belaUniversidade Federal do Rio de Janeiro

Centro de Ciências Matemáticas e da Natureza – CCMN

Instituto de Química – IQ

Departamento de Bioquímica

Análise Funcional e Caracterização Estrutural *in silico* da Xilose Isomerase (*XylA*) de *Burkholderia cenocepacia*

Igor Patrick Vasconcelos Vieira

Orientadoras:

Elis Cristina Araújo Eleutherio

Bianca Cruz Neves

Rio de Janeiro

2017

Análise Funcional e Caracterização Estrutural *in silico* da Xilose Isomerase (*XylA*) de *Burkholderia cenocepacia*.

Igor Patrick Vasconcelos Vieira

Dissertação de Mestrado submetida ao Programa de Pós Graduação em Bioquímica, Departamento de Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários para a obtenção do título de Mestre em Ciências.

Orientadoras:

Elis Cristina Araújo Eleutherio

Bianca Cruz Neves

Rio de Janeiro

2017

Análise Funcional e Caracterização Estrutural *in silico* da Xilose Isomerase (*XylA*) de *Burkholderia cenocepacia*.

Igor Patrick Vasconcelos Vieira

Orientadoras:

Elis Cristina Araújo Eleutherio

Bianca Cruz Neves

Dissertação de Mestrado submetida ao Programa de Pós Graduação em Bioquímica, Departamento de Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários para a obtenção do título de Mestre em Ciências.

Aprovado por:

Presidente: Prof. Dra. Elis Cristina Araújo Eleutherio

Instituto de Química / Universidade Federal do Rio de Janeiro

---

Prof. Dr. Anderson de Sá Pinheiro

Instituto de Química / Universidade Federal do Rio de Janeiro

---

Prof. Dra. Magaly Girão Albuquerque

Instituto de Química / Universidade Federal do Rio de Janeiro

Rio de Janeiro/2017

Ficha catalográfica:

F657

Vieira, Igor Patrick Vasconcelos.

Análise Funcional e Caracterização Estrutural *in silico* da Xilose Isomerase (*XylA*) de *Burkholderia cenocepacia*. / Igor Patrick Vasconcelos Vieira. Rio de Janeiro: UFRJ/IQ, 2017.

xi, 81f.:il;31cm

Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Instituto de Química, Programa de Pós graduação em Bioquímica, 2017.

Orientadora: Elis Cristina Araújo Eleutherio.

Coorientadora: Bianca Cruz Neves.

Referências Bibliográficas: f. 63-68.

1.Bioquímica. 2. Etanol. 3. Fermentação 4. Xilose isomerase. 5. Modelagem comparativa. I. Eleutherio, Elis Cristina Araújo (orient.). II. Neves, Bianca Cruz (coorient.). III. Universidade Federal do Rio de Janeiro. Instituto de Química. Programa de Pós Graduação em Bioquímica. IV. Título.

CDD: 572

Este trabalho foi realizado nas dependências do Departamento de Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro, sob orientação da Profa. Dra. Elis Cristina Araújo Eleutherio e coorientação da profa. Dra. Bianca Cruz Neves, com o apoio das seguintes agências de fomento:

- CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico.
- CAPES – Coordenação da Aperfeiçoamento de Pessoal de Nível Superior.

AGRADECIMENTOS

À minha família, que me deu apoio financeiro e emocional imensurável durante toda a minha jornada científica.

Às minhas orientadoras Elis Eleutherio e Bianca Neves, por dedicar seu tempo entre aulas, redação de projetos e reuniões para discutir os meus resultados e me guiar até esse momento.

À minha amiga (e aluna) Gabrielle Cordeiro por ter enfrentado todos esses desafios comigo e ter confiado a mim a missão de ajudá-la em sua formação. O seu lugar ao meu lado, seja dentro ou fora dos laboratórios, está garantido.

Aos meus amigos do LIFE, que fizeram com que este laboratório fosse o melhor lugar para se estar no Fundão. Todos os momentos de risada, de estresse e mau humor, de discussões técnicas e/ou filosóficas e de comilança certamente fizeram toda a diferença para que eu me mantivesse são.

Ao prof. Rafael Mesquita e seus alunos, por me acolherem como um integrante do seu laboratório e me dar todo o suporte técnico e a infraestrutura para gerar meus resultados.

Um agradecimento especial à Carolina Matos por estar ao lado da Gabi em todos os experimentos em que eu precisei de ajuda e por me tirar do tédio durante as longas tardes de atividade enzimática.

Ao prof. Pascutti e Diego Gomes, que entraram nessa colaboração comigo e tornaram esta conquista possível.

Aos colegas de corredor que sempre me forneceram reagentes nos momentos de necessidade. Sem vocês muitas coisas não seriam possíveis.

RESUMO

Com a crescente preocupação com as mudanças climáticas, há um esforço global para o desenvolvimento de novas tecnologias de produção energética a partir de matrizes renováveis. Neste cenário, a produção de etanol utilizando biomassa lignocelulósica ganha grande atenção. Um dos problemas associados à aplicação direta desta tecnologia está na incapacidade da levedura *Saccharomyces cerevisiae* em produzir etanol utilizando xilose. Isto se deve ao desbalanço redox nas duas primeiras etapas da via de catabolismo deste açúcar, criado pelas enzimas xilose redutase (XR) e xilose desidrogenase (XDH) que utlizam NAD^+ e NAD(P)H como cofatores. Na tentativa de compensar este efeito, a enzima xilose isomerase (XI) de *Burkholderia cenocepacia* foi clonada em *S. cerevisiae*, pois catalisa a conversão direta de D-xilose a D-xilulose utilizando Mg^{2+} como cofator. A expressão desta enzima recombinante aumenta a taxa de consumo de xilose durante a fermentação, assim como a produtividade de etanol, com o melhor resultado reportado até o momento. No entanto, pouco se sabe sobre o funcionamento desta enzima. A fim de melhor compreender os resultados obtidos, foi feita a expressão desta proteína fusionada a cauda His em *E. coli* e em seguida sua purificação. A análise funcional mostrou alta atividade específica a 37°C em pH 7,2, indicando que esta proteína consegue manter sua atividade em diversas condições. Os valores de Km e V_{max} para D-xilose foram determinados. O Km se mostrou 3 vezes menor que os melhores resultados reportados até o momento, consequentemente influenciando no fluxo metabólico para a produção de etanol durante a fermentação. Sendo assim, estes dados foram utilizados para a construção de um modelo teórico da proteína gerado por modelagem comparativa, onde foram feitas análises de estrutura global, predição dos sítios de ligação à íons metálicos, modo de ligação do substrato e geometria do sítio ativo. O refinamento dos complexos proteína-substrato foram feitos em simulações de dinâmica semelhantes às condições de bancada e revelaram grande estabilidade da estrutura global e estruturas secundárias, assim como da ligação da xilose ao sítio ativo. No entanto, percebe-se que regiões expostas ao solvente possuem um grau de flexibilidade maior em relação ao sítio ativo. Análises de superfície eletrostática revelaram que mutações próximas ao sítio ativo podem ser responsáveis pela mudança de carga na entrada deste motivo, que influencia na capacidade de interação da proteína com o substrato.

Palavras chave: Etanol. 2. Bioquímica. 3. Xilose Isomerase 4. Biomassa. 5. Fermentação. 6. Modelagem Comparativa.

ABSTRACT

The growing concern about climate change has created a global effort to develop new energy production technologies from renewable sources. In this scenario, the production of lignocellulosic ethanol receives great attention. One of the problems associated with the direct application of this technology lies in the inability of the yeast *Saccharomyces cerevisiae* to produce ethanol using xylose as sole carbon source. This is due to the redox imbalance in the first two steps of this sugar catabolism pathway, created by NAD^+-dependent xylose reductase (XR) and NAD(P)H-dependent xylose dehydrogenase (XDH). In order to bypass the native route the recombinant xylose isomerase (XI) from *Burkholderia cenocepacia* was cloned into the yeast *S. cerevisiae*, as it catalyzes the direct reaction of D-xylose to D-xylulose using the ion $Mg2^+$ as a cofactor. The expression of this enzyme was proven to increase the rate of xylose consumption as well as ethanol productivity, with no xylitol accumulation during fermentation. However, little is known about the functioning of this enzyme. In order to better understand the results obtained, the His-tagged recombinant XI was expressed in *E. coli* and further purified. Functional analysis showed high activity at 37ºC and pH 7.2, indicating that this protein is active in a wide range of environments. V_{max} and K_m parameters were determined. The apparent Km for xylose has shown to be 2-3 fold smaller than the best results published to date, and could be responsible for altering the metabolic flux towards ethanol production during fermentation. Thus, the *in vitro* assay conditions were used to build theoretical models by comparative modelling. The overall structure, metal binding sites, xylose binding sites and active site geometry were investigated and have shown to be highly conserved among species. Energy minimization and molecular dynamics simulation were performed in order to refine the models in conditions similar to the enzymatic assays. Data have shown that regions exposed to solvent are likely to be more flexible than regions closer to the active site. Furthermore, secondary and overall structure showed great stability as well as xylose conformation at the active site. Electrostatic surface analysis have shown that mutations close to the active site entrance might be responsible for altering the interaction between protein and substrate, thus interfering with enzyme activity.

Keywords: 1. Ethanol. 2. Biochemistry. 3. Xylose Isomerase. 4. Biomass. 5. Fermentation. 6. Comparative Modelling.

LISTA DE ABREVIAÇÕES

BSA	albumina sérica bovina
AMP	adenosina monofosfato
cAMP	adenosina monofosfato cíclico
ATP	adenosina trifosfato
DHAP	dihidroxiacetona-fosfato
GA3P	gliceraldeído-3-fosfato
IPTG	Isopropil β-D-1-tiogalactopiranosídeo
NADH	nicotinamida adenina dinucleotídeo
NADPH	nicotinamida adenina dinucleotídeo fosfato
PKA	proteína quinase A
PPP	via das pentoses fosfato (do inglês, pentose phosphate pathway)
XDH	Xilose Desidrogenase
XI	Xilose Isomerase
XK	Xiluloquinase
XR	Xilose Redutase

ÍNDICE

1. INTRODUÇÃO 16

1.1. O Brasil no contexto da produção de etanol mundial 16

1.2. O etanol lignocelulósico (etanol 2G) 18

1.3. Fermentação de glicose pela levedura *S. cerevisiae* 20

1.4. Fermentação de xilose por *S. cerevisiae* 22

1.5. Engenharia metabólica de *S. cerevisiae* para fermentação de xilose 24

1.6. Xilose Isomerase: estrutura e mecanismo de ação 27

2. JUSTIFICATIVA 34

3. OBJETIVO 34

4. MATERIAIS E MÉTODOS 35

4.1. Cepas, Plasmídeos e Meios de cultivo 35

4.2. Purificação da Xilose Isomerase 36

4.3. Ensaios enzimáticos 37

4.4. Predição da Estrutura Terciária de Proteínas 37

4.5. Simulações de dinâmica molecular e minimização de energia dos complexos proteicos 38

4.6. Cálculos de estado de protonação e superfície eletrostática 39

4.7. Análise de conservação de aminoácidos e estruturas 3D 40

4.8. Análise Estatística 40

5. RESULTADOS E DISCUSSÃO 40

5.1. Purificação e determinação dos parâmetros cinéticos ótimos 40

5.2. Predição da estrutura tridimensional da XI 47

5.3. Simulações de dinâmica molecular e minimização de energia dos complexos proteicos de *B. cenocepacia*, *C. phytofermentans e Piromyces* sp. 53

5.4. Cálculos de superfície eletrostática 59

6. CONCLUSÕES 62
7. REFERÊNCIAS 63
8. ANEXO 69

ÍNDICE DE FIGURAS

Figura 1: Esquema do efeito do pré tratamento na biomassa lignocelulósica. 19
Figura 2: Representação esquemática da ligação entre vias que sinalizam o metabolismo de glicose em levedura. 21
Figura 3: Via de catabolismo de xilose em S. cerevisiae. 23
Figura 4: Visão geral da estrutura tridimensional da família de proteínas Xilose Isomerase. 29
Figura 5: Visão geral da coordenação de íons metálicos na região do sítio ativo. 31
Figura 6: Esquema da dinâmica reacional de isomerização de D-xilose a D-xilulose. 32
Figura 7: Esquema planar da coordenação dos ións metálicos e da forma linear da xilose em estado transicional durante a catálise da reação. 33
Figura 8: Mapa do plasmídeo pTrcHis utilizado para a expressão heteróloga do gene xylA de B. cenocepacia. 36
Figura 9: Gel de acrilamida SDS-Page 12% com amostras das frações de lavagem da resina de níquel. 42
Figura 10: Separação das frações de proteína por volume molecular em resina Superdex 200. 44
Figura 11: Caracterização funcional da XI de *B. cenocepacia*. 46
Figura 12: Inibição da XI de B. cenocepacia na presença de diferentes concentrações de xilitol. 47
Figura 13: Gráfico de Ramachandran para os modelos teóricos selecionados após refino. 50
Figura 14: Conservação de resíduos na família de proteínas XI. 51
Figura 15: Alinhamento estrutural do sítio ativo da XI de B. cenocepacia, C. phytofermentas e Piromyces sp. 52
Figura 16: RMSD do esqueleto peptídico em relação à estrutura inicial durante a etapa de produção. 54
Figura 17: RMSF 3D para *B. cenocepacia* (A), *Piromyces sp.* (B) e *C. phytofermentans* (C) e gráfico de RMSF por resíduo (D). 55
Figura 18: Gráfico de variação do conteúdo total de estrutura secundária. 56
Figura 19: RMSF para os átomos do ligante. 57

Figura 20: Interação da xilose com o sítio ativo na proteína XI de B. cenocepacia. 58
Figura 21: Representação dos campos elétricos das proteínas em condições semelhantes às experimentais. 60
Figura 22: Valores dos vetores campo elétrico $(\vec{E})$ **na região de acesso ao sítio ativo.** 61

ÍNDICE DE TABELAS

Tabela 1: Valores de referência para a atividade enzimática da XI de três diferentes espécies...... 47
Tabela 2: Dados de BLAST dos moldes escolhidos para cada sequência alvo.48
Tabela 3: Dados comparativos de RMSD entre os modelos teóricos de XI. 53

1. INTRODUÇÃO

Com a crescente preocupação com as mudanças climáticas, majoritariamente causadas pela combustão de combustíveis fósseis, há um esforço global para o desenvolvimento de novas formas de produção energética a partir de matrizes renováveis (Jacobsen & Wyman, 2000). O principal candidato para a substituição da gasolina é o etanol, devido a sua menor emissão de gases após a combustão (Lopes *et al*., 2016). Dentro deste contexto, o Brasil destaca-se no cenário mundial por ser o segundo maior produtor de etanol. Com o crescimento populacional e a crescente demanda do mercado, políticas governamentais foram criadas para incentivar o uso do etanol, como a criação de variedades da cana de açúcar com maior rendimento e a fabricação de motores veiculares capazes de funcionar com diferentes misturas de gasolina e etanol (Lin & Tanaka, 2006).

1.1. O Brasil no contexto da produção de etanol mundial

O processo de produção de etanol pode variar majoritariamente no tipo de matéria prima utilizada. Os EUA e Europa utilizam majoritariamente fontes amiláceas (milho, trigo, batata) para a obtenção dos açúcares fermentescíveis, enquanto o processo brasileiro utiliza a cana de açúcar (Basso *et al.*, 2011). O processo pode ser dividido em três etapas principais: obtenção da mistura sacarídica, fermentação e purificação do etanol. A primeira etapa começa com a moagem das fibras de cana de açúcar para que se extraia o suco com alto conteúdo de sacarose (Cardona e Sanchez, 2007). Dependendo da concentração de açúcares presentes no suco, pode haver a adição de água ao xarope, a fim de evitar o estresse osmótico nos microrganismos responsáveis pela fermentação (Ivorra *et al.*, 1999). A levedura *Saccharomyces cerevisiae* é o microorganismo mais utilizado para a fermentação, devido a sua alta capacidade de conversão de hexoses à etanol. Com o objetivo de aumentar a concentração de etanol produzida, altas densidades celulares são utilizadas. Para isso, é feito o reciclo de até 90% da biomassa de leveduras de uma rodada de fermentação na rodada seguinte. O processo de reciclo consiste na utilização de um ácido forte (normalmente ácido sulfúrico) em uma diluição suficiente para atingir valores de pH tão baixos quanto 2,0 (Fischer *et al.,* 2008). Este processo ajuda na redução de contaminação bacteriana e reduz o custo associado à preparação de inóculos (Della-Bianca & Gombert, 2013; Della-Bianca *et al.*, 2014). Por consequência, atingem-se altas

densidades celulares dentro das dornas de fermentação (10-17% peso úmido/v), que contribui para um tempo de fermentação extremamente curto, variando entre 6 e 11 horas (Basso *et al.*, 2008). Após a fermentação, o etanol é separado da mistura através de destilação, produzindo o etanol hidratado utilizado como combustível. Para a produção de etanol anidro, usado na mistura da gasolina, é necessária uma segunda etapa de destilação seguida de uma etapa de desidratação (Mussato *et al.*, 2010).

Estima-se que na safra 2016/17 haja a produção de aproximadamente 694 milhões de toneladas de cana de açúcar, que representa um crescimento de 4,4% na safra anterior (2015/2016). Para isso, houve um aumento de 5,3% na área de colheita se comparado com a safra 2015/16. A destinação do suco da cana de açúcar é dependente da flutuação de preços do açúcar, do etanol combustível (hidratado) e da gasolina durante a safra. A projeção da produção de etanol é de cerca de 27,9 bilhões de litros, havendo um aumento na produção de etanol anidro (utilizado na mistura da gasolina) de 1,5% (CONAB, 2016). Por consequência, cerca de 222 milhões de toneladas de bagaço de cana de açúcar devem ser produzidos nesta safra. Este resíduo possui diversas finalidades, dentre elas alimentação do gado, cobertura da área de plantio para prevenção da erosão e cogeração de energia. O bagaço da cana de açúcar possui um alto poder calorífico e por isso a combustão deste material é feita em caldeiras utilizadas para a geração de energia em usinas. Cerca de 85% do bagaço é utilizado para esta finalidade, o que corresponde a 76,82% da biomassa utilizada para a geração de energia. Este montante chega a representar 8,93% da matriz de energia elétrica brasileira (Aneel, 2016). Quando este resíduo não possui destinação adequada, torna-se um poluente junto às cinzas provenientes da sua combustão.

Por outro lado, a eficiência média do processo de fermentação das usinas sucroalcooleiras brasileiras já encontra-se próximo do máximo teórico (Basso *et al*., 2008). Além disso, a expansão das áreas cultiváveis encontra embargos políticos e sociais, impedindo o aumento desmedido da área de cultivo (Brown et al., 2013). Nesse contexto, o uso de biomassa lignocelulósica, especialmente de resíduos agroindustriais, como o bagaço da cana de açúcar, para a produção de combustíveis líquidos ganha grande atenção.

1.2. O etanol lignocelulósico (etanol 2G)

A biomassa residual da cana de açúcar (bagaço e palha) é constituída de celulose, hemicelulose e lignina, majoritariamente. Este material possui um alto teor de carboidratos, presentes nos polímeros da celulose (40-50% da biomassa) e hemicelulose (20-40% da biomassa). Além disso, quase todo o conteúdo restante corresponde a lignina. A lignina é a responsável pela resistência física do vegetal e previne a ação microbiana. Já no processo de produção do etanol, este componente é um dos principais desafios da indústria pois dificulta o processamento após a separação dos polímeros e a total conversão a açucares livres por enzimas hidrolíticas. A conversão da biomassa em etanol consiste das seguintes etapas: pré tratamento, hidrólise e fermentação. O pré tratamento da biomassa tem por objetivo desfazer a estrutura cristalina dos polissacarídeos da celulose e hemicelulose, permitindo então a separação das frações de celulose e hemicelulose (**Figura *1***). Com este processo, o acesso de agentes químicos é facilitado, permitindo a quebra das ligações de hidrogênio intramoleculares do polímero (etapa conhecida como hidrólise ou sacarificação). Para a degradação da celulose, enzimas como as endoglucanases (EC 3.2.1.4) e exoglucanases (EC 3.2.1.91) são utilizadas para a liberação de glicose. Já no caso da hemicelulose, devida à sua grande complexidade estrutural, é necessário que se utilizem misturas enzimáticas que sejam capazes de quebrar os diferentes tipos de ligação presentes nesta molécula. No entanto, o custo de produção destas enzimas ser um fator limitante para a sua utilização em escala industrial. Agentes químicos, como ácidos fortes na presença de alta temperatura e pressão são comumente utilizados. (Vásquez, 2007). Porém, o uso de altas temperaturas e o baixo pH estão associados à produção de inibidores da fermentação em leveduras.

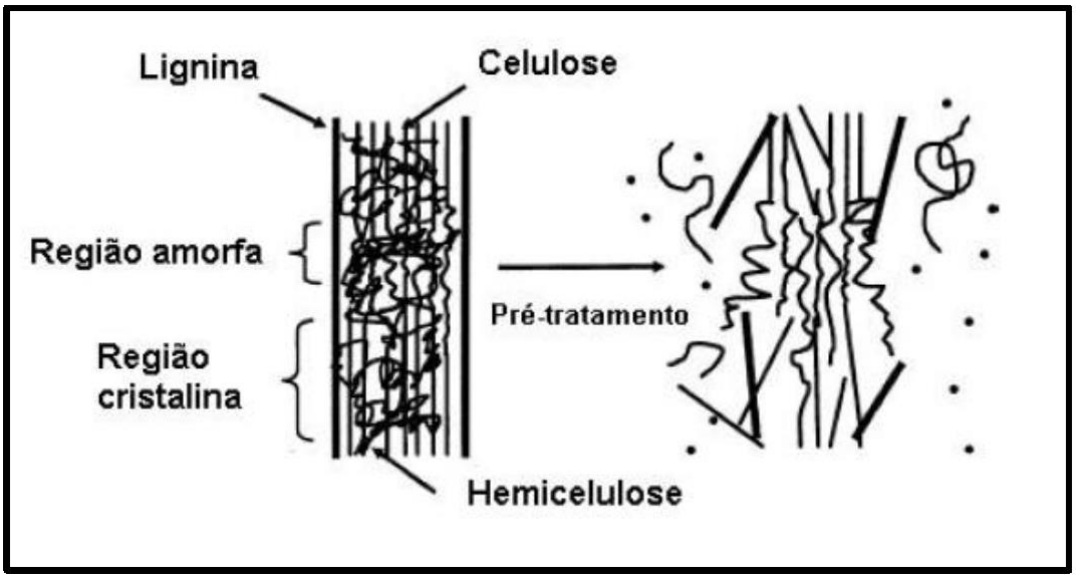

Figura 1: Esquema do efeito do pré tratamento na biomassa lignocelulósica. A estrutura cristalina dos polímeros sacarídicos é parcialmente destruída, permitindo o acesso de enzimas ou agentes químicos na etapa de hidrólise. Adaptado de (Mosier et al., 2005).

Várias construções de microorganismos recombinantes foram investigadas a fim de analisar sua capacidade de fermentação de hidrolisados lignocelulósicos na presença de inibidores (Zaldivar *et al.*, 1999). Dentre os principais inibidores gerados, destacam-se o ácido acético, furfural e hidroximetilfurfural (Binder *et al.*, 2010; Zaldivar *et al.*, 2000). Sabe-se que o furfural é um potente inibidor da produção de etanol, no entanto o seu mecanismo de ação ainda não está completamente esclarecido e ainda requer mais estudos. Já o ácido acético é capaz de atravessar a membrana plasmática da levedura e na presença do pH ácido citosólico é desprotonado. Desta maneira, a molécula não consegue ser carreada para fora da célula, levando à acidificação do citoplasma, que gera a inibição das proteínas deste compartimento e por consequência uma menor produtividade de etanol (Demeke et al., 2013).

Os hidrolisados provenientes da biomassa possuem diversas variedades de açúcares, sendo a glicose e a xilose os monossacarídeos predominantes. Sendo assim, para que se tenha um alto rendimento do processo fermentativo do etanol de segunda geração (2G), é necessário que se faça a fermentação simultânea dos dois açúcares.

1.3. Fermentação de glicose pela levedura *S. cerevisiae*

A fermentação da glicose se inicia pela difusão facilitada através dos transportadores da família HXT e pelo transporte ativo impulsionado pela força próton motriz, que permite o acúmulo intracelular da glicose (Stambuk *et al.*, 2008). Após a sua absorção, duas moléculas de ATP (adenosina trifosfato) são utilizadas para que haja a modificação a glicose-6-P, frutose-6-P e frutose-1,6-bifosfato. Imediatamente após a molécula é quebrada em dois composto de três carbonos: gliceraldeído 3-fosfato (G3P) e dihidroxiacetona fosfato (DHAP), sendo o último também convertido à G3P pela enzima triosefosfato isomerase. Sendo assim, esta molécula segue na segunda etapa da via glicolítica onde é convertida em múltiplas reações em duas moléculas de piruvato. Durante essas reações ocorre a redução de duas moléculas de NAD^+ a NADH e a formação de quatro moléculas de ATP a partir da fosforilação a nível de substrato. Com isso, obtém-se um saldo positivo de duas moléculas de ATP para cada molécula de glicose que é oxidada. Após a glicólise, as duas moléculas de piruvato são convertidas a duas moléculas de acetaldeído e CO_2. Este produto é então reduzido por duas moléculas de NADH, que formam o etanol, produto final da fermentação (Tortora *et al*., 2005).

A fermentação alcoólica pode ocorrer por duas vias distintas interconectadas: alta concentração de glicose e ausência e/ou baixa concentração de oxigênio (anaerobiose). Nas dornas de fermentação, o indutor da fermentação é a alta concentração de sacarose, que após ser hidrolisada libera uma molécula de glicose e uma de frutose. Ainda que haja a presença de oxigênio, a levedura irá metabolizar a glicose via fermentação, efeito conhecido como repressão catabólica ou efeito Crabtree (Crabtree, 1929). Desta forma, graças à inibição das vias respiratórias, a atividade glicolítica da célula pode aumentar em até 8 vezes (Oreb *et al.*, 2012). A repressão por glicose permite que usinas sucroalcooleiras possam utilizar dornas de fermentação abertas, o que reduz custos em equipamentos, sistemas de controle de pressão interna (gerado pela liberação de CO_2 durante o processo), sistemas de resfriamento e utilização de gases inertes para criação de atmosfera anaeróbica.

