

Sidney Jurado de Carvalho

Estudo Metodológico dos Aspectos
Eletrostáticos da Interação
Ligante - Biomolécula



1910001289



Universidade Estadual Paulista "Júlio de Mesquita Filho"
Instituto de Biociências, Letras e Ciências Exatas
Campus de São José do Rio Preto

Estudo Metodológico dos Aspectos Eletrostáticos da Interação Ligante-Biomolécula

Sidney Jurado de Carvalho

Dissertação apresentada para obtenção do grau de
Mestre em Biofísica Molecular, área de concentração
em Biofísica Molecular. Instituto de Biociências, Letras
e Ciências Exatas da Universidade Estadual Paulista
"Júlio de Mesquita Filho" – UNESP.

Orientador: Prof. Dr. Fernando Luís Barroso da Silva



São José do Rio Preto - SP

2003

7.1289/04



Carvalho, Sidney Jurado de.

Estudo metodológico dos aspectos eletrostáticos da interação ligante-biomolécula / Sidney Jurado de Carvalho. – São José do Rio Preto : [s.n.], 2004

104 f. : il. ; 30 cm.

Orientador: Fernando Luís Barroso da Silva

Dissertação (mestrado) – Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas

1. Teoria molecular. 2. Termodinâmica estatística. 3. Físico-química. 4. Método de Monte Carlo. 5. Poisson-Boltzmann, Equação de. I. Silva, Fernando Luís Barroso da. II. Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 577.32

Agradecimentos

Muito obrigado aos meus pais, especialmente ao meu pai, por todo o apoio e incentivo durante a minha vida acadêmica e profissional. Também agradeço aos meus amigos, especialmente aos meus colegas de trabalho, por todo o apoio e incentivo durante a minha vida acadêmica e profissional.

Muito obrigado aos meus pais, especialmente ao meu pai, por todo o apoio e incentivo durante a minha vida acadêmica e profissional. Também agradeço aos meus amigos, especialmente aos meus colegas de trabalho, por todo o apoio e incentivo durante a minha vida acadêmica e profissional.

Muito obrigado aos meus pais, especialmente ao meu pai, por todo o apoio e incentivo durante a minha vida acadêmica e profissional. Também agradeço aos meus amigos, especialmente aos meus colegas de trabalho, por todo o apoio e incentivo durante a minha vida acadêmica e profissional.

Dedicado a todos aqueles que se emocionam com o entendimento da natureza.

Muito obrigado aos meus pais, especialmente ao meu pai, por todo o apoio e incentivo durante a minha vida acadêmica e profissional. Também agradeço aos meus amigos, especialmente aos meus colegas de trabalho, por todo o apoio e incentivo durante a minha vida acadêmica e profissional.

Agradecimentos

Muitas pessoas estiveram envolvidas, direta ou indiretamente, no desenvolvimento deste trabalho. Muitas dessas pessoas talvez não percebam o quão importante foram neste período vital de minha vida. Reconheço e agradeço o carinho, dedicação, paciência, e cada gesto que, por menor que fosse, tiveram sua contribuição neste momento.

Mais especificamente, gostaria de começar agradecendo aos meus amigos da época de graduação: Zé, André, Marco Aurélio, Chico, Juliana, Vanessa, Macuco, Ariano, Cristiano, Júlio, Luciano, Robson, Emerson, entre outros, cuja convivência tornou meu aprendizado mais agradável e proveitoso.

Gostaria de agradecer também aos meus mestres na graduação: Renato, Chico, André, Pablo, Denise, Sardela, Vicente, Humberto, Momotaro, Rosa, Caluzi e Grandini, cujos ensinamentos foram fundamentais para eu chegar até aqui.

Aos amigos que fiz durante este período no departamento de Física no campus da UNESP de São José do Rio Preto (Carlos, Ésio, Sabrina, Márcia, Glaucia, Fausto, Priscila, André, Diogo, Leandro, Luciana, Ricardo, Patrícia, Henrique, Nelson, Ana Helena, Helen, entre outros), gostaria de dizer “MUITÍSSIMO OBRIGADO”. Fui extremamente bem recebido e ajudado por todos durante meu início em Rio Preto. As discussões de física com o Carlos e o Ésio, embora um pouco escassas, foram extremamente proveitosas, bem como as dicas de Linux, tanto do Carlos como do Leandro, que sempre contribuíram muito para a estabilidade da minha máquina. Durante a fase de representante discente, o auxílio da Sabrina, Márcia e Priscila sempre foram extremamente valiosas, me ajudando a ver algumas questões de maneira mais sensata. Finalizando, agradeço aos companheiros de bagunça Carlos, Ésio, André, Leandro e Diogo, cuja companhia foi extremamente prazerosa e divertida (levando



em conta que eu sempre ganhava na sinuca).

Duas pessoas que merecem um agradecimento especial são meus companheiros de sala Luíz e Lessandra, as quais sempre me ajudaram, principalmente na minha chegada a Rio Preto, e que foram sempre muito atenciosos comigo.

Este trabalho não teria sido possível sem o ensinamento dos meus mestres da pós-graduação: Zé, João, Augusto, Elói e Márcio. Obrigado pela paciência e dedicação que sempre demonstraram. Gostaria de agradecer também a todos os docentes e funcionários (Marcelo, Chahine, Elso e Ilva), do departamento de física que sempre foram bastante prestativos.

Durante a minha jornada em Rio Preto, tive o privilégio de morar com pessoas maravilhosas, cuja convivência foi extremamente agradável e instrutiva. Aprendi com essas pessoas a ser mais tolerante, a ver a vida de maneira mais alegre, e a me apegar a coisas que realmente importam. Obrigado Deni, Denis, Joaquim, João, Ademar, Tiago, Zé Eduardo e Ricardo. Por muitas vezes, uma brincadeira de voçes me fez olhar para mim mesmo e descobrir como eu podia ser uma pessoa melhor. VALEU MESMO !!!

Todas as vitórias de uma pessoa, na verdade não são só dela, e sim de todos aqueles que dedicaram o seu tempo em fazer essa pessoa cada vez melhor. Neste intuito, algumas pessoas fizeram isso com extrema EXCELÊNCIA: Meu avô Emílio e minha avó Cida, bases sólidas da minha família. Minha mãe, cuja garra e fortaleza foi fundamental para que eu chagasse até aqui. Meu padrasto Juvenal, cuja presença é imprescindível na estabilidade de minha família. Meu tio Sid, que sempre me prestou o apoio necessário com paciência e afeto, principalmente nos períodos de revolta. Minha irmã Fhá, que sempre foi minha companheira nas dificuldades que a vida nos impôs desde pequenos. A todas essas pessoas, obrigado pela dedicação e principalmente paciência. Por muitas vezes fui rude, ingrato, insensível, mas a convivência com voçes é o que guardo de mais sagrado em minha vida.

Gostaria de agradecer também aos membros da banca do exame geral de qualificação: Prof. Dr. José Roberto Ruggiero, Prof. Dr. Augusto Agostinho Neto e Prof. Dr. João Ruggiero Neto, e da banca da defesa de mestrado: Prof. Dr. Hernan Chaimovich e Prof. Dr. Bernhard Joachim Mokross, pela paciência e dedicação na leitura do meu trabalho, bem como nas orientações dadas, extremamente úteis para o aperfeiçoamento do trabalho.

Muitas vezes, a vida nos traz pessoas que podem transformá-la completamente. Na minha vida, uma dessas pessoas é a mulher que fez com que momentos de intensa luta se transformassem em momentos de alegres desafios, e momentos de alegria, se transformasse em momentos de “extrema” alegria. Camila, obrigado pela paciência, carinho, compreensão, dedicação, e tudo mais que você meu deu de maneira extremamente espontânea e alegre. Com certeza, cada pedaço deste trabalho, cada vitória e cada momento de satisfação, não teria o mesmo gosto se não fosse ao seu lado.

Fernando, poucas vezes na vida encontrei pessoas tão especiais como você. Muitas vezes fiquei refletindo sobre minha incapacidade em lhe agradecer por toda a dedicação e paciência que teve na minha orientação. Quanto mais amadureço e mais convivo com você, mais percebo a importância que você tem para mim. **MUITÍSSIMO OBRIGADO** por acreditar em mim, pelos ensinamentos, por saber “puxar minha orelha” nas horas certas, por ser compreensivo, por me dar todo o apoio necessário em todas as situações, por me incentivar, enfim, muito obrigado por TUDO. Se algum dia conseguir ser um pouco do ser humano que você é, minha vida estará completa.

Gostaria de agradecer a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) pelo suporte financeiro, sem o qual este trabalho não poderia ser desenvolvido. Muito obrigado também ao acessor FAPESP por acreditar neste trabalho.

Sumário

Resumo	1
Abstract	2
Lista de Abreviaturas	3
Lista de Símbolos	4
Lista de Figuras	9
Lista de Tabelas	13
1 Introdução	14
1.1 Níveis de Detalhamento do Modelo	17
1.2 Solução do Modelo	20
2 Objetivos do Trabalho	23
3 O Problema	
A Interação Ligante-Biomolécula	24
3.1 Proteínas da Família da <i>Calmodulina</i>	28
3.1.1 Calbindina D_{9k}	30
3.1.2 Calmodulina	31



4	Modelo do Sistema	34
4.1	Modelo da Macromolécula	37
4.2	Campos de Força	39
5	Solução do Modelo	
	A Mecânica Estatística Clássica	43
5.1	Função Distribuição Radial	45
5.2	Potencial de Força Média	48
5.3	A Equação de Poisson-Boltzmann	49
5.3.1	Aproximação de Debye-Hückel	51
5.3.2	Energia Livre Eletrostática	53
5.3.3	O Modelo de Tanford-Kirkwood	56
5.4	O Método Monte Carlo Metrópolis	59
5.5	Potenciais Químicos	61
6	Resultados	64
6.1	Efeito da Concentração da Proteína	64
6.2	Aproximação de Campo Médio	71
6.3	Dependência do Campo de Força	73
7	Conclusões	79
	Referências Bibliográficas	81
	A Campos de Força para o Acetato	91
	B Problemas com a solução numérica da Equação de <i>Poisson-Boltzmann</i>	92
	C Programas Desenvolvidos	95

Resumo

A interação entre ligantes e macromoléculas tem importância fundamental em uma variedade de processos bioquímicos. Modelos baseados em interações eletrostáticas vêm se mostrando bastante eficientes na descrição desses processos, incluindo boa concordância com resultados obtidos experimentalmente. Esses modelos são basicamente resolvidos de duas formas: utilizando aproximações de campo médio, como a teoria de *Poisson-Boltzmann* e sua forma linear, equivalente à teoria de *Debye-Hückel*, ou através de simulações computacionais (por exemplo, o método *Monte Carlo*), onde o efeito de correlação iônica pode ser explicitamente considerado. Entretanto, alguns aspectos metodológicos carecem de análises mais sistemáticas, como a arbitrariedade na escolha dos campos de força para modelagem das cargas atômicas situadas nas macromoléculas. Portanto, propomos nesse trabalho investigar aspectos metodológicos para as abordagens teóricas usadas no estudo dos aspectos eletrostáticos da interação ligante-biomolécula, e a aplicação destas ferramentas para sistemas de interesse biofísico, como as proteínas da família da *Calmodulina* que tem o cálcio como ligante.



Abstract

Ligand-protein interaction is important in biology. Models based on electrostatic interactions have provided a acceptable description and good agreement with experiments. These models are solved either by mean field approximations (as given by the *Poisson-Boltzmann* theory and its linear version – the *Debye-Hückel* theory) or computational simulations (e.g. *Monte Carlo*), where ion-ion correlations are explicitly taken into account. Nevertheless, some methodological issues as the choice of force field parameters for atomic charges need more systematical analysis. Therefore, we consider on this work both the study of such methodological issues on describing electrostatic features and its applications to describe interesting molecular biophysical systems, such as the *Calmodulin* protein family who has the calcium as a ligand.



Lista de Abreviaturas

AD	Aproximação Dipolar
BS	Beresford-Smith
CaM	Calmodulina
Cab	Calbindina D_{9k}
DH	Debye-Hückel
DM	Dinâmica Molecular
Hb	Hemoglobina
MD	Modelo Detalhado (descrição atomística)
MDE	Modelo Detalhado Esférico (descrição atomística inserida em uma superfície esférica)
ME	Modelo Esférico
MC	Monte Carlo
NS	Nível de Schrödinger
NBO	Nível de Born-Oppenheimer
NMM	Nível de McMillan-Mayer
PB	Poisson-Boltzmann (não linear)
SOD	Superóxido Dismutase
TK	Tanford-Kirkwood



Lista de Símbolos

ϵ_i	Constante dielétrica macromolecular.
ϵ_s	Constante dielétrica do solvente.
ϵ_0	Permissividade elétrica do vácuo.
e	Carga elementar.
h	Constante de <i>Planck</i> .
N_a	Número de <i>Avogadro</i> .
k	Constante de <i>Boltzmann</i> .
T	Temperatura.
Λ	Comprimento de onda térmico.
z_k	Valência dos íons da espécie k .
Z_k	Carga dos íons da espécie k .
N_k	Número de partículas da espécie k .
z_j	Valência do íon j .
Z_j	Carga do íon j .
z_p	Valência líquida da macromolécula.
Z_p	Valência líquida da macromolécula.
z_+	Valência dos cátions do sal.
z_-	Valência dos ânions do sal.
z_{ci}	Valência dos contra-íons da solução.
z_M	Valência da macromolécula.
N_p	Número de átomos da proteína.
N_s	Número de partículas móveis em solução.
N_+	Número de cátions móveis em solução.



N_-	Número de ânions móveis em solução.
N	Número total de partículas do sistema.
M	Número de partículas carregadas da proteína.
C_s	Concentração de sal.
R_j	Raio do íon j .
R_c	Raio da célula.
V	Volume da célula.
κ	Inverso do comprimento de <i>Debye</i> .
r_{jl}	distância entre as partículas j e l .
r_j	módulo do vetor que determina a posição da partícula j .
$\vec{r}_j$	Vetor que determina a posição da partícula j .
x	Fração molar da espécie neutra (célula contendo a macromolécula e íons móveis).
x_{MLn}	Fração molar da macromolécula na forma <i>holo</i> .
x_M	Fração molar da macromolécula na forma <i>apo</i> .
x_L	Fração molar do ligante livre em solução.
γ	Coefficiente de atividade da espécie neutra.
γ_{MLn}	Coefficiente de atividade da macromolécula na forma <i>holo</i> .
γ_M	Coefficiente de atividade da macromolécula na forma <i>apo</i> .
γ_L	Coefficiente de atividade do ligante livre em solução.
μ	Potencial químico da espécie neutra.
μ^o	Potencial químico padrão.
μ_v^o	Potencial químico padrão da macromolécula no vácuo.
μ_{MLn}^o	Potencial químico padrão da macromolécula na forma <i>holo</i> .
μ_M^o	Potencial químico padrão da macromolécula na forma <i>apo</i> .
μ_L^o	Potencial químico padrão do ligante livre em solução.
ΔG^o	Variação da energia livre de <i>Gibbs</i> no estado padrão.
ΔG_{ref}^o	Variação da energia livre de <i>Gibbs</i> no estado padrão em uma concentração de referência.
ΔG_B^o	Contribuição à ΔG^o devido apenas aos ligantes associados à macromolécula.
ΔG	Variação da energia livre de <i>Gibbs</i> no processo de ligação.



ΔG^{solv}	Energia livre de solvatação.
$\Delta G^{solv(ref)}$	Energia livre de solvatação em uma concentração de referência.
ΔG_{MLn}^{solv}	Energia livre de solvatação da macromolécula na forma <i>holo</i> .
ΔG_M^{solv}	Energia livre de solvatação da macromolécula na forma <i>apo</i> .
ΔG_L^{solv}	Energia livre de solvatação do ligante livre em solução.
$\Delta G_{MLn}^{solv(ref)}$	Energia livre de solvatação da macromolécula na forma <i>holo</i> em uma concentração de referência.
$\Delta G_M^{solv(ref)}$	Energia livre de solvatação da macromolécula na forma <i>apo</i> em uma concentração de referência.
$\Delta G_L^{solv(ref)}$	Energia livre de solvatação do ligante livre em solução em uma concentração de referência.
ΔG^{carg}	Variação da energia livre de <i>Gibbs</i> devido ao carregamento da macromolécula.
ΔG^{el}	Componente eletrostática da energia livre de solvatação.
$\Delta G^{el(ref)}$	Componente eletrostática da energia livre de solvatação em uma concentração de referência.
$\Delta \Delta G_B$	Variação de ΔG_B^o em relação a variação de concentração salina ou macromolecular.
G_{vacuo}^{el}	Componente eletrostática da energia livre da macromolécula no vácuo.
G_{conc}^{el}	Componente eletrostática da energia livre do sistema formado pela macromolécula em solução.
G^{mol}	Contribuição dos átomos da macromolécula à energia livre eletrostática.
G^{sal}	Contribuição dos íon móveis à energia livre eletrostática.
G^{mov}	Contribuição entrópica devido à reorganização dos íons móveis à energia livre eletrostática.
G^{solv}	Contribuição entrópica devido à reorganização das moléculas do solvente à energia livre eletrostática.
K	Constante de equilíbrio termodinâmica.
pK	Forma logarítma da constante de equilíbrio termodinâmica.
pK_{ref}	Forma logarítma da constante de equilíbrio termodinâmica em uma concentração de referência.



ΔpK	Variação do pK em relação a variação de concentração salina ou macromolecular.
A	Energia livre de <i>Helmholtz</i> .
A_o	Energia livre de <i>Helmholtz</i> do sistema com todas as partículas possuindo carga nula.
A_c	Energia livre de <i>Helmholtz</i> do sistema carregado.
$\langle \rho(\vec{r}) \rangle$	Densidade de carga média na posição $\vec{r}$.
$\rho(\vec{r})$	Densidade de carga na posição $\vec{r}$.
$\rho^f(\vec{r})$	Densidade de carga fixa da macromolécula na posição $\vec{r}$.
$\rho^m(\vec{r})$	Densidade de carga produzida pelos íons móveis da solução na posição $\vec{r}$.
$\rho_k(\vec{r})$	Densidade de carga produzida pelos íons móveis da espécie k na posição $\vec{r}$.
$\rho_k(bulk)$	Densidade de carga produzida pelos íons móveis da espécie k no <i>bulk</i> da solução.
$\langle C_k(\vec{r}) \rangle$	Concentração média dos íons móveis da espécie k na posição $\vec{r}$.
$C_k(\vec{r})$	Concentração dos íons móveis da espécie k na posição $\vec{r}$.
n_+	Densidade numérica dos cátions do sal.
n_-	Densidade numérica dos ânions do sal.
n_{ci}	Densidade numérica dos contra-íons.
n_M	Densidade numérica da macromolécula.
λ	Parâmetro de carregamento.
$\langle \varphi(\vec{r}) \rangle$	Potencial eletrostático médio na posição $\vec{r}$.
$\langle \varphi(\vec{r}, \lambda) \rangle$	Potencial eletrostático médio em um estado de carregamento λ na posição $\vec{r}$.
$\varphi(\vec{r})$	Potencial eletrostático em uma determinada posição $\vec{r}$.
$u(r_{jl})$	Energia potencial de interação entre duas partículas j e l .
$v(\vec{r}_j)$	Energia potencial de interação entre uma dada partícula j e a borda da célula.
$E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})$	Energia total do sistema.



$U(\vec{r}_1, \dots, \vec{r}_N)$	Energia potencial total do sistema.
ΔU	Variação na energia potencial total do sistema devido a inserção de uma partícula.
ΔU^{ext}	Contribuição à ΔU devido ao potencial produzido pela macromolécula.
ΔU^{el}	Contribuição à ΔU devido à interação eletrostática com os íons móveis da solução.
ΔU^{er}	Contribuição à ΔU devido à interação de esfera-rígida com os íons móveis da solução.
$E(x)$	Energia total de uma dada configuração x .
$P(x)$	Probabilidade do sistema se encontrar em uma dada configuração x .
$w(x \rightarrow y)$	Probabilidade de transição de uma configuração x para uma configuração y .
$g_k(\vec{r})$	Função distribuição radial da espécie k na posição $\vec{r}$.
$w_k(\vec{r})$	Potencial de força média de uma partícula da espécie k , localizada na posição $\vec{r}$.
$Q(N, V, T)$	Função partição canônica.
$Z(N, V, T)$	Integral de configuração do ensemble canônico.
$P(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})$	Probabilidade do sistema assumir um estado determinado pela posição $\vec{r}_1$ e momento linear $\vec{p}_1$ para a partícula 1, posição $\vec{r}_2$ e momento linear $\vec{p}_2$ para a partícula 2, e assim sucessivamente.
$P^j(\vec{r}_j)$	Probabilidade de encontrar uma dada partícula j na posição $\vec{r}_j$.
$P_k^j(\vec{r}_j)$	Probabilidade de encontrar uma dada partícula j da espécie k na posição $\vec{r}_j$.
$P_k^{id}(\vec{r})$	Probabilidade de encontrar uma partícula da espécie k na posição $\vec{r}$, estando essas partículas não submetidas a qualquer interação.



Lista de Figuras

1.1	Representação esquemática do mecanismo de ativação de enzimas dependentes de cálcio através da <i>Calmodulina</i>	18
1.2	Representação esquemática de um sistema formado por uma macromolécula imersa em uma solução eletrolítica, indicando as aproximações inerentes do nível de <i>Born-Oppenheimer</i> e <i>McMillan-Mayer</i>	20
1.3	Organograma mostrando os níveis de aproximação e as respectivas maneiras de resolvê-los.	23
3.1	Componentes idealizadas do potencial químico do sistema formado pela macromolécula imersa em uma solução eletrolítica.	28
3.2	Representação do sítio de ligação do Ca^{+2} e Mg^{+2}	31
3.3	Representação esquemática da estrutura terciária da forma <i>holo</i> da <i>Calbindina</i> D_{9k}	33
3.4	Representação esquemática da estrutura terciária da forma <i>holo</i> da <i>Calmodulina</i>	34
3.5	Resultados experimentais de ΔpK devido a variação da concentração salina para a ligação de dois íons Ca^{+2} à uma forma mutante da <i>Calbindina</i> D_{9k} (N56A), e de quatro íons de Ca^{+2} à <i>Calmodulina</i>	35
3.6	Resultados experimentais de ΔpK devido a variação da concentração macromolecular para a ligação de dois íons Ca^{+2} à uma forma mutante da <i>Calbindina</i> D_{9k} (N56A).	35



4.1	Representação esquemática do Modelo da Célula já incorporando uma biomolécula.	37
4.2	Modelos utilizados para a representação do macroíon.	43
5.1	Função distribuição radial $g_k(\vec{r})$ em função da distância $ \vec{r} $ a uma partícula fixa no centro de uma célula esférica.	50
5.2	Potencial eletrostático calculado na superfície de uma esfera carregada imersa em uma solução eletrolítica.	55
5.3	Diagrama esquemático para o modelo de <i>Tanford-Kirkwood</i>	60
6.1	Resultados de ΔpK devido a variação da concentração salina, para a ligação de dois íons bivalentes positivos, em concentração macromolecular de 20 μM , 200 μM e 2 mM. A proteína foi representada segundo o Modelo Esférico com $Z_p = -4$ e $Z_p = -24$	68
6.2	Resultados de $\Delta\Delta G_B$ devido a variação da concentração salina de 1 mM para 100 mM, para a ligação de dois íons bivalentes positivos, em concentração macromolecular de 20 μM e 2 mM. A proteína foi representada segundo o Modelo Esférico.	70
6.3	Resultados de ΔpK para a ligação de dois íons bivalentes positivos devido a variação da concentração salina de 1 mM para 100 mM em concentração macromolecular de 2 mM. A proteína foi representada segundo o Modelo Esférico.	71
6.4	Resultados de ΔpK para a ligação de dois íons bivalentes positivos à <i>Calbindina D_{9k}</i> utilizando o campo de força Reduzido.	71
6.5	Resultados de ΔpK para a ligação de íons bivalentes positivos à <i>Calmodulina</i> utilizando o campo de força Reduzido.	72



6.6	Resultados de ΔpK para a ligação de dois íons bivalentes positivos aos fragmentos da <i>Calmodulina</i> em concentração macromolecular de 20 μM , utilizando o campo de força Reduzido.	73
6.7	Resultados de ΔpK devido a variação da concentração salina de 1 mM para 100 mM em concentração macromolecular de 20 μM . A proteína foi representada segundo o Modelo Esférico.	74
6.8	Resultados de ΔpK devido a variação da concentração salina para a ligação de quatro íons bivalentes positivos à <i>Calmodulina</i> utilizando a aproximação de <i>Lin</i> , e para a ligação de dois íons bivalentes positivos à <i>Calbindina D_{9k}</i> utilizando a aproximação de DH.	76
6.9	Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons bivalentes positivos aos fragmentos da <i>Calmodulina</i> utilizando a aproximação de <i>Lin</i>	77
6.10	Resultados de ΔpK devido à variação da concentração salina para a ligação de dois íons bivalentes positivos aos fragmentos da <i>Calmodulina</i> em concentração macromolecular de 200 μM	77
6.11	Resultados de ΔpK devido à variação da concentração salina para a ligação de dois íons bivalentes positivos à <i>Calbindina D_{9k}</i> em concentração macromolecular de 200 μM	78
6.12	Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons de Ca^{+2} à <i>Calbindina D_{9k}</i> , e de quatro íons de Ca^{+2} à <i>Calmodulina</i>	80
6.13	Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons de Ca^{+2} aos fragmentos da <i>Calmodulina</i>	80
A.1	Representação esquemática do íon acetato	93



B.1	Cálculos de ΔpK devido a variação da concentração salina para a ligação de dois íons bivalentes positivos. A proteína foi representada segundo o Modelo Esférico.	96
C.1	Exemplo de um arquivo de saída do programa que resolve o modelo de TK. .	97



Lista de Tabelas

1.1	Resultados da energia livre de solvatação do íon acetato (ΔG^{solv}) obtidas a partir da solução linear da equação de PB.	20
4.1	Número de ânions obtidos apartir da relação 4.3 e 4.4, para algumas concentrações de sal e de proteína, utilizando uma macromolécula de carga líquida $Z_p = -4e$	37
A.1	Cargas parciais para os átomos do íon acetato.	91



Capítulo 1

Introdução

As interações entre ligantes e biomoléculas têm uma importância fundamental em um grande número de processos químicos e biológicos [1, 2, 3]. Um exemplo muito bem conhecido é o da *Hemoglobina* (Hb) [3]. Esta proteína tem a função de transportar oxigênio dos alvéolos pulmonares para os tecidos, e para isso ela deve ligar o O_2 em determinadas condições (pulmão) e liberá-lo em outras condições (tecidos). Para auxiliá-la no controle deste processo, a Hb também interage com o íon H^+ , cujo aumento em sua concentração (redução do pH) diminui a afinidade da Hb pelo oxigênio. Portanto, no pulmão, onde o pH é maior que nos tecidos, a Hb liga o O_2 e libera H^+ . Já nos tecidos, onde temos produção de H^+ , a liberação do O_2 é facilitada. Outro exemplo de extrema importância é o da catálise enzimática [3]. Neste processo, um substrato liga-se no sítio ativo de uma enzima formando o complexo enzima-substrato. Nesta situação, o substrato passa por uma série de reações químicas, se transformando em produto, com uma velocidade muito maior do que o faria sem a presença da enzima. Dessa forma, sem a ação das enzimas, a transformação do substrato em produto seria demasiadamente lenta, incapaz de fornecer produtos na proporção adequada para a manutenção do metabolismo [3].

