UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

JOSÉ FERNANDO RUGGIERO BACHEGA

**Estrutura cristalográfica da hemoglobina gigante de *Glossoscolex paulistus*, um
complexo de 3,6 mega Daltons**

São Carlos
2013

JOSÉ FERNANDO RUGGIERO BACHEGA

Estrutura cristalográfica da hemoglobina gigante de *Glossoscolex paulistus*, um complexo de 3,6 mega Daltons

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração :Física Aplicada
Opção: Física Biomolecular
Orientador: Richard Charles Garratt

Versão Corrigida
(Versão original disponível na Unidade que aloja o Programa)

São Carlos
2013

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Bachega, José Fernando Ruggiero
Estrutura cristalográfica da hemoglobina gigante
de *Glossoscolex paulistus*, um complexo de 3,6 mega
Daltons / José Fernando Ruggiero Bachega; orientador
Richard Charles Garratt - versão corrigida -- São
Carlos, 2013.
165 p.

Tese (Doutorado - Programa de Pós-Graduação em
Física Aplicada Biomolecular) -- Instituto de Física
de São Carlos, Universidade de São Paulo, 2013.

1. Hemoglobinas. 2. Eritrocrúorinas. 3.
Glossoscolex paulistus. 4. Cristalografia. I.
Garratt, Richard Charles, orient. II. Título.

Folha de aprovação

*Dedico este trabalho aos meus pais, Sônia Maria
Ruggiero Bachega e José Alberto Bachega, pelo
carinho e força.*

AGRADECIMENTOS

Ao Prof. Dr. Richard Charles Garratt, pela orientação, pelo vasto conhecimento compartilhado, pela amizade e pelo exemplo de ética e caráter.

Ao Prof. Dr. Marcel Tabak, pela coorientação e amizade, sem as quais esse trabalho jamais poderia ter sido realizado.

Ao prof. Dr. Eduardo Rebordero Horjales, pela amizade, trabalho em conjunto e conhecimento compartilhado.

Ao prof. Dr. Otavio Thiemann, pela ajuda e aconselhamento.

Ao meu amigo e colaborador Fernando Vasconcelos Maluf, cujas as habilidades e energias somadas a esse projeto resultaram em enormes avanços, dos quais muito me orgulho.

À Dra. Ana Cristina Puhl pelo companheirismo, paciência, carinho e amizade.

À Dra. Sandra Martha Gomes Dias e ao Dr. Marcelo Falsarella Carazzolle, que foram os grandes responsáveis pelo sequenciamento da HbGp.

Aos colaboradores, Dr. Allen Orville e Andi Babak do BNL.

À Dra. Adriana Paes Leme, Marcel Nakahira e Francieli Colussi, pela ajuda nos experimentos de espectrometria de massas.

À Dra. Inês Maria Basso Bernardi pelas análises de ICP

Ao pessoal da secretaria de pós-graduação, pela enorme ajuda e orientação, e às funcionárias da biblioteca, Cris e Neusa, pela grande ajuda na correção deste trabalho.

Aos técnicos e funcionários do CBME em especial a Dra. Susana Andréa e o Dr. Humberto Muniz, grande ajuda e amizade.

A todo o time de desenvolvedores do GTKDynamo, Dr. Martin J. Field, Dr. Troy Wymore, Lucas Assirati, Dr. Leonardo Ruggiero Bachega e Msc. Luis Fernando Saraiva.

Aos meus queridos amigos e companheiros de sala, Antônio Marinho, Lucas Bleicher, José Ribamar, Rinaldo Montalvão, Milton Sonoda, Alessandro Nascimento, Jademilson dos Santos, Leonardo França, Leonardo Ferreira, Evandro Araujo, Atílio Tomazini Jr. Michelle Fagundes, Nayara Cavalcante, César Camilo, Marco Kadowaki, Reginaldo Kuroshu, Eliane Bogu, Vanessa Arnoldi, Amanda, Laureana Stelmastchuk, João Jorge (o Joãozinho), Thiago Pereira, Renata Bueno, Ricardo Nascimento (Rick Gordão), Gustavo Lima (Udi), Matheus Postigo, Lívia Faim, Daiana Evelin, Lívia Manzini, Lis Miotto, Fernanda Costa, Viviane Serpa, Pablo Rosi, Mamé Dotta, Marcelo Liberato, Marco Túlio, Wanessa Altei, Ivani Pauli, Juliana Torini, Larissa Romanello, Maria Luiza Voltatódio, Renata Alves, Vitor Hugo Serrão,

Amanada Bernardes e João Renato.

A todos os meus amigos e companheiros de treino, em especial ao Marcos Rodrigues, o Marcão, que colaborou disponibilizando as amostras *E.andrei* para as etapas finais desse projeto.

A todo o pessoal da Minhobox, em especial ao Afrânio pela doação das mostras das espécies *E.andrei*, *P.escavatus* e *E.eugeni*.

“A melhor forma de ter uma boa ideia é ter muitas ideias.”

Linus Pauling

RESUMO

BACHEGA, J.F.R. **Estrutura cristalográfica da hemoglobina gigante de *Glossoscolex paulistus*, um complexo de 3.6 mega Daltons**. 2013. 165 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Eritrocruorinas são hemoglobinas gigantes compostas por uma bicamada hexagonal de massa molecular total entre 3,0 e 4,0 MDa. A sua estrutura é baseada em unidades básicas chamadas protômeros. Doze destes compõem a partícula inteira, seis em cada camada hexagonal, resultando numa estrutura contendo 180 subunidades. Nas eritrocruorinas do tipo I, o protômero é constituído por quatro tipos de cadeias de globina: *a*, *b*, *c*, e *d*, e por três tipos de cadeias de *linkers*, L1, L2 e L3. A compreensão atual do seu mecanismo da ação está atualmente prejudicada pela resolução limitada das estruturas cristalográficas disponíveis. Para abordar esta questão procuramos cristalizar uma série de eritrocruorinas de espécies diversas visando a determinação das suas estruturas a mais alta resolução. Na primeira parte deste trabalho a eritrocruorina de *Glossoscolex paulistus* (HbGp) teve suas subunidades sequenciadas e a estrutura da partícula inteira resolvida a uma resolução de 3,2Å, a mais alta até o momento. Na estrutura da HbGp as quatro cadeias de globina se associam na forma de um heterotetrâmero, que se repete três vezes formando uma estrutura dodecamérica denominada *cap*. A estrutura do *cap* associada a um heterotrîmero de *linkers* forma o então mencionado protômero. A estrutura completa permite uma descrição mais detalhada dos contatos entre subunidades que são essenciais para a manutenção da partícula como um todo. Além disto descrevemos sítios de ligação a metais (Zn^{2+} e Ca^{2+}) e sítios de glicosilação, alguns dos quais são inéditos. Em seguida, a subunidade *d* isolada da HbGp foi cristalizada e resolvida a 2,1Å. Uma análise dos contatos cristalinos demonstra um arranjo completamente diferente daquilo visto para a subunidade *d* no complexo inteira. Ao invés de associar-se na formar trîmeros, como acontece no complexo, a cadeia *d* isolada forma dímeros cristalográficos, com interface similar a aquela observada entre as cadeias *d* e *a*. Estas observações contribuíram para a compreensão de como são os possíveis mecanismos de associação das globinas para a montagem do *cap*. Finalmente, as eritrocruorinas de quatro outras espécies foram cristalizadas, que resultou na obtenção de uma estrutura cristalográfica preliminar para a eritrocruorina de *Eisenia andrei* a uma resolução de 4,7Å.

Palavras-chaves: Hemoglobinas. Eritrocruorinas. *Glossoscolex paulistus*. Cristalografia.

ABSTRACT

BACHEGA, J.F.R. **Crystallographic structure from the giant hemoglobin from *Glossoscolex paulistus*, a 3.6 mega dalton complex.** 2013. 165 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Erythrocruorins are giant haemoglobins in the form of a hexagonal bilayer of total molecular mass between 3 and 4 MDa. Their structures are based on a basic unit called the protomer. Twelve of these comprise the entire particle, six in each hexagonal layer, resulting in a structure containing 180 subunits. In the type I erythrocruorins the protomer is composed of four types of globin chains: *a*, *b*, *c* e *d*, and three types of linker; L1, L2 e L3. Our current understanding of their mechanism of action is limited by the resolution of the crystal structures available. To address this question we have attempted to crystallize a series of erythrocruorins from different species with a view to determining a crystal structure at higher resolution. In the first part of this thesis all chains of the erythrocruorin from *Glossoscolex paulistus* were completely sequenced and the structure of the full particle solved at 3.2Å, the highest reported to date. In the structure the four globin chains associate to form a heterotetramer, three of which unite to form a dodecameric cap. The latter associates with a heterotrimer of linkers to form the aforementioned protomer. The full structure permits a more detailed description of the contacts between subunits which are essential for particle stability. Furthermore, we describe metal binding sites (Zn^{2+} e Ca^{2+}) and glycosylation sites, some of which have not been reported previously. Subsequently the isolated *d* chain was crystallized and solved to 2.1Å. An analysis of the crystal contacts shows an arrangement which is completely different to that seen in the full particle. Instead of forming trimers, as seen in the complex, the isolated *d* chain associates to form a dimer across a crystallographic twofold axis making use of the interface normally used to associate with subunit *a*. These observations contributed to our understanding of the possible mechanisms of association of globin chains during the formation of the cap. Finally, the erythrocruorins from four other species were crystallized, which resulted in the preliminary determination of the structure of that from *Eisenia andrei* at 4.7Å.

Keywords: Haemoglobins. Erythrocruorins. *Glossoscolex paulistus*. Crystallography

LISTA DE FIGURAS

- Figura 1.1.1** - Os pioneiros da cristalografia de proteínas: (a) John Cowdery Kendrew e (b) Max Ferdinand Perutz. (c) Primeiro modelo da estrutura da subunidade beta da hemoglobina de cavalo (2dhb), apresentado por Perutz. Letras são atribuída a cada uma das 8 hélices do modelo, consolidando uma notação estrutural específica para globinas.32
- Figura 1.2.1** - Estrutura do grupo heme. O átomo de ferro é coordenado por 4 grupos pirrolínicos.....33
- Figura 1.2.2** - As diferentes classes de globinas. (a) globina (2dhb subunidade A), (b) globinas-sensor acoplada (2W31 – apenas o domínio globina está disponível) (c) globinas truncadas (1IDR) e (d) Flavo hemoglobina (4GLV).34
- Figura 1.2.3** - Estrutura de um globina (em representação Chainbow) contendo todas as 8 hélices descritas por Perutz, sítio de ligação do heme, que usualmente é um ambiente químico hidrofóbico, salvo pela presença das histidinas, proximal e distal.....35
- Figura 1.2.4** - (a), sobreposição da cadeia alfa da Hb humana com a globina de *L.luteus*. (b) um dendrograma demonstrando a divergência evolutiva entre as globinas mostradas no alinhamento (c) contendo espécies diversificadas.37
- Figura 1.3.1** - (a) Comparação entre o comportamento da hemoglobina e mioglobina humana. (b) Alterações no espectro de Uv-vis da Hb humana e virtude da entrada e saída do oxigênio. A transição da forma oxi para a forma deoxi é acompanhada de alterações no perfil das bandas Q.....38
- Figura 1.3.2** - Conformação dos grupos hemes na forma *oxi*(a) a *deoxi*(b). Note que existe um deslocamento do íon ferroso para fora do plano do anel com a saída do oxigênio (transição da forma *oxi* para *deoxi*). Na hemoglobina humana, as perturbações estruturais desses processo estendem-se para a estrutura terciária, com a reorientação da hélice F (c), até ao arranjo quaternário (d) resultando em uma rotação relativa entre os dímeros alfa-beta, como foi determinado por Perutz nos trabalhos pioneiros sobre estruturas de hemoglobinas.....40
- Figura 1.5.1** - (a) As duas teorias sobre a origem das hemoglobinas extracelulares em anelídeos. (b) Estrutura da HbLs. de *Oligobranchia mashikoi*. O complexo é formado por 24 subunidade de globinas.44
- Figura 1.5.2** - Uma comparação dos modelos gerados por microscopia eletrônica das eritrocruorinas do tipo I e II. (a) *Arenicola marina* (tipo II, eclipsada) e (b) *Lumbricus terrestris* (tipo I, não eclipsada)62.....45

Figura 1.6.1 - Classificação dos diferentes níveis de hierarquia das HBL Hbs/Eritrocruorinas. Unidade da HbLt (a) e (b), com simetria D6 (ou 622), contendo 12 protômeros. Ao todo são 144 cadeias de globinas e 36 cadeias tipo linkers. O protômero (c) por sua vez é composto por 4 diferentes tipos de globinas que primeiramente formam um heterotetrâmero <i>abcd</i> , e em seguida são associadas em uma estrutura dodecamérica (d) - cap ($[abc]_3[d]_3$). O <i>cap</i> por sua vez, está associado a um heterotrîmero de linkers.....	48
Figura 1.6.2 - Visão geral da estrutura dos linkers. A cadeia é constituída de 4 domínios, (1) Hélice 1, (2) hélice 2, (3) LDL-A e (4) β -barril. O domínio LDL-A possui um sítio de ligação de Ca^{2+} , assim como para o receptor de LDL humano, onde o Ca^{2+} é hexacoordenado.....	49
Figura 1.7.1 - Foto de um dos indivíduos coletados durante o período de 2010 e 2011. O <i>Glossosocolex paulistus</i> se destaca pelo seu tamanho incomum.....	52
Figura 1.7.2 - (a) <i>Perionyx excavatus</i> , também conhecida como violeta do Himalaia, (b) <i>Eudrilus eugeniae</i> - gigante africana, (c) <i>Eisenia andrei</i> – vermelha da Califórnia, é a que atinge o menor tamanho entre as espécies utilizadas neste trabalho.(d) <i>Rhinodrilus alatus</i> – minhocaçu, espécie nativa do vale do Paraopéba em Minas Gerais.....	54
Figura 2.3.1 - Representação do experimento de cristalização por difusão de vapor em gotas suspensas. A gota contendo a proteína em solução apresenta igual composição ao reservatório do poço, porém, com metade da concentração, dessa forma, esta, tende a perder solvente lentamente, até que as concentrações da gota e do poço adquiram uma condição de equilíbrio.....	62
Figura 3.1.1 - Etapas da purificação da HbGp. (a) corte na região ventral, (b) coleta da hemolinfa (c) centrifugação, eliminando o particulado sólido indesejável. (d) aplicação na coluna cromatográfica Superdex 200. (e) Perfil da curva obtida na separação pelo processo cromatográfico, evidenciando a saída da HbGp no volume de exclusão. (f) Gel de SDS PAGE (16%), demonstrando a composição do complexo da HbGp. Sete tipos básicos de cadeias de proteínas fazem parte do complexo. Três linkers (L1, L2 e L3) com massa de aproximadamente 25 kDa e quatro tipos de globinas, <i>a</i> , <i>b</i> , <i>c</i> e <i>d</i> , com massas entre 14 e 20 kDa.....	69
Figura 3.1.2 - Espectros de absorção da HbGp na forma <i>oxi</i> (vermelho) e na forma ciano (azul). O processo de oxidação de Fe^{2+} para Fe^{3+} é acompanhado de mudanças nas bandas Q, entre 500 e 600nm.....	70

- Figura 3.2.1** - Alinhamento geral das globinas obtidas para a HbGp e globinas de outro oligoquetas. Observa-se que as sequências da HbGp apresentam um padrão de resíduos similares ao das outras globinas. Isso inclui a conservação dos resíduos de cisteína, responsáveis pelas ligações dissulfeto entre subunidades (posições 3, 4, 129 e 134) presentes na HbGp. É importante mencionar que as sequências da HbEa e HbPe estão disponíveis apenas de forma parcial. A figura mostra também as regiões em conformação de hélice.....74
- Figura 3.2.2** - Alinhamento geral dos linkers obtidos para a HbGp e outros anelídeos. Observa-se que as sequências da HbGp apresentam um padrão de resíduos similares ao dos outros linkers. As posições 65 e 66 do alinhamento são marcadas pela abertura de um “gap” entre as sequências de eritrocruorinas do tipo I (não eclipsadas). Lmarinho corresponde as sequências de linkers de espécies marinhas, A.marina (HbAm) e A.pompojana (HbAp).....77
- Figura 3.3.1** - Cristais da ciano-HbGp obtidos em duas diferentes condições de cristalização (escala em μm). (a) 10% PEG8000 CaCl₂ 50mM Tris/HCl pH7,5, frágeis e (b) que apresentam uma vida útil muito curta (imagem coletada 5 dias após ao aparecimento dos cristais). (c) Cristais obtidos na condição 1,4M Tris/HCl 0,1M pH 8,0. (d) Mono cristal obtido na otimização da condição de cristalização (1,2M citrato de sódio 2,5mM CaCl₂ 50mM Tris/HCl pH7,5).80
- Figura 3.4.1** - Imagem do padrão de difração referente ao cristal da *ciano*-HbGp. Foram coletadas 220 imagens com uma oscilação de 0,5°. Cada imagem apresentou aproximadamente 13.800 reflexões e uma mosaicidade estimada de 0,55°. A predição de cela atribuída ao sistema cristalino ortorrômbico esta em boa concordância com as reflexões no padrão de difração, onde predominam reflexões parciais. A intensidade das reflexões sofre uma queda súbita, entre as regiões de 7 a 5Å no padrão de difração. Esses danos causados a rede são fruto do elevado teor de solvente (estimado em ~74%) que esses cristais possuem, maximizando os danos causados pelo resfriamento.....82
- Figura 3.5.1** - A unidade assimétrica é composta por três protômeros, cada um deles é colorido de uma cor diferente. Em branco estão representados os simétricos relacionados que formam a unidade biológica. (a) mostra a a unidade assimétrica ao longo do eixo b. (b) mostra o conteúdo da unidade assimétrica ao longo do eixo a. (c) mostra as operações de simetria presentes na unidade biológica isolada. Para ser compatível com a posição espacial que ocupa do grupo I222, a partícula se alinha de tal forma que um membro de cada família de eixos de ordem 2 é coincidente com os eixos cristalográficos paralelos a a e b e o eixo de ordem 6 (que possui um eixo de 2 como subelemento) ao longo do eixo c. (d) Cela unitária e a unidade biológica da HbGp.....84

Figura 3.5.2 - (a)Visão geral da densidade eletrônica (2Fo-Fc), onde é possível observar os elementos de estrutura secundária assumidos na HbGp. (b) Omit map calculado para a região das globinas, confirmando a presença do heme/ferro na região. (c) Omit map na região dos linker L3, correspondente a um sítio de ligação de cálcio (nível de contorno em 2 sigma).....	86
Figura 3.5.3 - Evolução do mapa de densidade eletrônica. A figura mostra a densidade mostra o mapa 2Fo – Fc, calculado em $1,5\sigma$ para a região do TRP132 da cadeia A. (a) mapa do modelo inicial e (b) mapa calculado para o modelo final, após o processo de refinamento.....	87
Figura 3.5.4 - Diagrama de Ramachandran para a estrutura da ciano-HbGp (1.01% de resíduos fora da região permitida).....	88
Figura 3.5.5 - Mapas calculados para algumas regiões das subunidades tipo a, ainda com a sequencia da isoforma a1 (estruturas em cinza), evidenciando que a sequência de a2 (estruturas em azul) melhor concorda com a densidades eletrônica. (2Fo-Fc – mapa cinza – sigma = 1.0 e Fo-Fc em verde – sigma = 3 / vermelho – sigma = -3).....	89
Figura 3.5.6 - Mapa de densidade eletrônica (2Fo-Fc) calculado para a região da globina a. Detalhes importantes podem ser observados como: (a), grupo heme, (b) ligações dissulfeto intramoleculares, (c) ligações dissulfeto intermoleculares. (d) Modificação no espectro de absorção após a exposição ao raio X, a amostra inicia-se como um típico espectro da forma <i>cianometa</i> , evoluindo para a uma configuração semelhante a da forma reduzida oxi-HbGp.	90
Figura 3.5.7 - A alta qualidade dos mapas de densidade eletrônica resultaram na determinação de novos grupos químicos ligados ao Linker L3: (a) Íon Zn^{2+} coordenado internamente no domínio LDL-A (sítio de zinco 1). (b) Íon Zn^{2+} coordenado na interface entre linker L3 e linker L2 do protômero adjacente (sítio de zinco 2), (c) Glicosilação – NAG; e (d) Íon de cálcio ligado ao domínio LDL-A.....	91
Figura 3.5.8 - Densidades eletrônicas do modelo cristalográfico da HbGp onde não foram atribuídos espelhadores. (a) região equivalente ao sítio de Zn^{2+} encontrado na subunidade L2 da estrutura da HbLt. A presença dos resíduos de arginina é conflitante com a coordenação de um cátion metálico nessa posição. (b) Região do β -barril das subunidades L3. (c) Interface entre o domínio LDL-A e o β -barril das subunidades L3. (d) Interface entre os domínios LDL-A das subunidades L1 e L2 do mesmo protômero.....	92
Figura 3.6.1 - Partícula inteira da HbGp (simetria 622), com um diâmetro de aproximadamente 286 Å e uma altura de 190, comparada a outras estruturas disponíveis no banco de dados.	93

- Figura 3.6.2** - A HbGp é uma bicamada hexagonal que contém 12 protômeros. Ao todo são 144 cadeias de globinas e 36 estruturas tipo linkers. Os protômeros são constituídos por um dodecâmero de globinas (cap) e um heterotrímero de linkers (L1, L2 e L3). As cores utilizadas para diferenciar os tipos de subunidades que compõem o complexo da HbGp foram adotadas como padrão neste trabalho.....94
- Figura 3.6.3** - Sobreposição das globinas da HbGp (em cores), com suas homólogas equivalentes da HbLt (em cinza). Logo abaixo estão representados os fatores de temperatura. As cores “quentes”, como amarelo, laranja e vermelho, representam as regiões de maior desordem.....95
- Figura 3.6.4** - Primeira esfera de resíduos ao redor do sexto sítio de coordenação do grupo heme. (a) globinas tipo a, (b) globinas tipo b, (c) globinas tipo c e (d) globinas tipo d. Existe uma predominância de resíduos hidrofóbicos.....96
- Figura 3.6.5** - Os dímeros formados *a/d* e *b/c*, possuem uma interface de contato conservada. Tal interface é conhecida para outras estruturas de hemoglobinas, como mencionado anteriormente⁸⁹. Área total entre *a/d*(HbGp) = 910,5Å² e área total entre *b/c* (HbLt) = 1146,9Å².....97
- Figura 3.6.6** - Formação do heterotetrâmero. Para a formação desta estrutura, cada dímero contribui com um sulco, formado pelas hélices B e G, e uma protuberância, formada pela hélice A. Os resíduos de cisteína presentes nos loops que conectam as hélices G e H das subunidades tipo a e b, formam uma ligação dissulfeto conectando os dois dímeros.....98
- Figura 3.6.7** - Hierarquia de associação das subunidades de globinas pra a formação do cap. As quatro subunidades básicas se associam inicialmente nos dímeros *a/d* e *b/c*, que posteriormente formam os heterotetrâmero *abcd*, onde a primeira ligação dissulfeto é formada. Três heterotetrâmeros se associam formando a estrutura do *cap*, onde a segunda ligação dissulfeto é formada.....99
- Figura 3.6.8** - Comparação entre as estruturas dos *caps* da HbGp, HbLt, Hb_LRp e HbLOm. Apesar da baixa identidade sequencial na comparação entre as sequências que formam esses sistemas, a estrutura quaternária é altamente conservada. As Hb_L apresentam sítios de ligação de metais que não são conservados para os *caps* das eritrocruorinas de oligoquetas.....101
- Figura 3.6.9** - Uma sobreposição dos pares de subunidades linkers da HbGp (cores) e HbLt (cinza). As estruturas revelam que as porções N-terminais apresentam as maiores divergências na sobreposição. Os domínios LDL-A e β-barril apresentam enovelamentos muito conservados, com as ligações dissulfeto e sítios de ligação de Ca²⁺ estruturalmente equivalentes.103
- Figura 3.6.10** - Comparação entre os domínios LDL-A das subunidades linkers e o domínio homólogo no receptor LR-5 humano.104

- Figura 3.6.11** - Sítios de ligação de grupos prostéticos na estrutura da HbGp. (a) sítio de zinco 1, interno ao domínio LDL-A e (b) sítio de zinco 2, interfacial entre a subunidade L2 e L3. (c) glicosilação e (d) sítio ativo da enzima Cu/ZnSOD humana. Os sítios de zinco encontrados na HbGp não são equivalentes/similares ao da enzima Cu/ZnSOD.....105
- Figura.3.6.12.-**Formação do heterotrímero de linkers, (a) associação das subunidades L1, L2 e L3. (b) Os principais contatos entre as subunidades que formam o heterotrímero acontecem na região do coiled-coil. Não há um padrão claro de β -ramificações e/ou resíduos não ramificados nas posições “a” e “d”, o que é consistente com a formação de um coiled-coil trimérico.....107
- Figura 3.6.13** - Sítios de interação entre as porções globulares dos linkers no heterotrímero. Os sítios de interação ocorrem através dos mesmos domínios, mas não há conservação dos resíduos envolvidos. Esses fracos contatos sugerem que a porção *coiled-coil* é a que de fato atua mantendo as subunidades dos linkers unidos.....108
- Figura 3.6.14** - O protômero é uma associação da estrutura dodecamérica de globinas (*cap*) com o heterotrímero de linkers. Na junção dessas duas estruturas oligoméricas forma-se uma grande cavidade com volume de $39.207 \text{ \AA}^3$, o que equivale ao volume de duas subunidades de globina.....109
- Figura 3.6.15** - Sítios de interação entre as porções globulares dos *linkers* no heterotrímero. Os sítios de interação ocorrem através dos mesmos domínios mas não há conservação dos resíduos envolvidos nesta região. A interação entre o domínio LDL de cada *linker* com a subunidade tipo *b* ocorre via a ponte salina formada entre os resíduos Arg34 da subunidade *b* com os resíduos Asp 88-L1/ 95-L2 / 85-L3 das subunidades linkers.....111
- Figura 3.6.16** - Os cinco grupos de canais de acesso a grande cavidade central formada pela interação entre a estrutura do *cap* e o heterotrímero de *linkers*. (a) canais de acesso formados na estrutura do *cap* (grupos 1 e 2). (b) canais de acesso formados no heterotrímero de linkers (grupos 3 e 4). (c) canais de acesso formados pela interação entre a estrutura do *cap* e o heterotrímero de *linkers*, grupo 5.....113
- Figura 3.6.17** - Interações entre protômeros de uma mesma camada hexagonal. (a) Estrutura da monocamada hexagonal onde metade dos *caps* foram removidos, expondo a porção globular dos *linkers* de cada protômero. (b) Interação entre protômeros vizinhos envolvendo a porção globular dos *linkers* L2 e L3, Além de compartilharem as fitas que constituem o β -barril a interface forma um sítio de ligação de zinco (sítio de zinco 2). (c) Contatos entre os *caps* de protômeros vizinhos.....115

- Figura 3.6.18** - Interação entre protômeros de camadas hexagonais diferentes. Para melhor visualização, parte das estruturas dos *caps* foram removidas. (a) visão frontal (b) visão lateral. (c) Interações entre os protômeros parcialmente sobrepostos, que ocorrem através do domínio β -barril das subunidades L1 e (d) através das porções *coiled-coil* da subunidades L2 e L3. (e) Interação entre um protômero com o protômero adjacente ao sobreposto da outra camada hexagonal. Este contato ocorre através dos resíduos Thr12, Gln17 e Asn16, presentes no primeiro domínio das subunidades L1.117
- Figura 4.1.1** - Resultados da purificação da subunidade d da HbGp (d-HbGp). (a) Gel filtração, com destaque para o pico coletado entre 70-85 mL que contém a d-HbGp. (b) SDS PAGE, 1 – marcador de massa molecular, 2 – HbGp na forma íntegra e 3 – Fração coletada entre os volumes 70 e 85 de gel filtração contendo a d-HbGp. (c) Análise espectroscópica da d-HbGp isolada, as bandas Q evidenciam que a forma reduzida (*oxi*) é a que predomina na amostra.....119
- Figura 4.2.1** - Cristais obtidos para a d-HbGp na condição 1,4 citrato de sódio, Tris/HCl pH7,5, com dimensões próximas a 0.5 μ m.....120
- Figura 4.3.1** - Imagem de difração do cristal da d-HbGp com resolução de 2.1 Å na borda, 0,4° de oscilação e mosaicidade inicial estimada de 0.64°. A imagem apresenta alguns pontos incoerentes com a predição, resultados da presença de um cristal satélite (pequeno cristal, ou fragmento de, anexado ao cristal que está sendo coletado) mas que não comprometeu a qualidade dos dados obtidos.121
- Figura 4.4.1** - Algumas das posições importantes nos mapas de densidade (cinza – 2Fo-Fc e $\delta = 1.0$, vermelho – Fo-Fc e $\delta = -3.0$, verde –Fo-Fc e $\delta = 3.0$) comprovando que a sequencia d1 melhor descreve o modelo. As fases dos mapas em questão foram calculadas com a sequencia d2.....123
- Figura 4.4.2** - Principais parâmetros que validam a qualidade do modelo final obtido. (a) diagrama de Ramachandran, onde nenhum resíduo encontra-se fora das regiões permitidas, e (b) Uma comparação dos valores obtidos ao final do refinamento com outras estruturas depositadas no banco de dados PDB. Os valores em cor vermelha referem-se ao mínimo é máximo encontrado para o dado parâmetro (considerando a faixa de resolução em questão). A estrutura da d-HbGp apresenta valores dentro dos considerados aceitáveis em todos os aspectos em questão.....124
- Figura 4.4.3** - Mapa de densidade eletrônica calculado (2Fo - Fc) para o modelo da d-HbGp. (a) visão geral, (b) região do grupo heme onde foi modelada a molécula de oxigênio ocupando o sexto sítio de coordenação do ferro. (c) ligação dissulfeto intramolecular.....125

- Figura 4.5.1** - Principais contatos cristalinos encontrados para a estrutura cristalográfica da *d*-HbGp. (a) ao longo do eixo a, (b) ao longo de eixo b, formando o dímero d/d, que é equivalente ao dímero a/d na estrutura da HbGp e (c) ao longo do eixo c, que resulta no rearranjo dos loops que conectam as hélices A e B, e G e H.127
- Figura 4.5.2** - Sobreposição da *d*-HbGp com a subunidade *d* presente na unidade biológica. (a) Existe grande coerência na sobreposição dessas duas estruturas com uma pequena diferença para a posição do resíduo Trp 32, que para a forma *oxi* da *d*-HbGp sofre um pequeno deslocamento. (b) Orientação do plano do anel imidazólico do resíduo de His94 em relação ao plano do grupo heme comparado com a Hb humana em (c).....129
- Figura 4.5.3** - O Mecanismo mostra as interfaces em que a subunidade d esta envolvida. Na estrutura do cap, o trímero d_3 faz contatos através das hélices B e G. Quando a subunidade *d* é isolada, seus frágeis contatos se desfazem, prevalecendo em solução a forma monomérica. Com o aumento da concentração, a d-HbGp se associa formando um dímero, que interage através das hélices E e F (incluindo o grupo heme). Essa interface é equivalente ao dímero a/d encontrado na estrutura do cap.....130
- Figura 5.1.1** - (a) Análise de SDS PAGE das frações purificadas contendo as hemoglobinas, HbPe, HbEe, HbEa e HbRa. A HbEa é a que apresenta um padrão de massa que mais diverge das outras Hbs purificadas neste trabalho. (b) Resultado da análise por UAC, a amostra de HbEa apresentam dois picos em S=10 (referentes ao *cap* isolado) e S~65, que equivale ao coeficiente de sedimentação da partícula inteira.....132
- Figura 5.2.1** - Cristais obtidos para as eritrocruorinas de 4 novas espécies. (a) Os cristais da HbRa, obtidos em 10% PEG8000 e 50mM Tris/HCl pH7,0. (b) Os cristais de HbEe, obtidos na condição 20 % MPD Tris/HCl pH7,0. (c) Os cristais de HbEa obtidos em sulfato de amônio 1,4 M, cacodilato pH 6,5. (d) Cristais de HbEa, obtidos em PEG 3350 10% cacodilato pH 7,0.134
- Figura 5.3.1** - Imagens dos padrões de difração coletados e estatística do processamento (Rmerge e Intensidade) por faixa de resolução para (a) HbPe, onde é possível ver a orientação do eixo maior, $c=608,17\text{\AA}$, que é caracterizado pela proximidade dos *spots*. (b) HbEa.....135
- Figura 5.4.1** - A unidade assimétrica é composta por quatro protômeros, cada um deles colorido por uma cor diferente. Em cinza claro estão representados os simétricos relacionados que formam a unidade biológica. (a) mostra a unidade assimétrica ao longo do eixo c e as operações de simetria que relacionam o grupo especial I23. (b) mostra o conteúdo da unidade assimétrica ao longo do eixo de simetria de ordem três que corta a diagonal da cela.....137

Figura 5.4.2 - Comparação entre os mapas de densidade eletrônica obtidos ($2F_o - F_c$ – $\sigma = 1,5$) para a HbEa (4,7Å) e ciano-HbGp (3,2Å). (a) região do coiled-coil. (b) domínio β -barril da subunidade L1 e (c) Região do sítio de ligação da globina tipo *d*.....**138**

Figura 5.4.3 - Mapa de densidade eletrônica ($2F_o - F_c$ – $\sigma = 1,5$) calculados para o modelo cristalográfico da HbEa onde as sequências das subunidades foram preditas como glicosiladas (a) Asn 86 das subunidades α , (b) Thr das subunidades L1 e (c) Asn 161 das subunidades L2.....**139**

ÍNDICE DE TABELAS

Tabela 3.2.1 - Comparação das propriedades físico-químicas preditas para as sequências da HbGp e as da HbLt. Os valores preditos pra a HbLt são mostrados entre parênteses. Nota-se que as isoformas <i>d2</i> e <i>L1b</i> apresentam as propriedades preditas mais divergentes em relação aos grupos as quais pertencem. As sequências de peptídeos sinais são caracterizadas pela alta concentração de resíduos hidrofóbicos, o símbolo “.” é usado para designar a posição de clivagem predita pelo servidor signalP4.1.....	73
Tabela 3.2.2 - Uma comparação da identidade entre as sequências homólogas de globinas pra as eritrocruorinas HbGp, HbLt, HbEa e HbPe. O simbolo * significa que a sequencia em questão esta incompleta.....	74
Tabela 3.2.3 - Uma comparação entre as identidades das sequências homólogas de linkers pra as eritrocruorinas HbGp, HbLt e HbEa. Não existe sequências disponíveis referentes aos linkers da HbPe.....	79
Tabela 3.4.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução (3.2 Å).....	84
Tabela 3.5.1 - Simétricos relacionados. O modelo da HbGp possui uma unidade assimétrica com 45 subunidades. As letras na tabela correspondem a nomenclatura das cadeias no arquivo PDB.	88
Tabela 3.5.2 - Estatísticas do refinamento e estereoquímica do modelo.....	89
Tabela 3.6.1 - Mostra os valores de rmsd resultantes da sobreposição entres as subunidades de globina. Os valores preditos para as globinas através da equação de Chothia e Lesk (mostrados entre parenteses) são maiores do que os encontrados na sobreposição das estruturas.....	97
Tabela 3.6.2 - Mostra os valores de rmsd resultantes da sobreposição entres as subunidades linkers. Valores com * indicam que não foi considerada a porção do coiled-coil para a sobreposição e cálculo do rmsd.....	107
Tabela 4.3.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução.....	123
Tabela 4.4.1 - Estatísticas do refinamento e geometria do modelo.....	125
Tabela 5.1.1 - Massas obtidas por UAC para as hemoglobinas gigantes utilizadas neste trabalho. A massa estimada para a HbEa é maior do que para a HbGp, HbRa e HbPe.	134

Tabela 5.3.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução.....	137
---	------------

Tabela 5.4.1 - Valores obtidos de LLG e R para cada etapa do processo de atribuição das fases por substituição molecular.....	138
--	------------

LISTA DE ABREVIATURAS E SIGLAS

Hb - Hemoglobina

Mb – Mioglobina

2,3-DPG - 2,3-difosfoglicerato

GSC - globinas-sensor acopladas

FHbs – flavohemoglobinas

LDL-A - *Low-density lipoprotein receptor domain class A*

SDS-PAGE - *sodium dodecyl sulfate polyacrylamide gel electrophoresis*

ICP-AES - *Inductively Coupled Plasma -Atomic Emission Spectrometry*

HbGp - Hemoglobina de *Glossoscolex paulistus*

d-HbGp - Subunidade *d* isolada da HbGp

HbEa - Hemoglobina de *Eisenia andrei*.

