

**Universidade Estadual Paulista “Júlio de Mesquita Filho”  
Instituto de Biociências de Botucatu**

**Hamine Cristina de Oliveira**

**ESTUDO COMPARATIVO ENTRE DIFERENTES TÉCNICAS BIOFÍSICAS  
VISANDO APROFUNDAR O ENTENDIMENTO DO MECANISMO DE  
IMPORTAÇÃO NUCLEAR**

Botucatu  
2024

Programa de pós-graduação em biotecnologia

**Hamine Cristina de Oliveira**

**ESTUDO COMPARATIVO ENTRE DIFERENTES TÉCNICAS BIOFÍSICAS  
VISANDO APROFUNDAR O ENTENDIMENTO DO MECANISMO DE  
IMPORTAÇÃO NUCLEAR**

Tese de doutorado apresentada ao Instituto  
de Biociências de Botucatu – UNESP.

Marcos Roberto de Mattos Fontes  
**Orientador**

Thiago Revers Dreyer  
**Coorientador**

Projeto financiado pela FAPESP  
Processo: 2019/05239-8

D418e

de Oliveira, Hamine Cristina

Estudos comparativos entre diferentes técnicas biofísicas visando  
aprofundar o entendimento do mecanismo de importação nuclear /  
Hamine Cristina de Oliveira. -- Botucatu, 2024

58 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),  
Instituto de Biociências, Botucatu

Orientador: Marcos Roberto de Mattos Fontes

Coorientador: Thiago Revers Dreyer

1. Biofísica molecular. I. Título.

**ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE HAMINE CRISTINA DE OLIVEIRA, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA, DO INSTITUTO DE BIOCIÊNCIAS - CÂMPUS DE BOTUCATU.**

Aos 29 dias do mês de julho do ano de 2024, às 14:00 horas, no(a) Departamento de biofísica e farmacologia, realizou-se a defesa de TESE DE DOUTORADO de HAMINE CRISTINA DE OLIVEIRA, intitulada **ESTUDO COMPARATIVO ENTRE DIFERENTES TÉCNICAS BIOFÍSICAS VISANDO APROFUNDAR O ENTENDIMENTO DO MECANISMO DE IMPORTAÇÃO NUCLEAR**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. MARCOS ROBERTO DE MATTOS FONTES (Orientador(a) - Participação Presencial) do(a) Departamento de Biofísica e Farmacologia / Instituto de Biociências de Botucatu Unesp, Prof. Dr. JOÃO RENATO CARVALHO MUNIZ (Participação Virtual) do(a) Departamento de Física e Ciência Interdisciplinar / Instituto de Física de São Carlos - USP, Prof. Dr. CARLOS ALEXANDRE HENRIQUE FERNANDES (Participação Virtual) do(a) Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC) - Sorbonne Université, Prof. Dr. JOEL MESA HORMAZA (Participação Presencial) do(a) Departamento de Biofísica e Farmacologia / Instituto de Biociências de Botucatu - Unesp, Prof. Dr. ANGELO JOSÉ MAGRO (Participação Presencial) do(a) Departamento de Bioprocessos e Biotecnologia / Faculdade de Ciências Agrônômicas de Botucatu - UNESP. Após a exposição pela doutoranda e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, a discente recebeu o conceito final: APROVADO. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Prof. Dr. MARCOS ROBERTO DE MATTOS FONTES

Dedico este trabalho a todas as  
mulheres que sonham e resistem.

## AGRADECIMENTOS

Ao professor Marcos Fontes, meu orientador, pelo apoio em todos os momentos, ensinamentos, disponibilidade, compreensão para além das obrigações acadêmicas, confiança e estímulo nesses 9 anos de jornada acadêmica. Ao Thiago, Ivan, Tainá e demais membros do laboratório de biologia molecular e estrutural que sempre me auxiliaram e tornaram a jornada ainda mais gratificante.

Ao Instituto de Biociências de Botucatu, seus docentes e funcionários, que desde 2015 quando ingressei na graduação, me acompanham e contribuem para tornar esse trabalho uma realidade.

À FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pela bolsa concedida, possibilitando a realização da pesquisa (Processo 2019/05239-8).

Agradeço a Ogum e todos Àqueles que emocionalmente me mantiveram em pé e no caminho.

À minha irmã Luany, meu pai Daniel, meu companheiro Luciano, ao Gustavo, aos meus amigos e irmãos pelo apoio, estímulo e companheirismo.

De forma incondicional à minha mãe Gislaine, que apesar de fisicamente ter me deixado no meio dessa jornada, é, e sempre será, minha maior fonte de inspiração e força.

“There have been lots of other women  
who had the talent and ability before me.  
I think this can be seen as an affirmation  
that we're moving ahead.  
And I hope it means that I'm just the first in  
a long line.”

Mae Jemison

## RESUMO

O estudo quantitativo da interação proteínas-ligantes vem se mostrando essencial em diversas áreas. No estudo da importação nuclear, particularmente na chamada via clássica de importação nuclear, estes dados são essenciais, pois a importina- $\alpha$  (Imp $\alpha$ ) se acopla às proteínas a serem transportadas ao núcleo a partir do reconhecimento de sequências de localização nuclear (NLS). Assim, o aprimoramento no uso de métodos biofísicos de análise da interação macromolécula/ligante tem um papel fundamental. Desta maneira, o presente trabalho tem duas vertentes principais: i) comparar a interação da Imp $\alpha$  com NLSs de proteínas cargo por diferentes técnicas, como a calorimetria por titulação isotérmica (ITC), a espectroscopia de fluorescência, a termoforese em microescala (MST) e o *docking* molecular, analisando as principais diferenças entre cada uma delas. Esta comparação é pouco abordada na literatura e se mostrou promissora, uma vez que apesar de utilizar princípios físicos diferentes as técnicas obtiveram valores na mesma ordem de grandeza. (ii) Desenvolvimento de um algoritmo de código aberto para a análise de dados de ITC, suprimindo a dificuldade que os *softwares* comerciais apresentam no cálculo dos parâmetros em sistemas mais complexos. Os dados de ITC entre a Imp $\alpha$  e diferentes NLSs foram utilizados como referência para a comparação entre as constantes de afinidade obtidas nas outras técnicas e como validação para o algoritmo. Os mesmos ligantes foram utilizados nos ensaios de espectroscopia de fluorescência, termoforese em microescala e *docking* molecular. Apesar de bastante distintas, todas as técnicas obtiveram valores da constante de dissociação na mesma ordem de grandeza. Os valores obtidos por *docking* molecular são superestimados. Como a fluorescência e o MST não diferenciam as constantes para cada sítio de ligação, o valor encontrado representa uma afinidade média de ambos os sítios. Nota-se, uma maior semelhança entre os dados de ITC e fluorescência. O algoritmo de ITC teve resultados similares aos obtidos pelo *software* do fabricante, sua implementação trará um avanço para a facilitação da análise de dados complexos.

## ABSTRACT

The quantitative study of protein-ligand interactions has been essential in several areas. In the study of nuclear import, particularly in the classical pathway of nuclear import, these data are essential, as importin- $\alpha$  (Imp $\alpha$ ) binds to proteins to be transported to the nucleus from the recognition of nuclear localization sequences (NLS). Thus, the improvement in the use of biophysical methods of analysis of the macromolecule/ligand interaction has a fundamental role. Thus, the present work has two main aspects: i) to compare the interaction of Imp $\alpha$  with NLSs of cargo proteins by different techniques, such as isothermal titration calorimetry (ITC), fluorescence spectroscopy, microscale thermophoresis (MST) and molecular docking, analyzing the main differences between each one of them. This comparison is less discussed in the literature and proved to be promising, since despite using different physical principles, the techniques obtained values in the same order of magnitude. ii) Development of an open source algorithm for analyzing ITC data, overcoming the difficulty that commercial software presents in calculating the parameters in more complex systems. The ITC data between Imp $\alpha$  and different NLSs were used as a reference for the comparison between the affinity constants obtained in the other techniques and as validation for the algorithm. The same ligands were used in the fluorescence spectroscopy, microscale thermophoresis and molecular docking assays. Despite being quite different, all techniques obtained dissociation constant values in the same order of magnitude. The values obtained by molecular docking are overestimated. As fluorescence and MST do not differentiate the constants for each binding site, the value found represents an average affinity for both sites. Note the greater similarity between the ITC and fluorescence data. The ITC algorithm had results similar to those obtained by the equipment software, its implementation will facilitate the analysis of complex data.

## LISTA DE FIGURAS

<b>Figura 1</b>	- Exemplo de um experimento padrão de calorimetria de titulação isotérmica indicado pelo fabricante	18
<b>Figura 2</b>	- Explicação de um experimento de MST. <b>A.</b> Equipamento utilizado. <b>B.</b> Representação esquemática da óptica MST. <b>C.</b> Sinal típico de um experimento. <b>D.</b> Dados formalizados e analisados.	26
<b>Figura 3</b>	- Visão geral da importação nuclear de proteínas	28
<b>Figura 4</b>	- Estrutura da Importina- $\alpha$ de <i>Mus musculus</i> (Mmlmp $\alpha$ ) destacando-se os dez motivos ARMs e o domínio IBB	29
<b>Figura 5</b>	- Verificação da indução e purificação da Mmlmp $\alpha$ em gel de poliacrilamida a 12% (SDS-PAGE), destacando a faixa da proteína recombinante, aproximadamente 55 kDa	34
<b>Figura 6</b>	- Cromatograma de purificação da proteína Mmlmp $\alpha$ por afinidade	35
<b>Figura 7</b>	- Termogramas de ITC da Mmlmp $\alpha$ complexada com as NLS: <b>A.</b> SV40. <b>B.</b> PMS2. <b>C.</b> MLH1	37
<b>Figura 8</b>	- Termogramas de ITC da Mmlmp $\alpha$ complexada com as NLS: <b>A.</b> XPG1. <b>B.</b> XPG2	37
<b>Figura 9</b>	- Termogramas de ITC da Mmlmp $\alpha$ complexada com as NLS: <b>A.</b> MLH1-R472K. <b>B.</b> MLH1-E475A. <b>C.</b> MLH1-R470A	38
<b>Figura 10</b>	- Gráfico de emissão de fluorescência em diferentes comprimentos de onda de excitação da Mmlmp $\alpha$ a 30 $\mu$ M	39
<b>Figura 11</b>	- Curvas da intensidade de absorvância por comprimento de onda. Cada cor representa a Mmlmp $\alpha$ depois de n injeções do peptídeo da MLH1, sendo o 0 a amostra isolada e as injeções indo de 1 a 20.	40
<b>Figura 12</b>	- Curvas da intensidade de absorvância normalizada por comprimento de onda. Cada cor representa a Mmlmp $\alpha$ depois de n injeções do peptídeo da MLH1, sendo o 0 a amostra isolada e as injeções indo de 1 a 20.	41
<b>Figura 13</b>	- Dados obtidos a partir da diferença entre os máximos das curvas de injeção de MLH1 NLS na Mmlmp $\alpha$ mostradas na figura 12, e a curva vermelha ajustada por uma função sigmóide.	42
<b>Figura 14</b>	- Dados obtidos a partir da diferença entre os máximos das curvas de injeção e a curva vermelha ajustada por uma função sigmóide da Mmlmp $\alpha$ com os ligantes. <b>A.</b> PMS2. <b>B.</b> SV40. <b>C.</b> XPG1. <b>D.</b> XPG2. <b>E.</b>	43

MLH1-R472K. **F.** MLH1-E475A. **G.** MLH1-R470A. **H.** MLH1-S467A. **I.** FEN1.

- Figura 15** - Dados de MST gerados com a Mmlmpα complexada aos ligantes. **A.** MLH1. **B.** MLH1<sup>a</sup>. **C.** SV40. **D.** XPG1<sup>a</sup>. **E.** XPG1. **F.** MLH1-S467A. **G.** FEN1. <sup>a</sup> Medida realizada com a proteína marcada pela cauda de histidina. 45
- Figura 16** - Gráfico de RMSD do peptídeo da MLH1 ligado no sítio principal da Mmlmpα gerado por simulação no GROMACS. 47
- Figura 17** - **A.** Gráfico de dispersão entre os valores de Kd obtidos via espectroscopia de fluorescência e ITC. **B.** Gráfico de dispersão entre os valores de Kd obtidos via MST e ITC. **C.** Gráfico de dispersão entre os valores de Kd obtidos via docagem molecular (PyRx) e ITC. **D.** Boxplot de distribuição das porcentagens de variação entre os valores obtidos em cada técnica e o ITC, variações acima de 1000% entre PyRx e ITC foram considerados *outliers* e desconsiderados na construção gráfica. 49
- Figura 18** - Gráficos de ITC gerados no R 51

**LISTA DE TABELAS**

<b>Tabela 1</b>	- Principais técnicas empregadas no estudo da interação proteína:ligante	13
<b>Tabela 2</b>	- Dados termodinâmicos de diferentes peptídeos complexados com a Mmlmpα	36
<b>Tabela 3</b>	- Dados de Kd gerados por espectroscopia de fluorescência de diferentes peptídeos complexados com a Mmlmpα	42
<b>Tabela 4</b>	- Dados de Kd gerados por MST de diferentes peptídeos complexados com a Mmlmpα	46
<b>Tabela 5</b>	- Constantes de dissociação de diferentes peptídeos complexados com a proteína Mmlmpα obtidas por diferentes técnicas	48
<b>Tabela 6</b>	- Constantes de dissociação calculadas pelo software do fabricante e pelo algoritmo	53

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	12
1.1 TÉCNICAS BIOFÍSICAS PARA OBTENÇÃO DE DADOS DE AFINIDADE ENTRE BIOMOLÉCULAS	12
1.2 ITC E O PROCESSAMENTO DE DADOS	16
1.3 TERMOFORESE EM MICROESCALA	25
1.4 VIA CLÁSSICA DE IMPORTAÇÃO NUCLEAR	26
1.5 ESTRUTURA CRISTALOGRÁFICA DA IMPORTINA- $\alpha$	28
1.6 SEQUÊNCIAS DE LOCALIZAÇÃO NUCLEAR	29
<b>2 OBJETIVO</b>	30
<b>3 MATERIAIS E MÉTODOS</b>	30
3.1 EXPRESSÃO E PURIFICAÇÃO DE IMPORTINA- $\alpha$ TRUNCADA (70-529)	30
3.2 CALORIMETRIA POR TITULAÇÃO ISOTÉRMICA	31
3.3 ESPECTROSCOPIA DE FLUORESCÊNCIA	32
3.4 TERMOFORESE EM MICROESCALA	32
3.5 TÉCNICAS DE BIOINFORMÁTICA PARA O CÁLCULO DE CONSTANTE DE AFINIDADE	33
3.6 DESENVOLVIMENTO DE UM ALGORITMO PARA ANÁLISE DE DADOS DE ITC	33
<b>4 RESULTADOS E DISCUSSÃO</b>	34
4.1 EXPRESSÃO E PURIFICAÇÃO	34
4.2 CALORIMETRIA POR TITULAÇÃO ISOTÉRMICA	35
4.3 ESPECTROSCOPIA DE FLUORESCÊNCIA	38
4.4 TERMOFORESE EM MICROESCALA	44
4.5 TÉCNICAS DE BIOINFORMÁTICA PARA O CÁLCULO DE CONSTANTE DE AFINIDADE	46
4.6 ANÁLISE COMPARATIVA ENTRE AS TÉCNICAS	47
4.7 DESENVOLVIMENTO DE UM ALGORITMO PARA ANÁLISE DE DADOS DE ITC	50
<b>5 CONCLUSÃO</b>	54
REFERÊNCIAS	56

## 1 INTRODUÇÃO

### 1.1 TÉCNICAS BIOFÍSICAS PARA OBTENÇÃO DE DADOS DE AFINIDADE ENTRE BIOMOLÉCULAS

Métodos biofísicos vêm sendo cada vez mais utilizados como abordagens complementares às tradicionais na busca do conhecimento aprofundado dos mecanismos moleculares. Assim, técnicas biofísicas aplicadas no estudo da interação proteína/ligantes ou proteínas/proteínas fornecem informações importantes, permitindo não só entender o processo como otimizar e listar compostos com maior afinidade (Holdgate, Anderson et al. 2010).

Existem diversas técnicas que predizem a afinidade entre duas ou mais moléculas em solução, entre elas estão calorimetria por titulação isotérmica (*Isothermal Titration Calorimetry* – ITC), ressonância de plasma de superfície (*Surface Plasmon Resonance* – SPR), espectrometria de massa (Mass Spectrometry – MS), termoforese em microescala (*MicroScale Thermophoresis* – MST), ultracentrifugação analítica (*Analytical Ultracentrifugation* - AUC), técnicas espectroscópicas (ressonância magnética nuclear (*Nuclear Magnetic Resonance* – NMR) e espectroscopia de fluorescência) entre outras. A tabela 1 apresenta essas técnicas com suas principais vantagens e desvantagens.