A figura 1.1 representa um outro caso clássico da interação entre ligantes e biomoléculas, o mecanismo de ativação de enzimas dependentes de cálcio [4, 5]. Neste processo, uma dada enzima é inativa devido a presença de um peptídeo inibidor. Quando íons de cálcio se ligam nos sítios de uma proteína intracelular chamada *Calmodulina* (CaM), ocorre uma



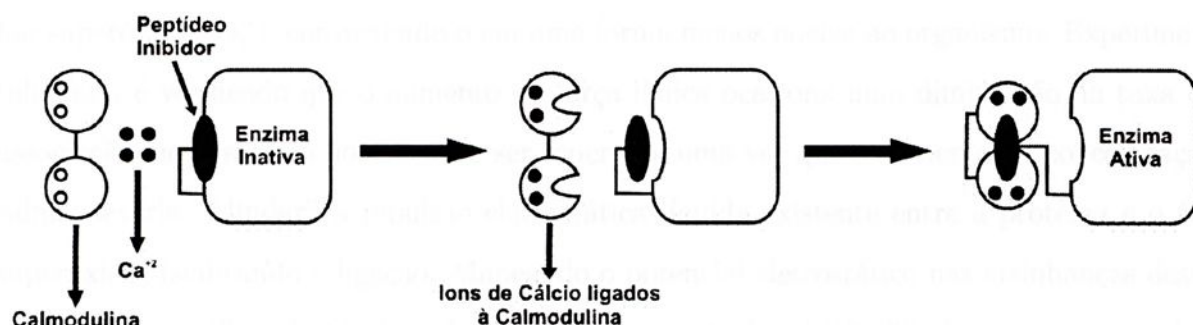


Figura 1.1: Representação esquemática do mecanismo de ativação de enzimas dependentes de cálcio através da *Calmodulina*. Esta figura foi adaptada da referência [4].

alteração conformacional nesta proteína (este processo será apresentado em mais detalhes na seção 3.1), fazendo com que ela interaja com o segmento inibidor, tornando a enzima ativa. Este mecanismo é equivalente ao mecanismo de ação de fármacos¹, como o antidepressivo trifluorperazina (TFP) que se liga a CaM [6].

O entendimento desses processos, incluindo seu detalhamento ao nível estrutural e energético, está diretamente relacionado ao desenvolvimento racional de fármacos (*drug design*) [1, 7, 8, 9]. Encontrar a localização precisa de sítios de ligação e estimar as características químicas do ligante, necessita, além de informações estruturais, do conhecimento das interações que direcionam o processo de associação [7, 8]. Dessa forma, o cálculo preciso de certas propriedades termodinâmicas é fundamental para a total eficiência das metodologias utilizadas na indústria farmacêutica e de biotecnologia, por exemplo. Além disso, todas essas moléculas são compostas de grupos ionizáveis, cujo equilíbrio ácido-base (associação e dissociação de prótons) influencia diretamente no seu comportamento [3, 10, 11, 12]. Devido a isso, a verificação e o conseqüente desenvolvimento de modelos teóricos, possibilitam avaliar quais aproximações são mais eficientes para o cálculo de propriedades físico-químicas de moléculas e complexos moleculares, e auxiliam no desenvolvimento de pacotes computacionais que realizam buscas em bancos de dados de ligantes [8], sugerindo os “melhores” ligantes.

Modelos teóricos baseados em interações eletrostáticas vêm conseguindo descrever estes sistemas de modo bastante eficiente [15, 16, 17, 18]. Um caso particularmente interessante é o da *Superóxido Dismutase* (SOD) [18, 19, 20], uma enzima de carga -4, que se associa ao

¹ Fármacos, na sua imensa maioria, são pequenas moléculas, geralmente orgânicas, que interagem e “modulam” a atividade biológica de determinadas biomoléculas (tipicamente enzimas) [13, 14].

íon superóxido (O_2^-), convertendo-o em uma forma menos nociva ao organismo. Experimentalmente, é verificado que o aumento da força iônica ocasiona uma diminuição na taxa de associação, ao contrário do que deve ser esperado, uma vez que o aumento da concentração salina deveria “blindar” a repulsão eletrostática líquida existente entre a proteína e o íon superóxido, facilitando a ligação. Mapeando o potencial eletrostático nas vizinhanças desta enzima [19] e utilizando técnicas de simulação computacional [19, 20], demonstrou-se, sobre certas condições, uma região de potencial positivo no sítio ativo da SOD, explicando este aparente paradoxo. Outros exemplos podem ser citados, como a descrição da regulação do equilíbrio alostérico entre as formas *oxi* e *desoxi* da Hb [15] e estudos da interação entre íons ligantes e proteínas [18], peptídeos e proteínas [21], e proteína-proteína [22].

Uma das principais dificuldades no estudo teórico da interação entre ligantes e biomoléculas, é o tratamento do ambiente (solvente e íons móveis da solução) no qual a associação ocorre [18]. As interações entre a molécula e o solvente (hidratação), e a distribuição dos íons móveis ao seu redor são fatores decisivos no processo de ligação, uma vez que esses íons tendem a se acumular próximos a regiões mais carregadas da molécula, atenuando a interação dessas regiões com o ligante. Além disso, a estrutura altamente organizada do solvente próximo a essas regiões carregadas, pode limitar a aproximação desses íons [15], influenciando em sua distribuição, e conseqüentemente, na interação efetiva entre a macromolécula e o ligante.

Podemos tratar esses sistemas utilizando ferramentas da Mecânica Estatística Clássica [23], de forma a obter as propriedades físico-químicas de interesse através da **média** sobre um *ensemble*² de configurações. Devido a complexidade envolvida nesses sistemas e ao grande número de graus de liberdade existentes, muitas vezes o cálculo explícito da *função partição*³ se torna impraticável matematicamente, necessitando de aproximações que podem comprometer a qualidade dos resultados. Alternativamente, métodos numérico-computacionais que permitem resolver “exatamente” o modelo, como o método Monte Carlo (MC) [18, 24], podem ser ineficientes devido ao grande tempo computacional necessário para obter uma convergência razoável, ou seja, para “equilibração” do sistema e obtenção de médias representativas das propriedades termodinâmicas de interesse.

² Ensemble é o conjunto de um número muito grande de sistemas macroscopicamente idênticos a um determinado sistema de interesse, mas não necessariamente idênticos do ponto de vista microscópico, de forma que as propriedades termodinâmicas do sistema de interesse possam ser consideradas equivalente a média dessas propriedades sobre todos os sistemas do ensemble.

³ A definição de *função partição* será vista no capítulo 5.



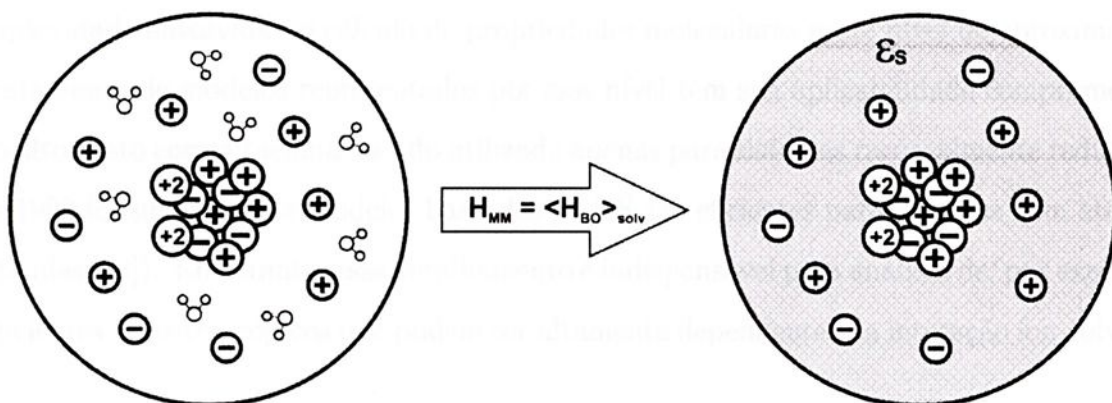


Figura 1.2: Representação esquemática de um sistema formado por uma macromolécula imersa em uma solução eletrolítica, indicando as aproximações inerentes do NBO e NMM. O sistema da esquerda mostra, além das espécies pertencentes ao soluto (macromolécula e íons de sal), as moléculas do solvente explicitamente (NBO). Já no sistema da direita, o comportamento explícito do solvente é substituído por sua constante dielétrica (NMM).

Assim, uma análise metodológica dos principais métodos utilizados para o estudo da interação entre ligantes e biomoléculas, mostrando as limitações de cada aproximação, e selecionando as mais adequadas e eficientes para cada tipo de aplicação, é viável e importante, uma vez que essas metodologias se tornam cada vez mais próximas da área tecnológica [8], necessitando de testes mais sistemáticos, e critérios para assegurar sua utilização.

1.1 Níveis de Detalhamento do Modelo

Para abordar este tipo de problema do ponto de vista teórico, primeiramente precisamos escolher o nível de detalhamento do sistema, ou seja, a *Hamiltoniana efetiva* [18, 25, 26] e seus parâmetros. Neste contexto, os átomos da biomolécula e da solução eletrolítica no qual esta molécula estará inserida, podem ser representados de forma bastante elaborada, onde se leva em conta o movimento eletrônico e interações do ponto de vista quântico, ou, em outro extremo, por modelos bastante simplificados e rudimentares, onde os parâmetros são apenas as coordenadas dos átomos da molécula e sua carga [27], porém, permitindo-nos capturar a essência da física destes sistemas.

Segundo *Friedmann*, no *nível de Schrödinger* (NS) [25, 26], os átomos são representados de maneira bem precisa, de forma que as variáveis da *Hamiltoniana* são as coordenadas dos núcleos e elétrons de cada componente do sistema, sendo sua forma funcional obtida de primeiros princípios (*ab initio*). Devido ao grande número de graus de liberdade e a

complexidade envolvida no cálculo de propriedades moleculares neste nível de aproximação, o tratamento de modelos representados por esse nível tem sua aplicabilidade comprometida pelo alto custo computacional, sendo utilizado apenas para sistemas razoavelmente reduzidos (em 1990 foi sugerido que modelos baseados no NS são eficientes para sistemas com até 100 partículas [28]). Entretanto, esse detalhamento é indispensável para análises de, por exemplo, coeficientes espectroscópicos que podem ser altamente dependentes da interação íon-solvente [26].

Uma forma de diminuir essa complexidade, é substituir o comportamento explícito de algumas espécies do sistema por seu comportamento médio, reduzindo o número de graus de liberdade e a necessidade da utilização de formas funcionais cada vez mais complexas para representar a *Hamiltoniana* do sistema. Um exemplo dessa aproximação é a substituição do comportamento eletrônico explícito por seu comportamento médio, de forma que as variáveis do modelo sejam as coordenadas e momentos dos átomos das moléculas do solvente e do soluto do sistema. Esse nível de aproximação é denominado de *nível de Born-Oppenheimer* (NBO) [18, 25, 26].

O *nível de McMillan-Mayer* (NMM) [18, 25, 26], baseando-se no NBO, substitui o comportamento explícito das moléculas do solvente pelo seu comportamento médio (figura 1.2), de forma que as interações entre as partículas do soluto sejam representadas em função da constante dielétrica estática do solvente. Essa aproximação, embora com bastante êxito e largamente empregada no cálculo de propriedades eletrostáticas macromoleculares [17, 15, 29], carece de eficiência quando utilizada para sistemas dependentes de potenciais de curto alcance, onde as características estruturais do solvente se tornam indispensáveis, sendo necessário a utilização de modelos no NBO, para a “correta” descrição dos fenômenos de hidratação [18].

Enquanto no NS as interações entre os constituintes do sistema (interação elétron-elétron, elétron-núcleo e núcleo-núcleo) são conhecidas por primeiros princípios, nos níveis mais simplificados a interação entre os átomos não é conhecida *a-priori* [25], uma vez que o comportamento explícito dos elétrons é suprimido. Uma forma de resolver este problema, é sugerir uma forma funcional *efetiva* para os termos referentes a energia potencial da *Hamiltoniana*, introduzindo certos parâmetros cujos valores possam ser **ajustados** de forma a garantir a concordância com resultados experimentais, ou obtidos através de cálculos no NS. Esse



Campo de Força	ΔG^{solv} (kT)	ΔG^{solv} (kcal/mol)
CHARMM	-121,48	-71,99
Reduzido	-125,80	-74,55
PARSE	-126,70	-75,08
GROMOS96 (43a1)	-128,63	-76,23

Tabela 1.1: Resultados da energia livre de solvatação do íon acetato (ΔG^{solv}) obtidas a partir da solução linear da equação de PB, utilizando pacote computacional *DelPhi* [41, 42]. Valores experimentais estão entre -79 a -87 kcal/mol [43].

conjunto de funções matemáticas e seus parâmetros é denominado **campo de força** [18].

Existe um número razoável de campos de força propostos na literatura (AMBER [30], PARSE [31, 32, 33], CHARMM [34], GROMOS [35, 36], entre outros [37, 38]), de forma que se torna importante verificar qual a dependência desses parâmetros na obtenção de propriedades termodinâmicas, apurando quais deles são mais apropriados para cada tipo de aplicação, diminuindo assim a arbitrariedade na sua escolha.

Um exemplo bastante ilustrativo da dependência do campo de força no cálculo de propriedades estruturais de macromoléculas, pode ser atribuído ao estudo da dupla hélice do DNA através de simulações por *Dinâmica Molecular* (DM) [18, 24, 39]. Alguns trabalhos têm mostrado a existência de diferentes estruturas de equilíbrio quando da utilização do campo de força AMBER ou CHARMM [40].

Outro caso pode ser verificado na Tabela 1.1, onde mostramos nossos próprios resultados reproduzindo a energia livre de solvatação do íon acetato (ΔG^{solv}) [44], através da utilização de quatro campos de força distintos⁴. Esses cálculos foram realizados utilizando o pacote computacional *DelPhi* [41, 42], resolvendo a equação linear de *Poisson-Boltzmann* através do método de diferenças finitas [27], atribuindo uma constante dielétrica $\epsilon_i = 2$ para o interior macromolecular, e $\epsilon_s = 78,7$ para o solvente (temperatura de 298 K), na ausência de força iônica. Verificamos uma diferença de até 7 unidades de kT entre os resultados obtidos com os diferentes campos de força, demonstrando que pequenas variações na carga parcial de alguns átomos desse íon, podem ter grande influência no cálculo de propriedades dependentes das

⁴ A descrição dos parâmetros do campo de força para o íon acetato é dada no apêndice A.



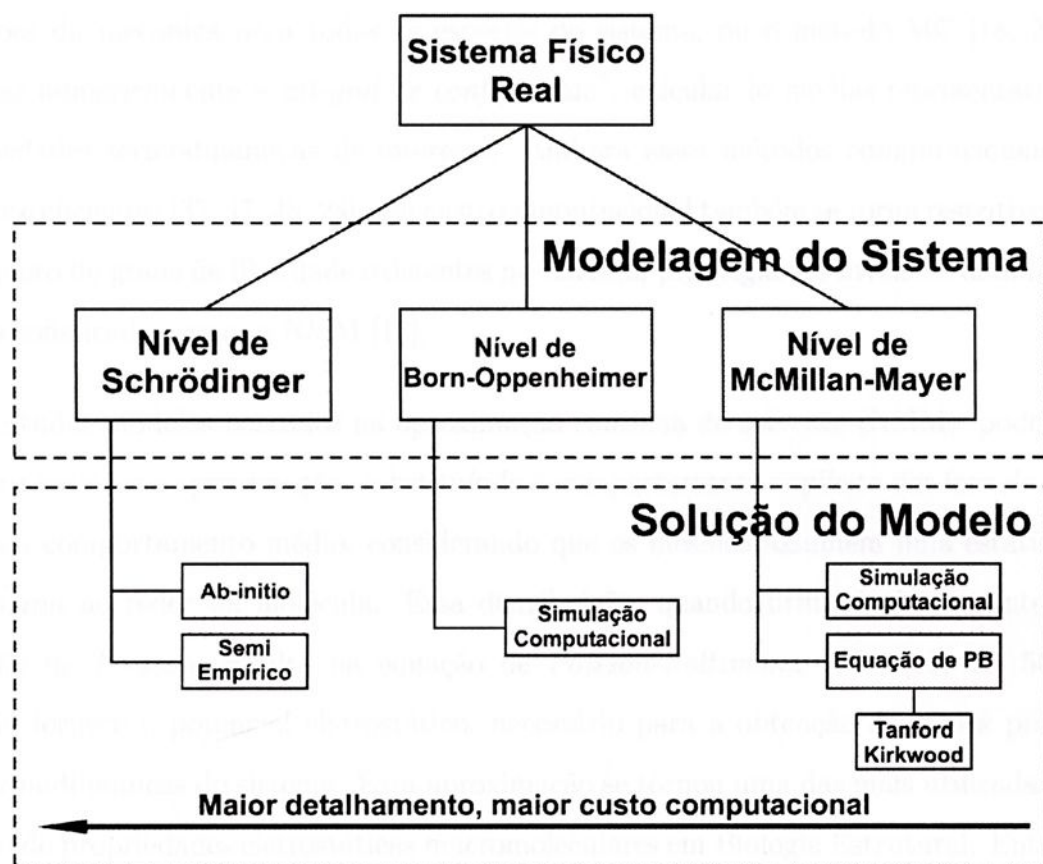


Figura 1.3: Organograma mostrando os níveis de aproximação e as respectivas maneiras de resolvê-los. Outras formas para solução destes modelos estão omitidas, como *Mean Spherical Approximation* e *Hypernetted Chain* [23]. As análises dos métodos de solução para o nível de McMillan-Mayer apresentados na figura são os objetos de estudo deste trabalho.

cargas de polarização. Veremos no capítulo 6 que outras grandezas, como, por exemplo, a variação da constante de ligação (ΔpK), não possuem essa dependência.

1.2 Solução do Modelo

A solução desses modelos também pode ser obtida em vários níveis de aproximação (figura 1.3), de forma que a relação custo computacional/qualidade dos resultados seja a melhor possível [28]. Para modelos que utilizam átomos e moléculas de forma detalhada, incluindo detalhes eletrônicos (NS), a solução deve ser realizada no nível quântico, resolvendo a equação de *Schrödinger* em cálculos *ab initio* ou *semi-empírico* [45].

Para modelos mais simplificados, onde a aproximação clássica se torna razoável, podemos utilizar técnicas de simulação computacional como o método de DM [46, 24], resolvendo as

⁵ A definição de *integral de configuração* será vista no capítulo 5.

equações da mecânica para todas as espécies do sistema, ou o método MC [18, 24] para integrar numericamente a *integral de configuração*⁵, calculando médias representativas das propriedades termodinâmicas de interesse. Embora esses métodos computacionais sejam bastante eficientes [37, 47, 48, 24], seu custo computacional também se torna restritivo devido ao número de graus de liberdade existentes no sistema, privilegiando níveis de detalhamento menos sofisticados como o NMM [18].

Para estudar modelos baseados na aproximação contínua do solvente (NMM), podemos introduzir mais uma aproximação, substituindo o comportamento explícito dos íons da solução pelo seu comportamento médio, considerando que os mesmos assumem uma estatística de *Boltzmann* ao redor da molécula. Essa distribuição, quando utilizada juntamente com a equação de *Poisson*, resulta na equação de *Poisson-Boltzmann* (PB) [18, 49, 50], cuja solução fornece o potencial eletrostático, necessário para a obtenção de várias propriedades termodinâmicas do sistema. Esta aproximação se tornou uma das mais utilizadas para o cálculo de propriedades eletrostáticas macromoleculares em Biologia Estrutural. Entretanto, devido a não linearidade dessa equação, mesmo sua solução numérica se torna extremamente complicada, tornando vantajoso a utilização da sua forma linear [51, 52], equivalente a teoria de *Debye-Hückel* (DH) [18, 23].

Dentre outras simplificações, modelos baseados em DH ou até mesmo PB não linear, assumem uma concentração macromolecular infinitamente pequena [53], inconsistente com algumas técnicas experimentais utilizadas, como NMR, onde a concentração macromolecular típica é de 1 mM [54], e a concentração de contraíons que acompanham a molécula pode ser uma ordem de magnitude maior. Assim, se torna interessante verificar qual a eficiência desses modelos na reprodução de resultados experimentais obtidos em concentração macromolecular finita.

A obtenção de soluções analíticas dessas equações só é possível para sistemas altamente simétricos, como no caso do modelo de *Tanford-Kirkwood* (TK) [18, 55, 56], onde a macromolécula é descrita atomisticamente e inserida em uma cavidade **esférica** de baixa constante dielétrica, e a solução eletrolítica é tratada de acordo com a teoria de DH. Embora a literatura tenha mostrado que várias propriedades físico-químicas de sistemas biomoleculares têm uma alta dependência com a descrição da superfície macromolecular [29], se torna muito



útil verificar a acuracidade do modelo de TK na previsão de propriedades termodinâmicas relacionadas com a interação entre ligantes e biomoléculas, uma vez que soluções analíticas são sempre preferíveis em relação a métodos numéricos, por não estarem limitados por problemas relacionados a convergência e custo computacional. Além disso, o modelo de TK tem uma importância histórica por servir de fundamentação para a área de eletrostática de proteínas.

Objetivos do Trabalho



Capítulo 2

Objetivos do Trabalho

A Interação Ligante-Biomolécula

Concentraremos nosso trabalho na análise e comparação dos métodos mais populares utilizados no cálculo de propriedades eletrostáticas da interação entre ligantes e biomoléculas. Mais especificamente trabalharemos com a interação de íons Ca^{+2} à proteínas da família da *Calmodulina*, como a própria *Calmodulina*, seus dois fragmentos, e a *Calbindina D_{9k}*. Podemos dividir nossos objetivos em basicamente duas partes:

1. A primeira parte está relacionada à formulação do modelo, onde a finalidade principal será verificar a influência da distribuição das cargas da macromolécula no cálculo da variação da constante de ligação em relação a variação da força iônica. Para isso, utilizaremos dois campos de força bastante conhecidos da literatura (GROMOS e PARSE) e outros modelos mais grosseiros para representar a macromolécula. Com essa análise, poderemos verificar qual o nível de detalhamento que precisamos inserir no modelo da macromolécula para obter resultados satisfatórios para a variação da constante de ligação em relação à concentração salina e/ou proteica.
2. A outra parte diz respeito a solução do modelo, principalmente na comparação das principais metodologias presentes na literatura. Será analisado o efeito das aproximações de campo médio, inerentes de modelos baseados na equação de *Poisson-Boltzmann*, superfície macromolecular esférica, que possibilita a solução analítica de *Tanford-Kirkwood*, e concentração macromolecular finita, que consegue ser incorporada através da utilização do método *Monte Carlo*.



Capítulo 3

O Problema

A Interação Ligante-Biomolécula

A função biológica da maioria das biomoléculas está diretamente relacionada com sua interação com ligantes [57]. Esses ligantes podem ser desde pequenos íons, como no caso de íons de Ca^{+2} [58] que se ligam a uma série de proteínas para controlar processos regulatórios intracelulares, até fármacos de complexa estrutura molecular [7, 8], cuja ligação pode promover alterações conformacionais no receptor, influenciando sua função. De um modo geral, a ligação de n ligantes L a uma macromolécula M pode ser representada pela equação [57]:



Este processo de ligação está associado a uma variação na energia livre de *Gibbs* do sistema (ΔG) [59], cuja análise nos fornece informações do quanto essa reação é favorecida nas condições em que ela é realizada. Podemos definí-la em função dos potenciais químicos μ_i das i espécies em solução:

$$\Delta G = \mu_{ML_n} - \mu_M - n\mu_L \quad (3.2)$$

Devido ao longo alcance da interação eletrostática, soluções compostas por macromoléculas



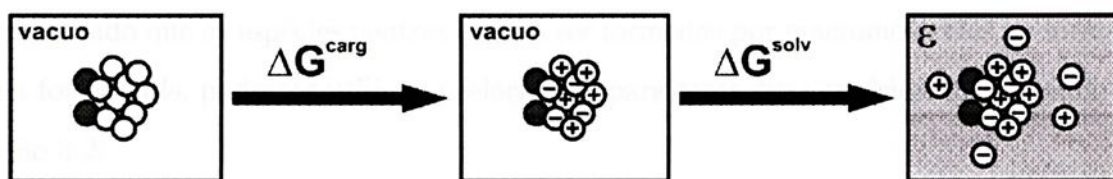


Figura 3.1: Componentes idealizadas do potencial químico do sistema formado pela macromolécula imersa em uma solução eletrolítica. Os círculos escuros representam os sítios de ligação. O particionamento da energia livre nas componentes acima é um artifício utilizado para o cálculo de certas propriedades termodinâmicas. Óbviamente, não existem macromoléculas carregadas no vácuo. As cargas são adquiridas em solução.

altamente carregadas e pequenos íons de baixa carga (Na^+ , K^+ e Cl^- , por exemplo), tendem a assumir uma estrutura, onde os macroíons se mantêm distantes devido a forte repulsão eletrostática existente entre eles, com os íons se distribuindo ao seu redor, de forma que em uma determinada região entre as macromoléculas, o campo eletrostático se anule [49]. Dessa forma, em baixa concentração macromolecular, onde os macroíons se mantêm suficientemente afastados, podemos considerar que cada espécie neutra formada por uma macromolécula e pelos íons distribuídos ao seu redor se comporte independente das outras [49, 60].

Sendo assim, o potencial químico dessa espécie neutra é dado por [61]:

$$\mu = \mu^o + kT \ln x + kT \ln \gamma \quad (3.3)$$

Onde k é a constante de *Boltzmann*, e T a temperatura. O segundo termo desta relação está associado a entropia de mistura, sendo x a fração molar da espécie neutra, enquanto o terceiro termo introduz os desvios do comportamento ideal através do coeficiente de atividade γ [59]. O potencial químico do estado padrão μ^o é a componente referente as interações entre a macromolécula, os íons móveis e o solvente. Podemos particioná-lo em três componentes (figura 3.1): Uma referente a todas as interações não eletrostáticas entre os átomos da macromolécula, considerando essa molécula no vácuo (μ_v^o), outra componente devido as interações eletrostáticas entre os átomos da macromolécula (ΔG^{carg}), e uma componente devido ao trabalho em transferir a macromolécula, já carregada, do vácuo para solução eletrolítica (ΔG^{solv}):

$$\mu^o = \mu_v^o + \Delta G^{carg} + \Delta G^{solv} \quad (3.4)$$

Considerando que as espécies neutras podem ser formadas por macromoléculas na forma *apo* ou na forma *holo*, podemos utilizar a relação 3.3 para essas duas espécies, substituindo-a na relação 3.2:

$$\Delta G = \mu_{MLn}^o - \mu_M^o - n\mu_L^o + kT \ln K \quad (3.5)$$

Onde K , a constante de equilíbrio termodinâmica (também denominada constante de ligação), é definida como [59]:

$$K = \frac{x_{MLn} \gamma_{MLn}}{x_M \gamma_M (x_L \gamma_L)^n} \quad (3.6)$$

Em uma solução em equilíbrio à temperatura e pressão constante, $\Delta G = 0$, portanto:

$$\mu_{MLn}^o - \mu_M^o - n\mu_L^o = \Delta G^o = -kT \ln K = 2,303 kT pK \quad (3.7)$$

Onde $pK = -\log K$ e ΔG^o é a variação da energia livre de *Gibbs* no estado padrão, definida como o trabalho isotérmico necessário para mover os ligantes da solução para os sítios de ligação, excluindo efeitos coletivos da solução. Dessa forma, variando as condições em que a ligação ocorre, e verificando a correspondente variação em ΔG^o , e consequentemente em K , podemos estimar como as interações entre os constituintes do sistema influenciam a ligação.

