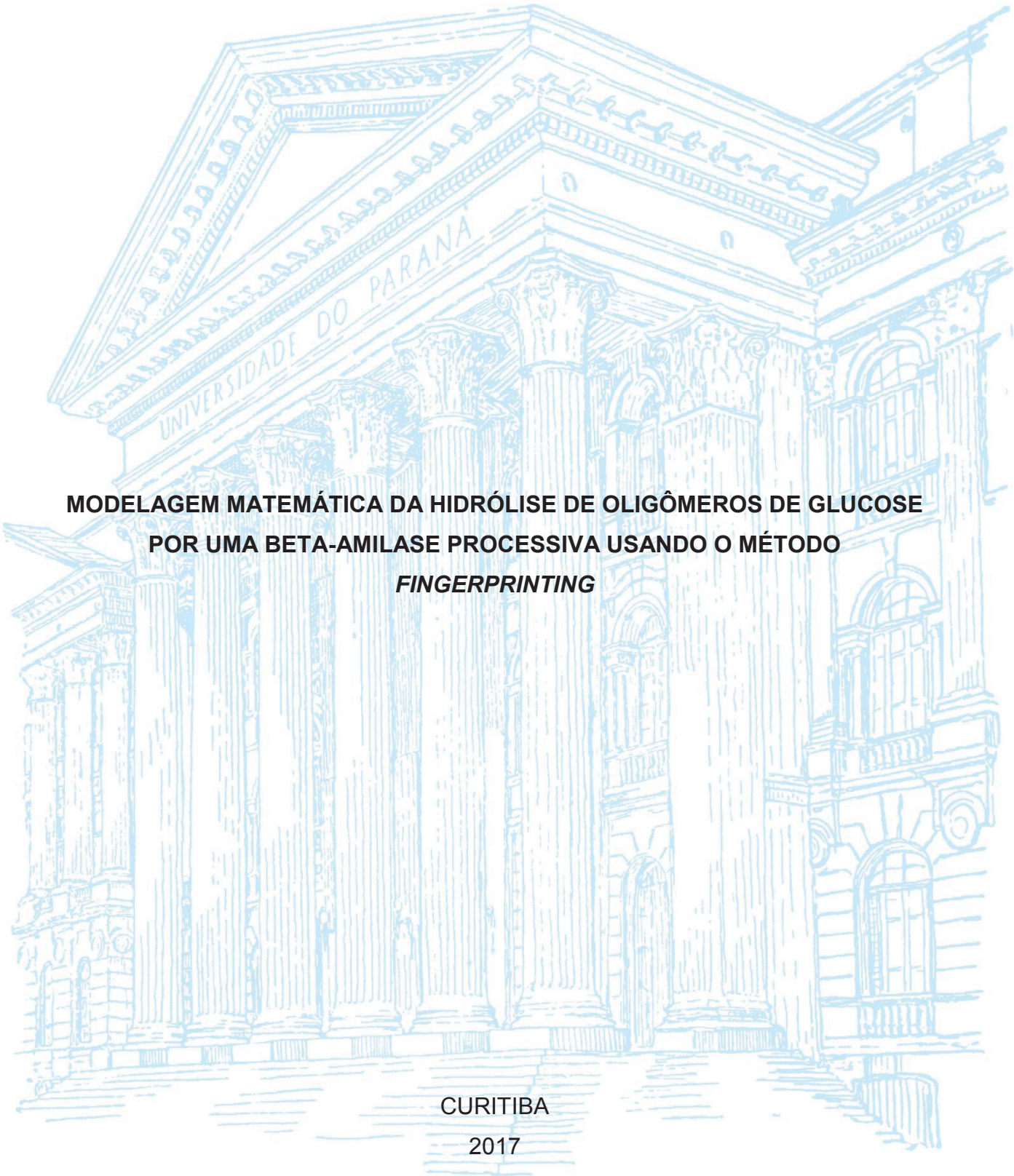
UNIVERSIDADE FEDERAL DO PARANÁ

GERSON RITZMANN DE MELLO



MODELAGEM MATEMÁTICA DA HIDRÓLISE DE OLIGÔMEROS DE GLUCOSE POR UMA BETA-AMILASE PROCESSIVA USANDO O MÉTODO *FINGERPRINTING*

CURITIBA

2017

GERSON RITZMANN DE MELLO

MODELAGEM MATEMÁTICA DA HIDRÓLISE DE OLIGÔMEROS DE GLUCOSE POR UMA BETA-AMILASE PROCESSIVA USANDO O MÉTODO *FINGERPRINTING*

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências (Bioquímica), pelo Programa de Pós-Graduação em Ciências (Bioquímica), do Departamento de Bioquímica e Biologia Molecular, Setor de Ciências Biológicas, da Universidade Federal do Paraná.

Orientador: Prof. Dr. David Alexander Mitchell
Co-orientadora: Profa. Dra. Nadia Krieger

CURITIBA

2017

Universidade Federal do Paraná
Sistema de Bibliotecas

Mello, Gerson Ritzmann de
Modelagem matemática da hidrólise de oligômeros de glucose por uma beta-amilase processiva usando o método *fingerprinting*. / Gerson Ritzmann de Mello. – Curitiba, 2017.
88 f.: il. ; 30cm.

Orientador: David Alexander Mitchell
Coorientador: Nadia Krieger

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Bioquímica.

1. Hidrolise 2. Glicose 3. Amilase Título II. Mitchell, David Alexander III. Krieger, Nadia IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Bioquímica.

CDD (20. ed.) 574.192

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
Setor CIÊNCIAS BIOLÓGICAS
Programa de Pós Graduação em CIÊNCIAS (BIOQUÍMICA)
Código CAPES: 40001016003P2

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS (BIOQUÍMICA) da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **GERSON RITZMANN DE MELLO**, intitulada: "**Modelagem matemática da hidrólise de oligômeros de glucose por uma beta-amilase processiva usando o método *fingerprinting***", após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua aprovação.

CURITIBA, 30 de Maio de 2017.

DAVID ALEXANDER MITCHELL
Presidente da Banca Examinadora (UFPR)

WELLINGTON BALMANT
Avaliador Externo (UFPR)

FERNANDO AUGUSTO PEDERSEN VOLL
Avaliador Externo (UFPR)

Dedico este trabalho aos meus pais, Edison e Cacilda, por toda a força que me deram, sem a qual eu não teria chegado aqui.

Dedico este trabalho também para Janaina, que tanto me ajudou, me inspirou e me fez ser uma pessoa melhor.

AGRADECIMENTOS

Ao Princípio Criador.

À Janaina, por todo o amor, carinho, apoio e ajuda nos momentos difíceis.

Aos meus pais, Edison e Cacilda, que sempre incentivaram e apoiaram meus estudos e me deram suporte.

Ao meu irmão e minha cunhada, Jefferson e Bruna, por toda a força que sempre me deram.

Aos meus amigos, antigos e novos, companheiros desta jornada com quem dividi momentos de alegria e dificuldades. Saibam que sempre poderão contar comigo, assim como tenho convicção que posso contar com vocês.

Ao Prof. David Mitchell, meu orientador, e à Profa. Nadia Krieger, minha co-orientadora, por todo o conhecimento transmitido. Obrigado por me emprestarem seus ombros, que me permitiram olhar mais longe e vislumbrar grandes oportunidades.

Aos professores e técnicos do Programa de Pós-Graduação em Ciências - Bioquímica. Um agradecimento especial à Profa. Joana Silveira e ao Prof. Miguel Noseda.

Ao CNPq, pelo auxílio financeiro.

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”

(Arthur Schopenhauer)

RESUMO

O fenômeno da processividade pode ser observado em muitas enzimas envolvidas na sacarificação de polissacarídeos. As probabilidades de ataque processivo, assim como as constantes de especificidade, variam de acordo com o tamanho do substrato. Quando os substratos são várias ordens de grandeza maiores que a enzima, como ocorre na hidrólise de polissacarídeos, as probabilidades de ataque processivo e a especificidade da enzima por moléculas de diferentes tamanhos pouco variam. Porém, quando a hidrolase está agindo em oligossacarídeos, a determinação destes parâmetros torna-se interessante, pois a enzima tem valores para estes parâmetros que são signficativamente diferentes para substratos que diferem no grau de polimerização por poucas unidades. A vantagem de se conhecer as constantes de especificidade relativas e as probabilidades de ataque processivo é que elas podem ser aplicadas em modelos matemáticos da reação de hidrólise que serviriam, por exemplo, para dimensionar biorreatores, para otimizar bioprocessos em uma biorrefinaria ou, ainda, para guiar experimentos de mutação em hidrolases, onde estes parâmetros seriam considerados na seleção do mutante desejado para ser submetido a novas rodadas de mutação. Entretanto, ainda não há um modelo matemático que seja válido para toda a extensão da reação de hidrólise, nem métodos que extraiam estes parâmetros simultaneamente a partir de perfis completos de hidrólise. Nos modelos disponíveis na literatura, os autores assumiram simplificações que os tornam válidos apenas para frações específicas da reação ou dificultam sua aplicação para sistemas mais complexos. O objetivo desta dissertação de mestrado foi, portanto, expandir o método Fingerprinting para extrair, além das constantes de especificidade relativas, as probabilidades de ataque processivo de perfis de hidrólise de oligossacarídeos lineares e elaborar modelos corretos para representar reações com processividade. A eficiência desta adaptação do método Fingerprinting foi comprovada pela sua aplicação em três estudos de caso de β-amilases hidrolisando maltooligossacarídeos, de complexidade crescente, utilizando dados experimentais retirados da literatura. O primeiro estudo de caso trata da hidrólise de maltoheptaose, o sistema mais simples onde se pode observar processividade. O segundo estudo de caso foi desenvolvido sobre resultados da hidrólise de maltooctaose, um sistema com duas possíveis etapas processivas, onde o produto final da extremidade redutora é o mesmo produto removido da extremidade não-redutora pela enzima após cada ataque, o que exige uma outra abordagem para calcular a fração de hidrólise. O terceiro estudo de caso aborda a hidrólise de maltoundecaose, um sistema com três possíveis etapas processivas, onde o modelo foi ajustado a mais de um perfil de hidrólise simultaneamente. Como resultados deste trabalho, foram desenvolvidos modelos matemáticos conceitualmente corretos e válidos para toda a extensão da reação, em função do tempo e da fração de reação, para sistemas onde uma exoenzima hidrolisa um oligômero linear de forma processiva. Obteve-se, também, um método para se determinar constantes de especificidades relativas e probabilidades de ataque processivo que é mais simples de se aplicar do que outros métodos disponíveis na literatura, pois exige um número menor de experimentos.

Palavras-chave: Processividade. Constantes de especificidade. Modelagem matemática. Fingerprinting. Hidrolases.

ABSTRACT

Processive action can be observed in many enzymes involved in the saccharification of polysaccharides. The probabilities of processive action, as well as the specificity constants, vary according to the size of the substrate. When substrates are several orders of magnitude larger than the enzyme, as occurs in polysaccharide hydrolysis, the probabilities of processive action and the specificity constants of the enzyme for molecules of different sizes do not vary much. However, when the hydrolase is acting on oligosaccharides, the determination of these parameters becomes interesting, since the enzyme has significantly different values of these parameters for substrates that differ in degree of polymerization by only a few units. The advantage of knowing the relative specificity constants and the probabilities of processive action is that they can be applied in mathematical models of the hydrolysis reaction that would serve, for example, to size bioreactors, to optimize bioprocesses in a biorefinery or to guide mutagenesis experiments in hydrolases, where these parameters would be considered in the selection of the desired mutant to undergo further rounds of mutation. However, prior to the current work, there was no valid mathematical model that described the full extent of the hydrolysis reaction, nor any method to extract the parameters simultaneously from complete hydrolysis profiles. In the models that were available in the literature, the authors made simplifying assumptions that limited the models to describing specific fractions of the reaction or prevented their application in more complex systems. The objective of this work was, therefore, to expand the Fingerprinting method to extract both relative specificity constants and the probabilities of processive action from hydrolysis profiles of linear oligosaccharides and to elaborate correct models to represent reactions with processivity. The suitability of this adaptation of the Fingerprinting method was proven by its application in three case studies of increasing complexity involving β-amylases hydrolysing maltooligosaccharides, using experimental data from the literature. The first case study deals with the hydrolysis of maltoheptaose, the simplest system in which processivity can be observed. The second case study involves the hydrolysis of maltooctaose, a system with two possible processive steps, where the end product of the reducing end is the same product as that removed from the non-reducing end by the enzyme in each attack, which requires a different approach to calculating the fractional reaction extent. The third case study addresses the hydrolysis of maltoundecaose, a system with three possible processive steps, where the model was fitted to more than one hydrolysis profile simultaneously. In this work, conceptually correct mathematical models were developed, being valid for the entire extent of the reaction, as functions of time and of fractional reaction extent, for systems where an exoenzyme hydrolyzes a linear oligomer processively. A method was also obtained to determine relative specificity constants and probabilities of processive action that is simpler to apply than other methods available in the literature, since it requires a smaller number of experiments.

Key-words: Processivity. Specificity constants. Mathematical model. Fingerprinting. Hydrolases.

LISTA DE FIGURAS

LISTA DE FIGURAS DO ARTIGO 1

LISTA DE TABELAS DO ARTIGO 1

LISTA DE SÍMBOLOS

® – marca registrada

Π – produtório de números

Σ – somatório de números

[Gi] – concentração de Gi

[S_i] – concentração do substrato genérico de grau de polimerização “i”

[Y] – concentração da espécie “Y”

D_m – denominador das equações de velocidade para uma reação iniciada com S_m

E – enzima

f – grau de polimerização do produto residual final (o menor produto residual, que não pode ser atacado, contém a extremidade redutora da molécula do substrato inicial)

F_m – fração de reação para uma reação iniciada com S_m

Gi – malto-oligossacarídeo genérico contendo "i" unidades de glucose

Gi – concentração relativa de Gi

$k_{cat(i)}$ – constante de velocidade da reação de conversão do substrato ligado à enzima e liberação do produto (s^{-1})

K_d – constante de dissociação do complexo enzima-substrato

k_i – constante de especificidade para o substrato genérico de grau de polimerização “i” (L mol^{-1} s^{-1})

K_{ID} – constante de inibição da enzima pelo produto D

$K_{M(i)}$ – constante de Michaelis-Menten para o substrato genérico de grau de polimerização “i” (mol L^{-1})

$k_{off(i)}$ – constante fundamental de velocidade para dissociação do complexo ES_i em S_i+E

$k_{on(i)}$ – constante fundamental de velocidade para formação do complexo ES_i de S_i+E

k_r – constante de especificidade para a reação escolhida como referência para o cálculo da constante de especificidade relativa (L mol^{-1} s^{-1})

m – grau de polimeriazação do maior substrato presente no sistema (do substrato usado no início da reação)

n – grau de polimerização do oligômero para o qual a equação diferencial está sendo escrita

N_k – número de moléculas do substrato de grau de polimerização “k”

o – valor inicial

P_i – probabilidade de ataque processivo proveniente do complexo ES_i (tal que a probabilidade do complexo ES_i liberar S_i em solução seja 1-P_i)

R_i – relação entre a constante de especificidade k_i e a constante de especificidade tomada como referência

S_i – substrato genérico de grau de polimerização “i”

S_i – concentração relativa do substrato genérico de grau de polimerização “i” (*S* em fonte itálica)

T – enzima total no sistema (livre e complexada)

t – tempo (s)

u – grau de polimerização do produto secundário (que provém da extremidade não redutora do substrato; produtos secundários idênticos são removidos do substrato inicial e os produtos residuais em ciclos sucessivos de hidrólise)

v_{oi} – velocidade inicial da reação do substrato “i” gerando produto

Y – substrato genérico

SUMÁRIO

1 INTRODUÇÃO

O mundo caminha para uma economia baseada na minimização de desperdícios e na substituição de derivados de petróleo por bioprodutos (DE JONG; JUNGMEIER, 2015). Esta mudança de paradigma exige a criação de ferramentas que auxiliem o desenvolvimento e a otimização de bioprocessos, para preservação do meio ambiente.

O futuro do desenvolvimento de processos em biorrefinarias está fortemente baseado no uso eficiente e sustentável da biomassa. Dentre os tipos de biomassa disponíveis para biorrefinarias, materiais ricos em polissacarídeos são os mais abundantes, compostos por polímeros como celulose, hemicelulose, pectina e amido. Resíduos agroindustriais contendo estes polissacarídeos são subaproveitados, sendo usados para alimentação animal, queimados, descartados em aterros ou simplesmente abandonados no campo (MOHANRAM et al., 2013). No entanto, são potenciais fontes de pentoses e hexoses, que podem ser transformadas em produtos valiosos em uma biorrefinaria.

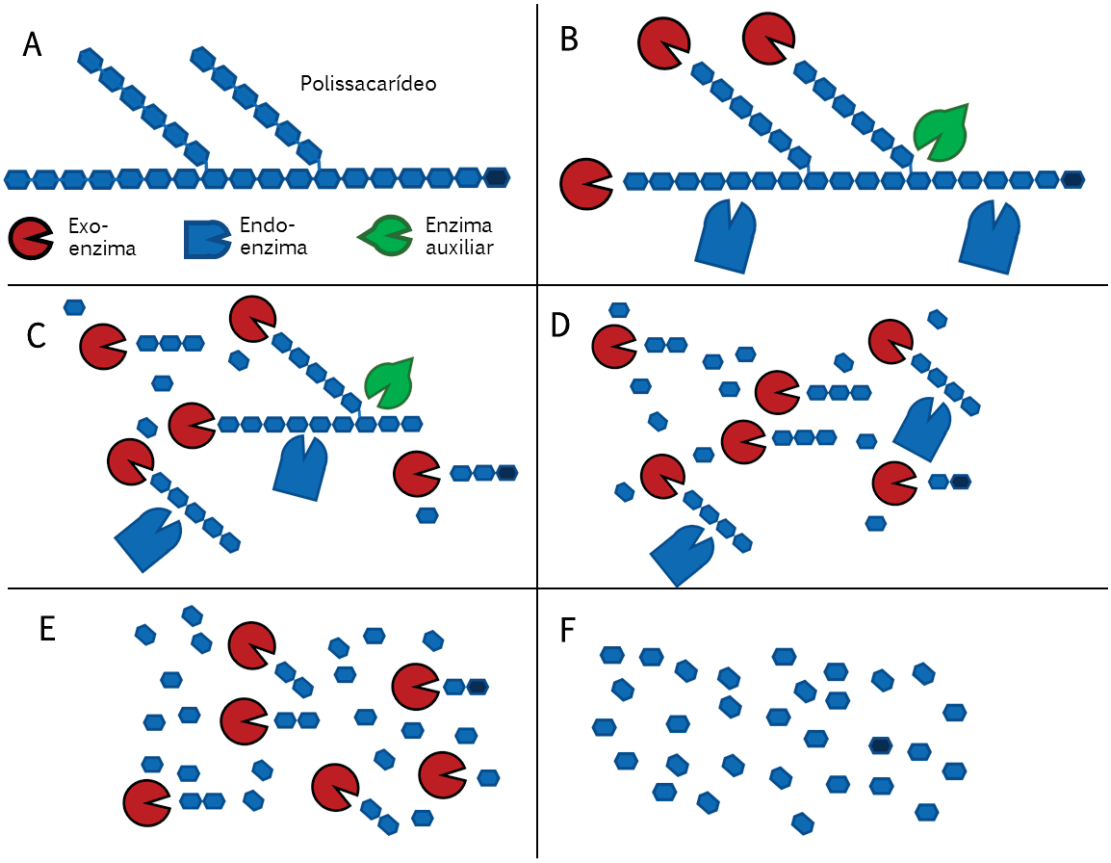
Os diferentes tipos de biomassa possuem composições e estruturas químicas bastante distintos. Por isso, a sacarificação de resíduos agroindustriais pode ser feita por meio de processos químicos ou enzimáticos. A hidrólise química tem a vantagem de ser um processo de baixo custo. Por outro lado, as condições de temperatura e pH usadas na hidrólise química podem promover reações paralelas ou destruir o monossacarídeo gerado, diminuindo o rendimento da reação (BASTOS, 2007).

A hidrólise enzimática é a alternativa para os polissacarídeos que não suportam as condições usadas na hidrólise química. Apesar de serem suscetíveis a bloqueios e inibições causados por produto de reação, são catalisadores muito específicos (BASTOS, 2007).

A sacarificação enzimática de um polissacarídeo envolve a ação de um consórcio de enzimas: as endoenzimas atuam no interior da cadeia, produzindo extremidades para a ação das exoenzimas, enquanto enzimas auxiliares removem eventuais grupos substituintes ou ramificações (FIGURA 1). Assim, durante esta sacarificação, um oligossacarídeo gerado como produto em uma reação serve de substrato para a reação seguinte. Estas moléculas menores, geradas tanto por endoenzimas quanto por exoenzimas, são chamadas de intermediários de reação

(PAOLUCCI-JEANJEAN et al., 2000). A reação acaba quando as cadeias ficam tão pequenas que não podem mais ser hidrolisadas.

FIGURA 1 – REPRESENTAÇÃO DA REAÇÃO DE SACARIFICAÇÃO ENZIMÁTICA DE UM POLISSACARÍDEO



FONTE: O autor (2017).
LEGENDA: Polissacarídeo e diferentes tipos de enzimas que atuam na sua sacarificação (A). As enzimas se diferenciam pelo ponto de ataque ao polissacarídeo: exoenzimas atacam pelas extremidades não redutoras, endoenzimas atacam ligações em posições intermediárias da cadeia e enzimas auxiliares removem ramificações da cadeia principal do polissacarídeo (B). A hidrólise do polissacarídeo gera oligômeros e monossacarídeos (C), que por sua vez servem de substrato para novas reações, reduzindo cada vez mais seu grau de polimerização (D). As exoenzimas do exemplo possuem baixa especificidade por dissacarídeos, que tendem a se acumular no meio junto com os monossacarídeos (E). Quando não há mais outro substrato disponível, os dímeros acabam sendo usados como substrato até que a sacarificação se complete, quando restam somente monossacarídeos no meio (F). O hexágono escuro representa a extremidade redutora do polissacarídeo inicial.

No decorrer da hidrólise completa de um polissacarídeo, ocorre o acúmulo de intermediários de reação de diferentes tamanhos. Do ponto de vista de determinada exoenzima, todos estes intermediários lhe servem de substrato.

Em um sistema onde uma enzima encontra diferentes possibilidades de reação no meio, a velocidade de consumo de determinado substrato não pode ser avaliada sem que se considere a presença de todos os possíveis substratos (CORNISH-BOWDEN, 2012). Assim, uma análise correta deste sistema deve

considerar a proporção entre uma determinada velocidade e a velocidade total de reação (v_{total}), que é dada pela soma das velocidades de todas as reações:

$$\frac{v_1}{v_{total}} = \frac{v_1}{v_1 + v_2 + v_3 + \cdots} = \frac{v_1}{\sum_1^n v_i} \quad (1)$$

A velocidade relativa de consumo de determinado substrato pode ser interpretada como a probabilidade de que a enzima catalise certa reação com determinado substrato em um meio reacional com vários tipos de substratos disponíveis. Esta probabilidade depende da preferência inata da enzima para catalisar a reação deste substrato (dado pela constante de especificidade) e da concentração relativa deste mesmo substrato em relação aos demais.

No início da reação de sacarificação, os substratos são majoriatariamente poliméricos e são ordens de grandeza maiores que a enzima. Assim, enquanto o avanço da reação ainda é pequeno, a enzima não apresenta diferença significativa entre as especificidades por estes substratos. Desta forma, as velocidades relativas de consumo de cada substrato seriam função apenas das concentrações destas moléculas no meio, pois as preferências relativas seriam iguais.

Com o avanço da fração de reação, os substratos são rapidamente reduzidos de polissacarídeos para oligossacarídeos. Substratos oligoméricos são moléculas menores e podem interagir de forma diferente com o sítio ativo, dependendo do tamanho. A enzima passa a apresentar preferência por determinados tamanhos de oligossacarídeos formados na reação (PEREIRA; KRIEGER; MITCHELL, 2016), mesmo quando intermediários diferem entre si por apenas algumas unidades.

Várias enzimas que agem em polissacarídeos demonstram o fenômeno da processividade, onde a enzima realiza reações sequenciais sem se desligar da cadeia principal do substrato (HORN et al., 2012). As enzimas apresentam diferentes probabilidades de ataque processivo, de acordo com o grau de polimerização dos intermediários ligados ao sítio ativo (COUMANS et al., 2006), e esta diferença é especialmente importante entre oligossacarídeos.