HbPe - Hemoglobina de *Perionyx excavatus*

HbRa - Hemoglobina de *Rhinodrilus alatus*

HbLt - Hemoglobina de *Lumbricus terrestris*

Hb_p - hemoglobinas “pesadas” - eritrocruorinas de vestimentíferos com ~3,6MDa

Hb_L - hemoglobinas “leves” de vestimentíferos com ~400kDa

SUMÁRIO

1 INTRODUÇÃO.....	31
1.1 Proteínas e a biologia estrutural.	31
1.2 Globinas.....	32
1.3 Hemoglobinas e mioglobinas.....	37
1.4 Hemoglobinas extracelulares de invertebrados.	41
1.5 Hemoglobinas gigantes em anelídeos.....	42
1.6 A hemoglobina gigante de <i>Lumbricus terrestris</i>	47
1.7 A hemoglobina gigante de <i>Glossoscolex paulistus</i> e outras espécies de interesse....	51
1.8 Hemoglobinas extracelulares como potenciais substituintes para sangue.	54
1.9 Objetivos.....	55
2 METODOLOGIA.....	57
2.1 Coleta dos animais utilizados.....	57
2.1.1 Purificação das oxi-Hbs íntegras e a subunidade d da HbGp.	57
2.1.2 Obtenção das formas meta- e ciano-meta da HbGp.....	58
2.1.3 Análise de plasma – composição das amostras.....	58
2.2 Sequenciamento.....	59
2.2.1 Extrações do RNA total do <i>G. Paulistus</i>	59
2.2.2 Sequenciamento Genômico no Illumina Genome Analyzer.....	59
2.2.3 Montagem do transcriptoma e identificação das sequências.....	60
2.3 Ensaios de cristalização.....	61
2.4 Coleta e processamentos dos dados.....	63
2.5 Substituição molecular.....	63
2.6 Refinamento.....	65
2.7 Análises e comparações estruturais.....	67
3 RESULTADOS E DISCUSSÃO – HbGp.....	69
3.1 Purificação	69
3.2 Sequenciamento.....	71
3.2.1 Sequências de globinas.....	73
3.2.2 Sequências - linkers.....	76
3.3 Cristalização.....	79
3.4 Coleta e processamentos dos dados.....	80
3.5 Substituição molecular e refinamento do modelo.....	83
3.6 Estrutura da HbGp – uma visão geral.....	92
3.6.1 Estrutura da HbGp – Globinas e a formação do Cap.....	94
3.6.2 Estrutura da HbGp – Linkers e a formação do coiled coil.....	102
3.6.3 Formação do protômero.....	109
3.6.4 Estrutura da HbGp – Interação entre protômeros de uma mesma camada hexagonal.....	114
3.6.5 Estrutura da HbGp – Interação entre protômeros de camadas hexagonais diferentes.....	116
4 RESULTADOS E DISCUSSÃO – HbGp subunidade d.....	119
4.1 Purificação.....	119
4.2 Cristalização.....	120
4.3 Coleta e processamentos dos dados.....	120
4.4 Substituição molecular e refinamento do modelo.....	122
4.5 Estrutura da subunidade d e seus contatos cristalinos.....	125
5 RESULTADOS E DISCUSSÃO – Outras Hemoglobinas	133
5.1 Purificação.....	133
5.2 Cristalização.....	134

5.3 Coleta e processamentos dos dados.....	135
5.4 Substituição molecular e estrutura preliminar da HbEa.....	137
6 CONCLUSÕES E Perspectivas FUTURAS.....	141
6.1 Conclusões.....	141
6.2 Perspectivas.....	144
REFERÊNCIAS.....	145
ANEXOS.....	159

1 INTRODUÇÃO

1.1 Proteínas e a biologia estrutural.

Em 1953, Max F. Perutz mostrou que o problema das fases em cristalografia de proteínas poderiam ser resolvido comparando-se os dados obtidos na presença e na ausência de átomos pesados na rede cristalina (método conhecido como substituição isomórfica múltipla)¹. Em março de 1958 John Kendrew e colaboradores reportaram a estrutura cristalográfica da mioglobina (Mb) de baleia cachalote^{2,3} (*sperm whale* - *Physeter macrocephalus* - código de acesso: 1MBO) com a modesta resolução de 6Å. Apesar disso foi possível pela primeira vez “enxergar” a estrutura tridimensional de uma proteína. Detalhes sobre o enovelamento permitiram responder várias questões propostas pela comunidade científica da época, como por exemplo a existência do modelo de estrutura secundária do tipo alfa hélice, proposta pelo químico norte americano Linus Pauling. Em 1959, Perutz reportou a estrutura da hemoglobina (Hb) de equinos^{4,5} (*Equus caballus* - código de acesso: 2DHB). Essa estrutura era ainda mais desafiadora do que a reportada anteriormente por Kendrew, uma vez que se trata de um tetrâmero. Perutz foi ainda mais longe, determinando as estruturas cristalográficas Hb na forma *oxi-Hb*, ligada a oxigênio e na forma *deoxi-Hb*⁶, não ligada a oxigênio. Os trabalhos pioneiros de Perutz e Kendrew, reportando pela primeira vez estruturas de proteínas, representaram o surgimento da biologia estrutural como importante ramo da ciência moderna. Em 1962, John Cowdery Kendrew e Max Ferdinand Perutz, ganharam o Premio Nobel de Química pelos estudos estruturais de proteínas globulares. Atualmente, 54 anos depois da data de publicação da estrutura da Mb por Kendrew, o banco de dados internacional de estruturas – *protein data bank* – PDB, conta com mais de 80.000 estruturas depositadas, e dessas, aproximadamente 88% foram resolvidas utilizando-se a técnica de difração de raios X.

Figura 1.1.1 - Os pioneiros da cristalografia de proteínas: (a) John Cowdery Kendrew e (b) Max Ferdinand Perutz. (c) Primeiro modelo da estrutura da subunidade beta da hemoglobina de cavalo (2dhb), apresentado por Perutz. Letras são atribuída a cada uma das 8 hélices do modelo, consolidando uma notação estrutural específica para globinas.

1.2 Globinas

As globinas constituem uma das famílias de proteínas mais estudadas. Essas proteínas incorporam um enovelamento característico (*globin folding*) bastante conservado. Devido a grande e variada quantidade de sequências que apresentam o mesmo enovelamento característico, globinas são um bom exemplo de que a estrutura terciária é uma propriedade mais conservada do que a identidade sequencial. Seu enovelamento consiste basicamente em um arranjo de seis/oito α -hélices, formando uma estrutura do tipo sanduíche α -hélice 3/3 (lê-se três sobre três). Um subgrupo especial de globinas, conhecidas com globinas truncadas, apresentam deleções em algumas regiões da estrutura e resultam em um enovelamento diferente, do tipo sanduíche α -hélice 2/2⁷. As globinas apresentam uma cor vermelha característica, devido à presença do grupo prostético heme (**figura 1.2.1**). São encontradas em todos os reinos conhecidos⁸, desempenhando funções diversificadas de acordo com o organismo em questão. O conceito mais aceito é que as globinas evoluíram de um ancestral comum em torno de 3,5 bilhões de anos atrás^{9,10}, dividindo-se em três subgrupos básicos (**figura 1.2.2**): globinas de um único domínio (incluem hemoglobinas, mioglobinas e globinas truncadas), flavohemoglobinas (FHbs) e globinas-sensor acopladas (GSC). Bactérias possuem todos os três tipos de globinas, enquanto *archaea* não dispõem de flavohemoglobinas, e por fim, eucariotos não dispõem de globinas-sensor acopladas¹¹.

Figura 1.2.1 - Estrutura do grupo heme. O átomo de ferro é coordenado por 4 grupos pirrolínicos.

Globinas contendo um único domínio, são majoritariamente conhecidas por desempenharem o papel de carreadoras de oxigênio (embora algumas também atuem em outras vias, como no metabolismo do óxido nítrico - NO). Em leguminosas, são encontradas subclasses de globinas denominadas leghemoglobinas¹². São responsáveis por armazenar o oxigênio que será posteriormente utilizado pela maquinaria simbiótica de fixação de nitrogênio. Apesar dessas globinas apresentarem em média somente 20% de identidade quando comparadas à suas homólogas nos vertebrados, estas compartilham um enovelamento muito similar. Foi reportado também a existência de globinas não simbióticas em raízes, caules e sementes germinadas de muitas outras plantas, mostrando que essas proteínas também estendem-se amplamente no reino vegetal^{14,15}. Globinas truncadas por sua vez, desempenham um papel diferente no meio biológico, como no caso do patógeno *Mycobacterium tuberculosis* (trHbN), onde estão associadas a patogenicidade desse organismo e atua como NO-dioxigenase, metabolizando óxido nítrico gerado pelos macrófagos¹⁶. As proteínas do subgrupo das FHbs são compostas de dois domínios, que podem fazer parte ou não de uma mesma cadeia. O primeiro domínio é o do tipo globina, e no caso de proteínas que apresentam os dois domínios em uma mesma cadeia, este encontra-se na porção N-terminal. O segundo domínio pertence à família das proteínas ferredoxina NADP- redutases (FNR), cujo enovelamento não tem relação com o apresentado pelas globinas. As FHbs com domínios fusionados possuem um tamanho médio de 400 aminoácidos. As FHbs atuam como potentes dioxigenases no metabolismo do óxido nítrico, com atividade vinte vezes maior do que a desempenhada pela Mb dos vertebrados para essa função^{17,18}.

Figura 1.2.2 - As diferentes classes de globinas. (a) globina (2dha subunidade A), (b) globinas-sensor acoplada (2W31 – apenas o domínio globina está disponível) (c) globinas truncadas (1IDR) e (d) Flavo hemoglobina (4GLV).

O subgrupo das GSC abrange globinas fusionadas a outros domínios, o que resultam em tamanhos variados (300 a 700 resíduos de aminoácidos). A região N-terminal apresenta o enovelamento característico de globinas (semelhante a Mb) e a região C-terminal apresenta um domínio homólogo ao domínio sinalizador citoplasmático de quimiorreceptores bacterianos. Sua possível função seria ligar-se ao oxigênio diatômico ou outros ligantes gasosos e transmitir o sinal através de um domínio sinalizador. Esse grupo de proteínas foi reportado recentemente e sua caracterização estrutural é ainda escassa. O banco de dados, PDB, conta com apenas uma única estrutura depositada, e esta, refere-se apenas ao domínio globina da globina-sensor acoplada de *Geobacter sulfurreducens*¹⁹.

Com a resolução da estrutura cristalográfica da Mb e logo em seguida, da Hb, Perutz propôs uma notação específica para o enovelamento característico de globinas. A estrutura é composta de oito hélices em um arranjo do tipo 3/3, que forma uma cavidade hidrofóbica onde o grupo heme se encaixa. As hélices são denotadas em A, B, C (onde C é uma hélice do

tipo 3₁₀), D, E, F (em mamíferos essa hélice é fragmentada em duas hélices consecutivas, F' e F), G e H. Várias globinas não apresentam a pequena hélice D. A vizinhança do grupo heme é comumente constituída de resíduos hidrofóbicos, com exceção de duas histidinas, a distal (E7, sétimo resíduo da hélice E) e a proximal (F8, oitavo resíduo da hélice F), como mostra a **figura 1.2.3**.

Figura 1.2.3 - Estrutura de uma globina (em representação *Chainbow*) contendo todas as 8 hélices descritas por Perutz, sítio de ligação do heme, que usualmente é um ambiente químico hidrofóbico, salvo pela presença das histidinas, proximal e distal.

A **figura 1.2.4** mostra um alinhamento envolvendo diferentes globinas e a divergência evolutiva entre elas. As cadeias *alfa* e *beta* da Hb humana partilham entre si, 42% de identidade. Quando comparadas com a sequência paróloga (homólogos dentro de um mesmo organismo) da Mb humana, as cadeias *alfa* e *beta* da Hb partilham somente 25% de identidade. A diferença torna-se ainda mais crítica a medida que sequências homólogas de espécies mais distantes evolutivamente são comparadas. A princípio, esperaríamos que estruturas de sequências mais conservadas entre si, fossem também conservadas, e que estas iriam se diferenciando conforme a diferenciação das sequências. Apesar de sequências pouco conservadas, as globinas partilham de enovelamentos muito próximos. A subunidade *alfa* da Hb humana por exemplo, partilha apenas 16,6% de identidade com a homóloga do vegetal *Lupinus luteus*, no entanto quando as estruturas dessas duas proteínas são sobrepostas espacialmente, mostram um enovelamento altamente conservado ($rmsd = 3.4\text{\AA}$).

Figura 1.2.4 - (a), sobreposição da cadeia alfa da Hb humana com a globina de *L. luteus*. (b) um dendrograma demonstrando a divergência evolutiva entre as globinas mostradas no alinhamento (c) contendo espécies diversificadas.

Então, como sequências tão pouco conservadas podem resultar em enovelamentos tão próximos? A resposta à essa pergunta está na própria maneira como ocorre o enovelamento de globinas. Estruturas de proteínas são estabilizadas pelo empacotamento de resíduos no seu interior (sendo estes majoritariamente hidrofóbicos). No caso das globinas, grande parte dos resíduos no interior da estrutura estão nas interfaces entre as hélices. Dessa forma cinco empacotamentos entre hélices comumente ocorrem: A/H, B/E, B/G, F/H e G/H, porém outros contatos menos gerais podem ocorrer. Dessa forma, globinas de domínio único do tipo 3/3, apresentam um padrão conservado para as interfaces entre hélices, que ocorre de acordo com o modelo denominado “cristas em sulcos” (do inglês: *ridges-into-grooves*), Empacotamentos de hélices em proteínas globulares geralmente obedecem a esse modelo²⁰.

1.3 Hemoglobinas e mioglobinas

Sem dúvida alguma, os membros mais proeminentes da família das globinas são a mioglobina (Mb) e a hemoglobina (Hb). Ambas apresentam como papel principal, ligar-se ao oxigênio de forma reversível, atuando como proteínas carreadoras. Apesar de partilharem de funções e características semelhantes, o comportamento bioquímico de cada uma é singular. A mioglobina, é uma proteína monomérica, contendo aproximadamente 150 resíduos de aminoácidos (massa em torno de 16 kDa) e é uma das proteínas mais abundantes nos tecidos musculares, o que confere a este, uma cor vermelha intensa. Nos mamíferos, ela armazena oxigênio nos tecidos, que posteriormente será utilizado no metabolismo aeróbico das células. Mamíferos que vivem em grande/total contato com ambientes aquáticos possuem mioglobinas em altas concentrações em seus tecidos musculares, restando oxigênio por grandes períodos durante o mergulho. A Mb não é uma enzima, mas se encaixa perfeitamente no comportamento cinético descrito por Michaelis–Menten²¹. A hemoglobina por sua vez, possui a função de transportar o oxigênio coletado do meio externo, dessa forma facilitando o processo de difusão, permitindo que o oxigênio supra a demanda exigida pelos tecidos do organismo. As Hbs são encontradas em diversos estados oligoméricos quando observamos gama ampla de organismos que as possui. Nos vertebrados inferiores, como lampreias por exemplo, a hemoglobina ligada ao oxigênio é encontrada na forma monomérica e torna-se dimérica com a liberação do mesmo²². Por outro lado, em vertebrados superiores, as hemoglobinas são simétricas. Nos seres humanos adultos, a hemoglobina é uma estrutura tetramérica, intracelular, composta por duas subunidades do tipo *alfa* e duas do tipo *beta*. Essas subunidades são globinas com enovelamento muito similar ao da Mb, porém estão associadas de tal forma que seu comportamento frente a disponibilidade de oxigênio não obedece o modelo proposto por Michaelis–Menten mas apresentam um efeito de cooperatividade entre suas subunidades²³, de forma que, quando uma das quatro cadeias de globina liga-se à uma molécula de oxigênio, as mudanças conformacionais resultantes dessa interação influenciam na afinidade pelo ligante das cadeias restantes. A recíproca é verdadeira, de forma que a desocupação de um dos sítios facilita a desocupação dos demais. Isso permite que a hemoglobina apresente alta afinidade por oxigênio nos pulmões, onde há um excesso do ligante (pressão parcial de $pO_2 = 100$ torrs), e baixa afinidade nos tecidos, onde há uma baixa concentração de oxigênio livre (pressão parcial de $pO_2 = 20$ torrs). Uma comparação entre o comportamento das Mbs e Hbs pode ser observado na **figura 1.3.1 - a**. A

forma ligada, ou *oxi*-Hb apresenta uma conformação designada como “relaxada” (R), ao mesmo tempo que a forma desligada, ou *deoxi*-Hb apresenta uma conformação denominada de “tensionada” (T).

Figura 1.3.1 - (a) Comparação entre o comportamento da hemoglobina e mioglobina humana. (b) Alterações no espectro de Uv-vis da Hb humana e virtude da entrada e saída do oxigênio. A transição da forma *oxi* para a forma *deoxi* é acompanhada de alterações no perfil das bandas Q.

Quando purificada, a Hb humana apresenta uma afinidade por oxigênio muito maior do que no interior das hemácias, essa diferença dramática se deve a presença do 2,3-difosfoglicerato (2,3-DPG). Esse composto, muito aniônico, está presente nas hemácias em concentração equivalente a de hemoglobinas (~2 mM). Sem o 2,3-DPG a Hb torna-se um transportador menos eficiente, liberando somente 8 % de sua carga nos tecidos. Isso acontece porque o 2,3-DPG se liga em um bolsão formado apenas na conformação T²⁴. Assim, quando ligado a Hb, o 2,3-DPG estabiliza a forma T, reduzindo de forma eficaz a afinidade pelo oxigênio. As Hbs são sensíveis a variação do pH, sendo esse efeito conhecido por: efeito Bohr. Em 1904, Christian Bohr demonstrou que a diminuição do pH de 7,4 para 7,2, diminui a afinidade da Hb humana pelo oxigênio. Dessa forma nos pulmões, onde o pH é 7,4, a Hb fica saturada e a medida que migra para os tecidos musculares em atividade, encontra um ambiente químico com pH em torno de 7,2. Essa pequena variação é capaz de alterar sensivelmente a capacidade de liberação do oxigênio de 66% para 77%. A entrada e saída do oxigênio causa alterações estruturais importantes nas globinas. Em condições normais, o íon de ferro ligado ao grupo heme está no estado ferroso (Fe²⁺). Este íon de ferro pode formar duas ligações adicionais, uma em cada lado do plano do heme. Esses sítios são chamados de

quinto e sexto sítios de coordenação. O quinto sítio é ocupado pela cadeia lateral da histidina presente na hélice F (histidina proximal). Na forma *deoxi*, o sexto sítio permanece desocupado, deslocando o íon ferroso aproximadamente 4° para fora do plano da porfirina. Isso acontece, pois o ferro nesta forma é um pouco maior do que a cavidade do centro do anel. A ligação de uma molécula de oxigênio ao sexto sítio de coordenação leva a um rearranjo dos elétrons do ferro, de modo que o íon torna-se efetivamente menor, permitindo mover-se para o plano do anel. As alterações na estrutura eletrônica são visíveis através de espectroscopia de absorção UV-vis, como mostra a **figura 1.3.1 - b**. O espectro de absorção das Hbs é caracterizado pela presença de uma banda intensa entre 400nm e 450nm, conhecida como banda de *Soret*, e outras duas bandas menos intensas, cujos máximos estão em 540nm e 580nm, denominadas bandas Q. A entrada e saída de ligantes no sexto sítio de coordenação resultam em bandas Q características de cada ligante coordenado. Na Hb humana (e de outros vertebrados), com a saída do oxigênio e deslocamento do átomo de ferro para fora do plano do anel, a histidina proximal e a hélice F, deslocam-se, alterando a orientação desta hélice. Esse efeito é peculiar das hemoglobinas, uma vez que o mesmo, não é observado na Mb humana²⁵. Detalhes do deslocamento da hélice na Hb humana, podem ser observados na **figura 1.3.2**. As mudanças estruturais devido à entrada e saída do oxigênio não estão restritas apenas a alterações locais. A transição da forma T para a forma R na Hb humana é acompanhada também de uma rotação de aproximadamente 15° entre os dímeros *alfa-beta* no complexo. A posição relativa entre as cadeias alfa e beta no dímero praticamente não é alterada, salvo as alterações locais do deslocamento efetuado pela hélice F. A **figura 1.3.2 - d** mostra a rotação relativa entre os dímeros *alfa-beta* na transição da forma T para forma R, como foi determinado por Perutz nos trabalhos pioneiras sobre estruturas de hemoglobinas^{5,6}.

Figura 1.3.2 - Conformação dos grupos hemes na forma *oxi*(a) a *deoxi*(b). Note que existe um deslocamento do íon ferroso para fora do plano do anel com a saída do oxigênio (transição d forma *oxi* para *deoxi*). Na hemoglobina humana, as perturbações estruturais desses processo estendem-se para a estrutura terciária, com a reorientação da hélice F (c), até ao arranjo quaternário (d) resultando em uma rotação relativa entre os dímeros *alfa-beta*, como foi determinado por Perutz nos trabalhos pioneiras sobre estruturas de hemoglobinas.

1.4 Hemoglobinas extracelulares de invertebrados.

Hemoglobinas de invertebrados por sua vez, variam de estruturas monoméricas até complexos gigantes com múltiplas subunidades, ocorrendo em uma diversificada variedade de meios biológicos, como por exemplo: citoplasma de tecidos específicos (músculo, nervos, gametas e etc.), células vermelhas ou livremente dissolvidas nos fluidos biológicos. Quando comparadas com as estruturas tetraméricas assumidas pelas hemoglobinas dos vertebrados, as Hbs dos invertebrados mostram uma grande variação de estruturas quartenárias e possuem uma alta variabilidade, refletindo a especialização e adaptação para uma maior gama de ambientes²⁶. Porém, pouco se sabe da relação entre a função biológica e suas estruturas moleculares ao nível atômico. As Hbs de invertebrados podem ser divididas em cinco grupos, de acordo com suas estruturas quartenárias:

1. Monoméricas, com um único domínio de massa ~17 kDa. Podem ser intracelulares ou extracelulares e são as que formam o maior grupo de Hbs de invertebrados.
2. O segundo grupo, contem Hbs diméricas de bactérias²⁷, dímeros e tetrâmeros de Hbs intracelulares, como as de ostras (*Scapharca*²⁸), e complexos de uma única cadeia, como a Hb polimérica intracelular de *Glycera*²⁹.
3. O terceiro grupo abrange complexos de múltiplas subunidades com massas que vão de 200 até 800 kDa, compostas de subunidades de globinas com dois domínios (~35 kDa), encontrada em artrópodes³⁰ e nematóides³¹.
4. O quarto grupo, contem Hbs com múltiplas subunidades, múltiplos domínios, constituídos de uma ou mais cadeias de 4 – 20 domínios de globinas ligados de forma covalente. Inclusos no quarto grupo, estão a Hb de camarão, com 250-kDa³², a Hb de lesma³³ com 1.700 a 2.300 – kDa, a Hb polimérica encontrada em ostras, com mais de 8.000 kDa³⁴, e a Hb de 124 e 153-kDa encontrada nos poliquetas *Branchiopolynoe* de hidrotermais³⁵.
5. O quinto grupo inclui as eritrocruorinas ou hemoglobinas de bicamadas hexagonais (*Hexagonal Bilayer Hemoglobins* - HBL Hbs) e clorocruorinas (muito similares as eritrocruorinas, porém possuem um grupo heme modificado). Essas estruturas apresentam massa em torno de 3.000 até 4.000 kDa, e são encontradas em anelídeos (incluindo vestimentíferos). São constituídas de dois tipos de subunidades básicas, globinas e *linkers*, que ao

todo somam entre 180 e 192 subunidades.

As Hbs extracelulares são invariavelmente sintetizadas no meio intracelular e então secretadas. Elas ocorrem em solução na hemolinfa de quatro filos de invertebrados (Annelida incluindo vestimentíferos e pogonóforos, Nematoda, Artropoda e Mollusca)^{36,37}. Devido à ausência de um micro ambiente intracelular, onde os níveis efetores podem ser regulados, como nas hemácias dos vertebrados, as Hbs extracelulares por sua vez, operam em condições que são mais dependentes de variações do ambiente do que as Hbs enclausuradas. A ocorrência de Hbs extracelulares é caracterizada por sistemas moleculares grandes que evitam serem excretados através de membranas. Uma exceção a essa regra é a Hb das larvas de mosquitos da família Chironomidae (libélulas) que está presente nas formas monomérica/dimérica. Essas estruturas pequenas são permitidas devido à presença de tubos de Malpighi, que desempenham a função de órgãos excretores ao invés de filtradores/coletores, como no caso de nefrídios/rins, encontrados em outros invertebrados e que poderiam resultar na perda de globinas livremente dissolvidas. Os anelídeos apresentam órgãos filtradores como nefrídios ou celomadutos (protonefrídios) o que está em acordo com a presença de complexos extracelulares gigantes, carregadores de oxigênio, como no caso das eritrocruorinas e clorocruorinas^{38,39}.

1.5 Hemoglobinas gigantes em anelídeos.

As eritrocruorinas e clorocruorinas presentes em anelídeos formam um único grupo de acordo com suas propriedades físico-químicas⁴⁰. Essas estruturas são formadas por dois tipos diferentes de cadeias de polipeptídeos, globinas e *linkers*⁴¹, que se associam formando uma estrutura quaternária característica em forma de bicamada hexagonal. Nos trabalhos primordiais, sobre evolução molecular de *hexagonal bilayer haemoglobins* (HBL Hbs) extracelulares em anelídeos, Gotoh et al⁴², dividiu-se as cadeias de globinas em dois grupos com uma origem em comum. Sucessivamente, Suzuki e Riggs⁴³ demonstraram que as cadeias de *linkers* apresentam um domínio homólogo ao receptor de lipoproteína de baixa densidade (*LDL receptor*). Dessa forma, os *linkers* não poderiam partilhar das mesmas funções que as globinas, uma vez que essas duas estruturas apresentam grande divergência evolutiva. Recentemente, Yuasa e colaboradores⁴⁴, propuseram um modelo geral de evolução para essas

Hbs extracelulares gigantes. De acordo com esse autor, o ancestral comum de todos esses complexos, era uma proteína formada apenas por cadeias de globinas, e então, durante o processo evolutivo em anelídeos, as cadeias de *linkers* foram adicionadas formando a estrutura final das Hbs extracelulares gigantes (**Fig. 1.5.1 - a**, parte superior). É importante mencionar, que os membros do grupo Siboglinidae que abrangem vestimentíferos e pogonóforos (vermes marinhos evolutivamente próximos aos poliquetas) apresentam variações em suas Hbs extracelulares. Vestimentíferos apresentam duas formas de Hbs extracelulares. A primeira, denominada pesada (Hb_P) com massa entre 3,0 e 4,0 MDa, é em realidade uma eritrocruorina, compostas por globinas e *linkers*. A segunda forma é chamada de leve (Hb_L), possui massa de aproximadamente 400kDa e é composta apenas por subunidades de globina. A forma Hb_L pode ser subdividida em duas de acordo com o meio extracelular onde se encontra. Uma coexiste com a forma pesada no meio vascular (Hb_{LV}), a outra é exclusiva de cavidades celômicas (Hb_{LC}). Pogonóforos por sua vez possuem apenas a forma Hb_L. Com base no sequenciamento das Hbs de vestimentíferos e pogonóforos, e com a classificação dessas espécies com um subgrupo dos poliquetas, Negrisolo e colaboradores⁴⁵, sugeriram uma origem evolutiva alternativa para as hemoglobinas extracelulares e clorocruorinas. Segundo esse autor, o ancestral comum de anelídeos, vestimentíferos e pogonóforos, conteria em seu genoma todo o conjunto de genes que codificam cadeias de globinas e *linkers*, constituindo uma HBL Hb primitiva. Dessa forma as Hbs_L, eritrocruorinas e clorocruorinas não haveriam coevoluído independentemente. Posteriormente o ancestral comum a vestimentíferos e pogonóforos, evoluiu para as Hbs_L, e por fim, pogonóforos perderam a Hb_P ao longo do processo evolutivo. Um esquema simplificado das duas teorias sobre a origem das hemoglobinas extracelulares em anelídeos pode ser observado na **figura 1.5.1 – a**. Recentemente, foram reportadas estruturas cristalográficas de Hbs_L de pogonóforos e vestimentíferos, *Oligobranchia mashikoi*^{46,47} (código PDB 2D2M, 2ZS0) e *R. pachyptila*⁴⁸ (código PDB 1YHU) respectivamente. Em ambos os casos, a composição básica compreende quatro tipos diferentes de globinas (A1, A2, B1, e B2). Essas subunidades formam uma estrutura dodecamérica em forma de cúpula. Dois dodecâmeros juntos formam um arranjo esférico contendo 24 subunidades de globinas com uma cavidade interior, como pode ser observado na **figura 1.5.1 - b**.

Figura 1.5.1 - (a) As duas teorias sobre a origem das hemoglobinas extracelulares em anelídeos. (b) Estrutura da Hb_{LS} de *Oligobranchia mashikoi*. O complexo é formado por 24 subunidade de globinas.

Os primeiros estudos, utilizando microscopia eletrônica, para as Hbs extracelulares de anelídeos, revelaram que tais Hbs são constituídas de duas camadas hexagonais superimpostas, com altura de ~20 nm e diâmetro de ~30nm em torno de uma cavidade central^{49,50}. Vários estudos também foram realizados utilizando-se espalhamento de raios-X a baixo ângulo para investigar o formato assumido por esses complexos⁵¹⁻⁵⁵. Quando a estrutura da eritrocruorina de *Lumbricus terrestris* (HbLt) foi caracterizada inicialmente, esta revelou um eixo local de simetria de ordem 3 para cada uma das doze unidades (protômeros) repetitivas que formavam o complexo hexagonal. Os doze protômeros da estrutura estão arranjados em duas camadas hexagonais sobrepostas, com uma rotação de 16° entre elas⁵⁶. Essa conformação chamada de **não eclipsada**, constitui o grupo das eritrocruorinas do tipo I. Uma estrutura quaternária muito semelhante foi encontrada para as Hbs gigantes de outros anelídeos; *Macrobdella decora* (um tipo de sangue suga), *Riftia pachyptila* (vestimentífero), *Alvinella pompejana* (poliqueta), e a clorocruorina de *Eudistylia vancouverii* (poliqueta)⁵⁷⁻⁶⁰. Porém, a hemoglobina de *Alvinella pompejana*, não apresenta a rotação entre as duas camadas hexagonais. Essa constatação se estendeu para varias outras estruturas de hemoglobinas de poliquetas, incluindo a da espécie *Arenicola marina* (HbAm)⁶¹, sugerindo que poliquetas,

incluindo os que habitam o oceano profundo, mas excluindo as quatro famílias que contêm clorocruorinas, apresentam estruturas sem a rotação entre as camadas. Essa conformação do tipo **eclipsada** constitui o grupo das eritrocruorinas do tipo II. Sendo assim todos os outros anelídeos (incluindo vestimentíferos) apresentam rotação relativa entre as camadas hexagonais. A **figura 1.5.2** mostra uma comparação dos modelos gerados por microscopia eletrônica das eritrocruorinas do tipo I e II.

Figura 1.5.2 - Uma comparação dos modelos gerados por microscopia eletrônica das eritrocruorinas do tipo I e II. (a) *Arenicola marina* (tipo II, eclipsada) e (b) *Lumbricus terrestris* (tipo I, não eclipsada)⁶².

Eritrocruorinas apresentam o fenômeno de cooperatividade com coeficientes de Hill da ordem de 7,0. Porém, a espectroscopia Raman de ressonância foi utilizada pra caracterizar a HbLt, mostrando que, em contraste com as com as Hbs de vertebrados, a ligação do O₂ não resulta em alterações espectrais envolvendo o modo de estiramento do ferro-His proximal⁶³. Este comportamento, análogo para a Mb, sugere que o mecanismo de cooperatividade nas eritrocruorinas é aparentemente diferente dos mecanismos do cooperatividade dos vertebrados^{64,65}. Dessa forma, estudos de espalhamento de raios X de baixo ângulo para as

formas *oxi*-Hb e *deoxi*-Hb das eritrocruorinas e também para a porção dodecamérica isolada (fração contendo apenas as subunidades de globinas, muito semelhante às Hbs leves de vestimentíferos e pogonóforos), mostram que não há alterações observáveis nos formatos dessas espécies com a entrada e saída do oxigênio^{66,67}. Esse resultado sugere que não existem alterações na estrutura quaternária maiores do que ~1%. Em contraste, a Hb de mamíferos onde o coeficiente de Hill é tipicamente constante frente aos níveis de saturação de oxigênio entre 10% e 90% em uma ampla variação de valores de pH, as eritrocruorinas variam grandemente de espécie pra espécie^{39,68-70}, e intrinsecamente, com a saturação de O₂, pH, concentração de cátions inorgânicos, e temperatura⁷¹⁻⁷⁵. Cooperatividade é comumente máxima em pH fisiológico, assim como para HbAm, onde o pico de cooperatividade é observado entre pH 7,2 e 7,6 (em 20°C) e para a *Alvinella caudata* e *Apompei*, onde o pico é em torno de pH 6,9, em concordância com o pH mais baixo apresentados pela hemolinfa desses organismos^{75,76}. No entanto a maior divergência entre pH fisiológico e o pH de máxima cooperatividade é encontrado para a HbLt, onde o coeficiente de Hill aumenta de 3 em pH baixo para um máximo de 9,5, perto de pH 7,8⁶⁶. Da mesma forma, a cooperatividade de Hbs dos oligoquetas *Pheretima hilgendorfi*⁷³ e *Eisenia foetida*⁷¹, e dos sangues sugas *Hirudo medicinalis*⁷⁷ e *Macrobdella decora*^{78,79}, possuem um máximo em pH acima de 8,0. A temperatura também pode ser um fator muito importante na cooperatividade. Na HBL Hb de *A.pompejana*, o coeficiente de Hill em pH 7,25 cai de 3 em 20°C, para 1,2–1,5 em 10 e 40°C⁷⁴. Isso indica pronunciadas diferenças no calor total de oxigenação das formas desoxigenada e oxigenada das eritrocruorinas. A cooperatividade das eritrocruorinas em sangue fresco aparece ser maior do que purificada pelo menos nos casos de *Octolasion complanatum*⁸⁰ e *L.terrestris*⁶⁶. A razão pela qual isso ocorre é a perda de cátions durante o processo de purificação. Diferente das hemoglobinas dos vertebrados, que são sensíveis a presença de compostos orgânicos aniônicos (2,3-DPG e ATP), as eritrocruorinas são sensíveis a presença de cátions, que aumentam a afinidade desses complexos por oxigênio. De uma forma geral, cátions divalente são mais efetivos ($Ba^{2+} > Ca^{2+} > Sr^{2+} > Mg^{2+} > Li^{+} > Na^{+} > K^{+}$), de forma que não depende apenas da carga do íon, mas também do raio iônico de cada espécie⁸¹. Isso está em acordo com os efeitos mais pronunciados do Ca^{2+} sobre o Mg^{2+} nas hemoglobinas de *Amphitrite ornata*⁸², *Eisenia foetida*^{71,73}, *Pheretima hilgendorfi*⁸³, e *Hirudo medicinalis*⁷⁷. Assim como o maior efeito do Ba^{2+} sobre os outros cátions divalentes na Hb de *Macrobdella decora*⁷⁸. No entanto algumas Hbs mostram-se insensíveis à presença de Ca^{2+} e Mg^{2+} , como no caso das Hbs de *Tubifex* e *Marphysa sanguinea*. As Hbs dos oligoquetas *Lumbricus terrestris*, *Tubifex tubifex*, e *Tylorrhynchus*, também apresentam cátions de Zn^{2+} e

Cu^{2+} como parte da sua estrutura^{84,85}. Na Hb de *Pheretima hilgendorf*, onde o sangue contém 0,9 mM Zn^{2+} , (o que significa entre 1 ou 2 átomos de Zn para cada 165 átomos de ferro) e não podem ser removidos por diálise. A adição de Zn^{2+} , aumenta a afinidade pelo O_2 , devido a alterações na constante de associação com o estado T (K_T)⁸⁶. Porém, esse fato aparentemente não altera a constante de associação de O_2 com o estado R (K_R), sugerindo um mecanismo de ação diferente dos outros cátions divalentes mencionados anteriormente.

1.6 A hemoglobina gigante de *Lumbricus terrestris*.

Com a resolução da estrutura cristalográfica da hemoglobina hexagonal gigante de *Lterrestris* (HbLt)^{87,88}, Royer e colaboradores, deram grande contribuição no esclarecimento sobre a estequiometria referente às subunidades constituintes, e também, sobre os níveis hierárquicos de empacotamento desse complexo. Na estrutura existem quatro tipos diferentes de globinas, denominadas *a*, *b*, *c* e *d*, que formam um heterotetrâmero *abcd*. O heterotetrâmero é repetido então três vezes, formando uma estrutura com simetria C_3 , denominada dodecâmero ou *cap*, com a seguinte estequiometria $([abc]_3[d]_3)$, onde $a_1b_1c_2$ (onde c_2 é a subunidade tipo *c* do heterotetrâmero adjacente), são unidos por ligações de dissulfeto (embora a presença das ligações de dissulfeto não se estende a todos os anelídeos). Especificamente a_1 forma duas pontes dissulfetos com b_1 e c_2 respectivamente mantendo esta unidades trimérica covalentemente ligada. Essa estrutura em forma de cúpula, com aproximadamente 200 kDa de massa, é muito semelhante as estruturas na forma de cúpula que formam a então denominada Hb “leve” de vestimentíferos e pogonóforos⁴⁶⁻⁴⁸. Em 2004 Strand e colaboradores resolveram a estrutura do dodecâmero da HbLt isoladamente⁸⁹, mostrando que este apresenta grande concordância estrutural com as estruturas de hemoglobinas leves de poliquetas resolvidas por difração de raios X. Porém na HbLt, assim como nas outras eritrocruorinas, o dodecâmero aparece associado a um héterotrímero de *linkers* denominados (L1, L2 e L3). Essa estrutura com estequiometria $([abc]_3[d]_3)(L1L2L3)$ forma o então mencionado protômero. O protômero é a unidade básica que se repete formando a bicamada hexagonal (unidade biologicamente ativa). Cada disco da bicamada é formado por seis protômeros, somando um total de doze (12) protômeros, contendo ao todo: 144 globinas e 36 *linkers*. A **figura 1.6.1** mostra os níveis de hierarquia adotados pela

estrutura da HbLt. Apesar da resolução moderada de 3,5Å, a estrutura da HbLt também revelou detalhes sobre o enovelamento das unidades constituintes. As cadeias de globinas apresentam um enovelamento similar ao das subunidades das globinas *alfa/beta* nas Hbs dos vertebrados, porém nenhuma das globinas da HbLt possui a hélice D. Detalhes como ligações de dissulfeto intra e inter subunidades foram observados, confirmando a presença e os locais de ligações entre as cadeias to tipo *a*, *b* e *c*. Grupos prostéticos como, cátions metálicos e a presença dos grupos heme também foram determinados.

