Universidade de São Paulo
Instituto de Física

O efeito da regulação de cargas nos potenciais de
interação proteína-proteína: uma abordagem teórica e
experimental

Fredderico Camilo Machado

Orientador: Prof. Dr. Leandro Ramos Souza Barbosa

Dissertação de Mestrado apresentada ao Instituto de Física
da Universidade de São Paulo, como requisito parcial para a
obtenção do título de Mestre em Ciências.

Banca Examinadora:

Prof. Dr. Leandro Ramos Souza Barbosa - Orientador (USP)

Prof. Dr. Fernando Luís Barroso da Silva (FCFRP/USP)

Prof. Dr. Luís Fernando Mercier Franco (UNICAMP)

São Paulo

2019

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Machado, Fredderico Camilo

O efeito de regulação de cargas nos potenciais de interação proteína-proteína: uma abordagem teórica e experimental. São Paulo, 2019.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. Física Geral.

Orientador: Prof. Dr. Leandro Ramos Souza Barbosa
Área de Concentração: Biofísica

Unitermos: 1. Biofísica; 2. Espalhamento de raios X à baixos ângulos; 3. Proteínas.

USP/IF/SBI-047/2019

University of São Paulo
Physics Institute

The effect of charge regulation on protein-protein
interaction potentials: a theoretical and experiental
approach

Fredderico Camilo Machado

Supervisor: Prof. Dr. Leandro Ramos Souza Barbosa

Dissertation submitted to the Physics Institute of the
University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:

Prof. Dr. Leandro Ramos Souza Barbosa - Supervisor (USP)

Prof. Dr. Fernando Luís Barroso da Silva (FCFRP/USP)

Prof. Dr. Luís Fernando Mercier Franco (UNICAMP)

São Paulo

2019

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao Prof. Dr. Leandro R. S. Barbosa pela oportunidade, além de sua disponibilidade para discussões, sempre dando sugestões para o melhor desenvolvimento deste trabalho.

À Profa. Dra. Rosangela Itri por seus valiosos conselhos durante as discussões de dados.

Ao Prof. Dr. Francesco Spinozzi pela implementação da teoria de regulação de cargas no seu programa, *GENFIT*, além das discussões acerca dos resultados obtidos com as simulações.

Ao Marcelo pela ajuda com os componentes de laboratório.

Ao amigo Gustavo Campos, por todos os conselhos, discussões e paciência para ensinar o funcionamento do programa *GENFIT*.

Ao amigo Fernando Tanouye por toda ajuda, discussões e resultados que obtivemos sobre regulação de carga.

À todos os amigos do departamento de biofísica: Mayra Lotierzo e Natália Fernandes, pela ajuda com as medidas de SAXS; Luiz Rodrigues, por toda a discussão sobre SAXS e proteínas; Raphael Castro, pela ajuda com programação; Juliana Raw, pela discussão sobre características da BSA; Bárbara Santana por todas as discussões e dicas.

À minha noiva Princess Ria Ponce, que apesar de morar nas Filipinas, sempre me deu amor, carinho, apoio e motivação para prosseguir com o mestrado.

À minha família: Jocely, Helio, Raphael, Carol, Julia e Matheus por todo amor e suporte necessário durante esses anos.

Aos meus vizinhos: Dona Carminha, Seu Clarindo, Dulce e João por toda ajuda emocional durante o mestrado.

Ao Laboratório Nacional de Luz Síncrotron (LNLS) pela realização das medidas de SAXS.

As agências CNPq e POSLATAM pelo apoio financeiro essencial para a participação de congressos e para o desenvolvimento desta pesquisa.

*“Eu mantenho o tema dos meus estudos sempre diante de mim, e espero até o amanhecer
iniciar gradualmente, pouco a pouco, numa luz clara e completa.”*

Sir Isaac Newton

Resumo

A regulação de cargas de proteínas é estudada por vários grupos de pesquisa desde os anos 50. Esta teoria considera que as cargas superficiais proteicas não tem um valor constante, mas sim uma flutuação devido as condições do sistema em que se encontram. Esta flutuação é dada pela variância de carga, também chamada de parâmetro de regulação de carga, podendo ser afetada pela presença de outras macromoléculas, como proteínas, ou mesmo pelas mudanças do meio, através da variação de pH, por exemplo. Com o objetivo de investigar a influência da flutuação de carga nos potenciais de interação proteína-proteína, foram utilizadas soluções de albumina de soro bovina (do inglês, *Bovine Serum Albumin*, *BSA*) e citocromo C em diferentes concentrações e valores de pH. Neste trabalho, os efeitos gerados pela regulação de cargas foram implementados nos potenciais de interação proteína-proteína, de modo que pudessem ser mensurados utilizando curvas de espalhamento de raio-X a baixo ângulo (do inglês, *Small Angle X-ray Scattering*, *SAXS*). Proteínas em solução interagem entre si através da função de correlação total, $h(r)$, sendo encontrada com a solução da equação de Ornstein-Zernike. Neste trabalho foi escolhido a relação de fechamento de Aproximação de Fase Aleatória (do inglês, *Random Phase Approximation*, *RPA*), de modo que os potenciais de interação proteína-proteína são considerados perturbações (já incluindo o termo de regulação de cargas) de um sistema de esferas duras. A partir desta equação, chega-se ao valor algébrico do fator de interferência, $S(q)$, que é um termo fundamental na equação de espalhamento a baixos ângulos.

Os resultados encontrados mostraram que apesar do desenvolvimento da teoria ser para sistemas diluídos, tanto para a BSA, quanto para o citocromo C, foram encontradas

soluções para sistemas concentrados, podendo ser a globularidade destas proteínas um fator para este resultado. Além disso, observa-se que o fator devido a regulação de cargas não consegue descrever todo o potencial atrativo do sistema quando considerado como único termo de potencial atrativo. Quando este fator é considerado juntamente com o atrativo tipo Yukawa, os resultados mostraram que o aumento do valor do parâmetro de regulação de cargas, não altera significativamente o termo de potencial de contato, J , indicando assim que o primeiro não pôde ser mensurado a partir de *SAXS* para as curvas de BSA e citocromo C por ser pequeno comparado à J . O próximo passo é testar essa teoria para sistemas micelares.

Palavras chave: Regulação de Cargas, Proteínas modelo, BSA, *SAXS*, interação proteína-proteína.

Abstract

Protein charge regulation has been studied by various research groups since the 1950s. This theory considers that protein charge surface is not a constant value but it fluctuates due to the conditions of the system in which they are found. This fluctuation is given by the charge variance, also called capacitance, and it may be affected by the presence of other macromolecules, such as proteins, or even by changes in the environment, through pH variation, for example. To investigate the influence of charge fluctuation on protein-protein interaction potential, solutions of bovine serum albumin (*BSA*) and cytochrome C were used in different concentrations and pH. In this work, the effects created by charge regulation are implemented in the protein-protein interaction potential, so that they can be measured through Small Angle X-ray Scattering (*SAXS*) curves. Proteins in solution interact with each other through the total correlation function, $h(r)$, which can be found with the solution of the Ornstein-Zernike equation. In this work we have chosen the Random Phase Approximation (*RPA*), so that protein-protein interaction potentials are considered perturbations (including the charge regulation term) of hard spheres system. From this equation, we get the algebraic value of the interference factor, $S(q)$, which is a fundamental term in the *SAXS* equation.

The results showed that despite the development of the theory for dilute systems, *BSA* and cytochrome C solutions were found for concentrated systems, and the globularity of these proteins may be a factor for this result. Moreover, it is observed that the factor due to charge regulation cannot describe the complete attractive potential of the system when it is considered the only attractive potential term. When this factor is coupled with an

attractive Yukawa potential, the results show that increasing the capacitance value does not significantly change the potential at contact term, J , thus indicating that the first could not be observed from *SAXS* for the BSA and cytochrome C curves because it is small compared to J . The next step is to test this theory for micellar systems.

Keywords: Charge regulation, proteins, BSA, *SAXS*, protein-protein interaction.

Sumário

1. <i>Preâmbulo</i>	15
2. <i>Introdução</i>	17
2.1 <i>Proteínas</i>	17
2.1.1 <i>Albumina de Soro Bovino - BSA</i>	20
2.1.2 <i>Citocromo C de coração equino</i>	21
2.2 <i>Interação entre proteínas</i>	21
2.3 <i>O efeito de Regulação de Carga</i>	23
3. <i>Justificativa e Objetivos</i>	25
3.1 <i>Justificativa</i>	25
3.2 <i>Objetivo Geral</i>	25
3.3 <i>Objetivos Específicos</i>	25
4. <i>Materiais e Métodos</i>	26
4.1 <i>Bases teóricas para o estudo dos potencial de interação entre proteínas</i>	26
4.2 <i>Desenvolvimento teórico do efeito de regulação de cargas</i>	29
4.2.1 <i>Parâmetro de regulação de carga em função do pH</i>	34
4.3 <i>Espalhamento de Raio-X a Baixos Ângulos - SAXS</i>	37
4.3.1 <i>Medidas de SAXS</i>	37
4.3.2 <i>Representação de Kratky</i>	38
4.3.3 <i>Lei de Porod</i>	39

4.3.4	Função de distribuição de pares de distâncias - $p(r)$	40
4.3.5	Aproximação de Guinier	41
4.3.6	Modelo Geométrico para $P(q)$	43
4.4	Fator de estrutura descrito a partir da função de correlação	45
4.4.1	Função de distribuição radial, $g(r)$	45
4.4.2	Função de correlação de pares, $c(r)$, e correlação total, $h(r)$	47
4.4.3	Cálculo do Fator de Estrutura, $S(q)$	52
4.4.4	Fator de estrutura do sistema de referência	54
4.5	Aquisição de medidas experimentais	57
4.5.1	Preparação de amostras	57
4.5.2	Medidas de SAXS	58
4.6	Cálculos Teóricos aplicados	59
4.6.1	Carga e parâmetro de regulação de carga em função do pH	59
4.6.2	Análise das curvas experimentais	60
5.	<i>Resultados e Discussões</i>	62
5.1	Carga e Capacitância em função do pH	62
5.2	Albumina de Soro Bovina	64
5.2.1	Caracterização estrutural da BSA	64
5.2.2	Ajuste das curvas com $C=0$	68
5.2.3	Diferenças dos fatores de estrutura	71
5.2.4	Ajuste das curvas com $J=0$	72
5.2.5	Ajuste das curvas com potencial de contato e capacitância	74
5.3	Citocromo C de coração equino	77
5.3.1	Ajuste das curvas com $C=0$	79
5.3.2	Ajuste das curvas para $C \neq 0$ com $J=0$ e para $C \neq 0$ com $J \neq 0$	81
5.4	Comparação dos resultados entre BSA e citocromo C	85
6.	<i>Conclusões</i>	89

<i>Apêndice</i>	99
<i>A. Tabela de funções e integrais</i>	100
<i>B. Transformada de Fourier de um Potencial do Tipo-Yukawa</i>	102

Preâmbulo

Neste trabalho estudamos o efeito do processo físico-químico chamado de regulação de cargas na análise das curvas de espalhamento de raios-X a baixos ângulos, SAXS, para sistemas compostos por proteínas em altas concentrações. Nestes casos, não é possível negligenciar a interação entre as partículas espalhadoras (neste caso as proteínas) durante a análise das curvas de espalhamento a baixos ângulos.

Para executar esta tarefa, realizamos uma busca bibliográfica acerca deste tema, estudando também em profundidade a teoria envolvida para descrever este processo. Um ponto que foi o núcleo deste trabalho e que consideramos como nossa maior contribuição dentro da área de análise de curvas de SAXS, se dá na proposição do processo de regulação de cargas na análise das curvas de SAXS de sistemas compostos por proteínas em altas concentrações. Neste contexto, esta dissertação encontra-se dividida em 5 capítulos, sendo eles:

- **Introdução**

Nesta sessão descrevemos a importância do estudo físico-químico de proteínas em solução, em condições mais próximas possíveis às encontradas na natureza. Começamos com uma breve descrição dos níveis estruturais e composição das proteínas. Neste capítulo, damos especial atenção às proteínas modelo: Albumina de Soro Bovino e o Citocromo C de coração de cavalo.

- **Objetivos**

Colocamos de modo direto os objetivos gerais e específicos deste trabalho.

- **Materiais e Métodos**

Nessa sessão descrevemos de modo pormenorizado a fundamentação teórica tanto do cálculo dos potenciais de interação entre proteínas em solução, seguindo a já tradicional área de SAXS, bem como uma descrição minuciosa do processo de regulação de cargas, com todo formalismo teórico necessário para a compreensão deste processo.

- **Resultados e Discussões**

Nesta sessão descrevemos tanto os principais resultados teóricos, quanto experimentais, obtidos ao longo do desenvolvimento deste trabalho.

- **Conclusões** Nesta sessão colocamos as principais conclusões obtidas no presente trabalho.

Introdução

2.1 Proteínas

Proteínas são macromoléculas biológicas que possuem uma grande variedade de estruturas, formas e funções dentro de todos os organismos vivos conhecidos. As proteínas podem ser constituídas desde pequenos peptídeos até grandes cadeias de aminoácidos, sendo conhecidos 20 diferentes tipos de aminoácidos naturais (Lehninger et al. [2005], Torres e Marzzoco [2007]), além de outros sintéticos. De um modo geral, são compostos que apresentam um grupo amino ($-\text{NH}_2$) e um grupo carboxila ($-\text{COOH}$), ligados a um carbono central. Os aminoácidos se diferem por sua cadeia lateral, chamado de resíduo ou também de grupo R, como mostra a figura (2.1). Além disso, alguns desses resíduos podem ser ionizáveis, tal como os resíduos ácidos (ácido aspártico, ácido glutâmico) e os resíduos básicos (lisina, arginina, histidina). Para a formação de uma cadeia polipeptídica, o grupo carboxila de um aminoácido liga-se ao grupo amino do seguinte através de ligações peptídicas, ocorrendo a liberação de uma molécula de água, figura (2.2).

Devido às diferenças entre aminoácidos, a sequência no qual estes estão ligados é muito importante pois definem a estrutura da proteína que, por sua vez, define sua função (figura 2.3). Os níveis estruturais de uma proteína são subdivididos em:

- (i) Estrutura primária: sequência de aminoácidos na cadeia polipeptídica;
- (ii) Estrutura secundária: organização da cadeia polipeptídica no qual podem ocorrer rotações entre o carbono C_α e os grupos amina e carboxila. Devido a essas rotações são formadas as chamadas estruturas α -hélice e folha- β que são estabilizadas através de ligações

Figura 2.1: Diferentes tipos de aminoácidos com seus respectivos resíduos. Figura retirada do site Reachingfordreams [2019]

de hidrogênio;

(iii) Estrutura terciária: dá a forma tridimensional da proteína ou ainda como é a organização das estruturas secundárias;

(iv) Estrutura quaternária: é o arranjo de subunidades peptídicas através de ligações não covalentes, sendo característica de proteínas com mais de um domínio. Esta estrutura nem sempre está presente.

Para que uma proteína desempenhe sua função, ela deve ter uma conformação bem definida. Para isso, acontece o processo de enovelamento, que pode acontecer espontane-

Figura 2.2: Formação de um peptídeo através de ligação peptídica. Retirada do site Biochemistry [2019]

amente ou sinteticamente, no qual a proteína irá se “dobrar” de modo a chegar em sua configuração funcional (Berg et al. [2012]). Devido às interações com outras partículas em solução, é possível que haja uma alteração no modo em que a proteína se dobra, podendo ser devida a mutações no material genético que a codifica, ou mesmo um desvio durante o processo de enovelamento, levando a conformações não-funcionais (Chaudhuri e Paul [2006]). Uma mutação devido à alteração de um aminoácido de sua cadeia, por exemplo, poderia gerar uma estrutura que se enovele de maneira diferente da nativa, acarretando na perda de função da proteína (Geiler [2010]). Segundo Bradbury [2003], o enovelamento incorreto de proteínas pode estar relacionado em quase metade das doenças humanas, sendo os males de Alzheimer, Parkinson e Creutzfeldt-Jacob alguns exemplos. Assim sendo, tanto a interação de proteínas e consequentemente o processo de enovelamento proteico estão sendo amplamente estudados na atualidade.

O interior da célula é um meio no qual há alta concentração de proteínas. Nestes casos, o volume relativo de solvente é menor do que em sistemas de baixa concentração *in vitro*, e portanto o volume das proteínas devem ser considerados no volume total. Isto resulta na teoria do “volume excluído” (Hill [1986]). Esses tipos de sistemas são chamados de sistemas aglomerados, ou *crowding systems*. A hemoglobina por exemplo, que tem como função o transporte de oxigênio e gás carbônico dos pulmões para os tecidos (Mairbäurl e Weber [2011]), é encontrada nas células sanguíneas nas concentrações 135-180 mg/ml para homens e 115-160 mg/ml para mulheres (Tresca [2019], GpNotebook [2016]). Por

Figura 2.3: Estruturas da proteína. Figura modificada do site Reachingfordreams [2019]

outro lado, a albumina que tem como função regular a pressão osmótica do sangue (Scott e Mercer [1996]), e é encontrada nas concentrações entre 35 e 50 mg/ml (Barbosa et al. [2010], Tresca [2011]).

2.1.1 Albumina de Soro Bovino - BSA

A BSA é uma proteína, globular, cuja massa molecular é aproximadamente 66.4kDa, figura (2.4). Ela é a proteína mais abundante no plasma sanguíneo, sendo > 50% (Peters Jr [1995], Scott e Mercer [1996]). Contém 583 aminoácidos (entrada do PDB- 4F5S) com 33

α -hélices (usando o *software PDBsum*), representando 66% de sua estrutura secundária (Carter et al. [1994]), com ponto isoelétrico entre 4.8 e 5.6 (Barbosa et al. [2010], Tanford e Buzzell [1956]). A albumina é uma proteína solúvel em água e tem grande capacidade de ligação com íons de Cálcio, Sódio e Potássio (Peters Jr [1995]), além de ser uma proteína facilmente cristalizada e altamente estável, possuindo excesso de aminoácidos ácidos. A sua estabilidade estrutural é devido as 17 pontes de dissulfetos geradas pelos 35 resíduos de cisteínas (Carter et al. [1994]). Normalmente a BSA é usada como proteína modelo devido ao extensivo número de estudos a cerca e sua estrutura e estabilidade.

2.1.2 Citocromo C de coração equino

O citocromo C, figura (2.4), é uma proteína vital na cadeia respiratória celular (Scott e Mercer [1996]), sendo um dos componentes de transporte de elétrons dentro da mitocôndria. É uma proteína pequena, globular, tendo massa molecular de aproximadamente 12,4kDa. Segundo Harding e Rowe [1982], as dimensões do citocromo C são determinadas a partir de um elipsoide de revolução com dimensões de 25 X 25 X 35 Å³. Possui 104 resíduos em sua composição, tendo um ponto isoelétrico por volta de pI = 10.0 - 10.5 (Miyashita et al. [2013], Malmgren et al. [1978]).

Figura 2.4: Estrutura do citocromo C (a) e, BSA (b).

2.2 Interação entre proteínas

No processo de enovelamento proteico é comum que as cargas fiquem na superfície, protegendo as regiões hidrofóbicas. Com a aproximação das macromoléculas em solução,

suas interações serão dadas por forças intermoleculares que incluem forças atrativas e repulsivas (Kurut e Lund [2013], Lund e Jönsson [2013, 2005]). Há autores, como por exemplo Narayanan e Liu [2003], que representam o potencial total de interação entre proteínas em solução a partir da soma de cinco componentes: em que U_{HS} é o potencial de esfera dura, U_{SC} é o potencial de Coulomb blindado, U_H é o potencial atrativo de Van der Waals (Hamaker [1937]), U_0 é potencial atrativo osmótico e U_a é o potencial atrativo de poço quadrado. Com esse potencial, os autores descreveram de modo satisfatório curvas de SAXS para Lisozima. Enquanto isso, há outros autores tais como Li et al. [2015] que utilizam três componentes na descrição do potencial:

$$U_{total}(r) = U_{SC}(r) + U_{LJ}(r) + U_{TS}(r) \quad (2.1)$$

sendo estes o Potencial de Coulomb blindado, Potencial de Leonard-Jones e Potencial devido aos sítios de cada aminoácido. Já Franco et al. [2015] também utilizam três componentes, mas diferentemente descrito na equação (2.1), o potencial é composto por:

$$U_{total}(r) = U_{HS}(r) + U_{SC}(r) + U_{VdW}(r) \quad (2.2)$$

sendo estes o potencial de esfera dura, Potencial de Coulomb blindado e Potencial de Van der Waals, respectivamente.

Para estudar interação entre proteínas, experimentos de espalhamento a baixos ângulos é uma técnica muito utilizada (Barbosa et al. [2013], Svergun e Koch [2002]). Uma boa característica desta técnica é que as amostras não precisam ter uma preparação especial sendo possível caracterizá-las para concentrações de até 200 mg/ml, em diferentes tampões, forças iônicas, temperatura, valores de pH, entre outros (Barbosa et al. [2013]). Por outro lado, as medidas de SAXS geram resultados de intensidade de raios-X no espaço recíproco (ou seja, não está no espaço real, diferentemente de experimentos de microscopia eletrônica, por exemplo), sendo sua interpretação bem descrita pela teoria de espalhamento. Também é possível fazer interpretações no espaço real, para isso é necessário um software que interprete os resultados a partir de transformações de Fourier, transformando resultados de

um espaço para o outro. A função de distribuição de distâncias, por exemplo (caracteriza a dimensão máxima da proteína), é uma das funções que necessitam dados no espaço real.

