



**UNIVERSIDADE ESTADUAL PAULISTA**  
**“JÚLIO DE MESQUITA FILHO”**  
**Campus de São José do Rio Preto**

**GUILHERME VOLPE BOSSA**

**DETERMINAÇÃO DO GRAU DE IONIZAÇÃO DE**  
**AMINOÁCIDOS POLARES CARREGADOS**

**SÃO JOSÉ DO RIO PRETO – SÃO PAULO**

**2013**

**GUILHERME VOLPE BOSSA**

**DETERMINAÇÃO DO GRAU DE IONIZAÇÃO DE  
AMINOÁCIDOS POLARES CARREGADOS**

Dissertação apresentada para obtenção do título de Mestre em Biofísica Molecular, área de Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus São José do Rio Preto.

Orientadores: Prof. Dr. Augusto Agostinho Neto (*in memoriam*)

Prof. Dr. Elso Drigo Filho

Co-orientadora: Dra. Tereza Pereira de Souza

**São José do Rio Preto, 22 de Março de 2013**

**GUILHERME VOLPE BOSSA**

**DETERMINAÇÃO DO GRAU DE IONIZAÇÃO DE  
AMINOÁCIDOS POLARES CARREGADOS**

Dissertação apresentada para obtenção do título de Mestre em Biofísica Molecular, área de Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus São José do Rio Preto.

**BANCA EXAMINADORA**

Prof. Dr. Elso Drigo Filho

Professor Titular

UNESP - São José do Rio Preto -SP

Orientador

Profa. Dra. Iolanda Midea Cuccovia

Professor Associado

USP - São Paulo -SP

Prof. Dr. Marcelo Andres Fossey

Professor Assistente Doutor

UNESP - São José do Rio Preto - SP

**São José do Rio Preto, 22 de Março de 2013**

*Ao meu saudoso mestre e amigo Augusto*

*Não é nossa função controlar todas as marés do mundo, mas sim fazer o que pudermos para socorrer os tempos em que estamos inseridos, erradicando o mal dos campos que conhecemos, para que aqueles que viverem depois tenham terra limpa para cultivar. Que tempo encontrarão não é nossa função determinar.*

Gandalf, the gray.

# Agradecimentos:

Agradeço a Deus por tudo e todo o companheirismo desde sempre, pois toda a minha formação e o desenvolvimento deste trabalho só se tornaram possíveis graças a Sua presença, a qual devo infinita gratidão.

Sou grato aos meus pais, Clavio e Rosana e a minha irmã, Carol, pelo carinho e motivação dados, pela atenção dispensada e pelo esforço em me auxiliar e tornar muitas coisas mais fáceis para mim. Obrigado.

Agradeço enormemente a Bruna Marti Trevisan por, praticamente durante todo o meu mestrado, ter sido presente, atenciosa e incentivadora. Muito obrigado pelas revisões, sugestões e pelas conversas sempre úteis e, principalmente, pela ajuda e carinho demonstrados.

Meu profundo e sincero "muito obrigado" é para o meu orientador e amigo Augusto Agostinho Neto que, ao longo de nossa convivência nestes seis anos, se demonstrou um verdadeiro Mestre, colaborando com tudo o que foi possível. Professor, obrigado por manter a porta da sua sala sempre aberta a quem precisasse de ajuda e por ser uma pessoa de altíssima competência e nobreza.

Como infelizmente a morte faz parte da vida, em meados desse ano nosso querido mestre faleceu. Em decorrência disso, muitas pessoas aqui do departamento demonstraram, evidentemente, apoio e preocupação com os alunos do Augusto. Dentre tais pessoas, agradeço aos professores Vitor, Zé Ruggiero, Fátima e Chahine. Agradeço muitíssimo aos professores Elso e Tereza por me adotarem como orientando. Pelas excelentes perguntas, respostas e gargalhadas vocês dois tem sido cruciais para a realização disso tudo.

Muito obrigado, também, aos meus amigos: Daniel por ser, sinceramente, o (co) orientador/mãezona de toda a nossa turma; aos “finados” Rodolfo e Vinicius Andrea pelas palavras comicamente sábias; ao Goiás pela autoconfiança passada e pelo lendário drible da lombada propagado às garotas do campus; ao Davi pelas mais variadas discussões e companheirismo; ao Tiba pelo apoio, idéias compartilhadas, AGEs e classic rock; ao PV, por ser um dos caras mais animados e capacitados na arte de falar borracha; ao Jesus, herdeiro de Shannon, pela cultura transmitida, pela companhia em festas e pela casa a qual retornávamos geralmente embaixo de chuva; ao Carlão,

detentor da peixeira mais rápida do oeste; ao Flávio, Viniboy, Miriam e Musa com quem tive a sorte de partilhar a mesma quente sala e realizar diversas conversas sobre vários assuntos. Agradeço ao Pedro Garcia, novato em nossa sala, mas um amigo de grande estima e belas amigas.

Agradeço ao contínuo e efetivo apoio financeiro recebido da agência FAPESP, essencial para a conclusão deste trabalho.

## Resumo:

Proteínas e peptídeos são constituídos por subunidades estruturalmente mais simples chamadas aminoácidos. Uma importante propriedade destes é que, dependendo das características do meio (tais como pH e concentração iônica), os seus grupos ionizáveis podem ceder prótons e, assim, adquirir carga elétrica não nula. Tal carga influenciará na eficiência da formação de ligações peptídicas e em interações proteína-ligante, por exemplo. Partindo da hipótese de que a diferença entre os valores de pK dos grupos ionizáveis isolados e destes como partes constituintes de um aminoácido é devida, principalmente, à interação eletrostática adicional que se atribui à presença de grupos vizinhos, elaborou-se um modelo que emprega a forma linearizada da equação de Poisson-Boltzmann para o estudo de propriedades físico-químicas de moléculas com três grupos ionizáveis. Neste trabalho tal modelo foi aplicado aos aminoácidos: Aspartato, Glutamato, Cisteína, Tirosina, Arginina, Lisina e Histidina. Calcularam-se os valores de pK e as respectivas cargas elétricas médias de tais moléculas. Como os resultados obtidos concordaram com aqueles oriundos de trabalhos experimentais, o modelo teórico foi expandido para tratar de di, tetra, pentapeptídeos e de resíduos de lisina e glutamato da proteína *Staphylococcal Nuclease*. Os valores do Fator de Correlação de Pearson calculados para ambas proteínas e peptídeos são superiores a 0,99, fato este que evidencia a eficiência e versatilidade do modelo ao reproduzir valores de pK reportados por outros autores.

**Palavras-chave:** pK de biomoléculas, Aproximação de campo médio, Equação de Poisson-Boltzmann, Carga elétrica média.



# Abstract:

Proteins and peptides are composed of subunits structurally simpler called amino acids. An important property of these is that, depending on the medium characteristics (such pH and ionic concentration), its ionizable groups may provide protons and thereby acquire a nonzero electric charge. Such charge will affect the formation of peptide bond and protein-ligand interactions, for example. Assuming that the difference between pK values of the isolates ionizable groups and of these as constituents parts of an amino acid is mainly due to the extra electrostatic interaction attributed to the presence of neighboring groups, was developed a structure-based model that employs the linearized form of the Poisson-Boltzmann equation for the study of physicochemical properties of molecules with three ionizable groups. In this work it was applied to the amino acids: aspartate, glutamate, cysteine, tyrosine, arginine, lysine and histidine. The pK values and respective mean electric charges were calculated. As the calculated values agreed with those from experimental studies, the theoretical model has been expanded to the treatment of di, tetra, pentapeptides and *Staphylococcal Nuclease* residues. The Pearson Correlation Factor calculated for both proteins and peptides are above 0.99, what points to the effectiveness and versatility of the model to reproduce pK values reported by other works.

**Key-words:** pK of biomolecules, Mean field approximation, Poisson-Boltzmann equation, Mean electric charge.

# Lista de Tabelas e Figuras:

<i>Figura 1.1: Representação de um aminoácido. A cadeia lateral é designada por R. ....</i>	<i>20</i>
<i>Figura 1.2: Pro, Val, Ala, Gly, Ile e Leu, respectivamente da esquerda para a direita. Átomos de carbono em preto, oxigênio em vermelho, nitrogênio em azul e hidrogênio em branco. Correspondência mantida em todas as figuras. ....</i>	<i>211</i>
<i>Figura 1.3: Ser e Thr, respectivamente da esquerda para a direita.....</i>	<i>211</i>
<i>Figura 1.4: Trp, Tyr e Phe, respectivamente da esquerda para a direita.....</i>	<i>222</i>
<i>Figura 1.5: His, Lys e Arg, respectivamente da esquerda para a direita.....</i>	<i>222</i>
<i>Figura 1.6: Glu e Asp, respectivamente da esquerda para a direita.....</i>	<i>222</i>
<i>Figura 1.7: Gln e Asn, respectivamente da esquerda para a direita.....</i>	<i>233</i>
<i>Figura 1.8: Cys e Met, respectivamente da esquerda para a direita. Átomo de Enxofre em amarelo. ....</i>	<i>233</i>
<i>Tabela 1.1: Valores das concentrações iônicas para os meios intra e extracelular. Adaptada de Alberts.....</i>	<i>244</i>
<i>Figura 1.9: Aminoácidos estudados e respectivos valores de pK. ....</i>	<i>277</i>
<i>Tabela 1.2: pK e estrutura dos grupos amina e ácido carboxílico.....</i>	<i>277</i>
<i>Figura 2.1: Representação da modelagem dos aminoácidos: <math>\epsilon_1</math> é a constante dielétrica do meio 1 (com valor igual a 2,0) <math>\epsilon_2</math>, é a constante dielétrica da água, cujo valor é dado pela eq. (2.1). Os outros parâmetros serão explicados logo a seguir.....</i>	<i>30</i>
<i>Figura 2.2: Variação da <math>\epsilon_2</math> com a temperatura para duas distintas concentrações de NaCl.....</i>	<i>31</i>
<i>Tabela 2.1: Valores de <math>pK_1</math> e <math>pK_2</math> intrínsecos para os grupos 1 e 2.....</i>	<i>322</i>
<i>Tabela 2.2: Aminoácido na forma completamente ionizada e tratamento dado a suas cadeias laterais com os respectivos valores de pK intrínseco. Em verde, cadeias ácidas e, em rosa, cadeias básicas.....</i>	<i>333</i>
<i>Tabela 2.3: Aminoácidos com os respectivos valores de raios dos grupos ionizáveis e distâncias entre tais. Os índices seguem a notação definida nesta mesma seção. Valores calculados pelo PyMOL. ....</i>	<i>344</i>
<i>Figura 2.3: Esfera que engloba cada par de grupos ácido e base e delimita dois meios distintos. No meio 2 há água e íons em solução.....</i>	<i>35</i>
<i>Figura 3.1: Distâncias entre um ponto P qualquer e os centros dos grupos ácido e base. ....</i>	<i>433</i>
<i>Figura 4.1: Representação de um sistema de duas cargas dentro de uma esfera de raio R. A carga <math>q_1</math> dista a da origem e, a carga <math>q_2</math>, b.....</i>	<i>51</i>

<i>Tabela 5.1: Aminoácidos e respectivos valores de constante dielétrica efetiva em 25°C</i>	555
<i>Figura 5.1: Cargas de sinais contrários dispostas 2Å da origem e internas à esfera de raio 5Å. No ponto P, distante r da origem, é onde se calculou o potencial.</i>	56
<i>Figura 5.2: Constante Dielétrica Efetiva versus Posição (Å) para um sistema de duas cargas. 25°C.</i>	57
<i>Figura 5.3: Constante Dielétrica Efetiva versus Posição (Å). Zoom na interface de um sistema de duas cargas. 25°C.</i>	57
<i>Figura 5.4: Constante Dielétrica Efetiva versus Posição (Å). Comparação entre temperaturas.</i>	58
<i>Figura 5.5: Constante Dielétrica Efetiva versus Posição (Å). Zoom na interface e comparação entre temperaturas.</i>	58
<i>Figura 6.1.1: Arginina (Arg). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.2: Ácido Aspártico (Asp). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.3: Cisteína (Cys). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.4: Ácido Glutâmico (Glu). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.5: Histidina (His). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.6: Lisina (Lys). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	60
<i>Figura 6.1.7: Tirosina (Tyr). Grau de Ionização (<math>\alpha</math>) versus pH.</i>	61
<i>Tabela 6.1: Valores de pK* calculados pelo modelo, provenientes da literatura e diferença entre ambos. Solução com 150mM de NaCl e 25°C. Diferença (%) entre os Calculados e os provenientes da Literatura.</i>	62
<i>Figura 6.2.1: Variação de pK1* com a Conc. Iônica para Tyr, Glu, Cys e Asp.</i>	63
<i>Figura 6.2.2: Variação de pK1* com a Conc. Iônica para Arg, His e Lys.</i>	63
<i>Figura 6.2.3: Variação de pK2* com a Conc. Iônica para Tyr, Glu, Cys e Asp.</i>	64
<i>Figura 6.2.4: Variação de pK2* com a Conc. Iônica para Arg, Lys e His.</i>	64
<i>Figura 6.2.5: Variação de pK3* com a Conc. Iônica para Tyr, Glu, Cys e Asp.</i>	64
<i>Figura 6.2.6: Variação de pK3* com a Conc. Iônica para Arg e Lys.</i>	64
<i>Figura 6.2.7: Variação de pK3* com a Conc. Iônica para His.</i>	64
<i>Figura 6.3.1: Carga Elétrica Média Total versus pH. 36,5°C.</i>	66
<i>Figura 6.3.2: Carga Elétrica Média versus pH para a Arginina. 25°C.</i>	66
<i>Figura 6.3.3: Carga Elétrica Média versus pH para a Histidina. 25°C.</i>	67
<i>Figura 6.3.4: Carga Elétrica Média versus pH para a Lisina. 25°C.</i>	67
<i>Figura 6.3.5: Carga Elétrica Média versus pH para o Ácido Aspártico. 25°C.</i>	67
<i>Figura 6.3.6: Carga Elétrica Média versus pH para o Ácido Glutâmico. 25°C.</i>	67
<i>Figura 6.3.7: Carga Elétrica Média versus pH para a Cisteína. 25°C.</i>	67
<i>Figura 6.3.8: Carga Elétrica Média versus pH para a Tirosina. 25°C.</i>	67

<i>Figura 6.3.9: Carga Elétrica Média Total versus pH para Tyr, His e Cys. 150mM de NaCl e 36,5°C.....</i>	<i>68</i>
<i>Figura 6.3.10: Carga Elétrica Média Total versus pH para Glu e Asp. 150mM de NaCl e 36,5°C.....</i>	<i>68</i>
<i>Figura 6.3.11: Carga Elétrica Média Total versus pH para Lys e Arg. 150mM de NaCl e 36,5°C.....</i>	<i>69</i>
<i>Tabela 6.2: Valores de Carga Elétrica Média Total em 36,5°C.....</i>	<i>70</i>
<i>Tabela 7.1: Valores de pK* calculados pelo modelo e aqueles provenientes da literatura nas correspondentes condições de temperatura e concentração iônica.....</i>	<i>74</i>
<i>Figura 7.1: Gráfico de correlação entre valores de pK* calculados e presentes na literatura: Gurd para gly-pentapeptídeos e Thurlkil para ala-pentapeptídeos.....</i>	<i>74</i>
<i>Figura 8.1: Representação de parte da proteína SNase. As cadeias laterais ionizáveis, algumas destas indicadas por setas, serão as levandas em consideração pelo modelo.....</i>	<i>77</i>
<i>Tabela 8.1: Resultados para a Constante Dielétrica Efetiva (CDef) e pKi* dos resíduos na proteína Staphylococcal Nuclease (SNase). Comparação com valores fornecidos por Isom e Castañeda.....</i>	<i>79</i>
<i>Figura 8.2: Correlação entre valores de pK* calculados para os resíduos de SNase e comparação de distâncias de cut-off. Referências: Isom para esíduos de Lys; Castañeda para resíduos de Glu.....</i>	<i>80</i>

# Lista de Símbolos e Abreviações:

[A]	Concentração da espécie ácida.
Ala	Alanina.
Asn	Asparagina.
Arg	Arginina.
Asp	Aspartato ou Ácido Aspártico.
[B]	Concentração da espécie básica.
C <sub>α</sub>	Carbono alfa.
Cys	Cisteína.
Gln	Glutamina.
Glu	Glutamato ou Ácido Glutâmico.
Gly	Glicina.
CD <sub>1</sub>	Constante dielétrica do meio 1.
CD <sub>2</sub>	Constante dielétrica do meio 2.
CD <sub>ef</sub>	Constante dielétrica efetiva do aminoácido em solução.
<i>e</i>	Carga elementar (1,6.10 <sup>-19</sup> C).
$\vec{E}_1$	Campo elétrico médio nas vizinhanças do grupo ácido.
$\vec{E}_2$	Campo elétrico médio nas vizinhanças do grupo básico.
$\vec{E}_3$	Campo elétrico médio nas vizinhanças do grupo ionizável da cadeia lateral.
[H <sup>+</sup> ]	Concentração de prótons no bulk.
[H <sup>+</sup> ] <sub>1</sub> <sup>*</sup>	Concentração de prótons nas vizinhanças do grupo ácido acético, corrigida pelo potencial eletrostático médio.
[H <sup>+</sup> ] <sub>2</sub> <sup>*</sup>	Concentração de prótons nas vizinhanças do grupo amina, corrigida pelo potencial eletrostático médio.
[H <sup>+</sup> ] <sub>3</sub> <sup>*</sup>	Concentração de prótons nas vizinhanças do grupo ionizável da cadeia lateral, corrigida pelo potencial eletrostático médio.
His	Histidina.
I <sub>m</sub>	Função de Bessel modificada de primeira espécie.
Ile	Isoleucina.

k	Constante de Boltzmann.
$K_1$	Constante intrínseca de equilíbrio para o grupo ácido acético.
$K_1^*$	Constante de equilíbrio para o grupo ácido acético, corrigida pelo potencial eletrostático médio.
$K_2$	Constante intrínseca de equilíbrio para o grupo amina.
$K_2^*$	Constante de equilíbrio para o grupo amina, corrigida pelo potencial eletrostático médio.
$K_3$	Constante intrínseca de equilíbrio para o grupo ionizável da cadeia lateral.
$K_3^*$	Constante de equilíbrio para o grupo ionizável da cadeia lateral, corrigida pelo potencial eletrostático médio.
$K_m$	Função de Bessel modificada de segunda espécie.
Leu	Leucina.
Lys	Lisina.
Met	Metionina.
$n_i$	Densidade numérica de íons.
$n_{i0}$	Concentração iônica no bulk (em mM).
$P_m$	Função de Legendre de primeira espécie.
$pK_i$	$-\log_{10} K_i$
$pK_i^*$	$-\log_{10} K_i^*$
Phe	Fenilalanina.
Pro	Prolina.
$Q_1$	Valor médio da carga elétrica média do grupo ácido acético.
$Q_2$	Valor médio da carga elétrica média do grupo amina.
$Q_3$	Valor médio da carga elétrica média do grupo ionizável da cadeia lateral.
$Q_m$	Função de Legendre de segunda espécie.
$R_{esf}$	Raio da esfera que completamente engloba o par de grupos ácido e base.
$R_1$	Distância entre os grupos ácido acético e amina ligados ao $C_\alpha$ .
$R_2$	Distância entre o grupo ácido acético e o grupo ionizável da cadeia lateral.
$R_3$	Distância entre o grupo amina e o grupo ionizável da cadeia lateral.
$r_a$	Raio do grupo ácido acético ligado ao $C_\alpha$ .
$r_b$	Raio do grupo amina ligado ao $C_\alpha$ .
$r_c$	Raio do grupo ionizável da cadeia lateral.
Ser	Serina.
T	Temperatura absoluta (°K).

Thr	Treonina.
Trp	Triptofano.
Tyr	Tirosina.
Val	Valina.
$z_i$	Valência dos íons da espécie $i$ .
$\alpha_1$	Grau de ionização do grupo ácido acético ligado ao $C_\alpha$ .
$\alpha_2$	Grau de ionização do grupo amina ligado ao $C_\alpha$ .
$\alpha_3$	Grau de ionização do grupo ionizável da cadeia lateral.
$\delta$	Parâmetro relacionado à contribuição iônica para o valor da constante dielétrica.
$\epsilon_1$	Constante dielétrica do meio 1.
$\epsilon_2$	Constante dielétrica do meio 2.
$\epsilon_{ef}$	Permissividade Efetiva.
$\theta$	Ângulo polar.
$\lambda$	Comprimento de Debye.
$\xi'$	Parâmetro adimensional resultante da divisão da coordenada $r$ por $\lambda$ .
$\rho$	Densidade de cargas.
$\rho_{FCP}$	Fator de Correlação de Pearson.
$\varphi_{(r)}$	Potencial reduzido médio.
$\varphi_{1(r)}$	Potencial reduzido médio calculado nas vizinhanças do grupo ácido.
$\varphi_{2(r)}$	Potencial reduzido médio calculado nas vizinhanças do grupo básico.
$\varphi_{3(r)}$	Potencial reduzido médio calculado nas vizinhanças do grupo ionizável da cadeia lateral.
$\Psi_{(r)}$	Potencial eletrostático médio.
$\Psi_{1(r)}$	Potencial eletrostático médio nas vizinhanças do grupo ácido ligado ao $C_\alpha$ .
$\Psi_{2(r)}$	Potencial eletrostático médio nas vizinhanças do grupo básico ligado ao $C_\alpha$ .
$\Psi_{3(r)}$	Potencial eletrostático médio do grupo ionizável da cadeia lateral.

# Sumário:

Capítulo 1: Introdução	18
1.1: Bioquímica de Aminoácidos e Peptídeos	20
1.2 Características físico-químicas	24
1.3 Hipótese Adotada	27
Capítulo 2: Modelo e Fundamentação Teórica	29
2.1 Modelo para o Cálculo da Constante Dielétrica Efetiva	30
2.2 Modelagem dos Aminoácidos	31
2.3 Aproximação de Campo Médio e Teoria de Debye-Hückel	34
Capítulo 3: Cálculo do Potencial Eletrostático Médio	39
3.1 Solução da Equação de Poisson-Boltzmann	40
3.2 Determinação das Constantes de Integração $a, b, c$ e Cálculo do Grau de Ionização	46
Capítulo 4: Modelo para o Cálculo da Constante Dielétrica Efetiva dos Aminoácidos	51
Capítulo 5: Resultados e Discussões: Constante Dielétrica Efetiva	55
Capítulo 6: Resultados e Discussões: Grau de Ionização e Grandezas Correlatas	59
6.1: Grau de Ionização ( $\alpha$ ) e $pK^*$	59
6.2: Efeito da Concentração Iônica Sobre os Valores de $pK^*$ Calculados	63



6.3: Estudo da Carga Elétrica Média Total em Condições Próximas às Fisiológicas de pH e Concentração Iônica	66
Capítulo 7: Pentapeptídeos: Modelo e Resultados	72
7.1: Adaptação ao estudo de pentapeptídeos	72
7.2: Resultados: $pK^*$ de Gly e Ala pentapeptídeos	73
Capítulo 8: Aplicação aos resíduos carregados da proteína <i>Staphylococcal Nuclease</i>	76
8.1: Interação entre resíduos carregados	76
8.2: Estudo sobre a distância de <i>cut-off</i> : Resultados para resíduos de Lys e Glu	78
Capítulo 9: Conclusões	82
Referências Bibliográficas	84

# Capítulo 1

## Introdução

Grande parte das atividades celulares é realizada por proteínas, macromoléculas compostas por subunidades chamadas aminoácidos. As distintas características destas subunidades, tais como hidrofobicidade e polaridade, são as responsáveis pelas propriedades conformacionais e estruturais que as referidas macromoléculas adquirem<sup>1</sup>, sendo um exemplo disto a formação de *motifs* de glicosilação em anticorpos recombinantes<sup>2</sup>.

Por si só, os aminoácidos podem desempenhar importantes funções. A glicina, por exemplo, mesmo sendo o aminoácido de estrutura mais simples, exerce a importante função de neurotransmissor inibitório<sup>3</sup>. Tocante à polaridade das subunidades, um estudo realizado por Trevino<sup>4</sup> mostra que ambos, glutamato e aspartato, contribuem mais para a solubilidade da *RNase* (Ribonuclease) do que os outros aminoácidos hidrofílicos.

Devido às dimensões e massas envolvidas em sistemas biológicos, a interação eletrostática é importantíssima<sup>5</sup> logo, a presença das cargas elétricas é um dos principais fatores que ditam a dinâmica de vários processos fisiológicos, dentro dos quais podem ser citados a transferência de elétrons entre peptídeos<sup>6</sup> e a adsorção destes em membranas<sup>7</sup>. A elevada complexidade tanto dos processos quanto das estruturas moleculares faz com que a determinação destas cargas em proteínas seja muito difícil. No entanto, como os aminoácidos são monômeros de estrutura relativamente mais simples e bem conhecida, há a possibilidade de se desenvolverem modelos para estes e, havendo sucesso, estendê-los para proteínas.