Para otimizar processos de sacarificação, modelos matemáticos são ferramentas úteis e, para estes modelos, é necessário saber os parâmetros que caracterizam as enzimas envolvidas. No caso de enzimas capazes de ação processiva, estes parâmetros incluem não somente as constantes de especificidade

para as reações com differentes substratos, mas também as probabilidades de ação processiva.

O intuito desta dissertação foi o de desenvolver um método para a determinação destes parâmetros a partir de perfis experimentais de hidrólise. Foram feitos estudos de caso de crescente complexidade, utilizando dados da literatura de β-amilases agindo sobre maltooligossacarídeos.

2 REVISÃO DA LITERATURA

2.1 IMPORTÂNCIA DA CONSTANTE DE ESPECIFICIDADE NA HIDRÓLISE DE OLIGOSSACARÍDEOS

Segundo os estudos clássicos de cinética enzimática, desenvolvidos por Victor Henri, Leonor Michaelis e Maud Leonora Menten, as reações enzimáticas podem ser divididas em dois passos: uma primeira etapa onde o substrato liga-se de forma reversível à enzima, formando um complexo enzima-substrato, e uma segunda etapa onde ocorre a transformação química catalisada pela enzima, com liberação do produto (NELSON; COX, 2014, p. 203). Para cara reação parcial, há uma constante de velocidade fundamental (FIGURA 2).

FIGURA 2 – ESQUEMA DE REAÇÃO ENZIMÁTICA

FONTE: O autor (2017).
LEGENDA: k_{on} é a constante de velocidade da reação de associação da enzima com o substrato; k_{off} é a constante de velocidade da reação de dissociação do complexo enzima-substrato; k_{cat} é a constante de velocidade da reação de conversão do substrato ligado à enzima e liberação do produto.

O esquema mostrado na FIGURA 2 pode ser aplicado a muitas reações com mecanismos de reação mais complexos. Independentemente de k_{cat} não representar apenas uma etapa do mecanismo, ela tem propriedades de uma constante de velocidade de primeira ordem. Esta constante define a capacidade do complexo enzima-substrato, uma vez formado, gerar produto (CORNISH-BOWDEN, 2012).

Nos estudos clássicos de cinética enzimática, trabalha-se com a constante de Michaelis-Menten (K_{M}), que é uma combinação de constantes de velocidade fundamentais da reação. Para o esquema representado na FIGURA 2, esta constante seria:

$$K_{\mathrm{M}} = \frac{k_{cat} + k_{off}}{k_{on}} \qquad (2)$$

A eficiência da catálise pode ser analisada pela constante de especificidade (k), dada pela razão k_{cat}/K_M. k pode ser expresso em termos das constantes de velocidade fundamentais, para o esquema mostrado na FIGURA 2, como:

$$k = \frac{k_{cat}}{K_M} = k_{cat}\left(\frac{k_{on}}{k_{cat} + k_{off}}\right) = k_{on}\left(\frac{k_{cat}}{k_{cat} + k_{off}}\right) \quad (3)$$

O rearranjo das constantes fundamentais resulta na constante de segunda ordem para a ligação do substrato à enzima (k_{on}) multiplicando o termo $\frac{k_{cat}}{k_{cat}+k_{off}}$, que é a probabilidade de que o complexo enzima-substrato proceda com a catálise ao invés de simplesmente se dissociar de novo (JOHNSON, 2008). Esta forma de expressar k é mais informativa, pois mostra que a eficiência da conversão de substrato a produto depende das eficiências de duas etapas: a formação do complexo enzima-substrato e a conversão do substrato ligado à enzima em produto.

A constante de especificidade é especialmente interessante para caracterizar as preferências quando uma mesma enzima age em diferentes substratos (PEREIRA; KRIEGER; MITCHELL, 2016). Este tipo de situação ocorre, por exemplo, com as enzimas que atuam na sacarificação de polissacarídeos. Estas enzimas podem executar mais de uma reação com o mesmo substrato ou realizar reações sequenciais, onde o produto da reação anterior serve de substrato para a reação seguinte. Como resultado, é gerado no meio uma mistura de oligômeros de tamanho diferentes, cujas concentrações relativas variam ao longo do tempo (PEREIRA; KRIEGER; MITCHELL, 2016).

A especificidade de uma enzima para um determinado substrato reflete o quanto ela consegue discriminar aquele substrato de outros, quando eles estão presentes na reação (CORNISH-BOWDEN, 2012). Neste contexto, quando vários substratos competem pela mesma enzima, um acaba atuando como inibidor competitivo do outro. Em um sistema onde uma enzima (E) catalisa a reação sequencial de conversão do substrato A aos intermediários B e C e ao produto final D ($A \rightarrow B \rightarrow C \rightarrow D$), as equações que representam as velocidades das reações são:

$$v_A = \frac{k_A[\mathrm{E}]_\mathrm{T}[\mathrm{A}]}{1 + \frac{[\mathrm{A}]}{K_\mathrm{MA}} + \frac{[\mathrm{B}]}{K_\mathrm{MB}} + \frac{[\mathrm{C}]}{K_\mathrm{MC}} + \frac{[\mathrm{D}]}{K_\mathrm{ID}}} \quad (4)$$

$$v_B = \frac{k_B[\mathrm{E}]_\mathrm{T}[\mathrm{B}]}{1 + \frac{[\mathrm{A}]}{K_\mathrm{MA}} + \frac{[\mathrm{B}]}{K_\mathrm{MB}} + \frac{[\mathrm{C}]}{K_\mathrm{MC}} + \frac{[\mathrm{D}]}{K_\mathrm{ID}}} \quad (5)$$

$$v_C = \frac{k_C[\mathrm{E}]_\mathrm{T}[\mathrm{C}]}{1 + \frac{[\mathrm{A}]}{K_\mathrm{MA}} + \frac{[\mathrm{B}]}{K_\mathrm{MB}} + \frac{[\mathrm{C}]}{K_\mathrm{MC}} + \frac{[\mathrm{D}]}{K_\mathrm{ID}}} \quad (6)$$

onde k_A, k_B e k_C são as constantes de especificidade e K_{MA}, K_{MB} e K_{MC} são as constantes de Michalis-Menten para os substratos A, B e C; K_{ID} é a constante de inibição da enzima pelo produto D (sendo a constante de dissociação do complexo ED). Como as três reações estão acontecendo no mesmo meio reacional, as equações de velocidade para cada reação têm o mesmo denominador.

Neste sistema, a velocidade total de reação (v_{total}) é dada pela soma das velocidades para cada reação:

$$v_{total} = v_A + v_B + v_C \quad (7)$$

Para deteminar como uma enzima discrimina um substrato dos outros, é necessario considerar a razão da velocidade de consumo do substrato de interesse à velocidade total de reação. Analisando o substrato A, a razão de velocidades será:

$$\frac{v_A}{v_{total}} = \frac{v_A}{v_A + v_B + v_C} \quad (8)$$

Substituindo as equações (4), (5) e (6) em (8) e eliminando os termos em comum:

$$\frac{v_A}{v_{total}} = \frac{k_A[\mathrm{A}]}{k_A[\mathrm{A}] + k_B[\mathrm{B}] + k_C[\mathrm{C}]} \quad (9)$$

A equação (9) mostra que os valores relativos das constantes de especificidade vão determinar as velocidades relativas com que as reações ocorrem, quando os substratos estão presentes em concentrações equimolares no meio de reação (CORNISH-BOWDEN, 1993).

Para substratos poliméricos, a especificidade da enzima pouco varia com o grau de polimerização da cadeia (MERTENS, 2013). Porém, esta especificidade é bastante sensível ao tamanho da cadeia quando se compara substratos oligoméricos, e este é o cenário que caracteriza grande parte do processo de sacarificação (PEREIRA; KRIEGER; MITCHELL, 2016).

2.2 PROCESSIVIDADE

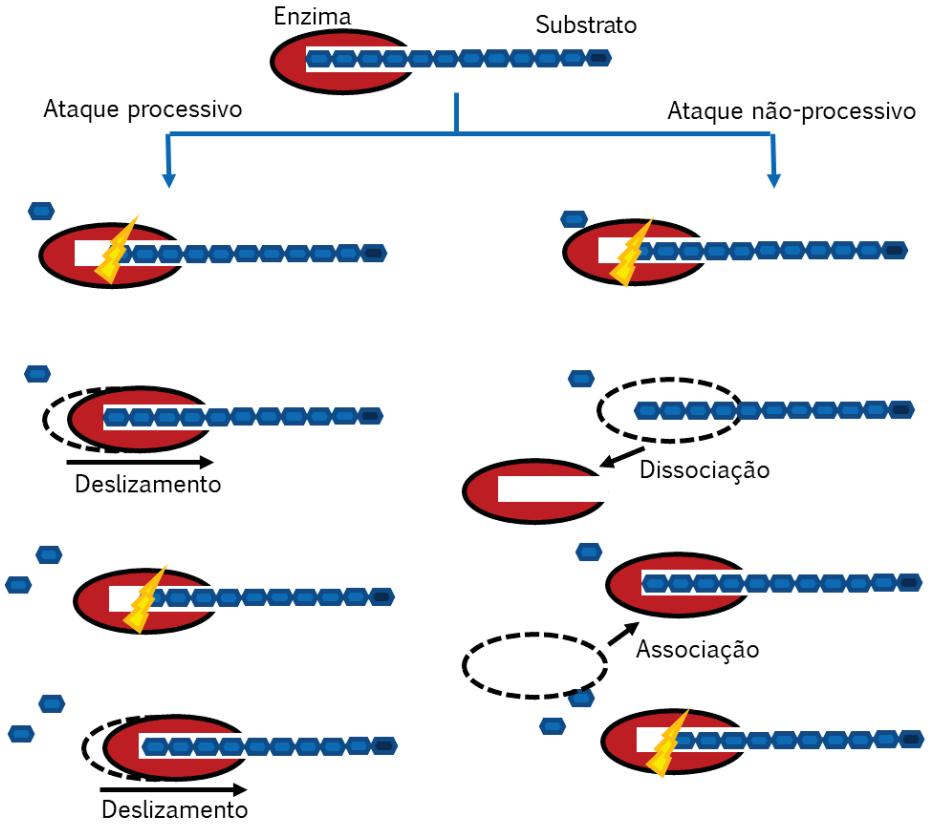
Muitas enzimas que atuam na hidrólise de polissacarídeos apresentam processividade, que ocorre quando uma enzima continua associada com seu substrato polimérico depois de catalisar uma reação hidrolítica, realizando uma série de hidrólises de ligações glicosídicas antes de se dissociar do polímero (HAMRE et al., 2014), conforme ilustrado na FIGURA 3. Um produto que ainda pode ser hidrolisado é chamado de intermediário e, no contexto da processividade, inclui tanto a molécula que está livre na solução quanto a que está ligada à enzima. Este fenômeno, detectado primeiramente em amilases, também ocorre em muitas β-glucanases, celulases, quitinases e poligalacturonases (MEDVE et al., 1998; MICHEL et al., 2003; ROBYT; FRENCH, 1967; SØRBOTTEN et al., 2005).

Diversos trabalhos mostram que resíduos aromáticos presentes no sítio ativo têm papel fundamental na processividade em hidrolases que atuam em polissacarídeos (HORN et al., 2006; KNOTT et al., 2014; TAYLOR et al., 2013). Subsítios vazios e, possivelmente, movimento de alças na proteína geram a força motriz do deslocamento da cadeia de substrato no sítio ativo, sem a liberação do polímero ao meio de reação (DAVIES; HENRISSAT, 1995).

Quando se analisa uma molécula de substrato na reação de sacarificação enzimática, a processividade pode ser caracterizada pela probabilidade de ataque processivo (P), que é a probabilidade de a enzima prosseguir, após uma etapa catalítica, para uma próxima rodada de hidrólise sem se dissociar da cadeia principal do substrato (BAILEY; FRENCH, 1957; LUCIUS et al., 2003; MCCLURE; CHOW, 1980). Esta probabilidade é dada pela seguinte expressão:

$$P = \frac{k_{cat}}{k_{cat} + k_{off}} \quad (10)$$

FIGURA 3 – COMPARAÇÃO ENTRE ATAQUES PROCESSIVO E NÃO-PROCESSIVO À UMA ENZIMA



FONTE: O autor (2017).

Em uma escala maior, as concentrações de substratos tipicamente se encontram na ordem de grandeza de micromolar ou mais, o que representa pelo menos 10^{17} moléculas por litro de meio reacional. Como cada enzima neste sistema obedece à uma probabilidade de seguir pela rota processiva (P) ou pela não-processiva ($1 - P$), ocorre que, em um meio reacional com um grande número de moléculas, as probalidades se traduzem em frações determinísticas, ou seja, uma fração determinada de moléculas segue uma rota enquanto a fração restante segue a outra rota. O grande número de repetições da mesma reação garante a robustez na delimitação do sistema necessária para que o ele seja considerado determinístico (STREVENS, 2011). Cruys-Bagger et al. (2013), trabalhando com probabilidades em sistema determinístico, seguiu esta linha de raciocínio ao fazer uma ressalva que seu modelo, apesar de explicar reações com substratos bastante diluídos, são mais

seguramente aplicados quando há um excesso de substrato em relação à concentração de enzima.

A probabilidade de ataque processivo depende do grau de polimerização do substrato, principalmente quando este é composto de oligossacarídeos (COUMANS et al., 2006). É importante, então, determinar as probabilidades de ataque processivo que uma enzima apresenta pelos oligossacarídeos, pois estas probabilidades afetam o perfil de intermediários de reação durante a hidrólise.

Os métodos convencionais para obtenção das constantes de especificidade e das probabilidades de ataque processivo partem da determinação de constantes de velocidade fundamentais (k_{on}, k_{off} e k_{cat}). Porém, esta determinação de constantes de velocidade é bastante trabalhosa e sujeita a interferências (HORN et al., 2012).

Para casos específicos, é possível assumir simplificações que permitem, a partir de experimentos de curta duração, extrair dados suficientes para estimar as constantes cinéticas. Modelos simples foram desenvolvidos por Song et al. (2015), Praestgaard et al. (2011) e Cruys-Bagger et al. (2013) para descrever a hidrólise processiva de celulose. Estes modelos são restritos apenas ao início da reação, ao assumir que o substrato, no sistema, é majoritariamente polimérico. Com esta restrição, é razoável assumir que não há inibição pelo produto, um fenômeno comum nestes sistemas. Neste caso, também é aceitável pressupor que k_{cat} e k_{off} são independentes do comprimento do polissacarídeo. Mas, este pressuposto não é válido para toda a reação, porque com o decorrer da hidrólise o substrato é reduzido à oligômeros, cujos valores de k_{cat} e k_{off} variam de acordo com o tamanho dos oligossacarídeos (PEREIRA; KRIEGER; MITCHELL, 2016).

As constantes cinéticas podem ser usadas na construção de modelos matemáticos da hidrólise enzimática. Estes modelos podem utilizar métodos estocásticos ou determinísticos. Métodos estocásticos são baseados em probabilidades, onde a hidrólise é transformada em eventos discretos e a cada iteração ocorre uma tomada de decisão de qual evento discreto vai ocorrer no modelo (KUMAR; MURTHY, 2013). A abordagem estocástica permite inserir parâmetros, por exemplo, para representar o modo de ação de cada enzima e descrever a estrutura dos substratos. Sua principal desvantagem é o tempo de otimização.

Nakatani (1997) e Ishikawa et al. (2007) usaram o método de Monte Carlo, um tipo de método estocástico, para representar a hidrólise de maltooligossacarídeos

por uma β-amilase, uma exoenzima processiva que remove unidades de maltose da extremidade não-redutora de seu substrato. Em ambos os casos, o modelo de ataque processivo da β-amilase considera duas probabilidades, que descrevem eventos diferentes. A primeira probabilidade é de a enzima escolher determinado substrato do meio de reação, entre todos os possíveis substratos, e efetuar a hidrólise (P_{k}). P_{k} é definido como:

$$P_{\mathrm{k}} = \frac{N_k / K_{d(k)}}{\sum_{i=4}^{m} N_i / K_{d(i)}} \tag{11}$$

onde N_k é o número de moléculas do substrato de grau de polimerização “k”, “m” é o número de unidades do maior substrato presente no sistema e, para dado substrato de grau de polimerização “i”, $K_{d(i)}$ é a constante de dissociação do complexo enzima-substrato. Depois da hidrólise, o produto residual ainda permanece no sítio ativo; este produto residual contém a extremidade redutora da molécula original. A segunda probabilidade é P, que é a probabilidade de a enzima executar um ataque processivo no produto residual, que faz com que a probabilidade do complexo enzima-substrato se dissociar seja igual a $1 - P$.

Nakatani (1997) simulou, separadamente, reações de hidrólise total de maltoundecaose (G11), maltononaose (G9), maltooctaose (G8) e maltoheptaose (G7) por β-amilase. Nestes sistemas, o autor assumiu que as constantes de dissociação eram iguais para todos os oligossacarídeos, com exceção para maltotetraose (G4). Assim, a probabilidade de um substrato ser escolhido dependeria apenas de sua fração molar no meio. Para maltotetraose, Nakatani (1997) precisou adicionar um parâmetro extra para dimunir a probabilidade de G4 ser escolhido pela enzima e permitir o ajuste da curva predita aos dados experimentais. Quanto às probabilidades de ataque processivo, o autor considerou que são iguais, independente do tamanho do substrato. Estes pressupostos não são apropriados e geram probabilidades que não são confiáveis.

Ishikawa et al. (2007) simularam, separadamente, a hidrólise de maltoheptaose por três β-amilases (a enzima nativa, um mutante com processividade diminuída e outro mutante sem processividade). Os autores basearam seu modelo em Nakatani (1997), repetindo o mesmo pressuposto não apropriado de que a probabilidade de um substrato ser escolhido depende apenas de sua fração molar no

meio, independente das constantes de especificidade dos oligossacarídeos. Os modelos de Ishikawa et al. (2007) e Nakatani (1997) não se baseiam nas equações de cada etapa da reação, o que compromete sua capacidade preditiva.

Os modelos determinísticos, baseados em equações diferenciais, necessitam de métodos numéricos para integrar estas equações e para predizer as concentrações das espécies da reação. Para cada substrato, intermediário e produto, pode se desenvolver uma equação diferencial ordinária (EDO) baseada no mecanismo da reação. A limitação desta estratégia é que o número de EDOs pode ser muito grande e tedioso de se desenvolver quando a reação envolve muitos substratos (GRIGGS; STICKEL; LISCHESKE, 2012).

2.3 FINGERPRINTING

O método Fingerprinting foi desenvolvido para analisar perfis de reação de enzimas que catalisam diversas reações sequenciais com o mesmo substrato, permitindo que sejam estimadas as constantes de especificidades relativas da enzima para os vários intermediários gerados ao longo da reação (MITCHELL et al., 2008a, 2008b; MITCHELL; CARRIÈRE; KRIEGER, 2008). Este método permite que estas constantes sejam determinadas a partir de um único e bem caracterizado perfil de reação, que inclua todas as espécies envolvidas na reação (substrato inicial, intermediários, produtos secundários e produto final) e que cubra a reação do início ao fim. A vantagem do método Fingerprinting sobre outras abordagens é que estas requerem um grande número de ensaios de velocidade inicial.

As equações diferenciais utilizadas no método Fingerprinting são expressas em termos de grau de reação como a variável independente, que é definido como o número de ligações quebradas em relação ao total de ligações quebráveis inicialmente presentes no meio de reação. Mudar a variável independente de tempo para grau de reação tem a grande vantagem de remover da análise as interferências causadas por inibição ou inativação das enzimas, facilitando a obtenção das constantes de especificidade relativas (MITCHELL; CARRIÈRE; KRIEGER, 2008).

O conjunto de constantes de especificidade relativas que resulta da análise podem ser usados para selecionar enzimas para aplicações específicas, monitorar o efeito de mutações introduzidas em determinada enzima e fornecer dados que

permitam correlacionar estrutura com especificidade da enzima (MITCHELL; CARRIÈRE; KRIEGER, 2008).

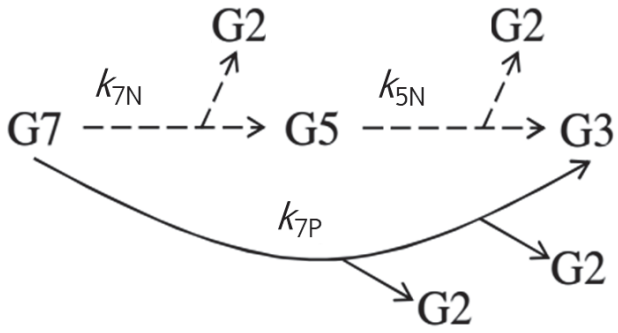
Apesar do método fornecer valores relativos para as constantes de especificidade, os resultados podem ser complementados com um experimento cinético simples, conduzido em paralelo, que forneça o valor absoluto da constante de especificidade para um dos substratos, o que já seria suficiente para o cálculo dos valores absolutos de todas as constantes de especificidade.

O método Fingerprinting já foi utilizado para obtenção de valores das constantes relativas de especificidade para diversas reações de hidrólise, como a de triacilglicerol, fitato, maltoheptaose, β-1-6-N-acetilglucosamina e oligômeros de ácido galacturônico (MITCHELL et al., 2008a; MITCHELL; CARRIÈRE; KRIEGER, 2008; PEREIRA; KRIEGER; MITCHELL, 2016).

Por outro lado, até o presente momento, o método Fingerprinting não foi satisfatoriamente aplicado em reações nas quais ocorre processividade. Mitchell, Carrière e Krieger (2008), analisando os dados experimentais de Ishikawa et al. (2007) não incluíram o fenômeno da processividade em seu modelo, que acabou não se ajustando aos dados experimentais para a β-amilase processiva. Pereira, Krieger e Mitchell (2016), trabalhando com estes mesmos dados, aplicaram o método Fingerprinting considerando processividade. Porém, os pressupostos que os autores adotaram para modelar o fenômeno induzem a interpretações equivocadas, pois o modelo assume que a enzima possui duas constantes de especificidade para hidrolisar determinado substrato (FIGURA 4), uma para a reação processiva e outra para a não processiva (k_{7P} e k_{7N}, respectivamente). Esse pressuposto assume que há uma escolha entre a rota processiva e a dissociação do produto residual no momento de ligação ao substrato, o que só seria possível caso o substrato pudesse se ligar de diferentes formas ao sítio ativo.

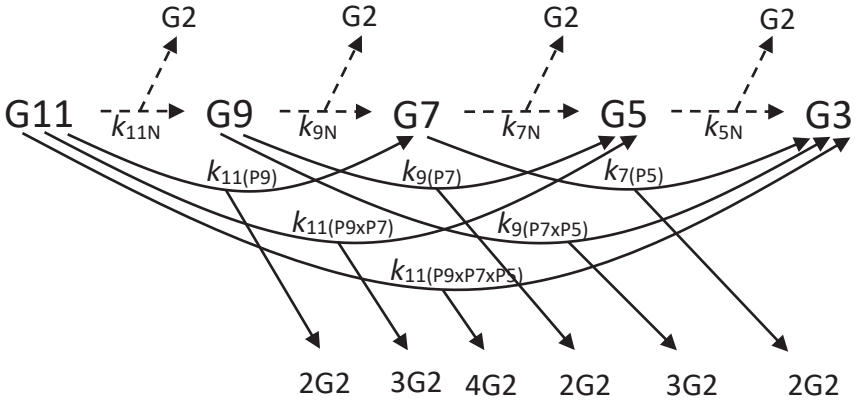
Embora esta abordagem seja conceitualmente inusitada, ela é matematicamente válida e pode, de fato, ser usada satisfatoriamente em reações com poucas etapas catalíticas. Mas, à medida que o número de intermediários aumenta, a reação fica mais complicada de ser representada matematicamente (FIGURA 5). Outro problema são as constantes de especificidade parciais para cada rota (não-processiva e processivas), que não seriam as probabilidades de ação processiva e constantes de especificadde de reações individuais e sim “produtos destas”.

FIGURA 4 – HIDRÓLISE ENZIMÁTICA DE MALTOHEPTAOSE ANALISADA PELA ABORDAGEM DE PEREIRA, KRIEGER E MITCHELL (2016).



FONTE: Pereira, Krieger e Mitchell (2016).
LEGENDA: As linhas tracejadas representam a hidrólise pela rota não-processiva, enquanto a linha contínua representa a rota processiva. G7 é maltoheptaose, G5 é maltopentaose, G3 é maltotriose e G2 é maltose. k_{7N}, k_{5N} e k_{7P} representam as contantes de especificidade da enzima para cada uma das reações.