Podemos particionar ΔG^o em duas componentes: uma componente devido apenas aos ligantes associados à macromolécula (ΔG_B^o), e uma componente devido apenas aos ligantes livres na solução:

$$\Delta G^o = \Delta G_B^o + n\mu_L^o \quad (3.8)$$

Onde definimos:



$$\Delta G_B^o = \mu_{MLn}^o - \mu_M^o \quad (3.9)$$

Assim, a variação em ΔG^o em relação a uma concentração salina ou macromolecular de referência (representada pelo índice *ref*) é dada por:

$$\Delta \Delta G = \Delta G^o - \Delta G_{ref}^o = \Delta \Delta G_B + n \Delta \mu_L^o \quad (3.10)$$

O termo $\Delta \Delta G_B$ será utilizado para verificarmos apenas o efeito da distribuição iônica ao redor da macromolécula, uma vez que a inclusão do potencial químico do ligante livre em solução, como para o cálculo do ΔpK , pode trazer compensações energéticas, podendo esconder alguns importantes efeitos.

Analogamente, a variação da constante de ligação é obtida facilmente através da equação:

$$\Delta pK = pK - pK_{ref} = \frac{\Delta \Delta G}{2,303 kT} \quad (3.11)$$

Substituindo a equação 3.7 na relação acima, utilizando a relação 3.4 para as espécies em solução, e se considerando que os termos devido às interações no vácuo independem das variações realizadas no sistema, obtemos ΔpK apenas em função da energia livre de solvatação das espécies em solução:

$$\Delta pK = \frac{1}{2,303 kT} \left\{ \Delta G_{MLn}^{solv} - \Delta G_M^{solv} - \Delta G_L^{solv} - \left[\Delta G_{MLn}^{solv(ref)} - \Delta G_M^{solv(ref)} - \Delta G_L^{solv(ref)} \right] \right\} \quad (3.12)$$

A energia livre de solvatação pode ser dividida, basicamente, em duas partes:

$$\Delta G^{solv} = \Delta G^{el} + \Delta G^{outras} \quad (3.13)$$

1. Uma parte eletrostática (ΔG^{el}), referente às interações eletrostáticas entre o solvente, os íons móveis da solução, e os grupos carregados existentes na macromolécula, bem como

as componentes entrópicas devido a redistribuição das partículas móveis da solução, causada pelo grande campo elétrico gerado pela macromolécula.

2. Uma parte referente a todas as outras contribuições, como mudanças estruturais da proteína, interações de curto alcance, e interações entre grupos não polares, etc.

Neste trabalho, estamos considerando que as componentes não eletrostáticas (ΔG^{outras}) são suficientemente desprezíveis em relação às alterações realizadas no sistema (alteração da concentração salina e macromolecular), sendo canceladas no cálculo da variação da constante de ligação (relação 3.12) devido a diferença entre as energias livre de solvatação [18]:

$$\Delta G^{solv} - \Delta G^{solv(ref)} \approx \Delta G^{el} - \Delta G^{el(ref)} \quad (3.14)$$

3.1 Proteínas da Família da *Calmodulina*

As proteínas da família da *Calmodulina* (CaM) (*Troponina C*, *Parvalbumina*, *Calbindina D_{28k}*, entre outras), embora relacionadas a funções bastante distintas, como a contração muscular e a regulação de processos intracelulares [58], possuem a estrutura do sítio de ligação do cálcio bastante conservada [62]. Nesta estrutura, conhecida como *EF-hand*, duas hélices são conectadas por um *loop* contendo um grande número de resíduos ácidos, que interagem fortemente com íons metálicos positivos, como o Ca^{+2} e Mg^{+2} [58, 62, 4]. Podemos observar através da figura 3.2, a grande quantidade de aminoácidos ácidos (aspartato e glutamato) que formam um dos sítios de ligação da *Calbindina D_{9k}* (Cab), produzindo uma região de alto potencial eletrostático negativo [63], que acomoda os íons metálicos positivos.

O termo “*EF-hand*” foi proposto por *Kretsinger* e *Nockolds* em 1973 [66], em um trabalho onde é descrita a estrutura tridimensional da *Parvalbumina*. Eles verificaram que o sítio de ligação do cálcio era formado por um par de hélices ligadas por um *loop*, cuja estrutura se assemelhava à uma mão com o dedo polegar e o dedo indicador estendidos e os outros três dedos fechados, sendo as hélices representada por cada um dos dois dedos estendidos, e o interior da proteína voltada para o lado da palma da mão. Além de estrutural, esta representação também descreve o movimento das hélices, uma vez que a ligação do cálcio

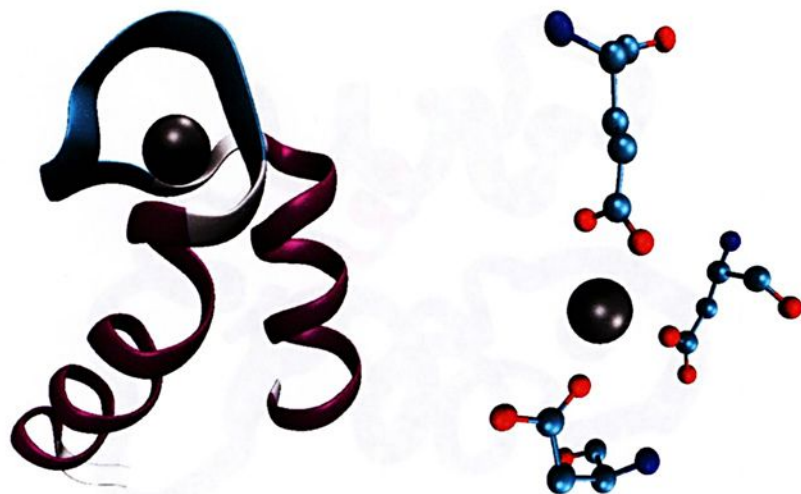


Figura 3.2: Representação do sítio de ligação do Ca^{+2} e Mg^{+2} (esfera maior). A figura da **esquerda** mostra a posição do sítio de ligação em relação às duas hélices e o *loop* que formam a *EF-hand*. Na figura da **direita** mostramos alguns dos aminoácidos pertencentes a um dos sítios de ligação da *Calbindina D_{9k}* (PDB-ID 3icb [64]). Esta ilustração foi produzida com o pacote computacional VMD [65].

induz uma mudança conformacional onde a hélice representada pelo polegar se move na direção da hélice representada pelo indicador, como se estes dedos estivessem inicialmente afastados, e se unissem. A *Parvalbumina* possui dois sítios de ligação cujas hélices foram designadas por C e D para o primeiro sítio, e E e F para o segundo (foi verificado que as hélices A e B, embora possua uma estrutura similar aos outros dois pares de hélices, não constitui um sítio de Ca^{+2}). Como o sítio formado pelo *loop* EF podia ser representado pela mão direita, esta estrutura recebeu o nome de *EF-hand* [67].

A função de cada uma dessas proteínas está diretamente relacionada com a afinidade com que esses íons são ligados (constante de ligação $K = 10^{-9}$ a $10^{-5} M$) [58], e com a mudança conformacional produzida pela ligação deles [4, 58, 62]. A substituição de um ácido glutâmico por um ácido aspártico (E101D) no sítio ativo da *Parvalbumina*, diminui a afinidade do cálcio em 100 vezes e aumenta a do magnésio em 10 vezes, devido a alterações estruturais que modificam a geometria de coordenação do cálcio, mas mantém a do magnésio [68], demonstrando a importância, não só da estrutura *EF-hand*, mas também da estrutura local do sítio de ligação para a função biológica dessas proteínas.

Neste trabalho, estamos interessados no estudo da ligação de íons de cálcio à Cab e a própria CaM. Essas proteínas foram escolhidas devido a grande distinção em relação a sua carga líquida e estabilidade estrutural, se tornando casos limites para verificação da validade das

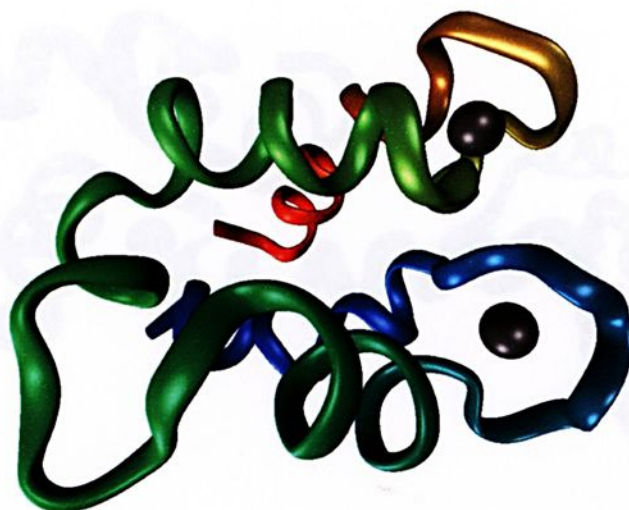


Figura 3.3: Representação esquemática da estrutura terciária da forma *holo* da Cab (PDB-ID 3icb). As esferas representam os íons de cálcio. Esta ilustração foi produzida com o pacote computacional VMD [65].

aproximações utilizadas por nós, além de sua importância biológica.

3.1.1 Calbindina D_{9k}

A Cab (PDB-ID 3icb [64]) é uma proteína intracelular regulatória formada por 75 resíduos de aminoácidos, possuindo peso molecular de 8.500. A carga líquida da forma *apo* em pH 7.5 é de -7e [69, 70]. Essa proteína possui duas estruturas *EF-hand*, responsáveis pela ligação de dois íons de Ca^{+2} com alta afinidade e cooperatividade positiva [70].

Sua estrutura revela uma forma bastante globular (figura 3.3), com todos os resíduos carregados próximos a superfície [70], fazendo com que modelos mais simplificados se tornem consideravelmente razoáveis para seu estudo. Outro fato que a faz um excelente objeto de estudo para modelos simplificados, é a sua enorme estabilidade estrutural [63, 69, 70]. Estudos de 1H NMR indicam que sua estrutura em solução é muito similar a da forma cristalina [63, 69], além de mostrar uma resposta conformacional muito pequena devido a ligação dos íons de cálcio, em relação a CaM e Troponina C, por exemplo [70].

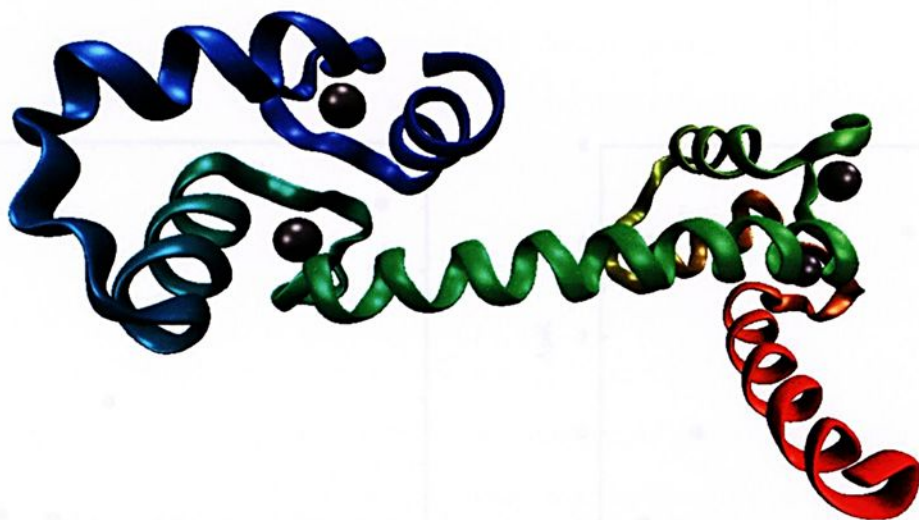


Figura 3.4: Representação esquemática da estrutura terciária da forma *holo* da CaM (PDB-ID 3cln). As esferas representam os íons de cálcio. Esta ilustração foi produzida com o pacote computacional VMD [65].

3.1.2 Calmodulina

A CaM (PDB-ID 3cln) é também uma proteína intracelular que regula a atividade biológica de diversas proteínas celulares e de transporte de cálcio. É formada por 148 resíduos de aminoácidos, possuindo peso molecular de 16.689. A carga líquida da forma *apo* em pH 7.4 é de -23e [69, 70]. Essa proteína possui dois domínios com duas estruturas *EF-hand* em cada um, ligando quatro íons de Ca^{+2} [4, 71].

O segmento de hélice que liga os dois domínios é bastante flexível, permitindo uma grande mudança conformacional quando da ligação dos cálcios. O rearranjo das hélices expõe certos resíduos hidrofóbicos, aumentando a afinidade desta proteína por uma série de proteínas alvo [58]. Essa grande flexibilidade (em relação à *Calbindina D_{9k}*), além de sua alta carga líquida, faz da CaM um excelente teste crítico para os modelos e as aproximações utilizadas no estudo de propriedades eletrostáticas da interação entre ligantes e biomoléculas. A forma *holo* da CaM é mostrada na figura 3.4.

Estudos realizados utilizando ^1H NMR, demonstram que o aumento da concentração salina produz uma diminuição da afinidade dessas proteínas pelo íon de Ca^{+2} [71, 54], ou seja, temos um efeito de sal no pK do cálcio. Podemos verificar esse comportamento através dos resultados mostrados na figura 3.5 (ver relação 3.11 para definição do ΔpK). Este comportamento sugere a existência de uma **blindagem eletrostática** produzida pelos íons

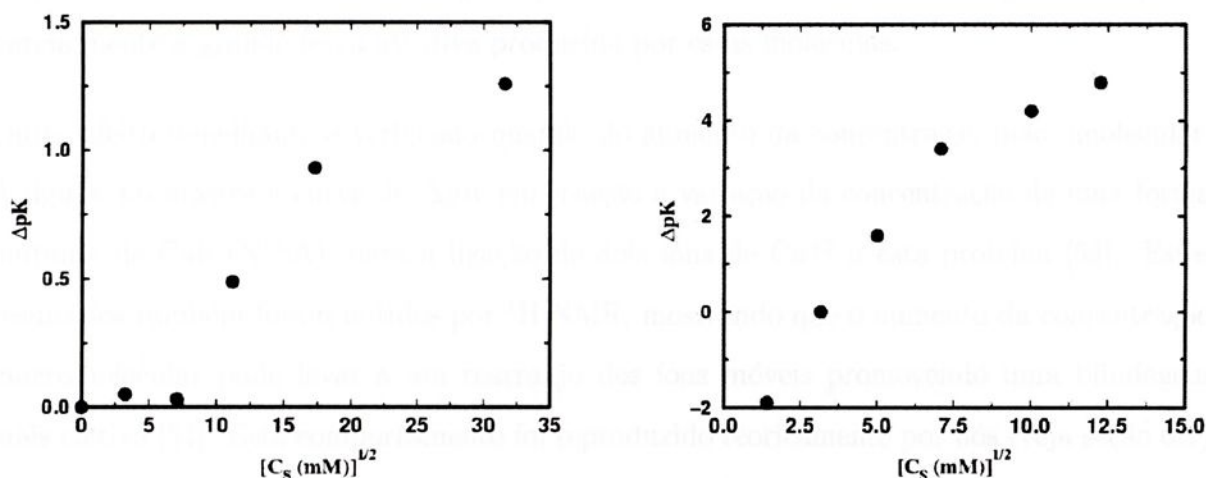


Figura 3.5: Resultados experimentais de ΔpK devido a variação da concentração salina para a ligação de dois íons Ca^{+2} à uma forma mutante da Cab (N56A) (gráfico da esquerda), e de quatro íons Ca^{+2} à CaM (gráfico da direita), obtidos por ^1H NMR. Para a Cab, foi utilizada uma concentração macromolecular de 3,3 mM, e uma concentração salina nula de referência [54]. Já para a CaM, a concentração macromolecular foi de 20 μM , e concentração salina de 10 mM foi utilizada com referência [71].

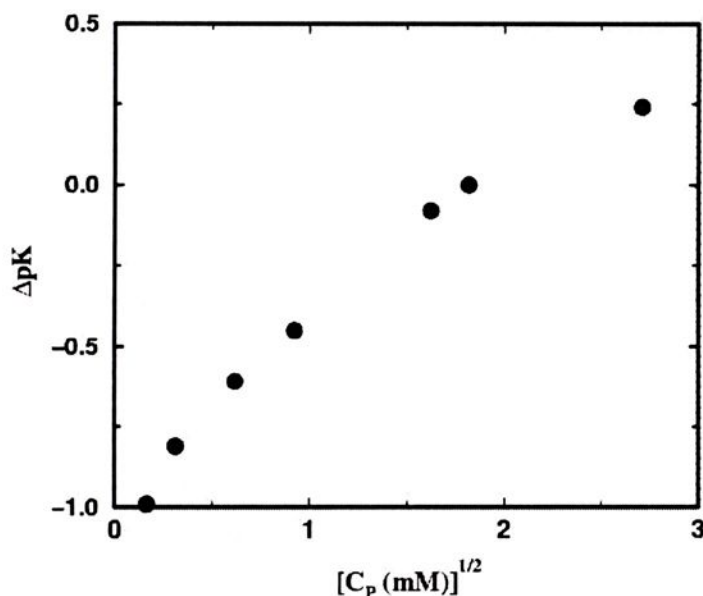


Figura 3.6: Resultados experimentais de ΔpK devido a variação da concentração macromolecular para a ligação de dois íons Ca^{+2} à uma forma mutante da Cab (N56A), obtidos por ^1H NMR. Foi utilizado um regime de concentração salina nula, com concentração macromolecular de referência de 3,3 mM [54].

livres da solução. Nesta situação, os contraíons móveis da solução se distribuem ao redor da proteína, de forma que a interação repulsiva entre esses contraíons e o ligante compense parcialmente a grande força atrativa produzida por estas moléculas.

Outro efeito semelhante é verificado quando do aumento da concentração macromolecular. A figura 3.6 mostra a curva de ΔpK em relação a variação da concentração de uma forma mutante de Cab (N56A), para a ligação de dois íons de Ca^{+2} a esta proteína [54]. Estes resultados também foram obtidos por ^1H NMR, mostrando que o aumento da concentração macromolecular pode levar a um rearranjo dos íons móveis promovendo uma blindagem mais efetiva [54]. Este comportamento foi reproduzido teoricamente por nós (veja seção 6.1) através de um modelo baseado em interações eletrostáticas, confirmando mais uma vez, a importância dessas interações na formação dos complexos moleculares estudados.



Figura 3.6. Representação da distribuição de íons móveis em torno de uma proteína. A esfera representa o volume de solução. Os íons positivos representam a proteína e os íons negativos representam os contraíons móveis.

Capítulo 4

Modelo do Sistema

Baseado na consideração feita no capítulo anterior, trataremos cada espécie neutra formada por uma macromolécula e pelos íons de sal e contraíons distribuídos ao seu redor, de maneira independente, utilizando o *Modelo da Célula* [49, 60, 72, 73]. Neste modelo, consideraremos cada uma dessas espécies neutras confinadas em uma célula de raio R_c , definido de acordo com a concentração macromolecular, com a macromolécula rígida⁶ em seu centro, de acordo com a figura 4.1.

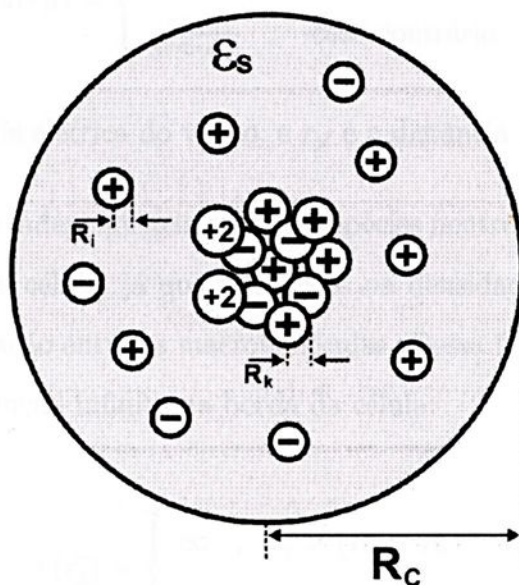


Figura 4.1: Representação esquemática do Modelo da Célula já incorporando uma biomolécula. Os círculos com inscrição “+2.” representam os ligantes.

⁶ Embora a macromolécula esteja sendo tratada como rígida, em um trabalho paralelo realizado em conjunto com o aluno de doutorado Carlos Emídio Sampaio Nogueira, pretendemos verificar o efeito da flexibilidade da proteína, através do cálculo das propriedades termodinâmicas relacionadas a interação entre íons cálcio e proteínas da família da *Calmodulina*, utilizando várias configurações geradas através de simulações por *Dinâmica Molecular*.

Além da baixa concentração macromolecular, uma outra condição para que essas espécies neutras se comportem independente das outras, é a presença apenas de íons monovalentes na solução eletrolítica. Teríamos problemas com a utilização de soluções assimétricas, uma vez que trabalhos têm mostrado que o forte efeito de correlação entre íons multivalentes pode causar atração entre as macromoléculas [49].

A solução eletrolítica foi tratada no *nível de McMillan-Mayer* [25, 26], de acordo com o *Modelo Primitivo Restrito* [23, 74], onde os íons móveis de valência z_j foram representados por esferas rígidas de raio $R_j = 2 \text{ \AA}$ e carga central $Z_j = z_j e = \pm 1e$, sendo e a carga elementar, imersos em um solvente representado apenas por sua constante dielétrica ($\epsilon_s = 78,7$ na temperatura de 298 K). Esta aproximação é usual e comumente usada no estudo de soluções eletrolíticas [75, 76, 77].

A interação entre quaisquer duas partículas j e l carregadas do sistema, é dada simplesmente de acordo com a *Lei de Coulomb* [78], adicionando uma restrição para evitar o colapso entre elas:

$$u(r_{jl}) = \begin{cases} \infty & , \quad r_{jl} < (R_j + R_l) \\ \frac{z_j z_l e^2}{4\pi\epsilon_0\epsilon_s r_{jl}} & , \quad \text{caso contrário} \end{cases} \quad (4.1)$$

Onde ϵ_0 é a permissividade elétrica do vácuo, e r_{jl} é a distância entre as duas partículas.

Além disso, a condição de independência entre as espécies neutras, sugere que os íons devem ficar confinados dentro da célula, já que flutuações na densidade de carga das atmosferas iônicas podem causar atração entre as macromoléculas. Dessa forma, os íons móveis devem ser submetidos a um potencial infinito na borda da célula:

$$v(\vec{r}_j) = \begin{cases} \infty & , \quad r_j > (R_c - R_i) \\ 0 & , \quad \text{caso contrário} \end{cases} \quad (4.2)$$

Dada uma solução salina simétrica monovalente (-1:1), com concentração C_s , o número de íons positivos (N_+) e negativos (N_-) em solução, são dados por:



$$N_+ = \left(\frac{4}{3}\pi R_c^3\right) 1000 N_a C_s + |z_p| \quad (4.3)$$

$$N_- = \left(\frac{4}{3}\pi R_c^3\right) 1000 N_a C_s \quad (4.4)$$

Onde N_a é o número de Avogrado, e Z_p é a “carga líquida” (soma das cargas parciais de todos os átomos) da macromolécula. O termo adicional na relação 4.3, representa o número de contraíons adicionados ao sistema para manter a eletroneutralidade da célula, uma vez que todas as proteínas que estamos considerando possuem carga líquida negativa, e o termo $1000 N_a$ é utilizado para conversão de unidades para o sistema internacional, uma vez que estamos considerando as concentrações em unidades de molar, e o raio R_c em metros.

Alguns exemplos do número de partículas utilizados em nossos cálculos podem ser obtidos na tabela abaixo:

Concentração Macromolecular (μM)	Raio da Célula ($\text{\AA}$)	Concentração Salina (mM)	Número de ânions (N_-)	Número de cátions (N_+)
20	270,63	1	50	54
20	270,63	25	1250	1254
20	270,63	100	4999	5003
200	125,62	1	5	9
200	125,62	25	125	129
200	125,62	100	500	504
2000	58,31	1	1	5
2000	58,31	25	13	17
2000	58,31	100	50	54

Tabela 4.1: Número de ânions obtidos apartir da relação 4.3 e 4.4, para algumas concentrações de sal e de proteína, utilizando uma macromolécula de carga líquida $Z_p = -4e$.

Sendo nosso sistema composto por um total de $N = N_p + N_s$ partículas, onde N_p é o número de átomos da proteína, e $N_s = N_+ + N_-$ é o número de íons móveis, a energia potencial total



do sistema é dada por:

$$U(\vec{r}_1, \dots, \vec{r}_N) = \sum_{j=1}^{N_s} v(\vec{r}_j) + \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^N u(r_{jl}) \quad (4.5)$$

4.1 Modelo da Macromolécula

O macroíon foi representado através de sua estrutura tridimensional, considerando seus detalhes atomísticos, onde cada um dos N_p átomos foi representado por uma pequena esfera rígida de raio $R_k = 2 \text{ \AA}$, com uma carga central Z_k . A posição tridimensional de cada átomo foi obtida da estrutura cristalográfica das proteínas em estudo: PDB-ID 3icb para a *Calbindina D_{9k}* (Cab) [79] e PDB-ID 3cln para *Calmodulina* (CaM) e seus fragmentos [80].

Ambas as estruturas utilizadas foram obtidas utilizando métodos de difração de raio-X. A utilização desta técnica experimental traz dificuldades na localização dos átomos de hidrogênio da molécula, uma vez que estes, quando ligados a outros átomos, não possuem sua densidade eletrônica centrada na posição do núcleo. Além disso, os átomos de hidrogênio tem uma amplitude vibracional muito maior do que a dos outros átomos, dificultando ainda mais sua localização [81]. Por isso, os hidrogênios ausentes foram adicionados através de critérios geométricos, com o auxílio do pacote computacional para simulação biomolecular GROMACS [82, 83], o qual segue o critério abaixo [83]:

Um Hidrogênio Simples: Supondo um átomo de hidrogênio H que se liga a um determinado átomo i , como na hidroxila, por exemplo. Este átomo i é ligado em um outro átomo j , que por sua vez é ligado a um átomo k . Este hidrogênio é adicionado a $0,1 \text{ nm}$ do átomo i , com o ângulo Hij de $109,5$ graus, e diedral $Hijk$ assumindo a conformação *trans*.

Um Hidrogênio Planar: No caso de um átomo de hidrogênio H ligado a algum átomo em um anel, ou ao nitrogênio da ligação peptídica, por exemplo, o hidrogênio se posiciona no plano formado pelos átomos i , j e k , onde j e k estão ligados em i , sendo adicionado a $0,1 \text{ nm}$ do átomo i , na linha bissetriz do ângulo jik , de forma que os ângulos Hij e Hik sejam maior que $90,0$ graus.

Dois Hidrogênios Planares: Dois hidrogênios H_1 e H_2 ligados a um determinado átomo i como no grupo NH_2 , os hidrogênios são adicionados a 0,1 nm do átomo i , de forma que os ângulos H_1ij e H_2ij assumam 120,0 graus, sendo o átomo i ligado ao átomo j , e os diedrais H_1ijk e H_2ijk assumam uma conformação *cis* e *trans*, respectivamente, onde o átomo k é ligado ao átomo j .

Um Hidrogênio Tetraédrico: No caso de grupos como o C_3CH , o hidrogênio se liga a um átomo i que também se liga aos átomos j , k e l , formando uma estrutura tetraédrica, de forma que o hidrogênio é adicionado a 0,1 nm do átomo i , tal que os ângulos H_1ij , H_2ik e H_3il assumam 109,5 graus.

Dois Hidrogênios Tetraédricos: Grupos como $C - CH_2 - C$, onde dois carbonos e dois hidrogênios se ligam a um outro carbono formando uma conformação tetraédrica, tem os hidrogênios adicionados a uma distância de 0,1 nm do átomo i , no plano formado pela bissetriz do ângulo ijk com os ângulos H_1ij , H_2ij e H_3ij assumindo 109,5 graus, sendo i o átomo no qual os outros quatro átomos estão ligados, e j e k os átomos que não são hidrogênios e estão ligados ao átomo i formando a estrutura tetraédrica.

