UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

CECÍLIA MIDORI IKEGAMI

**Caracterização funcional dos resíduos centrais
da rede estrutural da β -glicosidase
Sf β gli de *Spodoptera frugiperda***

Versão original da Tese

São Paulo

Data do Depósito na SPG:

10/04/2013

CECÍLIA MIDORI IKEGAMI

**Caracterização funcional dos resíduos centrais da
rede estrutural da β -glicosidase Sf β gli de
*Spodoptera frugiperda***

*Tese apresentada ao Instituto de Química da
Universidade de São Paulo para obtenção do
Título de Doutor em Ciências (Bioquímica)*

Orientador (a): Prof (a). Dr (a). Sandro Roberto Marana

São Paulo

2013

AGRADECIMENTOS

Ao Prof. Sandro Roberto Marana por prontamente aceitar me orientar e pela oportunidade de trabalhar em seu laboratório. Agradeço pela orientação, suporte, compreensão e paciência.

Aos membros e agregados do laboratório de enzimologia pela amizade, apoio, discussões e troca de idéias: Daniela Beton, Fábio Tamaki (Pererê), José Galeote, Larissa Gonçalves, Lúcio Mendonça, Maíra Frutuoso, Raphael Sayegh (Ceará), Roberto Rozenberg, Valquiria Souza.

Aos meus amigos e colegas do IQ pela amizade, convívio e por tornarem esses anos de pós mais felizes e agradáveis: Mariana Duarte, Mariana Manzini, Gabriela Avelar, Micheli Mainardi, Arquimedes Cheffer, Florêncio Freitas, Raphael Teixeira, Alexsandra Scalfio, Daniela Gonzalez, Armando Henrique, Maíra Nagai, Tiago Almeida, Eliezer Stefanello, Juliana Galvão, Juliana Osaki, Kátia Neves, Sophia Oliveira.

Às minhas amigas de república obrigada pela ótima amizade, super companheiras, que me deram a maior força: Kátia Kreutz, Márcia Pincerati, Tina Smuc, Ana García, Fabiana Missau

Aos amigos que conheci e aos que o destino afortunadamente nos reaproximou: Celina Takemura, Márcia Pincerati, Rosemeri Morokawa, Luana da Costa, Ana Simões, Luiz Bonani, Beatriz Stranski, Andresa Guimarães.

À minha família que embora distante sempre me ajudou, me incentivou e pelo amor incondicional. Em especial ao meu irmão César Augusto Ikegami que me mostrou que o importante nesta vida é procurar aquilo que te realiza e te faz feliz. Aos meus pais Mário e Midori Ikegami que sempre me incentivaram a alçar vôos mais altos e superar os desafios e dificuldades que a vida nos impõe. À minha cunhada/irmã Lisani Falavigna pelo apoio, força, carinho e amizade singela. E por último, mas tão importante quanto, ao meu marido Caetano Carezzato pelo companheirismo, lealdade, amizade, por sempre estar presente nos momentos em que mais preciso, por sua dedicação e amor.

À Coordenação de Aperfeiçoamento de Pessoal Nível Superior (CAPES), à Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP)

E à todas pessoas e amigos que não foram aqui citados mas que fizeram parte dessa caminhada.

RESUMO

Ikegami, C.M. **Caracterização funcional dos resíduos centrais da rede estrutural da β -glicosidase Sf β gli de *Spodoptera frugiperda***. 2013. 103 p. Tese (Doutorado) - Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

Na última década, a análise da estrutura proteica baseada em teoria de redes/grafos tem emergido. A abstração da estrutura tridimensional proteica em forma de uma rede, leva em consideração os resíduos de aminoácidos e suas interações através do espaço, e apresenta um conjunto de conexões e propriedades mais complexas do que aquelas visualizadas apenas com a estrutura covalente. A análise da estrutura proteica identificou que as proteínas pertencem às redes de classes de “mundo pequeno” (small-world) e “sem escala” (scale-free), o que significa que seus resíduos de aminoácidos são altamente agregados e que existem poucas conexões entre 2 resíduos quaisquer da proteína. A identificação dos resíduos com alto grau de conexão, chamados centrais (“resíduos hubs”), é feita pela determinação do caminho mais curto que conecta um dado resíduo aos demais compreendidos nesta rede. A remoção destes resíduos centrais (hubs) afeta a integridade da rede de forma mais contundente diferentemente da remoção de resíduos que não são centrais. Até o momento estes “resíduos hubs” ainda não foram experimentalmente correlacionados com as propriedades enzimáticas de proteínas. Para tal finalidade, a estrutura terciária de uma β -glicosidase de *Spodoptera frugiperda* (Sf β gli) foi analisada como uma rede. Após calcular-se os caminhos médios entre todos os pares de aminoácidos da β -glicosidase, encontrou-se 11 resíduos centrais (“resíduos hubs”). Alinhamento de sequências e comparações estruturais indicaram alta conservação destes “resíduos hubs”. Nosso objetivo foi produzir esta β -glicosidase mutando-se a maioria dos “resíduos hubs” e 3 aminoácidos não centrais (“não hubs”), expressar estes mutantes em *E. coli*, determinar suas propriedades enzimáticas como atividade catalítica e preferência pelo substrato e verificar a estabilidade destes mutantes em experimentos de inativação térmica. Os resultados obtidos sugerem que mutações nos “resíduos hubs” não afetam as propriedades catalíticas, contudo as enzimas com mutações nos “resíduos hubs” apresentaram uma menor estabilidade térmica. Estes resultados sugeriram que os “resíduos hubs” são relevantes na difusão da energia cinética (vibração) introduzida na estrutura desta β -glicosidase pelo seu aquecimento.

Palavras-chave: β -glicosidase, rede, aminoácidos centrais, enzima mutante, atividade enzimática, inativação térmica

ABSTRACT

Ikegami, C.M. **Functional characterization of the central residues of the structural network of β -glucosidase Sf β gly from *Spodoptera frugiperda*. 2013.** 103 p. PhD Thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

In recent years, graph-theoretic approaches have established that protein structures can be modeled as complex networks of interacting residues. Proteins structures can be represented as small-world and scale-free networks that are usually highly clustered with few links connecting any pair of nodes. The identification of nodes with high connection degrees, called hubs, is made by determining the shortest path linking one amino acid to the further nodes comprising the network. Targeted removal of the hubs has greater affect on the integrity of the network structure in contrast to a random removal of amino acid residues comprising the network. Nevertheless these hubs had not previously been correlated with enzymatic properties. The tertiary structure of β -glycosidase from *S. frugiperda* (Sf β gly) was analyzed as a network. After calculating the averaged paths between all pairs of amino acid residues of Sf β gly, we defined 11 hubs, which have the highest centrality on the network. Sequence alignment and structural comparison showed that these hubs residue are conserved among β -glycosidases. Our goal was to mutate most hubs and 3 'non-hubs' residues from Sf β gly, express these mutant enzymes in *E. coli*, test their enzymatic properties as catalytic efficiency and substrate preference, and verify the thermal stability of these mutants. The results implied that mutations in these hubs do not cause changes in catalytic properties although enzymes containing mutations in hubs showed lower thermal stability. Based on that, it was suggested that hub residues are important in the diffusion of kinetic energy (vibrations) introduced in the Sf β gly structure by heating.

Keywords: β -glucosidase, network, hubs, mutation, enzyme activity, thermal inactivation

LISTA DE ABREVIATURAS E SIGLAS

Abs – absorbância

BLAST - ‘basic local alignment search tool’

C – coeficiente de agregação

Celo – celobiose

CLT – mutante (N329A) inativo da enzima Sfβgli

EDTA – ácido etilenodiaminotetracético

EL – proteína eluída da resina após lavagens, incubação com tampão contendo 300 mM de imidazol e dessalinização

ES – complexo enzima substrato

ES[‡] - complexo enzima substrato no estado de transição

F – fluorescência máxima na presença de acrilamida.

F_o – fluorescência máxima na ausência de acrilamida

FT – flow-through que corresponde às proteínas não retidas pela resina Ni-NTA agarose obtido durante a etapa de ligação na resina

GH 1 – família 1 das glicosídeo hidrolases

I – cultivo bacteriano induzido

IPTG – isopropil-β-D-tio-galactosídeo

k – número de conexões feitas por um nó (resíduo de aminoácido)

k - constante da taxa de inativação térmica

K_{sv} – constante de Stern-Volmer

Δ*K_{sv}* – diferenças de *K_{sv}* entre a Sfβgli selvagem e mutantes

L – via mínima média

L_{kretirado} – via mínima média da rede após a remoção de um resíduo de aminoácido

lac Z – deleção de parte da sequência codificadora da enzima β-galactosidase da estirpe bacteriana NovaBlue(DE3).

LB – meio de cultivo Luria Broth

Lis – fração solúvel do lisado obtido após lise das bactérias induzidas e centrifugação

Lv1 – primeira lavagem das proteínas não retidas pela resina

Lv5 – última lavagem das proteínas não retidas pela resina

Mu β glc – metilumbeliferil- β -D-glucopiranosídeo

ÑI – cultivo bacteriano não induzido

NP β fuc – p-nitrofenil β fucosídeo

NP β gal – p-nitrofenil β galactosídeo

NP β glc - p-nitrofenil β glucosídeo

PCR – reação em cadeia com DNA polimerase

PDB – ‘protein data bank’

PEL – “pellet” oriundo da centrifugação do extrato após a lise bacteriana

PMSF – phenylmethanesulfonyl fluoride

Q – concentração de supressor (acrilamida)

rpm – rotações por minuto

Res – proteína retida pela resina mesmo após a etapa de eluição

SDS – dodecil sulfato de sódio

SDS-PAGE – eletroforese em gel de poliacrilamida na presença de SDS

Sf β gli – β -glicosidase de *Spodoptera frugiperda* (lagarta do cartucho)

Sumário

1. Introdução	10
1.1 β -glicosidases.....	11
1.2 Teoria de redes e proteínas.....	13
1.3. Motivação e justificativa	17
2. Objetivos	19
3. Materiais e métodos	20
3.1 Modelagem da estrutura tridimensional de SfbgI	20
3.2 Identificação dos nós centrais (“resíduos hubs”) da rede estrutural da enzima SfbgI	20
3.3 – Experimentos de mutagênese sítio-dirigida	22
3.4 – Sequenciamento das SfbgI mutantes.....	22
3.5 – Expressão de SfbgI em bactérias	22
3.6 – Purificação da SfbgI recombinante.....	23
3.6.1 - Lise das bactérias induzidas.....	23
3.6.2 - Cromatografia de afinidade.....	23
3.7 – Diálise e dessalinização da SfbgI purificada (‘desalting’)	24
3.8 – SDS-PAGE	24
3.9 – Determinação da conformação proteica por meio de espectros de fluorescência intrínseca e ensaios de supressão de fluorescência intrínseca	25
3.10 - Obtenção dos parâmetros cinéticos para as SfbgI mutantes e selvagem.....	25
3.11 – Determinação da estabilidade térmica a 45°C por meio da determinação de atividade enzimática remanescente.	26
3.12 – Cálculos de parâmetros termodinâmicos	27
4. Resultados e Discussão	28
4.1 Identificação dos resíduos centrais (hubs) da rede estrutural de SfbgI.....	28
4.2 Construção dos mutantes a partir da sequência de SfbgI clonada no vetor pET46-EK/LIC	47
4.3 - Expressão da proteína SfbgI recombinante e purificação	52
4.4 – Avaliação da estrutura global das SfbgI mutantes.	59
4.4.1 – Espectros de fluorescência intrínseca das SfbgI mutantes.....	59
4.4.2 Ensaio de apagamento de fluorescência das SfbgI mutantes	63
4.5 - Determinação dos parâmetros cinéticos das SfbgI mutantes	68
4.6 Estabilidade térmica das SfbgI mutantes	87
4.7 – Análise comparativa do papel de resíduos “hubs” e “não-hubs”	94
5. Conclusão.....	99

6. Referências Bibliográficas	100
<i>SÚMULA CURRICULAR</i>	104

1. Introdução

Os processos da vida são regidos pelas funções precisas desempenhadas pelas proteínas, que adotam estruturas tridimensionais únicas. Vasta quantidade de pesquisa experimental e teórica tem sido realizada nas últimas décadas para elucidar os princípios envolvidos na estrutura, estabilidade e enovelamento proteico. A análise de genomas completos, identificação de novas biomoléculas e proteínas gera uma quantidade cada vez maior de material para estudo e conseqüentemente maior complexidade na elucidação das interações e funções biológicas das proteínas (Whisstock e Lesk, 2003; Banaszynski e Wandless, 2006). Embora grandes avanços tenham ocorrido no campo da biologia molecular, ainda há espaço para o desenvolvimento de novas técnicas e diferentes perspectivas para a compreensão da singularidade e funções das proteínas. A disponibilidade de um grande número de estruturas proteicas no banco de dados de proteínas (PDB) tem auxiliado incisivamente o entendimento de alguns desses aspectos. A correlação entre estrutura-função em várias proteínas vem sendo elucidada em termos globais assim como a disposição das cadeias laterais de aminoácidos específicos nos sítios ativos. Contudo, o papel (função) da maioria dos aminoácidos nas estruturas proteicas ainda não é totalmente compreendido (Whisstock e Lesk, 2003; Lee *et al*, 2007). A sequência de aminoácidos e a configuração tridimensional adotada por estas determinará a disposição de seus resíduos em posições específicas que por sua vez provê a estabilidade, funcionalidade e os sítios de interação da estrutura resultante. Acredita-se que é possível apenas com a sequência primária de uma proteína prever suas prováveis funções e portanto quais seriam os seus resíduos críticos. Todavia há muitos exemplos de proteínas com estruturas similares (homólogas), porém com funções completamente não relacionadas e, inversamente, proteínas não homólogas que apresentam padrão de enovelamento muito distinto com funções similares (Whisstock e Lesk,

2003). Logo há outros fatores, em parte desconhecidos, que estão envolvidos na funcionalidade proteica.

Várias ferramentas computacionais vêm sendo desenvolvidas para auxiliar a compreensão das relações entre estrutura e função proteicas. Dentre estas novas técnicas, a aplicação da teoria de redes e grafos às proteínas tem sido explorada, auxiliando no entendimento das propriedades dessas estruturas (Csermely 2008; Greene 2012).

Nesta concepção a proteína é interpretada como uma rede, um conjunto de pontos (nós ou vértices) que mantêm alguma relação entre si (conexão ou aresta) (Barabási e Oltvai, 2004; Greene e Higman, 2003). Deste modo, os aminoácidos são nós (ou vértices) e as interações covalentes e não-covalentes são conexões (ou arestas) entre dois nós. Assim em uma rede protéica observa-se que de um único nó (aminoácido) podem partir múltiplas conexões (interações). Este tipo de representação da estrutura protéica gera uma intrincada rede com enorme número de conexões e com propriedades que vão além daquelas visualizadas apenas com a estrutura covalente. Neste sentido, a teoria de redes ajuda-nos a entender o comportamento de sistemas complexos, como as proteínas, pela redução deste sistema a um conjunto de elementos interagindo e unidos por conexões.

Considerando que nos últimos anos o conceito de redes estruturais proteicas tem sido explorado, analisou-se neste trabalho a β -glicosidase de *Spodoptera frugiperda* (Sf β gli) sob esta perspectiva, aplicando-se uma das técnicas de predição de aminoácidos importantes à funcionalidade proteica embasada nesta teoria e avaliou-se experimentalmente o impacto da alteração destes resíduos de aminoácidos em propriedades funcionais desta enzima.

1.1 β -glicosidases

As β -glicosidases são enzimas que hidrolisam ligações glicosídicas e atuam no metabolismo de glicosídeos e glicolípídeos em animais, defesa, lignificação da parede celular, ativação de

fito-hormônios, liberação de compostos aromáticos em plantas e conversão da biomassa em microorganismos (Cairns e Esen, 2010). Estas enzimas são amplamente usadas em processos biotecnológicos e industriais e também estão envolvidas em patologias humanas (Cairns e Esen, 2010; Service, 2007; Dvir *et al*, 2003). As β -glicosidases constituem o maior grupo entre as glicosídeo hidrolases e catalisam a hidrólise de ligações β -1,4 glicosídicas do tipo O ou S em uma variedade de glicosídeos pela remoção de monossacarídeos da extremidade não-redutora do substrato. São encontrados em vários organismos de Archaea a Eucariotos (Catarel *et al*, 2009).

Todos os membros da família 1 das β -glicosidases compartilham a mesma estrutura do tipo $(\alpha/\beta)_8$ barril com o sítio ativo localizado na porção C-terminal do barril. São glicosidases retentoras, isto é, mantêm a configuração do carbono anomérico no produto. O sítio ativo das β -glicosidases apresenta um subsítio de ligação da extremidade não-redutora do substrato, conhecido como subsítio -1 que interage com o monossacarídeo da porção não-redutora do substrato (glicone), enquanto a extremidade redutora do substrato é acomodada nos subsítios +1, +2, +3, etc... também denominados de subsítio do aglicone. (Davies *et al*, 1997). A hidrólise do substrato ocorre entre os subsítios -1 e +1 através de um mecanismo que é dividido em 2 etapas (glicosilação e deglicosilação) e requer a participação de 2 ácidos glutâmicos (E) catalíticos. A reação envolve estes 2 aminoácidos catalíticos presentes em regiões conservadas da enzima. Na etapa de glicosilação o ácido glutâmico do motivo I/VTENG atua como nucleófilo e o ácido glutâmico do motivo TFNEP age como ácido/base catalítico. Contudo nas mirosinases este último motivo é substituído por TINQL, no qual o ácido glutâmico catalítico é substituído por uma glutamina (Burmeister *et al*, 1997).

