bond is cleaved, the β_{re} bond is transformed into a G_2 bond. In other words, the bond



closest to the non-reducing end in a cellotriose molecule has been cleaved to yield one molecule of glucose and one of cellobiose. As before, the probability that the bond next to a β_{nre} bond is of β_{re} type is calculated as $\beta_{re}/(\beta_u+\beta_{re}+\beta_{snr})$. Now we have to account for the fact that in this model, both the endoglucanase and cellobiase are assumed to cleave β_{nre} bonds. The instantaneous rate at which one enzyme cleaves β_{nre} is considered independent of the other enzyme rate. The total rate of cleavage is assumed additive for enzyme type. And since the assumptions for this model include G_2, β_{nre} and β_{nres} as substrate for the E_1 enzyme, the cleavage rate of β_{nre} by the two enzymes is

$$\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$+\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

where the first line is the effect by E_x and the second line by E_1. Then the rate of change of β_{re} due to the cleavage of β_{nre} bonds is

$$-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

All the bond reactions that affect the rate of change of β_{re} with time have been discussed. Assemble together all the parts of the equation.

$$
\begin{aligned}
&\left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_u}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]
\end{aligned}
$$

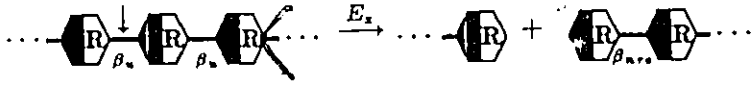
It all simplifies to

$$
\begin{aligned}
\frac{d\beta_{re}}{dt} = &\left[\frac{\beta_u\cdot(\beta_u+\beta_{re})}{\beta_u+\beta_{nre}+\beta_{sr}}-\beta_{re}-\frac{\beta_{re}\cdot(\beta_u+\beta_{nre})}{\beta_{re}+\beta_u+\beta_{snr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]
\end{aligned}
\tag{35}
$$

3.3.6 Rate Model for β_{nre}

Action on β_{nre} is very similar to that on β_{re} The difference is that we are now considering a bond between two unsubstituted anhydroglucose units at the non-reducing end of the cellulose derivative.

A β_{nre} bond is created when a β_u bond is cleaved if the adjacent bond to that being cleaved, said adjacent bond being on the reducing side of the cleaved β_u bond,



is also a β_u bond. The rate at which β_u bonds are cleaved is given by

$$
-\frac{d\beta_u}{dt} = \left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]
$$

And the probability that the bond on the reducing side of the β_u bond being cleaved is another β_u bond is given by the ratio of β_u bonds in the reactor to the total

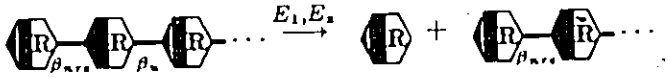
number of bonds which are eligible to be positioned on the reducing side of a β_u bond. Those eligible types of bonds are β_u, β_{re} and β_{snr}. Then the rate at which β_{re} bonds are produced because a β_u bond adjacent to another β_u bond has been cleaved is given by

$$\underbrace{\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]}_{\text{prob. of contiguous } \beta_u}\underbrace{\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]}_{\text{rate of cleavage of } \beta_u \text{ bonds}}$$

The rate at which β_{nre} bonds are directly cleaved after dissociation of the enzyme-β_{nre} complex is the sum of the action of the two enzymes on the β_{nre} bond. Recall here that it is assumed in the development of this model that non-reducing end bonds between two unsubstited anhydroglucose units are an eligible substrate of both cellobiase and endo-glucanase.

$$\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$+\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

This is the total rate of cleavage of β_{nre} bonds. However, if there was a β_u bond contiguous to the β_{nre} bond cleaved then the net effect on the number of β_{nre} bonds cleaved is nil because that β_u bond has in turn been transformed into a β_{nre} type bond.

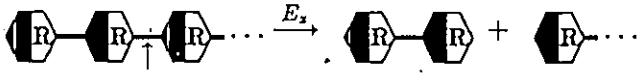


Again the probability that there was a β_u type bond adjacent to a β_{nre} type bond is calculated as the ratio of the quantity of β_u bonds to the quantity of all eligible bonds to be adjacent to a β_{nre} bond i.e., $\beta_u + \beta_{re} + \beta_{snr}$. The rate at which the enzymes cleave the β_{nre} bonds has already been accounted for thus this is a positive correction for the presence of some β_u bonds adjacent to cleaved β_{nre}

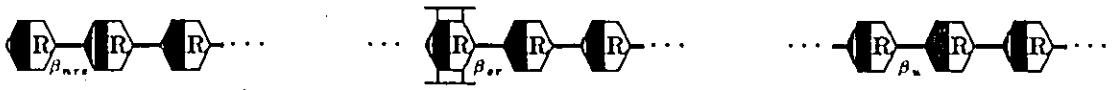
bonds.

$$+\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\cdot\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$+\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\cdot\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

One β_{nre} bond is transformed into a cellobiose bond if an adjacent β_u bond is



cleaved. On the non-reducing side, the only types of bonds that can be adjacent to a β_u bond are:



non-reducing $\Longleftarrow$ reducing

This rate of transformation of β_{nre} is calculated as the rate of cleavage of β_u bonds multiplied by the probability that a β_{nre} bond was adjacent.

$$-\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

where you will recognize the first (probability) term as the ratio of the number of β_{nre} bonds to the number of bonds in the mixture which are eligible to be adjacent to a β_u bond on its non-reducing side.

The last bond cleavage that affects the rate of change of β_{nre} with time is a β_{nre} bond transformed into a cellobiose bond from its position as a non-reducing end bond in a cellotriose molecule when the reducing end bond, β_{re}, in that cellotriose molecule is cleaved. The possible types of bond that are found on the non-reducing side of a β_{re} bond are β_u, β_{nre} or β_{sr}. Then the rate of change of β_{nre} due to this transformation is

$$-\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

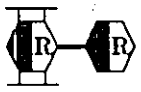
In the same order as the possibilities have been enumerated in this section, here is the overall equation for the rate of change of β_{nre} with time:

$$\begin{aligned}\frac{d\beta_{nre}}{dt} = &\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]\\
&+\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_u}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]\\
&-\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&-\left[\frac{\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\end{aligned} \tag{36}$$

Grouping of terms gives

$$\begin{aligned}\frac{d\beta_{nre}}{dt} = &\left[\frac{\beta_u\cdot(\beta_u+\beta_{nre})}{\beta_u+\beta_{re}+\beta_{snr}}-\beta_{nre}-\frac{\beta_{nre}\cdot(\beta_u+\beta_{re})}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\
&+\left[\frac{\beta_u\cdot\beta_{nre}}{\beta_u+\beta_{re}+\beta_{snr}}-\beta_{nre}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned} \tag{37}$$

3.3.7 Rate Model for β_{res}



Another end type bond considered in the development of this hydrolysis model is β_{res}. This is found at the reducing end of the polymer which may be as small as 2 units. The condition is that the reducing end unit be unsubstituted and the next unit be substituted.

A β_{res} bond is not cleaved by either E_1 or E_x. This assumption applies to the basic model presently considered. A β_{res} bond cannot be transformed into another type of bond because of the cleavage of a contiguous bond. There are no bonds on the reducing side, i.e., it is a reducing *end* bond. And according to the nomenclature set up in this work, a dimer of substituted cellulose is a β_{res} bond if only the reducing end unit is unsubstituted. Thus for this model's assumed

mechanism, the concentration β_{res} is strictly increasing. Or to put it in another way, the rate of change of β_{res} with time is positive throughout the reaction.

There are two ways in which a β_{res} bond is created. Either a β_u bond contiguous to a β_{sr} bond is cleaved or a β_{rc} bond contiguous to a β_{sr} bond is cleaved. In

either case, the β_{sr} bond is transformed into a β_{res} bond. Note also that it is not necessary to specify that β_u and β_{rc} are on the reducing side of β_{res}. They can be positioned only on the reducing side because the anhydroglucosidic unit of β_{res} closest to the molecule's non-reducing end is substituted. Since only E_x can cleave β_u and β_{rc} type bonds, the equation relating the rate of change of β_{res} with time includes two terms.

$$\frac{d\beta_{res}}{dt} = \left[\frac{\beta_{sr}}{\beta_u + \beta_{nrc} + \beta_{sr}}\right]\left[\frac{\beta_u}{G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{rc} + \beta_{nrc})}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right] + \left[\frac{\beta_{sr}}{\beta_u + \beta_{nrc} + \beta_{sr}}\right]\left[\frac{\beta_{rc}}{G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right]\left[\frac{E_x \cdot k_{1E_x} \cdot (G_2 + \beta_u + \beta_{rc} + \beta_{nrc})}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right] \quad (38)$$

where again the first term in each line of the equation represents the probability of having a β_{sr} type bond adjacent to a β_u or a β_{rc} bond. The second and third terms of each line taken together are the rates of cleavage of β_u and β_{rc} by E_x. Equation 38 simplifies to

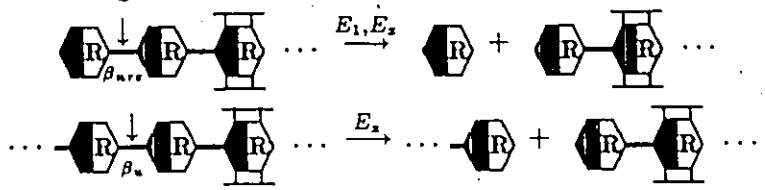
$$\frac{d\beta_{res}}{dt} = \left[\frac{\beta_u + \beta_{rc}}{\beta_u + \beta_{nrc} + \beta_{sr}}\right][\beta_{sr}]\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + G_2 + \beta_u + \beta_{rc} + \beta_{nrc}}\right] \quad (39)$$

3.3.8 Rate Model for β_{nres}

β_{nres} bonds, as opposed to β_{res}, can be cleaved direcly by E_1. This bond type does not accumulate to a fixed concentration as the reaction proceeds in the reactor in the presence of cellobiase. It becomes totally depleted.

The transformation of a β_{snr} bond into a β_{nres} bond is the only mechanism by which the quantity of β_{nres} in the solution can increase. Cleavage of a β_{nrc} or a β_u

bond which was contiguous to a β_{snr} bond transformed the β_{snr} bond into a β_{nres}



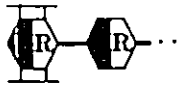
bond. Mathematically, the change in β_{nres} is calculated as

$$\begin{aligned}\frac{d\beta_{nres}}{dt} = &-\left[\frac{\beta_{nres}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\\ &+\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &+\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &+\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned}$$

The first line of the equation is the rate of direct cleavage of the β_{nres} bonds by cellobiase. The second line is the expression for the rate of cleavage of β_u bonds multiplied by the probability (first term) that there was a β_{snr} bond adjacent to that β_u being cleaved. The third and fourth lines are expression for the probability of finding a β_{snr} bond besides a β_{nre} bond multiplied by the rate of cleavage of that β_{nre} bond by the endoglucanase and the cellobiase respectively. Simplifications yield:

$$\begin{aligned}\frac{d\beta_{nres}}{dt} = &\ (\beta_u+\beta_{nre})\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &+\left[\frac{\beta_{nre}\cdot\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}-\beta_{nres}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned} \tag{40}$$

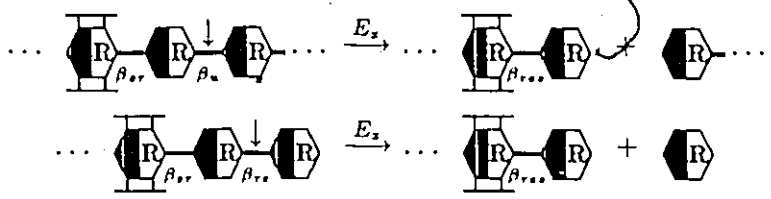
3.3.9 Rate Model for β_{sr}



β_{sr} bonds are mostly non-end bonds at the onset of the hydrolysis reaction. The non-reducing side unit of a β_{sr} bond is substituted. The reducing side unit is unsubstituted and cannot be at the end of the polymer chain or else the bond would be type β_{res}.

For the assumptions listed in this model, there is no way that a β_{sr} bond can be created through a transformation due to the cleavage of any type of bond. And

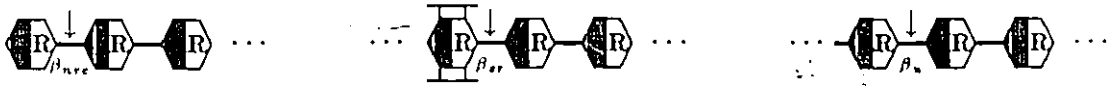
neither E_x nor E_1 can form a complex with this type of bond, followed by cleavage. Yet β_{sr} in the reactor is constantly decreasing. At reaction completion, there has been a partial depletion of β_{sr} bonds becoming reducing-end bonds. Those β_{sr} bonds will thus have been transformed into β_{res} bonds. The only way that this transformation can happen is through the cleavage of a β_u or a β_{re} type bond



which was adjacent to the β_{sr} bond in question. Again, one can go back to the nomenclature and figure out that by definition, β_u and β_{re} bonds can be linked to a β_{sr} bond on its reducing side only. The rate at which the concentration of β_{sr} bonds is changing in the mixture is given by

$$\frac{d\beta_{sr}}{dt} = -\left[\frac{\beta_{sr}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$-\left[\frac{\beta_{sr}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{\beta_{re}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

where the first term on each line is the probability of finding a β_{sr} bond besides any given β_u or β_{re} bond. The probability is the ratio of β_{sr} to the concentration of all bond types which may be on the non-reducing side of β_u or β_{re}. Only three bond types may be placed there...

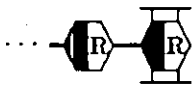


Simplifications yield

$$\frac{d\beta_{sr}}{dt} = -[\beta_u+\beta_{re}]\left[\frac{\beta_{sr}}{\beta_u+\beta_{nre}+\beta_{sr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \qquad (41)$$

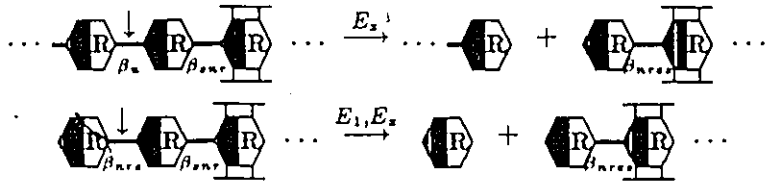
E_x is the only enzyme which directly comes into play in the equation because only it can cleave β_u or β_{re} bonds.

3.3.10 Rate Model for β_{snr}



An unsubstituted anhydroglucose unit and a substituted anhydroglucose unit are linked by a β_{snr} type bond if the unsubstituted unit is closer to the non-reducing end without being the non-reducing end's end unit. β_{snr} bonds are mostly non-end bonds at the onset of the hydrolysis reaction.

The reasoning leading to an expression for the changing β_{snr} is similar as that for β_{sr}. No direct loss of β_{snr} can occur as a direct result of an enzyme attack. However an indirect loss of β_{snr} occurs if a β_{snr} bond happens to be contiguous to a β_u or a β_{nre} bond in the process of being cleaved. Prior to cleavage of an adjacent



β_u or β_{nre} bond, the β_{snr} bond considered is anywhere on the CMC molecule but not at the non-reducing end. In the transformation considered, the β_{snr} bond becomes a β_{nres} bond, the non-reducing end bond of one of the two fragments of the original CMC molecule.

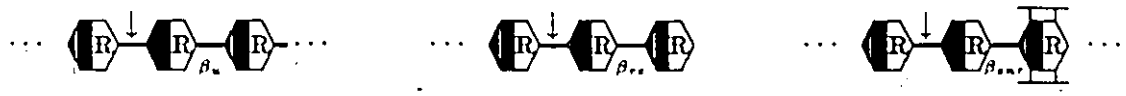
The rate at which β_u bonds are cleaved is

$$\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$

The rate at which β_{nre} bonds are cleaved is the sum of the individual rates of E_1 and E_x which both complex and cleave β_{nre}.

$$\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]$$
$$+\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_x}+G_2+\beta_{nre}+\beta_{nres}}\right]$$

The probability of finding a β_{snr} besides any given β_u or β_{nre} bond is equal to the ratio of β_{snr} to $\beta_u+\beta_{re}+\beta_{snr}$. This last sum is the concentration of all bond



types which may be adjacent to and on the reducing side of β_u and β_{nre}.

$$\begin{aligned}\frac{d\beta_{snr}}{dt} = &-\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_u}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &-\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\left[\frac{E_x\cdot k_{1E_x}\cdot(G_2+\beta_u+\beta_{re}+\beta_{nre})}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &-\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{\beta_{nre}}{G_2+\beta_{nre}+\beta_{nres}}\right]\left[\frac{E_1\cdot k_{1E_1}\cdot(G_2+\beta_{nre}+\beta_{nres})}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned}$$

Simplifying yields

$$\begin{aligned}\frac{d\beta_{snr}}{dt} = &-\left[\frac{\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right][\beta_u+\beta_{nre}]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right]\\ &-\left[\frac{\beta_{nre}\cdot\beta_{snr}}{\beta_u+\beta_{re}+\beta_{snr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right]\end{aligned} \tag{42}$$

3.3.11 Rate Model for β_{ss}

It is simple to track β_{ss} during the enzymatic hydrolysis reaction. These bonds remain unchanging through the course of the reaction. Neither E_1 nor E_x can cleave these bonds between two substituted anhydroglucose units. Even if adjacent bonds are cleaved to yield an end bond of this configuration, it is still considered to be of type β_{ss}. In the extreme case, a β_{ss} bond can be a simple dimer consisting of two substituted anhydroglucose units. The governing equation in this case is thus:

$$\frac{d\beta_{ss}}{dt} = 0 \quad \text{or} \quad \beta_{ss} = \beta_{ss}(0) \tag{43}$$

In the set of variables introduced to describe the hydrolysis reaction of cellulose derivatives, β_{ss} is not part of the set of simultaneous differential equations that are solved to estimate concentrations of G_1 and G_2. In an alternate model however, an assumption that β-1,4 bonds between substituted units inhibit the enzymes, would reintroduce β_{ss} into the differential equation system.

3.3.12 Rate Model for β

The rate at which the total concentration of bonds in the solution is changing depends only on how fast the enzyme system is cleaving bonds. Bond *transformations* do not have a *net* effect on total bond concentration. At any time t therefore, the instantaneous rate of change in bond concentration is calculated as

$$\begin{aligned}\frac{d\beta}{dt} = &-[G_2 + \beta_u + \beta_{re} + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right] \\ &-[G_2 + \beta_{nre} + \beta_{nres}]\left[\frac{E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right]\end{aligned} \tag{44}$$

The first line of equation 44 is the bond cleavage rate of G_2, β_u, β_{re} and β_{nre} by endoglucanase. The second line is the bond cleavage rate of G_2, β_{nre} and β_{nres} by cellobiase.

3.3.13 DP

The final quantity included in the model may give the reader a feel for the extent of reaction. It is possible to develop an equation to calculate the DP of the hydrolysis reaction mixture. Define the DP as a 'number-average degree of polymerization' that takes the concentration of monomer—substituted or unsubstituted glucose—into consideration.

For example, the reaction mixture has a degree of polymerization of 1000 if there is but one CMC molecule of length 1000 units. Should an enzyme cleave one unit off of one end of that molecule, the mixture will now contain two CMC fragments, one of length 1 unit and the other of length 999 units. The DP of the mixture is calculated as the average length of all CMC fragments. i.e., in this case, the DP is $\frac{1}{2}(999 + 1) = 500$ units. In other words, the DP is equal to the ratio of the total number of glucosidic units divided by the total number of molecules in the reactor.

$$\text{DP} = \frac{\text{number of glucosidic units}}{\text{total number of molecules}} \tag{45}$$

This definition of DP implies that there is a direct relationship between DP and the number of bonds cleaved by any enzyme, independent of the specificity of the enzyme in cleaving a particular bond.

The number of molecules equals the number of glucosidic units minus the number of β-1,4 bonds, β. Thus,

$$\text{units} = \text{number of glucosidic units} = \beta + \text{number of molecules} \tag{46}$$

Comparing equations 45 and 46 to eliminate 'number of molecules' and solving for the number of units,

$$\text{units} = \frac{\beta \cdot \text{DP}}{\text{DP} - 1} \tag{47}$$

At time $t = 0$,

$$\text{units} = \frac{\beta(0) \cdot \text{DP}(0)}{\text{DP}(0) - 1} \tag{48}$$

Eliminating 'units', which is unchanging through the reaction, from the last two equations,

$$\frac{\beta \cdot \text{DP}}{\text{DP} - 1} = \frac{\beta(0) \cdot \text{DP}(0)}{\text{DP}(0) - 1}$$

and rearranging to give

$$\text{DP} = \frac{\beta(0) \cdot \text{DP}(0)}{(\beta(0) - \beta) \cdot \text{DP}(0) + \beta} \tag{49}$$

The initial values of both DP and β as well as the concentration of bonds β, at time t, must be known. Equation 49 describes the number average degree of polymerization during the hydrolysis reaction.

3.4 Summary of the Rate Equations for the Basic Model (Model 1)

$$\frac{dG_1}{dt} = +[2G_2+\beta_{re}+\beta_{nre}]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] + [2G_2+\beta_{nre}+\beta_{nres}]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{50}$$

$$\frac{dG_2}{dt} = +\left[\frac{\beta_u\cdot\beta_{re}+\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{enr}}+\frac{\beta_u\cdot\beta_{nre}+\beta_{re}\cdot\beta_{nre}}{\beta_u+\beta_{nre}+\beta_{er}}-G_2\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] + \left[\frac{\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{enr}}-G_2\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{51}$$

$$\frac{d\beta_u}{dt} = -\left[\frac{\beta_u+\beta_{nre}}{\beta_u+\beta_{re}+\beta_{enr}}+\frac{\beta_u+\beta_{re}}{\beta_u+\beta_{nre}+\beta_{er}}+1\right][\beta_u]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] - \left[\frac{\beta_u\cdot\beta_{nre}}{\beta_u+\beta_{re}+\beta_{enr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{52}$$

$$\frac{d\beta_{re}}{dt} = +\left[\frac{\beta_u\cdot(\beta_u+\beta_{re})}{\beta_u+\beta_{nre}+\beta_{er}}-\beta_{re}-\frac{\beta_{re}\cdot(\beta_u+\beta_{nre})}{\beta_{re}+\beta_u+\beta_{enr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] - \left[\frac{\beta_{nre}\cdot\beta_{re}}{\beta_u+\beta_{re}+\beta_{enr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{53}$$

$$\frac{d\beta_{res}}{dt} = +\left[\frac{\beta_u+\beta_{re}}{\beta_u+\beta_{nre}+\beta_{er}}\right][\beta_{er}]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \tag{54}$$

$$\frac{d\beta_{nre}}{dt} = +\left[\frac{\beta_u\cdot(\beta_u+\beta_{nre})}{\beta_u+\beta_{re}+\beta_{enr}}-\beta_{nre}-\frac{\beta_{nre}\cdot(\beta_u+\beta_{re})}{\beta_u+\beta_{nre}+\beta_{er}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] + \left[\frac{\beta_u\cdot\beta_{nre}}{\beta_u+\beta_{re}+\beta_{enr}}-\beta_{nre}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{55}$$

$$\frac{d\beta_{nres}}{dt} = +[\beta_u+\beta_{nre}]\left[\frac{\beta_{enr}}{\beta_u+\beta_{re}+\beta_{enr}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] + \left[\frac{\beta_{nre}\cdot\beta_{enr}}{\beta_u+\beta_{re}+\beta_{enr}}-\beta_{nres}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{56}$$

$$\frac{d\beta_{er}}{dt} = -[\beta_u+\beta_{re}]\left[\frac{\beta_{er}}{\beta_u+\beta_{nre}+\beta_{er}}\right]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] \tag{57}$$

$$\frac{d\beta_{enr}}{dt} = -\left[\frac{\beta_{enr}}{\beta_u+\beta_{re}+\beta_{enr}}\right][\beta_u+\beta_{nre}]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] - \left[\frac{\beta_{nre}\cdot\beta_{enr}}{\beta_u+\beta_{re}+\beta_{enr}}\right]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{58}$$

$$\frac{d\beta_{ss}}{dt} = 0 \quad \text{or} \quad \beta_{ss} = \beta_{ss}(0) \tag{59}$$

$$\frac{d\beta}{dt} = -[\beta_u+G_2+\beta_{re}+\beta_{nre}]\left[\frac{E_x\cdot k_{1E_x}}{K_{mE_x}+G_2+\beta_u+\beta_{re}+\beta_{nre}}\right] - [G_2+\beta_{nre}+\beta_{nres}]\left[\frac{E_1\cdot k_{1E_1}}{K_{mE_1}+G_2+\beta_{nre}+\beta_{nres}}\right] \tag{60}$$

$$\mathrm{DP} = \frac{\beta(0)\cdot\mathrm{DP}(0)}{(\beta(0)-\beta)\cdot\mathrm{DP}(0)+\beta} \tag{61}$$

3.5 Initial Conditions of Model Equations

When applying the model to experimental data, the known initial conditions, by measurement, are $G_1(0)$, $G_2(0)$, the DS and DP of the CMC and the initial mass concentration of CMC in the reactor (CMC_0). The initial concentrations of the types of bonds from CMC, i.e., $\beta_u(0)$, $\beta_{re}(0)$ ect. as well as the initial degree of polymerization of the reaction mixture must be calculated.

First, consider the initial DP of the mixture. Define here, $\mathrm{DP_{CMC}}$, as the degree of polymerization of the CMC initially in the reactor, not that of the reaction mixture, which may be corrected in some cases to account for the presence of glucose and cellobiose. Similarly, define β_{CMC} as the concentration of β-1,4 bonds associated with the CMC initially in the reactor which can be calculated to be

$$\beta_{\mathrm{CMC}} = \frac{\mathrm{DP_{CMC}} - 1}{\mathrm{DP_{CMC}}} \cdot \underbrace{\frac{CMC_0}{162 + 80 \cdot \mathrm{DS} + 18/\mathrm{DP_{CMC}}}}_{\text{avg. CMC unit molecular weight}} \tag{62}$$

If only CMC is placed into the reaction mixture, then $\mathrm{DP}(0) = \mathrm{DP_{CMC}}$. However, in some cases, glucose (G_1) and cellobiose (G_2) may also be introduced into the reactor. Then, in general, the initial number of glucosidic units is given by

$$\text{units} = \left[\frac{CMC_0}{162 + 80 \cdot \mathrm{DS} + 18/\mathrm{DP_{CMC}}} + G_1(0) + 2G_2(0)\right] \cdot [\text{Reaction Volume}] \tag{63}$$

The concentration of molecules due to the initial CMC is

$$\left[\frac{CMC_0}{162 + 80 \cdot \mathrm{DS} + 18/\mathrm{DP_{CMC}}} - \beta_{\mathrm{CMC}}\right]$$

Consequently the total number of molecules initially in the mixture is given by

$$\text{molecules}(0) = \left[\frac{CMC_0}{162 + 80 \cdot \mathrm{DS} + 18/\mathrm{DP_{CMC}}} - \beta_{\mathrm{CMC}} + G_1(0) + G_2(0)\right] [\text{Volume}] \tag{64}$$

Thus the initial DP of the reaction mixture in the general case can be calculated as

$$\mathrm{DP}(0) = \frac{\left[\frac{CMC_0}{162+80\cdot\mathrm{DS}+18/\mathrm{DP_{CMC}}} + G_1(0) + 2G_2(0)\right]}{\left[\frac{CMC_0}{162+80\cdot\mathrm{DS}+18/\mathrm{DP_{CMC}}} - \beta_{\mathrm{CMC}} + G_1(0) + G_2(0)\right]} \tag{65}$$

Calculation of the initial concentrations of the different types of bonds will require the determination of the number of end bonds and non-end bonds. Considering the average CMC molecule of length DP_{CMC} units, the ratio of end bonds to total bonds is $[2/DP_{CMC} - 1]$. Thus the concentration of CMC end bonds in the reactor is $\beta_{CMC} \cdot [2/(DP_{CMC} - 1)]$ and the concentration of CMC non-end bonds is $\beta_{CMC} \cdot (DP_{CMC} - 3)/(DP_{CMC} - 1)$. The fraction of anhydroglucose units of the CMC which are unsubstituted, c_0, will also be required and can be calculated from either of two equations presented earlier, equations based on equal selectivities of carbon positions or on Timell's experimental determination [47].

The initial amount of β_u bonds in the solution is equal to the amount of CMC mid-molecule bonds multiplied by the probability that a given mid-molecule bond links two unsubstituted anhydroglucose units.

$$\beta_u(0) = \beta_{CMC} \cdot \left(\frac{DP_{CMC} - 3}{DP_{CMC} - 1}\right) \cdot c_0^2 \tag{66}$$

The initial amounts of β_{re} and of β_{nre} bonds are equal and calculated as the amount of CMC end bonds divided by two—to account for the fact that one of the two ends is of the reducing type while the other is a non-reducing end—multiplied by the probability that the end unit and the next-to-end unit of a given CMC end are both unsubstituted. The mathematical representation is

$$\beta_{re}(0) = \beta_{nre}(0) = \beta_{CMC} \cdot \left(\frac{2}{DP_{CMC} - 1}\right) \cdot \frac{1}{2} \cdot c_0^2 \tag{67}$$

The initial amounts of β_{res} and of β_{nres} bonds are equal and calculated as the amount of CMC end bonds divided by two, then multiplied by the probability that the end unit of the CMC end considered is unsubstituted and the next-to-end unit is substituted and is given by

$$\beta_{res}(0) = \beta_{nres}(0) = \beta_{CMC} \cdot \left(\frac{2}{DP_{CMC} - 1}\right) \cdot \frac{1}{2} \cdot c_0 \cdot (1 - c_0) \tag{68}$$

The initial amounts of β_{sr} and of β_{snr} bonds are equal and calculated as the amount of CMC mid-molecule bonds multiplied by the probability that one of the two units linked by the bond is unsubstituted whereas the other linked unit is substituted—noting that order is important since β_{sr} type bonds have the unsubstituted unit closer to the reducing end than is the substituted unit. Add to the above quantity the situation when an end bond is classified as a β_{sr} or a β_{snr} bond. For example, if the non-reducing end's end anhydroglucose unit is substituted and the unit next to it is unsubstituted, that end bond is considered to be a β_{sr} bond. The same situation can happen in reverse at the reducing end of the molecule, with the reducing end bond being of type β_{snr}. The mathematical representation is

$$\beta_{sr}(0) = \beta_{snr}(0) = \beta_{\mathrm{CMC}} \cdot \left(\frac{\mathrm{DP_{CMC}} - 3}{\mathrm{DP_{CMC}} - 1}\right) \cdot c_0 \cdot (1 - c_0) + \beta_{\mathrm{CMC}} \cdot \left(\frac{2}{\mathrm{DP_{CMC}} - 1}\right) \cdot \frac{1}{2} \cdot c_0 \cdot (1 - c_0)$$

which can be simplified to give

$$\beta_{sr}(0) = \beta_{snr}(0) = \beta_{\mathrm{CMC}} \cdot \left(\frac{\mathrm{DP_{CMC}} - 2}{\mathrm{DP_{CMC}} - 1}\right) \cdot c_0 \cdot (1 - c_0) \tag{69}$$

The β_{ss} bonds are not a function of position on the CMC molecule. They can be found in a mid-molecule position or as the end bond of the CMC. A bond is of type β_{ss} when the two units linked by the bonds are substituted. The probability that any given bond in the CMC chain is of type β_{ss} is equal to the probability that both units linked by that bond are substituted. And since the probability that one unit is substituted is given by $1 - c_0$, the probability that both units are substituted is $(1 - c_0)^2$. Multiply the probability of a CMC bond being of type β_{ss} by the concentration of bonds, β_{CMC} to obtain the initial—and unchanging—concentration of β_{ss} bonds.

$$\beta_{ss} = \beta_{\mathrm{CMC}} \cdot (1 - c_0)^2 \tag{70}$$

The total number of bonds initially in the reaction mixture is calculated as

$$\beta(0) = \beta_{\mathrm{CMC}} + G_2(0) \tag{71}$$

The consistency of equations 67 to 70 can be verified by substituting these equations into the relationship

$$\beta_u(0) + \beta_{re}(0) + \beta_{nre}(0) + \beta_{res}(0) + \beta_{nres}(0) + \beta_{sr}(0) + \beta_{snr}(0) + \beta_{ss}(0) = \beta_{\mathrm{CMC}}$$

The sum of the initial types of bonds should and does here equal the total number of bonds in the solution.

3.6 Final Conditions of Model Equations

Prediction of the final concentrations of the various species resulting from total hydrolysis (at $t = \infty$) of the CMC by a complete set of cellulases does not require any knowledge of the kinetic constants. These predictions can be used to provide a preliminary evaluation of the proposed mechanisms and the accuracy of the numerical procedure for solving the rate equations. Rather than solving the system of non-linear algebraic equations resulting from setting the derivative terms in the model equal to zero, a procedure based upon the distribution of bond types is used.

If we assume that there is no formation of side products due to the degradation of glucose, the glucose concentration in the reactor will constantly increase with time to reach its final value (at $t = \infty$). The harvested glucose comes from three sources:

1. Each glucose molecule initially in the reactor will go through the reaction unchanged.
2. Every cellobiose molecule initially in the reactor will be hydrolyzed to produce two glucose molecules.
3. Some glucose will be produced from the hydrolysis of the β_{CMC} bonds.

All CMC bonds that have their non-reducing side unit unsubstituted and the 'next-to-non-reducing-side' unit also unsubstituted will yield that non-reducing side unit,

the middle unit in the diagrams, in the form of a molecule of glucose at complete reaction. From this theory, the amount of glucose extracted from mid-molecule units is given as

$$\beta_{CMC} \cdot \left(\frac{DP_{CMC} - 2}{DP_{CMC} - 1} \right) \cdot c_0^2 \tag{72}$$

A non-reducing end unit will eventually become a glucose molecule if that unit is unsubstituted.

$$\beta_{CMC} \cdot \left(\frac{1}{DP_{CMC} - 1} \right) \cdot c_0 \tag{73}$$

And finally, a reducing end unit will eventually become a glucose molecule if that unit and its adjacent unit are both unsubstituted.

$$\beta_{CMC} \cdot \left(\frac{1}{DP_{CMC} - 1} \right) \cdot c_0^2 \tag{74}$$

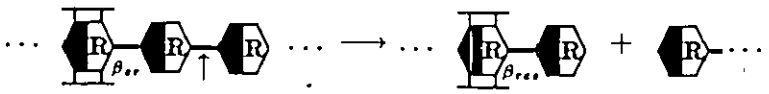
Add the three preceeding equations to the glucose from $G_1(0)$ and $G_2(0)$ to obtain

$$G_1(\infty) = G_1(0) + 2G_2(0) + \frac{\beta_{CMC} \cdot DP_{CMC}}{DP_{CMC} - 1} \cdot \left[\frac{c_0}{DP_{CMC}} + \frac{c_0^2}{DP_{CMC}} + \frac{(DP_{CMC} - 2)}{DP_{CMC}} \cdot c_0^2 \right] \tag{75}$$

$$G_1(\infty) = G_1(0) + 2G_2(0) + \beta_{CMC} \cdot c_0 \cdot \left(c_0 + \frac{1}{DP_{CMC} - 1} \right) \tag{76}$$

Many of the bond types are cleaved directly by the enzyme. Those bond types are either depleted as the reaction goes on or they reach a maximum before they start depleting. Conditions favor the attainment of a maximum if their production through the transformation of another bond type is occuring at a high rate. Here are all the bond types depleted as $t \to \infty$: $G_2(\infty) = \beta_u(\infty) = \beta_{re}(\infty) = \beta_{nre}(\infty) = \beta_{nres}(\infty) = 0$. This equation is valid under the condition that both endoglucanase and cellobiase are present in the reactor.

β_{res} bonds strictly accumulate as the reaction goes on. All the β_{res} bonds initially in the reactor remain unaffected by the reaction. And more β_{res} bonds are 'created' as the reducing-side bond adjacent to a β_{sr} bond is cleaved. This cleavage can only take place if the unit next to the β_{sr}, on its reducing side, is unsubstituted.

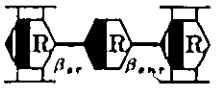


Thus at $t = \infty$,

$$\beta_{rcs}(\infty) = \beta_{rcs}(0) + c_0 \cdot \beta_{sr}(0) \tag{77}$$

$$\beta_{rcs}(\infty) = \beta_{\mathrm{CMC}} \cdot \frac{c_0 \cdot (1 - c_0)}{\mathrm{DP_{CMC}} - 1} + c_0 \cdot \beta_{\mathrm{CMC}} \cdot c_0 \cdot (1 - c_0) \cdot \left(\frac{\mathrm{DP_{CMC}} - 2}{\mathrm{DP_{CMC}} - 1} \right)$$

$$\beta_{rcs}(\infty) = \beta_{\mathrm{CMC}} \cdot c_0 \cdot (1 - c_0) \cdot (1 + (\mathrm{DP_{CMC}} - 2) \cdot c_0) / (\mathrm{DP_{CMC}} - 1) \tag{78}$$



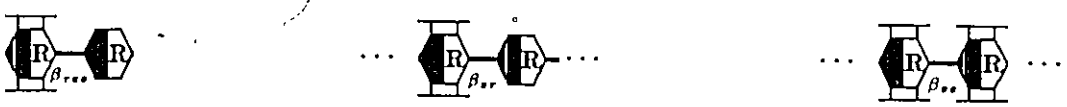
$\beta_{sr}(\infty)$ and $\beta_{snr}(\infty)$ are equal and their actual concentration depends on the probability that the unit next to the unsubstituted unit at the onset of the reaction is substituted. The final quantities of bond types β_{sr} and β_{snr} equals their initial quantity multiplied by the probability that the unit is substituted, $1 - c_0$.

$$\beta_{sr}(\infty) = \beta_{snr}(\infty) = \beta_{\mathrm{CMC}} \cdot \left(\frac{\mathrm{DP_{CMC}} - 2}{\mathrm{DP_{CMC}} - 1} \right) \cdot c_0 \cdot (1 - c_0)^2 \tag{79}$$

There is no change in β_{ss}.

$$\beta_{ss}(\infty) = \beta_{\mathrm{CMC}} \cdot (1 - c_0)^2 \tag{80}$$

The total quantity of bonds β at the completion of the reaction may also be calculated from the initial data. The question one must ask is 'What bonds in the reactor will not be cleaved by either E_x or E_1?' The answer is 'All bonds with a substituted unit on the non-reducing side',

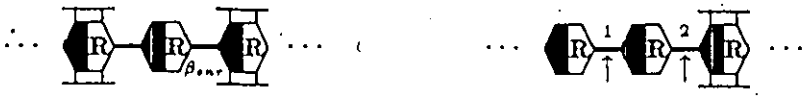


This category of bonds includes β_{ss}, β_{sr} and β_{rcs} bonds initially in the reactor. Their total concentration is $\beta_{\mathrm{CMC}} \cdot (1 - c_0)$.

The bonds that have no substituted units as part of their linked groups are not considered the calculation of $\beta(\infty)$ since they are cleavable by at least E_x.

Which of the bonds with a substituted unit on the reducing side only will escape cleavage? β_{nrcs} is directly depleted by the cellobiase thus our attention must turn to

β_{snr}. If the unit adjacent to the β_{snr} bond, on its non-reducing side is substituted, the β_{snr} bond cannot be cleaved because the adjacent β_{sr}-bond is uncleavable. On



the other hand, were the non-reducing side adjacent unit unsubstituted, then the adjacent bond to β_{snr} would first be cleaved, resulting in the transformation of our β_{snr} bond into a β_{nres} bond. This would lead to its eventual cleavage by cellobiase, as assumed in the enzyme mechanism. The contribution of $\beta_{snr}(0)$ to $\beta(\infty)$ is thus equal to its initial concentration multiplied by the probability, $1 - c_0$, that the next non-reducing side unit is substituted. The overall number of bonds in the reaction mixture at $t = \infty$ is

$$\beta(\infty) = \beta_{\mathrm{CMC}} \cdot \left[(1 - c_0) + c_0 \cdot (1 - c_0) \cdot \left(\frac{\mathrm{DP_{CMC}} - 2}{\mathrm{DP_{CMC}} - 1}\right) \cdot (1 - c_0)\right]$$

which may be rewritten as

$$\beta(\infty) = \beta_{\mathrm{CMC}} \left[\frac{(-c_0 + 2c_0^2 - c_0^3)}{\mathrm{DP_{CMC}} - 1} + (1 - 2c_0^2 + c_0^3)\right] \tag{81}$$

The sum of the individual types of bonds in the solution at $t = \infty$ should, and does, equal the total quantity of bonds β in the reactor as calculated in equation 81.

$$\beta(\infty) = \beta_{ss}(\infty) + \beta_{snr}(\infty) + \beta_{sr}(\infty) + \beta_{res}(\infty) \tag{82}$$

3.7 Development of Rival Mechanisms to Model CMC Hydrolysis

The first model developed, model 1, assumes ideal Michaelis-Menten type kinetics with no inhibition. Enzymatic reactions with inhibition of the enzyme by any of the various species involved in the reaction are more common. The enzymatic hydrolysis of cellulose substrates by the cellulase system is no exception according to the literature. The beauty of the approach of modelling bond hydrolysis is that it allows for inhibition of any of the enzymes in the mixture by any of the products or reactants. There is another reason for trying other kinetic models; the mechanism of the enzymatic hydrolysis is not known.

In the following subsections, the modifications to the present model necessary to accommodate the observations made in building up to our proposed model will be described. Note that the model numbering scheme is consistent with the presentation of results.

3.7.1 Model 2: No G_2 cleavage by E_x

In the development of the basic model, it was assumed that cellobiose is hydrolyzed by endoglucanase, E_x, as well as cellobiase, E_1. However it is often assumed that E_x does not cleave G_2 bonds. This assumption is essential in the use of cellobiose as a substrate in hydrolysis reactions to estimate E_1 activity. A model incorporating this assumption is therefore of interest. Comparison of experimental results with both models offers an opportunity to check the validity of either mechanism.

Elimination of G_2 as a substrate for E_x alters the denominator of the "Michaelis-Menten" term for E_x in model 1 so that

$$\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + G_2 + \beta_u + \beta_{re} + \beta_{nre}}\right] \quad \text{becomes} \quad \left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + \beta_u + \beta_{re} + \beta_{nre}}\right]$$

The three differential rate equations in model 1 describing G_1, G_2 and β will also change. With no direct cleavage of G_2 by E_x, the rate of production of G_1 and the

depletion rates of G_2 and β will contain one less term each.

$$\frac{dG_1}{dt} = +[\beta_{re} + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + \beta_u + \beta_{re} + \beta_{nre}}\right] + [2G_2 + \beta_{nre} + \beta_{nres}]\left[\frac{E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right] \quad (83)$$

$$\frac{dG_2}{dt} = +\left[\frac{\beta_u \cdot \beta_{re} + \beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} + \frac{\beta_u \cdot \beta_{nre} + \beta_{re} \cdot \beta_{nre}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + \beta_u + \beta_{re} + \beta_{nre}}\right] + \left[\frac{\beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} - G_2\right]\left[\frac{E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right] \quad (84)$$

$$\frac{d\beta}{dt} = -[\beta_u + \beta_{re} + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + \beta_u + \beta_{re} + \beta_{nre}}\right] - [G_2 + \beta_{nre} + \beta_{nres}]\left[\frac{E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right] \quad (85)$$

Application of model 2 will show this modification to be beneficial and it is retained in the development of all other models.

3.7.2 Model 3: Non Michaelis-Menten Kinetics for E_x

Enzymatic reactions generally follow Michaelis-Menten kinetics. Other modes of action can be assumed. The most popular alternate mechanisms are zero and first-order kinetics. The reason for deviating from Michaelis-Menten kinetics is that the data does not contain sufficient information to permit the evaluation of kinetic constants with any kind of precision.

In model 3, it is assumed that E_x follows zero-order kinetics with no inhibition. The rate of substrate cleavage by E_x is constant in kinetics of order zero. Model 2 is modified in one way: $K_{mE_x} = 0$. In practice, the integration routine for the computer program is modified. The hydrolysis rate by E_x is set to 0 when $\beta_u + \beta_{re} + \beta_{nre} = 0$ to prevent exponential overflow—there is then no substrate for E_x in the reactor.

$$\left[\frac{E_x \cdot k_{1E_x}}{K_{mE_x} + \beta_u + \beta_{re} + \beta_{nre}}\right] \quad \text{becomes} \quad \left[\frac{E_x \cdot k_{1E_x}}{\beta_u + \beta_{re} + \beta_{nre}}\right]$$

3.7.3 Model 4: Non Michaelis-Menten Kinetics for E_1

Cellobiase kinetics are first-order. A simple Michaelis-Menten rate equation is downgraded to first-order kinetics by eliminating the denominator. For example, in model 4, K_{mE_1} is assumed to be $\gg G_2 + \beta_{nre} + \beta_{nres}$ during the reaction. Then the denominator of the Michaelis expression is a constant. The rate constant k_{1E_1} is totally correlated with the Michaelis constant. In the fitting routine, they are combined into one parameter to be evaluated. That parameter will still be referred to as k_{1E_1}.

$$\left[\frac{E_1 \cdot k_{1E_1}}{K_{mE_1} + G_2 + \beta_{nre} + \beta_{nres}}\right] \quad \text{becomes} \quad [E_1 \cdot k_{1E_1}]$$

3.7.4 Model 5: Inhibition of E_1 by G_1

A major contribution of the Michaelis-Menten approach to modelling the hydrolysis kinetics is accounting quantitatively for the presence of inhibition. The modifications to the basic equation are developed elsewhere [28]. We will apply the derived results to our overall hydrolysis model. If we are fortunate, we may be able to identify the inhibitors of each particular enzyme.

The rate of reaction of a Michaelis-Menten reaction with inhibition depends on whether the inhibition is competitive or non-competitive. A competitive inhibitor blocks the active site. No other substrate or inhibitor can form a complex with the enzyme if the active site is already blocked by another substrate or inhibition. For total competition,

$$v = \frac{ke_0 s}{s + K_m(1 + i/K_i)} \tag{86}$$

where s and i are the substrate and inhibitor concentrations respectivelely. K_i is the dissociation constant of the enzyme-inhibitor complex.

Totally non-competitive inhibition arises when the inhibitor binds to the enzyme at a point other than the active site. A ternary enzyme-inhibitor-substrate complex is now possible. However the binding of the inhibitor halts the breakdown of the

ternary complex into products. The rate equation for non-competitive inhibition is given by

$$v = \frac{ke_0 s}{(s + K_m)(1 + i/K_i)} \tag{87}$$

After fitting the data to model 4, it was decided to keep the assumption that $K_{mE_1} \gg G_2 + \beta_{nrc} + \beta_{nres}$. Applied to equations 86 and 87, this means that $K_m \gg s$. Thus, it turns out that with a large Michaelis constant, the equations describing the kinetics of inhibition by a substrate are identical.

To obtain model 5, a change was made to the denominator of the cellobiase rate expression of model 4.

$$[E_1 \cdot k_{1E_1}] \quad \text{becomes} \quad \left[\frac{E_1 \cdot k_{1E_1}}{1 + G_1/K_{iG_1}}\right]$$

With this last modification, it is possible to fit the model that assumes G_1 to be an inhibitor of E_1. It is not possible however to distinguish whether the inhibition is competitive or non-competitive.

Model 5 fitting will show that no improvement is obvious in the model by assuming E_1 inhibition by G_1. That assumption is not retained in the overall hydrolysis models yet to be developed.

3.7.5 Model 6: Inhibition of E_x by its Non-Substrate Bonds

Model 6 checks for inhibition of E_x by non-substrate β-1,4 bonds in the solution. If we assume that the endoglucanase enzyme forms a complex with all β-1,4 glucosidic bonds in the solution and that the dissociation rate of that bond-enzyme complex does not depend on the type of bond, the model is easily modified. In this case the inhibition is assumed to be competitive since the complex would form at the active point of the enzyme. By assuming equal dissociation rates for the complex, whether it is a complex with a cleavable or an uncleavable β-1,4 linkage, a new parameter is not necessary in the model. The denominator of the E_x rate

term in the rate equations is transformed as follows:

$$\left[\frac{E_x \cdot k_{1E_x}}{\beta_u + \beta_{rc} + \beta_{nrc}}\right] \quad \text{becomes} \quad \left[\frac{E_x \cdot k_{1E_x}}{\beta}\right]$$

Equal dissociation rates in equation 86 means that $K_i = K_m$. Consequently the substrate and inhibitor concentrations are summed up in the denominator of the rate expression of E_x.

Some phenomena can be explained by the above equation, if it is indeed representative of the mechanism of the enzymatic hydrolysis of CMC. Highly-substituted methylcellulose has been found to be an inhibitor of the cellulase enzyme system. In adding methylcellulose to a reaction mixture, one is in fact dumping an important quantity of β-1,4 bonds which are not cleavable. These are mostly β_{ss}, β_{sr}, β_{snr}, β_{rcs} and β_{nrcs} type bonds which are direct inhibitors of the endoglucanase.

With the new assumptions concerning inhibition by the above species, the overall hydrolysis model is rewritten as on page 94.

3.7.6 Models 7–10: Cellobiose Hydrolysis by E_1

Models 7–10 are fitted to the 28 run locations which have no CMC as substrate. E_x and E_2 are again assumed to be incapable of cleaving cellobiose. Thus models 7–10 look at the kinetics of cellobiose hydrolysis by cellobiase.

The four models in question look at four different mechanisms for the mode of action of E_1. Model 7 is a first order kinetic model. Model 8 is Michaelis-Menten type kinetics with no inhibition. Model 9 is Michaelis-Menten type kinetics with reversible non-competitive inhibition by glucose. Model 10 is Michaelis-Menten type kinetics with reversible competitive inhibition by glucose. All of the rate equations describing these mechanisms are developed elsewhere [28] and have been discussed earlier. The four rate equations are as follows:

$$\text{Model 7} \qquad \frac{dG_1}{dt} = -\frac{2dG_2}{dt} = 2 \cdot k_{1E_1} \cdot E_1 \cdot G_2 \tag{100}$$

Equations of Model 6 and Model 12

$$\frac{dG_1}{dt} = [\beta_{re} + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + [2G_2 + \beta_{nre} + \beta_{nres}][E_1 \cdot k_{1E_1}] \tag{88}$$

$$\frac{dG_2}{dt} = \left[\frac{\beta_u \cdot \beta_{re} + \beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} + \frac{\beta_u \cdot \beta_{nre} + \beta_{re} \cdot \beta_{nre}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} - G_2\right][E_1 \cdot k_{1E_1}] \tag{89}$$

$$\frac{d\beta_u}{dt} = -\left[\frac{\beta_u + \beta_{nre}}{\beta_u + \beta_{re} + \beta_{snr}} + \frac{\beta_u + \beta_{re}}{\beta_u + \beta_{nre} + \beta_{sr}} + 1\right][\beta_u]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_u \cdot \beta_{nre}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{90}$$

$$\frac{d\beta_{re}}{dt} = \left[\frac{\beta_u \cdot (\beta_u + \beta_{re})}{\beta_u + \beta_{nre} + \beta_{sr}} - \beta_{re} - \frac{\beta_{re} \cdot (\beta_u + \beta_{nre})}{\beta_{re} + \beta_u + \beta_{snr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{91}$$

$$\frac{d\beta_{res}}{dt} = \left[\frac{\beta_u + \beta_{re}}{\beta_u + \beta_{nre} + \beta_{sr}}\right][\beta_{sr}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] \tag{92}$$

$$\frac{d\beta_{nre}}{dt} = \left[\frac{\beta_u \cdot (\beta_u + \beta_{nre})}{\beta_u + \beta_{re} + \beta_{snr}} - \beta_{nre} - \frac{\beta_{nre} \cdot (\beta_u + \beta_{re})}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_u \cdot \beta_{nre}}{\beta_u + \beta_{re} + \beta_{snr}} - \beta_{nre}\right][E_1 \cdot k_{1E_1}] \tag{93}$$

$$\frac{d\beta_{nres}}{dt} = [\beta_u + \beta_{nre}]\left[\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_{nre} \cdot \beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}} - \beta_{nres}\right][E_1 \cdot k_{1E_1}] \tag{94}$$

$$\frac{d\beta_{sr}}{dt} = -[\beta_u + \beta_{re}]\left[\frac{\beta_{sr}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] \tag{95}$$

$$\frac{d\beta_{snr}}{dt} = -\left[\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right][\beta_u + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_{nre} \cdot \beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{96}$$

$$\frac{d\beta_{ss}}{dt} = 0 \quad \text{or} \quad \beta_{ss} = \beta_{ss}(0) \tag{97}$$

$$\frac{d\beta}{dt} = -[\beta_u + \beta_{re} + \beta_{nre}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - [G_2 + \beta_{nre} + \beta_{nres}][E_1 \cdot k_{1E_1}] \tag{98}$$

$$\mathrm{DP} = \frac{\beta(0) \cdot \mathrm{DP}(0)}{(\beta(0) - \beta) \cdot \mathrm{DP}(0) + \beta} \tag{99}$$

Model 8
$$\frac{dG_1}{dt} = -\frac{2dG_2}{dt} = \frac{2 \cdot k_{1E_1} \cdot E_1 \cdot G_2}{K_{mE_1} + G_2} \quad (101)$$

Model 9
$$\frac{dG_1}{dt} = -\frac{2dG_2}{dt} = \frac{2 \cdot k_{1E_1} \cdot E_1 \cdot G_2}{(G_2 + K_{mE_1}) \cdot (1 + G_1/K_{iG_1})} \quad (102)$$

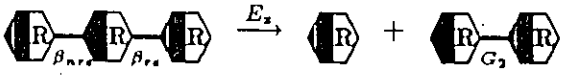
Model 10
$$\frac{dG_1}{dt} = -\frac{2dG_2}{dt} = \frac{2 \cdot k_{1E_1} \cdot E_1 \cdot G_2}{K_{mE_1} \cdot (1 + G_1/K_{iG_1}) + G_2} \quad (103)$$

3.7.7 Model 11: E_x Does Not Cleave β_{nre}

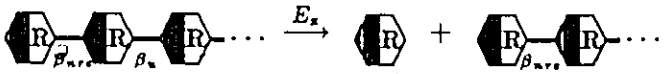
Model 11 is obtained by making one further assumption about the possible substrates of endoglucanase. Modify model 6 by assuming that β_{nre} is not a substrate of E_x. That mechanism may explain the concentration of cellobiose in the reactor. All models fitted thus far have underpredicted the concentration of cellobiose produced by the hydrolysis of CMC.

The two enzymes now do not have a common substrate. E_x cleaves β_u and β_{re}. E_1 cleaves G_2, β_{nre} and β_{nres}. Endoglucanase needs three consecutive units to cleave a given bond. And it must be positioned properly i.e., to be cleaved by E_x, a bond must lie between two unsubstituted anhydroglucose units and there must be a bond on the non-reducing side of the bond being cleaved. Most of the differential rate equations of model 6 are affected by the new assumption. Here are the bond reactions which are different from model 6 and affect the form of model 11.

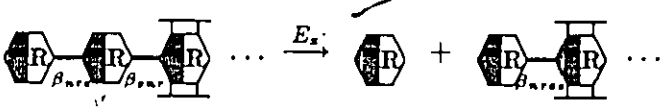
1. There is no direct cleavage of β_{nre} bonds by E_x.
2. β_{re} cannot be transformed into G_2 by cleavage of an adjacent β_{nre} by E_x.



3. β_u cannot be transformed into β_{nre} by cleavage of an adjacent β_{nre} by E_x.



4. β_{snr} cannot be transformed into β_{nres} by cleavage of an adjacent β_{nre} by E_x.



The step-by-step development of model 11 is not shown here. The new set of equation can be found on the following page.

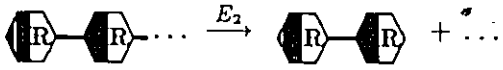
Making the assumption that β_{nre} is not a substrate of E_x is, in fact, not the solution to explaining the high concentration of cellobiose in the hydrolysis mixture. Consequently this mechanism is rejected.

3.7.8 Model 12: Application of Model 6 to Revised Data

Model 12 is, in all ways, similar to model 6. The equations are listed previously. The difference will be in the data set fitted to model 12. It is an adjusted version of the original data set. The data set adjustement is necessary because the activity tests do not give unbiased estimates of the concentration of a cellulase enzyme component in a solution.

3.7.9 Model 13: Introducing E_2 in the Hydrolysis Kinetics

If there is indeed a third enzyme, exoglucanase, in the reactor and it cleaves off cellobiose units from the non-reducing end of a polymer, we have to account for this mathematically. Using our nomenclature, the action of E_2, exoglucanase is described as the transformation of β_{nre} bonds into G_2. In order not to increase



the number of parameters in the model, E_2 will be assumed to act at a much faster rate than E_x or E_1.

β_{nre} bonds are transformed into G_2. There has to be direct cleavage of a bond by E_2 in order to release a cellobiose molecule into solution. Three types of bonds can be contiguous to β_{nre}:

The extra terms to account for the presence of E_2 are the rates of direct cleavage

Equations of Model 11

$$\frac{dG_1}{dt} = [\beta_{re}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + [2G_2 + \beta_{nre} + \beta_{nres}][E_1 \cdot k_{1E_1}] \tag{104}$$

$$\frac{dG_2}{dt} = \left[\frac{\beta_u \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} + \frac{\beta_u \cdot \beta_{nre} + \beta_{re} \cdot \beta_{nre}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} - G_2\right][E_1 \cdot k_{1E_1}] \tag{105}$$

$$\frac{d\beta_u}{dt} = -\left[\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} + \frac{\beta_u + \beta_{re}}{\beta_u + \beta_{nre} + \beta_{sr}} + 1\right][\beta_u]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_u \cdot \beta_{nre}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{106}$$

$$\frac{d\beta_{re}}{dt} = \left[\frac{\beta_u \cdot (\beta_u + \beta_{re})}{\beta_u + \beta_{nre} + \beta_{sr}} - \beta_{re} - \frac{\beta_{re} \cdot \beta_u}{\beta_{re} + \beta_u + \beta_{snr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_{nre} \cdot \beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{107}$$

$$\frac{d\beta_{res}}{dt} = \left[\frac{\beta_u + \beta_{re}}{\beta_u + \beta_{nre} + \beta_{sr}}\right][\beta_{sr}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] \tag{108}$$

$$\frac{d\beta_{nre}}{dt} = \left[\frac{\beta_u \cdot \beta_u}{\beta_u + \beta_{re} + \beta_{snr}} - \frac{\beta_{nre} \cdot (\beta_u + \beta_{re})}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_u \cdot \beta_{nre}}{\beta_u + \beta_{re} + \beta_{snr}} - \beta_{nre}\right][E_1 \cdot k_{1E_1}] \tag{109}$$

$$\frac{d\beta_{nres}}{dt} = [\beta_u]\left[\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] + \left[\frac{\beta_{nre} \cdot \beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}} - \beta_{nres}\right][E_1 \cdot k_{1E_1}] \tag{110}$$

$$\frac{d\beta_{sr}}{dt} = -[\beta_u + \beta_{re}]\left[\frac{\beta_{sr}}{\beta_u + \beta_{nre} + \beta_{sr}}\right]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] \tag{111}$$

$$\frac{d\beta_{snr}}{dt} = -\left[\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right][\beta_u]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - \left[\frac{\beta_{nre} \cdot \beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}}\right][E_1 \cdot k_{1E_1}] \tag{112}$$

$$\frac{d\beta_{ss}}{dt} = 0 \quad \text{or} \quad \beta_{ss} = \beta_{ss}(0) \tag{113}$$

$$\frac{d\beta}{dt} = -[\beta_u + \beta_{re}]\left[\frac{E_x \cdot k_{1E_x}}{\beta}\right] - [G_2 + \beta_{nre} + \beta_{nres}][E_1 \cdot k_{1E_1}] \tag{114}$$

$$\mathrm{DP} = \frac{\beta(0) \cdot \mathrm{DP}(0)}{(\beta(0) - \beta) \cdot \mathrm{DP}(0) + \beta} \tag{115}$$

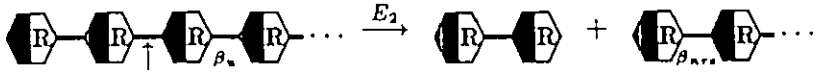
of β_u, β_{re} and β_{snr} by E_2.

$$d\beta_u = -\beta_{nre} \cdot \left(\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} \right)$$

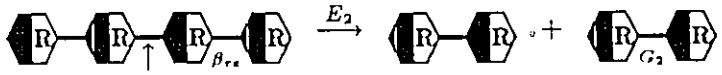
$$d\beta_{re} = -\beta_{nre} \cdot \left(\frac{\beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} \right)$$

$$d\beta_{snr} = -\beta_{nre} \cdot \left(\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}} \right)$$

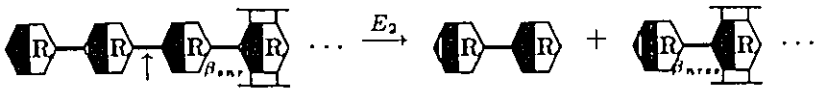
There are also transformations taking place when it was β_u that was cleaved by E_2. Then the bond next to β_u on its reducing side will be transformed into either β_{nre}, G_2 or β_{nres} from its initial state as a β_u or a β_{re} or a β_{snr}.



$$d\beta_{nre} = -d\beta_u = \beta_{nre} \cdot \left[\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} \right] \left[\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} \right] \qquad (116)$$



$$dG_2 = -d\beta_{re} = \beta_{nre} \cdot \left[\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} \right] \left[\frac{\beta_{re}}{\beta_u + \beta_{re} + \beta_{snr}} \right] \qquad (117)$$



$$d\beta_{nres} = -d\beta_{snr} = \beta_{nre} \cdot \left[\frac{\beta_u}{\beta_u + \beta_{re} + \beta_{snr}} \right] \left[\frac{\beta_{snr}}{\beta_u + \beta_{re} + \beta_{snr}} \right] \qquad (118)$$

Chapter 4

EXPERIMENTAL PROCEDURE

4.1 DS and Purity of the CMC

To determine the DS of carboxymethyl cellulose, the uranyl precipitation method, as described in Hercules bulletin FF-318 [49], was employed. The procedure, with minor modifications, is repeated here.

A sample of approximately 0.4 grams of CMC (let weight=g of sample) was weighed, moistened with anhydrous methanol and then dissolved in 400 mL of water. The temperature of the solution was brought up to 70°C. 50.0 mL of 2.5% uranyl nitrate solution were added, holding the pipette tip just above the liquid surface near the surface of the beaker. The sodium carboxymethyl cellulose precipitates at the solution pH of 3.5 to 4.0 as insoluble uranyl carboxymethyl cellulose, UCMC. Care must be taken when working with the slightly radioactive reagent uranyl nitrate. The 2.5% standard solution has a radioactivity count of 6 microcurries per litre. Note here that we used the 2.5% solution whereas Hercules procedures call for a 4.0% standard uranyl nitrate solution. However, in our case, 50.0 mL of standard solution were added per determination instead of Hercules' 25.0 mL. The extra uranyl nitrate added helped the extent to which the CMC precipitates. This is above the excess uranyl nitrate in the Hercules procedure.

After precipitation, the UCMC was washed with distilled water and ethanol to a filtering crucible. The UCMC was dried in an oven at 130°C, cooled in a desiccator and weighed. The residue's weight was recorded as g UCMC. As much of the UCMC as possible was transferred to a porcelain crucible and reweighed. This was recorded as 'g of UCMC used'. It was then covered and ignited at 800°C in

a furnace for 30 minutes till the combustion reaction to U_3O_8 was complete. The combustion residue was weighed ('g of U_3O_8'). The weight of U_3O_8 and the other measured data permit the calculation of both the purity of the CMC and its degree of substitution using the following formulae:

$$\mathrm{UF} = \frac{g \text{ of } U_3O_8 \cdot 0.962}{g \text{ of UCMC used}} \tag{119}$$

$$\mathrm{DS} = \frac{162 \cdot \mathrm{UF}}{135 - (192 \cdot \mathrm{UF})} \tag{120}$$

$$\left[\frac{\mathrm{CMC}}{\mathrm{UO_2(CMC)_2}}\right] = \frac{162 + (80 \cdot \mathrm{DS})}{162 + (192 \cdot \mathrm{DS})} \tag{121}$$

$$\text{Purity CMC}(\%) = \frac{g \,\mathrm{UCMC} \cdot \left[\frac{\mathrm{CMC}}{\mathrm{UO_2(CMC)_2}}\right]}{g \text{ of sample}} \cdot 100\% \tag{122}$$

4.2 Cellulase Activity

Determination of Cellulase Activity by Filter Paper Activity

The procedure of Mandels [32] has been modified for use by our biochemical engineering group at the University of Ottawa. The procedure of Mandels uses 0.5 mL of enzyme mixed with 1 mL of buffer whereas we use 1 mL of each. Also, the colorimetric determination of reducing sugars is carried out at a wavelength of 600 nm instead of the 550 nm of the Mandels procedure. The details of the procedure that is used in this research follow:

- Place 1 mL of diluted enzyme sample in a test tube
- Add 1 mL of 0.1 M citrate buffer, pH 4.8 and vortex
- Coil a 1 x 6 cm strip of filter paper no. 1
- Add coiled filter paper and vortex again
- Cover tubes with marbles and incubate at 50°C
- After 1 hour incubation, add 3 mL of 3,5-dinitrosalicylic acid reagent
- Boil tubes for 10 minutes
- Cool and centrifuge if there are suspended solids in solution

- Read absorbance at 600 nm against a blank

Substrate blanks are included for each sample to correct for the presence of reducing sugars. They are subjected to all of the above treatments except that no filter paper is added.

The blank standard is prepared with 1 mL distilled water, 1 mL buffer and 3 mL DNS reagent.

A series of standards containing between 0 and $1\,g/\mathrm{L}$ of glucose is treated the same way as the samples but without the inclusion of filter paper.

The filter paper activity is calculated as follows:

$$\mathrm{FPA} = (G_1 \text{ sample} - G_1 \text{ blank}) \cdot \mathrm{DF} \cdot 180/10.8 \qquad (123)$$

The units of FPA are μmol glucose/min/mL of enzyme solution. If the filter paper activity is a function of the dilution, use an interpolation or a modelling technique to calculate $\widehat{\mathrm{FPA}}$ at a DF that yields $0.667\,g/\mathrm{L}$ of glucose.

4.3 Cellobiase Activity

The test of Matteau et al. [20] has been considerably modified for our purposes because a chromatograph was at our disposition. Both the substrate used and the analysis method for the product concentration have been changed. Their test used salicin as a substrate and a colorimetric method to determine the rate of reaction for a given quantity of enzyme. The substrate used here is cellobiose. A colorimetric test responds to both G_1 and G_2 and hence is of little use with the cellobiose substrate. With a chromatograph (see following section), the extent of reaction is followed by measuring the uptake of cellobiose and/or the production of glucose.

- Prepare a solution of $10\,g/\mathrm{L}$ cellobiose in 0.1 M acetate buffer, pH 4.8
- Place 25.0 mL of the solution in a 250 mL erlenmeyer
- Heat to 50°Celsius

- Add 25.0 mL of diluted enzyme solution
- Incubate for one hour at 50°Celsius and a shaker speed Ω=100 rpm
- Boil for 15 minutes to quench the reaction
- Filter sample and run it through the chromatograph to determine the concentrations of glucose and cellobiose

NOTE. A cellobiose blank, where 25.0 mL of water has been added instead of the diluted enzyme, is included. The cellobiase activity is calculated as follows:

$$\text{CBA} = (G_1 \text{ sample} - G_1 \text{ blank}) \cdot \text{DF} \cdot 180/5.4 \qquad (124)$$

The units of CBA are μmol glucose/min/mL of enzyme solution. If the cellobiase activity is a function of the dilution, use an interpolation or a modelling technique to calculate $\widehat{\text{CBA}}$ at a DF that yields 0.667 g/L of glucose.

4.4 Hydrolysis of CMC

A 250 mL erlenmeyer was used as a batch reactor. To prepare for the hydrolysis run, the required amounts of CMC powder, citrate or acetate buffer, distilled water and enzyme were calculated. The total volume of the solution had to be known to calculate the concentrations. Because CMC is in powder form, its volume in solution was calculated from its weight, making use of the CMC bulking value in solution, 0.0652 gal/lb (=0.544 mL/g) [5].

The magnitude of the CMC bulking value was verified in three experiments, the details of which are not described here. At room temperature, samples were left sitting 10 days in a closed graduated cylinder.

Once the component amounts in the hydrolysis reactor were calculated, the water, buffer and CMC were mixed together at high speed for three minutes in a Waring blender. Then the solution was heated to 50°C in the shaker. The enzyme was then added to the erlenmeyer to start the hydrolysis reaction.

Table 4. *Bulking Value of CMC*

CMC (g)	Water (mL)	0.5 M Cit. buf. (mL)	Volume (10 days) (mL)	Bulking Value (mL/g)
5.13	100.0	20.0	123.0	0.58
3.53	100.0	20.0	121.7	0.48
1.91	100.0	20.0	121.1	0.58

The optimal pH for the hydrolysis reaction was 4.8. However as long as the pH does not vary by more than 0.03 away from its optimal, the enzyme shows little sensitivity to the pH.

Table 5. *pH of the Buffer Solutions*

Buffer Conc.	Citrate	Acetate
0.5 M	4.64	4.62
0.1 M	4.93	4.74

Table 5 shows some experimental measurements of the pH of the buffer stock and of diluted buffer. The pH was assayed to verify that the hydrolysis pH did not wander far from its optimal—at least initially.

Each erlenmeyer was one run location except for runs H29081–H29088 from which samples were withdrawn at irregular intervals. The withdrawn samples were boiled for 15 minutes immediately after withdrawl to deactivate the enzyme. Then each sample was filtered through a 0.45 micron filter and stored at 5°C. The sample was removed from cold storage as soon as possible to measure G_1 and G_2 by high performance liquid chromatography.

4.5 Concentrations of Glucose and Cellobiose

A Waters Associates liquid chomatograph with attached carbohydrate analysis column was used to determine the concentrations of glucose and cellobiose in the reaction mixture. This was the same method as used by Gritzali and Brown [9]. The chromatograph consisted of a Model 6000A solvent delivery system, a Model U6K injector, a Model 440 absorbance detector and a Model R401 differential refractometer connected to a Hewlett-Packard Model 3380S integrator.Three columns were available:

1. Waters' carbohydrate analysis column P/N 84038 S/N, serial no. T20811D 02
2. Waters' carbohydrate analysis column P/N 84038, serial no. T33181 S24.
3. Brownlee Labs' HPLC AMINO SPHERI-10 column, serial no. 1165, catalog no. H2-10A, tested at Brownlee labs Feb. 28, 1984.

The three columns reacted differently. Elution times for the glucose, cellobiose and for the citrate or the acetate buffer were not equal from one column to the next. Standards of glucose and cellobiose were injected between runs to constantly monitor the column efficiency.

The mobile phase is 80% acetonitrile and 20% water by volume. The Waters documentation strongly recommended against the use of citrate as a buffer in the mobile phase flow. Citrate has a chelating effect on the column packing material.

There was no citrate in our mobile phase. However, citrate was used as the buffer in the reactor until April 12th, 1985. The reaction mixture samples passing through the chromatograph contained citrate which eluted at 50 minutes. The late elution increased twofold the time necessary to analyze the samples hence an acetate buffer was used in the hydrolysis runs following April 12th, 1985.

Prior to March 1985, concentrations of the substances were calculated by comparing peak *heights* of the said substance to that of known standards. Starting on

March 14th, 1985, a Hewlett-Packard model 3380S integrator became available to do peak area measurements in the chromatograph runs. Thus from that date on, the glucose concentrations were measured by comparing both the peak height and the peak area for a sample to the peak height and area of the standards injected in the chromatograph. The cellobiose concentration however, could be calculated 'accurately' only when comparing peak heights of samples and standards. In most hydrolysis runs, the concentration of cellobiose is so low that the integrator cannot detect the peaks. For those runs no area measurement is available. But by using a low attenuation on the recorded value of the peak representing cellobiose, its height could be measured and compared to the height of the peaks of cellobiose standards.

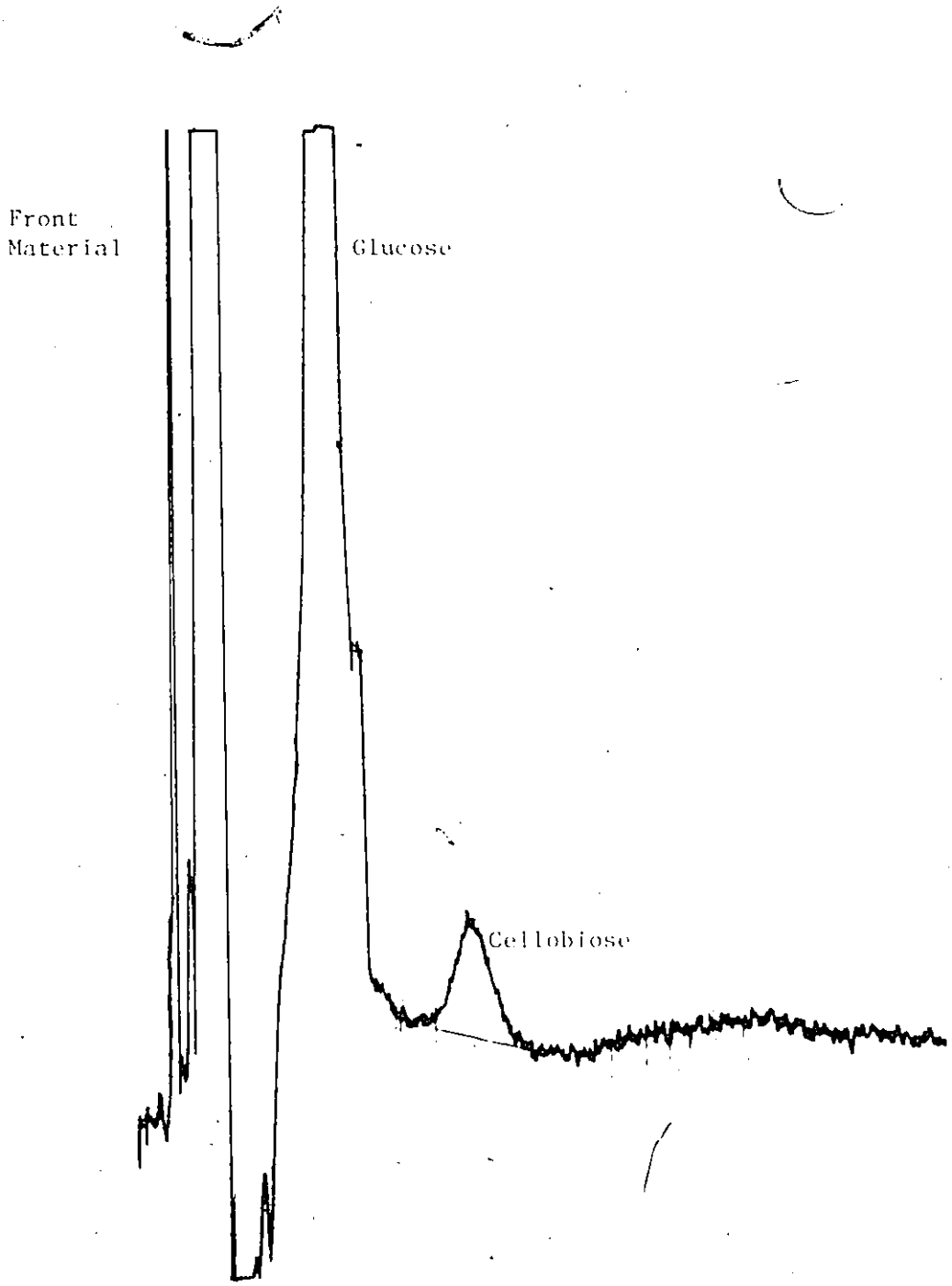


Figure 3: Low Attenuation Chromatogram: Run H90910

Chapter 5

RESULTS

5.1 Assay of the CMC Degree of Substitution and Purity

5.1.1 CMC Degree of Substitution

The degree of substitution of the CMC was determined according to the method of Hercules Inc. [49] described earlier. Table 6 lists the collected data and the calculation of DS and UF for each sample as per the procedure.

Table 6. *Raw Data and General Calculation of the Experimental CMC DS*

Obs	Run date	CMC type	*g* CMC	*g* UCMC	UCMC used	*g* U_3O_8	UF	DS
1	05SEP85	4H	0.47210	0.61343	0.37706	0.1078	0.27503	0.54208
2	09OCT85	4H	0.45400	0.57426	0.26210	0.0686	0.25179	0.47070
3	24OCT85	4H	0.39097	0.51393	0.43900	0.1223	0.26800	0.51968
4	03MAR85	7L	0.52902	0.72588	0.71908	0.25737	0.34432	0.80966
5	25MAR85	7L	0.49889	0.69031	0.69214	0.25679	0.35691	0.86982
6	05SEP85	7L	0.48186	0.65596	0.64722	0.22778	0.33856	0.78358
7	09OCT85	7L	0.47635	0.71085	0.66720	0.22739	0.32786	0.73717
8	24OCT85	7L	0.45886	0.65423	0.57111	0.19744	0.33258	0.75728
9	05SEP85	9M	0.48570	0.70430	0.69686	0.26938	0.37187	0.94722
10	09OCT85	9M	0.42220	0.65504	0.41654	0.15693	0.36243	0.89758
11	24OCT85	9M	0.49406	0.74877	0.46836	0.17635	0.36222	0.89650

The SAS [50] computer software package was useful in obtaining a statistical treatment of the data. The computer program that compiles the statistics associated with the determination of the DS is listed here:

Program 1. DS SAS

```
TITLE 'CALCULATION OF THE DEGREE OF SUBSTITUTION OF CARBOXYMETHYL CELLULOSE';
```

```
DATA DS_CALC;
    CMS FILEDEF MYDATA DISK DS DATA;
    INFILE MYDATA;
    FORMAT RUNDATE DATE7.;
    INPUT @ 7 RUNDATE DDMMYY6./* DATE: 2 DIGITS EACH FOR DA, NO, YR   */
          @14 CMCTYPE  $ 2.     /* CMC TYPE: 4H, 7L OR 9M              */
          @14 CMCDS      1.     /* CMC THEORETICAL DS: 4, 7 OR 9       */
          @15 CMCDP    $ 1.     /* CMC THEORETICAL DP: L, M OR H       */
          @17 GCMC       7.     /* WEIGHT OF CELLULOSE GUM: GRAMS      */
          @25 GUCMC      7.     /* WEIGHT OF URANYL CMC: GRAMS         */
          @33 UCMCUSED   7.     /* WEIGHT OF URANYL CMC USED: GRAMS    */
          @41 GU308      7.     /* WEIGHT OF URANIUM OXIDE: GRAMS      */;
    UF=GU308*0.962/UCMCUSED;
    DS=162.*UF/(135.-192.*UF);
PROC SORT DATA=DS_CALC;
    BY CMCDS;
PROC PRINT DATA=DS_CALC;
    BY CMCDS;
PROC MEANS DATA=DS_CALC;
    VAR UF DS;
    BY CMCDS;
    OUTPUT OUT=SUMMARY MEAN=MEANUF MEANDS;
PROC PRINT DATA=SUMMARY;
DATA PUR_CALC;
    MERGE DS_CALC SUMMARY;
    BY CMCDS;
    KEEP CMCDS GCMC GUCMC DS MEANDS CMC_UCMC CMCPURE RUNDATE;
    CMC_UCMC=(162.+80.*MEANDS)/(162.+192.*MEANDS);
    CMCPURE=GUCMC*CMC_UCMC*100./GCMC;
PROC PRINT DATA=PUR_CALC;
    TITLE 'CALCULATION OF THE PURITY OF CARBOXYMETHYL CELLULOSE';
    BY CMCDS;
PROC MEANS DATA=PUR_CALC;
    VAR CMCPURE;
    BY CMCDS;
```

The output from the DS SAS program is shown in table 7. The two variables of interest are the uranium fraction, UF, and the DS for each of the three types of CMC used. The best estimate of UF or DS is the mean value. It is to be used in subsequent calculations involving UF and DS in this work. The other entries in the table are variables representing the dispersion in the data.

The sample variance calculated for the DS of the three types of CMC can be found in table 7. Because the samples are subject to the same procedure, it is now possible to test if the three estimates of variance are not so statistically dissimilar that they can be considered to be equal. If it is plausible that the value of the ratio of the sample variances is "1", then pooling of the variances of the three types of

Table 7. *Statistical Treatment of the Experimental UF & DS Measurements*

	n	mean	$\hat{\sigma}$	Min	Max	Std error of mean	Sum	$\hat{\sigma}^2$	c.v.(%)
					CMC type 4H				
UF	3	0.26493	0.01192	0.25178	0.27503	0.006883	0.79481	0.00014212	4.500
DS	3	0.51081	0.03650	0.47069	0.54207	0.021076	1.53245	0.00133254	7.146
					CMC type 7L				
UF	5	0.34004	0.01128	0.32786	0.35691	0.005045	1.70022	0.00012728	3.318
DS	5	0.79150	0.05159	0.73717	0.86981	0.023075	3.95750	0.00266219	6.519
					CMC type 9M				
UF	3	0.36550	0.00551	0.36221	0.37187	0.003184	1.09652	0.00003040	1.509
DS	3	0.91376	0.02897	0.89649	0.94721	0.016730	2.74129	0.00083966	3.171

CMC will increase the accuracy of a test for the plausible values of the DS of the CMC.

The ratio $S_1^2\sigma_2^2/S_2^2\sigma_1^2$ has an F_{ν_1,ν_2} distribution. A confidence interval for the ratio σ_1^2/σ_2^2 is calculated as follows [51]:

$$1-\alpha = \text{Prob}\left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}_j^2 \cdot F_{\nu_i,\nu_j,\alpha/2}} \text{ to } \frac{\hat{\sigma}_i^2}{\hat{\sigma}_j^2 \cdot F_{\nu_i,\nu_j,1-\alpha/2}}\right) \tag{125}$$

where $\hat{\sigma}_i^2$ is the sample variance for sample i. The F-distribution is obtained from tables at the appropriate probability level α, and ν_i is the degrees of freedom associated with sample i. Here, σ_i^2 is the true but unknown value of the variance for the data considered. The sample variances for the CMC types 7L and 9M are put to the test because they have highest and lowest values respectively. The DS data for the CMC of type 7L also has more degrees of freedom than the other data. In our case, a 90% confidence interval for the ratio of the true sample variances is given by:

$$\left(\frac{\hat{\sigma}_7^2}{\hat{\sigma}_9^2 \cdot F_{\nu_7,\nu_9,\alpha/2}} < \frac{\sigma_{DS=7}}{\sigma_{DS=9}} < \frac{\hat{\sigma}_7^2}{\hat{\sigma}_9^2 \cdot F_{\nu_7,\nu_9,1-\alpha/2}}\right)_{\alpha=0.1} \tag{126}$$

Insert the applicable numbers, $\nu_7 = 4$, $\nu_9 = 2$, $\alpha = 0.1$

$$\left(\frac{0.0026622}{0.0008397 \cdot 19.20} < \frac{\sigma_{DS=7}}{\sigma_{DS=9}} < \frac{0.0026622}{0.0008397 \cdot 0.14409}\right) \tag{127}$$

$$0.165 < \frac{\sigma_{DS=7}}{\sigma_{DS=9}} < 22.0 \tag{128}$$

Thus the sample variances are not significantly different and they can be pooled to provide a better estimate of the true variance.

$$\hat{\sigma}_p^2 = \frac{\sum_i \nu_i \cdot \hat{\sigma}_i^2}{\sum_i \nu_i} \tag{129}$$

$$\hat{\sigma}_p^2 = \frac{2 \cdot 0.0013325 + 4 \cdot 0.0026622 + 2 \cdot 0.0008397}{2 + 4 + 2} \tag{130}$$

The pooled estimate of the variance of the DS is 0.001874. Calculate the 95% confidence interval for the DS values.

$$DS \pm t_{\nu,\alpha/2} \cdot \hat{\sigma}_p / \sqrt{n} \tag{131}$$

Here, ν is the degrees of freedom of the pooled variance. At 95% confidence, $t_{8,0.025} = 2.306$. With the 8 degrees of freedom of the pooled variance, the size of the confidence intervals is reduced—on average—compared to a procedure where the sample variances are left unpooled.

$$DS_{4H} \pm 2.306 \cdot \sqrt{0.001874/2}$$

$$DS_{7L} \pm 2.306 \cdot \sqrt{0.001874/4}$$

$$DS_{9M} \pm 2.306 \cdot \sqrt{0.001874/2}$$

$$\text{which comes out to } \begin{cases} DS_{4H} = 0.511 \pm 0.071 \\ DS_{7L} = 0.792 \pm 0.050 \\ DS_{9M} = 0.914 \pm 0.071 \end{cases}$$

The DS determined above are the independent variables in the modelling. Although least squares modelling techniques assume that the error associated with independent variables is virtually nil—i.e., the DS is perfectly known—in the hydrolysis system for CMC, there is a significant uncertainty concerning the true but unknown values of DS. Nonetheless, the standard least squares criteria are used to estimate parameters and their confidence.

Table 8. *General Calculation of the Experimental CMC Purity*

Obs	Run date	g CMC	g UCMC	DS	Mean DS	$\frac{CMC}{UO_2(CMC)_2}$	CMC purity
				CMC type 4H			
1	05SEP85	0.47210	0.61343	0.542075	0.510819	0.78002	101.353
2	09OCT85	0.45400	0.57426	0.470699	0.510819	0.78002	98.664
3	24OCT85	0.39097	0.51393	0.519683	0.510819	0.78002	102.534
				CMC type 7L			
4	03MAR85	0.52902	0.72588	0.809664	0.791502	0.71765	98.471
5	25MAR85	0.49889	0.69031	0.869817	0.791502	0.71765	99.301
6	05SEP85	0.48186	0.65596	0.783575	0.791502	0.71765	97.695
7	09OCT85	0.47635	0.71085	0.737170	0.791502	0.71765	107.094
8	24OCT85	0.45886	0.65423	0.757283	0.791502	0.71765	102.321
				CMC type 9M			
9	05SEP85	0.48570	0.70430	0.947219	0.913765	0.69671	101.029
10	09OCT85	0.42220	0.65504	0.897579	0.913765	0.69671	108.095
11	24OCT85	0.49406	0.74877	0.896498	0.913765	0.69671	105.590

Table 9. *Statistical Treatment of the Experimental CMC Purity Measurements*

n	Mean	$\hat{\sigma}$	Minimum	Maximum	Std error of mean	Sum	$\hat{\sigma}^2$	c.v.(%)
				CMC type 4H				
3	100.850	1.9832	98.663	102.53	1.1450	302.550	3.93323	1.967
				CMC type 7L				
5	100.976	3.8435	97.694	107.09	1.7188	504.881	14.7727	3.806
				CMC type 9M				
3	104.904	3.5825	101.028	108.09	2.0684	314.713	12.8350	3.415

5.1.2 CMC Purity

Tables 8 and 9 show the calculated values of the purity of the CMC and the statistical treatment of these data respectively.

Notice that the average experimental value of the purity of the CMC is greater than 100% for all types of CMC. Thus in our calculations, the purity of the CMC powder that was put into the reactor after drying was always assumed to be 100%.

5.2 Assay of the Enzyme Activities

Enzyme activity tests are used to estimate the concentration of various enzymes in a solution. The procedures specific to the two enzymes of interest to our work have been described in Chapter 4. Four activities had to be determined, one for each of the two enzymes, E_x and E_1, in each of the two commercial solutions, Celluclast and Novozym. It was known a *priori* that Celluclast contains a large concentration of E_x and a trace of E_1 while Novozym contains a large concentration of E_1 and a small concentration of E_x. These amounts of enzyme need to be evaluated as accurately as possible because they have a significant effect on the overall hydrolysis model parameter estimates.

From our survey of the literature, it was clear that there is room for improving the test procedures now in use. We have found that the amount of information extracted from the activity test data is insufficient due to poor reliability. By modifying the method of data analysis, the procedure of Mandels [32] can yield more information on the activity of an enzyme. Plots of the activity—FPA or CBA—as a function of enzyme dilution (DF), found in appendix 2, show that there is a trend of the activity with DF, even at a high DF. This may be specific to the DNS colorimetric method, as discussed earlier.

Consequently a strict modelling procedure was applied to the test data. The full details of the modelling procedure can be found in Appendix 2. From Appendix 2 the important information is the four empirical models that describe the behavior of FPA and CBA as a function of DF over the data range. Here are the four models deemed to best represent the activity data; activity units are μmoles gluc/min/mL.

Celluclast FPA=94.3+0.0048·DF

Novozym FPA=1.88+0.0103·DF

Celluclast CBA=5.88

Novozym CBA=$0.836 \cdot \mathrm{DF} \cdot (1 - e^{-2826/\mathrm{DF}})$

These models now were to estimate the enzyme activity at the standard conditions for Mandels's test, i.e., a glucose production of $0.667\,g/\mathrm{L}$ for a test of duration 1 hour.

To distinguish between the enzyme activities at standard and non-standard conditions, the nomenclature relative to enzyme activity tests must be modified somewhat. Let FPA and CBA be the enzyme activities calculated at each data point. The units of FPA and CBA are μmoles glucose/min/mL. FPA and CBA are functions of the DF of the starting commercial enzyme solution. Let $\widehat{\mathrm{FPA}}$ and $\widehat{\mathrm{CBA}}$ be the particular values of FPA and CBA that produce $0.667\,g/\mathrm{L}$ of glucose under the test conditions. Thus $\widehat{\mathrm{FPA}}$ and $\widehat{\mathrm{CBA}}$ are unique values of the activities for a particular enzyme solution, not a function of the DF. $\widehat{\mathrm{FPA}}$ and $\widehat{\mathrm{CBA}}$ become the estimated values of the concentration of E_x and E_1 respectively in a solution. The units of $\widehat{\mathrm{FPA}}$ and $\widehat{\mathrm{CBA}}$, in keeping with current enzyme methods, are I.U./mL, where I.U. is μmoles glucose/min, but is a unique value for an enzyme solution.