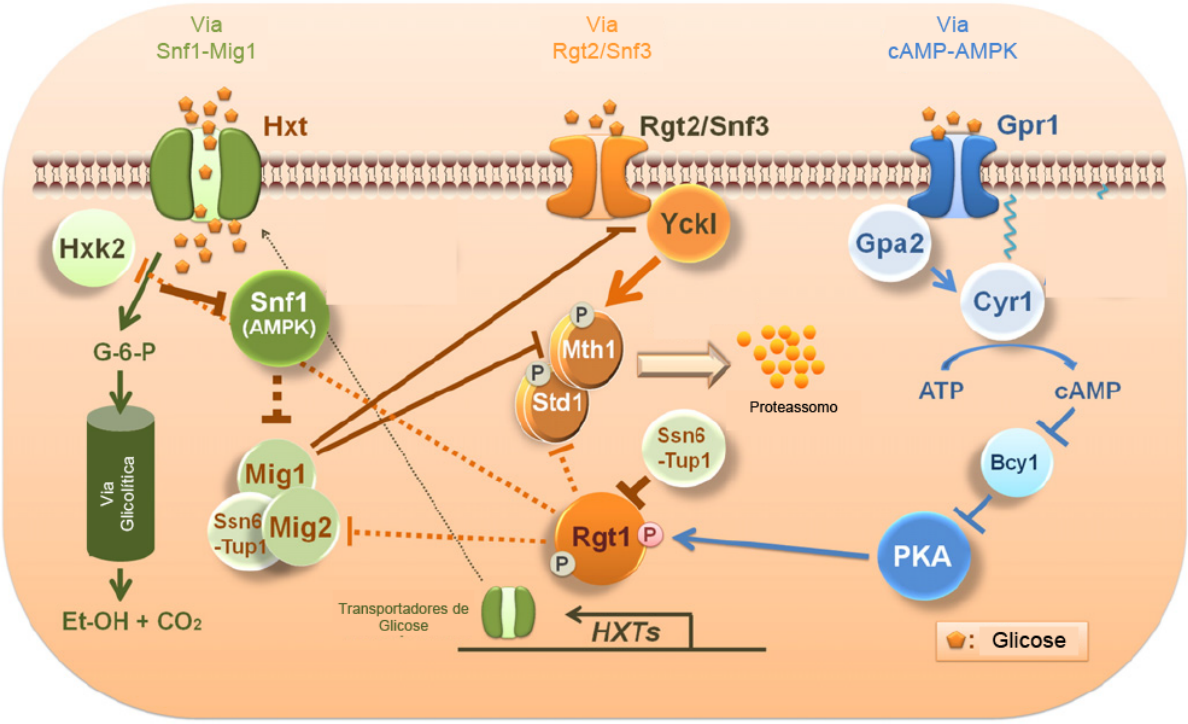

Figura 2: Representação esquemática da ligação entre vias que sinalizam o metabolismo de glicose em levedura. A glicose ativa os sensores de glicose Rgt2/Snf3, que por sua vez, ativam YckI levando à fosforilação de Mth1. O complexo Mth1/Std1 é degradado pelo proteassomo quando fosforilado. Uma cascata de sinalização dependente de PKA é ativada via Gpr1/Gpa2. Este complexo é responsável por ativar a adenilato ciclase (Cyr1), produz cAMP responsável por regular a ligação da subunidade Bcy1 à PKA. Por sua vez, PKA é responsável pela fosforilação de Rgt1, que nesta forma é dissociado do complexo Ssn6/Tup1, impedindo sua ligação ao DNA e desreprimindo seus genes alvo, como Hxt e Hxk2. A presença de glicose estimula a repressão da quinase Snf1 e de genes da respiração através de Mig1. A expressão de Mig2 é induzida pela via Rgt2/Snf3. A repressão do gene *snf3* por Mig1 indica que Snf3p possui um papel de sensor de glicose de alta afinidade, representando a integração entre a sinalização de glicose e vias de repressão. Adaptado de (Kim *et al.*, 2013).

O efeito Crabtree é resultado da conexão entre duas vias de sinalização da glicose: (i) a via Rgt2/Snf3, induzida por glicose, responsável pela absorção do açúcar; (ii) a via Snf1/Mig1, que regula positivamente os genes envolvidos na oxidação da glicose e faz a regulação negativa dos genes do metabolismo de açúcares alternativos (**Figura 2**). Uma vez que a geração de energia através da glicólise é ineficiente, se comparada à fosforilação oxidativa, células de leveduras compensam aumentando o número de transportadores Hxt em sua membrana plasmática, permitindo que se ultrapasse uma das etapas limitantes do metabolismo de glicose, sua absorção. Esta levedura possui ao menos 6 transportadores de glicose desta família (Hxt1, 2, 3, 4, 6 e 7), onde todos apresentam afinidades diferentes

para hexoses. Sendo assim, a célula expressa somente os transportadores mais adequados para a concentração de glicose disponível no ambiente. Uma vez que haja o acúmulo intracelular de glicose, o fator de transcrição Mig1 é translocado do citoplasma para o núcleo e liga-se aos repressores Tup1 e Ssn6, promovendo a repressão de genes regulados por glicose. Este mecanismo é dependente da sinalização intracelular via Rgt2/Snf3, que ativam YckI, quinase responsável pela fosforilação de Mth1. Quando Mth1 está fosforilada e complexada a Std1, o complexo é ubiquitinado e degradado no proteassomo. Quando este processo acontece, Rgt1 se torna disponível para fosforilação por PKA, que nesta forma está dissociado do complexo Ssn6/Tup1, impedindo sua ligação ao DNA e desreprimindo seus genes alvo, como *Hxt* e *Hxk2*. A cascata de sinalização dependente de PKA é ativada via Gpr1/Gpa2. Este complexo é responsável por ativar a adenilato ciclase (Cyr1), produzindo AMP cíclico (cAMP), responsável por regular a ligação da subunidade Bcy1 à PKA. A regulação de Mig1 é dependente da expressão de Mig2, gene esse induzido pela via Rgt2/Snf3. A repressão do gene *snf3* por Mig1 indica que Snf3 possui um papel de sensor de glicose de alta afinidade. Com isso, em poucos instantes, o processo de produção de ATP leva à diminuição de AMP (adenosina 3',5'-monofosfato) e inativa a proteína quinase Snf1 (homólogo à AMPK). Quando a concentração extracelular de glicose é diminuída ou esgotada, Snf1p é ativado via fosforilação, levando a translocação de Mig1 de volta do núcleo ao citoplasma, consequentemente desreprimindo seus alvos (Kim et al., 2013).

1.4. Fermentação de xilose por *S. cerevisiae*

A fermentação simultânea de xilose e glicose se faz necessária para que o processo de produção do etanol 2G seja economicamente viável. É necessário que seja utilizado somente um microrganismo capaz de fermentar na presença de misturas de açúcares e forme apenas um produto final (van Maris *et al.*, 2006 ; Harun *et al.*, 2010). A cadeia principal da hemicelulose presente na biomassa de cana de açúcar é composta majoritariamente por xilana, um polímero feito de unidades de β-D-Xilose em ligações β-1,4. A incapacidade da levedura *Saccharomyces cerevisiae* de produzir etanol a partir de xilose é um dos problemas associados com a aplicação direta desta tecnologia (Moysés *et al.*, 2016). Apesar de ser capaz de fazer a fermentação alcóolica de hexoses, como glicose e frutose, com alta eficiência, este microrganismo é ineficiente na produção de etanol a partir da

xilose. No entanto, em condições de anaerobiose, esta levedura é capaz de produzir etanol a partir de D-xilulose, isômero da xilose.

O metabolismo de xilose em *S. cerevisiae* requer a captação do açúcar por transportadores de hexose não específicos da família *HXT* (Kruckeberg, 1996). Então, a xilose é reduzida a xilitol pela enzima Xilose Redutase (XR), seguida de uma oxidação a D-xilulose pela enzima xilitol desidrogenase (XDH). Estas reações são NAD(P)H e NAD^+ dependentes, respectivamente. A D-xilulose precisa ser fosforilada pela xiluloquinase (XK) para que seja metabolizada na via das pentose fosfato (PPP) e subsequentemente na via glicolítica. No entanto, a utilização destes dois cofatores leva a um desbalanço redox citosólico. Isto se deve ao fato de *S. cerevisiae* não possuir atividade da enzima transhidrogenase, e por isso, não converte NADH a NAD(P)H, causando um acúmulo de NADH citosólico (**Figura *3***) (Bruinenberg, 1985). Uma consequência disso é o acúmulo de intermediários das vias, como xilitol e glicerol. Tentativas de construção de clones através da substituição de uma das enzimas para que houvesse a utilização de somente um cofator [NAD(P)H ou NAD^+] falharam em produzir uma cepa capaz de fazer a fermentação alcoólica eficiente de glicose e xilose simultaneamente, sem acúmulo de subprodutos (Matsushika, 2009).

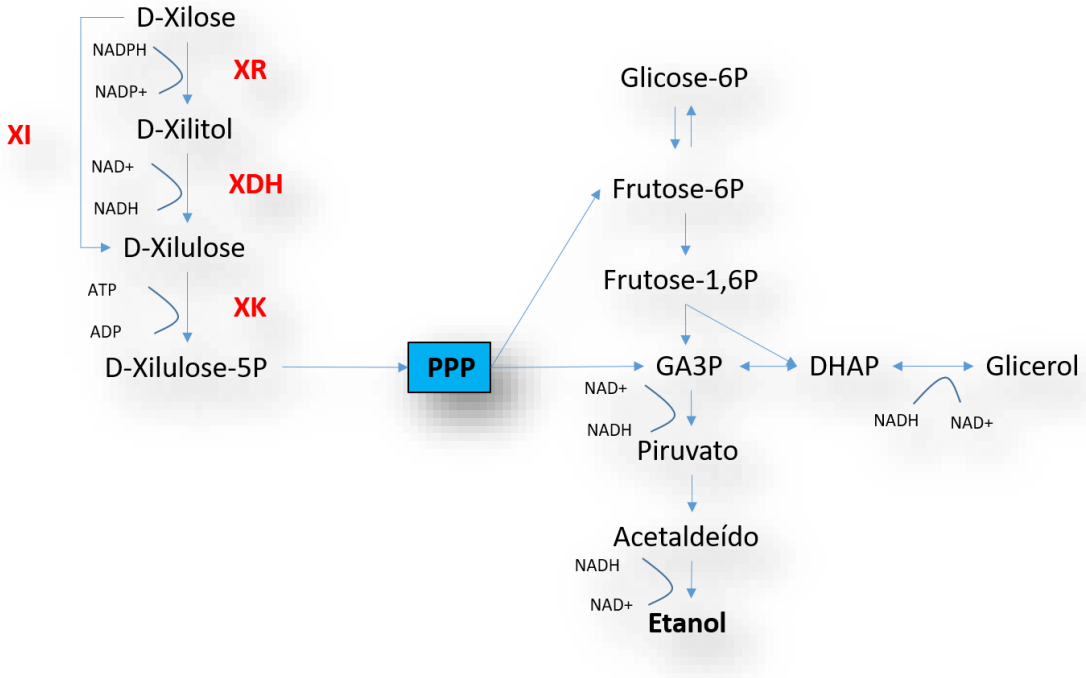

Figura 3: Via de catabolismo de xilose em *S. cerevisiae*. A D- xilose transportada para o citoplasma é convertida a D-xilulose através de duas etapas oxirredutivas pelas enzimas XR e XDH, respectivamente. A D-xilulose é fosforilada e utilizada na via das PPP, onde

produz os intermediários da via glicolítoca (Frutose-6P e GA3P). Estes intermediários então são totalmente oxidados a etanol. A inserção de uma XI recombinante permite a redução dos efeitos do desbalanço redox o acúmulo de subprodutos. PPP = via das pentoses fosfato ; XI = xilose isomerase ; XR = xilose redutase ; XDH = xilose desidrogenase ; XK = xiluloquinase ; GA3P = gliceraldeído-3-fosfato ; DHAP = dihidroxiacetona-fosfato.

Por outro lado, algumas espécies bacterianas utilizam outra via para a metabolização de xilose. A via da Xilose Isomerase (XI) faz a conversão direta da xilose à D-xilulose, utilizando metais divalentes como cofatores (Hahn-Hägerdal, 2007). A primeira XI recombinante expressa em *S. cerevisiae* funcional foi clonada de *Thermus Thermophilus* e mostrou baixa atividade a 30°C, pois sua atividade máxima ocorria a 85°C (Moysés, 2016). Desde então, outras XI procarióticas foram expressas em *S. cerevisiae*.

1.5. Engenharia metabólica de *S. cerevisiae* para fermentação de xilose

Diferentes estratégias estão sendo utilizadas para criar uma cepa capaz de fermentar xilose. Alguns grupos de pesquisa tentam fazer a inserção de uma enzima recombinante do sistema XR-XDH que utilizem os mesmos cofatores, equilibrando a produção de NADH durante a fermentação (Xiong *et al.*, 2011). No entanto, várias tentativas de clonagem de XR ou XDH de diferentes espécies em *S. cerevisiae* não resultaram em sucesso. Atribui-se o fracasso ao incorreto enovelamento das proteínas, modificações pós traducionais e ao pH interno da levedura (Gardonyi *et al*., 2003). Também foram feitas tentativas de suplementação do meio de cultivo com diferentes aceptores finais de elétrons, conexão da via de redução de furaldeído ao metabolismo de xilose e alteração da assimilação de amônio, produzindo cepas com baixas taxas de produção de xilitol. No entanto, a falta de estudos de fluxo metabólico em cima destes modelos e o alto grau de complexidade criado pelo grande número de alterações em vias controladoras do estado redox celular, desencorajam a sua utilização (Cai *et al.,* 2012; Moysés et al., 2016). Klimacek *et al.*, 2010 sugeriram, após análises de metabolômica e termodinâmica, que os fatores limitantes para a fermentação da xilose em etanol estariam na atividade de enzimas da via das pentoses fosfato e na concentração de frutose-6-fosfato livre, pois as concentrações de ATP e coenzimas não alteraram o fluxo metabólico das enzimas envolvidas na via da xilose. Sendo assim, conclui-se que seriam necessárias múltiplas modificações genéticas para que

se pudesse regular corretamente o fluxo metabólico e obter altos rendimentos na produção de etanol.

Por outro lado, a utilização da xilose isomerase para conversão direta da xilose a xilulose tem ganhado grande atenção pois são necessárias menos manipulações genéticas para que se obtenham resultados semelhantes aos descritos em estudos da via XR/XDH. A primeira XI recombinante expressa e com atividade significativa foi extraída do fungo anaeróbico *Piromyces* sp (Kuyper *et al*., 2003). A cepa recombinante mostrou alta atividade de XI, porém a levedura não mostrou crescimento significativo em meio contendo xilose como única fonte de carbono. Após fazer-se a adaptação em meio de xilose como única fonte de carbono, em anaerobiose, foi possível obter uma cepa com crescimento específico relativamente maior (0.03 h^{-1}) e rendimento de etanol de 0,42 $g.g^{-1}$ com uma produtividade de 0,2 g etanol / g célula / h (Kuyper *et al.*, 2004). Utilizando a mesma estratégia, (Madhavan et al., 2009) testaram a XI de *Orpinomyces* e obteve resultados semelhantes, mas com crescimento vagaroso em xilose. Todavia, a expressão de XI da bactéria anaeróbica *Clostridium phytofermentans* em uma cepa industrial de *S. cerevisiae* obteve um K_m similar comparado ao da XI de *Piromyces*, porém um K_i para xilitol três vezes menor, aproximadamente (Brat, *et al*., 2009).

Tabela 1: Produção de etanol por diferentes cepas expressando XI recombinante.

Espécie de origem da XI	**Crescimento específico (μ) (h^{-1})**	**Produtividade (g etanol / g célula / h)**	**Rendimento ($g.g^{-1}$)**	**Referência**
Piromyces sp.	0,03	0,2	0,42	(Kuyper *et al.,* 2003)
B. cenocepacia	-	0,19	0,42	(de Figueiredo Vilela et al., 2013)
C. phytofermentans	0,039	0,03	0,43	(Brat et al., 2009)

De forma similar, (Vilela et al., 2015) conseguiu expressar funcionalmente a XI de *Burkholderia cenocepacia*. Após fazer a engenharia adaptativa desta cepa em meio contendo xilose como fonte única de carbono, a levedura recombinante apresentou um dos melhores resultados até o momento, conseguindo atingir um rendimento de 0,42 g etanol/ g açúcar e uma produtividade de 0.19 g etanol / g célula / h. Em condições anaeróbicas, glicose e xilose foram simultaneamente consumidas enquanto não houve acúmulo de xilitol. Estes resultados mostram o grande potencial desta estratégia, considerando que não foram necessárias múltiplas modificações genéticas como ocorrido em outras construções (Cai *et al.*, 2012).

Ainda que a levedura *S. cerevisiae* mostre um grande potencial para a aplicação direta desta tecnologia, ainda existem algumas etapas limitantes a serem resolvidas para que haja a conversão simultânea de xilose e glicose em etanol. A levedura modificada para fazer este processo, o faz lentamente e não é capaz de consumir xilose até que a glicose seja completamente consumida (Vilela et al., 2015). Uma possível explicação para este fenômeno estaria na regulação dos genes que atuam no metabolismo de xilose, como a transaldolase. A reação de formação de eritrose-4-fosfato e frutose-6-fosfato a partir de sedoheptulose-7-fosfato e gliceraldeído-3-fosfato é catalisada por Tal1p e tem seu equílibrio fortemente alterado, em condições anaeróbias, indicando que esta seria uma etapa limitante da conversão de xilose a etanol (Klimacek et al., 2010). Além disso, há a competição entre glicose e xilose pelos transportadores de baixa afinidade Hxt, onde sua expressão é dependente da concentração de glicose no meio de cultivo (Cai et al., 2012). Sendo assim, é necessário que haja uma concentração mínima de glicose para que a expressão dos transportadores seja estimulada e a xilose também possa ser carreada para a célula. Caso contrário, em baixas concentrações de glicose, somente transportadores de alta afinidade serão expressos e estes possuem uma atividade extremamente reduzida para xilose. Esta teoria é corroborada por (Vilela et al., 2015) que publicaram as taxas de consumo de xilose e glicose de *S. cerevisiae* recombinante com o gene *XylA* de *B. cenocepacia*. Células crescidas em meio mínimo contendo diferentes misturas de xilose e glicose mostraram diferentes taxas de consumo dos açúcares durante a fermentação, onde cepas crescidas na proporção 3:1 glicose-xilose tiveram um consumo de 0,47 g/L/h para glicose e apenas 0,08 g/L/h para xilose. Estes resultados reforçam os dados publicados por (de Figueiredo Vilela et al., 2013) onde é comprovado que células fermentando glicose

possuem um rendimento de 0,51 g etanol / g açúcar, enquanto células fermentando xilose possuem um rendimento de apenas 0,23 g etanol / g açúcar.

Não se deve considerar que somente a inserção da xilose isomerase é suficientemente eficiente para que se promova a conversão eficiente da xilose em etanol. Isto se deve ao fato de um fenótipo desejado normalmente possuir uma característica pleiotrópica, ou seja, depende de múltiplos fatores (Baley, 1999). Diferentes vias metabólicas interagem para a formação de uma rede que leve à mudanças intracelulares responsáveis pela maior produção de determinado produto. De forma reversa, a simples deleção ou expressão aumentada de um gene pode não ser suficiente para obtenção do resultado esperado. Sendo assim, a engenharia evolutiva é uma ferramenta que possibilita a reprogramação do transcriptoma celular frente à uma determinada pressão evolutiva. Para o caso do etanol 2G, esta técnica normalmente envolve o cultivo de um microorganismo em meio de cultivo contendo xilose como única fonte de carbono. Sendo assim, há uma seleção de uma população capaz de crescer nestas condições. Quando este processo é feito por um número suficiente de vezes, mutações genéticas são acumuladas e a população remanescente é testada para o consumo de xilose, acúmulo de xilitol e produção de etanol, em condições de fermentação. Resultados mostram que cepas recombinantes com o gene *XylA* de *C. phytofermentans* submetidas à mutagênese química e engenharia evolutiva foram capazes de produzir 32% mais etanol que suas cepas parentais (Scalcinati *et al.*, 2012 ; Klimacek *et al.*, 2014).

Apesar de sua grande utilidade, poucos trabalhos publicados descrevem a utilização desta técnica para a obtenção de cepas com maior taxa de consumo de xilose, uma vez que o tempo necessário para a obtenção do fenótipo pode ser prolongado e variar significativamente entre diferentes espécies.

1.6. Xilose Isomerase: estrutura e mecanismo de ação

A enzima xilose isomerase (E.C.5.3.1.5) (Interpro: IPR001998) é uma enzima encontrada em microorganismos e plantas (Kristo *et al*., 1996) e catalisa a conversão de D-xilose à D-xilulose. É descrito que em algumas espécies também há a catálise de D-ribose à D-ribulose e D-glicose à D-frutose, o que inicialmente a levou esta proteína a ser catalogada como uma glicose isomerase (Marshall & Kooi, 1957). Na maioria dos casos, esta enzima é um homotetrâmero estabilizado por metais divalentes como cobalto (Co^{2+}),

magnésio (Mg^{2+}) e manganês (Mn^{2+}). Cada subunidade possui três domínios: (i) o domínio central que contém um barril-β circundado por α-hélices; (ii) o domínio C-terminal que consiste de uma estrutura de braço composta por 5 a 8 α-hélices que formam uma estrutura do tipo *domain swapped dimer*, responsável pela oligomerização; (iii) o sítio ativo encontra-se dentro de um compartimento próximo ao C-terminal e aos domínios de barril. Esta enzima possui atividade tanto em sua forma dimérica assim como na tetramérica (**Figura *4***) (Dauter *et al*., 1989; Henrick *et al.*, 1957).

Em termos de substrato, a xilose isomerase possui maior afinidade à xilose, originando então seu nome atual. Este monossacarídeo é considerado uma aldopentose, significando que possui cinco carbonos e um grupamento funcional formil. Como vários açúcares, este também pode adotar diversas estruturas dependendo das condições físico químicas da solução em que se encontra. Sua forma acíclica é pouco comum. No entanto, existem duas formas cíclicas prevalentes: (i) as piranoses, que consistem em anéis C_5O e; as furanoses, que são caracterizadas por anéis C_4O com um grupamento CH_2OH ligado ao carbono C4. Estas duas formas dependem da orientação relativa do grupamento hidroxila anomérico. Acredita-se que a D-xilopiranose seja o isômero catalisado por estas enzimas.

Figura 4: Visão geral da estrutura tridimensional da família de proteínas Xilose Isomerase. A) Tetrâmero colorido por subunidades; esferas magenta correspondem ao íon Mg^{2+} ; a região em vermelho corresponde ao domínio C-terminal. Código PDB: 1a0d. Xilose Isomerase de *Bacillus stearothermophilus*. Figuras geradas pelo software PyMol v1.7.4.5.

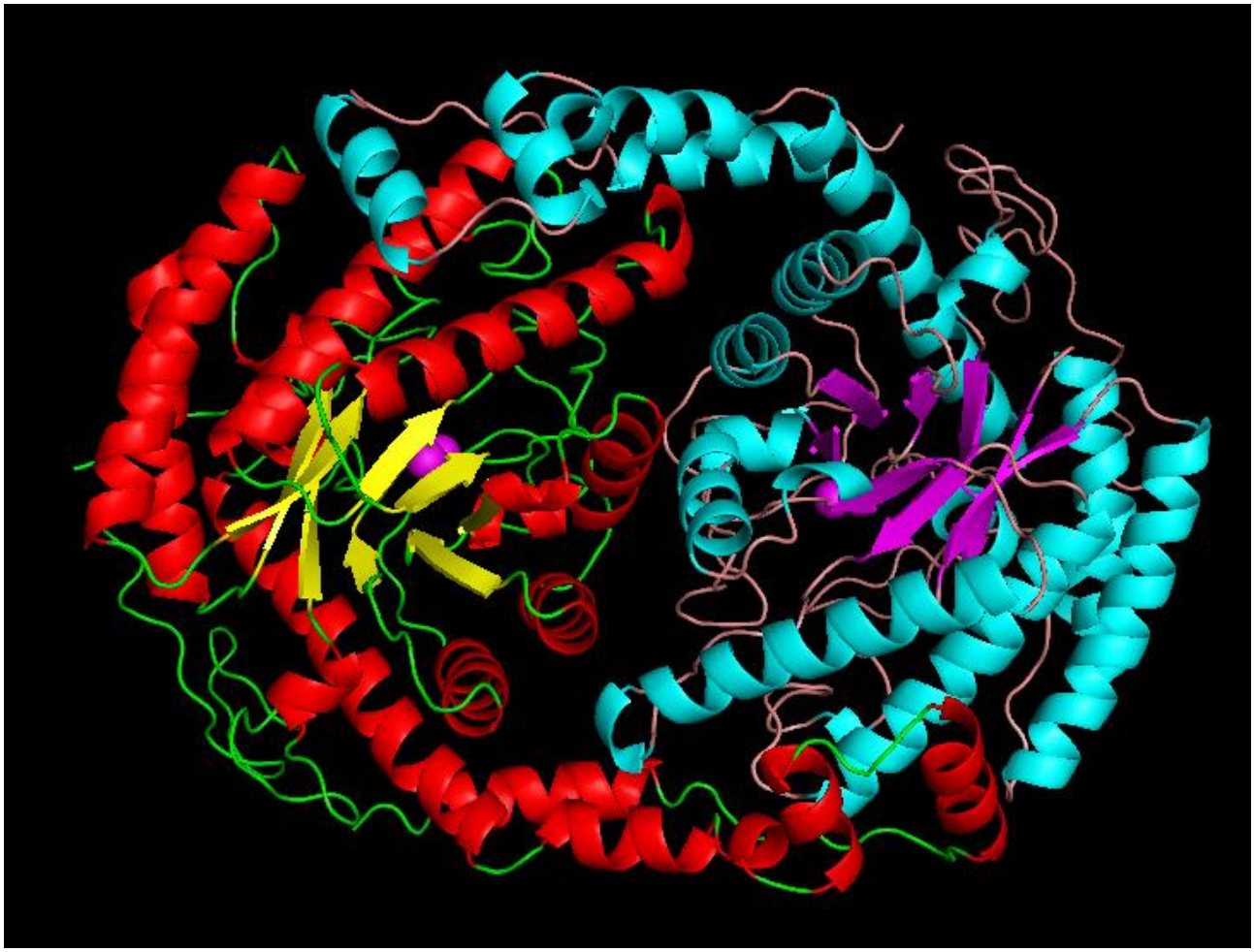

Figura 5: Visão longitudinal de cima para baixo contendo as subunidades A e B da Xilose Isomerase de *Bacillus stearothermophilus*. As subunidades estão representadas por um conjunto triplo de cores diferentes para cada estrutura secundária. Esferas magenta representam os íons Mg^{2+}. A região do sítio ativo encontra-se demarcada pelos domínios barril-β em amarelo e roxo. Código PDB: 1a0d. Figura gerada pelo software PyMol v1.7.4.5.