Figura 1.6.1 - Classificação dos diferentes níveis de hierarquia das HBL Hbs/Eritrocrutorinas. Unidade da HbLt (a) e (b), com simetria D6 (ou 622), contendo 12 protômeros. Ao todo são 144 cadeias de globinas e 36 cadeias tipo linkers. O protômero (c) por sua vez é composto por 4 diferentes tipos de globinas que primeiramente formam um heterotetrâmero *abcd*, e em seguida são associadas em uma estrutura dodecamérica (d) - cap ([abc]₃[d]₃). O cap por sua vez, está associado a um heterotrímero de *linkers*.

As globinas que formam o heterotetrâmero ([*abc*][*d*]), estão primeiramente associadas como heterodímeros, *a/d* e *b/c* e estes dímeros interagem entre si, através das hélices E e F. Nessa interface, resíduos presentes na hélice F, interagem diretamente com o grupo heme da subunidade complementar, formando uma ponte salina através dos grupos propionato que o

mesmo possui. Um arranjo quaternário semelhante é observado para a hemoglobina do molusco *Scapharca inaequivalvis*⁹⁰. A estrutura da HbLt possui três distintas subunidades de *linkers* (L1, L2 e L3), que são necessárias para a formação da bicamada hexagonal⁹¹⁻⁹³ onde cada uma, ocupa 1/15 da unidade protomérica. A densidade eletrônica obtida para a estrutura da HbLt, não é ambígua para duas das cadeias de *linkers*, L1 e L2, porém a terceira posição, aparentemente, pode ser ocupada pela cadeia L3 ou L4 (onde L4 é uma isoforma da cadeia L3⁹⁴). No modelo cristalográfico apenas a cadeia L3 foi modelada, uma vez que foi constatado experimentalmente que esta é a isoforma mais abundante na HbLt⁹⁵. Os três *linkers* são estruturalmente semelhantes entre si, e segundo Royer e colaboradores, podem ser divididos em quatro domínios diferentes (**figura 1.6.2**). O primeiro domínio, na porção N-terminal, é uma hélice- α contendo entre 21-30 resíduos. Esta estrutura em hélice interage com as outras homólogas dos dois *linkers* vizinhos formando uma estrutura do tipo *coiled coil* trimérico.

Figura 1.6.2 - Visão geral da estrutura dos *linkers*. A cadeia é constituída de 4 domínios, (1) Hélice 1, (2) hélice 2, (3) LDL-A e (4) β -barril. O domínio LDL-A possui um sítio de ligação de Ca^{2+} , assim como para o receptor de LDL humano, onde o Ca^{2+} é hexacoordenado.

A estrutura segue com um arranjo de 1-7 resíduos não helicoidais, que conectam com o segundo domínio. Esse por sua vez é uma pequena hélice contendo 12–17 resíduos que participam de um curto *coiled coil* também trimérico, dentro da “cabeça” do heterotrímero de *linkers*. O terceiro domínio é rico em cisteínas, homólogo ao receptor de lipoproteína de baixa densidade (LDL-A: *Low-density lipoprotein receptor domain class A*). Este apresenta um sítio de ligação de Ca^{2+} de alta afinidade, estimado em ~ 70 nM, o que explica parte da demanda por cátions divalente dessa natureza para a estabilização da estrutura. Por fim, o quarto e último domínio, é uma estrutura do tipo β -barril, formado por 8 fitas antiparalelas, que desempenha papel importante na interação com a porção dodecamérica de globinas. Os *linkers* desempenham um papel estrutural importante no arranjo quaternário das eritrocruínas, não apenas na formação do protômero, mas também realizando os principais contatos entre protômeros na unidade biológica. A HbLt apresenta atividade superóxido dismutase⁹⁶, o que está em concordância com a presença dos metais Zn e Cu na composição química determinada experimentalmente. Uma vez que a atividade de superóxido dismutase (SOD) não foi constatada para as cadeias de globinas na HbLt, essa atividade fica então, restrita para a porção dos *linkers*. No entanto, essa Hb apresenta apenas $\sim 10\%$ da atividade superóxido dismutase de CuZnSODs⁹⁷. Dado que existem quatro *linkers* diferentes e nenhum contribui com mais do que 10% da massa total da partícula, a atividade de cada *linker* isoladamente seria comparável à enzima de mamíferos. Para o modelo cristalográfico da HbLt, apenas um único átomo de Zn foi modelado na estrutura, e encontra-se em um ambiente químico que não sugere nenhuma semelhança com os sítios de Zn nas CuZnSODs⁹⁷. Não foram encontrados sítios/densidades eletrônicas, referentes a Cu na estrutura da HbLt, mas vale lembrar que essa é uma estrutura cristalográfica de baixa resolução, onde esse tipo de interpretação é bastante limitada. Sendo assim a presença desses íons metálicos e seus sítios de coordenação, ainda são uma questão em aberto para essas estruturas.

1.7 A hemoglobina gigante de *Glossoscolex paulistus* e outras espécies de interesse.

O *Glossoscolex paulistus* (figura 1.7.1) é um anelídeo de vida livre, endêmico da região centro-sul do estado de São Paulo, facilmente encontrado nos arredores das cidades de Rio Claro e Piracicaba. Os trabalhos pioneiros de Righi⁹⁸, classificam essa peculiar espécie dentro da classe Oligochaeta, ordem Haplotaxida e família Glossoscolecidae. Membros dessa família são encontrados do México até a Argentina, sendo que as espécies nativas do Brasil são em geral, endêmicas de regiões específicas, o que explica a preferência do *G.paulistus* por solos ricos em calcário. Assim como outros membros dessa família, o *G.paulistus* destaca-se pelo tamanho diferenciado, pois este apresenta um comprimento que varia de 280 a 350 mm e o maior diâmetro, na região mediana do corpo, alcança de 10 a 11 mm. O *G.paulistus* é uma espécie geófaga, alimentando-se de matéria orgânica do solo sem estrutura celular reconhecível. A Hb de *G.paulistus* (HbGp), é uma eritrocruorina cuja massa foi inicialmente determinada em 3,1 MDa⁹⁹. Recentemente uma cuidadosa investigação foi feita através da técnica de ultracentrifugação analítica, chegando a um novo valor, estimado entre 3,6 e 3,7 MDa¹⁰⁰. Esses valores são compatíveis com os obtidos para a HbLt. As semelhanças com a HbLt estendem-se em relação as propriedades físico-químicas e a estequiometria. Tanto a HbLt quanto a HbGp, são dissociadas em ambientes químicos alcalinos, pH >9,0, resultando na liberação da cadeia *d* (que não está ligada de forma covalente com outras subunidades)^{101,102}. Em 1996 Agostinho¹⁰³ e colaboradores mostraram que esse efeito de dissociação frente ao ambiente químico alcalino é mais pronunciado para a forma oxidada da HbGp. A possibilidade de isolamento da subunidade *d* foi também a razão pela qual esta é a porção mais estudada e caracterizada da HbGp. Dentre as subunidades que constituem a HbGp (número estimado entre 7 a 9 tipos de cadeias), apenas a subunidade *d* foi completamente sequenciada¹⁰⁴, e uma comparação com sua ortóloga na HbLt, mostra uma identidade sequencial de 57%. Esses dados sugerem que a HbGp apresenta uma hierarquia de empacotamento similar ao da HbLt. Tal hipótese também é sustentada pelos dados obtidos através de espalhamento de raios-X a baixo ângulo¹⁰⁵ assim como por análises de espectrometria de massas¹⁰⁶. O comportamento da HbGp também é sensível à presença de cátions metálicos divalentes. Os dados apontam que tanto as espécie Ca^{2+} quanto Mg^{2+} são capazes de exercer os mesmos efeitos sobre a afinidade da HbGp pelo O_2 . Porém, é sob a influencia da espécie Mn^{2+} que a HbGp apresentaria maior afinidade pelo O_2 ¹⁰⁷.

Figura 1.7.1 - Foto de um dos indivíduos coletados durante o período de 2010 e 2011. O *Glossosocolex paulistus* se destaca pelo seu tamanho incomum.

O interesse no comportamento físico-químico desses sistemas não se estende apenas a função de oxigenação que os mesmos apresentam, mas também sobre suas possíveis funções secundárias, como atividade de superóxido dismutase e/ou como carregadores de moléculas pequenas. A ampla caracterização das propriedades físico-químicas da HbGp, coloca esse sistema como uma das eritrocruorinas mais estudadas da literatura, ao lado da HbLt de *L. terrestris*.

O *Rhinodrilus alatus* é uma espécie geófaga, membro da família Glossoscolecidae, endêmica da região das Sete Lagoas/Paraopeba no estado de Minas Gerais⁹⁸. Esta espécie foi considerada como sendo ameaçada de extinção devido a seu uso extensivo como fertilizantes naturais e também como iscas para pesca. Uma caracterização detalhada da anatomia desses animais foi feita especialmente sobre seu sistema circulatório. O *R. alatus* é uma minhoca de tamanho pronunciado que atinge entre 56 e 63 cm de comprimento, e 11 ou 12 mm de largura¹⁰⁸. Nos períodos de seca, entre Abril e Setembro, os indivíduos dessa espécie entram em um estágio de dormência, em uma câmara protegida por muco. O *R. alatus* pertence à mesma família do *G. paulistus*, porém diferente da HbGp, que vem sendo extensivamente caracterizada, as propriedades da Hb de *R. alatus* (HbRa) ainda são totalmente desconhecidas.

O *Perionyx excavatus* é uma espécie detritívora, membro da família Mesgascolecidae e gênero *Perionyx*. Este verme comumente encontrado em uma vasta área da região tropical da Ásia, e devido ao seu interesse econômico foi levada pra a Europa e América do Norte¹⁰⁹. O *P. excavatus* é uma espécie que vive/alimenta-se de restos orgânicos com alta umidade. Dessa forma tal espécie requer quantidades adequadas de matéria orgânica disponível para que uma população se estabeleça e então processe a matéria orgânica de forma eficiente. A capacidade de processar matéria orgânica em adubo para o crescimento de plantas, torna esta

espécie uma alternativa para aplicações tecnológicas (produção de humos). Embora parte de seu RNA mensageiro tenha sido sequenciado¹¹⁰, e portanto são conhecidas diversas proteínas expressas por esse organismo, pouco é conhecido sobre a Hb desta espécie (HbPe).

O verme conhecido como gigante africano, *Eudrilus eugeniae*, é uma espécie detritívora, membro da ordem dos Haplotaxida e família Eudrilidae. É originário do oeste da África tropical. Embora seu nome popular sugira um verme de proporções excessivas, os indivíduos dessa espécie atingem um comprimento médio entre 15 e 20 cm, e um diâmetro entre 4 e 8 mm. Esta espécie apresenta interesse comercial devido à suas características alimentares^{111,112}. Entretanto outras aplicações biotecnológicas tem sido reportadas, como por exemplo modelo para estudos de efeitos mutagênicos causados por raios X, metais pesados e outros contaminantes¹¹³.

A espécie *Eisenia andrei*, é uma espécie detritívora, membro da ordem dos Haplotaxida e família Lumbricidae. Essa espécie é conhecida como vermelha da Califórnia, nome popular devido ao fato desta minhoca de cor avermelhada ter sido inicialmente produzida de forma sistemática nos Estados Unidos da América, mas sua origem encontra-se no leste europeu. Este pequeno verme tem entre 5 e 10 cm de comprimento e entre 4 e 6 mm de diâmetro e é a mais bem caracterizada espécie na aplicação de decomposição de material orgânico¹¹⁴. A fácil adaptação e reprodução dessa espécie justifica seu amplo uso em diversas atividades no mundo inteiro. Alguns estudos também apontam esses vermes terrestres como indicadores biológicos de contaminação do solo. Vampré e colaboradores¹¹⁵ demonstraram que a contaminação do solo por hexaclorobenzeno (HCB) pode ser identificado por análise do solo e tecidos extraídos desses vermes. Estes ainda são importantes no gerenciamento dos nutrientes contidos no solo, desempenhando importante papel na estabilidade/filtração do solo e também no ciclo do fósforo e nitrogênio. Esta espécie conta com uma ampla caracterização de seu papel ecológico e descrição anatômica. É uma das espécies de oligoquetas que conta com a maior quantidade de genes transcritos sequenciados, graças aos trabalhos de Cho e colaboradores¹¹⁰. Embora não exista nenhuma caracterização estrutural ou físico-química de sua hemoglobina, todas as suas subunidades possuem sequências disponíveis no banco de dados EST. As espécies *P.excavatus*, *E.eugeniae*, *E.andrei* e *R.alatus*, podem ser vistas na **figura 1.7.2**.

Figura 1.7.2 - (a) *Perionyx excavatus*, também conhecida como violeta do Himalaia, (b) *Eudrilus eugeniae* - gigante africana, (c) *Eisenia andrei* – vermelha da Califórnia, é a que atinge o menor tamanho entre as espécies utilizadas neste trabalho. (d) *Rhinodrilus alatus* – minhocucu, espécie nativa do vale do Paraopéba em Minas Gerais.

1.8 Hemoglobinas extracelulares como potenciais substituintes para sangue.

Em 2011, três diferentes abordagens na tentativa de obtenção de um substituinte para as células vermelhas do sangue chegaram a fase três, onde são realizados testes clínicos. Os métodos envolvendo o uso da Hb humana polimerizada e a Hb bovina polimerizada foram descontinuados, uma vez que os resultados apontaram um aumento na mortalidade e outras complicações. Porém, a utilização da Hb humana em superfície conjugada de polietilenoglicol, ainda se encontra em fase de testes. O maior problema associado a esses tipos de carregadores de oxigênio (instabilidade, estresse oxidativo e controle dos níveis de NO) podem ser atribuídos a retirada da Hb do ambiente intracelular. As hemácias contêm enzimas que previnem o estresse oxidativo, a membrana celular que reduz a interação com NO, os efeitos alostéricos que modulam a liberação de oxigênio¹¹⁶, e uma alta concentração de Hb que minimiza o processo de dissociação do tetrâmero em dímeros. Dessa forma, o uso de Hbs extracelulares de outros organismos podem ser uma atrativa estratégia na resolução desse problema, uma vez que esses carregadores de oxigênio são devidamente adaptados as condições extracelulares. Dentre as Hbs extracelulares conhecidas e caracterizadas, a grande

maioria pertence aos organismos marinhos de regiões profundas, o que dificulta a viabilidade para aplicações tecnológicas/industriais. Dessa forma, existe uma considerável preferência pelo uso de organismos terrestres como fonte dessas Hbs. O oligoqueta *Lterrestis*, assim com sua Hb, é um das mais estudadas espécies de oligoquetas conhecidos. Este organismo também demonstra baixo custo de produção em larga escala (~\$75/1,000 minhocas). Estima-se que em torno de mil minhocas são necessárias para a produção de 5 a 10g de HbLt (mas a eficiência pode variar de acordo com a técnica de purificação empregada). As Hbs de *Amarina* e *Lterrestres*, estão sendo testadas como substituintes para células vermelhas em camundongos. Os resultados preliminares mostram que apesar de não reportar efeitos colaterais, a HbAm dissocia-se rapidamente no plasma dos animais, enquanto os mesmo experimentos conduzidos com a HbLt apontam uma estrutura mais estável, sem a produção de efeitos adversos, assim como também sem a produção de resposta imune com repetidas aplicações¹¹⁷. Assim com a HbLt, a HbGp é uma Hb extracelular amplamente caracterizada, onde a os resultados apontam uma considerável semelhança com as propriedades identificadas na HbLt, o que faz da HbGp um potencial substituinte para as células vermelhas do sangue.

1.9 Objetivos

A HbLt foi a primeira proteína cristalizada que se tem notícia¹¹⁸. No entanto, foi preciso mais de 150 anos para que sua estrutura fosse de fato resolvida⁸⁷. Embora tendo representado um importante avanço para o entendimento do funcionamento das eritrocruorinas, o modelo cristalográfico da HbLt obtido é limitado, não apenas pela resolução (3,5Å), mas também por uma série de outros parâmetros, como por exemplo a baixa completeza na região de mais alta resolução (80%) e a presença de duas unidades biológicas inteiras na unidade assimétrica (7,2 Mda, o que equivale a 360 subunidades). Esses fatores restringiram o refinamento do modelo, obrigando Royer e colaboradores a utilizarem procedimentos de simplificação do modelo (simetria não-cristalográfica), reduzindo drasticamente o número de parâmetros a serem refinados (neste caso, um protômero representa todos os vinte e quatro protômeros da unidade assimétrica). A limitada resolução do modelo deixa em aberto questões de como e quais espécies químicas, como íons ou pequenas moléculas, estariam ligadas a estrutura (embora alguns íons coordenados a estrutura

foram identificados). Vale lembrar que a HbLt apresenta glicosilações determinadas experimentalmente, porém seu modelo cristalográfico não apresentou pistas de onde e como estas ocorrem. Considerando a significativa, porém limitada, contribuição do modelo cristalográfico da HbLt, este trabalho tem o objetivo de:

- Cristalizar e resolver a estrutura de novas eritrocruorinas, na busca por estruturas de mais alta resolução e portanto capazes de enriquecer o conhecimento de como esses sistemas se comportam em nível molecular, principalmente no que diz respeito ao comportamento alostérico das eritrocruorinas. As espécies utilizadas neste trabalho são: *Glossoscolex paulistus*, *Rhinodrilus alatus*, *Eisenia andrei*, *Eudrilus eugeniae* e *Perionyx excavatus*.
- Cristalizar e resolver a estrutura da subunidade *d* da HbGp isolada, que apesar da sua ampla caracterização, em termos de propriedades físico-químicas, ainda carece de informações estruturais.
- Sequenciar as subunidades constituintes da HbGp e outras Hbs de interesse, (utilizando a plataforma Solexa – Illumina), assim permitindo a obtenção de modelos cristalográficos devidamente completos.

2 METODOLOGIA

2.1 Coleta dos animais utilizados.

Os indivíduos da espécie *Glossoscolex paulistus* e *Rhinodrilus alatus* foram coletados em campo. Os exemplares coletados são mantidos em uma porção de terra da qual são extraídos e então transportados ao laboratório onde são posteriormente sacrificados. A espécie *Gpaulistus* é obtida nos arredores do distrito de Assistência, próximo a cidade de Rio Claro, estado de São Paulo. Os indivíduos da espécie *Rhinodrilus alatus* foram coletados na região do vale do Paraopeba, no norte do estado de Minas Gerais. As espécies *Eisenia andrei*, *Eudrilus eugeniae* e *Perionyx excavatus* foram doadas pela empresa Minhobox, que comercializa essas espécies para minhocultura e produção de humos.

2.1.1 Purificação das oxi-Hbs íntegras e a subunidade α da HbGp.

Os exemplares obtidos de cada espécie foram lavados em água corrente e realizado um corte no dorso de cada indivíduo. A hemolinfa foi coletada em um becker de 100 ml contendo 5 ml de tampão TRIS-HCl 100 mmol L⁻¹, pH 7,0, com a ajuda de pipetas e armazenada em tubos Falcon de 50mL. Os tubos foram mantidos em gelo durante toda a coleta. As amostras foram centrifugadas a 4000g por 10 minutos a 4° para separação de qualquer particulado sólido. O sobrenadante foi então concentrado até um volume de 2mL, utilizando a tecnologia Amicon - Millipore, com membrana de 50 kDa de corte de massa. A etapa final de purificação foi feita por cromatografia de exclusão molecular em sistema de purificação AKTA (AKTA Basic; General Electrics), utilizando-se a coluna preparativa Superdex 200 16/60 (GE Healthcare) previamente lavada e equilibrada com tampão TRIS-HCl 100 mmol/L, pH 7,0. O eluato coletado no volume de exclusão contendo oxi-Hb, foi novamente concentrado até uma concentração de 20 mg/mL, e mantida a 4° para uso posterior¹⁰³. As amostras das diferentes hemoglobinas foram submetidas à eletroforese desnaturante, em gel de poliacrilamida a 15%

(SDS-PAGE), confirmando purificação. Para o isolamento da subunidade *d* da HbGp (*d*-HbGp), uma amostra íntegra da *oxi-Hb* foi colocada em tampão TRIS-HCl 100 mmol/L, pH 9,0 por um período de 24 horas. A amostra foi então submetida a uma cromatografia de exclusão molecular em sistema de purificação AKTA, utilizando-se a coluna preparativa Superdex 200 16/60 previamente lavada e equilibrada com tampão TRIS-HCl 100 mmol/L, pH 9,0. Após coletadas, a *d*-HbGp foi recolocada em pH 7,0, com a adição do tampão, TRIS-HCl 100 mmol L⁻¹, pH 7,0 em excesso. Por fim a amostra foi concentrada a uma concentração final de 5 mg L⁻¹ e armazenada a 4 °C para posteriores teses de cristalização.

2.1.2 Obtenção das formas meta- e ciano-meta da HbGp.

A forma meta-HbGp foi obtida pela adição de ferricianeto de potássio a uma porção da *oxi-Hb* na proporção estequiométrica calculada de 1 mol de heme de (HbGp) para 4 mols de ferricianeto de potássio (K₃Fe(CN)₆). seguido de incubação por 45 minutos. Em seguida, foi realizada diálise contra tampão Tris-HCl 100 mmol/L pH 7,0 para remover o excedente de íons ferro e ferricianeto do meio. A obtenção da forma ciano-HbGp, é feita através da adição de cianeto de potássio (KCN) a um alíquota de meta-HbGp, na mesma proporção utilizada para o (K₃Fe(CN)₆) anteriormente, seguido de incubação por um período de 45 minutos e diálise em tampão Tris-HCl 100 mmol/L pH 7,0 por 2h, para remoção de excesso de íons ferro- e ferricianeto do meio. As mudanças no estado de oxidação e troca dos ligantes, foram monitoradas por espectroscopia de UV-vis, através de mudanças espectrais nas bandas Q, e bandas de Soret. As concentrações da proteína foram determinadas pelo método de Bradford (Bio-Rad) utilizando-se uma curva padrão de albumina bovina¹¹⁹.

2.1.3 Análise de plasma – composição das amostras.

A técnica de ICP-AES (*Inductively Coupled Plasma -Atomic Emission Spectrometry*) foi utilizada para a determinação da composição elementar das amostras de Hbs purificadas. O equipamento utilizado foi o ICP-AES-Simultâneo CCD-VISTA-MPX (Varian, Mulgrave, Austrália). Nesse procedimento, uma alíquota das amostras de cada Hb em questão foi

digerida em ácido nítrico HNO_3 (60%) e então submetida para análises

2.2 Sequenciamento

2.2.1 Extrações do RNA total do *G.Paulistus*.

A extração do RNA total do *G.Paulistus* (indivíduos coletados na região de Assitência – SP em agosto de 2011) foi realizada utilizando-se o *kit* RNAeasy (*Qiagen*), seguindo as instruções fornecidas pelo fabricante. A concentração e qualidade do RNA obtido foram verificadas por espectrofotometria a 260 nm e em gel de agarose 1% corrido em condições desnaturantes. Uma amostra de dez microgramas de RNA total foi enviado por correio (armazenadas em gel seco) ao laboratório de *High Throughput Sequencing Facility* (HTSF) da Universidade de Carolina do Norte em Chapel Hill sob a direção do Prof. Dr. Piotr Mieczkowski, para ser sequenciada utilizando-se a técnica *Pair End* x 100 pares de bases na versão 3.0 do Hiseq2000 (Illumina/Solexa). Um total de quatro meses foram gastos entre a preparação da amostra, envio para os EUA e devolução das sequências.

2.2.2 Sequenciamento Genômico no Illumina Genome Analyzer

Consta aqui uma visão geral do procedimento realizado em Chapel Hill. O sequenciamento realizado neste projeto envolveu somente os mRNAs. Desta maneira, separou-se o mRNA do RNA total por meio de um *kit* de purificação específico (Dynabeads mRNA *purification kit* - Invitrogen) o qual foi então fragmentado na presença de cátions divalentes e temperaturas elevadas. Após o isolamento do mRNA, realizou-se a síntese de cDNA empregando-se *primers* randômicos pelo *kit Superscript II Reverse Transcriptase* (Invitrogen). A segunda fita foi sintetizada utilizando-se o DNA polimerase I e Rnase H e as enzimas T4 DNA polimerase e *Klenow fragment* foram usadas para se obter extremidades de cDNA abruptas. Ao final, uma adenina foi adicionada às extremidades 3' das moléculas de

cDNA usando a *Klenow* exo-enzima, seguido da adição de adaptadores contendo uma timina na extremidade 5'. Adaptadores que não se ligaram foram removidos por eletroforese e fragmentos entre 200-300 pares de bases foram purificados, correspondendo aos cDNAs ligados aos adaptadores. Finalmente, as bibliotecas foram enriquecidas por 15 ciclos de PCR usando *primers* complementares às sequências dos adaptadores. Avaliação da qualidade e quantificação das bibliotecas foram feitas empregando-se o Bionalyzer Chip DNA 1000 series II (Agilent). Durante o sequenciamento, foram gerados pontes de amplificação, onde as moléculas desnaturadas de fita dupla são aderidas na superfície sólida de uma placa e amplificadas *in situ*, gerando agrupamentos de moléculas de DNA idênticas (*reads*). O sequenciamento da Illumina utiliza quatro nucleotídeos análogos fluorescentes, os quais funcionam como terminadores do sequenciamento. Quando o nucleotídeo correto é incorporado à cadeia crescente, a sua identidade é revelada pela cor de sua fluorescência. O grupo 3-OH do nucleotídeo é então bloqueado para evitar mais incorporação. Após a etapa de imagem, a placa foi lavada e o bloqueio invertido, permitindo assim a adição do nucleotídeo seguinte.

2.2.3 Montagem do transcriptoma e identificação das sequências.

A montagem *de novo* dos *reads* com extremidade pareáveis, foi executada utilizando-se o programa Trinity¹²⁰. Foi escolhido um tamanho do comprimento de inserção máximo de 300 pb (pares de base) e considerando o tamanho mínimo para os *contigs* obtidos de 500 pb. As sequências disponíveis para a HbLt foram utilizadas para identificar as sequências das subunidade homólogas no transcriptoma do *G.paulistus*. A comparação foi feita através do programa BLAST¹²¹, utilizando-se um valor para *e-value* = 1e-5. As sequências selecionadas foram então manualmente inspecionadas. Esse procedimento foi realizado em colaboração com o Dr. Marcelo Falsarella Carazzolle do LNBio.

2.3 Ensaios de cristalização

A cristalização de macromoléculas biológicas é um procedimento empírico uma vez que tal processo dependente de inúmeras variáveis tais como: tempo, pressão, temperatura, homogeneidade e pureza da amostra, ponto isoelétrico, agente precipitante, pH, força iônica, etc. Devido à complexidade de uma proteína e aos inúmeros parâmetros envolvidos, a determinação teórica das condições de cristalização é até hoje imprevisível. Nesse sentido, torna-se necessária uma variação sistemática das condições nesse espaço multiparamétrico visando o estabelecimento de combinações favoráveis de variáveis envolvidas na cristalização. Uma série de técnicas foram desenvolvidas para a cristalização de macromoléculas, destacando-se difusão de vapor em gotas suspensas “*hanging drop*” (**figura 2.3.1**), e gotas sentadas “*sitting drop*”. Independente da técnica empregada, o princípio básico da cristalização envolve uma transição de fase, da proteína em solução para um estado ordenado (cristalino). Em geral, tal transição ocorre uma vez que moléculas de água que solvavam as proteínas são abstraídas pelos agentes precipitantes, induzindo a formação de agregados ordenados. Uma descrição completa das técnicas, bem como nos parâmetros envolvidos na cristalização de macromoléculas pode ser encontrado em Bergfors (1999)¹²².

Neste trabalho, os ensaios iniciais de cristalização foram realizados utilizando a técnica de difusão de vapor em gota sentada. Para isso foi utilizado placas de cristalização CrystalQuick de 96 poços e as soluções de cristalização das empresas, Hampton Research (3 kits de 96 condições), e Qiagen (6 kits de 96 condições) somando um total de 864 condições diferentes. Essa abordagem (matriz esparsa) baseia-se na seleção de condições de cristalização conhecidas para macromoléculas. Dessa forma, foi possível realizar vários testes de cristalização, abrangendo uma grande variedade de condições. Esse procedimento foi realizado utilizando um robô Honeybee (Genomic Solutions), produzindo gotas com volume total de 1,4 μL , (0,7 μL de solução de proteína a 20 mg/mL e 0,7 μL de solução precipitante). As condições que apresentaram resultados promissores, isto incluem, cristais, *clusters*, micro cristais ou mesmo esferolitas, foram posteriormente otimizadas. Para isso, foram conduzidos experimentos de gota suspensa montados manualmente, em placas de cristalização com 24 poços. Para as Hbs na forma integral, foram produzidas gotas com um volume final de 4 μL (2,0 μL de solução de proteína a 20 mg/mL e 2,0 μL de solução precipitante). Para os ensaios de cristalização envolvendo a subunidade *d*-HbGp, foram produzidas gotas com um volume final de 3 μL (1,5 μL de solução de proteína a ~5 mg/mL e 1,5 μL de solução precipitante).

Figura 2.3.1 - Representação do experimento de cristalização por difusão de vapor em gotas suspensas. A gota contendo a proteína em solução apresenta igual composição ao reservatório do poço, porém, com metade da concentração, dessa forma, esta, tende a perder solvente lentamente, até que as concentrações da gota e do poço adquiram uma condição de equilíbrio.

Os experimentos de otimização das condições de cristalização consistiram em uma varredura (matriz incompleta) em torno das condições inicialmente obtidas. Para isso variou-se parâmetros como: concentração do agente precipitante, pH, tampão utilizado e temperatura ambiente. Para a crio proteção, foram preparadas soluções de mesma composição das que originalmente são obtidos os cristais, com a adição de 10% de etilenoglicol ou 10% de glicerol. Os cristais foram transferidos das soluções originais para a solução crio protetora com a ajuda de um *loop* (pequena alça de náilon), e em seguida, resfriados em nitrogênio líquido, onde permaneceram até a hora da coleta. Um procedimento alternativo foi utilizado para os cristais da HbGp, onde a uma solução crio protetora contendo 20% etilenoglicol foi adicionada gradativamente (em gotas de 0,5 μL) à gota de solução original contendo os cristais. Esse procedimento minimizou o efeito conhecido como choque osmótico, sofrido pelo cristal devido à transferência do meio original para uma ambiente químico hipertônico da solução crio protetora.

2.4 Coleta e processamentos dos dados

Os cristais da HbGp, HbEa e HbPe, foram coletados no National Sincrontron Light Source - NSLS (Brookhaven National Laboratory, USA). Os dados obtidos para a HbGp foram coletados no ano de 2009, na linha de luz X29, utilizando um detector ADSC Quantum 315, comprimento de onda de 1,00 Å, intervalo de oscilação de 0,5° e exposição de 1s por imagem. Os cristais de HbEa e HbPe, foram coletados no ano de 2011, na linha de luz X25, equipada com detector Pilatus 6M, comprimento de onda de 1,00 Å, intervalo de oscilação de 0,5°, com exposição de 1s por imagem. Os cristais de *d*-HbGp foram coletados no Diamond Light Source em Oxfordshire – Inglaterra, na linha de luz I02, equipada com o detector Pilatus 2M. Foi utilizado um comprimento de onda de 1,00 Å, intervalo de oscilação de 0,4° e exposição de 1s por imagem. A temperatura de coleta foi de 80 K em todos os casos. Os conjuntos de dados foram processados no programa Imosflm¹²³, o qual executa as seguintes rotinas: indexação das reflexões, refinamento dos parâmetros do cristal e detector, e integração das intensidades das reflexões. Para os cristais de HbGP e HbEa, os valores de mosaicidade foram mantidos fixos nos valores inicialmente preditos. Em sequência, o programa Scala¹²⁴ foi utilizado para reduzir à mesma escala relativa as intensidades medidas em diferentes imagens e simetricamente relacionadas. Nos conjuntos de dados de HbGp e HbEa, 5% das reflexões foram separadas para o cálculo do parâmetro R_{free} . Para os cristais da subunidade *d* da HbGp, que apresentaram um baixo número total de reflexões, foi reservado um valor de 10% das reflexões únicas para o cálculo do parâmetro R_{free} .

2.5 Substituição molecular

Dentre os métodos comumente utilizados para a solução do problema das fases em cristalografia, podemos destacar:

1. Métodos diretos, que são usados para a solução de estruturas pequenas, com até aproximadamente 200 átomos (embora em alguns casos podem chegar a 1000 átomos) e dados de difração a alta resolução, superior ou próxima a 1,0 Å, ainda hoje impraticáveis em cristalografia de proteínas¹²⁵.

2. Método dos átomos pesados, onde as fases $\alpha(hkl)$ são obtidas devido ao espalhamento diferenciado dos raios-X por átomos de elevado número atômico. Neste caso destaca-se o método MIR (*Multiple isomorphous replacement*).
3. Método do espalhamento anômalo, que depende da presença de átomos espalhadores anômalos no cristal, como selênio, por exemplo em uma fonte de comprimento de onda variável.
4. Método da substituição molecular, MR (*molecular replacement*), que utiliza um modelo tridimensional já conhecido para resolver a estrutura de interesse, supondo que exista semelhança entre elas.

A substituição molecular, é atualmente o método mais utilizado e também o mais rápido na resolução de estruturas cristalográficas de proteínas. Para tal, dois pré-requisitos são essenciais: um conjunto de dados cristalográficos de uma estrutura desconhecida, molécula alvo, e as coordenadas atômicas da estrutura com razoável identidade estrutural em relação à estrutura desconhecida, molécula molde. A resolução estrutural por MR envolve o posicionamento da molécula molde dentro da cela unitária do cristal da molécula alvo, de tal forma que o padrão de difração teórico emergente desse modelo assemelha-se com o experimental. Resumidamente, esse método envolve a determinação de seis parâmetros (três rotacionais e três translacionais), os quais descrevem totalmente a maneira como a molécula molde deve ser orientada dentro da cela unitária do cristal da molécula alvo. Em princípio, uma simples busca nesse espaço de seis dimensões poderia determinar a correta posição da molécula alvo, o que resultaria na melhor concordância entre os fatores de estrutura observados e calculados. No entanto, tal busca consome um elevado tempo computacional, inviabilizando o processo (com exceções para alguns casos onde a simetria do cristal reduz o número de parâmetros¹²⁶). Tal problema foi solucionado por Rossmann & Blow (1962) analisando as propriedades da função de Patterson¹²⁷ e dividindo o problema em duas etapas (rotação e translação) que são resolvidas separadamente. O mapa de Patterson é um mapa de vetores interatômicos. Self-Patterson vetores (que são relativamente curtos) conectam átomos pertencentes à mesma molécula e são usados na primeira etapa para descobrir a orientação correta da molécula na cela unitária (ou seja, resolver o problema da rotação). Na prática, isto é feito estabelecendo um raio máximo para a função de Patterson a ser considerada no cálculo da função de rotação de tal forma a eliminar ao máximo “cross vectors” (vetores entre átomos pertencentes as moléculas diferentes). Se uma das funções for mantida fixa e a outra rotacionada, podemos encontrar a orientação correta quando houver a

maior concordância entre os dois mapas. Na segunda etapa, “cross vectors” são usados para localizar a molécula na cela unitária, através de uma busca de translação em três dimensões. Uma vez que a molécula alvo encontra-se corretamente posicionada, fases podem ser calculadas desse modelo e combinadas com as amplitudes dos fatores de estrutura determinados experimentalmente. Este modelo inicial deverá ser cuidadosamente analisado no intuito de identificar se há ou não choques entre as moléculas na rede. Para a HbGp e HbEa, o processo de substituição molecular foi realizado através do programa Phaser (pacote CCP4)¹²⁸, utilizando-se como estrutura molde de procura o protômero da HbLt (código pdb – 2gtl) excluindo-se os grupos hemes e íons. No caso da estrutura da *d*-HbGp isolada, o molde de procura utilizado foi a subunidade *d* isolada do pdb referente a estrutura da HbGp íntegra.