Com os dados de SAXS também é possível caracterizar muitos elementos dos sistemas: Enovelamento de proteínas, utilizando a representação de Kratky; Raio de giro, medindo a compactação da proteína; Formato da proteína a partir do fator de forma; Interação entre proteínas no sistema, a partir do fator de estrutura; Massa molecular, utilizando a invariante de Porod, entre outras informações. Esses métodos serão detalhados nas próximas sessões.

2.3 O efeito de Regulação de Carga

Como descrito anteriormente, proteínas são constituídas por diferentes sequências de aminoácidos que por sua vez se diferenciam pelo resíduo ligado ao carbono *alpha*. Dependendo do pH do sistema, os resíduos desta proteína podem protonar ou desprotonar, ou seja, ganhar ou perder um próton (H^+), como mostrado na figura (2.5). Devido a esse fato, as cargas superficiais da proteína não serão constantes e portanto a carga total da proteína será dada por uma função de distribuição. A flutuação do valor da carga total é o que dá origem ao efeito de regulação de cargas.

Figura 2.5: Teoria de regulação de carga. Modificada do artigo de Avni et al. [2018].

Quando uma segunda proteína é aproximada de uma proteína de referência, a flutuação de carga de cada proteína irá alterar as forças presentes no sistema a partir de um potencial eletrostático atrativo, associado com a flutuação em número e configurações de ligações de prótons com a cada molécula (Kirkwood e Shumaker [1952]).

Kirkwood e Shumaker [1952] mostraram que a partir de interações eletrostáticas entre moléculas surge um potencial atrativo devido a flutuações de carga e descrevem essa teoria de forma quantitativa utilizando teorias de perturbação. Para isso, eles consideram duas proteínas quem contém uma quantidade de grupos básicos e quando estes ligam-se a um próton, terão sua carga aumentada em uma unidade. utilizando um potencial Coulombiano não blindado mostram que o potencial total será influenciado também por um termo devido a flutuação de cargas das proteínas.

Desde então, muitos pesquisadores contribuíram de maneira significativa para o estudo de interação de proteínas (e consequentemente para a teoria de regulação de carga). Recentemente, mais precisamente nos anos de 2005 (Lund e Jönsson [2005]) e 2013 (Lund e Jönsson [2013]), Jönsson e Lund descreveram a regulação de cargas a partir de um potencial externo ϕ gerado como consequência da interação entre as proteínas no sistema, além de alteração do pH da solução.

Justificativa e Objetivos

3.1 *Justificativa*

Nesta dissertação estudamos como o efeito da regulação de cargas pode afetar o potencial de interação proteína-proteína, com especial atenção as análises das curvas de SAXS de sistemas compostos por proteínas modelo em diversas concentrações e valores de pH.

3.2 *Objetivo Geral*

Estudar o efeito de regulação de cargas nos potenciais de interação de proteína em solução, utilizando técnicas de SAXS

3.3 *Objetivos Específicos*

- Estudar a fundamentação teórica do processo de regulação de cargas;
- Adicionar o potencial de regulação de cargas em modelos existentes para análise de curvas de SAXS.
- Investigar como a regulação de cargas pode alterar a análise de curvas de SAXS de proteínas em solução.

Materiais e Métodos

Para entendermos como o potencial gerado pelo efeito de regulação de cargas afetará as análises de SAXS, é importante estudar como um potencial pode afetar a descrição estatística (no limite termodinâmico) do sistema. Para isso, iniciamos este capítulo com esta descrição.

4.1 Bases teóricas para o estudo dos potencial de interação entre proteínas

Considerando um sistema fechado de proteínas, com volume total (V) e temperatura (T) constante, devido à interação entre os componentes desse sistema (proteína-proteína, proteína-íons), a carga média de cada proteína não é necessariamente zero, podendo ela ser alterada por fatores como pH e concentração de sais, e portanto, haverá uma energia livre de interação $F(r)$ (Hill [1986], Salinas [2013])(será escrita como valor médio do potencial). Na aproximação adotada neste trabalho, o potencial total de interação do sistema pode ser descrito como a soma de um potencial de referência, $U_0(\mathbf{r}^n)$, mais um termo perturbativo, $\omega(\mathbf{r}^n)$:

$$U(\mathbf{r}^n) = U_0(\mathbf{r}^n) + \omega(\mathbf{r}^n) \quad (4.1)$$

em que $\mathbf{r}^n$ representa as coordenadas de n partículas no sistema, ou seja, $\mathbf{r}^n = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_n)$. De agora em diante, sempre que uma coordenada estiver em negrito, esta estará relacionada com o vetor tridimensional de posição, enquanto para coordenadas que não estejam

em negrito estarão relacionadas com o módulo do vetor. Pela integral de configuração do *ensemble* canônico (A.2) para esse potencial temos:

$$Z_n = \int \dots \int e^{-\beta(U_0(\mathbf{r}^n) + \omega(\mathbf{r}^n))} d\mathbf{r}_1 \dots d\mathbf{r}_n \quad (4.2)$$

que pode ser reescrito como:

$$Z_n = \int \dots \int e^{-\beta U_0(\mathbf{r}^n)} d\mathbf{r}_1 \dots d\mathbf{r}_n \frac{\int \dots \int e^{-\beta(U_0(\mathbf{r}^n) + \omega(\mathbf{r}^n))} d\mathbf{r}_1 \dots d\mathbf{r}_n}{\int \dots \int e^{-\beta U_0(\mathbf{r}^n)} d\mathbf{r}_1 \dots d\mathbf{r}_n} \quad (4.3)$$

Podemos assumir aqui:

$$\frac{\int \dots \int e^{-\beta(U_0(\mathbf{r}^n) + \omega(\mathbf{r}^n))} d\mathbf{r}_1 \dots d\mathbf{r}_n}{\int \dots \int e^{-\beta U_0(\mathbf{r}^n)} d\mathbf{r}_1 \dots d\mathbf{r}_n} = \frac{\int \dots \int e^{-\beta \omega(\mathbf{r}^n)} e^{-\beta U_0(\mathbf{r}^n)} d\mathbf{r}_1 \dots d\mathbf{r}_n}{\int \dots \int e^{-\beta U_0(\mathbf{r}^n)} d\mathbf{r}_1 \dots d\mathbf{r}_n} = \langle e^{-\beta \omega(\mathbf{r}^n)} \rangle_0 \quad (4.4)$$

em que $\langle \dots \rangle_0$ é a média sobre o sistema de referência. O primeiro termo da equação (4.3) é a integral de configuração para o sistema não perturbado, $Z_n^{(0)}$, e juntamente com o resultado da equação (4.4) tem-se:

$$Z_n = Z_n^{(0)} \langle e^{-\beta \omega(\mathbf{r}^n)} \rangle_0 \quad (4.5)$$

Devido a natureza da função de partição, há um problema ao analisá-la no limite termodinâmico, $N, V \rightarrow \infty$ com $V/N = v$ (Hill [1986], Salinas [2013]). Para fazer essa correção é necessário adicionar o fator de contagem de Boltzmann $N!$ e também um fator h^{3N} , onde h é a constante de Plank, de modo que função de partição não dependa da dimensão do espaço de fase (Hill [1986], Salinas [2013]) Além disso, a integração sobre os momentos pode ser colocada junto com o fator da constante de Planck, gerando assim o termo Λ^{3N} . Portanto, a função de partição será dada como:

$$Z = \frac{Z_n}{N! \Lambda^{3N}} \quad (4.6)$$

A energia livre de Helmholtz é dada por $F = -k_B T \ln Z$, onde k_B é a constante de Boltzmann, T é a temperatura absoluta do sistema e Z a função de partição corrigida,

descrita pela equação (4.6). Dividindo a equação (4.5) pelos fatores de correção $N!\Lambda^{3N}$ e, aplicando a função logarítmica natural de ambos os lados obtém-se:

$$-\beta F = -\beta F^{(0)} + \ln(\exp(-\langle \beta \omega(\mathbf{r}^n) \rangle_0)) \quad (4.7)$$

Em ambos os lados da equação estão sendo retratados valores de energia livre, então o segundo fator do lado direito está relacionado com a energia livre de interação do sistema perturbado, definido por:

$$\beta F_p(\mathbf{r}^n) = -\ln(\exp(-\langle \beta \omega(\mathbf{r}^n) \rangle_0)) \quad (4.8)$$

Considerando interação entre pares de partículas (por exemplo partícula A interagindo com partícula B), $\mathbf{r}^n \rightarrow |\mathbf{r}_A - \mathbf{r}_B| = r$, a equação acima pode ser reescrita como (o índice p de perturbação será omitido a partir desse ponto):

$$\beta F(r) = -\ln(\exp(-\langle \beta \omega(r) \rangle_0)) \quad (4.9)$$

Como o termo dentro da função exponencial se trata de uma perturbação no sistema, é de se esperar que esse valor seja pequeno comparado com o potencial do sistema de referência, $U_0(r)$ e assim, é possível fazer a expansão da função exponencial, equação (A.4), resultando em:

$$\beta F(r) = -\ln\left(1 - \langle \beta \omega(r) \rangle_0 + \frac{1}{2} \langle (\beta \omega(r))^2 \rangle_0 + O\left(\langle (\beta \omega(r))^3 \rangle_0\right)\right) \quad (4.10)$$

Considerando os termos até segunda ordem e que os fatores dentro do logaritmo natural são próximos a zero, por se tratarem de perturbações, pode-se expandir a função $\ln(1-x)$,

utilizando $x = \langle \beta\omega(r) \rangle_0 - \frac{1}{2} \langle (\beta\omega(r))^2 \rangle_0$, equação (A.5), e portanto:

$$\begin{aligned} \beta F(r) &\approx -\ln \left(1 - \langle \beta\omega(r) \rangle_0 + \frac{1}{2} \langle (\beta\omega(r))^2 \rangle_0 \right) \\ \beta F(r) &\approx \langle \beta\omega(r) \rangle_0 - \frac{1}{2} \langle (\beta\omega(r))^2 \rangle_0 + \frac{1}{2} \left(\langle \beta\omega(r) \rangle_0 - \frac{1}{2} \langle (\beta\omega(r))^2 \rangle_0 \right)^2 = \\ &= \langle \beta\omega(r) \rangle_0 - \frac{1}{2} \langle (\beta\omega(r))^2 \rangle_0 + \frac{1}{2} \langle \beta\omega(r) \rangle_0^2 - \frac{1}{2} \langle \beta\omega(r) \rangle_0 \langle (\beta\omega(r))^2 \rangle_0 + \\ &\quad + \frac{1}{8} \langle (\beta\omega(r))^2 \rangle_0^2 \end{aligned} \quad (4.11)$$

O potencial de perturbação será descrito na sessão seguinte a partir dos termos de regulação de cargas.

4.2 Desenvolvimento teórico do efeito de regulação de cargas

A teoria de regulação de cargas reflete interações entre moléculas carregadas e, portanto, a princípio, será considerado apenas o potencial repulsivo. Jönsson e Lund descrevem primeiramente um sistema de duas proteínas em uma solução com baixa concentração de sais, interagindo por um potencial de Coulomb não blindado. Em seguida, adicionam a blindagem ao potencial Coulombiano de modo a simular um sistema cuja concentração de sal é importante (Lund e Jönsson [2013, 2005]). Neste presente trabalho foi seguido a mesma metodologia, mas considerando o potencial Coulombiano blindado obtido através da teoria do DLVO (Derjaguin, Landau, Verwey e Overbeek).

Seja um sistema de proteínas em que o potencial de referência é dado pelo potencial de esfera dura, $U_{HS}(r)$, relacionado com a não penetrabilidade das proteínas (Barbosa et al. [2010], Narayanan e Liu [2003], McDonald e Hansen [1976])

$$U_{HS}(r) = \begin{cases} 0, & \text{se } r \geq \sigma \\ \infty, & \text{se } r < \sigma \end{cases} \quad (4.12)$$

sendo σ o diâmetro da esfera (de modo que ao dizer diâmetro da esfera, deve ser entendido como o diâmetro de uma esfera com volume equivalente ao volume da proteína), perturbado

por um potencial de Coulomb blindado de pares (4.13), dado por Vervey e Overbeek [1948]:

$$\omega_{AB}(r) = \frac{\langle Q_A \rangle \langle Q_B \rangle e^2}{4\pi\epsilon(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} \quad (4.13)$$

em que $k^2 = \frac{2e^2 N_A I}{\epsilon k_B T}$ é o quadrado do inverso do comprimento de Debye ($k \approx \sqrt{I(M)}/3.04$, Narayanan e Liu [2003], Israelachvili [2011]), $\langle Q_{A(B)} \rangle$ é a carga média da proteína A(B), $\sigma_{A(B)}$ é o diâmetro da proteína A(B), e é a carga elementar ($= 1.6 \times 10^{-19}C$), ϵ ($= \epsilon_0 \epsilon_r$) é a constante dielétrica do meio, N_A é o número de Avogadro, k_B é a constante de Boltzmann, T é a temperatura absoluta e I , a Força Iônica, determinada por íons e contra-íons na solução como $I_C = 0.5n_p|\langle Q_i \rangle|$ e adição de sais I_S como $I_S = \frac{1}{2} \sum_{k=1}^n c_k q_k^2$, em que c_k é a molaridade e q_k é o número de carga do íon. Considerando que a interação de pares seja dada entre pares de diferentes elementos das proteínas, o potencial total perturbado é calculado como a soma desses pares de potenciais:

$$\omega_{AB}(r_{ij}) = \sum_i \sum_j \frac{q_i q_j e^2}{4\pi\epsilon(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r_{ij}-0.5(\sigma_A+\sigma_B))}}{|\mathbf{r}_{ij}|} \quad (4.14)$$

onde $\mathbf{r}_{ij}$ é o vetor posição entre os elementos i (que está dentro da proteína A) e j (proteína B). Multiplicando a equação acima por $\beta = 1/k_B T$:

$$\beta\omega_{AB}(r_{ij}) = \sum_i \sum_j \frac{q_i q_j l_B}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r_{ij}-0.5(\sigma_A+\sigma_B))}}{|\mathbf{r}_{ij}|} \quad (4.15)$$

onde l_B é o comprimento de Bjerrum, $l_B = \beta e^2 / 4\pi\epsilon$. Para resolver essa equação de forma que ela dependa da distância entre os elementos de cada proteína, pode-se utilizar a expansão multipolar no termo $\mathbf{r}_{ij}$. Considerando as interações entre duas proteínas em solução, distantes r do centro da primeira para a segunda, no qual cada elemento da proteína terá sua respectiva posição, $\mathbf{r}_i$ (ou $\mathbf{r}_j$) e, distribuição de carga q_i (ou q_j), o vetor distância $\mathbf{r}_{ij}$ será escrito como $\mathbf{r}_{ij} = \mathbf{r} + \mathbf{r}_j - \mathbf{r}_i = \mathbf{r} + \mathbf{R}$, em que $\mathbf{R} = \mathbf{r}_j - \mathbf{r}_i$, $\mathbf{r}$ é o vetor distância centro a centro entre as distribuições, que será considerado constante, como mostrado na figura 4.1. Para simplificar a notação, $|\mathbf{r}_{ij}| = r_{ij}$.

Figura 4.1: Duas distribuições de carga distantes r uma da outra

$$r_{ij} = \sqrt{|\mathbf{r} + \mathbf{R}|^2} = \sqrt{r^2 + 2\mathbf{r} \cdot \mathbf{R} + R^2}$$

$$r_{ij} = r \sqrt{1 + 2\frac{\mathbf{r} \cdot \mathbf{R}}{r^2} + \left(\frac{R}{r}\right)^2} \quad (4.16)$$

Escrevendo o vetor unitário $\mathbf{u} = \mathbf{r}/r$, chega-se à

$$r_{ij} = r \sqrt{1 + \frac{2\mathbf{u} \cdot \mathbf{R}}{r} + \left(\frac{R}{r}\right)^2} \rightarrow \frac{1}{r_{ij}} = \frac{1}{r \sqrt{1 + \frac{2\mathbf{u} \cdot \mathbf{R}}{r} + \left(\frac{R}{r}\right)^2}} \quad (4.17)$$

Para uma solução diluída, há poucas partículas por unidade de volume, e considerando um sistema homogeneamente distribuído, a distância média proteína-proteína r é muito maior do que as dimensões da proteína, então é possível fazer a expansão do inverso da raiz quadrada até primeira ordem, a partir da equação (A.6), usando $x = \frac{2\mathbf{u} \cdot \mathbf{R}}{r} + \left(\frac{R}{r}\right)^2$:

$$\frac{1}{r_{ij}} \approx \frac{1}{r} \left[1 - \frac{1}{2} \left(\frac{2\mathbf{u} \cdot \mathbf{R}}{r} + \left(\frac{R}{r}\right)^2 \right) \right] = \frac{1}{r} \left[1 - \frac{\mathbf{r} \cdot \mathbf{R}}{r^2} - \frac{R^2}{2r^2} \right] \quad (4.18)$$

A equação acima tem solução em termos dos polinômios de Legendre. Serão considerados termos até a primeira ordem, sendo esta proporcional a r^{-2} , por ser o suficiente para o propósito deste estudo. Multiplicando o fator fora do colchete com os fatores internos, teremos termos proporcionais a r^{-3} , que podem ser desconsiderados. Assim sendo, $1/r_{ij} \approx 1/r$ e, conseqüentemente $r_{ij} \approx r$. Podemos então reescrever a equação (4.15) como:

$$\beta\omega_{AB}(r) \approx \sum_i \sum_j \frac{q_i q_j l_B}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r - 0.5(\sigma_A + \sigma_B))}}{r} \quad (4.19)$$

Com essa equação podem ser calculados os termos da equação (4.11):

$$\langle \beta \omega_{AB}(r) \rangle_0 \approx \left\langle \sum_{i,j} \frac{l_B q_i q_j}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} \right\rangle_0$$

Todos os fatores dentro do colchetes são constantes, exceto os valores de carga, então eles podem ser extraídos do cálculo do valor médio e, portanto, será obtido a equação:

$$\langle \beta \omega_{AB}(r) \rangle_0 = \frac{l_B \langle Q_A \rangle_0 \langle Q_B \rangle_0}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} \quad (4.20)$$

onde usou-se o fato que $\sum_{i(j)} q_{i(j)} = Q_{A(B)}$ e também a independência entre a média de diferentes cargas de diferentes proteínas. É importante ressaltar neste ponto que considerando termos até r^{-2} os termos de interação íon-dipolo e dipolo-dipolo não estão sendo avaliados. Barroso da Silva et al. [2006] mostram que para lisozima, alfa-lactoalbumina e beta-lactoglobulina, a interação devido ao efeito de regulação de carga é da magnitude do efeito íon-dipolo. Além disso, indicam que o fator íon-dipolo pode ser mais importante que o termo de regulação de carga quando o momento de dipolo for muito maior que o parâmetro de regulação de carga, apesar de o segundo ter maior alcance ($\propto r^{-2}$) em relação ao primeiro ($\propto r^{-4}$).

Calculando o valor quadrático da equação (4.11), através do resultado da equação (4.19) e $\sum_{i(j)} q_{i(j)} = Q_{A(B)}$:

$$(\beta \omega_{AB}(r))^2 = \frac{l_B^2 Q_A^2 Q_B^2}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} \quad (4.21)$$

Tomando a média sobre esse valor:

$$\langle (\beta \omega_{AB}(r))^2 \rangle_0 = \frac{l_B^2 \langle Q_A^2 \rangle_0 \langle Q_B^2 \rangle_0}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} \quad (4.22)$$

Substituindo os valores encontrados em (4.20) e (4.22) na equação (4.11), pode-se por fim

escrever a energia livre de interação como:

$$\begin{aligned} \beta F(r) \approx & \frac{l_B \langle Q_A \rangle_0 \langle Q_B \rangle_0}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} - \\ & - \frac{1}{2} \frac{\langle Q_A^2 \rangle_0 \langle Q_B^2 \rangle_0 l_B^2}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} + \\ & + \frac{1}{2} \frac{\langle Q_A \rangle_0^2 \langle Q_B \rangle_0^2 l_B^2}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} \end{aligned} \quad (4.23)$$

que pode ser reescrita como:

$$\begin{aligned} \beta F(r) \approx & \frac{l_B \langle Q_A \rangle_0 \langle Q_B \rangle_0}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} - \\ & - \frac{1}{2} \frac{l_B^2 [\langle Q_A^2 \rangle_0 \langle Q_B^2 \rangle_0 - \langle Q_A \rangle_0^2 \langle Q_B \rangle_0^2]}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} \end{aligned} \quad (4.24)$$

Esse fato mostra que não só o potencial Coulombiano estará envolvido na energia livre do sistema, mas que também haverá outros termos agindo de modo a afetar esse valor. Devido a definição de variância, $V_{ar}(x) = \langle x^2 \rangle - \langle x \rangle^2 \geq 0$, é de se esperar que o segundo termo da equação (4.24) seja sempre negativo, resultando em um potencial atrativo.