A especificidade do subsítio do glicone está relacionada com a presença de uma rede de interações não-covalentes envolvendo as hidroxilas do substrato e um ácido glutâmico, uma histidina, um triptofano e uma glutamina altamente conservados (E451, H142, W444 e Q39,

respectivamente na β -glicosidase de *Spodoptera frugiperda*) (Marana *et al.*, 2002; Marana, 2006). A especificidade pelo aglicone no subsítio +1 é determinada pelo balanço de interações não-covalentes envolvendo o substrato e os resíduos E190, E194 e K201 na β -glicosidase de *Spodoptera frugiperda*. Contudo, estes resíduos não são conservados na família GH1, dado coerente com a grande variabilidade na preferência pelo aglicone exibida por estas enzimas (Mendonça e Marana, 2008). Assim, nas β -glicosidases GH1 este conjunto de aminoácidos do interior do sítio ativo (subsítio -1 e +1) e os dois ácidos glutâmicos catalíticos participam da ligação do substrato e estabilização do complexo enzima-substrato no estado de transição. Logo, são conjuntamente responsáveis pela atividade e especificidade destas enzimas (Marana, 2006).

Como já mencionado as β -glicosidases da família 1 são enzimas com papéis relevantes em diversos processos biológicos e consequentemente com muitas aplicações industriais e biotecnológicas. Por conseguinte, o entendimento dessas enzimas é relevante para os processos em que estejam envolvidas. Assim o estudo das β -glicosidases com uma nova abordagem (teoria de redes) se torna interessante e pode complementar o estudo destas enzimas obtido por técnicas tradicionais, como por exemplo, comparação de resíduos entre proteínas homólogas (Whisstock e Lesk; 2003; Jones e Thornton; 2004) e mutações seletivas obtidas por experimentos de evolução dirigida (Kuchner e Arnold; 1997; Chica *et al.*, 2005).

1.2 Teoria de redes e proteínas

Nos últimos anos houve um crescente interesse e muitas pesquisas em torno de redes complexas, ou seja, redes cuja estrutura é irregular, complexa e dinâmica modificando-se ao longo do tempo. Estas redes incluem redes de transporte, telefônicas, redes sociais, internet, redes neurais ou genéticas, metabólicas e redes proteicas. A pesquisa em torno de redes complexas iniciou-se com esforços para se definir novos conceitos e métricas para caracterizar

a topologia das ‘redes reais’. O resultado desses estudos tem sido a identificação de princípios unificadores e propriedades estatísticas comuns à maioria das ‘redes reais’ (Boccaletti *et al.* 2006).

Para estudo das propriedades de redes há 3 parâmetros básicos que podem ser determinados: a curva de distribuição de graus (k) dos nós, a via mínima média (L) e o coeficiente de agregação (C). O termo grau (k) refere-se ao número de conexões feitas por um nó, ou o grau de um nó é simplesmente o número de outros nós no grafo ao qual ele está conectado. Assim para uma rede é possível determinar a porcentagem de nós com diferentes graus ($k = 1, 2$, etc...) e construir um gráfico representando a curva desta distribuição. Baseado neste tipo de curva, uma rede pode ser classificada dentro de pelo menos 3 grandes classes: randômica, “mundo pequeno” (“small-world”) e regular (Figura 1). Na rede randômica, a curva de graus segue uma distribuição de Poisson, ou seja, a maioria dos nós apresenta o mesmo grau de conectividade e nós extremamente conectados (alto k) são muito raros. Já nas redes de “mundo pequeno”, a curva de graus segue distribuições do tipo exponencial, indicando uma maior presença de nós altamente conectados, também chamados de nós centrais ou “hubs”. Por fim, as redes regulares apresentam todos os nós com o mesmo grau de conexão, o que indica que nós centrais são inexistentes.

Figura 1 – Representação esquemática de 3 tipos de redes: regular, mundo pequeno e randômica (da esquerda para direita, respectivamente). Baseado em Watts e Strogatz (1998).

Os parâmetros via mínima média (L) e coeficiente de agregação (C) são usados para caracterização de qualquer uma das classes de rede. Assim, sendo que em uma rede a distância entre nós é medida em número de conexões (via), pode-se determinar que entre dois nós

especificados há uma via mínima, ou seja, uma via que contém o menor número de conexões possível. Portanto, calculadas todas as vias mínimas para todos os pares de nós existentes em uma rede, pode-se determinar a via mínima média (L), que é a média das distâncias mínimas entre todos os pares de nós de um grafo. Este parâmetro indicará a “navegabilidade” de uma rede, ou seja, quanto menor L mais facilmente se atinge outro ponto da rede. Por fim, o coeficiente de agregação (C) indica a probabilidade de três nós da rede serem interconectados, C é a média de quão conectados estão entre si os vizinhos de cada nó da rede. Logo, quanto maior C, maior a presença de agregados no interior da rede. Significaria, por exemplo, que numa rede de pessoas conhecidas, 2 indivíduos com um amigo em comum provavelmente se conheçam. (Boccaletti *et al*, 2006; Barabási e Oltvai, 2004).

Outra forma de se caracterizar o tipo de grafo (rede) utiliza os parâmetros L e C; isto é grafos regulares são caracterizados por valores de L e C altos, enquanto grafos randômicos tem valores pequenos de L e C e as redes mundo pequeno (“small-world”) está no meio destes dois extremos, com valores de L pequenos e de C grandes (del Sol e O’Meara, 2005).

A análise da estrutura espacial e contatos em centenas de proteínas pertencentes a 9 diferentes tipos de dobramento evidencia que redes protéicas possuem uma curva de distribuição de graus de conexão do tipo exponencial, um alto índice de agregação (C), uma extensão de via mínima média (L) pequena e a presença de nós centrais, características que indicam que não se tratam de redes randômicas, sendo classificadas dentro do grupo das redes de mundo pequeno, se aproximando de um grupo particular chamado de redes sem escala (sem referencial; “scale-free”) (Greene e Higman, 2003). Estas propriedades indicam que os nós da rede (aminoácidos) encontram-se conectados em grupos e que estes grupos são interconectados através de poucos nós (aminoácidos) dos quais partem muitas conexões. A presença destes nós contribui para determinação de vias (conjuntos de conexões) mais curtas entre quaisquer outros pontos da rede. Devido a estas propriedades propõe-se que a rede protéica é “robusta”, ou seja,

mutações aleatórias tendem a atingir os nós com baixo grau de conexão e portanto causam apenas pequeno efeito local e não se propagam causando dano global na estrutura. Por outro lado, postula-se que as mutações direcionadas para os nós com alto grau de conexão ou sua proximidade podem se propagar e modificar significativamente uma proteína (Greene e Higman, 2003). Há diferentes trabalhos baseados em teoria de redes que propõem métodos para identificar quais seriam os nós centrais da rede estrutural proteica (del Sol e Meara, 2005; del Sol *et al*, 2006; Csermely, 2008).

Segundo trabalhos de del Sol *et al* (2006) e Amitai *et al* (2004), a identificação dos nós com alto grau de conexão, também chamados nós centrais da rede (hubs) é feita determinando-se a via mínima que conecta um específico aminoácido a cada um dos demais de uma proteína. Logo, aminoácidos com valores de via mínima média menores do que a média geral são pontos centrais, pois tendem a “encurtar” a via de conexão entre dois quaisquer aminoácidos. Esta análise pode também ser feita calculando a “via mínima média” da rede como um todo ($L_{\text{médio}}$) na proteína selvagem e em mutantes onde cada um dos aminoácidos é artificialmente removido dos cálculos ($L_{\text{k retirado}}$). Consequentemente os aminoácidos que quando deletados tendem a aumentar de maneira significativa a “via mínima média” ($L_{\text{médio}}$) são nós centrais na rede. Outra forma de se analisar a importância (centralidade) de um elemento da rede é verificar-se se este é necessário para um grande número de vias mínimas, onde o termo via mínima significa rota mais curta entre dois elementos da rede (Barabasi e Oltvai, 2004; Boccaletti *et al*, 2006; Csermely, 2006).

Ruth Nussinov *et al* (2006) reduziram a rede de estruturas proteicas de 7 grandes famílias pela remoção de segmentos que não afetavam a via mínima média (L) de forma contundente. Com este método eles construíram um ‘arcabouço’ estrutural contendo somente os aminoácidos que eram suficientemente centrais para o fluxo de informação da proteína como um todo. De fato eles relataram que esses ‘determinantes de interconectividade conservados’

ou nós centrais (hubs) eram elementos chave na comunicação entre os sítios alostéricos e sítios ativos (eg sítios catalíticos). Interessantemente, encontraram resíduos (hubs) da protease de HIV-1 fora do sítio ativo cujas mutações levam à resistência a drogas.

Outros trabalhos analisando conjuntos de proteínas também relatam que os nós centrais da rede são encontrados em regiões funcionalmente importantes como sítios ativos, sítios de ligação de efetadores alostéricos e interfaces de interação proteína-proteína (Amitai *et al.*, 2004; Böde *et al.*, 2007).

De fato, é no estudo da alosteria que o conceito de redes estruturais proteicas tem sido frequentemente aplicado e submetido a testes experimentais (del Sol *et al.*, 2009). Por exemplo, na caspase 1 foi detectada uma rede de ligações de hidrogênio envolvendo cadeias laterais de 9 resíduos de aminoácidos conectando o sítio ativo a um sítio alostérico. A substituição destes resíduos por alaninas evidenciou que dois deles eram essenciais para propagação de efeitos através da rede de contatos (del Sol *et al.*, 2009, Data *et al.*, 2008). Redes de contatos envolvidas em alosteria também foram experimentalmente analisadas nas proteínas PDZ e CBP (Lockless e Ranganathan, 1999; Fuentes *et al.*, 2004; Bruschweiler *et al.*, 2009). Contudo, o papel dos nós centrais (“resíduos hubs”) de redes estruturais de proteínas em propriedades enzimáticas específicas fora do contexto de alosteria não foi ainda experimentalmente analisado.

1.3. Motivação e justificativa

Apesar de todas β -glicosidases da família 1 possuírem o mesmo tipo de estrutura básica, muito similares entre si, sua identidade de sequência nem sempre é igualmente alta (Husebye *et al.*, 2005; Czjzek, 2001). Além disso, os aminoácidos que formam os subsítios de interação com o substrato são um tanto variáveis (Czjzek *et al.*, 2000; Verdoucq *et al.*, 2003; Mendonça e Marana, 2008 e 2011) e existem resíduos fora do sítio ativo que estão envolvidos na especificidade das β -glicosidases (Verdoucq *et al.*, 2003; Marana, 2006; Mendonça e Marana,

2011). Adicionalmente, vários estudos postulam a existência de grupos de aminoácidos centrais na rede proteica que são críticos no fluxo da informação e influenciam a atividade proteica (Amitai *et al*, 2005; del Sol *et al*, 2006; Csermely 2008). Estas observações nos levaram a empregar uma forma de predição de aminoácidos centrais (resíduos hubs) na rede estrutural de proteínas e verificar se modificações nestes resíduos poderiam afetar algumas propriedades da β -glicosidase da ‘lagarta do cartucho’ *Spodoptera frugiperda* (Sf β gli).

2. Objetivos

O objetivo geral desta tese foi caracterizar experimentalmente a importância dos resíduos de aminoácidos que constituem os nós centrais (“resíduos hubs”) da rede estrutural da β -glicosidase Sf β gli na sua atividade catalítica, especificidade pelo substrato e estabilidade térmica. Para isso foram traçados alguns objetivos estratégicos:

- a) Substituir os 10 principais “resíduos hubs” de Sf β gli por alaninas através de mutagênese sítio dirigida, reduzindo assim pela “remoção quase total” da cadeia lateral o grau de conectividade destes resíduos.
- b) Substituir alguns “resíduos não hubs” por alaninas através de mutagênese sítio dirigida.
- c) Expressar estas Sf β gli mutantes e a selvagem em bactéria e purificá-las.
- d) Avaliar a integridade estrutural das Sf β gli mutantes através de experimentos de fluorescência intrínseca e de supressão de fluorescência por acrilamida
- e) Determinar os parâmetros cinéticos (k_{cat} e K_m) para hidrólise de *p*-nitrofenil β -glicosídeos (fucosídeos, glucosídeos e galactosídeos) e celobiose para estas Sf β gli mutantes e compará-los à enzima selvagem.
- f) Determinar a estabilidade das Sf β gli mutantes e selvagem em experimentos de inativação térmica.

3. Materiais e métodos

3.1 Modelagem da estrutura tridimensional de Sfβgli

O modelo estrutural de Sfβgli foi construído baseado na homologia com a mirosinase do afídeo *Brevicoryne brassicae* (1WCG) (Husebye, *et al.* 2005). A modelagem por homologia foi realizada com o servidor Swiss Model e a estrutura visualizada com os programas DeepView/Swiss PDB Viewer v3.7 (Guex, *et al.* 2007) e PyMol (DeLano Scientific LCC). A estabilidade deste modelo foi avaliada por dinâmica molecular e determinação do raio de giração e rmsd ao longo de 30.000 ps. Esta avaliação foi conduzida pelo Prof. Dr. Guilherme M. Arantes (IQUSP – Departamento de Bioquímica). Considerando que os parâmetros avaliados mantiveram relativamente estáveis ao longo deste período de tempo o modelo foi aceito.

Figura 2 – Parâmetros de avaliação da estabilidade do modelo estrutural de Sfβgli. Rg, raio de giração (nm); RMSD, root mean square deviation.

3.2 Identificação dos nós centrais (“resíduos hubs”) da rede estrutural da enzima Sfβgli

A identificação dos ‘resíduos hubs’ foi baseada no trabalho de del Sol *et al.* (2006). Inicialmente a estrutura terciária de Sfβgli foi modelada como um grafo, onde os resíduos de

aminoácidos eram os vértices ou nós e as interações não-covalentes entre os aminoácidos (contatos entre os aminoácidos) eram as arestas. Para identificação destes contatos, com base na estrutura de SfbgI foi construído um mapa de contatos entre os seus resíduos de aminoácidos, sendo que a distância limite usada para considerar que dois dados aminoácidos mantinham contato foi 5 Å. Em seguida, calculou-se a via mínima média (L) da rede estrutural de SfbgI empregando a seguinte equação:

$$L = [\sum d(i,j)] / N_p$$

onde N_p é o número de pares de aminoácidos (nós) que têm contato (distância menor que 5 Å), $d(i,j)$ é a menor distância entre os aminoácidos (nós) i e j medida em número de contatos (arestas).

Finalizada esta etapa, a via mínima média foi recalculada após a remoção individual da rede estrutural de cada resíduo de aminoácido e seus contatos ($L_{rem.k}$).

Assim, com base neste novo parâmetro foram calculadas as diferenças de L causadas pela remoção de cada resíduo (ΔL_k).

$$\Delta L_k = L - L_{rem.k}$$

Com base nestas diferenças foi determinado o ΔL_k médio e o seu desvio padrão (σ).

Finalmente foram calculados os valores de z (z-score) para a remoção de cada aminoácido, ou seja, os dados de ΔL_k foram normalizados em termos de números de desvios padrão (σ) de ΔL_k médio:

$$z_k = (\Delta L_k - \Delta L_k \text{ médio}) / \sigma$$

Aqueles resíduos cuja remoção causaram as maiores mudanças na via mínima média (L), variações maiores que 2,5 z ($z > 2,5$), foram considerados “resíduos hubs”.

Os programas computacionais necessários para estes cálculos foram desenvolvidos pelo Prof. Dr. Guilherme M. Arantes (IQUSP – Departamento de Bioquímica).

3.3 – Experimentos de mutagênese sítio-dirigida

Os experimentos de mutagênese foram efetuados utilizando-se o “kit” “Quick Change site directed Mutagenesis (Stratagene)” tendo como molde um plasmídeo pET46-Sfβgli que codifica para a Sfβgli selvagem.

Esse experimento consistiu na amplificação por PCR de toda a extensão do vetor utilizando dois “primers” mutagênicos, que durante a amplificação foram incorporados aos produtos da reação, resultando em vetores contendo a mutação desejada. Cada mutação foi feita com um par de “primers” específicos para os sítios a serem mutados. Após as reações de mutagênese sítio-dirigida, foram feitas ‘minipreps’ para obtenção de uma quantidade maior de plasmídeos codificadores das Sfβgli mutantes. Assim, cultivou-se bactérias XL1BLUE transformadas com estes plasmídeos em um volume de 10 ml de meio LB contendo 50 µg/µl de carbenicilina durante 16 horas. Após este período, sedimentou-se essas bactérias e usou-se o “kit” da Promega (Wizard Plus Minipreps DNA Purification System) seguindo as instruções do fabricante.

3.4 – Sequenciamento das Sfβgli mutantes

Após a obtenção dos vetores codificadores das Sfβgli mutantes, realizou-se reações de sequenciamento a fim de se averiguar a presença das mutações desejadas. Foram utilizados “primers” posicionados em diferentes pontos da sequência codificadora de Sfβgli, de modo a sequenciá-la integralmente.

3.5 – Expressão de Sfβgli em bactérias

A indução da expressão das Sfβgli selvagem e mutantes foi feita em células *E. coli* BL21(DE3) (Stratagene) e NovaBlue-DE3 (Novagen). Inicialmente foi preparado um pré-

inóculo que consistiu da adição de uma colônia previamente transformada com o plasmídeo pET 46-Sfβgli em 50 mL de meio LB com ampicilina (50μg/mL). Este pré-inóculo foi então incubado a 37 °C por 16 horas a 150 rpm. Após este período, o pré-inóculo foi adicionado a 1,1 L do mesmo meio, voltando a ser incubado a 20 °C a 200 rpm até atingir uma densidade óptica em 600 nm entre 0,3 e 0,8. Ao atingir valores neste intervalo foi adicionado IPTG (isopropil-β-D-tio-galactosídeo 1 mM em água) como indutor da expressão da proteína recombinante. A seguir as células foram incubadas por 16-20 horas a 150 rpm a 20 °C. Terminado este período, as células foram sedimentadas por centrifugação por 30 minutos a 7.000 xg, 4 °C, e congeladas a -20 °C.

