

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

Rodrigo Villares Portugal

**Estudos de complexos macromoleculares por crio-
microscopia eletrônica e técnicas biofísicas**

São Carlos

2006

**Estudos de complexos macromoleculares por crio-
microscopia eletrônica e técnicas biofísicas**

Rodrigo Villares Portugal

Tese apresentada ao Instituto de Física de São
Carlos da Universidade de São Paulo para
obtenção do título de Doutor em Física.

Área de Concentração: Física Biomolecular

Orientador: Prof. Dr. Igor Polikarpov

São Carlos

2006

Portugal, Rodrigo Villares

Estudos de complexos macromoleculares por criomicroscopia eletrônica e técnicas biofísicas / Rodrigo Villares Portugal. --

São Carlos, 2006.

133 p.

Tese (Doutorado) -- Instituto de Física de São Carlos, 2006.

Área: Física Biomolecular

Orientador: Prof. Dr. Igor Polikarpov

1. Crio-Microscopia Eletrônica. 2. Complexos macromoleculares. 3. Receptores Nucleares. 4. Hemocianina. I. Título.

DEDICATÓRIA

A Nalva, aos meus pais e irmãos.

AGRADECIMENTOS

Agradeço a todos que estiveram ao meu lado, durante esta etapa da minha vida. Infelizmente, não tenho como citar os nomes de todos que me marcaram ao longo destes anos: amigos, familiares, professores, colegas de grupo e de estudos, funcionários, enfim, todos que de alguma maneira, passaram pela minha vida e deixaram sua marca. Mas posso citar alguns, e através deles agradecer aos outros que não estarei citando aqui. Primeiramente agradeço a Nalva, Companheira, sempre. Aos meus pais, Ceição, João Bosco, Adelina e Roberto, minhas irmãs, Dandan, Marina e Rê, meus irmãos Beto e Caique, e através deles, toda minha família, que mesmo longe, sempre me apoiou e esteve ao meu lado.

Ao Igor por ter me orientado neste trabalho e me dado a oportunidade de conhecer diferentes técnicas e laboratórios. Ao Marin e Maurício meus segundos orientadores e amigos. A todos os amigos do grupo, mas em especial à Maria, "mãinha", que além de excelente profissional e professora, sempre foi muito, muito amiga. A todos os amigos aqui do grupo, Niko, Sandra Krau, Nádia, Lucas, João Renato, Fabão, Fred, Alessandro, Wanius, Marcel, Leandro Rizzi, Ribamar, Carol Gatinha, Carol Hoff, Carol Guzzi, Carol Figueira, Ney, Tavin, Suzana, Liliana, Lisandra, Thiago, Dani, Ana Letícia, Lívia, Daniel, Fernando Falvo, Bachega, Augusto, Gê, Mário Minime, Sonoda, Alécio, Caqui, Andressa, Bel, Napoleão, Luciana

Lavezzo, Alexandre, Bianca, Fernanda, Aderson, Renata, Cláudia Alves. Aos amigos que já encerraram sua participação no grupo, mas deixaram saudades, Bete, Mário Sanches, Marcão, Marcos Calgaro, Monique, Dulce, Aparício, Sandra Martha, André Ambrósio, Cláudia Munte, Jaciara, Pavão, Adriana Rojas. Aos mais antigos ainda, Márcio, PH, Peninha, Diet e Jorge. Aos professores do grupo, Gláucius, Otávio e Richard, pelo apoio que sempre tive. Ao Javier, pelas conversas de fim de churrasco. Aos amigos Rogério, Norba, Henrique, João, Alexandre, Léo e tantos outros da física.

Aos membros do LAPA, do Departamento de Bioquímica Médica da UFRJ, que desde o início da colaboração entre o Prof. Igor e a Profa. Débora Foguel, em meados de 2002, sempre me receberam muito bem. Em especial à Profa. Débora e aos Profs. Victor Moreau e Astria Ferrão-Gonzales, hoje ambos professores do Núcleo de Biotecnologia da Faculdade de Tecnologia e Ciências, em Salvador, BA. Também aos colegas de doutorado Léo, Carol e demais colegas do laboratório, pelo apoio que sempre me deram.

Ao Prof. Maurício T. R. Lima, da Faculdade de Farmácia da UFRJ, que desde março de 2005, quando foi instalado o espectrofluorímetro ISS-PC1 (ISS) no Grupo de Cristalografia, tem me orientado nos estudos biofísicos de receptores nucleares.

Aos amigos que conheci na Inglaterra, Rosi, Osnat, Tim, Cecília, Aris, Graeme, Frank, Sarah, Ardan, Gijs, Jorien, Willem, Alexis, Costis, Rüdi, Hans aos quais sou muito grato pela maneira que me

receberam. Em especial ao Prof. Marin van Heel, que me deu a oportunidade de aprender a técnica de crio-microscopia eletrônica e que durante minhas estadias no Centre for Biomolecular Electron Microscopy (CBEM), do Imperial College London, sempre esteve ao meu lado, me ensinando e me dando todo apoio necessário ao trabalho.

Ao Dr. Daniel Lorenzini do Departamento de Parasitologia da USP–São Paulo, pelo fornecimento da amostra de hemocianina de *Acanthoscurria gomesiana* para os estudos de crio-microscopia eletrônica. Ao Dr. Willem Tichelaar para a obtenção de micrografias das amostras de hemocianina e de receptores nucleares, estudados durante o primeiro estágio no CBEM. Ao Gijs, pelo apoio e ensinamentos dados na utilização dos microscópios e ao Tim, Cecília e Aris, pelas discussões e ensinamentos sobre o processamento dos dados.

E a todos, todos mesmo, que de alguma maneira contribuíram para que eu esteja aqui hoje, meu muito obrigado!

RESUMO

Este trabalho apresenta o estudo e caracterização de dois complexos moleculares, hRXR α Δ AB e hemocianina de *Acanthoscurria gomesiana*, através de técnicas estruturais e biofísicas. O uso da técnica de crio-microscopia eletrônica para o estudo da hemocianina de *Acanthoscurria gomesiana*, resultou em um modelo estrutural com resolução de 17 Å – pelo método de *Fourier Shell Correlation* com o critério de 1/2 bit. Neste limite de resolução, já é possível observar detalhes estruturais que o mostram como sendo compatível com outros modelos de hemocianinas. Com relação ao estudo de hRXR α Δ AB, mostrou-se, através das técnicas de cromatografia analítica de exclusão por tamanho, eletroforese de gel de poliacrilamida e SAXS, que a proteína pode se apresentar no estado dimérico em solução, mesmo na ausência do seu ligante, 9-cis-RA. Também foi estudado a associação de hRXR α Δ AB a elementos responsivos: DR1, DR4, F2 e PAL. Suas constantes de dissociação foram calculadas através da técnica de espectroscopia por anisotropia de fluorescência. Os resultados obtidos mostram maior afinidade por DR1 e DR4 e indicam uma origem entrópica para o processo de associação.

ABSTRACT

This work presents the study and characterization of two molecular complexes, hRXR α Δ AB and hemocyanin from *Acanthoscurria gomesiana*, using structural and biophysical techniques. The use of electron cryo-microscopy to study hemocyanin from *Acanthoscurria gomesiana* resulted in a structural model at 17 Å resolution as calculated by Fourier Shell Correlation using 1/2 bit criteria. This resolution allows us to see some structural details that show the correctness of that model when compared to other hemocyanin structures. As for the study of hRXR α Δ AB, it was shown, using analytical size exclusion chromatography, polyacrylamide gel electrophoresis and SAXS, that the protein forms dimers in solution, even in the absence of its ligand, 9-cis-RA. The association mechanism of hRXR α Δ AB was studied with several responsive elements: DR1, DR4, F2 and PAL. Their dissociation constants were measured using fluorescence anisotropy spectroscopy. Results show that DR1 and DR4 have the highest affinity for the protein and indicate that the association process is entropy driven.

LISTA DE FIGURAS

Figura 1. Estrutura da hemocianina de <i>Panulirus interruptus</i>	7
Figura 2. Estrutura modular dos receptores nucleares.....	9
Figura 3. Imagem de rosto humano à qual foi adicionado ruído.....	19
Figura 4. Resultado da somatória de imagens com ruído.....	20
Figura 5. Soma total de imagens de rosto humano.	21
Figura 6. Imagens desalinhadas translacionalmente e rotacionalmente.....	22
Figura 7. Imagens mostradas na Figura 6 após MRA.....	22
Figura 8. Auto-imagens do conjunto de rostos humanos.....	23
Figura 9. Partículas da hemocianina de <i>A. gomesiana</i>	50
Figura 10. Auto-imagens da hemocianina de <i>A. gomesiana</i>	52
Figura 11. Referências utilizadas para o primeiro MRA.....	53
Figura 12. Projeções típicas utilizadas para a primeira reconstituição angular	54
Figura 13. Comparação entre partículas, médias de partículas da mesma classe e reprojeções do modelo da hemocianina de <i>A. gomesiana</i>	55
Figura 14. Planos do modelo estrutural de hemocianina de <i>A. gomesiana</i> , ao longo do eixo de simetria.....	56
Figura 15. Resolução obtida para o modelo de hemocianina de <i>A.</i> <i>gomesiana</i>	56
Figura 16. Visão estéreo do modelo de hemocianina de <i>A. gomesiana</i> em diferentes orientações.....	60

Figura 17. Visão estéreo da provável organização das sub-unidades da hemocianina de <i>A. gomesiana</i>	61
Figura 18. Estrutura cristalográfica do hexâmero da hemocianina de <i>P. Interruptus</i> ajustado ao modelo obtido de hemocianina de <i>A. gomesiana</i>	62
Figura 19. Interfaces de oligomerização da hemocianina de <i>A. gomesiana</i> .	63
Figura 20. Caracterização da expressão de hRXR α Δ AB por SDS-PAGE 15%	65
Figura 21. Caracterização da purificação de hRXR α Δ AB em cromatografia de afinidade ao Co ²⁺	66
Figura 22. Etapa de purificação de hRXR α Δ AB em cromatografia de exclusão por tamanho.....	68
Figura 23. Estados oligoméricos da hRXR α Δ AB após a cromatografia de exclusão por tamanho.....	71
Figura 24. Desnaturação de hRXR α Δ AB do Pico 2 por uréia acompanhada por PAGE.....	72
Figura 25. Desnaturação de hRXR α Δ AB do Pico 2 por uréia acompanhada por dicróismo circular.....	73
Figura 26. Cromatografia analítica de exclusão por tamanho do Pico 2 da hRXR α Δ AB.....	74
Figura 27. Gráfico de Guinier para hRXR α Δ AB e BSA.....	75
Figura 28. Ligação de hRXR α Δ AB (Pico 1) ao elemento responsivo DR1 e ao ligante 9-cis-RA.....	79
Figura 29. Ligação de hRXR α Δ AB (Pico 2) ao elemento responsivo DR1 e	

ao ligante 9-cis-RA.....	80
Figura 30. Ligação de hRXR α Δ AB (Pico 1 e 2) ao elemento responsivo DR1 e ao ligante 9-cis-RA, após equilíbrio de 2 horas.....	82
Figura 31. Associação de hRXR α Δ AB a diferentes oligonucleotídeos, sem a presença do ligante 9-cis-RA.....	85
Figura 32. Associação de hRXR α Δ AB a diferentes oligonucleotídeos, na presença do ligante 9-cis-RA.....	86
Figura 33. Constantes de dissociação para hRXR α Δ AB e diferentes elementos responsivos (HRE) na presença (+9-cis-RA) e ausência de ligante (-9-cis-RA).....	87
Figura 34. Curvas de ligação hRXR α Δ AB–DR1 em diferentes temperaturas.	89
Figura 35. Gráfico de van't Hoff para a hRXR α Δ AB.....	91

LISTA DE TABELAS

Tabela 1. Descrição dos elementos responsivos utilizados.....	42
Tabela 2. Caracterização das imagens utilizadas.....	51
Tabela 3. Amostras de hRXR α Δ AB incubadas com elemento responsivo e ligante.....	75
Tabela 4. Constantes de dissociação (Kd) para hRXR α Δ AB e diferentes elementos responsivos (HRE) na presença (+9-cis-RA) e ausência de ligante (-9-cis-RA).....	87
Tabela 5. Constantes termodinâmicas da associação hRXR α Δ AB-DR1.....	90

LISTA DE ABREVIATURAS

6-mero	hexâmero
9- <i>cis</i> -RA	ácido retinóico 9- <i>cis</i>
Å	Angstrom
bis-ANS	bis-Anilinaftaleno sulfonado
CBEM	<i>Centre for Biomolecular Electron Microscopy</i>
CCD	<i>Charge-coupled device</i>
CD	Dicroísmo Circular
CTF	Função de transferência de contraste
DBD	Domínio de ligação ao DNA (<i>DNA Binding Domain</i>)
DNA	Ácido desoxirribonucléico
DR1	Repetição direta com 1 espaçamento
DR4	Repetição direta com 4 espaçamentos
DTT	<i>Dithiothreitol</i>
EDTA	<i>Ethylenediaminetetraacetic acid</i>
F2	Palíndromo invertido com 6 espaçamentos
FEG	<i>Field Emission Gun</i>
FSC	<i>Fourier Shell Correlation</i>
HPLC	<i>High performance liquid chromatography</i>
HRE	elementos responsivos ao hormônio
hRXR α Δ AB	Receptor de ácido retinóico 9- <i>cis</i> humano, isoforma α , sem a porção N-terminal
IFSC	Instituto de Física de São Carlos

IPTG	isopropylthio-b-D-galactoside
kDa	kilodalton
KLH1	<i>keyhole limpet hemocyanin type 1</i>
KLH2	<i>keyhole limpet hemocyanin type 2</i>
LBD	Domínio de ligação ao ligante (<i>Ligand Binding Domain</i>)
L	litro
LNLS	Laboratório Nacional de Luz Síncrotron
MDa	megadalton
ml	mililitro
μl	microlitro
mM	milimolar
μM	micromolar
MRA	<i>Multi reference alignment</i>
mRXR α Δ AB	Receptor de ácido retinóico 9-cis de camundongo, isoforma α , sem a porção N-terminal
MSA	<i>Multistatistical analysis</i>
NaCl	Cloreto de sódio
ncRNA	RNAs não codificadores
nM	nanomolar
PAGE	Eletroforese em gel de poliacrilamida
PAL	Palíndromo sem espaçamento
pET-28a(+)	Plasmídio pET-28a(+)
RA	Ácido retinóico
RAR	Receptor de ácido retinóico
RNA	Ácido ribonucléico
RXR	Receptor de ácido retinóico X

SAXS	<i>Small-angle X-ray scattering</i>
SDS	<i>Sodium dodecyl sulphate</i>
SPCEM	<i>Single Particle Cryo Electron Microscopy</i>
TR	Receptor de hormônio tireoidiano

SUMÁRIO

1. Introdução.....	1
1.1. Complexos Macromoleculares.....	1
1.1.1. Hemocianinas.....	6
1.1.2. Receptor de ácido retinóico 9-cis.....	9
1.2. Microscopia Eletrônica.....	14
1.2.1. Formação da imagem.....	14
1.2.2. Processamento das imagens.....	18
1.3. Termodinâmica de proteínas.....	24
1.4. Espectroscopia de fluorescência.....	26
1.4.1. Fluorescência em proteínas.....	26
1.4.2. Ainsotropia de fluorescência.....	27
1.5. SAXS – Espalhamento de raios-X a baixo ângulo.....	29
2. Objetivos.....	32
3. Materiais e Métodos.....	33
3.1. Microscopia Eletrônica de Hemocianina de <i>Acanthoscurria gomesiana</i>	33
3.1.1. Preparação da amostra e microscopia.....	33
3.1.2. Processamento das imagens.....	35
3.2. Expressão e purificação de hRXR α Δ AB.....	40
3.3. Oligonucleotídeos.....	42
3.4. Estudos de oligomerização	43

3.4.1. Ensaio em gel nativo com gradiente de poliacrilamida.....	43
3.4.2. Desnaturação de hRXR α Δ AB por uréia.....	44
3.4.3. Cromatografia analítica de exclusão por tamanho.....	45
3.4.4. Espalhamento de raios-X a baixo ângulo.....	46
3.5. Reconhecimento de elementos responsivos.....	46
4. Resultados e discussões.....	49
4.1. Modelo estrutural da hemocianina de <i>Acanthoscurria gomesiana</i>	49
4.2. Oligomerização de hRXR α Δ AB e especificidade no reconhecimento de elementos responsivos.....	65
4.2.1. Expressão e purificação de hRXR α Δ AB.....	65
4.2.2. Oligomerização de hRXR α Δ AB.....	69
4.2.3. Especificidade no reconhecimento de elementos responsivos...	83
5. Conclusões.....	92
6. REFERÊNCIAS.....	95
7. APÊNDICE A.....	105

1 Introdução

1.1 Complexos Macromoleculares

Nas últimas décadas, o avanço na biologia molecular tem trazido um maior entendimento sobre diversos processos celulares, com impactos científicos e tecnológicos para a sociedade. A hemoglobina foi a primeira macromolécula a ter sua estrutura resolvida, pela técnica de cristalografia de raios-X (PERUTZ et al., 1964), sendo um marco do início da biologia molecular estrutural. Atualmente, esta técnica tem permitido a elucidação de estruturas de macromoléculas com resolução atômica, como feito para o caso do ribossomo (BAN et al., 2000; WIMBERLY et al., 2000, WARREN et al., 2000). Além da cristalografia por raios-X, diversas outras técnicas vêm contribuindo da construção do conhecimento nesta área, entre elas a espectroscopia por ressonância magnética nuclear, microscopia eletrônica, espectroscopia por fluorescência, espectroscopia de massa e modelagem por homologia. Através destas e de outras técnicas estruturais e

biofísicas, obteve-se avanços no entendimento da estrutura e função de diversos constituintes celulares. Entretanto, sabe-se hoje, que a maior parte das funções biológicas tem origem na interação de diversos constituintes do sistema celular, tais como proteínas, DNAs, RNAs e pequenas moléculas. No fim da década passada, Alberts (1998) caracterizou a célula como sendo um *conjunto de máquinas macromoleculares pertencentes a diferentes linhas de produção interligadas*, e não apenas um conjunto de compartimentos, separados por membranas, onde proteínas específicas são confinadas e interagem quase ao acaso para realizar tarefas específicas:

"But instead of a cell dominated by randomly colliding individual protein molecules, we now know that nearly every major process in a cell is carried out by assemblies of 10 or more protein molecules. And, as it carries out its biological functions, each of these protein assemblies interacts with several other large complexes of proteins. Indeed, the entire cell can be viewed as a factory that contains an elaborated network of interlocking assembly lines, each of which is composed of a set of large protein machines."

A existência de complexos moleculares, que atuam conforme máquinas necessárias a um processo específico, sugere a existência de um nível organizacional superior ao que se supunha a algumas décadas atrás. Um ano após o artigo de Alberts, Hatwell et al. (1999), corroboraram a idéia através do reconhecimento da importância de módulos funcionais na organização biológica e de que suas características não podem ser facilmente entendidas através do estudo individual de suas partes constituintes:

"We argue here for the recognition of functional 'modules' as a critical level of biological organization. Modules are composed of many types of molecule. They have discrete functions that arise from interactions among their components (proteins, DNA, RNA and small molecules), but these functions cannot easily be predicted by studying the properties of isolated components."

No ambiente celular, além de complexos protéicos, também são formados complexos macromoleculares através da associação de proteínas a DNAs e RNAs. Receptores nucleares são um exemplo de proteínas que formam complexos, não somente entre si, originando oligômeros, mas também em associações à molécula de DNA (PICARD, 1999; LAUDET; GRONEMEYER, 2002). Estas múltiplas associações de constituintes celulares, para o desempenho de funções específicas, indicam a possível existência de um grande número de complexos macromoleculares dentro do ambiente celular. Estudos sobre o conteúdo de células de leveduras mostraram a existência de cerca de 200 a 300 complexos moleculares formados pela associação de duas ou mais proteínas (GAVIN et al., 2002; HO et al., 2002).

Além dos complexos que vêm sendo identificados através de técnicas correntes, há a possibilidade de existência de outros processos no ambiente celular que ainda não foram identificados. O nível de complexidade dos mecanismos celulares que atuam através da formação de complexos moleculares pode ser ainda maior do que o que se observa hoje.

Recentemente, Mattick (2004) sugeriu a existência de mecanismos de sinalização celular envolvendo RNAs não codificadores (ncRNA), que são originados principalmente por *introns*. A simples existência de um mecanismo como este aumentaria sobremaneira o número de possíveis complexos no ambiente celular. Isto porque, a maioria do genoma de organismos superiores codifica ncRNAs, sendo estes cerca de 98% nos humanos.

O estudo de complexos moleculares tem se mostrado, dessa maneira, como uma nova fronteira com uma clara perspectiva de ser fundamental para o aumento do entendimento dos processos celulares. Segundo Sali et al. (2003), o estudo de grandes complexos macromoleculares, juntamente com a genômica estrutural, é uma tendência da biologia molecular estrutural:

"In recent years, two trends have emerged in structural biology: efforts to achieve a comprehensive coverage of individual protein structures (so-called structural genomics) and efforts to analyse the structures of large complexes."

Embora seja clara a importância de uma contínua busca de entendimento dos constituintes celulares individuais, cada vez mais, o estudo de complexos macromoleculares será fundamental para o aumento da compreensão dos processos celulares. Hartwell et al. (1999) sugerem que a ciência atual está em um movimento de transição entre os estudos de proteínas e genes individuais para o entendimento do que definem como

módulos funcionais:

"Cell biology is in transition from a science that was preoccupied with assigning functions to individual proteins or genes, to one that is now trying to cope with the complex sets of molecules that interact to form functional modules"

Sali et al (2003), poeticamente, definiram a necessidade do estudo de complexos moleculares da seguinte maneira:

"Whatever the timeframe may be, there is no doubt that structural genomics will make a major contribution to the proteomics dictionary of words and phrases. But words and phrases alone do not make literature."

Sendo assim, o estudo de complexos moleculares tem se mostrado necessário para um melhor entendimento dos processos biológicos. Neste sentido, para se ter acesso ao real funcionamento e dinâmica com que complexos moleculares atuam nos organismos, torna-se necessário, não só o correto entendimento de suas partes, mas também, todo um conjunto de informações estruturais e biofísicas a respeito dos diversos complexos necessários à execução de um processo.

Neste trabalho, a técnica estrutural de crio-microscopia eletrônica e técnicas biofísicas, entre elas espectroscopia por fluorescência, foram utilizadas para a análise de complexos moleculares. Foram estudadas a hemocianina de *Acanthoscurria gomesiana* (*A. gomesiana*) e o receptor nuclear humano de ácido retinóico 9-cis (hRXR α Δ AB).

1.1.1 Hemocianinas

Hemocianinas são complexos proteicos que fazem parte do sistema respiratório de espécies de moluscos e artrópodes, sendo responsáveis pelo transporte de oxigênio. São proteínas extracelulares, presentes na hemolinfa, cujos complexos moleculares têm massa de até 9 MDa (ORLOVA et al., 1997). Estes complexos apresentam uma organização estrutural dependente do organismo, podendo ser decâmeros, didecâmeros, hexâmeros, ou complexos de hexâmeros (ORLOVA et al., 1997; VOLBEDA et al., 1989; BELTRAMINI et al., 2005). Cada um destes domínios liga uma única molécula de oxigênio, através de dois átomos de cobre que são coordenados por seis histidinas.

As hemocianinas de moluscos se organizam em homodecâmeros ou heterodecâmeros cilíndricos, compostos por subunidades com cerca de 350 kDa a 450 kDa. Cada uma destas é formada por sete ou oito domínios globulares com cerca de 44 kDa a 55 kDa. Estas subunidades se agrupam formando decâmeros que atingem massa em torno de 3,5 MDa a 4,5 MDa. Estes, por sua vez, podem estar associados dois a dois, formando didecâmeros com massa em torno de 8 MDa a 9 MDa (ORLOVA et al. 1997; KELLER et al. 1999). Um exemplo de um destes complexos é a hemocianina do *keyhole limpet* californiano, *Megathura crenulata*. Este organismo possui duas formas de hemocianina, chamadas KLH1 e KLH2, ambas formadas por um didecâmero cilíndrico. A primeira estrutura deste complexo, obtida por crio-microscopia eletrônica, possuía resolução de 45 Å

(DUBE et al., 1995), impossibilitando a separação entre as duas formas, mas evidenciando dados estruturais comuns a ambas. De Ioannes et al. (2004) mostraram que a hemocianina do gastrópode chileno *Concholepas concholepas* é formada por hetero-didecâmeros com massa de 8 MDa.

No caso dos artrópodes, estas proteínas são compostas por sub-unidades distintas com massa em torno de 75 kDa, que se combinam em hexâmeros conforme Figura 1. Estes hexâmeros podem ser encontrados isolados (1 × 6-mero) ou formando complexos de até oito hexâmeros (2 × 6-mero, 4 × 6-mero, 6 × 6-mero, 8 × 6-mero), dependendo da espécie e das condições fisiológicas (DUBE et al., 1994). Modelos estruturais de sub-unidades de hemocianinas de *Panulirus interruptus* (VAN SCHAICK et al. 1982; VOLBEDA; HOL, 1989) e de *Limulus polyphemus* (HAZES et al., 1993; TAVEAU, et al., 1997) foram feitos por difração de raios-X.

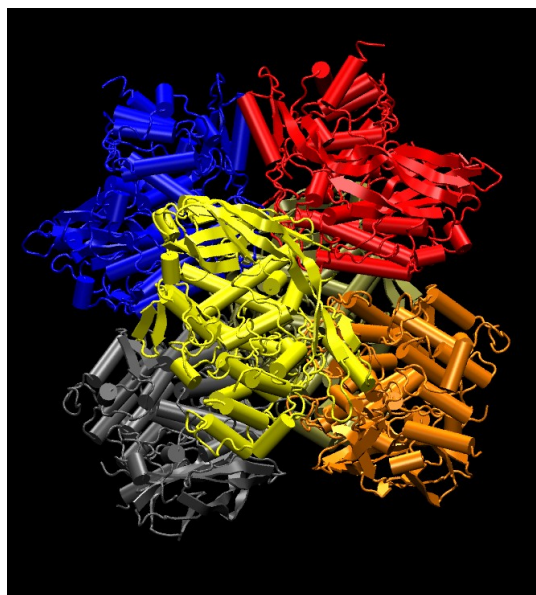


Figura 1. Estrutura da hemocianina de *Panulirus interruptus* obtida por cristalografia de raios-X a 3,2 Å de resolução (VOLBEDA; HOL, 1989), PDBs 1HC1, 1HC2, 1HC3, 1HC4, 1HC5 e 1HC6.

No caso de *P. interruptus*, a molécula de hemocianina é um hexâmero heterogêneo, formado por três sub-unidades: *a*; *b* e *c*. A estrutura resolvida por Volbeda e Hol (1989) possui duas destas: *a* e *b*. No caso da hemocianina de *Eurypelma californicum*, uma espécie de tarântula, o complexo é formado por quatro hexâmeros heterogêneos compostos por sete diferentes sub-unidades (VOIT et al., 2000). As sub-unidades são denominadas *a*, *b*, *c*, *d*, *e*, *f* e *g*. Para formação do 24-mero são necessárias quatro sub-unidades dos tipos *a*, *d*, *e*, *f*, *g*, e duas dos tipos *b* e *c* (a estequiometria do 24-mero é 4*a*, 2*b*, 2*c*, 4*d*, 4*e*, 4*f* e 4*g*). As sete sub-unidades são necessárias para a formação de um 4 × 6-mero completo e estável. Este é formado por dois 2 × 6-mero idênticos com estequiometria 2*a*, 1*b*, 1*c*, 2*d*, 2*e*, 2*f* e 2*g*. Os dois hexâmeros constituintes do 2 × 6-mero são, portanto, diferentes pois um deles possui a sub-unidade *b* enquanto o outro possui a sub-unidade *c*. Estas duas sub-unidades possuem uma forte associação, formando um heterodímero que exerce um importante papel na estabilização do complexo. Esta ligação não só estabiliza a associação entre os dois hexâmeros de cada 2 × 6-mero, como também é responsável pela formação de uma ligação central entre os dois 2 × 6-meros que formam o 4 × 6-mero. As diferentes sub-unidades são necessárias não só para a estabilização do complexo, como também para o seu correto funcionamento.

Assim como a hemocianina de *E. californicum*, a hemocianina da tarântula brasileira *A. gomesiana*, investigada neste trabalho, mostrou-se um complexo composto por 24 sub-unidades agrupadas em quatro hexâmeros (4 × 6-mero) com massa total em torno de 1,7 MDa. Embora

tenham sido identificadas oito sequências de sub-unidades pertencentes à hemocianina de *A. gomesiana* (LORENZINI et al., 2006), a análise do complexo através de espectrometria de massa apresentou peptídios de apenas sete das sequências originalmente identificadas (LORENZINI, comunicação pessoal).

1.1.2 Receptor de ácido retinóico 9-cis

A comunicação entre as células de animais superiores se dá através da interação entre diferentes tipos de moléculas de sinalização (proteínas, nucleotídeos, retinóides, entre outras) e proteínas específicas, denominadas receptores. Estes, estão presentes nas células alvos, podendo estar localizados em diferentes regiões, em sua superfície ou no interior da mesma.

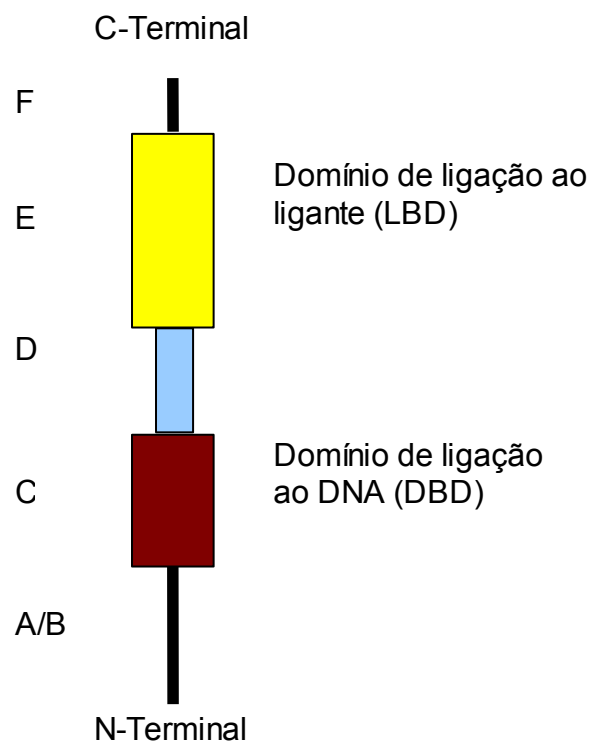


Figura 2. Estrutura modular dos receptores nucleares.

Os receptores nucleares (LAUDET; GRONEMEYER, 2002) compreendem, uma família de receptores que agem no núcleo da célula e estão envolvidos em praticamente todas as funções fisiológicas do organismo. Essa família é composta por receptores para hormônios lipofílicos: esteróides, retinóides, vitamina D e tiroidianos. Entre eles, há também um grupo de receptores para os quais não foram identificados ligantes naturais. Estes são denominados “receptores órfãos” (KLIEWER et al., 1999).

Todos receptores nucleares possuem uma estrutura modular (Figura 2) composta por três grandes domínios: o domínio N-terminal (região A/B), um domínio localizado centralmente (região C), responsável pela ligação da proteína ao DNA (DBD – *DNA Binding Domain*) e o domínio C-terminal (região E), responsável pela ligação ao ligante (LBD – *Ligand Binding Domain*). Há também uma região responsável pela ligação do DBD ao LBD, chamada de dobradiça (região D). Diferentes receptores nucleares se ligam seletivamente à diferentes regiões do DNA (MADER et al., 1993), chamadas de elementos responsivos ao hormônio (HRE, *Hormone Responsive Element*).

Os ácidos retinóicos (RA), metabólitos funcionais da vitamina A, apresentam-se nas formas isoméricas all-trans e 9-cis, e exercem função central em vários processos biológicos, sendo mais conhecidos como cromóforos no processo visual. Entretanto, é sabido que retinóides estão envolvidos em diversos outros processos, como diferenciação de células

epiteliais; desenvolvimento fetal e regulação da proliferação e diferenciação de diversos tipos de células.

O receptor do ácido retinóico 9-cis (RXR, *retinoic X receptor*) foi inicialmente descrito como um receptor órfão e seu ligante, o ácido retinóico 9-cis, (9-cis-RA), foi descrito posteriormente por Mangelsdorf et al. (1990). O RXR está envolvido em vários mecanismos de transcrição, atuando como homodímero e heterodímero, sendo o principal parceiro de outros receptores (ZHANG et al., 1992). Os receptores para o ácido retinóico 9-cis apresentam diferentes isoformas: RXR α , β e γ , os quais diferem na região N-terminal (região A/B) e possivelmente controlam a expressão de diferentes genes (LEID et al., 1992).

Foi demonstrado que RXRs, podem se apresentar em diferentes estados oligoméricos (CHEN, 1998; KERSTEN, 1995a), monômeros, dímeros e tetrâmeros, em função da concentração, da presença de seu ligante, ácido 9-cis-retinóico e de elementos responsivos. Estes estudos foram realizados com uma forma não fisiológica da proteína RXR α , possuindo uma deleção na região A/B (RXR $\alpha\Delta$ AB).

A existência de dímeros de RXR $\alpha\Delta$ AB em solução, na ausência do ligante 9-cis-RA, é observada desde os primeiros trabalhos sobre oligomerização destas proteínas. Chen et al. (1994) mostraram que a produção de RXR $\alpha\Delta$ AB de camundongo (mRXR $\alpha\Delta$ AB), através de expressão em *E. coli*, gerava proteínas ativas com relação à ligação a elementos responsivos e ao ligante, 9-cis-RA. Kersten et al. (1995a)

utilizando este protocolo de purificação, mostraram que na concentração de 50 μ M, mRXR α Δ AB apresenta monômeros, dímeros e tetrâmeros, em PAGE 10%. Também por *crosslinking*, foi observada a presença das três espécies oligoméricas. Neste mesmo trabalho, utilizando-se de medidas de anisotropia por fluorescência, e da análise através do programa BIOEQS (ROYER et al., 1990; ROYER, 1993), as constantes de dissociação de tetrâmeros em dímeros e dímeros em monômeros são calculadas como sendo 155 nM e 4.4 nM. Com isso, os autores mostram que tetrâmeros estariam presente em solução em concentrações a partir de 20 nM e seriam a espécie dominante acima de 70 nM. Sendo assim, a existência de dímeros em solução seria observável apenas em uma faixa de concentração da ordem de poucas dezenas de nM. A inconsistência destes dados com o que foi observado nos experimentos anteriores, de PAGE 10% e *crosslinking*, é justificada em função de perturbações que ocorreriam no equilíbrio entre as espécies devido à eletroforese. Segundo os autores, "*This quantitative discrepancy is likely to stem from perturbation of the equilibrium distribution between the various protein species by prolonged electrophoresis*".

Em trabalho posterior de Kersten et al. (1995b), a existência de dímeros, em solução e na ausência de ligante, também é observada. Os autores mostram que a ligação do 9-cis-RA ao tetrâmero de mRXR α Δ AB faz com que este se dissocie e que dímeros passem a ser a espécie dominante. Em experimento de *crosslinking* e eletroforese em gel nativo, dímeros são observados na ausência do ligante. A observação de dímeros da proteína e

a discussão do efeito da migração em gel é retomada em trabalhos posteriores (KERSTEN et al., 1997a; KERSTEN et al., 1998):

"(...) resolution of RXR by gel mobility-shift assays and even more so by non-denaturing PAGE showed a significant dimeric fraction even at high receptor concentrations. It should be noted, however, that electrophoresis tends to lead to dissociation of protein complexes resulting in shifting of the equilibrium position."

Nestes trabalhos, a purificação da proteína era feita em uma única etapa de cromatografia por afinidade conforme definido por CHEN et al. (1994). Em 1998, Chen et al. mostraram que a introdução de uma etapa de cromatografia de exclusão por tamanho, após a etapa de afinidade, separa a população em duas espécies, com massa de 165 kDa e 42 kDa, respectivamente tetrâmeros e monômeros, calculadas a partir do volume de eluição. Neste trabalho, a espécie resultante da associação do tetrâmero ao ligante 9-cis-RA é caracterizada como sendo monomérica. Anteriormente, Kersten et al. (1995b) havia caracterizado essa espécie como sendo um dímero compacto, com mobilidade, em gel nativo, intermediária entre a de monômero e o de dímero. Essa caracterização foi reforçada por medidas de anisotropia por fluorescência, que mostraram um valor de anisotropia intermediário para dímeros e monômeros. Ainda em 1998, Kersten et al. sugerem que a ligação do 9-cis-RA dissocia tetrâmeros em dímeros e os estabiliza em uma conformação favorável a ligação ao DNA.

A existência de dímeros, segundo estes trabalhos, fica praticamente descartada em concentrações de μM na ausência de ligante e elemento responsivo. Entretanto, Fischer et al. (2003) observaram a existência de dímeros de RXR $\alpha\Delta\text{AB}$ humana (hRXR $\alpha\Delta\text{AB}$) em solução, para concentrações bem acima destes valores. Isso foi feito através do cálculo de raio hidrodinâmico, por cromatografia de exclusão por tamanho e DLS, e raio de giro, por SAXS. Também foram propostos modelos estruturais para o dímero, a 30 Å de resolução, e o tetrâmero, a 17 Å.

Os RXR ligam-se ao DNA como homo- ou heterodímeros, regulando a transcrição de diversos genes. As sequências AGGTCA e AGTTCA são preferencialmente reconhecidas por receptores de retinóides, de hormônio tiroideano e vitamina D. RXRs, atuando como homodímero reconhece preferencialmente o motivo AGGTCA, em elementos responsivos do tipo repetição direta com um espaçamento (DR1) e, no caso da isoforma β , também em palíndromos sem espaçamento (GLASS, 2005). Foi mostrado que a ligação de tetrâmeros de RXRs a DNAs também pode ocorrer (CHEN et al., 1995; LIN et al., 1997), destacadamente para elementos responsivos formados por quatro repetições diretas do motivo AGGTCA, separadas por uma base C.

1.2 Microscopia Eletrônica

1.2.1 Formação da imagem

Assim como toda matéria, elétrons apresentam um

comportamento de partícula ou onda, dependendo das condições à qual são submetidos e observados (EISBERG; RESNICK, 1988). Sendo assim, em circunstâncias controladas, um feixe de elétrons pode se comportar como uma onda. Neste caso, a energia dos elétrons irá determinar o comprimento de onda, que poderá, portanto, ser ajustado em função da aceleração dos mesmos. Deste modo, a análise do funcionamento de um microscópio eletrônico e da formação da imagem pode ser feita considerando-se o feixe de elétrons como uma onda, com comprimento de onda definido pela sua energia.

Na microscopia por transmissão eletrônica, um feixe de elétrons é produzido e focalizado em uma amostra, para análise da mesma. Isto é feito através de um conjunto de bobinas que produzem campos magnéticos, atuando como lentes para a onda de elétrons. A passagem do feixe através da amostra, produz uma imagem de projeção, ampliada, do objeto analisado. Estas imagens podem ser observadas através de uma tela fosforescente ou uma câmera CCD, também podendo ser gravadas em um filme fotográfico ou adquiridas digitalmente.

A formação da imagem, na microscopia eletrônica, se dá pela interação dos elétrons com o objeto analisado, podendo ser através de um espalhamento elástico ou inelástico. No caso de amostras biológicas, a formação da imagem se dá através da utilização de contraste por fase. A onda eletrônica sofre deslocamentos de fase em função da interação com a amostra, ocasionando interações destrutivas e construtivas no plano da

imagem. Embora não seja possível analisar diretamente as diferenças de fase originadas, essas dão origem à diferenças de intensidades, que podem ser observadas.

Formalmente, para uma onda incidente, que se desloca na direção z , o espalhamento elástico é descrito como sendo um deslocamento de fase $\Phi(\vec{r})$, dado por

$$\Phi(\vec{r}) = \int C(\vec{r}, z) dz \quad (1)$$

onde \vec{r} é um vetor bidimensional e $C(\vec{r}, z)$ é a distribuição do potencial Coulombiano dentro do objeto. Sendo assim, uma onda plana incidente $\Psi_0 = \exp(ikz)$, ao interagir com uma amostra, terá sua fase modificada, dando origem à onda espalhada

$$\Psi(\vec{r}) = \Psi_0 \exp[i\Phi(\vec{r})] \quad (2)$$

Utilizando-se a aproximação de fase fraca, onde $\Phi(\vec{r}) \ll 1$, a Equação 2 pode ser reescrita através da expansão

$$\Psi(\vec{r}) = \Psi_0 \left[1 + i\Phi(\vec{r}) - \frac{1}{2}\Phi(\vec{r})^2 + \dots \right] \quad (3)$$

mostrando que após sofrer um espalhamento elástico, a onda resultante possuirá dois termos, o primeiro é a própria onda incidente e outro é a onda espalhada. Nesse caso, desprezando-se os termos de segunda ordem e maiores, obtêm-se

$$\Psi(\vec{r}) = \Psi_0 + \Psi_0 i\Phi(\vec{r}) \quad (4)$$

Na microscopia de transmissão, a função de onda, no plano

focal posterior às lentes objetivas, é a transformada de Fourier da equação 3 ou mesmo da sua forma simplificada na equação 4.

$$\Psi_{pf}(\vec{k}) = F\{\Psi(\vec{r})\} \quad (5)$$

onde F indica a transformada de Fourier do argumento.

Entretanto, a aberração esférica das lentes e o defoco causam um deslocamento de fase adicional na onda espalhada, sendo

$$\gamma(\vec{k}) = 2\pi\chi(\vec{k}) \quad (6)$$

onde $\chi(\vec{k})$ é a função de aberração da onda (*wave aberration function*).

A aberração esférica das lentes faz com que uma onda plana incidente seja transformada em uma onda esférica deformada, que não irá convergir perfeitamente para um único foco. Devido a esse fenômeno, as regiões externas da onda terão sua distância focal diminuída com relação às suas regiões internas.

A função de onda no plano focal posterior será, portanto,

$$\Psi_{pf}(\vec{k}) = F\{\Psi(\vec{r})\} \exp[i\gamma(\vec{k})]$$

A onda ainda sofrerá uma modificação devido à passagem por uma abertura, com uma função $A(\vec{k})$, onde

$$A(\vec{k}) = 1 \text{ para } |k| = \theta/\lambda \leq \theta_1/\lambda \quad (7)$$

$$A(\vec{k}) = 0 \text{ para os demais valores de } |k|$$

A função de onda no plano da imagem será, portanto

$$\Psi_i(\vec{r}) = F^{-1} \left\{ F \left\{ \Psi(\vec{r}) \right\} A(\vec{k}) \exp[i\gamma(\vec{k})] \right\} \quad (8)$$

onde F^{-1} indica a transformada inversa de Fourier. A distribuição de intensidades no plano da imagem, que é a imagem propriamente dita, será então

$$I(\vec{r}) = |\Psi_i(\vec{r})|^2 \quad (9)$$

1.2.2 Processamento das imagens

Após a aquisição e digitalização das imagens, inicia-se o ciclo de processamento. Uma ampla discussão sobre o processamento de dados, para a obtenção de estruturas de macromoléculas por microscopia eletrônica, pode ser encontrada em Frank (2006). Como primeiro procedimento, imagens de partículas devem ser selecionadas nas imagens adquiridas. Esse procedimento pode ser interativo ou semi-automatizado. Programas como Boxer (LUDTKE et al., 1999) e Tyson (PLAISIER et al., 2004), auxiliam na seleção de partículas. Diferentes metodologias de detecção de partículas podem ser utilizadas nestes programas. Uma revisão sobre esse assunto é apresentada por Nicholson e Glaeser (2001). Em seguida é feita a correção do CTF (*Contrast Transfer Function*), que é necessária para inversão das fases das imagens obtidas. Nessa etapa, as imagens também são filtradas e normalizadas de maneira a se ter um conjunto de dados homogêneo.



Figura 3. Imagem de rosto humano à qual foi adicionado ruído, gerando a figura ao lado.

Após estas etapas, inicia-se o processamento propriamente dito. As imagens obtidas possuem baixa relação sinal-ruído, sendo assim, é necessário, como primeiro passo que se some imagens semelhantes no sentido de se aumentar a relação sinal-ruído. A Figura 3 mostra uma imagem utilizada como exemplo, à qual foi adicionado ruído. A imagem original pode ser recuperada através da soma de diversas imagens com baixa relação sinal-ruído, como mostrado na Figura 4.

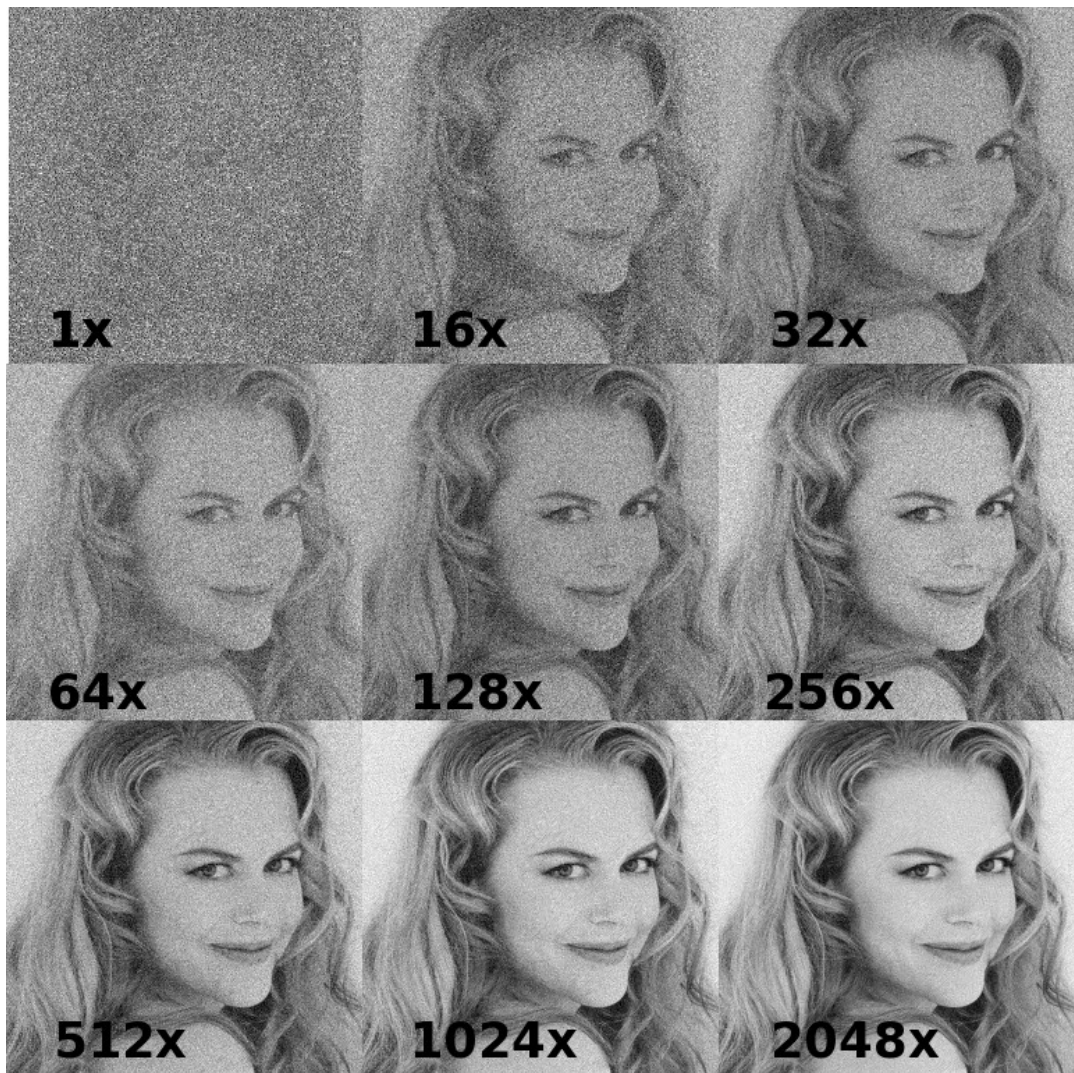


Figura 4. Resultado da somatória de imagens com ruído, como a da Figura 3. Os números indicam o total de imagens somadas.