Now proceed to calculate $\widehat{\mathrm{FPA}}$ and $\widehat{\mathrm{CBA}}$ for Celluclast and Novozym.

5.2.1 Celluclast $\widehat{\mathrm{FPA}}$

The activity of the endoglucanase in the Celluclast commercial enzyme solution was tested as per the procedure described earlier. At various dilution factors (DF), the activity of the enzyme in hydrolyzing a solution, under conditions of the test, was measured. The raw data are listed in Table 10. The DF and the enzyme volume are calculated with respect to the undiluted, commercial enzyme solution.

The modelling procedure—the details of which can be found in Appendix 2—for this data set shows that there is a significant relationship between FPA and DF. The model chosen to describe this relationship is

$$\mathrm{FPA} = \theta_0 + \theta_1 \mathrm{DF}$$

Table 10. *Celluclast FPA Data*

Obs #	Run date da mo yr	Enzyme mL	DF	$G1$ g/L	FPA μmoles/min/mL
1	11SEP84	0.001000	1000	1.381	127.9
2	29JAN85	0.001000	1000	0.918	85.0
3	29JAN85	0.000500	2000	0.703	130.2
4	29JAN85	0.000200	5000	0.295	136.6
5	29JAN85	0.000100	10000	0.181	167.6
6	11APR85	0.001000	1000	1.147	106.2
7	11APR85	0.000500	2000	0.494	91.5
8	11APR85	0.000250	4000	0.282	104.4
9	11APR85	0.000125	8000	0.134	99.3
10	29JUN85	0.001000	1000	1.159	107.3
11	29JUN85	0.000500	2000	0.537	99.4
12	29JUN85	0.000250	4000	0.323	119.8
13	29JUN85	0.000125	8000	0.153	113.4
14	23OCT85	0.001000	1000	1.094	101.3
15	23OCT85	0.000500	2000	0.603	111.7
16	23OCT85	0.000250	4000	0.320	118.5
17	23OCT85	0.000125	8000	0.206	152.6
18	23NOV85	0.001000	1000	0.939	86.9
19	23NOV85	0.000500	2000	0.531	98.3
20	23NOV85	0.000250	4000	0.260	96.3
21	23NOV85	0.000125	8000	0.196	145.2
22	25NOV85	0.001000	1000	1.005	93.1
23	25NOV85	0.000500	2000	0.504	93.3
24	25NOV85	0.000250	4000	0.274	101.5
25	25NOV85	0.000125	8000	0.164	121.5

where $\theta_0 = 94.3 \pm 11.2$ and $\theta_1 = 0.0048 \pm 0.0024$

The confidence intervals are listed for a 95% significance level.

To calculate $\widehat{\text{FPA}}$, i.e., the FPA at a glucose production level of $0.667\,g/\text{L}$, a further relationship is required. By definition,

$$\text{FPA} = G1 \cdot \text{DF}/10.8$$

$$\widehat{\text{FPA}} = 0.667 \cdot \widehat{\text{DF}}/10.8$$

And using the second algebraic equation in the set of two, the model determined from the Celluclast FPA tests:

$$\text{FPA} = 94.3 + 0.0048 \cdot \text{DF}$$

Eliminating FPA by comparing the two equations yields

$$0.667 \cdot \widehat{\mathrm{DF}}/10.8 = 94.30 + 0.0048 \cdot \widehat{\mathrm{DF}}$$

$$\widehat{\mathrm{DF}} = 1656.$$

Calculating $\widehat{\mathrm{FPA}}$,

$$\widehat{\mathrm{FPA}} = 94.30 + 0.0048 \cdot 1656.$$

$$= 102.2 \text{ I.U./mL}$$

The precision of $\widehat{\mathrm{FPA}}$ may be calculated by taking the variance of both sides of the fitted model for FPA as a function of DF at $\widehat{\mathrm{DF}}$ [51]. Thus,

$$\mathrm{Var}(\widehat{\mathrm{FPA}}) = \mathrm{Var}(\theta_0) + \widehat{\mathrm{DF}}^2 \cdot \mathrm{Var}(\theta_1) + 2 \cdot \widehat{\mathrm{DF}} \cdot \mathrm{Cov}(\theta_0, \theta_1) \qquad (132)$$

where $\mathrm{Var}(\theta_0)$, $\mathrm{Var}(\theta_1)$ and $\mathrm{Cov}(\theta_0, \theta_1)$ are obtained from the variance-covariance matrix for the parameters.

Actually, solving equation 132 served as a verification that the SAS PROC REG linear fitting routine made use of conventional methods to calculate the precision of predicted values. The computer procedure does provide a predicted value of the dependent variable, with associated precision, for data points input to the program without a response. In the case of Celluclast $\widehat{\mathrm{FPA}}$, at a confidence level of 95%, $\widehat{\mathrm{FPA}} = 102.2 \pm 8.4$ I.U./mL. This value was used as the estimate of the concentration of E_x in Celluclast.

Table 11. *Novozym FPA Data*

Obs #	Rundate da mo yr	Enzyme mL	DF	$G1$ g/L	FPA μmoles/min/mL
1	29JUN85	0.01000	100	0.6048	5.60
2	23OCT85	0.01000	100	0.2052	1.90
3	23OCT85	0.00500	200	0.2322	4.30
4	23OCT85	0.00250	400	0.1566	5.80
5	23NOV85	0.01000	100	0.2041	1.89
6	23NOV85	0.00500	200	0.1771	3.28
7	23NOV85	0.00250	400	0.1939	7.18
8	25NOV85	0.00500	200	0.1960	3.63
9	25NOV85	0.00250	400	0.1310	4.85
10	25NOV85	0.00125	800	0.1391	10.30

5.2.2 Novozym $\widehat{\text{FPA}}$

The data in Table 11 was obtained in FPA tests of Novozym. A strict modelling procedure was applied to the data to determine if a trend between FPA and DF was significant. The modelling procedure can be found in Appendix 2.

As for the Celluclast\FPA test, there is statistical evidence that the FPA of the Novozym solution is also a function of the enzyme dilution. The model best describing that function and the 95% confidence intervals for the parameters:

$$\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}$$

$$\theta_0 = 1.88 \pm 1.58$$

$$\theta_1 = 0.0103 \pm 0.0044$$

To calculate the enzyme activity, solve the set of two equations relating FPA and DF. By definition, at the point at which $0.667\,g/\text{L}$ glucose is produced,

$$\text{FPA} = G1 \cdot \text{DF}/10.8$$

$$\widehat{\text{FPA}} = 0{:}667 \cdot \widehat{\text{DF}}/10.8$$

Compare the latter equation with the model, $\text{FPA} = 1.88 + 0.0103 \cdot \text{DF}$.

$$0.667 \cdot \widehat{\text{DF}}/10.8 = 1.88 + 0.0103 \cdot \widehat{\text{DF}}$$

$$\widehat{\text{DF}} = 36.5$$

At that dilution, the predicted value of FPA is $\widehat{\text{FPA}}$.

$$\begin{aligned}\widehat{\text{FPA}} &= 1.88 + 0.0103 \cdot 36.5 \\ &= 2.26 \pm 1.45\,\text{I.U./mL}\end{aligned}$$

The 95% confidence interval on $\widehat{\text{FPA}}$, the predicted FPA, was included in the output listing. $\widehat{\text{FPA}}$ of Novozym is an estimate of the concentration of E_x in Novozym.

To calculate $\widehat{\text{FPA}}$ it is necessary to extrapolate the model. It is important to note the severe experimental limitations in estimating $\widehat{\text{FPA}}$ of Novozym. The dilution range of the enzyme is limited at the upper range by the precision of the spectrophotometer. So little glucose is produced at dilutions greater than 500 that measured values would have to be considered unreliable. At the lower DF, the solution which contains the glucose is too concentrated—in enzyme. The sample is not clear enough to be within the measurement range of the spectrophotometer. Hence the model is extrapolated and the precision of the predicted FPA is low.

5.2.3 Novozym $\widehat{\text{CBA}}$

The enzyme activity of the cellobiase was tested as per the method described in the experimental procedure section of this thesis. As can be seen in Table 12, the range of DF tested was particularly large.

Table 12. *Novozym CBA Data*

Obs #	Run date da mo yr	Run #	Enzyme mL/mL	DF	$G1$ g/L	CBA μmoles/min/mL
1	22AUG85	C22808	0.00400000	250	5.35	247.7
2	22AUG85	C22809	0.00200000	500	5.12	474.1
3	22AUG85	C22810	0.00100000	1000	4.69	868.5
4	22AUG85	C22811	0.00050000	2000	3.40	1259.
5	22AUG85	C22812	0.00020000	5000	1.70	1574.
6	22AUG85	C22813	0.00010000	10000	0.88	1630.
7	22AUG85	C22814	0.00005000	20000	0.61	2259.
8	01OCT85	C02106	0.00020000	5000	2.12	1963.
9	01OCT85	C02107	0.00010000	10000	1.20	2222.
10	01OCT85	C02108	0.00005000	20000	0.60	2222.
11	01OCT85	C02109	0.00003333	30000	0.41	2278.
12	01OCT85	C02110	0.00002500	40000	0.33	2444.

A non-linear model was fitted to the data and found to provide a satisfactory fit. The full details of the model discrimination procedure can be found in Appendix 2. The non-linear model is

$$\text{CBA} = \theta_0 \cdot \text{DF} \cdot (1 - \exp(-\theta_1/\text{DF}))$$

$$\text{where } \theta_0 = 0.836 \pm 0.303$$

$$\text{and } \theta_1 = 2826 \pm 1180$$

The wide range of enzyme DF for this particular data set emphasizes the trend of CBA being a function of DF. And, as for the FPA test, the estimate of the concentration of E_1 in Novozym is obtained by comparing two algebraic equations. The first is the model and the second is the defining relationship between $G1$, DF and FPA,

$$\text{CBA} = G1 \cdot \text{DF}/5.4 \quad \text{where} \quad G1 = 0.667\, g/\text{L} \quad \text{at} \quad \widehat{DF}$$

Compare the two equations to eliminate CBA.

$$0.667 \cdot \widehat{\mathrm{DF}}/5.4 = 0.8359 \cdot \widehat{\mathrm{DF}} \cdot (1 - \exp(-2826/\widehat{\mathrm{DF}}))$$

$$\widehat{\mathrm{DF}} = 17680$$

Calculate $\widehat{\mathrm{CBA}}$ at this DF which corresponds to 0.667 g/L glucose produced under the test conditions.

$$\widehat{\mathrm{CBA}} = 0.8359 \cdot \mathrm{DF} \cdot (1 - \exp(-2827/17680))$$

$$= 2184 \pm 158 \text{ I.U./mL}$$

The confidence intervals for both the parameters of the non-linear model and $\widehat{\mathrm{CBA}}$ are taken directly from the output listing of the SAS NLIN procedure. Note that the confidence intervals for a non-linear model are not calculated with the same formula as for linear models. The linear models, in this work, described FPA behavior as a function of DF. A non-linear model is appropriate for a wide ranging DF because cryptic substrate limitation and product inhibition, both affecting the product concentration non-linearly, become important factors.

5.2.4 Celluclast $\widehat{\mathrm{CBA}}$

Both the cellobiose and glucose concentrations of the test samples were measured with the HPLC. The glucose concentration data listed in Table 13 were calculated by difference because the cellobiose measurements were considered more reliable than the glucose measurements. Thus the reacted cellobiose at each test condition was calculated from the knowledge of the initial cellobiose and the final cellobiose. All of the reacted cellobiose is transformed to glucose. For the details of the run conditions in collecting the cellobiase activity data, please consult the Raw Hydrolysis Data tables, parts 1 and 2 in Appendix 3.

We attempted fitting several models to the data in Table 13. The objective was to detect a trend of CBA as a function of the DF in the test. But for Celluclast, at the level of DF tested, no trend was found. Four points are considered unreliable

Table 13. *Celluclast CBA Data*

Obs #	Rundate da mo yr	Run #	Enzyme mL/mL	DF	$G1$ g/L	CBA μmoles/min/mL
1	22AUG85	C22801	0.0100	100	0.38	7.037
2	22AUG85	C22802	0.0050	200	0.08	2.963
3	22AUG85	C22803	0.0020	500	0.07	6.482
4	22AUG85	C22804	0.0010	1000	0.15	27.778
5	22AUG85	C22805	0.0010	1000	0.13	24.074
6	22AUG85	C22806	0.0005	2000	0.06	22.222
7	22AUG85	C22807	0.0002	5000	0.00	0.000
8	01OCT85	C02101	0.1000	10	2.15	3.982
9	01OCT85	C02102	0.0500	20	1.19	4.407
10	01OCT85	C02103	0.0200	50	0.53	4.907
11	01OCT85	C02104	0.0100	100	0.51	9.444
12	01OCT85	C02105	0.0050	200	0.21	7.778

because of a high dilution factor. The high DF leads to a released glucose concentration which is below the limit of detection of the chromatograph. The four points in questions are observations 4, 5, 6 and 7 in Table 13.

After eliminating the four data points for those runs where the enzyme was diluted 1000 times or more, the average CBA was calculated. This average becomes the estimate of the concentration of E_1 in Celluclast.

$$\widehat{\mathrm{CBA}} = 5.88 \pm 1.82\,\mathrm{I.U./mL}$$

The model discrimination procedure is carried out in full detail in Appendix 2.

The confidence interval for $\widehat{\mathrm{CBA}}$ is at a level of 95% for the predicted value of the mean. The theoretical $\widehat{\mathrm{DF}}$ which corresponds to the production of $0.667\,g$/L glucose in the CBA test can be calculated from

$$\widehat{\mathrm{DF}} = \widehat{\mathrm{CBA}} \cdot 5.4/G1$$

$$= 5.88 \cdot 5.4/G1$$

$$\widehat{\mathrm{DF}} = 47.6$$

5.2.5 Summary

Using a strict modelling procedure, the details of which are found in Appendix 2, the estimated concentrations of the two enzymes of interest, E_x and E_1 in the two commercial solutions have been calculated.

E_x in Celluclast= Celluclast $\widehat{\text{FPA}}$ =102.2±8.4 I.U./mL

E_x in Novozym = Novozym $\widehat{\text{FPA}}$ =2.26±1.45 I.U./mL

E_1 in Celluclast=Celluclast $\widehat{\text{CBA}}$ =5.88±1.82 I.U./mL

E_1 in Novozym = Novozym $\widehat{\text{CBA}}$ =2184±158 I.U./mL

5.3 The Hydrolysis Data Used for Modelling

The data used in our modelling studies of the hydrolysis of CMC are listed in table 14. In total, 189 runs were used.

In the table, the reaction time, t, the initial concentration of the glucose, the cellobiose, the cellobiase, the endoglucanase and the β-1,4 bond concentration contributed by the CMC are listed, as well as the concentration of the response variables.

Table 14. *Hydrolysis Data Used for Modelling*

Run #	t hours	$G_1(0)$ 10^{-2}M	$G_2(0)$ 10^{-2}M	E_1 I.U./L	E_x I.U./L	DS	DP_{CMC}	β_{CMC} 10^{-2}M	G_1 10^{-2}M	G_2 10^{-2}M
H0507A	90.0	0.0	0.0	$0.6184 \cdot 10^{+1}$	$0.1076 \cdot 10^{+3}$	0.792	400	15.58	1.272	0.196
H0507B	90.0	0.0	0.0	$0.1156 \cdot 10^{+4}$	$0.1088 \cdot 10^{+3}$	0.792	400	15.58	1.756	0.0
H0507C	90.0	0.0	0.0	$0.6184 \cdot 10^{+0}$	$0.1076 \cdot 10^{+2}$	0.792	400	15.58	0.694	0.173
H29081	2.9	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	0.772	0.0
H29081	5.3	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	0.867	0.044
H29081	12.8	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	0.822	0.0
H29081	27.6	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	0.933	0.079
H29081	51.7	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	0.911	0.029
H29081	141.3	0.0	0.0	$0.1130 \cdot 10^{+4}$	$0.1412 \cdot 10^{+3}$	0.792	400	7.09	1.556	0.029
H29082	2.0	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	0.606	0.149
H29082	4.1	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	0.556	0.199
H29082	12.8	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	0.639	0.111
H29082	27.6	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	0.756	0.079
H29082	51.7	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	0.817	0.061
H29082	141.3	0.0	0.0	$0.8048 \cdot 10^{+1}$	$0.1400 \cdot 10^{+3}$	0.792	400	7.09	1.089	0.0
H29083	2.0	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.422	0.173
H29083	4.1	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.467	0.222
H29083	12.8	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.567	0.044
H29083	27.6	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.756	0.096
H29083	51.7	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.661	0.0
H29083	141.3	0.0	0.0	$0.4024 \cdot 10^{+1}$	$0.7000 \cdot 10^{+2}$	0.792	400	7.09	0.767	0.0
H29084	2.0	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.55	0.0
H29084	4.1	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.628	0.0
H29084	12.8	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.772	0.135
H29084	27.6	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.667	0.0
H29084	51.7	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.756	0.061
H29084	141.3	0.0	0.0	$0.1122 \cdot 10^{+4}$	$0.1402 \cdot 10^{+3}$	0.792	400	5.33	0.856	0.061
H29085	2.0	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.372	0.129

Table 14. *(continued)*

Run #	t hours	$G_1(0)$ 10^{-3}M	$G_2(0)$ 10^{-3}M	E_1 I.U./L	E_x I.U./L	DS	DP_{CMC}	β_{CMC} 10^{-3}M	G_1 10^{-3}M	G_2 10^{-3}M
H29085	4.1	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.45	0.073
H29085	5.3	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.411	0.111
H29085	12.8	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.594	0.225
H29085	51.7	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.683	0.0
H29085	141.3	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.778	0.061
H29086	5.3	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.389	0.082
H29086	12.8	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.433	0.178
H29086	27.6	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.622	0.0
H29086	51.7	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.583	0.0
H29086	141.3	0.0	0.0	$0.7993 \cdot 10^{+1}$	$0.1390 \cdot 10^{+3}$	0.792	400	5.33	0.856	0.029
H29087	2.0	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.372	0.123
H29087	4.1	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.483	0.173
H29087	5.3	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.517	0.091
H29087	12.8	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.617	0.0
H29087	27.6	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.667	0.038
H29087	51.7	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	0.7	0.0
H29087	141.3	0.0	0.0	$0.1599 \cdot 10^{+2}$	$0.2781 \cdot 10^{+3}$	0.792	400	5.33	1.033	0.029
H29088	4.1	0.0	0.0	$0.1114 \cdot 10^{+4}$	$0.1153 \cdot 10^{+1}$	0.792	400	5.33	0.128	0.0
H29088	51.7	0.0	0.0	$0.1114 \cdot 10^{+4}$	$0.1153 \cdot 10^{+1}$	0.792	400	5.33	0.356	0.044
H29088	141.3	0.0	0.0	$0.1114 \cdot 10^{+4}$	$0.1153 \cdot 10^{+1}$	0.792	400	5.33	0.444	0.0
H21091	65.9	0.0	0.0	$0.1781 \cdot 10^{+5}$	$0.1843 \cdot 10^{+2}$	0.792	400	10.57	1.611	0.0
H21091	87.0	0.0	0.0	$0.1781 \cdot 10^{+5}$	$0.1843 \cdot 10^{+2}$	0.792	400	10.57	1.75	0.0
H21091	118.4	0.0	0.0	$0.1781 \cdot 10^{+5}$	$0.1843 \cdot 10^{+2}$	0.792	400	10.57	1.733	0.0
H14031	0.2	0.0	0.0	$0.1530 \cdot 10^{+1}$	$0.2661 \cdot 10^{+2}$	0.792	400	8.9	0.211	0.0
H14032	1.0	0.0	0.0	$0.1530 \cdot 10^{+1}$	$0.2661 \cdot 10^{+2}$	0.792	400	8.9	0.347	0.088
H14033	6.3	0.0	0.0	$0.1530 \cdot 10^{+1}$	$0.2661 \cdot 10^{+2}$	0.792	400	8.9	0.453	0.091
H14034	25.5	0.0	0.0	$0.1530 \cdot 10^{+1}$	$0.2661 \cdot 10^{+2}$	0.792	400	8.9	0.903	0.018
H14035	102.5	0.0	0.0	$0.1530 \cdot 10^{+1}$	$0.2661 \cdot 10^{+2}$	0.792	400	8.9	0.817	0.0
H14036	0.2	0.0	0.0	$0.1224 \cdot 10^{+2}$	$0.2129 \cdot 10^{+3}$	0.792	400	8.9	0.528	0.073
H14037	1.0	0.0	0.0	$0.1224 \cdot 10^{+2}$	$0.2129 \cdot 10^{+3}$	0.792	400	8.9	0.95	0.137
H14038	6.3	0.0	0.0	$0.1224 \cdot 10^{+2}$	$0.2129 \cdot 10^{+3}$	0.792	400	8.9	0.811	0.029
H14039	25.5	0.0	0.0	$0.1224 \cdot 10^{+2}$	$0.2129 \cdot 10^{+3}$	0.792	400	8.9	0.978	0.035
H21031	0.25	0.0	0.0	$0.1901 \cdot 10^{+1}$	$0.3307 \cdot 10^{+2}$	0.792	400	21.27	0.261	0.158
H21032	11.5	0.0	0.0	$0.1901 \cdot 10^{+1}$	$0.3307 \cdot 10^{+2}$	0.792	400	21.36	0.556	0.085
H21033	81.6	0.0	0.0	$0.1901 \cdot 10^{+1}$	$0.3307 \cdot 10^{+2}$	0.792	400	21.13	1.447	0.079
H21034	121.0	0.0	0.0	$0.1901 \cdot 10^{+1}$	$0.3307 \cdot 10^{+2}$	0.792	400	21.31	1.5	0.012
H21035	0.25	0.0	0.0	$0.7067 \cdot 10^{+3}$	$0.7313 \cdot 10^{+0}$	0.792	400	21.36	0.161	0.085
H21036	11.5	0.0	0.0	$0.7067 \cdot 10^{+3}$	$0.7313 \cdot 10^{+0}$	0.792	400	21.22	0.269	0.0
H21037	81.6	0.0	0.0	$0.7067 \cdot 10^{+3}$	$0.7313 \cdot 10^{+0}$	0.792	400	21.31	0.828	NR
H21038	121.0	0.0	0.0	$0.7067 \cdot 10^{+3}$	$0.7313 \cdot 10^{+0}$	0.792	400	21.22	0.853	0.0
H12041	0.8	0.0	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	0.678	0.269

Table 14. *(continued)*

Run #	t hours	$G_1(0)$ 10^{-3}M	$G_2(0)$ 10^{-3}M	E_1 I.U./L	E_2 I.U./L	DS	DP_{CMC}	β_{CMC} 10^{-3}M	G_1 10^{-3}M	G_2 10^{-3}M
H12042	3.8	0.0	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.467	0.254
H12043	16.5	0.0	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.256	0.187
H12044	39.3	0.0	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.552	0.099
H12045	54.0	0.0	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.78	0.053
H12046	0.8	0.594	0.0	$0.1230 \cdot 10^{+2}$	$0.2140 \cdot 10^{+3}$	0.792	400	22.43	1.361	0.281
H12047	3.8	0.594	0.0	$0.1230 \cdot 10^{+2}$	$0.2140 \cdot 10^{+3}$	0.792	400	22.43	2.239	0.269
H12048	16.5	0.594	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.807	0.091
H12049	39.3	0.589	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	3.041	0.064
H13041	54.0	0.589	0.0	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	3.656	0.041
H13042	0.8	0.0	0.313	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	0.786	0.36
H13043	3.8	0.0	0.319	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.094	0.313
H13044	16.5	0.0	0.316	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	2.856	0.082
H13045	39.3	0.0	0.335	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	3.528	0.076
H13046	54.0	0.0	0.33	$0.1237 \cdot 10^{+2}$	$0.2152 \cdot 10^{+3}$	0.792	400	22.43	3.363	0.07
H13047	0.8	0.0	0.0	$0.2474 \cdot 10^{+2}$	$0.4303 \cdot 10^{+3}$	0.792	400	22.43	1.072	0.275
H13048	16.5	0.0	0.0	$0.2474 \cdot 10^{+2}$	$0.4303 \cdot 10^{+3}$	0.792	400	22.43	2.359	0.137
H13049	54.0	0.0	0.0	$0.2474 \cdot 10^{+2}$	$0.4303 \cdot 10^{+3}$	0.792	400	22.43	3.056	0.0
C13041	3.8	0.0	0.301	$0.1175 \cdot 10^{+2}$	$0.2044 \cdot 10^{+3}$	0.0	0	0.0	0.204	0.19
C13042	39.3	0.0	0.275	$0.1175 \cdot 10^{+2}$	$0.2044 \cdot 10^{+3}$	0.0	0	0.0	0.653	0.044
H09501	1.0	0.0	0.0	$0.2463 \cdot 10^{+2}$	$0.4285 \cdot 10^{+3}$	0.792	400	15.49	0.744	0.187
H09502	3.9	0.0	0.0	$0.2463 \cdot 10^{+2}$	$0.4285 \cdot 10^{+3}$	0.792	400	15.49	1.644	0.345
H09503	12.0	0.0	0.0	$0.2463 \cdot 10^{+2}$	$0.4285 \cdot 10^{+3}$	0.792	400	15.49	1.626	0.146
H09504	27.7	0.0	0.0	$0.2463 \cdot 10^{+2}$	$0.4285 \cdot 10^{+3}$	0.792	400	15.49	2.148	0.354
H09505	120.0	0.0	0.0	$0.2463 \cdot 10^{+2}$	$0.4285 \cdot 10^{+3}$	0.792	400	15.49	2.304	0.19
H09506	1.0	0.0	0.0	$0.1232 \cdot 10^{+2}$	$0.2143 \cdot 10^{+3}$	0.792	400	15.49	0.786	0.222
H09507	3.9	0.0	0.0	$0.1232 \cdot 10^{+2}$	$0.2143 \cdot 10^{+3}$	0.792	400	15.49	1.719	0.287
H09508	12.0	0.0	0.0	$0.1232 \cdot 10^{+2}$	$0.2143 \cdot 10^{+3}$	0.792	400	15.49	1.611	0.275
H09509	27.7	0.0	0.0	$0.1232 \cdot 10^{+2}$	$0.2143 \cdot 10^{+3}$	0.792	400	15.49	1.939	0.161
H09510	120.0	0.0	0.0	$0.1232 \cdot 10^{+2}$	$0.2143 \cdot 10^{+3}$	0.792	400	15.49	2.42	0.19
H09511	1.0	0.0	0.0	$0.4579 \cdot 10^{+4}$	$0.4738 \cdot 10^{+1}$	0.792	400	15.49	0.394	0.026
H09512	3.9	0.0	0.0	$0.4579 \cdot 10^{+4}$	$0.4738 \cdot 10^{+1}$	0.792	400	15.49	1.019	0.0
H09513	12.0	0.0	0.0	$0.4579 \cdot 10^{+4}$	$0.4738 \cdot 10^{+1}$	0.792	400	15.49	1.106	0.015
H09514	27.7	0.0	0.0	$0.4579 \cdot 10^{+4}$	$0.4738 \cdot 10^{+1}$	0.792	400	15.49	1.072	NR
H09515	120.0	0.0	0.0	$0.4579 \cdot 10^{+4}$	$0.4738 \cdot 10^{+1}$	0.792	400	15.49	1.561	0.0
H01701	1.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	15.49	0.764	0.178
H01702	4.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	15.49	1.297	0.143
H01703	17.9	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	15.49	1.528	0.076
H01704	43.5	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	15.49	1.856	0.167
H01705	117.3	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	15.49	2.225	0.164
H01706	1.0	0.0	0.0	$0.1223 \cdot 10^{+1}$	$0.2127 \cdot 10^{+2}$	0.792	400	15.49	0.0	0.164
H01707	4.0	0.0	0.0	$0.1223 \cdot 10^{+1}$	$0.2127 \cdot 10^{+2}$	0.792	400	15.49	0.497	0.225

Table 14. *(continued)*

Run #	t hours	$G_1(0)$ 10^{-3}M	$G_2(0)$ 10^{-3}M	E_1 I.U./L	E_x I.U./L	DS	DP_{CMC}	β_{CMC} 10^{-3}M	G_1 10^{-3}M	G_2 10^{-3}M
H01708	17.9	0.0	0.0	$0.1223 \cdot 10^{+1}$	$0.2127 \cdot 10^{+2}$	0.792	400	15.49	1.033	0.213
H01709	43.5	0.0	0.0	$0.1223 \cdot 10^{+1}$	$0.2127 \cdot 10^{+2}$	0.792	400	15.49	1.203	0.108
H01710	117.3	0.0	0.0	$0.1223 \cdot 10^{+1}$	$0.2127 \cdot 10^{+2}$	0.792	400	15.49	1.374	0.076
H01711	1.0	0.0	0.0	$0.1223 \cdot 10^{+0}$	$0.2127 \cdot 10^{+1}$	0.792	400	15.49	0.078	0.015
H01712	4.0	0.0	0.0	$0.1223 \cdot 10^{+0}$	$0.2127 \cdot 10^{+1}$	0.792	400	15.49	0.044	0.0
H01713	17.9	0.0	0.0	$0.1223 \cdot 10^{+0}$	$0.2127 \cdot 10^{+1}$	0.792	400	15.49	0.161	0.178
H01714	43.5	0.0	0.0	$0.1223 \cdot 10^{+0}$	$0.2127 \cdot 10^{+1}$	0.792	400	15.49	0.259	0.269
H01715	117.3	0.0	0.0	$0.1223 \cdot 10^{+0}$	$0.2127 \cdot 10^{+1}$	0.792	400	15.49	0.283	0.287
H01716	17.9	0.0	0.0	$0.1223 \cdot 10^{-1}$	$0.2127 \cdot 10^{+0}$	0.792	400	15.49	0.003	0.0
H01717	43.5	0.0	0.0	$0.1223 \cdot 10^{-1}$	$0.2127 \cdot 10^{+0}$	0.792	400	15.49	0.017	0.006
H01718	117.3	0.0	0.0	$0.1223 \cdot 10^{-1}$	$0.2127 \cdot 10^{+0}$	0.792	400	15.49	0.044	0.0
H20701	1.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	0.824	0.392
H20702	4.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	1.656	0.24
H20703	4.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	1.7	0.535
H20704	16.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	2.225	0.398
H20705	45.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	2.806	0.535
H20706	124.0	0.0	0.0	$0.1223 \cdot 10^{+2}$	$0.2127 \cdot 10^{+3}$	0.792	400	23.26	3.947	NR
H20707	1.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	0.25	0.275
H20708	4.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	0.572	0.351
H20709	16.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	1.061	0.316
H20710	16.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	1.393	0.392
H20711	45.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	1.681	0.327
H20712	124.0	0.0	0.0	$0.1221 \cdot 10^{+1}$	$0.2123 \cdot 10^{+2}$	0.792	400	23.26	2.169	0.234
H20713	1.0	0.0	0.0	$0.1221 \cdot 10^{+0}$	$0.2123 \cdot 10^{+1}$	0.792	400	23.26	0.011	0.029
H20714	4.0	0.0	0.0	$0.1221 \cdot 10^{+0}$	$0.2123 \cdot 10^{+1}$	0.792	400	23.26	0.0	0.0
H20715	16.0	0.0	0.0	$0.1221 \cdot 10^{+0}$	$0.2123 \cdot 10^{+1}$	0.792	400	23.26	0.039	0.269
H20716	45.0	0.0	0.0	$0.1221 \cdot 10^{+0}$	$0.2123 \cdot 10^{+1}$	0.792	400	23.26	0.094	0.272
H20717	124.0	0.0	0.0	$0.1221 \cdot 10^{+0}$	$0.2123 \cdot 10^{+1}$	0.792	400	23.26	0.217	0.371
C22800	1.0	0.0	1.462	$0.0000 \cdot 10^{+0}$	$0.0000 \cdot 10^{+0}$	0.0	0	0.0	0.039	1.424
C22801	1.0	0.0	1.462	$0.2937 \cdot 10^{+2}$	$0.5110 \cdot 10^{+3}$	0.0	0	0.0	0.239	1.292
C22802	1.0	0.0	1.462	$0.1469 \cdot 10^{+2}$	$0.2555 \cdot 10^{+3}$	0.0	0	0.0	0.147	1.424
C22803	1.0	0.0	1.462	$0.5875 \cdot 10^{+1}$	$0.1022 \cdot 10^{+3}$	0.0	0	0.0	0.128	1.418
C22804	1.0	0.0	1.462	$0.2937 \cdot 10^{+1}$	$0.5110 \cdot 10^{+2}$	0.0	0	0.0	0.106	1.368
C22805	1.0	0.0	1.462	$0.2937 \cdot 10^{+1}$	$0.5110 \cdot 10^{+2}$	0.0	0	0.0	0.269	1.456
C22806	1.0	0.0	1.462	$0.1469 \cdot 10^{+1}$	$0.2555 \cdot 10^{+2}$	0.0	0	0.0	0.083	1.409
C22807	1.0	0.0	1.462	$0.5875 \cdot 10^{+0}$	$0.1022 \cdot 10^{+2}$	0.0	0	0.0	0.039	1.518
C22808	1.0	0.0	1.462	$0.4368 \cdot 10^{+4}$	$0.4520 \cdot 10^{+1}$	0.0	0	0.0	3.033	0.0
C22809	1.0	0.0	1.462	$0.2184 \cdot 10^{+4}$	$0.2260 \cdot 10^{+1}$	0.0	0	0.0	2.844	0.038
C22810	1.0	0.0	1.462	$0.1092 \cdot 10^{+4}$	$0.1130 \cdot 10^{+1}$	0.0	0	0.0	2.786	0.181
C22811	1.0	0.0	1.462	$0.5460 \cdot 10^{+3}$	$0.5650 \cdot 10^{+0}$	0.0	0	0.0	1.842	0.471
C22812	1.0	0.0	1.462	$0.2184 \cdot 10^{+3}$	$0.2260 \cdot 10^{+0}$	0.0	0	0.0	0.967	0.953

Table 14. *(continued)*

Run #	t hours	$G_1(0)$ 10^{-3}M	$G_2(0)$ 10^{-3}M	E_1 I.U./L	E_x I.U./L	DS	DP_{CMC}	β_{CMC} 10^{-3}M	G_1 10^{-3}M	G_2 10^{-3}M
C22813	1.0	0.0	1.462	$0.1092 \cdot 10^{+3}$	$0.1130 \cdot 10^{+0}$	0.0	0	0.0	0.617	1.225
C22814	1.0	0.0	1.462	$0.5460 \cdot 10^{+2}$	$0.5650 \cdot 10^{-1}$	0.0	0	0.0	0.378	1.254
H90901	1.0	0.0	0.0	$0.8004 \cdot 10^{+1}$	$0.1392 \cdot 10^{+3}$	0.511	3200	3.65	0.411	0.184
H90902	1.0	0.0	0.0	$0.8004 \cdot 10^{+1}$	$0.1392 \cdot 10^{+3}$	0.511	3200	3.65	0.497	0.222
H90903	4.0	0.0	0.0	$0.7519 \cdot 10^{+3}$	$0.1400 \cdot 10^{+3}$	0.511	3200	3.65	1.081	0.091
H90904	16.0	0.0	0.0	$0.8004 \cdot 10^{+1}$	$0.1392 \cdot 10^{+3}$	0.511	3200	3.65	0.847	0.199
H90905	120.0	0.0	0.0	$0.8004 \cdot 10^{+1}$	$0.1392 \cdot 10^{+3}$	0.511	3200	3.65	1.3	0.0
H90906	4.0	0.0	0.0	$0.2001 \cdot 10^{+1}$	$0.3481 \cdot 10^{+2}$	0.511	3200	3.65	0.433	0.211
H90907	16.0	0.0	0.0	$0.2001 \cdot 10^{+1}$	$0.3481 \cdot 10^{+2}$	0.511	3200	3.65	0.603	0.178
H90908	45.0	0.0	0.0	$0.7459 \cdot 10^{+3}$	$0.3558 \cdot 10^{+2}$	0.511	3200	3.65	1.122	0.0
H90909	45.0	0.0	0.0	$0.2001 \cdot 10^{+1}$	$0.3481 \cdot 10^{+2}$	0.511	3200	3.65	0.872	0.123
H90910	120.0	0.0	0.0	$0.2001 \cdot 10^{+1}$	$0.3481 \cdot 10^{+2}$	0.511	3200	3.65	1.036	0.088
H90911	4.0	0.0	0.0	$0.1236 \cdot 10^{+1}$	$0.2149 \cdot 10^{+2}$	0.792	400	15.49	0.606	0.272
H90912	16.0	0.0	0.0	$0.1236 \cdot 10^{+1}$	$0.2149 \cdot 10^{+2}$	0.792	400	15.49	0.797	0.213
H90913	1.0	0.0	0.0	$0.2471 \cdot 10^{+2}$	$0.4299 \cdot 10^{+3}$	0.792	400	15.49	1.119	0.155
H90914	45.0	0.0	0.0	$0.2471 \cdot 10^{+2}$	$0.4299 \cdot 10^{+3}$	0.792	400	15.49	2.244	0.0
H90915	1.0	0.0	0.0	$0.1236 \cdot 10^{+2}$	$0.2149 \cdot 10^{+3}$	0.792	400	15.49	0.878	0.088
H90916	4.0	0.0	0.0	$0.1236 \cdot 10^{+2}$	$0.2149 \cdot 10^{+3}$	0.792	400	15.49	1.272	0.225
H90917	16.0	0.0	0.0	$0.1161 \cdot 10^{+4}$	$0.2151 \cdot 10^{+3}$	0.792	400	15.49	2.025	0.0
H90918	45.0	0.0	0.0	$0.1236 \cdot 10^{+2}$	$0.2149 \cdot 10^{+3}$	0.792	400	15.49	2.031	0.104
H90919	120.0	0.0	0.0	$0.1236 \cdot 10^{+1}$	$0.2149 \cdot 10^{+2}$	0.792	400	15.49	1.531	0.33
H90920	120.0	0.0	0.0	$0.2471 \cdot 10^{+2}$	$0.4299 \cdot 10^{+3}$	0.792	400	15.49	2.883	NR
H90921	4.0	0.0	0.0	$0.1194 \cdot 10^{+2}$	$0.2077 \cdot 10^{+3}$	0.914	1100	25.92	1.044	0.225
H90922	16.0	0.0	0.0	$0.1151 \cdot 10^{+2}$	$0.2002 \cdot 10^{+3}$	0.914	1100	25.92	1.489	0.272
H90923	45.0	0.0	0.0	$0.1304 \cdot 10^{+2}$	$0.2269 \cdot 10^{+3}$	0.914	1100	25.92	1.856	0.272
H90924	45.0	0.0	0.0	$0.3152 \cdot 10^{+1}$	$0.5483 \cdot 10^{+2}$	0.914	1100	25.92	1.261	0.181
H90925	122.0	0.0	0.0	$0.2778 \cdot 10^{+2}$	$0.4832 \cdot 10^{+3}$	0.914	1100	25.92	2.125	0.12
C02100	1.0	0.0	1.462	$0.0000 \cdot 10^{+0}$	$0.0000 \cdot 10^{+0}$	0.0	0	0.0	0.044	1.491
C02101	1.0	0.0	1.462	$0.2938 \cdot 10^{+3}$	$0.5110 \cdot 10^{+4}$	0.0	0	0.0	4.45	0.868
C02102	1.0	0.0	1.462	$0.1469 \cdot 10^{+3}$	$0.2555 \cdot 10^{+4}$	0.0	0	0.0	3.353	1.132
C02103	1.0	0.0	1.462	$0.5875 \cdot 10^{+2}$	$0.1022 \cdot 10^{+4}$	0.0	0	0.0	1.342	1.316
C02104	1.0	0.0	1.462	$0.2938 \cdot 10^{+2}$	$0.5110 \cdot 10^{+3}$	0.0	0	0.0	0.625	1.322
C02105	1.0	0.0	1.462	$0.1469 \cdot 10^{+2}$	$0.2555 \cdot 10^{+3}$	0.0	0	0.0	0.239	1.404
C02106	1.0	0.0	1.462	$0.2184 \cdot 10^{+3}$	$0.2260 \cdot 10^{+0}$	0.0	0	0.0	1.225	0.836
C02107	1.0	0.0	1.462	$0.1092 \cdot 10^{+3}$	$0.1130 \cdot 10^{+0}$	0.0	0	0.0	0.711	1.12
C02108	1.0	0.0	1.462	$0.5460 \cdot 10^{+2}$	$0.5650 \cdot 10^{-1}$	0.0	0	0.0	0.381	1.345
C02109	1.0	0.0	1.462	$0.3625 \cdot 10^{+2}$	$0.3752 \cdot 10^{-1}$	0.0	0	0.0	0.272	1.371
C02110	1.0	0.0	1.462	$0.2752 \cdot 10^{+2}$	$0.2848 \cdot 10^{-1}$	0.0	0	0.0	0.228	1.442

5.4 Experimental Evidence Supporting Reese's Theory

Reese [17] has theorized that endoglucanase splits bonds between two unsubstituted units and that β-glucosidase splits off glucose at the non-reducing end of the chain. The validity of this mechanism can be tested, in part, by comparing theoretical and experimental values of the the ultimate concentration of glucose obtained in hydrolyzing CMC.

Runs at three DSs were examined. These runs, with a high enzyme concentration and long reaction time, were chosen since it could be safely assumed that the reaction had proceeded to completion. The experimental results are summarized below:

DS = 0.511

Runs H90905 and H90908.

Initial $\beta_{CMC} = 0.0365$ M.

Glucose produced = 0.01300 and 0.01122 M respectively for an averaged estimate of 0.01211 M.

DS = 0.792

Runs H09505, H09510 and H90920.

Initial $\beta_{CMC} = 0.1549$ M.

Glucose produced = 0.02304, 0.02420 and 0.02883 M respectively for an average of 0.02536 M.

DS = 0.914

Run H90925.

Initial $\beta_{CMC} = 0.2592$ M.

Glucose produced = 0.02125 M.

The approach for DS = 0.511 is now presented.

Assuming Timell's distribution of the substituents:

$$c_0 = (1 - \tfrac{1}{2} \cdot \mathrm{DS})^2$$
$$= (1 - \tfrac{1}{2} \cdot (0.511))^2$$
$$c_0 = 0.5543$$

Assuming Random distribution, $Se_2 = Se_3 = Se_6$:

$$c_0 = (1 - \tfrac{1}{3} \cdot \mathrm{DS})^3$$
$$= (1 - \tfrac{1}{3} \cdot (0.511))^3$$
$$c_0 = 0.5711$$

Note that the choice for a distribution of the substituents on the polymer has more effect on c_0—the fraction of anhydroglucosidic units which are unsubstituted—at a higher DS.

If we assume that only bonds between unsubstituted units are cleaved, then any unsubstituted unit initially between 2 unsubstituted units will, at reaction completion, have become a glucose molecule.

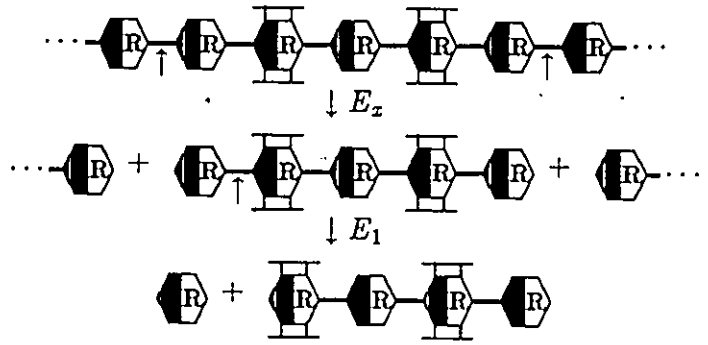
$$\text{Moles of anhyd. units} = \beta_{\mathrm{CMC}} \cdot \frac{\mathrm{DP}}{\mathrm{DP} - 1}$$
$$= 0.0365 \cdot \frac{3200}{3199}$$
$$= 0.0365\,\mathrm{M}$$

The probability that a mid-molecule unit is unsubstituted and surrounded by unsubstituted units on both sides is c_0^3. The probability that an end unit is unsubstituted and the unit next to it is unsubstituted is c_0^2.

Thus glucose at completion of reaction for Timell's distribution is

$$\hat{G}_1(\infty) = \text{units} \cdot \left(\frac{\mathrm{DP} - 2}{\mathrm{DP}} \cdot c_0^3 + \frac{2}{\mathrm{DP}} \cdot c_0^2 \right)$$
$$= 0.0365 \cdot \left(\frac{3198}{3200} \cdot 0.5543^3 + \frac{2}{3200} \cdot 0.5543^2 \right)$$
$$= 0.00622\,\mathrm{M}$$

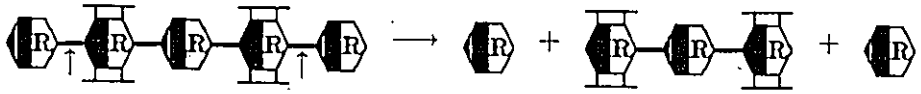
If we assume the validity of Reese's mechanism that unsubstituted units at the non-reducing end of the polymer are cleaved even though the contiguous unit may be substituted, then the expected glucose concentration will be higher. An unsubstituted non-reducing end unit will end up as glucose. All other units on the initial CMC polymer will end up as glucose at $t = \infty$ if they are unsubstituted and



the unit on their non-reducing side is also unsubstituted. The transformation of those eligible units is a two-step reaction. First the bonds are cleaved randomly along the polymer chain. Secondly end bonds are cleaved, throwing glucose into solution.

$$\begin{aligned}\hat{G}_1(\infty) &= \text{units} \cdot \left(\frac{\text{DP}-1}{\text{DP}} \cdot c_0^2 + \frac{1}{\text{DP}} \cdot c_0\right) \\ &= 0.0365 \left(\frac{3199}{3200} \cdot 0.5543^2 + \frac{1}{3200} \cdot 0.5543\right) \\ &= 0.01122\,\text{M}\end{aligned}$$

Now if we assume that unsubstituted units at both the reducing and non-reducing ends of the polymer chain are cleaved, more glucose would be expected at the completion of the reaction.



- End units of the initial polymer are transformed into glucose if they are unsubstituted.
- All other unsubstituted units of the initial polymer are to end up as glucose if any contiguous unit is unsubstituted.

$$\begin{aligned}\hat{G}_1(\infty) &= \text{units} \cdot \left[\frac{2}{\text{DP}} \cdot c_0 + \frac{\text{DP}-2}{\text{DP}} \cdot (c_0 \cdot (c_0 + (1-c_0) \cdot c_0))\right] \\ &= 0.0365 \left[\frac{2}{3200} \cdot 0.5543 + \frac{3198}{3200} \cdot 0.5543^2 \cdot (1+1-0.5543)\right] \\ &= 0.01622\,\text{M}\end{aligned}$$

The final case we are considering is when all unsubstituted units on the initial polymer are assumed to end up as glucose. Then the predicted maximum glucose

concentration is calculated from c_0 directly.

$$\begin{aligned}\hat{G}_1(\infty) &= \text{units} \cdot c_0 \\ &= 0.0365 \cdot 0.5543 \\ &= 0.02023\end{aligned}$$

At the three DSs, the same calculations expressing the theoretical $\hat{G}_1(\infty)$ were made for this, Timell's distribution, and for a random distribution. The results are listed in Table 15.

Table 15. *Comparison of $\hat{G}_1(\infty)$ According to Several Mechanisms*

	DS = 0.511	DS = 0.792	DS = 0.914
Initial β_{CMC}	0.0365	0.1549	0.2592
Experimental $G_1(\infty)$	0.01211	0.02536	0.02125
Bonds between unsubstituted units only are cleaved:			
Timell's distribution	0.00622	0.00761	0.00688
$Se_2 = Se_3 = Se_6 = 1/3$	0.00680	0.00992	0.00989
Reese's mechanism:			
Timell's distribution	0.01122	0.02076	0.02260
$Se_2 = Se_3 = Se_6 = 1/3$	0.01190	0.02478	0.02937
All unsubstituted end units end up as glucose:			
Timell's distribution	0.01622	0.03391	0.03853
$Se_2 = Se_3 = Se_6 = 1/3$	0.01702	0.03964	0.04886
All unsubstituted units end up as glucose:			
Timell's distribution	0.02024	0.05665	0.07649
$Se_2 = Se_3 = Se_6 = 1/3$	0.02085	0.06191	0.08722

Comparing the experimental data to all the other lines of the table shows that Reese's mechanism comes closest to predicting the amount of glucose that can be obtained from the enzymatic hydrolysis of CMC.

The samples chosen at DS = 0.511 and 0.792 have average measured amounts of glucose slightly greater than the maximum theoretically produceable under the assumption of Reese's mechanism. This may be explained by two things. The

variance of the experimental G_1 is high; these runs are still close enough to be considered similar to the predicted glucose production from Reese's mechanism. In the second place, the etherification reaction of the cellulose to produce CMC may not have been perfect. In other words the substituent groups may have bunched up along the polymer chain. This would increase the fraction of unsubstituted anhydroglucose units in the chain. Hence the glucose yield at reaction completion would be higher. We speculate that this unpredictable distribution is more likely for the 7L type CMC because it is a manufacturing grade cellulose gum.

The distribution of the substituents has a minor effect on the expected concentration of glucose at reaction completion when compared to the variation with assumed mechanism. It is worthy to note that the hydrolysis kinetics will be affected more by changes in proposed mechanisms than by a substituent group distribution in error.

5.5 Modelling Procedure

In this work, 189 observations of the CMC hydrolysis reaction were obtained from experiments described earlier. Then we proposed several different mechanisms for the manner of action of the endoglucanase and the cellobiase. Each of the proposals led to a model consisting of simultaneous non-linear differential equations.

The next step in the modelling of the enzymatic hydrolysis of CMC consisted of estimating the parameters, i.e., the kinetic constants. This step included the discrimination procedure used to decide which mechanism best represents the hydrolysis reaction.

Interpretation of plots of the residuals and estimation of confidence intervals for the parameters is part of a good approach to modelling. The parameter precision tells us whether a parameter should be tested further for its influence on the model or whether the variance in the data is too high to extract values for the parameter.

The residual plots were necessary for several reasons in our case:

1. Variables such as shaker speed, volume of the reactor, buffer type, buffer concentration etc... are assumed to have little effect on the reaction. Plots of the residuals as a function of these variables will confirm (or deny) this.
2. Model inadequacies may show up in the form of trends in the residuals with respect to the responses or the independent variables. Wrong assumptions concerning the mechanism can be detected and corrected.
3. Data points which appear as outliers can be detected and dealt with appropriately.

5.5.1 The Estimation Problem

To evaluate the proposed mechanisms, the models developed earlier must be fitted to the 189 data obtained from the hydrolysis experiments. This involves the estimation of the various kinetic constants in the model, least squares estimation being employed. The extent of the estimation problem is exemplified in considering the basic Michaelis-Menten model. Here the model consists of nine non-linear ordinary differential equations containing the rate parameters k_{1E_1}, K_{mE_1}, k_{1E_2} and K_{mE_2}. At each iteration of the non-linear least squares search, this set of equations must be solved. The need for an accurate, efficient solution algorithm for these differential equations is clear.

5.5.2 Solving the Set of Non-Linear Differential Equations

Many methods are available to solve a set of nine simultaneous non-linear ordinary differential equations. However the analytical solution is not one of those methods. The solution to these equations has to be generated numerically.

Euler's method and the fourth-order Runge-Kutta method were compared. The integration step size was reduced until there was no more change in the dependent variable (G_1). At equal step sizes, the Runge-Kutta solution was more accurate

than the Euler method's solution, as expected. However for the same amount of cpu time, the Euler answer was closer to the true value. Maximum accuracy with minimal computing time was the objective in choosing the integration method to be used. Truncation error was of little consequence as the program converts all data to, and does all calculations with, extended precision numbers. In general the Euler method was found to be more computationally efficient and provided accurate solutions. Consequently it was employed throughout the work.

Each proposed mechanism for the kinetics leads to a different model. The numerical integration is programmed in a different subroutine for each model. Whether it is desired to simulate the hydrolysis or to obtain the least squares estimates (LSE) of the parameters, the integration subroutine for a given model is appended to the main program. For example, the subroutine containing the integration routine for model 13 has been appended to the source listing of NLINIMSL FORTRAN in the appendices.

5.5.3 On Estimating the LSE of the Parameters

SAS software was used to generate the LSE of the parameters when the enzyme activities were modelled. Originally, it was intended to use SAS to calculate least squares estimates of the parameters and related statistical information for the overall hydrolysis model. However SAS programming consumes comparatively large amounts of computer resources [52]. To run the non-linear parameter estimation program with SAS would likely have taken several cpu hours.

For this reason, two IMSL routines were integrated with the main program NLINIMSL FORTRAN. The IMSL routine ZXSSQ locates the minimum SSR using a finite difference Levenberg-Marquardt algorithm. IMSL routine LINV3P inverts the J^TJ matrix. The covariance matrix of the parameters is estimated from $(J^TJ)^{-1}$. All statistical information could thus be calculated.

5.5.4 Hardware

The University of Ottawa has currently 2 mainframe Amdahl 470/V8 computers used for batch processing. The second Amdahl can be used to submit queued jobs written in `FORTRAN` or `WATFIV` for execution at a discount rate. Both mainframe Amdahl V8s have VM/CMS as their operating system. The execution of `NLINIMSL FORTRAN` was conducted at night and on weekends on the VM batch subsystem of the second Amdahl to minimize computing costs. This was necessary because the complete execution of `NLINIMSL FORTRAN` used between 5 and 30 minutes of central processing unit time.

5.6 History of Hydrolysis Reaction Modelling

5.6.1 Summary of Fitted Model Forms

The models examined in this study are summarized below:

1. Model 1
   - Basic Michaelis-Menten kinetics, with no inhibition, has 4 parameters
2. Model 2
   - Model 1 with added assumption that E_x does not cleave or form a complex with G_2
   - 4 parameters
3. Model 3
   - Model 2 with added assumption that $K_{mE_x} = 0$
   - 3 parameters
4. Model 4
   - A kinetic model modified to accomodate the observations from fitting models 1–3...
   - E_x does not cleave or form a complex with G_2
   - $K_{mE_x} = 0$ i.e., endoglucanase kinetics are a zero-order reaction
   - $K_{mE_1} \gg G_2 + \beta_{nrc} + \beta_{nres}$ i.e., cellobiase kinetics are first-order
   - 2 parameters
5. Model 5
   - Model 4 with added assumption that G_1 is a reversible inhibitor of E_1, no distinction being made between competitive and non-competitive inhibition
   - 3 parameters

6. Model 6
    - Model 4 with added assumption that all β-1,4 glucosidic bonds form a complex with E_x. Thus non-substrate β bonds are inhibitors of E_x
    - 2 parameters
7. Model 7
    - The data is only the 28 runs with no initial CMC
    - Fits a first-order kinetic model to the cellobiase activity
    - E_1 is the only enzyme cleaving G_2
    - 1 parameter
8. Model 8
    - This expands on model 7 to assume Michaelis-Menten kinetics instead of first-order kinetics for E_1
    - 2 parameters
9. Model 9
    - Model 8 with added assumption that G_1 is a reversible non-competitive inhibitor of E_1
    - 3 parameters
10. Model 10
    - Model 8 with added assumption that G_1 is a reversible competitive inhibitor of E_1
    - 3 parameters

11. Model 11

    - All 189 hydrolysis data points are included
    - The $\widehat{\text{CBA}}$ of Celluclast has been multiplied fivefold as a correction.
    - β_u and β_{re} are the E_x substrates
    - G_2, β_{nre} and β_{nres} are the E_1 substrates
    - Non-substrate β-1,4 bonds are inhibitors of E_x
    - $K_{mE_x} = 0$ i.e., endoglucanase kinetics are a zero-order reaction
    - $K_{mE_1} \gg G_2 + \beta_{nre} + \beta_{nres}$ i.e., cellobiase kinetics are first-order
    - 2 parameters

12. Model 12

    - This is model 6 fitted to the corrected—multiplied fivefold—cellobiase activity of Celluclast

13. Model 13

    - Model 12 with 1 addition ...
    - E_2 in the reactor is assumed to split off G_2 from the non-reducing end of the polymer chain
    - E_2 activity is assumed to be much faster than E_1 or E_x action so that β_{nre} bonds in the reactor are *almost* instantaneously transformed into G_2

5.6.2 Model 1 Fitting

Model 1 is the basic Michaelis-Menten model with no inhibition. As previously stated, the IMSL library routine ZXSSQ was used to obtain the least squares' estimates of the parameter. In this and all modelling, several integration step sizes were tried to ensure that there was no effect of step size on the sum of squared residuals and/or the least squares values of the parameters. As well, several initial estimates of the parameters were tried to ensure that the non-linear routine had converged to a global, and not a local, minimum.

In carrying out the parameter estimation, severe convergence problems were encountered. Depending upon the initial guesses for the parameters the integration step size and the convergence criterion the estimation routine failed to converge, converged at physically meaningless values or at different, possibly meaningful values.

The limiting factor of computer resources was ever present. In most cases, the estimate of the Hessian of the objective function was not positive definite thus reliable estimates of the precision of the parameter estimates could not be obtained. All the evidence lead to the suspicion that insufficient information was available to estimate the parameters. Nevertheless, the results for the best fit were examined and are summarized below:

MODEL 1

Non-linear routine specifications: $\Delta t = 100\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_s}	$0.5 \cdot 10^{-7}$	$0.2276 \cdot 10^{+0}$	$\pm 0.2276 \cdot 10^{-1}$
K_{mE_s}	$0.5 \cdot 10^{-1}$	$0.2785 \cdot 10^{+6}$	$\pm 0.1106 \cdot 10^{-8}$
k_{1E_i}	$0.5 \cdot 10^{-7}$	$0.5569 \cdot 10^{-2}$	$\pm 0.5798 \cdot 10^{-1}$
K_{mE_i}	$0.5 \cdot 10^{+1}$	$0.5729 \cdot 10^{+4}$	$\pm 0.5637 \cdot 10^{-7}$
Sum of squared residuals $= 0.003760\,\mathrm{M}^2$			

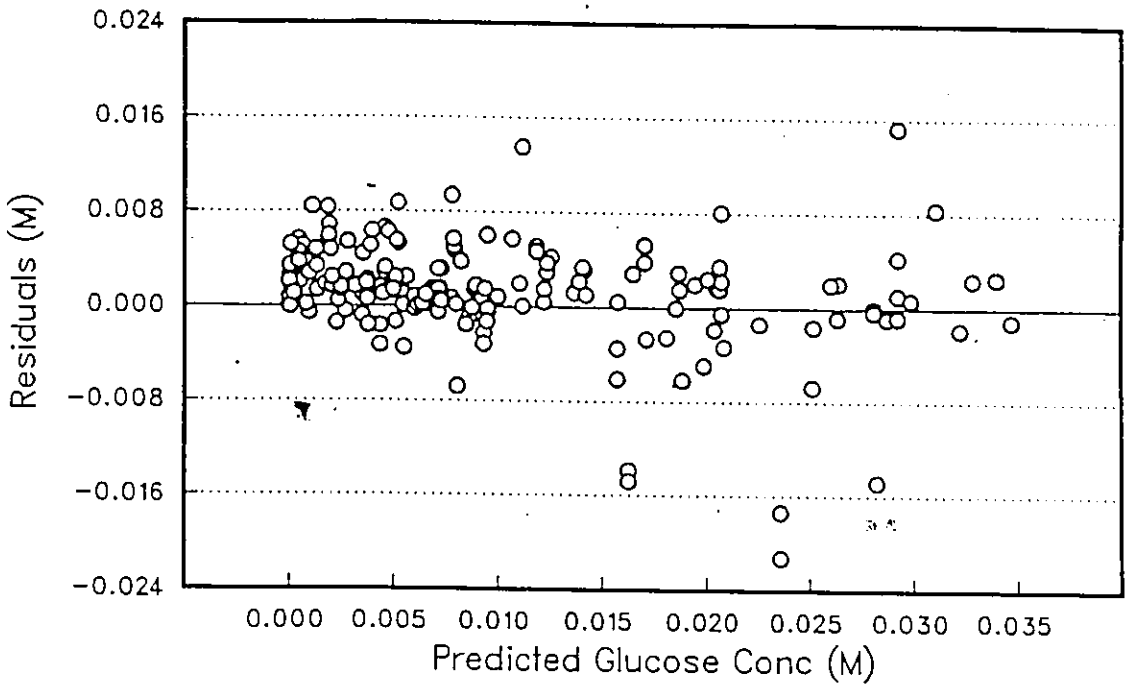
The covariance matrix of the parameters:

$$\begin{pmatrix} +0.13346 \cdot 10^{-03} & & & \\ -0.25299 \cdot 10^{-11} & +0.31490 \cdot 10^{-18} & & \\ +0.38029 \cdot 10^{-05} & -0.31084 \cdot 10^{-11} & +0.86633 \cdot 10^{-03} & \\ -0.36967 \cdot 10^{-11} & +0.30216 \cdot 10^{-17} & -0.84214 \cdot 10^{-09} & +0.81863 \cdot 10^{-15} \end{pmatrix}$$

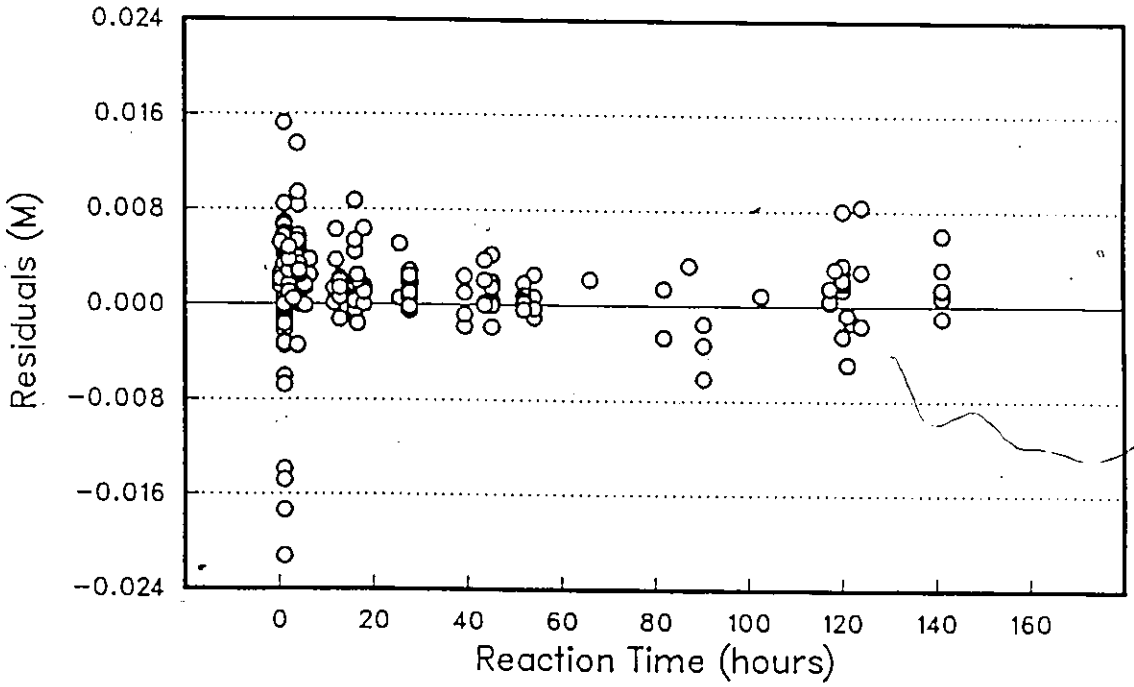
The correlation matrix of the parameters $= \begin{pmatrix} 1.000 & & & \\ -0.390 & 1.000 & & \\ 0.011 & -0.188 & 1.000 & \\ -0.011 & 0.188 & -1.000 & 1.000 \end{pmatrix}$

The 189 residuals for these parameter values of model 1 have been plotted as a function of 15 of the variables of interest in the hydrolysis reaction. Those residual plots are found in Figure 4. Plots of the residuals as a function of the two buffer types and buffer concentration indicate that they do not affect the enzyme activity. Figure 4(k), the residuals as a function of the date on which the runs were performed does not show any trends hence it is probable that the enzymes have not deactivated significantly in storage at 5°C. Shaker speed and reaction mixture volume are also insignificant variables at the levels tested. The remaining plots are used to detect model inadequacy yet little is shown unless outliers are examined.

Upon examination of the outliers, it was found useful to plot the residuals of runs with only cellobiose as the substrate. The trend shown in Figure 5 indicates model inadequacy. Furthermore the plots of observed versus predicted values of G_1 and G_2, Figures 6 and 7 reveal that the glucose concentrations are comparable but the predicted cellobiose concentrations are all low. This could indicate that E_x is incapable of cleaving cellobiose. This should logically improve the model adequacy because the CBA test results are based on this assumption. We cannot assume in the CBA test that E_x does not cleave G_2, then, in the overall hydrolysis model assume that E_x cleaves G_2.

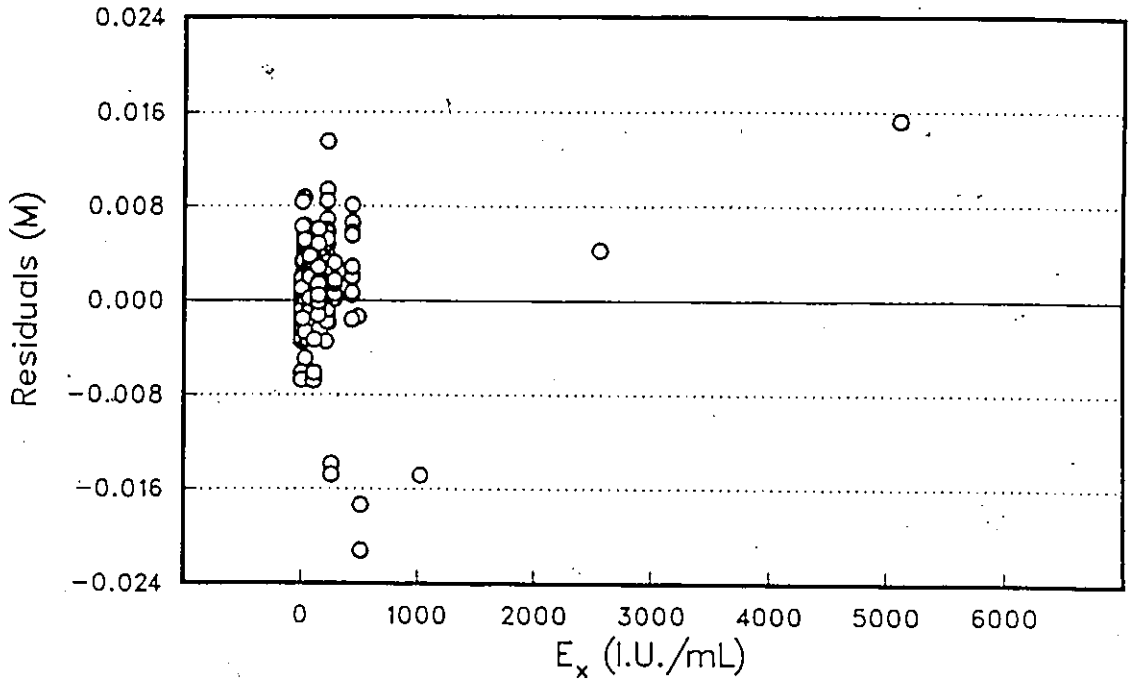


(a) Residuals vs Predicted Glucose

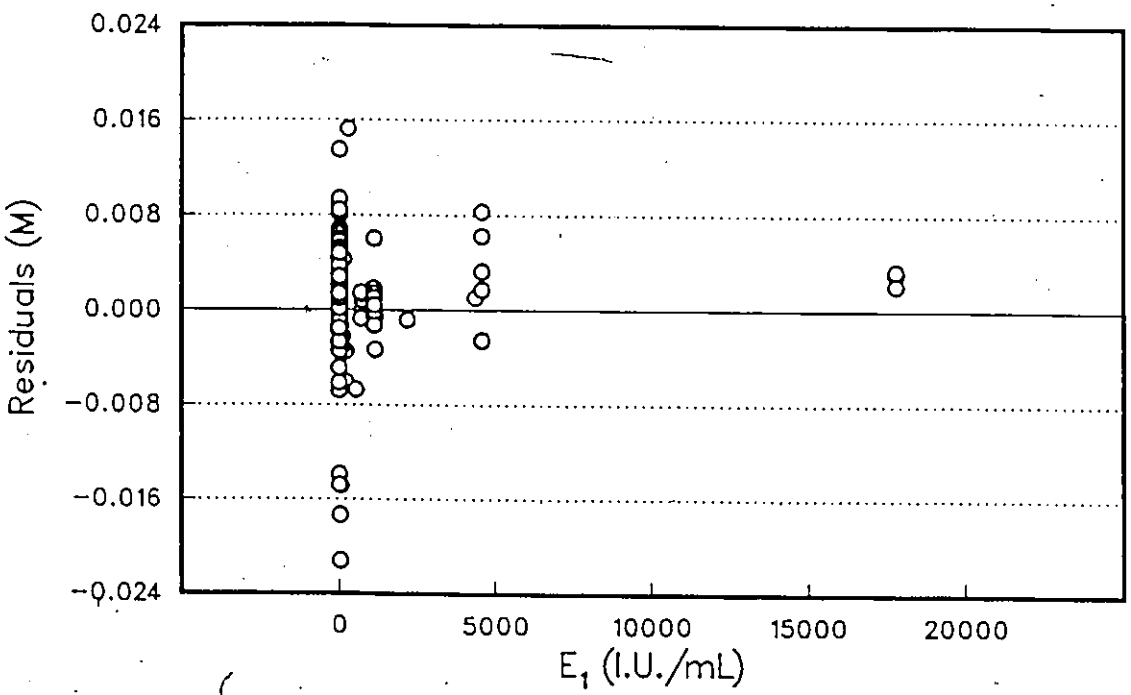


(b) Residuals vs Reaction Time

Figure 4: Model 1 Residuals

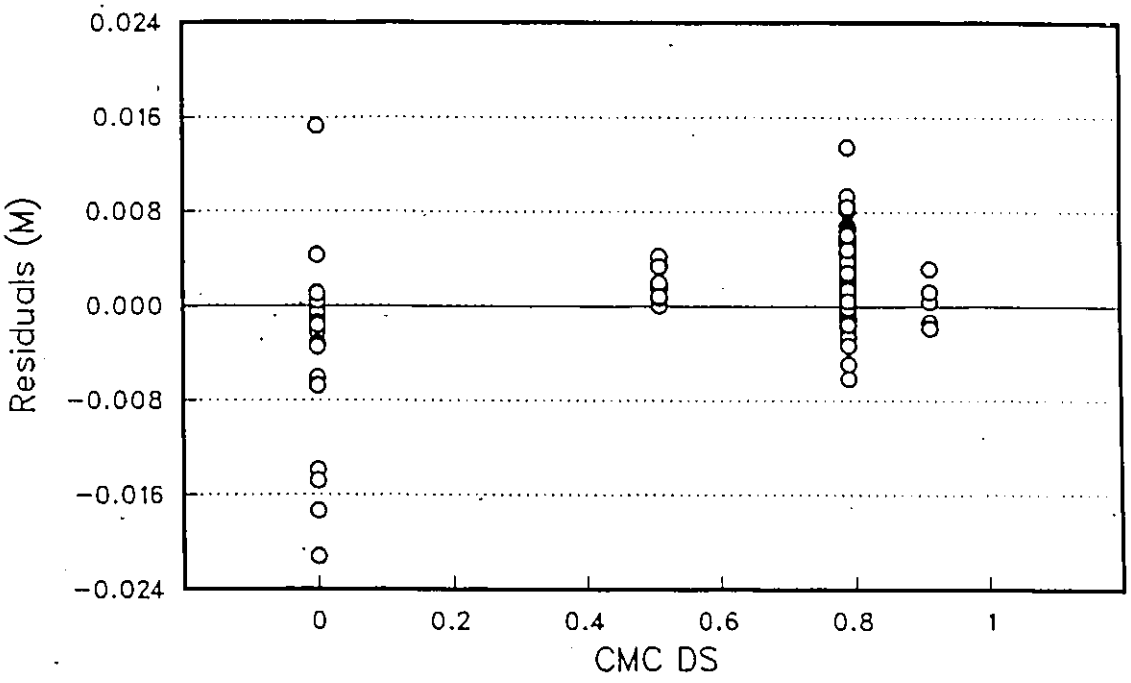


(c) Residuals vs E_x

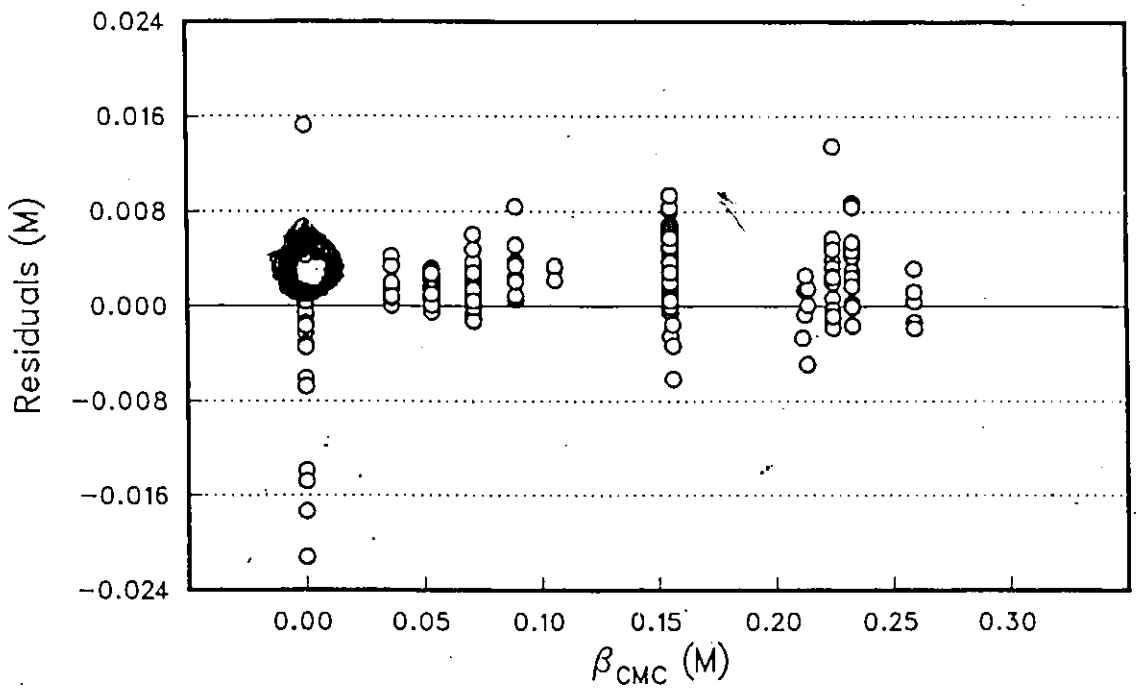


(d) Residuals vs E_1

Figure 4: Model 1 Residuals

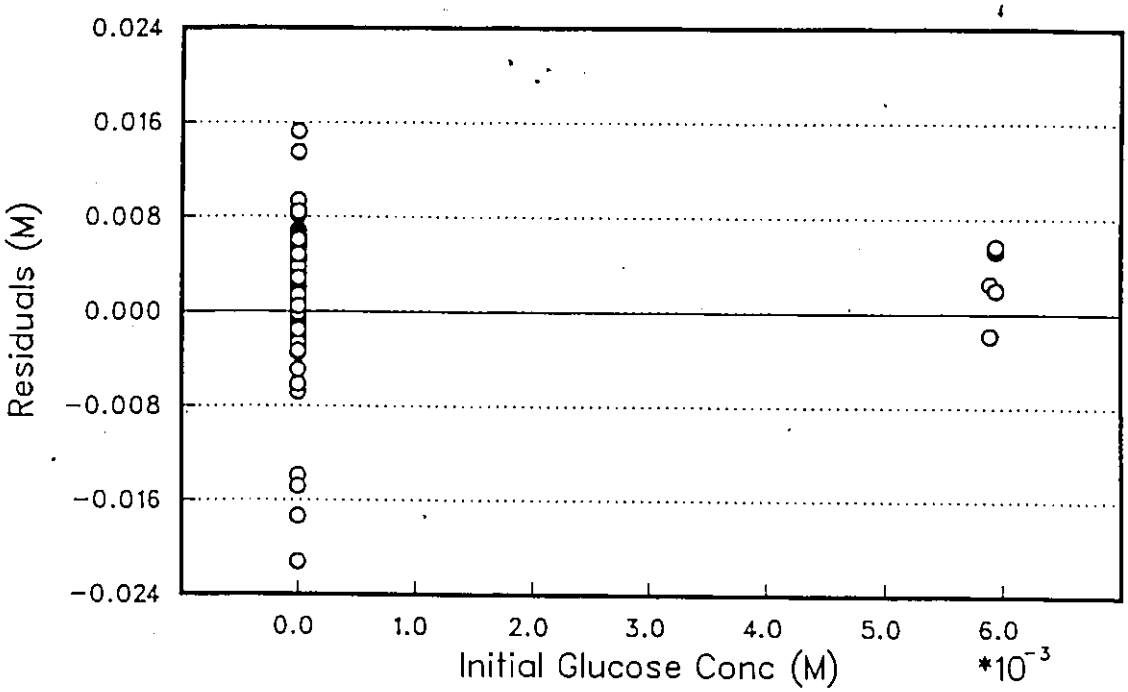


(e) Residuals vs CMC DS

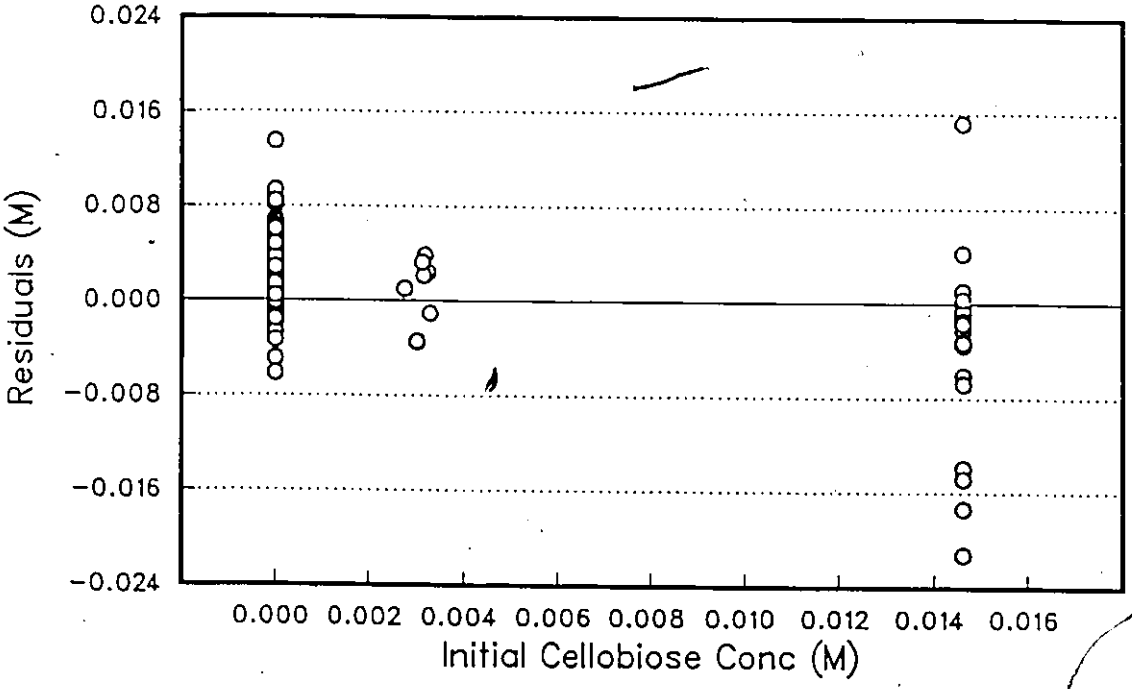


(f) Residuals vs β_{CMC}

Figure 4: Model 1 Residuals

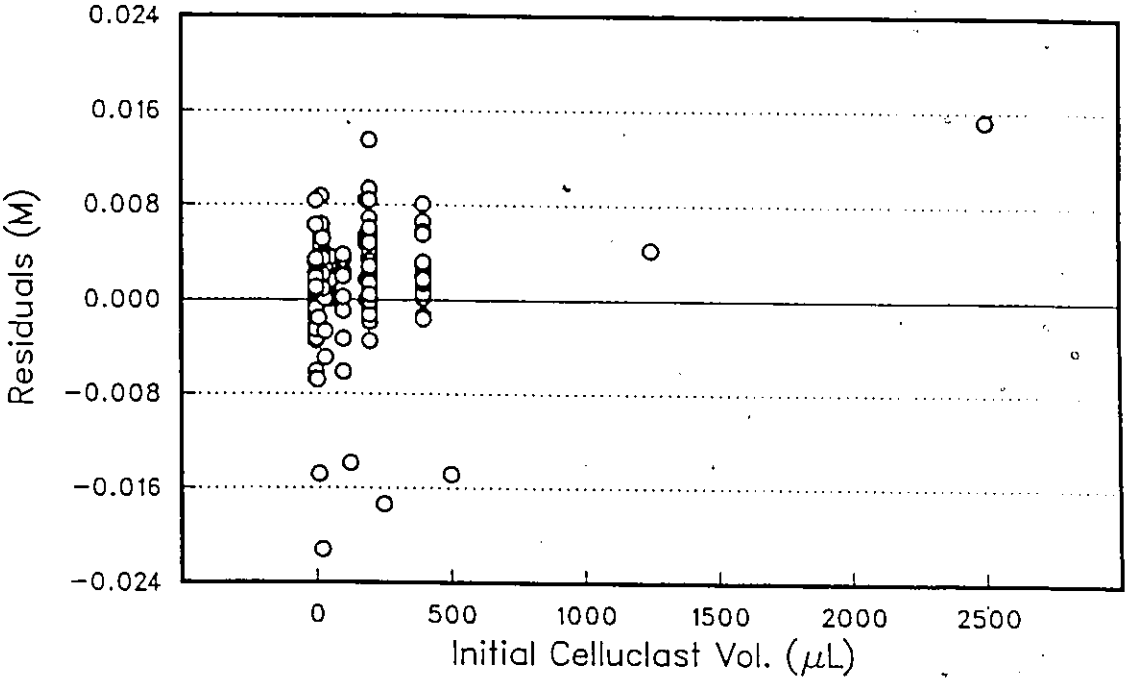


(g) Residuals vs Initial Glucose Concentration

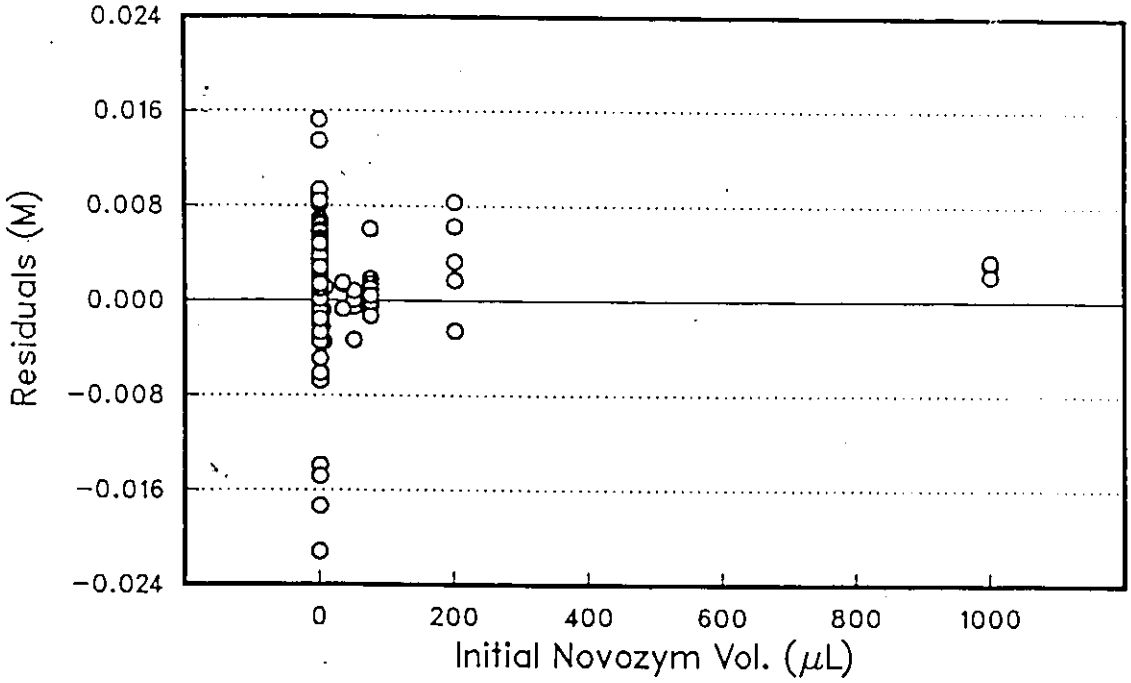


(h) Residuals vs Initial Cellobiose Conc.

Figure 4: Model 1 Residuals

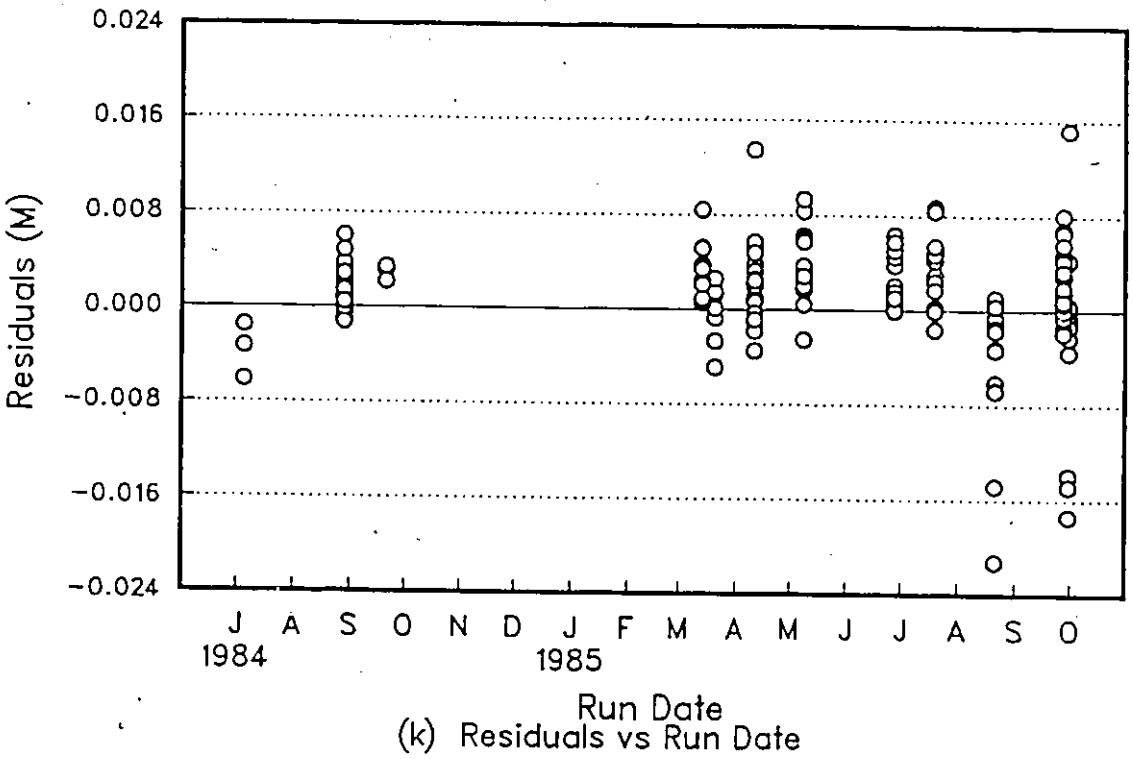


(i) Residuals vs Initial Celluclast Vol.



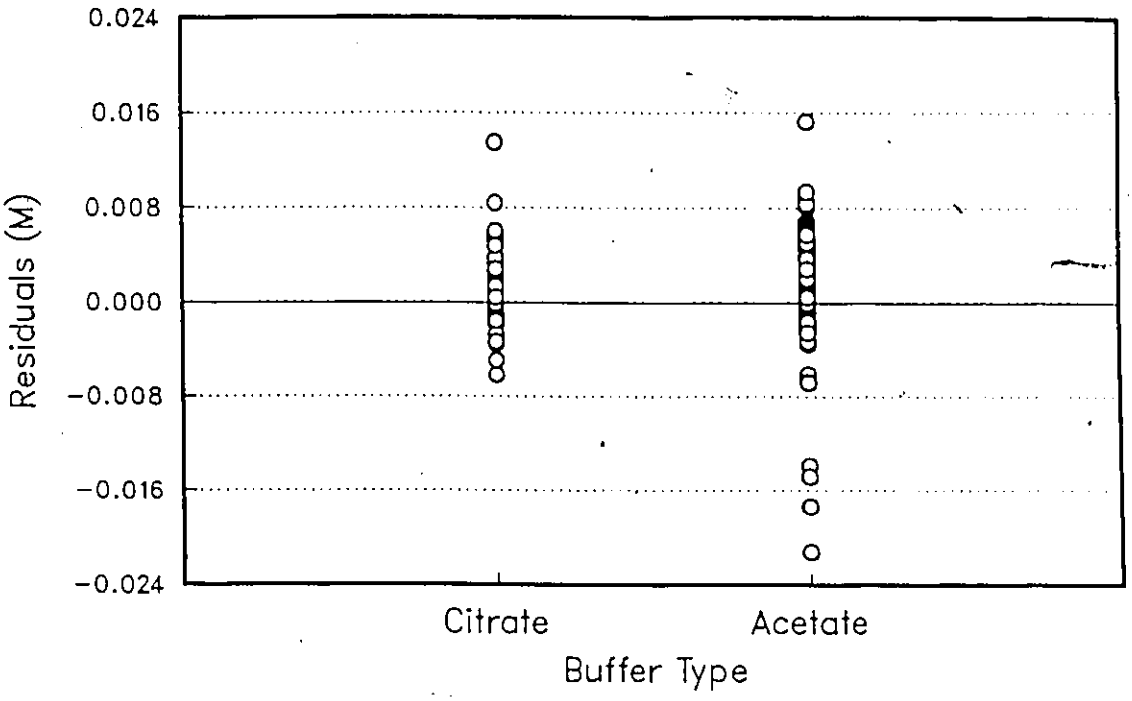
(j) Residuals vs Initial Novozym Vol.