3.6 – Purificação da Sfβgli recombinante

3.6.1 - Lise das bactérias induzidas.

O sedimento (“pellet”) das bactérias induzidas a expressar a Sfβgli selvagem e mutantes – proveniente de 0,5 L de cultivo - foi ressuspensão em 15 mL de tampão de lise (fosfato de sódio 20 mM, pH 7,4; contendo 0,5 M NaCl, 20 mM de imidazol e o inibidor de proteases PMSF). A suspensão foi sonicada (6 ciclos de sonicação de 20 segundos em “output” 3 (Sonicador Branson adaptado com “microtip”) com intervalos de 1 minuto de incubação em gelo). Na sequência, o material lisado foi centrifugado a 10.000 x g, 4°C, por 30 minutos. O sobrenadante foi coletado e utilizado durante os procedimentos de purificação.

3.6.2 - Cromatografia de afinidade.

O vetor, no qual a proteína Sfβgli foi clonada (pET46), codifica a proteína de interesse fusionado a uma cauda de 6 histidinas em sua porção N-terminal, permitindo a purificação da proteína de interesse através de cromatografia de afinidade por Níquel.

A purificação das Sfβgli recombinante foi realizada por meio de uma cromatografia de afinidade em 'batch'. Em microtubos plásticos de 1,5 mL, aproximadamente 1mL de sobrenadante do lisado das bactérias induzidas foi incubado com 200μL resina de Ni⁺²-NTA agarose (Qiagen) por 1 hora em homogeneizador (Heavy dietyrotator, Cole Parmer) em baixa velocidade. Foram realizadas 5 lavagens com 850 μL tampão de lise para remoção das proteínas não retidas pela resina. As proteínas retidas especificamente pela resina foram eluídas com 2 ciclos de lavagem com 150 μL de tampão de lise contendo 300 mM de imidazol, seguindo-se incubação por 7 minutos em gelo, centrifugação a 16.000 g e coleta do sobrenadante. As alíquotas de cada etapa da lise e purificação foram analisadas por SDSPAGE.

3.7 – Diálise e dessalinização da Sfβgli purificada ('desalting')

Algumas frações que foram obtidas após a eluição da proteína recombinante foram submetidas à diálise ou alternativamente à dessalinização com o uso de colunas pré-empacotadas HiTrap desalting column (GE Healthcare).

No caso de diálise, a fim de se retirar o imidazol que se encontra na solução utilizada na eluição da proteína recombinante, as frações obtidas após a purificação que continham atividade β-glicosidásica foram dialisadas em tampão citrato-fosfato 50 mM pH 6,0 por 6 a 12 horas a 4 °C. Foi utilizado um volume de tampão $6,7 \cdot 10^6$ vezes maior que o volume da amostra de proteína. Alternativamente optou-se pelo uso das colunas de "desalting" de 5 ml (HiTrap desalting column – GE Healthcare), isto é, estas frações eluídas foram passadas pelas colunas pré equilibradas com tampão citrato-fosfato 50 mM pH 6,0, segundo as instruções do fabricante.

3.8 – SDS-PAGE

Eletroforeses em gel de poliacrilamida 12 % na presença de SDS 0,1% (SDS-PAGE) foram executadas segundo Laemmli (1970). As amostras contendo proteínas (20-35 μL)

juntamente com 10 µL de tampão de amostra (tris-HCl 60 mM pH 6,8, SDS 2,5% (p/v), β-mercaptoetanol 0,36 mM, EDTA 0,5 mM, glicerol 10 % (v/v) e azul de bromofenol 0,05 % (p/v)) foram previamente fervidas por 5 min.

A coloração dos géis foi feita com Comassie Blue R-250 0,1% (p/v) preparado em uma solução de metanol 40% (v/v) e ácido acético 10% (v/v).

3.9 – Determinação da conformação proteica por meio de espectros de fluorescência intrínseca e ensaios de supressão de fluorescência intrínseca

Espectros de fluorescência intrínseca da Sfβgli selvagem e mutantes foram coletados a 30 °C utilizando-se excitação de 295 nm e emissão na faixa de 305 até 450 nm (Espectrofluorímetro F4500, Hitachi). Foram empregadas cubetas de quartzo de 10 mm.

Em seguida foram feitos ensaios de supressão de fluorescência intrínseca das amostras da Sfβgli selvagem e mutantes por meio da adição do apagador de fluorescência acrilamida (Sigma-Aldrich) em uma faixa de concentração de 0 a 0,47 M. A coleta de espectros na presença das diferentes concentrações de acrilamida foi feita seguindo os mesmos parâmetros das coletas anteriores. A partir destes dados foram construídos gráficos de F_0/F *versus* concentração de acrilamida, onde F_0 e F correspondem à fluorescência máxima na ausência e na presença de acrilamida, respectivamente. As leituras de fluorescência (F) foram corrigidas para a diluição ocorrida devido à adição de acrilamida. Com as inclinações destes gráficos lineares determinou-se a constante de Stern-Volmer (K_{sv}) para as enzimas selvagem e mutantes baseando-se na equação $F_0/F = K_{sv} [Q] + 1$, onde Q é a concentração de acrilamida (Eftink and Ghiron, 1976).

3.10 - Obtenção dos parâmetros cinéticos para as Sfβgli mutantes e selvagem

A caracterização cinética das Sfβgli mutantes consistiu na determinação das velocidades iniciais de hidrólise para diferentes concentrações dos substratos *p*-nitrofenil β-fucosídeo

(NP β fuco), *p*-nitrofenil β -glucosídeo (NP β glc), *p*-nitrofenil β -galactosídeo (NP β gala) e celobiose (celo). As concentrações de substrato utilizadas foram: 0,05; 0,1; 0,15; 0,2; 0,25; 0,3; 0,35; 0,4; 0,45; 0,5; 1; 1,5; 2; 2,5; 4; 5; 6; 7; 8; 9; 10 mM diluídos em tampão citrato-fosfato 100 mM pH 6,0. Estes ensaios foram feitos misturando-se 50 μ L de amostra contendo enzima e 50 μ L dos substratos nas concentrações acima descritas. A mistura de reação foi incubada a 30°C por diferentes intervalos de tempo para cada substrato e Sf β gli. No caso dos *p*-nitrofenil β -glicosídeos, as reações foram interrompidas com 100 μ L de uma solução de carbonato-bicarbonato 250 mM pH 11 e o produto *p*-nitrofenolato formado foi detectado por meio da absorbância em 420 nm. No caso da celobiose, a glicose formada foi detectada com um reagente comercial contendo glicose oxidase (Analisa). Em todos os casos, as absorbâncias foram convertidas em moles de produto com uso de uma curva padrão. Finalmente, os valores de v_o foram calculados e ajustados na equação de Michaelis-Menten usando o software Enzfitter, sendo assim obtidos os parâmetros cinéticos K_m e k_{cat} para as enzimas selvagem e mutantes.

3.11 – Determinação da estabilidade térmica a 45°C por meio da determinação de atividade enzimática remanescente.

Foram realizados ensaios de estabilidade térmica da Sf β gli selvagem e mutantes incubando amostras a 45°C, sendo retiradas alíquotas após 0, 5, 10, 15, 20, 25, 30 e 35 minutos de incubação (em alguns casos retirou-se alíquotas a cada 0,5 min no intervalo de tempo entre 0 a 5 minutos de incubação). Com estas alíquotas fez-se ensaios de atividade enzimática com metilumbeliferil- β -D-glucopiranosídeo (Mu β -gluco) 4 mM em tampão citrato fosfato 100 mM, pH 6,2, os quais foram interrompidos pela adição de 1 mL de solução glicina-NH₄OH 200 mM, pH 10,5, sendo detectada a fluorescência do produto metilumbeliferona com uma excitação de 360 nm e emissão de 450 nm. Na sequência construiu-se gráficos do logaritmo da atividade enzimática remanescente relativa *versus* o tempo de incubação a 45 °C (Segel, 1993). A

inclinação da reta obtida representa a constante de velocidade (k) do processo de inativação térmica da enzima analisada.

3.12 – Cálculos de parâmetros termodinâmicos

Diferenças de energia do complexo $ES^\ddagger$ ($\Delta\Delta G^\ddagger$) correspondendo à hidrólise de um mesmo substrato pela enzima selvagem e uma mutante foram calculadas com a equação $\Delta\Delta G^\ddagger = RT \ln(k_{\text{cat wt}}/k_{\text{cat mut}})$ (Fersht, 1999). Já as diferenças de energia de formação do complexo ES ($\Delta\Delta G_b$) foram calculadas com a equação $\Delta\Delta G_b = RT \ln(K_{\text{m mut}}/K_{\text{m wt}})$. Nestas equações $R = 8,3144 \text{ J.mol}^{-1}.\text{K}^{-1}$ e $T = 303 \text{ K}$.

4. Resultados e Discussão

4.1 Identificação dos resíduos centrais (hubs) da rede estrutural de Sfβgli

Com base no seu mapa de contatos em uma distância limite de 5 Å a estrutura terciária de Sfβgli foi transformada em um grafo (rede) (Figura 3).

Figura 3 – Representação da rede estrutural de Sfβgli. A linha preta no perímetro do círculo representa cada uma das posições da estrutura primária da proteína, enquanto que as linhas cinza representam contatos entre resíduos de aminoácidos. Representação obtida com o software AminoNet 2.0.

Com base na análise no “software” AminoNet 2.0 observou-se uma via mínima média (L) de 4,8 e um índice de agregação (C) de 0,485. Para redes regulares de semelhante dimensão os valores esperados de L e C seriam 29 e 0,643, respectivamente (Bagler e Sinha, 2005), enquanto que redes randômicas com o mesmo número de nós teriam $L = 2,6$ e $C = 0,022$. Deste modo, a grande diferença nos valores de L para rede estrutural de Sfβgli e uma rede regular, sugere que os contatos entre os resíduos de aminoácidos não seguem um padrão repetitivo, ou seja, a rede proteica não é regular. Já com base no critério de que redes com $L \geq L_{\text{randômico}}$ e $C \gg C_{\text{randômico}}$ apresentam propriedades de “mundo pequeno” (small world) (Watts e Strogatz, 1998), assume-se tal classificação para rede estrutural de Sfβgli, indicando que em geral

quaisquer pares de resíduos são conectados por um número relativamente pequeno de contatos (4,8 em média) e que há uma alta agregação local entre resíduos. Redes randômicas e redes regulares não exibem estas propriedades (Atilgan *et al.*, 2004).

Estas observações para Sfbglic são coerentes com a análise estrutural de outras proteínas. Vendruscolo *et al.* (2002) estudaram um conjunto de 978 proteínas e detectaram $L = 4,1$ e $C = 0,58$. Greene e Higman (2003) analisaram 65 estruturas proteicas distribuídas pelos nove tipos de dobramento mais abundantes, entre eles β/α barril como as β -glicosidases, obtendo L em torno de 4 e C entre 0,5 e 0,6. Bagler e Sinha (2005) analisaram 80 proteínas distribuídas igualmente pelas quatro classes estruturais de proteínas (α , β , $\alpha+\beta$ e α/β) obtendo valores similares, $L = 6,8$ e $C = 0,533$. Finalmente, Atilgan *et al.* (2004) empregando um conjunto de 595 proteínas não homólogas observaram que em média $L/L_{\text{randômico}} = 1,8$ e $C/C_{\text{randômico}} = 13$, também classificando estas proteínas como redes de “mundo pequeno”. Aliás, no caso da rede estrutural de Sfbglic, $L/L_{\text{randômico}} = 1,8$ e $C/C_{\text{randômico}} = 29$. Deste modo, com base no exposto acima, nota-se que a rede estrutural de Sfbglic comporta-se como uma rede proteica típica.

Outro aspecto analisado nos trabalhos citados acima foi a chamada “curva de graus”, ou seja, o gráfico que correlaciona a probabilidade de um determinado resíduo possuir certo número k de contatos *versus* o número de contatos k . Greene e Higman (2003), Atilgan *et al.* (2004) e Bagler e Sinha (2005) notaram que este gráfico segue uma distribuição em formato de sino, indicativo de uma distribuição de Poisson. Porém, uma análise apenas das interações de longo alcance (entre resíduos separados por mais de 10 posições da estrutura primária) mostrou uma distribuição semelhante a uma curva exponencial. Atilgan *et al.* (2004) e Greene e Higman (2003) observam que o ajuste a uma função exponencial não é perfeito, pois há um número máximo de contatos que um resíduo de aminoácido pode formar devido à impedimentos estéricos. Coerentemente, a “curva de graus” elaborada com o “software” AminoNet 2.0 para Sfbglic seguiu o mesmo padrão discutido acima, distribuição em forma de sino para todas as

interações e uma distribuição semelhante a uma curva exponencial para interações de longo alcance (Figura 4).

Figura 4 – Curva de distribuição de graus para rede estrutural de SfbgI. **A** – Curva considerando todas as interações entre resíduos de SfbgI. **B** – Curva considerando apenas as interações entre resíduos de SfbgI separados por mais de 10 posições na sua estrutura primária (interações de longo alcance).

Este tipo de distribuição na “curva de graus” sugere que as redes estruturais proteicas teriam propriedades de redes “sem escala” (scale free), nas quais um número muito pequeno de nós altamente conectados (hubs) é responsável pela determinação dos caminhos “mais curtos” entre os demais nós da rede. Indo mais além, Greene e Higman (2003) e Altigan *et al.* (2004) sugerem que a existência de resíduos centrais (hubs) garantiriam robustez a estrutura proteica que não seria significativamente alterada por mutações randômicas que não atinjam os “resíduos hubs”.

Deste modo, considerando que SfbgI comporta-se como um representante típico das redes estruturais proteicas, optou-se pela identificação dos seus resíduos centrais (“resíduos hubs”). Para tal, foi empregado o método da perturbação de L apresentado em del Sol *et al.* (2006). De fato, a remoção de somente um pequeno número de resíduos (11 em 473) afetou a conectividade da rede aumentando significativamente (mais do que 3 vezes o desvio padrão da média; $z_k \geq 3$) o valor da via mínima média L da rede de SfbgI, enquanto a remoção da maior parte dos resíduos e seus contatos correspondentes não alterou significativamente L (Figura 5).

Figura 5 – Distribuição dos efeitos de perturbação de L causados pela remoção de resíduos da rede estrutural de SfbgI.

Esta curva de distribuição dos efeitos de perturbação de L é muito semelhante àquela observada por del Sol *et al.* (2006) que empregou um conjunto de sete famílias proteicas distintas em sua análise, sugerindo mais uma vez que SfbgI comporta-se como um representante típico das redes estruturais proteicas (Figura 6).

Figura 6 - Distribuição dos efeitos de perturbação de L causados pela remoção de resíduos da rede estrutural proteica. **A** – Amostra de proteínas pertencentes a sete famílias distintas. **B** – Sfβgli. Dados do painel A foram extraídos de del Sol *et al.* (2006).

Os 11 resíduos que quando retirados aumentaram significativamente L, chamados “resíduos hubs”, estão distribuídos ao longo da estrutura de Sfβgli, concentrando-se na metade C-terminal, com exceção de R97 (Tabela 1; Figura 7). Dentre estes resíduos, R97, E399 e Y331 já foram caracterizados como resíduos do sítio ativo de Sfβgli. Sendo, que a arginina e tirosina modulam o pK_a de E399, o nucleófilo catalítico de Sfβgli (Marana *et al.*, 2003). Todos os demais resíduos encontram-se fora do sítio ativo.

Tabela 1: Resíduos cuja remoção causou as maiores alterações em L e seus respectivos valores de z_k

Resíduo (hubs)	z_k
TYR (Y) 331	2,92
TYR (Y) 420	3,39
ASN (N) 249	3,58
SER (S) 247	3,63
PHE (F) 334	3,65
LYS (K) 366	4,08
THR (T) 245	4,20
GLU (E) 399	4,21
SER (S) 358	5,11
PHE (F) 251	6,07
ARG (R) 97	6,42

Figura 7 – Mapeamento dos “resíduos hubs” na estrutura terciária de Sfβgli. Os “resíduos hubs” estão representados por esferas vermelhas posicionadas sobre seus Cα.

A fim de observar se haveria alguma particularidade nestes resíduos, nove outras β -glicosidases da família GH1 que possuíam estrutura cristalográfica determinada e pertenciam a diferentes grupos de seres vivos (Bacteria, Archaea, Plantas e Animais) foram também analisadas como redes e seus “resíduos hubs” identificados pelo mesmo método usado para Sf β gli (Tabela 2).

Tabela 2 – Parâmetros das redes estruturais de β -glicosidases da família GH1

β -glicosidase	L	L _r	C	C _r
1NP2	4,3	2,4	0,48	0,028
1UG6	4,3	2,4	0,48	0,028
1EI4	4,5	2,4	0,49	0,027
2O9R	4,5	2,4	0,49	0,027
1E56	4,6	2,5	0,49	0,025
1V03	4,6	2,5	0,49	0,025
1E6S	4,6	2,5	0,48	0,025
3GNO	4,6	2,5	0,49	0,025
Sf β gli	4,8	2,6	0,48	0,022
1WCG	4,4	2,4	0,49	0,026
1VFF	4,4	2,4	0,49	0,026

As β -glicosidases estão identificadas por seus códigos no PDB Data Bank, sendo originárias de Bactérias: 1NP2, *Thermus nonproteolyticus*; 1UG6, *Thermus thermophilus*; 1EI4, *Paenibacillus polymyxa* A; 2O9R, *Paenibacillus polymyxa* B; Eucariotos: 1E56, *Zea mays* (milho); 1V03, *Sorghum bicolor* (sorgo); 1E6S, *Sinapis alba* (trevo); 3GNO, *Oryza sativa* (arroz); tr, *Spodoptera frugiperda* (lagarta); 1WCG, *Brevicorine brassicae* (pulgão, afídeo); Archea: 1VFF, *Pyrococcus horikoshii*.