O ITC é a técnica estabelecida como padrão ouro para medir diretamente a afinidade de ligação e a termodinâmica da reação (Freyer and Lewis 2008). A técnica permite que a afinidade, a entalpia e a estequiometria de uma interação sejam medidas em um único experimento. Além disso, avanços recentes na sensibilidade, a não necessidade de um marcador e a utilização de um baixo volume da amostra em solução trazem vantagens para a escolha do ITC. A técnica envolve a quantificação do calor produzido ou absorvido durante a reação de ligação à temperatura constante (Ladbury 2010). Porém, uma dificuldade na utilização do ITC é o fato do processamento dos dados, em sistemas mais complexos, não ser tão bem resolvidos pelos softwares comerciais disponíveis (Velazquez Campoy and Freire 2005). Outra desvantagem é que proteína e ligante devem estar diluídos no mesmo tampão. Ademais, a utilização de solventes orgânicos em concentrações superiores a 5-10% (volume:volume) para diluição das espécies titulantes ou

tituladas não são recomendadas pois podem induzir a erros na determinação das constantes de afinidade e termodinâmicas da reação.

**Tabela 1** – Principais técnicas empregadas no estudo da interação proteína:ligante

Técnica	Princípio Físico	Vantagens	Desvantagens
ITC	Quantificação do calor gerado ou absorvido na reação	Pouca amostra, livre de marcadores, obtenção direta da estequiometria e entalpia da reação	Processamento de dados não é tão simples em sistemas complexos e limitação no tampão
SPR	Alteração do índice de refração devido à massa	Pouca amostra e automatizado	Apenas proteínas de massa molecular maior e necessário acoplamento ao filme metálico
MS	Identificação e análise do ligante; concentração do produto da reação	Reconhecimento do ligante no caso dele não ser conhecido e há não necessidade de purificar a proteína	Utilização de outras técnicas previamente, custo e não há quantificação
MST	Alterações na mobilidade termoforética	Utilização da amostra em qualquer tampão e pouca amostra	Necessidade de marcação e uso de capilares para manuseio da amostra
AUC	Sedimentação de proteínas em solução	Alta resolução e estudo do equilíbrio	Custo e tempo de aquisição de dados

NMR	Campos magnéticos locais em torno de núcleos atômicos	Informação estrutural	Muita amostra e limitação da massa molecular
Espectroscopia de Fluorescência	Variação na emissão e absorção de luz	Não destrutivo e interação de baixa afinidade	Necessidade da presença de fluoróforos

A Ressonância Plasmônica de Superfície (SPR) é uma técnica poderosa para a análise de afinidade entre biomoléculas, baseada na detecção de mudanças no índice de refração próximo à superfície de um filme metálico quando este é iluminado por luz polarizada sob condições de reflexão interna total. A principal vantagem do SPR é a capacidade de monitorar interações biomoleculares em tempo real, sem a necessidade de marcadores, e de fornecer informações cinéticas e termodinâmicas, como constantes de associação e dissociação (RICH; MORTON, 2005). No entanto, a técnica é mais adequada para moléculas grandes, como proteínas, e apresenta limitações na sensibilidade para moléculas menores. Além disso, a orientação imprecisa das biomoléculas na superfície metálica pode introduzir erros significativos na análise, comprometendo a precisão dos resultados (HUBER; MUELLER, 2006). Apesar dessas desvantagens, o SPR continua sendo amplamente utilizado em estudos de interação molecular devido à sua versatilidade e capacidade de operar em condições próximas ao ambiente fisiológico (SCHASFOORT et al., 2003).

A espectrometria de massa (MS) pode ser utilizada no estudo proteína-ligante de quatro formas diferentes: (a) identificação direta do ligante, (b) detecção do complexo proteína:ligante, (c) análise do sítio de ligação e, (d) monitoração do substrato ou a concentração do produto da reação (Holdgate, Anderson et al. 2010). A MS é realizada com o produto de outras técnicas, como por exemplo, ensaios enzimáticos ou ultracentrifugação, isso gera um gasto de amostra e custo muito maiores além de não gerar valores que quantificam a reação.

A termoforese em microescala (MST) é uma técnica importante para quantificar as interações biomoleculares. É baseado no movimento dirigido de moléculas em um gradiente de temperatura, que depende fortemente de diversas propriedades moleculares, como tamanho, carga, camada de hidratação ou

conformação. Assim, é altamente sensível a qualquer alteração nas propriedades moleculares, permitindo uma quantificação precisa de eventos moleculares independentes do tamanho ou natureza da espécie investigada (Jerabek-Willemsen, Wienken et al. 2011). O MST permite uma determinação precisa de constantes de ligação, estequiometria da interação, modos de ligação, análise de desdobramento de proteínas, termodinâmica e cinética enzimática. Uma vantagem é que o MST pode ser realizado em praticamente qualquer tampão, mesmo no plasma e no lisado celular. Desvantagens dessa técnica são a necessidade, na maioria das vezes, de marcar a proteína com fluoróforos e a necessidade de capilares para manuseio da amostra, gerando um custo considerável para sua utilização.

A ultracentrifugação analítica (AUC) monitora e interpreta a evolução dos perfis de concentração macromolecular após a aplicação de uma força centrífuga (Zhao, Ghirlando et al. 2015). Atualmente, há três tipos de sistemas ópticos disponíveis para AUC: o espectrofotômetro convencional, o interferômetro Rayleigh e a fluorescência, que oferecem uma ampla gama para a investigação das interações entre proteínas sendo possível calcular constantes de equilíbrio de dissociação e a estequiometria da reação (Zhao, Ghirlando et al. 2015). É um método vantajoso no estudo de interações de proteínas, principalmente porque o forte processo de sedimentação dependente do tamanho da macromolécula leva a uma alta resolução. Além disso, no modelo experimental padrão, apesar de sua velocidade de sedimentação mais alta, os complexos podem se associar novamente durante o processo, de um modo que reflita seu equilíbrio (Zhao, Ghirlando et al. 2015). Porém, apesar de ter sofrido melhorias substanciais em instrumentação, teoria e análise de dados computacionais, o processo de AUC ainda gera um alto custo e um tempo de centrifugação, em muitos casos, longo (Zhao, Ghirlando et al. 2015).

Estudos de ligação com NMR podem ser realizados em uma ampla gama de afinidades, no entanto, a maior utilidade dessa técnica é quando há a ligação com baixa afinidade que podem ser detectados de forma confiável (Holdgate, Anderson et al. 2010). NMR é uma técnica espectroscópica que observa campos magnéticos locais em torno de núcleos atômicos. A amostra é colocada sob ação de um campo magnético e o sinal de NMR é produzido pela excitação dos prótons da amostra com ondas eletromagnéticas, que é detectado com receptores de rádio frequência. É uma ferramenta comum para a determinação das relações de

atividade de conformação onde a estrutura antes e depois da interação com um ligante é comparada. As desvantagens são a utilização de uma grande quantidade de amostra e limitações quanto à massa da proteína (por exemplo 40 kDa dependendo do campo magnético aplicado e outros fatores).

Na espectroscopia de fluorescência o sinal é gerado pela quantificação da onda eletromagnética emitida pela amostra excitada por uma luz monocromática em um comprimento de onda específico que é absorvido pela molécula. Em proteínas, a análise na variação do sinal emitido pelos aminoácidos fluorescentes traz informações importantes sobre o estado proteico antes e após a interação (Jameson and Seifried 1999). Essa técnica é adequada para a medição de interações de baixa afinidade (mais especificamente, interações com taxas de dissociação rápidas). Uma vantagem da técnica é que o experimento não é destrutivo, permitindo assim medições repetitivas da mesma amostra sob diferentes condições (por exemplo, temperaturas diferentes) (Rossi and Taylor 2011). Uma das maiores desvantagens da espectroscopia de fluorescência é a necessidade da presença de fluoróforos na amostra.

Apesar de existirem, como mencionando, diversas técnicas biofísicas capazes de estimar a estequiometria e as constantes termodinâmicas de uma interação proteica, deve-se ter cautela ao comparar as afinidades obtidas por diferentes técnicas, já que podem não gerar valores similares, assim, um estudo comparativo, para o mesmo maquinário biológico, ganha relevância, podendo, então, entender as especificidades e similaridades de cada técnica em relação à outra. Atualmente, poucos estudos foram realizados visando à comparação de dados obtidos com diferentes técnicas em sistemas biológicos similares (Holdgate, Anderson et al. 2010; Zhao, Ghirlando et al. 2015).

## 1.2 ITC E O PROCESSAMENTO DE DADOS

Calorimetria por titulação isotérmica (ITC) é a metodologia que oferece maior número de parâmetros relativos a uma interação molecular e permite a medida direta da entalpia de qualquer reação em condições isotérmica e isobárica. É o único método capaz de determinar a entalpia, entropia e a energia livre de ligação em apenas uma titulação, em uma única temperatura (Ladbury 2010). Dentre as técnicas experimentais empregadas para estudar a ligação entre

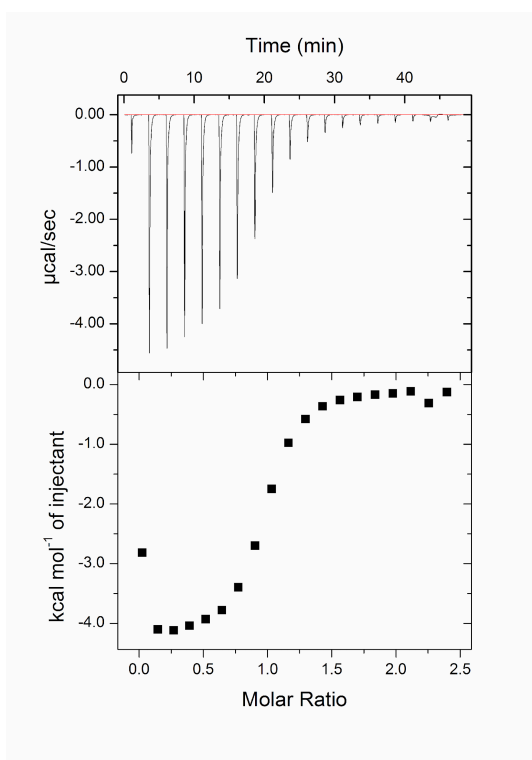
macromoléculas, o ITC apresenta várias características que o tornam uma ferramenta experimental única: evita o uso de marcadores espectroscópicos não naturais; as amostras estão em solução; e permite determinar simultaneamente a constante de dissociação, a entalpia e a estequiometria de ligação em um único experimento (Freire, Schon et al. 2009). Consequentemente, os dados do ITC são bastante adequados para serem analisados utilizando o formalismo polinomial de ligação.

No calorímetro, anteriormente à titulação, uma potência constante é aplicada à célula de referência. Este sinal aciona o aquecedor da célula de reação, e esta potência aplicada representa a linha de base (Figura 1, painel superior, linha vermelha). A quantidade diretamente mensurada em um experimento de ITC é a potência de entrada dependente do tempo para manter as células de referência e de reação à mesma temperatura. Durante a injeção do titulante na célula de reação, calor é absorvido ou liberado dependendo se a natureza da interação entre as macromoléculas é endo ou exotérmica, respectivamente. Para uma reação exotérmica, a temperatura na célula de reação tenderia a subir e com isso, a potência de *feedback* vai ser diminuída, para manter as duas células à mesma temperatura (Figura 1, painel superior, linha preta). Nas reações endotérmicas, a potência vai ser aumentada para a manutenção da temperatura constante, uma vez que a interação vai retirar energia do ambiente (Velazquez Campoy and Freire 2005). O calor gerado ou absorvido durante a titulação é proporcional à fração de ligação da espécie titulante ao titulado. Por isso, é de extrema importância a correta determinação das concentrações iniciais das espécies titulante e titulada para as injeções iniciais, fazendo com que todo ou quase todo o titulante se ligue ao titulado, resultando em considerável sinal exo ou endotérmico, dependendo da natureza da associação no início da titulação (Pierce, Raman et al. 1999; Freire, Schon et al. 2009).

No termograma gerado em um experimento típico de ITC, os calores iniciais decorrentes da reação entre ambas as espécies são maiores que os gerados nas injeções subsequentes do titulante, pois no início da titulação há um excesso de sítios de ligação não ocupados na espécie titulada. Com isso, os calores gerados no início da titulação são resultados da reação “completa” do titulante adicionado. Com o prosseguimento da titulação, menores quantidades de titulante se ligam à espécie

titulada e novas espécies passam a aparecer na célula de reação: titulante livre, titulado com sítios de ligação não ocupados e complexo formado. Já nas últimas injeções é de fundamental importância que a espécie titulada esteja toda saturada e que o calor gerado ou absorvido seja proporcional ao calor de diluição do titulante (Biswas and Tsodikov 2010), acabando a titulação majoritariamente com moléculas do titulante livres e moléculas do complexo titulante-titulado formadas.

No tratamento dos dados gerados pelo calorímetro é possível calcular a variação de calor das reações de titulação através da integração da potência aplicada durante a titulação ( $\mu\text{cal/s}$ ) normalizada pela concentração do titulante em função da razão molar das espécies titulante e titulada (painel inferior, Figura 1). Com base nesse termograma é possível determinar as variações de entalpia ( $\Delta H$ ), a estequiometria da titulação ( $n$ ) e a constante de dissociação ( $K_d$ ) entre as espécies tituladas. Com base nesses parâmetros é possível calcular a variação da entropia ( $\Delta S$ ) e a energia livre de ligação ( $\Delta G$ ).



**Figura 1** – Exemplo de um experimento padrão de calorimetria de titulação isotérmica indicado pelo fabricante. Titulação de íons cálcio (5 mM) em EDTA (0.4 mM), ambos diluídos em tampão MES (10 mM) à 25 °C. (Painel superior) Dados brutos de calorimetria de

titulação isotérmica (linha preta) e linha de base (linha vermelha). (Painel inferior) Isoterma da titulação, que representa a energia medida em cada injeção normalizada pela molaridade do titulante (eixo das ordenadas) em função da razão molar de titulante por titulado (eixo das abcissas)

Considerando-se uma reação química do tipo  $M + L \rightleftharpoons ML$ , em que a substância  $M$  é uma macromolécula e está alocada na célula de reação, também chamada de titulado,  $L$  é um ligante e está colocado na pipeta, chamado titulante, e  $ML$  é o complexo formado durante a titulação, o equilíbrio entre as espécies pode ser descrito matematicamente pelas equações (Velazquez Campoy and Freire 2005):

$$[M]_T = [M] + [ML] \quad (1)$$

$$[L]_T = [L] + [ML] \quad (2)$$

em que  $[M]_T$  e  $[L]_T$  representam as concentrações totais do titulado e do titulante e  $[M]$ ,  $[L]$  e  $[ML]$  as concentrações na condição de equilíbrio do titulado, titulante e do complexo titulado-titulante, respectivamente. A condição de equilíbrio da reação é dada por:

$$[ML] = K_a [M][L] \quad (3)$$

em que  $K_a$  é a constante de afinidade entre as espécies.

Resolvendo as equações (1) e (2) em função da concentração do complexo ( $[ML]$ ):

$$[M]_T = [ML] + \frac{[ML]}{(K_a([L]_T - [ML]))} \quad (4)$$

obtém-se a equação quadrática a seguir:

$$[ML]^2 - [ML] \left( [M]_T + [L]_T + \frac{1}{K_a} \right) + [M]_T [L]_T = 0 \quad (5)$$

que dá duas possíveis soluções, sendo a fisicamente relevante, a solução descrita abaixo:

$$[ML] = \frac{[M]_T + [L]_T + \frac{1}{K_a} - \sqrt{\left([M]_T + [L]_T + \frac{1}{K_a}\right)^2 - 4[M]_T[L]_T}}{2} \quad (6)$$

Usando o parâmetro adimensional "r" introduzido por Wiseman e colaboradores (Wiseman, Williston et al. 1989) a equação seguinte é obtida:

$$[ML] = [M]_T \frac{1 + r + X_R - \sqrt{(1 + r + X_R)^2 - 4X_R}}{2} \quad (7)$$

em que  $r = 1/(K_a[M]_T)$  e  $X_R = [L]_T/[M]_T$ .