FIGURA 5 – ESQUEMA DA HIDRÓLISE ENZIMÁTICA DE MALTOUNDECAOSE PROPOSTO SEGUNDO A ABORDAGEM DE PEREIRA, KRIEGER E MITCHELL (2016)



FONTE: O autor (2017).
LEGENDA: As linhas tracejadas representam a hidrólise pela rota não-processiva, enquanto a linha contínua representa a rota processiva. G11 é maltoundecaose, G9 é maltononaose, G7 é maltoheptaose, G5 é maltopentaose, G3 é maltotriose e G2 é maltose. k_{11N}, k_{9N}, k_{7N}, k_{5N}, $k_{11(P9)}$, $k_{9(P7)}$, $k_{7(P5)}$, $k_{11(P9xP7)}$, $k_{9(P7xP5)}$ e $k_{11(P9xP7xP5)}$ representam as contantes de especificidade da enzima para cada uma das reações.

3 JUSTIFICATIVA

Este trabalho tem como justificativa a necessidade de um modelo da hidrólise de oligossacarídeos lineares, para situações onde a enzima possa atuar de forma processiva, que seja mecanisticamente melhor do que os encontrados na literatura. Além disso, ainda não há um método que possa extrair as contantes de especificidade relativas e probabilidades de ataque processivo de perfis completos de reação de hidrólise. Estas constantes são fundamentais para simulações computacionais da sacarificação de polissacarídeos, além de serem úteis para monitorar os efeitos de mutações em estudos de melhoramento de enzimas.

4 OBJETIVOS

Este trabalho tem como objetivo geral estender o método Fingerprinting, para que possa ser aplicado a reações catalisadas por enzimas processivas. Desta forma, pretende-se obter uma forma de determinar, além das constantes de especificidades relativas, as probabilidades de ataque processivo para cada intermediário gerado na hidrólise de um polissacarídeo por uma exoenzima, com base em perfis completos de reação.

Os objetivos específicos do trabalho são:

a) Desenvolver um modelo geral da processividade de exoenzimas para incorporar ao método Fingerprinting.
b) Demonstrar o modelo no contexto de reações catalisadas por β-amilases, utilizando, para isso, dados experimentais da literatura.

5 ARTIGO 1

Manuscrito do artigo, a ser submetido à revista *Biochemical Engineering Journal*, Fator de Impacto 2016: 2,892 (JCR). *Qualis* Ciências Biológicas II: B1

Fingerprinting processive β-amylases

Gerson Ritzmann de Mello[a]
Aline Bescrovaine Pereira[a]
Nadia Krieger[b]
David Alexander Mitchell[a]

[a] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná, Cx. P. 19046 Centro Politécnico, Curitiba 81531-980, Paraná, Brazil

[b] Departamento de Química, Universidade Federal do Paraná, Cx. P. 19081 Centro Politécnico, Curitiba 81531-980, Paraná, Brazil

Corresponding author

Dr. David Alexander Mitchell.
Departamento de Bioquímica e Biologia Molecular. Universidade Federal do Paraná. Cx. P. 19046 Centro Politécnico, Curitiba 81531-980, Paraná, Brazil
E-mail: davidmitchell@ufpr.br
Tel: +55 41 33611536

Abstract

Many polysaccharide hydrolases are capable of processive action, where they perform repeated attacks without dissociating from the substrate molecule. However, a mathematical model that characterize processivity, for different substrate sizes, in the whole extention of the hydrolysis, remains to be estabilished. Previous models contain simplifications that make them valid only to a specific fraction of a given reaction or that make it difficult to apply them to more complex systems. In this paper, we extend the “Fingerprinting method”, a method that can be used to determine the relative specificities that an enzyme has for the various reactions that it can catalyze in a multi-step reaction scheme, to reaction schemes involving processive action. We do this for three case studies for the hydrolysis of linear malto-oligosaccharides by β-amylase, an exoenzyme that successively removes maltose units from the non-reducing ends of the molecules. In these case studies we demonstrate that it is possible to use reaction profiles to determine, simultaneously, the relative specificities that the enzyme has for the different species in the reaction mixture and the probabilities of processive action occurring. These parameters can be used in models for designing and optimizing hydrolysis processes and also to guide protein engineering programs.

Keywords: Processivity, Specificity constant, Fingerprinting, Mathematical model.

5.1 INTRODUCTION

Over the past few decades, there has been increasing interest in using biorefineries to process crops into bioproducts that can substitute petroleum-based fuels and chemicals [1]. One important processing step in such biorefineries will be the enzymatic saccharification of the polysaccharides contained in these crops, using enzyme cocktails. There is currently much interest in optimizing the efficiency of these enzyme cocktails through the selection of new enzymes or the engineering of current enzymes [2].

Many endo- and exoenzymes that are involved in the enzymatic saccharification of polysaccharides can catalyze reactions with polymers of different degrees of polymerization. Endo-acting enzymes can also catalyze various different

reactions with the same substrate. In a reaction medium in which an enzyme can carry out several different reactions, the relative rates at which these reactions proceed depends on two factors: the concentrations of the different substrates and the innate preferences that the enzyme has for catalyzing the various reactions. These innate preferences are determined by the relative values of the specificity constants ($k_{\mathrm{cat}}/K_{\mathrm{M}}$) for the reactions.

Although the specificity constants for the various reactions catalyzed by endo- and exoenzymes in the hydrolysis of polysaccharides are typically not much different [3], the situation is different for oligosaccharides, where specificity constants can vary significantly with the degree of polymerization of the substrate [4-11]. This variation is important in the middle and later stages of saccharification processes, when oligosaccharides predominate. It is desirable, then, to characterize the set of values of the specificity constants of an enzyme for the various reactions that it is capable of catalyzing with oligosaccharides. Knowledge of this set of values is useful in the selection of new enzymes during screening programs and also enables quantitative evaluation of the success of changes in the enzyme introduced by protein engineering methods [12].

Amongst the various possible strategies for obtaining the set of relative specificity constants of an enzyme, the so-called “Fingerprinting method” [12] is highly advantageous: the set of relative specificity constants can be determined based on a single, well-characterized reaction profile and the method is not affected by phenomena such as inhibition and denaturation of the enzyme [4]. In contrast, most other methods are tedious, requiring a large number of initial velocity assays [4].

The Fingerprinting method starts from a set of time-based differential equations for the various reaction intermediates. Changing the independent variable from time to fractional reaction extent results in a set of differential equations in which the only parameters are the relative values of the specificity constants of the enzyme for the various reactions that it can catalyze. These specificity constants are then determined by fitting the equation set to the reaction profiles for the various intermediates, plotted in terms of fractional reaction extent. This method has already been demonstrated for linear and branched reaction schemes [4,12-14]. However, it has not yet been applied adequately to systems in which processive attack occurs.

Processive attack means that, after having catalyzed one reaction, the enzyme does not dissociate from the substrate backbone that remains in the active site, but rather slides along it, catalyzing a second reaction somewhere else in the substrate backbone [15]. In the case of exoenzymes that hydrolyze polysaccharides and oligosaccharides, processive action means that various units (either monosaccharides or disaccharides) are removed successively from the end of the oligomer or polymer before the enzyme finally dissociates. The number of units removed in a round of processive attack can vary, with there being a certain probability that the enzyme will dissociate after each catalytic step [16].

Pereira et al. [4] did apply the Fingerprinting method to the data of Ishikawa et al. [17] for the attack of a soybean β-amylase on maltoheptaose, which involves the possibility of processive action. In their model, they assumed that there are two relative specificity constants for the attack of the β-amylase on maltoheptaose (Fig. 1), one associated with the non-processive route (k_{7N}) and one associated with the processive route (k_{7P}). However, in many cases, there can be multiple processive steps before the enzyme dissociates from the substrate backbone. In this case, application of the approach of Pereira et al. [4] would generate a complex branched scheme.

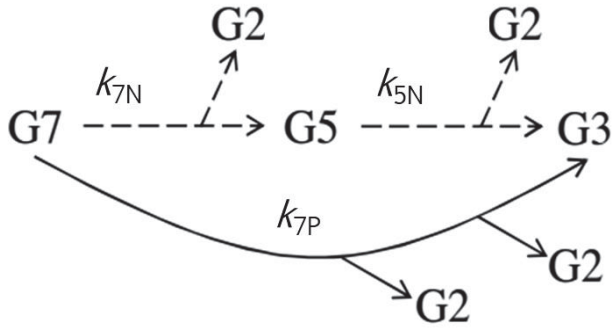


Fig. 1. Branched reaction scheme of Pereira et al. [4] representing the hydrolysis of maltoheptaose following a processive route (continuous arrow) and a non-processive route (dashed arrow).

A better way of treating processivity was suggested by the work of Cruys-Bagger et al. [18], who modeled the processive action of a cellobiohydrolase, an exoenzyme that removes cellobiose units from cellulose chains. In this case, the representation of the system does not involve multiple routes branching from each substrate. Rather, the scheme is linear (Fig. 2). The key point is that, after catalysis,

there is a certain probability that the enzyme will slide along the substrate and catalyze another reaction (i.e. act processively), otherwise it will dissociate.

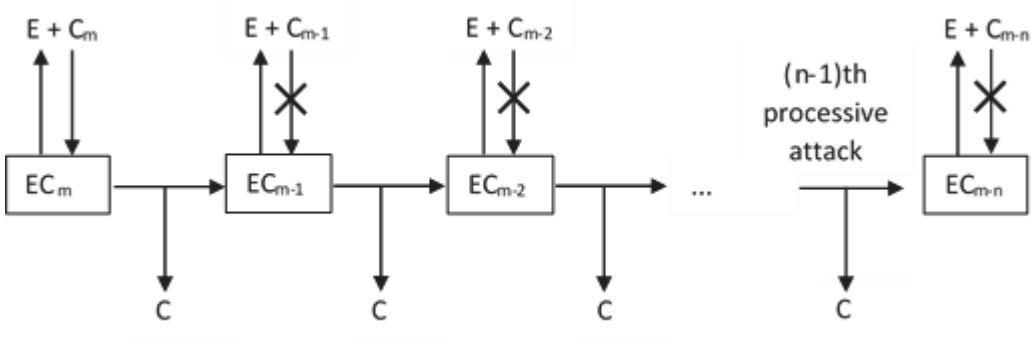


Fig. 2. Scheme proposed by Cruys-Bagger et al. [18] for the processive removal of cellobiose units (C) from a cellulose chain initially containing m cellobiose units. Although a free enzyme should theoretically be able to bind to a partially hydrolyzed substrate molecule, Cruys-Bagger et al. [18] did not represent this possibility in their model, as indicated by the arrows with crosses.

The aim of the present work is to incorporate the approach of Cruys-Bagger et al. [18] to describing processivity of exoenzymes into the Fingerprinting method. We do this in the form of three case studies involving β-amylase, an exoenzyme that removes maltose units from the non-reducing end of maltooligosaccharides. The first case study involves the hydrolysis of maltoheptaose, using the data of Ishikawa et al. [17], with the aim of demonstrating the new method in the same simple system as that analyzed by Pereira et al. [4]. The second case study involves the hydrolysis of maltooctaose, in a system in which two successive processive steps are possible. The third case study involves the fitting of three models simultaneously to data from three experiments that started with different substrates, maltoheptaose, maltononaose and maltoundecoase. The second and third case studies use data from French and Youngquist [19].

5.2 MATHEMATICAL MODELING AND METHODS

5.2.1 Description of the system and the data available

In the system under study (Fig. 3), the enzyme (E), a β-amylase, acts on an initial malto-oligosaccharide substrate, Gm, where “m” is the degree of polymerization. This attack generates two products. One product is a unit of maltose (G2) that is removed from the non-reducing end of the oligosaccharide and released directly into the reaction medium. It is referred to in this work as the “side product”.

The other product is an intermediate G(m-2) that contains the reducing end of the original substrate molecule and which is referred to in this work as the “residual product”. Immediately after the release of the side product, the residual product G(m-2) is still bound to the active site of the enzyme. The reaction can proceed from the complex EG(m-2) through either of two routes. In the processive route, the enzyme slides along the residual product G(m-2), and removes another G2, leaving the residual product G(m-4) in the active site. In the non-processive route, the enzyme simply releases the residual product G(m-2) into the reaction medium.

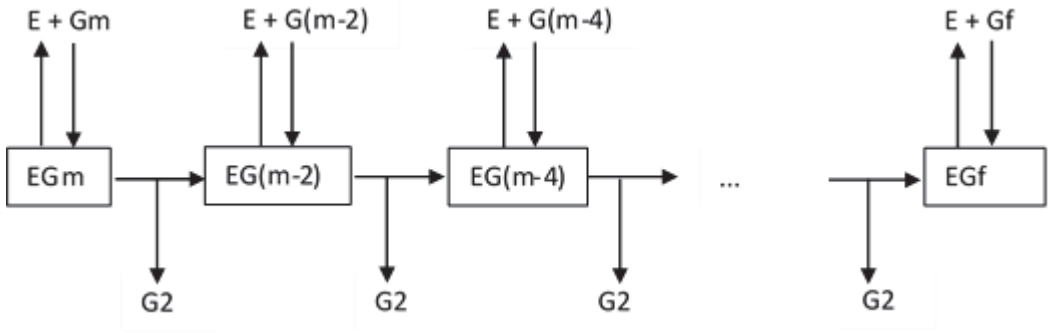


Fig. 3. General scheme for the attack of a processive β-amylase on a malto-oligosaccharide initially containing “m” glucose units.

Successive hydrolysis reactions occur in a similar manner, with all residual products except for the last one being potential substrates for the enzyme. The last residual product is called the final residual product and is a molecule that contains the reducing end of the original substrate molecule and which cannot be further hydrolyzed. It is represented by the symbol Gf, where “f” represents its degree of polymerization. For hydrolysis of a malto-oligosaccharide by a β-amylase, the final residual product will be maltose (G2) if the initial chain contains an even number of glucose units and maltotriose (G3) if the initial chain contains an uneven number of glucose units.

During the hydrolysis process, any complex EGi can follow either the processive route, with probability P_{i}, or the non-processive route, with probability $1 - P_{\mathrm{i}}$. Any free enzyme in the system can bind to any remaining molecules of the initial substrate or to any molecules of residual products that have been released into the medium. We assume that both G3 and G2 can bind reversibly to the β-amylase, forming nonproductive complexes and thereby causing product inhibition.

The three case studies are based on profiles for the initial substrate and the residual products. These profiles are plotted in terms of fractional reactional extent as the independent variable. Data profiles were provided in the original sources for the initial substrate, residual products and final residual products, identified in Table 1. Data profiles were not provided for G2, which is the side product in all case studies and also the final residual product of Case Study 2.

Table 1. Reaction species for which data profiles were given in the original sources used in the case studies

Case Study	Initial substrate	Residual products	Final residual product	Source
1	G7	G5	G3	[17]
2	G8	G6, G4	-	[19]
3 – Experiment 1	G7	G5	G3	[19]
3 – Experiment 2	G9	G7, G5	G3	[19]
3 – Experiment 3	G11	G9, G7, G5	G3	[19]

5.2.2 Overview of the model

The model is a set of differential equations representing the behavior of the mole ratios of all species, during the reaction, as a function of the fractional reaction extent. The nomenclature used in the equations is presented in Table 2.

The general procedure for deduction of all differential equations and transformation of the independent variable from time to fractional reaction extent was described by Mitchell et al. [12]. In the present study, the final equations are different, due to the processivity of the enzyme (more details are given in the Supplementary Material). For processive action of a β-amylase on a malto-oligosaccharide, the accumulation of an intermediate residual product Gn will be given by:

$$\frac{d[\mathrm{Gn}]}{dt} = k_{n+2}[\mathrm{G(n+2)}]\frac{[\mathrm{E}]_T}{D_m}(1 - P_{\mathrm{n}}) + k_{n+4}[\mathrm{G(n+4)}]\frac{[\mathrm{E}]_T}{D_m}P_{\mathrm{n+2}}(1 - P_{\mathrm{n}}) + k_{n+6}[\mathrm{G(n+6)}]\frac{[\mathrm{E}]_T}{D_m}P_{\mathrm{n+4}}P_{\mathrm{n+2}}(1 - P_{\mathrm{n}}) + \cdots - k_n[\mathrm{Gn}]\frac{[\mathrm{E}]_T}{D_m} \quad (1)$$

where "D_m" is a term considering all the possible forms of the enzyme in the system: free, inhibited, or complexed. It should be noted that, in any particular hydrolysis experiment, D_m has a generic form, independently of which substrate Gn is being considered, although it is different for experiments initiated with substrates of

different degrees of polymerization. A full explanation is provided in the Supplementary Material.

Table 2. Nomenclature used in the models for attack of a processive β-amylase on linear malto-oligosaccharides

Symbol	Meaning
Species	
E	Enzyme
Gi	Generic malto-oligosaccharide containing "i" glucose units
Variables	
D_{m}	Denominator of the rate equations for a reaction initiated with Gm
F_{m}	Fractional reaction extent for a reaction initiated with Gm
[Gi]	Concentration of Gi
$G\mathrm{i}$	Mole ratio of Gi
t	Time (s)
Parameters	
$k_{\mathrm{cat(i)}}$	Catalytic constant for Gi (s^{-1})
$K_{\mathrm{M(i)}}$	Michaelis-Menten constant for Gi (mol L^{-1})
k_{i}	Specificity constant for Gi (L mol^{-1} s^{-1})
k_{r}	Specificity constant for the reaction chosen as a reference for calculation of the relative specificity constants (L mol^{-1} s^{-1})
P_{i}	Probability of processive attack proceeding from the complex EGi when this complex is formed by reaction (such that the probability of the complex EGi liberating Gi into solution is $1 - P\mathrm{i}$)
R_{i}	Relative specificity constant for Gi (i.e. ratio of the specificity constant k_{i} to the specificity constant taken as reference, k_{r})
Subscripts	
o	Initial value
T	Total enzyme in the system (free and complexed)
m	Degree of polymerization of the largest substrate present in the system (i.e. of the substrate used to initiate the reaction)
f	Degree of polymerization of the final residual product (the smallest residual product, one that cannot be further attacked; it contains the reducing end of the initial substrate molecule)
n	Degree of polymerization of the oligomer for which the differential equation is being written

For this system, it is possible to arrive at general equations for the initial substrate (Eq. (2)), the residual products that are reaction intermediates (Eq. (3)), the last residual product (Eq. (4)) and for D_{m} (Eq. (5)), when an experiment is initiated with a malto-oligosaccharide containing "m" units of glucose:

$$\frac{d[\mathrm{Gm}]}{dt} = -k_m[\mathrm{Gm}]\frac{[\mathrm{E}]_T}{D_m} \tag{2}$$

$$\frac{d[\mathrm{Gn}]}{dt} = \frac{[\mathrm{E}]_T}{D_m}\left(\sum_{i=n+2}^{m}\left(k_i[\mathrm{Gi}](1 - P_\mathrm{n})\prod_{j=n+2}^{k-2} P_j\right) - k_n[\mathrm{Gn}]\right) \tag{3}$$

$$\frac{d[\mathrm{Gf}]}{dt} = \frac{[\mathrm{E}]_T}{D_m}\sum_{i=f+2}^{m}\left(k_i[\mathrm{Gi}]\prod_{j=f+2}^{i-2} P_j\right) \tag{4}$$

$$D_m = 1 + \sum_{i=f+2}^{m}\frac{[\mathrm{Gi}]}{K_{M(i)}}\left(1 + k_{cat(i)}\sum_{j=f}^{i-2}\frac{\left(\prod_{g=j+2}^{i-2} P_g\right)}{\left(k_{cat(j)} + k_{off(j)}\right)}\right) + \frac{[G2]}{K_{\mathrm{M2}}} + \frac{[Gf]}{K_{\mathrm{Mf}}} \tag{5}$$

where $k_{\mathrm{cat(f)}} = 0$. In these equations, “m” is the degree of polymerization of the initial substrate, “f” is the degree of polymerization of the final residual product (either 2 or 3, depending on whether the initial malto-oligosaccharide is of even or odd chain length) and “n” is the size of the oligomer under consideration (between $\mathrm{m} - 2$ and $\mathrm{f} + 2$, inclusive). In Eq. (5), K_{Mf} and K_{M2} are dissociation constants. In this equation, $\mathrm{f} = 2$ when the initial substrate is of even chain length, such that $K_{\mathrm{Mf}} = K_{\mathrm{M2}}$.

When the independent variable of the differential equations is changed from time to fractional reaction extent (F), the enzyme concentration and “D_m” cancel out of the equations [12]. Additionally, all concentrations are converted to mole ratios (i.e. $G\mathrm{i}$, with the “G” in italic font):

$$G\mathrm{i} = \frac{[\mathrm{Gi}]}{[\mathrm{Gm}]_\mathrm{o}} \tag{6}$$

Also, since specificity constants appear in all numerators and denominators of the final equation set, it is not possible to determine the absolute values of the specificity constants, only their relative values [12]. Therefore, one specificity constant is chosen as a reference for expression of the relative values of the others. The i[th] relative specificity constant is then given by:

$$R_\mathrm{i} = \frac{k_\mathrm{i}}{k_\mathrm{r}} \tag{7}$$

where k_r indicates the specificity constant that is chosen as the reference (by definition, R_r = 1.0).

The final equation set used in the fitting procedure is then given by:

$$D_{\mathrm{m}}^{\#} = \sum_{\mathrm{i=f+2}}^{\mathrm{m}} \left(R_{\mathrm{i}} G\mathrm{i} \left(1 + \sum_{\mathrm{g=f+2}}^{\mathrm{i-2}} \left(\prod_{\mathrm{j=g}}^{\mathrm{i-2}} P_{\mathrm{j}} \right) \right) \right) \tag{8}$$

$$\frac{dG\mathrm{m}}{dF} = -\frac{\left(\frac{\mathrm{m-f}}{2}\right) R_{\mathrm{m}} G\mathrm{m}}{D_{\mathrm{m}}^{\#}} \tag{9}$$

$$\frac{dG\mathrm{n}}{dF} = \frac{\left(\frac{\mathrm{m-f}}{2}\right) \left(\sum_{\mathrm{i=n+2}}^{\mathrm{m}} \left(R_{\mathrm{i}} G\mathrm{i} (1 - P_{\mathrm{n}}) \prod_{\mathrm{j=n+2}}^{\mathrm{i-2}} P_{\mathrm{j}} \right) - R_{\mathrm{n}} G\mathrm{n} \right)}{D_{\mathrm{m}}^{\#}} \tag{10}$$

where there are equations of the form of Eq. (10) for $\mathrm{n} = \mathrm{m} - 2$ to $\mathrm{n} = \mathrm{f} + 2$, inclusive, and

$$\frac{dG\mathrm{f}}{dF} = \frac{\left(\frac{\mathrm{m-f}}{2}\right) \left(\sum_{\mathrm{i=f+2}}^{\mathrm{m}} \left(R_{\mathrm{i}} G\mathrm{i} \prod_{\mathrm{j=f+2}}^{\mathrm{i-2}} P_{\mathrm{j}} \right) \right)}{D_{\mathrm{m}}^{\#}} \tag{11}$$

The specific equation sets used in each case study are shown in the individual case studies. The deductions of these equation sets are shown in the Supplementary Material.

5.2.3 Consistency tests of the data sets

The consistency of the data used in Case Studies 1 and 3 was checked. The consistency test tracked the sum of the molar fractions of the initial substrate and all residual products during the hydrolysis process. For each fractional reaction extent for which data were available, a consistency index was calculated. This index is represented by CI_m, where the subscript indicates the degree of polymerization of the initial substrate used in the particular experiment being analyzed.

For experiments initiated with G7, the consistency index (CI_7) was calculated as:

$$CI_7 = \frac{[\mathrm{G7}] + [\mathrm{G5}] + [\mathrm{G3}]}{[\mathrm{G7}]_{\mathrm{o}}} \tag{12}$$

For experiments initiated with G9, the consistency index (CI_9) was calculated as:

$$CI_9 = \frac{[\mathrm{G9}] + [\mathrm{G7}] + [\mathrm{G5}] + [\mathrm{G3}]}{[\mathrm{G9}]_o} \quad (13)$$

For experiments initiated with G11, the consistency index (CI_{11}) was calculated as:

$$CI_{11} = \frac{[\mathrm{G11}] + [\mathrm{G9}] + [\mathrm{G7}] + [\mathrm{G5}] + [\mathrm{G3}]}{[\mathrm{G11}]_o} \quad (14)$$

For all reaction profiles in Case Studies 1 and 3, the consistency index remained between 0.95 and 1.05 during the whole hydrolysis (i.e. over all fractional reaction extents), meaning that it was not necessary to apply any correction to the data [4].

In Case Study 2, there was insufficient data to undertake the consistency test since the reaction profile for the final residual product (a G2 molecule containing the reducing end of the original oligosaccharide) was not provided.