Figure 4: Model 1 Residuals

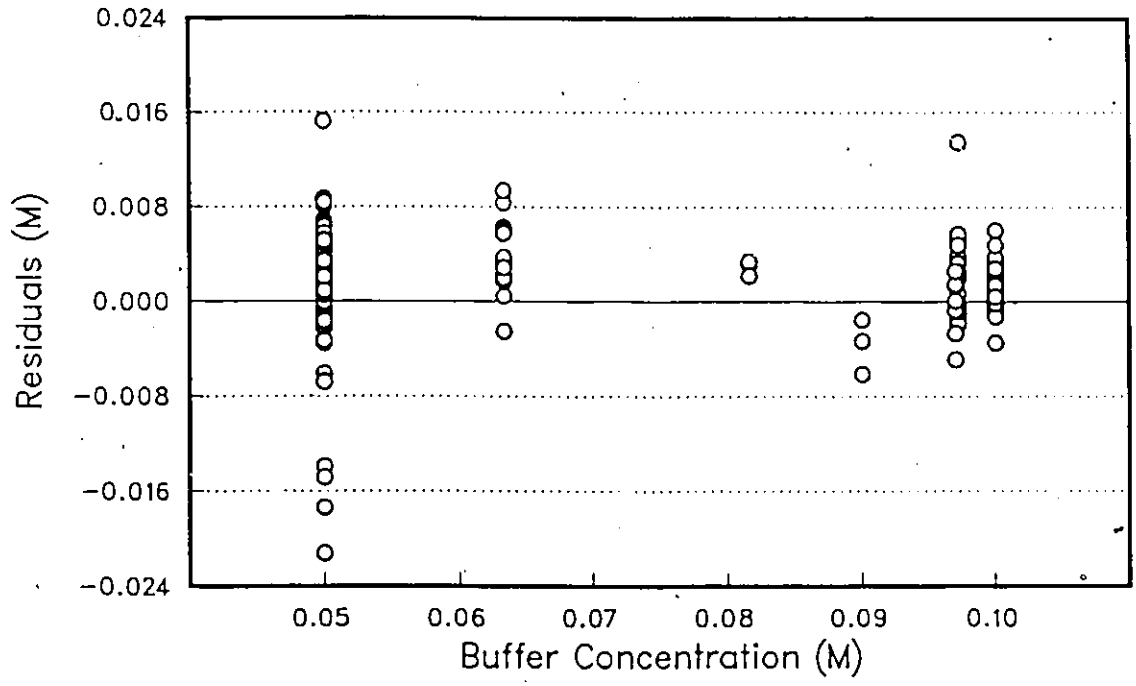


(k) Residuals vs Run Date

Figure 4: Model 1 Residuals

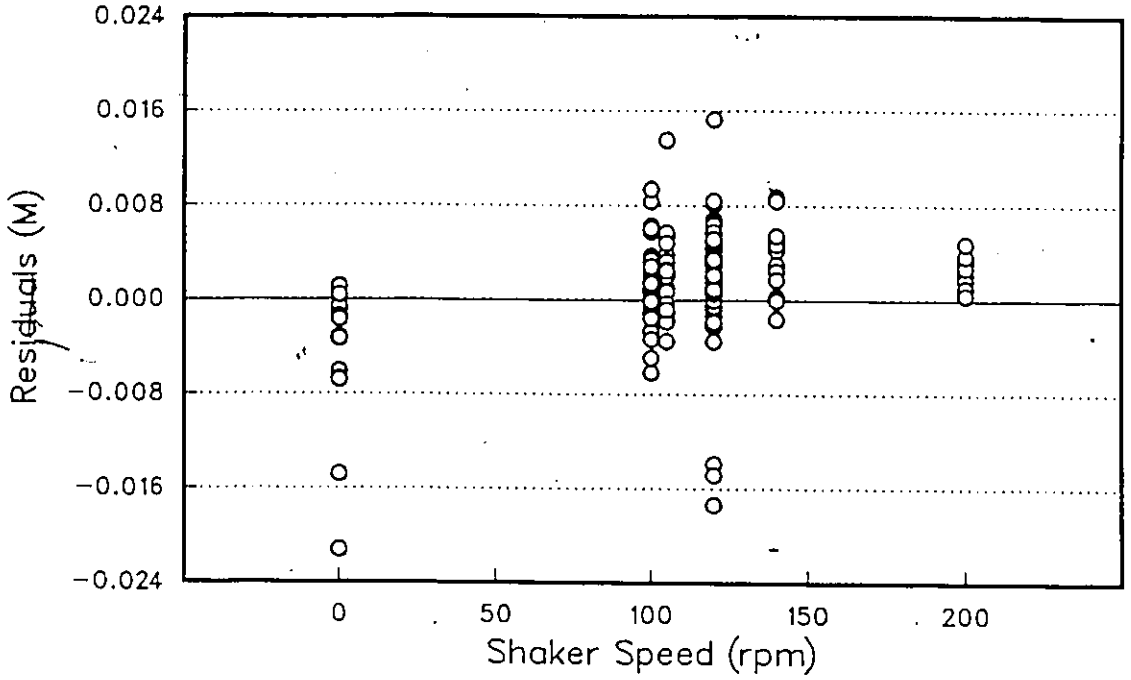


(l) Residuals vs Buffer Type

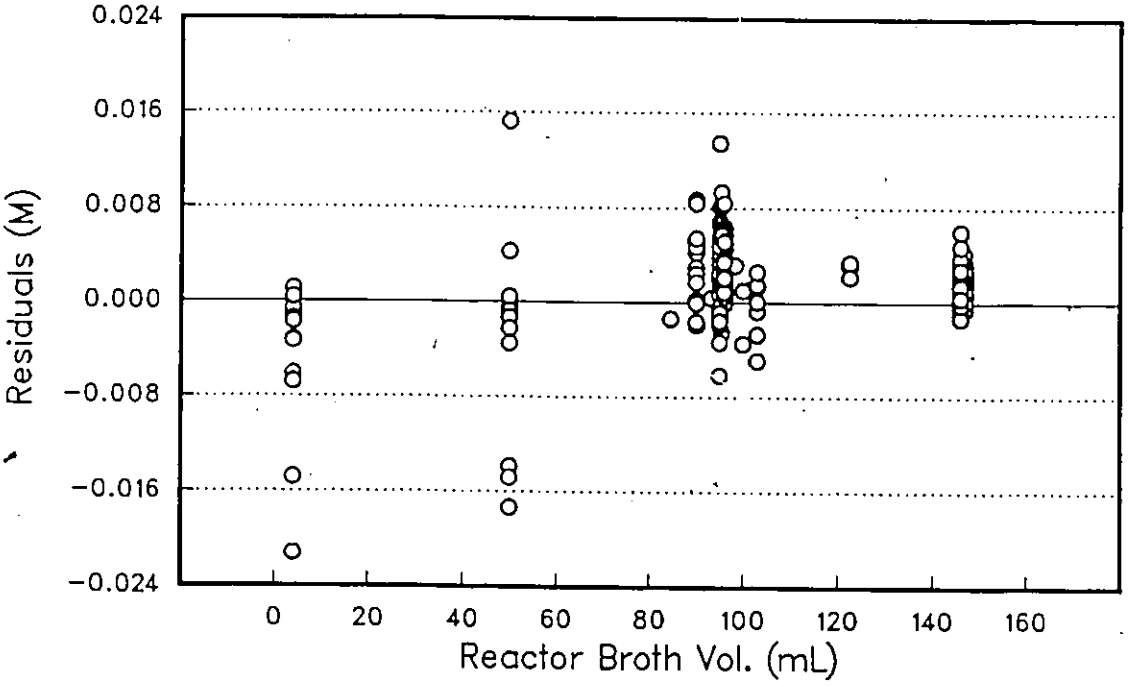


(m) Residuals vs Buffer Concentration

Figure 4: Model 1 Residuals

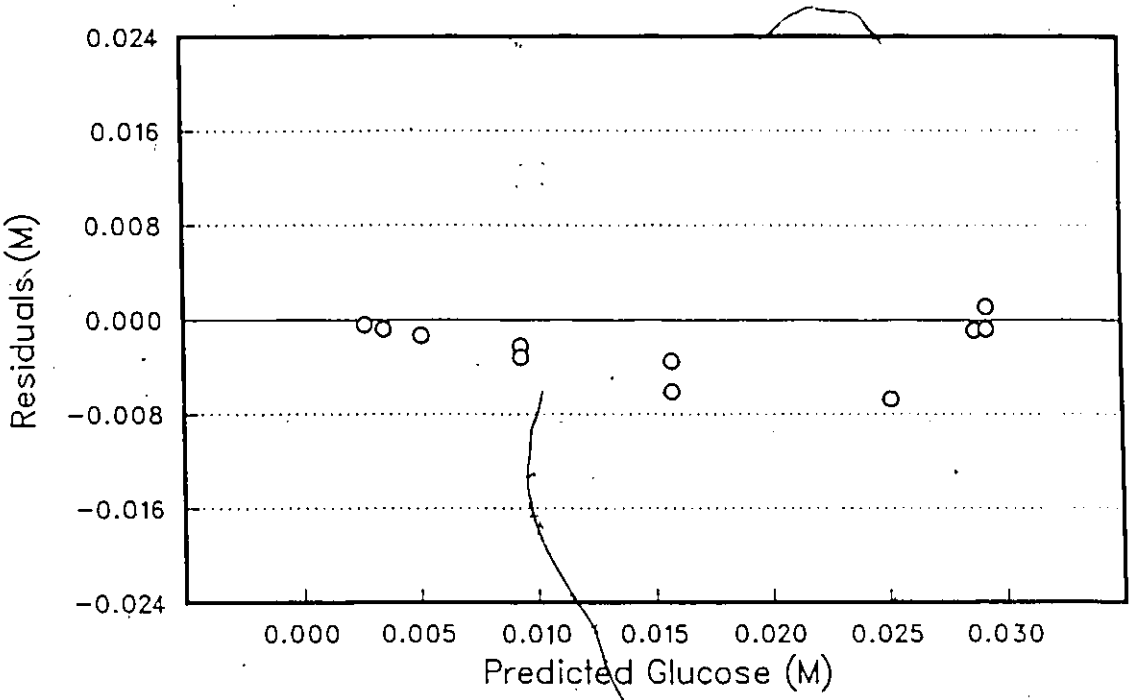


(n) Residuals vs Shaker Speed

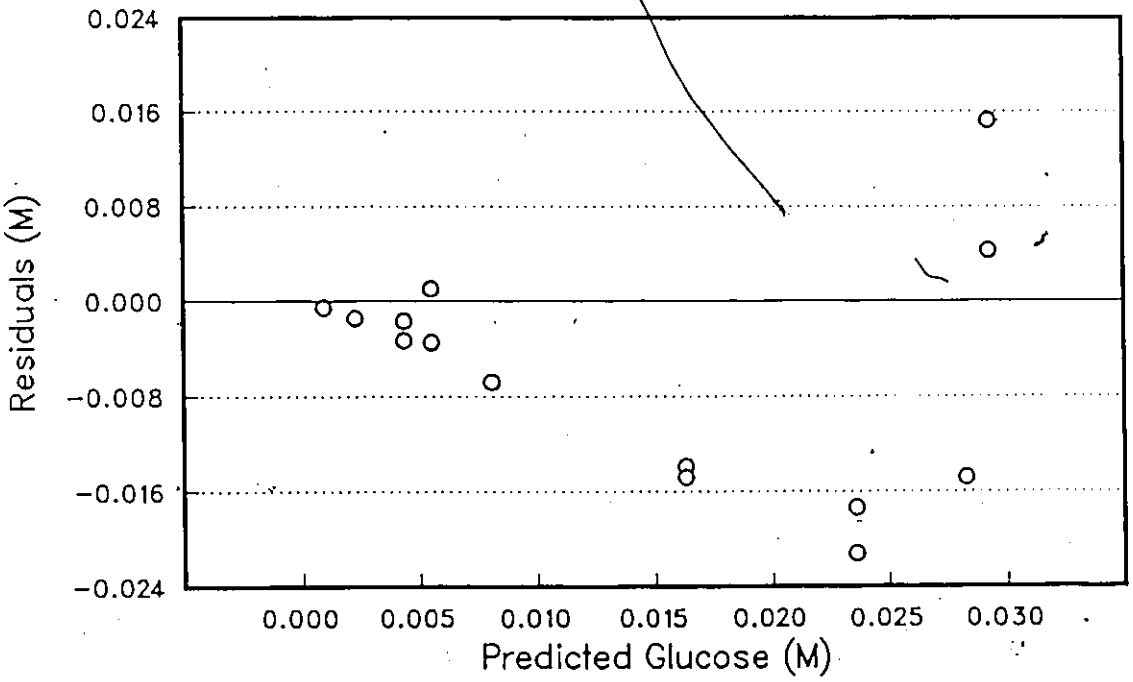


(o) Residuals vs Reactor Broth Vol.

Figure 4: Model 1 Residuals



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 5: Model 1 Residuals for Runs with no Initial CMC

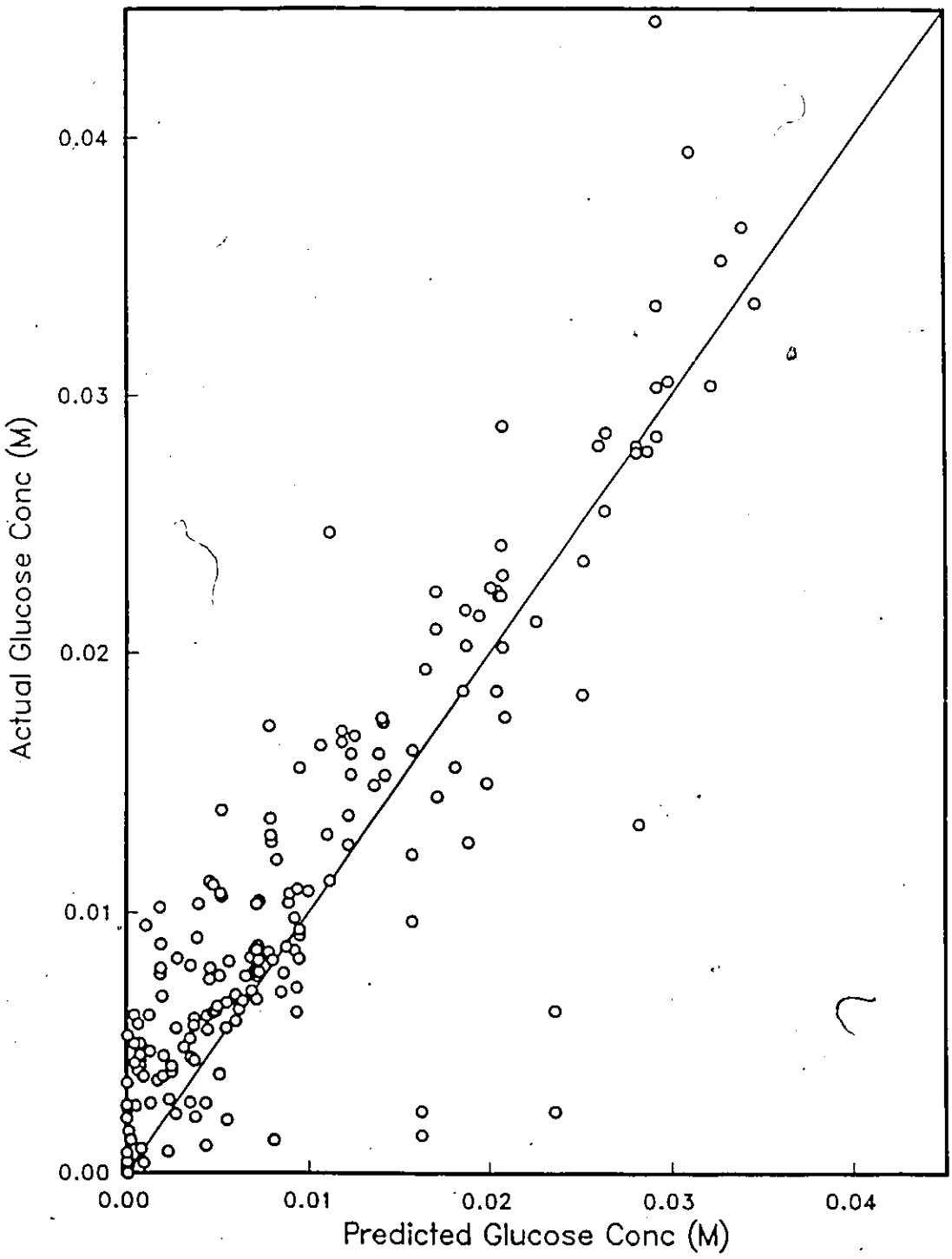


Figure 6: Actual vs Predicted Glucose for Model 1

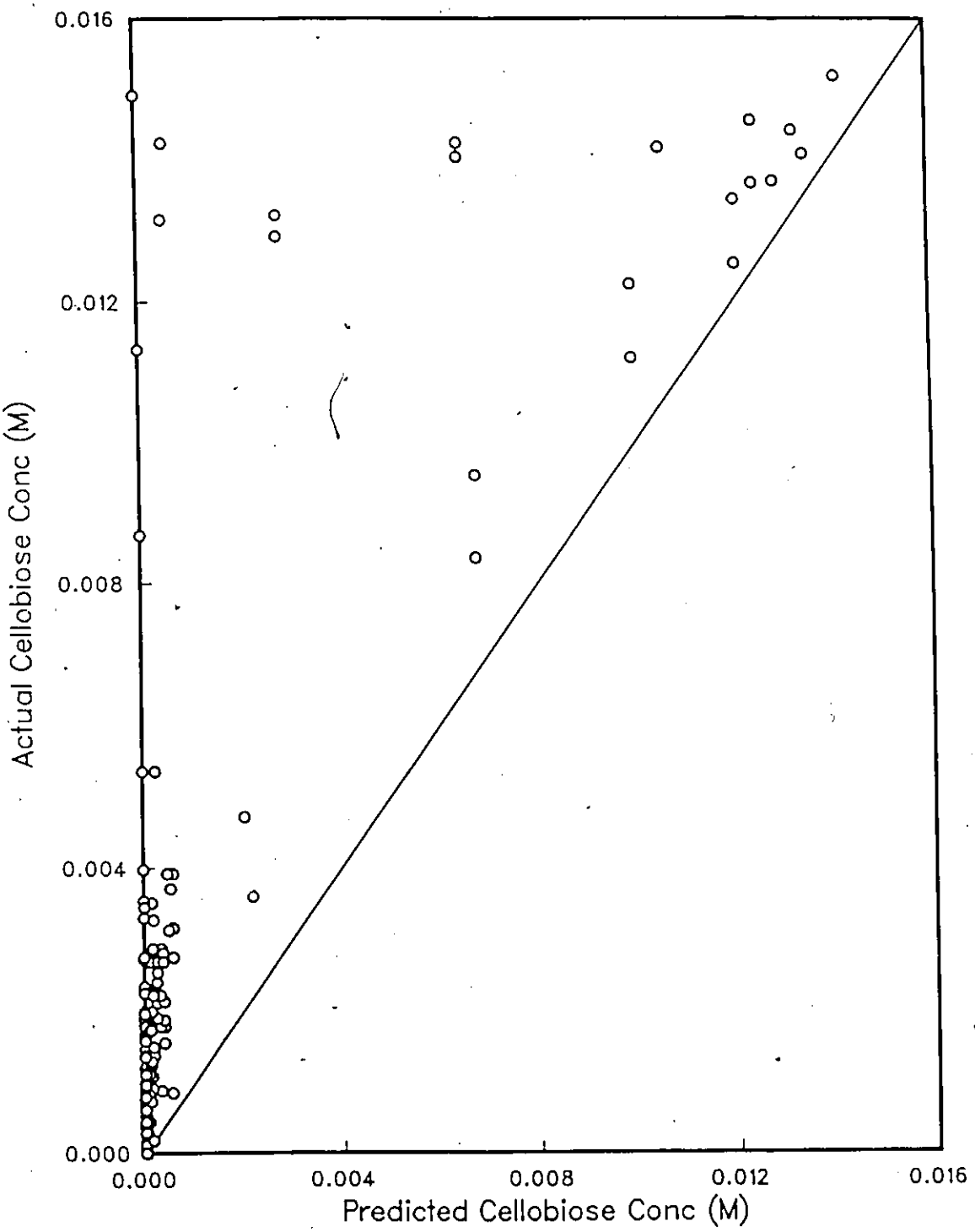


Figure 7: Actual vs Predicted Cellobiose for Model 1

5.6.3 Model 2 Fitting

There is only one difference between this model 2 and the basic model; E_x is assumed to be incapable of cleaving G_2. This is still a four-parameter model. A typical run is presented below:

MODEL 2

Non-linear routine specifications: $\Delta t = 200\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_x}	$0.5 \cdot 10^{-7}$	$0.3716 \cdot 10^{-7}$	$\pm 0.2140 \cdot 10^{-8}$
K_{mE_x}	$0.5 \cdot 10^{-1}$	$0.2587 \cdot 10^{-4}$	$\pm 0.1221 \cdot 10^{-3}$
k_{1E_1}	$0.5 \cdot 10^{-7}$	$0.1134 \cdot 10^{-6}$	$\pm 0.4240 \cdot 10^{-6}$
K_{mE_1}	$0.5 \cdot 10^{+1}$	$0.6648 \cdot 10^{-1}$	$\pm 0.2879 \cdot 10^{+0}$
Sum of squared residuals $= 0.002616\,\mathrm{M}^2$			

The correlation matrix of the parameters $= \begin{pmatrix} 1.000 & & & \\ -0.008 & 1.000 & & \\ 0.031 & 0.034 & 1.000 & \\ 0.046 & 0.021 & 0.999 & 1.000 \end{pmatrix}$

The assumption that endoglucanase does not cleave cellobiase is retained. It was the one difference in the attempt to improve on the base model. There was indeed an improvement because the SSR was lower for model 2 than for model 1. This was due to lowering the deviation from the predicted value for the data points which, as residuals of the basic fitted model, were outliers. Consequently that assumption is left unchanged.

What appears to be evident is that K_{mE_x} might be zero. Thus model 3 is obtained by dropping K_{mE_x} from model 2. The modelling thus far has shown that with our data, we cannot estimate K_{mE_x} to be significantly different than zero. Recall that K_{mE_x} is part of a sum in the denominator of the differential equations that make up the model: $K_{mE_x} + \beta_u + \beta_{rc} + \beta_{nrc}$. The variance of the glucose data may be so high and K_{mE_x} may be so small when compared to the sum of its substrate in the reactor that we cannot evaluate K_{mE_x} with any precision. Assuming that

K_{mE_x} equals zero is equivalent to saying that the endoglucanase enzyme cleaves bonds with reaction kinetics of order zero with respect to substrate concentration.

5.6.4 Model 3 Fitting

Here are typical results of fitting model 3 to the data:

MODEL 3

Non-linear routine specifications: $\Delta t = 70\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_x}	$0.36 \cdot 10^{-7}$	$0.3724 \cdot 10^{-7}$	$\pm 0.3246 \cdot 10^{-8}$
k_{1E_1}	$0.65 \cdot 10^{-7}$	$0.5558 \cdot 10^{-6}$	$\pm 0.9722 \cdot 10^{-5}$
K_{mE_1}	$0.42 \cdot 10^{-1}$	$0.3533 \cdot 10^{+0}$	$\pm 0.6363 \cdot 10^{+1}$
Sum of squared residuals $= 0.002632\,\mathrm{M}^2$			

$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1.000 & & \\ -0.008 & 1.000 & \\ -0.006 & 1.000 & 1.000 \end{pmatrix}$$

The SSR being identical in dropping K_{mE_x}, this change, making the model parsimonious, was kept.

However, results of fitting again reveal over parameterization. Closer examination reveals perfect correlation between k_{1E_1} and and K_{mE_1} and their ratio is virtually constant for the various starting points of the non-linear routine applied to model 3. The LSE of K_{mE_1} is greater than the sum of its substrates hence the denominator $K_{mE_1} + G_2 + \beta_{nrc} + \beta_{nrcs}$ in model 3 reduces to a constant in the model.

5.6.5 Model 4 Fitting

Model 4 is a simple two-parameter model. The model can now be fitted where the endoglucanase reaction is of order zero and the cellobiase reaction follows first-order kinetics. In this case, no convergence problems were encountered and the results are summarized below.

MODEL 4

Non-linear routine specifications: $\Delta t = 100\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_x}	$0.36805 \cdot 10^{-7}$	$0.3704 \cdot 10^{-7}$	$\pm 0.2105 \cdot 10^{-8}$
k_{1E_1}	$0.15000 \cdot 10^{-5}$	$0.1507 \cdot 10^{-5}$	$\pm 0.2473 \cdot 10^{-6}$
Sum of squared residuals $= 0.002615\,\mathrm{M}^2$			

The covariance matrix of the parameters:

$$\begin{pmatrix} +0.11415 \cdot 10^{-17} & \\ -0.18600 \cdot 10^{-16} & +0.15759 \cdot 10^{-13} \end{pmatrix}$$

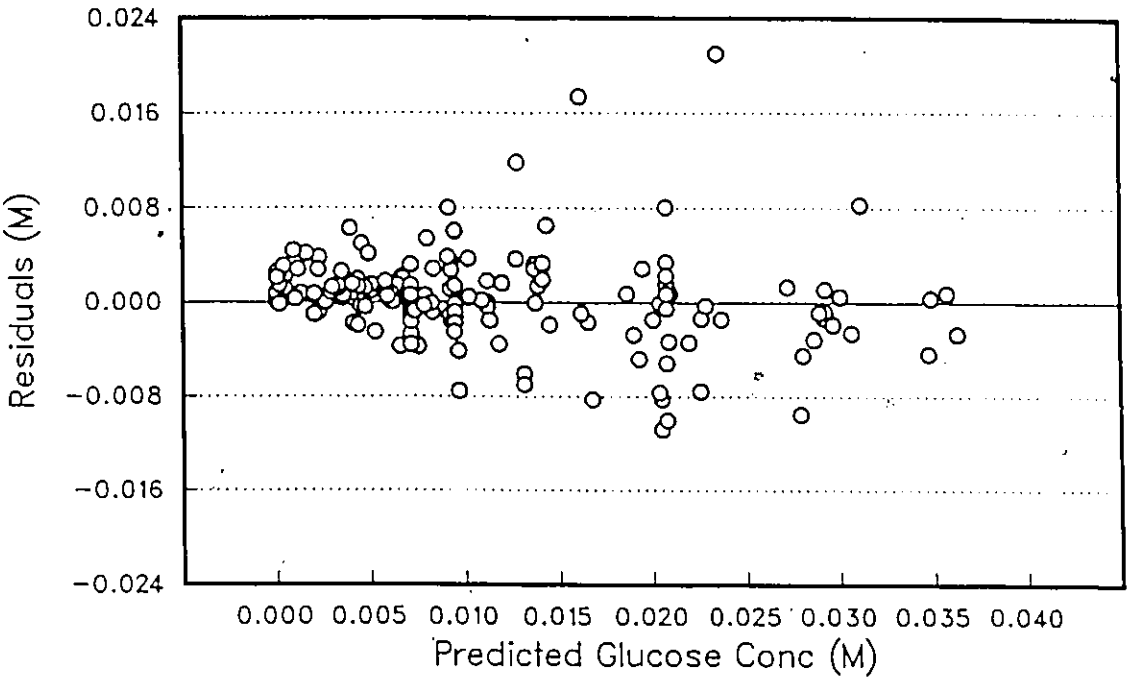
$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1.000 & \\ -0.139 & 1.000 \end{pmatrix}$$

The SSR remained the same in combining K_{1E_1} and K_{mE_1} from model 3. Thus subsequent models keep this combined form of the E_1 parameters.

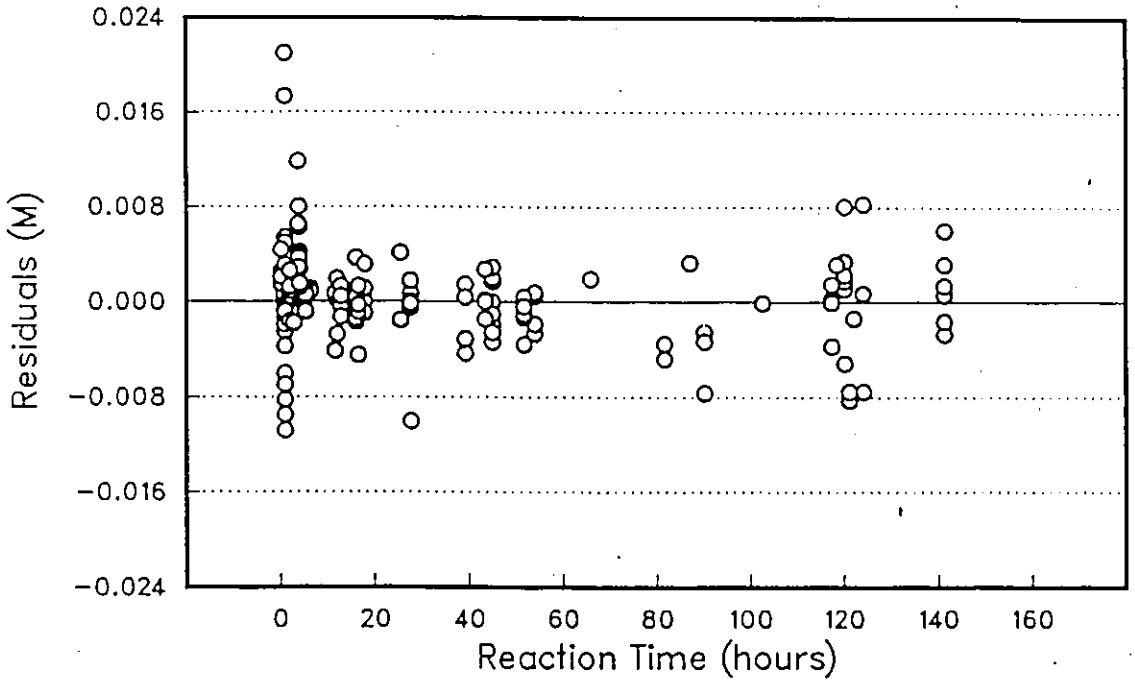
Four different residual plots may be found in Figure 8. The residual plots for runs with no initial CMC can be found in Figure 9. Plots of actual vs predicted for glucose and cellobiose can be found in figures 10 and 11.

The residuals for runs consisting of cellobiose hydrolysis have, in all models presented thus far, including model 4, shown a pattern. (This is seen in figure 9). Possible reasons may be that inhibition of either E_1 or E_x is significant. This is examined in models 5 and 6.

Continued underprediction of the cellobiose concentration as seen in figure 11 will not be explained by inhibition but will eventually be dealt with in later models.

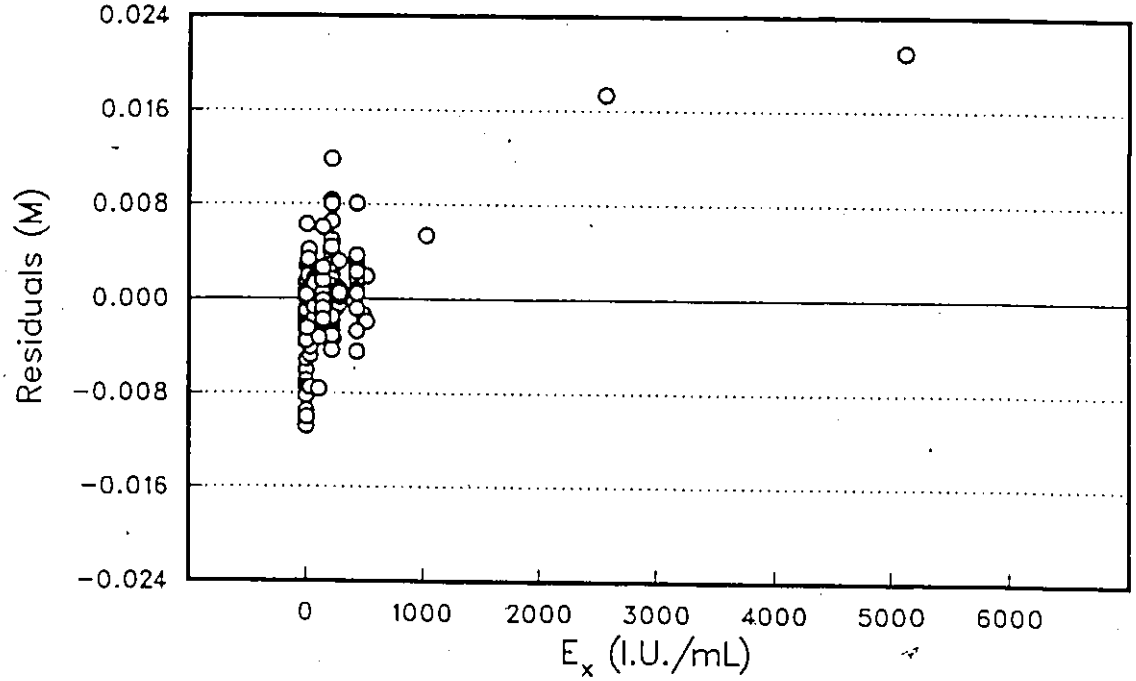


(a) Residuals vs Predicted Glucose

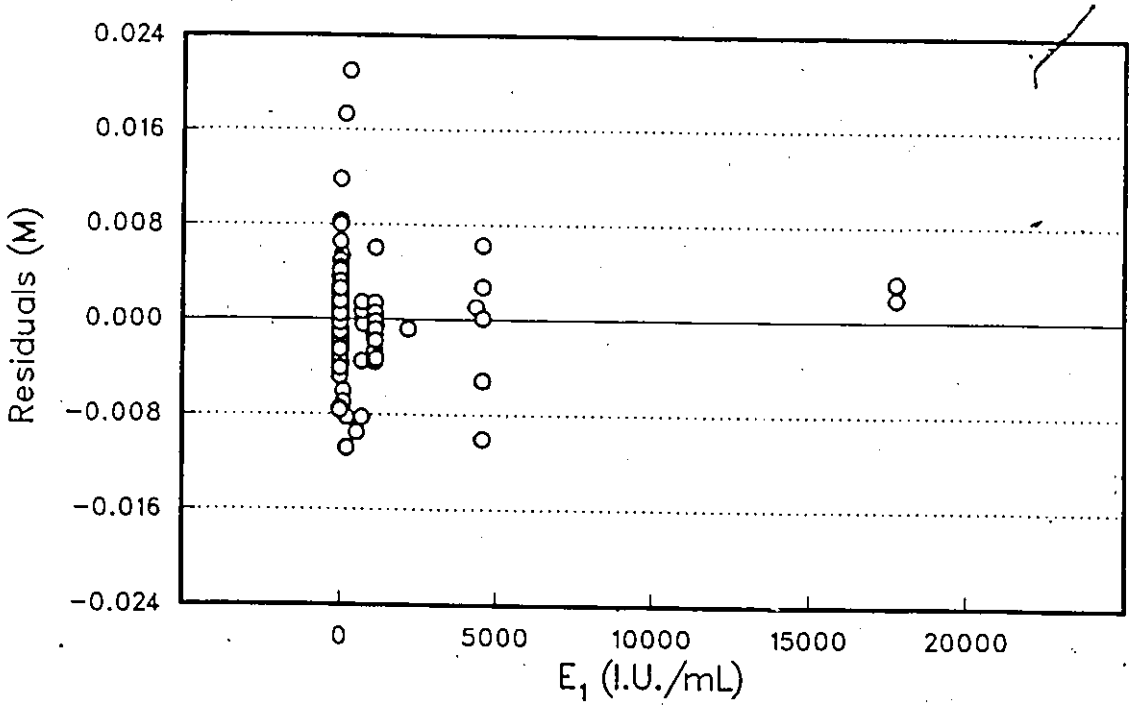


(b) Residuals vs Reaction Time

Figure 8: Model 4 Residuals

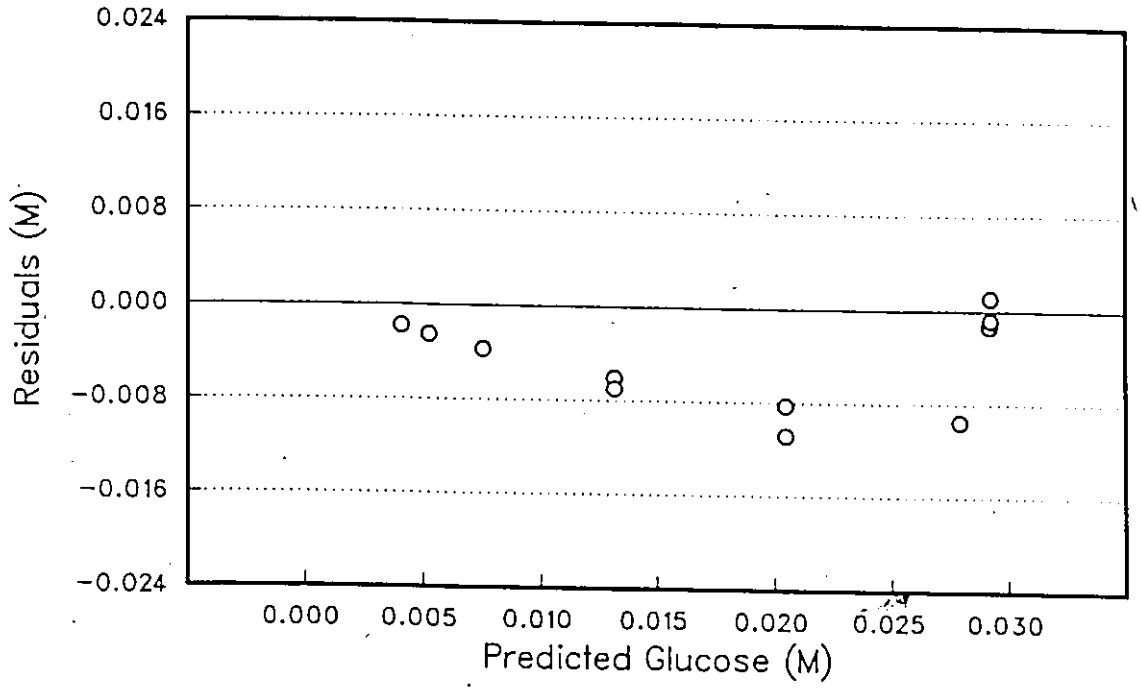


(c) Residuals vs E_x

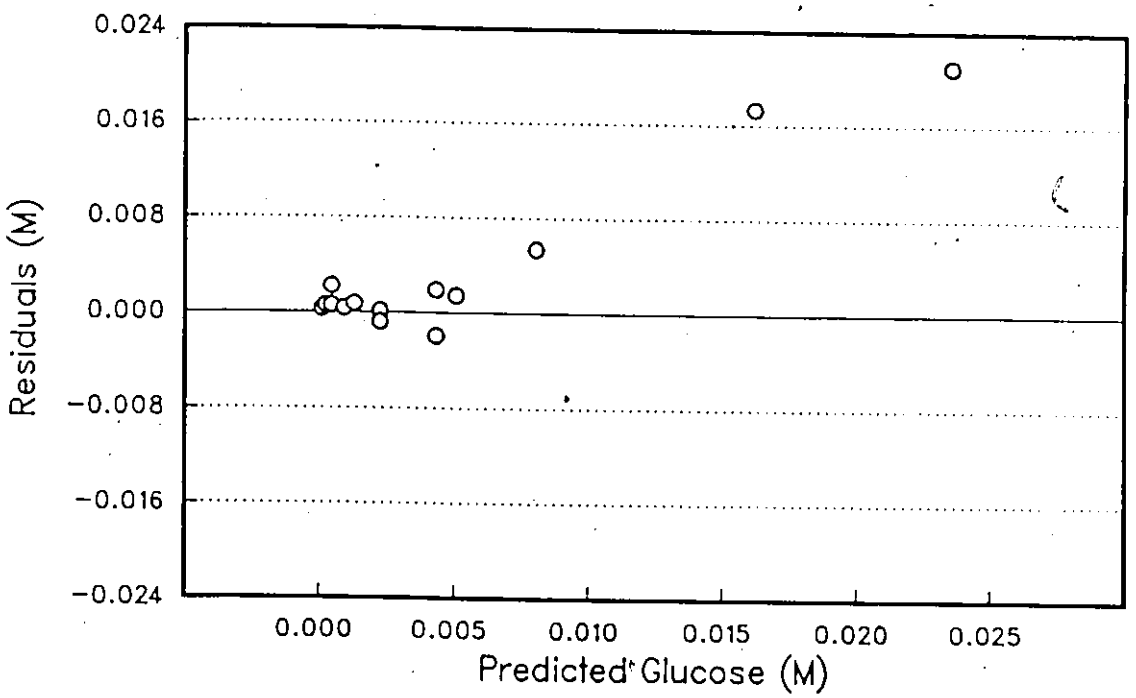


(d) Residuals vs E_1

Figure 8: Model 4 Residuals



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 9: Model 4 Residuals for Runs with no Initial CMC

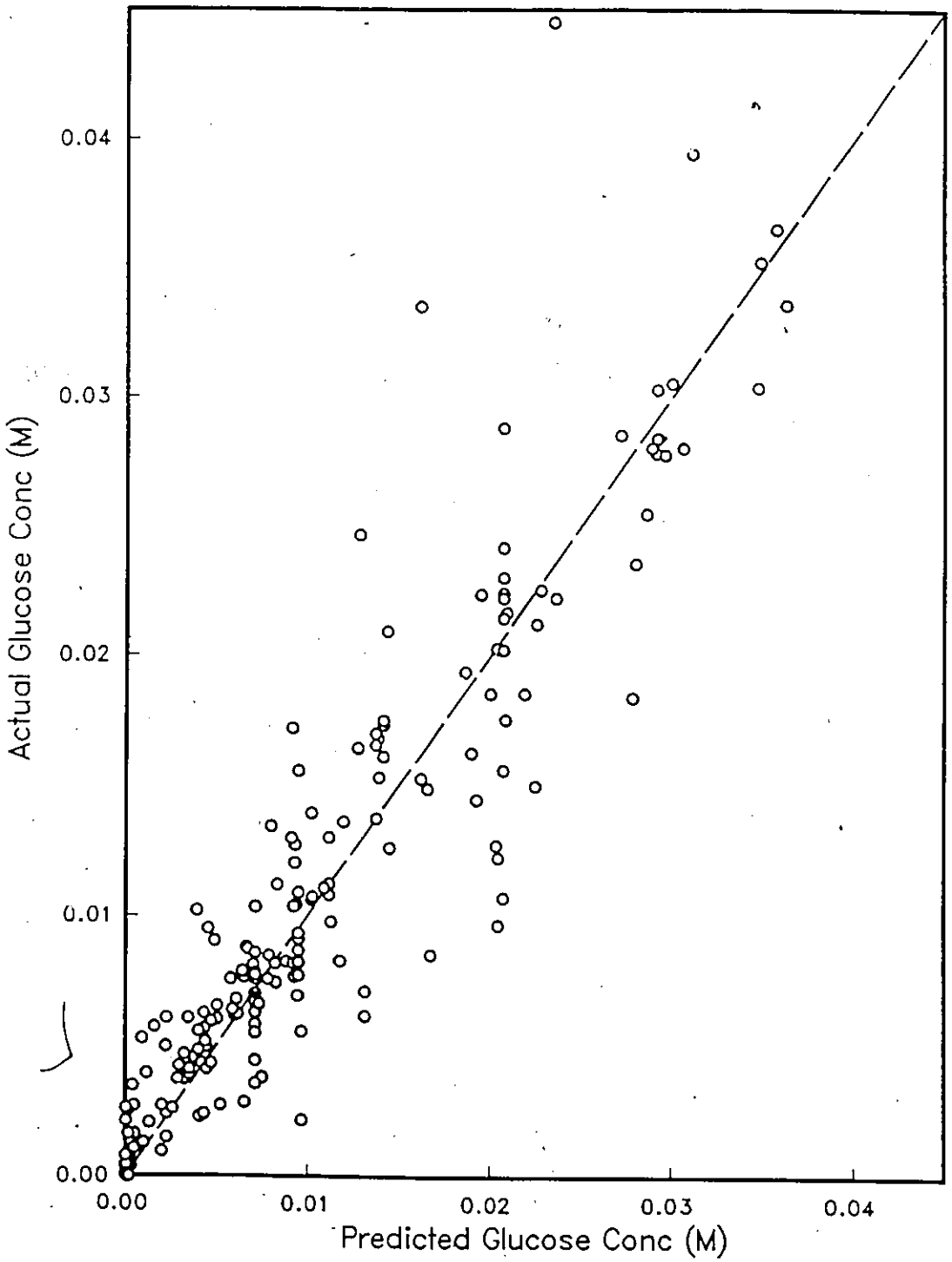


Figure 10: Actual vs Predicted Glucose for Model 4

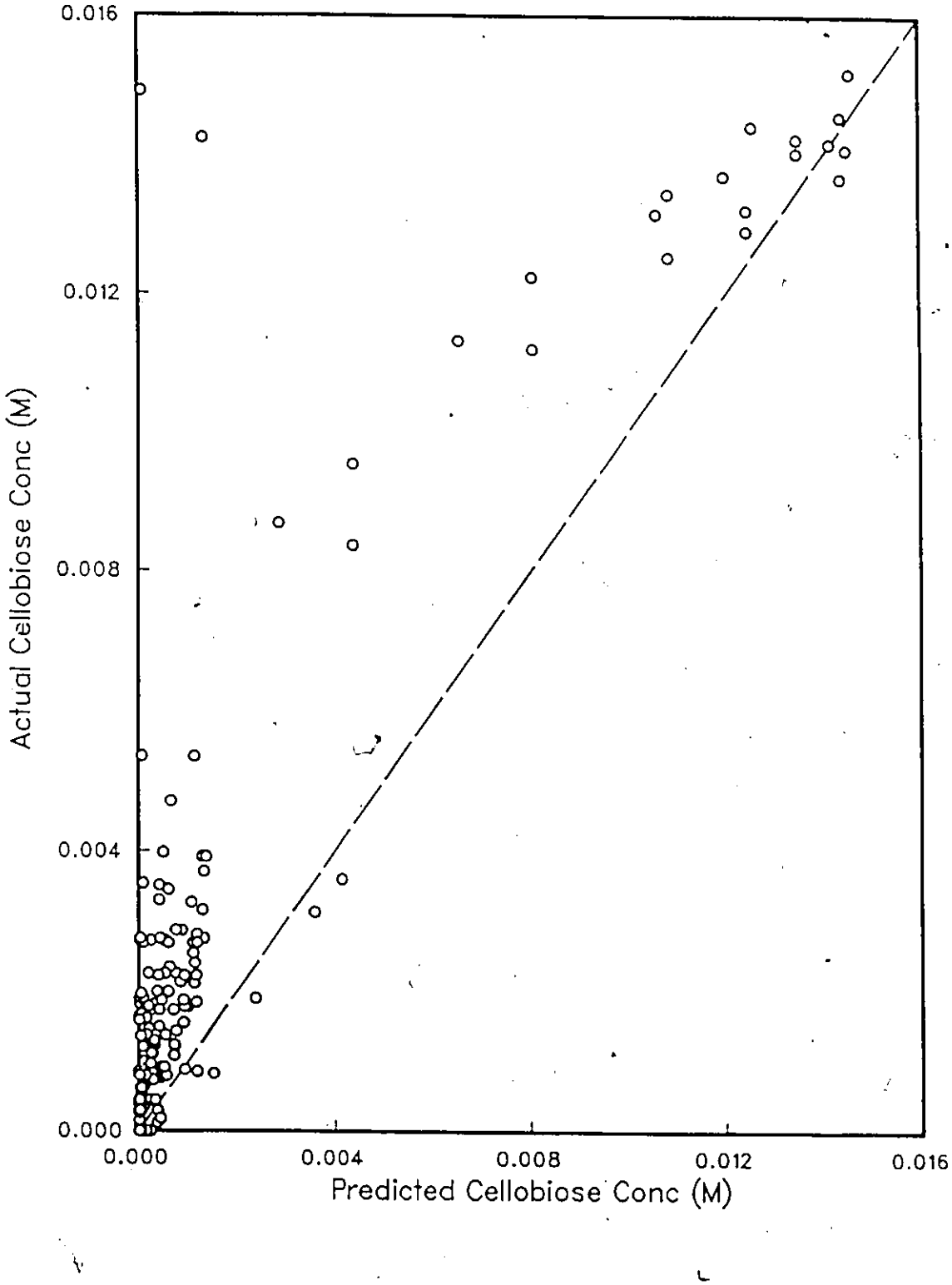


Figure 11: Actual vs Predicted Cellobiose for Model 4

5.6.6 Model 5 Fitting

In model 5, the previous model has been modified to account for inhibition of E_1 by glucose. Because of the model form, we cannot distinguish between competitive and non-competitive inhibition. The results, summarized below, show that K_{iG_1}, the term accounting for inhibition, is plausibly zero. Thus the data do not support this hypothesis. However, this may be the result of insufficient reproducibility and location of the experiments rather than an improper hypothesis.

MODEL 5

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, $\mathrm{TOL} = 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_2}	$0.37046 \cdot 10^{-7}$	$0.2844 \cdot 10^{-7}$	$\pm 0.4583 \cdot 10^{-8}$
k_{1E_1}	$0.35066 \cdot 10^{-5}$	$0.2274 \cdot 10^{-5}$	$\pm 0.8523 \cdot 10^{-6}$
K_{iG_1}	$0.50000 \cdot 10^{-2}$	$0.2491 \cdot 10^{-1}$	$\pm 0.2911 \cdot 10^{-1}$
Sum of squared residuals $= 0.002564\,\mathrm{M}^2$			

The covariance matrix of the parameters:

$$\begin{pmatrix} +0.54122 \cdot 10^{-17} & & \\ -0.25175 \cdot 10^{-15} & +0.18716 \cdot 10^{-12} & \\ +0.89297 \cdot 10^{-11} & -0.56837 \cdot 10^{-08} & +0.21835 \cdot 10^{-1} \end{pmatrix}$$

The correlation matrix of the parameters $= \begin{pmatrix} 1.000 & & \\ -0.250 & 1.000 & \\ 0.260 & -0.889 & 1.000 \end{pmatrix}$

5.6.7 Model 6 Fitting

Model 4 was modified in another way. The one new assumption that is made in model 6 is that endoglucanase forms a complex with all the β-1,4 glucosidic bonds in the reaction mixture, i.e., the β-1,4 bonds that endoglucanase is not capable of cleaving are considered inhibitors of E_2. Furthermore this inhibition is considered reversible. The enzyme's active site is assumed to bind with the uncleavable bond.

And furthermore the dissociation of this uncleavable bond–E_x complex is assumed to take place at the same rate as the other (cleavable) bond–E_x complexes. This is purely an intuitive feeling but it means that the model still has only 2 parameters.

MODEL 6

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_x}	$0.4 \cdot 10^{-6}$	$0.4002 \cdot 10^{-6}$	$\pm 0.1278 \cdot 10^{-6}$
k_{1E_1}	$0.2 \cdot 10^{-5}$	$0.1610 \cdot 10^{-5}$	$\pm 0.2678 \cdot 10^{-6}$
Sum of squared residuals $= 0.002609\,\mathrm{M}^2$			

The covariance matrix of the parameters:

$$\begin{pmatrix} +0.42087 \cdot 10^{-14} & \\ -0.11208 \cdot 10^{-14} & +0.18473 \cdot 10^{-13} \end{pmatrix}$$

$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1.000 & \\ -0.127 & 1.000 \end{pmatrix}$$

There is virtually no difference in the diagnostics for models 4 and 6. Both have the same number of parameters and the same SSR. However, since the literature reports inhibition of E_x by highly substituted methyl cellulose, model 6 is kept as our preferred form over model 4. There are still some concerns however.

Fifteen different residual plots may be found in Figure 12. The residual plots for runs with no initial CMC can be found in Figure 13. Plots of actual vs predicted for glucose and cellobiose can be found in figures 14 and 15.

The residual plots with respect to run date, buffer type, buffer concentration, shaker speed and reaction mixture volume indicate that none of these variables, in the ranges examined, influence the kinetics. Those plots will not be shown again until a definitive model is accepted. The other plots in Figure 12 also show no trends. However when the data are segregated and residuals for those runs where

only cellobiose is hydrolyzed are plotted, the same trends as on Figure 9 can be seen in Figure 13.

There are two inadequacies revealed by the residual plots thus far: The representation of the data by the model for runs where the only substrate is cellobiose is inadequate. There is a bias which can be attributed to the enzyme source. Runs from the Celluclast solution are, in general, overpredicted. Runs from the Novozym commercial solution are underpredicted for the glucose concentration. The second inadequacy concerns the model prediction of the cellobiose concentration. There is an obvious trend seen in Figure 15. We will address the former problem first, by modelling for only those runs with no initial CMC to find a fitted model describing cellobiose hydrolysis kinetics.

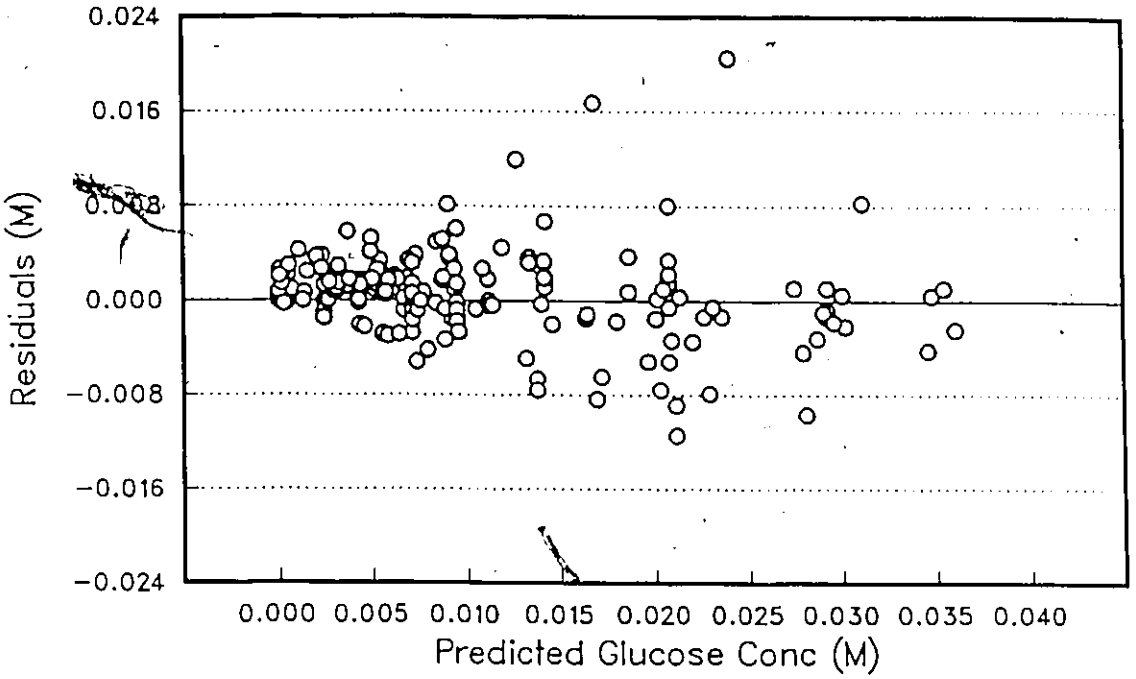
5.6.8 Model 7 Fitting

Model 7 is identical to model 6 but the parameter estimates are based on only those data points where the initial carboxymethyl cellulose concentration is nil. Since only E_1 is assumed to cleave cellobiose, this is equivalent to calculating the LSE of K_{1E_1} based on the runs with no initial CMC.

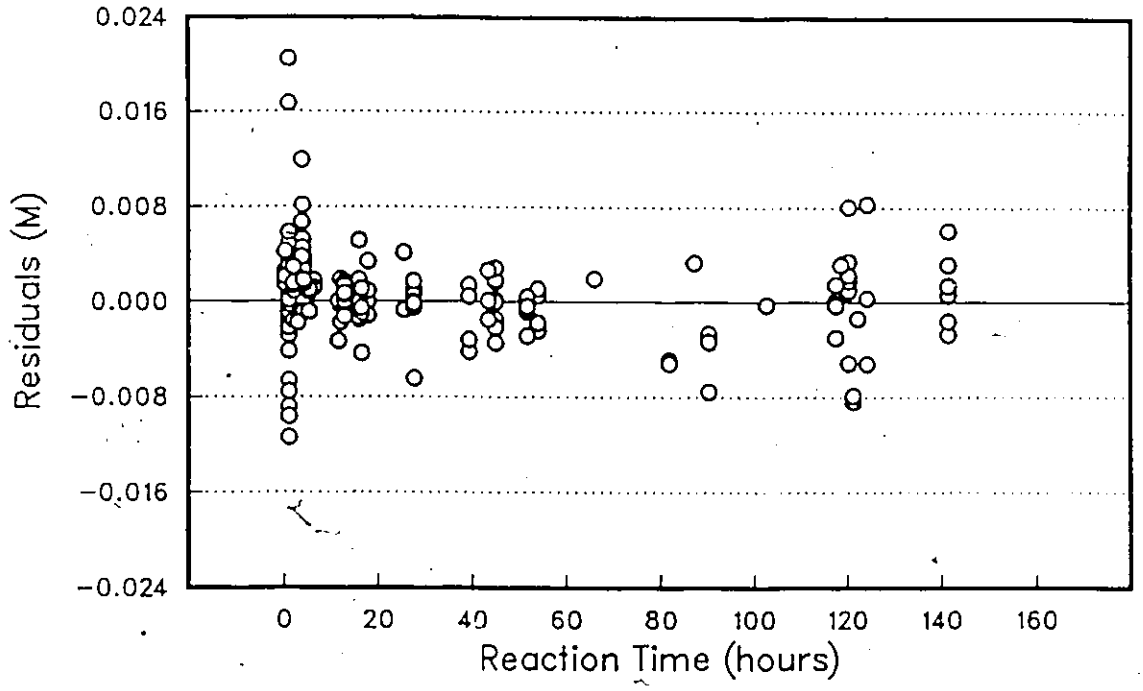
MODEL 7

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, $\mathrm{TOL} = 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_1}	$0.16102 \cdot 10^{-6}$	$0.1491 \cdot 10^{-5}$	$\pm 0.6813 \cdot 10^{-6}$
Sum of squared residuals $= 0.001189\,\mathrm{M}^2$			

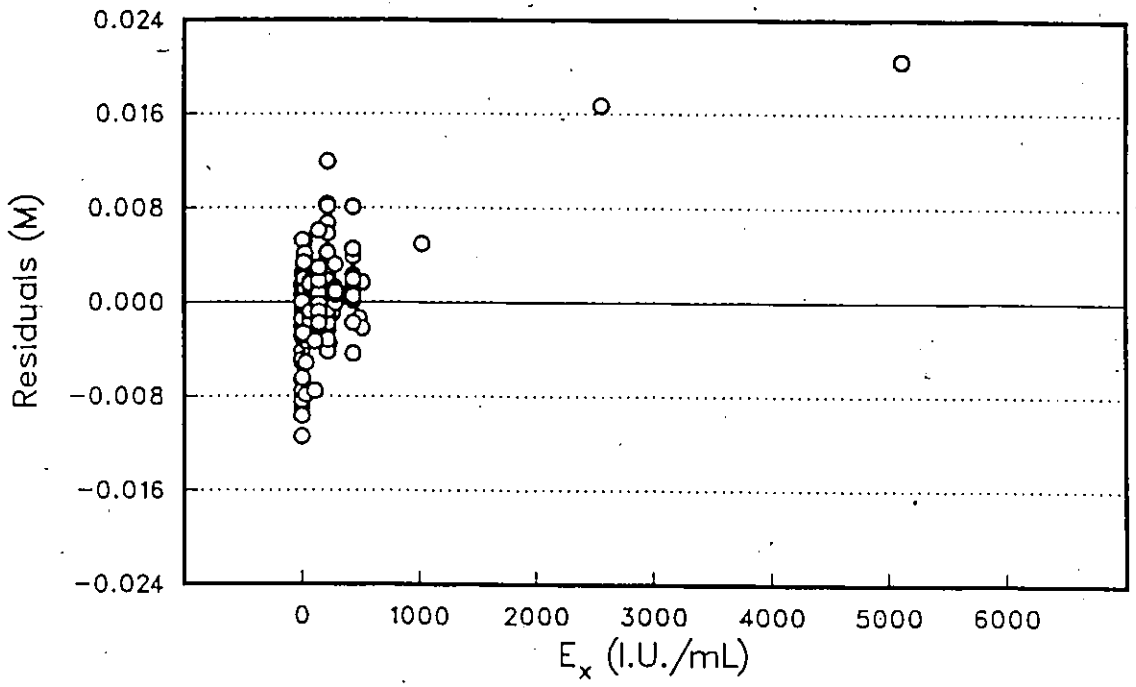


(a) Residuals vs Predicted Glucose

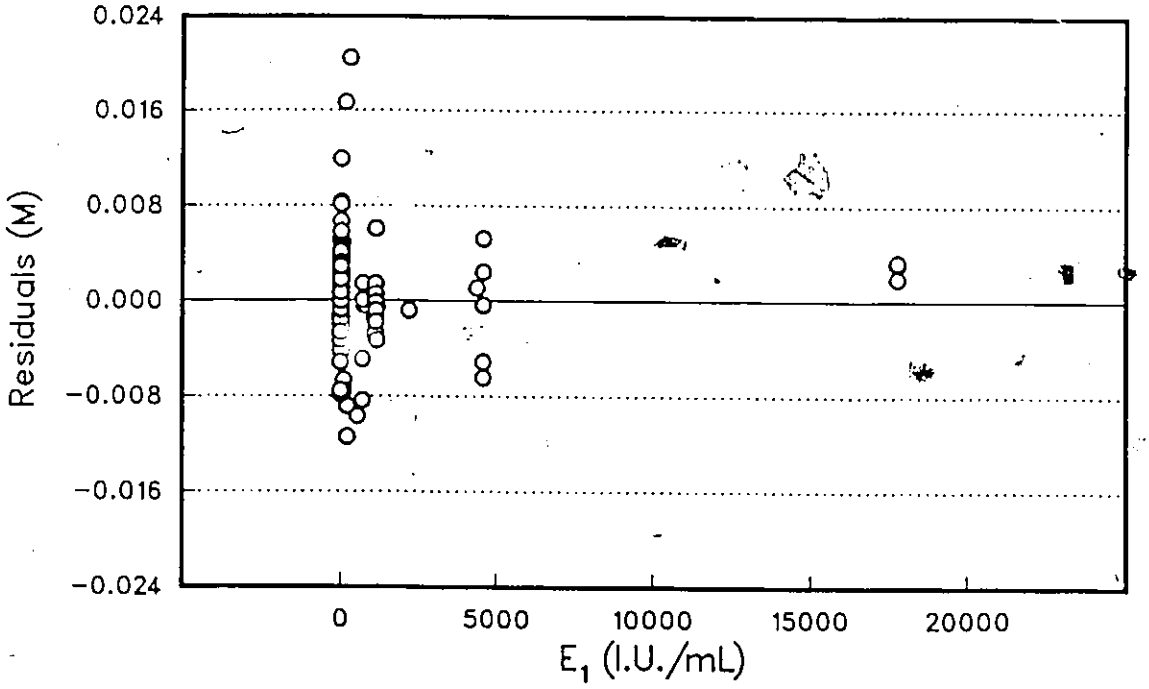


(b) Residuals vs Reaction Time

Figure 12: Model 6 Residuals

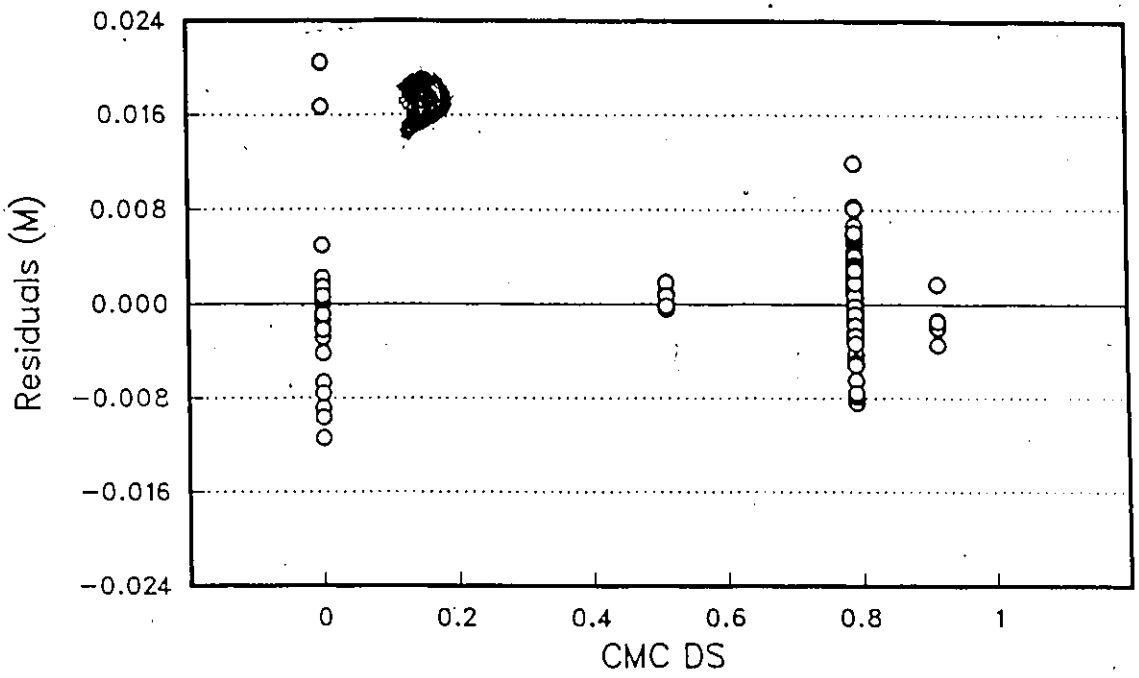


(c) Residuals vs E_x

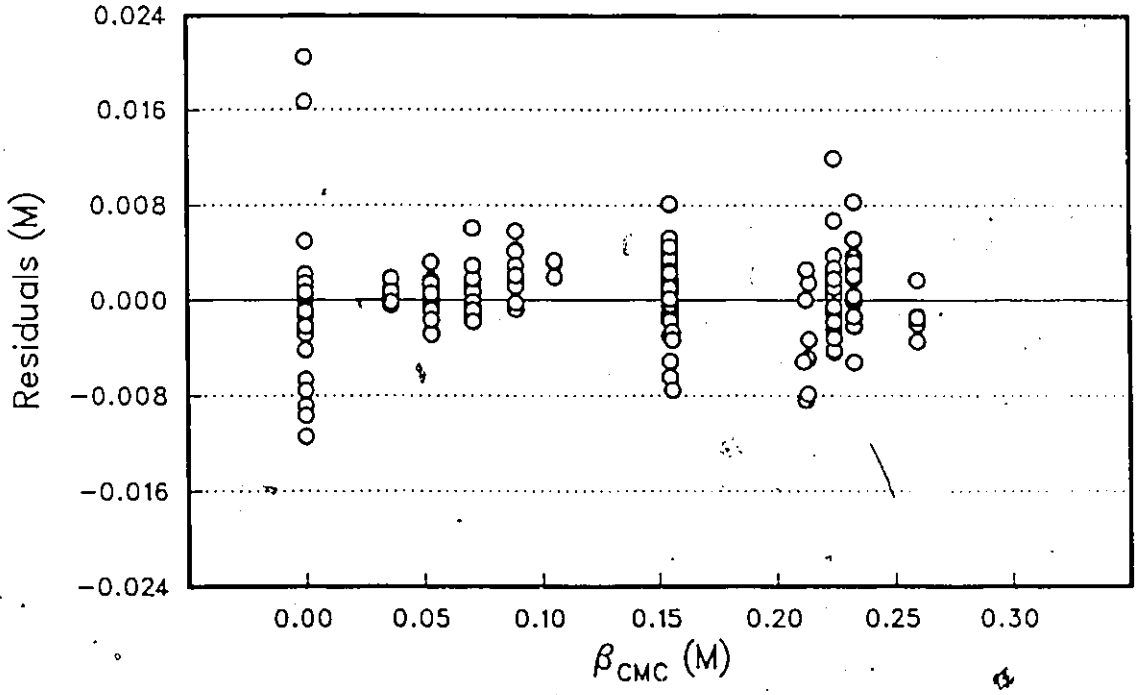


(d) Residuals vs E_1

Figure 12: Model 6 Residuals

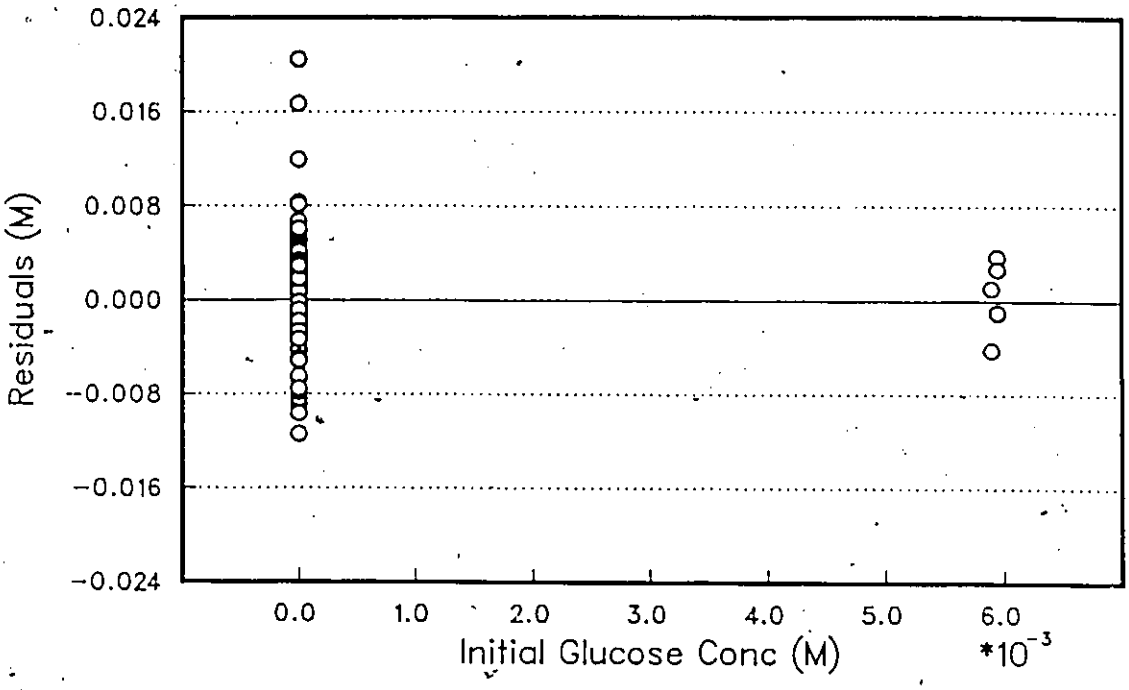


(e) Residuals vs CMC DS

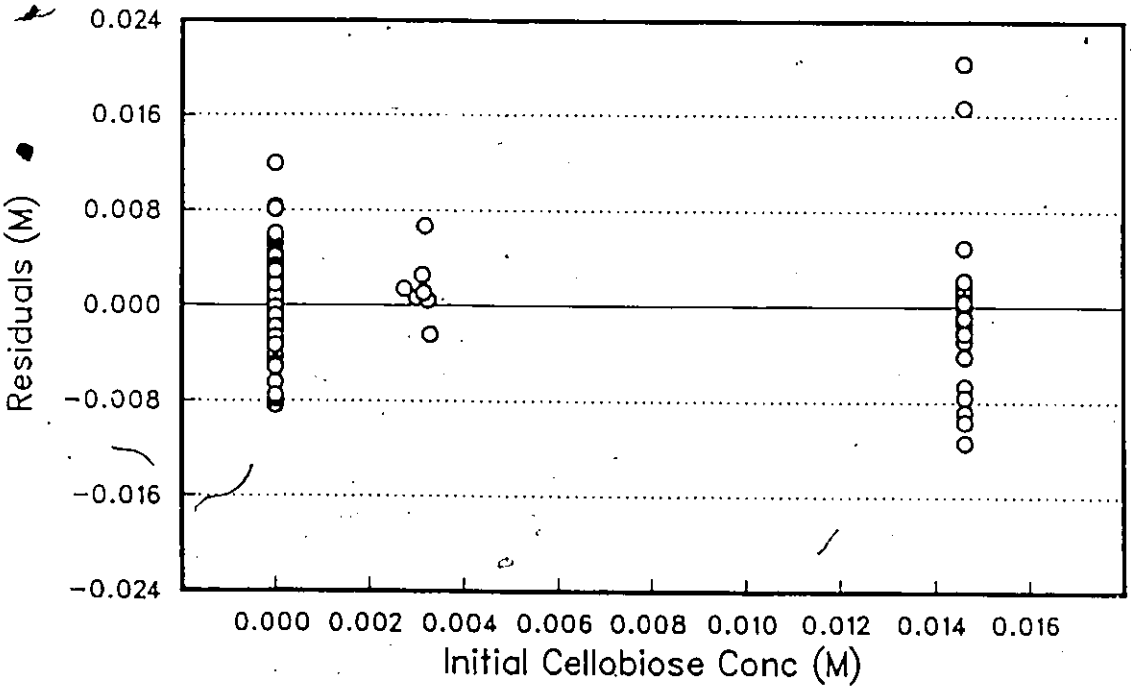


(f) Residuals vs β_{CMC}

Figure 12: Model 6 Residuals

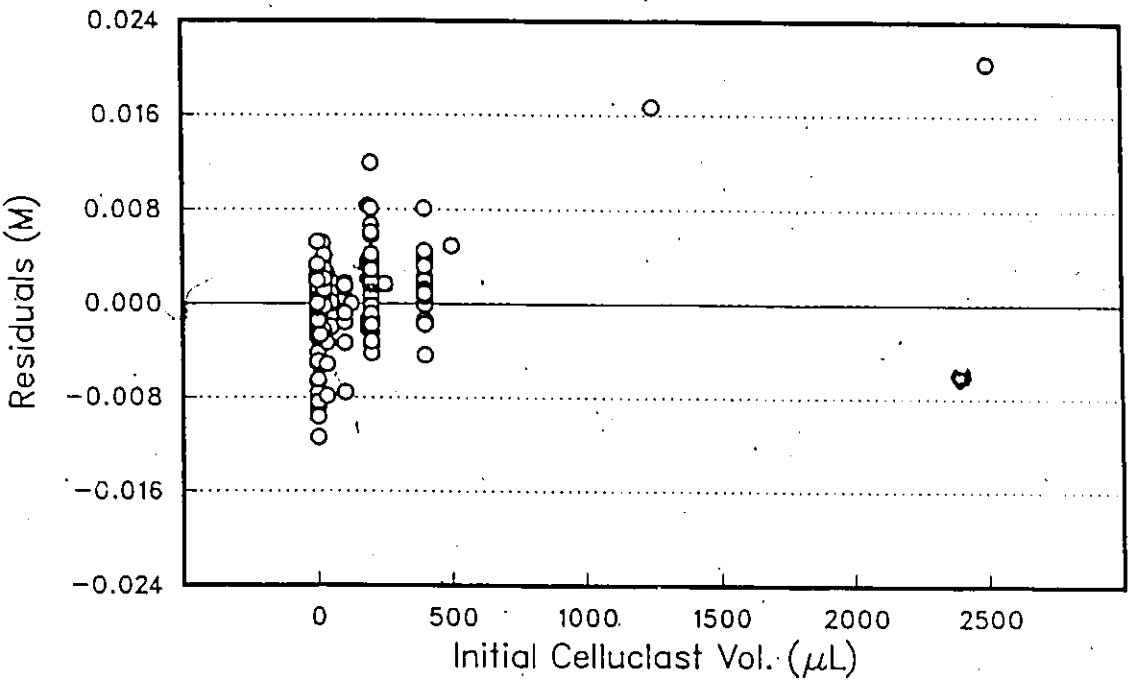


(g) Residuals vs Initial Glucose Concentration

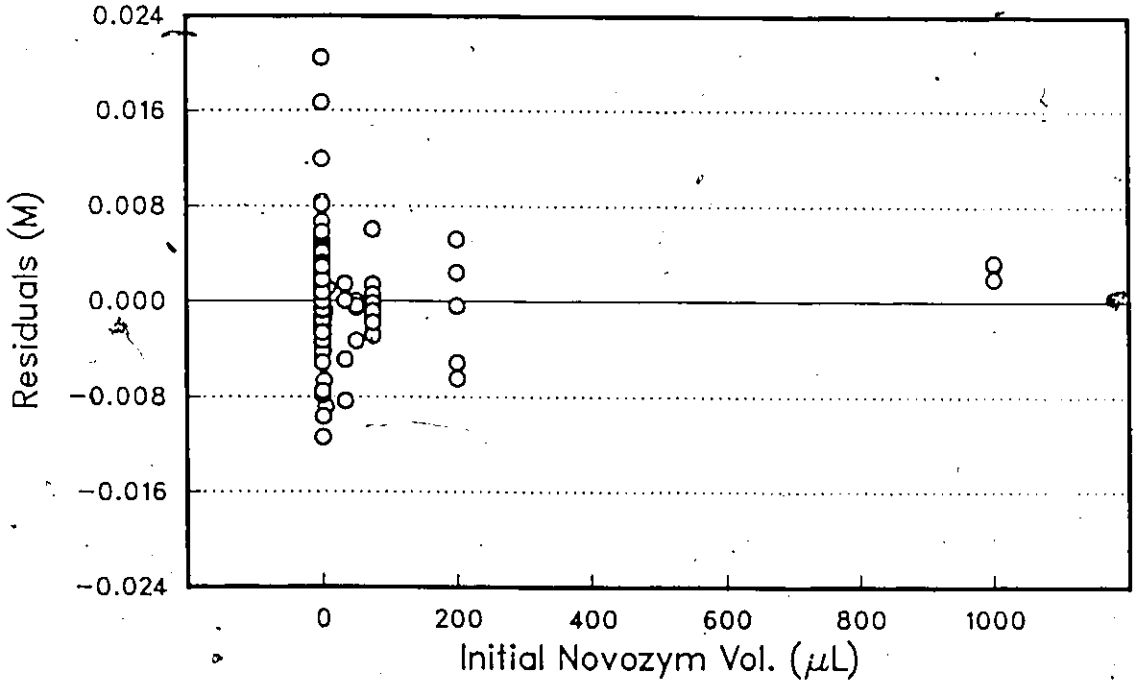


(h) Residuals vs Initial Cellobiose Conc.

Figure 12: Model 6 Residuals

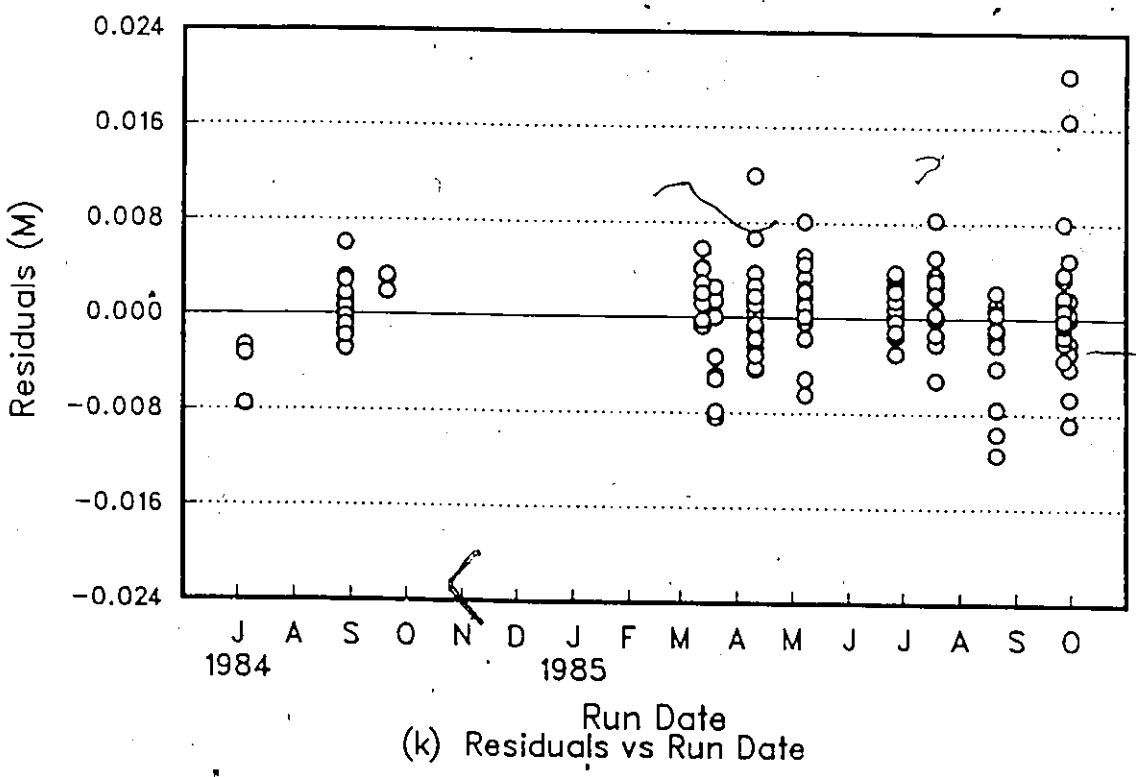


(i) Residuals vs Initial Celluclast Vol.



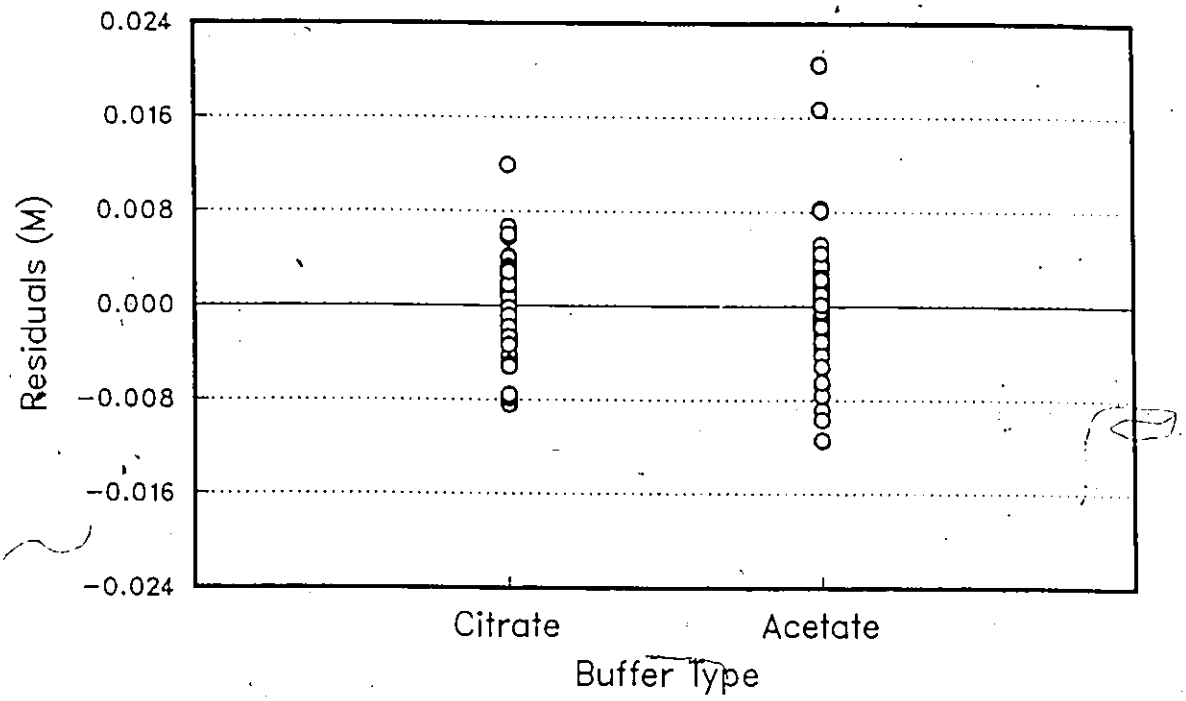
(j) Residuals vs Initial Novozym Vol.