Em seguida, foram elaborados um alinhamento de sequência de aminoácidos e uma comparação estrutural de Sf β gli e as outras β -glicosidases da família GH1 (Figuras 8 e 9).

O alinhamento de sequências revela que três dos “resíduos hubs” das β -glicosidases, R97, Y331 e W444 (numeração em Sf β gli), são altamente conservados, estão no interior do sítio e participam do mecanismo de catálise e ligação do substrato (Marana *et al.*, 2001 e 2003). Esta observação é coerente com o trabalho de Amitai *et al.* (2004) que notou que “resíduos hubs” têm alta probabilidade de ocorrerem no sítio ativo de enzimas. Os demais “resíduos hubs” das β -glicosidases, apesar de ocorrerem em regiões semelhantes da estrutura primária, não coincidem com resíduos envolvidos na ligação do substrato ou catálise (Figura 9).

Figura 8 – Representação da rede estrutural de β -glicosidases da família GH1. A linha preta no perímetro do círculo representa cada uma das posições da estrutura primária da proteína, enquanto que as linhas cinza representam contatos entre resíduos de aminoácidos. **A** - 1E14, *Paenibacillus polymyxa* A; **B** - 1E6S, *Sinapis alba* (trevo); **C** - 1E56, *Zea mays* (milho); **D** - 1NP2, *Thermus nonproteolyticus*; **E** - 1UG6, *Thermus thermophilus*; **F** - 1V03, *Sorghum bicolor* (sorgo); **G** - 1VFF, *Pyrococcus horikoshii*. **H** - 1WCG, *Brevicorine brassicae* (pulgão, afídeo); **I** - 2O9R, *Paenibacillus polymyxa* B; **J** - 3GNO, *Oryza sativa* (arroz); **K** - *Spodoptera frugiperda* (lagarta). Representação obtida com o software AminoNet 2.0.

O exemplo mais interessante de coincidência em posicionamento são os resíduos T245, S247, N249 e F251, que apesar de não serem conservados quanto ao tipo de aminoácido, ocorrem em uma região de alta conservação dos “resíduos hubs” nas β -glicosidases. De fato, a conservação estrutural destes resíduos, bem como dos demais “resíduos hubs” em β -glicosidases pode ser observada na Figura 10. Em Sf β gli estes quatro resíduos ocorrem alternadamente em uma fita β que forma o barril central da estrutura desta proteína, partindo do sua base para o topo (Figuras 7 e 10). A ocorrência alternada em uma fita β indica que as cadeias laterais destes “resíduos hubs” sempre estão voltadas para a mesma face da fita, no caso o interior proteico. Nas demais β -glicosidases, embora o número de “resíduos hubs” nesta região seja variável, sempre ocorre pelo menos um deles (Figuras 9 e 10). A comparação estrutural também mostra que os “resíduos hubs” tendem a ocorrer em regiões equivalentes da estrutura das β -glicosidases.

Considerando que as β -glicosidases comparadas provêm de organismos muito distintos e estas enzimas guardam 6,5% de identidade e 13% de similaridade geral, a semelhança dos parâmetros da rede estrutural (Tabela 2), o alinhamento de posicionamento dos “resíduos hubs” nas sequências destas β -glicosidases (Figura 9) e a similaridade da distribuição estrutural dos seus “resíduos hubs” (Figura 10) mostra que o conjunto de contatos (interações entre resíduos) e as redes estruturais destas enzimas são muito semelhantes. Assim, se Sf β gli apresentava-se como um exemplar típico frente a grandes populações de redes estruturais de proteínas diversas, Sf β gli também mostra-se um exemplar representativo dentre as β -glicosidases da família GH1.

```

1NP2|Q9L794|Thermus      -----
1UG6|Q53W75|Thermus      -----
1EI4|P22073|APaenibacillus -----
1E56|P49235|Zea          MAPLLAAAMNHAAAHPLGRSHLVGPNNESFSRHHLPPSSSPQSSKRRCNLS 50
1V03|Q41290|Sorghum      MALLLASAINHTAHPAGLRSH---PNNESFSRHHLCPSSPQNISKRRSNLS 47
1E6S|P29736|Sinapis      -----DEEITCQEN 9
3GNO|Q8L7J2|Oryza        -----MGRIKSSSGRCSTARLEAVAV 21
tr|O61594|Spodoptera     -----MKLLVVLS 8
1WCG|Q95X01|Brevicoryne -----
1VFF|O58104|Pyrococcus   -----

```

```

1NP2|Q9L794|Thermus      -----MTEN----AEKFLWGVATSAYQIEGATQE 25
1UG6|Q53W75|Thermus      -----MTEN----AEKFLWGVATSAYQIEGATQE 25
1EI4|P22073|APaenibacillus -----MTIFQ--FPQDFMWGTATAAYQIEGAYQE 27
1E56|P49235|Zea          FTTRSARVGSQ-NGVQMLSPSEIPQRDWFPSDFTFGAATSAYQIEGAWNE 99
1V03|Q41290|Sorghum      FRPRAQTISSESAGIHRSLPWEIPRRDWFPPSFLFGAATSAYQIEGAWNE 97
1E6S|P29736|Sinapis      LPFTCGNT-----DALNSSSFSSDFI[REDACTED]GVASSAYQIEGTI-- 44
3GNO|Q8L7J2|Oryza        LVVVFGVASSSLRGCIQQSGGGLTRGSFPEGFVFGTASAAYQYEGAVKE 71
tr|O61594|Spodoptera     LVAVACNAS-----IVRQQRFPDDFLFGTATASYQIEGAWDE 46
1WCG|Q95X01|Brevicoryne -----MDYKFPKDFMFGTSTASYQIEGGWNE 26
1VFF|O58104|Pyrococcus   -----MPLKFPEMFLFGTATSSHQIEGNR- 25

```

. * :*.:::.* **

```

1NP2|Q9L794|Thermus      DGRGPSIWDTFARRPGAIRD--GSTGEPAC[REDACTED]DHYHRYEEDIALMQSLGVGV 73
1UG6|Q53W75|Thermus      DGRGPSIWDAFAQRPGAIRD--GSTGEPAC[REDACTED]DHYRRYEEDIALMQSLGVRA 73
1EI4|P22073|APaenibacillus DGRGLSIWDTFAHTPGKVFN--GDNGNVACDSYHRYEEDIRLMKELGIRT 75
1E56|P49235|Zea          DGKGESNWDHFCHNHPERIL-DGSNSDIGANSYHMYKTDVRLKEMGMDA 148
1V03|Q41290|Sorghum      DGKGPSTWDHFCHNFPWIV-DRSNGDVAADSYHMYAEDVRLKEMGMDA 146
1E6S|P29736|Sinapis      -GRGLNIWDGFTHRYPNKS GPDHGNGDTTCDSFSYWQKDIDVLDELNATG 93
3GNO|Q8L7J2|Oryza        DGRGQTIWDTFAHTFGKIT--DFSNAADVAVDQYHRFEEDIQLMADMGMDA 119
tr|O61594|Spodoptera     DGKGENIWDYMVHNTPEVIR-DLSNGDIAADSYHNYKRDVEMMRELGLDA 95
1WCG|Q95X01|Brevicoryne DGKGENIWDRLVHTSPEVIK-DGTNGDIACDSYHKYKEDVAIIKDLNLKF 75
1VFF|O58104|Pyrococcus   ----WNDWWYYEQIGKLPYR----SGKACNHWELYRDDIQLMTSLGYNA 66

```

. * : . : : * : : . .

1NP2 Q9L794 Thermus	YRFSVAWPRILPEG--RGRINPKGLAFYDRLVDRLLAAGITPFLTLYHWD	121
1UG6 Q53W75 Thermus	YRFSVAWPRILPEG--RGRINPKGLAFYDRLVDRLLASGITPFLTLYHWD	121
1EI4 P22073 APaenibacillus	YRFSVSWPRIFPNG--DGEVNQEGLDYYHRVVDLLNDNGIEPFCTLYHWD	123
1E56 P49235 Zea	YRFSISWPRILPKGTKEGGINPDGIKYYRNLINLLLENGIEPYVTIFHWD	198
1V03 Q41290 Sorghum	YRFSISWPRILPKGTLAGGINEKGV EY YNKLIDLLLENGIEPYITIFHWD	196
1E6S P29736 Sinapis	YRFSIAWSRIIPRGKRSRGVNEKGIDYYHGLISGLIKKGITPFVTLFHW	143
3GNO Q8L7J2 Oryza	YRFSIAWSRIYPNG--VGQVNQAGIDHYNKLIDALLAKGIQPYVTLYHWD	167
tr O61594 Spodoptera	YRFSLSWARILPTG-MANEVNPAGIAFYNNYIDEMLKYNITPLITLYHWD	144
1WCG Q95X01 Brevicoryne	YRFSISWARIAPSG-VMNSLEPKGIAYNNLINELIKNDIIPLVTMHWD	124
1VFF O58104 Pyrococcus	YRFSIEWSRLFP EE---NKFNE DAFMKYREIIDLLLTRGITPLVTLHHFT	113
	****: *.*: * .: .. * :. : .: * *:.*:	

1NP2 Q9L794 Thermus	LPQALEDRG-GWRSRET---AFAFAEYAEAVARALADRVPFFATLNEEWC	167
1UG6 Q53W75 Thermus	LPLALEERG-GWRSRET---AFAFAEYAEAVARALADRVPFFATLNEEWC	167
1EI4 P22073 APaenibacillus	LPQALQDAG-GWGNRRT---IQAFVQFAETMFREFHGKIQH WLT FNEEWC	169
1E56 P49235 Zea	VPQALEEKYGGFLDKSHKSIVEDYTYFAKVCFDNFGDKVKNWLT FNEEQT	248
1V03 Q41290 Sorghum	TPQALVDAYGGFLDEED---YKDYTDFAKVCFEKFGKTVKNWLT FNEEET	243
1E6S P29736 Sinapis	LPQTLQDEYEGFLDPQI---IDDFKDYADLCFEEFGDSVKYWLTINQIYS	190
3GNO Q8L7J2 Oryza	LPQALEDKYKGLDRQI---VDDFAAYAETCFREFGDRVKHWITLNEEHT	214
tr O61594 Spodoptera	LPQKLQELG-GFANPLIS---DWFEDYARVVFENFGDRVKMFIT FNEERE	190
1WCG Q95X01 Brevicoryne	LPQYLQDLG-GWVNPIMS---DYFKEYARVLFYFGDRVKWWIT FNEEIA	170
1VFF O58104 Pyrococcus	SPLWFMKKG-GLREEN---LKHWEKYIEKVAELL-EKVKL VAT FNEEMV	158
	* : . *: : : : *:.*:	

1NP2 Q9L794 Thermus	SAFLGHWTGEHAPG-----LRNLEAALRAAHLLLGHGLAVEA	205
1UG6 Q53W75 Thermus	SAFLGHWTGEHAPG-----LRNLEAALRAAHLLLGHGLAVEA	205
1EI4 P22073 APaenibacillus	IAFLSNMLGVHAPG-----LTNLQTAIDVGHHLLVAHGLSVRR	207
1E56 P49235 Zea	FTSFSYGTGVFAPGRCS PGLDCAYPTGNSLVEPYTAGHNILLAHAEAVDL	298
1V03 Q41290 Sorghum	FCSVSYGTGV LAPGRCS PGVSCAVPTGNSLSEPYIVAHNLLRAHAETVDI	293
1E6S P29736 Sinapis	VPTRGYGSALDAPGRCSPTVDPSCYAGNSSTEPYIVAHHQLLAHAKVVDL	240
3GNO Q8L7J2 Oryza	VAIQGYDAGLQAPGRCSVLLHLYCKAGNSGTEPYVVAHHFILAHAAAASI	264
tr O61594 Spodoptera	ICFEGYGSATKAP-----ILNATAMGAYLCAKNLVT AHAKAYYL	229
1WCG Q95X01 Brevicoryne	VC-KGYSIKAYAP-----NLNLKTTGHYLAGHTQLIAHGKAYRL	208
1VFF O58104 Pyrococcus	YVMMGYLTAYWEP-----FIRSPFKAFKVAANLLKAHAIA YEL	196
	. . * . : *	

1NP2 Q9L794 Thermus	LRAAG---AR-RVGIVL	NFAP--AYGEDP-EAVDVADR---YHNRYFLDP	245
1UG6 Q53W75 Thermus	LRAAG---AR-RVGIVL	NFAP--AYGEDP-EAVDVADR---YHNRFFLDP	245
1EI4 P22073 APaenibacillus	FRELG---TSGQIGIAP	NVSWAVPYSTSE-EDKAACARTISLHSDWFLQP	253
1E56 P49235 Zea	YNKHY-KRDDTRIGLAF	DVMGRVPYGTSTF-LDKQAEERSWDINLGFLEP	346
1V03 Q41290 Sorghum	YNKYH-KGADGRIGLALNV	FCRVPYTNTF-LDQQAQERSMDKCLGFLEP	341
1E6S P29736 Sinapis	YRKNY-THQGGKIGPTM	ITRWFPLPYNDTDRHSIAATERMKEFFLGWFMGP	289
3GNO Q8L7J2 Oryza	YRTKYKATQNGQLGIAFDVMW	FEPMSNTT-IDIEAAKRAQEFQLGWFADP	313
tr O61594 Spodoptera	YDREFRPVQGGQCGIT	ISVNWFGPATPTPEDEMAAELRR-QGEWGIYAFP	278
1WCG Q95X01 Brevicoryne	YEEMFKPTQNGKISISIS	SGVFFMPKNAESDDDIETAERANQFE	258
1VFF O58104 Pyrococcus	LHGKF-----KVGIVKNI	FIILPASDKE-RDRKAAEKADNLFNWHFLDA	239
	:	.	.

1NP2 Q9L794 Thermus	IILG--RGYPESPFQDPPPAP-----	IILSRDLEAIARPLDPLGLV	281
1UG6 Q53W75 Thermus	IILG--KGYPESPFQDPPPVP-----	IILSRDLELVARPLDPLGLV	281
1EI4 P22073 APaenibacillus	IYQ--GSYPQFLVDWFQEG-----	ATVPIQDGDMDIIGEPIDMIGI	293
1E56 P49235 Zea	VVR--GDYPFSMRSLARER-----	LPFFKDEQKEKLAGSYNMLGL	384
1V03 Q41290 Sorghum	VVR--GDYPFSMRVSARDR-----	VPYFKEKEQEKLVSYSYDMIGI	379
1E6S P29736 Sinapis	LTN--GTYPQIMIDTVGER-----	LPSFSPEESNLVKGSYDFLGL	327
3GNO Q8L7J2 Oryza	FFF--GDYPATMRARVGER-----	LPRFTADEAAVVKGALDFVGI	351
tr O61594 Spodoptera	IFSAEGGFPKELSDKIAEKSAQQGY	PWSRLPEFTEEEKAFVRGTSDLIGV	328
1WCG Q95X01 Brevicoryne	VY--KGDYPPIMKKWVDQKSKEEGL	PWSKLPKFTKDEIKLLKGTADFYAL	306
1VFF O58104 Pyrococcus	IWS--GKYRGVFKT-----	YRIPQSDADFIGV	264
	.	:	::

1NP2 Q9L794 Thermus	NY	YAP--VRVAPGTGPLPVRYLPPEGPVTAMG-----	311
1UG6 Q53W75 Thermus	NY	YAP--VRVAPGTGTLPVRYLPPEGPATAMG-----	311
1EI4 P22073 APaenibacillus	NY	YSMSVNRFNPEAGFLQSEEINMGLPVTDT	325
1E56 P49235 Zea	NY	YSRFSKNIDISPNYSPVLNTDDAYASQEVNGPDGKPIG-----	428
1V03 Q41290 Sorghum	NY	YSTFSKHIDLSPNNSPVLNTDDAYASQETKGPDGNA	423
1E6S P29736 Sinapis	NY	YFTQYAQPSPNPVNSTNHTAMMDAGAKLTYINASGHYIGPL	377
3GNO Q8L7J2 Oryza	NE	YTYTRHNNTNIIGTLLNNTLADTGTVSLPFKNGKPIG-----	396
tr O61594 Spodoptera	NE	YTAFLVSATERKGPYPVPSLLDDVDTG	367
1WCG Q95X01 Brevicoryne	NE	YSRLVTF----GSDPNPNFNPDASYVTSVDEA	342
1VFF O58104 Pyrococcus	NY	YTASEVRHTWNPLKFFFEVKLADISERKTQMG-----	298
	*	:	*