5.2.4 Fractional reaction extent

All data sets were presented in the original sources with the fractional reaction extent (F) as the independent variable. The fractional reaction extent is defined as the ratio of the number of hydrolyzed bonds to the total number of hydrolysable bonds present in the initial substrate. In Case Studies 1 and 3, in which the original substrates are malto-oligosaccharides with uneven numbers of glucose units, one (and only one) molecule of G2 is generated for every bond that is hydrolyzed. In this case, the fractional reaction extent for a reaction that started with a substrate molecule containing "m" glucose units, denoted F_m, is given by

$$F_7 = \frac{G5 + 2G3}{2} \quad (15)$$

$$F_9 = \frac{G7 + 2G5 + 3G3}{3} \quad (16)$$

$$F_{11} = \frac{G9 + 2G7 + 3G5 + 4G3}{4} \quad (17)$$

where the denominator represents the number of attackable bonds in the original substrate molecule.

In Case Study 2, the reaction profiles did not allow calculation of the *G2* released as a side product, so the fractional reaction extent was calculated on the basis of the number of hydrolysable bonds that were still present in the reaction mixture, as done by Pereira et al. [4]:

$$F_{\mathrm{m}} = 1 - \frac{\sum_{\mathrm{j=f}}^{\mathrm{m}} \left\{ \left(\frac{\mathrm{j} - 2}{2} \right) G\mathrm{j} \right\}}{3} \quad (18)$$

where “m” is the degree of polymerization of the initial substrate ($\mathrm{m} = 8$ in this case), and “f” is the degree of polymerization of the final residual product.

5.2.5 Determination of the model parameters

The only parameters in the final equation sets (each equation set contains specific equations corresponding to Eqs (8) to (11)) are the relative specificity constants and the probabilities of processive attack.

The numerical solution for each model was found with the function *ode45* of MATLAB®. In each of Case Studies 1 and 2, one set of differential equations was integrated. In Case Study 3, three sets of differential equations were integrated, namely one set for each of the three differential initial substrates. Parameter values were determined by non-linear optimization, as described by Pereira et al. [4], using the function *fminsearch* of MATLAB®, which uses a Nelder-Mead simplex algorithm. The objective function was the sum of the squares of the smallest “diagonal errors” between the predicted curve and the experimental points [4].

5.3 RESULTS

5.3.1 Case Study 1: processive attack on maltoheptaose

The first case study is the simplest system where processive attack can be observed; in other words, it is a system where only one processive step is possible. It involves the same system as that modeled by Pereira et al. [4], namely the hydrolysis of a maltoheptaose by soybean β-amylase, using the experimental data of Ishikawa et al. [17].

The system was represented by the reaction scheme given in Fig. 4. Initially, the free enzyme binds to G7, producing the complex EG7. This can dissociate to liberate free enzyme and G7 or it can hydrolyze the G7 in the active site to G5, liberating G2 and producing the complex EG5. This complex may either dissociate to release free enzyme and G5 (i.e. non-processive action, with probability $1 - P_5$) or hydrolyze the G5 in the active site directly to G3 (i.e. processive action, with probability P_5), producing the complex EG3 and liberating another G2. Free enzyme can also bind to G5, producing the complex EG5. This complex can dissociate to liberate free enzyme and G5 again or it can hydrolyze the G5 in the active site to G3, producing the complex EG3. This complex can dissociate, producing free enzyme and G3. It was considered that free enzyme can also bind to G3 and G2, leading to product inhibition.

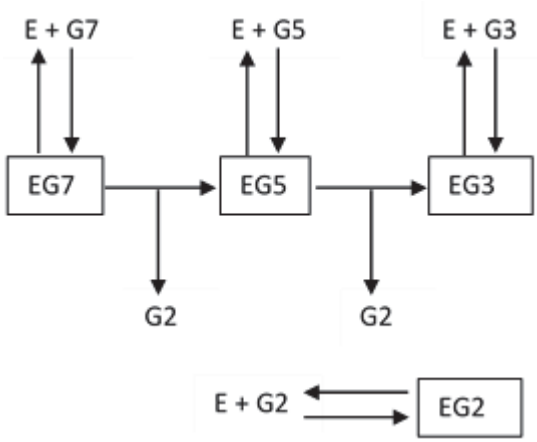


Fig. 4. The hydrolysis of a “maltoheptaose” by a β-amylase, representing the system studied by Ishikawa et al. [17].

The equation set that describes this system is:

$$\frac{dG7}{dF} = 2\left(\frac{-R_7G7}{R_7G7(1+P_5)+R_5G5}\right) \quad (19)$$

$$\frac{dG5}{dF} = 2\left(\frac{R_7G7(1-P_5)-R_5G5}{R_7G7(1+P_5)+R_5G5}\right) \quad (20)$$

$$\frac{dG3}{dF} = 2\left(\frac{R_7G7P_5+R_5G5}{R_7G7(1+P_5)+R_5G5}\right) \quad (21)$$

The specificity constants (*R*-values) are expressed relative to k_7, such that, by definition, $R_7 = 1$.

The detailed deduction of this equation set is given in the Supplementary Material.

The model represented by Eqs. (19), (20) and (21) fits well to the data of Ishikawa et al. [17] for three different enzymes, namely the native soybean β-amylase (Fig. 5a), the mutant enzyme D53A (Fig. 5b), which has reduced processivity, and the mutant enzyme W55R (Fig. 5c), in which processivity is almost completely abolished. The fitted constants of the model are given in Table 3. These constants show that the native free enzyme has almost identical specificities for G7 and G5 (i.e. the relative specificity for G5 is almost equal to 1). The value of P_5 indicates the probability of the complex EG5 following the processive route after having been produced by reaction, namely the probability of the enzyme sliding along the residual product G5 in the reaction site and then cleaving it to produce EG3 and G2. Note that P_5 says nothing about the fate of the complex EG5 when it is produced by the enzyme binding to free G5 in the reaction medium.

Table 3. Values of the parameters obtained by fitting the model represented by Eqs. (19) to (21) to the experimental data from Ishikawa et al. [17].

Enzyme	Native	D53A	W55R
R_5	0.975	0.889	0.668
P_5	0.557	0.151	0.003

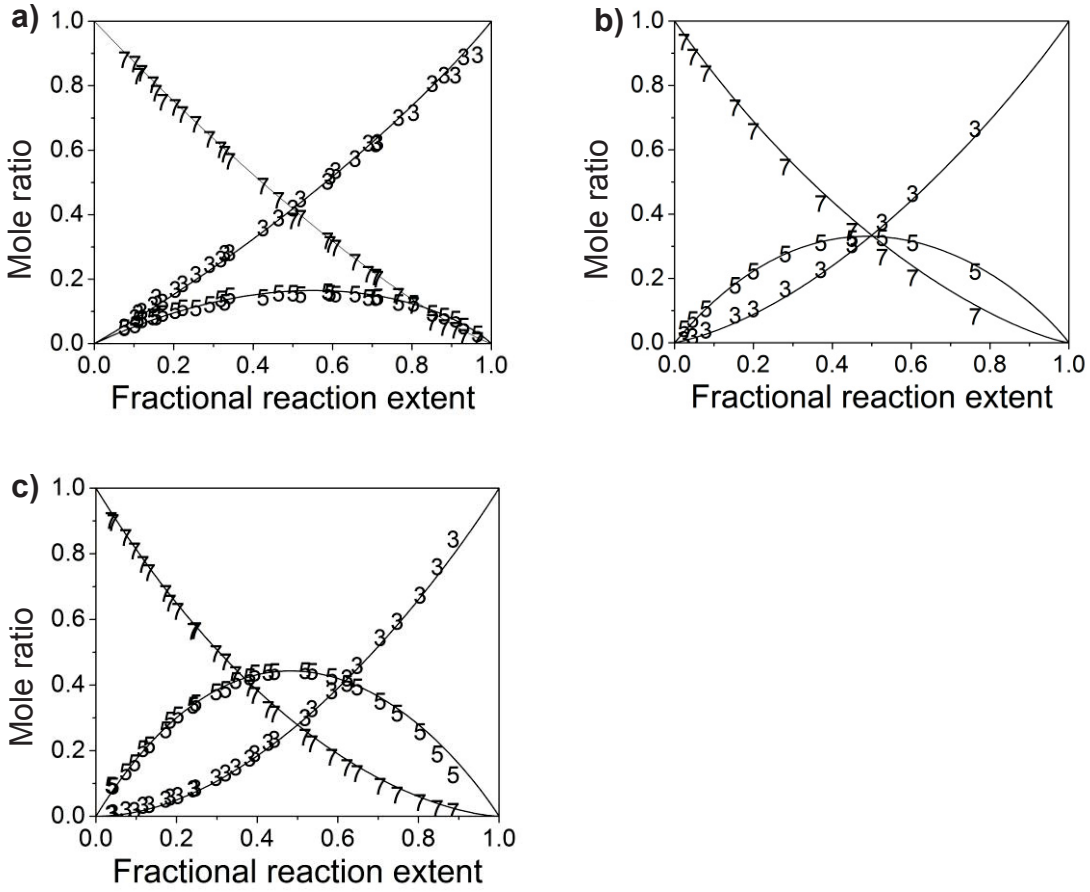


Fig. 5. Experimental data and fitted model curves for the hydrolysis of maltoheptaose by soybean β-amylase. a) Native enzyme; b) Mutant enzyme D53A; c) Mutant enzyme W55R. The mole ratio on the ordinate is calculated using Eq. (6), with $[\mathrm{G7}]_\mathrm{o}$ as the denominator. Solid lines represent the predictions of the fitted model. The number plotted as a symbol represents the length of the malto-oligomer. The experimental data were obtained from Ishikawa et al. [17].

5.3.2 Case Study 2: processive attack on maltooctaose

The second case study involves a slightly more complex reaction scheme, namely the hydrolysis of maltooctaose (G8) by a β-amylase from sweet potato. The experimental data are those of French and Youngquist [19]. This system has a complication that did not occur in the first case study: maltose is released not only as the side product, but also as the final residual product. Since French and Youngquist [19] did not provide the data for G2, the consistency of the data could not be checked. This lack of data for G2 also meant that the fractional reaction extent had to be calculated from the data for G8, G6 and G4 using Eq. (18), which, for $\mathrm{m} = 8$, can be written as:

$$F_8 = 1 - \left(\frac{3G8 + 2G6 + G4}{3}\right) \quad (22)$$

In this equation, the first term on the right-hand side represents the total normalized concentration of “attackable bonds” in the original maltooctaose present at the beginning of the reaction. The term within parentheses represents the normalized value of the attackable bonds remaining at a particular time during the reaction. In both terms, the normalization was done by dividing by $3[\mathrm{G8}]_{\mathrm{o}}$ (more details are given in the Supplementary Material).

The model in this case study is based on Fig. 6. The enzyme can bind to the initial substrate (G8) and molecules of the residual products (G6 and G4) that are free in solution. In these cases, reaction can occur. The enzyme can also bind to the maltose (G2) that is released as side product and final residual product, in this case product inhibition occurs.

When the complexes EG6 or EG4 are produced by reaction there are two possibilities. First, processive action is possible, with the enzyme proceeding directly to remove another G2 from the residual product that is still bound in the active site. Second, non-processive action is possible, with the complex dissociating to release free enzyme and the corresponding residual product molecule. It should be noted that, in the case of an enzyme hydrolyzing G8, two successive processive actions are possible after EG6 is formed (EG6→ EG4→ EG2).

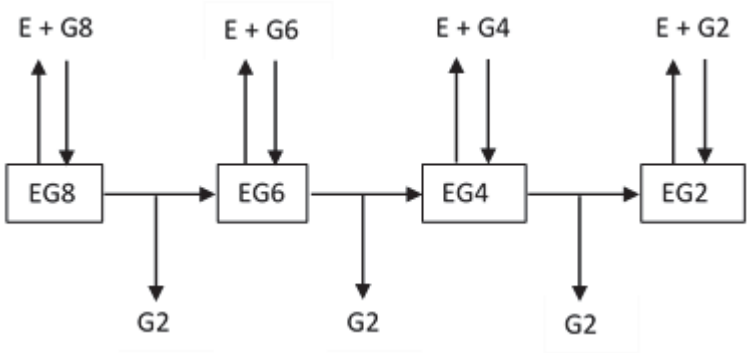


Fig. 6**.** The hydrolysis of a maltooctaose by a processive β-amylase, representing one of the systems studied by French and Youngquist [19].

The equation set that describes the scheme shown in Fig. 6 is:

$$\frac{dG8}{dF} = 3\frac{-R_8G8}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad (23)$$

$$\frac{dG6}{dF} = 3\frac{R_8G8(1-P_6)\ -R_6G6}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad (24)$$

$$\frac{dG4}{dF} = 3\frac{R_8G8P_6(1-P_4)+R_6G6(1-P_4)-R_4G4}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad (25)$$

The specificity constants (*R*-values) are expressed relative to k_8, such that, by definition, $R_8 = 1$.

The detailed deduction of the model is given in the Supplementary Material.

The model represented by Eqs. (23), (24) and (25) fits well to the data of French and Youngquist [19] for the hydrolysis of maltooctaose (Fig. 7). The values of the parameters of the model obtained in this fitting are shown in Table 4.

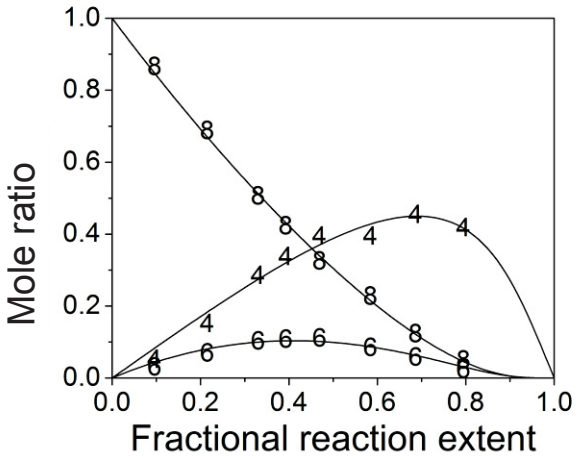


Fig. 7**.** Experimental data and fitted model curves for the hydrolysis of maltooctaose by sweet potato β-amylase. The mole ratio on the ordinate is calculated using Eq. (6), with $[\mathrm{G8}]_o$ as the denominator. Solid lines represent the predictions of the fitted model. The number plotted as a symbol represents the length of the malto-oligomer. The experimental data were obtained from French and Youngquist [19].

Table 4**.** Values of the parameters obtained by fitting the model represented by Eqs. (23) to (25) to the experimental data for even-length malto-oligomers from French and Youngquist [19].

Substrate	G8	G6	G4
R-values	1.000	1.249	0.274
P-values	**	0.684	0.223

** Processive attack is not possible for the longest malto-oligosaccharide.

The results show that the relative specificity constant varies with the degree of polymerization of the malto-oligomer. Relative to the specificity for G8, the free enzyme has a 1.25-fold higher specificity for G6 and an almost 4-fold lower specificity for G4. When formed by a reaction, the probability of the complex EG6 proceeding

processively is three-fold greater than the probability of the complex EG4 proceeding processively.

5.3.3 Case Study 3: processive attack on maltoundecaose

The third case study involves the hydrolysis of maltooligomers of uneven chain lengths by the same β-amylase from sweet potato as that analyzed in the second case study. In this case, French and Youngquist [19] provided data for the hydrolysis of three different oligomers, namely G11, G9 and G7.

Three different models were developed, one each for G11, G9 and G7 as the initial substrates. The scheme for G11 as the initial substrate is shown in Fig. 8 (for reactions initiated with G9 or G7, one can just start from the appropriate place in this scheme, ignoring the part that deals with oligosaccharides longer than the initial substrate).

The free enzyme can bind with G11, G9, G7 and G5, producing enzyme-substrate complexes that react or dissociate. The free enzyme can bind to G3 and G2, causing product inhibition. After a reaction, the enzyme-substrate complex may dissociate or proceed with the next hydrolysis processively. There may be one or more consecutive processive steps (the maximum possible number is 3, after EG9 is formed, corresponding to EG9→ EG7→ EG5→ EG3). Different probabilities of processive attack are associated with the complexes EG9 (P_9), EG7 (P_7) and EG5 (P_5).

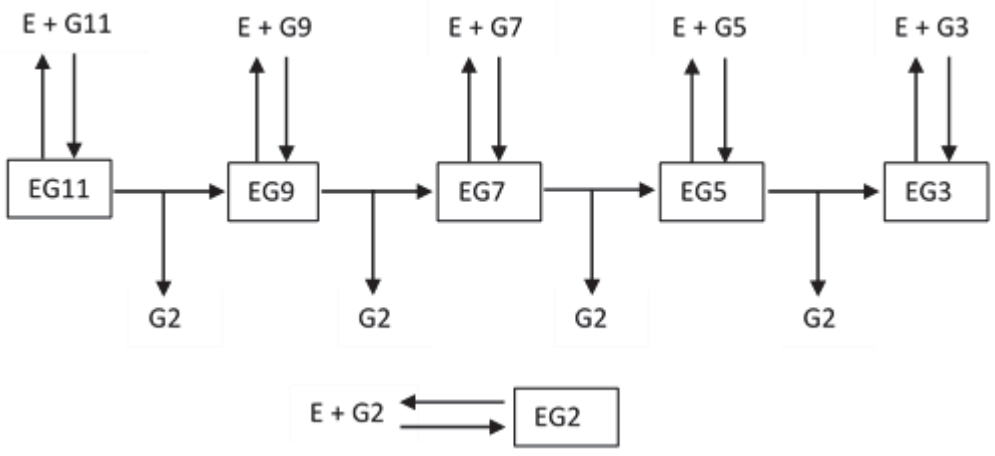


Fig. 8. The hydrolysis of a maltoundecaose by a processive β-amylase.

The three set of equations were normalized to k_7 (further explanations are given in the Supplementary Material). The equation set for the experiment starting with G11 is:

$$\frac{dG11}{dF} = 4\frac{-R_{11}G11}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (26)$$

$$\frac{dG9}{dF} = 4\frac{R_{11}G11(1-P_9)-R_9G9}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (27)$$

$$\frac{dG7}{dF} = 4\frac{R_{11}G11P_9(1-P_7)+R_9G9(1-P_7)-R_7G7}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (28)$$

$$\frac{dG5}{dF} = 4\frac{R_{11}G11P_9P_7(1-P_5)+R_9G9P_7(1-P_5)+R_7G7(1-P_5)-R_5G5}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (29)$$

$$\frac{dG3}{dF} = 4\frac{R_{11}G11P_9P_7P_5+R_9G9P_7P_5+R_7G7P_5+R_5G5}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (30)$$

The equation set for the reaction started with G9 is:

$$\frac{dG9}{dF} = 3\frac{-R_9G9}{R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (31)$$

$$\frac{dG7}{dF} = 3\frac{R_9G9(1-P_7)-R_7G7}{R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (32)$$

$$\frac{dG5}{dF} = 3\frac{R_9G9P_7(1-P_5)+R_7G7(1-P_5)-R_5G5}{R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (33)$$

$$\frac{dG3}{dF} = 3\frac{R_9G9P_7P_5+R_7G7P_5+R_5G5}{R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \quad (34)$$

Finally, the equation set for the reaction started with G7 is represented by Eqs. (19), (20) and (21). In all three equation sets, the specificity constants (*R*-values) are expressed relative to k_7, such that, by definition, $R_7 = 1$. The detailed deductions of these three equation sets are given in the Supplementary Material.

The three equation sets were fitted, simultaneously, to the three reaction profiles of French and Youngquist [19]. The fitted curves are shown in Fig. 9, while the parameters obtained in the fitting are given in the Table 5. For all three experiments, the fitted curves agreed reasonably well with the experimental data. A fitting was also done for a situation corresponding to the assumptions made by Nakatani [20], namely that the free enzyme shows no preference for any of the species (which translates to $k_5 = k_7 = k_9 = k_{11}$) and that the probability of processivity is independent of chain length (i.e. $P_5 = P_7 = P_9$). The parameters obtained from this fitting are also given in Table 5.

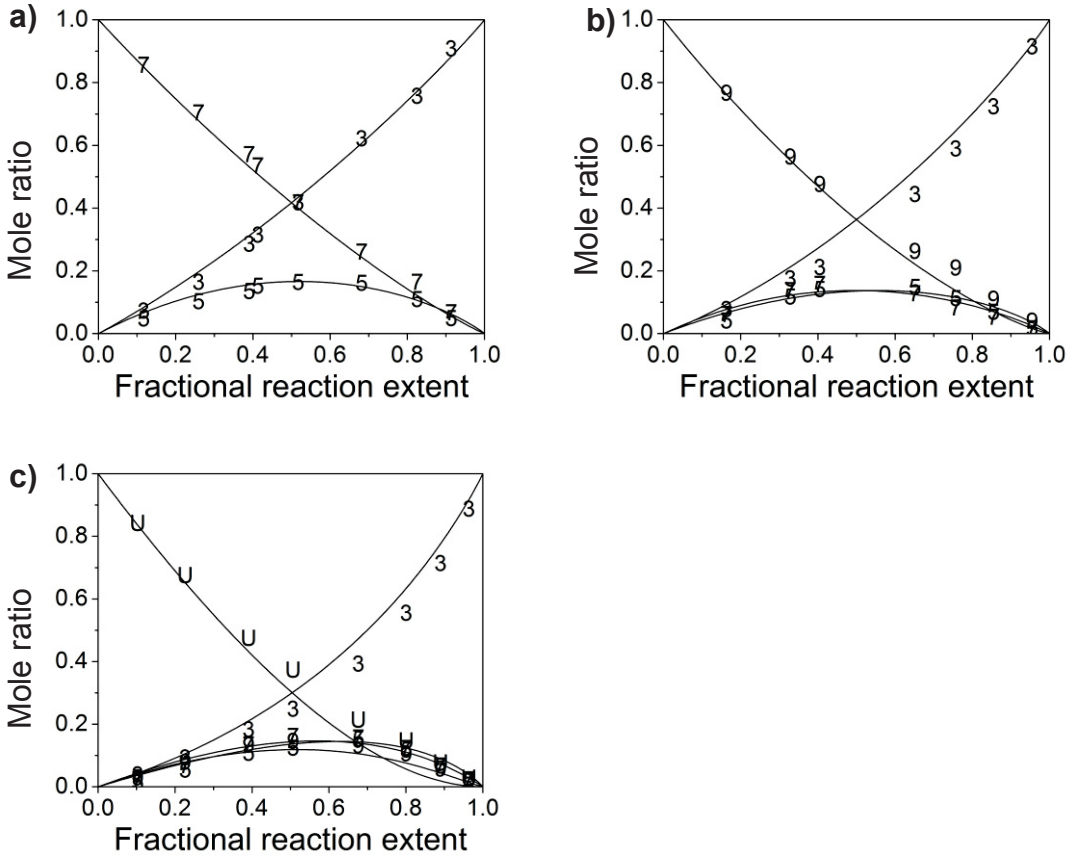


Fig. 9**.** Experimental data and fitted model curves for the hydrolysis of uneven-length maltooligomers by sweet potato β-amylase. a) maltoheptaose; b) maltononaose; c) maltoundecaose. The mole ratio on the ordinate is calculated using Eq. (6), with a) $[\mathrm{G7}]_{\mathrm{o}}$ as the denominator, b) $[\mathrm{G9}]_{\mathrm{o}}$ as the denominator and c) $[\mathrm{G11}]_{\mathrm{o}}$ as the denominator. Solid lines represent the predictions of the fitted model. The number plotted as a symbol represents the length of the malto-oligomer; U represents G11. The experimental data were obtained from French and Youngquist [19].

Table 5**.** Values of the parameters obtained by fitting the three equation sets (Eqs. (19)-(21), Eqs. (26)-(30) and Eqs. (31) – (34)) to the experimental data for uneven-length malto-oligomers from French and Youngquist [19].

Substrate	G11	G9	G7	G5
R-values‡	1.591	0.999	1.000	1.164
P-values‡	**	0.739	0.626	0.512
R-values for assumptions of Nakatani [20]*	1.000	1.000	1.000	1.000
P-values for assumptions of Nakatani [20]*	**	0.574	0.574	0.574

‡ Using assumptions of this work; * The assumptions of Nakatani [20] are $k_5 = k_7 = k_9 = k_{11}$ and $P_5 = P_7 = P_9$ ** Processive attack is not possible for the longest malto-oligosaccharide.

The results in Table 5 show that, relative to its specificity for G7, the enzyme has a 1.6-fold higher specificity for G11, but similar specificities for G9 and G5. When an enzyme-substrate complex is formed by a reaction, the probability of processive attack increases as the degree of polymerization of the bound malto-oligossacharide increases.

5.4 DISCUSSION

Four different approaches have been proposed previously for modeling processivity: both Ishikawa et al. [17] and Pereira et al. [4] developed models to describe the data of Ishikawa et al. [17], Nakatani [20] developed a model to describe the data of French and Youngquist [19] and Cruys-Bagger et al. [18] developed a model to describe the release of cellobiose during the processive attack by cellobiohydrolase on cellulose chains. Our modeling approach has advantages over the approaches used in these previous models, as we discuss in the subsections below.