Figure 12: Model 6 Residuals

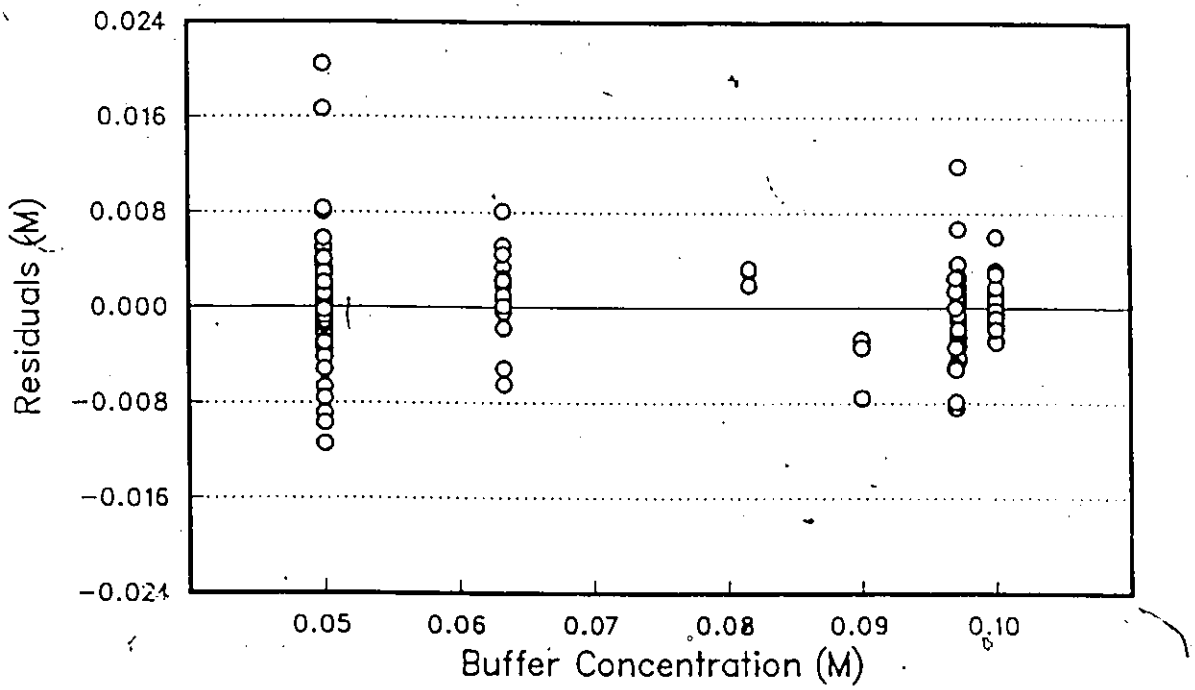


(k) Residuals vs Run Date

Figure 12: Model 6 Residuals

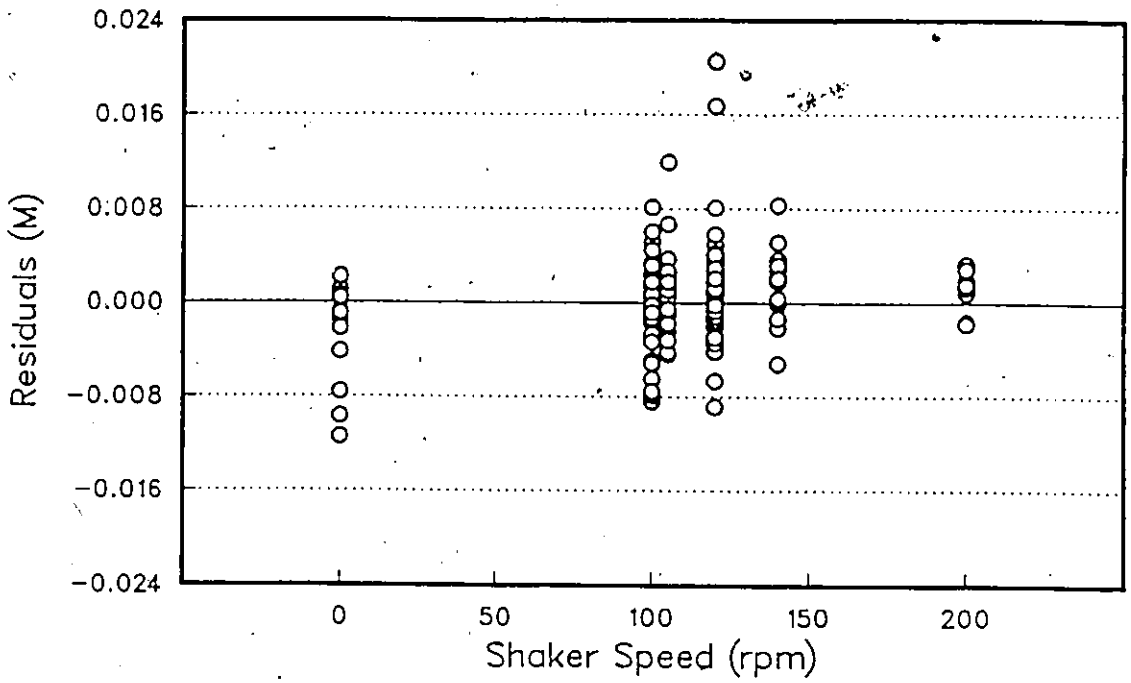


(l) Residuals vs Buffer Type

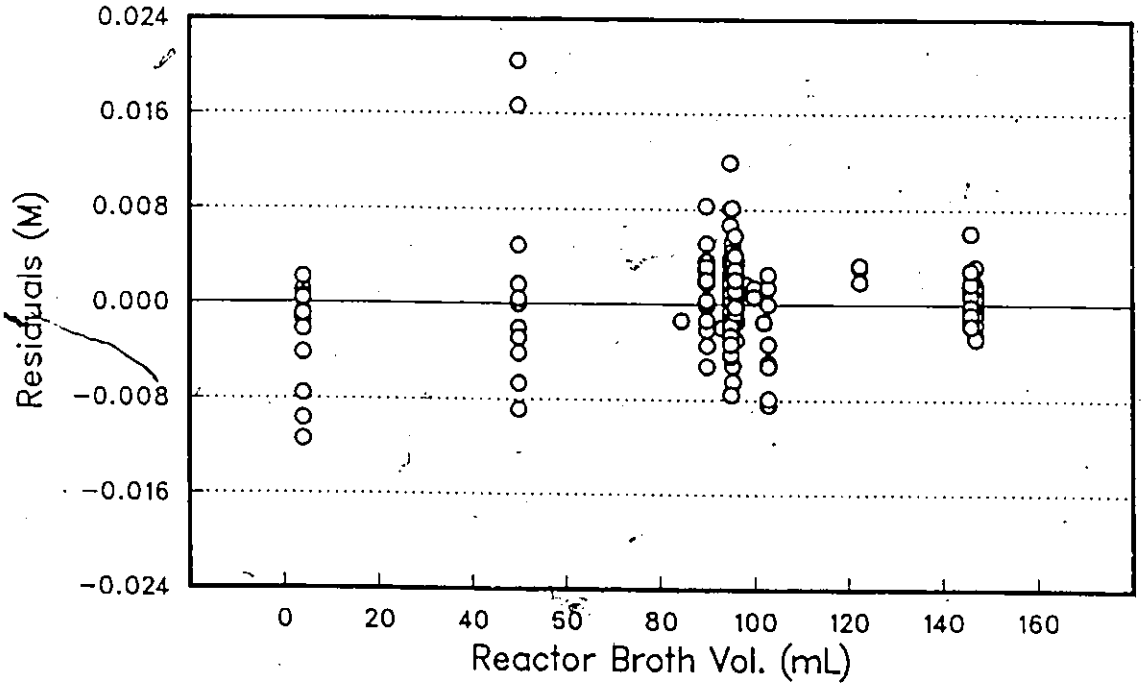


(m) Residuals vs Buffer Concentration

Figure 12: Model 6 Residuals

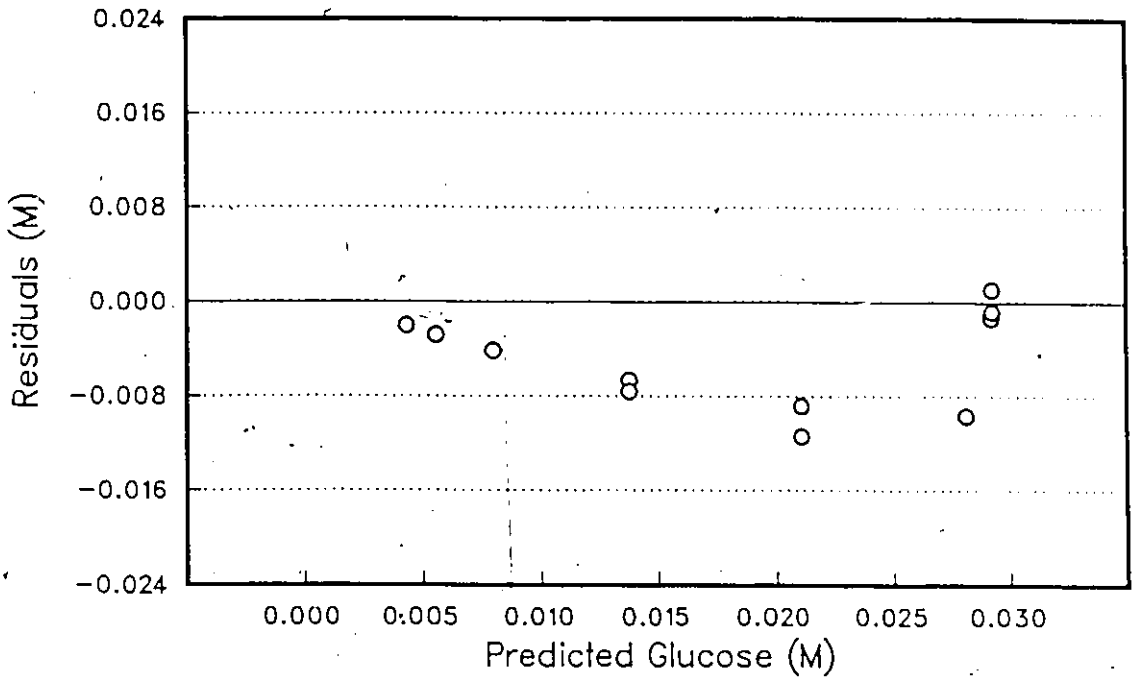


(n) Residuals vs Shaker Speed

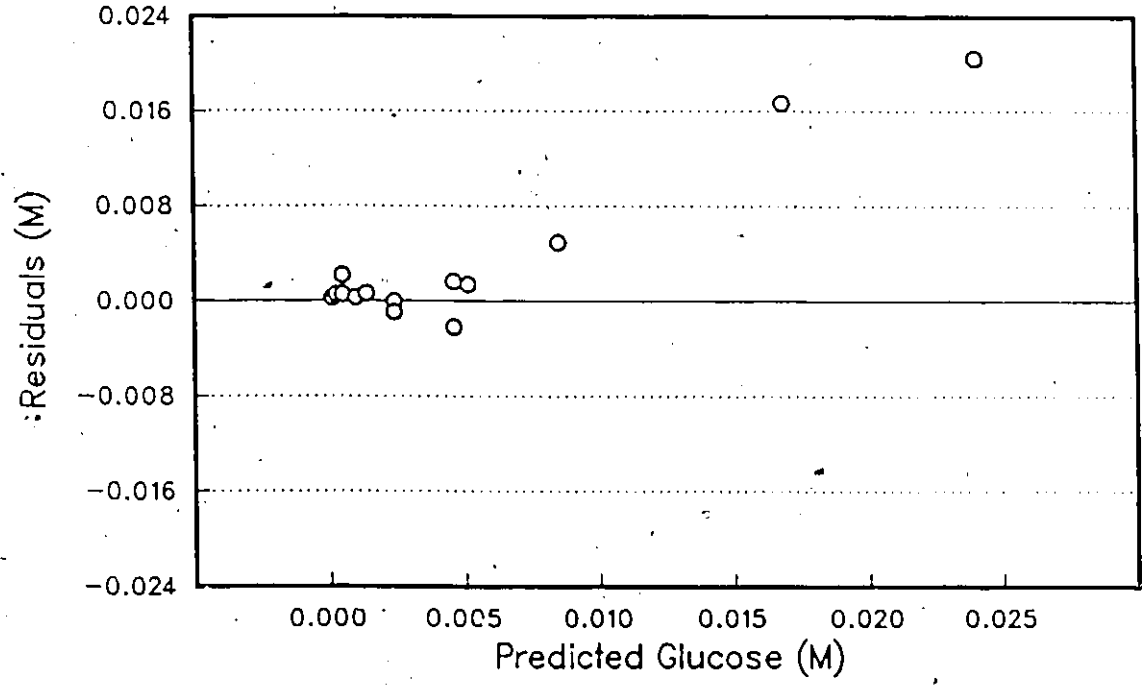


(o) Residuals vs Reactor Broth Vol.

Figure 12: Model 6 Residuals



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 13: Model 6 Residuals for Runs with no initial CMC

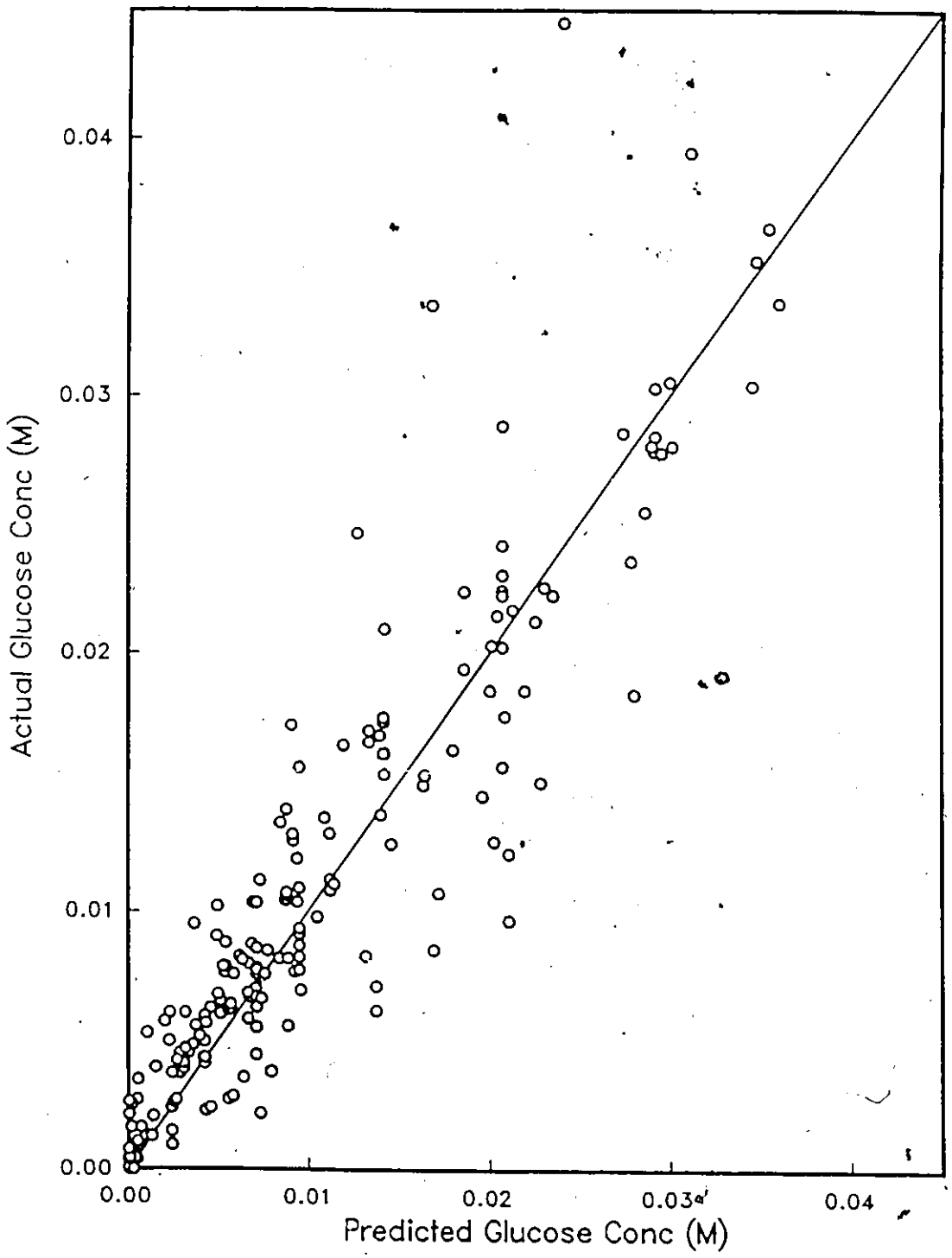


Figure 14: Actual vs Predicted Glucose for Model 6

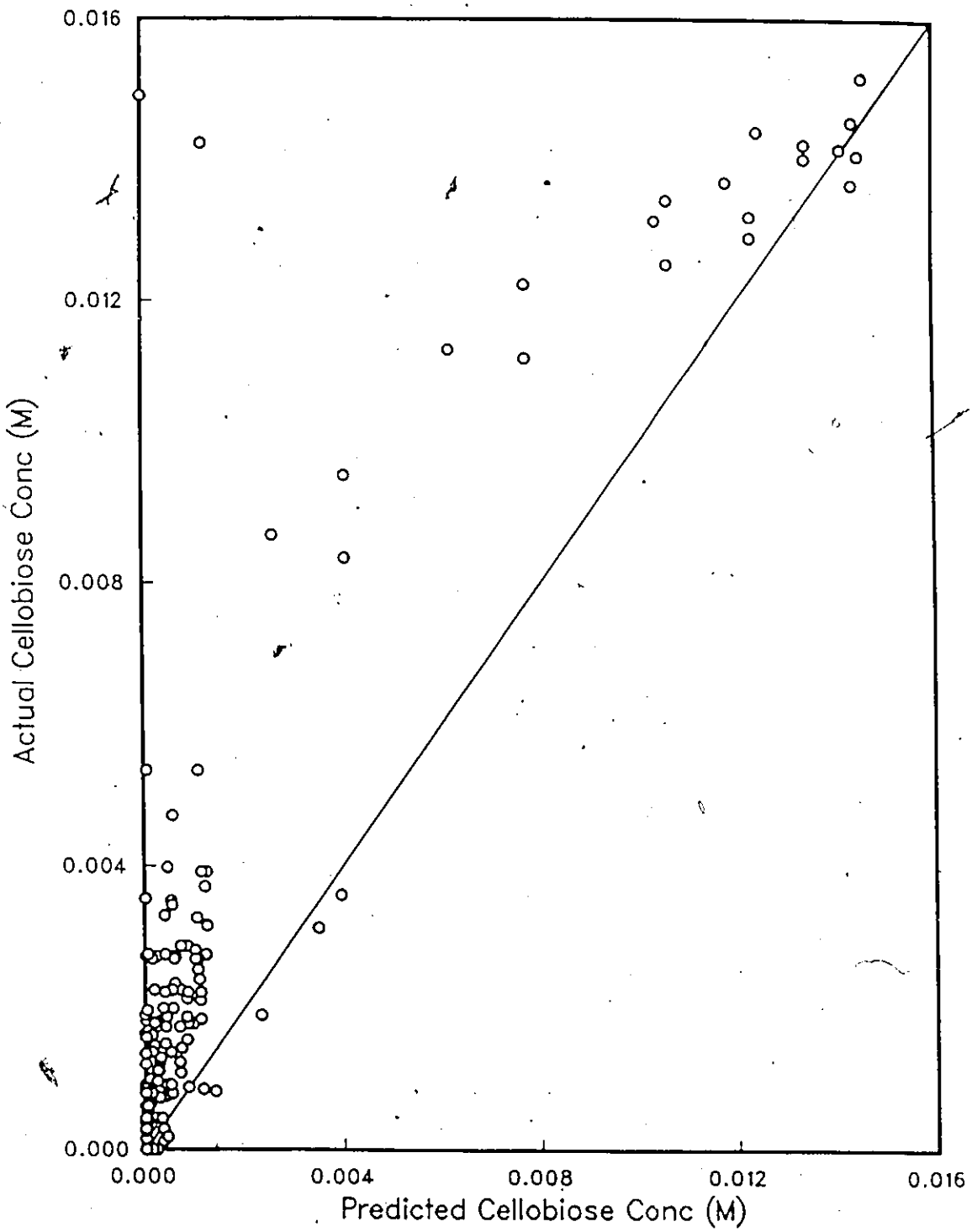


Figure 15: Actual vs Predicted Cellobiose for Model 6

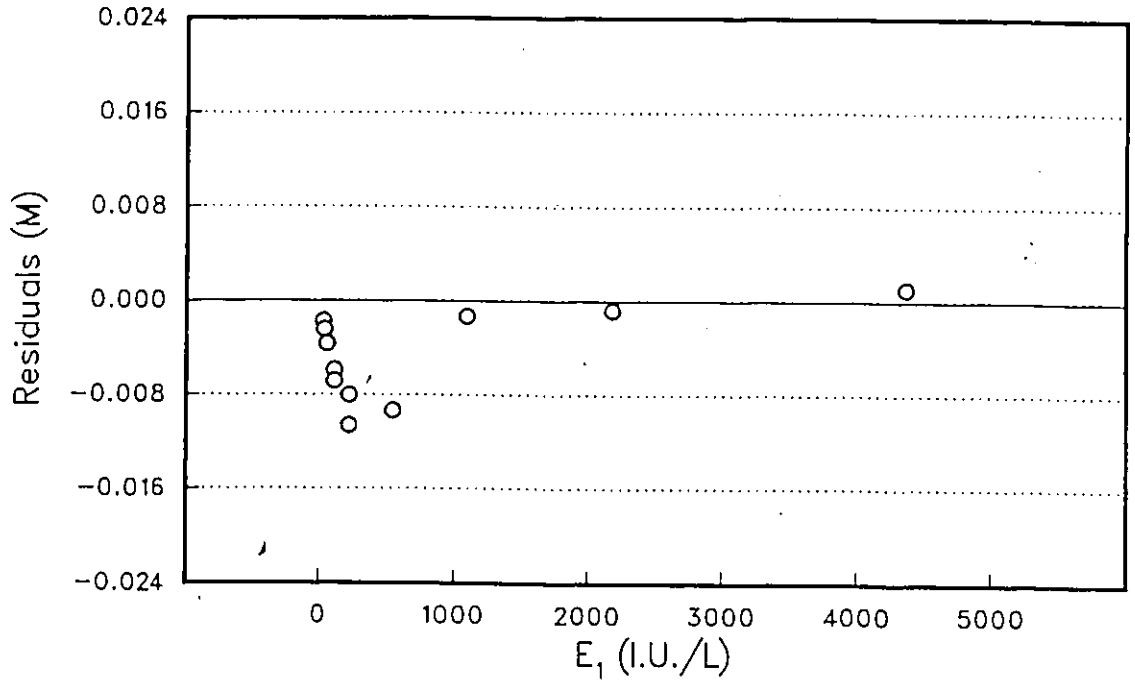
Segregated list of residuals for runs with		
Celluclast only	no enzyme	Novozym only
$0.75678 \cdot 10^{-3}$	$0.39000 \cdot 10^{-3}$	$0.10900 \cdot 10^{-2}$
$0.14909 \cdot 10^{-2}$	$0.44000 \cdot 10^{-3}$	$-0.79992 \cdot 10^{-3}$
$-0.18783 \cdot 10^{-2}$		$-0.13153 \cdot 10^{-2}$
$-0.74784 \cdot 10^{-3}$		$-0.93518 \cdot 10^{-2}$
$0.37226 \cdot 10^{-3}$		$-0.10601 \cdot 10^{-1}$
$0.60268 \cdot 10^{-3}$		$-0.68363 \cdot 10^{-2}$
$0.22327 \cdot 10^{-2}$		$-0.36599 \cdot 10^{-2}$
$0.60037 \cdot 10^{-3}$		$-0.80208 \cdot 10^{-2}$
$0.29795 \cdot 10^{-3}$		$-0.58963 \cdot 10^{-2}$
$0.21197 \cdot 10^{-1}$		$-0.36299 \cdot 10^{-2}$
$0.17524 \cdot 10^{-1}$		$-0.24557 \cdot 10^{-2}$
$0.54980 \cdot 10^{-2}$		$-0.17386 \cdot 10^{-2}$
$0.19804 \cdot 10^{-2}$		
$0.17216 \cdot 10^{-3}$		

Trends are still seen in Figures 16(a) and (b). Figure 16(a) is of more concern to us than Figure 16(b); The two highest residuals of (b) can be attributed to experimental error. The measured values of G_1 are greater than those predicted assuming all the cellobiose has been hydrolyzed. The two runs are C02101 and C02102. 52% more glucose than theoretically possible has been recorded for run C02101. These runs will not be referred to in subsequent analyses.

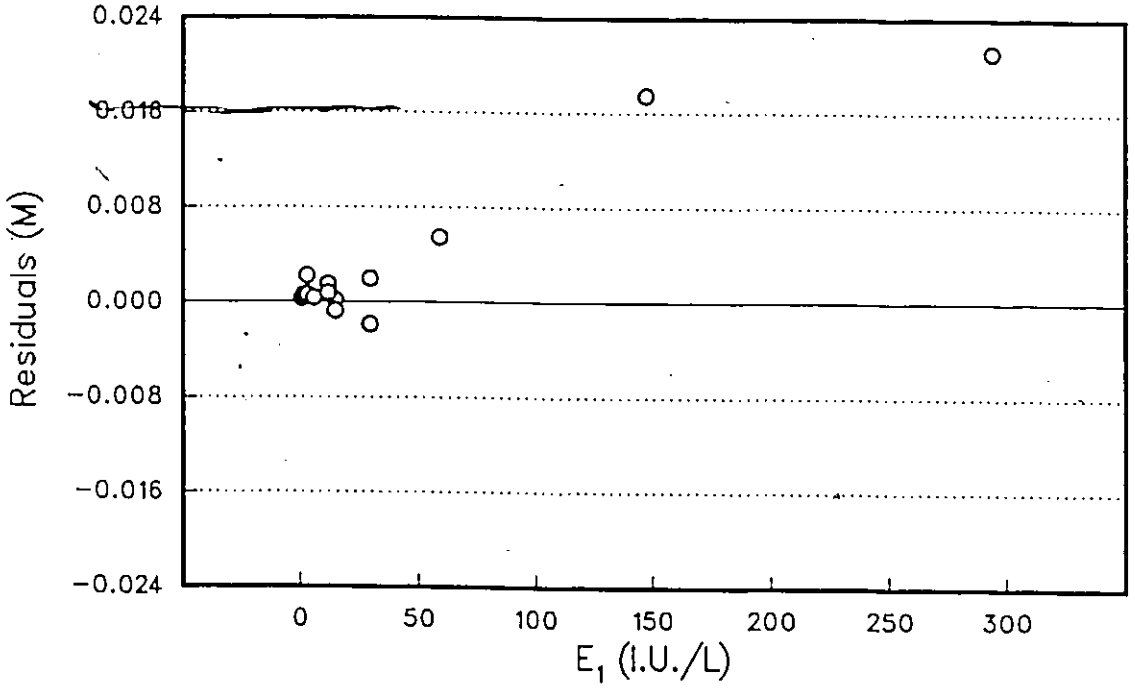
5.6.9 Model 8 Fitting

Model 8 is identical to model 2 except that it was fitted to just the 28 data points that had an initial concentration of CMC equal to zero. In this case, the objective was to find out if Michaelis-Menten kinetics with no inhibition could describe the hydrolysis of cellobiose by cellobiase. This is a two parameter model. E_x is assumed to have no effect on the hydrolysis of G_2.

In this case, the estimation routine converged at times to a physically meaningless, negative value for K_{mE_1} and at other times to positive parameter values. Nonetheless, a bias was still present in the residuals. The attempt at explaining the bias was taken one step further by introducing inhibition in the cellobiose hydrolysis



(a) Residuals vs E_1 (Novozym Runs Only Included)



(b) Residuals vs E_1 (Celluclast Runs Only Included)

Figure 16: Model 7 Residuals (Runs with no Initial CMC)

model.

5.6.10 Model 9 Fitting

This model fitted the 28 data points with no initial CMC to a model that again assumed only cellobiase cleaved cellobiose. Glucose was assumed to be a non-competitive inhibitor of cellobiase. Again, the results are not reported here but the trend in the segregated residuals was ever present.

5.6.11 Model 10 Fitting

This model fitted the 28 data points with no initial CMC to a model that again assumed only cellobiase cleaved cellobiose. This time, glucose was assumed to be a competitive inhibitor of cellobiase. Typical results are shown below.

MODEL 10

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, TOL $= 10^{-18}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_1}	$0.1 \cdot 10^{-7}$	$+0.1359 \cdot 10^{-7}$	$\pm 0.7698 \cdot 10^{-5}$
K_{mE_1}	$0.3 \cdot 10^{-2}$	$-0.5901 \cdot 10^{-4}$	$\pm 0.8136 \cdot 10^{-2}$
K_{iG_1}	$0.5 \cdot 10^{-2}$	$-0.3278 \cdot 10^{+19}$	$\pm 0.0000 \cdot 10^{+0}$
Sum of squared residuals $= 0.001070\,\mathrm{M}^2$			

The covariance matrix of the parameters:

$$\begin{pmatrix} +0.14815 \cdot 10^{-10} & & \\ -0.19659 \cdot 10^{-03} & 0.16549 \cdot 10^{-4} & \\ +0.00000 \cdot 10^{+00} & 0.00000 \cdot 10^{+0} & 0.00000 \cdot 10^{0} \end{pmatrix}$$

$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1.000 & & \\ \mathrm{NA} & 1.000 & \\ 0.000 & 0.000 & 0.000 \end{pmatrix}$$

Segregated list of residuals for runs with		
Celluclast only	no enzyme	Novozym only
$-0.24934 \cdot 10^{-2}$	$0.39000 \cdot 10^{-3}$	$0.10900 \cdot 10^{-2}$
$0.10300 \cdot 10^{-2}$	$0.44000 \cdot 10^{-3}$	$-0.80000 \cdot 10^{-3}$
$-0.49598 \cdot 10^{-3}$		$-0.13800 \cdot 10^{-2}$
$0.26675 \cdot 10^{-4}$		$-0.10820 \cdot 10^{-1}$
$0.70280 \cdot 10^{-3}$		$-0.11855 \cdot 10^{-1}$
$0.77146 \cdot 10^{-3}$		$-0.45686 \cdot 10^{-2}$
$0.24015 \cdot 10^{-2}$		$-0.15863 \cdot 10^{-2}$
$0.68568 \cdot 10^{-3}$		$-0.92751 \cdot 10^{-2}$
$0.33228 \cdot 10^{-3}$		$-0.36286 \cdot 10^{-2}$
$0.15260 \cdot 10^{-1}$		$-0.15563 \cdot 10^{-2}$
$0.19076 \cdot 10^{-1}$		$-0.84222 \cdot 10^{-3}$
$0.76456 \cdot 10^{-2}$		$-0.42416 \cdot 10^{-3}$
$0.33630 \cdot 10^{-2}$		
$-0.94667 \cdot 10^{-3}$		

The segregated residuals for model 10 are shown in figure 17. This attempt at explaining the poor prediction of the hydrolysis of cellobiose by cellobiase failed, just as models 7–9 had failed; A trend was still evident, especially in the case of the runs where celluclast was the only solution involved.

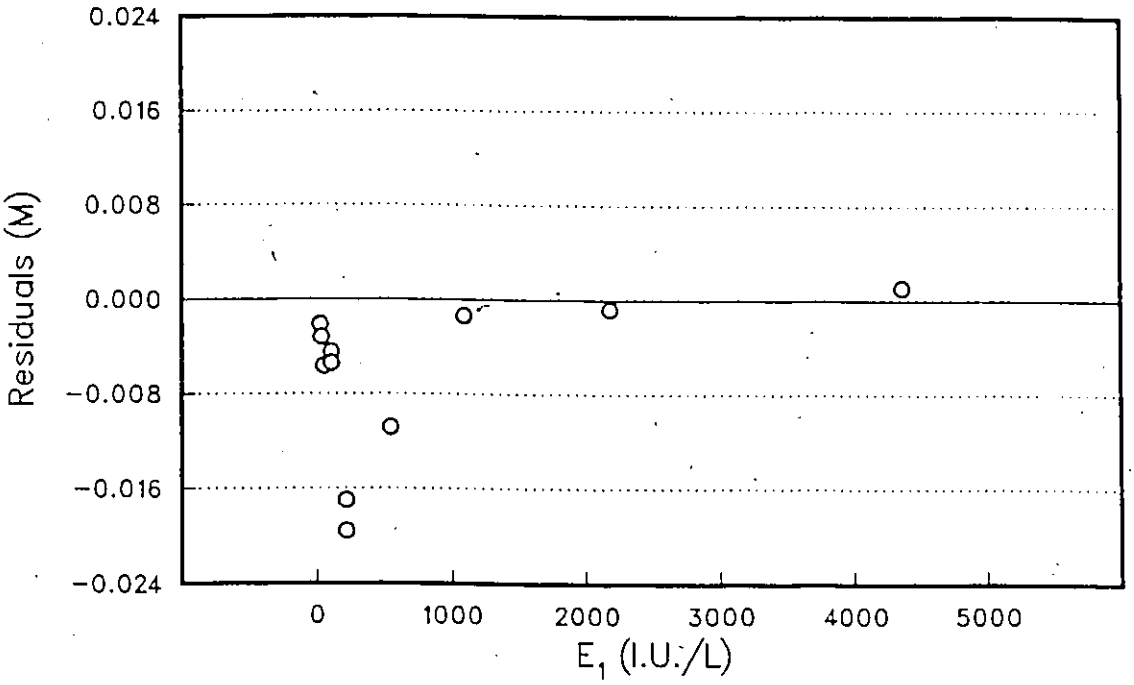
5.6.12 Correction to the Celluclast $\widehat{\text{CBA}}$

The analysis to this point has indicated that the estimation of the concentration of E_1 in either Celluclast or Novozym may be inaccurate. If the ratio of $\widehat{\text{CBA}}$ of Novozym to Celluclast has been overestimated the trend seen in all models thus far would be explained. To correct the situation, the factor by which the cellobiase activity is off is to be estimated from the 28 data points that have only cellobiose as a substrate. Furthermore, all of the error is allocated to the Celluclast $\widehat{\text{CBA}}$. Define the correction factor of $\widehat{\text{CBA}}$ as the factor by which the cellobiase activity of Celluclast must be multiplied to eliminate trends in the residuals.

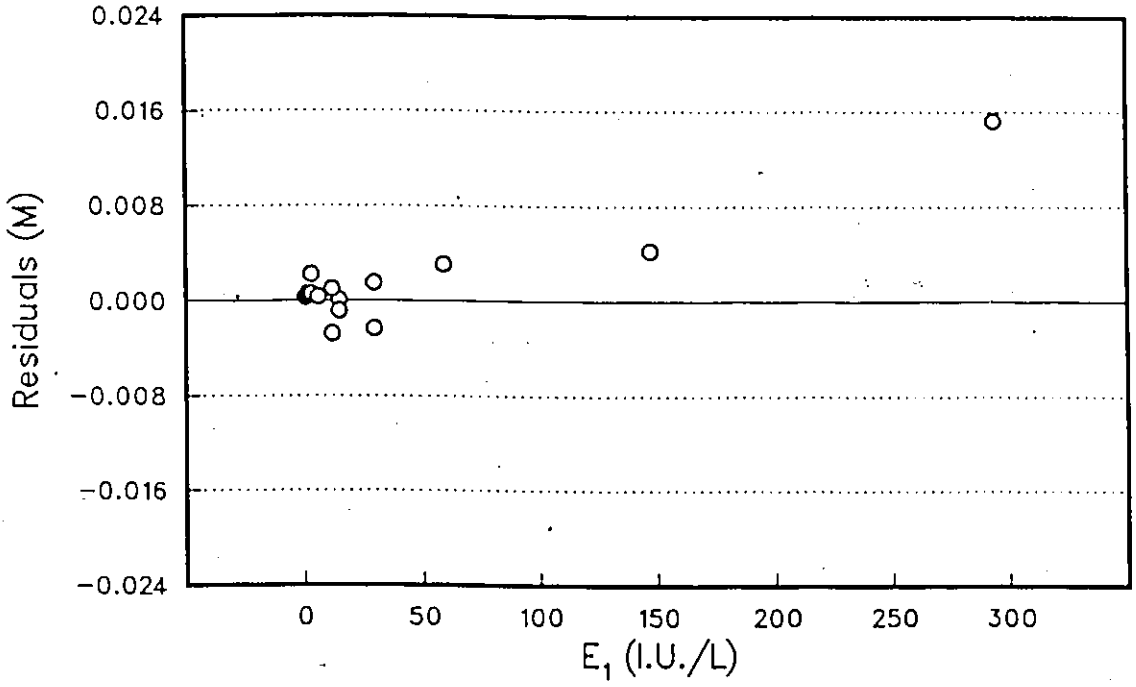
Here are some typical results of fitting to obtain a correction factor:

$\widehat{\text{CBA}}$ CORRECTION

Non-linear routine specifications: $\Delta t = 100\,\text{s}$, $\text{TOL} = 10^{-18}\,\text{M}$



(a) Residuals vs E_1 (Novozym Runs Only Included)



(b) Residuals vs E_1 (Celluclast Runs Only Included)

Figure 17: Model 10 Residuals (Runs with no Initial CMC)

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_1}	$0.5 \cdot 10^{-7}$	$0.5914 \cdot 10^{-6}$	$\pm 0.2396 \cdot 10^{-6}$
$\widehat{\text{CBA}}$ Correction	0.1	5.839	± 3.444
Sum of squared residuals = 0.000419 M^2			

The covariance matrix of the coded parameters:

$$\begin{pmatrix} +0.14347 \cdot 10^{-13} & \\ -0.14167 \cdot 10^{-13} & 0.29654 \cdot 10^{-13} \end{pmatrix}$$

The correlation matrix of the parameters $= \begin{pmatrix} 1.000 & \\ -0.687 & 1.000 \end{pmatrix}$

Segregated list of residuals for runs with		
Celluclast only	no enzyme	Novozym only
$-0.52811 \cdot 10^{-3}$	$0.39000 \cdot 10^{-3}$	$0.10906 \cdot 10^{-2}$
$0.10475 \cdot 10^{-2}$	$0.44000 \cdot 10^{-3}$	$-0.59870 \cdot 10^{-3}$
$-0.65914 \cdot 10^{-2}$		$0.12640 \cdot 10^{-2}$
$-0.34217 \cdot 10^{-2}$		$-0.18493 \cdot 10^{-2}$
$-0.78133 \cdot 10^{-3}$		$-0.12581 \cdot 10^{-2}$
$0.11197 \cdot 10^{-4}$		$0.87188 \cdot 10^{-4}$
$0.16412 \cdot 10^{-2}$		$0.56637 \cdot 10^{-3}$
$0.30076 \cdot 10^{-3}$		$0.13219 \cdot 10^{-2}$
$0.17721 \cdot 10^{-3}$		$0.10272 \cdot 10^{-2}$
$0.15882 \cdot 10^{-1}$		$0.59637 \cdot 10^{-3}$
$0.87784 \cdot 10^{-2}$		$0.54608 \cdot 10^{-3}$
$-0.18388 \cdot 10^{-2}$		$0.61479 \cdot 10^{-3}$
$-0.27339 \cdot 10^{-2}$		
$-0.25017 \cdot 10^{-2}$		

Several fits of the data and several proposed mechanisms for E_1 were tested. Although the specific example shown above found a correction factor of 5.839 to best fit the data, the average correction factor for all trials was found to be 5.1. This approach of assuming a wrong $\widehat{\text{CBA}}$ for celluclast has eliminated the trend in the residuals for the cellobiase hydrolysis runs. Consequently, a correction factor of 5 is used in subsequent modelling.

5.6.13 Model 11 Fitting

Model 11 uses the adjusted (by a factor of 5) $\widehat{\text{CBA}}$ of Celluclast. And a new assumption has been made to account for the relatively high concentration of G_2. It is assumed that E_x cleaves only β_u and β_{re} but forms complexes with all β bonds. Furthermore this is assumed to be a zero-order reaction. The other cleavable bonds, G_2, β_{nre} and β_{nres} are all cleaved by E_1 with first-order kinetics. No inhibitors of cellobiase are assumed to be present in the reaction mixture. The results follow:

MODEL 11

Non-linear routine specifications: $\Delta t = 50\,\text{s}$, TOL $= 10^{-19}\,\text{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_x}	$0.2 \cdot 10^{-6}$	$0.4301 \cdot 10^{-6}$	$\pm 0.1487 \cdot 10^{-6}$
k_{1E_1}	$0.8 \cdot 10^{-6}$	$0.6537 \cdot 10^{-6}$	$\pm 0.9897 \cdot 10^{-7}$
Sum of squared residuals $= 0.002288\,\text{M}^2$			

The covariance matrix of the parameters:

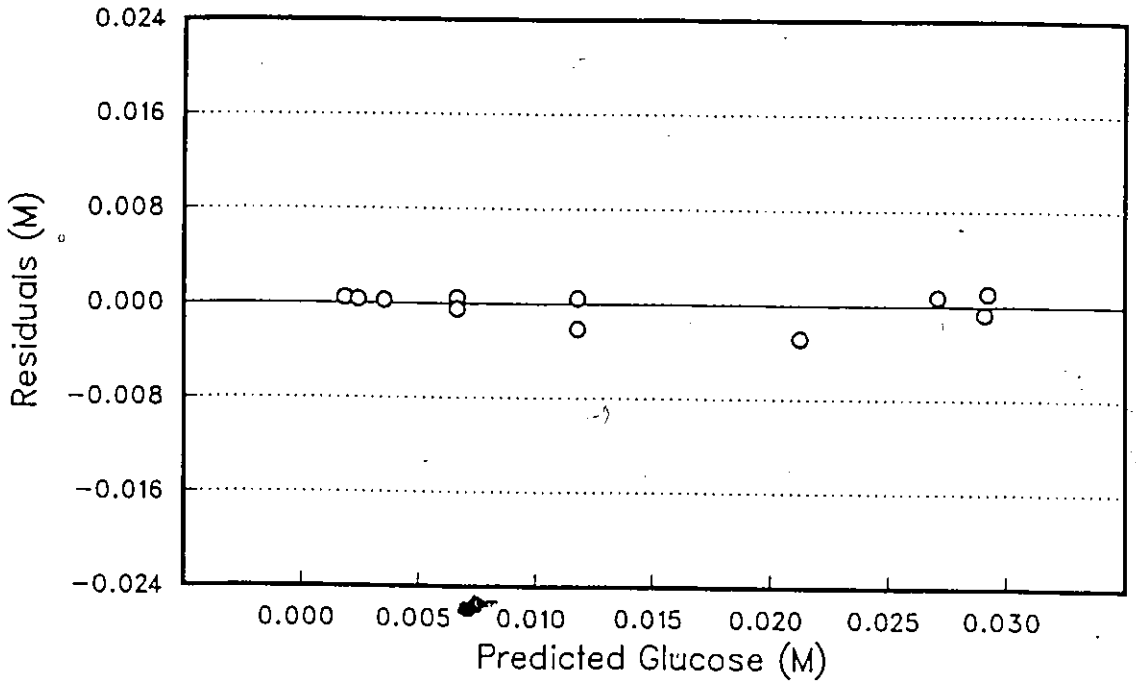
$$\begin{pmatrix} +0.56938 \cdot 10^{-14} & \\ -0.79447 \cdot 10^{-15} & 0.25239 \cdot 10^{-14} \end{pmatrix}$$

$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1 & \\ -0.210 & 1 \end{pmatrix}$$

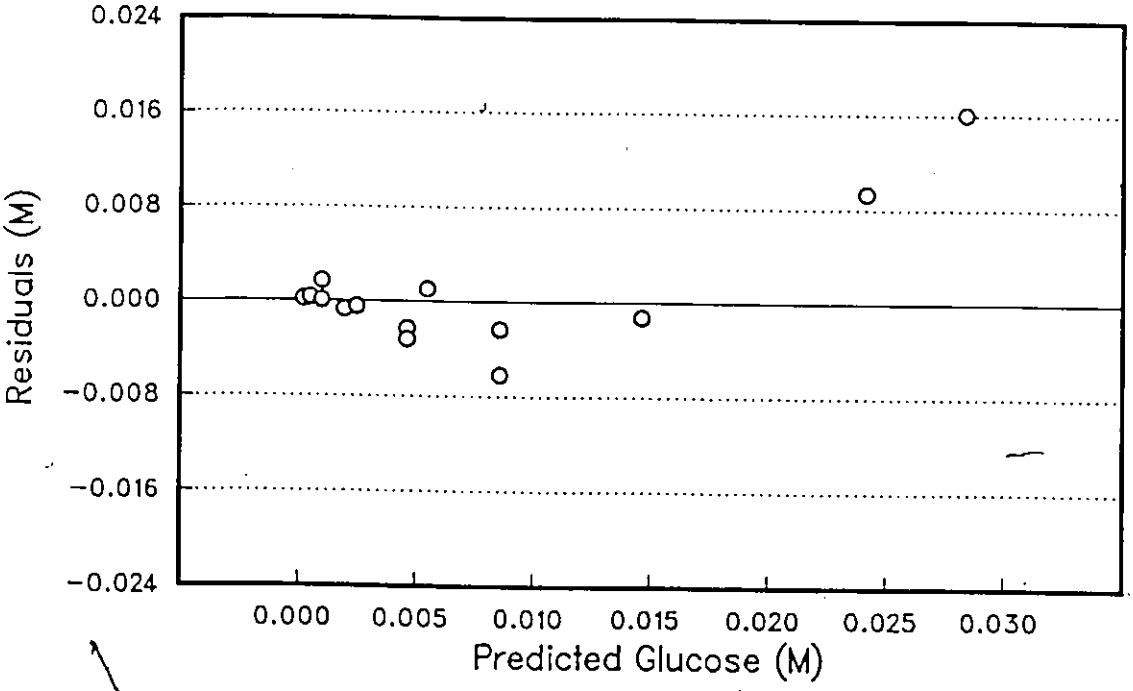
Note the reduction in the SSR for model 11 compared with models 1–6. The reduction can be attributed to the better fit for the runs with G_2 as the only substrate. However, the trend in the cellobiose prediction has still not been eliminated. Pertinent plots for model diagnostic can be found in Figures 18, 19 and 20.

5.6.14 Model 12 Fitting

Model 12 is equivalent to model 6 but with the modified $\widehat{\text{CBA}}$ for Celluclast. The data adjustment does not mean that trends in the cellobiose residuals of model 6 (now model 12) will be eliminated. Model 13, our final model will attempt to account for that underprediction of the cellobiose concentration. Results of the estimation and residual plots follow.



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 18: Model 11 Residuals for Runs with no Initial CMC

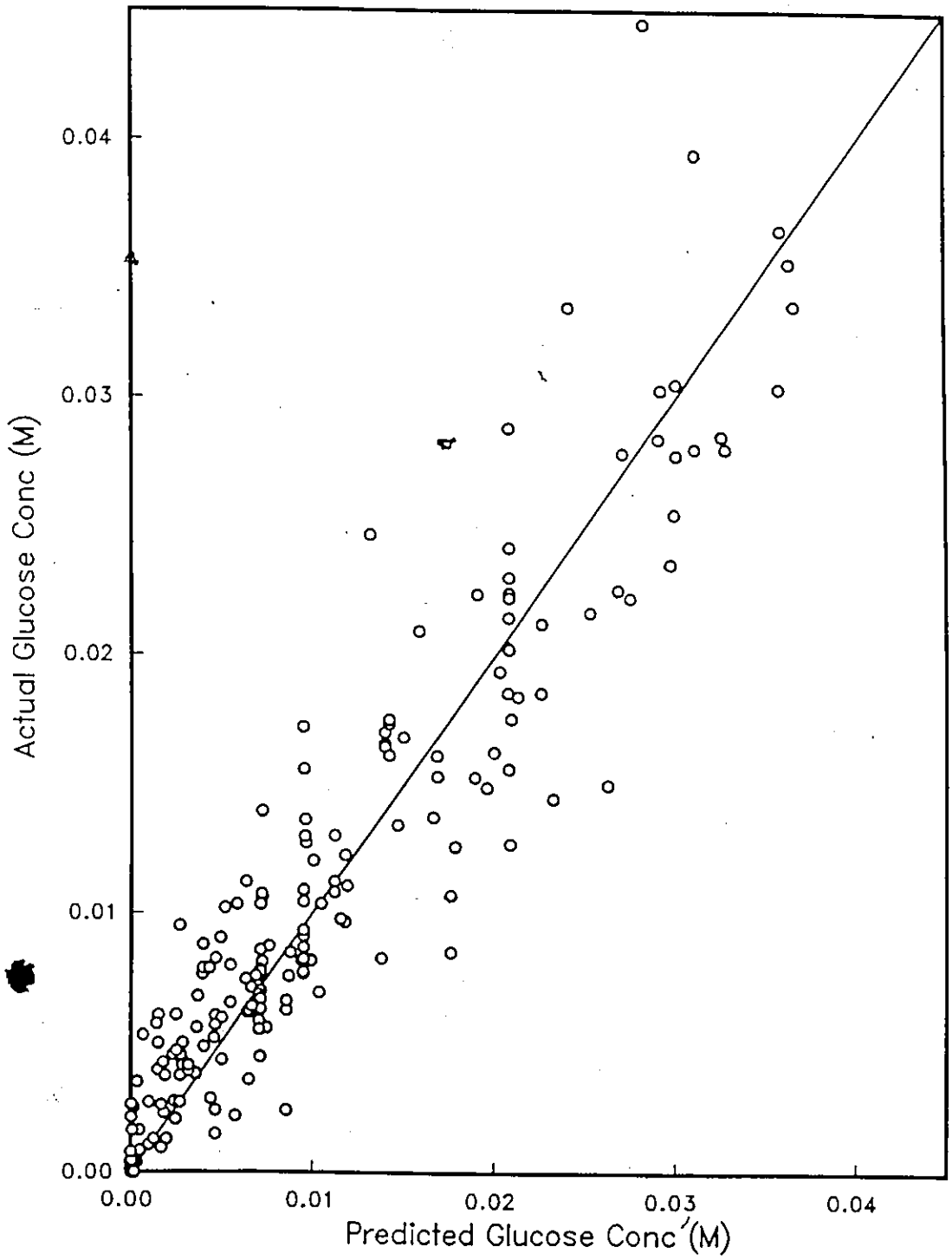


Figure 19: Actual vs Predicted Glucose for Model 11

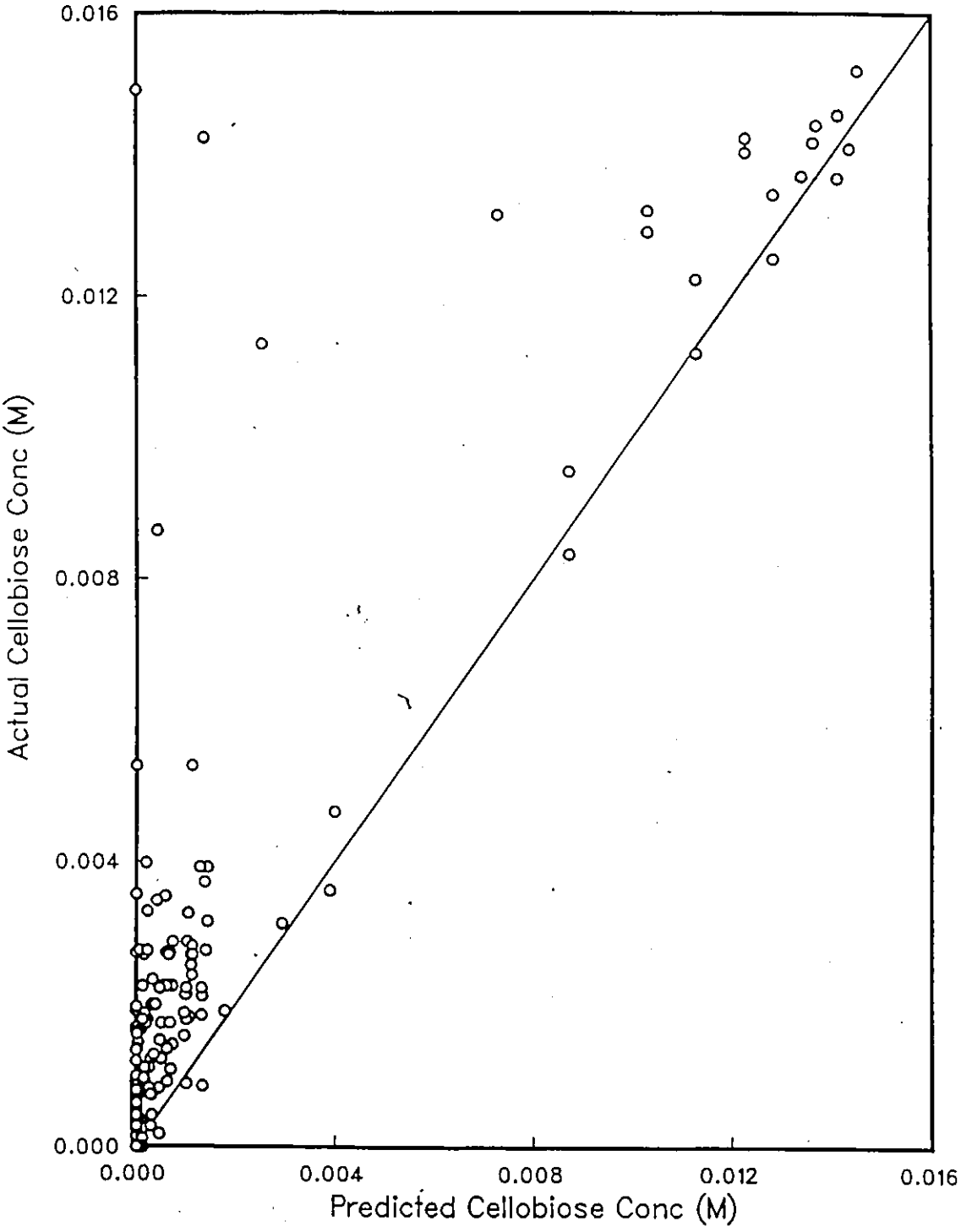


Figure 20: Actual vs Predicted Cellobiose for Model 11

MODEL 12

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, $\mathrm{TOL} = 10^{-19}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_2}	$0.8 \cdot 10^{-6}$	$0.3630 \cdot 10^{-6}$	$\pm 0.9856 \cdot 10^{-7}$
k_{1E_1}	$0.8 \cdot 10^{-6}$	$0.5196 \cdot 10^{-6}$	$\pm 0.8223 \cdot 10^{-7}$
Sum of squared residuals $= 0.001979\,\mathrm{M}^2$			

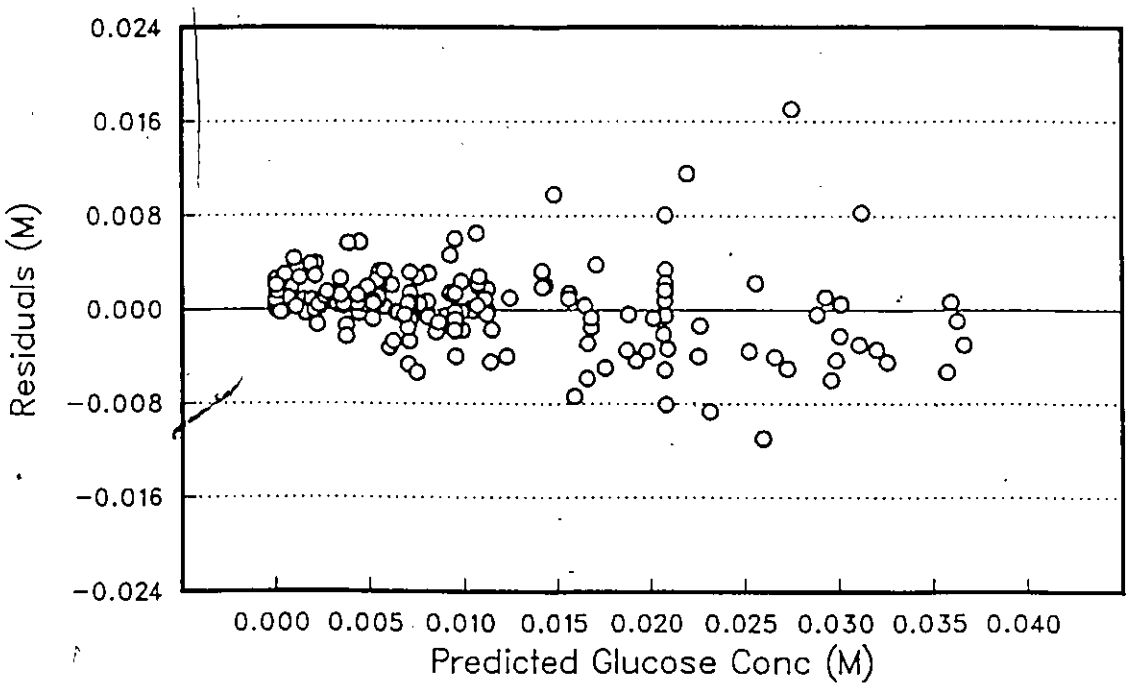
Two different starting points—the other being $(0.43008 \cdot 10^{-6}, 0.65368 \cdot 10^{-6})$—gave exactly the same results as shown here.

The covariance matrix of the parameters:

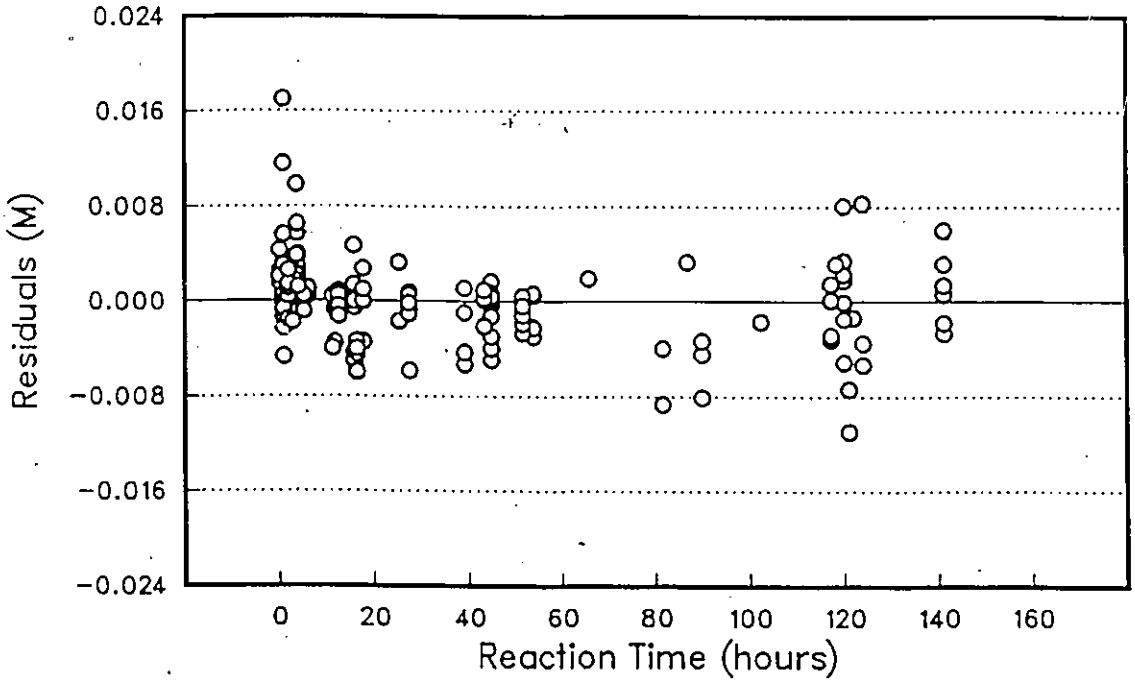
$$\begin{pmatrix} +0.25029 \cdot 10^{-14} & \\ -0.43937 \cdot 10^{-15} & 0.17421 \cdot 10^{-14} \end{pmatrix}$$

$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1 & \\ -0.210 & 1 \end{pmatrix}$$

Again, there are no trends in the residuals save for the fact that cellobiose concentration is underpredicted by the model, as can be seen in figure 24.

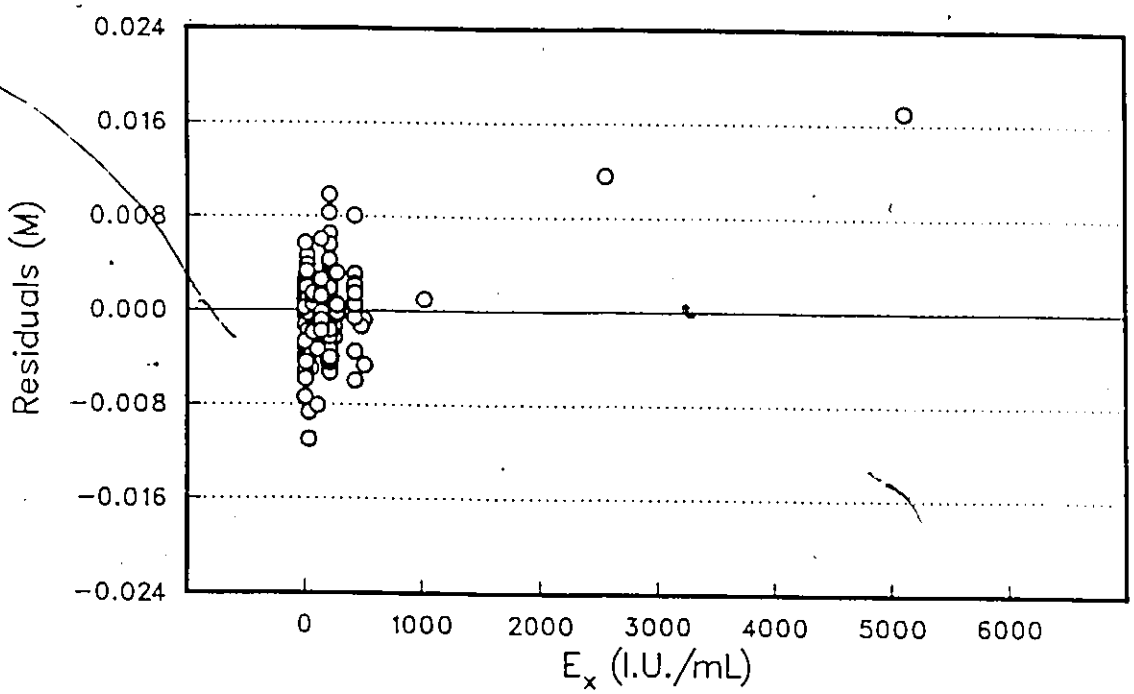


(a) Residuals vs Predicted Glucose

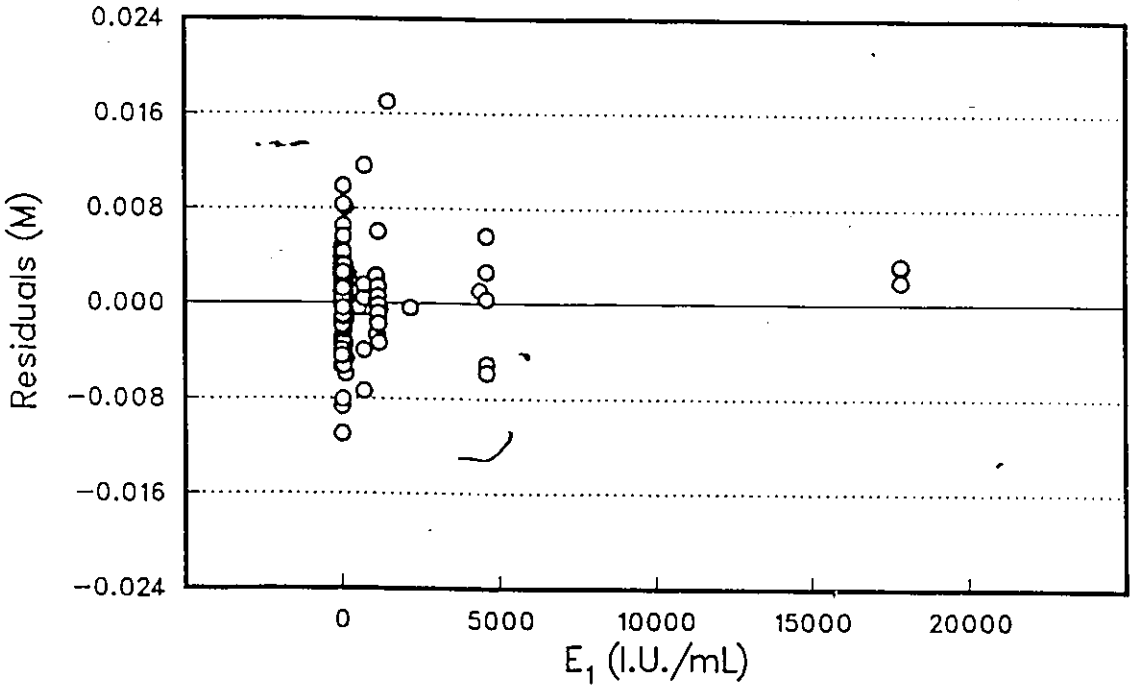


(b) Residuals vs Reaction Time

Figure 21: Model 12 Residuals

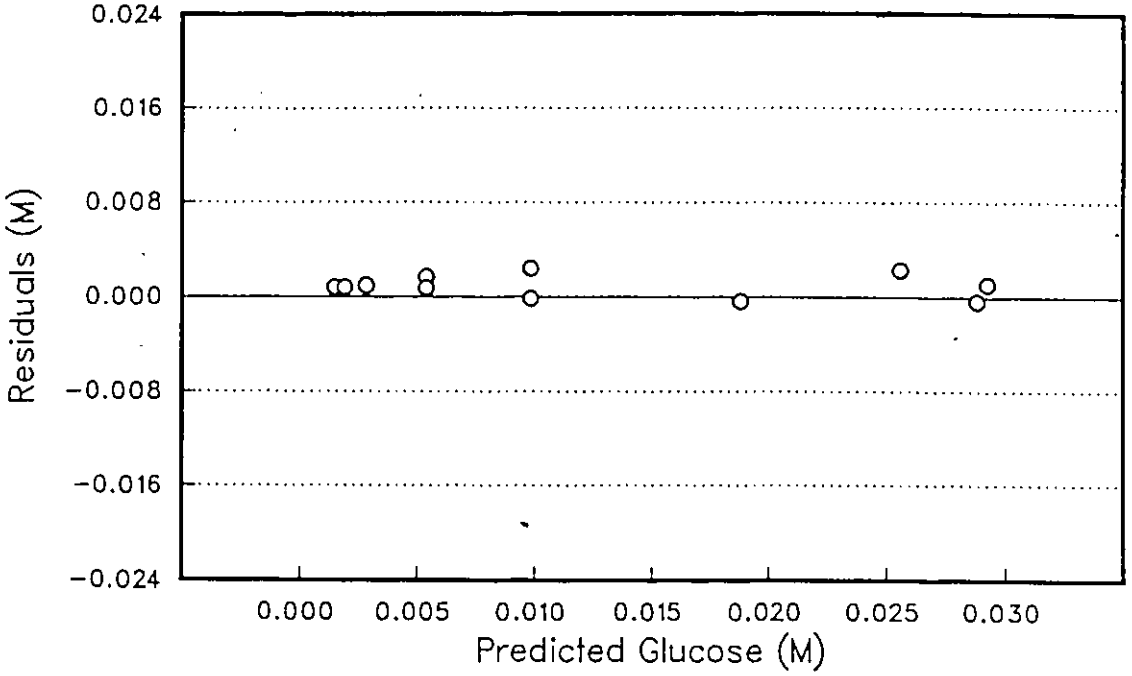


(c) Residuals vs E_x

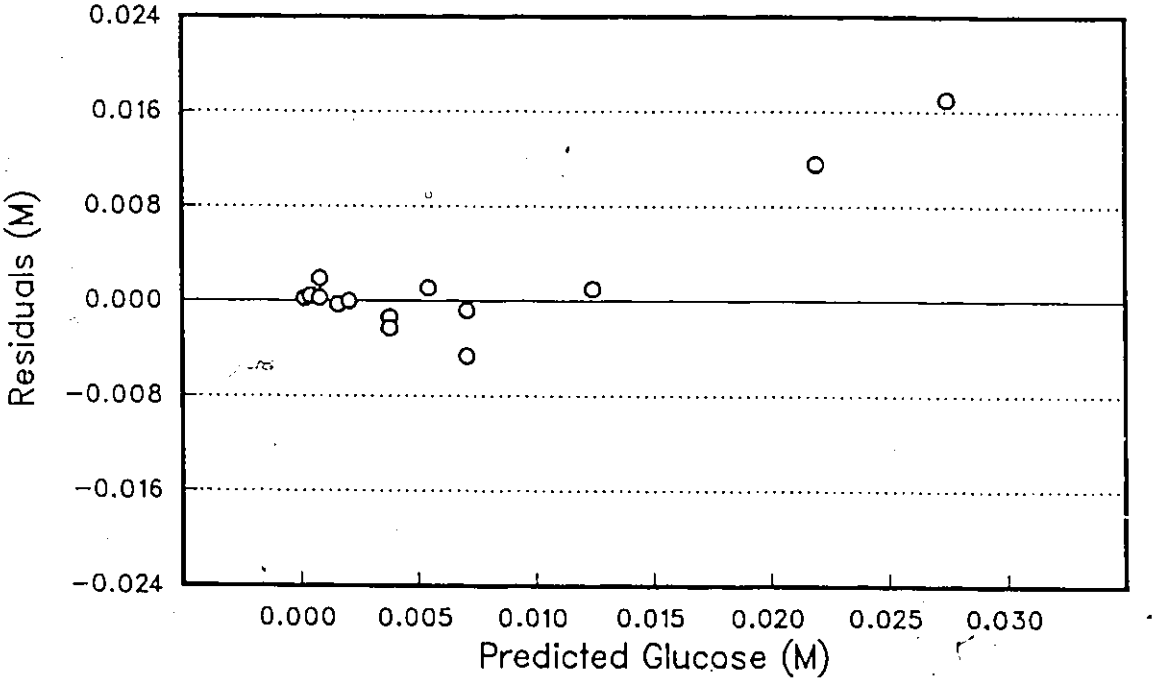


(d) Residuals vs E_1

Figure 21: Model 12 Residuals



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 22: Model 12 Residuals for Runs with no Initial CMC

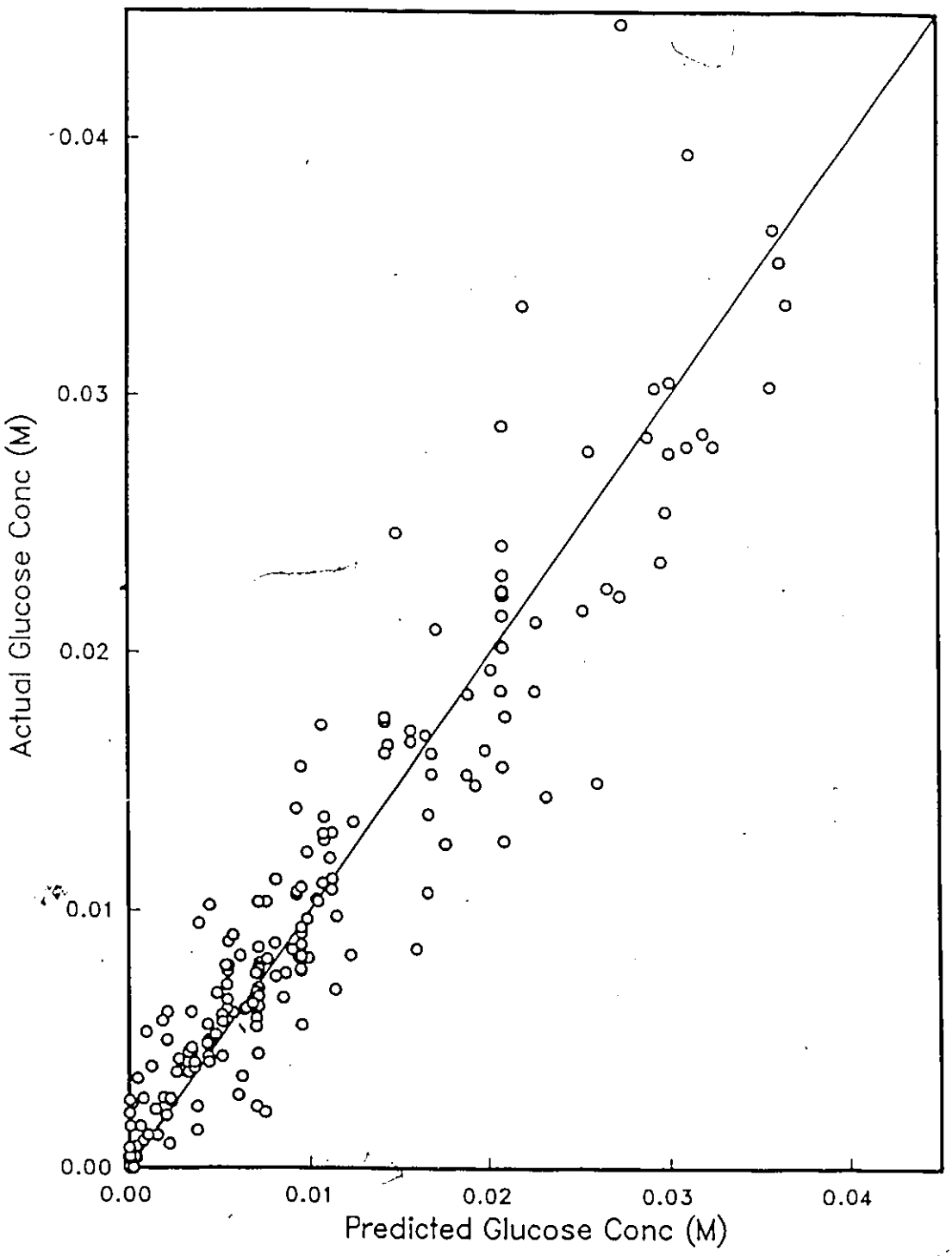


Figure 23: Actual vs Predicted Glucose for Model 12

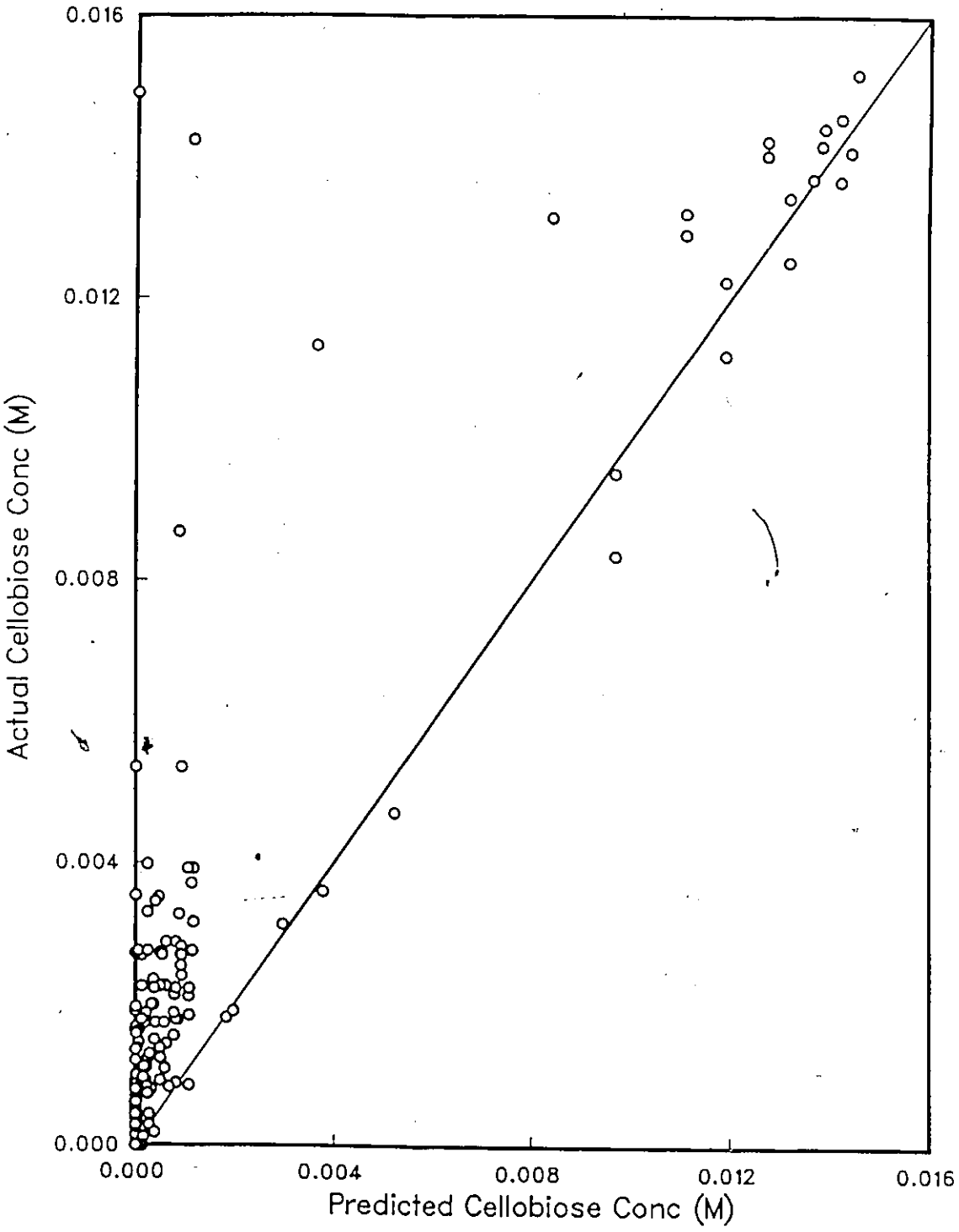


Figure 24: Actual vs Predicted Cellobiose for Model 12

5.6.15 Model 13 fitting

Model 13 attempts to account for the high cellobiose concentration by assuming that all β_{nre}-bonds are transformed into G_2 by E_2 in solution as quickly as they are produced. Model 12, the best glucose prediction model thus far, has been modified accordingly. Consult the Theory section of this thesis for the mathematical development of this mechanism. The results follow:

MODEL 13

Non-linear routine specifications: $\Delta t = 50\,\mathrm{s}$, TOL $= 10^{-19}\,\mathrm{M}$

Parameter	Initial Value	Converged Value	95% Confidence
k_{1E_2}	$0.35 \cdot 10^{-6}$	$0.3660 \cdot 10^{-6}$	$\pm 0.1371 \cdot 10^{-6}$
k_{1E_1}	$0.60 \cdot 10^{-6}$	$0.6588 \cdot 10^{-6}$	$\pm 0.1005 \cdot 10^{-6}$
Sum of squared residuals $= 0.002328\,\mathrm{M}^2$			

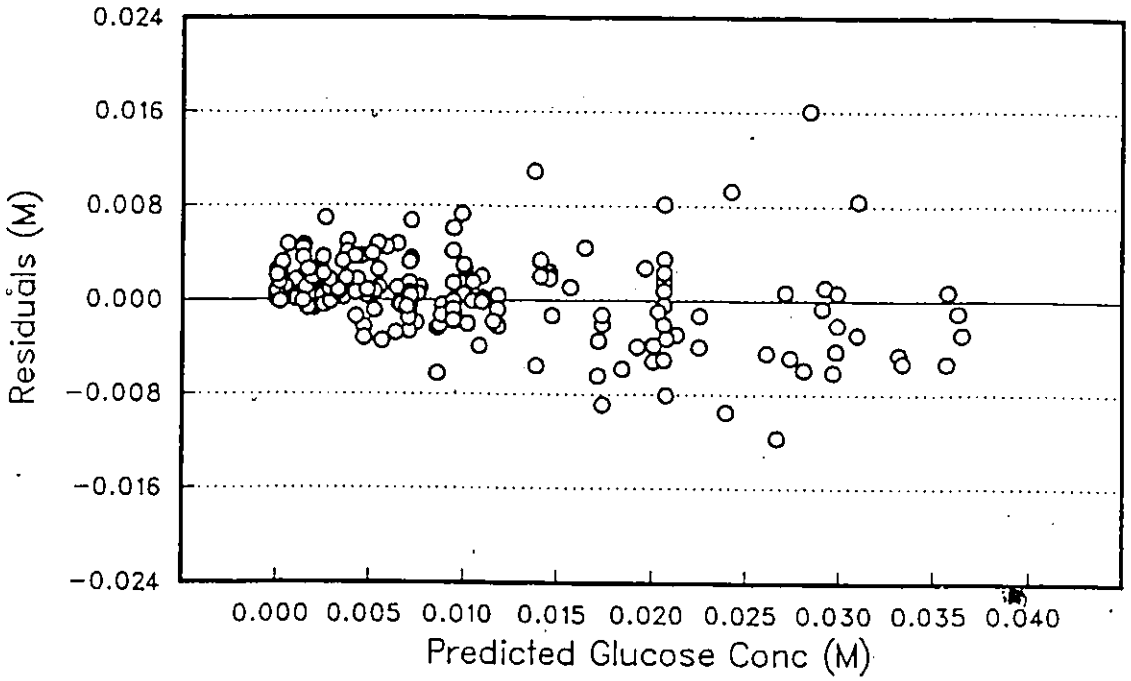
The covariance matrix of the parameters:

$$\begin{pmatrix} +0.48419 \cdot 10^{-14} & \\ -0.84981 \cdot 10^{-15} & 0.26021 \cdot 10^{-14} \end{pmatrix}$$

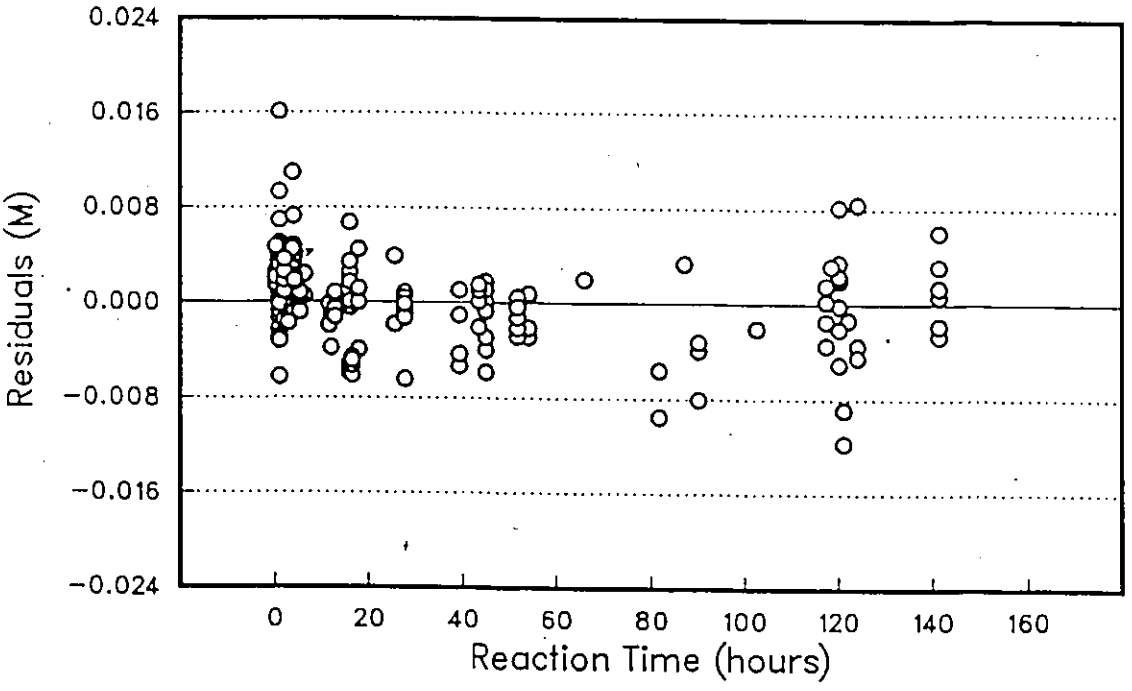
$$\text{The correlation matrix of the parameters} = \begin{pmatrix} 1 & \\ -0.239 & 1 \end{pmatrix}$$

The full complement of residual plots can be found in Figures 25(a)-(o). No trends can be seen in the residual plots of Figure 25. Thus it is unlikely that effect of these variables on the kinetics has been misjudged. The lack of trends in Figures 26(a) and (b) shows that the $\widehat{\mathrm{CBA}}$ of celluclast correction accounted for the previous trends in those residuals.

Finally, the actual vs predicted plots for glucose and cellobiose, Figures 27 and 28, indicate good correlation between the data and the model, although the latter plot illustrates the high relative variation in the G_2 data.

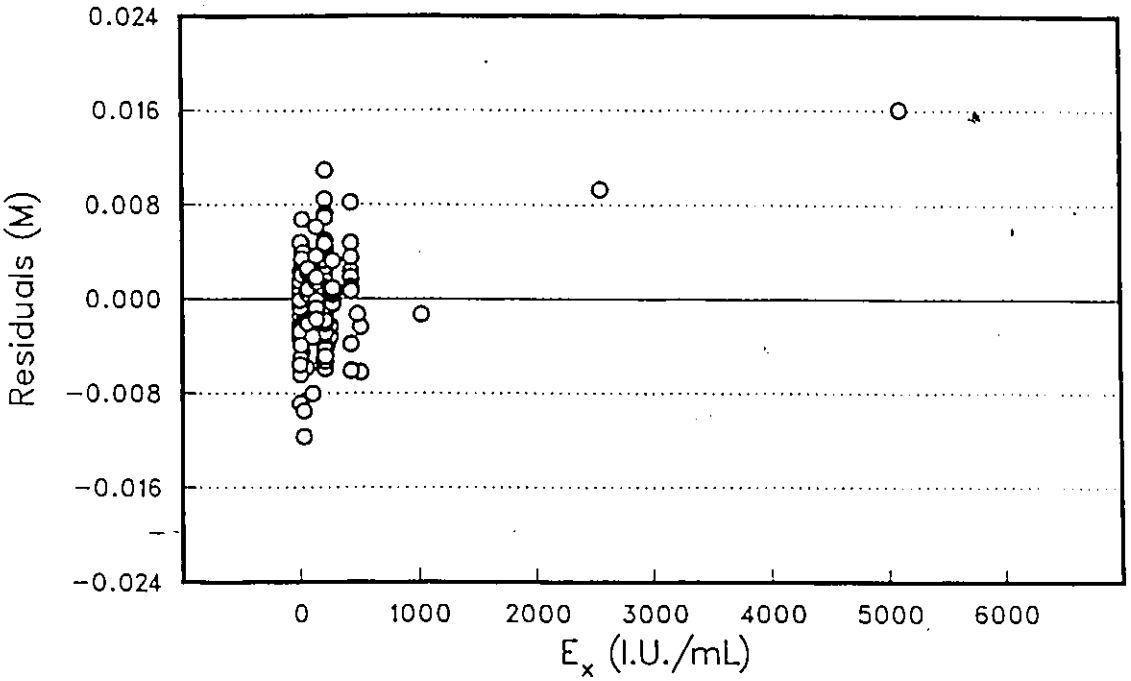


(a) Residuals vs Predicted Glucose

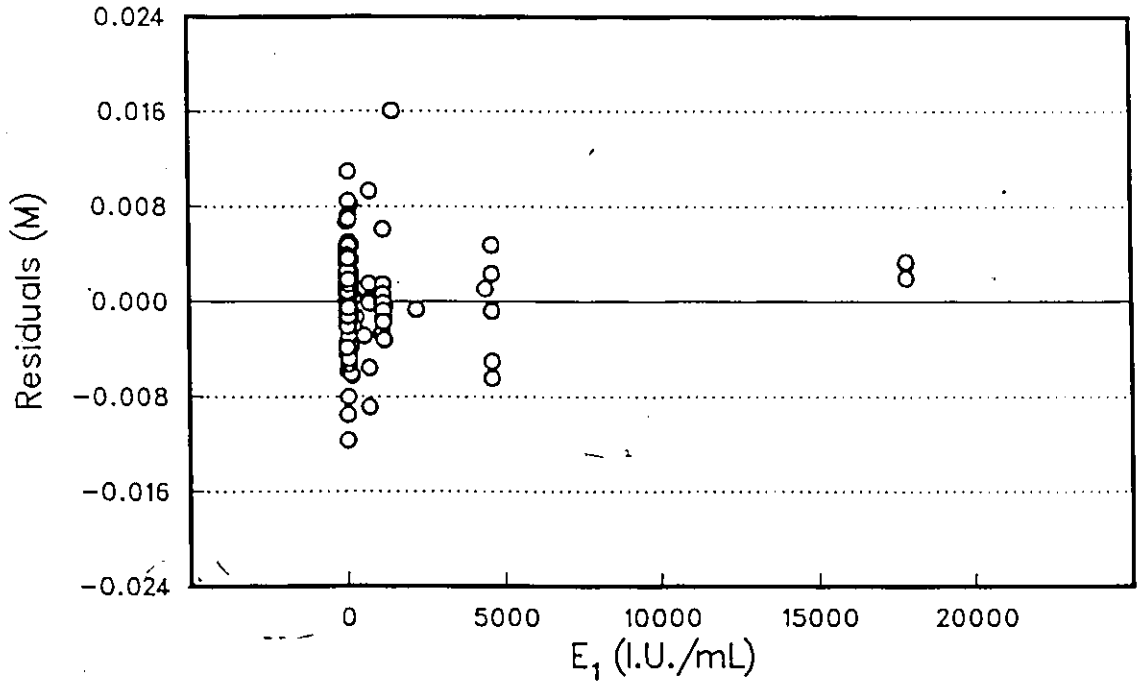


(b) Residuals vs Reaction Time

Figure 25: Model 13 Residuals

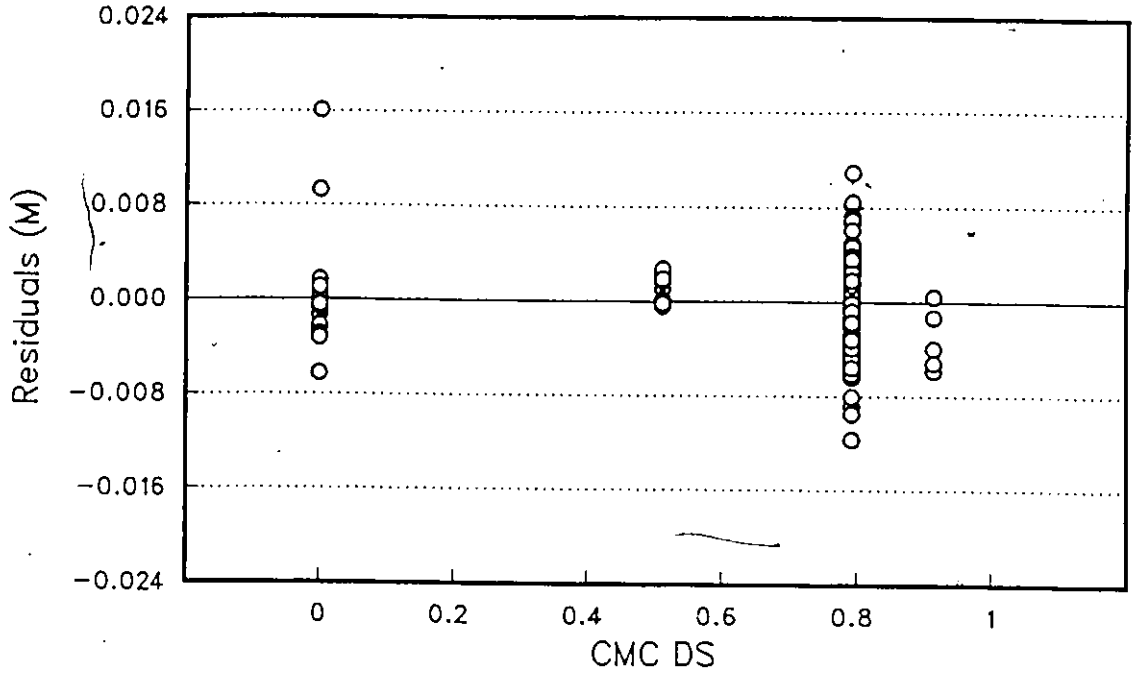


(c) Residuals vs E_x

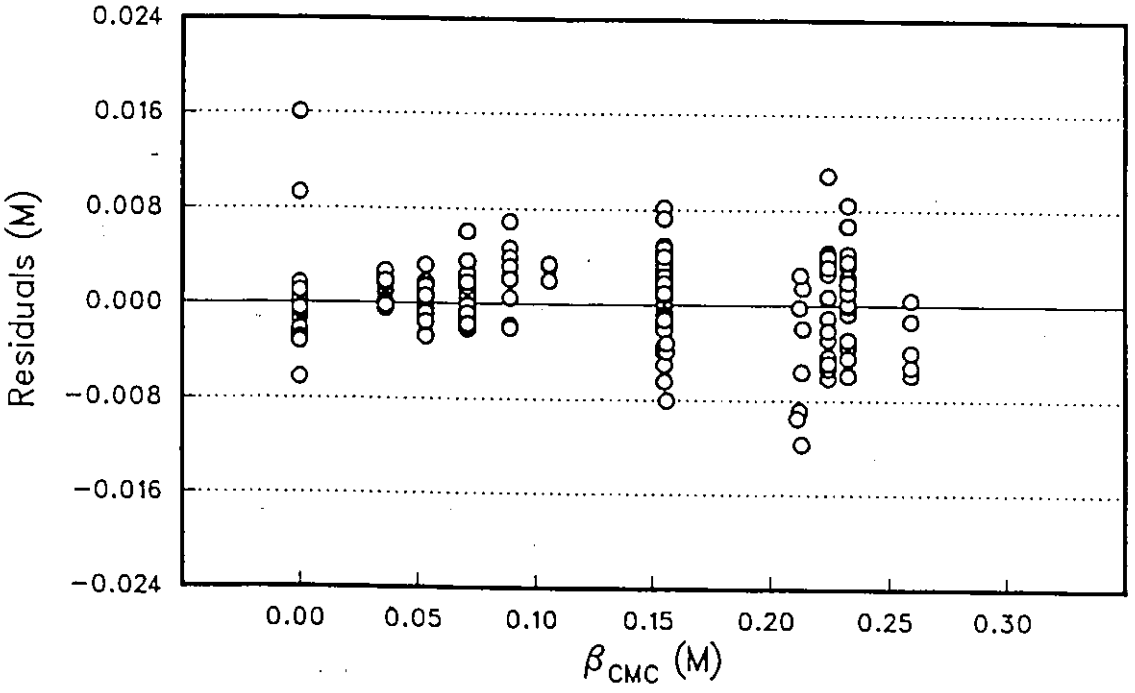


(d) Residuals vs E_1

Figure 25: Model 13 Residuals

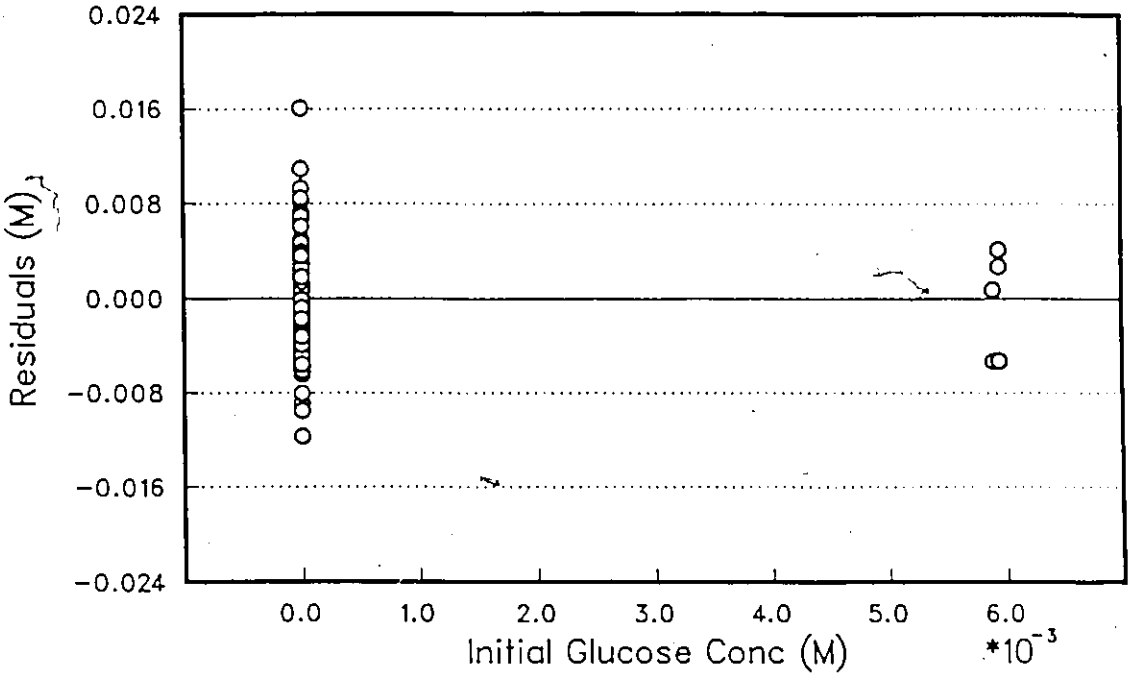


(e) Residuals vs CMC DS

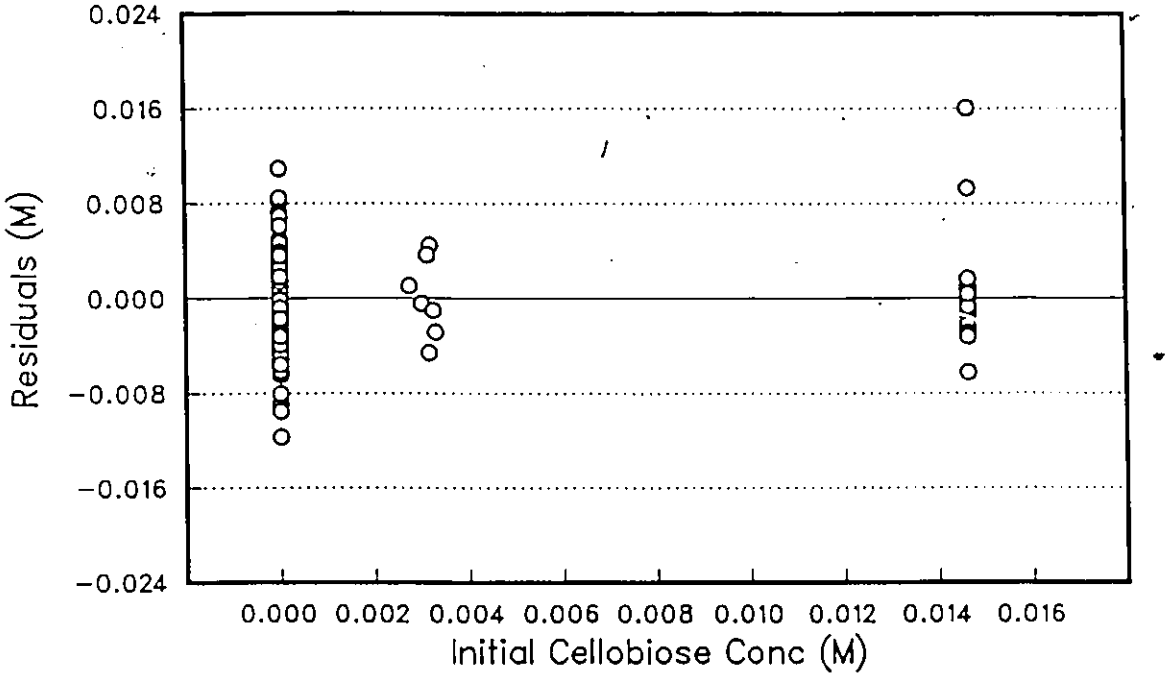


(f) Residuals vs β_{CMC}

Figure 25: Model 13 Residuals

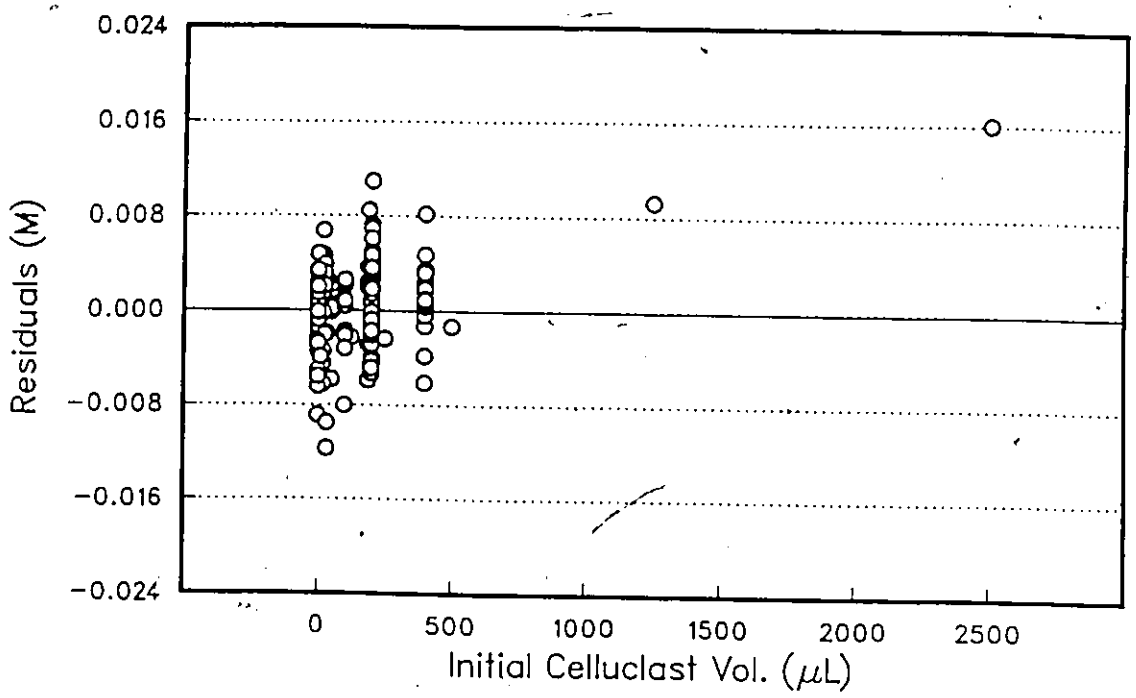


(g) Residuals vs Initial Glucose Concentration

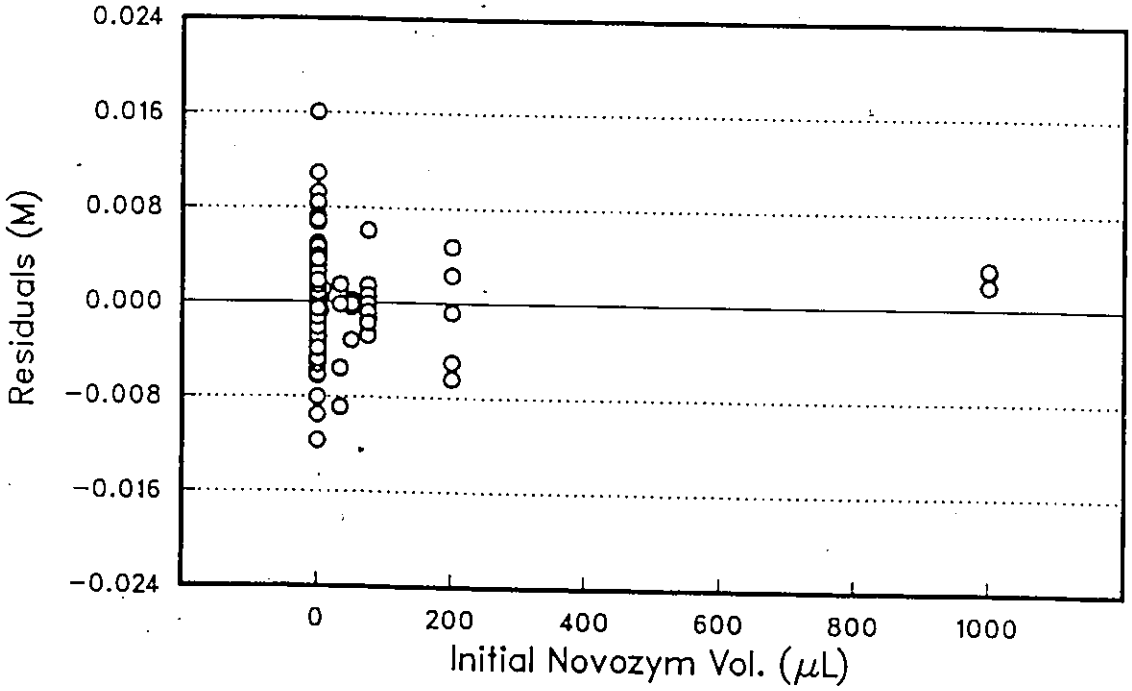


(h) Residuals vs Initial Cellobiose Conc.