2.6 Refinamento

A partir da substituição molecular, um modelo aproximado para a estrutura da unidade assimétrica, pode ser obtido, no qual as características mais marcantes da arquitetura molecular tornam-se aparentes. Isto é, o mapa de densidade eletrônica inicial, mostra evidências da posição das cadeias principais e laterais, e muitas vezes a presença de átomos pesados na rede. Fatores de estrutura calculados com base nesse molde geralmente discordam dos fatores de estrutura observados. O índice de acordo entre fatores de estrutura calculados e observados é comumente representado por um parâmetro denominado R_{factor} (ou R_{work} no programa Phenix), ou simplesmente “R”, definido por :

$$R = \frac{\sum ||F_{obs}|-|F_{calc}||}{\sum |F_{obs}|} \times 100 \quad (1)$$

De forma similar R_{free} é um parâmetro quantitativo que também indica quão bom é o modelo gerado. Ele é calculado da mesma forma que R, porém com uma porcentagem das reflexões separadas propositalmente durante o processo de escalonamento para essa finalidade (em torno de 5% a 10%). As reflexões utilizadas para calcular R_{free} não são utilizadas no processo

de refinamento e, portanto, este sempre assume valores maiores do que o de R. O parâmetro R_{free} é muito importante para monitorar o processo de refinamento, permitindo verificar se existe um “super refinamento” da estrutura ou não. O refinamento consiste no ajuste do modelo para que as amplitudes dos fatores de estrutura calculados cheguem o mais próximo possível as amplitudes dos fatores de estrutura observados, isto é, a obtenção de valores de R e R_{free} mais baixo possíveis. Os ajustes nesse modelo consistem na mudança de parâmetros posicionais (x,y e z além da troca, inserção ou remoção dos espalhadores atribuídos ao modelo) e em fatores de temperatura para todos os átomos na estrutura. Comumente, devido à limitação imposta pela resolução obtida, é preciso aplicar restrições estereoquímicas, onde o conhecimento prévio de distâncias e ângulos de ligações químicas etc. é imposto ao modelo, garantindo que o procedimento de ajuste de distâncias durante o refinamento não leve a um modelo conflitante com os conhecimentos de química estrutural. No refinamento da estrutura de HbGp, foi adotada a seguinte metodologia: substituição dos aminoácidos a partir da estrutura de *Lumbricus terrestris* usado para a substituição molecular com base nas sequências obtidas pelo sequenciamento massivo do transcriptoma do *G.paulistus*. Adicionalmente foram excluídos os grupos prostéticos; grupos heme, íons e glicosilações. Esse procedimento foi realizado no programa Coot, através de inspeção de mapas de densidade eletrônica $2|F_o|-|F_c|$, contornado a 1σ , e $|F_o|-|F_c|$, contornado a 3σ , alternadamente com ciclos de refinamento executados no programa PHENIX¹²⁹. Devido ao tamanho do modelo e a limitada resolução obtida, alguns cuidados especiais foram tomados durante o refinamento.

- A aplicação de simetria não cristalográfica (NCS – *non crystallographic symmetry*) foi essencial para impedir o sobre refinamento do modelo (que ocorre quando o número de informações experimentais, reflexões, é menor do que o número de parâmetros a serem refinados). Este procedimento exigiu considerável capacidade computacional (ocupando aproximadamente 13 Gb de memória virtual). Para isso foi utilizado um computador com processador Core i7 intel 2600 (Sandbridge) com 16Gb de memória RAM e Linux Ubuntu 11.04. Nestas condições, o ajuste do peso *X-ray/stereochemistry* foi de 0,1, para otimizar simultaneamente a geometria do modelo aos dados de difração de raios X.
- O programa ReadySet (pacote PHENIX) foi utilizado para estabelecer as restrições especiais de modificações químicos ligados a aminoácidos, tal como ocorre para as glicosilação e também para grupos prostéticos como heme.

Para a verificação da qualidade do modelo durante o refinamento foi utilizado o programa MolProbity¹³⁰. Foram levados em consideração parâmetros como: *rmsd bonds*, *rmsd angles*, digrama de Ramachandran, índices de choques e valores de R e R_{free}. Para a adição dos íon, foi levado em consideração a intensidade do mapa de diferença (Fo - Fc) para a região em questão, o ambiente químico nas redondezas do íon, e a composição de metais na amostra da ciano-HbGp utilizada nos ensaios de cristalização. As figuras geradas de estruturas tridimensionais e representações de densidade eletrônica foram feitas com o programa PyMOL(*The PyMOL Molecular Graphics System, Versão 1.4 Schrödinger, LLC*) .

2.7 Análises e comparações estruturais

Os alinhamentos múltiplos realizados nesse trabalho foram feitos usando o programa *muscle*¹³¹ e editados no programa Jalview¹³². A predição de propriedades físico-químicas das sequências obtidas foi realizado no servidor expasy (ProtParam tool). A predição do ponto de remoção do peptídeo sinal, foi feita no servidor SignalP4.1 e os pontos de glicosilação nas sequências foram preditos pelo servidor NetNGlyc 1.0. A análise dos sítios de metais com base no ambiente químico da região em questão foi feita no servidor MESPEUS (<http://mespeus.bch.ed.ac.uk/MESPEUS/>) Os cálculos de *rmsd* posicionais advindos das sobreposições estruturais foram feitos com o programa Coot (pacote CCP4). Os valores esperados para o *rmsd* (desvio médio quadrático) da sobreposição entre sequências homólogas foram calculados segundo a equação de Chothia e Lesk¹³³.

$$\Delta(\text{\AA}) = 0.4 e^{1.87(1-i)} \quad (2)$$

Neste caso, *i*, é fração de resíduos conservados para o alinhamento entre as duas sequências comparadas.

3 RESULTADOS E DISCUSSÃO – HBGP

3.1 Purificação

A *oxi*-HbGp foi extraída da hemolinfa dos indivíduos coletados em campo. O rendimento médio obtido com esse processo de extração foi de 10 mg de proteína por indivíduo sacrificado. Como esperado, a *oxi*-HbGp na sua forma íntegra, foi coletada no volume de exclusão durante a purificação por gel filtração. A análise por SDS PAGE (16%) das proteínas presentes no volume de exclusão revelou a presença de 7 bandas distintas, claramente divididas em dois grupos. Esse padrão de bandas caracteriza a HbGp na forma íntegra (contendo globinas e *linkers*). O grupo contendo três bandas, com massas entre 24 e 29 kDa, é atribuído às subunidades de *linkers*. O conjunto contendo as quatro bandas, com massas entre 14 e 20 kDa é atribuído às cadeias de globinas. A **figura 3.1.1** mostra imagens do processo de purificação e as bandas obtidas na caracterização inicial por SDS PAGE.

Figura 3.1.1 - Etapas da purificação da HbGp. (a) corte na região ventral, (b) coleta da hemolinfa (c) centrifugação, eliminando o particulado sólido indesejável. (d) aplicação na coluna cromatográfica Superdex 200. (e) Perfil da curva obtida na separação pelo processo cromatográfico, evidenciando a saída da HbGp no volume de exclusão. (f) Gel de SDS PAGE (16%), demonstrando a composição do complexo da HbGp. Sete tipos básicos de cadeias de proteínas fazem parte do complexo. Três *linkers* (L1, L2 e L3) com massa de aproximadamente 25 kDa e quatro tipos de globinas, a, b, c e d, com massas entre 14 e 20 kDa.

O estado de oxidação da fração obtida após o processo cromatográfico e a obtenção da forma *cianometa* (ou simplesmente *ciano*), através do procedimento descrito anteriormente, foram monitoradas por espectroscopia de absorção UV-Vis. O processo de oxidação e substituição do ligante (saída do oxigênio e entrada do cianeto), resulta claramente em modificações nas bandas Q, que originalmente são típicas da forma *oxi*. As alterações no espectro de absorção UV-Vis, podem ser observadas na **figura 3.1.2**.

Figura 3.1.2 - Espectros de absorção da HbGp na forma *oxi* (vermelho) e na forma *ciano* (azul). O processo de oxidação de Fe^{2+} para Fe^{3+} é acompanhado de mudanças nas bandas Q, entre 500 e 600nm.

A análise qualitativa por ICP-AES para a composição das amostras preparadas de *ciano*-HbGp usadas nos experimentos de cristalização, comprovou a presença dos seguintes elementos: C, N, O, Fe, Ca e Zn e descartou a presença dos elementos: Mg, Cu e P. Embora a presença de Cu já tenha sido determinada para amostras de HbLt¹³⁴, vale lembrar que a amostra em questão (*ciano*-HbGp) foi submetida a diversas etapas de purificação, onde os processos de diluição em tampão sem a adição de íons, pode ter removido íons ligados fracamente ao complexo da HbGp.

3.2 Sequenciamento

O sequenciamento massivo utilizando a plataforma Solexa/Illumina (HiSeq 2000), gerou em torno de 353 milhões de *reads* com tamanho 250 pb cada. A montagem do transcriptoma foi feita através da associação dos *reads* obtidos por meio do pacote Trinity. A montagem do transcriptoma foi feita no LNLS com a colaboração do Dr. Marcelo Falsarella Carazzolle. Foram obtidas 97.639 sequências únicas, maiores ou iguais a 500 bp. Os transcritos referentes as subunidades da HbGp foram identificados por comparação com as subunidades sequenciadas da HbLt, usando o programa BLASTx. Foram encontrados 13 transcritos com significativa similaridade. Seis transcritos são referentes as subunidades de globinas, cinco são referentes as subunidades de *linkers*, e outros dois são referentes ao receptor de lipoproteína de baixa densidade (LDL) (que possui um domínio homólogo ao presente nas subunidades de *linkers*). As sequências obtidas para globinas e *linkers*, apresentam uma porção N-terminal rica em resíduos hidrofóbicos. Essa região foi predita pelo servidor signalP 4.1 (www.cbs.dtu.dk/services/SignalP/) como uma região de peptídeo sinal, responsáveis pela migração dessas proteínas para o meio extracelular. O servidor é capaz de estimar o ponto de clivagem para o processo de remoção do peptídeo sinal de cada sequência. Dessa forma a numeração adotada para as subunidades da HbGp neste trabalho, referem-se às sequências editadas pelo processo de remoção do peptídeo sinal. As sequências de globinas obtidas foram nominadas *a1*, *a2*, *b*, *c*, *d1* e *d2*, onde *a1* e *a2* são isoformas da subunidade tipo *a*, partilhando 80% de identidade. A sequência *d1* refere-se a uma subunidade do tipo *d*, a qual a sequência já havia sido obtida experimentalmente em trabalhos anteriores pelo método de degradação de Edman¹³⁵. A sequência *d2* é aparentemente uma isoforma da subunidade *d1*, onde a identidade sequencial entre *d1* e *d2* é de apenas 58%. Para as sequências identificadas como *linkers*, três foram assinaladas como sendo as sequências L1a, L2a, L3a. Essas sequências compartilham em torno de 60 % de identidade com as subunidades equivalentes na HbLt, porém apresentam entre 26% e 35%, quando comparadas entre si. A quarta sequência encontrada é similar a sequência L4 da HbLt. Esta é considerada uma isoforma da subunidade L3a (acredita-se que poderia substituir L3a na estrutura do complexo⁸⁸), sendo assim, esta é denominada aqui, por L3b, sendo a identidade sequencial entre L3a e L3b de 43% (não partilhando mais do que 30% com as demais sequências de *linkers*). A quinta e última sequência identificada foi denominada L1b, partilhando 35% de identidade com L1a. É importante mencionar que a mesma apresenta ambiguidade, possuindo identidade quase que

equivalente quando comparada a subunidade L2a (34%). Uma comparação das propriedades físico-químicas previstas para as sequências da HbGp e da HbLt, mostram que não há grandes divergências com relação à massa total e ponto isoelétrico. As propriedades físico-químicas calculadas podem ser observadas na **tabela 3.2.1**.

Tabela 3.2.1 - Comparação das propriedades físico-químicas previstas para as sequências da HbGp e as da HbLt. Os valores previstos para a HbLt são mostrados entre parênteses. Nota-se que as isoformas *d2* e L1b apresentam as propriedades previstas mais divergentes em relação aos grupos a quais pertencem. As sequências de peptídeos sinais são caracterizadas pela alta concentração de resíduos hidrofóbicos, o símbolo “.” é usado para designar a posição de clivagem prevista pelo servidor signalP4.1.

Sequencia	Massa	pI	Glicosilações	Peptídeo sinal
Globina <i>a1</i>	17139,6 (17236,7)	6,39 (7,27)	0	MKFGLCCLVLVAVLGYAFA.DDDCCS
Globina <i>a2</i>	17095,2(17236,7)	5,96 (7,27)	0	FGLCCLVLVAVLGYAFA.DDDCCS
Globina <i>b</i>	16356,4 (15983,8)	6,85 (5,89)	0	MKTLVVLLALVAVAMA.SECDVL
Globina <i>c</i>	16780,3 (16921,4)	6,32 (6,06)	0	MLRLVLGLLAACSMAARA.HQFCC
Globina <i>d1</i>	16099,2 (15964,2)	6,19 (5,74)	0	MKVALVFLLLGVAAFASAD.DCSILE
Globina <i>d2</i>	16539,6 (15964,2)	7,22 (5,74)	0	MKPQIVVFIVLVSYVSGQ.CSILESLE
Linker L1a	25612,4 (25775,5)	5,34 (5,21)	0	MWYILLVGLAAAA.SDTYKDR
Linker L1b	25726,5 (-)	7,66 (-)	1	MRCLGKLLLLLLPVGMA.DFGS
Linker L2a	26729,0 (26064,1)	6,35 (5,51)	1	MWRLILSALVGLILA.DEPQGT
Linker L3a	24693,1 (24911,4)	4,83 (4,71)	1	MKSLRLLLPALAVITAVLA.DHH
Linker L3b	24518,5 (24113,8)	5,27 (5,21)	0	MRISTRILDAGTLAVLTLVLSFQRVDAA.EDP

As isoformas encontradas apresentam divergências significativas quando comparadas as suas homólogas canônicas (homólogos diretos das sequências apresentadas na estrutura da HbLt). A globina *d2* apresenta uma diferença de massa de 450 Da e um ponto isoelétrico muito próximo da neutralidade. O *linker* L1b também apresentou um ponto isoelétrico divergente. Diferente de todos os *linkers*, que possuem pI abaixo de 6.50, esta isoforma apresenta um ponto isoelétrico de 7.66. Todas as ferramentas de bioinformática utilizadas na predição de glicosilação em sequências, apontam que para as subunidades de globinas da HbGp, não existem glicosilações. Para as sequências de *linkers*, foram previstas, como possuindo uma única glicosilação, as subunidades L1b (Asn44), L2a (Asn243) e L3a (140).

3.2.1 Sequências de globinas

A presença de isoformas está em concordância com os dados apresentados na literatura para outras eritrocruorinas^{137,138}. Como mencionado anteriormente, a HbLt também possui uma isoforma *d2* reportada, com identidade sequencial *d1/d2* de 80%. Um alinhamento geral contendo as globinas da HbGp e de Hbs de outros oligoquetos é apresentado na **figura 3.2.1**. Observa-se que as sequências da HbGp apresentam um padrão de resíduos conservados em relação as outras globinas. Isso inclui a conservação dos resíduos de cisteína, responsáveis pela ligação dissulfeto intramolecular (caixas em cor cinza - **figura 3.2.1**), que todas as globinas de oligoquetos apresentam (posições 4 e 137 do alinhamento). O padrão de resíduos conservados também inclui os resíduos de cisteína, responsáveis pelas ligações dissulfeto entre as subunidades. Estas cisteínas, encontram-se nas posições 3 do alinhamento, para as globinas *a1*, *a2* e *c*, e posição 129 para as globinas *a1*, *a2* e *b* (caixas em cor vermelha - **figura 3.2.1**). Dessa forma, as sequências obtidas satisfazem as condições necessárias para a formação da estrutura dodecamérica ($[abc]_3[d]_3$), onde *abc* estão ligados de forma covalente. Nas globinas, 5 resíduos formam a primeira camada de aminoácidos ao redor do ligante no sexto ponto de coordenação, na região distal ao grupo heme. Esses resíduos ocupam as posições: 37, 51, 67, 71 e 113 do alinhamento. Com exceção do resíduo 67, que é a chamada histidina distal, as outras quatro posições são ocupadas por resíduos hidrofóbicos que se conservam no alinhamento. Esse ambiente hidrofóbico ao redor do ligante é conservado até mesmo quando as globinas da HbGp são comparadas as globinas humanas (não mostradas no alinhamento). A posição 37 é marcada pela presença de resíduos hidrofóbicos grandes, para as subunidades tipo *a*, *b* e *d*. Nas globinas humanas e subunidades tipo *c* dos oligoquetos, essa posição é ocupada por um resíduo de cadeia lateral alifática. Maillett e colaboradores¹³⁹, descrevem esta posição como importante regulador da afinidade dessas moléculas por O₂ e NO, alterando-se as constantes de afinidade em várias ordens de grandeza devido a mutações que inserem resíduos hidrofóbicos volumosos nesta posição. Na região proximal 3 resíduos hidrofílicos mostram-se altamente conservados entre as globinas de oligoquetos, ocupando as posições 94, 98 e 99 do alinhamento (caixas em cor azul - **figura 3.2.1**). As posições 94 e 98 são ocupadas por resíduos de histidina e glutaminas, respectivamente, e desempenham um papel importante nas interações entre as subunidade de globinas no complexo (formação dos dímero *a/d* e *b/c*). A posição 99 é ocupada pela chamada histidina proximal, que ocupa o quinto sítio de coordenação do grupo heme.

globinas do tipo *a*, *b* e *c*, partilham em torno de 70% de identidade. Isso demonstra que a HbGp e HbLt são mais similares do que a previsão feita com base somente na sequência *d1*¹⁰⁴.

Tabela 3.2.2 - Uma comparação da identidade entre as sequências homólogas de globinas pra as eritrocruorinas HbGp, HbLt, HbEa e HbPe. O símbolo * significa que a sequência em questão está incompleta.

	HbGp <i>a1</i>	HbGp <i>a2</i>	HbLt <i>a</i>	HbEa <i>a</i>	HbPe <i>a</i>
HbGp <i>a1</i>	X	80.1	70.27	66.22	58.78
HbGp <i>a2</i>	-	X	72.30	85.81	59.46
HbLt <i>a</i>	-	-	X	85.81	59.46
HbEa <i>a</i>	-	-	-	X	54.73
HbPe <i>a</i>	-	-	-	-	100

	HbGp <i>b</i>	HbLt <i>b</i>	HbEa <i>b</i> *	HbPe <i>b</i>
HbGp <i>b</i>	X	72.54	68.67	-
HbLt <i>b</i>	-	X	86,8	-
HbEa <i>b</i>	-	-	X	-
HbPe <i>b</i>	-	-	-	-

	HbGp <i>c</i>	HbLt <i>c</i>	HbEa <i>c</i>	HbPe <i>c</i>
HbGp <i>c</i>	X	70.00	72.00	62.00
HbLt <i>c</i>	-	X	89.33	60.00
HbEa <i>c</i>	-	-	X	60.00
HbPe <i>c</i>	-	-	-	X

	HbGp <i>d1</i>	HbGp <i>d2</i>	HbLt <i>d</i>	HbEa <i>d</i>	HbPe <i>d</i>
HbGp <i>d1</i>	X	58.87	57.14	53.90	43.55
HbGp <i>d2</i>		X	58.57	59.57	48.39
HbLt <i>d</i>			X	70.00	50.00
HbEa <i>d</i>				X	50.00
HbPe <i>d</i>					X

É justo mencionar que entre os *reads* obtidos, os quais referem-se às subunidades da HbGp, é possível encontrar sequências de variantes. Isso mostra que apesar da robusta montagem do transcriptoma, devemos considerar a existência de *reads* (fragmentos de sequências) referentes a variantes com altíssima identidade em relação às sequências obtidas (90% ou mais). Portanto pequenas variações nestas sequências são esperadas.

3.2.2 Sequências - *linkers*

A região N-terminal das subunidades *linkers* é caracterizada por assumir a conformação de longas hélices que posteriormente irão se associar em estruturas oligoméricas do tipo *coiled-coil*, que por sua vez, desempenham papel importante na formação da bicamada hexagonal das eritrocruorinas e clorocruorinas. Essa região apresenta um padrão característico de resíduos hidrofóbicos que formam a região interna do *coiled-coil*. A **figura 3.2.2** mostra o alinhamento das sequências de *linkers* da HbGp comparadas a outros anelídeos, com destaque para os resíduos que formam o *coiled-coil* na porção N-terminal. Nesta região se encontra os septetos típicos de *coiled coils* (*abcdefg*) onde as posições *a* e *d* são geralmente ocupadas por resíduos hidrofóbicos (destacados pelas caixas em preto na **figura 3.2.2**). As posições 65 e 66 do alinhamento são marcadas pela abertura de um “*gap*” entre as sequências de eritrocruorinas do tipo I (não eclipsadas), que resulta também na quebra da fase dos resíduos hidrofóbicos que formam o *coiled-coil*. Por isso, os *linkers* da estrutura da HbLt formam *coiled-coils* “imperfeitos”, fragmentados em duas porções, já que essa estrutura característica se desfaz na região onde ocorre a quebra de fase. Por outro lado, as sequências de *linkers* de eritrocruorinas do tipo II (HbAp e HbAm) são marcadas pela presença de resíduos na posição 65 e 66, os quais preservam uma única fase para o *coiled-coil*, estando em pleno acordo com os mapas de baixa resolução obtidos para a estrutura da HbAm⁶¹. Nesta estrutura, a densidade eletrônica na região dos *linkers* aponta o domínio N-terminal como uma longa hélice formando o *coiled-coil* sem quebras. As sequências obtidas para os *linkers* da HbGp (incluindo as isoformas L1b e L3b) apresentam o padrão conservado de resíduos hidrofóbicos igual ao da HbLt, incluindo a presença do *gap* nas posições 65 e 66, sugerindo fortemente que a HbGp é uma eritrocruorina do tipo I. Apesar de obedecerem aos padrões de resíduos hidrofóbicos que caracterizam a formação do *coiled-coil* (hidrofóbicos nas posições *a* e *d* do septeto), as isoformas L1b e L3b apresentam variações (incluindo a troca por resíduos de carga oposta) ao longo das posições dos septetos formados. As posições *g* e *e* dos septetos são comumente ocupadas por resíduos carregados, de forma que os resíduos na posição *e* interagem com os resíduos de carga oposta na posição *g* da hélice vizinha (interações secundárias). Dessa forma, analisando o padrão de resíduos carregados nas posições *g* e *e* de L1b, observa-se que nas posições 41g, 53e, 55g 62g obedecem aos padrões de carga de L2a, e apenas a posição 48g possui cargas equivalentes a L1a. As demais posições *e* e *g* do septeto não apresentam coerência entre L1b e L1a ou L2a. Considerando que o *coiled-coil* é a porção mais

significativa de contatos entre os *linkers* que formam o heterotrímero (L1 L2 L3), do ponto de vista das interações secundárias que formam essa estrutura, as sequência L1b aparentemente encontra-se mais adequada para representar a subunidade L2 no complexo.

Figura 3.2.2 - Alinhamento geral dos *linkers* obtidos para a HbGp e outros anelídeos. Observa-se que as sequências da HbGp apresentam um padrão de resíduos similares ao dos outros *linkers*. As posições 65 e 66 do alinhamento são marcadas pela abertura de um “gap” entre as sequências de eritrocruorinas do tipo I (não eclipsadas). L_{marinho} corresponde as sequências de *linkers* de espécies marinhas, *A.marina* (HbAm) e *A.pompojana* (HbAp).

O domínio LDL-A é a região mais conservada entre os *linkers*. Esse domínio, que se estende entre as posições 89 e 125 do alinhamento da **figura 3.2.2**, é rico em cisteínas (posições 85, 94, 101, 108, 114 e 125) que formam 3 pares de ligações dissulfeto. Existe ainda a conservação dos resíduos característicos do sítio de ligação de cálcio, sugerindo que a HbGp, assim como as outras sequências de anelídeos, apresenta esse sítio de ligação de metal. O domínio seguinte, que forma a estrutura β -barril na porção C-terminal, apresenta um padrão de resíduos hidrofóbicos que alternam com resíduos hidrofílicos. Esse um padrão conversado é característico de sequências que adquirem o enovelamento β -barril. O alinhamento também mostra que a ligação dissulfeto presente neste domínio, realizada entre os resíduos Cys145 e Cys247, é conservada, incluindo para as isoformas L1b e L3b. Essas características sugerem que as isoformas encontradas também devem assumir enovelamentos semelhantes aos de seus homólogos, e a princípio, poderiam ocupar as posições dos mesmos na estrutura. A **tabela 3.2.3** mostra as identidades das sequências comparadas no alinhamento da **figura 3.2.2**.

Tabela 3.2.3 - Uma comparação entre as identidades das sequências homólogas de linkers pra as eritrocruorinas HbGp, HbLt e HbEa. Não existe sequências disponíveis referentes aos *linkers* da HbPe.

	HbGp <i>L1a</i>	HbGp <i>L1b</i>	HbLt <i>L1</i>	HbEa <i>a</i>
HbGp <i>L1a</i>	X	35.40	56.00	52.21
HbGp <i>L1b</i>	-	X	32.30	31.39
HbLt <i>L1</i>	-	-	X	78.68
HbEa <i>L1</i>	-	-	-	X

	HbGp <i>L2a</i>	HbLt <i>L2</i>	HbEa <i>L2</i> *
HbGp <i>L2a</i>	X	59.32	61.11
HbLt <i>L2</i>	-	X	79.31
HbEa <i>L2</i>	-	-	X

	HbGp <i>L3a</i>	HbGp <i>L3b</i>	HbLt <i>L3</i>	HbLt <i>L3b</i>	HbEa <i>L3</i>
HbGp <i>L3a</i>	X	43.06	64.22	51.16	66.52
HbGp <i>L3b</i>	-	X	40.09	47.00	42.59
HbLt <i>L3</i>	-	-	X	47.22	77.17
HbLt <i>L3b</i>	-	-		X	47.91
HbEa <i>L3</i>					X

3.3 Cristalização

Os experimentos iniciais de cristalização envolvendo a ciano-HbGp, resultaram em duas condições com cristais bem formados. Na primeira, condição contendo 8% PEG 8000 Tris 0,1M pH 8,0, os cristais foram obtidos em 28 horas. Uma otimização variando-se a concentração de PEG 8000, pH e incluindo o cátion Ca^{2+} como aditivo, levou a condição aprimorada (cristais mais bem formados): 10% PEG8000, CaCl_2 50mM Tris/HCl pH7,5. Os cristais obtidos nesta condição apresentam um comportamento de auto degradação, deteriorando-se em um período de aproximadamente 4 dias. Devido à fragilidade excessiva, sua manipulação mecânica era inviável, resultando na quebra e perda do cristal. A segunda condição apresentou a seguinte composição: citrato de sódio 1,4M Tris/HCl 0,1M pH 8,0. Esses cristais foram obtidos entre 24 e 72 horas após a montagem dos experimentos de cristalização. Essa condição foi refinada, variando-se a concentração de precipitante e pH, chegando-se à condição otimizada: 1,2 M citrato de sódio 50mM Tris/HCl pH7,5. Esses cristais, por sua vez, apresentaram uma resistência mecânica muito superior em relação aos das condições mencionada anteriormente. Isso permitiu que os mesmos fossem manipulados com relativa facilidade. Devido à sensibilidade apresentada por esses cristais com relação ao choque osmótico, no processo de transferência dos mesmos para a solução de crio proteção, a própria solução crio protetora foi adicionada gradativamente na gota original contendo os cristais. Isso foi feito até que a concentração final da gota apresentasse 10% de etilenoglicol, ou 10% glicerol, dependendo da solução crio em questão. Os cristais mais resistentes, ou seja, aqueles que não apresentaram danos mecânicos aparentes neste processo, foram então submetidos ao processo de resfriamento. A **figura 3.3.1** mostra os cristais obtidos, resultados do processo de otimização.

Figura 3.3.1 - Cristais da *ciano-HbGp* obtidos em duas diferentes condições de cristalização (escala em μm). (a) 10% PEG8000 CaCl_2 50mM Tris/HCl pH7,5, frágeis e (b) que apresentam uma vida útil muito curta (imagem coletada 5 dias após ao aparecimento dos cristais). (c) Cristais obtidos na condição 1,4M Tris/HCl 0,1M pH 8,0. (d) Mono cristal obtido na otimização da condição de cristalização (1,2M citrato de sódio 2,5mM CaCl_2 50mM Tris/HCl pH7,5).

3.4 Coleta e processamentos dos dados

Os cristais obtidos na condição 10% PEG8000 CaCl_2 50mM Tris/HCl pH7,5 apresentaram um poder de difração muito limitado, resumindo-se a poucos pontos difusos na região de baixa resolução. Aparentemente os cristais são destruídos no processo de resfriamento. Já os cristais obtidos na condição citrato de sódio 1,4M Tris/HCl 0,1M pH 8,0 apresentaram padrões de difração com reflexões bem definidas e intensas. Nos primeiros experimentos de difração, ficou claro que parâmetros como; tamanho do cristal, colimação do feixe e condições de crio proteção, seriam importantes para a obtenção de bons conjuntos de dados. A condição que utiliza etilenoglicol como agente crio protetor foi a que melhor preservou a integridade do cristal no processo de resfriamento. A etapa de crio resfriamento do cristal foi considerada um dos principais “gargalos” para a obtenção de boas imagens de

difração. O processo de resfriamento abala significativamente a estrutura dos cristais de HbGp. Para os cristais maiores (0.20 mm os mais) os padrões de difração apresentaram imagens com pontos intensos, difusos e sobrepostos na região central da imagem. Este efeito de perda na qualidade dos dados obtidos para cristais maiores é associado ao resfriamento não homogêneo que ocorre nesse sistema. Assim os cristais menores, com tamanho entre 0,05 e 0,1 mm, combinados com etilenoglicol como agente crio protetor, mostraram-se mais resistentes ao processo de resfriamento em nitrogênio líquido. Dezenas de cristais foram testados até que um padrão de difração razoável fosse obtido. Cristais menores tem menor poder de espalhamento, e por isso, a disponibilidade de um feixe colimado (50 x 50 μ m) foi importante para minimizar o espalhamento incoerente causado pelo solvente ao redor do cristal, minimizando o ruído de fundo (*background*) nas imagens obtidas.

O espectrofotômetro de Uv-Vis, acoplado a linha de luz X26 (NSLS) permitiu coletar espectros do cristal simultaneamente ao processo de difração, e assim constatar quais são os efeitos da radiação ionizante nos cristais de *ciano*-HbGp. Logo nos primeiros segundos de exposição ao raio X, o cristal sofreu modificações significativas na região das bandas Q, adquirindo um perfil de Hb na forma reduzida (semelhante a forma *oxi*). Foram coletadas 240 imagens de difração, com oscilação de 0.5° para cristais da *ciano*-HbGp. A **figura 3.4.1** mostra um perfil dos pontos de difração gerados pelos cristais de *ciano*-HbGp e a predição de cela, realizada com boa concordância em relação às posições das reflexões na imagem. O processo de indexação, apontou o grupo espacial I222 (não descartando a possibilidade do grupo I2₁2₁2₁) pertencente ao sistema cristalino ortorrômbico. Os parâmetros de cela unitária preditos foram $a = 272.68$, $b = 319.90$ e $c = 333.18$ Å. Neste caso, ambos os grupos espaciais apresentam elementos centrados na cela unitária, o que resulta em ausências sistemáticas iguais nos padrões de difração. Assim não foi possível resolver essa ambiguidade apenas com base nas reflexões obtidas.

Figura 3.4.1 - Imagem do padrão de difração referente ao cristal da *ciano-HbGp*. Foram coletadas 220 imagens com uma oscilação de $0,5^\circ$. Cada imagem apresentou aproximadamente 13.800 reflexões e uma mosaicidade estimada de $0,55^\circ$. A predição de cela atribuída ao sistema cristalino ortorrômbico está em boa concordância com as reflexões no padrão de difração, onde predominam reflexões parciais. A intensidade das reflexões sofre uma queda súbita, entre as regiões de 7 a 5 Å no padrão de difração. Esses danos causados a rede são fruto do elevado teor de solvente (estimado em $\sim 74\%$) que esses cristais possuem, maximizando os danos causados pelo resfriamento.

Um total de 3.038.803 reflexões (que são majoritariamente parciais) foram integradas pelo programa Imosflm com uma média de 13.800 reflexões por imagem. Neste processo o parâmetro mosaicidade foi mantido fixo no valor de predição inicial de $0,55^\circ$, evitando que um valor superestimado para mosaicidade resultasse na sobreposição de *spots* preditos (pontos de difração). Os principais dados obtidos durante indexação, integração e processamento dos dados podem ser vistos na **tabela 3.4.1**. O escalonamento do conjunto de dados foi realizado com o programa Scala. Foi obtido um total de 1.105.247 reflexões completas, sendo 237.062 tratam-se reflexões únicas com um valor de R_{merge} de 11,2% (50,0% na última faixa de resolução). Isso resultou em um conjunto de dados processado com 99,8% (99,2% na última faixa de resolução) de completeza e uma resolução na última camada de $3,2\text{\AA}$. O conteúdo da unidade assimétrica foi analisado com base no coeficiente de Matthews C_M^{143} , que permite uma estimativa da porcentagem de solvente e número que elementos que compõem a unidade assimétrica. Para uma unidade assimétrica composta por três protômeros

(assim como foi encontrados pelo programa Phaser - ver adiante), o C_M calculado foi de 4.1 $\text{\AA}^3/\text{Da}$, e a porcentagem de solvente estimada de 74%. Essa elevada taxa de solvente na composição do cristal está nos limites dos valores esperados para cristais de proteínas (que oscila entre 40 e 70%), e normalmente esta associado a cristais que difratam a baixa resolução. Esta é também uma das razões pelas quais esses cristais são pouco resistentes ao processo de crioproteção¹⁴³.

Tabela 3.4.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução (3.2 $\text{\AA}$).

Grupo espacial	I222
Dimensões de cela a, b, c ($\text{\AA}$)	272,68, 319,90, 333,18
Detector	ADSC Quantum 315
Fonte de raios-X	NSLS X29
Comprimento de onda ($\text{\AA}$)	1,00
Resolução	50,0–3,2 (3,37–3,20)
Multiplicidade	4,7 (4,2)
Rmerge(%)	11,2 (50,0)
Completeza	99,8 (99,2)
Reflexões totais	1105247 (145116)
Reflexões únicas	237062 (34189)
Intensidade ($I/\sigma I$)	11,2 (3,1)
B-fator no Wilson plot ($\text{\AA}^2$)	60,4

3.5 Substituição molecular e refinamento do modelo

O problema das fases e a ambiguidade do grupo espacial foram resolvidos por substituição molecular utilizando a estrutura homóloga do protômero da HbLt (código de acesso - *2gtl*) como molde (utilizando o programa Phaser¹²⁸). O modelo de procura representa 1/12 da unidade biológica e os grupos heme e metais ligados a estrutura foram excluídos do modelo original. Posteriormente, as regiões contendo esses grupos prostéticos foram usadas como critérios de avaliação para a solução obtida (avaliado-se o mapa da diferença e as deficiências na densidade eletrônica em virtude de ausência de espalhadores). Ao todo, foram localizados três protômeros na unidade assimétrica, o que representa um quarto da unidade

biológica. Dessa forma a unidade biológica (com simetria 622 **figura 3.5.1 - c**), é gerada pelas operações de simetrias presentes no grupo especial I222. Isso obriga que a partícula esteja “sentada” sobre uma posição especial, a saber a intersecção dos três eixos de ordem 2 que o grupo espacial possui. Essa configuração só é possível para o grupo I222, já que no grupo $I2_12_12_1$ os eixos de simetria não se tocam, e dessa forma não é possível acomodar a unidade assimétrica em questão e ao mesmo tempo gerar a unidade biológica por operações de simetria cristalográfica. A **figura 3.5.1** mostra a posição da unidade assimétrica na cela unitária e as operações de simetria, referentes ao grupo especial I222.

Figura 3.5.1 - A unidade assimétrica é composta por três protômeros, cada um deles é colorido de uma cor diferente. Em branco estão representados os simétricos relacionados que formam a unidade biológica. (a) mostra a a unidade assimétrica ao longo do eixo **b**. (b) mostra o conteúdo da unidade assimétrica ao longo do eixo **a**. (c) mostra as operações de simetria presentes na unidade biológica isolada. Para ser compatível com a posição espacial que ocupa do grupo I222, a partícula se alinha de tal forma que um membro de cada família de eixos de ordem 2 é coincidente com os eixos cristalográficos paralelos a a e b e o eixo de ordem 6 (que possui um eixo de 2 como subelemento) ao longo do eixo c. (d) Cella unitária e a unidade biológica da HbGp.

