

**MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO TRIANGULO MINEIRO  
PÓS-GRADUAÇÃO MULTICÊNTRICO EM QUÍMICA DE MINAS GERAIS**

**PAMELA APARECIDA CANDIDO**

*Modelos Teóricos Simplificados Aplicados no  
Desenvolvimento de Proteínas para Diagnóstico e  
Imunoterapia de Alergia Respiratória*

**Orientador:** *Prof. Dr. Ronaldo Junio de Oliveira*  
*DF/ICENE/UFTM*

**UBERABA - MG  
2016**

**MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO  
PÓS-GRADUAÇÃO MULTICÊNTRICO EM QUÍMICA DE MINAS GERAIS**

*Modelos Teóricos Simplificados Aplicados no  
Desenvolvimento de Proteínas para Diagnóstico e  
Imunoterapia de Alergia Respiratória*

*Projeto de Mestrado apresentado ao Programa Pós-graduação Multicêntrico em Química de Minas Gerais da Universidade Federal do Triângulo Mineiro em colaboração com o Instituto de Ciências Exatas, Naturais e Educação / Departamento de Física da Universidade Federal do Triângulo Mineiro (ICENE/UFTM). Com a finalidade para ser utilizado na Dissertação de Mestrado para a obtenção do Título de Mestre.*

**Orientador:** Prof. Dr. Ronaldo Junio de Oliveira.

**UBERABA - MG  
2016**

**Catálogo na fonte: Biblioteca da Universidade Federal do  
Triângulo Mineiro**

C223m Candido, Pamela Aparecida  
Modelos teóricos simplificados aplicados no desenvolvimento  
de proteínas para diagnóstico e imunoterapia de alergia respirató-  
ria / Pamela Aparecida Candido. -- 2016.  
53 f. : il.

Dissertação (Mestrado em Química) -- Universidade Federal  
do Triângulo Mineiro, Uberaba, MG, 2016

Orientador: Prof. Dr. Ronaldo Junio de Oliveira

1. Alergia respiratória. 2. *Blomia tropicalis*. 3. Imunoterapia.  
4. Métodos de simulação. I. Oliveira, Ronaldo Junio de. II. Univer-  
sidade Federal do Triângulo Mineiro. III. Título.

CDU 616.2

PÂMELA APARECIDA CÂNDIDO

Modelos Teóricos Simplificados Aplicados no Desenvolvimento de Proteínas para Diagnóstico e Imunoterapia de Alergia Respiratória

Dissertação apresentada ao Curso de Pós-Graduação Multicêntrico em Química de Minas Gerais, área de concentração Química Teórica Computacional, da Universidade Federal do Triângulo Mineiro, como requisito parcial para a obtenção do título de mestre em Química.

Uberaba, 11 de Outubro de 2016.

Banca examinadora:



---

Prof. Dr. Ronaldo Junio de Oliveira

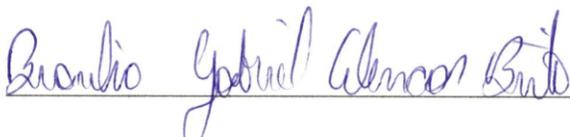
Universidade Federal do Triângulo Mineiro



---

Prof. Dr. Luis Paulo Barbour Scott

Universidade Federal do ABC



---

Prof. Dr. Bráulio Gabriel Alencar Brito

Universidade Federal do Triângulo Mineiro

## **DEDICATÓRIA**

*Aos meus pais, Maria Helena e Luiz Roberto, razões do meu viver, por todo apoio que me deram durante toda a minha vida e por sempre acreditarem que a Educação é essencial na formação do ser humano. Que sempre se esforçaram ao máximo para que eu fosse educada corretamente e que sempre foram meu exemplo de doçura, amor incondicional e dedicação.*

*Ao meu querido esposo, Mário Sérgio, pelo apoio constante e por me complementar em todas as horas, pela compreensão, carinho e dedicação.*

*Aos meus queridos irmãos, Paulo e Pricila, por todo carinho e amizade, mesmo na distância para conseguir dar um passo no longo caminho da vida profissional e também pela grande contribuição de vocês no meu crescimento como pessoa, amo vocês sempre...*

## **AGRADECIMENTOS**

*Agradeço a DEUS em primeiro lugar pela fortaleza, inteligência e coragem que tem me dado para o objetivo que hoje estou cumprindo!*

*Agradeço ao meu orientador Ronaldo Junio de Oliveira pela oportunidade, pelos ensinamentos constantes, pelas discussões científicas de excelente nível, por me incentivar sempre à busca pelo conhecimento, por ser sempre um apoio e agente motivador desse trabalho e, por fim, por continuar sendo uma fonte de inspiração intelectual. Minha admiração e respeito por sua carreira profissional e postura ética são indescritíveis.*

*Agradeço em especial ao meu companheiro de trabalho Wilian Verro pelos ensinamentos e paciência, que tão prontamente se disponibilizou a me ensinar a trabalhar com simulação computacional além de participar ativamente em discussões a respeito do meu projeto. Obrigado pelas horas dedicadas e por um pouco dos ensinamentos que eu pude absorver daquilo que você me passou!*

*Ao Vinicius Contessoto e Vinicius Moraes pela intensa colaboração em meu trabalho.*

*Aos meus companheiros, Angélica Nakagawa, Eric, Pedro Túlio que passaram continuamente a energia e o brio pelo conhecimento na III Escola de Modelagem Molecular ao qual foram reforçados durante todo esse período a nossa amizade.*

*A todos os amigos do grupo Cléa, Ederson, Frederico, Geraldo, Lívia, Lorena, Luciana, Luis Guilherme, Renato, Rogério, pela boa convivência, pelo companheirismo e pelas experiências.*

*Aos amigos da Pós-Graduação em especial Luiza Pomarolli pelo companheirismo e atenção.*

*À secretária Mayla do Curso de Programa de Pós-Graduação Multicêntrico em Química de Minas Gerais (PPGMQ-MG), que sempre estive presente quando preciso e que me dedicou grande atenção.*

*Às pessoas que de maneira tão especial fizeram parte do meu cotidiano durante esse período, contribuindo direta ou indiretamente para o meu amadurecimento intelectual e pessoal.*

*Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.*

## RESUMO

Ácaros da família *Blomia tropicalis* são predominantes na poeira domiciliar e são os principais responsáveis por doenças respiratórias alérgicas no mundo, geralmente em regiões tropicais e subtropicais. Neste estudo, foi aplicado simulações de dinâmica molecular com modelos simplificados e cálculos de interações eletrostáticas para a caracterização termodinâmica da proteína do alérgeno *Blomia tropicalis* do grupo 5 (rBlo t 5). O objetivo foi o estudo de um novo alérgeno com a mesma característica termodinâmica da *Blomia tropicalis* na sua forma nativa (rBlo t 5) mas com propriedades imunológicas modificadas com a finalidade de promover a redução da alergenicidade por IgE e aumentar a imunogenicidade da resposta para uso em imunoterapia. Foi realizado a simulação de modelos baseados em estrutura apenas com os carbonos alfa da proteína e a análise termodinâmica mostrou que a Blo t 5 nativa e a modificada são similares apresentando uma termo-estabilidade de  $\Delta F = 2.2$  kT. A interação carga-carga pela nova técnica Tanford-Kirkwood Surface Accessibility - Monte Carlo (TKSA-MC) tem sido empregado para compreensão da termoestabilidade de proteínas e para a modelagem por meio da otimização da interação eletrostática e foi aplicada nas proteínas em estudo. Os resultados de TKSA-MC mostram uma interação eletrostática mais favorável na região do epitopo de mBlo t 5 com relação ao rBlo t 5. Estes resultados teóricos corroboraram com os dados experimentais que mostram que o alérgeno modificado de rBlo t 5 tem uma mudança de isotipo para um aumento da antigenicidade do anticorpo IgG4, tornando assim a sua aplicação adequada como candidato potencial em estudos imunoterapia com alérgenos específicos.

**Palavras-Chave:** Alergia respiratória. Modelo baseado na estrutura. Interação eletrostática.

## **ABSTRACT**

*Blomia tropicalis* family mites are prevalent in house dust and they are primarily responsible for allergic respiratory diseases in the world, generally in tropical and subtropical regions. In this study, we applied molecular dynamics simulations with simplified models and electrostatic calculations for the characterization of the recombinant *Blomia tropicalis* allergen protein group 5 (rBlo t 5). The objective was produce a new allergen with the same thermodynamic characteristic of *Blomia tropicalis* in their native form (rBlo t 5), therefore with immunological properties modified in order to promote the reduction of allergenicity by IgE and increased the immunogenicity response for use immunotherapy. It was performed structure-based model simulations only with the protein alpha carbons. Thermodynamic analysis showed that the native and the mutant Blo t 5 have similar thermalstability of  $\Delta F = 2.2$  kT. The charge-charge interaction by the new technique Tanford-Kirkwood Surface Accessibility – Monte Carlo (TKSA-MC) has been employed for the understanding of protein thermostability and modeling by means of optimization of the electrostatic interaction and it was applied in the proteins under study. The TKSA-MC results show an electrostatic interaction more favourable in the epitope region of mBlo T 5 with relation to rBlo t 5. These theoretical results corroborated with the experimental data, which shows that the modified allergen rBlo t 5 has a shift isotype to an increased antigenicity for the antibody IgG4, thus making suitable its application as potential candidate for immunotherapy studies with specific allergens.

**Key words:** Respiratory allergy. Structure-based model. Electrostatic interaction

## LISTA DE ABREVIATURAS

3D – Três Dimensões

APC – Células Apresentadoras de Antígenos

ARG – Arginina

ASP – Aspartato

Blo t – *Blomia tropicalis*

CHP – Complexo de Histocompatibilidade Principal

CSU – Algoritmo Contact of Structural Units

C $\alpha$  – Carbono alfa

Der p – *Dermatophagoides pteronyssinus*

Fc $\epsilon$ RI – Receptor I de IgE de alta afinidade

GLU – Glutamato

GROMACS - GRONingen MACHine for Chemical Simulations

HIS – Histidina

IgE – Imunoglobulina de classe E

IgG - Imunoglobulina de classe G

IL-13 – Interleucina 13

IL-4 – Interleucina 4

IL-5 – Interleucina 5

kDa – Quilodaltos

LYS – Lisina

PDB – Protein Data Bank

RMSD - Root Mean Square Deviation

RMSF - Root Mean Square Fluctuation

SBM – Modelo baseado em estrutura

SDS – Dodecil Sulfato de Sódio

SDS-PAGE – Eletroforese em gel de poliacrilamina na presença de SDS

SMOG - Structure-Based Models in GROMACS

Th2 – Linfócito T helper 2

TKSA-MC - Tanford-Kirkwood com área acessível ao solvente por Monte Carlo

VMD - Virtual Molecular Dynamics

WHAM - Weighted Histogram Analysis Method

## LISTA DE FIGURAS

Figura 1 -	Mecanismo do Sistema Imune quando em contato com o alérgeno.....	18
Figura 2 -	Representação da rBlo t 5 obtida pelo VMD.....	24
Figura 3 -	Representação do modelo Tanford-Kirkwood.....	28
Figura 4 -	Representações das sequências de aminoácidos do alérgeno <i>Blomia tropicalis</i> do grupo 5 recombinante, a sequência nativa (rBlo t 5) e o alérgeno modificado (mBlo t 5).....	30
Figura 5 -	Contatos nativos (Q) em função do tempo computacional em unidades reduzidas.....	32
Figura 6 -	Calor específico da estrutura mBlo t 5_MH (verde), mBlo t 5_IT (azul) e rBlo t 5 (vermelho).....	32
Figura 7 -	Os perfis de energia livre em função da reação de coordenadas fração de contatos nativos (Q) à temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo $C\alpha$ .....	34
Figura 8 -	Otimização de interação eletrostática calculado na proteína nos resíduos carregados pelo modelo de Acessibilidade Superfície Tanford-Kirkwood implementados usando simulações de Monte Carlo (TKSA-MC).....	36
Figura 9 -	Correlação entre os dados teóricos e experimentais.....	38
Figura 10 -	Calor específico das Estruturas Blo t 12 (vermelho), Blo t 21 (amarelo), Der f 13 (violeta), Der p 2 (laranja), Der p 23 (azul turquesa), Der p 5 (azul) e Blo t 5 (verde).....	52
Figura 11 -	Os perfis de energia livre em função da reação de coordenadas fração de contatos nativos (Q) à temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo $C\alpha$ para as proteínas Blo t 5 (verde), Der p 2 (laranja) e Der f 13 (violeta).....	52

Figura 12 -	Os perfis de energia livre em função da reação de coordenadas fração de contatos nativos (Q) à temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo $C\alpha$ para as proteínas <i>Der p 5</i> (azul), <i>Blo t 21</i> (amarelo), <i>Blo t 12</i> (vermelho) e <i>Der p 23</i> (azul turquesa).....	53
-------------	--	----

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	14
1.1	<i>Reações alérgicas e as imunoglobulinas G e E</i>	16
1.2	<i>Mecanismo da resposta imune da alergia</i>	19
1.3	<i>Modelo baseado na estrutura</i>	21
1.3.1	<i>Modelo carbono alfa</i>	21
1.3.2	<i>Cálculo de interação carga-carga – TKSA-MC</i>	22
2	<b>MATERIAL E MÉTODOS</b>	23
2.1	<i>Modelagem da estrutura 3D da proteína modificada mBlo t 5</i>	23
2.2	<i>Simulação computacional</i>	23
2.3	<i>Cálculo de interação carga-carga – TKSA-MC</i>	27
3	<b>RESULTADOS E DISCUSSÕES</b>	29
3.1	<i>Estrutura da proteína Blomia tropicalis do grupo 5 (Blo t 5)</i>	29
3.2	<i>Análise termodinâmica</i>	31
3.2.2	<i>Energia livre versus contatos nativos (Q) para modelo C-alfa</i>	33
3.3	<i>Otimização de interação eletrostática dos alérgenos recombinantes</i>	35
3.4	<i>Correlação das predições teóricas com os dados experimentais</i>	37
3.5	<i>Discussões</i>	39
4	<b>CONCLUSÕES</b>	40
	<b>REFERÊNCIAS</b>	41
	<b>APÊNDICE A – Trabalho em andamento</b>	50

---

## 1 INTRODUÇÃO

Doenças alérgicas são consideradas como a quarta doença mais importante do mundo, de acordo com a Organização Mundial de Saúde, e estima-se que de 10% a 40% da população é afetada por doenças alérgicas causadas por ácaros presentes em poeira domiciliar. O ácaro é considerado como o principal fator desencadeador de asma e outras doenças alérgicas na população. Os ácaros das espécies *Blomia ssp* e *Dermatophagoides ssp* estão entre os mais comumente encontrados (CHUA et al., 2007).

A espécie *Blomia tropicalis* (*Blo t*) pertence à superfamília *Glycyphagoidea*, e família *Echimyopodidae*, e está presente em regiões tropicais e subtropicais, onde apresentam uma maior importância clínica. Os organismos da espécie *Dermatophagoides pteronyssinus* (*Der p*) são encontrados em poeira domiciliar de todo o mundo e pertencem à superfamília *Analgoidea*, da família *Pyroglyphidae*. (CHUA et al., 2007).

Os ácaros encontrados na poeira domiciliar se alimentam da descamação de células mortas da pele humana e existem mais de 40 espécies e entre essas estão presentes as proteínas do Grupo 5 como a *Blo t 5* e a *Der p 5*, responsáveis pela reação alérgica. Geralmente, são encontrados em colchões, cobertores, travesseiros (FERNANDEZ-CALDAS et al., 1997; YONG et al., 2009).