1NP2 Q9L794 Thermus	----WEVYPEGLYHLLKRLGREVP--WPLYITENGGAAYPDLWTG----EA	351
1UG6 Q53W75 Thermus	----WEVYPEGLYHLLKRLGREVP--WPLYVTENGGAAYPDLWTG----EA	351
1EI4 P22073 APaenibacillus	----WPVESRGLYEVLHYLQKYGN--IDIYITENGACINDEVV----NG	364
1E56 P49235 Zea	GNPWIYMYPEGLKDLLMIMKNKYGN-PPIYITENGIGDVDTKETPLPMEA	477
1V03 Q41290 Sorghum	GNAWINMYPKGLHDILMTMKNKYGN-PPMYITENGMGDIDKGD--LPKPV	470
1E6S P29736 Sinapis	STDNIYYYPKGIYSVMDFKNKYYN-PLIYVTENGIS--TPGD--ENRNQ	422
3GNO Q8L7J2 Oryza	SIW-LYIVPRGMRLMNYVKERYNS-PPVYITENGMD--DSNNPFISIKD	442
tr O61594 Spodoptera	ASAWLTAPNSIHTALTHLNNLYNK-PVFYITENGWSTDESRE-----N	410
1WCG Q95X01 Brevicoryne	ETPYIIPVPEGLRKLLIWLKNEYGN-PQLLITENGYGDD-----G	381
1VFF O58104 Pyrococcus	----WSVYPKGIYMALKKASRYGRP---LYITENGIAT-----	329
	...: : . . :****	

1NP2 Q9L794 Thermus	VVEDPERVAYLEAHVEAALRARE-EGVDLRGYFVWSLMDNFEWAFGYTRR	400
1UG6 Q53W75 Thermus	VVEDPERVAYLEAHVEAALRARE-EGVDLRGYFVWSLMDNFEWAFGYTRR	400
1EI4 P22073 APaenibacillus	KVQDDRRISYMQQHLVQVHRTIH-DGLHVKGYMAWSLLDNFEWAEGYNMR	413
1E56 P49235 Zea	ALNDYKRLDYIQRHIATLKESID-LGSNVQGYFASWSLLDNFEWFAGFTER	526
1V03 Q41290 Sorghum	ALEDHTRLDYIQRHLSVLKQSID-LGADVRYGYFASWSLLDNFEWSSGYTER	519
1E6S P29736 Sinapis	SMLDYTRIDYLCSHLCFLNKVIKEKDVNVKGYLAWSALGDNYEFNKGFTVR	472
3GNO Q8L7J2 Oryza	ALKDSKRIKYHNDYLTNLAASIKEDGCDVRGYFASWSLLDNFEWAAGYSSR	492
tr O61594 Spodoptera	SLIDDDRIQYRASMESLLNCL-DDGINLKGYMAWSLMDNFEWMEGYIER	459
1WCG Q95X01 Brevicoryne	QLDDFEKISYLNKYNLNLATLQAMYEDKCNVIGYTVWSLLDNFEWFGYSIH	431
1VFF O58104 Pyrococcus	-LDDEWRVEFIIQHLQYVHKAIE-DGLDVRGYFYWSFMDNFEWKEGFGPR	377
	: * : : : .: ** *: *: *: *: *: *	

1NP2 Q9L794 Thermus	FGLYYVDFPS--QRRIPKRSALWYRERIARAQTGGSAH-----	436
1UG6 Q53W75 Thermus	FGLYYVDFPS--QRRIPKRSALWYRERIARAQT-----	431
1EI4 P22073 APaenibacillus	FGMIHVDVRT--QVRTPKESYYWYRNVVSNNWLETRR-----	448
1E56 P49235 Zea	YGIVYVDRNN-NCTRYMKESAKWLKEFN-TAKKP-SKKILTPA-----	566
1V03 Q41290 Sorghum	FGIVYVDREN-GCERTMKRSARWLQEFNGAAKKVENNKILTPAGQLN---	565
1E6S P29736 Sinapis	FGLSYIDWNN-VTDRDLKKSQWYQSFISP-----	501
3GNO Q8L7J2 Oryza	FGLYFVDYKD-NLKRYPKNSVQWFKALLKT-----	521
tr O61594 Spodoptera	FGLYEVDVSDPARTRTPRKA AFVYKHIIKHRVVDY EYEPETMVM TIDE GH	509
1WCG Q95X01 Brevicoryne	FGLVKIDFNDPQRTTRTKRESYTYFKNVSTGKP-----	464
1VFF O58104 Pyrococcus	FGLVEVDYQT--FERRPRKSAYVYGEIARSKEIKDELLKRYGLPELQL--	423
	: *: : : *	

Figura 9 – Alinhamento da estrutura primária de um grupo de β -glicosidases. Os resíduos centrais da rede estrutural (hubs) destas enzimas estão marcados em vermelho com fundo azul. Caixas azuis indicam aminoácidos envolvidos na ligação do substrato, enquanto que caixas pretas marcam aqueles envolvidos na hidrólise da ligação glicosídica. Asteriscos indicam posições totalmente conservadas, dois pontos indicam posições conservadas na maioria das sequências e pontos indicam posições conservadas em metade das sequências. As β -glicosidases estão identificadas por seus códigos no PDB Data Bank, sendo originárias de Bactérias: 1NP2, *Thermus nonproteolyticus*; 1UG6, *Thermus thermophilus*; 1EI4, *Paenibacillus polymyxa*; Eucariotos: 1E56, *Zea mays* (milho); 1V03, *Sorghum bicolor* (sorgo); 1E6S, *Sinapis alba* (trevo); 3GNO, *Oryza sativa* (arroz); tr, *Spodoptera frugiperda* (lagarta); 1WCG, *Brevicorine brassicae* (pulgão, afídeo); Archea: 1VFF, *Pyrococcus horikoshii*.

NP2

1UG6

1EI4

1E56

1VO3

1E6S

3GNO

1VFF

Figura 10 – Alinhamento da estrutura terciária de um grupo de β -glicosidases. Os “resíduos hubs” estão representados por esferas vermelhas posicionadas sobre seus $C\alpha$. As β -glicosidases estão identificadas por seus códigos no PDB Data Bank, sendo originárias de Bactérias: 1NP2, *Thermus nonproteolyticus*; 1UG6, *Thermus thermophilus*; 1EI4, *Paenibacillus polymyxa*; Eucariotos: 1E56, *Zea mays* (milho); 1V03, *Sorghum bicolor* (sorgo); 1E6S, *Sinapis alba* (trevo); 3GNO, *Oryza sativa* (arroz); SfbgI, *Spodoptera frugiperda* (lagarta); 1WCG, *Brevicorine brassicae* (pulgão, afídeo); Archea: 1VFF, *Pyrococcus horikoshii*.

Com base nesta conclusão, para estudar o papel dos “resíduos hubs” nas propriedades funcionais destas β -glicosidases, mutantes simples correspondente a cada “resíduo hub” de SfbgI foram construídos, exceto E399, pois sendo este o nucleófilo catalítico certamente a mutação levaria a inativação da enzima. Estes “resíduos hubs” foram todos substituídos por alanina (A) a fim de remover a maioria dos contatos de suas cadeias laterais, reduzindo o seu grau de conectividade. Alanina foi o resíduo de substituição escolhido, pois este não possui a cadeia lateral além do carbono β e também não causa alterações na conformação da cadeia principal (ao contrário da substituição por glicina e deleção de aminoácidos) (Cunningham e Wells, 1989). Para fins comparativos, além dos “resíduos hubs” outros 3 “resíduos não-hubs” (com $z_k = 0$) D84, L100 e L350 também foram alterados para alanina. Além disso, para aumentar a amostragem de mutantes com substituição em “resíduos não-hubs” foram incluídos na comparação dados previamente descritos na tese de doutorado de Lúcio M.F. Mendonça (2009; <http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/pt-br.php>) para as SfbgI mutantes Y56A, M57A, S198A, P344A, P346A e S349A. Assim, a Figura 11 traz a distribuição na estrutura de SfbgI dos “resíduos hubs” e “não hubs” analisados nesta presente tese.

Figura 11 – Distribuição dos resíduos “hubs” e “não-hubs” estudados nesta tese na estrutura de Sfgli. Resíduos “hubs” estão marcados em amarelo e “não-hubs” em azul. O resíduo E399, nucleófilo catalítico, está marcado em púrpura.

4.2 Construção dos mutantes a partir da sequência de SfbgIi clonada no vetor

pET46-EK/LIC

A sequência codificadora da SfbgIi selvagem clonada no vetor pET46-Ek/LIC (Novagen), serviu como molde das reações de mutagênese sítio-dirigida. Os 'primers' cujas sequências continham as mutações desejadas, foram usados nas reações de mutagênese com o kit QuikChange Site Directed Mutagenesis gerando os plasmídeos contendo os insertos codificadores das SfbgIi mutantes. A região contendo os insertos foi integralmente sequenciada com o uso de 4 'primers'. Após a reação de sequenciamento confirmou-se a presença das mutações de interesse e a manutenção da integridade do restante da sequência codificadora da SfbgIi sem alterações. A seguir mostra-se o cromatograma do sequenciamento das regiões contendo as mutações, bem como o alinhamento destas com a sequência codificadora da SfbgIi selvagem (Figura 12).

Figura 12 – Eletroferogramas com destaque (caixas) nas mutações introduzidas na sequência de DNA que codifica SfbgIi e resultam na troca de aminoácidos apresentada no alinhamento abaixo.

S247A

Q C G I T I **S** V N W F G P
 Sfβgli selvagem 661 CAGTGC GGTATTACCATC**AGT**GTAAACTGGTTTGGACCAG 700
 mutante S247A 231 CAGTGC GGTATTACCATC**GCT**GTAAACTGGTTTGGACCAG 253
 A

N249A

G I T I S V **N** W F G P A
 Sfβgli selvagem 667 GGTATTACCATCAGTGTAA**ACT**GGTTTGGACCAGCCAC 704
 mutante N249A 238 GGTATTACCATCAGTGTAG**CTG**GGTTTGGACCAGCCAC 275
 A

F251A

I S V N W **F** G P A T P
 Sfβgli selvagem 674 CCATCAGTGTAAACTGG**TTT**GGACCAGCCACGCCCAC 710
 Mutante F251A 223 CCATCAGTGTAAACTGG**GCT**GGACCAGCCACGCCCAC 259
 A

Y331A

N H **Y** T A F L V S A T E

Sfβgli selvagem 930 GAACCACT**AC**ACAGCATTCTGGTATCGGCGACTGAAC 961

Mutante Y331A 488 GAACCAC**GC**ACAGCATTCTGGTATCGGCGACTGAAC 525
A

F334A

G V N H Y T A **F** L V S A T E

Sfβgli selvagem 919 GGAGTGAACCACTACACAGCA**TT**CCTGGTATCGGCGACTGAAC 961

mutante F334A 198 GGAGTGAACCACTACACAGCA**GC**CCTGGTATCGGCGACTGAAC 240
A

L350A

Y P V P S **L** L D D V D

Sfβgli selvagem 934 CGTACCCTGTGCCATCT**TT**ACTGGATGATGTGGACAC 970

mutante L350A 510 CGTACCCTGTGCCATCT**GC**ACTGGATGATGTGGACAC 546
A

S358A

D D V D T G **S** W A D D

Sfβgli selvagem 994 GATGATGTGGACACTGGCT**TC**CTGGGCTGATGAT 1026

mutante S358A 414 GATGATGTGGACACTGGC**GC**CTGGGCTGATGAT 446

A

K366A

A D D S W L **K** S A S A W

Sfβgli selvagem 1018 GCTGATGATAGCTGGCTC**AG**TCTGCTTCAGCTTGGC 1052

mutante K366A 536 GCTGATGATAGCTGGCTC**CG**TCTGCTTCAGCTTGGC 600

A

Y420A

D D R I Q **Y** Y R A S M

Sfβgli selvagem 1181 GATGACAGGATCCAGT**ACT**ACAGGGCTTCCATGG 1206

mutante Y420A 120 GATGACAGGATCCAG**GC**CTACAGGGCTTCCATGG 155

A

4.3 - Expressão da proteína SfbgIi recombinante e purificação

Após a construção dos vetores para expressão das SfbgIi mutantes, passou-se para a etapa de produção destas enzimas. Inicialmente a SfbgIi selvagem foi usada para padronização do protocolo. O vetor que codifica esta enzima foi introduzido em bactérias BL21(DE3) e NovaBlueDE3 e as colônias transformadas empregadas para expressão. Inicialmente utilizou-se a estirpe BL21(DE3) para a expressão das SfbgIi, contudo ensaios enzimáticos com as proteínas recombinantes revelaram a presença de uma β -galactosidase endógena desta estirpe que era co-purificada e clivava os substratos utilizados nos ensaios. Para contornar este problema, optou-se por utilizar somente a estirpe NovaBlueDE3 para as expressões, pois esta possui uma mutação que leva à deleção de parte da sequência codificadora da β -galactosidase bacteriana. Essa deleção inativa a β -galactosidade endógena desta bactéria, evitando assim a interferência desta enzima endógena nos ensaios enzimáticos empregados neste trabalho.

Para averiguar a efetiva produção das proteínas SfbgIi selvagem e mutantes submeteu-se alíquotas de 30 μ l do cultivo bacteriano não-induzido e induzido à SDS-PAGE. O padrão de bandas obtidos nos géis e ensaios de Western blot confirmaram a expressão de SfbgIi selvagem e todos os mutantes produzidos (os géis e ensaios revelaram a presença de uma proteína cuja massa corresponde à da proteína SfbgIi ($\approx$ 56.000 Da) que não está presente no cultivo não-induzido (resultados não mostrados).

Após a verificação da expressão das proteínas recombinantes, procedeu-se à lise bacteriana. A seguir o lisado obtido foi submetido à purificação de SfbgIi por afinidade com a resina Ni-NTA agarose. Separou-se frações de 30 μ l de cada etapa da purificação e submeteu-se à eletroforese em SDS-PAGE 12% (Figura 13).

Embora tenha se observado a purificação das SfbgIi mutantes (raia eluição - EL), observa-se que parte da enzima recombinante permanece no “pellet” obtido após a lise e

centrifugação, indicando que grande quantidade da mesma encontra-se provavelmente em corpos de inclusão. Outra parcela significativa da enzima recombinante fica retida na resina mesmo após duas etapas de eluição com alta concentração de imidazol (300 mM), resultado observado para todas as S β gli mutantes purificadas. Ademais, realizou-se várias tentativas de expressão e de purificação em diferentes condições para os mutantes L100A e T245A, como por exemplo, temperaturas distintas de expressão à 30, 25, 20, 15°C, menor concentração de IPTG nas induções (0,5 e 0,1 mM), lise e extração da proteína recombinante com uso de tampão de lise contendo detergentes como NP40 0,5%, etanol 10%. Contudo estes mutantes permaneceram em corpos de inclusão, o que impossibilitou sua purificação e caracterização.

Como controle da pureza e procedência da atividade enzimática das S β gli mutantes purificadas (Figura 13), uma forma inativa de S β gli (que possui a mutação N329A) e que contém em seu N-terminal uma etiqueta de calmodulina ao invés das 6 histidinas também foi expressa e purificada sob as mesmas condições das demais enzimas. A amostra eluída do procedimento de purificação desta S β gli inativa foi denominada controle (CTL) e foi utilizada nos ensaios de atividade enzimática para verificar se outras proteínas com atividade β -glicosidásica eram co-purificadas com as S β gli mutantes. Pelo resultado obtido da purificação deste mutante inativo (N329A), inferiu-se que o protocolo de purificação foi eficiente, pois não se observou a presença de nenhuma proteína contaminantes após a etapa de eluição (Figura 13), bem como nos ensaios enzimáticos empregando os substratos p-nitrofenil β -glicosídeos e celobiose com amostras obtidas da eluição (EL) da purificação do mutante controle (CTL) não houve a clivagem dos substratos e detecção dos respectivos produtos mesmo após 48h de ensaio à 30° C.

Ao final dos experimentos de purificação de Sfβgli foi observada grande perda da atividade das enzimas eluídas em tampão contendo 300 mM imidazol. Para contornar este problema, foram introduzidos procedimentos de desalinização das frações oriundas da etapa de eluição de todos mutantes purificados. Contudo a recuperação da atividade após a diálise também não foi satisfatória (menor que 50%). Também não era incomum a precipitação proteica. No entanto o emprego de cromatografia de filtração em gel na coluna HiTrap desalting minimizou a perda da atividade enzimática, pois as recuperações superaram 90%.

Figura 13 - SDS-PAGE das diferentes etapas da purificação das Sfβgli selvagem e mutantes. Aplicou-se cerca de 30 µl de amostra de cada etapa: Amostra bacteriana **não induzida (ÑI)** e **induzida (I)**, fração solúvel do **lisado (Lis)** obtido após lise das bactérias induzidas e centrifugação, **pellet (PEL)** oriundo da centrifugação do extrato após a lise bacteriana, **'flow-through'(FT)** correspondendo às proteínas não retidas pela resina Ni-NTA agarose obtido durante a etapa de ligação na resina, **Lv 1** e **Lv 5**, primeira e última lavagens das proteínas não retidas pela resina, **eluição (EL)** – proteína eluída da resina após lavagens, incubação com tampão contendo 300 mM de imidazol e dessalinização, **resina (Res)** – proteína retida pela resina mesmo após a etapa eluição. A seta indica a proteína recombinante purificada (56 kDa). O gel de poliacrilamida 12% foi corado com Coomassie Blue R-250.

D84A **R97A** **L100A**

T245A **S247A** **N249A**

F251A **Y331A** **F334A**

L350A **S358A** **K366A**

4.4 – Avaliação da estrutura global das Sfβgli mutantes.

4.4.1 – Espectros de fluorescência intrínseca das Sfβgli mutantes

Considerando que foram introduzidas substituições de resíduos em posições que poderiam ser importantes no processo de dobramento de Sfβgli, fez-se necessária uma avaliação estrutural destas enzimas mutantes.