A quantidade de calor total ( $Q_T$ ) gerada ou absorvida no calorímetro devido à formação do complexo  $ML$  é dada por

$$Q_T = V_0[ML]\Delta H = V_0\Delta H[M]_T \frac{1 + r + X_R - \sqrt{(1 + r + X_R)^2 - 4X_R}}{2} \quad (8)$$

em que  $V_0$  é o volume da cela de reação e  $\Delta H$  é a entalpia de ligação.

A representação da isoterma de Wiseman, forma mais recomendada para apresentação dos resultados de ITC, é obtida pelo cálculo da derivada  $dQ_T/V_0d[L]_T$ , que é a forma diferencial equivalente ao calor incremental por injeção,  $\Delta Q_T$ , normalizado pelo incremento molar do ligante injetado na cela de reação,  $\Delta n_{LT}$ ,

$$\frac{\Delta Q_T}{\Delta n_{LT}} \approx \frac{dQ_T}{V_0d[L]_T} = \frac{d[ML]}{d[L]_T} \Delta H \quad (9)$$

A derivada com respeito a  $[L]_T$  pode ser convertida em uma derivada relacionada à razão molar,  $X_R$ , aplicando a relação

$$\frac{d}{d[L]_T} = \frac{d}{dX_R} \frac{dX_R}{d[L]_T} = \frac{1}{[M]_T} \frac{d}{dX_R} \quad (10)$$

que considera a concentração da macromolécula constante durante a titulação. Aplicando-se as equações (9) e (10) à equação (8), obtém-se a isoterma de Wiseman:

$$\frac{dQ_T}{V_0 d[L]_T} = \left( \frac{1}{2} + \frac{1}{2} \frac{1-r-X_R}{\sqrt{X_R^2 - 2X_R(1-r) + (1+r)^2}} \right) \Delta H \quad (11)$$

A isoterma de Wiseman (Wiseman, Williston et al. 1989) é a representação da  $dQ_T/V_0 d[L]_T$  em função da razão molar,  $X_R$ , que é calculada assumindo que não há diluição nem da macromolécula nem do ligante utilizado na titulação que, segundo Velazquez-Campoy (Velazquez Campoy and Freire 2005), é igual à obtida usando-se um quociente incremental ao invés da derivada em função de  $[L]_T$ , e da diluição da macromolécula e do ligante em função das injeções. Isso se deve ao fato que a diluição da macromolécula está implicitamente considerada quando do cálculo dos valores da razão molar ao longo da análise dos dados.

Os processos de ajustes dos dados são realizados tendo como base o conhecimento do tipo de interação entre a espécie titulada e titulante, sendo que alguns parâmetros devem ser incluídos no ajuste dos dados baseados nos modelos matemáticos que descrevem o equilíbrio da ligação: (i) estequiometria da reação; (ii) constante de afinidade (associação) ou dissociação,  $K_a$  ou  $K_d$ , respectivamente; (iii) caracterização das contribuições entálpica e entrópica para a energia livre de ligação; (iv) caracterização de múltiplos sítios e cooperatividade positiva ou negativa para os casos de múltiplos sítios; (v) caracterização da dependência do equilíbrio de ligação com variáveis como pH, força iônica do tampão, entre outros (Freire, van Osdol et al. 1990). Os parâmetros acima relacionados estão abaixo explicados.

A estequiometria,  $n$ , de uma reação descreve as relações quantitativas entre duas ou mais substâncias em uma reação física ou química. Na calorimetria de titulação isotérmica, a determinação da estequiometria de ligação tem um papel fundamental na caracterização dos mecanismos de ligação de macromoléculas biológicas (Doyle 1997).

Com base nas concentrações de ambas as espécies em interação, a estequiometria de ligação pode ser determinada através da razão molar das

espécies no ponto de inflexão do termograma. No processo de análise de dados, o valor da estequiometria pode ser utilizado como um valor variável, a ser ajustado no processo de minimização de erros dos parâmetros experimentais, como pode ser mantido fixo, caso se tenha informação do sistema em estudo. Uma grande quantidade de fatores pode influenciar o valor da estequiometria, fazendo com que seu valor seja um valor não inteiro. Os principais determinantes para isso são: erro nas concentrações da espécie titulada (Freyer and Lewis 2008), ligação inespecífica, degradação ou baixa qualidade do titulado (estado inativo, sem *folding* correto) (Tellinghuisen 2004; Tellinghuisen 2004; Gruner, Neeb et al. 2014). Por isso, a análise da estequiometria da reação também pode ser considerada um excelente parâmetro de qualidade na análise de ensaios de ligação de diferentes lotes de proteínas purificadas em diferentes condições, bem como na análise de proteínas recombinantes expressas em diferentes linhagens celulares ou mesmo em testes de estabilidade de congelamento-descongelamento (Doyle 1997).

Outra característica que, assim como a variação de entalpia, é medida diretamente na isoterma de ligação é a constante de afinidade ou associação ( $K_a$ ), que está relacionada com a constante de dissociação ( $K_d$ ) através da equação

$$K_d = \frac{1}{K_a} \quad (12)$$

sendo estas diretamente relacionadas à energia livre da ligação ( $\Delta G$ ) através da equação

$$\Delta G = -RT \ln K_a = RT \ln K_d \quad (13)$$

A energia livre de ligação é o parâmetro termodinâmico mais importante de uma ligação, uma vez que seu valor, em condições específicas de concentrações dos reagentes, direciona o equilíbrio biomolecular. Se o valor de  $\Delta G$  é negativo, a reação de ligação ou transição conformacional será espontânea numa medida ditada pela magnitude do  $\Delta G$ . Se o sinal é positivo, a magnitude da variação específica a quantidade de energia necessária para a formação dos produtos da reação. Além disso, este parâmetro quantifica a afinidade da ligação entre 2 moléculas em uma interação (Ladbury 2010), determina a estabilidade do complexo formado e ainda é

uma ferramenta analítica de caracterização fenomenológica da relação estrutura-função (Perozzo, Folkers et al. 2004).

Para determinar os parâmetros termodinâmicos relevantes às interações entre as moléculas, geralmente seguem-se alguns passos (Velazquez-Campoy, Ohtaka et al. 2004):

1- Primeiramente, ajusta-se a linha de base no termograma e integra-se os picos.

2- Em seguida, verifica-se as concentrações das espécies tituladas e titulantes, que inicialmente podem ser estimadas pela estequiometria da ligação. Se necessário, realiza-se correções nas concentrações.

3- Depois, estima-se os calores de diluição, seja pelas últimas injeções da titulação realizada ou através de experimentos controle, caso o calor de diluição da espécie titulante dependa da concentração. Este calor de diluição, ou mesmo o de interações inespecíficas, deve ser subtraído dos calores calculados nas injeções da titulação para uma melhor estimativa dos dados termodinâmicos.

4- O passo final é determinar o modelo de ligação a ser utilizado para ajustar os dados da isoterma de ligação da titulação (Velazquez-Campoy, Ohtaka et al. 2004). Os modelos de ligação são descrições matemáticas dos processos biológicos, físicos e químicos que descrevem a interação entre as espécies tituladas. As variáveis dependentes (calor ou fluxo de calor) são definidas como função das variáveis independentes (molaridade de titulante) e de outros parâmetros do modelo. Por isso, usa-se o modelo mais simples possível que corrobore as informações já conhecidas do sistema em estudo, de modo que as equações do modelo sejam utilizadas para determinar os melhores valores para os parâmetros de ajuste à isoterma, como a constante de associação ou dissociação ( $K_a$  ou  $K_d$ , respectivamente), variação de entalpia ( $\Delta H$ ) e estequiometria ( $n$ ) (Freyer and Lewis 2008).

5- Caso não se tenham informações acerca do sistema em estudo, diversos modelos podem ser ajustados aos dados experimentais, e o modelo mais adequado, ou seja, o que apresentar melhor ajuste dos dados, pode ser escolhido estatisticamente.

Entretanto, o modo padrão de análise dos dados possui algumas limitações significativas. Dentre os problemas frequentemente encontrados na análise dos dados obtidos estão: (a) problemas na determinação da linha de base; (b) análise de dados com modelos de alta complexidade de um sistema sobre o qual não se tenham outras informações obtidas por metodologias ortogonais; (c) variações encontradas na realização do mesmo experimento entre diferentes laboratórios e operadores.

Uma macromolécula com dois sítios de ligação apresenta pelo menos seis possíveis mecanismos de ligação diferentes: *i)* dois sítios de ligação idênticos e independentes; *ii)* dois sítios idênticos e competitivos; *iii)* dois sítios idênticos e positivamente cooperativos; *iv)* dois sítios diferentes e independentes; *v)* dois sítios diferentes e competitivos e *vi)* dois sítios diferentes e positivamente cooperativos. Nem todos esses casos são distinguíveis experimentalmente, ou seja, alguns dão origem exatamente à mesma curva de ligação e assim informações extras termodinâmicas são necessárias para elucidar o mecanismo de ligação. Por isso, a análise de dados ainda se apresenta, nesses casos, com limitações significativas. Assim, os experimentos devem ser realizados de modo que as fontes de incertezas sejam minimizadas ou determinadas e subtraídas na análise de dados. As principais fontes de incertezas são: quantificação do calor liberado pelo equipamento, erros nas concentrações das espécies titulada e titulante, erros na linha de base por problema no equipamento ou variações ambientais, variações das soluções tampão utilizadas; problemas nas operações com o equipamento, como bolhas de ar ou uso de velocidades de agitação não ótimas (Tellinghuisen and Nowak 2003; Tellinghuisen 2003; Tellinghuisen 2005; Baranauskiene, Petrikaite et al. 2009; Pethica 2015).

O *software* comercial de análise de dados que acompanha o equipamento (Origin v.7.0, OriginLab) apresenta poucos modelos de ligação já implementados e, por isso, algumas vezes o usuário deve desenvolver suas próprias rotinas de ajuste ou usar as já implementadas assumindo algumas simplificações (Vega, Abian et al. 2015).

A proteína Importina- $\alpha$ , envolvida no processo de importação nuclear, é um exemplo de sistema complexo que pode ser interessante para o processamento de dados calorimétricos. Esta proteína possui dois sítios de ligação independentes com

afinidades diferentes pelos ligantes, trazendo desafios quanto ao processamento de dados calorimétricos.

Assim, a criação de um algoritmo que analise os dados experimentais de vários graus de complexidade, de interações simples às mais complexas, se faz cada vez mais necessário, no intuito de tornar a interpretação de dados de qualquer sistema mais simplificada.

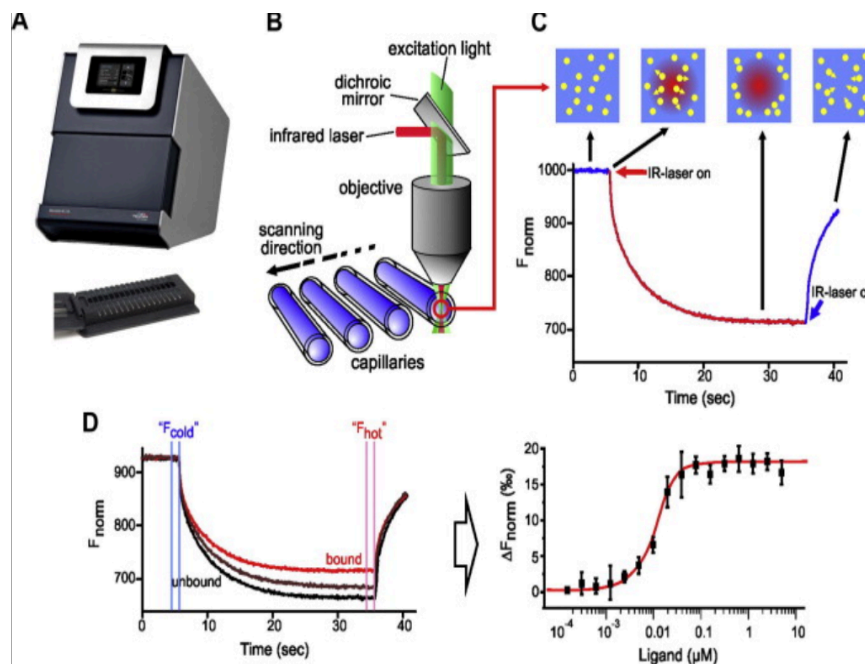
### 1.3 TERMOFORESE EM MICROESCALA

Moléculas fluorescentes são usadas em ciências biológicas para rastrear e analisar interações moleculares. As técnicas baseadas em fluorescência tornaram-se inestimáveis para uma compreensão quantitativa dos processos biológicos e a termoforese em microescala (MST) é uma abordagem totalmente óptica para caracterizar as propriedades das biomoléculas (Seidel, Dijkman et al. 2013).

O fenômeno da termoforese descreve o movimento direcionado de moléculas através de um gradiente de temperatura. A termoforese em microescala (MST) é um método para monitorar o movimento direcionado de moléculas fluorescentes através de gradientes microscópicos de temperatura, permitindo a análise precisa de eventos de ligação em uma solução de poucos microlitros de praticamente qualquer molécula - independente de tamanho ou propriedades físicas (Jerabek-Willemsen, Wienken et al. 2011). Sob condições tampão constantes, a termoforese investiga o tamanho, a carga e a entropia de solvatação das moléculas. A termoforese de uma proteína normalmente difere significativamente da termoforese de um complexo proteína-ligante devido às alterações induzidas pela ligação no tamanho, carga e energia de solvatação. Mesmo que uma ligação não altere significativamente o tamanho ou a carga de uma proteína, a MicroScale Thermoforese ainda pode detectar a ligação devido a alterações induzidas pela ligação na entropia de solvatação das moléculas (Seidel, Dijkman et al. 2013).

O funcionamento do MST está desenhado na figura 2. No equipamento (figura 2A), a termoforese é induzida e detectada em pequenos capilares de vidro que contêm uma solução da molécula marcada e um gradiente de concentração do ligante. Um laser infravermelho (IR) produz um gradiente de temperatura, assim os

fluoróforos em solução são excitados e a fluorescência emitida é coletada (Figura 2B). Esta configuração permite acompanhar a depleção de fluoróforos dentro do gradiente de temperatura induzido (Figura 2C). Mudanças na termoforese das moléculas fluorescentes devido à ligação ao ligante podem então ser usadas para calcular constantes de ligação de equilíbrio (Figura 2D).



**Figura 2** – Explicação de um experimento de MST. **A.** Equipamento utilizado. **B.** Representação esquemática da óptica MST. **C.** Sinal típico de um experimento. **D.** Dados formalizados e analisados. (Jerabek-Willemsen, Wienken et al. 2011).

#### 1.4 VIA CLÁSSICA DE IMPORTAÇÃO NUCLEAR

A via clássica de importação nuclear de proteínas permite o transporte de um grande número de proteínas, utilizando o heterodímero importina- $\alpha$  (Imp $\alpha$ ) e importina- $\beta$  (Imp $\beta$ ) como transportador, e tem sido estudada em detalhes biológicos e estruturais.

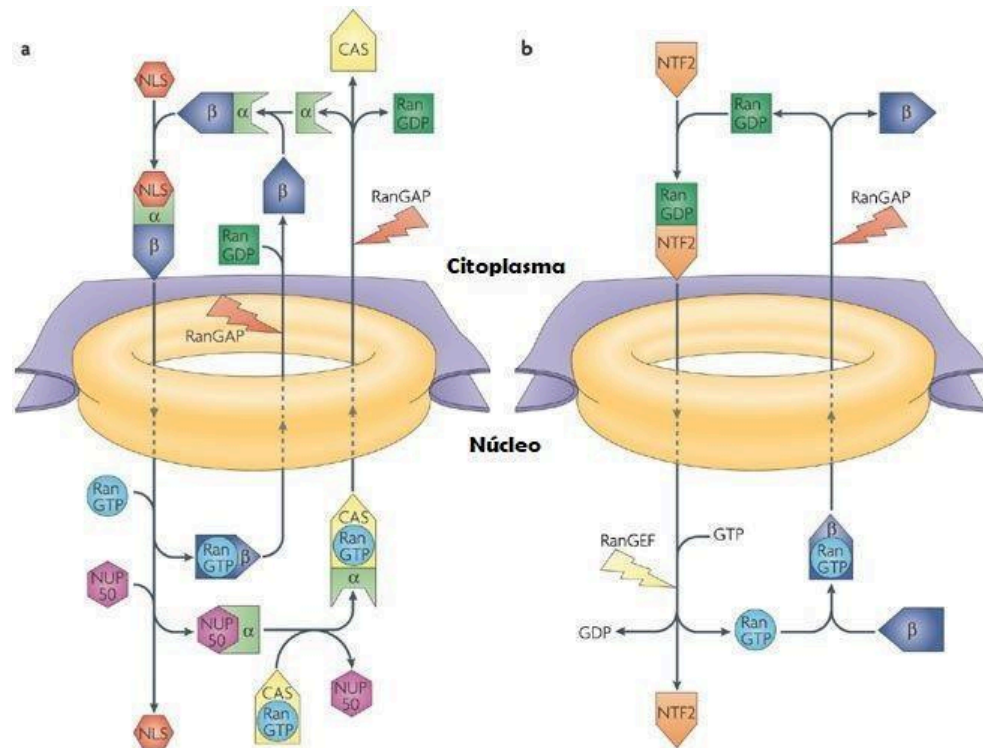
O processo de importação nuclear tem como princípio uma cadeia de interações, que precisam ser constantemente reguladas, esse ciclo pode ser convenientemente dividido em quatro etapas: I) Formação do complexo de importação carga proteica/transportador no citoplasma; II) Translocação através de

complexos poro-nucleares (Nuclear Pore Complex - NPCs); III) Desligamento do complexo de importação no núcleo; IV) Reciclagem das proteínas transportadoras.