Para introduzir a flutuação de carga na equação (4.24), primeiramente é necessário definir esse fator. Seja C o parâmetro de regulação de carga definida como a variância de carga da proteína, podemos escrever (Kurut e Lund [2013], Lund e Jönsson [2013, 2005], Jönsson et al. [2007]):

$$C_A \equiv \langle Q_A^2 \rangle - \langle Q_A \rangle^2 \quad (4.25)$$

$$\langle Q_A^2 \rangle = C_A + \langle Q_A \rangle^2$$

$$\langle Q_A^2 \rangle \langle Q_B^2 \rangle = (C_A + \langle Q_A \rangle^2) (C_B + \langle Q_B \rangle^2)$$

$$\langle Q_A^2 \rangle \langle Q_B^2 \rangle = C_A C_B + C_A \langle Q_B \rangle^2 + C_B \langle Q_A \rangle^2 + \langle Q_A \rangle^2 \langle Q_B \rangle^2$$

$$\langle Q_A^2 \rangle \langle Q_B^2 \rangle - \langle Q_A \rangle^2 \langle Q_B \rangle^2 = C_A C_B + C_A \langle Q_B \rangle^2 + C_B \langle Q_A \rangle^2 \quad (4.26)$$

O valor mostrado na equação (4.26) pode ser substituído na equação (4.24) e por consequência, a energia livre do sistema pode ser escrita em função da distribuição de

carga.

$$\beta F(r) \approx \frac{l_B \langle Q_A \rangle_0 \langle Q_B \rangle_0}{(1 + 0.5k\sigma_A)(1 + 0.5k\sigma_B)} \frac{e^{-k(r-0.5(\sigma_A+\sigma_B))}}{r} - \frac{1}{2} \frac{l_B^2 (C_A C_B + C_A \langle Q_B \rangle^2 + C_B \langle Q_A \rangle^2)}{(1 + 0.5k\sigma_A)^2 (1 + 0.5k\sigma_B)^2} \frac{e^{-2k(r-0.5(\sigma_A+\sigma_B))}}{r^2} \quad (4.27)$$

4.2.1 Parâmetro de regulação de carga em função do pH

Um outro modo para descrever o parâmetro de regulação de carga é através do pH da solução e isso pode ser feito através da equação de Henderson-Hassenbach (Lund e Jönsson [2013], Torres e Marzzoco [2007]). Seja uma solução formada por um ácido fraco HA , sendo dissociado em H^+ e A^- . Pode-se achar a constante de dissociação e o pH da solução como:

$$K_a = \frac{[H^+][A^-]}{[HA]} \Rightarrow pK_a = pH - \log \frac{[A^-]}{[HA]} \quad (4.29)$$

$$pK_a = pH - \log \frac{\gamma_{A^-} \rho_{A^-}}{\gamma_{HA} \rho_{HA}}$$

em que γ_x e ρ_x são o coeficiente de atividade e a concentração média do elemento x . Para uma solução ideal, $\gamma_{A^-} = \gamma_{HA} = 1$:

$$pK_a = pH - \log \frac{\rho_{A^-}}{\rho_{HA}} \quad (4.30)$$

Usando a diferença de energia livre, $\beta\epsilon$ (Reed e Reed [1992], Sassi et al. [1992], Barroso da Silva et al. [2006]), a equação (4.30) pode ser reescrita como:

$$\beta\epsilon = -\ln \frac{\rho_{HA}}{\rho_{A^-}} = -\log \frac{\rho_{HA}}{\rho_{A^-}} \ln 10$$

$$\beta\epsilon = (pH - pK_a) \ln 10 \quad (4.31)$$

Com esta expressão, um sistema de dois estados pode ser considerado para um único sítio de titulação. Para isso deve-se rescrever a função de partição do sistema (A.1) como

uma soma dos dois estados, onde o estado fundamental (desprotonação) e a valência da titulação são dados por $\epsilon_0 = 0$ e q_{i_T} ($= +1$ para próton), respectivamente. Além disso, a equação (4.31) é a energia do primeiro estado (protonação).

$$Z = \sum_k e^{-\beta\epsilon_k} = e^{-\beta\epsilon_0} + e^{-\beta\epsilon_1} = 1 + e^{-\beta\epsilon_1} = 1 + e^{-(pH-pK_a)\ln 10} \quad (4.32)$$

Assim, tanto $\langle q_i \rangle$ quanto $\langle q_i^2 \rangle$ podem ser reescritos como:

$$\langle q_i \rangle = \frac{\sum_k q_{i_k} e^{-\beta\epsilon_k}}{\sum_k e^{-\beta\epsilon_k}} = \frac{q_{i_0} + (q_{i_0} + q_{i_T})e^{-(pH-pK_a)\ln 10}}{Z} = q_{i_0} + \frac{q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z} \quad (4.33)$$

$$\langle q_i^2 \rangle = \frac{q_{i_0}^2 + (q_{i_0} + q_{i_T})^2 e^{-(pH-pK_a)\ln 10}}{Z} = q_{i_0}^2 + \frac{2q_{i_0}q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z} + \frac{q_{i_T}^2 e^{-(pH-pK_a)\ln 10}}{Z} \quad (4.34)$$

A partir da equação (4.25) e usando os valores de $\langle q_i \rangle$ e $\langle q_i^2 \rangle$:

$$c_i = q_{i_0}^2 + \frac{2q_{i_0}q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z} + \frac{q_{i_T}^2 e^{-(pH-pK_a)\ln 10}}{Z} - \left(q_{i_0} + \frac{q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z} \right)^2$$

$$c_i = \frac{q_{i_T}^2 e^{-(pH-pK_a)\ln 10}}{Z^2} \quad (4.35)$$

Por outro lado, a carga está relacionada com o pH , e para uma pequena variação no pH pode ocorrer uma variação na carga de modo que:

$$\frac{\partial \langle q_i \rangle}{\partial pH} = -\frac{q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z} \ln 10 + \frac{q_{i_T} e^{-2(pH-pK_a)\ln 10}}{Z^2} \ln 10$$

$$\frac{\partial \langle q_i \rangle}{\partial pH} = -\frac{q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z^2} (1 + e^{-(pH-pK_a)\ln 10}) + \frac{q_{i_T} e^{-2(pH-pK_a)\ln 10}}{Z^2} \ln 10$$

$$\frac{\partial \langle q_i \rangle}{\partial pH} = -\frac{q_{i_T} e^{-(pH-pK_a)\ln 10}}{Z^2} \ln 10 \quad (4.36)$$

Esta equação é exatamente a equação (4.35) com o fator $\ln 10$ dividido por q_{i_T} e então

$$\frac{\partial \langle q_i \rangle}{\partial pH} = -\frac{c_i}{q_{i_T}} \ln 10 \implies c_i = -\frac{q_{i_T}}{\ln 10} \frac{\partial \langle q_i \rangle}{\partial pH} = -\frac{1}{\ln 10} \frac{\partial \langle q_i \rangle}{\partial pH} \quad (4.37)$$

Generalizado esse resultado para a carga total da molécula, chega-se a (Lund e Jönsson

[2005, 2013], Barroso da Silva et al. [2006]):

$$C = -\frac{1}{\ln 10} \frac{\partial \langle Q \rangle}{\partial pH} \quad (4.38)$$

Os resultados das equações (4.25) e (4.38) mostram que a a flutuação em uma distribuição de carga é uma função do pH .

Para duas proteínas idênticas, tem-se $\sigma_A = \sigma_B = \sigma$, $\langle Q_A \rangle = \langle Q_B \rangle = \langle Q \rangle$, e portanto, $C_A = C_B = C$, levando a equação (4.27) a ser reescrita como (Kurut e Lund [2013], Lund e Jönsson [2013, 2005], Jonsson et al. [2007])

$$\beta F(r) \approx \frac{l_B \langle Q \rangle_0^2}{(1 + 0.5k\sigma)^2} \frac{e^{-k(r-\sigma)}}{r} - \frac{l_B^2 (C^2 + 2C \langle Q \rangle^2)}{(1 + 0.5k\sigma)^4} \frac{e^{-2k(r-\sigma)}}{2r^2} \quad (4.39)$$

Se $pH = pI$, onde o pI é o ponto isoelétrico ($\langle Q \rangle = 0$):

$$\beta F(r) \approx -\frac{l_B^2 C^2}{(1 + 0.5k\sigma)^4} \frac{e^{-2k(r-\sigma)}}{2r^2} \quad (4.40)$$

Esse resultado mostra que mesmo quando a carga efetiva total da proteína for zero, poderá existir uma energia livre de interação atrativa, devido às flutuações de cargas. Um modo muito mais simplificado para chegar a essa mesma conclusão seria utilizar o potencial Coulombiano sem blindagem e alterar o comprimento de Bjerrum para um comprimento aparente l'_B de forma a considerar os sais na solução, $l'_B \rightarrow l_B \frac{\exp(-k(r-\sigma))}{(1+0.5\kappa\sigma)^2}$. Kurut e Lund [2013] empregam parecido, mas utilizando $l'_B \rightarrow l_B \frac{\exp(-k(r-\sigma))}{(1+0.5\kappa\sigma)}$.

Assim sendo, o termo de regulação mostra-se como consequência da interação devido à flutuação de carga. Assim como o potencial blindado, este termo é uma perturbação do sistema e portanto, deve ser incluído no termo $\omega(r)$ como:

$$\omega(r) = \omega_{SC}(r) + \omega_{CR} \quad (4.41)$$

4.3 Espalhamento de Raio-X a Baixos Ângulos - SAXS

4.3.1 Medidas de SAXS

Uma radiação é gerada a partir de uma fonte (sendo essa uma fonte convencional de raios-X ou síncrotron) e é direcionada para uma amostra sendo colimada dentro de uma câmara de vácuo. Quando a radiação chega à amostra, parte desta radiação será defletida da trajetória inicial e a outra parte manterá sua trajetória até ambas chegarem ao anteparo. Os vetores da radiação defletida e não defletida são medidos com um ângulo de 2θ , e o vetor de espalhamento é dado por $|\mathbf{q}| = 4\pi \sin(\theta) / \lambda$. Essa relação é devido ao vetor $\mathbf{k}$ de radiação ser descrito como $\mathbf{k} = 2\pi / \lambda$. Assim como pode ser visto na figura (4.2), o vetor de espalhamento é a diferença entre os vetores e portanto seu módulo será:

$$\begin{aligned}
 |\mathbf{q}|^2 &= (\mathbf{k}_{out} - \mathbf{k}_{in})^2 = \mathbf{k}_{out}^2 + \mathbf{k}_{in}^2 - 2\mathbf{k}_{out}\mathbf{k}_{in}\cos(2\theta) \\
 |\mathbf{q}|^2 &= \left(\frac{2\pi}{\lambda}\right)^2 2(1 - \cos(2\theta)) \\
 |\mathbf{q}|^2 &= \left(\frac{2\pi}{\lambda}\right)^2 2(2\sin^2(\theta)) = \frac{16\pi^2 \sin^2(\theta)}{\lambda^2} \\
 |\mathbf{q}| &= \frac{4\pi \sin(\theta)}{\lambda}
 \end{aligned} \tag{4.42}$$

Figura 4.2: Esquema do experimento de espalhamento de Raio-X a baixos ângulos. Retirada de Barbosa et al. [2013].

A solução geral do experimento de espalhamento a baixos ângulos está relacionada com

seção de choque transversal (Spinozzi et al. [1998], Melnichenko [2016]), dado pela equação:

$$\frac{d\Sigma}{d\Omega}(\mathbf{q}) = \frac{1}{V} \left\langle \left| \int_V \delta n_p(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \right|^2 \right\rangle \quad (4.43)$$

Para um sistema isotrópico, monodisperso e centro simétrico a equação acima pode ser escrita como (Barbosa et al. [2010], Melnichenko [2016])

$$\frac{d\Sigma}{d\Omega}(\mathbf{q}) = I(\mathbf{q}) = \gamma n_p P(\mathbf{q}) S_M(\mathbf{q}) + B \quad (4.44)$$

γ é o fator relacionado ao arranjo experimental, B é a radiação de fundo do sistema (ou *Background*), n_p , $P(\mathbf{q})$ e $S_M(\mathbf{q})$ são a densidade numérica, o fator de forma e o fator de estrutura medido, respectivamente. O fator de forma é definido como (Franco et al. [2015]):

$$P(\mathbf{q}) = \langle |F(\mathbf{q})|^2 \rangle \quad (4.45)$$

em que $F(\mathbf{q})$ é a amplitude do vetor de espalhamento $\mathbf{q}$ (McDonald e Hansen [1976], Barbosa et al. [2013], Ortore et al. [2009]). O fator de estrutura medido, mostrado na equação (4.44), é escrito como (Franco et al. [2015]):

$$S_M(\mathbf{q}) = 1 + \frac{\langle F(\mathbf{q}) \rangle^2}{\langle F^2(\mathbf{q}) \rangle} (S(\mathbf{q}) - 1) \quad (4.46)$$

Para proteínas globulares, o fator de estrutura medido pode ser substituído pelo fator de estrutura teórico por um fator de $< \pm 1\%$ (Barbosa et al. [2010], Franco et al. [2015]). Isso ocorre devido ao fato que para proteínas globulares, $\langle F(\mathbf{q}) \rangle^2 \approx \langle F^2(\mathbf{q}) \rangle$ e portanto, $S_M(\mathbf{q}) \approx S(\mathbf{q})$, equação (4.46).

4.3.2 Representação de Kratky

O gráfico de Kratky que é definido como a curva de $I(q)q^2 \times q$. A partir de dados de SAXS é possível saber o estado conformacional da proteína sem nenhum tipo de análise prévia. Para proteínas globulares, a intensidade decai com $q^{-3.6}$ e portanto no gráfico de

Kratky será observado uma formato de “sino” (Barbosa et al. [2013], Svergun et al. [2013]), curva azul na figura (4.3). Já para proteínas desenoveladas, essa relação não ocorre porque sua curva de intensidade decai mais lentamente, resultando em um platô no gráfico de Kratky (curva preta). Enquanto isso, para proteínas parcialmente desenoveladas será observado um comportamento intermediário entre enovelado e desenovelado, sendo que é possível saber se a proteína está mais próxima de um estado ou de outro (estado mais ou menos enovelado - curva vermelha).

Figura 4.3: Gráfico de Kratky para BSA em diferentes estados conformacionais, sendo a curva azul referente a proteína compacta e enovelada, curva em vermelho referente a proteína parcialmente desenovelada e curva preta, referente a proteína totalmente desenovelada. Retirada de Barbosa et al. [2013].

A partir da figura (4.3), também é possível perceber que a curva vermelha em seu início possui um formato de sino e depois sua curva começa a crescer, indicando que esta proteína está enovelada, mas já entrou em processo de desenovelamento, ou seja, se encontra parcialmente desenovelada.

4.3.3 Lei de Porod

A Lei de Porod descreve o comportamento assintótico da curva de intensidade para altos valores do vetor de espalhamento. Nesta região, é possível descrever a intensidade

como (Valcárcel e Santos [2015], Barbosa et al. [2013]):

$$I(q)_{q \rightarrow \infty} = n_p P(q) = \frac{2\pi (\Delta\rho)^2 S}{V q^4} \quad (4.47)$$

em que $\Delta\rho$ é o contraste de densidade eletrônica, S é a área superficial e V é o volume da proteína. A lei de Porod é aplicada a sistemas fractais quando a densidade no interior da partícula não é homogênea (assim como proteínas). Neste caso a intensidade de espalhamento não será proporcional à q^{-4} , mas a um fator q^{-n} com $3 < n < 4$ (Brumberger [2013]).

Apesar da Lei de Porod ser comumente usada para ângulos maiores, há casos no qual ela também é observada no limite $\lim qR_g \rightarrow \infty$. Quando há grandes agregados de proteínas, por exemplo, é possível observar um decaimento de intensidade, para baixos ângulos, $\sim 0.1 \text{ nm}^{-1}$, proporcional à q^{-4} (Valcárcel e Santos [2015]).

4.3.4 Função de distribuição de pares de distâncias - $p(r)$

A função de intensidade depende do estado conformacional da proteína. Para partículas não interagentes (Svergun et al. [2013])(usando o fato da solução estar bem diluída), a intensidade é dada pela equação:

$$I(q) = 4\pi \int_0^{D_{max}} p(r) \frac{\sin(qr)}{qr} dr \quad (4.48)$$

A função $p(r)$, também chamada de função de distribuição de distâncias, é dada como a probabilidade de encontrar um par de elementos espalhadores, a uma certa distância r , em todo o volume da proteína. A partir dela também é possível saber o tamanho máximo da proteína. Esse valor é dado quando a função $p(r)$ vai a zero, ou seja, $p(D_{max}) = 0$. A figura (4.4) mostra que para diferentes geometrias, como o mesmo valor de D_{max} , os gráficos de acabam no mesmo valor, $D_{max} = 10 \text{ nm}$. Esse fato mostra que também é possível estimar a forma da proteína a partir desse fator.

Pode-se fazer uma Transformada de Fourier na equação (4.48), para obter $p(r)$ em

Figura 4.4: Gráfico da função de distribuição geometrias de proteínas. Retirada de Svergun et al. [2013]

função da intensidade:

$$p(r) = \frac{1}{2\pi^2} \int_0^\infty I(q) (qr)^2 \frac{\text{sen}(qr)}{qr} dq \quad (4.49)$$

Essa fórmula é bem útil quando se tem a curva de intensidade. Um fato importante de ressaltar é que para proteínas globulares, a função $p(r)$ é esperada de ser simétrica a partir do seu máximo, ou seja, $r_{p_{max}} \approx D_{max}/2$. Para outras geometrias não é tão trivial fazer uma relação, sendo esta diferente para cada caso.

4.3.5 Aproximação de Guinier

Usando a intensidade dependente da função de distribuição das distâncias, equação (4.48), é possível fazer a aproximação de Guinier para encontrar o Raio de Giro da proteína. Fazendo a aproximação da função $\text{sen}(x)/x$, equação (A.8), e substituindo na equação (4.48) obtém-se:

$$\begin{aligned} I(q) &= 4\pi \int_0^{D_{max}} p(r) \left(1 - \frac{(qr)^2}{3!} + \dots \right) dr \\ &= 4\pi \left(\int_0^{D_{max}} p(r) dr - \int_0^{D_{max}} p(r) \frac{(qr)^2}{3!} dr + \dots \right) \end{aligned} \quad (4.50)$$

Colocando o primeiro termo em evidência:

$$I(q) = 4\pi \int_0^{D_{max}} p(r) dr \left(1 - \frac{q^2}{3} \frac{\int_0^{D_{max}} p(r) r^2 dr}{2 \int_0^{D_{max}} p(r) dr} + \dots \right) \quad (4.51)$$

Definindo o raio de giro como sendo a distribuição de matéria da proteína com relação ao centro de massa, tem-se a equação:

$$R_g^2 \equiv \frac{\int_0^{D_{max}} p(r) r^2 dr}{2 \int_0^{D_{max}} p(r) dr} \quad (4.52)$$

Usando a equação (4.51), o fator $\lim_{q \rightarrow 0} \frac{\sin(x)}{x} = 1$ e portanto:

$$I(0) = \lim_{q \rightarrow 0} 4\pi \int_0^{D_{max}} p(r) \frac{\sin(qr)}{qr} dr = 4\pi \int_0^{D_{max}} p(r) dr \quad (4.53)$$

Nota-se que esse termo é exatamente o fator multiplicativo da equação (4.51) e portanto, juntando essa equação até o segundo ordem, juntamente com as equações (4.52) e (4.53) resultará em:

$$I(q) = I(0) \left(1 - \frac{q^2 R_g^2}{3} + \dots \right) \quad (4.54)$$

O termo em parênteses é muito parecido com os termos da expansão da exponencial mostrada pela equação (A.4), e portanto substituindo pode ser escrita como:

$$I(q)_{q \rightarrow 0} \approx I(0) \exp \left(-\frac{q^2 R_g^2}{3} \right) \quad (4.55)$$

Com os dados experimentais da equação de intensidade é possível achar uma expressão mais simples do raio de giro, mas esta só será válida para os casos em que o sistema está diluído e as partículas do sistema homogêneas. Para isso, basta aplicar o logaritmo da

equação (4.55) em ambos os lados e aproximar por uma equação de reta:

$$\log(I(q)) = \log(I(0)) - \frac{R_g^2}{3}q^2 \quad (4.56)$$

Fazendo o gráfico de $\log(I(q))$ x q^2 tem-se uma reta, com coeficiente angular dado por $-R_g^2/3$ e o coeficiente linear dado por $\log(I(0))$. Essa aproximação pode ser usada independente do formato da proteína, sendo válida limite de $q_{max}R_g < 1.3$ (Barbosa et al. [2013]).