Um parâmetro crucial para o estudo de cargas elétricas médias, entre outras grandezas físico-químicas, é a determinação do valor de pK de cada grupo ionizável dos aminoácidos e, para tanto, abordagens teóricas e experimentais tem sido amplamente desenvolvidas. Včeláková<sup>8</sup> lançou mão da técnica de eletroforese capilar para mensurar a mobilidade efetiva da forma catiônica de aminoácidos e aplicou a aproximação de

McInnes e a teoria de Onsager para a determinação das constantes termodinâmicas de ionização  $K$  (com  $K$  sendo igual a  $10^{-pK}$ ) para cada um deles. Utilizando a técnica anterior juntamente com *UV* e *Coupled Contactless Conductivity Detection (C<sup>4</sup>D)*, Henchó<sup>9</sup> determinou os valores de  $pK$  para todos os grupos ionizáveis dos 20 aminoácidos de ocorrência natural (proteínogênicos).

Alguns programas computacionais como o SPARC<sup>10,11</sup> e o PKAS<sup>12</sup> conseguem, pelo emprego de relações de energia livre e atividade química, obter o  $pK$  em moléculas polipróticas. No entanto, eles se caracterizam pela intrincada especificação, por parte do usuário, das condições (tais como temperatura e concentração iônica) nas quais a molécula analisada se encontra.

Para a obtenção dos valores de  $pK$  dos resíduos de Lys, Arg, Asp e Glu internos à *Staphylococcal Nuclease*, Wallace<sup>13</sup> exemplifica a aplicabilidade de Dinâmica Molecular enquanto que Gunner<sup>14</sup> utiliza o Método de Monte Carlo em conjunto com a solução da equação de Poisson-Boltzmann obtida pelo método de Diferenças-Finitas.

Ainda na linha computacional, Antosiewicz<sup>15</sup> calcula o  $pK$  dos resíduos de aminoácidos em proteínas solucionando a equação de Poisson-Boltzmann por meio do método de Diferenças-Finitas<sup>16</sup>. Nota-se que os resultados apresentados foram, segundo o próprio autor, satisfatórios quando assumido que, nas regiões analisadas, a *constante dielétrica efetiva*<sup>17,18</sup> possuía valores acima de 20. Pela conjunta aplicação desta *constante efetiva* e do modelo de Born<sup>19</sup>, Isom calculou os valores de  $pK$  de resíduos de Lys<sup>20</sup> e de Glu<sup>21</sup> em regiões do *core* da *Staphylococcal Nuclease* (SNase). Devido ao papel desempenhado por esta constante, a sua definição é de extrema importância e, embora seu estudo remonte<sup>22-24</sup> aos trabalhos de Debye, Sack e Lorentz, ainda hoje modelos continuam a ser elaborados. No trabalho aqui apresentado, desenvolve-se, de maneira detalhada no Capítulo 4, uma abordagem que permite obter a constante dielétrica efetiva em termos da posição na qual se calcula o potencial eletrostático.

Quanto aos trabalhos teórico-analíticos para o cálculo de constantes de ionização, Agostinho Neto<sup>25</sup> empregou a aproximação de campo médio juntamente com a equação de Poisson-Boltzmann para obter o grau de ionização do ácido galacturônico em função da sua concentração.

Seguindo esse mesmo tratamento, construiu-se em um projeto de iniciação científica<sup>26</sup> um modelo que, utilizando a forma linearizada da supracitada equação, parâmetros espaciais e valores de  $pK$  de moléculas estruturalmente mais simples, possibilitava a determinação dos valores de  $pK$  e das cargas elétricas médias dos

aminoácidos apolares, i.e., aminoácidos com apenas dois grupos ionizáveis, que são: glicina, alanina, leucina, isoleucina e valina. Devido ao fato de os valores calculados anteriormente<sup>26</sup> concordarem com os da literatura<sup>8,9,27</sup>, o trabalho aqui descrito é a extensão deste modelo para os aminoácidos com três grupos ionizáveis: Asp, Glu, Arg, Lys, His, Cys e Tyr. Tal extensão, além de completar o modelo desenvolvido, possibilita estimar a carga elétrica média, o grau de ionização e os valores de pK em função da concentração iônica e do pH do meio, sendo que, por tal modelo, há a direta especificação destes parâmetros por parte do usuário. Objetiva-se, também, analisar os resultados obtidos em condições próximas às condições fisiológicas<sup>28</sup> e testar a aplicabilidade do modelo a peptídeos e proteínas.

## 1.1 Bioquímica de Aminoácidos e Peptídeos

Todas as proteínas, desde as pertencentes às linhagens mais antigas de bactérias até as presentes nos seres humanos, são construídas a partir do mesmo conjunto de 20 aminoácidos ligados em seqüências lineares. São os componentes e a ordem de tais seqüências os responsáveis pela variedade de proteínas observadas na natureza.

Apesar de serem constituídos basicamente por carbono (elemento de baixíssima abundância no meio interestelar), estudos conduzidos por Pizzarello<sup>29</sup> e Sephton<sup>30</sup> encontraram, respectivamente, vestígios dos aminoácidos isovalina e glicina no meteorito *Murchison*. Tais resultados colaboram para a hipótese de que a formação de moléculas orgânicas em nosso planeta só foi possível pela vinda não só de água, mas também de moléculas orgânicas em cometas e asteróides.

Aminoácidos são moléculas que contêm simultaneamente grupos funcionais amina e ácido carboxílico ligados ao mesmo átomo de carbono (denominado carbono- $\alpha$ ,  $C_\alpha$ ). Excetuando-se a glicina, nos demais 19 aminoácidos a este carbono se ligam quatro diferentes grupos substituintes: os grupos funcionais supracitados, um hidrogênio e uma cadeia lateral cuja estrutura será a responsável pela diferenciação entre os aminoácidos.

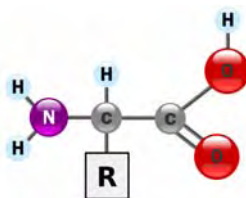


Figura 1.1: Representação de um aminoácido. A cadeia lateral é designada por R.

Dependendo da estrutura da cadeia lateral, o aminoácido poderá ser classificado pelas seguintes classes<sup>1,31</sup>:

✚ *Aminoácidos de cadeia lateral não-polar e alifática*: possuem caráter hidrofóbico e apolar. Leucina (Leu), isoleucina (Ile), alanina (Ala) e valina (Val) detêm cadeias laterais volumosas, com seus átomos arranjos de tal forma que contribuem para a estabilização da estrutura protéica. A glicina (Gly) é o aminoácido mais simples e a sua cadeia lateral (constituída por um único átomo de hidrogênio) não contribui para interações hidrofóbicas. A prolina (Pro), por apresentar um grupo imino ao invés da amina, leva à redução da flexibilidade estrutural das cadeias polipeptídicas da qual faz parte. A composição desses aminoácidos é mostrada na figura 1.2.

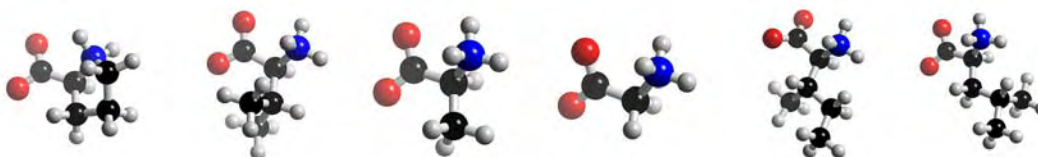


Figura 1.2: Pro, Val, Ala, Gly, Ile e Leu, respectivamente da esquerda para a direita. Átomos de carbono em preto, oxigênio em vermelho, nitrogênio em azul e hidrogênio em branco. Correspondência cor-elemento mantida em todas as figuras desta seção. Fonte: <http://www.3dchem.com/search>

✚ *Aminoácidos com grupo hidroxila na cadeia lateral*: serina (Ser) e a treonina (Thr) são caracterizadas por, em suas cadeias laterais pequenas e alifáticas, possuírem um grupo hidroxila o qual não é mais reativo do que o etanol. Veja a figura 1.3.

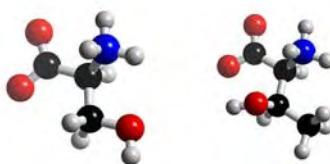


Figura 1.3: Ser e Thr, respectivamente da esquerda para a direita. Fonte: <http://www.3dchem.com/search>

✚ *Aminoácidos de cadeia lateral aromática*: os aminoácidos desta classe, fenilalanina (Phe), tirosina (Tyr) e triptofano (Trp), são relativamente hidrofóbicos. O grupo hidroxila da tirosina e o anel indol do triptofano atribuem a estes aminoácidos uma característica significativamente mais polar do que a fenilalanina. Vide figura 1.4.

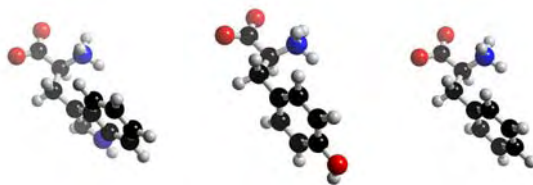


Figura 1.4: Trp, Tyr e Phe, respectivamente da esquerda para a direita. Fonte: <http://www.3dchem.com/search>

✚ *Aminoácidos de cadeia lateral carregada positivamente:* a presença de carga confere a estes aminoácidos características de basicidade e hidrofilicidade. Em pH 7, os aminoácidos que possuem carga (total) positiva são a lisina (Lys), cuja cadeia lateral possui um grupo amina terminal; a arginina (Arg), que tem um grupo guanidina carregado positivamente e a histidina (His), constituída por um imidazol que confere à cadeia lateral deste aminoácido um valor de pK próximo a 6,0.

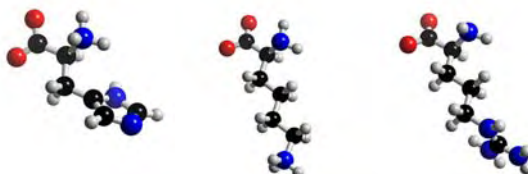


Figura 1.5: His, Lys e Arg, respectivamente da esquerda para a direita. Fonte: <http://www.3dchem.com/search>

✚ *Aminoácidos de cadeia lateral carregada negativamente:* em pH 7, os dois únicos aminoácidos cujas cadeias laterais se encontram carregadas negativamente são o aspartato (Asp) e o glutamato (Glu), cada um deles possuindo um segundo grupo carboxila na cadeia e representados na figura 1.6.

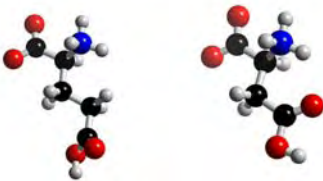


Figura 1.6: Glu e Asp, respectivamente da esquerda para a direita. Fonte: <http://www.3dchem.com/search>

✚ *Aminoácidos de cadeia lateral amidada:* os aminoácidos pertencentes a este grupo, asparagina (Asn) e glutamina (Gln), apesar de serem as formas amidadas de Asp e Glu, respectivamente, ocorrem naturalmente e são diretamente incorporados às

proteínas. Suas cadeias laterais não se ionizam e não são quimicamente muito reativas. Veja a figura 1.7.

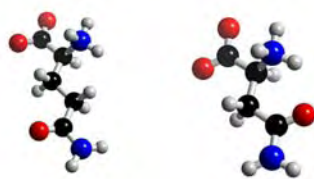


Figura 1.7: Gln e Asn, respectivamente da esquerda para a direita. Fonte: <http://www.3dchem.com/search>

✚ **Aminoácidos sulfurados:** a metionina (Met) detém uma longa e apolar cadeia lateral na qual se encontra um átomo de enxofre. Este, por sua vez, possui um caráter nucleofílico mas, ao contrário dos outros nucleófilos presente em proteínas, ele não pode ser ionizado. Por outro lado, o grupo tiol presente na cadeia lateral da cisteína (Cys) é, entre todas as cadeias laterais dos aminoácidos, o mais reativo. Em algumas proteínas, ligações dissulfeto ocorrem entre resíduos de Cys e este par de resíduos assim conectados recebe o nome de *resíduo de cistina*. Estes dois aminoácidos estão representados na figura 1.8.

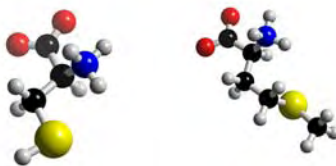


Figura 1.8: Cys e Met, respectivamente da esquerda para a direita. Átomo de Enxofre em amarelo. Fonte: <http://www.3dchem.com/search>

Como mencionado anteriormente, neste trabalho foram estudados os aminoácidos com três grupos ionizáveis: Asp, Glu, Cys, Tyr, Arg, Lys e His.

Um dos objetivos do modelo desenvolvido consistiu em aplicá-lo a situações próximas às condições fisiológicas e, sempre que possível, buscou-se confrontar os resultados calculados com aqueles presentes na literatura e obtidos experimentalmente. Desta forma, no capítulo 6 serão apresentados alguns resultados para a temperatura de 309,65 K (36,5°C). Para a análise dos efeitos da concentração iônica, tomaram-se por base os valores apresentados por Alberts<sup>28</sup> para os meios intra e extracelulares, mostrados na tabela 1.1.

Tabela 1.1: Valores das concentrações iônicas para os meios intra e extracelular. Concentrações de outras espécies aniônicas foram omitidas. Adaptada de Alberts<sup>28</sup>

Íons	Concentração intracelular (mM)	Concentração extracelular (mM)
Na <sup>+</sup>	5-15	145
K <sup>+</sup>	140	5
Mg <sup>2+</sup>	0,5	1-2
Ca <sup>2+</sup>	10 <sup>-4</sup>	1-2
H <sup>+</sup>	7.10 <sup>-5</sup>	4.10 <sup>-5</sup>
Cl <sup>-</sup>	5-15	110

No capítulo 6 são apresentados gráficos da variação de pK dos grupos ionizáveis com o acréscimo da concentração iônica bem como gráficos e tabelas de valores de carga elétrica média dos aminoácidos em condições próximas as fisiológicas de pH e temperatura.

## 1.2 Características Físico-Químicas:

Um dos tipos de reação química que tem extrema importância para as células é aquela que ocorre quando uma molécula, possuidora de alguma ligação covalente altamente polar entre um hidrogênio e um segundo átomo, se dissolve em água. Quando isto ocorre, o átomo de hidrogênio da molécula polar pode doar, ao átomo a ele associado, um elétron restando assim um núcleo de hidrogênio (H<sup>+</sup>), i.e., um próton. Como o oxigênio é um elemento de elevada eletronegatividade, este próton é por ele atraído formando, assim, um íon hidrônio (H<sub>3</sub>O<sup>+</sup>). A reação inversa ocorre muito prontamente e, portanto, em um estado de equilíbrio, muitos prótons estão continuamente transitando entre moléculas polares e a água.

Ácidos e bases, pela definição de Svante Arrhenius<sup>32,33</sup>, são, respectivamente, substâncias capazes de doar prótons e íons hidroxila (OH<sup>-</sup>). Todavia, tal definição é muito limitada, pois, por exemplo, não considera o fato de que o grupo NH<sub>3</sub>, mesmo não possuindo a hidroxila, apresenta propriedades básicas. Uma definição mais geral foi formulada em 1923 por Johannes Brønsted e Thomas Lowry<sup>33</sup>. Segundo eles, um *ácido* é uma substância que pode *doar prótons* e uma *base* é uma substância capaz de *aceitar prótons*. Em símbolos tem-se:





Na expressão (1.1), um ácido de Brønsted (HA) reage com uma base de Brønsted (H<sub>2</sub>O) formando uma base conjugada (A<sup>-</sup>) e um ácido conjugado (íon hidrônio, H<sub>3</sub>O<sup>+</sup>). Como exemplos têm-se o íon acetato (base conjugada do ácido acético) e o íon amônio que é o ácido conjugado da amônia.

Uma vez que os prótons de um íon hidrônio podem facilmente passar para o interior das células<sup>1,28</sup>, a concentração de tal partícula deve ser rigidamente controlada de tal forma que a acidez do meio celular seja mantida constante. Com este intuito, as moléculas com tendência a doar prótons o farão se a concentração de H<sub>3</sub>O<sup>+</sup> for baixa. Por outro lado, quando a concentração deste íon for alta, as moléculas com tendência a absorver prótons o farão com maior facilidade. Pelo fato de o íon HO<sup>-</sup> combinar-se com um íon hidrônio para formar duas moléculas de água, um aumento na concentração do primeiro força uma diminuição na concentração do segundo e vice-versa.

O valor da concentração de prótons pode ser expresso de uma maneira experimentalmente mais convencional usando uma escala logarítmica. Assim, a concentração de prótons é escrita como:

$$pH = -\log_{10} [H^+] \quad (1.2)$$

Em soluções aquosas de pH neutro (pH 7), a maioria dos aminoácidos estará em uma forma ionizada, situação em que o grupo amina se encontra protonado (—CH<sub>3</sub>NH<sub>3</sub><sup>+</sup>) e o ácido carboxílico desprotonado a carboxilato (—COO<sup>-</sup>), denominando-se esta forma de **zwitterionica** (do alemão *zwitter*, que significa hermafrodita). Devido à estrutura da cadeia lateral da histidina, que possui o grupo imidazol, ela é o único aminoácido capaz de exercer, significativamente, função tamponante próximo ao pH 7.

O processo de ionização é caracterizado por uma constante de equilíbrio que, para uma reação ácido-base, é conhecida como constante de ionização (K). Com o uso da lei de ação das massas<sup>32</sup> definem-se, respectivamente, para os grupos ácido e básico de um aminoácido:

$$\frac{[H_3CCOO^-][H^+]}{[H_3CCOOH]} = K_1 \quad (1.3)$$

$$\frac{[CH_3NH_2][H^+]}{[CH_3NH_3^+]} = K_2 \quad (1.4)$$

Em aminoácidos de cadeia lateral polar carregada, além dos dois grupos mencionados acima, há mais um grupo ionizável e este se localiza em sua cadeia lateral. Portanto, para o estudo desta classe de aminoácidos se torna necessária a definição de mais uma constante de ionização que será denotada por  $K_3$  e pode ser exemplificada pelo íon butilamínio da cadeia lateral da lisina onde tem-se:

$$\frac{[H(CH_2)_4NH_2][H^+]}{[H(CH_2)_4NH_3^+]} = K_3 \quad (1.5)$$

Ao longo deste trabalho é adotada a seguinte notação para os sub-índices: 1 para as grandezas referentes ao ácido etanóico (grupo 1) ligado ao carbono  $\alpha$ ; 2 para a base (grupo 2) ligada ao carbono  $\alpha$ ; e 3 para o grupo ionizável presente na cadeia lateral do aminoácido (grupo 3).

O grau de ionização ( $\alpha$ ) é a relação entre o número (concentração) de grupos ionizados e o número total de grupos. Seguindo a notação acima e tomando novamente a cadeia lateral da lisina como exemplo obtém-se:

$$\frac{[H_3CCOO^-]}{[H_3CCOO^-] + [H_3CCOOH]} = \alpha_1 \rightarrow \text{ácido} \quad (1.6)$$

$$\frac{[CH_3NH_2]}{[CH_3NH_3^+] + [CH_3NH_2]} = \alpha_2 \rightarrow \text{base (amina)} \quad (1.7)$$

$$\frac{[H(CH_2)_4NH_2]}{[H(CH_2)_4NH_3^+] + [H(CH_2)_4NH_2]} = \alpha_3 \rightarrow \text{cadeia lateral.} \quad (1.8)$$

A partir das seis últimas equações e com o uso da definição<sup>33</sup> de pH e pK consegue-se correlacionar o grau de ionização de cada grupo ionizável com a sua respectiva constante intrínseca<sup>34</sup> de ionização (pK):

$$\alpha_1 = \frac{1}{1 + \frac{[H^+]}{K_1(T)}} = \frac{1}{1 + 10^{(-pH + pK_1)}} \quad (1.9)$$

$$\alpha_2 = \frac{1}{1 + \frac{[H^+]}{K_2(T)}} = \frac{1}{1 + 10^{(-pH + pK_2)}} \quad (1.10)$$

$$\alpha_3 = \frac{1}{1 + \frac{[H^+]}{K_3(T)}} = \frac{1}{1 + 10^{(-pH + pK_3)}} \quad (1.11)$$

## 1.3 Hipótese Adotada

Como mencionado anteriormente, há casos em que os aminoácidos se encontram na forma zwitterionica, i.e., uma molécula globalmente neutra em termos de carga elétrica, mas que possui cargas locais devido à presença de grupos ionizados. Por consequência disto, há um potencial eletrostático resultante não nulo.

Apesar de os grupos ácidos (grupos 1) dos aminoácidos serem estruturalmente semelhantes ao ácido etanóico, o valor de  $pK_1$  deste é relativamente maior do que os correspondentes valores dos primeiros. Comparando-se, também, os valores de  $pK_2$  definidos para os grupos básicos (grupos 2) dos aminoácidos com os de um grupo semelhante, tal como a metilamina, percebe-se que o  $pK_2$  desta também é maior do que os valores de  $pK_2$  dos aminoácidos. Para explicitar o que foi dito, veja a figura 1.9 e tabela 1.2:

	Glu	Asp	Tyr	Cys	His	Lys	Arg
$pK_1$ - ácido	2,28	2,02	2,33	2,05	1,54	2,21	1,95
$pK_2$ - base	9,87	10,03	9,15	10,20	9,18	9,19	9,00
$pK_3$ - cadeia lateral	4,17	3,66	10,57	8,44	6,07	10,75	12,50

Figura 1.9: Aminoácidos estudados e respectivos valores de  $pK$ . O valor de  $pK_3$  da Arginina é fornecido por Willians<sup>35</sup> enquanto todos os outros são reportados por Henchoz<sup>9</sup>.

Tabela 1.2:  $pK$  e estrutura dos grupos amina e ácido carboxílico.

Grupo	estrutura	$pK$
Ácido Etanóico	$\begin{array}{c} \text{H} \\   \\ \text{H}-\text{C}-\text{COOH} \\   \\ \text{H} \end{array}$	$pK = 4,76$
Metilamina	$\begin{array}{c} \text{H} \quad \text{H} \\   \quad   \\ \text{H}-\text{N}-\text{C}-\text{H} \\   \\ \text{H} \end{array}$	$pK = 10,65$

A hipótese adotada neste trabalho é que a diferença entre os valores de pK entre os grupos ionizáveis constituintes dos aminoácidos e dos mesmos quando “não combinados” se deve à interação eletrostática adicional devida aos grupos vizinhos e aos íons em solução. Como citado anteriormente e seguindo esta hipótese, a presença (ou ausência) de cargas em um dos grupos gera um potencial que irá influenciar na ionização dos outros, alterando assim os respectivos valores de pK.

Para realizar o teste desta hipótese, compararam-se as grandezas calculadas pelo modelo teórico com aquelas oriundas de trabalhos experimentais. Os resultados encontram-se no Capítulo 6.

## Capítulo 2

### Modelo e Fundamentação Teórica

Considerando que os aminoácidos estudados encontram-se em solução aquosa e cada um destes é tratado como um “*aminoácido de ensemble*” representando, assim, os valores médios das grandezas referentes às moléculas iguais a ele.

Os aminoácidos são principalmente constituídos por carbono, elemento responsável por conferir a eles características apolares. Como tais se encontram em solução com um líquido polar (água), para estudos que visam descrever as suas propriedades em termos da interação eletrostática<sup>36</sup>, a *constante dielétrica* é um parâmetro de mister importância.

Devido à dificuldade experimental de se obter a constante dielétrica (CD) em regiões internas das proteínas, alguns trabalhos empregam a fenomenologia em conjunto com o modelo de Born para estimá-la<sup>20,21</sup>. Nestas regiões, tanto trabalhos experimentais quanto analítico-computacionais apresentam valores entre 6 e 38<sup>37</sup>. No entanto, na interface proteína-solvente, estudos experimentais fornecem valores que variam de 40<sup>38</sup> a 120<sup>19,39</sup>. É digno de nota que todos esses resultados se referem à Constante Dielétrica Efetiva ( $CD_{ef}$ ). Segundo Isom<sup>20</sup>, tal grandeza é um parâmetro em cujo valor estão inclusas as contribuições não explicitadas<sup>15,18</sup> pelo modelo empregado para o cálculo do pK, por exemplo<sup>20,21</sup>.

A variedade de valores apresentados para  $CD_{ef}$  em diferentes regiões sugere que ela seja uma função dependente da posição onde é calculada e que possua um caráter sigmoidal<sup>40-42</sup>. Ao longo deste trabalho é apresentado um modelo teórico que resolve um problema de condições de contorno para um sistema constituído por dois meios dielétricos lineares distintos. Este modelo, além de fornecer resultados em termos da posição onde se calcula a  $CD_{ef}$ , justifica alguns dos valores supracitados.