Uma das primeiras propostas de mecanismo de ligação de cofatores e substratos foram descritas por Asbóth & Náray-szabó, 2000. Foi proposto que as estruturas das subunidades monoméricas das XI de diversas espécies eram semelhantes, ainda que sua estrutura quaternária pudesse variar. Através de técnicas de difração de raio X, pode-se determinar o modo de ligação dos cofatores à proteína. Existem dois sítios de ligação a metais, denominados M1 e M2. O sítio M1 participa na ligação e coordenação do substrato enquanto o sítio M2 possui um papel na catálise da reação. O sítio M2 é conhecido por ocupar diferentes posições dentro do sítio ativo, denominados subsítios, durante a reação de isomerização. Deste momento em diante, irá se referir a estes subsítios como $M2_1$, onde o metal está em sua posição inicial e; $M2_2$ quando se desloca durante a reação. No subsítio

$M2_1$ o metal pode estar quadri-, penta- ou hexacoordenado por resíduos de ác. Glutâmico (Glu), Histidina (His), ác. Áspartico (Asp) e moléculas de água. Já no subsítio $M2_2$, o metal está coordenado por ác. Glutâmico e ác. Aspártico (**Figura *6***).

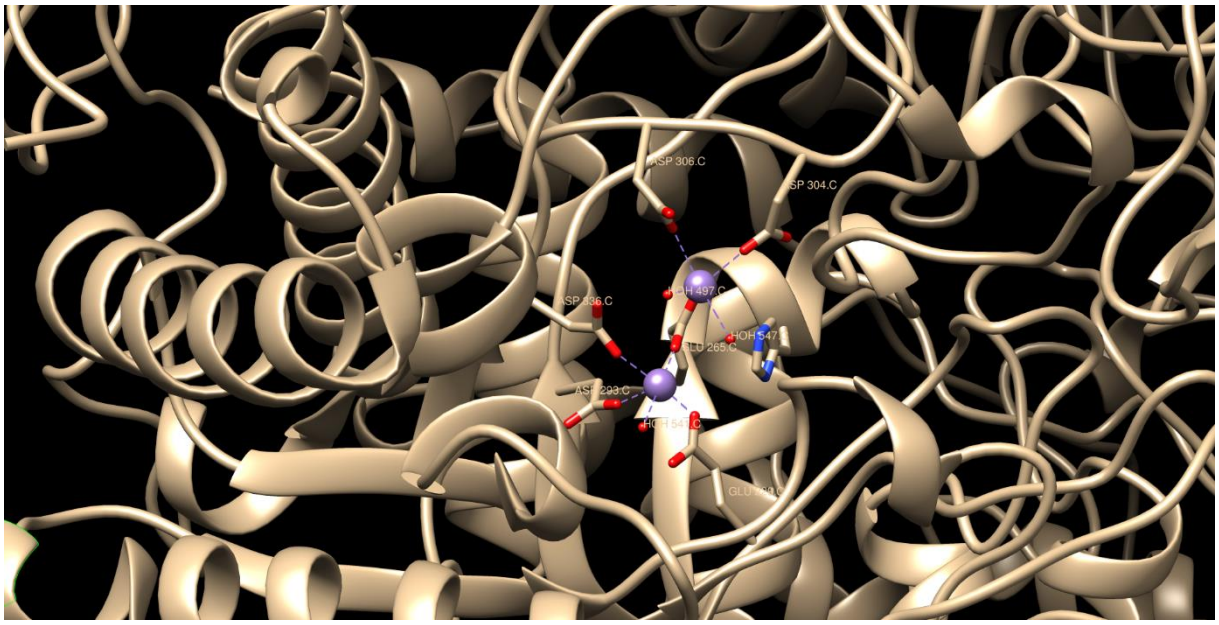

Figura 6: Visão geral da coordenação de íons metálicos na região do sítio ativo. Código PDB: 1a0d. Xilose Isomerase de *Bacillus stearothermophilus*.

Com isto, a teoria mais aceita para o mecanismo de ação envolve uma reação de migração de hidreto entre carbonos adjacentes. No entanto, a presença de cofatores e diferentes formas de substrato tornam mais complexa a cinética desta reação. Ainda não se sabe qual é a sequência de ligação entre proteína, substrato e cofatores. (Asbóth & Náray-szabó, 2000). No entanto, estudos de mutagênese dirigida indicam o papel individual de cada resíduo de aminoácido presente no sítio ativo, que indicam que os grupamentos carboxilato das cadeias laterais de Glu e Asp possuem um papel de ligação com os cofatores iônicos metálicos (Batt *et al.*, 1990; Cha et al., 1994; Kuyper et al., 2003). A dinâmica de reação possui três etapas: a reação começa com a abertura do anel de xilose, seguido da isomerização via migração de hidreto e, por último, o produto acíclico é estabilizado pelo fechamento do anel (**Figura *7***). Durante a reação, o metal catalítico move-se continuamente entre os sítios $M2_1$ e $M2_2$.

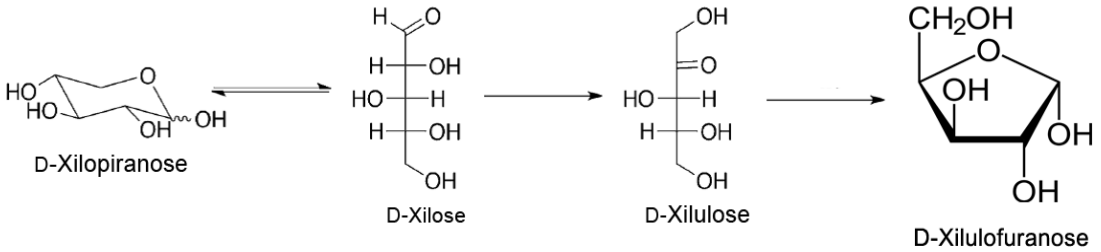

Figura 7: Esquema da dinâmica reacional de isomerização de D-xilose a D-xilulose. Adaptado de (Toteva *et al.*, 2011).

Sugere-se que a ligação da forma cíclica do substrato induz o deslocamento do íon metálico, e não a sua forma linear como sugerido anteriormente (Meng *et al.,* 1993). Existem evidências suficientes para suportar a teoria que descreve a etapa limitante da reação, sendo ela a etapa de transferência de prótons para a abertura do anel. Sugere-se que o átomo O1 da hidroxila do açúcar seja doadora de prótons para o átomo O5 do anel do açúcar, sendo transferido através do átomo NE2 da His (REFERENCIA MENG). Sendo assim, cria-se uma distribuição de cargas do tipo ânioca-catiônica-aniônica que pode contribuir significativamente para a atividade catalítica (**Figura *8***). Este tipo de distribuição de cargas também pode ser encontrado em lisozimas (Blake *et al.*, 1967) e serino proteases. Além disso, a estrutura tridimensional estruturada pelo íon metálico possui um importante papel na determinação do caminho do potencial eletrostático e na estabilização da estrutura em estado transicional durante a etapa de abertura do anel de xilose.

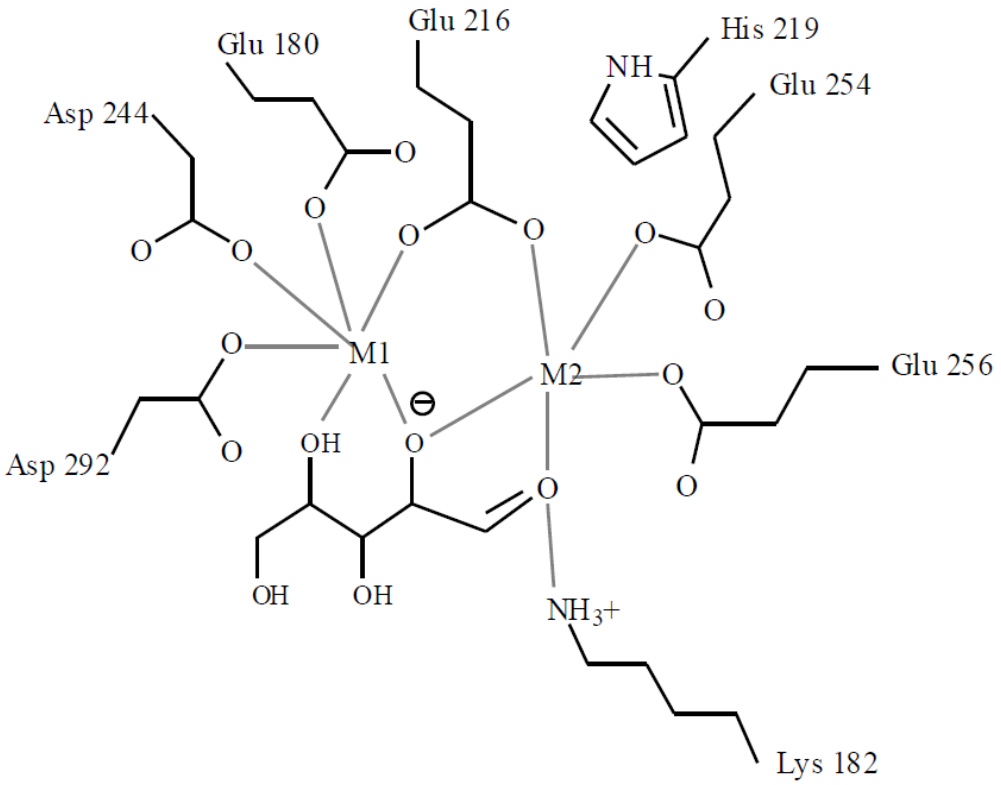

Figura 8: Esquema planar da coordenação dos íons metálicos e da forma linear da xilose em estado transicional durante a catálise da reação. Retirado de: (Asbóth & Náray-szabó, 2000).

Sendo assim, neste trabalho foi avaliada a performance da XI de *B. cenocepacia* através da obtenção de amostras purificadas da enzima, a fim de analisar as razões pelas quais somente a expressão desta enzima foi capaz de obter resultados equivalentes a cepas submetidas à extensivas adaptações metabólicas. Para isso, foi feita sua caracterização funcional através da medida de atividade enzimática específica *in vitro*, em diferentes condições de temperatura e pH. Além disso foram determinados parâmetros cinéticos como Km e V_{max} aparente para a xilose. As informações experimentais obtidas nestes ensaios foram utilizadas para a geração de modelos teóricos através de modelagem comparativa por homologia. Sendo assim, foram determinadas as regiões do sítio ativo e seus domínios de ligação a metais. Simulações de dinâmica molecular foram utilizadas para observar o comportamento do complexo proteína-substrato em condições semelhantes às dos ensaios de bancada. Além disso, foi investigada a conservação de resíduos de aminoácidos e suas cargas em diferentes regiões da proteína, a fim de determinar

diferenças significativas que pudessem explicar os resultados encontrados na análise funcional.

2. JUSTIFICATIVA

As novas políticas governamentais de incentivo ao uso do etanol criaram uma crescente demanda deste combustível líquido. No entanto, o aumento das áreas de plantio é desestimulado. Sendo assim, a fim de aumentar a produtividade por hectare plantado, novas tecnologias como o desenvolvimento de cultivares de cana de açúcar com maior conteúdo de sacarose e fibras alongadas estão sendo desenvolvidos. Ademais, há um esforço global para o desenvolvimento de tecnologias economicamente viáveis que permitam o uso da biomassa de bagaço para a produção de açúcares fermentescíveis.

O presente trabalho contextualiza-se nesta área de interesse, pois descreve importantes conceitos envolvendo a tecnologia de cofermentação de hexoses e pentoses pela levedura *S. cerevisiae*, uma vez que esta levedura não é capaz de fermentar xilose em meios contendo o açúcar como única fonte de carbono. É necessário que se façam estudos aprofundados na fisiologia de leveduras recombinantes que possuam o gene *xyl*A, codificante para a enzima xilose isomerase. O melhor entendimento de seu mecanismo de ação e propriedades físico químicas irá permitir que modificações racionais sejam feitas a fim de melhorar seu desempenho e, consequentemente, aumentar a produtividade de etanol pela levedura.

3. OBJETIVO

Caracterização bioquímica e estrutural da proteína Xilose Isomerase (XylA) de *Burkholderia cenocepacia*.

Metas:

1. Obter uma amostra purificada da proteína XI recombinante, expressa em *E. coli*;

2. Determinar a atividade enzimática específica *in vitro*;
3. Determinar os parâmetros cinéticos da XI de *B. cenocepacia* e mensurar sua atividade enzimática específica;
4. Construir modelo estrutural *in silico* das XI de *B. cenocepacia, C. phytofermentans* e *Piromyces sp.*;
5. Analisar a ligação da xilose no sítio ativo das XI citadas em 4;
6. Caracterizar o modo de ligação do substrato das proteínas das espécies citadas em 4.

4. MATERIAL E MÉTODOS

4.1. Cepas, Plasmídeos e Meios de cultivo

Uma reação de PCR foi utilizada para obter o gene *xylA* que codifica a enzima Xilose Isomerase de *Burkholderia cenocepacia* J2315 (GenBank: AM747722.1) , de acordo com (de Figueiredo Vilela et al., 2013). O gene foi clonado no plasmídeo pTrcHis-B (Thermo Scientific, EUA). O plasmídeo foi transfectado em *E. coli* TOP10 [F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*Z ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*Gλ] e selecionados em meio LB adicionado de 200 µg/mL de ampicilina. A expressão de proteínas foi feita em meio LB (1% p/v BactoTriptona, 0,5% p/v extrato de levedura e 1% p/v NaCl) suplementado com 500 µM Isopropil β-D-1-tiogalactopiranosídeo (IPTG) a 18ºC sob agitação de 160 rpm por 16 horas.

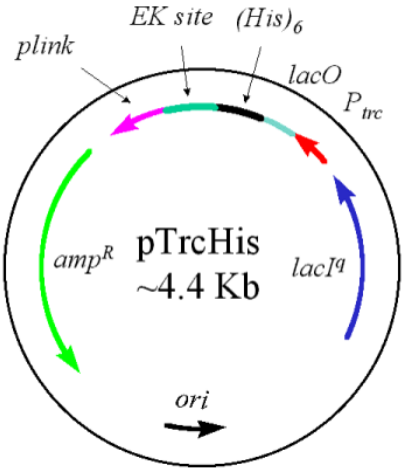

Figura 9: Mapa do plasmídeo pTrcHis-B utilizado para a expressão heteróloga do gene *xylA* de *B. cenocepacia*. P_{trc} : promotor *trc* ; *lacO*: operon *lac* ; *lacI*^{q} : ORF da proteína repressora do operon *lac* ; *ori* : origem de replicação PBR322 ; *amp*^{R} : ORF do gene de resistência à ampicilina ; *plink* : região de *polylinker* ; EK site: sítio de reconhecimento da enteroquinase (EK) ; $(His)_6$: região polihistidina.

4.2. Purificação da Xilose Isomerase

Células induzidas por IPTG foram coletadas por centrifugação a 2.900 x g por 30 minutos a 4ºC e ressuspendidas em tampão de lavagem (50 mM Fosfato de Sódio, 300 mM NaCl, pH 7) e congeladas até o momento da extração. Após descongelamento a temperatura ambiente, células foram rompidas em sonicador com 10 ciclos de 1 minuto intervalado de 1 minuto em banho de água e gelo, usando uma microponteira de titânio de 4 mM (QR 300, Ultronique, Brasil). O lisado celular foi centrifugado a 9.100 x g por 30 minutos a 4ºC e o sobrenadante foi coletado. Uma resina de Níquel (Novagen®, Merck, EUA) foi equilibrada em tampão de lavagem e as amostras foram incubadas com a resina por 30 minutos a 4ºC sob leve agitação. A resina contendo as amostras foi transferida para uma coluna descartável (Bio-Rad) e lavada duas vezes com 10 volumes de tampão de lavagem. O eluente foi coletado. Proteínas marcadas com o tag His foram eluídas com tampão de eluição (50 mM Fosfato de Sódio, 300 mM Cloreto de Sódio, pH 7,5) contendo 100, 200 e 300 mM de imidazol.

A eficiência de expressão e isolamento da XI foi analisada em gel de acrilamida SDS-PAGE (5% stacking e 12% gel de resolução) com tampão de amostra contendo: 100 mM Tris-HCl pH 6,8, 8M Uréia, 20% Glicerol, 4% SDS, 0,2% azul de bromofenol e 100 mM β-mercaptoetanol. O tampão de corrida Tris-Glicina-SDS foi utilizado para a passagem de 60mA por cerca de 45 minutos para a separação das proteínas em gel. Amostras contendo

XI foram concentradas utilizando a coluna Amicon Ultra 30K (Merck) por centrifugação a 4.000 x g a 4ºC por 10 min. A amostra concentrada foi injetada em uma coluna de resina Superdex 200 (GE Healthcare) com fluxo de 1 mL/min de tampão 100 mM Tris-HCl, 20 mM NaCl, pH 7.5. A concentração de proteína após a purificação foi feita através do método de BCA (Smith et al., 1985). A curva padrão foi feita utilizando albumina sérica bovina (BSA) como proteína reativa, variando sua concentração entre 25 a 2000 µg/µL.

4.3. Ensaios enzimáticos

A atividade da enzima XI foi mensurada de acordo com (Brat *et al.*, 2009). Uma solução contendo a enzima xilose isomerase a uma concentração de 195 µg/mL foi incubada a 37 ºC antes do início da reação. A mistura reacional contendo 0,23 mM NADH, 10 mM $MgCl_2$, 2U/mL sorbitol desidrogenase e 30 µg de XI em 100 mM Tris-HCl (pH 7.2) foi feita no momento do ensaio e incubada por 5 minutos à temperatura de ensaio em um volume total de 1 mL em cubeta de quartzo de caminho ótico de 1 cm. A reação é iniciada com a adição de 500 mM D-xilose e acompanhada por 30 minutos através da medida da oxidação do NADH em espectrofotômetro a 340 nm até completa oxidação à NAD^+. Variações na concentração de D-xilose entre 10 e 500 mM foram usadas para determinar os parâmetros cinéticos. O ponto isoelétrico teórico (pI) da proteína foi calculado pelo software ProtParam Tool (http://web.expasy.org/protparam/). A partir disso, o pH ótimo da proteína foi estimado através da variação do pH do tampão reacional entre 5,8 e 7,6. Para valores de pH entre 5,8 e 7,2, foi utilizado o tampão 100 mM Bis-Tris + 70 mM NaCl. Entre 7,2 e 7,8 foi utilizado o tampão 100 mM Tris-HCl + 20 mM NaCl. A força iônica dos tampões foi calculada de forma a se manter constante em todas as condições (87 mM). Ensaios de inibição foram feitos na presença de 125 mM de xilose e variando a concentração de xilitol de 10 a 50 mM em pH 7,2 e temperatura a 37ºC.

4.4. Predição da Estrutura Terciária de Proteínas

Para a modelagem por homologia, a sequência da XI de *B. cenocepacia* foi comparada com todas as sequências publicadas no banco de dados PDB (Protein Databank) utilizando a ferramenta BLAST (National Center for Biotechnology, NCBI). As sequências dos cristais com melhor valor de identidade e similaridade foram escolhidos e alinhados através da ferramenta MUSCLE (Edgar, 2004). Os alinhamentos foram manualmente verificados para

assegurar a qualidade e editados manualmente quando necessário. A modelagem da estrutura terciária foi feita no software MODELLER v9.17, gerando 500 modelos (Martí-Renom et al., 2000). A fim de escolher o melhor modelo gerado, foram utilizados os valores de função objetiva do programa, DOPE e Z-score (DOPE normalizado). A qualidade da geometria estrutural dos modelos escolhidos foi analisada no software PROCHECK (Laskowski, 1993). Já para a estatística de interações não covalentes foi utilizado o software ERRAT (Bowie, 1991). A verificação da compatibilidade do modelo atômico da estrutura terciária e a estrutura primária foi feita pelo software VERIFY 3D (Lüthy, 1992). Anomalias identificadas em regiões de loop foram refinadas utilizando a heurística de refino de loops do MODELLER. Os cálculos de RMSD dos carbonos α da cadeia principal foram feitos pelo software PyMol v1.8.4.

4.5. Simulações de dinâmica molecular e minimização de energia dos complexos proteicos

A investigação das estruturas cristalográficas usadas para derivar o modelo tridimensional do tetrâmero (1a0c, 1a0d), bem como das estruturas complementares (1s5m, 1s5n) revelaram a importância estrutural das águas, resolvidas na cristalografia de raio X, para a manutenção da estrutura do sítio ativo. Durante a preparação das estruturas para as simulações, ficou claro que a determinação dos estados de protonação dos aminoácidos depende fortemente da consideração de tais moléculas de água.

Os modelos estruturais para a *B. cenocepacia*, *C. phytofermentans e Piromyces* sp. foram sobrepostos às estruturas 1a0c, 1a0d, 1s5m e 1s5n. A estrutura cíclica da xilose foi obtida pela modificação da estrutura da glicose sobreposta na estrutura 1s5m (troca do grupamento metil por H_2). A orientação da glicose foi escolhida por se assemelhar ao modelo proposto para a catálise da xilose. As moléculas de água cristalográficas presentes no cristal 1a0c foram preservadas, exceto as moléculas de água 494, 495, 496, 639 (cadeia A), 527, 528, 529, 671 (cadeia B), 525, 526, 527, 670 (cadeia C), 558, 559, 560, 702 (cadeia D) que estavam sobrepostas a xilose no sítio ativo.

Os estados de protonação foram designados de acordo com o cálculo do PROPKA (Olsson *et al.*, 2011) para o pH 7,0 e investigação visual (incluindo as moléculas de água). As estruturas foram preparadas com o pacote de programas AMBER14 (Case *et*

al., 2014). Os parâmetros para a construção das topologias das proteínas foram obtidas do campo de forças AMBER 14SB (Maier *et al.*, 2015). Já os parâmetros para a xilose foram obtidos do campo de força GLYCAM 06j-1 (Kirschner *et al.*, 2008). Os sistemas foram envolvidos em caixa d'água octaédrica com 15 A de distância das extremidades e neutralizados com íons Na^+ ou Cl^- quando necessário.

O sistema foi relaxado antes das simulações de dinâmica molecular de acordo com o protocolo de equilíbrio a seguir: minimização da estrutura com restrições de posição a 100 kcal/mol para os átomos pesados da proteína e xilose, seguidos por uma minimização da estrutura com restrições de posição a 10 kcal/mol para a cadeia principal (*backbone*) da proteína e átomos pesados da xilose, seguidos por minimização do sistema completo com restrições de posição fracas a 1 kcal/mol para os átomos pesados do complexo. Em seguida o sistema foi aquecido de 0-100 K no "ensemble" canônico (NVT) por 20 ps, passando para o ensemble isotérmico-isobárico (NPT) por 250 ps de 100-298,15 K, ambos com controle de temperatura Berendsen (Berendsen *et al.*, 1984). Após o aquecimento, o sistema passou por mais 250 ps para equilíbrio da densidade das restrições de posição de 1 kcal/mol para os átomos pesados do complexo e finalmente equilíbrio dos demais parâmetros por 250 ps com restrições de posição de 0,1 kcal/mol para os átomos pesados do complexo, antes de iniciar a etapa de produção por cerca de 20 ns

4.6. Cálculos de superfície eletrostática e vetor campo elétrico ($\vec{E}$)

A preparação dos arquivos para análise dos campos elétricos foi feita utilizando o servidor online PDB2PQR (Dolinsky *et al.*, 2004). O arquivo .pdb das proteínas foi analisado pelo campo de força PARSE e o arquivo de saída foi colocado no formato .pqr, com nomenclatura de átomos no padrão internacional (Markley *et al.*, 1998). O estado de protonação das cadeias laterais dos aminoácidos foi designado com o cálculo do PROPKA nos pHs indicados: 7,2 (*B. cenocepacia*) e 7,5 (*C. phytofermentans* e *Piromyces* sp.). Os campos eletrostáticos foram calculados no programa VMD v1.9.3 com o plugin APBS v1.4.1 para a resolução da equação de Poisson-Boltzman. As caixas foram calculadas automaticamente de acordo com o volume da proteína e mantidas a 298,15 K (25ºC), exceto para *B. cenocepacia* (310,15 K ou 37ºC).