Sinais específicos contribuem para a fidelidade desses eventos, dentre os quais as sequências de localização nuclear (*Nuclear Localization Signal* - NLS) (Stewart 2007). Proteínas que possuem NLS são importadas para o núcleo via heterodímero Imp $\alpha$ /Imp $\beta$ , o qual se liga a estas proteínas através do adaptador Imp $\alpha$  (Gorlich and Kutay 1999; Chook and Blobel 2001; Macara 2001; Weis 2003), a qual possui sítios de reconhecimento específicos (Conti, Uy et al. 1998; Conti and Kuriyan 2000; Fontes, Teh et al. 2000; Stewart 2006; Stewart 2007). Esse processo ocorre inicialmente no citoplasma, após a formação do heterodímero (Imp $\alpha$ /Imp $\beta$ ), é possível a ligação da proteína a ser importada, via reconhecimento do NLS (Gorlich, Vogel et al. 1995), como ilustrado na figura 1. Em seguida a Imp $\beta$  ancora o complexo aos filamentos citoplasmáticos do NPC e ocorre o transporte através do poro (NPC) (Moore and Blobel 1993; Gorlich and Mattaj 1996).

No núcleo, o complexo proteína-substrato se dissocia com a ligação da proteína Ran ligada a GTP (Ran-GTP) à subunidade  $\beta$ , alterando sua conformação e liberando o domínio auto-inibitório (*Importin beta binding* - IBB) da Imp $\alpha$ . Esse domínio então passa a competir com a NLS da proteína importada, promovendo uma diminuição de afinidade, facilitando a liberação dessa proteína no núcleo (Kobe 1999). Uma otimização do processo de desacoplamento da carga e reciclagem da Imp $\alpha$  ocorre com auxílio da nucleoporina (Nup2 em levedura, Nup50 ou Npap60 em vertebrados) (Lange, Mills et al. 2007), que se liga a uma das regiões de ligação do NLS e uma porção C-terminal da Imp $\alpha$ , em virtude de maior afinidade com a mesma. A nucleoporina liga-se à superfície estendida do domínio de repetição ARM da Imp $\alpha$ , incluindo o sítio de ligação secundário de NLS, liberando a proteína transportada e o sítio de ligação (Matsuura and Stewart 2005).

Após a carga ter sido liberada pelo deslocamento do NLS pela nucleoporina, a mesma é deslocada pela ligação com a exportina (Cse1/RanGTP em leveduras ou CAS/RanGTP em vertebrados) na região C-terminal da Imp $\alpha$ , e assim a ligação sequencial das nucleoporinas, seguido por exportinas assegura que Imp $\alpha$  será reciclada para o citoplasma, após a liberação da sua carga (Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003; Lange, Mills et al. 2007) (Figura 3).



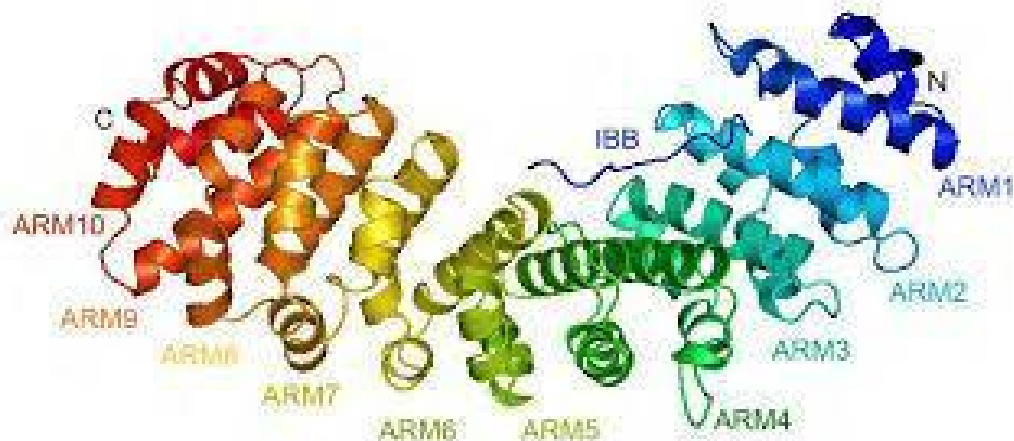
**Figura 3** - Visão geral da importação nuclear de proteínas. **A.** No citoplasma, um complexo de importação entre NLS e o heterodímero Imp $\alpha$ /Imp $\beta$  é formado e **B.** Ciclos da Ran entre núcleo e citoplasma (Stewart 2007, adaptado).

### 1.5 ESTRUTURA CRISTALOGRÁFICA DA IMPORTINA- $\alpha$

A Imp $\alpha$  pertence à família das proteínas chamadas carioferinas e apresenta cerca de 55 kDa. A Imp $\alpha$  apresenta uma estrutura conservada em diferentes organismos, sendo que alguns destes organismos apresentam mais de uma isoforma. Pela análise das sequências dessas proteínas, foi possível classificá-las em três famílias: Kap- $\alpha$ 1, Kap- $\alpha$ 2, e Kap- $\alpha$ 3 (Kohler, Haller et al. 1999). As diferenças entre as proteínas de cada família podem estar relacionadas com especificidades da Imp $\alpha$  em cada organismo.

Com a elucidação da estrutura molecular da Imp $\alpha$  de organismos modelo como *Saccharomyces cerevisiae* ((Conti, Uy et al. 1998); ScImp $\alpha$ ) e *Mus musculus* ((Kobe 1999); MmImp $\alpha$ ), foi possível observar dois principais domínios funcionais responsáveis pela atuação da proteína no mecanismo de transporte celular via reconhecimento de NLSs: o primeiro consiste em uma curta sequência N-terminal envolvida na ligação com a Imp $\beta$  e formação do complexo de transporte (domínio IBB) (Moore and Blobel 1993; Gorlich and Mattaj 1996) e outro, formado por dez

repetições em armadilha (ARM) que formam os dois sítios de ligação para sequências NLSs (Gorlich, Prehn et al. 1994; Peifer, Pai et al. 1994), o sítio principal e secundário (Figura 4). Houve também a elucidação da estrutura da Imp $\alpha$  de outros organismos além dos citados: *Homo sapiens* (Cutress, Whitaker et al. 2008; Chang, Counago et al. 2012), *Oryza sativa* (Chang, Counago et al. 2012) e *Neurospora crassa* (Bernardes, Takeda et al. 2015).



**Figura 4** - Estrutura da Importina- $\alpha$  de *Mus musculus* (MmImp $\alpha$ ) destacando-se os dez motivos ARMs e o domínio IBB (Kobe 1999).

## 1.6 SEQUÊNCIAS DE LOCALIZAÇÃO NUCLEAR

Para a importação nuclear ocorrer a Imp $\alpha$  discrimina a proteína a ser transportada de outras proteínas celulares. Proteínas destinadas ao transporte para o núcleo contêm sequências de aminoácidos chamados de sinais ou sequências de localização nuclear (NLSs) (Lange, Mills et al. 2007).

As NLSs são sequências que por si só são suficientes para designar uma proteína a ser importada para o núcleo (Dingwall, Sharnick et al. 1982). Os sinais melhor caracterizados são denominados clássicos (cNLS), apresentando um ou mais grupos de aminoácidos básicos (Lange, Mills et al. 2007), essencialmente lisinas e argininas (Christophe, Christophe-Hobertus et al. 2000) em que duas classes podem ser destacadas. As cNLS monopartidas requerem uma lisina na posição P2, seguido de resíduos básicos nas posições P3 e P5 para produzir uma sequência de consenso KR/KXR/K (Lanford and Butel 1984; Lanford, Kanda et al. 1986; Fanara, Hodel et al. 2000; Fontes, Teh et al. 2003; Fontes, Teh et al. 2003), e

as cNLSs bipartidas, constituídas por dois conjuntos de aminoácidos básicos separados por 10-12 aminoácidos (Knudsen, Andersen et al. 2009), apresentando como consenso KRX10-12KRXK (Dingwall 1991; Fanara, Hodel et al. 2000; Fontes, Teh et al. 2003; Lange, Mills et al. 2007).

## 2 OBJETIVO

O objetivo é caracterizar e comparar a interação da MmImp $\alpha$  com sequências de localização nuclear através de diferentes técnicas experimentais e computacionais, como a calorimetria por titulação isotérmica (ITC), a espectroscopia de fluorescência e a termoforese em microescala (MST), analisando as principais diferenças entre cada uma delas. Um segundo objetivo é a criação de um algoritmo que analise de maneira mais assertiva os termogramas gerados pelo ITC, visto que há uma dificuldade para o cálculo da afinidade principalmente em sistemas mais complexos, como, por exemplo, dois sítios de ligação na proteína com diferentes afinidades, como é geralmente o caso da Imp $\alpha$ /NLSs.

## 3 MATERIAIS E MÉTODOS

### 3.1 EXPRESSÃO E PURIFICAÇÃO DE IMPORTINA- $\alpha$ TRUNCADA (70-529)

Para a expressão em larga escala da MmImp $\alpha$  truncada (70-529), utilizamos vetores pET-30a (Novagen) contendo a sequência que codifica a isoforma 2 da MmImp $\alpha$ . Esses vetores foram gentilmente cedidos pelo nosso colaborador, Prof. Kobe, da Universidade de Queensland, Austrália. As bactérias *E. coli* da linhagem BL21(DE3) pLysS foram transformadas com o vetor fornecido, e os clones selecionados.

A bactéria, armazenada como cultura permanente em freezer a  $-80\text{ }^{\circ}\text{C}$ , foi pré-inoculada em tubos tipo Falcon contendo meio LB (Luria Bertani) líquido e antibiótico canamicina ( $1\text{ }\mu\text{L/mL}$ ). O pré-inóculo foi mantido sob agitação a  $37\text{ }^{\circ}\text{C}$  durante a noite para permitir o crescimento da cultura. Em seguida, o conteúdo dos tubos *Falcon* foi transferido para *erlenmeyers* contendo meio LB líquido e o

antibiótico, até atingir uma Densidade Óptica Abs<sub>600</sub> de 0,6, medida por espectrofotômetro UV-Vis (nanodrop). Após essa etapa, o indutor IPTG (Isopropil β-d-tiogalactopiranosídeo) foi adicionado aos *erlenmeyers* e incubado por 4 horas. Ao final desse período, o conteúdo dos frascos foi centrifugado para sedimentar as células bacterianas. Os *pellets* obtidos após a centrifugação foram submetidos à lise celular para ruptura das células bacterianas e obtenção da proteína solúvel. A porção sobrenadante, onde se encontrava a proteína solúvel, foi armazenada para a etapa posterior de purificação por cromatografia líquida.

A purificação da proteína foi realizada por cromatografia de afinidade, utilizando uma coluna de níquel His-TrapHP (*GE Healthcare*), acoplada ao aparelho ÄKTA Purifier System (*GE Healthcare*). Após a injeção do extrato na coluna, realizou-se a eluição de moléculas com ligação inespecífica através da passagem de tampão contendo 15 mM de Imidazol. Posteriormente, a proteína Mmlmpα foi eluída e coletada, através do fluxo de tampão junto a um gradiente de 0 a 100 % de tampão contendo 150 mM de Imidazol. Os picos de eluição observados através do cromatograma obtido foram coletados separadamente e amostras de cada pico foram submetidas à análise por gel SDS-PAGE 12 % para verificação das frações mais puras. Em seguida, a fração mais pura foi submetida à diálise overnight para troca do tampão com Imidazol para tampão contendo 20 mM de Tris HCl pH 8,0 e 100 mM de NaCl, mantendo a estabilidade da proteína. A concentração da proteína foi realizada por centrifugação em *vivaspin* 30 kDa (coluna de rotação de membrana).

### 3.2 CALORIMETRIA POR TITULAÇÃO ISOTÉRMICA

Realizamos medidas de calorimetria por titulação isotérmica (ITC) utilizando o microcalorímetro MicroCal iTC200 (*GE Healthcare*) disponível em nosso laboratório. As amostras de Mmlmpα, a uma concentração de 30 μM, foram tituladas com peptídeos em concentrações 10 ou 20 vezes maiores em relação à proteína. A titulação do peptídeo foi conduzida no mesmo tampão da amostra de proteína para compensar o calor de diluição. No experimento, realizamos 20 injeções: a primeira de 0,4 μL e as demais de 2 μL, com intervalos de 240 s, mantendo a temperatura a 20 °C.

### 3.3 ESPECTROSCOPIA DE FLUORESCÊNCIA

As medidas de espectroscopia de fluorescência foram realizadas em um equipamento HITACHI 700 e cubeta de quartzo de caminho óptico de 10 mm. A amostra foi excitada por um comprimento de onda de 280 nm e a emissão lida de 310 a 450 nm. Foi realizada uma medida da proteína isolada e depois 19 titulações de 2  $\mu$ L de peptídeo com concentração 10 vezes maior que da proteína (m/m), com o intuito de mimetizar as proporções do experimento de ITC, a cada injeção era realizada uma nova medição.

### 3.4 TERMOFORESE EM MICROESCALA

Para os ensaios de termoforese em microescala (MST) foi utilizado o equipamento Monolith NT.115 (NanoTemper) instalado na *startup BioLinker* incubada no Centro de Inovação, Empreendedorismo e Tecnologia (CIETEC) da USP (Universidade de São Paulo). Para geração das curvas de fluorescência, todos os experimentos envolvendo titulação foram feitos em duplicata ou triplicata.

A Mmlmp $\alpha$  foi marcada de duas formas diferentes, a primeira utilizando-se o kit RED-tris-NTA (NanoTemper), para marcação não covalente de cauda dehexahistidinas, e a segunda com um corante fluorescente RED utilizando o kit Monolith NT ProteinLabeling RED - NHS (NanoTemper), para marcação de aminas. A segunda técnica de marcação necessita de um passo de remoção de todo o corante livre, para isso se utilizou uma coluna de fluxo por gravidade inclusa no kit. Foram realizados testes de fluorescência para confirmar a marcação da proteína.

Os ensaios de MST foram realizados pré-incubando por 5 min (temperatura ambiente no escuro) Mmlmp $\alpha$  a 50 nM de concentração e os ligantes com uma concentração máxima 20 vezes maior que o Kd estimado, o ligante sofreu uma série de 15 diluições usando o tampão MST 1x (NanoTemper). Para melhor ajustar a concentração do ligante e confirmar a afinidade, ensaios de apenas dois pontos foram realizados, utilizando uma alta concentração de ligante (na ordem de mM) e uma pequena concentração (em nM). As 16 misturas foram transferidas individualmente para tubos capilares e a termoforese foi medida usando uma potência de LED de 20% e uma potência do MST de 60%. O processo foi repetido para cada ligante testado. A constante de dissociação (Kd) foi determinada usando

um ajuste não linear das respostas de termoforese usando T-jump como estratégia no *software* PALMIST v.1.5.8.

### 3.5 TÉCNICAS DE BIOINFORMÁTICA PARA O CÁLCULO DE CONSTANTE DE AFINIDADE

Os estudos de *docking* molecular entre a MmImp $\alpha$  e as NLSs foram realizados utilizando a ferramenta para *Virtual Screening PyRx*, os complexos mais favoráveis escolhidos são submetidos em experimentos de simulações de dinâmica molecular, com a intenção de analisar a estabilidade, conformação após a simulação e a posição dos ligantes, com os dados obtidos do *docking* é possível estimar um valor de  $K_d$ .