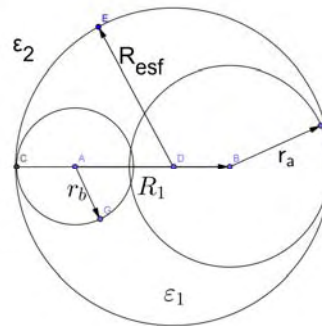
## 2.1 Modelo para o Cálculo da Constante Dielétrica Efetiva

Inicialmente considera-se que, nos aminoácidos, cada *par* de grupos *ácido* e *base* está envolto por uma superfície esférica cujo interior (*meio 1*) é caracterizado pela constante dielétrica  $\epsilon_1$ . Devido à constituição molecular deste meio, admiti-se que ele é semelhante a um hidrocarboneto. Para vários hidrocarbonetos, Richards<sup>43</sup> apresenta valores de constante dielétrica bem próximos de 2,0. Com base nisto e para efeitos de cálculo, o valor de  $\epsilon_1$  utilizado é 2,0.

O meio exterior à(s) superfície(s) supracitada(s) é constituído por água e íons em solução. A constante dielétrica deste meio, denotado por *meio 2*, é representada pelo parâmetro  $\epsilon_2$ , cujo valor se obtém por uma interpolação de termos relacionados à temperatura centígrada<sup>44</sup>(T, em °C) e à concentração iônica ( $n_{i0}$ )<sup>45</sup>:

$$\begin{aligned} \epsilon_2 = & 87,7404 - 0,40014T + 0,940425 \cdot 10^{-3}T^2 - \\ & - 0,140165 \cdot 10^{-5}T^3 - 0,929212 \cdot 10^{-10}T^4 + 2\delta n_{i0} \end{aligned} \quad (2.1)$$

O parâmetro  $\delta$  está envolvido com a contribuição dos íons para o valor da constante dielétrica e é expresso em litros por mol. A figura 2.1 é uma representação do sistema e do modelo:



*Figura 2.1: Representação da modelagem dos aminoácidos:  $\epsilon_1$  é a constante dielétrica do meio 1 (com valor igual a 2,0) e  $\epsilon_2$  é a constante dielétrica da água, cujo valor é dado pela eq. (2.1). Os outros parâmetros serão explicados logo a seguir.*

Para a visualização de como se dá variação de  $\epsilon_2$  com a temperatura, elaborou-se a figura 2.2 na qual são apresentados os valores calculados por (2.1) para dois distintos valores de concentração iônica.

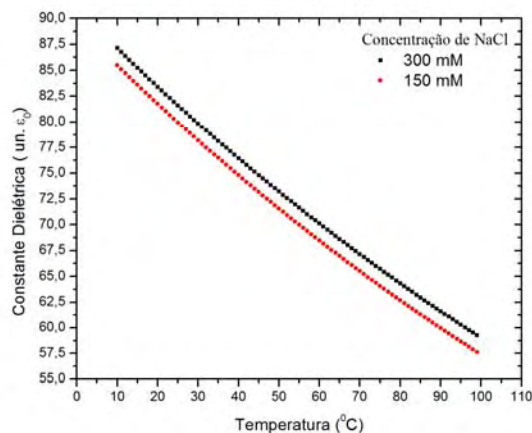


Figura 2.2: Variação da  $\epsilon_2$  com a temperatura para duas distintas concentrações de NaCl.

Os principais parâmetros que influenciam no valor da constante dielétrica efetiva são o raio ( $R_{\text{esf}}$ ) da superfície que engloba cada par de grupos e a posição do ponto onde ela é calculada. O desenvolvimento matemático do modelo se encontra no Capítulo 4.

## 2.2 Modelagem dos Aminoácidos

Os grupos ionizáveis são tratados como esferas que, quando carregados, possuem uma carga pontual central. A presença de um oxigênio, átomo de elevada eletronegatividade, no ácido carboxílico possibilita o efeito de ressonância da dupla ligação e o nitrogênio do grupo básico, que também é eletronegativo, pode atrair fortemente o elétron de um dos seus átomos de hidrogênio. Estes fatores fazem com que os “raios” de tais grupos sejam difíceis de serem determinados e, portanto, optou-se por trabalhar com os valores da distância entre o centro de cada grupo e o início da sua respectiva camada de hidratação<sup>46</sup>.

Para as grandezas referentes aos grupos, tais como grau de ionização e distâncias entre eles, é adotada ao longo de todo o trabalho a seguinte notação para os grupos e respectivos sub-índices:

- 1 para o grupo ácido (de raio  $r_a$ ) ligado ao carbono  $\alpha$ ;
- 2 para o grupo básico (de raio  $r_b$ ) ligado ao carbono  $\alpha$ ;
- 3 para o grupo ionizável (de raio  $r_c$ ) presente na cadeia lateral do aminoácido;

- ✚ R<sub>1</sub>: Distância entre o grupos 1 e 2;
- ✚ R<sub>2</sub>: Distância entre os grupos 1 e 3;
- ✚ R<sub>3</sub>: Distância entre os grupos 2 e 3.

Foram estudados os seguintes aminoácidos: ácido aspártico (Asp), ácido glutâmico (Glu), arginina (Arg), lisina (Lys), histidina (His), tirosina (Tyr) e cisteína (Cys). Como todos eles apresentam o mesmo número de grupos ionizáveis (três), não houve a necessidade de se adaptar o modelo para cada um deles. Devido a essa versatilidade, pode-se estendê-lo a outras moléculas que também possuam três grupos ionizáveis tais como certos oligopeptídeos (vide Capítulo 7).

Quando os grupos ionizáveis não se encontram combinados em um aminoácido eles possuem constantes de ionização denominadas de *constantes de ionização intrínsecas*, K<sub>i</sub>, com *i* dado pela notação acima. Baseando-se na semelhança estrutural que os grupos 1 possuem com o ácido acético (etanóico), em todos os aminoácidos eles foram tratados como tal. Pelo mesmo critério de semelhança, todos os grupos 2 foram tratados como metilamina, conforme presente na tabela 2.1.

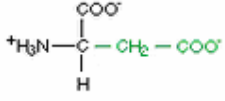
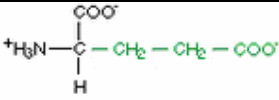
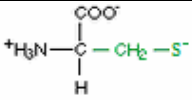
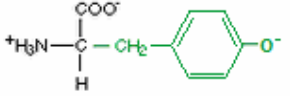
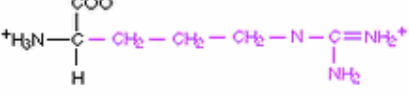
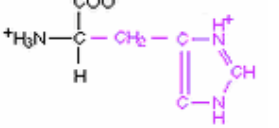
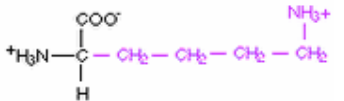
Tabela 2.1: Valores<sup>34</sup> de pK<sub>1</sub> e pK<sub>2</sub> intrínsecos para os grupos 1 e 2.

Grupo		pK
Ácido Etanóico	$\begin{array}{c} \text{H} \\   \\ \text{H}-\text{C}-\text{COOH} \\   \\ \text{H} \end{array}$	pK = 4,76
Metilamina	$\begin{array}{c} \text{H} \quad \text{H} \\   \quad   \\ \text{H}-\text{N}-\text{C}-\text{H} \\   \\ \text{H} \end{array}$	pK = 10,65

O que distingue cada um dos aminoácidos é o grupo presente em sua cadeia lateral e cada um destes também foi tratado como a molécula estruturalmente a ele mais semelhante. A tabela 2.2 a seguir detalha a molécula e o respectivo valor de pK intrínseco usado para cada aminoácido.



Tabela 2.2: Aminoácido na forma completamente ionizada e tratamento dado a suas cadeias laterais com os respectivos valores de pK intrínseco. Em verde, cadeias ácidas e, em rosa, cadeias básicas.

Aminoácidos	Cadeia lateral tratada como: e Respectivo pK (intrínseco):
Ácido Aspártico 	Ácido Propanóico $pK_3=4,87^{34,47}$
Ácido Glutâmico 	Ácido Butanóico $pK_3=4,83^{47}$
Cisteína 	Etil-mercaptano $pK_3=10,50^{35}$
Tirosina 	p-etano-fenol $pK_3=10,2^{35}$
Arginina 	Butil-guanidina $pK_3=12,65^{35}$
Histidina 	Metil-Imidazol $pK_3=6,95^{35}$
Lisina 	Pentilamina $pK_3=10,63^{34,47}$

Depois de uma série de trabalhos, Momany<sup>48</sup> apresenta o desenvolvimento de potenciais interatômicos empíricos para a obtenção dos parâmetros estruturais dos *resíduos* dos 20 aminoácidos de ocorrência natural na sua forma energeticamente mais favorável. Dentre os vários resultados apresentados, tanto na forma de figuras quanto em tabelas, há os valores dos ângulos de torção e comprimento de ligações entre os átomos constituintes dos aminoácidos estudados. Outras abordagens que possibilitam tal estimativa dimensional, embora que indiretamente, são fornecidas por Sakurai<sup>49,50</sup> e Meiler<sup>51</sup>, sendo que este último fornece os valores dos volumes das cadeias laterais.

No entanto, para uma melhor determinação tanto das distâncias entre os grupos quanto das dimensões destes nos aminoácidos estudados, utilizou-se o programa PyMOL<sup>52</sup> em conjunto com os arquivos oriundos do *Protein Data Bank* (PDB). Seguem

na tabela 2.3 os valores desta fonte provenientes e utilizados no modelo aqui desenvolvido:

*Tabela 2.3: Aminoácidos com os respectivos valores de raios dos grupos ionizáveis e distâncias entre eles. Os índices seguem a notação definida nesta mesma seção. Valores calculados pelo PyMOL.*

Aminoácidos	$r_a$ (Å) <sup>48</sup>	$r_b$ (Å) <sup>48</sup>	$r_c$ (Å) <sup>48,51</sup>	$R_1$ (Å) <sup>48-50</sup>	$R_2$ (Å) <sup>48-50</sup>	$R_3$ (Å) <sup>48-50</sup>
Ácido Aspártico	2,4	1,4	2,81	3,82	2,67	4,18
Ácido Glutâmico	2,4	1,4	5,08	2,70	5,88	5,38
Cisteína	2,4	1,4	2,76	2,71	4,83	4,05
Tirosina	2,4	1,4	6,53	3,75	8,44	6,51
Arginina	2,4	1,4	7,83	3,75	6,83	8,53
Histidina	2,4	1,4	2,91	3,75	3,99	3,91
Lisina	2,4	1,4	6,41	3,75	5,94	7,56

A hipótese que alicerça o modelo é que a diferença entre os valores de pK de um grupo ionizável isolado e deste como parte constituinte de um aminoácido se deve à interação eletrostática adicional exercida pelos grupos vizinhos e pelos íons em solução, sendo o potencial eletrostático médio quem representa todas essas contribuições.

O cálculo deste potencial nas superfícies de cada grupo ionizável é crucial para a obtenção das grandezas de interesse e, conseqüentemente, teste da hipótese. Tal determinação é possibilitada pela Aproximação de Campo Médio e pela teoria de Debye-Hückel<sup>53,54</sup>, que são de grande valia para o estudo de sistemas eletrolíticos. Alguns detalhes destas abordagens são apresentados logo a seguir.

## 2.3 Aproximação de Campo Médio e Teoria de Debye-Hückel

Trabalhar com um sistema contendo um grande número de partículas interagindo fortemente entre si graças à interação eletrostática requer um tratamento estatístico. Admitindo-se que o sistema possa ser considerado como constituído de subsistemas independentes, é possível empregar a aproximação de campo médio<sup>53</sup>. Tal aproximação permite que, em um sistema de muitos corpos interagentes, toda a interação que um corpo “sinta” devido aos outros seja representada por um campo

efetivo. Procedendo assim, todas as flutuações nas interações são desconsideradas, o que caracteriza esta aproximação como uma aproximação de ordem zero.

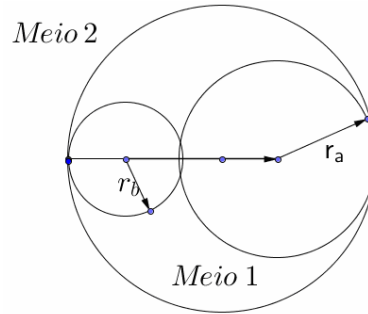


Figura 2.3: Esfera que engloba cada par de grupos ácido e base e delimita dois meios distintos. No meio 2 há água e íons em solução.

No meio exterior à esfera que engloba cada par de grupos ácido e base (meio 2), há  $N_i$  íons em uma solução aquosa de volume  $V$ . A figura 2.2 representa esse sistema. A densidade numérica de íons,  $n_i$ , é dada pela seguinte expressão:

$$n_i = \frac{N_i}{V} \quad (2.2)$$

Os íons são tratados como esferas rígidas de raio  $a$ . Sendo  $z_i$  e  $z_j$  as valências dos íons das espécies  $i$  e  $j$ , respectivamente, e levando em conta a “impenetrabilidade” destas esferas, a energia potencial “interiônica” pode ser expressa por<sup>53</sup>:

$$u_{ij}(r) = \left\{ \begin{array}{l} +\infty \rightarrow r < a \\ \frac{z_i z_j e^2}{\epsilon_2 r} \rightarrow r \geq a \end{array} \right\} \quad (2.3)$$

Na realidade, a presença de cargas faz com que o dielétrico não seja um meio contínuo quando  $r$  é pequeno e os íons não se comportam como esferas rígidas. Contudo, apesar da expressão anterior ser apenas uma aproximação, ela fornece resultados exatos para o caso em que a solução é muito diluída (ou seja,  $n_i \rightarrow 0$ ).

Em 1923, Peter Debye e Erich Hückel<sup>53</sup> formularam uma teoria para as soluções iônicas. Tal teoria tem como base a consideração de que a interação coulombiana, caracterizada por ser de longo alcance e de grande intensidade em diminutas distâncias, é a principal responsável por fazer com que o comportamento das soluções iônicas se afaste do ideal.

A solução iônica como um todo é eletricamente neutra. Todavia, devido ao fato de íons de cargas contrárias se atraírem mutuamente, ao redor de um íon há um excesso de íons de cargas opostas (chamados de contra-íons). Tomando-se a média temporal do arranjo iônico em solução, os contra-íons se encontram com maior probabilidade nas proximidades dos íons. Essa nuvem de íons é denominada *atmosfera iônica* e apresenta uma simetria esférica em torno do íon central.

Apesar de o número de contra-íons ser maior do que o número de íons com mesma carga que a do íon central, a carga líquida da atmosfera iônica corresponde à carga do próprio íon central. A energia e, conseqüentemente, o potencial químico de qualquer que seja o íon central acaba se reduzindo devido à interação deste íon com a atmosfera iônica. Em um sistema em que não há interação entre as partículas, o potencial químico ( $\mu_i$ ) é:

$$\mu_i = \mu_i^0(\vec{r}) + kT \ln(n_{i0}) \quad (2.4)$$

No entanto, como há interação entre as partículas que constituem o sistema, a densidade iônica,  $n_i$ , é dependente da posição (dada pelo vetor posição  $\vec{r}$ ) do íon  $i$  e, por causa disto, o potencial químico pode ser reescrito já se considerando esta correção de primeira ordem<sup>55</sup>:

$$\mu_i = \mu_i^0(\vec{r}) + kT \ln(n_i(\vec{r})) + kT \ln(\gamma_i(\vec{r})) \quad (2.5)$$

Em condições de equilíbrio termodinâmico,  $\mu_i$  deve assumir o mesmo valor em toda e qualquer região da solução. O termo  $\mu_i^0(\vec{r})$  está relacionado às propriedades de um íon poder ser hidratado. Se as moléculas de solvente tiverem um arranjo na camada de hidratação idêntico ao observado em qualquer outro lugar da solução,  $\mu_i^0$  será independente de  $\vec{r}$  possuindo, assim, um valor constante.

O termo  $kT \ln(\gamma_i(\vec{r}))$ , onde  $\gamma_i$  é o coeficiente de atividade da espécie  $i$ , representa o trabalho isotérmico reversível requerido para se adicionar ao sistema mais um íon da espécie  $i$  (que estava anteriormente em  $\vec{r}_1$ ) na posição dada por  $\vec{r}_2$ . Denota-se este trabalho por  $W$ .

Se a estrutura do íon hidratado não apresentar variações<sup>55</sup>,  $W$  corresponde à diferença de energia potencial eletrostática decorrente da transferência da molécula de  $\vec{r}_1$  para  $\vec{r}_2$ . Sendo  $\Psi(r)$  a função *potencial eletrostático médio*, tem-se:

$$\vec{E}(r) = -\nabla\Psi(r) \quad (2.6)$$

Onde  $\vec{E}(r)$  representa o *campo elétrico médio*. Considerando-se que o produto escalar entre o momento de dipolo do sistema e  $\vec{E}(r)$  seja muito menor que  $z_i e \Psi(r)$ , a energia potencial eletrostática do íon hidratado vale  $z_i e \Psi(r)$ . Tal resultado é o mesmo que o obtido para uma carga pontual  $z_i e$  na posição  $r$ . No tocante aos íons monovalentes que, por sua vez, apresentam certa simetria esférica, pode-se aproxima-los a cargas pontuais.

Com base nessa discussão, a expressão (2.5) ganha nova forma:

$$\mu_i = \mu_i^0(\vec{r}) + kT \ln(n_i(\vec{r})) + z_i e \Psi(\vec{r}) \quad (2.7)$$

Igualando as expressões (2.7) e (2.4) tem-se que:

$$kT \ln(n_i(\vec{r})) + z_i e \Psi(\vec{r}) = kT \ln(n_{i0}) \quad (2.8)$$

O símbolo  $n_{i0}$  representa a densidade numérica de íons da espécie  $i$  nas regiões onde o potencial eletrostático médio é nulo. A esta região dá-se o nome de *bulk*. Aplicando-se a função exponencial em ambos os termos da equação (2.8), chega-se a uma expressão que informa como a densidade numérica de íons é influenciada pela energia potencial eletrostática média:

$$n_i(\vec{r}) = n_{i0} \exp\left(\frac{-z_i e \Psi(r)}{kT}\right) \quad (2.9)$$

Com este resultado, pode-se expressar a densidade de cargas ( $\rho$ ) por:

$$\rho = \sum z_i e n_i(r) = \sum z_i e n_{i0} \exp\left(\frac{-z_i e \Psi(r)}{kT}\right) \quad (2.10)$$

A Lei de Gauss afirma que:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_1} \quad (2.11)$$

Empregando-se a identidade dada pela equação (2.6) tem-se que:

$$\nabla^2 \Psi = -\frac{\rho}{\epsilon_1} \quad (2.12)$$

Desta forma, pode-se relacionar o potencial eletrostático médio com a densidade de cargas dada pela equação (2.10):

$$\nabla^2 \Psi = -\frac{1}{\epsilon} \sum_i z_i e n_{i0} \exp(-z_i e \Psi / kT) \quad (2.13)$$

Esta equação é conhecida com *Equação de Poisson-Boltzmann*. A consideração de que  $z_i e \Psi \ll kT$  constitui a segunda hipótese adotada na teoria de Debye-Hückel.

# Capítulo 3

## Cálculo do Potencial Eletrostático Médio.

Considerando a hipótese de que o valor da constante de ionização de cada um dos grupos ionizáveis de um aminoácido seja alterado pela interação eletrostática adicional devida à presença dos outros grupos e dos íons em solução, podem-se escrever estas “novas” constantes de ionização (que serão denotadas por  $K_1^*$ ,  $K_2^*$  e  $K_3^*$ ) da seguinte forma:

$$K_1^* = K_1 \exp\left(\frac{-z_1 e \Psi_1}{kT}\right) \quad (3.1)$$

$$K_2^* = K_2 \exp\left(\frac{-z_2 e \Psi_2}{kT}\right) \quad (3.2)$$

$$K_3^* = K_3 \exp\left(\frac{-z_3 e \Psi_3}{kT}\right) \quad (3.3)$$

O símbolo  $\Psi$  é o potencial eletrostático médio nas vizinhanças do grupo indicado pelo sub-índice, cuja correspondência com os grupos ionizáveis dos aminoácidos é dada pela notação apresentada na seção 2.2 e mantida ao longo de todo o texto. Para cada grupo, este potencial leva em consideração a interação devida ao próprio grupo, aos grupos vizinhos e aos íons em solução.  $T$  é a temperatura absoluta,  $k$  a constante de Boltzmann e  $e$  é a carga elétrica elementar.

As equações (3.1) e (3.2) podem ser exemplificadas pelo caso de um grupo ácido e outro básico. Para ácidos, a valência ( $z_1$ ) é -1 enquanto que para bases tal parâmetro ( $z_2$ ) é +1. Assim sendo, chega-se a:

$$K_1^* = K_1 \exp\left(\frac{e \Psi_1}{kT}\right) \longrightarrow \text{para o grupo ácido} \quad (3.4)$$

$$K_2^* = K_2 \exp\left(\frac{-e \Psi_2}{kT}\right) \longrightarrow \text{para o grupo básico} \quad (3.5)$$

Nota-se que se ambos os potenciais forem positivos, por exemplo, a constante de ionização do ácido irá aumentar o que significa que ele cede prótons mais facilmente, porém, a base terá sua constante de ionização diminuída, ou seja, ela apresentará uma maior dificuldade em se ionizar.

Os mesmos potenciais mencionados também influenciam na concentração de prótons nas vizinhanças dos grupos ionizáveis. Deste modo, nas expressões ao longo deste trabalho, o termo pH corresponde a  $-\log_{10}[\text{H}^+]$ , onde  $[\text{H}^+]$  é a concentração de prótons em uma região (*bulk*) na qual os potenciais são nulos.

Como as constantes de ionização são alteradas pelos potenciais eletrostáticos médios, o grau de ionização de cada grupo também é influenciado. Com base nas equações do primeiro capítulo e por (3.1), (3.2) e (3.3) tem-se:

$$\alpha_1 = \frac{1}{1 + 10^{(-\text{pH} + \text{p}K_1)} \exp\left(-\frac{e\Psi_1}{kT}\right)} \quad (3.6)$$

$$\alpha_2 = \frac{1}{1 + 10^{(-\text{pH} + \text{p}K_2)} \exp\left(\frac{e\Psi_2}{kT}\right)} \quad (3.7)$$

$$\alpha_3 = \frac{1}{1 + 10^{(-\text{pH} + \text{p}K_3)} \exp\left(\frac{e\Psi_3}{kT}\right)} \quad (3.8)$$

A expressão para  $\alpha_3$  foi escrita para uma cadeia lateral básica. Para o caso de uma cadeia ácida, o argumento da exponencial é negativo.

### 3.1 Solução da Equação de Poisson-Boltzmann:

A carga efetiva de cada grupo ionizável interage com íons e moléculas nas vizinhanças. A partir daí, para se descrever o potencial eletrostático médio lança-se mão da equação de Poisson-Boltzmann:

$$\nabla^2\Psi = -\frac{1}{\epsilon_{ef}} \sum_i z_i e n_{i0} \exp(-z_i e\Psi / kT) \quad (3.9)$$

Duas condições de contorno foram usadas na determinação do potencial eletrostático médio sendo que a primeira é que tanto o campo quanto o potencial



elétricos sejam nulos em uma região “distante” dos grupos ionizáveis. Devido ao fato de a carga de um grupo ionizável depender do grau de ionização deste, a segunda condição de contorno empregada é que, nas vizinhanças dos grupos ionizáveis, o campo elétrico médio seja dado em função do grau de ionização ( $\alpha$ ) do grupo. Levando em conta a fenomenologia do sistema, tal condição é matematicamente expressa por:

$$E_1 = -\frac{\alpha_1 e}{4\pi\epsilon_{ef} r_a^2} \quad (3.10)$$

$$E_2 = \frac{(1-\alpha_2)e}{4\pi\epsilon_{ef} r_b^2} \quad (3.11)$$

$$E_3 = \begin{cases} -\frac{\alpha_3 e}{4\pi\epsilon_{ef} r_c^2} \rightarrow \text{Se o grupo da cadeia lateral for ácido} \\ \frac{(1-\alpha_3)e}{4\pi\epsilon_{ef} r_c^2} \rightarrow \text{Se o grupo da cadeia lateral for básico} \end{cases} \quad (3.12)$$

$$\epsilon_{ef} = \epsilon_0 CD_{ef} \quad (3.13)$$

Os parâmetros  $r_a$ ,  $r_b$  e  $r_c$  são, respectivamente, os raios do grupo ácido ligado ao  $C_a$ , do grupo básico ligado ao  $C_a$  e do grupo ionizável da cadeia lateral. Como o sistema em estudo é constituído por dois meios de constantes dielétricas distintas, nos cálculos é empregada uma constante dielétrica efetiva denotada pelo símbolo  $CD_{ef}$ . O valor desta constante é determinado no capítulo 4. O símbolo  $\epsilon_{ef}$  representa a *permissividade efetiva*.