4.3.6 Modelo Geométrico para $P(q)$

O fator de forma descreve a geometria de uma proteína em solução. Ele pode ser entendido como o espalhamento de uma única partícula (proteína) no sistema, dependendo de sua forma e tamanho. Para diferentes geometrias, têm-se diferentes valores de fator de forma, assim como mostrado na figura (4.5). Devido a sua importância na análise de curvas de SAXS, encontra-se na literatura alguns resultados de fator de forma para diferentes geometrias (Narayanan e Liu [2003], Pedersen [1997], Melnichenko [2016]):

(i) Esfera homogênea de raio r :

$$P(q, r) = 9 \left[\frac{\sin(qr) - qr \cos(qr)}{(qr)^3} \right]^2 \quad (4.57)$$

(ii) Disco fino de raio r :

$$P(q, r) = \frac{2}{(qr)^2} \left[1 - \frac{J_1(2qr)}{q\sigma} \right] \quad (4.58)$$

em que $J_1(x)$ é a função de Bessel de ordem 1 (A.10).

(iii) Barra fina de comprimento L :

$$P(q, L) = \frac{2S_i(qL)}{qL} - \frac{\sin^2(qL/2)}{(qL/2)} \quad (4.59)$$

em que $S_i(x)$ é a função seno integral (A.12).

(iv) Elipsóide de revolução, com eixos $(r, r, \epsilon r)$:

$$P(q, r, \epsilon) = \int_0^{\pi/2} P_{esfera}(q, r'(r, \epsilon, \alpha)) \sin(\alpha) d\alpha \quad (4.60)$$

onde:

$$r'(r, \epsilon, \alpha) = r \left(\sin^2(\alpha) + \epsilon^2 \cos^2(\alpha) \right)^{1/2} \quad (4.61)$$

Figura 4.5: Gráfico do fator de forma para diferentes geometrias de proteínas. Retirada de Svergun et al. [2013]

Foi mostrado na sessão da função de distribuição de distâncias, $p(r)$, que a partir dessa função é possível descobrir a geometria da proteína. Portanto é possível relacionar essa função com o fator de forma através de uma transformação de Fourier (Guinier e Fournet [1955]):

$$P(q) = \int_0^\infty p(r) \frac{\sin(qr)}{qr} dr \quad (4.62)$$

Por métodos de Monte Carlo, é possível encontrar a função de $p(r)$ e por consequência, encontrar o fator de forma (Spinozzi et al. [2002]). Há também a possibilidade de calcular o fator de forma através da estrutura cristalográfica da proteína (Spinozzi et al. [2014]).

Nesses casos, a amplitude do espalhamento ($F(q)$) é descrita em termos de harmônicos esféricos ($Y_{l,m}(\omega_r)$), sendo esse método melhor detalhado no artigo de Spinozzi et al. [1998].

4.4 Fator de estrutura descrito a partir da função de correlação

4.4.1 Função de distribuição radial, $g(r)$

A função de densidade de probabilidade para um sistema de $\mathbf{N}$ partículas, $P_N^N(\mathbf{r}^N)$, é definida como a probabilidade de, simultaneamente, encontrar a partícula 1 em um volume $d\mathbf{r}_1$ que contém $\mathbf{r}_1$, partícula 2 em um volume $d\mathbf{r}_2$ que contém $\mathbf{r}_2$, etc. Escolhendo um subsistema de $\mathbf{n}$ partículas, $P_N^n(\mathbf{r}^n)$, esta probabilidade é dada por (McDonald e Hansen [1976]):

$$P_N^n(\mathbf{r}^n) = \int \cdots \int P_N^N(\mathbf{r}^N) d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N = \frac{1}{Z_n} \int \cdots \int \exp(-\beta U_N(\mathbf{r}^N)) d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N \quad (4.63)$$

onde Z_n está mostrado na equação (A.1). Como a equação (4.63) é uma função de densidade de probabilidade, é possível escrever a média de uma função $f(\mathbf{r}^n)$ como:

$$\langle f(\mathbf{r}^n) \rangle = \int P_N^n(\mathbf{r}'^n) f(\mathbf{r}'^n) d\mathbf{r}'^n = \frac{1}{Z_n} \int \exp(-\beta U_N(\mathbf{r}^N)) f(\mathbf{r}^n) d\mathbf{r}^N \quad (4.64)$$

Substituindo essa função geral, por um produto de funções delta até a dimensão $\mathbf{r}'^n$, $f(\mathbf{r}^n) = \delta(\mathbf{r}_1 - \mathbf{r}'_1) \delta(\mathbf{r}_2 - \mathbf{r}'_2) \cdots \delta(\mathbf{r}_n - \mathbf{r}'_n)$, quando aplicado na equação (4.64) obtém-se novamente a função $P_N^n(\mathbf{r}^n)$ da equação (4.63).

$$\begin{aligned}
\langle \delta(\mathbf{r}_1 - \mathbf{r}'_1) \dots \delta(\mathbf{r}_n - \mathbf{r}'_n) \rangle &= \int P_N^n(\mathbf{r}'^n) \delta(\mathbf{r}_1 - \mathbf{r}'_1) \dots \delta(\mathbf{r}_n - \mathbf{r}'_n) d\mathbf{r}'^n = \\
&= \int \dots \int P_N^n(\mathbf{r}'_1, \dots, \mathbf{r}'_n) \delta(\mathbf{r}_1 - \mathbf{r}'_1) \dots \delta(\mathbf{r}_n - \mathbf{r}'_n) d\mathbf{r}'_1 \dots d\mathbf{r}'_n \\
&= \int \dots \int P_N^n(\mathbf{r}_1, \mathbf{r}'_2, \dots, \mathbf{r}'_n) \delta(\mathbf{r}_2 - \mathbf{r}'_2) \dots \delta(\mathbf{r}_n - \mathbf{r}'_n) d\mathbf{r}'_2 \dots d\mathbf{r}'_n \\
&\quad \vdots \\
&= \int P_N^n(\mathbf{r}_1, \dots, \mathbf{r}_{n-1}, \mathbf{r}'_n) \delta(\mathbf{r}_n - \mathbf{r}'_n) d\mathbf{r}'_n = P_N^n(\mathbf{r}_1, \dots, \mathbf{r}_n) \\
\langle \delta(\mathbf{r}_1 - \mathbf{r}'_1) \dots \delta(\mathbf{r}_n - \mathbf{r}'_n) \rangle &= P_N^n(\mathbf{r}_1, \dots, \mathbf{r}_n)
\end{aligned} \tag{4.65}$$

Se as partículas são diluídas na solução, isso significa que as partículas estão longe umas das outras, $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$ para todos os i, j , então a densidade de probabilidade das n partículas pode ser aproximada em produtos singulares da densidade de probabilidade

$$P_N^n(\mathbf{r}^n) \approx P_N^1(\mathbf{r}_1) P_N^1(\mathbf{r}_2) \dots P_N^1(\mathbf{r}_n) \tag{4.66}$$

Deste modo, a função de distribuição radial para um subsistema de n partículas pode ser definida, usando os fatores descritos acima:

$$g_N^{(n)}(\mathbf{r}^n) \equiv \frac{P_N^n(\mathbf{r}^n)}{\prod_{i=1}^n P_N^1(\mathbf{r}_i)} \tag{4.67}$$

Pela definição, equação (4.67), quando as partículas estão suficientemente longe uma das outras, a função de distribuição tem a propriedade, $g_N^{(n)}(\mathbf{r}^n) \rightarrow 1$, pois pela equação (4.66) as probabilidades podem ser separadas em produtos das densidades de probabilidade. No limite onde o número de partículas é muito grande, haverá infinitas distâncias mútuas a serem consideradas, e assim neste caso, deve-se utilizar o limite termodinâmico. A densidade de probabilidade de uma única partícula é dada por (usando $n = 1$ na equação

4.63, e também Z_n da equação A.1)

$$P_N^1(\mathbf{r}_1) = \frac{\int \cdots \int \exp(-\beta U_N(\mathbf{r}^N)) d\mathbf{r}_2 \cdots d\mathbf{r}_N}{\int \cdots \int \exp(-\beta U_N(\mathbf{r}^N)) d\mathbf{r}_1 \cdots d\mathbf{r}_N} \quad (4.68)$$

Usando a origem do sistema na partícula 1 e também considerando um sistema homogêneo com invariância translacional:

$$P_N^1(\mathbf{r}_1) = \frac{\int \cdots \int \exp(-\beta U_N(\mathbf{r}^N)) d\mathbf{r}_2 \cdots d\mathbf{r}_N}{\int d\mathbf{r}_1 \int \cdots \int \exp(-\beta U_N(\mathbf{r}^N)) d\mathbf{r}_2 \cdots d\mathbf{r}_N} = \frac{1}{V} \quad (4.69)$$

Estendendo essa teoria para as n partículas do subsistema, a função de distribuição radial pode ser reescrita como:

$$g_N^{(n)}(\mathbf{r}^n) = V^n P_N^n(\mathbf{r}^n) = V^n \langle \delta(\mathbf{r}_1 - \mathbf{r}'_1) \delta(\mathbf{r}_2 - \mathbf{r}'_2) \cdots \delta(\mathbf{r}_n - \mathbf{r}'_n) \rangle \quad (4.70)$$

Para um subsistema de 2 partículas (ou proteínas), $n=2$ e, assumindo um sistema isotrópico (McDonald e Hansen [1976]):

$$g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = g_N^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|) = V^2 \langle \delta(\mathbf{r}_1 - \mathbf{r}'_1) \delta(\mathbf{r}_2 - \mathbf{r}'_2) \rangle \quad (4.71)$$

$$g_N^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|) \rightarrow g(r) \quad (4.72)$$

Esta é a função de distribuição radial (ou do inglês, *Radial Distribution Function* - RDF) está relacionado com a probabilidade de encontrar uma partícula distante de r de uma partícula referencial. Pela equação (4.67), fazendo $n = 2$, pode ser visto que para uma partícula distante de uma partícula de referência, $g(\mathbf{r}) \rightarrow 1$ quando $r \rightarrow \infty$.

4.4.2 Função de correlação de pares, $c(r)$, e correlação total, $h(r)$

Quando um sistema é composto por mais de uma partícula, elas estão correlacionadas a partir de uma função de correlação total $h(r)$, em que $r = |\mathbf{r}_1 - \mathbf{r}_2|$. Isso significa que as posições entre as partículas dependem uma das outras. Por definição, o fator de correlação total está relacionando com a função de distribuição radial (McDonald e Hansen [1976],

Zhang et al. [2007]):

$$h(\mathbf{r}) = g(\mathbf{r}) - 1 \quad (4.73)$$

Outra maneira de escrever a correlação total é pela equação de Ornstein-Zernick, e neste caso a correlação total é uma função da correlação de pares $c(\mathbf{r})$ (McDonald e Hansen [1976]):

$$h(\mathbf{r}) = c(\mathbf{r}) + n_p \int c(|\mathbf{r} - \mathbf{r}'|) h(\mathbf{r}') d\mathbf{r}' \quad (4.74)$$

Existem diferentes maneiras de resolver a equação de Ornstein-Zernick (4.74), e as soluções são dadas pelas relações de fechamento tais como: Perkus-Yevick, Aproximação de Esfera Média (ou do inglês, *Mean Spherical Approximation*), *Hipernetted-Chain* e a relação de fechamento utilizada neste projeto, Aproximação de Fase Aleatória (ou do inglês, *Random Phase Approximation*- RPA) (McDonald e Hansen [1976], Bostrom et al. [2005]). As relações de fechamento relacionam a função de correlação de pares $c(\mathbf{r})$ com a função de distribuição radial $g(\mathbf{r})$, atuando de maneiras diferentes conforme a relação escolhida. A RPA é a mais conveniente para o objetivo deste estudo porque, para as proteínas utilizadas, será considerado um potencial de esfera dura perturbado por interações devido a presença de outras proteínas. Devemos salientar no entanto que a escolha da RPA é arbitrária, ou seja, a escolha de outra relação de fechamento deveria levar a um outro resultado.

A função de correlação total pode ser rescrita a partir de expansões de *clusters* (McDonald e Hansen [1976], Andersen e Chandler [1972], Mayer e Montroll [1941]). De acordo com essa definição:

$$h(1, 2) = h(\mathbf{r}) = \{\text{Soma de todos os diagramas distintos consistindo de 2 círculos brancos, com índices 1 e 2, círculos pretos e ligações-f, e que são livres de articulações}\} \quad (4.75)$$

$$= \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} + \dots$$

The diagrams represent terms in a series expansion. Each diagram consists of two white circles labeled 1 and 2. Diagram 1 shows a direct bond between 1 and 2. Diagram 2 shows a bond between 1 and a black circle, which is then bonded to 2. Diagram 3 shows a bond between 1 and a black circle, which is bonded to another black circle, which is then bonded to 2. Ellipses indicate further terms in the series.

As ligações-f são substituídas por ligações- h_0 , onde h_0 é a função de correlação total do sistema de referência. Expandindo a função de correlação de pares considerando para um sistema isotrópico:

$$n_p^2 h(1, 2) = n_p^2 h_0(1, 2) + n_p^2 C(1, 2) \quad (4.76)$$

O fator $C(1, 2) = C(\mathbf{r})$ é chamado de potencial renormalizado, ou cadeia de ligação, e tem valor dado por:

$$n_p^2 C(\mathbf{r}) = n_p^2 \sum_{n=1}^{\infty} C^{(n)}(\mathbf{r}) \quad (4.77)$$

Substituindo os valores da função de correlação total, usando a equação (4.73), chega-se a relação:

$$g(\mathbf{r}) = g_0(\mathbf{r}) + C(\mathbf{r}) \quad (4.78)$$

A definição do potencial renormalizado, $n_p^2 C(\mathbf{r})$ é a soma de todas as cadeias que tenham 2 círculos brancos (com índices 1 e 2) e círculos pretos com ligações do tipo ϕ e h_0 , de modo a não haver ligações h_0 consecutivas, assim como mostrado na figura (4.6).

$$\begin{aligned} \rho^2 C^{(1)}(1, 2) = & \text{Diagram 1} + \text{Diagram 2} \\ & + \text{Diagram 3} + \text{Diagram 4} \\ \rho^2 C^{(2)}(1, 2) = & \text{Diagram 5} + \text{Diagram 6} \\ & + \text{Diagram 7} + \text{Diagram 8} \\ & + \text{Diagram 9} + \text{Diagram 10} \\ & + \text{Diagram 11} \end{aligned}$$

The diagrams show the first two terms of the expansion for the renormalized potential. The first term, $\rho^2 C^{(1)}$, includes four diagrams: a direct bond between white circles 1 and 2; a bond between 1 and a black circle, then to 2; a bond between 1 and a black circle, then to another black circle, then to 2; and a bond between 1 and a black circle, then to a black circle, then to another black circle, then to 2. The second term, $\rho^2 C^{(2)}$, includes seven diagrams: a bond between 1 and a black circle, then to 2; a bond between 1 and a black circle, then to another black circle, then to 2; a bond between 1 and a black circle, then to a black circle, then to another black circle, then to 2; a bond between 1 and a black circle, then to a black circle, then to another black circle, then to a black circle, then to 2; a bond between 1 and a black circle, then to a black circle, then to another black circle, then to a black circle, then to a black circle, then to 2; a bond between 1 and a black circle, then to a black circle, then to another black circle, then to a black circle, then to a black circle, then to a black circle, then to 2; and a bond between 1 and a black circle, then to a black circle, then to another black circle, then to a black circle, then to a black circle, then to a black circle, then to a black circle, then to 2.

Figura 4.6: Primeiros 2 termos da equação (4.77)

A interpretação desses diagramas é mais simples do que a interpretação por integrais. Os círculos brancos de índices 1 e 2 são devido a ligação de cadeias $C(1, 2)$ que foi definido anteriormente como $C(r)$. Caso fosse definida como $C(1, 2, \dots, n)$, então haveriam n círculos brancos. As ligações com linhas pretas mostram que os círculos que tem essa ligação, estão relacionados por uma interação do tipo $\tilde{\phi}(\mathbf{q})$. O número total dessas linhas é dada pelo índice n , sobreescrito em $C^{(n)}$. Já ligações rachuradas, são devido a interações do tipo h_0 dadas pela função $\tilde{F}_0(\mathbf{q})$ e de acordo com a definição, não podem haver duas linhas rachuradas consecutivas, sendo $n+1$ o número máximo dessas linhas. O número de círculos pretos está de acordo com o número máximo de linhas pretas e rachuradas e portanto o número total de diagramas diferentes será 2^{n+1} . Conforme o aumento do número n , é mais complicado escrever todos os fatores do diagrama e por isso ele pode ser representado de maneira mais compacta a partir de hipervértices, assim como mostrado na figura (4.7). Esses diagramas são compostos por círculos maiores e círculos menores (brancos e pretos), onde os maiores relacionam com os menores através de funções $\tilde{F}_0(\mathbf{q})$. Já ligações entre círculos grandes são dadas pela função $\tilde{\phi}(\mathbf{q})$, que representa a transformada de Fourier do potencial de perturbação ($\tilde{\phi}(\mathbf{q}) = -\beta F(q)$), em que $\beta F(q)$ está descrito na equação (4.39).

Figura 4.7: Generalização da equação (4.77) dada por hipervértices

Os termos da equação (4.77), usando a notação de hipervértices, podem ser escritos

como:

$$n_p^2 C^{(1)}(\mathbf{r}) = \iint F_0(1, 3) \phi(3, 4) F_0(4, 2) d3 d4 \quad (4.79)$$

$$n_p^2 C^{(2)}(\mathbf{r}) = \iiint F_0(1, 3) \phi(3, 4) F_0(4, 5) \phi(5, 6) F_0(6, 2) d3 d4 d5 d6 \quad (4.80)$$

$\vdots$

$$n_p^2 C^{(n)}(\mathbf{r}) = \int \cdots \int F_0(1, 3) \phi(3, 4) F_0(4, 5) \phi(5, 6) \cdots \\ \cdots F_0(2n, 2n+1) \phi(2n+1, 2n+2) F_0(2n+2, 2) d3 d4 \cdots d(2n+1) d(2n+2) \quad (4.81)$$

Aplicando a transformada de Fourier na equação generalizada (4.81), obtém-se:

$$n_p^2 \tilde{C}(\mathbf{q}) = \sum_{n=1}^{\infty} \left[\tilde{F}_0(\mathbf{q}) \tilde{\phi}(\mathbf{q}) \right]^n \tilde{F}_0(\mathbf{q}) \quad (4.82)$$

Quando $|n_p^2 \tilde{C}^{(n)}(\mathbf{q})| < 1$, a equação (4.82) pode ser resolvida através da série geométrica, equação (A.7), com primeiro termo $a_1 = (\tilde{F}_0(\mathbf{q}))^2 \tilde{\phi}(\mathbf{q})$ e razão $q = \tilde{F}_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})$:

$$n_p^2 \tilde{C}(\mathbf{q}) = \frac{(\tilde{F}_0(\mathbf{q}))^2 \tilde{\phi}(\mathbf{q})}{1 - \tilde{F}_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} \quad (4.83)$$

Tomando a transformada inversa desse fator:

$$C(\mathbf{r}) = \frac{1}{n_p^2 (2\pi)^3} \int e^{i\mathbf{q} \cdot \mathbf{r}} \frac{(\tilde{F}_0(\mathbf{q}))^2 \tilde{\phi}(\mathbf{q})}{1 - \tilde{F}_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} d\mathbf{q} \quad (4.84)$$

Esse fator é muito importante para descrever a função de interferência entre as funções, também chamada de fator de estrutura e será usado nas próximas sessões.