Os arquivos de entrada contendo as coordenadas do complexo em uma caixa de água hexagonal são gerados utilizando o servidor CHARMM-GUI. As simulações são feitas utilizando o *software* GROMACS (*Groningen Machine for Chemical Simulation*), aplicando-se um campo de força do tipo CHARMM36. Inicialmente, é realizada uma minimização de energia (EM – energy minimization) até o sistema atingir uma energia menor que 100 kJ/mol/nm. Após este passo, são realizadas duas simulações de 1 ns de duração cada do tipo NVT e NPT (considerando-se a temperatura e pressão, respectivamente) seguidas de uma simulação sem restrições de 100 ns de duração. Os resultados obtidos são analisados por meio de ferramentas presentes no *software* GROMACS, sendo possível gerar gráficos de RMSD (*Root Mean Square Deviation* - raiz quadrada do desvio quadrático médio) dos modelos.

### 3.6 DESENVOLVIMENTO DE UM ALGORITMO PARA ANÁLISE DE DADOS DE ITC

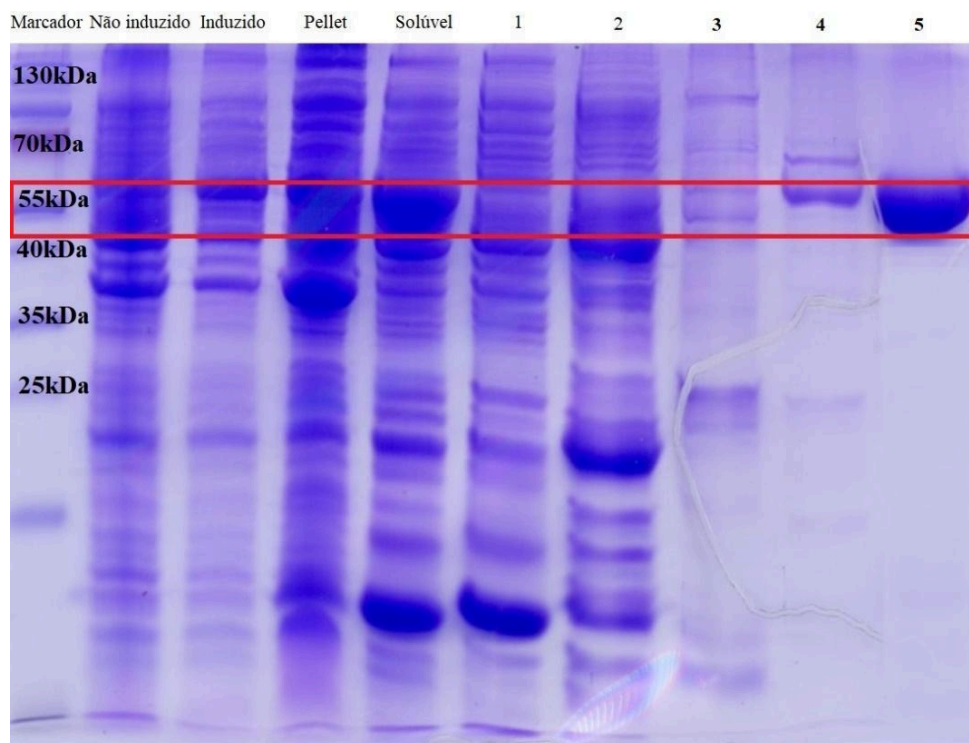
Os polinômios de ligação que descrevem a isoterma de ligação serão descritos tendo como base os modelos já descritos anteriormente (Freire, Schon et al. 2009; Vega, Abian et al. 2015), sendo que serão implementados em algoritmo de código aberto na linguagem R.

O algoritmo será validado com dados experimentais previamente analisados em um *software* comercial (Add-on Origin (*GE Healthcare*)).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 EXPRESSÃO E PURIFICAÇÃO

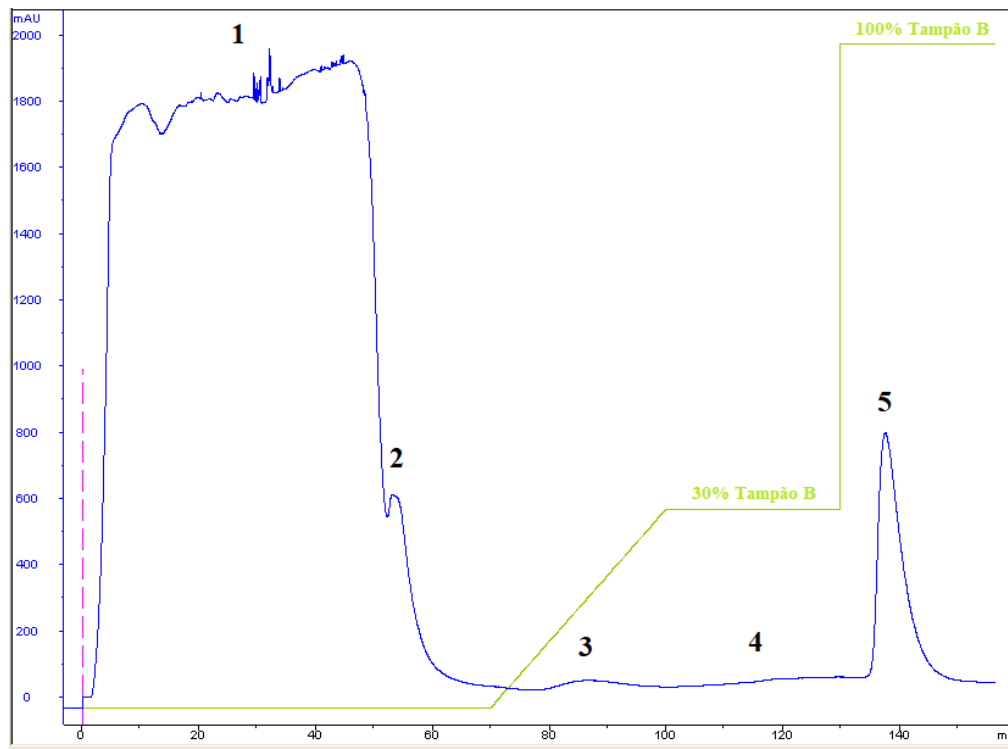
A indução da proteína recombinante foi confirmada através do gel de eletroforese SDS Page 12% (Figura 5). Além disso, após lise celular, apenas a fração solúvel do extrato total foi utilizada na purificação, para garantir que a proteína presente na amostra apresentasse integridade estrutural.



**Figura 5** - Verificação da indução e purificação da MmImp $\alpha$  em gel de poliacrilamida a 12% (SDS-PAGE), destacando a faixa da proteína recombinante, aproximadamente 55 kDa.

A fração solúvel foi purificada em uma coluna de afinidade por níquel, que separou a MmImp $\alpha$  de outras moléculas. O processo de purificação envolveu um gradiente de concentração de imidazol, resultando no cromatograma mostrado na Figura 6. Através do gel de eletroforese em SDS-PAGE a 12% (Figura 5) e do cromatograma, observamos que outras proteínas do metabolismo bacteriano não apresentaram afinidade com o níquel, como evidenciado pelo pico 1 no cromatograma e pela alíquota denominada 1 no gel. Por outro lado, as proteínas que aderiram à coluna não demonstraram afinidade significativa com a matriz cromatográfica e foram eluídas durante a lavagem com o tampão A (identificadas

como número 2 no cromatograma e no gel). As substâncias com afinidade moderada pela matriz foram eluídas com concentrações de 17% e 30% de tampão B, correspondendo aos números 3 e 4 no cromatograma e no gel, respectivamente. No caso do tampão B a 100%, a MmImp $\alpha$  foi eluída, resultando no pico 5 do cromatograma e na alíquota “5” do gel.



**Figura 6** - Cromatograma de purificação da proteína MmImp $\alpha$  por afinidade.

Em seguida foi utilizada a técnica de diálise em tampão C (Tris HCl pH 8.0 e 100 mMNaCl) para que as condições da amostra continuassem viáveis para outros experimentos e, posteriormente, realizou-se a concentração da proteína.

#### 4.2 CALORIMETRIA POR TITULAÇÃO ISOTÉRMICA

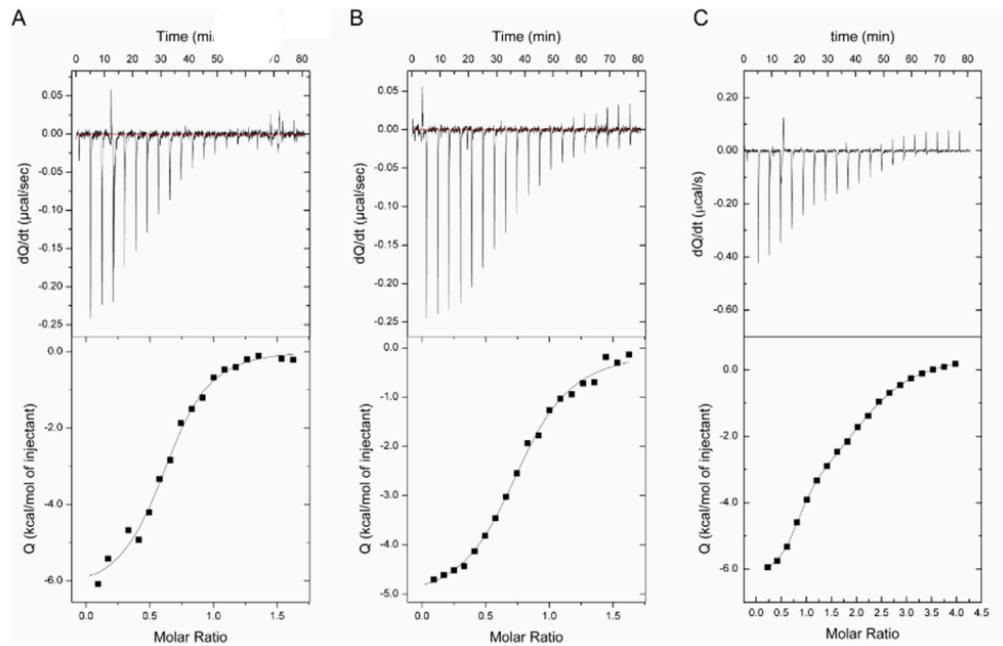
As medidas de ITC geraram os dados termodinâmicos mostrados na tabela 2.

**Tabela 2** – Dados termodinâmicos de diferentes peptídeos complexados com a Mmlmp $\alpha$ 

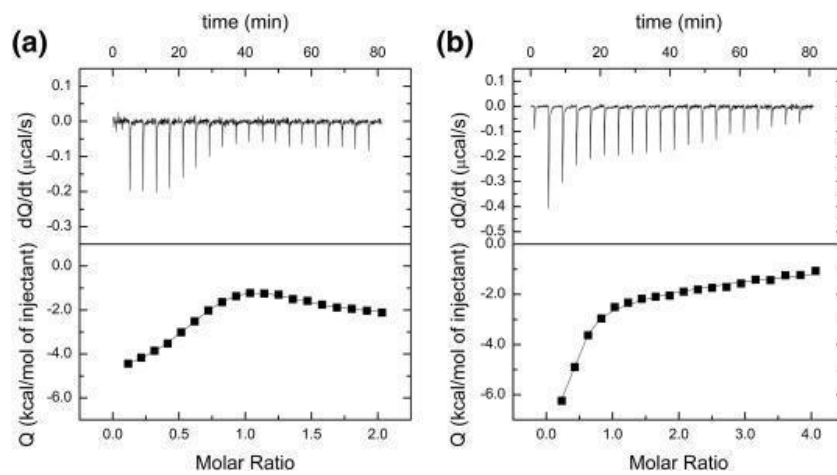
Peptídeo	Kd ( $\mu$ M)	$\Delta$ H (kcal/mol)
MLH1 (de Barros, Takeda et al. 2018)	Sítio Principal $0,10 \pm 0,02$	$-6,5 \pm 0,7$
	Sítio secundário $1,5 \pm 0,20$	$-2,9 \pm 0,7$
PMS2 (de Barros, Takeda et al. 2018)	$2,03 \pm 0,05$	$-5,78 \pm 0,1$
SV40 (de Barros, Takeda et al. 2018)	Sítio Principal $1,8 \pm 0,4$	$-9,5 \pm 0,7$
	Sítio secundário $21 \pm 1,6$	$4,1 \pm 0,4$
XPG1 (Barros, Takeda et al. 2016)	Sítio Principal $0,5 \pm 0,01$	$-3,2 \pm 0,1$
	Sítio secundário $3,4 \pm 0,7$	$3,1 \pm 0,8$
XPG2 (Barros, Takeda et al. 2016)	$1,0 \pm 0,02$	$-6,7 \pm 0,3$
MLH1-R472K (de Oliveira, da Silva et al. 2021)	$2,2 \pm 0,41$	$-5,21 \pm 0,18$
MLH1-E475A (de Oliveira, da Silva et al. 2021)	Sítio Principal $0,69 \pm 0,07$	$-11,08 \pm 0,33$
	Sítio secundário $3,76 \pm 0,04$	$-1,42 \pm 0,36$
MLH1-R470A (de Oliveira, da Silva et al. 2021)	Sítio Principal $7,7 \pm 0,54$	$-10,05 \pm 0,94$
	Sítio secundário $3,54 \pm 0,39$	$-3,26 \pm 1,04$

Os peptídeos PMS2, XPG2 e MLH1-R472K se ligaram a proteína em uma razão estequiométrica próxima a 1, ou seja, um peptídeo interagindo com uma proteína, corroborando com dados cristalográficos já publicados no grupo, assim há apenas um valor para afinidade e entalpia (Barros, Takeda et al. 2016; de Barros, Takeda et al. 2018; de Oliveira, da Silva et al. 2021). Para os demais ligantes a estequiometria foi de 2 peptídeos para uma proteína, ou seja, houve ocupação do sítio principal e secundário da Mmlmp $\alpha$  (Barros, Takeda et al. 2016; de Barros, Takeda et al. 2018; de Oliveira, da Silva et al. 2021).

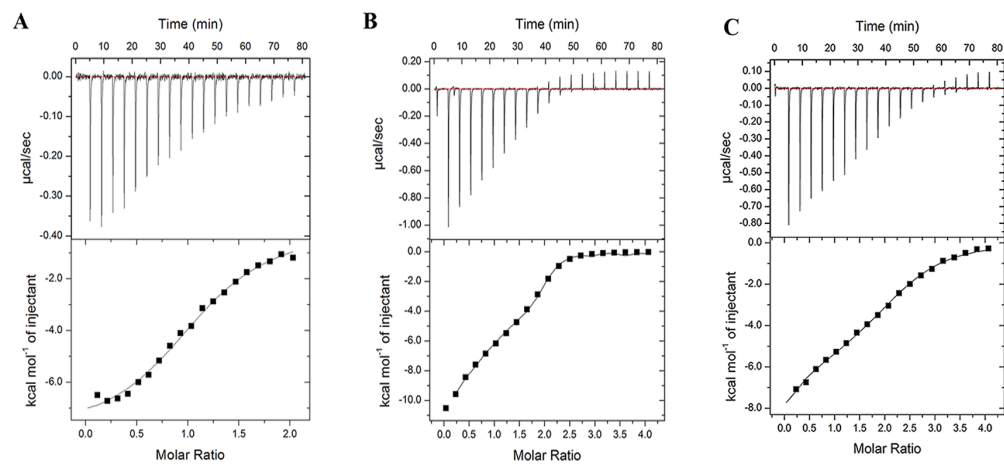
No termograma gerado (Figuras 7, 8 e 9), o painel superior é referente à titulação, enquanto o painel inferior mostra a isoterma de ligação (quadrados pretos) e o ajuste preliminar dos dados (linha contínua) realizado no *software* do fabricante.