Assumindo que os grupos tenham simetria esférica e que  $\Psi$  possuísse apenas dependência radial, pode-se escrever a equação de Poisson-Boltzmann como:

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\Psi(r)}{dr} \right) = -\frac{1}{\epsilon_{ef}} \sum_i z_i e n_{i0} \exp(-z_i e \Psi(r) / kT) \quad (3.14)$$

A equação para o potencial eletrostático usada nesse trabalho, admitido como potencial da força média, é a forma linearizada da equação (3.14). Considerando a segunda hipótese da teoria de Debye-Hückel (i.e.,  $z_i e \Psi \ll kT$ ) pode se expandir a função exponencial em uma série de Taylor que, truncada nos dois primeiros termos, dá nova forma a equação (3.14):

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d}{dr} \Psi(r) \right) = -\frac{1}{\epsilon_{ef}} \sum_i z_i e n_{i0} + \frac{1}{\epsilon_{ef}} \sum_i \frac{z_i^2 e^2 \Psi(r) n_{i0}}{kT} \quad (3.15)$$

Como o sistema é eletricamente neutro, ou seja, a carga elétrica total contida em um volume infinitesimal, porém macroscópico, é zero, o primeiro termo do lado direito da equação acima é nulo. Assim, a equação a ser considerada é:

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d}{dr} \Psi(r) \right) = \frac{1}{\epsilon_{ef}} \sum_i \frac{z_i^2 e^2 \Psi(r) n_{i0}}{kT} \quad (3.16)$$

Esta equação pode ser expressa em termos de parâmetros adimensionais. Para tanto, definem-se o *comprimento de Debye* ( $\lambda^2$ ), o potencial reduzido ( $\phi_{(r)}$ ) e a variável  $\xi'$  que são dados, respectivamente, por:

$$\lambda^2 = \frac{\epsilon_{ef} kT}{\sum_i n_{i0} z_i^2 e^2} \quad (3.17)$$

$$\phi(r) = \frac{e\Psi(r)}{kT} \quad (3.18)$$

$$\xi' = r / \lambda \quad (3.19)$$

Novamente, o símbolo  $n_{i0}$  está relacionado com a concentração de íons em uma região (bulk) na qual o potencial eletrostático é nulo. Com estas três identidades inseridas na equação (3.16) e considerando-se uma solução que contenha apenas uma espécie iônica, chega-se a:

$$\frac{1}{\xi'^2} \frac{d}{d\xi'} \xi'^2 \frac{d}{d\xi'} \phi = \phi \quad (3.20)$$

A solução geral desta última equação diferencial pode ser escrita como uma combinação de seno hiperbólico e co-seno hiperbólico ou em termos de exponenciais, como se segue:

$$\phi(r) = \frac{a' \exp(-r/\lambda)}{(r/\lambda)} + \frac{b' \exp(r/\lambda)}{(r/\lambda)} \quad (3.21)$$

Os símbolos  $a'$  e  $b'$  são constantes a serem determinadas pelo uso de condições de contorno. Uma destas condições é que o potencial não divirja para “grandes” valores de  $r$  e para concordar com isso a constante  $b'$  tem que ser tomada como nula. Assim, a solução geral da equação (3.20) é dada por:

$$\phi(r) = \frac{a' \exp(-r/\lambda)}{(r/\lambda)} \quad (3.22)$$

Considere um ponto P cuja distância em relação aos centros dos grupos ionizáveis seja dada com base na figura 3.1 indicada a seguir.

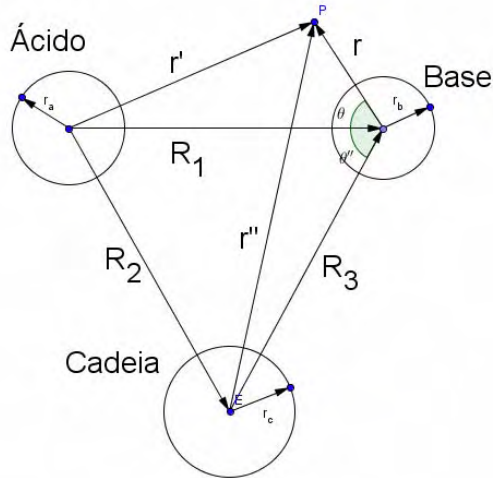


Figura 3.1: Distâncias entre um ponto  $P$  e os centros dos grupos ácido e base.

O potencial reduzido “total”,  $\varphi(r)$ , calculado no ponto  $P$  é o potencial que leva em conta as contribuições de todos os três grupos ionizáveis e dos íons em solução. Este potencial é dado por:

$$\varphi(r) = \varphi_a + \varphi_b + \varphi_c \quad (3.23)$$

$\varphi_a$ ,  $\varphi_b$ ,  $\varphi_c$  correspondem ao potencial no ponto  $P$  calculado a partir dos grupos ácido, base e da cadeia lateral, respectivamente. Esses potenciais são escritos na forma indicada pela expressão (3.22), ou seja:

$$\varphi_a = \frac{a \exp(-|\vec{r}'|/\lambda)}{(|\vec{r}'|/\lambda)} = \frac{a\lambda \exp(-|\vec{R}_1 + \vec{r}|/\lambda)}{|\vec{R}_1 + \vec{r}|} \quad (3.24)$$

$$\varphi_b = \frac{b \exp(-|\vec{r}|/\lambda)}{(|\vec{r}|/\lambda)} \quad (3.25)$$

$$\varphi_c = \frac{c \exp(-|\vec{r}''|/\lambda)}{(|\vec{r}''|/\lambda)} = \frac{c\lambda \exp(-|\vec{R}_3 + \vec{r}|/\lambda)}{|\vec{R}_3 + \vec{r}|} \quad (3.26)$$

A equação (3.24) informa o potencial reduzido exercido pelo grupo ácido, a equação (3.25) é o potencial reduzido devido ao grupo básico e a (3.26) é o potencial reduzido exercido pelo grupo ionizável da cadeia lateral. As constantes de integração  $a$ ,  $b$  e  $c$  são determinadas pelas condições de contorno relacionadas com os campos elétricos, i.e., equações (3.10), (3.11) e (3.12).

Com o intuito de se obter o potencial reduzido em termos da variável  $r$  e seguindo a definição de ângulos e vetores dada pela figura 3.1, as expressões (3.24) e (3.26) foram escritas como:

$$\varphi_a = \frac{a\lambda \exp(-|\vec{R}_1 + \vec{r}|/\lambda)}{|\vec{R}_1 + \vec{r}|} = a\lambda \frac{\exp\{-(r^2 + R_1^2 - 2rR_1 \cos\theta)^{1/2} / \lambda\}}{\left((r^2 + R_1^2 - 2rR_1 \cos\theta)^{1/2}\right)} \quad (3.27)$$

$$\varphi_c = \frac{c\lambda \exp(-|\vec{R}_3 + \vec{r}|/\lambda)}{|\vec{R}_3 + \vec{r}|} = c\lambda \frac{\exp\{-(r^2 + R_3^2 - 2rR_3 \cos\theta)^{1/2} / \lambda\}}{\left((r^2 + R_3^2 - 2rR_3 \cos\theta)^{1/2}\right)} \quad (3.28)$$

Os argumentos das duas últimas exponenciais podem ser expressos em termos de funções de Bessel e polinômios de Legendre<sup>56</sup>:

$$\begin{aligned} & \frac{\exp\{-(r^2 + R_1^2 - 2rR_1 \cos\theta)^{1/2} / \lambda\}}{\left((r^2 + R_1^2 - 2rR_1 \cos\theta)^{1/2}\right)} = \\ & = \sum_{m=0}^{\infty} (2m+1) \frac{1}{\sqrt{R_1}} K_{m+1/2}\left(\frac{R_1}{\lambda}\right) \frac{1}{\sqrt{r}} I_{m+1/2}\left(\frac{r}{\lambda}\right) P_m(\cos\theta) \end{aligned} \quad (3.29)$$

$$\begin{aligned} & \frac{\exp\{-(r^2 + R_3^2 - 2rR_3 \cos\theta)^{1/2} / \lambda\}}{\left((r^2 + R_3^2 - 2rR_3 \cos\theta)^{1/2}\right)} = \\ & = \sum_{m=0}^{\infty} (2m+1) \frac{1}{\sqrt{R_3}} K_{m+1/2}\left(\frac{R_3}{\lambda}\right) \frac{1}{\sqrt{r}} I_{m+1/2}\left(\frac{r}{\lambda}\right) P_m(\cos\theta) \end{aligned} \quad (3.30)$$

$I_{m+1/2}$  é a função de Bessel modificada de primeira espécie e  $K_{m+1/2}$  é a função de Bessel modificada de segunda espécie. De acordo com o teorema de adição dos harmônicos esféricos<sup>57</sup>, o polinômio de Legendre,  $P_m(\cos\theta)$ , é igual a:

$$P_m(\cos\theta) = \frac{4\pi}{2m+1} \sum_{n=-m}^m (-1)^n Y_m^{n*}(\hat{R}_1) Y_m^n(\hat{r}) \quad (3.31)$$

Combinando as equações (3.31) e (3.29), a expressão (3.27) se torna:

$$\varphi_a = a\lambda 4\pi \sum_{m=0}^{\infty} \sum_{n=-m}^m (-1)^n \frac{1}{\sqrt{R_1}} K_{m+1/2}\left(\frac{R_1}{\lambda}\right) \frac{1}{\sqrt{r}} I_{m+1/2}\left(\frac{r}{\lambda}\right) Y_m^{n*}(\hat{R}_1) Y_m^n(\hat{r}) \quad (3.32)$$

As funções de Bessel de primeira e segunda espécies podem ser respectivamente definidas em termos dos símbolos  $i_m$  e  $k_m$ :

$$K_{m+1/2}\left(\frac{R_1}{\lambda}\right) = k_m \left(\frac{R_1}{\lambda}\right) \cdot \left(\frac{2R_1}{\lambda\pi}\right)^{1/2} \quad (3.33)$$

$$I_{m+1/2}\left(\frac{r}{\lambda}\right) = i_m \left(\frac{r}{\lambda}\right) \cdot \left(\frac{2r}{\lambda\pi}\right)^{1/2} \quad (3.34)$$

Assim, inserindo (3.33) e (3.34) em (3.32) chega-se a:

$$\varphi_a = 8a \sum_{m=0}^{\infty} \sum_{n=-m}^m (-1)^n k_m \left( \frac{R_1}{\lambda} \right) i_m \left( \frac{r}{\lambda} \right) Y_m^{n*}(\hat{R}_1) Y_m^n(\hat{r}) \quad (3.35)$$

Com base no texto de Watson<sup>56</sup>, novamente, definem-se os seguintes termos de monopolo:

$$k_0 \left( \frac{R_1}{\lambda} \right) = \sqrt{\frac{\pi}{2}} \frac{\exp(-R_1 / \lambda)}{(R_1 / \lambda)} \quad (3.36)$$

$$i_0 \left( \frac{r}{\lambda} \right) = \sqrt{\frac{\pi}{2}} \frac{\sinh(r / \lambda)}{(r / \lambda)} \quad (3.37)$$

$$Y_0^0 = \frac{1}{\sqrt{4\pi}} \quad (3.38)$$

Considerando apenas os termos de monopolo indicados acima, as somatórias da equação (3.35) fornecem a seguinte expressão para  $\varphi_a$ :

$$\varphi_a = \frac{a \exp(-R_1 / \lambda)}{(R_1 / \lambda)} \frac{\sinh(r / \lambda)}{(r / \lambda)} \quad (3.39)$$

Utilizando um desenvolvimento análogo ao descrito acima e trocando  $R_1$  por  $R_3$  se obtém a seguinte expressão para  $\varphi_c$ :

$$\varphi_c = \frac{c \exp(-R_3 / \lambda)}{(R_3 / \lambda)} \frac{\sinh(r / \lambda)}{(r / \lambda)} \quad (3.40)$$

Finalmente, usando o princípio da superposição, o potencial eletrostático reduzido médio no ponto P levando em consideração a interação exercida pelos três grupos ionizáveis é dado pela soma das equações (3.40), (3.39) e (3.25):

$$\varphi(r) = \varphi_a + \varphi_b + \varphi_c = \frac{b \exp(-r / \lambda)}{(r / \lambda)} + \frac{\sinh(r / \lambda)}{(r / \lambda)} \left[ \frac{c \exp(-R_3 / \lambda)}{(R_3 / \lambda)} + \frac{a \exp(-R_1 / \lambda)}{(R_1 / \lambda)} \right] \quad (3.41)$$

Para o cálculo das constantes de ionização é necessário conhecer o potencial nas vizinhanças de cada grupo ionizável. Tendo por base a figura 3.1 e nomeando  $r_a$ ,  $r_b$  e  $r_c$  os raios dos grupos ácido, base e da cadeia lateral, respectivamente, têm-se:

$$\varphi_1 = \frac{a \exp(-r_a / \lambda)}{(r_a / \lambda)} + \frac{\sinh(r_a / \lambda)}{(r_a / \lambda)} \left[ \frac{b \exp(-R_1 / \lambda)}{(R_1 / \lambda)} + \frac{c \exp(-R_2 / \lambda)}{(R_2 / \lambda)} \right] \quad (3.42)$$

$$\varphi_2 = \frac{b \exp(-r_b / \lambda)}{(r_b / \lambda)} + \frac{\sinh(r_b / \lambda)}{(r_b / \lambda)} \left[ \frac{a \exp(-R_1 / \lambda)}{(R_1 / \lambda)} + \frac{c \exp(-R_3 / \lambda)}{(R_3 / \lambda)} \right] \quad (3.43)$$

$$\varphi_3 = \frac{c \exp(-r_c / \lambda)}{(r_c / \lambda)} + \frac{\sinh(r_c / \lambda)}{(r_c / \lambda)} \left[ \frac{a \exp(-R_2 / \lambda)}{(R_2 / \lambda)} + \frac{b \exp(-R_3 / \lambda)}{(R_3 / \lambda)} \right] \quad (3.44)$$

Aplicando-se apenas o negativo da parte radial do gradiente em coordenadas esféricas nas equações acima, o campo elétrico ( $\vec{E}$ ) nas vizinhanças de cada grupo ionizável é:

$$\vec{E}_1 = \frac{kT}{e} \left\{ a \frac{\exp(-r_a / \lambda)}{r_a} \left[ 1 + \frac{\lambda}{r_a} \right] + \left[ \frac{b\lambda}{R_1} \exp(-R_1 / \lambda) + \frac{c\lambda}{R_2} \exp(-R_2 / \lambda) \right] \left[ \frac{\lambda}{r_a^2} \sinh\left(\frac{r_a}{\lambda}\right) - \frac{\cosh\left(\frac{r_a}{\lambda}\right)}{r_a} \right] \right\} \hat{r} \quad (3.45)$$

$$\vec{E}_2 = \frac{kT}{e} \left\{ b \frac{\exp(-r_b / \lambda)}{r_b} \left[ 1 + \frac{\lambda}{r_b} \right] + \left[ \frac{a\lambda}{R_1} \exp(-R_1 / \lambda) + \frac{c\lambda}{R_3} \exp(-R_3 / \lambda) \right] \left[ \frac{\lambda}{r_b^2} \sinh\left(\frac{r_b}{\lambda}\right) - \frac{\cosh\left(\frac{r_b}{\lambda}\right)}{r_b} \right] \right\} \hat{r} \quad (3.46)$$

$$\vec{E}_3 = \frac{kT}{e} \left\{ c \frac{\exp(-r_c / \lambda)}{r_c} \left[ 1 + \frac{\lambda}{r_c} \right] + \left[ \frac{b\lambda}{R_3} \exp(-R_3 / \lambda) + \frac{a\lambda}{R_2} \exp(-R_2 / \lambda) \right] \left[ \frac{\lambda}{r_c^2} \sinh\left(\frac{r_c}{\lambda}\right) - \frac{\cosh\left(\frac{r_c}{\lambda}\right)}{r_c} \right] \right\} \hat{r} \quad (3.47)$$

O termo  $kT/e$  provém da própria definição do potencial reduzido.  $\vec{E}_1$  é o campo elétrico médio nas vizinhanças do grupo ácido (que possui raio igual a  $r_a$ );  $\vec{E}_2$  é o campo elétrico médio nas vizinhanças do grupo básico, de raio igual a  $r_b$ ; e  $\vec{E}_3$  é o campo elétrico médio nas vizinhanças do grupo ionizável da cadeia lateral, de raio  $r_c$ .

### 3.2 Determinação das constantes de integração $a$ , $b$ , $c$ e cálculo do grau de ionização:

As equações de (3.10) a (3.12) permitem relacionar o campo elétrico na superfície de um determinado grupo com o seu grau de ionização. Seguindo este raciocínio e pela fenomenologia do sistema estudado, as equações supracitadas juntamente com (3.45), (3.46) e (3.47) fornecem:

$$\frac{-\alpha_1 e^2}{4\pi\epsilon_{ef} r_a^2 kT} = a \frac{\exp(-r_a / \lambda)}{r_a} \left[ 1 + \frac{\lambda}{r_a} \right] + \left[ \frac{b\lambda}{R_1} \exp(-R_1 / \lambda) + \frac{c\lambda}{R_2} \exp(-R_2 / \lambda) \right] \left[ \frac{\lambda}{r_a^2} \sinh\left(\frac{r_a}{\lambda}\right) - \frac{\cosh\left(\frac{r_a}{\lambda}\right)}{r_a} \right] \quad (3.48)$$

$$\frac{(1 - \alpha_2) e^2}{4\pi\epsilon_{ef} r_b^2 kT} = b \frac{\exp(-r_b / \lambda)}{r_b} \left[ 1 + \frac{\lambda}{r_b} \right] + \left[ \frac{a\lambda}{R_1} \exp(-R_1 / \lambda) + \frac{c\lambda}{R_3} \exp(-R_3 / \lambda) \right] \left[ \frac{\lambda}{r_b^2} \sinh\left(\frac{r_b}{\lambda}\right) - \frac{\cosh\left(\frac{r_b}{\lambda}\right)}{r_b} \right] \quad (3.49)$$

$$c \frac{\exp(-r_c / \lambda)}{r_c} \left[ 1 + \frac{\lambda}{r_c} \right] + \left[ \frac{\lambda}{r_c^2} \sinh\left(\frac{r_c}{\lambda}\right) - \frac{\cosh\left(\frac{r_c}{\lambda}\right)}{r_c} \right]. \quad (3.50)$$

$$\cdot \left[ \frac{b\lambda}{R_3} \exp(-R_3 / \lambda) + \frac{a\lambda}{R_2} \exp(-R_2 / \lambda) \right] =$$

$$= \left\{ \begin{array}{l} -\frac{\alpha_3 e^2}{4\pi\epsilon_{ef} r_c^2 kT} \rightarrow \text{Se o grupo da cadeia lateral for ácido} \\ \frac{(1 - \alpha_3) e^2}{4\pi\epsilon_{ef} r_c^2 kT} \rightarrow \text{Se o grupo da cadeia lateral for básico} \end{array} \right\} \quad (3.51)$$

As equações (3.6), (3.7) e (3.8) relacionam os graus de ionização com os potenciais eletrostáticos. Inserindo em cada uma delas (3.42), (3.43) e (3.44), respectivamente, obtém-se:

$$\alpha_1 = \frac{1}{1 + 10^{(-pH + pK_1)} \exp\left(-\left\{ \frac{a \exp(-r_a / \lambda)}{(r_a / \lambda)} + \frac{\sinh(r_a / \lambda)}{(r_a / \lambda)} \left[ \frac{b \exp(-R_1 / \lambda)}{(R_1 / \lambda)} + \frac{c \exp(-R_2 / \lambda)}{(R_2 / \lambda)} \right] \right\} \right)} \quad (3.52)$$

$$\alpha_2 = \frac{1}{1 + 10^{(-pH + pK_2)} \exp\left(\frac{b \exp(-r_b / \lambda)}{(r_b / \lambda)} + \frac{\sinh(r_b / \lambda)}{(r_b / \lambda)} \left[ \frac{a \exp(-R_1 / \lambda)}{(R_1 / \lambda)} + \frac{c \exp(-R_3 / \lambda)}{(R_3 / \lambda)} \right] \right)} \quad (3.53)$$

$$\alpha_3 = \frac{1}{1 + 10^{(-pH + pK_3)} \exp\left(\frac{c \exp(-r_c / \lambda)}{(r_c / \lambda)} + \frac{\sinh(r_c / \lambda)}{(r_c / \lambda)} \left[ \frac{a \exp(-R_2 / \lambda)}{(R_2 / \lambda)} + \frac{b \exp(-R_3 / \lambda)}{(R_3 / \lambda)} \right] \right)} \quad (3.54)$$

Novamente, a expressão para o grau de ionização do grupo ionizável da cadeia lateral foi escrita como se este fosse um grupo básico. Para o caso de se ter um grupo ácido, o argumento da exponencial da expressão (3.54) deve ser negativo, obtendo assim uma relação semelhante a (3.52).

Trabalhando com as sete últimas equações é possível se obter as seguintes relações entre os campos elétricos médios e os graus de ionização corrigidos pelos potenciais elétricos médios ( $\varphi$ ):

#### 🚩 Grupo 1:

$$a \frac{e^{-r_1 / \lambda}}{r_1} \left[ 1 + \frac{\lambda}{r_1} \right] + \left[ b \frac{e^{-R_1 / \lambda}}{R_1 / \lambda} + c \frac{e^{-R_2 / \lambda}}{R_2 / \lambda} \right] \left[ \frac{\lambda}{r_1^2} \sinh(r_1 / \lambda) - \frac{\cosh(r_1 / \lambda)}{r_1} \right] = \quad (3.55)$$

$$= \frac{(z_1 e)^2}{4\pi\epsilon_{ef} r_1^2 kT} \left[ \frac{1}{1 + 10^{(-pH + pK_1)} \exp\left(-\left\{ \frac{a \exp(-r_a / \lambda)}{(r_a / \lambda)} + \frac{\sinh(r_a / \lambda)}{(r_a / \lambda)} \left[ \frac{b \exp(-R_1 / \lambda)}{(R_1 / \lambda)} + \frac{c \exp(-R_2 / \lambda)}{(R_2 / \lambda)} \right] \right\} \right)} \right]$$

✚ Grupo 2:

$$\begin{aligned}
 & b \frac{e^{-r_2/\lambda}}{r_2} \left[ 1 + \frac{\lambda}{r_2} \right] + \left[ a \frac{e^{-R_1/\lambda}}{R_1/\lambda} + c \frac{e^{-R_3/\lambda}}{R_3/\lambda} \right] \left[ \frac{\lambda}{r_2^2} \sinh(r_2/\lambda) - \frac{\cosh(r_2/\lambda)}{r_2} \right] = \\
 & = \frac{(z_2 e)^2}{4\pi \varepsilon_{ef} r_2^2 kT} \left[ \frac{1}{1 + 10^{(-pH + pK_2)} \exp \left( -\frac{b \exp(-r_b/\lambda)}{(r_b/\lambda)} - \frac{\sinh(r_b/\lambda)}{(r_b/\lambda)} \left[ \frac{a \exp(-R_1/\lambda)}{(R_1/\lambda)} + \frac{c \exp(-R_3/\lambda)}{(R_3/\lambda)} \right] \right)} \right] \quad (3.56)
 \end{aligned}$$

✚ Grupo 3:

$$\begin{aligned}
 & c \frac{e^{-r_3/\lambda}}{r_3} \left[ 1 + \frac{\lambda}{r_3} \right] + \left[ a \frac{e^{-R_2/\lambda}}{R_2/\lambda} + b \frac{e^{-R_3/\lambda}}{R_3/\lambda} \right] \left[ \frac{\lambda}{r_3^2} \sinh(r_3/\lambda) - \frac{\cosh(r_3/\lambda)}{r_3} \right] = \\
 & = \frac{(z_3 e)^2}{4\pi \varepsilon_{ef} r_3^2 kT} \left[ \frac{1}{1 + 10^{(-pH + pK_3)} \exp \left( -\frac{c \exp(-r_c/\lambda)}{(r_c/\lambda)} - \frac{\sinh(r_c/\lambda)}{(r_c/\lambda)} \left[ \frac{a \exp(-R_2/\lambda)}{(R_2/\lambda)} + \frac{b \exp(-R_3/\lambda)}{(R_3/\lambda)} \right] \right)} \right] \quad (3.57)
 \end{aligned}$$