Figure 25: Model 13 Residuals

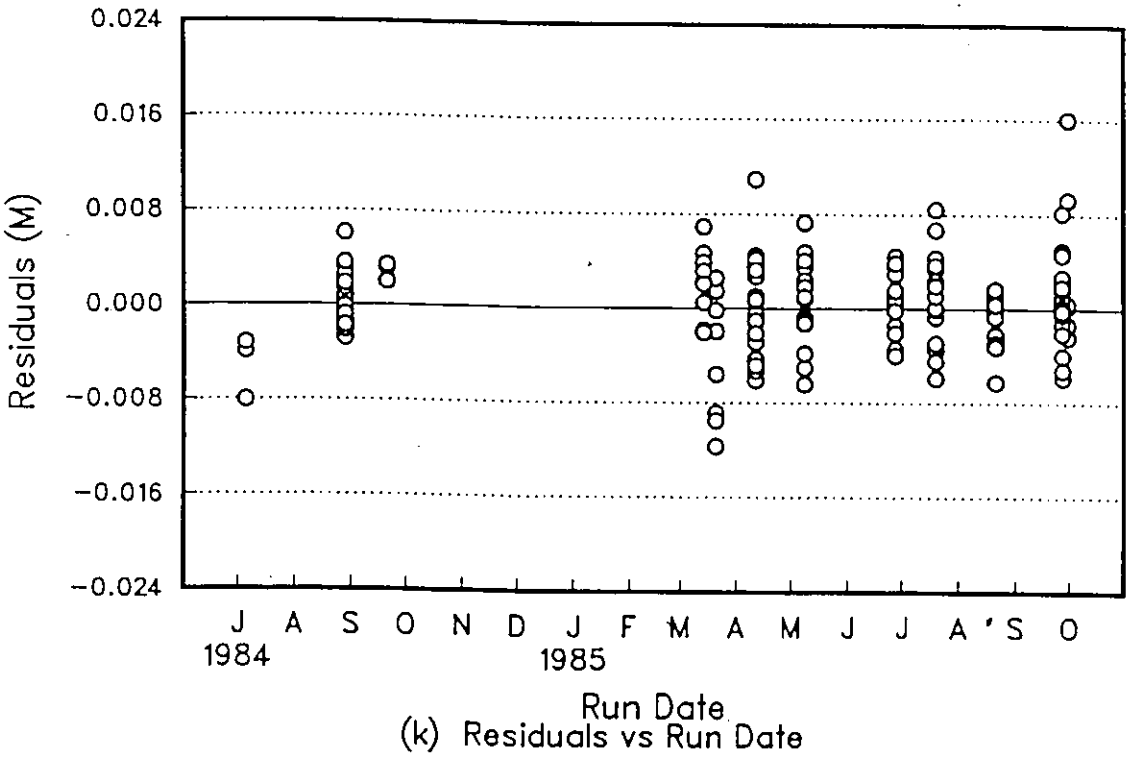


(i) Residuals vs Initial Celluclast Vol.



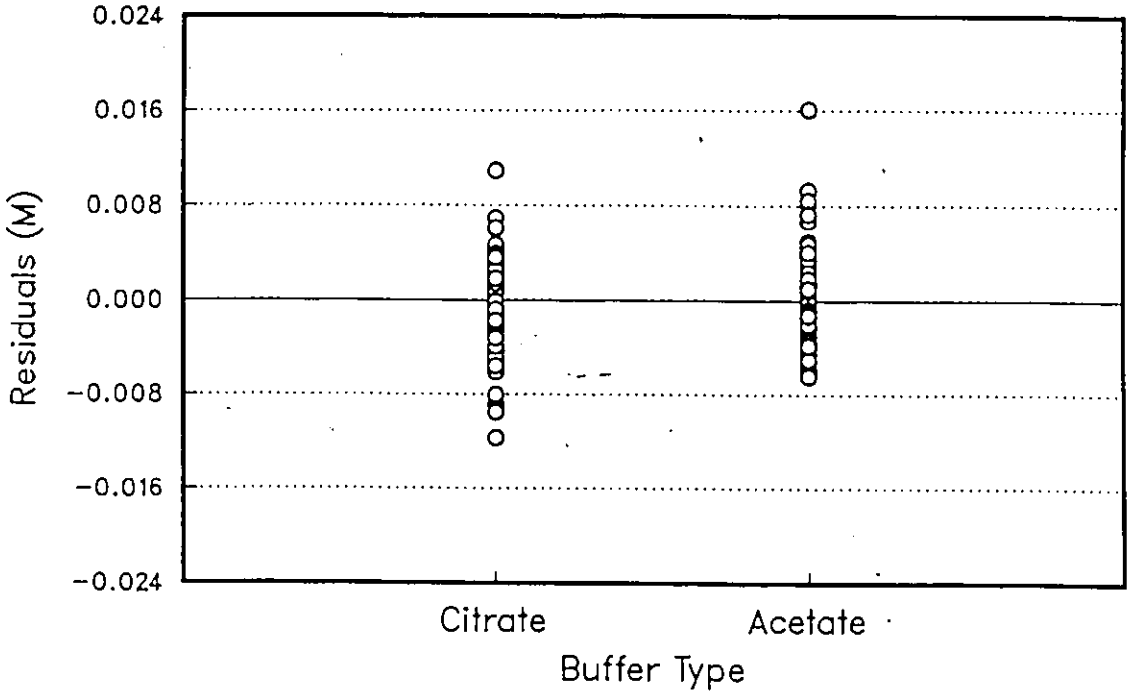
(j) Residuals vs Initial Novozym Vol.

Figure 25: Model 13 Residuals

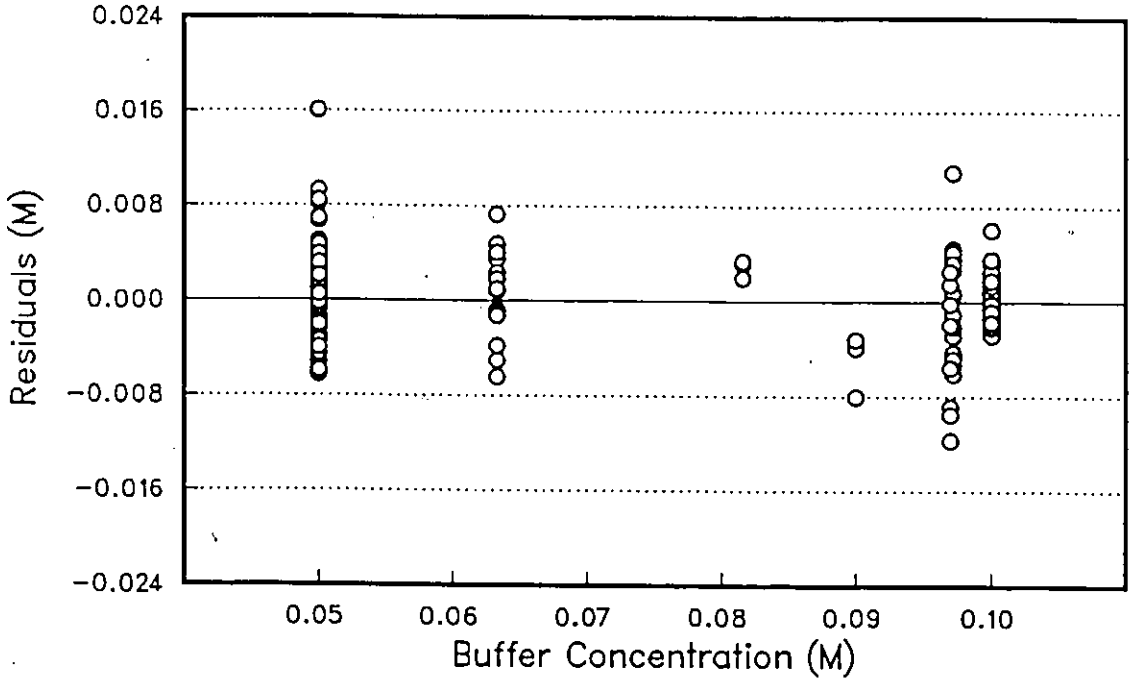


(k) Residuals vs Run Date

Figure 25: Model 13 Residuals

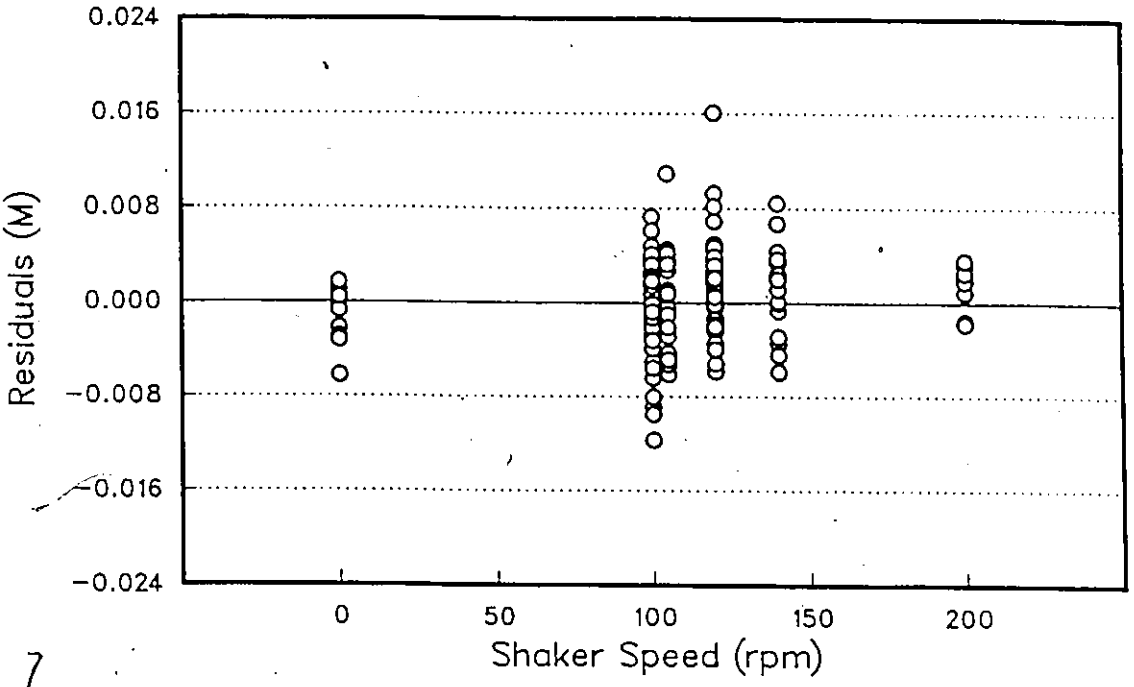


(l) Residuals vs Buffer Type

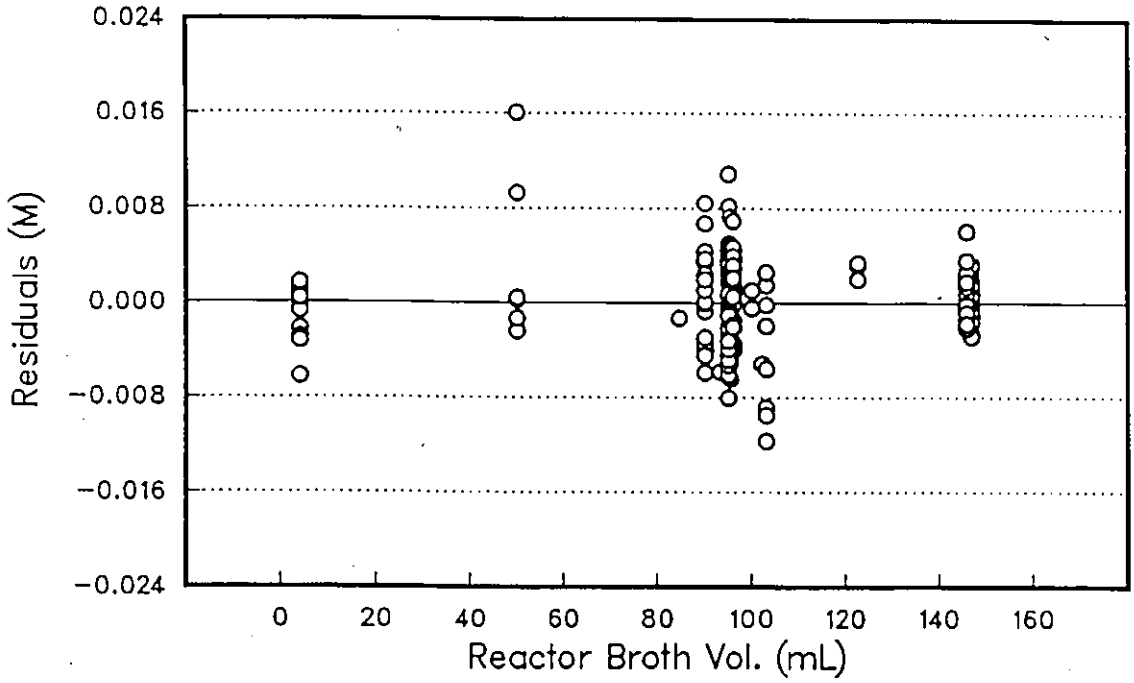


(m) Residuals vs Buffer Concentration

Figure 25: Model 13 Residuals

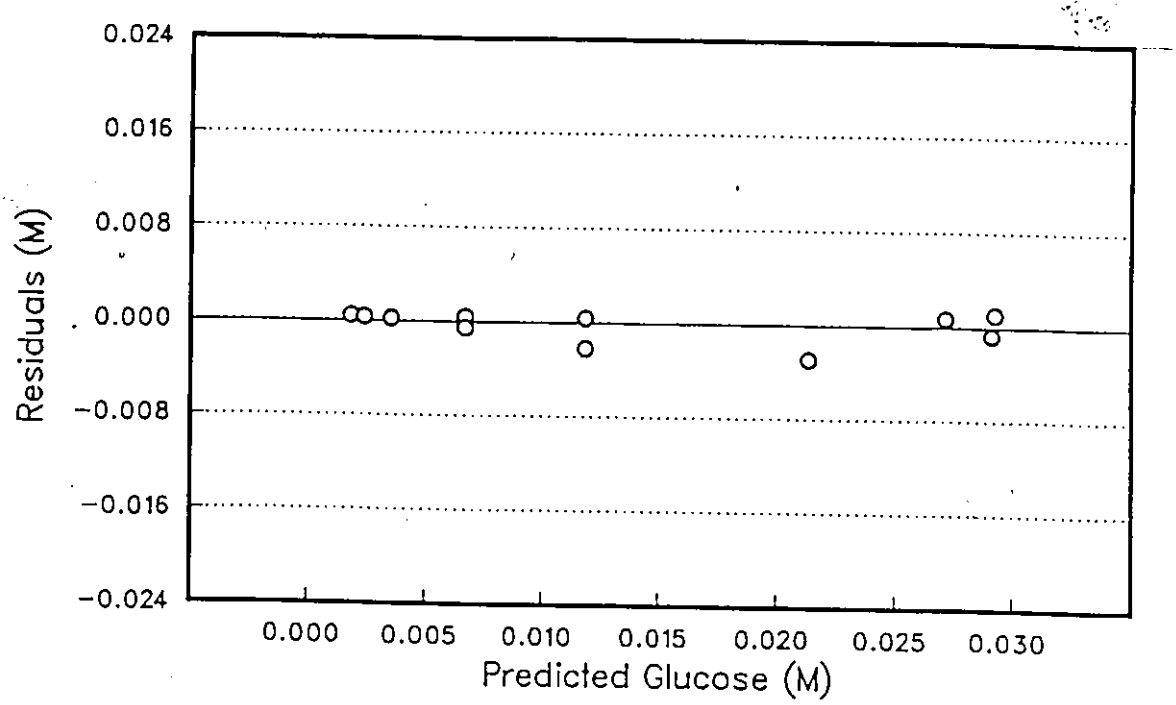


(n) Residuals vs Shaker Speed

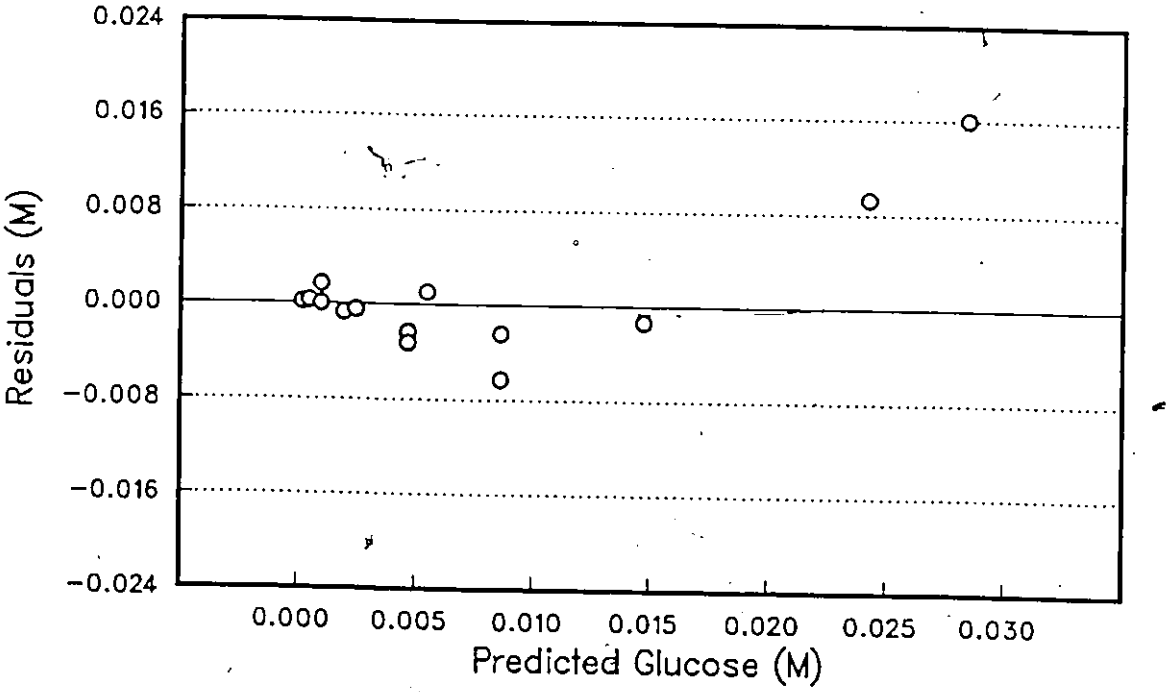


(o) Residuals vs Reactor Broth Vol.

Figure 25: Model 13 Residuals



(a) Residuals vs Predicted (Novozym Runs Only Included)



(b) Residuals vs Predicted (Celluclast Runs Only Included)

Figure 26: Model 13 Residuals for Runs with no Initial CMC

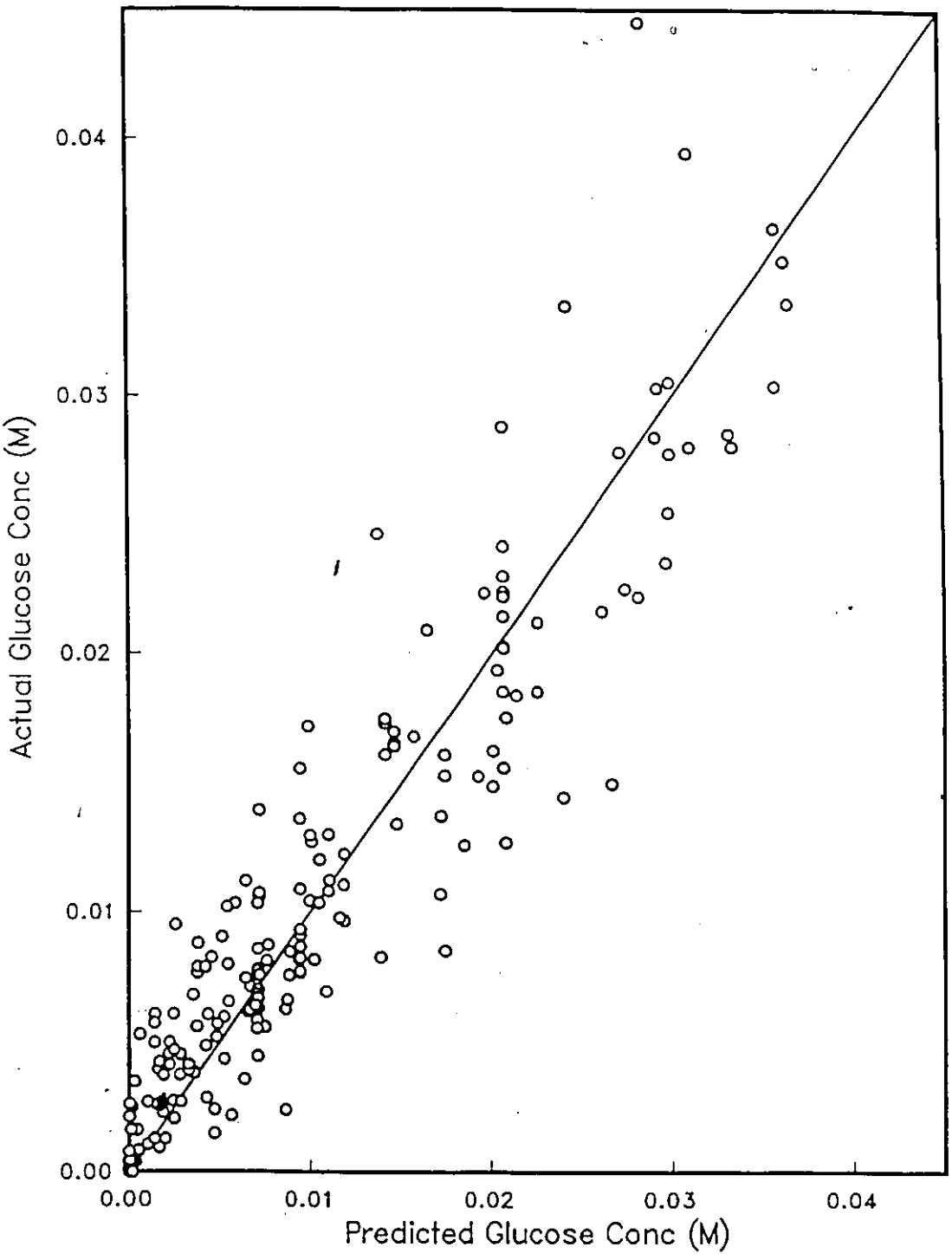


Figure 27: Actual vs Predicted Glucose for Model 13

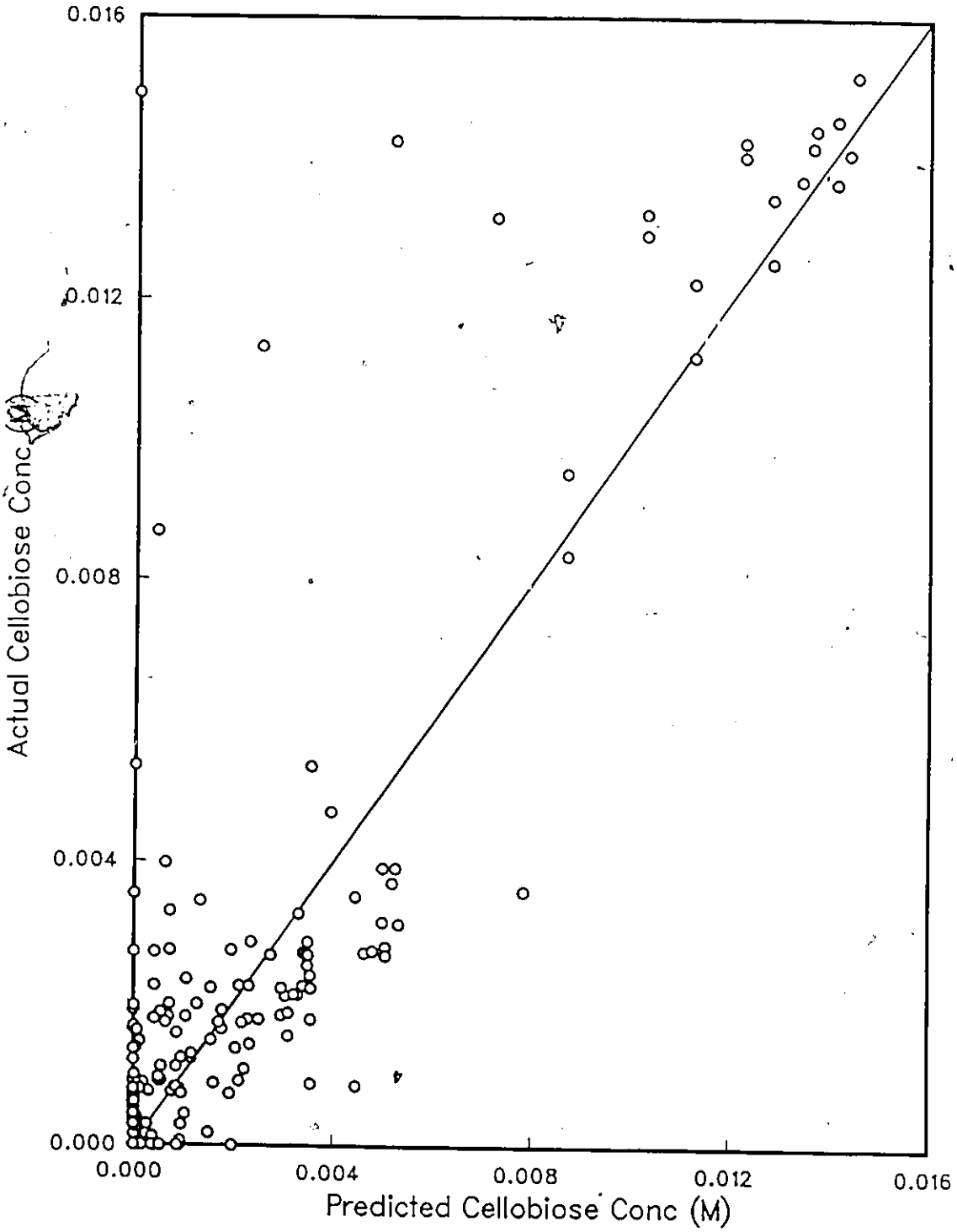


Figure 28: Actual vs Predicted Cellobiose for Model 13

Chapter 6

DISCUSSION OF RESULTS

6.1 Model of the CMC Enzymatic Hydrolysis Reaction

The models were presented in chronological order of fitting. Model 1, the two-enzyme Michaelis-Menten model with no inhibition was gradually improved by making changes in order to

- reduce the sum of squared residuals,
- eliminate trends in the residual plots,
- parsimoniously parameterize the model and
- closely predict the second response variable, G_2.

The models that followed model 1 were developed to better represent some phenomena of the actual hydrolysis reaction. By coming closer to representing the actual mechanism of the enzymatic reaction, the SSR was lowered and trends in the residuals were eliminated.

One assumption reduced the SSR from 0.0037 to 0.0026: endoglucanase does not cleave G_2. Unfortunately, we do not have complete confidence in accepting this as a conclusion. The CBA activity test may be forcing us to make that change to the overall model because the results of the CBA test itself are dependent on the assumption that endoglucanase does not cleave G_2. Recall here that the concentration of E_1 is calculated from the uptake rate of cellobiose in the CBA test. The result is valid only if there is no influence by E_x or E_2 on cellobiose hydrolysis kinetics. Nonetheless, the assumption was kept because of the significantly improved fit.

However, this assumption may have led to the trend in the residuals for the runs where cellobiose was the only substrate. The overall models no. 1–6 had difficulty in predicting G_2 hydrolysis by Celluclast and Novozym solutions. By multiplying fivefold the cellobiase activity of Celluclast, the trend in the residuals was eliminated. This points to the fact that either the CBA test was not sensitive enough or at least two enzymes in the solution may be cleaving G_2 at different rates.

Enzyme technology is not advanced enough at this point to be able to calculate the concentration of a given cellulase enzyme in a mixture. Specifically in our case, use of a substrate specific to cellobiase, if one can be identified, would be very beneficial to the modelling. This substrate may be *p*-nitrophenyl-β-D-glucoside. More research is needed in this area.

Of course, the calculation of the endoglucanase in the solution also left some doubt as to the accuracy of its results. The FPA test is, in reality, specific for the whole enzyme system. A viscometric method, as proposed by Demeester [31] is perhaps the most accurate here. The CMCase activity test based on sugar determination must be rejected. The results of this work have shown that the production rate of sugars in CMC hydrolysis is a function of the whole enzyme system. Through simulations presented later in the discussion, the high non-linearity of the enzyme dilution curve can be shown to be the result of the high rate at which endoglucanase cleaves its substrate bonds. To eliminate the non-linearity of the test, it would be necessary to both highly dilute the enzyme solution to be assayed and to add some cellobiase. The added cellobiase would keep the reducing sugar concentration to a detectable level. The cellobiase added must be in excess of the level at which it—E_1—still influences the rate of production of reducing sugars i.e., E_x must be the limiting enzyme for the CMCase reducing sugar activity test to provide a good estimate of E_x's concentration in the solution.

It has been shown by the modelling exercise that the high concentration of

cellobiose in the reaction mixture may be explained by the presence of an endwise-acting enzyme removing G_2 dimers from the non-reducing end of the CMC. In practice, the assumption that E_2 is responsible for this action and that its hydrolysis rate is infinitely faster than the hydrolysis rates of E_x and E_1 do allow us to predict the concentration of cellobiose in the reaction mixture. However, the kinetic parameters associated with E_2 have not been evaluated.

The concentration of exoglucanase in a solution is more difficult to estimate than the concentrations of endoglucanase and cellobiase. Such a method must be developed in the future to permit the building of an overall hydrolysis model based on a three-enzyme cellulase system, E_x, E_2 and E_1.

There are some limitations to the information extracted from the modelling exercise caused by the variance in the data and the limitations in the experimental run locations i.e., no design. Some parameters could not be evaluated with sufficiently high precision and had to be dropped from the model. Or a parameter would be highly correlated with another, leading us to combine the two. For example, the Michaelis constant of E_x is too small, and the variance in the data too high to evaluate K_{mE_x}. The Michaelis constant of E_1 cannot be evaluated; K_{mE_1} is highly correlated with k_{1E_1}. The two parameters were combined into one. This results in the approximation of first-order kinetics for E_1.

Furthermore, we could not conclusively show that inhibition plays a major role in the enzymatic degradation of CMC. A dissociation constant for competitive or non-competitive inhibition of E_1 by G_1 could not be calculated because of high variance in the data and lack of revealing run locations i.e., high $G_1(0)$.

Inhibition of E_x by all β-1,4 non-substrate bonds has been included in hydrolysis models 6, 11, 12 and 13. The difference with G_1 inhibition is that a new parameter is not added to the model if the dissociation rate of the enzyme-substrate complex is assumed equal to the dissociation rate of the enzyme-inhibitor complex. The

change in this assumed E_x inhibition did not improve or worsen the model. It was kept because some papers have reported methyl cellulose to be an inhibitor. This assumption could be easily verified by adding methyl cellulose of DS = 3 to the reaction mixture.

Because the modelling was presented here in chronological order, it should not come as a surprise that our choice for representing the kinetics of the enzymatic hydrolysis of carboxymethyl cellulose is model 13.

Model 13 in fact is the only one that can reasonably predict glucose and cellobiose. This is important since we are trying to determine the kinetics of the overall enzymatic hydrolysis of cellulose derivatives.

Because of a lack of information in the data, the Michaelis-Menten kinetics which usually describe enzymatic reactions have had to be approximated—we think—by zero-order kinetics for the endoglucanase and first-order kinetics for cellobiase. This is what gives the CMC hydrolysis curve its characteristic shape; The endoglucanase, and the exoglucanase, as has been approximated, cleave β-1,4 bonds very quickly. This action accounts for a rapid increase in the concentration of glucose at the onset of the reaction. It also provides non-reducing end unsubstituted units as a substrate to cellobiase. The cellobiase, following first-order kinetics uses the substrate provided by E_x with a rate of production of glucose slowing down as the reaction wears on and the glucose concentration eventually reaches its maximum.

To account for the significantly high cellobiose concentration in the reactor, the E_2 enzyme is assumed to act at an infinitely faster rate than E_x or E_1 transforming any β_{nre} bonds into G_2 as the reaction proceeds. This provides further substrate for the cellobiase in solution.

We could not find a two-enzyme mechanism to explain the cellobiose concentration. The attempt in model 11 to exclude β_{nre} as a substrate of E_x provided only a marginal improvement in the Actual vs Predicted Cellobiose graph. As one

source of glucose is eliminated, the fitted—to glucose concentration—rate constant of E_1 was higher to maintain the good prediction of G_1. Cellobiose uptake by E_1 was consequently higher nullifying most of the advantage gained by excluding β_{nre} as a substrate of E_x.

6.2 Simulating the CMC Hydrolysis; Substrate Multiplicity

The progress curve of the CMC hydrolysis reaction for the conditions of runs H09506–H09510 can be followed for the first 40 hours on Figure 29(a)–(f). The data for the first 50 hours of hydrolysis can be found in Tables 16 and 17.

Model 13, our best model, was used to simulate the hydrolysis and obtain concentrations of all the species in the reactor, as we have defined them in the nomenclature.

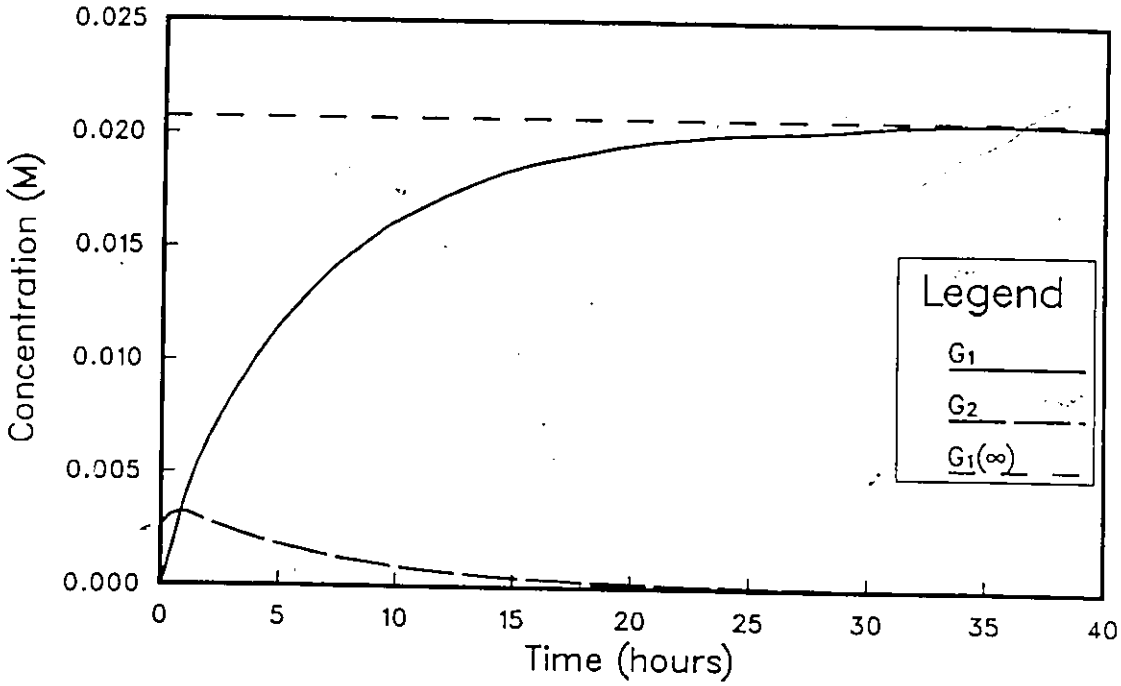
The simulation parameters are $\Delta t = 2$ seconds and TOL $= 10^{-14}$. The kinetic constants from model 13, you will recall, are $k_{1E_x} = 3.66 \cdot 10^{-7}$ moles/I.U./s and $k_{1E_1} = 6.59 \cdot 10^{-7}$ L/I.U./s. The CMC has a DS of 0.792 and a DP of 400.

Table 16. *CMC Time Course Hydrolysis, Runs H09506–H09510: Part 1*

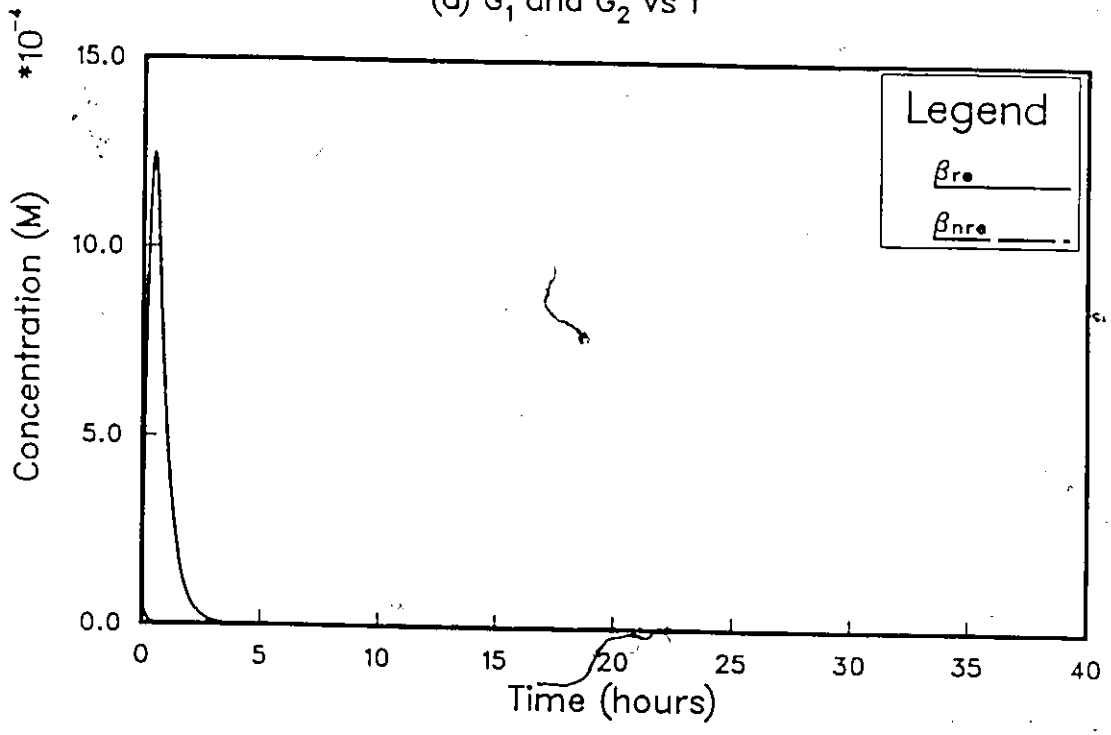
Time hours	β_u mM/L	G_2 mM/L	G_1 mM/L	β_{ss} mM/L	β mM/L	DP
0.0	20.6	0	0	62.4	155	400
0.5	4.20	3.13	1.82	62.4	140	11
1.0	1.19	3.23	3.78	62.4	136	8.2
1.5	0.384	3.03	5.24	62.4	134	7.4
2.0	0.129	2.82	6.46	62.4	133	7.0
2.5	0.044	2.63	7.46	62.4	132	6.7
3.0	0.0149	2.44	8.40	62.4	131	6.5
3.5	0.00506	2.27	9.26	62.4	130	6.3
4.0	0.00170	2.11	10.1	62.4	130	6.1
4.5	0.000571	1.96	10.8	62.4	129	6.0
5.0	0.000191	1.82	11.5	62.4	129	5.9
5.5	0.000063	1.69	12.1	62.4	128	5.8
6.0	0.000021	1.58	12.7	62.4	128	5.7
6.5	0.000007	1.46	13.3	62.4	127	5.6
7.0	0.000002	1.36	13.8	62.4	127	5.5
7.5	0.000001	1.27	14.3	62.4	126	5.4
8.0	0	1.18	14.7	62.4	126	5.4
8.5	0	1.09	15.1	62.4	126	5.3
9.0	0	1.02	15.5	62.4	125	5.2
9.5	0	0.945	15.9	62.4	125	5.2
10.0	0	0.878	16.2	62.4	125	5.1
12.5	0	0.610	17.5	62.4	124	5.0
15.0	0	0.423	18.5	62.4	123	4.8
17.5	0	0.294	19.1	62.4	123	4.8
20.0	0	0.204	19.6	62.4	122	4.7
22.5	0	0.142	19.9	62.4	122	4.7
25.0	0	0.0982	20.1	62.4	122	4.7
27.5	0	0.0682	20.2	62.4	122	4.6
30.0	0	0.0473	20.4	62.4	122	4.6
32.5	0	0.0328	20.4	62.4	122	4.6
35.0	0	0.0228	20.5	62.4	121	4.6
37.5	0	0.0158	20.5	62.4	121	4.6
40.0	0	0.0110	20.5	62.4	121	4.6
42.5	0	0.00762	20.5	62.4	121	4.6
45.0	0	0.00529	20.6	62.4	121	4.6
47.5	0	0.00367	20.6	62.4	121	4.6
50.0	0	0.00255	20.6	62.4	121	4.6
∞	0	0	20.7	62.4	121	4.6

Table 17. *CMC Time Course Hydrolysis, Runs H09506–H09510: Part 2*

Time hours	β_{rc} mM/L	β_{rcs} mM/L	β_{nrc} mM/L	β_{nrcs} mM/L	β_{sr} mM/L	β_{snr} mM/L
0.0	0.0514	0.0896	0.0514	0.0896	35.9	35.9
0.5	1.24	8.18	0.0000163	7.07	27.7	26.1
1.0	0.543	11.2	0.0000002	8.61	24.7	24.1
1.5	0.201	12.3	0	8.69	23.6	23.1
2.0	0.0708	12.7	0	8.31	23.2	22.8
2.5	0.0244	12.8	0	7.81	23.1	22.8
3.0	0.00835	12.9	0	7.28	23.0	22.8
3.5	0.00283	12.9	0	6.78	23.0	22.8
4.0	0.000955	12.9	0	6.31	23.0	22.8
4.5	0.000320	12.9	0	5.86	23.0	22.8
5.0	0.000107	12.9	0	5.45	23.0	22.8
5.5	0.0000356	12.9	0	5.07	23.0	22.8
6.0	0.0000127	12.9	0	4.71	23.0	22.8
6.5	0.0000039	12.9	0	4.38	23.0	22.8
7.0	0.0000013	12.9	0	4.07	23.0	22.8
7.5	0.0000004	12.9	0	3.78	23.0	22.8
8.0	0.0000001	12.9	0	3.52	23.0	22.8
8.5	0	12.9	0	3.27	23.0	22.8
9.0	0	12.9	0	3.04	23.0	22.8
9.5	0	12.9	0	2.82	23.0	22.8
10.0	0	12.9	0	2.63	23.0	22.8
12.5	0	12.9	0	1.82	23.0	22.8
15.0	0	12.9	0	1.27	23.0	22.8
17.5	0	12.9	0	0.878	23.0	22.8
20.0	0	12.9	0	0.609	23.0	22.8
22.5	0	12.9	0	0.423	23.0	22.8
25.0	0	12.9	0	0.294	23.0	22.8
27.5	0	12.9	0	0.204	23.0	22.8
30.0	0	12.9	0	0.141	23.0	22.8
32.5	0	12.9	0	0.0982	23.0	22.8
35.0	0	12.9	0	0.0682	23.0	22.8
37.5	0	12.9	0	0.0473	23.0	22.8
40.0	0	12.9	0	0.0328	23.0	22.8
42.5	0	12.9	0	0.0228	23.0	22.8
45.0	0	12.9	0	0.0158	23.0	22.8
47.5	0	12.9	0	0.0110	23.0	22.8
50.0	0	12.9	0	0.00762	23.0	22.8
∞	0	13.1	0	0	22.8	22.8

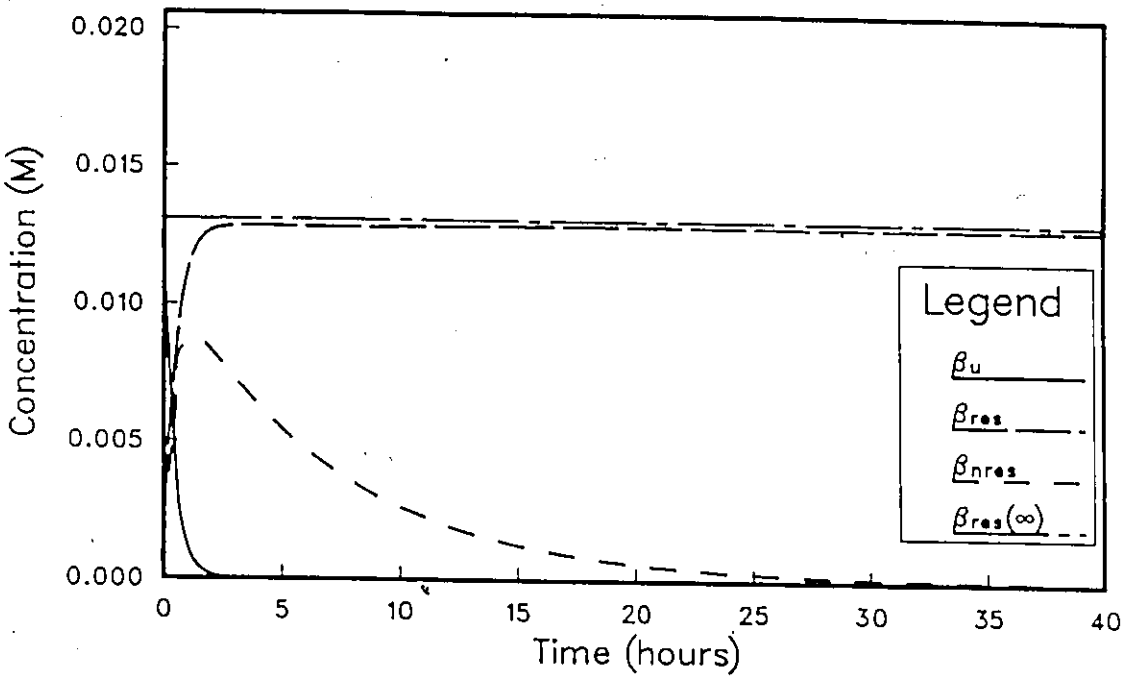


(a) G_1 and G_2 vs t

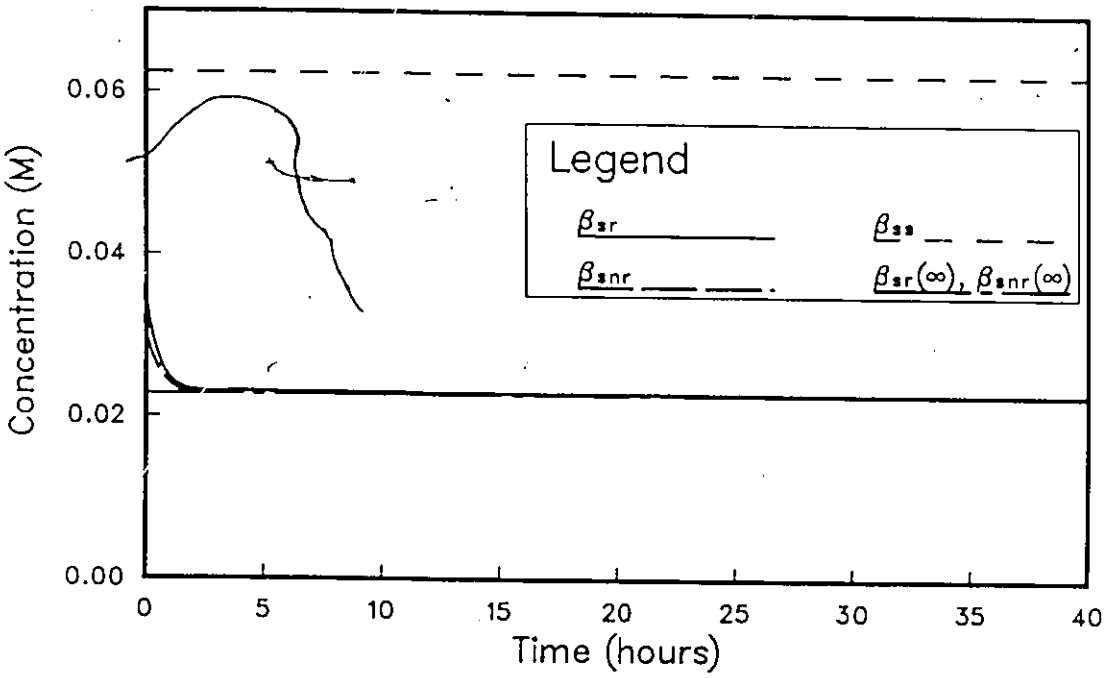


(b) β_{re} and β_{nre} vs t

Figure 29: CMC Time Course Hydrolysis
at Conditions of Run H09510

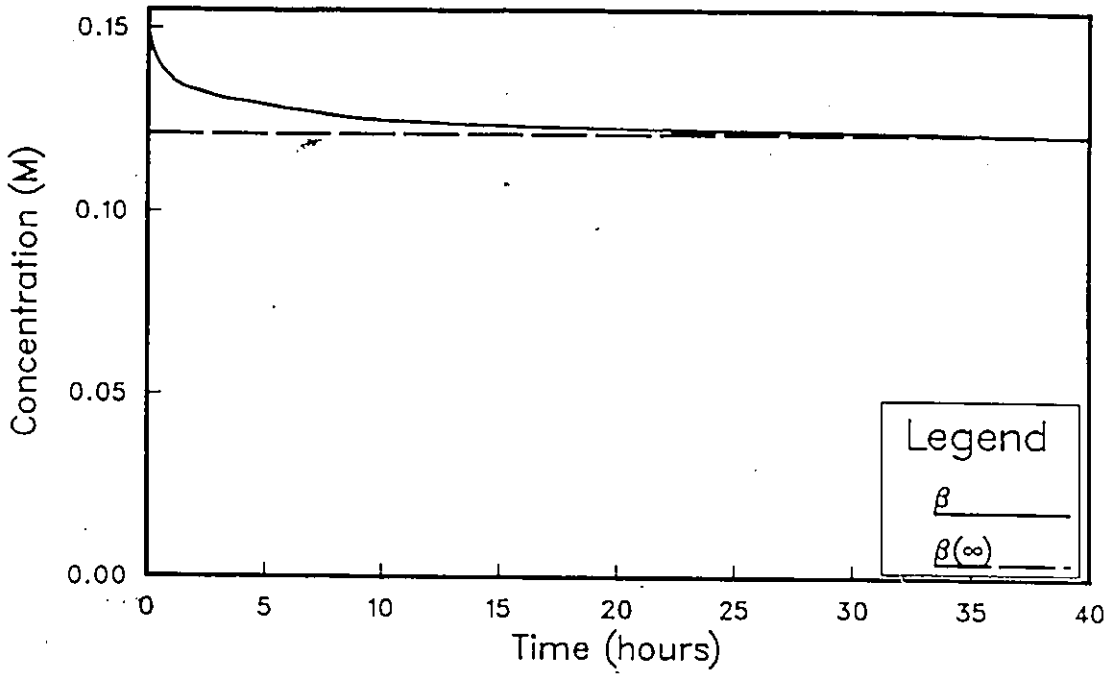


(c) β_u, β_{res} and β_{nres} vs t

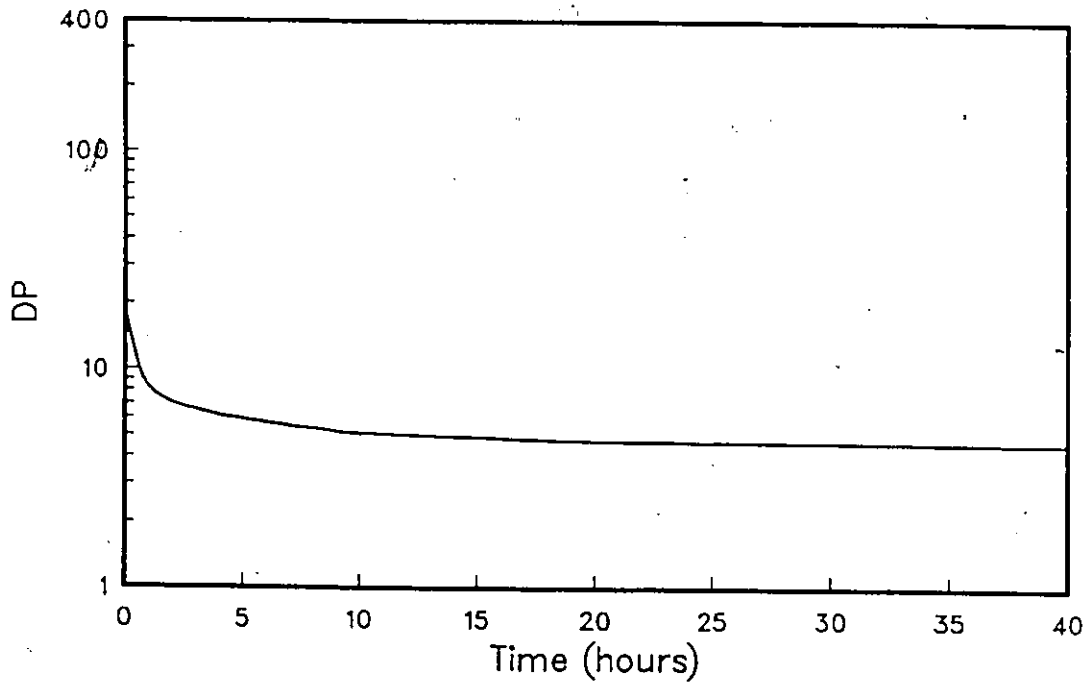


(d) β_{sr}, β_{snr} and β_{ss} vs t

Figure 29: CMC Time Course Hydrolysis at Conditions of Run H09510



(e) β vs t



(f) Number–Average DP vs t

Figure 29: CMC Time Course Hydrolysis at Conditions of Run H09510

The initial rate of production of glucose is typical of the CMC progress curve. E_x is mostly responsible for that high initial rate. As the reaction proceeds, substrates available to E_x are depleted quickly and the slower acting E_1 gives the glucose curve its characteristic levelling off. It cleaves the remaining non-reducing end bonds if the end unit is unsubstituted. A close inspection of the data reveals that the rate of production of glucose is increasing for the first hour of hydrolysis. This trend has not been noticed by anyone investigating CMC hydrolysis. It is not evident even on our plot. This trend cannot be seen by someone measuring the reducing power of a mixture. However, it is predicted by the model. It is logical that progress curve of glucose should be S-shaped. The depolymerization reaction initially should be releasing very little sugar. But as the CMC fragments become smaller, the rate at which glucose is produced increases. More substrate is available for the endwise-acting enzymes. It can be shown that the ratio of the three types of enzymes in CMC hydrolysis is crucial to the shape of the glucose progress curve. A reaction mixture with a moderate concentration of E_1 and with little E_x will yield a distinct S-shape.

The cellobiose progress curve shows a maximum concentration approximately 1 hour into the reaction. The concentration of cellobiose is accounted for by the assumption that E_2 cleaves off unsubstituted dimers from the non-reducing end of the CMC at a (nearly) infinite rate. In reality, the E_x cellobiose production rate depends on its concentration, and on the concentration of β_{nrc}. However, there is too much variance in our data to be able to estimate the kinetic parameters of E_2. Our measured cellobiose concentration shows that there is a component of the cellulase enzyme system that is producing more cellobiose than would be expected if the hydrolysis reaction were random. The component, we assume, is E_2.

The exoglucanase action keeps the concentration of β_{nrc} bonds virtually nil during the reaction. Since we have assumed that E_2 acts (almost) infinitely quickly,

a β_{nre} bond in solution is immediately transformed into G_2.

β_u and β_{re}, two more substrates of E_x are exhausted rapidly by the endoglucanase.

E_x acts rapidly at the onset of the reaction to produce chain ends. One of those is β_{res}. Since E_1 does not have any effect on β_{res}, the concentration of the latter reaches its steady-state level early in the reaction. On the other hand, β_{nres} is similar to β_{res} until its production by E_x stops and its depletion by E_1 accelerates. The maximum in β_{nres} occurs after approximately 1.5 hours at the simulated run conditions.

The concentrations of β_{sr} and β_{snr} falls rapidly to reach steady state. The rate of transformation of β_{sr} depends strictly on the rate at which adjacent bonds, β_u and β_{re} are cleaved. The endoglucanase does most of the work initially in the reaction. By the same token, the rate of transformation of β_{nre} depends on the rate at which β_u and β_{nre} are cleaved. But the direct action of E_2 on β_{snr}, if β_{snr} and β_{nre} are contiguous, accelerates the transformation rate of β_{snr} relative to β_{sr}.

Bonds that link two substituted units are unaffected by the enzyme hydrolysis. The fraction of bonds which undergo cleavage is 22% at the DS of 0.792. β_{ss} is a major contributor to that statistic. By comparison, the number average DP plunges to near its value at reaction completion, 4.6. The model reflects the observation of a rapid loss in viscosity when soluble cellulose derivatives are exposed to the cellulases.

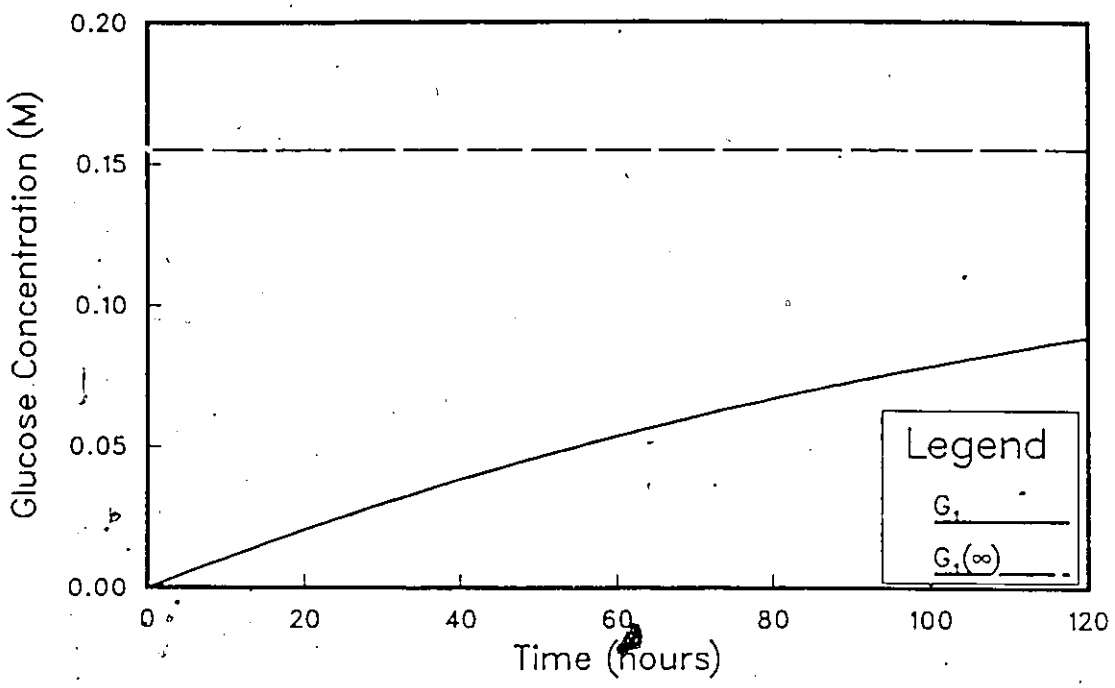
The steady-state DP—or average molecule length at reaction completion—is 4.6 at a DS of 0.792. The steady-state bond concentrations, like the DP, are a strong function of the DS of the initial CMC. This has been shown when we showed that Reese's proposed mechanism was the most likely to explain which bonds are cleaved by cellulases.

6.3 Enzyme Deactivation

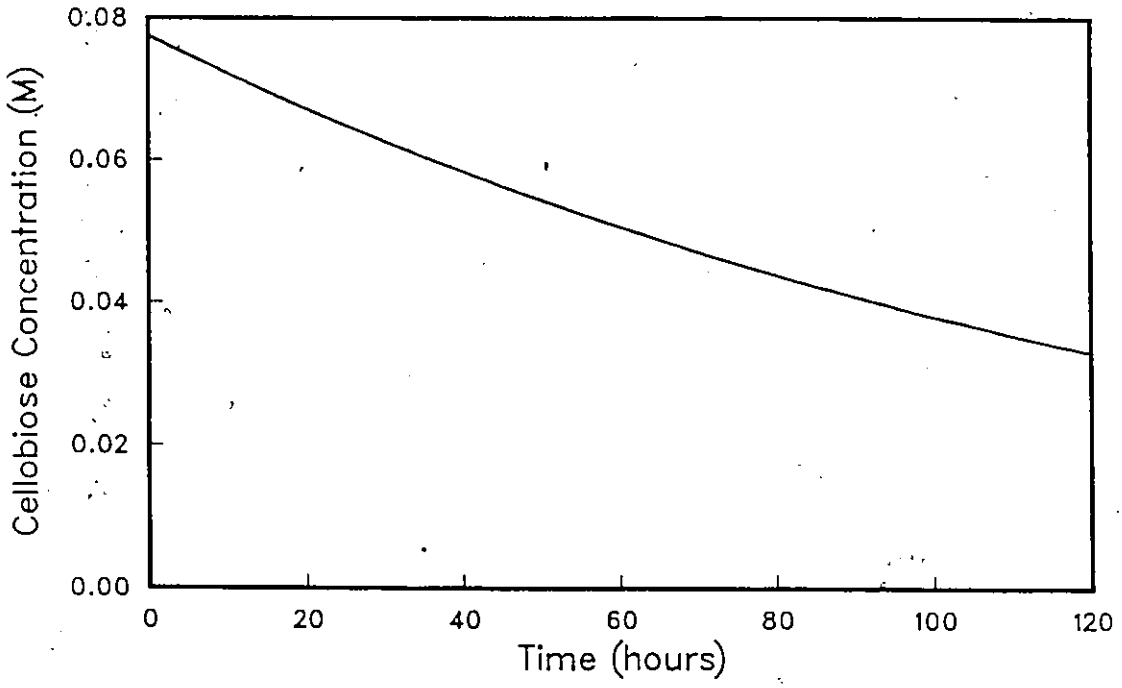
We are examining the kinetics of the cellulase enzyme system where the enzymes are free in solution. Enzyme deactivation in free solution is a potential concern according to our literature survey, especially when the reaction time is several days. However, it has not been taken into account in model-building. In fact no laboratory tests were performed to track deactivation of the enzyme at a pH of 4.8 and temperature 50°Celsius. All users of the model should be aware of possible pitfalls associated with disregarding deactivation, should it be important in this case. However no indication of any deactivation was revealed in the residual plots. Of little consolation is the fact that only two out of fifteen hydrolysis models reviewed by Gilbert and Tsao [3] have included a deactivation parameter.

6.4 Extrapolation to Soluble Cellulose

One of the significant contributions of this work to cellulose hydrolysis is the possibility of using our CMC hydrolysis model to extrapolate to, and predict the hydrolysis rate of soluble cellulose. Now, cellulose is not soluble in practice, however it may be useful to separate and identify the diffusional resistance of the solid cellulose to enzyme attack [39]. The model discrimination procedure has led to model 13 being chosen as best representing the CMC hydrolysis reaction at all values of DS. By extrapolating our model to DS = 0, we can calculate the rate of hydrolysis of soluble cellulose. One assumption will be deleted from the model 13 development; E_2 will not be included in the cellulase system. In CMC hydrolysis, the assumption of (almost) infinite cleavage rate by E_2 influences the cellobiose concentration in an acceptable, yet limited, way. But because cellulose has no substituent groups, an infinitely fast-acting E_2 would mean instant transformation of the cellulose to cellobiose. The soluble cellulose hydrolysis is simulated by assuming the presence of E_x and E_1 only.

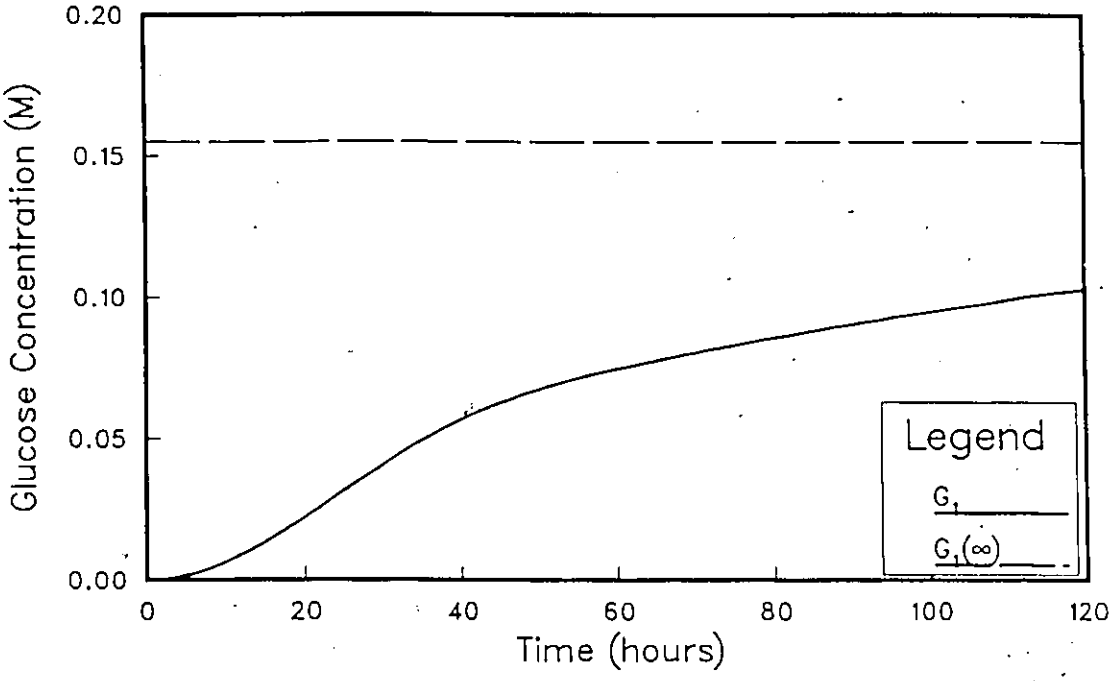


(a) Glucose vs Time

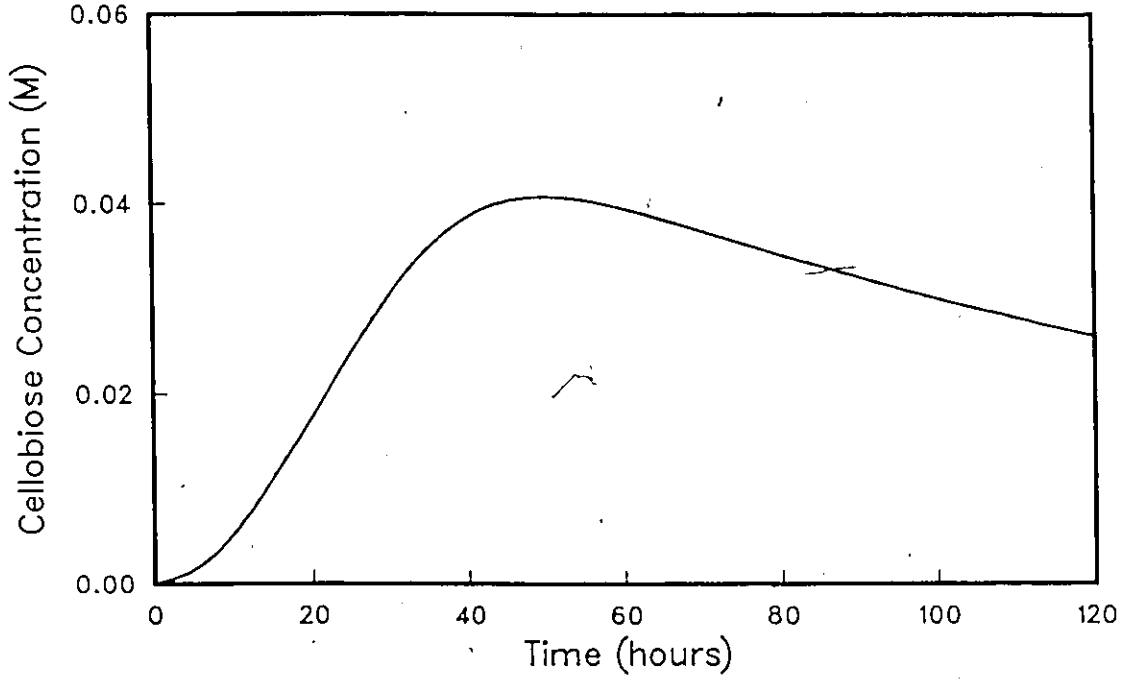


(b) Cellobiose vs Time

Figure 30: Cellobiose Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

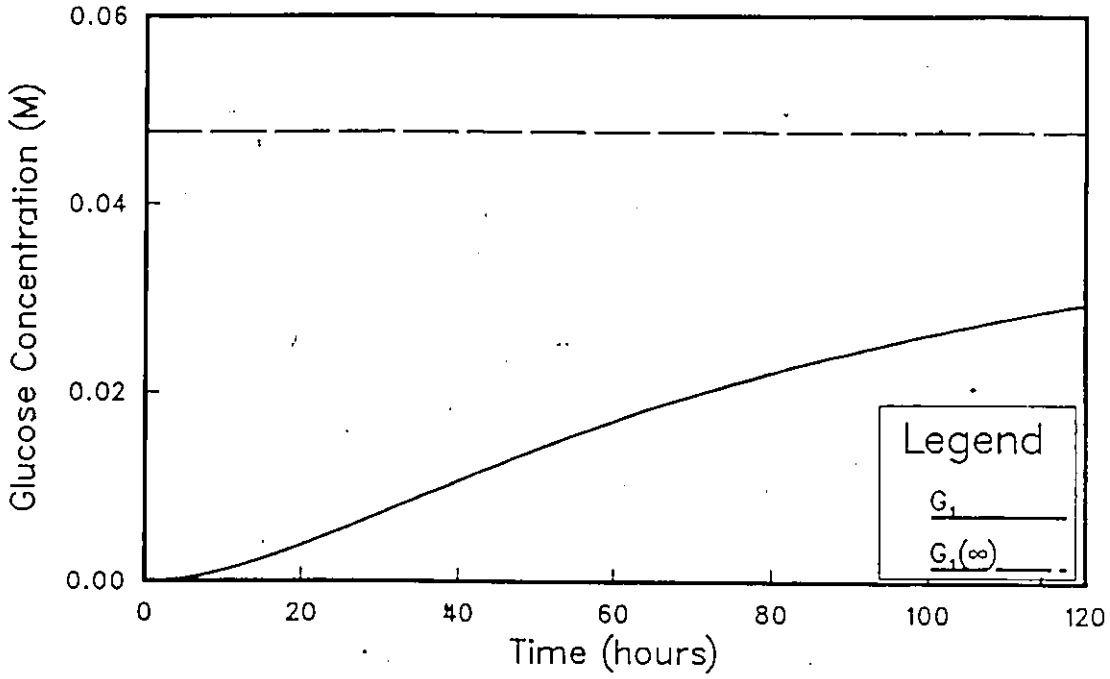


(a) Glucose vs Time

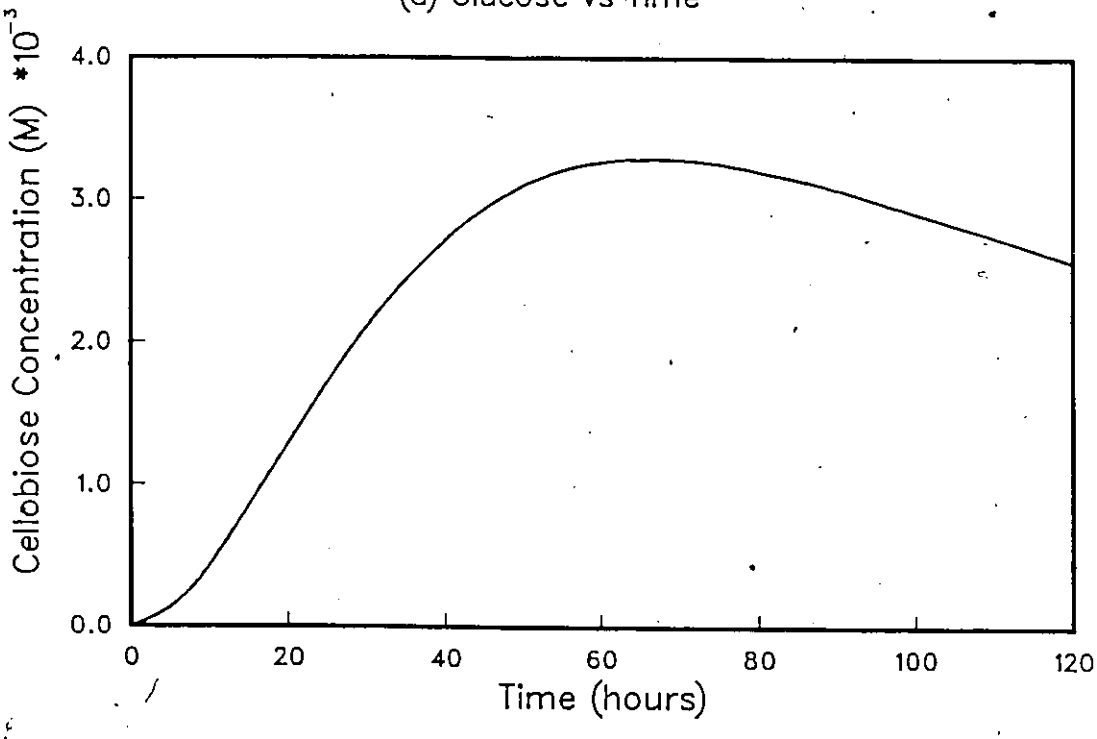


(b) Cellobiose vs Time

Figure 31: Soluble Cellulose Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

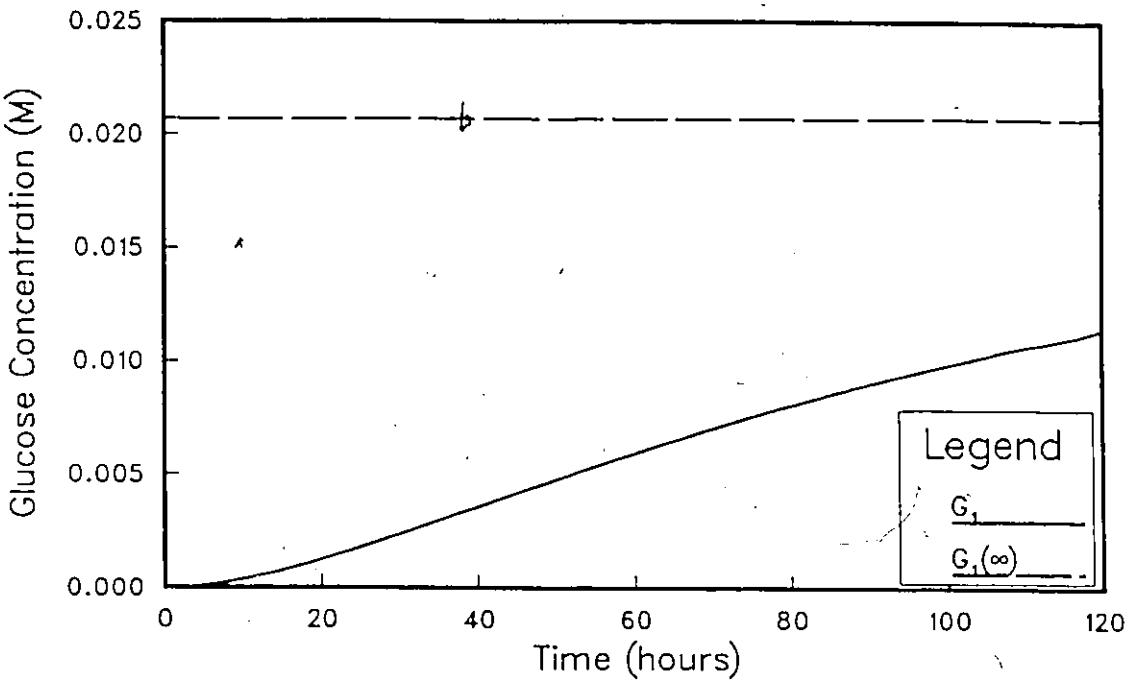


(a) Glucose vs Time

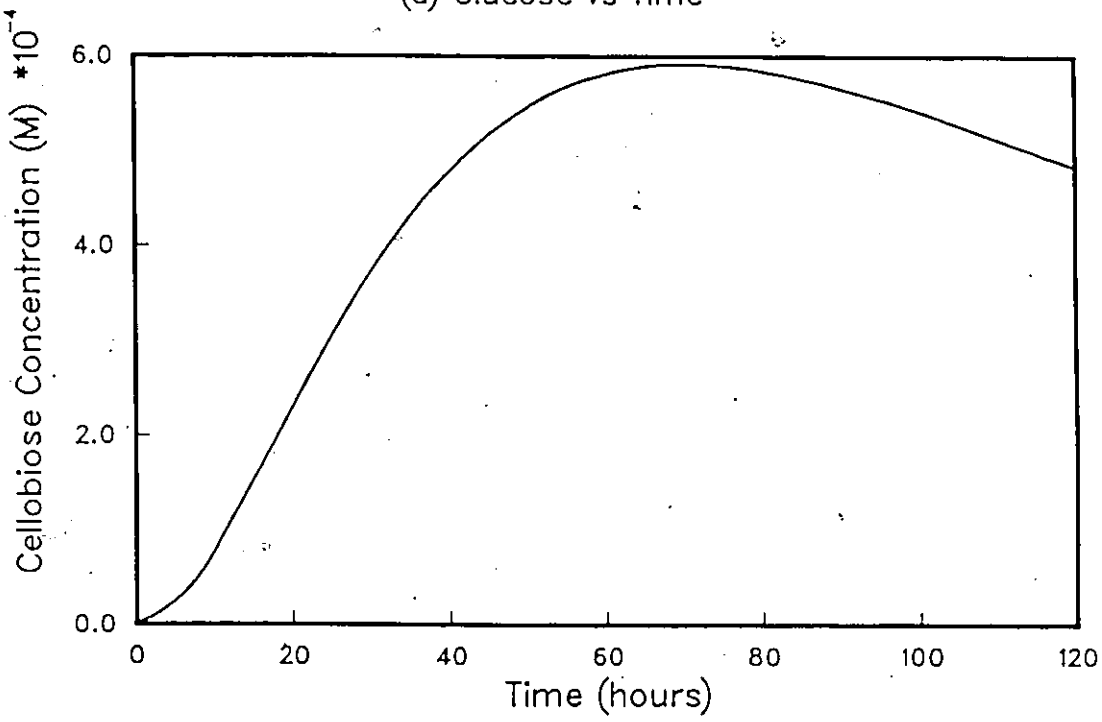


(b) Cellobiose vs Time

Figure 32: CMC DS 0.511 Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

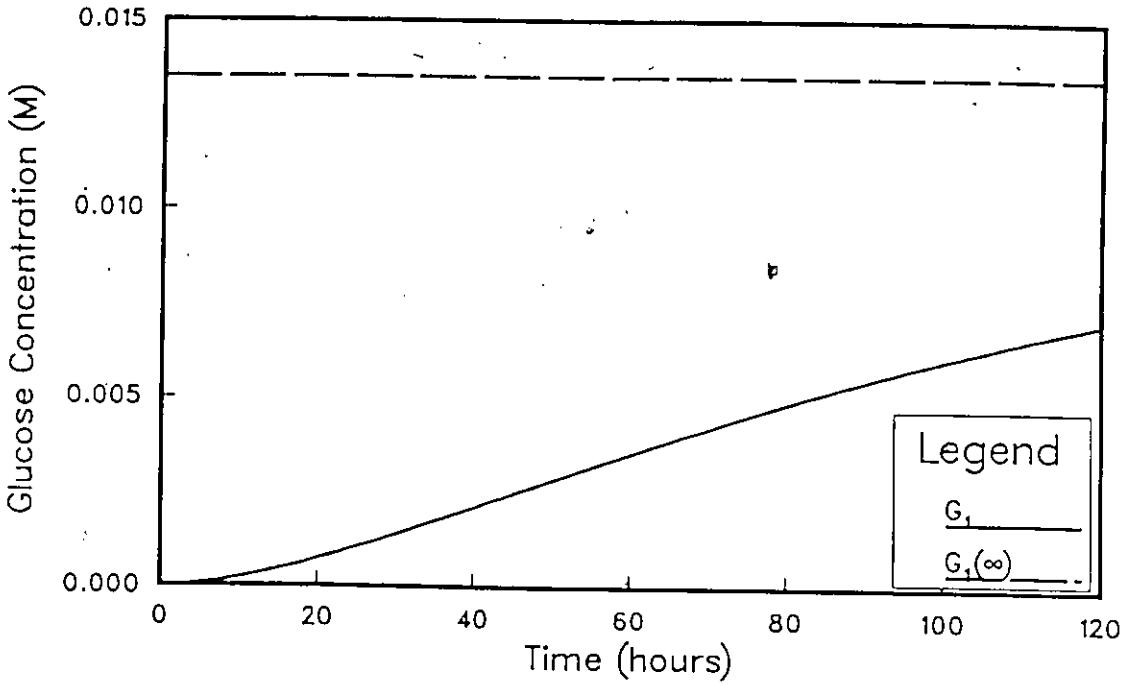


(a) Glucose vs Time

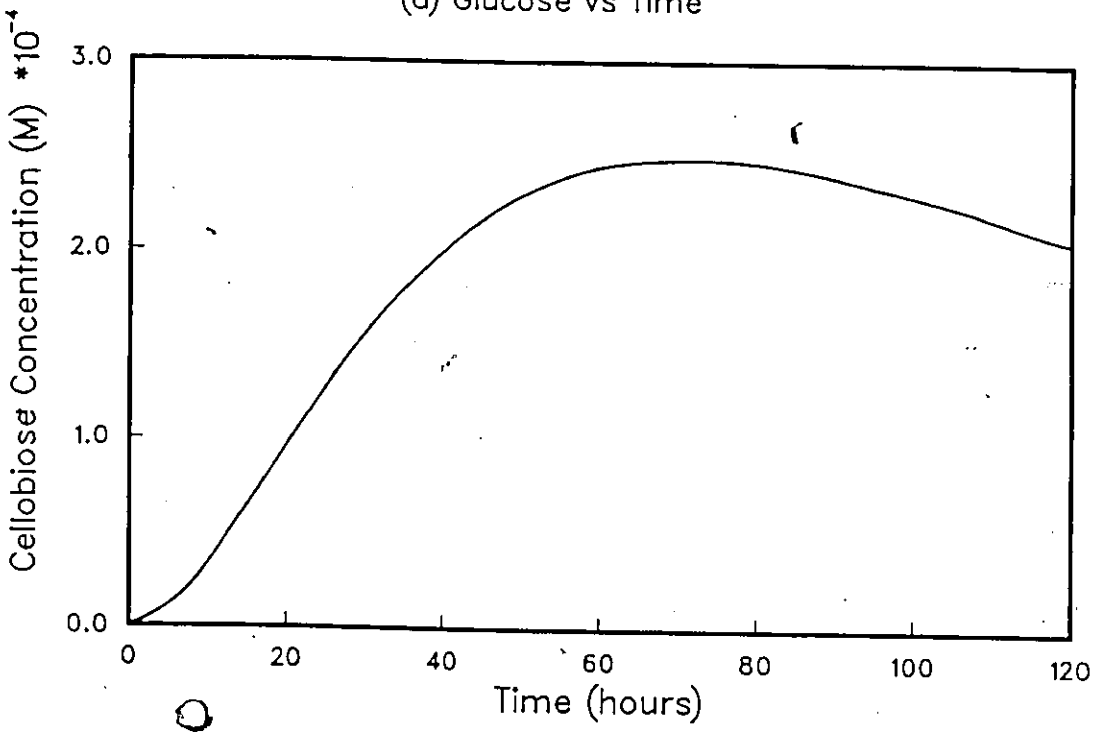


(b) Cellobiose vs Time

Figure 33: CMC DS 0.792 Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L



(a) Glucose vs Time



(b) Cellobiose vs Time

Figure 34: CMC DS 0.914 Hydrolysis Simulation, Model 13 Kinetics Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