Dois ou Três Hidrogênios Tetraédricos: Para grupos CH_2 e CH_3 , onde dois ou três átomos de hidrogênio são ligados a um determinado átomo i , que por sua vez se liga a um átomo j , assumindo uma conformação tetraédrica, os hidrogênios são adicionados a 0,1 nm do átomo i , de forma que os ângulos H_1ij , H_2ij e H_3ij formam 109,5 graus, com os diedrais H_1ijk assumindo uma conformação *trans* e H_3ijk assumindo um ângulo 240,0 graus além da conformação *trans*, sendo k o átomo que se liga ao átomo j .

Para determinação da constante dielétrica macromolecular, utilizamos o modelo do dielétrico contínuo e **uniforme** [18], ou seja, consideramos a constante dielétrica do interior macromolecular igual a do solvente ($\epsilon_i = \epsilon_s$). Tem se mostrado que essa aproximação é bastante eficiente na obtenção de resultados concordantes com resultados experimentais para a ligação de íons metálicos à proteínas [51, 84]. Além disso, um trabalho recente mostra resultados fisicamente impossíveis na predição de efeitos mutacionais para a ligação do Ca^{+2} à Cab, quando da utilização de constantes dielétricas diferentes para o interior e o exterior macromolecular [51].



Para a CaM foi necessário a inclusão manual dos quatro primeiros resíduos, bem como do último, através do pacote computacional *SwissPDB-Viewer* [85], já que a estrutura resolvida cristalograficamente não os possuíam, devido as incertezas nas suas posições [71]. Este pacote computacional obtém a estrutura de uma determinada proteína através de seu arquivo PDB, permitindo que se selecione um determinado resíduo existente, adicionando, a partir deste, um novo resíduo. As coordenadas dos átomos do resíduo incluído são determinadas de forma a minimizar a energia da macromolécula.

As coordenadas atômicas dos dois fragmentos foram obtidas da estrutura da CaM intacta, sendo os primeiros 74 resíduos correspondentes ao primeiro fragmento, com carga líquida total $Z_p = -12e$, e os resíduos de número 78 ao 148 correspondentes ao segundo fragmento, com carga líquida total $Z_p = -13e$ [71].

Como forma de generalizar nossa análise, utilizamos largamente outros dois modelos para a macromolécula, os quais podem ser vistos como uma simplificação do modelo proposto inicialmente (figura 4.2):

- Um modelo equivalente ao resolvido por *Tanford e Kirkwood* [86], onde a macromolécula é inserida em uma cavidade esférica de raio R_p , impenetrável aos íons móveis (figura 4.2b). Para a CaM, utilizamos $R_p = 38,00 \text{ \AA}$, enquanto para a Cab, $R_p = 18,05 \text{ \AA}$. O primeiro e segundo fragmento da CaM receberam $R_p = 26,41$ e $22,79 \text{ \AA}$, respectivamente.
- Foi também utilizado um outro modelo mais simplificado para o macroíon, desprezando seus detalhes atômicos e o representando por uma esfera rígida de raio R_p , com sua carga líquida Z_p concentrada pontualmente em seu centro (figura 4.2c). Os valores de R_p para esse modelo são equivalentes aos do modelo anterior.

4.2 Campos de Força

Para o **Modelo Detalhado** e **Modelo Detalhado Esférico**, onde os átomos da macromolécula são descritos explicitamente, necessitamos conhecer a carga residual Z_k de cada um



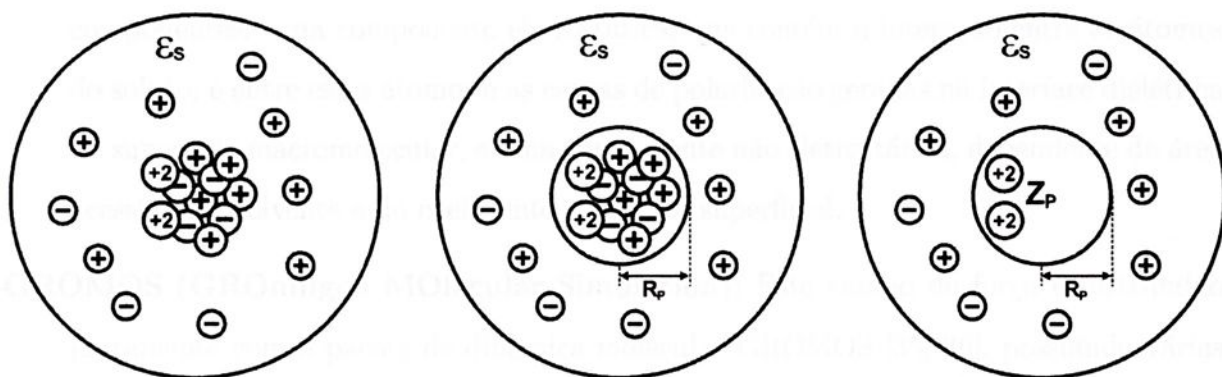


Figura 4.2: Modelos utilizados para a representação do macroíon. A figura da esquerda (figura 4.2a) representa o modelo mais realista para a macromolécula, onde sua superfície é definida pelos raios dos átomos que se localizam na sua superfície. Chamaremos este modelo de **Modelo Detalhado**. Na figura central (figura 4.2b), a macromolécula contém uma descrição atomística, entretanto, sua superfície é considerada esférica. Chamaremos esse modelo de **Modelo Detalhado Esférico**. Já na figura da direita (figura 4.2c), temos a representação mais simplificada para o macroíon, representando-o por uma esfera rígida impenetrável à solução. Este modelo será denominado **Modelo Esférico**.

deles, de forma que o potencial eletrostático gerado pela macromolécula possa ser calculado através da lei de *Coulomb* presente na relação 4.1.

Como cargas atômicas não são observáveis quânticos, sua obtenção através de cálculos teóricos realizados no *nível de Schrödinger* sofrem de arbitrariedades, sendo altamente dependentes da metodologia utilizada [87]. Devido a isso, essas cargas podem constituir parâmetros de um modelo, e serem obtidas de forma a garantir a concordância com resultados empíricos. Entretanto, essas cargas não são obtidas de maneira única, sendo dependentes do modelo utilizado e da propriedade física utilizada para comparação [38].

Neste trabalho, utilizamos as cargas parciais propostas por três campos de força bastante conhecidos da literatura:

PARSE (PARAMeters for Solvation Energy): Este campo de força [31, 32] é parametrizado de forma que a energia livre de solvatação de um grupo de 67 moléculas orgânicas, obtidas experimentalmente, concorde com os valores calculados através de um modelo baseado na aproximação contínua do solvente. Neste modelo, as moléculas foram tratadas segundo sua estrutura tridimensional, onde cada átomo foi considerado uma esfera rígida com carga central z_k . Esta molécula foi inserida em uma cavidade de baixa constante dielétrica ($\epsilon_i = 2$), cuja forma é definida de acordo com a superfície molecular. A energia livre calculada a partir deste modelo foi particionada em duas

componentes: uma componente eletrostática, que contém a interação entre os átomos do soluto, e entre esses átomos e as cargas de polarização geradas na interface dielétrica da superfície macromolecular, e uma componente não eletrostática, dependente da área acessível ao solvente e do coeficiente de tensão superficial.

GROMOS (GRONingen MOlecular Simulation): Este campo de força é distribuído juntamente com o pacote de dinâmica molecular GROMOS [35, 36], possuindo várias versões. A versão utilizada por nós é 43a1 do GROMOS96.

Campo de Força Reduzido: No campo de força Reduzido, as cargas são atribuídas de forma mais simplificada, apenas aos átomos mais polares de cada aminoácido carregado. Assim, atribui-se uma carga de $-0,5e$ para cada oxigênio carboxílico localizado no *aspartato*, *glutamato* e no aminoácido C-terminal, uma carga unitária positiva para o ξ -nitrogênio da *lisina* e para o aminoácido N-terminal, e $0,33e$ para os ϵ_1 e ϵ_2 nitrogênios da *arginina*. Embora as moléculas utilizadas por nós não contenham *histidinas*, critério semelhante pode ser seguido para atribuição das cargas a seus átomos [88].

Uma outra aproximação mais grosseira para a representação da macromolécula, que chamaremos de *Aproximação Dipolar*, atribui a carga líquida positiva e negativa às coordenadas (x^+, y^+, z^+) e (x^-, y^-, z^-) , respectivamente, formando um dipolo. Essas coordenadas são calculadas da seguinte forma:

$$x^\pm = \frac{1}{Q^\pm} \sum_{j^\pm} x_j Z_j \quad y^\pm = \frac{1}{Q^\pm} \sum_{j^\pm} y_j Z_j \quad z^\pm = \frac{1}{Q^\pm} \sum_{j^\pm} z_j Z_j \quad (4.6)$$

onde $Q^\pm$ é a carga total positiva (índice +) ou negativa (índice -). A somatória se aplica a todas as j^+ cargas positivas ou j^- cargas negativas da molécula, sendo x_j , y_j e z_j as coordenadas de cada átomo com carga Z_j .

Uma outra maneira bastante promissora de atribuir cargas parciais aos átomos da proteína é a utilização do método de *equalização de eletronegatividades* [89, 90]. Como o campo de força PARSE, esse método também foi desenvolvido para cálculos de energia de solvatação utilizando modelos com solvente contínuo. Entretanto, tem apresentado resultados mais



precisos, se tornando adequado para estudos desta natureza. Dessa forma, futuramente pretendemos, em colaboração com o Dr. Luís Gustavo Dias, aplicar esta metodologia para estudos relacionados a este trabalho.

Capítulo 3

Solução do Modelo

A. Mecânica Estatística Clássica

Para a descrição estatística de um sistema físico, é necessário conhecer as propriedades físicas do sistema, como a energia, o momento, a carga, etc. A mecânica estatística clássica é a teoria que descreve o comportamento estatístico de um sistema físico em equilíbrio térmico. Ela é baseada na mecânica clássica e na teoria da probabilidade.

$$A = \frac{1}{N} \sum_{i=1}^N \frac{U_i}{k_B T} \quad (3.1)$$

Onde A é a entropia por partícula, N é o número de partículas, U_i é a energia da partícula i , k_B é a constante de Boltzmann e T é a temperatura. A equação (3.1) pode ser escrita como $A = \frac{1}{N} \sum_{i=1}^N \ln \Omega_i$, onde Ω_i é o número de estados acessíveis à partícula i . A mecânica estatística clássica é baseada na mecânica clássica e na teoria da probabilidade.

Em geral, a mecânica estatística clássica é aplicada a sistemas em equilíbrio térmico. Ela é baseada na mecânica clássica e na teoria da probabilidade. A mecânica estatística clássica é a teoria que descreve o comportamento estatístico de um sistema físico em equilíbrio térmico.

Capítulo 5

Solução do Modelo

A Mecânica Estatística Clássica

Para a descrição microscópica do nosso sistema através do modelo descrito no capítulo anterior, a Mecânica Estatística fornece as ferramentas necessárias para a obtenção da distribuição dos íons móveis, e consequentemente das propriedades termodinâmicas de interesse [18]. Dessa forma, uma determinada propriedade macroscópica A , é obtida através da média sobre todos os i estados microscópicos do sistema [23]:

$$A = \frac{\sum_i A_i P_i}{\sum_i P_i} \quad (5.1)$$

Onde P_i é a probabilidade do sistema assumir um determinado estado i . Para o nosso sistema, onde podemos utilizar uma abordagem clássica, esses estados são completamente determinados pela posição $\vec{r}_j$ das N partículas, e pelos momentos lineares $\vec{p}_j$ das N_s partículas móveis, assim, $P_i = P(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})$. Estaremos considerando que as partículas móveis da solução serão representadas pelos índices 1 até N_s , e os átomos da macromolécula, pelos índices $N_s + 1$ até N .

De acordo com o modelo que adotamos, o volume da célula e o número de partículas no seu interior serão constantes. Além disso, de forma a manter uma compatibilidade com a situação experimental, consideraremos nosso sistema em contato com um reservatório térmico man-



tendo a temperatura constante. Nesta situação (correspondente ao *ensemble canônico* [23]), a probabilidade do sistema assumir um estado determinado pela posição $\vec{r}_1$ e momento linear $\vec{p}_1$ para a partícula 1, posição $\vec{r}_2$ e momento linear $\vec{p}_2$ para a partícula 2, e assim sucessivamente, é representada, em função da energia total do sistema $E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})$, através da relação [23]:

$$P(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s}) = \frac{\exp \left[-\frac{E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})}{kT} \right]}{\int \dots \int \exp \left[-\frac{E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_N d\vec{p}_1 \dots d\vec{p}_{N_s}} \quad (5.2)$$

Esta relação pode ser obtida em termos da *função partição canônica* $Q(N, V, T)$, onde o termo a frente da integral é uma constante de proporcionalidade responsável em garantir a concordância com a *função partição quântica* no limite de $h \rightarrow 0$, sendo h a constante de Plank [24].

$$Q(N, V, T) = \frac{1}{[N_+]! [N_-]! h^{3N_s}} \int \dots \int \exp \left[-\frac{E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_N d\vec{p}_1 \dots d\vec{p}_{N_s} \quad (5.3)$$

A função partição estabelece uma relação entre as características microscópicas do sistema e suas propriedades termodinâmicas macroscópicas, uma vez que esta função está diretamente relacionada a energia livre de *Helmholtz* [23]:

$$A = -kT \ln Q(N, V, T) \quad (5.4)$$

Como em nosso caso a energia potencial U independe dos momentos lineares das partículas móveis, podemos considerar $E(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s}) = U(\vec{r}_1, \dots, \vec{r}_N) + K(\vec{p}_1, \dots, \vec{p}_{N_s})$, onde K , a energia cinética total do sistema, é dada por $K = \sum_{j=1}^{N_s} \frac{p_j^2}{2m_j}$, sendo m_j a massa da partícula j . Dessa forma, podemos integrar a relação 5.3 nos momentos, obtendo:

$$Q(N, V, T) = \frac{Z(N, V, T)}{[N_+]! [N_-]! \Lambda^{3N_s}} \quad (5.5)$$

Onde Λ é o comprimento de onda térmico, dado por $\Lambda = \frac{h}{(2\pi mkT)^{1/2}}$, e $Z(N, V, T)$, a *integral de configuração*, é definida como [18]:

$$Z(N, V, T) = \int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_N \quad (5.6)$$

Sendo assim, a probabilidade $P^j(\vec{r}_j)$ de encontrar uma dada partícula j em uma determinada posição $\vec{r}_j$, independente do seu momento linear e da posição e momento linear de todas as outras partículas, é dada da seguinte forma:

$$P^j(\vec{r}_j) = \frac{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{j-1} d\vec{r}_{j+1} \dots d\vec{r}_N}{Z(N, V, T)} \quad (5.7)$$

Baseado nessas relações, podemos definir certas grandezas que são muito úteis na análise da distribuição dos íons móveis ao redor da macromolécula, como a *função distribuição radial* e o *potencial da força média*, além de propriedades termodinâmicas necessárias para o cálculo da constante de ligação. Entretanto, o cálculo da integral de configuração, necessária para determinação dessas grandezas, é impraticável analiticamente, necessitando de métodos numéricos ou de aproximações que viabilizem sua solução [18].

5.1 Função Distribuição Radial

Quando utilizada em conjunto com o modelo da célula, a *função distribuição radial* $g_k(\vec{r})$ de uma determinada espécie k (em nosso caso $k = 2$, pois temos apenas duas espécies de partículas móveis: cátions e ânions da solução), é definida como a razão entre a probabilidade $P_k(\vec{r})$ de encontrar uma das N_k partículas da espécie k na posição $\vec{r}$, e a probabilidade P_k^{id} de encontrar uma das N_k partículas na posição $\vec{r}$, estando essas partículas não submetidas a qualquer interação (sistema ideal) [23]:

$$g_k(\vec{r}) = \frac{P_k(\vec{r})}{P_k^{id}(\vec{r})} = \frac{V}{N_k} P_k(\vec{r}) \quad (5.8)$$



Segundo essa definição, a *função distribuição radial* corrige a distribuição dos íons ao redor da molécula, incluindo os efeitos de correlação iônica responsáveis pelo desvio do comportamento ideal.

Para proceder com a dedução de $g_k(\vec{r})$, seja $P_k^j(\vec{r}_j)$ a probabilidade de encontrar uma determinada partícula j da espécie k na posição $\vec{r}_j$, independente da posição das outras partículas:

$$P_k^j(\vec{r}_j) = \frac{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{j-1} d\vec{r}_{j+1} \dots d\vec{r}_N}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_N} \quad (5.9)$$

Assim, a probabilidade de encontrar qualquer uma das N_k partículas da espécie k na posição $\vec{r}_j$ é dada por $P_k(\vec{r}_j) = N_k P_k^j(\vec{r}_j)$, sendo portanto possível, utilizando a relação 5.8, obter $g_k(\vec{r})$ (a notação foi modificada fazendo $\vec{r}_j \equiv \vec{r}$):

$$g_k(\vec{r}) = V \frac{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{j-1} d\vec{r}_{j+1} \dots d\vec{r}_N}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_N} \quad (5.10)$$

A figura 5.1 mostra a característica estrutural da função distribuição radial. Para distâncias pequenas $g_k(\vec{r}) \rightarrow 0$, devido as restrições que evitam a sobreposição de duas partículas (potencial de esfera rígida). Já para grandes distâncias, os desvios do comportamento ideal se tornam desprezíveis, e consequentemente $g_k(\vec{r}) \rightarrow 1$ [18, 23].

O gráfico da esquerda, onde as partículas interagem apenas através de um potencial de esfera-rígida, contém dois picos que indicam a existência de camadas onde as partículas têm uma grande probabilidade de serem encontradas (essas camadas são chamadas *camadas de coordenação* [91]). Entre essas duas camadas há uma vale, onde a probabilidade de se encontrar partículas é reduzida devido a tendência das partículas se excluírem. Sendo assim, podemos inferir que a existência de um raio finito para as partículas (desvio da situação ideal) faz com que certas regiões próximas ao macroíon sejam “estatisticamente privilegiadas”, devido a existência de uma estrutura local mais ordenada.

O gráfico da direita mostra os resultados referentes a um sistema formado por um macroíon esférico de carga $-8e$ e raio $14 \text{ \AA}$, fixo no centro de uma célula esférica de raio $125 \text{ \AA}$ (con-

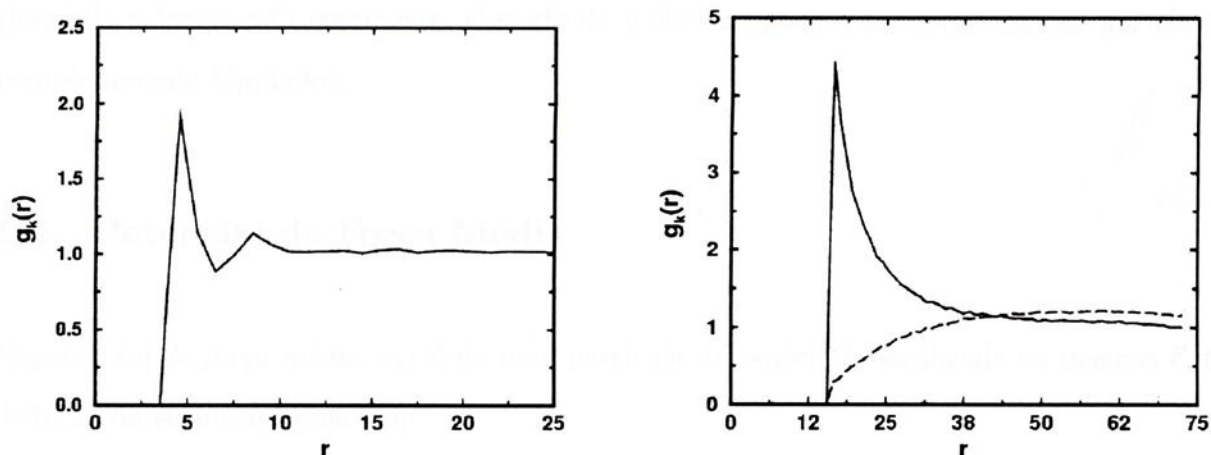


Figura 5.1: Função distribuição radial $g_k(\vec{r})$ em função da distância $|\vec{r}|$ a uma partícula fixa no centro de uma célula esférica, obtida através do programa desenvolvido por nós. O gráfico da **esquerda** mostra resultados para um sistema formado por 884 esferas rígidas de raio 2 Å, confinados em uma célula de raio 30 Å, com uma dessas esferas fixa no centro dessa célula. O gráfico da **direita** mostra a função distribuição radial para os cátions e ânions de uma solução eletrolítica (-1:1) em 50 mM, confinada em uma célula de raio 125 Å, com macroíon central de carga -8e e raio 14 Å. A constante dielétrica do solvente foi $\epsilon_s = 78,7$, em temperatura de 298 K. A linha contínua representa os cátions, enquanto a linha tracejada, os ânions.

centração macromolecular de 200 μM), e imerso em uma solução eletrolítica simétrica monovalente (-1:1) modelada segundo o *modelo primitivo restrito* [23, 92], em concentração salina de 50 mM. Devido ao efeito do grande campo elétrico produzido pelo macroíon, as camadas de coordenação desaparecem criando uma nuvem iônica bastante “densa” ao redor do macroíon, responsável pela blindagem do seu potencial eletrostático. Podemos observar um maior acúmulo de coíons em regiões próximas à borda da célula, indicando a existência de um **efeito de borda**, decorrente do potencial descrito pela relação 4.2. Nesta situação, os coíons, fortemente repelidos pelo macroíon central, tendem a se localizar em regiões afastadas do macroíon, se localizando próximos a borda da célula devido ao raio finito da mesma.

Dessa forma, considerando que em um sistema ideal as partículas são uniformemente distribuídas, podemos inferir que a concentração média $\langle C_k(\vec{r}) \rangle$ de partículas da espécie k em uma dada posição $\vec{r}$, pode ser dada por:

$$\langle C_k(\vec{r}) \rangle = C_s g_k(\vec{r}) \quad (5.11)$$

Obviamente, a relação acima é válida apenas para soluções eletrolíticas simétricas, visto que para soluções não simétricas as espécies possuem concentrações diferentes no *bulk* da solução

(local da solução suficientemente afastado da molécula, onde o potencial gerado por ela é completamente blindado).

5.2 Potencial de Força Média

O *potencial de força média* $w_k(\vec{r})$ de uma partícula da espécie k , localizada na posição $\vec{r}$, é definido da seguinte forma [23]:

$$g_k(\vec{r}) = \exp \left[-\frac{w_k(\vec{r})}{kT} \right] \quad (5.12)$$

Tomando o logaritmo de ambos os lados, substituindo a relação 5.10, e tomando o gradiente com respeito a posição $\vec{r}$, obtemos:

$$-\vec{\nabla} w_k(\vec{r}) = \frac{\int \dots \int \left[-\vec{\nabla} U(\vec{r}_1, \dots, \vec{r}_N) \right] \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{j-1} d\vec{r}_{j+1} \dots d\vec{r}_N}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{j-1} d\vec{r}_{j+1} \dots d\vec{r}_N} \quad (5.13)$$

O lado direito da equação 5.13 é a força média agindo em uma dada partícula da espécie k , estando essa partícula na posição $\vec{r}$. Dessa forma, $w_k(\vec{r})$ é o potencial cujo gradiente negativo é a força média, ou seja, é a energia potencial de interação média entre uma partícula da espécie k , localizada na posição $\vec{r}$, e a macromolécula.

Em concentrações muito baixas, $w_k(\vec{r}) \approx \sum_{m=N_s+1}^N u(|\vec{r} - \vec{r}_m|)$, já que o pequeno número de partículas móveis influenciam muito pouco essa interação [23]. Com o aumento da concentração, $w_k(\vec{r})$ é influenciado pela posição média de todas as partículas móveis, justificando a relação entre $w_k(\vec{r})$ e $g_k(\vec{r})$, uma vez que a *função distribuição radial* é diretamente relacionada com a posição média das partículas.



5.3 A Equação de Poisson-Boltzmann

Dada a relação 5.11, o cálculo de $g_k(\vec{r})$ é suficiente para obtermos a distribuição dos íons móveis ao redor do macroíon. Entretanto, o cálculo exato das integrais existente nas relações 5.10 e 5.13 é extremamente complicado, necessitando de aproximações que viabilizem sua obtenção. A principal aproximação utilizada é a substituição do *potencial de força média* pela *energia potencial eletrostática média* $z_k e \langle \varphi(\vec{r}) \rangle$ [18, 23, 50, 52]:

$$g_k(\vec{r}) = \exp \left[-\frac{z_k e \langle \varphi(\vec{r}) \rangle}{kT} \right] \quad (5.14)$$

O potencial eletrostático médio $\langle \varphi(\vec{r}) \rangle$, em uma posição $\vec{r}$, pode ser escrito como:

$$\langle \varphi(\vec{r}) \rangle = \frac{\int \dots \int \varphi(\vec{r}) \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{N_s}}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{N_s}} \quad (5.15)$$

Aplicando o *Laplaciano* em relação a posição $\vec{r}$, verificamos que $\langle \varphi(\vec{r}) \rangle$ satisfaz a equação de *Poisson* [23]:

$$\nabla^2 \langle \varphi(\vec{r}) \rangle = \frac{\int \dots \int \nabla^2 \varphi(\vec{r}) \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{N_s}}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_{N_s})}{kT} \right] d\vec{r}_1 \dots d\vec{r}_{N_s}} = -\frac{\langle \rho(\vec{r}) \rangle}{\epsilon_0 \epsilon_s} \quad (5.16)$$

Dessa forma, conhecendo a distribuição média dos íons, e consequentemente a densidade de carga média $\langle \rho(\vec{r}) \rangle$, podemos resolver a equação de *Poisson* e obter o potencial eletrostático médio em qualquer posição $\vec{r}$. A densidade de carga média pode ser dada por:

$$\langle \rho(\vec{r}) \rangle = \rho^f(\vec{r}) + \rho^m(\vec{r}) = \rho^f(\vec{r}) + \sum_k 1000 \langle C_k(\vec{r}) \rangle N_a e z_k \quad (5.17)$$

Onde $\rho^f(\vec{r})$ é a densidade de carga fixa da macromolécula, e $\rho^m(\vec{r})$ é a densidade de carga móvel relacionada aos íons da solução. O termo $1000 N_a$ é utilizado para converter a densidade de carga móvel para unidades do sistema internacional (estamos considerando as concentrações em unidades de molar).

Substituindo $\langle \rho(\vec{r}) \rangle$ na relação 5.16, obtemos a *Equação de Poisson-Boltzmann* (PB) [18, 23, 52, 50]:

$$\nabla^2 \varphi(\vec{r}) = -\frac{\rho^f(\vec{r})}{\epsilon_0 \epsilon_s} - \frac{1}{\epsilon_0 \epsilon_s} \sum_k \rho_k(bulk) \exp \left[-\frac{z_k e \varphi(\vec{r})}{kT} \right] \quad (5.18)$$

Onde $\rho_k(bulk) = 1000 C_s N_a e z_k$ é a densidade de carga produzida por íons da espécie k , no *bulk* da solução. A notação foi simplificada para $\langle \varphi(\vec{r}) \rangle \equiv \varphi(\vec{r})$.

A substituição do *potencial de força média* pela *energia potencial eletrostática média* desconsidera as interações de *esfera-rígida* existente entre as partículas, ou seja, os íons móveis são tratados como cargas pontuais podendo se localizar mais próximos um do outro. Assim, podemos ter um maior acúmulo dos íons na superfície da macromolécula prevendo uma maior “blindagem eletrostática” [49]. Versões modificadas da equação de PB [93, 94] tem conseguido eliminar esta limitação, incorporando o efeito do tamanho finito dos íons móveis. Entretanto, neste trabalho estamos utilizando apenas a versão tradicional da equação de PB, uma vez que nossa ênfase está nos aspectos **eletrostáticos** da interação entre íons e biomoléculas, sendo interessante a utilização de modelos com um número mínimo de parâmetros.