**Figura 7** - Termogramas de ITC da Mmlmp $\alpha$  complexada com as NLS: **A.** SV40. **B.** PMS2. **C.** MLH1 (de Barros, Takeda et al. 2018)



**Figura 8** - Termogramas de ITC da Mmlmp $\alpha$  complexada com as NLS: **A.** XPG1. **B.** XPG2 (Barros, Takeda et al. 2016)

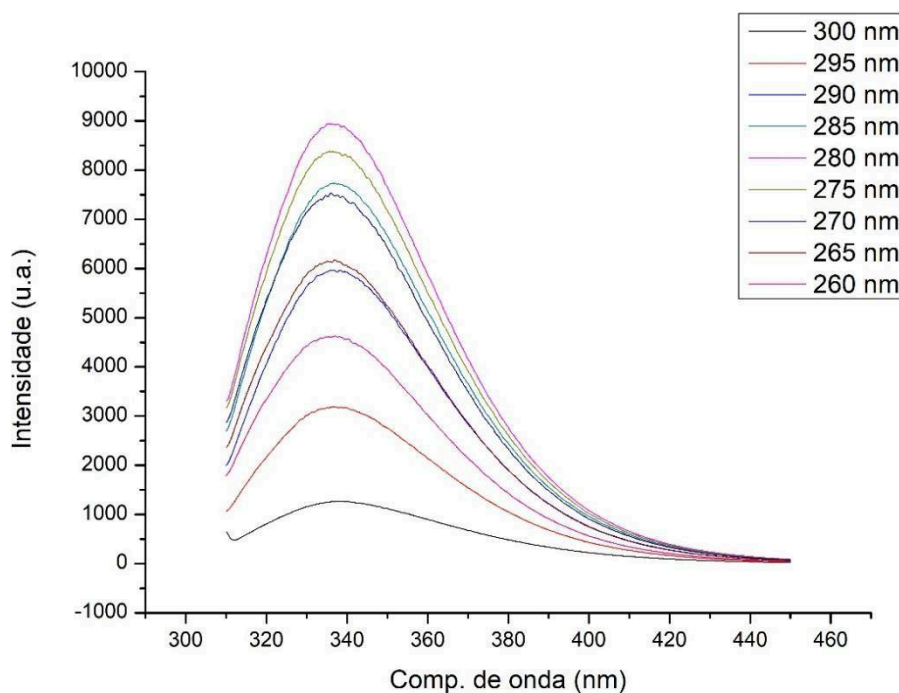


**Figura 9** - Termogramas de ITC da Mmlmp $\alpha$  complexada com as NLS: **A.** MLH1-R472K. **B.** MLH1-E475A. **C.** MLH1-R470A (de Oliveira, da Silva et al. 2021)

Esses dados serão utilizados para comparação com outras técnicas e como validação do algoritmo de análise que está sendo desenvolvido.

#### 4.3 ESPECTROSCOPIA DE FLUORESCÊNCIA

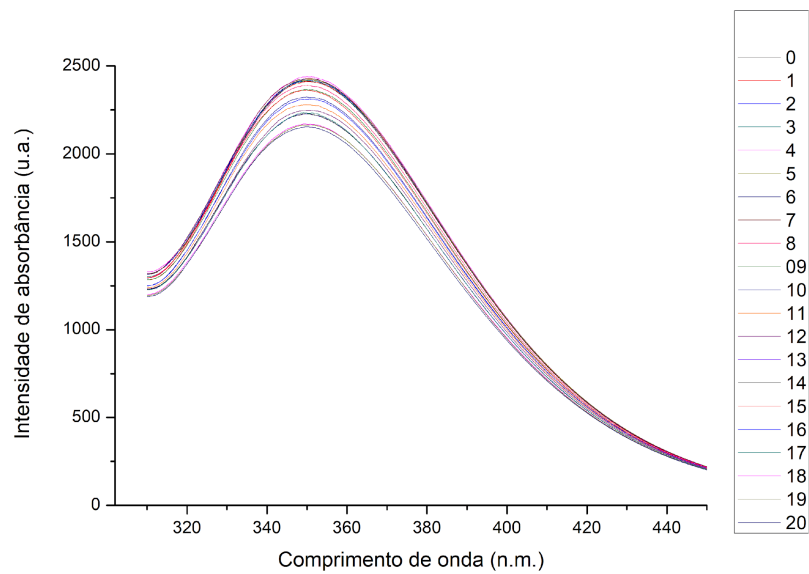
O primeiro ensaio realizado com o fluorímetro foi uma varredura entre diferentes comprimentos de onda de excitação a fim de certificar-se do comprimento ótimo de resposta. As curvas estão desenhadas na figura 10, onde cada cor representa um comprimento de onda em que a proteína isolada foi exposta.



**Figura 10** - Gráfico de emissão de fluorescência em diferentes comprimentos de onda de excitação da Mmlmp $\alpha$  a 30  $\mu$ M.

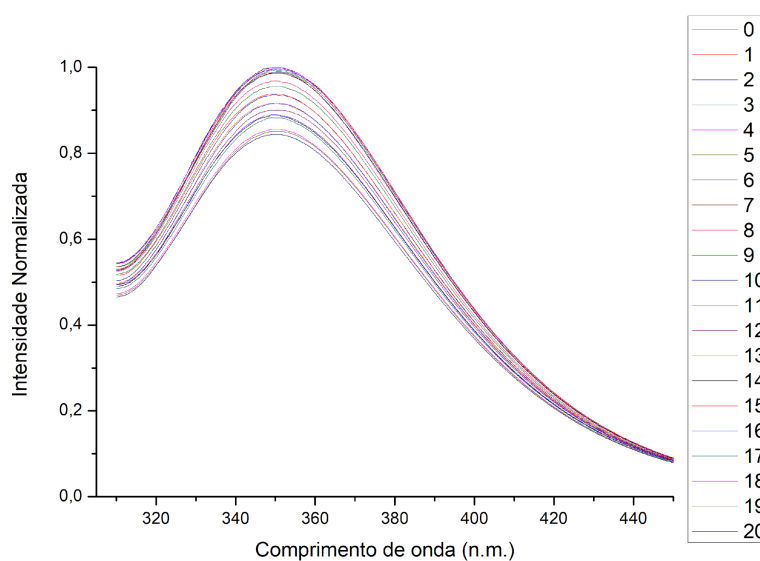
Como era esperado, a máxima emissão se dá quando a proteína é excitada a 280 nm, que é o comprimento de onda máximo de absorção do triptofano. Além de ser o aminoácido com maior rendimento quântico, o triptofano está presente na proteína principalmente nos sítios de ligação, o que para o caso do projeto é bem interessante, uma vez que os experimentos de fluorescência serão utilizados para aferir afinidade entre a Mmlmp $\alpha$  e diferentes peptídeos. Tendo isso em vista, todas as medidas com ligantes foram realizadas excitando o complexo a 280 nm.

Para a utilização da espectroscopia de fluorescência para medidas quantitativas da afinidade entre proteína e ligante não há um protocolo estabelecido, no laboratório utiliza-se a técnica, de maneira mais ampla, para a confirmação qualitativa da ligação. O protocolo escolhido realiza diversas medidas seguido de injeções de ligante, as curvas de absorbância geradas estão expressas na figura 11.



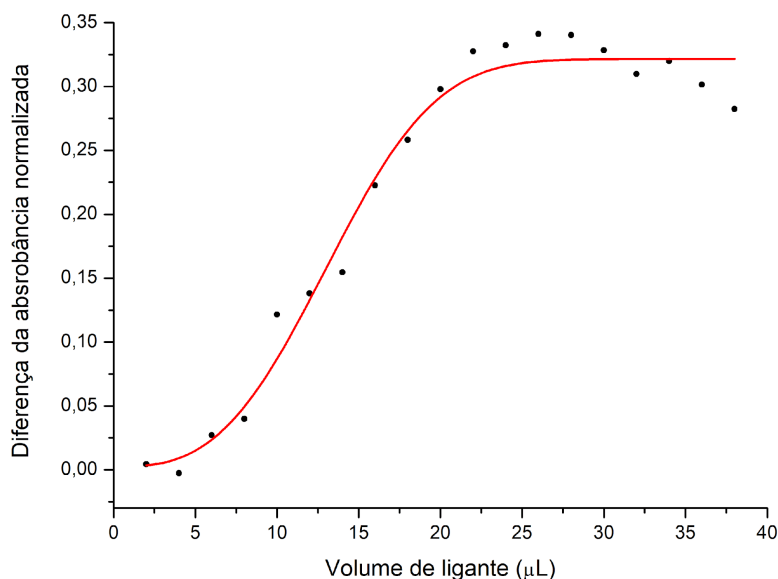
**Figura 11** - Curvas da intensidade de absorvância por comprimento de onda. Cada cor representa a Mmlp $\alpha$  depois de n injeções do peptídeo da MLH1, sendo o 0 a amostra isolada e as injeções indo de 1 a 20.

As curvas expressas na figura 11 precisam ser normalizadas pela concentração da proteína a cada medida, uma vez que a injeção do ligante acarreta também uma diluição da amostra, evitando assim que essa redução no sinal causada pela diluição dê um falso sinal de ligação. As curvas normalizadas estão expressas na figura 12.



**Figura 12** - Curvas da intensidade de absorvância normalizada por comprimento de onda. Cada cor representa a Mmlmpa depois de n injeções do peptídeo da MLH1, sendo o 0 a amostra isolada e as injeções indo de 1 a 20.

A partir dos dados normalizados e utilizando a diferença entre os máximos da curva da amostra isolada e das demais injeções, obtêm-se os pontos do gráfico da figura 13. Os dados devem seguir um comportamento de uma curva sigmóide, uma vez que a disposição de fluoróforos é maior no início da reação, onde temos muitos sítios não ocupados, ou seja, uma diferença muito pequena entre essa medida e a amostra sem ligante, essa diferença tende a aumentar com a diminuição dos triptofanos do sítio de ligação disponíveis ao solvente. No *software* Origin é feito o ajuste da curva com uma função sigmóide de ponto de inflexão livre, os pontos e a curva ajustada estão na figura 13.



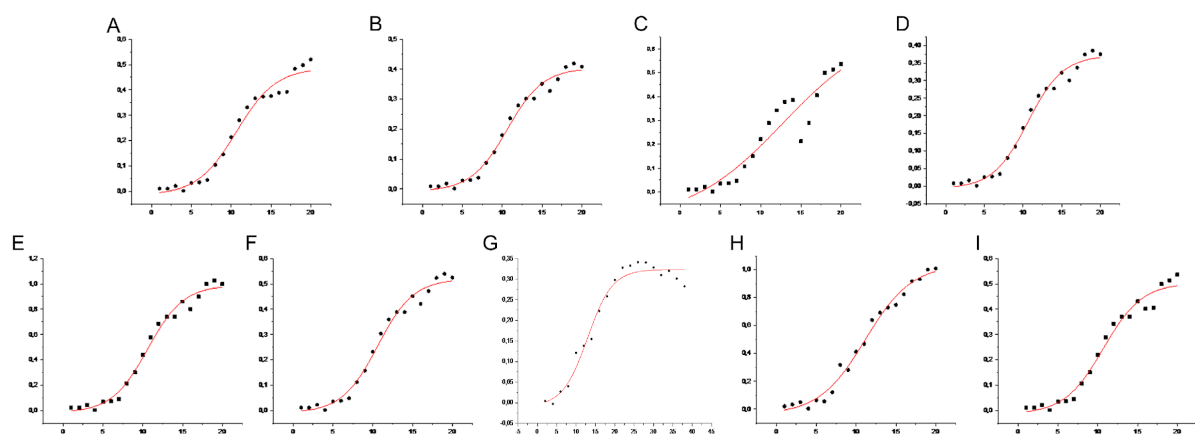
**Figura 13** - Dados obtidos a partir da diferença entre os máximos das curvas de injeção de MLH1 NLS na Mmlmp $\alpha$  mostradas na figura 12, e a curva vermelha ajustada por uma função sigmóide.

A informação da constante de dissociação pode ser obtida através da inclinação da reta tangente do ponto de inflexão. Os dados de ajuste da curva já dão informações como o ponto de inflexão e a adequação da equação com a curva mínima de 95%. Traçando uma reta tangente ao ponto de inflexão obtêm-se a declividade da curva, que expressa à constante de afinidade, que é o inverso da constante de dissociação ( $K_d$ ). Esse mesmo procedimento foi realizado em duplicata para todos os ligantes e as informações de  $K_d$  estão expressos na tabela 3 e na figura 14 para os demais ligantes.

**Tabela 3** – Dados de  $K_d$  gerados por espectroscopia de fluorescência de diferentes peptídeos complexados com a Mmlmp $\alpha$

Peptídeo	$K_d$ ( $\mu\text{M}$ )
MLH1	$0,80 \pm 0,09$
PMS2	$1,60 \pm 0,41$

SV40	$7,98 \pm 0,92$
XPG1	$2,30 \pm 1,17$
XPG2	$5,40 \pm 0,02$
MLH1-R472K	$1,10 \pm 0,41$
MLH1-E475A	$2,97 \pm 0,70$
MLH1-R470A	$3,40 \pm 0,56$
MLH1-S467A	$0,30 \pm 0,09$
FEN1	$4,20 \pm 1,20$

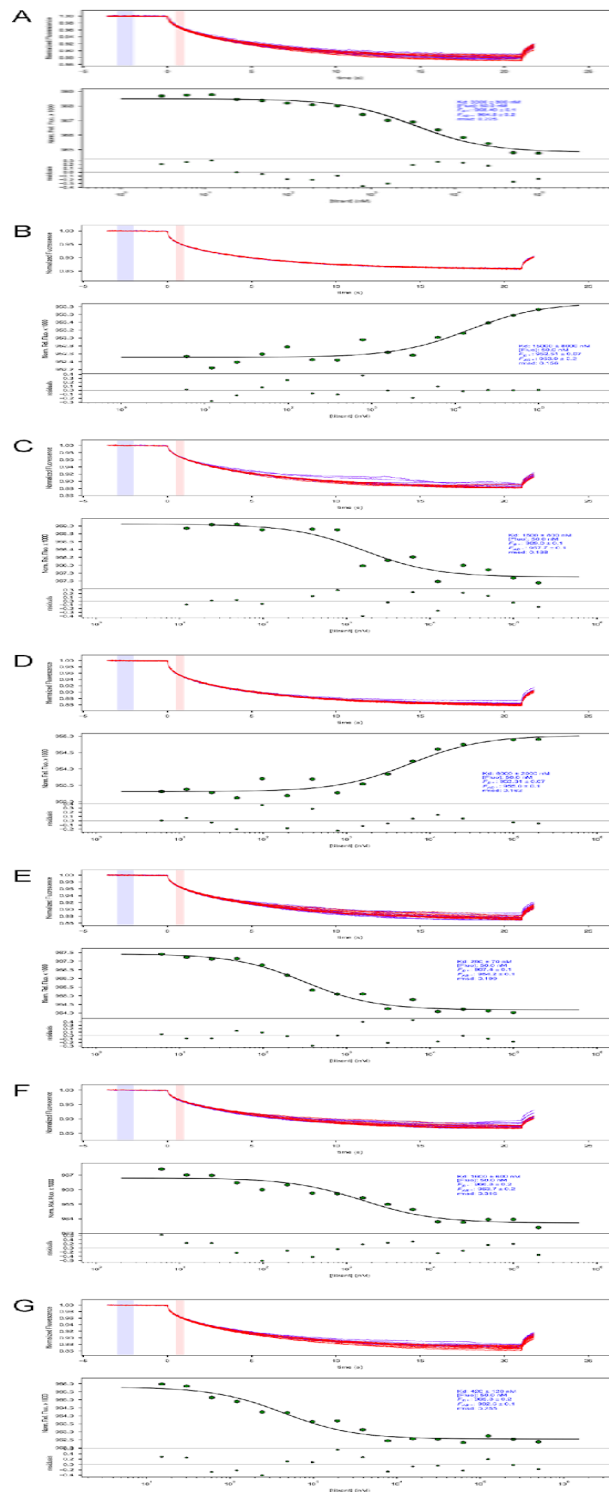


**Figura 14** – Dados obtidos a partir da diferença entre os máximos das curvas de injeção e a curva vermelha ajustada por uma função sigmóide da  $MmImp\alpha$  com os ligantes. **A.** PMS2.

**B.** SV40. **C.** XPG1. **D.** XPG2. **E.** MLH1-R472K. **F.** MLH1-E475A. **G.** MLH1-R470A. **H.** MLH1-S467A. **I.** FEN1.

#### 4.4 TERMOFORESE EM MICROESCALA

Os experimentos de MST foram realizados com os seguintes ligantes: XPG1, SV40, MLH1, MLH1-S467A e FEN1 (figura 15), todos demonstraram afinidade pela  $MmImp\alpha$ . Na tabela 4 consta os valores de constante de dissociação. Na figura 15 é possível observar três gráficos, o primeiro gráfico sendo o sinal típico do experimento MST em cada capilar, onde, inicialmente, as moléculas são distribuídas homogeneamente e uma fluorescência inicial constante é detectada, após a ativação do laser infravermelho, observa-se o “T-Jump”, que corresponde a uma rápida mudança nas propriedades do fluoróforo devido à rápida mudança de temperatura, assim o movimento termoforético das moléculas marcadas com fluorescência pode ser detectado e após a desativação do laser, ocorre um T-Jump inverso. No segundo gráfico, a mudança na termoforese é expressa como a mudança na fluorescência normalizada para produzir uma curva de ligação, que pode ser ajustada com uma curva que fornece a constante de dissociação. O terceiro gráfico expressa o erro de cada ponto em relação a curva ajustada.