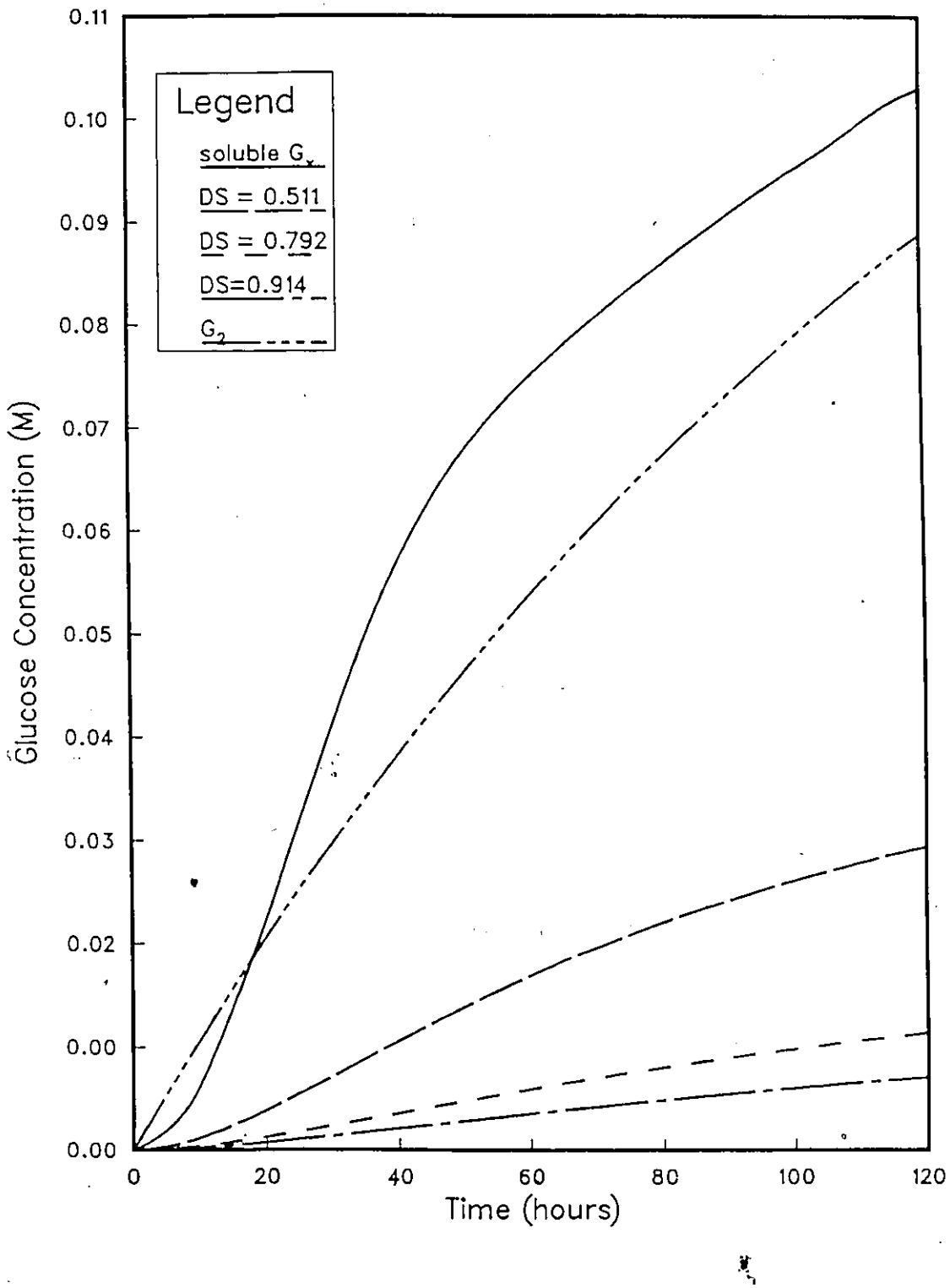


Figure 35: Enzymatic Hydrolysis Simulation, Glucose Response, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

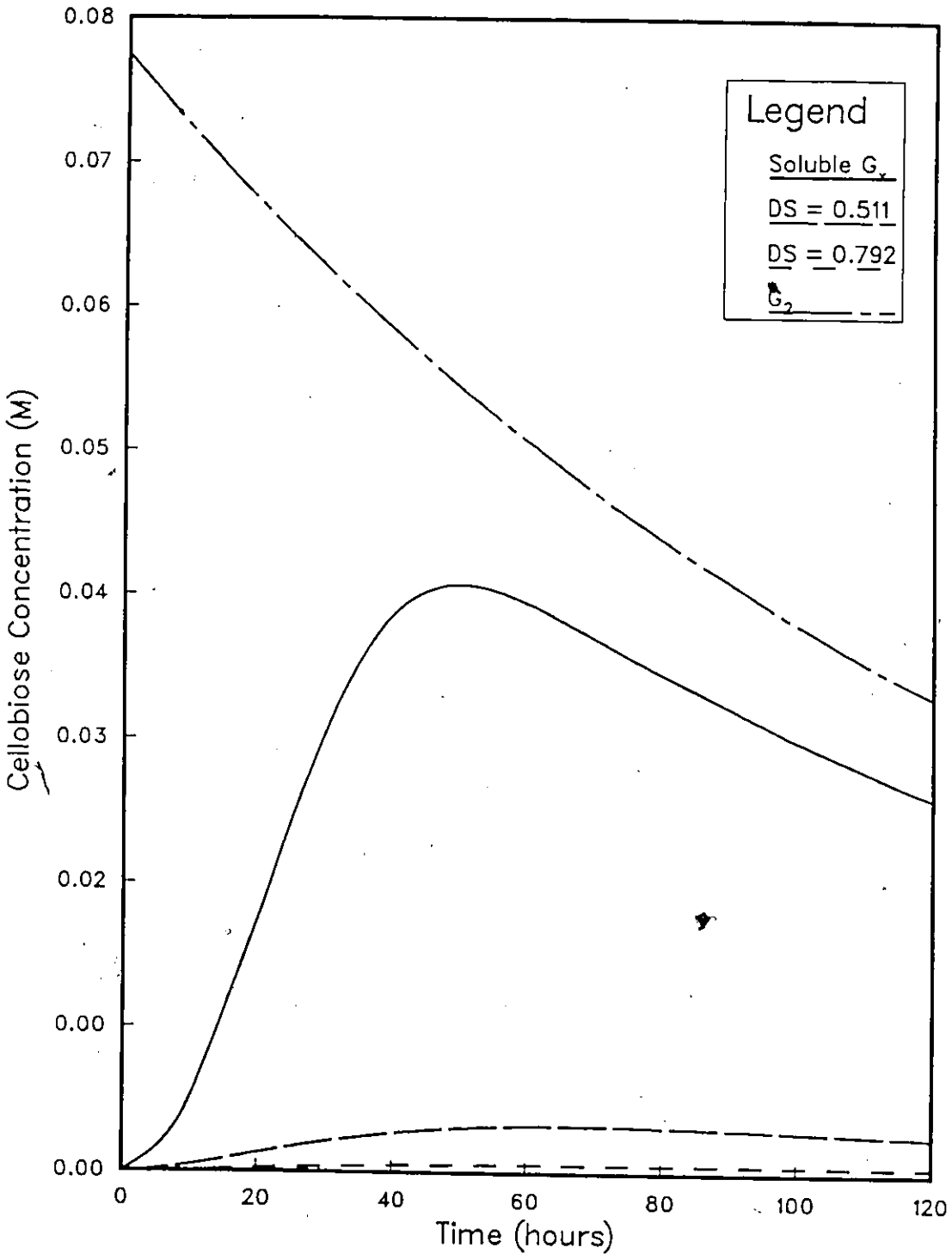
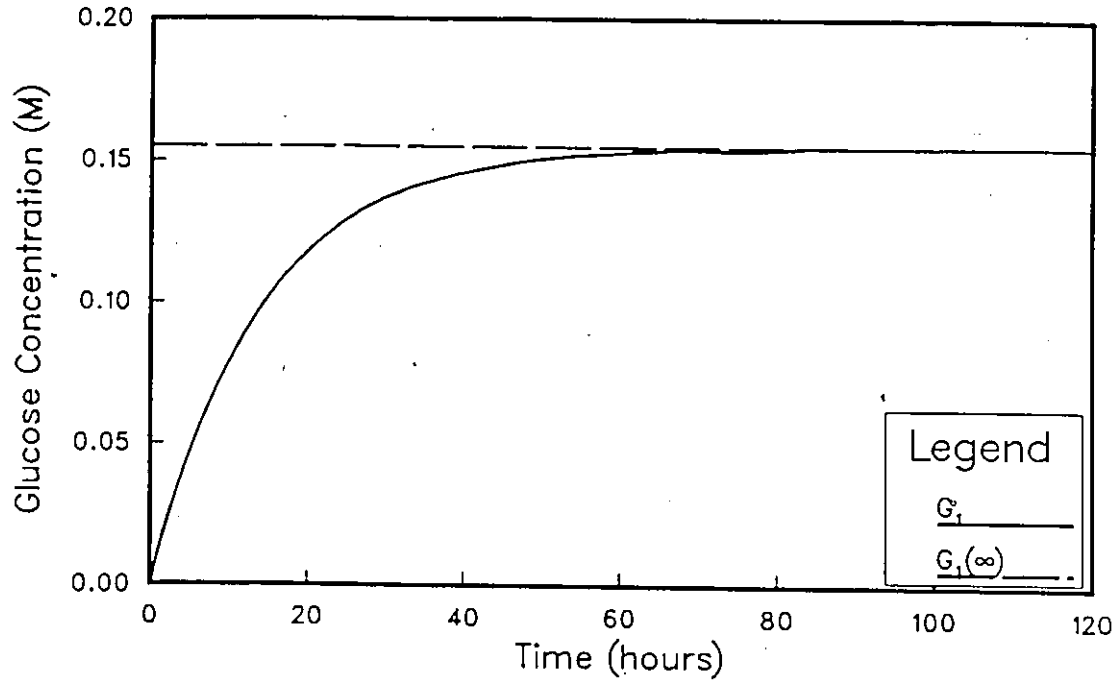


Figure 36: Enzymatic Hydrolysis Simulation, Cellobiose Response, Model 13 Kinetics
Units=0.155 M, E_1=3.0 I.U./L, E_2=0 I.U./L, E_x=2.0 I.U./L

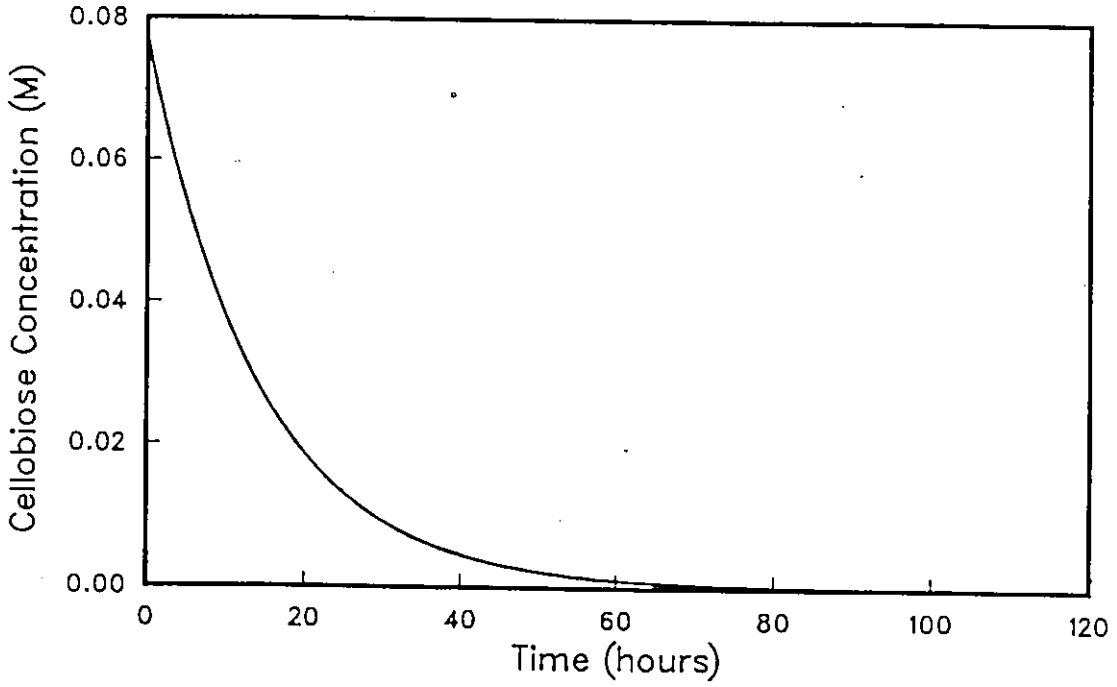
The large number of variables that influence the enzymatic hydrolysis of cellulose and its derivatives prohibits showing the effect of each one here. E_1, E_x, $G_1(0)$, $G_2(0)$, CMC_0, DS, DP all have an effect on G_1, G_2, and the βs. In figures 30–34 are plots of G_1 and G_2 as a function of time for polymers of varying degrees of substitution. Figure 35 is a common graph illustrating the effect of DS on glucose response whereas Figure 36 is a common graph illustrating the effect of DS on cellobiose response. Note that the initial amount of β_{CMC} is constant for all simulations. The initial concentration of cellobiose in Figure 30 has been calculated to yield $G_1(\infty)$ equal to that for the soluble cellulose polymer. Also note that the initial DP of the polymers is set to 1000 in all cases.

The glucose progress curve for soluble cellulose is S-shaped and the cellobiose progress curve exhibits a maximum. Those results are analogous to experimental hydrolyses carried out on crystalline cellulose. Note also the S-shaped glucose curves for the CMCs. They are not as pronounced. As was mentioned earlier, some combinations of enzyme concentrations will yield an S-shaped glucose response curve. We repeat the warning that these hydrolysis curves are simulated and they apply only under the conditions described on the graphs.

In figures 37–43—although not obvious on the latter—are plots of G_1 and G_2 as a function of time for a larger amount and a different ratio of each of the two enzymes. In figures 35 and 42, at the onset of hydrolysis, the glucose concentration is greater for cellulose hydrolysis than for cellobiose hydrolysis. This is a direct result of assuming that E_x cannot cleave G_2 but can cleave β_{re} and β_{nre}. An enzyme cleaving cellulose bonds *randomly* would yield equimolal amounts of glucose and cellobiose. The initial cleavage rate by E_x is so high that much of the monomer and dimer shows up in the solution in the first hours of the reaction. After the polymer has been solubilized to G_1 and G_2, only E_1 becomes important in the kinetic behavior of the hydrolysis reaction. By that time E_x has cleaved many β_{re}

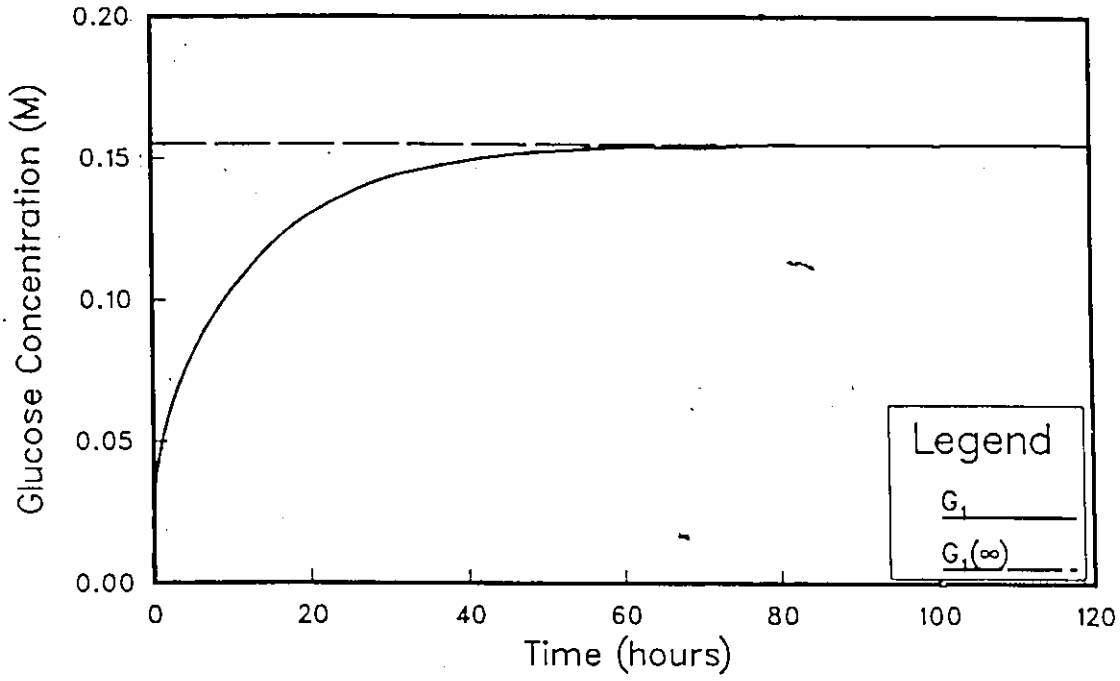


(a) Glucose vs Time

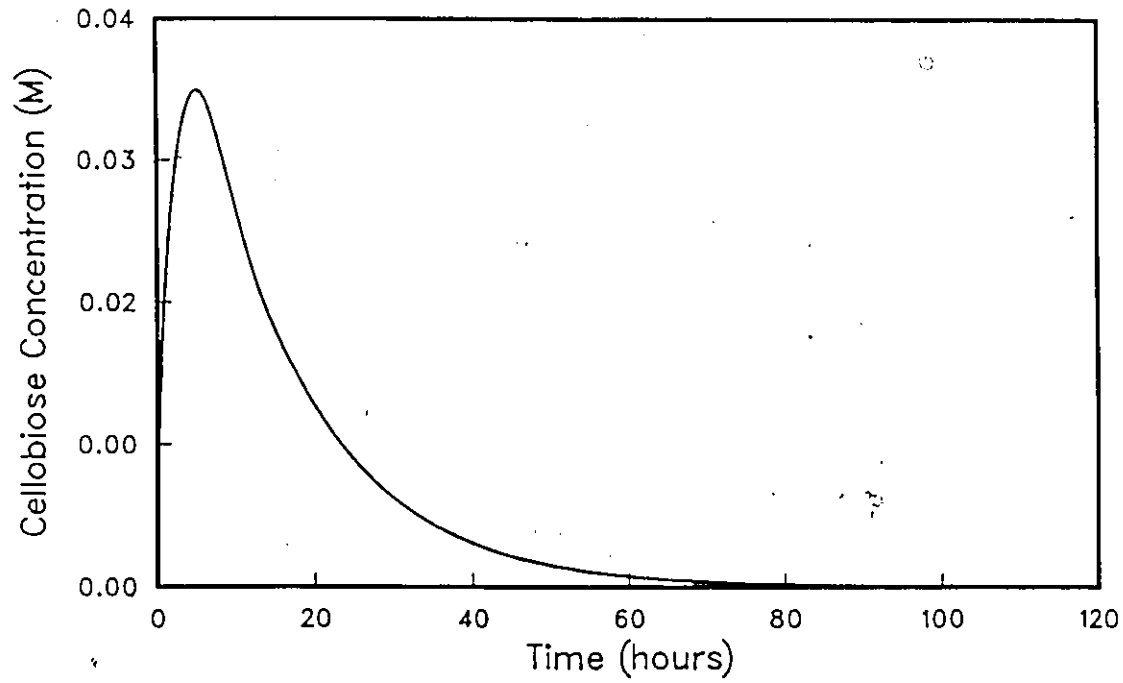


(b) Cellobiose vs Time

Figure 37: Cellobiose Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

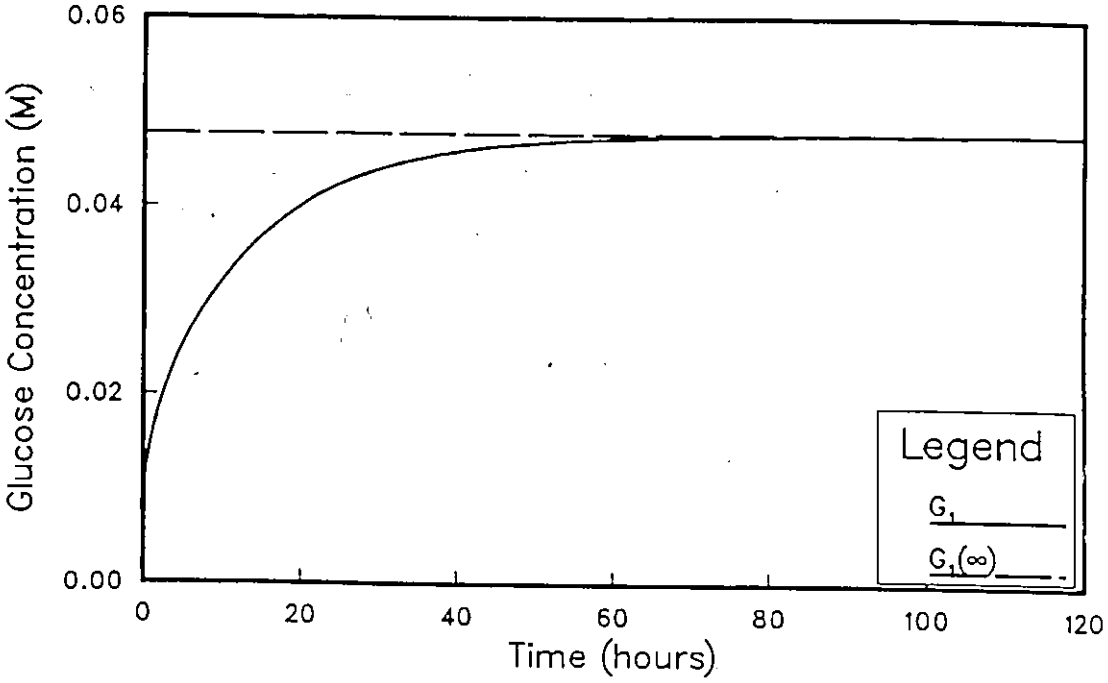


(a) Glucose vs Time

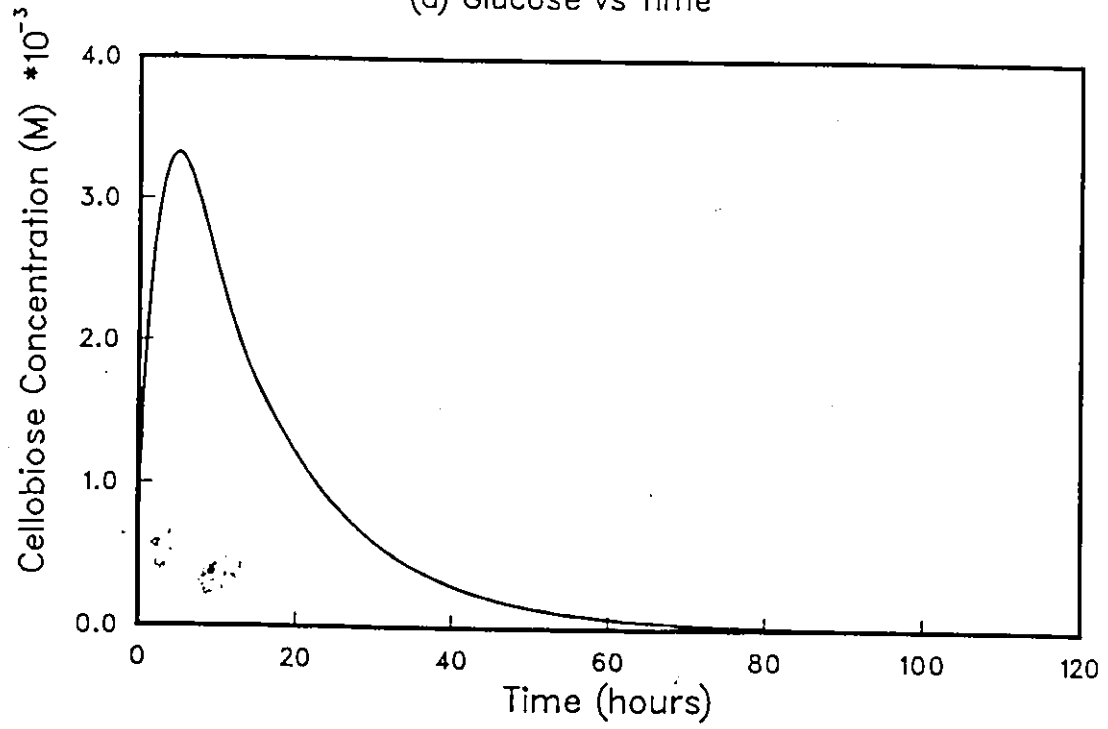


(b) Cellobiose vs Time

Figure 38: Soluble Cellulose Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

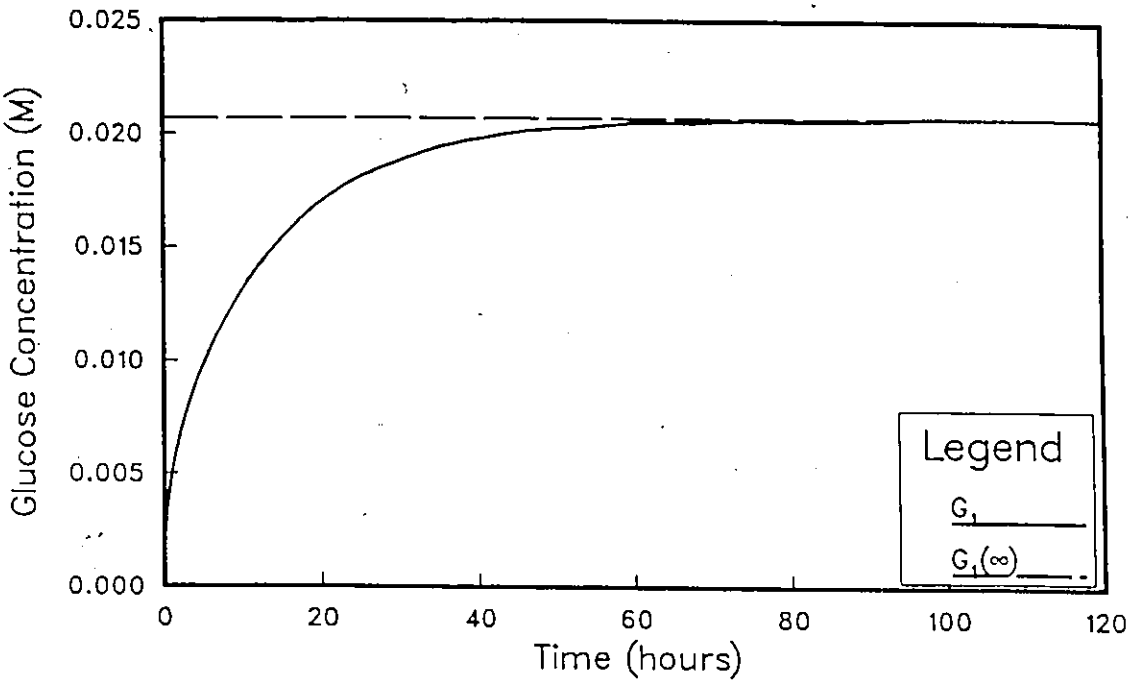


(a) Glucose vs Time

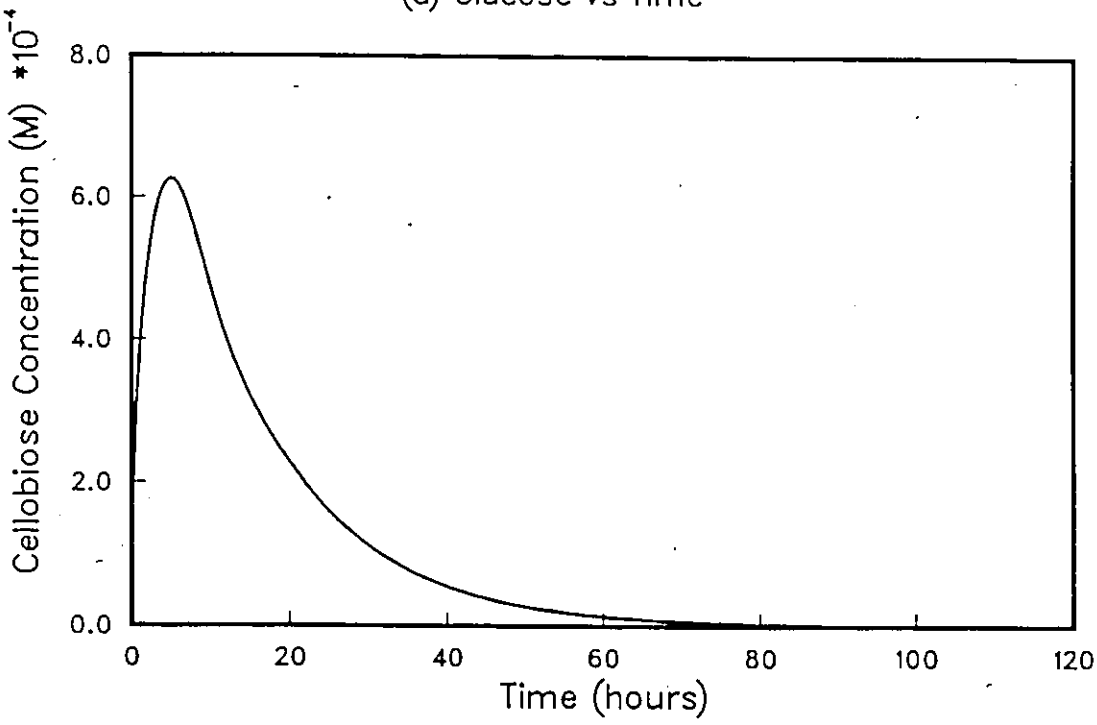


(b) Cellobiose vs Time

Figure 39: CMC DS 0.511 Hydrolysis Simulation, Model 13 Kinetics Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

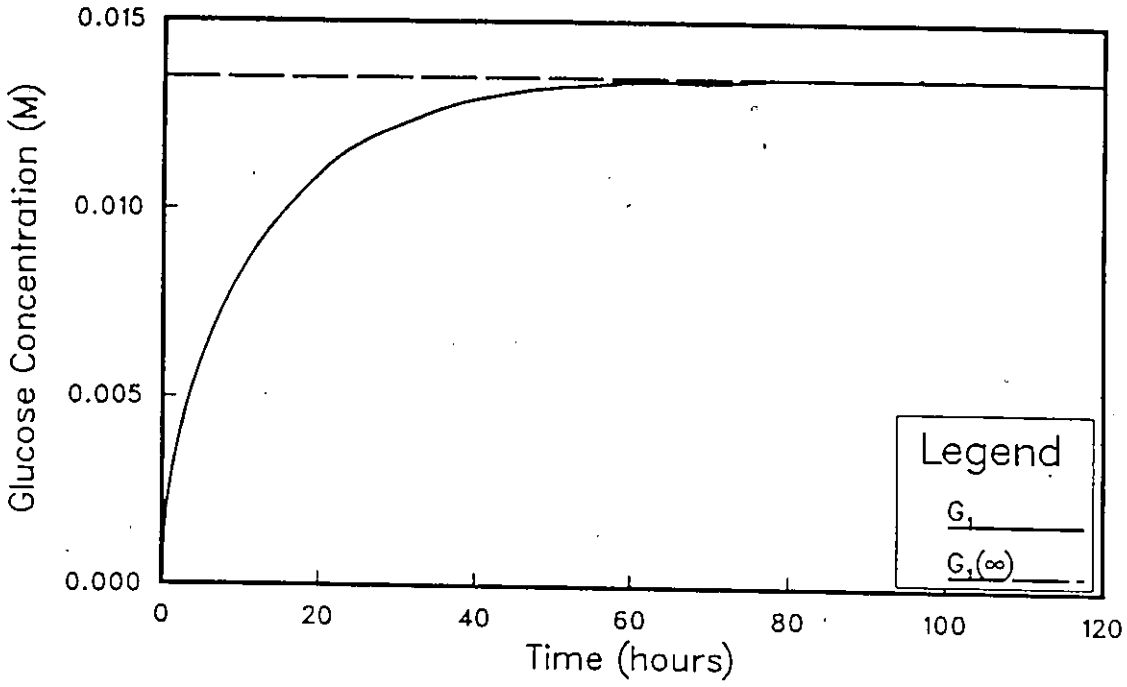


(a) Glucose vs Time

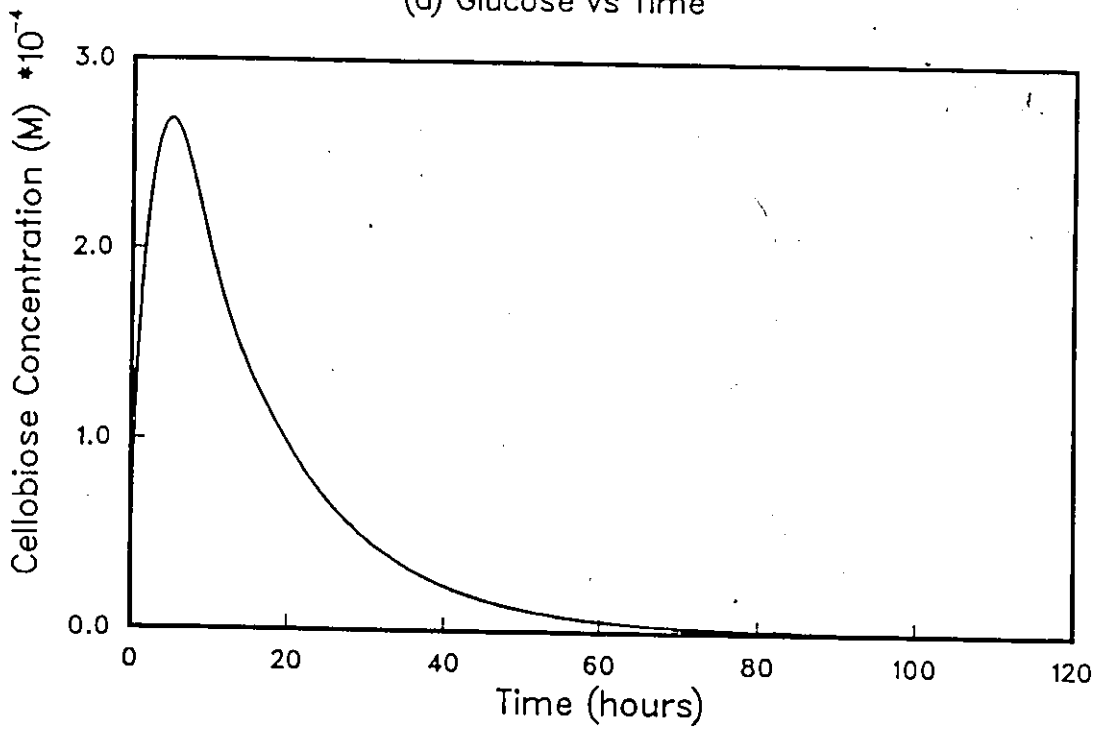


(b) Cellobiose vs Time

Figure 40: CMC DS 0.792 Hydrolysis Simulation, Model 13 Kinetics
Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L



(a) Glucose vs Time



(b) Cellobiose vs Time

Figure 41: CMC DS 0.914 Hydrolysis Simulation, Model 13 Kinetics Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

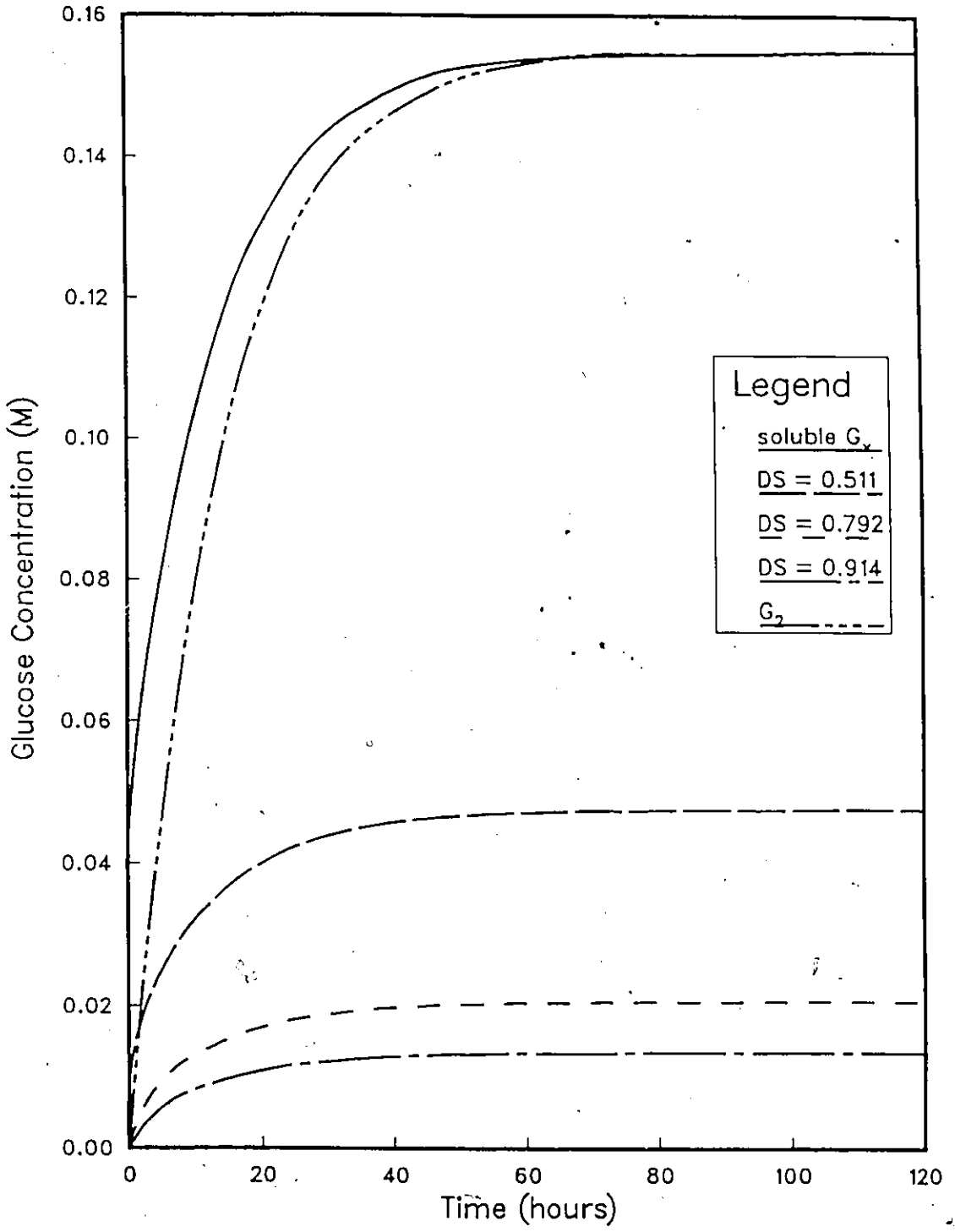


Figure 42: Enzymatic Hydrolysis Simulation, Glucose Response, Model 13 Kinetics
Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

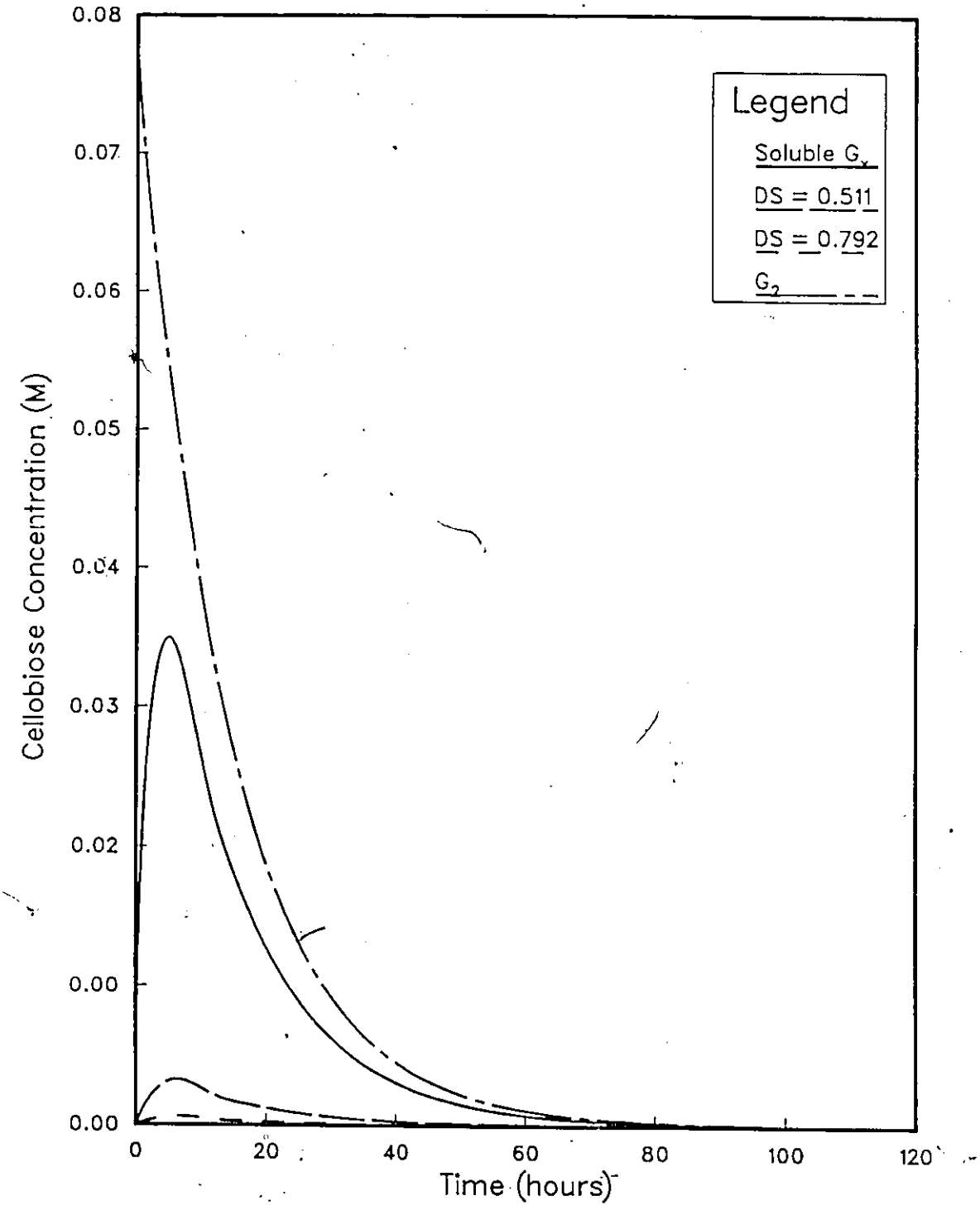


Figure 43: Enzymatic Hydrolysis Simulation, Cellobiose Response, Model 13 Kinetics
Units=0.155 M, E_1=30 I.U./L, E_2=0 I.U./L, E_x=200 I.U./L

and β_{nrc} bonds, increasing the glucose concentration in the reactor faster than if cellobiose had been the only substrate and hence E_1 the only important enzyme. This behavior weakens the assumption that E_x is unable to cleave G_2.

The simulations also reflect the theory that was applied in developing a bond hydrolysis model. The DS affects the glucose response at reaction completion. For example, going from a DS of 0.511 to 0.792, $G_1(\infty)$ is 2.3 times higher at the lower DS. For the same step change in DS, the maximum cellobiose concentration is 5.6 times higher at the lower DS. DS affects the cellobiose progress curve more than the glucose progress curve.

6.5 Modelling the Enzymatic Hydrolysis of Insoluble Cellulose

The hydrolysis of the insoluble substrate cellulose needs to be modelled using our approach. There will have to be modifications taken into account when developing the model though. The lack of substituents makes the nomenclature easy. Of the ten species traced by the CMC model, only five are found in the reactor during pure cellulose hydrolysis, G_1, G_2, β_u, β_{rc} and β_{nrc}. β_{rcs}, β_{nrcs}, β_{sr}, β_{snr} and β_{ss} are not present in the reactor. However, we suggest that more nomenclature must be introduced because of the substrate crystallinity. The β_u bonds will be segregated into an accessible fraction and an inaccessible fraction. No cellulase is capable of hydrolyzing the inaccessible β_u. But the accessible β_u is assumed to be hydrolyzed at the same rate as the β_u bonds in CMC. This is to be consistent with the assumption that E_x—or E_1 or E_2—action cannot distinguish between substrates i.e., the substrate type is transparent to the enzyme as long as it is a legitimate substrate. Since it has been shown that the crystallinity of the cellulose changes with hydrolysis and that total enzymatic hydrolysis of cellulose is possible, there has to be a mechanism whereby the inaccessible β_u bonds are transformed into accessible β_u bonds. Such an assumption would be consistent with most of the reported

observations concerning insoluble cellulose hydrolysis.

In a review [3], eight factors are reported to influence the suceptibility of cellulose to enzymatic hydrolysis. The moisture content of the fiber, the diffusivity and size of the enzyme molecules, the unit cell-dimension and the nature of substances with which the cellulose is associated are, in our opinion, not important in modelling G_z hydrolysis. They are physically unchanging from one hydrolysis to another or they can be a subset of the following four quantities that are more important to chemical engineers: the crystallinity of the cellulose, the conformation of the anhydroglucose unit, the degree of polymerization and, if applicable, the nature and distribution of the substituent groups.

Gilbert and Tsao [3] suggest that the following physical events must be taken into account when modelling cellulose hydrolysis:

- Mass transfer of the enzyme from the bulk solution to the cellulose and diffusion within pores.
- Enzyme adsorption on the cellulose surface.
- The actual kinetics of the hydrolysis reaction.
- The mobility of the cellulase in the adsorbed layer.

The observations leading to believe the importance of all the physical events can be explained by a change in nomenclature, similar to what has been done here to mechanistically describe the overall reaction in a parsimonious way. It is hoped this research will provide a mathematically structured way of analyzing the hydrolysis of both cellulose and its derivatives. Several researchers in the field of hydrolysis have supported the concept of bonds as the object of attack by cellulases. Here, we have presented the modelling tool to treat the kinetics of the hydrolysis of β-1,4 bonds linking anhydroglucose units.

Our data has a high variance yet some important points were shown: the types of

bonds hydrolyzed has been identified, the relative rate of two enzymes is estimated, there is likely a third enzyme releasing cellobiose from the non-reducing end. It is hoped that more accurate data and standardized enzyme activity tests can lead to the definite model describing, not the kinetics of cellulose hydrolysis, but the kinetics of β-1,4 glucosidic bonds' hydrolyses, regardless of their parent molecule.

Chapter 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

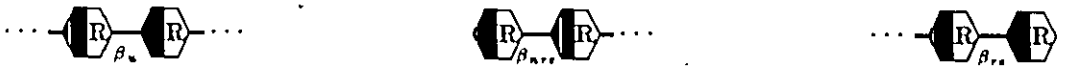
1. A new approach to modelling the kinetics of the enzymatic hydrolysis of soluble cellulose derivatives has been developed. There are two important aspects in the approach. The first is the assumption that the enzymes in the cellulase system, E_x, E_2 and E_1, form a complex with selected β-1,4 glucosidic bonds. Some of these bonds are the substrate bonds to each of the enzymes. The bonds that the enzymes are not capable of cleaving are non-substrate bonds. The assumption of an enzyme-bond complex lets us use Michaelis-Menten kinetics to represent the reaction kinetics. The second important aspect of our work is the identification of the nomenclature to develop the mechanistic model. This nomenclature is particular to the cellulase system acting on derivatives of cellulose. There are indeed ten distinct species in a CMC hydrolysis reaction mixture: G_1, G_2, β_u, β_{rc}, β_{nrc}, β_{rcs}, β_{nrcs}, β_{sr}, β_{snr} and β_{ss}.

   The reaction kinetics are complex because there are three enzymes acting on nine potential substrates—G_1 being a reaction product. The model development shows that a sound modelling procedure can reveal the mechanism of enzymatic hydrolyses.

2. After developing the theory, it was put to the test by fitting 189 data points to several models. The modelling procedure verified some assumptions that were discussed in the Literature Survey. Mechanistic assumptions were kept if they improved the basic Michaelis-Menten model. The information regarding the mechanism is affected in part by the high variance of our data. However,

modelling leads to the following observations.

Endoglucanase, E_x, cleaves β-1,4 glucosidic bonds between two unsubstituted anhydroglucose units with the exception of cellobiose. But another conclusion



regarding enzyme activity test results will explain why confidence is not high for this exception. E_x also forms a complex with all bonds in the reaction mixture; Non-substrate bonds of E_x are its inhibitors, including G_2. This effect is included in the 'β' term in equation 133. The equation describing the rate of direct action by E_x on its substrate bonds, β_u, β_{re} and β_{nre}, is:

$$\text{Rate of cleavage of bonds by } E_x = -\frac{k_{1E_x} \cdot E_x \cdot s}{\beta} \qquad (133)$$

where s is the concentration of all E_x substrate (M), E_x is the enzyme concentration in the reaction mixture measured with an FPA test, in units of I.U./L and β is the total concentration of β-1,4 bonds in the reactor (M). The kinetic constant, k_{1E_x}, was found to be $3.66 \cdot 10^{-7} \pm 1.37 \cdot 10^{-7}$ moles/I.U./s.

Cellobiase, E_1, cleaves bonds at the non-reducing end of the polymer. The size of the polymer is unimportant as E_1 cleaves β_{nre} and β_{nres} in CMC as well as the bond of cellobiose, G_2. The last unit in the chain, at the non-reducing end must however be unsubstituted. With the high variance in our data, we could

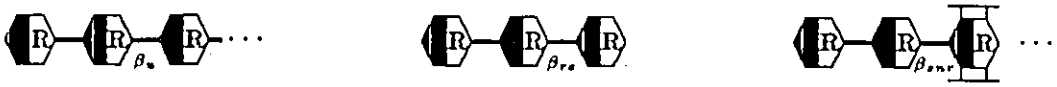
not find inhibition of E_1 by any species in the reactor. The equation describing the rate of direct action by E_1 on its substrate bonds, G_2, β_{nre} and β_{nres}, is:

$$\text{Rate of cleavage of bonds by } E_1 = -k_{1E_1} \cdot E_1 \cdot s \qquad (134)$$

where $s = G_2 + \beta_{nre} + \beta_{nres}$ is the concentration of all E_1 substrate (M) and E_1 is the enzyme concentration (I.U./L) in the reaction mixture, measured

with a CBA test described in the Experimental Procedure section. The kinetic constant, k_{1E_1}, was found to be $6.59 \cdot 10^{-7} \pm 1.01 \cdot 10^{-7}$ L/I.U./s.

In the reaction mixture, the much higher observed concentration of cellobiose than predicted by a random attack on the hydrolyzable bonds was attributed by default to exoglucanase, E_2, since we could not change the assumed mechanism of E_1 and E_x to explain it. This time, high variance in the data prohibits us from estimating the kinetic parameters. Nonetheless, we made the assumption



that E_2 in the reactor acted at a virtually infinite rate in cleaving the β_u, β_{nre} and β_{snr} bonds that are second from the non-reducing end. This eliminated the trend in the cellobiose residuals. Thus it was likely that the presence of a third enzyme, E_2, contributed to the high cellobiose concentration. An added condition for the action of E_2 is that the two consecutive end units must be unsubstituted. More precision in measuring G_2 will allow the estimation of E_2's kinetic parameters.

3. Our CBA test, using G_2 as substrate gives an estimate of the amount of cellobiase in a mixture. However the test results are valid if no other enzyme in the mixture cleaves the β-1,4 glucosidic bond of cellobiose. For example, E_x may be a random acting enzyme that cleaves the β-1,4 glucosidic bond of cellobiose at the same rate as it cleaves all β-1,4 bonds between unsubstituted anhydroglucose units in CMC. Perhaps because of the high variance in our data and perhaps because we are making the assumption that E_x does not cleave G_2 in drawing conclusions from the CBA test, the overall hydrolysis results cannot conclusively prove or disprove that E_x binds and cleaves the β-1,4 bond of cellobiose. Further research on cellulase enzymes must identify substrate bonds that are specific to each enzyme in a solution. A substrate can then be used

to assay the particular enzyme that attacks it. From the Literature Survey, in retrospect, salicin and *p*-nitrophenyl-β-glucoside may be specific substrates of E_1.

4. There is insufficient information for the runs where the only substrate is cellobiose. Runs should have been done for ratios of the enzyme components other than for the injection of pure Celluclast or pure Novozym.

5. According to the theory, the specificity of the FPA test is at least as suspect as the CBA test. From the proposed mode of action of E_x, a viscometric method based on hydrolysis of a highly substituted derivative would have been more accurate [31]. The release of reducing sugars from the hydrolysis of filter paper depends on the combined action of three enzymes. However, as long as the measured ratio of E_x in Celluclast to E_x in Novozym is close to the true value, the model will not show a bias, regardless of the true, individual activities.
   From our use of the FPA test, it appears that the DNS reagent ma be in part reacting with the reducing sugars. Using a strict modelling procedure, we found evidence of non-linearity in the enzyme dilution series. This violates the requirement of a linear enzyme dilution series in assaying enzymes. We did not notice the non-linearity in the CBA test except where it was due to cryptic substrate limitation In our wide range of DF.

6. A strict modelling approach to assaying enzyme activity allows the calculation of confidence intervals for the activity. Non-linearities of enzyme dilutions can be identified under conditions that do not make them obvious under the present procedure. This strict statistical approach indicates that more information may be drawn from an FPA vs DF plot than the presently used Glucose vs Enzyme plot.

7. The mechanism on which the overall hydrolysis model is based assumes that the initial DP of the polymer is high. The probability of a certain type of bond

being adjacent to another of the same or different type is calculated as the fraction of that bond to all eligible bonds for the position considered. If the initial length of a polymer falls below a threshold level, say 8 units, the model will fail. For example, slight modifications would have to be made to the model in order to describe the hydrolysis of cellotetraose because the initial probability of finding a β_{nre} bond besides a β_{re} is nil—all β_{re} and β_{nre} bonds are separated

by a β_u bond. The model would however calculate the probability to be 0.5 since the two types of bonds allowed are in equal concentrations. Cellotriose is an exception to this model failure at low DP.

At high DP, the actual DP of the polymer is of no importance. i.e., CMC of DP 1000 gives virtually the same progress curve as CMC of DP 10000, provided that one does not specify $E_x = 0$.

7.2 Recommendations

1. With cellulase research having slowed down since the early 80's, it is hoped that this work can rekindle interest in the subject. Some recommendations are possible with respect to future work in this field. One assumption of our model is that the substrate type is transparent to a given enzyme component i.e., the enzyme acts at the same rate on all substrate bonds when it cleaves them. The data and modelling techniques were applied to only one of several possible substrates. Other cellulose derivatives such as hydroxyethyl cellulose, cellulose acetate, methyl cellulose and soluble short chains of cellulose should be tested to verify that the model applies universally to all substrates containing β-1,4 bonds. Of course each of these substrates has a particular distribution of the substituent group on the cellulose chain. The characteristics of this distribution must be identified (or included as parameters in the nonlinear estimation routine) in

each case. Cellodextrins do not have substituents.

2. Cellulose is a special case of bond hydrolysis since it is insoluble. It has been shown that the initial rate of G_x hydrolysis is lower than for CMC but that the steady-state concentration of product is higher [7]. We believe that only a change in the substrate definition is necessary to model cellulose saccharification with the bond hydrolysis approach. Some of the β-1,4 bonds of cellulose are available to the enzymes; the bonds that are not immediately available will eventually undergo a transformation to a hydrolyzable form, not unlike the bond transformations that we explained in our CMC hydrolysis model development. Some attention—this is obvious—of biomass research groups must turn toward identifying the mechanism by which bonds in cellulose are transformed to a hydrolyzable form by the cellulases. It may not be diffusion that limits the extent of cellulose hydrolysis. That is all that will be said here about a future extension of the bond hydrolysis model to cellulose saccharification.

3. Finally, the present tests to determine the activity of the cellulase components must be redesigned. First, specific bond substrates of individual enzymes must be identified. They will be eligible assay substrates of the new activity tests. In this way, interference by enzymes of no interest will be eliminated. Second, the test results are strictly valid at infinite dilution. (This is analogous to the measurement of the intrinsic viscosity of a polymer). Hence a plot of the activity as a function of the dilution factor must be extrapolated to infinite dilution (of the enzyme). The two recommendations put together should yield an unbiased, if not absolute, estimate of the concentration of an individual enzyme component in a solution.

REFERENCES

1. Knuth, D.E., *The TEXbook*, Addison-Wesley Publishing Co. ©1984.

2. TELLAGRAF, *ISSCO Visual Information System Software*, 10505 Sorrento Valley Road, San Diego, CA 92121-1698, ©1984.

3. Gilbert, I.G. and Tsao, G.T., *Interaction Between Solid Substrate and Cellulase Enzymes in Cellulose Hydrolysis*, *Annual Reports on Fermentation Processes* **6**, Academic Press (1983), 323–358.

4. Humphrey, A.E., *The Hydrolysis of Cellulosic Materials to Useful Products*, *Advances in Chemistry Series* **181** (1979), 25–53.

5. *Hercules Cellulose Gum—Chemical and Physical Properties and Uses*, *Hercules Inc. Product Information Bulletin No. 250-10*, (1984).

6. Mandels, Mary., *Cellulases*, *Annual Reports on Fermentation Processes* **5**, Academic Press (1982), 35–78.

7. Zabriskie, D.W., Qutubuddin, S.A.S.M. and Downing, K.W., *Production of Ethanol from Cellulose Using a Soluble Cellulose Derivative as an Intermediate*, *Biotechnology & Bioengineering Symp. No. 10* (1980), 149–162.

8. Wood, T.M. and McCrae, S.I., *Synergism Between Enzymes Involved in the Solubilization of Native Cellulose* *Advances in Chemistry Series* **181** (1979), 181–209.

9. Gritzali, M. and Brown, R.D.Jr., *The Cellulase System of* Trichoderma, *Advances in Chemistry Series* **181** (1979), 237–260.

10. Lee, S.E., Armiger, W.B., Watteeuw, C.M. and Humphrey, A.E., *A Theoretical Model for Enzymatic Hydrolysis of Cellulose*, *Biotechnology & Bioengineering* **20** (1978), 141–144.

11. Ladisch, M.R., Lin, K.W., Voloch, M. and Tsao, G.T., *Process Considerations*

in the Enzymatic Hydrolysis of Biomass, *Enzyme Microb. Technol.* **5**, March (1983).

12. Gong, C.-S., Ladisch, M.R. and Tsao, G.T., *Biosynthesis, Purification and Mode of Action of Cellulases from Trichoderma reesei*, *Advances in Chemistry Series* **181** (1979), 261–287.

13. Poulsen, O.M. and Petersen, L.W., *A Standard Formula for the Determination of the Initial Rate of Hydrolysis of Carboxymethylcellulose*, *Biotechnology and Bioengineering* **27** (1985), 409–414.

14. Darbyshire, J., *Large Scale Enzyme Extraction and Recovery*, *Topics in Enzyme and Fermentation Biotechnology* **5**, Wiseman, A., editor, Halsted Press (1981), 147–186.

15. Reese, E.T., Siu, R.G.H. and Levinson, H.S., *The Biological Degradation of Soluble Cellulose Derivatives and its Relationship to the Mechanism of Cellulose Hydrolysis*, *J. Bacteriol.* **59** (1950), 485–493.

16. Huang, A.A., *Kinetic Studies on Insoluble Cellulose-Cellulase System*, *Biotechnology & Bioengineering* **27** (1975), 1421–1433.

17. Reese, E.T., *Biological Degradation of Cellulose Derivatives*, *Industrial and Engineering Chemistry* **49**(1), (1957).

18. Fan, L.T. and Lee, Y.-H., *Kinetic Studies of Enzymatic Hydrolysis of Insoluble Cellulose: Derivation of a Mechanistic Kinetic Model*, *Biotechnology & Bioengineering* **25** (1983), 2707–2733.

19. Lee, Y.-H. and Fan L.T., *Properties and Mode of Action of Cellulases*, *Advances in Biochemical Engineering* **17** (1980), 101–129.

20. Matteau, P.P. and Saddler, J.N., *Glucose Production Using Immobilized Mycelial-Associated β-Glucosidase of Trichoderma E58*, *Biotechnology Letters* **4**(8) (1982), 513–518.

21. Montenecourt, B.S. and Eveleigh, D.E. *Selective Screening Method for the Isolation of High Yielding Cellulase Mutants of Trichoderma reesei*, *Advances in Chemistry Series* **181** (1979), 289–301.

22. Ohmine, K., Ooshima, H. and Harano, Y., *Kinetic Study on Enzymatic Hydrolysis of Cellulose by Cellulase from Trichoderma Viride*, *Biotechnology & Bioengineering* **25** (1983), 2041–2053.

23. Howell, J.A. and Stuck, J.D., *Kinetics of Solka Floc Cellulose Hydrolysis by Trichoderma Viride Cellulase*, *Biotechnology & Bioengineering* **17** (1975), 873–893.

24. Kanda, T., Shingo, N., Wakabayashi, K. and Nisizawa, K., *The Mode of Enzymatic Degradation of Cellulose Based on the Properties of Cellulase Components*, *Advances in Chemistry Series* **181** (1979), 211–236.

25. Okazaki, M. and Moo-Young, M., *Kinetics of Enzymatic Hydrolysis of Cellulose: Analytical Description of a Mechanistic Kinetic Model*, *Biotechnology and Bioengineering* **20** (1978), 637–663.

26. Canevascini, G. and Gattlen, C., *A Comparative Investigation of Various Cellulase Essay Procedures*, *Biotechnology & Bioengineering* **23** (1981), 1573–1590.

27. Suga, K., Van Dedem, G. and Moo-Young, M., *Degradation of Polysaccharides by Endo and Exo Enzymes: A Theoretical Analysis*, *Biotechnology and Bioengineering* **17** (1975), 433–439.

28. Bailey, James E. and Ollis, David F. *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company (1977).

29. Heptinstall, J., Stewart, J.C. and Seras, M., *Fluorimetric Estimation of Exocellobiohydrolase and β-D-glucosidase activities in cellulase from Aspergillus fumigatus fresenius*, *Enzyme and Microbial Technology* **8** (1986), 70–74.

30. Lindner, W.A., Dennison, C. and Quicke, G.V., *Pitfalls in the Assay of Carbo-*

xymethylcellulase Activity, Biotechnology & Bioengineering **25** (1983), 377–385.

31. Demeester, J., Bracke, M. and Lauwers, A., *Absolute Viscosimetric Method for the Determination of Endocellulase Activities Based Upon Light-Scattering Interpretations of Gel Chromatographic Fractionation Data*, *Advances in Chemistry* **181** (1979), 91–125.

32. Mandels, M., Andreotti, R. and Roche, C., *Measurement of Saccharifying Cellullase*, *Biotechnology and Bioengineering Symp. No. 6* (1976), 21–33.

33. Somogyi, M., *A New Reagent for the Determination of Sugars*, *Journal of Biological Chemistry* **160** (1945), 61–63.

34. Nelson, N., *A Photometric Adaptation of the Somogyi Method for the Determination of Glucose*, *Journal of Biological Chemistry* **153** (1944), 375–380.

35. Somogyi, M., *Notes on Sugar Determination*, *Journal of Biological Chemistry* **195** (1952), 19–23.

36. Savage, A.B., *Temperature-Viscosity Relationships for Water-Soluble Cellulose Ethers*, *Industrial and Engineering Chemistry* **49**(1), (January 1957).

37. Hägerdal, B., Ferchak, J., Pye, E.K. and Forro, J.R., *The Cellulolytic Enzyme System of Thermoactinomyces*, *Advances in Chemistry Series* **181** (1979), 331–345.

38. Gianfreda, L., Marrucci, G., Grizzuti, N. and Greco, G.Jr., *Series Mechanism of Enzyme Deactivation. Characterization of Intermediate Forms*, *Biotechnology and Bioengineering* **27** (1985), 877–882.

39. André, G. and McLean, D.D., *Modelling the Enzymatic Hydrolysis of Cellulose: Some Thoughts*, Internal Paper, Dept. of Chemical Engineering, University of Ottawa, (Sept. 1983).

40. Huebner, A., Ladisch, M.R. and Tsao, G.T., *Preparation of Cellodextrins: An Engineering Approach*, *Biotechnology and Bioengineering* **20** (1978), 1669–1678.

41. Hsu, T.-A., Gong, C.-S. and Tsao, G.T., *Kinetic Studies of Cellodextrin Hydrolyses by Exocellulase from T. reesei*, Biotechnology and Bioengineering **22**(11) (1980), 2305–2320.

42. Timell, T.E., *A Method for Estimating the Distribution of the Substituents in Partially Substituted Carboxymethylcelluloses*, Svensk Papperstidning **55**(17) (1952), 649–660.

43. Lee, Y.-H., Fan, L.T. and Fan, L.-S., *Kinetics of Hydrolysis of Insoluble Cellulose by Cellulase Advances in Biochemical Engineering* **17** (1980), 131–168.

44. Attala, R.H., *Conformational Effects in the Hydrolysis of Cellulose*, Advances in Chemistry Series **181** (1979), 55–69.

45. Banergee, D.K. and Laidler, K.J., *Stopped Flow Kinetic Studies of the Cellulase-Catalyzed Conversion of Cellulose to Glucose*, Biotechnology and Bioengineering **25** (1983), 2413–2418.

46. Spurlin, H.M., *Arrangement of Substituents in Cellulose Derivatives*, Journal of the American Chemical Society **61** (1939), 2222–2227.

47. Timell, T.E., *The Distribution of the Substituents in Partially Substituted Carboxymethylcelluloses*, Svensk Papperstidning **55**(18) (1952), 700–708.

48. DeButts, E.H., Hudy, J.A. and Elliot, J.H., *Rheology of Sodium Carboxymethyl Cellulose Solutions*, Industrial and Engineering Chemistry **49**(1),(1957).

49. *Hercules Cellulose Gum—Analytical Procedures for its Assay and Determination in Formulations*, Hercules Bulletin No. *FF-318.*

50. *SAS User's Guide: Statistics*, Version 5 Edition, 1985, SAS Institute Inc., Cary, NC.

51. Bacon, D.W. *Collection and Interpretation of Industrial Data*

52. Chatelain, R., Operations Dept., Computing Centre, University of Ottawa, *Personal Communication.*

Appendix 1

REAGENT SPECIFICATIONS

1. Sodium carboxymethyl cellulose, Hercules "cellulose gum", type 7L, lot 66745.
2. Sodium carboxymethyl cellulose, Hercules CMC, type 9M31F, lot no. 65662, date 03-83.
3. Sodium carboxymethyl cellulose, Hercules CMC, type 4H1, lot no. 61410, date 06-82.
4. Dextrose (anhydrous D-Glucose), Fisher Scientific, lot 733808, specific rotation +53.0°.
5. D(+) cellobiose, No. C-7252, Sigma Chemical Company, lot 43F-0366
6. D-cellobiose, cat. 115 4509, Eastman-Kodak company, lot no. A14A.
7. Uranyl nitrate (depleted), $(UO_2(NO_3)_2 \cdot 6H_2O)$, Fisher Scientific Company, lot. 742437, Fisher Assay: 98.0%, F.W. 502.13, radioactivity 6 microcuries in 25 grams.
8. 3,5 Dinitrosalicylic acid, Sigma Chemical Company, lot 112F-0031, anhydrous molecular weight 228.1.
9. Acetonitrile HPLC grade, Fisher Scientific, F.W. 41.05, lot 851714, Filtered through a 0.5 micron filter, Fisher assay (as CH_3CN) 99.9%
10. Cellulase, Celluclast^{TM}, Novo industries, in concentrated solution form, from *Trichoderma reesei*, nominal activity: 2000 CcmcU/g.
11. Cellobiase, Novozym^{TM}, Novo industries, in concentrated solution form, from *Aspergillus niger*, nominal activity: 250 CBU/g.

Appendix 2

MODELLING THE ENZYME ACTIVITIES

2.1 Introduction

The usual assay procedure for the enzymes consists in conducting the standard test at or near the DF at which 0.667 g/L of glucose is produced [32]. If the amount of glucose obtained initially in the test is too high, the enzyme solution is diluted until the amount of glucose is low enough to satisfy the test criterion i.e., a dilution series of the enzyme should be linear. The details of the assay procedure can be found in the Experimental Procedure chapter. When the proper data have been collected, interpolate between data points to obtain $\widehat{\text{FPA}}$ or $\widehat{\text{CBA}}$ at the $G1$ yield of 0.667 g/L. The interpolation can be done graphically on a $G1$ vs enzyme volume graph. Regression may even be applied if the $G1$ vs enzyme volume curve is linear.

In this work, we apply a strict modelling procedure to the enzyme activity data. In this way, confidence intervals can be calculated for the predicted enzyme activity. In fact, the statistical analysis will reveal if there is an insufficient number of test runs. In this latter case, a thorough analysis of the data could indicate the location of DF in future runs to optimize the modelling procedure. In this work, the data were collected first and then analyzed. No attempt was made to find optimal run locations to be eventually followed by more data collection.

In the course of the statistical analysis to be presented in this appendix, please note that the modelling was conducted with the aid of a computer software package, "Statistical Analysis System" [50] henceforth referred to as "SAS". This software package consists in part of a linear curve fitting routine, `PROC REG` and a non-linear curve fitting routine `PROC NLIN`. Both routines provide, on output, the statistical diagnostics related to a least squares' fitted model.

2.2 FPA of Celluclast

The data collected for the FPA test of Celluclast are plotted on Figure 44. The model shown on the Figure is derived from considering mechanisms of enzyme kinetics and possible non-linearities of the enzyme dilution series. See the derivation in Appendix 4. Here are the parameters and their 95% confidence intervals:

$$G1 - \theta_0 \cdot \mathrm{Enz} - \theta_1 \cdot \ln\left((1 - \theta_2 \cdot G1)^2\right) = 0 \qquad \text{Celluclast}$$

$$\theta_0 = 1071 \pm 375\,\mathrm{g/L/mL\ enzyme}$$

$$\theta_1 = 0.0007 \pm 0.079\,\mathrm{g/L}$$

$$\theta_2 = -1.8 \cdot 10^9 \pm 3.8 \cdot 10^{12}\,\mathrm{L/g}$$

Evidently the model must be rejected on the basis that the confidence intervals for parameters θ_1 and θ_2 include the value 0.

The more important fact to notice from Figure 44 is the heteroscedasticity of the data. The residual plots of Figure 45 confirm the presence of heteroscedasticity; As the quantity of enzyme in the test is increased, the variance associated with the measure of collected glucose also increases. Heteroscedasticity violates an assumption of least squares' modelling. The violated assumption states that the variance of the random error associated with the measure of the dependent variable is constant over the whole region of the operating variable [51].

In the first attempt to get rid of heteroscedasticity, the dependent variable was transformed. The new dependent variable is FPA. It is calculated from the equation $\mathrm{FPA} = G1 \cdot \mathrm{DF}/10.8$. The independent variable is also changed—it is now DF—even though it has no effect on correcting the heteroscedastic behavior. DF is directly related to the volume of enzyme: $\mathrm{DF} = (\text{enzyme vol. in mL})^{-1}$. The transformed data is shown in Figure 46. A trend of increasing FPA with DF is apparent. Modelling techniques will show this trend to be significant.

The fitted curves $\mathrm{FPA} = \theta_0 + \theta_1 \cdot \mathrm{DF}$ and $\mathrm{FPA} = \theta_0 + \theta_1 \cdot \mathrm{DF}^2$ are shown along

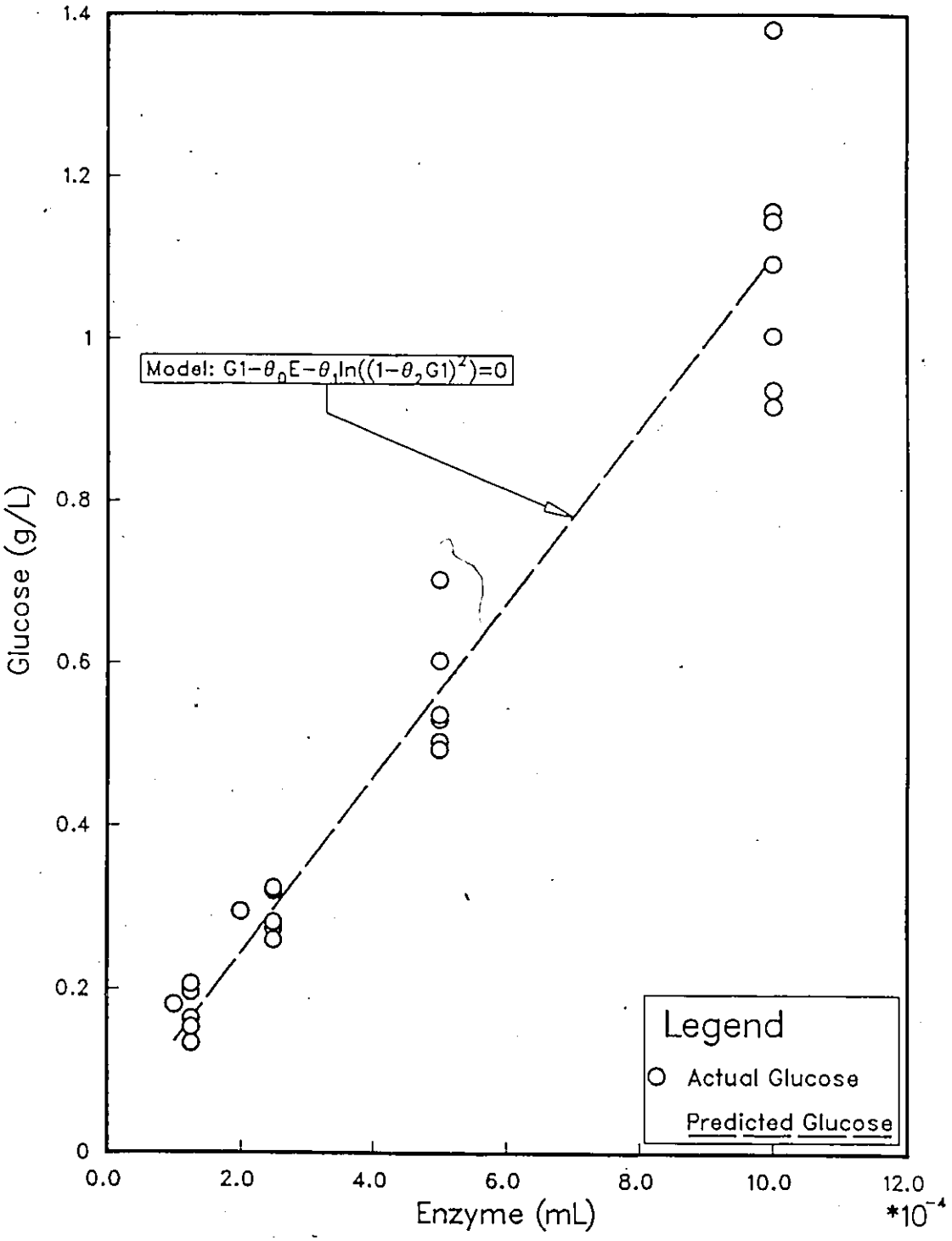
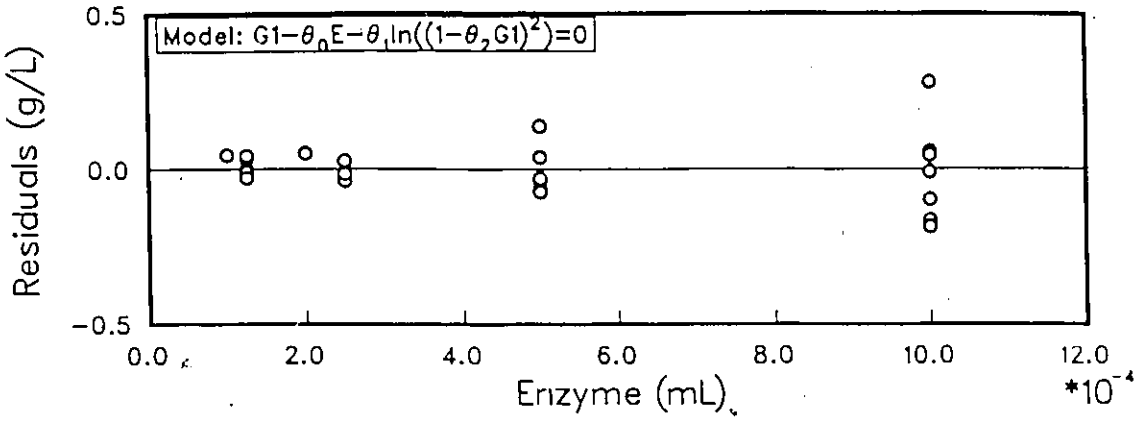
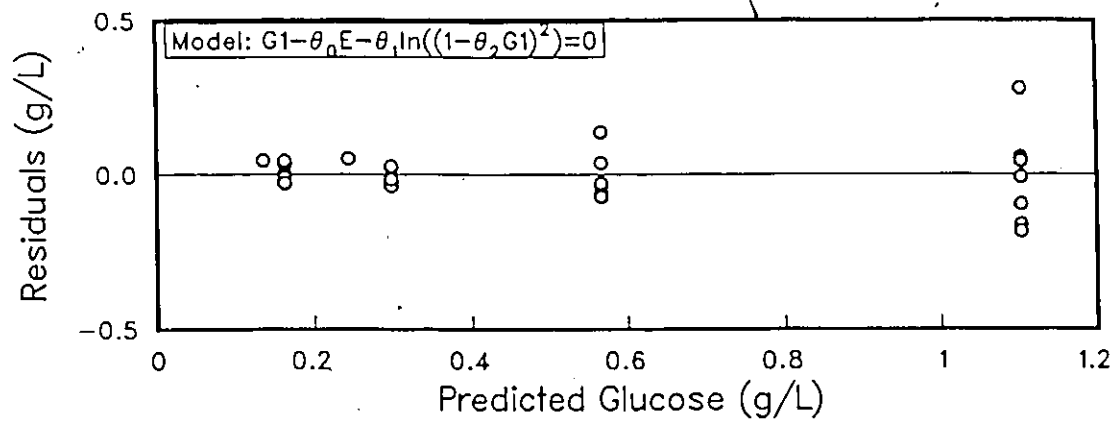


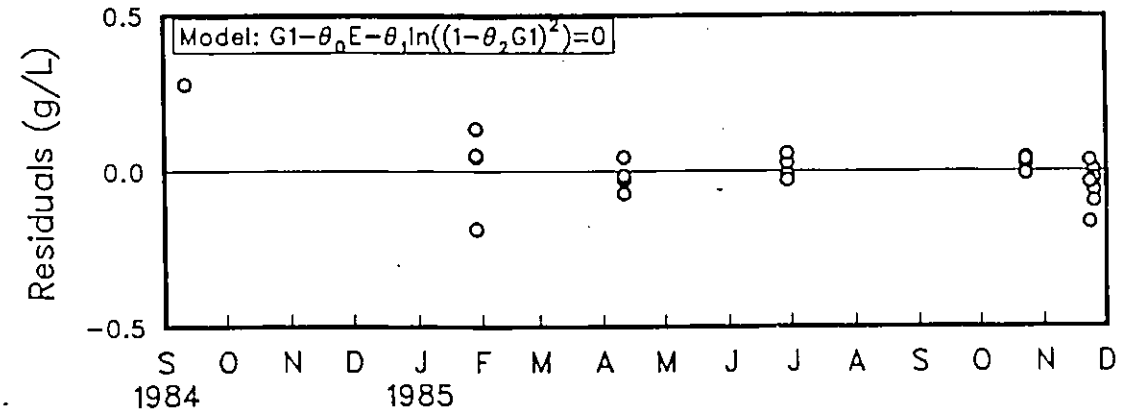
Figure 44: Actual and Predicted Gluc. vs Enz. for Celluclast FPA



(a) Residuals vs Enzyme Concentration



(b) Residuals vs Predicted Glucose



(c) Residuals vs Run Date

Figure 45: Celluclast FPA Residuals Plots
Model: $G1-\theta_0E-\theta_1\ln((1-\theta_2G1)^2)=0$

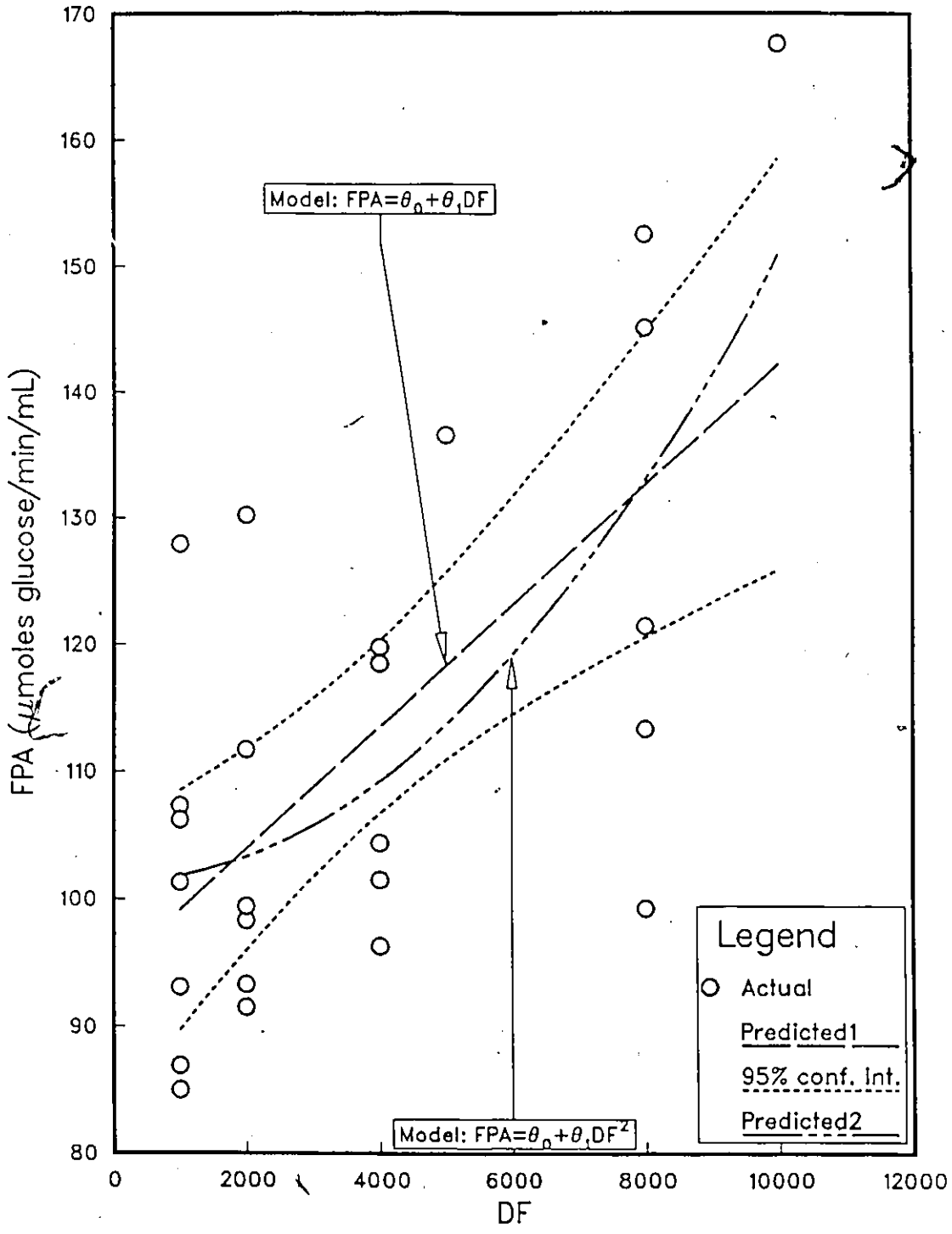


Figure 46: Actual and Predicted Celluclast FPA for 2 Models

with the actual data. Both curves are empirical. Here are the diagnostics for these two models:

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}}$ Celluclast

$$\theta_0 = 94.3 \pm 11.2$$

$$\theta_1 = 0.0048 \pm 0.0024$$

The correlation matrix of the parameters:

$$\begin{bmatrix} 1 & -0.797 \\ -0.797 & 1 \end{bmatrix}$$

$$\text{SSR} = 6134$$

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}^2}$ Celluclast

$$\theta_0 = 101.3 \pm 8.4$$

$$\theta_1 = 4.97 \cdot 10^{-7} \pm 2.32 \cdot 10^{-7}.$$

The correlation matrix of the parameters:

$$\begin{bmatrix} 1 & -0.617 \\ -0.617 & 1 \end{bmatrix}$$

$$\text{SSR} = 5824$$

The $100(1-\alpha)$ per cent confidence interval for θ_i is given by the expression

$$\theta_i \pm t_{\nu,\alpha/2}\sqrt{\text{estimated variance of } \theta_i}$$

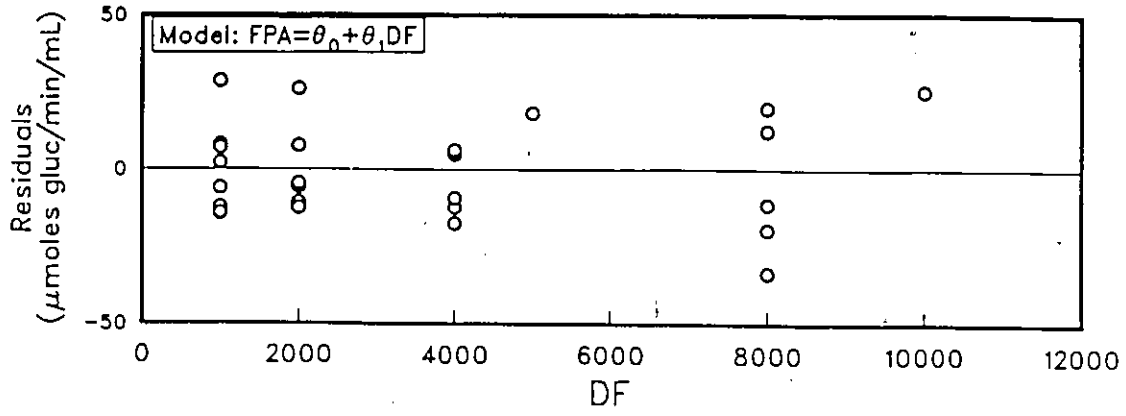
$$\theta_i \pm t_{23,.05/2} \cdot \text{std error of } \theta_i$$

For example, for the straight line model,

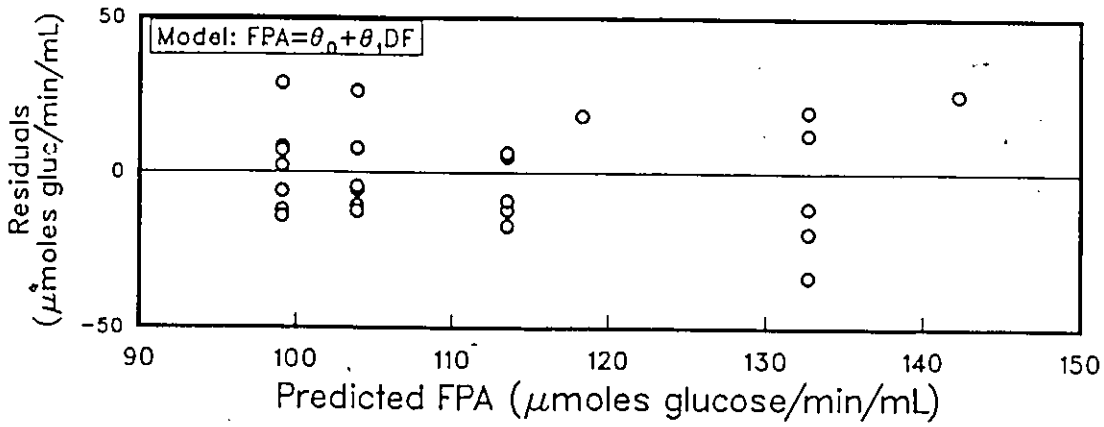
$$\theta_0 = 94.3 \pm 2.069 \cdot 5.411 \qquad \theta_1 = 0.0048 \pm 2.069 \cdot 0.00115$$

$$\theta_0 = 94.3 \pm 11.2 \qquad \theta_1 = 0.0048 \pm 0.0024$$

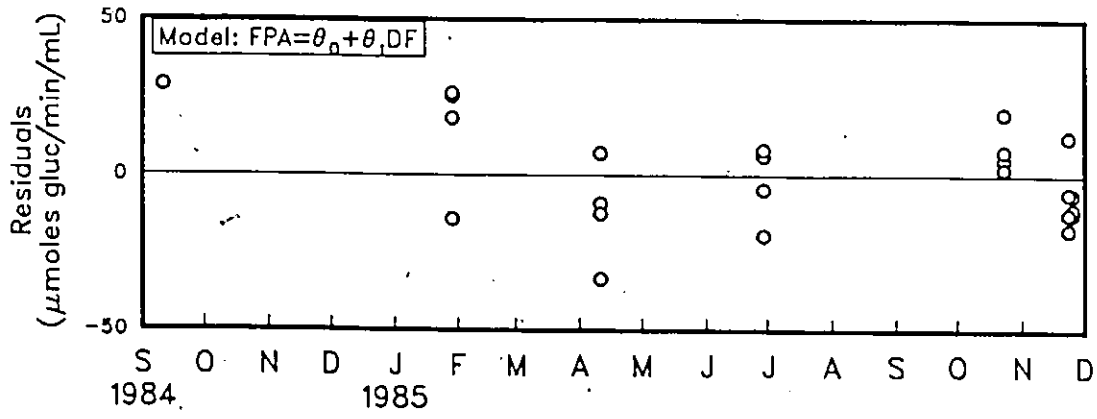
The choice of a model to describe the data was between the two equations above. All the statistical information indicates that $\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}^2$ is the better model. The confidence in the parameters is higher, the parameters are less correlated and the SSR is lower. The residual plots for the two proposed models are found



(a) Residuals vs DF

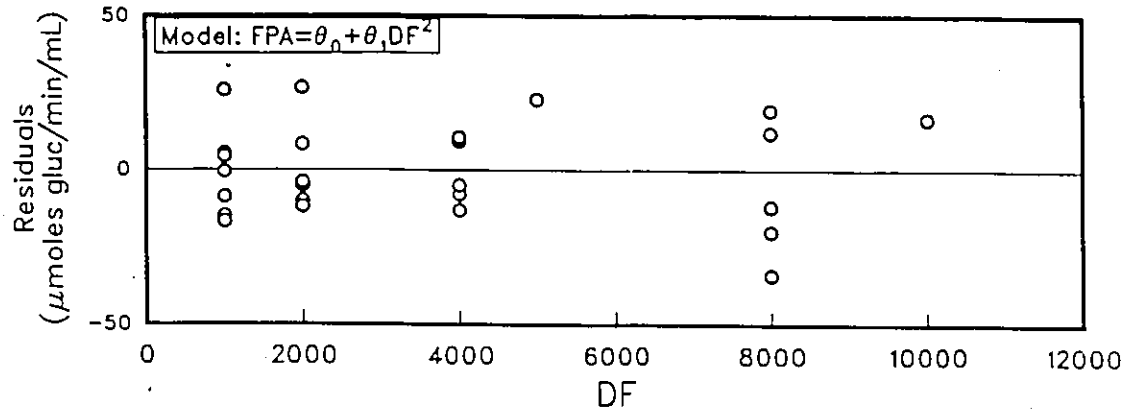


(b) Residuals vs Predicted FPA

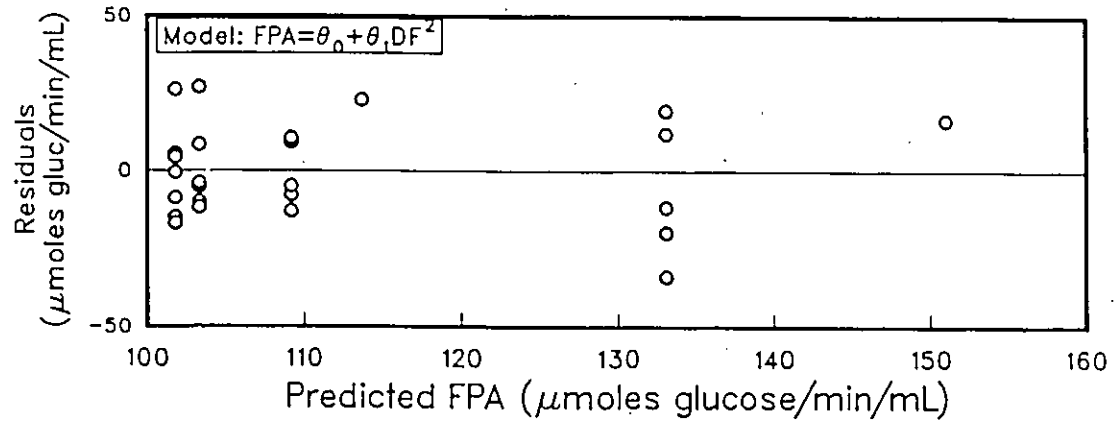


(c) Residuals vs Run Date

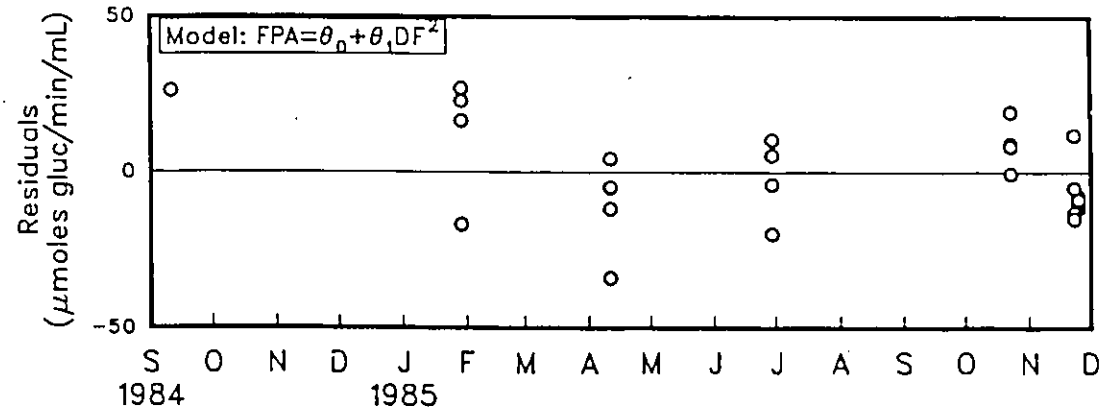
Figure 47: Celluclast FPA Residual Plots
Model: FPA=$\theta_0+\theta_1$DF



(a) Residuals vs DF



(b) Residuals vs Predicted FPA



(c) Residuals vs Run Date

Figure 48: Celluclast FPA Residual Plots
Model: FPA=$\theta_0+\theta_1 DF^2$

in figures 47 and 48. There is little difference between the two sets of residual plots. Despite those statistical indicators that favor the 2 parameter, 2nd degree polynomial, our choice for the model to represent the data was the straight line model $\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}$. Justify this choice as follows; as the DF tends to ∞, the FPA should level out to a constant value. FPA being the ratio of glucose produced to enzyme in the reactor, this means that the glucose produced as a function of enzyme concentration exhibits a linear behavior at a very high DF or low enzyme concentration. The model $\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}^2$ is not consistent with the expected glucose behavior. The higher the DF, the higher the rate of change in FPA becomes whereas it should level out. Thus the model $\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}$ is chosen. The 95% confidence interval for values predicted by this model are shown in Figure 46.