Entretanto, para que as imagens possam ser somadas, de maneira a melhorar a relação sinal-ruído do resultado e preservar o seu conteúdo, é necessário que estejam alinhadas, translacionalmente e rotacionalmente. Para isso, as imagens precisam ser divididas em pequenos grupos (classes) que contenham imagens com a mesma orientação da molécula. Somas de imagens de uma mesma classe são chamadas *classsums*.

O alinhamento inicial das imagens deve ser feito sem

referências, de maneira a não introduzir tendências no conjunto de dados. Um dos processos de alinhamento livre de referência é o chamado alinhamento por classificação. Neste processo, as imagens são inicialmente centralizadas, através de translações com relação a soma total de todas as imagens do conjunto. A Figura 5 mostra a soma total de 25 imagens de rosto humano. Em seguida, procedimentos de classificação por estatística multivariada (MSA – *Multivariate Statistical Analysis*) (VAN HEEL; FRANK, 1981) são utilizados para achar imagens similares, que estejam em orientações similares, visto que nenhum alinhamento rotacional foi realizado até esta etapa. As médias de imagens das classes resultantes desse procedimento são utilizadas como referências para um primeiro alinhamento multi referência (MRA – *Multireference Alignment*) (VAN HEEL; STÖFFLER-MEILICKE, 1985).



Figura 5. Soma total de imagens de rosto humano.

O MRA é um processo que irá alinhar translacionalmente e

rotacionalmente um conjunto de imagens, baseado em um conjunto menor de referências. A Figura 6 mostra um conjunto de imagens antes de ser submetido a alinhamento, utilizando o processo de MRA, tendo como referências o conjunto de imagens da Figura 5. O resultado pode ser visto na Figura 7.



Figura 6. Imagens desalinhadas translacionalmente e rotacionalmente.



Figura 7. Imagens mostradas na Figura 6 após MRA utilizando como referência o conjunto de imagens mostrado na Figura 5.

Após este alinhamento, serão realizados novos ciclos de: classificação por MSA; cálculo de médias das imagens de uma mesma classe; seleção de médias características para serem utilizadas como referências e MRA (SCHATZ et al., 1997).

No procedimento de classificação por MSA as imagens do conjunto são consideradas como combinação linear de outras imagens, chamadas auto-imagens (*eigen-images*). A Figura 8 mostra as auto-imagens obtidas através de MSA para o conjunto de imagens mostrado na Figura 6.

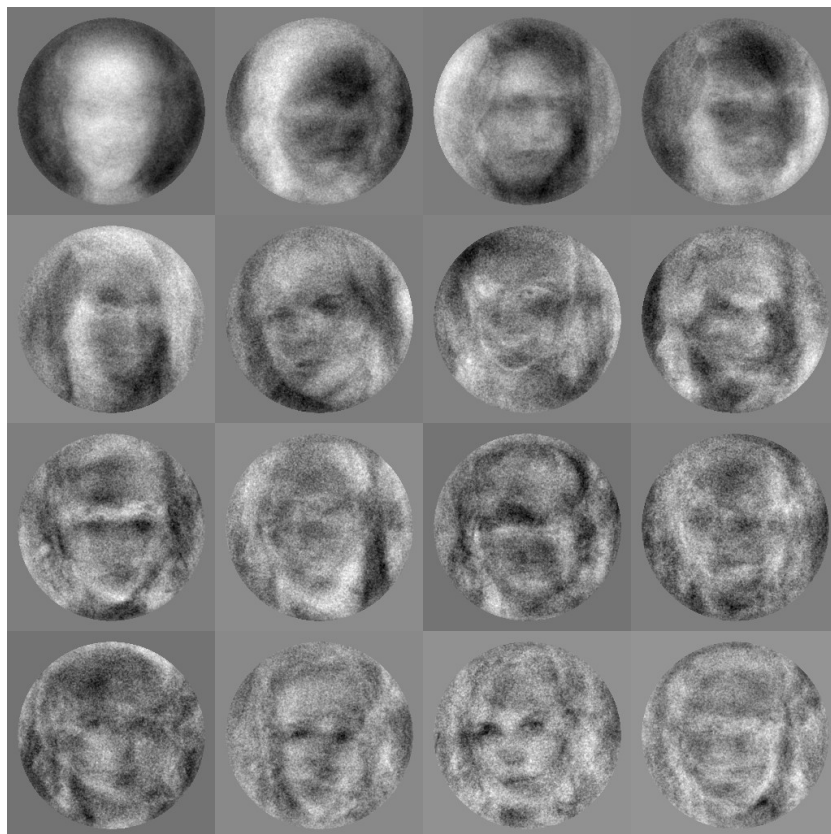


Figura 8. Auto-imagens do conjunto de rostos humanos.

Após ciclos de MSA-MRA, conforme descrito anteriormente, serão obtidas boas imagens médias (*classums*), com relação sinal-ruído

melhorada em relação ao conjunto de dados. Estas podem ser utilizadas na etapa seguinte de reconstituição angular (SERYSHEVA et al., 1995; SCHATZ et al., 1995, VAN HEEL et al., 2000). Neste procedimento, os ângulos de Euler das projeções, que definem sua orientação espacial, são determinados, obtendo-se uma relação angular entre as imagens escolhidas.

Baseado na relação angular entre as imagens, o modelo estrutural pode ser obtido. Este primeiro modelo, será utilizado para gerar novas projeções bidimensionais, que serão utilizadas, tanto na validação do modelo, quanto na atribuição de ângulos e também como referências para novos MRAs. A melhora na qualidade das referências, faz com que o alinhamento seja mais eficiente, melhorando a qualidade das classes obtidas e, em última instância, do modelo estrutural. Ciclos de MSA-MRA com construção do modelo e reprojeções, são então utilizados para o refinamento da estrutura.

1.3 Termodinâmica de proteínas

Em proteínas em solução, vários processos de equilíbrio podem estar presentes simultaneamente (CANTOR; SCHIMMEL, 1969). A constante de equilíbrio de cada um destes processos está relacionada com uma variação da energia livre de Gibbs padrão (ΔG^0) da reação, pela seguinte equação:

$$\Delta G^0 = - RT \ln(K) \quad (10)$$

ou

$$\Delta G^0 = RT \ln(K_d) \quad (11)$$

onde K_d é a constante de dissociação do processo, R é a constante universal dos gases ($1,985 \text{ cal.mol}^{-1}.\text{K}^{-1}$) e T é a temperatura (K). A equação 11 pode ser reescrita como sendo:

$$\Delta G^0 / T = R \ln(K_d) \quad (12)$$

A variação da energia livre de Gibbs padrão, para uma temperatura constante, é definida por:

$$\Delta G^0 = \Delta H^0 - T \Delta S^0 \quad (13)$$

que pode ser reescrito como:

$$\Delta G^0 / T = \Delta H^0 (1/T) - \Delta S^0 \quad (14)$$

onde ΔH^0 é a variação da entalpia e ΔS^0 é a variação da entropia. Através da equação 12, observa-se que, dado o K_d de um processo, pode-se calcular $\Delta G^0 / T$. Uma vez que essa constante seja medida em diferentes temperaturas, pode-se analisar graficamente a variação do valor de $\Delta G^0 / T$ em função de $1 / T$, utilizando-se equação 14. A reta obtida terá como coeficiente linear, $-\Delta S^0$, e como coeficiente angular, ΔH^0 . O gráfico de $\Delta G^0 / T$ em função de $1 / T$ é chamado de gráfico de van't Hoff.

Neste trabalho, a formação do complexo entre $\text{hRXR}\alpha\Delta\text{AB}$ e elementos responsivos será analisada em diferentes temperaturas para se investigar as contribuições entrópicas e entálpicas no processo de formação do complexo.

1.4 Espectroscopia de fluorescência

1.4.1 Fluorescência em proteínas

A espectroscopia de fluorescência utiliza-se da fluorescência intrínseca e/ou extrínseca de proteínas para análise de diversos processos, entre eles, desenovelamento, oligomerização e ligação de ligantes. A técnica de fluorescência e suas aplicações são abordadas de maneira ampla por Lakowicz (1983). A fluorescência intrínseca é propiciada pelos resíduos de triptofano e tirosina que possuem a propriedade de fluorescência natural. A fluorescência extrínseca é obtida através da ligação de sondas à proteína, tais como bis-Anilinaftaleno sulfonado (bis-ANS) e fluoresceína. Estas sondas são utilizadas com diferentes grupos reativos, para que haja uma ligação seletiva a diferentes sítios de interesse na proteína. Estes fluoróforos, intrínsecos ou extrínsecos, quando excitados com luz em comprimento de onda específicos, emitem luz, sob a forma de fluorescência, em energia proporcional à constante dielétrica do meio onde os mesmos se encontram.

O resíduo de triptofano é hidrofóbico, permanecendo preferencialmente no interior das proteínas, quando na forma nativa. Desta forma, através do monitoramento do sinal de fluorescência do triptofano, é possível monitorar a exposição destes resíduos ao solvente aquoso e, conseqüentemente a desestruturação da proteína. Esta medida é feita obtendo-se o espectro de emissão de fluorescência da proteína e

analisando-se os dados que podem ser extraídos desta curva, tipicamente o pico de máxima emissão e, mais convenientemente, o centro de intensidade de emissão do espectro, também chamado centro de massa do espectro de emissão.

Alterações no espectro de emissão são consequência de alterações locais nas proximidades dos resíduos fluorescentes. Entretanto, estas alterações costumam refletir, com certa precisão, o comportamento da estrutura terciária da proteína, podendo ser utilizada para monitorar, também, alterações de estrutura quaternária, desde que os resíduos observados estejam na interface das subunidades do oligômero.

Por sua vez, a utilização de sondas fluorescentes extrínsecas, ligadas em regiões de interesse da proteína, proporciona a observação de diversos parâmetros tais como estado de enovelamento, interação proteína-proteína, proteína-ligante e proteína-DNA. Utilizando-se diferentes técnicas de fluorescência, estes fenômenos podem ser acompanhados não só qualitativamente, mas também quantitativamente, podendo-se determinar, entre outros parâmetros, as constantes de associação/dissociação de oligômeros, proteína-ligante e proteína-DNA.

1.4.2 Ainsotropia de fluorescência

A fluorescência extrínseca pode ser utilizada de diversas maneiras, entre elas, para a realização de experimentos de anisotropia de fluorescência. Medidas de anisotropia podem ser utilizadas em análises

bioquímicas e fornecem informações sobre o tamanho e forma de proteínas ou sobre a mobilidade de diferentes ambientes moleculares. Este tipo de medida é utilizado também para investigar a associação proteína-proteína ou proteína-DNA.

Experimentos de anisotropia estão baseados no princípio de excitação fotoseletiva de fluoróforos pela luz polarizada. Fluoróforos absorvem preferencialmente fótons cujos vetores elétricos estejam alinhados paralelamente ao momento de transição do fluoróforo. Este, por sua vez, tem sua orientação definida com relação à molécula. Numa solução isotrópica, os fluoróforos estão orientados aleatoriamente. Com a incidência da luz polarizada, aqueles fluoróforos que tiverem seu dipolo de transição alinhados paralelamente ao vetor elétrico da luz incidente serão seletivamente excitados. Esta excitação seletiva resulta em uma população de fluoróforos parcialmente orientada (fotoseleção) e uma emissão fluorescente parcialmente polarizada. Esta emissão também ocorre com a luz polarizada ao longo de um eixo fixo do fluoróforo. O ângulo relativo entre estes momentos determina a máxima anisotropia. Anisotropia e polarização são definidas por (WEBER, 1952):

$$r = \frac{(I_{||} - I_{\perp})}{(I_{||} + 2 I_{\perp})} \quad (15)$$

$$P = \frac{(I_{||} - I_{\perp})}{(I_{||} + I_{\perp})} \quad (16)$$

onde r é a anisotropia, P é a polarização, $I_{||}$ e I_{\perp} são as intensidades de fluorescência da emissão polarizada verticalmente ($||$) e

horizontalmente (\perp), quando a amostra é excitada com luz polarizada verticalmente. Alguns fenômenos podem diminuir a máxima anisotropia medida em uma amostra, sendo o mais comum a difusão rotacional. Essa difusão ocorre durante o tempo de vida do estado excitado e desloca o dipolo de emissão do fluoróforo. A intensidade deste deslocamento dependerá da difusão e portanto, do tamanho e/ou mobilidade da proteína a qual a molécula está ligada. Sendo assim, a observação da anisotropia de fluoróforos ligados ao DNA pode ser utilizada para estudos de associação à proteínas.

Neste trabalho, estudos de anisotropia de fluorescência serão utilizados para investigar a associação de hRXR α Δ AB aos elementos responsivos DR1, DR4, F2 e PAL, na presença e ausência de ligante (9-cis-RA) e em diferentes temperaturas. Para isso, serão utilizados oligonucleotídeos marcados covalentemente com fluoresceína na terminação 5'.

1.5 SAXS - Espalhamento de raios-X a baixo ângulo

A técnica de espalhamento de raios-X a baixo ângulo, SAXS, permite a obtenção de informações da estrutura de diversos materiais. Informações como flutuações ou heterogeneidades da densidade eletrônica com dimensões características da ordem de 10 a 1000 Å podem ser obtidas, por exemplo, para soluções de macromoléculas biológicas (proteínas, ribossomos, DNAs, etc).

A interpretação de dados de espalhamento a baixos ângulos de solução de macromoléculas permite a determinação do envelope molecular e, conseqüentemente, de estados de oligomerização. Dependendo da concentração destas soluções, podemos ter macromoléculas interagindo umas com as outras, efeito inter-partícula (GLATER, 1982). Isso gera uma correlação espacial que altera a intensidade espalhada, gerando um artefato no sinal, dificultando a análise dos dados. Sendo assim, embora o uso de mais altas concentrações seja desejável, para diminuir a dispersão do sinal, é importante que os sinais de baixas e altas concentrações sejam proporcionais, descartando a possibilidade de estar ocorrendo um efeito inter-partícula.

As soluções de macromoléculas biológicas apresentam pouca diferença na densidade eletrônica entre seus constituintes, que são essencialmente compostas por átomos leves (H, C, N, O e P no caso de ácidos nucleicos), e o solvente. Estes tipos de soluções produzem um espalhamento muito fraco necessitando usar feixes de raios-X produzidos por fontes síncrotron, que possuem um fluxo de fótons maior que aparelhos de raios-X convencionais.

A diferença entre a energia dos fótons de raios-X incidentes e a energia de ligação dos elétrons permite tratar estes últimos como elétrons livres. Desta forma, partículas livres carregadas oscilantes produzem ondas espalhadas coerentes com a onda incidente em todas as direções.

Em solução, as partículas de macromoléculas estão em

movimento browniano e podem assumir todas as orientações possíveis em relação ao feixe incidente. Assim, a intensidade $I(q)$, produzida por esse sistema, é proporcional a amplitude de espalhamento de uma macromolécula, promediado para todas as orientações. Sendo $q = 4\pi \sin(\theta)/\lambda$, onde 2θ é o ângulo entre o feixe incidente e o feixe espalhado e λ é o comprimento de onda utilizado. Quando os valores de q são pequenos, vale a aproximação de Guinier (GUINIER; FOURNET, 1955):

$$I(q) = I(0) \exp(-1/3 R_g^2 q^2) \quad (17)$$

sendo R_g o raio de giro. R_g é um parâmetro estrutural importante utilizado na análise de mudanças conformacionais de macromoléculas em solução. A equação 17 pode ser reescrita na forma:

$$\ln [I(q)] = - (1/3 R_g^2) q^2 + \ln [I(0)] \quad (18)$$

Sendo assim, ao se fazer a análise gráfica da variação logarítmica de intensidade produzida por um determinado sistema, $\ln [I(q)]$, em função de q^2 , obtém-se uma reta, onde o coeficiente angular é $-(1/3 R_g^2)$ e o coeficiente linear é $-\ln [I(0)]$. A partir do valor de $-\ln [I(0)]$, possuindo-se um parâmetro de comparação, pode-se calcular a massa molecular do sistema.

2 Objetivos

Este trabalho tem como objetivos o estudo e caracterização de diferentes complexos moleculares, formados por hRXR α Δ AB e hemocianina de *A. gomesiana*, através de técnicas biofísicas e estruturais.

Com relação à hemocianina de *A. gomesiana*, o objetivo é construir um modelo estrutural, através da técnica de crio-microscopia eletrônica, deste complexo molecular.

Com relação ao estudo de hRXR α Δ AB, os objetivos deste trabalho são a caracterização biofísica dos estados oligoméricos da proteína e da sua ligação a diferentes elementos responsivos, na presença e ausência de seu ligante. Em especial, pretende-se investigar a dimerização de hRXR α Δ AB em solução, na ausência de seu ligante.

3 Materiais e Métodos

3.1 Microscopia Eletrônica de Hemocianina de Acanthoscurria gomesiana

3.1.1 Preparação da amostra e microscopia

No presente trabalho, será apresentado e discutido um modelo estrutural para a hemocianina de *A. gomesiana*. Este modelo é resultado de uma colaboração entre o Departamento de Parasitologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, Laboratório de Artrópodes do Instituto Butantan, Centre for Biomolecular Electron Microscopy do Imperial College London (CBEM) e o Grupo de Cristalografia de Proteínas do Instituto de Física de São Carlos da Universidade de São Paulo.

A purificação da amostra foi feita pelo Dr. Daniel Macedo Lorenzini no Departamento de Parasitologia, Instituto de Ciências

Biomédicas, Universidade de São Paulo, através de hemolinfa fornecida pelo Dr. Pedro Ismael da Silva Jr. do Laboratório de Artrópodes do Instituto Butantan. O protocolo de purificação foi adaptado do utilizado por Decker (1998). A hemolinfa é extraída através de punção dorsal do coração, sendo imediatamente diluída em tampão Tris 200 mM, pH 8,0, CaCl₂ 10 mM, MgCl₂ 10 mM na proporção 1 (hemolinfa) : 4 (tampão) (v/v). Após a coleta, a hemolinfa é centrifugada a 800 g por 10 min. O sobrenadante é transferido e 200 µl do mesmo são utilizados para purificação por exclusão por tamanho em uma coluna Sepharose 6, com fluxo de 0,5 ml/min. São coletadas frações de 0,5 ml em um sistema HPLC (GE Healthcare) (LORENZINI et al., 2003). Em seguida a amostra é armazenada à 4°C, permanecendo nessa temperatura durante o transporte para o CBEM.

Após o recebimento da amostra no laboratório do CBEM, os ensaios de microscopia foram conduzidos pelo Dr. Willem Tichelaar para a obtenção de micrografias da amostra em gelo vítreo (crio-microscopia eletrônica). Uma gota da amostra purificada é aplicada a um filme fino de carbono sem furos. O excesso da amostra é absorvido com um papel de filtro e a amostra é imediatamente mergulhada e congelada em etano líquido. A partir do congelamento, as amostras são manipuladas e analisadas em temperatura de nitrogênio líquido (DUBOCHET et al., 1982; ADRIAN et. al., 1984). Foram feitas seis sessões de microscopia para aquisição de dados, sendo que duas destas resultaram em bons conjuntos de imagens. Outras sessões haviam sido conduzidas anteriormente para elaboração e otimização do protocolo a ser utilizado.

Para coleta dos dados foi utilizado um microscópio de transmissão eletrônica CM200 FEG (Philips), equipado com uma entrada lateral para porta amostras criogênico e câmera CCD 1024 x 1204 (Tietz Video Systems GmbH), operando com uma voltagem de aceleração de 200 kV. Foi utilizado um porta amostra criogênico, com reservatório de nitrogênio líquido (Gatan, UK). As amostras foram analisadas em condições de baixa dose, com magnificação real de 49.000 vezes e defoco entre 0,5 μm e 2,2 μm .

Duas sessões de aquisição de imagens forneceram conjuntos de imagens viáveis. Na primeira delas foram adquiridas 54 micrografias, das quais 16 foram eliminadas por apresentarem astigmatismo ou artefatos de deslocamento, restando 38 imagens viáveis. Destas, 12 micrografias foram utilizadas para a seleção de partículas, de onde se obteve um conjunto de dados com cerca de 11.000 partículas. Os valores de defoco utilizados na obtenção destas micrografias foram de -2,2 μm ; -1,9 μm ; -1,7 μm . Na outra sessão de microscopia, foram obtidas 53 imagens, com valores de defoco entre -0,5 μm e -2,2 μm . Estas imagens serão utilizadas em uma etapa futura do processamento para melhoria da resolução do modelo obtido, através do aumento do número de imagens no conjunto de dados.

3.1.2 Processamento das imagens

Após inspeção visual, as micrografias foram digitalizadas em um digitalizador Nikon 8000 Scanner com um passo de amostragem de 6.35

µm, o que representa, para uma magnificação de 49.000 vezes, um pixel de 1,3 Å na amostra.

Após digitalizadas, as imagens foram armazenadas no *cluster* de computadores do CBEM. O processamento foi realizado alternadamente no *cluster*, para uso dos programas que utilizam processamento paralelo, e em um *notebook*, para uso dos programas que não utilizam processamento paralelo. O *cluster* do CBEM é composto por cerca de 20 computadores com processadores AMD Athlon – 32 bits, 1 GB de memória RAM e sistema operacional Linux, distribuição Fedora Core³¹. O *notebook* utilizado foi um Compaq Presario R3000, com processador AMD 2,4 GHz 64bits, 1,3 GB de memória RAM – sistema operacional Linux, distribuição Mandriva 2006² – 64 bits. Posteriormente, no Grupo de Cristalografia do IFSC, o refinamento do modelo estrutural obtido foi feito utilizando-se o mesmo *notebook* e um *cluster* de cinco computadores Dell Dimension Intel Xeon 64 bits, 1 GB de memória RAM, sendo que um deles conta com dois processadores e 3 GB de memória. O *cluster* possui um total de 1,4 TB de armazenamento em disco e todos os computadores operam com o Sistema operacional Linux, distribuição Gentoo³ 64 bits. O processamento dos dados, em todos os computadores citados, foi feito utilizando-se o pacote de processamento IMAGIC-V (Image Science Software GmbH, Alemanha) (VAN HEEL et al., 1996; VAN HEEL et al., 2000) em suas versões para 32 e 64 bits.

Durante a etapa de seleção de partículas, foi utilizado o

1 Disponível em <http://fedora.redhat.com/>

2 Disponível em <http://www.mandriva.com/>

3 Disponível em <http://www.gentoo.org/>

programa Boxer do pacote de processamento EMAN⁴ (LUDTKE et al., 1999). As partículas foram selecionadas semi-automaticamente, utilizando uma metodologia baseada no cálculo da correlação cruzada entre partículas selecionadas pelo usuário e a imagem completa, para identificação da posição de outras partículas. As partículas foram selecionadas com tamanhos de 512 x 512 pixels e inspecionadas visualmente após a seleção para eliminação de partículas que estivessem próximas à borda da imagem ou a outras partículas. A correção de CTF foi feita através da inversão de fases, utilizando-se o pacote de processamento IMAGIC-V e outros módulos de programas desenvolvidos no CBEM. O valor do defoco da imagem, é calculado para 72 posições da imagem e então ajustados para uma equação de um plano que representa a variação dos valores de defoco ao longo da imagem. Esta equação é utilizada para cálculo aproximado do valor de defoco para cada partícula, em função da sua posição na imagem, que é utilizado na correção de CTF de cada partícula obtida individualmente (VAN HEEL et al., 2000). Após a correção de CTF, as imagens foram novamente cortadas para um tamanho final de 300 x 300 pixels.

O processamento das partículas, para obtenção do modelo estrutural, seguiu uma abordagem que utiliza reconstituição angular associada com análise estatística multivariada (MSA) para classificação das partículas (VAN HEEL, 1987; SCHATZ et al., 1997; VAN HEEL et al. 2000).

A dimensão das imagens de partículas foi alterada, através de promediação, para 150 x 150 pixels, tornando o tamanho do pixel de 2,6 Å.

4 Disponível em <http://ncmi.bcm.tmc.edu/~stevel/EMAN/doc>

Isso é feito no sentido de se diminuir o esforço computacional das etapas posteriores, sem implicar em diminuição da resolução final do modelo. Uma soma de todas as partículas é feita para estimativa de dimensões máximas da mesma. As partículas também foram submetidas a um filtro passa-banda com corte em 17 Å e 240 Å para eliminação de componentes de baixa frequência que podem interferir no processamento dos dados e também componentes de alta frequência que não são importantes para a obtenção de um primeiro modelo, em baixa resolução.

A técnica de MSA (VAN HEEL; FRANK, 1981; VAN HEEL et al., 1992; VAN HEEL et al., 2000) foi utilizada para classificação das partículas de maneira a identificar as projeções características da molécula. Essas projeções características são obtidas somando-se imagens que pertençam a uma mesma classe, após classificadas, chamadas em inglês de *classsums*. Primeiramente, o conjunto de partículas foi dividido em dois subconjuntos por MSA. Um deles possuía uma das projeções laterais, muito abundante no conjunto de dados, correspondendo a cerca de 20% do total de partículas. O outro subconjunto possuía as demais partículas. Novos MSA foram realizados para classificação interna dos dois subconjuntos e obtenção de suas projeções características.

Ciclos de MSA, MRA, atribuição de ângulos e reconstrução do modelo tridimensional, sem imposição de simetria, foram utilizados para a melhoria da qualidade das projeções características. As projeções utilizadas como referências, para a primeira atribuição de ângulos, foram extraídas de

um modelo 4 x 6-mero de hemocianina, derivado de uma estrutura de hemocianina de *L. Polyphemus* (8 x 6-mero). Para isso, foram obtidas 30 projeções do modelo, as quais foram filtradas para não possuírem detalhes de alta resolução e em seguida a qualidade das imagens foi diminuída através da adição de ruído.

Após obtenção de um modelo com resolução de 30 Å, foi imposta simetria C2 ao mesmo, e novas projeções, estas com um eixo de simetria de ordem 2, foram calculadas e utilizadas para o reinício do processamento dos dados. As partículas originais foram novamente filtradas, com parâmetros de corte em 280 Å e 5,2 Å, sendo o tamanho do pixel de 2,6 Å. O processamento dos dados foi novamente realizado através de ciclos de MSA, MRA, atribuição de ângulos e reconstrução do modelo tridimensional, desta vez com imposição de simetria C2. Esse modelo foi refinado até a resolução de 17 Å.

A estrutura do hexâmero (1 x 6-mero) da hemocianina de *P. interruptus* foi então ajustada ao modelo obtido. Para isso foi utilizado o programa Situs⁵ (WRIGGERS et al., 1999; WRIGGERS; CHACÓN, 2001).

Todas as figuras foram feitas utilizando-se o programa VMD – Visual Molecular Dynamics (HUMPHREY et al., 1996). Estimou-se um erro de 10 Å para as medidas de distâncias feitas sobre as imagens. Isto baseado no fato de que as imagens foram filtradas a 5,2 Å e de que cada medida de distância depende da definição de dois pontos.

⁵ Disponível em <http://www.biomachina.org/>

3.2 Expressão e purificação de hRXR α Δ AB

A proteína hRXR α Δ AB com cauda de poli-histidina foi expressada em *E. coli*, cepa BL21(DE3), após indução com 0,5 mM de IPTG por 3 horas, em meio de cultura LB na presença de 10 μ M ZnSO₄. As células hospedeiras são coletadas através de centrifugação e ressuspensas em tampão 50 mM fosfato de potássio, pH 8,2, 100 mM NaCl, 10% glicerol, 10 mM imidazol, 2 mM β -mercapto etanol, 1mM PMSF, para serem sonicadas a 4°C. Para cada 5 L de meio de cultura utilizado na expressão, 100 mL de tampão são utilizados para ressuspensão do *pellet*. Este é incubado em lisozima (concentração final de 500 μ g/ml) por 30 min a 4°C, para posterior sonicação por cerca de 7 ciclos de 10 segundos com intervalos de 10 segundos, mantendo a proteína 4°C durante o processo. Após esta etapa, a proteína permanece solúvel e é separada através de centrifugação.

Para cada 100 ml de sobrenadante da etapa anterior, são adicionados 15 ml de resina Talon Superflow e em seguida incubado por 1 hora a 4°C. Após incubação, a resina é empacotada numa coluna 16/10 (Amersham Bioscience). A resina é lavada com tampão 50 mM fosfato de potássio, pH 8,2, 50 mM NaCl, 10% glicerol, 10 mM imidazol, 2 mM β -mercapto etanol. A lavagem da resina e eluição é feita a 4°C, 1 ml/min, frações de 0,5 ml e acompanhada através do perfil cromatográfico por absorção em 280 nm. Após a lavagem com dois volumes de coluna e estabilização da absorção a 280 nm, a proteína é eluída com o mesmo tampão, contendo 300 mM de imidazol. As frações que contêm a proteína

recombinante são separadas através da análise do perfil cromatográfico do HPLC.

A amostra, em seguida à etapa de cromatografia de afinidade, é submetida a uma etapa de purificação por exclusão por tamanho, onde é eluída em frações de 5,5 ml. Para isso, utiliza-se uma coluna HiLoad Superdex 200 26/60 (Amersham Bioscience) e o perfil cromatográfico é utilizado para se isolar as frações de interesse. A eluição é feita em tampão 20 mM HEPES, pH8,0, 200 mM NaCl, 5% glicerol, 1 mM DTT. A proteína pode então ser congelada através de processo de resfriamento rápido em nitrogênio líquido e armazenada -80°C. Para clivar a cauda de poli-histidina, a proteína é quantificada pelo método de Bradford (BRADFORD, 1976) e, em seguida, incubada por 8 horas à 18°C com trombina (Amersham Bioscience), numa proporção de 0,5 unidade de trombina para 100 µg de proteína. A pureza das amostras é analisada através de SDS-PAGE 15%. Para ambas etapas cromatográficas de purificação, foi utilizado um sistema AKTA Purifier da Amersham Bioscience.

Para descongelamento da proteína, as alíquotas são deixadas em equilíbrio a 4°C, até que se liquefaçam. Em seguida são centrifugadas por 10 minutos a 13.000g em uma centrífuga mantida a 4°C. O sobrenadante é retirado e a concentração é medida pelo método de Bradford (BRADFORD, 1976).

3.3 Oligonucleotídeos

Foram utilizados cinco diferentes elementos responsivos: DR1, repetição direta com um espaçamento, DR4, repetição direta com 4 espaçamentos, F2, palíndromo invertido com seis espaçamentos, e PAL, palíndromo sem espaçamento, sendo que o elemento responsivo DR1 foi utilizado em duas formas, chamadas de *curto* e *longo*, conforme descrito na Tabela 1, onde a sequência consenso está sublinhada:

Tabela 1. Descrição dos elementos responsivos utilizados

HRE	Sequência
DR1 longo	5' – AGCTTCAGGTCAGAGGTCAGAG – 3'
DR1 curto	5' – TAGGTCAAAGGTCAG – 3'
DR4	5' – AGCTTCAGGTCACAGGAGGTCAGAG – 3'
F2	5' – AGCTTATTGACCCCAGCTGAGGTCAAGTTACG – 3'
PAL	5' – ATATTCAGGTCATGACCTGAATAT – 3'

Estes oligonucleotídeos foram utilizados nos estudos de oligomerização e também para a determinação da especificidade de reconhecimento de elementos responsivos por parte de hRXR α Δ AB. Para os estudos de oligomerização foram utilizados DR1 *longo* e *curto*. Nos estudos de determinação de especificidade foram utilizados DR1 *longo*, referido apenas como DR1, DR4, F2 e PAL, todos modificados, possuindo fluoresceína ligada covalentemente à sua porção 5'.

3.4 Estudos de oligomerização

3.4.1 Ensaio em gel nativo com gradiente de poliacrilamida

Para a realização de gel nativo com gradiente de poliacrilamida foi utilizado o PastSystem (Amersham Pharmacia) com géis de gradiente de 8% a 25%. O sistema possui controle de temperatura e todas as medidas foram realizadas em temperaturas entre 4°C e 8°C. Para os ensaios com DNA, os géis foram corados primeiramente com brometo de etídio e em seguida, com *coomassie blue*. Para os demais experimentos em gel, estes foram corados apenas com *coomassie blue*.

Para os estudos de migração em gel, foram utilizados as alíquotas provenientes do Pico 1 e Pico 2, de maneira a se estudar separadamente, cada uma das populações. Para o primeiro conjunto de medidas as amostras foram descongeladas, segundo procedimento de descongelamento, e em seguida incubadas na presença de 9-cis-RA (2 mM); DR1 curto (400 pmol/μl); 9-cis-RA (2 mM) e DR1 curto (400 pmol/μl); e na ausência de ambos, ligante e elemento responsivo. A proteína foi incubada a 50 μM. A incubação foi feita a 4°C. Decorrida 1 hora de incubação, 4 μl de cada amostra foram aplicados em gel no sistema PhastSystem. O sistema permite que se trabalhe com dois géis simultaneamente, sendo assim, os géis para análise do Pico 1 e Pico 2 foram feitos ao mesmo tempo. As amostras continuaram sendo incubadas a 4°C, até completarem um

período de incubação de 12 horas, quando novos géis foram preparados seguindo o mesmo procedimento. Estes géis foram corados com brometo de etídio e em seguida, com *coomassie blue*.

Num segundo conjunto de medidas, as amostras foram descongeladas e, antes de serem incubadas, permaneceram em equilíbrio a 4°C, durante 2 horas. Em seguida foram incubadas por 1 hora nas mesmas condições anteriores, exceto pela preparação de duas outras amostras. Foram preparadas também: DR1 longo (400 pmol/ μ l); 9-cis-RA (2 mM) e DR1 longo (400 pmol/ μ l). Todas estas medidas foram feitas no tampão de eluição da cromatografia de exclusão por tamanho: 20 mM Hepes, pH 8,0, 200 mM NaCl, 5% glicerol, 3 mM DTT.

A análise da migração das espécies oligoméricas, e associações a monômeros, dímeros e tetrâmeros foi feita por comparação com a proteína hTR β Δ AB, que foi anteriormente estudado em nosso grupo (DIAS, 2004), isto por se tratar de uma proteína homóloga, com geometria similar, que apresenta os três estados oligoméricos.

Todas estes experimentos foram realizados em colaboração com a Dra. Maria Auxiliadora Morim Santos do Grupo de Cristalografia do IFSC.

3.4.2 Desnaturação de hRXR α Δ AB por uréia

A desnaturação de hRXR α Δ AB foi acompanhada por dicroísmo circular (CD) (FASMAN, 1996) e também por eletroforese de gel de

gradiente de acrilamida. As amostras foram equilibradas em concentrações crescentes de uréia, sendo 0 M, 1 M, 2 M, 3 M e 4 M para as amostras analisadas por gel de acrilamida e 0 M, 1 M, 2 M, 3 M, 4 M, 5 M, 6 M, 7 M e 7,2 M, para o dicroísmo circular. As amostras (40 μ M para o gel de poliacrilamida e 5 μ M para o dicroísmo circular) foram equilibradas por 2 horas a 18° C. As medidas de CD foram realizadas a temperatura ambiente e o gel a 4° C.

Para os géis nativos foi utilizado o sistema PhastSystem (Amersham Pharmacia) e para as medidas de dicroísmo circular o epectropolarímetro Jasco J-720 (Jasco) do grupo de biofísica molecular do IFSC. As medidas de CD foram realizadas durante um estágio no LAPA, Departamento de Bioquímica Médica, UFRJ, em colaboração com a Profa. Débora Foguel, o Prof. Victor Moreau e a Profa. Astria Ferrão-Gonzales.

3.4.3 Cromatografia analítica de exclusão por tamanho

A cromatografia analítica de exclusão por tamanho foi realizada em uma coluna TSK-3000 com hXR α Δ AB (7 μ M) em tampão 10 mM Tris-HCl pH 8,0, 50 mM NaCl, 0,5 mM EDTA e 1 mM DTT. Foram utilizados para calibração, proteínas com as seguintes massas moleculares: 13,5 kDa; 16,0 kDa; 18,0 kDa; 18,0 kDa; 33,0 kDa; 33,0 kDa; 55,0 kDa, 67,0 kDa. Estas medidas foram realizadas durante um estágio no LAPA, Departamento de Bioquímica Médica, UFRJ, em colaboração com a Profa. Débora Foguel, o

Prof. Victor Moreau e a Profa. Astria Ferrão-Gonzales.

3.4.4 Espalhamento de raios-X a baixo ângulo

As medidas de espalhamento de raios-X a baixo ângulo foram realizadas no LNLS, Campinas –SP, onde utilizou-se a linha de SAXS1. Utilizou-se como comprimento de onda incidente $\lambda = 1,48 \text{ \AA}$. Para aquisição de dados utilizou-se um detector 1D sensível a posição (DSP). Com ele e com uma eletrônica associada multicanal, que faz a discretização dos canais, adquiriu-se a curva de intensidade de espalhamento, $I(q)$, em função do número de canal, q . O espalhamento parasita do ar, janelas e das fendas foi subtraído da intensidade total espalhada. Estes estudos foram realizados em colaboração com o aluno de doutorado Mário Oliveira Neto, do Grupo de Cristalografia, do Instituto de Física da USP/São Carlos.

Os resultados obtidos foram analisados utilizando-se a aproximação de Guinier, para valores pequenos de q , segundo a equação 18. foram medidas amostras de hRXR α Δ AB a 3,5 mg/ml, 7,2 mg/ml e BSA a 1 mg/ml, ambas a 4°C. Os valores obtidos para ambas as concentrações de hRXR α Δ AB se mostraram proporcionais à concentração, permitindo o uso da mais alta concentração para a análise dos dados.

3.5 Reconhecimento de elementos responsivos

A interação da proteína hRXR α Δ AB com diferentes elementos

responsivos foi analisada na presença e ausência do seu ligante, 9-cis-RA, através da técnica de espectroscopia por anisotropia de fluorescência. Quatro diferentes elementos responsivos foram utilizados: DR1, repetição direta com um espaçamento; DR4, repetição direta com quatro espaçamentos; F2, palíndromo invertido com seis espaçamentos; PAL, palíndromo sem espaçamento, no caso de DR1, foi utilizado o *longo*, conforme Tabela 1.

As medidas foram realizadas em um espectrômetro de fluorescência ISS – PC1 (ISS), equipado com polarizadores. Para controle da temperatura foi utilizado um banho de circulação RTE-740 (Thermo Neslab). Estas medidas contaram com a colaboração do aluno de iniciação científica Leandro Rizzi, IFSC - USP.

Todas as medidas foram realizadas em tampão 20 mM HEPES, pH 8,0, 5% glicerol, 3 mM DTT. Foram utilizadas diferentes concentrações de NaCl do tampão: 50 mM, 100 mM, 200 mM e 400 mM. A excitação foi feita em 480 nm e a leitura de intensidade era feita sem o uso de monocromadores, com um filtro OG515, com corte em 515 nm. A fonte foi utilizada em 18 A e a temperatura do recipiente da cubeta foi mantida a 10°C.

A proteína foi titulada com concentrações entre 0 e 200 nM de monômero, na presença de 1 nM de DNA marcado. No caso das medidas na presença de ligante, foi utilizado 300 nM de 9-cis-RA. As constantes de dissociação (K_d) foram calculadas para o sistema hRXR α Δ AB e os diferentes

elementos responsivos através do ajuste da equação 19 (LIMA; SILVA, 2004), onde Af e Ai são os valores de anisotropia final e inicial; RXR é a concentração molar dimérica da proteína e D é a concentração total de DNA, correspondente à soma do DNA livre com o DNA ligado.

$$An = Ai + \frac{(Af - Ai) * [K_d + RXR + D - \sqrt{(K_d + RXR + D)^2 - 4RXR * D}]}{2D} \quad (19)$$

4 Resultados e discussões

4.1 Modelo estrutural da hemocianina de *Acanthoscurria gomesiana*

Nesta parte do trabalho, será apresentado e discutido o modelo estrutural de hemocianina de *A. gomesiana*. Este trabalho foi desenvolvido em colaboração, sendo que, a extração da hemolinfa foi feita pelo Dr. Pedro Ismael da Silva Jr., do Laboratório de Artrópodes do Instituto Butantan, a purificação da proteína pelo Dr. Daniel Macedo Lorenzini, no Departamento de Parasitologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, e a obtenção de imagens de microscopia pelo Dr. Willem Tichelaar, do CBEM, Imperial College London. Coube ao autor deste trabalho as etapas posteriores à aquisição das micrografias, sendo elas: seleção e digitalização das mesmas e processamento dos dados.

O modelo aqui apresentado é um complexo formado por quatro hexâmeros (4 x 6-mero), conforme já observado para outra espécie de

tarântula, *E. californicum* (BIJHOLT et al., 1982; DE HAAS; VAN BRUGGEN, 1994; DECKER et al., 1996; VOIT et al., 2000). A Figura 9 mostra algumas partículas obtidas a partir das imagens de microscopia eletrônica. Cada imagem de partículas têm dimensões de 390 Å x 390 Å, sendo que a vista superior das partículas, e portanto do modelo obtido, tem dimensões aproximadas de 244 ± 10 Å x 244 ± 10 Å e as duas vistas laterais de 244 ± 10 Å x 105 ± 10 Å. Estes valores são compatíveis com os valores observados por de Haas e van Bruggen (1994) para a hemocianina de *E. californicum*, sendo 240 Å x 240 Å e 240 Å x 100 Å, respectivamente.

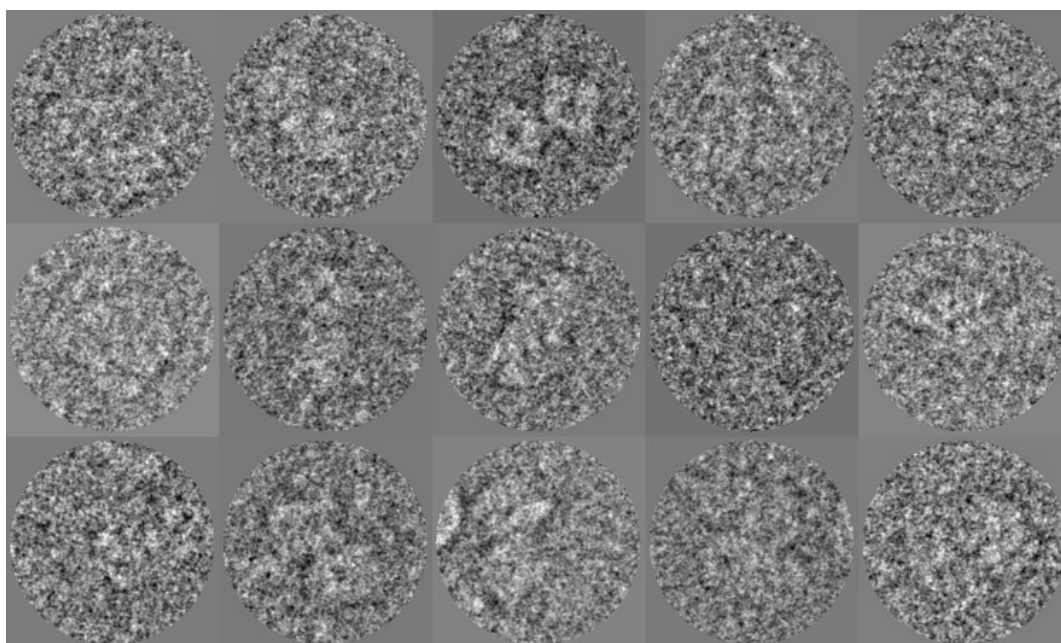


Figura 9. Partículas da hemocianina de *A. gomesiana*. Cada imagem possui 150 pixels x 150 pixels, com dimensões de 390 Å x 390 Å. As dimensões das partículas observadas são de aproximadamente 244 ± 10 Å x 244 ± 10 Å, para a vista superior e 244 ± 10 Å x 105 ± 10 Å, para as duas vistas laterais.

Nesta etapa do trabalho foram utilizadas 12 micrografias com valores de defoco entre -1,7 µm -2,2 µm, num total de 11.547 imagens. A

Tabela 2 mostra as características das imagens utilizadas.

Tabela 2: Caracterização das imagens utilizadas.

Número	Número de Partículas	Defoco	
		Nominal (μm)	Aferido (μm)
1	781	-1,15	-2,2
2	1004	-1,15	-2,2
3	922	-1,15	-2,2
4	536	-1,15	-2,2
5	754	-0,96	-1,9
6	1038	-0,77	-1,7
7	1077	-0,77	-1,7
8	933	-0,77	-1,7
9	1144	-0,77	-1,7
10	1018	-0,77	-1,7
11	1032	-0,96	-1,9
12	1308	-0,96	-1,9

As imagens foram submetidas à ciclos de MSA, MRA, atribuição de ângulos e reconstrução do modelo. Inicialmente, após a centralização das imagens, foi feita a classificação por MSA. A Figura 10 mostra as 20 primeiras auto-imagens obtidas de um total de 70 resultantes do MSA. A primeira auto-imagem mostra uma média de todas as partículas do conjunto. Por se tratarem de dados não alinhados rotacionalmente, a segunda e terceira imagens estão rotacionadas uma com relação à outra. Estas duas auto-imagens também indicam a presença de simetrias no conjunto de dados, neste caso, um eixo de ordem 2.