Devido à forma das três últimas equações, nota-se que não é possível escrever a constante de integração  $a$  em função de  $b$ ,  $b$  em função de  $c$  e  $c$  em função de  $a$ . As equações (3.55), (3.56) e (3.57) são conhecidas como *equações algébricas transcendentais acopladas* e, para a sua resolução, foi utilizado o método das aplicações contratantes (vide Kolmogorov<sup>58</sup>). Trabalhou-se com estas três últimas equações de modo a se chegar a igualdades entre as constantes de integração e funções denotadas por  $f(a,b,c)$ ,  $g(a,b,c)$  e  $h(a,b,c)$  delas dependentes, definidas como se segue:

$$a = f(a,b,c)$$

Onde:

$$\begin{aligned}
 f(a,b,c) = & \frac{\frac{(z_1 e)^2}{4\pi \varepsilon_{ef} r_a^2 kT} \left[ \frac{1}{1 + 10^{-pH + pK_1} \exp \left[ -a \frac{e^{-r_1/\lambda}}{r_1/\lambda} - \frac{\sinh(r_1/\lambda)}{(r_1/\lambda)} \left[ b \frac{e^{-R_1/\lambda}}{R_1/\lambda} + c \frac{e^{-R_2/\lambda}}{R_2/\lambda} \right] \right]} \right]}{\frac{e^{-r_1/\lambda}}{r_1} \left[ 1 + \frac{\lambda}{r_1} \right]} \\
 & \frac{\left[ b \frac{e^{-R_1/\lambda}}{R_1/\lambda} + c \frac{e^{-R_2/\lambda}}{R_2/\lambda} \right] \left[ \frac{\lambda}{r_1^2} \sinh(r_1/\lambda) - \frac{\cosh(r_1/\lambda)}{r_1} \right]}{\frac{e^{-r_1/\lambda}}{r_1} \left[ 1 + \frac{\lambda}{r_1} \right]} \quad (3.58)
 \end{aligned}$$



$$b = g(a, b, c)$$

Onde:

$$g(a, b, c) = \frac{\frac{(z_2 e)^2}{4\pi\epsilon_{ef} r_2^2 kT} \left[ \frac{1}{1 + 10^{-pH + pK_2} \exp \left[ -b \frac{e^{-r_2/\lambda}}{r_2/\lambda} - \frac{\sinh(r_2/\lambda)}{(r_2/\lambda)} \left[ a \frac{e^{-R_1/\lambda}}{R_1/\lambda} + c \frac{e^{-R_3/\lambda}}{R_3/\lambda} \right] \right]}{\frac{e^{-r_2/\lambda}}{r_2} \left[ 1 + \frac{\lambda}{r_2} \right]} \right]}{\frac{\left[ a \frac{e^{-R_1/\lambda}}{R_1/\lambda} + c \frac{e^{-R_3/\lambda}}{R_3/\lambda} \right] \left[ \frac{\lambda}{r_2^2} \sinh(r_2/\lambda) - \frac{\cosh(r_2/\lambda)}{r_2} \right]}{\frac{e^{-r_2/\lambda}}{r_2} \left[ 1 + \frac{\lambda}{r_2} \right]}} \quad (3.59)$$

$$c = h(a, b, c)$$

Onde:

$$h(a, b, c) = \frac{\frac{(z_3 e)^2}{4\pi\epsilon_{ef} r_3^2 kT} \left[ \frac{1}{1 + 10^{-pH + pK_3} \exp \left[ c \frac{e^{-r_3/\lambda}}{r_3/\lambda} + \frac{\sinh(r_3/\lambda)}{(r_3/\lambda)} \left[ a \frac{e^{-R_2/\lambda}}{R_2/\lambda} + b \frac{e^{-R_1/\lambda}}{R_1/\lambda} \right] \right]}{\frac{e^{-r_3/\lambda}}{r_3} \left[ 1 + \frac{\lambda}{r_3} \right]} \right]}{\frac{\left[ a \frac{e^{-R_2/\lambda}}{R_2/\lambda} + b \frac{e^{-R_1/\lambda}}{R_1/\lambda} \right] \left[ \frac{\lambda}{r_3^2} \sinh(r_3/\lambda) - \frac{\cosh(r_3/\lambda)}{r_3} \right]}{\frac{e^{-r_3/\lambda}}{r_3} \left[ 1 + \frac{\lambda}{r_3} \right]}} \quad (3.60)$$

Com a aplicação do método das aplicações contratantes, primeiramente são atribuídos um valor inicial aos parâmetros  $c$  e  $b$  em  $f(a, b, c)$ , por exemplo, e, a partir disto, determina-se um valor para  $a$ . Com  $a$  determinado, tal valor é usado em  $g(a, b, c)$  para se calcular  $b$ . Este valor de  $b$  é inserido em  $h(a, b, c)$  para se calcular  $c$ . Quando o novo valor assim obtido para  $c$  difere do valor  $a$  ele atribuído no passo anterior, este novo valor de  $c$  é utilizado em  $f(a, b, c)$  e repete-se o processo até que se chegue à convergência, que ocorre quando satisfeito o seguinte critério:

$$\left| \frac{\beta_{i+1} - \beta_i}{\beta_i} \right| < 10^{-7} \quad (3.61)$$

Sendo  $\beta_{i+1}$  o valor de saída de cada uma das constantes ( $a$ ,  $b$  e  $c$ ) após o cálculo de  $h(a, b, c)$  e  $\beta_i$  o valor de entrada destas no passo anterior. Utilizou-se  $10^{-7}$  como parâmetro de "corte", pois tais cálculos foram realizados computacionalmente e, dessa forma e devido à precisão das variáveis adotadas, grandezas que difiram  $10^{-7}$  entre si podem ser consideradas como "iguais". As constantes de integração são consideradas determinadas quando todos os parâmetros  $a$ ,  $b$  e  $c$  satisfizerem (3.61).

Utilizando o método introduzido neste capítulo, elaborou-se um programa na linguagem FORTRAN para calcular as constantes de integração e, conseqüentemente, os potenciais eletrostáticos médios e as grandezas deles dependentes em função do pH, da concentração iônica e da temperatura para todos os grupos ionizáveis de cada um dos aminoácidos estudados. A equação de Henderson-Hasselbach permite correlacionar o grau de ionização de cada grupo com o seu pK. Por base na fenomenologia, a carga elétrica média de cada grupo e a carga total dos aminoácidos também podem ser inferidas a partir do grau de ionização,  $\alpha$ .

A próxima tarefa é a determinação da constante dielétrica efetiva ( $CD_{ef}$ ) a ser empregada nos cálculos. Tal processo está detalhadamente descrito no capítulo seguinte.

# Capítulo 4

## Modelo para o Cálculo da Constante Dielétrica Efetiva dos Aminoácidos

O cálculo da constante dielétrica efetiva dos aminoácidos é realizado considerando-se um sistema constituído por dois meios dielétricos lineares distintos, conforme mostrado nas figuras 2.1, 2.2 e 4.1.

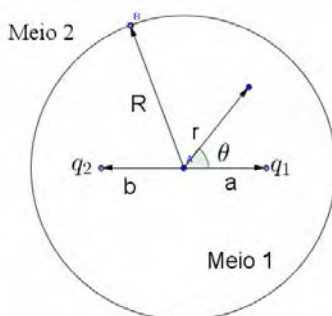


Figura 4.1: Representação de um sistema de duas cargas dentro de uma esfera de raio  $R$ . A carga  $q_1$  dista  $a$  da origem e, para a carga  $q_2$ , esta distância vale  $b$ .

O meio 1 é circunscrito por uma superfície esférica de raio  $R$  e caracterizado pela constante dielétrica  $\epsilon_1$  enquanto que o meio 2 corresponde a toda a região externa a tal superfície e possui constante dielétrica  $\epsilon_2$ . Visando a aplicação deste modelo a moléculas biológicas, como os aminoácidos, e devido à própria constituição molecular destes, considerou-se o meio 1 como semelhante a um hidrocarboneto possuindo, assim,  $\epsilon_1$  igual a  $2^{43,59}$ . No meio 2 tem-se água e íons em solução e para a determinação de  $\epsilon_2$  em função da temperatura ( $^{\circ}\text{C}$ ) e concentração iônica ( $n_{i0}$ , em mols por litro)<sup>45</sup> utiliza-se a seguinte interpolação<sup>44,45</sup>:

$$\begin{aligned} \epsilon_2 = & 87,7404 - 0,40014T + 0,940425 \cdot 10^{-3}T^2 - \\ & - 0,140165 \cdot 10^{-5}T^3 - 0,929212 \cdot 10^{-10}T^4 + 2\delta n_{i0} \end{aligned} \quad (4.1)$$

para NaCl, o parâmetro  $\delta$  é 5,5 litros/mol.

No interior da esfera existem duas cargas, as quais foram denotadas por  $q_1$  e  $q_2$ . Tomando o centro da esfera como origem, a posição de tais cargas é indicada na figura 4.1. O potencial eletrostático no meio 1 ( $\Psi_1$ ) satisfaz a equação de Poisson a qual, pela disposição das cargas apresentada, pode ser escrita como:

$$\nabla^2 \Psi_1 = \frac{-q_1 \delta(\vec{r} - \vec{a}) - q_2 \delta(\vec{r} - \vec{b})}{\epsilon_0 \epsilon_1} \quad (4.1)(4.2)$$

Escrevendo  $x = \cos \theta$ ,  $y = \cos(\pi - \theta)$  e sendo  $P_l$  o polinômio de Legendre de ordem  $l$ , obtém-se a seguinte solução, válida para  $|\vec{r}| \geq (|\vec{a}|, |\vec{b}|)$ :

$$\Psi_1 = \frac{q_1}{4\pi\epsilon_0\epsilon_1} \sum_{l=0}^{\infty} \frac{a^l}{r^{l+1}} P_l(x) + \sum_{l=0}^{\infty} T_l r^l P_l(x) + \frac{q_2}{4\pi\epsilon_0\epsilon_1} \sum_{l=0}^{\infty} \frac{b^l}{r^{l+1}} P_l(y) + \sum_{l=0}^{\infty} V_l r^l P_l(y) \quad (4.3)$$

Exterior ao meio 1 tem-se o meio 2, região na qual o potencial ( $\Psi_{2DH}$ ) é descrito pela forma linearizada da equação de Poisson-Boltzmann (3.16):

$$\nabla^2 \Psi_{2DH} = \frac{\Psi_{2DH}}{\lambda^2} \quad (4.4)$$

com:

$$\lambda^2 = \frac{\epsilon_0 \epsilon_2 kT}{\sum_i n_{i0} z_i^2 e^2} \quad (4.5)$$

onde o símbolo  $n_{i0}$  é a concentração de íons em uma região (bulk) em que o potencial eletrostático é nulo,  $z_i$  é valência iônica e  $e$  é a carga elementar. Considerando simetria azimutal, a solução da equação (4.4), expressa em termos das funções de Bessel<sup>15,60</sup>  $i_l$  e  $k_l$ , é:

$$\Psi_{2DH} = \sum_{l=0}^{\infty} X_l \frac{2(2l+1)}{\pi} i_l \left( \frac{a}{\lambda} \right) k_l \left( \frac{r}{\lambda} \right) P_l(x) + \sum_{l=0}^{\infty} Y_l \frac{2(2l+1)}{\pi} i_l \left( \frac{b}{\lambda} \right) k_l \left( \frac{r}{\lambda} \right) P_l(y) \quad (4.6)$$

Para a determinação dos parâmetros  $T_l$ ,  $V_l$ ,  $X_l$  e  $Y_l$ , observa-se as seguintes condições de contorno na interface entre os meios (i.e.,  $r=R$ ):

✚ Continuidade da componente normal do vetor deslocamento elétrico:

$$-\epsilon_1 \nabla \Psi_1(\hat{r}) \Big|_{r=R} = -\epsilon_2 \nabla \Psi_{2DH}(\hat{r}) \Big|_{r=R} \quad (4.7)$$

✚ Continuidade do potencial eletrostático:

$$\Psi_1 \Big|_{r=R} = \Psi_{2DH} \Big|_{r=R} \quad (4.8)$$

Através destas duas condições, é possível obter:

$$X_l = \frac{q_1 a^l}{8\varepsilon_0 R^{l+2} i_l \left(\frac{a}{\lambda}\right) \left[ \frac{l}{R} k_l \left(\frac{R}{\lambda}\right) (\varepsilon_1 - \varepsilon_2) + \frac{\varepsilon_2}{\lambda} K_{l+1} \left(\frac{R}{\lambda}\right) \right]} \quad (4.9)$$

$$T_l = \frac{q_1 a^l (2l+1) K_l \left(\frac{R}{\lambda}\right)}{4\pi\varepsilon_0 R^{2l+2} \left[ \frac{l}{R} k_l \left(\frac{R}{\lambda}\right) (\varepsilon_1 - \varepsilon_2) + \frac{\varepsilon_2}{\lambda} K_{l+1} \left(\frac{R}{\lambda}\right) \right]} - \frac{q_1 a^l}{4\pi\varepsilon_0 \varepsilon_1 R^{2l+1}} \quad (4.10)$$

$$Y_l = \frac{q_2 b^l}{8\varepsilon_0 R^{l+2} i_l \left(\frac{b}{\lambda}\right) \left[ \frac{l}{R} k_l \left(\frac{R}{\lambda}\right) (\varepsilon_1 - \varepsilon_2) + \frac{\varepsilon_2}{\lambda} K_{l+1} \left(\frac{R}{\lambda}\right) \right]} \quad (4.11)$$

$$V_l = \frac{q_2 b^l (2l+1) K_l \left(\frac{R}{\lambda}\right)}{4\pi\varepsilon_0 R^{2l+2} \left[ \frac{l}{R} k_l \left(\frac{R}{\lambda}\right) (\varepsilon_1 - \varepsilon_2) + \frac{\varepsilon_2}{\lambda} K_{l+1} \left(\frac{R}{\lambda}\right) \right]} - \frac{q_2 b^l}{4\pi\varepsilon_0 \varepsilon_1 R^{2l+1}} \quad (4.12)$$

Conhecidos tais parâmetros, tem-se a solução completa para  $\Psi_1$  e  $\Psi_{2DH}$ . A constante dielétrica efetiva, dependente do meio onde é calculada, é então dada pela seguinte expressão:

$$CD_{ef} = \begin{cases} \frac{1}{\Psi_{1DH}} \left[ \frac{q_1}{4\pi\varepsilon_0 |\vec{r} - \vec{a}|} + \frac{q_2}{4\pi\varepsilon_0 |\vec{r} - \vec{b}|} \right] & \rightarrow \text{pontos internos à esfera} \\ \frac{1}{\Psi_{2DH}} \left[ \frac{q_1}{4\pi\varepsilon_0 |\vec{r} - \vec{a}|} + \frac{q_2}{4\pi\varepsilon_0 |\vec{r} - \vec{b}|} \right] & \rightarrow \text{pontos externos à esfera} \end{cases} \quad (4.13)$$

Utilizando-se dos valores fornecidos pela tabela 2.3, o raio ( $R_{esf}$ ) da esfera que engloba cada par de grupos analisados é estabelecido pela seguinte expressão:

$$R_{esf} = \frac{r_i + r_j + R_{i \rightarrow j}}{2} \quad (4.14)$$

onde  $r_i$  e  $r_j$  são os raios dos grupos envolvidos ( $i$  e  $j$  podem ser 1, 2 ou 3) e  $R_{i \rightarrow j}$  é a distância entre esses grupos ( $R_1$ ,  $R_2$  ou  $R_3$ ).

Com base nas expressões desenvolvidas, construiu-se um programa computacional para o cálculo da constante dielétrica efetiva em termos da temperatura, concentração iônica e posição do ponto desejado. Para a determinação da  $CD_{ef}$  dos aminoácidos estudados, considera-se nula uma das cargas e, sob o ponto onde esta se encontraria, calculou-se o potencial exercido pela remanescente. Neste caso, seguindo a

figura 4.1, o parâmetro  $r$  assume o valor de  $R_{\text{esf}}$  menos o raio do grupo carregado correspondente à carga anulada, ou seja:

$$r = \frac{R_{i \rightarrow j} + r_i - r_j}{2} \quad (4.15)$$

# Capítulo 5

## Resultados: Constante Dielétrica Efetiva

Para 25°C e 150mM de NaCl, os valores de constante dielétrica efetiva calculados estão mostrados na tabela 5.1.

*Tabela 5.1: Aminoácidos e respectivos valores de constante dielétrica efetiva calculados em 25°C e 150mM de NaCl*

<b>Aminoácidos</b>	<b>Grupos Envolvidos</b>	<b>Constante Dielétrica Efetiva (CD<sub>ef</sub>)</b>
Arginina	Ácido-Base	15,66
	Ácido-Cadeia	79,95
Histidina	Ácido-Base	15,66
	Ácido-Cadeia	26,11
Lisina	Ácido-Base	15,66
	Ácido-Cadeia	79,95
Cisteína	Ácido-Base	25,27
	Base-Cadeia	14,77
Tirosina	Ácido-Base	15,66
	Base-Cadeia	79,95
Ácido Aspártico	Ácido-Base	16,70
	Base-Cadeia	25,25
Ácido Glutâmico	Ácido-Base	25,25
	Base-Cadeia	29,01

Os elevados valores de CD<sub>ef</sub> observados nas regiões ácido-cadeia e base-cadeia dos aminoácidos Arg, Lys e Tyr se devem à própria geometria destas moléculas, caracterizadas por suas extensas cadeias laterais (veja as figuras 1.4, 1.5 e os valores de R<sub>2</sub> e R<sub>3</sub> na tabela 2.3). Devido a estas cadeias serem praticamente “lineares” e a predominância de moléculas de água na região entre os grupos, a interação exercida pelas cargas destes será atenuada apenas por uma constante dielétrica cujo valor é

praticamente o mesmo que o de  $\epsilon_2$  (i.e., 79,95). Para cada aminoácido, o valor da  $CD_{ef}$  correspondente à região não mencionada na tabela 5.1 é o mesmo de  $\epsilon_2$ .

As regiões ácido-base da His, Lys Arg e Tyr possuem o mesmo valor de constante dielétrica efetiva, pois a distância entre estes grupos ( $R_1 = 3,75\text{\AA}$ ) é a mesma em tais aminoácidos. O mesmo argumento vale para a igualdade observada entre Cys e Glu.

Para as outras regiões, os valores da constante dielétrica efetiva estão entre 15 e 29. Tal faixa de valores concorda com os trabalhos que empregam tal parâmetro em cálculos que envolvem a interação eletrostática e a variação dos valores de pK de resíduos de aminoácidos em proteínas<sup>38</sup>.

Lys apresenta, segundo o modelo aqui desenvolvido,  $CD_{ef}$  próxima a 16 para a região ácido-base. Para este mesmo aminoácido, Isom<sup>20</sup> estimou valores entre 8 e 26. Já para as regiões envoltas do ácido glutâmico, nas quais a  $CD_{ef}$  calculada está entre 25 e 29, o mesmo autor reporta<sup>21</sup> valores entre 9 e 38. Estudando resíduos de Asp na *Staphylococcal Nuclease*, Castañeda<sup>61</sup> utilizou  $CD_{ef}$  igual a 20. Novamente percebe-se que os resultados aqui apresentados estão consistentes com os presentes na literatura e os valores desta oriundos são, pelo modelo aqui desenvolvido, justificados.

Para visualizar a variação da constante dielétrica efetiva com a posição do ponto onde se calcula o potencial, considerou-se o caso de duas cargas de sinais contrários e equidistantes  $2\text{\AA}$  do centro de uma esfera de raio  $5\text{\AA}$ . Adotou-se o meio externo como tendo uma concentração de 150mM de NaCl. O potencial foi calculado sobre a linha que une as cargas e a partir de  $1\text{\AA}$  da carga negativa (nesta situação,  $r$  é igual a  $3\text{\AA}$  e  $\theta$  é igual à zero).

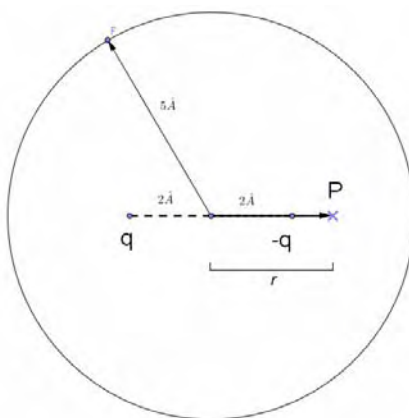


Figura 5.1: Cargas de sinais contrários dispostas  $2\text{\AA}$  da origem e internas à esfera de raio  $5\text{\AA}$ . No ponto  $P$ , distante  $r$  da origem, é onde se calculou o potencial.



Os resultados são mostrados nas figuras 5.2 e 5.3.

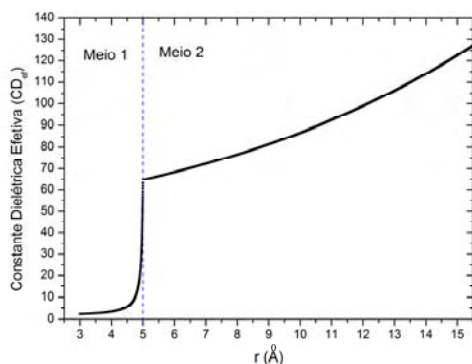


Figura 5.2: Constante Dielétrica Efetiva versus Distância ( $r$ ) para um sistema de duas cargas. 25°C

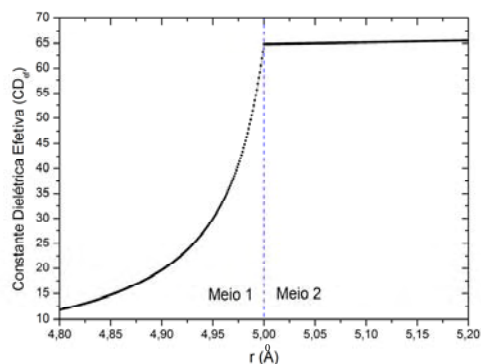


Figura 5.3: Constante Dielétrica Efetiva versus Distância ( $r$ ). Zoom na interface de um sistema de duas cargas. 25°C

Nas figuras 5.2 e 5.3, a linha azul pontilhada foi usada para delimitar os meios 1 e 2. Percebe-se, principalmente pelo *zoom* da figura 5.3, que próximo à interface há uma grande variação da constante dielétrica em um curto espaço (4,8Å a 5,20Å). Um dos principais motivos disto é a proximidade de um meio de elevada constante dielétrica ( $\epsilon_2$ ), cerca de 40 vezes o valor do meio interno<sup>62</sup>.

Valores de  $CD_{ef}$  entre 44 e 128 foram encontrados por Rees<sup>39</sup> ao estudar, no *citocromo c*, a interação entre um átomo de ferro e resíduos de lisina. Nas duas últimas figuras, apesar de representarem uma situação com parâmetros distintos aos usados por ele, observa-se que tais valores são encontrados em regiões próximas à interface o que é, novamente, justificado pelo papel exercido pelo solvente no qual, além de água, há íons monovalentes.

No meio 2, a curva para constante dielétrica efetiva apresenta o comportamento sempre crescente, pois tal constante é calculada pela divisão do potencial coulombiano pelo potencial de Debye-Hückel (expressão semelhante a (3.22)) o qual, por envolver um termo exponencial, vai a zero “mais rapidamente” do que  $1/r$ . Outro fator, também relacionado com a distância  $r$ , e que contribui para este comportamento crescente, é que para altos valores de  $r$  a expansão em série de Taylor que origina a forma linearizada da equação de Poisson-Boltzmann acarreta grandes erros.

Como em seções posteriores é descrita a aplicação do modelo a condições próximas às fisiológicas de temperatura e concentração iônica, realizaram-se os cálculos de constante dielétrica efetiva também para 36,5°C. Seguem gráficos (figuras 5.4 e 5.5)

análogos aos anteriores e que permitem a comparação entre os valores de  $CD_{ef}$  obtidos para duas temperaturas distintas.

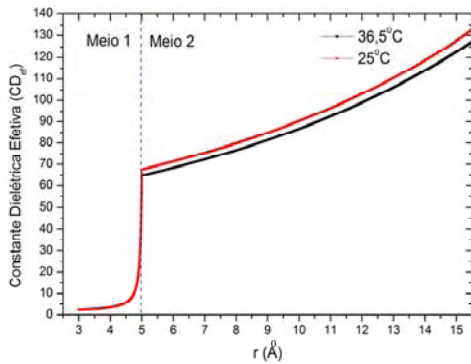


Figura 5.4: Constante Dielétrica Efetiva versus Distância ( $r$ ). Comparação entre temperaturas.

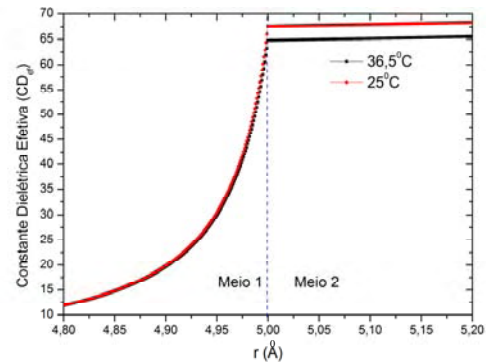


Figura 5.5: Constante Dielétrica Efetiva versus Distância ( $r$ ). Zoom na interface e comparação entre temperaturas.

Nas figuras 5.4 e 5.5 nota-se que o aumento da temperatura resulta em valores ligeiramente menores de  $CD_{ef}$ . Tal comportamento é devido ao fato de que o valor da constante dielétrica do meio 2, constituído por água e íons monovalentes, é regido pela expressão (4.1) que, com o acréscimo da temperatura, fornece menores valores de  $\epsilon_2$ .

# Capítulo 6

## Resultados e Discussões: Grau de Ionização e Grandezas Correlatas

No modelo adotado neste trabalho, os aminoácidos são tratados como aminoácidos de *ensemble*, conseqüentemente, os resultados apresentados se caracterizam como valores médios.