Além disso, essa aproximação negligencia os efeitos produzidos pela correlação explícita entre os íons [49]. Trabalhos mostram que esses efeitos são importantes para solução compostas por íons bivalentes ([49] e referências aí citadas).

A equação de PB é uma equação diferencial parcial não linear, cuja solução é extremamente complicada. Analiticamente, existe solução apenas para casos muito simples, como sistemas dotados de simetria cilíndrica com uma única espécie iônica na solução [95], e sistemas com simetria plana contendo misturas de íons mono e bivalentes [96]. Mesmo os métodos numéricos disponíveis muitas vezes carecem de estabilidade, sendo a solução muito dependente de parâmetros inerentes desses métodos.



5.3.1 Aproximação de Debye-Hückel

A solução da equação de PB se torna bem complicada devido a sua não linearidade. *Debye* e *Hückel* propuseram a linearização desta equação através da expansão em série de *Taylor* do termo exponencial da distribuição iônica, e da utilização apenas do termo de primeira ordem [18, 23, 52]:

$$\exp \left[-\frac{z_k e \varphi(\vec{r})}{kT} \right] \approx 1 - \frac{z_k e \varphi(\vec{r})}{kT} \quad (5.19)$$

Essa aproximação se torna razoável para $kT \gg z_k e \varphi(\vec{r})$, ou seja, para sistemas composto de macromoléculas com baixa carga [52], ou suficientemente “blindadas”, de forma que a energia potencial eletrostática em qualquer ponto da solução seja desprezível em relação a energia térmica.

Substituindo a relação 5.19 na equação 5.18, obtemos:

$$\nabla^2 \varphi(\vec{r}) = -\frac{\rho^f(\vec{r})}{\epsilon_0 \epsilon_s} - \frac{1}{\epsilon_0 \epsilon_s} \sum_k \rho_k(bulk) + \frac{1}{\epsilon_0 \epsilon_s} \sum_k \frac{\rho_k(bulk) z_k e \varphi(\vec{r})}{kT} \quad (5.20)$$

A somatória do segundo termo se anula devido a condição de eletroneutralidade do sistema ($\sum_k \rho_k(bulk) = 0$). Já o terceiro termo pode ser escrito como:

$$\frac{1}{\epsilon_0 \epsilon_s} \sum_k \frac{\rho_k(bulk) z_k e \varphi(\vec{r})}{kT} = \left[\frac{e^2}{kT \epsilon_0 \epsilon_s} \sum_k n_k z_k^2 \right] \varphi(\vec{r}) = \kappa^2 \varphi(\vec{r}) \quad (5.21)$$

Onde $n_k = 1000 C_s N_a$ é a densidade numérica da espécie k no *bulk* da solução, e κ^{-1} é denominado **comprimento de Debye** [23, 97].

O termo κ^{-1} tem unidade de comprimento e está relacionada com a efetividade da blindagem produzida pela solução eletrolítica. O aumento da concentração salina causa um decréscimo nesta grandeza, sugerindo que podemos interpretá-la como um parâmetro de escala que indica a distância em que o potencial produzido pelo macroíon é efetivamente blindado [18, 23].

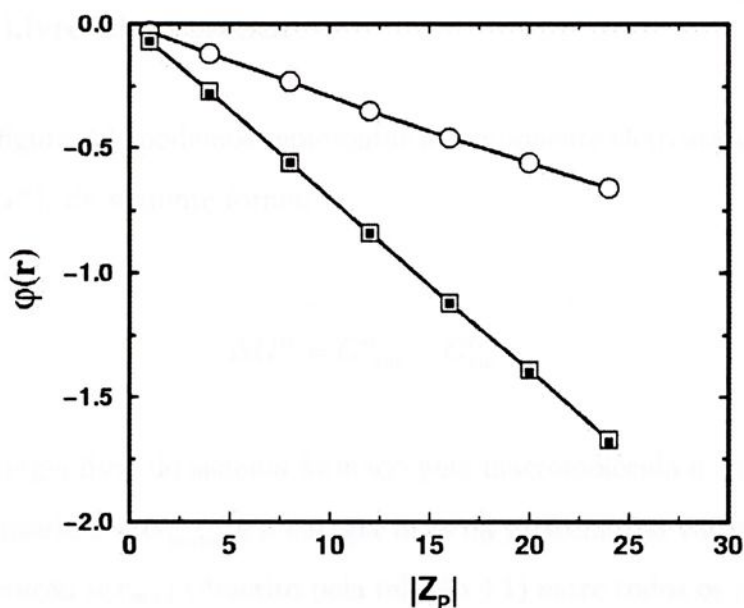


Figura 5.2: Potencial eletrostático, em unidade de $\frac{kT}{e}$ calculado a 2 Å da superfície de uma esfera de raio 14 Å e carga central $-|Z_p|$. Estes resultados foram obtidos através da solução numérica da equação de PB, em um regime de concentração salina de 50 mM ($\kappa^{-1} = 13,63$ Å). As linhas representam os resultados obtidos com o pacote computacional APBS [98], utilizando o método de *elementos finitos multigrid* [99, 100]. Os círculos e quadrados conectados a essas linhas representam a solução da equação não linear e linear de PB, respectivamente. Os quadrados negros representam resultados obtidos com o pacote computacional MEAD [33], utilizando o método de *diferenças finitas* [27].

Substituindo a relação 5.21 na expressão 5.20, obtemos:

$$\nabla^2 \varphi(\vec{r}) = -\frac{\rho^f(\vec{r})}{\epsilon_0 \epsilon_s} + \kappa^2 \varphi(\vec{r}) \quad (5.22)$$

A figura 5.2 mostra a variação do potencial eletrostático com o aumento da carga de um macroíon imerso em uma solução eletrolítica. Podemos observar que no regime de baixa carga Z_p , as soluções obtidas com as formas linear e não linear da equação de PB têm boa concordância, devido ao baixo potencial gerado pelo macroíon, que faz com que a aproximação representada pela relação 5.19 seja razoável. Com o aumento da carga do macroíon, essa aproximação se torna precária, prevendo uma densidade iônica, e consequentemente uma blindagem eletrostática, menor do que a prevista pela solução não linear.

5.3.2 Energia Livre Eletrostática

De acordo com a figura 3.1, podemos representar a componente eletrostática da energia livre de solvatação (ΔG^{el}), da seguinte forma:

$$\Delta G^{el} = G_{conc}^{el} - G_{vacuo}^{el} \quad (5.23)$$

Onde G_{conc}^{el} é a energia livre do sistema formado pela macromolécula e a solução eletrolítica no qual ela está inserida, e G_{vacuo}^{el} é a energia livre da molécula no vácuo, dada pela soma da energia de interação $u(r_{mn})$ (descrito pela relação 4.1) entre todos os pares de átomos m e n da macromolécula:

$$G_{vacuo}^{el} = \frac{1}{2} \sum_{mn} u(r_{mn}) \quad (5.24)$$

Podemos obter o termo G_{conc}^{el} submetendo o sistema a um processo de carregamento adiabático hipotético, como sugerido por *Onsager* em 1933 [101]. Em 1990, *Sharp* e *Honig* obtiveram a energia livre eletrostática utilizando o princípio variacional [102]. Entretanto, este método sofre de algumas inconsistências, apontadas por *Fogolari* e *Briggs* em 1997 [103]. *Zhou* em 1994 [104] mostrou que a energia livre eletrostática, obtida como sugerida por *Onsager*, além de independente do modo como o sistema é carregado, concorda com os resultados obtidos por *Honig*. Assim, a despeito dessa incerteza, decidimos utilizar o método de carregamento do sistema, uma vez que esse método, juntamente com o proposto por *Honig* e colaboradores, tem sido o mais utilizado. Entretanto, permanece a controvérsia na literatura.

Se durante o processo de carregamento o volume da solução não se alterar, podemos considerar que $G_{conc}^{el} = A_c - A_o$ [105], onde A_c é a energia livre de *Helmholtz* do sistema carregado, e A_o do sistema com todas as partículas possuindo carga nula.

Consideremos a energia livre de *Helmholtz* A dada em função do número de partículas N , do volume V , da temperatura T e das cargas z_j de todas as partículas do sistema. Em N , V e T constantes, a diferencial de A é dada por:



$$dA = \sum_j \left(\frac{\partial A}{\partial z_j} \right)_{N,V,T,z_l \neq j} dz_j \quad (5.25)$$

Pode-se mostrar [23] que esta grandeza está relacionada ao potencial eletrostático médio agindo em uma determinada partícula j , localizada na posição $\vec{r}_j$ ⁷:

$$\langle \varphi(\vec{r}_j) \rangle = \frac{\int \dots \int \varphi(\vec{r}_j) \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1, \dots, d\vec{r}_N}{\int \dots \int \exp \left[-\frac{U(\vec{r}_1, \dots, \vec{r}_N)}{kT} \right] d\vec{r}_1, \dots, d\vec{r}_N} = \left(\frac{\partial A}{\partial z_j} \right)_{N,V,T,z_l \neq j} \quad (5.26)$$

Assim, substituindo a relação 5.26 na relação 5.25, obtemos:

$$dA = \sum_j \langle \varphi(\vec{r}_j) \rangle dz_j \quad (5.27)$$

Introduzindo um parâmetro de carregamento λ , de forma que $dz_j = z_j d\lambda$ ($\lambda = 0$ representa o sistema totalmente descarregado, enquanto $\lambda = 1$ representa o estado final carregado), e reescrevendo a equação 5.27 em função de uma distribuição contínua de cargas $\rho(\vec{r})$, obtemos a integral de carregamento do sistema [104]:

$$A_c - A_o = \sum_j z_j \int_0^1 \langle \varphi(\vec{r}_j, \lambda) \rangle d\lambda = \int_V \int_0^1 \langle \varphi(\vec{r}, \lambda) \rangle \rho(\vec{r}) d\vec{r} d\lambda \quad (5.28)$$

A equação de PB para um determinado estado λ de carregamento do sistema, é dada por [104]:

$$\nabla^2 \varphi(\vec{r}, \lambda) = -\frac{\lambda}{\epsilon_0 \epsilon_s} \rho^f(\vec{r}) - \frac{1}{\epsilon_0 \epsilon_s} \sum_j \rho_j(bulk) \exp \left[-\frac{\lambda z_j e \varphi(\vec{r}, \lambda)}{kT} \right] \quad (5.29)$$

Derivando em relação a λ e rearranjando, obtemos:

⁷ Isso pode ser mostrado apenas quando da utilização da equação de PB, uma vez que a aproximação de campo-médio desconsidera o tamanho das partículas, permitindo que a interação entre quaisquer duas partículas j e l seja dada apenas por $u(r_{jl}) = \frac{z_j z_l e^2}{4\pi \epsilon_0 \epsilon_s r_{jl}}$.

$$\rho^f(\vec{r}) = -\epsilon_0 \epsilon_s \nabla^2 \left[\frac{\partial \varphi(\vec{r}, \lambda)}{\partial \lambda} \right] - \sum_j \rho_j(bulk) \frac{\partial}{\partial \lambda} \left\{ \exp \left[-\frac{\lambda z_j e \varphi(\vec{r}, \lambda)}{kT} \right] \right\} \quad (5.30)$$

Da mesma forma, a relação para a densidade de carga $\rho_j(\vec{r})$ dos íons móveis da espécie j (relação 5.11), pode ser reescrita:

$$\rho_j(\vec{r}) = \rho_j(bulk) \exp \left[-\frac{\lambda z_j e \varphi(\vec{r}, \lambda)}{kT} \right] \quad (5.31)$$

Sendo assim, dado $\rho(\vec{r}) = \rho^f(\vec{r}) + \sum_j \rho_j(\vec{r})$, podemos substituir as expressões 5.30 e 5.31 na expressão 5.28, obtendo G_{conc}^{el} :

$$G_{conc}^{el} = G^{mol} + G^{sal} + G^{mov} + G^{solv} \quad (5.32)$$

Onde G^{mol} e G^{sal} são as componentes referente a energia potencial eletrostática da macromolécula e dos íons móveis da solução, respectivamente:

$$G^{mol} = \frac{1}{2} \int_V \rho^m(\vec{r}) \varphi(\vec{r}) d\vec{r} \quad (5.33)$$

$$G^{sal} = \frac{1}{2} \int_V \sum_j \rho_j(\vec{r}) \varphi(\vec{r}) d\vec{r} \quad (5.34)$$

Durante o carregamento do sistema há uma redistribuição dos íons da solução, de forma que os contraíons se concentrarão próximos à molécula, enquanto os coíons tenderão a se afastar. A contribuição entrópica devido a este rearranjo dos íons móveis é representada por G^{mov} :

$$G^{mov} = kT \int_V \sum_j C_j(\vec{r}) \ln \left[\frac{C_j(\vec{r})}{C_s} \right] d\vec{r} \quad (5.35)$$

O último termo da relação 5.32 (G^{solv}) está relacionado à reorganização das moléculas do solvente, devido ao “efeito osmótico” gerado pelo gradiente de concentração salina através

do volume da solução:

$$G^{solv} = kT \int_V \sum_j [C_s - C_j(\vec{r})] d\vec{r} \quad (5.36)$$

No caso da versão linear da equação de PB, o acréscimo de contraíons é sempre balanceado pelo decréscimo de coíons, fazendo com que a concentração iônica total seja constante por todo o volume da solução, resultado em $G^{solv} = 0$. Além disso, existe uma compensação entre efeitos entálpicos e entrópicos dos íons livres, de forma que a energia livre para um sistema descrito pela versão linear da equação de PB se resume a:

$$G_{conc}^{el} = G^{mol} \quad (5.37)$$

5.3.3 O Modelo de Tanford-Kirkwood

O modelo de *Tanford-Kirkwood* [86, 55, 56, 51] foi desenvolvido no nível de *McMillan-Mayer* [18], onde a proteína é modelada como uma cavidade de baixa constante dielétrica inserida em uma solução eletrolítica, como mostrado na figura 5.3. A região 1, uma cavidade esférica de raio b e constante dielétrica ϵ_i , representa a nossa proteína. A região 2 representa a área inacessível ao eletrólito ($\epsilon_s = 78,7$), e a região 3 é a nossa solução, tratada de acordo com a teoria de *Debye-Hückel* (DH).

Tratando esse problema do ponto de vista puramente eletrostático, consideramos a origem do nosso sistema de referência no centro das três esferas que representam as regiões 1, 2 e 3. O potencial φ_1 na região 1 será descrito, exceto na posição exata das cargas, pela solução da equação de *Laplace* [78]. Portanto, considerando que esse potencial deve ser finito na origem, temos:

$$\varphi_1(r < b, \theta, \varphi) = \sum_{j=1}^M \frac{Z_j}{\epsilon_i |\vec{r} - \vec{r}_j|} + \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} B_{nm} r^n P_n^m(\cos \theta) e^{im\varphi} \quad (5.38)$$

Onde M é o número de cargas Z_j da proteína posicionadas na posição $\vec{r}_j$, e r , θ e φ são as

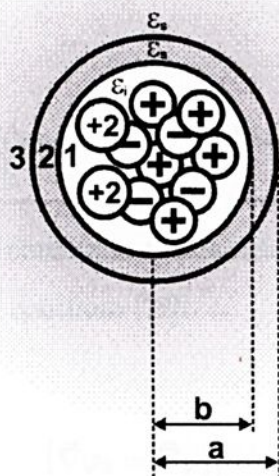


Figura 5.3: Diagrama esquemático para o modelo de *Tanford-Kirkwood*. Modelo baseado na descrição dada na figura 4.2b

coordenadas usuais do sistema de coordenadas esféricas. O primeiro termo de φ_1 refere-se à auto energia da configuração das cargas da proteína, e o segundo termo à interação dessas cargas com a solução eletrolítica e as cargas de polarização geradas na interface entre as regiões 2 e 3.

O potencial da região 2 também deve satisfazer a equação de *Laplace*:

$$\varphi_2(b < r < a, \theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \left\{ \frac{C_{nm}}{r^{n+1}} + G_{nm} r^n \right\} P_n^m(\cos \theta) e^{im\varphi} \quad (5.39)$$

Já para a região 3, o potencial deve satisfazer a equação linear de PB [86] ($\nabla^2 \varphi_3 - \kappa^2 \varphi_3 = 0$).

Portanto, considerando que esse potencial deve ser nulo no infinito, temos:

$$\varphi_3(r > a, \theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \frac{A_{nm}}{r^{n+1}} e^{-\kappa r} K_n(\kappa r) P_n^m(\cos \theta) e^{im\varphi} \quad (5.40)$$

Onde $P_n^m(\cos \theta)$ são os polinômios associados de *Legendre* [78, 106] e $K_n(x)$ é dado por:

$$K_n(x) = \sum_{s=0}^n \frac{2^s n! (2n-s)!}{s! (2n)! (n-s)!} x^n \quad (5.41)$$

Na interface entre as três regiões nossa solução deve satisfazer a continuidade do potencial [78]:

$$[\varphi_3 = \varphi_2]_{r=a} \quad [\varphi_2 = \varphi_1]_{r=b} \quad (5.42)$$

Além disso, entre as regiões 2 e 3, como não temos descontinuidade da constante dielétrica, o campo elétrico também deve ser contínuo [78]:

$$[\vec{\nabla}\varphi_3 = \vec{\nabla}\varphi_2]_{r=a} \quad (5.43)$$

Já na interface entre as regiões 1 e 2, devido a descontinuidade da constante dielétrica, devemos ter a continuidade do vetor deslocamento elétrico e das componentes tangencial e azimutal do campo elétrico [78]:

$$\left[\frac{1}{r} \frac{\partial \varphi_2}{\partial \theta} = \frac{1}{r} \frac{\partial \varphi_1}{\partial \theta} \right]_{r=b} \quad \left[\frac{1}{r \sin \theta} \frac{\partial \varphi_2}{\partial \varphi} = \frac{1}{r \sin \theta} \frac{\partial \varphi_1}{\partial \varphi} \right]_{r=b} \quad \left[\epsilon_s \frac{\partial \varphi_2}{\partial r} = \epsilon_i \frac{\partial \varphi_1}{\partial r} \right]_{r=b} \quad (5.44)$$

Para calcularmos a energia livre descrita na seção acima (G^{mol}), precisamos do potencial na posição de cada carga, portanto basta calcular a constante B_{nl} através das condições de contorno acima, obtendo [55]:

$$G = \frac{e^2}{8\pi\epsilon_0} \sum_{j=1}^M \sum_{l=1}^M z_j z_l (A_{jl} - B_{jl} - C_{jl})$$

$$A_{jl} = \frac{1}{\epsilon_i |\vec{r}_j - \vec{r}_l|} \quad B_{jl} = \frac{1}{b\epsilon_i} \sum_{n=0}^{\infty} \frac{(\epsilon_s - \epsilon_i)}{(\epsilon_s + n\epsilon_i/(n+1))} \left(\frac{r_j r_l}{b^2} \right)^n P_n(\cos \theta_{jl}) \quad (5.45)$$

$$C_{jl} = \frac{1}{a\epsilon_s} \left[\frac{x}{1+x} + x^2 \sum_{n=1}^{\infty} \frac{\frac{2n+1}{2n-1} \left(\frac{\epsilon_s}{(n+1)\epsilon_s + n\epsilon_i} \right)^2 \left(\frac{r_j r_l}{a^2} \right)^n P_n(\cos \theta_{jl})}{\frac{K_{n+1}(x)}{K_{n-1}(x)} + \frac{n(\epsilon_s - \epsilon_i)}{(n+1)\epsilon_s + n\epsilon_i} \left(\frac{x^2}{4n^2 - 1} \right) \left(\frac{b}{a} \right)^{2n+1}} \right]$$

Onde o termo A_{jl} representa o trabalho de montar a configuração de cargas em um meio infinito de constante dielétrica ϵ_i . O termo B_{jl} é a energia devido a inserção da interface dielétrica entre as regiões 1 e 2, e o termo C_{jl} , a energia de interação com a solução eletrolítica, sendo $x = \kappa a$.

Como toda a carga dos átomos está concentrada em um ponto, os termos A_{jl} para $j = l$ (auto energia de cada fonte de carga) divergem. Entretanto, esses termos se cancelam no cálculo de ΔG^{el} (equação 3.13), podendo ser desconsiderados do cálculo. O mesmo acontece para B_{jl} quando temos cargas na interface dielétrica. Isso também não será um problema, pois utilizamos dielétrico uniforme ($\epsilon_i = \epsilon_s$) em nossos cálculos.

Substituindo a proteína por apenas um íon de carga central Z_j e raio b , podemos calcular, através da expressão acima, a energia livre de solvatação desse íon. Portanto, podemos interpretar essa expressão como uma generalização da energia de *Born* [107], onde aqui, consideramos a inclusão de efeitos de concentração salina e distribuição assimétrica de carga.

5.4 O Método Monte Carlo Metrópolis

A equação de PB, seja na sua forma linear ou não linear, nos fornece um tratamento completo, porém aproximado, para a solução de modelos representando soluções polieletrólíticas [108]. Entretanto, as aproximações inerentes deste método, em sua forma tradicional, carecem de eficiência para macroíons altamente carregados [102, 29], bem como na previsão de efeitos de concentração macromolecular finita [53]. Uma forma de resolver estes modelos de maneira “exata” (em relação ao erro estatístico), é gerar um número razoável de configurações de acordo com a probabilidade $P(\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_{N_s})$ dada pela relação 5.2, utilizando-as para o cálculo das médias das propriedades termodinâmicas de interesse. Este método é chamado método *Monte Carlo* (MC) [18, 24].

Cada uma dessas configurações corresponde a um ponto no espaço de fases. Dessa forma, a média de uma determinada propriedade A é dada por [109]:

$$\langle A \rangle = \frac{1}{M} \sum_{i=1}^M A_i \quad (5.46)$$

Onde M é o número de configurações geradas, e A_i é o valor da propriedade A em uma determinada configuração i .

A geração dessas configurações é realizada através da construção de um processo *Markoviano* [109], cuja probabilidade de transição $w(x \rightarrow y)$ de uma configuração x para uma configuração y , no equilíbrio, é dada pela equação do balanço detalhado [109, 110]:

$$P(x)w(x \rightarrow y) = P(y)w(y \rightarrow x) \quad (5.47)$$

Onde $P(x)$ é a probabilidade do sistema se encontrar em uma dada configuração x . Dessa maneira, devemos obter as probabilidades de transição, de forma que [110]:

$$\frac{w(x \rightarrow y)}{w(y \rightarrow x)} = \exp \left[-\frac{E(y) - E(x)}{kT} \right] \quad (5.48)$$

Onde $E(x)$ e $E(y)$ são as energias totais de uma determinada configuração x e y , respectivamente.

Segundo a relação 5.48, as probabilidades de transição não podem ser obtidas univocamente [110], nos possibilitando escolhas que sejam mais razoáveis do ponto de vista computacional. Uma das soluções disponíveis na literatura é a proposta por Metropolis [18, 24, 109, 110]:

$$w(x \rightarrow y) = \begin{cases} \exp \left[-\frac{E(y) - E(x)}{kT} \right] & , \quad E(y) - E(x) > 0 \\ 1 & , \quad E(y) - E(x) \leq 0 \end{cases} \quad (5.49)$$

Computacionalmente, a geração deste processo *Markoviano* deve seguir o algoritmo abaixo [109]:

1. Escolher arbitrariamente uma configuração qualquer x , e calcular sua energia $E(x)$.
2. A partir da atual configuração x , escolher uma nova configuração y , calculando sua energia $E(y)$.
3. Calcular $\Delta E = E(y) - E(x)$.
 - (a) Se $\Delta E \leq 0$ aceitar a configuração y .
 - (b) Senão, calcular $w(x \rightarrow y) = \exp \left[-\frac{\Delta E}{kT} \right]$ e gerar um número aleatório ξ igualmente distribuído entre 0 e 1.
 - Se $w(x \rightarrow y) \leq \xi$, aceitar a configuração y .
 - Senão, manter a configuração x .
4. Voltar ao passo 2.

Depois de um número suficientemente grande de passos, a relação 5.47 será satisfeita, possibilitando a utilização das configurações seguintes para o cálculos das propriedades médias de interesse, como por exemplo, o potencial químico [109]. Neste caso, muitos passos devem ser gerados até uma convergência razoável ser obtida.

5.5 Potenciais Químicos

Um modo bastante elegante de se obter potenciais químicos através do método MC ou DM, é a utilização do **método de inserção de Widom** [111]. Basicamente, a idéia deste método reside em realizar a inserção de uma partícula (representada pelo índice será $N + 1$) em uma determinada posição, calculando a interação desta partícula “virtual” com todas as outras N partículas, obtendo o potencial químico em um sistema com densidade ρ , através da relação:

$$\mu = \mu^o + kT \ln \rho - kT \ln \left\langle \exp \left[-\frac{\Delta U(\vec{r}_1, \dots, \vec{r}_{N+1})}{kT} \right] \right\rangle_N \quad (5.50)$$

Onde os dois primeiros termos estão relacionados com a componente ideal do potencial químico. O terceiro termo é o potencial químico de excesso μ^{ex} , sendo ΔU a energia potencial



de interação entre a partícula $N + 1$ e todas as outras N partículas do sistema, e a média é realizada sobre as configurações apenas das N partículas móveis.

Para sistemas iônicos com número finito de partículas (como no *modelo da célula*), este método se torna inconsistente devido a perda da condição de eletroneutralidade [112]. Uma forma de corrigir este problema é calcular o μ^{ex} através de um processo de carregamento hipotético, introduzindo um parâmetro de escala de tal forma que o sistema se mantenha neutro em todo esse processo [112].

Podemos particionar ΔU em três contribuições:

$$\Delta U = \Delta U^{ext} + \Delta U^{el} + \Delta U^{er} \quad (5.51)$$

Onde ΔU^{ext} é a contribuição devido ao potencial produzido pela macromolécula sobre a partícula $N+1$. Os outros dois termos são as contribuições referente a interação eletrostáticas e de esfera-rígida com todas as outras N partículas móveis, respectivamente.

Substituindo na relação 5.50, obtemos:

$$\mu^{ex} = -kT \ln \left\{ \exp \left[-\frac{\Delta U^{ext}}{kT} \right] \left\langle \exp \left[-\frac{\Delta U^{el} + \Delta U^{er}}{kT} \right] \right\rangle_N \right\} \quad (5.52)$$

Introduzindo um parâmetro de carregamento λ , de forma que $dZ_{N+1} = Z_{N+1} d\lambda$, podemos obter o potencial químico de excesso através de um processo de carregamento hipotético, resultando em:

$$\mu^{ex} = \frac{\Delta U^{ext}}{kT} - kT \ln \left\langle \exp \left[-\frac{\Delta U^{er}}{kT} \right] \right\rangle_N + kT \int_0^1 \frac{\left\langle \frac{\Delta U^{el}}{kT} \exp \left[-\frac{\Delta U^{er} + \lambda \Delta U^{el}}{kT} \right] \right\rangle_N}{\left\langle \exp \left[-\frac{\Delta U^{er} + \lambda \Delta U^{el}}{kT} \right] \right\rangle_N} d\lambda \quad (5.53)$$

Podemos introduzir um parâmetro de escala ξ , de forma a manter a eletroneutralidade do sistema durante todo o processo de carregamento. A carga excessiva λZ_{N+1} deve ser distribuída para todos os íons móveis, de forma que o termo ΔU^{el} deve ser substituído por

$\xi \Delta U^{el}$ na relação 5.53, onde $\xi = 1 - \frac{\lambda Z_{N+1}}{Z_i N_i}$ [112] para a interação da partícula $N + 1$ com cada uma das partículas da espécie i .