O peptídeo de sinal geralmente é composto de uma sequência de 15 a 60 aminoácidos localizados na região N-terminal ou amino-terminal de proteínas (ALBERTS et al., 2008). As proteínas são sintetizadas em compartimentos celulares e mesmo assim acabam não desempenhando sua função biológica onde foram sintetizadas e precisam ser carregadas para a região específica onde executarão suas funções. A função da sequência do peptídeo sinal é de dirigir a proteína para determinados locais, como por exemplo, o ambiente extracelular (CLARK et al., 2005).

Quando a proteína chega ao seu local específico são reconhecidas por meio do peptídeo sinal, mas depois que acontece esse reconhecimento da proteína automaticamente esse peptídeo pode ser removido por intermédio da enzima protease ou não, mantendo-se como parte integrante da proteína quando localizados internamente (CLARK et al., 2005).

---

Na *Blot 5*, esse peptídeo de sinal consiste em 17 resíduos de aminoácidos que estão entre os 117 aminoácidos da proteína desenvolvida, enquanto este mesmo peptídeo na *Der p 5* consiste em 19 resíduos de aminoácidos que estão entre os 113 aminoácidos da proteína (CHUA et al., 2007). Estas proteínas são homólogas e possuem uma identidade de 42%.

Os aminoácidos dessas proteínas que apresentam carga positiva são lisina e arginina existindo 18 desses resíduos para a *Blot 5* e 21 resíduos para a *Der p 5*. Os aminoácidos de carga negativa são aspartato e glutamato e existem 26 desses resíduos tanto para a *Blot 5* quanto para a *Der p 5* (CHUA et al., 2007). Além dessa similaridade, a reação para reatividade por ligação à imunoglobulina E (IgE) é de 43-92% para a *Blot 5* e de 50-70% para a *Der p 5* (CHUA et al., 2007).

A biologia estrutural das proteínas desses alérgenos é utilizada para sua classificação e é importante para a compreensão desses alérgenos, descrevendo as propriedades básicas dos mesmos para determinar a capacidade de ligação à IgE e prever a possibilidade de reação cruzada com outros alérgenos relacionados, contribuindo, dessa forma, para o desenvolvimento de uma imunoterapia contra esses alérgenos (CUI et al., 2013).

Estudos são desenvolvidos com a finalidade de compreender os mecanismos imunológicos, moleculares e, principalmente, as reações clínicas desses alérgenos que são desencadeadas em pacientes que apresentam sensibilidade à IgE (CUI et al., 2013).

Estudos baseados na estrutura do alérgeno tem promovido grandes avanços no campo da tecnologia computacional a fim de permitir a produção em larga escala de alérgenos recombinantes. Com isso, obtém-se uma produção de tratamentos cada vez mais eficazes para a redução da doença alérgica em pacientes alérgicos (CUI et al., 2013).

Os estudos realizados no início dos anos 90 usaram algumas técnicas moleculares e imunológicas que são utilizadas até hoje para promover informações de caracterização sobre os alérgenos como: cromatografia de afinidade para a purificação das proteínas dos alérgenos, SDS-PAGE para determinar características como determinar a mobilidade de proteínas e o tamanho ao qual dá-se o nome de peso molecular; e ELISA Radioimunoensaio para quantificar a alergenicidade, a capacidade do alérgeno se ligar à IgE.

### 1.1 Reações alérgicas e as imunoglobulinas G e E

Há evidências de que os alérgenos de *Blomia tropicalis* não são semelhantes aos alérgenos de *Dermatophagoides pteronyssinus* e apresentam uma moderada reatividade cruzada entre eles. Isso pode ser observado pelo fato de que anticorpos de pacientes alérgicos contra o alérgeno *B. tropicalis*, que são proteínas com peso de 14,3 e 27,3 quilodaltons (kDa), não inibem a ligação do anticorpo anti-*D. pteronyssinus* para antígenos de *D. pteronyssinus* (ARRUDA et al., 1997).

Estudos relatam que esta reatividade cruzada, ou seja, a ligação do anti-*D. pteronyssinus* para antígenos de *D. pteronyssinus* pode estar relacionada com epitopos presentes no segmento C-terminal ou carboxi-terminal destes alérgenos. Pelo fato de existir um elevado grau de homologia deste segmento entre as duas moléculas, sugere-se que a inibição de ligação a IgE é obtido devido a existência de um epitopo comum (CARABALLO et al., 1997). O epitopo é uma pequena parte de um antígeno (proteína), ou seja, uma pequena sequência de aminoácidos que se liga às células T e B, que participam da defesa do organismo e desenvolvem os mecanismos imunológicos (PLATTS-MILLS et al., 1997).

Pacientes alérgicos, quando expostos aos alérgenos, desencadeiam uma reação por mecanismos imunológicos, produzindo altos níveis de anticorpos da classe IgE. Isso ocorre porque a excreção dos ácaros contém certo número de substâncias proteicas e quando aspiramos essas substâncias ou quando elas entram em contato com a pele, nosso organismo produz anticorpos (PLATTS-MILLS et al., 1997). Neste caso, a proteína *Blo t 5*, quando em contato com o organismo por meio das mucosas, desencadeia uma reação alérgica que se inicia pela ligação de moléculas de IgE a um antígeno específico, no caso a *Blo t 5*, culminando na ativação de processos inflamatórios (Figura 1).

Para ativação do sistema imune, são necessárias algumas etapas. Num primeiro momento ocorre a inalação de partículas ou fragmentos dos alérgenos pelas vias respiratórias por meio da mucosa das vias respiratórias e os alérgenos entram em contato com o sistema imune no organismo. Apesar das partículas fecais dos ácaros serem pequenos, eles podem chegar até os pulmões (CALDERON et al., 2015).

No epitélio, os alérgenos são capturados pelas células dendríticas e transportados para a drenagem dos gânglios linfáticos e posteriormente são

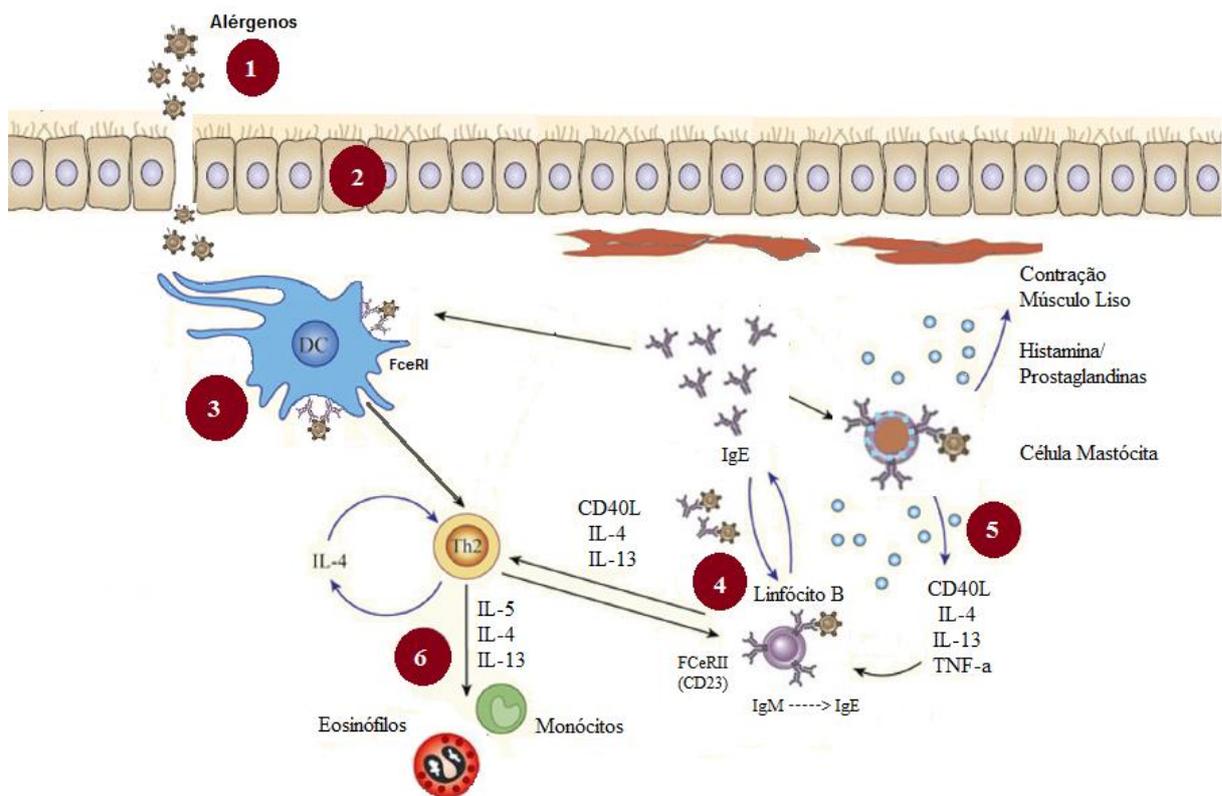
---

apresentados aos linfócitos (CALDERON et al., 2015). Nos gânglios, linfáticos ocorre a apresentação de antígenos, momento em que as células dendríticas interagem com alérgeno específico das células T. Fragmentos de alérgeno são parcialmente degradadas e apresentados para a célula T no Complexo de Histocompatibilidade Principal (CHP). Em indivíduos atópicos, as células dendríticas são distorcidas para promover diferenciação de células do perfil Th2 em células T, e clivagem proteolítica solúvel de baixa afinidade do receptor de IgE por proteases derivado do ácaro podendo aumentar ainda mais a diferenciação do perfil Th2 (CALDERON et al., 2015).

Células de perfil Th2 alérgeno-específicas produzem IL-4, IL-5 e IL-13. Estas citocinas induzem a uma mudança de classe em células B e produção de IgE em células B ativadas. A exposição contínua ao alérgeno levam a produção contínua de IgE específico para o mesmo, bem como a manutenção contínua de inflamação alérgica das vias aéreas. E isto torna o indivíduo cada vez mais sensibilizado e predisponente a reações alérgicas (CALDERON et al., 2015).

A IgE solúvel secretada pelas células B ativadas se liga aos receptores FcεRI de alta afinidade específica em basófilos e mastócitos. Estas células são degranuladas com a ação de enzimas e produtos químicos, histamina, que pode ser liberada imediatamente após repetido encontro com alérgeno. Ações diretas de proteases do antígeno em basófilos e mastócitos podem aumentar a degranulação e intensificar esta resposta. Estas proteases podem também atuar em células musculares lisas e aumentar a contração (CALDERON et al., 2015).

Respostas inflamatórias sustentadas são mantidas por exposição contínua ao alérgeno. A IL-5 produzida por células do perfil Th2 alérgeno-específicos ativam e recrutam granulócitos dos eosinófilos. Eosinófilos potentes produzem enzimas e produtos químicos de condução e remodelamento das vias aéreas para progressão da asma (CALDERON et al., 2015).



**Figura 1:** Mecanismo do Sistema Imune quando em contato com o alérgeno. Figura adaptada de CALDERON et al., 2015. 1 – Mucosa das vias respiratórias, 2 – Epitélio, 3 - Apresentação de antígenos, 4 - Produção de IgE, 5 - Reação alérgica imediata, 6 - Resposta inflamatória sustentada são mantidas por exposição contínua ao alérgeno.

Dessa forma, se faz necessário a compreensão, em nível molecular, da região onde o alérgeno é reconhecido pelos anticorpos IgE e IgG, chamada epitopo, e os mecanismos pela preferência do antígeno por IgE em pacientes alérgicos. Em indivíduos não predispostos, o alérgeno é reconhecido pelo sistema imune ligando-se, principalmente, aos anticorpos das classes IgG1 e IgG4 que não desencadeiam manifestação alérgica (HOLGATE et al., 2008). A indução de IgG4 desempenha um papel bloqueador pela inibição de eventos inflamatórios mediados por IgE (HOLGATE et al., 2008). Vários estudos frequentemente citam a alta prevalência de alergia causada por ácaros da poeira domiciliar (HDM<sup>1</sup>), apesar de que uma estimativa global provou ser ilusório.

Estudos de alergia sugerem que 1% a 2% da população do mundo pode ser afetada, o que equivale a cerca de 130 milhões de pessoas e isso ainda depende da variação geográfica que, por sua vez, complica o quadro. Apesar desse tipo de alergia ser consistentemente encontrado em países ocidentais, a variação entre

<sup>1</sup> Do inglês House Dust Mite

---

países, regiões e até mesmo laboratórios de teste individuais é significativa. Isto sugere que as discrepâncias no uso de ferramentas de diagnóstico podem confundir estudos epidemiológicos (CALDERON et al., 2015).

## 1.2 Mecanismo da resposta imune da alergia

A IgE é um anticorpo encontrado no CHP, essa imunoglobulina é a principal responsável capaz de mediar as respostas de Hipersensibilidade do Tipo I quando se liga aos antígenos. Quando o organismo entra em contato com um alérgeno, estes são imediatamente internalizados e processados pelas Células Apresentadoras de Antígenos (APC) para apresentação aos linfócitos T.

Os antígenos são apresentados para as células T auxiliares CD4<sup>+</sup> pelas APC ocorrendo um processamento desses antígenos pelas Células T auxiliares CD4<sup>+</sup> e posterior migração para os gânglios linfáticos onde os Linfócitos B serão ativados para a produção de IgE específicos para alérgenos e citocinas do perfil Th2 como a IL-4, IL-5 e IL-13.

Após estimulação, os Linfócitos B apresentam na membrana receptores específicos para o antígeno *Blo t 5* e, em seguida, acontece uma diferenciação dos Linfócitos B em células plasmócitos e células denominadas células de memória. Os plasmócitos segregam anticorpos específicos para cada antígeno IgE e as Células de Memória ficam inativas, prontas para serem ativadas novamente quando se tem um novo contato com o mesmo antígeno com uma maior rapidez, maior intensidade e dura mais tempo a reação alérgica.

Posteriormente, a IgE será liberada na corrente sanguínea e se ligará na célula mastócita por meio do receptor *FcεRI* provocando uma sensibilização dos mastócitos ou basófilos. Esse mecanismo ocorre na primeira exposição ao antígeno, e numa segunda exposição, como os mastócitos estão sensibilizados pela IgE, automaticamente os antígenos *Blo t 5* se ligarão na célula mastócita promovendo uma resposta de hipersensibilidade do tipo I, pois as moléculas como a histamina e outras substâncias mediadoras são produzidas e liberadas na corrente sanguínea onde desenvolvem as reações alérgicas e, neste caso, conhece-se como Resposta Secundária, resultando em vasodilatação e aumento da permeabilidade dos vasos, constrição das vias respiratórias, edema, dor e calor (CUI et al., 2014).

---

Resumindo, as reações alérgicas ocorrem em duas fases: a fase de sensibilização do mastócito, na primeira exposição ao antígeno, e a resposta secundária, na segunda exposição ao antígeno. No caso das Células B, os epitopos que são capazes de promover alergia de uma proteína são reconhecidas por IgE e, no caso das células T, esses epitopos podem estimular indiretamente a produção de IgE por meio de um processo mediado por células T auxiliares CD4<sup>+</sup> (CUI et al., 2014).

A análise de um alérgeno envolve a clonagem, expressão e a caracterização molecular seguido da avaliação de alergenicidade. Para a caracterização molecular, pode ser realizado uma comparação de outras sequências homólogas e a análise estrutural tridimensional (3D). É por meio da estrutura 3D da proteína que se determina a alergenicidade, sendo assim, o conhecimento estrutural pode ajudar a fornecer parâmetros sobre a base molecular da causa de reações alérgicas e ainda a elucidação da função biológica de alérgenos (BRUSIC et al., 2003).