A unidade biológica apresenta simetria do grupo pontual 622, que possui duas famílias independentes de eixos de ordem 2 ambas perpendiculares ao eixo de ordem 6. Para ser compatível com a posição espacial que ocupa do grupo I222, a partícula se alinha de tal forma

que um membro de cada família de eixos de ordem 2 é coincidente com os eixos cristalográficos paralelos a **a** e **c** e o eixo de ordem 6 (que possui um eixo de 2 como subelemento) ao longo do eixo **b**. A localização do primeiro protômero na unidade assimétrica gerou um valor de $R = 57,2\%$ e um LLG (*log likelihood*) de 1446. A localização do segundo protômero resultou em $R = 53,5\%$ e $LLG = 5509$, e por fim a localização do terceiro e último protômero resultou em $R = 49,6\%$ e $LLG = 12050$. Não existe um valor absoluto para o parâmetro LLG que indique uma correta atribuição das fases por substituição molecular. Os valores obtidos dependem do número de átomos no sistema e de quão próximas são as estruturas molde e de interesse, porém, a progressiva queda dos valores de R e o aumento dos valores de LLG, indicam que procedimento foi feito corretamente. Os mapas de densidade eletrônica iniciais ($2F_o - F_c$) claramente mostram as evidências de cadeia laterais e grupos prostéticos, excluídos no processo de substituição molecular, o que reforça o êxito na correta atribuição das fases. A **figura 3.5.2** mostra regiões de destaque para o mapa de densidade eletrônica do modelo inicial. A unidade assimétrica é composta por um total de 45 subunidades, sendo 36 globinas, e 9 *linkers* (somando 7.108 resíduos). Mesmo com a determinação de sequências de isoformas, tanto para globinas como para *linkers*, optou-se por utilizar apenas uma das possíveis isoformas (havendo isoformas) como representante de uma dada subunidade. O processo de escolha foi feito com base na sequência que melhor apresentasse coerência com a densidade eletrônica calculada. Dessa forma, todas as subunidades de globinas tipo *a* e *d*, são representadas pelas sequências *a2*, e *d1*, respectivamente. Para os *linkers*, as subunidades L1 e L3, são descritas pelas sequências L1a e L3a, que são as sequências ortólogas das subunidades L1 e L3 do modelo da HbLt.

Figura 3.5.2 - (a) Visão geral da densidade eletrônica (2Fo-Fc), onde é possível observar os elementos de estrutura secundária assumidos na HbGp. (b) *Omit map* calculado para a região das globinas, confirmando a presença do heme/ferro na região. (c) *Omit map* na região dos *linker* L3, correspondente a um sítio de ligação de cálcio (nível de contorno em 2 sigma).

Essa abordagem favoreceu o uso de simetria não cristalográfica aplicada ao modelo. Uma vez que tal modelo apresenta um elevado número de átomos (59.091), e um limitado número de reflexões (237.062), é preciso reduzir tanto quando possível o número de parâmetros a serem refinados. Como consequência, no modelo final, as subunidades do mesmo tipo (relacionadas por simetria não cristalográfica) são todas estruturalmente equivalentes ($rmsd = 0.06\text{\AA}$ ou menores). A **tabela 3.5.1** mostra as subunidades relacionadas por simetria não cristalográfica. O uso de NCS, combinado ao refinamento dos fatores de temperatura por grupos (um valor associado a cadeia principal e um associado a cadeias laterais), reduziu consideravelmente o número de parâmetros a serem refinados, obtendo-se, ao final do processo de refinamento, valores de R e R_{free} iguais a 21,65 e 23,52, respectivamente.

Tabela 3.5.1 - Simétricos relacionados. O modelo da HbGp possui uma unidade assimétrica com 45 subunidades. As letras na tabela correspondem a nomenclatura das cadeias no arquivo PDB.

Tipo	Protômero 1			Protômero 2			Protômero 3		
<i>a</i>	A	E	I	P	T	X	e	i	m
<i>b</i>	B	F	J	Q	U	Y	f	j	n
<i>c</i>	C	G	K	R	V	Z	g	k	o
<i>d</i>	D	H	L	S	W	a	h	l	p
L1	M			b			q		
L2	N			c			r		
L3	O			d			s		

Foi possível observar uma clara evolução nos mapas de densidade eletrônica em consequência dos ciclos de refinamento aplicados ao modelo. A **figura 3.5.3** mostra a evolução do mapa de densidade eletrônica ($2F_o - F_c$) contornado em $1,5\sigma$, na região do resíduo TRP132 da cadeia A. Os mapas evidenciam não apenas a presença das cadeias laterais, mas também a correta orientação das mesmas. Devido a baixa resolução, não foram incluídas moléculas de solvente no modelo.

Figura 3.5.3 - Evolução do mapa de densidade eletrônica. A figura mostra a densidade mostra o mapa $2F_o - F_c$, calculado em $1,5\sigma$ para a região do TRP132 da cadeia A. (a) mapa do modelo inicial e (b) mapa calculado para o modelo final, após o processo de refinamento.

Os pesos impostos para manter e estereoquímica padrão do modelo durante o refinamento podem ser considerados bastante satisfatórios, resultando simultaneamente em valores aceitáveis para R, como comentado anteriormente, e em baixos valores de *rmsd* para comprimentos de ligações e ângulos de ligações. Os valores dos ângulos Phi e Psi, também estão de acordo com os valores esperados para estruturas com esta resolução. Os principais dados obtidos para o processo de refinamento podem ser observados na **tabela 3.5.2**. O diagrama de Ramachandran para o modelo da HbGp pode ser observado na **figura 3.5.4**.

Figura 3.5.4 - Diagrama de Ramachandran para a estrutura da *ciano-HbGp* (1,01% de resíduos fora da região permitida).

Tabela 3.5.2 - Estatísticas do refinamento e estereoquímica do modelo.

Reflexões usadas no refinamento	236866
R e R_{free}	21,65 e 23,52
B-fator médio.	59,336
No. de átomos (proteína)	59091
No. de átomos (ligantes)	1629
Ramachandran plot	
Região permitida (%)	95,58% 6775/ 7108
Região adic. permitida (%)	3,40% 229/7108
Outliers (%)	1,01% 104/7108
R.m.s. bond lengths (Å)	0,01
R.m.s. bond angles (°)	0,96

Como consequência dos bons mapas obtidos, detalhes importantes nas subunidades da HbGp puderam ser observados. Estes incluem; grupos heme, glicosilações, íons coordenados, ligações dissulfeto etc. As cadeias laterais internas apresentam em geral, melhor definição, suficiente para sugerir algumas incoerências nas sequências inicialmente propostas. Dessa

forma, algumas poucas divergências entre sequências e densidades eletrônicas foram encontradas, permitindo uma escolha mais criteriosa de quais sequências de isoformas melhor representariam a região em questão. A **figura 3.5.5**, mostra os mapas calculados para vários resíduos das subunidades do tipo *a*, onde a sequência *a1* foi inicialmente modelada, constatando que em realidade a sequência *a2* concorda melhor com as densidades eletrônicas do modelo. O mesmo critério foi aplicado para as subunidades tipo *d*, descartando a possibilidade das mesmas serem representadas pela sequência *d2*. Neste caso as incoerências da sequência *d2* modeladas nas densidades eletrônicas, são somadas ao fato de que a sequência *d1* é comprovadamente isolada do complexo em pH alcalino¹⁰³.

Figura 3.5.5 - Mapas calculados para algumas regiões das subunidades tipo *a*, ainda com a sequência da isoforma *a1* (estruturas em cinza), evidenciando que a sequência de *a2* (estruturas em azul) melhor concorda com a densidades eletrônica. (2Fo-Fc – mapa cinza – sigma = 1.0 e Fo-Fc em verde – sigma = 3 / vermelho – sigma = -3)

A densidade eletrônica na região dos grupos heme, é suficiente para evidenciar a presença e orientação dos mesmos. Uma vez que, a amostra cristalizada foi comprovadamente a *ciano-HbGp* (espectro de Uv-vis – **figura 3.1.2**) a espécie química CN^- foi modelada como o sexto ligante nos átomos de ferro. É justo mencionar que tal densidade é inconclusiva para estabelecer qual espécie química majoritariamente esta ligada ao centro metálico. As medidas espectroscópicas realizadas simultaneamente com a coleta de dados resultam em modificações na configuração eletrônica do centro metálico, apontando uma aparente evolução da forma oxidada, para a forma reduzida (de Fe^{3+} para Fe^{2+}) logo após a exposição ao raio X (**figura 3.5.6 - d**). Assim, a questão de qual espécie química, e como esta encontra-se ligada ao sexto sítio de coordenação do heme permanece ambígua. A densidade eletrônica calculada também confirmou a presença de íons Ca^{2+} coordenados ao domínio tipo receptor-

LDL das subunidades *linkers*. Estes estão presentes em posições equivalentes aos íons Ca^{2+} encontrados nas subunidade *linkers* da estrutura da HbLt. Foram modelados 9 íons de Ca^{2+} (um para cada *linker* da unidade assimétrica). As subunidades tipo L3 também apresentaram outros sítios com grande deficiência nos mapas de diferença (visíveis nos mapas de $F_o - F_c$ a um $\sigma = 10.0$), indicando a presença de outros metais/átomos pesados coordenados a estrutura. Para estas regiões da estrutura foram modelados íons Zn^{2+} , considerando a natureza dos resíduos que compõem o ambiente químico (particularmente a primeira esfera de coordenação), e a comprovada presença de Zn^{2+} na amostra por análise de ICP.

Figura 3.5.6 - Mapa de densidade eletrônica ($2F_o - F_c$) calculado para a região da globina α . Detalhes importantes podem ser observados como: (a), grupo heme, (b) ligações dissulfeto intramoleculares, (c) ligações dissulfeto intermoleculares. (d) Modificação no espectro de absorção após a exposição ao raio X, a amostra inicia-se como um típico espectro da forma *cianometa*, evoluindo para a uma configuração semelhante a da forma reduzida *oxi-HbGp*.

Um total de 6 íons Zn^{2+} foram adicionados ao modelo cristalográfico, dois para cada uma das três subunidades L3 existentes. Os sítios de coordenação formados estão: 1 - no domínio LDL-A (esse sítio específico da sequência L3a da HbGp) e 2 - na região de contato entre os domínios β -barril das subunidades L2 e L3 de protômeros diferentes. O único íon Zn^{2+} presente na estrutura cristalográfica da HbLt encontra-se na cavidade (“boca”) formada pelo β -barril da subunidade L2. Este é um sítio conservado na para as subunidades tipo L2 da HbGp, porém, embora os mapas de diferença ($F_o - f_c$) para o modelo cristalográfico da HbGp indiquem a presença de um espalhador nesta posição, o ambiente químico rico em resíduos positivamente carregados (Arg123 e Arg209) é inconsistente com a presença de um cátion Zn^{2+} nesta posição. Dessa forma, não foram adicionados íons Zn^{2+} para as subunidades tipo L2 da HbGp. Dentre as sequências previstas como glicosiladas na estrutura da HbGp, estão as

subunidades L2 e L3. No entanto, apenas as subunidades tipo L3 apresentam densidades eletrônicas correspondentes a glicosilações na posição predita. Nesta posição foi modelado o grupo N-acetilglicosamina (NAG) ligado de forma covalente ao resíduo Asn121. A glicosilação e a presença do sítio de Zn^{2+} , são evidências fortes de que a sequência L3a é a dominante para o modelo cristalográfico da HbGp resolvido neste trabalho, já que a sua isoforma (L3b) não apresenta conservação para os sítios de glicosilação e ligação de Zn^{2+} . **figura 3.5.7** mostra algumas regiões dos mapas de densidade eletrônica calculados para as subunidades de *linkers* L3 da estrutura da HbGp. Várias são as evidências nos mapas de densidade eletrônica apontando que, coordenado a estrutura, existe moléculas pequenas e/ou íons.

Figura 3.5.7 - A alta qualidade dos mapas de densidade eletrônica resultaram na determinação de novos grupos químicos ligados ao *Linker* L3: (a) Íon Zn^{2+} coordenado internamente no domínio LDL-A (sítio de zinco 1). (b) Íon Zn^{2+} coordenado na interface entre *linker* L3 e *linker* L2 do protômero adjacente (sítio de zinco 2), (c) Glicosilação – NAG; e (d) Íon de cálcio ligado ao domínio LDL-A.

Porém, a grande maioria dessas regiões é inconclusiva, e portanto, permaneceu não modelada. Vale lembrar que essas densidades são mais uma evidência de que essa estrutura pode atuar com carregadora de outras espécies químicas, estendendo sua função além da de carregadora de oxigênio. A **figura 3.5.8** mostra algumas das densidades eletrônicas as quais não foram atribuídos nenhum tipo de espalhador (átomo).

Figura 3.5.8-Densidades eletrônicas do modelo cristalográfico da HbGp onde não foram atribuídos espalhadores. (a) região equivalente ao sítio de Zn^{2+} encontrado na subunidade L2 da estrutura da HbLt. A presença dos resíduos de arginina é conflitante com a coordenação de um cátion metálico nessa posição. (b) Região do β -barril das subunidades L3. (c) Interface entre o domínio LDL-A e o β -barril das subunidades L3. (d) Interface entre os domínios LDL-A das subunidades L1 e L2 do mesmo protômero.

3.6 Estrutura da HbGp – uma visão geral

A estrutura determinada da HbGp é uma estrutura hexagonal, composta por duas camadas, cada uma contendo 6 protômeros. Assim como a HbLt, a HbGp, possui uma rotação relativa de 16° entre as duas camadas hexagonais. Isso caracteriza a HbGp como uma eritrocruorina tipo I, ou “não eclipsada”, reforçando a ideia de que todos os oligoquetas

apresentariam eritrocruorinas com camadas hexagonais não eclipsadas⁸. A unidade biológica é uma estrutura com 180 subunidades (144 globinas e 36 *linkers*), com um diâmetro de aproximadamente 286 Å e uma altura de 190 Å. A **figura 3.6.1** mostra uma visão geral da estrutura da HbGp e suas dimensões comparadas a outras estruturas do banco de dados PDB.

Figura 3.6.1 - Partícula inteira da HbGp (simetria 622), com um diâmetro de aproximadamente 286 Å e uma altura de 190, comparada a outras estruturas disponíveis no banco de dados.

Os protômeros que formam as duas camadas hexagonais são estruturas oligoméricas que assumem a forma de um cogumelo com aproximadamente 130 Å de comprimento. Essas estruturas oligoméricas podem ser divididas em duas partes; um dodecâmero de globinas, e um trímero de *linkers*. O dodecâmero de globinas, ou *cap*, é uma estrutura em forma de cúpula, assim como presente na HbLt, composta por 3 heterotetrâmero de globinas, onde as subunidades tipo a_1 , b_1 , e c_2 (c do tetrâmero vizinho – ver adiante) estão conectadas por ligações dissulfeto. O trímero de *linkers* por sua vez, é uma associação das subunidades tipo L1, L2 e L3, formando uma estrutura em forma de funil, onde a “cabeça dos *linkers*” (porção C-terminal) interage com o dodecâmero de globinas. A **figura 3.6.2** mostra o padrão de cores adotado neste trabalho para designar os vários tipos de subunidades e os vários os níveis de hierarquia de associação das mesmas na HbGp.

Figura 3.6.2 - A HbGp é uma bicamada hexagonal que contém 12 protômeros. Ao todo são 144 cadeias de globinas e 36 estruturas tipo *linkers*. Os protômeros são constituídos por um dodecâmero de globinas (*cap*) e um heterotrímero de *linkers* (L1, L2 e L3). As cores utilizadas para diferenciar os tipos de subunidades que compõem o complexo da HbGp foram adotadas como padrão neste trabalho.

3.6.1 Estrutura da HbGp – Globinas e a formação do Cap

Na HbGp, as globinas apresentam o típico enovelamento descrito por Perutz (3/3) contendo todas as hélices descritas, com exceção da pequena hélice D, que é substituída por um curto *loop*. A hélice C está presente na forma de hélice 3_{10} , assim como nas globinas de mamíferos. A maior divergência nos enovelamentos das globinas surge nos *loops* que conectam as hélices A e B, entre as posições 19 e 30 do alinhamento, **figura 3.2.1**. Esta posição é caracterizada por uma baixa conservação e também por inserções e deleções em algumas sequências. Assim como as globinas de outros anelídeos, as globinas da HbGp, quando comparadas com as subunidades da Hb/Mb humana, possuem a hélice F consideravelmente maior. Nesta região, existe uma inserção de 3 resíduos para sequências de

anelídeos, resultando em um curto *loop* (4 resíduos) conectando as duas hélices E e F, e um aumento de 7 resíduos na conformação helicoidal da hélice F. Por consequência, a histidina proximal coordenada ao átomo de ferro (posição 100 – **figura 3.2.1**), ocupa a posição 15 da hélice F. As subunidades de globina da HbGp apresentam um enovelamento muito similar entre si, apesar de apresentarem baixa identidade sequencial. Uma sobreposição das coordenadas atômicas mostra grande coerência entre os enovelamentos. Essa semelhança é ainda maior quando as subunidades da HbGp são comparadas com as suas equivalentes na HbLt. A **figura 3.6.3** mostra uma sobreposição das globinas da HbGp e HbLt.

Figura 3.6.3 - Sobreposição das globinas da HbGp (em cores), com suas homólogas equivalentes da HbLt (em cinza). Logo abaixo estão representados os fatores de temperatura. As cores “quentes”, como amarelo, laranja e vermelho, representam as regiões de maior desordem.

A **tabela 3.6.1**, mostra os valores de *rmsd* resultantes da sobreposição entre as subunidades. Note que os valores de *rmsd* obtidos da sobreposição das cadeias homólogas entre HbGp e HbLt resulta em valores ligeiramente menores do que os estimados pela equação de Chothia e Lesk¹³³, evidenciando que o enovelamento dessas estruturas é altamente conservado.

Tabela 3.6.1 - Mostra os valores de *rmsd* resultantes da sobreposição entre as subunidades de globina. Os valores preditos para as globinas através da equação de Chothia e Lesk (mostrados entre parênteses) são maiores do que os encontrados na sobreposição das estruturas.

Tipo	<i>HbGp a</i>	<i>HbGp b</i>	<i>HbGp c</i>	<i>HbGp d</i>	<i>HbLt a</i>	<i>HbLt b</i>	<i>HbLt c</i>	<i>HbLt d</i>
<i>HbGp a</i>	X				0,5 (0.6)	1,2	0,9	1,1
<i>HbGp b</i>	1,1	X			1,1	0,4(0.6)	1,2	0,9
<i>HbGp c</i>	0,9	1,2	X		0,9	1,3	0,4(0.6)	1,1
<i>HbGp d</i>	1,1			X	1,2	0,8	1,3	0,7(0.9)

A presença de resíduos hidrofóbicos na vizinhança dos grupos heme também é conservada para as subunidades da HbGp, destacando-se a presença de hidrofóbicos volumosos para as subunidade *a*, *b* e *d*, na posição 38 do alinhamento. A **figura 3.6.4**, mostra como estão dispostos os resíduos na primeira esfera de coordenação ao redor do sexto sítio de coordenação do grupo heme.

Figura 3.6.4 - Primeira esfera de resíduos ao redor do sexto sítio de coordenação do grupo heme. (a) globinas tipo *a*, (b) globinas tipo *b*, (c) globinas tipo *c* e (d) globinas tipo *d*. Existe uma predominância de resíduos hidrofóbicos.

O heterotetrâmero pode ser descrito como sendo composto por dois heterodímeros, cujos monômeros interagem diretamente pelos grupos hemes e resíduos situados nas hélices E e F. São formados pares entre as subunidades *a/d* e *b/c*, onde os contatos são majoritariamente

de natureza polar (ligações de hidrogênio ou pontes salinas), **figura 3.6.5**. Dois resíduos altamente conservados entre as globinas de anelídeos, presentes na hélice F, são responsáveis pela interação direta com os grupos propionato do heme da subunidade complementar no dímero. Esses resíduos são Gln98 e His99 (alinhamento, **figura 3.2.1**).

Figura 3.6.5 - Os dímeros formados *a/d* e *b/c*, possuem uma interface de contato conservada. Tal interface é conhecida para outras estruturas de hemoglobinas, como mencionado anteriormente⁸⁹. Área total entre *a/d*(HbGp) = 910,5Å² e área total entre *b/c* (HbLt) = 1146,9Å²

Os dímeros formados *a/d* e *b/c*, associam-se em uma estrutura tetramérica, que apresenta um pseudo eixo de ordem dois. Para a formação desta estrutura, cada dímero contribui com um sulco, formado pelas hélices B e G, e uma protuberância, formada pela hélice A. Essas estruturas são complementares, de forma que a hélice A da subunidade tipo *d* (dímero *a/d*) interage com o sulco entre as hélices B e G da subunidade tipo *c*, no dímero *b/c*. Da mesma forma, a hélice A da subunidade tipo *b*, no dímero *b/c*, interage com o sulco, formado pelas hélices B e G, na subunidade tipo *a*, do dímero *a/d*. Isso resulta também em uma interação cruzada entre os resíduos que conectam as hélices A e B das subunidades tipo *a*

e *c*. Por fim, resíduos de cisteína presentes nos *loops* que conectam as hélices G e H das subunidades tipo *a* e *b*, formam uma ligação dissulfeto, conectando os dois dímeros. Essas interações são representadas na **figura 3.6.6**. Note que o tetrâmero formado tem uma pseudo eixo de simetria de ordem 2 que passa entre a interação cruzada das hélices B das subunidades *a* e *c*.

Figura 3.6.6 - Formação do heterotetrâmero. Para a formação desta estrutura, cada dímero contribui com um sulco, formado pelas hélices B e G, e uma protuberância, formada pela hélice A. Os resíduos de cisteína presentes nos *loops* que conectam as hélices G e H das subunidades tipo *a* e *b*, formam uma ligação dissulfeto conectando os dois dímeros.

A estrutura tetramérica é repetida 3 vezes formando um trímero de tetrâmeros. Essa estrutura é denominada “*cap*” ou dodecâmero e possui simetria de ordem 3. A hierarquia de empacotamento das subunidades que formam o *cap* pode ser vista na **figura 3.6.7**. Os contatos entre tetrâmeros que geram a estrutura do *cap* ocorrem entre as subunidades a_1 e c_2 e d_1 e d_2 (onde a_1 é a subunidade tipo *a* do primeiro tetrâmero escolhido arbitrariamente, e c_2 é a subunidade tipo *c* do tetrâmero adjacente).

Figura 3.6.7 - Hierarquia de associação das subunidades de globinas pra a formação do *cap*. As quatro subunidades básicas se associam inicialmente nos dímeros *a/d* e *b/c*, que posteriormente formam os heterotetrâmero *abcd*, onde a primeira ligação dissulfeto é formada. Três heterotetrâmeros se associam formando a estrutura do *cap*, onde a segunda ligação dissulfeto é formada.

O contato envolvendo as subunidades a_1 e c_2 ocorre entre as hélicas A de ambas as subunidades. Esse contato inclui as cisteínas N-terminais das duas subunidades (posição 3 do alinhamento, **figura 3.2.1**) resultando em uma ligação dissulfeto entre tetrâmeros diferentes. As demais interações são majoritariamente polares, e envolvem o *loop* que conecta as hélices E e F da subunidade tipo a_1 e início da hélice H da subunidade c_2 . Os contatos entre as subunidades tipo d na estrutura do *cap* formam um trímero (d_3) que está relacionado pelo mesmo eixo de simetria de ordem três que caracteriza a estrutura do *cap*. As subunidades do tipo d , são as únicas subunidades dentro do *cap* que fazem contatos entre si. Essas interações ocorrem entre a hélice B em d_1 , com a hélice G, em d_2 , e entre o *loop* que conecta as hélices C e F em d_1 , com o início da hélice H, em d_2 . As interações entre subunidades do tipo d , são de natureza polar, destacando-se a ponte salina feita entre os resíduos Asp 37 em d_1 e Arg 34 em d_2 . O trímero de globinas *abc* isolado a partir da partícula inteira da HbLt exibe um pequeno comportamento cooperativo ($C_{Hill} \sim 1,3$), enquanto que a fração contendo apenas a subunidade d não exibe nenhuma cooperatividade. Porém, quando o trímero *abc* é reassociado com as subunidades d formando novamente a estrutura do *cap*, o sistema apresenta em pH 6,8¹³⁶, um comportamento cooperativo que é indistinguível comparado a partícula inteira ($C_{Hill} = 7,8$). Dessa forma a estrutura do *cap* é apontada como unidade básica com comportamento cooperativo dentro da HbLt. Esse fato é sustentado também pela presença das Hbs leves (Hb_L), encontradas em vestimentíferos e pogonóforos. Essas estruturas de 400 kDa são formadas apenas por subunidades de globinas, que primordialmente se associam formando a estrutura do *cap*, equivalente aos encontrados nos oligoquetas. As interfaces do tipo *a/d* e *b/c* nessas estruturas, são altamente conservadas quando comparadas à HbLt ou HbGp, indicando que os mecanismos de comunicação entre essas subunidades dever ser semelhante nessas estruturas. Nas Hb_L os *caps* estão associados em uma estrutura dimérica com formato globular que contém 24 subunidades de globinas. Um empacotamento muito similar foi encontrado para a estrutura cristalográfica do *cap* da HbLt isolado. As estruturas de Hb_L são formadas por subunidades de globinas que apresentam baixa conservação em relação às globinas que constituem a HbGp (~30% identidade entre os homólogos equivalentes), porém, a sobreposição dessas estruturas (*caps*) com a estrutura cristalográfica da HbGp resulta em valores de *rmsd* baixos (entre 1,3 e 1,6 Å). Esse fato reforça que não apenas o enovelamento local dessas estruturas (de cada subunidade) é altamente conservado, mas que ao se associarem, mantém conservada a orientação das subunidades. A **figura 3.6.8** mostra uma comparação entre a estrutura do *cap* que forma a HbGp e as estruturas equivalentes em *L. terrestris*, *O. mashikoi* (Hb_{LOm}) e *R. pachyptila* (Hb_{LRp}). As Hb_L s apresentam sítios de

ligação de metais que não são conservados para os *caps* das eritrocruorinas de oligoquetas. Na estrutura da Hb_LOm são encontrados quatro sítios de Ca²⁺, onde três estão ligados nas regiões inter- ou intra- tetrâmeros. Para Hb_LRp, são formados sítio de íons Zn²⁺ nas interfaces entre subunidades tipo *d*, sugerindo diferenças na influencia dos metais sobre essas estruturas.

Figura 3.6.8 - Comparação entre as estruturas dos *caps* da HbGp, HbLt, Hb_LRp e Hb_LOm. Apesar da baixa identidade sequencial na comparação entre as sequências que formam esses sistemas, a estrutura quaternária é altamente conservada. As Hb_L apresentam sítios de ligação de metais que não são conservados para os *caps* das eritrocruorinas de oligoquetas.

3.6.2 Estrutura da HbGp – Linkers e a formação do coiled coil.

A estrutura apresenta três subunidades *linkers* distintos para cada um dos doze protômeros que formam a unidade biológica. Os *linkers* apresentam um enovelamento muito similar ao de seus homólogos na estrutura da HbLt, e as três subunidades de *linkers*, L1, L2 e L3, apresentam a mesma estrutura básica, que é dividida em 4 domínios (**figura 3.6.9**). Os dois primeiros domínios dizem respeito as duas hélices que são conectadas por uma alça que possui tamanho variável entre os *linkers*. Uma análise dos fatores de temperatura nesta região indica que a extremidade da primeira hélice e a alça que conecta o domínio 1 e 2 são as regiões mais móveis para a estrutura do *coiled-coil*. O terceiro é o domínio rico em cisteínas, similar aos domínios LDL-A, assim como o domínio LR5 do receptor LDL humano¹⁴⁰. Uma sobreposição da estrutura do domínio LR5 humano e dos domínios LDL-A dos *linkers* da HbGp apresentam um *rmsd* de 1,3–1,4 Å para 35–37 Ca's. Da mesma forma que observado para as outras estruturas de domínios LDL-A, intensos picos no mapa da diferença (Fo-Fc) são observados, indicando a presença de íons Ca²⁺ em cada uma das subunidades de *linkers*. Vale lembrar que, os sítios de ligação de Ca²⁺ são formados por 6 oxigênios (formando um complexo octaédrico distorcido) o que é um sítio também compatível com o íon Mg²⁺. Sendo assim, embora a análise de ICP para a amostra purificada da *ciano*-HbGp exclui a presença do íon Mg²⁺, dados disponíveis na literatura mostram que o mesmo também exerce um efeito muito similar ao Ca²⁺ sobre a HbGp. A **figura 3.6.10** mostra uma comparação entre os domínios LDL-A da HbGp e HbLt comparados ao LR5 humano. O quarto e último domínio tem um típico enovelamento de barril de oito fitas antiparalelas, que desempenha uma função importante na formação do protômero, sendo um dos principais sítios de interação entre subunidades de *linkers* e globinas. O mapa de densidade eletrônica (2Fo-Fc) na região do *linker* L1 da HbGp, possui uma pequena extensão da região N-terminal, o que permitiu modelar sete resíduos a mais do que a extremidade da subunidade L1 da HbLt. Isso resultou em um prolongamento da hélice N-terminal de L1 em duas voltas. Para as subunidades tipo L2, não são observados nos mapas de densidade eletrônica os 17 primeiros resíduos da porção N-terminal, semelhante ao que ocorreu para estrutura da HbLt, onde uma porção grande da região N-terminal, rica em glicinas, não pode ser observada para a subunidade L2.

Figura 3.6.9 - Uma sobreposição dos pares de subunidades *linkers* da HbGp (cores) e HbLt (cinza). As estruturas revelam que as porções N-terminais apresentam as maiores divergências na sobreposição. Os domínios LDL-A e β -barril apresentam enovelamentos muito conservados, com as ligações dissulfeto e sítios de ligação de Ca^{2+} estruturalmente equivalentes.

O domínio β -barril da subunidade L2, apresenta uma densidade eletrônica que faz uma ponte entre os resíduos Arg130 e Arg216 situado internamente na cavidade oposta ao domínio LDL-A. Esse sítio também foi descrito no modelo de L2 da HbLt com um sítio de íon Zn^{2+} . Devido ao conflito de cargas entre o suposto Zn^{2+} e o ambiente químico, optou-se por não atribuir uma espécie química para esta região do modelo (**figura 3.5.8 - a**). A subunidade L2 foi predita como possuindo uma glicosilação no resíduo Asn 227 (posição 249 do alinhamento da **figura 3.2.2**). Porém, não existem evidências nos mapas de diferença (Fo-Fc) que tal glicosilação de fato está presente nesta subunidade (o que pode ser resultado da desordem nesta região da estrutura). Para as subunidade L3, além da confirmação da presença do íon Ca^{2+} no domínio LDL-A, foram encontradas duas outras regiões (**figura 3.6.9**) de picos intensos na densidade eletrônica, atribuídas aos íons Zn^{2+} . O primeiro sítio (sítio de zinco 1) é formado entre o domínio LDL-A e o β -barril da região C-terminal. Neste caso o íon Zn^{2+} aparece coordenado aos resíduos His64, Asp90 e His94. Ao quarto sítio de coordenação foi atribuído uma molécula de água. A **figura 3.6.11 - a** mostra o íon Zn^{2+} coordenado à estrutura, próximo ao sítio de Ca^{2+} . Com base no alinhamento (**figura 3.2.2**) que contém as sequências de *linkers* conhecidos para oligoquetas, o sítio de ligação de Zn^{2+} no domínio LDL-A mostra-se específico da espécie *G.paulistus*.

Figura 3.6.10 - Comparação entre os domínios LDL-A das subunidades *linkers* e o domínio homólogo no receptor LR-5 humano.

O segundo sítio de Zn^{2+} encontrado para a subunidade L3 (sítio de zinco 2) é em realidade um ligante interfacial, entre a subunidade L3 e a subunidade L2 do protômero vizinho dentro da mesma camada hexagonal (interações entre protômeros diferentes serão descritas adiante). O sítio é formado pelos resíduos His142 e His144 da subunidade L3 e pela His 236 da subunidade L2 (**figura 3.6.11 - b**). O quarto sítio de coordenação foi modelado como sendo uma molécula de água. Segundo Liochev e colaboradores⁹⁶, a HbLt possui atividade superóxido dismutase (SOD), o que está em pleno acordo com a presença das espécies Cu^{2+} e Zn^{2+} determinadas experimentalmente para este sistema^a. Essa atividade é atribuída única e exclusivamente as subunidades *linkers*, uma vez que a fração de globinas isoladas não apresentou atividade de SOD. Nas enzimas Cu/ZnDOS, sete resíduos participam do sítio ativo onde estão coordenados aos íons Cu^{2+} e Zn^{2+} . Seis deles são resíduos de histidinas, e o sétimo resíduo é um aspártico. Uma das histidinas é compartilhada pelos dois íons (**figura 3.6.11 - d**) tendo papel crucial no mecanismo proposto para esse sistema⁹⁷. Dessa forma os sítios de Zn^{2+} encontrados na estrutura HbGp, e o sítio presente na estrutura da HbLt, não preenchem os requisitos da formação de um sítio de CuZnSOD tal como observados nas SODs. Atividade de SOD não foi determinada para a HbGp, no entanto, considerando sua similaridade com a HbLt, para que esta venha a apresentar tal comportamento, outros sítios de metais não encontrados podem existir, preenchendo as características necessárias a formação de um sítio

^a O autor sugere que sítios em questão seriam formados por Cu e Zn assim como nas Cu/ZnSODs

de Cu/ZnSOD, e/ou de que esses sistemas obedecem mecanismos diferentes dos encontrados nas SODs.

Figura 3.6.11 - Sítios de ligação de grupos prostéticos na estrutura da HbGp. (a) sítio de zinco 1, interno ao domínio LDL-A e (b) sítio de zinco 2, interfacial entre a subunidade L2 e L3. (c) glicosilação e (d) sítio ativo da enzima Cu/ZnSOD humana. Os sítios de zinco encontrados na HbGp não são equivalentes/similares ao da enzima Cu/ZnSOD.

A densidade eletrônica na região do resíduo Asn 121 da cadeia L3 (152 do alinhamento – **figura 3.2.2**) é bastante coerente com a presença da glicosilação predita para esta posição. Este resíduo encontra-se na segunda fita que forma o domínio β -barril, onde foi atribuída uma unidade de NAG para a densidade eletrônica da região. Porém, vale mencionar que tal densidade é abrangente, sugerindo que a glicosilação pode estender-se em mais unidades de sacarídeos. A **figura 3.6.11 - c** mostra a glicosilação encontrada para a subunidade L3 do protômero 1. Com exceção para os domínios 1 e 2, que assumem diferentes orientações, é possível sobrepor as estruturas dos *linkers*, demonstrando a alta conservação estrutural das porções globulares (LDL-A e β -barril). As subunidade L1 e L2 são mais

similares, apresentando uma *rmsd* de 1.0 Å para a sobreposição das mesmas. A **tabela 3.6.2** mostra os desvios quadráticos médios para as sobreposições das subunidades de *linkers* da HbGp e HbLt.

Tabela 3.6.2 - Mostra os valores de *rmsd* resultantes da sobreposição entre as subunidades *linkers*. Valores com * indicam que não foi considerada a porção do *coiled-coil* para a sobreposição e cálculo do *rmsd*.

Tipo	HbGp L1	HbGp L2	HbGp L3	HbLt L1	HbLt L2	HbLt L3
HbGp L1	X	1,0*	1,5*	0,6	1,0*	4,5
HbGpL2	1,4*	X	1,4*	1,1	0,63	1,7*
HbLt L3	1,5*	1,4*	X	3,7	1,2*	0,5

Os *linkers* interagem entre si formando o heterotrímero. A formação desse oligômero depende majoritariamente das interações hidrofóbicas que cada *linker* faz com seu vizinho através dos domínios helicoidais do N-terminal de cada subunidade, formando então a estrutura *coiled-coil*. A região em *coiled-coil* é interrompida a cerca do resíduo 45, dividindo-a em duas porções helicoidais. O primeiro, é a hélice maior, e corresponde a cerca de quatro repetições septeto (a **figura 3.2.2** mostra a fase do *coiled-coil*). A hélice tripla do *coiled-coil* é, portanto, principalmente estabilizado por oito conjuntos de contatos hidrofóbicos (quatro na posição “a” e quatro na posição “d”, de cada um destes septetos). Não há um padrão claro de β -ramificações e/ou resíduos não ramificados nas posições “a” e “d”, o que é consistente com a formação de um trímero, ao invés de um dímero ou tetrâmero¹⁴¹. Leucina e isoleucina, são os resíduos mais abundantes, mas no primeiro nível “d”, dois resíduos hidrofílicos, Gln17 e Asn24, são observadas nas subunidades L1 e L2 respectivamente.

Figura 3.6.12 - Formação do heterotrímero de *linkers*, (a) associação das subunidades L1, L2 e L3. (b) Os principais contatos entre as subunidades que formam o heterotrímero acontecem na região do *coiled-coil*. Não há um padrão claro de β -ramificações e/ou resíduos não ramificados nas posições “a” e “d”, o que é consistente com a formação de um *coiled-coil* trimérico.