Computacionalmente, escolhemos uma determinada posição dentro da célula no qual a partícula será inserida. Esta posição poderá ser a posição dos sítios de ligação, quando do cálculo do potencial químico de excesso do ligante no sítio, ou uma posição escolhida aleatoriamente dentro da célula, para o cálculo do potencial químico de excesso do ligante livre em solução. Daí, calcula-se a interação desta partícula com a macromolécula (ΔU^{ext}), e a interação eletrostática (ΔU^{el}) e de *esfera-rígida* (ΔU^{er}) com as partículas móveis, obtendo o potencial químico a partir da expressão 5.53, sendo a média dada sobre as várias configurações geradas.

5.3 Cálculo da Interação da Partícula

Para o cálculo da interação da partícula com a macromolécula, é necessário conhecer a posição da partícula e a posição dos sítios de ligação. A posição da partícula é dada pela coordenada (x, y, z) e a posição dos sítios de ligação é dada pela coordenada (x_i, y_i, z_i) . A interação da partícula com a macromolécula é dada pela expressão 5.53, onde ΔU^{ext} é o potencial químico de excesso do ligante no sítio, ΔU^{el} é o potencial eletrostático e ΔU^{er} é o potencial de esfera-rígida.

No cálculo da interação da partícula com a macromolécula, é necessário conhecer a posição da partícula e a posição dos sítios de ligação. A posição da partícula é dada pela coordenada (x, y, z) e a posição dos sítios de ligação é dada pela coordenada (x_i, y_i, z_i) .

Podemos considerar que para o cálculo da interação da partícula com a macromolécula, é necessário conhecer a posição da partícula e a posição dos sítios de ligação. A posição da partícula é dada pela coordenada (x, y, z) e a posição dos sítios de ligação é dada pela coordenada (x_i, y_i, z_i) . A interação da partícula com a macromolécula é dada pela expressão 5.53, onde ΔU^{ext} é o potencial químico de excesso do ligante no sítio, ΔU^{el} é o potencial eletrostático e ΔU^{er} é o potencial de esfera-rígida.



Capítulo 6

Resultados

6.1 Efeito da Concentração da Proteína

Começaremos nossas análises verificando o efeito da concentração macromolecular, não considerado na formulação original de *Tanford-Kirkwood* (TK) e nos principais pacotes computacionais que resolvem a equação de *Poisson-Boltzmann* (PB). Verificaremos esse efeito utilizando o método *Monte Carlo* (MC), unicamente através da variação do raio da célula. Todos os cálculos foram realizados em temperatura de 298 K, considerando demais detalhes de acordo com o capítulo 4.

Na figura 6.1, temos resultados da solução do Modelo Esférico (ME), descrito na seção 4.1, através do método MC, utilizando três regimes de concentrações macromoleculares diferentes (20 μM , 200 μM e 2 mM), e através da formulação do modelo de TK.

Podemos verificar que para baixas concentração e carga macromolecular, os resultados obtidos pelo método MC e pela formulação de TK são bastante concordantes. Entretanto, o aumento da concentração proteica desloca as curvas para menores valores de ΔpK , mostrando que energia livre de ligação alcança uma saturação mais rapidamente, indicando o aumento da blindagem eletrostática. Como o número de contraíons adicionados ao sistema para manter a eletroneutralidade de célula (segundo termo da relação 4.3) está relacionado apenas a carga macromolecular, o aumento da concentração proteica leva a um aumento da concentração desses contraíons, responsável por essa blindagem mais efetiva. Em um regime



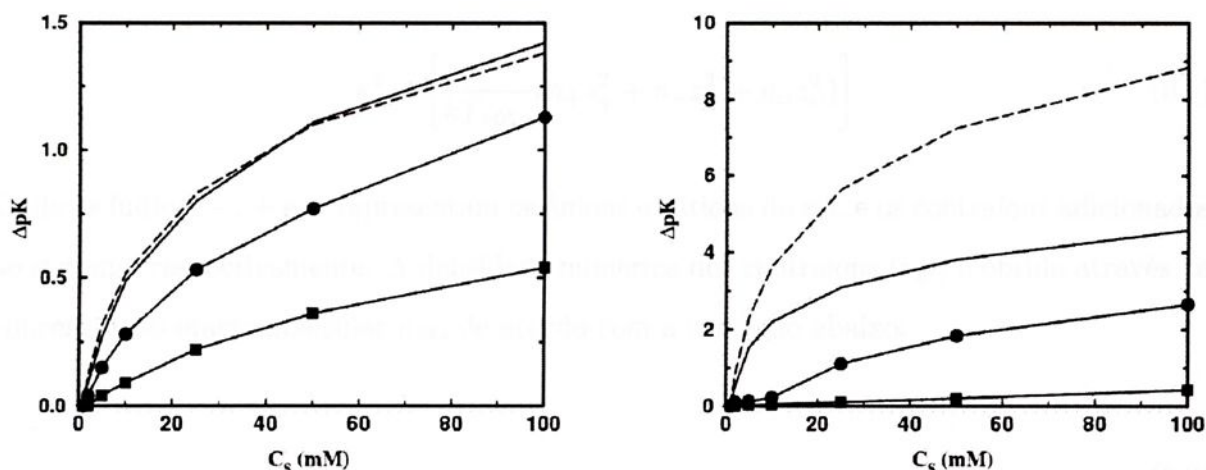


Figura 6.1: Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons bivalentes positivos. A proteína foi representada segundo o ME (veja seção 4.1) com raio de 14 Å e carga $Z_p = -4$ (gráfico da esquerda) e $Z_p = -24$ (gráfico da direita). A posição dos sítios de ligação foi obtida do arquivo PDB-ID 3icb (Calbindina D_{9k}). A linha tracejada demonstra os resultados obtidos com o modelo de TK. As linhas sólidas sem símbolos, com círculos e com quadrados representam cálculos através do método MC nas concentrações macromoleculares de 20 μM , 200 μM e 2 mM, respectivamente. A força iônica de 1 mM foi tomada como referência. Detalhes dos cálculos realizados podem ser encontrados nas referências [53, 51].

de maior diluição macromolecular, esses contraíons estão mais dispersos, blindando menos efetivamente a interação ligante-proteína.

Todos esses efeitos são aumentados quando verificamos o comportamento de ΔpK para sistemas altamente carregados ($Z_p = -24e$). Nesse regime, mesmo em baixa concentração macromolecular, o modelo de TK não consegue reproduzir o efeito da grande blindagem eletrostática observada nos resultados obtidos pelo método MC. Devido a alta carga da proteína, a aproximação de *Debye-Hückel* (DH), utilizada no modelo de TK, se torna precária, subestimando a blindagem eletrostática produzida pelos íons móveis da solução, resultando na superestimação do ΔpK .

Baseados na literatura da área de física e química de sistemas colóidais [97, 113, 114], propomos uma modificação no modelo de TK de forma a recuperar os efeitos de alta blindagem eletrostática causada pelo aumento da concentração e carga macromolecular. Primeiramente, a contribuição dos contraíons adicionados ao sistema se torna mais importantes a medida que se aumenta a concentração macromolecular, ou seja, a força iônica se torna cada vez mais dependente desses contraíons, e por isso eles devem ser incluídos no comprimento de *Debye*:

$$\kappa^2 = \left[\frac{e^2}{kT\epsilon_0\epsilon_s} (n_+ z_+^2 + n_- z_-^2 + n_{ci} z_{ci}^2) \right] \quad (6.1)$$

Onde os índices $-$, $+$ e ci representam os ânions e cátions do sal, e os contraíons adicionados ao sistema, respectivamente. A densidade numérica dos contraíons (n_{ci}) é obtida através da concentração macromolecular n_M , de acordo com a condição abaixo:

$$n_{ci} z_{ci} = n_M z_M \quad (6.2)$$

Essa inclusão foi proposta primeiramente por *Beresford-Smith* (BS) e colaboradores [113], onde obtiveram um κ dependente da concentração macromolecular, diferente da tradicional teoria de DH.

Para cálculos de $\Delta\Delta G_B$ (definido na relação 3.10), em baixa concentração macromolecular, a aproximação de BS reproduz praticamente o mesmo resultado de TK tradicional, como podemos verificar na figura 6.2. Já para altas concentrações, essa aproximação, além de reproduzir o comportamento qualitativo não linear, consegue uma excelente concordância quantitativa com resultados obtidos através do método MC, inclusive para proteínas altamente carregadas.

Verificando os resultados de ΔpK apresentados na figura 6.3, onde é incluído os efeitos dos ligantes livres na solução, observamos um desvio considerável entre os resultados obtidos através do método MC e a aproximação de BS. Isso pode ser explicado, considerando que o potencial químico do ligante livre em solução (μ_L^o) obtido através do método MC, contém a contribuição do potencial eletrostático da macromolécula, uma vez que, utilizando o método de *Widom*, realizamos inserções em posições aleatórias dentro da célula contendo a macromolécula. Quantitativamente, temos obtido uma diferença de aproximadamente 4 unidades de kT entre o potencial químico do ligante livre para carga macromolecular de $Z_p = -8e$ e $Z_p = -24e$, em baixa concentração salina e proteica. Entretanto, quando utilizamos a aproximação de BS essa diferença é de apenas 0,02 unidades de kT . Essa discrepância leva ao desvio observado.



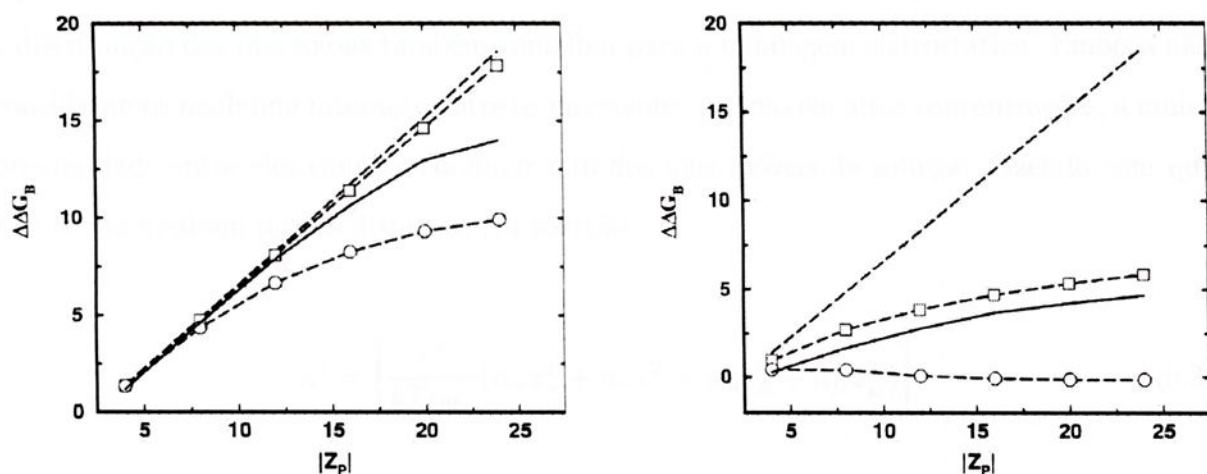


Figura 6.2: Resultados de $\Delta\Delta G_B$ devido a variação da concentração salina de 1 mM para 100 mM. A proteína possui carga $-|Z_p|$. No gráfico da **esquerda** foi considerada uma concentração macromolecular de 20 μM , enquanto para o gráfico da **direita** essa concentração foi de 2 mM. A linha sólida representa resultados obtidos através do método MC. As linhas tracejadas sem símbolo, com quadrados e com círculos, demonstram os valores obtidos com o modelo de TK na aproximação tradicional, de BS e de Lin, respectivamente. Outros detalhes de acordo com a figura 6.1.

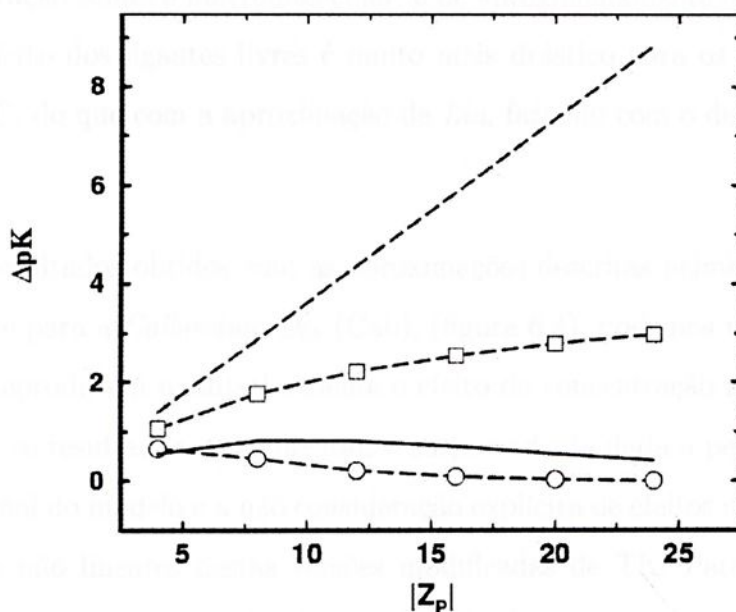


Figura 6.3: Resultados de ΔpK para a ligação de dois íons bivalentes positivos devido a variação da concentração salina de 1 mM para 100 mM em concentração macromolecular de 2 mM. Outros detalhes de acordo com a figura 6.2.

Outra modificação, proposta por *Lin* e colaboradores [114], considera a inclusão da própria espécie macromolecular no κ , considerando que para altas concentrações macromoleculares a distribuição dos macroíons também contribui para a blindagem eletrostática. Embora não consideramos nenhuma interação entre os macroíons, mesmo em altas concentrações, a maior proximidade entre eles causa o confinamento dos íons móveis da solução, fazendo com que eles se mantenham menos dispersos em solução.

$$\kappa^2 = \left[\frac{e^2}{kT\epsilon_0\epsilon_s} (n_+ z_+^2 + n_- z_-^2 + n_{ci} z_{ci}^2 + n_M z_M^2) \right] \quad (6.3)$$

Podemos observar o sucesso dessa aproximação analisando os resultados da figura 6.3 e comparando-os com resultados obtidos por MC. Entretanto, a figura 6.2 nos mostra que a aproximação de *Lin* subestima os resultados de $\Delta\Delta G_B$, devido ao termo z_M^2 na relação 6.3 que acarreta um grande decréscimo do comprimento de *Debye*, principalmente para alta carga macromolecular. Isso indica que há uma cancelamento de erros quando da inclusão da contribuição dos ligantes livres em solução. Para a aproximação de *Lin*, a diferença entre o potencial químico do ligante livre para carga macromolecular de $Z_p = -8e$ e $Z_p = -24e$, em baixa concentração salina e macromolecular, é de aproximadamente 0,4 unidades de kT . Dessa forma, o efeito dos ligantes livres é muito mais drástico para os resultados obtidos com o método MC, do que com a aproximação de *Lin*, fazendo com o desvio em $\Delta\Delta G$ seja compassado.

Comparando os resultados obtidos com as aproximações descritas acima e com os obtidos experimentalmente para a *Calbindina D9k* (Cab), (figura 6.4), podemos verificar que ambas as aproximações reproduzem qualitativamente o efeito da concentração macromolecular. A concordância com os resultados experimentais é surpreendente dada a pequena modificação na versão tradicional do modelo e a não consideração explícita de efeitos de correlação iônica e efeitos térmicos não lineares nestas versões modificadas de TK. Para os resultados de variação de concentração macromolecular (gráfico da direita), a aproximação de BS tem uma boa concordância com os resultados experimentais, sendo o maior desvio de 0,5 pK. A aproximação de *Lin* para esse caso superestima o efeito de blindagem, dada a ausência completa de sal no sistema. Já para os resultados de variação da força iônica, a aproximação



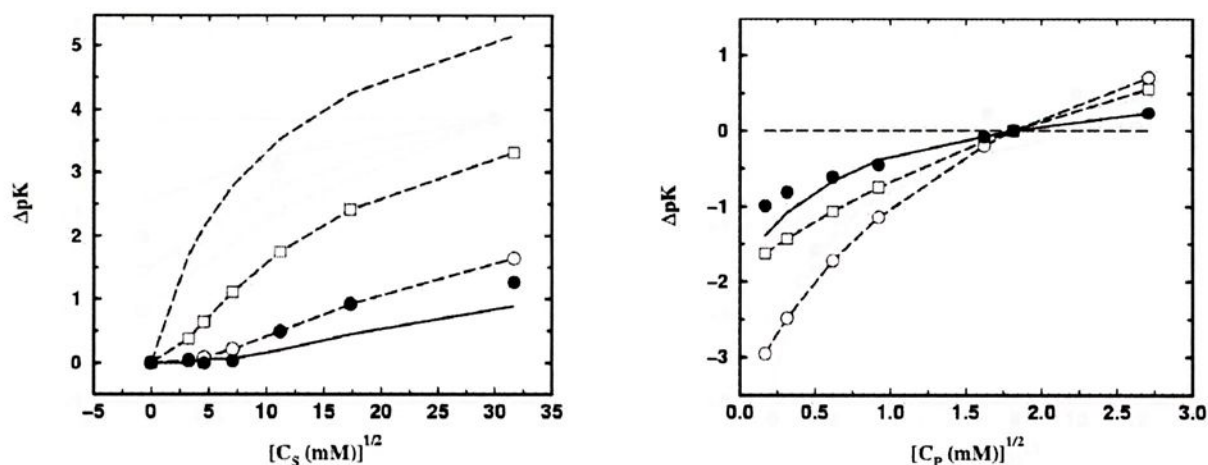


Figura 6.4: Resultados de ΔpK para a ligação de dois íons bivalentes positivos à *Calbindina D_{9k}* utilizando o campo de força Reduzido. O gráfico da esquerda se refere a concentração macromolecular de 3,3 mM e concentração salina de referência de 0 mM. O gráfico da direita se refere a concentração salina nula e concentração macromolecular de referência de 3,3 mM. Os círculos pretos representam os resultados experimentais [54]. Outros detalhes de acordo com a figura 6.2.

de *Lin* consegue reproduzir muito bem os efeitos de blindagem eletrostática, tendo em vista as condições críticas do sistema (concentração macromolecular de 3,3 mM).

Para proteínas altamente carregadas, como a *Calmodulina* (CaM), verificamos através da figura 6.5 que ambas as aproximações propostas (BS e *Lin*), conseguem reproduzir de maneira bastante razoável os resultados obtidos experimentalmente. Para baixas concentrações macromoleculares (gráfico da direita), a aproximação de *Lin* superestima a blindagem, sendo o maior desvio em relação aos resultados experimentais, aproximadamente de 1,6 pK. Ainda nesse caso, a aproximação de BS concorda muito bem com a tradicional teoria de DH, e a concordância com os resultados experimentais é surpreendente. Já para sistemas menos diluídos, ao contrário do que acontece para Cab (figura 6.4), a aproximação de *Lin* consegue ser mais eficiente quando comparada a BS, mesmo com concentração salina nula, conseguindo recuperar o efeito dos contraíons que se aglomeram ao redor da proteína devido a sua alta carga.

Essas conclusões são confirmadas nos resultados apresentados na figura 6.6, onde, para os dois fragmentos da CaM, a aproximação de *Lin* superestima a blindagem eletrostática e a aproximação de BS concorda muito bem com a tradicional teoria de DH e com os resultados experimentais. Observamos um desvio entre os resultados experimentais e a aproximação de BS, um pouco maior para o segundo fragmento, sendo porém, menor que o desvio padrão

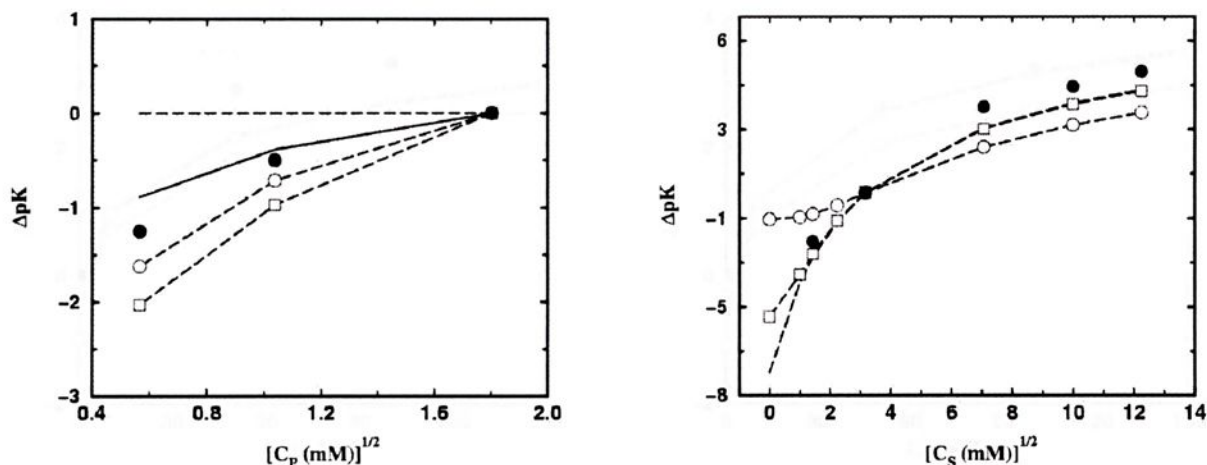


Figura 6.5: Resultados de ΔpK para a *Calmodulina* utilizando o campo de força Reduzido. O gráfico da esquerda se refere a ligação de dois íons bivalentes (sítios I e II) mantendo os dois outros ligados aos sítios III e IV, em concentração salina nula e concentração macromolecular de referência de 3,25 mM. O gráfico da direita se refere a ligação de quatro íons bivalentes, em concentração macromolecular de 20 μM e concentração salina de referência de 10 mM. Os resultados experimentiais foram obtidos, para os gráficos da esquerda e da direita, através das referências [54] e [71], respectivamente.

experimental ($\pm 0,6$ unidades de pK).

Baseados em todos esses resultados, percebemos que a inclusão de termos adicionais no κ , responsável pela blindagem decorrente do aumento da concentração macromolecular, contém toda a informação física necessária para recuperar as previsões de MC. Porém, ainda não conseguimos determinar o limite de validade para a utilização de uma ou outra aproximação. Podemos especular que para regimes de alta concentração, tanto salina quanto macromolecular, a aproximação de *Lin* é mais eficiente. Entretanto, para situações mais críticas de baixa concentração salina, a aproximação de BS se torna mais recomendada em baixa carga macromolecular. A compreensão mais precisa desses limites será alcançada em uma etapa posterior de trabalho, onde cálculos com o método MC serão utilizados para uma análise mais detalhada das componentes da energia livre que influenciam a interação da proteína com o ligante.

6.2 Aproximação de Campo Médio

Analisaremos agora a influência da aproximação de campo médio, intrínsecas das abordagens por PB e DH, no cálculo do ΔpK . Neste intuito, utilizaremos o ME, resolvendo-o utilizando três métodos: (a) aproximação de DH, (b) solução da equação não linear de PB, e (c) o

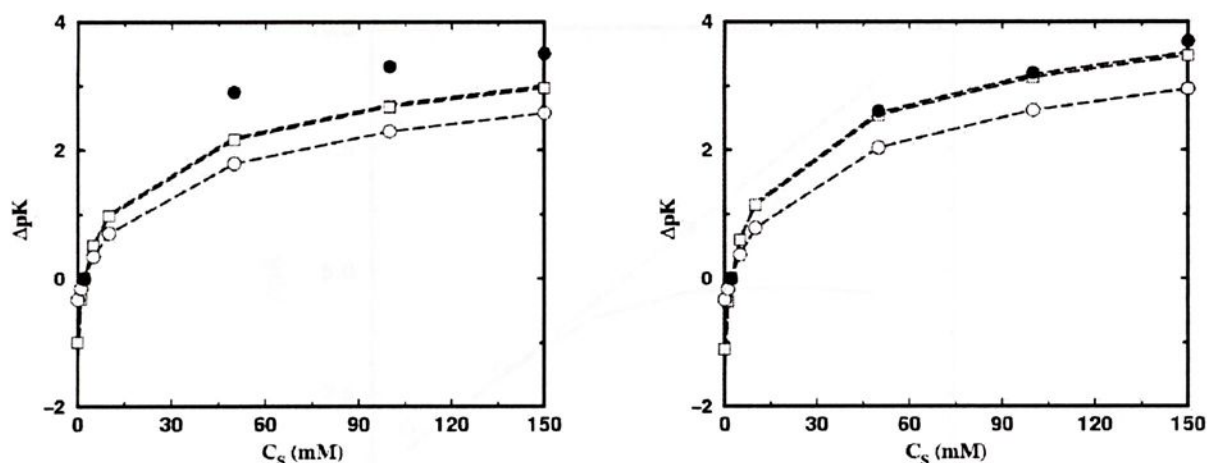


Figura 6.6: Resultados de ΔpK para a ligação de dois íons bivalentes positivos aos fragmentos da *Calmodulina* em 20 μM , utilizando o campo de força Reduzido. A concentração salina de referência foi de 2 mM. Outros detalhes de acordo com a figura 6.4. os resultados experimentais foram obtidos através da referência [71].

método MC.

Na figura 6.7, temos resultados para a ligação de dois íons bivalentes positivos a uma proteína esférica de raio 14 Å e carga central $-|Z_p|$, a qual foi variada de -4 a -24e. Os resultados de ΔpK obtidos pelas três metodologias apresentam uma concordância muito boa para proteínas de baixa carga macromolecular, como já era esperado. Isto confirma que para pequenos valores de $|Z_p|$, embora não incluídos explicitamente, os efeitos de correlação iônica entre os íons do sal e a interação entre esses íons e a molécula, são bem descritos pela aproximação de campo médio. Entretanto, o aumento da carga macromolecular causa a “quebra” desta aproximação, principalmente na solução prescrita por DH.

Resultados obtidos através da solução da equação de PB não linear, demonstram uma concordância qualitativa com os resultados gerados por MC. De fato, ambas as aproximações preveem uma saturação no ΔpK para macromoléculas altamente carregadas. Devido a problemas de estabilidade, intrínsecos de equações diferenciais não lineares, os resultados obtidos por nós através da equação não linear de PB podem sofrer de inconsistências quantitativas, inviabilizando uma análise mais profunda desses resultados. Entretanto, observando o comportamento qualitativo da curva obtida por PB (concordante com resultados obtidos por outros modelos semelhantes), e comparando com os resultados obtidos pelo método MC, podemos afirmar que as divergências mostradas no gráfico da direita da figura 6.1 ($Z_p = -24$) entre DH e MC, podem ser apenas atenuadas através da utilização da solução da equação

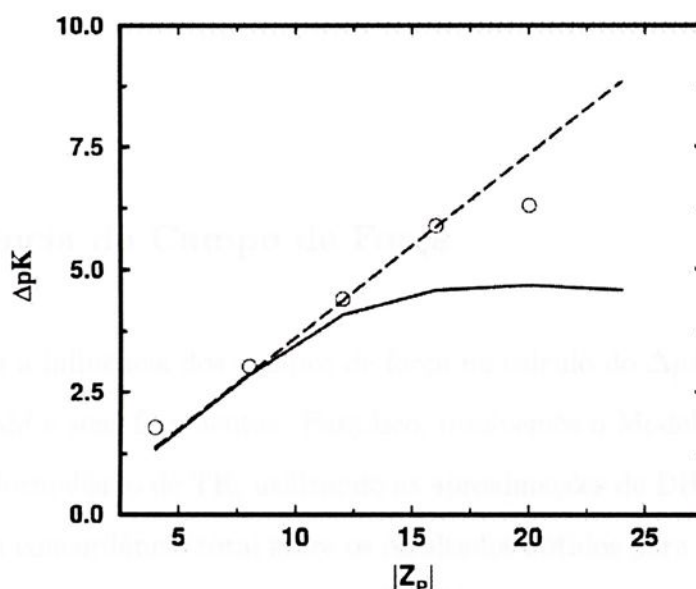


Figura 6.7: Resultados de ΔpK devido a variação da concentração salina de 1 mM para 100 mM em função da carga da proteína. A proteína possui carga $-|Z_p|$ em concentração proteica de 2 μM . A posição dos sítios de ligação foi obtida do arquivo PDB-ID 3icb (Calbindina D_{9k}). A linha sólida representa resultados obtidos através do método MC, enquanto a linha tracejada demonstra os valores obtidos com o modelo de TK. Os círculos representam resultados obtidos através da solução da equação não linear de PB, utilizando o pacote computacional APBS [98], através do método de *elementos finitos multigrid* [99, 100].

não linear, indicando que os fortes efeitos de correlação iônica são fundamentais para a boa descrição da interação entre íons e macromoléculas altamente carregadas.