4.4.3 Cálculo do Fator de Estrutura, $S(q)$

A definição do fator de estrutura teórica é dada pela transformação espacial da densidade de partículas (n_p), usando o vetor de espalhamento $\mathbf{q}$ (McDonald e Hansen [1976]):

$$S(\mathbf{q}) = \frac{1}{N} \langle n_p(\mathbf{q}) n_p(-\mathbf{q}) \rangle \quad (4.85)$$

$$S(\mathbf{q}) = \frac{1}{N} \sum_i^N \sum_j^N \langle e^{-i\mathbf{q} \cdot \mathbf{r}_i} e^{i\mathbf{q} \cdot \mathbf{r}_j} \rangle \quad (4.86)$$

separando os termos $i = j$:

$$S(\mathbf{q}) = 1 + \frac{1}{N} \sum_{i \neq j}^N \langle e^{-i\mathbf{q} \cdot \mathbf{r}_i} e^{i\mathbf{q} \cdot \mathbf{r}_j} \rangle \quad (4.87)$$

A exponencial dentro da soma pode ser reescrito em termos de funções delta como:

$$e^{-i\mathbf{q} \cdot \mathbf{r}_i} e^{i\mathbf{q} \cdot \mathbf{r}_j} = \iint e^{-i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_j) d\mathbf{r} d\mathbf{r}'$$

$$S(\mathbf{q}) = 1 + \frac{1}{N} \sum_{i \neq j}^N \left\langle \iint e^{-i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_j) d\mathbf{r} d\mathbf{r}' \right\rangle \quad (4.88)$$

Usando a definição na equação (4.71) e a dispersão de n partículas como $N!/(N-n)! = N(N-1)\dots(N-n+1)$, em que para $n=2$, $N(N-1)$:

$$S(\mathbf{q}) = 1 + \frac{N(N-1)}{NV^2} \iint e^{-i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} g(\mathbf{r} - \mathbf{r}') d\mathbf{r} d\mathbf{r}' \quad (4.89)$$

Considerando no limite termodinâmico, $V, N \rightarrow \infty \implies (N-1) \rightarrow N$, $n_p = N/V$, através de mudanças de variáveis chega-se a (McDonald e Hansen [1976]):

$$S(\mathbf{q}) = 1 + n_p \int e^{-i\mathbf{q} \cdot \mathbf{r}} g(\mathbf{r}) d\mathbf{r} \quad (4.90)$$

Utilizando a equação (4.73) na equação (4.90) é possível obter o fator de estrutura em

função da correlação total como:

$$S(\mathbf{q}) = 1 + n_p \int e^{-i\mathbf{q}\cdot\mathbf{r}} (h(\mathbf{r}) + 1) d\mathbf{r} = 1 + n_p \int e^{-i\mathbf{q}\cdot\mathbf{r}} h(\mathbf{r}) d\mathbf{r} + n_p \int e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$$

$$S(\mathbf{q}) = 1 + n_p \tilde{h}(\mathbf{q}) + (2\pi)^3 n_p \delta(\mathbf{q}) \quad (4.91)$$

O último termo se trata de radiação não desviada e como esta não é acessível experimentalmente, será ignorada a partir desse ponto. Consequentemente, o fator de estrutura será tratado por:

$$S(\mathbf{q}) = 1 + n_p \tilde{h}(\mathbf{q}) \quad (4.92)$$

$$S(\mathbf{q}) = 1 + n_p \int (g(\mathbf{r}) - 1) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (4.93)$$

Utilizando a equação (4.92) é possível reescrever a RDF como função do fator de estrutura como: $n_p(\tilde{g}(\mathbf{q}) - 1) = S(\mathbf{q}) - 1$. Aplicando a transformada de Fourier na equação (4.78) e usando a equação (4.83), chega-se:

$$\begin{aligned} \tilde{g}(\mathbf{q}) &= \tilde{g}_0(\mathbf{q}) + \tilde{C}(\mathbf{q}) \\ \tilde{g}(\mathbf{q}) - 1 &= \tilde{g}_0(\mathbf{q}) - 1 + \tilde{C}(\mathbf{q}) \\ n_p(\tilde{g}(\mathbf{q}) - 1) &= n_p(\tilde{g}_0(\mathbf{q}) - 1) + n_p \tilde{C}(\mathbf{q}) \\ S(\mathbf{q}) - 1 &= S_0(\mathbf{q}) - 1 + n_p \tilde{C}(\mathbf{q}) \\ S(\mathbf{q}) &= S_0(\mathbf{q}) + \frac{1}{n_p} \frac{(\tilde{F}_0(\mathbf{q}))^2 \tilde{\phi}(\mathbf{q})}{1 - \tilde{F}_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} \end{aligned} \quad (4.94)$$

As ligações- ϕ estão relacionadas com o potencial de perturbação, enquanto as ligações h_0 estão relacionadas com a função $\tilde{F}_0(\mathbf{q})$ por (McDonald e Hansen [1976])

$$F_0(\mathbf{r}) = n_p \delta(\mathbf{r}) + n_p^2 h_0(\mathbf{r})$$

$$\tilde{F}_0(\mathbf{q}) = \int \exp(-i\mathbf{q}\cdot\mathbf{r}) F_0(\mathbf{r}) d\mathbf{r} = n_p S_0(\mathbf{q}) \quad (4.95)$$

$S_0(\mathbf{q})$ é o fator de estrutura para o sistema de referência (sem perturbação). Substituindo

a equação (4.95) na equação (4.94), o fator de estrutura será:

$$\begin{aligned}
 S(\mathbf{q}) &= S_0(\mathbf{q}) + n_p \frac{[S_0(\mathbf{q})]^2 \tilde{\phi}(\mathbf{q})}{1 - n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} = \\
 &= \frac{S_0(\mathbf{q}) (1 - n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q}))}{1 - n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} + n_p \frac{[S_0(\mathbf{q})]^2 \tilde{\phi}(\mathbf{q})}{1 - n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})} \\
 S(\mathbf{q}) &= \frac{S_0(\mathbf{q})}{1 - n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q})}
 \end{aligned} \tag{4.96}$$

A equação do fator de estrutura usando a relação de fechamento RPA mostra que haverá uma descontinuidade se, $n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q}) = 1$ mas, se a perturbação do sistema é baixa comparado com o potencial de referência ou se a densidade de partículas é baixa (poucas moléculas por unidade de volume), será garantido para todos valores do vetor de espalhamento, $\mathbf{q}$:

$$n_p S_0(\mathbf{q}) \tilde{\phi}(\mathbf{q}) < 1 \tag{4.97}$$

Satisfazendo assim a condição que $S(\mathbf{q})$ seja positivo para todos valores de vetor de espalhamento $\mathbf{q}$.

4.4.4 Fator de estrutura do sistema de referência

O sistema de referência é regido por potencias do tipo esfera dura e portanto, suas partículas estarão correlacionadas através da correlação de Perkus-Yevick (McDonald e Hansen [1976], Thiele [1963], Wertheim [1963]):

$$c_{PY}(r) = \begin{cases} -\frac{(1+2\eta)^2}{(1-\eta)^4} + \frac{6\eta(1+0.5\eta)^2}{(1-\eta)^4} \left(\frac{r}{\sigma}\right) - \frac{\eta(1+2\eta)^2}{2(1-\eta)^4} \left(\frac{r}{\sigma}\right)^3, & \text{se } r \geq \sigma \\ 0, & \text{se } r < \sigma \end{cases} \tag{4.98}$$

em que η é o fator de empacotamento, $\eta = n_p \pi \sigma^3 / 6$, e σ é o diâmetro da proteína. Esse fator será aproximado por uma correlação de núcleo vazio (do inglês, *empty-core*) (Croxtton

[1975]), onde:

$$c_{EC}(r) = c_{PY}(\sigma) = \frac{\eta(3 - \eta^2) - 2}{2(1 - \eta)^4} \quad (4.99)$$

Aplicando a transformada de Fourier na equação de Ornstein-Zernick (4.74), chega-se a $\tilde{h}(\mathbf{q}) = \tilde{c}(\mathbf{q}) + n_p \tilde{c}(\mathbf{q}) \tilde{h}(\mathbf{q})$ e portanto a correlação de pares pode ser escrita como:

$$\tilde{c}(\mathbf{q}) = \frac{\tilde{h}(\mathbf{q})}{1 + n_p \tilde{h}(\mathbf{q})} \quad (4.100)$$

A partir da equação (4.93), também é possível obter a correlação total como (lembrando que está sendo feito para S_0 , pois estamos no sistema de referência):

$$\tilde{h}(\mathbf{q}) = \frac{S_0(\mathbf{q}) - 1}{n_p} \quad (4.101)$$

Substituindo esse valor na equação (4.100) chega-se na relação do fator de estrutura com a correlação de pares:

$$\begin{aligned} \tilde{c}(\mathbf{q}) &= \frac{S_0(\mathbf{q}) - 1}{n_p S_0(\mathbf{q})} \\ S_0(\mathbf{q}) &= \frac{1}{1 - n_p \tilde{c}(\mathbf{q})} \end{aligned} \quad (4.102)$$

Utilizando a correlação de pares de núcleo vazio, e fazendo a transformada de Fourier nos respectivos intervalos:

$$n_p \tilde{c}(\mathbf{q}) = 4\pi n_p \int_0^\infty c(r) \frac{\text{sen}(qr)}{qr} r^2 dr \quad (4.103)$$

$$\begin{aligned} n_p \tilde{c}(\mathbf{q}) &= 4\pi n_p \int_0^\sigma c(r) \frac{\text{sen}(qr)}{qr} r^2 dr + 4\pi n_p \int_\sigma^\infty c(r) \frac{\text{sen}(qr)}{qr} r^2 dr \\ n_p \tilde{c}(\mathbf{q}) &= \frac{4\pi n_p}{q} \frac{[\eta(3 - \eta^2) - 2]}{2(1 - \eta)^4} \int_0^\sigma \text{sen}(qr) r dr \end{aligned} \quad (4.104)$$

A solução dessa integral está na equação (A.13), resultando em:

$$\begin{aligned} n_p \tilde{c}(\mathbf{q}) &= \frac{2\pi n_p}{q} \frac{[\eta(3 - \eta^2) - 2]}{(1 - \eta)^4} J_1(q\sigma) \sigma^2 \\ n_p \tilde{c}(\mathbf{q}) &= 2\pi n_p \sigma^3 \frac{[\eta(3 - \eta^2) - 2]}{(1 - \eta)^4} \frac{J_1(q\sigma)}{q\sigma} \\ n_p \tilde{c}(\mathbf{q}) &= 12\eta \frac{[\eta(3 - \eta^2) - 2]}{(1 - \eta)^4} \frac{J_1(q\sigma)}{q\sigma} \end{aligned} \quad (4.105)$$

Foi usado a definição de fração volume, onde o valor 12η é igual ao primeiro fator da equação acima. Substituindo a equação (4.105) na equação (4.102) pode ser encontrado o fator de estrutura da partícula de referência:

$$[S_0(q)]^{-1} = 1 - \frac{12\eta [\eta(3 - \eta^2) - 2]}{(1 - \eta)^4} \frac{J_1(q\sigma)}{q\sigma} \quad (4.106)$$

Esta equação, será utilizada na equação (4.96). Basta encontrar a transformada de Fourier da perturbação e assim tem-se o fator de estrutura:

$$\tilde{\phi}(q) = \mathcal{F}(-\beta F(r)) = - \int \beta F(r) e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} = -4\pi \int_0^\infty \beta F(r) \frac{\sin(qr)}{qr} r^2 dr \quad (4.107)$$

Para a última igualdade foi usado que o vetor de espalhamento está em uma direção onde o ângulo entre $\mathbf{q}$ e $\mathbf{r}$ é o mesmo ângulo θ das coordenadas esféricas, e portanto a exponencial pode ser substituída pelo fator $\sin(qr)/qr$. Guinier e Fournet [1955] explicam essa substituição a partir de sistemas isotrópicos, onde $\langle e^{i\mathbf{q} \cdot \mathbf{r}} \rangle = \sin(qr)/qr$.

Inicialmente o potencial de perturbação será descrito como a soma de um potencial de Coulomb blindado, equação (4.13), e um potencial atrativo tipo Yukawa, descrito no trabalho de Narayanan e Liu [2003] como a soma dos potenciais atrativos osmótico e de Van der Waals:

$$U_Y(r) = -J\sigma \frac{e^{-\frac{(r-\sigma)}{d}}}{r} \quad (4.108)$$

em que J é o potencial atrativo no contato ($r = \sigma$) e d é o alcance deste potencial. Esse método também foi aplicado em outros artigos encontrados na literatura (Barbosa et al. [2010, 2013]). Após estas análises, será acrescentado o potencial devido a regulação de

cargas e portanto o potencial total de perturbação será descrito por como a soma destes três componentes. Assim para resolver a equação (4.107), deve-se obter a transformada do potencial total, dado por (vide apêndice B):

$$\tilde{\phi}(q) = -\beta \left(\tilde{U}_{SC}(q) + \tilde{U}_{CR}(q) + \tilde{U}_Y(q) \right) \quad (4.109)$$

onde:

$$\beta \tilde{U}_{SC}(q) = \frac{4\pi}{q} \frac{\langle Q \rangle^2 l_B}{(1 + 0.5\kappa\sigma)^2} \frac{\kappa}{\kappa^2 + q^2} \left(\sin(q\sigma) + \frac{q}{\kappa} \cos(q\sigma) \right) \quad (4.110)$$

$$\beta \tilde{U}_{CR}(q) = -\frac{iE(q)\pi}{q} \frac{l_B^2}{(1 + 0.5\kappa\sigma)^4} (C^2 + 2C \langle Q \rangle^2) e^{2\kappa\sigma} \quad (4.111)$$

i é o número imaginário tal que $i^2 = -1$, e $E(q)$ é dado por:

$$E(q) = E_1((2\kappa + iq)\sigma) - E_1((2\kappa - iq)\sigma) \quad (4.112)$$

$E_1(z)$ é a função exponencial integral definida em (A.14). O potencial $\tilde{U}_{CR}(q)$ é devido a regulação de cargas, enquanto o $\tilde{U}_{SC}(q)$ é devido ao potencial Coulombiano. A parte atrativa da transformada de Fourier é dada por:

$$\beta \tilde{U}_Y(q) = -\frac{4\pi J\sigma}{q} \frac{d}{1 + (qd)^2} (\sin(q\sigma) + qd \cos(q\sigma)) \quad (4.113)$$

Estas equações, juntamente com as equações (4.96) e (4.106) darão o valor do fator de estrutura através da relação de fechamento RPA.

4.5 Aquisição de medidas experimentais

4.5.1 Preparação de amostras

As amostras de BSA e citocromo C foram preparadas da mesma maneira. Primeiramente, 100mg de proteína foram pesados na balança da empresa *Bel engineering* modelo M214A e colocados em diferentes eppendorfs. A seguir, tampões compostos de **Acetato**

Fosfato Borato a 20 mM, $pH \approx 7$, foram inserido em diferentes tubos Falcon de 15 ml e ajustado pH, adicionando NaOH (5M) para tampões básicos e HCl(17.5%) para tampões ácidos. Os valores de pH foram conferidos utilizando o pHmetro da empresa Bel *engineering* modelo W3B. O processo de mistura será explicado a seguir. A proteína foi colocada em solução somente momentos antes da realização do experimento, pois proteínas nesse estado tem curta duração (pode se quebrar facilmente).

Quando proteínas são imersas em líquidos, esta irá ocupar um determinado volume dentro dessa solução. Devido a isso, deve-se excluir esse volume ocupado pela proteína. Assim pode-se escrever o volume total (V_T) como a soma do volume ocupado pela proteína (V_p) com o volume de líquido (V_l) adicionado:

$$V_T = V_p + V_l \rightarrow V_l = V_T - V_p \quad (4.114)$$

Para proteínas globulares é utilizado uma aproximação válida para o volume específico, dada por $\nu = 0.74$ ml/g (Tardieu et al. [1999]) e portanto, o volume ocupado de proteína será $V_p(\text{ml}) = 0.74m$, onde m é a massa de proteína adicionada na solução (em gramas). Para uma massa, $m=0.1$ g (100 mg), terá-se um volume de proteínas $V_p = 74 \mu\text{l}$. Assim, para que se tenha uma solução de 100 mg/ml de concentração proteica, deve-se adicionar 926 μl de tampão. Para a concentração de 50 mg/ml, a solução de 100 mg/ml foi diluída no tampão na proporção 1:1. Para BSA, as concentrações de 25 e 5 mg/ml foram diluídos das soluções de 50 e 25 mg/ml com tampão nas proporções 1:1 e 1:4, respectivamente. Já para o citocromo C, as concentrações de 25, 10 e 5 mg/ml foram diluídos das soluções de 50, 25 e 10 mg/ml com tampão nas proporções 1:1 e 2:3 e 1:1, respectivamente.

4.5.2 Medidas de SAXS

As medidas de SAXS foram realizadas no Laboratório Nacional de Luz Síncrotron (LNLS- Campinas) na linha SAXS-1. Para esse trabalho, a distância entre o porta amostra e o detector foi de aproximadamente 1000 mm, conforme figura (figura 4.2), resultando em um vetor de espalhamento no intervalo $0.1 - 5.0 \text{ nm}^{-1}$. O porta amostra foi colocado em

temperatura fixa de 298 K, sendo controlado por um banho circulante de água.

O porta amostra foi montado usando mica e teflon para a vedação. Ele foi desenvolvido pelo LNLS e é chamado de PALMI, Porta Amostra de Líquidos com Mica (LNLS [2019]), figura (4.8). Após cada coleta de dados, ele era limpo bombeando, de 30 à 50 ml de água Milli-Q na parte interna. O volume de amostra utilizado em cada experimento foi de $\sim 300 \mu\text{l}$.

Figura 4.8: Diferentes tipos de porta amostra.

As medidas de espalhamento das proteínas foram obtidas por 100 ou 300 segundos de exposição para a radiação, dependendo da concentração proteica. A correção das curvas de espalhamento foi feita pela contribuição do tampão e também pela atenuação da respectiva mostra. As curvas mostrados ao longo deste trabalho não estão em escala absoluta, mas sim em escala relativa, do modo como sai do programa de aquisição de dados.

4.6 Cálculos Teóricos aplicados

4.6.1 Carga e parâmetro de regulação de carga em função do pH

Para se obter o valor teórico da carga das proteínas, há autores como por exemplo, Lund e Jönsson [2013], Teixeira et al. [2010], Franco e de Alcântara Pessoa Filho [2011] que mostraram expressões com as quais eram possível obter a carga total de uma proteína dependendo do pK_a de cada grupo ionizável e do pH da solução. Assim sendo, para calcular o parâmetro de regulação de carga em função do pH basta utilizar a equação (4.38). Em contrapartida a isso, neste trabalho foi utilizado o programa Propka 3.0 (Olsson et al.

[2011]). O programa utiliza um arquivo PDB de uma proteína, calcula o valor de pK_a para cada resíduo (Li et al. [2005]) e a partir da equação de Henderson-Hasselbach, calcula-se a carga total, tanto para o caso no qual a proteína está enovelada quanto para o caso desenovelado. Com os dados de carga em função do pH, foi feito um ajuste de uma função na curva e a partir da equação (4.38) foi obtido o valor do parâmetro de regulação de carga.

Vale ressaltar que com este programa, diferentemente do simulado através de Monte Carlo, não era possível obter os dados de carga em função do pH para distintas concentrações de sal e de proteínas. Devido a isso, os resultados neste trabalho foram obtidos a partir das predefinições do próprio programa (Olsson et al. [2011]).

As estruturas para a BSA e citocromo C foram obtidos a partir do banco de dados de proteínas do site de Pesquisa Colaborativa para Bioinformática Estrutural (do inglês, *Research Collaboratory for Structural Bioinformatics- RCSB* [2000]), com o primeiro tendo identificação **4F5S** enquanto para o segundo, **1AKK**. O programa também proporciona o valor do ponto isoelétrico das proteínas, sendo este o valor de pH para o qual a carga efetiva média da proteína é nula.

4.6.2 Análise das curvas experimentais

Para esse trabalho, foi desenvolvido um programa para fazer análises de curvas de SAXS. Com o programa é possível calcular a curva de intensidade teórica para geometrias esféricas e elipsoidais.

Além do programa desenvolvido e utilizado inicialmente nas análises das curvas de SAXS das proteínas modelo estudadas, também foi utilizado o programa *GENFIT* (Spinuzzi et al. [2014]), desenvolvido pelo prof Dr. Francesco Spinuzzi, da *Università Politecnica delle Marche*, Ancona, Itália, sendo utilizado por diversos pesquisadores no meio científico. Com ajuda do Prof. Spinuzzi, foram implementados dois novos modelos em seu programa:

i) *Interacting PDBs: All-Atoms Form Factor With Solvation Shell Of Dummy Atoms And Multipole Expansion Average And HS+Double Yukawa+Charge Fluctuation Potential*

Under RPA- Esse modelo calcula o fator de forma da proteína a partir do seu arquivo e calcula o fator de estrutura utilizando a relação de fechamento RPA, considerando o termo de regulação de cargas, explicitado nas sessões anteriores.

ii) *Two density-level tri-axial ellipsoid and HS+Double Yukawa+Charge Fluctuation Potential Under RPA*- O fator de forma da proteína é calculada a partir de um elipsoide com dupla densidade eletrônica e calcula o fator de estrutura utilizando a relação de fechamento RPA, considerando o termo de regulação de cargas.

Este segundo modelo não foi utilizado na realização deste trabalho, mas está descrito de modo a exibir que também é possível simular as curvas de SAXS com regulação de carga utilizando um elipsóide como formato de proteína. A vantagem na utilização do *GENFIT* (a partir desses modelos) é que a determinação dos parâmetros é feita a partir de um intervalo (designado pelo usuário) e, a melhor solução é obtida através da minimização do funcional χ^2 reduzido (Ortore et al. [2009]), definido na equação (A.11). Além disso, é possível relacionar variáveis para diferentes curvas através de um ajuste global, onde múltiplas curvas são ajustadas simultaneamente, sendo esse método utilizado nesta dissertação. É importante ressaltar também que esses modelos não foram escritos para a interface do programa, sendo possível utilizar apenas partir do *prompt* de comando do computador. Tanto para BSA quanto para o citocromo C foi aplicado o modelo que utiliza a estrutura cristalográfica da proteína para o cálculo do fator de forma.

No programa, o parâmetro de alcance de potencial foi fixado em aproximadamente 10% do valor do diâmetro da proteína. Esse método também foi utilizado no artigo de Barbosa et al. [2010]. Quanto ao diâmetro da proteína, por esta ser globular, ou seja, com um volume específico de $\nu_{esp} = 0.74 \text{ cm}^3/\text{g}$ (Tardieu et al. [1999]), foi utilizado a seguinte função para o ajuste:

$$\sigma(\text{\AA}) = 2 \left(\frac{3}{4\pi} \frac{MM}{0.6022\nu_{esp}} \right)^{1/3} \quad (4.115)$$

em que, MM é a massa molecular e ν_{esp} é o volume específico da proteína.