4.7. Análise de conservação de aminoácidos e estruturas 3D

As sequências de aminoácidos de XI foram preditas através da tradução das sequências de nucleotídeos do gene *XylA* de diferentes espécies através do programa ConSurf (Ashkenazy *et al.,* 2010 ; Celniker *et al.,* 2013). A busca das sequências foi feita com a heurística do CSI-BLAST contra a base de dados curada do SWISS-PROT. Um máximo de 150 sequências únicas foi selecionado com intervalos de identidade entre 35 e 95%. O alinhamento foi realizado com o algoritmo MUSCLE. As árvores filogenéticas foram geradas utilizando o método de cálculo Bayesiano com o modelo de substituição JTT. A partir de uma estrutura tridimensional e o cálculo de conservação evolutiva, o programa mostra em um mapa de cores a conservação e funcionalidade dos resíduos na determinada proteína.

4.8. Análise Estatística

Todos os valores obtidos nos ensaios de bancada foram analisados pelo programa GraphPad Prism 6 e os valores se mostraram estatisticamente relevantes dentro de um intervalo de confiança de 95%.

5. RESULTADOS E DISCUSSÃO

5.1. Purificação e análise funcional da xilose isomerase de *B. cenocepacia*

Até o presente momento, todas as tentativas de medir a atividade de xilose isomerase foram feitas utilizando extratos celulares. Porém, a metodologia mais difundida na literatura utiliza uma reação acoplada para a medida da atividade enzimática na presença de todo o conteúdo citoplasmático, o NADH pode ser oxidado por diversas proteínas diferentes da xilose isomerase. Sendo assim, é de grande interesse que se façam novas análises das propriedades cinéticas destas enzimas tanto *in vivo* quanto *in vitro.* Considerando que (Vilela et al., 2015) demonstraram a funcionalidade deste modelo *in vivo*, este trabalho teve por objetivo avaliar a funcionalidade desta proteína em condições ótimas.

Para isso, a expressão da XI de *Burkholderia cenocepacia* foi feita utilizando o plasmídeo pTRC-His (Thermo Scientific) transfectado na bactéria *Escherichia coli* TOP10. Este plasmídeo possui o promotor do operon Triptofano e o operador *lac*, que permite a modulação da expressão gênica por lactose ou também pelo indutor fortuito isopropil β-D-

1-tiogalactopiranosídeo (IPTG). Além disso, este vetor possui uma cauda de 6 histidinas *in tandem* conectadas ao N-terminal da proteína. Isto confere à proteína características físico química e estruturais de interesse para as etapas seguintes de purificação.

Primeiramente, a proteína é expressa a 18°C a fim de manter a sua solubilidade em altas concentrações, pois nesta temperatura o metabolismo da bactéria é reduzido e suas funções ocorrem de maneira mais vagarosa. Sendo assim, o enovelamento de proteínas possui mais tempo para ocorrer e se torna mais eficiente (Wu *et al.,* 2016). A seguir, as células são lisadas por ultrassom e mantidas a 4 ºC a fim de evitar sua degradação por ação de proteases contaminantes. Este lisado celular é então incubado na presença de uma resina de sefarose contendo níquel (Ni^{2+}) reticulado, permitindo a separação de proteínas com alto conteúdo de His, efeito da interação do anel imidazólico presente na cadeia lateral deste aminoácido com o íon. A eluição da proteína é feita em tampão contendo imidazol livre (100 a 300 mM), que compete pela ligação ao níquel com a proteína, fazendo com que a mesma seja eluída junto à fase móvel. Em seguida, as diferentes frações de lavagem são analisadas em gel de acrilamida desnaturante SDS-Page 12% como visto na **Figura 10**, onde busca-se obter proteínas de peso molecular aproximado de 51 kDa, equivalente ao peso de 50 kDa da XI e 1 kDa da cauda poli histidina. Através da regressão linear observou-se inicialmente que o peso aproximado da banda corresponde à XI era de 55 kDa.

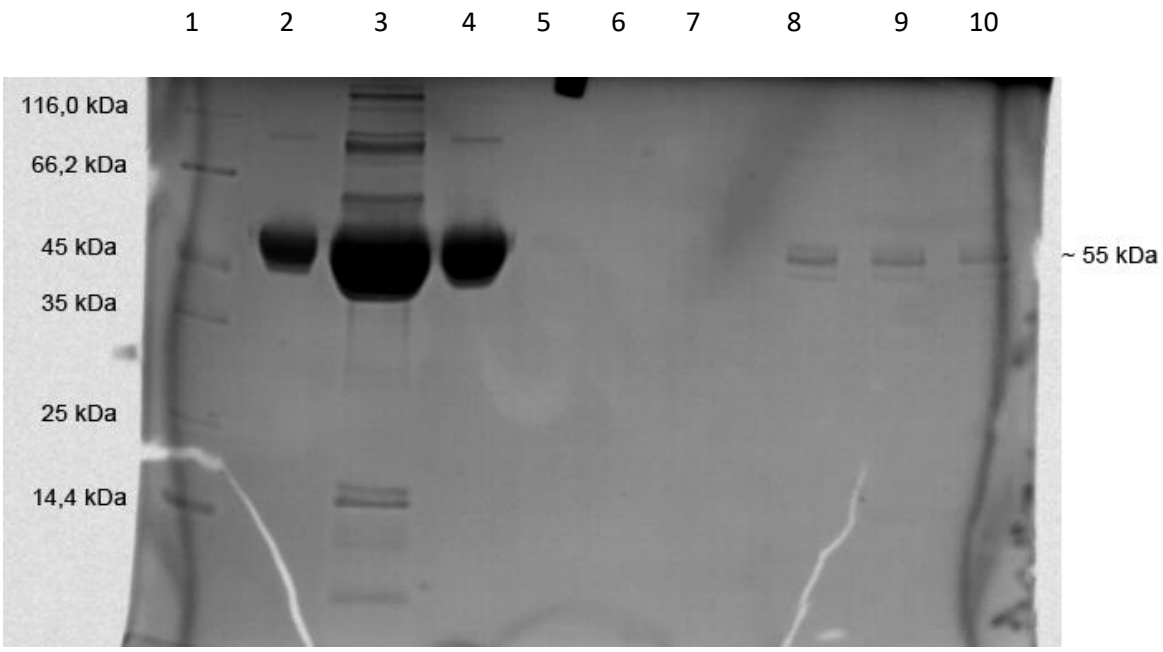

Figura 10: Gel de acrilamida SDS-Page 12% com amostras das frações de lavagem da resina de níquel. Poços: (1) padrão de peso molecular; (2) primeira eluição com tampão de lavagem; (3) segunda eluição com tampão de lavagem; (4) terceira eluição com tampão de lavagem; (5 e 6) 100 mM imidazol; (7 e 8) 200 mM imidazol; (9 e 10) 300 mM imidazol.

A amostra enriquecida é separada do restante e suas frações são concentradas por exclusão molecular na membrana Amicon Ultra 30K, permitindo somente a passagem de moléculas menores que 30.000 Da de massa molecular. A amostra concentrada é aplicada em uma coluna contendo a resina Superdex 200 de volume de cama 200 mL, onde há a separação das amostras de acordo com seu tamanho e volume. O espectro de proteínas é acompanhado através da absorbância sob a incidência de luz ultravioleta e o cromatograma é formado como indicado na **Figura 11**. Por último, os picos são analisados individualmente em gel SDS-Page 12% e a atividade enzimática é mensurada. A concentração de proteínas foi mensurada antes e após a etapa de gel filtração, onde partiu de 2 mg/mL e foi reduzida a aproximadamente 1 mg/mL. É importante notar que a proteína apresenta duas bandas com pesos moleculares equivalentes a 55 e 110 kDa. Isto é um indicativo de que esta proteína poderia ser capaz de formar homodímeros, característica comum entre proteínas desta família (Peng *et al.*, 2015).

A caracterização funcional da enzima xilose isomerase de *B. cenocepacia* pode ser visualizada na **Figura 12**. Para isso, a partir de uma amostra purificada de XI, a atividade enzimática específica foi determinada em condições variantes de pH e temperatura. Com

a finalidade de determinar a temperatura ótima, o mix reacional foi mantido tamponado a pH 7,5 seguindo protocolos já publicados anteriormente (Brat et al., 2009). A atividade enzimática foi determinada a 25°C, 30°C e 37°C. A **Figura 12A** mostra que a temperatura possui um grande impacto na atividade desta proteína. O maior valor de atividade foi encontrado a 37°C, sendo ele 22,47 $\mu mol.min^{-1}.mg\ ptn^{-1}$. Este resultado mostra que esta enzima possui um potencial para utilização em condições semelhantes às encontradas em dornas de fermentação nas indústrias, onde durante a fermentação podem haver grandes variações de temperatura, entre 35°C e 42°C. (Basso *et al*., 2011).

Para avaliar a influência do pH na atividade de XI, diversos mix reacionais foram preparados para tamponar a reação em um intervalo de 5,8 a 7,8. Para isso, foi necessário fazer a mudança da proteína para os diferentes tampões através de lavagens e centrifugações na Amicon Ultra 30K. No intervalo de pH entre 5,8 e 7,8 foi utilizado o tampão Bis-Tris contendo 70 mM de NaCl para a manutenção da força iônica (87 mM) constante em todos os ensaios. Os experimentos foram realizados em temperatura ótima de 37 °C.

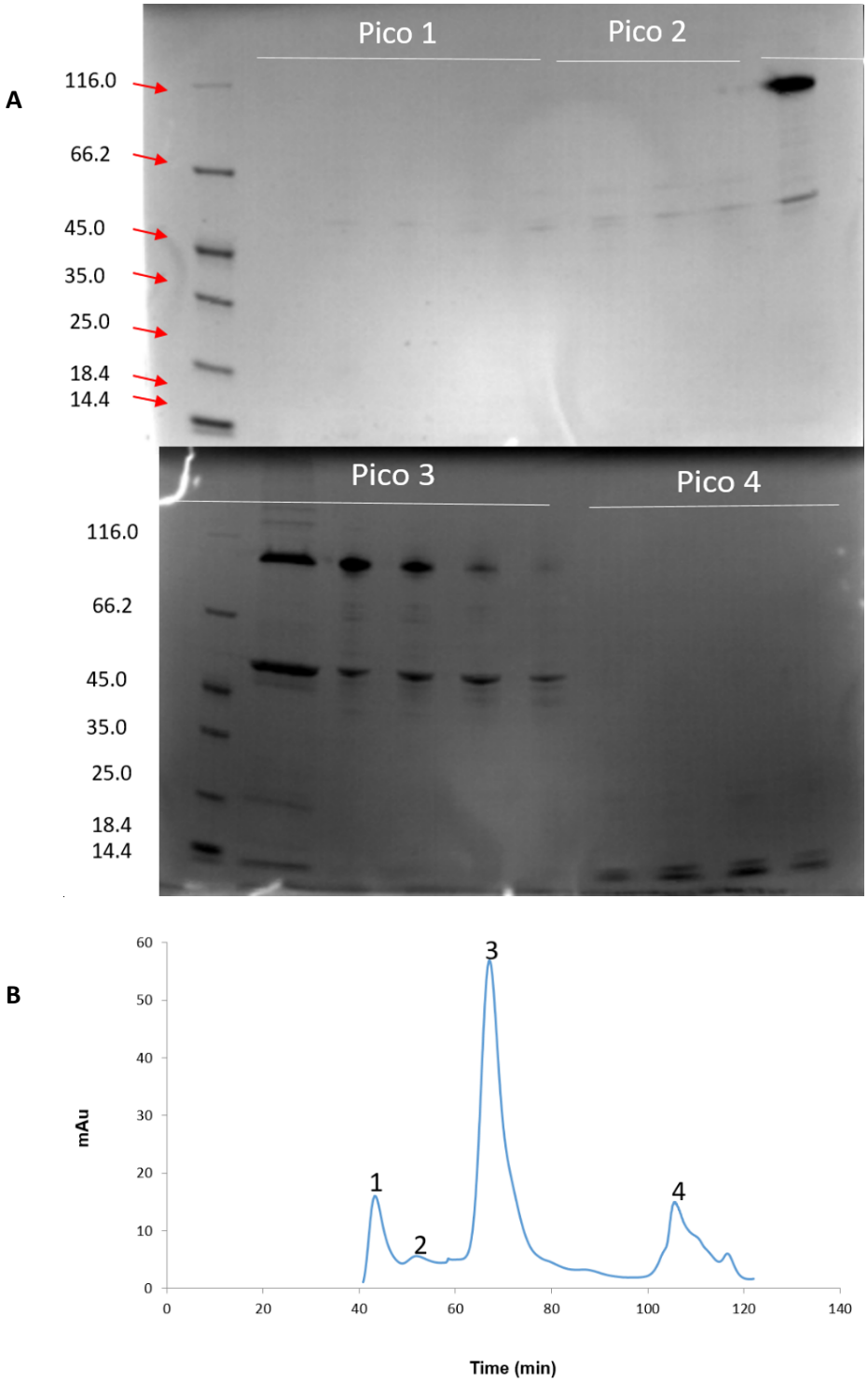

Figura 11: Separação das frações de proteína por volume molecular em resina Superdex 200. As frações de proteínas são coletadas de acordo com seu tempo de retenção demonstrado no cromatograma e diferentes porções dos picos são analisadas em gel de acrilamida SDS-Page 12%. A correlação entre as amostras no gel em A e os picos em B está sinalizada pela barra branca acima dos poços.

A **Figura 12B** mostra que a atividade enzimática é fortemente influenciada por pequenas mudanças no pH da solução, tendo sua maior taxa de conversão (5,16 μmol.min^{-1}.mg ptn^{-1}) a pH 7,2. Portanto, uma vez estabelecida as condições ótimas de reação, o K_m e o V_{max} da XI de *B. cenocepacia* foi calculado a partir da regressão não linear da curva de Michaelis-Menten (**Figura 12C)** e da regressão linear do gráfico de Lineweaver-Burk, também conhecido como duplo recíproco (**Figura 12D**). Calculou-se que o K_m desta proteína para a D-xilose é de 26,43 ± 8,71 mM e V_{max} é 50,65 ± 3,83 μmol.min^{-1}.mg ptn^{-1}. Estes são resultados significativos se comparados aos melhores dados já publicados **Tabela 2**. (Brat *et al*., 2009) concluíram que o K_m da XI de *Piromyces* sp. possui um valor de 49,85 ± 2,82 mM e o de *C. phytofermentas* é de aproximadamente 61,85 ± 3,41 mM. Desta forma, a XI de *B. cenocepacia* possui um K_m aproximadamente três vezes menor para a xilose, característica que pode ter um resultado notável no metabolismo de xilose da levedura *S. cerevisiae*. Conclui-se que esta proteína necessita de uma concentração intracelular deste açúcar significativamente menor para atingir ½ V_{max} e consequentemente competir pela xilose com a XR (Km para xilose 14,8 ± 2,4 mM ; Jeong *et al.*, 2001), assim prevenindo o acúmulo de xilitol.

O xilitol é um açúcar comumente utilizado pela indústria alimentícia e está presente em baixas concentrações em fibras e frutas. Normalmente a produção deste carboidrato se dá pela extração da xilana, proveniente da hemiceluose, que é hidrolisada em xilose e hidrogenada cataliticamente a xilitol. Já na via biológica em *S. cerevisiae*, este é um subproduto comumente encontrado em modelos de fermentação de etanol 2G. Isso se dá pela ineficiência das enzimas das etapas subsequentes em manter o fluxo metabólico até a formação de etanol (Klimacek et al., 2010). Além disso, este carboidrato é uma das maiores preocupações para pesquisadores que utilizam a estratégia da XI, pois é um natural potente inibidor desta enzima. Então, decidiu-se testar o impacto da presença do xilitol na reações contendo a XI de *B. cenocepacia*. Para isso, manteve-se a concentração mínima de xilose necessária para atingir V_{max} (125 mM) e variou-se a concentração de xilitol de 10 a 50 mM de acordo com a metodologia utilizada por (Brat et al., 2009). Na **Figura 13** observa-se que até mesmo pequenas concentrações de xilitol (10 mM) podem afetar drasticamente a atividade enzimática, reduzindo-a em cerca de 50%. A teoria mais aceita é de que este açúcar é um inibidor competitivo da xilose (Fenn *et al.*, 2004).

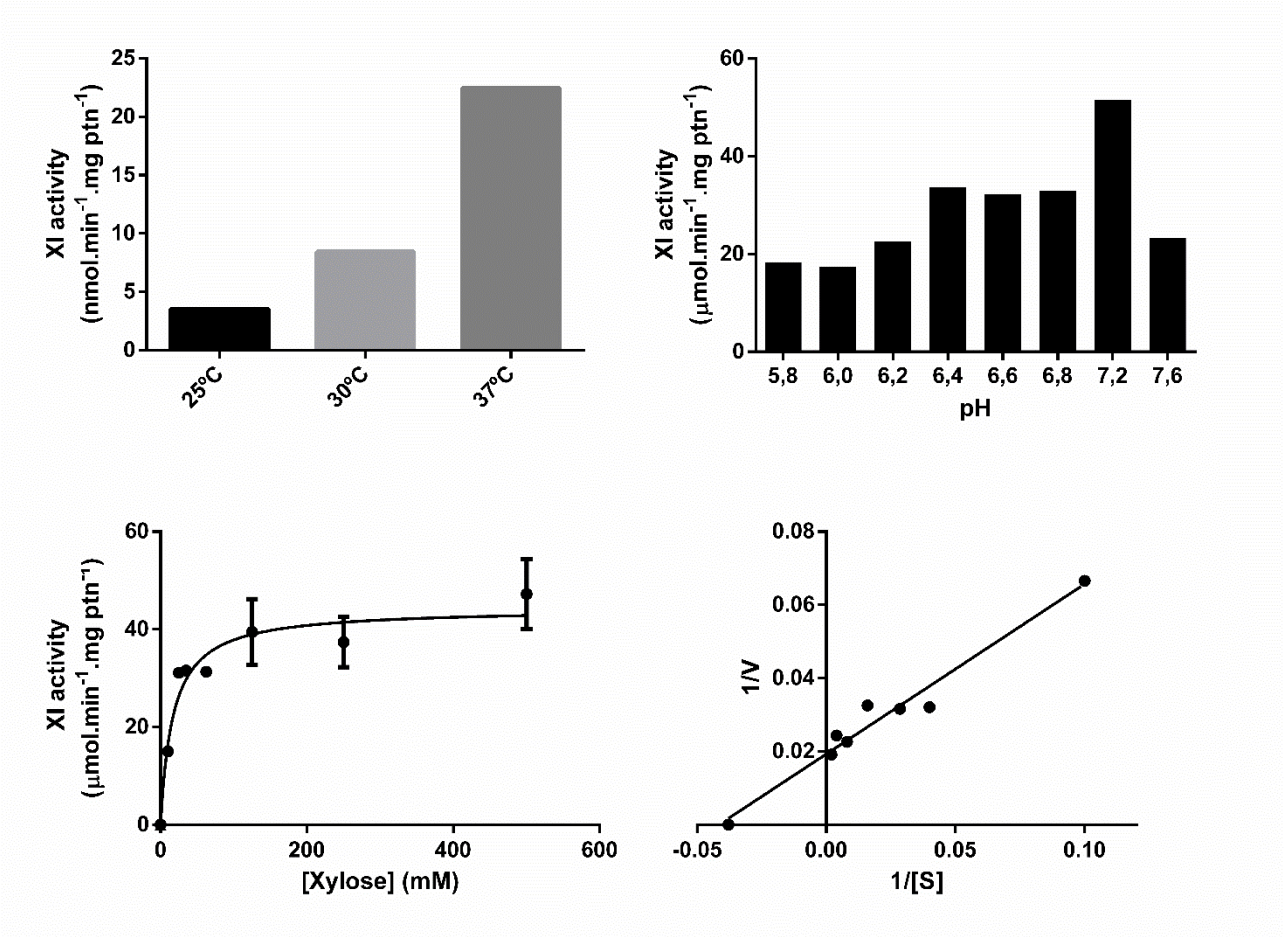

Figura 12: Caracterização funcional da XI de *B. cenocepacia*. (A) Atividade enzimática em diferentes temperaturas. (B) Atividade enzimática em tampões com variação de pH. (C) Atividade enzimática com diversas concentrações de xilose em condições ótimas. (D) Representação gráfica de Lineweaver-Burk para os dados apresentados em C.

Porém, maiores concentrações deste inibidor não parecem ter um efeito dose-dependente linear. Ademais, estudos sugerem que a inibição por xilitol pode não ser tão severa *in vivo* quanto é demonstrada *in vitro*, pois a capacidade de difusão destas moléculas no citoplasma é bastante reduzida devido à pressão osmótica (Ha *et al*., 2011). Além disso, este efeito não se mostra significativo frente ao resultado apresentado por (Vilela et al., 2015), que demonstra que neste modelo não há acúmulo de xilitol durante a cofermentação de glicose e xilose, mostrando assim que a XI de *B. cenocepacia* possui um grande potencial para futuras aplicações na indústria do etanol 2G.

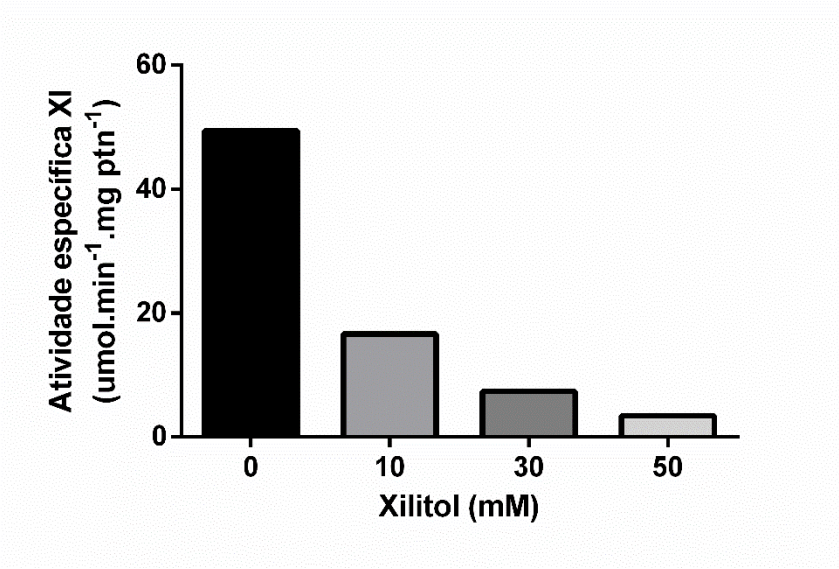

Figura 13: Inibição da XI de *B. cenocepacia* na presença de diferentes concentrações de xilitol. Os ensaios foram feitos com uma concentração de 125 mM de xilose, 37°C, pH 7,2.

Tabela 2: Valores de referência para a atividade enzimática da XI de três diferentes espécies.

Espécie	Km (mM)	Referência
B. cenocepacia	17,08 ± 5,445	Este trabalho
C. phytofermentans	61,85 ± 3,41	(Brat et al., 2009)
Piromyces sp.	49,85 ± 2,82	(Brat et al., 2009)

5.2. Predição da estrutura tridimensional da XI

A busca por homólogos da enzima xilose isomerase de *B. cenocepacia* foi feita contra as sequências publicadas no Protein Data Bank (PDB) usando o software online BLAST alocado no servidor do NCBI. As sequências provenientes dos cristais com identidade acima de 50% foram considerados os mais adequados para a modelagem com o software Modeller v9.17 (**Tabela *3***). Sendo assim, o alinhamento automatizado das sequências foi feito pelo módulo sAlign do Modeller em suas configurações padrão das matrizes de pontuação [overhang=30, gap penalties 1d= (-450,-50)] . A fim de obter uma estrutura tridimensional de melhor qualidade, os alinhamentos gerados foram editados manualmente para melhor adequar o posicionamento dos resíduos próximos aos terminais da proteína (**Figura S 2**, **Figura S 3** e **Figura S 4**). Sendo assim, foram gerados 500 modelos teóricos para cada proteína e o mais provável selecionado como descrito no item **4.4**. Para fins de

comparação do mecanismo de ação, a mesma metodologia foi utilizada para gerar modelos das XI de *C. phytofermentans* e *Piromyces* sp. Estas espécies foram escolhidas pois possuem um dos melhores resultados já publicados até o momento (Brat et al., 2009).

Tabela 3: Dados de BLAST dos moldes escolhidos para cada sequência alvo.

Sequência alvo	Cobertura (%)	Identidade (%)	Cód. de acesso
B. cenocepacia	99	52	1a0c
	98	53	1a0d
	98	51	1a0e
C. phytofermentans	99	62	1a0c
	99	61	1a0d
	99	60	1a0e
Piromyces sp.	99	49	1a0c
	99	52	1a0d

Apesar da proteína de código de acesso PDB: 4xkm possuir uma cobertura de 99% e identidade de 84% com a sequência alvo da XI de *Piromyces* (resultado não mostrado), o modelo biológico construído (*biological assembly*) para este cristal possui falhas fundamentais na representação de cadeias e rotâmeros que poderiam inserir um viés na construção dos modelos teóricos. Sendo assim, a utilização desta proteína como molde foi descartada.

Em relação à semelhança entre moldes e modelos, os valores de RMSD entre carbonos α da cadeia principal se mostraram baixos para a XI de *B. cenocepacia* (1a0c=0,230 Å; 1a0d=0,276 Å), *C. phytofermentans* (1a0c=0,467 Å ; 1a0d=0,378 Å ; 1a0e=0,379 Å) e *Piromyces* (1a0c=0,581 Å; 1a0d=0,584 Å). Além disso, os gráficos de Ramachandran indicam que menos de 0,5% dos resíduos de aminoácidos possuem rotâmeros em posições proibidas, 93% em posições permitidas e 0,5% em regiões generosamente permitidas, levando a crer que as estruturas secundárias formadas estão em sua conformação correta. Ademais, também foi confirmada a correta inserção das

estruturas secundárias em cada subunidade. Para isso, foi utilizado o software PSIPRED, que faz a predição da formação de estruturas secundárias a partir da sequência primária de aminoácidos, utilizando dados de sistemas de rede neural e do alinhamento gerado pelo PSI-BLAST, obtendo um acerto comprovado de aproximadamente 81,6% dos casos (Jones, 1999). O resultado mostra que cerca de 95% dos aminoácidos estavam com a predição de estrutura secundária correta. Os 5% restantes correspondiam aos aminoácidos em regiões de transição entre estruturas randômicas e secundárias (alças e voltas) (**Figura *S 7***, **Figura *S 8*** e **Figura *S 9***).