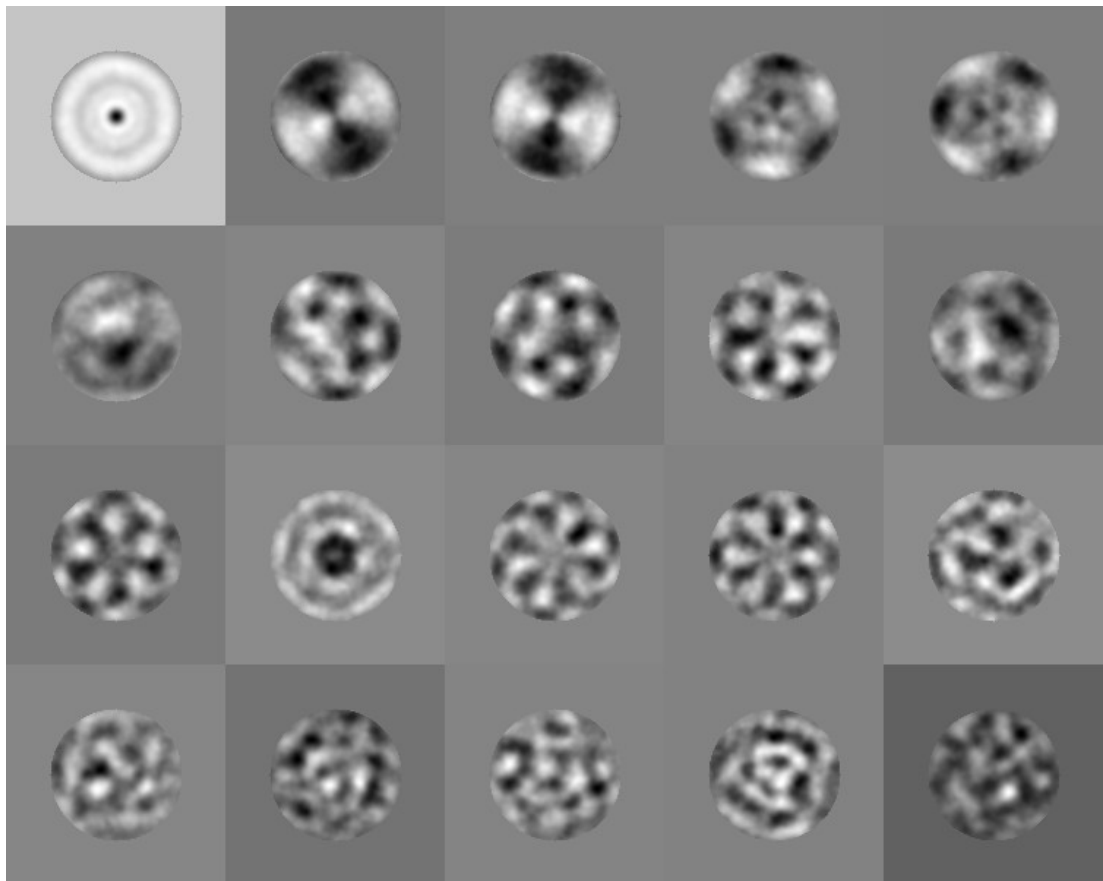


Figura 10. Auto-imagens da hemocianina de *A. gomesiana*. 20 primeiras auto-imagens do primeiro ciclo de processamento dos dados, sem o primeiro alinhamento rotacional.

O conjunto de partículas foi então agrupado em 400 classes com uma média de aproximadamente 30 imagens por classe. Os membros das mesmas classes foram somados para obtenção de projeções características da hemocianina (*classums*). Dentre as somas obtidas, foram selecionadas algumas para serem imagens de referência para o MRA, realizando o primeiro alinhamento rotacional. A Figura 11 mostra as imagens utilizadas para isso.

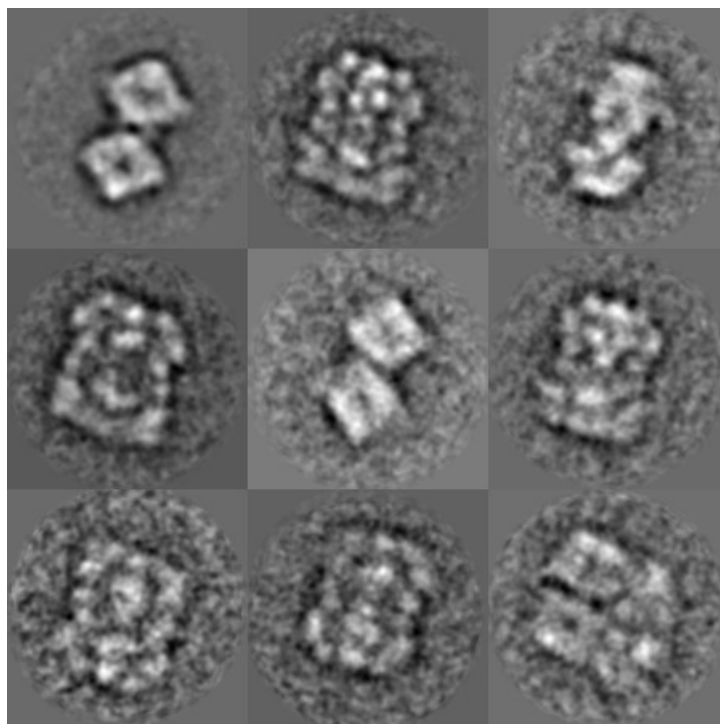


Figura 11. Referências utilizadas para o primeiro MRA. As demais partículas do conjunto de dados foram alinhadas translacionalmente e rotacionalmente com relação a estas referências.

Após alguns ciclos de MSA-MRA um conjunto de projeções típicas foi selecionado para reconstrução de um modelo da proteína. Para isso, foi feita a primeira reconstituição angular, utilizando-se um outro conjunto de imagens como referência angular. Foram utilizadas 30 projeções derivadas de um modelo, 4 x 6-mero, extraído de um modelo de hemocianina de *L. Polyphemus* (8 x 6-mero). Estas referências foram filtradas, para diminuir o nível de detalhes de alta resolução, e em seguida a qualidade das imagens foi diminuída através da adição de ruído. Após a atribuição de ângulos às partículas, o modelo foi reconstruído. A Figura 12 mostra o conjunto de projeções típicas que foram utilizadas para reconstrução do primeiro modelo da hemocianina de *A. gomesiana*.

Após ciclos de refinamento, que envolvem novos MSA, MRA,

reconstituição angular e construção, foi obtido um modelo com resolução de 30 Å, sem utilização de imposição de simetria. A este modelo, foi imposta simetria C2 e em seguida novas projeções, estas com um eixo de simetria de ordem 2, foram calculadas e utilizadas para o reinício do processamento dos dados.

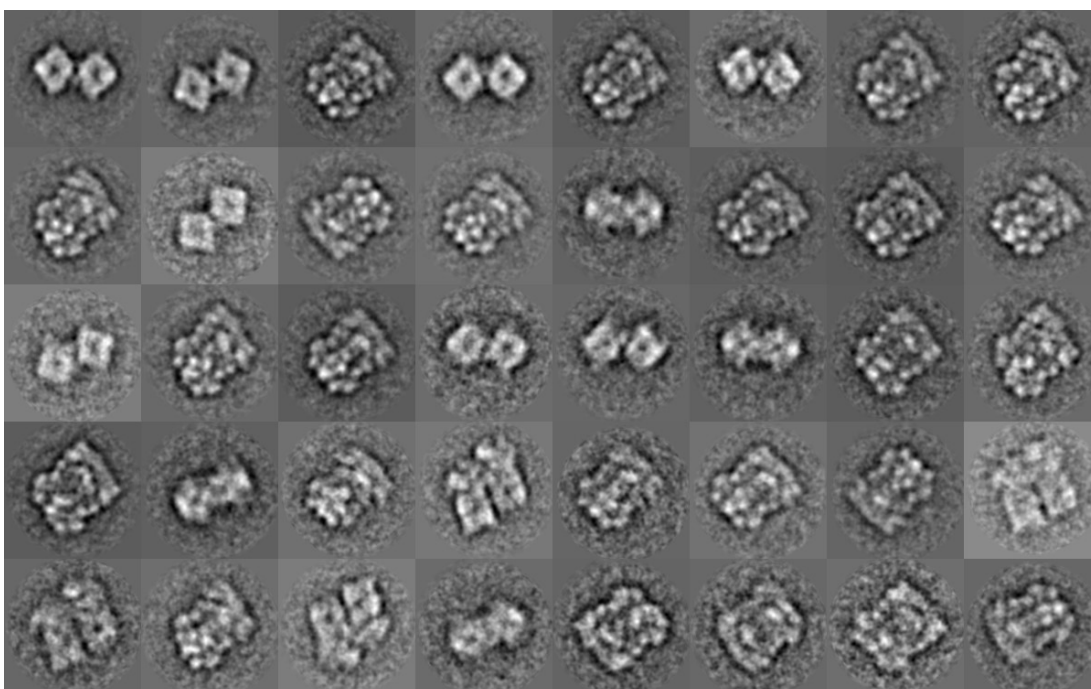


Figura 12. Projeções típicas utilizadas para a primeira reconstituição angular.

Para o início de um novo ciclo de processamento, as partículas originais foram novamente filtradas, dessa vez com maior resolução, 280 Å a 5,2 Å, sendo o tamanho do pixel de 2,6 Å. Em seguida, o processamento dos dados foi realizado, desta vez com imposição de simetria C2, obtendo-se um modelo com resolução de 17 Å.

A Figura 13 mostra, em cada uma das linhas, projeções da

hemocianina de *A. gomesiana* em diferentes orientações. A primeira coluna mostra imagens de uma partícula única, retirada do conjunto de dados. Na segunda coluna, está o resultado da soma de partículas pertencentes a mesma classe, após o MSA. Na terceira coluna, estão as reprojeções obtidas a partir do modelo tridimensional. A comparação entre as reprojeções do modelo e as somas de partículas permite a validação da reconstrução tridimensional.

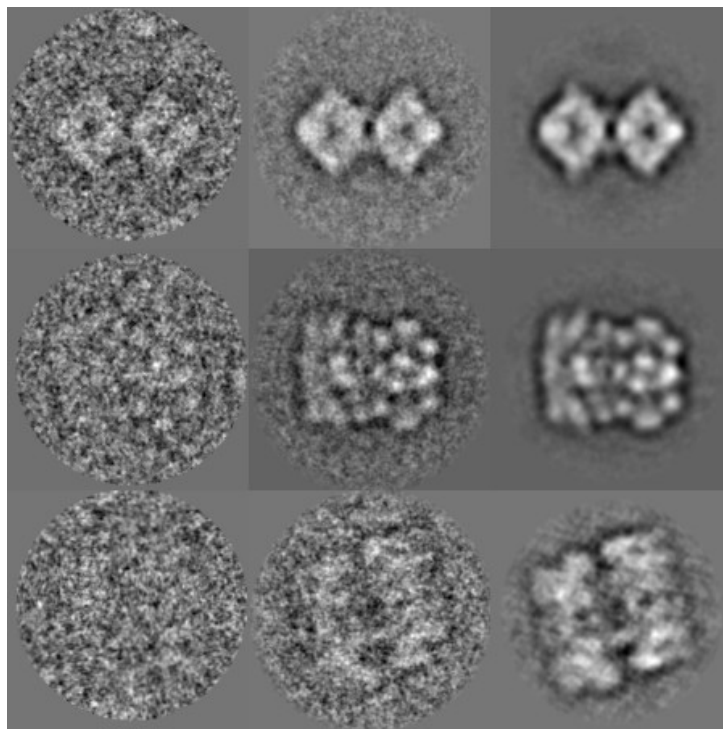


Figura 13. Comparação entre partículas, médias de partículas da mesma classe e reprojeções do modelo da hemocianina de *A. gomesiana*. Cada linha mostra uma determinada projeção em diferentes orientações. A primeira coluna mostra imagens individuais retiradas do conjunto de dados. A segunda coluna mostra as somas de imagens alinhadas, pertencentes à mesma classe, com clara melhoria na relação sinal-ruído. A terceira coluna mostra as reprojeções, geradas à partir do modelo tridimensional nas mesmas orientações atribuídas às imagens anteriores.

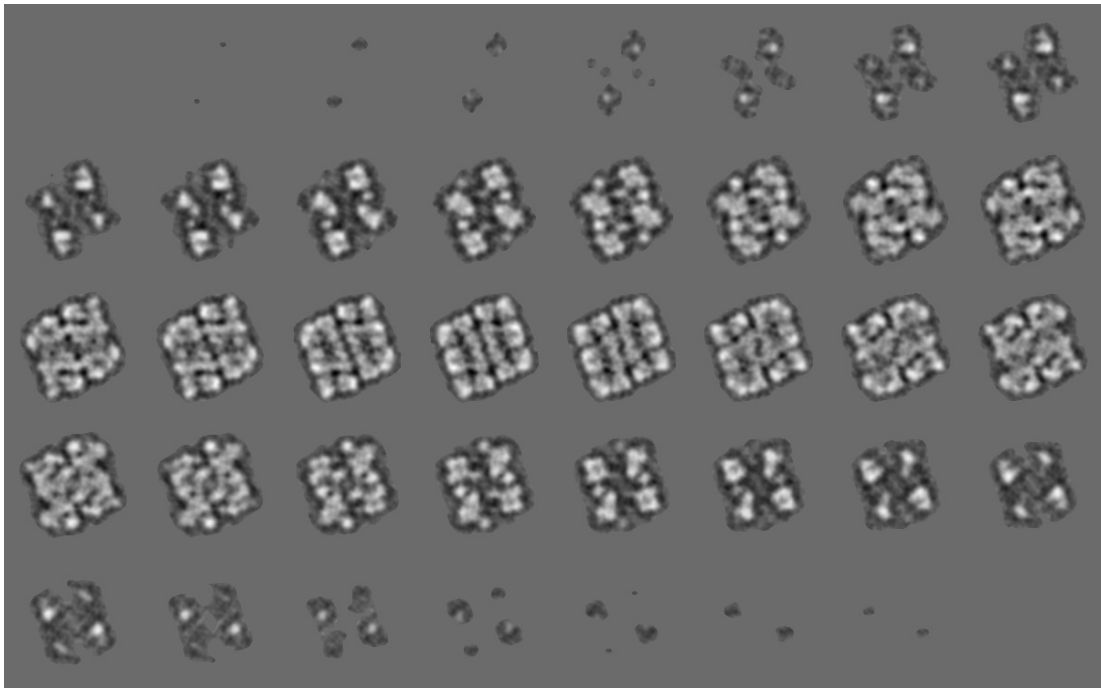


Figura 14. Planos do modelo estrutural de hemocianina de *A. gomesiana*, ao longo do eixo de simetria.

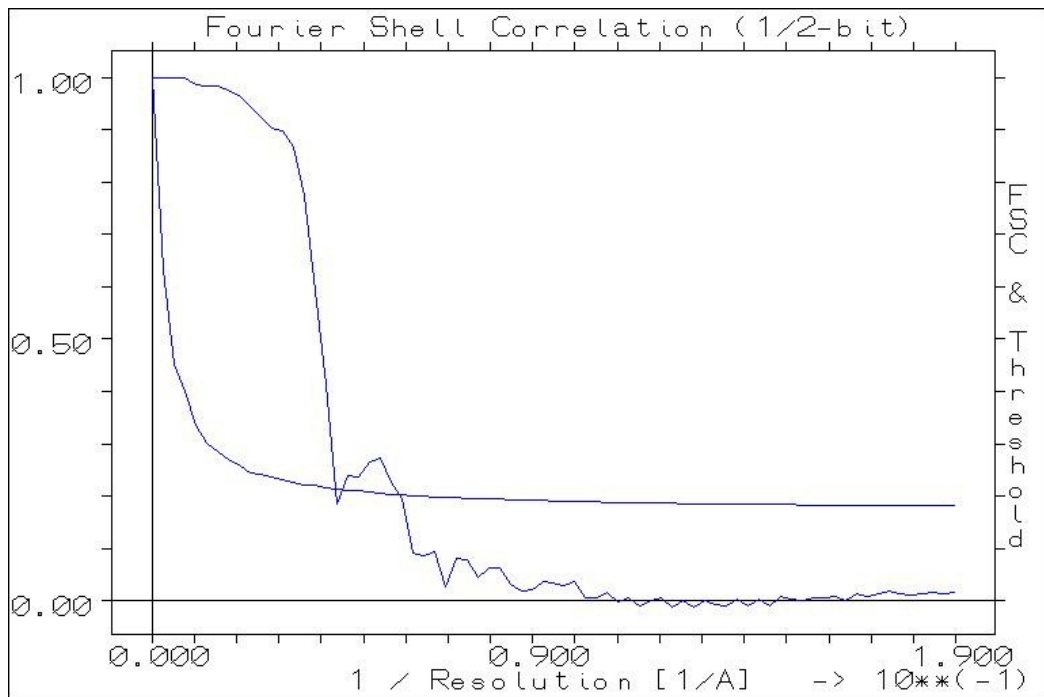


Figura 15. Resolução obtida para o modelo de hemocianina de *A. gomesiana*. O método utilizado foi o de Fourier Shell Correlation com o critério de $\frac{1}{2}$ bit. A resolução atingida pode ser calculada através do ponto de interseção entre as duas curvas. A primeira interseção em $0,043 \text{ 1/Å}$ pode ser descartada por ter origem em um artefato gerado pelos valores de defoco utilizados para obtenção do conjunto de dados. O valor utilizado para o cálculo da resolução foi de $0,059 \text{ 1/Å}$ que equivale à aproximadamente 17 Å .

A Figura 14 mostra o modelo final obtido, através de planos ao longo do eixo de simetria. A resolução do modelo foi calculada pelo método de *Fourier Shell Correlation* (FSC) usando o critério de $\frac{1}{2}$ bit (VAN HEEL; SCHATZ, 2005). A Figura 15 mostra o resultado da análise de resolução utilizando-se o programa de FSC⁶ do pacote IMAGIC. A primeira interseção, que ocorre em 0,043 1/Å, não foi considerada por se tratar de um artefato gerado pelos valores de defoco utilizados na obtenção das imagens. O conjunto de imagens, utilizado para esta etapa do processamento, possui valores de defoco próximos uns aos outros. Isso faz com que, mesmo após a soma de imagens obtidas com diferentes defocos, a função de transferência de contraste ainda possua zeros. Essas regiões produzem falhas de informação, ou seja, para determinadas resoluções não há informação disponível. Isso pode ser observado no gráfico de FSC através de falhas na função, como a que pode ser vista na região entre 0,04 1/Å e 0,05 1/Å. Sendo assim, a primeira interseção entre as curva a ser considerada é a que ocorre em torno de 0,059 1/Å, que representa uma resolução de aproximadamente 17 Å.

A Figura 16 mostra uma visão estéreo do modelo obtido em diferentes orientações. (A) visão em perspectiva; (B) vista superior, com o eixo z saindo do plano, x na direção horizontal e y na vertical; (C) vista lateral com o eixo x na horizontal, z na vertical e y saindo do plano; (D) vista lateral com o eixo y na horizontal, z na vertical e x saindo do plano. As dimensões observadas para esse modelo são aproximadamente um

⁶ Disponível em <http://www.imagescience.de/>

paralelepípedo de $244\pm 10 \text{ \AA} \times 244\pm 10 \text{ \AA} \times 105\pm 10 \text{ \AA}$. A distância observada entre os eixos do dodecâmero é de $117\pm 10 \text{ \AA}$. Estas dimensões são compatíveis com valores observados anteriormente (DE HAAS; VAN BRUGGEN, 1994; DECKER et al., 1996), sendo 119 \AA para a distância interdodecâmero e $240 \text{ \AA} \times 240 \text{ \AA} \times 100 \text{ \AA}$ para o paralelepípedo. Sabe-se no entanto que hemocianinas possuem uma variação de estrutura quaternária apreciável, quando comparados seus estados oxi e deoxi (DECKER et al., 1996; ERKER et al., 2005). Neste estudo, trabalhou-se em atmosfera de N_2 durante o congelamento da amostra em gelo vítreo, o que é inerente à técnica. Sendo assim, embora nenhum outro controle adicional tenha sido implementado para garantir que a amostra se encontrasse no estado deoxi, é razoável supor que a maioria da população tenha sido congelada neste estado.

Para verificar a consistência estrutural do modelo, foi feito o ajuste da estrutura do hexâmero da hemocianina de *P. interruptus* (PDB 1HC1, 1HC2, 1HC3, 1HC4, 1HC5 e 1HC6) resolvida a $3,2 \text{ \AA}$ por Volbeda e Hol (1989). Foram feitos quatro ajustes independentes para cada par de hexâmeros simétricos do modelo. Com isso, buscou-se que dentre os quatro melhores ajustes, houvesse dois ajustados, um em cada dodecâmero. Os ajustes foram feitos usando o programa Situs⁷ (WRIGGERS et al., 1999; WRIGGERS; CHACÓN, 2001). De fato, para um dos pares de hexâmeros simétricos do modelo, os dois primeiros melhores ajustes eram, um em cada hexâmero do modelo e simétricos um ao outro. Para o outro par de

⁷ Disponível em <http://www.biomachina.org/>

hexâmeros do modelo, das quatro melhores soluções, o segundo e terceiro ajustes foram em hexâmeros opostos e ocorreram de maneira simétrica. Embora a estrutura do hexâmero utilizada não corresponda exatamente a que se espera para *A. gomesiana*, o fato de se obter um bom ajuste, mostra um relativo sucesso no refinamento do modelo, embora esse ainda possa ser refinado. Considerando-se todas as micrografias coletadas e a média de imagens adquiridas até o momento por micrografia, podemos estimar um conjunto de cerca de 50.000 imagens disponíveis para o trabalho. Destas, apenas cerca de 10.000 foram utilizadas até o momento, com valores de defoco próximos uns aos outros. Ao se adicionar as demais imagens, espera-se obter uma melhoria da resolução do modelo em função da melhoria da relação sinal-ruído.

Após o ajuste do hexâmero de hemocianina de *P. interruptus* ao modelo obtido, o posicionamento das sub-unidades foi atribuído conforme estudos anteriores (DUBE et al., 1995; VOIT et al. 2000). A Figura 17 mostra o posicionamento proposto das sub-unidades ajustadas ao modelo de hemocianina da *A. gomesiana*. As sete sub-unidades podem ser vistas, sendo que todos os hexâmeros possuem as sub-unidades *a* (verde), *d* (azul-claro), *e* (marrom), *f* (azul-escuro) e *g* (vermelho). As sub-unidades *b* (cinza) e *c* (amarelo) aparecem simetricamente em apenas dois hexâmeros.

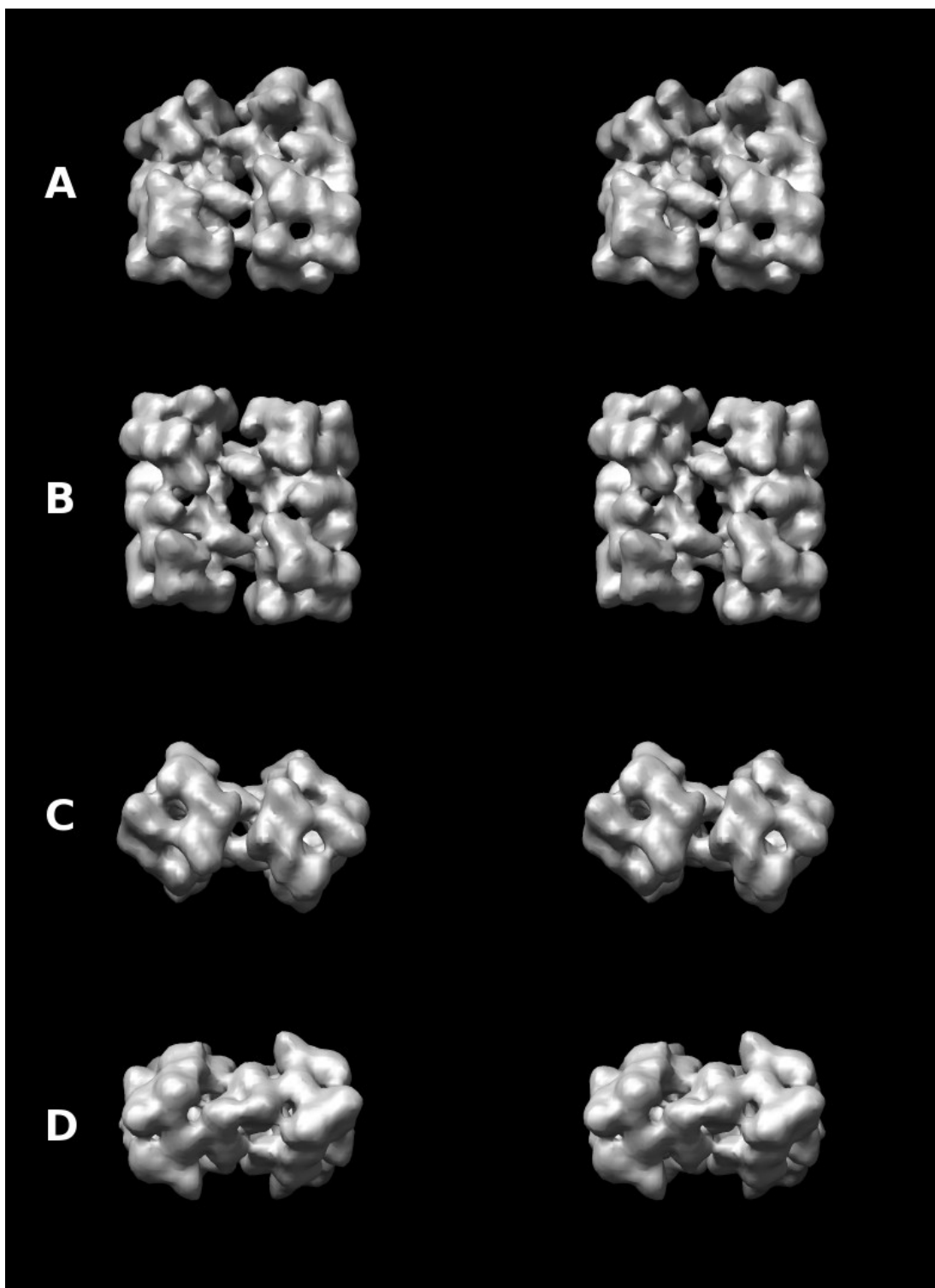


Figura 16. Visão estéreo do modelo de hemocianina de *A. gomesiana* em diferentes orientações. (A) visão em perspectiva. (B) vista de superior, com o eixo z saindo do plano, x na direção horizontal e y na vertical. (C) vista lateral com o eixo x na horizontal, z na vertical e y saindo do plano. (D) vista lateral com o eixo y na horizontal, z na vertical e x saindo do plano.

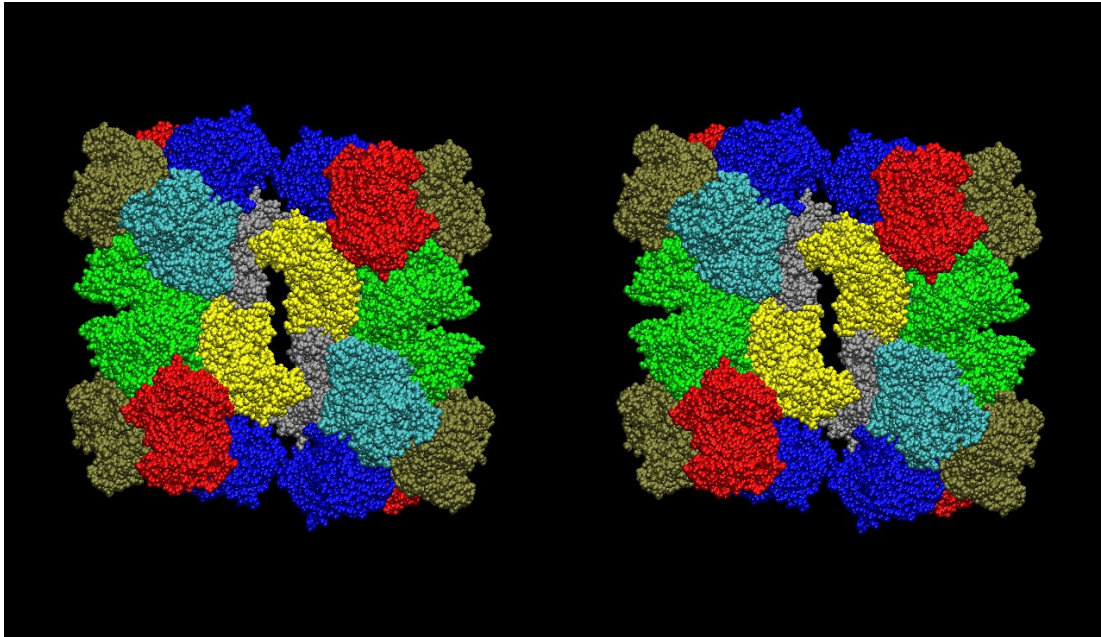


Figura 17: Visão estéreo da provável organização das sub-unidades da hemocianina de *A. gomeisana*. Sub-unidade *a* (verde), sub-unidade *b* (cinza), sub-unidade *c* (amarelo), sub-unidade *d* (azul-claro), sub-unidade *e* (marrom) sub-unidade *f* (azul-escuro) e sub-unidade *g* (vermelho).

A Figura 18 mostra o ajuste proposto das estruturas cristalográfica ao modelo obtido, nas diferentes orientações mostradas na Figura 16. Na primeira linha, tem-se o modelo proposto, em seguida, na segunda linha, uma sobreposição do modelo e da estrutura cristalográfica e na terceira linha somente a estrutura cristalográfica. As sub-unidades estão representadas nas mesmas cores citadas anteriormente.

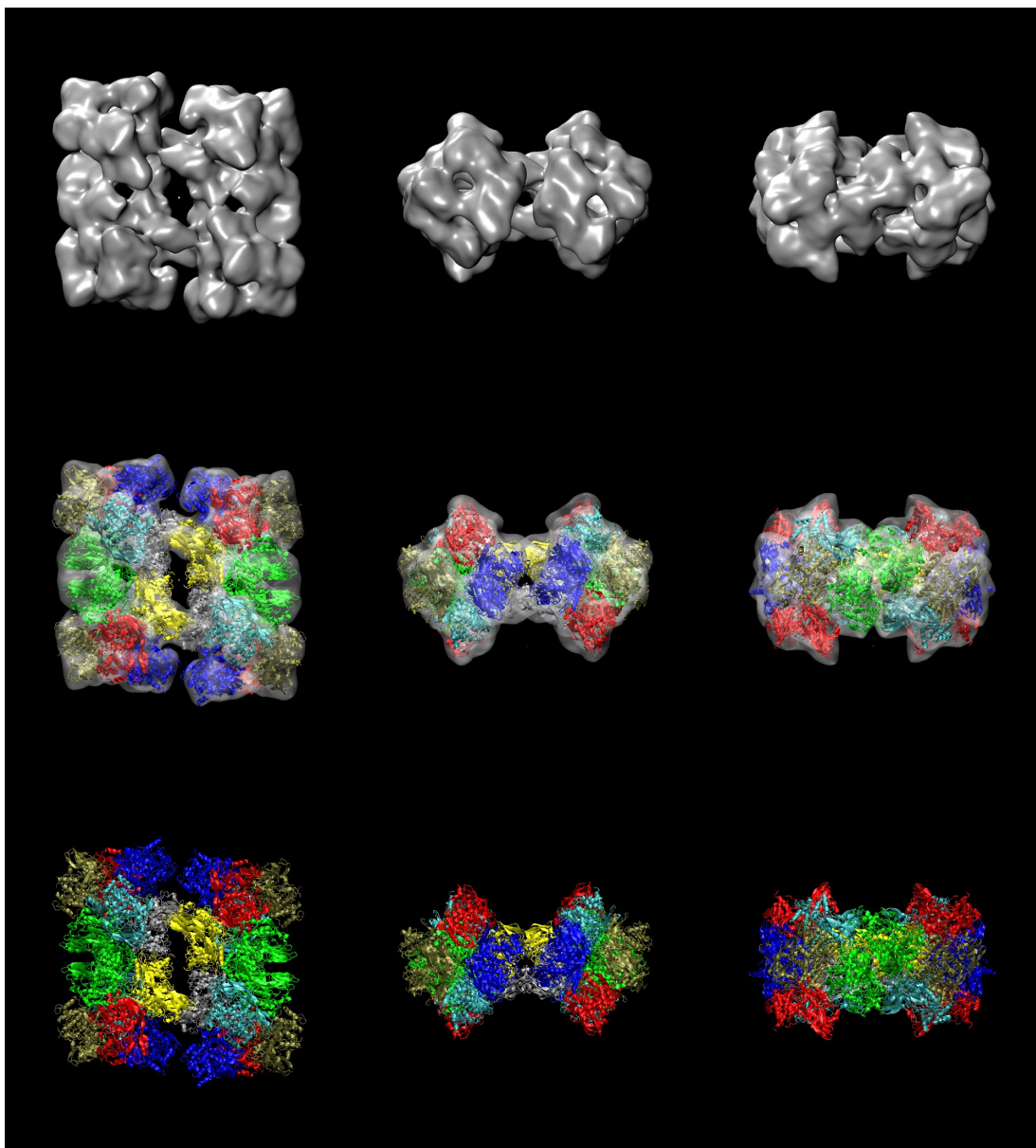


Figura 18. Estrutura cristalográfica do hexâmero da hemocianina de *P. Interruptus* (VOLBEDA; HOL, 1989), Figura 1, ajustado ao modelo obtido de hemocianina de *A. gomesiana*. As três colunas mostram diferentes orientações do modelo. A primeira linha mostra as diferentes vistas do modelo, conforme descrito na Figura 16. A segunda linha mostra o ajuste de quatro hexâmeros ao modelo do complexo obtido. A terceira linha mostra apenas as estruturas cristalográficas ajustadas.

Após o ajuste, as interfaces de oligomerização foram analisadas. Como a estrutura cristalográfica utilizada foi a de um hexâmero, as interfaces intra-hexaméricas não foram analisadas, somente as formadas entre os hexâmeros, conforme mostrado na Figura 19.

A associação entre os dois dodecâmeros é feita por duas pontes (Figura 19B). Cada uma das pontes é formada por duas ligações, sendo uma entre o domínio 3 da sub-unidade *b* e o domínio 2 da sub-

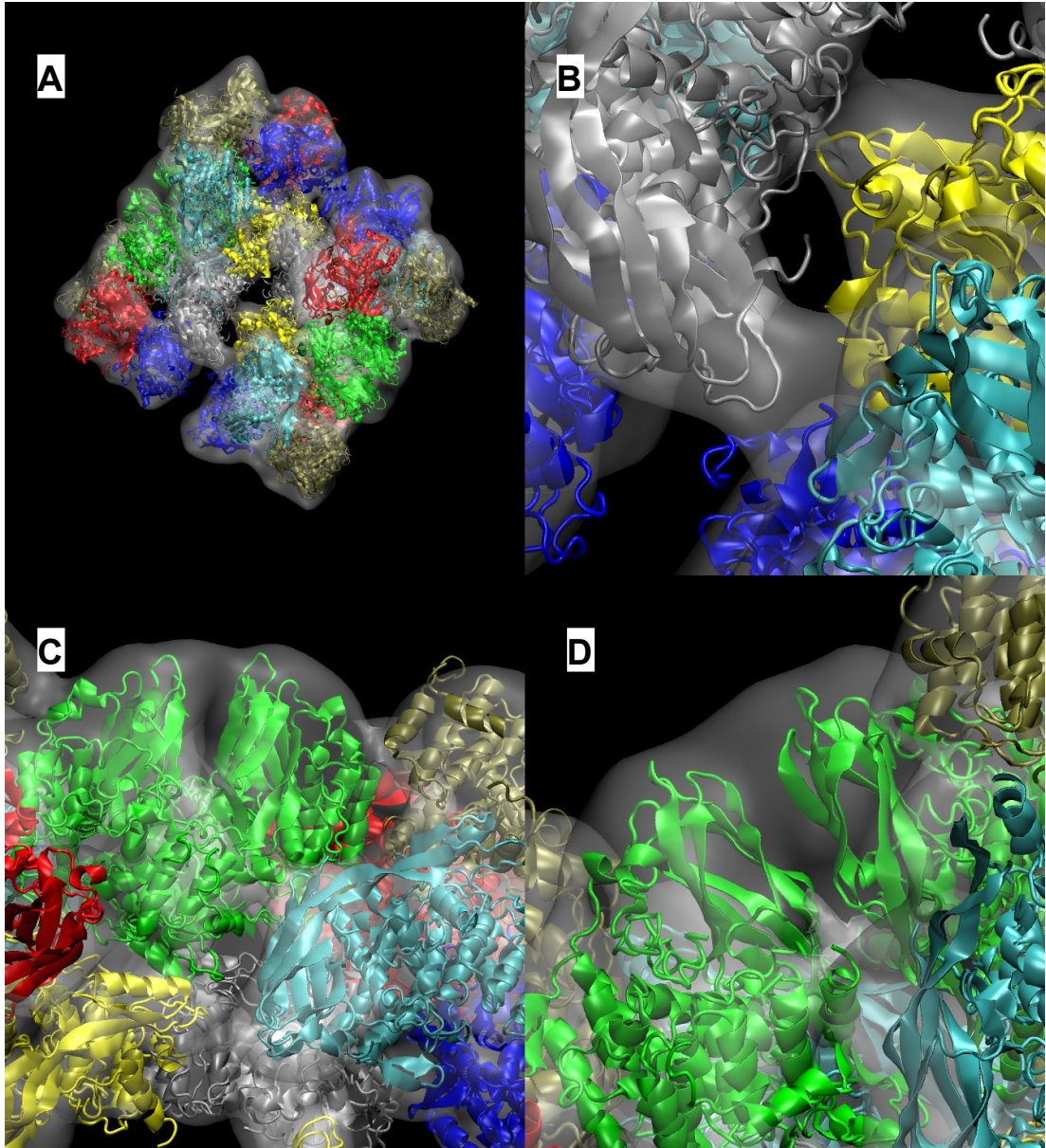


Figura 19. Interfaces de oligomerização da hemocianina de *A. gomesiana*. (A) Perspectiva do modelo com indicações das interfaces analisadas. (B) Interface de associação dos dois 2x6-mero: domínio 3 da sub-unidade *c* (cinza) e domínio 2 da sub-unidade *b* (amarelo); domínio 3 da sub-unidade *b* (amarelo) e domínio 2 da sub-unidade *c* (cinza). (C) interface de oligomerização dos hexâmeros pertencentes ao mesmo 2x6-mero: domínios 3 das sub-unidades *a* (verde) de hexâmeros diferentes; domínio 1 da sub-unidade *a* (verde) e domínio 3 da sub-unidade *d* (azul-claro); domínio 1 das sub-unidades *b* (amarelo) sobrepondo o domínio 1 da sub-unidade *c* (cinza). (D) Detalhe da interface entre os domínios 3 das sub-unidades *a* (verde).

unidade *c* e a outra entre o domínio 3 da sub-unidade *c* e o domínio 2 da sub-unidade *b*. Em (C) podemos observar, os contatos formados entre dois hexâmeros do mesmo dodecâmero. Esta associação é feita através das regiões de contato: o domínio 1 das sub-unidades *b* (amarelo) e *c* (cinza); o domínio 3 das sub-unidades *a* (verde); o domínio 1 da sub-unidade *a* (verde) e o domínio 3 da sub-unidade *d* (azul-claro). Exceto por essa última, as demais interfaces já foram descritas anteriormente por de Haas e van Bruggen (1994) para a tarântula *E. californicum*.

Embora esse padrão de associação derive da escolha do posicionamento das sub-unidades conforme estudos anteriores, o correto posicionamento entre os domínios das sub-unidades e o ajuste entre o modelo proposto e a estrutura cristalográfica, mostram consistência entre as informações provenientes deste trabalho e as disponíveis na literatura.

As próximas etapas propostas para este trabalho são: melhoria da resolução do modelo através do aumento do número de imagens e utilização de imagens obtidas em diferentes defocos; modelagem por homologia das estruturas das sub-unidades de *A. gomesiana* e ajuste das estruturas obtidas no modelo proposto.

4.2 Oligomerização de hRXR α Δ AB e especificidade no reconhecimento de elementos responsivos

4.2.1 Expressão e purificação de hRXR α Δ AB

A proteína hRXR α Δ AB foi expressada em *E. coli*, cepa BL21(DE3), utilizando-se o plasmídio pET-28a(+). A Figura 21 mostra o SDS-PAGE 15%, com as seguintes colunas: PM, com os marcadores de massa molecular, sendo 67,00 kDa, 45,00 kDa, 30,00 kDa, 22,46 kDa e 13,37 kDa; Não Induzida, com BL21 (DE3)/pET-28a(+):hRXR α Δ AB antes da indução com IPTG; Induzida Lisado, BL21 (DE3)/pET-28a(+):hRXR α Δ AB induzida por 3 horas com 0,5 mM de IPTG. A seta preta indica uma banda de aproximadamente 40 kDa, correspondente a hRXR α Δ AB recombinante com cauda de poli-histidina. A indicação PM será utilizada nos demais SDS-PAGE apresentados neste capítulo, indicando as mesmas

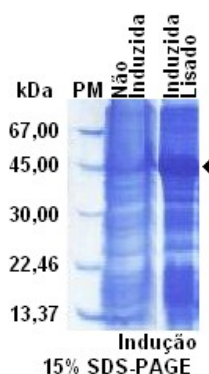


Figura 20. Caracterização da expressão de hRXR α Δ AB por SDS-PAGE 15%. Coluna PM - Marcadores de pesos moleculares (67,00; 45,00; 30,00; 22,46; 13,37 kDa). Coluna Não Induzida - BL21 (DE3)/pET-28a(+):hRXR α Δ AB antes da indução. Coluna Induzida Lisado - BL21 (DE3)/pET-28a(+):hRXR α Δ AB induzida por 3 horas com 0,5 mM IPTG. A seta preta mostra uma banda correspondente a hRXR α Δ AB recombinante com cauda de poli-histidina.

massas moleculares.

Após a expressão, uma primeira etapa de purificação é feita através de cromatografia por afinidade à cauda de poli-histidina, utilizando resina de cobalto (Co^{2+}) Talon Superflow. A Figura 22 mostra o perfil cromatográfico típico (B), juntamente com um SDS-PAGE 15% do resultado obtido (A). No SDS-PAGE, PM é o marcado de massa molecular, como descrito anteriormente e Af apresenta o resultado obtido após eluição com 300 mM de imidazol. Em Af observa-se uma banda intensa referente à proteína hRXR α Δ AB.

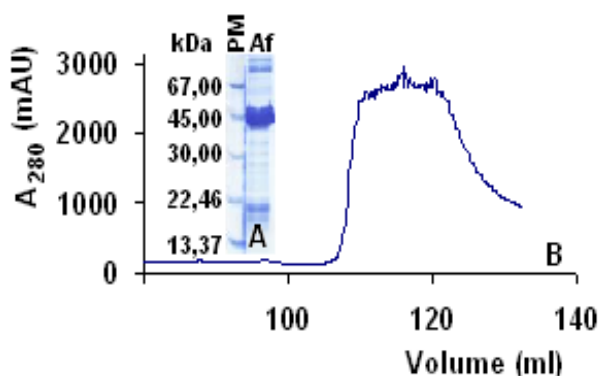


Figura 21. Caracterização da purificação de hRXR α Δ AB em cromatografia de afinidade ao Co^{2+} . (A) SDS-PAGE 15%: coluna PM - Marcadores de massa molecular (67,00; 45,00; 30,00; 22,46; 13,37 kDa). Coluna Af – Após purificação a principal espécie presente é hRXR α Δ AB. (B) Perfil de eluição cromatográfico típico da hRXR α Δ AB após eluição com 300 mM de imidazol. O eixo vertical indica a absorbância no ultravioleta a 280nm e o eixo horizontal o volume de eluição em ml.

O segundo passo de purificação é feito utilizando-se cromatografia por exclusão de tamanho, onde o produto do passo anterior é aplicado em uma coluna HiLoad Superdex 200 26/60. Os resultados dessa etapa são mostrados na Figura 22. Na cromatografia (B), são observados dois picos de absorção, com volumes de eluição de 176 ml e 199 ml,

respectivamente para os denominados Pico 1 (P1) e o Pico 2 (P2). A pureza dos picos obtidos é mostrada em SDS-PAGE 15% (A), onde: PM indica os marcadores de massa molecular, conforme descrito anteriormente; Af indica o resultado da etapa anterior de purificação, utilizado aqui apenas para comparação do grau de pureza da proteína; P1 indica o resultado da purificação por cromatografia por exclusão de tamanho para a fração central do Pico 1 e P2 indica o mesmo resultado para a fração central do Pico 2. O cálculo dos raios hidrodinâmicos (Rh) obtidos a partir destes volumes de eluição foram 50 Å e 40 Å, atribuídos às formas tetramérica e dimérica de hRXR α Δ AB, respectivamente. Estes valores são condizentes com valores medidos anteriormente por Fischer et al. (2003) onde, para o tetrâmero, o raio de giro, medido por SAXS, é 53 Å e o raio hidrodinâmico, medido por cromatografia de exclusão por tamanho, é 52 Å a 57 Å, e para o dímero, o raio de giro, medido por SAXS, é 34 Å e o raio hidrodinâmico, medido por cromatografia de exclusão por tamanho, é 45 Å.