Resultados e Discussões

5.1 Carga e Capacitância em função do pH

Os gráficos de carga em função do pH para BSA e para o citocromo C calculados através do programa *PROPKA 3.0* estão ilustrados na figura (5.1a) e (5.1c), respectivamente. Enquanto que suas respectivas capacitâncias, calculadas a partir do gráfico de cargas e através da equação (4.38), estão na figura (5.1b - BSA) e (5.1d - citocromo C). As retas sólidas em vermelho indicam os valores de pH em que a carga total das proteínas é nulo. A este valor de pH chama-se ponto isoelétrico da proteína, pI. Os valores de pI calculado para a BSA foi de 5.8, figura (5.1a), enquanto que para Citocromo C de 10.4, figura (5.1c).

Para BSA, é reportado na literatura um valor de ponto isoelétrico entre 4.8 e 5.6 (Barbosa et al. [2010], Tanford e Buzzell [1956]). Rezwan et al. [2004], por exemplo, calcularam o ponto isoelétrico teórico da BSA como sendo pI=5.56. Assim, o valor calculado pelo programa está em acordo ao encontrado por esses autores. Já para o citocromo C de coração equino, é possível encontrar valores de ponto isoelétrico entre 10.0 e 10.5 (Miyashita et al. [2013], Malmgren et al. [1978]), mostrando novamente a concordância entre o valor calculado aqui e os reportados na literatura.

Apesar disso, para o gráfico de carga da BSA, os valores calculados pelo programa PropKA 3.0 estão superestimados comparando com os obtidos por simulação de Monte Carlo por Barroso da Silva et al. [2014] em diferentes concentrações de sais. Como consequência dessa diferença, os valores de capacitância também são maiores do que os calculados por eles. Isso se dá porque o programa PropKa não considera interações eletrostáticas

Figura 5.1: Representação teórica de carga e capacitância em função do pH para BSA (a), (b) e Citocromo C (c), (d).

dos sítios tituláveis com cargas presentes no sistema (Barroso da Silva e MacKernan [2017]), enquanto que através de simulações de Monte Carlo, essas interações são consideradas.

A carga efetiva das proteínas mostram um comportamento similar entre elas. Tanto no início quanto no final das curvas a carga está chegando a um valor limite mostrando um comportamento assintótico. Como o aumento do pH, os valores de carga diminuem, sendo que para valores próximos do ponto isoelétrico, a BSA e o citocromo C tem um comportamento parecido. Isso pode ser analisado através dos gráficos da capacitância, figuras (5.1b e 5.1d), onde para esses valores de pH são encontrados valores de capacitância próximos à 5. Nestes gráficos também é possível perceber a ocorrência de picos gerados pela inclinação na curva de carga em função do pH. Eles estão relacionados com a quantidade de um respectivo resíduo em cada proteína.

5.2 Albumina de Soro Bovina

5.2.1 Caracterização estrutural da BSA

Para a BSA, foram realizados experimentos de *SAXS* em diferentes condições, variando-se o pH (5.8, 6.8 e 8.6) e a concentração da proteína (5, 25, 50 e 100 mg/ml). Os dados são mostrados na figura (5.2a, 5.2c, 5.2e). Para os três valores de pH utilizados, as curvas estão dispostas conforme as concentrações de proteínas em solução. Esse resultado está de acordo com o esperado, já que a intensidade medida é proporcional a concentração de proteínas (equação 4.44). Já as figuras (5.2b, 5.2d e 5.2f) mostram as mesmas curvas de *SAXS*, mas agora normalizadas pelas respectivas concentrações. A normalização das curvas mostra que além da concentração, existem outros fatores influenciando seu formato, já que existe uma inversão na disposição das curvas. Esse efeito ocorre devido a presença do fator de estrutura, $S(q)$, que depende da concentração proteica (será discutido a seguir).

A figura (5.3) mostram os gráficos de Kratky para os mesmos sistemas reportados anteriormente, ou seja, BSA em diferentes valores de pH e concentração crescente. Percebe-se que todos os gráficos tem o formato de “sino”, indicando que a BSA está com forma globular. Independentemente da concentração e do pH, além da forma de sino, nota-se que há valores negativos de intensidade para pH 5.8 na concentração 5 mg/ml. Isso pode ser explicado, por exemplo, pela subtração incorreta do tampão deste. É interessante perceber que as curvas de Kratky para todas as concentrações mostraram o mesmo comportamento e, portanto, os dados serão analisados de modo independente da concentração, ou seja, o fator de forma da proteína se manterá igual. Também deve ser ressaltado que nessas curvas estão contidas informações sobre a $S(q)$ sendo que para os cálculos de raio de giro e tamanho, a proteína deve estar diluída em solução. Assim, serão utilizados apenas os dados na concentração de 5 mg/ml.

Utilizando o programa *Primus* do pacote *ATSAS* (Konarev et al. [2003]), foram calculadas as funções $p(r)$, para os sistemas compostos de BSA, dos quais foram calculados valores de raio de giro e tamanho da proteína. O programa calcula o raio de giro por dois métodos: (i) Utilizando a curva da função $p(r)$, equação (4.52); (ii) Utilizando a aproximação de Gui-

Figura 5.2: Curvas experimentais de BSA para pH=5.8, 6.8 e 8.6 para intensidade x vetor de espalhamento (lado esquerdo) e intensidade normalizado (pela concentração) x vetor de espalhamento (lado direito).

nier, tomando pontos experimentais nas condições onde $qR_g < 1.3$ (condição de Guinier), equação (4.56). Os resultados gerados encontram-se na tabela (5.1), com os respectivos gráficos de $p(r)$.

Nagy et al. [2016] encontram o valor de raio de giro da BSA de 27.6 Å para pH 6.9,

Figura 5.3: Curvas de Kratky nos valores de pH 5.8, 6.8 e 8.6 para BSA nas concentrações (a) 5 mg/ml; (b) 25 mg/ml; (c) 50 mg/ml; (d) 100 mg/ml.

Tabela 5.1 - Parâmetros calculados para raio de giro e D_{max} da BSA.

pH	Raio de giro (nm)-Guinier	Raio de giro (nm)-p(r)	D_{max} (nm)
5.8	3.05(5)	3.05(5)	11(1)
6.8	2.98(5)	2.98(5)	10.9(1.1)
8.6	2.89(4)	2.89(4)	9.1(1.0)

Legenda: Valores encontrados pelo programa *Primus* a partir das curvas de intensidades.

enquanto que Bendedouch e Chen [1983] encontram um valor de 30.5 Å em pH 5.1. Comparando esses dados com os da tabela (5.1), pode-se perceber que os resultados experimentais obtidos no presente trabalho estão de acordo com os obtidos pelos dois métodos. Também é possível notar que dentro do erro estatístico, os valores são equivalentes entre si, mostrando que apesar de variar o pH, tanto o raio de giro, quanto o tamanho da proteína não são alterados.

Figura 5.4: Cálculo de $p(r)$, D_{max} e Raio de Giro utilizando o programa *Primus* para BSA em pH 5.8 (topo), 6.8 (meio) e 8.6 (baixo).

Referente aos gráficos das funções $p(r)$, figura (5.4), é possível notar que as funções são praticamente simétricas com relação ao valor r , indicando assim que a proteína pode ser aproximada por uma esfera. Além disso, o valor calculado de D_{max} está entre 9.1 à 11 nm, sendo que Barbosa et al. [2013] encontram $D_{max} = 120(10)$ Å para pH 5.4. Assim será utilizado o monômero da BSA para calcular o fator de forma da proteína.

A curva de distribuição de pares de distâncias de concentração 5 mg/ml e pH 5.8, figura (5.4) teve sua função calculada até $q_{max} = 1.71 \text{ nm}^{-1}$, pois a partir desse valor, os dados eram negativos e não aparecem no gráfico de logaritmo. Contudo, deve-se prestar atenção que a função $p(r)$, descrita na equação (4.49), será truncada em um valor de q_{max} , calculando a integral no intervalo $[0, q_{max}]$, e portanto, quanto maior for o valor de q_{max} , menos informações serão perdidas por essa aproximação.

Para as concentrações de BSA maiores que 5 mg/ml, não é possível negligenciar as funções de interferência entre as proteínas. Para realizar esta etapa, as análises serão separadas em 3 partes:

- (i) Análise sem capacitância, $C = 0$;
- (ii) Análise sem potencial de contato, $J = 0$;
- (iii) Análise com capacitância e potencial de contato, $C, J \neq 0$.

5.2.2 Ajuste das curvas com $C=0$

Foi utilizado o programa *GENFIT* de modo a realizar o cálculo de intensidade por vetor de deslocamento, através dos fatores de forma e de interferência, sendo o primeiro calculado a partir da estrutura cristalográfica e o segundo a partir dos potenciais de interação. Foram efetuados ajustes globais das três curvas com diferentes concentrações para cada pH. A escolha foi feita de modo a explicitar o resultado para altas concentrações (25, 50 e 100 mg/ml) nos valores de pH 5.8 e 6.8, enquanto que para o pH 8.6 foi escolhido também a baixa concentração (5 mg/ml). As curvas de SAXS com os respectivos ajustes estão expostos na figura (5.5), enquanto que os parâmetros dos ajustes estão mostradas na

tabela (5.2).

Figura 5.5: Curvas experimentais para BSA em (a)pH=5.8; (b)pH=6.8; (c) pH=8.6, com os respectivos ajustes.

Tabela 5.2 - Tabela com os parâmetros encontrados a partir do ajuste dos dados da BSA para concentrações 5, 25, 50 e 100 mg/ml para pH entre 5.8 e 8.7

pH	5.8			6.8			8.6		
[BSA](mg/ml)	25	50	100	25	50	100	5	25	50
$J_0(k_B T)$	4.0(7)	2.0(7)	0.1(7)	3.0(4)	2.0(5)	1.0(5)	13.0(5)	2.0(5)	0.1(5)
$d_0(\text{Å})$	6.1(3)			6.7(3)			6.1(3)		
$\sigma_0(\text{Å})$	65(1)			68.7(1.3)			65.2(7)		
$ Z_0 (e)$	7.9(1.3) / 8(1)			10.2(6) / 13(2)			12.7(6) / 20(2)		
$C_0(e^2)$	0			0			0		
$ Z_{teorico} (e)$	0			11.4			22.8		
$C_{teorico}(e^2)$	5.8			5			1.7		

Legenda: Os valores de carga em vermelhos estão relacionados aos encontrados por Barbosa et al. [2010] para os valores de pH 5.4, 7 e 8 respectivamente.

Para realizar os ajustes, a massa molecular da BSA foi fixada em 66.4kDa e o volume

específico da proteína, ν_{esp} , foi ajustado para um intervalo entre 0.69 e $0.76 \text{ cm}^3/\text{g}$ de modo que o diâmetro da esfera efetiva varia entre 65.2 a 67.3 Å. Já o potencial no contato (J) e o módulo da carga superficial média ($|Z|$) foram deixados como parâmetros livres, o primeiro definido em um intervalo entre 0 e $40(k_B T)$, enquanto o segundo, variando entre 0 e $30(e)$.

Os resultados da figura (5.5) mostram alguns fatores interessantes para serem discutidos: (i) Deve-se notar que para fazer a expansão multipolar, equação (4.18), foi suposto que o sistema deveria estar diluído, mas a partir dos resultados obtidos, mostrados na figura (5.5), é observado que as equações também se aplicam à sistemas concentrados. Esse fato também foi demonstrado na literatura por Barbosa et al. [2010]; (ii) O modelo adotado para a análise é capaz de descrever satisfatoriamente as curvas de SAXS até $q \sim 0.5 \text{ Å}^{-1}$.

Utilizando o gráfico de carga em função do pH, mostrado na figura (5.1a) é possível adquirir os valores (teóricos) de carga para os valores de pH utilizados no experimento. Assim, é possível perceber que somente para o dado no qual o pH é quase neutro o resultado se adequou, enquanto que para outros valores de pH os valores teóricos encontrados não retratam o resultado experimental. Comparando os ajustes encontrados em Barbosa et al. [2010] com os valores encontrados neste projeto, a carga obtida para pH=5.4 foi $|Z| = 8(1)$, e esse resultado se mostrou equivalente ao valor apresentado nesse trabalho para pH=5.8, $|Z| = 7.9(1.3)$. Enquanto isso, para valores de pH mais distantes do pI da proteína, por exemplo 8.0, o valor encontrado no artigo para a carga foi de $|Z| = 20(2)$, bem diferente do valor obtido nas curva da figura (5.5), $|Z| = 12.7(6)$. Uma possível razão para essa diferença seria a presença de diferentes sais na solução que alteram o valor de carga, assim como mostrado por Medda et al. [2012]. Bostrom et al. [2005] calcula a carga superficial da BSA (em módulo) de 9.8, 17.8 e 22.7 para os valores de pH 5.5, 7.0 e 8.0 respectivamente. Já Rezwan et al. [2004] obtiveram cargas (em módulo) de 7.8, 18.0 e 28.9 para os valores de pH 6.0, 7.0 e 8.0. Comparando esses resultados com os encontrados na tabela (5.2) é possível perceber que os valores encontrados para pH perto do pI são equivalentes enquanto que para pH mais distantes do pI há uma notável diferença nos valores. Todos os dados encontrados pelos autores foram colocados na figura (5.6).

Figura 5.6: Comparações entre os valores de carga obtido em função do pH para os diferentes autores.

5.2.3 Diferenças dos fatores de estrutura

No programa *GENFIT* também é gerado um arquivo contendo os dados de fator de estrutura, calculado segundo a equação (4.96). Os resultados estão representados na figura (5.7). A partir dos gráficos é possível notar diferenças entre o fator de estrutura medido (que depende do fator de amplitude) e o fator de estrutura e como essas funções variam com a mudança de pH. Um primeiro aspecto a notar nos gráficos é a pequena diminuição do primeiro pico ($0.05 < q < 0.1$) para a curva de fator de estrutura medido com relação ao fator de estrutura. Como esta redução é menor do que 5% do valor, podemos inferir que a diferença entra dentro da incerteza experimental, indicando que os valores podem ser considerados iguais. Observando a curva de 5 mg/ml à pH 8.6, também pode-se inferir o mesmo, indicando que para essa concentração o fator de estrutura é igual a 1, ou seja, esse sistema pode ser considerado diluído. Os casos nos quais o valor do fator de interferência está no intervalo $0.95 < S(q) < 1.0$ indicam que o fator de forma é predominante na curva de espalhamento e portanto as curvas de intensidade podem ser aproximadas por:

$$I(q) = \gamma n_p P(q) + B \quad (5.1)$$

Figura 5.7: Curvas do fator de estrutura (lado esquerdo) e fator de estrutura medido para BSA em diferentes concentrações.

5.2.4 Ajuste das curvas com $J=0$

Nesta sessão as curvas de SAXS são descritas de modo que todo o potencial atrativo do sistema será inferido na flutuação de carga (ou seja, ao termo de regulação de carga).

Tabela 5.3 - Parâmetros encontrados a partir do ajuste dos dados da BSA para concentrações 25, 50 e 100 mg/ml para pH entre 5.8 e 8.6.

pH	5.8			6.8			8.6		
[BSA](mg/ml)	25	50	100	25	50	100	5	25	50
$J_1(k_B T)$	0			0			0		
$\sigma_1(\text{\AA})$	65.3(1)			67.3(1)			65.2(7)		
$ Z_1 (e)$	9.5(1)			12.0(1)			11.3(1)		
$C_1(e^2)$	94.8(1.1)	77.5(5)	41.0(1.6)	111.5(1.1)	89.2(6)	62.7(7)	145(3)	44(2)	0.01(1)

Legenda: Potencial atrativo devido exclusivamente à regulação de carga. O valor de d não foi colocado, pois independente do seu valor ($d > 0$), o potencial atrativo tipo Yukawa será nulo devido à $J=0$.

Utilizando os mesmos intervalos para todos os parâmetros, foi calculado a curva que melhor se ajustava aos dados experimentais. Para isso, o fator J do potencial atrativo de Yukawa foi anulado, enquanto o valor da capacitância foi escolhido para estar no intervalo de 0 à 200 (e^2). Os resultados dos ajustes estão mostrados nas figuras (5.8 e 5.9) já os parâmetros dos ajustes descritos na tabela (5.3).

Como pode ser observado nas figuras (5.8 e 5.9) os ajustes encontrados se mostram de acordo com os dados experimentais para todos os valores de pH e concentrações estudadas. Comparando os parâmetros como o modelo descrito na sessão anterior onde $C = 0$ com o apresentado agora onde $J = 0$, pode-se perceber algumas variações quanto ao valor do módulo da carga encontrado. Apesar disso, como os valores estão compreendidos para em um intervalo de confiança de até 3σ , podem ser considerados equivalentes (nível de confiança de aproximadamente 99%, Vuolo [2002]). Vale notar que o ajuste encontrado para pH 8.6 resultou em um valor de carga menor do que para pH 6.8 indicando que apesar dos ajustes encontrados se adequarem as curvas, a solução pode não ter validade, pois como mostrado na figura 5.1, a carga diminui com o aumento do pH. Deste modo, o programa *GENFIT* encontrou uma solução matemática para a curva mas sem sentido físico claro.

Já os parâmetros encontrados para a capacitância mostraram-se altos, quando comparados com os valores teóricos calculados segundo a figura (5.1b). Os valores de capacitância capazes de descrever os dados experimentais deveriam ser até 20 vezes maiores que o esperado. Isso ocorre pois neste caso não foi considerado o efeito de dipolo (que para a BSA será um valor alto comparado com o de outras proteínas), e como mostrado por Barroso da

Figura 5.8: Curvas experimentais com ajustes sem capacitância (linha preta) e sem par potencial atrativo (linha amarela) com pH=5.8 para BSA nas concentrações (a) 25 mg/ml; (b) 50 mg/ml ; (c) 100 mg/ml

Silva et al. [2006, 2014], no pI, o valor do momento de dipolo é $\mu = 297$.

Devido a esses resultados, foi testado o comportamento dos ajustes para valores intermediários entre $C=0$ e os valores dispostos na tabela (5.3). Esses dados serão descrito na próxima sessão.

5.2.5 Ajuste das curvas com potencial de contato e capacitância

Valendo-se dos ajustes encontrados nas seções anteriores foram realizados novos ajustes utilizando os mesmos parâmetros, e fixando a capacitância nos valores $C=16$, 25 e 36 , de modo que o valor não ultrapassasse o valor de capacitância descrito na tabela (5.3). Os parâmetros relativos aos melhores ajustes encontram-se na tabela (5.4). Os dados

Figura 5.9: Curvas experimentais da BSA com ajustes para pH=6.8 nas concentrações (a) 25 mg/ml; (c) 50 mg/ml ; (e) 100 mg/ml e para pH=8.6 nas concentrações (b) 5 mg/ml; (d) 25 mg/ml ; (f) 50 mg/ml.

encontrados estão de acordo com os dados experimentais, assim como mostrado as figuras (5.8 e 5.9). Novamente é possível perceber uma alteração nos valores da carga, mas assim como explicado anteriormente, por estarem dentro de uma incerteza de até 3σ pode-se inferir que dos resultados são equivalentes.

Tabela 5.4 - Parâmetros encontrados a partir do ajuste dos dados da BSA para concentrações 5,25, 50 e 100 mg/ml para pH entre 5.8 e 8.7.

pH	5.8			6.8			8.6		
[BSA](mg/ml)	25	50	100	25	50	100	5	25	50
$J_2(k_B T)$	3.0(5)	2.0(1)	0.1(5)	4(1)	2.0(8)	1.0(5)	11.0(1.4)	3.0(8)	1.0(4)
$d_2(\text{\AA})$	6.1(4)			6.5(3)			6.1(3)		
$\sigma_2(\text{\AA})$	65(1)			67.4(6)			65(2)		
$ Z_2 (e)$	8.3(4)			11(1)			13.5(9)		
$C_2(e^2)$	16			16			16		
$J_3(k_B T)$	3.0(5)	2.0(4)	0.1(1)	3.0(4)	2.0(5)	1.0(5)	8(1)	2.0(8)	0.0(4)
$d_3(\text{\AA})$	6.1(1)			6.5(2)			6.1(3)		
$\sigma_3(\text{\AA})$	65.2(3)			68.7(1)			68.7(1)		
$ Z_3 (e)$	8.9(3)			11.4(9)			12(1)		
$C_3(e^2)$	25			25			25		
$J_4(k_B T)$	2.0(5)	1.0(5)	0.1(5)	3.0(6)	1.0(8)	0.1(4)	8.0(4)	1.0(1)	0.0(1)
$d_4(\text{\AA})$	6.8(4)			6.7(2)			6.5(1)		
$\sigma_4(\text{\AA})$	65.8(8)			68.1(2)			65.2(9)		
$ Z_4 (e)$	8.4(6)			12.6(3)			12.3(1)		
$C_4(e^2)$	36			36			36		

Legenda: Resultados consideram potenciais atrativos devido à regulação de carga e potencial de tipo Yukawa.