A alta semelhança da cadeia principal das estruturas terciárias levantaram questões em relação à metodologia utilizada para a modelagem das proteínas. Trata-se da: (i) criação de viés por utilização dos mesmos moldes; (ii) alta semelhança entre sequências. Para responder a primeira pergunta, a heurística utilizada pelo software Modeller foi analisada. Segundo o autor (Martí-Renom et al., 2000), o programa utiliza as coordenadas dos átomos presentes no arquivo de extensão .pdb para gerar um arquivo secundário de restrições espaciais a serem respeitadas pelo programa. Caso o rotâmero de um átomo ou resíduo de aminoácido não consiga ser modelado nestas condições, o programa então tenta encaixá-lo de diferentes maneiras que sirvam àquele ambiente comandado pelas nuvens eletrônicas. Sendo assim, elimina-se o viés que poderia ser gerado pela utilização dos mesmos moldes, pois o posicionamento de cada átomo pode ser modificado para melhor acomodar cada átomo, que levaria a diferenças nos modelos construídos.

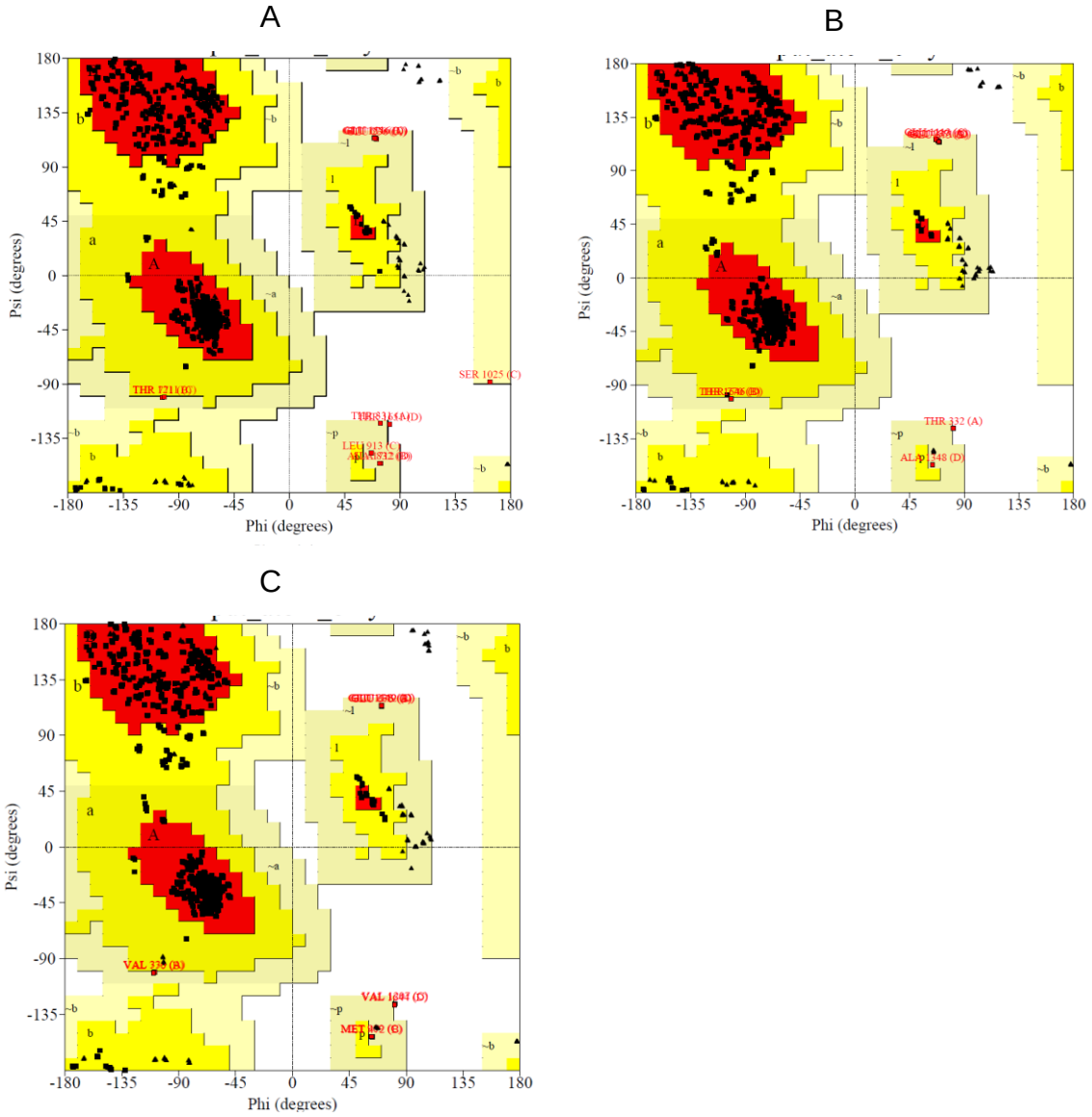

Figura 14: Gráfico de Ramachandran para os modelos teóricos selecionados. (A) *B. cenocepacia* (B) *C. phytofermentans* (C) *Piromyces sp.* Pontos pretos indicam resíduos de aminoácidos. *Triângulos* representam Glicinas e Prolinas. Regiões favoráveis estão representadas em vermelho. Regiões permitidas estão representadas em um gradiente de amarelo, onde o menos intenso indica menor probabilidade.

Já no segundo caso, é interessante o fato de estruturas tão similares possuírem atividades enzimáticas distintas mesmo quando possuem uma sequência primária tão distinta. A análise de conservação dos resíduos de aminoácidos e da estrutura indicam que as regiões externas de expostas ao solvente possuem uma taxa de substituição maior em relação à região de barril do sítio ativo (**Figura *15***).

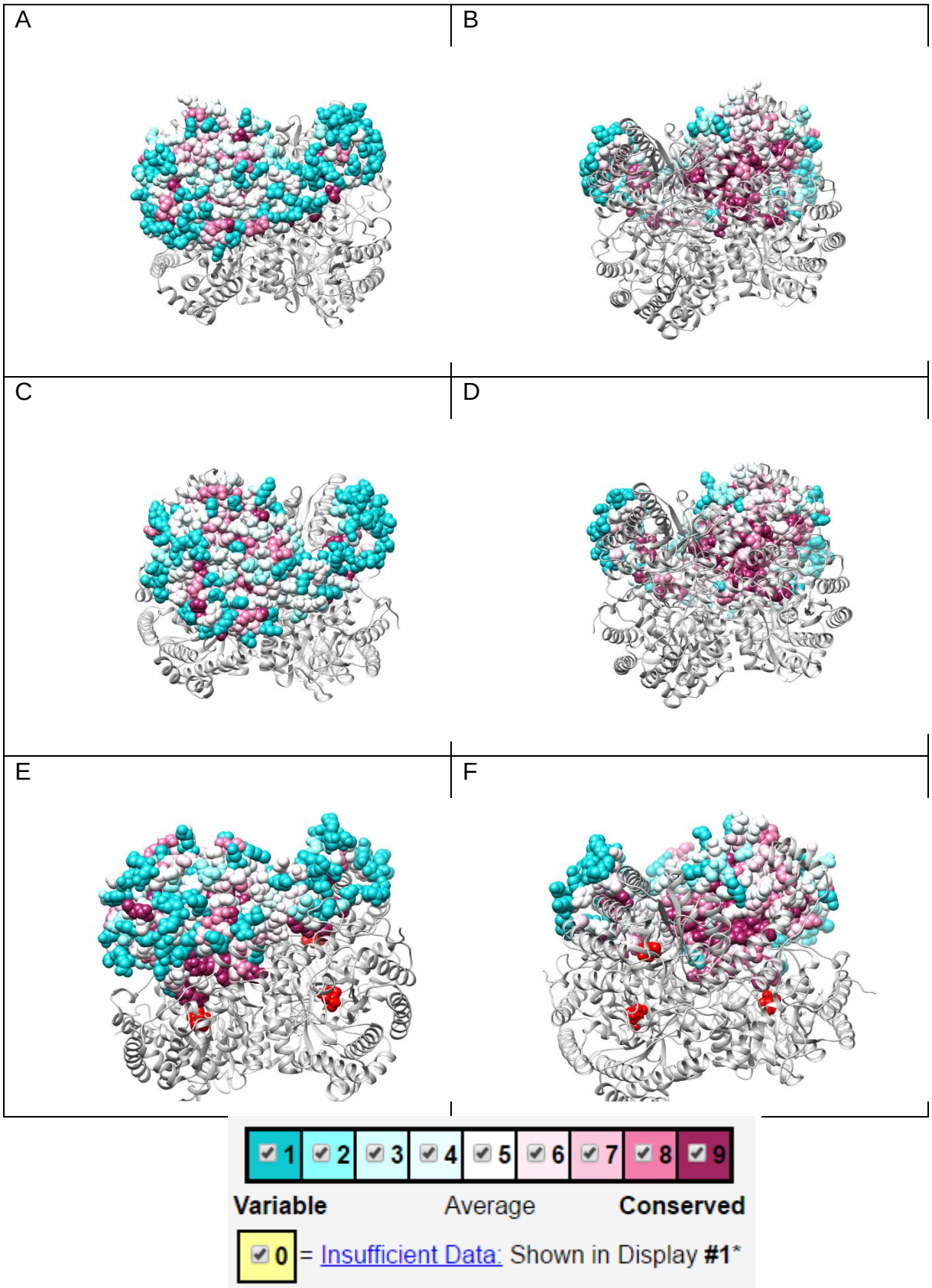

Figura 15: Conservação de resíduos na família de proteínas XI. (A e B) *B. cenocepacia*, (C e D) *C. phytofermentans* e (E e F) *Piromyces*. A árvore filogenética contendo 150 genes homólogos à *XylA* das espécies indicadas foi utilizada para avaliar a conservação de

aminoácidos na estrutura tridimensional das proteínas. É levado em consideração a posição dos resíduos, característica físico química e taxa de substituição.

Outra hipótese considerada estaria baseada na diferença de dobra das cadeias laterais do sítio ativo. De forma geral, sabe-se que a sequência primária de uma proteína fornece as características físico químicas que irão permitir a dobra de estruturas secundárias específicas, a partir da interação com o solvente. Por sua vez, a estrutura secundária comanda a formação de estruturas globais como alças, barris e etc. Desta forma, inicialmente levou-se a pensar que a estrutura tridimensional do sítio ativo poderia ser diferente entre estas proteínas. No entanto, o alinhamento estrutural mostrado na **Figura 16** indica que até mesmo esta região da proteína é muito similar em todos os casos. Os valores de RMSD estão presentes na **Tabela 4**. Porém, para que esta afirmação seja verdadeira, é necessário que se faça o relaxamento da cadeia principal e dos ângulos das cadeias laterais, que irão alterar levemente a estrutura tridimensional do sítio ativo e, portanto, a ligação aos íons metálicos e ao substrato.

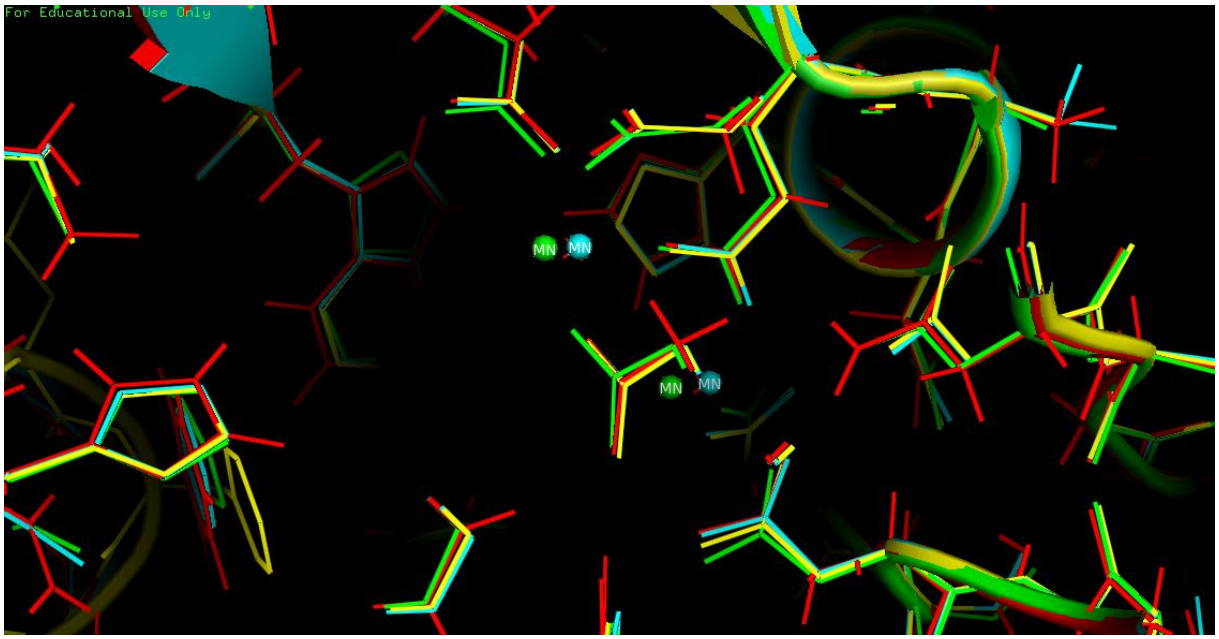

Figura 16: Alinhamento estrutural do sítio ativo da XI de *B. cenocepacia, C. phytofermentas* e *Piromyces* sp. O alinhamento tridimensional mostra grande sobreposição entre os resíduos de diferentes proteínas, assim como o posicionamento dos íons metálicos. Isto é um indicativo do grau de conservação mantido nesta região da proteína.

Tabela 4: Dados comparativos de RMSD entre os modelos teóricos de XI.

Modelo	Alvo	RMSD global Cα (A)
B. cenocepacia	*C. phytofermentans*	0,339
B. cenocepacia	*Piromyces*	0,493
C. phytofermentans	*Piromyces*	0,269

Considerando que a metodologia de criação dos modelos utilizada pelo software Modeller se utiliza majoritariamente da teoria dos corpos rígidos, optou-se por fazer o refino das regiões de estrutura randômica através do algoritmo presente no módulo "Loopmodel" do Modeller. Esta etapa é fundamental para que se possa corrigir os rotâmeros em posições não favoráveis. Além disso, isso evita que haja a criação de um movimento artificial durante as simulações de dinâmica molecular. Após a etapa de refinamento, a qualidade de modelo aumentou em cerca de 15%. Estes modelos foram então inseridos em um campo de simulação de dinâmicas molecular para se estudar a interação do complexo proteína-substrato.

5.3. Simulações de dinâmica molecular e minimização de energia dos complexos proteicos de *B. cenocepacia*, *C. phytofermentans e Piromyces* sp.

Através de simulações de dinâmica molecular pretendemos refinar os modelos moleculares das XI de *B. cenocepacia*, *C. phytofermentans e Piromyces* sp. a fim de compreender melhor as diferenças do mecanismo de ação das proteínas e, consequentemente, nas atividades enzimáticas pelo substrato xilose. A partir das estruturas dos modelos, foram preparados complexos com uma molécula de xilose em sua forma cíclica (β-D-xilopiranose). As simulações foram realizadas mimetizando as condições dos ensaios de bancada. Para isso, os sistemas foram feitos na presença de água a pH 7,2 para *B. cenocepacia* e pH 7,5 para *C. phytofermentans* e *Piromyces* sp., a uma temperatura de 298,15 K (25ºC) e 1atm, otimizados e equilibrados. As simulações de produção foram realizadas por 20 ns e analisadas em duas etapas: (i) foco no comportamento geral dos modelos, verificando a estabilidade da estrutura e convergência das propriedades termodinâmicas; (ii) geometria do sítio ativo, descrevendo a interação com a xilose e investigando as possíveis razões para a diferença de Km para xilose entre as enzimas.

5.3.1. Análise da estabilidade dos complexos proteína-substrato

Análises globais*:** As simulações foram inspecionadas visualmente e não foram observadas transições conformacionais marcantes. A molécula de xilose permaneceu estável nos sítios ativos. A estabilidade global foi monitorada usando como métrica para as análises de RMSD, RMSF e manutenção do conteúdo de estrutura secundária. A análise de RMSD (**Figura *17) revelou um relaxamento semelhante para as estruturas, em torno de 2 A em relação ao modelo inicial, consistente com o comportamento em água, que geralmente varia entre 1 a 4 A. A análise de RMSF (**Figura 18**) revelou que as regiões de maior flutuação na estrutura estão concentradas nas extremidades e *loops* (área de estrutura randômica) expostos ao solvente.

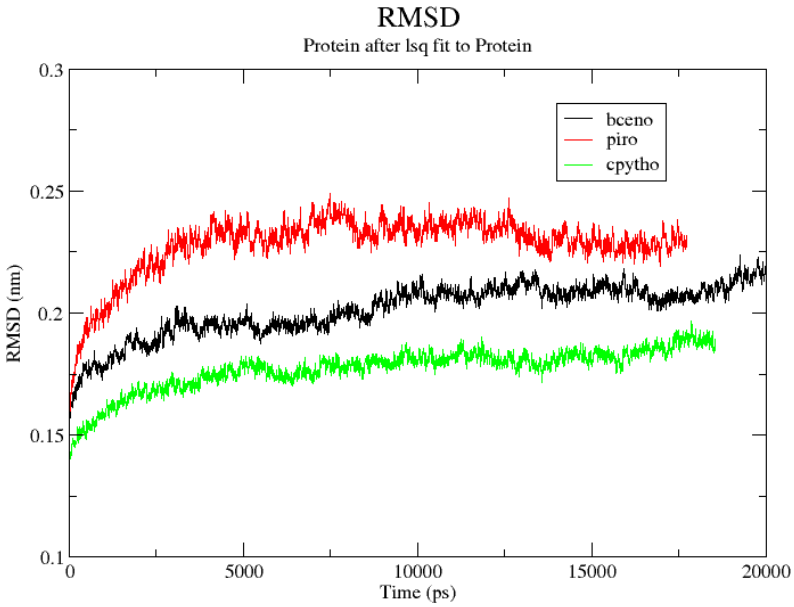

Figura 17: RMSD do esqueleto peptídico em relação à estrutura inicial durante a etapa de produção. O gráfico indica que houve uma pequena variação global da estrutura durante toda simulação, sem haver grandes flutuações.

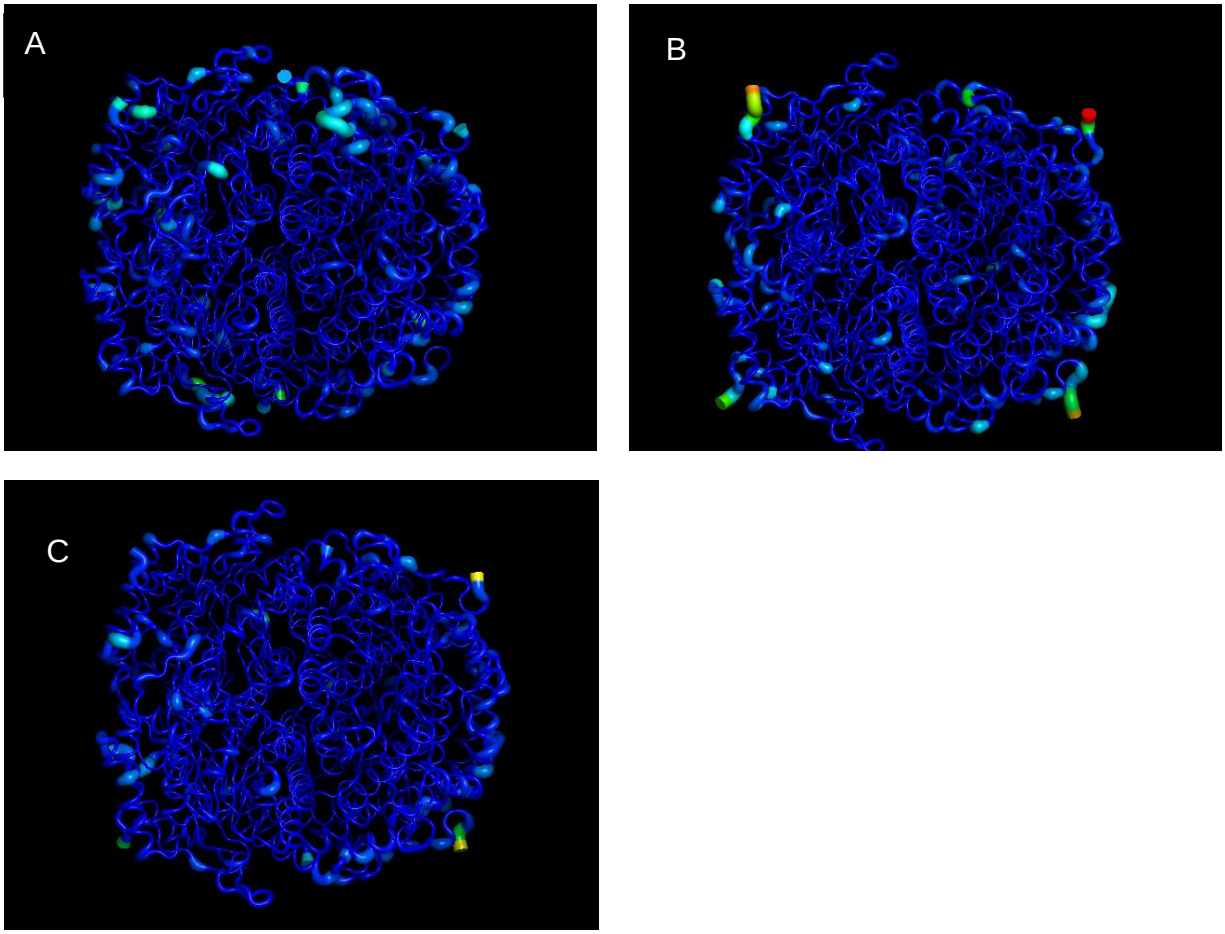

Figura 18: RMSF 3D para *B. cenocepacia* (A), *Piromyces sp.* (B) e *C. phytofermentans* (C). As estruturas estão representadas em tubos, onde a espessura e cor indicam o grau de flexibilidade. O cromatograma indica o grau de flututação das regiões indicadas em A,B e C ao longo do tempo. As regiões em azul correspondem à estruturas com pouca movimentação enquanto regiões coloridas em tons vermelhos possuem maior movimentação.

Estruturas secundárias*:** A análise da conservação de estrutura secundária (**Figura *19) indicou leve redução do conteúdo global de estruturas secundárias (cerca de 2,5%, em média). O detalhamento da variação de cada tipo de estrutura secundária está presente na **Figura S 10.** Por fim, tais análises indicam que os modelos construídos são estáveis e viáveis, apresentando uma convergência aparente em torno de 10 ns de simulação.

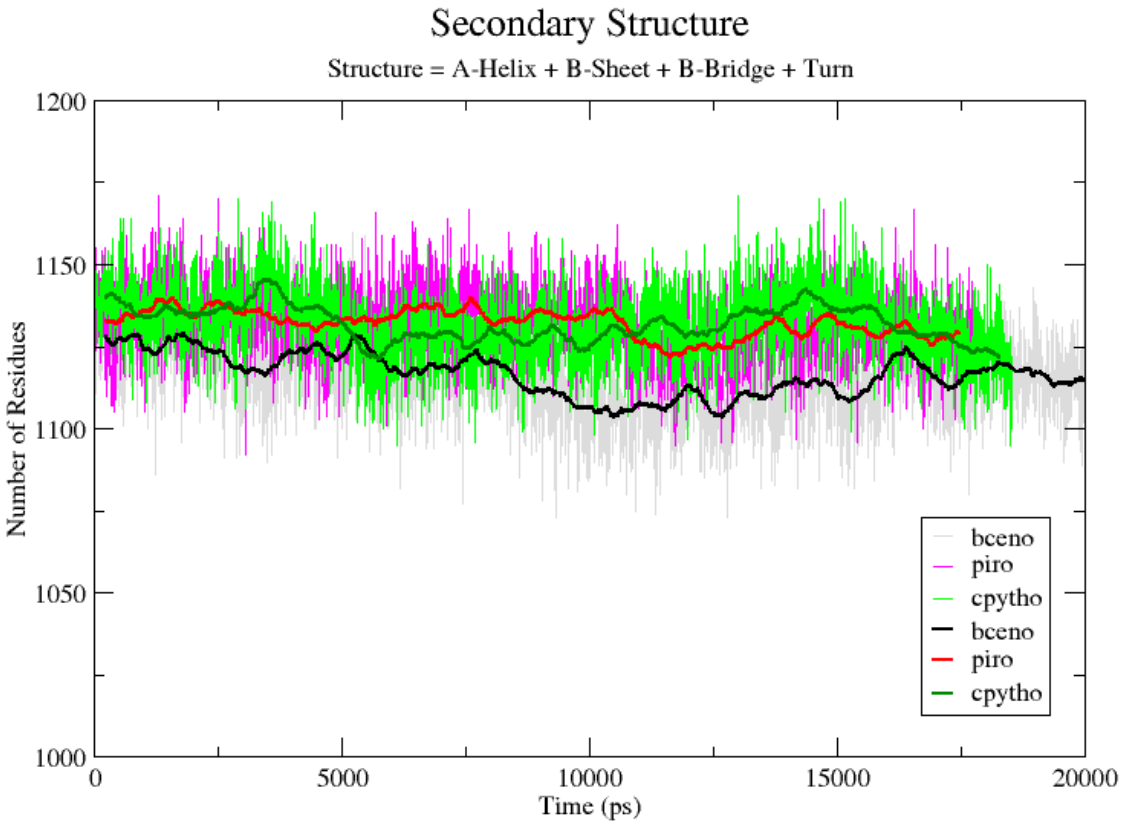

Figura 19: Gráfico de variação do conteúdo total de estrutura secundária. O cromatograma indica o número total de resíduos que variaram sua estrutura secundária ao longo do tempo. As linhas representam a média numérica destes valores.