### 6.1 Grau de ionização ( $\alpha$ ) e $pK^*$ :

Conhecida a constante dielétrica efetiva de cada aminoácido, ela foi inserida no cálculo do comprimento de Debye ( $\lambda$ ) que, por sua vez, é dado pela equação (3.17). Com as constantes de integração determinadas por um programa computacional (conforme descrito na seção 3.2) obtiveram-se valores dos potenciais reduzidos na superfície de cada grupo ionizável e, a partir dessa grandeza, calcularam-se os respectivos grau de ionização,  $pK^*$  e cargas elétricas médias. Salvo outra indicação, a temperatura empregada nos cálculos foi 298,15K e a concentração iônica como igual a 150mM de NaCl. *A posteriori* são apresentados resultados para 309,65K, valor este correspondente, aproximadamente, à temperatura fisiológica.

A seguir estão os resultados obtidos, a começar pelos gráficos de grau de ionização *versus* pH (figuras de 2.1.1 a 2.1.7).

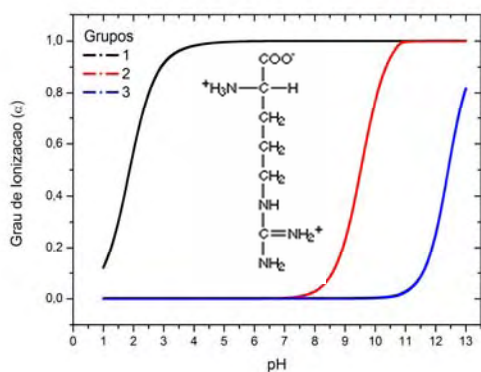


Figura 6.1.1: Arginina (Arg). Grau de Ionização ( $\alpha$ ) versus pH.

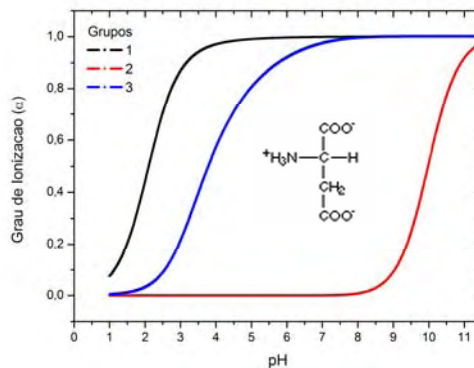


Figura 6.1.2: Ácido Aspártico (Asp). Grau de Ionização ( $\alpha$ ) versus pH.

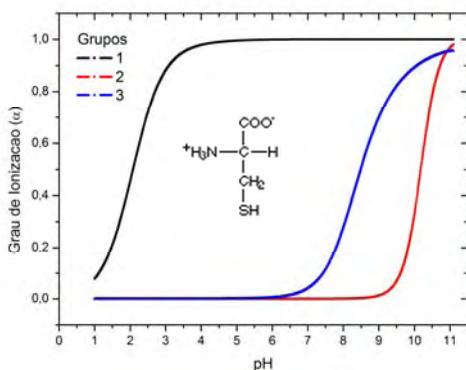


Figura 6.1.3: Cisteína (Cys). Grau de Ionização ( $\alpha$ ) versus pH.

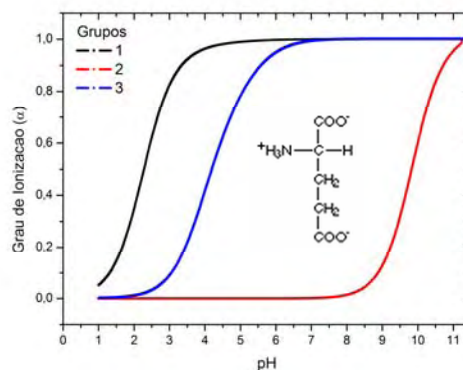


Figura 6.1.4: Ácido Glutâmico (Glu). Grau de Ionização ( $\alpha$ ) versus pH.

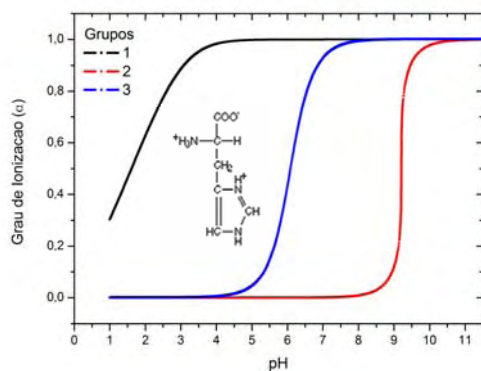


Figura 6.1.5: Histidina (His). Grau de Ionização ( $\alpha$ ) versus pH.

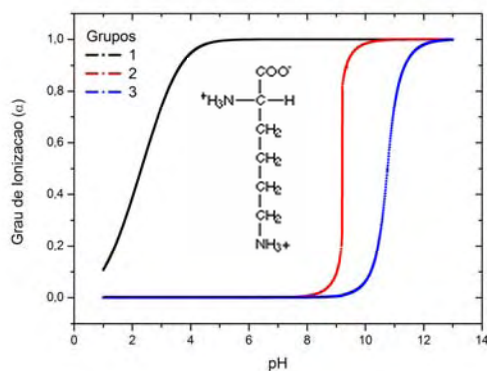


Figura 6.1.6: Lisina (Lys). Grau de Ionização ( $\alpha$ ) versus pH.

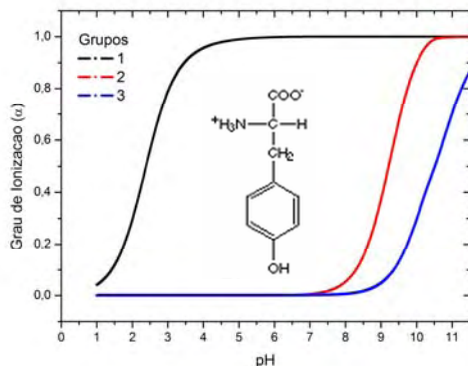


Figura 6.1.7: Tirosina (Tyr). Grau de Ionização ( $\alpha$ ) versus pH.

As curvas dos gráficos mostrados nas figuras de 6.1.1 a 6.1.7 permitem acompanhar, considerando o pH, o processo de desprotonação dos grupos ionizáveis. Como os valores de pK intrínseco dos grupos ionizáveis da cadeia lateral de Asp e Glu são muito parecidos, as curvas de grau de ionização destes aminoácidos também são semelhantes. Devido aos grupos ionizáveis das cadeias laterais de Arg e Lys terem sido tratados como bases de elevado pK intrínseco (12,65 e 10,63, respectivamente) e pelo potencial exercido sob suas superfícies ser relativamente baixo por causa da grande distância entre tais grupos e os restantes, os grupos 3 iniciam o processo de desprotonação (curvas em azul) em elevados valores de pH. Este mesmo argumento explica o comportamento do grupo 3 da tirosina.

Nas curvas apresentadas, as quais se caracterizam pela continuidade, têm-se os valores médios para o grau de ionização ( $\alpha$ ), pois os aminoácidos estudados são tratados como aminoácidos de ensemble.

Para o cálculo dos valores de pK\* lançou-se mão da equação de Henderson-Hasselbach:

$$pH = pK_i^* + \log_{10} \left( \frac{\alpha}{1-\alpha} \right) \quad (6.1)$$

a qual informa que, quando o pH da solução é igual ao pK\* do grupo ionizável, este possui  $\alpha$  igual a 0,5.

No programa computacional desenvolvido, o cálculo do grau de ionização é feito em intervalos de 0,01 unidades de pH. Realizado isto, buscou-se o ponto onde o  $\alpha$  de cada grupo ionizável é 0,5 e, encontrados estes pontos, utilizou-se a equação (6.1)

para se determinar os valores de  $pK^*$  de todos os grupos ionizáveis dos aminoácidos estudados. Estes valores estão transcritos na tabela 6.1, bem como os resultados experimentais<sup>9</sup> para comparação.

Tabela 6.1: Valores de  $pK^*$  calculados pelo modelo, provenientes da literatura. Solução com 150mM de NaCl e 25°C.  $\Delta$  é diferença entre os valores de  $pK^*$  calculados e os provenientes da literatura.

Aminoácidos	$pK^*$	Calculado	Literatura	$\Delta$
Ácido Aspártico (Asp)	$pK_1^*$	2,14	2,02 <sup>9</sup>	0,08
	$pK_2^*$	9,98	10,03 <sup>9</sup>	0,05
	$pK_3^*$	3,70	3,66 <sup>9</sup>	0,04
Ácido Glutâmico (Glu)	$pK_1^*$	2,21	2,29 <sup>9</sup>	0,08
	$pK_2^*$	9,89	9,87 <sup>9</sup>	0,02
	$pK_3^*$	4,17	4,17 <sup>9</sup>	0,00
Cisteína (Cys)	$pK_1^*$	2,23	2,05 <sup>9</sup>	0,18
	$pK_2^*$	10,10	10,19 <sup>9</sup>	0,09
	$pK_3^*$	8,48	8,44 <sup>9</sup>	0,04
Tirosina (Tyr)	$pK_1^*$	2,26	2,33 <sup>9</sup>	0,07
	$pK_2^*$	9,32	9,15 <sup>9</sup>	0,17
	$pK_3^*$	10,05	10,57 <sup>9</sup>	0,52
Arginina (Arg)	$pK_1^*$	2,12	1,95 <sup>9</sup>	0,17
	$pK_2^*$	8,98	9,00 <sup>9</sup>	0,02
	$pK_3^*$	12,49	12,50 <sup>35</sup>	0,01
Histidina (His)	$pK_1^*$	1,60	1,54 <sup>9</sup>	0,06
	$pK_2^*$	9,20	9,18 <sup>9</sup>	0,02
	$pK_3^*$	6,44	6,07 <sup>9</sup>	0,37
Lisina (Lys)	$pK_1^*$	2,28	2,21 <sup>9</sup>	0,07
	$pK_2^*$	9,19	9,19 <sup>9</sup>	0,00
	$pK_3^*$	10,45	10,75 <sup>9</sup>	0,30

Pela análise de todos os valores da tabela 6.1, constata-se que os resultados fornecidos pelo modelo estão bem próximos daqueles obtidos experimentalmente por Henchoz<sup>9</sup> em 150mM de NaCl e 25°C. A diferença média entre ambas as abordagens está por volta de 2,7%. Esta concordância suporta a hipótese deste trabalho no qual se

admite que a interação eletrostática, exercida pelos grupos vizinhos e íons em solução, é o principal fator que influencia as constantes de ionização dos grupos ionizáveis<sup>63</sup>.

Os resultados mostram, também, que o tratamento dado aos grupos 1 e 2 (i.e., ácido e base ligados ao  $C_\alpha$  sempre modelados como ácido acético e metilamina), conferiu ao modelo eficiência e versatilidade, pois para distinguir cada aminoácido necessita-se apenas dos valores dos parâmetros do grupo 3. Esta mesma modelagem por semelhança estrutural é empregada por Makowski<sup>64</sup> no desenvolvimento de um campo de força para descrever a interação entre cadeias laterais ionizáveis.

Para detalhar o papel exercido pelos íons em solução, empregou-se a mesma metodologia desta seção para determinar quais valores de  $pK^*$  os grupos ionizáveis assumiam com o acréscimo da concentração iônica.

## 6.2 Efeito da Concentração Iônica Sobre os Valores de $pK^*$ Calculados:

O programa computacional desenvolvido foi implementado para permitir o cálculo do grau de ionização de todos os grupos ionizáveis quando variadas o pH e a concentração iônica. Toda vez que o grau de ionização de um dado grupo é igual a meio, o programa, por base na equação de Henderson-Hasselbach, gravava em um arquivo o pH e a concentração iônica. Os resultados de  $pK_1^*$ ,  $pK_2^*$  e  $pK_3^*$  para todos os aminoácidos estudados estão mostrados nos gráficos presentes nas figuras de 6.2.1 a 6.2.6.

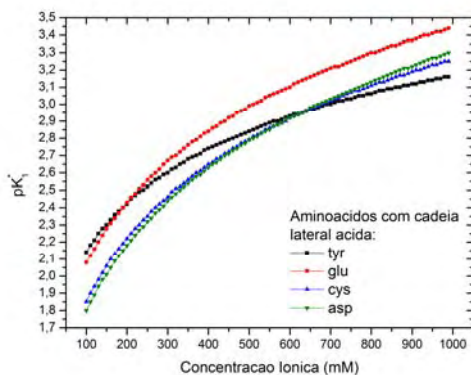


Figura 6.2.1: Variação de  $pK_1^*$  com a Conc. Iônica para Tyr, Glu, Cys e Asp.

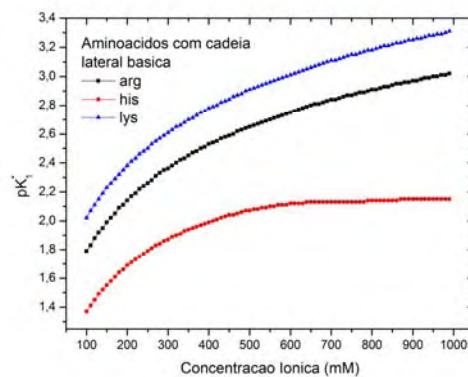


Figura 6.2.2: Variação de  $pK_1^*$  com a Conc. Iônica para Arg, His e Lys.

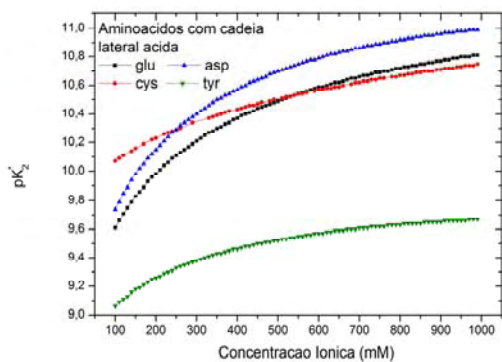


Figura 6.2.3: Variação de  $pK_2$  com a Conc. Iônica para Tyr, Glu, Cys e Asp.

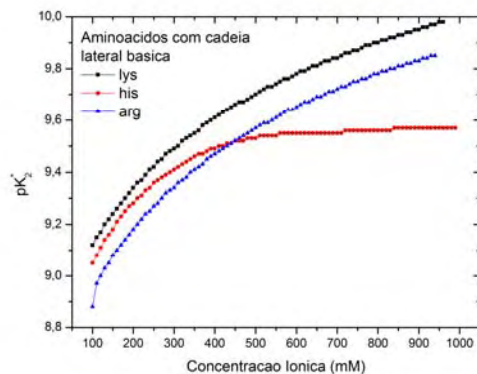


Figura 6.2.4: Variação de  $pK_2^*$  com a Conc. Iônica para Arg, Lys e His.

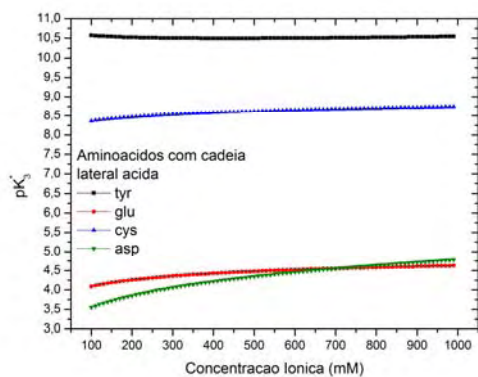


Figura 6.2.5: Variação de  $pK_3^*$  com a Conc. Iônica para Tyr, Glu, Cys e Asp.

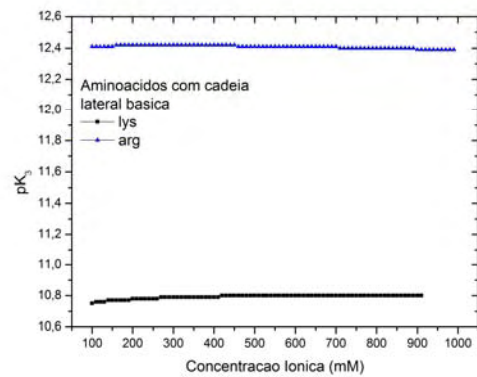


Figura 6.2.6: Variação de  $pK_3^*$  com a Conc. Iônica para Arg e Lys.

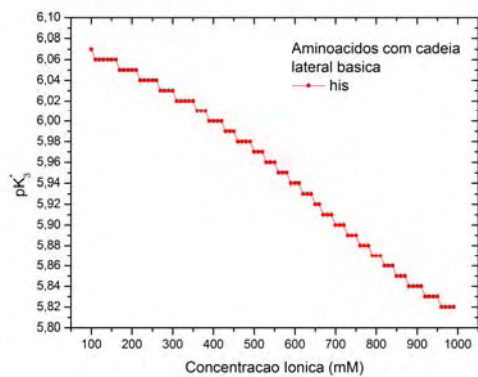


Figura 6.2.7: Variação de  $pK_3^*$  com a Conc. Iônica para His.



Nos gráficos de 6.2.1 a 6.2.4 percebe-se que, com o acréscimo da concentração iônica há, também, o aumento dos valores de  $pK_1^*$  e  $pK_2^*$ . Este comportamento se deve ao fato de que com mais íons em solução, maiores serão os efeitos de blindagem eletrostática, a qual fará com que a interação oriunda dos grupos vizinhos seja atenuada (veja a equação (3.17)). Quão mais atenuada for a interação, menor a influência de um grupo sobre os outros<sup>53</sup> e, por isso, os valores de  $pK^*$  crescem e tendem aos respectivos valores intrínsecos, ou seja, os valores que ocorrem quando cada grupo não está combinado em um aminoácido.

Algumas curvas apresentam "degraus", ou seja, valores de  $pK^*$  que se mantêm constantes mesmo quando variada a concentração iônica, pois no procedimento computacional adotado e devido à precisão das variáveis empregues, determinava-se o valor de  $pK^*$  de um dado grupo  $i$  quando seu grau de ionização era maior que 0,495 e menor que 0,505.

Novamente, as diferenças entre as curvas de cada aminoácido se devem ao potencial eletrostático médio que é influenciado pelas correspondentes constantes dielétricas efetivas. Este parâmetro, juntamente com a distância do grupo aos vizinhos, explica a baixíssima variação de  $pK_3^*$  do grupo ionizável da cadeia lateral da Tyr, Arg e Lys, conforme pode ser visualizado nas figuras 6.2.5 e 6.2.6.

O intrigante comportamento decrescente observado para o  $pK_3^*$  da histidina e da arginina (vide figuras 6.2.6 e 6.2.7, respectivamente) se deve aos efeitos de reestruturação resultantes do processo de desprotonação dos grupos ionizáveis. Quando isto ocorre, a conformação da molécula se altera e as distâncias entre os grupos interagentes também é modificada. Além disso, para a histidina, a  $CD_{ef}$  do meio ácido-cadeia é cerca de três vezes menor do que a do meio base-cadeia assim, na superfície do grupo 3 a interação predominante é a exercida pelo grupo ácido, cuja carga é negativa. Por causa disto, o potencial nas vizinhanças do grupo da cadeia lateral ( $\phi_3$ ) é aquele que, dentre todos os aminoácidos básicos, alcança os valores mais negativos, principalmente com o acréscimo de íons que atenuarão as interações com ambos os vizinhos. Da expressão (3.3) pode-se obter que:

$$pK_3^* = pK_3 + \phi_3 \log_{10} e \quad (6.2)$$

assim, quanto mais negativo o valor de  $\phi_3$ , menor o valor de  $pK_3^*$ .

## 6.3 Carga Elétrica Média Total em Condições Fisiológicas.

Dando seguimento a análise das propriedades físico-químicas dos aminoácidos estudados, segue uma análise sobre a carga elétrica média de tais moléculas. Tendo determinado o grau de ionização de cada um deles, a fenomenologia permite correlacioná-lo com a carga elétrica média por meio das seguintes expressões:

$$Q_1 = -\alpha_1 e \quad (6.3)$$

$$Q_2 = (1 - \alpha_2) e \quad (6.4)$$

$$Q_3 = \begin{cases} -\alpha_3 e \rightarrow \text{Se o grupo da cadeia lateral for ácido} \\ (1 - \alpha_3) e \rightarrow \text{Se o grupo da cadeia lateral for básico} \end{cases} \quad (6.5)$$

onde  $Q_1$  é o valor médio da carga do grupo 1, ácido acético ligado ao  $C_\alpha$ ;  $Q_2$  a do grupo 2, metilamina ligada ao  $C_\alpha$ ; e  $Q_3$  a do grupo ionizável da cadeia lateral. Seguem os resultados obtidos com 150mM de NaCl. A carga elétrica média total corresponde à soma das três cargas médias parciais  $Q_1$ ,  $Q_2$  e  $Q_3$ . Os valores de carga elétrica média são inicialmente apresentados na forma de gráficos contidos nas figuras de 6.3.1 a 6.3.8. Nas figuras de 6.3.2 a 6.3.8 a temperatura utilizada foi de 25°C. Para a obtenção dos dados da figura 6.3.1 utilizou-se a temperatura de 36,5°C e, conseqüentemente, os valores de  $CD_{ef}$  a ela correspondentes.

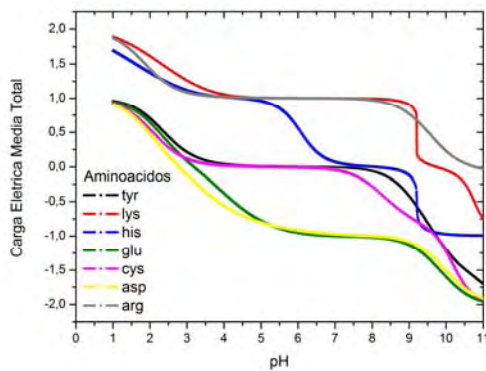


Figura 6.3.1: Carga Elétrica Média Total versus pH. 36,5°C

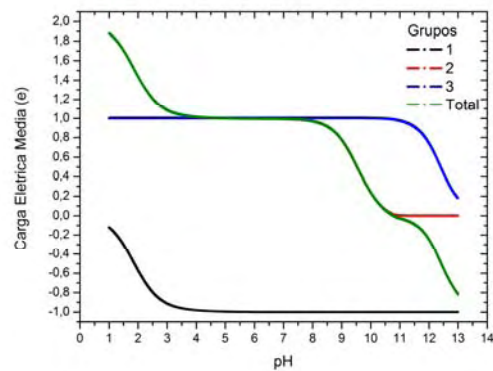


Figura 6.3.2: Carga Elétrica Média versus pH para a Arginina. 25°C

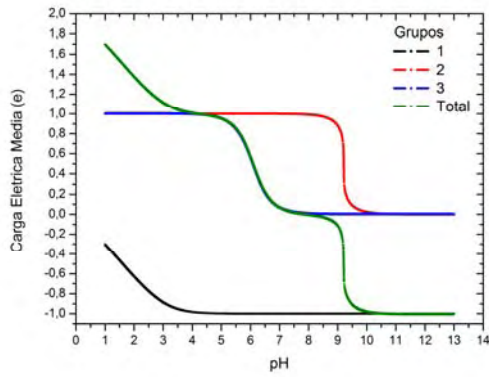


Figura 6.3.3: Carga Elétrica Média versus pH para a Histidina. 25°C

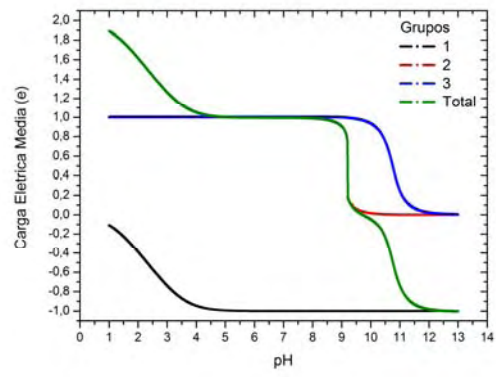


Figura 6.3.4: Carga Elétrica Média versus pH para a Lisina. 25°C

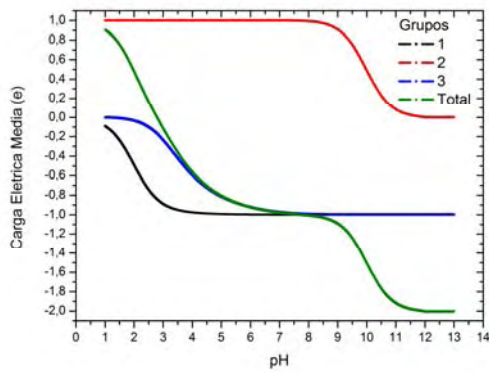


Figura 6.3.5: Carga Elétrica Média versus pH para o Ácido Aspártico. 25°C

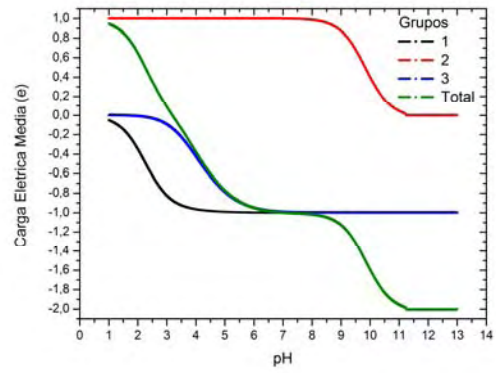


Figura 6.3.6: Carga Elétrica Média versus pH para o Ácido Glutâmico. 25°C

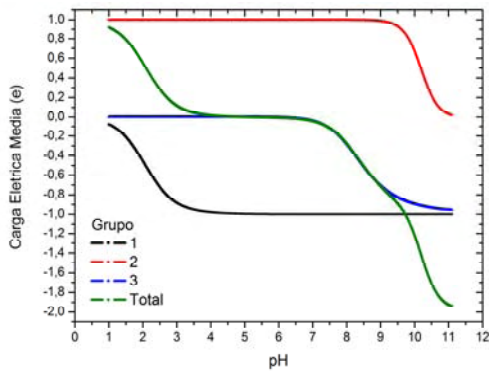


Figura 6.3.7: Carga Elétrica Média versus pH para a Cisteína. 25°C

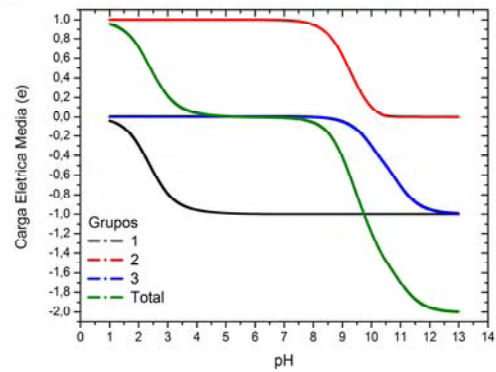


Figura 6.3.8: Carga Elétrica Média versus pH para a Tirosina. 25°C

Nos gráficos de 6.3.2 a 6.3.6 visualiza-se um comportamento que está diretamente ligado à variação do potencial eletrostático nas vizinhanças dos grupos ionizáveis, pois com o aumento de pH os grupos se desprotonam e, conseqüentemente, os potenciais se tornam menos positivos. É digno de nota que essas curvas são contínuas, pois correspondem à carga elétrica média de um aminoácido que tem suas propriedades descritas em termos de valores médios.