Another model, the 3-parameter, 2nd degree polynomial was also fitted then rejected. The fitted curve for a full second degree polynomial is shown in Figure 49. The residuals as a function of DF for this 3 parameter model are plotted in Figure 50. This model was rejected because of the lack of precision of the parameter estimates. Here are the estimates and 95% confidence intervals:

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF} + \theta_2 \cdot \text{DF}^2}$ Celluclast

$$\theta_0 = 102.4 \pm 19.1$$

$$\theta_1 = -0.00073 \pm 0.011$$

$$\theta_2 = 5.7 \cdot 10^{-7} \pm 10.8 \cdot 10^{-7}$$

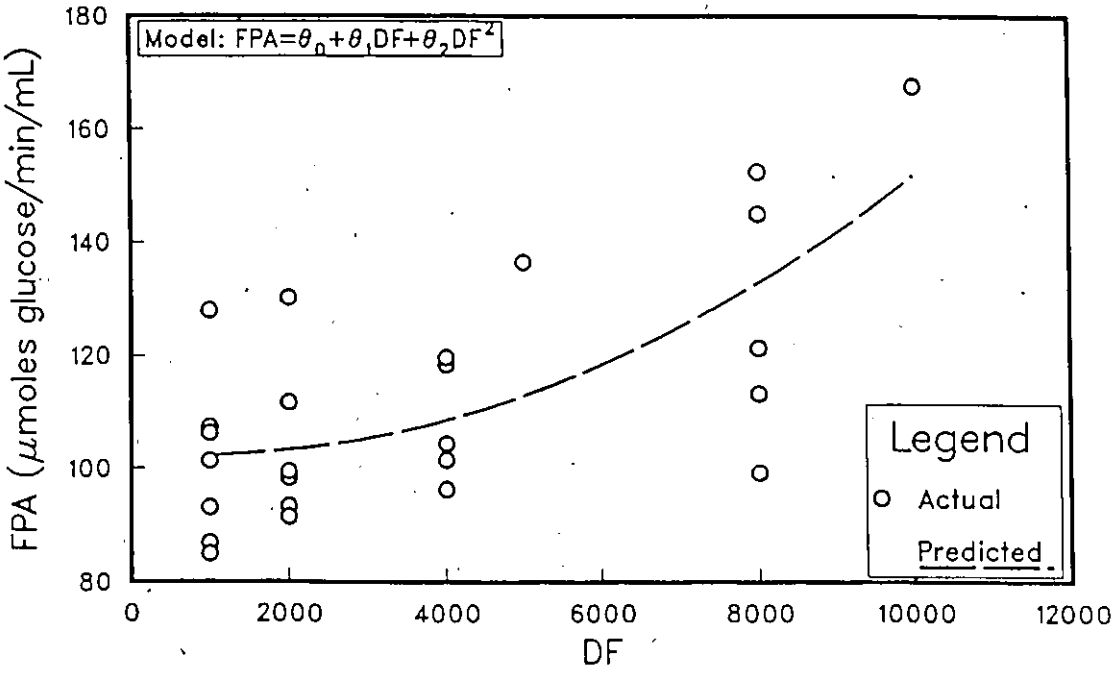


Figure 49: Actual and Predicted FPA vs DF for Celluclast

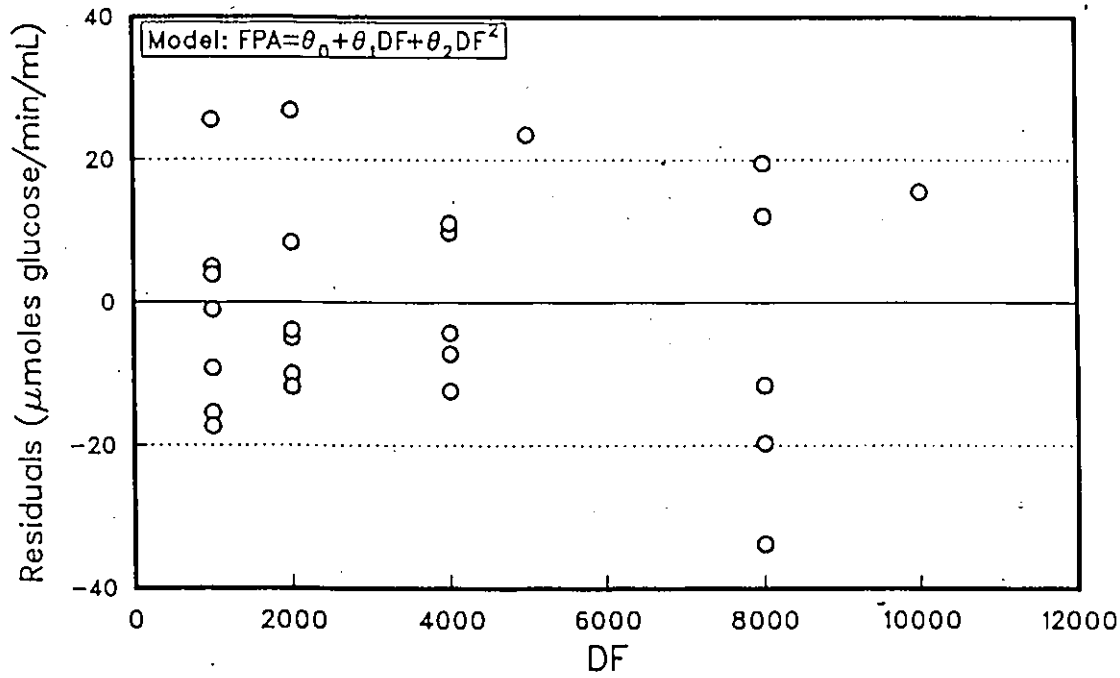


Figure 50: Residuals vs DF for Celluclast

2.3 FPA of Novozym

Since the G1 vs enzyme curve of Celluclast was shown to be heteroscedastic and since the transformed curve—activity vs DF—seems to contain more information, it was decided to also model FPA as a function of DF for the for the Novozym filter paper activity test. The units of FPA are μmoles gluc./min/mL in all modelling exercises. The ten data points and three empirical curves proposed to represent the data are shown in Figure 51.

As before, the full second order polynomial is rejected because of the lack of confidence in the LSE of the parameters.

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF} + \theta_2 \cdot \text{DF}^2}$ Novozym

$$\theta_0 = 2.19 \pm 2.99$$

$$\theta_1 = 0.0081 \pm 0.018$$

$$\theta_2 = 2.6 \cdot 10^{-6} \pm 20.7 \cdot 10^{-6}$$

The two other models to be fitted to the data are reduced from a full second degree polynomial.

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}}$ Novozym

$$\theta_0 = 1.88 \pm 1.58$$

$$\theta_1 = 0.0103 \pm 0.0044$$

Correlation matrix of the parameters:

$$\begin{bmatrix} 1 & -0.814 \\ -0.814 & 1 \end{bmatrix}$$

$$\text{SSR} = 12.7$$

$\underline{\text{FPA} = \theta_0 + \theta_1 \cdot \text{DF}^2}$ Novozym

$$\theta_0 = 3.41 \pm 1.20$$

$$\theta_1 = 1.15 \cdot 10^{-5} \pm 0.54 \cdot 10^{-5}$$

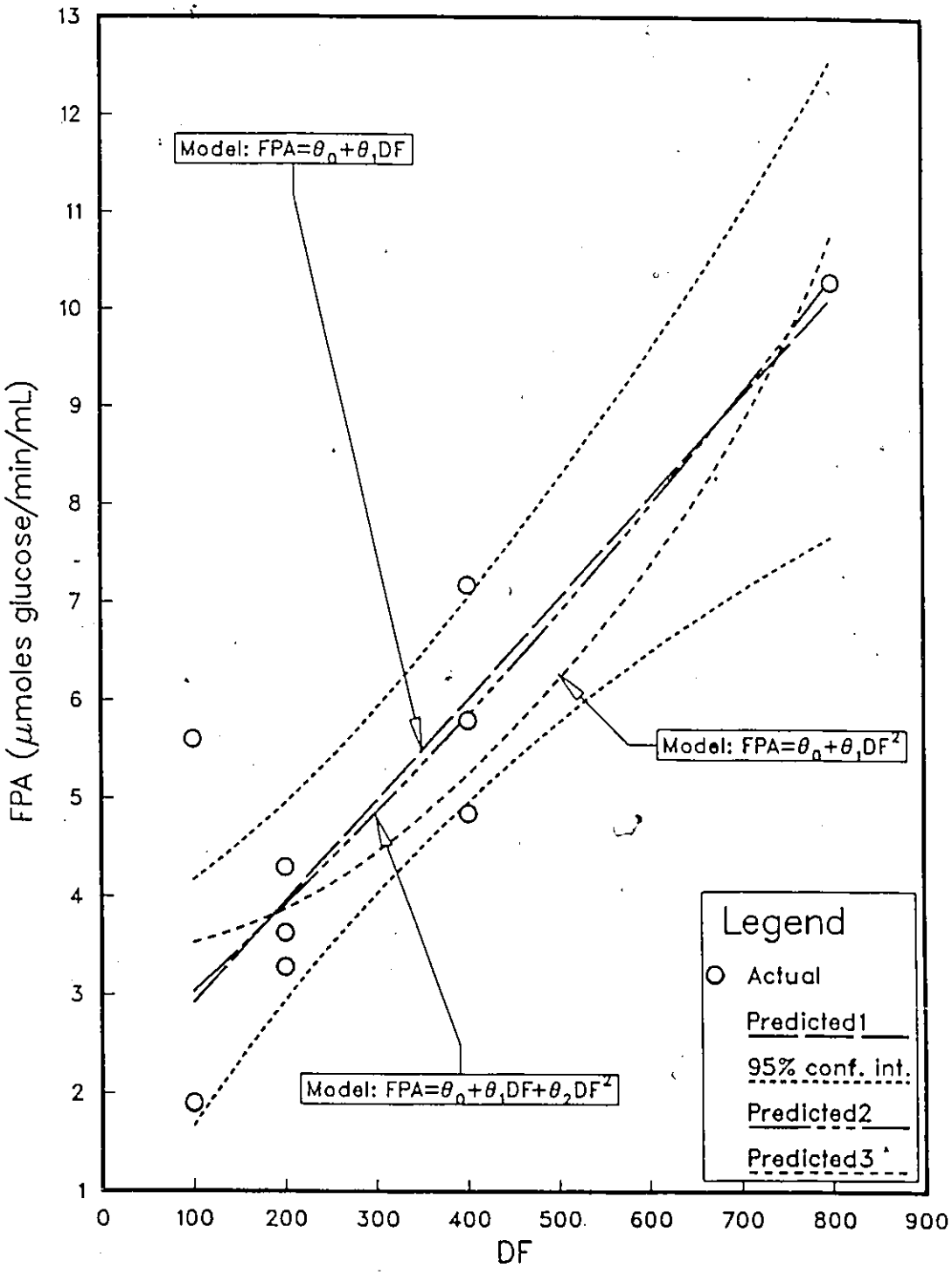


Figure 51: Actual and Predicted FPA vs DF for Novozym

Correlation matrix of the parameters:

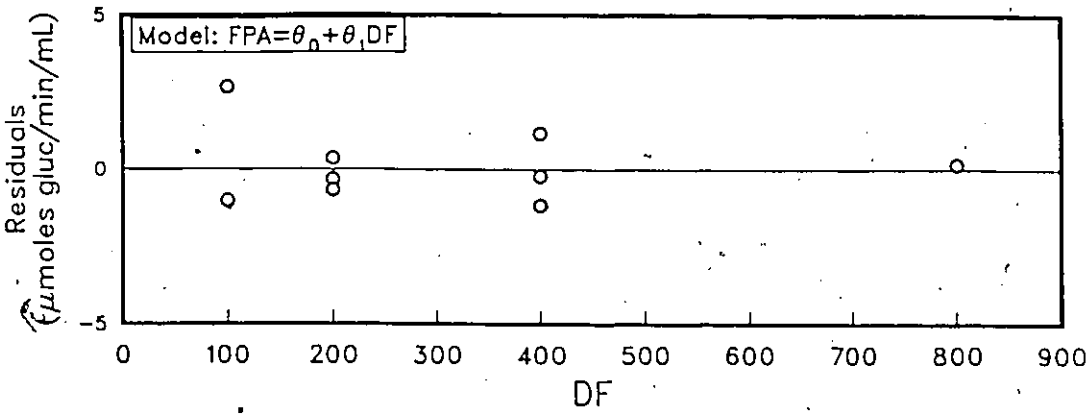
$$\begin{bmatrix} 1 & -0.573 \\ -0.573 & 1 \end{bmatrix}$$

$$\text{SSR} = 14.6$$

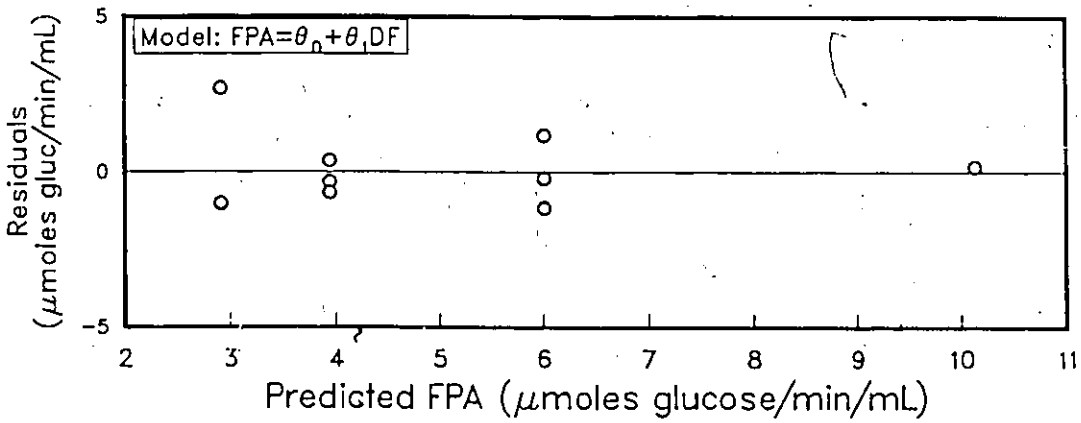
The residuals for all three models are shown in figures 52, 53 and 54. Note the residual for the run of June 29, 1985. The reason for plotting the residuals of the enzyme activity models as a function of run date is to detect downward trends with time. If the enzyme is losing a significant amount of activity in cold storage, then this trend would show up on the residuals vs run date plot. In particular for this, the FPA test of Novozym, there may be deactivation of the enzyme in storage. However, based on one point, that is not a proper conclusion. There is a lack of information relative to the deactivation in storage. FPA tests should have been conducted in the months of July, August and September 1985. Then deactivation, if present would have shown up as a trend in several data points. In view of this lack of information, it is assumed that the enzyme did not deactivate in storage.

Notwithstanding this lack of information on enzyme deactivation, a choice must be made between the two rival models. There is little difference in how both models represent the Novozym FPA data set. The straight line model was chosen because of its lower SSR. The same arguments that were put forward in the preceeding section, where the straight line was chosen to represent the FPA of Celluclast, also apply here. The 95% confidence interval for predicted values of the straight line model is shown in Figure 51.

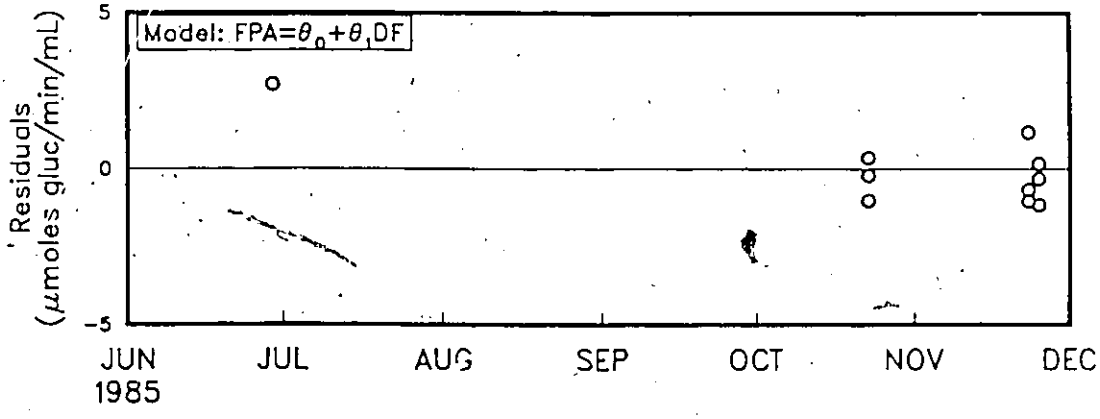
Thus the trend of the FPA increasing with higher DF was proven significant for both the Celluclast and Novozym enzyme solutions. Next, verify if the CBA is a function of the DF in the cellobiase activity test.



(a) Residuals vs DF

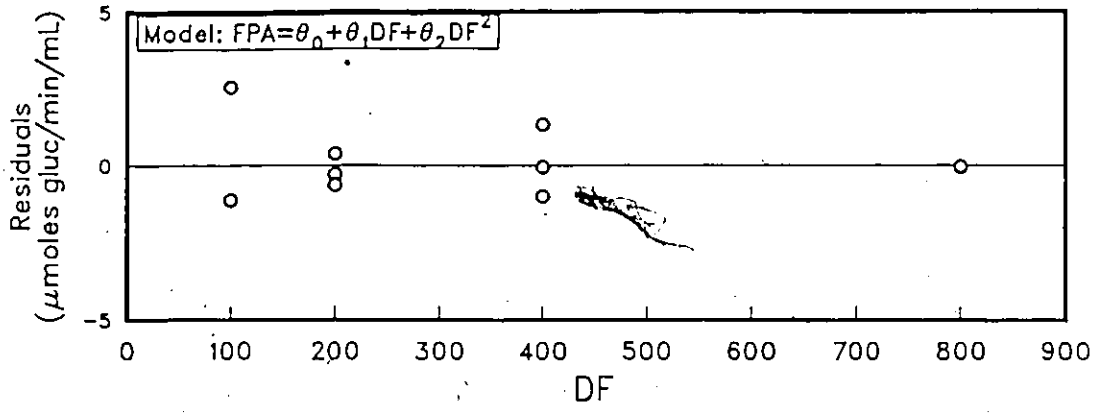


(b) Residuals vs Predicted FPA

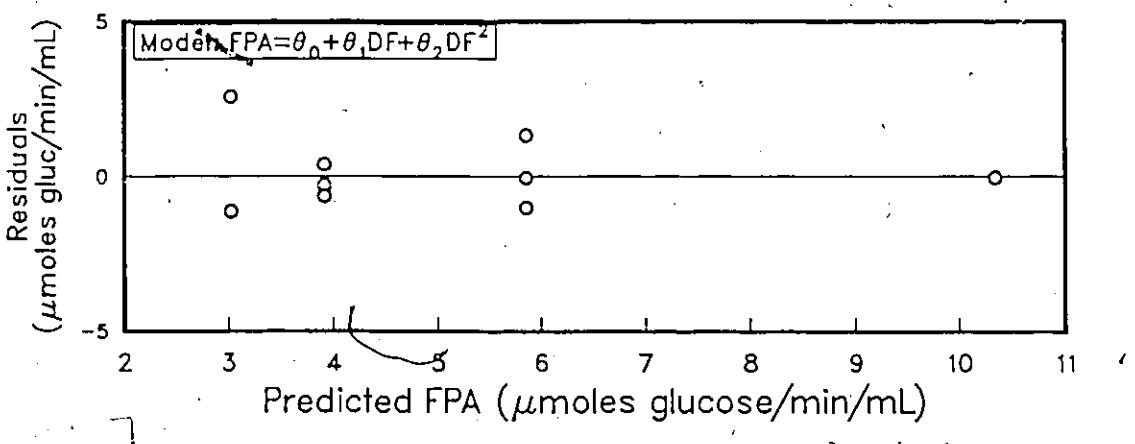


(c) Residuals vs Run Date

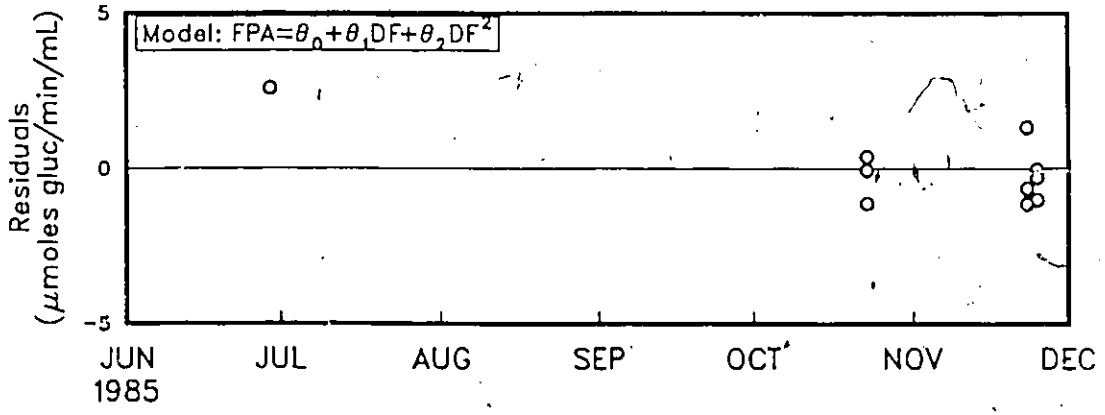
Figure 52: Novozym FPA Residual Plots
Model: FPA=$\theta_0+\theta_1$DF



(a) Residuals vs DF

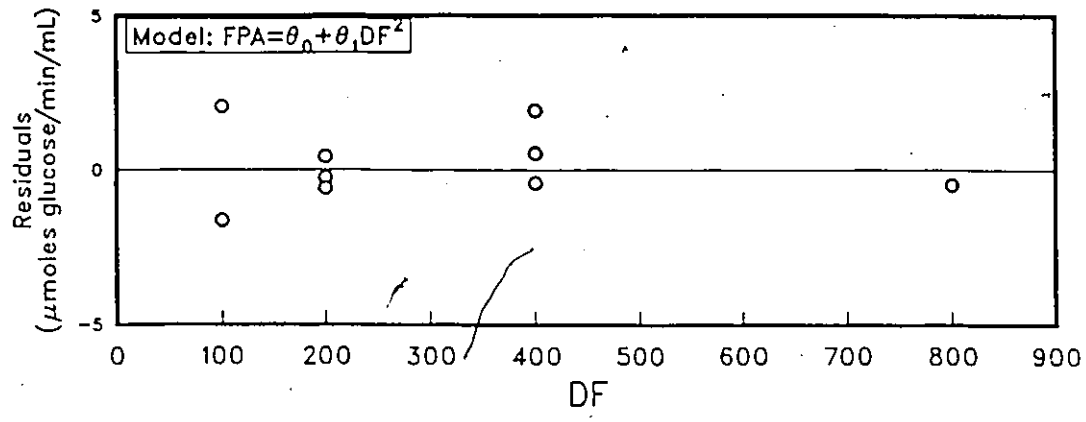


(b) Residuals vs Predicted FPA

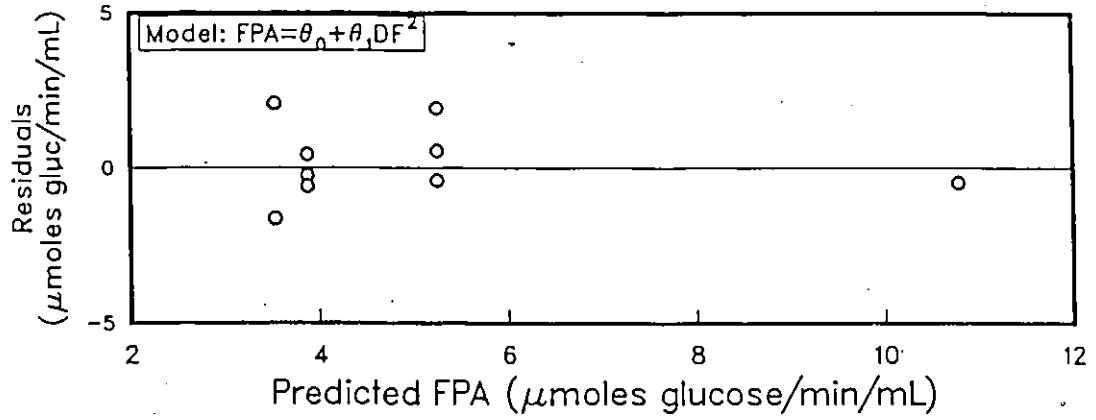


(c) Residuals vs Run Date

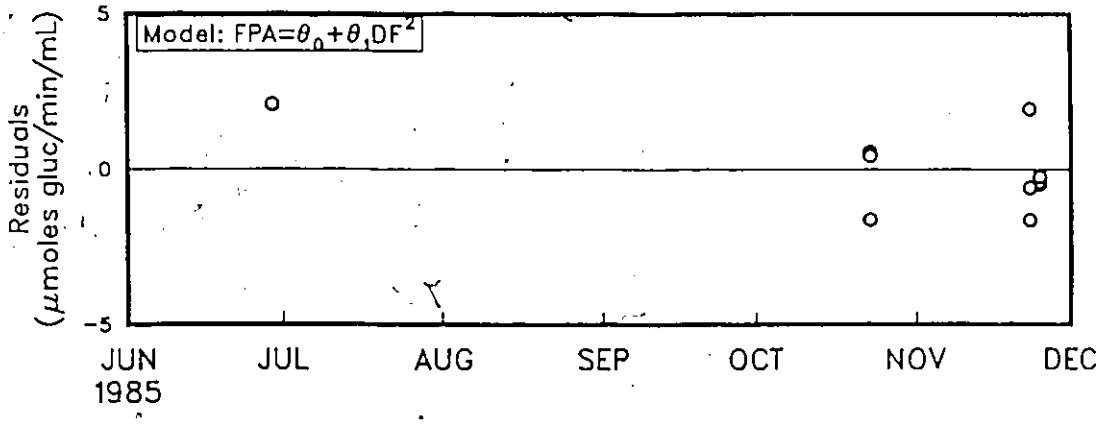
Figure 53: Novozym FPA Residual Plots
Model: FPA=$\theta_0+\theta_1 DF+\theta_2 DF^2$



(a) Residuals vs DF for Novozym



(b) Residuals vs Predicted FPA for Novozym



(c) Residuals vs Run Date for Novozym

Figure 54: Novozym FPA Residual Plots
Model: FPA=$\theta_0+\theta_1 DF^{-2}$

2.4 CBA of Novozym

The relationship between glucose, dilution factor and CBA is slightly different than for FPA. The independent variable, DF = 1/(enzyme vol. in mL), is unchanged from before. However the independent variable CBA is calculated as $\text{CBA} = G1 \cdot \text{DF}/5.4$.

The data collected for the purpose of testing the cellobiase activity of Novozym covers a wide range of the DF. Because of this wide range, the first model to be fitted is a mechanistic model (see Appendix 4).

$$\text{CBA} - \theta_0 - \theta_1 \cdot \text{DF} \cdot \ln((1 - 1.026 \cdot \text{CBA}/\text{DF})^2) = 0$$

There is a difficulty in modelling to obtain the LSE of the parameters. There are two roots to this equation i.e., at each value of the DF, two possible CBAs solve the mechanistic implicit equation. The actual collected data and the predicted CBA at the discrete test points for the above mechanistic model are plotted in Figure 55. Note that the non-linear SAS routine `PROC NLIN` was used to find the LSE of the parameters even though this is a linear, in the parameters, estimation problem.

The behavior of the predicted CBA for every datum point is inconsistent. At each point, the non-linear estimation routine picks a root which is not consistent with a particular locus of one set of solutions. We have not found the solution to this technical problem. Different starting parameter values for the non-linear program or different starting points for the trial and error routine did not result in a satisfactory fit. The residuals plotted as a function of the DF in Figure 56 clearly show a trend. For the sake of completeness, here are the particular parameter values for the plots in figures 55 and 56.

$\underline{\text{CBA} - \theta_0 - \theta_1 \cdot \text{DF} \cdot \ln((1 - 1.026 \cdot \text{CBA}/\text{DF})^2) = 0}$ Celluclast

$$\theta_0 = 1159 \pm 133$$

$$\theta_1 = 0.150 \pm 0.086$$

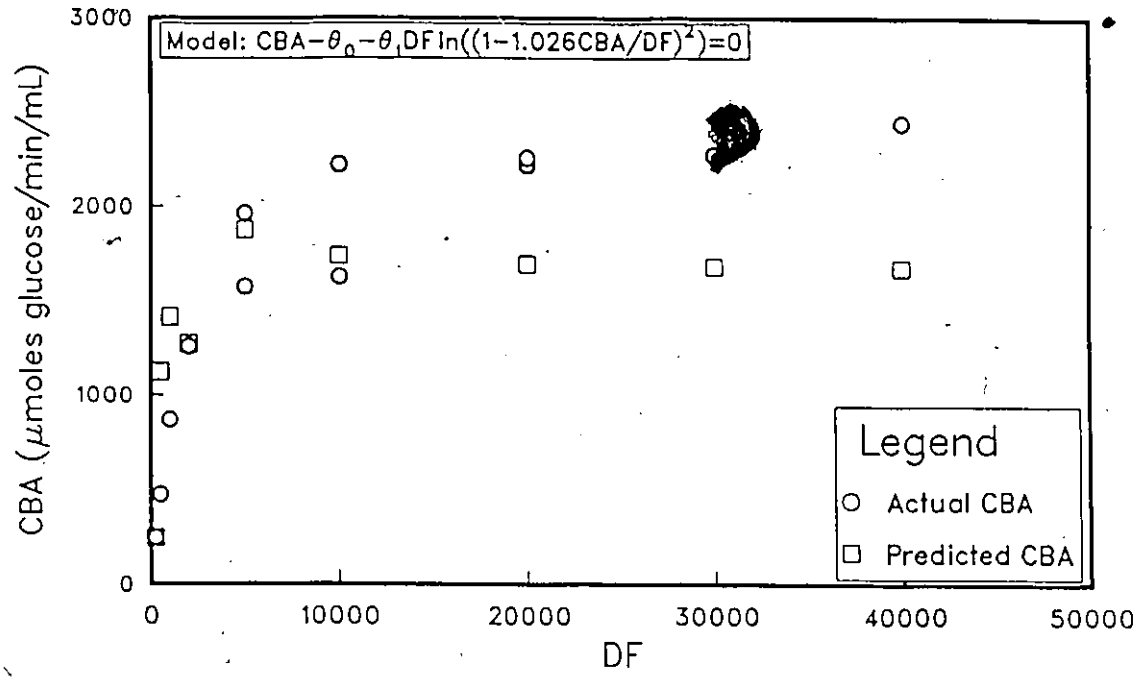


Figure 55: Actual and Predicted CBA vs DF for Novozym

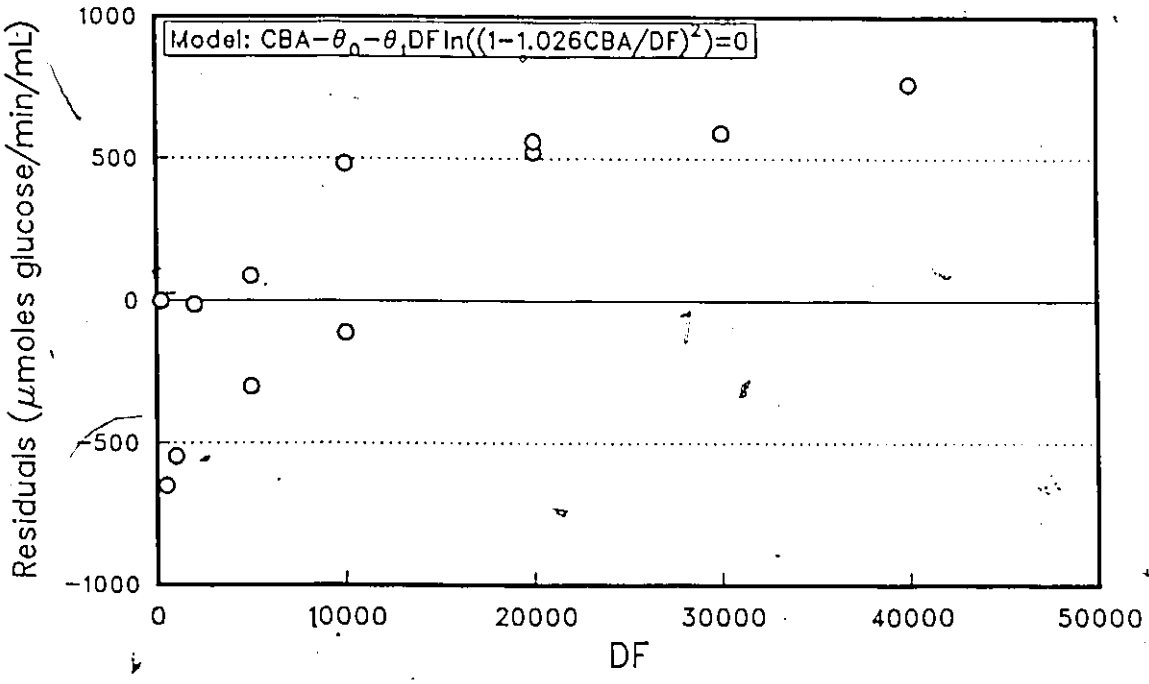


Figure 56: Residuals vs DF for Novozym

The correlation matrix of the parameters:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{SSR} = 2.13 \cdot 10^6$$

Two modifications were made to the above implicit mechanistic model: The first 'CBA' term was removed and the '1.026' term was allowed to float as a parameter. Recall that 1.026 is a theoretical value for $5.4/p_\infty$. Since θ_0 and θ_1 become one parameter after the elimination of the 'CBA' term, the new model is a two-parameter, non-linear model.

$$\text{CBA} = \theta_0 \cdot \text{DF} \cdot \left(1 - e^{-\theta_1/\text{DF}}\right)$$

There is no theoretical basis for the modifications made to the initial implicit model. This new model cannot then be considered mechanistic.

It is fortunate that the model represents the data so well; Figure 57 is a plot of the actual data and the behavior predicted by the model. Also shown are the 95% confidence limits for predicted values of the CBA. Figure 58 shows no important trends in the behavior of residuals with the dilution factor, predicted CBA or the date on which the test was performed. Here are the statistical diagnostics for the model deemed to well represent the data.

$\text{CBA} = \theta_0 \cdot \text{DF} \cdot \left(1 - e^{-\theta_1/\text{DF}}\right)$ Novozym

$$\theta_0 = 0.836 \pm 0.303$$

$$\theta_1 = 2826 \pm 1180$$

Correlation matrix of the parameters:

$$\begin{bmatrix} 1 & -0.983 \\ -0.983 & 1 \end{bmatrix}$$

$$\text{SSR} = 3.31 \cdot 10^5$$

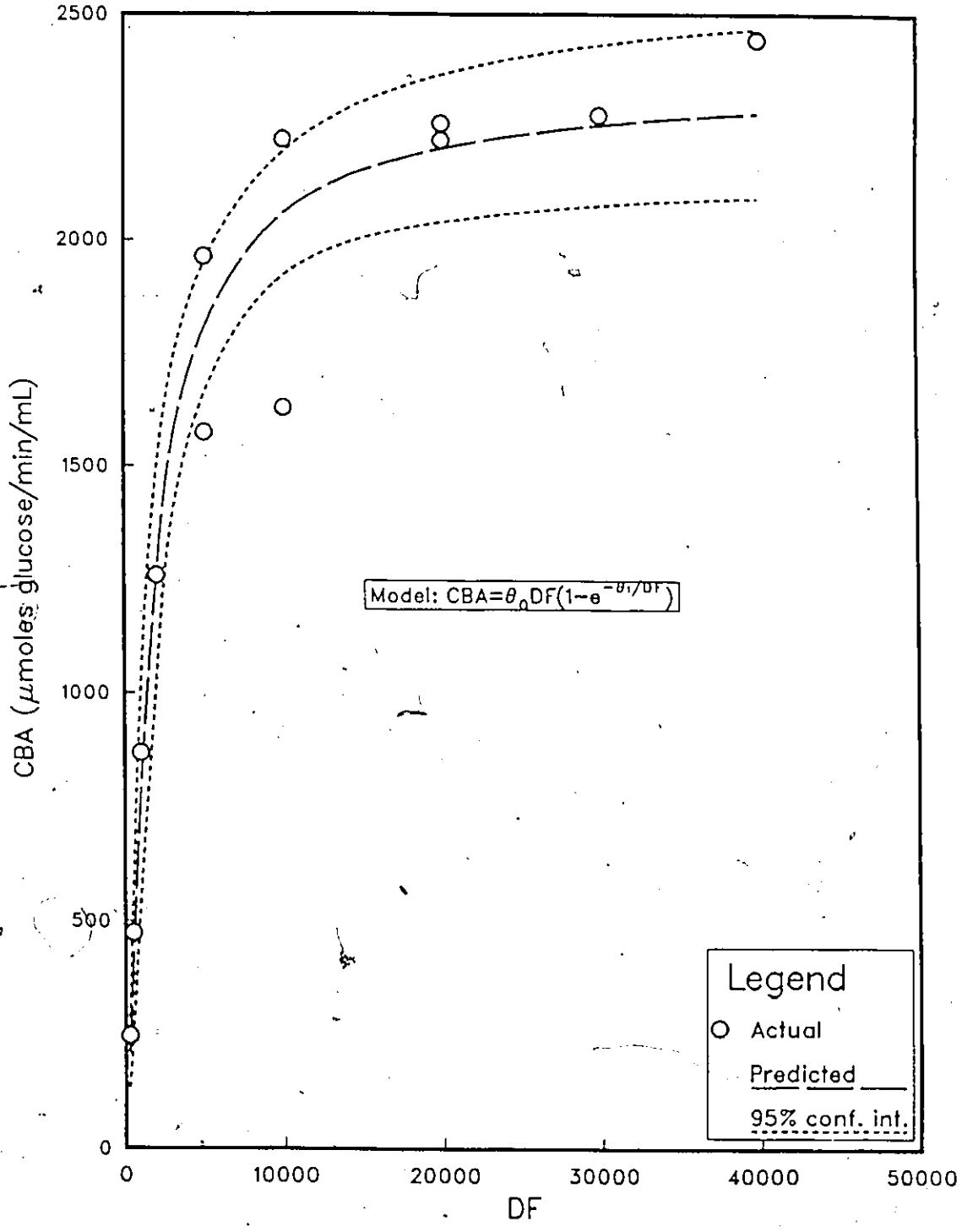
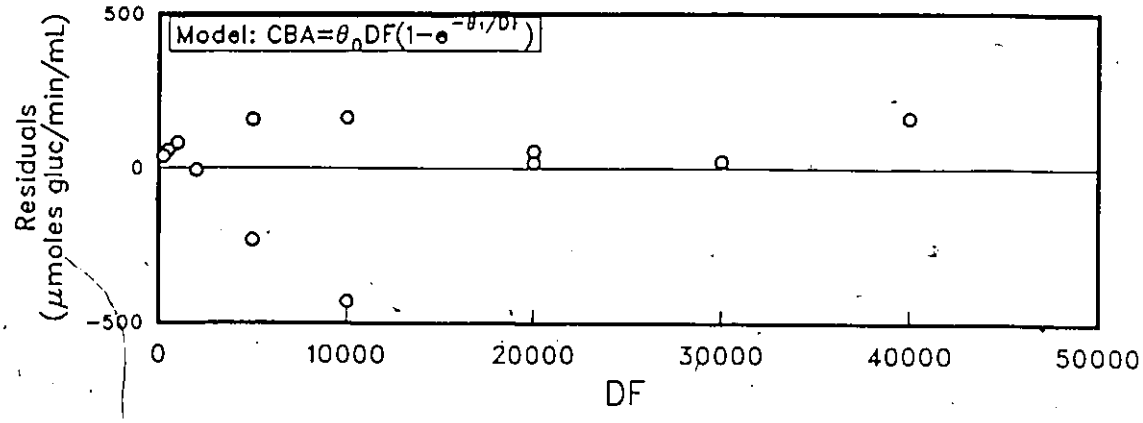
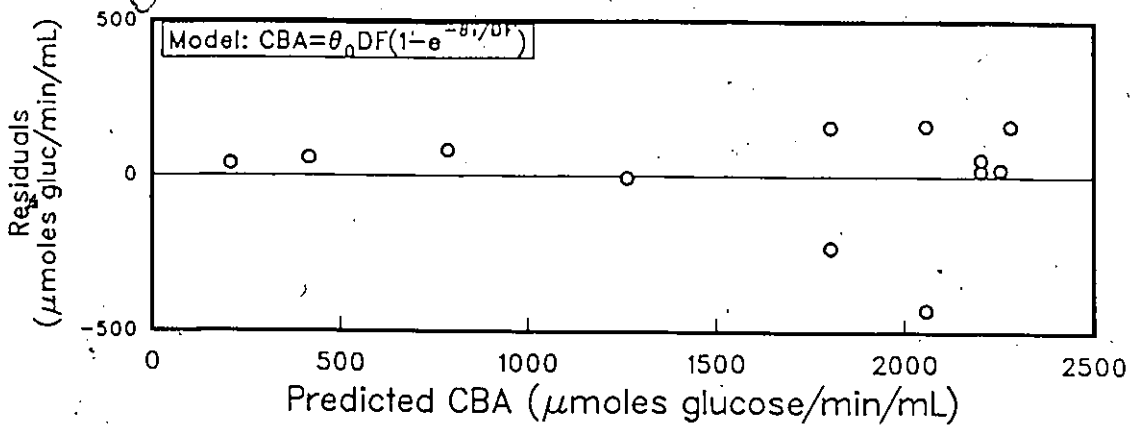


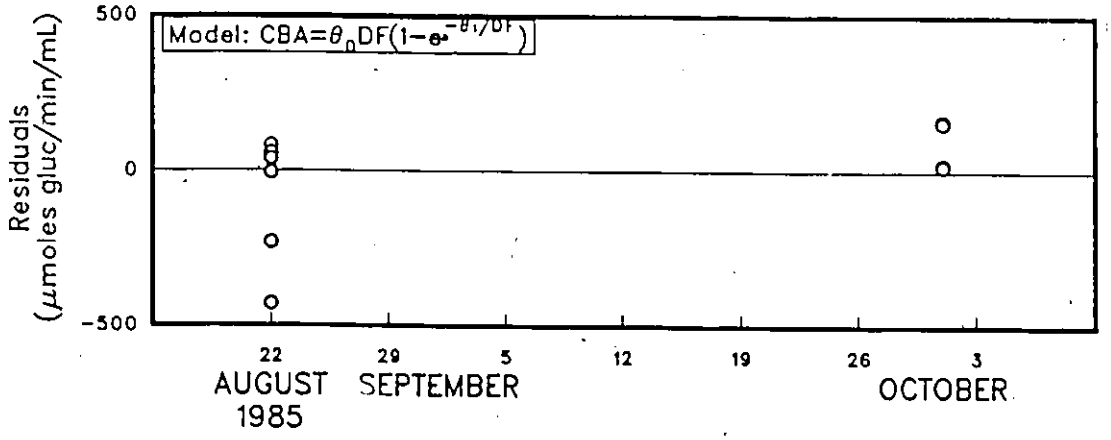
Figure 57: Actual and Predicted CBA vs DF for Novozym



(a) Residuals vs DF



(b) Residuals vs Predicted CBA



(c) Residuals vs Run Date

Figure 58: Novozym CBA Residual Plots
Model: CBA=θ_0DF(1−e$^{-\theta_1/DF}$)

2.5 CBA of Celluclast

The data from tests to determine the CBA activity of Celluclast can be seen on Figure 59. As well, the model CBA $= \theta_0 \cdot \mathrm{DF} \cdot (1 - e^{-\theta_1/\mathrm{DF}})$ predicted curve is shown. The parameter estimates are not significantly different from zero at a confidence level of 95%.

$$\theta_0 = 0.072 \pm 0.136$$

$$\theta_1 = 234 \pm 528$$

Hence this model which successfully represented Novozym CBA is rejected.

The next model to be fitted is the straight line. This linear relationship between enzyme activity and enzyme dilution was the chosen model for representing FPA of both commercial enzyme solutions. However, when the Celluclast CBA data are fitted to the model, the diagnostics show the model to be inadequate. One of the parameter confidence intervals includes 0.

$\underline{\mathrm{CBA} = \theta_0 + \theta_1 \cdot \mathrm{DF}}$ Celluclast

$$\theta_0 = 10.1 \pm 7.3$$

$$\theta_1 = -.00004 \pm 0.005$$

After unsuccessful attempts at fitting the whole data set, it was decided to discard 4 data points. For the runs where the dilution factor is 1000 or greater, the amount of glucose obtained in this CBA test is so low that it is considered to be below the spectrophotometer's ability to measure. The highest measured glucose concentration for the four rejected points is $0.15\,g/\mathrm{L}$. In fact at a DF of 5000, no glucose was measured. This makes the test results erroneous. There is further justification in dropping the 4 data points at the high DF; The region of interest to us is near a product concentration of $0.667\,g/\mathrm{L}$. That product concentration, for this set of data will occur at a DF less than 500.

Figure 63 shows the smaller range of data on an expanded scale. As well the straight line model was refitted to the reduced data set.

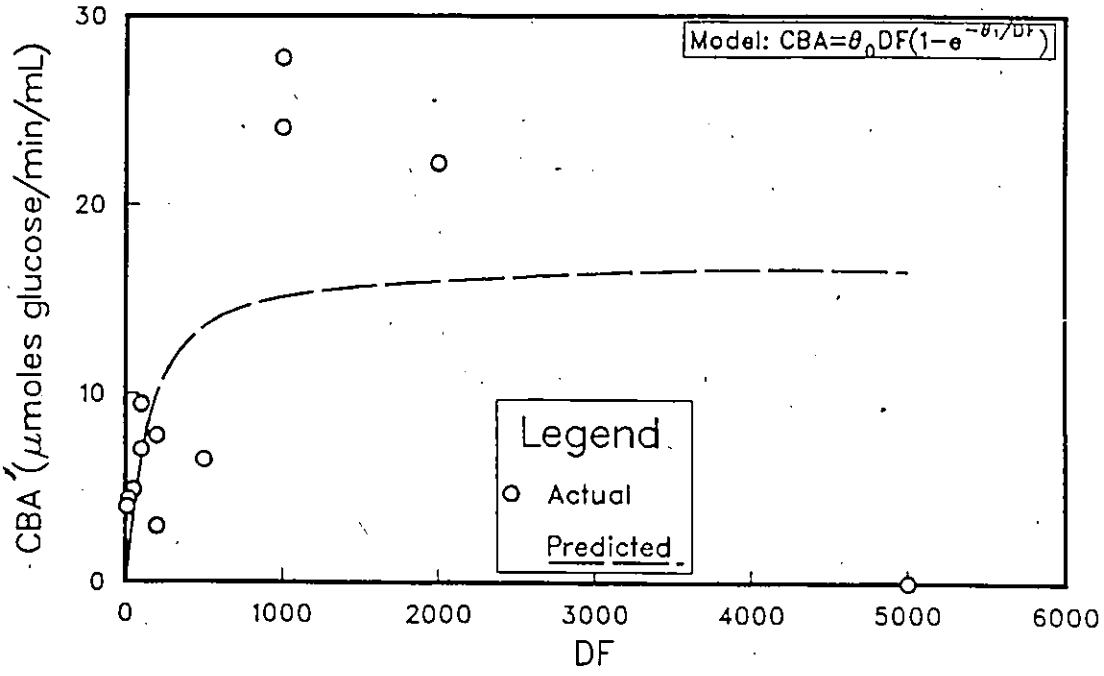


Figure 59: Actual and Predicted CBA vs DF for Celluclast

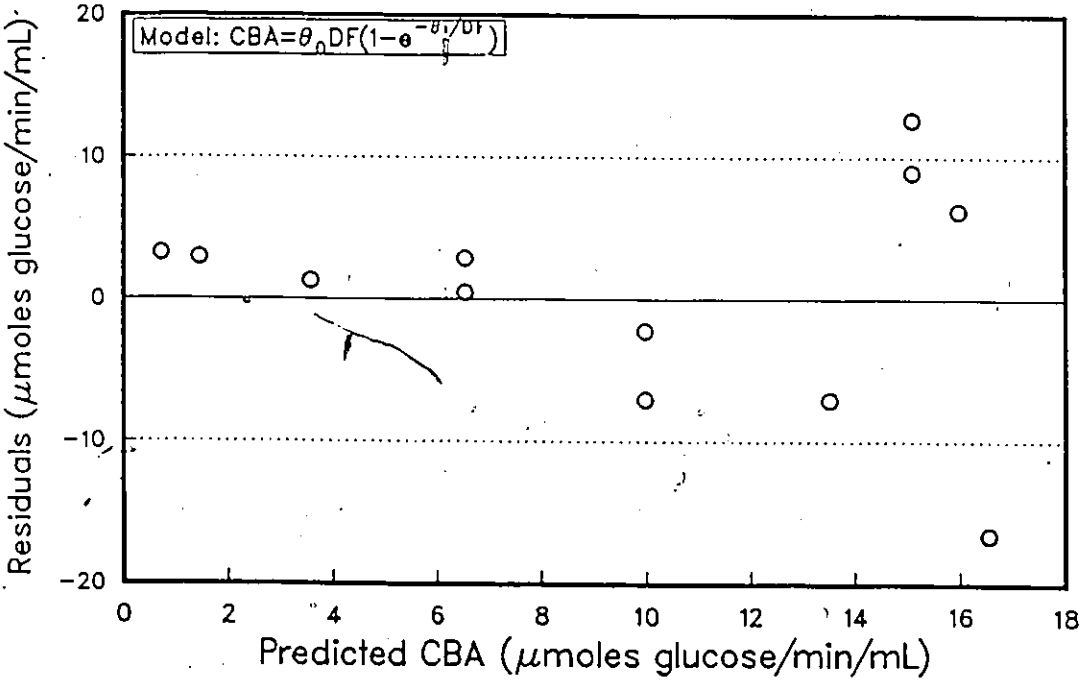


Figure 60: Residuals vs Predicted CBA for Celluclast

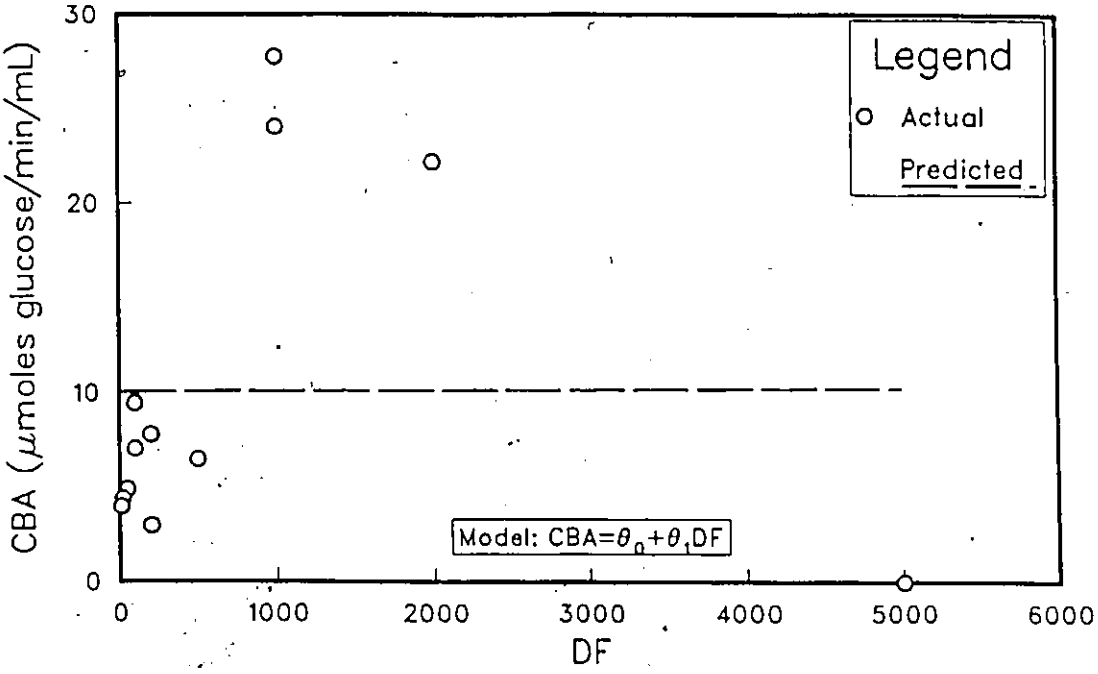


Figure 61: Actual and Predicted CBA vs DF for Celluclast

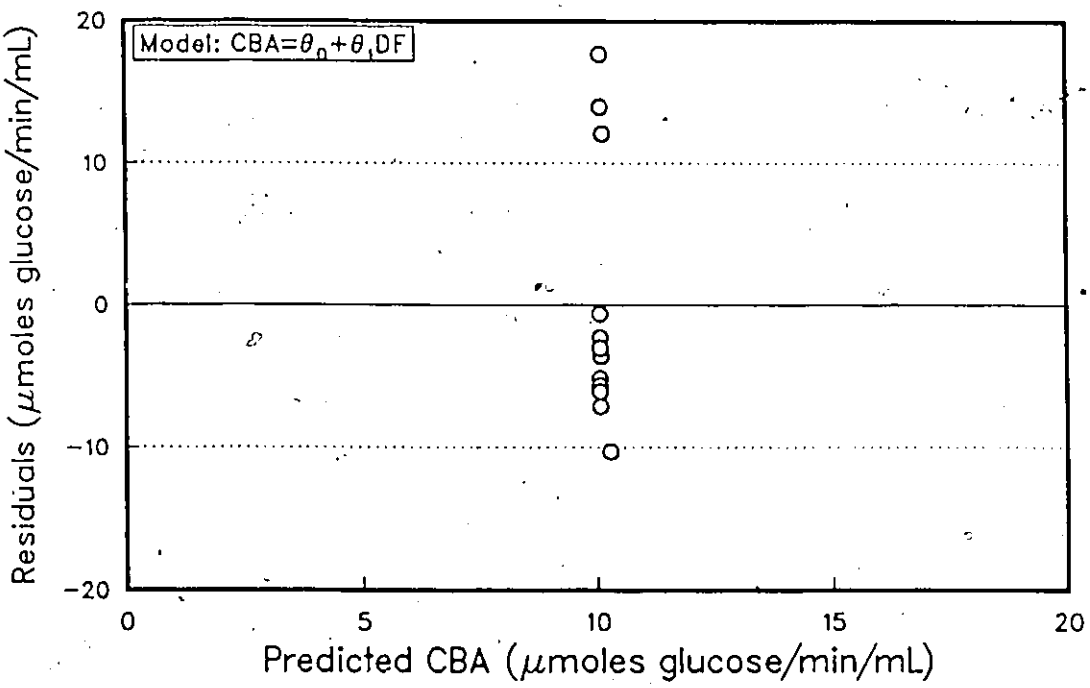


Figure 62: Residuals vs Predicted CBA for Celluclast

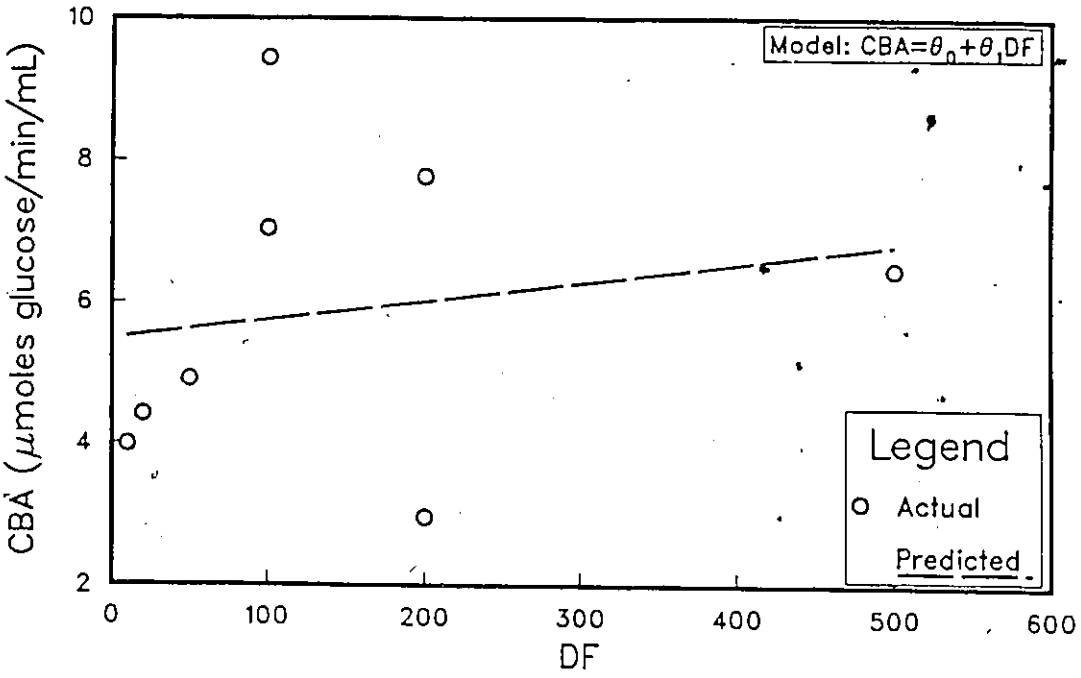


Figure 63: Actual and Predicted CBA vs DF for Celluclast Fitted for Reduced Data Set

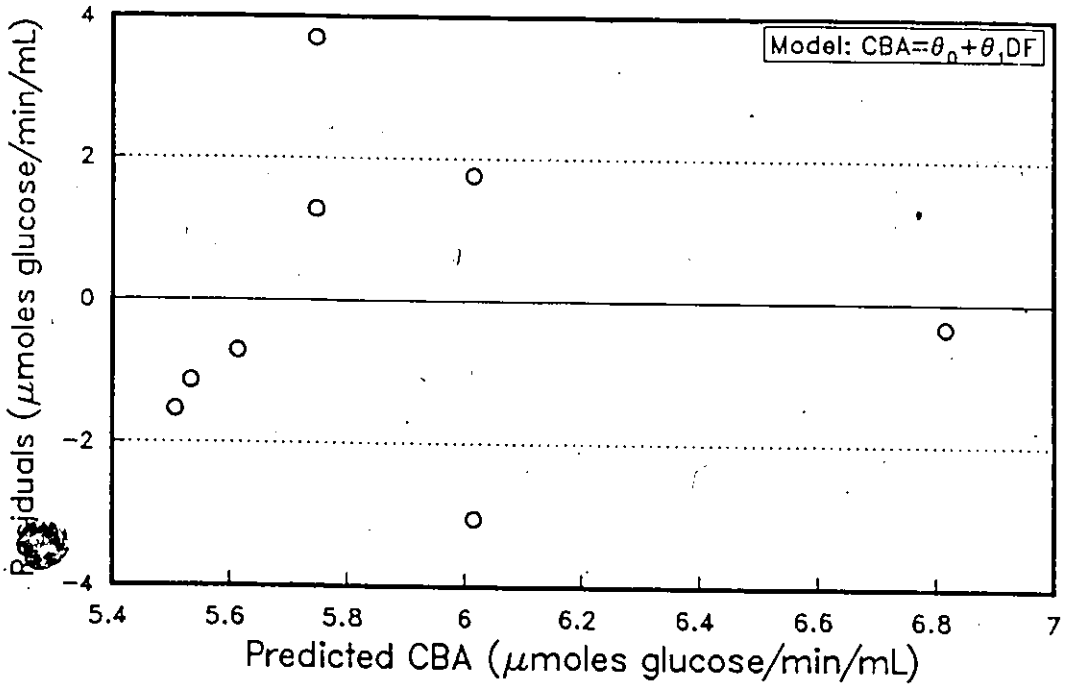


Figure 64: Residuals vs Predicted CBA for Celluclast Fitted for Reduced Data Set

$\underline{\text{CBA} = \theta_0 + \theta_1 \cdot \text{DF}}$ Celluclast

$$\theta_0 = 5.48 \pm 2.80$$

$$\theta_1 = 0.0026 \pm 0.0134$$

The correlation matrix of the parameters:

$$\begin{bmatrix} 1 & -0.702 \\ -0.702 & 1 \end{bmatrix}$$

$$\text{SSR} = 32.0$$

This model is rejected because at a 95% confidence level, we cannot be sure that θ_1 is not 0.

All the models fitted to the Celluclast data have failed to give an adequate representation of the relationship between CBA and DF. There are no trends in the data that suggest any relationship can be found. Thus the best estimate of the CBA of Celluclast here will be obtained by taking an average CBA. The average CBA for the 8 data points at a low DF is 5.88 μmoles gluc/min/mL.

Taking the mean of the CBA data is equivalent to saying that FPA = θ_0 is the best model in view of the lack of trend. See Figure 65 for a comparison of the actual data, the mean CBA and the 95% confidence intervals for the predicted value of the mean CBA.

$\underline{\text{CBA} = \theta_0}$ Celluclast

$$\theta_0 = 5.88 \pm 1.82$$

$$\text{SSR} = 33.2$$

See Figure 66 for the residual plots. There are no evident trends of the residuals with DF, predicted CBA or run date.

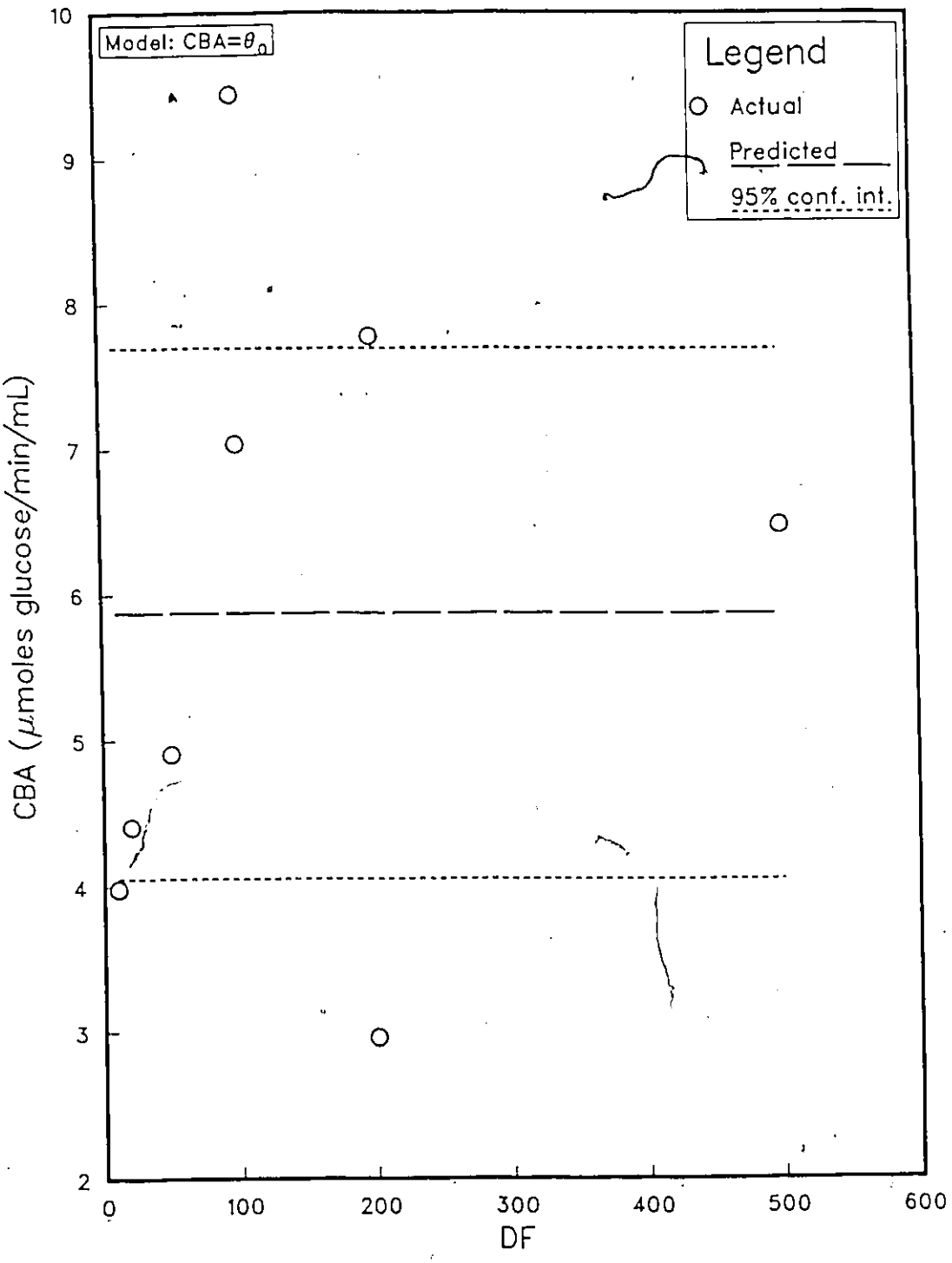
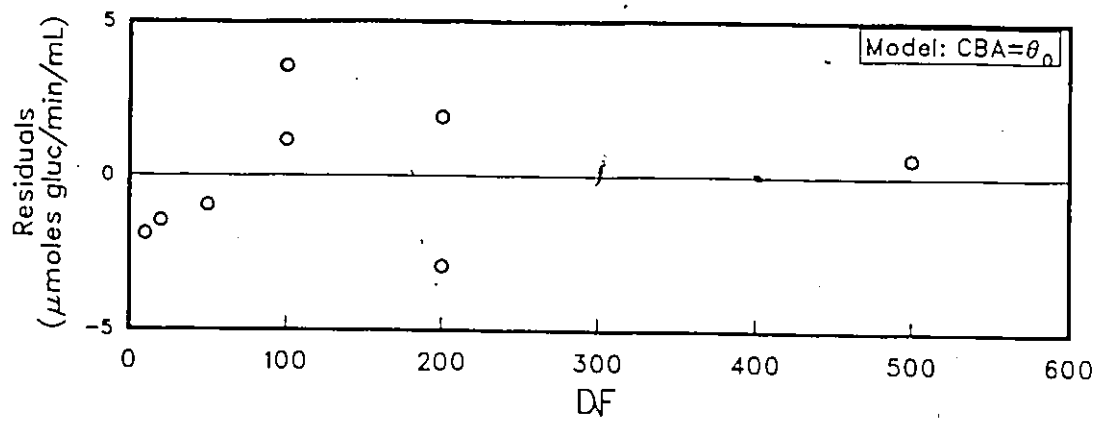
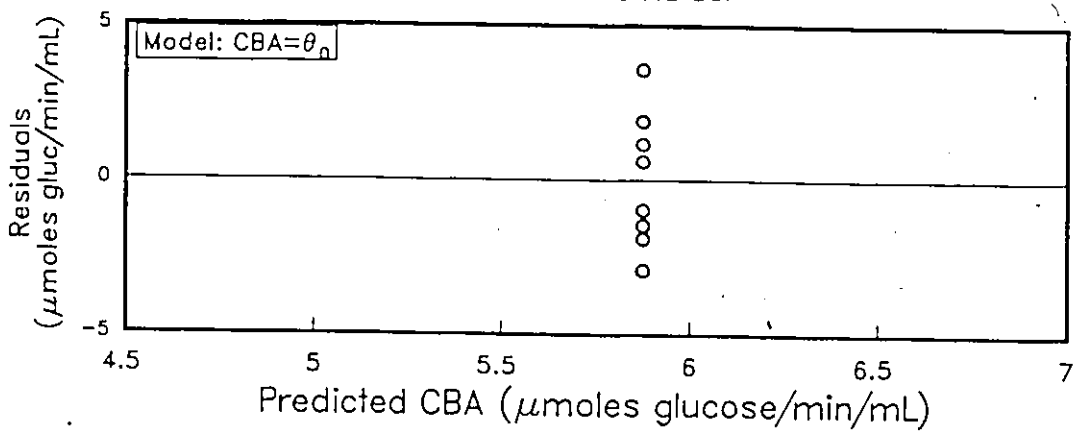


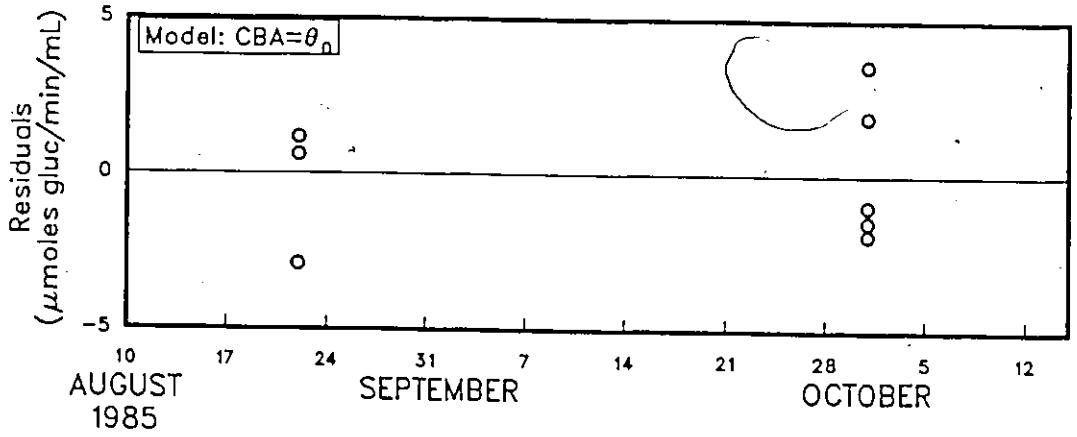
Figure 65: Actual and Predicted CBA vs DF for Celluclast Fitted for Reduced Data Set



(a) Residuals vs DF
Fitted for Reduced Data Set



(b) Residuals vs Predicted CBA
Fitted for Reduced Data Set



(c) Residuals vs Run Date
Fitted for Reduced Data Set

Figure 66: Celluclast CBA Residual Plots
Model: CBA=θ_0

2.6 Summary

In summary, equations have been used to model the behavior of the enzyme activity—FPA and CBA—as a function of the enzyme dilution factor, DF. More information is extracted from the activity vs DF plot than from the glucose vs enzyme plot. This G1 vs enz. plot is the plot normally used when calculating the activity of an enzyme mixture. Any procedure should conduct the activity tests at relatively high dilution factors. If the DF is high enough, then the dependency of FPA or CBA on DF is not evident on a G1 vs enz. plot. Then the G1 vs enz. plot becomes a straight line through the origin. The procedure in use in enzyme activity measurements then consists in interpolating or using regression to obtain the activity at standard conditions.

With our procedure, all of the information from activity tests is used. Glucose measurements that are in the non-linear portion of the G1–enz. curve are included in the measurement of $\widehat{\text{FPA}}$ or $\widehat{\text{CBA}}$.

It is necessary to distinguish between FPA and $\widehat{\text{FPA}}$. FPA is the average rate at which glucose is produced for one run during the test. Its units are μmoles/min/mL. At each experimental run condition, G1 and enzyme volume are the measured data. These data are transformed into FPA, the new dependent variable and DF, the new independent variable using the following formulae:

$$\text{DF} = 1/(\text{enzyme vol. in mL})$$

$$\text{FPA} = G1 \cdot \text{DF}/10.8$$

In the case of cellobiase activity, the following formulae are used:

$$\text{DF} = 1/(\text{enzyme vol. in mL})$$

$$\text{CBA} = G1 \cdot \text{DF}/5.4$$

Apply a strict modelling procedure to the plots of activity as a function of DF. The proposed models to be tested may be empirical, semi-empirical or mechanistic. If the data was collected at such a high dilution factor that the activity is independent

of DF then the appropriate model is FPA (or CBA)= θ_0. The modelling procedure would then reduce to the method presently used in cellulase enzyme kinetics.

In our case, four models were needed, one for each of two enzymes E_x and E_1, in each of two commercial enzyme solutions, Celluclast and Novozym. The concentration of E_x is estimated by $\widehat{\text{FPA}}$. The concentration of E_1 is estimated by $\widehat{\text{CBA}}$. Before going on to calculate $\widehat{\text{FPA}}$ and $\widehat{\text{CBA}}$, the FPA and CBA behaviors as a function of DF are described by a model.

Celluclast is a purified cellulase enzyme. Novozym is a purified cellobiase enzyme. The filter paper activity of Celluclast is high and the cellobiase activity of Novozym is also high. However Celluclast contains a trace of cellobiase and Novozym contains a trace endoglucanase. The actual trace amounts of enzymes in the commercial enzyme solutions are important in the overall CMC hydrolysis modelling. The models show the vast difference in the concentration of either E_x or E_1 in the two commercial solutions. Here are the four models deemed to best represent the activity data:

Celluclast FPA=94.3+0.0048·DF

Novozym FPA=1.88+0.0103·DF

Novozym CBA=0.836·DF·$(1 - e^{-2826/\text{DF}})$

Celluclast CBA=5.88

$\widehat{\text{FPA}}$ and $\widehat{\text{CBA}}$ are to be calculated at the conditions where 0.667 g/L of glucose is produced.

This appendix demonstrates the model discrimination procedure applied to our four data sets of enzyme activities. See the Results section of this thesis for the calculation of $\widehat{\text{FPA}}$ and $\widehat{\text{CBA}}$ from the FPA and CBA models.