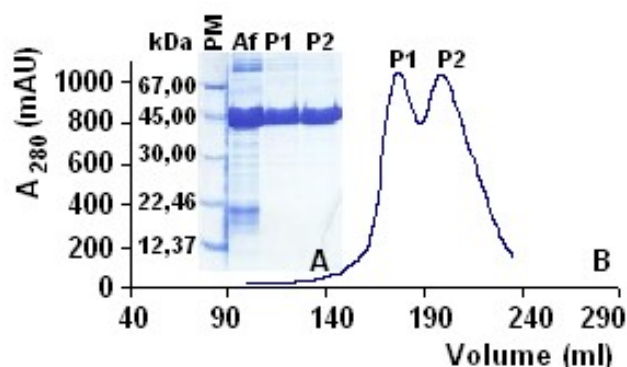


Figura 22. Etapa de purificação de hRXR α Δ AB em cromatografia de exclusão por tamanho. (A) SDS-PAGE 15%. Coluna PM – Marcadores de massa molecular (67,00; 45,00; 30,00; 22,46; 13,37 kDa). Coluna Af – material proveniente da coluna de afinidade e aplicado na coluna de gel filtração. Colunas P1 e P2 – frações centrais dos picos 1 e 2, respectivamente, mostrando uma única banda, ambas contendo a proteína recombinante hRXR α Δ AB purificada. (B) Perfil cromatográfico de eluição isocrática típico, obtido após passagem pela Superdex 200 HL 26/60. O eixo vertical indica a absorvância no ultravioleta a 280nm e o eixo horizontal o volume de eluição em ml. A análise dos volumes de eluição mostram o Pico 1 constituído por tetrâmeros e o Pico 2 por dímeros de hRXR α Δ AB.

É importante ressaltar que protocolo de purificação apresentado neste trabalho difere do utilizado em estudos anteriores que investigaram mRXR α Δ AB (CHEN et al., 1994; KERSTEN et al., 1995a; KERSTEN et al., 1995b; KERSTEN et al., 1996; KERSTEN et al., 1997a; KERSTEN et al., 1997b; KERSTEN et al., 1998). Chen et al. (1994), e trabalhos subsequentes de Kersten, utilizam uma única etapa de purificação por afinidade, sem a etapa final de cromatografia de exclusão por tamanho. Além disso, a cauda de poli-histidina não é clivada em nenhum dos estudos realizados por estes autores. Em 1998, Chen et al. introduzem uma etapa de purificação por gel filtração, após a afinidade, que separa a população em duas espécies, caracterizadas como monômeros e tetrâmeros.

Neste trabalho, após a etapa de purificação por afinidade,

também segue-se uma etapa de cromatografia de exclusão por tamanho, onde são isolados os diferentes estados oligoméricos da proteína, para serem estudados separadamente. Além disso, a cauda de poli-histidina, pode ser eliminada através de protocolo de clivagem por trombina. Os estudos de oligomerização de hRXR α Δ AB foram realizados com proteína com cauda de poli-histidina. Já os estudos de especificidade de reconhecimento a elementos responsivos foi realizado com proteína sem a presença de cauda de poli-histidina.

4.2.2 Oligomerização de hRXR α Δ AB

Os estudos, apresentados a seguir, foram feitos utilizando-se proteína com cauda de poli-histidina, assim como em estudos anteriores sobre mRXR α Δ AB (CHEN et al., 1994; KERSTEN et al., 1995a; KERSTEN et al., 1995b; KERSTEN et al., 1996; KERSTEN et al., 1997a; KERSTEN et al., 1997b; KERSTEN et al., 1998, CHEN et al., 1998). Entretanto, ao contrário desses estudos, as populações iniciais de dímeros e tetrâmeros foram analisadas separadamente assim como em Chen et al. (1998).

Para investigar-se os diferentes estados oligoméricos apresentados pela hRXR α Δ AB, após a purificação por exclusão de tamanho foi realizado uma eletroforese em gel nativo com gradiente de poliacrilamida de 8% a 25%, conforme mostrado na Figura 23. Nestes experimentos, devido às características da técnica, na qual a forma e a carga da proteína influenciam na mobilidade, os marcadores servem apenas como referência

para posicionamento no gel, não podendo ser usados para aferição de massa molecular. Adotou-se como parâmetro de comparação de massa, as posições obtidas para monômeros, dímeros e tetrâmeros de hTR β Δ AB, que foi anteriormente estudado em nosso grupo (DIAS, 2004). Foram utilizadas quatro frações da eluição, numeradas de 1 a 4 na figura, sendo elas: as frações centrais de ambos os picos, onde a absorbância a 280 nm é máxima (1 e 3); a fração intermediária aos picos (2) e a porção final de P2 (4). Observou-se que, na fração central de P1 (1), a forma oligomérica predominante é tetramérica, embora uma banda de menor intensidade seja observada na região de dímeros. Para a fração correspondente à região intermediária entre P1 e P2 (2) observou-se uma banda de tetrâmeros e outra, de mesma intensidade, de dímeros. Na fração central de P2 (3) a forma predominante é de dímeros embora uma banda de menor intensidade, de tetrâmeros, também esteja presente. Para a última fração, proveniente da parte final de P2 (4), observa-se um padrão semelhante ao apresentado pela fração anterior, embora com bandas de menor intensidade, conforme esperado. Estas observações confirmam os resultados obtidos a partir da análise do volume de eluição na cromatografia por exclusão de tamanho, onde P1 é composto principalmente por tetrâmeros e P2 por dímeros, quando comparados com a proteína padrão hTR β Δ AB.

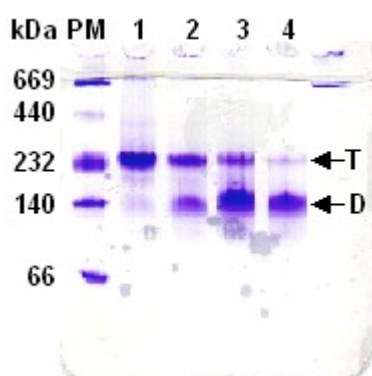


Figura 23. Estados oligoméricos da hRXR α Δ AB após a cromatografia de exclusão por tamanho. Gel de gradiente nativo 8-25%: Coluna PM – Marcadores de pesos moleculares (669; 440; 232; 140; 66 kDa). Colunas 2, 3, 4, 5 – Contém as frações: central do Pico 1; intermediária entre o Pico 1 e Pico 2; central do Pico 2; final do Pico 2. O experimento foi realizado logo após a purificação em coluna de gel filtração. As setas pretas apontam para as posições de tetrâmero (T) e dímero (D) de hRXR α Δ AB.

Utilizando-se hRXR α Δ AB proveniente do Pico 2, uma fração predominantemente dimérica, sem a presença de tetrâmeros, foi feito um outro gel de gradiente nativo 8-25%, com as amostras tendo sido incubadas em uréia (Figura 24). Foram utilizados os mesmos marcadores de massa molecular, conforme indicado na coluna PM. As amostras (40 μ M) foram equilibradas por 2 horas, a 18°C, na presença de 0 M, 1 M, 2 M, 3 M e 4 M de uréia. As colunas 0, 1, 2, 3 e 4 indicam, respectivamente, a concentração molar de uréia. Em 4 M de uréia, o gel apresenta duas bandas, atribuídas a monômeros (M) e dímeros (D).

Em condições similares de concentração de uréia, a desnaturação da proteína, proveniente do Pico 2, foi analisada por dicroísmo circular (FASMAN, 1996). Observa-se que na concentração de 4 M de uréia, o sinal de CD praticamente não apresenta diferença com relação à proteína

nativa, indicando estabilidade da estrutura secundária, nessas condições (Figura 25).

Sendo assim, observa-se em 4 M de uréia, a existência de uma espécie de maior mobilidade do que as duas espécies purificadas na cromatografia de exclusão por tamanho. Esta tem sua estrutura secundária preservada, o que pode ser um indicativo de que a espécie observada na altura do marcador de 140 kDa é realmente dimérica e que, ao dissociar na presença de uréia, origina uma banda de monômeros.

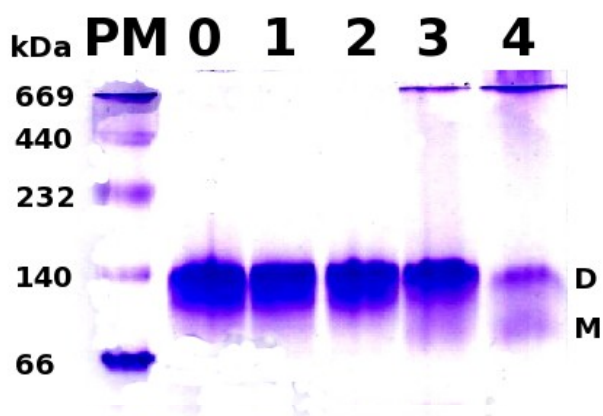


Figura 24. Desnaturalização de hRXR α Δ AB do Pico 2 por uréia acompanhada por PAGE. Gel de gradiente nativo 8-25% de hRXR α Δ AB incubada em uréia. As colunas 0, 1, 2, 3 e 4 indicam, respectivamente, a concentração molar de uréia, 0 M, 1 M, 2 M, 3 M e 4 M.

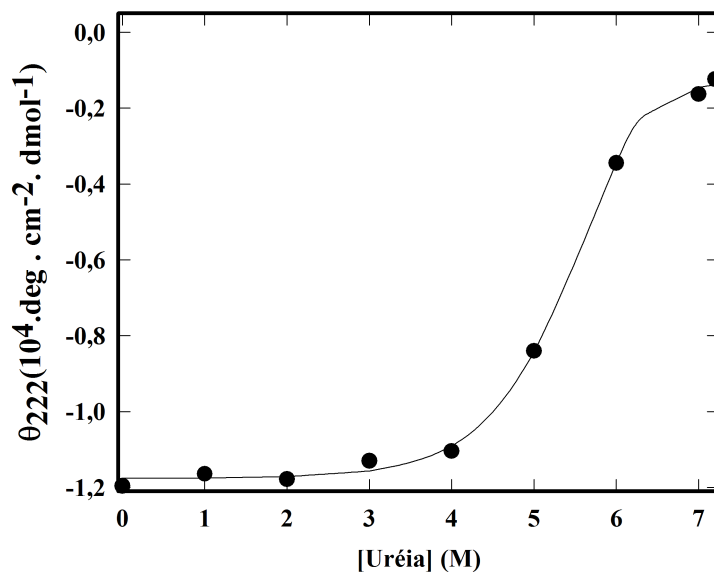


Figura 25. Desnaturalização de hRXR α Δ AB do Pico 2 por uréia acompanhada por dicroísmo circular.

Utilizando-se ainda o Pico 2 da purificação, realizou-se uma cromatografia analítica de exclusão por tamanho em uma coluna TSK-3000 (Figura 26). Foram utilizados para calibração proteínas com as seguintes massas moleculares: 13,5 kDa; 16,0 kDa; 18,0 kDa; 18,0 kDa; 33,0 kDa; 33,0 kDa; 55,0 kDa, 67,0 kDa, que são indicadas na figura por círculos pretos. O valor de eluição da hRXR α Δ AB (7 μ M) proveniente do Pico 2 (dimérico) foi de 7,2 ml, o que corresponde a uma massa molecular de 61 kDa, indicado na figura por um círculo aberto.

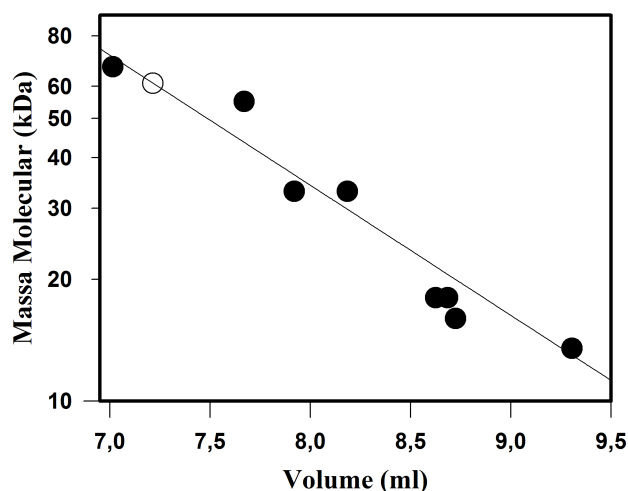


Figura 26. Cromatografia analítica de exclusão por tamanho do Pico 2 da hRXR α Δ AB. Foi utilizada uma coluna TSK-3000, calibrada utilizando-se proteínas com massas moleculares de 13,5 kDa; 16,0 kDa; 18,0 kDa; 18,0 kDa; 33,0 kDa; 33,0 kDa, 55,0 kDa e 67,0 kDa. O eixo vertical indica a massa molecular das proteínas e o eixo horizontal o volume de eluição em ml. O volume de eluição e massa molecular das proteínas utilizadas como padrão são indicadas com um círculo preto. O círculo aberto mostra o volume de eluição correspondente a hRXR α Δ AB, que indica uma massa molecular de 61 kDa.

Os dados apresentados pela cromatografia analítica corroboram os dados de Fischer et al. (2003), no sentido de que há uma predominância de dímeros em proteína proveniente do Pico 2, nessa concentração. Entretanto, o protocolo de purificação utilizado por Fischer et al. (2003) difere do apresentado neste trabalho. No sentido de se verificar a predominância de dímeros nas frações de purificação oriundas do Pico 2, novas medidas de SAXS foram realizadas. Para estas, a cauda de polihistidina da proteína foi clivada. A Figura 27 mostra o gráfico de Guinier para hRXR α Δ AB (7,2 mg/ml) e BSA (1 mg/ml). Utilizando-se a equação 18, calculou-se o valor do raio de giro do hRXR α Δ AB como 39 Å, compatível com o valor de 40 Å, observado para o raio hidrodinâmico na cromatografia de exclusão por tamanho. A massa foi calculada comparando-se as

intensidades iniciais de ambas as curvas, obtendo-se o valor de 76 kDa de massa para o hRXR α Δ AB, também compatível com a massa do dímero.

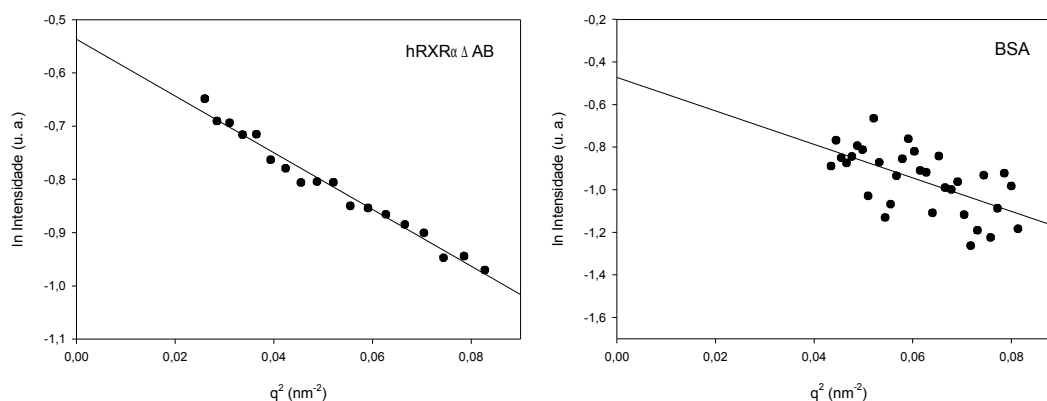


Figura 27. Gráfico de Guinier para hRXR α Δ AB e BSA.

No sentido de se fazer uma análise comparativa entre os picos 1 e 2, as frações centrais de ambos foram escolhidas para serem investigadas quanto à sua resposta ao ligante, 9-cis-RA, e ao elemento responsivo DR1. Com isso, investigou-se a influência destes nos estados oligoméricos observados. Foram preparadas amostras conforme descrito na Tabela 3. São indicadas as concentrações finais obtidas.

Tabela 3: Amostras de hRXR α Δ AB incubadas com elemento responsivo e ligante.

Amostra	Pico	Ligante	DR1
(50 μ M)		(2 mM)	(400 pmol/ μ l)
RXR 1	1	-	-
RXR 1 – 9cis	1	9-cis-RA	-
RXR 1 – curto	1	-	Curto
RXR 1 – longo	1	-	Longo

Amostra	Pico	Ligante	DR1
(50 μM)		(2 mM)	(400 pmol/μl)
RXR 1 – 9cis – curto	1	9-cis-RA	Curto
RXR 1 – 9cis – longo	1	9-cis-RA	Longo
RXR 2	2	-	-
RXR 2 – 9cis	2	9-cis-RA	-
RXR 2 – curto	2	-	Curto
RXR 2 – longo	2	-	Longo
RXR 2 – 9cis – curto	2	9-cis-RA	Curto
RXR 2 – 9cis – longo	2	9-cis-RA	Longo

Estas amostras foram descongeladas e imediatamente incubadas, na presença do elemento responsivo e ligante, por 1h e 12h a 4°C. Em seguida foram analisadas em gel de gradiente nativo 8-25%. Em outro conjunto de medidas, as amostras foram descongeladas e, antes de serem incubadas, permaneceram em equilíbrio a 4°C, durante 2 horas. As amostras incubadas com DR1 com extremidades longas foram utilizadas apenas no segundo conjunto de medidas, após verificar-se que a ligação a este era mais efetiva do que ao DR1 com extremidades curtas. As figuras a seguir mostram os resultados obtidos.

A Figura 28 mostra os resultados para o Pico 1 da hRXR α Δ AB. Foram feitos dois géis de gradiente nativo 8-25%, da fração central do pico, com tempos de incubação de 1 hora e 12 horas, a 4°C. Em (A) e (A') os géis foram corados com *coomassie blue* e em (B) e (B') os mesmos géis foram corados com brometo de etídio para visualização do DNA. A coluna PM, em todos os géis, mostra o marcador de massa molecular, conforme descrito

anteriormente. As concentrações utilizadas foram: proteína (50 μ M); 9-cis-RA (2 mM); DR1 (400 pmol/ μ l).

Observa-se que a incubação das amostras por um período de 12 horas parece não alterar a intensidade relativa das bandas, indicando que a incubação por 1 hora, a 4°C, é suficiente para observar-se a ligação ao 9-cis-RA e ao DR1. Na ausência de ligante e do elemento responsivo, a proteína, proveniente do Pico 1, mostra-se predominantemente tetramérica, ainda que apresentando um menor população de dímeros (A1 e A'1). A adição do ligante, 9-cis-RA, faz com que parte da população de tetrâmeros se dissocie em dímeros. Entretanto, a forma tetramérica continua predominante (A2 e A'2). A associação ao DR1 ocorre sem a necessidade da presença do ligante, conforme esperado (CHEN et al., 1994) (A3, A4, A'3, A'4, B3, B4, B'3 e B'4). A presença do elemento responsivo causa um maior deslocamento do dímero, que pode estar associado à alteração da carga do complexo hRXR α Δ AB-DR1 e também a um enovelamento mais compacto do dímero. Apesar de haver, claramente uma população de dímeros ligados ao DR1, há também uma população que apresenta deslocamento semelhante ao de dímeros e tetrâmeros não ligados, ainda que indicando a presença do elemento responsivo (B3 e B4). Parte da população de tetrâmeros, na presença de DR1, não se dissocia, apresentando apenas uma maior difusão na banda (A3T e A4T). Sendo assim, concluímos que a espécie de tetrâmero predominante neste pico, é bastante estável, apresentando pouca dissociação em função da presença do ligante e

embora, dissocie-se na presença do DR1, parte significativa da população permanece tetramérica.

A Figura 29 mostra os resultados obtidos para o Pico 2, nas mesmas condições utilizadas para o estudo do Pico 1. Assim como para o Pico 1, as medidas com 1 hora e 12 horas de incubação parecem indicar os mesmos resultados. Para o Pico 2, a população predominante é dimérica, na ausência de ligante e elemento responsivo, apresentando uma menor população de tetrâmeros (A1 e A'1). Diferentemente dos tetrâmeros do Pico 1, a maior parte da população de tetrâmeros deste pico responde à presença de 9-cis-RA, se dissociando em dímeros (A2 e A'2). Na presença de DR1, observa-se que parte da população de hRXR α Δ AB se liga como dímero, não dependendo da presença do ligante para isso (A3, A4, B3, B4, A'3, A'4, B'3 e B'4). Assim como para o Pico 1, na presença de DNA, a migração é mais intensa, apresentando bandas em posições inferiores. Entretanto, não aparecem bandas de tetrâmero, após ligação ao DR1, e sim bandas na posição de dímeros não ligados (A3, A4, A'3 e A'4). Parte destes apresenta, ainda que fracamente, bandas quando corado com brometo de etídio, indicando a presença de DNA (B3, B4, B'3 e B'4). Não é possível determinar qual o estado oligomérico desse complexo, podendo ser um dímero com migração retardada ou um tetrâmero com uma migração intensificada. A ligação de tetrâmeros de RXR (isoformas α , β , γ) a elementos responsivos foi anteriormente relatada (CHEN; PRIVALSKY, 1995; LIN et al., 1997).

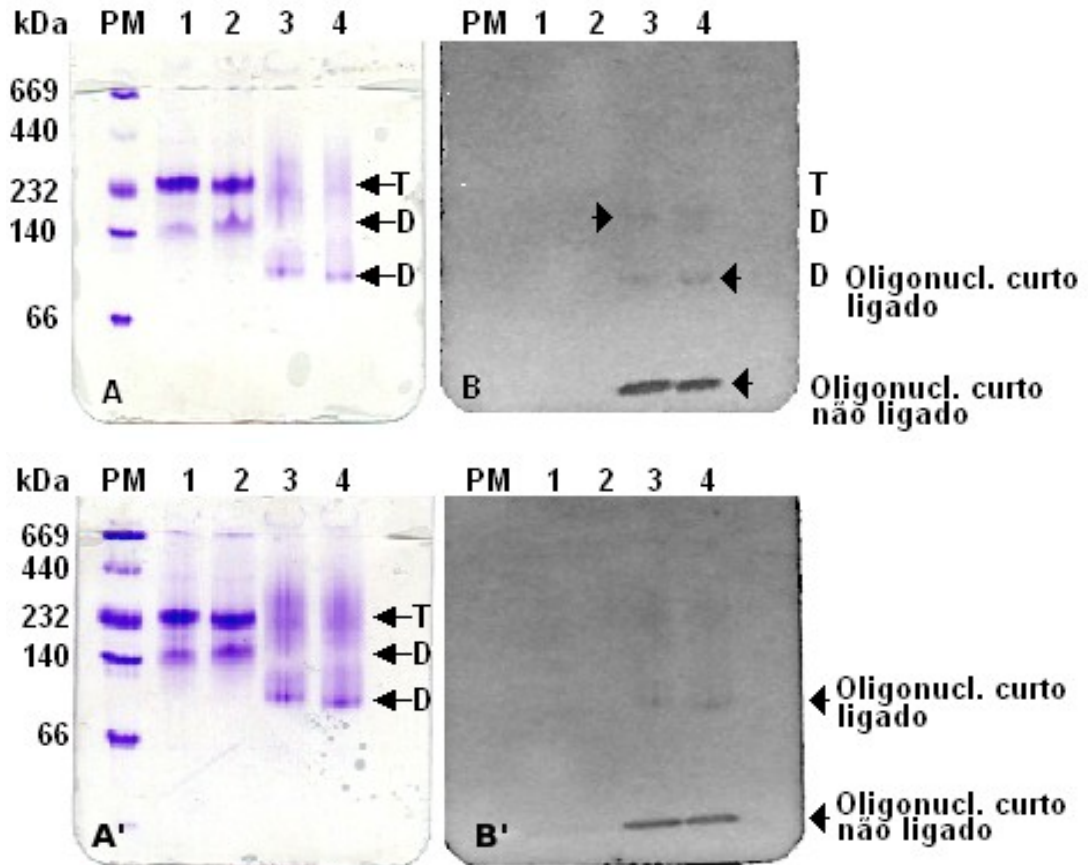


Figura 28. Ligação de hRXR α Δ AB (Pico 1) ao elemento responsivo DR1 e ao ligante 9-cis-RA. Gel de gradiente nativo 8-25% da fração central do Pico 1 de hRXR α Δ AB após incubação por 1 hora e 12 horas a 4°C – (A) Gel corado com *coomassie blue* - Coluna PM – Marcadores de massa molecular (669; 440; 232; 140; 66 kDa). Coluna 1 – proteína (50 μ M) equilibrada por 1 hora a 4° C. Coluna 2 – proteína (50 μ M) incubada por 1 h com 9-cis-RA (2 mM). Coluna 3 – proteína (50 μ M) incubada por 1 h com HRE DR1 de extremidades curtas (400 pmol/ μ l). Coluna 4 – proteína (50 μ M) incubada por 1 h com ambos, RA 9-cis (2 mM) e HRE DR1 de extremidades curtas (400 pmol/ μ l). (B) O mesmo gel anterior, corado com brometo de etídio para visualização do DNA. Os oligonucleotídeos ligados à proteína podem ser vistos em bandas cinzas fracas. As setas pretas apontam para as posições de tetrâmero, indicado por T, e dímero, indicado por D, de hRXR α Δ AB na ausência ou presença de HRE DR1 de extremidades curtas. (A') e (B') referem-se às mesmas amostras utilizadas em (A) e (B), porém com incubação por 12 horas.

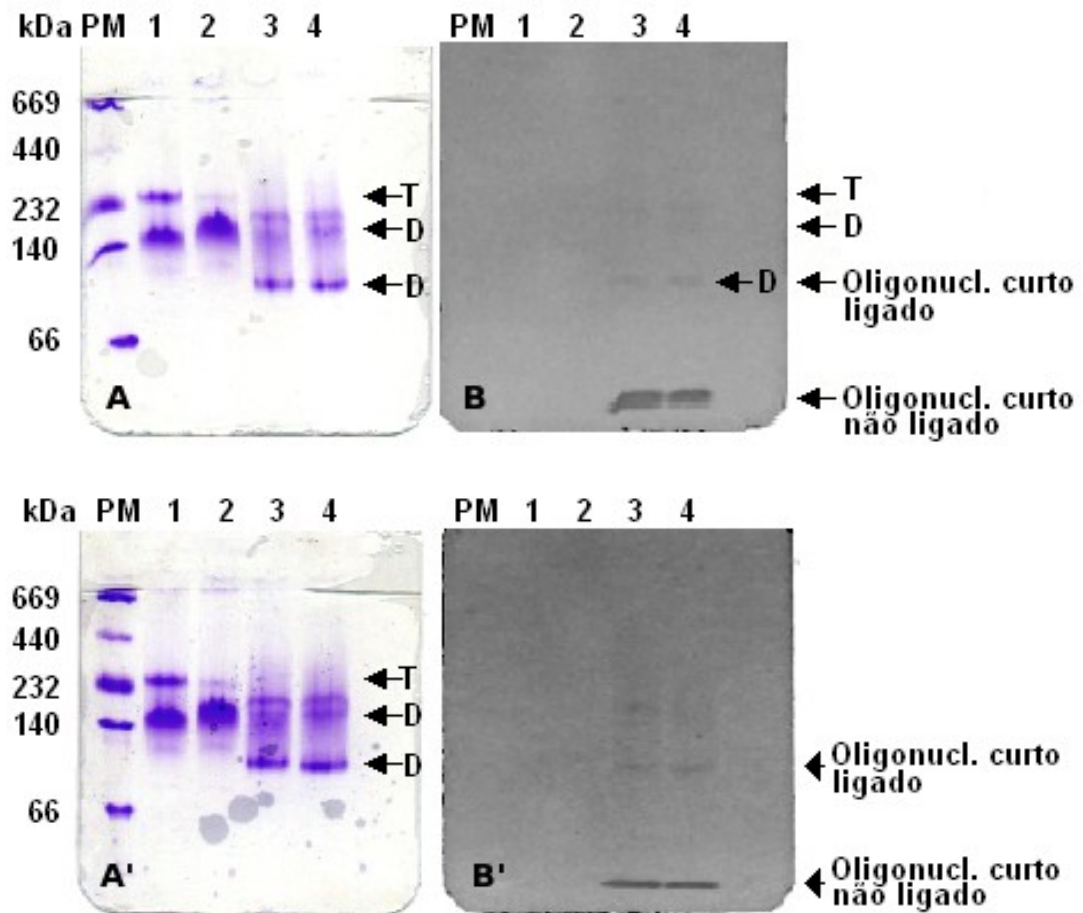


Figura 29. Ligação de hRXR α Δ AB (Pico 2) ao elemento responsivo DR1 e ao ligante 9-cis-RA. Gel de gradiente nativo 8-25% da fração central do Pico 2 de hRXR α Δ AB após incubação por 1 hora e 12 horas a 4°C – (A) Gel corado com *coomassie blue* - Coluna PM – Marcadores de massa molecular (669; 440; 232; 140; 66 kDa). Coluna 1 – proteína (50 μ M) equilibrada por 1 hora a 4° C. Coluna 2 – proteína (50 μ M) incubada por 1 h com 9-cis-RA (2 mM). Coluna 3 – proteína (50 μ M) incubada por 1 h com HRE DR1 de extremidades curtas (400 pmol/ μ l). Coluna 4 – proteína (50 μ M) incubada por 1 h com ambos, RA 9-cis (2 mM) e HRE DR1 de extremidades curtas (400 pmol/ μ l). (B) O mesmo gel anterior, corado com brometo de etídio para visualização do DNA. Os oligonucleotídeos ligados à proteína podem ser vistos em bandas cinzas fracas. As setas pretas apontam para as posições de tetrâmero, indicado por T, e dímero, indicado por D, de hRXR α Δ AB na ausência ou presença de HRE DR1 de extremidades curtas. (A') e (B') referem-se às mesmas amostras utilizadas em (A) e (B), porém com incubação por 12 horas.

Foram realizados em seguida ensaios onde as amostras permaneceram em equilíbrio a 4°C por 2 horas e foram incubadas nas mesmas condições anteriores, incluindo-se mais dois tipos de amostras, hRXR α Δ AB com DR1 longo e hRXR α Δ AB com DR1 longo e 9-cis-RA (Figura 30). Os resultados obtidos são os mesmos já observados no experimento anterior, exceto pela constatação de que o DR1 longo parece apresentar maior afinidade para com a proteína (colunas 4 e 6 de A, B, A', B') quando comparado com DR1 curto (colunas 3 e 5 de A, B, A', B'). Este resultado foi importante no sentido de se definir pelo uso de elementos responsivos "longos" para as medidas de determinação da especificidade de reconhecimento de elementos responsivos, que se seguiram.

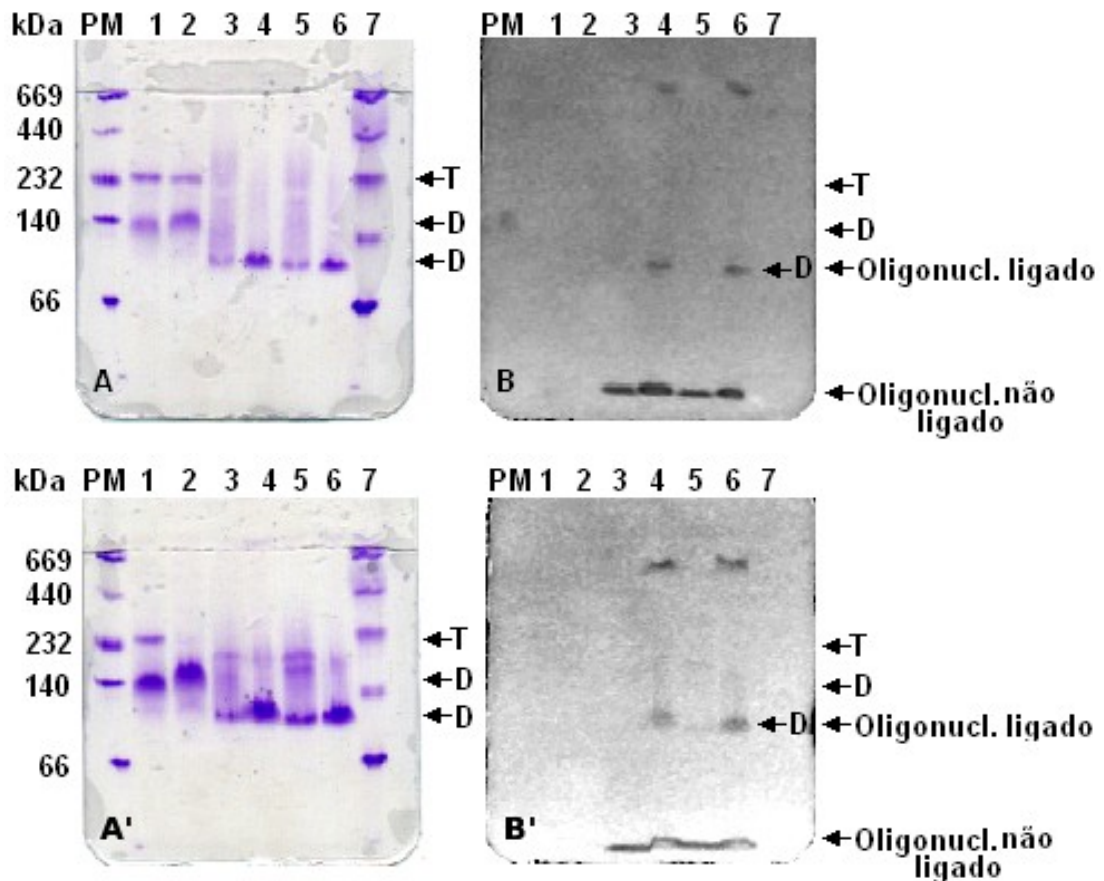


Figura 30. Ligação de hRXR α Δ AB (Pico 1 e 2) ao elemento responsivo DR1 e ao ligante 9-cis-RA, após equilíbrio de 2 horas. Gel de gradiente nativo 8-25% da fração central do Pico 1 de hRXR α Δ AB após equilíbrio por 2 horas, a 4°C, e incubação por 1 hora a 4°C – (A) Gel corado com *coomassie blue* - Coluna PM – Marcadores de massa molecular (669; 440; 232; 140; 66 kDa). Coluna 1 – proteína (50 μ M) equilibrada por 3 horas a 4° C. Coluna 2 – proteína (50 μ M) incubada por 1 h com 9-cis-RA (2 mM). Colunas 3 e 4 – proteína (50 μ M) incubada por 1 h com HRE DR1 de extremidades curtas e longas (400 pmol/ μ l), respectivamente. Colunas 5 e 6 – proteína (50 μ M) incubada por 1 h com ambos, 9-cis-RA (2 mM) e HRE DR1 de extremidades curtas e longas (400 pmol/ μ l), respectivamente. (B) O mesmo gel anterior, corado com brometo de etídio para visualização do DNA. Os oligonucleotídeos ligados à proteína podem ser vistos em bandas cinzas fracas. A ligação ao DR1 de extremidades longas é notadamente mais forte do que ao de extremidades curtas. As setas pretas apontam para as posições de tetrâmero, indicado por T, e dímero, indicado por D, de hRXR α Δ AB na ausência ou presença de HRE DR1. (A') e (B') mostram os resultados obtidos, segundo o mesmo protocolo, para o Pico 2 de hRXR α Δ AB. As mesmas indicações acima foram utilizadas.

4.2.3 Especificidade no reconhecimento de elementos responsivos

Nesta etapa do trabalho foi analisada a interação da proteína hRXR α Δ AB com diferentes elementos responsivos na presença e ausência do seu ligante, 9-cis-RA. Utilizando a técnica de anisotropia de fluorescência, foram analisadas as constantes de dissociação do sistema proteína-DNA. Quatro diferentes elementos responsivos foram utilizados: DR1, repetição direta com um espaçamento; DR4, repetição direta com quatro espaçamentos; F2, palíndromo invertido com seis espaçamentos; PAL, palíndromo sem espaçamento. Os resultados estão apresentados na forma de variação da anisotropia (Δ Anisotropia), considerando-se o valor observado para o oligonucleotídeo livre, como valor inicial.

A Figura 31 mostra os resultados obtidos para a ligação proteína-DNA na ausência do ligante 9-cis-RA. Foram realizados três experimentos independentes em tampão contendo 50 mM de NaCl e um experimento em tampão com 200 mM de NaCl. Observou-se não haver interação significativa da hRXR α Δ AB com nenhum dos elementos responsivos na presença de 200 mM de NaCl, para concentrações de proteína de até 200 nM. Para as mesmas condições de concentração de proteína, também não foi observada interação para o elemento responsivo PAL, em nenhuma das concentrações de NaCl utilizadas. Para os elementos responsivos DR1, DR4 e F2 foi observada interação proteína-DNA para a concentração de 50 mM de NaCl, não sendo observada interação na condição de 200 mM de NaCl. Os valores

de K_d obtidos foram 13 ± 4 nM, 26 ± 10 nM e 39 ± 21 nM, para DR1, DR4 e F2, respectivamente.

A Figura 32 mostra os resultados obtidos para a interação de hRXR α Δ AB com os elementos responsivos DR1, DR4 e F2, em tampão com 50 mM de NaCl, porém adicionando-se 300 nM de 9-cis-RA. Foram realizados três experimentos independentes para cada elemento responsivo. Em todos os experimentos foi observada interação entre a proteína e os oligonucleotídeos. Os valores de K_d obtidos foram 8 ± 3 nM, 18 ± 6 nM e 92 ± 40 nM, para DR1, DR4 e F2, respectivamente.

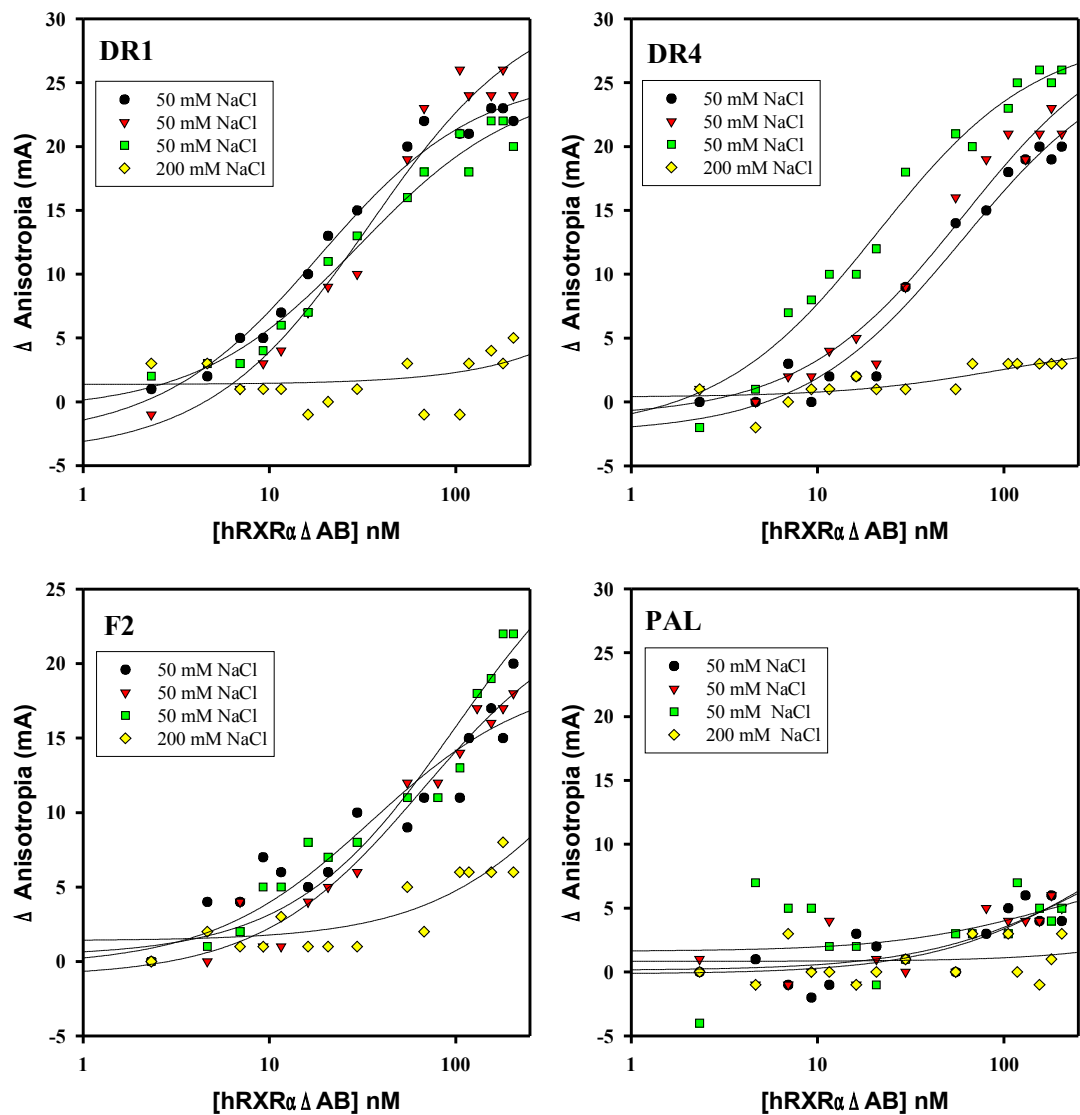


Figura 31. Associação de hRXR α Δ AB a diferentes oligonucleotídeos, sem a presença do ligante 9-cis-RA. Foram realizados três experimentos independentes em tampão contendo 50 mM de NaCl e um experimento em tampão com 200 mM de NaCl, para cada um dos elementos responsivos: DR1, DR4, F2 e PAL.

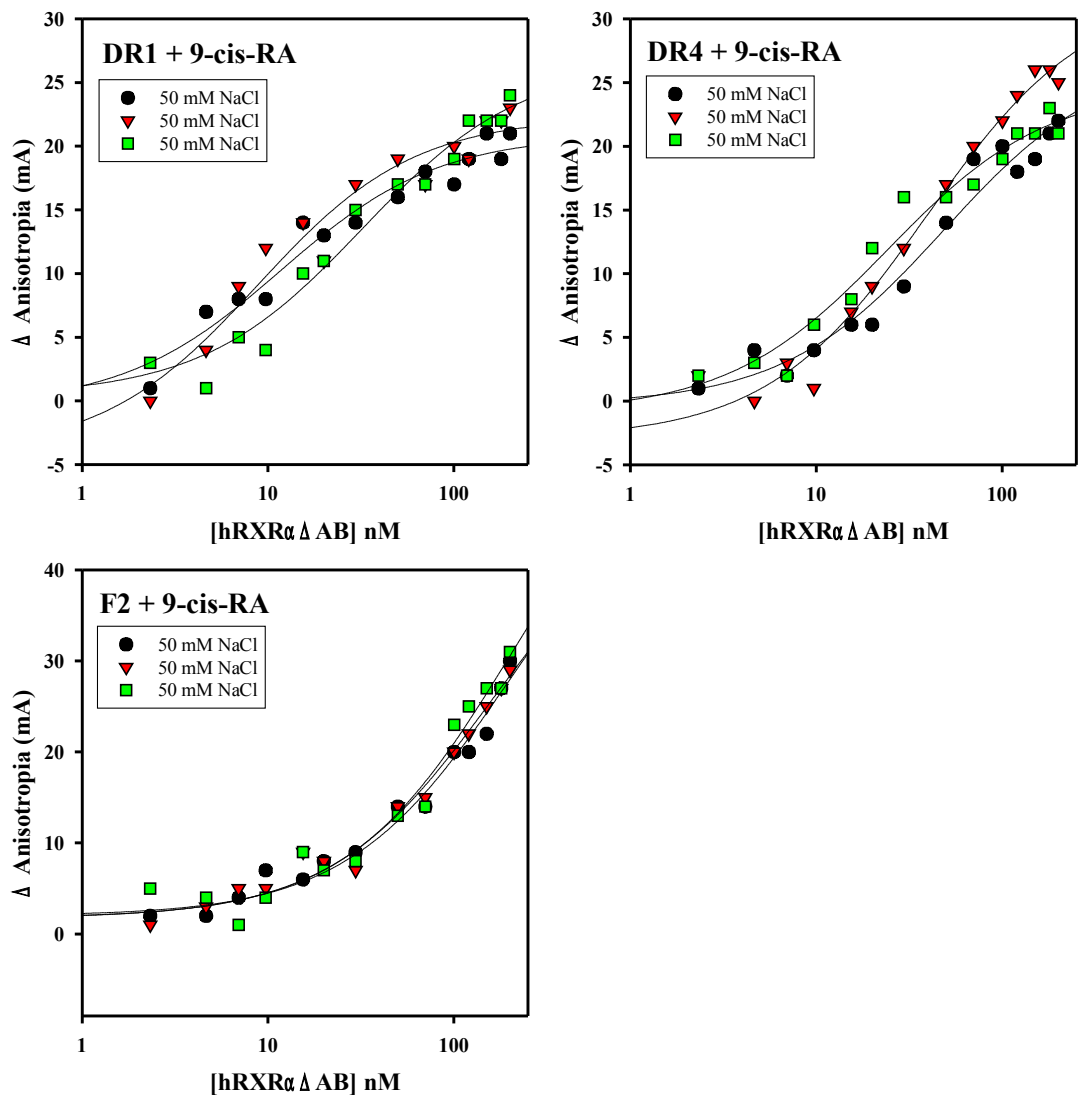


Figura 32. Associação de hRXR α Δ AB a diferentes oligonucleotídeos, na presença do ligante 9-cis-RA. Foram realizados três experimentos independentes em tampão contendo 50 mM de NaCl e um experimento em tampão com 200 mM de NaCl, para cada um dos elementos responsivos: DR1, DR4, F2 e PAL.

A Tabela 4 mostra os valores de constantes de dissociação obtidos para a interação proteína-DNA, para os diferentes elementos responsivos utilizados e na ausência e presença de ligante. Os mesmos resultados também são apresentados na Figura 33. Os experimentos foram conduzidos a 10°C. Os resultados mostrados são a média das constantes calculadas para experimentos independentes e o erro indicado é o desvio padrão obtido.

Tabela 4: Constantes de dissociação (K_d) para hRXR α Δ AB e diferentes elementos responsivos (HRE) na presença (+9-cis-RA) e ausência de ligante (-9-cis-RA).

HRE	- 9-cis-RA	+ 9-cis-RA
DR1	13 \pm 4 nM	8 \pm 3 nM
DR4	26 \pm 10 nM	18 \pm 6 nM
F2	39 \pm 21 nM	92 \pm 40 nM

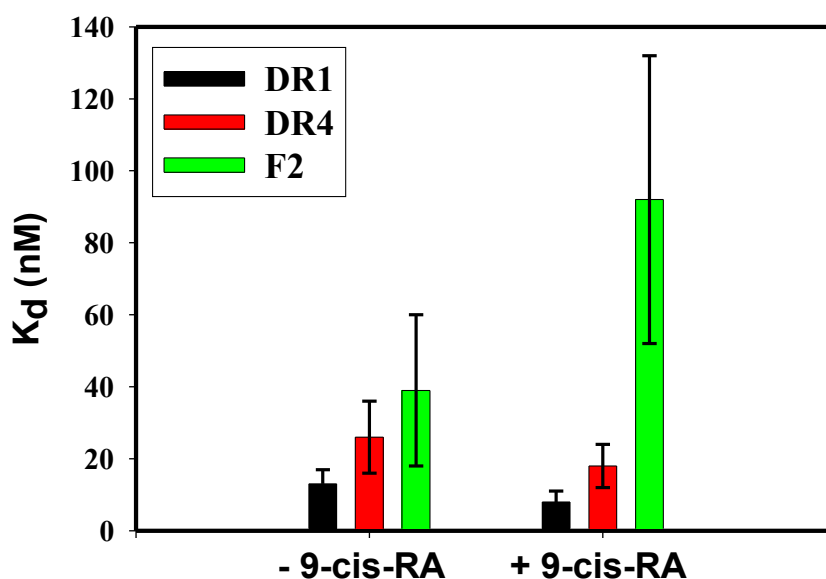


Figura 33. Constantes de dissociação para hRXR α Δ AB e diferentes elementos responsivos (HRE) na presença (+9-cis-RA) e ausência de ligante (-9-cis-RA).