5.4.1 Comparison with the approach of Ishikawa et al. [17]

Ishikawa et al. [17] proposed a stochastic model based on the Monte Carlo method to describe the reaction scheme shown in Fig. 4. However, a key assumption that they made is questionable. They define P5 as the probability of the free enzyme forming a productive complex with G5 and P7 as the probability of the free enzyme forming a productive complex with G7. In their fitting procedure, they treat the ratio P5/P7 as a constant. This is not valid, since as Nakatani [20] points out, the probability of an enzyme binding to a particular substrate within a mixture is not a simple constant, rather, it is given by an expression in which the numerator is the product of the innate preference of the enzyme for binding to that substrate and the concentration of that substrate. Since the concentrations of both G5 and G7 change over the course of the reaction, P5/P7 does not remain constant. Our model recognizes this fact, as the differential equations that we used in our model (Eqs. (19), (20) and (21)) contain the relative mole ratios, $G5$ and $G7$.

5.4.2 Comparison with the approach of Pereira et al. [4]

The approaches used in the current work and in the work of Pereira et al. [4] to model the data of Ishikawa et al. [17] are mathematically consistent with one another. In other words, it is possible to use combinations of the parameters determined by Pereira et al. [4] to estimate parameters that were estimated in the current work. Table 6 shows that, when this is done, the same estimates are obtained for the native enzyme and the mutated enzyme D53A (Pereira et al. [4] did

not fit their model to the data for the essentially non-processive mutated enzyme, W55R).

Table 6. Comparison of the fitting of the models of this work and the model of Pereira et al. [4] to the data Ishikawa et al. [17].

		This work	**Pereira et al. [4]***
Specificity of free enzyme for G5 relative to that for G7	Expression	k_5/k_7 (i.e. R_5)	$k_{5N}/(k_{7P}+k_{7N})$
	Value for native	0.97	0.97
	Value for D53A	0.89	0.89
Probability of EG5 proceeding processively #	Expression	P_5	$k_{7P}/(k_{7P}+k_{7N})$
	Value for native	0.56	0.56
	Value for D53A	0.15	0.15

* the *k*'s for Pereira et al. [4] are the specificity constants for the reactions identified in Fig. 1
# Probability of the enzyme continuing without releasing G5 after having attacked G7

However, despite the mathematical equivalence, the present work provides a conceptually better description of processivity than does the scheme of Pereira et al. [4] that is shown in Fig. 1. The scheme of Pereira et al. [4] requires the enzyme to choose between the processive and non-processive routes at the time of binding of the original substrate, in other words, before any reaction occurs. In our scheme (see Fig. 4), the choice between the processive and non-processive routes occurs for the complex EG5, after the release of the side product (G2) from the active site.

Further, there is a question of the practicality of proposing the equation set. In the case of the system studied by Ishikawa et al. [17], where there is only one processive step, there is not much difference in the effort taken between the approach of Pereira et al. [4] and that used in the current work. However, the approach of Pereira et al. [4] becomes more and more difficult to formulate when the original substrate becomes longer and more successive processive actions become possible: It would be necessary to draw a separate reaction to represent each combination of a different number of processive steps from each of the intermediates, as Fig. 10 demonstrates for attack of a processive β-amylase on maltoundecaose.

5.4.3 Comparison with the approach of Nakatani [20]

The modeling approach undertaken in the present paper has some key differences with the modeling approach of Nakatani [20], who used a stochastic model based on the Monte Carlo method to describe the reaction profile in terms of the fractional reaction extent.

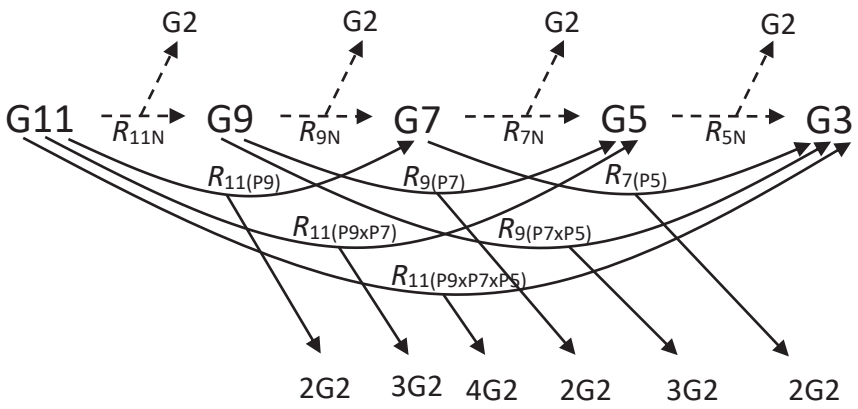


Fig. 10. Branched reaction scheme that would need to be proposed if the strategy of Pereira et al. [4] were used to represent the hydrolysis of maltoundecaose through both processive actions (continuous arrows) and non-processive actions (dashed arrows).

The first difference between our work and that of Nakatani [20] is that Nakatani [20] based the probability of the enzyme binding to an *n*-mer and then catalyzing its hydrolysis directly on the probability of formation of the *n*-mer complex. This probability is given by:

$$P_n = \frac{N_n/K_{d(n)}}{\sum_{i=f+2}^{m}\left(N_i/K_{d(i)}\right)} \tag{35}$$

where N_i is the number of molecules of the i-mer, $K_{d(i)}$ the dissociation constant for the complex between the enzyme and the i-mer, m is the degree of polymerization of the initial substrate and $f+2$ is the degree of polymerization of the smallest substrate in the reaction.

However, it is the specificity constant, and not K_{M}, that determines the relative rates of reaction. It must be remembered that two events must happen in order for the reaction to occur: first, the complex must be formed and, second, catalysis must occur (rather than the enzyme-substrate complex simply dissociating). This means that it is possible to write the following expression for the probability of reaction with species S_n, denoted as $P_{rxn(n)}$:

$$P_{rxn(n)} = \frac{k_{on(n)}[\mathrm{S}_n]}{\sum_{i=f+2}^{m} k_{on(i)}[\mathrm{S}_i]} \times \frac{k_{\mathrm{cat}(n)}}{k_{off(n)} + k_{\mathrm{cat}(n)}} \tag{36}$$

where $k_{on(n)}$ and $k_{off(n)}$ are the fundamental constants for formation and dissociation of the complex ES_n, $k_{\mathrm{cat}(n)}$ is the fundamental constant for the reaction $ES_n \rightarrow ES_{n-1} + P$, m is the degree of polymerization of the initial substrate and $f + 2$ is the degree of polymerization of the smallest substrate in the reaction.

This relationship can be rearranged to show that:

$$P_{rxn(n)} \propto \frac{k_{on(n)} k_{\mathrm{cat}(n)}}{k_{off(n)} + k_{\mathrm{cat}(n)}} [\mathrm{S}_n] \tag{37}$$

The combination of fundamental constants is equivalent to the specificity constant, so that it is possible to show that, for the scheme of Nakatani [20]:

$$P_{rxn(n)} = \frac{k_n N_n}{\sum_{i=f+2}^{m} k_i N_i} \tag{38}$$

Our analysis is consistent with the statement of Johnson [21] that "... the specificity constant is best understood as the second-order rate constant for substrate binding times the probability that, once bound, the substrate continues forward to form product".

The second difference between our work and the work of Nakatani [20] is that the values that we used on the abscissa of Fig. 7 are different from the values on the abscissa of the corresponding figure of Nakatani [20]. This is due to the fact that Nakatani [20] calculated F_8 as being equal to $[G2]/(4[G8]o)$, presumably based on the fact that each initial molecule of G8 can generate four molecules of G2. However, the fractional reaction extent is based on the number of bonds that can be hydrolyzed, and each initial molecule of G8 can only suffer 3 attacks by a maltose-removing exoenzyme (i.e. a β-amylase): the first and second attacks each liberate one molecule of G2, but the third attack liberates two molecules of G2. Our approach for calculating F_8 (see Eq. (18)) is preferable since it estimates the number of hydrolyzed bonds based on the reaction course profiles for G8, G6 and G4: at any time, the concentrations of these species give direct information about the number of hydrolysable bonds that have not yet been hydrolyzed.

The third difference between our work and the work of Nakatani [20] is that Nakatani [20] made two simplifying assumptions in modeling the hydrolysis of the

uneven oligomers that we did not make. First, he assumed that the probability of the free enzyme binding to any of the uneven oligomers (G11, G9, G7 and G5) was equal. Second, he assumed that the probabilities of processive action were independent of the length of the oligomer. Our analysis of the data of French and Youngquist [19] suggest that the specificity for the different oligomers is not independent of oligomer length. Although our estimates of R_5, R_7 and R_9 are reasonably similar, R_{11} is around 1.4- to 1.6-fold greater than these other values (Table 5). Likewise, our analysis of the data of French and Youngquist [19] suggests that the probabilities of processive action increase with increase in the length of the oligomer. The increase is not small: P_7 is approximately 20% greater than P_5 while P_9 is approximately 20% greater than P_7 (Fig 11).

The fourth difference between our work and the work of Nakatani [20] is that, in modeling the hydrolysis of the even oligomers, he did not allow for different "probabilities of processive action" for the complexes EG6 (represented in our work by P_6) and EG4 (represented in our work by P_4). Rather, he set these values to the single value that he had estimated for "the probability of processive action" in the experiment in which he started with G7 as the initial substrate. Our work suggests that P_4 is significantly lower than P_6 (Fig. 11).

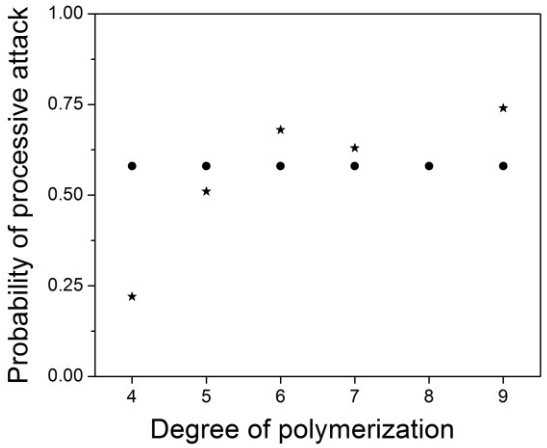


Fig. 11. Variation of the probability of processive attack as a function of the degree of polymerization of the molecule of residual product held within the enzyme-substrate complex. Key: (★) Using the assumptions made during the application of the Fingerprinting method in the current work; (●) using the assumptions made by Nakatani [20] in his stochastic model based on the Monte Carlo method.

The fifth difference between our work and the work of Nakatani [20] is that he fitted the model independently to the reaction course profiles for G11, G9 and G7 whereas we fitted all three reaction profiles simultaneously. Given that Nakatani [20]

assumed that the probability of following the processive route was independent of oligomer length, he arrived at three different estimates of this same parameter: 0.58 from the G7 data, 0.57 from the G9 data and 0.61 from the G11 data. The relative specificity constant for a particular substrate and the probability of processivity when that molecule is a residual product in the enzyme-substrate complex are independent of the length of the original substrate used in the experiment. For example, the probability of the enzyme following the processive route from EG7 through EG5 to EG3 is P_5. This processive route occurs in all three experiments (the one that starts with G11, the one that starts with G9 and the one that starts with G7). P_5 should have exactly the same value in all three experiments since it is an intrinsic parameter of the enzyme, but in the work of Nakatani [20] it does not. Our approach, in which the model was fitted to three sets of data simultaneously, generating a single estimate, is preferable to the "experiment-by-experiment" fitting approach of Nakatani [20].

5.4.4 Comparison with the approach of Cruys-Bagger et al. [18]

Cruys-Bagger et al. [18] developed an equation to describe the steady-state production of reducing sugars during the attack of a processive cellobiohydrolase on cellulose chains. They proposed the scheme shown in Fig. 2. For this scheme, they arrived at the following equation:

$$ {}_{P}V_{SS} = \frac{S_0 E_0 \; {}_{P}k_{\mathrm{on}} \; {}_{P}k_{\mathrm{cat}} \left(1 - \left(\frac{{}_{P}k_{\mathrm{cat}}}{{}_{P}k_{\mathrm{cat}} + \; {}_{P}k_{\mathrm{off}}} \right)^{n} \right)}{{}_{P}k_{\mathrm{off}} + S_0 \; {}_{P}k_{\mathrm{on}}} \tag{39} $$

where ${}_{P}V_{SS}$ is the "steady-state rate of processive hydrolysis", which they defined as the rate of cellobiose production at the start of the reaction (i.e. before the amount of cellulose available is significantly reduced), n is the number of processive reactions, E_0 and S_0 are the initial amounts of enzyme and substrate, respectively, and the fundamentals constants refer to a processive mechanism.

Although Cruys-Bagger et al. [18] made so many simplifying assumptions that it is not possible to compare their equation directly with our model (represented by Eqs. (8) to (11)), our Eq. (10) (which describes the residual products) is similar to their equation in two key aspects. First, our "P" variables are exactly equivalent to

their "$_{P}k_{\mathrm{cat}}/(\ _{P}k_{\mathrm{cat}} + \ _{P}k_{\mathrm{off}})$" term. In other words, both this term of their model and our "P" variables express the probability that a complex, when formed as the product of a previous reaction, will proceed through a new round of catalysis (rather than dissociating to produce free enzyme and product). Second, if, in a particular case, the probability of processive action is independent of the length of the substrate molecule in the active site (in our case, this probability would be denoted simply as a single variable, P), our model also describes the probability of n successive rounds of processive action occurring as being equal to the nth power of the probability of one processive action. In other words, the equation of Cruys-Bagger et al. [18] does capture the essence of successive processive actions.

However, since we did not make any simplifying assumptions, our approach is much more flexible than that of Cruys-Bagger et al. [18]. For example, the equation of Cruys-Bagger et al. [18] focuses solely on the liberation of side product, while our equation system describes the variation of concentrations of shorter intermediate substrates. Additionally, in the scheme of Cruys-Bagger et al. [18], once the enzyme finishes processive action by dissociating from the residual substrate chain, it cannot rebind to this free residual substrate chain or to any other shorter substrate chains that might be present. This means that the equation of Cruys-Bagger et al. [18] is limited to predicting the initial rate of side product formation in situations where the major substrate is the original C_m. Our model is capable of describing the whole course of the reaction.

5.5 CONCLUSION

In the current work, we developed a method for determining the relative specificity constants and probabilities of processive action for enzymes that are capable of processive actions. Our method is conceptually superior to that developed previously by Pereira et al. [4] and is more convenient to apply to reaction schemes in which multiple successive processive actions are possible. Our method is also more flexible than the model developed by Cruys-Bagger et al. [18] to describe the initial reaction rate of processive cellulases: our method is capable of describing the profiles for all intermediate-length substrates over the whole course of the reaction.

5.6 ACKNOWLEDGEMENTS

This research was supported by a “Universal” Grant from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), a Brazilian government agency for the advancement of science and technology. Research scholarships were granted to Nadia Krieger, David Mitchell and Gerson Ritzmann de Mello by CNPq, and to Aline Pereira by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), a Brazilian government agency for the development of personnel in higher education.

5.7 REFERENCES

[1] A.J. Ragauskas, C.K. Williams, B.H. Davison, G. Britovsek, J. Cairney, C.A. Eckert, W.J. Frederick, J.P. Hallett, D.J. Leak, C.L. Liotta, J.R. Mielenz, R. Murphy, R. Templer, T. Tschaplinski, The path forward for biofuels and biomaterials, Science. 311 (2006) 484–489. doi:10.1126/science.1114736.

[2] S. Mohanram, D. Amat, J. Choudhary, A. Arora, L. Nain, Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries, Sustain. Chem. Process. 1 (2013) 15. doi:10.1186/2043-7129-1-15.

[3] J.A. Mertens, Kinetic properties of two Rhizopus exo-polygalacturonase enzymes hydrolyzing galacturonic acid oligomers using isothermal titration calorimetry, Appl. Biochem. Biotechnol. 170 (2013) 2009–2020. doi:10.1007/s12010-013-0336-6.

[4] A.B. Pereira, N. Krieger, D.A. Mitchell, Fingerprinting of oligosaccharide-hydrolyzing enzymes that catalyze branched reaction schemes, Biochem. Eng. J. 113 (2016) 93–101. doi:10.1016/j.bej.2016.05.012.

[5] G.J. Davies, V. Ducros, R.J. Lewis, T. V Borchert, M. Schülein, Oligosaccharide specificity of a family 7 endoglucanase: insertion of potential sugar-binding subsites, J. Biotechnol. 57 (1997) 91–100.

[6] M. Hrmova, E.A. MacGregor, P. Biely, R.J. Stewart, G.B. Fincher, Substrate binding and catalytic mechanism of a barley beta-D-Glucosidase/(1,4)-beta-D-glucan exohydrolase, J. Biol. Chem. 273 (1998) 11134–43.

[7] E. Bonnin, A. Le Goff, R. Körner, G.W. Van Alebeek, T.M. Christensen, A.G. Voragen, P. Roepstorff, C. Caprari, J.F. Thibault, Study of the mode of action of

endopolygalacturonase from *Fusarium moniliforme*, Biochim. Biophys. Acta. 1526 (2001) 301–309.

[8] L.D. Kluskens, G.J.W.M. Van Alebeek, J. Walther, A.G.J. Voragen, W.M. De Vos, J. Van Der Oost, Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium *Thermotoga maritima*, FEBS J. 272 (2005) 5464–5473. doi:10.1111/j.1742-4658.2005.04935.x.

[9] J.A. Mertens, J.D. Braker, D.B. Jordan, Catalytic properties of two *Rhizopus oryzae* 99-880 glucoamylase enzymes without starch binding domains expressed in *Pichia pastoris*, Appl. Biochem. Biotechnol. 162 (2010) 2197–2213. doi:10.1007/s12010-010-8994-0.

[10] M. Sakaguchi, Y. Matsushima, T. Nankumo, J. Seino, S. Miyakawa, S. Honda, Y. Sugahara, F. Oyama, M. Kawakita, Glucoamylase of *Caulobacter crescentus* CB15: cloning and expression in *Escherichia coli* and functional identification, AMB Express. 4 (2014) 5. doi:10.1186/2191-0855-4-5.

[11] M. Kamon, J. Sumitani, S. Tani, T. Kawaguchi, M. Kamon, J. Sumitani, S. Tani, T. Kawaguchi, Characterization and gene cloning of a maltotriose-forming exo-amylase from *Kitasatospora* sp. MK-1785, Appl. Microbiol. Biotechnol. 99 (2015) 4743–4753. doi:10.1007/s00253-015-6396-5.

[12] D.A. Mitchell, F. Carrière, N. Krieger, An analytical method for determining relative specificities for sequential reactions catalyzed by the same enzyme: General formulation, Biochim. Biophys. Acta - Proteins Proteomics. 1784 (2008) 705–715. doi:10.1016/j.bbapap.2008.01.015.

[13] D.A. Mitchell, J.A. Rodriguez, F. Carrière, N. Krieger, Determination of the quantitative stereoselectivity fingerprint of lipases during hydrolysis of a prochiral triacylglycerol, J. Biotechnol. 135 (2008) 168–73. doi:10.1016/j.jbiotec.2008.03.013.

[14] D.A. Mitchell, J.A. Rodriguez, F. Carrière, J. Baratti, N. Krieger, An analytical method for determining relative specificities for sequential reactions catalyzed by the same enzyme: Application to the hydrolysis of triacylglycerols by lipases, J. Biotechnol. 133 (2008) 343–350. doi:10.1016/j.jbiotec.2007.10.012.

[15] A.K. Mazur, H. Nakatani, Multiple attack mechanism in the porcine pancreatic α-amylase hydrolysis of amylose and amylopectin, Arch. Biochem. Biophys. 306 (1993) 29–38. doi:10.1006/abbi.1993.1476.

[16] A.L. Lucius, N.K. Maluf, C.J. Fischer, T.M. Lohman, General methods for analysis of sequential “n-step” kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding, Biophys. J. 85 (2003) 2224–2239.

[17] K. Ishikawa, H. Nakatani, Y. Katsuya, C. Fukazawa, Kinetic and structural analysis of enzyme sliding on a substrate: multiple attack in beta-amylase, Biochemistry. 46 (2007) 792–798. doi:10.1021/bi061605w.

[18] N. Cruys-Bagger, J. Elmerdahl, E. Praestgaard, K. Borch, P. Westh, A steady-state theory for processive cellulases, FEBS J. 280 (2013) 3952–3961. doi:10.1111/febs.12397.

[19] D. French, R.W. Youngquist, Die wirkungsweise von β-amylase auf stärkeoligosaccharide, Starch - Stärke. 15 (1963) 425–431. doi:10.1002/star.19630151202.

[20] H. Nakatani, Monte Carlo simulation of multiple attack mechanism of beta-amylase-catalyzed reaction, Biopolymers. 42 (1997) 831–6.

[21] K.A. Johnson, Role of induced fit in enzyme specificity: a molecular forward/reverse switch, J. Biol. Chem. 283 (2008) 26297–301. doi:10.1074/jbc.R800034200.

5.8 SUPPLEMENTARY MATERIAL

“Fingerprinting Processive Enzymes”

by

Gerson Ritzmann de Mello, Aline Bescrovaine Pereira, Nadia Krieger, David Alexander Mitchell

The reference numbers used in this supplementary material are those of the main article.

[12] D.A. Mitchell, F. Carrière, N. Krieger, An analytical method for determining relative specificities for sequential reactions catalyzed by the same enzyme: General formulation, Biochim. Biophys. Acta - Proteins Proteomics. 1784 (2008) 705–715. doi:10.1016/j.bbapap.2008.01.015.

[17] K. Ishikawa, H. Nakatani, Y. Katsuya, C. Fukazawa, Kinetic and structural analysis of enzyme sliding on a substrate: multiple attack in beta-amylase, Biochemistry. 46 (2007) 792–8. doi:10.1021/bi061605w.

[19] D. French, R.W. Youngquist, Die wirkungsweise von β-amylase auf stärkeoligosaccharide, Starch - Stärke. 15 (1963) 425–431. doi:10.1002/star.19630151202.

5.8.1 A general model for attack of a processive exoenzyme on linear molecules

5.8.1.1 System, variables and parameters

This section develops a general model for the processive action of exoenzymes on linear substrates. Two versions are developed, a kinetic model, namely a model describing variations of system variables in time, and a model suitable for determining specificity constants by the “Fingerprinting” method, namely a model describing variations of system variables as a function of fractional reaction extent.

The system involves the attack of an exoenzyme on a linear oligosaccharide or polysaccharide containing m repeated units, denoted S_m. Each attack removes an oligomer of u units in length (S_u) from the non-reducing end of the chain (Fig. S1). The hydrolysis of the initial substrate (S_m) therefore generates two products. One is S_u,

which is liberated directly into the reaction medium. The other product of this reaction is $\mathrm{S_{m-u}}$, which, immediately after catalysis, remains complexed in the active site. In this work, the product that is liberated directly into the reaction medium after catalysis is referred to as the “side product”, while the product that remains complexed in the active site is referred to as the “residual product”. Two alternative routes are possible from the complex $\mathrm{ES_{m-u}}$. In the non-processive route, the enzyme dissociates from the substrate, liberating E and $\mathrm{S_{m-u}}$ into the reaction medium. In the processive route, the enzyme slides along $\mathrm{S_{m-u}}$ without dissociating and catalyzes another hydrolysis reaction, producing $\mathrm{ES_{m-2u}}$ and $\mathrm{S_u}$.

Successive hydrolysis reactions occur in a similar manner. Any free enzyme in the system can bind to any remaining molecules of the initial substrate or to any molecules of residual products that have been released into the medium. In general, after a reaction $\mathrm{ES_{i+u}} \rightarrow \mathrm{ES_i}$, either processive action or non-processive action is possible. The probability that the complex $\mathrm{ES_i}$ will follow the processive route is denoted P_i. The probability that the complex will follow the non-processive route is then $1 - P_\mathrm{i}$. The only exception is when the complex $\mathrm{ES_f}$ is formed, with $\mathrm{S_f}$ being the smallest possible residual product, referred to as the final residual product. In other words, the final residual product is a molecule containing the reducing end of the original substrate $\mathrm{S_m}$ and which cannot suffer further attack. Since catalysis is not possible, the binding of the enzyme to a free molecule of $\mathrm{S_f}$ represents product inhibition. Product inhibition may also occur if the enzyme binds to the side product, $\mathrm{S_u}$. Note that it is possible for the final residual product to be identical to the side product, although, if desired, it would be possible to distinguish the two experimentally by marking the reducing end units of the original oligomer or polymer in some manner (e.g. by radiolabeling).