Este arranjo é estabilizado por ligações de hidrogênio, que também se estende para incluir Asn22 na posição “e” do septeto equivalente na subunidade L3. Estes resíduos hidrofílicos são conservados em sequências de outras espécies (alinhamento **figura 3.2.2**), sugerindo que eles representam uma característica comum. No C-terminal da primeira hélice, o resíduo Arg42 que ocupa a posição “d” (fase do *coiled-coil*) da subunidade L3, tem sua carga compensada pelos grupos carbonílicos vindos da dos resíduos Leu38(L1) e Ala38(L2). Esta arginina é um resíduo conservado para as subunidade L3 dos oligoquetas e seu papel na interrupção da hélice é presumivelmente de importância estrutural. No entanto, a sequência do *linker* disponível para *A. marina* (poliqueta) apresenta nesta região um padrão contínuo de septetos, consistente com a presença de uma hélice contínua observada no modelo cristalográfico de baixa resolução da HbAm⁶¹. A interrupção do *coiled-coil* pode, portanto, ser necessária para a determinação do deslocamento relativo dos protômeros de um disco em relação ao outro, o que é a principal diferença observada entre as estruturas do tipo I, tais como partícula de *G. paulistus* aqui apresentada, e do tipo II como vistas em *A. marina*. Resíduos de ácido aspártico são encontrados em todos os três elementos de ligação no início

da segunda hélice. No caso das cadeias de L3 e L2, estes são bem conservados e formam clássicos *N-caps* (onde a cadeia lateral do resíduo Asp (posição n) interage com o nitrogênio da cadeia principal do resíduo n+3). A segunda hélice é mais curta e composta apenas por duas repetições septeto. No primeiro nível, as hélices são espaçadas de tal modo que o quarto resíduo (Phe46 (L2) que ocupa uma posição “g”) contribui para o núcleo hidrofóbico no interior do *coiled-coil*. Na extremidade C-terminal, o *coiled-coil* é estabilizado por uma clássica interação salina, com Arg (em “g”) - Glu (em “e”). O *coiled-coil* representa ~90% da superfície de contato entre os *linkers* na formação do heterotrímero. Cada *linker* também realiza um pequeno contato através de seu domínio LDL-A com o domínio β -barril do *linker* adjacente. As regiões de contato entre os domínios globulares não são conservados e não apresentam um padrão de complementariedade (como resíduos de subunidades diferentes que realizam ligações de hidrogênio ou pontes salinas) com exceção da interação entre os resíduos Asp72 de L1 com Arg 196 e His216 em L3. Isso sugere que a interação entre esses domínios é uma consequência da formação do heterotrímero e é majoritariamente guiada pela formação de *coiled-coil*.

Figura 3.6.13 - Sítios de interação entre as porções globulares dos *linkers* no heterotrímero. Os sítios de interação ocorrem através dos mesmos domínios, mas não há conservação dos resíduos envolvidos. Esses fracos contatos sugerem que a porção *coiled-coil* é a que de fato atua mantendo as subunidades dos *linkers* unidos.

3.6.3 Formação do protômero

O protômero é formado por uma estrutura dodecamérica de globinas (*cap*) e um heterotrímero de *linkers*, produzindo uma estrutura com pseudo simetria de ordem três. Os contatos entre o *cap* e heterotrímero de *linkers* acontecem pelas bordas externas da cabeça do heterotrímero com a borda interna do *cap*. Em consequência do formato de cúpula que o *cap* assume, é formada uma grande cavidade dentro da estrutura do protômero, que para a HbGp não apresentou densidades eletrônicas que sugerissem a presença de moléculas pequenas ligadas nesta região (**figura 3.6.14**).

Figura 3.6.14 - O protômero é uma associação da estrutura dodecamérica de globinas (*cap*) com o heterotrímero de *linkers*. Na junção dessas duas estruturas oligoméricas forma-se uma grande cavidade com volume de $39.207 \text{ \AA}^3$, o que equivale ao volume de duas subunidades de globina.

Os domínios LDL-A e β -barril de cada *linker* interagem com o trímero de globinas [$a_1b_1c_2$] (unido por ligações dissulfeto). A interação entre o heterotrímero de *linkers* e o *cap*, resulta em uma área total de contato calculada de $3.655 \text{ \AA}^2$. O domínio beta barril na porção C-terminal dos *linkers* realiza a maior área de contato entre o heterotrímero e a estrutura do *cap*. Isso equivale a $\sim 75\%$ da superfície que os *linkers* compartilham (os outros 25% referem-se ao domínio LDL-A). Para o dodecâmero de globinas, as subunidades tipo *b* contribuem com

cerca de 60% da área de superfície de contato, seguida das subunidades tipo *a*, que contribuem com cerca de 25% e por fim, subunidades tipo *c*, contribuindo com de cerca de 15% da área de superfície de contato. As subunidades do tipo *d* não participam do contato, estando unidas em torno do eixo de ordem três na borda superior externa da partícula. Apesar da simetria de ordem 3 apresentada pelo *cap*, o heterotrímero de *linkers* apresenta algumas diferenças nos resíduos de contato, mostrando que a interface entre as subunidades do *cap* e os *linkers* não é conservada (**figura 3.6.15**). Um exemplo desta variabilidade envolve os contatos entre o domínio β -barril e as subunidade tipo *b*. Nas subunidades tipo L1, um agregado de resíduos aromáticos presentes nos *loops* que conectam as fitas 2 e 3 do domínio β -barril, prevalecem nas interações entre L1 e *b*. Esses resíduos são F140, F141, F170, e W146, no entanto, apenas as posições F140 e F14 são conservadas em L2 e L3. A dissociação dos *caps* em experimentos de espectrometria de massas já foi reportada para eritrocruorinas de várias espécies. Aparentemente a estrutura dimérica entre os dois *caps* que formam a Hb_LRp (*R.pachyptila*) é mais estável do que a interação entre os *caps* e os *linkers* das eritrocruorinas. Um possível efeito dessas interações mais fracas entre *caps* e *linkers* nas eritrocruorinas é de que estas permitem um maior grau de liberdade nas alterações de estrutura quaternária em virtude das transições alosterias desses sistemas. Isso é consistente com o fato de que a Hb_LRp apresenta um efeito Bohr reduzido comparada com a eritrocruorina de *R.pachyptila* (Hb_pRp). (*Obs: vestimentíferos possuem os dois tipos de hemoglobinas, ver introdução página - 42*).

Embora a contribuição da área de contato entre a subunidade *b* e o domínio LDL-A seja menor do que a interação desta globina com o domínio β -barril, uma interação mostra-se potencialmente importante, evidenciando a influência dos *linkers* sobre o comportamento apresentado pelas globinas. O resíduo Arg 34 da subunidade *b*, (hélice B – posição 38 do alinhamento, **figura 3.2.2**), realiza uma ponte salina com o resíduo Asp88 da subunidade L1, e influencia a orientação do resíduo Asp84 (posição 113 do alinhamento, **figura 3.2.2**) que são resíduos diretamente coordenados aos íons Ca^{2+} . Vale lembrar que, o resíduo Trp33 (vide **figura 3.6.15**) da subunidade *b* é um dos 4 resíduos hidrofóbicos que cercam o sexto sítio de coordenação do grupo heme. Dessa forma, o resíduo Asp88 da subunidade L1 e seus equivalentes nos outros *linkers*, são capazes de exercer certa influência no comportamento da hélice B das globinas tipo *b*, via interação iônica com o resíduo Arg34. Essa interação é conservada não apenas para os três *linkers* da HbGp, mas também esta presente na estrutura da HbLt.

Figura 3.6.15 - Sítios de interação entre as porções globulares dos *linkers* no heterotrímero. Os sítios de interação ocorrem através dos mesmos domínios mas não há conservação dos resíduos envolvidos nesta região. A interação entre o domínio LDL de cada *linker* com a subunidade tipo *b* ocorre via a ponte salina formada entre os resíduos Arg34 da subunidade *b* com os resíduos Asp 88-L1/ 95-L2 / 85-L3 das subunidades *linkers*.

Como mencionado anteriormente, a estrutura do *cap* é sugerida como o “caroço” alostérico da HbLt (e é razoável pensar que o mesmo seja verdade pra a HbGp), exibindo cooperatividade igual ao da partícula inteira em pH 6,8. Porém, o comportamento alostérico do *cap* e da partícula inteira divergem para outras condições de pH, o que indica que as propriedades

alostéricas desses sistemas também são moduladas pela ação dos *linkers* e ou contatos entre *caps* de protômeros diferentes (ver adiante – interações entre protômeros da mesma camada hexagonal). Na estrutura da HbLt, a interface entre os *linkers* L1 e subunidades tipo *c*, apresentam um pico de densidade intenso, que foi atribuído a íons de Ca^{2+} . Porém o sítio presente nas globinas tipo *c*, responsável pela ligação direta com íon, não é conservado para a globinas tipo *c* da HbGp. Dessa forma, nem a sequência, nem a densidade eletrônica, sugerem a presença de íons dessa espécie na interface entre os *linkers* e globinas da HbGp. A cavidade formada na interface entre o heterotrímero de *linkers* e o *cap*, possui um volume de $39.207 \text{ \AA}^3$ (**figura 3.6.14**), que é aproximadamente o dobro de volume ocupado por uma subunidade de globina isolada ($\sim 20.000 \text{ \AA}^3$). Ao todo são encontradas 13 canais de acesso ao interior do protômero, divididos aqui em 5 grupos de acordo a localização. O primeiro grupo de canais de acesso é em realidade um único canal no topo da estrutura do *cap*, na interface entre as subunidades *d*, onde se encontra o eixo de simetria de ordem 3. Essa imperfeição no empacotamento das subunidades *d* gera um canal com diâmetro aproximado de $6,5 \text{ \AA}$. O segundo grupo corresponde aos canais formados na interface entre tetrâmeros *abcd* vizinhos, com aproximadamente $7,0 \text{ \AA}$ de diâmetro, envolvem as subunidades a_1 e d_1 e c_2 e d_2 . Na estrutura da HbLom esse canal é ocupada por um sítio de ligação de Ca^{2+} que não é conservado na HbGp/HbLt. O terceiro grupo é formado internamente nas subunidades tipo *linkers* (interação entre domínio da mesma subunidade), e são consequência de contatos envolvendo o segundo domínio helicoidal, domínio LDL-A e o domínio β -barril. O quarto grupo de canais de acesso à cavidade é formado na interface do heterotrímero de *linkers*. A interação entre o β -barril de uma subunidade de *linker* com o domínio LDL-A da subunidade *linker* adjacente, resulta em pequenos espaços que dão acesso à cavidade central. Os canais de acesso do grupo 3 e 4 possuem um diâmetro médio de $10 \text{ \AA}$. O quinto e último grupo de canais de acesso é fruto da interação entre o *cap* e o heterotrímero de *linkers*. Esse canal encontra-se na interface entre o dímero de globinas *bc* e a interface entre o β -barril de um dado *linker* com o domínio LDL-A do *linker* adjacente. São os maiores canais de acesso ao interior da cavidade com diâmetros entre $13 - 16 \text{ \AA}$. **Afigura 3.6.16** mostra os canais de acesso à cavida de acordo com sua localização na estrutura do tetrâmero. Não existe função atribuída à cavidade formada na interação entre o *cap* e o heterotrímero de *linkers*, porém, com base no volume total que esta ocupa e dos vários canais de acesso, é bastante razoável consideramos que esta cavidade pode ter função carreadora.

Figura 3.6.16 - Os cinco grupos de canais de acesso a grande cavidade central formada pela interação entre a estrutura do cap e o heterotrímero de *linkers*. (a) canais de acesso formados na estrutura do cap (grupos 1 e 2). (b) canais de acesso formados no heterotrímero de *linkers* (grupos 3 e 4). (c) canais de acesso formados pela interação entre a estrutura do cap e o heterotrímero de *linkers*, grupo 5.

3.6.4 Estrutura da HbGp – *Interação entre protômeros de uma mesma camada hexagonal.*

As subunidades *linkers* proporcionam os contatos primários entre os protômeros que formam a unidade biológica das eritrocruorinas. O contato lateral entre protômeros vizinhos de um mesmo disco inclui interações entre as subunidades dos *linkers* L2 e L3, através do domínio β -barril e também envolve as globinas do *cap* de cada protômero. A contribuição dos *linkers* para o contacto lateral envolve as ligações de hidrogênio proporcionadas pela cadeia principal de duas fitas beta antiparalelas; o início da fita 4 da subunidade L3 interage com a extremidade carboxílica da fita 2 de L2 do protômero vizinho (**figura 3.6.17 - b**) tornando as fitas *beta* envolvidas no contato lateral mais extensas para as subunidades L2 e L3. Também contribuem para este contato os resíduos no início do domínio β -barril da subunidade L2, incluindo Lys 113 que forma uma ponte salina com o resíduo Glu 159 da subunidade L3. Tanto o compartilhamento das fitas beta como a interação iônica entre as cadeias laterais mencionadas anteriormente são conservadas na estrutura da HbLt. Porém, a estrutura da HbGp apresenta nesta interface a formação do sítio de zinco 2. Os resíduos His 142 e His 144, presentes na fita 3 da subunidade L3, combinados com a His 236 no C-terminal da subunidade L2, formam o sítio de Zn^{2+} que faz uma ponte entre protômeros. Até então, acreditava-se que os protômeros estavam unidos apenas por interações envolvendo a parte peptídica da estrutura, como interações hidrofóbicas e interações polares (ligações de hidrogênio e pontes salinas). Porém, a presença do íon metálico na interface entre as subunidades L2 e L3 dos protômeros de uma mesma camada hexagonal adiciona uma interação importante. O Zn^{2+} é um metal descrito por exercer influência sobre a afinidade dessas estruturas de uma forma diferente dos metais alcalinos terrosos. Segundo Ochiai e colaboradores⁸⁶ na espécie *Pheretima hilgendorfi*, o Zn^{2+} liga-se a um sítio diferente dos descritos para os metais alcalinos terrosos, e atua aumentando a afinidade forma tensionada pelo oxigênio. O zinco também é apontado como um agente capaz de aumentar a estabilidade da partícula em condições que levam a fragmentação do complexo, com pH alcalino por exemplo.

Nos contatos laterais entre protômeros envolvendo os *caps*, duas interações entre subunidades de globinas são formadas; a primeira envolve subunidades do tipo *a*. Neste caso, a subunidade *a* do primeiro protômero interage, através das hélices B e F, com o curto *loop* que conecta as hélices B e C e começo da hélice G da subunidade *a* do segundo protômero

(figura 3.6.17 - c). Nesse contato, prevalecem interações polares, como para ligação de hidrogênio formada entre His42 (hélice B) da subunidade *a* do primeiro protômero com Gln107 (hélice G) da subunidade *a* do segundo protômero. Segundo os fatores de temperatura refinados, o final da hélice F e começo da hélice G, são as regiões que apresentam grande mobilidade na estrutura dessa subunidade. Os contatos polares envolvendo o resíduo Gln107 (hélice G) podem influenciar na mobilidade dessas hélices, e por consequência, exercer influência sobre o ambiente químico do grupo heme (uma vez que a histidina proximal pertence à hélice F).

Figura 3.6.17 - Interações entre protômeros de uma mesma camada hexagonal. (a) Estrutura da monocamada hexagonal onde metade dos *caps* foram removidos, expondo a porção globular dos *linkers* de cada protômero. (b) Interação entre protômeros vizinhos envolvendo a porção globular dos *linkers* L2 e L3. Além de compartilharem as fitas que constituem o β -barril a interface forma um sítio de ligação de zinco (sítio de zinco 2). (c) Contatos entre os *caps* de protômeros vizinhos.

O segundo contato envolvendo os *caps* de diferentes protômeros ocorre através de uma subunidade tipo *b* com uma subunidade tipo *c* do protômero vizinho. Tal interação envolve o *loop* que conecta as hélices F e G da globina *b*, com o N-terminal da hélice G e C-Terminal da hélice H, da globina tipo *c*. Nesta interação podemos destacar a ponte salina entre os resíduos Asp 102 da subunidade *b* e Lys 108 da subunidade *c*. Resumindo, há contatos importantes entre *caps* da mesma camada hexamérica envolvendo partes das cadeias globinas próximas aos grupos heme. Existe portanto a possibilidade de comunicação entre globinas de protômeros diferentes que poderiam estar envolvidos em cooperatividade, uma vez que a partícula inteira apresenta um comportamento diferente do *cap* isolado.

3.6.5 Estrutura da HbGp – *Interação entre protômeros de camadas hexagonais diferentes.*

Os contatos mais expressivos entre protômeros de camadas hexagonais diferentes ocorrem entre protômeros parcialmente sobrepostos (uma vez que as bicamadas hexagonais possuem uma rotação relativa de 16° entre si). Esses protômeros estão relacionados por um eixo de ordem 2 (que representa uma das famílias de eixos de ordem dois que a simetria 622 possui – **figura 3.5.1 - c**). Essa interação envolve a região globular dos *linkers* L1 e as hélices do primeiro domínio dos *linkers* L2 e L3. É importante mencionar que não existem interações entre globinas, de forma que as duas camadas hexagonais interagem unicamente através das subunidades *linkers*. A interação entre os domínios β -barril das subunidades L1 resulta em uma superfície de contato de 1102 Å² com a formação de 16 ligações de hidrogênio. Esse contato inclui a participação das fitas B1, B2, B3, B4 e B5 e os *loops* nas redondezas **figura 3.6.18 - c**. A interação entre as estruturas *coiled-coil* incluem as subunidades L2 e L3 dos dois protômeros, formando um ângulo de 100° entre os dois feixes de hélices. Essa interface é caracterizada por uma alta concentração de resíduos carregados que são complementares. Devido ao eixo de simetria que relaciona os dois protômeros, estas interações são duplicadas resultando em seis pontes salinas. A **figura 3.6.18 - d**, mostra os resíduos envolvidos nas interações entre os *coiled-coils* de protômeros parcialmente sobrepostos. A última interação característica entre protômeros de diferentes camadas hexagonais ocorre entre um dado protômero com o protômero adjacente ao sobreposto da outra camada hexagonal.

Figura 3.6.18 - Interação entre protômeros de camadas hexagonais diferentes. Para melhor visualização, parte das estruturas dos *caps* foram removidas. (a) visão frontal (b) visão lateral. (c) Interações entre os protômeros parcialmente sobrepostos, que ocorrem através do domínio β -barril das subunidades L1 e (d) através das porções *coiled-coil* da subunidades L2 e L3. (e) Interação entre um protômero com o protômero adjacente ao sobreposto da outra camada hexagonal. Este contato ocorre através dos resíduos Thr12, Gln17 e Asn16, presentes no primeiro domínio das subunidades L1.

Essa interação ocorre entre subunidades tipo L1, através dos resíduos Thr12, Gln17 e Asn16, presentes no primeiro domínio. O contato gera um ângulo de 97° entre as hélices, muito próximo ao observado para o contato mencionado anteriormente para as hélices de L2 e L3. As duas subunidades estão relacionadas pela segunda família de eixos de ordem 2 que a partícula inteira possui. Assim, os três resíduos de cada subunidade atuam de forma complementar, resultando em um contato com a formação de 4 ligações de hidrogênio (**figura 3.6.18 - e**) com destaque para o par de ligações de hidrogênio que envolvem os resíduos Asn16. Neste caso, o nitrogênio δI da cadeia lateral e o oxigênio da carbonila do resíduo simétrico formam a interação. O resultado é que os *linkers* tem um papel fundamental em estabilizar a partícula como um todo e especificamente em manter os dois discos hexagonais juntos. O *coiled-coil* de cada protômero somente faz contatos com *coiled-coils* da outra camada e de uma forma interdigitada. Ou seja, cada *coiled-coil* fica intercalado entre outros dois *coiled-coils* provindo da outra camada, o protômero sobreposto e o seu vizinho. Dessa forma os *coiled-coils* estão costurados em uma rede que dá sustentação ao complexo. Essa interação pouco ou nada depende das globinas (já que a estrutura base da bicamada hexagonal sem as globinas foi observada em experimentos de microscopia eletrônica - dados coletados no LNNano e não reportados na literatura).

4 RESULTADOS E DISCUSSÃO – HBGP SUBUNIDADE D

4.1 Purificação

Uma amostra da *oxi*-HbGp foi dissociada em solução alcalina e a fração monomérica, a *d*-HbGp, foi purificada por gel filtração (Sephadex G 200). A fração contendo a *d*-HbGp é eluída e coletada entre 70 e 85ml. Esse resultado está em concordância com os obtidos por Cabral e colaboradores¹⁰⁴. Um espectro de UV-vis comprova a predominância da forma *oxi* para a *d*-HbGp purificada. A **figura 4.1.1** mostra a análise de SDS PAGE do volume coletado através do processo de gel filtração e o espectro de UV-vis que caracteriza essa amostra como *oxi-d*-HbGp.

Figura 4.1.1 - Resultados da purificação da subunidade *d* da HbGp (*d*-HbGp). (a) Gel filtração, com destaque para o pico coletado entre 70-85 mL que contém a *d*-HbGp. (b) SDS PAGE, 1 – marcador de massa molecular, 2 – HbGp na forma íntegra e 3 – Fração coletada entre os volumes 70 e 85 de gel filtração contendo a *d*-HbGp. (c) Análise espectroscópica da *d*-HbGp isolada, as bandas Q evidenciam que a forma reduzida (*oxi*) é a que predomina na amostra.

4.2 Cristalização

Os cristais de *d*-HbGp foram inicialmente obtidos na seguinte condição: citrato de sódio 1,6M Tris/HCl 0,1M pH 7,5. O tempo de crescimento médio para esses cristais foi de aproximadamente 24 horas. Uma otimização da condição inicial, variando-se a concentração de citrato de sódio, produziu cristais maiores e mais bem formados na condição: citrato de sódio 1,4M Tris/HCl 0,1M pH 7,5 (**figura 4.2.1**). Apesar da dimensão média reduzida (0,05mm), os cristais da *d*-HbGp não apresentaram problemas com relação à manipulação, sendo facilmente coletados do ambiente de cristalização e transferidos para a solução crioprotetora, onde não apresentaram danos físicos aparentes em virtude do choque osmótico. A solução crioprotetora mais eficiente para este sistema foi: 1,4M Tris/HCl 0,1M pH 7,5 10% etilenoglicol.

Figura 4.2.1 - Cristais obtidos para a *d*-HbGp na condição 1,4 citrato de sódio, Tris/HCl pH7,5, com dimensões próximas a 0.5μm.

4.3 Coleta e processamentos dos dados

Os cristais de *d*-HbGp apresentaram padrões de difração com reflexões bem definidas e intensas. O conjunto coletado apresenta alguns sinais de anéis de gelo formados durante o processo de crio proteção, mas que não comprometeram os dados de difração obtidos. Ao todo foram coletadas 300 imagens com um intervalo de 0.4° por imagem (**figura 4.3.1**, uma

das imagens obtidas para o conjunto de dados da *d*-hbGP). O grupo espacial predito no processo de indexação foi o I222 (não excluindo a possibilidade de I2₁2₁2₁), e uma mosaicidade inicial estimada de 0.65°. Durante o processo de integração, apenas os parâmetros de cela foram fixados. Não houve oscilações significativas para os parâmetros de mosaicidade e centro de feixe. Um total de 214.504 reflexões foram integrados pelo programa Imosflm.

Figura 4.3.1 - Imagem de difração do cristal da *d*-HbGp com resolução de 2.1 Å na borda, 0,4° de oscilação e mosaicidade inicial estimada de 0.64°. A imagem apresenta alguns pontos incoerentes com a predição, resultados da presença de um cristal satélite (pequeno cristal, ou fragmento de, anexado ao cristal que está sendo coletado) mas que não comprometeu a qualidade dos dados obtidos.

O conjunto de dados possui um total 7.937 reflexões únicas, das quais 100 foram rejeitadas na fase de redução dos dados (uma média de 0,3 reflexões rejeitadas por imagem). A resolução obtida foi de 2.10 Å na ultima faixa de resolução com uma completude de 99.1% e multiplicidade de 4,1. O coeficiente de Mathews determinado para esse sistema aponta um conteúdo de solvente de apenas 37,5%. Esse valor está nos limites mínimos de solvente estimado para um cristais de proteínas¹⁴³. Diferente do conjunto de dados coletados para a *ciano*-HbGp, que sofre uma queda significativa na intensidade das reflexões a partir de 7.0Å, o conjunto de dados para a *d*-HbGp mostrou um suave queda na intensidade até a região de

alta resolução, o que está em pleno acordo com o baixo conteúdo de solvente estimado para o cristal. A **tabela 4.3.1** mostra os principais parâmetros avaliados durante o processo de integração e redução dos dados (escalonamento). Para o cálculo do parâmetro R_{free} foram separadas aleatoriamente 10% das reflexões únicas do sistema.

Tabela 4.3.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução.

Grupo espacial	I222
Dimensões de cela a, b, c (Å)	53,10, 63,13, 78,36
Detector	Pilatus 2M
Fonte de raios-X	I04-1
Comprimento de onda (Å)	1,00
Resolução	31,57(2,10)
Multiplicidade	4,2 (4,3)
Rmerge(%)†	2,6 (41,7)
Completeza	99,1 (99,8)
Reflexões totais	32.296 (4.887)
Reflexões únicas	7.891 (1.125)
Intensidade (I/σI)	24,6 (3,1)
B-fator no Wilson plot (Å ²)	29,21

4.4 Substituição molecular e refinamento do modelo

O problema das fases foi resolvido por substituição molecular utilizando-se a subunidade *d* (na qual foi modelada a sequência *dI*) da estrutura cristalográfica da HbGp como molde de procura. Esse processo foi realizado no programa Phaser, onde foi feita uma busca exaustiva pela solução do problema nos dois possíveis grupos especiais (I222 e I2₁2₁2₁). O programa Phaser apontou o grupo espacial I222 como sendo a solução correta para esse sistema (LLG = 2413, R = 36.22%). O grupo heme foi excluído neste processo, e a região a qual ocupa foi usada como critério para avaliar a correta atribuição das fases por substituição molecular. Um mapa de densidade eletrônica (2Fo-Fc) interpretável foi obtido e detalhes como a ligação dissulfeto e o grupo heme foram facilmente identificados. A unidade assimétrica conta com apenas uma única subunidade. Uma questão relevante levantada durante o processo de refinamento foi a possibilidade dos mapas obtidos serem igualmente

adequados para a sequência *d2*. Dessa forma, as duas sequências foram modeladas sobre a densidade eletrônica e então refinadas. O refinamento não deixou dúvidas sobre a melhor concordância dos dados com a sequência *d1*. Os valores de *R* e *R*_{free} obtidos ao final do refinamento são iguais a 20,32 e 24,53, para a sequência *d1* e 28,14 e 33,36 para o modelo contendo a sequência *d2*. Não apenas os valores de *R* e *R*_{free} foram decisivos para a escolha da sequência correta, mas tal diferença também se manifestou radicalmente nos mapas de diferença calculados (*F*_o – *F*_c), onde a grande maioria das trocas de resíduos que acontece na comparação entre as duas sequências, perturba o mapa. Podemos destacar as posições 22, 50, 75, 96, 107 e 110 como as mais significativas. A **figura 4.4.1** mostra as densidades eletrônicas nas posições mencionadas anteriormente, destacando regiões onde a sequência *d2* é incoerente com os mapas calculados. Os principais dados obtidos para o processo de refinamento podem ser observados na **tabela 4.4.1**.

Figura 4.4.1 - Algumas das posições importantes nos mapas de densidade (cinza – $2F_o - F_c$ e $\delta = 1.0$, vermelho – $F_o - F_c$ e $\delta = -3.0$, verde – $-F_o - F_c$ e $\delta = 3.0$) comprovando que a sequência *d1* melhor descreve o modelo. As fases dos mapas em questão foram calculadas com a sequência *d2*.

Tabela 4.4.1 - Estatísticas do refinamento e geometria do modelo.

Reflexões usadas no refinamento	
R e R_{free}	20,32 e 24,53
B-fator médio.	24,4
No. de átomos (proteína)	2245
No. de átomos (ligantes)	117
Ramachandran plot	
Região permitida (%)	97,83
Região adicionalmente permitida (%)	2,17
Outliers (%)	0
R.m.s. bond lengths (Å)	0,01
R.m.s. bond angles (°)	1,08

Os parâmetros impostos para a regularização de geometria foram bastante satisfatórios. Foram obtido baixos valores de *rmsd* para os comprimentos de ligações e ângulos de ligações. Nenhum resíduo encontra-se fora das regiões permitas do diagrama de Ramachandran, e a distribuição dos fatores de temperatura também encontra-se dentro dos valores médios esperados para uma estrutura com esta resolução¹⁴⁷(figura 4.4.2).

Figura 4.4.2 - Principais parâmetros que validam a qualidade do modelo final obtido. (a) diagrama de Ramachandran, onde nenhum resíduo encontra-se fora das regiões permitidas, e (b) Uma comparação dos valores obtidos ao final do refinamento com outras estruturas depositadas no banco de dados PDB. Os valores em cor vermelha referem-se ao mínimo é máximo encontrado para o dado parâmetro (considerando a faixa de resolução em questão). A estrutura da *d*-HbGp apresenta valores dentro dos considerados aceitáveis em todos os aspectos em questão.

A densidade eletrônica na região do grupo heme é bastante clara, evidenciando que o ligante ocupando o sexto sítio de coordenação do ferro é de fato uma molécula de oxigênio. A **figura 4.4.3** mostra o mapa de densidade eletrônica ($2F_o - F_c$) calculado para a estrutura cristalográfica da *d*-HbGp na região do grupo heme, onde foi modelado a molécula de oxigênio.

Figura 4.4.3 - Mapa de densidade eletrônica calculado ($2F_o - F_c$) para o modelo da *d*-HbGp. (a) visão geral, (b) região do grupo heme onde foi modelada a molécula de oxigênio ocupando o sexto sítio de coordenação do ferro. (c) ligação dissulfeto intramolecular.

4.5 Estrutura da subunidade *d* e seus contatos cristalinos.

A subunidade *d* isolada apresenta um enovelamento global muito próximo do observado na estrutura do complexo da HbGp, preservando todos os elementos de estrutura secundária. As maiores diferenças encontram-se nas conformações dos *loops* que conectam as

hélices A e B, e G e H. Na Os *loops* em questão não participam de contatos importantes. O *loop* que conecta A e B aponta para o interior da grande cavidade formada na junção do heterotrímero de *linkers* e a estrutura do *cap*, e o *loop* que conecta G e H faz um contato hidrofóbico não conservado com a subunidade *c* no heterotetrâmero. No empacotamento cristalino da *d*-HbGp, o rearranjo destes *loops* ocorre na região do eixo de ordem 2 coincidente com o eixo *c* do cristal. Não existem interações nesta região que indicam o favorecimento de tal contato, em realidade, tal conformação leva a uma orientação de cargas de mesma natureza, aproximando o resíduo Glu 21 e seu simétrico relacionado. Uma área total de $421.3 \text{ \AA}^2$ é compartilhada nesta interação, com uma energia predita (servidor PISA) para o contato de apenas 2,8 kcal/mol (menos do que uma ligação de hidrogênio) sugerindo que este contato é meramente um artefato de cristalização. Dessa forma, a mudança conformacional dos *loops* é provavelmente um efeito único do empacotamento cristalino. A sobreposição da estrutura da *d*-HbGp com a subunidade *d* presente na HbGp, apresenta um *rmsd* de 0,44 Å. Este valor é muito próximo ao valor estimado para a sobreposição de diferentes estruturas de uma mesma proteína¹³³. Os contatos cristalinos para a estrutura da *d*-HbGp podem ser vistos na **figura 4.5.1**. Aparentemente, essas estruturas (subunidades de globinas das eritrocruorinas) são sensíveis a pequenas alterações conformacionais dentro do bolsão que envolve o grupo heme^{145,146}, em especial, envolvendo os resíduos aromáticos volumosos na posição 37 do alinhamento – **figura 3.2.2**. Embora exista uma considerável diferença na resolução dos modelos obtidos (HbGp e *d*-HbGp) tal análise deve ser interpretada com cautela, uma comparação entre os resíduos de primeira esfera de coordenação do grupo heme para as subunidades *d* e subunidade *d* isolada (*d*-HbGp) mostra que existe um pequeno rearranjo dos resíduos que fazem parte do sítio hidrofóbico. Na estrutura da *d*-HbGp existe um deslocamento de aproximadamente 0,5 Å do resíduo Trp32 em direção ao grupo heme, tornando o sítio mais fechado em relação à estrutura do complexo. Este é apontado como um resíduo chave, capaz de afetar radicalmente a afinidade pelo oxigênio, restringindo o acesso e a mobilidade do ligante dentro do sítio. O pequeno deslocamento do resíduo Trp também reflete na orientação das ligações de hidrogênio próximas que formam a hélice B de forma que parte da hélice é ligeiramente deslocada. Nas globinas *d* da HbLt o plano do anel imidazólico da histidina proximal forma um ângulo de 76-80° com o plano do grupo heme (em contraste com o que ocorre nas hemoglobinas de vertebrados onde ambos são perpendiculares).

Figura 4.5.1 - Principais contatos cristalinos encontrados para a estrutura cristalográfica da *d*-HbGp. (a) ao longo do eixo **a**, (b) ao longo de eixo **b**, formando o dímero *d/d*, que é equivalente ao dímero *a/d* na estrutura da HbGp e (c) ao longo do eixo **c**, que resulta no rearranjo dos loops que conectam as hélices A e B, e G e H.

Na estrutura da *d*-HbGp essa mesma conformação é encontrada, porém, com um ângulo de 70°. Segundo Strand e colaboradores⁸⁹, essa variação angular em relação ao plano do heme seria potencialmente diminuidora da afinidade pelo ligante, se esta no caso, atuasse

restringindo a capacidade do ferro em mover-se para fora do plano do anel. No entanto, segundo Vidugiris e colaboradores⁶³, uma investigação por espectroscopia Raman de ressonância, envolvendo a troca do ligante na HbLt, sugere que não existem variações nas distâncias entre o nitrogênio da histidina proximal e o átomo de ferro do grupo heme. Isso aponta que os mecanismos de regulação para esses sistemas obedecem a um mecanismo diferente do determinado para os vertebrados.

Figura 4.5.2 - Sobreposição da *d*-HbGp com a subunidade *d* presente na unidade biológica. (a) Existe grande coerência na sobreposição dessas duas estruturas com uma pequena diferença para a posição do resíduo Trp 32, que para a forma *oxi* da *d*-HbGp sofre um pequeno deslocamento. (b) Orientação do plano do anel imidazólico do resíduo de His94 em relação ao plano do grupo heme comparado com a Hb humana em (c)

A subunidade *d* da HbLt (e tudo indica que o mesmo seja verdade para a subunidade *d* da HbGp) não possui cooperatividade quando isolada¹³⁶. Uma das razões é porque em solução, a mesma está presente majoritariamente na forma de monômero. No entanto, uma pequena porção da *d*-HbGp em solução associa-se em oligômeros. Esse estado é enriquecido com o aumento da concentração. Uma das sugestões é de que essa pequena fração oligomérica poderia estar associada na forma de um trímero, que atuaria como um agente nucleador, agregando outras globinas para a formação do *cap*. Porém os contatos cristalinos

da estrutura da *d*-HbGp não levam a formação de estruturas triméricas relacionadas por simetria de ordem três (devido a própria limitação das operações de simetria do cristal). Na estrutura cristalográfica da *d*-HbGp, uma única subunidade *d* forma a unidade assimétrica e um dos contatos cristalinos gerados leva a formação de uma estrutura dimérica, *d/d*, cuja interface é equivalente ao dímero *a/d* (interagindo através das hélices E e F) encontrada na estrutura do complexo. A área estimada para a interface *d/d* é de 646,8 Å², menor do que a apresentada pelo do dímero *a/d*, que é igual a 914,5 Å². Porém, as duas subunidades possuem energias de interação próximas, estimadas de -11,0 e -13,7 kcal/mol respectivamente. Por sua vez, a estrutura trímica, *d_x3*, presente na partícula inteira, possui uma interface entre subunidades adjacentes de 331,4 Å² e uma energia de interação estimada em -3,6 kcal/mol. Esses dados, combinados com as interfaces encontradas na rede cristalina do modelo da *d*-HbGp apontam que a existência da espécie trimérica em solução é desfavorecida, sugerindo então que as formas oligoméricas predominantes seriam dímeros ou múltiplos de dímeros. Isso torna remota a possibilidade de que a estrutura do *cap* seria formada a partir da estrutura *d_x3*, onde esta atuaria orientando as outras subunidades, ou oligômeros pré-formados, a se associarem na estrutura com simetria correta. Dessa forma, o mecanismo de montagem do *cap* envolveria primeiramente a formação do tetrâmero *abdc*, que posteriormente se associa na forma trimérica. Outra possibilidade é que a partir dos trímero *abc*, é formada uma estrutura *(abc)₃* onde as subunidades *d* seriam então incorporadas posteriormente. A existência do arranjo dodecamérico sugere que os contatos feitos pelas demais cadeias globina são suficiente para estabilizar os frágeis contatos entre cadeias *d* em torno do eixo de ordem 3. É fato que, as globinas que constituem a HbLt tem capacidade de se reassociarem após a fragmentação do complexo em pH alcalino. Porém, a estrutura dos *linkers* que formam a base da bicamada hexagonal da HbGp, foi observada separadamente em experimentos de microscopia eletrônica, o que sugere que esta estrutura poderia atuar como um agente nucleador, agregando as estruturas dos *caps* para a formação do complexo inteiro. Isso mostra mais uma vez que as interfaces de proteínas são muito vezes promíscuas, e altamente dependentes do ambiente químico em questão. A **figura 4.5.3** mostra as interfaces das estruturas oligoméricas em que a subunidade *d* esta envolvida.