As constantes de dissociação calculadas para DR1 e DR4, no limite do erro experimental, são respectivamente 17 nM e 16 nM na presença do ligante, e 11 nM e 12 nM na sua ausência. Sendo assim, embora estes sejam valores limite, não é possível afirmar que a ligação de hRXR α Δ AB ocorra de maneira diferenciada entre estes elementos, na presença ou ausência do ligante. Por outro lado, comparando-se as constantes obtidas para F2, na presença e ausência de ligante com as obtidas para DR1 e DR4, observa-se que a adição do 9-cis-RA torna a proteína seletiva com relação à DR1 e DR4 em comparação à F2.

A dependência negativa da associação à F2, na presença de 9-cis-RA, pode indicar a existência de um mecanismo de resposta que, na presença do ligante, faz com que o receptor perca afinidade com relação a elementos responsivos não específicos que se ligam à forma apo da proteína. Isso faria com que o maquinário de transcrição, que é requisitado através da ligação de 9-cis-RA ao receptor, só viesse a atuar quando este estivesse associado ao DNA nas regiões de interesse. Do ponto de vista estrutural, espera-se que a organização do dímero para ligação aos elementos responsivos com repetições diretas (DR1, DR4) seja diferente do que para a ligação à F2. Isto porque, F2 é um palíndromo invertido, e portanto, necessita que os DBDs da proteínas se orientem em sentidos opostos, ao contrário dos elementos de repetição direta. Sendo assim, a ligação do 9-cis-RA pode estar impondo uma configuração estrutural à proteína que impossibilita a inversão dos seus DBDs.

Para a caracterização da ligação entre a proteína e os elementos responsivos foram realizados experimentos em diferentes temperaturas. A Figura 34 mostra três experimentos de titulação de hRXR α Δ AB em DR1, realizados segundo o mesmo protocolo anterior, nas temperaturas de 10°C, 15°C e 20°C. Utilizando-se o resultado de K_d já obtido anteriormente para a temperatura de 10°C e calculando-se os valores para 15°C e 20°C à partir das curvas abaixo, é possível montar o gráfico de van't Hoff mostrado na Figura 35. Desse gráfico obtêm-se os valores de $\Delta H^0 = 24 \pm 5$ kcal/mol e $\Delta S^0 = 0,121 \pm 0,017$ kcal/mol.K. Os erros indicados para ΔH^0 e ΔS^0 são derivados da regressão linear utilizada para o ajuste dos dados. Os resultados obtidos para as constantes termodinâmicas da reação de associação proteína-DNA são mostrados na Tabela 5.

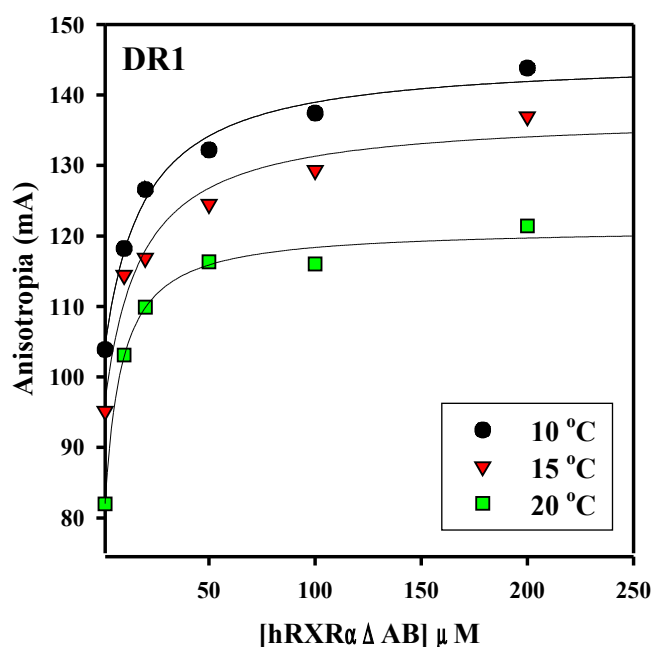


Figura 34. Curvas de ligação hRXR α Δ AB-DR1 em diferentes temperaturas.

Tabela 5: Constantes termodinâmicas da associação hRXR α Δ AB-DR1.

Temperatura	K_d (nM)	ΔH^0 (kcal/mol)	$T\Delta S^0$ (kcal/mol)
10°C	13 \pm 4	24 \pm 5	34 \pm 5
15°C	8 \pm 6	24 \pm 5	35 \pm 5
20°C	3 \pm 2	24 \pm 5	35 \pm 5

Analisando-se os resultados mostrados na Tabela 5, observa-se que a componente entrópica da associação da hRXR α Δ AB ao elemento responsivo DR1, considerando-se a margem de erro, é maior ou igual à componente entálpica observada. Visto que a associação da proteína ao DNA é observada espontaneamente ($\Delta G^0 < 0$) é possível supor que o processo de associação possui uma maior componente entrópica. Além disso, o processo mostra uma variação entálpica no sentido contrário à variação da energia livre de Gibbs ($\Delta H^0 > 0$), como pode ser visto também na Figura 35. Sendo assim, estes dados indicam uma natureza entrópica para a associação de hRXR α Δ AB ao elemento responsivo DR1. Esse mecanismo é também observado para outras proteínas que interagem com DNA (KE et al, 2002; JEN-JACOBSON et al., 2000; FOGUEL; SILVA, 1994).

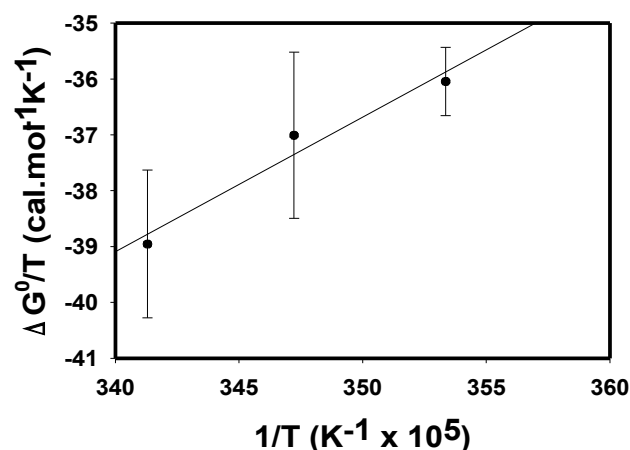


Figura 35. Gráfico de van't Hoff para a hRXR α Δ AB. As temperaturas mostradas são 10°C, 15°C e 20°C.

O mecanismo estrutural de reconhecimento de DR1 é conhecido (ZHAO et al., 2000) e apresenta aminoácidos com característica hidrofóbica na região de ligação. A associação proteína-DNA, quando de natureza entrópica, depende da existência de regiões hidrofóbicas que estejam expostas ao solvente, organizando-o em torno destas. A ligação da proteína ao DNA, reorganiza ambos sistemas, fazendo com que estas regiões se tornem protegidas pelo DNA, ou mesmo pela proteína, liberando o solvente (JEN-JACOBSON, 2000) e ocasionando a variação da entropia. Sendo assim, existe a possibilidade estrutural de que o mecanismo de associação entre o hRXR α Δ AB e o elemento responsivo DR1 tenha origem entrópica.

5 Conclusões

O trabalho desenvolvido teve como objetivo o estudo e caracterização de diferentes complexos formados pelo receptor de ácido retinóico 9-cis e também pela hemocianina de *Acanthoscurria gomesiana*, através de técnicas biofísicas e estruturais.

No estudo da hemocianina de *A. gomesiana*, obteve-se um modelo estrutural com resolução de 17 Å. A análise da estrutura mostrou coerência com as informações disponíveis para outras hemocianinas, notadamente a da tarântula *E. californicum*. Com a análise dos resultados obtidos pôde-se confirmar a participação de domínios específicos das sub-unidades constituintes na oligomerização do complexo. Esperamos, como perspectiva deste trabalho, realizar o refinamento do modelo proposto, a partir dos dados disponíveis de *A. gomesiana*, juntamente com a modelagem por homologia das sub-unidades do complexo.

Com relação à hRXR α Δ AB, este trabalho mostrou, através do uso das técnicas de cromatografia analítica de exclusão por tamanho,

eletroforese em gel de poliacrilamida e SAXS, que dímeros dessa proteína podem ser encontrados em solução, não só na presença do ligante 9-cis-RA e do elemento responsivo DR1, mas também em sua forma não ligada.

Com relação ainda à proteína hRXR α Δ AB, também foi estudado a sua associação à diferentes elementos responsivos, DR1, DR4, F2 e PAL, através da técnica de anisotropia de fluorescência. Isso foi feito na presença e na ausência do ligante 9-cis-RA. Conclui-se que a proteína estudada possui maior afinidade por DR1 e DR4 do que por F2 na presença do ligante ($K_d = 8\pm 3$ nM, $K_d = 18\pm 6$ nM e $K_d = 92\pm 40$ nM, respectivamente), o que não ocorre na sua ausência ($K_d = 13\pm 4$ nM, $K_d = 26\pm 10$ nM e $K_d = 39\pm 21$ nM, respectivamente). Também se observa que a ligação da proteína aos elementos responsivos DR1 e DR4 ocorre com afinidades semelhantes, tanto na presença quanto na ausência do ligante. Devido à magnitude dos erros experimentais obtidos, não é possível observar diferenças significativas de afinidade entre os elementos DR1 e DR4. No experimento conduzido, não foi observada ligação ao elemento responsivo PAL, na ausência do ligante.

A dependência negativa da associação da proteína ao elemento F2, com relação à presença de 9-cis-RA, pode indicar a existência de um mecanismo, que na presença do ligante, diminui a afinidade do receptor para ligação a elementos responsivos não específicos, ainda que possa se ligar a estes quando na forma apo. Isso faria com que a transcrição só ocorresse quando o receptor estivesse associado ao DNA nas regiões de

interesse.

Os estudos conduzidos, ainda com a técnica de anisotropia de fluorescência, para verificar a associação de hRXR α Δ AB ao elemento responsivo DR1 em diferentes temperaturas mostram uma possível origem entrópica para essa associação. As constantes termodinâmicas de associação observadas para temperaturas entre 10°C e 20°C são $\Delta H^0 = 24 \pm 5$ kcal/mol e $\Delta S^0 = 0,121 \pm 0,017$ kcal/mol.K. Nesta faixa de temperaturas o valor de ΔG é negativo, oposto à variação da entropia, mostrando uma contribuição desfavorável de ΔH^0 , e indicando uma origem entrópica para a associação. Este tipo de mecanismo de associação proteína-DNA já foi observado anteriormente e parece estar relacionando à regiões hidrofóbicas do domínio de ligação da proteína ao DNA. Sendo assim, no caso de hRXR α Δ AB, este estudo merece um aprofundamento, inclusive com a investigação do mecanismo de associação existente para a proteína em sua forma fisiológica, com a região N-terminal.

6 REFERÊNCIAS

- ADRIAN, M.; DUBOCHET, L.; LEPAULT, J.; MCDOWALL, A. W. Cryoelectron microscopy of viruses. *Nature*, v. 308, p. 32-6, 1984.
- ALBERTS, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, v. 92, n. 3, p. 291-4. feb 1998.
- BAN, N.; NISSEN, P.; HANSEN, J.; MOORE, P. B.; STEITZ, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. *Science*, v. 289, n. 5481, p. 905-20. 2000.
- BELTRAMINI, M.; COLANGELO, N.; GIOMI, F.; BUBACCO, L.; DI MURO, P.; HELLMANN, N.; JAENICKE, E.; DECKER, H.. Quaternary structure and functional properties of *Penaeus monodon* hemocyanin. *FEBS J.*, v. 272, n. 8, p. 2060-75, 2005.
- BIJHOLT, M. M. C., VAN HEEL, M., VAN BRUGGEN, E. F. J. Comparison of 4 x 6-meric hemocyanins from three different arthropods using computer alignment and correspondence analysis. *J. Mol. Biol.*, v. 161, p. 139-53, 1982.
- BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, v. 72, p. 248-54, may 1976.
- CANTOR, C. R.; SCHIMMEL, P. R. *Biophysical Chemistry: Part III*. New

- York, W. H. Freeman and Company, 1969.
- CHEN, H.; PRIVALSKY, M. L. Cooperative Formation of High-order oligomers by Retinoid X Receptors: an unexpected mode of DNA recognition. *Proc. Natl. Acad. Sci. USA*, v. 92, p. 422-6, jan. 1995.
- CHEN, Y.; NOY, N. Retinoic Specificity of Interphotoreceptor Retinoic-Binding Protein. *Biochem.*, v. 33, p. 10658-65, 1994.
- CHEN, Z. P.; IYER, J.; BOURGUET, W.; HELD, P.; MIOSKOWSKI, C.; LEBEAU, L.; NOY, N.; CHAMBON, P.; GRONEMEYER, H. Ligand- and DNA-induced dissociation of RXR tetramers. *J. Mol. Biol.*, v. 275, n. 1, p. 55-65, jan. 1998.
- DECKER, H.; HARTMANN, H.; STERNER, R.; SCHWARZ, E.; PILZ, I. Small-angle X-ray scattering reveals differences between the quaternary structures of oxygenated and deoxygenated tarantula hemocyanin. *FEBS Lett.*, v. 393, n. 2-3, p. 226-30, 1996.
- DECKER, H.; RINKE, T. Tarantula hemocyanin shows phenoloxidase activity. *J. Mol. Chem.*, v. 273, n. 40, p. 25889-92, 1998.
- DE HAAS, F.; VAN BRUGGEN, E. F. J. The Interhexameric contacts in the 4-Hexameric Hemocyanin from the Tarantula *Eurypelma californicum* – A Tentative Mechanism for Cooperative Behavior. *J. Mol. Biol.*, v. 237, n. 4, p. 464-78, 1994.
- DE IOANNES, P.; MOLTEDO, B.; OLIVA, H.; PACHECO, R.; FAUNES, F.; DE IOANNES, A. E.; BECKER, M. I. Hemocyanin of the molluscan *Concholepas concholepas* exhibits an unusual heterodecameric array of subunits. *J. Biol. Chem.*, v. 279, n. 25, p. 26134-42, 2004.
- DIAS, S. M. G. *Estudos estruturais dos receptores nucleares humanos dos hormônios tireoidianos isoforma 'beta'1 (hTR'beta'1) e do ácido retinóico 9-cis isoforma 'alfa' (hRXR'alfa)*. São Carlos, 2004, p. 207. Tese (Doutorado) – Instituto de Física de São Carlos, Universidade de São Paulo.
- DUBE, P.; ORLOVA, E. V.; ZEMLIN, F.; VAN HEEL, M.; HARRIS, J. R.; MARKL, J. Three-dimensional structure of keyhole limpet hemocyanin by cryoelectron microscopy and angular reconstitution. *J. Struct. Biol.*, v. 115, n. 3, p. 226-32, 1995.
- DUBOCHET, L.; LEPAULT, J.; FREEMAN, R.; BERRIMAN, J. A.; HOMO, J. C. Electron microscopy of frozen water and aqueous solutions. *J.*

- Microsc.*, v. 128, p. 219-37, 1982.
- EISBERG, R.; RESNICK, R. *Física Quântica: átomos, moléculas, sólidos, núcleos e partículas*. Trad. de Paulo Costa Ribeiro, Enio Frota da Silveira e Marta Feijó Barroso. Rio de Janeiro, Campus, 1988.
- ERKER, W.; BEISTER, U.; DECKER, H. Cooperative transition in the conformation of 24-mer tarantula hemocyanin upon oxygen binding. *J. Biol. Chem.*, v. 280, n. 13, p. 12391-396, 2005.
- FASMAN, G. D., *Circular Dichroism and the Conformational Analysis of Biomolecules*, New York, Plenum Press, 1996.
- FISCHER, H.; DIAS, S. M. G.; SANTOS, M. A. M.; ALVES, A. C.; ZANCHIN, N.; CRAIEVICH, A. F.; APRILETTI, J. W.; BAXTER, J. D.; WEBB, P.; NEVES, F. A. R.; RIBEIRO, R. C. J.; POLIKARPOV, I. Low resolution structures of the retinoid X receptor DNA-binding and ligand-binding domains revealed by synchrotron x-ray solution scattering. *J. Biol. Chem.*, v. 278, n. 18, p. 16030-38, 2003.
- FOGUEL, D.; SILVA, J. L. Cold denaturation of a repressor-operator complex – the role of entropy in protein-DNA recognition. *Proc. Nat. Acad. Sci.*, v. 91, n. 17, p. 8244-47, 1994.
- FRANK, J. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*. New York, Oxford University Press, 2006.
- GAVIN, A. C.; BOSCHE, M.; KRAUSE, R.; GRANDI, P.; MARZIOCH, M.; BAUER, A.; SCHULTZ, J.; RICK, J. M.; MICHON, A. M.; CRUCIAT, C. M.; REMOR, M.; HOFERT, C.; SCHEIDER, M.; BRAJENOVIC, M.; RUFFNER, H.; MERINO, A.; KLEIN, K.; HUDAK, M.; DICKSON, D.; RUDI, T.; GNAU, V.; BAUCH, A.; BASTUCK, S.; HUHSE, B.; LEUTWEIN, C.; HEURTIER, M. A.; COPLEY, R. R.; EDELMANN, A.; Querfurth, E. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, v. 415, n. 6868, p. 123-4, 2002.
- GLASS, C. K. Differential Recognition of Target Genes by Nuclear Receptor Monomers, Dimers and Heterodimers. *Endocrine Reviews*, v. 15, n. 3, p. 391-407, sept. 2005.
- GLATER, G. *Small Angle X-ray Scattering*. London, Academic Press, 1982.
- GUINIER, A.; FOURNET, G. *Small Angle Scattering of X-rays*. New York, John Wiley and Sons, 1955.

- HARTWELL, L. H.; HOPFIELD, J. J.; LEIBLER, S.; MURRAY, A. W. From molecular to modular cell biology. *Nature*, v. 402. n. 6761, p. C47-C52, 1999.
- HAZES, B.; MAGNUS, K. A.; BONAVENTURA, C.; BONAVENTURA, J.; DAUTER, Z.; KALK, K. H.; HOL, W. G. J. Crystal-Structure of Deoxygenated Limulus-Polyphemus Subunit-II Hemocyanin at 2.18 Å Resolution. Clues for a Mechanism for Allosteric Regulation. *Protein Sci.*, v. 2, n. 4, p. 597-619, 1993.
- HENDERSON, R. Realizing the potential of electron cryo-microscopy. *Quarterly Reviews of Biophysics*, v. 37, n. 1, p. 3-13. 2004.
- HO, Y.; GRUHLER, A.; HEILBUT, A.; BADER, G. D.; MOORE, L.; ADAMS, S. L.; MILLAR, A.; TAYLOR, P.; BENNETT, K.; BOUTILIER, K.; YANG, L.; WOLTING, C.; DONALDSON, I.; SCHANDORFF, S.; SHEWNARANE, J.; VO, M.; TAGGART, J.; GOUDREAU, M.; MUSKAT, B.; ALFARANO, C.; DEWAR, D.; LIN, Z.; MICHALICKOVA, K.; WILLEMS, A. R.; SASSI, H.; NIELSEN, P. A.; RASMUSSEN, K. J.; ANDERSEN, J. R.; JOHANSEN, L. E.; HANSEN, L. H.; JESPERSEN, H.; PODTELEJNIKOV, A.; NIELSEN, E.; CRAWFORD, J.; POULSEN, V.; SORENSEN, B. D.; MATTHIESEN, J.; HENDRICKSON, R. C.; GLEESON, F.; PAWSON, T.; MORAN, M. F.; DUROCHER, D.; MANN, M.; HOGUE, C. W.; FIGEYS, D.; TYERS, M. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, v. 415, n. 6868, p. 123-4, 2002.
- HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics*, v. 14, p. 33-8, 1996.
- JEN-JACOBSON, L.; ENGLER, L. E.; JACOBSON, L. A. Structural and thermodynamic strategies for site-specific DNA binding proteins. *Structure*, v. 8, n. 10, p. 1015-23, 2000.
- KE, A. L.; MATHIAS, J. R.; VERSHON, A. K.; WOLBERGER, C. Structural and thermodynamic characterization of the DNA binding properties of a triple alanine mutant of MAT alpha 2. *Structure*, v. 10, n. 7, p. 961-71, 2002.
- KELLER, H.; LIEB, B.; ALTENHEIN, B.; GEBAUER, D.; RICHTER, S.; STRICKER, S.; MARKL, J. Abalone (*Haliotis tuberculata*) hemocyanin type 1 (HtH1) - Organization of the approximate to 400 kDa subunit, and amino acid sequence of its functional units f, g and h. *Eur. J. Biochem.*, v. 264, n. 1, p. 27-38, 1999.
- KERSTEN, S.; KELLEHER, D.; CHAMBON, P.; GRONEMEYER, H.; NOY,

- N. Retinoic X receptor alpha forms tetramers in solution. *Proc. Natl. Acad. Sci. USA*, v. 92, p. 8645-9,1995a.
- KERSTEN, S.; PAN, L.; CHAMBON, P.; GRONEMEYER, H.; NOY, N. On the role of ligand in retinoid signaling. 9-cis-retinoic acid modulates the oligomeric state of the retinoid X receptor. *Biochem.*, v. 34, n. 42, p. 13717-21,1995b.
- KERSTEN, S.; DAWSON, M. I.; LEWIS, B. A.; NOY, N. Individual Subunits of Heterodimers Comprised of Retinoic Acid and Retinoic X Receptors Interact with Theirs Ligands Independently. *Biochem.*, v. 35, p. 3816-24,1996.
- KERSTEN, S.; GRONEMEYER, H.; NOY, N. The DNA binding pattern of the retinoid X receptor is regulated by ligand-dependent modulation of its oligomeric state. *J. Biol. Chem.*, v. 272, n. 19, p. 12771-7,1997a.
- KERSTEN, S.; RECZEK, P. R.; NOY, N. The tetramerization region of the retinoid X receptor is important for transcriptional activation by the receptor. *J. Biol. Chem.*, v. 272, n. 47, p. 29759-68,1997b.
- KERSTEN, S.; DONG, D.; LEE, W.; RECZEK P. R.; NOY, N. Auto-silencing by the Retinoic X Receptors. *J. Mol. Biol.*, v. 284, p. 21-32,1998.
- KLIEWER, S. A. ; UMESONO, K.; MANGELSDORF, D. J.; EVANS, R. M. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature*, v. 355, N. 6359, P. 446-9, 1992.
- LAKOWICZ, J. R. *Principles of Fluorescence Spectroscopy*. 2.ed. New York, Plenum Press. 1983.
- LAUDET, V.; GRONEMEYER, H. *The Nuclear Receptor*. London, Academic Press, 2002.
- LEID, M.; KASTNER, P.; LYONS, R.; NAKSHATRI, H.; SAUNDERS, M.; ZACHAREWSKI, T.; CHEN, J. Y.; STAUB, A.; GARNIER, J. M.; MADER, S. Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently. *Cell*, v. 68, n. 2, p. 377-95, 1992.
- LIMA, L. M. T. R.; SILVA, J. L. Positive contribution of hydration on DNA binding by E2c protein from papillomavirus. *J. Biol. Chem.*, v. 279, n. 46, p. 47968-74, 2004.

- LIN, B. C.; WONG, C.; CHEN, H.; PRIVALSKY, M. L. Plasticity of Tetramer Formation by Retinoid X Receptors: an alternative paradigm for DNA recognition. *J. Biol. Chem*, v. 272, n. 15, p. 9860-7, 1997.
- LORENZINI, D. M.; FUKUZAWA, A. H.; DA SILVA, P. I.; MACHADO-SANTELLI, G.; BIJOVSKY, A. T.; DAFFRE, S. Molecular cloning, expression analysis and cellular localization of gomesin, an anti-microbial peptide from hemocytes of the spider, *Acanthoscurria gomesiana*. *Insect Biochem. Mol. Biol.*, v. 33, n. 10, p. 1011-16, 2003.
- LORENZINI, D. M.; SILVA, P. I. D.; SOARES, M. B.; ARRUDA, P.; SETUBAL, J.; DAFFRE, S. Discovery of immune-related genes expressed in hemocytes of the tarantula spider *Acanthoscurria gomesiana*. *Dev. Comp. Immunol.*, v. 30, n. 6, p. 545-56, 2006.
- LUDTKE, S. J.; BALDWIN, P. R.; CHIU, W. EMAN: Semiautomated software for high-resolution single-particle reconstruction. *J. Struct. Biol.*, v. 128, n. 1, p. 82-97, dec 1999.
- MADER, S.; LEROY, P.; CHEN, J. Y.; CHAMBON, P. Multiple parameters control the selectivity of nuclear receptors for their response elements - selectivity and promiscuity in response element recognition by retinoic acid receptors and retinoid-X receptors. *J. Biol. Chem*, v.268, n.1, p. 591-600, 1993.
- MANGELSDORF, D. J.; ONG, E. S.; DYCK, J. A.; EVANS, R. M. Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature*, v. 345, n. 6272, p. 224-9, 1990.
- MATTICK JS. RNA regulation: a new genetics? *Nat Rev Genet.*, v. 5, n. 4, p. 316-23, 2004.
- NICHOLSON, W. V.; GLAESER, R. M. Review: Automatic Particle Detection in Electron Microscopy, *J. Struct. Biol.*, v. 133, p. 90-101, 2001.
- ORLOVA, E. V.; DUBE, P.; HARRIS, J. R.; BECKMAN, E.; ZEMLIN, F.; MARKL, J.; VAN HEEL, M. Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 angstrom resolution by electron cryomicroscopy and angular reconstitution. *J. Mol. Biol.*, v. 271, n. 3, p. 417-37, 1997.
- PERUTZ, M. F.; BOLTON, W.; DIAMOND, R.; MUIRHEAD, H.; WATSON, H. C. Structure of haemoglobin. An x-ray examination of reduced horse haemoglobin. *Nature*, v. 203, p. 687-90, 1964.

- PICARD, D. *The Nuclear Receptor: A Practical Approach*. New York, Oxford University Press, 1999.
- PLAISIER, J. R.; KONING, R. I.; KOERTEN, H. K.; VAN HEEL, M.; ABRAHAMMS, J. P. TYSON: Robust searching, sorting, and selecting of single particles in electron micrographs. *J. Struct. Biol.*, v. 145, p. 76-83, 2004.
- ROYER, C. A. Improvements in the Numerical Analysis of Thermodynamic Data from Biomolecular Complexes. *Anal. Biochem.*, v. 210, p. 91-7, 1993.
- ROYER, C. A.; SMITH, W. R.; BEECHEM, J. M. Analysis of Binding in Macromolecular Complexes: A Generalized Numerical Approach. *Anal. Biochem.*, v. 191, p. 287-94, 1990.
- SALI A, GLAESER R, EARNEST T, BAUMEISTER W. From words to literature in structural proteomics. *Nature*, v. 422, n. 6928, p. 216-25, 2003.
- SCHATZ, M.; ORLOVA, E. V.; DUBE, P.; JAGER, J.; VAN HEEL, M. Structure of Lumbricus-Terrestris Hemoglobin at AT 30 Å Resolution Determined using Angular Reconstitution. *J. Struct. Biol.*, v. 114, n. 1, p. 28-40, 1995.
- SCHATZ, M.; ORLOVA, E. V.; DUBE, P.; STARK, H.; ZEMLIN, F.; VAN HEEL, M.; Angular Reconstitution in Three-dimensional Electron Microscopy: Practical and Technical Aspects. *Scanning Microscopy.*, v. 11, p. 179-93. 1997.
- SERYSHEVA, I. I.; ORLOVA, E. V.; CHIU W.; SHERMAN, M. B.; HAMILTON, S. L.; VAN HEEL, M. Electron cryomicroscopy and angular reconstitution used to visualize the skeletal muscle calcium release channel. *Nature struct. Biol.*, v. 2, p. 18-24, 1995.
- TAVEAU, J. C.; BOISSET, N.; LAMY, J.; LAMBERT, O.; LAMY, J. N. Three-dimensional reconstruction of Limulus polyphemus hemocyanin from cryoelectron microscopy. *J. Mol. Biol.*, v. 266, n. 5, p. 1002-15, 1997
- VAN HEEL, M. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*. v. 21, n. 2, p. 111-23, 1987.
- VAN HEEL, M.; DUBE, P. Quaternary Structure of Multihexameric of Arthropod Hemocyanins. *Micron.*, v. 25, n. 4, p. 387-418, 1994.

- VAN HEEL, M.; FRANK, J. Use of multivariate statistical statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, v. 6, p. 187-94, 1981.
- VAN HEEL, M.; GOWEN, B.; MATADEEN, R.; ORLOVA, E. V.; FINN, R.; PAPE, T.; COHEN, D.; STARK, H.; SCHMIDT, R.; SCHATZ, M.; PATWARDHAN, A. Single-particle electron cryo-microscopy: towards atomic resolution. *Quarterly Reviews of Biophysics*, v. 33, n.4, p. 307-69. 2000.
- VAN HEEL, M.; HARAUZ, G.; ORLOVA, E. V.; SCHMIDT, R.; SCHATZ, M. A new generation of IMAGIC image processing system. *J. Struct. Biol.*, v. 116, n. 1, p. 17-24. jan-feb 1996.
- VAN HEEL, M.; SCHATZ, M. Fourier Shell Correlation Threshold Criteria. *Struct. Biol.*, v. 151, n. 3, p. 250-62. jsep 2005.
- VAN HEEL, M.; STÖFFLER-MEILICKE, M. The characteristic views of *E. Coli* and *B. Stearothermophilus* 30S ribosomal subunit in the electron microscope. *EMBO J.*, v. 4, p. 2389-95, 1985.
- VAN HEEL, M.; WINKLER, H.; ORLOVA, E. V.; SCHATZ, M. Structure analysis of ice-embedded single particles, scanning microsc. *Suppl.*, v. 66, p. 23-42, 1992.
- VAN SCHAICK, E. J. M.; SCHUTTER, W. G.; GAYKEMA, W. P. J.; SCHEPMAN, A. M. H.; HOL, W. G. J. Structure of *Panulirus interruptus* Haemocyanin at 5 Å Resolution. *J. Mol. Biol.*, v. 158, p. 457-85, 1982.
- VOIT, R.; FELDMAIER-FUCHS, G.; SCHWEIKARDT, T.; DECKER, H.; BURMESTER, T. Complete sequence of the 24-mer hemocyanin of the tarantula *Eurypelma californicum* - Structure and intramolecular evolution of the subunits. *J. Biol. Chem.*, v. 275, n. 50, p. 39339-44, 2000.
- VOLBEDA, A.; HOL, W. G. J. Crystal Structure of Hexameric Haemocyanin from *Panulirus interruptus* Refined at 3.2 Å Resolution. *J. Mol. Biol.*, v. 209, p. 249-79, 1989.
- WEBER, G. Polarization of the fluorescence of macromolecules. I. Theory and experimental method. *Biochem. J.*, v. 51, p. 145-55, 1952.
- WIMBERLY, B. T.; BRODERSEN, D. E.; CLEMONS, W. M.; MORGAN-WARREN, R. J.; CARTER, A. P.; VONRHEIN, C.; HARTSCH, T.; RAMAKRISHNAN, V. Structure of the 30S ribosomal subunit. *Nature*, v. 407, n. 6802, p. 327-39, 2000.

WRIGGERS, W.; MILLIGAN, R. A.; MCCAMMON, J. A. SITUS: A Package for Docking Crystal Structures into Low-Resolution Maps from Electron Microscopy. *J. Struct. Biol.*, v. 125, p. 185-95, 1999.

WRIGGERS, W.; CHACÓN, P. Modeling Tricks and Fitting Techniques for Multi-Resolution Structures. *Struct.*, v. 9, p. 779-88, 2001.

ZHANG, X. K.; HOFFMANN, B.; TRAN, P. B.; GRAUPNER, G.; PFAHL, M. Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. *Nature*, v. 355, n. 6359, p. 441-6, 1992.

ZHAO, Q.; CHASSE, S. A.; DEVARAKONDA, S.; SIERK, M. L.; AHVAZI, B.; RASTINEJAD, F. Structural Basis of RXR-DNA Interactions. *J. Mol. Biol.*, v. 296, p. 509-20, 2000.

APÊNDICE

7 APÊNDICE A

Trabalhos publicados em colaboração

GRIMM, E. D.; PORTUGAL, R. V.; OLIVEIRA NETO, M.; MARTINS, N. H.; POLIKARPOV, I.; ZAHA, A.; FERREIRA, H. B. Structural analysis of an *Echinococcus granulosus* actin fragmenting protein by SAXS studies and molecular modeling. *Biophysical Journal.*, v. 90, n. 9, p. 3216-23, 2006.

SANTOS, M. A. M.; SANTOS, S. M.; MATOZO, H. C.; PORTUGAL, R. V.; IULIANO, R.; FUSCO, A.; POLIKARPOV, I. Expression, purification, and characterization of rat protein tyrosine phosphatase eta catalytic domain. *Protein Expression and Purification*, v. 41, n. 1, p. 113-20, 2005.

NUNES, F. M.; APARICIO, R.; SANTOS, M. A. M.; PORTUGAL, R. V.; DIAS, S. M. G.; NEVES, F. A. R.; SIMEONI, L. A.; BAXTER, J. D.; WEBB, P.; POLIKARPOV, I. Crystallization and preliminary X-ray diffraction studies of isoform alpha1 of the human thyroid hormone receptor ligand-binding domain. *Acta Crystallographica Section D-Biological Crystallography*, v. 60, p. 1867-70, 2004.

RAZZERA, G.; VERNAL, J.; CALGARO, M. R.; PORTUGAL, R. V.; FERNANDEZ, P.; ZAKIN, M. M.; POLIKARPOV, I.; TEREZI, H. Expression, purification, and initial structural characterization of rat orphan nuclear receptor NOR-1 LBD domain. *Protein Expression And Purification*, V. 37, N. 2, P. 443-49, 2004.

Structural Analysis of an *Echinococcus granulosus* Actin-Fragmenting Protein by Small-Angle X-Ray Scattering Studies and Molecular Modeling

Eliana D. Grimm,* Rodrigo V. Portugal,[†] Mário de Oliveira Neto,[†] Nádia H. Martins,[†] Igor Polikarpov,[†] Arnaldo Zaha,*[‡] and Henrique B. Ferreira*[‡]

*Laboratório de Biologia Molecular de Cestódeos, Centro de Biotecnologia, and [‡]Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre RS, Brazil; and [†]Grupo de Cristalografia, Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos SP, Brazil

ABSTRACT The *Echinococcus granulosus* actin filament-fragmenting protein (EgAFFP) is a three domain member of the gelsolin family of proteins, which is antigenic to human hosts. These proteins, formed by three or six conserved domains, are involved in the dynamic rearrangements of the cytoskeleton, being responsible for severing and capping actin filaments and promoting nucleation of actin monomers. Various structures of six domain gelsolin-related proteins have been investigated, but little information on the structure of three domain members is available. In this work, the solution structure of the three domain EgAFFP has been investigated through small-angle x-ray scattering (SAXS) studies. EgAFFP exhibits an elongated molecular shape. The radius of gyration and the maximum dimension obtained by SAXS were, respectively, 2.52 ± 0.01 nm and 8.00 ± 1.00 nm, both in the absence and presence of Ca^{2+} . Two different molecular homology models were built for EgAFFP, but only one was validated through SAXS studies. The predicted structure for EgAFFP consists of three repeats of a central β -sheet sandwiched between one short and one long α -helix. Possible implications of the structure of EgAFFP upon actin binding are discussed.

INTRODUCTION

The larval stage of the cestode tapeworm *Echinococcus granulosus* is the causative agent of cystic hydatid disease or hydatidosis, recognized as one of the world's major zoonoses (1). This parasite requires two mammalian hosts for completion of its life cycle. Adult tapeworms develop in the small intestine of definitive hosts (domestic dogs and wild canids), whereas the metacestode or hydatid cyst usually develops in the liver or lungs of intermediate hosts (mainly in ungulates, and accidentally in humans). The pathological effect of the disease is caused by the pressure exerted by the hydatid cyst on the intermediate host's viscera. Within the cyst, protoscoleces are produced by asexual reproduction and develop into the adult worm when ingested by the definitive host.

The *E. granulosus* complex life cycle involves important changes in cell morphology and physiology (2). The molecular and cellular mechanisms involved in *E. granulosus* development are still largely unknown but are likely to require extensive cytoskeleton reorganization (3,4).

The actin cytoskeleton is a vital component of several key cellular and developmental processes in eukaryotes, such as motility, cytokinesis, cytoplasmic organization, and endocytosis (5). In cells, the assembly and disassembly of actin filaments, in addition to their organization into functional three-dimensional (3D) networks, are regulated by a variety

of actin-binding proteins (6–10). Among these proteins, those from the gelsolin superfamily control actin organization by severing filaments, capping filament ends, and nucleating actin assembly (11).

The best-studied members of this protein family are severin (12–14) and fragmin (15,16) from *Dictyostelium discoideum* and *Physarum polycephalum*, respectively, and gelsolin (17,18) and villin (19,20) from higher organisms. A common feature of this family is the segmental organization into three (severin, fragmin) or six (gelsolin, villin) homologous domains that might have evolved from an ancestral one domain protein through a stepwise process, involving a gene triplication followed by an additional duplication event (21,17). The activities of these proteins are often modulated by signaling molecules, such as Ca^{2+} or phosphorylated phosphoinositides (22). Based on gelsolin (23,24), the most extensively studied member of the family, it is generally accepted that the second domain binds F-actin, whereas the first domain (and the fourth one, for six domain proteins) binds G-actin.

So far, only a few proteins of the gelsolin superfamily have solved 3D structures. A search in the database of protein structures indicates that of all known members of this family to date, gelsolin is the only protein that has its full-length (six domains) structure solved (25). Structures of other proteins, like villin (26) and severin (27,28), have been determined only for the first or second domains, usually bound to ligands (Ca^{2+} and/or actin). Domain comparisons between the known structures show that they all share a common fold built around a central five-stranded mixed

Submitted June 1, 2005, and accepted for publication December 12, 2005.

Address reprint requests to Dr. Henrique B. Ferreira, Laboratório de Biologia Molecular de Cestódeos, Centro de Biotecnologia, UFRGS, Caixa Postal 15005, Porto Alegre RS 91501-970, Brazil. Tel.: 55-51-33167768; Fax: 55-51-33167309; E-mail: henrique@cbiot.ufrgs.br.

© 2006 by the Biophysical Society

0006-3495/06/05/3216/08 \$2.00

doi: 10.1529/biophysj.105.067801

β -sheet, which is flanked by a long α -helix running parallel to the sheet and a short perpendicular running α -helix (25–29).

Our laboratory has previously cloned and functionally characterized a 42 kDa actin filament-fragmenting protein from *E. granulosus* (EgAFFP) (30). The recombinant EgAFFP protein is recognized by sera of ~69% of human hydatid disease patients (31) and, in vitro, was able to induce actin polymerization and sever actin filaments, confirming that it belongs to the gelsolin superfamily (30). According to sequence analysis, EgAFFP presents three repeated domains and is similar (36% identity) to the gelsolin NH₂-terminal half (G1–G3). The lack of structural data for full-length three domain members of the gelsolin superfamily, such as EgAFFP, represents an obstacle to the understanding of structure-function relationships of these smaller proteins, which are functionally equivalent to their six domain counterparts. Structural characterization of EgAFFP might help to understand how three domain members function and how they are regulated by calcium.

There is a relatively high degree of sequence identity between gelsolin and orthologous proteins, such as villin and severin, particularly in regions that are buried or forming secondary structure (25). Therefore, it has been suggested that the domain folds in gelsolin may be used to predict the geometries for the as of yet unsolved structures of the corresponding domains in other members of the gelsolin superfamily. Homology or comparative protein structure modeling allows the construction of a 3D model for a given protein sequence based on its similarity to one or more known structures. It is the most reliable method to predict the 3D structure of a protein, with an accuracy that can be comparable to that of a low resolution, experimentally determined structure (32). The accuracy of comparative modeling is correlated with sequence identity between the template structure and the modeled sequence (33). High accuracy models can be obtained when template structures with >50% sequence identity are available, medium accuracy with 30% \pm 50% sequence identity, and low accuracy with <30% sequence identity.

Small-angle x-ray scattering (SAXS) is a well-established technique for obtaining structural information on macromolecules in solution under close to physiological conditions (34) and is particularly useful when high resolution structural information, such as that from NMR or x-ray crystallography, is not yet available (35) as is the case for EgAFFP. Previous work has shown that SAXS data can be conveniently used to confirm computer-modeled protein structures (36,37).

In this study, we constructed a homology model for EgAFFP both in the presence and in the absence of Ca²⁺, based on its similarity with horse plasma gelsolin (Protein Data Bank (PDB) id: 1RGI (38) and 1D0N (25), respectively). We also obtained an ab initio low resolution model of EgAFFP through SAXS studies, which is in agreement with the structure of our homology model based on Ca²⁺-free gelsolin (PDB id: 1D0N) but differs significantly from the

homology model generated from Ca²⁺-gelsolin bound to actin (PDB id: 1RGI). We also examined changes in EgAFFP conformation caused by Ca²⁺ and found no significant variations in the molecular envelope of the protein upon addition of Ca²⁺. The functional implications of the proposed structural model upon actin binding are discussed.

MATERIALS AND METHODS

Protein expression and purification

Full-length EgAFFP (370 aa) (30) was expressed as a fusion with glutathione-S-transferase using the pGEX system (Amersham Biosciences, Uppsala, Sweden) and purified by affinity chromatography according to Smith and Johnson (39). The fusion protein was cleaved with thrombin (10 units/mg of matrix-bound protein) for 12 h at 4°C, and the recombinant EgAFFP was eluted in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4) plus 3 mM dithiothreitol (DTT) and 1 mM phenylmethylsulfonyl fluoride (PMSF) and concentrated in Centricon YM-10 microconcentrators (Millipore, Bedford, MA) to 12 mg/ml. Protein purity was evaluated by SDS-PAGE (40), and concentration was determined using the Bio-Rad dye assay (Bio-Rad, Hercules, CA), with BSA as standard. Typically, 30 mg of recombinant protein were obtained from 1 liter of culture.

Pyrene actin preparation

Pyrene actin was prepared according to Cooper et al. (41) with minor modifications. Monomeric actin (G-actin) from bovine skeletal muscle (Sigma-Aldrich, St. Louis, MO) was diluted in buffer G (2 mM Tris-HCl, pH 8.0, 0.2 mM ATP, 0.5 mM DTT, 0.2 mM CaCl₂, 0.01% NaN₃) to a concentration of 1 mg/ml. G-actin was polymerized at room temperature for 1 h by adding KCl and MgCl₂ to 100 mM and 2 mM, respectively. *N*-(1-pyrene) iodoacetamide (Molecular Probes, Eugene, OR) freshly dissolved in dimethyl formamide was added from a 5 mM stock solution to the F-actin (final concentration of 160 μ M) and incubated overnight at 4°C in the dark. The pyrene F-actin was collected by centrifugation at 163,000 \times *g* for 90 min at 4°C, resuspended in 1 ml of buffer G, and dialyzed against the same buffer for 48 h with several changes of buffer to depolymerize the actin and remove free pyrenyl. The pyrene G-actin was then clarified at 163,000 \times *g* for 90 min at 4°C and divided into 50 μ l aliquots, which were frozen in liquid nitrogen and stored at –70°C. Before use, the pyrene G-actin was thawed on ice and then centrifuged at 16,000 \times *g* for 15 min at 4°C to remove larger precipitates. After the absorbance at 344 nm and 290 nm was determined, the total actin concentration was calculated using mM actin = (A₂₉₀ – 0.127 A₃₄₄)/26.6 mM^{–1} cm^{–1}. The pyrene/actin ratio was calculated using an extinction coefficient of 22,000 M^{–1} cm^{–1} for pyrene with pyrene/actin = (A₃₄₄/22.0)/mM actin.

Fluorescence spectroscopy

All fluorescence measurements with pyrene actin were performed at 20°C using an ISS K2 fluorimeter (ISS, Fluorescence, Analytical and Biomedical Instrumentation; Champaign, IL). The excitation and emission wavelength were set at 365 nm and 407 nm, respectively, with sample volumes of 300 μ l. To avoid bleaching the fluorophore, pyrene actin was prepared and stored in the dark and, during fluorescence experiments, it was exposed to the light source only intermittently. The data were plotted as arbitrary relative fluorescence values.

For nucleation assays, actin (6 μ M final concentration, 10% of which was pyrene labeled) was preincubated for 10 min with EgAFFP (250 nM). The polymerization was initiated by the addition of buffer F (buffer G with 100

mM KCl and 2 mM MgCl₂) from a 10× concentrated stock solution, and the increase in fluorescence was recorded as a function of time.

For severing assays, actin (8 μM, 25% pyrene labeled) was polymerized in buffer F for 15 min. Subsequently, the filaments were prepacked with recombinant gelsolin (23) (40 nM final concentration) to avoid measuring nucleation activity. The mixture was stored overnight on ice. This solution was then diluted in buffer G to a final concentration of 400 nM in the absence or presence of EgAFFP (10 nM), and fluorescence was recorded.

Modeling of EgAFFP

The PDB database was searched for proteins homologous to EgAFFP with solved atomic structures using the BLAST program (42). The selected sequences were aligned with CLUSTALW (43), and model building by homology was performed with the MODELLER 6 program (44). For a given alignment, 20 model structures were built and evaluated with the PROCHECK software suite (45). Only the best-evaluated model was retained after the analysis.