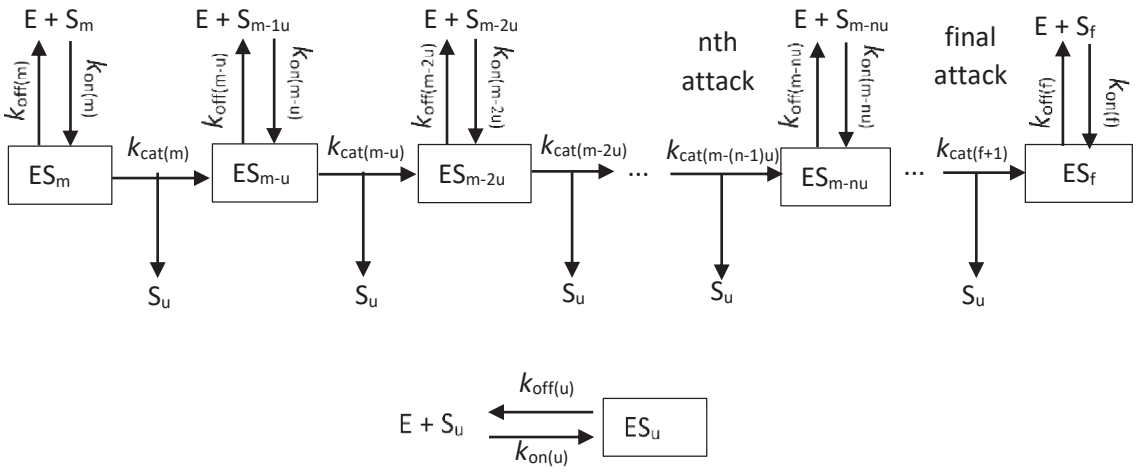


Fig. S1. General scheme for the attack of a processive exoenzyme on a linear oligosaccharide or polysaccharide containing m monosaccharide units, with the repeated removal of units containing u monosaccharide units. For each i-mer, $k_{off(i)}$ is the fundamental rate constant for dissociation of the complex, $k_{on(i)}$ is the fundamental rate constant for formation of the complex from free enzyme and the i-mer and $k_{cat(i)}$ is the catalytic rate constant.

The nomenclature used in the models is presented in Table S1.

5.8.1.2 A general kinetic model for processive action

The kinetic model is a set of differential equations representing the behavior of the mole ratios of all species, during the reaction over time.

The model begins with the following set of equations. First, balances are written for the complexes of the enzyme with the original substrate and the residual products:

$$\frac{d[\mathrm{ES_m}]}{dt} = k_{\mathrm{on(m)}}[\mathrm{E}][\mathrm{S_m}] - k_{\mathrm{off(m)}}[\mathrm{ES_m}] - k_{\mathrm{cat(m)}}[\mathrm{ES_m}] \quad \text{(S1)}$$

$$\frac{d[\mathrm{ES_{m-u}}]}{dt} = k_{\mathrm{on(m-u)}}[\mathrm{E}][\mathrm{S_{m-u}}] - k_{\mathrm{off(m-u)}}[\mathrm{ES_{m-u}}] + k_{\mathrm{cat(m)}}[\mathrm{ES_m}] - k_{\mathrm{cat(m-u)}}[\mathrm{ES_{m-u}}] \quad \text{(S2)}$$

$$\frac{d[\mathrm{ES_{m-2u}}]}{dt} = k_{\mathrm{on(m-2u)}}[\mathrm{E}][\mathrm{S_{m-2u}}] - k_{\mathrm{off(m-2u)}}[\mathrm{ES_{m-2u}}] + k_{\mathrm{cat(m-u)}}[\mathrm{ES_{m-u}}] - k_{\mathrm{cat(m-2u)}}[\mathrm{ES_{m-2u}}] \quad \text{(S3)}$$

Similar balances are written for each complex in which the enzyme is capable of catalyzing. For the nth residual product:

$$\frac{d[\mathrm{ES_n}]}{dt} = k_{\mathrm{on(n)}}[\mathrm{E}][\mathrm{S_n}] - k_{\mathrm{off(n)}}[\mathrm{ES_n}] + k_{\mathrm{cat(n+u)}}[\mathrm{ES_{n+u}}] - k_{\mathrm{cat(n)}}[\mathrm{ES_n}] \quad \text{(S4)}$$

Table S1. Nomenclature used in the general model for attack of a processive exoenzyme on a linear oligosaccharide or polysaccharide

Symbol	Meaning
Species	
E	Enzyme
$\mathrm{S_i}$	Generic substrate i-mer
Variables	
D_m	Denominator of the rate equations for a reaction initiated with $\mathrm{S_m}$
F_m	Fractional reaction extent for a reaction initiated with $\mathrm{S_m}$
S_i	Mole ratio of the generic substrate i-mer (S in italic font)
$[\mathrm{S_i}]$	Concentration of the generic substrate i-mer
t	Time (s)
Parameters	
$k_{\mathrm{cat(i)}}$	Catalytic constant for the generic substrate i-mer (s^{-1})
$k_{\mathrm{off(i)}}$	Fundamental rate constant for dissociation of the complex $\mathrm{ES_i}$ into $\mathrm{S_i}$ and E
$k_{\mathrm{on(i)}}$	Fundamental rate constant for formation of the complex $\mathrm{ES_i}$ from $\mathrm{S_i}$ and E
$K_{\mathrm{M(i)}}$	Michaelis-Menten constant for the generic substrate i-mer (mol L^{-1})
k_i	Specificity constant for the generic substrate i-mer (L mol^{-1} s^{-1})
P_i	Probability of processive attack proceeding from the complex $\mathrm{ES_i}$ (such that the probability of the complex $\mathrm{ES_i}$ liberating $\mathrm{S_i}$ into solution is $1 - P_\mathrm{i}$)
R_i	Relative specificity constant for the i-mer (i.e. ratio of the specificity constant k_i to the specificity constant taken as a reference)
Subscripts	
o	Initial value
T	Total enzyme in the system (free and complexed)
m	Degree of polymerization of the largest substrate present in the system (i.e. of the substrate used to initiate the reaction)
f	Degree of polymerization of the final residual product (the smallest residual product, one that cannot be further attacked; it contains the reducing end of the initial substrate molecule)
n	Degree of polymerization of the oligomer for which the differential equation is being written
u	Degree of polymerization of the side product (which comes from the non-reducing end of the substrate; identical side products are removed from the initial substrate and the residual products in successive rounds of hydrolysis)

When the enzyme is complexed with the final residual product or with the side product, it cannot catalyze a reaction, so the balances for the complexes $\mathrm{ES_f}$ and $\mathrm{ES_u}$ are given by:

$$\frac{d[\mathrm{ES_f}]}{dt} = k_{\mathrm{on(f)}}[\mathrm{E}][\mathrm{S_f}] - k_{\mathrm{off(f)}}[\mathrm{ES_f}] + k_{\mathrm{cat(f+u)}}[\mathrm{ES_{f+u}}] \quad \text{(S5)}$$

$$\frac{d[\mathrm{ES_u}]}{dt} = k_{\mathrm{on(u)}}[\mathrm{E}][\mathrm{S_u}] - k_{\mathrm{off(u)}}[\mathrm{ES_u}] \quad \text{(S6)}$$

The total enzyme in the system is either free or in one of the complexes, therefore:

$$[\mathrm{E}]_{\mathrm{T}} = [\mathrm{E}] + \sum_{\mathrm{i=f}}^{\mathrm{m}} [\mathrm{ES_i}] + [\mathrm{ES_u}] \tag{S7}$$

The balance on the original substrate is given by:

$$\frac{d[\mathrm{S_m}]}{dt} = k_{\mathrm{off(m)}}[\mathrm{ES_m}] - k_{\mathrm{on(m)}}[\mathrm{E}][\mathrm{S_m}] \tag{S8}$$

The balance on the nth residual product is given by:

$$\frac{d[\mathrm{S_n}]}{dt} = k_{\mathrm{off(n)}}[\mathrm{ES_n}] - k_{\mathrm{on(n)}}[\mathrm{E}][\mathrm{S_n}] \tag{S9}$$

The balance on the final residual product is given by:

$$\frac{d[\mathrm{S_f}]}{dt} = k_{\mathrm{off(n)}}[\mathrm{ES_f}] - k_{\mathrm{on(n)}}[\mathrm{E}][\mathrm{S_f}] \tag{S10}$$

Finally, the balance on the side product is given by:

$$\frac{d[\mathrm{S_u}]}{dt} = k_{\mathrm{off(u)}}[\mathrm{ES_u}] - k_{\mathrm{on(u)}}[\mathrm{E}][\mathrm{S_u}] + \sum_{\mathrm{i=f+u}}^{\mathrm{m}} \left(k_{\mathrm{cat(i)}}[\mathrm{ES_i}]\right) \tag{S11}$$

The quasi-steady state assumption is used to deduce the final kinetic equations for the various species from the balance equations presented above. During this deduction, various different combinations of the fundamental rate constants (i.e. k_{on}, k_{off} and k_{cat}) are replaced by parameters that will appear in the final equation set.

The probability of processive action, P_{i}, which can vary with the degree of polymerization of the substrate, is given by:

$$P_{\mathrm{i}} = \frac{k_{\mathrm{cat(i)}}}{k_{\mathrm{cat(i)}} + k_{\mathrm{off(i)}}} \tag{S12}$$

It should be noted that:

$$1 - P_{\mathrm{i}} = \frac{k_{\mathrm{off(i)}}}{k_{\mathrm{cat(i)}} + k_{\mathrm{off(i)}}} \tag{S13}$$

$K_{\mathrm{M(f)}}$ and $K_{\mathrm{M(u)}}$ are dissociation constants for the complexes $\mathrm{ES_f}$ and $\mathrm{ES_u}$, respectively:

$$K_{\mathrm{M(f)}} = \frac{k_{\mathrm{off(f)}}}{k_{\mathrm{on(f)}}} \tag{S14}$$

$$K_{\mathrm{M(u)}} = \frac{k_{\mathrm{off(u)}}}{k_{\mathrm{on(u)}}} \tag{S15}$$

The other $K_{\mathrm{M(i)}}$ parameters are saturation constants for binding of the i-mer and are given by:

$$K_{\mathrm{M(i)}} = \frac{k_{\mathrm{cat(i)}} + k_{\mathrm{off(i)}}}{k_{\mathrm{on(i)}}} \tag{S16}$$

The specificity constant of the enzyme for the i-mer, k_{i}, is given by:

$$k_{\mathrm{i}} = \frac{k_{\mathrm{cat(i)}}}{K_{\mathrm{M(i)}}} = \frac{k_{\mathrm{cat(i)}} k_{\mathrm{on(i)}}}{k_{\mathrm{cat(i)}} + k_{\mathrm{off(i)}}} \tag{S17}$$

From Eqs. (S1) to (S17), it can be deduced that the variation in the concentration of the original substrate is given by:

$$\frac{d[\mathrm{S_m}]}{dt} = -k_{\mathrm{m}}[\mathrm{S}_m]\frac{[\mathrm{E}]_{\mathrm{T}}}{D_{\mathrm{m}}} \tag{S18}$$

The variation in the concentration of an intermediate residual product $\mathrm{S_n}$ is given by:

$$\frac{d[\mathrm{S_n}]}{dt} = \frac{[\mathrm{E}]_\mathrm{T}}{D_\mathrm{m}}\left(\sum_{\mathrm{i=n+u}}^{\mathrm{m}}\left(k_\mathrm{i}[\mathrm{S_i}]\left(\prod_{\mathrm{j=n+u}}^{\mathrm{i-u}} P_\mathrm{j}\right)(1 - P_\mathrm{n})\right) - k_n[\mathrm{S_n}]\right) \tag{S19}$$

There is one such equation for each intermediate $\mathrm{S_n}$ (between $\mathrm{S_{m-u}}$ and $\mathrm{S_{f+u}}$, inclusive). Figure S2 demonstrates the logic of Eq. (S19) for the case $\mathrm{n = m - 3u}$. Within the sum, it may be reasonable to ignore terms in which the products of the probabilities of processivity (i.e. $\prod_{\mathrm{j=n+u}}^{\mathrm{i-u}} P_\mathrm{j}$) become very small, such that the term makes an insignificant contribution to the generation of $\mathrm{S_n}$.

The accumulation of the final residual product $\mathrm{S_f}$ and the side product $\mathrm{S_u}$ are given by:

$$\frac{d[\mathrm{S_f}]}{dt} = \frac{[\mathrm{E}]_\mathrm{T}}{D_\mathrm{m}}\sum_{\mathrm{i=f+u}}^{\mathrm{m}}\left(k_\mathrm{i}[\mathrm{S_i}]\left(\prod_{\mathrm{j=f+u}}^{\mathrm{i-u}} P_\mathrm{j}\right)\right) \tag{S20}$$

$$\frac{d[\mathrm{S_u}]}{dt} = \frac{[\mathrm{E}]_\mathrm{T}}{D_\mathrm{m}}\left(\sum_{\mathrm{i=f+u}}^{\mathrm{m}}\left(k_\mathrm{i}[\mathrm{S_i}]\left(1 + \sum_{\mathrm{g=f+u}}^{\mathrm{i-u}}\left(\prod_{\mathrm{j=g}}^{\mathrm{i-u}} P_\mathrm{j}\right)\right)\right)\right) \tag{S21}$$

In these equations the denominator D_m is a term considering all the possible forms of the enzyme in the system: free, inhibited, or complexed. It is given by:

$$D_\mathrm{m} = 1 + \sum_{\mathrm{i=f+2}}^{\mathrm{m}}\frac{[\mathrm{S}_i]}{K_{\mathrm{M(i)}}}\left(1 + k_{\mathrm{cat(i)}}\sum_{\mathrm{j=f}}^{\mathrm{i-u}}\frac{\left(\prod_{\mathrm{g=j+u}}^{\mathrm{i-u}} P_\mathrm{g}\right)}{\left(k_{\mathrm{cat(j)}} + k_{\mathrm{off(j)}}\right)}\right) + \frac{[\mathrm{S_u}]}{K_{\mathrm{M(u)}}} + \frac{[\mathrm{S}_f]}{K_{\mathrm{M(f)}}} \tag{S22}$$

where $k_{\mathrm{cat(f)}} = 0$. It should be noted that, in any particular hydrolysis experiment, D_m is identical, independently of which substrate $\mathrm{S_n}$ is being considered, although it is different for experiments initiated with substrates of different degrees of polymerization.

production of S_{m-3u} from S_m, in a 3-step reaction (2 successive processive actions)

production of S_{m-3u} from S_{m-u}, in a 2-step reaction (1 processive action).

$$\frac{d[S_{m-3u}]}{dt} = k_m[S_m]\frac{[E]_T}{D_m}P_{m-u}P_{m-2u}(1 - P_{m-3u}) + k_{m-u}[S_{m-u}]\frac{[E]_T}{D_m}P_{m-2u}(1 - P_{m-3u})$$

$$+ k_{m-2u}[S_{m-2u}]\frac{[E]_T}{D_m}(1 - P_{m-3u}) - k_{m-3u}[S_{m-3u}]\frac{[E]_T}{D_m}$$

production of S_{m-3u} from S_{m-2u} in a 1-step reaction (non-processive)

total flux from $E + S_{m-3u}$ through ES_{m-3u} to ES_{m-4u}, without specifying how many processive steps will be undertaken, if any, after ES_{m-4u} is formed

$$P_i = \frac{k_{cat(i)}}{k_{cat(i)} + k_{off(i)}} \qquad 1 - P_i = \frac{k_{off(i)}}{k_{cat(i)} + k_{off(i)}}$$

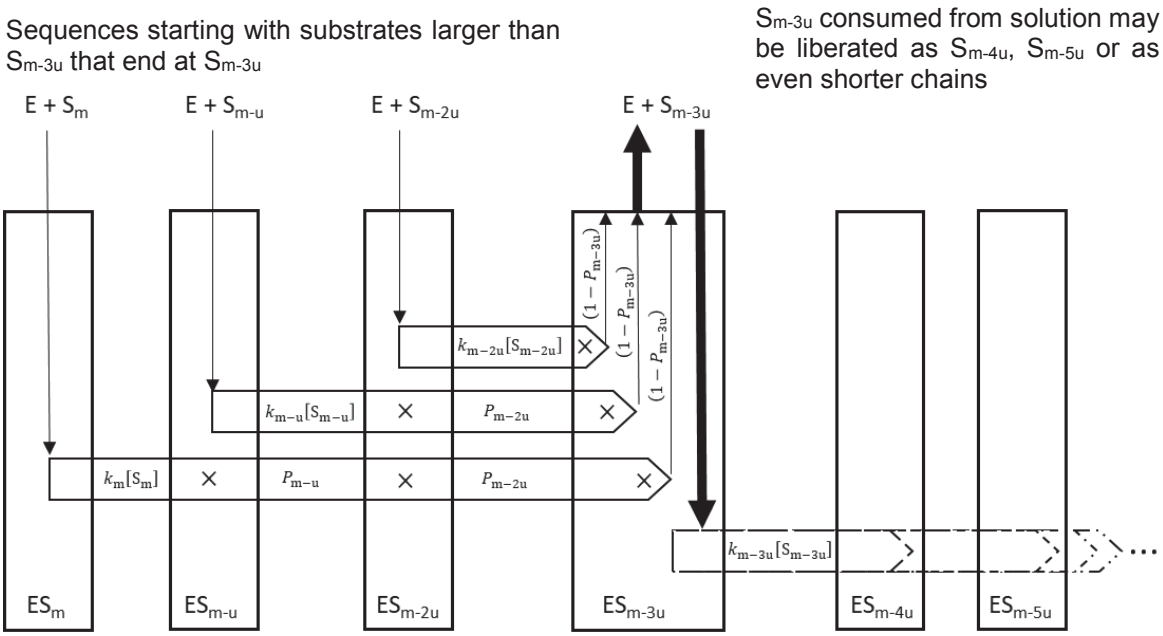


Fig. S2. Demonstration of the logic of the construction of Eq. (S19) (which is reproduced within the figure), shown for the particular case of the balance on the residual product S_{m-3u}. Many different reaction sequences are possible between the intermediates that are shown, but the equation contains only terms that affect the concentration of S_{m-3u}. The origins of these terms are represented in the diagram. Three sequences produce ES_{m-3u}, which then dissociates, liberating S_{m-3u} into solution. Reaction sequences that start with substrates longer than S_{m-3u}, but that pass processively through $ES_{m-3u} \rightarrow ES_{m-4u}$ do not appear in Eq. (S19) since they do not affect the concentration of S_{m-3u} in the reaction medium. They are not shown in the figure. The thick down arrow represents the total consumption of S_{m-3u} from solution, regardless of the number of processive steps that occur after the complex ES_{m-4u} is produced.

5.8.1.3 A general model based on fractional reaction extent

Following the Fingerprinting method [12], the independent variable can be transformed from time to fractional reaction extent (F_{m}). F_{m} is defined as the ratio of the number of bonds that have been hydrolyzed to the total number of bonds in the initial substrate:

$$F_{\mathrm{m}} = \frac{[\mathrm{S_u}]}{\left(\frac{\mathrm{m}-\mathrm{f}}{\mathrm{u}}\right)[\mathrm{S_m}]_{\mathrm{o}}} \tag{S23}$$

If data is not provided for $[\mathrm{S_u}]$ but data for the original substrate, the residual products and the final residual product are available, then $[\mathrm{S_u}]$ can be calculated as:

$$[\mathrm{S_u}] = \left(\frac{\mathrm{m}-\mathrm{f}}{\mathrm{u}}\right)[\mathrm{S_m}]_{\mathrm{o}} - \sum_{\mathrm{j=f}}^{\mathrm{m}}\left\{\left(\frac{\mathrm{j}-\mathrm{f}}{\mathrm{u}}\right)[\mathrm{S_j}]\right\} \tag{S24}$$

The derivative of Eq. (S23) with respect to time is:

$$\frac{dF_{\mathrm{m}}}{dt} = \left(\frac{1}{\left(\frac{\mathrm{m}-\mathrm{f}}{\mathrm{u}}\right)[\mathrm{S_m}]_{\mathrm{o}}}\right)\frac{d[\mathrm{S_u}]}{dt} \tag{S25}$$

The chain rule can then be applied. This rule states that, for any species $\mathrm{S_i}$:

$$\frac{d[\mathrm{S_i}]}{dF} = \left(\frac{d[\mathrm{S_i}]}{dt}\right)\Big/\left(\frac{dF}{dt}\right) \tag{S26}$$

When the division on the right-hand side of Eq. (S26) is performed, the total enzyme concentration ($[\mathrm{E}]_{\mathrm{T}}$) and the denominator that is common to all kinetic equations (D_{m}) both cancel out of all equations in the final equation set [12].

It is often convenient to convert all concentrations to mole ratios (i.e. S_{i}, namely with the "S" in italic font), by dividing all concentrations by the initial concentration of the original substrate:

$$S_i = \frac{[\mathrm{S_i}]}{[\mathrm{S_m}]_\mathrm{o}} \tag{S27}$$

Also, since specificity constants appear in all numerators and denominators of the final equation set, it is not possible to determine the absolute values of the specificity constants, only their relative values [12]. Therefore, one specificity constant is chosen as the basis for expression of the relative values of the others. The i^{th} relative specificity constant is then given by:

$$R_\mathrm{i} = \frac{k_\mathrm{i}}{k_\mathrm{r}} \tag{S28}$$

where k_r indicates the specificity constant that is chosen as the basis for normalizing the others. By definition, R_r = 1.0.

The final equation set, which can be used in the fitting procedure of the fingerprinting method, is then given by:

$$\frac{dS_\mathrm{m}}{dF} = \frac{\left(\frac{\mathrm{m-f}}{\mathrm{u}}\right)}{D_\mathrm{m}^{\#}}\left(-R_\mathrm{m}S_\mathrm{m}\right) \tag{S29}$$

$$\frac{dS_\mathrm{n}}{dF} = \frac{\left(\frac{\mathrm{m-f}}{\mathrm{u}}\right)}{D_\mathrm{m}^{\#}}\left(\sum_{\mathrm{i=n+u}}^{\mathrm{m}}\left(R_\mathrm{i}S_\mathrm{i}\left(1-P_\mathrm{n}\right)\prod_{\mathrm{j=n+u}}^{\mathrm{i-u}}P_\mathrm{j}\right)-R_\mathrm{n}S_\mathrm{n}\right) \tag{S30}$$

$$\frac{dS_\mathrm{f}}{dF} = \frac{\left(\frac{\mathrm{m-f}}{\mathrm{u}}\right)}{D_\mathrm{m}^{\#}}\left(\sum_{\mathrm{i=f+u}}^{\mathrm{m}}\left(R_\mathrm{i}S_\mathrm{i}\prod_{\mathrm{j=f+u}}^{\mathrm{i-u}}P_\mathrm{j}\right)\right) \tag{S31}$$

$$\frac{dS_\mathrm{u}}{dF} = \left(\frac{\mathrm{m-f}}{\mathrm{u}}\right) \tag{S32}$$

$$D_\mathrm{m}^{\#} = \sum_{\mathrm{i=f+u}}^{\mathrm{m}}\left(R_\mathrm{i}S_\mathrm{i}\left(1+\sum_{\mathrm{g=f+u}}^{\mathrm{i-u}}\left(\prod_{\mathrm{j=g}}^{\mathrm{i-u}}P_\mathrm{j}\right)\right)\right) \tag{S33}$$

5.8.2 Case Study 1

The first case study is developed for the hydrolysis of maltoheptaose by β-amylase, using the experimental data of Ishikawa et al. [17]. The proposed scheme for the reaction is shown in Fig. S3.

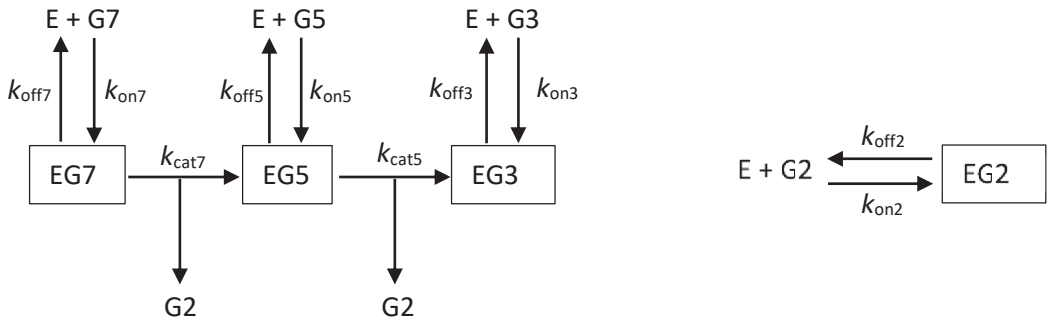


Fig. S3. Reaction scheme for the hydrolysis of maltoheptaose by β-amylase. For each i-mer, $k_{off(i)}$ is the fundamental rate constant for dissociation of the complex; $k_{on(i)}$ is the fundamental rate constant for formation of the complex from free enzyme and the i-mer and $k_{cat(i)}$ is the catalytic rate constant.