Pelo fato de os valores de  $pK^*$  dos grupos básicos serem elevados, percebe-se que os aminoácidos Arginina e Lisina permanecem com carga elétrica média total positiva até pH próximo a 10. A Histidina, apesar de ter um grupo básico em sua cadeia lateral, assume valores mais negativos de carga total em pH 7,5, pois o  $pK^*$  do referido grupo (i.e., o  $pK_3^*$ ) é baixo, sendo igual a 6,05 segundo os valores da tabela 6.1.

A figura 6.3.1 representa a variação dessas cargas com o pH e foi feita para dar uma visão geral do processo de ionização. No entanto, o objetivo principal desta figura, a única para elaborada em 36,5°C, é permitir uma análise dessas cargas elétricas médias quando cada aminoácido se encontrava em condições próximas às condições fisiológicas de temperatura e concentração iônica. Para uma análise mais detalhada dos valores assumidos nesta situação, construíram-se os gráficos apresentados nas figuras 6.3.9, 6.3.10 e 6.3.11.

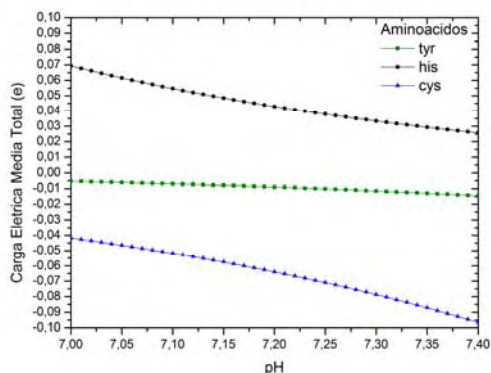


Figura 6.3.9: Carga Elétrica Média Total versus pH para Tyr, His e Cys. 150mM de NaCl e 36,5°C

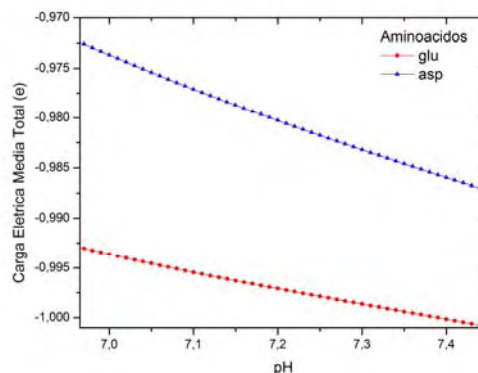
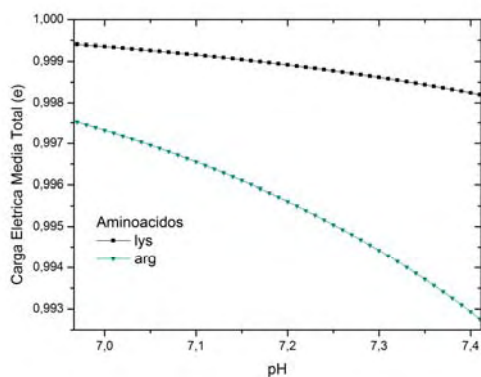


Figura 6.3.10: Carga Elétrica Média Total versus pH para Glu e Asp. 150mM de NaCl e 36,5°C



*Figura 6.3.11: Carga Elétrica Média Total versus pH para Lys e Arg. 150mM de NaCl e 36,5°C*

Para os aminoácidos histidina, cisteína e tirosina percebe-se que, devido aos valores de  $pK^*$  dos seus grupos ionizáveis, a carga elétrica média total é da ordem de  $10^{-2}$ . Assim, considerando as aproximações envolvidas no desenvolvimento do modelo, é possível afirmar que em condições fisiológicas estes aminoácidos tem carga total nula.

Arginina e lisina apresentam, no intervalo de pH analisado, carga elétrica média total praticamente igual a um, pois os seus valores de  $pK_3^*$  são superiores a 10,7, o que faz com que a cadeia lateral esteja ainda completamente ionizada. Os ácidos aspártico e glutâmico, cujos valores tanto de  $pK_1^*$  quanto de  $pK_3^*$  não superam 4,5, possuem seus ácidos completamente desprotonados e, embora haja uma carga positiva na base, como tais aminoácidos tem dois grupos ácidos cada, a carga elétrica total é, na referida faixa de pH, praticamente igual a -1. Para se analisar como se dá a variação da carga elétrica média total com o acréscimo da concentração iônica, construiu-se tabela 6.2.

Tabela 6.2: Valores de Carga Elétrica Média Total em 36,5°C.

Aminoácido	pH	Concentração iônica (mM)							Carga elétrica média total do aminoácido (e)
		10	20	50	100	150	160	170	
Arginina	7,0	0,992	0,994	0,997	0,998	0,999	0,999	0,999	
	7,2	0,987	0,991	0,995	0,997	<b>0,998</b>	0,998	0,998	
	7,4	0,980	0,985	0,992	0,995	0,997	0,997	0,997	
Ácido Aspártico	7,0	-0,990	-0,988	-0,984	-0,978	-0,973	-0,972	-0,972	
	7,2	-0,996	-0,993	-0,988	-0,984	<b>-0,980</b>	-0,979	-0,978	
	7,4	-1,004	-0,999	-0,993	-0,989	-0,986	-0,985	-0,984	
Ácido Glutâmico	7,0	-1,006	-1,002	-0,997	-0,995	-0,993	-0,993	-0,993	
	7,2	-1,017	-1,010	-1,002	-0,998	<b>-0,997</b>	-0,997	-0,996	
	7,4	-1,034	-1,021	-1,008	-1,002	-1,000	-0,999	-0,999	
Cisteína	7,0	-0,068	-0,063	-0,053	-0,045	-0,040	-0,039	-0,038	
	7,2	-0,103	-0,095	-0,081	-0,069	<b>-0,061</b>	-0,060	-0,059	
	7,4	-0,153	-0,141	-0,122	-0,104	-0,094	-0,092	-0,090	
Histidina	7,0	0,056	0,005	0,060	0,062	0,062	0,062	0,062	
	7,2	0,031	0,033	0,036	0,037	<b>0,038</b>	0,038	0,038	
	7,4	0,013	0,016	0,019	0,022	0,227	0,223	0,061	
Lisina	7,0	0,996	0,997	0,998	0,999	0,999	0,999	1,000	
	7,2	0,995	0,996	0,997	0,998	<b>0,998</b>	0,998	0,999	
	7,4	0,991	0,993	0,995	0,997	0,998	0,998	0,998	
Tirosina	7,0	-0,022	-0,017	-0,011	-0,007	-0,006	-0,005	-0,005	
	7,2	-0,036	-0,028	-0,019	-0,012	<b>-0,010</b>	-0,009	-0,009	
	7,4	-0,057	-0,046	-0,030	-0,020	-0,016	-0,015	-0,015	

Com o acréscimo da concentração iônica nota-se que as cargas elétricas totais se tornam mais positivas. Isto é explicado pelo resultado da seção 6.2 onde foi visto que com o aumento da concentração iônica os valores de  $pK^*$  tendem aos valores de  $pK$  intrínseco. Com os valores de  $pK_1^*$  aumentando, a completa ionização dos grupos ácidos é “retardada” e isto acarreta uma diminuição da contribuição negativa para a carga elétrica média total.

Na tabela 6.2 estão, em negrito, os valores calculados em condições próximas às fisiológicas. Estes valores permitem visualizar a participação da cadeia lateral na determinação da carga elétrica média total de cada aminoácido. A histidina, apesar de ser um dos aminoácidos com cadeia lateral básica, é dentre estes o que apresenta o menor valor de carga elétrica média total. Isto se deve ao seu  $pK_3^*$  ser em torno de seis, o que lhe confere um caráter tamponante.<sup>31</sup>

Devido à concordância obtida entre os valores de  $pK^*$  calculados pelo modelo teórico e aqueles provenientes de outros estudos, conforme mostrado na tabela 6.1, decidiu-se expandir o modelo desenvolvido e aplicá-lo a outras moléculas com três grupos ionizáveis, tais como peptídeos e proteínas. Os capítulos seguintes mostram as considerações para tanto adotadas e os resultados obtidos.

# Capítulo 7




## Pentapeptídeos: Modelo e Resultados

Neste capítulo são apresentados a expansão modelo teórico, necessária para o cálculo dos valores de  $pK^*$  dos grupos ionizáveis e da constante dielétrica efetiva de pentapeptídeos, e os resultados assim obtidos.

### 7.1 Adaptações ao Estudo de Pentapeptídeos

Nos capítulos 2 e 3 foram apresentados as aproximações e o desenvolvimento matemático para a obtenção do potencial eletrostático na superfície de cada grupo ionizável. Tal potencial leva em consideração a interação devida à carga do próprio grupo, dos grupos vizinhos e dos íons em solução. Devido às considerações adotadas e à forma generalizada como tal modelo foi concebido para os aminoácidos, o estudo de peptídeos com três grupos ionizáveis necessitou apenas das dimensões características destas moléculas e das distâncias entre os grupos, os quais permaneceram sendo modelados como esferas. A disposição dos grupos estudados segue a figura 3.1.

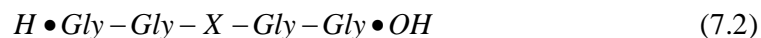
Os grupos e os índices das suas grandezas correspondentes seguem a notação:

-  1 para o grupo ácido, de raio  $r_a$ , na extremidade C-terminal;
-  2 para o grupo básico, de raio  $r_b$ , na extremidade N-terminal;
-  3 para o grupo ionizável, de raio  $r_c$ , do resíduo de aminoácido a ser estudado;

Todos os parâmetros espaciais, ou seja, raio dos grupos e distância entre eles, foram novamente calculados pelo programa *PyMol*<sup>52</sup> a partir de arquivos PDB. Os grupos C e N terminais permaneceram sendo tratados como ácido acético e metilamina, respectivamente com  $pK$  intrínsecos de 4,76 e 10,65. Os valores de  $pK$  intrínseco dos grupos ionizáveis das cadeias laterais são mostrados na tabela 2.2.



Na seqüência deste trabalho foram estudados pentapeptídeos com as seguintes estruturas:



A representação (7.1) foi chamada de Ala-pentapeptídeos e a (7.2) de Gly-pentapeptídeos. Em ambas, é na posição X onde se encontra os resíduos (chamados de X-resíduos) dos aminoácidos com cadeia lateral ionizável: Asp, Glu, His, Cys, Tyr ou Lys.

O cálculo do valor de  $pK^*$  de cada grupo e, conseqüente, grau de ionização, foi feito por meio de expressões análogas a (3.1), (3.2) e (3.3):

$$K_1^* = K_1 \exp\left(\frac{e\Psi_1}{kT}\right) \quad (7.3)$$

$$K_2^* = K_2 \exp\left(\frac{-e\Psi_2}{kT}\right) \quad (7.4)$$

$$K_3^* = K_3 \exp\left(\frac{-e\Psi_3}{kT}\right) \quad (7.5)$$

Sendo que a expressão (7.5) foi escrita para um grupo ionizável básico. Se o grupo ionizável da cadeia lateral fosse ácido, ter-se-ia um argumento positivo na exponencial.

Nos peptídeos estudados, os grupos nas extremidades C e N terminais estão aproximadamente dispostos linearmente. Graças a isto e seguindo o mesmo argumento da disposição geométrica dos grupos adotado nos capítulos 4 e 5, a constante dielétrica efetiva para o meio entre os supracitados grupos (região 1-2) foi tomada como igual a  $\epsilon_1$ , ou seja, tal região foi tratada como sendo constituído unicamente por hidrocarbonetos de  $CD_{ef}$  igual a 2. Nas outras regiões, nas quais há a supremacia da solução eletrolítica, o valor da  $CD_{ef}$  é fornecido pela expressão (4.1), sendo que, para 25°C e 150mM de NaCl, ele corresponde a 79,95.

## 7.2 Resultados: $pK^*$ de Gly e Ala Pentapeptídeos

Na tabela 7.1 são apresentados os valores de  $pK^*$  obtidos a partir das considerações descritas neste capítulo e, para verificação da validade destas, os resultados obtidos pelo modelo teórico são confrontados com os obtidos

experimentalmente por outros autores<sup>65-67</sup>. Construiu-se, também, um gráfico de correlação, presente na figura 7.1.

Tabela 7.1: Valores de  $pK^*$  calculados pelo modelo e aqueles provenientes da literatura nas correspondentes condições de temperatura e concentração iônica.

	X-Resíduo	$pK^*$			
		Calculado na Respec. Temp.	Gurd <sup>65</sup> 100mM KCl e 26°C	Thurkill <sup>66</sup> 100mM KCl e 25°C	Creighton <sup>67</sup>
Ala Pentapeptídeos	Asp	3,97	-	3,67	3,9-4,0
	Glu	4,2	-	4,25	4,3-4,5
	His	6,71	-	6,54	6,0-7,0
	Cys	9,03	-	8,55	9,0-9,5
	Tyr	10,25	-	9,84	10,0-10,30
	Lys	10,35	-	10,40	10,4-11,1
Gly Pentapeptídeos	Asp	3,81	3,9	-	3,9-4,0
	Glu	3,99	4,2	-	4,3-4,5
	His	6,70	6,8	-	6,0-7,0
	Tyr	10,25	10,0 (33°C)	-	10,0-10,30
	Lys	10,33	10,5	-	10,4 - 11,1

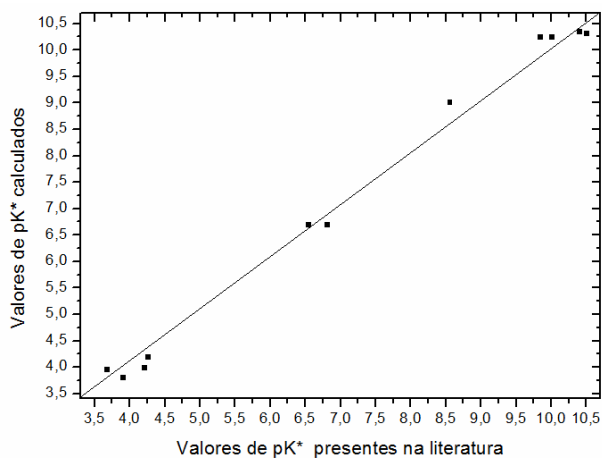


Figura 7.1: Gráfico de correlação entre valores de  $pK^*$  calculados e os presentes na literatura: Gurd<sup>65</sup> para Gly-pentapeptídeos e Thurkill<sup>66</sup> para Ala-pentapeptídeos.

Pela análise da tabela 7.1 percebe-se que os valores gerados pelo modelo se encontram próximos aos determinados pelos outros três autores, sendo que os maiores desvios, em ambos os tipos de pentapeptídeos, ocorrem para resíduos de tirosina. Uma possível causa disto é que o processo de desprotonação dos grupos ionizáveis faz com que haja uma reestruturação da conformacional dos aminoácidos, o que acarreta na alteração das distâncias entre os grupos e, conseqüentemente, nos valores da constante

dielétrica efetiva. No caso da tirosina, é provável que a hidroxila fique mais próxima do grupo básico e, assim, esse fortalecimento da contribuição decorrente de uma carga positiva produz a alteração do  $pK^*$ .

Praticamente todos os valores de  $pK^*$  calculados encontram-se na faixa apresentada por Creighton<sup>67</sup>, os quais correspondem aos valores de  $pK^*$  de resíduos localizados na superfície de proteínas. Pelo fato de nesta região os resíduos carregados se encontrarem plenamente expostos ao solvente, a consideração de que as interações sofridas pelo grupo ionizável do X-resíduo sejam atenuadas por uma constante dielétrica efetiva igual a do solvente é plenamente justificada.

O gráfico de correlação (figura 7.1) evidencia a concordância entre os resultados obtidos teoricamente e aqueles fornecidos por outros trabalhos<sup>65-67</sup>. Uma medida quantitativa que explicita a concordância e eficácia da abordagem desenvolvida é dada pelo Fator de Correlação de Pearson<sup>68</sup>, denotado por  $\rho_{FCP}$ , e que para os dados do referido gráfico vale 0,9966.

Outro fato importante é que, quando se comparam os valores de  $pK^*$  calculados com os valores de  $pK$  intrínseco atribuído a cada grupo, constata-se que a diferença entre tais é menor no caso de pentapeptídeos do que nos aminoácidos. Isto se dá, pois os grupos vizinhos que, juntamente com os íons em solução, são os responsáveis pelas interações eletrostáticas adicionais que alteram o valor de  $pK^*$ , se encontram mais afastados no caso dos pentapeptídeos (tenha em mente que o potencial eletrostático calculado pelo modelo decai exponencialmente com a distância, vide expressão (3.41)).

Os resultados deste capítulo indicam que o modelo desenvolvido pode ser caracterizado como generalizado e versátil, sendo capaz de produzir resultados satisfatórios tanto para aminoácidos quanto para pentapeptídeos em variadas condições de temperatura e concentração iônica. Estes bons indicadores motivaram o estudo de resíduos carregados em proteínas, o que será discutido no capítulo a seguir.

# Capítulo 8

## Aplicações aos resíduos carregados da proteína *Staphylococcal Nuclease*

Devido à variedade de trabalhos, tanto experimentais quanto analítico-computacionais, relacionados ao estudo de propriedades físico-químicas de resíduos carregados em proteínas, este trabalho foi estendido para tratar de resíduos da proteína *Staphylococcal Nuclease (SNase)*, a qual é amplamente estudada por ser uma proteína consideravelmente “estável”, ou seja, mesmo com a alteração de alguns de seus resíduos ela continua se enovelando, o que permite que sejam analisados os efeitos de troca de um resíduo apolar por um polar, por exemplo<sup>20,21,61,69,70</sup>.

Neste oitavo capítulo são mostradas as considerações acrescentadas ao desenvolvimento matemático, presente nos capítulos 2 e 3, necessárias a uma maior generalização do modelo. Os resultados assim obtidos são, quando possível, comparados com aqueles reportados por outros autores.

### 8.1 Interação entre resíduos carregados

Após todo o desenvolvimento matemático apresentado em detalhes no capítulo 3 obtiveram-se expressões para a interação entre três grupos ionizáveis sendo que o potencial eletrostático nas vizinhanças de um grupo  $i$  era decorrente das interações devidas à carga do próprio grupo, dos íons em solução e dos dois vizinhos carregados mais próximos o que, nos aminoácidos e pentapeptídeos analisados, correspondiam, respectivamente, ao grupo ácido acético e metilamina e extremidades C e N terminais. Agora, em proteínas, as quais costumam ter vários grupos carregados, trabalha-se com uma aproximação semelhante à de *primeiros vizinhos*, ou seja, são consideradas, inicialmente, sobre o grupo  $i$ , as interações decorrentes dos dois grupos ionizáveis mais próximos. Isto será estudado em mais detalhes na seção 8.2.

De forma análoga à obtida nos capítulos anteriores, a expressão para o potencial eletrostático nas vizinhanças do grupo  $i$  é:

$$\varphi_i(r) = \frac{a_i \exp(-r_i / \lambda)}{(r_i / \lambda)} + \frac{\sinh(r_i / \lambda)}{(r_i / \lambda)} \sum_{j=2}^3 \frac{a_j \exp(-R_{i-j} / \lambda)}{(R_{i-j} / \lambda)} \quad (8.1)$$

Na qual o índice  $j$  representa os vizinhos considerados e os símbolos  $a_i$  e  $a_j$  denotam as constantes de integração cuja determinação se dá pelo método das operações contratantes<sup>58</sup>, vide seção 3.2. Para uma situação na qual fossem levados em conta quatro grupos vizinhos, o somatório da equação (8.1) se daria até 5, pois para um conjunto de  $N$  resíduos tem-se  $N-1$  interações cruzadas.

A aplicação do modelo a proteínas difere dos casos anteriores pois agora os grupos carregados considerados não se encontram nas extremidades da macromolécula e nem distribuídos de forma simétrica tal como nos Gly e Ala-pentaptideos. Devido a isto, trabalharam-se apenas com as cadeias laterais ionizáveis e os valores de pK intrínseco a cada uma delas atribuído continuou seguindo o critério de semelhança estrutural (vide tabelas 2.2 e 7.1).

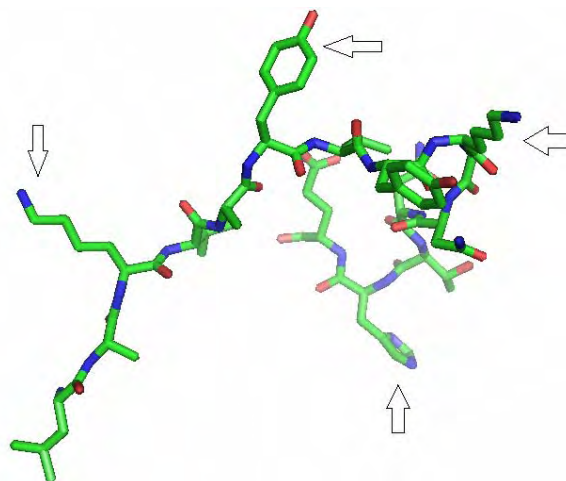


Figura 8.1: Representação de parte da proteína SNase. Somente as cadeias laterais ionizáveis, algumas destas indicadas por setas, serão as consideradas pelo modelo.

Seguindo a aproximação de primeiros vizinhos, todos os grupos considerados (i.e., as cadeias ionizáveis) foram modeladas como esferas sendo que o raio destas e as distancias entre elas foram novamente calculados pelo programa *PyMol*<sup>52</sup> com base em arquivos PDB. Em tal tratamento, tais esferas quando carregadas apresentavam carga pontual central. Como buscou-se comparar os resultados do modelo com aqueles

reportados em outros trabalhos, usaram-se os os arquivos PDB fornecido pelas correspondentes referências. A variedade deste arquivos por elas apresetados se deve à existência de muitas “variantes” da *SNase* pois esta, como mencionado anteriormente, é uma proteína caracterizada pelo fato de que, mesmo quando trocados alguns de seus resíduos, ela continua se enovelando<sup>69</sup>. Os valores de  $pK_i^*$  calculados estão na tabela 8.1.

O cálculo da constante dielétrica efetiva das regiões entre cada par de grupos ionizáveis se deu da mesma forma que a descrita no capítulo 4. As condições de temperatura e concentração iônica usadas foram as mesmas que as apresentadas pelas fontes com as quais se comparam os resultados do modelo.