SAXS measurements and data analysis

SAXS data were collected at the small-angle scattering beamline on the LNLS (National Synchrotron Light Laboratory, Campinas, Brazil) using a one-dimensional position-sensitive detector. EgAFFP samples at a concentration of 4, 8, and 12 mg/ml were prepared in phosphate-buffered saline (pH 7.4) in the absence or presence of 0.2 mM CaCl₂. EGTA was added to the Ca²⁺-free samples to a final concentration of 1 mM. Samples were measured at a wavelength of 1.488 Å for a sample-detector distance of 1634.5 mm covering the momentum transfer range 0.21 nm⁻¹ < *q* < 2.35 nm⁻¹ (*q* = 4πsinθ/λ, where 2θ is the scattering angle). The scattering curves of the protein solutions and the corresponding solvents were collected in 20 frames of 100 s each to monitor radiation damage and beam stability. The data were normalized to the intensity of the incident beam and corrected for detector response. The scattering of the buffer was subtracted, and the difference curves were scaled for concentration. The distance distribution functions *p*(*r*) and the radii of gyration *R*_g were evaluated by indirect Fourier transform using GNOM (46), which also corrected the smearing effect.

Ab initio molecular shape determination from SAXS measurements

The low resolution particle shape was restored using the ab initio procedure described by Petoukhov et al. (47), as implemented in the CREDO program. The known part of the high resolution structure was fixed, and missing portions, such as disordered loops and N-terminal residues, were then modeled to fit the experimental scattering data obtained from the intact particle. A simulated annealing procedure was used to minimize a scoring function that depends on the discrepancy between the experimental and calculated patterns and relevant penalty terms. The final shape restoration was performed using 436 residues and 378 waters assuming no molecular symmetry. The values for *R*_g, *D*_{max}, and the scattering curve simulated from the atomic coordinates of the EgAFFP homology models based both on Ca²⁺-free and Ca²⁺-bound gelsolin and the discrepancies between the homology model and experimental data were calculated using the CRY SOL program (48). The SUPCOMB program (49) was used to superpose the low resolution ab initio model to the structure of the homology model.

RESULTS

Functional integrity of recombinant EgAFFP

Recombinant EgAFFP functionality has been previously assessed and, compared to gelsolin, EgAFFP nucleates actin at a similar rate, whereas fragmentation is 20% less efficient

(30). To assess recombinant EgAFFP functionality before its structural characterization, the pyrene actin fluorescence assay was used to indirectly determine its actin-severing and -nucleating activities by monitoring the polymerization/depolymerization rate of actin. The ability of recombinant EgAFFP to nucleate actin assembly (Fig. 1 A) and sever actin filaments (Fig. 1 B) was confirmed by its direct effect on the rate of increase and decrease, respectively, in the fluorescence of pyrene-labeled actin. These results indicated that the recombinant EgAFFP samples used for structural characterization retained these Ca²⁺-dependent activities, which are typical of proteins of the gelsolin superfamily.

EgAFFP modeling

Sequence alignments with the database of structures revealed that the EgAFFP sequence has the highest sequence identity, 36% of identical residues and 58% of conserved residues, with the N-terminal half (amino acids 27–367) of horse plasma gelsolin (in the Ca²⁺-free form, PDB id: 1DON; (25)). Relatively high homology and the fact that 1DON is the only full-length structure of the gelsolin superfamily available justified its use as a template for EgAFFP modeling. For modeling, we aligned the EgAFFP target sequence to that of amino acids 27–367 of the 1DON template. The first 35

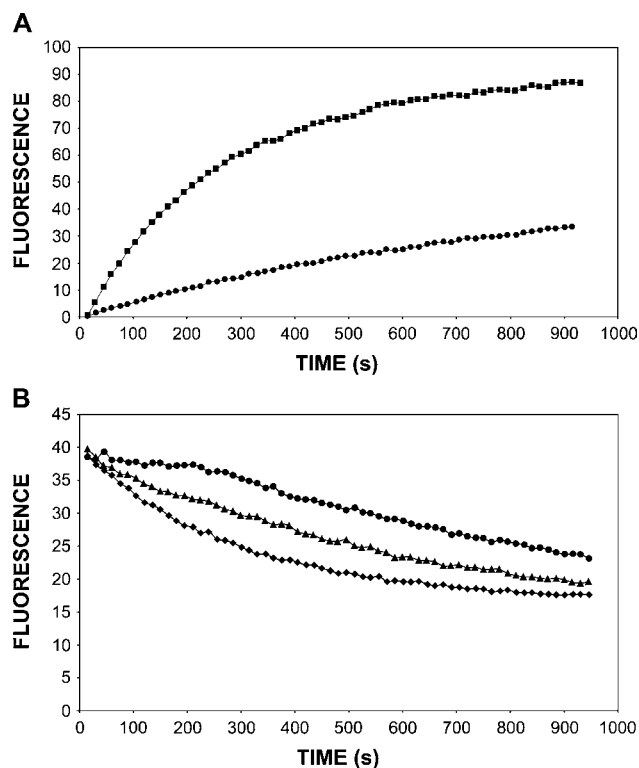


FIGURE 1 (A) Nucleating and (B) severing activity of recombinant EgAFFP. Assays were performed in the presence of Ca²⁺ as described in Materials and Methods. Nucleating activity of 60 nM EgAFFP (□). Severing activity of 4 nM (△) and 10 nM (◇) EgAFFP. Actin alone (○) was used as a control in both assays.

amino acids of the EgAFFP sequence, which would consist of two short α -helices, were eliminated, since they did not align with the template sequence. The obtained alignment was then used to generate a homology model (Fig. 2, *shaded*). As expected from the building procedure, the model resembles the N-terminal half of the 1D0N structure (domains G1–G3) and suggests that EgAFFP is composed of three independent domains (E1, Phe⁴⁹–Tyr¹⁴³; E2, Leu¹⁴⁷–His²⁵⁵; and E3, Ser²⁷⁸–Trp³⁶⁸), as predicted by sequence analysis (30). Each domain adopts a similar geometric fold composed of a core of mixed β -sheet comprising five or six strands sandwiched between a long helix (helix 1) running roughly parallel to the strands in the sheet, and a short helix (helix 2) running approximately perpendicular to the strands. The sheets of E1 and E3 form a continuous 10-stranded β -sheet. E1 also contributes a single strand to the sheet of E2, forming a continuous seven-stranded β -sheet. The C-terminus of E1 immediately enters the N-terminus of E2, whereas there is an extended 23-residue loop connecting E2 and E3, which runs over the central sheet of E1. EgAFFP was also modeled in the presence of Ca²⁺, based on the crystallographic structure of the Ca²⁺-bound N-terminal half (G1–G3 domains) of gelsolin:actin complex (PDB id: 1RGI; (38)), from which the actin fragment was discarded. The obtained model (Fig. 2, *solid*) shows that the two halves of the β -sheet that runs continuously from E1 to E3 in the Ca²⁺-free model are separated in the Ca²⁺-bound model, and E1 no longer interacts with E3. In contrast, the E2–E3 linker shortens and brings E2 close to E3, whereas the first strand of E2 extends, forming a long linker between E1 and E2.

SAXS data analysis

Experimental scattering curves from EgAFFP are presented in Fig. 3, and the structural parameters derived from these curves are given in Table 1. Guinier plots (50) of the data obtained at 4 mg/ml exhibited linear profiles, indicating satisfactory monodispersity of EgAFFP both in the absence and in the presence of Ca²⁺ (Fig. 3 A, *inset*). The scattering

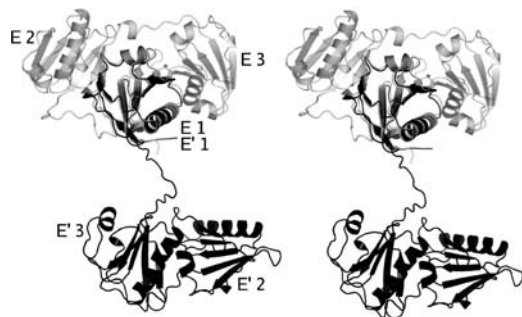


FIGURE 2 Stereo view of superposition of the EgAFFP homology model based on Ca²⁺-bound gelsolin (PDB id 1RGI) (*solid*) and EgAFFP homology model based on Ca²⁺-free gelsolin (PDB id 1D0N) (*shaded*). Domains E1, E2, and E3 are indicated in each model. The relative positions of both models were based on the superposition of the corresponding E1 domains.

TABLE 1 Structural parameters derived from SAXS data and models

Parameter/sample	EgAFFP absence of Ca ²⁺			EgAFFP presence of Ca ²⁺		
	Exp.*	Mod.†	DR‡	Exp.*	Mod.†	DR‡
D_{\max} (nm)	8.00 ± 1.00	7.68	8.53	8.00 ± 1.00	9.02	8.47
R_g (nm)	2.53 ± 0.01	2.37	2.35	2.51 ± 0.01	2.71	2.29
Discrepancy χ	–	1.29	1.15	–	6.00	1.13
Resolution (nm)	2.67	–	2.67	2.67	–	2.67

*Exp., calculated from the experimental data.

†Mod., parameters obtained from the EgAFFP homology models.

‡DR, parameters of the experimental SAXS low resolution structure, averaged over 10 independent DR models.

curves obtained in the presence or in the absence of Ca²⁺ were virtually identical at all measured protein concentrations (Fig. 3 A). The radius of gyration (R_g), estimated using the Guinier approximation, was 2.53 ± 0.01 nm in the absence of Ca²⁺ and 2.51 ± 0.01 nm in the presence of Ca²⁺. Since scattering curves at higher concentrations showed nonnegligible interference effects, we combined a low resolution part of the scattering curve obtained at 4 mg/ml (<0.5 nm⁻¹) with a high resolution part of the scattering curve measured at 12 mg/ml in a composed scattering curve, which was used in ab initio model reconstruction (Fig. 3 B). The pair distance distribution function $p(r)$ for both protein samples is shown in Fig. 4 and was calculated from the GNOM fitted data. These curves also provided the EgAFFP maximum dimension (D_{\max}), determined from the value of r for which the function $p(r)$ drops to zero. The determined EgAFFP D_{\max} was 8.0 ± 1.0 nm, both in the absence and in the presence of Ca²⁺. The ratio of the radius of gyration to the maximum dimension indicates an elongated shape for EgAFFP in solution, also indicated by the shape of the $p(r)$ function, typical of an asymmetric particle (51). To evaluate our computer-designed EgAFFP models, based on the structure of Ca²⁺-free and Ca²⁺-bound gelsolin domains G1–G3, the SAXS experimental curves were compared to the scattering intensity and the distance distribution $p(r)$ function calculated from their atomic coordinates. The scattering intensity and the distance distribution $p(r)$ curve calculated for the EgAFFP Ca²⁺-free homology model (presented in Figs. 3 B and 4, respectively) are in good agreement with the SAXS experimental data, indicating that scattering data are fully compatible with this structural model. Conversely, both scattering intensity and distance distribution function for the EgAFFP Ca²⁺-bound model grossly deviate from the scattering intensity curves and $p(r)$ function obtained from SAXS experiments (Figs. 3 B and 4).

Ab initio molecular shape determination from SAXS experiments

The low resolution ab initio models of EgAFFP in the absence and presence of calcium were both built using the

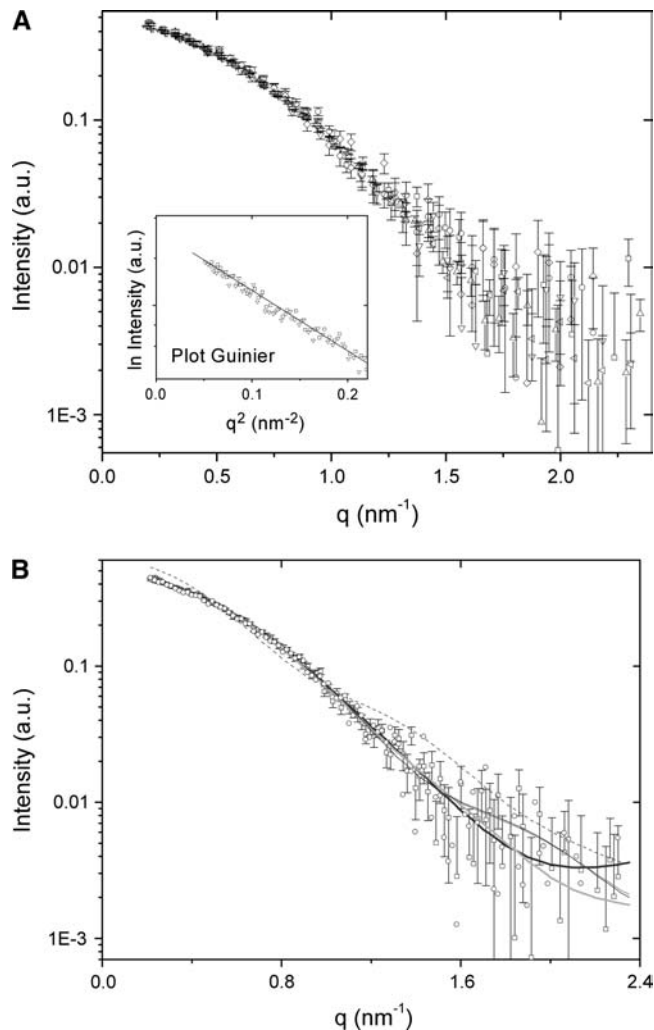


FIGURE 3 Experimental scattering curves for EgAFFP. (A) Experimental scattering curves corrected by the smearing effect of EgAFFP in the absence of Ca^{2+} at 4 mg/ml (\square), 8 mg/ml (\circ), and 12 mg/ml (\triangle) and in the presence of Ca^{2+} at 4 mg/ml (∇), 8 mg/ml (\diamond), and 12 mg/ml (\triangleleft). The Guinier plot with linear fit profiles is shown in the inset, allowing the estimation of $R_g = 2.53$ nm in the absence of Ca^{2+} , and $R_g = 2.51$ nm in the presence of Ca^{2+} . (B) Composed scattering curve of EgAFFP corrected by the smearing effect in the absence of Ca^{2+} (\square), in the presence of Ca^{2+} (\circ), and scattering intensities from the DR models (*thick solid line*), from the homology model based on the N-terminal half of Ca^{2+} -free gelsolin (PDB id 1D0N, *thick shaded line*), from the homology model based on the N-terminal half of Ca^{2+} -bound gelsolin (PDB id 1RGI, *traced line*), from the crystallographic structure of the Ca^{2+} -bound C-terminal half of gelsolin (PDB id 1NPH, *thin shaded line*), and from the x-ray structure of the C-terminal half of gelsolin bound to Ca^{2+} and actin (PDB id 1H1V, *thin solid line*). The actin structure was extracted from the later model.

CREDO program from the composed x-ray scattering curves. The models were derived from experimental data assuming P1 symmetry. To verify the validity of the dummy residue (DR) models, several independent constructions were made from different starting conditions and, in all cases, very similar envelopes were obtained. Virtually identical low resolution models were also obtained when the GASBOR

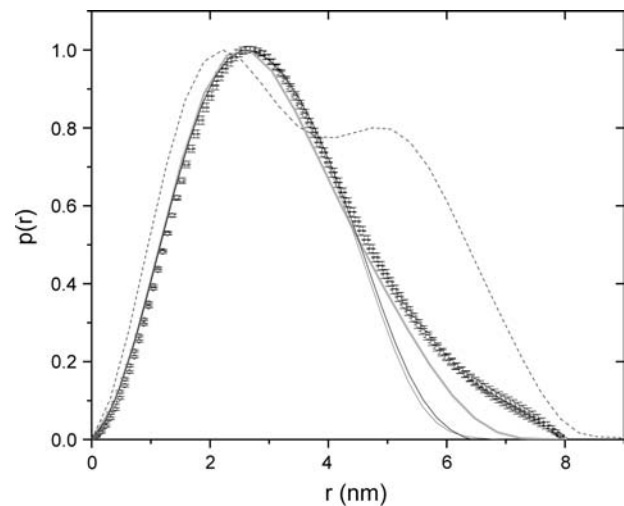


FIGURE 4 Comparison of distance distribution functions for EgAFFP. The $p(r)$ curves obtained were calculated using the GNOM program, and the experimentally derived distributions for EgAFFP are given in the absence (\square) and in the presence (\circ) of Ca^{2+} . Distributions were calculated for the homology model based on the N-terminal half of Ca^{2+} -free gelsolin (PDB id 1D0N, *thick shaded line*), for the homology model based on the N-terminal half of Ca^{2+} -bound gelsolin (PDB id 1RGI, *traced line*), for the crystallographic structure of the Ca^{2+} -bound C-terminal half of gelsolin (PDB id 1NPH, *thin shaded line*), and for the x-ray structure of the C-terminal half of gelsolin bound to Ca^{2+} and actin (PDB id 1H1V, *thin solid line*). The actin structure was extracted from the later model.

program (52) was used. The homology model of EgAFFP based on Ca^{2+} -free gelsolin, but not the model based on the Ca^{2+} /gelsolin:actin complex, could be unambiguously positioned inside the SAXS-derived molecular envelopes of EgAFFP both in the absence and in presence of Ca^{2+} (Fig. 5). The structural parameters calculated for the DR models and the Ca^{2+} -free EgAFFP homology model and the structural parameters obtained from experimental SAXS measurements (all presented in Table 1) show very good agreement, further supporting our computer-designed Ca^{2+} -free EgAFFP structure. Furthermore, molecular envelopes obtained from the SAXS curves measured for EgAFFP in the absence and presence of Ca^{2+} were virtually identical, as one would expect given the extremely high similarity of the initial scattering curves (Fig. 3). At the same time, structural parameters derived from the EgAFFP homology model based on the Ca^{2+} -bound gelsolin:actin complex are very different from the experimental data (Table 1). Moreover, fitting of the predicted x-ray scattering curve based on this model with the experimental scattering data (with and without Ca^{2+}) indicates gross differences between these curves (Fig. 3 B), which results in a discrepancy parameter $\chi = 6.0$. Given the fact that the structure of the N-terminal half of gelsolin (domains G1–G3) closely resembles its C-terminal half (domains G4–G6), we also compared the predicted x-ray scattering intensity curves and distance distribution for crystallographic models of G4–G6 domains of Ca^{2+} -bound gelsolin (PDB id: 1NPH (53)) and Ca^{2+} /actin:gelsolin



FIGURE 5 Stereo views showing the superposition of the Ca^{2+} -free EgAFFP homology model with the envelope obtained from the 10 DR models (spheres). Only DR models retrieved from the Ca^{2+} -free EgAFFP scattering curves are shown since these are virtually identical to the DR models based on Ca^{2+} -bound EgAFFP scattering curves. Middle and bottom images are rotated by 90° around the y and x axes, respectively, compared to the top image.

complex (PDB id: 1H1V (54)). Actin structure was removed from the later model. Both G4–G6 gelsolin structures agree with the experimental x-ray scattering curves significantly better than the EgAFFP model based on Ca^{2+} -bound G1–G3 domains of gelsolin (Fig. 3 B; $\chi = 1.77$ for 1NPH and $\chi = 1.78$ for 1H1V crystallographic models, respectively) but not as well as the EgAFFP model based on the Ca^{2+} -free G1–G3 gelsolin structure. Distance distribution functions for G4–G6 gelsolin domains also show smaller deviations from the experimentally derived curve as compared to the EgAFFP model based on Ca^{2+} -bound G1–G3 gelsolin domains but

not when compared with the model based on the Ca^{2+} -free G1–G3 gelsolin structure (Fig. 4).

DISCUSSION

We have previously cloned and expressed a functional three domain EgAFFP and demonstrated that this protein has two actin-binding sites and is regulated by calcium (30). To understand how this protein functions and is regulated by calcium, we started its structural characterization. Our attempts to solve its structure by x-ray crystallography have not been successful so far due to problems with protein crystallization. Therefore, we have used the alternative approaches of SAXS studies and molecular homology modeling to obtain structural information on EgAFFP.

Overall SAXS results (R_g , D_{max} , and DR models) indicate that EgAFFP in solution both in the presence and in the absence of Ca^{2+} is an asymmetric elongated protein. This is in agreement with the homology model based on the Ca^{2+} -free gelsolin structure (PDB id 1D0N) in which domains E1 and E3 are gathered at one side of the protein, forming an intimate contact through a continuous β -sheet, whereas E2 is relatively apart from them, on the other side. Small differences in R_g and D_{max} between the SAXS measurements and the homology model are probably due to subtle variations in position and orientation of the three distinct domains and their connecting loops in the model.

Comparison of x-ray data obtained in the absence and in the presence of Ca^{2+} indicates that Ca^{2+} does not significantly change x-ray scattering curves, rendering practically identical radii of gyration, maximum dimensions, and overall shapes of the protein in both Ca^{2+} -free and Ca^{2+} -bound states. This is surprising, since the homologous six domain protein, gelsolin, is known to suffer drastic domain shifts upon interaction with Ca^{2+} alone (53–56) or with Ca^{2+} and actin (38). As a consequence, EgAFFP homology models based on these two different structural templates (Ca^{2+} -free gelsolin and Ca^{2+} /gelsolin:actin complex, domains G1–G3) have very different molecular shapes (Fig. 2). However, only one of them (the Ca^{2+} -free gelsolin-based model) is in agreement with the experimental SAXS data measured from the EgAFFP samples both in the absence and in the presence of Ca^{2+} , suggesting that addition of Ca^{2+} alone is not sufficient to promote conformational changes in EgAFFP similar to those observed in the Ca^{2+} -bound N-terminal half of gelsolin: actin complex.

EgAFFP is more similar to the N-terminal half (36% of identical residues and 58% of conserved residues) than to the C-terminal half of gelsolin (29% of identical residues and 43% of conserved residues). Therefore, the G1–G3 gelsolin structure (in its Ca^{2+} -free form) was used for homology modeling. The predicted x-ray data for G4–G6 gelsolin structures also agree with the experimental x-ray scattering curves but less consistently than the model based on the G1–G3 Ca^{2+} -free gelsolin structure.

In EgAFFP, at least one actin-binding site is known to be Ca^{2+} dependent (30), but apparently the mere binding of Ca^{2+} does not determine structural changes large enough to be detected by SAXS. It is possible that more significant domain shifts would depend upon actin binding to EgAFFP, characterizing a different behavior from that of gelsolin, in which calcium activation triggers large domain movements both in the N- and C-terminal halves (53–59). This issue remains to be experimentally addressed for EgAFFP.

In gelsolin, Ca^{2+} activation is mediated by type 1 and type 2 Ca^{2+} -binding sites (54). Type 1 Ca^{2+} -binding sites are shared between actin and gelsolin domains, allowing the prediction that those domains will bind G-actin in a Ca^{2+} -sensitive manner. Type 2 Ca^{2+} -binding sites are completely contained within gelsolin domains, facilitating disruption of interactions between domains to release latches and render the activated gelsolin able to bind actin (59). Results obtained through sequence alignment and the molecular homology model (data not shown) indicate that EgAFFP has one type 1 Ca^{2+} -binding site, located on the first domain (E1), and three type 2 Ca^{2+} -binding sites, one on each domain. Considering this is the same organization found in each half of gelsolin, the different EgAFFP behavior upon Ca^{2+} binding cannot be simply attributed to the nature of Ca^{2+} -binding sites.

It has been shown that an acidic pH could induce conformational changes in gelsolin similar to those predicted to occur upon Ca^{2+} binding (60). Therefore, we cannot discard, at least formally, the hypothesis that the conformational differences between Ca^{2+} -bound EgAFFP and Ca^{2+} -bound gelsolin are due, at least to some extent, to differences in pH conditions in which EgAFFP SAXS measurements (this work) and gelsolin crystallization (38) were performed. It would be possible that gelsolin crystallization at an acidic pH (4.7) could have induced the large conformational change observed in its N-terminal (G1–G3) domains.

In conclusion, based on the low resolution ab initio structures and the Ca^{2+} -free homology model, we can infer that EgAFFP is a compact protein very similar to the N-terminal half of Ca^{2+} -free gelsolin. In contrast to what has been described for six domain members of the gelsolin family, the addition of Ca^{2+} does not result in major conformational changes in EgAFFP, suggesting a distinct functioning mechanism for three domain members of the gelsolin family. Subtle structural changes of regulatory relevance upon Ca^{2+} binding are likely to occur, but their study will possibly require EgAFFP crystallization. It also remains to be clarified how actin-binding would affect EgAFFP conformation and how EgAFFP binds to and severs F-actin.

We thank Dr. Elizabeth Cortez-Herrera for the gift of purified gelsolin.

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul, and the Network for Research and Training in Parasitic Diseases at the Southern Cone of Latin America/Swedish International

Development Agency. E.D.G. was a recipient of a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico. R.V.P. was a recipient of a fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. M.O.N. and N.H.M. were recipients of fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo.

REFERENCES

1. Thompson, R. C. A., and D. P. McManus. 2001. Aetiology: parasites and life-cycles. In WHO/OIE Manual on Echinococcosis in Humans and Animals: A Public Health Problem of Global Concern. J. Eckert, M. A. Hemmell, F.-X. Meslin, and Z. S. Pawlowski, editors. World Organization for Animal Health, Paris. 1–19.
2. Howell, M. J., and J. D. Smyth. 1995. Maintenance and cultivation of *Echinococcus* species in vivo and in vitro. In *Echinococcus* and Hydatid Disease. R. C. A. Thompson and A. J. Lymbery, editors. Allen & Unwin, London. 201–232.
3. Galindo, M., M. J. González, and N. Galanti. 2002. *Echinococcus granulosus* protoscolex formation in natural infections. *Biol. Res.* 35:365–371.
4. Martínez, C., R. Paredes, R. P. Stock, A. Saralegui, M. Andreu, C. Cabezon, R. Ehrlich, and N. Galanti. 2005. Cellular organization and appearance of differentiated structures in developing stages of the parasitic plathelminth *Echinococcus granulosus*. *J. Cell. Biochem.* 94:327–335.
5. Pollard, T. D., L. Blanchoin, and R. D. Mullins. 2000. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Biomol. Struct.* 29:545–576.
6. Pollard, T. D., and J. A. Cooper. 1986. Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. *Annu. Rev. Biochem.* 55:987–1035.
7. Stossel, T. P. 1993. On the crawling of animal cells. *Science.* 260:1086–1094.
8. Welch, M. D., D. A. Holtzman, and D. G. Drubin. 1994. The yeast actin cytoskeleton. *Curr. Opin. Cell Biol.* 6:110–119.
9. Schafer, D. A., and J. A. Cooper. 1995. Control of actin assembly at filament ends. *Annu. Rev. Cell Dev. Biol.* 11:497–518.
10. Sun, H. Q., K. Kwiatkowska, and H. L. Yin. 1995. Actin monomer binding proteins. *Curr. Opin. Cell Biol.* 7:102–110.
11. Yin, H. L. 1987. Gelsolin: calcium- and polyphosphoinositide-regulated actin-modulating protein. *Bioessays.* 7:176–179.
12. André, E., F. Lottspeich, M. Schleicher, and A. Noegel. 1988. Severin, gelsolin, and villin share a homologous sequence in regions presumed to contain F-actin severing domains. *J. Biol. Chem.* 263:722–727.
13. Eichinger, L., A. A. Noegel, and M. Schleicher. 1991. Domain structure in actin-binding proteins: expression and functional characterization of truncated severin. *J. Cell Biol.* 112:665–676.
14. Eichinger, L., and M. Schleicher. 1992. Characterization of actin- and lipid-binding domains in severin, a Ca^{2+} -dependent F-actin fragmenting protein. *Biochemistry.* 31:4779–4787.
15. Ampe, C., and J. Vandekerckhove. 1987. The F-actin capping proteins of *Physarum polycephalum*: cap42(a) is very similar, if not identical, to fragmin and is structurally and functionally very homologous to gelsolin; cap42(b) is *Physarum* actin. *EMBO J.* 6:4149–4157.
16. Gettemans, J., Y. De Ville, J. Vandekerckhove, and E. Waelkens. 1993. Purification and partial amino acid sequence of the actin-fragmin kinase from *Physarum polycephalum*. *Eur. J. Biochem.* 214:111–119.
17. Way, M., and A. G. Weeds. 1988. Nucleotide sequence of pig plasma gelsolin. Comparison of protein sequence with human gelsolin and other actin severing proteins shows strong homologies and evidence for large internal repeats. *J. Mol. Biol.* 203:1127–1133.
18. Cunningham, C. C., T. P. Stossel, and D. J. Kwiatkowski. 1991. Enhanced motility in NIH 3T3 fibroblasts that overexpress gelsolin. *Science.* 251:1233–1236.

19. Janmey, P. A., and P. T. Matsudaira. 1988. Functional comparison of villin and gelsolin: effect of Ca^{++} , KCl and polyphosphoinositides. *J. Biol. Chem.* 263:16738–16743.
20. Friederich, F., K. Vancompemolle, C. Huet, M. Goethals, J. Finidori, J. Vandekerckhove, and D. Louvard. 1992. An actin binding site containing a conserved motif of charged amino acid residues is essential for morphogenic effect of villin. *Cell.* 70:81–92.
21. Kwiatkowski, D. J., T. P. Stossel, S. H. Orkin, J. E. Mole, H. Colten, and H. L. Yin. 1986. Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin binding domain. *Nature.* 323:455–458.
22. Janmey, P. A. 1994. Phosphoinositides and calcium as regulators of cellular actin assembly and disassembly. *Annu. Rev. Physiol.* 56: 169–191.
23. Way, M., J. Gooch, B. Pope, and A. G. Weeds. 1989. Expression of human plasma gelsolin in *Escherichia coli* and dissection of actin binding sites by segmental deletion mutagenesis. *J. Cell Biol.* 109: 593–605.
24. Pope, B., M. Way, and A. G. Weeds. 1991. Two of the three actin-binding domains of gelsolin bind to the same subdomain of actin. *FEBS Lett.* 280:70–74.
25. Burtnick, L. D., E. K. Koepf, J. M. Grimes, E. Y. Jones, D. I. Stuart, P. J. McLaughlin, and R. C. Robinson. 1997. The crystal structure of plasma gelsolin: implications for actin severing, capping and nucleation. *Cell.* 90:661–670.
26. Markus, M. A., P. Matsudaira, and G. Wagner. 1997. Refined structure of villin-14T and a detailed comparison with other actin-severing proteins. *Protein Sci.* 6:1197–1209.
27. Schnuchel, A., R. Wiltschek, L. Eichinger, M. Schleicher, and T. A. Holak. 1995. Structure of severin domain 2 in solution. *J. Mol. Biol.* 247:21–27.
28. Puius, Y. A., E. V. Fedorov, L. Eichinger, M. Schleicher, and S. C. Almo. 2000. Mapping the functional surface of domain 2 in the gelsolin superfamily. *Biochemistry.* 39:5322–5331.
29. McLaughlin, P. J., J. T. Gooch, H.-G. Mannherz, and A. G. Weeds. 1993. Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. *Nature.* 364:685–692.
30. Cortez-Herrera, E., R. R. Yamamoto, J. J. S. Rodrigues, S. E. Farias, H. B. Ferreira, and A. Zaha. 2001. *Echinococcus granulosus*: cloning and functional in vitro characterization of an actin filament fragmenting protein. *Exp. Parasitol.* 97:215–225.
31. Virginio, V. G., A. Hernández, M. B. Rott, K. M. Monteiro, A. F. Zandonai, A. Nieto, A. Zaha, and H. B. Ferreira. 2003. A set of recombinant antigens from *Echinococcus granulosus* with potential for use in the immunodiagnosis of human cystic hydatid disease. *Clin. Exp. Immunol.* 132:309–315.
32. Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali. 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29:291–325.
33. Baker, D., and A. Sali. 2001. Protein structure prediction and structural genomics. *Science.* 294:93–96.
34. Trewthella, J. 1997. Insights into biomolecular function from small-angle scattering. *Curr. Opin. Struct. Biol.* 7:702–708.
35. Svergun, D. I., and M. H. J. Koch. 2003. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 66: 1735–1782.
36. Casadio, R., E. Polverini, P. Mariani, F. Spinazzi, F. Carsughi, A. Fontana, P. Polverino de Lauro, G. Matteucci, and C. M. Bergamini. 1999. The structural basis for the regulation of tissue transglutaminase by calcium ion. *Eur. J. Biochem.* 262:672–679.
37. Occhipinti, E., P. L. Martelli, F. Spinazzi, F. Corsi, C. Formantici, L. Molteni, H. Amenitsch, P. Mariani, P. Tortora, and R. Casadio. 2003. 3D structure of *Sulfolobus solfataricus* carboxypeptidase developed by molecular modeling is confirmed by site-directed mutagenesis and small angle x-ray scattering. *Biophys. J.* 85:1165–1175.
38. Burtnick, L. D., D. Urosev, E. Irobi, K. Narayan, and R. C. Robinson. 2004. Structure of the N-terminal half of gelsolin bound to actin: roles in severing, apoptosis and FAF. *EMBO J.* 23:2713–2722.
39. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *E. coli* as fusions with glutathione S-transferase. *Gene.* 67:31–40.
40. Laemmli, U. K. 1970. Cleavage of structured proteins during assembly of head of bacteriophage T4. *Nature.* 227:680–685.
41. Cooper, J. A., S. B. Walker, and T. D. Pollard. 1983. Pyrene actin: documentation of the validity of a sensitive assay for actin polymerization. *J. Muscle Res. Cell Motil.* 4:253–262.
42. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
43. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.
44. Šali, A., and T. L. Blundell. 1993. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.
45. Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.
46. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.
47. Petoukhov, M. V., N. A. J. Eady, K. A. Brown, and D. I. Svergun. 2002. Addition of missing loops and domains to protein models by x-ray solution scattering. *Biophys. J.* 83:3113–3125.
48. Svergun, D. I., C. Barberato, and M. H. Koch. 1995. CRY SOL: a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773.
49. Kozin, M. B., and D. I. Svergun. 2001. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* 34:33–41.
50. Guinier, A., and G. Fournet. 1955. *Small-Angle Scattering of X-Rays*. John Wiley and Sons, New York.
51. Glatter, O., and O. Kratky. 1982. *Small Angle X-Ray Scattering*. Academic Press, London, New York.
52. Svergun, D. I., M. V. Petoukhov, and M. H. J. Koch. 2001. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* 80:2946–2953.
53. Kolappan, S., J. T. Gooch, A. G. Weeds, and P. J. McLaughlin. 2003. Gelsolin domains 4–6 in active, actin-free conformation identifies sites of regulatory calcium ions. *J. Mol. Biol.* 329:85–92.
54. Choe, H., L. D. Burtnick, M. Mejillano, H. L. Yin, R. C. Robinson, and S. Choe. 2002. The calcium activation of gelsolin: insights from the 3 Å structure of the G4–G6/actin complex. *J. Mol. Biol.* 324:691–702.
55. Pope, B. J., J. T. Gooch, and A. G. Weeds. 1997. Probing the effects of calcium on gelsolin. *Biochemistry.* 36:15848–15855.
56. Robinson, R. C., M. Mejillano, V. P. Le, L. D. Burtnick, H. L. Yin, and S. Choe. 1999. Domain movement in gelsolin: a calcium-activated switch. *Science.* 286:1939–1942.
57. Gremm, D., and A. Wegner. 2000. Gelsolin as a calcium-regulated actin filament-capping protein. *Eur. J. Biochem.* 267:4339–4345.
58. Lin, K.-M., M. Mejillano, and H. L. Yin. 2000. Ca^{2+} regulation of gelsolin by its C-terminal tail. *J. Biol. Chem.* 275:27746–27752.
59. Narayan, K., S. Chumnarnsilpa, H. Choe, E. Irobi, D. Urosev, U. Lindberg, C. E. Schutt, L. D. Burtnick, and R. C. Robinson. 2003. Activation in isolation: exposure of the actin-binding site in the C-terminal half of gelsolin does not require actin. *FEBS Lett.* 552:82–85.
60. Lamb, J. A., P. G. Allen, B. Y. Tuan, and P. A. Janmey. 1993. Modulation of gelsolin function. Activation at low pH overrides Ca^{2+} requirement. *J. Biol. Chem.* 268:8999–9004.

Expression, purification, and characterization of rat protein tyrosine phosphatase η catalytic domain

Maria A.M. Santos^a, Silvia M. Santos^a, Huita C. Matozo^a, Rodrigo V. Portugal^a,
Rodolfo Iuliano^b, Alfredo Fusco^c, Igor Polikarpov^{a,*}

^a Instituto de Física de São Carlos, Departamento de Física e Informática, Universidade de São Paulo, Avenida Trabalhador São Carlense, 400, CEP 13566-590 São Carlos, SP, Brazil

^b Dipartimento di Medicina Sperimentale e Clinica, Facoltà di Medicina e Chirurgia, Università di Catanzaro, 88100 Catanzaro, Italy

^c Centro di Endocrinologia ed Oncologia Sperimentale del Consiglio Nazionale delle Ricerche clo Dipartimento di Biologia e Patologia Cellulare e Molecolare, Facoltà di Medicina e Chirurgia, Università di Napoli "Federico II," 80131 Napoli, Italy

Received 3 November 2004, and in revised form 23 January 2005

Abstract

Receptor-like protein tyrosine phosphatases generally contain one or two conserved intracellular catalytic domains with a conserved sequence motif ([I/V]HCXAGXXR[S/T]G), a single transmembrane domain, and an external highly variable part. Here, we describe cloning of the intracellular catalytic domain of the rat protein tyrosine phosphatase η (rPTP η CD) into pET28a(+) vector, its expression in *Escherichia coli*, purification and initial characterization. The purification of His₆-tagged rPTP η CD to near homogeneity was achieved by a combination of affinity and size exclusion chromatography. The His-tag was subsequently removed by thrombin digestion. PhastGel IEF electrophoresis demonstrated that the isoelectric point of this 41 kDa His₆-tag free recombinant protein was 7.3, which is just slightly higher than the theoretically predicted value of 7.2. To assess the functionality of the rPTP η CD we used the pNPP hydrolysis assay and observed that the enzyme has a specific activity of 9 nmol/min/ μ g. The secondary structure and stability of the recombinant protein was also analyzed by circular dichroism and fluorescence spectroscopy. In summary, the rPTP η CD is stable at 18 °C, properly folded, and fully active, which makes it a suitable candidate for structural and functional studies.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Rat protein tyrosine phosphatases η ; Intracellular catalytic domain; *E. coli* expression; pNPP hydrolysis assay; Fluorescence; Circular dichroism

The balance between protein tyrosine kinases (PTKs) and protein tyrosine phosphatases (PTPs) activities regulates the intracellular levels of protein tyrosine phosphorylation. The PTPs dephosphorylation of tyrosyl residues plays, in vivo, a major role in controlling cell activities, proliferation, differentiation, and neoplastic transformation [1]. The PTPs represent a diverse family of enzymes that exist in both soluble cytosolic and receptor-like tyrosine phosphatase (RPTP) forms. Gen-

erally, the RPTPs contain one or two conserved intracellular catalytic domains of approximately 240 amino acids with a conserved motif ([I/V]HCXAGXXR[S/T]G), a single transmembrane domain, and an external segment that is highly variable. This ectodomain is believed to harbor binding sites for putative physiological ligands, although the majority of them remains unknown [2]. Based upon their extracellular domains, the receptor-like protein tyrosine phosphatases have been subdivided into eight different classes constituted of various combining functional motifs such as: the immunoglobulin (Ig)-like repeat, the fibronectin type III (FnIII)

* Corresponding author. Fax: +55 16 33739881.

E-mail address: ipolikarpov@ifsc.usp.br (I. Polikarpov).

repeat, the carbonic anhydrase-like domain (CA) [3], and the *Xenopus* A5-like antigen (“MAM” domain) [4].

The RPTPs involved in intercellular communication such as LAR-PTP and PTP δ , bear in their extracellular region the immunoglobulin and the fibronectin III domains [5,6]. As the neural cell adhesion molecules (N-CAM, L1, TAG-1, Ng-CAM, Nr-CAM, and F3/F11) [7], these receptor-like tyrosine phosphatases have been implicated in the processes of nervous system differentiation, development, and axonal guidance. The PTP μ , PTP κ , and PTP λ belong to a subset of adhesion-like RPTPs which carry a third extracellular domain, the MAM domain. These proteins intermediate cells aggregation, in neurons, via homophilic interactions between cell–cell or cell–matrix [8–10]. In the non-neural cells, the rat protein tyrosine phosphatase η (rPTP η) gene is engaged in the process of thyroid cells differentiation and neoplastic cells suppression [11]. The rPTP η gene is highly homologous to human DEP-1/HPTP η and mouse protein phosphatase η , Ptp η [12–15]. The protein structure of rPTP η can be subdivided in only one intracellular catalytic phosphatase domain, a single transmembrane domain, and eight fibronectin type III-like repeats in the extracellular region [16–18]. Similar to thyroid specific genes whose expression is induced by TSH, rPTP η is positively regulated by thyrotropin through the protein kinase A pathway and negatively regulated by protein kinase C activation.

Differently from PTKs whose contribution to tumorigenesis is a subject of intense research, the potential role of PTPs in malignant cell growth is yet poorly explored. Studies involving LAR-PTP and human breast carcinoma cells showed that overexpression of *LAR-PTP* gene is able to slow down the cell proliferation rate [19]. Likewise, the transfection of PTP α in breast carcinoma cells also leads to growth inhibition, suggesting that specific protein tyrosine phosphatases are playing a suppressive regulatory role in malignant cells [20]. Another evidence of the involvement of PTPs in neoplastic transformations is observed with rPTP η . A reduction in the protein expression occurs in all oncogene-transformed thyroid cell lines and is absent in highly malignant thyroid cells [21]. However, it has been shown that inhibition of the tumorigenic process is achieved when *rPTP η* gene expression is reestablished. The mechanism involved in this process includes increasing levels of the cell-cycle inhibitor p27^{kip1} protein and dephosphorylation of PLC γ 1, a substrate of DEP-1/HPTP η [22]. More recent studies have also implicated the mouse homologue of *rPTP η* , *Ptp η* , in susceptibility to mouse colon cancer and missense mutations of this gene have been commonly observed in human colon, lung, and breast cancers [23].

Given the importance of the rat tyrosine phosphatases η and the need to further study its molecular mechanism of signal transduction, we aimed in this work to

establish an expression and purification system capable of providing a large-scale quantity of the functional intracellular catalytic domain of the rPTP η (rPTP η CD). For this purpose, the phosphatase domain of *rPTP η* gene was cloned into the expression vector pET-28a(+). The His₆-tagged rPTP η CD was expressed at high levels in BL21 (DE3) and the purified protein was obtained by a combination of metal chelate affinity and size exclusion chromatography. The His-tag was removed by thrombin cleavage. To confirm the native conformation of the pure His₆-tag free rPTP η CD, the *p*-nitrophenyl phosphate (pNPP) hydrolysis assay was performed, in addition to fluorescence and circular dichroism spectroscopy studies.

Materials and methods

Cloning of the 3' end of the rPTP η gene

The expression vector pGEX-2TK (Promega) containing the coding sequence of the intracellular region (1026 bp) of rat tyrosine phosphatase η fused in frame to glutathione-S-transferase (GST) gene was obtained from Dr. Alfredo Fusco, Università di Napoli “Federico II,” Napoli, Italy. To isolate the rPTP η insert for cloning into pET-28a(+) vector (Novagen), we digested the plasmids with *Bam*HI and *Eco*RI. The digested products were purified by electrophoresis in agarose gel using Perfectprep Gel Clean-up kit (Eppendorf). The ligation was performed overnight to *Bam*HI–*Eco*RI digested pET-28a(+) in the presence of T4-DNA ligase (New England Biolabs). The ligated samples containing a nucleotide sequence coding for six histidines, placed upstream and in frame with the rPTP η CD insert, were transformed into chemically competent BL21 (DE3) cells. A single colony obtained on Luria–Bertani (LB)—kanamycin plates was then selected and grown for sequencing using ABI PRISM 377 DNA Sequencer (AB Applied Biosystems) and for preparation of glycerol freezer stocks.

Expression of His₆-tagged rPTP η catalytic phosphatases domain

BL21 (DE3) cells harboring the plasmid containing the rPTP η CD insert were grown overnight in 100 ml LB medium supplemented with 30 μ g/ml kanamycin. After growth at 37 °C, the pre-culture was used to inoculate 500 ml of 2 \times YT media plus kanamycin in a flask and the cells were grown at 30 °C under continuous shaking, until the absorbance at 600 nm reached 0.6–0.8. At this point, 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG) was added to induce rPTP η CD expression and cells were incubated at 30 °C for 1–6 h. The induced bacteria were harvested by centrifugation at 6000g in a Sorvall RC-5C Plus centrifuge at 4 °C for

20 min. Then, cell pellets were frozen in liquid nitrogen and stored at -80°C .

Purification of recombinant rPTP η CD and thrombin treatment

Prior to purification, bacterial pellets from 5 L of culture were thawed and resuspended in 100 ml of lysis buffer (50 mM sodium phosphate buffer at pH 7.8; 100 mM NaCl; 10% glycerol; 10 mM imidazol; 2 mM β -mercaptoethanol) containing 1 mM PMSF and 0.5 mg/ml of lysozyme (Sigma). The suspension was incubated on ice for 30 min to lyse cells. The lysate was further disrupted by sonication at approximately 60 W for 10×30 s pulses (with 1 min delay between pulses) on ice, with a 550 Sonic Dismembrator (Fisher Scientific) to reduce viscosity. Centrifugation was done at 14,000g for 1 h to remove the insoluble cell debris and the clear supernatant obtained constituted the crude protein preparation. The supernatant from the above step was mixed with 20 ml Talon Superflow resin (Clontech), pre-equilibrated with equilibration buffer (50 mM sodium phosphate buffer at pH 7.8; 300 mM NaCl; 10% glycerol; 10 mM imidazol; 2 mM β -mercaptoethanol), and left rotating at 4°C for 1 h. The mixture of resin and supernatant was poured into a c16/10 glass column (Amersham Biosciences) and the flow-through of the soluble fraction was collected. The column was connected to a HPLC ÄkTA purifier (Amersham Biosciences) and washed with equilibration buffer until the absorbance reading reached the baseline. The tightly bound proteins were eluted with elution buffer (50 mM sodium phosphate buffer, pH 7.8; 50 mM NaCl; 10% glycerol; 300 mM imidazol; 2 mM β -mercaptoethanol) and the protein peak was collected in fractions of 1 ml each. At this stage, the rPTP η CD was 70% pure as judged on the basis of 15% SDS–PAGE. It was further purified to >96% by size exclusion chromatography on a Superdex 200 HL 26/60 column (Amersham Biosciences) using as eluent Hepes buffer (20 mM Hepes, pH 7.8; 200 mM NaCl; 5% glycerol; and 1 mM DTT). All purification procedures were carried out at 4°C . The purified protein fractions were visualized on 15% SDS–PAGE. Soluble His $_6$ -rPTP η CD (molecular weight of 43 kDa) was concentrated to 1 mg/ml and incubated, according to manufacturer's recommendation, with 0.5 U/ml bovine thrombin protease for 1–18 h at 18°C followed by dialysis against Hepes buffer. The thrombin cleaved rPTP η CD was, then, frozen in liquid nitrogen and stored at -80°C .