Os valores obtidos pela aproximação de DH mostram divergências, mesmo qualitativas, com os resultados encontrados pelos cálculos com os outros métodos (PB e MC), para o regime de alta carga macromolecular. Isto acontece porque neste regime os contraíons da solução se acumulam muito próximo ao macroíon, de forma que a relação entre o potencial eletrostático e a densidade iônica deixa de ser linear, ficando em desacordo com a aproximação de DH. Essa aproximação considera um menor acúmulo desses íons em regiões próximas a macromolécula¹⁰, prevendo uma menor blindagem da proteína, em relação ao que deveria ser encontrado no sistema real ou previsto pelo método MC, facilitando a ligação do cálcio. Esta subestimação da blindagem eletrostática é mais drástica para baixa força iônica, fazendo com que a diferença entre as energias livre de ligação para duas concentrações de sal seja superestimada.

6.3 Dependência do Campo de Força

Verificaremos agora a influência dos campos de força no cálculo do ΔpK da ligação de íons Ca^{+2} à Cab, e à CaM e seus fragmentos. Para isso, resolvemos o Modelo Esférico Reduzido (MED) através do formalismo de TK, utilizando as aproximações de DH, BS e *Lin*. A figura 6.8 nos mostra uma concordância total entre os resultados obtidos para a CaM, utilizando a aproximação de *Lin*, com os campos de força GROMOS, PARSE e Reduzido, independente da concentração salina ou macromolecular utilizada.

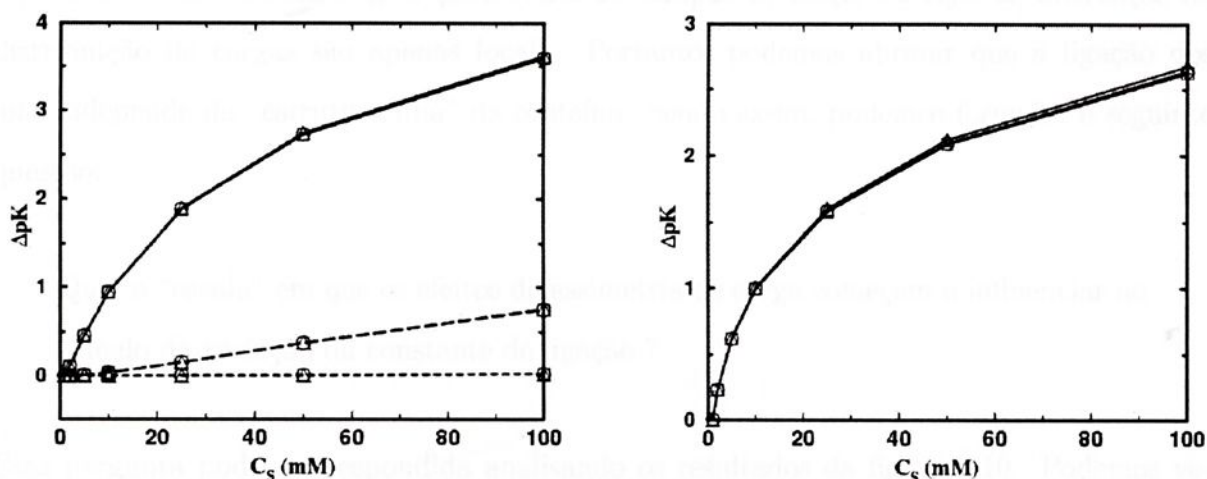


Figura 6.8: No gráfico da **esquerda** demonstramos os resultados de ΔpK devido a variação da concentração salina para a ligação de quatro íons bivalentes positivos à *Calmodulina* utilizando a aproximação de *Lin*. O gráfico da **direita** se refere ao ΔpK em função da concentração salina para a ligação de dois íons bivalentes positivos à *Calbindina D9k* utilizando a aproximação de DH. A linha contínua, tracejada longa e tracejada curta representam concentrações macromoleculares de 20 μM , 200 μM e 2 mM, respectivamente. As linhas com círculos, quadrados e triângulos se referem aos campos de força GROMOS, PARSE e Reduzido, respectivamente. A força iônica de 1 mM foi considerada como referência.

Poderíamos argumentar que, devido a elevada carga existente na CaM, o κ para a aproximação de *Lin* é bastante reduzido, de forma que o potencial obtido próximo a superfície macromolecular seja baixo, podendo atenuar os desvios devido a variação da distribuição de carga. Este efeito é equivalente a considerar uma aproximação mais efetiva dos íons de sal à molécula de forma a cancelar os efeitos da assimetria na distribuição de carga. Entre-

¹⁰ Veja a figura 5.2, onde apresentamos o potencial eletrostático para as aproximações de DH e PB não linear, identificando uma menor blindagem produzida por DH.

tanto, essa concordância é a mesma para proteínas menos carregadas como a Cab, também mostrada na figura 6.8, que mesmo em um regime de baixa concentração macromolecular e utilizando a aproximação tradicional de DH, onde o sal se torna mais disperso em solução, não demonstra nenhuma característica peculiar de um determinado campo de força.

O mesmo comportamento mostrado acima é confirmado quando verificamos os resultados para os dois fragmentos da CaM (figura 6.9). Da mesma forma que para o dímero, esses resultados foram obtidos em um grande intervalo de força iônica e concentração macromolecular, confirmando que a influência quase inexistente dos campos de força, independe das condições em que a ligação ocorre.

A justificativa a esse efeito quase nulo pode ser dada se baseando na idéia de que a carga líquida dos aminoácidos é igual para todos os campos de força, ou seja, as diferenças na distribuição de cargas são apenas locais. Portanto, podemos afirmar que a ligação dos íons independe da “estrutura fina” da proteína. Sendo assim, podemos formular a seguinte questão:

Qual a “escala” em que os efeitos de assimetria de carga começam a influenciar no cálculo da variação da constante de ligação ?

Essa pergunta pode ser respondida analisando os resultados da figura 6.10. Podemos verificar, para todas as aproximações propostas, a concordância entre o campo de força Reduzido, que já foi mostrado ser equivalente à campos de força razoavelmente sofisticados como GROMOS e PARSE, e aproximações bastante grosseiras, como o ME e Aproximação Dipolar (AD). Isso pode ser confirmado através dos resultados da figura 6.11, onde, embora não tenhamos uma concordância tão boa ($\Delta\Delta pK = 0,5$) como para os fragmentos da CaM, o desvio pode ser considerado pequeno, considerando a simplicidade da aproximação.

Sendo assim, podemos considerar o processo de ligação de íons à proteínas um fenômeno global, dependendo mais de características gerais da proteína e do ambiente no qual esse sistema esta imerso (solvente e sal) do que de suas características locais e peculiaridades estruturais. Acreditamos que a característica de **longo alcance da força eletrostática**,



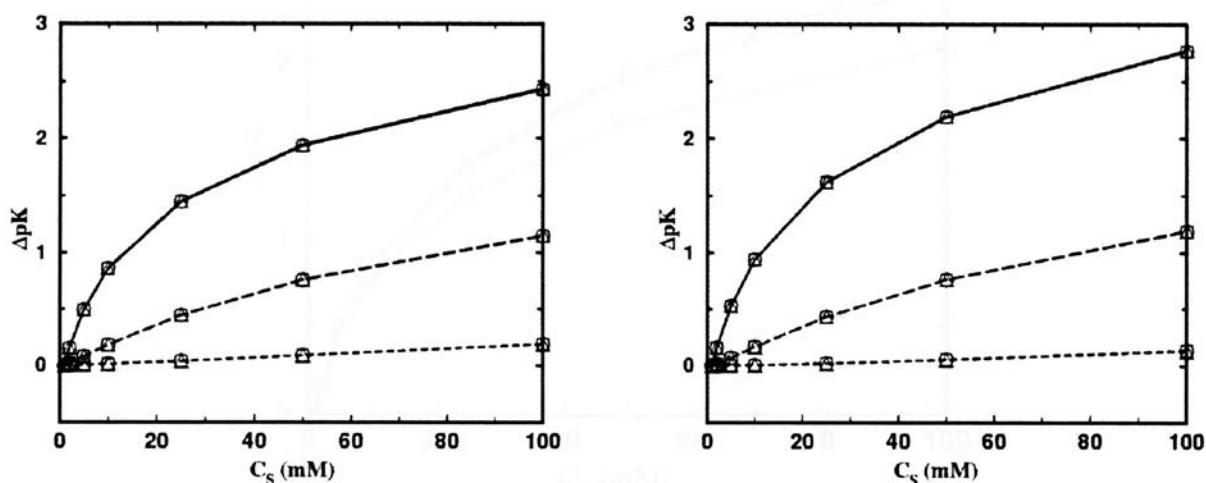


Figura 6.9: Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons bivalentes positivos aos fragmentos da *Calmodulina* utilizando a aproximação de *Lin*. Outros detalhes de acordo com a figura 6.8.

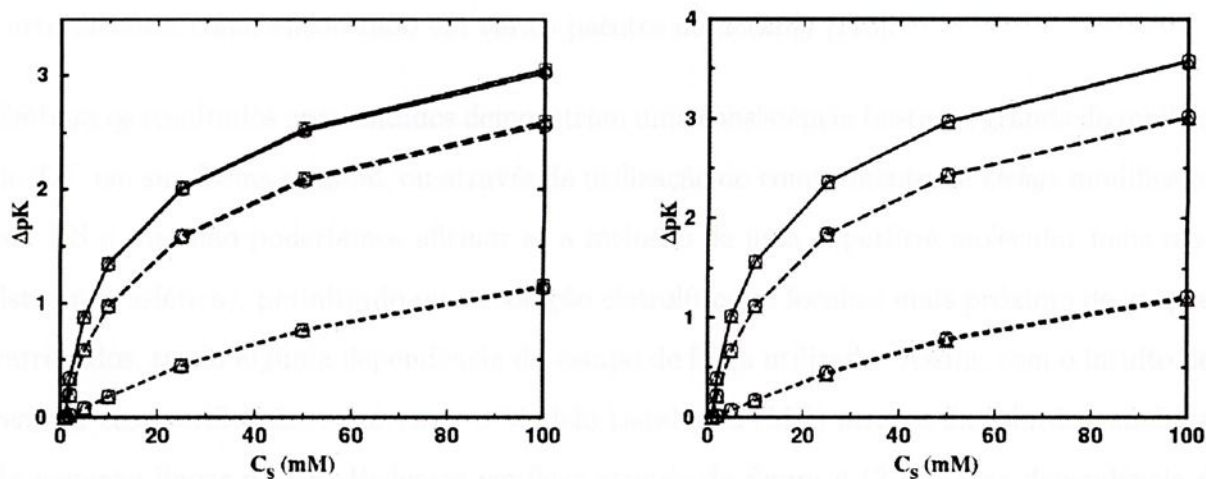


Figura 6.10: Resultados de ΔpK devido à variação da concentração salina para a ligação de dois íons bivalentes positivos aos fragmentos da *Calmodulina* em concentração macromolecular de $200 \mu M$. A linha contínua, tracejada longa e tracejada curta representam as aproximações de DH, BS e *Lin*, respectivamente. As linhas com círculos, quadrados e triângulos se referem ao ME, AD e MED utilizando o campo de força Reduzido, respectivamente. A força iônica de 1 mM foi considerada como referência.

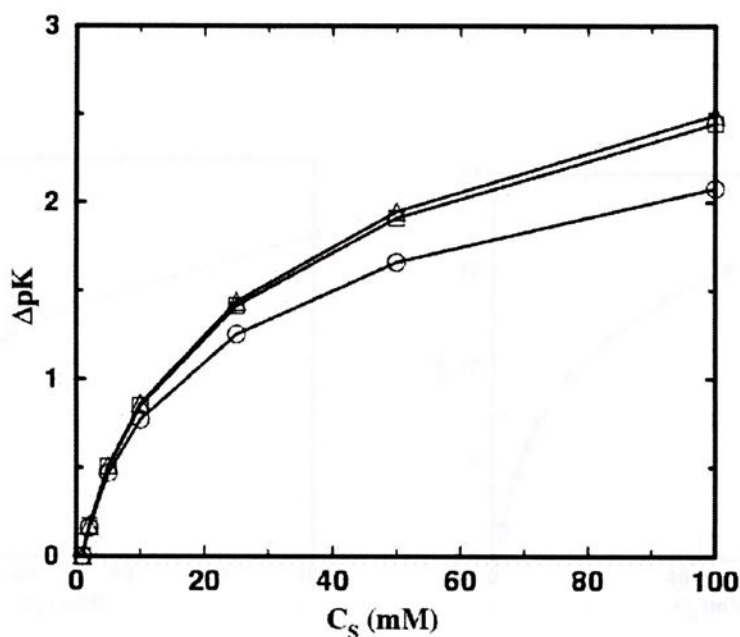


Figura 6.11: Resultados de ΔpK devido à variação da concentração salina para a ligação de dois íons bivalentes positivos à *Calbindina D_{9k}* em concentração macromolecular de 200 μ M. Outros detalhes de acordo com a figura 6.10.

seja a principal causa desse comportamento, fazendo com que o efeito de toda a proteína no sítio de ligação seja mais importante do que efeitos locais, onde a diferença na distribuição de cargas é particular de cada campo de força. Dessa maneira, as conclusões obtidas aqui podem não ser válidas para modelos mais sofisticados que também consideram forças de curto alcance, como encontrado em vários pacotes de *docking* [115].

Embora os resultados apresentados demonstrem uma consistência bastante grande do modelo de TK, em sua forma original, ou através da utilização do comprimento de *Debye* modificado por BS e *Lin*, não poderíamos afirmar se a inclusão de uma superfície molecular mais realista (não esférica), permitindo que a solução eletrolítica se localize mais próximo de grupos carregados, traria alguma dependência do campo de força utilizado. Assim, com o intuito de realizar essa verificação, resolvemos o Modelo Detalhado (MD) através da solução numérica da equação linear de PB. Podemos verificar através da figura 6.12 que essa dependência é nula para a Cab e a CaM, bem como para seus os dois fragmentos (figura 6.13)

Comparando os resultados das figuras 6.12 e 6.11, observamos para a Cab, um pequeno desvio nos resultados de ΔpK para concentração salina de 100 mM e concentração macromolecular de 20 μ M, quando da utilização do MED ($\Delta pK = 2,676$) e MD ($\Delta pK = 3,242$). Assim, devido a forma bastante globular da Cab, a aproximação esférica para a superfície molecular

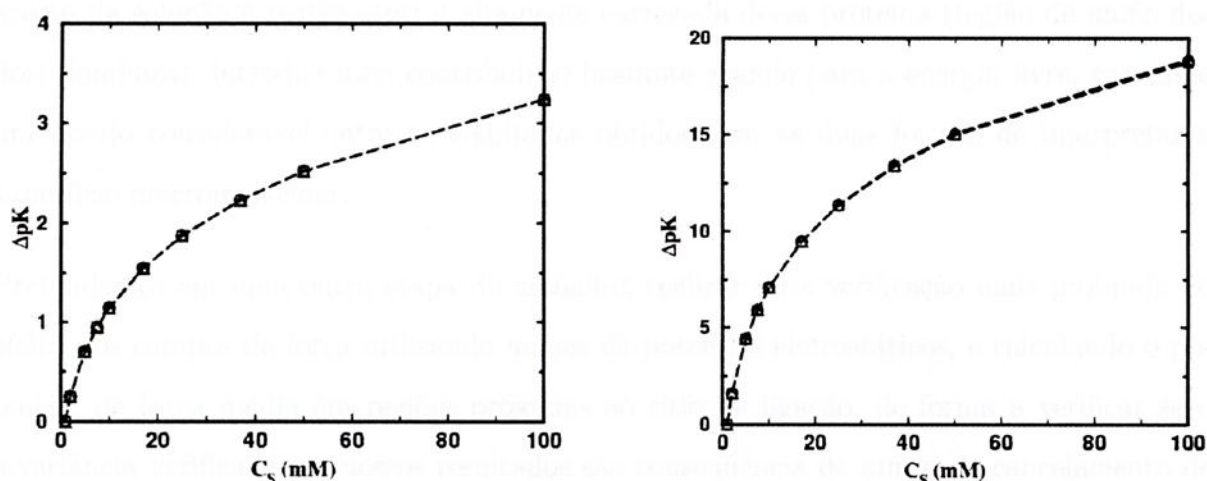


Figura 6.12: Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons de Ca^{+2} à *Calbindina D9k* (gráfico da esquerda) e de quatro íons de Ca^{+2} à *Calmodulina*, obtidos através da solução numérica da equação linear de PB, utilizando o pacote computacional APBS [98], que emprega o método de *elementos finitos multigrid* [99, 100]. As linhas com círculos, quadrados e triângulos se referem aos campos de força GROMOS, PARSE e Reduzido, respectivamente. A força iônica de 1 mM foi tomada como referência.

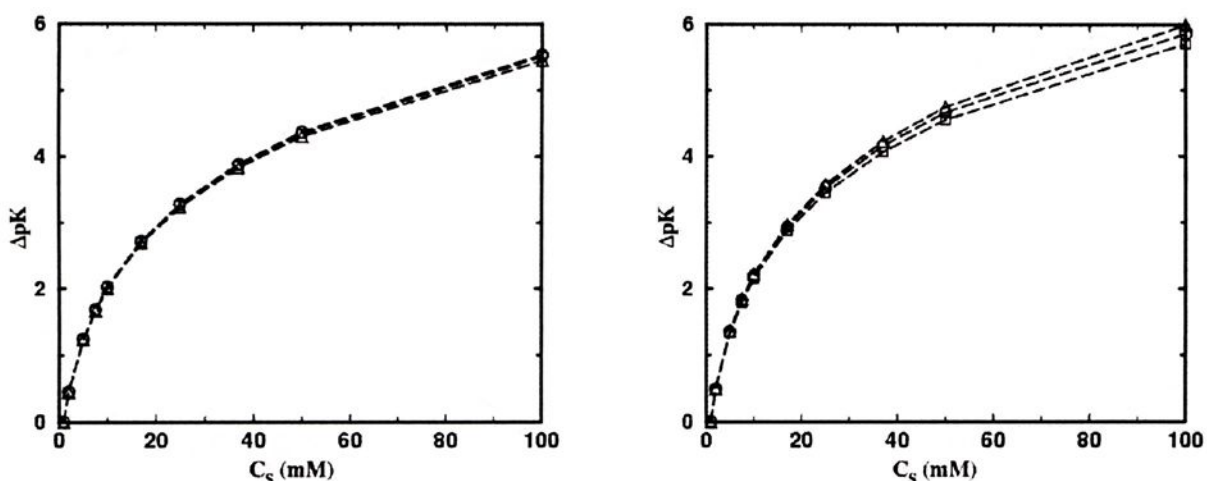


Figura 6.13: Resultados de ΔpK devido a variação da concentração salina para a ligação de dois íons de Ca^{+2} aos fragmentos da *Calmodulina*. O gráfico da esquerda demonstra os resultados do primeiro fragmento, enquanto o gráfico da direita, mostra os do segundo. Outros detalhes de acordo com a figura 6.12

se torna uma boa alternativa, uma vez que soluções analíticas podem ser obtidas. Entretanto, a forma mais irregular da CaM, faz com que essa aproximação não seja razoável, já que o acesso da solução à região central altamente carregada dessa proteína (região de união dos dois domínios), introduz uma contribuição bastante grande para a energia livre, causando um desvio considerável entre os resultados obtidos com as duas formas de interpretar a superfície macromolecular.

Pretendemos em uma outra etapa do trabalho, realizar uma verificação mais profunda do efeito dos campos de força utilizando mapas de potencial eletrostáticos, e calculando o potencial de força média em regiões próximas ao sítio de ligação, de forma a verificar se a invariância verificada em nossos resultados são consequência de um mero cancelamento de erros devido a diferença entre os ΔG s, ou realmente indica que cálculos de propriedades eletrostáticas são realmente independentes de peculiaridades da distribuição de carga da proteína.



Capítulo 7

Conclusões

O presente trabalho confirma a existência de um efeito da concentração macromolecular na constante de ligação de íons à macromoléculas. Esse efeito está relacionado diretamente ao aumento da blindagem eletrostática causada pela existência de outras espécies carregadas em solução (contraíons adicionados para manutenção da eletroneutralidade da célula). Uma maneira de incorporar esse efeito nos modelos baseados na tradicional teoria de *Debye-Hückel* (DH) é a utilização das aproximações de *Beresford-Smith* e *Lin*, que conseguem recuperar esse mecanismo de maneira bastante satisfatória e surpreendente, devido a alta concordância com resultados experimentais, considerando às simplificações realizadas no sistema (continuidade dielétrica, negligenciamento de mudanças estruturais na molécula e interações de curto alcance, etc).

Para baixa carga e concentração macromolecular, verificamos uma ótima concordância entre os resultados da variação da constante de ligação de íons à macromoléculas, utilizando a solução da equação linear e não linear de *Poisson-Boltzmann* (PB), e o método *Monte Carlo* (MC). Entretanto, para proteínas de elevada carga, existe uma divergência entre os resultados de PB e DH, mesmo para baixa concentração dessas proteínas, devido a uma subestimação da blindagem eletrostática prevista pela aproximação de DH. Além disso, a aproximação de DH não consegue recuperar sequer o comportamento qualitativo de ΔpK em relação ao aumento da carga macromolecular, obtido por MC e PB não linear, embora os resultados obtidos através da solução da equação não linear de PB também não concordem



quantitativamente com os resultados obtidos através do método MC. Devido a isso, podemos afirmar que os desvios obtidos entre o modelo de TK e os resultados obtidos pelo método MC, para macromoléculas altamente carregadas, podem ser apenas atenuados pela utilização da equação não linear de PB, podendo ser atribuídos então, ao grande efeito de correlação iônica desprezado pela aproximação de campo médio.

De nossas análises sistemáticas dos campos de força, verificamos que a distribuição de cargas na macromolécula tem uma influência muito pequena no cálculo da variação da energia livre de ligação, sendo que esses cálculos podem ser realizados, com bastante eficiência, utilizando modelos simplificados de macromoléculas, onde detalhes estruturais e distribuições de carga em pequena “escala” (em **átomos** dos aminoácidos, como sugerido pelo campo de força GROMOS, por exemplo) podem ser desprezados. Portanto, a arbitrariedade existente na escolha da utilização de um ou outro campo de força não exerce uma influência significativa nos resultados obtidos nesse nível de aproximação, sugerindo, portanto, a utilização do campo de força reduzido como uma forma de reduzir custos computacionais.

Referências Bibliográficas

- [1] R. E. Babine e S. L. Bender. Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem. Rev.*, 97:1359, 1997.
- [2] R. Hanna e J. A. Doudna. Metal ions in ribozyme folding e catalysis. *Curr. Opin. Chem. Biol.*, 4:166, 2000.
- [3] M. Campbell. *Bioquímica*. Artmed, Brasil, 3th edition, 2000.
- [4] A. Malmendal. *Shaking the EF-Hand – Dynamics e Ion-Binding of Calmodulin e Calbindin D_{9k}*. PhD thesis, Lund University, Lund, Sweden, 1999.
- [5] S. W. Vetter e E. Leclerc. Novel aspects of Calmodulin target recognition e activation. *Eur. J. Biochem.*, 270:404–414, 2003.
- [6] B. G. Vertessy et. al. Simultaneous binding of drugs with different chemical structures to Ca⁺²-Calmodulin: Crystallographic and spectroscopic studies. *Biochemistry*, 37:15300–15310, 1998.
- [7] P. Bamborough e F. E. Cohen. Modeling protein-ligand complexes. *Curr. Opin. Struct. Biol.*, 6:236–241, 1996.
- [8] R. E. Hubbard. Can drugs be designed ? *Curr. Opin. Biotechnol.*, 8:696, 1997.
- [9] N. C. Cohen. *Molecular Modeling in Drug Design*. Academic Press, San Diego, 1996.
- [10] J. Antonsiewicz J. A. McCammon e M. K. Gilson. Prediction of pH-dependence properties of proteins. *J. Mol. Biol.*, 238:415–436, 1994.
- [11] A. H. Elcock e J. A. McCammon. Electrostatic contributions to the stability of halophilic proteins. *J. Mol. Biol.*, 280:731–748, 1998.



- [12] A. H. Juffer. Theoretical calculations of acid-dissociation constants of proteins. *Biochem. Cell Biol.*, 76:198–209, 1998.
- [13] A. R. Leach. *Molecular Modelling – Principles e Applications*. Addison Wesley Longman Limited, Singapore, 1st edition, 1996.
- [14] T. M. Devlin, editor. *Textbook of Biochemistry with Clinical Correlations*. Wiley-Liss, New York, 1997.
- [15] M. F. Perutz. Electrostatic effects in proteins. *Science*, 201:1187–1191, 1978.
- [16] M. K. Gilson e B. H. Honig. Calculation of electrostatic potentials in an enzyme active site. *Nature*, 330:84–86, 1987.
- [17] B. Honig e A. Nicholls. Classical electrostatics in biology e chemistry. *Science*, 268:1144–1149, 1995.
- [18] F. L. B. da Silva. *Statistical Mechanics Studies of Aqueous Solutions and Biomolecular Systems*. PhD thesis, Lund University, Sweden, 2000.
- [19] K. A. Sharp R. Fine e B. Honig. Computer simulations of the diffusion of a substrate to an active site of an enzyme. *Science*, 236:1460–1463, 1987.
- [20] C. E. Woodward e B. Svensson. Potentials of mean force in charged systems: Application to Superoxide Dismutase. *J. Phys. Chem.*, 95:7471–7477, 1991.
- [21] N. Froloff A. Windemuth e B. Honig. On the calculation of binding free energies using continuum methods: Application to MHC class I protein-peptide interactions. *Prot. Sci.*, 6:1293, 1997.
- [22] A. J. McCoy V. Chandana Epa e P. M. Colman. Electrostatic complementarity at protein/protein interfaces. *J. Mol. Biol.*, 268:570–584, 1997.
- [23] D. A. McQuarrie. *Statistical Mechanics*. Harper Collins, New York, 1976.
- [24] D. Frenkel e B. Smit. *Understanding Molecular Simulation: From algorithms to Applications*. Academic Press, San Diego, 1996.
- [25] H. L. Friedman. Ion-ion e ion-solvent interactions – Introduction. *Faraday Discuss. of the Chem. Soc.*, 64:7–15, 1977.