Pelos valores calculados na concentração de 25 mg/ml para o pH 6.8 é possível perceber um aumento no potencial de contato comparando-o com o modelo $C=0$. Apesar disso, ele está dentro do erro estatístico do ajuste. Além do mais, embora haja um aumento no valor da capacitância, o potencial de contato teve apenas uma pequena alteração (quando alterado) comparado ao valor de referência ($C=0$). O único dado que teve uma mudança mais significativo foi o de concentração de 5 mg/ml em pH 8.6. Mas como explicado anteriormente, neste caso a proteína está diluída neste sistema, o fator de estrutura é igual à 1 e portanto não sendo necessário utilizá-lo. Outro resultado a se notar é o fato de uma solução com $C = 16 e^2$ para $J = 1.0(4) k_B T$ ajustar a curva de 50 mg/ml para pH=8.6 apesar do ajuste para $J = 0 k_B T$ resultar em uma capacitância de $C = 0.01(1) e^2$. Este resultado proporciona a ideia de haver uma inconsistência nos dados encontrados para esse pH.

Deve ser notado que quanto maior a concentração proteica, menor o efeito do potencial de contato e, portanto, maior será a influência do potencial repulsivo na $S(q)$. Além disso, para baixa concentração (5 mg/ml) não é necessário um ajuste para os potenciais atrativos,

pois neste caso $S(q) \sim 1$.

5.3 Citocromo C de coração equino

Para o citocromo C foram realizados experimentos de SAXS em diferentes concentrações da proteína (5, 10, 25, 50 e 100 mg/ml) e valores de pH (7.1, 8.8, 10.5 e 11.1). As curvas para esses sistemas estão retratadas nas figuras (5.10 e 5.11). Observando os gráficos nota-se que assim como para a BSA, as intensidades medidas dependem da concentração mas também de outros fatores tal como o fator de interferência, $S(q)$. Observando mais atentamente, há um aumento no início da curva de espalhamento e este poderia ser explicado devido a alta interação atrativa entre as proteínas ou agregação. Para mostrar o efeito desse aumento abrupto, foram calculados os raios de giro em dois segmentos das curvas de 5 mg/ml, figura (5.12). Só foram calculadas para essas curvas pois a aproximação de Guinier só é válida para sistemas diluídos.

Os valores encontrados de raio de giro para as duas regiões são diferentes entre si, como por exemplo, para pH 11.1, onde o $R_g = 2.89(12)$ nm para q^2 no intervalo entre 0.02nm^{-2} e 0.05nm^{-2} , enquanto que para a curva, para q^2 no intervalo entre 0.05nm^{-2} e 1.0nm^{-2} , o raio de giro é dado por $R_g = 1.02(1)$ nm. É importante salientar que em ambos intervalos a condição de Guinier é válida. Além disso, como mostrado nas figuras (5.10 e 5.11) esse início abrupto se torna mais evidente para altas concentrações. Comparando os valores encontrados de R_g da curva, todos estão na região entre 0.93 e 1.02 nm, indicando uma pequena diferença entre os valores para diferentes pH. Devido a pequena diferença, pode-se inferir que a proteína não muda seu formato entre esses valores de pH. Kataoka et al. [1993] e Gast et al. [1994] encontraram para o citocromo C, a pH 7, $R_g = 1.35(1)$ nm, valor diferente do encontrado neste trabalho. As diferenças entre os raios de giro encontradas nas duas regiões indicam agregação de proteínas nos sistemas estudados.

Assim como aplicado para a BSA, a interação de citocromo C será estudada em 3 partes para os sistemas de 5, 10 e 25 mg/ml:

Figura 5.10: Curvas experimentais de citocromo C para pH=7.1 e 8.8 para intensidade x vetor de espalhamento (lado esquerdo) e intensidade normalizado (pela concentração) x vetor de espalhamento (lado direito).

- (i) Análise sem capacitância, $C = 0$;
- (ii) Análise sem potencial de contato, $J = 0$;
- (iii) Análise com capacitância e potencial de contato, $C \neq 0$ e $J \neq 0$.

Deve ser ressaltado que devido ao indício de agregação, não foi possível fazer o cálculo da $p(r)$ com o programa *Primus*. Apesar disso, assim como foi feito para a BSA, foi utilizada a estrutura cristalográfica do monômero do citocromo C para calcular seu fator de forma.

Figura 5.11: Curvas experimentais de citocromo C para pH=10.5 e 11.1 para intensidade x vetor de espalhamento (lado esquerdo) e intensidade normalizado (pela concentração) x vetor de espalhamento (lado direito).

5.3.1 Ajuste das curvas com $C=0$

Assim como para BSA, os cálculos foram realizados utilizando o mesmo modelo do GENFIT, sendo que para o citocromo C foram executados ajustes globais para as curvas de 5, 10 e 25 mg/ml. Devido à globularidade da proteína, foi utilizado o volume específico (ν_{esp}) entre 0.67 a 0.76 cm³/g, referentes ao diâmetro entre 37.3 e 38.9 Å. Os gráficos estão mostrados na figura (5.13) com os ajustes na tabela (5.5).

Devido à agregação mostrada no início da curva, foi adicionado o termo com lei de Porod ao modelo do *PDB* de modo a ajustar os dados experimentais. Os resultados utilizando a soma dos modelos, apresentados na figura (5.13), mostram que o termo referente à lei de Porod ajusta de maneira correta parte da curva, enquanto que o modelo do *PDB* e RPA conseguem ajustar para valores maiores do vetor de espalhamento.

Figura 5.12: Curvas experimentais de citocromo C para pH=10.5 e 11.1 para intensidade x vetor de espalhamento (lado esquerdo) e intensidade normalizado (pela concentração) x vetor de espalhamento (lado direito).

Diferentemente dos gráficos da BSA, os gráficos de citocromo C foram colocados a um vetor de espalhamento de até $0.35 \text{ \AA}^{-1}$, visto que até esse valor já é possível caracterizar a curva. Além disso, é possível perceber que para regiões maiores que $0.3 \text{ \AA}^{-1}$ há um aumento no valor da intensidade que, assim como mostrado para o caso da BSA, pode ser devido a subtração do tampão.

Comparando os valores de carga obtidos pelos ajustes das curvas com os valores simulados, figura (5.1c), é possível perceber que para pH distante do pI, os valores de cargas ajustados pelo *GENFIT* estão subestimado aos valores simulados pelo *Propka*. Horn e Heuck [1983] calcularam a carga do citocromo C para pH 7 e 9, encontrando os valores de 8.9 e 7.9 respectivamente. Esses dados se mostram próximos aos calculados teoricamente mas diferentes comparados com os encontrados pela simulação deste trabalho.

(a) Gráfico de citocromo C para pH=7.1

(b) Gráfico de citocromo C para pH=8.8

(a) Gráfico de citocromo C para pH=10.5

(b) Gráfico de citocromo C para pH=11.1

Figura 5.13: Curvas experimentais para citocromo C em (a) pH=7.1; (b) pH=8.8; (c) pH=10.5; (d) pH=11.1, com os respectivos ajustes.

5.3.2 Ajuste das curvas para $C \neq 0$ com $J=0$ e para $C \neq 0$ com $J \neq 0$.

Foram feitas simulações, supondo que todo potencial atrativo do sistema era devido a capacitância, ou seja, $J=0$. Em seguida, foram simulado curvas tanto para capacitância quanto para o potencial de contato diferentes de zero $C, J \neq 0$. Os ajustes são encontrados nas figuras (5.14 e 5.15), enquanto os parâmetros são dados na tabela (5.6)

Para as curvas onde tanto $J = 0$ quanto onde $J \neq 0$, podem-se perceber que os valores encontrados de carga são bem parecidos com os da tabela (5.5). Enquanto isso, os valores calculados de capacitância, para $J = 0$ são muitas vezes maiores do que os simulados na figura (5.1d), sendo que para pH 8.8 era de se esperar $C_{teorico} = 0.4e^2$, enquanto que os

Tabela 5.5 - Parâmetros encontrados a partir do ajuste dos dados do citocromo C para concentrações 5,10 e 25 mg/ml para pH entre 7.1 e 11.1

pH	11.1			10.5			8.8			7.1		
[Citocromo C](mg/ml)	5	10	25	5	10	25	5	10	25	5	10	25
$J_0(k_B T)$	13.0(1.8)	8(1)	5.0(6)	8.0(1.7)	6.0(8)	5(1)	11.0(1.5)	7(1)	6.0(1.2)	6.9(2)	6.0(2)	4.6(2)
$d_0(\text{\AA})$	3.5(1)			3.5(1)			3.5(1)			3.9(2)		
$\sigma_0(\text{\AA})$	38.5(6)			37.9(4)			37.1(3)			37.3(7)		
$ Z_0 (e)$	1.8(9)			1.8(1.1)			3.5(7)			4.0(1)		
$C_0(e^2)$	0			0			0			0		
$ Z_{teorico} (e)$	6.8			1.0			8.5			9.6		
$C_{teorico}(e^2)$	3.9			4.6			0.4			0.4		

Legenda: Valores calculados considerando somente o potencial atrativo tipo Yukawa.

valores encontrados variam de 30 a 44 e^2 . Por isso os ajustes onde $J \neq 0$ foram escolhidos valores fixos de capacitância, de modo que este não ultrapasse o maior valor de capacitância calculado pelas curvas onde $J = 0$.

Pelas figuras (5.14 e 5.15) é possível perceber que todas as curvas, independente de pH ou concentração de citocromo C, foram bem ajustadas com os parâmetros descritos na tabela (5.6). Contudo, apesar do aumento da capacitância, indo de 4 e^2 para 16 e^2 , os valores de J mostraram uma pequena alteração, sendo que para o resultado total, os valores encontrados são iguais dentro da incerteza estatística.

Tabela 5.6 - Parâmetros encontrados a partir do ajuste dos dados do citocromo C para concentrações 5,10 e 25 mg/ml para pH entre 7.1 e 11.1

pH	11.1			10.5			8.8			7.1		
[Citocromo C](mg/ml)	5	10	25	5	10	25	5	10	25	5	10	25
$J_1(k_B T)$	0			0			0			0		
$d_1(\text{\AA})$	3.5(2)			4.0(2)			3.9(1)			3.9(1)		
$\sigma_1(\text{\AA})$	37.2(1)			37.2(1)			37.2(1)			37.2(3)		
$ Z_1 (e)$	1.7(3)			1.9(2)			4.0(2)			4.9(1)		
$C_1(e^2)$	58.5(1.3)	42.9(1.4)	34.8(1.3)	60(3)	42.3(1.6)	33(1)	44(1)	31(1)	30(1)	39.3(9)	36.3(8)	34(1)
$J_2(k_B T)$	11(2)	7(1)	5.0(1.3)	10.0(1)	7(1)	5.0(6)	10(2)	7(1)	5(1)	7.6(1)	7.0(1)	5.1(1)
$d_2(\text{\AA})$	3.8(1)			3.5(1)			3.6(1)			3.5(1)		
$\sigma_2(\text{\AA})$	37.2(4)			37.4(1)			37.3(5)			37.6(2)		
$ Z_2 (e)$	1.1(1.1)			1.8(7)			3.1(1)			4.1(1)		
$C_2(e^2)$	4			4			4			4		
$J_3(k_B T)$	13(1)	6.0(5)	4.0(5)	9.0(5)	6.0(1)	5.0(5)	10(1)	7.0(5)	5.0(5)	7.3(3)	6.7(1)	4.7(1)
$d_3(\text{\AA})$	3.7(1)			3.5(1)			3.5(2)			3.5(1)		
$\sigma_3(\text{\AA})$	37.8(6)			38.0(3)			37.3(3)			37.2(1)		
$ Z_3 (e)$	1.2(9)			2.2(5)			3.5(3)			4.2(1)		
$C_3(e^2)$	9			9			9			9		
$J_4(k_B T)$	11(1)	5.0(1)	4.0(6)	8(1)	5.0(6)	3.0(6)	8.0(8)	5.0(1)	4.0(5)	6.0(3)	5.3(3)	3.8(1)
$d_4(\text{\AA})$	3.5(2)			4.1(1)			3.5(3)			3.5(1)		
$\sigma_4(\text{\AA})$	37.2(1)			37.2(3)			37.5(1)			37.2(1)		
$ Z_4 (e)$	1.3(7)			1.5(8)			3.1(1)			4.2(1)		
$C_4(e^2)$	16			16			16			16		

Legenda: Resultados consideram potenciais atrativos devido à regulação de carga e potencial de tipo Yukawa.

Figura 5.14: Curvas experimentais do citocromo C com ajustes para pH=7.1 e 8.8 nas concentrações (a) e (b) 5 mg/ml; (c) e (d) 10 mg/ml ; (e) e (f) 25 mg/ml respectivamente.

Figura 5.15: Curvas experimentais do citocromo C com ajustes para pH=10.5 e 11.1 nas concentrações (a) e (b) 5 mg/ml; (c) e (d) 10 mg/ml ; (e) e (f) 25 mg/ml respectivamente.

5.4 Comparação dos resultados entre BSA e citocromo C

Os resultados obtidos para os dados de BSA e citocromo C mostraram-se adequados aos dados experimentais obtidos. Com eles é possível perceber algumas semelhanças entre os parâmetros. O potencial de contato (J), por exemplo, está relacionado com a concentração, de modo que para ambas proteínas estudados, conforme a concentração foi aumentada, houve uma diminuição valor de J . Esse é um fato interessante de mostrar pois, esse resultado indica que conforme há um aumento na quantidade de proteínas em um sistema, estas estarão interagindo menos entre si. Isso contrapõe o fato de que as partículas (proteínas) estarão mais perto umas das outras e portanto, tanto o potencial de Lennard-Jones quanto o potencial de depleção deveriam ser maiores (gerando mais atração no sistema).

Além disso, alguns resultados mostraram que apesar de correspondentes bons ajustes de curvas nos dados experimentais, os valores encontrados são bem diferentes dos valores simulados de cargas. Para a BSA, por exemplo, o valor encontrado de carga efetiva a pH 8.6 foi encontrada como sendo $|Z| \sim 12e$, enquanto que o valor simulado deveria ser de 22.8. Já para o citocromo C, a carga para pH 7.1 foi encontrada como $|Z| \sim 4e$, enquanto o valor esperado era 9.6.

A inclusão do termo de capacitância não gerou alterações significativas nos valores de cargas obtidos nas simulações, enquanto que para J houve uma pequena diminuição em seu valor, de forma esta foi compensada pela própria incerteza do valor. Uma maneira de observar esse fato é através dos gráficos de potenciais gerados tanto pelos dados de BSA quanto para os dados de citocromo C.

A figura (5.16) mostra simulações para BSA a partir dos potenciais atrativos gerados. As curvas das figuras (5.16a, c, e) foram simuladas utilizando a tabela (5.2), ou seja, dos dados sem capacitância para as respectivas concentrações, enquanto que as curvas da figura (5.16b, d, f) foram simuladas segundo dos valores de capacitância utilizados na tabela (5.4). A partir dos gráficos é possível perceber que $U_Y \geq -4k_B T$ para concentrações começando em 25 mg/ml. Já os valores teóricos calculados levam a um potencial devido a regulação

de cargas de $U_{CR} \geq -0.25 k_B T$, enquanto que para $C = 16e^2$ haverá um potencial atrativo de até $U_{CR} \sim -0.5 k_B T$. Assim, o termo de regulação de carga é muito baixo comparado com o termo atrativo tipo Yukawa.

Os potenciais atrativos gerados pelos dados do citocromo C estão na figura (5.17), onde as figuras (5.17a, c, e) representam o potencial atrativo do tipo Yukawa para as respectivas concentrações, enquanto as figuras (5.17b, d, f) representam o potencial atrativo devido a regulação de cargas. Para o valor de pH 10.5, onde as cargas calculadas foi parecida com a carga obtida por simulação, pode-se perceber que até mesmo para os dados teóricos, o valor de potencial atrativo não será menor do que $\sim -0.5 k_B T$, resultando em fator de $\sim 10\%$ do potencial atrativo tipo Yukawa.

Figura 5.16: Simulações dos potenciais atrativos para BSA utilizando os dados da tabela 5.2 para os gráficos (a), (c) e (e), e dados da tabela 5.4 para os gráficos (b), (d) e (f) com os respectivos valores de pH.

Figura 5.17: Simulações dos potenciais atrativos para citocromo C utilizando os dados da tabela 5.5 para os gráficos (a), (c), (e) e (g), e dados da tabela 5.6 para os gráficos (b), (d), (f) e (h) com os respectivos valores de pH.

Conclusões

Com esse trabalho, o potencial atrativo, gerado a partir da teoria de regulação de cargas, foi implementado nos potenciais de interação proteína-proteína de modo que pudesse ser calculado através das curvas de SAXS. Os resultados mostram que considerando ou não este termo, é possível caracterizar sistemas concentrados, mesmo que o desenvolvimento da teoria seja para sistemas diluídos. A inclusão do termo de regulação de cargas nas análises das curvas mostrou que o potencial devido a esse fator, por si só, não consegue descrever todo o potencial atrativo do sistema, pois, assim como apresentado neste trabalho, os valores deste fator seriam muito altos comparados aos calculados da curva de carga superficial por pH. Esse fato indica que tanto para BSA quanto para o citocromo C, também deveriam ser considerados termos de ordens mais altos que r^{-2} , incluindo assim interação íon-dipolo e dipolo-dipolo, já que o momento de dipolo delas é alto (comparado com outras proteínas). Além disso, quando o termo de regulação de carga foi simulado concomitantemente com o potencial atrativo tipo Yukawa, seu aumento não resultou em alterações significativas do valor do potencial de contato, J . Assim sendo, o efeito provocado pelo potencial de regulação de cargas é pequeno quando comparado ao potencial atrativo total e, por isso, apesar de observado, não é expressivo para soluções de BSA e citocromo C.

Deve-se levar em conta também, que todo o processo descrito foi utilizado valores de carga gerada pelo programa Propka e, por isso, não foram considerados como as condições do sistema afetam esses valores. Para se aprofundar melhor nessas situações, na continuidade deste trabalho, serão realizados simulações de Monte Carlo, em que diferentes

concentrações de sais e de proteínas. Além disso, foram utilizadas soluções com uma mesma proteína e, portanto, também serão testado resultados para diferentes proteínas em solução.

A metodologia utilizada pode ser empregada em outros sistemas, como por exemplo, sistemas micelares. Comparada a sistemas proteicos, as micelas apresentam uma distribuição de carga homogênea em sua superfície, enquanto que para proteínas essa homogeneidade é apenas uma aproximação.

Acreditamos que este trabalho fornece mais informações teóricas acerca da regulação de cargas tal como os valores do termo de regulação de carga em função do pH. Além disso, deve-se fazer novos testes sobre este termo e também sobre carga total, em função das condições do sistema, para poder concluir com maior exatidão a influência dos seus valores nos potenciais de interação do sistema.