Enzymatic activity

The phosphatases activity of purified rPTP η CD was tested in vitro by pNPP (Sigma) hydrolysis assay. The assay was performed mixing in each microtiter well 5 μl of BSA (Sigma) solution (5 mg/ml) with different

concentrations of purified enzyme. The final volume was adjusted to 80 μl with pNPP buffer (25 mM Hepes, pH 7.2; 50 mM NaCl; 5 mM dithiothreitol; and 2.5 mM EDTA). Each concentration of protein was tested in triplicate. The blanks were prepared the same as samples minus the enzyme. To initiate the reaction, 120 μl pNPP (1.5 mg/ml in pNPP buffer) was added as substrate solution and the material was incubated for 5–60 min at room temperature. The reaction was stopped by the addition of 20 μl of 13% K_2HPO_4 solution and the assay was read against blank at 405 nm in a Spectra max Plus³⁸⁴ (Molecular Devices). To calculate the specific activity of rPTP η we used Eq. (1), where ϵ is the extinction coefficient for pNPP ($1.78 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$). The results were expressed in nmoles/min/ μg by multiplying the outcome from Eq. (1) by a factor of 10^9 .

$$\begin{aligned} \text{Specific activity} = & [(\text{sample vol. in liters}) \\ & \times (A_{405})]/[(\epsilon \text{ in } \text{M}^{-1} \text{ cm}^{-1}) \\ & \times (\text{pathlength in cm}) \\ & \times (\text{time in min}) \\ & \times (\mu\text{g of enzyme})]. \end{aligned} \quad (1)$$

Protein concentration measurements and SDS–PAGE

Protein concentration was determined using the Bio-Rad protein assay according to manufacturer's instructions (Bio-Rad Laboratories). The assay is based on the method of Bradford [24] and uses bovine serum albumin (Sigma) as reference standard. The SDS–PAGE of protein samples was carried out in Laemmli's buffer [25] with 12% polyacrylamide. Electrophoresis was performed in a Gibco-BRL Vertical Cell Electrophoresis Apparatus (Life Technologies) together with standard molecular mass markers (Sigma). For protein visualization, the gels were stained with Coomassie brilliant blue R-250 (Bio-Rad Laboratories) and destained with 10% acetic acid.

Western blot assays

The nitro-cellulose membrane, on which the purified recombinant His $_6$ -rPTP η CD protein band had been transferred, was blocked at room temperature for 1:30 h with 5% non-fat dried milk in Tris-buffered saline (TBS, pH 8.0) and then incubated for 1 h with the Ni-NTA AP conjugate (Qiagen). After washes in TBS containing 0.05% Tween 20 (TBS-T), the membrane was incubated with AP conjugate substrate (kit Bio-Rad Laboratories) to visualize the reaction.

Isoelectric focusing (IEF)

Isoelectric focusing was accomplished in a PhastSystem (Amersham Biosciences). The method consisted of

three steps: In the first step the pH gradient is formed on the PhastGel IEF media 3–9 (pH range 3–9) containing pharmalyte carrier ampholytes; at the second step, the samples are applied to the gel for 15 V h at the voltage of 200 V; and finally, at the third step, the proteins migrate to their isoelectric points. Separations took approximately 30 min after which the gel was stained with 0.1% PhastGel Blue R (Amersham Biosciences) solution in 30% methanol and 10% acetic acid. To visualize the bands the gel was destained in a solution containing 30% methanol and 10% acetic acid until the background was clear.

Circular dichroism spectroscopy

Purified rPTP η CD, with His-tag removed by thrombin treatment was dialyzed against 5 mM phosphate buffer, pH 7.8; 12.5 mM NaCl; 0.25 mM DTT. CD spectra were measured in a quartz 1 mm path length cuvette at 18 °C and protein concentration of 0.08 mg/ml, using a Jasco J-720 spectropolarimeter (Jasco). The protein far-UV spectra were recorded over a wavelength range of 190–250 nm by signal averaging of 64 spectra. The instrument was set at scanning speed of 100 nm/min and step resolution of 1 nm. The secondary structure content of the protein was monitored following the 208 nm peak during thermal denaturation (10–80 °C) induced by a PFD-425S Peltier temperature control system with a slope of temperature increase of 1 °C/min. CD measurements of thermal denaturation were performed in quartz 1 mm path length cuvettes at protein concentration of 0.08 mg/ml in the same buffer as described above. For standardization, the baseline spectrum (buffer alone) was subtracted from the spectra and the results were expressed as mean residue ellipticity $[\theta] = (\theta \cdot MRW)/(10 \cdot c \cdot d)$; where θ is observed ellipticity; MRW = protein mean weight/number of residues; d is an optical path length in cm; and c is protein concentration in mg/ml.

Fluorescence spectroscopy

The intrinsic fluorescence of His₆-tag free rPTP η CD was measured in a quartz 10 × 2 mm cuvette at 18 °C, using an K2 Multifrequency Phase Fluorometer (ISS) equipped with a temperature control water bath RTE-111 (Thermo-Neslab). Emission spectra in the range of 300–450 nm were recorded using an excitation wavelength of 280 nm. The fluorescence emission spectra of rPTP η CD at a concentration of 0.3 mg/ml were monitored at a scanning interval of 1 nm and integration time of 1 s. After each measurement the spectrum was corrected for buffer (100 mM Hepes, pH 7.8; 50 mM NaCl, and 1 mM DTT) contribution.

Results and discussion

Cloning and expression of the catalytic phosphatase domain of rPTP η in Escherichia coli

The insert coding for the phosphatase domain of rPTP η which shows, respectively, 95 and 98% identity to the intracellular domains of human and mouse homologues (HPTP η /Ptp η), was excised with *Bam*HI and *Eco*RI from pGEX-2TK and ligated into pET-28a(+) vector. The ligation product was then transformed into BL21 (DE3) competent cells. To confirm the presence of pET-rPTP η recombinant plasmids, DNA obtained from different recombinant colonies was digested with both restriction enzymes and examined by electrophoresis in 1% agarose gel. Sequencing of the cloned vectors revealed the open reading frame (ORF) of the intracellular region of rPTP η (W875-A1216) plus the expected 34 additional amino acid residues derived from pET-28a(+) vector at its amino terminus (MGSSHHHHH HSSGLVPRGSHMASMTGGQQMGRGS), including the cluster of six histidine residues for protein purification by metal affinity chromatography.

After several attempts to determine the optimal conditions for rPTP η CD expression in *E. coli* BL21 (DE3), the His₆-rPTP η CD production was optimized at 30 °C with 0.5 mM of IPTG used for induction. Cells were grown until their optical density reached the levels of absorbance between 0.6 and 0.8 at 600 nm. To evaluate the time course for rPTP η CD expression, bacteria were harvested every hour for 6 h following induction and total cell lysates analyzed on SDS-PAGE. A band at the expected molecular mass (43 kDa) was detected and reached high level of expression after 3 h of induction. This band was absent in the control total lysate of *E. coli* BL21 (DE3)/pET-rPTP η CD (Fig. 1A). Since under these conditions the recombinant protein was well over-produced, we established an induction time of 3 h for the large-scale production of His₆-rPTP η CD.

Purification of His₆-tagged rPTP η CD

The over-expressed recombinant His₆-rPTP η CD was solubilized in lysis buffer before sonication to reduce the size of cell debris and nucleic acid. The clear lysate obtained after centrifugation constituted the crude protein solution and contained the His₆-rPTP η CD soluble fraction. His₆-rPTP η CD was successfully bound to and competitively eluted from the metal-chelating affinity column by application of high concentration of imidazole (300 mM) in the elution buffer. The nonspecific binding of nontagged proteins was reduced by including a low concentration of imidazole (10 mM) in the equilibration buffer as well as 300 mM NaCl to prevent undesirable ionic interactions. The purification yielded significant amount of protein with considerable enrich-

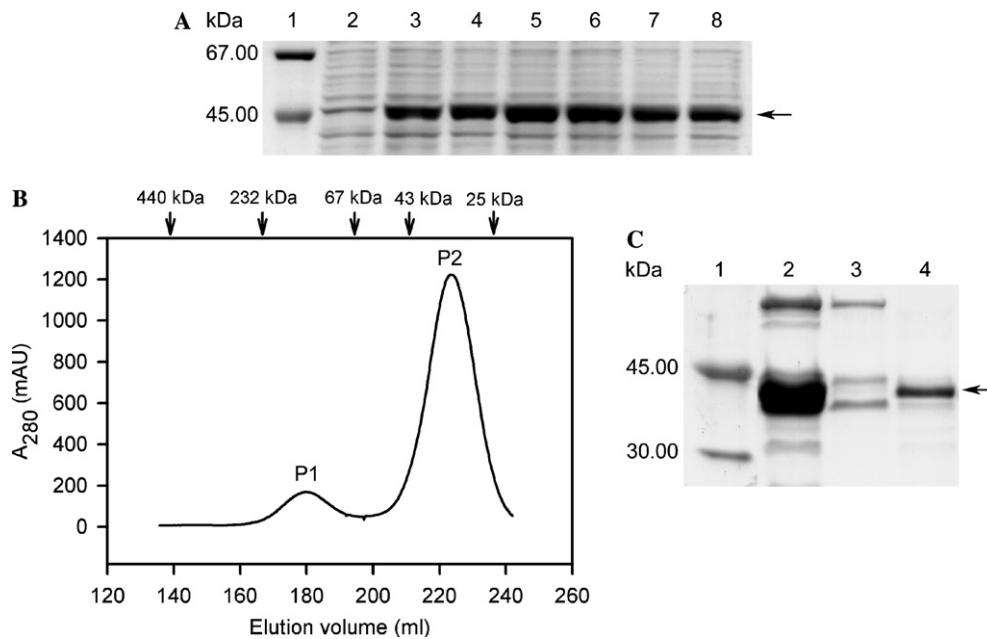


Fig. 1. His₆-rPTP η CD expression and purification. (A) SDS-PAGE was carried out on a Mini V8-10 Vertical Gel Electrophoresis Apparatus (Life Technologies-Gibco-BRL) and proteins detected by Coomassie brilliant blue staining. Lane 1, molecular weight standards: 67.00 kDa, bovine serum albumin; 45.00 kDa, egg albumin (Sigma); lane 2, total lysate of control BL21 (DE3) culture before induction with IPTG; lanes 3–8, total lysate of BL21 (DE3) culture 1, 2, 3, 4, 5, and 6 h, respectively, after IPTG addition. Black arrow indicates a band of approximately 43 kDa corresponding to His₆-rPTP η CD. (B) HPLC elution profile of gel filtration chromatography of His₆-rPTP η CD with two major protein peaks, P1 and P2, and Superdex 200 HL 26/60 column calibration with ferritin (440 kDa), catalase (232 kDa), albumin (67 kDa), ovalbumin (43 kDa), and chymotrypsinogen A (25 kDa) (Amersham Biosciences). (C) SDS-PAGE analysis of gel filtration chromatographic peaks. Lane 1, molecular weight standards: 45.00 kDa, albumin from egg; 30.00 kDa, carbonic anhydrase (Sigma); lane 2, pool of His₆-rPTP η CD after affinity chromatography injected on the Superdex 200 HL 26/60 column (Amersham Biosciences); lane 3, gel filtration peak P1, consisting of only protein contaminants; and lane 4, peak P2, containing the purified His₆-rPTP η CD of about 43 kDa (arrow).

ment of the His₆-rPTP η CD. Peak protein fractions from the chromatographic separation were resolved by SDS-PAGE and showed approximately 70% purity (Fig. 1C). To remove the rest of the contaminants, the affinity purified His₆-rPTP η CD was submitted to gel filtration chromatography. In order to determine the molecular size of the eluted protein, the Superdex 200 HL 26/60 column was previously calibrated with standard proteins (Amersham Biosciences) and the calibration curve established by relating their elution volume data to a molecular size parameter, Stokes' radius (Rs). His₆-rPTP η CD and the ovalbumin marker have the same molecular mass of 43 kDa and elute from the column with Rs of, approximately, 2.97 and 3.05 nm, respec-

tively, suggesting that His₆-rPTP η CD elutes from the gel filtration column as a monomer. SDS-PAGE analysis of the purified protein indicates that the preparation was more than 96% pure (Figs. 1B and C). The efficiencies of affinity and gel filtration purifications are shown in Table 1. The total protein yield at the last purification step was around 4 mg/L of cell culture.

Thrombin treatment of purified His₆-rPTP η CD

The amino acid sequence LVPRGS is the target for thrombin and the cleavage site is the peptide bond between the amino acids R and G. Thus, digestion of the recombinant protein by thrombin resulted in elimination of 17 amino acid residues, MGSSHHHHHSSGLVPR, at the amino terminus of His₆-rPTP η CD, including the His₆-Tag sequence that was not encoded by the *Bam*HI-*Eco*RI fragment (W875–A1216). Complete thrombin digestion occurred after the first hour of incubation at 18 °C. A single protein band with a molecular mass of approximately 41 kDa can be visualized at the SDS-PAGE (Fig. 2A). The His₆-tag removal from the recombinant protein was also observed by PhastGel 8–25% native PAGE followed by Western Blot assay where His₆-rPTP η CD but not rPTP η CD was recognized by Ni-NTA AP conjugate (Fig. 2B).

Table 1
Purification protocol of rPTP η from 1 L culture

Fraction (%)	Yield (mg) ^a	Total protein
I	10.0 ^b	100
II	4.0 ^b	40
III	3.5 ^c	35

Note. I, His-rPTP η protein after imidazole elution; II, His-rPTP η protein after gel filtration; III, rPTP η protein after thrombin cleavage. ^aYield was determined using the Bio-Rad protein assay based on the method of Bradford, and corresponded to ^bHis-tagged rPTP η or ^cHis-tag free rPTP η .

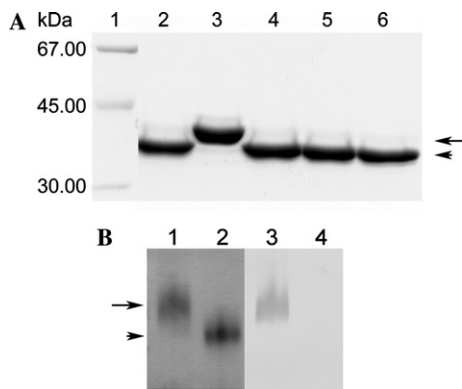


Fig. 2. Thrombin protease treatment of His₆-rPTP η CD. (A) SDS-polyacrylamide gel electrophoresis. Lane 1, molecular weight standard: 67.00 kDa, bovine serum albumin; 45.00 kDa, egg albumin; 30.00 kDa, carbonic anhydrase (Sigma); lanes 3, untreated His₆-rPTP η CD after 18 h incubation at 18 °C, showing a band of approximately 43 kDa (arrow); lanes 2 and 4–6, His₆-rPTP η CD after thrombin treatment for 1, 3, 6, and 18 h at 18 °C. The arrow head indicates a band of approximately 41 kDa corresponding to His₆-tag free rPTP η CD. (B) Lanes 1 and 2, PhastGel gradient 8–25% native PAGE of purified His₆-rPTP η CD untreated and treated with 0.5 U of thrombin/100 μ g of recombinant protein for 18 h at 18 °C, respectively; lanes 3 and 4, the same gel transferred onto nitro-cellulose membrane prior Coomassie blue staining. The Western blot was performed using Ni-NTA coupled to calf intestinal alkaline phosphatase to detect recombinant His₆-rPTP η CD (lane 3).

Isoelectric focusing and phosphatase activity of recombinant rPTP η CD

The isoelectric point and enzyme activity of recombinant rPTP η CD were measured as described in the Materials and methods. The *pI* was found to be 7.3 which was just slightly higher than the value predicted theoretically for thrombin treated rPTP η by the ProtParam tool (*pI* = 7.2) [26].

The biological activity of rPTP η CD was accessed by measuring the ability of the protein to hydrolyze *p*-nitrophenyl phosphate. This reaction releases a yellow colored compound, *p*-nitrophenol, which absorbs at 405 nm. The purified His₆-tag free rPTP η CD, subjected to pNPP hydrolysis assay, showed phosphatase activity for pNPP in a dose-dependent manner, whereas the control failed to exhibit phosphatase activity (Fig. 3). In the assay, pure rPTP η CD (0.12 and 0.25 μ M) showed a time-dependent increase in release of *p*-nitrophenol and phosphate up to 60 min. Based on these results, the mean rPTP η CD specific activity was calculated to be 9 nmol/min/ μ g suggesting that after purification and thrombin treatment rPTP η CD maintains its native structure and remains intact, with no signs of proteolytic degradation which could occur during the His₆-tag cleavage procedure. The experimentally determined rPTP η CD specific activity is in the same range as the previously reported activities for PTP μ [27], PTP α [28] and for commercial preparations of recombinant PTP-B1 [29] and SHP2 [30].

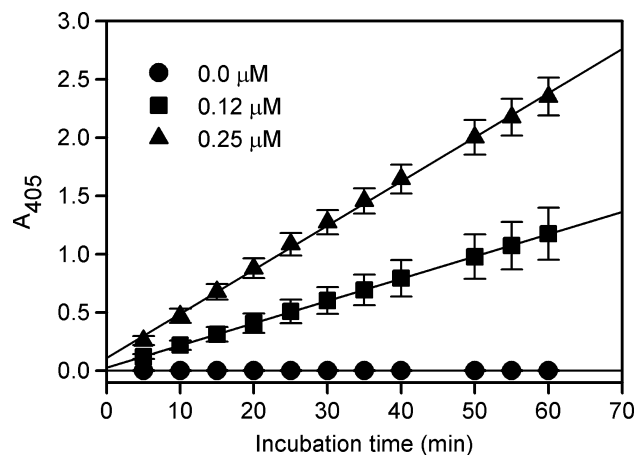


Fig. 3. rPTP η CD specific activity. The purified recombinant rPTP η CD was assayed in triplicates for phosphatase activity toward a substrate, pNPP, at two enzyme concentrations (0.12 and 0.25 μ M) and different incubation times (5–60 min). The activity (free phosphate release) is expressed as A_{405} . Samples containing no enzyme were used as negative control (filled circles).

Circular dichroism and fluorescence spectroscopy

The secondary structure of rPTP η CD was investigated by CD and fluorescence measurements. The far-UV CD spectrum at 18 °C displayed two negative bands at 222 and 208 nm and a positive band near 195 (Fig. 4A). Analysis of the spectrum was done using DichroWeb [31] with the programs SELCON3 [32], CONTIN [33], and K2D [34]. The amount of α -helix and β -strands content ranged from 25 to 37% and from 15 to 22%, respectively. It is in a good agreement with the secondary structure prediction using PSIPRED (Protein Structure Prediction Server) [35] and APSSP (Advanced Protein Secondary Structure Prediction Server) [36] which found, respectively, for α -helix content 31–36% and for β -strands content 14–16%. Furthermore, the CD results are in a good agreement with published reports on the crystal structures of the catalytic domains of different members of the PTP superfamily. For example, the structure of the membrane proximal domain (D1) of RPTP α determined by Bilwes et al. [37] has 29.80% of α -helix and 17.88% of β -strands. Similar content of α -helices (31.65%) and β -strands (20.14%) was found for the RPTP μ D1 domain [38].

The secondary structure thermal denaturation of rPTP η CD shows a clear transition from a stable conformation to a partially denatured state, occurring between 45 and 55 degC. The CD signal at 208 nm was monitored from 10 to 80 °C (Fig. 4B). The temperature effect on the rPTP η CD secondary structure started at approximately 45 °C and reached a plateau at about 55 °C. The sigmoidal profile of the thermal transition is an indication of cooperativity as well as a native-like packing of the protein. The CD spectrum, monitored before and after

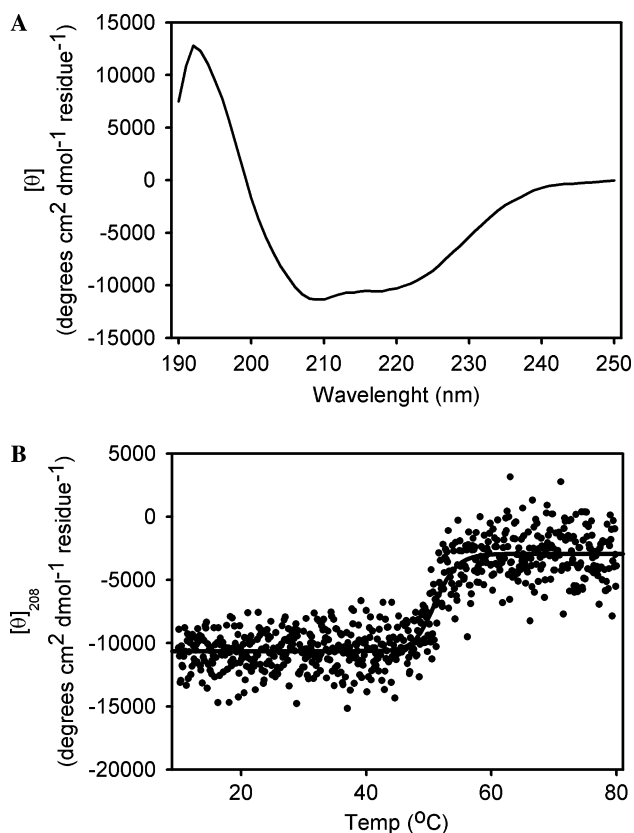


Fig. 4. Circular dichroism analysis of recombinant rPTP η CD. (A) CD spectra of native rPTP η CD at concentration of 0.08 mg/ml in 5 mM phosphate buffer (pH 7.8), 12.5 mM NaCl, 0.25 mM DTT. The spectra were recorded at 18 °C from 190 to 250 nm in a quartz 1 mm path length cuvette using Jasco J-720 spectropolarimeter as an average of 64 scans. (B) Thermal unfolding of purified rPTP η CD. The secondary structure content of the protein was monitored following the 208 nm peak during thermal denaturation (10–80 °C) with a temperature slope of 1 °C/min. Measurement were carried out in a quartz 1 mm path-length cuvette using protein at concentration of 0.08 mg/ml in 5 mM phosphate buffer (pH 7.8), 12.5 mM NaCl, and 0.25 mM DTT.

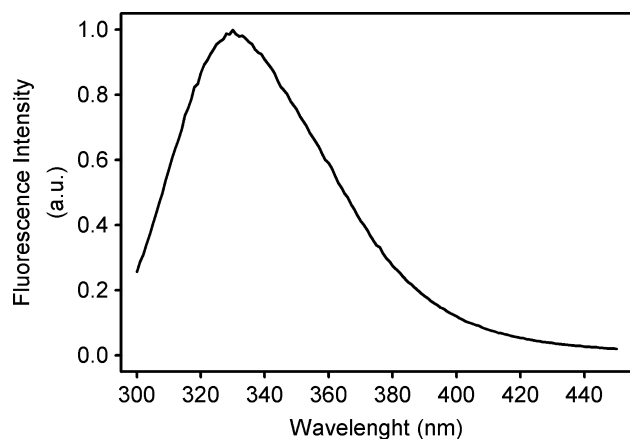


Fig. 5. rPTP η CD intrinsic fluorescence spectrum. At the excitation wavelength of 280 nm and protein concentration of 0.3 mg/ml in 100 mM Hepes (pH 7.8), 50 mM NaCl, and 1 mM DTT buffer. The emission spectra of rPTP η CD were recorded with 1 nm interval and 1 s of integration time. Experiments were carried out in a 10 × 2 mm fluorescence cuvette, at 18 °C, and corrected for buffer contribution.

thermal denaturation, reveals clear conformational changes in the protein secondary structure (data not shown).

The rPTP η CD intrinsic fluorescence emission spectrum has a maximum fluorescence wavelength of 330 nm upon excitation at 280 nm (Fig. 5). Two spectra, recorded in an interval of 1 h, are virtually identical which indicated that the protein is stable at 18 °C.

Conclusions

The protein tyrosine phosphatases comprehend a diverse family of enzymes which elicit, in different cell types, a plethora of fundamental biological effects such as gene expression, cell division and differentiation, and neoplastic transformation [1]. The expression of PTPs has been observed in various areas of the nerve system and non-neural cells, where their functions remain largely unknown [5–10]. From previous reports, it is known that the rat protein tyrosine phosphatases η gene expression inhibits the tumorigenic process in oncogene-transformed thyroid cell lines [11], whereas mutations in the human and mouse homologues *HPTP η* /*Ptprj* are related to the susceptibility in developing colon, lung, and breast cancers [23]. In the present paper we describe heterologous expression, purification of intracellular phosphatase domain of rPTP η and its characterization by SDS–PAGE, native gel electrophoresis, analytical gel filtration chromatography, far-UV circular dichroism, and intrinsic fluorescence techniques. In addition, the enzymatic activity of rPTP η CD has been tested. Taken together all the results of these investigations, we conclude that the purified catalytic phosphatase domain of rPTP η was properly folded, fully active, and therefore suitable for further structural and functional studies.

Acknowledgments

The authors thank the Brazilian Agencies: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 99/03387-4), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support.

References

- [1] J.M. Bishop, Molecular themes in oncogenesis, *Cell* 64 (1991) 235–248.
- [2] H. Charbonneau, N.K. Tonks, S. Kumar, C.D. Diltz, M. Harrylick, D.E. Cool, E.G. Trebs, E.H. Fisher, K.A. Walsh, Human placenta protein–tyrosine-phosphatase: amino acid sequence and relationship to a family of receptor-like proteins, *Proc. Natl. Acad. Sci. USA* 86 (1989) 5252–5256.

- [3] K.M. Walton, J.E. Dixon, Protein tyrosine phosphatases, *Annu. Rev. Biochem.* 62 (1993) 101–120.
- [4] Y.P. Jiang, H. Wang, P.D. Eustachio, J.M. Musacchio, J. Schlessinger, J. Sap, Cloning and characterization of R-PTP-kappa, a new member of the receptor protein tyrosine phosphatase family with a proteolytically cleaved cellular adhesion molecule-like extracellular region, *Mol. Cell. Biol.* 13 (1993) 2942–2951.
- [5] M. Sahin, J.J. Dowling, S. Hockfield, Seven protein tyrosine phosphatases are differentially expressed in the developing rat brain, *J. Comp. Neurol.* 351 (1995) 617–631.
- [6] L.P. Shock, D.J. Bare, S.G. Klinz, P.F. Maness, Protein tyrosine phosphatases expressed in developing brain and retinal Müller glia, *Mol. Brain Res.* 28 (1995) 110–116.
- [7] R.O. Hynes, A.D. Lander, Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons, *Cell* 68 (1992) 303–322.
- [8] S.M. Brady Kalnay, A.J. Flint, N.K. Tonks, Homophilic binding of PTP mu, a receptor-type protein tyrosine phosphatase, can mediate cell–cell aggregation, *J. Cell Biol.* 122 (1993) 961–972.
- [9] J. Sap, Y.P. Jiang, D. Friedlander, M. Grumet, J. Schlessinger, Receptor tyrosine phosphatase RPTP kappa mediates homophilic binding, *Mol. Cell. Biol.* 14 (1994) 1–9.
- [10] M.F.B.G. Gebbink, G.C.M. Zondag, R.W. Wubbolts, R.L. Beijersbergen, I. van, W.H. Moolenaar, Homophilic interactions mediated by receptor tyrosine phosphatases μ and κ , *J. Biol. Chem.* 268 (1993) 16101–16104.
- [11] M.L. Martelli, F. Trapasso, P. Bruni, M.T. Berlingieri, C. Battaglia, M.T. Vento, B. Belletti, R. Iuliano, M. Santoro, G. Viglietto, A. Fusco, Protein tyrosine phosphatase- η expression is upregulated by the PKA-dependent and is downregulated by the PKC-dependent pathways in thyroid cells, *Exp. Cell Res.* 245 (1998) 195–202.
- [12] L. Zhang, M.L. Martelli, C. Battaglia, F. Trapasso, J. Cerutti, D. Tramontano, G. Viglietto, A. Porcellini, M. Santoro, A. Fusco, Thyroid cell transformation inhibits the expression of a novel rat protein tyrosine phosphatase, *Exp. Cell Res.* 235 (1997) 62–70.
- [13] H. Honda, J. Inazawa, J. Nishida, Y. Yazaki, H. Hirai, Molecular cloning, characterization, and chromosomal localization of a novel protein-tyrosine phosphatase, HPTP eta, *Blood* 15 (1994) 4186–4194.
- [14] A. Ostman, Q. Yang, K. Tonks, Expression of DEP-1, a receptor-like protein tyrosine phosphatase, is enhanced with increasing cell density, *Proc. Natl. Acad. Sci. USA* 91 (1994) 9680–9684.
- [15] C.A.L. Ruivenkamp, T. van Wezel, C. Zanon, A.P.M. Stassen, C. Vlcek, Tamás Csikós, A.M. Klous, N. Tripodis, A. Perrakis, L. Boerrigter, P.C. Groot, J. Lindeman, W.J. Mooi, G.A. Meijjer, G. Scholten, H. Dauwerse, V. Paces, N. van Zandwijk, G.J.B. van Ommen, P. Demant, *Ptprj* is a candidate for the mouse colon-cancer susceptibility locus *Sccl* and is frequently deleted in human cancers, *Nat. Genet.* 31 (2002) 295–300.
- [16] N.X. Krueger, M. Streuli, H. Saito, Structural diversity and evolution of human receptor-like protein tyrosine phosphatases, *EMBO J.* 9 (1990) 3241–3252.
- [17] E.H. Fischer, H. Charbonneau, N.K. Tonks, Protein tyrosine phosphatases: a diverse family of intracellular and transmembrane enzymes, *Science* 253 (1991) 401–406.
- [18] H. Saito, Structural diversity of eukaryotic protein tyrosine phosphatases: functional and evolutionary implications, *Semin. Cell Biol.* 4 (1993) 379–387.
- [19] Y. Zhai, J. Wirth, S. Kang, C.W. Welsch, W.J. Esselman, LAR-PTPase cDNA transfection suppression of tumor growth of neu oncogene-transformed human breast carcinoma cells, *Mol. Carcinog.* 14 (1995) 103–110.
- [20] E. Ardini, R. Agresti, E. Tagliabue, M. Greco, P. Aiello, L.T. Yang, S. Ménard, J. Sap, Expression of protein tyrosine phosphatase alpha (RPTP α) in human breast cancer correlates with low tumor grade, and inhibits tumor cell growth in vitro and in vivo, *Oncogene* 19 (2000) 4979–4987.
- [21] K. Okazaki, N. Sagato, MAP kinase activation is essential for oncogenic transformation of NIH3T3 cells by Mos, *EMBO J.* 14 (1995) 5048–5059.
- [22] F. Trapasso, R. Iuliano, A. Boccia, A. Stella, R. Visconti, P. Bruni, G. Baldassarre, M. Santoro, G. Viglietto, A. Fusco, Rat protein tyrosine phosphatase η suppresses the neoplastic phenotype of retrovirally transformed thyroid cells through the stabilization of p27Kip1, *Mol. Cell. Biol.* 20 (2000) 9236–9246.
- [23] C.A. Ruivenkamp, T. van Wezel, C. Zanon, A.P. Stassen, C. Vlcek, T. Csikos, A.M. Klous, N. Tripodis, A. Perrakis, L. Boerrigter, P.C. Groot, J. Lindeman, W.J. Mooi, G.A. Meijjer, G. Scholten, H. Dauwerse, V. Paces, N. van Zandwijk, G.J. van Ommen, P. Demant, *Ptprj* is a candidate for the mouse colon-cancer susceptibility locus *Sccl* and is frequently deleted in human cancers, *Nat. Genet.* 31 (2002) 295–300.
- [24] M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [25] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685.
- [26] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res.* 31 (2003) 3784–3788.
- [27] M.F.B.G. Gebbink, M.H.G. Verheijen, G.C.M. Zondag, I. van Etten, W.H. Moolenaar, Purification and characterization of the cytoplasmic domain of the human receptor-like protein tyrosine phosphatase rPTP μ , *Biochemistry* 32 (1993) 13516–13522.
- [28] K.L. Lim, D.S.Y. Lai, M.B. Kalousek, Y. Wang, C.J. Pallen, Kinetic analysis of two closely related receptor-like protein-tyrosine-phosphatases, PTP α and PTP ϵ , *Eur. J. Biochem.* 245 (1997) 693–700.
- [29] Available from: <<http://www.upstate.com/browse/productdetail.q.ProductID.e.14-358>>.
- [30] Available from: <<http://www.purelyproteins.com/products/datasheets/ptpn11domain.html>>.
- [31] A. Lobley, L. Whitmore, B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics* 18 (2002) 211–212.
- [32] N. Sreerema, S.Y. Venyaminov, R.W. Woody, Estimation of the number of helical and strand segments in proteins using CD spectroscopy, *Protein Sci.* (1999) 370–380.
- [33] I.H.M. Van Stokkum, H.J.W. Spoelder, M. Bloemendal, R. Van Grondelle, F.C.A. Groen, Estimation of protein secondary structure and error analysis from CD spectra, *Anal. Biochem.* 191 (1999) 110–118.
- [34] M.A. Andrade, P. Chacón, J.J. Merelo, F. Morán, Evaluation of secondary structure of proteins from UV circular dichroism using an unsupervised learning neural network, *Protein Eng.* 6 (1993) 383–390.
- [35] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics* 16 (2000) 404–405.
- [36] G.P.S. Raghava, Protein secondary structure prediction using nearest neighbor and neural network approach, *CASP4* (2000) 75–76.
- [37] A.M. Bilwes, J. den Hertog, T. Hunter, J.P. Noel, Structural basis for inhibition of receptor protein-tyrosine phosphatases-alpha by dimerization, *Nature* 382 (1996) 555–559.
- [38] K.M.V. Hoffmann, N.K. Tonks, D. Barford, The crystal structure of domain 1 of receptor protein-tyrosine phosphatases μ , *J. Biol. Chem.* 272 (1997) 27505–27508.

Crystallization and preliminary X-ray diffraction studies of isoform $\alpha 1$ of the human thyroid hormone receptor ligand-binding domain

F. M. Nunes,^a R. Aparicio,^a ‡
M. A. M. Santos,^a R. V. Portugal,^a
S. M. G. Dias,^a F. A. R. Neves,^b
L. A. Simeoni,^b J. D. Baxter,^c
P. Webb^c and I. Polikarpov^{a*}

^aInstituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador SãoCarlense 400, São Carlos, SP, 13560-970, Brazil,

^bUniversidade de Brasília, Faculdade de Ciências da Saúde, Laboratório de Farmacologia Molecular, Brasília, DF, Brazil, and ^cUniversity of California, Metabolic Research Unit, San Francisco, USA

‡ Current address: Universidade Estadual de Campinas, Instituto de Química, CP6154, Campinas, SP, 13084-862, Brazil

Correspondence e-mail:
ipolikarpov@if.sc.usp.br

Thyroid hormone receptors (TR) play critical roles in virtually all tissues. The TR ligand-binding domain (LBD) participates in important activities, such as transcriptional activation and repression, through conformational changes induced by hormone binding. Two crystal forms of isoform $\alpha 1$ of the human thyroid hormone receptor LBD (hTR $\alpha 1$) in complex with the thyroid hormones T3 and Triac were obtained. The hTR $\alpha 1$ -T3 complex was crystallized in a previously unobserved crystal form (space group $P2_12_12_1$, $a = 59.98$, $b = 80.80$, $c = 102.21$ Å), with diffraction patterns extending to 1.90 Å resolution on a rotating-anode X-ray source, and in space group $C2$ ($a = 117.54$, $b = 80.66$, $c = 62.55$ Å, $\beta = 121.04^\circ$), with data extending to 2.32 Å resolution. The hTR $\alpha 1$ -Triac complex was also crystallized in the new space group $P2_12_12_1$, with unit-cell parameters $a = 60.01$, $b = 80.82$, $c = 102.39$ Å; its resolution limit extended to 2.20 Å on a home source. Phasing was carried out by the molecular-replacement method and structural refinement is currently in progress. The refined structures may provide insight into the design of new thyromimetics.

Received 5 May 2004

Accepted 20 July 2004

1. Introduction

Thyroid hormone receptors (TRs), members of a superfamily of eukaryotic transcription factors, are ligand-activated transcription factors that bind to thyroid hormone-response elements (TREs) in the regulatory region of target genes and mediate the biological effects of thyroid hormones. The receptors exhibit a modular structure with functionally separable domains. The most highly conserved domains are the DNA-binding domain (DBD) and the ligand-binding domain (LBD) (Evans, 1988; Laudet *et al.*, 1992). The LBD participates in several types of activity, including hormone binding, homo- and/or heterodimerization, molecular interactions with heat-shock proteins and transcriptional activation and repression (Tsai & O'Malley, 1994; Ribeiro *et al.*, 1995). Hormone binding induces conformational changes which control these properties and influence gene expression (Tsai & O'Malley, 1994; Ribeiro *et al.*, 1995). Therefore, the three-dimensional structure of a liganded LBD is critical for understanding the structural mechanisms of hormone action.

Thyroid hormones, namely 3,5,3'-triiodo-L-thyronine (T3), 3,5,3',5'-tetraiodo-L-thyronine (T4) and 3,5,3'-triiodothyroacetic acid (Triac), play critical roles in the differentiation, growth, metabolism and physiological function of virtually all tissues. Two major subtypes of thyroid hormone receptors (TR $\alpha 1$ and TR $\alpha 2$, and TR $\beta 1$ and TR $\beta 2$) are encoded by two

different genes (Ribeiro *et al.*, 1995). TR $\alpha 1$, TR $\beta 1$ and TR $\beta 2$ are ligand-binding isoforms of TR, whereas TR $\alpha 2$ does not bind thyroid hormones. Differences in affinity towards thyroid hormones are observed amongst the different ligand-binding isoforms. TRs mediate distinct physiological effects owing to differences in tissue abundance and receptor-specific activity (Forrest & Vennström, 2000). Studies in patients with the syndrome of resistance to thyroid hormones, in which abnormal TR β is present, and with TR $\alpha 1$ ^{-/-} mice suggest that TR α is the major TR regulating heart rate (Johansson *et al.*, 1998; Forrest & Vennström, 2000; Yen, 2001; Gloss *et al.*, 2001). TR β is critical in controlling hepatic cholesterol metabolism and thyroid-stimulating hormone (TSH) suppression, which may be because of the high expression of TR β in liver (70–80% of total TR) and pituitary (Schwartz *et al.*, 1992; Wikström *et al.*, 1998; Gloss *et al.*, 2001). In particular, TR $\beta 1$ is widely distributed in the tissues and regulates metabolic rate. Identifying thyromimetics that interact selectively with the isoforms TR $\alpha 1$ and TR $\beta 1$ may be crucial in treating important diseases such as obesity and lipid disorders.

The available TR structural data include rat TR $\alpha 1$ LBD in complex with T3 (Wagner *et al.*, 1995) and Triac (Wagner *et al.*, 2001) and human TR $\alpha 1$ LBD (hTR $\alpha 1$) in complex with a synthetic thyromimetic (Ye *et al.*, 2003). However, despite this previous work, the structural basis for the selectivity of the TR

isoforms in thyroid ligand binding is not yet well established. To further investigate the molecular mechanism of hTR α 1 specificity, X-ray crystallographic studies were recently initiated. In this work, we report the purification, crystallization, data collection and molecular-replacement solutions obtained for hTR α 1 in complex with T3 and Triac. Data to higher resolution than previously published (Ye *et al.*, 2003; Dow *et al.*, 2003) have been obtained, which is likely to be a consequence of a different molecular packing. These new studies may be important for better understanding of the structural basis of TR isoform selectivity and for the design of more potent isoform-specific thyromimetics.

2. Materials and methods

2.1. Expression and purification

The human TR α 1 LBD construct including amino-acid residues Glu148–Val410 (NCBI protein accession No. A40917) fused in frame to the C-terminus of a polyhistidine (His) tag was expressed in *Escherichia coli* strain B834 harbouring a pET28a(+) plasmid (Novagen), as illustrated in Fig. 1. A Luria Broth [LB; 1.6% (w/v) tryptose, 1% (w/v) yeast extract, 0.5% (w/v) NaCl] starter culture was inoculated with a single colony of an LB-agar culture and grown overnight at 310 K. The initial culture was inoculated at 1% in a major LB culture and grown at 293 K in kanamycin medium until the $A_{600\text{nm}}$ reached 1.7. After this, isopropyl thio- β -D-galactoside (IPTG) was added to a concentration of 0.5 mM and culture growth was continued for 6 h at 293 K.

The cells were harvested by centrifugation and the pellet resuspended in 50 mM Tris–HCl buffer pH 7.5 containing 150 mM NaCl, 0.05% Tween-20, 1 mM phenylmethylsulfonylfluoride (PMSF) and 20 mM 2-mercaptoethanol. The culture was incubated on ice with 0.5 mg ml⁻¹ lysozyme and disrupted by sonication. The lysate was centrifuged for 20 min at 14 000 rev min⁻¹ in a Sorvall SS34 rotor at 277 K and the obtained supernatant was incubated for 30 min with a 20-fold molar excess of T3 (Sigma) or Triac (Sigma).

To purify the holo hTR α 1, the supernatant was mixed with Talon Superflow Metal Affinity Resin (Clontech) and shaken at 277 K for 1 h. The resin was washed twice with 50 mM sodium phosphate buffer pH 8.0 containing 300 mM NaCl, 10% glycerol, 10 mM 2-mercaptoethanol, 25 mM imidazole, 1 mM PMSF and 0.05% Tween-20 and

Table 1
Crystal parameters and data-collection statistics.

	T3 complex		Triac complex
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>C</i> 2	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Unit-cell parameters			
<i>a</i> (Å)	59.98	117.54	60.01
<i>b</i> (Å)	80.80	80.66	80.82
<i>c</i> (Å)	102.21	62.55	102.39
β (°)		121.04	
Solvent content (%)	68.3	69.1	68.4
ASU content (molecules)	1	1	1
Resolution range (Å)	31.47–1.90 (2.00–1.90)	19.92–2.32 (2.45–2.32)	48.22–2.20 (2.32–2.20)
No. images	260	77	291
No. observed reflections	323718	32461	267634
No. unique reflections	39272	18188	26000
Multiplicity	8.2 (7.8)	1.8 (1.7)	10.3 (11.0)
Completeness (%)	98.7 (98.9)	83.7 (87.3)	100 (100)
R_{sym} (%)	5.9 (38.3)	7.9 (37.5)	5.9 (37.5)
$\langle I/\sigma(I) \rangle$	7.8 (2.0)	7.7 (2.0)	11.2 (2.1)

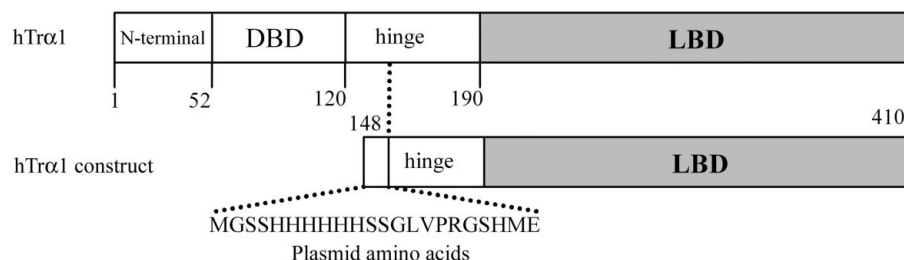


Figure 1
Schematic diagrams of human TR α 1 (hTR α 1) and the construct used in the crystallographic experiments (hTR α 1 construct), showing the N-terminal, DBD (DNA-binding), hinge (connecting region) and LBD (ligand-binding) domains. Numbers indicate the amino-acid positions in the receptors. The amino acids of the plasmid pET28a(+) that is coexpressed with hTR α 1 construct are shown.

twice with the same solution without Tween-20. The protein was eluted in a single step with 50 mM sodium phosphate buffer pH 8.0 containing 300 mM NaCl, 10% glycerol, 10 mM 2-mercaptoethanol, 1 mM PMSF, 0.05% Tween-20 and 500 mM imidazole.

After the affinity column, the fractions were pooled and corrected for the conductivity of the initial phenyl solution. The protein was applied onto a Phenyl 5PW 8/75 (TosoHaas) column pre-equilibrated with 20 mM Na HEPES buffer pH 8.0 containing 0.5 mM EDTA, 700 mM (NH₄)₂SO₄, 3 mM dithiothreitol (DTT). The column was washed with the previous solution and eluted with a 90 min 0–100% gradient of 20 mM Na HEPES buffer pH 8.0 containing 0.5 mM EDTA, 20% glycerol, 10% acetonitrile, 3 mM DTT at 0.75 ml min⁻¹. After this step, the protein was loaded onto a HL Superdex 200 or 75 26/60 gel-filtration column (Amersham Bioscience) equilibrated with 20 mM Na HEPES buffer pH 8.0 containing 1 mM EDTA, 3 mM DTT, 0.01% Tween-20 and 200 mM NaCl. The protein recovered was concentrated by ultrafiltration (Amicon Ultra 10 kDa MWCO, Millipore), resuspended in a solution of 600 mM NaCl and 3 mM DTT and further concentrated to 10 mg ml⁻¹.

The protein content and purity of all chromatographic fractions were checked by Coomassie Blue-stained SDS–PAGE. The average yield of the protein was about 6 mg per litre of culture, with purity higher than 99%. Protein concentrations were determined using the Bradford dye assay (Bio-Rad) using bovine serum albumin as standard.

2.2. Crystallization

Initial crystallization conditions were screened at temperatures of 277 and 291 K by the sparse-matrix method (Jancarik & Kim, 1991) using the macromolecular crystallization reagent kits Crystal Screens I and II (Hampton Research). In each trial, a hanging drop of 1 μ l of protein solution containing either T3 or Triac (see §2.1) was mixed with 1 μ l precipitant solution and equilibrated against a reservoir containing 500 μ l precipitant solution. At both temperatures, evidence for crystals was found in Hampton Crystal Screen I solution No. 07, formulated with 1.4 M sodium acetate trihydrate (NaH₃OAc) and 0.1 M sodium cacodylate (NaCac) pH 6.5. For both the T3 and Triac complexes, further optimization at 291 K led to crystallization condi-

tions similar to those reported for human TR β LBD complexes (Wagner *et al.*, 2001). A reservoir solution containing 1.0 M NaCac and 0.1 M NaH₃OAc pH 7.2 was mixed with protein solution in equal amounts and equilibrated against reservoir solution at 291 K. Well formed crystals grew within 12–24 h.