Interações com a xilose*:** O gráfico de RMSF mostra o grau de flutuação do ligante dentro do sítio ativo ao longo do tempo de simulação (**Figura *20). Percebe-se que há um baixo grau de flutuação do anel piranosídico durante toda a simulação. No entanto, as hidroxilas das posições 0, 15, 17 e 19 possuem um alto grau de variação. Este resultado é esperado, uma vez que estes radicais possuem um grau de liberdade maior dado pelos ângulos de torção da ligação C-OH.

No entanto, há uma clara orientação do substrato durante a interação com as cadeias laterais dentro do sítio ativo e o íon metálico. Essa orientação espacial se dá por interações eletrostáticas, as águas presentes ajudam na coordenação do metal através de uma interação íon-dipolo gerada pelo momento dipolar da água e os íons metálicos, por sua vez, fazem a coordenação de duas hidroxilas, fazendo seu ancoramento. Esse

posicionamento obriga que o oxigênio piranosídico (O13) esteja sempre apontando na direção da histidina presente no sítio ativo (**Figura *21***). Este mecanismo de ligação é semelhante ao descrito por (Fenn et al., 2004; Toteva et al., 2011), onde a histidina pode possuir papel relevante de duas maneiras: (i) em seu estado desprotonado, esta cadeia lateral poderia servir como uma carga líquida formal para cadeias laterais adjacentes promoverem a hidrólise do anel piranosídico; ou (ii) pode assumir o papel de resíduo catalítico ao ser o ácido envolvido na doação de prótons para o oxigênio piranosídico.

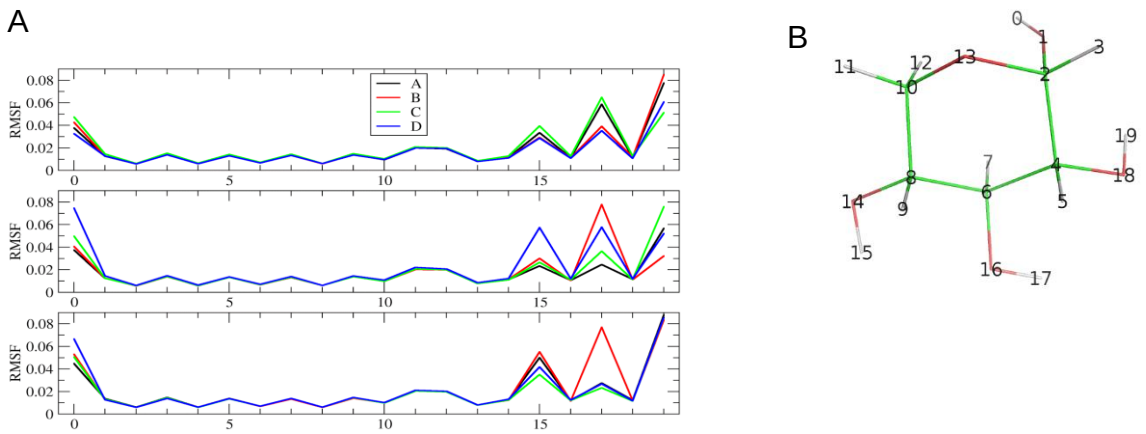

Figura 20: RMSF para os átomos do ligante. A) RMSF para cada átomo de xilose. (1A) *B. cenocepacia*; (2A) P*iromyces* sp.; (3A) *C. phytofermentans*. B) Mapa da numeração dos átomos de xilose.

Desta forma, propõe-se que o distanciamento entre a xilose e a histidina causa uma mudança na probabilidade a qual o próton pode ser doado entre os átomos de oxigênio e nitrogênio. No entanto, este mecanismo não é suficiente para promover a total inatividade da enzima pois através da presença de moléculas de água no sítio ativo seria possível que o *shift* de hidreto ocorra através da ionização da água, que pode assumir a forma intermediária de íon hidrônio (H_3O^+).

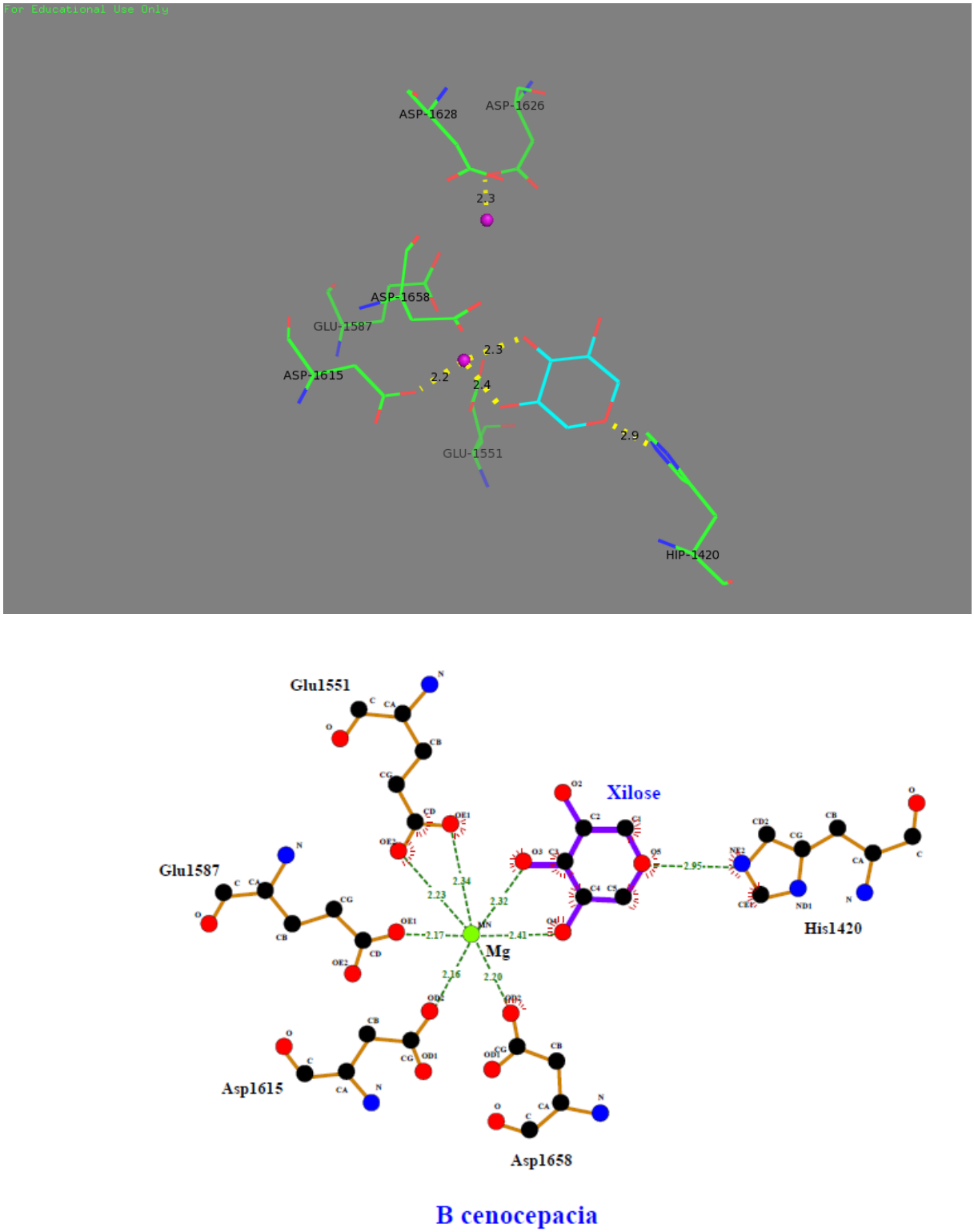

Figura 21: Interação da xilose com o sítio ativo na proteína XI de *B. cenocepacia.* (A) Representação tridimensional do sítio ativo. (B) Representação planar do sítio ativo. Os hidrogênios e um dos íons metálicos estão escondidos para melhor visualização. A distância entre átomos está indicada em Angstrons.

***Interações entre íons metálicos e aminoácidos*:** Através da visualização de movimentação dos íons metálicos durante a simulação, percebe-se que o M_1 possui uma variação de posição muito menor em relação a M_2. Este grau de liberdade pode estar diretamente relacionado com seu papel dentro do sítio catalítico. Acredita-se que M_1 possua um papel estrutural na manutenção da correta geometria do sítio ativo, assim como na orientação do substrato no momento anterior à catálise. A interação da xilose com M_1 através dos radicais hidroxila gera um ancoramento de parte da molécula, que faz com que o lado oposto da molécula aumente a amplitude do movimento. Pode-se hipotetizar, então, que este metal é essencial para a ligação do substrato no sítio ativo, pois o mesmo torna mais fraca a estrutura em torno do oxigênio piranosídico, facilitando a reação de linearização do açúcar. Com esta afirmativa, busca-se contribuir com uma pista relevante para a elucidação do mecanismo de ligação do substrato, ainda não descrito completamente.

5.4. Cálculos de superfície eletrostática

Nota-se que a família das XI possuem uma grande variação em sua sequência primária. No entanto, percebe-se que quase todas as substituições ocorrem por aminoácidos de características físico químicas semelhantes. Sendo assim, deve-se considerar que uma mutação que cause uma modificação por um aminoácido de características diferentes pode afetar significativamente na atividade da proteína. Para analisar este efeito, foram feitos cálculos de predição da superfície eletrostática nos modelos teóricos com energia minimizada após as simulações de dinâmica, a fim de avaliar a interação dos motivos destas proteínas com o substrato.

Percebe-se que as três proteínas possuem uma carga líquida negativa variando sua distribuição pelos resíduos da camada de solvatação. Porém, na região de acesso ao sítio ativo existe uma concentração de cargas positivas, que podem indicar que a interação entre a proteína e seu substrato se dá pela interação do substrato com estes campos elétricos. Sendo assim, variações no vetor campo elétrico ($\vec{E}$) podem ter grandes impactos na atividade enzimática, pois estariam regulando a acessibilidade do substrato ao sítio ativo.

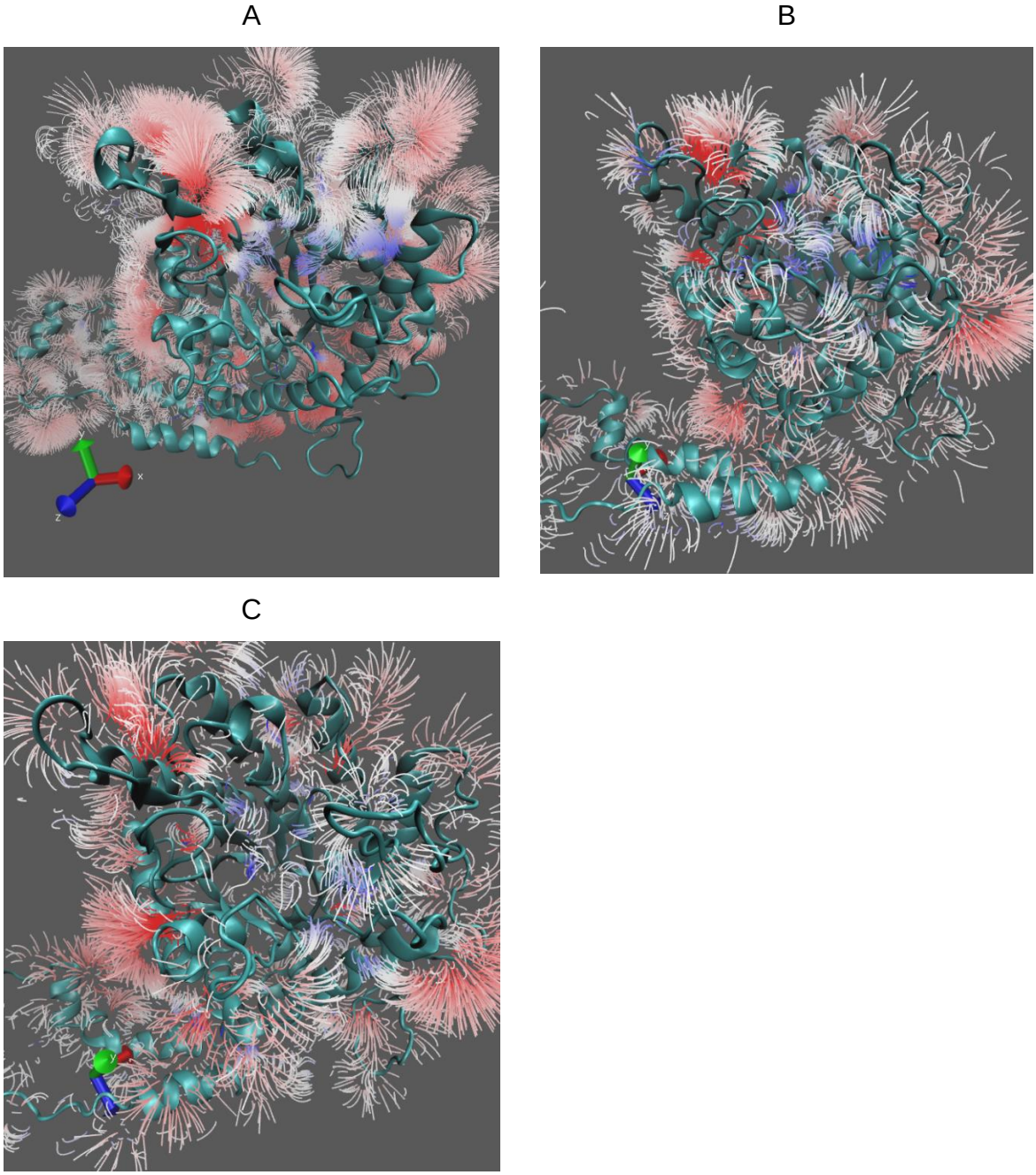

Figura 22: Representação dos campos elétricos das proteínas em condições semelhantes às experimentais. (A) *B. cenocepacia.* (B) *C. phytofermentans.* (C) *Piromyces* sp. Linhas vermelhas representam campos elétricos negativos. Linhas azuis representam campos elétricos positivos. Linhas brancas representam cargas neutras.

As proteínas apresentam uma variação de cargas distribuída entre -10 e 10 kT/e e estão representadas na **Figura *22***. As análises foram feitas em condições semelhantes às encontradas nos experimentos *in vitro* e os campos elétricos foram calculados e distribuídos espacialmente. Os dados apresentados foram convertidos em valores numéricos apresentados na **Figura *23***. Os resultados mostram que pequenas diferenças na

intensidade deste vetor podem influenciar significativamente na atividade enzimática. Estes resultados corroboram com os dados apresentados na **Tabela *2***, onde a XI de *C. phytofermentans* possui o maior Km aparente para xilose e a menor V_{max}.

Surpreendentemente, a XI de P*iromyces* possui uma intensidade maior que a XI de *B. cenocepacia.* No entanto, os resultados *in vitro* não são equiparáveis. Isto pode ser causado por uma leve mudança na direção do vetor, que pode levar a mudanças nas interações possíveis entre proteína e substrato. Além disso, estes resultados podem estar sendo subestimados pela técnica utilizada para mensurar a atividade enzimática destas proteínas. É necessário que novos estudos façam análises detalhadas destas proteínas em condições ótimas, para que se possam definir mais parâmetros que permitam a comparação entre estas proteínas e outras já estudadas por cristalografia de difração de raio X e nêutrons.

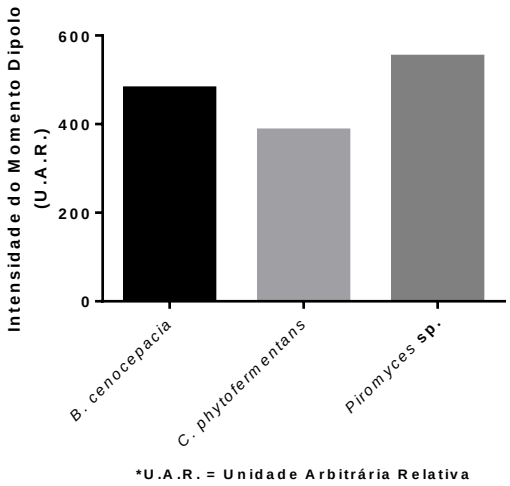

Figura 23: Valores dos vetores campo elétrico $(\vec{E})$ na região de acesso ao sítio ativo.

A caracterização funcional da xilose isomerase de *Burkholderia cenocepacia* em sua forma recombinante, purificada por cromatografia, revelou que esta enzima possui sua atividade ótima a 37°C e pH 7,2. Estes são bons indicativos da versatilidade desta proteína, que é capaz de manter sua atividade nas mais diversas condições as quais as leveduras podem ser expostas no processo de fermentação industrial. Além disso, os valores de Km e V_{max} apresentados neste trabalho corroboram com os resultados obtidos em artigos publicados anteriormente, onde a XI de *B. cenocepacia* possui um Km aparente para a xilose cerca de 3 vezes menor em relação aos melhores resultados publicados. Este é

possivelmente o motivo que explica a ausência de xilitol, em condições de fermentação, pela levedura *S. cerevisiae* recombinante, expressando o gene xylA de *B. cenocepacia*. Acredita-se que esta enzima seria capaz de desviar o fluxo metabólico da via nativa XR/XDH para a via inserida por engenharia metabólica.

Os modelos teóricos gerados para as XI de *B. cenocepacia, C. phytofermentans* e *Piromyces* sp. se mostraram plausíveis em condições biológicas através dos testes de qualidade. Além disso, é possível perceber que mesmo com sequências primárias tão distintas, esta família de proteínas possui uma alta conservação de resíduos de aminoácidos em regiões próximas ao sítio ativo. Com as simulações de dinâmica foi possível perceber que as regiões expostas ao solvente possuem um grau maior de agitação, enquanto regiões próximas ao sítio ativo sofrem poucas modificações estruturais. Já os íons metálicos apresentam diferentes padrões de movimentação entre si que pode estar relacionados com seu papel na catálise da reação de isomerização. Em relação ao sítio de ligação à xilose, este se mostrou altamente conservado entre as espécies analisadas, assim como seu mecanismo de ligação. No entanto, diferenças de carga apresentada por resíduos presentes na entrada do sítio ativo podem alterar a carga líquida do campo elétrico apresentado ao solvente, que influencia diretamente na acessibilidade do substrato ao sítio ativo e pode ser o mecanismo por trás dos resultados observados nos ensaios enzimáticos.

Portanto, se faz necessário estudos futuros com a substituição racional de aminoácidos (mutação sítio dirigida) nos domínios do sítio ativo e de interação com o substrato, para que se possa ter o completo entendimento dos papéis da reação em diferentes escalas, desde a entrada da xilose no sítio ativo até a finalização da reação de isomerização.

6. CONCLUSÕES

- A caracterização funcional da xilose isomerase de *Burkholderia cenocepacia* revelou que esta enzima possui sua atividade ótima a 37°C e pH 7,2.
- Foram encontrados os seguintes valores aparentes de Km 26,43 ± 8,71 mM e V_{max} é 50,65 ± 3,83 μmol.min^{-1}.mg ptn^{-1} para D-xilose para a XI de *B. cenocepacia*.
- Os modelos teóricos gerados para as XI de *B. cenocepacia, C. phytofermentans* e *Piromyces* sp. se mostraram plausíveis em condições biológicas.

- A enzima XI de diversas espécies de possuem grande distinção em suas sequências primárias, porém grande conservação de resíduos de aminoácidos em regiões próximas ao sítio ativo.
- As simulações de dinâmica no tempo estudado mostram que as regiões expostas ao solvente possuem um grau maior de flexibilidade, enquanto regiões próximas ao sítio ativo sofrem poucas modificações estruturais.
- O sítio de ligação à xilose se mostrou altamente conservado entre as três espécies analisadas, assim como seu modo de ligação.
- Diferenças em resíduos presentes na entrada do sítio ativo podem alterar a carga líquida do campo elétrico apresentado ao solvente, que influencia diretamente na acessibilidade do substrato.

7. REFERÊNCIAS

Asbóth, B., & Náray-szabó, G. (2000). Mechanism of Action of D -Xylose Isomerase. *Current Protein and Peptide Science*, *1*(3), 237–254.

Basso, L. C., De Amorim, H. V., De Oliveira, A. J., & Lopes, M. L. (2008). Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Research*, *8*(7), 1155–1163. https://doi.org/10.1111/j.1567-1364.2008.00428.x

Batt, C. A., Jamieson, A. C., & Vandeyar, M. A. (1990). Identification of essential histidine residues in the active site of Escherichia coli xylose (glucose) isomerase. *Biochemistry*, *87*, 618–622.

Binder, J., Blank, J., Cefali, A., & Raines, R. (2010). Synthesis of Furfural from Xylose and Xylan. *ChemSusChem*, *3*(11), 1268–1272. https://doi.org/10.1002/cssc.201000181.Synthesis

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science (New York, N.Y.)*, *253*(5016), 164–70. https://doi.org/10.1126/science.1853201

Brat, D., Boles, E., & Wiedemann, B. (2009). Functional expression of a bacterial xylose isomerase in Saccharomyces cerevisiae. *Applied and Environmental Microbiology*, *75*(8), 2304–2311. https://doi.org/10.1128/AEM.02522-08

Brown, N. a., de Castro, P. a., de Castro Pimentel Figueiredo, B., Savoldi, M., Buckeridge, M. S., Lopes, M. L., … Goldman, G. H. (2013). Transcriptional profiling of Brazilian Saccharomyces cerevisiae strains selected for semi-continuous fermentation of sugarcane must. *FEMS Yeast Research*, *13*(3), 277–290. https://doi.org/10.1111/1567-1364.12031

Bruinenberg, P. M., Jonker, R., Dijken, J. P. Van, & Scheffers, W. A. (1985). Utilization of formate as an additional energy source by glucose-limited chemostat cultures of *Candida utilis* CBS 621 and *Saccharomyces cerevisiae* CBS 8066. *Archives of Microbiology*, *1442*(i 985), 302–306. https://doi.org/10.1007/BF00693408

Cai, Z., Zhang, B., & Li, Y. (2012a). Engineering Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: Reflections and perspectives. *Biotechnology Journal*, *7*(1), 34–46. https://doi.org/10.1002/biot.201100053

Cai, Z., Zhang, B., & Li, Y. (2012b). Engineering Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: Reflections and perspectives. *Biotechnology Journal*, *7*(1), 34–46. https://doi.org/10.1002/biot.201100053

Cha, J., Cho, Y., Whitaker, R. D., Carrell, H. L., Glusker, J. P., Karplus, P. A., & Batt, C. A. (1994). Perturbing the Metal Site in D-Xylose Isomerase: Effect of Mutations of His-220 on Enzyme Stability. *Journal of Biological Chemistry*, *269*(4), 2687–2694.