Devido ao grande número de pessoas alérgicas e ao rápido crescimento da população com predisposição para a alergia respiratória, os métodos de bioinformática como Dinâmica Molecular com simulações computacionais estão se tornando cada vez mais indispensáveis para promover uma análise detalhada da sequência desses alérgenos, assim como suas estruturas, propriedades funcionais e alergenicidade (BRUSIC et al., 2003). Diante de uma grande necessidade, vários métodos computacionais foram introduzidos como ferramentas cruciais para prever potenciais alergênicos de proteínas a partir de sua sequência nativa ou de uma sequência homóloga à sequência de interesse (BRUSIC et al., 2003).

No entanto, deve-se considerar que os alérgenos homólogos possuem sequência de aminoácidos e estrutura 3D com diferenças significativas na região proximal dos resíduos carregados do epitopo e apresenta uma afinidade diferente para a ligação com IgE (TAN et al., 2012; CHRUSZCZ et al., 2009). Sendo assim, a estabilidade de proteínas e as interações eletrostáticas dos resíduos carregados na região do epitopo deve ser levada em consideração para calibrar a atividade antialérgica.

A motivação deste trabalho em estudar o alérgeno *Blomia tropicalis in silico* foi o estudo experimental do alérgeno modificado pela Msc. Bárbara Godoi Melo Ávila sob a orientação do Prof. Dr. Ernesto A. Taketomi e coorientação do Dr. Jair P.

---

da Cunha Jr., ambos do Laboratório de Alergia e Imunologia Clínica, Instituto de Ciências Biomédicas, Universidade Federal de Uberlândia, Uberlândia, MG (ÁVILA et al., 2012).

### 1.3 Modelo baseado na estrutura

A simulação computacional de modelos minimalistas tem contribuído para o entendimento do enovelamento de proteínas (CLEMENTI et al., 2000; WHITFORD et al., 2009), dimerização (LU et al., 2007), mudanças conformacionais funcionais e reações enzimáticas (WHITFORD et al., 2007; BEST et al., 2005; ZUCKERMAN et al., 2004; LU et al., 2008), entre outros sistemas biomoleculares como os estudados pelo grupo (CHAHINE et al., 2007; OLIVEIRA et al., 2010; OLIVEIRA et al., 2010; WANG et al., 2012; XU et al., 2012; CONTESSOTO et al., 2013; MANDELLI et al., 2013). Isso porque, a simulação de modelos mais realísticos de proteínas envolve mais parâmetros e interações atômicas e, como consequência, o tempo computacional necessário para se extrair informações relevantes aumenta dramaticamente.

#### 1.3.1 Modelo carbono alfa

A forma mais importante dos aminoácidos são os chamados  $\alpha$ - aminoácidos que formam as proteínas, pois, tem como estrutura um carbono central representados como carbono alfa ( $C\alpha$ ), quase sempre quiral, ao qual se ligam a 4 grupos sendo: grupo amina, grupo carboxílico, hidrogênio e um radical constituído por uma cadeia orgânica, exceto na glicina onde esse radical é outro átomo de hidrogênio. Sendo assim, com exceção da glicina, os aminoácidos apresentam atividade óptica pelo fato de apresentarem 4 grupos diferentes ligados ao carbono central. O que demonstra que esse carbono é quiral é o fato dele ser assimétrico e isso permite a formação de estereoisômeros desses aminoácidos devido aos diferentes arranjos espaciais opticamente ativos. Esses carbonos são assimétricos, formam imagem especular um do outro e não se sobrepõem a outro.

A simulação computacional baseada na dinâmica molecular do modelo  $C\alpha$  (CLEMENTI et al., 2010) utiliza uma simplificação fundamental. Todos os átomos de um aminoácido da cadeia polipeptídica são substituídos por uma esfera centrada na posição do  $C\alpha$  correspondente, mantendo seu raio de volume excluído. Dessa forma,

---

apenas o C $\alpha$  da cadeia principal é representado explicitamente e interagem entre si pelo potencial dado pela equação 1 com uma pequena adaptação. No modelo C $\alpha$ , existem somente os termos de ligação covalente, angular e diedral para manter a geometria da cadeia principal. Os ângulos diedrais são formados entre quatro átomos C $\alpha$  adjacentes e os contatos não-locais interagem via potencial 12-10 ao invés do 12-6 do Lennard-Jones típico do modelo de todos-os-átomos (WHITFORD et al., 2009).

#### 1.4 Cálculo de interação carga-carga – TKSA-MC

O método calcula a média de interação carga-carga e tem sido utilizado para compreender a termo-estabilização de proteínas e enzimas para redesenhar por otimização das interações eletrostáticas de superfície (IBARRA-MOLERO et al., 1999; PETERSEN et al., 1997; IBARRA-MOLERO et al., 2002; MAKATHADZE ET AL., 2004; PERMYAKOV et al., 2005; STRICKLER et al., 2006; GRIBENKO et al., 2009). Neste estudo, foi estabilizado as interações eletrostáticas entre cada resíduo ionizável na proteína modificada, *mBlt 5*, (menores valores de energia negativa) (NAIK et al., 2008) em relação a proteína nativa, *rBlt 5*. As interações carga-carga das proteínas foram determinadas pelo método TKSA-MC, modificado para ser mais eficiente em termos computacional.

A aplicação deste método é de fundamental importância quando se pretende realizar uma mutação visando o aumento da termoestabilidade, mas deve-se levar em consideração que não envolve apenas fatores energéticos ao proceder com a escolha dos resíduos. É assim o resíduo escolhido deve ter a sua cadeia lateral exposta ao solvente em mais de 50% e apresentar uma contribuição energética desfavorável. Nesse caso, a mutação não pode ser próxima do sítio catalítico para não influenciar na atividade da proteína e também não deve alterar de forma significativa a estrutura da proteína.

---

## 2 MATERIAIS E MÉTODOS

### 2.1 Modelagem da estrutura tridimensional da proteína modificada m*Blo t 5*

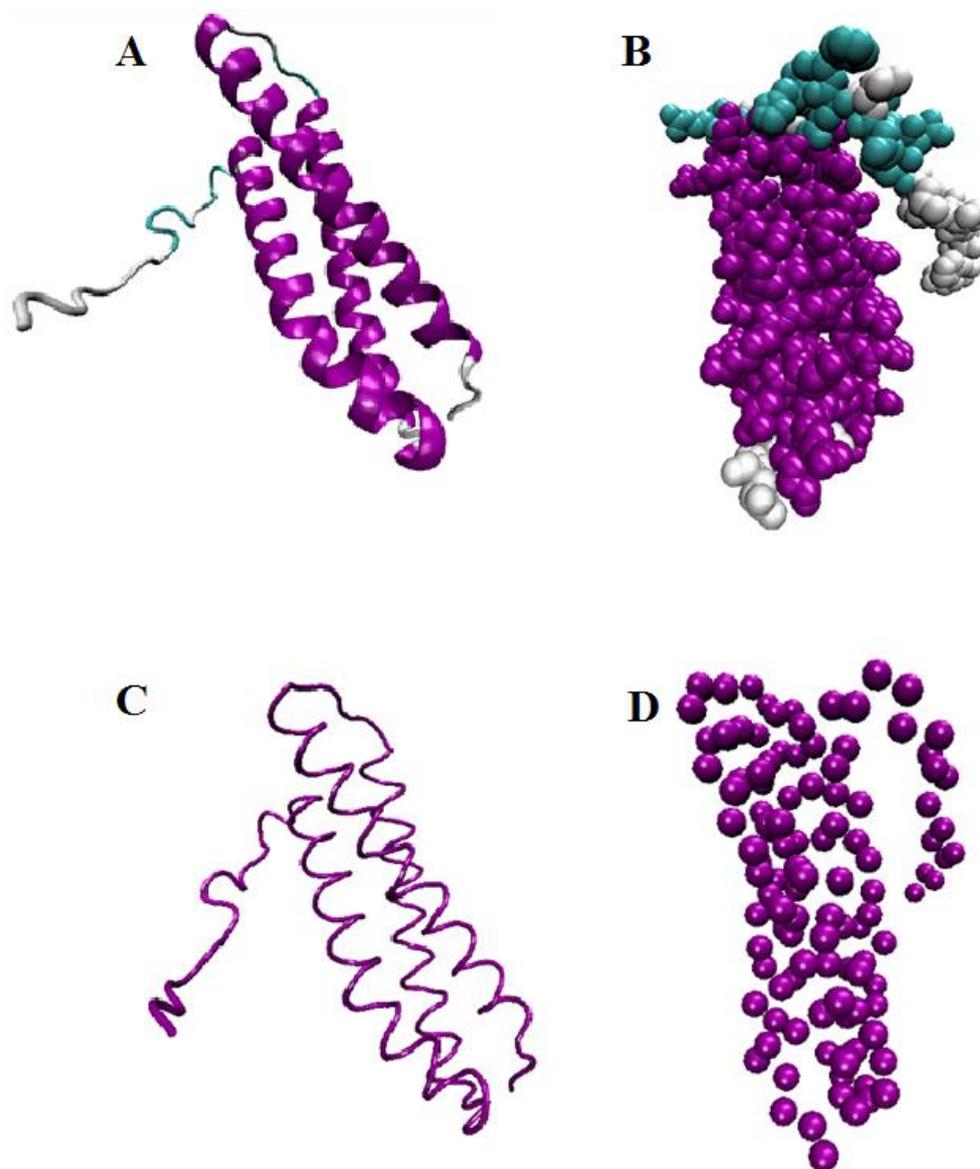
A predição da estrutura tridimensional do alérgeno recombinante modificado (m*Blo t 5*) a partir da *B. tropicalis* foi realizada pela modelagem por homologia com o servidor da web MHOLline (CAPRILES et al., 2010) e com o servidor web I-Tasser, que consiste em um fluxo de trabalho biológico que combina um conjunto específico de programas automatizados para a predição e anotação da função da estrutura da proteína. Otimização da estrutura e minimização de energia foram realizadas com campo de FoldX (SCHYMKOWITZ et al., 2005; VAN DURME et al., 2011). A modelagem por homologia de m*Blo t 5* utilizou-se a *Blo t 5* de Ressonância Magnética Nuclear (RMN) e coordenadas da estrutura (Protein Data Bank código 2JMH, modelo 1) como o modelo do modelo (Figura 3B).

### 2.2 Simulação computacional

Os arquivos de entrada para a simulação, topologia e coordenadas iniciais, foram criados pela interface gráfica do servidor SMOG-Server (NOEL et al., 2010). As simulações foram realizadas com o pacote gratuito de dinâmica molecular GROMACS, version 4.6.7 (VAN DER SPOEL et al., 2005), as simulações foram realizadas utilizando os recursos computacionais do “Núcleo de computação científica da UNESP – Gridunesp”, em seguida depois de finalizada as simulações com os dados obtidos foram feitos os cálculos de interação carga-carga onde foram usados scripts e algoritmos criados pelo grupo e padronizado para cada tipo de análise.

Ambos perfis de energia livre e calor específico termodinâmicos da *Blo t 5* foram analisados utilizando o método de análise de múltiplo histograma ponderado (WHAM) (FERRENBURG et al., 1988; FERRENBURG et al., 1989; KUMAR et al., 1992; HESS et al., 2008) ao longo de 24 temperaturas constantes variando de 0,8 a 1,5 que abrange a temperatura de enovelamento ( $T_f \sim 1.0$ ).  $T_f$  foi definida como a temperatura no pico de calor específica de cada proteína simulada a temperatura é dada em unidades reduzidas do Gromacs.

A proteína representada na Figura 2 usada no estudo é a *Blomia tropicalis* do grupo 5, *Blo t 5* (PDB: 2JMH) e que contém 976 átomos e 117 Ca.



**Figura 2:** Representação da *Blt 5* obtida pelo VMD<sup>2</sup>, a Figura A e B é constituído por todos os átomos que constituem a proteína (976 átomos) e a Figura C e D representa a proteína só com os carbonos- $\alpha$  da proteína (117 C $\alpha$ ).

Neste estudo observa-se que abordagens baseadas em metodologias computacionais têm sido amplamente aplicada em conjunto com estudos experimentais e clínicos, e que esta combinação é importante pois, desenvolve um papel fundamental nas pesquisas científicas como em estudos de imunoterapia (ZHANG et al., 2013; TAI et al., 2013; CARAVELLA et al., 2010). Estudos In Silico mediante métodos computacionais podem reduzir o tempo gasto com a realização

<sup>2</sup> Do inglês Visual Molecular Dynamics

---

de experimentos (HAKKAART et al., 1998) e triagens clínicas para se obter os dados experimentais (MOREIRA et al., 2015).

Os modelos minimalistas, por possuírem uma representação simplificada do problema, nos permitem explorar uma grande faixa de parâmetros dos sistemas em razoável tempo computacional. No caso dos modelos minimalistas baseados na estrutura, a primeira simplificação se refere ao potencial de interação entre os componentes do sistema. A segunda se refere aos próprios componentes do sistema: todos-os-átomos ou somente carbonos alfas (C $\alpha$ ). Aplica-se o modelo com todos-os-átomos quando as informações sobre as cadeias laterais são importantes para o entendimento do problema em estudo, como por exemplo, empacotamento e interação entre os resíduos de aminoácidos. Caso contrário, aplica-se o modelo C $\alpha$  pois, além de ser o modelo mais testado e utilizado pela comunidade científica, permite uma economia de tempo computacional de até 10 vezes se comparado ao modelo com todos-os-átomos.

O potencial do modelo baseado na estrutura é construído a partir da estrutura nativa do monômero ou do dímero, ou seja, os parâmetros da expressão do potencial são obtidos da estrutura nativa. O mínimo de energia corresponde à conformação da estrutura nativa da proteína depositada no Protein Data Bank (PDB) obtidas por técnicas experimentais (cristalografia de raios-X e ressonância magnética nuclear).

Os modelos cujos potenciais são construídos a partir da estrutura nativa são conhecidos na literatura como modelos G $\phi$  (UEDA et al., 1975). Com isso, ao iniciar uma simulação com a cadeia completamente aberta, no final se obtém a estrutura nativa enovelada. No caso da mudança conformacional, pode-se incluir no potencial informações da estrutura aberta e fechada e a simulação terá dois mínimos para visitar. Isso permite estudar os mecanismos que governam o enovelamento ou uma mudança conformacional induzida por um ligante e, particularmente, o aspecto do estado de transição desses processos.