Referências Bibliográficas

- Andersen H. C., Chandler D., Optimized Cluster Expansions for Classical Fluids. I. General Theory and Variational Formulation of the Mean Spherical Model and Hard Sphere Percus-Yevick Equations, *The Journal of Chemical Physics*, 1972, vol. 57, p. 1918
- Avni Y., Andelman D., Podgornik R., Charge Regulation with Fixed and Mobile Charges, *Current Opinion in Electrochemistry*, 2018
- Barbosa L. R., Ortore M. G., Spinozzi F., Mariani P., Bernstorff S., Itri R., The importance of protein-protein interactions on the pH-induced conformational changes of bovine serum albumin: a small-angle X-ray scattering study, *Biophysical journal*, 2010, vol. 98, p. 147
- Barbosa L. R. S., Spinozzi F., Mariani P., Itri R., Small-angle X-ray scattering applied to proteins in solution, *Proteins in Solution and at Interfaces*. John Wiley & Sons, Inc, 2013, pp 49–72
- Barroso da Silva F. L., Bostrom M., Persson C., Effect of charge regulation and ion–dipole interactions on the selectivity of protein–nanoparticle binding, *Langmuir*, 2014, vol. 30, p. 4078
- Barroso da Silva F. L., Lund M., Jönsson B., Åkesson T., On the complexation of proteins and polyelectrolytes, *The Journal of Physical Chemistry B*, 2006, vol. 110, p. 4459
- Barroso da Silva F. L., MacKernan D., Benchmarking a fast proton titration scheme in im-

- PLICIT solvent for biomolecular simulations, *Journal of chemical theory and computation*, 2017, vol. 13, p. 2915
- Bendedouch D., Chen S. H., Structure and interparticle interactions of bovine serum albumin in solution studied by small-angle neutron scattering, *The Journal of Physical Chemistry*, 1983, vol. 87, p. 1473
- Berg J. M., Tymoczko J. L., Stryer L., et al., 2012 *Biochemistry*/Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; with Gregory J. Gatto, Jr.
- Biochemistry, 2019 Peptide Bond Formation <https://www.drawittoknowit.com/course/biochemistry/glossary/biochemical-pathway/peptide-bond-formation>. Acesso em: 01 Julho 2019
- Bostrom M., Tavares F. W., Bratko D., Ninham B., Specific ion effects in solutions of globular proteins: comparison between analytical models and simulation, *The Journal of Physical Chemistry B*, 2005, vol. 109, p. 24489
- Bradbury J., Chaperones: keeping a close eye on protein folding, *The Lancet*, 2003, vol. 361, p. 1194
- Brumberger H., *Modern aspects of small-angle scattering*. vol. 451, Springer Science & Business Media, 2013
- Carter D. C., Chang B., Ho J. X., Keeling K., Krishnasami Z., Preliminary crystallographic studies of four crystal forms of serum albumin, *European Journal of Biochemistry*, 1994, vol. 226, p. 1049
- Chaudhuri T. K., Paul S., Protein-misfolding diseases and chaperone-based therapeutic approaches, *The FEBS journal*, 2006, vol. 273, p. 1331
- Croxton C. A., *Introduction to liquid state physics*. Wiley, 1975
- Franco L. F. M., de Alcântara Pessoa Filho P., On the solubility of proteins as a function of pH: Mathematical development and application, *Fluid Phase Equilibria*, 2011, vol. 306, p. 242

- Franco L. F. M., de Oliveira C. L. P., de Alcântara Pessoa Filho P., Thermodynamics of protein aqueous solutions: From the structure factor to the osmotic pressure, *AIChE Journal*, 2015, vol. 61, p. 2871
- Gast K., Damaschun H., Misselwitz R., Müller-Frohne M., Zirwer D., Damaschun G., Compactness of protein molten globules: temperature-induced structural changes of the apomyoglobin folding intermediate, *European biophysics journal*, 1994, vol. 23, p. 297
- Geiler K., , 2010 Protein folding: the good, the bad, and the ugly <http://sitn.hms.harvard.edu/flash/2010/issue65/>. Acesso em: 01 Julho 2019
- GpNotebook, 2016 hemoglobin (reference range) <http://www.gpnotebook.co.uk/simplepage.cfm?ID=1026883654>. Acesso em: 27 Junho 2019
- Guinier A., Fournet G., Small angle X-ray scattering, New York, 1955
- Hamaker H. C., The London—van der Waals attraction between spherical particles, *physica*, 1937, vol. 4, p. 1058
- Harding S. E., Rowe A. J., Modelling biological macromolecules in solution: 1. The ellipsoid of revolution, *International Journal of Biological Macromolecules*, 1982, vol. 4, p. 160
- Hill T. L., An introduction to statistical thermodynamics. Dover Publications, 1986
- Horn D., Heuck C.-C., Charge determination of proteins with polyelectrolyte titration., *Journal of Biological Chemistry*, 1983, vol. 258, p. 1665
- Israelachvili J. N., Intermolecular and surface forces. Academic press, 2011
- Jonsson B., Lund M., da Silva F. L. B., Electrostatics in macromolecular solutions. Royal Society of Chemistry: Cambridge, UK, 2007, 129
- Kataoka M., Hagihara Y., Mihara K., Goto Y., , 1993 Molten globule of cytochrome c studied by small angle X-ray scattering

- Kirkwood J. G., Shumaker J. B., Forces between protein molecules in solution arising from fluctuations in proton charge and configuration, *Proceedings of the National Academy of Sciences of the United States of America*, 1952, vol. 38, p. 863
- Konarev P. V., Volkov V. V., Sokolova A. V., Koch M. H., Svergun D. I., PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *Journal of applied crystallography*, 2003, vol. 36, p. 1277
- Kurut A., Lund M., Solution electrostatics beyond pH: a coarse grained approach to ion specific interactions between macromolecules, *Faraday discussions*, 2013, vol. 160, p. 271
- Lehninger A. L., Nelson D. L., Cox M. M., Cox M. M., et al., *Lehninger principles of biochemistry*. Macmillan, 2005
- Li H., Robertson A. D., Jensen J. H., Very fast empirical prediction and rationalization of protein pKa values, *Proteins: Structure, Function, and Bioinformatics*, 2005, vol. 61, p. 704
- Li W., Persson B. A., Morin M., Behrens M. A., Lund M., Oskolkova M. Z., Charge-induced patchy attractions between proteins, *The Journal of Physical Chemistry B*, 2015, vol. 119, p. 503
- LNLS, 2019 SAXS Overview <https://www.lnls.cnpem.br/linhas-de-luz/saxs1-en/>. Acesso em: 27 Junho 2019
- Lund M., Jönsson B., On the charge regulation of proteins, *Biochemistry*, 2005, vol. 44, p. 5722
- Lund M., Jönsson B., Charge regulation in biomolecular solution, *Quarterly reviews of biophysics*, 2013, vol. 46, p. 265
- McDonald I. R., Hansen J., *Theory of simple liquids*. Academic, 1976
- Mairbäurl H., Weber R. E., Oxygen transport by hemoglobin, *Comprehensive physiology*, 2011, vol. 2, p. 1463

- Malmgren L., Olsson Y., Olsson T., Kristensson K., Uptake and retrograde axonal transport of various exogenous macromolecules in normal and crushed hypoglossal nerves, *Brain research*, 1978, vol. 153, p. 477
- Mayer J. E., Montroll E., Molecular distribution, *The Journal of Chemical Physics*, 1941, vol. 9, p. 2
- Medda L., Barse B., Cugia F., Bostrom M., Parsons D. F., Ninham B. W., Monduzzi M., Salis A., Hofmeister challenges: ion binding and charge of the BSA protein as explicit examples, *Langmuir*, 2012, vol. 28, p. 16355
- Melnichenko Y. B., *Small-Angle Scattering from Confined and Interfacial Fluids*. Springer, 2016
- Miyashita Y., Wazawa T., Mogami G., Takahashi S., Sambongi Y., Suzuki M., Hydration-state change of horse heart cytochrome c corresponding to trifluoroacetic-acid-induced unfolding, *Biophysical journal*, 2013, vol. 104, p. 163
- Nagy B., Tóth A., Savina I., Mikhalovsky S., Mikhalovska L., Grillo I., Geissler E., László K., Small angle neutron scattering study of globular proteins confined in porous carbons, *Carbon*, 2016, vol. 106, p. 142
- Narayanan J., Liu X., Protein interactions in Undersaturated and supersaturated solutions: A study using light and X-Ray scattering., *Biophys. J.*, 2003, vol. 84, p. 523
- Olsson M. H., Søndergaard C. R., Rostkowski M., Jensen J. H., PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions, *Journal of chemical theory and computation*, 2011, vol. 7, p. 525
- Ortore M. G., Spinozzi F., Mariani P., Paciaroni A., Barbosa L. R., Amenitsch H., Steinhart M., Ollivier J., Russo D., Combining structure and dynamics: non-denaturing high-pressure effect on lysozyme in solution, *Journal of The Royal Society Interface*, 2009, vol. 6, p. S619

- Pedersen J. S., Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting, *Advances in colloid and interface science*, 1997, vol. 70, p. 171
- Peters Jr T., All about albumin: biochemistry, genetics, and medical applications. Academic press, 1995
- RCSB, 2000 Protein Data Bank <https://www.rcsb.org/>. Acesso em: 27 Junho 2019
- Reachingfordreams, 2019 A4 structures of proteins and their building blocks – 20 amino acids <http://biology.reachingfordreams.com/biology/biomolecules/27-20-amino-acids-and-protein-structures>. Acesso em: 27 Junho 2019
- Reed C. E., Reed W. F., Monte Carlo study of titration of linear polyelectrolytes, *The Journal of chemical physics*, 1992, vol. 96, p. 1609
- Rezwan K., Meier L. P., Rezwan M., Vörös J., Textor M., Gauckler L. J., Bovine serum albumin adsorption onto colloidal Al₂O₃ particles: a new model based on zeta potential and UV- Vis measurements, *Langmuir*, 2004, vol. 20, p. 10055
- Salinas S. R., Introdução a física estatística. Edusp, 2013
- Sassi A. P., Beltrán S., Hooper H. H., Blanch H. W., Prausnitz J., Siegel R. A., Monte Carlo simulations of hydrophobic weak polyelectrolytes: Titration properties and pH-induced structural transitions for polymers containing weak electrolytes, *The Journal of chemical physics*, 1992, vol. 97, p. 8767
- Scott T. A., Mercer E. I., Concise Encyclopedia Biochemistry and Molecular Biology. Walter de Gruyter Berlin, 1996
- Spinozzi F., Carsughi F., Mariani P., Particle shape reconstruction by small-angle scattering: Integration of group theory and maximum entropy to multipole expansion method, *The Journal of chemical physics*, 1998, vol. 109, p. 10148

- Spinozzi F., Ferrero C., Ortore M. G., Antolinos A. d. M., Mariani P., GENFIT: software for the analysis of small angle X-ray and neutron scattering data of macromolecules in solution, *Jornal of applied crystallography*, 2014, vol. 47, p. 1132
- Spinozzi F., Gazzillo D., Giacometti A., Mariani P., Carsughi F., Interaction of proteins in solution from small-angle scattering: a perturbative approach, *Biophysical journal*, 2002, vol. 82, p. 2165
- Svergun D. I., Koch M. H., Advances in structure analysis using small-angle scattering in solution, *Current opinion in structural biology*, 2002, vol. 12, p. 654
- Svergun D. I., Koch M. H., Timmins P. A., May R. P., Small angle X-ray and neutron scattering from solutions of biological macromolecules. vol. 19, Oxford University Press, 2013
- Tanford C., Buzzell J. G., The viscosity of aqueous solutions of bovine serum albumin between pH 4.3 and 10.5., *The Journal of Physical Chemistry*, 1956, vol. 60, p. 225
- Tardieu A., Le Verge A., Malfois M., Bonneté F., Finet S., Ries-Kautt M., Belloni L., Proteins in solution: from x-ray scattering intensities to interaction potentials, *Journal of Crystal Growth*, 1999, vol. 196, p. 193
- Teixeira A. A. R., Lund M., Barroso da Silva F. L., Fast proton titration scheme for multiscale modeling of protein solutions, *Journal of chemical theory and computation*, 2010, vol. 6, p. 3259
- Thiele E., Equation of state for hard spheres, *The Journal of Chemical Physics*, 1963, vol. 39, p. 474
- Torres B. B., Marzzoco A., *Bioquímica básica*, Rio de Janeiro, RJ: Guanabara Koohan, 2007
- Tresca A. J., , 2011 Agreed Adult Clinical Biochemistry Reference Intervals <http://www.acb.org.uk/docs/default-source/committees/scientific/>

guidelines/acb/pathology-harmony-ii-biochemistry.pdf?sfvrsn=2. Acesso em: 27 Junho 2019

Tresca A. J., , 2019 What Is a Hemoglobin Test? <https://www.verywellhealth.com/hemoglobin-level-test-1942658>. Acesso em: 27 Junho 2019

Valcárcel C. A., Santos F. P., Estrategias y Avances en el Estudio de Toxinas de Interés para la Biomedicina. Benemérita Universidad Autonoma de Puebla, 2015

Vervey E. J. W., Overbeek J. G., , 1948 Theory of stability of Lyophobic Colloids,

Vuolo J., Fundamentos da teoria de erros. 2a Edição, Ed, Edgard Blücher Ltda., São Paulo, 2002

Wertheim M., Exact solution of the Percus-Yevick integral equation for hard spheres, Physical Review Letters, 1963, vol. 10, p. 321

Zhang F., Skoda M. W., Jacobs R. M., Martin R. A., Martin C. M., Schreiber F., Protein interactions studied by SAXS: effect of ionic strength and protein concentration for BSA in aqueous solutions, The Journal of Physical Chemistry B, 2007, vol. 111, p. 251

Apêndices

Apêndice A

Tabela de funções e integrais

- Função de partição canônica:

$$Z = \sum_i e^{-\beta E_i} \quad (\text{A.1})$$

- Integral de configuração:

$$Z_n = \int \cdots \int e^{-\beta U(\mathbf{r}^n)} d\mathbf{r}_1 \cdots d\mathbf{r}_n \quad (\text{A.2})$$

- Energia livre de Boltzmann:

$$F = -k_B T \ln Z \quad (\text{A.3})$$

- Expansão exponencial

$$e^{-x} = 1 - x + \frac{x^2}{2!} + O(x^3) \quad (\text{A.4})$$

- Expansão logarítmica natural:

$$\ln(1-x) = -x - \frac{x^2}{2} + O(x^3) \quad (\text{A.5})$$

- Expansão inversa da Raiz Quadrada

$$(1+x)^{-1/2} = 1 - \frac{x}{2} + O(x^2) \quad (\text{A.6})$$

- Série geométrica

$$\sum_{n=1}^{\infty} a_n = \frac{a_1}{1-q}, \quad \text{se } |q| < 1 \quad (\text{A.7})$$

- Expansão $\text{sen}(x)/x$

$$\frac{\text{sen}(x)}{x} \approx 1 - \frac{x^2}{3!} + O(x^4) \quad (\text{A.8})$$

- Função de Bessel de primeira espécie e ordem m , $J_m(x)$:

$$J_m(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n! (n+m)!} \left(\frac{x}{2}\right)^{2n+m} \quad (\text{A.9})$$

- Função de Bessel de primeira ordem, $J_1(x)$:

$$J_1(x) = \frac{\text{sen}(x) - x \cos(x)}{x^2} \quad (\text{A.10})$$

- χ^2 reduzido:

$$\chi^2 = \frac{1}{N_c} \sum_{c=1}^{N_c} \frac{1}{N_{q,c}} \sum_{i=1}^{N_{q,c}} \left\{ \frac{I_{exp,c}(q_i) - I_c(q_i)}{\sigma_c(q_i)} \right\}^2 \quad (\text{A.11})$$

- Função Seno integral, $S_i(x)$:

$$S_i(x) = \int_0^x \frac{\text{sen}(t)}{t} dt \quad (\text{A.12})$$

- Integral $x \text{sen}(ax)$:

$$\int_0^b x \text{sen}(ax) dx = \frac{\text{sen}(ab) - ab \cos(ab)}{a^2} = J_1(ab)b^2 \quad (\text{A.13})$$

- Função exponencial integral:

$$E_1(z) = \int_z^{\infty} \frac{e^{-t}}{t} dt, \quad |Arg(z)| < \pi \quad (\text{A.14})$$

- Integral seno exponencial :

$$\int_{\sigma}^{\infty} e^{-ar+b} \text{sen}(qr) dr = \frac{a}{a^2 + b^2} \left(\text{sen}(q\sigma) + \frac{q}{a} \cos(q\sigma) \right) e^{-a\sigma+b} \quad (\text{A.15})$$

- Integral $x \text{sen}(ax)/a$:

$$\int_0^{\sigma} \frac{x \text{sen}(qx)}{q} dx = J_1(q\sigma) \frac{\sigma^2}{q} \quad (\text{A.16})$$

- Integral $x^2 \text{sen}(ax)/a$:

$$\int_0^{\sigma} \frac{x^2 \text{sen}(qx)}{q} dx = \frac{(2 - \sigma^2 q^2) \cos(\sigma q) + 2\sigma q \text{sen}(\sigma q) + 2}{q^4} \quad (\text{A.17})$$

- Integral $x^4 \text{sen}(ax)/a$:

$$\int_0^{\sigma} \frac{x^4 \text{sen}(qx)}{q} dx = \frac{4\sigma q(\sigma^2 q^2 - 6) \text{sen}(\sigma q) - (\sigma^4 q^4 - 12\sigma^2 q^2 + 24) \cos(\sigma q) + 24}{q^6} \quad (\text{A.18})$$

Apêndice B

Transformada de Fourier de um Potencial do Tipo-Yukawa

O potencial de Perturbação foi definido na equação (4.41)

$$\omega(\mathbf{r}) = U_{SC}(r) + U_{CR}(r) + U_Y(r) \quad (\text{B.1})$$

onde cada uma dos potenciais foram definidos nas equações (4.13) e (4.108), enquanto a componente referente a regulação de carga é descrito como o segundo termo da equação (4.39) dividido por beta. Para a transformada de Fourier então:

$$\phi(q) = \int \beta\omega(r) \frac{\text{sen}(qr)}{qr} r^2 d\mathbf{r} = 4\pi \int_{\sigma}^{\infty} \beta\omega(r) \frac{\text{sen}(qr)}{qr} r^2 dr \quad (\text{B.2})$$

Substituindo os valores para $\beta\omega(r)$:

$$\phi(q) = \frac{4\pi}{q} \left[\int_{\sigma}^{\infty} A e^{-\kappa(r-\sigma)} \text{sen}(qr) dr - \int_{\sigma}^{\infty} B e^{-2\kappa(r-\sigma)} \frac{\text{sen}(qr)}{r} dr - \int_{\sigma}^{\infty} D e^{-\frac{(r-\sigma)}{d}} \text{sen}(qr) dr \right] \quad (\text{B.3})$$

onde $A = \frac{\langle Q \rangle^2 l_B}{(1+0.5\kappa\sigma)^2}$, $B = \frac{l_B^2 (C^2 + 2C\langle Q \rangle^2)}{2(1+0.5\kappa\sigma)^4}$ e $D = J\sigma$.

A primeira e última integral são bem parecidas, pois são do tipo:

$$\int_{\sigma}^{\infty} e^{-ar+b} \text{sen}(qr) dr \quad (\text{B.4})$$

onde para a primeira equação, $a = \kappa$, e $b = \kappa\sigma$. Já para a terceira equação, $a = 1/d$, e $b = \sigma/d$. Pela tabela de integrais no apêndice, equação (A.15), a solução da equação (B.3) para o primeiro e terceiro termos serão:

$$\begin{aligned}\beta U_{SC}(q) &= \frac{4\pi A}{q} \int_{\sigma}^{\infty} e^{-\kappa r + \kappa\sigma} \text{sen}(qr) dr = \frac{4\pi A}{q} \frac{\kappa}{\kappa^2 + q^2} \left(\text{sen}(q\sigma) + \frac{q}{\kappa} \cos(q\sigma) \right) \\ \beta U_{SC}(q) &= \frac{4\pi}{q} \frac{\langle Q \rangle^2 l_B}{(1 + 0.5\kappa\sigma)^2} \frac{\kappa}{\kappa^2 + q^2} \left(\text{sen}(q\sigma) + \frac{q}{\kappa} \cos(q\sigma) \right)\end{aligned}\quad (\text{B.5})$$

$$\begin{aligned}\beta U_Y(q) &= -\frac{4\pi D}{q} \int_{\sigma}^{\infty} e^{-\frac{r}{d} + \frac{\sigma}{d}} \text{sen}(qr) dr = -\frac{4\pi D}{q} \frac{d}{1 + (qd)^2} (\text{sen}(q\sigma) + qd \cos(q\sigma)) \\ \beta U_Y(q) &= -\frac{4\pi J\sigma}{q} \frac{d}{1 + (qd)^2} (\text{sen}(q\sigma) + qd \cos(q\sigma))\end{aligned}\quad (\text{B.6})$$

Já o cálculo do segundo termo, referente a regulação de carga, é diferente devido ao fator r no dividendo.

$$\beta U_{CR}(q) = -\frac{4\pi B}{q} \int_{\sigma}^{\infty} e^{-2\kappa(r-\sigma)} \frac{\text{sen}(qr)}{r} dr \quad (\text{B.7})$$

Escrevendo a função seno como exponenciais, $\text{sen}(x) = (e^{ix} - e^{-ix})/2i$:

$$\begin{aligned}\beta U_{CR}(q) &= -\frac{4\pi B}{q} e^{2\kappa\sigma} \int_{\sigma}^{\infty} \frac{e^{-2\kappa r}}{2ir} (e^{iqr} - e^{-iqr}) dr \\ &= -\frac{2i\pi B}{q} e^{2\kappa\sigma} \left[\int_{\sigma}^{\infty} \frac{e^{-(2\kappa+iq)r}}{r} dr - \int_{\sigma}^{\infty} \frac{e^{-(2\kappa-iq)r}}{r} dr \right]\end{aligned}\quad (\text{B.8})$$

Para a primeira integral será feito a mudança de variável $z = (2\kappa + iq)r$, já para a segunda integral $z = (2\kappa - iq)r$ resultando em:

$$\beta U_{CR}(q) = -\frac{2i\pi B}{q} e^{2\kappa\sigma} \left[\int_{(2\kappa+iq)\sigma}^{\infty} \frac{e^{-z}}{z} dz - \int_{(2\kappa-iq)\sigma}^{\infty} \frac{e^{-z}}{z} dz \right] \quad (\text{B.9})$$

Essas integrais são resolvidas através de da função exponencial integral definido na equação

(A.14). A parte de regulação de carga será dada por:

$$\beta U_{CR}(q) = -\frac{2i\pi B}{q} e^{2\kappa\sigma} [E_1((2\kappa + iq)\sigma) - E_1((2\kappa - iq)\sigma)] \quad (\text{B.10})$$

Escrevendo o valor entre colchetes como $E(q) = E_1((2\kappa + iq)\sigma) - E_1((2\kappa - iq)\sigma)$ e substituindo o valor de B :

$$\beta U_{CR}(q) = -\frac{iE(q)\pi}{q} \frac{l_B^2}{(1 + 0.5\kappa\sigma)^4} (C^2 + 2C\langle Q \rangle^2) e^{2\kappa\sigma} \quad (\text{B.11})$$