Appendix 3

RAW DATA

Table 18. *Raw Hydrolysis Data: Part 1*

Run #	Date damoyr	DS theoretical	DP	Buffer type	Buffer M	Ω rpm	Volume mL	CLCLST µL	NOVOZM µL
H0507A	5JUL84	7	L	CIT	0.09	100	95.0	100.0	0.0
H0507B	5JUL84	7	L	CIT	0.09	100	95.0	100.0	50.0
H0507C	5JUL84	7	L	CIT	0.09	100	95.0	10.0	0.0
H29081	29AUG84	7	L	CIT	0.10	200	146.0	200.0	75.0
H29081	29AUG84	7	L	CIT	0.10	100	146.0	200.0	75.0
H29081	29AUG84	7	L	CIT	0.10	100	146.0	200.0	75.0
H29081	29AUG84	7	L	CIT	0.10	100	146.0	200.0	75.0
H29081	29AUG84	7	L	CIT	0.10	100	146.0	200.0	75.0
H29081	29AUG84	7	L	CIT	0.10	100	146.0	200.0	75.0
H29082	29AUG84	7	L	CIT	0.10	200	146.0	200.0	0.0
H29082	29AUG84	7	L	CIT	0.10	100	146.0	200.0	0.0
H29082	29AUG84	7	L	CIT	0.10	100	146.0	200.0	0.0
H29082	29AUG84	7	L	CIT	0.10	100	146.0	200.0	0.0
H29082	29AUG84	7	L	CIT	0.10	100	146.0	200.0	0.0
H29082	29AUG84	7	L	CIT	0.10	100	146.0	200.0	0.0
H29083	29AUG84	7	L	CIT	0.10	200	146.0	100.0	0.0
H29083	29AUG84	7	L	CIT	0.10	100	146.0	100.0	0.0
H29083	29AUG84	7	L	CIT	0.10	100	146.0	100.0	0.0
H29083	29AUG84	7	L	CIT	0.10	100	146.0	100.0	0.0
H29083	29AUG84	7	L	CIT	0.10	100	146.0	100.0	0.0
H29083	29AUG84	7	L	CIT	0.10	100	146.0	100.0	0.0
H29084	29AUG84	7	L	CIT	0.10	200	147.0	200.0	75.0
H29084	29AUG84	7	L	CIT	0.10	100	147.0	200.0	75.0
H29084	29AUG84	7	L	CIT	0.10	100	147.0	200.0	75.0
H29084	29AUG84	7	L	CIT	0.10	100	147.0	200.0	75.0
H29084	29AUG84	7	L	CIT	0.10	100	147.0	200.0	75.0
H29084	29AUG84	7	L	CIT	0.10	100	147.0	200.0	75.0
H29085	29AUG84	7	L	CIT	0.10	200	147.0	200.0	0.0
H29085	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29085	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29085	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29085	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29085	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29086	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0

Table 18. *(continued)*

Run #	Date da mo yr	DS theoretical	DP	Buffer type	M	Ω rpm	Volume mL	CLCLST μL	NOVOZM μL
H29086	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29086	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29086	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29086	29AUG84	7	L	CIT	0.10	100	147.0	200.0	0.0
H29087	29AUG84	7	L	CIT	0.10	200	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29087	29AUG84	7	L	CIT	0.10	100	147.0	400.0	0.0
H29088	29AUG84	7	L	CIT	0.10	100	147.0	0.0	75.0
H29088	29AUG84	7	L	CIT	0.10	100	147.0	0.0	75.0
H29088	29AUG84	7	L	CIT	0.10	100	147.0	0.0	75.0
H21091	21SEP84	7	L	CIT	0.0816	200	122.6	0.0	1000.0
H21091	21SEP84	7	L	CIT	0.0816	200	122.6	0.0	1000.0
H21091	21SEP84	7	L	CIT	0.0816	200	122.6	0.0	1000.0
H14031	14MAR85	7	L	CIT	0.05	120	96.0	25.0	0.0
H14032	14MAR85	7	L	CIT	0.05	120	96.0	25.0	0.0
H14033	14MAR85	7	L	CIT	0.05	120	96.0	25.0	0.0
H14034	14MAR85	7	L	CIT	0.05	120	96.0	25.0	0.0
H14035	14MAR85	7	L	CIT	0.05	120	96.0	25.0	0.0
H14036	14MAR85	7	L	CIT	0.05	120	96.0	200.0	0.0
H14037	14MAR85	7	L	CIT	0.05	120	96.0	200.0	0.0
H14038	14MAR85	7	L	CIT	0.05	120	96.0	200.0	0.0
H14039	14MAR85	7	L	CIT	0.05	120	96.0	200.0	0.0
H21031	21MAR85	7	L	CIT	0.097	100	103.0	33.33	0.0
H21032	21MAR85	7	L	CIT	0.097	100	103.0	33.33	0.0
H21033	21MAR85	7	L	CIT	0.097	100	103.0	33.33	0.0
H21034	21MAR85	7	L	CIT	0.097	100	103.0	33.33	0.0
H21035	21MAR85	7	L	CIT	0.097	100	103.0	0.0	33.33
H21036	21MAR85	7	L	CIT	0.097	100	103.0	0.0	33.33
H21037	21MAR85	7	L	CIT	0.097	100	103.0	0.0	33.33
H21038	21MAR85	7	L	CIT	0.097	100	103.0	0.0	33.33
H12041	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12042	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12043	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12044	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12045	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12046	12APR85	7	L	CIT	0.0972	105	95.5	200.0	0.0
H12047	12APR85	7	L	CIT	0.0972	105	95.5	200.0	0.0

Table 18. *(continued)*

Run #	Date da mo yr	DS theoretical	DP	Buffer type	Buffer M	Ω rpm	Volume mL	CLCLST µL	NOVOZM µL
H12048	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H12049	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13041	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13042	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13043	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13044	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13045	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13046	12APR85	7	L	CIT	0.0972	105	95.0	200.0	0.0
H13047	12APR85	7	L	CIT	0.0972	105	95.0	400.0	0.0
H13048	12APR85	7	L	CIT	0.0972	105	95.0	400.0	0.0
H13049	12APR85	7	L	CIT	0.0972	105	95.0	400.0	0.0
C13041	12APR85	0	0	CIT	0.10	105	100.0	200.0	0.0
C13042	12APR85	0	0	CIT	0.10	105	100.0	200.0	0.0
H09501	9MAY85	7	L	ACE	0.0633	100	95.4	400.0	0.0
H09502	9MAY85	7	L	ACE	0.0633	100	95.4	400.0	0.0
H09503	9MAY85	7	L	ACE	0.0633	100	95.4	400.0	0.0
H09504	9MAY85	7	L	ACE	0.0633	100	95.4	400.0	0.0
H09505	9MAY85	7	L	ACE	0.0633	100	95.4	400.0	0.0
H09506	9MAY85	7	L	ACE	0.0633	100	95.4	200.0	0.0
H09507	9MAY85	7	L	ACE	0.0633	100	95.4	200.0	0.0
H09508	9MAY85	7	L	ACE	0.0633	100	95.4	200.0	0.0
H09509	9MAY85	7	L	ACE	0.0633	100	95.4	200.0	0.0
H09510	9MAY85	7	L	ACE	0.0633	100	95.4	200.0	0.0
H09511	9MAY85	7	L	ACE	0.0633	100	95.4	0.0	200.0
H09512	9MAY85	7	L	ACE	0.0633	100	95.4	0.0	200.0
H09513	9MAY85	7	L	ACE	0.0633	100	95.4	0.0	200.0
H09514	9MAY85	7	L	ACE	0.0633	100	95.4	0.0	200.0
H09515	9MAY85	7	L	ACE	0.0633	100	95.4	0.0	200.0
H01701	28JUN85	7	L	ACE	0.05	120	96.1	200.0	0.0
H01702	28JUN85	7	L	ACE	0.05	120	96.1	200.0	0.0
H01703	28JUN85	7	L	ACE	0.05	120	96.1	200.0	0.0
H01704	28JUN85	7	L	ACE	0.05	120	96.1	200.0	0.0
H01705	28JUN85	7	L	ACE	0.05	120	96.1	200.0	0.0
H01706	28JUN85	7	L	ACE	0.05	120	96.1	20.0	0.0
H01707	28JUN85	7	L	ACE	0.05	120	96.1	20.0	0.0
H01708	28JUN85	7	L	ACE	0.05	120	96.1	20.0	0.0
H01709	28JUN85	7	L	ACE	0.05	120	96.1	20.0	0.0
H01710	28JUN85	7	L	ACE	0.05	120	96.1	20.0	0.0
H01711	28JUN85	7	L	ACE	0.05	120	96.1	2.0	0.0
H01712	28JUN85	7	L	ACE	0.05	120	96.1	2.0	0.0
H01713	28JUN85	7	L	ACE	0.05	120	96.1	2.0	0.0

Table 18. *(continued)*

Run #	Date da mo yr	DS theoretical	DP	Buffer type	Buffer M	Ω rpm	Volume mL	CLCLST μL	NOVOZM μL
H01714	28JUN85	7	L	ACE	0.05	120	96.1	2.0	0.0
H01715	28JUN85	7	L	ACE	0.05	120	96.1	2.0	0.0
H01716	28JUN85	7	L	ACE	0.05	120	96.1	0.2	0.0
H01717	28JUN85	7	L	ACE	0.05	120	96.1	0.2	0.0
H01718	28JUN85	7	L	ACE	0.05	120	96.1	0.2	0.0
H20701	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20702	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20703	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20704	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20705	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20706	20JUL85	7	L	ACE	0.05	140	90.0	187.3	0.0
H20707	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20708	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20709	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20710	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20711	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20712	20JUL85	7	L	ACE	0.05	140	90.0	18.7	0.0
H20713	20JUL85	7	L	ACE	0.05	140	90.0	1.87	0.0
H20714	20JUL85	7	L	ACE	0.05	140	90.0	1.87	0.0
H20715	20JUL85	7	L	ACE	0.05	140	90.0	1.87	0.0
H20716	20JUL85	7	L	ACE	0.05	140	90.0	1.87	0.0
H20717	20JUL85	7	L	ACE	0.05	140	90.0	1.87	0.0
C22800	22AUG85	0	0	ACE	0.05	0	4.0	0.0	0.0
C22801	22AUG85	0	0	ACE	0.05	0	4.0	20.0	0.0
C22802	22AUG85	0	0	ACE	0.05	0	4.0	10.0	0.0
C22803	22AUG85	0	0	ACE	0.05	0	4.0	4.0	0.0
C22804	22AUG85	0	0	ACE	0.05	0	4.0	2.0	0.0
C22805	22AUG85	0	0	ACE	0.05	0	4.0	2.0	0.0
C22806	22AUG85	0	0	ACE	0.05	0	4.0	1.0	0.0
C22807	22AUG85	0	0	ACE	0.05	0	4.0	0.4	0.0
C22808	22AUG85	0	0	ACE	0.05	0	4.0	0.0	8.0
C22809	22AUG85	0	0	ACE	0.05	0	4.0	0.0	4.0
C22810	22AUG85	0	0	ACE	0.05	0	4.0	0.0	2.0
C22811	22AUG85	0	0	ACE	0.05	0	4.0	0.0	1.0
C22812	22AUG85	0	0	ACE	0.05	0	4.0	0.0	0.4
C22813	22AUG85	0	0	ACE	0.05	0	4.0	0.0	0.2
C22814	22AUG85	0	0	ACE	0.05	0	4.0	0.0	0.1
H90901	28SEP85	4	H	ACE	0.05	120	146.8	200.0	0.0
H90902	28SEP85	4	H	ACE	0.05	120	146.8	200.0	0.0
H90903	28SEP85	4	H	ACE	0.05	120	146.8	200.0	50.0
H90904	28SEP85	4	H	ACE	0.05	120	146.8	200.0	0.0

Table 18. *(continued)*

Run #	Date da mo yr	DS theoretical	DP	Buffer type	M	Ω rpm	Volume mL	CLCLST µL	NOVOZM µL
H90905	28SEP85	4	H	ACE	0.05	120	146.8	200.0	0.0
H90906	28SEP85	4	H	ACE	0.05	120	146.8	50.0	0.0
H90907	28SEP85	4	H	ACE	0.05	120	146.8	50.0	0.0
H90908	28SEP85	4	H	ACE	0.05	120	146.8	50.0	50.0
H90909	28SEP85	4	H	ACE	0.05	120	146.8	50.0	0.0
H90910	28SEP85	4	H	ACE	0.05	120	146.8	50.0	0.0
H90911	28SEP85	7	L	ACE	0.05	120	95.1	20.0	0.0
H90912	28SEP85	7	L	ACE	0.05	120	95.1	20.0	0.0
H90913	28SEP85	7	L	ACE	0.05	120	95.1	400.0	0.0
H90914	28SEP85	7	L	ACE	0.05	120	95.1	400.0	0.0
H90915	28SEP85	7	L	ACE	0.05	120	95.1	200.0	0.0
H90916	28SEP85	7	L	ACE	0.05	120	95.1	200.0	0.0
H90917	28SEP85	7	L	ACE	0.05	120	95.1	200.0	50.0
H90918	28SEP85	7	L	ACE	0.05	120	95.1	200.0	0.0
H90919	28SEP85	7	L	ACE	0.05	120	95.1	20.0	0.0
H90920	28SEP85	7	L	ACE	0.05	120	95.1	400.0	0.0
H90921	28SEP85	9	M	ACE	0.05	120	98.4	200.0	0.0
H90922	28SEP85	9	M	ACE	0.05	120	102.1	200.0	0.0
H90923	28SEP85	9	M	ACE	0.05	120	90.1	200.0	0.0
H90924	28SEP85	9	M	ACE	0.05	120	93.2	50.0	0.0
H90925	28SEP85	9	M	ACE	0.05	120	84.6	400.0	0.0
C02100	1OCT85	0	0	ACE	0.05	120	50.0	0.0	0.0
C02101	1OCT85	0	0	ACE	0.05	120	50.0	2500.0	0.0
C02102	1OCT85	0	0	ACE	0.05	120	50.0	1250.0	0.0
C02103	1OCT85	0	0	ACE	0.05	120	50.0	500.0	0.0
C02104	1OCT85	0	0	ACE	0.05	120	50.0	250.0	0.0
C02105	1OCT85	0	0	ACE	0.05	120	50.0	125.0	0.0
C02106	1OCT85	0	0	ACE	0.05	120	50.0	0.0	5.0
C02107	1OCT85	0	0	ACE	0.05	120	50.0	0.0	2.5
C02108	1OCT85	0	0	ACE	0.05	120	50.0	0.0	1.25
C02109	1OCT85	0	0	ACE	0.05	120	50.0	0.0	0.83
C02110	1OCT85	0	0	ACE	0.05	120	50.0	0.0	0.63

Table 19. *Raw Hydrolysis Data: Part 2*

Run #	t hours	CMC_0 g/L	$G1_0$ g/L	$G2_0$ g/L	$G1$ via area g/L	ω_{h1}	$G1$ via height g/L	ω_{h1}	Cellobiose g/L	ω_{h1}
H0507A	90.0	35.2	0.0	0.0	0.0	0	2.29	1	0.67	1
H0507B	90.0	35.2	0.0	0.0	0.0	0	3.16	1	0.0	1
H0507C	90.0	35.2	0.0	0.0	0.0	0	1.25	1	0.59	1
H29081	2.9	16.02	0.0	0.0	0.0	0	1.39	1	0.0	1
H29081	5.3	16.02	0.0	0.0	0.0	0	1.56	1	0.15	1
H29081	12.8	16.02	0.0	0.0	0.0	0	1.48	1	0.0	1
H29081	27.6	16.02	0.0	0.0	0.0	0	1.68	1	0.27	1
H29081	51.7	16.02	0.0	0.0	0.0	0	1.64	1	0.1	1
H29081	141.3	16.02	0.0	0.0	0.0	0	2.8	1	0.1	1
H29082	2.0	16.02	0.0	0.0	0.0	0	1.09	1	0.51	1
H29082	4.1	16.02	0.0	0.0	0.0	0	1.0	1	0.68	1
H29082	12.8	16.02	0.0	0.0	0.0	0	1.15	1	0.38	1
H29082	27.6	16.02	0.0	0.0	0.0	0	1.36	1	0.27	1
H29082	51.7	16.02	0.0	0.0	0.0	0	1.47	1	0.21	1
H29082	141.3	16.02	0.0	0.0	0.0	0	1.96	1	0.0	1
H29083	2.0	16.02	0.0	0.0	0.0	0	0.76	1	0.59	1
H29083	4.1	16.02	0.0	0.0	0.0	0	0.84	1	0.76	1
H29083	12.8	16.02	0.0	0.0	0.0	0	1.02	1	0.15	1
H29083	27.6	16.02	0.0	0.0	0.0	0	1.36	1	0.33	1
H29083	51.7	16.02	0.0	0.0	0.0	0	1.19	1	0.0	1
H29083	141.3	16.02	0.0	0.0	0.0	0	1.38	1	0.0	1
H29084	2.0	12.04	0.0	0.0	0.0	0	0.99	1	0.0	1
H29084	4.1	12.04	0.0	0.0	0.0	0	1.13	1	0.0	1
H29084	12.8	12.04	0.0	0.0	0.0	0	1.39	1	0.46	1
H29084	27.6	12.04	0.0	0.0	0.0	0	1.2	1	0.0	1
H29084	51.7	12.04	0.0	0.0	0.0	0	1.36	1	0.21	1
H29084	141.3	12.04	0.0	0.0	0.0	0	1.54	1	0.21	1
H29085	2.0	12.04	0.0	0.0	0.0	0	0.67	1	0.44	1
H29085	4.1	12.04	0.0	0.0	0.0	0	0.81	1	0.25	1
H29085	5.3	12.04	0.0	0.0	0.0	0	0.74	1	0.38	1
H29085	12.8	12.04	0.0	0.0	0.0	0	1.07	1	0.77	1
H29085	51.7	12.04	0.0	0.0	0.0	0	1.23	1	0.0	1
H29085	141.3	12.04	0.0	0.0	0.0	0	1.4	1	0.21	1
H29086	5.3	12.04	0.0	0.0	0.0	0	0.7	2	0.28	2
H29086	12.8	12.04	0.0	0.0	0.0	0	0.78	1	0.61	1
H29086	27.6	12.04	0.0	0.0	0.0	0	1.12	1	0.0	1
H29086	51.7	12.04	0.0	0.0	0.0	0	1.05	1	0.0	1
H29086	141.3	12.04	0.0	0.0	0.0	0	1.54	1	0.1	1
H29087	2.0	12.04	0.0	0.0	0.0	0	0.67	1	0.42	1
H29087	4.1	12.04	0.0	0.0	0.0	0	0.87	1	0.59	1
H29087	5.3	12.04	0.0	0.0	0.0	0	0.93	1	0.31	1

Table 19. *(continued)*

Run #	t hours	CMC_0 g/L	$G1_0$ g/L	$G2_0$ g/L	$G1$ via area g/L	ω_{h1}	$G1$ via height g/L	ω_{h1}	Cellobiose g/L	ω_{h1}
H29087	12.8	12.04	0.0	0.0	0.0	0	1.11	1	0.0	1
H29087	27.6	12.04	0.0	0.0	0.0	0	1.2	1	0.13	1
H29087	51.7	12.04	0.0	0.0	0.0	0	1.26	1	0.0	1
H29087	141.3	12.04	0.0	0.0	0.0	0	1.86	1	0.1	1
H29088	4.1	12.04	0.0	0.0	0.0	0	0.23	1	0.0	1
H29088	51.7	12.04	0.0	0.0	0.0	0	0.64	1	0.15	1
H29088	141.3	12.04	0.0	0.0	0.0	0	0.8	2	0.0	2
H21091	65.9	23.88	0.0	0.0	0.0	0	2.9	1	0.0	1
H21091	87.0	23.88	0.0	0.0	0.0	0	3.15	1	0.0	1
H21091	118.4	23.88	0.0	0.0	0.0	0	3.12	1	0.0	1
H14031	0.2	20.11	0.0	0.0	0.0	0	0.38	1	0.0	1
H14032	1.0	20.11	0.0	0.0	0.56	2	0.69	2	0.3	2
H14033	6.3	20.11	0.0	0.0	0.62	1	1.01	1	0.31	1
H14034	25.5	20.11	0.0	0.0	2.08	1	1.17	1	0.06	1
H14035	102.5	20.11	0.0	0.0	0.0	0	1.47	1	0.0	1
H14036	0.2	20.11	0.0	0.0	1.3	1	0.6	1	0.25	1
H14037	1.0	20.11	0.0	0.0	0.0	0	1.71	1	0.47	1
H14038	6.3	20.11	0.0	0.0	0.0	0	1.46	1	0.1	1
H14039	25.5	20.11	0.0	0.0	1.85	2	1.58	1	0.12	2
H21031	0.25	48.06	0.0	0.0	0.0	0	0.47	1	0.54	1
H21032	11.5	48.27	0.0	0.0	0.78	1	1.22	1	0.29	1
H21033	81.6	47.76	0.0	0.0	2.74	2	2.47	2	0.27	1
H21034	121.0	48.16	0.0	0.0	2.89	2	2.51	2	0.04	1
H21035	0.25	48.27	0.0	0.0	0.26	1	0.32	1	0.29	1
H21036	11.5	47.96	0.0	0.0	0.37	1	0.6	1	0.0	1
H21037	81.6	48.16	0.0	0.0	1.55	2	1.43	2	0.0	0
H21038	121.0	47.96	0.0	0.0	1.61	2	1.46	2	0.0	1
H12041	0.8	50.68	0.0	0.0	0.0	0	1.22	1	0.92	1
H12042	3.8	50.68	0.0	0.0	4.44	1	0.0	0	0.87	1
H12043	16.5	50.68	0.0	0.0	4.06	1	0.0	0	0.64	1
H12044	39.3	50.68	0.0	0.0	5.01	2	3.76	1	0.34	1
H12045	54.0	50.68	0.0	0.0	5.29	2	4.43	1	0.18	2
H12046	0.8	50.68	1.07	0.0	2.45	1	0.0	0	0.96	1
H12047	3.8	50.68	1.07	0.0	4.03	1	0.0	0	0.92	1
H12048	16.5	50.68	1.07	0.0	5.41	2	4.34	1	0.31	1
H12049	39.3	50.68	1.06	0.0	5.87	2	4.68	1	0.22	2
H13041	54.0	50.68	1.06	0.0	7.16	2	5.42	1	0.14	2
H13042	0.8	50.68	0.0	1.07	1.68	1	1.15	1	1.23	1
H13043	3.8	50.68	0.0	1.09	3.77	1	0.0	0	1.07	1
H13044	16.5	50.68	0.0	1.08	5.14	2	0.0	0	0.28	1
H13045	39.3	50.68	0.0	1.11	6.35	2	0.0	0	0.26	2

Table 19. *(continued)*

Run #	t hours	CMC_0 g/L	$G1_0$ g/L	$G2_0$ g/L	$G1$ via area g/L	ω_{h1}	$G1$ via height g/L	ω_{h1}	Cellobiose g/L	ω_{h1}
H13046	54.0	50.68	0.0	1.13	6.49	2	5.18	1	0.24	1
H13047	0.8	50.68	0.0	0.0	1.93	1	0.0	0	0.94	1
H13048	16.5	50.68	0.0	0.0	4.46	2	3.82	1	0.47	2
H13049	54.0	50.68	0.0	0.0	5.9	2	4.7	1	0.0	1
C13041	3.8	0.0	0.0	1.03	0.38	2	0.36	3	0.65	2
C13042	39.3	0.0	0.0	0.94	1.33	1	1.02	1	0.15	1
H09501	1.0	35.0	0.0	0.0	1.34	1	1.34	1	0.64	1
H09502	3.9	35.0	0.0	0.0	3.3	1	2.62	1	1.18	1
H09503	12.0	35.0	0.0	0.0	3.06	2	2.66	1	0.5	2
H09504	27.7	35.0	0.0	0.0	4.2	2	3.2	1	1.21	2
H09505	120.0	35.0	0.0	0.0	4.02	2	4.4	1	0.65	3
H09506	1.0	35.0	0.0	0.0	1.36	1	1.47	1	0.76	1
H09507	3.9	35.0	0.0	0.0	3.31	2	2.66	1	0.98	2
H09508	12.0	35.0	0.0	0.0	3.21	1	2.59	1	0.94	1
H09509	27.7	35.0	0.0	0.0	3.71	2	3.05	1	0.55	2
H09510	120.0	35.0	0.0	0.0	4.47	2	4.13	1	0.65	2
H09511	1.0	35.0	0.0	0.0	0.59	1	0.77	2	0.09	2
H09512	3.9	35.0	0.0	0.0	1.86	2	1.78	1	0.0	2
H09513	12.0	35.0	0.0	0.0	2.04	2	1.89	1	0.05	2
H09514	27.7	35.0	0.0	0.0	1.74	1	2.12	1	0.0	0
H09515	120.0	35.0	0.0	0.0	3.14	1	2.48	1	0.0	1
H01701	1.0	35.0	0.0	0.0	1.52	1	1.23	1	0.61	1
H01702	4.0	35.0	0.0	0.0	2.51	1	2.16	1	0.49	1
H01703	17.9	35.0	0.0	0.0	2.86	1	2.64	1	0.26	1
H01704	43.5	35.0	0.0	0.0	3.47	1	3.21	1	0.57	1
H01705	117.3	35.0	0.0	0.0	3.89	1	4.12	1	0.56	1
H01706	1.0	35.0	0.0	0.0	0.0	3	0.0	3	0.56	3
H01707	4.0	35.0	0.0	0.0	0.89	1	0.9	1	0.77	1
H01708	17.9	35.0	0.0	0.0	1.85	1	1.87	1	0.73	1
H01709	43.5	35.0	0.0	0.0	2.02	2	2.31	2	0.37	2
H01710	117.3	35.0	0.0	0.0	2.37	2	2.68	1	0.26	2
H01711	1.0	35.0	0.0	0.0	0.0	0	0.14	1	0.05	1
H01712	4.0	35.0	0.0	0.0	0.0	0	0.08	1	0.0	1
H01713	17.9	35.0	0.0	0.0	0.0	0	0.29	1	0.61	1
H01714	43.5	35.0	0.0	0.0	0.4	1	0.5	2	0.92	2
H01715	117.3	35.0	0.0	0.0	0.59	1	0.47	2	0.98	1
H01716	17.9	35.0	0.0	0.0	0.0	1	0.01	1	0.0	1
H01717	43.5	35.0	0.0	0.0	0.0	0	0.03	1	0.02	1
H01718	117.3	35.0	0.0	0.0	0.0	0	0.08	1	0.0	1
H20701	1.0	52.55	0.0	0.0	1.39	2	1.67	1	1.34	2
H20702	4.0	52.55	0.0	0.0	2.87	2	3.2	1	0.82	2

Table 19. *(continued)*

Run #	t hours	CMC_0 g/L	$G1_0$ g/L	$G2_0$ g/L	$G1$ via area g/L	ω_{h1}	$G1$ via height g/L	ω_{h1}	Cellobiose g/L	ω_{h1}
H20703	4.0	52.55	0.0	0.0	2.96	2	3.16	2	1.83	1
H20704	16.0	52.55	0.0	0.0	3.67	1	4.34	1	1.36	1
H20705	45.0	52.55	0.0	0.0	4.95	2	5.15	2	1.83	1
H20706	124.0	52.55	0.0	0.0	8.07	1	6.14	1	0.0	0
H20707	1.0	52.55	0.0	0.0	0.59	1	0.31	1	0.94	1
H20708	4.0	52.55	0.0	0.0	1.04	1	1.02	1	1.2	1
H20709	16.0	52.55	0.0	0.0	1.92	1	1.9	1	1.08	1
H20710	16.0	52.55	0.0	0.0	2.45	2	2.62	1	1.34	2
H20711	45.0	52.55	0.0	0.0	3.01	1	3.04	1	1.12	1
H20712	124.0	52.55	0.0	0.0	3.91	1	3.9	1	0.8	1
H20713	1.0	52.55	0.0	0.0	0.01	1	0.03	1	0.1	1
H20714	4.0	52.55	0.0	0.0	0.0	1	0.0	1	0.0	1
H20715	16.0	52.55	0.0	0.0	0.0	0	0.07	1	0.92	1
H20716	45.0	52.55	0.0	0.0	0.0	0	0.17	1	0.93	1
H20717	124.0	52.55	0.0	0.0	0.3	1	0.48	1	1.27	2
C22800	1.0	0.0	0.0	5.0	0.0	0	0.07	1	4.87	1
C22801	1.0	0.0	0.0	5.0	0.4	1	0.46	1	4.42	1
C22802	1.0	0.0	0.0	5.0	0.26	1	0.27	1	4.87	1
C22803	1.0	0.0	0.0	5.0	0.0	0	0.23	1	4.85	1
C22804	1.0	0.0	0.0	5.0	0.19	1	0.19	1	4.68	1
C22805	1.0	0.0	0.0	5.0	0.48	1	0.49	1	4.98	1
C22806	1.0	0.0	0.0	5.0	0.0	0	0.15	1	4.82	1
C22807	1.0	0.0	0.0	5.0	0.0	0	0.07	1	5.19	1
C22808	1.0	0.0	0.0	5.0	5.46	1	0.0	0	0.0	1
C22809	1.0	0.0	0.0	5.0	5.12	1	0.0	0	0.13	1
C22810	1.0	0.0	0.0	5.0	5.01	1	5.02	1	0.62	1
C22811	1.0	0.0	0.0	5.0	3.3	1	3.33	1	1.61	1
C22812	1.0	0.0	0.0	5.0	1.71	1	1.77	1	3.26	1
C22813	1.0	0.0	0.0	5.0	1.09	1	1.13	1	4.19	1
C22814	1.0	0.0	0.0	5.0	0.69	1	0.67	1	4.29	1
H90901	1.0	7.4	0.0	0.0	0.7	1	0.78	1	0.63	1
H90902	1.0	7.4	0.0	0.0	0.87	1	0.92	1	0.76	1
H90903	4.0	7.4	0.0	0.0	1.9	1	1.99	1	0.31	1
H90904	16.0	7.4	0.0	0.0	1.51	1	1.54	1	0.68	1
H90905	120.0	7.4	0.0	0.0	2.32	1	2.36	1	0.0	1
H90906	4.0	7.4	0.0	0.0	0.74	1	0.82	1	0.72	1
H90907	16.0	7.4	0.0	0.0	1.04	1	1.13	1	0.61	1
H90908	45.0	7.4	0.0	0.0	1.95	1	2.09	1	0.0	1
H90909	45.0	7.4	0.0	0.0	1.53	1	1.61	1	0.42	2
H90910	120.0	7.4	0.0	0.0	1.86	1	1.87	1	0.3	2
H90911	4.0	35.0	0.0	0.0	1.03	1	1.15	1	0.93	1

Table 19. *(continued)*

Run #	t hours	CMC_0 g/L	$G1_0$ g/L	$G2_0$ g/L	$G1$ via area g/L	ω_{h1}	$G1$ via height g/L	ω_{h1}	Cellobiose g/L	ω_{h1}
H90912	16.0	35.0	0.0	0.0	1.43	1	1.44	1	0.73	1
H90913	1.0	35.0	0.0	0.0	2.02	1	2.01	1	0.53	1
H90914	45.0	35.0	0.0	0.0	4.25	1	3.83	1	0.0	1
H90915	1.0	35.0	0.0	0.0	1.57	1	1.59	1	0.3	1
H90916	4.0	35.0	0.0	0.0	2.35	1	2.23	1	0.77	2
H90917	16.0	35.0	0.0	0.0	3.66	1	3.63	1	0.0	1
H90918	45.0	35.0	0.0	0.0	3.78	1	3.53	1	0.56	2
H90919	120.0	35.0	0.0	0.0	2.8	1	2.71	1	1.13	1
H90920	120.0	35.0	0.0	0.0	5.19	1	0.0	0	0.0	0
H90921	4.0	61.0	0.0	0.0	1.88	1	0.0	0	0.77	1
H90922	16.0	61.0	0.0	0.0	2.85	1	2.51	1	0.93	1
H90923	45.0	61.0	0.0	0.0	3.34	1	0.0	0	0.93	1
H90924	45.0	61.0	0.0	0.0	2.35	1	2.19	1	0.62	2
H90925	122.0	61.0	0.0	0.0	3.98	1	3.67	1	0.41	1
C02100	1.0	0.0	0.0	5.0	0.0	0	0.08	1	5.1	1
C02101	1.0	0.0	0.0	5.0	0.0	0	8.01	1	2.97	1
C02102	1.0	0.0	0.0	5.0	6.21	1	5.86	1	3.87	1
C02103	1.0	0.0	0.0	5.0	2.56	2	2.27	2	4.5	1
C02104	1.0	0.0	0.0	5.0	1.1	1	1.15	1	4.52	1
C02105	1.0	0.0	0.0	5.0	0.4	1	0.46	1	4.8	1
C02106	1.0	0.0	0.0	5.0	2.16	1	2.25	1	2.86	1
C02107	1.0	0.0	0.0	5.0	1.24	1	1.32	1	3.83	1
C02108	1.0	0.0	0.0	5.0	0.62	1	0.75	1	4.6	1
C02109	1.0	0.0	0.0	5.0	0.44	1	0.54	1	4.69	1
C02110	1.0	0.0	0.0	5.0	0.34	1	0.48	1	4.93	1

Appendix 4

MINOR DERIVATIONS

4.1 Modelling the Non-Linearities of Enzyme Dilution Series

The normal procedure for estimating the concentration of a given enzyme in a solution consists in measuring the quantity of product obtained or the uptake of substrate as the enzyme of interest catalyzes a reaction on a specific substrate.

In this work, the concentration of endo-glucanase in each of the Novozym and Celluclast solutions is calculated as being proportional to the amount of glucose produced when a piece of filter paper #1 of size $1 \times 6\,\text{cm}$ is hydrolyzed. See the experimental procedure for details.

Cellobiose is the substrate we chose to determine the concentration of cellobiase enzyme in a given solution. The cellobiase activity test with cellobiose as the substrate can be used only when the product—glucose—concentration can be measured separately from the cellobiose. The DNS spectrophotometric method of Matteau and Saddler [20] uses a salicin solution as the substrate because the glucose and cellobiose affect the adsorbance measurements and hence the product concentrations. The HPLC available to us discerns between the glucose and the cellobiose. That is why we use cellobiose as a substrate while slightly modifying the salicinase activity test.

The theoretical problem to be adressed in this section concerns the non-linearity of the product obtained as a function of the enzyme in the solution. In other words, a dilution series of the enzyme should yield a linear product concentration. If the product curve is not linear, the enzyme must be further diluted until the linear region is found.

Mandels et al. [32] have standardized the filter paper activity (FPA) test so that the enzyme activity is picked off for the run which yields the glucose concentration of 0.667g/L. At that glucose yield, for the conditions of [32], the enzyme dilution curve is linear.

For an unknown concentration of enzyme in a mixture, it may be necessary to try many volumetric concentrations of the enzyme before the proper amount of glucose is obtained. The present data analysis procedure in the FPA and CBA tests makes use of the results in the linear section of the dilution curve. It is to our advantage to apply a strict modelling procedure. A model can be developed to represent the data over a wide range of dilutions of the enzyme. There are two advantages to this approach: 1) all the activity test data influences the estimate of the concentration of the enzyme of interest and 2) it is possible to calculate the confidence in the values of the concentrations of cellobiase and endoglucanase in the Novozym and Celluclast solutions.

The non-linearity of product obtained as a function of enzyme has been discussed earlier in the Literature Survey. The following are possible causes.

1. Inhibition of enzyme by the product.
2. Initial enzyme-to-substrate ratio is too high.
3. Available substrate is depleted quickly.
4. Deactivation of the enzyme during the test.

The second reason is not plausible in our case; by definition, the test must take place at high enzyme dilution and a fixed initial substrate concentration. The fourth item is assumed negligible because the test hydrolysis lasts one hour. Over that period, at 50°C, the specification sheet accompanying the Celluclast mentions that the enzyme endoglucanase retains 92% of its activity. Note also that the fraction of deactivated enzyme will be the same at all initial enzyme concentration. It is assumed here,

without the specification sheet for Novozym, that cellobiase deactivation is also negligible.

A model must be developed that represents enzyme activity data over a wide range of dilutions, accounting for non-linearities with items 1 and 3.

Huang [16] has developed a kinetic model that considers the adsorption of one enzyme on the insoluble amorphous cellulose followed by the slow step. This slow step is the decomposition of the enzyme-substrate complex. Furthermore the model assumes competitive inhibition by the products. In the case of cellulose hydrolysis both the reducing sugars are considered inhibitors. Here is Huang's model, using the nomenclature in his paper:

$$\frac{t}{(P)} = \frac{1 + K_1(E)_0 + K_3(S)_0}{k_2 X_{1m} K_1 (E)_0} \left\{ \frac{1}{(P)} \ln \frac{(S)_0}{(S)_0 - P} \right\} + \frac{X_{1m} K_1 - K_3}{k_2 X_{1m} K_1 (E)_0} \qquad (135)$$

Let us transform Huang's model using our own nomenclature and some simplifications that apply to enzyme activity tests. Lump the constants into parameters to be evaluated in the modelling. The only variables of interest in the CBA and FPA modelling are:

- e_0 the initial enzyme concentration, the independent variable and
- the product concentration, p.

The test temperature is constant at 50°C hence the kinetic rate constants do not vary during the test. The initial substrate concentration and the time of hydrolysis during the test are unchanging. These data may be lumped with the parameters. Finally, set $(S)_0$ to p_∞. This is permissible because, like Huang, we set $(S) = (S)_0 - (P)$ ignoring any stoichiometry. Huang's equation reduces to:

$$1 + \left(\theta_0 + \frac{\theta_1}{e_0}\right) \ln \left(1 - \frac{p}{p_\infty}\right) - \theta_2 \cdot \frac{p}{e_0} \qquad (136)$$

where θ_0 and θ_1 are positive modelling parameters. θ_2 is also a parameter, but not necessarily positive.

The substrate of the cellobiase activity test is soluble. Huang's model was developed for an insoluble substrate onto which the enzyme is adsorbing. We nevertheless attempt to fit Huang's model to the range of enzyme concentrations involved. After all, Poulsen and Petersen [13] have attempted to fit the progress curves of the hydrolysis of CMC to Huang's model. That fit appears excellent even though the CMC is soluble in solution whereas the model assumes adsorption. It may then be possible to fit Huang's model to the hydrolysis of cellobiose by cellobiase.

Our own model can be developed in a manner similar to Huang's. The assumptions would take into account the fact that the substrate in the CBA test, cellobiose, is soluble. If we assume that Michaelis-Menten kinetics hold, and that there is competitive inhibition by the product and significant depletion of substrate [28], then

$$v = -\frac{ds}{dt} = \frac{dp}{dt} = \frac{ke_0 s}{s + K_s(1 + p/K_p)} \tag{137}$$

As Huang [16], set $s = s_0 - p$ and substitute in 137.

$$\frac{dp}{dt} = \frac{ke_0(s_0 - p)}{s_0 - p + K_s(1 + p/K_p)} \tag{138}$$

$$\frac{dp}{dt} = \frac{ks_0e_0 - ke_0p}{(s_0 + K_s) + (K_s/K_p - 1)p} \tag{139}$$

$$\frac{(s_0 + K_s) + (K_s/K_p - 1)p}{ks_0e_0 - ke_0p}\, dp = dt \tag{140}$$

Integrate from $t = 0$ to $t = t$

$$\int_0^p \left(\frac{(s_0 + K_s)}{ks_0e_0 - ke_0p} + \frac{(K_s/K_p - 1)p}{ks_0e_0 - ke_0p} \right) dp = \int_0^t dt \tag{141}$$

$$t = \left[\frac{(s_0 + K_s)\ln(ks_0e_0 - ke_0p)}{-ke_0} \right] + \left[\frac{K_s}{K_p} - 1 \right] \left[\frac{-1}{ke_0} \right] \left[p + \frac{ks_0e_0}{ke_0} \ln\left(p - \frac{ks_0e_0}{ke_0} \right) \right] \Bigg|_{p=0}^{p=p} \tag{142}$$

Simplify before evaluating.

$$-ke_0t = (s_0 + K_s)\left[\ln(ke_0) + \ln(s_0 - p)\right] + \left(\frac{K_s}{K_p} - 1 \right) \left[p + s_0 \ln(p - s_0) \right] \Big|_{p=0}^{p=p} \tag{143}$$

$$-ke_0t = (s_0 + K_s)\left[\ln(s_0 - p) - \ln s_0\right] + \left(\frac{K_s}{K_p} - 1\right)\left[p + s_0\ln(p - s_0) - s_0\ln(-s_0)\right] \tag{144}$$

$$-ke_0t = (s_0 + K_s)\ln\left(1 - \frac{p}{s_0}\right) + \left(\frac{K_s}{K_p} - 1\right)\left[p + s_0\ln\left(1 - \frac{p}{s_0}\right)\right] \tag{145}$$

$$-ke_0t = \left(s_0 + K_s + \frac{K_s}{K_p}s_0 - s_0\right)\ln(1 - \frac{p}{s_0}) + \left(\frac{K_s}{K_p} - 1\right)p \tag{146}$$

Let $s_0 = p_\infty$

$$-kte_0 = K_s\left(1 + \frac{p_\infty}{K_p}\right)\ln\left(1 - \frac{p}{p_\infty}\right) + \left(\frac{K_s}{K_p} - 1\right)p \tag{147}$$

Reparameterize equation 147 for the purpose of modelling.

$$p = \theta_0 e_0 + \theta_1 \ln\left(1 - \frac{p}{p_\infty}\right) \tag{148}$$

One more parameter may be added to the model by letting p_∞ float.

The equation derived here and Huang's model are both useful in enzyme activity tests where the DF covers a large range and the dilution series of the enzyme is not linear. Both our equation and Huang's equation assume that the non-linearity of the dilution series is due to product inhibition and cryptic substrate depletion. A strict modelling approach leads to a more accurate estimation of $\widehat{\text{FPA}}$ and $\widehat{\text{CBA}}$.

4.2 Action of an Enzyme on Multiple Substrates

It is apparent that the mode of action of a multi-substrate enzyme on one of its substrates can be calculated from the knowledge of the concentration of the individual substrate considered and of the total available substrate to that enzyme. In Bailey & Ollis [28], on page 114, the kinetics of a 'one-enzyme, two-substrate reaction' are examined. The substrates both bind differently to the enzyme hence the tabulation of two different dissociation constants. Their application of the Michaelis-Menten enzyme equilibria yields the two simultaneous equations

$$\frac{-ds_1}{dt} = v_1 = \frac{k_1 e_0 s_1/K_1}{1 + s_1/K_1 + s_2/K_2} \tag{149}$$

$$\frac{-ds_2}{dt} = v_2 = \frac{k_2 e_0 s_2 / K_2}{1 + s_1/K_1 + s_2/K_2} \tag{150}$$

An important assumption of this work is that enzymes of the cellulase system and β-glucosidase act at the same rate on the different bond substrates. In equations 149 and 150, this is equivalent to saying that $k_1 = k_2$ and $K_1 = K_2$. Simplification yields

$$\frac{-ds_1}{dt} = v_1 = \frac{k \cdot e_0 \cdot s_1}{K + s_1 + s_2} \tag{151}$$

$$\frac{-ds_2}{dt} = v_2 = \frac{k \cdot e_0 \cdot s_2}{K + s_1 + s_2} \tag{152}$$

Now, consider the more general case of one enzyme acting on multiple substrates. These are the relevant reactions for the n substrates of the one enzyme in the system.

$$E + S_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_1 \xrightarrow{k_{p1}} P_1$$

$$E + S_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ES_2 \xrightarrow{k_{p2}} P_2$$

$$\cdots$$

$$E + S_n \underset{k_{-n}}{\overset{k_n}{\rightleftharpoons}} ES_n \xrightarrow{k_{pn}} P_n$$

The substrate uptake equations are given by:

$$-\frac{ds_1}{dt} = k_1 \cdot e \cdot s_1 - k_{-1} \cdot (es_1)$$

$$-\frac{ds_2}{dt} = k_2 \cdot e \cdot s_2 - k_{-2} \cdot (es_2)$$

$$\cdots$$

$$-\frac{ds_n}{dt} = k_n \cdot e \cdot s_n - k_{-n} \cdot (es_n) \tag{153}$$

where s_i is the concentration of substrate i in the reactor, e is the free enzyme concentration in the reactor and (es_i) is the concentration of the enzyme-substrate i complex. The rate equations describing the concentration of (es_i) are:

$$-\frac{d(es_1)}{dt} = -k_1 \cdot e \cdot s_1 + (k_{-1} + k_{p1}) \cdot (es_1)$$

$$-\frac{d(es_2)}{dt} = -k_2 \cdot e \cdot s_2 + (k_{-2} + k_{p2}) \cdot (es_2)$$

$$\cdots$$

$$-\frac{d(es_n)}{dt} = -k_n \cdot e \cdot s_n + (k_{-n} + k_{pn}) \cdot (es_n) \qquad (154)$$

We now make the *quasi-steady-state approximation*, $\frac{d(es_i)}{dt} = 0$. This assumption can be proved valid for the case where e_0/s_0 is sufficiently small [28]. The equations relating the concentrations of the enzyme-substrate complex thus become:

$$(es_1) = \frac{k_1 \cdot e \cdot s_1}{k_{-1} + k_{p1}}$$

$$(es_2) = \frac{k_2 \cdot e \cdot s_2}{k_{-2} + k_{p2}}$$

$$\cdots$$

$$(es_n) = \frac{k_n \cdot e \cdot s_n}{k_{-n} + k_{pn}} \qquad (155)$$

Substitute into the substrate uptake equations to eliminate (es_i).

$$-\frac{ds_1}{dt} = k_1 \cdot e \cdot s_1 - k_{-1} \cdot \left[\frac{k_1}{k_{-1} + k_{p1}}\right] \cdot e \cdot s_1 = k_{p1} \cdot e \cdot s_1 \cdot \left[\frac{k_1}{k_{-1} + k_{p1}}\right]$$

$$-\frac{ds_2}{dt} = k_2 \cdot e \cdot s_2 - k_{-2} \cdot \left[\frac{k_2}{k_{-2} + k_{p2}}\right] \cdot e \cdot s_2 = k_{p2} \cdot e \cdot s_2 \cdot \left[\frac{k_2}{k_{-2} + k_{p2}}\right]$$

$$\cdots$$

$$-\frac{ds_n}{dt} = k_n \cdot e \cdot s_n - k_{-n} \cdot \left[\frac{k_n}{k_{-n} + k_{pn}}\right] \cdot e \cdot s_n = k_{pn} \cdot e \cdot s_n \cdot \left[\frac{k_n}{k_{-n} + k_{pn}}\right] \qquad (156)$$

Apply the law of conservation of mass to the enzyme.

$$\begin{aligned} e_0 &= e + (es_1) + (es_2) + \cdots + (es_n) \\ &= e + \left[\frac{k_1}{k_{-1} + k_{p1}}\right] \cdot e \cdot s_1 + \left[\frac{k_2}{k_{-2} + k_{p2}}\right] \cdot e \cdot s_2 + \cdots + \left[\frac{k_n}{k_{-n} + k_{pn}}\right] \cdot e \cdot s_n \end{aligned}$$

Thus $$e = \frac{e_0}{1 + \sum_{i=1}^{n} \frac{k_i \cdot s_i}{k_{-i} + k_{pi}}} \qquad (157)$$

Substitute equation 157 into 156 to eliminate e.

$$-\frac{ds_1}{dt} = k_{p1} \cdot e_0 \cdot s_1 \cdot \frac{\frac{k_1}{k_{-1}+k_{p1}}}{1+\sum_{i=1}^{n} \frac{k_i \cdot s_i}{k_{-i}+k_{pi}}}$$

$$-\frac{ds_2}{dt} = k_{p2} \cdot e_0 \cdot s_2 \cdot \frac{\frac{k_2}{k_{-2}+k_{p2}}}{1+\sum_{i=1}^{n} \frac{k_i \cdot s_i}{k_{-i}+k_{pi}}}$$

$$\cdots$$

$$-\frac{ds_n}{dt} = k_{pn} \cdot e_0 \cdot s_n \cdot \frac{\frac{k_n}{k_{-n}+k_{pn}}}{1+\sum_{i=1}^{n} \frac{k_i \cdot s_i}{k_{-i}+k_{pi}}}$$

Now, let $\frac{k_{-i}+k_{pi}}{k_i} = K_i$ for all i

The rate of disappearance of substrate i is given by

$$\frac{ds_i}{dt} = \frac{k_{pi} \cdot e_0 \cdot s_i/K_i}{1+\sum_{j=1}^{n} s_j/K_j} \tag{158}$$

An important assumption of this work is that the enzymes of the cellulase system, including β-glucosidase act at the same rate on their own bond substrates. In equation 158, this is equivalent setting $k_{p1} = k_{p2} = \cdots = k_{pn} = k$ and $K_1 = K_2 = \cdots = K_n = K$. The substrate depletion rate equations for a one-enzyme, multi-substrate reaction with the substrate type being transparent to the enzyme are:

$$\frac{-ds_1}{dt} = v_1 = \frac{k \cdot e_0 \cdot s_1}{K + s_1 + s_2 + \cdots + s_n}$$

$$\frac{-ds_2}{dt} = v_2 = \frac{k \cdot e_0 \cdot s_2}{K + s_1 + s_2 + \cdots + s_n}$$

$$\cdots$$

$$\frac{-ds_n}{dt} = v_n = \frac{k \cdot e_0 \cdot s_n}{K + s_1 + s_2 + \cdots + s_n} \tag{159}$$

Appendix 5

COMPUTER PROGRAMS

Program 2. FPA SAS

```
TITLE1'FILTER PAPER ACTIVITY OF CELLUCLAST AND NOVOZYM';
DATA FPA_CALC;
    CMS FILEDEF MYDATA DISK FPA DATA;
    INFILE MYDATA;
    FORMAT RUNDATE DDMMYY6.;
    INPUT @ 7 RUNDATE DDMMYY6./* DATE: 2 DIGITS EACH FOR DA, MO, YR   */
          @14 ENZTYPE  $ 6.     /* ENZYME BROTH TYPE: NOVOZM OR CLCLST  */
          @21 DF         5.     /* DILUTION FACTOR FOR THE FPA TEST     */
          @27 FPA        6.     /* FILTER PAPER ACTIVITY FOR GIVEN RUN  */;
    GLUC=FPA*10.8/DF;
    ENZ=1./DF;
    DFSQ=DF*DF ;
PROC SORT DATA=FPA_CALC;
    BY ENZTYPE;
PROC PRINT DATA=FPA_CALC;
    BY ENZTYPE;
DATA NOVOZM CLCLST ;
    SET FPA_CALC;
    IF ENZTYPE='NOVOZM' THEN OUTPUT NOVOZM;
    IF ENZTYPE='CLCLST' THEN OUTPUT CLCLST;
RUN;
TITLE1'FILTER PAPER ACTIVITY OF CELLUCLAST';
TITLE2'MODEL: GLUC2-THETA0*ENZ-THETA1*LOG((1.-THETA2*GLUC)**2)';
PROC NLIN DATA=CLCLST METHOD=DUD EFORMAT;
    PARMS CTEA=1069.0E+0 CTEB=0.00180E-0 CTEC=11359.;
    GLUC2=GLUC;
    DO I=1 TO 15 ;
      FUNC=GLUC2-CTEA*ENZ-CTEB*LOG((1.-CTEC*GLUC2)**2) ;
      DERIV=1.+CTEC*CTEB*2./(1.-CTEC*GLUC2) ;
      GLUC1=GLUC2-FUNC/DERIV;
      EPS=ABS((GLUC1-GLUC2)/GLUC1) ;
      IF EPS<1.E-12 THEN I=15 ;
      GLUC2=GLUC1 ;
    END ;
MODEL GLUC=GLUC2 ;
OUTPUT OUT=FPARES P=PRED R=RESID ;
PROC PLOT DATA=FPARES;
  PLOT GLUC*ENZ='A' PRED*ENZ='P' / OVERLAY VPOS=50;
  PLOT RESID*ENZ / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
PROC PRINT DATA=FPARES NOOBS UNIFORM;
RUN;
TITLE2'MODEL: FPA=THETA0+THETA1*DF';
PROC REG DATA=CLCLST;
  MODEL FPA=DF / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
```

```
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: FPA=THETA0+THETA1*DF+THETA2*DF**2';
PROC REG DATA=CLCLST;
  MODEL FPA=DF DFSQ / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: FPA=THETA0+THETA1*DF**2';
PROC REG DATA=CLCLST;
  MODEL FPA=DFSQ / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE1'FILTER PAPER ACTIVITY OF NOVOZYM';
TITLE2'MODEL: FPA=THETA0+THETA1*DF';
PROC REG DATA=NOVOZM;
  MODEL FPA=DF / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: FPA=THETA0+THETA1*DF+THETA*DFSQ';
PROC REG DATA=NOVOZM;
  MODEL FPA=DF DFSQ / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: FPA=THETA0+THETA1*DFSQ';
PROC REG DATA=NOVOZM;
  MODEL FPA=DFSQ / COVB CORRB CLM CLI P;
  OUTPUT OUT=FPARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=FPARES;
  PLOT FPA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
```

Program 3. CBA SAS

```
TITLE1'CELLOBIASE ACTIVITY OF NOVOZYM AND CELLUCLAST';
DATA CBA_CALC;
    CMS FILEDEF MYDATA DISK CBA DATA;
    INFILE MYDATA;
    FORMAT RUNDATE DDMMYY6.;
    INPUT @ 2 RUNDATE DDMMYY6./* DATE: 2 DIGITS EACH FOR DA, MO, YR   */
          @10 RUN       $ 6.     /* RUN NUMBER                          */
          @20 ENZTYPE   $ 6.     /* ENZYME BROTH TYPE: NOVOZM OR CLCLST  */
          @30 DF          8.     /* DILUTION FACTOR FOR THE FPA TEST     */
          @40 GLUC        8.     /* GLUCOSE CONCENTRATION FOR GIVEN RUN  */;
    CBA=GLUC*DF/5.4 ;
    DFSQ=DF*DF;
    DROP GLUC;
PROC SORT DATA=CBA_CALC;
    BY ENZTYPE;
PROC PRINT DATA=CBA_CALC;
    BY ENZTYPE;
DATA NOVOZM CLCLST ;
    SET CBA_CALC;
    IF ENZTYPE='NOVOZM' THEN OUTPUT NOVOZM;
    IF ENZTYPE='CLCLST' THEN OUTPUT CLCLST;
RUN;
TITLE1'CELLOBIASE ACTIVITY OF NOVOZYM';
TITLE2 'MODEL CBA-THETA0-THETA1*DF*LN((1-1.026*CBA/DF)**2)=0' ;
PROC NLIN DATA=NOVOZM METHOD=DUD EFORMAT;
    PARMS CTEA=970. CTEB=0.0031 ;
    CBA2=CBA ;
    DO I=1 TO 15 ;
      FUNC=CBA2-CTEA-CTEB*DF*LOG((1.-5.4/5.35*CBA2/DF)**2) ;
      DERIV=1.+5.4/5.35*CTEB*2./(1.-5.4/5.35*CBA2/DF) ;
      CBA1=CBA2-FUNC/DERIV ;
      EPS=ABS((CBA1-CBA2)/CBA1) ;
      IF EPS<1.E-13 THEN I=15 ;
      CBA2=CBA1 ;
    END ;
MODEL CBA=CBA2 ;
OUTPUT OUT=CBARES P=PRED R=RESID;
PROC PLOT DATA=CBARES;
  PLOT CBA*DF='A' PRED*DF='P' / VREF=0 OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
PROC PRINT DATA=CBARES NOOBS UNIFORM;
RUN;
TITLE2'MODEL: CBA=THETA0*DF*(1-EXP(-THETA1/DF))';
PROC NLIN DATA=NOVOZM METHOD=MARQUARDT EFORMAT;
  PARMS CTEA=.020 CTEB=2000.0 ;
  MODEL CBA=CTEA*DF*(1.-EXP(-CTEB/DF));
  DER.CTEA=DF*(1.-EXP(-CTEB/DF));
  DER.CTEB=CTEA*EXP(-CTEB/DF);
  OUTPUT OUT=CBARES P=PRED R=RESID U95=U95 L95=L95 U95M=U95M L95M=L95M;
```

```
PROC PLOT DATA=CBARES;
  PLOT CBA*DF='A' PRED*DF='P' U95M*DF='-' L95M*DF='-' /
  VREF=0 OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
PROC PRINT DATA=CBARES NOOBS UNIFORM;
RUN;
TITLE1'CELLOBIASE ACTIVITY OF CELLUCLAST';
TITLE2'MODEL: CBA=THETA0*DF*(1-EXP(-THETA1/DF))';
PROC NLIN DATA=CLCLST METHOD=MARQUARDT EFORMAT;
  PARMS CTEA=.100 CTEB=200.0 ;
  MODEL CBA=CTEA*DF*(1.-EXP(-CTEB/DF));
  DER.CTEA=DF*(1.-EXP(-CTEB/DF));
  DER.CTEB=CTEA*EXP(-CTEB/DF);
  OUTPUT OUT=CBARES P=PRED R=RESID U95=U95 L95=L95 U95M=U95M L95M=L95M;
PROC PLOT DATA=CBARES;
  PLOT CBA*DF='A' PRED*DF='P' U95M*DF='-' L95M*DF='-' /
  VREF=0 OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
PROC PRINT DATA=CBARES NOOBS UNIFORM;
RUN;
TITLE2'MODEL: CBA=THETA0+THETA1*DF ...ALL EXPERIMENTAL DATA';
PROC REG DATA=CLCLST ;
    MODEL CBA=DF / COVB CORRB CLM CLI P;
OUTPUT OUT=CBARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=CBARES;
  PLOT CBA*DF='A' PRED*DF='P' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: CBA=THETA0 ...ALL EXPERIMENTAL DATA';
PROC MEANS DATA=CLCLST ;
RUN;
TITLE2' ';
DATA CLCLST2;
  SET CLCLST;
  IF DF<900.;
RUN;
TITLE2'MODEL: CBA=THETA0+THETA1*DF ...SELECTED EXPERIMENTAL DATA';
PROC REG DATA=CLCLST2;
    MODEL CBA=DF / COVB CORRB CLM CLI P;
OUTPUT OUT=CBARES P=PRED R=RESID L95=L95 U95=U95;
PROC PLOT DATA=CBARES;
  PLOT CBA*DF='A' PRED*DF='P' U95*DF='-' L95*DF='-' / OVERLAY VPOS=50;
  PLOT RESID*DF / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*PRED / VREF=0 OVERLAY VPOS=50 ;
  PLOT RESID*RUNDATE / VREF=0 OVERLAY VPOS=50 ;
RUN;
TITLE2'MODEL: CBA=THETA0 ...SELECTED EXPERIMENTAL DATA';
PROC MEANS DATA=CLCLST2;
RUN;
```

Program 4. NLINIMSL FORTRAN

```
C     THE MAIN PROGRAM NLINIMSL READS IN THE HYDROLYSIS DATA AND CALLS  NLI00010
C     ON THE SUBROUTINES TO BEGIN THE PARAMETER OPTIMIZATION.           NLI00020
C                                                                       NLI00030
C     2 FILEDEFS: 1 IS PROCESS DATA, 2 IS THE OUTPUT FILE.              NLI00040
C                                                                       NLI00050
C                                                                       NLI00060
C     DICTIONARY OF VARIABLES                                           NLI00070
C                                                                       NLI00080
C     BCMC     BOND CONTRIBUTION OF CMC INITIAL EXPERIMENTAL            NLI00090
C     BNRE     BETA NRE TYPE BONDS (MOLES/L)                            NLI00100
C     BNREO(189)   BETA NRE TYPE BONDS (MOLES/L)                        NLI00110
C     BNRES    BETA NRES TYPE BONDS (MOLES/L)                           NLI00120
C     BNRESO(189)  BETA NRES TYPE BONDS (MOLES/L)                       NLI00130
C     BRE      BETA RE TYPE BONDS (MOLES/L)                             NLI00140
C     BREO(189)    BETA RE TYPE BONDS (MOLES/L)                         NLI00150
C     BRES     BETA RES TYPE BONDS (MOLES/L)                            NLI00160
C     BRESIN(189)  MAXIMUM NUMBER OF BRES BONDS (MOLES/L)               NLI00170
C     BRESO(189)   BETA RES TYPE BONDS (MOLES/L)                        NLI00180
C     BSNR     BETA SNR TYPE BONDS (MOLES/L)                            NLI00190
C     BSNRO(189)   BETA SNR TYPE BONDS (MOLES/L)                        NLI00200
C     BSR      BETA SR TYPE BONDS (MOLES/L)                             NLI00210
C     BSRINF(189)  MIN NUMBER OF BSR AND BSNR BONDS (MOLES/L)           NLI00220
C     BSRO(189)    BETA SR TYPE BONDS (MOLES/L)                         NLI00230
C     BSSO(189)    BETA SS TYPE BONDS (M)                               NLI00240
C     BU       BETA U TYPE BONDS (MOLES/L)                              NLI00250
C     BUO(189)     BETA U TYPE BONDS (MOLES/L)                          NLI00260
C     COV(28)      THE COVARIANCE MATRIX OF THE PARAMETERS              NLI00270
C     CORR(28)     THE CORRELATION MATRIX OF THE PARAMETERS             NLI00280
C     DELTA    THE 3RD CRITERION OF CONVERGENCE (IGNORED HERE)          NLI00290
C     DPCMC    DEGREE OF POLYMERIZATION OF INITIAL CMC                  NLI00300
C     DS       DEGREE OF SUBSTITUTION OF INITIAL CMC                    NLI00310
C     DT       STEP SIZE IN EULER DIF. EQ. SOLVER (S)                   NLI00320
C     DTO      NORMAL STEP SIZE IN EULER DIF. EQ. SOLVER (S)            NLI00330
C     EPS      MINIMUM RELATIVE DIFFERENCE IN SSQ BETWEEN 2 ITERATIONS  NLI00340
C     EXK1EX   MODEL TERM CONTAINING THE ENDO-GLUCANASE PARAMETERS      NLI00350
C     E1K1E1   MODEL TERM CONTAINING THE CELLOBIASE PARAMETERS          NLI00360
C     EXO(189)     ENDO-GLUCANASE CONC (I.U./L)                         NLI00370
C     E1O(189)     CELLOBIASE CONC (I.U./L)                             NLI00380
C     F(189)       RESIDUALS' VECTOR                                    NLI00390
C     GLUC(189)    GLUCOSE PRODUCED (EXP. VALUE OF RESPONSE VARIABLE)   NLI00400
C     G1       GLUCOSE CONC (MOLES/L)                                   NLI00410
C     G1INF(189)   GLUCOSE CONC. IN REACTOR AS TIME GOES TO INFINITY    NLI00420
C     G1O(189)     GLUCOSE CONC (MOLES/L)                               NLI00430
C     G2       CELLOBIOSE CONC (MOLES/L)                                NLI00440
C     G2O(189)     CELLOBIOSE CONC (MOLES/L)                            NLI00450
C     IDPCMC(189)  DEGREE OF POLYMERIZATION OF THE CMC (INTEGER)        NLI00460
C     IER      ERROR RETURN FROM IMSL ROUTINES                          NLI00470
C     INFER    INDICATES CRITERION(A) TYPE(S) TO SATISFY CONVERGENCE    NLI00480
C     IOPT     TYPE OF DESCENT (TRY 1)                                  NLI00490
C     IXJAC    INPUT ROW DIMENSION OF MATRIX XJAC                       NLI00500
C     KMEX     MICHAELIS CONSTANT OF ENDOGLUCANASE (MOLES/L)            NLI00510
C     KME1     MICHAELIS CONSTANT OF CELLOBIASE (MOLES/L)               NLI00520
C     KIEX     RATE CONSTANT OF ENDOGLUCANASE (MOLES/(I.U.*S))          NLI00530
C     KIG1     DISSOCIATION CONSTANT (M)                                NLI00540
C     K1E1     RATE CONSTANT OF CELLOBIASE (MOLES/(I.U.*S))             NLI00550
C     M        THE NUMBER OF DATA POINTS                                NLI00560
```

```
C     MAXFN   MAXIMUM NUMBER OF FUNCTION EVALUATIONS                    NLI00570
C     N       THE NUMBER OF PARAMETERS IN THIS MODEL                    NLI00580
C     NOVER(189)   NUMBER OF ITERATIONS WHERE G1(I)>G1INF(I)            NLI00590
C     NSIG    NUMBER OF SIGNIFICANT DIGITS DESIRED IN PARAMETERS        NLI00600
C     PARM(4)      ZXSSQ CONVERGENCE PARAMETERS                         NLI00610
C     TIME(189)    TIME OF REACTION EXPERIMENTAL (SECONDS)              NLI00620
C     TMSAVE  STORAGE OF THE ELAPSED SIMULATION TIME IN SIMILAR RUNS (S)NLI00630
C     TOL     SMALLEST TOLERABLE CONC. BEFORE ASSUMING 0 MOLES/L        NLI00640
C     UNRESR  MODEL TERM "BU+BNRE+BSR"                                  NLI00650
C     URESNR  MODEL TERM "BU+BRE+BSNR"                                  NLI00660
C     WORK    WORK VECTOR                                               NLI00670
C     X(N)    AN ARRAY OF THE PARAMETERS TO BE UNCHANGED BY THIS ROUTINENLI00680
C     XJAC    OUTPUT MATRIX(M BY N) CONTAINING THE APPROXIMATE JACOBIAN NLI00690
C     XJTJ    OUTPUT VECTOR CONTAINING  XJAC TRANSPOSED-- XJAC          NLI00700
      EXTERNAL EULERS                                                   NLI00710
      REAL*8 BNREO(189), BNRESO(189), BREO(189), BRESO(189), BSNRO(189),NLI00720
     1  BSRO(189), BUO(189), G10(189), G20(189), G1INF(189), BSSO(189)  NLI00730
      REAL*8 EXO(189), E10(189), GLUC(189), TIME(189), BSRINF(189)      NLI00740
      REAL*8 DS, DPCMC, CO, BCMC, BRESIN(189)                           NLI00750
      REAL*8 TOL, DTO, K1EX, KMEX, K1E1, KME1, KIG1                     NLI00760
      INTEGER I, IDPCMC(189), NOVER(189)                                NLI00770
      INTEGER M, N, IXJAC, NSIG, MAXFN, IOPT, INFER, IER                NLI00780
C     IF N=2                                                            NLI00790
      REAL*8 PARM(4), X(2), F(189), XJAC(189,2), XJTJ( 3), WORK(391)    NLI00800
C     IF N=3                                                            NLI00810
C     REAL*8 PARM(4), X(3), F(189), XJAC(189,3), XJTJ( 6), WORK(399)    NLI00820
C     IF N=4                                                            NLI00830
C     REAL*8 PARM(4), X(4), F(189), XJAC(189,4), XJTJ(10), WORK(408)    NLI00840
C     IF N=5                                                            NLI00850
C     REAL*8 PARM(4), X(5), F(189), XJAC(189,5), XJTJ(15), WORK(418)    NLI00860
      REAL*8 EPS, DELTA, SSQ, COV(28), CORR(28)                         NLI00870
      COMMON /ZSQ/ DTO,TOL,TIME,G10,G20,E10,EXO,GLUC,G1INF,BSRINF,BUO,  NLI00880
     1 BSSO,BREO,BNREO,BRESO,BNRESO,BSRO,BSNRO,BRESIN,NOVER             NLI00890
     2 /ZSQI/ IDPCMC                                                    NLI00900
C                                                                       NLI00910
C     READ THE PROGRAM PARAMETERS FOR THE ZXSSQ CONVERGENCE ROUTINE.    NLI00920
C                                                                       NLI00930
      READ(1,898) M, N, IXJAC, EPS, DELTA, IOPT                         NLI00940
  898 FORMAT(38X,I4,1X,I2,1X,I4,1X,D9.2,1X,D9.2,1X,I1)                  NLI00950
C                                                                       NLI00960
C     INPUT THE STARTING PARAMETER VALUES.                              NLI00970
C                                                                       NLI00980
      READ(1,899) K1EX, KMEX, K1E1, KME1, KIG1                          NLI00990
  899 FORMAT(5(6X,D12.5,1X))                                            NLI01000
      X(1) = K1EX                                                       NLI01010
C     X(2) = KMEX                                                       NLI01020
      X(2) = K1E1                                                       NLI01030
C     X(3) = KME1                                                       NLI01040
C     X(3) = KIG1                                                       NLI01050
C                                                                       NLI01060
C     INPUT INTEGRATION STEP SIZE, TOLERANCE AND MAX NO OF FUNC EVAL.   NLI01070
```

```
C                                                                       NLIO1080
      READ(1,903) DTO, TOL, NSIG, MAXFN                                 NLIO1090
  903 FORMAT(6X,D11.4,7X,D11.4,10X,I1,10X,I3)                           NLIO1100
      WRITE(2,904) DTO, TOL, NSIG, MAXFN                                NLIO1110
  904 FORMAT(' NONLINEAR LEAST SQUARES PARAMETER ESTIMATOR ZXSSQ',/,    NLIO1120
     1 ' DTO=',D11.4,'  TOL=',D11.4,'  NSIG=',I1,'  MAXFN=',I3)         NLIO1130
      WRITE(2,905)                                                      NLIO1140
  905 FORMAT(' ',74('-'))                                               NLIO1150
      WRITE(2,906)                                                      NLIO1160
  906 FORMAT(        4X,'K1EX',10X,'KMEX',10X,'K1E1',10X,'KME1',        NLIO1170
     1       10X,'KIG1',10X,'SSQ')                                      NLIO1180
      WRITE(2,905)                                                      NLIO1190
C                                                                       NLIO1200
C     READ IN THE DATA DOING CALCULATIONS AS EACH POINT IS INPUT.       NLIO1210
C                                                                       NLIO1220
      DO 900 I=1,189                                                    NLIO1230
        READ(1,901) TIME(I), G10(I), G20(I), E10(I), EX0(I),            NLIO1240
     1     DS, IDPCMC(I), BCMC, GLUC(I)                                 NLIO1250
        TIME(I)=TIME(I)*3600.D0                                         NLIO1260
        DPCMC = DFLOAT(IDPCMC(I))                                       NLIO1270
        IF(IDPCMC(I).GT.0) THEN                                         NLIO1280
          C0 = (1.-DS/2.)**2                                            NLIO1290
          G1INF(I) = G10(I)+2.D0*G20(I)+BCMC*C0*(C0+1.D0/(DPCMC-1.D0))  NLIO1300
          BSRINF(I) = BCMC*(DPCMC-2.D0)/(DPCMC-1.D0)*C0*(1.D0-C0)**2    NLIO1310
          BUO(I) = BCMC*(DPCMC-3.D0)/(DPCMC-1.D0)*C0**2                 NLIO1320
          BREO(I) = BCMC/(DPCMC-1.D0)*C0**2                             NLIO1330
          BNREO(I) = BREO(I)                                            NLIO1340
          BRESO(I) = BCMC/(DPCMC-1.D0)*C0*(1.D0-C0)                     NLIO1350
          BNRESO(I) = BRESO(I)                                          NLIO1360
          BSRO(I) = BCMC*C0*(1.D0-C0)*(DPCMC-2.D0)/(DPCMC-1.D0)         NLIO1370
          BSNRO(I) = BSRO(I)                                            NLIO1380
          BRESIN(I) = BRESO(I)+C0*BSRO(I)                               NLIO1390
          BSSO(I) = BCMC*(1.D0-C0)**2                                   NLIO1400
        ELSE                                                            NLIO1410
          G1INF(I) = G10(I)+2.D0*G20(I)                                 NLIO1420
          BSRINF(I)=0.D0                                                NLIO1430
          BUO(I)=0.D0                                                   NLIO1440
          BREO(I)=0.D0                                                  NLIO1450
          BNREO(I)=0.D0                                                 NLIO1460
          BRESO(I)=0.D0                                                 NLIO1470
          BNRESO(I)=0.D0                                                NLIO1480
          BSRO(I)=0.D0                                                  NLIO1490
          BSNRO(I)=0.D0                                                 NLIO1500
          BRESIN(I)=0.D0                                                NLIO1510
          BSSO(I)=0.D0                                                  NLIO1520
        END IF                                                          NLIO1530
        NOVER(I) = 0                                                    NLIO1540
  900 CONTINUE                                                          NLIO1550
  901 FORMAT(1X,F6.2,1X,F7.5,1X,F7.5,1X,D11.4,1X,D11.4,1X,F5.3,1X,I4,   NLIO1560
     1 1X,F8.5,1X,F7.5)                                                 NLIO1570
      CALL ZXSSQ(EULERS,M,N,NSIG,EPS,DELTA,MAXFN,IOPT,PARM,X,SSQ,F,XJAC,NLIO1580
     1 IXJAC,XJTJ,WORK,INFER,IER)                                       NLIO1590
C                                                                       NLIO1600
C     WRITE THE PARAMETER VALUES AND SSQ AT THE SSQ MINIMUM.            NLIO1610
C                                                                       NLIO1620
      WRITE(2,1602) X(1),       X(2),             SSQ                   NLIO1630
 1602 FORMAT('-PROGRAM HAS CONVERGED...HERE ARE THE PARMS:',            NLIO1640
     1         /,'OK1EX = ',D12.5,' MOLES*I.U.*S/L'                     NLIO1650
```

```
C     2         /,' KMEX = ',D12.5,' MOLES/L'                                NLI01660
      3         /,' K1E1 = ',D12.5,' MOLES*I.U.*S/L'                         NLI01670
C     4         /,' KME1 = ',D12.5,' MOLES/L'                                NLI01680
C     5         /,' KIG1 = ',D12.5,' MOLES/L'                                NLI01690
      6         /,' WITH A SUM OF SQ. RESID. = ',D12.5,' (MOLES/L)**2')      NLI01700
C                                                                            NLI01710
C     WRITE THE (JACOBIAN-TRANSPOSED JACOBIAN) MATRIX                        NLI01720
C                                                                            NLI01730
      WRITE(2,1612)                                                          NLI01740
 1612 FORMAT(' THE JACOBIAN-TRANSPOSED JACOBIAN MATRIX:')                    NLI01750
      DO 1705 I=1,N                                                          NLI01760
        WRITE(2,1605) (XJTJ(J),J=(I*(I-1)/2+1),(I*(1+I)/2))                  NLI01770
 1705 CONTINUE                                                               NLI01780
 1605 FORMAT (8(1X,D12.5))                                                   NLI01790
C                                                                            NLI01800
C     CALL THE IMSL ROUTINE LINV3P TO INVERSE THE XJTJ MATRIX                NLI01810
C                                                                            NLI01820
      CALL LINV3P(XJTJ,1.DO,1,N,IER)                                         NLI01830
C                                                                            NLI01840
C     CALCULATE AND WRITE THE COVARIANCE MATRIX OF THE PARAMETERS            NLI01850
C                                                                            NLI01860
      DO 1706 I=1,(N*(N+1)/2)                                                NLI01870
 1706   COV(I) = XJTJ(I)*SSQ/DFLOAT(M-N)                                     NLI01880
      WRITE(2,1606)                                                          NLI01890
 1606 FORMAT (' VARIANCE-COVARIANCE MATRIX OF THE PARAMETERS:')              NLI01900
      DO 1707 I=1,N                                                          NLI01910
        WRITE(2,1605) (COV(J),J=(I*(I-1)/2+1),(I*(1+I)/2))                   NLI01920
 1707 CONTINUE                                                               NLI01930
C                                                                            NLI01940
C     CALCULATE AND WRITE THE CORRELATION MATRIX OF THE PARAMETERS.          NLI01950
C                                                                            NLI01960
      WRITE(2,1607)                                                          NLI01970
 1607 FORMAT (' CORRELATION MATRIX OF THE PARAMETERS:')                      NLI01980
      DO 1710 I=1,N                                                          NLI01990
      DO 1709 J=1,I                                                          NLI02000
        CORR(I*(I-1)/2+J)=COV(I*(I-1)/2+J)/                                  NLI02010
     1  DSQRT(COV(J*(1+J)/2)*COV(I*(1+I)/2))                                 NLI02020
 1709   CONTINUE                                                            NLI02030
        WRITE(2,1805) (CORR(JJ),JJ=(I*(I-1)/2+1),(I*(1+I)/2))               NLI02040
 1805   FORMAT (8(1X,F6.3))                                                 NLI02050
 1710 CONTINUE                                                               NLI02060
C                                                                            NLI02070
C     CALCULATE THE PARAMETER PRECISION (95% CONFIDENCE INTERVALS)           NLI02080
C                                                                            NLI02090
      WRITE(2,1611)                                                          NLI02100
 1611 FORMAT(' PARAMETERS AND THEIR ASSOCIATED 95% CONF. INTERVALS:')        NLI02110
      DO 1609 I=1,N                                                          NLI02120
        WRITE(2,1610) I, X(I), DSQRT(COV(I*(1+I)/2))*1.97                    NLI02130
 1609 CONTINUE                                                               NLI02140
 1610 FORMAT(' THETA',I1,' = ',D12.5,' PLUS OR MINUS ',D12.5)                NLI02150
      STOP                                                                   NLI02160
      END                                                                    NLI02170
C                                                                            NLI02180
C     INSERT THE PROPER EULER SUBROUTINE HERE                                NLI02190
      SUBROUTINE EULERS(X,N,N,F)                                             NLI02200
C                                                                            NLI02210
C     SUBROUTINE EULER13: SOLVES FOR THE 13TH MODEL PARM ESTIMATES           NLI02220
C     SUBROUTINE CALLED BY IMSL ROUTINE ZXSSQ.                               NLI02230
```

```
C     SUBROUTINE ASSUMES ONLY G1 CLEAVES CELLOBIOSE.                    NLIO2240
C     SUBROUTINE ASSUMES KNEX IS ZERO.                                  NLIO2250
C     SUBROUTINE ASSUMES KME1 IS INFINITELY LARGER THAN G2+BNRE+BNRES   NLIO2260
C     SUBROUTINE ASSUMES THAT ALL BETA BONDS FORM A COMPLEX WITH EX     NLIO2270
C     SUBROUTINE ASSUMES THAT E2 TRANSFORMS BNRE INTO G2 INF. QUICKLY   NLIO2280
C     THE SUBROUTINE FCN SOLVES THE DIFFERENTIAL EQUATIONS              NLIO2290
C     WITH AN EULER ONE-STEP METHOD                                     NLIO2300
C     THEN CALCULATES THE RESIDUALS OF A MODEL                          NLIO2310
C     GIVEN THAT THE PARAMETER VECTOR X(2) IS KNOWN.                    NLIO2320
C     SSQ IS BASED ON 189 DATA POINTS COLLECTED FOR CMC HYDROLYSIS.     NLIO2330
C                                                                       NLIO2340
      REAL*8 BNRE, BNRES, BRE, BRES, BSNR, BSR, BU, G1, G2              NLIO2350
      REAL*8 DBNRE, DBNRES, DBRE, DBRES, DBSNR, DBSR, DBU, DG1, DG2     NLIO2360
      REAL*8 BNREO(189), BNRESO(189), BREO(189), BRESO(189), BSNRO(189),NLIO2370
     1 BSRO(189), BUO(189), G1O(189), G2O(189), G1INF(189), BSSO(189)   NLIO2380
      REAL*8 EXO(189), E1O(189), GLUC(189), TIME(189), BSRINF(189)      NLIO2390
      REAL*8 KMEX, KME1, K1EX, K1E1, BRESIN(189), RESID(189), PRED2(189)NLIO2400
      REAL*8 URESNR, UNRESR, EXK1EX, E1K1E1                             NLIO2410
      REAL*8 DT, TOL, TMSAVE, DTO, X(2), SSR, F(189), DG22              NLIO2420
      INTEGER I, N, IDPCMC(189) , ITERS, NOVER(189), M                  NLIO2430
      COMMON /ZRESID/ PRED2                                             NLIO2440
      COMMON /ZSQ/ DTO,TOL,TIME,G1O,G2O,E1O,EXO,GLUC,G1INF,BSRINF,BUO,  NLIO2450
     1 BSSO,BREO,BNREO,BRESO,BNRESO,BSRO,BSNRO,BRESIN,NOVER             NLIO2460
     2 /ZSQI/ IDPCMC                                                    NLIO2470
C                                                                       NLIO2480
      SSR = 0.DO                                                        NLIO2490
      K1EX = X(1)                                                       NLIO2500
      KMEX = 0.DO                                                       NLIO2510
      K1E1 = X(2)                                                       NLIO2520
      KME1 = 0.DO                                                       NLIO2530
      DO 2000 I=1,189                                                   NLIO2540
C                                                                       NLIO2550
C       TEST FOR E1, E2 AND EX NIL IN WHICH CASE, G1=G1(0)              NLIO2560
C                                                                       NLIO2570
        IF(I.EQ.139.OR.I.EQ.179) THEN                                   NLIO2580
          G1 = 0.DO                                                     NLIO2590
          GOTO 1999                                                     NLIO2600
        END IF                                                          NLIO2610
        DT = DTO                                                        NLIO2620
C                                                                       NLIO2630
C       TEST FOR A RUN THAT IS IDENTICAL TO PRECEEDING BUT FOR TIME(I)  NLIO2640
```

```
C                                                                       NLI02650
      IF(I.GT.1) THEN                                                   NLI02660
        IF(TIME(I).GE.TIME(I-1)                                         NLI02670
     1    .AND.G10(I).EQ.G10(I-1).AND.G20(I).EQ.G20(I-1)                NLI02680
     2    .AND.EXO(I).EQ.EXO(I-1).AND.E10(I).EQ.E10(I-1)                NLI02690
     3    .AND.BUO(I).EQ.BUO(I-1).AND.BNRESO(I).EQ.BNRESO(I-1)) THEN    NLI02700
          ITERS=IFIX(SNGL((TIME(I)-TIME(I-1))/DT))                      NLI02710
          TMSAVE=TIME(I-1)+DT*DFLOAT(ITERS)                             NLI02720
        ELSE                                                            NLI02730
          ITERS=IFIX(SNGL(TIME(I)/DT))                                  NLI02740
          TMSAVE=DT*DFLOAT(ITERS)                                       NLI02750
          G1=G10(I)                                                     NLI02760
          G2 = G20(I)                                                   NLI02770
          BU=BUO(I)                                                     NLI02780
          BRE=BREO(I)                                                   NLI02790
          BNRE=BNREO(I)                                                 NLI02800
          BRES=BRESO(I)                                                 NLI02810
          BNRES=BNRESO(I)                                               NLI02820
          BSR=BSRO(I)                                                   NLI02830
          BSNR=BSNRO(I)                                                 NLI02840
          EX=EXO(I)                                                     NLI02850
          E1=E10(I)                                                     NLI02860
        END IF                                                          NLI02870
      ELSE IF(I.EQ.1) THEN                                              NLI02880
        ITERS=IFIX(SNGL(TIME(I)/DT))                                    NLI02890
        TMSAVE=DT*DFLOAT(ITERS)                                         NLI02900
        G1=G10(I)                                                       NLI02910
        G2 = G20(I)                                                     NLI02920
        BU=BUO(I)                                                       NLI02930
        BRE=BREO(I)                                                     NLI02940
        BNRE=BNREO(I)                                                   NLI02950
        BRES=BRESO(I)                                                   NLI02960
        BNRES=BNRESO(I)                                                 NLI02970
        BSR=BSRO(I)                                                     NLI02980
        BSNR=BSNRO(I)                                                   NLI02990
        EX=EXO(I)                                                       NLI03000
        E1=E10(I)                                                       NLI03010
      END IF                                                            NLI03020
      DG22=0.D0                                                         NLI03030
C                                                                       NLI03040
C       TEST FOR A BROTH CONTAINING NO CMC                              NLI03050
C                                                                       NLI03060
      IF(IDPCMC(I).EQ.0) THEN                                           NLI03070
        IF(ITERS.EQ.0) GOTO 1200                                        NLI03080
        DO 1000 J=1,ITERS                                               NLI03090
          E1K1E1=E1*K1E1                                                NLI03100
          DG2 = -G2*E1K1E1                                              NLI03110
          DG1 = -2.D0*DG2                                               NLI03120
          G1=G1+DG1*DT                                                  NLI03130
C                                                                       NLI03140
C         TEST FOR GLUCOSE CONCENTRATION HAVING REACHED ITS MAXIMUM.    NLI03150
```

```
C                                                                       NLI03160
          IF(G1.GE.G1INF(I)) THEN                                       NLI03170
            G1 = G1INF(I)                                               NLI03180
            NOVER(I) = NOVER(I)+1                                       NLI03190
            GOTO 1999                                                   NLI03200
          END IF                                                        NLI03210
          G2=G2+DG2*DT                                                  NLI03220
          IF(G2.LT.TOL) G2=0.DO                                         NLI03230
 1000     CONTINUE                                                      NLI03240
C                                                                       NLI03250
C     ,   ADJUST THE FINAL INTERVAL SIZE TO ARRIVE AT THE EXACT TIME.   NLI03260
C                                                                       NLI03270
 1200     DT = TIME(I)-TMSAVE                                           NLI03280
          E1K1E1=E1*K1E1                                                NLI03290
          DG2 = -G2*E1K1E1                                              NLI03300
          DG1 = -2.DO*DG2                                               NLI03310
        ELSE                                                            NLI03320
          IF(ITERS.EQ.0) GOTO 1201                                      NLI03330
          DO 1001 J=1,ITERS                                             NLI03340
            URESNR = BU+BRE+BSNR                                        NLI03350
            UNRESR = BU+BNRE+BSR                                        NLI03360
            E1K1E1=E1*K1E1                                              NLI03370
            DBU=-BU*BNRE/URESNR*E1K1E1                                  NLI03380
            DG2=(BNRE*BRE/URESNR-G2)*E1K1E1                             NLI03390
            DG1=(2.DO*G2+BNRE+BNRES)*E1K1E1                             NLI03400
            DBRE=-BNRE*BRE/URESNR*E1K1E1                                NLI03410
            DBRES=0.DO                                                  NLI03420
            DBNRE=BNRE*(BU/URESNR-1.DO)*E1K1E1                          NLI03430
            DBNRES=(BNRE*BSNR/URESNR-BNRES)*E1K1E1                      NLI03440
            DBSR=0.DO                                                   NLI03450
            DBSNR=-BSNR*BNRE/URESNR*E1K1E1                              NLI03460
            IF(BRE+BNRE+BU.GT.0.DO) THEN                                NLI03470
            EXK1EX=EX*K1EX/(G2+BNRE+BRE+BU+BRES+BNRES+BSR+BSNR+BSSO(I)) NLI03480
            DBU=DBU-BU*((BU+BNRE)/URESNR+(BU+BRE)/UNRESR+1.DO)*EXK1EX   NLI03490
            DG2=DG2+((BU*BRE+BNRE*BRE)/URESNR+(BU*BNRE+BRE*BNRE)/UNRESR)NLI03500
     1      *EXK1EX                                                    NLI03510
            DG1=DG1+(BRE+BNRE)*EXK1EX                                   NLI03520
            DBRE=DBRE+                                                  NLI03530
     1      (BU*(BU+BRE)/UNRESR-BRE-BRE*(BU+BNRE)/URESNR)*EXK1EX       NLI03540
            DBRES=DBRES+(BU+BRE)/UNRESR*BSR*EXK1EX                      NLI03550
            DBNRE=DBNRE+(BU*(BU+BNRE)/URESNR-BNRE-BNRE*(BU+BRE)/UNRESR  NLI03560
     1           )*EXK1EX                                               NLI03570
            DBNRES=DBNRES+(BU+BNRE)/URESNR*BSNR*EXK1EX                  NLI03580
            DBSR=DBSR-(BU+BRE)/UNRESR*BSR*EXK1EX                        NLI03590
            DBSNR=DBSNR-(BU+BNRE)/URESNR*BSNR*EXK1EX                    NLI03600
            END IF                                                      NLI03610
            BU=BU+DBU*DT                                                NLI03620
            BRE=BRE+DBRE*DT                                             NLI03630
            BRES=DMIN1(BRES+DBRES*DT,BRESIN(I))                         NLI03640
            BNRE=BNRE+DBNRE*DT                                          NLI03650
            BNRES=BNRES+DBNRES*DT                                       NLI03660
            BSR=DMAX1(BSR+DBSR*DT,BSRINF(I))                            NLI03670
            BSNR=DMAX1(BSNR+DBSNR*DT,BSRINF(I))                         NLI03680
            IF(BU.LT.TOL) BU=0.DO                                       NLI03690
            IF(BRE.LT.TOL) BRE=0.DO                                     NLI03700
            IF(BNRE.LT.TOL.AND.DBNRE.LE.0.DO) BNRE=0.DO                 NLI03710
            IF(BNRES.LT.TOL.AND.DBNRES.LE.0.DO) BNRES=0.DO              NLI03720
C                                                                       NLI03730
```

```
C          NOW ADD TO THE PROGRAM TO ACCOUNT FOR THE PRESENCE OF E2    NLI03740
C                                                                      NLI03750
           IF(BNRE.GT.0.D0) THEN                                       NLI03760
             URESNR=BU+BRE+BSNR                                        NLI03770
             DBRE=-BNRE*BRE/URESNR-BNRE*BU/URESNR*BRE/URESNR           NLI03780
             DBSNR=-BNRE*BSNR/URESNR-BNRE*BU/URESNR*BSNR/URESNR        NLI03790
             DBU=-BNRE*BU/URESNR-BNRE*BU/URESNR*BU/URESNR              NLI03800
             DBNRE=-BNRE+BNRE*BU/URESNR*BU/URESNR                      NLI03810
             DG22=BNRE+BNRE*BU/URESNR*BRE/URESNR                       NLI03820
             DBNRES=BNRE*BU/URESNR*BSNR/URESNR                         NLI03830
             BRE=BRE+DBRE                                              NLI03840
             BSNR=BSNR+DBSNR                                           NLI03850
             BU=BU+DBU                                                 NLI03860
             BNRES=BNRES+DBNRES                                        NLI03870
             BNRE=BNRE+DBNRE                                           NLI03880
             BSNR=DMAX1(BSNR+DBSNR*DT,BSRINF(I))                       NLI03890
             IF(BU.LT.TOL) BU=0.D0                                     NLI03900
             IF(BRE.LT.TOL) BRE=0.D0                                   NLI03910
             IF(BNRE.LT.TOL) BNRE=0.D0                                 NLI03920
             IF(BNRES.LT.TOL) BNRES=0.D0                               NLI03930
           ELSE                                                        NLI03940
             DG22=0.D0                                                 NLI03950
           END IF                                                      NLI03960
           G1=G1+DG1*DT                                                NLI03970
C                                                                      NLI03980
C          TEST FOR GLUCOSE CONCENTRATION HAVING REACHED ITS MAXIMUM.  NLI03990
C                                                                      NLI04000
           IF(G1.GE.G1INF(I)) THEN                                     NLI04010
             G1 = G1INF(I)                                             NLI04020
             NOVER(I) = NOVER(I)+1                                     NLI04030
             GOTO 1999                                                 NLI04040
           END IF                                                      NLI04050
           G2=G2+DG2*DT+DG22                                           NLI04060
           IF(G2.LT.TOL.AND.(DG2*DT+DG22).LE.0.D0) G2=0.D0             NLI04070
 1001    CONTINUE                                                      NLI04080
C                                                                      NLI04090
C        ADJUST THE FINAL INTERVAL SIZE TO ARRIVE AT THE EXACT TIME.   NLI04100
```

```
C                                                                       NLI04110
 1201     DT = TIME(I)-TMSAVE                                           NLI04120
          URESNR = BU+BRE+BSNR                                          NLI04130
          UNRESR = BU+BNRE+BSR                                          NLI04140
          E1K1E1=E1*K1E1                                                NLI04150
          DBU=-BU*BNRE/URESNR*E1K1E1                                    NLI04160
          DG2=(BNRE*BRE/URESNR-G2)*E1K1E1                               NLI04170
          DG1=(2.DO*G2+BNRE+BNRES)*E1K1E1                               NLI04180
          DBRE=-BNRE*BRE/URESNR*E1K1E1                                  NLI04190
          DBRES=0.DO                                                    NLI04200
          DBNRE=BNRE*(BU/URESNR-1.DO)*E1K1E1                            NLI04210
          DBNRES=(BNRE*BSNR/URESNR-BNRES)*E1K1E1                        NLI04220
          DBSR=0.DO                                                     NLI04230
          DBSNR=-BSNR*BNRE/URESNR*E1K1E1                                NLI04240
          IF(BRE+BNRE+BU.GT.0.DO) THEN                                  NLI04250
          EXK1EX=EX*K1EX/(G2+BNRE+BRE+BU+BRES+BNRES+BSR+BSNR+BSSO(I))   NLI04260
          DBU=DBU-BU*((BU+BNRE)/URESNR+(BU+BRE)/UNRESR+1.DO)*EXK1EX     NLI04270
          DG2=DG2+((BU*BRE+BNRE*BRE)/URESNR+(BU*BNRE+BRE*BNRE)/UNRESR)  NLI04280
     1     *EXK1EX                                                       NLI04290
          DG1=DG1+(BRE+BNRE)*EXK1EX                                     NLI04300
          DBRE=DBRE+                                                    NLI04310
     1     (BU*(BU+BRE)/UNRESR-BRE-BRE*(BU+BNRE)/URESNR)*EXK1EX          NLI04320
          DBRES=DBRES+(BU+BRE)/UNRESR*BSR*EXK1EX                        NLI04330
          DBNRE=DBNRE+(BU*(BU+BNRE)/URESNR-BNRE-BNRE*(BU+BRE)/UNRESR    NLI04340
     1          )*EXK1EX                                                NLI04350
          DBNRES=DBNRES+(BU+BNRE)/URESNR*BSNR*EXK1EX                    NLI04360
          DBSR=DBSR-(BU+BRE)/UNRESR*BSR*EXK1EX                          NLI04370
          DBSNR=DBSNR-(BU+BNRE)/URESNR*BSNR*EXK1EX                      NLI04380
          END IF                                                        NLI04390
        BU=BU+DBU*DT                                                    NLI04400
        BRE=BRE+DBRE*DT                                                 NLI04410
        BRES=DMIN1(BRES+DBRES*DT,BRESIN(I))                             NLI04420
        BNRE=BNRE+DBNRE*DT                                              NLI04430
        BNRES=BNRES+DBNRES*DT                                           NLI04440
        BSR=DMAX1(BSR+DBSR*DT,BSRINF(I))                                NLI04450
        BSNR=DMAX1(BSNR+DBSNR*DT,BSRINF(I))                             NLI04460
        IF(BU.LT.TOL) BU=0.DO                                           NLI04470
        IF(BRE.LT.TOL.AND.DBRE.LE.0.DO) BRE=0.DO                        NLI04480
        IF(BNRE.LT.TOL.AND.DBNRE.LE.0.DO) BNRE=0.DO                     NLI04490
        IF(BNRES.LT.TOL.AND.DBNRES.LE.0.DO) BNRES=0.DO                  NLI04500
C                                                                       NLI04510
C       NOW ADD TO THE PROGRAM TO ACCOUNT FOR THE PRESENCE OF E2        NLI04520
```

```
C                                                                       NLI04530
          IF(BNRE.GT.0.DO) THEN                                         NLI04540
            URESNR=BU+BRE+BSNR                                          NLI04550
            DBRE=-BNRE*BRE/URESNR-BNRE*BU/URESNR*BRE/URESNR             NLI04560
            DBSNR=-BNRE*BSNR/URESNR-BNRE*BU/URESNR*BSNR/URESNR          NLI04570
            DBU=-BNRE*BU/URESNR-BNRE*BU/URESNR*BU/URESNR                NLI04580
            DBNRE=-BNRE+BNRE*BU/URESNR*BU/URESNR                        NLI04590
            DG22=BNRE+BNRE*BU/URESNR*BRE/URESNR                         NLI04600
            DBNRES=BNRE*BU/URESNR*BSNR/URESNR                           NLI04610
            BRE=BRE+DBRE                                                NLI04620
            BSNR=BSNR+DBSNR                                             NLI04630
            BU=BU+DBU                                                   NLI04640
            BNRES=BNRES+DBNRES                                          NLI04650
            BNRE=BNRE+DBNRE                                             NLI04660
            BSNR=DMAX1(BSNR+DBSNR*DT,BSRINF(I))                         NLI04670
            IF(BU.LT.TOL) BU=0.DO                                       NLI04680
            IF(BRE.LT.TOL) BRE=0.DO                                     NLI04690
            IF(BNRE.LT.TOL) BNRE=0.DO                                   NLI04700
            IF(BNRES.LT.TOL) BNRES=0.DO                                 NLI04710
          ELSE                                                          NLI04720
            DG22=0.DO                                                   NLI04730
          END IF                                                        NLI04740
        END IF                                                          NLI04750
        G2=G2+DG2*DT+DG22                                               NLI04760
        G1=G1+DG1*DT                                                    NLI04770
        IF(G1.GE.G1INF(I)) THEN                                         NLI04780
          G1=G1INF(I)                                                   NLI04790
          NOVER(I)=NOVER(I)+1                                           NLI04800
        END IF                                                          NLI04810
        IF(G2.LT.TOL.AND.(DG2*DT+DG22).LE.0.DO) G2=0.DO                 NLI04820
 1999   F(I)=GLUC(I)-G1                                                 NLI04830
        PRED2(I)=G2                                                     NLI04840
        SSR=SSR+F(I)**2                                                 NLI04850
 2000 CONTINUE                                                          NLI04860
C                                                                       NLI04870
C     PRINT THE PARAMETERS AND THE SSR FOR THIS ENTER OF THE EULERS SUB.NLI04880
C                                                                       NLI04890
      WRITE(2,1603) K1EX,       K1E1,             SSR                   NLI04900
 1603 FORMAT(1X,D13.6,15X,D13.6,  14X      ,15X,D13.6)                  NLI04910
      RETURN                                                            NLI04920
      END                                                               NLI04930
```