Para  $C\alpha$

O potencial de uma determinada configuração  $\Gamma$  de uma proteína, tendo  $\Gamma_0$  como sua configuração no estado nativo, será dado pela expressão a seguir:

$$\begin{aligned}
 V(\Gamma, \Gamma_0) = & \sum_{bonds} \epsilon_r (r - r_0)^2 \\
 & + \sum_{angles} \epsilon_\theta (\theta - \theta_0)^2 \\
 & + \sum_{backbone} \epsilon_\phi \left\{ [1 - \cos(\phi - \phi_0)] + \frac{1}{2} [1 - \cos(3(\phi - \phi_0))] \right\} \\
 & + \sum_{contacts} \epsilon_C \left[ 5 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 6 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{10} \right] + \sum_{non-contacts} \epsilon_{NC} \left( \frac{\sigma_{NC}}{r_{ij}} \right)^{12} \quad (1)
 \end{aligned}$$

onde no primeiro termo representa a ligação entre dois carbonos  $\alpha$  adjacentes, formando um potencial harmônico,  $r_0$  é a distancia entre dois carbonos  $\alpha$  da conformação nativa ligados entre si. O segundo termo se refere ao potencial harmônico angular formado por três carbonos  $\alpha$  consecutivos, onde  $\theta_0$  é o ângulo formado pelos resíduos  $i$ ,  $i+1$  e  $i+2$  da conformação nativa. O terceiro termo contabiliza a torção realizada na cadeia, o termo utiliza quatro carbonos  $\alpha$  em sequência, os três primeiros carbonos  $\alpha$  formam um plano que possui um ângulo  $\phi_0$  com o plano formado entre três últimos carbonos  $\alpha$ . O quarto termo representa a interação entre o carbono  $\alpha$   $i$  e o carbono  $\alpha$   $j$  que realizam um contato na estrutura nativa. O parâmetro  $\sigma_{ij}$  do termo de van der Waals (última linha na equação) é determinado por meio do mapa de contatos entre os átomos na estrutura nativa da proteína. O mapa de contatos é obtido utilizando o algoritmo Contact of Structural Units (CSU) (SOBOLEV et al., 1996). Para esta informação é atribuído um potencial 10 – 12, onde o parâmetro  $d_{ij}$  é o valor da distancia entre estes carbonos que realizam um contato nativo. O último termo parâmetro  $\sigma_{NC}$  é um potencial repulsivo utilizado para manter a distância máxima de aproximação entre os carbonos  $\alpha$ , este termo é utilizado para todos os carbonos  $\alpha$  que não possuem um contato nativo. O valor do parâmetro  $\sigma_{NC}$  é de 4Å, este valor caracteriza aproximadamente o volume ocupado por um carbono do modelo. As constantes  $\epsilon_r$ ,  $\epsilon_\theta$ ,  $\epsilon_\phi$ , e  $\epsilon_{NC}$  utilizadas no potencial possuem valores de 100, 20, 1 e 1, respectivamente, todas em unidades de  $\epsilon_C$ .

### 2.3 Cálculo de interação carga-carga – TKSA-MC

A energia de interação carga-carga entre dois resíduos carregados foi calculada utilizando o modelo Tanford-Kirkwood com a contribuição da área acessível ao solvente (TKSA) (TANFORD e KIRKWOOD et al., 1957; MATTHEW e GURD et al., 1986). Neste trabalho, foi desenvolvido uma metodologia que consome menos memória e tempo computacional que utiliza a simulação de Monte Carlo (METROPOLIS et al., 1953) para a distribuição de carga. Este procedimento será posteriormente referido como a Tanford-Kirkwood com área acessível ao solvente por Monte Carlo ou simplesmente como o procedimento TKSA-MC.

Neste modelo simples, a proteína é representada por uma esfera com constante dielétrica baixa  $\epsilon = 4$  e raio  $b$  que representa o raio da proteína. A energia de interação é entre as cargas de unidade positivo colocados nos sítios de protonação de grupos ionizáveis. A energia de interação entre pares de resíduos  $i$  e  $j$  é descrito por

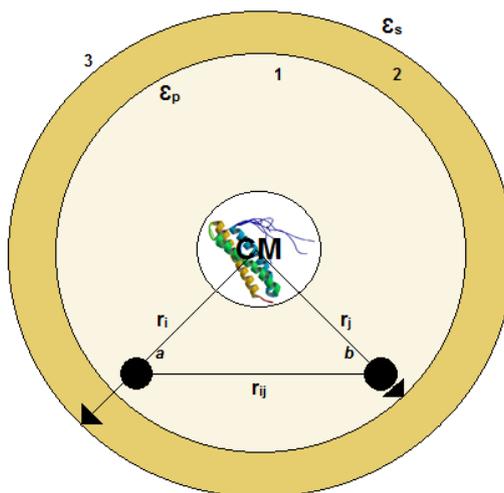
$$U_{ij}^{el} = e^2 \left( \frac{A_{ij} - B_{ij}}{2b} - \frac{C_{ij}}{2a} \right) (1 - SA_{ij}) \quad (2)$$

onde  $e$  é a unidade de carga,  $A_{ij}$ ,  $B_{ij}$  e  $C_{ij}$  são parâmetros obtidos a partir da solução Tanford-Kirkwood, que estão associados com a constante dielétrica da proteína e solvente, a distância entre as cargas  $e$ , o termo  $C_{ij}$ , a força iônica (TANFORD e KIRKWOOD et al., 1957; HAVRANEK et al., 1999; IBARRA-MOLERO et al., 1999).  $SA_{ij}$  é a média da área de superfície acessível da cadeia lateral do resíduo  $i$  e  $j$  e um íon é o raio de exclusão da esfera com  $a = b + 1,4 \text{ \AA}$  (média cátion e ânion raios típicos de um sal NaCl).

Na representação mostrada na Figura 3 a seguir, tem-se CM que indica o centro de massa da proteína,  $r_i$  e  $r_j$  que são as distância das cargas  $i$  e  $j$  até o centro de massa respectivamente,  $\epsilon_p$  é constante dielétrica da proteína e  $\epsilon_s$  é a constante dielétrica da solução.

Pode-se observar na representação também a descrição dos parâmetros  $A_{ij}$ ,  $B_{ij}$  e  $C_{ij}$  presentes na equação 2 que foi utilizado no modelo Tanford kirkwood para a realização do cálculo. O trabalho realizado por Kirkwood envolve um sistema composto por uma solução eletrolítica com uma constante dielétrica maior que envolve uma cavidade esférica com uma constante dielétrica menor e dentro dessa cavidade contém as cargas inseridas. Em nosso modelo de forma análoga ao

modelo de Kirkwood, pode-se considerar a proteína como uma cavidade esférica que apresenta uma baixa constante dielétrica  $\epsilon_p$  inserida em uma solução eletrolítica de constante dielétrica  $\epsilon_s$  (Figura 3). Esse sistema envolve 3 regiões, a região 1 será a cavidade esférica (proteína) de raio  $b$  e constante dielétrica  $\epsilon_p$ , a região 2 representa a área de exclusão de íons entre o raio  $a$  e o raio  $b$  e a região 3 representa a solução eletrolítica (solvente) (Figura 3).



**Figura 3:** Representação do modelo Tanford Kirkwood

A distribuição de carga da proteína é alterada em cada passo de Monte Carlo por titulação, em que os valores da valência dos resíduos são alterados ( $0 \leftrightarrow +1$ , para ARG / LYS / HIS,  $-1 \leftrightarrow 0$  para ASP / GLU). A aceitação das etapas de Monte Carlo seguem o critério de Metropolis considerando a variação de energia (ULLNER et al., 1996; De OLIVEIRA et al., 2014)

$$\Delta U = \Delta U^{sl} \pm K_b t \ln(10) (Ph - PK_a) \quad (3)$$

em que o sinal positivo é utilizado quando o resíduo é protonado e o sinal negativo é utilizado quando o resíduo é desprotonado durante a simulação. Os valores intrínsecos de pKa dos resíduos são: ASP = 4.0; GLU = 4.5; HIS = 6.3; LYS = 10.6; ARG = 12.0 (FERSHT et al., 1985; CREIGHTON et al., 1992). Foi realizada 105 passos de Monte Carlo para equilibração e 108 passos para a produção. A área de superfície acessível, foi calculada utilizando o software de superfície piloto (RECORD et al., 2002). Os diagramas de superfície foram preparados com o pacote Virtual Molecular Dynamics (VMD) e o servidor PDB2PQR (HUMPHREY et al., 1996; BAKER et al., 2001; DOLINSKY et al., 2007).

## 3 RESULTADOS E DISCUSSÕES

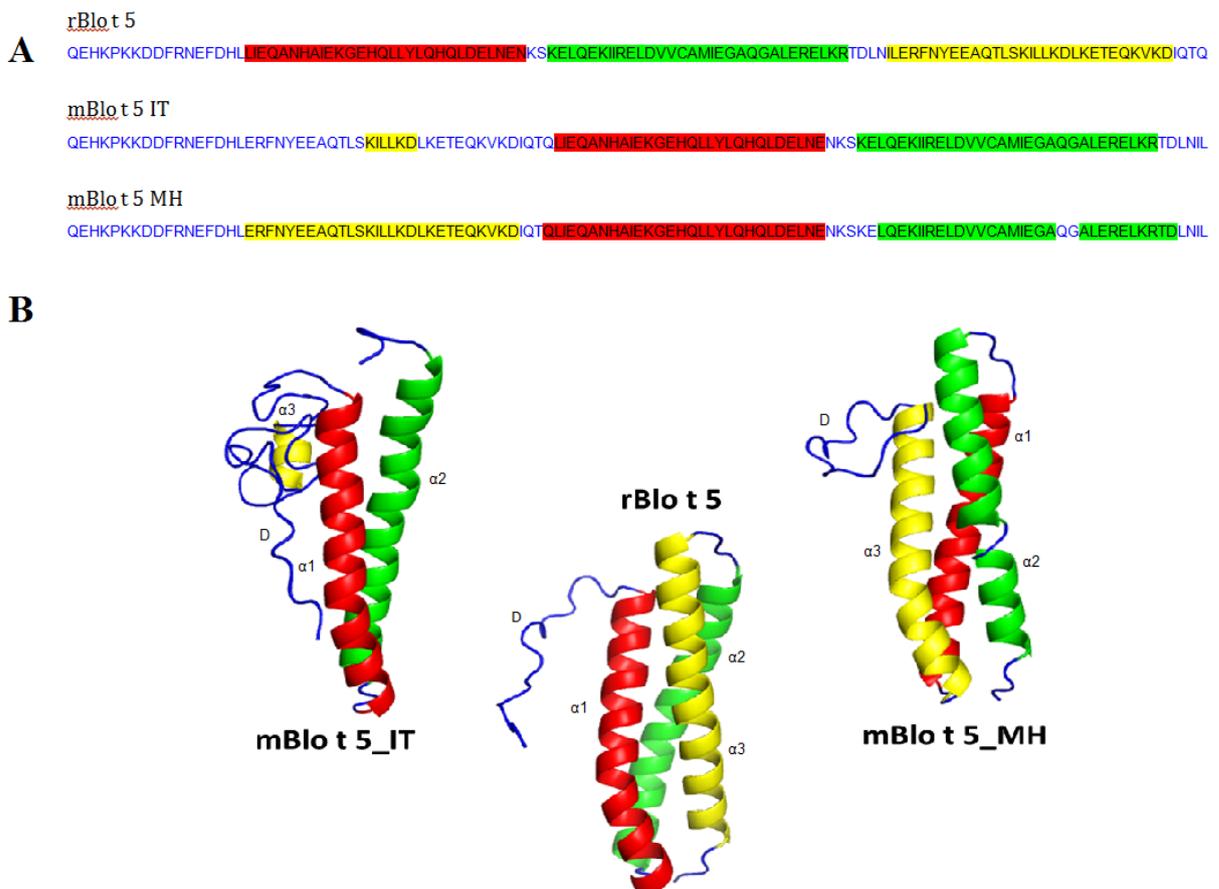
### 3.1 Estrutura da proteína *Blomia tropicalis* do grupo 5 (*Blo t 5*)

Os alérgenos que pertencem ao grupo 5 representam um importante grupo devido à alta capacidade de resposta dos pacientes a poeiras derivadas de extratos de ácaros contendo essas proteínas. Entretanto, ao passo que alguma reatividade cruzada existe entre os alérgenos do grupo 5, a ligação de IgE parece ser específico da espécie (CUI et al., 2013).

A *Blo t 5* adota uma estrutura na forma de feixe helicoidal que compreende três  $\alpha$ -hélices disposta num modo antiparalelo. A representação conformacional da *Blo t 5* é mostrada na Figura 1, onde a *Blo t 5* é representada na forma de um feixe de cadeia tripla em espiral na qual a primeira hélice ( $\alpha_1$ , contém os resíduos de 18 a 46) e a terceira hélice ( $\alpha_3$ , contém os resíduos de 84 a 113) estão em arranjo paralelo e a segunda hélice ( $\alpha_2$ , contém os resíduos de 49 a 77) está antiparalela com as outras duas hélices. Todas as três hélices de *Blo t 5* parecem ser ligeiramente inclinada, com ângulos inter-helicoidal de aproximadamente  $20^\circ$ . Cada hélice em *Blo t 5* contém um mínimo de 28 resíduos ou 4 repetições de septeto - um padrão de 7 resíduos indicados por (abcdefg) n), embora as posições de um septeto para a segunda hélice na estrutura são diferentes (NAIK et al., 2008). Na sequência da região flexível N-terminal, três  $\alpha$ -hélices antiparalelas formam um feixe de hélice separadas por curvas fechadas (CHAN et al., 2008).

As sequências e as estruturas das proteínas de ligação a IgE na forma nativa (*rBlo t 5*) e modificada (*mBlo t 5*) são mostradas na Figura 4. A sequência de aminoácidos da proteína *rBlo t 5* foi obtida no Genbank pelo número de acesso O96870, e a *mBlo t 5*, foi modelada por um deslocamento da hélice  $\alpha_3$  da *rBlo t 5* (indicado pela sequência amarela), onde na *rBlo t 5* encontra-se no final da proteína ( $\alpha_3$ ) e na *mBlo t 5\_IT* e *mBlo t 5\_MH* encontra-se no início da proteína ( $\alpha_1$ ) como mostrados na Figura 4A. Na Figura 4A, os números árabes ao longo das sequências referem-se à posição de aminoácido nas proteínas, o número total de aminoácidos está indicado no final das sequências (134 aminoácidos) e os 17 aminoácidos iniciais, em cada sequência, que correspondem ao peptídeo sinal, que não foram incluídos na análise de modelagem e simulação. A estrutura da proteína *rBlo t 5* (Protein Data Bank código 2JMH) consiste em três  $\alpha$ -hélices ( $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ ), formando um feixe de hélice, e um domínio desordenado no N-terminal (D) como mostrado na

Figura 4B. Duas estruturas semelhantes foram modeladas para *mBlo t 5* uma no servidor I-Tasser e outra no MHOLline, com as três  $\alpha$ -hélices rearranjadas tal como na Figura 4A ( $\alpha_3$ ,  $\alpha_1$ ,  $\alpha_2$ ) e o N-terminal (D), também em um estado desordenado (Fig. 4B). O valor do RMSD para os modelos gerados em relação à estrutura nativa *rBlo t 5* foram de 1,73 Å para *mBlo t 5\_IT* e de 6,19 Å para *mBlo t 5\_MH*. A proteína modificada *mBlo t 5* tem 72% de identidade da sequência em relação à proteína do tipo nativo ou recombinante, *rBlo t 5*.