Figura 4.5.3 - O Mecanismo mostra as interfaces em que a subunidade d esta envolvida. Na estrutura do *cap*, o trímero d_3 faz contatos através das hélices B e G. Quando a subunidade d é isolada, seus frágeis contatos se desfazem, prevalecendo em solução a forma monomérica. Com o aumento da concentração, a d -HbGp se associa formando um dímero, que interage através das hélices E e F (incluindo o grupo heme). Essa interface é equivalente ao dímero a/d encontrado na estrutura do *cap*.

O dímero d/d encontrado para a estrutura da d -HbGp possui uma interface presente em outras hemoglobinas, como encontrada em *Scapharca inaequalvis*. Esse molusco possui

uma Hb composta por dois tipos de subunidades (*a* e *b*), onde a estrutura quaternária é um tetrâmero, equivalente a estrutura do tetrâmero *abcd* que forma o *cap*. Por esse motivo, a Hb de *S. inaequalvis* vem sendo utilizada como um modelo em menor escala para entender os efeitos de cooperatividade que ocorrem entre as subunidades de globinas da HbLt. Porém, a interface dimérica da Hb de *S. inaequalvis* apresenta uma série de resíduos não conservados quando comparados com as sequências do oligoquetas, onde alguns já foram determinados como cruciais para atividade dessa proteína^{145,146}. Sendo assim, a interface *d/d* encontrada para a *d*-HbGp é um modelo em menor escala alternativo para o estudo do comportamento das subunidades globinas de eritrocruorinas.

5 RESULTADOS E DISCUSSÃO – OUTRAS HEMOGLOBINAS

5.1 Purificação

As hemoglobinas das três espécies comerciais, *E.andrei*, *P.escavatus* e *E.eugeniae*, e da espécie nativa do Vale do Paraopeba, *R.alatus*, foram extraídas e purificadas. O rendimento obtido na purificação foi variado, assim como o tamanho das espécies em questão. Destacando-se o máximo rendimento obtido por indivíduo sacrificado para a espécie *R.alatus*, entre 15 – 20 mg, e um mínimo rendimento por indivíduo sacrificado para a espécie *E.andrei*, ~ 0,75 mg. As diferentes espécies também apresentaram Hbs com propriedades físico-químicas diferentes. A **figura 5.1.1 - a** é uma análise por SDS PAGE demonstrando que, assim como a HbGp, todas as diferentes eritrocruorinas purificadas nesta etapa são compostas de dois conjuntos de bandas, *linkers* e globinas, porém, o perfil das bandas referentes a cada conjunto varia de acordo com a espécie. Dentre as Hbs analisadas, as HbEa é a que apresenta maior discrepância nas massas relativas das subunidades. Essas diferenças se refletem na massa total do complexo que foi estimada posteriormente por ultracentrifugação analítica (UAC – *ultra analytic centrifugation*) (**figura 5.1.1 - b**). Os valores das massas obtidos por UAC podem ser vistos na **tabela 5.1.1**.

Figura 5.1.1 - (a) Análise de SDS PAGE das frações purificadas contendo as hemoglobinas, HbPe, HbEe, HbEa e HbRa. A HbEa é a que apresenta um padrão de massa que mais diverge das outras Hbs purificadas neste trabalho. **(b)** Resultado da análise por UAC, a amostra de HbEa apresentam dois picos em S=10 (referentes ao *cap* isolado) e S ~65, que equivale ao coeficiente de sedimentação da partícula inteira.

Tabela 5.1.1 - Massas obtidas por UAC para as hemoglobinas gigantes utilizadas neste trabalho. A massa estimada para a HbEa é maior do que para a HbGp, HbRa e HbPe.

Espécie/Hemoglobina		Massa estimada (kDa)
<i>Glossoscolex paulistus</i>	HbGp	3600 ± 80
<i>Rhinodrilus alatus</i>	HbRa	3500 ± 70
<i>Eisenia andrei</i>	HbEa	3850 ± 50
<i>Perionyx excavatus</i>	HbPe	3500 ± 50

Os resultados de UAC apontam que a HbEa possui a maior massa entre as hemoglobinas aqui estudadas, estimada em 3850 kDa. Isto é em torno de 250-350kDa acima dos valores obtidos para as demais espécies. Pelo padrão de bandas no gel, aparentemente este aumento na massa para a HbEa é devido a diferenças tanto nas cadeias *linkers* quanto nas globinas.

5.2 Cristalização

A motivação que levou ao estudo de hemoglobinas de novas espécies neste projeto, foi a busca por cristais de melhor qualidade e que difratem a mais alta resolução. Dessa forma temos o objetivo de vislumbrar melhor os detalhes das estruturas de cada uma das cadeias além dos contatos entre elas que poderiam, por sua vez, contribuir para modelos mais apurados sobre o comportamento biológico desses complexos. Foram obtidos cristais das Hbs de todas as espécies em questão. Os cristais da HbRa foram obtidos na condição: 10% PEG8000 e 50mM Tris/HCl pH7,0. Esta condição é muito similar aos primeiros cristais obtidos para a HbGp, o que sugere uma similaridade entre essas duas estruturas. Os cristais de HbEe foram obtidos na condição: 20 % MPD Tris/HCl pH7,0. Nestes dois casos os cristais gerados eram de difícil manipulação devido à fragilidade mecânica apresentada pelos mesmos. Para HbEa e HbPe, as condições de cristalização foram; sulfato de amônio 1,4 M, cacodilato pH 6,5 e PEG 3350 10%, cacodilato pH 7,0, respectivamente. Os cristais de HbEa e HbPe eram de fácil manipulação e apresentaram boa resistência na transferência para a condição crioprotetora. A **figura 5.2.1** mostra os cristais obtidos para as quatro espécies em questão.

Figura 5.2.1 - Cristais obtidos para as eritrocruorinas de 4 novas espécies. (a) Os cristais da HbRa, obtidos em 10% PEG8000 e 50mM Tris/HCl pH7,0. (b) Os cristais de HbEe, obtidos na condição 20 % MPD Tris/HCl pH7,0. (c) Os cristais de HbEa obtidos em sulfato de amônio 1,4 M, cacodilato pH 6,5 . (d) Cristais de HbEa, obtidos em PEG 3350 10% cacodilato pH 7,0.

5.3 Coleta e processamentos dos dados

Os cristais obtidos de HbRa e HbEe não apresentaram difração significativa, limitando-se a poucos pontos de difração na zona de baixa resolução, impossibilitando a coleta de dados. Já os cristais de HbPe e HbEa, apresentaram poder de difração considerável, com reflexões bem definidas e intensas. A análise preliminar dos dados coletados para o cristal de HbPe sugerem que o mesmo faz parte do grupo espacial P3 (portanto sistema trigonal), com parâmetros de cela de $a = b = 317,76\text{\AA}$ e $c = 608,17\text{\AA}$. Um total de 604.209 reflexões foram coletadas, sendo 107.283 reflexões únicas, limitando o conjunto de dados a uma resolução de $7,0\text{\AA}$. A **figura 5.3.1 - a** mostra uma imagem de difração do cristal de HbPe, onde é possível ver o conjunto de pontos aproximados devido a grande dimensão do eixo c . O conteúdo da unidade assimétrica não pôde ser determinado para esse sistema.

Figura 5.3.1 - Imagens dos padrões de difração coletados e estatística do processamento (R_{merge} e Intensidade) por faixa de resolução para (a) HbPe, onde é possível ver a orientação do eixo maior, $c=608,17\text{\AA}$, que é caracterizado pela proximidade dos *spots*. (b) HbEa.

Os cristais de HbEa pertencem ao grupo I23, cúbico, com parâmetros de cela $a = b = c = 489,59\text{\AA}$. Um total de 300.603 reflexões foram integradas, destas 111.634 são reflexões únicas. A resolução obtida para o conjunto de dados da HbEa foi de $4,7\text{\AA}$. Nos dois casos, o processamento foi feito com parâmetro mosaicidade fixo ($0,5^\circ$) adotando-se a mesma estratégia usada nos conjuntos de dados da HbGp. Os principais dados obtidos durante

indexação, integração e processamento dos dados podem ser vistos na **tabela 5.3.1**. Uma análise do C_M mostra que para o conteúdo da unidade assimétrica encontrado (4 protômeros - ver adiante) a porcentagem de solvente do cristal é estimado em 75%, o que esta em pleno acordo com as baixas resoluções obtidas.

Tabela 5.3.1 - Principais resultados obtidos nos procedimentos de indexação, integração e processamento dos dados. Os valores entre parênteses referem-se a última faixa de resolução.

	HbPe	HbEa
Grupo espacial	P3	I23
Dimensões de cela a, b, c (Å)	317,76, 317,76 , 608,17	489,59, 489,59, 489,59
Detector	ADSC Quantum 315	Pilatus 6M
Fonte de raios-X	NSLS X29	NSLS X25
Comprimento de onda (Å)	1,00	1,00
Resolução	109,85 (7,38–7,00)	130,85 (4,74–4,7)
Multiplicidade	5,6 (5,6)	2,7 (2,7)
Rmerge(%)†	53,7 (75,1)	4,3 (56,20)
Completeza	99,1 (99,6)	94,2 (98,3)
Reflexões totais	604.209(88.947)	300.603 (44.661)
Reflexões únicas	107.283 (15.815)	111.634(16.353)
Intensidade (I/σI)	3,8 (1,7)	15,7 (1,6)
B-fator no Wilson plot (Å ²)	258.8	96.9

5.4 Substituição molecular e estrutura preliminar da HbEa

O problema das fases para a estrutura da HbEa foi resolvido por substituição molecular utilizando a estrutura homóloga da HbLt como molde (código de acesso – 2GTL). Para esse procedimento foi utilizado o programa Phaser. Os metais e grupos heme foram excluídos da estrutura molde e as regiões em que ocupam, usadas posteriormente como critério de avaliação da solução obtida (mesmo procedimento adotado anteriormente para a HbGp e *d*-HbGp). Foram localizados quatro protômeros na unidade assimétrica, o que representa um terço da unidade biológica (60 subunidades ao todo). Como consequência, a unidade biológica com 180 subunidades esta situada sob o eixo de ordem três que corta a diagonal de corpo (*body diagonal*) da cela unitária, como mostra na figura 5.4.1. Os valores de R e LLG obtidos para os sucessivos protômeros encontrados durante o processo de atribuição das fases pode sem vistos na **tabela 5.4.1**.

Figura 5.4.1 - A unidade assimétrica é composta por quatro protômeros, cada um deles colorido por uma cor diferente. Em cinza claro estão representados os simétricos relacionados que formam a unidade biológica. (a) mostra a unidade assimétrica ao longo do eixo **c** e as operações de simetria que relacionam o grupo espacial I23. (b) mostra o conteúdo da unidade assimétrica ao longo do eixo de simetria de ordem três que corta a diagonal da cela.

O modelo cristalográfico da HbEa apresenta um mapa de densidade eletrônica de baixa resolução que é limitado majoritariamente a interpretação dos elementos de estrutura secundária da estrutura. Isso é suficiente para constatar que a HbEa também trata-se de uma eritrocruorina do tipo I (o que já era esperado considerando o alto grau de conservação sequencial com a HbLt – alinhamento da **figura 3.2.2**). Em algumas regiões específicas do mapa de densidade eletrônica (2Fo-Fc) existem indícios da presença de cadeias laterais, como no do bolsão hidrofóbico do grupo heme.

Tabela 5.4.1 - Valores obtidos de LLG e R para cada etapa do processo de atribuição das fases por substituição molecular.

	LLG	R
1 ^a protômero	257	62
2 ^a protômero	3025	55
3 ^a protômero	6710	48
4 ^a protômero	10283	42

O mapa de densidade eletrônica (2Fo-Fc) na região dos grupos heme comprova a presença dos mesmos, mas não dá pistas a respeito da orientação (logo, estes foram modelados com base na orientação em que ocupam nas estruturas da HbLt/HbGp). Um pico significativo no mapa de diferença (Fo-Fc - $\sigma = 7$) comprova a presença do íon Ca^{2+} nas estruturas dos *linkers*. Os mapas de densidade eletrônica/diferença (2Fo-Fc e Fo-Fc) da HbEa não sugerem a presença de íons Zn^{2+} equivalentes aos encontrados na estrutura da *ciano-HbGp*. A **figura 5.4.2** mostra regiões do mapa de densidade eletrônica comparadas as regiões equivalentes no modelo da HbGp.

Figura 5.4.2 - Comparação entre os mapas de densidade eletrônica obtidos (2Fo-Fc – sigma = 1,5) para a HbEa (4,7Å) e *ciano-HbGp* (3,2Å). (a) região do *coiled-coil*. (b) domínio β -barril da subunidade L1 e (c) Região do sítio de ligação da globina tipo *d*.

A HbEa possui sequências que representam a todas as suas subunidades, porém, as sequências referentes as subunidades *b* e L2 estão disponíveis apenas de forma parcial (**figura 3.2.2**). Dessa forma, as regiões desconhecidas para as sequencias de HbEa foram preenchidas com as sequências das subunidades homólogas da HbLt (que conservam 80% identidade com a HbEa). Três subunidades da HbEa são preditas como glicosiladas; a subunidade *a*, na posição 86 (posição 88 do **alinhamento 3.2.1**), subunidade L1, na posição Thr103 e subunidade L2, na posição 161 (posições 128 e 188 do **alinhamento 3.2.2**). Essas predições estão de acordo com os mapas (2Fo-Fc) calculados nas respectivas regiões. Na posição 88 da subunidade *a* há um pequeno prologamento da densidade eletrônica na região da cadeia lateral, coerente com a presença de uma unidade de sacarídeo ligada à estrutura. A posição Thr 103 da subunidade L1 apresenta uma prologamento de densidade eletrônica que equivalente a 3-4 unidades de sacarídeos. Por fim, a posição 161 da subunidade L2, possui um prolongamento na densidade eletrônica correspondente à duas ou mais unidades de sacarídeos. As densidades eletrônicas referentes às glicosilações encontradas para a estrutura da HbEa podem ser observadas na **figura 5.4.3**. Embora as sequências das subunidade *b* e L2 estejam incompletas, os mapas calculados (2Fo-Fc e Fo-Fc) não apresentam nenhum indício de que as subunidades da HbEa apresentam prologamentos nas porções N/C-terminais. Dessa forma, as principais diferenças nas massas encontradas nas análises de SDS PAGE e UAC, são atribuídas as glicosilações e não para diferenças nas massas das cadeias polipeptídicas.

Figura 5.4.3 - Mapa de densidade eletrônica (2Fo-Fc – sigma = 1,5) calculados para o modelo cristalográfico da HbEa onde as sequências das subunidades foram preditas como glicosiladas (a) Asn 86 das subunidades *a*, (b) Thr das subunidades L1 e (c) Asn 161 das subunidades L2.

6 CONCLUSÕES E PERSPECTIVAS FUTURAS

6.1 Conclusões

Na primeira etapa deste trabalho a hemoglobina gigante de *Glossoscolex paulistus* foi purificada, cristalizada e sua estrutura cristalográfica resolvida. Para a obtenção de um modelo cristalográfico completo, foi realizado um sequenciamento massivo do cDNA da espécie *G.paulistus* com objetivo de identificar as sequências desconhecidas de 6 das 7 subunidades básicas que constituem o complexo da HbGp. O sequenciamento massivo resultou em um total de 97.639 sequências únicas, das quais 11 são referentes às subunidades da HbGp. Foram identificadas seis sequências de globinas, $\alpha 1$, $\alpha 2$, β , γ , $\delta 1$ e $\delta 2$ e cinco sequências de *linkers*, L1a, L1b, L2a, L3a e L3b, onde $\delta 1$ é mesma sequência obtida para a subunidade δ por Cabral e colaboradores em 2001¹⁰⁴. O trabalho de sequenciamento da HbGp resultou na obtenção das sequências de todas as subunidades constituintes da HbGp, incluindo algumas de suas isoformas. A comparação entre as sequências obtidas para a HbGp e as suas homólogas na estrutura da HbLt, mostra que a conservação entre essas estruturas é 15 – 20% maior do que a estimativa feita por Cabral e colaboradores com base apenas na sequência $\delta 1$. A estrutura da HbGp foi obtida com uma resolução de 3.2 Å, com valores de $R = 21,65$ e $R_{\text{free}} = 23,52$ em amplo acordo com os critérios que avaliam a qualidade dos modelos depositados no banco de dados. Um total de 45 subunidades (7.108 resíduos) constituem a unidade assimétrica. A estrutura cristalográfica da HbGp apresentou melhor resolução e melhores estatísticas que a estrutura resolvida para a HbLt. O conteúdo reduzido na unidade assimétrica permitiu que o refinamento feito para a HbGp não fizesse uso dos mesmos recursos para a simplificação do modelo imposto para a HbLt. Isso permitiu que ao final do refinamento, os mapas de densidade eletrônica calculados ($2F_o - F_c$ e $F_o - F_c$) revelassem detalhes importantes sobre essa estrutura. Além dos íons Ca^{2+} que já haviam sido reportados para a estrutura da HbLt e que estão conservados na HbGp, outros dois sítios de ligação de metais foram encontrados. O sítio de zinco 1, formado na interface entre o domínio LDL-A e o β -barril, e o sítio de zinco 2 na interface entre as subunidades L3 e L2 de protômeros vizinhos dentro da mesma camada hexagonal. A presença desses íons está relacionada a dois aspectos importantes no comportamento das eritrocruorinas: (i) esse cátion é apontado como um

regulador da afinidade desses complexos por oxigênio, atuando de forma diferente dos metais alcalinos terrosos; (ii) sua presença está associada à atividade superóxido dismutase que esses sistemas apresentam. Os dois sítios de zinco encontrados para a HbGp não formam um ambiente químico equivalente/semelhante ao encontrado nos sítios ativos das enzimas Cu/ZnSODs. Isso indica que, caso esses sítios de zinco encontrados na HbGp estejam relacionados com a atividade de SOD, estes provavelmente fazem uso de mecanismos alternativos aos encontrados nas Cu/ZnSODs. Uma glicosilação predita para o resíduo Asn 121 das subunidades L3 apresentou grande coerência com o mapa (2Fo-Fc) nesta região. Embora apenas uma unidade NAG foi atribuída a esta glicosilação, a densidade eletrônica sugere que pelo menos mais uma unidade de sacarídeo estaria presente. O enovelamento das subunidades constituintes da HbGp é muito similar aos apresentados na HbLt, resultando em sobreposições com valores de *rmsd* abaixo do esperado. O mesmo acontece para a orientação dos oligômeros formados o que inclui a posição dos *caps* e protômeros, mostrando que as estruturas são de fato muito semelhantes. A HbGp e HbLt apresentam uma mesma hierarquia de associação das subunidades de globinas para a montagem do *cap*, que é apontado como o caroço alostérico para esses sistemas. Isso torna-se mais evidente com a grande semelhança que as estruturas dos *caps* da HbGp e HbLt apresentam quando comparadas as Hb_L de vestimentíferos e pogonóforos, porém, esses organismos marinhos apresentam sítios de ligação de metais nas regiões interfaciais das globinas que não são conservados para a HbGp e HbLt. Isso mostra que a influência dos metais sobre as eritrocruorinas de oligoquetas acontece majoritariamente através de sua interação com as subunidades *linkers*. Sem dúvida, a mais proeminente dessas interações é a conservada ponte salina entre a Arg 34 da subunidade *b* com a o Asp88 de L1 (e seus equivalentes nos outros *linkers*) reforçando que o papel da espécie Ca²⁺ não é meramente estrutural na formação do domínio LDL-A. Foram encontrados 13 canais que dão acesso a grande cavidade formada dentro da estrutura do protômero, permitindo que exista um fluxo de entrada e saída de solvente. Isso leva a crer que pequenas moléculas podem migrar para dentro da cavidade, sugerindo que as eritrocruorinas atuem como estruturas carreadores dessas moléculas. A interação entre protômeros de uma mesma camada hexagonal inclui contatos formados por globinas de *caps* diferentes, sugerindo que essas interações também contribuem para o comportamento diferenciado que a unidade biológica demonstra em relação ao *cap* isolado. Por fim, os *linkers* desempenham exclusivamente os contatos entre protômeros de camadas hexagonais diferentes, possuindo um papel fundamental na estabilização do complexo ao manter os dois discos hexagonais unidos.

A estrutura da *d*-HbGp apresentou um enovelamento similar ao encontrado para a subunidade *d* presente no complexo, com exceção ao rearranjo do *loop* AB em virtude dos contatos cristalinos na estrutura da *d*-HbGp. A caracterização da *d*-HbGp em solução, combinada com a análise dos contatos cristalinos, demonstram que o trímero d_3 , presente na estrutura do *cap* é desfeito quando esta subunidade é isolada do resto do complexo. Ao se reassociar, a subunidade *d* forma uma dímero *d/d*, com interface equivalente ao dímero *a/d* encontrado na estrutura do *cap*. Isso torna remota a possibilidade de que o *cap* seja formado a partir de um trímero d_3 , onde este atuaria como núcleo, orientando e agregando as outras subunidades para a formação do *cap*. Dessa forma, concluímos que o trímero d_3 é uma consequência das interfaces que a subunidade *d* desempenha com as outras subunidades no complexo (especialmente com a subunidade *a*).

A busca sistemática por cristais de outras espécies com o objetivo de obter modelos cristalográficos de maior resolução foi bem sucedida para a cristalização das 4 novas eritrocruorinas disponíveis até presente momento; *E.andrei* (HbEa), *P.escavatus* (HbPe) *R.alatus* (HbRa) e *E.eugenia* (HbEe). Somente os cristais de HbEa e HbPe, apresentaram difração considerável, e por fim, apenas os dados coletados para a HbEa permitiram um correto faseamento e obtenção da estrutura preliminar. A estrutura cristalográfica da HbEa foi resolvida a uma resolução de 4,7 Å, possui uma unidade assimétrica com 60 subunidades (o que equivale a 1/3 da unidade biológica) e alguns de suas sequências estão incompletas. Nestas condições a interpretação do modelo é bastante limitada quando comparada aos resultados obtidos anteriormente para a HbGp, mas a orientação dos protômeros constituintes comprovam que esta é uma eritrocruorina do tipo I. Uma análise dos mapas da densidade eletrônica (2Fo-Fc) e dos mapas de diferença (Fo-Fc) demonstram claramente a presença de glicosilações nas subunidades *a*, L1 e L2. As subunidades tipo *a* apresenta suas glicosilações orientadas para dentro da grande cavidade formadas entre o *cap* e a trímero de *linkers*. Não é possível dizer quantas unidades de sacarídeos exatamente compõem cada uma das glicosilações. Não existem evidências de que as porções N/C-terminais das subunidades da HbEa apresentem prolongamentos, de forma que as diferenças de massa que caracterizam essa eritrocruorina é atribuída majoritariamente as glicosilações encontradas.

Foram apresentadas aqui 2 novas estruturas de hemoglobinas gigantes (eritrocruorinas), com destaque para a estrutura cristalográfica da HbGp que é o modelo obtido com mais alta resolução para esses sistemas até hoje. O modelo cristalográfico da *d*-HbGp revelou detalhes importantes do comportamento oligomérico dessas subunidades

quando isoladas, auxiliando na compreensão de como é a cascata de eventos que ocorre para a montagem desses sistemas. A caracterização estrutural da HbGp e suas frações isoladas, combinados com a ampla caracterização de suas propriedades físico-químicas, coloca essa eritrocruorina ao lado da HbLt como uma das mais bem caracterizadas hemoglobinas gigantes conhecidas.

6.2 Perspectivas

A obtenção das estruturas cristalográficas apresentadas neste trabalho representam um importante, porém limitado, passo para o entendimento dos sistemas em questão. Consideramos que os seguintes itens merecem esforços a serem realizados no futuro:

- Obtenção de estruturas cristalográficas da HbGp/HbEa nas formas a *oxi*, *deoxi*, *aquometa*, e *hemecromo*.
- Para os cristais nas diferentes formas obtidas, intensificar as análises espectroscópicas durante os experimentos de difração, afim de entender melhor como os efeitos da radiação alteram o ambiente químico do grupo heme.
- Obtenção das estruturas cristalográficas de outras frações da HbGp/HbEa, como *cap* e/ou o trímero *abc*, que irá auxiliar na determinação do mecanismo de empacotamento da unidade biológica.
- Sequenciamento completo das subunidades da HbEa por métodos de biologia molecular/química de proteínas, o que permitirá a completa construção do modelo cristalográfico da HbEa.
- Obtenção de cristais que difratem a mais a alta resolução, em especial para a HbEa, afim de entender melhor o comportamento e importância dos *oligossacarídeos* ligados a estrutura.

REFERÊNCIAS

- 1 PERUTZ, M. F. An optical method for finding the molecular orientation in different forms of crystalline haemoglobin; changes in dichroism accompanying oxygenation and reduction. **Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character.** Royal Society, v.141, n.902, p.69-71, 1953.

- 2 KENDREW, J. C. Architecture of a protein molecule. **Nature**, v.182, n.4638, p.764-7, 1958.

- 3 KENDREW, J. C. et al. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. **Nature**, v.181, n.4610, p.662-6, 1958.

- 4 PERUTZ, M. F. Structure of hemoglobin. **Brookhaven Symposia in Biology**, v.13, p.165-83, 1960.

- 5 PERUTZ, M. F. et al. Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis. **Nature**, v.185, n.4711, p.416-22, 1960.

- 6 FERMI, G. et al. The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. **Journal of Molecular Biology**, v.175, n.2, p.159-74, 1984.

- 7 PESCE, A. et al. Protein structure in the truncated (2/2) hemoglobin family. **IUBMB Life**, v.59, n.8-9, p.535-41, 2007.

- 8 WEBER, R. E.; VINOGRADOV, S.N. Nonvertebrate hemoglobins: functions and molecular adaptations. **Physiological Reviews**, v.81, n.2, p.569-628, 2001.

- 9 GOODMAN, M. et al. An evolutionary tree for invertebrate globin sequences. An evolutionary tree for invertebrate globin sequences. **Journal of Molecular Evolution**, v.27, n.3, p.236 –49, 1988.

- 10 MOENS, L. et al. Globins in nonvertebrate species: dispersal by horizontal gene transfer and evolution of the structure-function relationships. **Molecular Biology and Evolution**, v.13, n.2, p.324 –33, 1996.

- 11 VINOGRADOV, S. N. et al. A phylogenomic profile of globins. **BMC Evolutionary Biology**, v.6, p.31-48, 2006. DOI:10.1186-1471-2148-631.

12 APPLEBY, C. A. Leghemoglobin and rhizobium respiration. **Annual Review of Plant Physiology**, v.345, p.443-478, 1984. DOI: 10.1146/annurev.pp.35.060184.002303

13 APPLEBY, C. A.; TJEPKEMA, J. D.; TRINICK, M. J. Hemoglobin in a nonleguminous plant, parasponia: possible genetic origin and function in nitrogen fixation. **Science**, v.220, n.4600, p.951-3, 1983.

14 BROWN, G. G. et al. The evolution of a plant globin gene family. **Journal of Molecular Evolution**, v.21, n.1, p.19-32, 1984.

15 LANDSMANN, J. et al. Common evolutionary origin of legume and non-legume plant haemoglobin. **Nature**, v. 324, p.166-168, 1986. DOI:10.1038/324166a0

16 BIDON-CHANAL, A. et al. Ligand-induced dynamical regulation of NO conversion in Mycobacterium tuberculosis truncated hemoglobin-N. **Proteins**, v.64, n.2, p.457-64, 2006.

17 POOLE, R. K.; HUGHES, M. N. New functions for the ancient globin family: bacterial responses to nitric oxide and nitrosative stress. **Molecular Microbiology**, v.36, n.4, p.775–783, 2000.

18 MUKAI, M., et al. Flavohemoglobin, a globin with a peroxidase-like catalytic site. **The Journal of Biological Chemistry**, v.276, n.10, p.7272-7, 2001.

19 PESCE, A. et al. HisE11 and HisF8 provide bis-histidyl heme hexa-coordination in the globin domain of Geobacter sulfurreducens globin-coupled sensor. **Journal of Molecular Biology**, v.386, n.1, p.246-60, 2009.

20 CHOTHIA, C.; LEVITT, M.; RICHARDSON, D. Helix to helix packing in proteins. **Journal of Molecular Biology**, v.145, n. 1, p.215-50, 1981.

21 GARDNER, J. D.; SCHUBERT, R.W. Evaluation of myoglobin function in the presence of axial diffusion. **Advances in Experimental Medicine and Biology**, v.411, p.157-69, 1997.

22 QIU, Y. et al. Lamprey hemoglobin structural basis of the bohr effect. **The Journal of Biological Chemistry**, v. 275, n.18, p.13517-28, 2000.

23 PERUTZ, M. F. et al. The stereochemical mechanism of the cooperative effects in hemoglobin revisited. **Annual Review of Biophysics and Biomolecular Structure**, v.27, p.1-34, 1998. DOI:10.1146/annurev.biophys.27.1.1

24 RICHARD, V.; DODSON, G. G.; MAUGUEN, Y. Human deoxyhaemoglobin-2,3-diphosphoglycerate complex low-salt structure at 2.5 Å resolution. **Journal of Molecular Biology**, v.233, n.2, p.270-4, 1993.

25 NICHOLS, W. L.; , ZIMM, B. H.; TEN EYCK, L. F. Conformation-invariant structures of the $\alpha 1\beta 1$ human hemoglobin dimer. **Journal of Molecular Biology**, v.270, n.4, p.598-615, 1997.

26 WEBER, R. E. Functions of invertebrate hemoglobins with special reference to adaptations to environmental hypoxia. **American Zoologist**, v.20, n. 1, p.79 –101, 1980.

27 WAKABAYASHI, S.; MATSUBARA, H.; WEBSTER, D. A. Primary sequence of a dimeric bacterial haemoglobin from *Vitreoscilla*. **Nature**, v.322, n.6078, p.481– 83, 1986.

28 ROYER JR, W. E.; HENDRICKSON, W. A.; CHIANCONE, E. The 2.4-Å crystal structure of *Scapharca* dimeric hemoglobin: cooperativity based on directly communicating hemes at a novel subunit. **Journal of Biological Chemistry**, v.264, n.35, p.21052–61, 1989.

29 VINOGRADOV, S. N.; MACHLIK, C. A., CHAO, L. L. The intracellular hemoglobins of a polychaete: some properties of the hemoglobins of *Glycera dibranchiata*. **Journal of Biological Chemistry**, v.245, n.24, p.6533– 38, 1970.

30 ILAN, E.; DANIEL, E. Structural diversity of arthropod extracellular haemoglobins. **Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology**, v.63, n. 3, p.303–8, 1979.

31 BLAXTER, M. L. Nemoglobins: divergent nematode globins. **Parasitology Today**, v.9, n.10, p.353–60, 1993.

32 MANNING, A. M.; TROTMAN, C. N.; TATE, W. P. Evolution of a polymeric globin in the brine shrimp *Artemia*. **Nature**, v.348, n.6302, p.653– 56, 1990.

33 BONAVENTURA, C.; BONAVENTURA, J. Respiratory pigments: structure and function. In: HOCHACHKA, P.W. **The Mollusca: environmental biochemistry and physiology**. New York: Academic Press, 1983, v. 12, p.1–362.

34 TERWILLIGER, R. C.; TERWILLIGER, N. B.; SCHABTACH, E. Extracellular hemoglobin of the clam, *Cardita borealis* (Conrad): an unusual polymeric hemoglobin. **Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology**, v.59, n.1, p. 9 –14, 1978.

35 HOURDEZ, S. et al. Hemoglobins from deep-sea hydrothermal vent scaleworms of the genus *Branchipolynoe*: a new type of quaternary structure. **Proteins, Structure, Function and Genetics**, v.34, n.4, p.427–34, 1999.

36 TERWILLIGER, N. B. Molecular structure of the extracellular heme proteins. In: GILLES, R.(Ed.) **Advances in comparative and environmental physiology**. Berlin: Springer-Verlag, 1992. v.13, p.193–229.

37 TOULMOND, A. Properties and functions of extracellular heme pigments. In: GILLES, R. (Ed.) **Advances in comparative and environmental physiology**. Berlin: Springer-Verlag, 1992. v.13, p.231–256.

38 WEBER, R. E. Oxygenational properties of vascular and coelomic haemoglobins from *Nephtys hombergii* (Polychaeta) and their functional significance. **Netherlands Journal of Sea Research**, v.5, n. 2, p. 240 –51, 1971.

39 WEBER, R. E. Respiratory pigments. In: MILL PJ (Ed.) **Physiology of Annelids**. New York: Academic Press, 1978. p. 393– 446.

40 VINOGRADOV, S. N. The structure of invertebrate extracellular hemoglobins (erythrocrurins and chlorocrurins). **Comparative Biochemistry and Physiology - Part B: Biochemistry & Molecular Biology**, v.82, n.1, p.1–15, 1985.

41 LAMY, J. N. Giant hexagonal hemoglobins. **Chemical Reviews**, v.96, n.8, p.3113– 124, 1996.

42 GOTOH, T. et al. Two globin strains in the giant annelid extracellular haemoglobins. **The Biochemical Journal**, v.241, n.2, p.441-5, 1987.

43 SUZUKI, T.; RIGGS, A. F. Linker chain L1 of earthworm hemoglobin: structure of gene and protein: homology with low density lipoprotein receptor. **The Journal of Biological Chemistry**, v.268, n.18, p.13548-55, 1993.

44 YUASA, H. J. et al. Electrospray ionization mass spectrometric composition of the 400 kDa hemoglobin from the pogonophoran *Oligobranchia mashikoi* and the primary structures of three major globin chains. **Biochimica et Biophysica Acta**, v.1296, n.2, p.235-44, 1996.

45 NEGRISOLO, E. et al. The evolution of extracellular hemoglobins of annelids, vestimentiferans, and pogonophorans. **The Journal of Biological Chemistry**, v.276, n.28, p.26391-7, 2001.

46 NUMOTO, N. et al. Structure of an extracellular giant hemoglobin of the gutless beard worm *Oligobranchia mashikoi*. **Proceedings of the National Academy of Sciences of the United States of America**, v.102, n.41, p.14521-6, 2005.

47 NUMOTO, N. et al. Structure of the partially unliganded met state of 400 kDa hemoglobin: insights into ligand-induced structural changes of giant hemoglobins. **Proteins**, v.73, n.1, p.113-25, 2008.

48 FLORES, J. F. et al. Sulfide binding is mediated by zinc ions discovered in the crystal structure of a hydrothermal vent tubeworm hemoglobin. **Proceedings of the National Academy of Sciences of the United States of America**, v.102, n.8, p.2713-8, 2005.

49 ROCHE J. Electron-microscope studies on high molecular weight erythrocrurins (invertebrate haemoglobins) and chlorocrurins of annelids. In: MUNDAY, KA (Ed.) **Studies in comparative biochemistry**. Oxford, UK: Pergamon Press, 1965. p. 62– 80.

50 VINOGRADOV SN, KAPP OH, AND OHTSUKI M. The extracellular haemoglobins and chlorocrurins of annelids. In: HARRIS J (Ed.) **Electron microscopy of proteins**. New York: Academic Press, 1982. v. 3, p.135–163.

51 MESSERSCHMIDT, U. et al. The molecular size and shape of the extracellular hemoglobin of *Nephtys incisa*. **Biochimica and Biophysica Acta**, v. 742, n. 2, p. 366 –373, 1983.

52 PILZ, I. et al. Small angle X-ray scattering studies of giant hemoglobin from the annelid *Tylorrhynchus heterochaetus*. **International Journal of Biology Macromolecules**, v.10, n. 6, p. 356 –360, 1988.

53 PILZ, I.; SCHWARZ, E.; VINOGRADOV, S. N. Small angle X-ray studies of *Lumbricus terrestris* haemoglobin. **International Journal of Biology Macromolecules**, v.2, n.5, p.279 – 83, 1980.

54 HEUER, M., et al. The molecular size and shape of the native and reassociated forms of the extracellular hemoglobin of *Tubifex tubifex*. **International Journal of Biological Macromolecules**, v.7, n.1, p.25–9, 1985.

55 WILHELM, P.; PILZ, I.; VINOGRADOV, S. N. Comparison of *Arenicola marina* and *Lumbricus terrestris* haemoglobins by small-angle X-ray scattering. **International Journal of Biological Macromolecules**, v.2, n. 2, p.383–4, 1980.

56 DE HAAS, F., et al. Three-dimensional reconstruction of native and reassembled *Lumbricus terrestris* extracellular hemoglobin: localization of the monomeric globin chains. **Biochemistry**, v.36, n. 24, p.7330–8, 1997.

57 DE HAAS, F. et al. Three-dimensional reconstruction of *Macrobdella decora*(Leech) hemoglobin by cryoelectron microscopy. **Biophysical Journal**, v.70, n.4, p.1973–84, 1996.

58 DE HAAS, F., et al. Three-dimensional reconstruction of the chlorocruorin of the polychaete annelid *Eudistylia vancouverii*. **Journal of Molecular Biology**, v.255, n.1, p.140–153, 1996.

59 DE HAAS, F. et al. Threedimensional reconstruction of the hexagonal bilayer hemoglobin of the hydrothermal vent tube worm *Riftia pachyptila* by cryoelectron microscopy. **Proteins**, v.26, n.3, p.241–256, 1996.

60 DE HAAS, F. Three-dimensional reconstruction by cryoelectron microscopy of the giant hemoglobin of the polychaete worm *Alvinella pompejana*. **Journal of Molecular Biology**, v.264, n.1, p.111–120, 1996.