Dado o grande número de proteínas mutantes, uma abordagem viável foi a determinação dos espectros de fluorescência intrínseca de triptofano. Eventuais alterações estruturais drásticas e ausência de enovelamento podem ser evidenciados por deslocamento do espectro em direção à comprimentos de onda mais longos devido à

exposição dos resíduos de triptofano ao ambiente mais polar do solvente. Na estrutura de SfbgI distribuem-se 14 triptofanos, sendo assim os espectros de fluorescência são compostos da contribuição de resíduos em diferentes posições e ambientes. Isto por um lado permite uma visão global da proteína, porém por outro não revela informações de regiões específicas. Como visto na figura 14, os espectros das enzimas mutantes, exceto F251A, se sobrepõem à enzima selvagem revelando proteínas enoveladas. A mesma conclusão foi obtida para os mutantes em resíduos 'não hubs' (Y56A, M57A, S198A, P344A, P346A, S349A) previamente caracterizados na tese de Lúcio M. F. Mendonça (2009; <http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/en.php>). No caso de F251A, o deslocamento do espectro para região do vermelho sugere maior exposição dos triptofanos. Porém, como o deslocamento é parcial, é provável que a enzima siga enovelada, mas com uma estrutura significativamente diferente da selvagem. Em suma, as mutações introduzidas em resíduos 'hubs' ou 'não hubs' não impediram o enovelamento de SfbgI.

Atentos ao fato de que as enzimas mutantes poderiam estar enoveladas, mas com estrutura distinta da SfbgI selvagem, buscou-se outro ensaio que permitisse uma rápida caracterização estrutural das várias enzimas mutantes.

Figura 14 - Espectro de fluorescência intrínseca das Sfbgls selvagem e mutantes. O espectro da enzima selvagem está representado com linha vermelha, enquanto o dos mutantes está na linha azul. Sfbgls mutantes com trocas em “resíduos não hubs” estão grifadas.

4.4.2 Ensaios de apagamento de fluorescência das Sfbgli mutantes

A fim de verificar se as mutações dos aminoácidos centrais ('hubs') da rede estrutural de Sfbgli causariam alguma alteração conformacional nesta proteína, realizou-se ensaios de apagamento de fluorescência intrínseca por adição de acrilamida. Eventuais alterações de conformação da proteína, mesmo que em pequena escala, podem ocasionar mudanças na exposição e/ou acesso do supressor dos resíduos de triptofano. Deste modo, os experimentos de apagamento de fluorescência tendo pequenas moléculas como supressor, são uma ferramenta eficaz para sondar alterações estruturais decorrentes de mutações. Levando em consideração que há 14 resíduos de triptofano distribuídos pela proteína Sfbgli, tal experimento pode ser utilizado para averiguar globalmente a estrutura de Sfbgli evidenciando até mesmo pequenas alterações localizadas em regiões distantes do sítio ativo. Assim foram determinados os espectros de fluorescência intrínseca para Sfbgli selvagem e mutantes na presença de adições sucessivas do supressor acrilamida.

Considerando o ponto máximo de fluorescência (λ_{max}) observou-se que a redução da fluorescência expressa por F_0/F é diretamente proporcional à concentração do supressor acrilamida (Figura 15), sugerindo que o fenômeno de apagamento é resultado das colisões entre os triptofanos e a acrilamida (Eftink and Ghiron, 1976). A probabilidade destes encontros pode ser expressa pela constante Sterner-Volmer (K_{sv}) calculada a partir da inclinação dos gráficos F_0/F *versus* Q (concentração de acrilamida) (Figura 15). Aumentos ou diminuições nos valores de K_{sv} indicam alterações, maior ou menor exposição dos resíduos de triptofano, respectivamente.

Foram analisadas a enzima selvagem, 9 mutantes com trocas dos resíduos 'hubs' (todos mutantes construídos, exceto T245A) e 8 mutantes com substituições de resíduos 'não-hubs' (D84A• e L350A• produzidos neste trabalho e outros 6 já caracterizados na tese de doutoramento de Lúcio M. F. Mendonça (2009; <http://www.teses.usp.br/teses/>

disponiveis/46/46131/tde-10122009-143601/en.php). Foram calculadas as diferenças de K_{sv} entre a Sfbgli selvagem e mutantes (ΔK_{sv}) e seu respectivo desvio padrão ($\sigma = 0,5$). Assim, pelos valores de ΔK_{sv} encontrados (Tabela 3) apenas 5 Sfbgli mutantes, N249A, F251A, S358A (‘hubs’), M57A• e D84A• (‘não-hubs’), parecem ter uma estrutura terciária diferente da enzima selvagem, pois os K_{sv} para estes mutantes variaram em pelo menos 1 desvio padrão quando comparado à enzima selvagem.

Tabela 3 - Constantes de Stern-Volmer (K_{sv}) para o apagamento da fluorescência intrínseca da Sfbgli selvagem e mutantes por acrilamida.

Sfbgli	K_{sv}	ΔK_{sv}
Selvagem	3,29±0,02	0
Y56A•	3,3±0,1	0,01
M57A•	2,76±0,03	-0,53
<u>D84A•</u>	4,1±0,2	0,81
R97A	3,61±0,08	0,32
S198A•	3,3±0,1	0,01
S247A	4,01±0,07	0,72
N249A	2,94±0,06	-0,35
F251A	3,91±0,09	0,62
Y331A	3,2±0,1	-0,09
F334A	4,7±0,1	1,41
P344A•	3,38±0,03	0,09
P346A•	3,0±0,1	-0,29
S349A•	3,4±0,1	0,11
<u>L350A•</u>	3,1±0,1	-0,19
S358A	3,46±0,05	0,17
K366A	3,20±0,04	-0,09
Y420A	3,16±0,06	-0,13

• Mutantes com troca em resíduos ‘não-hubs’. Os dados dos mutantes em resíduos “não hubs” que não estão sublinhados foram retirados da tese de Lúcio M. F. Mendonça (2009; <http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/en.php>).

No caso do mutante F251A, observou-se a tendência da relação F_0/F versus Q tornar-se curva nas concentrações mais altas de acrilamida, revelando ocorrência de supressão de fluorescência pelo mecanismo estático, além do colisional (Eftink and

Ghiron, 1976). Este resultado também é indicação de alteração estrutural. Para o mutante F251A apenas o trecho inicial da curva, que mais se aproxima de uma reta, foi utilizado para o cálculo de K_{sv} . O aumento no valor de K_{sv} para os mutantes D84A•, S247A, F251A, F334A e diminuição para M57A• são indicativos de alteração estrutural com maior e menor exposição dos triptofanos, respectivamente. Os demais mutantes tiveram valores de K_{sv} muito similares ao da selvagem.

Inferiu-se que todas SfbgIi purificadas se encontraram enoveladas, pois o espectro de fluorescência intrínseca de todas se sobrepôs ao da enzima selvagem (Figura 14), sendo que as alterações de K_{sv} refletem apenas diferenças estruturais, mas não desnaturação.

Deve ser notado que as diferenças significativas de K_{sv} ocorrem nas enzimas mutantes com trocas de resíduos ‘hubs’ e ‘não hubs’. Logo, o papel dos resíduos na rede estrutural de SfbgIi não tem relação com as alterações estruturais. Apesar da alteração em K_{sv} refletir mudanças conformacionais, devido ao grande número de triptofanos não há como prever como e quais regiões da proteína foram alteradas. Apenas no caso do mutante F251A há indicação de aumento de exposição dos triptofanos pelo aumento do K_{sv} e pelo deslocamento do espectro de fluorescência em direção ao vermelho. É possível que nos demais mutantes que exibem aumento de K_{sv} , os triptofanos com aumento de exposição não contribuam majoritariamente com o espectro.

Em suma, estes resultados sugerem que as SfbgIi mutantes encontram-se enoveladas, embora a estrutura terciária de algumas delas não seja idêntica à da selvagem. Com base nesta conclusão, considerou-se válido avaliar as enzimas mutantes quanto à sua estabilidade estrutural e propriedades funcionais.

Figura 15 - Supressão de fluorescência intrínseca das S β gli selvagem e mutantes por acrilamida.

4.5 - Determinação dos parâmetros cinéticos das S β gli mutantes

Para avaliar o efeito das mutações em resíduos ‘hubs’ e ‘não hubs’ sobre a eficiência catalítica e preferência pelo substrato de S β gli, os parâmetros cinéticos (k_{cat} e K_m) para hidrólise de *p*-nitrofenil β -glicosídeos e de celobiose foram medidos (Figura 16 e Tabela 4). Os parâmetros cinéticos de S β gli mutantes com substituições em 8 resíduos que são ‘hubs’ foram determinados. O mutante Y331A com substituição em resíduo ‘hub’ apesar de apresentar-se solúvel e enovelado, conforme resultado inferido pelo espectro de fluorescência dos triptofanos (Figura 14), estava inativo. Já o mutante D84A•, por sua vez, apresentou atividade baixíssima frente aos *p*-nitrofenil β -glicosídeos e celobiose, o que permitiu apenas a análise da sua atividade relativa. Paralelamente, foram reunidos os parâmetros cinéticos para 6 S β gli mutantes com substituições em resíduos ‘não hubs’ (Tabela 5), os quais encontram-se depositadas na tese de doutoramento de Lúcio M. F. Mendonça (2009) (<http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/en.php>).

Para inferir o quanto a especificidade pelo substrato das S β gli mutantes difere em relação à enzima selvagem, comparou-se a eficiência catalítica (k_{cat}/K_m) sobre os diferentes substratos. Observou-se que as suas especificidades, avaliadas pelo k_{cat}/K_m relativo, são similares. Assim, nenhuma das mutações, seja em resíduo ‘hub’ ou ‘não hub’ exibiu especificidade alterada. Deve ser notado que há claro aumento da preferência por celobiose para os mutantes N249A e R97A, mas estas tendências não se confirmaram para os outros mutantes (Figura 16 e Tabela 4). Portanto, o papel do resíduo na rede estrutural de S β gli não está relacionado com a especificidade pelo substrato.

Tabela 4 - Parâmetros cinéticos para a hidrólise de *p*NPβ-glicosídeos pela Sfbglic selvagem e mutantes.

Enzima	Substrato	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m	k_{cat}/K_m relativo
Sfbglic selvagem	NPβglc	2,3±0,2	0,61±0,01	0,26±0,02	19,4
	NPβgal	10±1	0,16±0,01	0,016±0,002	1,2
	NPβfuco	0,98±0,07	1,35±0,03	1,3±0,1	100
	celo	10±1	0,82±0,06	0,08±0,01	6
D84A●	NPβglc	#	#	#	47,9
	NPβgal	#	#	#	13,9
	NPβfuco	#	#	#	100
	celo	#	#	#	32,8
R97A	NPβglc	2,8±0,3	$5,5 \cdot 10^{-6} \pm 0,2 \cdot 10^{-6}$	$2,0 \cdot 10^{-7} \pm 0,2 \cdot 10^{-7}$	13,7
	NPβgal	7,8±0,6	$9,0 \cdot 10^{-7} \pm 0,4 \cdot 10^{-7}$	$1,1 \cdot 10^{-7} \pm 0,1 \cdot 10^{-7}$	0,8
	NPβfuco	1,8±0,2	$30,0 \cdot 10^{-6} \pm 0,7 \cdot 10^{-6}$	$14,0 \cdot 10^{-6} \pm 0,04 \cdot 10^{-6}$	100
	celo	0,40±0,09	$1,6 \cdot 10^{-6} \pm 0,1 \cdot 10^{-6}$	$4,1 \cdot 10^{-6} \pm 1,0 \cdot 10^{-6}$	29,1
S247A	NPβglc	0,45±0,03	0,411±0,005	0,91±0,06	24,9
	NPβgal	7,7±0,2	0,19±0,03	0,025±0,004	0,7
	NPβfuco	0,46±0,03	1,69±0,03	3,7±0,2	100
	celo	3,1±0,1	0,7±0,1	0,21±0,04	5,8
N249A	NPβglc	0,40±0,06	0,36±0,01	0,9±0,1	23,1
	NPβgal*	-----	-----	0,000181±0,0000108	0,004
	NPβfuco	0,30±0,04	1,17±0,03	3,9±0,5	100
	celo	0,80±0,08	0,79±0,02	0,9±0,1	25,3
F251A	NPβglc	1,2±0,1	0,176±0,005	0,14±0,01	17,9
	NPβgal	-----	-----	0,0070±0,0002	0,9
	NPβfuco	1,2±0,1	1,04±0,02	0,82±0,06	100
	celo	4,9±0,6	0,32±0,01	0,066±0,008	8,1

(Tabela 4 continua na próxima página)

(continuação da tabela 4)

Enzima	Substrato	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m	k_{cat}/K_m relativo
Sfβgl selvagem	NPβglc	2,3±0,2	0,61±0,01	0,26±0,02	19,4
	NPβgal	10±1	0,16±0,01	0,016±0,002	1,2
	NPβfuco	0,98±0,07	1,35±0,03	1,3±0,1	100
	celo	10±1	0,82±0,06	0,08±0,01	6
F334A	NPβglc	1,1±0,1	0,76±0,02	0,69±0,06	57
	NPβgal	1,6±0,2	0,045±0,002	0,028±0,004	2,3
	NPβfuco	0,69±0,07	0,84±0,02	1,2±0,1	100
	celo	1,8±0,1	0,48±0,01	0,27±0,02	5,8
L350A●	NPβglc	0,98±0,03	0,800±0,007	0,82±0,03	24,8
	NPβgal	2,1±0,1	0,087±0,002	0,042±0,002	1,26
	NPβfuco	0,52±0,03	1,71±0,02	3,3±0,2	100
	celo	4,6±0,3	0,37±0,01	0,080±0,005	2,4
S358A	NPβglc	4,1±0,3	0,35±0,01	0,085±0,007	25,2
	NPβgal*	-----	-----	0,00289±0,00006	0,9
	NPβfuco	0,70±0,06	0,238±0,007	0,34±0,03	100
	celo	#	#	#	#
K366A	NPβglc	1,4±0,1	0,75±0,02	0,53±0,04	25
	NPβgal	8,6±0,6	0,117±0,004	0,013±0,001	0,6
	NPβfuco	1,29±0,04	2,76±0,03	2,14±0,07	100
	celo	5,4±0,8	0,45±0,03	0,08±0,01	3,9
Y420A	NPβglc	1,29±0,07	0,55±0,01	0,42±0,02	14,1
	NPβgal	3,2±0,2	0,106±0,004	0,033±0,002	1,1
	NPβfuco	0,92±0,09	2,79±0,07	3,0±0,3	100
	celo	4,7±0,9	0,51±0,03	0,11±0,02	3,6

Mutantes com substituições em resíduos que não são 'hubs' estão grifados e marcados com •. Mutantes com substituições em resíduos 'hubs' são apresentados sem qualquer marcação. Para o mutante em resíduo 'não hub' **D84A●** foram apresentados os dados de atividade relativa

Os parâmetros foram determinados utilizando-se no mínimo 17 concentrações diferentes de substratos e os ajustes dos dados na equação de Michaelis-Menten foram feitos através do Software Enzfitter (Leatherbarrow R.J. Elsevier – Biosoft – 1987).

* O gráfico de v versus $[S]$ apresentou-se linear, portanto estimou-se apenas k_{cat}/K_m a partir da inclinação da reta não sendo possível determinar os valores de K_m e k_{cat} separadamente.

Não foi possível a determinação

Tabela 5- Parâmetros cinéticos do estado estacionário para a hidrólise de pNPβ-glicosídeos pela Sfbglic selvagem e mutantes. Todas as Sfbglic mutantes apresentam substituição em resíduos que não são hubs.

Enzima	Substrato	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m	k_{cat}/K_m relativo
Sfbglic selvagem	NPβglc	0,74	0,2	0,26±0,01	26
	NPβgal	3	0,0	0,0092±0,0004	0,9
	NPβfuco	0,54	0,5	1,01±0,04	100
Y56A●	NPβglc	0,83±0,05	0,122±0,002	0,147±0,008	29
	NPβgal	2,8±0,1	0,0109±0,0001	0,0039±0,0002	0,8
	NPβfuco	0,72±0,03	0,359±0,003	0,49±0,02	100
M57A●			0,278		
	NPβglc	1,10±0,05	0,141±0,001	0,128±0,006	26
	NPβgal	2,9±0,1	0,0150±0,0002	0,0052±0,0002	1
	NPβfuco	0,66±0,04	0,3289±0,004	0,49±0,03	100
S198A●	NPβglc	0,54±0,02	0,099±0,001	0,182±0,008	29
	NPβgal	1,22±0,04	0,0120±0,0001	0,0098±0,0004	1,6
	NPβfuco	0,45±0,02	0,278±0,002	0,61±0,03	100
P344A●	NPβglc	0,88±0,05	0,227±0,003	0,25±0,01	25
	NPβgal	3,8±0,2	0,0199±0,0003	0,0052±0,0003	0,5
	NPβfuco	0,44±0,03	0,459±0,005	1,04±0,07	100
P346A●	NPβglc	1,64±0,07	0,0374±0,0006	0,023±0,001	20
	NPβgal	5,6±0,4	0,0036±0,0001	0,00063±0,00005	0,5
	NPβfuco	0,8±0,01	0,097±0,003	0,11±0,02	100
S349A●	NPβglc	1,08±0,04	0,273±0,003	0,252±0,009	28
	NPβgal	3,7±0,1	0,0544±0,0005	0,0144±0,0004	1,6
	NPβfuco	0,51±0,2	0,455±0,003	0,87±0,03	100

Os parâmetros foram determinados utilizando-se no mínimo 10 concentrações diferentes de substratos e os ajustes dos dados na equação de Michaelis-Menten foram feitos através do Software Enzfitter (Leatherbarrow R.J. Elsevier – Biosoft – 1987).

* dados obtidos na tese de doutoramento de Lúcio M. F. Mendonça (2009; <http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/en.php>).