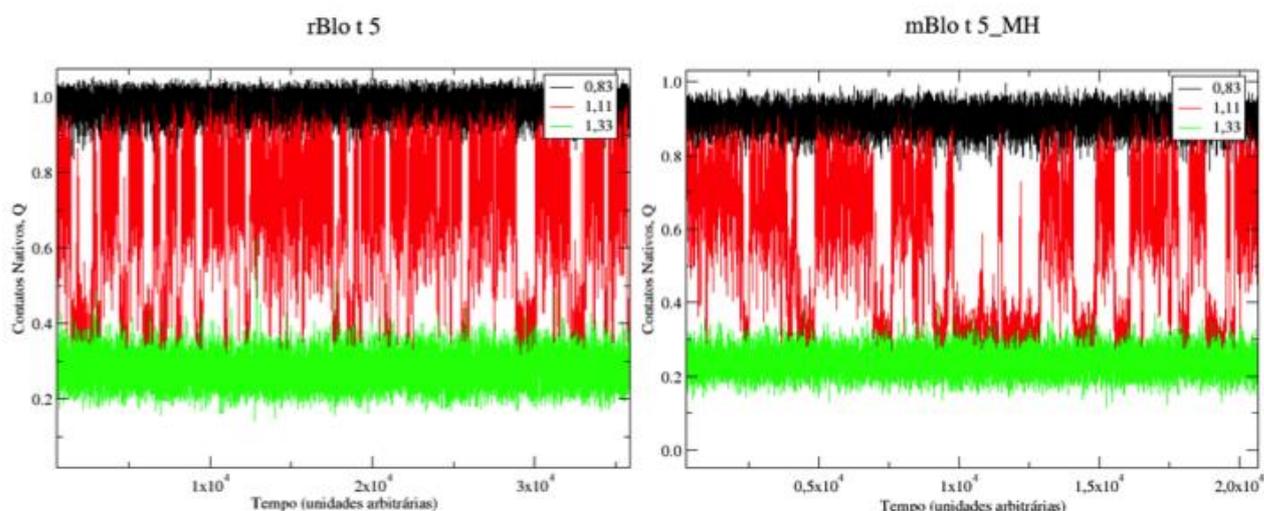


**Figura 4:** (A) Observa-se as sequências de aminoácidos do alérgeno *Blomia tropicalis* do grupo 5 recombinante, a sequência nativa (*rBlo t 5*), o alérgeno modificado construído por homologia no servidor I-Tasser (*mBlo t 5\_IT*) e, o alérgeno modificado construído por homologia no servidor MHOLline (*mBlo t 5\_MH*). Ambos os modelos gerados foram construído por um deslocamento da hélice  $\alpha_3$  da *rBlo t 5* (indicado pela sequência amarela). As letras coloridas representam os quatro domínios, uma desordem/desestruturado no domínio N seguido por três hélices (D,  $\alpha_1$ ,  $\alpha_2$  e  $\alpha_3$ ), rearranjados no alérgeno *rBlo t 5* para produzir a proteína modificada (D,  $\alpha_3$ ,  $\alpha_1$  e  $\alpha_2$ ). Em (B), observa-se a estrutura tridimensional da *rBlo t 5* (código no Protein Data Bank 2JMH, modelo 1) e os alérgenos modificados *mBlo t 5\_IT* e *mBlo t 5\_MH*. Os domínios da desordem são marcados em azul e as  $\alpha$ -hélices  $\alpha_1$ ,  $\alpha_2$  e  $\alpha_3$  são coloridos em vermelho, verde, e amarelo, respectivamente.

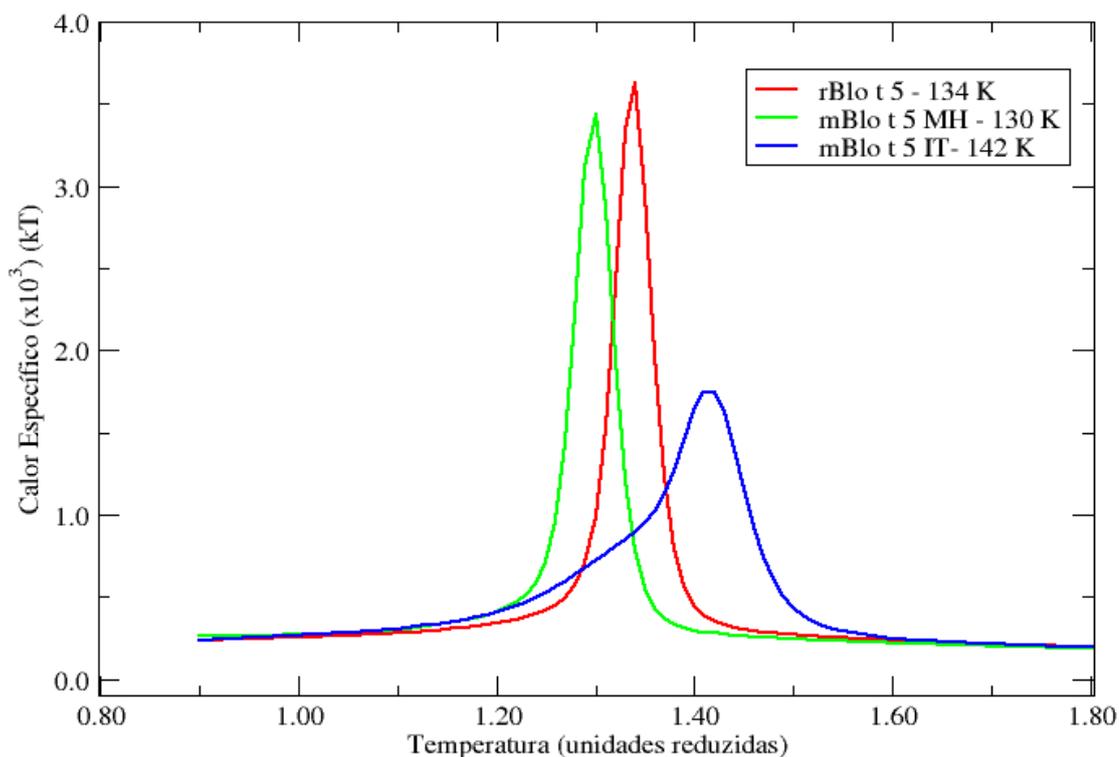
### 3.2 Análise Termodinâmica

Realizou-se o estudo estatístico das trajetórias oriundas das dinâmicas, utilizando o algoritmo WHAM (FERRENBURG et al., 1988; FERRENBURG et al., 1989; KUMAR et al., 1992; HESS et al., 2008) por meio do qual se obteve os contatos nativos (Q) em função do tempo medido em unidades reduzidas (Figura 5), calor específico das proteínas *rBlo t 5* e *mBlo t 5* em função da temperatura (Figura 6), assim como os perfis de energia livre em função da temperatura e em seguida em função do parâmetro de ordem (Figura 7).

Ao realizar as simulações computacionais com as proteínas *rBlo t 5* e *mBlo t 5*, observou-se sua forte dependência com a temperatura como mostrado na Figura 6. Nota-se que não houve diferença significativa na formação dos contatos nativos entre a proteína *rBlo t 5* e *mBlo t 5\_MH*, e isso demonstra uma estabilidade da proteína modificada. Com relação à *mBlo t 5\_IT*, não houve diferença significativa, pois, deve-se considerar que este modelo apresenta uma estrutura desordenada maior no início com relação à *rBlo t 5* e *mBlo t 5\_MH*. Foi também calculado o calor específico ( $C_v$ ) de ambas as proteínas (*mBlo t 5* e *rBlo t 5*) em função da temperatura. No que diz respeito às proteínas purificadas em solução, assim como os estados enovelado e desenovelado tem diferentes entalpias, e eles também têm diferentes capacidades de calor. Em relação a estabilidade da *Blo t 5*, pode-se notar na Figura 6, que mostra o calor específico de cada estrutura que não houve diferença significativa nos picos das temperaturas das estruturas *rBlo t 5* e *mBlo t 5\_MH*, pois, elas apresentaram um calor específico de 134K e 130K respectivamente. O pico da curva foi definido com a temperatura de enovelamento da proteína (CLEMENTI et al., 2000). Observa-se na Figura 6, que a temperatura de enovelamento é bem próxima para as duas estruturas *rBlo t 5* e *mBlo t 5\_MH*, e isso demonstra uma estabilização da estrutura. No modelo I-Tasser o pico é menor e isso demonstra que este modelo se enovela numa faixa de temperatura maior, mas isso deve-se ao fato da não modelagem de uma das hélices que foi gerado pelo servidor como estrutura desordenada da proteína, sendo assim ela se enovela com uma temperatura maior. Esses dados corroboram com os dados experimentais realizados em que após a modificação da *Blo t 5* também apresenta estabilização na sua estrutura, resultados bem similares do encontrado por meio das simulações.



**Figura 5:** Contatos nativos ( $Q$ ) em função do tempo computacional em unidades reduzidas. Em preto ( $T = 0,83$ ), a proteína encontra-se enovelada (estado nativo), em vermelho ( $T = 1,11$ ) esta em seu estado de transição. Em verde ( $T = 1,33$ ), encontra-se desenovelada. Temperatura em unidades reduzidas.



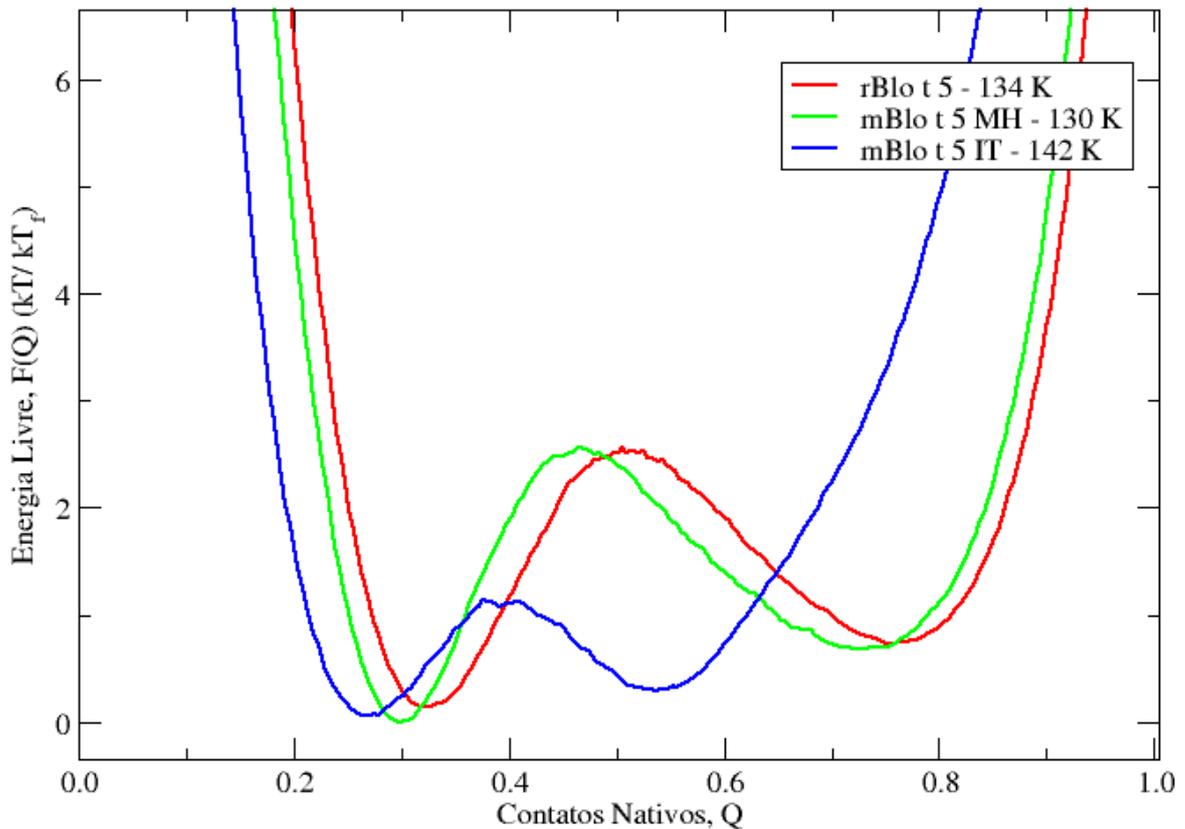
**Figura 6:** Calor específico da estrutura *mBlo t 5\_MH* (verde), *mBlo t 5\_IT* (azul) e *rBlo t 5* (vermelho).

### 3.2.2 Energia livre versus contatos nativos ( $Q$ ) para o modelo C-alfa

Na reação de enovelamento de proteínas é empregado a coordenada de reação  $Q$ , que corresponde a fração de contatos nativos formados na estrutura. Para monitorar a termodinâmica do sistema, deve-se agrupar as configurações obtidas durante uma simulação como uma função de  $Q$ . Sendo assim, o parâmetro  $Q$  representa a quantidade de contatos correspondentes na forma nativa, demonstrando-se como uma boa medida do grau de semelhança com a estrutura nativa (CLEMENTI et al., 2000).

Nota-se que, por meio do perfil de energia livre em função de  $Q$ , se torna de fácil localização os agrupamentos desenovelados, enovelados e o estado de transição, como é mostrado na Figura 7. Por estes modelos considerarem as sequências não-frustradas, eles podem não gerar as mesmas energias das proteínas reais, assim como o valor das alturas de barreira e a estabilidade dos estados intermediários; porém há uma ótima concordância com resultados experimentais (CLEMENTI et al., 2000).

As energias livres são medidas em unidades de  $k_B T_f$ , em que  $T_f$  é a temperatura de enovelamento. Para um rápido enovelamento, mesmo que a barreira termodinâmica inerente é baixa, ou similar, a energia térmica  $k_B T$  sobre a cinética pode ser significativa.



**Figura 7:** Os perfis de energia livre em função da coordenada de reação fração de contatos nativos ( $Q$ ) na temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo C- $\alpha$ .  $T_f$  foi definida como a temperatura no pico de calor específico de cada proteína simulada. Ambas as proteínas estão no estado desenovelado quando  $Q \sim 0,3$  e no estado enovelado quando pelo  $Q \sim 0,8$ , com exceção da *mBlo t 5\_IT* que alcança o seu estado enovelado quando  $Q \sim 0,55$ . A proteína modificada *mBlo t 5\_MH* tem a mesma estabilidade térmica e um perfil de energia livre muito semelhante termodinamicamente como o da proteína recombinante nativa *rBlo t 5*.

As Figuras 6 e 7 mostram as simulações do modelo baseado em estrutura (SBM<sup>3</sup>) utilizando o modelo C $\alpha$ , para a análise cinética e termodinâmica. Os resultados com o modelo C $\alpha$  da Figura 6 apresenta calor específico ( $C_v(t)$ ) e da Figura 7 os perfis termodinâmicos de energia livre em função da coordenada de reação ( $F(Q)$ ) em  $T_f$ . A reação de enovelamento/desenovelamento para *rBlo t 5* e *mBlo t 5\_MH* foi monitorado pelo parâmetro de ordem  $Q$  com a forma enovelada ( $Q \sim 0,8$ ) e populações desenoveladas ( $Q \sim 0,3$ ), e separados pela barreira do estado de transição em ( $Q \sim 0,5$ ). Enquanto que para *mBlo t 5\_IT* a reação para a forma enovelada foi ( $Q \sim 0,55$ ) e desenoveladas ( $Q \sim 0,27$ ) separados pelo estado de transição em ( $Q \sim 0,4$ ). Na Figura 6, apesar do fato de que a temperatura de transição de enovelamento para o estado desenovelado ( $T_f$ ), a temperatura no pico