Crabtree, H. G. (1929). Observations on the carbohydrate metabolism of tumours. *The Biochemical Journal*, *23*(3), 536–45. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1254097&tool=pmcentrez&rendertype=abstract

Dauter, Z., Dauter, M., Hemker, J., Witzel, H., & Wilson, K. S. (1989). Crystallisation and preliminary analysis of glucose isomerase from Streptomyces albus. *FEBS Letters*, *247*(1), 1–8. https://doi.org/10.1016/0014-5793(89)81227-X

de Figueiredo Vilela, L., de Mello, V. M., Reis, V. C. B., Bon, E. P. D. S., Gonçalves Torres, F. A., Neves, B. C., & Eleutherio, E. C. A. (2013). Functional expression of Burkholderia cenocepacia xylose isomerase in yeast increases ethanol production from a glucose-xylose blend. *Bioresource Technology*, *128*, 792–796. https://doi.org/10.1016/j.biortech.2012.10.014

Della-Bianca, B. E., de Hulster, E., Pronk, J. T., van Maris, A. J. A., & Gombert, A. K. (2014). Physiology of the fuel ethanol strain Saccharomyces cerevisiae PE-2 at low pH indicates a context-dependent performance relevant for industrial applications. *FEMS Yeast Research*, *14*(8), 1196–1205. https://doi.org/10.1111/1567-1364.12217

Della-Bianca, B. E., & Gombert, a. K. (2013). Stress tolerance and growth physiology of yeast strains from the Brazilian fuel ethanol industry. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *104*(6), 1083–1095. https://doi.org/10.1007/s10482-013-0030-2

Demeke, M. M., Dietz, H., Li, Y., Foulquié-Moreno, M. R., Mutturi, S., Deprez, S., … Thevelein, J. M. (2013). Development of a D-xylose fermenting and inhibitor tolerant industrial Saccharomyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnology for Biofuels*, *6*(1), 89. https://doi.org/10.1186/1754-6834-6-89

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Fenn, T. D., Ringe, D., & Petsko, G. A. (2004). Xylose isomerase in substrate and inhibitor michaelis states: Atomic resolution studies of a metal-mediated hydride shift. *Biochemistry*, *43*(21), 6464–6474. https://doi.org/10.1021/bi049812o

Ha, S. J., Kim, S. R., Choi, J. H., Park, M. S., & Jin, Y. S. (2011). Xylitol does not inhibit xylose fermentation by engineered Saccharomyces cerevisiae expressing xylA as severely as it inhibits xylose isomerase reaction in vitro. *Applied Microbiology and Biotechnology*, *92*(1), 77–84. https://doi.org/10.1007/s00253-011-3345-9

Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I., & Gorwa-Grauslund, M. F. (2007). Towards industrial pentose-fermenting yeast strains. *Applied Microbiology and Biotechnology*, *74*(5), 937–953. https://doi.org/10.1007/s00253-006-0827-2

Henrick, K., Collyer, C. A., & Blow, D. M. (1989). Structures of d-xylose isomerase from Arthrobacter strain B3728 containing the inhibitors xylitol and d-sorbitol at 2.5 ?? and 2.3 ?? resolution, respectively. *Journal of Molecular Biology*, *208*(1), 129–157. https://doi.org/10.1016/0022-2836(89)90092-2

Jacobsen, S. E., & Wyman, C. E. (2000). Cellulose and Hemicellulose Hydrolysis Models for Application to Current and Novel Pretreatment Processes. In *Twenty-First Symposium on Biotechnology for Fuels and Chemicals* (pp. 81–96). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-4612-1392-5_6

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, *292*, 195–202. https://doi.org/10.1006/jmbi.1999.3091

Kim, J., Roy, A., & Ii, D. J. (2013). The glucose signaling network in yeast. *Biochimica et Biophysica Acta*, *1830*(11), 5204–5210. https://doi.org/10.1016/j.bbagen.2013.07.025

Klimacek, M., Krahulec, S., Sauer, U., & Nidetzky, B. (2010). Limitations in xylose-fermenting saccharomyces cerevisiae, made evident through comprehensive metabolite profiling and thermodynamic analysis. *Applied and Environmental Microbiology*, *76*(22), 7566–7574. https://doi.org/10.1128/AEM.01787-10

Kristo, P., Saarelainen, R., Fagerström, R., Aho, S., & Korhola, M. (1996). Protein purification, and cloning and characterization of the cDNA and gene for xylose isomerase of barley. *European Journal of Biochemistry / FEBS*, *237*(1), 240–6. https://doi.org/10.1111/j.1432-1033.1996.0240n.x

Kruckeberg, A. L. (1996). The hexose transporter family of Saccharomyces cerevisiae. *Archives of Microbiology*, *166*(5), 283–292. https://doi.org/10.1007/s002030050385

Kuyper, M., Harhangi, H. R., Stave, A. K., Winkler, A. A., Jetten, M. S. M., de Laat, W. T. A. M., … Pronk, J. T. (2003). High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by Saccharomyces cerevisiae? *FEMS Yeast Research*, *4*(1), 69–78. https://doi.org/10.1016/s1567-1356(03)00141-7

Kuyper, M., Winkler, A. A., Van Dijken, J. P., & Pronk, J. T. (2004). Minimal metabolic engineering of Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: A proof of principle. *FEMS Yeast Research*, *4*(6), 655–664. https://doi.org/10.1016/j.femsyr.2004.01.003

Laskowski, R. a., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(November), 283–291.

https://doi.org/10.1107/S0021889892009944

Lin, Y., & Tanaka, S. (2006). Ethanol fermentation from biomass resources: current state and prospects. *Applied Microbiology and Biotechnology*, *69*(6), 627–642. https://doi.org/10.1007/s00253-005-0229-x

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, *356*(6364), 83–85. https://doi.org/10.1038/356083a0

Madhavan, A., Tamalampudi, S., Srivastava, A., Fukuda, H., Bisaria, V. S., & Kondo, A. (2009). Alcoholic fermentation of xylose and mixed sugars using recombinant Saccharomyces cerevisiae engineered for xylose utilization. *Applied Microbiology and Biotechnology*, *82*(6), 1037–1047. https://doi.org/10.1007/s00253-008-1818-2

Marshall, R. O., & Kooi, E. R. (1957). Enzymatic Conversion of D-Glucose to D-Fructose. *Science*, *125*(3249), 648–649. https://doi.org/10.1126/science.125.3249.648

Martí-Renom, M. A., Stuart, A. C., Fiser, A., Sánchez, R., Melo, F., & Šali, A. (2000). Comparative Protein Structure Modeling of Genes and Genomes. *Annual Review of Biophysics and Biomolecular Structure*, *29*(1), 291–325. https://doi.org/10.1146/annurev.biophys.29.1.291

Matsushika, A., Inoue, H., Kodaki, T., & Sawayama, S. (2009). Ethanol production from xylose in engineered Saccharomyces cerevisiae strains: Current state and perspectives. *Applied Microbiology and Biotechnology*, *84*(1), 37–53. https://doi.org/10.1007/s00253-009-2101-x

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M., & Ladisch, M. (2005). Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*. https://doi.org/10.1016/j.biortech.2004.06.025

Moysés, D., Reis, V., Almeida, J., Moraes, L., & Torres, F. (2016). Xylose Fermentation by Saccharomyces cerevisiae: Challenges and Prospects. *International Journal of Molecular Sciences*, *17*(3), 207. https://doi.org/10.3390/ijms17030207

Peng, B., Huang, S., Liu, T., & Geng, A. (2015). Bacterial xylose isomerases from the mammal gut Bacteroidetes cluster function in Saccharomyces cerevisiae for effective xylose fermentation. *Microbial Cell Factories*, *14*(1), 70. https://doi.org/10.1186/s12934-015-0253-1

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., … Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, *150*(1), 76–85. https://doi.org/10.1016/0003-2697(85)90442-7

Stambuk, B. U., Eleutherio, E. C. A., Florez-Pardo, L. M., Souto-Maior, A. M., & Bon, E. P. S. (2008). Brazilian potential for biomass ethanol: Challenge of using hexose and pentose cofermenting yeast strains. *Journal of Scientific and Industrial Research*, *67*(11), 918–926.

Toteva, M. M., Silvaggi, N. R., Allen, K. N., & Richard, J. P. (2011). Binding Energy and Catalysis by D-Xylose Isomerase: Kinetic, Product and X-Ray Crystallographic Analysis of Enzyme- Catalyzed Isomerization of (R)-Glyceraldehyde. *Biochemistry2*, *50*(46), 10170–10181. https://doi.org/10.1021/bi201378c.Binding

Vásquez, M. P., da Silva, J. N. C., de Souza, M. B., & Pereira, N. (2007). Enzymatic Hydrolysis Optimization to Ethanol Production by Simultaneous Saccharification and Fermentation. In *Applied Biochemistry and Biotecnology* (pp. 141–153). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-60327-181-3_13

Vilela, L. D. F., de Araujo, V. P. G., Paredes, R. D. S., Bon, E. P. D. S., Torres, F. A. G., Neves, B. C., & Eleutherio, E. C. A. (2015). Enhanced xylose fermentation and ethanol production by engineered Saccharomyces cerevisiae strain. *AMB Express*, *5*(1). https://doi.org/10.1186/s13568-015-0102-y

Ye Jeong, E., Sopher, C., Seon Kim, I., & Lee, H. (2001). Mutational study of the role of tyrosine-49 in theSaccharomyces cerevisiae xylose reductase. *Yeast*, *18*(11), 1081–1089. https://doi.org/10.1002/yea.758

Zaldivar, J., Martinez, A., & Ingram, L. O. (1999). Effect of selected aldehydes on the growth and fermentation of ethanologenic Escherichia coli. *Biotechnology and Bioengineering*, *65*(1), 24–33. https://doi.org/10.1002/(SICI)1097-0290(19991005)65:1<24::AID-BIT4>3.0.CO;2-2

Zaldivar, J., Martinez, A., & Ingram, L. O. (2000). Effect of alcohol compounds found in hemicellulose hydrolysate on the growth and fermentation of ethanologenic Escherichia coli. *Biotechnology and Bioengineering*, *68*(5), 524–530. https://doi.org/10.1002/(SICI)1097-0290(20000605)68:5<524::AID-BIT6>3.0.CO;2-T

8. ANEXO I

```
CLUSTAL O(1.2.4) multiple sequence alignment


bcenocepacia_xyla          --MSYFEHIPAIRYEGPQSDNPLAYHHYDPDKRVLGKTLAEHLRIAVCYWHTFVWPGHDI      58
C_phytofermentans_xyla     -MKNYFPNVPEVKYEGPNSTNPFAFKYYDANKVVAGKTMKEHCRFALSWWHTLCAGGADP      59
piromyces_xyla             MAKEYFPQIQKIKFEGKDSKNPLAFHYYDAEKEVMGKKMKDWLRFAMAWWHTLCAEGADQ      60
                              .** ::  :::** :* **:*:::** :* * **.: :  *:*:.:***:   * *

bcenocepacia_xyla          FGQAAFRRPWQQPGDALERARMKADAAFEFFTKLGTPFYTFHDTDVAPEGDSLREYAANF      118
C_phytofermentans_xyla     FGVTTMDRTYGNITDPMELAKAKVDAGFELMTKLGIEFFCFHDADIAPEGDTFEESKKNL      119
piromyces_xyla             FGGGTKSFPWNEGTDAIEIAKQKVDAGFEIMQKLGIPYYCFHDVDLVSEGNSIEEYESNL      120
                           **  :    : :  * :* *: *.**.**:: ***  :: ***.*:. **:::.*   *:

bcenocepacia_xyla          ARMVDYLGERQQASGVRLLWGTANLFSHPRFAAGAATNPNPDVFAWAATQVCHALDATHR      178
C_phytofermentans_xyla     FEIVDYIKEKMDQTGIKLLWGTANNFSHPRFMHGASTSCNADVFAYAAAKIKNALDATIK      179
piromyces_xyla             KAVVAYLKEKQKETGIKLLWSTANVFGHKRYMNGASTNPDFDVVARAIVQIKNAIDAGIE      180
                             :* *: *: . :*::***.*** *.* *:  **:*. : **.* * .:: :*:**  .

bcenocepacia_xyla          LGGENYVLWGGREGYETLLNTDLKRERDQFARFLSMVVEHKHRIGFKGALLIEPKPQEPT      238
C_phytofermentans_xyla     LGGKGYVFWGGREGYETLLNTDLGLELDNMARLMKMAVEYGRANGFDGDFYIEPKPKEPT      239
piromyces_xyla             LGAENYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARSKGFKGTFLIEPKPMEPT      240
                           **.:.**:******* :*****   * :::* ::.*. :: :  **.* : ***** ***

bcenocepacia_xyla          KHQYDYDVATVHGFLVQYGLQNEIRVNIEANHATLAGHSFHHEIANAFALGVFGSVDANR      298
C_phytofermentans_xyla     KHQYDFDTATVLAFLRKYGLEKDFKMNIEANHATLAGHTFEHELAMARVNGAFGSVDANQ      299
piromyces_xyla             KHQYDVDTETAIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDANR      300
                           ***** *. *. .**  :.*::::::***.********:*.**:* *   * :**:***:

bcenocepacia_xyla          GDPQNGWDTDQFPNSVEELTLAFYEILRHGGFTTGGMNFDAKVRRQSIDPEDLFYGHVGA      358
C_phytofermentans_xyla     GDPNLGWDTDQFPTDVHSATLAMLEVLKAGGFTNGGLNFDAKVRRGSFEFDDIAYGYIAG      359
piromyces_xyla             GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGFVTGGTNFDAKTRRNSTDLEDIIIAHVSG      360
                           ** : ******** .  . . *  *::: ***..** *****.** * : :*:  .::..

bcenocepacia_xyla          IDVLALALERAAVLVENDRLDALRRQRYAQWDDAFGRKILAGGYTLESLAADALARGVDP      418
C_phytofermentans_xyla     MDTFALGLIKAAEIIDDGRIAKFVDDRYASYKTGIGKAIVDGTTSLEELEQYVLT-HSEP      418
piromyces_xyla             MDAMARALENAAKLLQESPYTKMKKERYASFDSGIGKDFEDGKLTLEQVYEYGKK-NGEP      419
                           :*.:* .* .** ::::.    :  :***.:. .:*: :  *  :**.:         :*

bcenocepacia_xyla          QHASGAQERLENIVNQAIYGLR      440
C_phytofermentans_xyla     VMQSGRQEVLETIVNNILFR--      438
piromyces_xyla             KQTSGKQELYEAIVAMYQ----      437
                              ** **  * **
```

Figura S 1 Alinhamento múltiplo entre as sequências alvos da modelagem por homologia.

Tabela S 1: Matriz de identidade entre as sequências de XI.

%	*B. cenocepacia*	*C. phytofermentans*	*Piromyces*
B. cenocepacia	100,00	51,95	50,34
C. phytofermentans	51,95	100,00	53,67
Piromyces	50,34	53,67	100,00

```
CLUSTAL O(1.2.4) multiple sequence alignment


bcenocepacia_xyla                MSYFEHIPAIRYEGPQSDNPLAYHHYDPDKRVLGKTLAEHLRIAVCYWHTFVWPGHDIFG   60
1A0D:A|PDBID|CHAIN|SEQUENCE      -PYFDNISTIAYEGPASKNPLAFKFYNPEEKVGDKTMEEHLRFSVAYWHTFTGDGSDPFG   59
1A0C:A|PDBID|CHAIN|SEQUENCE      NKYFENVSKIKYEGPKSNNPYSFKFYNPEEVIDGKTMEEHLRFSIAYWHTFTADGTDQFG   60
                                   **:::  * **** *.** :::.*:*:: : .**: ****:::.*****.  * * **

bcenocepacia_xyla                QAAFRRPWQQPGDALERARMKADAAFEFFTKLGTPFYTFHDTDVAPEGDSLREYAANFAR   120
1A0D:A|PDBID|CHAIN|SEQUENCE      AGNMIRPWNKYS-GMDLAKARVEAAFEFFEKLNIPFFCFHDVDIAPEGETLKETYKNLDI   118
1A0C:A|PDBID|CHAIN|SEQUENCE      KATMQRPWNHYTDPMDIAKARVEAAFEFFDKINAPYFCFHDRDIAPEGDTLRETNKNLDT   120
                                  . : ***::    :: *: :.:****** *:. *:: *** *:****::*:*   *:

bcenocepacia_xyla                MVDYLGERQQASGVRLLWGTANLFSHPRFAAGAATNPNPDVFAWAATQVCHALDATHRLG   180
1A0D:A|PDBID|CHAIN|SEQUENCE      IVDMIEEYMKTSKTKLLWNTANLFTHPRFVHGAATSCNADVFAYAAAKVKKGLEIAKRLG   178
1A0C:A|PDBID|CHAIN|SEQUENCE      IVAMIKDYLKTSKTKVLWGTANLFSNPRFVHGASTSCNADVFAYSAAQVKKALEITKELG   180
                                 :*  : :  ::* .::**.*****::***. **:*. * ****::*::* :.*: ::.**

bcenocepacia_xyla                GENYVLWGGREGYETLLNTDLKRERDQFARFLSMVVEHKHRIGFKGALLIEPKPQEPTKH   240
1A0D:A|PDBID|CHAIN|SEQUENCE      AENYVFWGGREGYETLLNTDMKLELDNLARFLHMAVDYAKEIGFDGQFLIEPKPKEPTKH   238
1A0C:A|PDBID|CHAIN|SEQUENCE      GENYVFWGGREGYETLLNTDMEFELDNFARFLHMAVDYAKEIGFEGQFLIEPKPKEPTKH   240
                                 .****:**************:: * *::**** *.*:: :.***.* :******:*****

bcenocepacia_xyla                QYDYDVATVHGFLVQYGLQNEIRVNIEANHATLAGHSFHHEIANAFALGVFGSVDANRGD   300
1A0D:A|PDBID|CHAIN|SEQUENCE      QYDFDVATALAFLQTYGLKDYFKFNIEANHATLAGHTFEHELRVARIHGMLGSVDANQGD   298
1A0C:A|PDBID|CHAIN|SEQUENCE      QYDFDVANVLAFLRKYDLDKYFKVNIEANHATLAFHDFQHELRYARINGVLGSIDANTGD   300
                                 ***:***.. .**  *.*.. ::.********** * *.**:  *   *::**:*** **

bcenocepacia_xyla                PQNGWDTDQFPNSVEELTLAFYEILRHGGFTTGGMNFDAKVRRQSIDPEDLFYGHVGAID   360
1A0D:A|PDBID|CHAIN|SEQUENCE      MLLGWDTDEFPTDLYSTTLAMYEILKNGGLGRGGLNFDAKVRRGSFEPEDLFYAHIAGMD   358
1A0C:A|PDBID|CHAIN|SEQUENCE      MLLGWDTDQFPTDIRMTTLAMYEVIKMGGFDKGGLNFDAKVRRASFEPEDLFLGHIAGMD   360
                                    *****:**..:   ***:**::: **:  **:******** *::***** .*:..:*

bcenocepacia_xyla                VLALALERAAVLVENDRLDALRRQRYAQWDDAFGRKILAGGYTLESLAADALARGVDPQH   420
1A0D:A|PDBID|CHAIN|SEQUENCE      SFAVGLKVAHRLIEDRVFDEFIEERYKSYTEGIGREIVEGTADFHKLEAHALQLGE-IQN   417
1A0C:A|PDBID|CHAIN|SEQUENCE      AFAKGFKVAYKLVKDRVFDKFIEERYASYKDGIGADIVSGKADFRSLEKYALERSQ-IVN   419
                                  :* .:: *  *:::  :* : .:** .: :.:* .*: *   :..*   **  .    :

bcenocepacia_xyla                ASGAQERLENIVNQAIYGLR---        440
1A0D:A|PDBID|CHAIN|SEQUENCE      QSGRQERLKTLLNQYLLEVCAAR        440
1A0C:A|PDBID|CHAIN|SEQUENCE      KSGRQELLESILNQYLFAE----        438
                                  ** ** *:.::** :
```

Figura S 2: Alinhamento entre sequência alvo (*B. cenocepacia)* e moldes.

```
CLUSTAL O(1.2.4) multiple sequence alignment


C_phytofermentans_xyla          MKNYFPNVPEVKYEGPNSTNPFAFKYYDANKVVAGKTMKEHCRFALSWWHTLCAGGADPF   60
1A0D:A|PDBID|CHAIN|SEQUENCE     --PYFDNISTIAYEGPASKNPLAFKFYNPEEKVGDKTMEEHLRFSVAYWHTFTGDGSDPF   58
1A0C:A|PDBID|CHAIN|SEQUENCE     -NKYFENVSKIKYEGPKSNNPYSFKFYNPEEVIDGKTMEEHLRFSIAYWHTFTADGTDQF   59
                                   ** *:  : **** *.** :**:*: :: : .***:** **::::***: ..*:* *

C_phytofermentans_xyla          GVTTMDRTYGNITDPMELAKAKVDAGFELMTKLGIEFFCFHDADIAPEGDTFEESKKNLF   120
1A0D:A|PDBID|CHAIN|SEQUENCE     GAGNMIRPWNKYS-GMDLAKARVEAAFEFFEKLNIPFFCFHDVDIAPEGETLKETYKNLD   117
1A0C:A|PDBID|CHAIN|SEQUENCE     GKATMQRPWNHYTDPMDIAKARVEAAFEFFDKINAPYFCFHDRDIAPEGDTLRETNKNLD   119
                                *  .* * :.: :  *::***:*:*.**:: *:.  :***** ******:*:.*: ***

C_phytofermentans_xyla          EIVDYIKEKMDQTGIKLLWGTANNFSHPRFMHGASTSCNADVFAYAAAKIKNALDATIKL   180
1A0D:A|PDBID|CHAIN|SEQUENCE     IIVDMIEEYMKTSKTKLLWNTANLFTHPRFVHGAATSCNADVFAYAAAKVKKGLEIAKRL   177
1A0C:A|PDBID|CHAIN|SEQUENCE     TIVAMIKDYLKTSKTKVLWGTANLFSNPRFVHGASTSCNADVFAYSAAQVKKALEITKEL   179
                                 **  *:: :. :  *:**.*** *::***:***:**********:**::*:.*: : .*

C_phytofermentans_xyla          GGKGYVFWGGREGYETLLNTDLGLELDNMARLMKMAVEYGRANGFDGDFYIEPKPKEPTK   240
1A0D:A|PDBID|CHAIN|SEQUENCE     GAENYVFWGGREGYETLLNTDMKLELDNLARFLHMAVDYAKEIGFDGQFLIEPKPKEPTK   237
1A0C:A|PDBID|CHAIN|SEQUENCE     GGENYVFWGGREGYETLLNTDMEFELDNFARFLHMAVDYAKEIGFEGQFLIEPKPKEPTK   239
                                *.:.****************: :****:**:::***:*.:  **:*:* **********

C_phytofermentans_xyla          HQYDFDTATVLAFLRKYGLEKDFKMNIEANHATLAGHTFEHELAMARVNGAFGSVDANQG   300
1A0D:A|PDBID|CHAIN|SEQUENCE     HQYDFDVATALAFLQTYGLKDYFKFNIEANHATLAGHTFEHELRVARIHGMLGSVDANQG   297
1A0C:A|PDBID|CHAIN|SEQUENCE     HQYDFDVANVLAFLRKYDLDKYFKVNIEANHATLAFHDFQHELRYARINGVLGSIDANTG   299
                                ******.*..****:.*.*.. **.********** * *:***  **::* :**:*** *

C_phytofermentans_xyla          DPNLGWDTDQFPTDVHSATLAMLEVLKAGGFTNGGLNFDAKVRRGSFEFDDIAYGYIAGM   360
1A0D:A|PDBID|CHAIN|SEQUENCE     DMLLGWDTDEFPTDLYSTTLAMYEILKNGGLGRGGLNFDAKVRRGSFEPEDLFYAHIAGM   357
1A0C:A|PDBID|CHAIN|SEQUENCE     DMLLGWDTDQFPTDIRMTTLAMYEVIKMGGFDKGGLNFDAKVRRASFEPEDLFLGHIAGM   359
                                *  ******:****:  :**** *::* **: .***********.*** :*:  .:****

C_phytofermentans_xyla          DTFALGLIKAAEIIDDGRIAKFVDDRYASYKTGIGKAIVDGTTSLEELEQYVLTHSEPVM   420
1A0D:A|PDBID|CHAIN|SEQUENCE     DSFAVGLKVAHRLIEDRVFDEFIEERYKSYTEGIGREIVEGTADFHKLEAHALQLGEIQN   417
1A0C:A|PDBID|CHAIN|SEQUENCE     DAFAKGFKVAYKLVKDRVFDKFIEERYASYKDGIGADIVSGKADFRSLEKYALERSQIVN   419
                                *:** *:  * .::.*  : :*:::** **. ***  **.*.:.:..** :.*  .:

C_phytofermentans_xyla          QSGRQEVLETIVNNILFR-----        438
1A0D:A|PDBID|CHAIN|SEQUENCE     QSGRQERLKTLLNQYLLEVCAAR        440
1A0C:A|PDBID|CHAIN|SEQUENCE     KSGRQELLESILNQYLFAE----        438
                                :***** *::::*: *:
```

Figura S 3: Alinhamento entre sequência alvo (*C. phytofermentas)* e moldes.

```
CLUSTAL O(1.2.4) multiple sequence alignment


piromyces_xyla                  MAKEYFPQIQKIKFEGKDSKNPLAFHYYDAEKEVMGKKMKDWLRFAMAWWHTLCAEGADQ   60
1A0D:A|PDBID|CHAIN|SEQUENCE     ---PYFDNISTIAYEGPASKNPLAFKFYNPEEKVGDKTMEEHLRFSVAYWHTFTGDGSDP   57
1A0C:A|PDBID|CHAIN|SEQUENCE     --NKYFENVSKIKYEGPKSNNPYSFKFYNPEEVIDGKTMEEHLRFSIAYWHTFTADGTDQ   58
                                    ** ::..* :**  *:** :*::*: *: : .*.*:: ***::*:***: .:*:*

piromyces_xyla                  FGGGTKSFPWNEGTDAIEIAKQKVDAGFEIMQKLGIPYYCFHDVDLVSEGNSIEEYESNL   120
1A0D:A|PDBID|CHAIN|SEQUENCE     FGAGNMIRPWNKYS-GMDLAKARVEAAFEFFEKLNIPFFCFHDVDIAPEGETLKETYKNL   116
1A0C:A|PDBID|CHAIN|SEQUENCE     FGKATMQRPWNHYTDPMDIAKARVEAAFEFFDKINAPYFCFHDRDIAPEGDTLRETNKNL   118
                                ** ..   ***. :  ::** :*:*.**:::*:. *::**** *:. **:::.*  .**

piromyces_xyla                  KAVVAYLKEKQKETGIKLLWSTANVFGHKRYMNGASTNPDFDVVARAIVQIKNAIDAGIE   180
1A0D:A|PDBID|CHAIN|SEQUENCE     DIIVDMIEEYMKTSKTKLLWNTANLFTHPRFVHGAATSCNADVFAYAAAKVKKGLEIAKR   176
1A0C:A|PDBID|CHAIN|SEQUENCE     DTIVAMIKDYLKTSKTKVLWGTANLFSNPRFVHGASTSCNADVFAYSAAQVKKALEITKE   178
                                . :*  :::  * :  *:**.***:* : *:::**:*. : **.* : .::*:.::   .

piromyces_xyla                  LGAENYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARSKGFKGTFLIEPKPMEPT   240
1A0D:A|PDBID|CHAIN|SEQUENCE     LGAENYVFWGGREGYETLLNTDMKLELDNLARFLHMAVDYAKEIGFDGQFLIEPKPKEPT   236
1A0C:A|PDBID|CHAIN|SEQUENCE     LGGENYVFWGGREGYETLLNTDMEFELDNFARFLHMAVDYAKEIGFEGQFLIEPKPKEPT   238
                                **.************ :***** : * :::* :* ** ***:. **.* ******* ***

piromyces_xyla                  KHQYDVDTETAIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDANR   300
1A0D:A|PDBID|CHAIN|SEQUENCE     KHQYDFDVATALAFLQTYGLKDYFKFNIEANHATLAGHTFEHELRVARIHGMLGSVDANQ   296
1A0C:A|PDBID|CHAIN|SEQUENCE     KHQYDFDVANVLAFLRKYDLDKYFKVNIEANHATLAFHDFQHELRYARINGVLGSIDANT   298
                                *****.*. ..:.**: :.*.. **.***.****** * *:***  *   *:***:***

piromyces_xyla                  GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGFVTGGTNFDAKTRRNSTDLEDIIIAHVSG   360
1A0D:A|PDBID|CHAIN|SEQUENCE     GDMLLGWDTDEFPTDLYSTTLAMYEILKNGGLGRGGLNFDAKVRRGSFEPEDLFYAHIAG   356
1A0C:A|PDBID|CHAIN|SEQUENCE     GDMLLGWDTDQFPTDIRMTTLAMYEVIKMGGFDKGGLNFDAKVRRASFEPEDLFLGHIAG   358
                                **   *****:** *    . *  *::: **:  ** *****.** * : **:: .*::*

piromyces_xyla                  MDAMARALENAAKLLQESPYTKMKKERYASFDSGIGKDFEDGKLTLEQVYEYGKKNGEPK   420
1A0D:A|PDBID|CHAIN|SEQUENCE     MDSFAVGLKVAHRLIEDRVFDEFIEERYKSYTEGIGREIVEGTADFHKLEAHALQLGEIQ   416
1A0C:A|PDBID|CHAIN|SEQUENCE     MDAFAKGFKVAYKLVKDRVFDKFIEERYASYKDGIGADIVSGKADFRSLEKYALERSQIV   418
                                **::* .:: * :*:::  : :: :*** *: .*** :: .*.  :..:  :. : .:

piromyces_xyla                  QTSGKQELYEAIVAMYQ-------       437
1A0D:A|PDBID|CHAIN|SEQUENCE     NQSGRQERLKTLLNQYLLEVCAAR       440
1A0C:A|PDBID|CHAIN|SEQUENCE     NKSGRQELLESILNQYLFAE----       438
                                : **:**  ::::  *
```