**Figura 15** – Dados de MST gerados com a Mmlmp $\alpha$  complexada aos ligantes. **A.** MLH1. **B.** MLH1<sup>a</sup>. **C.** SV40. **D.** XPG1<sup>a</sup>. **E.** XPG1. **F.** MLH1-S467A. **G.** FEN1. <sup>a</sup>Medida realizada com a proteína marcada pela cauda de histidina.

Os experimentos com os complexos Mmlmp $\alpha$ -XPG1NLS e Mmlmp $\alpha$ -MLH1NLS também foram realizados com a proteína marcada pela cauda de histidina. Em ambos os casos, o Kd estimado foi maior que com a proteína marcada pelas aminas.

**Tabela 4** – Dados de Kd gerados por MST de diferentes peptídeos complexados com a Mmlmp $\alpha$

Peptídeo	Kd ( $\mu$ M)
MLH1	3,3 $\pm$ 0,8
MLH1 <sup>a</sup>	15,0 $\pm$ 8,0
SV40	1,6 $\pm$ 0,8
XPG <sup>a</sup>	6,0 $\pm$ 2,0
XPG1	0,29 $\pm$ 0,07
MLH1-S467A	1,6 $\pm$ 0,6
FEN1	0,42 $\pm$ 0,12

<sup>a</sup> Medida realizada com a proteína marcada pela cauda de histidina.

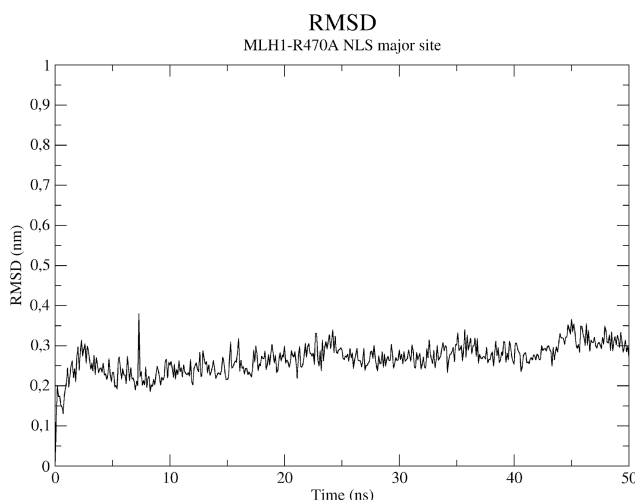
#### 4.5 TÉCNICAS DE BIOINFORMÁTICA PARA O CÁLCULO DE CONSTANTE DE AFINIDADE

Os ensaios de *docking* molecular foram realizados e 20 soluções foram geradas. Analisando cada complexo gerado e selecionando primeiramente os com o ligante antiparalelo em relação à proteína, escolheu uma solução para cada sítio, quando houver, com o critério de similaridade com as posições dos aminoácidos consenso já estudados na literatura. A conformação antiparalela de ligação da Mmlmp $\alpha$  com a NLS garante a formação de contatos favoráveis com os resíduos da proteína conservados, assim, todos os complexos depositados no banco de dados do PDB têm o ligante antiparalelo a hélice da proteína. O PyRx retorna o valor da energia de ligação do complexo ( $\Delta G$ ), com esse dado é possível calcular o Kd, através da equação:

$$\Delta G = RT \ln(K_d)$$

Sendo T a temperatura em Kelvin (298,15 K) e R a constante universal dos gases ( $R=1,9872 \cdot 10^{-3}$  kcal/K.mol). Como o valor de  $\Delta G$  obtido pelo *software* tem apenas dois algarismos significativos (por exemplo: 6,7 kcal/mol) há uma faixa de valores de Kd para cada  $\Delta G$ , assim há uma imprecisão quanto ao valor exato, porém a unidade de grandeza sempre é preservada. Os valores estão na tabela 5.

Para certificarmos da estabilidade do peptídeo no complexo, uma dinâmica molecular de 50 ns foi realizada com o programa GROMACS. O RMSD do peptídeo ao longo do tempo foi esboçado em um gráfico como ilustrado na figura 16. Para todos os peptídeos testados o RMSD teve variações menores que 1 nm, o que reafirma a estabilidade da ligação.



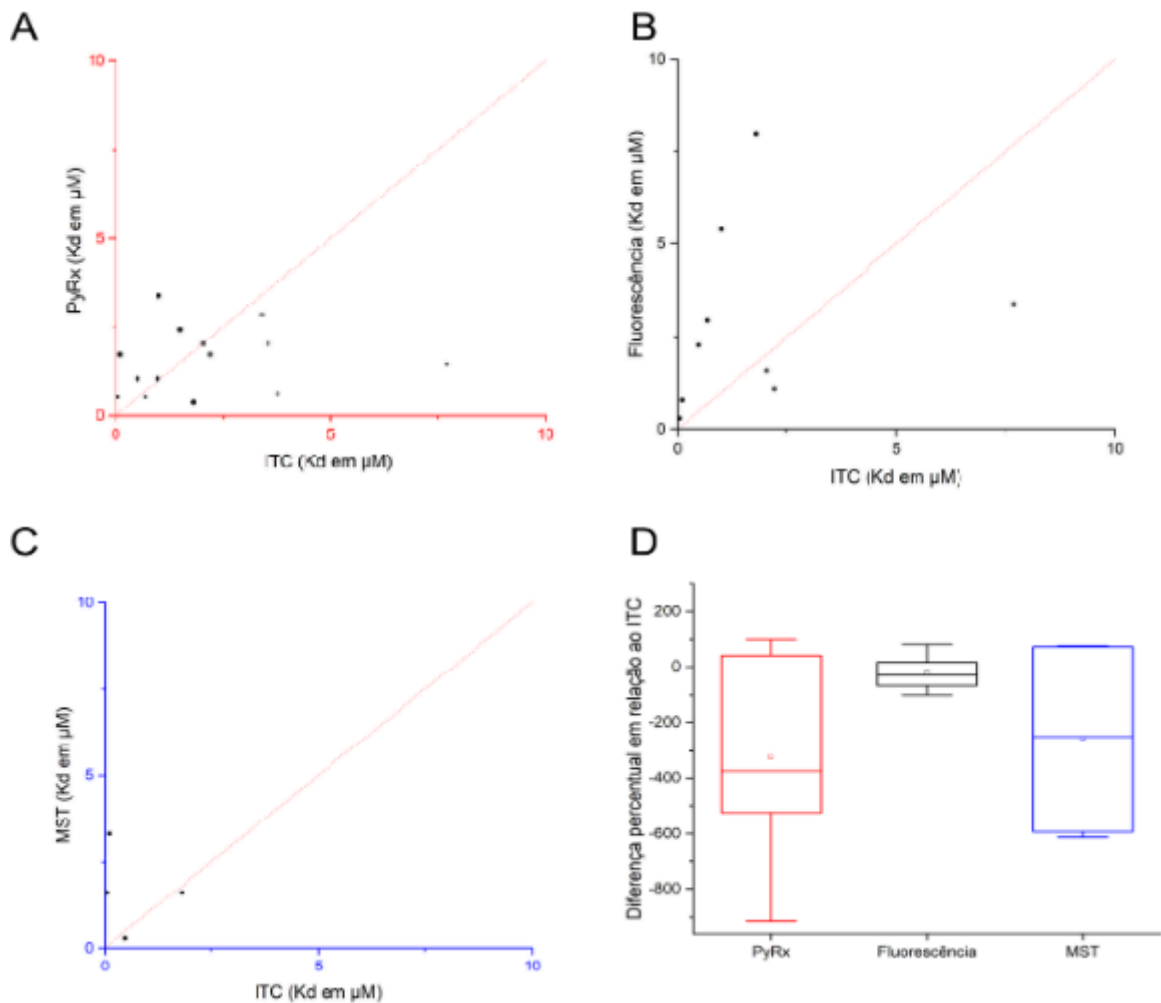
**Figura 16** - Gráfico de RMSD do peptídeo da MLH1 ligado no sítio principal da Mmlmpa gerado por simulação no GROMACS.

#### 4.6 ANÁLISE COMPARATIVA ENTRE AS TÉCNICAS

A tabela 5 apresenta os valores de constante de dissociação dos complexos obtidos por diferentes técnicas. A figura 17 expressa graficamente às diferenças entre os valores em relação ao ITC.

**Tabela 5** - Constantes de dissociação de diferentes peptídeos complexados com a proteína Mmlmpa obtidas por diferentes técnicas

Peptídeo	ITC	Docking		Fluorescência	MST
	Kd ( $\mu\text{M}$ )	$\Delta\text{G}$ (kcal/mol)	Kd ( $\mu\text{M}$ )	Kd ( $\mu\text{M}$ )	Kd ( $\mu\text{M}$ )
MLH1	Sítio Principal $0,10 \pm 0,02$	-6,5	0,17	$0,8 \pm 0,09$	$3,3 \pm 0,8$
	Sítio secundário $1,5 \pm 0,20$	-6,3	0,24		
PMS2	$2,03 \pm 0,05$	-6,4	0,20	$1,6 \pm 0,41$	-
SV40	Sítio Principal $1,8 \pm 0,4$	-7,4	3,80	$7,98 \pm 0,92$	$1,6 \pm 0,8$
	Sítio secundário $21 \pm 1,6$	-7,1	6,20		
XPG1	Sítio Principal $0,5 \pm 0,01$	-6,8	0,10	$2,3 \pm 1,17$	$0,29 \pm 0,07$
	Sítio secundário $3,4 \pm 0,7$	-6,2	0,29		
XPG2	Sítio Principal $1,0 \pm 0,02$	-6,1	0,38	$1,4 \pm 0,02$	-
	Sítio secundário -	-7,1	6,24		
MLH1-R472K	$2,2 \pm 0,41$	-6,5	0,17	$1,1 \pm 0,41$	-
MLH1-R475A	Sítio Principal $0,69 \pm 0,07$	-7,2	5,27	$2,97 \pm 0,70$	-
	Sítio secundário $3,76 \pm 0,04$	-7,1	6,24		
MLH1-R470A	Sítio Principal $7,7 \pm 0,54$	-6,6	0,15	$3,4 \pm 0,56$	-
	Sítio secundário $3,54 \pm 0,39$	-6,4	0,20		
MLH1-S467A	Sítio Principal $0,04 \pm 0,01$	-7,2	5,27	$0,3 \pm 0,09$	$1,6 \pm 0,6$
	Sítio secundário $0,97 \pm 0,04$	-6,8	0,10		
FEN1	-	-7,9	1,62	$4,2 \pm 1,2$	$0,42 \pm 0,12$



**Figura 17** – **A.** Gráfico de dispersão entre os valores de Kd obtidos via espectroscopia de fluorescência e ITC. **B.** Gráfico de dispersão entre os valores de Kd obtidos via MST e ITC. **C.** Gráfico de dispersão entre os valores de Kd obtidos via docagem molecular (PyRx) e ITC. **D.** Boxplot de distribuição das porcentagens de variação entre os valores obtidos em cada técnica e o ITC, variações acima de 1000% entre PyRx e ITC foram considerados outliers e desconsiderados na construção gráfica.

Diversos autores já relataram a importância da escolha correta para a melhor técnica de medição de afinidade, considerando as vantagens e limitações particulares, principalmente pelo avanço no desenvolvimento de fármacos com alvos proteicos específicos (Watzig, Oltmann-Norden et al. 2015; Kairys, Baranauskiene et al. 2019). Porém, estudos que abrangem diversas combinações de interações entre macromoléculas se deparam com o desafio de não encaixar todos os ensaios necessários em uma única técnica, então a comparação entre diferentes ferramentas se faz necessária. Nessa comparação, nota-se que apesar das técnicas

utilizarem diferentes princípios físicos, não houve variações na unidade de grandeza das afinidades mensuradas com a proteína estudada. Os valores obtidos por *docking* molecular são os únicos dados nesse estudo que não experimentais e sim teóricos, assim, na maioria dos casos, esse valor de afinidade é superestimado (Figura 17 C) uma vez que não considera fatores como concentração e agregação da amostra. Considerando os valores de ITC como referência para a análise comparativa (Figura 17 D), podemos observar que todas as técnicas estimaram valores próximos para a constante de afinidade. Como a fluorescência e a termoforese em microescala (MST) não diferenciam as constantes para cada sítio de ligação, nos casos de estequiometria de dois ligantes para uma proteína, o valor encontrado representa a afinidade média de ambos os sítios. Nota-se, porém, uma maior semelhança entre os dados de ITC e fluorescência (Figura 17 A e D).

#### 4.7 DESENVOLVIMENTO DE UM ALGORITMO PARA ANÁLISE DE DADOS DE ITC

Para o desenvolvimento do algoritmo de análise dos dados de ITC (calorimetria por titulação isotérmica), o primeiro passo foi desenhar o gráfico de pontos a partir da conversão do arquivo .csv gerado pelo *Add-on* do *software* Origin em um *dataframe*, e desenhado através de um pacote de gráfico do R (ggplot2), também é possível construir o gráfico de calor das injeções. Pela sintaxe do R os comentários são seguidos do símbolo #, ou seja, tudo após esse símbolo na linha não será executado. O código para a execução dos gráficos, bem como o gráfico gerado são mostrados abaixo e na figura 18, respectivamente. Para que os gráficos fiquem em uma mesma página, foram utilizados os comandos de grid.

```

#Pacote para construção de gráficos
library(ggplot2)
require(gridExtra)

#dataframes carregados: itc_graf1 e itc_graf2

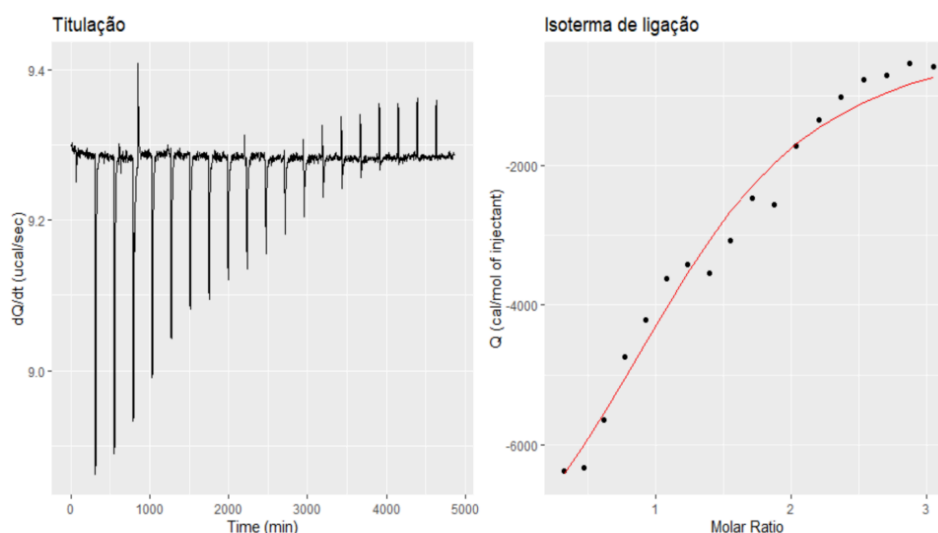
#Remover o primeiro dado de injeção
itc_graf2 <- itc_graf2[2:20,]

#Gráfico 1
plot1 <- ggplot(itc_graf1, aes(x = ...1, y = ...2)) +
  geom_line() +
  labs(x = 'Time (min)', y = 'dQ/dt (ucal/sec)')

#Gráfico 2
plot2 <- ggplot(itc_graf2, aes(x = XMt)) +
  geom_point(aes(y = NDH)) +
  geom_line(aes(y = Fit), color = "red") +
  labs(x = 'Molar Ratio', y = 'Q (cal/mol of injectant)')

#Organizar a exibição dos gráficos
grid.arrange(plot1, plot2, ncol = 2)

```



**Figura 18** - Gráficos de ITC gerados no R

Para o ajuste das curvas e o correto processamento de cada tipo de isoterma de ligação, é recomendável seguir os seguintes passos: enumerar todos os estados de cada composto presente na titulação e construir o polinômio de ligação; calcular as concentrações das macromoléculas e dos ligantes na célula de reação, durante a

titulação; solucionar as equações de ligação, baseadas em polinômios de ligação, que são modelos capazes de identificar todas as espécies presentes no decorrer da titulação, e calcular as concentrações de todos complexos após cada injeção; cálculo do calor normalizado pela concentração da espécie titulante em cada injeção da titulação; estimar os parâmetros termodinâmicos da ligação através de análise de regressão não linear. Para isso os seguintes valores de entrada foram necessários:

- ✓ Mt: Concentração da célula de amostra
- ✓ Lt: Concentração da seringa de amostra
- ✓ Vo: Volume da célula de reação
- ✓ Vi: Volume das injeções
- ✓ Ni: número de injeções
- ✓ Lz: Concentração de ligante na célula

O usuário não necessita mais inserir os dados de execução do experimento manualmente, apenas um aviso é gerado para conferência das informações extraídas do arquivo de texto com os dados

A equação de diluição da amostra na célula de amostra e depósito de ligante a cada injeção e a isoterma de Wiseman já foram desenhados no R como mostradas na caixa de código abaixo.

```
for (i in 1:ni){
  Li<- Lt*(Lz+Vi*(i-1))/(Vo+(i*Vi)) #Concentração do ligante na célula a cada injeção
  print(Li)
  Mi<-Mt*(Vo+Vi*(i-1))/(Vo+Vi*i) #Concentração de proteína na célula
  print(Mi)
}
```

```

#Isoterma de Wiseman

#Concentração da célula de amostra (usuário dependente)
Mt<-0.03
#Concentração da seringa de amostra (usuário dependente)
Lt<-0.3
#Volume da célula de reação (usuário dependente)
Vo<-200
#Para cada linha de dados eu tenho um conjunto de pontos (Dh é o delta H)
Dh<-itc_graf2$DH

#Quantidade de calor devido a formação do complexo (Qt):
Qt<-Vo*Dh*Mt*((1+(1/ka*Mt)+(Lt/Mt)+sqrt((1+(1/ka*Mt)+(Lt/Mt))^2-(4*Lt/Mt)))/2)

```

Os valores de saída são: estequiometria, delta H e as constantes de dissociação. As constantes de dissociação de alguns ligantes são mostrados na tabela 6. Para validação do modelo foi analisado a variação percentual entre o dado do algoritmo e do software do fabricante, considerando o valor sem os erros associados, e a maioria dos dados teve variação inferior a 10%.