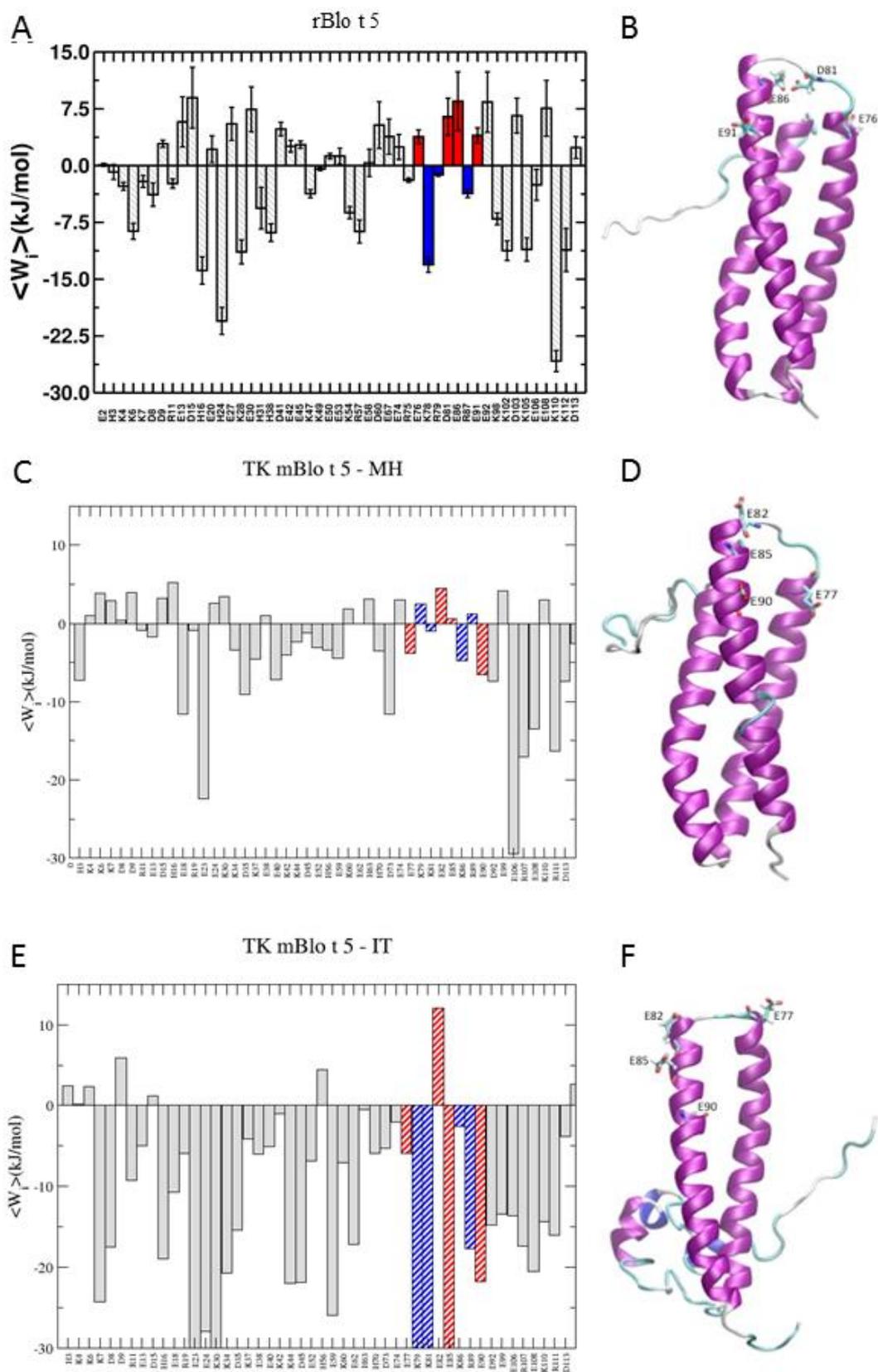
<sup>3</sup> Do inglês Structure-based model

---

de  $C_v(t)$ , de *rBlo t 5* e *mBlo t 5\_MH* serem ligeiramente diferentes, ambas as proteínas mostraram-se duas proteínas estáveis, com um perfil muito semelhante e uma estabilidade térmica de  $\Delta F = 2.2kT$  obtido a partir de simulações e para o modelo *mBlo t 5\_IT*, essa estabilidade reduziu-se pela metade em torno de  $\Delta F = 1.1kT$ . Esta é uma indicação de que a sequência modelada *mBlo t 5\_MH* tem uma estrutura tridimensional estável, possivelmente, uma proteína funcional, devido à mesma característica termodinâmica da *rBlo t 5*.

### 3.3 Otimização de interação eletrostática dos alérgenos recombinantes

Chan e colaboradores realizam mutações pontuais na estrutura nativa de *Blo t 5*, alterando os resíduos ionizáveis na região do epitopo por meio de um aminoácido neutro (alanina) (CHAN et al., 2008). Alguns mutantes dos alérgenos foram capazes de reduzir a ligação para IgE e também reduziram a resposta ao teste subcutâneo. É interessante observar nos cálculos eletrostáticos que estes efeitos foram mais significativos quando a mutação foi feita nos resíduos que mostram uma energia eletrostática média positiva (instáveis) de interação com o resto dos resíduos carregados determinado pelo nosso método. Os resíduos específicos para ligação nos epitopos da IgE carregados nos alérgenos nativos (Glu76, Asp81, Glu86 e Glu91), determinados por RMN de acordo com o estudo de (CHAN et al., 2008), mostrou a estabilização (energias negativas) da interação eletrostática no epitopo de ligação do novo alérgeno devido a inserção de uma carga positiva. Nesse estudo foram realizados utilizando todos os átomos da proteína e não somente os  $C\alpha$ .



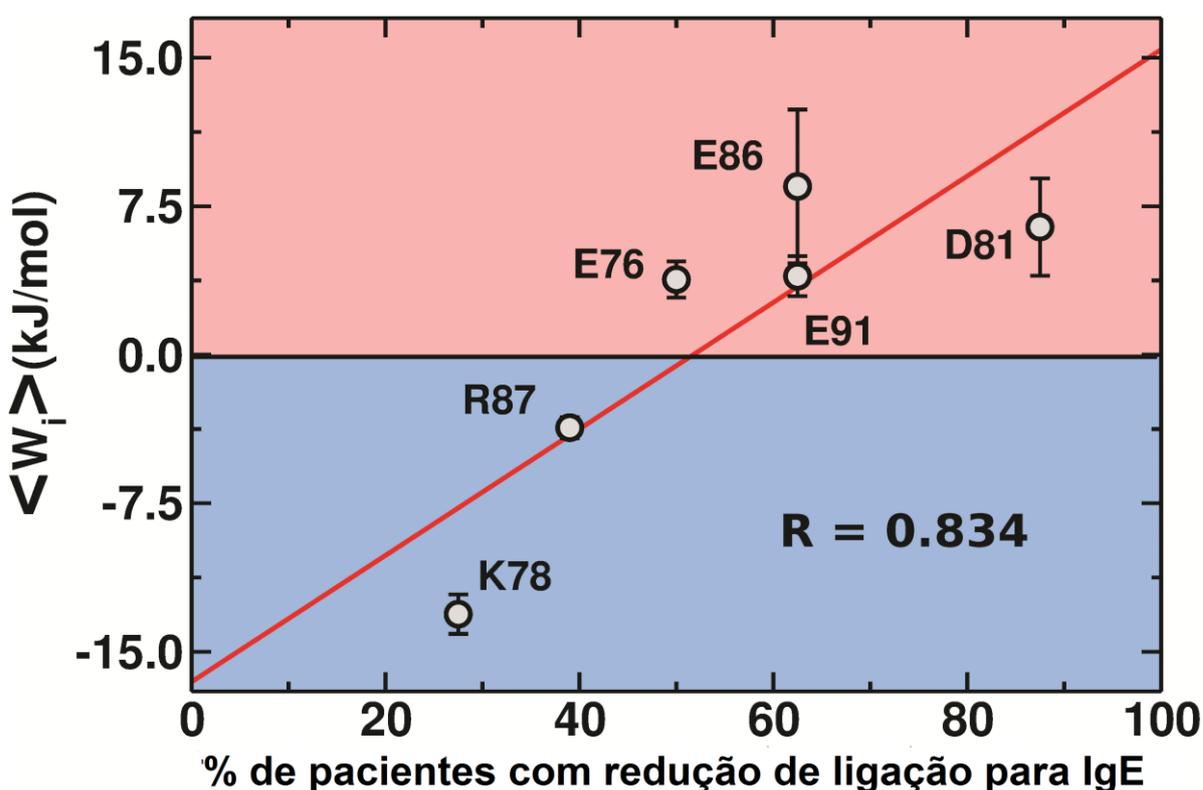
**Figura 8:** Otimização de interação eletrostática calculado na proteína nos resíduos carregados pelo modelo de Acessibilidade Superfície Tanford-Kirkwood implementados usando simulações de Monte Carlo (TKSA-MC).

A contribuição da energia eletrostática para cada resíduo ionizável dos alérgenos *rBlo t 5*, *mBlo t 5\_IT* e *mBlo t 5\_MH* foi investigada usando as simulações TKSA-MC. Na Figura 8A, a região do epitopo de ligação da *rBlo t 5* (do resíduo de 76 para 91), a média da interação carga-carga de energia eletrostática dos resíduos Glu76, Asp81, Glu86 e Glu91 tem um valor de energia positiva, o que significa que a contribuição eletrostática devido a estes resíduos não ser energeticamente favorável se torna uma região menos estável em comparação com o resto da proteína (barras vermelhas na Figura 8A). Esta instabilidade está associada com a proximidade dos grupos ionizáveis negativos. A distância média entre os grupos com carga negativa é 12,88 Å, enquanto que nos grupos positivos a média é de 30,60 Å. Esta diferença é causada principalmente por os resíduos Glu76, Glu86 e Asp81, que estão mais próximos entre si do que os outros resíduos ionizáveis no epitopo de ligação. Esta concentração de carga negativa também provoca um potencial eletrostático negativo nesta região (diagrama de superfície na Figura 8B). A mesma análise eletrostática foi aplicada no alérgeno *mBlo t 5*. A energia eletrostática da região do epitopo do *mBlo t 5* modificado para os dois modelos IT e MH (alinhado com *rBlo t 5*) mostrou-se mais estável do que a região do epitopo do *rBlo t 5* (Figura 8C e 8E). Esta estabilização da energia eletrostática em *mBlo t 5* está associado com a inclusão de um resíduo positivo na região do epitopo (Figura 8D e 8F), o que provoca uma distribuição de carga mais homogênea e um potencial eletrostático mais positivo quando comparado com o epitopo de ligação da *rBlo t 5*. Isto sugere que o aumento de estabilidade e a diminuição da energia eletrostática média na região do epitopo de *mBlo t 5* pode ser uma das razões do aumento de antigenicidade para a IgG4 e a diminuição para ligação de IgE em relação ao alérgénio nativo *rBlo t 5*.

### 3.4 Correlação das predições teóricas com os dados experimentais

Existe uma forte correlação entre a redução da ligação de IgE pelas mutações pontuais em *rBlo t 5* (dados a partir da experiência de RMN CHAN et al., 2008) e o valor médio da energia eletrostática ( $\langle W \rangle$ ) de cada resíduo ionizável desenvolvido neste estudo (Figura 9). A percentagem de redução da ligação de IgE é promovida por cada uma das seis mutações nos resíduos do epitopo entre oito pacientes (apenas as mutações induzidas em que mais do que 20% de redução da

ligação de IgE foram consideradas para análise). A correlação  $R = 0,834$  sugere que a substituição dos resíduos que possuem uma interação eletrostática favorável, não sendo um resíduo neutro, aumenta a estabilidade eletrostática na região do epitopo e induz a redução da interação entre *rBlo t 5* e IgE.



**Figura 9:** Correlação entre os dados teóricos e experimentais. Média da interação carga-carga de energia eletrostática  $\langle W_i \rangle$  dos resíduos de epitopos da *rBlo t 5* em função da porcentagem de pacientes com redução de ligação a IgE induzida pela mutação destes resíduos carregados para alanina (dados experimentais extraídos de CHAN et al., 2008). Os quatro principais resíduos de ligação a IgE (E76, D81, E86 e E91) são carregados negativamente e mostram uma interação de energia eletrostática desfavorável ( $\langle W_i \rangle > 0$ ) opondo-se à ligação de menores resíduos (K78 e R87), que são carregados positivamente com uma interação de energia eletrostática favorável ( $\langle W_i \rangle < 0$ ). O ajuste de correlação linear de cálculo teórico aos dados experimentais mostra uma forte correlação positiva com  $R = 0,83$  (ajuste linear obtido pela aplicação da correlação de Pearson). A energia é expressa em unidades de kJ/mol e as barras de erro são o desvio padrão da média das simulações TKSA-MC.

### 3.5 Discussões

Em primeiro lugar, foram caracterizados o estudo termodinâmico e a cinética por meio da modificação do alérgeno *mBlo t 5*. Foi certificado que o alérgeno projetado com as modificações pelo servidor MHOLline tem as mesmas características globais da proteína no seu estado nativo: apresentou um perfil quase idêntico de energia livre e termoestabilidade ( $\Delta F = 2.2$  kT), enquanto o alérgeno projetado pelo servidor I-Tasser apresentou uma redução de 50% em relação ao perfil de energia livre e termoestabilidade ( $\Delta F = 1.1$  kT). Esses resultados confirmam que o alérgeno modificado pelo MHOLline tem o seu grupo helicoidal dinamicamente estável e o epitopo de ligação altamente exposto ao solvente devido ao mesmo perfil de flutuação apresentado para *rBlo t 5*, que corrobora com a descrição original de dinâmica de proteínas para o alérgeno conforme estudos de RMN (CHAN et al., 2008; NAIK et al., 2008).

Em segundo lugar, as interações eletrostáticas de resíduos de carga localizados no epitopo e no restante da proteína modificada foram otimizadas. Trabalhos anteriores mostram que os resíduos específicos carregados localizados na superfície dos alérgenos HDM desempenham um papel crucial na interação entre o alérgeno-anticorpo (CHAN et al., 2006; CHAN et al., 2008; TAN et al., 2012). Neste estudo, foi estabilizado (menores valores de energia negativa) as interações eletrostáticas entre cada resíduo ionizável na *mBlo t 5* modificado em relação ao *rBlo t 5*.

A correlação apresentada de  $R = 0.83$  é considerada como uma boa validação do modelo computacional de como a simulação computacional e experimental corroboram, considerando que é um modelo simplificado de baixo custo computacional.

---

## 4 CONCLUSÕES

Conclui-se que a simulação por dinâmica molecular é capaz de detalhar os mecanismos dos processos biológicos envolvidos. Neste estudo, caracterizamos um novo protocolo para avaliar os epitopos projetados por computador por meio de modelos simplificados que podem levar a um aumento da antigenicidade do alérgeno para o isotipo IgG reduzindo a ligação e a consequente alergenicidade para o isotipo IgE.

A correlação apresentada é uma boa validação do nosso modelo de como a simulação experimental e computacional convergiram em grande acordo. A estabilização da energia eletrostática na região do epitopo, bem como no resto da *mBlo t 5*, sugere uma possível explicação para a mudança de antigenicidade para o isotipo IgG4 quando comparado com o *rBlo t 5*, como descrito no ensaio experimental (ÁVILA et al., 2012). Sendo assim, o modelo baseado em estrutura, juntamente com a simulação TKSA-MC parece ser um protocolo poderoso e promissor para investigar no futuro mutações pontuais *in silico* e propor novos candidatos em estudos de imunoterapia alérgeno específicos. Os resultados teóricos corroboraram com os dados experimentais, o que mostra que o alérgeno modificado *rBlo t 5* tem uma mudança de isotipo para a antigenicidade do anticorpo IgG4, tornando assim a sua aplicação adequada como candidato potencial em estudos imunoterapia com alérgenos específicos.

Outros estudos teóricos, estudos experimentais e clínicos são necessários para avaliar a atividade potencial e funcional hipoalergênico deste alérgeno modificado (*mBlo t 5*) em pacientes atópicos com sensibilidade adquirida ao *Blomia tropicalis* após estimulação *in vitro* com *rBlo t 5* e *mBlo t 5*.

Por fim, o Anexo A apresenta o trabalho em andamento sobre o estudo de outros alérgenos (*Blo t 12*, *Blo t 21*, *Der p 1*, *Der p 2*, *Der p 5*, *Der p 23* e *Der f 13*) com o propósito de verificar propriedades globais que definem a interação antígeno-anticorpo.