Figura 15 - Efeito da concentração do substrato na velocidade inicial (v_o) de reações de hidrólise catalisadas pela enzima S β glu selvagem e mutantes frente aos diferentes substratos

Na tentativa de observar alguma tendência de variação de K_m entre os mutantes foi feita uma análise global dos valores assumidos por este parâmetro para todos os substratos testados. Observou-se que entre as Sf β gli com mutações em resíduos “hubs” há uma maior ocorrência de diminuições do K_m , com 22 casos contra apenas 6 aumentos (Tabela 6; Figura 17). Deve ser notado que não foram incluídos nesta análise os dados de K_m do substrato NP β gal para as enzimas mutantes F251A, N249A e S358A. Nestes casos há provável aumento de K_m , pois as concentrações de substrato empregadas nos ensaios não superaram K_m e deste modo o gráfico $v_0 \times [S]$ mostrou-se linear para estes três mutantes (Tabela 4; Figura 16). Combinando estes dados, seriam observados 22 casos de diminuição e 9 casos de aumento de K_m . Já entre as mutações em resíduos “não-hubs”, há maior equilíbrio, com 10 casos de diminuição e 11 de aumento.

Contudo, as faixas de variação de K_m nestas duas populações não diferem marcadamente. Dentre as mutações em resíduos “não-hubs” as variações vão de uma diminuição de 4,8 vezes a um aumento de 2,2 vezes. Já entre as mutações em resíduos “hubs” as variações vão de uma diminuição de 6,3 vezes a um aumento de 1,9 vezes, excluindo o substrato celobiose que consta apenas para esta população. No caso do substrato celobiose há dois casos de redução mais drástica, 25 e 12 vezes, para os mutantes R97A e N249A, respectivamente, enquanto que os demais (6 casos) se inserem na faixa de variações descrita acima. Porém, não há dados para comparação com a população de mutantes em resíduos “não-hubs” (Tabela 6; Figura 17).

Portanto, independentemente do substrato analisado, as mutações em resíduos “hubs” e “não-hubs” causam efeitos similares e de baixa amplitude no K_m de Sf β gli.

Tabela 6 – Variação do K_m para SfbglI mutantes em relação à enzima selvagem

	<u>Y56A</u>	<u>M57A</u>	<u>R97A</u>	<u>S198A</u>	<u>S247A</u>	<u>N249A</u>	<u>F251A</u>	<u>F334A</u>	<u>P344A</u>	<u>P346A</u>	<u>S349A</u>	<u>L350A</u>	<u>S358A</u>	<u>K366A</u>	<u>Y420A</u>
NPβglc	1,1	1,5	1,2	1,4	5,1	5,8	1,9	2,1	1,2	2,2	1,5	2,3	1,8	1,6	1,8
NPβgal	1,1	1,0	1,3	2,5	1,3	*	*	6,3	1,3	1,9	1,2	4,8	*	1,2	3,1
NPβfuco	1,3	1,2	1,8	1,2	2,1	3,3	1,2	1,4	1,2	1,5	1,1	1,9	1,4	1,3	1,1
celo	-	-	25	-	3,2	12	2,0	5,6	-	-	-	2,2	-	1,9	2,1

Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifados.

Fundo azul indica redução; Fundo vermelho indica aumento; Fundo cinza indica ausência de variação.

- K_m não determinado para a enzima mutante

* gráfico $v_0 \times [S]$ para a enzima mutante é linear; não é possível determinar K_m

Figura 17 - Variação do K_m para Sfβgli mutantes em relação à enzima selvagem. Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifados.

Para que as variações de K_m assumam sentido físico mais concreto, foram calculados os efeitos das mutações nas variações da energia de associação para formação do complexo ES ($\Delta\Delta G_b$) (Figura 18). Nota-se que os maiores valores de $\Delta\Delta G_b$ tendem a ocorrer entre os mutantes com substituições em resíduos “hubs”. Contudo, observa-se que dentre as duas populações o efeito mutacional circunscreve-se na grande maioria dos casos à faixa de $\pm 3,0$ kJ/mol, não indo além de 4,6 kJ/mol mesmo no caso mais extremo (mutante F334A).

Assim, considerando que ligações de hidrogênio envolvendo participantes não carregados em enzimas e seus ligantes apresentam energia na faixa de 2,0 a 6,0 kJ/mol (Fersht *et al.*, 1985) e que as ligações de hidrogênio envolvendo resíduos do sítio ativo e hidroxilas do glicone do substrato estão na faixa de 3,0 a 33 kJ/mol (Badiéyan *et al.*, 2012; Marana *et al.*, 2002; Namchuk e Withers, 1995), pode-se assumir que caso as mutações tenham provocado rompimento, formação, enfraquecimento ou fortalecimento de interações não-covalentes entre a enzima e o substrato, o balanço destes efeitos foi no máximo igual a uma ligação de hidrogênio, ficando na maioria dos casos analisados abaixo deste limite.

Portanto, os efeitos das mutações nos resíduos “hubs” e “não-hubs” não alteraram significativamente a afinidade de Sf β gli pelos *p*-nitrofenil β -glicosídeos.

Figura 18 – Efeito mutacional sobre a energia de associação da enzima e substrato na formação do complexo ES ($\Delta\Delta G_b$). Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifadas. As linhas tracejadas representam o limite inferior de energia esperado para uma ligação de hidrogênio. Os erros relativos estão na faixa dos 7 a 17%.

Uma análise global do efeito mutacional sobre a constante de velocidade de hidrólise dos substratos (k_{cat}) completou o quadro de comparações dos parâmetros cinéticos (Tabela 7; Figura 19). Notou-se que em ambas as populações de mutantes predominam diminuições do k_{cat} . No caso das mutações em resíduos “hubs” ocorrem 19 casos de diminuição contra 6 ocorrências de aumento do k_{cat} . Já dentre os mutantes com alterações em resíduos “não-hubs”, ocorrem 17 diminuições *versus* 5 aumentos. As faixas de variação de k_{cat} também são semelhantes nas duas populações de mutantes, indo de uma diminuição de 5,7 vezes a um aumento de 2,1 vezes dentre os mutantes para resíduos “hubs” e oscilando entre diminuição de 7,8 vezes a aumento de 1,9 vezes dentre os mutantes para resíduos “não-hubs”. Nota-se também que as variações de k_{cat} para o substrato celobiose, determinado apenas para os mutantes com alterações nos resíduos “hubs”, são reduções em todos os casos, atingindo no máximo 2,6 vezes.

O mutante R97A foi excluído da determinação desta faixa de amplitude de variação de k_{cat} , pois foi demonstrado anteriormente que R97 modula através de interações não-covalentes o pK_a do nucleófilo catalítico E399 (Marana *et al.*, 2003). Logo, esta arginina tem papel essencial na catálise, sendo que os minúsculos valores de k_{cat} observados para R97A sugerem que seu o papel catalítico é mais preponderante no efeito mutacional do que seu posicionamento como “hub” na rede estrutural.

Portanto, as mutações em resíduos “hubs” e “não hubs” causam efeitos similares no k_{cat} da hidrólise dos *p*-nitrofenil β -glicosídeos.

Tabela 7 – Variação do k_{cat} para Sf β gli mutantes em relação à enzima selvagem

	<u>Y56A</u>	<u>M57A</u>	<u>R97A</u>	<u>S198A</u>	<u>S247A</u>	<u>N249A</u>	<u>F251A</u>	<u>F334A</u>	<u>P344A</u>	<u>P346A</u>	<u>S349A</u>	<u>L350A</u>	<u>S358A</u>	<u>K366A</u>	<u>Y420A</u>
NP β glc	1,6	1,3	110909	1,9	1,5	1,7	3,5	1,2	1,2	5,1	1,4	1,3	1,7	1,2	1,1
NP β gal	2,6	1,9	177777	2,3	1,2	-	-	3,6	1,4	7,8	1,9	1,8	-	1,4	1,5
NP β fuco	1,5	1,6	45000	1,9	1,3	1,2	1,3	1,6	1,2	5,6	1,2	1,3	5,7	2,0	2,1
celo	-	-	512500	-	1,2	1,0	2,6	1,7	-	-	-	2,2	-	1,8	1,6

Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifados.

Fundo azul indica redução; Fundo vermelho indica aumento; Fundo cinza indica ausência de variação.

- k_{cat} não determinado para a enzima mutante

Figura 18 - Variação do k_{cat} para Sfβgli mutantes em relação à enzima selvagem.

Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifadas. * Os dados para o mutante R97A não foram incluídos para que a escala não fosse achatada dificultando a visualização dos demais. Dados para este mutante estão na Tabela 7.

Mais uma vez para que a variação deste parâmetro cinético assuma sentido mais concreto, foram calculados os efeitos mutacionais sobre a estabilidade do complexo $ES^\ddagger$ ($\Delta\Delta G^\ddagger$; Figura 20). Neste contexto, reduções do k_{cat} indicam que o complexo $ES^\ddagger$ formado com a enzima mutante é menos estável que aquele formado com a enzima selvagem, o que indicaria a perda ou enfraquecimento de interações não-covalentes formadas entre a enzima e o substrato no estado de transição ($S^\ddagger$) (Wilkinson *et al.*, 1983; Fersht, 1999).

Nota-se que exceto pelo mutante P346A, todas as demais mutações causaram alterações na estabilidade do complexo $ES^\ddagger$ que se enquadram na faixa de $\pm 3,0$ kJ/mol. Portanto, as mutações em resíduos “hubs” e “não hubs” causam apenas pequena desestabilização do complexo $ES^\ddagger$, ou seja, as interações não-covalentes destes complexos não são drasticamente alteradas.

Deve ser destacado que R97A e Y331A não se enquadram nesta tendência. No primeiro caso, observa-se desestabilizações do complexo $ES^\ddagger$ na faixa de 30 kJ/mol para todos os substratos analisados. Já no caso de Y331A a desestabilização é ainda maior, resultando em uma enzima inativa. Estes resultados são coerentes com aqueles observados anteriormente para os mutantes R97M e Y331F (Marana *et al.*, 2003), quando foi demonstrado que R97 e Y331 modulam através de interações não-covalentes o pK_a do nucleófilo catalítico E399. Logo, os resíduos R97 e Y331 tem papel essencial na catálise, além de serem “hubs” da rede estrutural de SfbgI. A análise do resultado para as mutações nos demais resíduos “hubs” sugere que o papel catalítico de R97 e Y331 é mais preponderante no efeito mutacional do que seus posicionamentos como “hubs” na rede estrutural.

Figura 20 – Efeito mutacional sobre a energia do complexo $ES^\ddagger$ ($\Delta\Delta G_b$). Mutações em resíduos “não hubs” estão grifadas. Mutações em resíduos “hubs” não estão grifadas. As linhas tracejadas representam o limite inferior de energia esperado para uma ligação de hidrogênio. Os erros relativos estão na faixa de 1 a 17%.

Em suma, os parâmetros cinéticos para hidrólise dos *p*-nitrofenil β -glicosídeos são afetados de modo similar por mutações em resíduos “hubs” e “não hubs” da rede estrutural de SfbgIi. Os efeitos são de pequena intensidade, levando a alterações na estabilidade de ES e ES[‡] da ordem de 3,0 kJ/mol, não superando assim a energia de uma ligação de hidrogênio. Considerando que os espectros de fluorescência intrínseca de triptofano (Figura 14) indicam que todas as SfbgIi encontram-se enoveladas, pode-se sugerir que apesar das substituições em resíduos “hubs” resultarem em redes estruturais significativamente diferentes, estas enzimas mutantes se enovelaram de modo a apresentarem sítios ativos com estrutura espacial muito semelhante à enzima selvagem, os quais apresentam as mesmas propriedades funcionais. Esta similaridade funcional mesmo frente a uma rede estrutural onde o “caminho mínimo médio” (L) é significativamente maior também sugere que na ligação do substrato e catálise eventuais movimentações da estrutura terciária de SfbgIi são localizadas ou não se propagam amplamente pela rede estrutural da enzima. Conclui-se, assim, que os resíduos “hubs” não são determinantes nas propriedades catalíticas de SfbgIi.

Esta conclusão contrasta com o trabalho de Amitai *et al.* (2004) que demonstrou que o posicionamento como “hub” da rede estrutural proteica é um parâmetro com alto poder preditivo para localização de sítios ativos, do qual se depreenderia que estes resíduos seriam funcionalmente importantes em enzimas.

4.6 Estabilidade térmica das SfbgIi mutantes

Para avaliar a estabilidade térmica das SfbgIi mutantes foram realizados experimentos de inativação térmica a 45°C (item 3.11). Os dados obtidos nestes ensaios foram tratados assumindo uma cinética de primeira ordem para o mecanismo de desnaturação, assim construiu-se gráficos do logaritmo do percentual de atividade remanescente *versus* o tempo de incubação a 45°C (Figura 21). As inclinações de tais gráficos são proporcionais à constante de velocidade

para o processo de desnaturação térmica (k) e assim refletem a estabilidade térmica a 45°C para SfbgI selvagem e mutantes (Tabela 8) (Segel, 1993).

Baseando-se em uma análise comparativa nota-se que a inativação de todos mutantes com trocas nos resíduos ‘não-hubs’ é muito similar àquela mostrada pela enzima selvagem. O k relativo variou entre 0,8 a 1,3 vezes. Coerentemente, uma publicação prévia (Mendonça e Marana, 2011) já havia apresentado que a inativação térmica a 50° C de mais de 30 aminoácidos de SfbgI que não são ‘hubs’ ($z_k < 1.5$) mostram processos de inativação idênticos à enzima selvagem, com cinética de primeira ordem a 50°C e com k próximo a 0,03. Em contraste, com exceção do mutante S247A ($k_{mut}/k_{wt} = 1,1$) e R97A ($k_{mut}/k_{wt} = 0,5$), todos os mutantes com troca de resíduos ‘hubs’ apresentaram um k muito maior que a enzima selvagem. De fato, o k relativo destes mutantes variou entre 2,5 a 8,8 vezes. Os mutantes N249A, F334A e S358A inativaram-se segundo um mecanismo de primeira ordem atingindo níveis de atividade remanescente muito baixo (cerca de 3% ou menos) que se mantiveram estáveis ao longo do tempo (Figura 21). É provável que o processo de inativação destes mutantes tenha atingido o equilíbrio e a baixa atividade remanescente corresponda à pequena quantidade de enzima nativa nesta situação. Já os mutantes F251A, K366A e Y420A, que seguem um mecanismo de inativação com duas etapas, exibem uma delas com k também muito superior à enzima selvagem. Neste caso, a segunda componente de inativação partiria em $t = 0$ de cerca de 10% da atividade inicial. Assim, parece improvável que a segunda componente represente o equilíbrio do processo de desnaturação. Assim, este mecanismo pode indicar que a enzima nativa é rapidamente convertida em uma forma com menor atividade que se inativa mais lentamente. Uma interpretação alternativa seria que cada uma das fases de inativação corresponde à desnaturação de porções distintas da enzima, sendo que estas porções apresentariam estabilidades térmicas diferentes.

Dentre as exceções comentadas acima, o mutante R97A ($k_{mut}/k_{wt} = 0,5$) exibe uma atividade catalítica muito baixa, o que possivelmente dificultou a análise precisa do processo de inativação. Assim, o mutante S247A parece ser realmente uma exceção à tendência de mutações em “resíduos hubs” aumentaram a velocidade do processo de inativação. Uma possível justificativa para este comportamento distinto seria que a troca S por A foi relativamente conservativa, não gerando impacto como no caso das demais mutações de “resíduos hubs”. De fato, as demais mutações são mais drásticas em termos de remoção da cadeia lateral, N249A, F334A, F251A, K366A e Y420A. Contudo, uma das mutações em “resíduo hub” é qualitativamente idêntica à S247A. Neste caso uma análise mais detida do modelo estrutural de Sfβgli empregando o software DeepView/Swiss-PdbViewer 3.7 mostra que das 11 interações estabelecidas por S247, apenas 1 envolve o grupo hidroxila da cadeia lateral, sendo perdida com a mutação. Já S358 estabelece 6 contatos, sendo que 2 deles envolvem o grupo hidroxila da cadeia lateral. Assim, a mutação também neste caso parece ter sido mais drástica para S358 do que S247, talvez explicando a ausência de efeito sobre o perfil de inativação da Sfβgli mutante S247A.

Tabela 8: Constante da taxa de inativação (k) derivada das curvas de inativação térmica de primeira ordem das SfbgIi selvagem e mutantes expostos a 45°C.

SfbgIi	k (min^{-1})	$k_{\text{mut}}/k_{\text{wt}}$ relativo
Selvagem	$0,026 \pm 0,001$	1,0
Y56A●	$0,0316 \pm 0,0006$	1,2
M57A●	$0,0323 \pm 0,0006$	1,2
<u>D84A●</u>	$0,0215 \pm 0,001$	0,8
R97A	$0,0121 \pm 0,001$	0,5
S198A●	$0,032 \pm 0,001$	1,2
S247A	$0,0279 \pm 0,001$	1,1
N249A	$0,10 \pm 0,01$	3,8
F251A	$0,10 \pm 0,01$	3,8
	$0,007 \pm 0,001$	
F334A	$0,066 \pm 0,002$	2,5
P344A●	$0,033 \pm 0,001$	1,3
P346A●	$0,032 \pm 0,0001$	1,2
S349A●	$0,032 \pm 0,001$	1,2
<u>L350A●</u>	$0,0216 \pm 0,0008$	0,8
S358A	$0,13 \pm 0,01$	5,0
K366A	$0,10 \pm 0,01$	3,8
	$0,022 \pm 0,002$	
Y420A	$0,23 \pm 0,02$	8,8
	$0,0183 \pm 0,002$	

Mutantes marcados com ● e não grifados foram retirados da tese de doutoramento de Lúcio M. F. Mendonça (2009). (<http://www.teses.usp.br/teses/disponiveis/46/46131/tde-10122009-143601/en.php>)

Figura 21: Inativação térmica das Sfβgli selvagem e mutantes a 45° C.