**Tabela 6** - Constantes de dissociação calculadas pelo software do fabricante e pelo algoritmo

	Software do fabricante	Algoritmo desenvolvido	Diferença
Peptídeo	Kd (µM)	Kd (µM)	(%)
MLH1	Sítio Principal 0,10 ± 0,02	0,098 ± 0,002	2,0
	Sítio secundário 1,5 ± 0,20	1,398 ± 0,121	6,8
PMS2	2,03 ± 0,05	2,032 ± 0,038	- 0,1
SV40	Sítio Principal 1,8 ± 0,4	1,675 ± 0,431	6,9
	Sítio secundário 21 ± 1,6	18,124 ± 0,870	13,7
XPG1	Sítio Principal 0,5 ± 0,01	0,531 ± 0,087	- 6,2
	Sítio secundário 3,4 ± 0,7	3,217 ± 0,464	5,4
XPG2	1,0 ± 0,02	0,986 ± 0,008	1,4
MLH1-R472 K	2,2 ± 0,41	2,192 ± 0,764	0,4
MLH1-R475 A	Sítio Principal 0,69 ± 0,07	0,732 ± 0,009	- 6,1
	Sítio secundário 3,76 ± 0,04	3,987 ± 0,054	- 6,0
MLH1-R470 A	Sítio Principal 7,7 ± 0,54	7,546 ± 0,003	2,0
	Sítio secundário 3,54 ± 0,39	3,688 ± 0,490	- 4,2
MLH1-S467 A	Sítio Principal 0,04 ± 0,01	0,052 ± 0,002	- 30,0
	Sítio secundário 0,97 ± 0,04	1,003 ± 0,087	- 3,4

A implementação do algoritmo permitirá uma maior exatidão e menor dependência do usuário na análise de dados calorimétricos de casos com mais de um sítio de ligação e/ou cooperação entre interações, trazendo assim uma menor complexidade na obtenção de dados confiáveis e com ajustes satisfatórios.

## 5 CONCLUSÃO

As três técnicas experimentais e a computacional escolhidas, apesar de bem distintas, obtiveram valores de constante de dissociação com a mesma ordem de grandeza para a interação entre a  $Imp\alpha$  e as NLSs, o que reforça a tese inicial de que esses valores podem ser cuidadosamente comparados. Nessa comparação, nota-se que apesar das técnicas utilizarem diferentes princípios físicos, não houve variações na unidade de grandeza das afinidades mensuradas com a proteína estudada. Os valores obtidos por *docking* molecular são superestimados. Como as outras técnicas experimentais não diferenciam as constantes para cada sítio de ligação, o valor encontrado representa uma afinidade média de ambos os sítios.

Os dados apresentados neste trabalho podem ser muito importantes para a comparação entre diferentes técnicas obtidas com um determinado sistema de proteínas. A literatura mostra (Salvador, Pinto et al. 2023) que por dificuldades experimentais inerentes a cada técnica (por exemplo a dificuldade de pareamento das condições de tampão e pH (*mismatching*) entre proteína e ligante para ITC), não é possível fazer a triagem de ligantes com a mesma técnica. Similarmente, resultados de interação obtidos por diferentes autores com diferentes técnicas, tem sua comparação colocada em dúvida. Neste trabalho, pela primeira vez, este tema é abordado de maneira sistemática com quatro técnicas diferentes.

O algoritmo se mostra promissor, afinal os dados calculados estão bem satisfatórios e próximos dos dados previamente calculados pelo *software* do fabricante, o que trará um avanço para a facilitação da análise de dados complexos e permitirá uma maior exatidão com menor dependência do usuário.

A calorimetria por titulação isotérmica é considerada por muitos autores uma das técnicas mais robustas para estudar afinidade entre duas moléculas, pois permite obter a constante de afinidade em mais de um sítio de interação, mesmo com interações diferentes, bem como dados que incluem estequiometria, entalpia,

entropia e energia livre de Gibbs. Porém, diversos usuários relatam que o processamento de dados usando esta técnica pode ser complexo, principalmente no caso de múltiplos sítios de interação. No presente trabalho foi desenvolvido um novo algoritmo que tenta melhorar a interface e os resultados obtidos por um usuário regular, permitindo obter dados mais exatos e uma melhor interface usuário-*software*. Desta forma, o uso da técnica pode se tornar bem mais corriqueiro na busca de novos fármacos, por exemplo, em empresas de biotecnologia, onde o tempo pode ser bem escasso na busca por um número grande de compostos.

## REFERÊNCIAS

- Baranauskiene, L., V. Petrikaite, et al. (2009). "Titration calorimetry standards and the precision of isothermal titration calorimetry data." *Int J Mol Sci* **10**(6): 2752-2762.
- Barros, A. C., A. A. Takeda, et al. (2016). "Structural and Calorimetric Studies Demonstrate that Xeroderma Pigmentosum Type G (XPG) Can Be Imported to the Nucleus by a Classical Nuclear Import Pathway via a Monopartite NLS Sequence." *J Mol Biol* **428**(10 Pt A): 2120-2131.
- Bernardes, N. E., A. A. Takeda, et al. (2015). "Structure of Importin-alpha from a Filamentous Fungus in Complex with a Classical Nuclear Localization Signal." *PLoS One* **10**(6): e0128687.
- Biswas, T. and O. V. Tsodikov (2010). "An easy-to-use tool for planning and modeling a calorimetric titration." *Anal Biochem* **406**(1): 91-93.
- Chang, C. W., R. L. Counago, et al. (2012). "Crystal structure of rice importin-alpha and structural basis of its interaction with plant-specific nuclear localization signals." *Plant Cell* **24**(12): 5074-5088.
- Chook, Y. M. and G. Blobel (2001). "Karyopherins and nuclear import." *Curr Opin Struct Biol* **11**(6): 703-715.
- Christophe, D., C. Christophe-Hobertus, et al. (2000). "Nuclear targeting of proteins: how many different signals?" *Cell Signal* **12**(5): 337-341.
- Cingolani, G., J. Bednenko, et al. (2002). "Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta." *Mol Cell* **10**(6): 1345-1353.
- Conti, E. and J. Kuriyan (2000). "Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha." *Structure* **8**(3): 329-338.
- Conti, E., M. Uy, et al. (1998). "Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha." *Cell* **94**(2): 193-204.
- Cutress, M. L., H. C. Whitaker, et al. (2008). "Structural basis for the nuclear import of the human androgen receptor." *J Cell Sci* **121**(Pt 7): 957-968.
- de Barros, A. C., A. A. S. Takeda, et al. (2018). "DNA mismatch repair proteins MLH1 and PMS2 can be imported to the nucleus by a classical nuclear import pathway." *Biochimie* **146**: 87-96.
- de Oliveira, H. C., T. D. da Silva, et al. (2021). "Structural and calorimetric studies reveal specific determinants for the binding of a high-affinity NLS to mammalian importin-alpha." *Biochem J* **478**(13): 2715-2732.
- Dingwall, C. (1991). "Transport across the nuclear envelope: enigmas and explanations." *Bioessays* **13**(5): 213-218.
- Dingwall, C., S. V. Sharnick, et al. (1982). "A polypeptide domain that specifies migration of nucleoplasmin into the nucleus." *Cell* **30**(2): 449-458.
- Doyle, M. L. (1997). "Characterization of binding interactions by isothermal titration calorimetry." *Curr Opin Biotechnol* **8**(1): 31-35.
- Fanara, P., M. R. Hodel, et al. (2000). "Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding." *J Biol Chem* **275**(28): 21218-21223.
- Fontes, M. R., T. Teh, et al. (2003). "Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha." *J Biol Chem* **278**(30): 27981-27987.
- Fontes, M. R., T. Teh, et al. (2000). "Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha." *J Mol Biol* **297**(5): 1183-1194.
- Fontes, M. R., T. Teh, et al. (2003). "Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-alpha." *Biochem J* **375**(Pt 2): 339-349.

- Freire, E., A. Schon, et al. (2009). "Isothermal titration calorimetry: general formalism using binding polynomials." *Methods Enzymol* **455**: 127-155.
- Freire, E., W. W. van Osdol, et al. (1990). "Calorimetrically determined dynamics of complex unfolding transitions in proteins." *Annu Rev Biophys Biophys Chem* **19**: 159-188.
- Freyer, M. W. and E. A. Lewis (2008). "Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions." *Methods Cell Biol* **84**: 79-113.
- Gorlich, D. and U. Kutay (1999). "Transport between the cell nucleus and the cytoplasm." *Annu Rev Cell Dev Biol* **15**: 607-660.
- Gorlich, D. and I. W. Mattaj (1996). "Nucleocytoplasmic transport." *Science* **271**(5255): 1513-1518.
- Gorlich, D., S. Prehn, et al. (1994). "Isolation of a protein that is essential for the first step of nuclear protein import." *Cell* **79**(5): 767-778.
- Gorlich, D., F. Vogel, et al. (1995). "Distinct functions for the two importin subunits in nuclear protein import." *Nature* **377**(6546): 246-248.
- Gruner, S., M. Neeb, et al. (2014). "Impact of protein and ligand impurities on ITC-derived protein-ligand thermodynamics." *Biochim Biophys Acta* **1840**(9): 2843-2850.
- Holdgate, G. A., M. Anderson, et al. (2010). "Affinity-based, biophysical methods to detect and analyze ligand binding to recombinant proteins: matching high information content with high throughput." *J Struct Biol* **172**(1): 142-157.
- Huber, W. and F. Mueller (2006). "Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology." *Curr Pharm Des* **12**(31): 3999-4021.
- Jameson, D. M. and S. E. Seifried (1999). "Quantification of protein-protein interactions using fluorescence polarization." *Methods* **19**(2): 222-233.
- Jerabek-Willemsen, M., C. J. Wienken, et al. (2011). "Molecular interaction studies using microscale thermophoresis." *Assay Drug Dev Technol* **9**(4): 342-353.
- Kairys, V., L. Baranauskiene, et al. (2019). "Binding affinity in drug design: experimental and computational techniques." *Expert Opin Drug Discov* **14**(8): 755-768.
- Knudsen, N. O., S. D. Andersen, et al. (2009). "Nuclear translocation contributes to regulation of DNA excision repair activities." *DNA Repair (Amst)* **8**(6): 682-689.
- Kobe, B. (1999). "Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha." *Nat Struct Biol* **6**(4): 388-397.
- Kohler, M., H. Haller, et al. (1999). "Nuclear protein transport pathways." *Exp Nephrol* **7**(4): 290-294.
- Ladbury, J. E. (2010). "Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design." *Biochem Soc Trans* **38**(4): 888-893.
- Lanford, R. E. and J. S. Butel (1984). "Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen." *Cell* **37**(3): 801-813.
- Lanford, R. E., P. Kanda, et al. (1986). "Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal." *Cell* **46**(4): 575-582.
- Lange, A., R. E. Mills, et al. (2007). "Classical nuclear localization signals: definition, function, and interaction with importin alpha." *J Biol Chem* **282**(8): 5101-5105.
- Lee, S. J., T. Sekimoto, et al. (2003). "Crystallization and preliminary crystallographic analysis of the importin-beta-SREBP-2 complex." *Acta Crystallogr D Biol Crystallogr* **59**(Pt 10): 1866-1868.
- Macara, I. G. (2001). "Transport into and out of the nucleus." *Microbiol Mol Biol Rev* **65**(4): 570-594, table of contents.
- Matsuura, Y. and M. Stewart (2005). "Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling." *EMBO J* **24**(21): 3681-3689.
- Moore, M. S. and G. Blobel (1993). "The GTP-binding protein Ran/TC4 is required for protein import into the nucleus." *Nature* **365**(6447): 661-663.
- Peifer, M., L. M. Pai, et al. (1994). "Phosphorylation of the Drosophila adherens junction protein Armadillo: roles for wingless signal and zeste-white 3 kinase." *Dev Biol* **166**(2): 543-556.

- Perozzo, R., G. Folkers, et al. (2004). "Thermodynamics of protein-ligand interactions: history, presence, and future aspects." *J Recept Signal Transduct Res* **24**(1-2): 1-52.
- Pethica, B. A. (2015). "Misuse of thermodynamics in the interpretation of isothermal titration calorimetry data for ligand binding to proteins." *Anal Biochem* **472**: 21-29.
- Pierce, M. M., C. S. Raman, et al. (1999). "Isothermal titration calorimetry of protein-protein interactions." *Methods* **19**(2): 213-221.
- Rossi, A. M. and C. W. Taylor (2011). "Analysis of protein-ligand interactions by fluorescence polarization." *Nat Protoc* **6**(3): 365-387.
- Salvador, G. H. M., E. K. R. Pinto, et al. (2023). "Structural basis of the myotoxic inhibition of the Bothrops pirajai PrTX-I by the synthetic varespladib." *Biochimie* **207**: 1-10.
- Seidel, S. A., P. M. Dijkman, et al. (2013). "Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions." *Methods* **59**(3): 301-315.
- Stewart, M. (2006). "Structural basis for the nuclear protein import cycle." *Biochem Soc Trans* **34**(Pt 5): 701-704.
- Stewart, M. (2007). "Molecular mechanism of the nuclear protein import cycle." *Nat Rev Mol Cell Biol* **8**(3): 195-208.
- Tellinghuisen, D. J. and E. J. Nowak (2003). "The inability to ignore auditory distractors as a function of visual task perceptual load." *Percept Psychophys* **65**(5): 817-828.
- Tellinghuisen, J. (2003). "A study of statistical error in isothermal titration calorimetry." *Anal Biochem* **321**(1): 79-88.
- Tellinghuisen, J. (2004). "Statistical error in isothermal titration calorimetry." *Methods Enzymol* **383**: 245-282.
- Tellinghuisen, J. (2004). "Volume errors in isothermal titration calorimetry." *Anal Biochem* **333**(2): 405-406.
- Tellinghuisen, J. (2005). "Statistical error in isothermal titration calorimetry: variance function estimation from generalized least squares." *Anal Biochem* **343**(1): 106-115.
- Vega, S., O. Abian, et al. (2015). "A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry." *Methods* **76**: 99-115.
- Velazquez-Campoy, A., H. Ohtaka, et al. (2004). "Isothermal titration calorimetry." *Curr Protoc Cell Biol* **Chapter 17**: Unit 17 18.
- Velazquez Campoy, A. and E. Freire (2005). "ITC in the post-genomic era...? Priceless." *Biophys Chem* **115**(2-3): 115-124.
- Watzig, H., I. Oltmann-Norden, et al. (2015). "Data quality in drug discovery: the role of analytical performance in ligand binding assays." *J Comput Aided Mol Des* **29**(9): 847-865.
- Weis, K. (2003). "Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle." *Cell* **112**(4): 441-451.
- Wiseman, T., S. Williston, et al. (1989). "Rapid measurement of binding constants and heats of binding using a new titration calorimeter." *Anal Biochem* **179**(1): 131-137.
- Zhao, H., R. Ghirlando, et al. (2015). "A multilaboratory comparison of calibration accuracy and the performance of external references in analytical ultracentrifugation." *PLoS One* **10**(5): e0126420.