---

---

## REFERÊNCIAS

1. Alberts, B; Johnson, A; Lewis, J; Raff, M; Roberts, K; Walter, P. Molecular Biology of the Cell. New York: Garland science, Taylor & Francis Group, 2008. 1268 p.
2. Arruda LK, Vailes LD, PlattsMills TAE, FernandezCaldas E, Montealegre F, Lin KL, et al. Sensitization to *Blomia tropicalis* in patients with asthma and identification of allergen Blo t 5. American Journal of Respiratory and Critical Care Medicine. 1997;155(1):343-50.
3. Ávila, BGM. Perfil de reatividade de anticorpos IgE, IgG1 e IgG4 ao alérgeno recombinante Blo t 5 de *Blomia tropicalis* e sua forma modificada em pacientes atópicos e não atópicos. 2012. 87 f. Dissertação (Mestrado em Ciências Biológicas) - Universidade Federal de Uberlândia, Uberlândia, 2012.
4. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(18):10037-41.
5. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. MOLECULAR-DYNAMICS WITH COUPLING TO AN EXTERNAL BATH. Journal of Chemical Physics. 1984;81(8):3684-90.
6. Best RB, Chen YG, Hummer G. Slow protein conformational dynamics from multiple experimental structures: The helix/sheet transition of arc repressor. Structure. 2005;13(12):1755-63.
7. Best RB, Hummer G. Coordinate-dependent diffusion in protein folding. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(3):1088-93.
8. Brusica V, Petrovsky N. Bioinformatics for characterisation of allergens, allergenicity and allergic crossreactivity. Trends in Immunology. 2003;24(5):225-8.
9. Calderon MA, Linneberg A, Kleine-Tebbe J, De Blay F, Hernandez Fernandez de Rojas D, Virchow JC, et al. Respiratory allergy caused by house dust mites: What do we really know? Journal of Allergy and Clinical Immunology. 2015;136(1):38-48.

- 
10. Capriles PVSZ, Guimaraes ACR, Otto TD, Miranda AB, Dardenne LE, Degraeve WM. Structural modelling and comparative analysis of homologous, analogous and specific proteins from *Trypanosoma cruzi* versus *Homo sapiens*: putative drug targets for chagas' disease treatment. *Bmc Genomics*. 2010;11.
  11. Caraballo L, Mercado D, Jimenez S, Moreno L, Puerta L, Chua KY. Analysis of the cross-reactivity between BtM and Der p 5, two group 5 recombinant allergens from *Blomia tropicalis* (Bt) and *Dermatophagoides pteronyssinus* (Dp). *Journal of Allergy and Clinical Immunology*. 1997;99(1):1417-.
  12. Caravella JA, Wang D, Glaser SM, Lugovskoy A. Structure-Guided Design of Antibodies. *Current Computer-Aided Drug Design*. 2010;6(2):128-38.
  13. Chahine J, Oliveira RJ, Leite VBP, Wang J. Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *Proceedings of the National Academy of Sciences of the United States of America*. 2007;104(37):14646-51.
  14. Chan SL, Ong ST, Ong SY, Chew FT, Mok YK. Nuclear magnetic resonance structure-based epitope mapping and modulation of dust mite group 13 allergen as a hypoallergen. *Journal of Immunology*. 2006;176(8):4852-60.
  15. Chan SL, Ong TC, Gao YF, Tiong YS, Wang DY, Chew FT, et al. Nuclear magnetic resonance structure and IgE epitopes of Blo t 5, a major dust mite allergen. *Journal of Immunology*. 2008;181(4):2586-96.
  16. Chavez LL, Onuchic JN, Clementi C. Quantifying the roughness on the free energy landscape: Entropic bottlenecks and protein folding rates. *Journal of the American Chemical Society*. 2004;126(27):8426-32.
  17. Chruszcz M, Chapman MD, Vailes LD, Stura EA, Saint-Remy J-M, Minor W, et al. Crystal Structures of Mite Allergens Der f 1 and Der p 1 Reveal Differences in Surface-Exposed Residues that May Influence Antibody Binding. *Journal of Molecular Biology*. 2009;386(2):520-30.
  18. Chua KY, Cheong N, Kuo IC, Lee BW, Yi FC, Huang C-H, et al. The *Blomia tropicalis* allergens. *Protein and Peptide Letters*. 2007;14(4):325-33.
  19. Clark, D. P.. *Molecular Biology: Understanding the Genetic Revolution*. Oxford: Elsevier Academic Press, 2005.

- 
20. Clementi C, Nymeyer H, Onuchic JN. Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *Journal of Molecular Biology*. 2000;298(5):937-53.
21. Contessoto VG, Lima DT, Oliveira RJ, Bruni AT, Chahine J, Leite VBP. Analyzing the effect of homogeneous frustration in protein folding. *Proteins-Structure Function and Bioinformatics*. 2013;81(10):1727-37.
22. Cui Y. Structural biology of mite allergens. *Molecular Biology Reports*. 2013;40(1):681-6.
23. Cui Y. Immunoglobulin E-Binding Epitopes of Mite Allergens: From Characterization to Immunotherapy. *Clinical Reviews in Allergy & Immunology*. 2014;47(3):344-53.
24. de Oliveira VM, de Carvalho SJ. Adsorption of pH-responsive polyelectrolyte chains onto spherical macroions. *European Physical Journal E*. 2014;37(8).
25. de Queirs MGJ, Silva DO, Alves R, Chiba HF, de Amaral VBS, de Almeida KC, et al. Mite-specific immunotherapy using allergen and/or bacterial extracts in atopic patients in Brazil. *Journal of Investigational Allergology and Clinical Immunology*. 2008;18(2):84-92.
26. de Sousa Moreira PF, Gangl K, Machado Vieira FdA, Ynoue LH, Linhart B, Flicker S, et al. Allergen Microarray Indicates Pooideae Sensitization in Brazilian Grass Pollen Allergic Patients. *Plos One*. 2015;10(6).
27. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research*. 2007;35:W522-W5.
28. FernandezCaldas E. Mite species of allergologic importance in Europe. *Allergy*. 1997;52(4):383-7.
29. Fernandezcaldas E, Puerta L, Mercado D, Lockey RF, Caraballo LR. Mite fauna, der-p-i, der-f-i and blomia-tropicalis allergen levels in a tropical environment. *Clinical and Experimental Allergy*. 1993;23(4):292-7.

- 
30. Ferrenberg AM, Swendsen RH. New monte-carlo technique for studying phase-transitions. *Physical Review Letters*. 1988;61(23):2635-8.
  31. Ferrenberg AM, Swendsen RH. Optimized monte-carlo data-analysis. *Physical Review Letters*. 1989;63(12):1195-8.
  32. Fersht AR. Characterizing transition-states in protein-folding - an essential step in the puzzle. *Current Opinion in Structural Biology*. 1995;5(1):79-84.
  33. Fukunaga A, Tsumoto K. Improving the affinity of an antibody for its antigen via long-range electrostatic interactions. *Protein Engineering Design & Selection*. 2013;26(12):773-80.
  34. Gribenko AV, Patel MM, Liu J, McCallum SA, Wang C, Makhatadze GI. Rational stabilization of enzymes by computational redesign of surface charge-charge interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(8):2601-6.
  35. Hakkaart GAJ, Aalberse RC, van Ree R. Epitope mapping of the house-dust-mite allergen Der p 2 by means of site-directed mutagenesis. *Allergy*. 1998;53(2):165-72.
  36. Havranek JJ, Harbury PB. Tanford-Kirkwood electrostatics for protein modeling. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96(20):11145-50.
  37. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*. 2008;4(3):435-47.
  38. Holgate ST, Polosao R. Treatment strategies for allergy and asthma. *Nature Reviews Immunology*. 2008;8(3):218-30.
  39. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics & Modelling*. 1996;14(1):33-8.
  40. Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry*. 1999;38(25):8138-49.

- 
41. Ibarra-Molero B, Sanchez-Ruiz JM. Genetic algorithm to design stabilizing surface-charge distributions in proteins. *Journal of Physical Chemistry B*. 2002;106(26):6609-13.
  42. Jorgensen WL, Maxwell DS, TiradoRives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*. 1996;118(45):11225-36.
  43. Kosmoliaptsis V, Dafforn TR, Chaudhry AN, Halsall DJ, Bradley JA, Taylor CJ. High-resolution, three-dimensional modeling of human leukocyte antigen class I structure and surface electrostatic potential reveals the molecular basis for alloantibody binding epitopes. *Human Immunology*. 2011;72(11):1049-59.
  44. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules .1. The method. *Journal of Computational Chemistry*. 1992;13(8):1011-21.
  45. Lu Q, Lu HP, Wang J. Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. *Physical Review Letters*. 2007;98(12).
  46. Lu Q, Wang J. Single molecule conformational dynamics of adenylate kinase: Energy landscape, structural correlations, and transition state ensembles. *Journal of the American Chemical Society*. 2008;130(14):4772-83.
  47. Makhatadze GI, Loladze VV, Gribenko AV, Lopez MM. Mechanism of thermostabilization in a designed cold shock protein with optimized surface electrostatic interactions. *Journal of Molecular Biology*. 2004;336(4):929-42.
  48. Mandelli F, Franco Cairo JPL, Citadini APS, Buechli F, Alvarez TM, Oliveira RJ, et al. The characterization of a thermostable and cambialistic superoxide dismutase from *Thermus filiformis*. *Letters in Applied Microbiology*. 2013;57(1):40-6.
  49. Matouschek A, Kellis JT, Serrano L, Fersht AR. Mapping the transition-state and pathway of protein folding by protein engineering. *Nature*. 1989;340(6229):122-6.
  50. Matthew JB, Gurd FRN. Calculation of electrostatic interactions in proteins. *Methods in Enzymology*. 1986;130:413-36.

- 
51. Matthew JB, Gurd FRN. Stabilization and destabilization of protein-structure by charge interactions. *Methods in Enzymology*. 1986;130:437-53.
52. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*. 1953;21(6):1087-92.
53. Mohamadi F, Richards NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, et al. MacroModel - an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *Journal of Computational Chemistry*. 1990;11(4):440-67.
54. Mueller GA, Randall TA, Glesner J, Pedersen LC, Perera L, Edwards LL, et al. Serological, genomic and structural analyses of the major mite allergen Der p 23. *Clinical and Experimental Allergy*. 2016;46(2):365-76.
55. Naganathan AN. A Rapid, Ensemble and Free Energy Based Method for Engineering Protein Stabilities. *Journal of Physical Chemistry B*. 2013;117(17):4956-64.
56. Naik MT, Chang C-F, Kuo IC, Kung CCH, Yi F-C, Chua K-Y, et al. Roles of structure and structural dynamics in the antibody recognition of the allergen proteins: An NMR study on *Blomia tropicalis* major allergen. *Structure*. 2008;16(1):125-36.
57. Neves Petersen MT, Martel P, Petersen EI, Drablos F, Petersen SB. Surface and electrostatics of cutinases. *Lipases, Part a: Biotechnology*. 1997;284:130-+.
58. Noel JK, Whitford PC, Sanbonmatsu KY, Onuchic JN. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Research*. 2010;38:W657-W61.
59. Oliveira RJ, Whitford PC, Chahine J, Leite VBP, Wang J. Coordinate and time-dependent diffusion dynamics in protein folding. *Methods*. 2010;52(1):91-8.
60. Oliveira RJ, Whitford PC, Chahine J, Wang J, Onuchic JN, Leite VBP. The Origin of Nonmonotonic Complex Behavior and the Effects of Nonnative Interactions on the Diffusive Properties of Protein Folding. *Biophysical Journal*. 2010;99(2):600-8.

- 
61. Pereira EAL, Silva DAO, Cunha JP, Almeida KC, Alves R, Sung SJ, et al. IgE, IgG1, and IgG4 antibody responses to *Blomia tropicalis* in atopic patients. *Allergy*. 2005;60(3):401-6.
62. Permyakov SE, Makhatadze GI, Owenius R, Uversky VN, Brooks CL, Permyakov EA, et al. How to improve nature: study of the electrostatic properties of the surface of alpha-lactalbumin. *Protein Engineering Design & Selection*. 2005;18(9):425-33.
63. Platts-Mills TAE, Vervloet D, Thomas WR, Aalberse RC, Chapman MD. Indoor allergens and asthma: Report of the Third International Workshop. *Journal of Allergy and Clinical Immunology*. 1997;100(6):S2-S24.
64. Pramanik A, Hauf W, Hoffmann J, Cernescu M, Brutschy B, Braun V. Oligomeric Structure of ExbB and ExbB-ExbD Isolated from *Escherichia coli* As Revealed by LILBID Mass Spectrometry. *Biochemistry*. 2011;50(41):8950-6.
65. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic Acids Research*. 2005;33:W382-W8.
66. Snow CD, Sorin EJ, Rhee YM, Pande VS. How well can simulation predict protein folding kinetics and thermodynamics? *Annual Review of Biophysics and Biomolecular Structure*. 2005;34:43-69.
67. Sobolev V, Wade RC, Vriend G, Edelman M. Molecular docking using surface complementarity. *Proteins-Structure Function and Genetics*. 1996;25(1):120-9.
68. Socci ND, Onuchic JN, Wolynes PG. Diffusive dynamics of the reaction coordinate for protein folding funnels. *Journal of Chemical Physics*. 1996;104(15):5860-8.
69. Spector S, Wang MH, Carp SA, Robblee J, Hendsch ZS, Fairman R, et al. Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry*. 2000;39(5):872-9.
70. Strickler SS, Gribenko AV, Keiffer TR, Tomlinson J, Reihle T, Loladze VV, et al. Protein stability and surface electrostatics: A charged relationship. *Biochemistry*. 2006;45(9):2761-6.

- 
71. Tai H-Y, Zhou J-K, Chou H, Tam MF, Chen Y-S, Sheu S-Y, et al. Epitope Mapping and In Silico Characterization of Interactions between Der p 7 Allergen and MoAb WH9. *Plos One*. 2013;8(8).
72. Taketomi H, Ueda Y, Go N. Studies on protein folding, unfolding and fluctuations by computer-simulation .1. Effect of specific amino-acid sequence represented by specific inter-unit interactions. *International Journal of Peptide and Protein Research*. 1975;7(6):445-59.
73. Tan KW, Jobichen C, Ong TC, Gao YF, Tiong YS, Wong KN, et al. Crystal Structure of Der f 7, a Dust Mite Allergen from *Dermatophagoides farinae*. *Plos One*. 2012;7(9).
74. Tan KW, Ong TC, Gao YF, Tiong YS, Wong KN, Chew FT, et al. NMR Structure and IgE Epitopes of Blo t 21, a Major Dust Mite Allergen from *Blomia tropicalis*. *Journal of Biological Chemistry*. 2012;287(41):34776-85.
75. Tanford C, Kirkwood JG. THEORY OF PROTEIN TITRATION CURVES .1. GENERAL EQUATIONS FOR IMPENETRABLE SPHERES. *Journal of the American Chemical Society*. 1957;79(20):5333-9.
76. Tsodikov OV, Record MT, Sergeev YV. Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *Journal of Computational Chemistry*. 2002;23(6):600-9.
77. Ullner M, Jonsson B. A Monte Carlo study of titrating polyelectrolytes in the presence of salt. *Macromolecules*. 1996;29(20):6645-55.
78. Valenta R, Niespodziana K, Focke-Tejkl M, Marth K, Huber H, Neubauer A, et al. Recombinant allergens: What does the future hold? *Journal of Allergy and Clinical Immunology*. 2011;127(4):860-4.
79. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*. 2005;26(16):1701-18.
80. Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J, Rousseau F. A graphical interface for the FoldX forcefield. *Bioinformatics*. 2011;27(12):1711-2.
81. Wang J, Oliveira RJ, Chu X, Whitford PC, Chahine J, Han W, et al. Topography of funneled landscapes determines the thermodynamics and kinetics of

---

protein folding. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(39):15763-8.

82. Whitford PC, Miyashita O, Levy Y, Onuchic JN. Conformational transitions of adenylate kinase: Switching by cracking. *Journal of Molecular Biology*. 2007;366(5):1661-71.

83. Whitford PC, Noel JK, Gosavi S, Schug A, Sanbonmatsu KY, Onuchic JN. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins-Structure Function and Bioinformatics*. 2009;75(2):430-41.