4.7 – Análise comparativa do papel de resíduos “hubs” e “não-hubs”

A determinação dos efeitos mutacionais da substituição de resíduos “hubs” e “não-hubs” sobre parâmetros cinéticos da atividade catalítica (item 4.5) e da inativação térmica de Sfβgli (item 4.6) permitiu uma análise comparativa da participação destas duas classes de resíduos na determinação de propriedades desta β-glicosidase.

No caso da atividade catalítica foi tomado como referencial o k_{cat} para hidrólise de *p*-nitrofenil β-glucosídeo. Este substrato foi escolhido, pois Sfβgli apresenta considerável atividade sobre o mesmo e seu glicone ocorre em substratos naturais desta enzima (Marana *et al.*, 2001). Além disso, parâmetros cinéticos para hidrólise deste substrato puderam ser determinados para todos os mutantes, o que permitiu uma comparação quantitativa em toda a série de mutantes estudados. Como já comentado anteriormente, foram consideradas variações significantes de k_{cat} aquelas que resultaram da alteração de estabilidade de $ES^\ddagger$ em 3,0 kJ/mol, pois tal variação de energia corresponderia a uma ligação de hidrogênio. Tal alteração de estabilidade corresponde a variações de 3,5 vezes no k_{cat} . Assim, baseando-se nos dados de $\Delta\Delta G^\ddagger$ para o substrato *p*-nitrofenil β-glucosídeo, observa-se variações significativas apenas para

as mutações em P346 e F251. A mutação no resíduo D84 também foi considerada significativa, pois reduziu drasticamente a atividade de Sfβgli inviabilizando a determinação dos parâmetros cinéticos. Os mutantes R97A e Y331A não foram considerados, pois é provável que seus efeitos mutacionais estejam ligados ao papel destes resíduos na modulação do pK_a do nucleófilo catalítico E399 e não a sua posição na rede estrutural de Sfβgli.

Paralelamente, no caso da cinética de inativação térmica, foram considerados efeitos mutacionais significativos aumentos ou reduções da constante de inativação (k) em duas vezes. Nesta categoria se encontram mutações em R97, N249, F251, F334, S358, K366 e Y420. Não causaram efeito significativo no processo de inativação térmica de Sfβgli, as mutações nos resíduos Y56, M57, D84, S198, S247, P344, P346, S349 e L350. Não foram classificados os mutantes T245A e Y331A que eram inativos.

Para facilitar a sua visualização e comparação, estes resultados foram mapeados na distribuição dos resíduos estudados na estrutura de Sfβgli, sendo que os resíduos cuja substituição por alanina gerou alteração significativa em k_{cat} ou k (constante da taxa de inativação térmica) foram marcados em vermelho e aqueles que geram efeitos mutacionais irrelevantes em verde (Figura 22). Uma inspeção deste mapeamento mostra que mutações em resíduos “hubs” tendem a modificar significativamente o processo de inativação térmica, enquanto que mutações em “não-hubs” são inócuas sob este ponto de vista. A única exceção a esta tendência é o resíduo S247, um “hub” cuja substituição por alanina não alterou o perfil de inativação térmica de Sfβgli. Contudo, como explicado acima esta exceção parece ser um efeito do tipo de mutação e não uma quebra legítima da tendência destacada. Por fim outra tendência observada é que tanto mutações em resíduos “hubs” como em “não-hubs” não afetam a atividade catalítica de Sfβgli. São apenas 3 mutações (2 em resíduos “não-hubs e 1 em resíduo “hub”) que levam a alterações significativas da atividade catalítica em um total de 17.

Figura 22 – Distribuição dos resíduos “hubs” e “não-hubs” na estrutura de Sfβgli e seus efeitos mutacionais sobre a atividade e estabilidade térmica desta enzima. **A** – Distribuição dos resíduos “hubs” (marcados em amarelo) e “não-hubs” (marcados em azul). O nucleófilo catalítico E399 está em púrpura. Para identificação dos resíduos ver figura 11. **B** – Efeito mutacional sobre a estabilidade térmica. Resíduos cuja substituição por alanina alterou significativamente (aumentos ou reduções em 2x) a velocidade de inativação térmica de Sfβgli à 45°C estão marcados em vermelho. Resíduos cuja substituição não causou efeito neste parâmetro estão em verde. O resíduo Y331 está em cinza. **C** – Efeito mutacional na atividade de Sfβgli sobre *p*-nitrofenil β-glucosídeo. Resíduos cuja substituição por alanina alterou significativamente a estabilidade do complexo ES[‡] (variações acima de 3 kJ/mol) estão marcados em vermelho. Resíduos cuja substituição não alterou significativamente este parâmetro estão em verde. Os resíduos R97 e Y331 estão em cinza.

Em resumo, os resultados sugerem que os resíduos “hubs” da rede estrutural de SfbgIi estão associados à estabilidade térmica e ao processo de inativação desta enzima, mas não determinam sua atividade catalítica.

Frente a tal conclusão cabe sugerir uma hipótese explicativa. O processo de inativação térmica de SfbgIi não está descrito em detalhes, ou seja, não é conhecido como as interações não-covalentes que sustentam sua estrutura terciária são rompidas pelo aquecimento. Contudo, é válido assumir que o incremento de energia cinética resultante do aquecimento se distribua pela estrutura desta proteína através das suas interações covalentes e não-covalentes aumentando a movimentação (vibração) dos átomos participantes até que se atinja um limite onde as interações não-covalentes são finalmente rompidas e a proteína se desnatura. Em uma alegoria simplificada os aminoácidos poderiam ser imaginados como esferas e suas interações não-covalentes seriam molas. O aumento de energia cinética (aquecimento) se propagaria por este sistema através das molas até o ponto em que a vibração das esferas atingiria um limite em que as molas seriam rompidas e todo sistema seria desfeito.

Neste modelo do processo de inativação térmica, os resíduos “hubs” devem desempenhar um papel importante, pois determinando as “rotas de contato mais curtas” na rede estrutural proteica devem influenciar decisivamente a distribuição da energia cinética através da mesma. Assim, a mutação dos resíduos “hubs” provavelmente modificou o modo como a energia cinética se distribui na estrutura de SfbgIi, alterando significativamente o modo como sua estrutura é desfeita (desnaturação). Aliás, os resultados observados, aumentos da constante de velocidade de inativação térmica, sugerem que na ausência dos resíduos “hubs”, a energia cinética não se distribui mais com a mesma homogeneidade pela estrutura de SfbgIi, concentrando-se em regiões que se desnaturam mais rapidamente e levam a enzima como um todo à falência e perda de função. Neste aspecto, os resultados aqui obtidos e esta hipótese concordam com o comportamento previsto para redes “sem escala” que submetidas à remoção

seletiva dos “hubs” se fragmentam em conjuntos de nós isolados e entram em falência funcional, enquanto que o mesmo grau de dano dirigido à “não-hubs” não danifica seu funcionamento (Albert *et al.*, 2000). Portanto, é possível que o mesmo comportamento observado para redes do tipo “mundo pequeno” e “sem escala”, como a internet, tenha sido aqui observado para uma rede estrutural proteica.

5. Conclusão

Na tentativa de elucidar se mudanças nos resíduos centrais (“resíduos hubs”) de uma glicosidase (Sfβgli) poderiam afetar algumas propriedades enzimáticas desta proteína, trocou-se seus “resíduos hubs” por alanina. Considerou-se hubs aqueles aminoácidos cuja remoção ocasionasse maiores aumentos na via mínima média da rede de Sfβgli. As medidas dos parâmetros cinéticos destes mutantes usando 4 tipos diferentes de substratos mostrou que não houve tendência na mudança da especificidade ou eficiência catalítica, pois as alterações destes parâmetros foram pequenas. Contudo ensaios de estabilidade térmica destas enzimas contendo mutações em hubs revelaram uma estabilidade diminuída quando comparada às mutações direcionadas a aminoácidos não-hubs, que por sua vez, comportaram-se de maneira similar à enzima selvagem. Assim, considerando que os resíduos hubs determinam vias curtas para a conexão entre quaisquer resíduos em uma proteína, mutações nestas posições alterariam a propagação de vibrações resultantes do aquecimento ao longo da estrutura proteica de Sfβgli. Logo, vibrações poderiam estar concentradas em regiões restritas da proteína acelerando a sua desnaturação.

6. Referências Bibliográficas

- Albert R., Jeong H., Barabási A-L. Error and attack tolerance of complex networks. *Nature*. 2000. (406) 378-382.
- Amitai G., Shemesh A., Sitbon F., Shklar M., Netanel D., Venger J., Pietrokovski S.. Network analysis of protein structures identifies functional residues. *J Mol Biol*. 2004 (344) 1135-1146.
- Banaszynski, L.A., Wandless, T.J. Conditional control of protein function. *Chem Biol*. 2006. 13, 11-21.
- Atilgan A.R., Akan P., Baysal C.. Small-world communication of residues and significance for protein dynamics. *Biophysical Journal*. 2004 (86) 85-91.
- Badieyan S., Bevan D.R., Zhang C. Probing the active site chemistry of β -glucosidases along the hydrolysis reaction pathway. *Biochemistry*. 2012 (51) 8907-8918.
- Barabasi, A.L. e Oltvai, Z.N. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet*. 2004. 5, 101-13.
- Boccaletti, S., Latora, V. Moreno, Y., Chavez, M e Hwang, D.U. Complex Networks: Structure and dynamics. *FEBS Lett*. 2006. 581, 2776-82.
- Bruschweiler, S., Schanda, P., Klobner, K., Brutscher, B., Kontaxis, G., Konrat, R., and Tollinger, M. (2009). Direct observation of the dynamic process underlying allosteric signal transmission. *J. Am. Chem. Soc*. 131, 3063-3068.
- Cairns J.R.K, Esen A. B-glycosidases. *Cell Mol Life Sci*. 2010 (67) 3389-3405.
- Cantarel B.L., Coutinho P.M, Rancurel C., Bernard T., Lombard V., Henrissat B.. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*. 2009 (37) 233-238.
- Chica R.A., Doucet N., Pelletier J.N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. 2005 (16) 378-384.
- Csermely, P. Weak links: a universal key for network diversity and stability. 2006. Springer Verlag, Heidelberg.
- Csermely P. Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends in Biochemical Sciences*. 2008. 33 (12) 569-576.
- Czjzek M. , Cicek M., Zamboni V., Bevan D.R., Henrissat B., Esen A. The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, - DIMBOA-Glc and – dhurrin complexes. *Proc. Natl Acad Sci. USA*. 2000 (97) 13555-13560.

Czjek M., Cicek M., Zamboni V., Burmeister W.P., Bevans D.R., Henrissat B., Esens A. Crystal structure of a monocotyledon (maize ZMGlu1) β -glucosidase and a model of its complex with p-nitrophenyl β -D-thioglucoside. *Biochem J.* 2001 (354) 37-46.

Cunningham B.C., Wells J.A. High resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science.* 1989 (244) 1081-1085.

Datta D., Scheer J.M., Romanowski M.J., Wells J.A. An allosteric circuit in caspase-1. *J. Mol. Biol.* 2008 (381) 1157-1167

Davies G., Henrissat B. Structures and mechanisms of glycosyl hydrolases. *Structure.* 1995 (3) 853-859.

del Sol A., Fujihashi H., Amoros D., Nussinov R.. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol Syst Biol.* 2006 (2) 1-12.

del Sol A., Tsai C-H, Ma, Nussinov R.. The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure* 2009 (17) 1042-1050.

Dvir H., Harel M., McCarthy A.A., .Toker L, Silman I., Futerman A.H.. X-ray structure of human acid- β -glucosidase, the defective enzyme in Gaucher disease. *EMBO reports.* 2003 (4)7. 704-709.

Eftink M.R., Ghiron C.A. Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. *Biochemistry.* 1976. (15) 672-680.

Fersht A.R., Shi J-P, Knill-Jones J., Lowe D.M., Wilkinson A.J., Blow D.M., Brick P., Carter P., Waye M.M.Y., Winter G. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature.* 1985 (314) 235-239.

Fersht A. Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. 1999. W.H. Freeman, New York, NY.

Fuentes E.J., Der C.J., Lee A.L. Ligand-dependent dynamics and intramolecular signaling in a PDZ domain. *J. Mol. Biol.* 2004 (335) 1105-1115.

Greene J., Higman V.. Uncovering network systems within protein structures. *J Mol Biol.* 2003 (334) 781-791.

Guex N., Peitsch M.C., Swiss-Model and Swiss-PDBViewer: an environment for comparative protein modeling. *Electrophoresis.* 1997. (18) 214-272.

Husebye H., Arzt S., Burmeister W.P., Härtel F.V., Brandt A., Rossiter J.T., Bones A.M. Crystal structure at 1.1 Å resolution of an insect myrosinase from *Brevicoryne brassicae* shows its close relationship to β -glucosidases. *Insect Biochem Mol Biol.* 2005 (35) 1311-1320.

Jones S. & Thornton J.M.. Searching for functional sites in protein structures. *Curr Opin Chem Biol.* 2004. (8) 3-7.

Kelley L.A. & Sternberg M.J.E.. Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols*. 2009. 4 (3) 363-371.

Kuchner O., Arnold F.H.. Directed evolution of enzyme catalysts. *Trends Biotechnol.* 1997. 15, (12) 523-530.

Lee, D., Redfern, O., Orengo, C. Predicting protein function from sequence and structure. *Nat Rev Mol Cell Bio.* 2007. 8, 995-1005.

Lockless S.W., Ranganathan R. Evolutionary conserved pathways of energetic connectivity in protein families. *Science* 1999 (286) 295-299.

Marana S.R. Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life – Critical Reviews*. 2006 (58) 67-73

Martin A.J.M., Vidotto M., Boscariol F., Di Domenico T., Walsh I., Tosatto S.C.E.. RING: networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics*. 2011. 27 (14) 2003-2005.

Marana S.R., Mendonça L.M.F., Andrade E.H.P., Terra W.R., Ferreira C. The role of residues R97 and Y331 in modulating the pH optimum of an insect β -glycosidase of Family 1. *European J Biochem*. 2003 (270) 4866-4875.

Marana S.R., Terra W.R., Ferreira C. The role of amino-acid residues Q39 and E451 in the determination of substrate specificity of the *Spodoptera frugiperda* β -glycosidase.

Mendonça L.M.F., Marana S.R. The role in the substrate specificity and catalysis of residues forming the substrate aglycone-binding site of a β -glycosidase. *FEBS J*. 2008 (275) 2536-2547.

Mendonça L.M.F., Marana S.R. Single mutation outside the active site affect the substrate specificity in a β -glycosidase. *Biochim Biophys Acta*. 2011 (1314) 1616-1623.

Namchuk M.N. & Withers S.G. Mechanism of *Agrobacterium* β -glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. 1995 (34) 16194-16202.

Service R.F.. Biofuel Researchers Prepare to Reap a New Harvest. *Science*. 2007. (315) 5818. 1488-1491.

Verdoucq L., Czjzek M., Morimete J., Bevan D.R., Esen A. Mutational and structural analysis of aglycone specificity in maize and sorghum β -glycosidase. *J Biol Chem*. 2003 (278) 25055-25062.

Watts D.J. & Strogatz S.H.. Collective dynamics of 'small-world' networks. *Nature*. 1998 (393) 440-442.

Whisstock J.C., Lesk A.M.. Prediction of protein function from protein. *Quarterly Reviews of Biophysics*. 2003. 36, (3) 307-340.

Wilkinson A.J., Fersht A.R., Blow D.M. & Winter G. Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. *Biochemistry*. 1983 (22) 3581-3586.

SÚMULA CURRICULAR

DADOS PESSOAIS

Nome: Cecília Midori Ikegami

Local e data de nascimento: Foz do Iguaçu, PR; 02 de outubro de 1981

EDUCAÇÃO

Colégio, local, ano. Colégio Anglo Americano, Foz do Iguaçu, Ensino médio término em 1998.

Universidade Federal do Paraná, Curitiba, PR, período 2000-2003.

Graduação (Modalidade): Ciências Biológicas (Bacharelado e Licenciatura).

Universidade Federal do Paraná, Curitiba, PR, conclusão 2006

Mestrado em Biologia Celular e Molecular – Setor de Ciências Biológicas

Orientação Prof Dr Sílvio Marques Zanata

Universidade de São Paulo, São Paulo, SP.

Doutoranda do Instituto de Química, Bioquímica sob orientação do Prof Dr Sandro Roberto Marana de 2010 até o presente momento (no ano de 2009 esteve sob orientação do Prof Dr Alexander Henning Ulrich).

OCUPAÇÃO

Bolsista CAPES do período de 04/2009 à 04/2012.

RESUMO EM CONGRESSOS

Ikegami, C.M.; Arantes, G.M.; Marana, S.R. Role of highly connected residues (hubs) of enzyme structural networks on the substrate specificity. XL Reunião Anual da Sbbq, 2011, Foz do Iguaçu, PR.

MONITORIAS

Participação do Estágio Supervisionado em docência do Programa de Aperfeiçoamento de Ensino (PAE) junto à disciplina de Bioquímica: Metabolismo e Biologia Molecular (QBQ 0214), ministrada aos alunos de nutrição noturno no segundo semestre de 2009.

Participação do Estágio Supervisionado em docência do Programa de Aperfeiçoamento de Ensino (PAE) junto à disciplina de Bioquímica Experimental (QBQ 216), ministrada aos alunos de farmácia no segundo semestre de 2010.

FICHA DE ALUNOS EGRESSOS

Nome: Cecília Midori Ikegami

Telefone: 11 3673 0668

Endereço:

E-mail: cikegami@yahoo.com.br

Observações e sugestões para CPG:

.....

.....

.....

.....

.....

.....

.....

.....

.....