Two different crystal forms of hTR α 1 in complex with T3 have been obtained (Table 1). Crystals were grown under the same crystallization conditions and exhibited similar morphology (Fig. 2). In the case of the hTR α 1–Triaic complex, only crystals belonging to the previously unobserved space group $P2_12_12_1$ were obtained, with unit-cell parameters similar to those of the corresponding crystal form of the hTR α 1–T3 complex; however, they had a somewhat different morphology (Table 1 and Fig. 3).

2.3. Data collection

X-ray diffraction experiments were performed with a MAR Research MAR345dtb image-plate detector mounted on a Rigaku UltraX 18 rotating-anode X-ray generator providing Cu $K\alpha$ radiation (1.5418 Å), operated at 50 kV and 100 mA and equipped with Osmic confocal Max-Flux optics.

To prevent radiation damage, cryocrystallographic techniques (Garman &

Schneider, 1997) were employed. Crystals were briefly soaked in a cryoprotectant solution containing 1.0 M NaCac, 0.1 M NaH₃OAc pH 7.2 and 20%(v/v) ethylene glycol and rapidly cooled in a gaseous nitrogen stream (Oxford Cryosystems). Data were collected by the oscillation method from single crystals maintained at 100 K during data collection. In all cases, the oscillation range was 1°, with exposure times of 6 min (T3 complex, space group $P2_12_12_1$), 15 min (T3 complex, space group $C2$)

and 20 min (Triaic complex) per image (Fig. 4).

A single data set was collected for each crystal form of hTR α 1–T3. X-ray data for the orthorhombic crystal were collected with a crystal-to-detector distance of 150 mm, giving an outer edge resolution of 1.86 Å. For the monoclinic form, the crystal-to-detector distance was set to 200 mm, with a maximum resolution of 2.21 Å at the detector edge. The hTR α 1–Triaic data set was collected using a crystal-to-detector

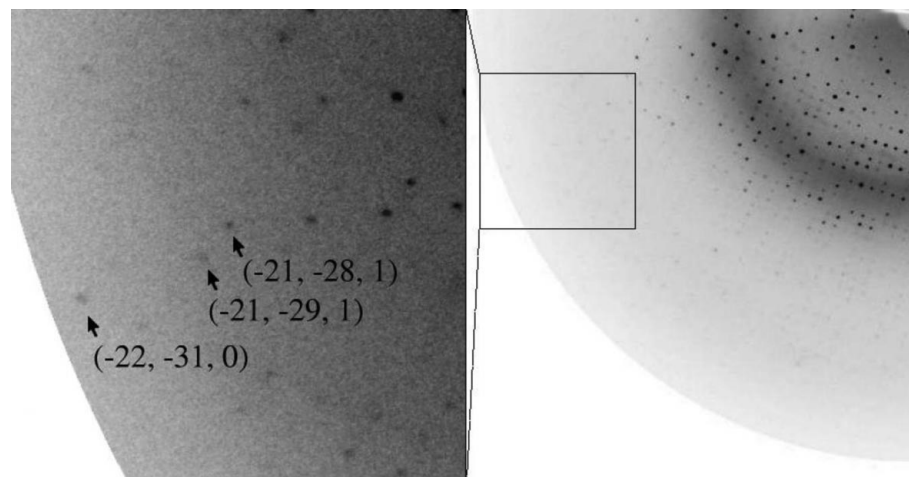


Figure 4

A 1° oscillation frame from a cryocooled hTR α 1–T3 crystal. Diffraction spots with hkl indices $(-21, -28, 1)$, $(-21, -29, 1)$ and $(-22, -31, 0)$ were marked at resolutions of 2.03, 2.00 and 1.88 Å, respectively. The resolution at the edge of the image is 1.87 Å.

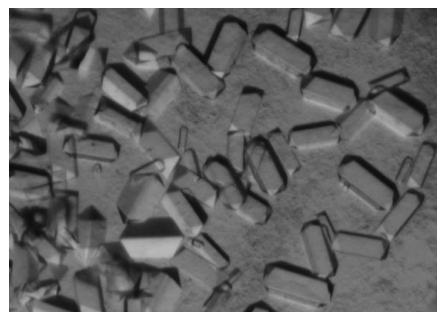


Figure 2

Single crystals of hTR α 1 in complex with T3. Typical dimensions are 0.5 × 0.4 × 0.3 mm.

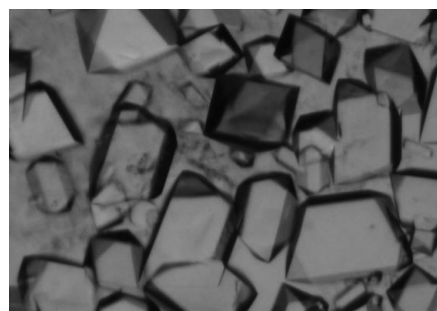


Figure 3

hTR α 1–Triaic crystals measuring approximately ~0.5 mm in the longest dimension.

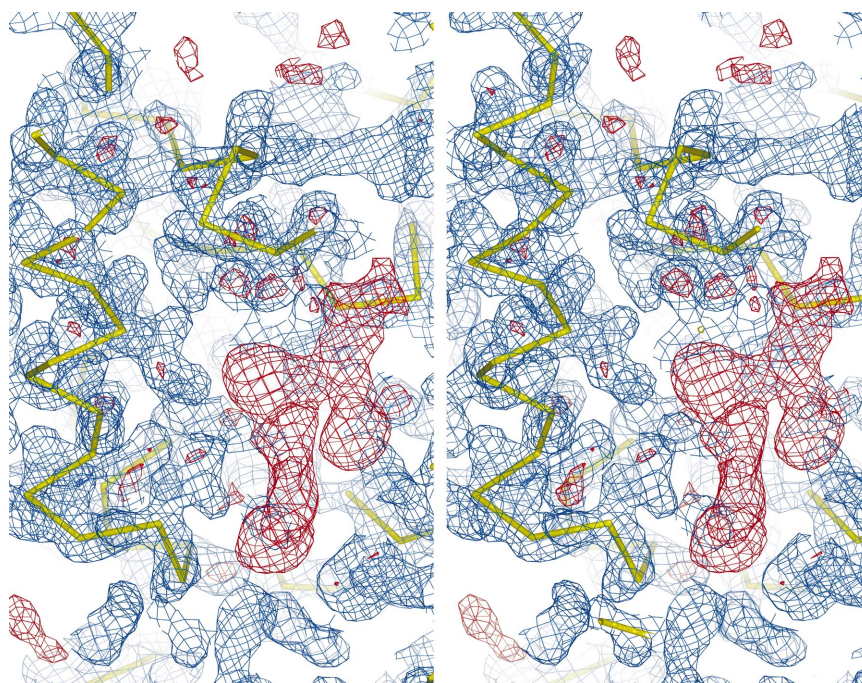


Figure 5

Stereoview of the hTR α 1–T3 complex ligand-binding region. Initial $2F_{\text{obs}} - F_{\text{calc}}$ (blue) and $F_{\text{obs}} - F_{\text{calc}}$ (red) electron-density maps were contoured at the 1.0σ and 3.0σ levels, respectively, around the model (C^{α} trace, coloured yellow). The figure unequivocally indicates the presence of the ligand. The drawing was prepared using *PyMOL* (DeLano Scientific, San Carlos, CA, USA; <http://www.pymol.org>).

distance of 170 mm, with an outer edge resolution of 2.00 Å.

3. Results and discussion

Data reduction was performed using *MOSFLM* and *SCALA* (Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 1997). Crystal parameters and data-collection statistics are summarized in Table 1. The solvent content was calculated using the total molecular weight of the hTR α 1 amino-acid sequence in addition to 21 amino acids from the initiation codon, yielding a polypeptide of 284 amino acids (MW = 32 205 Da; Table 1). The correct number of molecules present in the asymmetric unit (ASU) was determined during the molecular-replacement procedure.

Primary sequence search and sequence alignments were performed using *ENTREZ* and *BLAST* (Altschul *et al.*, 1997). A sequence identity of 83%, with 92% similarity, resulted from the alignment of hTR α 1 (Nakai *et al.*, 1988) and human TR β (Weinberger *et al.*, 1986; Sakurai *et al.*, 1990). Thus, chain A of the corresponding dimeric X-ray structure (PDB code 1bsx; Darimont *et al.*, 1998) was used as a search model (waters, the T3 molecule and a peptide fragment of GRIP1, a coactivator, were excluded from the search model). It is worth noting that the recently reported thyroid receptor structures (PDB codes 1nav and 1nax; Ye *et al.*, 2003) could equally well be used in this case.

Molecular-replacement calculations were carried out using a resolution range of 10.0–4.0 Å and default parameters in the program *AMoRe* (Navaza, 1994; Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 1997). In all cases, clear solutions were obtained for one molecule in the ASU. A search for a second molecule was not successful. After fitting, correlation coefficients of 60.2, 62.9 and 62.4% resulted for hTR α 1–T3 in the orthorhombic crystal form, hTR α 1–T3 in the monoclinic form and hTR α 1–Triac, respectively. The corre-

sponding *R* factors were 41.7, 39.5 and 41.9%, respectively.

The I atoms present in the T3 and Triac molecules provided significant anomalous signal, except in the case of the monoclinic form of hTR α 1–T3 complex, the data set of which contains a smaller number of images. Thus, to certify whether the ligands were bound to the active sites in the three structures, the following procedure was employed. For each complex, an initial model was built by application of the suitable molecular-replacement solution (*PDBSET*; Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 1997), from which structure factors were derived using the program *SFALL* (Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 1997). Electron-density maps ($2F_{\text{obs}} - F_{\text{calc}}$ and $F_{\text{obs}} - F_{\text{calc}}$) were calculated and extended around the initial model. The known ligand-binding region was assessed by visual inspection using the program *O* (Kleywegt *et al.*, 2001), contouring the $2F_{\text{obs}} - F_{\text{calc}}$ and $F_{\text{obs}} - F_{\text{calc}}$ maps at the 1.0 σ and 3.0 σ levels, respectively. This approach was successful in clearly identifying the presence of the ligand in all three initial structures, as expected. Fig. 5 illustrates this result for the hTR α 1–T3 complex in the orthorhombic crystal form. Structural refinement is in progress.

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via grants 02/13577-0 and 99/03387-4, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and National Institutes of Health Grant DK41842 (to JDB). JDB has proprietary interests in and serves as a consultant and Deputy Director to Karo Bio AB, which has commercial interests in this area of research.

References

Altschul, S. F., Madden, T., Schffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.

Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J. & Yamamoto, K. R. (1998). *Genes Dev.* **12**, 3343–3356.
Dow, R. L. *et al.* (2003). *Bioorg. Med. Chem. Lett.* **13**, 379–382.
Evans, R. M. (1988). *Science*, **240**, 889–895.
Forrest, D. & Vennström, B. (2000). *Thyroid*, **10**, 41–52.
Garman, E. F. & Schneider, T. R. (1997). *J. Appl. Cryst.* **30**, 211–237.
Gloss, B., Trost, S., Bluhm, W., Swanson, E., Clark, R., Winkfein, R., Janzen, K., Giles, W., Chas-sande, O., Samarut, J. & Dillmann, W. (2001). *Endocrinology*, **142**, 544–550.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Johansson, C., Vennström, B. & Thoren, P. (1998). *Am. J. Physiol.* **275**, R640–R646.
Kleywegt, G., Zou, J., Kjeldgaard, M. & Jones, T. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, ch. 17.1, pp. 353–356, 366–367. Dordrecht: Kluwer Academic Publishers.
Laudet, V., Hanni, C., Coll, J., Catzeflis, F. & Stehelin, D. (1992). *EMBO J.* **11**, 1003–1013.
Nakai, A., Sakurai, A., Bell, G. I. & DeGroot, L. J. (1988). *Mol. Endocrinol.* **2**, 1087–1092.
Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
Ribeiro, R. C., Kushner, P. & Baxter, J. D. (1995). *Annu. Rev. Med.* **46**, 443–453.
Sakurai, A., Nakai, A. & DeGroot, L. J. (1990). *Mol. Cell. Endocrinol.* **71**, 83–91.
Schwartz, H. L., Strait, K. A., Ling, N. C. & Oppenheimer, J. H. (1992). *J. Biol. Chem.* **267**, 11794–11799.
Tsai, M. & O'Malley, B. W. (1994). *Annu. Rev. Biochem.* **63**, 451–486.
Wagner, R. L., Apriletti, J. W., McGrath, M. E., West, B. L., Baxter, J. D. & Fletterick, R. J. (1995). *Nature (London)*, **378**, 690–697.
Wagner, R. L., Huber, B. R., Shiau, A. K., Kelly, A., Cunha Lima, S. T., Scanlan, T. S., Apriletti, J. W., Baxter, J. D., West, B. L. & Fletterick, R. J. (2001). *Mol. Endocrinol.* **15**, 398–410.
Weinberger, C., Thompson, C. C., Ong, E. S., Lebo, R., Gruol, D. J. & Evans, R. M. (1986). *Nature (London)*, **324**, 641–646.
Wikström, L., Johansson, C., Salto, C., Barlow, C., Campos Barros, A., Bass, F., Forrest, D., Thoren, P. & Vennström, B. (1998). *EMBO J.* **17**, 455–461.
Winn, M., Dodson, E. J. & Ralph, A. (1997). *Methods Enzymol.* **277**, 620–633.
Ye, L., Li, Y., Mellstrom, K., Mellin, C., Bladh, L. G., Koehler, K., Garg, N., Garcia Collazo, A. M., Litten, C., Husman, B., Persson, K., Ljunggren, J., Grover, G., Slep, P. G., George, R. & Malm, J. (2003). *J. Med. Chem.* **46**, 1580–1588.
Yen, P. M. (2001). *Physiol. Rev.* **81**, 1097–1142.

Expression, purification, and initial structural characterization of rat orphan nuclear receptor NOR-1 LBD domain

Guilherme Razzera^a, Javier Vernal^a, Rodrigo V. Portugal^b, Marcos R. Calgare^b, Pablo Fernandez^c, Mario M. Zakin^c, Igor Polikarpov^b, Hernán Terenzi^{a,*}

^a Laboratório de Expressão Gênica, Departamento de Bioquímica, Universidade Federal de Santa Catarina, 88040-900 Florianópolis, SC, Brazil

^b Instituto de Física de São Carlos, USP-Av. Trabalhador São-carlense 400, 13560-970 São Carlos, SP, Brazil

^c Institut Pasteur, Unite d' Expression des Genes Eucaryotes, Institut Pasteur, 75015 Paris, France

Received 7 May 2004, and in revised form 11 June 2004

Available online 4 August 2004

Abstract

NOR-1 is an orphan member of the nuclear receptor superfamily, which includes a group of transcription factors involved in the response to steroids, fatty acids, retinoic acids, and other lipophilic molecules. The NOR-1 subfamily (NR4), composed also of Nurr1 and Nurr77, has been implicated in cell proliferation, differentiation, apoptosis, chondrosarcomas, inflammation, and atherogenesis. The NOR-1 receptor is an orphan ligand receptor which acts over gene transactivation. No ligands, if such in fact exist, are known for this receptor. Recently, the three-dimensional structure of the homolog receptor Nurr1 has been solved using protein crystallography techniques. Surprisingly, the structure does not present either a typical cavity for ligand binding or a classical co-factor binding site in the ligand binding domain (LBD). To allow for structural studies of other members of NR4 subfamily, we have subcloned, overexpressed in *Escherichia coli* cells, purified, and characterized the rat orphan nuclear receptor NOR-1 LBD domain. We obtained NOR-1 LBD at a high degree of purity and with an overall yield of 3 mg/L of culture media. CD spectroscopic analysis shows a high α -helical secondary structure content (52%), similar to that of Nurr 1 LBD three-dimensional structure. Thermal denaturation monitored by UV absorption and CD spectroscopy suggests proper folding of recombinant NOR-1 LBD.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Nuclear receptor; NOR-1 receptor; NR4 family; Expression in *Escherichia coli*; LBD domain

The hormone nuclear receptor superfamily includes different proteins responsible for the direct link between transcriptional control and physiological responses. Nuclear receptors (NRs) are one of the largest groups of transcription factors [1]. The activity of many nuclear receptors is controlled by the binding of small lipophilic ligands that include hormones, metabolites such as fatty acids, bile acids, oxysterols, and xeno- and endobiotics. Several of these receptors were characterized before the identification of their ligands and, for this reason, they

were termed orphan receptors. As their natural and synthetic ligands have become known, many of these orphans have been adopted and included in other subfamilies [2]. The recognition of the importance of nuclear receptors as master regulators of gene transcription involved in metabolic control has resulted in the intensive search for novel ligands for these receptors that might be used in preventive and therapeutic strategies to combat common diseases such as atherosclerosis, diabetes, and obesity [3]. The importance of NRs in human physiology is underscored by the extensive range of therapeutics that has been created to treat disorders associated with dysfunctional hormone signalling [4].

* Corresponding author. Fax: +55-48-331-9672.

E-mail address: hterenzi@ccb.ufsc.br (H. Terenzi).

The NRs have a modular structure composed of three conserved domains. The N-terminal AF-1 region is a ligand-independent domain involved in transactivation and coactivator recruitment, which can be phosphorylated [4,5]; the highly conserved DNA binding domain (DBD) is composed of two zinc fingers recognizing specific DNA sequences [6,7]; and, finally, the ligand binding and dimerization domain (LBD) contains a ligand-dependent AF-2 transactivation region in its C-terminal portion [8].

The NR4 subfamily has been implicated in cell proliferation, differentiation, apoptosis, chondrosarcomas, inflammation, and atherogenesis [9–13]. The receptors of this subfamily share extensive homology in the DNA-binding domain (DBD) and the ligand binding domain (LBD), but diverge significantly in the N-terminal region [14]. The first low resolution model of nuclear receptor containing both DBD and LBD domains has been recently revealed by small-angle X-ray scattering technique using synchrotron radiation [15].

The rat orphan nuclear receptor NOR-1/NR4A3 and its close relatives Nurr1 and Nur77 have been shown to play a key role in regulating expression of various genes in hypothalamic–pituitary–adrenal axis. The NOR-1 and Nur77 amino acid sequence homology and the analysis of the *trans*-activation properties of NOR1 demonstrated that the overall structural and functional organization of the two proteins is similar [16]. Corticotropin-releasing hormone treatment of adrenal and pituitary cells induces NOR-1 and Nur77 expression, exemplifying their role in steroidogenesis. In accordance with the role of NOR-1 in inflammation and carcinogenesis, NOR-1 is activated by the anti-cancer and anti-inflammatory drug 6-mercaptopurine [4]. It was demonstrated that NOR-1 mediates transcription and co-factor recruitment operating in an AF-1 dependent manner. Furthermore, the AF-1 domain is required for NOR-1-dependent transactivation of target genes [4]. This receptor is also regulated by Ca²⁺ calmodulin-dependent protein kinase cascade [17] and is induced by parathyroid hormone in osteoblasts [18].

Many agonists of NRs have been identified, such as fatty acids, oxysterols, and bile acids, supporting the idea that most NRs are regulated by ligands. In contrast, the structural studies of orphan members indicate that various NRs like the hepatocyte nuclear factor 4 α (HNF4 α), and HNF4 γ , although retaining capacity to bind ligands, may not be ligand-regulated [19,20].

The crystal structure of homologous receptor Nurr1 LBD domain was solved at 2.2 Å resolution [21]. Two distinct features differ Nurr1 LBD domain from the characteristic canonical fold of other NR LBDs. The Nurr1 LBD domain contains no ligand binding cavity as a result of the tight packing of side chains of several bulky hydrophobic residues in the region normally occu-

ried by ligands. In addition, this LBD lacks a classical binding site for coactivators [21]. Other members of NR4 subfamily were not structurally characterized up to date [5].

To investigate the structural features of another NR4 subfamily member, we have subcloned, expressed, purified, and partially characterized the NOR-1 LBD domain. Here, we present the results of NOR-1 LBD purification and initial structural characterization, which confirms its proper folding, and provides the basis for further structural studies.

Materials and methods

Construction of pGEX-LBD vector

A fragment of 768 bp corresponding to the rat LBD domain of NOR-1 (amino acids 380–628) (Accession No. P51179) was amplified by PCR using rat NOR-1 cDNA as template and two synthetic oligonucleotides (LBDfor and LBDrev) as primers. The following oligonucleotides were used for PCR amplification of LBD: LBDfor, 5'-CCTGAATTCCCAAAGAGCCCCTACTAA-3' (*EcoRI*); LBDrev, 5'-TTCGAATTCTCAGAAAGGCAGGGTATC-3' (*EcoRI*), containing the restriction sites indicated. The resulting fragment was digested with *EcoRI* and ligated into the pGEX-2T –128/129 vector [22] previously digested with the same enzyme. The fragment was inserted downstream and in-frame with the glutathione *S*-transferase (GST) coding sequence plus a sequence of 18 amino acids corresponding to the *Bam*HI restriction site, a FLAG peptide, and an HMK (heart muscle kinase) recognition site (GS-DYKDDDDKARRASV) followed by glutamic acid and phenylalanine residues [22], and the construct was verified by DNA sequencing. *Escherichia coli* strain DH5 α was transformed with the obtained plasmid (named pGEX-LBD) to overexpress NOR-1 LBD.

Expression and purification of NOR-1 LBD domain

The LBD domain (30,138.5 Da) corresponding to amino acid residues 380–628, including the FLAG peptide (in italic):

(*gsdykdddldkarrasvefPKSPLQQEPSQSPSPSPICM*
 MNALVQRALTDATPRDLTDYSRYCPTDQATAGTD
 AEHVQFYNLLTASIDVSRSWAEKIPGFTDLPKE
 DQTLLESFAFLFVLRLSIRSNTAEDKVFVFCNG
 LVLHRLQCLRGFGEWLDSIKDFSLNLQSLNLDI
 QALACLALSAMITERHGLKEPKRVEELCNKITSS
 LKDHQRKQALEPSEPKVLRALVELRKICTQGL
 QRIFYLKLEDLVSPSPVIDKLFDLTLFPF) was expressed in *E. coli* and purified. The DH5 α bacteria cells were transformed with pGEX-LBD, grown in Luria–Bertani (LB) medium supplemented with ampicillin at

Table 1
Purification protocol of NOR-1 LBD from 1 L culture

Fraction	Yield (mg) ^a	Recovery (%)
I	19.2 ^b	100
II	11.5 ^b	59.8
III	3.25 ^c	16.9
IV	3 ^c	15.6

Note. I, GST-NOR-1 LBD protein before GSH elution; II, GST-NOR-1 LBD protein after first elution with GSH; III, NOR-1 LBD protein after thrombin treatment; and IV, NOR-1 LBD protein purified after second elution with GSH.

^a Yield was calculated from densitometric scanning of stained gels, and corresponded to ^bGST-NOR-1 LBD or ^cNOR-1 LBD.

0.1 mg/ml, to an absorbance of 1.7 at 600 nm, and were then induced with 1.0 mM isopropyl- β -D-thiogalactopyranoside (IPTG) for 16 h at 37°C under good aeration. Cells were harvested by centrifugation at 6000g. The pellet from 1000 ml of culture was suspended in 20 ml of phosphate-buffered saline (PBS) with 5 mM dithiothreitol (DTT) and protease inhibitor cocktail (Complete, Mini, Boehringer–Mannheim). The cells were disrupted by gentle sonication (4 cycles/20 s) in ice, Triton X-100 was added to a final concentration of 1% and the suspension was centrifuged for 20 min at 15000g. The supernatant was incubated for 90 min at 4°C with 2 ml of glutathione–Sepharose (GSH–Sepharose, Amersham). The resin was washed with 30 volumes of PBS (phosphate-buffered saline). The fusion protein was eluted with 10 ml of 10 mM reduced glutathione. The eluted fraction was concentrated by ultrafiltration on Amicon Membranes (Millipore) and the buffer was changed for cleavage with thrombin protease in the same concentrator tube (50 mM Tris–HCl, pH 8, 150 mM NaCl, 2.5 mM CaCl₂, and 5 mM DTT). In this manner, free GSH was eliminated. The fusion protein was digested with thrombin (1 U thrombin/100 μ g fusion protein) in a reaction mixture containing 50 mM Tris–HCl, pH 8, 150 mM NaCl, 2.5 mM CaCl₂, and 5 mM DTT, for 20 h at 4°C. The cleaved GST-LBD protein was incubated for 2 h at 4°C with 2 ml of glutathione–Sepharose and the eluted fractions were collected. The concentration of protein was estimated by the Bradford method [23]. Fractions corresponding to each purification step were analyzed by SDS–PAGE (sodium dodecyl sulfate–gel electrophoresis) [6]. The gel was stained with Coomassie brilliant blue R-250. The fusion protein GST-LBD and LBD were quantified by densitometric scanning of the gel shown in Fig. 2, using protein concentration in the final purified sample estimated by the Bradford method as the standard (see also Table 1).

NOR-1 LBD N-terminal sequencing

NOR-1 LBD protein was separated by SDS–PAGE and blotted onto a polyvinylidene fluoride (PVDF)

membrane. The protein band was excised and subjected to N-terminal protein sequencing by sequential Edman degradation (Applied Biosystems Procise Sequencer).

Mass spectrometry

Mass spectra were obtained by using a protein biochip technology from Ciphergen (SELDI: Surface Enhanced Laser Desorption/Ionization).

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were measured between 4 and 80°C in a Jasco J-715 spectropolarimeter equipped with a peltier unit. The α -helical content of the protein was monitored following the 222 nm peak during the thermal denaturation. All measurements were converted to molar ellipticity. The far-UV spectra of LBD protein were measured over a wavelength range of 190–240 nm as an average of 5 spectra with 100 nm/min scan speed and a step resolution of 0.1 nm. Temperature increase slope was 1°C/min. All measurements were carried out in a 0.1 mm path-length cuvette using protein at a concentration of 10 μ M in PBS buffer with NaCl concentration adjusted to 17 mM. Blank was subtracted from all spectra. The Bradford method [23] was used to determine the protein concentration using bovine serum albumin as the standard.

Denaturation under UV monitoring

The analysis was carried out at 55°C using a spectrophotometer Ultrospec 3100 pro (Amersham) at 280 nm and monitoring the signal between 0 and 2.5 h. The experiment was performed in triplicate [24].

Results and discussion

NOR-1 LBD domain (DNA sequence corresponding to amino acids 380–628 in full-length NOR-1) was amplified by PCR and subcloned in vector *EcoRI* site for expression in pGEX-2T 128/129 as a fusion protein with GST (Fig. 1A). The resulting plasmid of 5754 bp was named pGEX-LBD. The construct was checked by DNA sequencing of both strands. Alignments (Clustal W 1.82) of the NR4 subfamily LBD domains showed a high degree of similarity between NOR-1 and Nurrl (63%) and between NOR-1 and Nur77 (54%) (Fig. 1B). The three-dimensional structure of Nurrl LBD was already described [21]. This nuclear receptor is active in the absence of ligands, and its ligand binding pocket (LBP) is blocked by bulky residues, conserved in the three NR4 family members [21]. The coactivator interaction surface observed in Nurrl is also different to other nuclear receptor members,

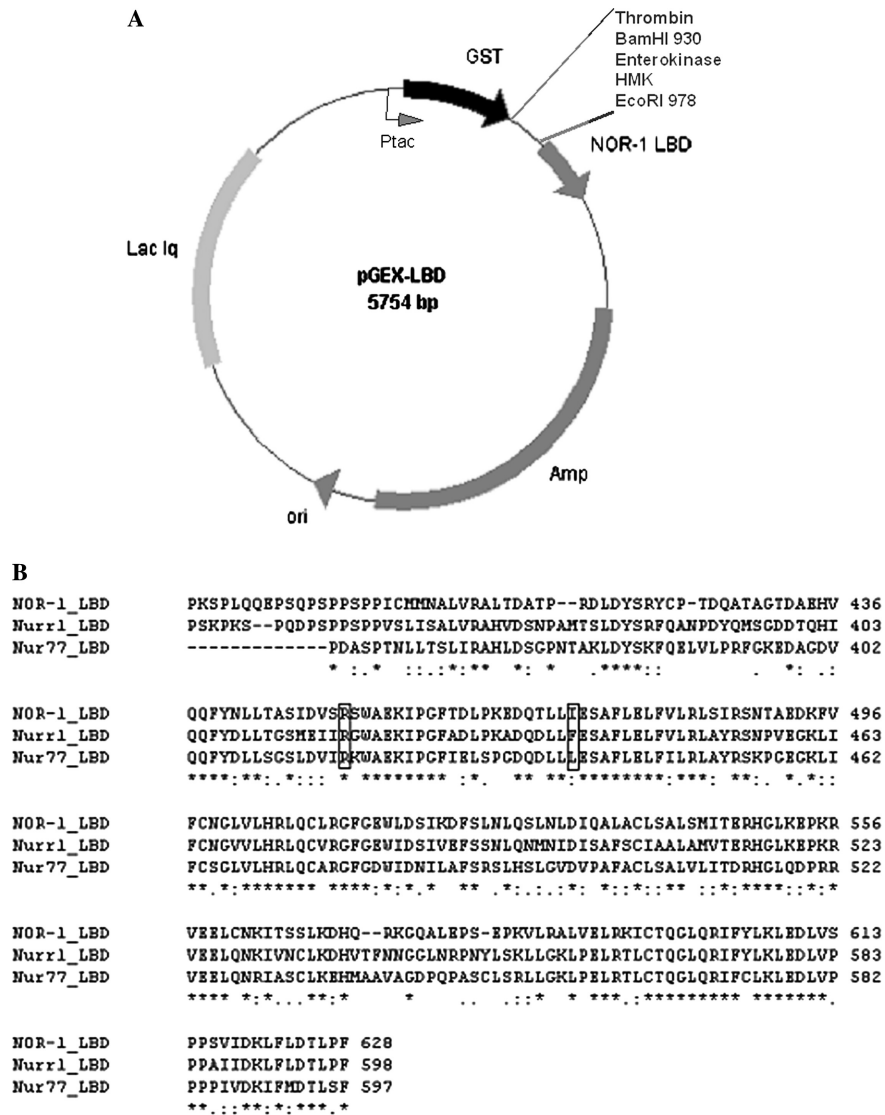


Fig. 1. (A) Schematic diagram of pGEX-LBD plasmid. (B) Amino acid sequence alignments between LBD domains of NR4 subfamily (Rat NOR-1, Nurr1, and Nur77). Boxed residues indicate the coactivator interaction surface described for Nurr-1.

and involves residues R418 and F439 (Fig. 1B, amino acid residues indicated in boxes). In NOR-1 LBD, R418 is conserved, however F439 is substituted for isoleucine, suggesting a functional or structural role for this region [21] (Fig. 1B).

Expression and purification of recombinant NOR-1 LBD

Several different conditions for expression of NOR-1 LBD cloned in pGEX-2T vector were tested. Initially, the induction period of fusion protein synthesis (GST-LBD) upon addition of IPTG was analyzed. The fusion protein was best produced after overnight incubations, particularly NOR-1 LBD was obtained at higher concentration from cultures induced at 1.7 O.D. at 600 nm. In this growing phase (stationary phase), *E. coli* was already shown to produce GST fusion proteins at

high yields [25]. The bacterial lysate, after clarification by centrifugation, was incubated with GSH–Sepharose. The resin was then washed until no protein was found in the eluate (Fraction I, Table 1). The fusion protein bound to the resin was eluted with a 10 mM GSH solution (Fraction II, Table 1) and then it was treated with thrombin (Fraction III, Table 1). The optimal condition for cleavage was 20 h at 4°C. Highly pure (at least 90%) NOR-1 LBD (Fraction IV, Table 1) was obtained (Fig. 2A). The yield, of GST-NOR-1 LBD, NOR-1 LBD after thrombin cleavage, and NOR-1 LBD after second GSH–Sepharose incubation are shown in Table 1. The protein yield at the last purification step was 3 mg/L of cell culture. The overall recovery was approximately 15% (Table 1), and the protein remained intact, without proteolytic degradation occurring during the purification procedure.

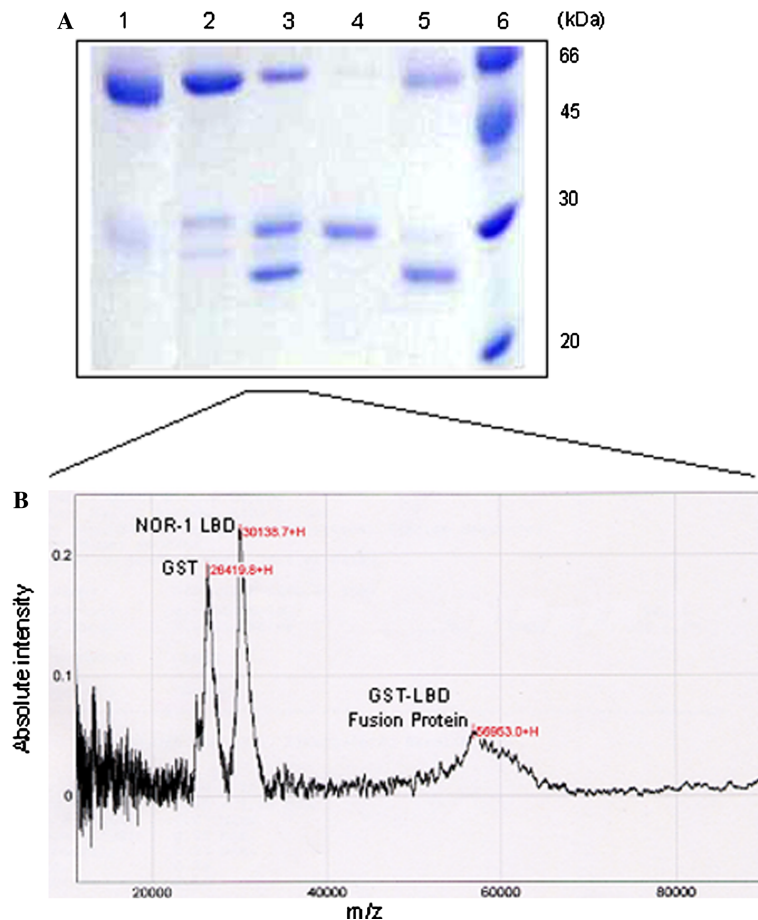


Fig. 2. (A) SDS-PAGE of NOR-1 LBD purification steps. Lanes: 1, proteins bound to glutathione-Sepharose before elution with reduced glutathione; 2, GST-NOR-1 LBD after elution from the column with reduced glutathione; 3, GST-NOR-1 LBD after treatment with thrombin; 4, NOR-1 LBD eluted from the second glutathione-Sepharose step; 5, proteins bound to second glutathione-Sepharose step after elution of NOR-1 LBD; and 6, molecular weight markers. (B) SELDI analysis of GST, GST-NOR-1 LBD, and NOR-1 LBD.

Characterization of purified recombinant NOR-1 LBD

The purified fraction was checked by N-terminal sequencing and mass spectrometry (Fig. 2B). Both methods confirmed the absence of contaminants, and also the coincidence between theoretical predicted recombinant NOR-1 LBD N-terminal sequence and molecular weight (30,138.5 Da), with that of the purified recombinant protein (30,138.7 Da).

Since NOR-1 LBD has no intrinsic catalytic activity, its correct folding was checked by thermal denaturation analysis under UV spectroscopy. We submitted NOR-1 LBD to thermal denaturation at 55°C for 2.5 h. The absorbance at 280 nm was measured, and a typical increase was detected over time. The increase in absorbance versus time followed a sigmoidal curve (Fig. 3), suggesting an unfolding event characteristic of increased tryptophan residue exposure to solvent [24].

The conformation of NOR-1 LBD was investigated by CD measurements. The far-UV CD spectrum at 4°C displayed two negative maxima at 222 and 208 nm (Fig. 4), suggesting the presence of α -helical secondary

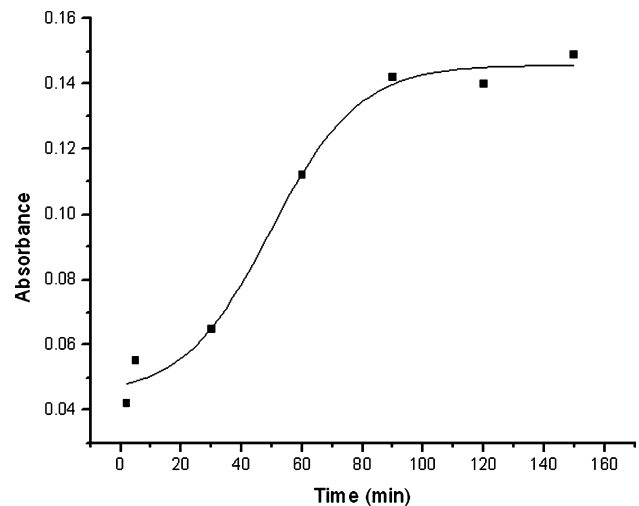


Fig. 3. Temperature induced unfolding. NOR-1 LBD absorbance at 280 nm versus incubation time at 55°C.

structure. Analysis of the spectrum with the software CDSSTR [26] gives the amount of α -helical structure of 52%, β -strands of 20%, and random coils of 28%.

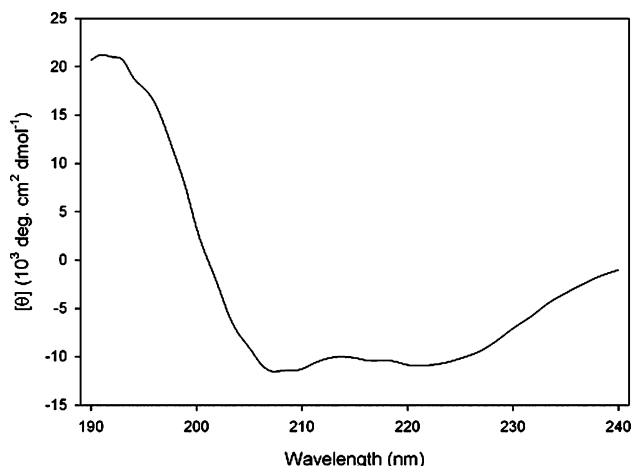


Fig. 4. CD spectrum of recombinant NOR-1 LBD domain at 10 μ M, 4°C.

These data are in good agreement to that observed in the crystal structure of a member of NR4 subfamily of nuclear receptors (Nurr1, PDB access 1OVL), in which the total α -helical content is about 52% [21]. CD spectrum at 80°C reveals a clear conformational change in the protein secondary structure. The CD signal at 222 nm was monitored from 5 to 80°C (Fig. 5). The temperature effect on the α -helical structure of NOR-1 LBD starts at approximately 45°C and reaches a plateau at about 55°C. The sigmoidal profile of the thermal transition is an indication of cooperativity as well as a native-like packing of the protein.

To fully understand NOR-1 transcriptional regulation, in the absence of a classical ligand binding pocket and coactivator interaction surface, structural and biochemical analyses will be necessary, and the use of highly purified and properly folded NOR-1 LBD protein, as described in this paper, would be an advantage. The

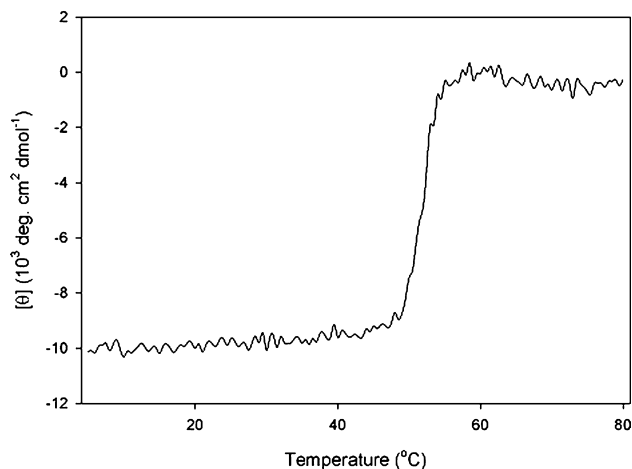


Fig. 5. Temperature induced unfolding of NOR-1 LBD. The decrease in the 222nm CD minima is plotted against temperature (°C).

high yield and purity of NOR-1 LBD described, as well as its proper folding, in comparison to the three-dimensional structure of a member of NR4 family (Nurr1), may be very useful to future structure–activity analysis of this nuclear receptor; and may eventually unmask the specific transcriptional regulation requirements in NR4 nuclear receptor family.

Acknowledgments

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Grant Nos. 471366/2003-0 and 550732/2002-0 (to H.T.) and 300220/96-0 (to I.P.) and by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Grant No. 99/03387-4 (to I.P.). CAPES (master degree fellowship to G.R.M.)

References

- [1] D.J. Mangelsdorf, C. Thummel, M. Beato, P. Herrlich, G. Schutz, The nuclear receptor superfamily: the second decade, *Cell* 83 (1995) 835–839.
- [2] A. Chawla, J.J. Repa, R.M. Evans, D.J. Mangelsdorf, Nuclear receptors and lipid physiology: opening the X-files, *Science* 294 (2001) 1866–1870.
- [3] A. Gordon, L. Francis, E. Fayard, F. Picard, J. Auwerx, Nuclear receptors and the control of metabolism, *Annu. Rev. Physiol.* 65 (2003) 261–311.
- [4] K.D. Senali Abayratna Wansa, J.M. Harris, G. Yan, P. Ordentlich, G.E.O. Muscat, The AF-1 domain of the orphan nuclear receptor NOR-1 mediates trans-activation, coactivator recruitment, and activation by the purine anti-metabolite 6-mercaptopurine, *J. Biol. Chem.* 278 (2003) 24736–24790.
- [5] H. Greschik, D. Moras, Structure–activity relationship of nuclear receptor–ligand interactions, *Curr. Top. Med. Chem.* 14 (2003) 1573–1599.
- [6] H. Terenzi, R.O. Cassia, M.M. Zakin, Expression, purification, and functional analysis of the DNA binding domain of nuclear receptor Rev-erb β , *Protein Expr. Purif.* 8 (1996) 313–318.
- [7] H. Terenzi, P.M. Alzari, M.M. Zakin, Structural features involved in the formation of a complex between the monomeric or the dimeric form of the Rev-erb β DNA-binding domain and its DNA reactive sites, *Biochemistry* 37 (1998) 11488–11495.
- [8] A. Aranda, A. Pascual, Nuclear hormone receptors and gene expression, *Physiol. Rev.* 81 (2001) 1269–1304.
- [9] E.P. Murphy, A. McEvoy, O.M. Conneely, B. Bresnihan, O. FitzGerald, Involvement of the nuclear orphan receptor NURR1 in the regulation of corticotropin-releasing hormone expression and actions in human inflammatory arthritis, *Arthritis Rheum.* 44 (2001) 782–793.
- [10] E.K. Arkenbout, V. de Waard, M. van Bragt, T.A. van Achterberg, J.M. Grimbergen, B. Pichon, H. Pannekoek, C.J. de Vries, Protective function of transcription factor TR3 orphan receptor in atherogenesis: decreased lesion formation in carotid artery ligation model in TR3 transgenic mice, *Circulation* 106 (2002) 1530–1535.
- [11] H. Monajemi, E.K. Arkenbout, H. Pannekoek, Gene expression in atherogenesis, *Thromb. Haemost.* 86 (2001) 404–412.
- [12] X.K. Zhang, Vitamin A and apoptosis in prostate cancer, *Endocr. Relat. Cancer* 9 (2002) 87–102.

- [13] A. Maltais, C. Filion, Y. Labelle, The AF2 domain of the orphan nuclear receptor TEC is essential for the transcriptional activity of the oncogenic fusion protein EWS/TEC, *Cancer Lett.* 183 (2002) 87–94.
- [14] T. Iwawaki, K. Kohno, K. Kobayashi, Identification of a potential nurr1 response element that activates the tyrosine hydroxylase gene promoter in cultured cells, *Biochem. Biophys. Res. Commun.* 274 (2000) 590–595.
- [15] H. Fischer, S.M.G. Dias, M.A.M. Santos, A.C. Alves, N. Zanchin, A.F. Craievich, J.W. Apriletti, J.D. Baxter, P. Webb, F.A.R. Neves, R.C.J. Ribeiro, I. Polikarpov, Low resolution structures of the retinoid X receptor DNA-binding and ligand-binding domains revealed by synchrotron X-ray solution scattering, *J. Biol. Chem.* 278 (2002) 16030–16038.
- [16] P.M. Fernandez, F. Brunel, M. Jimenez, J.M. Saez, S. Cereghini, M.M. Zakin, Nuclear receptors NOR-1 and NGFI-B/Nur77 play similar, albeit distinct, roles in the hypothalamus–pituitary–adrenal axis, *Endocrinology* 141 (2000) 2392–2400.
- [17] H. Inuzukaa, H. Tokumitsu, N. Ohkurab, R. Kobayashi, Transcriptional regulation of nuclear orphan receptor, NOR-1, by Ca^{2+} /calmodulin-dependent protein kinase cascade, *FEBS Lett.* 522 (2002) 88–92.
- [18] F.Q. Pirih, J.M. Nervina, L. Pham, T. Aghaloo, S. Tetradis, Parathyroid hormone induces the nuclear orphan receptor NOR-1 in osteoblasts, *Biochem. Biophys. Res. Commun.* 20 (2003) 144–150.
- [19] G.B. Wisely, A.B. Miller, R.G. Davis, A.D. Thomquest Jr., R. Johnson, T. Spitzer, A. Seffler, B. Shearer, J.T. Moore, A.B. Miller, T.M. Willson, S.P. Williams, Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids, *Structure* 10 (2002) 1225–1234.
- [20] S. Dhe-Paganon, K. Duda, M. Iwamoto, Y.I. Chi, S.E. Shoelson, Crystal structure of the HNF4 α ligand binding domain in complex with endogenous fatty acid ligand, *J. Biol. Chem.* 277 (2002) 37973–37976.
- [21] Z. Wang, G. Benoit, J. Liu, S. Prasad, P. Aarnisalo, X. Liu, H. Xu, N.P.C. Walker, T. Perlmann, Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors, *Nature* 423 (2003) 555–560.
- [22] M.A. Blanar, W.J. Rutter, Interaction cloning: identification of a helix-loop-helix zipper protein that interacts with c-Fos, *Science* 256 (1992) 1014–1018.
- [23] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 48–54.
- [24] A. Ginsburg, W.R. Carroll, Some specific ion effects on the conformation and thermal stability of ribonuclease, *Biochemistry* 4 (1965) 2159–2174.
- [25] Y.K. Chae, K.S. Cho, W. Chun, K. Lee, Protein production by stationary phase induction (Spi), *Protein Pept. Lett.* 10 (2003) 369–374.
- [26] L.A. Compton, W.C. Johnson Jr., Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication, *Anal. Biochem.* 155 (1986) 155–167.