## 8.2 Estudo sobre a distância de *cut-off*. Resultados para resíduos de Lys e Glu

No cálculo do potencial eletrostático nas vizinhanças de um grupo  $i$  foram consideradas as interações devidas aos íons em solução, computadas implicitamente no comprimento de Debye, e as exercidas pelos grupos vizinhos. Em todas as partes anteriores deste trabalho foram considerados apenas dois vizinhos, pois se lidou com aminoácidos e pentapeptídeos compostos por apenas três grupos ionizáveis. Agora, no estudo de resíduos ionizáveis em proteínas, buscou-se expandir o modelo e estabelecer quantos grupos devem ser levados em consideração para determinar o valor de  $pK_i^*$  de um dado resíduo  $i$ . Para tanto, foram realizados cálculos quando considerados dois ou quatro vizinhos:

$$\varphi_i(r) = \frac{a_i \exp(-r_i / \lambda)}{(r_i / \lambda)} + \frac{\sinh(r_i / \lambda)}{(r_i / \lambda)} \sum_{j=2}^3 \frac{a_j \exp(-R_{i-j} / \lambda)}{(R_{i-j} / \lambda)} \rightarrow 2 \text{ vizinhos} \quad (8.2)$$

$$\varphi_i(r) = \frac{a_i \exp(-r_i / \lambda)}{(r_i / \lambda)} + \frac{\sinh(r_i / \lambda)}{(r_i / \lambda)} \sum_{j=2}^5 \frac{a_j \exp(-R_{i-j} / \lambda)}{(R_{i-j} / \lambda)} \rightarrow 4 \text{ vizinhos} \quad (8.3)$$

Os cálculos realizados pela equação (8.2) e os resultados por ela obtidos são designados pelo símbolo **3** pois há um total de três grupos interagentes. Seguindo este mesmo raciocínio, o que for obtido pela equação (8.3) é designado pelo símbolo **5**. A tabela 8.1 fornece os valores de  $pK_i^*$  calculados para os resíduos de Lys e Glu na *SNase*. Nesta tabela também são apresentados os valores de constante dielétrica efetiva determinados e diferenças percentuais entre os resultados do modelo e aqueles

fornecidos por Isom<sup>20</sup>, para resíduos de lisina, e Castañeda<sup>61</sup>, para resíduos de glutamato. Outra informação importantíssima é a separação máxima entre os resíduos considerados. Tal dado é crucial para a determinação da distância de *cut-off*, que é a distância máxima entre os resíduos cuja interação é considerada e que reproduza resultados *concordantes* com a literatura. Construiu-se, também, a figura 8.2 na qual há um gráfico de correlação para os valores de  $pK_i^*$  obtidos quando considerados dois ou quatro grupos vizinhos e para várias distâncias de *cut-off*.

Tabela 8.1: Resultados para a Constante Dielétrica Efetiva ( $CD_{ef}$ ) e  $pK_i^*$  dos resíduos na proteína *Staphylococcal Nuclease (SNase)*. Comparação com valores fornecidos por Isom<sup>20</sup> e Castañeda<sup>61</sup>.

Resíduo <i>i</i>	$CD_{ef}$	$pK_i^*$ 3 <sup>§</sup> grupos	$pK_i^*$ 5 <sup>§</sup> grupos	$pK_i^*$ ref <sup>20,61</sup>	$\Delta^\ddagger$ 3 <sup>§</sup> - 5 <sup>§</sup>	Max R <sup>‡</sup>	
						3 <sup>§</sup>	5 <sup>§</sup>
Lys A132K	75	10,65	<u>10,50</u>	10,40 <sup>20</sup>	0,25-0,10	13,14	15,28
Glu 101 (0,1M)	20	4,08	<u>3,84</u>	3,81 <sup>61</sup>	0,27-0,03	13,75	18,04
Glu 101 (1M)	20	4,67	<u>4,49</u>	4,41 <sup>61</sup>	0,26-0,08	13,75	18,04
Lys A109K	24	10,82	10,62	9,20 <sup>20</sup>	1,62-1,42	23,0	32,15
Glu 135 (0,1M)	20	3,84	<u>3,81</u>	3,76 <sup>61</sup>	0,08-0,05	12,4	15,85
Glu 135 (1,0M)	20	<u>3,96</u>	4,40	4,08 <sup>61</sup>	0,12-0,32	12,4	15,85
Glu 10 (0,1M)	20	<u>2,84</u>	3,2	2,82 <sup>61</sup>	0,02-0,38	10,8	27,01
Glu 10 (1M)	20	<u>3,70</u>	4,02	3,45 <sup>61</sup>	0,25-0,57	10,8	27,01
Glu 129 (0,1M)	79	4,55	<u>3,85</u>	3,75 <sup>61</sup>	0,80-0,10	13,85	14,01
Glu 129 (1M)	79	5,12	<u>4,61</u>	4,32 <sup>61</sup>	0,8-0,29	13,85	14,01

§ Total de grupos considerados no cálculo do potencial eletrostático médio nas vizinhanças do resíduo *i*.

† $\Delta$  é a diferença entre os valores de  $pK^*$  calculados pelo modelo e aqueles reportados pelas referências. Diferença quando considerados 3 ou 5 grupos localizados dentro de uma distância de *cut-off* indicada na coluna Max R.

‡ Distância entre o resíduo *i* e o grupo ionizável mais distante cuja interação foi levada em conta para o cálculo do valor de  $pK_i^*$  quando considerados 3 ou 5 grupos.

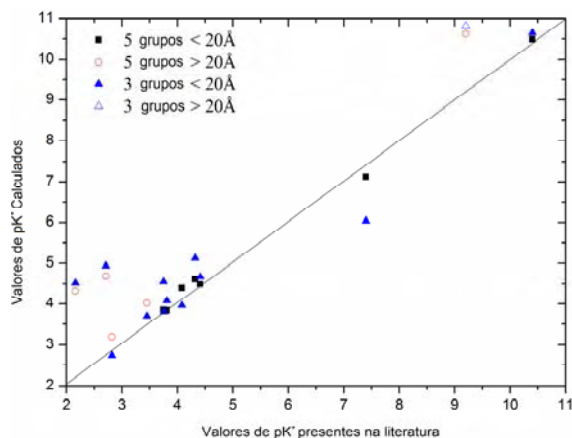


Figura 8.2: Correlação entre valores de  $pK^*$  calculados para os resíduos de SNase e comparação de distâncias de cut-off. Referências: Isom<sup>20</sup> para resíduos de Lys; Castañeda<sup>61</sup> para resíduos de Glu.

Os valores de constante dielétrica efetiva calculados para grande parte dos resíduos de glutamato e para Lys A109K estão na faixa de 20 a 24, o que concorda com o fato destes resíduos estarem localizados em posições internas (*buried residues*) da proteína. Tais resultados vão de encontro com os utilizados pelas correspondentes fontes. Valores de  $CD_{ef}$  superiores a 75 indicam que o resíduo está exposto ao solvente.

Ambos, tabela 8.1 e figura 8.2, evidenciam que a consideração de interações exercidas por grupos, distantes do resíduo  $i$  mais que 20Å, acarreta um aumento na diferença entre os valores de  $pK_i^*$  calculados e os oriundos de outros trabalhos. A explicação para isto é que para distâncias maiores que esta, a qual corresponde a aproximadamente dois comprimentos de Debye, efeitos negligenciados pelo modelo começam a se tornar mais importantes e o potencial eletrostático médio, assumido como potencial da força média, não é mais o principal determinante do grau de ionização do resíduo  $i$ . Na supracitada tabela, os valores sublinhados correspondem aos alcançados quando se trabalhou com resíduos distantes não mais que 20Å.

Os resultados obtidos para os resíduos de glutamato estão em melhor acordo com as referências do que os de lisina. Uma possível explicação para isto é a diferença entre os valores de  $CD_{ef}$  calculados, pois para Glu foi obtido para este parâmetro o mesmo valor que o reportado por Castañeda enquanto que para o resíduo Lys 132A obteve-se  $CD_{ef}$  igual a 75, ligeiramente menor à utilizada por Isom (78,5). O resíduo Lys A109K apresenta grande discrepância, pois resíduos distantes mais que 20Å foram levados em conta tanto no cálculo de potencial eletrostático com três quanto com cinco grupos. Quando considerados apenas resíduos distantes menos que 20Å, i.e., só um



vizinho, calculou-se o valor de  $pK_i^*$  deste resíduo, valor este que se encontra mais próximo a 9,2 fornecido por Isom<sup>20</sup>.

O gráfico de correlação na figura 8.2 foi construído para uma melhor visualização do papel exercido pela consideração de um número maior de grupos interagentes. Quando computados o máximo número de vizinhos localizados "dentro" destes 20Å, obteve-se uma Fator de Correlação de Pearson,  $\rho_{FCP}$ , igual a 0,9976. Este resultado, muito próximo à unidade, evidencia que, para o cálculo do valor de  $pK_i^*$  de um resíduo  $i$ , o modelo é eficiente quando considerados o máximo número de grupos vizinhos (neste caso, quatro) com uma distância de *cut-off* de 20Å.

Este resultado é um forte indicativo de que a abordagem aqui desenvolvida pode ser eficientemente aplicada a análise de moléculas complexas e à implementação de campos de força em simulações computacionais já que o resultado desta seção evita a necessidade de serem computados *explicitamente* a interação devida aos íons e a uma grande quantidade de grupos para se determinar os valores de  $pK_i^*$ , uma vez que resíduos carregados localizados além da distância de *cut-off* interferem "negativamente" no valor deste parâmetro, pois para grandes distâncias fatores negligenciados pelo modelo começam a se tornar mais relevantes, tais como a re-estruturação da proteína ocasionada pelas desprotonações de seus resíduos, por exemplo.

O estabelecimento do valor de  $pK_i^*$  em ambas proteínas e peptídeos viabiliza a determinação da carga elétrica média de toda a macromolécula ou de uma dada região, tal como o sítio ativo, na qual estão os resíduos considerados. Uma das mais importantes características do modelo desenvolvido neste trabalho é a capacidade de realizar os cálculos necessários de maneira rápida e em diversas condições de temperatura e concentração iônica.



# Capítulo 9

## Conclusões

Em linhas gerais, o trabalho aqui descrito consistiu no desenvolvimento de um modelo teórico para o cálculo da constante dielétrica efetiva e das propriedades físico-químicas dos aminoácidos. Em relação ao primeiro parâmetro, os valores apresentados nos capítulos 5, 7 e 8 se mostram concordantes com aqueles relatados em outros trabalhos e obtidos por abordagens tanto teóricas<sup>13,19,38</sup> quanto teórico-experimentais<sup>20,21,61</sup>. Pela forma como se dá o cálculo desta constante, o papel exercido pelos íons em solução é levado em conta implicitamente pela utilização do comprimento de Debye,  $\lambda$ , no cálculo dos potenciais eletrostáticos e conseqüente aplicação de condições de contorno.

Realizado o desenvolvimento matemático para o cálculo do potencial eletrostático médio e com os valores da supracitada constante, construiu-se um programa computacional para a determinação das constantes de integração e solução das equações transcendentais encontradas. Após uma série de aprimoramentos, esse programa permitiu calcular os valores de  $pK^*$ , grau de ionização e carga elétrica média em função da temperatura, pH e concentração iônica.

Os valores assim determinados de  $pK^*$  foram, no capítulo 6, comparados com os obtidos experimentalmente por Henchoz<sup>9</sup> e houve uma satisfatória concordância entre ambos, sendo que na maioria dos casos as diferenças ocorrem na segunda casa decimal. Além de indicar a versatilidade do modelo, este fato deu consistência à hipótese e aproximações adotadas e serviu de motivação para extensão do modelo a moléculas mais complexas, tais como peptídeos com três grupos ionizáveis, o que foi realizado no capítulo 7. Novamente lançou-se mão do critério de semelhança estrutural para atribuir os valores de  $pK$  intrínseco e determinou-se a constante dielétrica efetiva das regiões compreendidas entre cada par de grupos ácido-base. Os resultados de  $pK^*$  obtidos para gly e ala pentapeptídeos apresentaram um Fator de Correlação de Pearson igual a 0,9966.

Este cenário motivou a extensão do modelo para o estudo de resíduos ionizáveis da proteína *Staphylococcal Nuclease*. Usando uma metodologia bastante análoga a dos capítulos anteriores e com o desenvolvimento de uma generalização da expressão para o potencial eletrostático médio, no oitavo capítulo determinaram-se os valores de  $pK_i^*$  de resíduos de glutamato e lisina. Quando comparados os valores de  $pK_i^*$  calculados com os correspondentes fornecidos pelas referências, obteve-se uma correlação ( $\rho_{FCP}$ ) de 0,9976. Realizou-se também o estabelecimento de uma distância de *cut-off*, responsável por caracterizar a máxima distância entre os grupos ionizáveis a serem considerados no cálculo do potencial eletrostático. O valor encontrado para tal distância corresponde a, aproximadamente, o dobro do comprimento de Debye.

Além da publicação, em periódicos científicos, dos resultados provenientes do modelo desenvolvido, pretende-se disponibilizar os programas computacionais desenvolvidos em forma de uma *facility*, o que consistirá em uma importante ferramenta para o estudo de biomoléculas em variadas condições de temperatura e concentração iônica, por exemplo. Estes programas são caracterizados pela rapidez dos cálculos e pela direta especificação, por parte do usuário, das condições supracitadas.

## Referências Bibliográficas:

1. Nelson, D. L. & Cox, M. M. *Lehinger Princípios de Bioquímica*. (Savier: São Paulo, 2006).
2. Valliere-Douglass, J. F. *et al.* Glutamine-Linked and Non-Consensus Asparagine-Linked Oligosaccharides Present in Human Recombinant Antibodies Define Novel Protein Glycosylation Motifs. *J. Biol. Chem.* **285**, 16012–16022 (2010).
3. Mordel, J. *et al.* Activation of glycine receptor phase - shifts the circadian rhythm in neuronal activity in the mouse suprachiasmatic nucleus. *The Journal of Physiology* **589**, 2287–2300 (2011).
4. Trevino, S. R., Scholtz, J. M. & Pace, C. N. Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa. *Journal of Molecular Biology* **366**, 449–460 (2007).
5. Honig, B. & Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **268**, 1144–1149 (1995).
6. Gao, J. *et al.* Electron Transfer in Peptides: The Influence of Charged Amino Acids. *Angewandte Chemie International Edition* **50**, 1926–1930 (2011).
7. Nishikawa, M., Hasegawa, S., Yamashita, F., Takakura, Y. & Hashida, M. Electrical charge on protein regulates its absorption from the rat small intestine. *American Journal of Physiology - Gastrointestinal and Liver Physiology* **282**, G711–G719 (2002).
8. Vceláková, K., Zusková, I., Kenndler, E. & Gas, B. Determination of cationic mobilities and pKa values of 22 amino acids by capillary zone electrophoresis. *Electrophoresis* **25**, 309–317 (2004).
9. Henchoz, Y. *et al.* Rapid determination of pK a values of 20 amino acids by CZE with UV and capacitively coupled contactless conductivity detections. *Analytical and Bioanalytical Chemistry* **389**, 1869–1878 (2007).
10. Hilal, S. H., Karickhoff, S. W. & Carreira, L. A. A Rigorous Test for SPARC's Chemical Reactivity Models: Estimation of More Than 4300 Ionization pKas. *Quantitative Structure - Activity Relationships* **14**, 348–355 (1995).
11. Hilal, S. H. *et al.* Estimation of the ionization pKa of pharmaceutical substances using the computer program Sparc. *Talanta* **43**, 607–619 (1996).

12. Motekaitis, R. J. & Martell, A. E. Program PKAS: a novel algorithm for the computation of successive protonation constants. *Canadian Journal of Chemistry* **60**, 168–173 (1982).
13. Wallace, J. A. & Shen, J. K. Predicting pKa values with continuous constant pH molecular dynamics. *Meth. Enzymol.* **466**, 455–475 (2009).
14. Gunner, M. R., Zhu, X. & Klein, M. C. MCCE analysis of the pKas of introduced buried acids and bases in staphylococcal nuclease. *Proteins: Structure, Function, and Bioinformatics* **79**, 3306–3319 (2011).
15. Antosiewicz, J., McCammon, J. A. & Gilson, M. K. Prediction of Ph-dependent Properties of Proteins. *Journal of Molecular Biology* **238**, 415–436 (1994).
16. Smith, G. D. *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. (Oxford University Press: 1985).
17. Schutz, C. N. & Warshel, A. What are the dielectric ‘constants’ of proteins and how to validate electrostatic models? *Proteins* **44**, 400–417 (2001).
18. Warshel, A. & Aqvist, J. Electrostatic Energy and Macromolecular Function. *Annual Review of Biophysics and Biophysical Chemistry* **20**, 267–298 (1991).
19. García-Moreno, B. E. *et al.* Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophysical Chemistry* **64**, 211–224 (1997).
20. Isom, D. G., Castañeda, C. A., Cannon, B. R. & García-Moreno E., B. Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **108**, 5260–5265 (2011).
21. Isom, D. G., Castañeda, C. A., Cannon, B. R., Velu, P. D. & García-Moreno E., B. Charges in the hydrophobic interior of proteins. *Proceedings of the National Academy of Sciences* **107**, 16096–16100 (2010).
22. Mehler, E. L. The Lorentz-Debye-Sack theory and dielectric screening of electrostatic effects in proteins and nucleic acids. *Molecular Electrostatic Potentials Concepts and Applications* **Volume 3**, 371–405 (1996).
23. Debye, P. *Polar Molecules*. (Dover: New York, 1929).
24. Lorentz, H. A. *Theory of Electrons*. (Dover Publications: New York, 1952).
25. Agostinho Neto, A., Drigo Filho, E., Fossey, M. A. & Ruggiero Neto, J. pK Determination. A Mean Field, Poisson–Boltzmann Approach. *J. Phys. Chem. B* **103**, 6809–6814 (1999).
26. Bossa, G. V. & Agostinho Neto, A. *Determinação do Grau de Dissociação de Aminoácidos*. (Universidade Estadual Paulista ‘Júlio de Mesquita Filho’: São José do Rio Preto, 2010).
27. Kiss, T., Sóvágó, I. & Gergely, A. Critical Survey of Stability Constants of Complexes of Glycine. *Pure & Appl. Chem.* **63**, 597–638 (1991).

28. Alberts, B. et al. *Biologia Molecular da Célula*. (Artes Médicas: Porto Alegre, 2004).
29. Cronin, J. R. & Pizzarello, S. Enantiomeric Excesses in Meteoritic Amino Acids. *Science* **275**, 951–955 (1997).
30. Sephton, M. A. & Botta, O. Recognizing life in the Solar System: guidance from meteoritic organic matter. *International Journal of Astrobiology* **4**, 269–276 (2005).
31. Voet, D., Voet, J. G. & Pratt, C. W. *Fundamentos de Bioquímica*. (Armed: Porto Alegre, 2000).
32. Atkins, P. W. *Physical Chemistry*. (Oxford University Press: Oxford, 1994).
33. Atkins, P. W. & de Paula, J. *Físico-Química*. **1**, (LTC: Rio de Janeiro, 2008).
34. Solomons, T. W. G. *Química Orgânica*. (LTC: Rio de Janeiro, 2001).
35. Willians, R. pKa Data Compiled. (2004).
36. Voges, D. & Karshikoff, A. A model of a local dielectric constant in proteins. *The Journal of Chemical Physics* **108**, 2219–2227 (1998).
37. Mellor, B. L., Cruz Cortés, E., Busath, D. D. & Mazzeo, B. A. Method for Estimating the Internal Permittivity of Proteins Using Dielectric Spectroscopy. *J. Phys. Chem. B* **115**, 2205–2213 (2011).
38. Wisz, M. S. & Hellinga, H. W. An empirical model for electrostatic interactions in proteins incorporating multiple geometry - dependent dielectric constants. *Proteins: Structure, Function, and Bioinformatics* **51**, 360–377 (2003).
39. Rees, D. C. Experimental evaluation of the effective dielectric constant of proteins. *Journal of Molecular Biology* **141**, 323–326 (1980).
40. Arora, N. & Jayaram, B. Strength of hydrogen bonds in  $\alpha$  helices. *Journal of Computational Chemistry* **18**, 1245–1252 (1997).
41. Hassan, S. A. Liquid-Structure Forces and Electrostatic Modulation of Biomolecular Interactions in Solution. *J. Phys. Chem. B* **111**, 227–241 (2006).
42. Gong, H., Hocky, G. & Freed, K. F. Influence of Nonlinear Electrostatics on Transfer Energies Between Liquid Phases: Charge Burial Is Far Less Expensive Than Born Model. *PNAS* **105**, 11146–11151 (2008).
43. Richards, T. W. & Shipley, J. W. THE DIELECTRIC CONSTANTS OF TYPICAL ALIPHATIC AND AROMATIC HYDROCARBONS, CYCLOHEXANE, CYCLOHEXANONE, AND CYCLOHEXANOL. *J. Am. Chem. Soc.* **41**, 2002–2012 (1919).
44. Hamelin, J., Mehl, J. B. & Moldover, M. R. The Static Dielectric Constant of Liquid Water Between 274 and 418 K Near the Saturated Vapor Pressure. *International Journal of Thermophysics* **19**, 1359–1380 (1998).

45. Robinson, R. A., Stokes, R. H. & Bates, R. G. *Electrolyte Solutions: The Measurement and Interpretation of Conductance, Chemical Potential and Diffusion in Solutions of Simple Electrolytes* (Second Edition). *Journal of The Electrochemical Society* **107**, 205C–206C (1960).
46. Israelachvili, J. N. *Intermolecular And Surface Forces*. (Academic Press: 2010).
47. *CRC Handbook of Chemistry and Physics, 92nd Edition*. (CRC Press: 2011).
48. Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* **79**, 2361–2381 (1975).
49. Yasuda, Y., Tochio, N., Sakurai, M. & Nitta, K. Partial Molar Volumes and Isentropic Compressibilities of Amino Acids in Dilute Aqueous Solutions. *J. Chem. Eng. Data* **43**, 205–214 (1998).
50. Kikuchi, M., Sakurai, M. & Nitta, K. Partial Molar Volumes and Adiabatic Compressibilities of Amino Acids in Dilute Aqueous Solutions at 5, 15, 25, 35, and 45 .degree.C. *J. Chem. Eng. Data* **40**, 935–942 (1995).
51. Meiler, J., Zeidler, A., Schmüscke, F. & Müller, M. Generation and evaluation of dimension-reduced amino acid parameter representations by artificial neural networks. *Journal of Molecular Modeling* **7**, 360–369 (2001).
52. Delano, W. The PyMOL Molecular Graphics System. (2002).at <<http://www.pymol.org>>
53. Hill, T. L. *An Introduction to Statistical Thermodynamics*. (Dover Publications: New York, 1986).
54. Poon, W. C. K. & Andelman, D. *Soft Condensed Matter Physics in Molecular And Cell Biology*. (Taylor & Francis: 2006).
55. Drigo Filho, E. & Agostinho Neto, A. Equilibrium Ionic Distribution as a Boundary Value Problem. *Trends in Physical Chemistry* **10**, 65–77 (2004).
56. G. N. Watson *A Treatise of the Theory of Bessel Functions*. (Cambridge University Press: Cambridge, United Kingdom, 1944).
57. Arfken, G. B. & Weber, H. J. *Física Matemática: métodos matemáticos para engenharia e física*. (Campus Elsevier: Rio de Janeiro, 2007).
58. Kolmogorov, A. N. & Fomin, S. V. *Elementos da Teoria das Funções e de Análise Funcional*. (Mir: Moscou, 1982).
59. Montal, M. & Mueller, P. Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties. *PNAS* **69**, 3561–3566 (1972).

60. Feshbach, Herman, M., Philip McCord *Methods of Theoretical Physics*. **1,2**, (McGraw Hill: New York, 1953).
61. Castañeda, C. A. *et al.* Molecular determinants of the pKa values of Asp and Glu residues in staphylococcal nuclease. *Proteins: Structure, Function, and Bioinformatics* **77**, 570–588 (2009).
62. Mukerjee, P. & Desai, N. R. ‘Effective’ Dielectric Constant for Short-range Ionic Interactions at High Salt Concentrations. , *Published online: 06 September 1969; / doi:10.1038/2231056a0* **223**, 1056–1057 (1969).
63. Laurents, D. V. *et al.* Charge–Charge Interactions are Key Determinants of the pK Values of Ionizable Groups in Ribonuclease Sa (pI=3.5) and a Basic Variant (pI=10.2). *Journal of Molecular Biology* **325**, 1077–1092 (2003).
64. Makowski, M., Liwo, A., Sobolewski, E. & Scheraga, H. A. Simple Physics-Based Analytical Formulas for the Potentials of Mean Force of the Interaction of Amino-Acid Side Chains in Water. V. Like-Charged Side Chains. *J. Phys. Chem. B* **115**, 6119–6129 (2011).
65. Keim, P., Vigna, R. A., Morrow, J. S., Marshall, R. C. & Gurd, F. R. N. Carbon 13 Nuclear Magnetic Resonance of Pentapeptides of Glycine Containing Central Residues of Serine, Threonine, Aspartic and Glutamic Acids, Asparagine, and Glutamine. *J. Biol. Chem.* **248**, 7811–7818 (1973).
66. Thurlkill, R. L., Grimsley, G. R., Scholtz, J. M. & Pace, C. N. pK values of the ionizable groups of proteins. *Protein Science* **15**, 1214–1218 (2006).
67. Creighton, T. E. *Proteins: Structures and Molecular Properties*. (W. H. Freeman and Company: New York, 1993).
68. Lee Rodgers, J. & Nicewander, W. A. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician* **42**, 59–66 (1988).
69. Harms, M. J. *et al.* The pK(a) values of acidic and basic residues buried at the same internal location in a protein are governed by different factors. *J. Mol. Biol.* **389**, 34–47 (2009).
70. Chimenti, M. S. *et al.* Structural Reorganization Triggered by Charging of Lys Residues in the Hydrophobic Interior of a Protein. *Structure* **20**, 1071–1085 (2012).