- [26] H. L. Friedman. Electrolyte solutions at equilibrium. *Ann. Rev. Phys. Chem.*, 32:179–204, 1981.
- [27] M. E. Davis e J. A. McCammon. Electrostatics in biomolecular structure e dynamics. *Chem. Rev.*, 90:509–521, 1990.
- [28] D. B. Boyd. Aspects of molecular modeling. In K. B. Lipkowitz e D. B. Boyd, editor, *Reviews in Computational Chemistry*, volume 1, pages 321–354, New York, 1990. VCH Publishers, inc.
- [29] B. Honig K. Sharp e A. Yang. Macroscopic models of aqueous solutions: Biological and chemical applications. *J. Phys. Chem.*, 97:1101, 1993.
- [30] S. J. Weiner et al. A new force field for molecular mechanical simulation of nucleic acids e proteins. *J. Am. Chem. Soc.*, 106:765–784, 1984.
- [31] D. Sitkoff K. A. Sharp e B. Honig. Accurate calculation of hydration free energies using macroscopic solvent model. *J. Phys. Chem.*, 98:1978–1988, 1994.
- [32] D. Sitkoff N. Ben-Tal e B. Honig. Calculation of alkane to water solvation free energies using continuum solvent model. *J. Phys. Chem.*, 100:2744–2752, 1996.
- [33] D. Bashford e K. Gerwert. Electrostatic calculations of the pKa values of ionizable groups in Bacteriorhodopsin. *J. Mol. Biol.*, 224:473–486, 1992.
- [34] W. D. Cornell et. al. A second generation force-field for the simulation of proteins, nucleic acids, e organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.
- [35] W. F. van Gunsteren e H. J. C. Berendsen. *Groningen Molecular Simulation (GROMOS) Library Manual*. Biomos, Groningen, 1987.
- [36] W. R. P. Scott et al. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A*, 103:3596–3607, 1999.
- [37] W. F. van Gunsteren e H. J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications, e perspective in chemistry. *Angew. Chem. Int. Ed. Engl.*, 29:992–1023, 1990.
- [38] P.H. Hünenberger e W. F. Gunsteren. Empirical classical interaction functions for molecular simulation. In W. F. van Gunsteren e P. K. Weiner e A. Wilkinson, editor,



Computer Simulation of Biomolecular Systems, volume 3, pages 3–74, Dordrecht, 1997. ESCOM.

- [39] H. J. C. Berendsen. A glimpse of the holy grail? *Science*, 282:642–643, 1999.
- [40] M. Feig e B. Montgomery Petit. Structural equilibrium of DNA represented with different force fields. *Biophys. J.*, 75:134–149, 1998.
- [41] K. A. Sharp A. Nicholls e S. Sridharan. *Dephi - A macromolecular Electrostatics Modeling Package*. Columbia University, 1998.
- [42] W. Rocchia E. Alexov e B. Honig. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants e multivalent ions. *J. Phys. Chem.*, 105:6507–6514, 2001.
- [43] Y. K. Jang G. Nemethy e H. A. Scheraga. Free-energies of hydration of solute molecules. 3. Application of the hydration shell-model to charged organic molecules. *J. Phys. Chem.*, 91(15):4118–4120, 1987.
- [44] M. K. Gilson K. A. Sharp e B. Honig. Calculating the electrostatic potential of molecules in solution: Method e error assessment. *J. Comp. Chem.*, 9:327–335, 1987.
- [45] P. W. Atkins e R. S. Friedman. *Molecular Quantum Mechanics*. Oxford University Press, New York, 3th edition, 1997.
- [46] M. S. Skaf e F. L. B. da Silva. Explorando as propriedades estruturais e dinâmicas de solventes por simulação computacional. *Química Nova*, 17(6):507–512, 1994.
- [47] B. Jönsson e B. Svensson. Monte Carlo simulation of ion–protein binding. In W. F. van Gunsteren P. K. Weiner e A. Wilkinson, editor, *Computer Simulation of Biomolecular Systems*, volume 2, pages 464–482, Leiden, 1993. ESCOM.
- [48] B. Svensson e B. Jönsson. An efficient simulation technique for electrostatic free energies with applications to Azurim. *J. Comp. Chem.*, 18:370–377, 1995.
- [49] V. Vlachy. Ionic effects beyond Poisson-Boltzmann theory. *Annu. Rev. Phys. Chem.*, 50:145–165, 1999.
- [50] F. Fogolari A. Brigo e H. Molinari. The Poisson-Boltzmann equation for biomolecular electrostatics: A tool for structural biology. *J. Mol. Recognit.*, 15:377–392, 2002.



- [51] F. L. B. da Silva B. Jönsson e R. Penfold. A critical investigation of the Tanford-Kirkwood scheme by means of Monte Carlo simulations. *Prot. Sci.*, 10:1415–1425, 2001.
- [52] F. Fogolari et al. Biomolecular Electrostatics with the Linearized Poisson-Boltzmann Equation. *Biophys. J.*, 76:1–16, 1999.
- [53] S. J. Carvalho R. C. T. Ghiotto e F. L. B. da Silva. Monte Carlo e modified Tanford-Kirkwood results for macromolecular electrostatics calculations. *Biophys. J.* submetido.
- [54] S. Linse B. Jönsson e W. J. Chazin. The effect of protein concentration on ion binding. *Proc. Natl. Acad. Sci. USA*, 92:4748–4752, 1995.
- [55] C. Tanford e J. G. Kirkwood. Theory of protein titration curves I. General equations for impenetrable spheres. *J. Am. Chem. Soc.*, 79:5333–5339, 1957.
- [56] C. Tanford. Theory of protein titration curves II. Calculations for simple models at low ionic strength. *J. Am. Chem. Soc.*, 79:5340–5347, 1957.
- [57] C. R. Cantor e P. R. Schimmel. *The Behavior of Biological Macromolecules*. W. H. Freeman e Company, New York, 1980.
- [58] J. Evenäs A. Malmendal e S. Forsén. Calcium. *Curr. Opin. Chem. Biol.*, 2:293–302, 1998.
- [59] P. W. Atkins. *Physical Chemistry*. Oxford University Press, London, 5th edition, 1995.
- [60] H. Wennerström B. Jönsson e P. Linse. The cell model for polyelectrolyte systems. Exact statistical mechanics relations, Monte Carlo simulations, e the Poisson-Boltzmann approximation. *J. Chem. Phys.*, 76:4665, 1982.
- [61] K. A. Sharp. Polyelectrolyte electrostatics: Salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear Poisson-Boltzmann model. *Biopolymers*, 36:227–243, 1995.
- [62] A. M-Bentley e S. Réty. EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.*, 10:637–643, 2000.



- [63] T. Kesvatera et. al. Focusing the electrostatic potential at EF-hands of Calbindin D_{9K}. titration of acidic residues. *Proteins*, 45(2):129–135, 2001.
- [64] Protein data bank. <http://www.pdb.org>.
- [65] W. Humphrey A. Dalke e K. Schuten. VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.
- [66] R. H. Kretsinger e C. E. Nockolds. Carp muscle calcium-binding protein. *J. Biol. Chem.*, 218(9):3313–3326, 1973.
- [67] C. Branden e J. Tooze. *Introduction to Protein Structure*. Garland Publishing, New York, 1991.
- [68] M. S. Cates et. al. Metal-ion affinity e specificity in EF-hand proteins: Coordination geometry e domain plasticity in Parvalbumina. *Structure*, 7:1269–1278, 1999.
- [69] B. Svensson et. al. Ion binding properties of Calbindin D_{9k} — A Monte Carlo simulation study. *Biochemistry*, 30:5209–5217, 1991.
- [70] T. Kesvatera et. al. Binding of Ca⁺² to Calbindin D_{9k}: Structural stability e function at high salt concentration. *Biochemistry*, 33:14170–14176, 1996.
- [71] B. Svensson et. al. Binding of Ca⁺² to Calmodulin e its tryptic fragments: Theory e experiment. *Biochemistry*, 32:2828–2834, 1993.
- [72] Be. Jönsson. *The Thermodynamics of Ionic Amphiphilie–water systems — A Theoretical Analysis*. PhD thesis, Lund University, Lund, Sweden, 1981.
- [73] P. Linse e B. Jönsson. A Monte Carlo study of the electrostatic interaction between highly charged aggregates. A test of the cell model applied to micellar systems. *J. Chem. Phys.*, 78:3167–3176, 1983.
- [74] H. L. Friedman. *Ionic Solution Theory*. Interscience Publishers, New York, 1962.
- [75] D. Levesque J. J. Weis e J. P. Hansen. Simulation of classical fluids. In K. Binder, editor, *Monte Carlo Methods in Statistical Physics*, volume 5, pages 47–119, Berlin, 1986. Springer-Verlag.
- [76] J. P. Hansen e I. R. McDonald. *Theory of Simple Liquids*. Academic Press, London, 2nd edition, 1986.



- [77] L. Degrève et. al. Monte Carlo simulation for a symmetrical electrolyte next to a charged spherical colloid particle. *J. Chem. Phys.*, 98(11):8905–8909, 1993.
- [78] D. J. Griffiths. *Introduction to Eletrodynamics*. Prentice-Hall, Inc., New Jersey, 3rd edition, 1999.
- [79] D. M. E. Szebenyi e K. Moffat. The refined structure of vitamin-D-dependent calcium-binding protein from bovine intestine – molecular details, ion binding, e implications for the structure of other calcium-binding proteins. *J. Biol. Chem.*, 261:8761–8777, 1986.
- [80] Y. S. Babu C. E. Bugg e W. J. Cook. Structure of Calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.*, 204:191–204, 1988.
- [81] Shelxl-93 – A program for the refinement of crystal structures. <http://www.xray.ncsu.edu/shelxl/contents.html>.
- [82] E. Lindahl B. Hess e D. van der Spoel. Gromacs 3.0: A package for molecular simulation e trajectory analysis. *J. Mol. Model.*, 7(8):306–317, 2001.
- [83] D. van der Spoel et. al. *Gromacs User Manual version 3.0*. Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: <http://www.gromacs.org>, 2001.
- [84] R. Penfold J. Warwicker e B. Jönsson. Electrostatic models for calcium binding proteins. *J. Phys. Chem. B*, 108:8599–8610, 1998.
- [85] N. Guex e M. C. Peitsch. Swiss-model e the Swiss-PDBViewer: An environment for comparative protein modeling. *Electrophoresis*, 18:2714–2723, 1997.
- [86] J. G. Kirkwood. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.*, 2:351–361, 1934.
- [87] P. H. Guadagnini R. E. Bruns e A. A. de Souza. Cargas atômicas em moléculas. *Química Nova*, 19(2):148–155, 1996.
- [88] B. Svensson B. Jönsson e C. E. Woodward. Electrostatic contributions of the binding of Ca^{2+} in Calbindin mutants. A Monte Carlo study. *Biophys. Chem.*, 38:179–183, 1990.



- [89] L. G. Dias et. al. A simple method for the fast calculation of charge redistribution of solutes in an implicit solvent model. *Chem. Phys. Lett.*, 282(2):237–243, 2002.
- [90] K. Shimizu et. al. Calculation of the dipole moment for polypeptides using the Generalized Born-Electronegativity Equalization method: Results in vacuum e continuum-dielectric solvent. *J. Phys. Chem. B*, 2004. submetido.
- [91] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, Oxford, 1987.
- [92] S. L. Carnie e G. M. Torrie. The statistical mechanics of the electrical double layer. *Adv. Chem. Phys.*, LVI:141–253, 1984.
- [93] C.W. Outhwaite e L.B. Bhuiyan. A modified Poisson-Boltzmann analysis of the electric double layer around an isolated spherical macroion. *Mol. Phys.*, 74(2):367–381, 1991.
- [94] L. Degève e M. Lozada-Cassou e E. Sánchez and E. Gonzáles-Tovar. Monte carlo simulation for a symmetrical electrolyte next to a charged spherical colloid particle. *J. Chem. Phys.*, 98(11):8905, 1993.
- [95] B. Jayaran e D. L. Beveridge. Modeling DNA in aqueous solutions: Theoretical e computer simulation studies on the ion atmosphere of DNA. *Annu. Rev. Biophys. Biomol. Structure*, 25:367–394, 1996.
- [96] A. Teso E. Drigo Filho e A. Agostinho Neto. Solution of the Poisson-Boltzmann equation for a system with four ionic species. *J. Math. Biol.*, 35:814–824, 1997.
- [97] K. S. Schmitz, editor. *Macroion Characterization: From Dilute Solutions to Complex Fluids*. American Chemistry Society, Washington, 1994.
- [98] N. A. Baker et. al. Electrostatics of nanosystems: Application to microtubules e the ribosome. *Proc. Natl. Acad. Sci. USA*, 98:10037–10041, 2001.
- [99] M. Holst e F. Saied. Multigrid solution of the Poisson-Boltzmann equation. *J. Comp. Chem.*, 14(1):105–113, 1993.
- [100] M. Holst N. Baker e F. Wang. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. algorithms e examples. *J. Comp. Chem.*, 21(15):1319–1342, 2000.



- [101] L. Onsager. Theories of concentrated electrolytes. *Chem. Rev.*, 13:73–89, 1933.
- [102] K. A. Sharp e B. Honig. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.*, 94:7684–7692, 1990.
- [103] F. Fogolari e J. M. Briggs. On the variational approach to the Poisson-Boltzmann free energy. *Chem. Phys. Lett.*, 281:135–139, 1997.
- [104] H. Zhou. Macromolecular electrostatic energy within the nonlinear Poisson-Boltzmann equation. *J. Chem. Phys.*, 100(4):3152–3162, 1994.
- [105] B. Jönsson. *The Thermodynamics of Ionic Amphiphile-Water Systems. A Theoretical Analysis*. PhD thesis, Lund University, Sweden, 1981.
- [106] G. Arfken. *Mathematical Methods for Physicists*. Academic Press, London, 1970.
- [107] J. Israelachvili. *Intermolecular e Surface Forces*. Academic Press, London, 2nd edition, 1991.
- [108] F. Fogolari et. al. Electrostatic effects in Homeodomain-DNA interactions. *J. Mol. Biol.*, 267:368–381, 1997.
- [109] T. Tomé e M. J. de Oliveira. *Dinâmica Estocástica e Irreversibilidade*. Editora da Universidade de São Paulo, São Paulo, 2001.
- [110] S. R. A. Salinas. *Introdução à Física Estatística*. Editora da Universidade de São Paulo, São Paulo, 1999.
- [111] B. Widom. Some topics in the theory of fluids. *J. Chem. Phys.*, 39:2808–2812, 1963.
- [112] B. R. Svensson e C. E. Woodward. Widom's method for uniform e non-uniform electrolyte solutions. *Mol. Phys.*, 64:247–259, 1988.
- [113] B. Beresford-Smith e D. Y. C. Chan. Electrical double-layer interactions in concentrated colloidal systems. *Faraday Discuss. Chem. Soc.*, 76:65–75, 1983.
- [114] S. C. Lin W. I. Lee e J. M. Shurr. Brownian motion of highly charged poly(L-lysine). Effects of salt e polyion concentration. *Biopolymers*, 17:1041–1064, 1978.
- [115] G. M. Morris et. al. Automated docking using a Lamarckian genetic algorithm e an empirical binding free energy function. *J. Comp. Chem.*, 19(14):1639–1662, 1998.



[116] J. D. Madura et. al. Biological applications of electrostatic calculations e Brownian dynamics simulations. In K. B. Lipkowitz e D. B. Boyd, editor, *Reviews in Computational Chemistry*, volume 5, pages 229–267, New York, 1994. VCH Publishers, inc.

Apêndice A

Campos de Força para o Acetato

Aqui, a descrição de campo de força utilizado para a simulação das cargas pontuais de um sistema. A parametrização dos campos é a mesma que a utilizada no artigo principal.



Figura A.1: Representação estrutural do íon acetato.

Átomo	CHARMM	GROMOS	PAR-1	Reboredo
C ₁	-0.160	0.000	0.000	0.000
O ₁	0.360	0.270	0.100	0.000
O ₂	-0.600	-0.630	-0.500	-0.500
C ₂	-0.000	0.000	0.000	-0.000
H ₁	0.000	0.000	0.000	0.000
H ₂	0.000	0.000	0.000	0.000
H ₃	0.000	0.000	0.000	0.000

Tabela A.1: Cargas pontuais para os átomos do íon acetato.

Apêndice A

Campos de Força para o Acetato

Abaixo, a descrição do campo de força utilizado para atribuição das cargas parciais ao íon acetato. A nomenclatura dos átomos é definida de acordo com a figura abaixo:

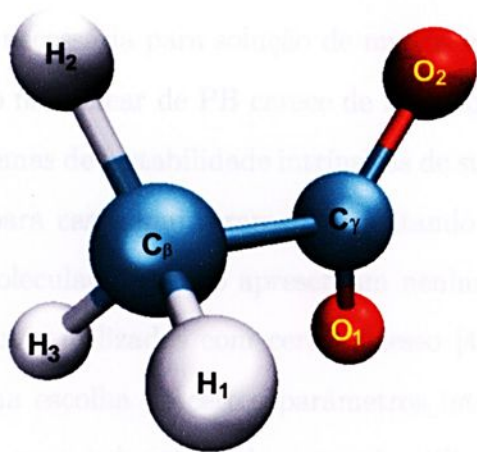


Figura A.1: Representação esquemática do íon acetato

Átomo	CHARMM	GROMOS	PARSE	Reduzido
C _β	-0,160	0,000	0,000	0,000
C _γ	0,360	0,270	0,100	0,000
O ₁	-0,600	-0,635	-0,550	-0,500
O ₂	-0,600	-0,635	-0,550	-0,500
H ₁	0,000	0,000	0,000	0,000
H ₂	0,000	0,000	0,000	0,000
H ₃	0,000	0,000	0,000	0,000

Tabela A.1: Cargas parciais para os átomos do íon acetato.

Apêndice B

Problemas com a solução numérica da Equação de *Poisson-Boltzmann*

Embora reconhecidamente necessária para solução de muitos modelos cuja teoria de DH se torna ineficiente, a equação não linear de PB carece de métodos robustos para obtenção de sua solução, devido a problemas de instabilidade intrínsecas de sua forma não linear. Soluções analíticas são disponíveis para casos muito raros, necessitando de métodos numéricos para o estudo de sistemas biomolecular, que não apresentam nenhum tipo de simetria. Muitos métodos tem sido propostos e utilizados com certo sucesso [42]. Entretanto, sua solução carece de arbitrariedades na escolha de certos parâmetros intrínsecos do método, como a distância entre os pontos de uma rede retangular, quando utilizando o método de *diferenças finitas* [13]. Neste apêndice pretendemos mostrar alguns resultados errôneos obtidos por nós na utilização de alguns pacotes computacionais utilizados por nós, bem como sugerir um protocolo para esses cálculos de forma a amenizar os efeitos da instabilidade numérica dessa equação.

Os dois métodos utilizados por nós para solução da equação não linear de PB, estão descritos abaixo:

Método de Diferenças finitas: Este método divide o espaço em uma rede, geralmente cúbica, atribuindo o potencial eletrostático, a densidade de carga, e a acessibilidade dos íons da solução, a cada um dos pontos dessa rede, e a permissividade dielétrica às linha



que unem esses pontos. A equação de PB é convertida em um conjunto de equações de diferenças para cada ponto da grade, de forma que um método iterativo é utilizado para garantir a consistência da solução em cada ponto da rede com os seus pontos vizinhos.

Método de Elementos Finitos: O método de *elementos finitos*, divide o espaço em um conjunto de domínios, os elementos finitos, de forma que uma solução expandida em série de potências é atribuída a cada nó destes domínios, formando um conjunto sistema de equações lineares, cuja solução, levando em conta certas condições de contorno, nos fornece o potencial em todos os nós desses domínios.

Ambos os métodos citados acima possuem dois parâmetros necessários para sua execução: (a) o espaçamento entre os pontos no qual a solução é atribuída e (b) o número desses pontos, diretamente relacionada ao tamanho da “caixa” no qual a molécula será inserida. Esses parâmetros são arbitrários e, por isso, alguns cuidados devem ser tomados:

- Para moléculas altamente carregadas o tamanho da “caixa” deverá ser grande, de forma que o potencial na borda dessa caixa possa satisfazer o potencial prescrito por *Debye-Huckel*. Dessa forma, esse potencial pode ser obtido analiticamente e usado como uma das condições de contorno do método.
- O espaçamento entre os pontos da rede deve ser variado em um pequeno intervalo, de forma a garantir que a solução é independente da forma como as cargas são atribuídas nos pontos da rede.
- Como o mesmo objetivo do item acima, a molécula deve ser rotacionada e transladada em distâncias não proporcionais ao espaçamento dos pontos da rede [116].

Um procedimento largamente utilizado para garantir a consistência dos resultados, e obter potências em regiões específicas da molécula, é o chamado **método de focagem** [13]. Nesta situação, um cálculo grosseiro é realizado (grande espaçamento entre os pontos da rede, e consequentemente grande distanciamento da superfície da molécula à borda da caixa), de forma que um mapeamento do potencial eletrostático é gerado e utilizado como condição de contorno para cálculos mais precisos.



Um exemplo da instabilidade desses método é mostrado no gráfico abaixo, onde uma variação no tamanho da caixa, resultam em divergências gritantes nos resultados finais. O resultado correto (linha contínua) foi obtido utilizando uma caixa de 48,3 Å, com um espaçamento de 0,3 Å entre os pontos da rede. Diminuindo o tamanho da caixa para 31,0 Å, obtemos resultados completamente divergentes (linha tracejada), inclusive qualitativamente, onde é verificada uma diminuição da blindagem eletrostática com o aumento da concentração salina.

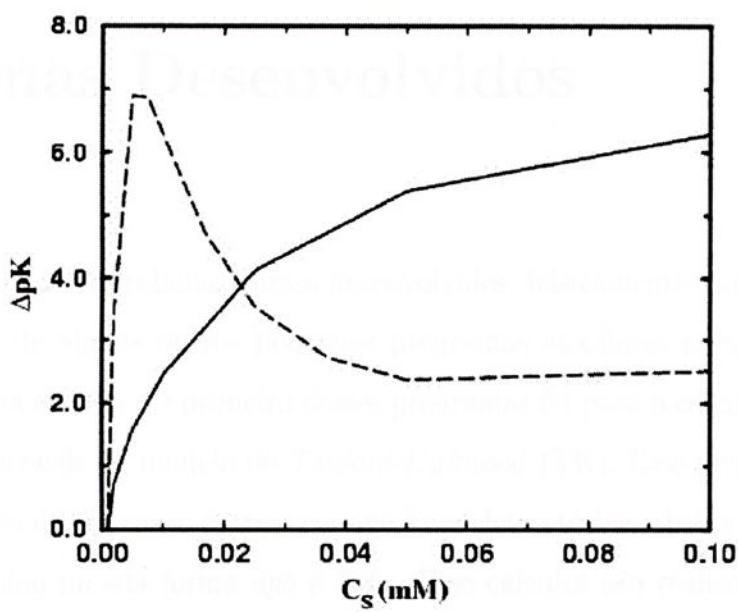


Figura B.1: Cálculos de ΔpK devido a variação da concentração salina para a ligação de dois íons bivalentes positivos. A proteína foi representada segundo o ME (veja seção 4.1) com raio de 14 Å e carga $Z_p = -20$. A posição dos sítios de ligação foi obtida do arquivo PDB-ID 3icb (Calbindina D_{9k}). Os cálculos foram obtidos através da utilização do pacote computacional APBS [98], através do método de *elementos finitos multigrid* [99, 100]. A linha tracejada demonstra os resultados obtidos com uma caixa de 31 Å e espaçamento entre os pontos da rede de 0,47 Å. A linha sólida demonstra os resultados obtidos com uma caixa de 48,3 Å e espaçamento entre os pontos da rede de 0,3 Å. A força iônica de 1 mM foi tomada como referência.



Apêndice C

Programas Desenvolvidos

Durante a execução deste trabalho, foram desenvolvidos, basicamente, dois programas computacionais, além de alguns outros pequenos programas auxiliares para manipulação dos arquivos de entrada e saída. O primeiro desses programas foi para o cálculo da energia livre de ligação (ΔG) através do modelo de *Tanford-Kirkwood* (TK). Esse programa obtém esta propriedade através da diferença entre a energia livre eletrostática, obtida através da relação 5.45, para a proteína na sua forma *apo* e *holo*. Esse cálculos são realizados para a versão tradicional de *Debye-Hückel* e para as aproximações de *Beresford-Smith* e *Lin*. Um exemplo do arquivo de saída deste programa é mostrado na figura abaixo:

=====	=====
Debye-Huckel	Lin
=====	=====
Concentracao de Sal: 0.00100	Concentracao de Sal: 0.00100
Concentracao da Biomolecula: 0.00002	Concentracao da Biomolecula: 0.00002
Carga da Biomolecula: -7.000	Carga da Biomolecula: -7.000
Kappa em 1/metro: 103844251.02503	Kappa em 1/metro: 129701427.95930
Pot. Qui. de Excesso dos Ligantes: 0.28415	Pot. Qui. de Excesso dos Ligantes: 0.35141
Energia Livre com Ligantes Livres: 21.30329	Energia Livre com Ligantes Livres: 20.98086
Energia Livre com Ligantes Ligados: -12.75124	Energia Livre com Ligantes Ligados: -12.76691
Energia Livre Eletrostatica: -34.05453	Energia Livre Eletrostatica: -33.74777
=====	=====
Beresford-Smith	
=====	
Concentracao de Sal: 0.00100	
Concentracao da Biomolecula: 0.00002	
Carga da Biomolecula: -7.000	
Kappa em 1/metro: 107417328.50856	
Pot. Qui. de Excesso dos Ligantes: 0.29352	
Energia Livre com Ligantes Livres: 21.25747	
Energia Livre com Ligantes Ligados: -12.75526	
Energia Livre Eletrostatica: -34.01272	
=====	

Figura C.1: Exemplo de um arquivo de saída do programa que resolve o modelo de TK.



O outro programa desenvolvido resolve o modelo proposto neste trabalho através do método de *Monte Carlo*. Esse programa segue, basicamente, o mesmo algoritmo apresentado na seção 5.4. Entretanto, introduzimos o conceito de *macro-passos* e *micro-passos*, de forma que o algoritmo pode ser considerado conforme abaixo:

1. Iniciar o contador de *macro-passos*.
2. Iniciar o contador de *micro-passos*.
3. Propor o movimento de uma partícula móvel.
4. Verificar a aceitação do movimento através do critério de *Metropolis*.
5. Voltar ao passo 3 até mover todas as partículas.
6. Chamar a rotina para cálculo do potencial químico (*Método de Widom*).
7. Voltar ao passo 2 até concluir todos os *micro-passos*.
8. Chamar a rotina para calcular *função distribuição radial*.
9. Calcula energia.
10. Voltar ao passo 1 até concluir todos os *macro-passos*.
11. Concluir *função distribuição radial*.
 Concluir cálculo do potencial químico.

O procedimento acima é realizado duas vezes. A primeira para a “equilibração” do sistema, onde os passos referentes ao *método de Widom* e *função distribuição radial* são suprimidos. Tendo o sistema atingido o estado de equilíbrio, o procedimento é repetido para o cálculo dessas propriedades. O estado de equilíbrio é alcançado quando a energia calculada no passo 9 oscila em torno em um valor médio, não tendo mais tendência de queda.

O cálculo do potencial químico é realizado da seguinte forma: Em cada *micro-passo* realizamos um procedimento de inserção, conforme descrito na seção 5.5, e realizados os cálculos de ΔG^{ext} , ΔG^{el} e ΔG^{hc} . No final de cada bloco de *micro-passo*, esses valores são utilizados para o cálculo das médias da relação 5.53. No final do programa, teremos vários valores de



potencial químico para cada *macro-passo* realizado. Assim, uma média aritmética é realizada para o cálculo do potencial químico final e de seu desvio padrão.

A *função distribuição radial* é calculada em cada *macro-passo*. Em cada um desses procedimentos verificamos o número de partículas móveis em cada “fatia” da nossa célula. No final do programa, dividimos esse número pelo volume de cada uma dessas “fátias”, pelo número de *macro-passos*, e pelo número de partículas dentro de cada “fatia”, caso o sistema fosse ideal.