84. Xu W, Lai Z, Oliveira RJ, Leite VBP, Wang J. Configuration-Dependent Diffusion Dynamics of Downhill and Two-State Protein Folding. *Journal of Physical Chemistry B*. 2012;116(17):5152-9.

85. Yang S, Onuchic JN, Levine H. Effective stochastic dynamics on a protein folding energy landscape. *Journal of Chemical Physics*. 2006;125(5).

86. Yong T-S, Jeong KY. Household Arthropod Allergens in Korea. *Korean Journal of Parasitology*. 2009;47:S143-S53.

87. Zhang D, Chen C-F, Zhao B-B, Gong L-L, Jin W-J, Liu J-J, et al. A Novel Antibody Humanization Method Based on Epitopes Scanning and Molecular Dynamics Simulation. *Plos One*. 2013;8(11).

88. Zuckerman DM. Simulation of an ensemble of conformational transitions in a united-residue model of calmodulin. *Journal of Physical Chemistry B*. 2004;108(16):5127-37.

---

## APÊNDICE A

### Trabalho em andamento

Sabe-se que existem diferenças sutis na composição e distribuição de cargas por conta do epitopo que demonstra uma diferença na especificidade de ligação a IgE entre os diferentes grupos de alérgenos em *B. tropicalis*, *D. pteronyssinus* e *D. farinae* do HDM. Devido a essas diferenças, está em andamento o estudo de outros alérgenos (*Blo t 12*, *Blo t 21*, *Der p 1*, *Der p 2*, *Der p 5*, *Der p 23* e *Der f 13*) com o propósito de verificar propriedades globais que definem a interação antígeno-anticorpo.

Muitos alérgenos provenientes de ácaros e outras substâncias contribuem promovendo vários sintomas alérgicos em humanos, incluindo a asma e a rinite como já vimos anteriormente no caso da *Blo t 5* (*Blomia tropicalis*) do grupo 5. Sendo assim, percebe-se que esses alérgenos pertencem a vários grupos de alérgenos e que podem ter características em comum. Uma das perspectivas futuras deste nosso estudo é estudar outros alérgenos que apresentam uma similaridade para a *Blo t 5* e uma proporção de ligação em IgE, aplicando a metodologia utilizada para a *Blo t 5*. Neste caso selecionamos alguns alérgenos que serão melhor descritos abaixo.

Alérgenos do GRUPO 1: Neste grupo, escolheu-se a *Dermatophagoides pteronyssinus* (*Der p 1*) que é um alérgeno que apresenta 50% de ligação à IgE. (CUI, 2013). Essa estrutura foi obtida no Protein Data Bank (PDB) pelo código: 2AS8, e contém 222 resíduos de aminoácidos.

Alérgenos do GRUPO 2: Neste grupo, também optou-se pela espécie *Dermatophagoides pteronyssinus* (*Der p 2*). É um alérgeno com 80% de ligação à IgE (CALDERON et al., 2015), obtida pelo PDB com o código 1KTJ, e composta por 129 resíduos de aminoácidos.

Alérgenos do GRUPO 5: Neste, trabalhou-se com duas espécies a *Blomia tropicalis* (*Blo t 5*) e a *Dermatophagoides pteronyssinus* (*Der p 5*), ambas apresentam uma identidade de 42% na sequência de aminoácidos. Este grupo apresenta uma grande importância clínica pelo fato da alta resposta alérgica de pessoas predispostas a estes alérgenos. A *Blo t 5* apresenta 90% de ligação à IgE enquanto que a *Der p 5* apresenta 72% de ligação. (CUI, 2013). A *Blo t 5* foi obtida do PDB (código 2JMH) e apresenta 117 resíduos de aminoácidos e a *Der p 5* foi

---

obtida pelo código 3MQ1 e apresenta um fragmento de resíduos de 32 a 132, totalizando 101 resíduos de aminoácidos. Essas duas são homologas e apresentam uma alta capacidade para reatividade cruzada.

Alérgeno do GRUPO 12: Neste grupo escolheu-se a *Blo t 12* por apresentar uma reatividade de ligação à IgE de 50%, (CHUA et al., 2007) e a mesma foi obtida através do código 2MFK, com 143 resíduos.

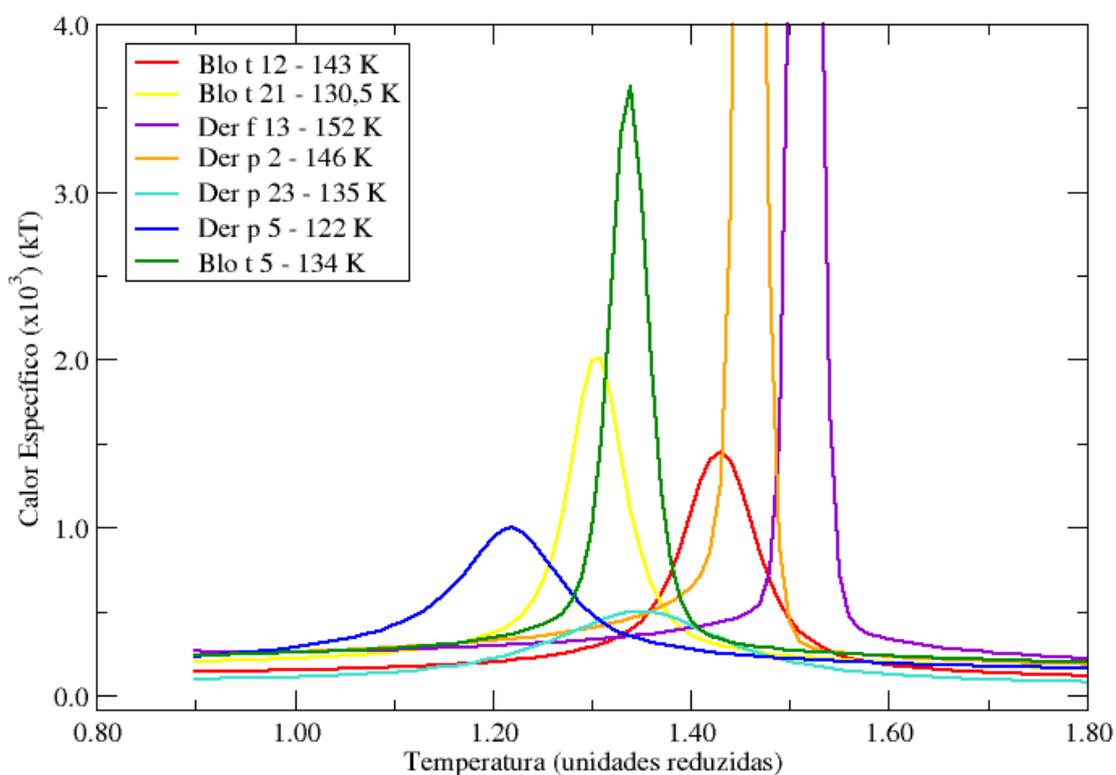
Alérgeno do GRUPO 13: A *Der f 13*, espécie *Dermatofagoides farinae* do grupo 13 obtida pelo código do PBD 2A0A, possui 131 resíduos e é um dos alérgenos que possui menos reatividade de ligação em IgE, que de acordo com CALDEIRA et al., 2015 é de 11%.

Alérgeno do GRUPO 21: A *Blo t 21* representa um novo grupo de alérgeno de grande importância clínica dentro da espécie *Blomia tropicalis*, ela apresenta uma capacidade de reatividade cruzada com *Blo t 5*, *Der p 5* e *Der f 21* e ainda apresenta mais que 75% de ligação à IgE, apresenta também uma moderada identidade da sequência com a *Blo t 5*, em torno de 40,7% . A estrutura foi obtida por meio do código 2LM9, apresenta-se um fragmento de 18 a 113, perfazendo um total de 96 resíduos de aminoácidos (TAN et al., 2012).

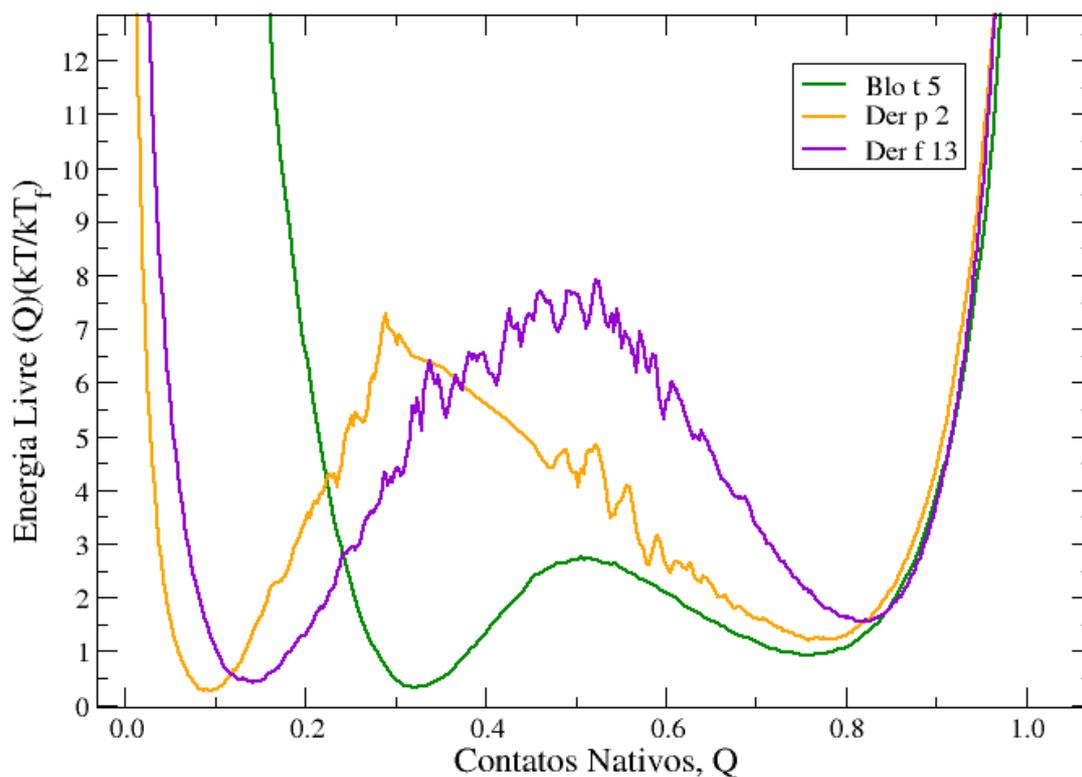
Alérgeno do GRUPO 23: Neste grupo, deu-se preferência pela *Der p 23* por apresentar uma grande reatividade de ligação para IgE, que apresenta 75% de ligação à IgE. A prevalência de quantidade de IgE específica é bastante elevada para a *Der p 23* e *Der p 2* em comparação com a expressão de outros alérgenos da espécie *Dermatophagoides pteronyssinus* (MUELLER et al., 2016). Este foi obtido pelo código PDB 4ZCE e apresenta um fragmento 43-90, num total de resíduos.

A análise termodinâmica das proteínas em estudo é apresentada nas Figuras 10 e 11.

Como continuação do trabalho, pretende-se calcular o perfil eletrostático com o método TKSA-MC e agrupar essas proteínas de acordo com suas características similares levando em conta a barreira de energia livre, interação carga-carga e reatividade com IgE, com o objetivo de compreender os mecanismos que levam a interação dos alérgenos com as imunoglobulinas.

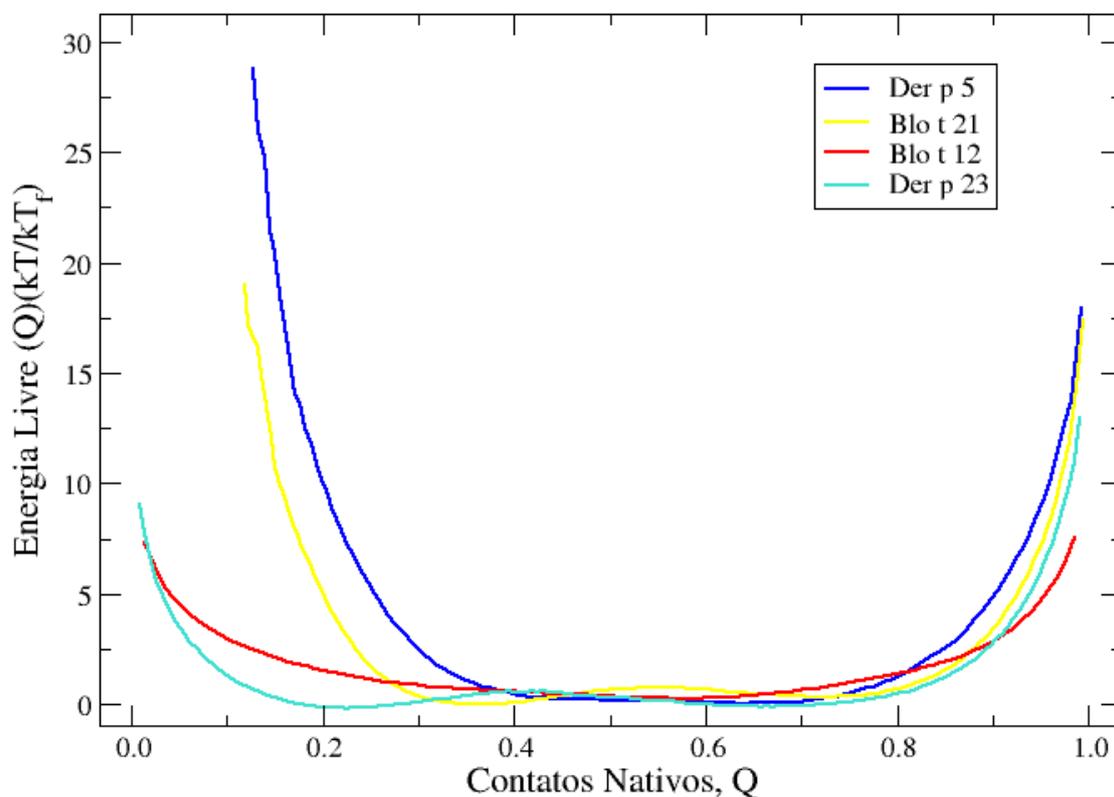


**Figura 10:** Calor específico das estruturas *Blo t 12* (vermelho), *Blo t 21* (amarelo), *Der f 13* (violeta), *Der p 2* (laranja), *Der p 23* (azul turquesa), *Der p 5* (azul) e *Blo t 5* (verde).



**Figura 11:** Os perfis de energia livre em função da coordenada de reação fração de contatos nativos ( $Q$ ) na temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo  $C\alpha$  para as proteínas *Blo t 5* (verde), *Der p 2* (laranja) e *Der f 13* (violeta).  $T_f$  foi definida através da temperatura no pico de calor específico de cada proteína simulada.

Ambas as proteínas estão no estado desenovelado quando  $Q \sim 0,1$  e no estado enovelado quando pelo  $Q \sim 0,8$ , com exceção da *Blo t 5* que alcança o seu estado desenovelado quando  $Q \sim 0,3$ . Essas são proteínas com barreira de energia livre no estado de transição.



**Figura 12:** Os perfis de energia livre em função da coordenada de reação fração de contatos nativos ( $Q$ ) na temperatura de enovelamento ( $T_f$ ) obtidos com as simulações em modelo C- $\alpha$  para as proteínas *Der p 5* (azul), *Blo t 21* (amarelo), *Blo t 12* (vermelho) e *Der p 23* (azul turquesa). Utilizou-se  $T_f$  como parâmetro definido através da temperatura no pico de calor específico de cada proteína simulada. Essas são proteínas com nenhuma ou baixíssima barreira de energia livre no estado de transição.

---