In this system, maltoheptahose (G7) is hydrolyzed to the intermediate maltopentaose (G5) and then to the final product maltotriose (G3), releasing one maltose (G2) in each step. The probability that the enzyme will follow the processive route after it has produced the complex EG5 through reaction is P_5. The side product (G2) and the final residual product (G3) are assumed to cause product inhibition.

The set of balance equations for all the complexes that the enzyme can form in the system is:

$$\frac{d[\mathrm{EG7}]}{dt} = k_{\mathrm{on7}}[\mathrm{G7}][\mathrm{E}] - k_{\mathrm{off7}}[\mathrm{EG7}] - k_{\mathrm{cat7}}[\mathrm{EG7}] \tag{S34}$$

$$\frac{d[\mathrm{EG5}]}{dt} = k_{\mathrm{on5}}[\mathrm{G5}][\mathrm{E}] - k_{\mathrm{off5}}[\mathrm{EG5}] - k_{\mathrm{cat5}}[\mathrm{EG5}] + k_{\mathrm{cat7}}[\mathrm{EG7}] \tag{S35}$$

$$\frac{d[\mathrm{EG3}]}{dt} = k_{\mathrm{on3}}[\mathrm{G3}][\mathrm{E}] - k_{\mathrm{off3}}[\mathrm{EG3}] + k_{\mathrm{cat5}}[\mathrm{EG5}] \tag{S36}$$

$$\frac{d[\mathrm{EG2}]}{dt} = k_{\mathrm{on2}}[\mathrm{G2}][\mathrm{E}] - k_{\mathrm{off2}}[\mathrm{EG2}] \tag{S37}$$

The total enzyme in the system ($[\mathrm{E}]_{\mathrm{T}}$) is represented by:

$$[\mathrm{E}]_{\mathrm{T}} = [\mathrm{E}] + [\mathrm{EG7}] + [\mathrm{EG5}] + [\mathrm{EG3}] + [\mathrm{EG2}] \quad \text{(S38)}$$

From Eqs. (S34) to (S38), it can be shown that the free enzyme in the system is:

$$[\mathrm{E}] = \frac{[\mathrm{E}]_{\mathrm{T}}}{D_7} \quad \text{(S39)}$$

where the denominator D_7 represents all the forms of the enzyme at any particular time during the reaction. It is given by:

$$D_7 = 1 + \left(\frac{[\mathrm{G7}]}{K_{\mathrm{M7}}} + \frac{k_7[\mathrm{G7}]}{k_{\mathrm{off5}} + k_{\mathrm{cat5}}} + \frac{k_7[\mathrm{G7}]P_5}{k_{\mathrm{off3}}}\right) + \left(\frac{[\mathrm{G5}]}{K_{\mathrm{M5}}} + \frac{k_5[\mathrm{G5}]}{k_{\mathrm{off3}}}\right) + \frac{[\mathrm{G3}]}{K_{\mathrm{M3}}} + \frac{[\mathrm{G2}]}{K_{\mathrm{M2}}} \quad \text{(S40)}$$

In Eq. (S40), K_{M7} and K_{M5} are the saturation constants for binding of G7 and G5, respectively, and K_{M3} and K_{M2} are the inhibition constants for G3 and G2, respectively. Table S2 explains the origins of the various terms in this denominator.

Table S2. Meanings of the terms within the denominator of the rate equations (i.e. within the expression for D_7)

Term of Eq. (S40)	Form of the enzyme represented by the term
1	free enzyme
$\frac{[\mathrm{G7}]}{K_{\mathrm{M7}}}$	complex EG7 originating from direct binding of G7 to free enzyme
$\frac{k_7[\mathrm{G7}]}{k_{\mathrm{off5}} + k_{\mathrm{cat5}}}$	complex EG5 originating from the sequence E+G7→EG7→EG5
$\frac{k_7[\mathrm{G7}]P_5}{k_{\mathrm{off3}}}$	complex EG3 originating from the sequence E+G7→EG7→EG5→EG3
$\frac{[\mathrm{G5}]}{K_{\mathrm{M5}}}$	complex EG5 originating from direct binding of G5 to free enzyme
$\frac{k_5[\mathrm{G5}]}{k_{\mathrm{off3}}}$	complex EG3 originating from the sequence E+G5→EG5→EG3
$\frac{[\mathrm{G3}]}{K_{\mathrm{M3}}}$	complex EG3 originating from direct binding of G3 to free enzyme
$\frac{[\mathrm{G2}]}{K_{\mathrm{M2}}}$	complex EG2 originating from direct binding of G2 to free enzyme

The balance equations for the original substrate, all residual products and the side product are:

$$\frac{d[\mathrm{G7}]}{dt} = -k_7[\mathrm{G7}]\frac{[\mathrm{E}]_\mathrm{T}}{D_7} \tag{S41}$$

$$\frac{d[\mathrm{G5}]}{dt} = k_7[\mathrm{G7}](1-P_5)\frac{[\mathrm{E}]_\mathrm{T}}{D_7} - k_5[\mathrm{G5}]\frac{[\mathrm{E}]_\mathrm{T}}{D_7} \tag{S42}$$

$$\frac{d[\mathrm{G3}]}{dt} = k_7[\mathrm{G7}]P_5\frac{[\mathrm{E}]_\mathrm{T}}{D_7} + k_5[\mathrm{G5}]\frac{[\mathrm{E}]_\mathrm{T}}{D_7} \tag{S43}$$

$$\frac{d[\mathrm{G2}]}{dt} = k_7[\mathrm{G7}](1+P_5)\frac{[\mathrm{E}]_\mathrm{T}}{D_7} + k_5[\mathrm{G5}]\frac{[\mathrm{E}]_\mathrm{T}}{D_7} \tag{S44}$$

The independent variable in Eqs. (S41) to (S44) is changed from time to fractional reaction extent (F_7). The scheme in Fig. S3 shows that, for every hydrolyzed bond, only one G2 is liberated. Therefore, F_7 can be written as:

$$F_7 = \frac{[\mathrm{G2}]}{2[\mathrm{G7}]_\mathrm{o}} \tag{S45}$$

where $[\mathrm{G7}]_\mathrm{o}$ is the initial amount of the substrate G7. The derivative of Eq. (S45) with respect to time is:

$$\frac{dF_7}{dt} = \left(\frac{1}{2[\mathrm{G7}]_\mathrm{o}}\right)\frac{d[\mathrm{G2}]}{dt} \tag{S46}$$

Replacing $d[G2]/dt$ in Eq. (S46) with the right-hand side of Eq. (S44) gives:

$$\frac{dF_7}{dt} = \frac{1}{2[\mathrm{G7}]_\mathrm{o}}(k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}])\frac{[\mathrm{E}]_\mathrm{T}}{D_7} \tag{S47}$$

Dividing Eqs. (S41) to (S44) by Eq. (S47) and applying the chain rule gives:

$$\frac{d[\mathrm{G7}]}{dF} = (2[\mathrm{G7}]_\mathrm{o})\frac{-k_7[\mathrm{G7}]}{k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S48}$$

$$\frac{d[\mathrm{G5}]}{dF} = (2[\mathrm{G7}]_\mathrm{o})\frac{k_7[\mathrm{G7}](1-P_5) - k_5[\mathrm{G5}]}{k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S49}$$

$$\frac{d[\mathrm{G3}]}{dF} = (2[\mathrm{G7}]_\mathrm{o})\frac{k_7[\mathrm{G7}]P_5 + k_5[\mathrm{G5}]}{k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S50}$$

$$\frac{d[\mathrm{G2}]}{dF} = 2[\mathrm{G7}]_\mathrm{o} \tag{S51}$$

It is convenient to express the concentrations of all species relative to the concentration of the initial substrate as mole ratios (in other words, as $Y = [\mathrm{Y}]/[\mathrm{G7}]_\mathrm{o}$, where Y (in italic font) is the "molar ratio" of species Y). Additionally, dividing both the denominator and the numerator on the right-hand sides of Eqs. (S48) to (S51) by k_7 leads to *R*-parameters that represent relative specificity constants. These two transformations lead to the following set of equations:

$$\frac{dG7}{dF} = 2\frac{\frac{-k_7}{k_7}G7}{\frac{k_7}{k_7}G7(1+P_5)+\frac{k_5}{k_7}G5} = 2\frac{-R_7G7}{R_7G7(1+P_5)+R_5G5} \tag{S52}$$

$$\frac{dG5}{dF} = 2\frac{\frac{k_7}{k_7}G7(1-P_5)-\frac{k_5}{k_7}G5}{\frac{k_7}{k_7}G7(1+P_5)+\frac{k_5}{k_7}G5} = 2\frac{R_7G7(1-P_5)-R_5G5}{R_7G7(1+P_5)+R_5G5} \tag{S53}$$

$$\frac{dG3}{dF} = 2\frac{\frac{k_7}{k_7}G7P_5+\frac{k_5}{k_7}G5}{\frac{k_7}{k_7}G7(1+P_5)+\frac{k_5}{k_7}G5} = 2\frac{R_7G7P_5+R_5G5}{R_7G7(1+P_5)+R_5G5} \tag{S54}$$

$$\frac{dG2}{dF} = 2 \tag{S55}$$

where R_7, by definition, is equal to 1.

5.8.3 Case Study 2

Case Study 2 is developed for the hydrolysis of maltooctaose (G8), using experimental data from French and Youngquist [19]. The scheme in Fig. S4 is slightly more complex than the scheme from Case Study 1: there is one more processive step in the complete hydrolysis scheme and the substrate has an even degree of polymerization.

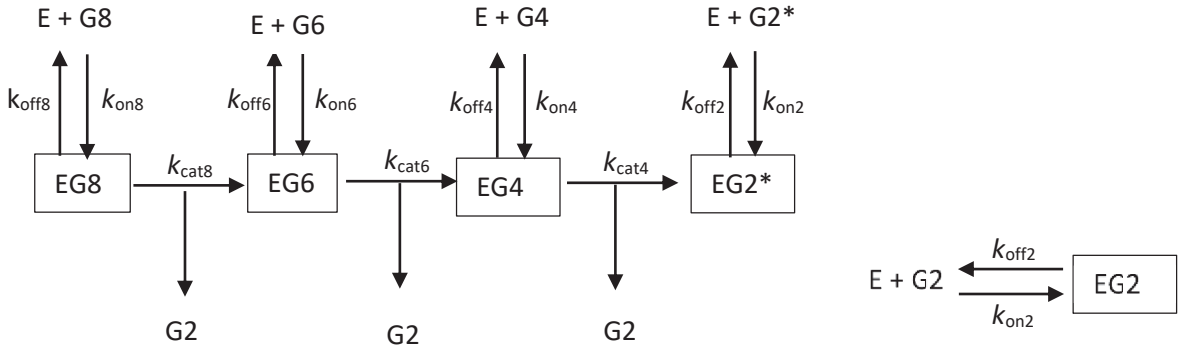


Fig. S4**.** Reaction scheme for the hydrolysis of maltooctaose by β-amylase. For each i-mer, $k_{off(i)}$ is the fundamental rate constant for dissociation of the complex; $k_{on(i)}$ is the fundamental rate constant for formation of the complex from free enzyme and the i-mer and $k_{cat(i)}$ is the catalytic rate constant. The symbol G2* denotes that this maltose corresponds to the maltose at the reducing end of the original G8 molecule.

The enzyme can bind to any of the species in the reaction mixture, forming a complex with it. After the reaction EG8→EG6, the enzyme can release the G6 from the active site. Alternatively, it can slide along it and catalyze another hydrolysis reaction (i.e. EG6→EG4). Likewise, after the reaction EG6→EG4, the enzyme can release the G4 from the active site. Alternatively, it can slide along it and catalyze another hydrolysis reaction (i.e. EG4→EG2*). Each hydrolysis reaction liberates one G2 from the non-reducing end of the oligomer. The reaction EG4→EG2* liberates an extra G2, the final residual product of which is the reducing-end of an original G8 molecule. This “extra G2” is denoted as G2*, allowing for the possibility that the reducing end of the original G8 were marked.

The equation set was developed in the same manner as that used in Case Study 1, following the same assumptions.

The balance equations for the complexes of the enzyme with the various species involved in the reaction are:

$$\frac{d[\mathrm{EG8}]}{dt} = k_{\mathrm{on8}}[\mathrm{G8}][\mathrm{E}] - k_{\mathrm{off8}}[\mathrm{EG8}] - k_{\mathrm{cat8}}[\mathrm{EG8}] \quad \text{(S56)}$$

$$\frac{d[\mathrm{EG6}]}{dt} = k_{\mathrm{on6}}[\mathrm{G6}][\mathrm{E}] - k_{\mathrm{off6}}[\mathrm{EG6}] - k_{\mathrm{cat6}}[\mathrm{EG6}] + k_{\mathrm{cat8}}[\mathrm{EG8}] \quad \text{(S57)}$$

$$\frac{d[\mathrm{EG4}]}{dt} = k_{\mathrm{on4}}[\mathrm{G4}][\mathrm{E}] - k_{\mathrm{off4}}[\mathrm{EG4}] - k_{\mathrm{cat4}}[\mathrm{EG4}] + k_{\mathrm{cat6}}[\mathrm{EG6}] \quad \text{(S58)}$$

$$\frac{d[\mathrm{EG2^*}]}{dt} = k_{\mathrm{on2}}[\mathrm{G2^*}][\mathrm{E}] - k_{\mathrm{off2}}[\mathrm{EG2^*}] + k_{\mathrm{cat4}}[\mathrm{EG4}] \quad \text{(S59)}$$

$$\frac{d[\mathrm{EG2}]}{dt} = k_{\mathrm{on2}}[\mathrm{G2}][\mathrm{E}] - k_{\mathrm{off2}}[\mathrm{EG2}] \quad \text{(S60)}$$

The total enzyme in the system is represented by:

$$[\mathrm{E}]_{\mathrm{T}} = [\mathrm{E}] + [\mathrm{EG8}] + [\mathrm{EG6}] + [\mathrm{EG4}] + [\mathrm{EG2^*}] + [\mathrm{EG2}] \quad \text{(S61)}$$

From Eqs. (S56) to (S61), the free enzyme in the system is:

$$[\mathrm{E}] = \frac{[\mathrm{E}]_{\mathrm{T}}}{D_8} \quad \text{(S62)}$$

where the denominator D_8 represents all the forms of the enzyme at any particular time during the reaction. It is given by:

$$D_8 = 1 + \left(\frac{[\mathrm{G8}]}{K_{\mathrm{M8}}} + \frac{k_8[\mathrm{G8}]}{k_{\mathrm{off6}} + k_{\mathrm{cat6}}} + \frac{k_8[\mathrm{G8}]P_6}{k_{\mathrm{off4}} + k_{\mathrm{cat4}}} + \frac{k_8[\mathrm{G8}]P_6P_4}{k_{\mathrm{off2}}}\right) + \left(\frac{[\mathrm{G6}]}{K_{\mathrm{M6}}} + \frac{k_6[\mathrm{G6}]}{k_{\mathrm{off4}} + k_{\mathrm{cat4}}} + \frac{k_6[\mathrm{G6}]P_4}{k_{\mathrm{off2}}}\right) + \left(\frac{[\mathrm{G4}]}{K_{\mathrm{M4}}} + \frac{k_4[\mathrm{G4}]}{k_{\mathrm{off2}}}\right) + \frac{[\mathrm{G2^*}]}{K_{\mathrm{M2}}} + \frac{[\mathrm{G2}]}{K_{\mathrm{M2}}} \quad \text{(S63)}$$

Table S3 explains the origins of the various terms in this denominator.

Table S3. Meanings of the terms within the denominator of the rate equations (i.e. within the expression for D_8)

Term of Eq. (S63)	Form of the enzyme represented by the term
1	free enzyme
$\frac{[\mathrm{G8}]}{K_{\mathrm{M8}}}$	complex EG8 originating from direct binding of G8 to free enzyme
$\frac{k_8[\mathrm{G8}]}{k_{\mathrm{off6}}+k_{\mathrm{cat6}}}$	complex EG6 originating from the sequence E+G8→EG8→EG6
$\frac{k_8[\mathrm{G8}]P_6}{k_{\mathrm{off4}}+k_{\mathrm{cat4}}}$	complex EG4 originating from the sequence E+G8→EG8→EG6→EG4
$\frac{k_8[\mathrm{G8}]P_6P_4}{k_{\mathrm{off2}}}$	complex EG2* originating from the sequence E+G8→EG8→EG6→EG4→EG2*
$\frac{[\mathrm{G6}]}{K_{\mathrm{M6}}}$	complex EG6 originating from direct binding of G6 to free enzyme
$\frac{k_6[\mathrm{G6}]}{k_{\mathrm{off4}}+k_{\mathrm{cat4}}}$	complex EG4 originating from the sequence E+G6→EG6→EG4
$\frac{k_6[\mathrm{G6}]P_4}{k_{\mathrm{off2}}}$	complex EG2* originating from the sequence E+G6→EG6→EG4→EG2*
$\frac{[\mathrm{G4}]}{K_{\mathrm{M4}}}$	complex EG4 originating from direct binding of G4 to free enzyme
$\frac{k_4[\mathrm{G4}]}{k_{\mathrm{off2}}}$	complex EG2* originating from the sequence E+G4→EG4→EG2*
$\frac{[\mathrm{G2^*}]}{K_{\mathrm{M2}}}$	complex EG2* originating from direct binding of G2* to free enzyme
$\frac{[\mathrm{G2}]}{K_{\mathrm{M2}}}$	complex EG2 originating from direct binding of G2 to free enzyme

The balance equations for the original substrate, all residual products and the side product are:

$$\frac{d[\mathrm{G8}]}{dt} = -k_8[\mathrm{G8}]\frac{[\mathrm{E}]_\mathrm{T}}{D_8} \tag{S64}$$

$$\frac{d[\mathrm{G6}]}{dt} = k_8[\mathrm{G8}](1-P_6)\frac{[\mathrm{E}]_\mathrm{T}}{D_8} - k_6[\mathrm{G6}]\frac{[\mathrm{E}]_\mathrm{T}}{D_8} \tag{S65}$$

$$\frac{d[\mathrm{G4}]}{dt} = k_8[\mathrm{G8}]P_6(1-P_4)\frac{[\mathrm{E}]_\mathrm{T}}{D_8} + k_6[\mathrm{G6}](1-P_4)\frac{[\mathrm{E}]_\mathrm{T}}{D_8} - k_4[\mathrm{G4}]\frac{[\mathrm{E}]_\mathrm{T}}{D_8} \tag{S66}$$

$$\frac{d[\mathrm{G2^*}]}{dt} = k_8[\mathrm{G8}]P_6P_4\frac{[\mathrm{E}]_\mathrm{T}}{D_8} + k_6[\mathrm{G6}]P_4\frac{[\mathrm{E}]_\mathrm{T}}{D_8} + k_4[\mathrm{G4}]\frac{[\mathrm{E}]_\mathrm{T}}{D_8} \tag{S67}$$

$$\frac{d[\mathrm{G2}]}{dt} = k_8[\mathrm{G8}](1+P_6+P_6P_4)\frac{[\mathrm{E}]_T}{D_8} + k_6[\mathrm{G6}](1+P_4)\frac{[\mathrm{E}]_T}{D_8} + k_4[\mathrm{G4}]\frac{[\mathrm{E}]_T}{D_8} \tag{S68}$$

The fractional reaction extent (F_8) for the scheme in Fig. S4 is given by the difference between the original concentration of attackable bonds, and the

concentration of remaining attackable bonds, both of these terms being normalized with respect to the original concentration of attackable bonds:

$$F_8 = \frac{3[\mathrm{G8}]_\mathrm{o}}{3[\mathrm{G8}]_\mathrm{o}} - \left(\frac{3[\mathrm{G8}] + 2[\mathrm{G6}] + [\mathrm{G4}]}{3[\mathrm{G8}]_\mathrm{o}}\right) \tag{S69}$$

The derivative of Eq. (S69) with respect to time is:

$$\frac{dF_8}{dt} = \left(-\frac{1}{3[\mathrm{G8}]_\mathrm{o}}\right)\left(3\frac{d[\mathrm{G8}]}{dt} + 2\frac{d[\mathrm{G6}]}{dt} + \frac{d[\mathrm{G4}]}{dt}\right) \tag{S70}$$

Replacing the terms for $d[\mathrm{G8}]/dt$, $d[\mathrm{G6}]/dt$ and $d[\mathrm{G4}]/dt$ in Eq. (S70) with Eqs. (S64), (S65) and (S66), respectively, gives:

$$\frac{dF_8}{dt} = \left(\frac{1}{3[\mathrm{G8}]_\mathrm{o}} \times \frac{[\mathrm{E}]_\mathrm{T}}{D_8}\right)(k_8[\mathrm{G8}](1 + P_6 + P_6P_4) + k_6[\mathrm{G6}](1 + P_4) + k_4[\mathrm{G4}]) \tag{S71}$$

The set of differential equations with fractional reaction extent as the independent variable is:

$$\frac{d[\mathrm{G8}]}{dF} = 3[\mathrm{G8}]_\mathrm{o}\frac{-k_8[\mathrm{G8}]}{k_8[\mathrm{G8}](1 + P_6 + P_6P_4) + k_6[\mathrm{G6}](1 + P_4) + k_4[\mathrm{G4}]} \tag{S72}$$

$$\frac{d[\mathrm{G6}]}{dF} = 3[\mathrm{G8}]_\mathrm{o}\frac{k_8[\mathrm{G8}](1 - P_6) - k_6[\mathrm{G6}]}{k_8[\mathrm{G8}](1 + P_6 + P_6P_4) + k_6[\mathrm{G6}](1 + P_4) + k_4[\mathrm{G4}]} \tag{S73}$$

$$\frac{d[\mathrm{G4}]}{dF} = 3[\mathrm{G8}]_\mathrm{o}\frac{k_8[\mathrm{G8}]P_6(1-P_4) + k_6[\mathrm{G6}](1 - P_4) - k_4[\mathrm{G4}]}{k_8[\mathrm{G8}](1 + P_6 + P_6P_4) + k_6[\mathrm{G6}](1 + P_4) + k_4[\mathrm{G4}]} \tag{S74}$$

$$\frac{d[\mathrm{G2}^*]}{dF} = 3[\mathrm{G8}]_\mathrm{o}\frac{k_8[G8]P_6P_4 + k_6[\mathrm{G6}]P_4 + k_4[\mathrm{G4}]}{k_8[\mathrm{G8}](1 + P_6 + P_6P_4) + k_6[\mathrm{G6}](1 + P_4) + k_4[\mathrm{G4}]} \tag{S75}$$

$$\frac{d[\mathrm{G2}]}{dF} = 3[\mathrm{G8}]_\mathrm{o} \tag{S76}$$

Again, it is convenient to express the concentrations of all species relative to the concentration of the initial substrate as mole ratios (i.e. as $Y = [Y]/[\mathrm{G8}]_\mathrm{o}$). Additionally, dividing both the denominator and the numerator on the right-hand sides

of Eqs. (S72) to (S76) by k_8 leads to *R*-parameters that represent relative specificity constants. These two transformations lead to the following set of equations:

$$\frac{dG8}{dF} = 3\frac{-R_8G8}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad \text{(S77)}$$

$$\frac{dG6}{dF} = 3\frac{R_8G8(1-P_6)-R_6G6}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad \text{(S78)}$$

$$\frac{dG4}{dF} = 3\frac{R_8G8(1-P_4)P_6+R_6G6(1-P_4)-R_4G4}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad \text{(S79)}$$

$$\frac{dG2^*}{dF} = 3\frac{R_8G8P_6P_4+R_6G6P_4+R_4G4}{R_8G8(1+P_6+P_6P_4)+R_6G6(1+P_4)+R_4G4} \quad \text{(S80)}$$

$$\frac{dG2}{dF} = 3 \quad \text{(S81)}$$

where R_8, by definition, is equal to 1.

This equation set describes the situation when the reducing end of the initial substrate is marked. When this is not the case, the equation for G2 will be the sum of the equations for G2 and G2*.

5.8.4 Case Study 3

The third case study is done for the hydrolysis of maltoundecaose (G11), maltononaose (G9) and maltoheptaose (G7), using data from French and Youngquist [19] obtained in three different experiments. Three different models were developed, one for each experiment, but since some parameters are common to all three models, the three models were adjusted simultaneously to their corresponding data sets.