61 ROYER, W. E. J.; OMARTIAN, M. N.; KNAPP, J. E. Low resolution crystal structure of *Arenicola erythrocrurin*: influence of coiled coils on the architecture of a megadalton respiratory protein. **Journal of Molecular Biology**, v.365, n.1, p.226-36, 2007.

62 JOUAN, L. et al. Occurrence of two architectural types of hexagonal bilayer hemoglobin in annelids: comparison of 3D reconstruction volumes of *Arenicola marina* and *Lumbricus terrestris* hemoglobins. **Journal of Molecular Biology**, v.305, n.4, p.757-71, 2001.

63 VIDUGIRIS, G. J. et al. Absence of ligand binding-induced tertiary changes in the multimeric earthworm *Lumbricus terrestris* hemoglobin. A resonance Raman study. **Journal of Biological Chemistry**, v.268, n.35, p.26190–92, 1993.

64 PERUTZ, M.; PERUTZ M. **Mechanisms of cooperativity and allosteric regulation in proteins**. Cambridge, UK: Cambridge University Press, 1990. p. 1–101.

65 PERUTZ, M. F. Myoglobin and haemoglobin: role of distal residues in reactions with haem ligands. **Trends in Biochemical Sciences**, v.14, n.2, p.42–44, 1989.

66 KREBS, A. et al. Molecular shape, dissociation, and oxygen binding of the dodecamer subunit of *Lumbricus terrestris* hemoglobin. **Journal of Biological Chemistry**, v.271, n.31, p.18695–704, 1996.

67 KREBS, A.; ZIPPER, P.; VINOGRADOV, S. N. Lack of size and shape alteration of oxygenated and deoxygenated *Lumbricus terrestris* hemoglobin? **Biochimica et Biophysica Acta**, v.1297, n.2, p.115–118, 1996.

68 CHUNG, M. C. M.; ELLERTON, H. D. The physico-chemical and functional properties of extracellular respiratory haemoglobins and chlorocruorins. **Progress in Biophysics and Molecular Biology**, v.35, n. 2, p.53–102, 1979.

69 GIARDINA, B.; CHIANCONE, E.; ANTONINI, E. Studies on erythrocrurorin. III. Oxygen equilibrium of earthworm erythrocrurorin. **Journal of Molecular Biology**, v.93, n.1, p.1–10, 1975.

70 TOULMOND, A. Properties and functions of extracellular heme pigments. In: GILLES, R. (Ed.) **Advances in comparative and environmental physiology**. Berlin: Springer-Verlag, 1992. v.13, p.231–56.

71 IGARASHI, Y.; KIMURA, K.; KAJITA, A. Analysis of oxygen equilibria of the giant hemoglobin from the earthworm *Eisenia foetida* using the Adair model. **Journal of Biochemistry**, v.109, n.2, p.256–261, 1991.

72 KROGH-RASMUSSEN, K.; WEBER, R. E. Respiratory properties of erythrocrurorin (extracellular hemoglobin) in the blood of the annelid *Arenicola marina* with special reference to the influences of salinity and temperature. **Ophelia**, v.18, n.2, p.151–170, 1979.

73 OCHIAI T. Dissociation and oxygen equilibrium properties of the extracellular hemoglobin of *Eisenia foetida*. **Archives of Biochemistry and Biophysics**, v.231, n.1, p.136–143, 1984.

74 TOULMOND, A. et al. Extracellular hemoglobins of hydrothermal vent annelids: structural and functional characteristics in three Alvinellid species. **The Biological Bulletin**, v.179, n.3, p.366–373, 1990.

75 WEBER, R. E. Functions of invertebrate hemoglobins with special reference to adaptations to environmental hypoxia. **American Zoologist**, v.20, n. 1, p.79–101, 1980.

76 TOULMOND, A.; TCHERNIGOVTZEFF, C. Ventilation and respiratory gas exchanges of the lugworm, *Arenicola marina* (L.) as functions of ambient PO₂ (20–700 Torr). **Respiratory Physiology**, v.57, n.3, p.349–363, 1984.

77 ILAN, E.; HAROUN, J. Oxygen-binding properties of extracellular hemoglobin from the leech, *Hirudo medicinalis*. Effects of pH, cations and temperature. **Biochimica et Biophysica Acta**, v.1162, n.1-2, p.77– 83, 1993.

78 ILAN, E.; HAROUN, J. Physicochemical features and oxygenation properties of extracellular hemoglobin from the leech, *Macrobdella decora*. **Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology**, v.104, n.4, p.833– 839, 1993.

79 WEBER, R. E., et al. Mass spectrometric composition, molecular mass and oxygen binding of *Macrobdella decora* hemoglobin and its tetramer and monomer subunits. **Journal of Molecular Biology**, v.251, n.5, p.703–720, 1995.

80 SANTUCCI, R.; CHIANCONE, E.; GIARDINA B. Oxygen binding to *Octolasion complanatum* erythrocrurin. Modulation of homo- and heterotrophic interactions by cations. **Journal of Molecular Biology**, v.179, n.4, p.713–727, 1984.

81 FUSHITANI, K.; IMAI, K.; RIGGS, A. F. Oxygenation properties of hemoglobin from the earthworm, *Lumbricus terrestris*: effects of pH, salts and temperature. **Journal of Biological Chemistry**, v.261, n.18, p.8414 – 23, 1986.

82 CHIANCONE, E. et al. A cooperative hemoglobin with directly communicating hemes. The *Scapharca inaequivalvis* homodimer. **Biophysical Chemistry**, v.37, n.1-3, p.287–292, 1990.

83 OCHIAI, T. Dissociation and oxygen equilibrium properties of earthworm (*Pheretima hilgendorfi*) hemoglobin. **Archives of Biochemistry and Biophysics**, v.226, n.1, p.111–17, 1983.

84 OCHIAI, T.; WEBER, R.E. Effects of magnesium and calcium on the oxygenation reaction of erythrocrurin from the marine polychaete *Arenicola marina* and the terrestrial oligochaete *Lumbricus terrestris*. **Zoological Science**, v.19, n.9, p.995-1000, 2002.

85 STANDLEY, P. R. et al. The calcium, copper and zinc content of some annelid extracellular haemoglobins. **Biochemical Journal**, v.249, n.3, p.915–16, 1988.

86 OCHIAI, T.; HOSHINA, S.; USUKI, I. Zinc as modulator of oxygenation function and stabilizer of quaternary structure in earthworm hemoglobin. **Biochimica et Biophysica Acta**, v.1203, n.2, p.310-4, 1993.

87 ROYER, W. E. J. et al. Structural hierarchy in erythrocrurin, the giant respiratory assemblage of annelids. **Proceedings of the National Academy of Sciences of the United States of America**, v.97, n.13, p.7107-11, 2000.

88 ROYER, W. E. J. et al. Lumbricus erythrocrurin at 3.5 Å resolution: architecture of a megadalton respiratory complex. **Structure**, v.14, n.7, p.1167-77, 2006.

89 STRAND, K. et al. Crystal structure of the hemoglobin dodecamer from Lumbricus erythrocrurin: allosteric core of giant annelid respiratory complexes. **Journal of Molecular Biology**, v.344, n.1, p.119-34, 2004.

90 ROYER, W. E. J. et al. The 2.0 Å crystal structure of Scapharca tetrameric hemoglobin: cooperative dimers within an allosteric tetramer. **Journal of Molecular Biology**, v.253, n.1, p.168-86, 1995.

91 GOTOH, T. et al. Hexagonal bilayer structuring activity of linker chains of an annelid giant hemoglobin from the polychaete Perinereis aibuhitensis. **Archives of Biochemistry and Biophysics**, v.360, n.1, p.75– 84, 1998.

92 KUCHUMOV, A. R., et al. The role of linkers in the reassembly of the 3.6MDa hexagonal bilayer hemoglobin from Lumbricus terrestris. **Journal of Molecular Biology**, v.289, n.5, p.1361–74, 1999.

93 ZHU, H. et al. Assembly of the gigantic hemoglobin of the earthworm Lumbricus terrestris: roles of subunit equilibria, non-globin linker chains, and valence of the heme iron. **Journal of Biological Chemistry**, v.271, n.47, p.30007–21, 1996.

94 KAO, W. et al. Linker chains of the gigantic hemoglobin of the earthworm Lumbricus terrestris: primary structures of linkers L2, L3, and L4 and analysis of the connectivity of the disulfide bonds in linker L1. **Proteins**, v.63, n.1, p. 174-87, 2006.

95 FUSHITANI, K. et al. Characterization of the constituent polypeptides of the extracellular hemoglobin from Lumbricus terrestris: heterogeneity and discovery of a new linker chain L4. **Biochimica et Biophysica Acta**, v.1292, n.2, p.273-80, 1996.

96 LIOCHEV, S. I. et al. Superoxide dismutase activity in the giant hemoglobin of the earthworm, Lumbricus terrestris. **Archives of Biochemistry and Biophysics**, v.330, n.2, p.281–284, 1996.

97 HART, P. J. et al. A structure-based mechanism for copper-zinc superoxide dismutase. **Biochemistry**, v.38, n.7, p. 2167-78, 1999.

98 RIGHI, G. Sobre a família Glossoscolecidae (Oligochaeta) no Brasil. **Arquivos de Zoologia**. v. 29, p. 1-95, 1971. ISSN 0066-7870.

99 COSTA, M. C. et al. : Sedimentation coefficient and minimum molecular weight of extracellular hemoglobin of *Glossoscolex paulistus* (Oligochaeta). **Brazilian Journal of Medical and Biological Research**, v. 21, n. 1, p. 115-8, 1988.

100 CARVALHO, F. A. O. et al. On the molecular mass of the extracellular hemoglobin of *Glossoscolex paulistus*: analytical ultracentrifugation reexamination. **Analytical Biochemistry**, v.385, n.2, p.257-63, 2009.

101 CARVALHO, F. A. O. et al. Molecular masses and sedimentation coefficients of extracellular hemoglobin of *Glossoscolex paulistus*: alkaline oligomeric dissociation. **International Journal of Biological Macromolecules**, v. 48, n.1, p.183-93, 2011.

102 CARVALHO, F. A. O. et al. Urea-induced unfolding of *Glossoscolex paulistus* hemoglobin, in oxy- and cyanomet-forms: a dissociation model. **International Journal of Biological Macromolecules**, v. 52, p. 340-8, 2013. DOI:10.1016-jbiomac.2012.09.23.

103 AGUSTINHO, S. C. et al. Spectroscopic studies of the met form of the extracellular hemoglobin from *Glossoscolex paulistus*. **Biochimica et Biophysica Acta**, v.1298, n.2, p.148-58, 1996.

104 BOSCH CABRAL, C. et al. Fluorescence properties of tryptophan residues in the monomeric d-chain of *Glossoscolex paulistus* hemoglobin: an interpretation based on a comparative molecular model. **Biophysical Chemistry**, v.97, n.2-3, p.139-57, 2002.

105 GELAMO, E. L.; ITRI, R.; TABAK, M. Small angle x-ray scattering (SAXS) study of the extracellular hemoglobin of *Glossoscolex paulistus*: effect of pH, iron oxidation state, and interaction with anionic SDS surfactant. **The Journal of Biological Chemistry**, v.279, n.32, p.33298-305, 2004.

106 OLIVEIRA, M. S.; MOREIRA, L. M.; TABAK, M. Partial characterization of giant extracellular hemoglobin of *Glossoscolex paulistus*: a MALDI-TOF-MS study. **International Journal of Biological Macromolecules**, v.40, n.5, p.429-36, 2007.

107 MARQUES, M. B.; MEIRELLES, N. C. Erythrocrurin of *Glossoscolex paulistus* (Righi) (Oligochaeta, Glossoscolecidae): effects of divalent ions, acid-alkaline transition and alkali and urea denaturation. **Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology**, v.111, n.2, p.311–318, 1995.

108 RIGHI, G. Bionomic considerations upon Glossoscolecidae (Oligochaeta). **Pedobiologia**, v.12, n.4, p.254-260, 1972.

109 MABOETA, M. S.; REINECKE, A. J.; REINECKE, S.A. Effects of low levels of lead on growth and reproduction of the Asian earthworm *Perionyx excavatus* (Oligochaeta). **Ecotoxicology and Environmental Safety**, v.44, n.3, p.236-40, 1999.

110 CHO, S. J. et al. Gene expression profile in the anterior regeneration of the earthworm using expressed sequence tags. **Bioscience, Biotechnology, and Biochemistry**, v.73, n.1, p.29-34, 2009.

111 CHAUDHARI, R. D.; DATAR, M. T.; BABOOKANI, M.R. Municipal solid waste management through vermicomposting employing exotic species of earthworm *Eudrilus eugeniae*. **Journal of Environmental Science & Engineering**, v.53, n.1, p.129-34, 2011.

112 KUMARI, M. et al. Bioconversion of herbal industry waste into vermicompost using an epigeic earthworm *Eudrilus eugeniae*. **Waste Management & Research : the Journal of the International Solid Wastes and Public Cleansing Association, ISWA**, v.29, n.11, p.1205-12, 2011.

113 YESUDHASON, B. V. et al. Unique phenotypes in the sperm of the earthworm *Eudrilus eugeniae* for assessing radiation hazards. **Environmental Monitoring and Assessment**, 2012. In press. DOI:10.1007-s10661-012-2901-9.

114 BUCH, A.C. **Pontoscolex Corethrurus (Müller, 1857) e Eisenia Andrei, Bouché 1972, como bioindicadoras de solos contaminados por agrotóxicos**. 2010. 63p. Dissertação (Mestrado em Ciência do Solo) - Setor de Ciências Agrárias, Universidade Federal do Paraná, Curitiba, 2010.

115 VAMPRÉ, T. M; FUCCILLO, R.; DE ANDRÉA, M. M. Oligochaeta *Eisenia andrei* como bioindicador de contaminação de solo por hexaclorobenzeno. **Revista Brasileira de Ciência do Solo**, v.34, n.1, p.59-66, 2010.

116 DANIEL, E. et al. Towards a resolution of the long standing controversy regarding the molecular mass of extracellular erythrocrurin of the earthworm *Lumbricus terrestris*. **Biochimica et Biophysica Acta- Protein and Proteomics**. v.1649, n.1, p.1–15, 2003.

117 HIRSCH, R.E. et al. A first evaluation of the natural high molecular weight polymeric *Lumbricus terrestris* hemoglobin as an oxygen carrier. **Artificial Cells, Blood Substitutes, and Immobilization Biotechnology**, v.25, n.5, p.429–44, 1997.

118 MCPHERSON, A. **Crystallization of biological macromolecules**. New York: Cold Spring HarborLab. Press, 1999.

119 BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v.72, p.248-54, 1976.

120 GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v.29, n.7, p. 644-52, 2011.

121 ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v.25, n.17, p.3389-402, 1997.

122 BERGFORS, T. M. **Protein crystallization**. 2nd.ed. Oxford: Internacional University Line, 2006.

123 LESLIE, A. G. The integration of macromolecular diffraction data. **Acta Crystallographica Section D: Biological Crystallography**, v.62, Pt. 1, p. 48-57, 2006.

124 EVANS P. Scaling and assessment of data quality. **Acta Crystallographica Section D**, v.62, n.10, p.72-82, 2005.

125 USÓN, I.; SHELDRIK, G.M. Advances in direct methods for protein crystallography. **Current Opinion in Structural Biology**, v.9, n.5, p.643-8, 1999.

126 BALDWIN, J. M. The structure of human carbonmonoxy haemoglobin at 2.7 Å resolution. **Journal of Molecular Biology**, v.136, n.2, p.103-28, 1980.

127 ROSSMAN, M. G.; BLOW, D. M. The detection of sub-units within the crystallographic asymmetric unit. **Acta Crystallographica**, v.15, Pt. 1, p.24-32, 1962.

128 COLLABORATIVE COMPUTATIONAL PROJECT, N. M. The CCP4 suite: programs for protein crystallography. **Acta Crystallographica Section D: Biological Crystallography**, v.50, Pt. 5, p.760-3, 1994.

129 ADAMS, P. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D: Biological Crystallography**, v.66, Pt. 2, p.213-21, 2010.

130 CHEN, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D: Biological Crystallography**, v.66, Pt. 1, p.12-21, 2010.

131 EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v.32, n.5, 1792-97, 2004.

132 CLAMP, M. et al. The Jalview Java alignment editor. **Bioinformatics**, v.20, n.3, p. 426-7, 2004.

133 CHOTHIA, C.; LESK, A. M. The relation between the divergence of sequence and structure in proteins. **The EMBO Journal**, v.5, n.4, p.823-6, 1986.

134 STANDLEY, P. R. The calcium, copper and zinc content of some annelid extracellular haemoglobins. **Biochemical Journal**, v.249, n.3, p.915-6, 1988.

135 NIAL, H. D. Automated Edman degradation: the protein sequencer. **Methods in Enzymology**, v.27, p.942-1010, 1973.

136 FUSHITANI, K.; RIGGS, A. F. The extracellular hemoglobin of the earthworm, *Lumbricus terrestris*: oxygenation properties of isolated chains, trimer, and a reassociated product. **Journal of Biological Chemistry**, v.266, n.16, p.10275-81, 1991.

137 OLIVEIRA, M. S.; MOREIRA, L. M.; TABAK, M. Partial characterization of giant extracellular hemoglobin of *Glossoscolex paulistus*: a MALDI-TOF-MS study. **International Journal of Biological Macromolecules**, v. 40, n. 5, p. 429-36, 2007.

138 OLIVEIRA, M. S.; MOREIRA, L. M.; TABAK, M. Interaction of giant extracellular *Glossoscolex paulistus* hemoglobin (HbGp) with ionic surfactants: a MALDI-TOF-MS study. **International Journal of Biological Macromolecules**, v.42, n. 2, p.111-9, 2008.

139 MAILLETT, D. H. et al. Interfacial and distal-heme pocket mutations exhibit additive effects on the structure and function of hemoglobin. **Biochemistry**, v.47, n. 40, p. 10551-63., 2008.

140 FASS, D., et al. Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. **Nature**, v.388, n.6643, p.691–3, 1997.

141 HARBURY, P.B. “A switch between two-, three- and four-stranded coiled coils in GCN4 leucine zipper mutants”, **Science**, v.262, n.5138, p.1401-7, 1993.

142 ROYER, W. et al. Nonvertebrate hemoglobins: structural bases for reactivity. **Progress in Biophysics and Molecular Biology**. v. 68, n. 1, p. 29–68, 1997.

143 KANTARDJIEFF, K.A.; RUPP, B. Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. **Protein Science**, v. 12, n. 9, p. 1865-1871, 2003.

144 STRAND, K. et al. Crystal structure of the hemoglobin dodecamer from *Lumbricus erythrocrur*: allosteric core of giant annelid respiratory complexes. **Journal of Molecular Biology**, v. 344, n.1, p.119-34, 2004.

145 PARDANANI, A. et al. Mutation of residue Phe97 to Leu disrupts the central allosteric pathway in *Scapharca* dimeric hemoglobin. **Journal Biological Chemistry**, v. 272, n. 20, p. 13171–13179, 1997.

146 ROYER, W. et al. Ordered water molecules as key allosteric mediators in a cooperativedimeric hemoglobin. **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, p.14526–14531, 1996.

147 URZHUMTSEVA, L. et al. Crystallographic model quality at a glance. **Acta Crystallographica. Section D: Biological Crystallography**, v. 65, p. 297-300, 2009. DOI:10.1107-50907444908044296

ANEXOS

TRABALHOS PUBLICADOS

Systematic structural studies of iron superoxide dismutases from human parasites and a statistical coupling analysis of metal binding specificity

Jose Fernando Ruggiero Bachega,¹ Marcos Vicente Albuquerque Salles Navarro,¹ Lucas Bleicher,¹ Raquel Kely Bortoleto-Bugs,¹ Daniel Dive,³ Pascal Hoffmann,² Eric Viscogliosi,³ and Richard Charles Garratt^{1*}

¹Institute of Physics of São Carlos, University of São Paulo, CEP 13560-970, São Carlos, SP, Brazil

²UMR/CNRS 5068, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France

³INSERM U547, Institut Pasteur, 1 rue du Professeur Calmette, BP 245, 59019 Lille cedex, France

ABSTRACT

Superoxide dismutases (SODs) are a crucial class of enzymes in the combat against intracellular free radical damage. They eliminate superoxide radicals by converting them into hydrogen peroxide and oxygen. In spite of their very different life cycles and infection strategies, the human parasites *Plasmodium falciparum*, *Trypanosoma cruzi* and *Trypanosoma brucei* are known to be sensitive to oxidative stress. Thus the parasite Fe-SODs have become attractive targets for novel drug development. Here we report the crystal structures of FeSODs from the trypanosomes *T. brucei* at 2.0 Å and *T. cruzi* at 1.9 Å resolution, and that from *P. falciparum* at a higher resolution (2.0 Å) to that previously reported. The homodimeric enzymes are compared to the related human MnSOD with particular attention to structural aspects which are relevant for drug design. Although the structures possess a very similar overall fold, differences between the enzymes at the entrance to the channel which leads to the active site could be identified. These lead to a slightly broader and more positively charged cavity in the parasite enzymes. Furthermore, a statistical coupling analysis (SCA) for the whole Fe/MnSOD family reveals different patterns of residue coupling for Mn and Fe SODs, as well as for the dimeric and tetrameric states. In both cases, the statistically coupled residues lie adjacent to the conserved core surrounding the metal center and may be expected to be responsible for its fine tuning, leading to metal ion specificity.

Proteins 2009; 77:26–37.
© 2009 Wiley-Liss, Inc.

Key words: *Plasmodium falciparum*; *Trypanosoma cruzi*; *Trypanosoma brucei*; malaria; Chagas' diseases; sleeping sickness; iron superoxide dismutase; crystal structure; statistical coupling analysis; metal specificity.

INTRODUCTION

Tropical protozoal parasite infections such as malaria, African sleeping sickness and Chagas' disease represent a major global health problem, causing the death of ~1.3 million people per year¹ and having serious consequences on economic development in affected countries. A variety of limiting factors associated with the current treatments, mainly drug resistance for malaria, low efficacy and toxicity² of drugs for sleeping sickness and Chagas' disease, and the limitation of future prospects in term of vaccination justifies the need for the continuing development of efficient drugs against these developing-world diseases. From among many plausible candidates, the enzymes involved in antioxidant defence have emerged as potential new drug targets.³ Despite distinct life cycles, the causative agents of malaria, African sleeping sickness and Chagas' disease in humans, *Plasmodium falciparum* (the most deadly *Plasmodium* species), *Trypanosoma brucei* and *Trypanosoma cruzi*, respectively, are all exposed to intense oxidative stress, during at least one stage of their life cycles.^{4,5} However, an imbalance in the levels of reactive oxygen species (ROS), such as those caused by some antimalarial and antitrypanosomal drugs,^{6–8} can lead to parasite death. Additionally, the antioxidant defences found in mammalian hosts present major differences when compared

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: FAPESP; Grant number: 98/14138-2; Grant sponsor: CAPES, CNPq.
Navarro Marcos Vicente Albuquerque Salles's current address is CA-173 Veterinary Medical Center, Cornell University, Ithaca, NY 14853.
Bortoleto-Bugs Raquel Kely's current address is Centro de tecnologia da informação (CTI) Renato Archer, Rodovia Dom Pedro I, km 143, 6, Bairro Amara, CEP 13069-901 Campinas, SP, Brazil.

Bachega Jose Fernando Ruggiero and Navarro Marcos Vicente Albuquerque Salles contributed equally to this work.

*Correspondence to: Garratt Richard Charles, Institute of Physics of São Carlos, University of São Paulo, CEP 13560-970, São Carlos, SP, Brazil. E-mail: richard@ifsc.usp.br

Received 8 October 2008; Revised 13 January 2009; Accepted 5 February 2009

Published online 3 March 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22412

Crystallization and preliminary structural analysis of the giant haemoglobin from *Glossoscolex paulistus* at 3.2 Å

J. F. R. Bachega,^{a*} L. Bleicher,^a E. R. Horjales,^a P. S. Santiago,^b R. C. Garratt^a and M. Tabak^{b*}

^aCentro de Biotecnologia Molecular Estrutural, Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos – SP, CEP 13566-590, Brazil, and ^bInstituto de Química de São Carlos, Universidade de São Paulo, São Carlos – SP, CEP 13566-590, Brazil.
E-mail: bachega@ursa.ifsc.usp.br, marceltabak2@yahoo.com.br

Glossoscolex paulistus is a free-living earthworm encountered in south-east Brazil. Its oxygen transport requirements are undertaken by a giant extracellular haemoglobin, or erythrocrurin (HbGp), which has an approximate molecular mass of 3.6 MDa and, by analogy with its homologue from *Lumbricus terrestris* (HbLt), is believed to be composed of a total of 180 polypeptide chains. In the present work the full 3.6 MDa particle in its cyanomet state was purified and crystallized using sodium citrate or PEG8000 as precipitant. The crystals contain one-quarter of the full particle in the asymmetric unit of the *I*222 cell and have parameters of $a = 270.8$ Å, $b = 320.3$ Å and $c = 332.4$ Å. Diffraction data were collected to 3.15 Å using synchrotron radiation on beamline X29A at the Brookhaven National Laboratory and represent the highest resolution data described to date for similar erythrocrurins. The structure was solved by molecular replacement using a search model corresponding to one-twelfth of its homologue from HbLt. This revealed that HbGp belongs to the type I class of erythrocrurins and provided an interpretable initial electron density map in which many features including the haem groups and disulfide bonds could be identified.

Keywords: haemoglobin; erythrocrurin; crystallization; earthworm.

1. Introduction

The annelid erythrocrurins are giant hexagonal bilayer haemoglobins with molecular masses within the megadalton range. They are multimeric assemblies built from the association of both globin and non-globin chains which typically present a highly cooperative behaviour (Weber & Vinogradov, 2001). The crystal structures of two types of such erythrocrurins have thus far been described, those from *Lumbricus terrestris* (HbLt) (Royer *et al.*, 2006) and *Arenicola marina* (HbAc) (Royer *et al.*, 2007). In both cases the architecture of the full particle of approximately 3.6 MDa is based on two hexagonal discs in which the most prominent substructure corresponds to one-twelfth of the particle (a protomer) and is composed of a dodecamer of globins together with three non-globin chains called linkers. Both structures possess 622 (or D_6) symmetry but present slightly different relative orientations of the two discs. In type I structures, such as that from HbLt, the hexagonal layers are staggered with respect to one another by approximately 16°, while in type II structures, such as that seen in HbAc, the two layers are eclipsed. These

arrangements as seen in the crystal structures are consistent with those observed in three-dimensional reconstructions based on cryo-electron microscopy images (Jouan *et al.*, 2001) and would therefore appear to be intrinsic structural differences inherent to the two classes of molecule rather than alternative quaternary states accessible to a single species. Besides earthworms, cryo-electron microscopy has also established that the type I architecture is also observed in both leeches (de Haas, Biosset *et al.*, 1996) and a hydrothermal vent tube worm (de Haas, Zal *et al.*, 1996), whereas the type II class is only well established for polychaetes (de Haas, Taveau *et al.*, 1996).

A full particle of HbLt is composed of 144 globin chains and 36 linkers. Four different types of globin chain (a , b , c and d) unite to form $abcd$ tetramers in which the a subunit is covalently linked to b and c by disulphide bridges. Three such tetramers form the dodecameric head of the one-twelfth protomers which are further stabilized by the presence of one copy of each of the three linker chains $L1$, $L2$ and $L3$. The latter protrude from the head into the centre of the particle giving the protomer a mushroom-like appearance and are

Recent New Characterizations on the Giant Extracellular Hemoglobin of *Glossoscolex paulistus* and Some Other Giant Hemoglobins from Different Worms

Marcel Tabak¹, Francisco A.O. Carvalho¹, José W.P. Carvalho¹,
Jose F.R. Bachega² and Patrícia S. Santiago^{1,3}

¹Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP,

²Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP,

³Campus Experimental de Registro, UNESP, SP,
Brazil

1. Introduction

Giant extracellular hemoglobins, also known as erythrocruorins, have been investigated as a model of extreme complexity in oxygen-binding heme proteins [1,2]. They are characterized by a very high molecular mass (MM) of several megadalton (MDa), and their oligomeric structure together with the crowded and protected heme environment are two of the main factors responsible for the high redox stability [3,4]. Superoxide dismutase (SOD)-like intrinsic activity, observed for hemoglobins of *Lumbricus terrestris* (HbLt) and of *Arenicola marina* (HbAm), is another important factor [5,6]. These hemoglobins present a highly cooperative oxygen binding and a peculiar behavior associated to their oligomeric dissociation into smaller subunits and possible rearrangement back into the native oligomeric structure [7,8].

Moreover, a strong motivation to study these giant hemoglobins is related to their potential use in medicine as blood substitutes. Studies have been performed in the past for HbLt [9], and are presently underway to test and validate the use of HbAm in this direction [6,10]. They seem to be very promising due to the lack of undesirable immunological reactions in tests with animals, explained by the absence of cell membranes as occurs with human hemoglobin in red blood cells [6,10]. Besides, the resistance to oxidation of extracellular hemoglobins, as noticed by their redox stability, is also an advantage as compared to the use of human hemoglobin in this medical application.

The giant extracellular hemoglobin of *Glossoscolex paulistus* (HbGp) is similar to several homologous proteins described in the literature [11-14]. These extracellular hemoglobins are constituted by a large number of globin-like subunits containing heme groups with MM in the range 15-19 kDa. These globin subunits form monomers of 16 kDa (*d*) and disulphide bound hetero-trimers of 51-52 kDa (*abc*), linked by non-heme structures (24-32 kDa), named linkers (*L*) [12,13]. Recent partial characterization of MM of HbGp by matrix assisted laser

research papers

Acta Crystallographica Section D

Biological
Crystallography

ISSN 0907-4449

Larissa Romanello,^a
 José Fernando Ruguiero
 Bachega,^a Alexandre Cassago,^b
 José Brandão-Neto,^c Ricardo
 DeMarco,^a Richard Charles
 Garratt^a and Humberto D'Muniz
 Pereira^{a*}

^aCentro de Biotecnologia Molecular Estrutural,
 Instituto de Física de São Carlos, Universidade
 de São Paulo, Avenida Trabalhador
 Saocarlense 400, São Carlos-SP 13566-590,
 Brazil, ^bLaboratório Nacional de Biotecnologia –
 UNBIO, CP 6192, Campinas-SP 13083-970,
 Brazil, and ^cDiamond Light Source, Harwell
 Science and Innovation Campus, Didcot,
 Oxfordshire OX11 0DE, England

Correspondence e-mail:
 hmuniz.pereira@gmail.com

Adenosine kinase from *Schistosoma mansoni*: structural basis for the differential incorporation of nucleoside analogues

In adult schistosomes, the enzyme adenosine kinase (AK) is responsible for the incorporation of some adenosine analogues, such as 2-fluoroadenosine and tubercidin, into the nucleotide pool, but not others. In the present study, the structures of four complexes of *Schistosoma mansoni* AK bound to adenosine and adenosine analogues are reported which shed light on this observation. Two differences in the adenosine-binding site in comparison with the human counterpart (I38Q and T36A) are responsible for their differential specificities towards adenosine analogues, in which the *Schistosoma* enzyme does not tolerate bulky substituents at the N7 base position. This aids in explaining experimental data which were reported in the literature more than two decades ago. Furthermore, there appears to be considerable plasticity within the substrate-binding sites that affects the side-chain conformation of Ile38 and causes a previously unobserved flexibility within the loop comprising residues 286–299. These results reveal that the latter can be sterically occluded in the absence of ATP. Overall, these results contribute to the body of knowledge concerning the enzymes of the purine salvage pathway in this important human parasite.

Received 11 July 2012
 Accepted 29 October 2012

PDB References: SmAK–
 adenosine, 3vas; 3vaq;
 SmAK–Ado–AMP, 3uq6;
 SmAK–tubercidin, 3uq9;
 SmAK–2-fluoroadenosine,
 4dc3

1. Introduction

Schistosoma mansoni is one of the parasitic species responsible for the disease schistosomiasis; in 2005, the World Health Organization estimated that approximately 200 million people were infected with it (World Health Organization, 2005). It is known that schistosomes (Senft, Crabtree *et al.*, 1973; Senft, Senft *et al.*, 1973) and schistosomules (Dovey *et al.*, 1984) are unable to synthesize purine bases *de novo* and therefore the purine-salvage pathway is exclusively used to supply purine bases for energy requirements and nucleic acid synthesis. This pathway was elucidated in schistosomes and schistosomules in the 1970s and 1980s by the work of Senft (Miech *et al.*, 1975; Senft & Crabtree, 1977, 1983; Senft, Crabtree *et al.*, 1973; Senft *et al.*, 1972; Senft, Senft *et al.*, 1973; Stegman *et al.*, 1973; Crabtree & Senft, 1974) and Dovey (Dovey *et al.*, 1984).

Schistosomes can incorporate adenine, adenosine, inosine and hypoxanthine into ATP. Adenine is anabolized to AMP *via* adenine phosphoribosyltransferase (APRT). Adenosine can be incorporated *via* two different mechanisms: the adenosine kinase reaction, yielding AMP, and the indirect pathway (adenosine → inosine → hypoxanthine → IMP → AMP), which requires the enzymes adenosine deaminase (ADA), purine nucleoside phosphorylase (PNP), hypoxanthine-guanine phosphoribosyltransferase (HGPRT), adenylosuccinate synthase (ADSS) and adenylosuccinate

GTKDynamo: A PyMOL Plug-In for QC/MM Hybrid Potential Simulations

José Fernando R. Bachega,^{[a]*} Luís Fernando S. M. Timmers,^[b] Lucas Assirati,^[a] Leonardo R. Bachega,^[c] Martin J. Field,^[d] and Troy Wymore^[e]

Hybrid quantum chemical/molecular mechanical (QCMM) potentials are very powerful tools for molecular simulation. They are especially useful for studying processes in condensed phase systems, such as chemical reactions that involve a relatively localized change in electronic structure and where the surrounding environment contributes to these changes but can be represented with more computationally efficient functional forms. Despite their utility, however, these potentials are not always straightforward to apply since the extent of significant electronic structure changes occurring in the condensed

phase process may not be intuitively obvious. To facilitate their use, we have developed an open-source graphical plug-in, GTKDynamo that links the PyMOL visualization program and the pDynamo QC/MM simulation library. This article describes the implementation of GTKDynamo and its capabilities and illustrates its application to QC/MM simulations. © 2013 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23346

Introduction

The molecular simulation of condensed-phase chemical reactions is a very challenging computational task. Quantum chemical (QC) approaches provide the most realistic description of such systems because they can determine molecular electronic structure from first principles and so permit the exploration of processes such as bond-breaking and forming, charge transfer, and electronic excitation. QC methods are, however, computationally expensive and are limited to relatively small systems of several hundred atoms or less, depending upon the method that is being employed. Alternatives to QC approaches are based on molecular mechanical (MM) potentials or force-fields. These can study systems of almost arbitrary size and are widely employed. In the area of biomacromolecular simulation, for example, they are used for structure refinement and prediction, the study of macromolecular dynamics and of protein-ligand interactions. Unfortunately, MM potentials suffer from limitations of their own, including their dependence on a large number of empirical parameters and their unsuitability, in many cases, for studying processes where electronic structure rearrangements occur, such as chemical reactions.

For many condensed-phase chemical reaction simulations, hybrid approaches that combine both QC and MM descriptions of a system are appropriate. Warshel and Levitt first introduced a method of this type in their study of the reaction catalyzed by the enzyme lysozyme.^[1] Since then, QC/MM potentials have become a standard tool for molecular simulation and dominate for some applications, such as in the investigation of the mechanisms of condensed-phase reactions for which a realistic description of the environment is required, as in enzymes. There are a wide variety of QC/MM methods and their development and application have been extensively documented.^[2–6]

Many molecular modeling packages now contain QC/MM capabilities. One of the earliest that was developed with these calculations in mind, however, was the Dynamo series of programs. The first publicly available version, named fDynamo, was written in Fortran 90/95 and was released under an open-source license in 1999.^[7,8] fDynamo was superseded in 2008 by a new version of the library, pDynamo that was written in the scripting language Python and in the compiled language C.^[9,10] pDynamo is much more capable than fDynamo as it is compatible with most of the common biomolecular force-fields, has a wide range of semiempirical and *ab initio* QC methods, and a large number of simulation algorithms. In addition, its use of Python makes it much easier to use and more powerful than fDynamo and other competing packages.

[a] J. F. R. Bachega, L. Assirati

Centro de Biotecnologia Molecular Estrutural, Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos-SP, CEP 13566-590, Brazil
E-mail: ferbachega@gmail.com

[b] L. F. S. M. Timmers

Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO)-Programa de Pós-Graduação em Biologia Celular e Molecular (PPGBCM) - Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre-RS, Brazil

[c] L. R. Bachega

School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, 47907

[d] M. J. Field

DYNAMO/DYNAMOP, Institut de Biologie Structurale-Jean-Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France (UMR 5075-CEA/CNRS/UJF Grenoble II)

[e] T. Wymore

Pittsburgh Supercomputing Center, National Resource for Biomedical Supercomputing Department, Pittsburgh

Contract grant sponsor: National Center for Research Resources at the National Institute of Health; Contract grant number: RR06009

© 2013 Wiley Periodicals, Inc.