Figura S 4: Alinhamento entre sequência alvo (*Piromyces* sp.*)* e moldes.

```
# Illustrates the SALIGN multiple sequence alignment
from modeller import *

log.verbose()
env = environ()
env.io.atom_files_directory = ['.', '../atom_files']

aln = alignment(env, file='malign_in.ali')

aln.salign(overhang=30, gap_penalties_1d=(-450, -50),
           alignment_type='tree', output='ALIGNMENT')

aln.write(file='malign.ali', alignment_format='PIR')
aln.write(file='malign.pap', alignment_format='PAP')
```

Figura S 5: Algoritmo utilizado para o alinhamento múltiplo das sequências antes da modelagem.

```
# Comparative modeling with multiple templates
from modeller import *              # Load standard Modeller classes
from modeller.automodel import *    # Load the automodel class
from modeller.parallel import *         #Load parallel job class

#Parallelize jobs
j = job(host='localhost')
for i in range(12):     #Define number of CPUs
    j.append(local_slave())

log.verbose()    # request verbose output
env = environ()  # create a new MODELLER environment to build this model in

# directories for input atom files
env.io.atom_files_directory = ['.', '../atom_files']

# Read in HETATM records from template PDBs
env.io.hetatm = True

#Read in WATER records from template PDBs
env.io.water = True

a = automodel(env,
              alnfile  = 'malign.ali', # alignment filename
              knowns   = ('1a0c', '1a0d'),     # codes of the templates
              sequence = ('xyla'),             # code of the target
              assess_methods=(assess.DOPE, assess.normalized_dope)) #DOPE and normalized DOPE score

a.starting_model= 1                 # index of the first model
a.ending_model  = 500                 # index of the last model
                                    # (determines how many models to calculate)

# Thorough MD optimization:
a.md_level = refine.slow

# Repeat the whole cycle 2 times and do not stop unless obj.func. > 1E6
a.repeat_optimization = 5
a.max_molpdf = 1e6

a.use_parallel_job(j)
a.make()                            # do the actual comparative modeling
```

Figura S 6: Algoritmo utilizado para a construção dos modelos teóricos no programa MODELLER v9.17.

```
# PSIPRED HFORMAT (PSIPRED V3.3)

Conf: 999999997123489999985224469767315977044432354543036999999999
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHCCCCCCCCCCC
  AA: MSYFEHIPAIRYEGPQSDNPLAYHHYDPDKRVLGKTLAEHLRIAVCYWHTFVWPGHDIFG
             10        20        30        40        50        60

Conf: 877788899999389999975777733366518973431147789999998999994999
Pred: CCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCHHHHHHHHHH
  AA: QAAFRRPWQQPGDALERARMKADAAFEFFTKLGTPFYTFHDTDVAPEGDSLREYAANFAR
             70        80        90       100       110       120

Conf: 999999987641893000033588661123798999984588985788789998876618
Pred: HHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHC
  AA: MVDYLGERQQASGVRLLWGTANLFSHPRFAAGAATNPNPDVFAWAATQVCHALDATHRLG
            130       140       150       160       170       180

Conf: 856122165322023202543898999999999999977740997003205699999888
Pred: CCCCEECCCCCCCHHHHCCCHHHHHHHHHHHHHHHHHHHCCCCCCEEECCCCCCCCCCC
  AA: GENYVLWGGREGYETLLNTDLKRERDQFARFLSMVVEHKHRIGFKGALLIEPKPQEPTKH
            190       200       210       220       230       240

Conf: 641568899999986199953211134451001488425899999980983221379999
Pred: CCHHHHHHHHHHHHHCCCCCCCCCCCCCCCCHHCCCCCHHHHHHHHHHCCCCCCCCCCCC
  AA: QYDYDVATVHGFLVQYGLQNEIRVNIEANHATLAGHSFHHEIANAFALGVFGSVDANRGD
            250       260       270       280       290       300

Conf: 989988788999949999999999980998999667534576789991146889875589
Pred: CCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHH
  AA: PQNGWDTDQFPNSVEELTLAFYEILRHGGFTTGGMNFDAKVRRQSIDPEDLFYGHVGAID
            310       320       330       340       350       360

Conf: 999999999994052377999986225861112356507789999999999972999989
Pred: HHHHHHHHHHHHHHCCHHHHHHHHHHCCCCHHHHHHHHCCCCCHHHHHHHHHHCCCCCC
  AA: VLALALERAAVLVENDRLDALRRQRYAQWDDAFGRKILAGGYTLESLAADALARGVDPQH
            370       380       390       400       410       420
```

Figura S 7: Predição de regiões de estrutura secundária para XI de *B. cenocepacia*.

```
# PSIPRED HFORMAT (PSIPRED V3.3)

Conf: 988799999712338999998522456975522687704442234464403798999988
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHCCCCCCCCCCC
  AA: MKNYFPNVPEVKYEGPNSTNPFAFKYYDANKVVAGKTMKEHCRFALSWWHTLCAGGADPF
              10        20        30        40        50        60

Conf: 988766568999998999997466665656751895243114677999999899998299
Pred: CCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCHHHHHHHHH
  AA: GVTTMDRTYGNITDPMELAKAKVDAGFELMTKLGIEFFCFHDADIAPEGDTFEESKKNLF
              70        80        90       100       110       120

Conf: 999999998665189311013358867335689888997458998689978999887671
Pred: HHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHH
  AA: EIVDYIKEKMDQTGIKLLWGTANNFSHPRFMHGASTSCNADVFAYAAAKIKNALDATIKL
             130       140       150       160       170       180

Conf: 886530216620023420255248799999999999997453199612000469999988
Pred: CCCCCEECCCCCHHHHHHCCCCHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCC
  AA: GGKGYVFWGGREGYETLLNTDLGLELDNMARLMKMAVEYGRANGFDGDFYIEPKPKEPTK
             190       200       210       220       230       240

Conf: 851255899999999619997431001321122027941799999998197221137999
Pred: CCCHHHHHHHHHHHHHCCCCCCCCHHCCCCHHHHCCCCHHHHHHHHHHCCCCCCCCCCC
  AA: HQYDFDTATVLAFLRKYGLEKDFKMNIEANHATLAGHTFEHELAMARVNGAFGSVDANQG
             250       260       270       280       290       300

Conf: 999998889789894999999999998099899965765467789988146798977759
Pred: CCCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHH
  AA: DPNLGWDTDQFPTDVHSATLAMLEVLKAGGFTNGGLNFDAKVRRGSFEFDDIAYGYIAGM
             310       320       330       340       350       360

Conf: 999998999999851203347799741269865321333187899999999983799989
Pred: HHHHHHHHHHHHHHHCCHHHHHHHHHHCCCCCCCCHHHHCCCCCHHHHHHHHHHCCCCC
  AA: DTFALGLIKAAEIIDDGRIAKFVDDRYASYKTGIGKAIVDGTTSLEELEQYVLTHSEPVM
             370       380       390       400       410       420

Conf: 984379998896632129
Pred: CCCHHHHHHHHHHHHCC
  AA: QSGRQEVLETIVNNILFR
             430
```

Figura S 8: Predição de regiões de estrutura secundária para XI de *C. phytofermentas*.

```
# PSIPRED HFORMAT (PSIPRED V3.3)

Conf: 975458999872233699999752344697562169751555434656520379899999
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHCCCCCCCCC
  AA: MAKEYFPQIQKIKFEGKDSKNPLAFHYYDAEKEVMGKKMKDWLRFAMAWWHTLCAEGADQ
              10        20        30        40        50        60

Conf: 899988656899993899999846665575876289613212566789899989999959
Pred: CCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCHHHHHHHH
  AA: FGGGTKSFPWNEGTDAIEIAKQKVDAGFEIMQKLGIPYYCFHDVDLVSEGNSIEEYESNL
              70        80        90       100       110       120

Conf: 999999999988628930101345888734548989999835888667898899987556
Pred: HHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHH
  AA: KAVVAYLKEKQKETGIKLLWSTANVFGHKRYMNGASTNPDFDVVARAIVQIKNAIDAGIE
             130       140       150       160       170       180

Conf: 189661221673101342124517889999999999999766419973310047999998
Pred: HCCCCCEECCCCCCHHHHHCCCHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCC
  AA: LGAENYVFWGGREGYMSLLNTDQKREKEHMATMLTMARDYARSKGFKGTFLIEPKPMEPT
             190       200       210       220       230       240

Conf: 985303488999999865999643200133302202794178999999819811302799
Pred: CCCCHHHHHHHHHHHHCCCCCCCCCCCCCCCHHHCCCCCHHHHHHHHHHCCCCCCCCCCC
  AA: KHQYDVDTETAIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAVDAGMLGSIDANR
             250       260       270       280       290       300

Conf: 998899888979989499999999999729988886675456778899814689998774
Pred: CCCCCCCCCCCCCCCHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHH
  AA: GDYQNGWDTDQFPIDQYELVQAWMEIIRGGGFVTGGTNFDAKTRRNSTDLEDIIIAHVSG
             310       320       330       340       350       360

Conf: 999999999999884048202789985236986434001128789999999998659999
Pred: HHHHHHHHHHHHHHHCCCCHHHHHHHHCCCCCCCCCCHHCCCCCHHHHHHHHHHCCCCC
  AA: MDAMARALENAAKLLQESPYTKMKKERYASFDSGIGKDFEDGKLTLEQVYEYGKKNGEPK
             370       380       390       400       410       420

Conf: 99824899998641089
Pred: CCCCHHHHHHHHHCCC
  AA: QTSGKQELYEAIVAMYQ
             430
```

Figura S 9: Predição de regiões de estrutura secundária para XI de *Piromyces* sp.

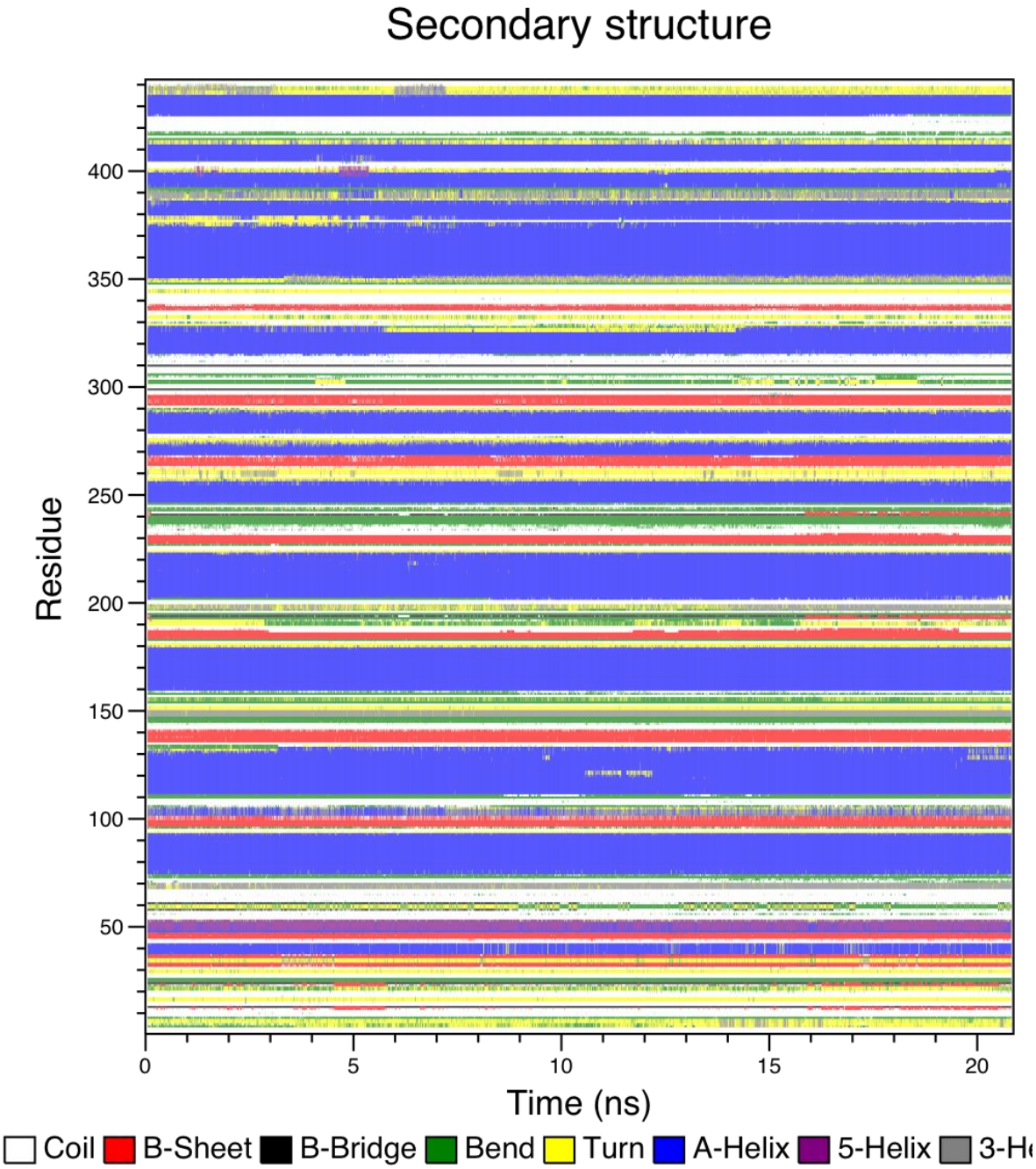

Figura S 10: Variação do conteúdo de estruturas secundárias ao longo da simulação.

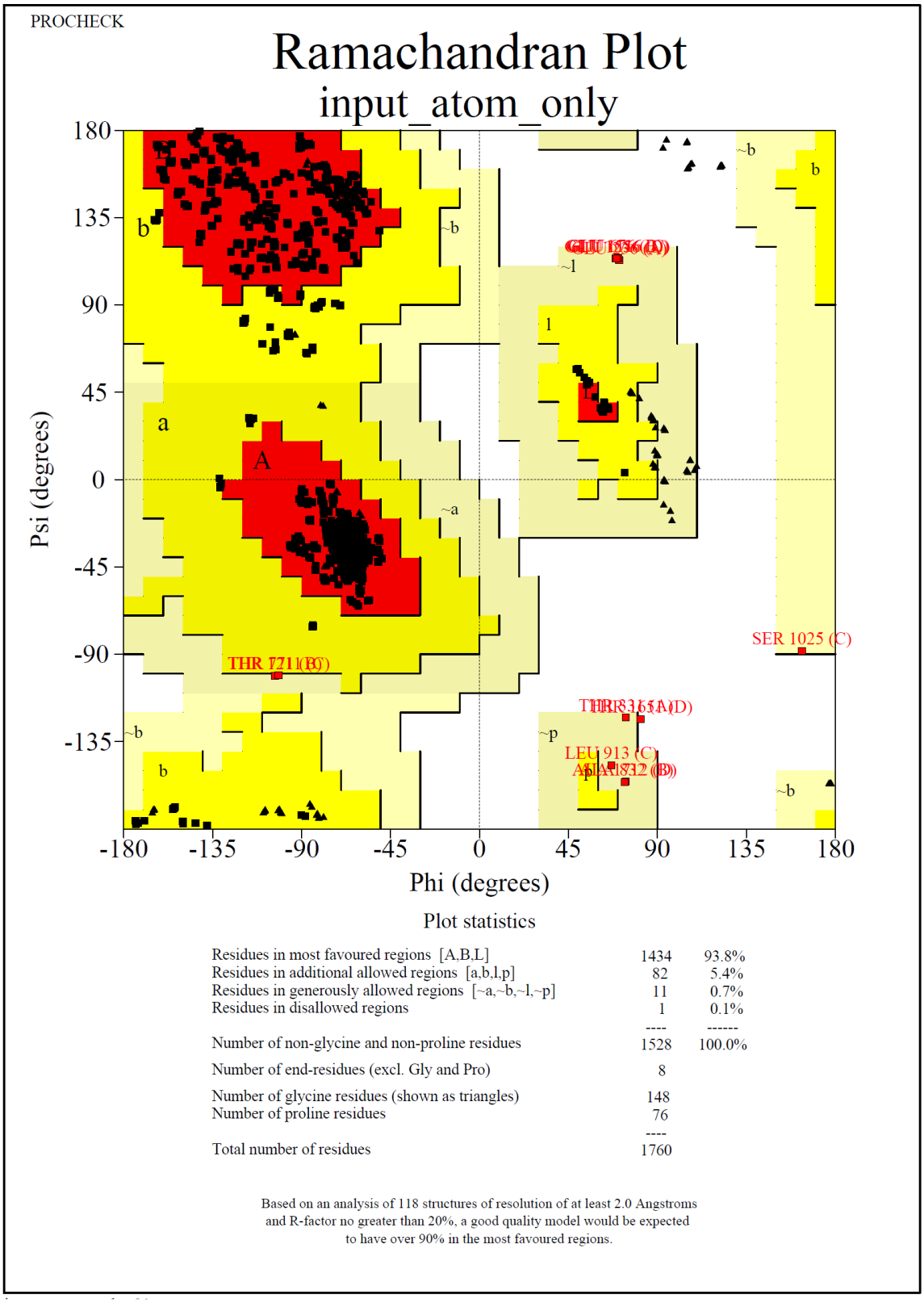

Figura S11: Plot de Ramachandran para o modelo de *B. cenocepacia*.

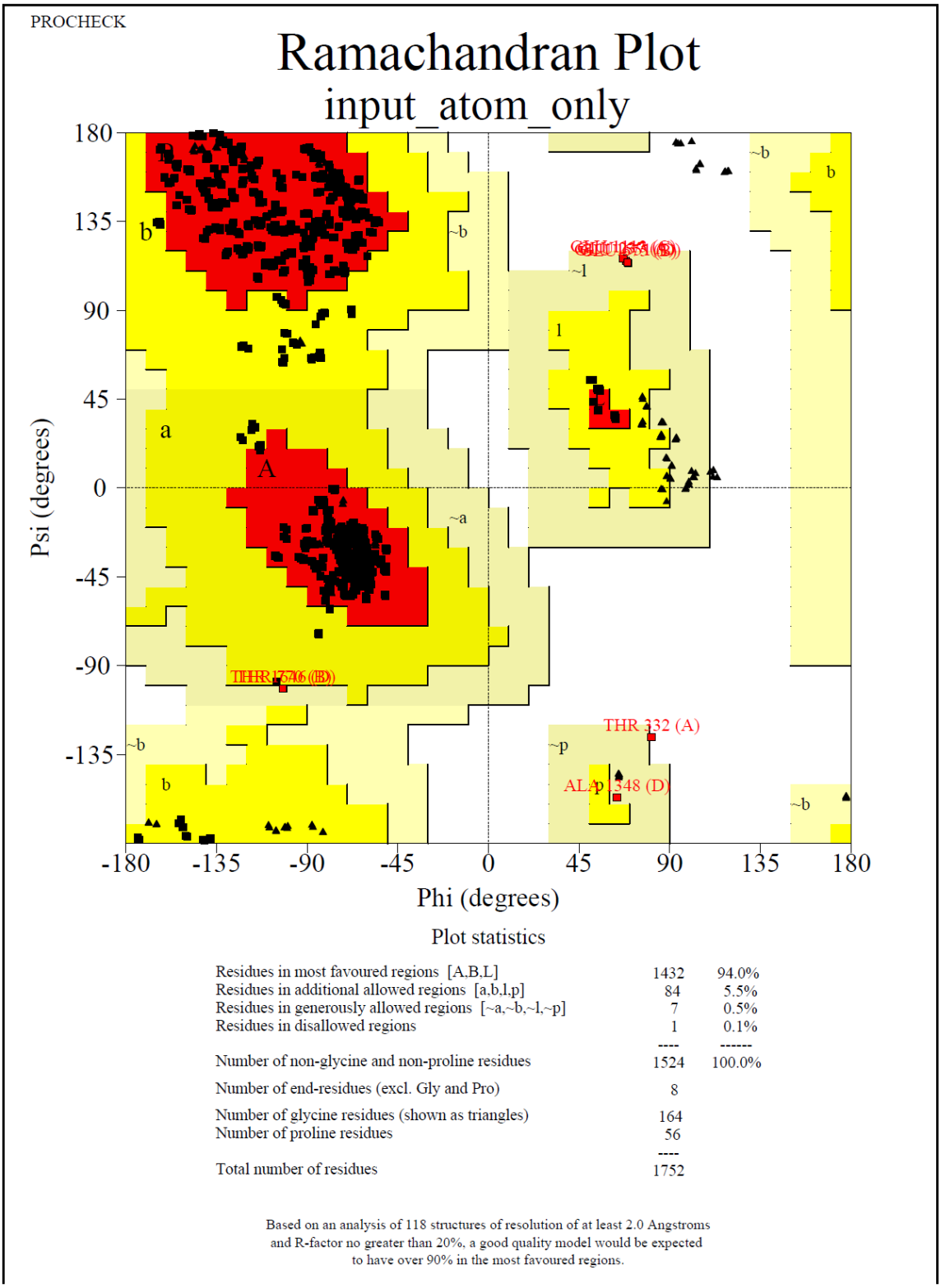

Figura S12: Plot de Ramachandran para o modelo de *C. phytofermentans.*

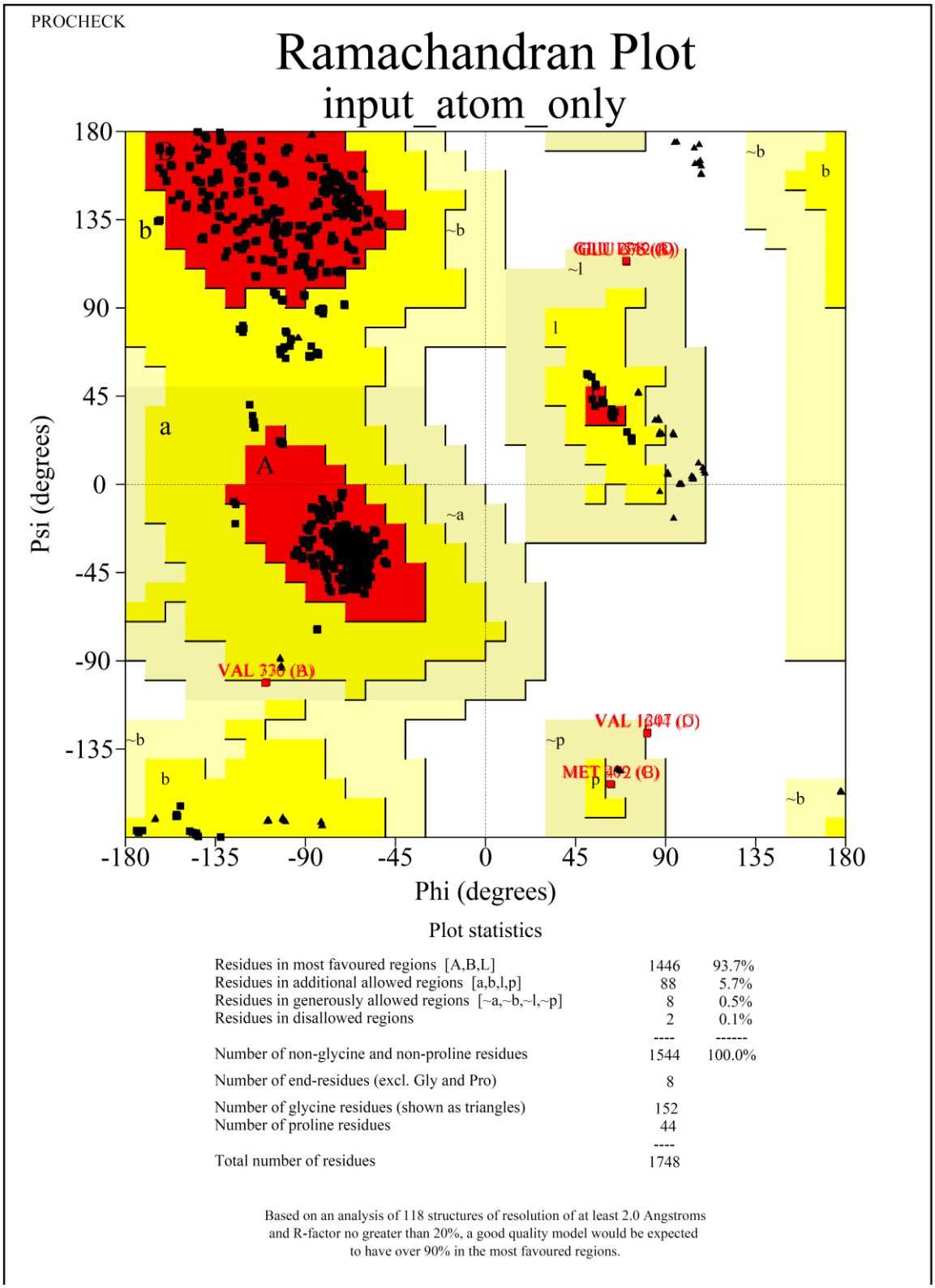

Figura S13: Plot de Ramachandran para o modelo de *Piromyces sp.*