Figure S5 shows the scheme for the hydrolysis of maltoundecaose (G11). Again, the enzyme hydrolyzes G11 to G3, removing G2 units in each step. In this scheme, there can be as many as three successive processive actions, with the complexes EG9, EG7 and EG5 being able to follow either processive or non-processive routes. It is assumed that enzyme can also bind to G3 and G2, giving rise to unproductive complexes. The schemes for the hydrolysis of G9 and G7 are subsets of this scheme.

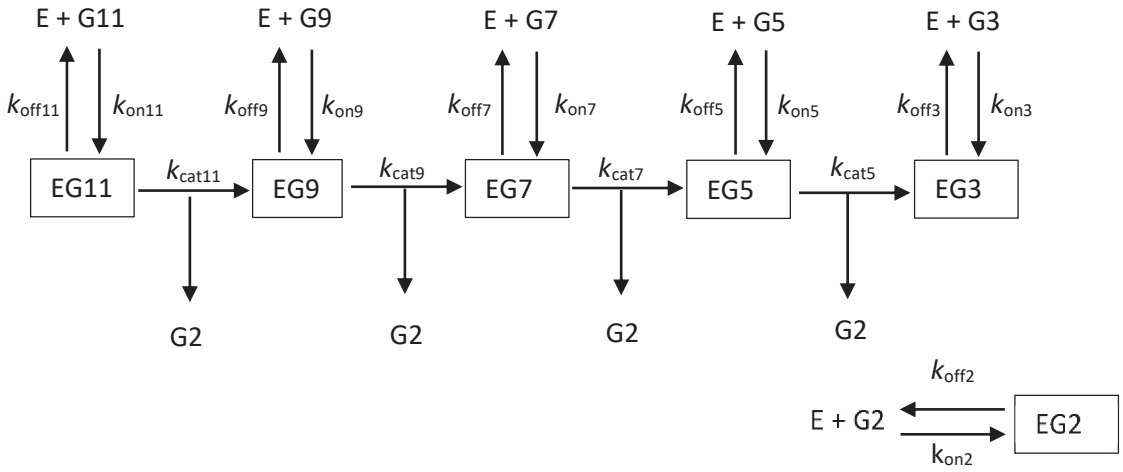


Fig. S5. Reaction scheme for the hydrolysis of G11 by β-amylase. For each i-mer, $k_{off(i)}$ is the fundamental rate constant for dissociation of the complex; $k_{on(i)}$ is the fundamental rate constant for formation of the complex from free enzyme and the i-mer and $k_{cat(i)}$ is the catalytic rate constant.

5.8.4.1 Reaction with G7 as initial substrate

Starting from the intermediate G7, the model is the same as that used in Case Study 1, where the fractional reaction extent is given by Eq. (S45) and the differential equations with F_7 as the independent variable are given by Eqs. (S52) to (S55).

Since the three experiments are based on the same reaction, all specificity constants should be normalized by dividing by a specificity constant that is common to

all experiments. The specificity constant k_7 was used to normalize all others in this case study (such that, in this case study, by definition, $R_7 = 1$).

5.8.4.2 Reaction with G9 as initial substrate

In this case, the reaction scheme corresponds to the part of Fig. S5 involving G9 and smaller oligosaccharides. The fractional reaction extent in this experiment is given by:

$$F_9 = \frac{[\mathrm{G2}]}{3[\mathrm{G9}]_\mathrm{o}} \tag{S82}$$

The set of differential equations with fractional reaction extent as the independent variable is:

$$\frac{d[\mathrm{G9}]}{dF} = 3[\mathrm{G9}]_\mathrm{o} \frac{-k_9[\mathrm{G9}]}{k_9[\mathrm{G9}](1+P_7+P_7P_5) + k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S83}$$

$$\frac{d[\mathrm{G7}]}{dF} = 3[\mathrm{G9}]_\mathrm{o} \frac{k_9[\mathrm{G9}](1-P_7) - k_7[\mathrm{G7}]}{k_9[\mathrm{G9}](1+P_7+P_7P_5) + k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S84}$$

$$\frac{d[\mathrm{G5}]}{dF} = 3[\mathrm{G9}]_\mathrm{o} \frac{k_9[\mathrm{G9}]P_7(1-P_5) + k_7[\mathrm{G7}](1-P_5) - k_5[\mathrm{G5}]}{k_9[\mathrm{G9}](1+P_7+P_7P_5) + k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S85}$$

$$\frac{d[\mathrm{G3}]}{dF} = 3[\mathrm{G9}]_\mathrm{o} \frac{k_9[\mathrm{G9}]P_7P_5 + k_7[\mathrm{G7}]P_5 + k_5[\mathrm{G5}]}{k_9[\mathrm{G9}](1+P_7+P_7P_5) + k_7[\mathrm{G7}](1+P_5) + k_5[\mathrm{G5}]} \tag{S86}$$

Normalizing all concentrations based on $[\mathrm{G7}]_\mathrm{o}$ and all specificity constants based on k_7 gives:

$$\frac{dG9}{dF} = 3\frac{-R_9G9}{R_9G9(1+P_7+P_7P_5) + R_7G7(1+P_5) + R_5G5} \tag{S87}$$

$$\frac{dG7}{dF} = 3\frac{R_9G9(1-P_7) - R_7G7}{R_9G9(1+P_7+P_7P_5) + R_7G7(1+P_5) + R_5G5} \tag{S88}$$

$$\frac{dG5}{dF} = 3\frac{R_9G9P_7(1-P_5) + R_7G7(1-P_5) - R_5G5}{R_9G9(1+P_7+P_7P_5) + R_7G7(1+P_5) + R_5G5} \tag{S89}$$

$$\frac{dG3}{dF} = 3\frac{R_9G9P_7P_5 + R_7G7P_5 + R_5G5}{R_9G9(1+P_7+P_7P_5) + R_7G7(1+P_5) + R_5G5} \tag{S90}$$

5.8.4.3 Reaction with G11 as initial substrate

The fractional reaction extent for the complete scheme in Figure S5 is:

$$F_{11} = \frac{[\mathrm{G2}]}{4[\mathrm{G11}]_\mathrm{o}} \tag{S91}$$

The set of differential equations with fractional reaction extent as independent variable with G11 as the initial substrate is:

$$\frac{d[\mathrm{G11}]}{dF} = 4[\mathrm{G11}]_\mathrm{o} \frac{-k_{11}[\mathrm{G11}]}{k_{11}[\mathrm{G11}](1+P_9+P_9P_7+P_9P_7P_5)+k_9[\mathrm{G9}](1+P_7+P_7P_5)+k_7[\mathrm{G7}](1+P_5)+k_5[\mathrm{G5}]} \tag{S92}$$

$$\frac{d[\mathrm{G9}]}{dF} = 4[\mathrm{G11}]_\mathrm{o} \frac{k_{11}[\mathrm{G11}](1-P_9)-k_9[\mathrm{G9}]}{k_{11}[\mathrm{G11}](1+P_9+P_9P_7+P_9P_7P_5)+k_9[\mathrm{G9}](1+P_7+P_7P_5)+k_7[\mathrm{G7}](1+P_5)+k_5[\mathrm{G5}]} \tag{S93}$$

$$\frac{d[\mathrm{G7}]}{dF} = 4[\mathrm{G11}]_\mathrm{o} \frac{k_{11}[\mathrm{G11}]P_9(1-P_7)+k_9[\mathrm{G9}](1-P_7)-k_7[\mathrm{G7}]}{k_{11}[\mathrm{G11}](1+P_9+P_9P_7+P_9P_7P_5)+k_9[\mathrm{G9}](1+P_7+P_7P_5)+k_7[\mathrm{G7}](1+P_5)+k_5[\mathrm{G5}]} \tag{S94}$$

$$\frac{d[\mathrm{G5}]}{dF} = 4[\mathrm{G11}]_\mathrm{o} \frac{k_{11}[\mathrm{G11}]P_9P_7(1-P_5)+k_9[\mathrm{G9}]P_7(1-P_5)+k_7[\mathrm{G7}](1-P_5)-k_5[\mathrm{G5}]}{k_{11}[\mathrm{G11}](1+P_9+P_9P_7+P_9P_7P_5)+k_9[\mathrm{G9}](1+P_7+P_7P_5)+k_7[\mathrm{G7}](1+P_5)+k_5[\mathrm{G5}]} \tag{S95}$$

$$\frac{d[\mathrm{G3}]}{dF} = 4[\mathrm{G11}]_\mathrm{o} \frac{k_{11}[\mathrm{G11}]P_9P_7P_5+k_9[\mathrm{G9}]P_7P_5+k_7[\mathrm{G7}]P_5+k_5[\mathrm{G5}]}{k_{11}[\mathrm{G11}](1+P_9+P_9P_7+P_9P_7P_5)+k_9[\mathrm{G9}](1+P_7+P_7P_5)+k_7[\mathrm{G7}](1+P_5)+k_5[\mathrm{G5}]} \tag{S96}$$

Normalizing all concentrations based on $[\mathrm{G11}]_\mathrm{o}$ and all specificity constants based on k_7 gives:

$$\frac{dG11}{dF} = 4\frac{-R_{11}G11}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \tag{S97}$$

$$\frac{dG9}{dF} = 4\frac{R_{11}G11(1-P_9)-R_9G9}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \tag{S98}$$

$$\frac{dG7}{dF} = 4\frac{R_{11}G11P_9(1-P_7)+R_9G9(1-P_7)-R_7G7}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \tag{S99}$$

$$\frac{dG5}{dF} = 4\frac{R_{11}G11P_9P_7(1-P_5)+R_9G9P_7(1-P_5)+R_7G7(1-P_5)-R_5G5}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \tag{S100}$$

$$\frac{dG3}{dF} = 4\frac{R_{11}G11P_9P_7P_5+R_9G9P_7P_5+R_7G7P_5+R_5G5}{R_{11}G11(1+P_9+P_9P_7+P_9P_7P_5)+R_9G9(1+P_7+P_7P_5)+R_7G7(1+P_5)+R_5G5} \tag{S101}$$

6 CONSIDERAÇÕES FINAIS

Este trabalho traz o primeiro modelo cinético de uma enzima processiva que considera a variação das probabilidades de ataque processivo em função dos tamanhos das cadeias dos vários oligossacarídeos no esquema de reação. Conforme mostrado na discussão do artigo (seção 5.4), o método Fingerprinting com processividade é uma abordagem superior aos métodos já apresentados na literatura, no que diz respeito à determinação das constantes de especificidade e probabilidades de ação processiva.

O modelo aqui apresentado não incorre nos erros conceituais e simplificações de Nakatani (1997) e Ishikawa et al. (2007). A abordagem usada para descrever processividade no presente trabalho é superior em relação à abordagem previamente desenvolvida por Pereira, Krieger e Mitchell (2016). Pode ser estendido mais facilmente para descrever esquemas de maior complexidade que envolvem reações processivas. Além disto, fornece as probabilidades de ataque processivo e as constantes de especificidade relativas diretamente, enquanto, na abordagem de Pereira, Krieger e Mitchell (2016), os parâmetros ajustados são combinações destes parâmetros.

Esta extensão do método Fingerprinting apresenta vários pontos em comum com trabalho de Cruys-Bagger et al. (2013), como o termo $k_{cat}/(k_{cat} + k_{off})$ representando a probabilidade de ataque processivo. Também, ambos os trabalhos consideram que múltiplas rodadas de ataque processivo podem ocorrer. Entretanto, a abordagem tomada neste trabalho é mais abrangente que a de Cruys-Bagger et al. (2013), pois permite que quando a enzima se dissocia de um oligossacarídeo, ela pode se ligar a outra molécula e isso faz com que o modelo possa descrever todo o curso da reação.

6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Segundo Fazekas et al. (2012), durante o processo de hidrólise catalisado por β-amilase de batata doce, podem ocorrer também reações de transglicosilação, que afetam o perfil de reação tanto quanto a processividade. Normalmente, após a hidrólise enzimática de um polímero de carboidratos, um dos produtos deixa o sítio ativo e o outro é transferido para uma molécula de água. Porém, em altas

concentrações de substrato, esta transferência pode ser feita para outra molécula de substrato (GENTA et al., 2007). Há, então, uma necessidade de expandir o método Fingerprinting que descreve processividade, desenvolvido no presente trabalho, para considerar esta reação paralela.

Como perspectiva futura, as constantes de especificidade relativas e as probabilidades de ataque processivo poderiam ser utilizadas em modelos matemáticos como ferrramentas no projeto e otimização de biorreatores enzimáticos. Para isso, o método Fingerprinting teria que ser complementado com outras estratégias para a determinação de constantes de saturação e constantes de inibição.

REFERÊNCIAS

BAILEY, J. M.; FRENCH, D. The significance of multiple reactions in enzyme-polymer systems. **Journal of Biological Chemistry**, n. 226, p. 1–14, 1957.

BASTOS, V. D. Etanol, alcoolquímica e biorrefinarias. **BNDES Setorial**, n. 25, p. 5–38, 2007.

BONNIN, E. et al. Study of the mode of action of endopolygalacturonase from Fusarium moniliforme. **Biochimica et biophysica acta**, v. 1526, n. 3, p. 301–309, 2001.

CORNISH-BOWDEN, A. Enzyme specificity in reactions of more than one co-substrate. **Biochemical Journal**, v. 291, n. 1, p. 323.2-324, 1 abr. 1993.

CORNISH-BOWDEN, A. **Fundamentals of Enzyme Kinectics**. 4. ed. Berlim: Wiley-Blackwell, 2012

COUMANS, R. G. E. et al. Processive enzyme mimic: Kinetics and thermodynamics of the threading and sliding process. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 52, p. 19647–51, 26 dez. 2006.

CRUYS-BAGGER, N. et al. A steady-state theory for processive cellulases. **FEBS Journal**, v. 280, n. 16, p. 3952–3961, ago. 2013.

DAVIES, G.; HENRISSAT, B. Structures and mechanisms of glycosyl hydrolases. **Structure**, v. 3, n. 9, p. 853–859, 1995.

DAVIES, G. J. et al. Oligosaccharide specificity of a family 7 endoglucanase: insertion of potential sugar-binding subsites. **Journal of biotechnology**, v. 57, n. 1–3, p. 91–100, 16 set. 1997.

DE JONG, E.; JUNGMEIER, G. Biorefinery Concepts in Comparison to Petrochemical Refineries. In: PANDEY, A. **Industrial Biorefineries and White Biotechnology**. Amsterdan: Elsevier, 2015. p. 3–33.

FAZEKAS, E.; KANDRA, L.; GYÉMÁNT, G. Model for b -1 , 6 -N- acetylglucosamine oligomer hydrolysis catalysed by DispersinB , a biofilm degrading enzyme. **CARBOHYDRATE RESEARCH**, v. 363, p. 7–13, 2012.

FRENCH, D.; YOUNGQUIST, R. W. Die Wirkungsweise von β-Amylase auf Stärkeoligosaccharide. **Starch - Stärke**, v. 15, n. 12, p. 425–431, 1963.

GENTA, F. A. et al. The interplay of processivity, substrate inhibition and a secondary substrate binding site of an insect exo-β-1,3-glucanase. **Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics**, v. 1774, n. 9, p. 1079–1091, 2007.

GRIGGS, A. J.; STICKEL, J. J.; LISCHESKE, J. J. A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics II: Cooperative enzyme action, solution kinetics, and product inhibition. **Biotechnology**

and Bioengineering, v. 109, n. 3, p. 676–685, mar. 2012.

HAMRE, A. G. et al. Enzyme processivity changes with the extent of recalcitrant polysaccharide degradation. **FEBS Letters**, v. 588, n. 24, p. 4620–4624, 20 dez. 2014.

HORN, S. J. et al. Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 48, p. 18089–18094, 2006.

HORN, S. J. et al. Chapter five – Measuring Processivity. In: GILBERT, H.J. **Methods in Enzymology**. San Diego: Academic Press, 2012. v. 510. p. 69–95.

HRMOVA, M. et al. Substrate binding and catalytic mechanism of a barley beta-D-Glucosidase/(1,4)-beta-D-glucan exohydrolase. **The Journal of biological chemistry**, v. 273, n. 18, p. 11134–43, 1 maio 1998.

ISHIKAWA, K. et al. Kinetic and structural analysis of enzyme sliding on a substrate: multiple attack in beta-amylase. **Biochemistry**, v. 46, n. 3, p. 792–798, 23 jan. 2007.

JOHNSON, K. A. Role of induced fit in enzyme specificity: a molecular forward/reverse switch. **The Journal of biological chemistry**, v. 283, n. 39, p. 26297–301, 26 set. 2008.

KAMON, M. et al. Characterization and gene cloning of a maltotriose-forming exo-amylase from Kitasatospora sp. MK-1785. **Applied Microbiology and Biotechnology**, v. 99, n. 11, p. 4743–4753, 27 jun. 2015.

KLUSKENS, L. D. et al. Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium Thermotoga maritima. **FEBS Journal**, v. 272, n. 21, p. 5464–5473, 2005.

KNOTT, B. C. et al. Carbohydrate–Protein Interactions That Drive Processive Polysaccharide Translocation in Enzymes Revealed from a Computational Study of Cellobiohydrolase Processivity. **Journal of the American Chemical Society**, v. 136, n. 24, p. 8810–8819, 18 jun. 2014.

KUMAR, D.; MURTHY, G. S. Stochastic molecular model of enzymatic hydrolysis of cellulose for ethanol production. **Biotechnology for Biofuels**, v. 6, n. 1, p. 63, 2 maio 2013.

LUCIUS, A. L. et al. General methods for analysis of sequential "n-step" kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. **Biophysical journal**, v. 85, n. 4, p. 2224–39, out. 2003.

MAZUR, A. K.; NAKATANI, H. Multiple Attack Mechanism in the Porcine Pancreatic α-Amylase Hydrolysis of Amylose and Amylopectin. **Archives of Biochemistry and Biophysics**, v. 306, n. 1, p. 29–38, out. 1993.

MCCLURE, W. R.; CHOW, Y. [11] The kinetics and processivity of nucleic acid polymerases. In: PURICH, D.L. **Methods in Enzymology**. Cambridge: Academic Press, 1980. p. 277–297.

MEDVE, J. et al. Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II fromTrichoderma reesei: Adsorption, sugar production pattern, and synergism of the enzymes. **Biotechnology and Bioengineering**, v. 59, n. 5, p. 621–634, 5 set. 1998.

MERTENS, J. A. Kinetic properties of two rhizopus Exo-polygalacturonase enzymes hydrolyzing galacturonic acid oligomers using isothermal titration calorimetry. **Applied Biochemistry and Biotechnology**, v. 170, n. 8, p. 2009–2020, 2013.

MERTENS, J. A.; BRAKER, J. D.; JORDAN, D. B. Catalytic Properties of Two Rhizopus oryzae 99-880 Glucoamylase Enzymes Without Starch Binding Domains Expressed in Pichia pastoris. **Applied Biochemistry and Biotechnology**, v. 162, n. 8, p. 2197–2213, 13 dez. 2010.

MICHEL, G. et al. The structural bases of the processive degradation of iota-carrageenan, a main cell wall polysaccharide of red algae. **Journal of molecular biology**, v. 334, n. 3, p. 421–33, 28 nov. 2003.

MITCHELL, D. A. et al. An analytical method for determining relative specificities for sequential reactions catalyzed by the same enzyme: Application to the hydrolysis of triacylglycerols by lipases. **Journal of Biotechnology**, v. 133, n. 3, p. 343–350, 1 fev. 2008a.

MITCHELL, D. A. et al. Determination of the quantitative stereoselectivity fingerprint of lipases during hydrolysis of a prochiral triacylglycerol. **Journal of Biotechnology**, v. 135, n. 2, p. 168–173, 1 jun. 2008b.

MITCHELL, D. A.; CARRIÈRE, F.; KRIEGER, N. An analytical method for determining relative specificities for sequential reactions catalyzed by the same enzyme: General formulation. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1784, n. 4, p. 705–715, abr. 2008.

MOHANRAM, S. et al. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. **Sustainable Chemical Processes**, v. 1, n. 1, p. 15, 2013.

NAKATANI, H. Monte Carlo simulation of multiple attack mechanism of beta-amylase-catalyzed reaction. **Biopolymers**, v. 42, n. 7, p. 831–6, dez. 1997.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger.** 6. ed. Porto Alegre: Artmed, 2014.

PAOLUCCI-JEANJEAN, D. et al. Kinetics of cassava starch hydrolysis with Termamyl(R) enzyme. **Biotechnology and Bioengineering**, v. 68, n. 1, p. 71–7, 5 abr. 2000.

PEREIRA, A. B.; KRIEGER, N.; MITCHELL, D. A. Fingerprinting of oligosaccharide-hydrolyzing enzymes that catalyze branched reaction schemes. **Biochemical Engineering Journal**, v. 113, p. 93–101, 2016.

PRAESTGAARD, E. et al. A kinetic model for the burst phase of processive cellulases. **FEBS Journal**, v. 278, n. 9, p. 1547–1560, maio 2011.

RAGAUSKAS, A. J. et al. The Path Forward for Biofuels and Biomaterials. **Science**, v. 311, n. 5760, 2006.

ROBYT, J. F.; FRENCH, D. Multiple attach hypothesis of alpha-amylase action: action of porcine pancreatic, human salivary, and Aspergillus oryzae alpha-amylases. **Archives of biochemistry and biophysics**, v. 122, n. 1, p. 8–16, out. 1967.

SAKAGUCHI, M. et al. Glucoamylase of Caulobacter crescentus CB15: cloning and expression in Escherichia coli and functional identification. **AMB Express**, v. 4, n. 1, p. 5, 27 jan. 2014.

SONG et al. Kinetics of enzymatic hydrolysis revealed by video rate AFM single molecule analysis. In: IEEE International Conference on Nanotechnology (IEEE-NANO), 15th, 2015, Rome. **Proceedings**…New York: IEEE, jul. 2015. p. 1288-1291.

SØRBOTTEN, A. et al. Degradation of chitosans with chitinase B from Serratia marcescens. **FEBS Journal**, v. 272, n. 2, p. 538–549, jan. 2005.

STREVENS, M. Probability out of determinism. In: BEISBART, C.; HARTMANN, S. **Probabilities in Physics.** Oxford: Oxford University Press, 2011.

TAYLOR, C. B. et al. Binding Site Dynamics and Aromatic–Carbohydrate Interactions in Processive and Non-Processive Family 7 Glycoside Hydrolases. **The Journal of Physical Chemistry B**, v. 117, n. 17, p. 4924–4933, 2 maio 2013.

ANEXO A

License Number	4103281320061
License date	May 06, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	FEBS Journal
Licensed Content Title	A steady- state theory for processive cellulases
Licensed Content Author	Nicolaj Cruys- Bagger,Jens Elmerdahl,Eigil Praestgaard,Kim Borch,Peter Westh
Licensed Content Date	Jul 12, 2013
Licensed Content Pages	10
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Scheme 1
Will you be translating?	No
Title of your thesis / dissertation	SIMULAÇÃO DA HIDRÓLISE DE OLIGÔMEROS DE GLUCOSE POR UMA BETA-AMILASE PROCESSIVA USANDO O MÉTODO “FINGERPRINTING”
Expected completion date	May 2017
Expected size (number of pages)	90
Requestor Location	Gerson Mello Rua Olivo Carnasciali, 311 Curitiba, Parana 82970-190 Brazil Attn: Gerson Mello
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Gerson Mello Rua Olivo Carnasciali, 311 Curitiba, Brazil 82970-190 Attn: Gerson Mello
Total	0.00 USD

License Number	4103280924211
License date	May 06, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Biochemical Engineering Journal
Licensed Content Title	Fingerprinting of oligosaccharide-hydrolyzing enzymes that catalyze branched reaction schemes
Licensed Content Author	Aline Bescrovaine Pereira,Nadia Krieger,David Alexander Mitchell
Licensed Content Date	15 September 2016
Licensed Content Volume	113
Licensed Content Issue	n/a
Licensed Content Pages	9
Start Page	93
End Page	101
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Original figure numbers	Figure 3
Title of your thesis/dissertation	SIMULAÇÃO DA HIDRÓLISE DE OLIGÔMEROS DE GLUCOSE POR UMA BETA-AMILASE PROCESSIVA USANDO O MÉTODO “FINGERPRINTING”
Expected completion date	May 2017
Estimated size (number of pages)	90
Elsevier VAT number	GB 494 6272 12
Requestor Location	Gerson Mello Rua Olivo Carnasciali, 311 Curitiba, Parana 82970-190 Brazil Attn: Gerson Mello
Total	0.00 USD

This figure was reprinted from Biochemical Engineering Journal, Vol 113. A.B. Pereira, N. Krieger and D.A. Mitchell. Fingerprinting of oligosaccharide-hydrolyzing enzymes that catalyze branched reaction schemes, 93-101, Copyright (2016), with permission from Elsevier.